

**Mechanistic Studies and Manipulation of the Enzymes
in Sialic Acid and Pseudaminic Acid Biosynthesis**

by

Feng Liu

B.Sc., Tianjin University, China, 1997

M.Sc., Tianjin University, China, 2001

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)

February, 2010

© Feng Liu, 2010

Abstract

The *Neisseria meningitidis* sialic acid synthase (NeuB) catalyzes the metal-dependent condensation of *N*-acetylmannosamine (ManNAc) and phosphoenolpyruvate (PEP) to generate *N*-acetylneuraminic acid. This work describes the synthesis and characterization of the first potent inhibitor of sialic acid synthase, a tetrahedral intermediate analogue as a mixture of stereoisomers at the key tetrahedral center. Inhibition studies demonstrate that one stereoisomer binds more tightly than the other. An X-ray crystallographic analysis of the NeuB·inhibitor·Mn²⁺ complex solved to a resolution of 1.75 Å shows that the more tightly bound stereoisomer bears a (2*R*)-configuration. This suggests that the tetrahedral intermediate formed in the NeuB reaction also bears a (2*R*)-configuration. This analysis demonstrates that the active site metal serves as a source of nucleophilic water and delivers it to the *si* face of the oxocarbenium intermediate to generate a tetrahedral intermediate with a (2*R*)-configuration.

The flagellin proteins in pathogenic bacteria such as *Campylobacter jejuni* and *Helicobacter pylori* are heavily glycosylated with the nine-carbon α -keto acid, pseudaminic acid. A key step in pseudaminic acid biosynthesis has been shown to involve the generation of 6-deoxy-AltdiNAc from its nucleotide-linked form, UDP-6-deoxy-AltdiNAc, by the action of a hydrolase that cleaves the glycosidic bond and releases UDP. This thesis describes the first characterization of a UDP-6-deoxy-AltdiNAc hydrolase, namely PseG (Cj1312) from *C. jejuni*. The activity of this enzyme is independent of the presence of divalent metal ions, and the values of the catalytic constants were found to be $k_{\text{cat}} = 27 \text{ s}^{-1}$ and $K_{\text{M}} = 174 \text{ }\mu\text{M}$. The enzyme was shown to hydrolyze the substrate with an overall inversion of stereochemistry at C1 and to utilize a C-O bond cleavage mechanism during catalysis.

The last part of the thesis describes the engineering of *C. jejuni*. We demonstrated that by feeding non-motile mutant *C. jejuni* bacteria with a neutral azide-labeled pseudaminic acid precursor, the mutants regained the ability to generate functional azido-bearing flagella and their motility was restored. The presence of the azido-pseudaminic acid on the surface of the flagella provides a bioorthogonal chemical handle that can be used to modify the flagellar proteins and to engineer bacteria for further studies.

Table of Contents

Abstract	ii
Table of Contents	iv
List of Tables	vii
List of Figures	viii
List of Symbols and Abbreviations	xiii
Acknowledgements	xix
Dedications	xx
Chapter 1 : Sialic Acid, Pseudaminic Acid and Related Higher Order Sugars	1
1.1 Sialic Acid-Related Higher Order Sugars	2
1.2 Sialic Acids	8
1.2.1 The Discovery of Sialic Acid.....	8
1.2.2 Sialic Acids in Mammals and Bacteria.....	9
1.2.3 Biosynthesis of Sialic Acid in Mammals and Bacteria.....	12
1.2.4 Hydrolyzing UDP-GlcNAc 2-Epimerase.....	14
1.2.5 Sialic Acid Synthase	17
1.3 Pseudaminic Acids	27
1.3.1 Flagella of <i>C. jejuni</i> and <i>H. pylori</i>	28
1.3.2 Flagellin Glycosylation with Pseudaminic Acids	28
1.3.3 Biosynthesis of Pseudaminic Acid.....	30
1.4 Glycoconjugate Modification Using Bio-orthogonal Reagents	33
1.4.1 Bio-orthogonal Chemical Reporters	34
1.4.2 Staudinger Ligation and Glycoconjugate-Labeling with Azides.....	40
1.5 Project Goals	44
Chapter 2 : The Inhibition of <i>Neisseria meningitidis</i> Sialic Acid Synthase by Intermediate Analogues	46
2.1 Introduction	47
2.2 Over-expression and Purification of <i>Neisseria meningitidis</i> Sialic Acid Synthase (NeuB) ...	51
2.3 Synthesis of, and Inhibition by, the Oxocarbenium Ion Intermediate Analogue (O1)	53
2.3.1 Synthesis of Analogue O1	53
2.3.2 Test for Inhibition	54

2.4 Syntheses of, and Inhibition by, Tetrahedral Intermediate Analogues	59
2.4.1 Synthesis of Analogue T1	62
2.4.2 Attempted Synthesis of Analogue T2	63
2.4.3 Inhibition Assay and Kinetic Evaluation of Analogue T1	67
2.4.4 Metal Dependency of Inhibition by Inhibitor T1	73
2.4.5 Crystallographic Analysis of the Complex Between Inhibitor T1 and NeuB	74
2.5 Conclusions	80
2.6 Future Directions.....	82
2.7 Experimental Procedures.....	85
2.7.1 Materials.....	85
2.7.2 General Methods	85
2.7.3 Over-expression and Purification of <i>Neisseria meningitidis</i> NeuB.....	86
2.7.4 Syntheses of Intermediate Analogues	88
2.7.5 Kinetic Studies	98
2.7.6 Crystallization, Data Collection, and Structure Refinement.....	100

Chapter 3 : PseG of Pseudaminic Acid Biosynthesis: A UDP-Sugar Hydrolase as a Masked Glycosyltransferase.....102

3.1 Introduction	103
3.2 Preparation of the Hydrolase Substrate UDP-6-deoxy-AltdiNAc	105
3.3 Identification of the UDP-6-deoxy-AltdiNAc Hydrolase from <i>C. jejuni</i>	108
3.3.1 Initial Attempts with NeuC2 (Cj1328).....	108
3.3.2 Identification of PseG (Cj1312) as the UDP-6-deoxy-AltdiNAc Hydrolase.....	117
3.4 Kinetic Analysis of the UDP-6-deoxy-AltdiNAc Hydrolase Reaction.....	123
3.5 Metal Dependency of PseG.....	125
3.6 Stereochemistry of the Hydrolase Reaction.....	127
3.7 Mechanistic Analysis of the Hydrolase Reaction	132
3.8 Conclusions	135
3.9 Future Directions.....	137
3.10 Experimental Procedures.....	140
3.10.1 Materials and General Methods	140
3.10.2 Chemoenzymatic Synthesis of UDP-6-deoxy-AltdiNAc	140
3.10.3 Chemical Hydrolysis and Characterization of the Product.....	141
3.10.4 Cloning, Over-expression and Purification of Enzymes.....	142
3.10.5 Enzyme Kinetics Using a Continuous Coupled Assay	144
3.10.6 Metal Dependency Experiment.....	145
3.10.7 Reaction Stereochemistry and Deuterium Incorporation.....	145
3.10.8 ¹⁸ O Incorporation Experiment.....	146

Chapter 4 : The Engineering of Bacteria Bearing Azido-Pseudaminic Acid-Modified Flagella	147
4.1 Introduction.....	148
4.2 Synthesis of Azido Sugars and Labeling Reagents.....	152
4.2.1 Chemo-enzymatic Synthesis of Azido-altrose.....	152
4.2.2 Syntheses of Staudinger Reagent 31	153
4.3 <i>In vitro</i> Tests.....	156
4.3.1 <i>In vitro</i> Conversion of 6-Deoxy-AltNAc4NAz to Az-Pse.....	156
4.3.2 <i>In vitro</i> Labeling of Az-Pse with Biotin-Staudinger Reagent (31).....	156
4.4 Motility Test.....	157
4.5 Probing the Azido Functionality on the Flagella	160
4.6 Conclusions.....	165
4.7 Future Directions.....	167
4.8 Experimental Procedures.....	168
4.8.1 Materials and General Methods.....	168
4.8.2 Over-expression and Purification of His ₆ -Tagged PseB, PseC, PseG and PseI	168
4.8.3 Chemical Syntheses	168
4.8.4 Motility Assays	174
4.8.5 Cell surface Tagging of Az-Pse with Staudinger Reagent, 31	175
4.8.6 Streptavidin HRP Detection of Biotinylated Az-Pse	175
Bibliography	176
Appendix	186

List of Tables

Table 2.1 Data Collection and Refinement Statistics.	101
---	-----

List of Figures

Figure 1.1	Sialic acids and sialic acid-related molecules.....	2
Figure 1.2	Sialic acid displayed on the cell surface.....	4
Figure 1.3	<i>Campylobacter jejuni</i> cell and its Pse-glycosylated flagellar filament.	4
Figure 1.4	The structural organization of lipopolysaccharide (LPS) in Gram-negative bacteria. ..	5
Figure 1.5	The shikimate pathway	6
Figure 1.6	Structure of <i>L. pneumophila</i> LPS.	7
Figure 1.7	Biosynthesis of Neu5Ac (sialic acid) in mammals and bacteria.	12
Figure 1.8	Three potential mechanisms for the hydrolyzing UDP-GlcNAc 2-epimerase.	14
Figure 1.9	Mechanism of the hydrolysis catalyzed by GDP-mannose mannosyl hydrolase.....	15
Figure 1.10	Positional isotopic exchange (PIX) catalyzed by RffE.....	16
Figure 1.11	Reactions catalyzed by PEP-condensing synthases..	18
Figure 1.12	Overall architecture of NeuB.....	20
Figure 1.13	Proposed mechanisms for NeuB reaction.....	22
Figure 1.14	Stereochemistry of addition: <i>si</i> -face of the ManNAc aldehyde approaches the <i>si</i> -face of PEP.	23
Figure 1.15	Active site of NeuB..	24
Figure 1.16	The mechanism of the reaction catalyzed by KDO 8-P synthase and structures of intermediate analogues.	25
Figure 1.17	Structures of Neu5Ac, Pse and PseAm	27
Figure 1.18	Biosynthesis of CMP-pseudaminic acid.....	30
Figure 1.19	General bio-orthogonal chemical reporter strategy.....	35
Figure 1.20	Ketone/aldehyde reporters and their ligation with hydrazide/aminooxy probes.....	35
Figure 1.21	Tetracysteine reporter and its ligation with FIAsh-/ReAsH-EDT ₂	37
Figure 1.22	The Staudinger ligation.	38

Figure 1.23 Azide-alkyne click chemistry.....	39
Figure 1.24 Classic Staudinger reaction and the reaction of aza-ylide with other electrophilic compounds.....	40
Figure 1.25 Mechanism of the Staudinger ligation.....	41
Figure 1.26 Metabolic delivery of azides to cell surfaces by unnatural sialic acid biosynthesis and the tagging of azides with a biotinylated phosphine reagent through the Staudinger ligation.....	42
Figure 2.1 Structures of KDO 8-P synthase intermediate analogues 1 and 2 and the one-pot direct reductive amination for the synthesis of compound 1	48
Figure 2.2 Intermediate analogues proposed as potential inhibitors of NeuB.....	49
Figure 2.3 SDS-PAGE for the over-expression and purification of His ₆ -NeuB.....	52
Figure 2.4 Synthesis of oxocarbenium analogue O1	54
Figure 2.5 Continuous coupled assay for measuring sialic acid synthase activity.....	55
Figure 2.6 Test of inhibition by analogue O1 (100 μM).....	55
Figure 2.7 A structural analysis of PEP binding in the active sites of NeuB and KDO 8-P synthase.....	57
Figure 2.8 Biosynthetic strategy for synthesizing analogues T1 and T2	60
Figure 2.9 Mechanism of indium-mediated allylation.....	61
Figure 2.10 Synthesis of analogue T1	62
Figure 2.11 Attempted synthesis of analogue T2	64
Figure 2.12 Potential mechanisms for the elimination of water from T2 under basic or acidic conditions.....	66
Figure 2.13 Inhibition of NeuB by analogue T1 as analyzed by ESI-MS.....	68
Figure 2.14 Assay for phosphate release by NeuB.....	69
Figure 2.15 A plot of initial velocity vs. PEP concentration for the reaction catalyzed by NeuB.....	70
Figure 2.16 Initial inhibition kinetics test using the continuous coupled phosphate assay.....	71
Figure 2.17 Inhibition kinetics of inhibitor T1 ((2 <i>S</i>)- T1 :(2 <i>R</i>)- T1 = 4.5:1).....	72
Figure 2.18 Metal dependency of the inhibition by T1	73
Figure 2.19 Electron density of inhibitor (2 <i>R</i>)- T1 in the NeuB• T1 •Mn ²⁺ complex.....	74

Figure 2.20 The active site of NeuB complexed with inhibitor (2 <i>R</i>)- T1	76
Figure 2.21 Revised mechanism of the reaction catalyzed by NeuB reflecting the proposed stereochemistry of the tetrahedral intermediate and the dual role played by Mn ²⁺	77
Figure 2.22 The octahedral Mn ²⁺ coordination sphere in A) NeuB• <i>N</i> -acetylmannosaminitol•PEP•Mn ²⁺ complex (regular), B) NeuB•inhibitor T1 • Mn ²⁺ complex (distorted).	79
Figure 2.23 Potential tetrahedral intermediate analogues.....	83
Figure 2.24 Other potential NeuB inhibitors.	84
Figure 3.1 Biosynthesis of CMP-pseudaminic acid.....	103
Figure 3.2 SDS-PAGE showing the purification of the PseB (Cj1293) and PseC (Cj1294) proteins.	106
Figure 3.3 Chemo-enzymatic synthesis of UDP-6-deoxy-AltdiNAc and its acidic hydrolysis to give 6-deoxy-AltdiNAc, the substrate for PseI.....	107
Figure 3.4 Comparison of the biosynthetic pathway of pseudaminic acid and sialic acid in bacteria.....	109
Figure 3.5 A sequence comparison of NeuC and NeuC2 and potential mechanisms.....	110
Figure 3.6 SDS-PAGE analysis showing the attempted purification of His ₆ -NeuC2.	112
Figure 3.7 Purification of recombinant protein using MBP as the affinity tag.....	114
Figure 3.8 SDS-PAGE analysis showing the purification of MBP-NeuC2.....	115
Figure 3.9 CMP-Leg5Ac7Ac biosynthesis and the proposed transformation from CMP-Leg5Ac7Ac to CMP-Leg5Am7Ac.....	116
Figure 3.10 Glycosyl-transfer reaction catalyzed by MurG.	118
Figure 3.11 Alignment of <i>C. jejuni</i> PseG and <i>E. coli</i> MurG. The alignment was obtained from an iterative PSI-BLAST homology search using PseG as the query sequence (four iterations).....	119
Figure 3.12 SDS-PAGE of purified PseG.....	120
Figure 3.13 ³¹ P NMR spectra monitoring the time course of the conversion of UDP-6-deoxy-AltdiNAc into UDP.	121
Figure 3.14 Continuous coupled assay monitoring the release of UDP.....	123
Figure 3.15 Enzyme kinetic plots of initial velocity vs. substrate concentration.	124
Figure 3.16 ³¹ P NMR spectra showing the metal-independency of PseG.	125

Figure 3.17 Mutarotation between anomers of pyranose monosaccharide through the open chain aldehyde form.	127
Figure 3.18 ¹ H NMR spectra monitoring the time course of the conversion of UDP-6-deoxy-AltdiNAc into α-6-deoxy-AltdiNAc and β-6-deoxy-AltdiNAc.	128
Figure 3.19 ⁴ C ₁ and ¹ C ₄ conformations of the α- and β-anomers of 6-deoxy AltdiNAc.	129
Figure 3.20 COSY spectrum of 6-deoxy-AltdiNAc.	130
Figure 3.21 NOE difference spectrum of 6-deoxy-AltdiNAc with irradiation at 5.28 ppm.	131
Figure 3.22 Potential mechanisms for the reaction catalyzed by UDP-6-deoxy AltdiNAc hydrolase (PseG).	132
Figure 3.23 UDP-6-deoxy-AltdiNAc and potential candidates for co-crystallization with PseG.	138
Figure 3.24 Proposed mechanism of LegG (Cj1328) and strategies to characterize it.	139
Figure 4.1 Biosynthesis of CMP-pseudaminic acid.	148
Figure 4.2 Potential mechanisms for the transport of sugar across the cell membrane.	150
Figure 4.3 6-Deoxy-AltNAc4NAz (27) and compound 28 as potential targets to introduce an azido functionality onto the <i>C. jejuni</i> flagella.	151
Figure 4.4 Chemo-enzymatic synthesis of 6-deoxy-AltNAc4NAz (27).	152
Figure 4.5 Staudinger reagents.	153
Figure 4.6 Synthesis of biotin-Staudinger reagent 31	153
Figure 4.7 Synthesis of fluorescent Staudinger reagent 32 using fluorescamine.	154
Figure 4.8 Synthesis of fluorescein-linked Staudinger reagent 33	155
Figure 4.9 <i>In vitro</i> conversion of 6-deoxy-AltNAc4NAz (27) to Az-Pse (41).	156
Figure 4.10 <i>In vitro</i> labeling of Az-Pse with biotin-Staudinger reagent 31	157
Figure 4.11 Motility of <i>C. jejuni</i> 81-176 and <i>C. jejuni</i> 81-176 <i>pseG::cat</i>	158
Figure 4.12 Generation of azido-pseudaminic acid-bearing bacteria and probing with biotin-linked Staudinger reagent, 31	160
Figure 4.13 Chemiluminescence detection by streptavidin-HRP.	161
Figure 4.14 Incorporation of Az-Pse into the flagellin glycoprotein.	162
Figure 4.15 Coomassie blue-stained SDS-PAGE gel of samples used in Western blot.	164

Figure A.1 ¹ H NMR spectrum of 2-acetamido-2-deoxy-3,4,5,6-tetra- <i>O</i> -acetyl-D-mannose <i>O</i> -methyloxime (7).....	187
Figure A.2 ¹ H NMR spectrum of methyl 5-acetamido-3,5-dideoxy-2-methylidene-4,6,7,8,9-penta- <i>O</i> -acetyl-D- <i>glycero</i> -D- <i>galacto</i> -2-nonulosonate (9).....	188
Figure A.3 ¹ H NMR spectrum of 5-acetamido-4,6,7,8,9-pentahydroxy-2-phosphorylnonanoic acid ditriethylammonium salt (T1).....	189
Figure A.4 ¹³ C NMR spectrum of 5-acetamido-4,6,7,8,9-pentahydroxy-2-phosphorylnonanoic acid ditriethylammonium salt (T1).....	190
Figure A.5 ¹ H NMR spectrum of UDP-6-deoxy-AltDiNAc.	191
Figure A.6 ¹ H NMR spectrum of 6-deoxy-AltDiNAc.....	192
Figure A.7 ¹ H NMR spectrum of UDP-6-deoxy-AltNAc4NAcCl (29).	193
Figure A.8 ¹ H NMR spectrum of 6-deoxy-AltNAc4NAcCl (30).....	194
Figure A.9 ¹ H NMR spectrum of 6-deoxy-AltDiNAc (27).....	195
Figure A.10 ¹ H NMR spectrum of Az-Pse (41).....	196

List of Symbols and Abbreviations

δ	chemical shift (ppm)
λ	wavelength
λ_{em}	emission wavelength
λ_{ex}	excitation wavelength
A	adenosine
Å	angstrom, 0.1 nm
A ₂₆₀	absorbance at 260 nm
A5P	D-arabinose 5-phosphate
<i>A. aeolicus</i>	<i>Aquifex aeolicus</i>
AAG	autoagglutination
ABC	ATP-binding cassette
Abs	absorbance
Ac	acetyl
Ac ₂ O	acetic anhydride
ACN	acetonitrile
Ag ₂ O	silver(I) oxide
AcCl	acetic chloride
AcCoA	acetyl coenzyme A
ADP	adenosine 5'-diphosphate
AFP	anti-freeze protein
ATP	adenosine 5'-triphosphate
Az	azido group
Az-Pse	azido-pseudaminic acid
BirA	biotin protein ligase
Bn	benzyl
BSA	bovine serum albumin
C	cytidine
<i>C. coli</i>	<i>Campylobacter coli</i>
<i>C. jejuni</i>	<i>Campylobacter jejuni</i>
CMP	cytidine 5'-monophosphate

CMP-Neu5Ac	cytidine 5'-monophospho- <i>N</i> -acetylneuraminic acid
CMP-Pse	cytidine 5'-monophospho-pseudaminic acid
CoA	coenzyme A
COSY	correlation spectroscopy
CTP	cytidine 5'-triphosphate
CV	column volume
d	doublet
dd	doublet of doublets
dt	doublet of triplets
D	deuterium (^2H)
Da	Dalton
DAH 7-P	3-deoxy- <i>D</i> - <i>arabino</i> -heptulosonate 7-phosphate
DCM	dichloromethane (CH_2Cl_2)
DMAP	<i>N,N'</i> -dimethyl-4-aminopyridine
DMF	dimethylformamide ($(\text{CH}_3)_2\text{NCOH}$)
DMSO	dimethylsulfoxide ($(\text{CH}_3)_2\text{SO}$)
DTT	dithiothreitol
DNA	deoxyribonucleic acid
E	enzyme
<i>E. coli</i>	<i>Escherichia coli</i>
EDT ₂	bis-1,2-ethanedithiol
EDTA	ethylene diamine tetraacetic acid
ESI-MS	electrospray ionization mass spectrometry
EtOH	ethanol
FGE	formylglycine generating enzyme
FITC	fluorescein isothiocyanate
FlaA	flagellin protein
FlaB	flagellin protein
FIAsH	fluorescein-based arsenical helix ligand
FRET	fluorescence resonance energy transfer
G	guanosine
GDP	guanosine 5'-diphosphate

GDP-Bac2Ac4Ac	guanosine 5'-diphospho- <i>N,N'</i> -diacetylbacillosamine
GFP	green fluorescent protein
GlcNAc	<i>N</i> -acetylglucosamine
GlcNAc 6-p	<i>N</i> -acetylglucosamine 6-phosphate
<i>H. pylori</i>	<i>Helicobacter pylori</i>
HRP	horse radish peroxidase
Hz	Hertz (s ⁻¹)
[I]	inhibitor concentration
IPTG	isopropyl β-D-1-thiogalactopyranoside
<i>J</i>	coupling constant (Hz)
<i>k</i> _{cat}	catalytic rate constant
<i>K</i> _M	Michaelis constant
<i>k</i> _{cat} / <i>K</i> _M	specificity constant; second-order rate constant
kDa	kiloDalton
KDN	2-keto-3-deoxy-D- <i>glycero</i> -D- <i>galacto</i> -2-nonulosonic acid
KDO	2-keto-3-deoxy-D- <i>manno</i> -octulosonic acid
KDO 8-P	2-keto-3-deoxy-D- <i>manno</i> -octulosonic acid 8-phosphate
<i>lac</i> operon	operon required for the transport and metabolism of lactose
LB medium	Luria-Bertani medium
LDH	lactate dehydrogenase
Leg	legionaminic acid
Leg5Am7Ac	5- <i>N</i> -acetimidoyl-7- <i>N</i> -acetyllegionaminic acid
LegG	Cj1328, putative GDP-Bac2Ac4Ac 2-epimerase, previously annotated as NeuC2
LIC	ligation-independent cloning
LOS	lipooligosaccharide
LPS	lipopolysaccharide
m	multiplet
ManLev	<i>N</i> -levulinoylmannosamine
ManNAc	<i>N</i> -acetylmannosamine
ManNAc 6-p	<i>N</i> -acetylmannosamine 6-phosphate
MBP	maltose binding protein

MeOH	methanol
MESG	methylthioguanosine or amino-6-mercapto-7-methylpurine ribonucleoside
MH	Mueller-Hinton
MOMP	major outer membrane protein
MurG	a glycosyltransferase involved in peptidoglycan biosynthesis
MurNAc	<i>N</i> -acetylmuramic acid
MW	molecular weight
MWCO	molecular weight cut-off
NAD ⁺ , NADH	nicotinamide adenine dinucleotide, and its reduced form
NADP ⁺ , NADPH	nicotinamide adenine dinucleotide phosphate, and its reduced form
NCAM	neural cell adhesion molecule
Neu5Ac	<i>N</i> -acetylneuraminic acid, or sialic acid
Neu5Ac 6-P	<i>N</i> -acetylneuraminic acid 6-phosphate
<i>neuA</i>	the gene encoding for CMP-sialic acid synthetase, the gene encoding for <i>H. pylori</i> CMP-pseudaminic acid synthetase, <i>pseF</i> , was also first annotated as <i>neuA</i>
NeuB	sialic acid synthase
NeuC	UDP GlcNAc 2-epimerase
<i>neuC</i>	the gene encoding for NeuC
NeuC2	the name previously assigned for Cj1328 or LegG
<i>neuC2</i>	the gene encoding for NeuC2
NDP	nucleotide diphosphate (any nucleotide)
<i>N. meningitidis</i>	<i>Neisseria meningitidis</i>
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NRC	National Research Council
OD ₆₀₀	optical density at 600 nm
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEO	poly (ethylene oxide)
PEP	phosphoenolpyruvate

PIX	positional isotope exchange
PLP	pyridoxal 5'-phosphate
PNP	purine nucleoside phosphorylase
ppm	parts per million
PSA	polysialic acid
Pse	5,7-diacetamido-2-keto-3,5,7,9-tetra-deoxy-L-glycero-L-manno-2-nonulosonic acid or pseudaminic acid
PseA	putative amidotransferase
PseAm	5-acetamidino derivative of pseudaminic acid
PseB	UDP-N-acetylglucosamine 5-inverting 4,6-dehydratase
<i>pseB</i>	the gene encoding for PseB
PseC	UDP-2-acetyl-2,6-dideoxy-β-L-arabino-4-hexulose aminotransferase
<i>pseC</i>	the gene encoding for PseC
PseF	CMP-pseudaminic acid synthase
<i>pseF</i>	the gene encoding for PseF
PseG	UDP-2,4-diacetamido-2,4,6-trideoxy-L-altrose hydrolase
<i>pseG</i>	the gene encoding for PseG
PseH	UDP-2-acetamido-2,4,6-trideoxy-L-altrose N-acetyltransferase
PseI	pseudaminic acid synthase
psi	pounds per square inch
Pyr	pyridine
PTS	phosphate transport system
R5P	ribose 5-phosphate
ReAsH	resorufin-based arsenical helix ligand
RffE	UDP-N-acetylglucosamine 2-epimerase
rManNAc	N-acetylmannosaminitol
rmsd	root-mean-square deviation
rpm	revolutions per minute
RT	room temperature
s	singlet (in reference to NMR)
[S]	substrate concentration
SAL	sialic acid lyase

SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
sialyl Le ^a	sialyl Lewis a
SiaLev	<i>N</i> -levulinoyl sialic acid
sialyl Le ^x	sialyl Lewis x
<i>t</i>	time
T	thymidine
<i>tac</i>	a highly efficient promoter, hybrid of <i>E. coli trp</i> and <i>lac</i> promoters
TBAI	tetrabutylammonium iodide
TEA	triethylamine
THF	tetrahydrofuran (C ₄ H ₈ O)
TLC	thin layer chromatography
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
tRNA	transfer RNA
<i>trp</i>	<i>trp</i> operon that is required for the biosynthesis of tryptophan
TyrRS	tyrosyl-tRNA synthase
U	units
UDP or OUDP	uridine 5'-diphosphate
UDP-Bac2Ac4Ac	uridine 5'-diphospho- <i>N,N'</i> -diacetylbacillosamine
UMP	uridine 5'-monophosphate
UTP	uridine 5'-triphosphate
UV	ultraviolet
V_{\max}	maximal reaction velocity
W_{ax}	axial metal-bound water
W_{eq}	equatorial metal-bound water

Acknowledgements

My success is completely attributable to my research supervisor Dr. Martin Tanner, as well as group members of the Tanner lab. I would like to thank Dr. Martin Tanner for his guidance, kindness and patience over the past few years. I would like to thank my predecessors for teaching me lab skills and all group members for their support and friendship. I would like to thank Ho Jun Lee, Dr. Natalie C. J. Strynadka, Dr. Ian Shoenhofen and Dr. Susan Logan for their collaborative works. I would like to acknowledge the shops and services in the Chemistry Department at UBC. In particular, I would like to thank current and former Biological Services staff, including Dr. Elena Polishchuk, Candice Martins, Helen Wright and Jessie Chen, as well as current and former Mass Spectrometry Centre staff, including Dr. Yun Ling, Minaz Lakha, Marshall Lapawa and David Wong, for their support and helpful discussions. I would also like to thank Dr. Warren Wakarchuk for donating the recombinant plasmids Cj1293 (*pseB*), Cj1294 (*pseC*) and Cj1328 (*legG*).

Dedicated to

My parents

&

My wife Huaiyu

**Chapter 1 : Sialic Acid, Pseudaminic Acid and Related Higher
Order Sugars**

1.1 Sialic Acid-Related Higher Order Sugars

Sugars not only play important roles in the storage and transport of energy, but also serve as starting materials in the biosyntheses of other organic compounds. Some sugars are also important structural components in living organisms and may play key roles in cellular recognition phenomena. Among the latter group of carbohydrates, the sialic acid-related higher order sugars form a unique family and are usually displayed on the cell surface or are located in cell membranes.

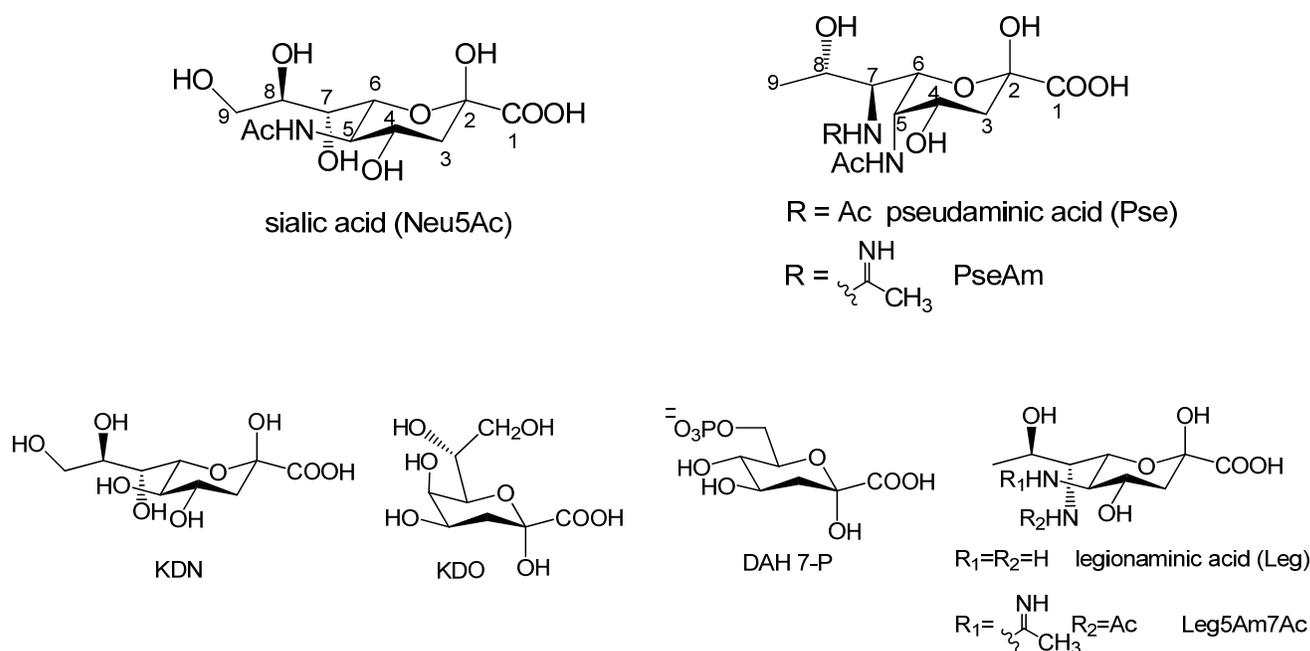


Figure 1.1 Sialic acids and sialic acid-related molecules.

The sialic acid-related higher order sugars are a group of naturally occurring 2-keto-3-deoxy polyhydroxylated acids containing more than six carbons. They include the sialic acids (derivatives of *N*-acetyl neuraminic acid, Neu5Ac, and KDN), 5,7-diacetamido-2-keto-3,5,7,9-

tetradecy-L-*glycero*-L-*manno*-2-nonulosonic acid (pseudaminic acid), 2-keto-3-deoxy-D-*manno*-octulosonic acid (KDO), 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAH 7-P) and 5,7-diamino-3,5,7,9-tetradecy-D-*glycero*-D-*galacto*-non-2-ulosonic acid (legionaminic acid) (Figure 1.1).¹

These α -keto acids share many biological and structural similarities. They are not ubiquitous in nature and are often only found in certain species of animals and bacteria. With the exception of DAH 7-P (a shikimate pathway precursor), these acids are found as components of cell surface polysaccharides or glycoconjugates. They are all synthesized from the condensation of a four- to six-carbon monosaccharide and phosphoenolpyruvate (PEP). In addition, unlike other monosaccharides that are usually activated in the nucleotide form of uridine or guanine diphosphates, most of these α -keto 3-deoxy acids are activated as cytidine 5'-monophosphate (CMP) nucleotides before being transferred onto cell surfaces.

The sialic acids are a family of over 60 derivatives of *N*-acetylneuraminic acid (Neu5Ac) and 2-keto-3-deoxy-D-*glycero*-D-*galacto*-2-nonulosonic acid (KDN), that may bear modifications at C-4, C-7, C-8 and C-9 with acetate, lactate, sulphate, phosphate and methyl ethers. Sialic acids are mainly found in the deuterostome lineage of animals (vertebrates, ascidians and echinoderms) and the pathogenic bacteria infecting them. They are often displayed on the cell surface (Figure 1.2) and are involved in the recognition and the mediation of cellular adhesion processes. The biological roles of sialic acid will be discussed in greater detail in Section 1.2.

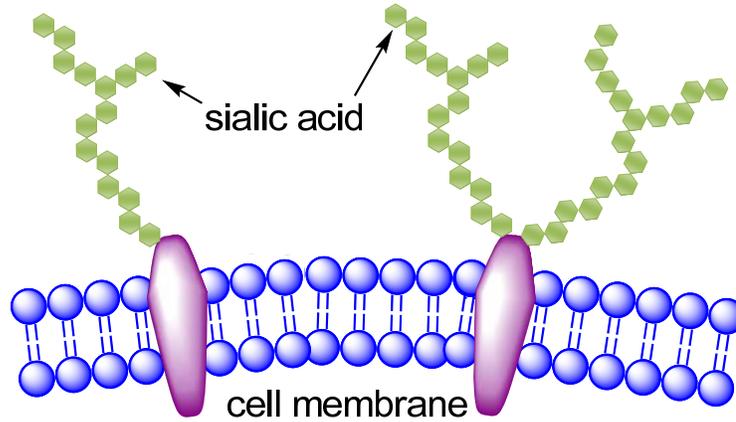


Figure 1.2 Sialic acid displayed on the cell surface.

Pseudaminic acid and its derivative (PseAm) are found as the major post-translational modification on the flagellin proteins of some enteric bacteria, such as *Campylobacter jejuni*, *Campylobacter coli* and *Helicobacter pylori* (Figure 1.3).²⁻⁴ They are very important for the assembly of flagellar filaments, which are essential to the motility, colonization and invasion of these bacteria. The biological roles of pseudaminic acid will be discussed in greater detail in Section 1.3.

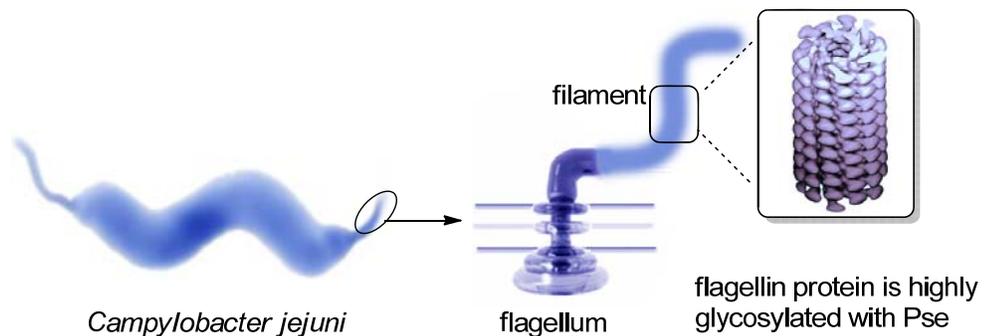


Figure 1.3 *Campylobacter jejuni* cell and its Pse-glycosylated flagellar filament.

KDO is a unique and essential component of the inner core oligosaccharide in the lipopolysaccharide (LPS) of Gram-negative bacteria (Figure 1.4).⁵ LPS is a major component of the Gram-negative bacterial outer membrane and is composed of O-antigen (an oligosaccharide chain compound of repeating subunits), core oligosaccharide and lipid A. KDO links the outermost O-antigen and outer part of the core oligosaccharide to lipid A. The enzymes responsible for the biosynthesis of KDO have become attractive targets in the design of antibiotics, as the disruption of KDO synthesis leads to the failure of the outer membrane in Gram-negative bacteria and finally to the death of bacteria.

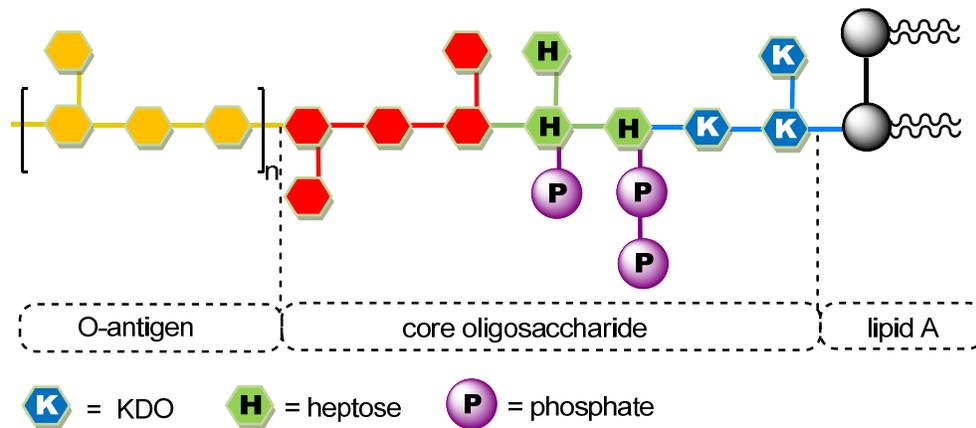


Figure 1.4 The structural organization of lipopolysaccharide (LPS) in Gram-negative bacteria.

DAH 7-P is the first product in the shikimate pathway which ultimately leads to the formation of shikimate (Figure 1.5).⁶ Shikimate is the starting material employed in the biosyntheses of many aromatic compounds, such as the aromatic amino acids and many secondary aromatic metabolites. The shikimate pathway only exists in plants and microorganisms, whereas animals do not have this biosynthetic machinery and hence have to be supplemented with aromatic amino acids from external sources. Because of this difference, DAH

7-P synthase and other enzymes involved in this pathway are potential targets for antimicrobial drugs to be used in humans and other animals.

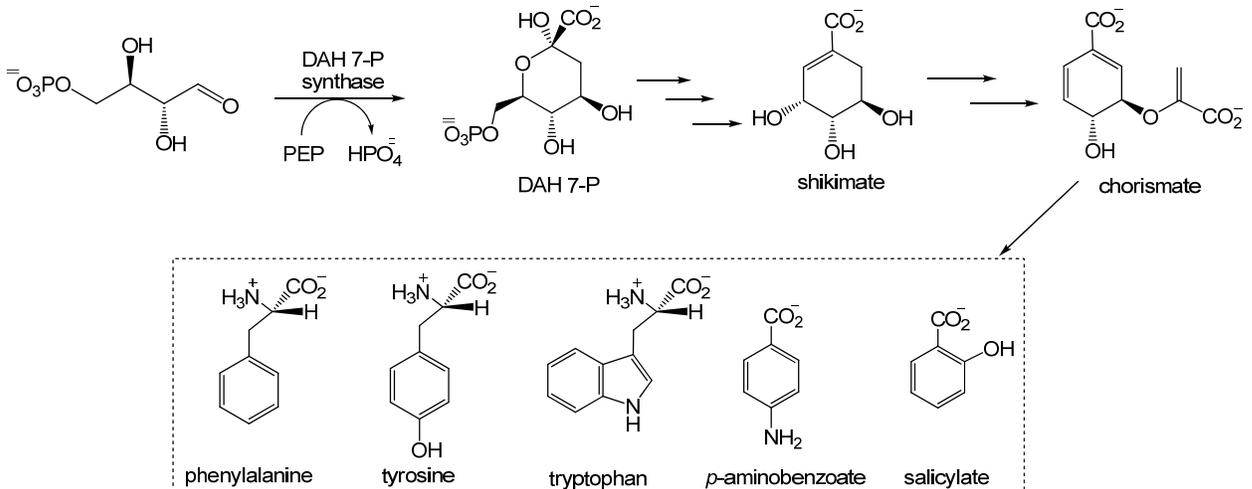


Figure 1.5 The shikimate pathway.

Legionaminic acids derived their name from *Legionella pneumophila*, a facultative intracellular parasitic bacterium that causes Legionnaires' disease.⁷⁻⁹ *Legionella pneumophila* contains a repeating homopolymer of α -(2 \rightarrow 4)-linked 5-N-acetimidoyl-7-N-acetyllegionaminic acid (Leg5Am7Ac) in the O-antigen part of its LPS (Figure 1.6).^{10, 11} Leg5Am7Ac is also produced in *Campylobacter coli* and is used to modify the flagellin protein as a monosaccharide unit via an O-linkage.¹²

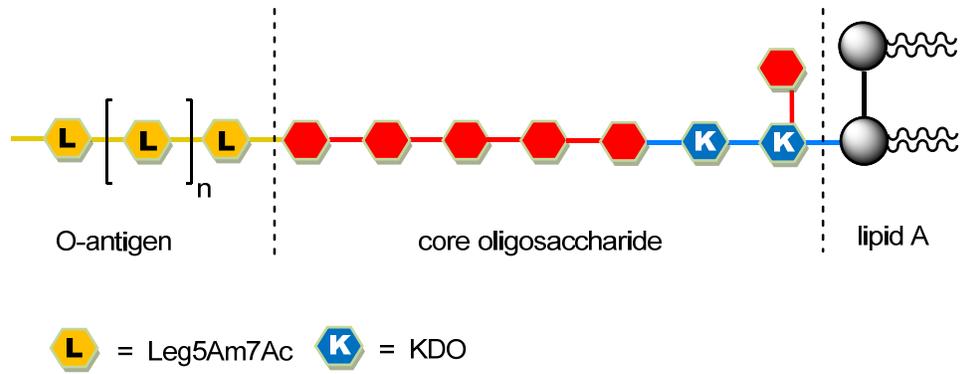


Figure 1.6 Structure of *L. pneumophila* LPS.

The research described in this thesis focuses on enzymes involved in the biosynthesis of sialic acid and pseudaminic acid. Therefore, the following introductory sections of this chapter will expand on these topics and introduce the goals of the research.

1.2 Sialic Acids

1.2.1 The Discovery of Sialic Acid¹³

Sialic acid was first discovered in 1927, when a sugar-like compound was identified in a glycolipid fraction extracted from animal kidney. This compound showed an unusual purple color upon treatment with Bial's reagent, whereas most other sugars gave a green color. In the 1930's Klenk found a brain glycolipid showing the same property. When the glycolipid was subjected to methanolysis, a crystalline substance which contained a carboxylic group and a free amino group, was isolated. He later named this compound 'methoxyneuraminsäure' (methoxyneuraminic acid) because of its brain glycolipid origin. We now know that the acidic methanolysis of lipid sialoside will generate the methylglycoside of neuraminic acid.

In 1936 Blix also isolated a crystalline substance, later named sialic acid by him, from the glycoprotein bovine submaxillary mucin. This compound appeared to have an *N*-acetyl group and was shown to have some chemical characteristics in common with the compound discovered by Klenk. Blix's sialic acid was later identified as *N*-acetylneuraminic acid and could be readily converted into the neuraminic acid methylglycoside isolated by Klenk upon treatment with methanolic HCl at 105 °C.

Further studies fully elucidated the structure of neuraminic acid and its *N*-acetyl derivative, sialic acid.^{14, 15} 'Sialic acid' was then suggested for use as the group name for all *N*- and *O*- substituted neuraminic acid derivatives.¹⁶ Many more derivatives have been discovered over the years and the sialic acids have become a large family of more than 60 members. However, the name 'sialic acid', in its narrow sense, still refers to *N*-acetylneuraminic acid

(Neu5Ac) in most cases. For example, sialic acid synthase, one of the focuses of this thesis, is the enzyme responsible for catalyzing the formation of Neu5Ac.

1.2.2 Sialic Acids in Mammals and Bacteria

In mammals, sialic acids are usually found at the distal end of oligosaccharide chains of cell surface glycoconjugates such as glycoproteins and glycolipids. Their C-1 carboxylate functionality is negatively charged at physiological pHs. Their outer most position and negative charge make them key determinants in regulating cellular processes such as cell adhesion and recognition.¹⁷

Sialic acid is known to play a key role in the development and regeneration of the vertebrate embryonic nervous system. Neural cell adhesion molecules (NCAMs) are cell surface glycoproteins expressed in neurons, muscles and glia, and are important in cell-cell adhesion, neurite growth and learning. The adhesion properties of NCAMs are mainly due to homophilic binding. During the early stages of neural cell development, surface-displayed NCAMs are glycosylated with polysialic acid (PSA), a homopolymer of up to 200 α -(2-8)-linked sialic acids. The high surface charge and hydration brought about by sialylation inhibits the homophilic binding of NCAMs. At later stages NCAMs are desialylated and regain their full adhesion function.^{18, 19}

Lectins are proteins that bind certain carbohydrates with high specificity. Endogenous and exogenous (from pathogens *etc.*) sialic acid-binding lectins play critical roles in an animal's life cycle, from pathogen invasion to the immune response and the regulation of cell death.^{1, 20} Selectins, for example, are a family of three calcium-dependent lectins and can be subcategorized into E-, L- and P-selectins. These cell adhesion proteins are expressed on the

surface of cells such as activated endothelial cells (E-selectin), leukocytes (L-selectin) or activated platelets (P-selectin) and are found to be essential factors in leukocyte trafficking during inflammation and the immune response.^{21, 22} The interactions of selectins with two particular tetrasaccharide-containing sialic acids, sialyl Lewis x (Le^x) and sialyl Lewis a (Le^a), are vital to transporting leukocytes from the flowing bloodstream to the site of inflammation. Selectins first tether leukocytes on the endothelial cell surface, enabling them to roll along the blood vessel wall. This triggers a cascade of molecular interactions that leads to leukocyte extravasation which is the starting point of leukocyte interstitial migration. High amounts of sialyl Le^x and sialyl Le^a are also found on the surface of certain cancer cells and are responsible for the adhesion of cancer cells to endothelium. The interactions of P- and E-selectins with sialyl Le^x and sialyl Le^a has been shown to mediate tumor metastasis.²³⁻²⁵ Blocking these interactions is a potential target for cancer therapy.

Pathogens such as viruses and bacteria also express sialic acid-binding lectins and use them to recognize and infect host cells. The viral sialic acid-recognizing lectins are usually referred to as hemagglutinins because of their capability to agglutinate red blood cells. Among these lectins, influenza virus hemagglutinins are the most well-known and well-studied molecules. They bind to certain sialic acid-containing side chains on the surface of host cells. Influenza viruses can adapt to new hosts by producing hemagglutinins that bind to new types of sialic acids through cross-species infections. Some bacteria also express soluble sialic acid-recognizing lectins that are called adhesins. Adhesins expressed by certain strains of bacteria usually specifically recognize carbohydrate side chains on the cell surface of the target tissue that is to be infected or colonized.

While some pathogens use sialic acid-specific lectins as virulence factors to invade host cells, animals also have similar molecules, such as complement factor H, that act as defensive systems to fight against invaders. Complement factor H is an inhibitory regulator that binds to sialic acid on the cell surface and prevents cells from being attacked by the host's own complement pathway, an essential component of the host innate immune system.²⁶ Foreign cells that are not covered by sialic acids will be exposed to the attack by complement, which triggers proteolytic cascade, phagocytosis, inflammation and finally the response of the whole immune system.

After hundreds of millions of years of evolution, bacteria have developed a lot of ways to circumvent the mammalian immune system. The use of sialic acid in a “Trojan horse” strategy is one of them. Although most bacteria do not produce sialic acids, several pathogenic strains are able to biosynthesize and display them on their cell surface to evade the host immune system by mimicking the sialylated mammalian cells. These include *Escherichia coli* K1, *Campylobacter jejuni* and *Neisseria meningitidis*. Neuroinvasive *Escherichia coli* K1 colonizes an infant's intestines and causes bacteremia, which leads to meningitis. *Campylobacter jejuni* causes food-borne gastroenteritis and will be addressed in Section 1.3.1 along with pseudaminic acid. *Neisseria meningitidis* is a Gram-negative bacterium that causes meningitis and can lead to mortality in childhood, and epidemics in underdeveloped countries with insufficient medical facilities. *Escherichia coli* K1 and *Neisseria meningitidis* produce a capsule composed of α -(2-8)-linked polysialic acid (PSA). The capsule not only serves as a shield outside of the cell wall to protect the bacterium against harmful chemicals and phagocytosis, but also helps the pathogen to evade the host's immune system by mimicking the mammalian NCAM-containing cells. *Campylobacter jejuni* and *Neisseria meningitidis* also have sialic acid in their

lipooligosaccharides (LOSs) which mimic human glycolipids.^{27, 28} PSA-containing capsules and sialic acid-containing LOSs are considered important virulence factors in these bacterial pathogens.²⁸⁻³²

In order to unveil bacterial pathogenesis and develop methods to fight diseases caused by bacteria and viruses, it is crucial to obtain a good understanding of the biosynthesis of sialic acid.

1.2.3 Biosynthesis of Sialic Acid in Mammals and Bacteria

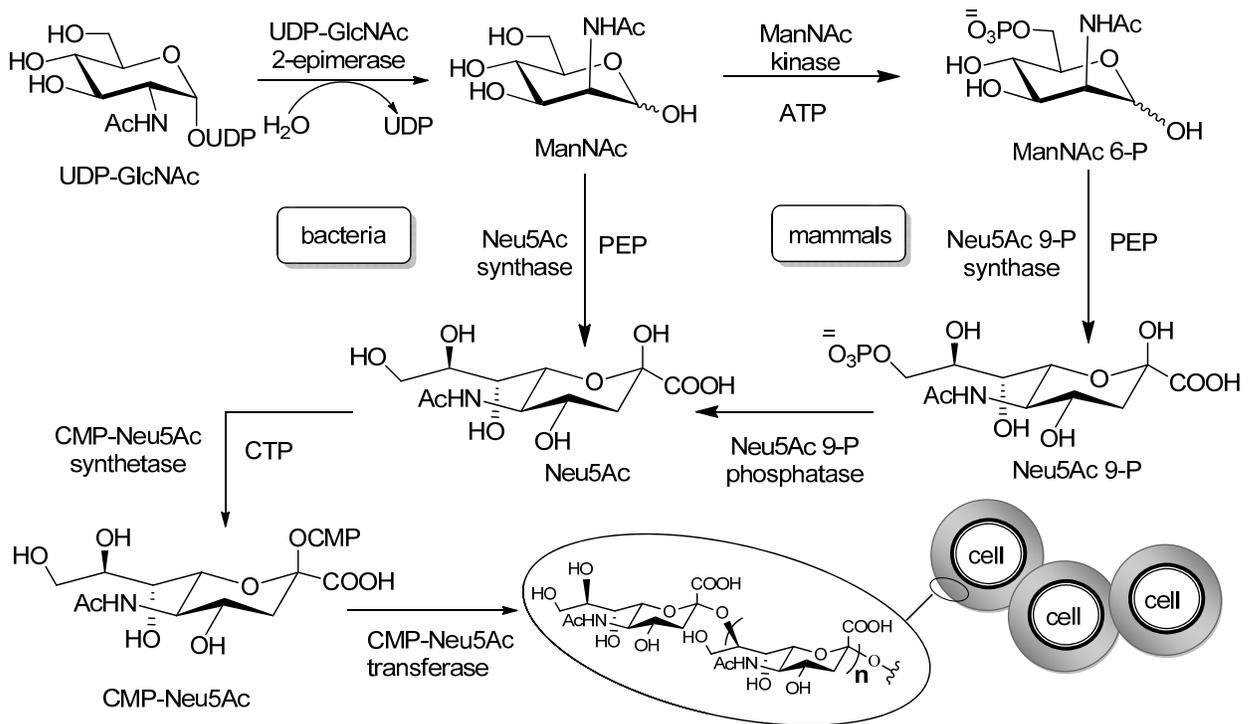


Figure 1.7 Biosynthesis of Neu5Ac (sialic acid) in mammals (outer pathway) and bacteria (inner pathway).

Sialic acids are biosynthesized slightly differently in animals and bacteria, however, both pathways start with the same sugar nucleotide uridine 5'-diphosphate *N*-acetyl-D-glucosamine (UDP-GlcNAc) (Figure 1.7). Sugar nucleotides are glycosylated nucleoside (purine or

pyrimidine) mono- or diphosphates with phosphate linked to the anomeric position of the monosaccharide. Sugar nucleotides are the activated forms of the monosaccharides because nucleotides (nucleoside phosphates) are excellent leaving groups. These high-energy glycosyl donors are substrates for glycosyl transferases and other sugar-modifying enzymes.

In mammals, UDP-GlcNAc is first converted to *N*-acetyl-D-mannosamine 6-phosphate (ManNAc 6-P) in two steps by a bifunctional enzyme, uridine 5'-diphosphate *N*-acetyl-D-glucosamine 2-epimerase/*N*-acetyl-D-mannosamine kinase (UDP-GlcNAc 2-epimerase/ManNAc kinase, Figure 1.7, outer pathway). The epimerase domain of this enzyme inverts the stereochemistry at C-2 and hydrolyzes the UDP-sugar linkage to form *N*-acetyl-D-mannosamine (ManNAc). ManNAc is then phosphorylated to give ManNAc 6-P by the kinase domain of the same enzyme. The next enzyme on the pathway, *N*-acetylneuraminic acid 9-phosphate synthase (Neu5Ac 9-P synthase), then catalyzes the condensation of ManNAc 9-P with phosphoenolpyruvate (PEP) to generate *N*-acetylneuraminic acid 9-phosphate (Neu5Ac 9-P). Neu5Ac 9-P is then dephosphorylated by *N*-acetylneuraminic acid 9-phosphate phosphatase (Neu5Ac 9-P phosphatase) to yield sialic acid (Neu5Ac). A cytidine 5'-monophosphate *N*-acetylneuraminic acid synthetase (CMP-Neu5Ac synthetase), the last enzyme on the pathway, then activates Neu5Ac as cytidine 5'-monophosphate *N*-acetylneuraminic acid (CMP-Neu5Ac) using cytidine 5'-triphosphate (CTP). Subsequent transferases are responsible for delivering it to glycoconjugates that end up on cell surfaces. CMP-Neu5Ac is known to be the substrate for all of the 18 known Neu5Ac transferases.

In the bacterial pathway, the phosphorylated intermediates are not formed. Instead, a mono-functional enzyme, UDP-GlcNAc 2-epimerase, converts UDP-GlcNAc to ManNAc (Figure 1.7, inner pathway). The second enzyme on the pathway, *N*-acetylneuraminic acid

synthase (Neu5Ac synthase) or sialic acid synthase (NeuB) takes ManNAc as the substrate and directly converts it to *N*-acetylneuraminic acid via condensation with PEP. The bacterial sialic acid synthase is found to be homologous to the mammalian Neu5Ac 9-P synthase with a sequence identity of 30~40%.³³ Finally, bacterial CMP-Neu5Ac synthetase forms CMP-Neu5Ac and activates the sugar for incorporation into capsular polysaccharides and lipooligosaccharides by bacterial sialyltransferases.

1.2.4 Hydrolyzing UDP-GlcNAc 2-Epimerase

The first step in sialic acid biosynthesis involves a combination of epimerization at C-2 and hydrolysis of the UDP-glycosidic bond catalyzed by UDP-GlcNAc 2-epimerase. Mammalian and bacterial UDP-GlcNAc 2-epimerases have been extensively studied in our research group.³⁴⁻³⁶ Three possible mechanisms can be envisioned to account for this conversion (Figure 1.8).

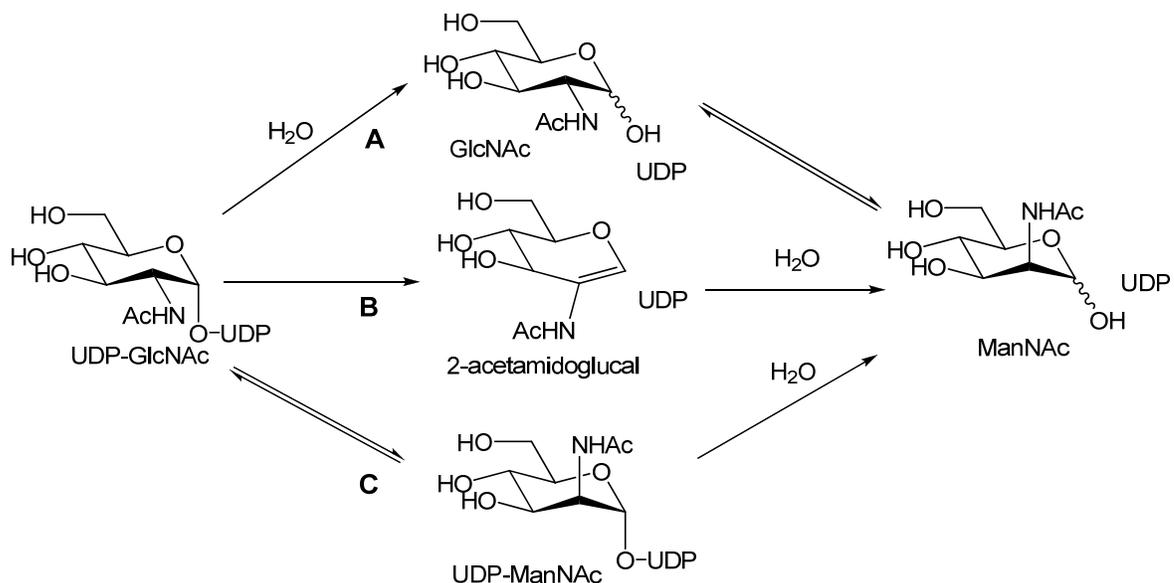


Figure 1.8 Three potential mechanisms for the hydrolyzing UDP-GlcNAc 2-epimerase.

In path A, hydrolysis occurs first to give free GlcNAc, which is then epimerized to ManNAc. The epimerization in Path A is likely to proceed via an open chain form of the free sugar and would involve deprotonation-reprotonation of the relatively acidic proton at C-2. In Path B, the reaction proceeds via the formation of a 2-acetamidoglucal intermediate. In this case, the glucal intermediate is formed by an *anti*-elimination of UDP from UDP-GlcNAc and a subsequent *syn*-hydration gives ManNAc.³⁷ Finally, in Path C, epimerization could occur before the hydrolysis takes place. The epimerization of UDP-GlcNAc would presumably proceed through the same 2-acetamidoglucal intermediate as described in Path B, however, a *syn*-addition of UDP, and not water, would give UDP-ManNAc. Precedent for this step comes from the activity of the known non-hydrolyzing UDP-GlcNAc 2-epimerase (RffE) that interconverts UDP-GlcNAc and UDP-ManNAc (Figure 1.10).³⁴ The hydrolysis steps in Paths A and C would very likely involve the direct nucleophilic attack of water at the anomeric position as seen in the reaction catalyzed by GDP-mannose mannosyl hydrolase (Figure 1.9).³⁸

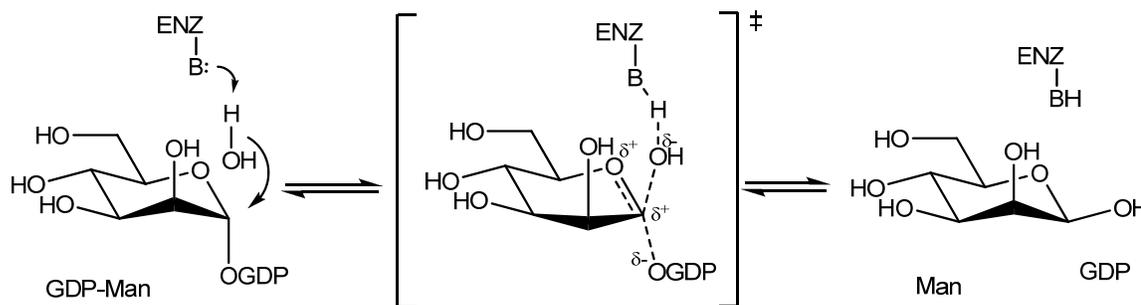


Figure 1.9 Mechanism of the hydrolysis catalyzed by GDP-mannose mannosyl hydrolase.

The mechanism was initially probed by an isotope incorporation study. When the reaction was carried out in tritiated water, the ManNAc produced was shown to have tritium incorporated at the C-2 position.³⁹ This result indicates that the epimerization is accompanied by a deprotonation-reprotonation process at C-2, and is consistent with all three proposed

mechanisms. Kinetic studies favored Paths A and B by showing that the reaction proceeds in an ordered manner in which UDP is released first, followed by the formation of ManNAc.³⁷ Moreover, the hydrolyzing UDP-GlcNAc 2-epimerase does not take GlcNAc as a substrate,⁴⁰ but was shown to be able to convert 2-acetamidoglucal into ManNAc.³⁵⁻³⁷ More evidence favoring the formation of 2-acetamidoglucal as the intermediate came from the study of the closely related non-hydrolyzing UDP-GlcNAc 2-epimerase (RffE).^{34,41} When incubating [$1''$ - ^{18}O]UDP-GlcNAc with RffE, a positional isotope exchange (PIX) was observed (Figure 1.10). The finding that the ^{18}O -label scrambled between the bridging and non-bridging positions in the β -phosphate of both epimers indicates that the glycosidic C-O bond is transiently cleaved during catalysis. 2-Acetamidoglucal and UDP was also found to be released into solution after extended incubation of RffE and UDP-GlcNAc. Together, these results strongly support Path B.

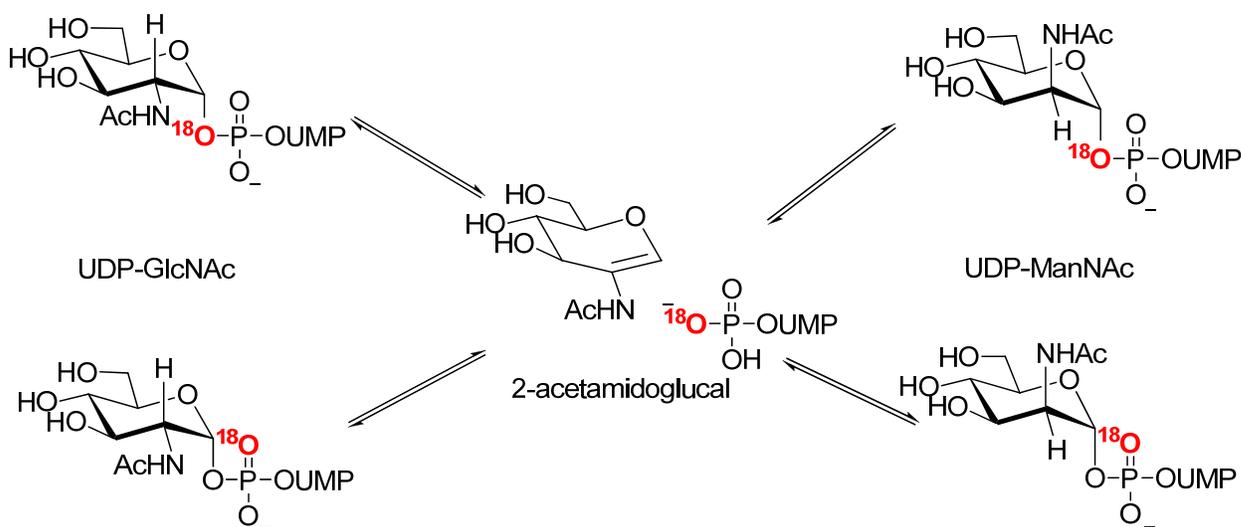


Figure 1.10 Positional isotopic exchange (PIX) catalyzed by RffE.

Mechanistic studies on the mammalian³⁵ and bacterial³⁶ hydrolyzing UDP-GlcNAc 2-epimerases also showed that UDP is cleaved via a C-O bond cleavage process to give α -

ManNAc. In addition, both enzymes could catalyze the hydration of 2-acetamidoglucal to give ManNAc. The mechanism is therefore confirmed to be Path B, where an *anti*-elimination of UDP forms the 2-acetamidoglucal intermediate and a following *syn*-addition of water gives α -ManNAc as the final product.

The mechanism of the bacterial hydrolyzing UDP-GlcNAc 2-epimerase will be mentioned again in Section 1.3.3 and Chapter 3 that outlines studies on a UDP sugar hydrolase involved in pseudaminic acid biosynthesis.

1.2.5 Sialic Acid Synthase

As described in Section 1.2.3, the mammalian and bacterial sialic acid synthases catalyze slightly different reactions. The mammalian Neu5Ac 9-P synthase catalyzes the condensation of ManNAc 6-P and PEP and produces Neu5Ac 9-P (Figure 1.11A). The bacterial sialic acid synthase, Neu5Ac synthase or NeuB, catalyzes the condensation between ManNAc and PEP to form Neu5Ac (Figure 1.11B). Both synthases require a divalent metal ion, such as Mn^{2+} or Mg^{2+} , for catalysis. Two other enzymes, 2-keto-3-deoxy-D-*manno*-2-octulosonic acid 8-phosphate synthase (KDO 8-P synthase) and 2-keto-3-deoxy-D-*arabino*-2-heptulosonic acid 7-phosphate synthase (DAH 7-P synthase) also catalyze very similar PEP-utilizing condensations to yield the

sialic acid-related molecules, 2-keto-3-deoxy-D-*manno*-2-octulosonic acid 8-phosphate (KDO 8-P, Figure 1.11C) and 2-keto-3-deoxy-D-*arabino*-2-heptulosonic acid 7-phosphate (DAH 7-P, Figure 1.11D), respectively. Extensive structural and mechanistic studies have previously been conducted on these two enzymes. Despite the fact that the sequence identity between sialic acid synthase and KDO 8-P/DAH 7-P synthase is very low (<10% for the *E. coli* enzymes),⁴² studies on the latter enzymes have proven to be useful in the study of sialic acid synthase.

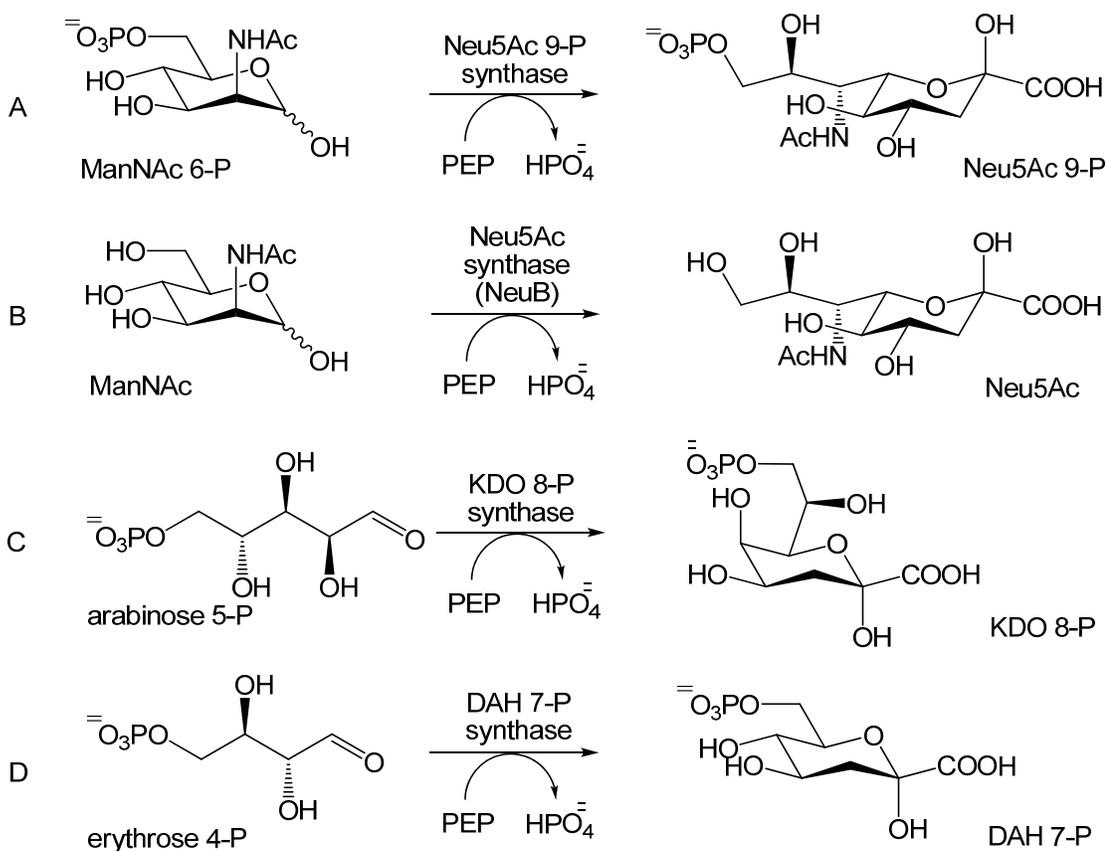


Figure 1.11 Reactions catalyzed by PEP-condensing synthases. A) Neu5Ac 9-P synthase, B) Neu5Ac synthase, C) KDO 8-P synthase, D) DAH 7-P synthase.

1.2.5.1 Structure of Sialic Acid Synthase (NeuB)

While structural and mechanistic studies have not yet been performed on the mammalian Neu5Ac 9-P synthase, the structure of its bacterial counterpart, NeuB, has been reported.⁴² A static light scattering experiment on purified *N. meningitidis* NeuB and a following X-ray crystallographic analysis of NeuB in a complex with *N*-acetylmannosaminitol (substrate that has been inactivated by reduction at C-1), PEP, and Mn²⁺ revealed its overall architecture. This enzyme was shown to exist as a domain-swapped homodimer (Figure 1.12A). Each monomer is composed of 349 residues and can be divided into two domains that are joined by a linker region (Figure 1.12B). The first domain consists of the N-terminal residues 1-273 and has the typical fold of a (β/α)₈ barrel (TIM barrel), which is a conserved protein folding of eight α -helices and eight parallel β -strands linked in an alternating manner. The tertiary structure of this domain is similar to those of KDO 8-P synthase and DAH 7-P synthase, indicating that these enzymes share a common fold and thus a similar active site conformation despite their low sequence identity. The second domain of NeuB, however, has never been seen in other PEP-utilizing enzymes such as KDO 8-P and DAH 7-P synthase. This “pretzel-shaped” domain consists of residues 285 to 349 and resembles the fish type III anti-freeze proteins (AFPs) as revealed by sequence and structural analysis.^{43, 44} It is located at the C-terminal end of NeuB and is linked to the N-terminal domain by a linker region consisting of residues 274-284.

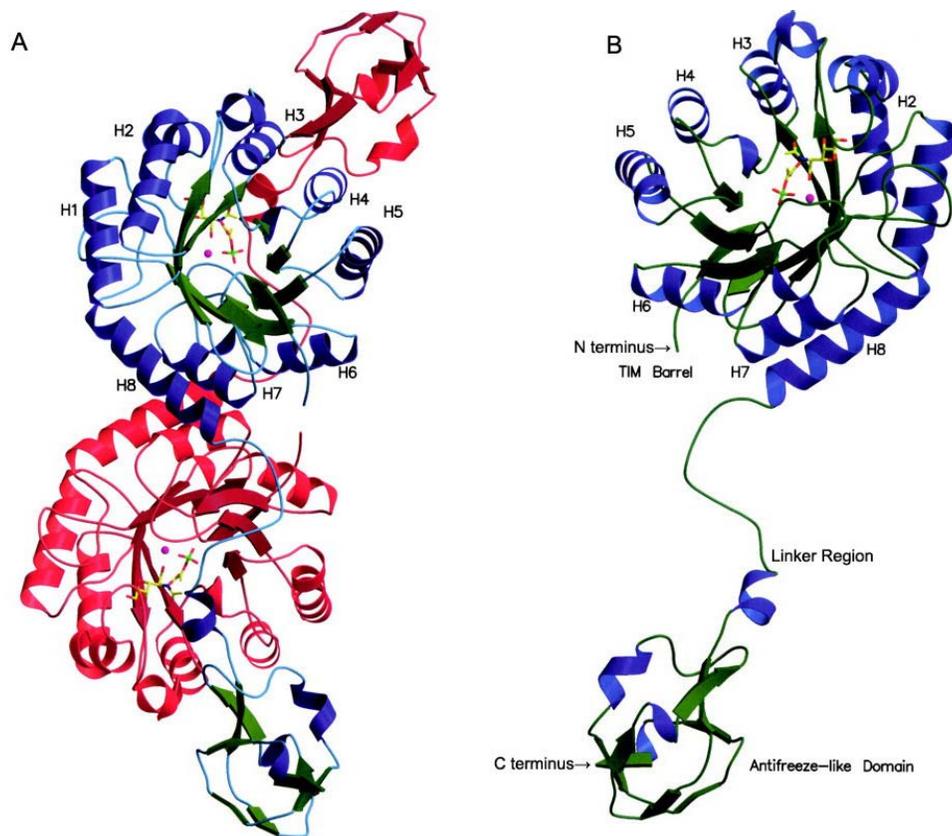


Figure 1.12 Overall architecture of NeuB. A) Overall structure of the NeuB domain-swapped homodimer with *N*-acetylmannosaminitol, PEP and Mn^{2+} in the active site. Monomers are shown in blue/green or red/orange. *N*-acetylmannosaminitol, PEP and Mn^{2+} are shown in stick representation with carbon atoms in yellow, oxygen atoms in red, nitrogen atoms in blue, phosphorus atoms in green, and Mn^{2+} in magenta. B) Structure of the NeuB monomer showing the TIM barrel and antifreeze-like domain (adapted from Gunawan, J. *et al.*, *J. Biol. Chem.* **2005**, 280 (5), 3555-3563, © 2009 American Chemical Society) .⁴²

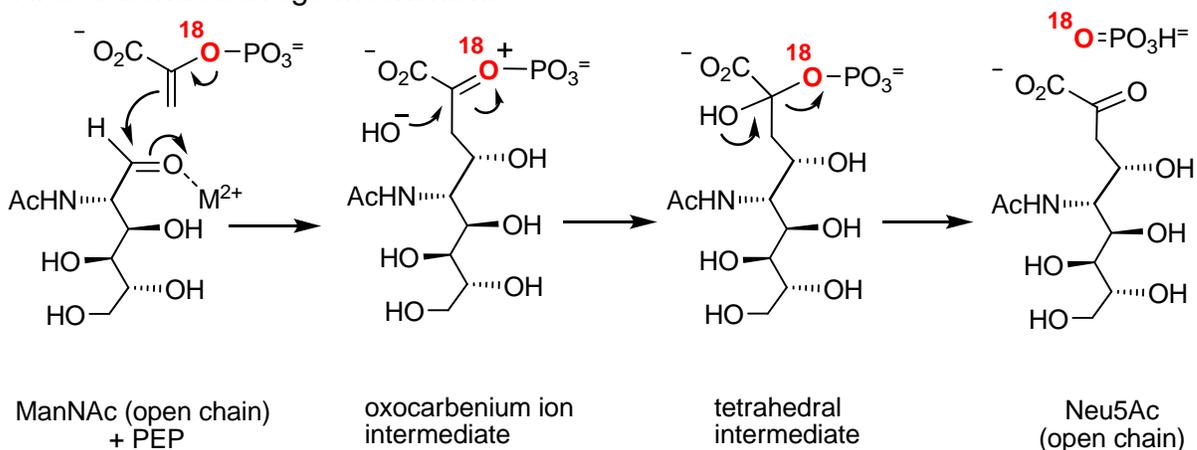
The homodimer of NeuB exists as two overlapping antiparallel monomers with the TIM barrel from one monomer closely packed with the antifreeze-like domain from the other monomer. The active site is formed by the C-terminal end of the TIM barrel, the helix and coiled region of the linker region and an extended 11-residue loop from the antifreeze-like domain. The structure of NeuB with *N*-acetylmannosaminitol, PEP and Mn^{2+} bound in the active site provides

valuable information that furthers our understanding of the mechanism of NeuB, which will be elaborated in the following section.

1.2.5.2 Mechanism of the Bacterial Sialic Acid Synthase (NeuB)

Two general mechanisms, C-O bond cleavage and P-O bond cleavage, can be proposed for the action of NeuB (Figure 1.13).⁴² In the C-O bond cleavage mechanism (Figure 1.13A), the enzyme first catalyzes the ring opening of ManNAc to expose the aldehyde functionality at C-1. The aldehyde is then attacked by the C-3 of PEP to give an oxocarbenium ion intermediate. Water then adds to the C-2 position of the oxocarbenium ion intermediate, resulting in the formation of a tetrahedral intermediate. The tetrahedral intermediate then collapses, releasing phosphate, and yielding the open chain form of Neu5Ac that readily cyclizes in solution. In contrast, the P-O bond cleavage mechanism begins with an initial attack of water at the phosphate group of PEP and release of the enolate anion of pyruvate. This highly reactive intermediate then adds to the C-1 aldehyde group of the open chain form of ManNAc to generate Neu5Ac (Figure 1.13B). Previous studies on the DAH 7-P and KDO 8-P synthases showed these enzymes utilized a C-O bond cleavage mechanism.^{45, 46} The NeuB mechanism was investigated by incubating the enzyme with [2-¹⁸O]-PEP, which contains an ¹⁸O-isotopic label in the phosphate bridging position (see labeled atoms in Figure 1.13).⁴² The ¹⁸O-label was found in the phosphate product but not in the Neu5Ac, confirming that the reaction proceeds through a C-O bond cleavage mechanism.

A. C-O bond cleavage mechanism



B. P-O bond cleavage mechanism

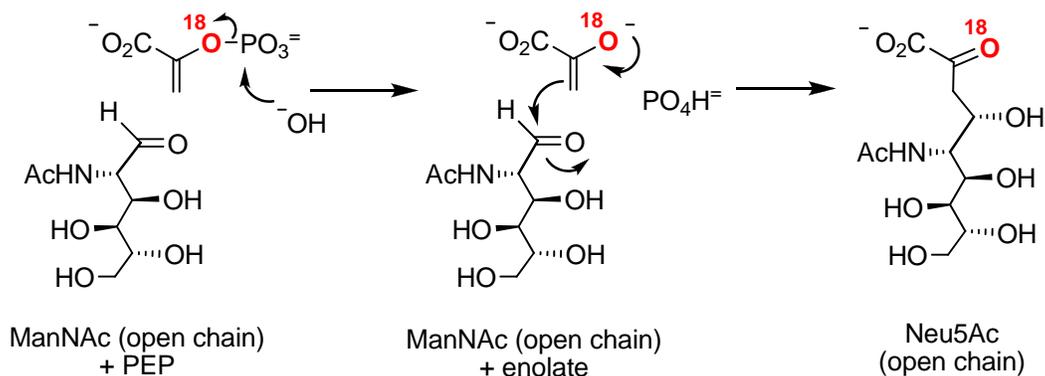


Figure 1.13 Proposed mechanisms for NeuB reaction.

Sundaram *et al.* uncovered the stereochemistry of the addition of PEP by using C-3-deuterated PEP analogues.⁴⁷ *C. jejuni* NeuB condenses *Z*-[3-²H]-PEP and *E*-[3-²H]-PEP with ManNAc and forms (3*S*)-3-²H-Neu5Ac and (3*R*)-3-²H-Neu5Ac, respectively. These results indicate that the *si*-face of the ManNAc aldehyde approaches the *si*-face of PEP during catalysis (Figure 1.14).

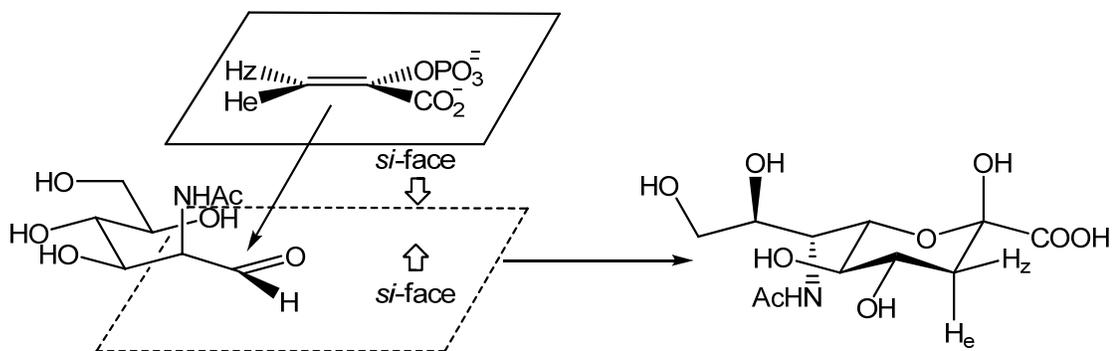


Figure 1.14 Stereochemistry of addition: *si*-face of the ManNAc aldehyde approaches the *si*-face of PEP.

A crystallographic analysis of the complex of *N. meningitidis* NeuB with *N*-acetylmannosaminitol (substrate reduced at C-1), PEP, and Mn^{2+} provided further support for the C-O bond cleavage mechanism and stereochemistry of this reaction (Figure 1.15). The crystal structure solved to a resolution of 2.2 Å serves as an excellent model for the Michaelis complex formed during catalysis. The C-1 hydroxyl of the substrate analogue is coordinated to Mn^{2+} , as would be expected if the metal ion activates the C-1 aldehyde via electrophilic catalysis in the C-O bond cleavage mechanism. The C-3 of PEP sits in the close proximity of the C-1 of *N*-acetylmannosaminitol. This is also consistent with the C-O bond cleavage mechanism where the first step involves C-C bond formation between those two carbons. In addition, the *si* face of the PEP was found positioned directly above the *pro-S* H-1 of *N*-acetylmannosaminitol, agreeing with the observed stereochemistry of the addition.

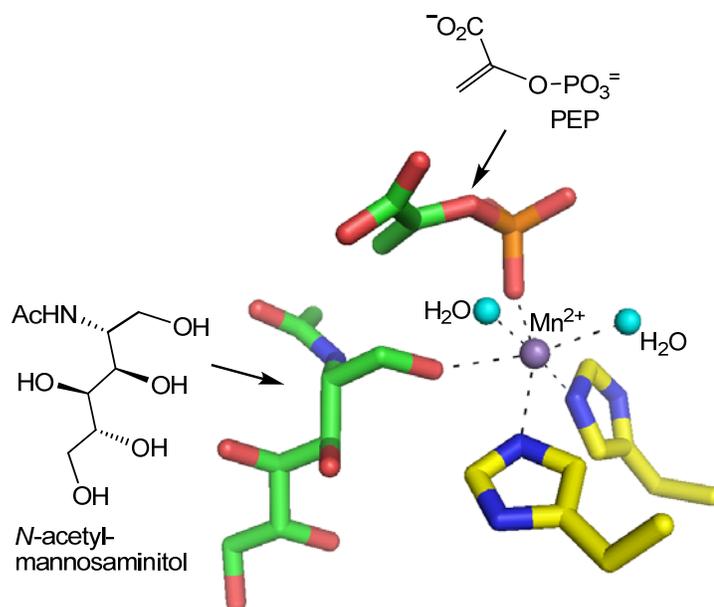


Figure 1.15 Active site of NeuB. *N*-acetylmannosaminitol, PEP and Mn^{2+} are shown in stick representation with carbon atoms in green, oxygen atoms in red, nitrogen atoms in blue, phosphorus atoms in orange, and Mn^{2+} in magenta. Water molecules are shown in cyan. Carbon atoms in active site residues are shown in yellow.

Evidence supporting the existence of the oxocarbenium ion intermediate was obtained in studies on KDO 8-P synthase.⁴⁸⁻⁵³ Aminophosphonate compounds (Figure 1.16, compound **1** and **2**) that mimic the oxocarbenium ion intermediate were synthesized and found to be micromolar inhibitors of KDO 8-P synthase. The transient tetrahedral intermediate in the KDO 8-P synthase-catalyzed reaction was also detected using time-resolved electrospray ionization mass spectrometry (time-resolved ESI-MS).⁵⁴ Grison *et al.* later synthesized mechanism-based tetrahedral intermediate analogues for KDO 8-P synthase (Figure 1.16, compound **3**) and DAH 7-P synthase (not shown in the figure); however, kinetics studies were not presented.⁵⁵

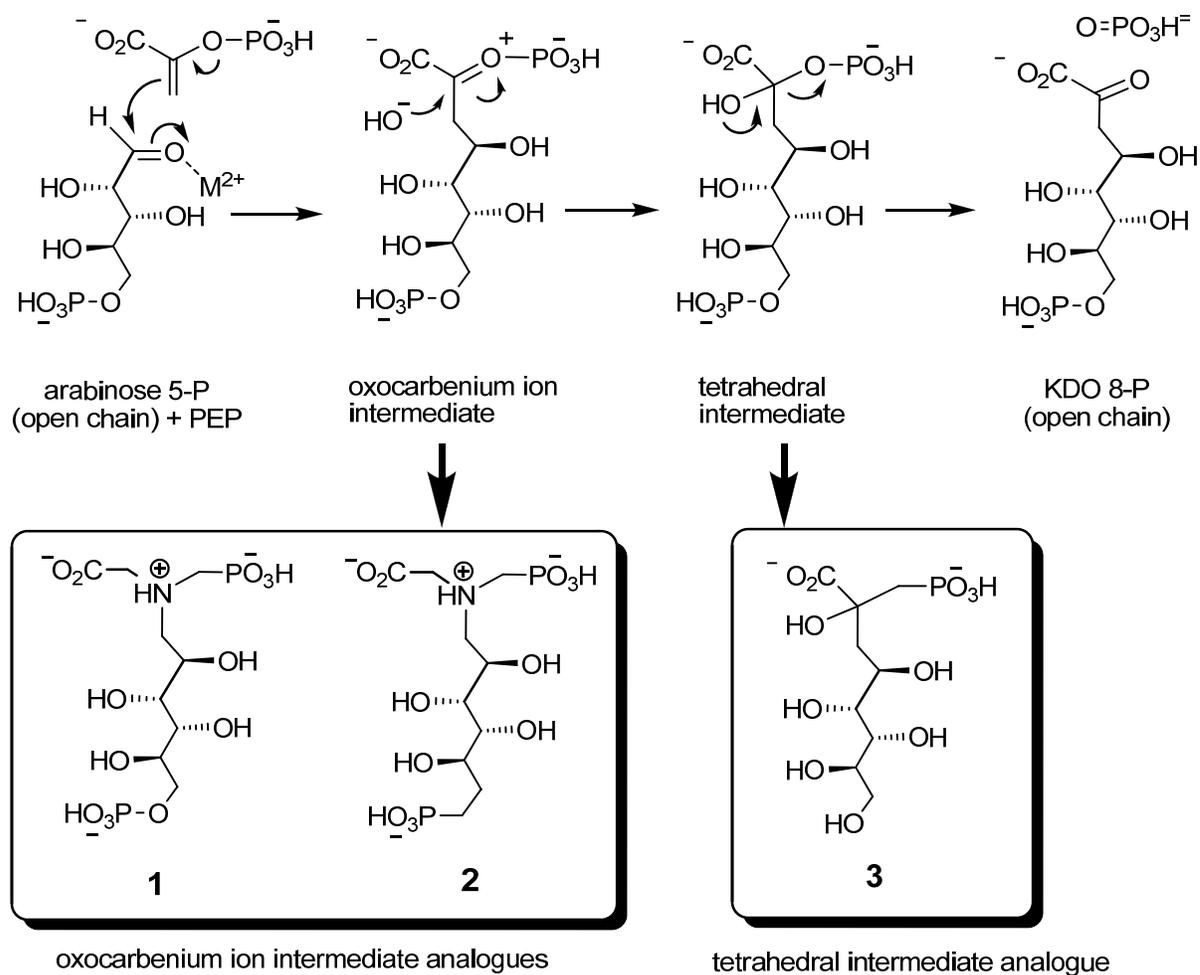


Figure 1.16 The mechanism of the reaction catalyzed by KDO 8-P synthase and structures of intermediate analogues.

Prior to the work described in this thesis, intermediate analogues had never been made for NeuB and at the time when my research project first started in 2004, the x-ray crystal structure of sialic acid synthase was not even available. Therefore, the goals of the project were to synthesize mechanism-based inhibitors and use them in the mechanistic and structural study of NeuB. Making an oxocarbenium ion intermediate analogue was the first goal of my research and will be elaborated in the first part of Chapter 2.

As the details of NeuB catalysis became clearer, one question remained to be solved: which stereochemical configuration does the tetrahedral intermediate have at C-2 (or from which face is the oxocarbenium ion intermediate attacked by the hydroxide)? This question was addressed by synthesizing and testing tetrahedral intermediate analogues and these studies will be presented in detail in the second half of Chapter 2.

1.3 Pseudaminic Acids

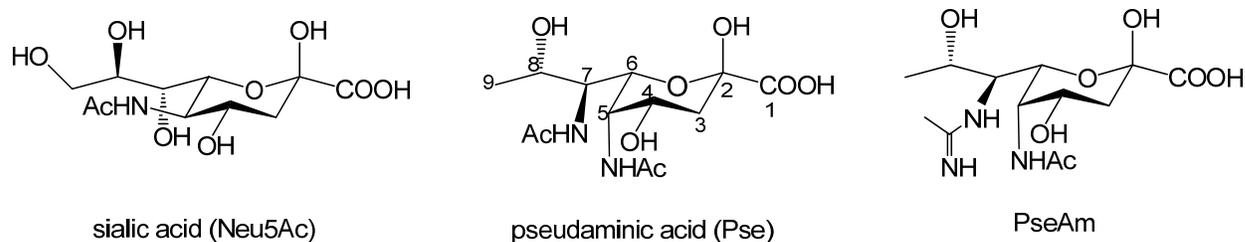


Figure 1.17 Structures of Neu5Ac, Pse and PseAm.

Pseudaminic acid (Pse) and its derivatives, such as the 5-acetamidino variation PseAm, are structurally similar to sialic acid (Figure 1.17). The structural differences are that pseudaminic acids have different stereochemical configurations at C-5, C-7 and C-8, an acetamido/acetamidino group at C-7, and no hydroxyl group at C-9. Pseudaminic acid was first discovered in the LPS of *Pseudomonas aeruginosa*.⁵⁶ Unlike the prevalence of sialic acid in mammals, pseudaminic acids are found exclusively in Gram-negative bacteria as components of important cell surface glycoconjugates, such as LPS, capsular polysaccharide⁵⁷, pili⁵⁸ and flagella^{59, 60}. In recent years, pseudaminic acids have received much attention as an important posttranslational modification of the flagellin protein. The *Campylobacter* flagellin glycosylation was first described in 1989⁶¹ and the carbohydrate moiety was then shown to be sensitive to periodic acid and capable of binding to a sialic-acid-specific lectin.⁶² Tandem mass spectrometry, preparative high performance liquid chromatography and nano-NMR spectroscopy later confirmed the identity of the modification on *Campylobacter* flagellin to be *O*-linked pseudaminic acid.⁵⁹

1.3.1 Flagella of *C. jejuni* and *H. pylori*

The flagella of *C. jejuni* and *H. pylori* not only confer motility to the bacteria and allow successful colonization in the alimentary tract, but are also shown to have roles in other processes such as biofilm formation and protein secretion.⁶⁰ The most important part of the flagellum is a glycoprotein called flagellin. Flagellin monomers are assembled as a hollow tube to form the filament which propels the cell by rotation. The filaments of *C. jejuni* and *H. pylori* are composed of two flagellin proteins, FlaA and FlaB, of which the former protein is the main component.² *C. jejuni* and *H. pylori* have unique mechanisms to mediate subunit interactions in the flagellar filament. The glycosylation of their flagellin proteins, unlike those of other bacteria, is absolutely required for the assembly and function of the flagella. *C. jejuni* and *H. pylori* mutants that cannot produce pseudaminic acid are found to be aflagellate and non-motile.^{60, 63}

1.3.2 Flagellin Glycosylation with Pseudaminic Acids

Protein glycosylation is considered to be an important post-translational modification in eukaryotic systems. Glycosylation alters the properties of proteins, by changing their stability, solubility and physical bulk. In addition, carbohydrates of glycoproteins act as recognition elements for proteins and other cells. While protein glycosylation is almost limited to eukaryotic systems, the flagellin proteins in some types of bacteria such as *C. jejuni* and *H. pylori* are found to be extensively glycosylated through an *O*-linkage. *C. jejuni* is among the most frequent causative agents of bacterial diarrhea worldwide. It is also the most common antecedent infection

that triggers Guillain-Barré syndrome, an acute autoimmune disorder causing neuromuscular paralysis.⁶⁴ The flagellin proteins of *C. jejuni* are heavily glycosylated with pseudaminic acids, which are responsible for a ~6000-Da (10%) enhancement in the molecular mass over the predicted protein mass from the DNA sequence.^{59, 65} *H. pylori*, a motile Gram-negative bacterium which colonizes more than half of the world's population,⁶⁶ also has extensively glycosylated flagellin.⁶⁰ This bacterium is of particular interest because it is the causative agent of chronic type B gastritis and is strongly associated with duodenal/gastric ulcers and gastric cancer.⁶⁷

In the FlaA or FlaB monomer of *C. jejuni*, up to 19 serine/threonine residues are found to be glycosylated with Pse, PseAm, and/or Leg5Am7Ac *etc.*⁵⁹ The sites of modification are restricted to the central domain of the sequence and are surface-exposed in the assembled filament. In *H. pylori*, up to 10 serine/threonine residues in the central region of the FlaA or FlaB molecule are glycosylated with Pse.⁶⁰ Compared to that of *C. jejuni*, the flagellin modification of *H. pylori* not only differs in the location of the modification, but also lacks heterogeneity in the composition of flagellar glycans.

Evidence suggests that the pseudaminic acid-containing *Campylobacter* flagellar glycans also play a key role in mediating autoagglutination (AAG) through surface-associated interactions of the glycan moiety.^{68, 69} AAG is considered the preliminary step in the formation of microcolonies⁷⁰ and biofilms,⁷¹ and is often associated with the virulence of bacteria.^{68, 72} A mutant strain of *C. jejuni* 81-176 that produces flagella containing Pse, but lacking PseAm, failed to autoagglutinate when compared to its parent strain. Furthermore, the adhesion and invasion abilities of the mutant strain were greatly compromised as evidenced by *in vitro* experiments and

in vivo studies using the ferret diarrheal disease model.⁶⁸ PseAm is thereby suggested to play a role in the pathogenesis of *C. jejuni* through microcolony formation.

As previously mentioned pseudaminic acids are only present in certain bacteria and are not produced or utilized by mammals. Since glycosylation with pseudaminic acid is critical for the assembly and consequent motility and colonization of certain types of gastrointestinal pathogens, elucidation of the pseudaminic acid biosynthesis pathway is of special interest and therapeutic importance.

1.3.3 Biosynthesis of Pseudaminic Acid

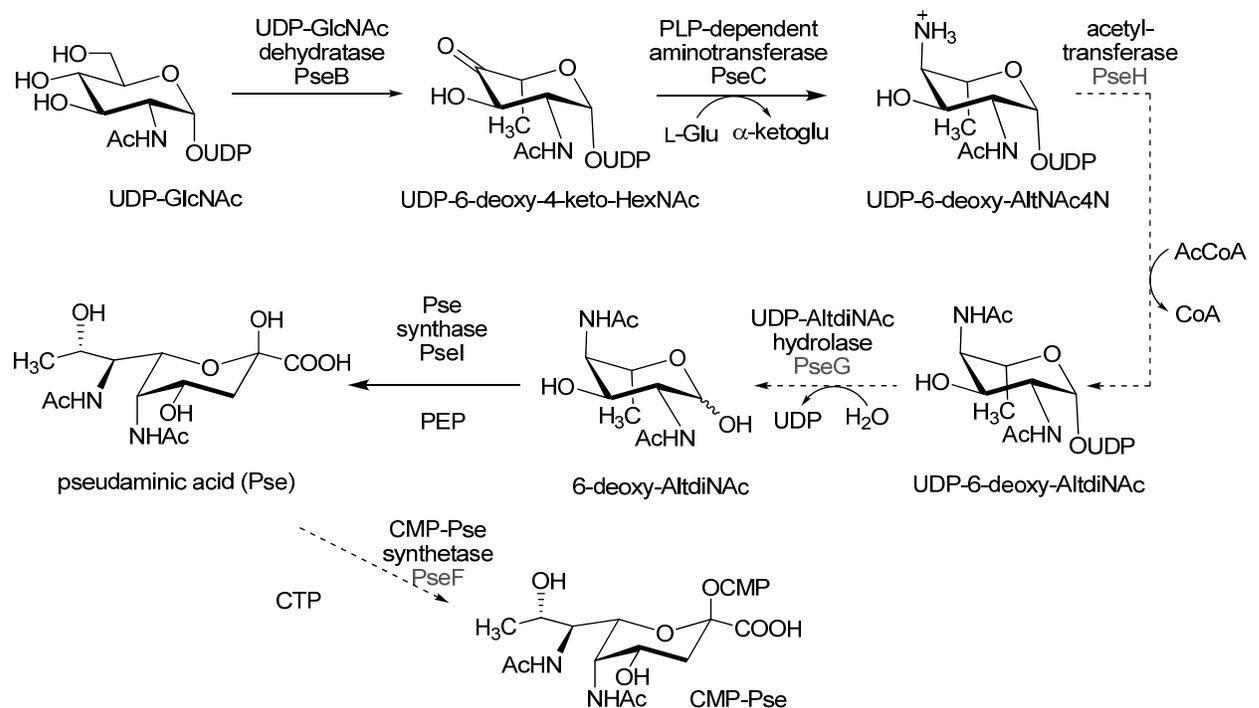


Figure 1.18 Biosynthesis of CMP-pseudaminic acid.

Comparative genomic analyses of *Campylobacter spp.* and *H. pylori* showed that a number of genes in the flagellar locus have significant homology to those involved in sialic acid biosynthesis.^{59, 73} Mutational knockout studies^{2, 63, 74} and *in vitro* characterization of recombinantly expressed gene products have now led to the full assignment of the enzymes involved in pseudaminic acid biosynthesis (Figure 1.18).^{75, 76} However, only three enzymes had been identified at the time the research project in this thesis started. Those not-yet-confirmed biosynthetic steps are shown in dashed arrows with the name of the enzymes in grey.

The biosynthesis of pseudaminic acid starts with the action of PseB, a UDP-*N*-acetylglucosamine (UDP-GlcNAc) dehydratase that installs a ketone group at C-4 and deoxygenates at C-6 to give UDP-2-acetamido-2,6-dideoxy-*L*-arabino-4-hexulose (UDP-6-deoxy-4-keto-HexNAc).⁷⁷ Unlike most known sugar nucleotide dehydratases, PseB also inverts the stereochemical configuration at C-5 and generates a sugar with the *L*-configuration. In the following step, a pyridoxal 5'-phosphate (PLP)-dependent aminotransferase, PseC, converts the C-4 carbonyl functionality to an amino group and produces UDP-4-amino-2,4,6-trideoxy-*L*-AltNAc (UDP-6-deoxy-AltNAc4N).^{78, 79} An acetyltransferase, PseH, then acetylates the 4-amino group using acetyl coenzyme A (AcCoA) and gives UDP-2,4-diacetamido-2,4,6-trideoxy-*L*-altrose (UDP-6-deoxy-AltDiNAc). After that, a UDP-AltDiNAc hydrolase cleaves the UDP-linkage and generates the free sugar 2,4-diacetamido-2,4,6-trideoxy-*L*-altrose (6-deoxy-AltDiNAc). Pseudaminic acid synthase, PseI, then takes the free sugar and condenses it with PEP to give pseudaminic acid (Pse).⁸⁰ Finally, a CMP-Pse synthetase, PseF, activates Pse as CMP-Pse and makes it ready for flagellar glycosylation.

We found the putative UDP-6-deoxy-AltDiNAc hydrolase particularly interesting, not only because it is an important enzyme in pseudaminic acid biosynthesis, but also because its

function somewhat resembles that of the hydrolyzing UDP-GlcNAc 2-epimerases that had been extensively studied in our research group. The latter enzyme also hydrolyzes the UDP linkage and provides the free sugars for the PEP-utilizing synthase in the next biosynthetic step. Considering the relative stereochemistry of sialic acid and pseudaminic acid, it was expected that UDP-6-deoxy-AltdiNAc hydrolase would not have the epimerizing ability. However, it may utilize a similar chemical mechanism involving glycal intermediates. We wished to discover whether the hydrolase is an ancestor of UDP-GlcNAc 2-epimerase that had lost its epimerizing ability during evolution, or if it originated from a completely different enzyme family? This riddle will be unraveled in Chapter 3.

1.4 Glycoconjugate Modification Using Bio-orthogonal Reagents

The discovery of UDP-6-deoxy-AltdiNAc hydrolase and the establishment of the Pse biosynthetic pathway inspired an idea: can we track pseudaminic acid in its native environment to gain further insight into related biological processes at a cellular level? Actually, scientists have long been seeking methods to track biomolecules in living systems. Some techniques have been developed to detect biomolecules such as proteins, glycans, lipids and nucleic acids within their native environments by tagging certain cellular components with bio-compatible reporters.

Tagging proteins with **green fluorescent protein (GFP)** and related protein fluorophores is one of the most popular strategies for cell imaging.^{81, 82} The gene of GFP can be genetically encoded and co-expressed as a fusion with that of the target protein. The fusion of the target protein with a GFP tag can be visualized by fluorescent microscopy and quantified by flow cytometry. GFP-tagging has become a powerful tool to monitor most cellular processes including the expression and transportation of proteins.⁸³⁻⁸⁶ GFP and related proteins have also been engineered to develop fluorescence resonance energy transfer (FRET) biosensors for the detection of Ca^{2+} ^{87, 88} and proteases^{89, 90}. Nonetheless, the use of GFP does have limitations. First of all it can only be used to visualize proteins and not other biomolecules that are not directly encoded in the genome. Secondly, GFP-like proteins are large in size and may affect the folding and functioning of the target protein.

Unlike GFP, **antibody reporters** can be applied to any kind of biomolecules that could potentially serve as epitopes. Proteins, glycans, lipids, nucleic acids and other organic metabolites can all be labeled with hybridoma or synthetic antibodies with high specificity.⁹¹⁻⁹³ However, because of the high specificity of the antigen-antibody interaction, a small change in

the epitope could greatly compromise the tagging efficiency. In this case, a new antibody conjugate needs to be regenerated to label every slightly different target molecule, and this process is expensive and time-consuming. Moreover, antibodies are proteins with relatively high molecular weight and may not have easy access to intercellular and extravascular target molecules.

In order to tag biomolecules within their native environment, a small reporter, that conjugates to its target under physiological conditions and does not interfere with any biological functionality, is highly desirable. In recent years, researchers have developed a family of small molecules that are named bio-orthogonal chemical reporters.

1.4.1 Bio-orthogonal Chemical Reporters

Bio-orthogonal chemical reporters are referred to as non-native and non-perturbing chemical modifications that easily conjugate to an exogenous probe in a specific manner.⁹⁴ Bio-orthogonal tagging typically includes two separate steps (Figure 1.19). The first step is to install a certain chemical functionality (reporter) onto the target molecule using the cell's own metabolic machinery such as protein expression or glycan biosynthesis. The chemical reporter is usually small so as to be tolerated by the cellular machinery. In the second step the reporter is covalently tagged with a small reactive probe through a rapid and highly selective chemical reaction under physiological conditions. Both chemical reporter and probe should be stable and inert to surrounding biomolecules to avoid any undesired side reactions. This strategy has been widely applied in tagging proteins, glycans, lipids and monitoring enzyme activities. Several chemical functionalities have been found to have ideal properties as bio-orthogonal reporters so far. These include ketone/aldehyde functional groups, tetracysteine peptides (CCxxCC), azides and terminal alkynes.

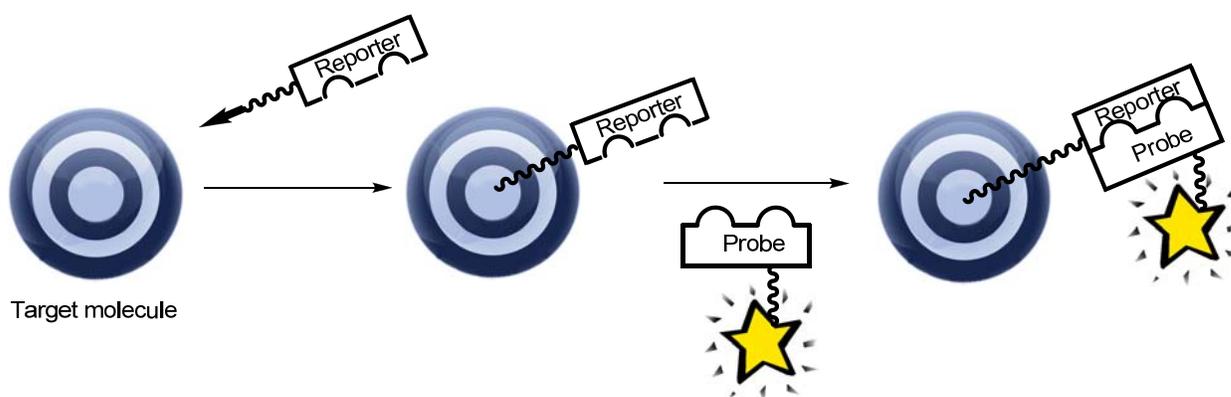


Figure 1.19 General bio-orthogonal chemical reporter strategy.

Ketones and aldehydes are not truly bio-orthogonal as there are abundant ketone- and aldehyde-containing metabolites present in biological fluids or within cells. However ketones and aldehydes are still considered suitable bio-orthogonal modifying functional groups in environments free of carbonyl electrophiles, such as cell surfaces or in extracellular locations. Ketones and aldehyde can be easily introduced into target molecules by feeding carbonyl-containing precursors to the cellular machinery. These electrophiles react readily with exogenous hydrazides and aminoxy groups to form Schiff bases that are stable under physiological conditions (Figure 1.20).

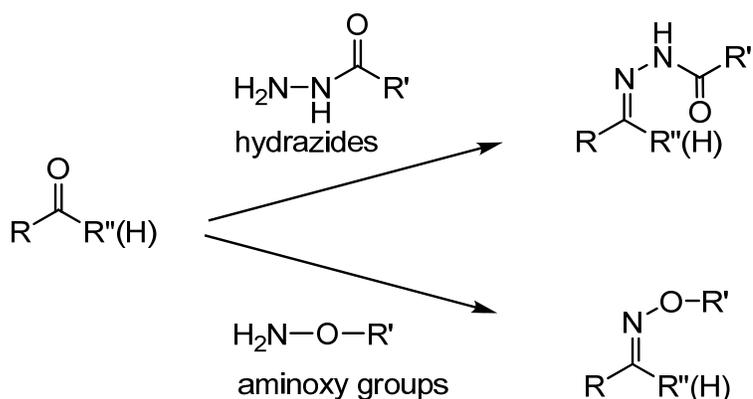


Figure 1.20 Ketone/aldehyde reporters and their ligation with hydrazide/aminoxy probes.

The most widely used strategy for labeling carbohydrate-containing biomolecules is to metabolically convert unnatural ketone-bearing precursors into the targets of interest. Bertozzi *et al.* fed *N*-levulinoylmannosamine (ManLev), a ketone-bearing derivative of ManNAc (a precursor of sialic acid), to cultured sialic-acid-producing human cells. The corresponding *N*-levulinoyl sialic acid (SiaLev) was displayed on the cell surface, labeled with biotin hydrazide, stained with fluorescein-avidin and analyzed by flow cytometry.^{95, 96} Sadamoto and coworkers also introduced ketones into the bacterial cell wall by using ketone-containing analogues of UDP-*N*-acetylmuramic acid-pentapeptide (UDP-MurNAc-pentapeptide).⁹⁷

For the labeling of proteins, genetic approaches are usually employed. Zhang and coworkers incorporated *m*- and *p*-acetyl-L-phenylalanine into a target protein by using a host-orthogonal tRNA-TyrRS (tyrosyl-tRNA synthase) pair that responds to the amber nonsense codon.⁹⁸ Chen *et al.* labeled recombinant proteins containing a 15-amino-acid acceptor peptide with a synthetic ketone-containing derivative of biotin using biotin ligase (BirA).⁹⁹ Recently, Rush and Bertozzi reported a new strategy to introduce an aldehyde group into target proteins without using any unnatural substrate.¹⁰⁰ In this case, a short peptide containing a CxPxR motif is firstly genetically installed on the target protein (x refers to a random amino acid). The formylglycine generating enzyme (FGE) is known to activate the enzyme sulfatase by converting cysteine present in this motif to a C α -fomylglycine (FGly) residue.¹⁰¹ It is also able to recognize the CxPxR motif in a genetically modified target protein and oxidizes the cysteine thiol group to an aldehyde.

Nevertheless, as mentioned earlier, ketone and aldehyde reporters are not suitable for use in cells and biological fluids that contain abundant keto and aldehydic metabolites. Moreover,

the optimum pH of the condensation between a ketone/aldehyde and hydrazides/aminooxy groups is 5 to 6, which is hard to achieve in many physiological conditions.

The **tetracysteine reporter** was developed by Tsien and coworkers to image proteins.¹⁰²⁻¹⁰⁴ They genetically incorporated a short peptide sequence with a CCxxCC motif into the target protein (x refers to a noncysteine amino acid). This short peptide reporter contains two pairs of thiols and covalently binds to nonfluorescent fluorescein/resorufin-based arsenical helix ligand (FAsH¹⁰³/ReAsH¹⁰⁴) to form strongly fluorescent complexes. Bis-1,2-ethanedithiol (EDT₂) adducts of FAsH and ReAsH are used in the labeling experiment to avoid tagging any isolated cysteine residues (Figure 1.21). The genetically encoded tetracysteine reporter causes minimal structural perturbation and is considered a good reporter for a wide range of proteins.

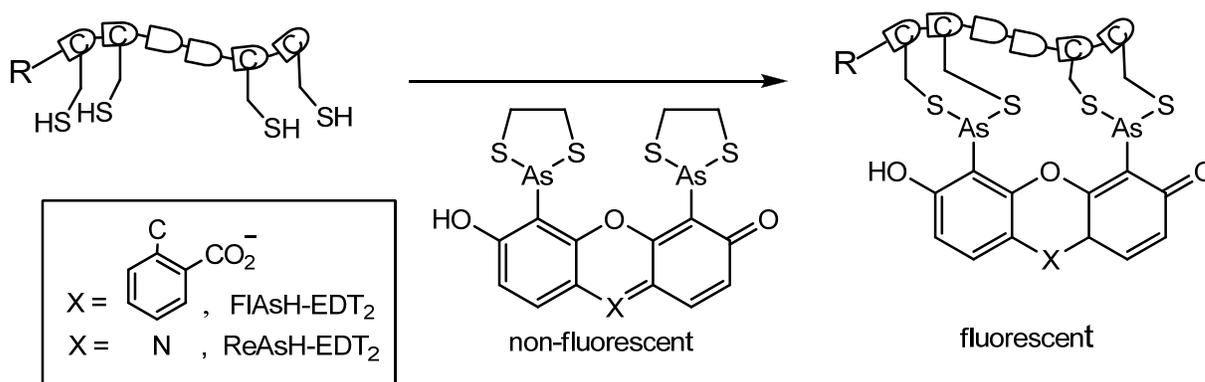


Figure 1.21 Tetracysteine reporter and its ligation with FAsH-/ReAsH-EDT₂.

Azides have received much attention in the past decade and are considered the most versatile chemical reporters that are suitable for labeling any kind of biomolecules in any biological environment. Azides are almost absent from all living systems. They are small in size and bring nearly no structural perturbation. Moreover, contrary to a general misconception,

organic azides are nontoxic.¹⁰⁵ They are stable and do not react with any naturally occurring biomolecules and other metabolites in physiological environments. In attempts to exploit the use of azides in biological labeling, Bertozzi and coworkers were inspired by the classic Staudinger reaction and designed a so-called Staudinger ligation where azides readily ligate with phosphine reagents under extremely mild conditions (Figure 1.22).¹⁰⁶ This method was used as a powerful tool to detect azide-bearing sugars on cell surfaces and will be further discussed in Section 1.4.2.

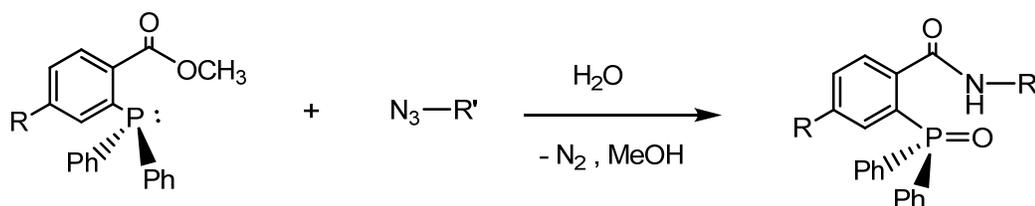


Figure 1.22 The Staudinger ligation.

Proteins can also be labeled with azides residue-specifically or site specifically. In a residue-specific labeling of a protein, auxotrophic cells that are unable to produce a certain kind of amino acid are used. The target protein can only be over-expressed when the cells are supplemented with azido-bearing analogues of the amino acid and will therefore be tagged with azides at sites that encode for the corresponding amino acid.^{107, 108} Proteins can also be labeled in a site-specific way where the azide-containing amino acid analogue is incorporated at a single position, rather than at multiple sites. In this case, a tRNA and aminoacyl-tRNA synthase pair is designed and incorporates the unnatural azido amino acid in response to an encoded amber nonsense codon.^{109, 110} In addition to glycans and proteins, lipids¹¹¹ and nucleic acid-derived cofactors¹¹² are also labeled with azides. These azide-labeled biomolecules can be subsequently tagged by reporters using Staudinger ligation with phosphines or by using “click chemistry” with terminal alkynes.

The concept of “click chemistry” was first proposed by Sharpless.¹¹³ It defines carbon-heteroatom bond-forming reactions that are simple, clean, fast, carried out under mild conditions and in benign solvents (such as water) and can be widely applied. One of the best known click reactions is the azide-alkyne Huisgen [2+3] cycloaddition (Figure 1.23A).¹¹⁴ Aliphatic azides rapidly react with terminal alkynes to form a triazole derivative in the presence of a Cu(I) catalyst. The reaction can be carried out in aqueous solution and neutral pH and is free of any byproduct. This method has been widely used to detect azide labels.^{108, 115-117} However, the Cu(I)-catalyzed azide-alkyne ligation is not an ideal click reaction for use in biological systems because of the drawback that Cu(I) is cytotoxic. Bertozzi and coworkers designed a catalyst-free, strain-promoted azide-alkyne cycloaddition that is suitable for the covalent modification of biomolecules in living systems (Figure 1.23B).¹¹⁸ A cyclooctyne derivative was shown to selectively react with azides, but at a slower rate than that of the Cu(I)-catalyzed reaction. Following research showed that the rate can be greatly enhanced by using a difluorinated cyclooctyne derivative.¹¹⁹

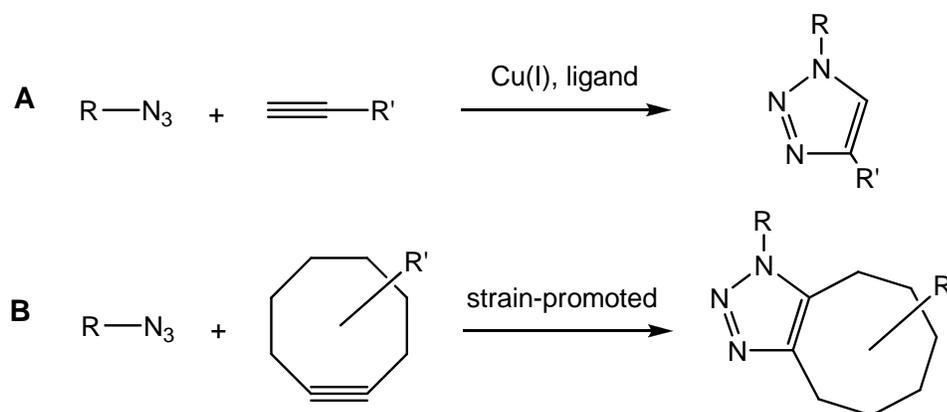


Figure 1.23 Azide-alkyne click chemistry. A) Cu(I) catalyzed [2+3] cycloaddition, B) strain-promoted [2+3] cycloaddition.

Alkynes, in turn, can also be used as bio-orthogonal reporters and be detected by azido-bearing probes.¹²⁰ They are inert in biological systems and, like azides, are well tolerated by the cellular machinery.

1.4.2 Staudinger Ligation and Glycoconjugate-Labeling with Azides

In the classic Staudinger reaction, an azide rapidly reacts with a phosphine to form an aza-ylide, which spontaneously hydrolyzes into a primary amine and the corresponding phosphate in the presence of water (Figure 1.24).¹²¹ As a highly reactive nucleophilic intermediate, the aza-ylide can also react with a variety of other electrophilic compounds, including aldehydes, amides and isocyanato compounds to give a variety of C-N ligated products.

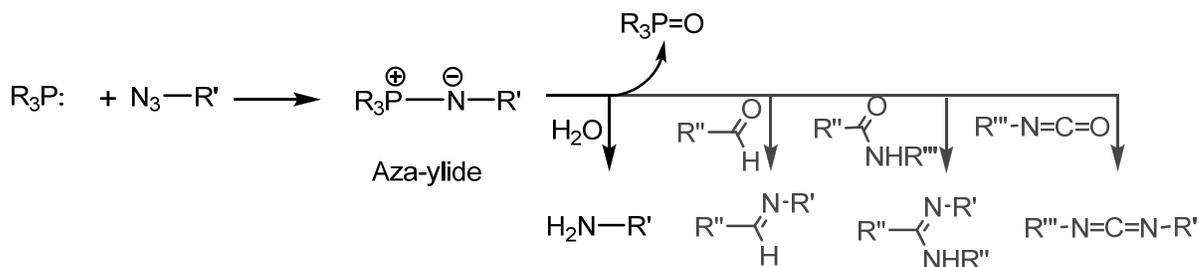


Figure 1.24 Classic Staudinger reaction (shown in black) and the reaction of aza-ylide with other electrophilic compounds (shown in grey).

Bertozzi and coworkers, while seeking a chemo-selective ligation method to be used in living systems, were inspired by this rapid and specific reaction. Both azides and phosphines are largely absent from biological systems and are unreactive towards biomolecules and other metabolites under physiological conditions. The small azido group is very likely to be tolerated by cellular machinery and therefore has the potential as a good bio-orthogonal reporter. In order

to overcome the instability of aza-ylide towards hydrolysis, a modified phosphine reagent containing a methyl ester within the proximity of the phosphine functionality was developed (Figure 1.25). The ester serves as an electrophilic trap to capture the aza-ylide intermediate by an intramolecular cyclization. A stable amide is formed as the ligation product as shown by *in vitro* model reactions.¹⁰⁶ The reaction was later shown to proceed via a bridged bicyclic oxazaphosphetane intermediate (Figure 1.25, lower route).¹²²

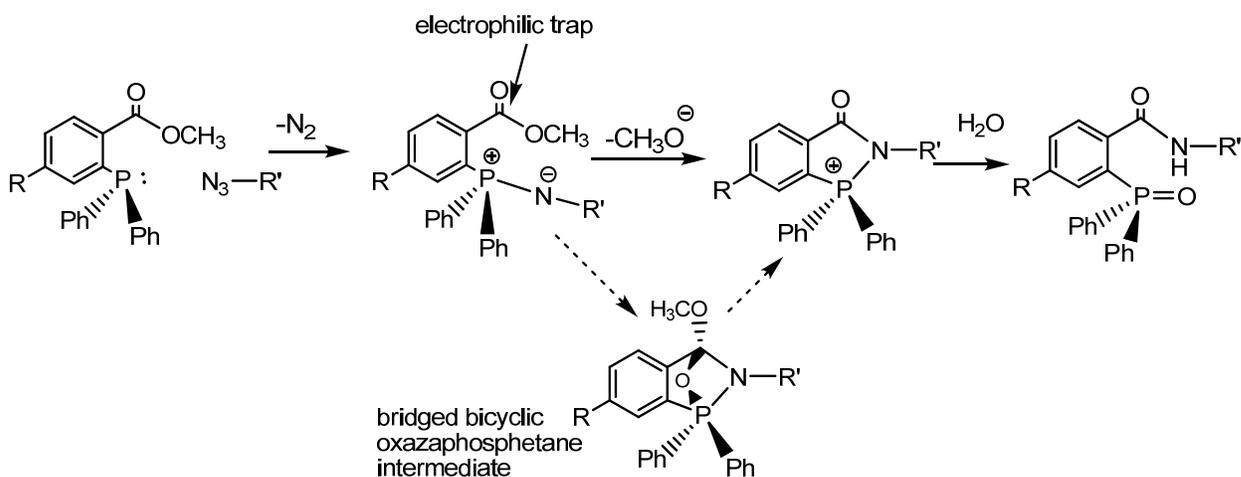


Figure 1.25 Mechanism of the Staudinger ligation.

The mammalian sialic acid biosynthetic machinery, a system of great interest, was used as the model to test the feasibility of this labeling strategy. A biotinylated tetraethyleneglycol linker bearing a terminal biotin functionality was installed on the phosphine reagent for enhanced water solubility and ease of detection (Figure 1.26). An azido analogue of ManNAc was peracetylated prior to incubation with Jurkat cells, to improve its trans-membrane migration. Once inside the cell, the peracetylated ManNAc can be deacetylated by cytosolic esterases.¹²³ After incubation with the azido analogue, the Jurkat cells were treated with the biotinylated

phosphine reagent, stained with fluorescein isothiocyanate (FITC)-avidin and analyzed by flow cytometry (Figure 1.26). It was shown that azido ManNAc was well tolerated by the sialic acid biosynthetic machinery and the resulting azido-bearing sialic acid was synthesized and displayed on the cell surface.¹²⁴ Similar studies using GlcNAc azido analogues have been used to detect *O*-GlcNAc-modified proteins in cells.¹²⁵

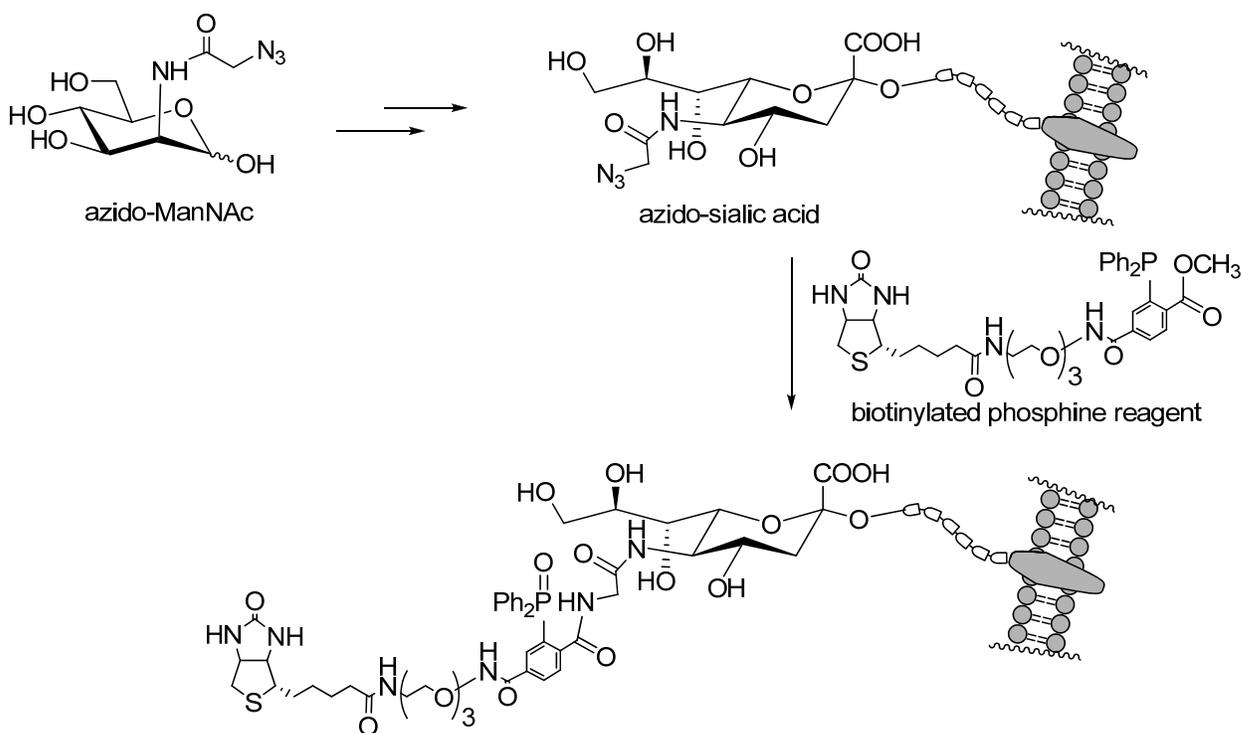


Figure 1.26 Metabolic delivery of azides to cell surfaces by unnatural sialic acid biosynthesis and the tagging of azides with a biotinylated phosphine reagent through the Staudinger ligation.

The Staudinger ligation is now widely used for the detection of azide labels in a variety of biomolecules. Although phosphine is liable to oxidation in air, the Staudinger ligation does not require any cytotoxic catalyst, as compared to the Cu(I)-catalyzed azide-alkyne [2+3] “click” cycloaddition, and still proceeds at a satisfactory rate. The recent commercial availability of 2-

(diphenylphosphino) terephthalic acid 1-methyl 4-pentafluorophenyl diester, a precursor of Bertozzi's Staudinger reagent, further makes this strategy a preferable option for many chemists.

As a part of my research, the *in vitro* and *in vivo* azide-labeling of pseudaminic acid followed by a Staudinger ligation-type tagging was employed in bacterial flagella engineering and will be presented in Chapter 4.

1.5 Project Goals

The aims of this thesis are to identify and/or study the mechanism of two unique enzymes involved in related biosynthetic pathways and to investigate the possibility of extending modern cell-labeling technology into the biosynthesis of other higher order sugars. The mechanistic study of sialic acid synthase using synthesized inhibitors will be the focus of Chapter 2. In Chapter 3, PseG, an enzyme involved in pseudaminic acid biosynthesis, will be identified and studied. Finally, as an offshoot of our studies on pseudaminic acid biosynthesis, flagella engineering will be elaborated in Chapter 4.

Sialic acid synthase (NeuB) catalyzes the condensation of ManNAc and PEP to give a nine-carbon α -keto acid, sialic acid, which plays important roles in mammals and some bacteria. Intermediate analogue inhibitors for NeuB have never been synthesized before. In Chapter 2, we will synthesize and test NeuB inhibitors that mimic the proposed oxocarbenium ion and tetrahedral intermediates formed during catalysis. Moreover, although much is understood about the mechanism of NeuB, the stereo configuration at C-2 of the tetrahedral intermediate still remains unclear. We plan to determine the stereochemistry of the tightest-binding isomer of the tetrahedral intermediate analogue using crystallographic analysis. The stereochemistry of the tetrahedral intermediate will be unveiled for the first time.

The work in Chapter 3 will be focused on PseG, the UDP-6-deoxy-AltdiNAc hydrolase involved in the biosynthesis of pseudaminic acid. This previously unidentified enzyme will be cloned and characterized. Its biochemical characteristics will be studied using chemo-enzymatically synthesized substrate, an isotope incorporation experiment, NMR spectroscopy,

mass spectrometry and a coupled kinetic assay. We wish to determine whether it is related to UDP-GlcNAc 2-epimerase or to glycosyl transferases.

Following previous examples of *in vivo* carbohydrate labeling, the elucidation and relevant studies of the pseudaminic biosynthetic pathway presents a new and exciting opportunity for cell surface engineering. The goal of Chapter 4 is to feed a mutant strain of *C. jejuni*, that is deficient in pseudaminic acid production, with a chemo-enzymatically synthesized azido-bearing Pse precursor in order to incorporate an azido functional group onto *C. jejuni* flagella. The surface-displayed chemical handles could provide endless possibilities for further studies.

Chapter 2 : The Inhibition of *Neisseria meningitidis* Sialic Acid Synthase by Intermediate Analogues^a

^a A version of this chapter has been published, some figures are reproduced with permission from:
Feng Liu, Ho Jun Lee, Natalie C. Strynadka, Martin E. Tanner, (2009) "Inhibition of Sialic acid Synthase by a
Tetrahydro Intermediate Analogue", *Biochemistry* 48: 9194-9201. © 2009 American Chemical Society.

2.1 Introduction

As described in Section 1.2.5, the reaction catalyzed by NeuB has been proposed to proceed through a mechanism involving the initial formation of an oxocarbenium ion intermediate and a subsequent formation of a tetrahedral intermediate. Evidence supporting the existence of the corresponding intermediates in the reactions catalyzed by KDO 8-P synthase and DAH 7-P synthase has been reported previously. Aminophosphonate analogues that mimic the structure of the oxocarbenium ion intermediate were synthesized and studied.⁴⁸⁻⁵³ Baasov and co-workers achieved the synthesis of an oxocarbenium ion analogue, aminophosphonate compound **1**, by employing a one-pot direct reductive amination between D-mannose 6-phosphate (Man 6-P) and commercially available phosphonomethyl glycine (glyphosate) in the presence of NaBH₃CN (Figure 2.1).^{48,49} This compound contains a positively charged nitrogen atom in the place of the corresponding C-2 carbon of the oxocarbenium ion intermediate. Although the nitrogen atom is sp³-hybridized instead of sp²-hybridized, this analogue presumably functions by mimicking the electrostatic properties of the oxocarbenium ion intermediate rather than the geometry of the tetrahedral intermediate. Other structural changes include the replacement of a phosphate oxygen atom with a methylene group to give a phosphonate functionality, and the addition of an extra methylene group between the carboxylate and the central nitrogen atom. These modifications impart stability to the analogue while introducing minimal changes in steric bulk. Compound **1**⁴⁸, as well as another analogue **2**⁵³ with the C-6 oxygen replaced by a carbon atom, proved to be effective inhibitors of KDO 8-P synthase from several species. Compound **1** was successfully co-crystallized with KDO 8-P synthase for structural analysis.

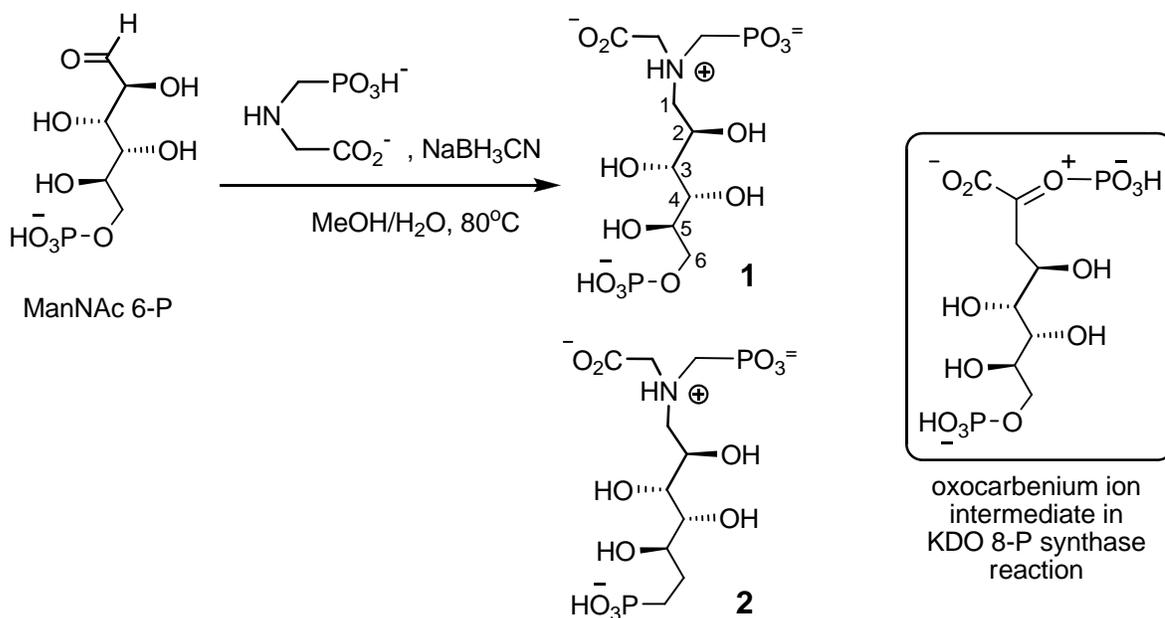


Figure 2.1 Structures of KDO 8-P synthase intermediate analogues 1 and 2 and the one-pot direct reductive amination for the synthesis of compound 1.

Prior to this study, inhibitors of NeuB had never been synthesized. An aminophosphonate-based oxocarbenium ion intermediate analogue (Figure 2.2, **O1**) that resembles the ones tested in the KDO 8-P synthase studies, was our first target. If analogue **O1** serves as a potent competitive inhibitor of the reaction catalyzed by NeuB, the proposed mechanism involving an oxocarbenium ion intermediate will be further supported. In addition, it may be possible to analyze the structure of a complex between NeuB and analogue **O1** using X-ray crystallography, which would help provide insights into the mechanism of catalysis. A second goal of this chapter was to synthesize analogues of the tetrahedral intermediate.

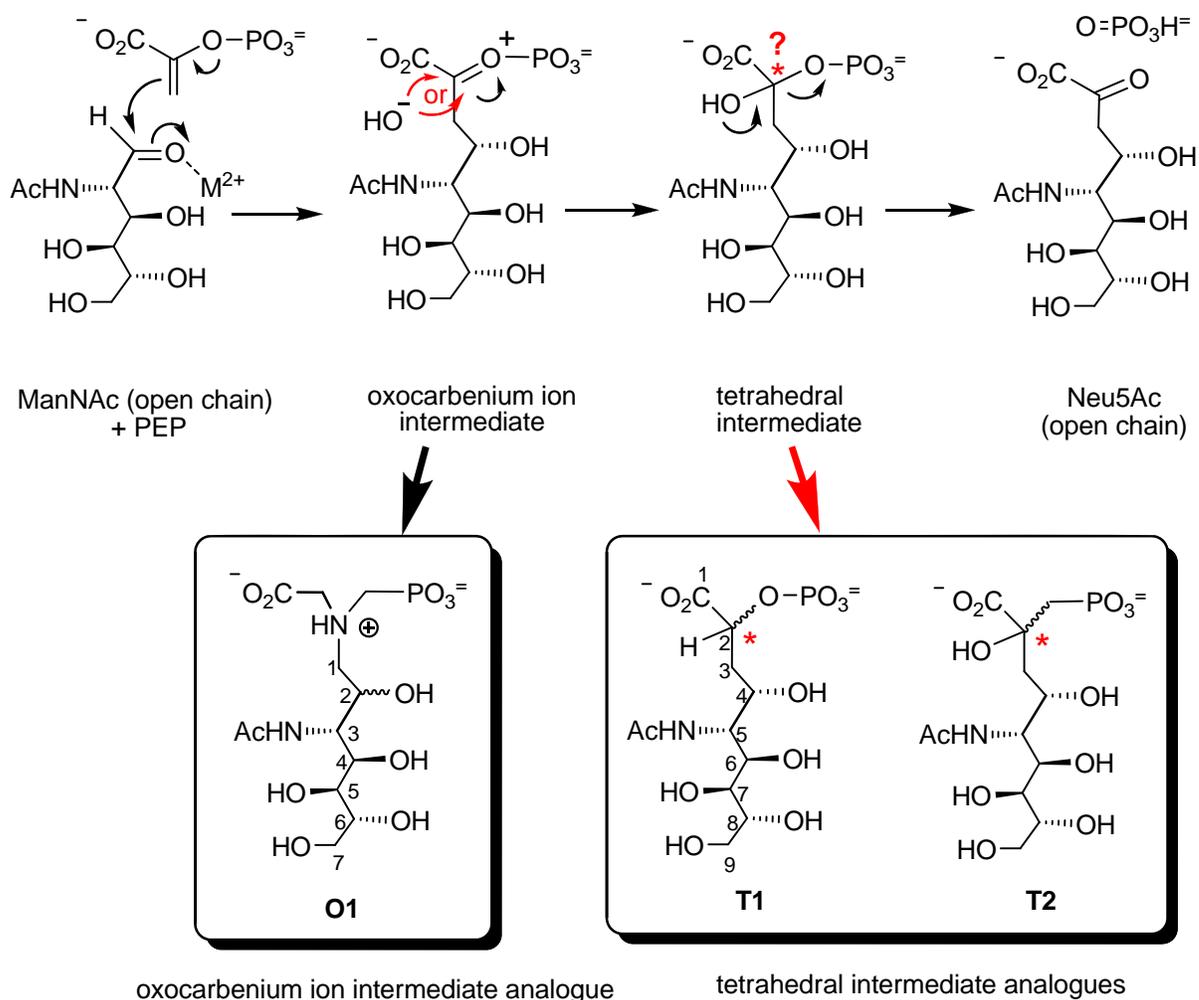


Figure 2.2 Intermediate analogues proposed as potential inhibitors of NeuB.

Two candidates, **T1** and **T2** (Figure 2.2), could be envisioned as analogues of the NeuB tetrahedral intermediate and would be expected to serve as inhibitors of NeuB. In **T1**, the C-2 hydroxyl group has been removed and in **T2**, a methylene has been installed in the place of the oxygen that connects C-2 and the phosphorus atom. These changes impart stability to **T1** and **T2** so that they will not undergo the spontaneous loss of phosphate that would be expected for the actual tetrahedral intermediate. The electrostatic and steric properties resemble those of the intermediate, giving these analogues good potential to serve as NeuB inhibitors. An additional

benefit of studying tetrahedral intermediate analogues is the potential opportunity to reveal the stereochemistry at C-2 in the actual tetrahedral intermediate. No previous studies have ever shown which face of the oxocarbenium ion is attacked by the solvent-derived hydroxide, and therefore, what stereochemistry the resultant tetrahedral intermediate has at C-2. In Section 2.3, the synthesis of analogue **T1** as a mixture of two epimers differing in stereochemistry at C-2 is described. One of the epimers is expected to bind to the NeuB active site much more tightly than the other. Analogue **T1** was found to be a NeuB inhibitor as determined by a kinetic study, and the NeuB•**T1** complex was subjected to crystallographic analysis. It will be shown that the more tightly binding epimer of **T1** is of the (2*R*)-configuration, suggesting that the tetrahedral intermediate also has the (2*R*)-configuration. The mechanistic implication of this finding is that the metal ion delivers a hydroxide ion to the *si* face of the oxocarbenium ion.

2.2 Over-expression and Purification of *Neisseria meningitidis* Sialic Acid Synthase (NeuB)

The *N. meningitidis* NeuB used in the determination of the inhibition kinetics and that used in the crystallographic studies were over-expressed and purified differently. A His₆-tagged NeuB was employed in the kinetic study due to the relative ease of purification. The gene encoding for the *N. meningitidis* NeuB had previously been inserted into a Novagen pET-30 Xa/LIC expression vector.⁴² The pET-30 Xa/LIC vector contains a gene encoding for kanamycin-resistance, a T7 promoter followed by a down-stream *lac* operator, and the site for insertion of the desired open reading frame (ORF). High level expression can be induced by isopropyl- β -D-1-thiogalactopyranoside (IPTG) in host strains that produce T7 DNA polymerase, such as the *E. coli* BL21 (DE3) used in this case. The over-expressed recombinant protein has a His₆-tag, an S-Tag and a Factor Xa cleavage site at the *N*-terminus of the target protein and can be easily purified by Ni-affinity chromatography. The purified His₆-NeuB was > 95% pure as shown by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, Figure 2.3).

As the wild-type NeuB is preferred for structural analysis, a NeuB construct without any tags was employed by our collaborator in the laboratory of Professor Strynadka at UBC for crystallographic studies. The NeuB gene had previously been cloned into the pCWOri+ vector, which contains an ampicillin-resistance gene and *tac* promoter (hybrid of the *E. coli trp* and *lac* promoters¹²⁶) for high-level expression.⁴² NeuB was over-expressed in the presence of ampicillin and IPTG and purified using ion-exchange and gel filtration chromatography.

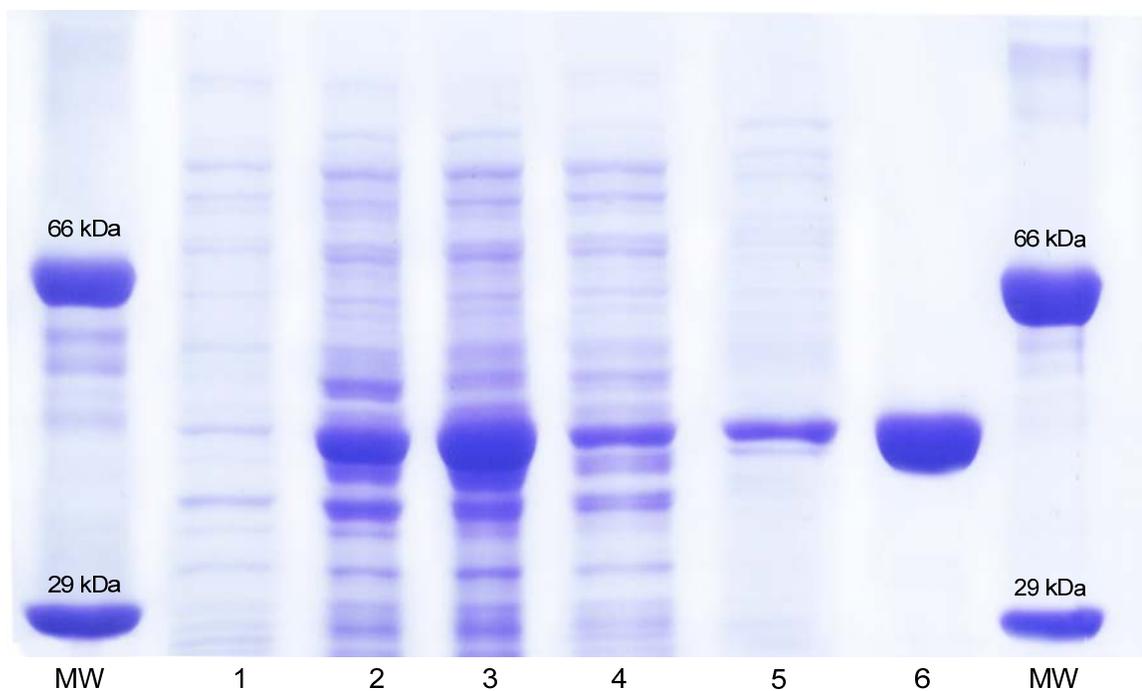


Figure 2.3 SDS-PAGE for the over-expression and purification of His₆-NeuB. 1) *E. coli* cells before induction, 2) after induction with IPTG, 3) crude cell lysate, 4) Ni-affinity column eluate with 20 mM imidazole, 5) column eluate with 100 mM imidazole, 6) column eluate with 500 mM imidazole.

2.3 Synthesis of, and Inhibition by, the Oxocarbenium Ion Intermediate Analogue (O1)

2.3.1 Synthesis of Analogue O1

Analogue **O1** was synthesized using a one-pot reductive amination procedure similar to the one previously described by Du⁴⁸ and Belakhov⁵³ in the preparation of KDO8P synthase inhibitors (Figure 2.4). ManNAc was first subjected to a Urech cyanohydrin reaction,¹²⁷ where the aldehyde group of the open-chain form of ManNAc underwent nucleophilic attack by potassium cyanide in an aqueous solution. This one-carbon-extension reaction produced a mixture (**4**) containing two C-2 epimers differing in stereochemistry at C-2. Compound **4** was hydrogenated to generate the corresponding aldehyde (**5**) as a mixture of isomers at C-2. In an attempt to isolate the (*2R*)-stereoisomer using silica gel column chromatography, compound **5** was peracetylated to give compound **6**. However, this problem was complicated by the fact that the peracetylation gave the pyranose form of the sugar as both α - and β -isomers for each of the two C-2 epimers. These four peracetylated stereoisomers could not be separated by silica gel column chromatography to give pure C-2 stereoisomers of compound **5** after deacetylation. Instead, heptose **5**, as a mixture of two C-2 epimers, was treated with glyphosate and was subsequently reduced by NaBH₃CN to yield analogue **O1** as a mixture of stereoisomers at C-2 (ratios vary from 1:1 to 1:9). Compound **O1** was then isolated from the reaction mixture by anion exchange chromatography.

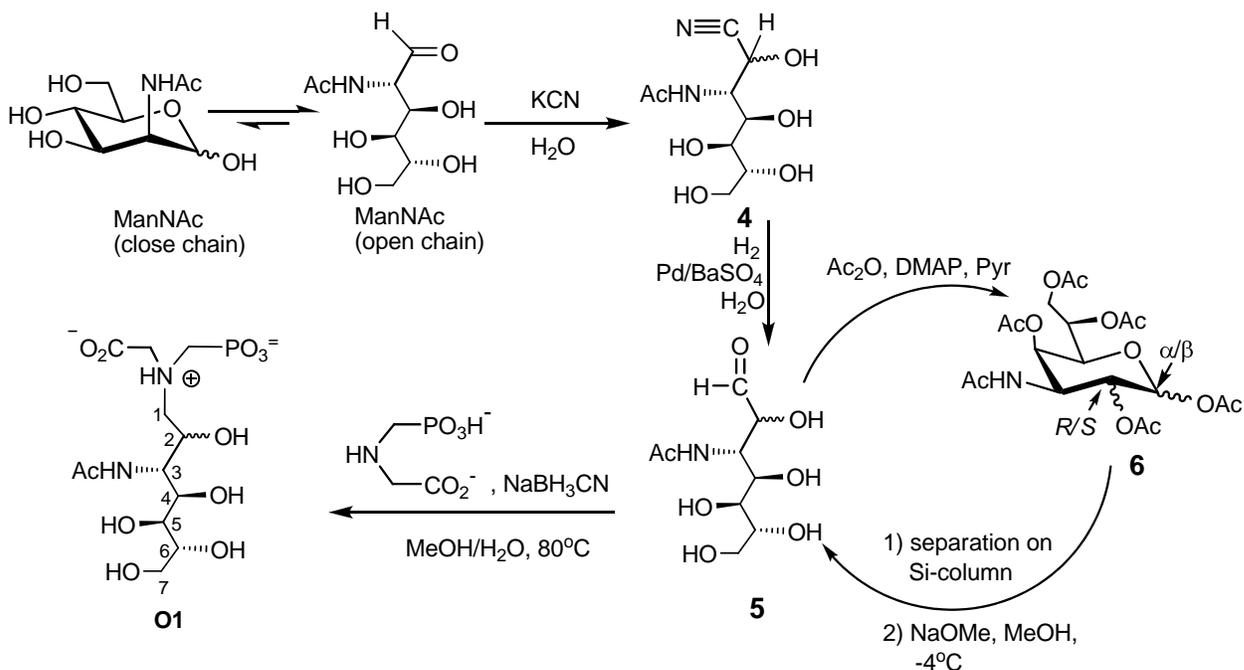


Figure 2.4 Synthesis of oxocarbenium analogue **O1**.

2.3.2 Test for Inhibition

Compound **O1**, prepared as a mixture of stereoisomers at C-2, was tested as an inhibitor of the NeuB reaction. A continuous coupled assay monitoring the formation of Neu5Ac was used in the kinetic study (Figure 2.5). As NeuB catalyzes the reaction between ManNAc and PEP to produce Neu5Ac, a coupling enzyme, sialic acid lyase (SAL), cleaves Neu5Ac into ManNAc and pyruvate. Pyruvate is then reduced to lactate by lactate dehydrogenase (LDH), which at the same time consumes nicotinamide adenine dinucleotide hydride (NADH) and converts it into nicotinamide adenine dinucleotide (NAD⁺). The rate of reaction catalyzed by NeuB can therefore be followed by the disappearance of the NADH absorption at 340 nm ($\epsilon = 6\,220\text{ M}^{-1}\text{ cm}^{-1}$).

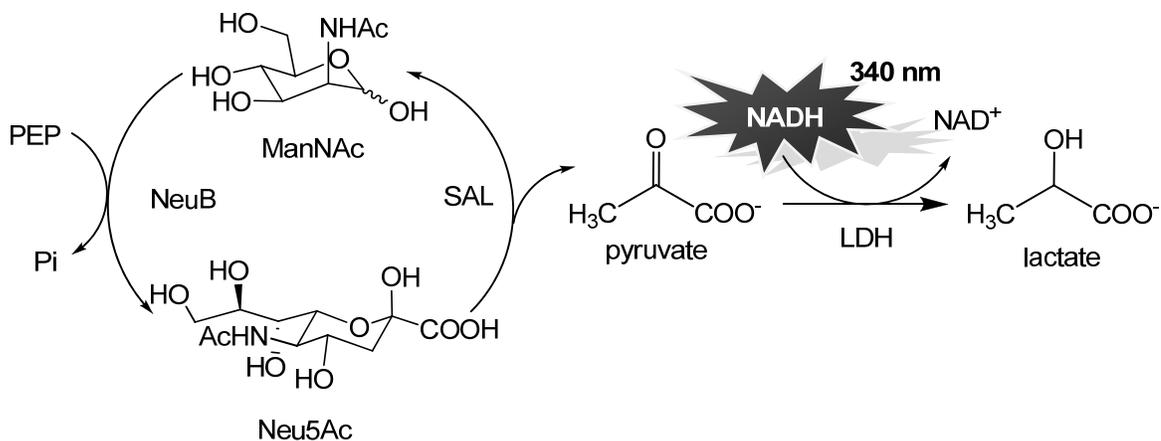


Figure 2.5 Continuous coupled assay for measuring sialic acid synthase activity.

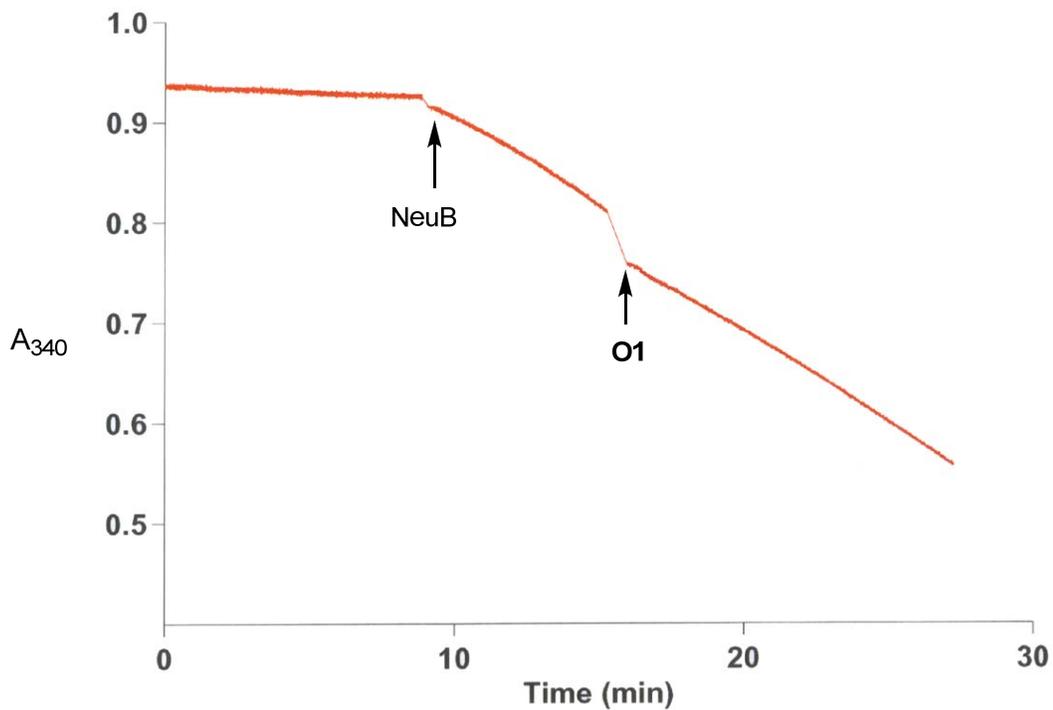


Figure 2.6 Test of inhibition by analogue **O1** (100 μ M). NeuB was added after approximately 9 min. Analogue **O1** was added after approximately 15 min. The assay was conducted at 37 $^{\circ}$ C.

In a typical kinetic trial, NeuB was added to a cuvette containing a pre-equilibrated solution of ManNAc, PEP, Mn^{2+} , SAL and LDH, and a linear decrease in absorbance at 340 nm, which reflects the rate of the reaction, was observed (Figure 2.6). Analogue **O1** was then added, and the inhibitory effect was evaluated by comparing the rate before and after its addition. However, no inhibition was observed, even when the final concentration of **O1** reached 1 mM.

The lack of inhibition by **O1** was somewhat surprising when one considers that similar analogues were shown to be potent inhibitors of both KDO 8-P synthase and DAH 7-P synthase. However, several reasons can be envisioned to account for the inability of analogue **O1** to bind to NeuB. First of all, the protonated tertiary amine is sp^3 hybridized and may not serve as a good mimic of the sp^2 hybridized oxocarbenium ion. Moreover, when compared to the oxocarbenium ion intermediate, analogue **O1** contains an extra methylene group separating the carboxylate from the positively charged nitrogen. While this does not appear to be a problem in the case of KDO 8-P synthase, the additional steric bulk may be much more significant within the active site of NeuB and greatly reduce the binding affinity. Evidence in support of this notion comes from a comparison of the crystal structures of the *N. meningitidis* NeuB and the *Aquifex aeolicus* KDO 8-P synthase. A close look at the crystal structure of NeuB•*N*-acetylmannosaminitol•PEP• Mn^{2+} (Figure 2.7, A) shows that the carboxylate group of PEP is positioned in a crowded cavity enclosed by Glu25, Lys53, Gln55, Lys129, Thr110, Phe112 and *N*-acetylmannosaminitol (rManNAc). In particular, the carboxylate group of PEP is hydrogen-bonded to the nearby Thr110 which is only 2.6 Å away. Upon replacing rManNAc/PEP with the oxocarbenium analogue **O1**, the corresponding carboxylate would be positioned much closer to Thr110 because of the extra methylene group. This may affect the formation of a hydrogen bond and even generate a steric clash that disfavors the binding of analogue **O1** to the active site. Moreover, the

change in position of the carboxylate brought about by the extra methylene group would likely lead to the loss of other stabilizing interactions (such as the electrostatic interactions with Lys53 and Lys129), and the accumulation of destabilizing interactions (such as electrostatic or steric repulsion with Glu25 or Gln55).

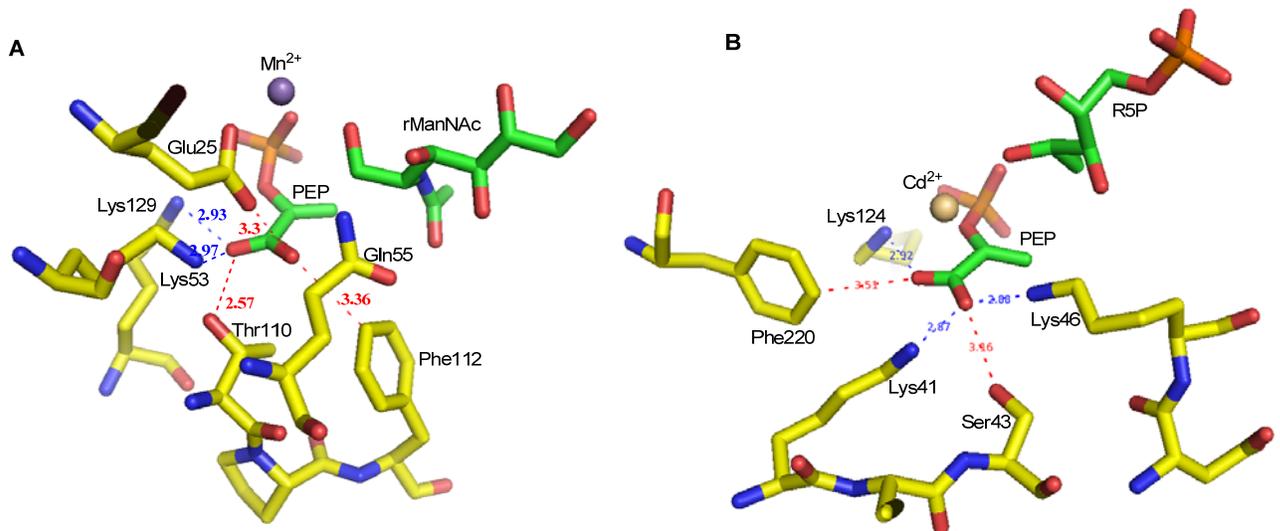


Figure 2.7 A structural analysis of PEP binding in the active sites of NeuB and KDO 8-P synthase.

A) Active site of *N. meningitidis* NeuB complexed with PEP and *N*-acetylmannosaminitol (rManNAc), B) Active site of *A. aeolicus* KDO 8-P synthase complexed with PEP and ribose 5-phosphate (R5P). *N*-acetylmannosaminitol, PEP are shown in stick representation with carbon atoms in green, oxygen atoms in red, nitrogen atoms in blue, phosphorus atoms in orange, Mn²⁺ in magenta and Cd²⁺ in sand color. Active residues are also shown in stick representation with carbon atoms in yellow. Favorable carboxylate-Lys interactions are shown in blue. Interactions with other nearby residues are shown in red.

In contrast, in the active site of KDO 8-P synthase•R5P•PEP•Mn²⁺ complex (R5P stands for ribose 5-phosphate, structurally close to the natural substrate A5P), the carboxylate of PEP is positioned in a relatively more spacious cavity (Figure 2.7, B). Compared to that in the NeuB•rManNAc•PEP•Mn²⁺ complex, the carboxylate group of PEP in the KDO 8-P

synthase•R5P•PEP•Mn²⁺ complex is located further away from nearby residues (Ser43, Phe220) that may potentially cause electrostatic repulsion or steric clashes with the carboxylate. Additionally, there are three basic residues (Lys41, Lys46, Lys124), as compared to the two lysine residues in the case of NeuB, that could form stabilizing electrostatic interactions with the carboxylate group. If R5P and PEP are replaced by a binary oxocarbenium analogue such as compound **1** (Figure 1.16), the carboxylate group can rotate away from Ser43 and be stabilized by the additional basic residue Lys124, that has no counterpart in the NeuB active site. This is confirmed by the crystallographic analysis of the *A. aeolicus* KDO 8-P synthase•**1**•Mn²⁺ complex, in which the carboxylate group is positioned closer to Lys124 and away from Ser43 with distances of 2.65 Å and 4.55 Å, respectively (not shown).⁵¹

2.4 Syntheses of, and Inhibition by, Tetrahedral Intermediate Analogues

Soon after the oxocarbenium ion intermediate analogue **O1** was prepared and found not to be an inhibitor, NeuB was crystallized as a complex with *N*-acetylmannosaminitol (rManNAc), PEP and Mn²⁺. Results from the crystallographic analysis greatly increased our understanding of the NeuB mechanism. However, in order to reveal more information regarding the changes that occur at the active site during catalysis, it was still desirable to investigate analogues that reflect the structural properties of intermediates. Analogues **T1** and **T2** closely mimic the structure of the tetrahedral intermediate and will help to elucidate the stereochemistry at C-2 as well as the potential role of the metal cofactor as a source of nucleophilic water. It is expected that analogue **T1** will bind in the active site of NeuB, because it is sterically smaller than the tetrahedral intermediate yet retains all of the functional groups except the C-2 hydroxyl. Analogue **T2** has a methylene group in the place of the bridging phosphate oxygen which may impart slight changes in both the steric and electrostatic interactions. Nevertheless, analogue **T2** is best suited to elucidate the interaction of the C-2 hydroxyl with surrounding active site residues and the metal cofactor since this group is retained.

Analogues **T1** and **T2** should be readily prepared from a suitably protected form of Neu5Ac. To do this, one must protect all the hydroxyls and the carboxylate while exposing the C-2 carbonyl group for further reaction (Figure 2.8). The initially adopted strategy was to protect the C-2 carbonyl group of Neu5Ac as an oxime, which may then be deprotected by ozonolysis. This method was previously used to prepare hydroxyl-protected ManNAc with a free aldehyde group.¹²⁸ However, the Neu5Ac 2-methyloxime and Neu5Ac 2-benzyloxime could not be

deprotected by any conventional methods such as ozonolysis, KMnO_4 oxidation, cross oximation¹²⁹ or zinc-induced deoximation.¹³⁰ This was probably due to the increased steric hindrance at the ketone carbonyl (as compared to the aldehyde in previous studies) and its conjugation with the adjacent carboxylate group. We then decided to switch to a strategy involving the ozonolysis of a 2-methylidene derivative of Neu5Ac, which can be synthesized by an indium-mediated coupling of ManNAc and bromomethylacrylate.

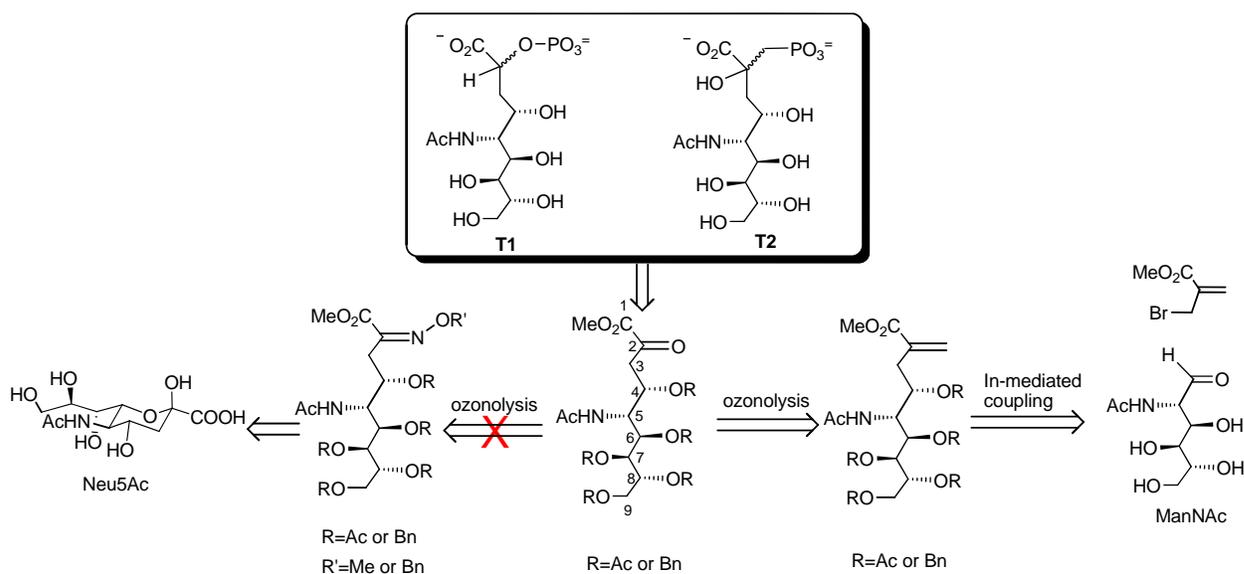


Figure 2.8 Biosynthetic strategy for synthesizing analogues T1 and T2.

Indium-mediated couplings between an aldehyde and an allylic halide were first reported by Li and Chan.¹³¹ Indium-mediated coupling reactions can be conducted in aqueous solutions, often avoiding tedious steps of protection-deprotection and the use of flammable and anhydrous solvents. Compared to previously developed aqueous allylation reactions mediated by zinc¹³², tin¹³³ and bismuth¹³⁴, the advantages of the indium-mediated allylation are that promoters are often not required and the production of byproducts is minimized. The combination of an

indium-mediated allylation with a subsequent ozonolysis, was previously used to synthesize either peracetylated or unprotected versions of the sialic acids KDN^{135, 136} and Neu5Ac¹³⁷. A variety of α -thiol and α -amino aldehydes, including ManNAc derivatives, have been subjected to this reaction.^{138, 139} Results suggested that the reaction between ManNAc and bromomethylacrylate proceeds through a chelated intermediate and the stereo-configuration (*syn*-relationship of C-4 hydroxyl and C-5 acetamido groups) of the product is consistent with the chelation-controlled Cram's rule (Figure 2.9). It was also shown that the major isomer formed by the indium-mediated allylation of protected or unprotected ManNAc using bromomethylacrylate, bears a (4*S*)-configuration as Neu5Ac does, and the extent of stereoselectivity is affected by the composition of solvent.

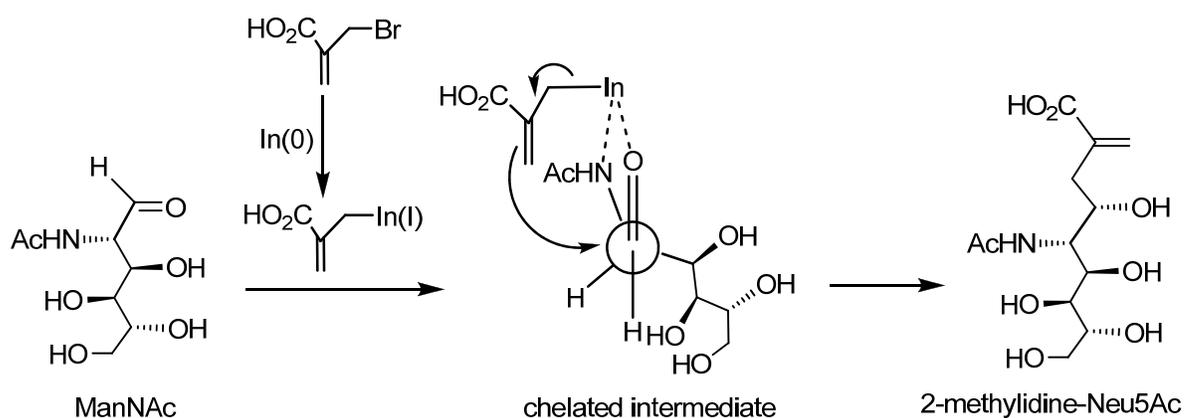


Figure 2.9 Mechanism of indium-mediated allylation.

The following section describes the use of a modified indium-mediated allylation and ozonolysis to build the protected Neu5Ac backbone bearing a free C-2 ketone.

2.4.1 Synthesis of Analogue T1

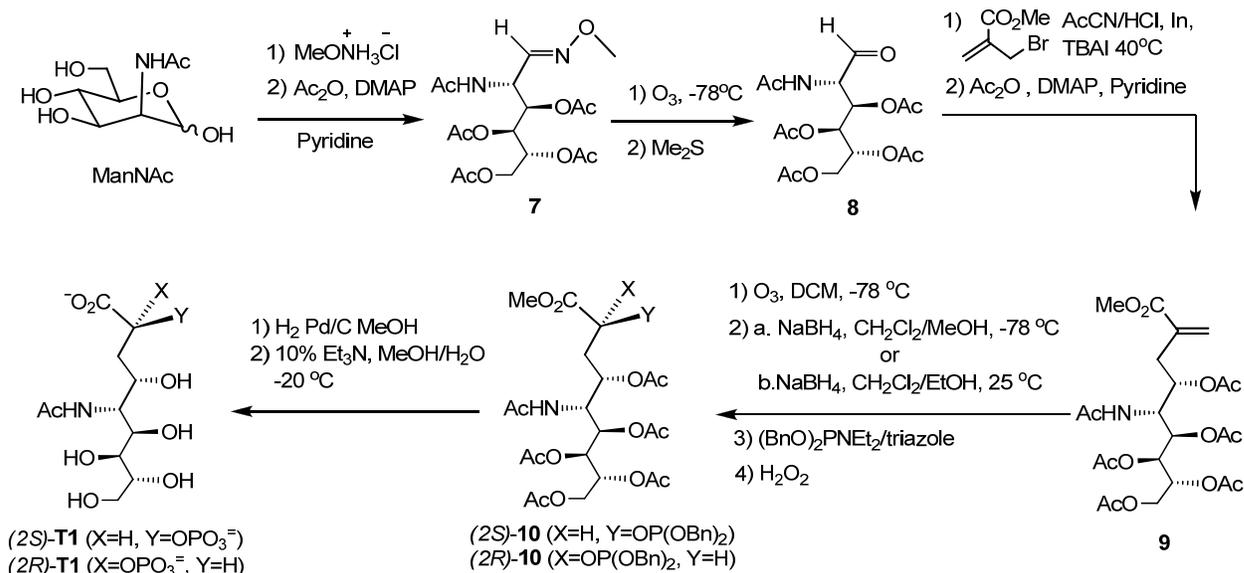


Figure 2.10 Synthesis of analogue **T1**.

The overall strategy used in the synthesis involved the allylation of the peracetylated open-chain form of ManNAc, **8** (Figure 2.10). The C-1 aldehyde group in the open chain form of ManNAc was first masked as an oxime using methoxyamine before the peracetylation of the hydroxyl groups was performed. The peracetylated ManNAc oxime **7** was then subjected to ozonolysis which cleaves the oxime functionality to give aldehyde **8** as previously described.¹²⁸ An indium-mediated allylation of aldehyde **8** using methyl bromomethylacrylate and the conditions described by Vorwerk and Vasella¹⁴⁰, was followed by an acetylation with acetic anhydride to give peracetylated compound **9** with a (4S):(4R) ratio of ~3:1. (4S)-**9** has the appropriate stereochemical configuration required to mimic the sialic acid backbone, and was isolated from its stereoisomer by recrystallization. The C-4 stereochemistry was ultimately confirmed by crystallographic analysis of the NeuB•**T1** complex (*vide infra*). Compound **9** was then subjected to ozonolysis, followed by reduction with NaBH_4 and phosphorylation with

dibenzyl phosphoramidite to generate compound **10** as a mixture of two C-2 epimers that were not separable by conventional silica gel column chromatography. The mixture of (2*S*)-**10** and (2*R*)-**10** could be produced in different ratios by conducting the NaBH₄ reduction at different temperatures. Reduction at -78 °C gave a 4.5:1 (2*S*)-**10**:(2*R*)-**10** whereas treatment with NaBH₄ in DCM/EtOH at 25 °C gave a 1.2:1 mixture of (2*S*)-**10**:(2*R*)-**10**. The mixtures were then subjected to debenzylation by hydrogenolysis and a mild deacetylation using triethylamine in MeOH/H₂O at -20 °C. The reaction mixtures were purified by size exclusion chromatography to yield analogue **T1** with a (2*S*)-**T1**: (2*R*)-**T1** ratio of either 4.5:1 or 1.2:1 depending on the reduction conditions used in the preparation of compound **10**. The ratio of the **T1** diastereomers was determined by ³¹P NMR spectroscopy. The C-2 configuration of the two stereoisomers was assigned by first comparing the inhibition due to each of the 4.5:1 and the 1.2:1 (2*S*)-**T1**:(2*R*)-**T1** mixtures in order to determine which stereoisomer is a better binder, and then determining the stereochemistry of the tighter binding stereoisomer by crystallographic analysis. More detailed information will be given in Section 2.4.3 and Section 2.4.5.

2.4.2 Attempted Synthesis of Analogue **T2**

The strategy employed to synthesize analogue **T2** involves the nucleophilic attack of the lithiated carbanion derived from methylphosphonate and a suitable α -ketoacid derivative (Figure 2.11). A similar approach was previously described by Grison *et al.* in the preparation of tetrahedral intermediate analogues for KDO 8-P synthase and DAH 7-P synthase.⁵⁵ Compound **9** prepared in the synthesis of analogue **T1** was used as the precursor of the α -ketoacid derivatives.

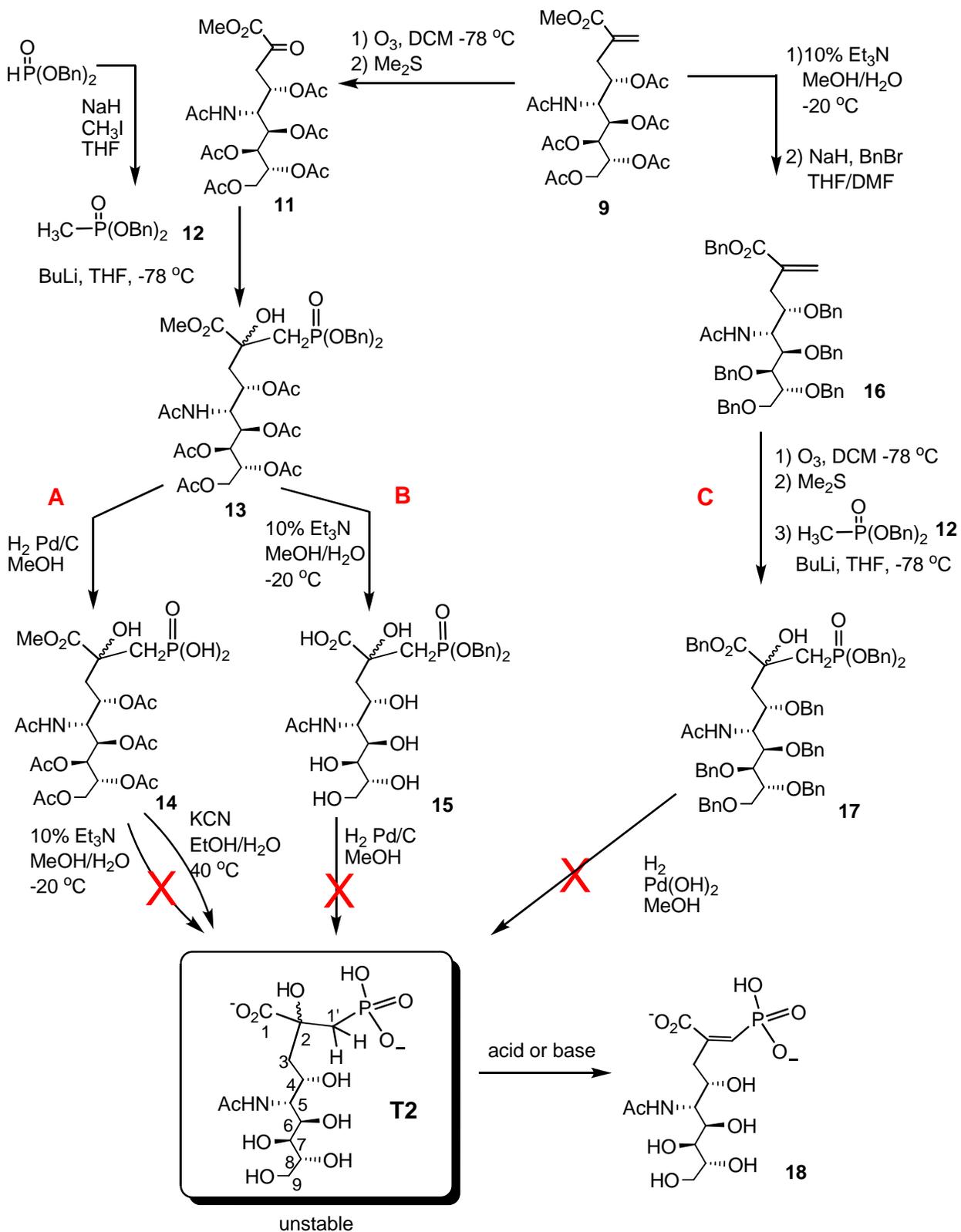


Figure 2.11 Attempted synthesis of analogue T2.

In Path A, compound **9** was first subjected to ozonolysis and a subsequent reduction by dimethyl sulfide to give the α -ketoacid methyl ester **11**, which was then reacted with the anion derived from dibenzyl methylphosphonate **12**¹⁴¹ to yield compound **13** as a mixture of two C-2 stereoisomers. As a direct precursor, compound **13** is expected to be converted to analogue **T2** by removal of the methyl, acetyl and benzyl protecting groups (Figure 2.11). Unfortunately, the attempts to deprotect compound **13** were unsuccessful. Hydrogenolysis gave compound **14**, however the mild ester hydrolysis with triethylamine in MeOH/H₂O at -20 °C used in the synthesis of analogue **T1** did not produce the desired product **T2** ($m/z = 404$, (M-H⁺)). Instead, the reaction yielded a product with m/z of 386 (M-H⁺), implying that a molecule of water had been eliminated. Since a lactone or an intramolecular phosphonate ester was unlikely to form under these basic hydrolytic conditions, it was proposed that compound **18** was generated as the result of base-promoted elimination of the C-2 hydroxyl and the C-1' proton. Other attempts, such as dealkylation of the ester via hydrolysis with lipase, were shown not to be able to fully deprotect compound **14**. Cyanide-mediated trans-esterification was also used as an alternative method to deacetylate compound **14**, and an analysis of the crude reaction mixture by ESI-MS ($m/z = 404$, (M-H⁺)) suggested that analogue **T2** had formed. Nevertheless, an attempt to purify **T2** led to the formation of compound **18** even under neutral aqueous conditions.

Other strategies were also employed, including an initial removal of the *O*-acetyl and *O*-methyl groups to give compound **15** and a subsequent deprotection of the benzyl groups of compound **15** by hydrogenolysis (Figure 2.11 Path B). Alternatively, compound **9** was converted into a perbenzylated precursor **16** which was carried on to give compound **17** and then subjected to hydrogenolysis, to ensure a non-basic environment in the final step (Figure 2.11 Path C). Neither of these attempts led to the formation of analogue **T2**. It was shown that the C-2

hydroxyl group was still intact when only one of the two phosphonate benzyl esters was removed by hydrogenolysis, but the elimination occurred to give compound **18** upon full debenzylation of the phosphonate group. Since the pH would decrease as a result of the formation of the phosphonic acid, we tried to control the pH using Tris-HCl buffer or triethylammonium bicarbonate. However, these attempts led to incomplete debenzylation.

A close look at the structure of **T2** shows that it has relatively acidic C-1' protons adjacent to the phosphonate, as well as a tertiary hydroxyl group on C-2. These structural characteristics render analogue **T2** liable to both base- and acid-catalyzed eliminations (Figure 2.12). In basic conditions, the elimination may happen in either a stepwise or concerted manner starting with the deprotonation of the C-1' proton (Figure 2.12, mechanisms shown with blue arrows).

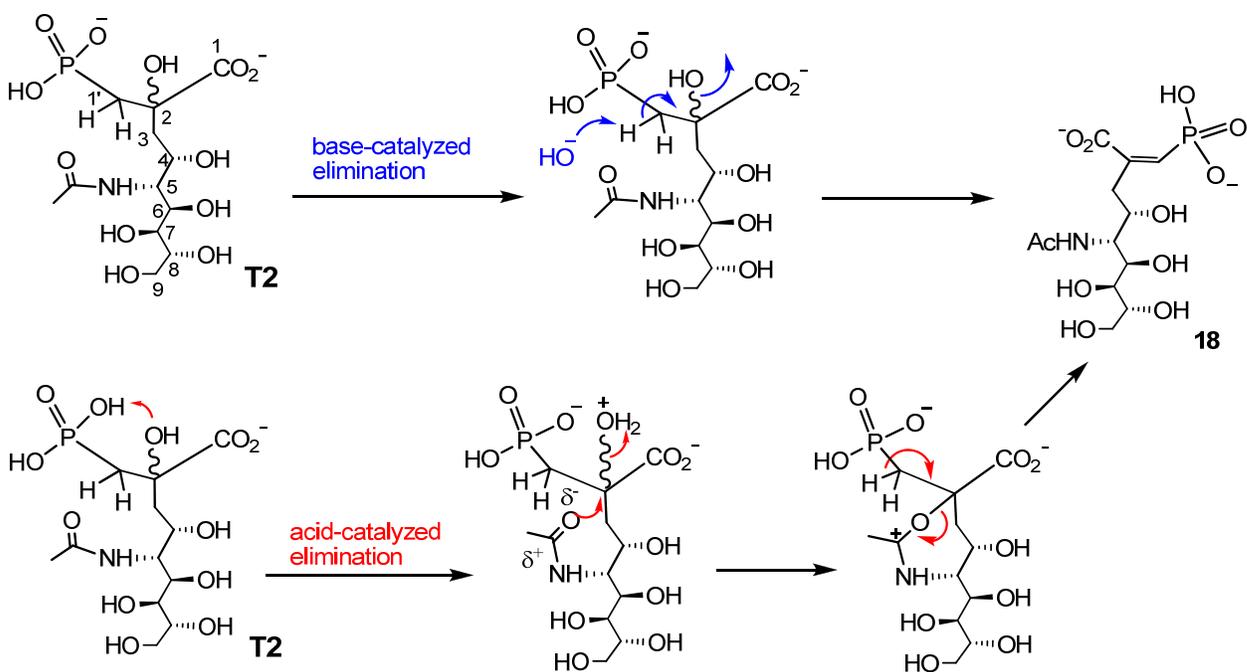


Figure 2.12 Potential mechanisms for the elimination of water from **T2** under basic or acidic conditions.

In acidic conditions, the elimination is likely triggered by the protonation of the C-2 hydroxyl group by the adjacent phosphonic acid (Figure 2.12, mechanisms shown with red arrows). In retrospect, a similar tetrahedral intermediate analogue **3** (Figure 1.16) and its protected precursor synthesized in the studies of KDO 8-P synthase and DAH 7-P synthase were shown to be relatively stable at controlled pH.⁵⁵ The instability of analogue **T2**, as well as the fact that its precursor, compound **13**, was found to be unstable to purification using silica gel column chromatography (as a result, it was purified by ion exchange chromatography), implied that the *N*-acetyl group at the C-5 position of analogue **T2** (which is a major structural difference between analogue **3** and **T1**) may play a role in facilitating the elimination under acidic or near neutral conditions. The *N*-acetyl carbonyl group could act as a nucleophile to promote the elimination of water at C-1' (Figure 2.12, mechanisms shown with red arrows). However, this would invoke a seven-membered ring transition state which may be energetically unfavorable. In summary, it was determined that analogue **T2** was not stable enough to be isolated and used for kinetic and crystallographic studies. Therefore, only analogue **T1** was tested for inhibition and crystallized in a complex with NeuB.

2.4.3 Inhibition Assay and Kinetic Evaluation of Analogue **T1**

The inhibitory effect of analogue **T1** towards the activity of *N. meningitidis* NeuB was preliminarily tested using ESI-MS (-). By comparing the relative intensity of formed Neu5Ac ($m/z = 308$, (M-H⁺)) and unreacted ManNAc ($m/z = 244$, (M+Na⁺)), it was shown that analogue **T1** was an effective inhibitor at sub-millimolar levels (Figure 2.13).

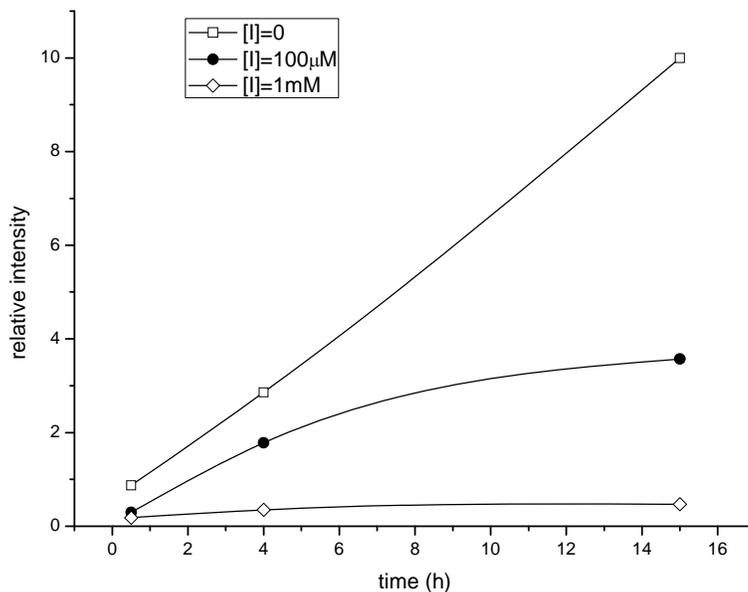


Figure 2.13 Inhibition of NeuB by analogue T1 as analyzed by ESI-MS. Y axis shows the relative signal intensities of formed Neu5Ac ($m/z = 308$, $(M-H^+)$) and unreacted ManNAc ($m/z = 244$, $(M+Na^+)$). Data were collected without **T1** or with 100 μ M or 1 mM of **T1** after 0.5, 4 and 15 h of incubation. The concentrations of PEP, ManNAc and Mn^{2+} , were 100 μ M, 10 mM and 1 mM, respectively.

The inhibition kinetics of analogue **T1** were then studied using a continuous coupled assay monitoring the release of free phosphate (Figure 2.14).¹⁴² Purine nucleoside phosphorylase (PNP) is an enzyme that cleaves purine nucleosides to produce a purine and ribose phosphate. 2-Amino-6-mercapto-7-methylpurine ribonucleoside (methylthioguanosine or MESHG), a guanosine analogue, shows a maximum UV absorbance at 330 nm at a neutral pH (Figure 2.14, curve **A**). In the presence of free phosphate, PNP cleaves MESHG and releases 2-amino-6-mercapto-7-methylpurine **19**, which has a max UV absorbance at 355 nm (Figure 2.14, curve **B**). At 360 nm and neutral pH, the difference between the UV absorbance of MESHG and that of compound **19** is maximized (Figure 2.14, difference spectrum shown as curve **C**). Therefore, the differential

extinction coefficient at 360 nm ($\Delta\epsilon_{360}$) can be used to determine the amount of phosphate formed by the NeuB-catalyzed reaction. This assay was previously used in the kinetic study of pseudaminic acid synthase by Chou *et al.*⁸⁰

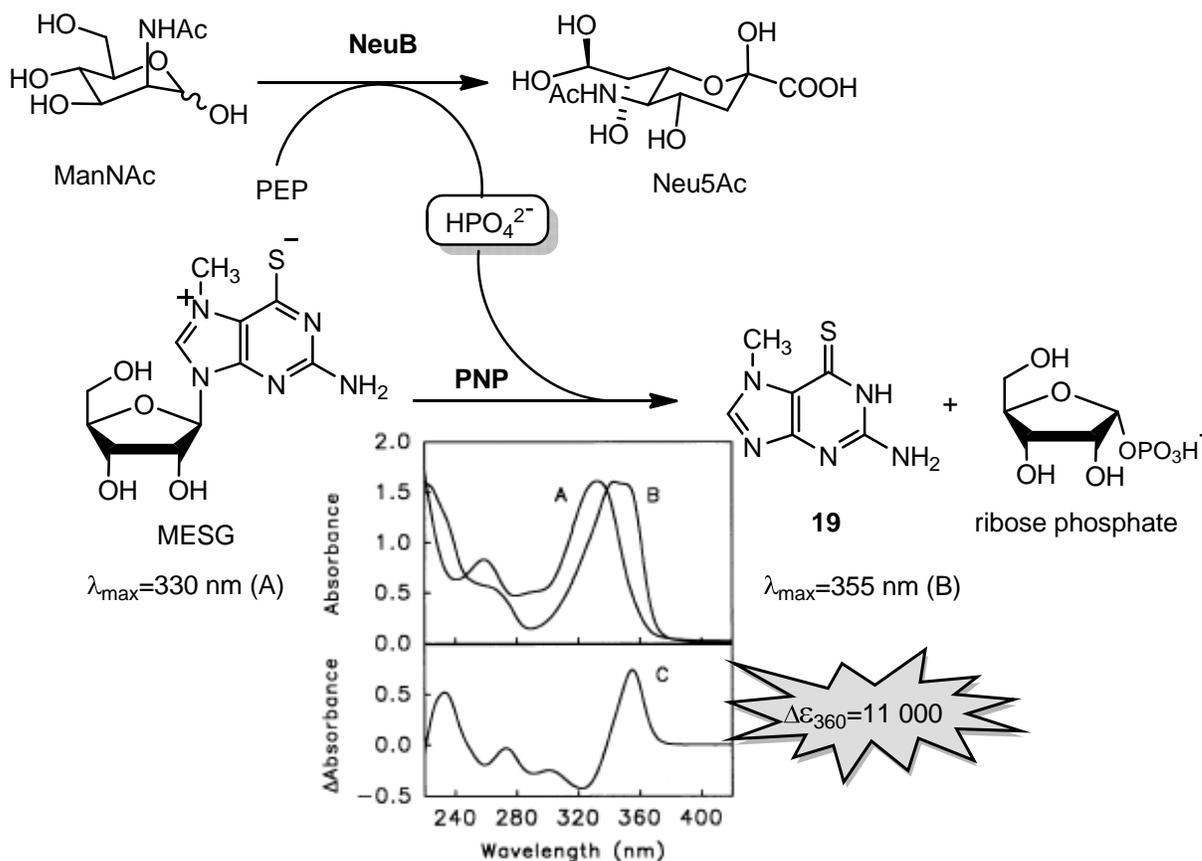


Figure 2.14 Assay for phosphate release by NeuB. Scheme showing the continuous coupled assay for phosphate release. The upper inset shows the wavelength-dependent absorbance curve of A) MESG, and B) compound 19. The lower inset shows the difference trace C at pH 7.6.¹⁴²

Before conducting the inhibition studies, the kinetic constants of the reaction catalyzed by the purified His₆-NeuB were obtained using the phosphate assay described above (Figure 2.15). The K_M value for PEP was determined to be $157 \pm 10 \mu\text{M}$ and the k_{cat} value was determined to be $3.4 \pm 0.2 \text{ s}^{-1}$. These results are consistent with previous kinetic studies on NeuB activity.⁴²

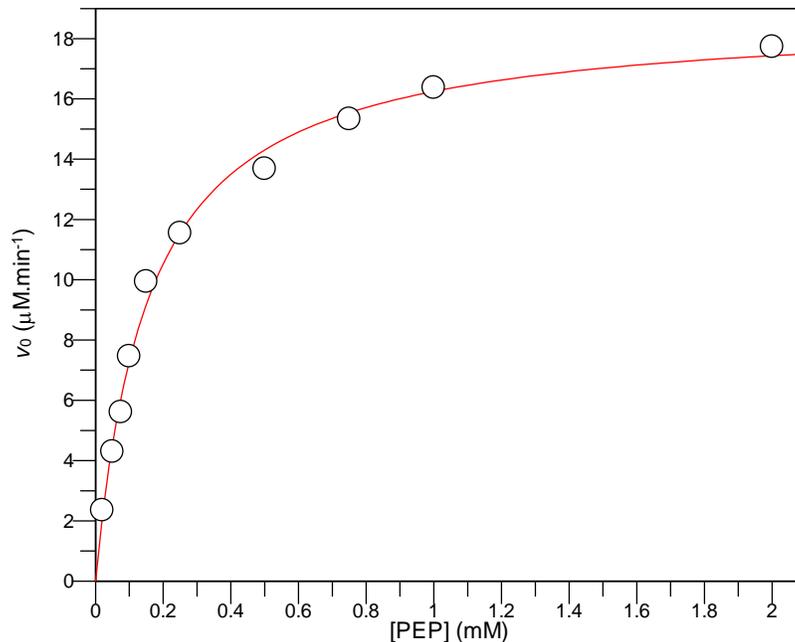


Figure 2.15 A plot of initial velocity vs. PEP concentration for the reaction catalyzed by NeuB. The kinetic parameters determined by fitting the data to the Michealis-Menten equation are as follows: $k_{\text{cat}} = 3.4 \pm 0.6 \text{ s}^{-1}$, $K_M = 157 \pm 10 \text{ }\mu\text{M}$. Kinetics were determined in Tris-HCl buffer (100 mM, pH 7.6), with $[\text{ManNAc}] = 10 \text{ mM}$ and $[\text{Mn}^{2+}] = 1 \text{ mM}$, 37 °C.

In an initial inhibition kinetic study using the phosphate assay, the reaction was initiated by the addition of NeuB to a pre-equilibrated mixture containing ManNAc, PEP and Mn^{2+} . After 3 min a sample of 4.5:1 mixture of (2*S*)-**T1**:(2*R*)-**T1** was added (Figure 2.16). The resulting decrease in rate indicated that **T1** was acting as an inhibitor. However, the extent of inhibition increased as a function of time, indicating that **T1** acts as a slow-binding inhibitor. This observation was not very surprising since the oxocarbenium ion intermediate analogue **1** used in the studies of KDO 8-P synthase was also shown to be a slow-binding inhibitor.⁴⁸ To avoid this problem, NeuB was pre-incubated with analogue **T1** and PEP for different periods of time and the reaction was initiated by the addition of ManNAc. It was determined that the time required for equilibration to occur, as evidenced by linear reaction kinetics, was about 15 min.

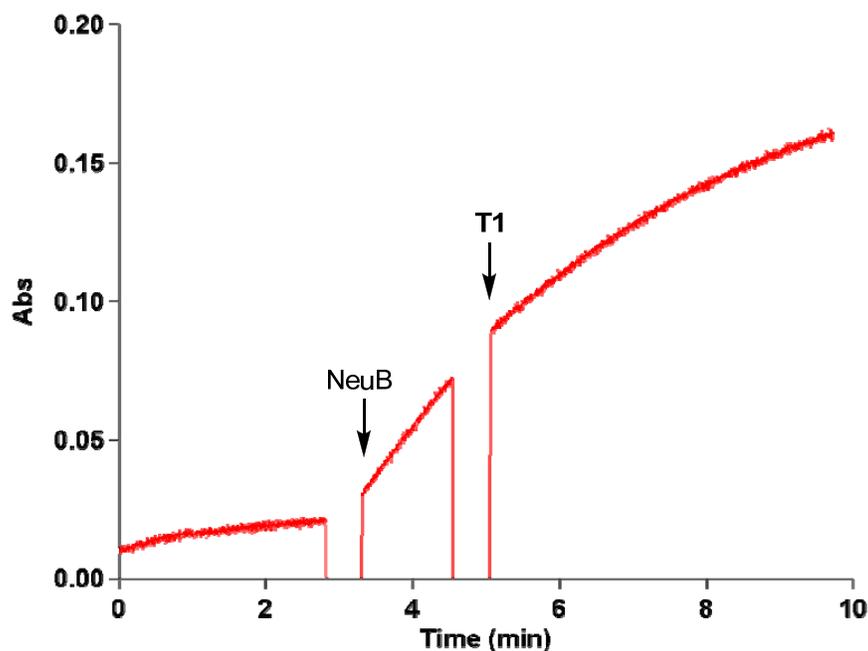


Figure 2.16 Initial inhibition kinetics test using the continuous coupled phosphate assay. The reaction was initiated by the addition of NeuB and after 2 min, **T1** (10 μM , total concentration) was added.

Accordingly, in the following kinetic analyses, the enzyme was preincubated with inhibitor **T1** (variable amounts) in a buffered solution ($\text{pH} = 7$) containing PEP (variable amounts) and Mn^{2+} (1 mM) for 20 minutes to ensure that binding equilibration had been established. The reactions were then initiated by the addition of ManNAc (10 mM, $\sim K_{\text{M ManNAc}}$). Analogue **T1** ((*2S*)-**T1**:(*2R*)-**T1** = 4.5:1, containing 18% (*2R*)-**T1**) was found to be a competitive inhibitor against PEP with an apparent K_i value of $3.1 \pm 0.1 \mu\text{M}$ (Figure 2.17). Identical kinetic runs using the 1.2:1 mixture of (*2S*)-**T1**:(*2R*)-**T1** (containing 45% (*2R*)-**T1** instead of 18%) were shown to require ~ 2.5 -fold lower total inhibitor concentration in order to bring about a comparable reduction of rate. Calculations show that the 4.5:1 mixture of (*2S*)-**T1**:(*2R*)-**T1** and the 1.2:1 mixture of (*2S*)-**T1**:(*2R*)-**T1** with ~ 2.5 -fold lower total inhibitor concentration

contained approximately the same amount of (2*R*)-**T1** but the latter one contained ~4-fold more (2*S*)-**T1**. Since the 4-fold increase in the concentration of (2*S*)-**T1** did not change the inhibition effect by any significant amount, it could be concluded that the minor stereoisomer (2*R*)-**T1** was responsible for the majority of the inhibition, and that it binds to NeuB much more tightly than (2*S*)-**T1**. The fact that the K_i for inhibitor **T1** is significantly lower than the K_M for PEP strongly supports the notion that the structure of **T1** mimics that of the tetrahedral intermediate formed in the NeuB reaction. Moreover, as one stereoisomer was found to bind to NeuB much more tightly than the other, it would likely be possible to use crystallographic analysis to address the stereochemistry of the tetrahedral intermediate and the positioning of the active site residues around C-2.

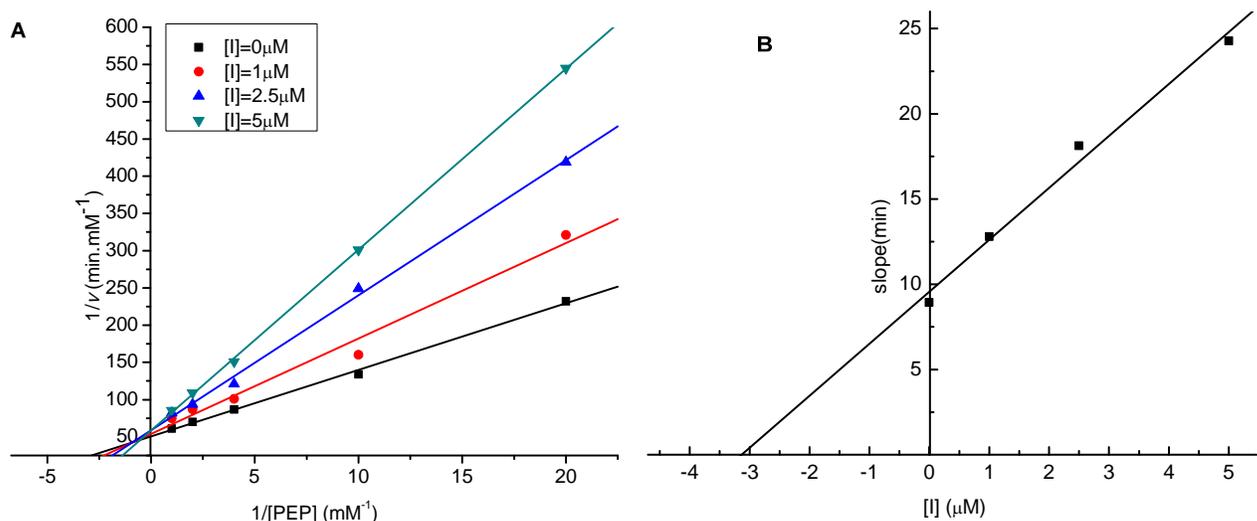


Figure 2.17 Inhibition kinetics of inhibitor **T1 ((2*S*)-**T1**:(2*R*)-**T1** = 4.5:1).** [ManNAc]=10 mM, [Mn²⁺] = 1 mM, [MESG] = 200 μM, PNP = 5 units, [NeuB] = 92 nM. 100 mM Tris-HCl buffer (pH 7.0), 37 °C. [I]= 0, 1, 2.5, 5 μM. Reaction was initiated with ManNAc after a 20-min pre-incubation. **A**) Lineweaver-Burk plot showed **T1** as a competitive inhibitor; **B**) Replot of data from Lineweaver-Burk plots for the determination of K_i . $K_i = 3.1 \pm 0.1 \mu\text{M}$ ((2*S*)-**T1**:(2*R*)-**T1**=4.5:1).

2.4.4 Metal Dependency of Inhibition by Inhibitor T1

The metal dependency in the inhibition of NeuB by inhibitor **T1** was investigated by varying the concentration of Mn^{2+} in the kinetic studies. It was shown that even in the presence of inhibitor **T1**, NeuB exhibited higher activity as the concentration of Mn^{2+} increased from 0.2 to 1 mM (Figure 2.18). The activity then slowly decreased when the concentration of MnCl_2 was further increased from 1 mM to 10 mM. The initial increase of activity reflects the dependency of NeuB on divalent metal ions, because metal cofactors such as Mn^{2+} are absolutely required for the catalysis to occur. The decrease of activity at higher Mn^{2+} concentrations was likely caused by the non-productive binding of the metal ions to an allosteric site, which decrease the affinity of the active site for the substrate. This was previously observed for other PEP-utilizing enzymes.^{143, 144} The similar trends exhibited by rate-*vs*- $[\text{Mn}^{2+}]$ curves, in the absence or in the presence of inhibitor **T1**, suggest that the metal ion does not bind competitively with the inhibitor.

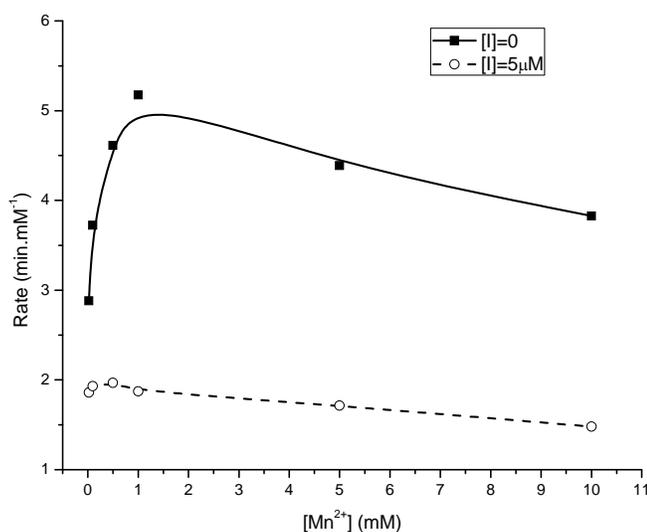


Figure 2.18 Metal dependency of the inhibition by T1. Correlation of NeuB activity and the metal concentration with or without inhibitor **T1**. [ManNAc]=10 mM, [PEP]=10 μM, [MESG] = 200 μM, PNP = 5 units, [NeuB] = 92 nM. 100 mM Tris-HCl buffer (pH 7.0), 37 °C. $[\text{Mn}^{2+}]$ = 0.2, 0.5, 1, 2, 5, 10 mM, [I]= 0 or 5 μM.

2.4.5 Crystallographic Analysis of the Complex Between Inhibitor T1 and NeuB.^b

Untagged NeuB was first crystallized using the hanging drop vapor diffusion method in the presence of the precipitant malic acid (1.50~1.55 M, pH 6.2). Crystals of malic acid/Mn²⁺ complex were obtained as described previously.⁴² The bound malate was then exchanged for the analogue **T1** by soaking the crystals in 2 M phosphate buffer (pH 6.2) containing 10 mM MnCl₂ and 3 mM of the 4.5:1 mixture of (2*S*)-**T1**:(2*R*)-**T1**. The crystal structure of the resultant NeuB•inhibitor **T1**• Mn²⁺ complex was obtained with high resolution (1.75 Å).

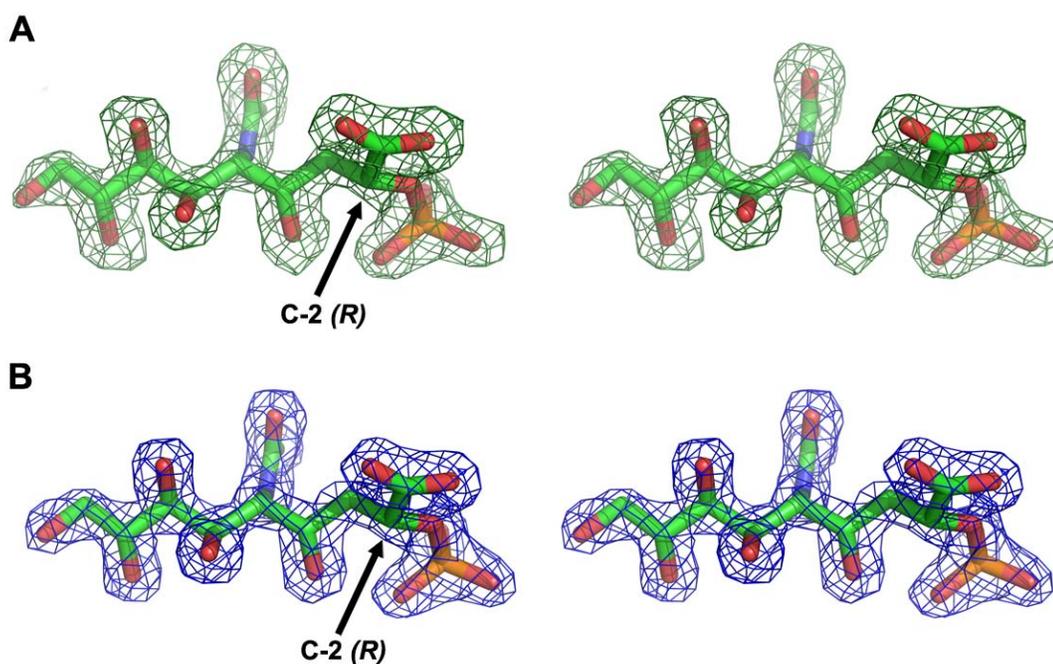


Figure 2.19 Electron density of inhibitor (2*R*)-**T1** in the NeuB• **T1**•Mn²⁺ complex. Stereoview of the observed electron density of the NeuB bound inhibitor **T1** in **A**) the initial F_o-F_c map contoured at 3.5σ and **B**) the refined $2F_o-F_c$ map contoured at 1.5σ . The (*R*) configuration of inhibitor **T1** is highlighted with an arrow pointing to the C-2 position. Carbon atoms are shown in green, oxygen atoms in red, nitrogen atoms in blue and phosphorus atoms in orange.

^b All of the crystallographic experiments were performed by Ho-Jun Lee in the research group of Dr. Natalie Strynadka in the Department of Biochemistry at UBC.

The electron density for the bound inhibitor **T1** was clearly visible and was found to be fully occupied within the active site (Figure 2.19). The fact that only a single stereoisomer of analogue **T1** was found to be bound in the active site implies that NeuB had selected the less prevalent but tighter-binding isomer from the incubation mixture. The stereochemistry of the bound stereoisomer was clearly identified to be *2R*. Along with the results of the kinetic study, this structural characterization provides the basis for the assignment of stereochemistry at C-2, and indicates that the major stereoisomer of **T1** bears a (*2S*)-configuration whereas the minor stereoisomer bears a (*2R*)-configuration.

The overall structure of the NeuB•**T1**•Mn²⁺ complex was found to be very similar to that of NeuB•*N*-acetylmannosaminitol•PEP•Mn²⁺. The two structures can be superpositioned with a root-mean-square deviation (rmsd) of 0.16 Å over all 345 C α atoms. The phosphate and carboxylate groups of **T1** are found to be anchored at similar positions in the active site by electrostatic or hydrogen bond interactions analogous to those previously observed in the NeuB•*N*-acetylmannosaminitol•PEP•Mn²⁺ complex (Figure 2.20A). The phosphate group of **T1** interacts with the conserved residues Lys129, Ser132, Ser154, Ser213 and Asn184 via hydrogen bonds, whereas the carboxylate group forms electrostatic interactions with Lys53 and Lys129 as well as a hydrogen bond with Thr110 (Figure 2.20B). A Mn²⁺ ion with partial (50%) occupancy as indicated by a reduced electron density, and two water molecules, W_{ax} (axial metal-bound water) and W_{eq} (equatorial metal-bound water), were also found in similar positions to those described previously.⁴² The hydroxyl groups of inhibitor **T1** are found to engage in hydrogen bonding interactions with Gln55, Asn74, Tyr86, Asp247 and water molecules. The *N*-acetyl group of inhibitor **T1** also interacts with Glu134 and Arg314', a neighboring residue from the other monomer in the homodimeric complex, as well as the C-7 hydroxyl group of inhibitor **T1**,

via water-mediated hydrogen bonds as also seen in the NeuB•*N*-acetylmannosaminitol•PEP•Mn²⁺ complex.

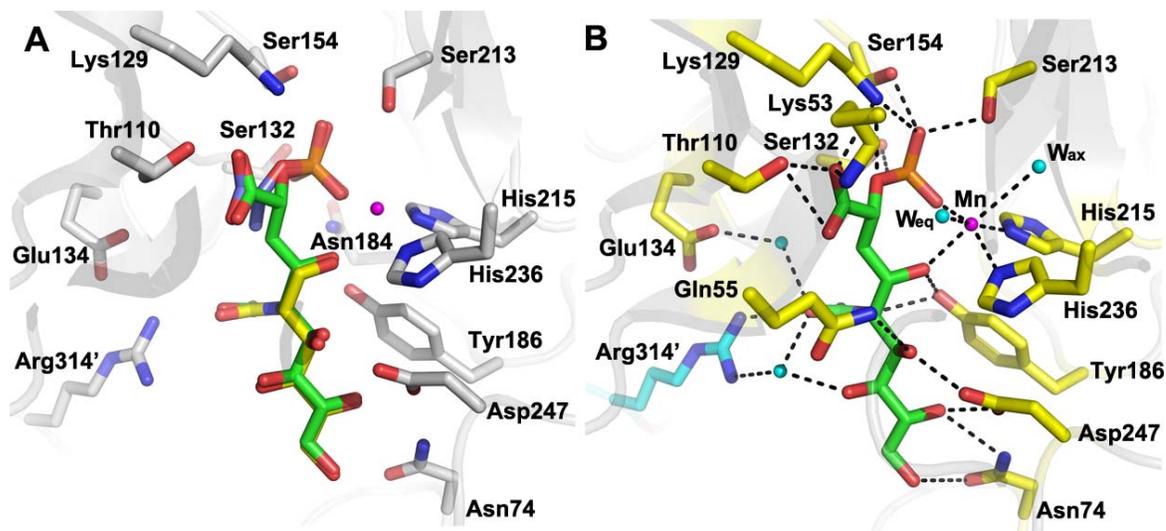


Figure 2.20 The active site of NeuB complexed with inhibitor (2*R*)-**T1**. (A) Comparison of active sites between the NeuB•**T1**•Mn²⁺ and the NeuB•*N*-acetylmannosaminitol•PEP•Mn²⁺ structures. Interacting amino acid residues are shown in CPK coloring with carbons in grey. Carbon atoms in the inhibitor **T1**, *N*-acetylmannosaminitol, and PEP are displayed with green, yellow and light blue, respectively. Lys53 and Gln55 have been omitted for clarity. (B) Active site of the NeuB• **T1**• Mn²⁺ complex. All interacting amino acid residues are shown in CPK with carbons in yellow from one monomer and in cyan from the adjacent monomer with black dotted lines representing hydrogen bonds. Carbon atoms in the inhibitor **T1** are displayed with green and non-carbon atoms are colored according to atom type (N blue, O red, P orange). Manganese ion and water molecules are represented as magenta and cyan spheres, respectively. Asn 184 has been omitted for clarity.

As described previously in this section, the structural analysis shows that the tighter-bound stereoisomer of the tetrahedral intermediate analogue mixture **T1** bears a (2*R*)-configuration. Assuming NeuB selectively binds the C-2 stereoisomer that more closely resembles that of the tetrahedral intermediate, one can infer that the tetrahedral intermediate also

has a (2*R*)-configuration. This suggests that the water-derived hydroxide is delivered to the *si* face of the oxocarbenium ion. The fact that the C-2 hydrogen in (2*R*)-**T1** is oriented towards the active site Mn²⁺ ion indicates that the C-2 hydroxyl group in the (2*R*)-tetrahedral intermediate normally serves as a ligand for the metal. This suggests that in addition to acting as an electrophilic catalyst that polarizes the ManNAc C-1 carbonyl group in the initial nucleophilic attack, the metal also plays a role in activating the water molecule that attacks the oxocarbenium ion intermediate (Figure 2.21).

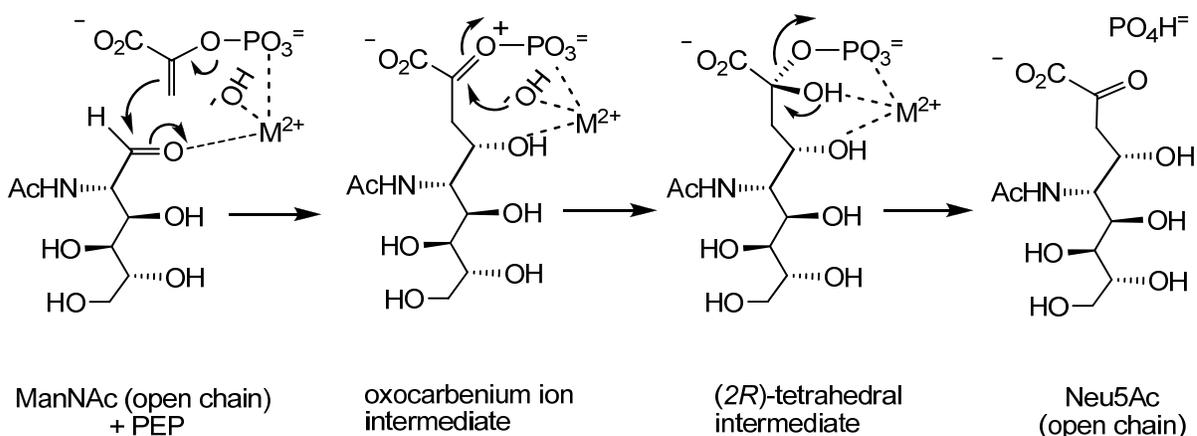


Figure 2.21 Revised mechanism of the reaction catalyzed by NeuB reflecting the proposed stereochemistry of the tetrahedral intermediate and the dual role played by Mn²⁺.

An inspection of the crystal structure of the NeuB•*N*-acetylmannosaminitol•PEP•Mn²⁺ complex reveals that among the two metal-bound water molecules, W_{eq} is 3.1 Å away from C-2 of PEP and is the most likely candidate to attack the oxocarbenium ion intermediate as a nucleophile (Figure 2.22 A). In addition, W_{eq} is hydrogen-bonded to Glu25 and Glu234, which are potential residues that could serve as the base to deprotonate the water molecule during

catalysis. However, in the crystal structure of the NeuB•inhibitor **T1**• Mn²⁺ complex, the Mn²⁺ ion and metal-bound water molecules were found to be bound and positioned slightly differently (Figure 2.22 B). The partial occupancy of Mn²⁺ ion at the active site indicates that metal ion binds weakly to the NeuB•**T1** complex. This is further supported by the observation of crystal structures of NeuB•**T1** that were devoid of metal obtained during preliminary soaking trials. Moreover, the Mn²⁺ coordination sphere typically has an octahedral symmetry. Compared to the regular octahedral geometry found in NeuB•*N*-acetylmannosaminitol•PEP•Mn²⁺ structure, the Mn²⁺ coordination sphere in the NeuB•inhibitor **T1**• Mn²⁺ complex exhibits a distorted octahedral arrangement. W_{eq} has moved 0.6 Å away from the bound sugar moiety and the distance between the metal and W_{ax} has increased by 0.8 Å (Figure 2.22 B). In addition, W_{ax} has lost contact with the carboxylate of Glu234 and now interacts with Ser213 and the main chain carbonyl of Asp214 and Glu234 instead. Its position rotates ~25° away from its position in the structure of NeuB•*N*-acetylmannosaminitol•PEP•Mn²⁺, resulting in the distortion of the regular octahedral geometry. The partial occupancy of Mn²⁺ and distorted geometry of the metal coordination sphere suggests that the binding of inhibitor **T1** disturbs metal binding in the active site. This may be due to the presence of the C-2 hydrogen of inhibitor **T1**, which is oriented towards W_{eq} and Mn²⁺, and may force W_{eq} away from its preferred position. It could be envisioned that the C-2 hydroxyl group of the normal tetrahedral intermediate would replace W_{eq} as the metal ligand and would not cause a steric clash. In this way the geometry of the octahedral coordination sphere of Mn²⁺ would be maintained. Alternatively, it is also possible that the formation of the tetrahedral intermediate is normally accompanied by a conformational change at the active site of NeuB, but that soaking NeuB with the inhibitor in the solid state was not sufficient to induce the same change. As a consequence, the active site may not be in the optimal

conformation to bind both inhibitor **T1** and the metal ion, and the metal binding could be impaired. Unfortunately, our attempts to obtain a structure by co-crystallization were unsuccessful.

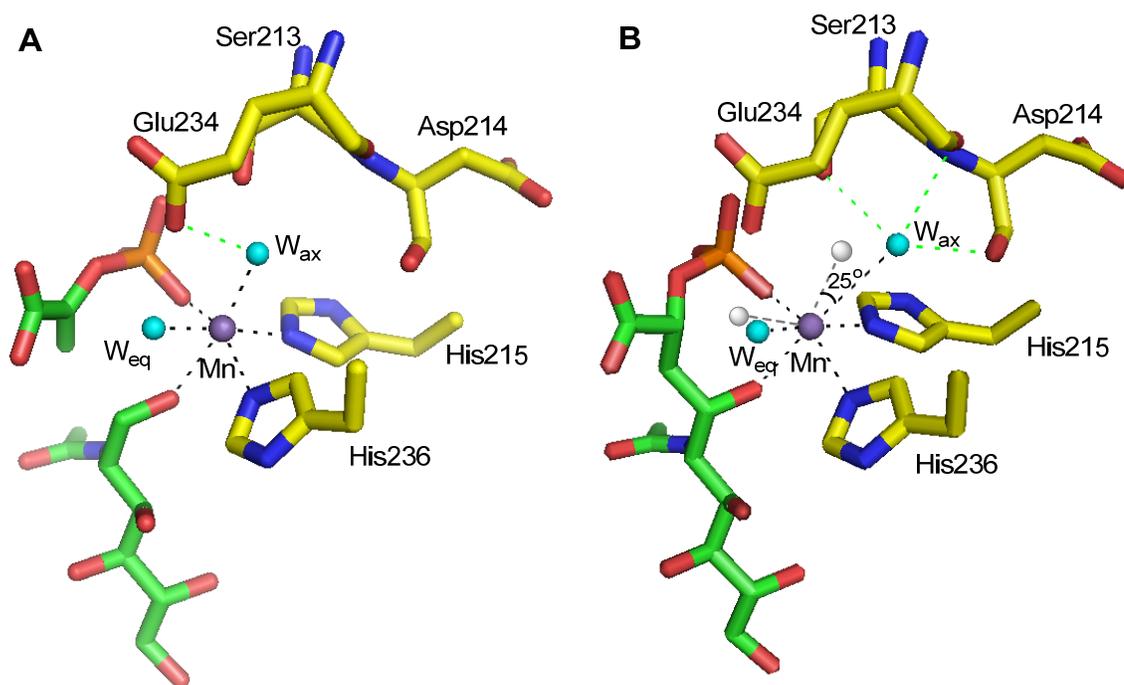


Figure 2.22 The octahedral Mn^{2+} coordination sphere in A) $NeuB \cdot N\text{-acetylmannosaminitol} \cdot PEP \cdot Mn^{2+}$ complex (regular), B) $NeuB \cdot \text{inhibitor T1} \cdot Mn^{2+}$ complex (distorted). The positions of W_{eq} and W_{ax} and their interactions with Mn^{2+} in the $NeuB \cdot N\text{-acetylmannosaminitol} \cdot PEP \cdot Mn^{2+}$ complex are shown in grey in B.

2.5 Conclusions

While much was known about the mechanism of the bacterial sialic acid synthase NeuB prior to this study, no NeuB inhibitors had ever been made before. This chapter described the syntheses and characterization of mechanism-based inhibitors of NeuB that have greatly contributed to our understanding of NeuB mechanism.

The first synthesized oxocarbenium ion intermediate analogue **O1** showed no inhibition against NeuB, even though the analogous compounds **1** and **2** were proven to be potent inhibitors of KDO 8-P synthase and DAH 7-P synthase. Considering the low sequence identity between NeuB and these enzymes, the marked difference between the binding of these analogues likely reflects the significant structural variations between their active sites. While well tolerated by KDO 8-P synthase and DAH 7-P synthase, the extended carbon chain in the structure of oxocarbenium ion intermediate analogue **O1** could cause steric clashes with the residues in the active site of NeuB so as to strongly disfavor its binding.

The tetrahedral intermediate analogue **T1** was prepared as a mixture of two C-2 stereoisomers and was shown to act as a micromolar competitive inhibitor against NeuB with an apparent K_i value of $3.1 \pm 0.1 \mu\text{M}$. The fact that the value of K_i is significantly lower than that of K_M for PEP indicates that analogue **T1** is a close mimic of the tetrahedral intermediate and binds to NeuB much more tightly than the substrate. As previously found for the oxocarbenium ion intermediate analogue of KDO 8-P synthase (**1**), analogue **T1** is also a slow binding inhibitor.

X-ray crystallographic studies showed that NeuB selectively bound the minor stereoisomer bearing the (2*R*)-configuration. Assuming the preference in intermediate analogue binding reflects the stereochemistry of the intermediate in the enzymatic reaction, the tetrahedral

intermediate should also bear a (2*R*)-configuration. This implies that the solvent-derived hydroxide attacks from the *si* face of the oxocarbenium ion intermediate. This was the first time that evidence has been presented outlining the stereochemistry of the tetrahedral intermediates of sialic acid synthase or other related PEP-utilizing enzymes. As the C-2 hydrogen in analogue **T1** was found oriented towards the Mn²⁺ in the active site, the C-2 hydroxyl group of the true tetrahedral intermediate likely serves as a metal ligand. Therefore, it can be postulated that the divalent metal ion not only acts as an electrostatic catalyst to polarize the ManNAc C-1 carbonyl group for nucleophilic attack, but also plays a role in activating the water molecule to attack the oxocarbenium ion intermediate (Figure 2.21).

The partial occupancy of Mn²⁺ ion at the active site, and the distorted geometry of the metal coordination sphere in the NeuB•**T1** complex suggested that the C-2 hydrogen in inhibitor **T1**, which is oriented towards W_{eq} and Mn²⁺, may force W_{eq} away from its preferred position and result in the distorted geometry of the coordination sphere. The C-2 hydroxyl group of the true tetrahedral intermediate would replace W_{eq} as a ligand and relieve this steric clash so that the octahedral coordination sphere of Mn²⁺ may be maintained.

2.6 Future Directions

The stereochemical assignment of the tetrahedral intermediate was based on the assumption that NeuB binds more tightly to the **T1** stereoisomer bearing the same C-2 stereochemistry as the tetrahedral intermediate. This hypothesis was supported by the fact that only (2*R*)-**T1** was found at the active site of NeuB in crystallographic studies. Nevertheless, it is conceivable that binding preferences could be inverted between the intermediate analogue and the true intermediate. 5-Enolpyruvylshikimate-3-phosphate synthase, for example, binds more tightly to phosphonate-based inhibitors bearing the non-natural configuration at the transient tetrahedral center than to those with the natural configuration.¹⁴⁵ Therefore, it would be helpful to synthesize and test more tetrahedral intermediates, especially those bearing a hydroxyl group at the position corresponding to the transient tetrahedral center, to provide further evidence in support of the stereochemical assignment.

Analogue **T2** retains the C-2 hydroxyl group and would likely provide more information regarding the bind of divalent metal ion in the active site. The synthesis of analogue **T2** was unsuccessful and the presence of the C-5 *N*-acetyl group was proposed to account for its decreased stability when compared to the similar tetrahedral intermediate analogue **3** employed in the study of KDO 8-P synthase. This can be tested by synthesizing 5-trifluoroacetamido, 5-amino and 5-hydroxyl variations of analogue **T2** (Figure 2.23, **20**, **21**, **22**) and determining whether they are stable and whether they still serve as inhibitors of NeuB. In particular, compound **20**, which bears a trifluoroacetamido group at C-5, may be the most likely to inhibit

as it still bears an amide group at C-5, yet not decompose via mechanisms requiring neighboring group participation.

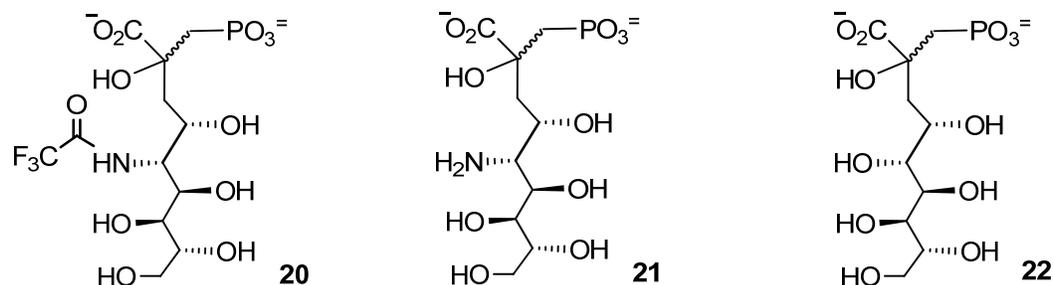


Figure 2.23 Potential tetrahedral intermediate analogues.

Other inhibitors can also be synthesized and tested for inhibition against NeuB (Figure 2.24). Compound **23** mimics the structure of the oxocarbenium intermediate with the phosphonate connected to a sp²-hybridized carbon. Tetrahedral analogue **24** has a hydroxyl group in the place of the carboxylate group of **T2** and would tend to be more stable under basic conditions. Although it will lose electrostatic interactions with the surrounding Lys53 and Lys126, analogue **24** would still be able to engage in favorable hydrogen bonding with Thr110. Furthermore, instead of having the phosphate/phosphonate group covalently connected to the central carbon atom, it would also be worthwhile synthesizing analogues bearing no phosphate/phosphonate and then soaking them into NeuB crystals in the presence of free phosphate. Analogue **25** lacks the phosphate at the C-2 tetrahedral center whereas the hydroxylammonium compound **26** is missing both the carboxylate group and the phosphate group. They would be relatively easy to prepare and would have a good chance to bind to the active site of NeuB. Analogues **24-26** all have hydroxyl groups at their tetrahedral centers and

will thus provide more insights into the interaction of the hydroxyl group and active site metal cofactor.

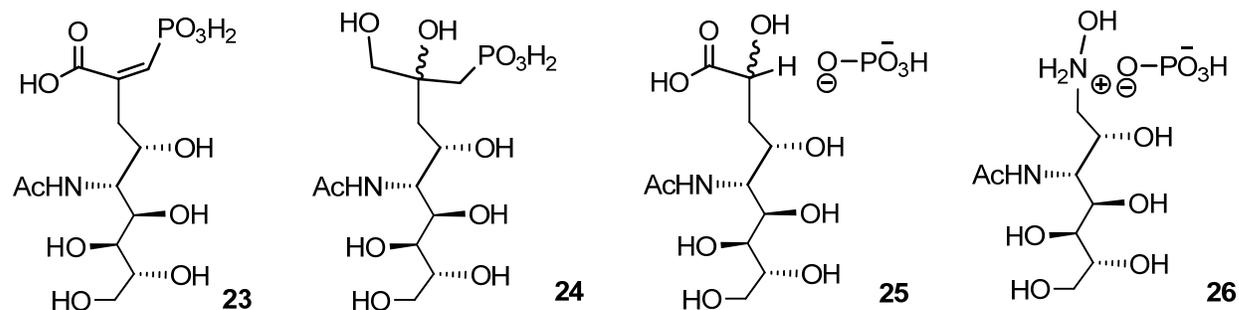


Figure 2.24 Other potential NeuB inhibitors.

Tetrahedral intermediate analogues **T2**, **24**, **25** and **26** would be expected to form complexes that maintain the regular octahedral geometry of the Mn²⁺ coordination sphere if the lack of a hydroxyl group was the main cause of the observed distortion in the NeuB•**T1** complex. At the same time, more studies can be conducted to establish the optimum conditions for the co-crystallization of NeuB with tetrahedral intermediate analogues, so as to address the hypothesis regarding a possible conformational change accompanying the formation of the tetrahedral intermediate as was proposed in the end of Section 2.5.

Lastly, the same strategy described in this chapter can also be applied to studies on other enzymes that are related to sialic acid synthase, such as KDO 8-P synthase, DAH 7-P synthase and pseudaminic acid synthase. Given that the tetrahedral intermediate analogues for KDO 8-P synthase and DAH 7-P synthase were previously synthesized and were reported to be stable, the determination of inhibition kinetics and crystallographic analysis should be straightforward.

2.7 Experimental Procedures

2.7.1 Materials

Chemicals and enzymes were purchased from Sigma-Aldrich and used without further purification unless otherwise noted. Aprotinin was purchased from Roche Diagnostics GmbH. Neu5Ac lyase (SAL) was purchased from Jülich Fine Chemicals (Germany). Bacterial purine nucleoside phosphorylase (PNP) was purchased from Sigma. 2-Amino-6-mercapto-7-methylpurine ribonucleoside (MESG) was purchased from Berry and Associates. Pyridine, triethylamine, methanol (MeOH) and methylene chloride (CH₂Cl₂) were distilled over CaH₂ under an atmosphere of N₂. Tetrahydrofuran (THF) was purified by distillation over sodium under an atmosphere of N₂. Chelex[®] 100 resin (200-400 mesh, Na⁺ form), AG[®] 1-X8 resin (100-200 mesh, formate form), Amberlite[®] IR 120 (H⁺ form) resin and Bio-Gel[®] P-2 resin were purchased from Bio-Rad Laboratories. Sepharose Fast Flow resin was purchased from Pharmacia Biotech. Amicon Ultra Centricons (4 mL or 15 mL, 10 000 MWCO) were purchased from Millipore.

2.7.2 General Methods

All proteins were handled at 4 °C unless otherwise stated. Protein concentrations were determined according to the method of Bradford¹⁴⁶ using a Cary3E UV/Vis spectrophotometer with bovine serum albumin (BSA) as the standard. Protein purity was determined using SDS-PAGE gel electrophoresis and visualized using Coomassie Blue stain according to the method

described by Laemmli.¹⁴⁷ Protein molecular masses were determined using BSA (66 kDa) and carbonic anhydrase (29 kDa) as standard masses.

¹H NMR spectra were obtained on Bruker AV300 or AV400 NMR spectrometers at field strengths of 300 or 400 MHz. ¹³C NMR spectra were obtained on Bruker AV400 NMR spectrometer at a field strength of 100 MHz. Proton-decoupled ³¹P NMR spectra were obtained on Bruker AV200 or AV400 NMR spectrometers at field strengths of 81 or 162 MHz, respectively. Mass spectra were obtained on a Waters Micromass LCT mass spectrometer using electrospray ionization (ESI-MS).

2.7.3 Over-expression and Purification of *Neisseria meningitidis* NeuB

2.7.3.1 Over-expression and purification of NeuB for kinetics study.

N. meningitidis NeuB was previously cloned into the pET 30 Xa/LIC vector and subsequently transformed into CaCl₂ competent *E. coli* BL21 (DE3) cells (Novagen) for expression. Cells were incubated in 10 mL of Lucia-Bertani (LB) medium containing 50 mg/L kanamycin at 37 °C/225 rpm for 10 hours. The overnight culture was then poured into 500 mL of LB medium containing 50 mg/L kanamycin at 37 °C/255 rpm until an OD₆₀₀ of 0.6 had been reached. Over-expression was carried out by induction with 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG) and incubation at 37 °C for 4 h. Cells were harvested by centrifuging at 6,000 × g for 30 min, and then resuspended in 10 mL of phosphate buffer (20 mM, pH 8.0) 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 6,000 × g for 1 h and passed through a 0.22 μm filter (Millipore).

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 Tris-HCl buffer (20 mM, pH 8.0, containing 0.5 M NaCl and 5 mM of imidazole). The clarified 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). Eluate fractions that were eluted with 500 mM imidazole and showed absorbance at 280 nm were collected. These fractions were concentrated using Amicon Ultra Centricons before flash freezing with liquid N₂ in the presence of 10% glycerol and 2 mM DTT.

2.7.3.2 Cloning, over-expression and purification of NeuB for crystallization.

All molecular biology procedures were performed as described previously.⁴² Briefly, the untagged NeuB enzyme was cloned into the pCWori+ vector and subsequently transformed into electrocompetent *E. coli* BL21 (DE3) cells (Novagen) for expression. Over-expression of NeuB was carried out by induction with 0.5 mM of IPTG at OD₆₀₀ ~0.6 with overnight shaking at 20 °C. Cells were harvested by centrifuging at 5,000 rpm (6,200 × g) for 15 min, resuspended, and lysed at 20,000 psi using a high-pressure homogenizer (Avestin) in the presence of EDTA-free protease inhibitor cocktail (Roche Applied Science). The lysate was subsequently centrifuged at 40,000 rpm (161,000 × g) for 35 min and the supernatant containing the target protein was purified by a series of chromatographic procedures including ion-exchange and gel filtration steps. Purified NeuB enzyme was concentrated to 10 mg/mL and used for crystallization.

2.7.4 Syntheses of Intermediate Analogues

2.7.4.1 Synthesis of analogue O1

A solution of potassium cyanide (1.76 g, 27 mmol) in 12 mL of distilled H₂O was adjusted to a pH of 7.5 using acetic acid. A solution of ManNAc (2.0 g, 9 mmol) in 12 mL of distilled H₂O was slowly added dropwise and the pH was maintained at ~7.5. After the reaction mixture was stirred for 20 min at room temperature, the pH was re-adjusted to ~4 with acetic acid. The solution was then bubbled with argon for 5 h before water was removed under reduced pressure. Compound **4** was obtained as white solid and was used without further purification. ¹H NMR (D₂O, 400 MHz) δ 2.03 (s, 3H, 2*S*-CH₃), 2.06 (s, 3H, 2*R*-CH₃), 3.50 (dd, 1H, *J*_{5,6} 9.1 Hz, 2*S*-H-5), 3.53 (dd, 1H, *J*_{5,6} 11.7 Hz, 2*R*-H-5), 3.67 (dd, 1H, *J*_{6,7a} 6.3 Hz, *J*_{7a,7b} 11.4 Hz, H-7a), 3.78 (m, 1H, H-6), 3.86 (dd, 1H, *J*_{6,7b} 2.3 Hz, *J*_{7a,7b} 11.4 Hz, H-7b), 3.98 (dd, 1H, *J*_{3,4} 10.3 Hz, 2*S*-H-4), 4.13 (dd, 1H, *J*_{3,4} 10.0 Hz, 2*R*-H-4), 4.38 (d, 3H, 2*R*-H-1), 4.40 (dd, 1H, *J*_{3,4} 10.3 Hz, 2*S*-H-3), 4.52 (dd, 1H, *J*_{3,4} 10.1 Hz, 2*R*-H-3), 4.61 (d, 3H, 2*S*-H-1). ESI-MS (+) *m/z* 271 (M+Na⁺).

5% Pd/BaSO₄ (0.57 g) was suspended in 6 mL of water solution containing compound **4** (2.24 g, 27 mmol), and hydrogenated on a Parr hydrogenator at 40 psi for 16 h. The reaction mixture was filtered through celite and passed through Amberlite[®] IR 120 (H⁺ form) resin. Compound **5** containing both C-2 epimers was obtained as a white solid (2.08 g, 92%). ESI-MS (+) *m/z* 274 (M+Na⁺).

As an attempt to separate the two C-2 epimers, compound **5** was peracetylated to be purified by silica gel column chromatography. To a flask containing compound **5** (2.08 g, 8 mmol) and pyridine (70 mL) was added acetic anhydride (40 mL, 0.42 mol) and *N,N'*-dimethyl-

4-aminopyridine (DMAP, 5 mg, 0.042 mmol). The reaction mixture was stirred for 16 h at room temperature before the removal of pyridine and excess acetic anhydride *in vacuo*. The resultant syrup was then dissolved in 150 mL of CH₂Cl₂, washed with saturated NaHCO₃ solution and brine, and dried over anhydrous Na₂SO₄. After removal of solvent, the peracetylated compound **6** was yielded as a mixture of four stereoisomers (α -2*R*, β -2*R*, α -2*S*, β -2*S*, total yield 2.6 g, 68%, ESI-MS (+) *m/z* 484 (M+Na⁺)). The four stereoisomers of compound **6** appeared as two spots (A and B) on TLC (R_{fA} = 0.26, R_{fB} = 0.32, CH₂Cl₂:EtOAc = 1:1), which were separated by silica gel column chromatography. The separated compounds corresponding to each spot were later proven to contain a (2*R*)- and a (2*S*)- stereoisomer with either α - or β -configuration.

Isolated compound **6** (1.3 g, 2.8 mmol) was dissolved in 3 mL of distilled methanol and cooled down to 0 °C in the ice bath. A 40 mM solution of sodium methoxide in methanol (1 mL) was then slowly added. The reaction mixture was then allowed to stand at -20 °C for 16 h before the sodium methoxide was quenched by bubbling with CO₂. It was then diluted with H₂O and passed through Amberlite[®] IR 120 (H⁺ form) resin. The regenerated compound **5** was obtained in high yield (0.64 g, 91%).

The final step was the one-pot reductive amination as described by Du *et al.*⁴⁸ Compound **5** (100 mg, 0.4 mmol, **A** or **B** or unseparated) was dissolved in MeOH/H₂O (4 mL/ 4 mL), to which glyphosate (200 mg, 1.18 mmol) was added and the pH was adjusted to 6.2 with triethylamine (TEA). While the solution was stirred at 80 °C, sodium cyanoborohydride (NaBH₃CN) was added in five batches (100 mg each, 1.6 mmol) over a period of 20 h. MeOH was then removed *in vacuo* and the water solution of the reaction mixture was treated with Amberlite[®] IR 120 (H⁺ form) resin. The resultant solution was then loaded on a column containing 10 mL of AG[®] 1-X8 resin (100-200 mesh, formate form) and subjected to a stepwise

elution with 100 mL of distilled H₂O, 0.4, 0.6, 0.8 and 1.0 M formic acid. Eluates were collected in 50 mL-fractions, concentrated and analyzed by ³¹P NMR spectroscopy. Those containing analogue **O1** ($\delta \sim 8.0$ and 7.0) but not glyphosate ($\delta \sim 60$) were collected and repeatedly rotovapped with distilled H₂O. After treatment with Amberlite[®] IR 120 (Na⁺ form) resin and lyophilization, analogue **O1** was obtained as a white solid (66 mg, 39%). ¹H NMR (D₂O, 300 MHz) δ 1.94 (s, 3H, CH₃), 3.20-3.35 (m, 2H, H-3 and H-4), 3.31-3.46 (m, 2H, H-1), 3.48 (dd, 1H, $J_{5,6}$ 8.9 Hz, H-5), 3.64 (dd, 1H, $J_{6,7a}$ 6.3 Hz, $J_{7a,7b}$ 11.6 Hz, H-7a), 3.76 (ddd, 1H, $J_{5,6}$ 8.9 Hz, $J_{6,7a}$ 6.0 Hz, $J_{6,7b}$ 2.3 Hz, H-6), 3.86 (dd, 1H, $J_{6,7b}$ 2.3 Hz, $J_{7a,7b}$ 11.8 Hz, H-7b), 3.91 and 4.17 (2 \times dd, 2H, $J_{H,P}$ 16.2 Hz, CH₂-P), 3.99 (s, 2H, CH₂-CO₂H), 4.42 and 4.61 (m, 1H, (*R*)- or (*S*)-H-2). ESI-MS (-) m/z 403 (M-H⁺).

2.7.4.2 Synthesis of inhibitor T1

2-Acetamido-2-deoxy-3,4,5,6-tetra-O-acetyl-D-mannose O-methyloxime (7)- Oxime **7** was synthesized as reported by Weitz and Bednarski¹²⁸. *N*-acetyl-D-mannosamine (ManNAc, 2.23 g, 10.1 mmol) and methoxyamine hydrochloride (1 g, 12 mmol) were dissolved in 40 mL of dry pyridine and stirred at room temperature for 24 hrs. The progress of the reaction was followed by ESI-MS(+). After ManNAc (m/z 244 (M+Na⁺)) was fully converted to 2-acetamido-2-deoxy-D-mannose methyloxime (m/z 273 (M+Na⁺)), acetic anhydride (15 mL, 159 mmol) and *N,N*-dimethyl-4-aminopyridine (DMAP, 3 mg, 0.025 mmol) were added. The reaction mixture was kept stirred for 16 h before the removal of pyridine and excessive acetic anhydride *in vacuo*. The resultant syrup was redissolved in 150 mL of ethyl acetate, washed with saturated NaHCO₃ solution and brine and dried over anhydrous Na₂SO₄. After removal of solvent *in vacuo*, oxime **7** was recrystallized from ethyl acetate/toluene (1:1) as white crystals (3.72 g, 88.9%) with a *syn/anti* ratio of 1:5. ¹H NMR (CDCl₃, 400 MHz) δ 2.00 (s, 3H, CH₃),

2.061 (s, 3H, CH₃), 2.063 (s, 3H, CH₃), 2.08 (s, 3H, CH₃), 2.13 (s, 3H, CH₃), 3.84 (s, 3H, *anti*-N=O-CH₃), 3.92 (s, 3H, *syn*-N=O-CH₃), 4.07 (dd, 1H, $J_{5,6}$ 5.7 Hz, $J_{6,6}$ 12.4 Hz, *syn*-H-6), 4.10 (dd, 1H, $J_{5,6}$ 5.7 Hz, $J_{6,6}$ 12.4 Hz, *anti*-H-6), 4.24 (dd, 1H, $J_{6,6}$ 12.4 Hz, $J_{5,6}$ 2.8 Hz, *syn*-H-6), 4.27 (dd, 1H, $J_{6,6}$ 12.4 Hz, $J_{2,3}$ 2.87 Hz, *anti*-H-6), 4.84 (ddd, 1H, $J_{1,2}$ 5.4 Hz, $J_{2,3}$ 6.2 Hz, $J_{\text{NH},2}$ 8.6 Hz, H-2), 5.13 (m, 1H, H-5), 5.37 (dd, 1H, $J_{2,3}$ 7.8 Hz, $J_{3,4}$ 2.6 Hz, *syn*-H-3), 5.40 (dd, 1H, $J_{2,3}$ 6.3 Hz, $J_{3,4}$ 3.1 Hz, *anti*-H-3), 5.44 (dd, 1H, $J_{3,4}$ 3.2 Hz, $J_{4,5}$ 7.8 Hz, *anti*-H-4), 5.60 (dd, 1H, $J_{3,4}$ 2.6 Hz, $J_{4,5}$ 7.8 Hz, *syn*-H-4), 6.02 (d, 1H, $J_{\text{NH},2}$ 8.7 Hz, *anti*-NH), 6.05 (d, 1H, $J_{\text{NH},2}$ 8.7 Hz, *syn*-NH), 6.73 (d, 1H, $J_{1,2}$ 6.35 Hz, *syn*-H-1), 7.33 (d, 1H, $J_{1,2}$ 5.36 Hz, *anti*-H-1). See Appendix Figure A.1 for detailed information. ESI-MS (+) m/z 441 (M+Na⁺).

2-Acetamido-2-deoxy-3,4,5,6-tetra-O-acetyl-D-mannose (8)- Oxime **8** was subjected to ozonolysis as previously described.¹²⁸ ESI-MS (+) showed that the reaction was completed. Compound **8** was obtained as colorless syrup and was used without further purification. ¹H NMR (CDCl₃, 400 MHz) δ 2.04 (s, 3H, CH₃), 2.060 (s, 3H, CH₃), 2.062 (s, 3H, CH₃), 2.073 (s, 3H, CH₃), 2.15 (s, 3H, CH₃), 4.13 (dd, 1H, $J_{5,6}$ 5.5 Hz, $J_{6,6}$ 12.4 Hz, H-6), 4.27 (dd, 1H, $J_{5,6}$ 2.9 Hz, $J_{6,6}$ 12.5 Hz, H-6), 4.77 (dd, 1H, $J_{2,3}$ 5.7 Hz, $J_{\text{NH},2}$ 7.5 Hz, H-2), 5.14 (m, 1H, H-5), 5.46 (dd, 1H, $J_{2,3}$ 5.6 Hz, $J_{3,4}$ 3.4 Hz, H-3), 5.49 (dd, 1H, $J_{3,4}$ 3.4 Hz, $J_{4,5}$ 7.7 Hz, H-4), 6.42 (d, 1H, $J_{\text{NH},2}$ 8.7 Hz, NH), 9.57 (s, 1H, H-1). ESI-MS (+) m/z 412 (M+Na⁺).

Methyl 5-acetamido-3,5-dideoxy-2-methylidene-4,6,7,8,9-penta-O-acetyl-D-glycero-D-galacto-2-nonulosonate (9). To a solution of compound **8** (3.72 g, 8.9 mmol) in acetonitrile (ACN)/HCl (60 mL, 20:1) were added methyl bromomethylacrylate (0.6 mL, 5 mmol) and a suspension of indium powder (100~200 mesh, 1.72 g, 15 mmol) in ACN/HCl (5 mL, 20:1). The mixture was vigorously stirred at 45 °C for 3.5 h and then the indium clump was removed. Fresh methyl bromomethylacrylate (0.6 mL, 5 mmol) and a suspension of indium (0.5 g, 4.3 mmol) in

ACN/HCl (5 mL, 20:1) were added and the mixture was vigorously stirred at 45 °C for an additional 3 h. The reaction mixture was filtered through celite and the solvent was removed *in vacuo*. The resultant solid was re-dissolved in pyridine (40 mL), and acetic anhydride (3.62 mL, 38 mmol) and DMAP (5 mg) were added. The reaction mixture was stirred at RT for 16 h and then the solvent was removed *in vacuo*. The resultant syrup was dissolved in ethyl acetate (100 mL) and washed with brine (100 mL). Silica gel column chromatography eluting with ethyl acetate ($R_f \sim 0.3$) gave a mixture of the two epimers ((4*S*):(*4R*) 3:1) as a white solid (2.01 g, 40%). Recrystallization from toluene/petroleum ether (1:1) gave the pure (*4S*)-isomer, **9**, as a white solid (1.35 g, 27%). $^1\text{H NMR}$ (CDCl_3 , 400 MHz) δ 1.98 (s, 3H, CH_3), 2.03 (s, 3H, CH_3), 2.04 (s, 3H, CH_3), 2.04 (s, 3H, CH_3), 2.06 (s, 3H, CH_3), 2.14 (s, 3H, CH_3), 2.34 (dd, 1H, $J_{3,3}$ 14.1 Hz, $J_{3,4}$ 9.0 Hz, H-3), 2.59 (dd, 1H, $J_{3,3}$ 14.0 Hz, $J_{3,4}$ 3.9 Hz, H-3), 3.98 (dd, 1H, $J_{8,9}$ 5.7 Hz, $J_{9,9}$ 12.5 Hz, H-9), 4.26 (dd, 1H, $J_{8,9}$ 3.1 Hz, $J_{9,9}$ 12.5 Hz, H-9), 4.49 (ddd, 1H, $J_{4,5}$ 1.5 Hz, $J_{5,6}$ 10.3 Hz, $J_{\text{NH},5}$ 10.4 Hz, H-5), 5.02 (ddd, 1H, $J_{7,8}$ 8.1 Hz, $J_{8,9}$ 3.1 Hz, $J_{8,9}$ 5.7 Hz, H-8), 5.14 (ddd, 1H, $J_{3,4}$ 4.1 Hz, $J_{3,4}$ 8.9 Hz, $J_{4,5}$ 1.5 Hz, H-4), 5.21 (dd, 1H, $J_{5,6}$ 10.3 Hz, $J_{6,7}$ 2.1 Hz, H-6), 5.35 (dd, 1H, $J_{6,7}$ 2.1 Hz, $J_{7,8}$ 8.1 Hz, H-7), 5.56 (d, 1H, $J_{\text{NH},5}$ 8.7 Hz, NH), 5.57 (d, 1H, $J_{1',1'}$ 1.0 Hz, H-1'), 6.15 (d, 1H, $J_{1',1'}$ 1.0 Hz, H-1'). See Appendix Figure A.2 for detailed information. ESI-MS (+) m/z 554 ($\text{M}+\text{Na}^+$).

Methyl 5-acetamido-4,6,7,8,9-pentaacetoxy-2-dibenzylphosphorylnonanoate (**10**).

Compound **9** (300 mg, 0.56 mmol) was dissolved in methylene chloride (30 mL) and O_3 was bubbled through the solution at -78 °C until a blue color persisted. Excess O_3 was purged by bubbling argon through the solution at -78 °C. A solution of sodium borohydride (68 mg, 1.8 mmol) in ethanol (30 mL) was then added. The mixture was allowed to stand at **a**) -78 °C for 24 h or **b**) 25 °C for 30 min. After removal of the solvent *in vacuo*, the solid was re-dissolved in

ethyl acetate, which was extracted with brine and dried over anhydrous Na₂SO₄. After removal of the solvent *in vacuo*, the reduced product was used in the next step without further purification. Dibenzyl *N,N*-diethylphosphoramidite (600 μL, 2 mmol) was added to a solution of the reduced product (400 mg, 1.3 mmol) and 1,2,4-triazole (40 mg, 0.59 mmol) in CH₂Cl₂ (10 mL) and stirred at RT for 16 h. After removal of the solvent *in vacuo* the reaction mixture was redissolved in diethyl ether (50 mL). The solution was then cooled to -78 °C and stirred for 30 min after 1 mL of 30% H₂O₂ was added. The organic layer was then washed with saturated Na₂S₂O₃ and dried over Na₂SO₄. After removal of the solvent *in vacuo*, the resultant syrup was washed with petroleum ether and purified using silica gel column chromatography (ethyl acetate, R_f = 0.2). Compound **10** was obtained as colorless syrup, and was found to be comprised of a 4.5:1 mixture of (2*S*)-**10**:(2*R*)-**10** using condition **a** or a 1.2:1 mixture of (2*S*)-**10**:(2*R*)-**10** using condition **b**. ¹H NMR (CDCl₃, 400 MHz) δ 1.92 (s, 3H, (2*S*)-CH₃), 1.93 (s, 3H, (2*R*)-CH₃), 1.98 (s, 3H, (2*S*)-CH₃), 2.00 (s, 3H, (2*R*)-CH₃), 2.019 (s, 3H, CH₃), 2.023 (s, 3H, CH₃), 2.031 (s, 3H, CH₃), 2.077 (s, 3H, (2*R*)-CH₃), 2.092 (s, 3H, (2*S*)-CH₃), 2.1 (m, 2H, H-3), 3.96 (dd, 1H, *J*_{8,9a} 5.7 Hz, *J*_{9a,9b} 12.4 Hz, (2*S*)-H-9a), 4.06 (dd, 1H, *J*_{8,9a} 6.6 Hz, *J*_{9a,9b} 6.7 Hz, (2*R*)-H-9a), 4.18 (dd, 1H, *J*_{8,9b} 2.5 Hz, *J*_{9a,9b} 6.7 Hz, (2*R*)-H-9b), 4.25 (dd, 1H, *J*_{8,9b} 2.8 Hz, *J*_{9a,9b} 12.4 Hz, (2*S*)-H-9b), 4.46 (ddd, 1H, *J*_{4,5} 1.5 Hz, *J*_{5,6} 10.4 Hz, *J*_{NH,5} 10.5 Hz, (2*S*)-H-5), 4.60 (ddd, 1H, *J*_{4,5} 1.5 Hz, *J*_{5,6} 10.4 Hz, *J*_{NH,5} 10.5 Hz, (2*R*)-H-5), 4.88 (ddd, 1H, *J*_{3,4} 4.6 Hz, *J*_{3,4} 8.2 Hz, *J*_{4,5} 1.5 Hz, H-4), 5.00 (ddd, 1H, *J*_{7,8} 8.1 Hz, *J*_{8,9a} 5.7 Hz, *J*_{8,9b} 3.0 Hz, H-8), 5.05 -5.17 (m, 5H, Bn-CH₂ and H-2), 5.22 (dd, 1H, *J*_{5,6} 10.4 Hz, *J*_{6,7} 1.8 Hz, (2*S*)-H-6), 5.25 (dd, 1H, *J*_{5,6} 10.4 Hz, *J*_{6,7} 1.8 Hz, (2*R*)-H-6), 5.33 (dd, 1H, *J*_{6,7} 1.8 Hz, *J*_{7,8} 8.1 Hz, (2*S*)-H-7), 5.37 (dd, 1H, *J*_{6,7} 1.8 Hz, *J*_{7,8} 8.1 Hz, (2*R*)-H-7), 5.65 (d, 1H, *J*_{NH,5} 10.5 Hz, (2*R*)-NH), 5.70 (d, 1H, *J*_{NH,5} 10.5 Hz, (2*S*)-NH), 7.34 (m, 10H, Bn), ESI-MS (+) *m/z* 818 (M+Na⁺).

5-acetamido-4,6,7,8,9-pentahydroxy-2-phosphorylnonanoic acid ditriethylammonium salt (T1). To a solution of **10** (80 mg, 0.10 mmol) in methanol (20 mL) was added Pd/C (10%, 30 mg) and the mixture was stirred under H₂ (1 atm) for 1 h. After filtration through celite and removal of the solvent *in vacuo*, the resulting solid was dissolved in 1:1 MeOH/H₂O (20 mL) containing 10% triethylamine (TEA), and was allowed to stand at -20 °C for 14 h. After removal of the methanol *in vacuo*, the remaining aqueous solution was diluted with distilled H₂O and lyophilized to dryness. The solid was then purified by passage through a Bio-Gel P-2 Column (2.5 cm × 44 cm) eluting with distilled water (0.2 mL min⁻¹). The fractions containing inhibitor **T1**, as analyzed by negative ESI-MS, were lyophilized twice with distilled water to give **T1** as a white solid (25 mg, 42%). The ratio of (2*S*)-**T1**:(2*R*)-**T1** reflected the ratio of (2*S*)-**10**:(2*R*)-**10** used in the reaction, as determined by ³¹P NMR spectroscopy. Due to the presence of many overlapping signals, ¹H and ¹³C NMR data are only given for the major isomer (2*S*)-**T1** obtained from deprotection of the 4.5:1 mixture of (2*S*)-**10**:(2*R*)-**10**. ¹H NMR ((2*S*)-**T1**, D₂O, 400 MHz) δ 1.29 (t, 2H, TEA CH₃), 1.68 (ddd, 1H, *J*_{3a,3b} 11.8 Hz, *J*_{3a,2} 11.0 Hz, H-3a), 1.94 (dd, 1H, *J*_{3a,3b} 12.1 Hz, *J*_{3b,4} 11.0 Hz, H-3b), 2.07 (s, 3H, CH₃), 3.21 (q, 2H, TEA CH₂) 3.48 (dd, 1H, *J*_{7,8} 8.9 Hz, H-7), 3.65 (dd, 1H, *J*_{8,9a} 6.4 Hz, *J*_{9a,9b} 11.8 Hz, H-9a), 3.79 (ddd, 1H, *J*_{7,8} 8.9 Hz, *J*_{8,9a} 6.3 Hz, *J*_{8,9b} 2.7 Hz, H-8), 3.85 (dd, 1H, *J*_{8,9b} 2.6 Hz, *J*_{9a,9b} 11.8 Hz, H-9b), 3.94 (dd, 1H, *J*_{5,6} 10.1 Hz, H-5), 4.01 (dd, 1H, *J*_{5,6} 10.0 Hz, H-6), 4.37 (dd, 1H, *J*_{3,4} 11 Hz, H-4), 4.54 (ddd, 1H, *J*_{2,3} 2.8 Hz, *J*_{2,3} 10 Hz, *J*_{2,p} 10 Hz, H-2). ¹³C NMR ((2*S*)-**1**, MeOD, 100 MHz) δ 9.3 (TEA), 22.6 (CH₃), 38.3 (C-3), 47.5 (TEA), 54.4 (C-5), 63.4 (C-9), 65.9 (C-4), 67.9 (C-6), 69.8 (C-7), 70.5 (C-8), 72.8 (C-2), 172.6 (C=O), 178.2 (O=C-OH). ³¹P NMR (D₂O, pD 7, 162 MHz) δ 2.30 ((2*R*)-**T1**), 2.75((2*S*)-**T1**). See Appendix Figure A.3 and Figure A.4 for detailed information. HRMS calculated for C₁₁H₂₁NO₁₂P (M-H⁺) 390.0801, found 390.0806.

2.7.4.3 Attempted synthesis of analogue T2

Yields and NMR spectroscopic information are not available for some compounds because they are not stable enough to be purified by conventional chromatographic methods.

Methyl 5-acetamido-3,5-dideoxy -4,6,7,8,9-penta-O-acetyl-D-glycero-D-galacto-2-nonate (**11**). Compound **9** (300 mg, 0.56 mmol) was dissolved in methylene chloride (30 mL) and O₃ was bubbled through the solution at -78 °C until a blue color persisted. Excess O₃ was purged by bubbling argon through the solution at -78 °C. Dimethyl sulfide (1 mL, 13.5 mmol) was added and the solution was warmed up to room temperature. After removal of CH₂Cl₂, dimethyl sulfide and generated dimethyl sulfoxide (DMSO) *in vacuo*, compound **11** was obtained as a colorless syrup (295 mg, 98%). ¹H NMR (CDCl₃, 400 MHz) δ 2.01 (s, 3H, CH₃), 2.039 (s, 3H, CH₃), 2.042 (s, 3H, CH₃), 2.048 (s, 3H, CH₃), 2.10 (s, 3H, CH₃), 2.11 (s, 3H, CH₃), 2.83 (dd, 1H, *J*_{3,3} 16.5 Hz, *J*_{3,4} 9.6 Hz, H-3), 3.16 (dd, 1H, *J*_{3,3} 16.48 Hz, *J*_{3,4} 6.4 Hz, H-3), 3.99 (dd, 1H, *J*_{8,9} 5.5 Hz, *J*_{9,9} 12.5 Hz, H-9), 4.26 (dd, 1H, *J*_{8,9} 2.9 Hz, *J*_{9,9} 12.5 Hz, H-9), 4.49 (ddd, 1H, *J*_{4,5} 1.2 Hz, *J*_{5,6} 10.3 Hz, *J*_{NH,5} 10.4 Hz, H-5), 5.06 (ddd, 1H, *J*_{7,8} 8.4 Hz, *J*_{8,9} 3.0 Hz, *J*_{8,9} 5.7 Hz, H-8), 5.13 (ddd, 1H, *J*_{3,4} 6.4 Hz, *J*_{3,4} 6.6 Hz, *J*_{4,5} 1.2 Hz, H-4), 5.31 (dd, 1H, *J*_{5,6} 10.4 Hz, *J*_{6,7} 2.0 Hz, H-6), 5.40 (dd, 1H, *J*_{6,7} 2.0 Hz, *J*_{7,8} 8.5 Hz, H-7), 5.56 (d, 1H, *J*_{NH,5} 10.5 Hz, NH). ESI-MS (+) *m/z* 556 (M+Na⁺).

Dibenzyl methylphosphonate (**12**). Compound **12** was synthesized as previously reported.¹⁴¹ ¹H NMR (CDCl₃, 400 MHz) δ 1.48 (3H, d, *J*_{H,P} 17 Hz, CH₃), 4.97 (4H, m, CH₂Ph), 7.25-7.37 (10H, m, Ph). ESI-MS (+) *m/z* 299 (M+Na⁺).

Methyl 5-acetamido-4,6,7,8,9-pentaacetoxy-2-dibenzylphosphonylmethyl-2-hydroxynonanoate (13). Compound **13** was synthesized using the method previously reported by Grison *et al.*⁵⁵ Dibenzyl methylphosphonate (92 mg, 0.33 mmol) in THF (5 mL) was added to a solution of *n*-butyllithium (222 μ L, 1.6 M in hexane) in THF (5 mL) at -78 °C. After stirring for 20 min, compound **11** in THF (5 mL) was slowly added. The reaction mixture was kept in a -85 °C freezer for 7 days before being quenched with a saturated HCl Et₂O solution (2 mL) at -35 °C. To the reaction mixture was then added H₂O (10 mL), and the mixture was warmed up to room temperature and extracted with CH₂Cl₂ (3 \times 15 mL). The organic layer was dried over anhydrous Na₂SO₄ and the solvent was removed *in vacuo* to give a yellowish syrup, which was then repeatedly washed with petroleum ether to give compound **13** as a solid. ESI-MS (+) *m/z* 832 (M+Na⁺).

General procedure for hydrogenolysis using Pd/C or PdOH in methanol. To a solution of **13** (100 mg) or **15** (50 mg) in methanol (20 mL) was added Pd/C (10%, 30 mg) and the mixture was stirred under H₂ (1 atm) for 1 h. After filtration through celite, the solvent was removed *in vacuo*. The procedure for the hydrogenolytic debenylation of compound **17** was the same except that PdOH (10%, 30 mg) was used in the place of Pd/C.

General procedure for hydrolysis using Et₃N in MeOH/H₂O. Compound **13** (100 mg) or **14** (50 mg) was dissolved in 1:1 MeOH/H₂O (20 mL) containing 10% TEA, and was allowed to stand at -20 °C for 14 h. After removal of the MeOH *in vacuo*, the remaining aqueous solution was diluted with distilled H₂O and lyophilized to dryness.

Benzyl 5-acetamido-3,5-dideoxy-2-methylidene-4,6,7,8,9-penta-O-benzyl-D-glycero-D-galacto-2-nonulosonate (16). Compound **9** (100 mg, 0.19 mmol) was first subjected to

hydrolysis using 10% TEA in MeOH/ H₂O as described above. The lyophilized solid was dissolved in dimethylformamide (DMF, 5 mL) and slowly added to a suspension of NaH (60%, 120 mg, 3 mmol) and tetrabutylammonium iodide (TBAI, 5 mg, 0.013 mmol) in THF (10 mL) under an atmosphere of argon. The reaction mixture was stirred at room temperature for 30 min. Benzyl bromide (BnBr, 340 μ L, 4 mmol) was then added in 3 portions over 24 h and the progress of the reaction was monitored by ESI-MS (+). The reaction was quenched by first pouring the reaction mixture over dry ice and then adding H₂O (5 mL). The resultant mixture was diluted in 100 mL of ethyl acetate, successively extracted with brine (3 \times 30 mL) and dried over anhydrous NaSO₄. After the removal of solvent *in vacuo*, the product was purified by silica gel column chromatography eluting with 1:9 ethyl acetate:dichlormethane (R_f=0.21). Compound **16** was obtained as a white solid (84.7 mg, 53%). ESI-MS (+) *m/z* 870 (M+Na⁺).

Benzyl *5-acetamido-4,6,7,8,9-pentabenzoxo-2-dibenzylphosphonylmethyl-2-hydroxynonanoate* (**17**). The procedures for the synthesis of compound **11** and **13** were followed. ESI-MS (+) *m/z* 1148 (M+Na⁺).

5-Acetamido-2-dibenzylphosphonylmethyl-2,4,6,7,8,9-hexahydroxynonanoic acid (analogue **T2**) formed by cyanide-mediated hydrolysis of compound **14**. Compound **14** (10 mg) was added to 95% EtOH (4 mL) containing NaCN (7 mg) and was stirred for 10 min at 65 °C. ESI-MS (-) indicated the formation of analogue **T2** (*m/z* 404 (M-H⁺)). After the removal of solvent *in vacuo*, the solid was re-dissolved in H₂O (3 mL), with the pH adjusted to ~7.0, and was then purified by passage through a Bio-Gel P-2 Column (2.5 cm \times 44 cm) eluting with distilled water (0.2 mL min⁻¹). The fractions containing analogue **T2**, as analyzed by negative ESI-MS spectrometry, were lyophilized twice with distilled water to give **T2** as a white solid

(0.5 mg, 16%). Analogue **T2** was dissolved in H₂O and re-examined using ESI-MS (-). It was shown that >70% of analogue **T2** had turned into compound **18** (m/z 386 (M-H⁺)).

2.7.5 Kinetic Studies

2.7.5.1 Test of inhibition by analogue **O1**

Inhibition kinetics were measured using a slight modification of a previously reported continuous coupled assay.⁴² A cuvette containing Tris-acetic acid (pH 8.0) (100 mM, 0.8 ml final volume), ManNAc (30 mM, saturating), PEP (1 mM, saturating), MnCl₂ (1 mM), NADH (160 μM), lactate dehydrogenase (LDH, 250 units), sialic acid lyase (SAL, 20 units), and dithiothreitol (DTT, 1 mM) was incubated for 5 min at 37 °C allow all the components to allow thermal equilibrium. The enzymatic reaction was initiated by the addition of His₆-tagged NeuB (1 μg), and the enzymatic rate was visualized by the decrease in absorbance at 340 nm ($\epsilon = 6,220 \text{ M}^{-1} \text{ cm}^{-1}$). Analogue **O1** was added after a linear decrease of absorption at 340 nm was observed. The inhibitive effect can then be evaluated by comparing the rate before and after the addition of **O1**.

2.7.5.2 Measurement of inhibition kinetics of analogue **T1**

Inhibition kinetics were measured using a slight modification of a previously reported continuous coupled assay for phosphate.^{80, 142} A cuvette containing Tris-HCl buffer (pH 7.0, 100 mM), MnCl₂ (1 mM), PEP (variable, 50 μM to 1 mM), His₆-tagged NeuB (5 μg), purine nucleoside phosphorylase (PNP, 5 units, previously buffer exchanged to 20 mM Tris-HCl, pH 7.0), 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG, 200 μM) and inhibitor **T1** (variable, 0 μM to 20 μM of the 4.5:1 mixture of (2*S*)-**T1**:(2*R*)-**T1**) was pre-incubated for 20 min at 37 °C. The enzymatic reaction was initiated by addition of ManNAc (10 mM). Rates were measured by monitoring the increase of absorption at 360 nm ($\epsilon = 11,000 \text{ M}^{-1} \text{ cm}^{-1}$). Kinetic parameters were determined by fitting initial velocities to the Michaelis-Menten equation using GraFit 5.0. The reported error range was calculated from the non-linear least square fitting on results of kinetic trials carried out on the same day, with enzymes, substrates and/or inhibitors from the same batch. However, systematic errors can occur due to factors such as use of different batches of enzymes and inhibitors, errors in concentration measurements, as well as temperature fluctuation. These typically lead to an additional 15% error when measurements were performed on different days.

Metal dependency of the inhibition by analogue **T1** was evaluated by using different concentrations of MnCl₂ (0.2, 0.5, 1, 5 and 10 mM), 100 mM of PEP and without or with analogue **T1** (5 μM) in the above described continuous coupled assay.

2.7.6 Crystallization, Data Collection, and Structure Refinement.

Purified NeuB enzyme was crystallized at 18 °C in the presence of 10 mM MnCl₂ and 1.50-1.55 M malic acid (pH 6.2) using the hanging-drop vapor diffusion technique, and subsequently soaked in 2 M sodium phosphate (pH 6.2) for 24 h with 10 mM MnCl₂ and 3 mM of the 4.5:1 mixture of (2*S*)-**T1**:(2*R*)-**T1**. The soaked crystal was transferred into the mother liquor containing 25% ethylene glycol for 10 s and frozen in liquid nitrogen prior to data collection. X-ray diffraction data were collected at 100 K under a nitrogen stream at the beamline 4.2.2 of the Advanced Light Source (Berkeley, CA) coupled to a NOIR-1 CCD detector. Collected data were processed by MOSFLM¹⁴⁸ and SCALA¹⁴⁹. The inhibitor **T1**-bound NeuB crystal belongs to the space group P2₁2₁2 with unit cell dimensions a=58.62, b=75.74, c=77.36 Å, and contains one molecule in the asymmetric unit. Scaled data were directly used for the structure refinement by REFMAC5¹⁵⁰ with the coordinate of NeuB as a starting model (PDB accession code: 1XUZ). The required parameter file for inhibitor **T1** was generated from the Dundee PRODRG2 server¹⁵¹ and all structural figures were produced using PyMOL. Statistics for data collection and refinement are summarized in Table 2.1.

Table 2.1 Data Collection and Refinement Statistics.

NeuB+Inhibitor T1	
Data Collection	
X-ray source	ALS 4.2.2
wavelength (Å)	0.9790
resolution (Å) ^a	40–1.75 (1.84–1.75)
space group	P2 ₁ 2 ₁ 2
a (Å)	58.62
b (Å)	75.74
c (Å)	77.36
α, β, γ (°)	90, 90, 90
R _{sym} (%) ^a	10.5 (46.1)
I/σ(I) ^a	12.4 (2.2)
completeness (%) ^a	99.3 (97.0)
unique reflections ^a	35,158 (4926)
redundancy ^a	4.0 (3.6)
Refinement Statistics	
average B-factors (Å ²)	
protein	9.0
ligand	13.8
water	18.4
ramachandran statistics	
favored regions (%)	98.9
allowed regions (%)	1.1
R _{work} (%)	16.3
R _{free} (%)	19.7
r.m.s. ^b bonds (Å)	0.011
r.m.s. angles (°)	1.202

^a Values in parentheses represent the highest resolution shell

^b r.m.s., root mean square

Chapter 3 : PseG of Pseudaminic Acid Biosynthesis: A UDP-Sugar Hydrolase as a Masked Glycosyltransferase^c

^c A version of this chapter has been published, some figures are reproduced with permission from:
Feng Liu, Martin E. Tanner, (2006) "PseG of Pseudaminic Acid Biosynthesis: A UDP-sugar hydrolase as a masked glycosyltransferase", *J. Biol. Chem.* 281: 20902-20909. © The American Society for Biochemistry and Molecular Biology.

3.1 Introduction

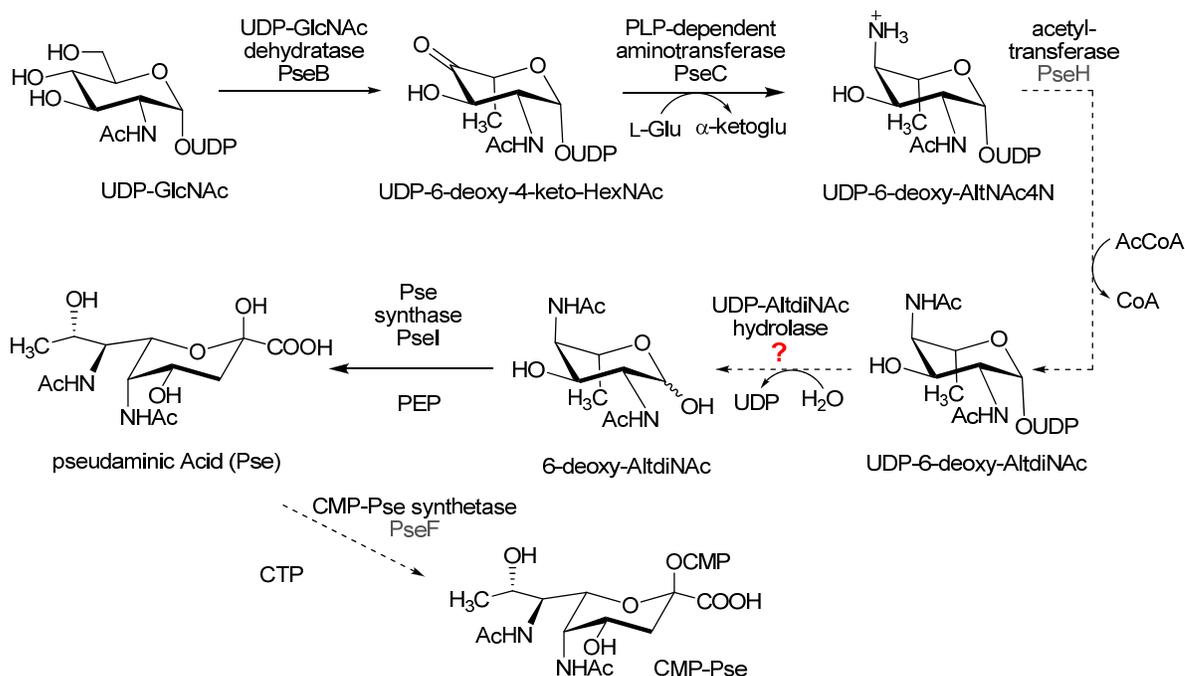


Figure 3.1 Biosynthesis of CMP-pseudaminic acid.

As outlined in Section 1.3.3, the flagella of *C. jejuni* and *H. pylori* are found to be extensively glycosylated with pseudaminic acid (Pse), and this modification is absolutely required for the assembly and function of flagella. In the biosynthesis of pseudaminic acid, a hydrolase is thought to cleave the UDP moiety from UDP-6-deoxy-AltdiNAc and generate 6-deoxy-AltdiNAc, which serves as a substrate for pseudaminic acid synthase PseI (Figure 3.1). In the corresponding biosynthesis of sialic acid, the enzyme NeuC plays a similar role by cleaving the UDP moiety from UDP-GlcNAc, but also catalyzes a C-2 epimerization in generating the substrate for sialic acid synthase (NeuB). It is possible that a NeuC homologue in *C. jejuni* performs the hydrolysis function, but not the epimerization function, in pseudaminic acid biosynthesis. The gene, *neuC2* or Cj1328, is found at the *C. jejuni* flagella glycosylation locus and shares 32.5% sequence identity with *Neisseria meningitidis neuC*. Given the similarity in the

overall strategy employed in Pse and sialic acid biosynthesis, it was reasonable to expect that the UDP-6-deoxy-AltdiNAc hydrolase may be an enzyme homologous to NeuC. The *C. jejuni* NeuC2 therefore became our first target in the identification of the hydrolase. Dr. Wakarchuk from NRC Canada provided us with a plasmid containing a recombinant *neuC2* (Cj1328 from *C. jejuni* NCTC11168) that expresses NeuC2 as a maltose-binding-protein (MBP) fusion protein.

In order to identify and characterize the UDP-6-deoxy-AltdiNAc hydrolase, the substrate, UDP-6-deoxy-AltdiNAc, first needed to be synthesized. Instead of pursuing a lengthy and low yielding chemical synthesis,^{80, 152} we established a chemo-enzymatic method to prepare the substrate, which after acidic hydrolysis gave 6-deoxy-AltdiNAc that was shown to be identical to previously chemically synthesized material. This synthesis is described in Section 3.2 of this chapter.

In Section 3.3.1 we outline our purification and studies with NeuC2. Unfortunately, we found that NeuC2 did not show any activity upon incubation with UDP-6-deoxy-AltdiNAc. This implies that NeuC2 is involved in a pathway that is distinct from Pse biosynthesis. A putative glycosyltransferase, PseG or Cj1312, then became our next target. A homologue of PseG in *H. pylori*, HP0326b, was known to be indispensable for Pse biosynthesis in *H. pylori*.^{60, 153} This implies that PseG also plays a similarly important role in *C. jejuni*. In Section 3.3.2, we describe the cloning of PseG gene from *Campylobacter jejuni* (strain NCTC 11168) genomic DNA and its expression in *E.coli* BL21 (DE3) cells. Ultimately, we show that PseG is responsible for the formation of 6-deoxy-AltdiNAc and we identify it as the UDP-6-deoxy-AltdiNAc hydrolase for the first time. In following mechanistic studies, PseG is shown to be a member of the GT-B superfamily of metal-independent glycosyltransferases and catalyzes the reaction via a C-O bond cleavage mechanism with inversion of stereochemistry at C-1.

3.2 Preparation of the Hydrolase Substrate UDP-6-deoxy-AltdiNAc

One of the main obstacles in identifying the enzymes involved in Pse biosynthesis is the difficulty in obtaining the substrates for the advanced steps of the pathway. In previous studies, 6-deoxy-AltdiNAc, the product of the hydrolase, was chemically prepared from L-fucose in a 12-step synthesis with an accumulative yield of ~0.5%.^{80, 152} We also tried to synthesize UDP-6-deoxy-AltdiNAc from L- threonine, however, the length of the synthesis and the number of low yielding steps made us realize that chemical approaches are undesirable for the preparation of this substrate. Since UDP-GlcNAc is commercially available, and the first two of the three enzymes involved in the conversion of UDP-GlcNAc to UDP-6-deoxy-AltdiNAc in *C. jejuni* were already identified and cloned, we decided that an enzymatic approach would be a superior route to obtain the substrate (Figure 3.3). As the third enzyme, a putative acetyltransferase which acetylates the 4-amino group of UDP-6-deoxy-AltNAc4N using acetyl-CoA (AcCoA), had not yet been identified at the time this research was conducted, a chemical acetylation was employed. In this fashion, UDP-6-deoxy-AltdiNAc was chemo-enzymatically prepared on a multimilligram scale.

The *C. jejuni* plasmids coding for the first two enzymes in the Pse biosynthetic pathway, the inverting dehydratase PseB and the aminotransferase PseC, were obtained from the laboratory of Dr. Wakarchuk. They were transformed into competent BL21 (DE3) *E. coli* cells and over-expressed. The His₆-tagged PseB and PseC were isolated using nickel affinity columns to >95% purity as indicated by the SDS-PAGE analysis shown below (Figure 3.2).

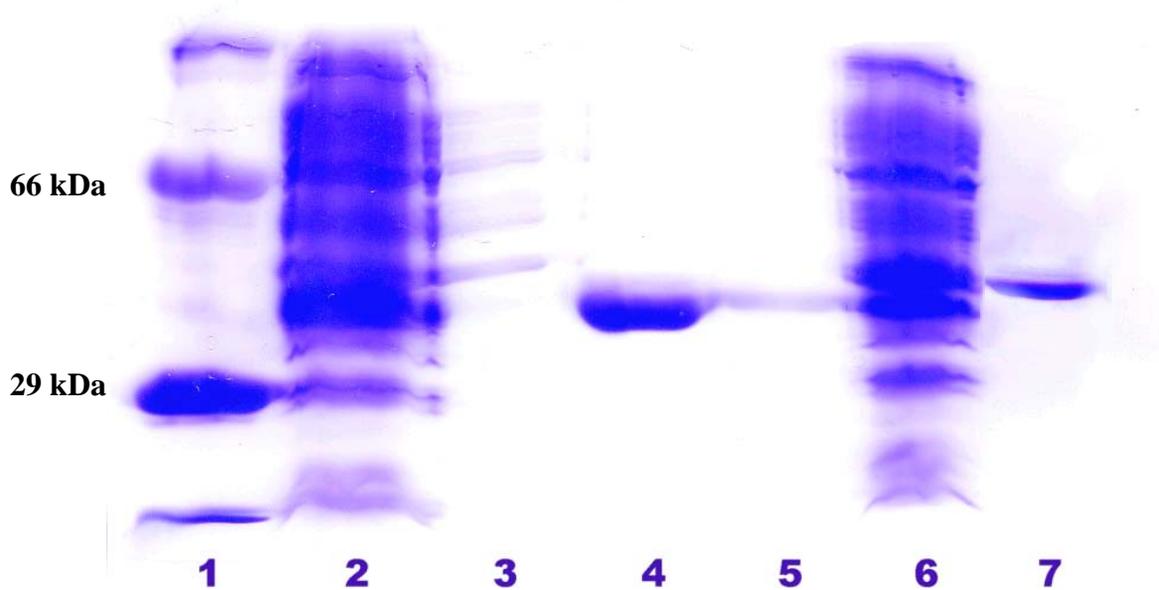


Figure 3.2 SDS-PAGE showing the purification of the PseB (Cj1293) and PseC (Cj1294) proteins. 1) Molecular weight (MW) standards, 2) crude cell lysate of PseB, 3) PseB column wash with 125 mM imidazole, 4&5) PseB column wash with 500 mM imidazole, 6) crude cell lysate of PseC, 7) PseC column wash with 500 mM imidazole.

The purified PseB and PseC were then incubated with UDP-GlcNAc in the presence of pyridoxal 5'-phosphate (PLP) and glutamate to give UDP-6-deoxy-AltNAc4N (Figure 3.3). This compound was purified by ion-exchange chromatography, and the C-4 amino group was then chemically acetylated using acetic anhydride to give the hydrolase substrate, UDP-6-deoxy-AltNAc. After purification, this material was characterized by ESI-MS(-) spectrometry and ^{31}P NMR/ ^1H NMR spectroscopy and was shown to bear the expected structure. In order to further confirm the structure of this compound, a sample of chemo-enzymatically synthesized UDP-6-deoxy-AltNAc was subjected to a mild hydrolysis under acidic conditions to cleave the UDP moiety from the glycoside (Figure 3.3). After removal of UDP by ion-exchange chromatography,

the ^1H NMR spectrum of the resulting free sugar was shown to be identical to that of chemically synthesized 6-deoxy-AltdiNAc.^{80, 152} This confirms that the stereochemical assignments previously made on the product of the PseB/C pair were correct.¹⁵⁴ Furthermore, upon incubation with PseI (Pse synthase) under previously reported conditions,⁸⁰ the hydrolysis product was converted to Pse as shown by ESI-MS and ^1H NMR spectroscopy.

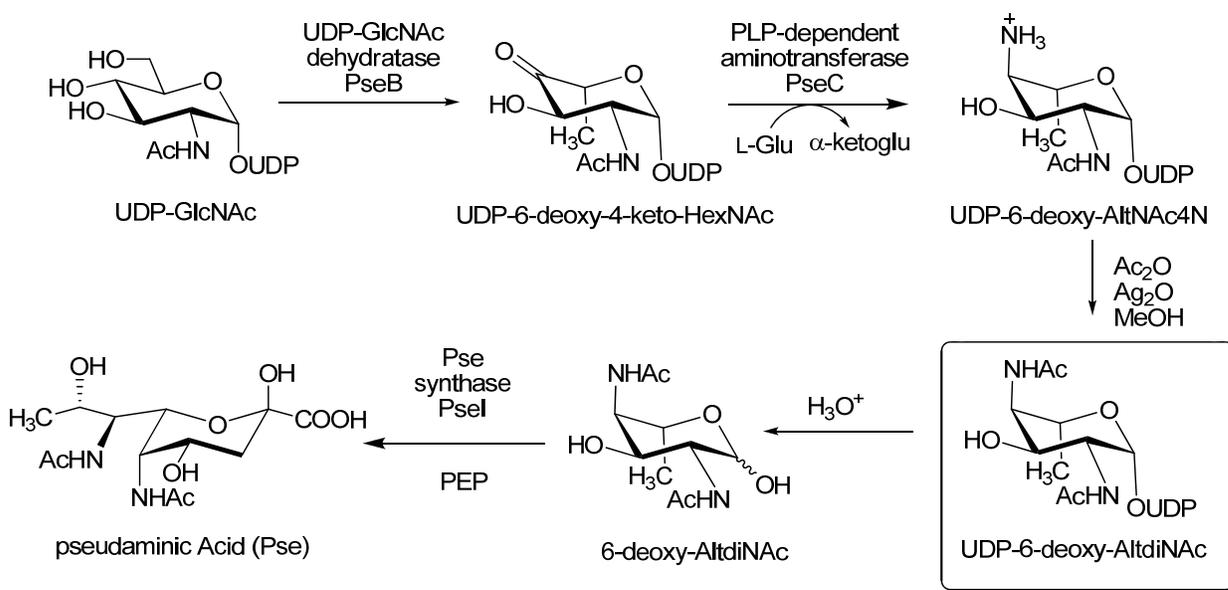


Figure 3.3 Chemo-enzymatic synthesis of UDP-6-deoxy-AltdiNAc and its acidic hydrolysis to give 6-deoxy-AltdiNAc, the substrate for PseI.

3.3 Identification of the UDP-6-deoxy-AltdiNAc Hydrolase from *C.*

jejuni

3.3.1 Initial Attempts with NeuC2 (Cj1328)

3.3.1.1 Targeting NeuC2 as the UDP-6-deoxy-AltdiNAc hydrolase

The search for enzymes involved in pseudaminic acid biosynthesis has been strongly influenced by studies on the biosynthesis of the related α -keto acid, sialic acid, due to the similarities between these two pathways. The overall strategy employed in pseudaminic acid biosynthesis, including the use of UDP-GlcNAc as the starting material, the condensation of a hexose with PEP, and the ultimate formation of a CMP-linked α -keto acid, was known to be shared with that of sialic acid biosynthesis. Since the key α -keto acid synthases are closely related (35% sequence identity),^{42, 47, 80} it is reasonable to suspect that the hydrolases may also be homologous (Figure 3.4). A *C. jejuni* gene, Cj1328 or NeuC2, was found to share 32% sequence identity with the *Neisseria meningitidis* NeuC and had been previously implicated in the biosynthesis of PseAm, an acetamidino-containing derivative of pseudaminic acid.⁶⁵ By using a similar mechanism to that of NeuC, this enzyme could catalyze the elimination of UDP from UDP-6-deoxy-AltdiNAc to give a glycal intermediate, and then hydrate the double bond of the glycal to give the net hydrolysis product 6-deoxy-AltdiNAc with overall retention of stereochemistry at C-2 (Figure 3.5C). The lack of C-2 epimerization would generate a precursor with the appropriate stereochemistry for Pse biosynthesis. This would be distinguished from the NeuC reaction in which the elimination/hydration process proceeds with an overall inversion of stereochemistry at C-2 (Figure 3.5B). Following the hydrolysis step, the resulting 6-deoxy-AltdiNAc would then be directly converted to Pse by the action of PseI (Figure 3.4). A CMP-Pse

synthetase (PseF) then activates Pse as CMP-Pse, which could be further converted to CMP-PseAm by a putative amidotransferase PseA. PseA was identified after this project was finished and is currently under study in our laboratory.

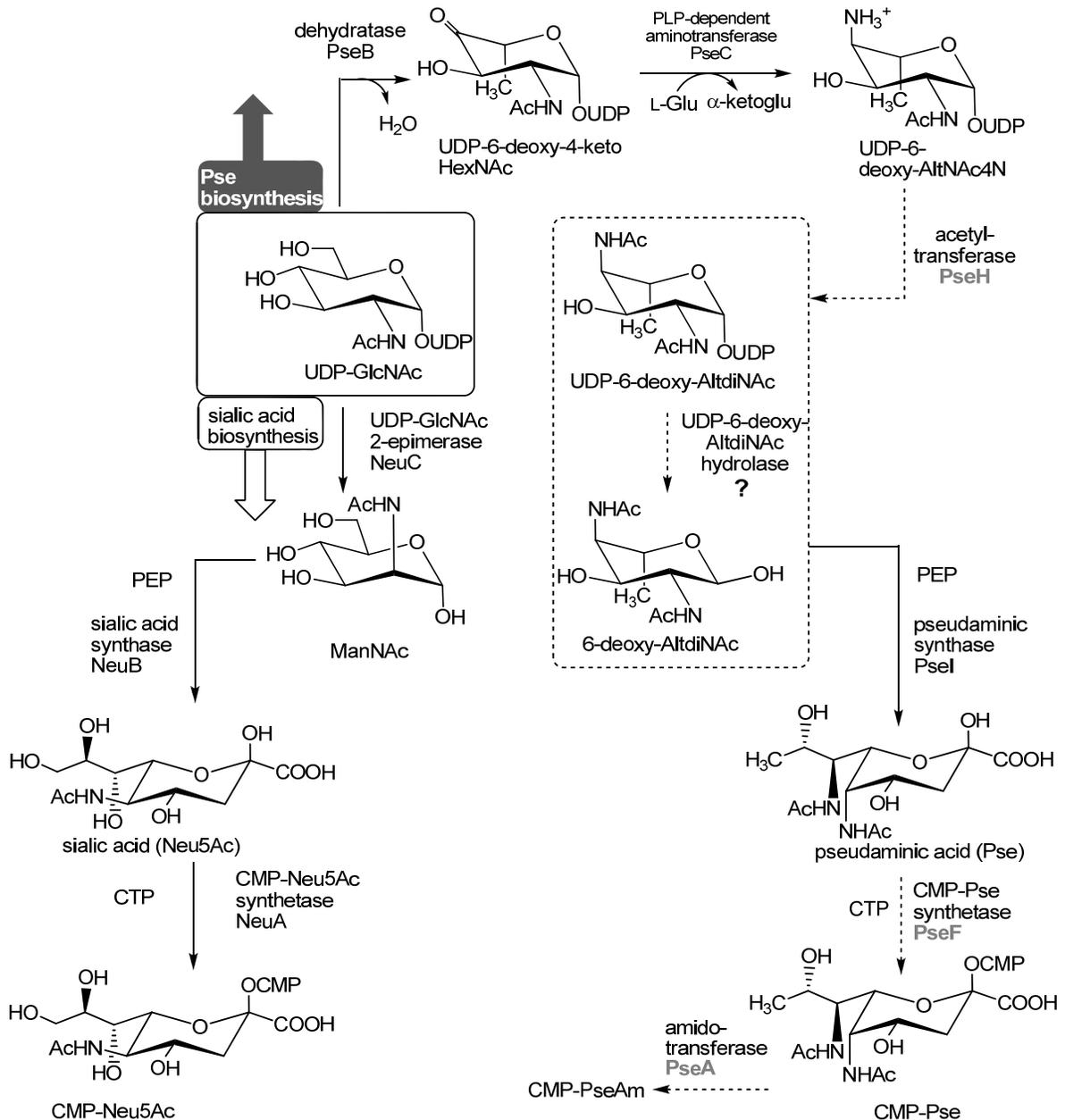


Figure 3.4 Comparison of the biosynthetic pathway of pseudaminic acid and sialic acid in bacteria.

Solid box shows the common starting material, UDP-GlcNAc. Dotted box shows the proposed hydrolase reaction in Pse biosynthesis that serves a similar function to NeuC of sialic acid biosynthesis.

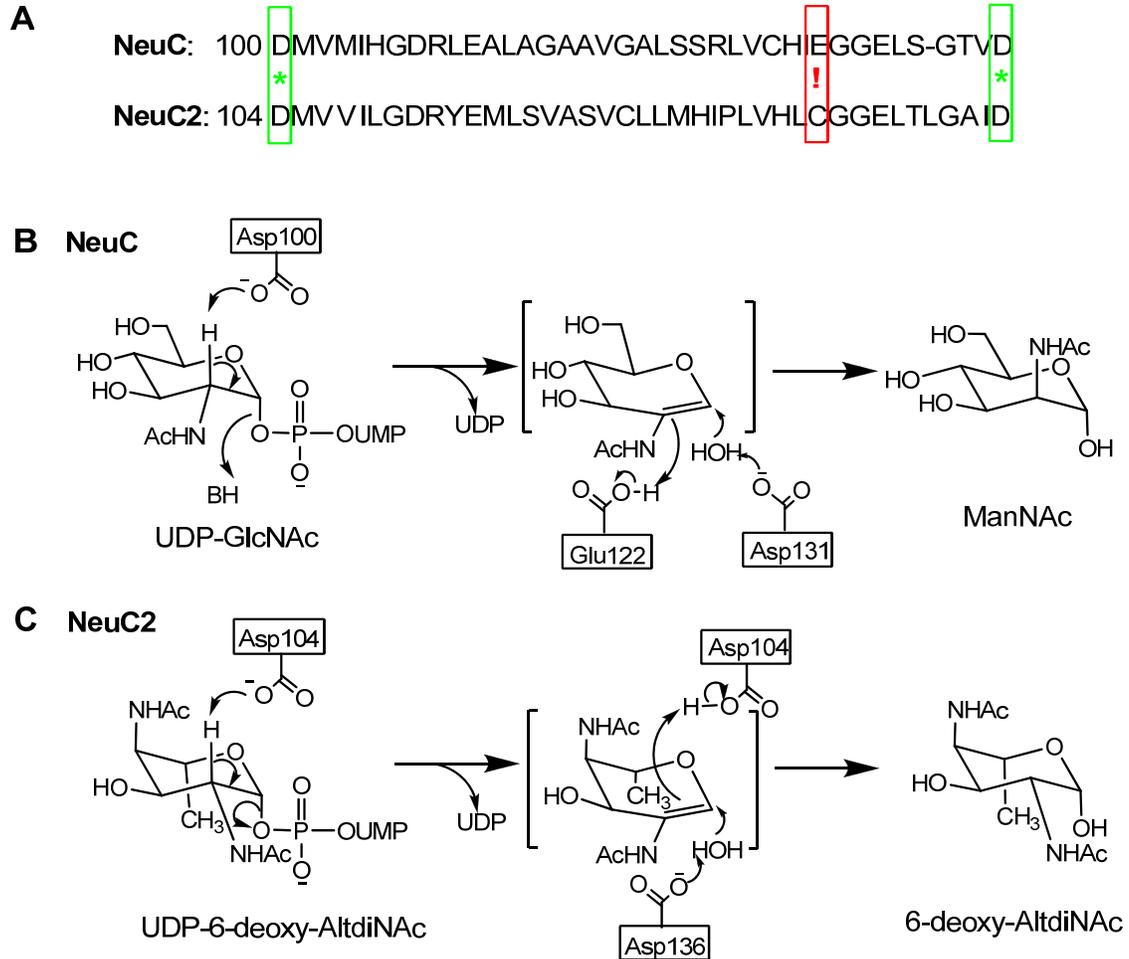


Figure 3.5 A sequence comparison of NeuC and NeuC2 and potential mechanisms. A) A sequence alignment showing key active site regions of NeuC and NeuC2. **B)** Putative roles of active site residues in NeuC **C)** Putative roles of active site residues in NeuC2.

If NeuC2 serves as the UDP-6-deoxy-AltdiNAc hydrolase, it would likely share a common ancestor with UDP GlcNAc 2-epimerase (NeuC), and the ability to epimerize would have been lost during evolution. To propose a reasonable explanation for the loss of epimerizing activity, differences in potential key catalytic residues were analyzed. The crystal structure of a homologous non-hydrolyzing UDP GlcNAc 2-epimerase from *E. coli* implicated three carboxylate residues in the active site of NeuC that could be involved in the transfer of a proton or in the stabilization of an oxocarbenium transition state (Figure 3.5B).³⁶ The D100N, E122Q and D131N mutants of NeuC showed significantly reduced activity, and the release of the 2-

acetamidoglucal intermediate was observed when incubating UDP GlcNAc with E122Q and D131N, but not with the wild type enzyme.³⁶ This strongly supported the mechanism involving a glycal intermediate. Glu122 and Asp131 may be involved in the *syn* hydration of the glycal intermediate since these mutations lead to the release of the glycal intermediate. It is interesting to note that when comparing the sequence of NeuC and NeuC2, Asp100 and Asp131 of NeuC align with Asp104 and Asp136 of NeuC2, while Glu122 of NeuC aligns with Cys126 of NeuC2 (Figure 3.5A). The latter mismatch can be explained by the notion that NeuC2 is only a hydrolase and does not require a third carboxylic acid to bring about an inversion of stereochemistry.

3.3.1.2 Over-expression of NeuC2 and Activity Test

The plasmid donated by Dr. Wakarchuk had *neuC2* (*Cj1328*) inserted into a pMALTM-c4X vector and encoded for NeuC2 as a fusion with the maltose binding protein (MBP). For the ease of purification, *neuC2* was cloned from this plasmid and re-inserted into a pET 30 Xa/LIC vector (Novagen) which encoded for NeuC2 as a His₆-tagged enzyme. The new plasmid was transformed into CaCl₂ competent *E. coli* BL21 (DE3) cells and expression was induced. Unfortunately, the vast majority of the expressed His₆-NeuC2 protein was insoluble and was found in the solid pellet after cytolysis using a French Pressure cell (Figure 3.6 lanes 2 and 3).

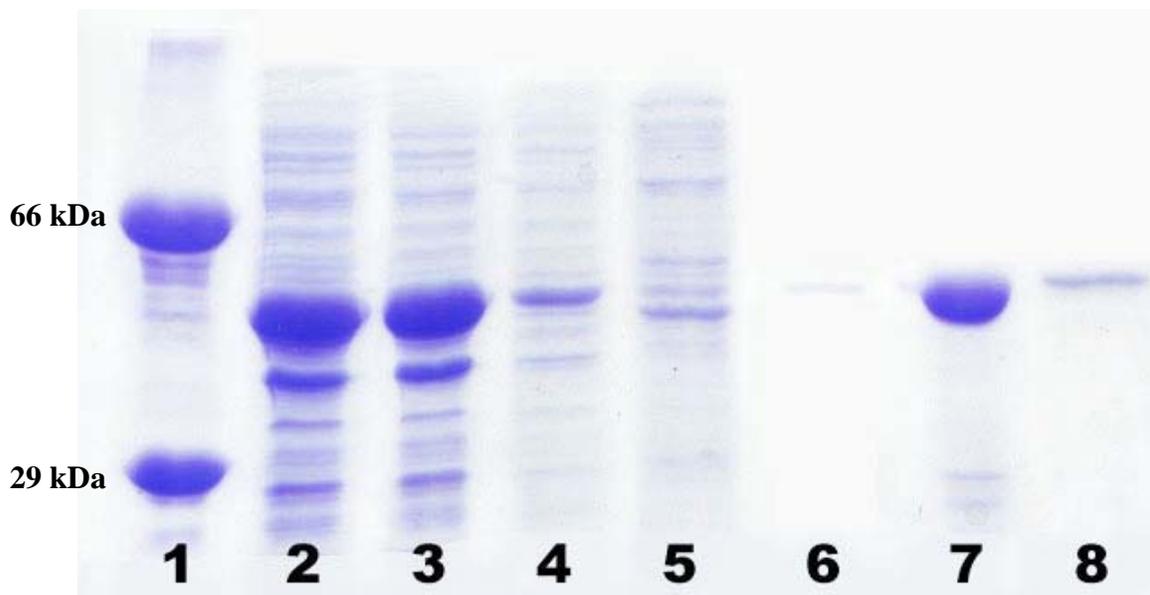


Figure 3.6 SDS-PAGE analysis showing the attempted purification of His₆-NeuC2. 1) Molecular weight (MW) standards, 2) after over-expression, 3) solid pellet after removal of crude lysate, 4) column wash with 5 mM imidazole, 5) column wash with 125 mM imidazole, 6) column wash with 500 mM imidazole, 7) precipitates after 2 h incubation with substrate at 37 °C, 8) enzyme remained in solution after 2 h incubation with substrate at 37 °C.

As the insolubility of the His₆-NeuC2 protein probably results from the formation of inclusion bodies, various methods, including over-expression at lower temperatures, induction with lower IPTG concentration, and the addition of sucrose and ethanol were explored to solve this problem. Nevertheless none of these attempts brought significant improvement in the production of more soluble protein. A small amount of soluble protein was found to be retained in a nickel affinity column and was purified using the methods described in Chapter 2 (Figure 3.6, lanes 4, 5 and 6). When the resultant purified His₆-NeuC2 was incubated with chemoenzymatically synthesized substrate UDP-6-deoxy-AltdiNAc, no activity was observed. The majority of the enzyme was observed to precipitate during the reaction (Figure 3.6, lanes 7 and 8). This implies that the His₆-NeuC2 has folding and/or stability problems. We then turned our attention back to over-expressing NeuC2 as a fusion with the maltose binding protein (MBP-NeuC2).

The maltose binding protein (MBP) has been found to serve as a superior solubility enhancer for peptides and proteins to which it is fused¹⁵⁵⁻¹⁵⁷ and is also widely used in the purification of recombinant proteins as an affinity tag. In a typical purification using the MBP system, the target protein is first expressed as a fusion construct where it is connected to MBP by an oligopeptide sequence with a protease cleavage site situated immediately next to its N-terminus, so that no vector-derived amino acids are attached to the target protein after cleavage (Figure 3.7). Crude cell lysate containing the MBP-target protein is loaded onto a column containing amylose resin. After washing with buffer, only the MBP-target protein is retained on the column. A subsequent elution with maltose-containing buffer will give the MBP-target fusion protein, which can be subjected to cleavage by a protease. The protease (Factor Xa in this case) recognizes its specific site (immediately after Ile-Glu-Gly-Arg in the case of Factor Xa)

and cleaves the target protein from MBP. Pure target protein can be obtained by first removing maltose from the buffer using membrane ultracentrifugation and then removing MBP using a second amylose column.

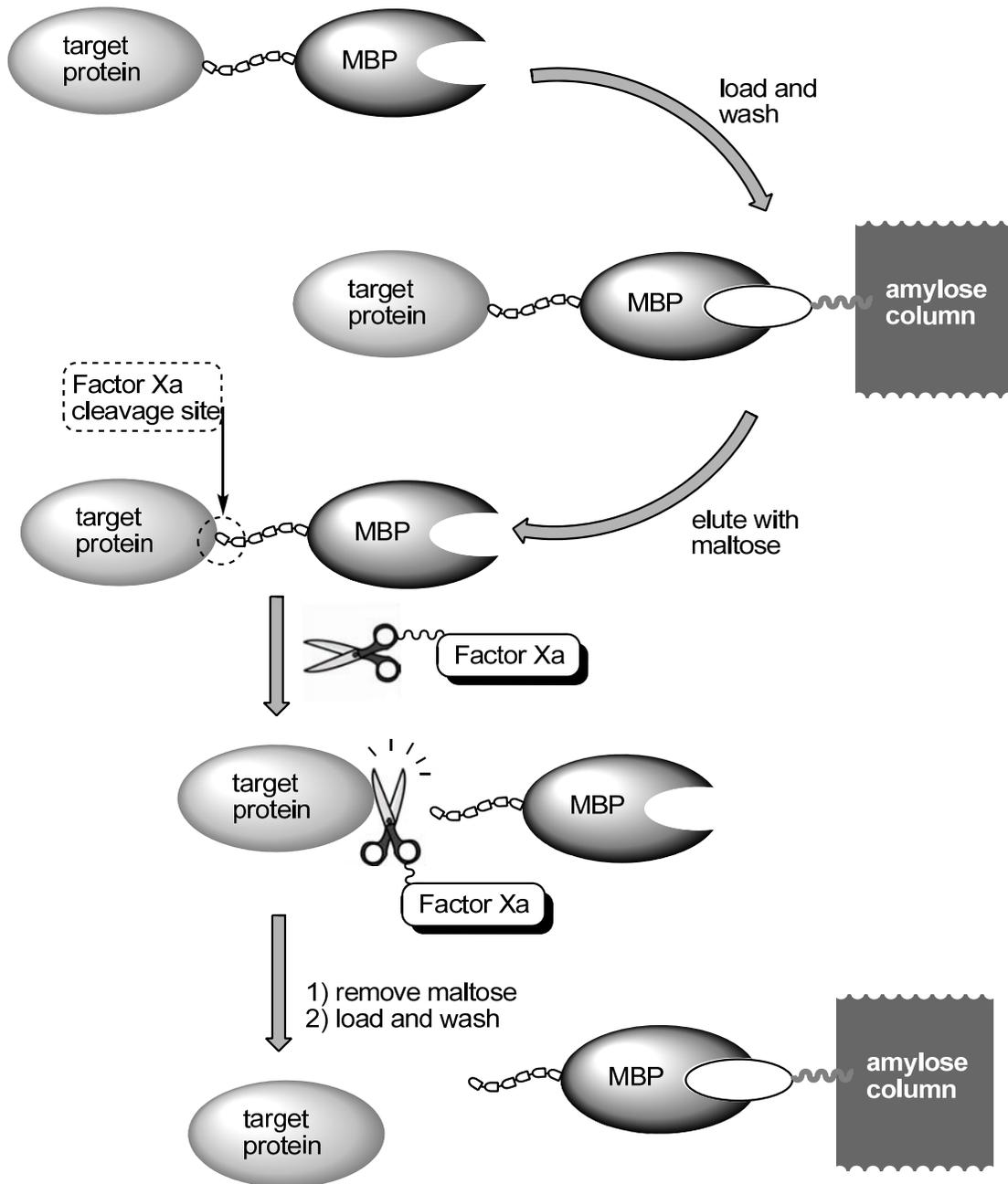


Figure 3.7 Purification of recombinant protein using MBP as the affinity tag.

The MBP-NeuC2 construct gave large quantities of soluble protein and was purified on an amylose column as described above (Figure 3.8 A). Both MBP-NeuC2 and wild type NeuC2 obtained after cleavage from MBP by Factor Xa (Figure 3.8 B) were subjected to activity test by incubation with UDP-6-deoxy-AltdiNac. To our disappointment, no hydrolase activity could be observed with MBP-NeuC2 or NeuC2.

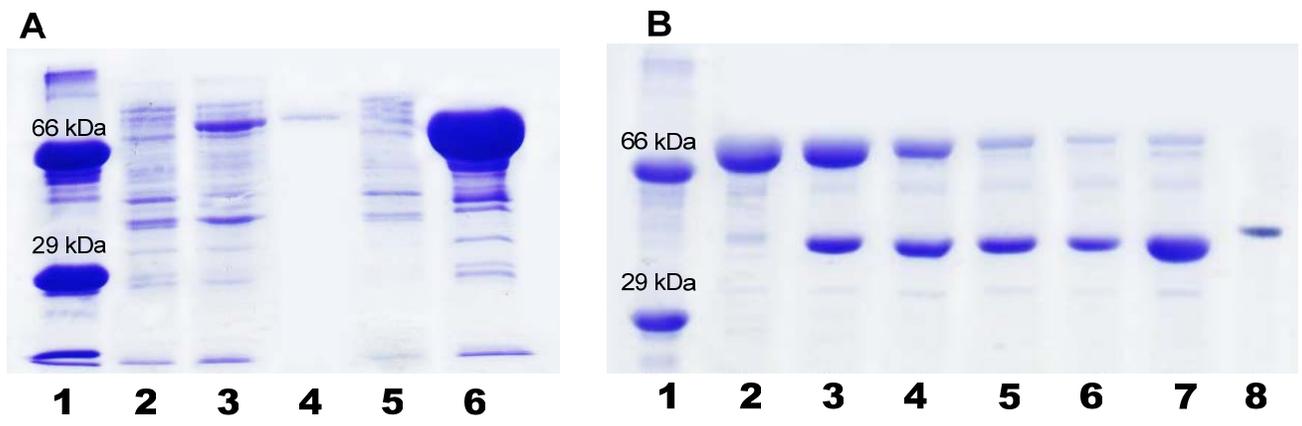


Figure 3.8 SDS-PAGE analysis showing the purification of MBP-NeuC2. A) Over-expression and purification on amylose column. 1) MW standards, 2) *E. coli* cells before induction, 3) after induction with IPTG, 4) crude lysate, 5) column wash with Tris-HCl buffer, 6) protein eluted with 10 mM maltose. B) cleavage of MBP from target protein. 1) MW standard, 2) MBP-NeuC2. 3) 2h, 4) 4h, 5) 6h, 6) 8h, 7) 18h after addition of Factor Xa, 8) MBP.

As mentioned at the beginning of this chapter, NeuC2 (Cj1328) was first proposed to be involved in the biosynthesis of PseAm, an acetamidino-derivative of pseudaminic acid based on a mass spectral analysis of metabolites of *C. coli* VT167.⁶⁵ However, given that UDP-6-deoxy-AltdiNac does not serve as a substrate for this enzyme, it appears that NeuC2 is responsible for catalyzing a different reaction (most likely one that involves both C-2 epimerization and hydrolysis) in a pathway that is distinct from Pse biosynthesis. This notion was supported by the interesting observation that a flagella-modifying sugar produced in *C. jejuni* 81-176, with *m/z* corresponding to that of PseAm, was found to be structurally different than that produced in *C.*

coli VT167.⁶⁵ Additionally, NeuC2 (Cj1328) was only found in the flagella glycosylation locus of *C. coli* VT167 but not in *C. jejuni* 81-176. Indeed, a paper emerged in 2007, indicating that the identity of the PseAm-like sugar produced in *C. coli* VT167 is Leg5Am7Ac, a 5-acetamidino derivative of legionaminic acid (Figure 3.9).¹² Leg5Am7Ac (produced in *C. coli* VT167 and *C. jejuni* NCTC 11168, but not in *C. jejuni* 81-176) and PseAm (produced in *C. jejuni* 81-176 and *C. jejuni* NCTC 11168, but not in *C. coli* VT167) share similar structural characteristics and only differ in the stereochemistry at C-8, C-7 and C-5. They have identical molecular weights and could not be discriminated by the early mass spectra-based studies. It can therefore be deduced that Cj1328 is likely involved in the biosynthesis of Leg5Am7Ac in *C. coli* VT167 and *C. jejuni* NCTC 11168, and acts as a hydrolyzing 2-epimerase. Research on the biosynthesis of CMP-legionaminic acid in *C. jejuni* NCTC 11168 was published in 2009 and confirmed this notion.¹⁵⁸ The authors also demonstrated that, as opposed to UDP-Bac2Ac4Ac, GDP-Bac2Ac4Ac is a better substrate for Cj1328 or LegG. It was suggested that *Campylobacter* utilizes a distinct pathway involving GDP-linked intermediates for legionaminic acid biosynthesis (Figure 3.9).

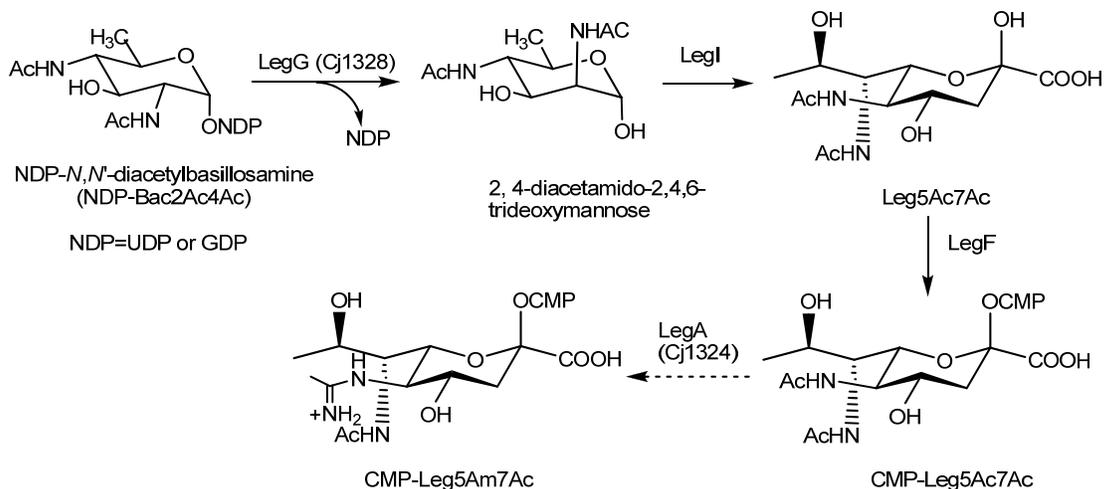


Figure 3.9 CMP-Leg5Ac7Ac biosynthesis and the proposed transformation from CMP-Leg5Ac7Ac to CMP-Leg5Am7Ac.

3.3.2 Identification of PseG (Cj1312) as the UDP-6-deoxy-AltdiNAc Hydrolase

3.3.2.1 Targeting PseG (Cj1312) as the UDP-6-deoxy-AltdiNAc hydrolase

Since NeuC2 (Cj1328 or LegG) was found not to serve as a UDP-6-deoxy-AltdiNAc hydrolase, we then started seeking a second target. As was mentioned in Section 1.3.2, *H. pylori* lacks heterogeneity in the composition of flagellar glycans and only produces Pse. Comparative studies on homologous enzymes in *H. pylori* would provide very useful clues to the identification of enzymes involved in Pse biosynthesis in *C. jejuni*. A report on proteins required for flagellar glycosylation in *H. pylori* attracted our attention.¹⁵³ Two overlapping genes, *neuA* (HP0326a) and *flmD* (HP0326b), encoded for enzymes related to protein-glycosylation and were controlled by a flagellar class 2 σ^{54} promoter. Mutations in these genes severely compromised the flagellar glycosylation process, and resultant mutants lacked flagellar hook and filament and were found to be non-motile.^{60, 153} The *neuA* (HP0326a) gene is homologous to the *neuA* of sialic acid biosynthesis and likely encodes for a CMP-pseudaminic acid synthetase. Indeed, in a paper later published in 2006, the gene product of *neuA* (HP0326a) in *H. pylori* was confirmed to be a CMP-pseudaminic acid synthetase (assigned as PseF).⁷⁶ The *flmD* gene product was annotated as a putative glycosyltransferase, which was proposed to be involved in the transfer of a carbohydrate onto the flagellin protein, basing on a bioinformatic analysis. Homologues of these two genes are also found to be similarly colocalized in several other strains of flagellated bacteria, including *C. jejuni* (Cj1311/Cj1312), although in this case they do not overlap. The gene product of Cj1311 was annotated as a putative CMP-pseudaminic acid synthetase and shared 40% sequence identity with HP0326a, whereas the gene product of Cj1312 was assigned as a putative glycosyltransferase and shared 35% sequence identity with HP0326b. Furthermore, mass spectral analyses revealed that mutation of HP0326b lead to the accumulation of mono- and

diacetamido trideoxyhexose UDP sugar nucleotides in the cytosolic extract. These metabolites were identified to be precursors of Pse. This implies that the gene product of HP0326b, as well as that of Cj1312, are very likely directly involved in the biosynthesis of pseudaminic acid.

The putative glycosyltransferase encoded by Cj1312 (or PseG) could represent a likely candidate for a sugar nucleotide hydrolase involved in Pse biosynthesis since transfer of the donor glycoside to water as an acceptor would constitute a hydrolysis reaction. Although the amino acid sequence of PseG does not show significant homology (> 25% identity) with any proteins of proven function, further homology searching using the Position Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST) on Cj1312 or PseG, indicated that it most closely resembled *N*-acetylglucosaminyltransferases in the metal-independent GT-B superfamily that includes the peptidoglycan biosynthetic enzyme MurG (Figure 3.10).¹⁵⁹⁻¹⁶¹

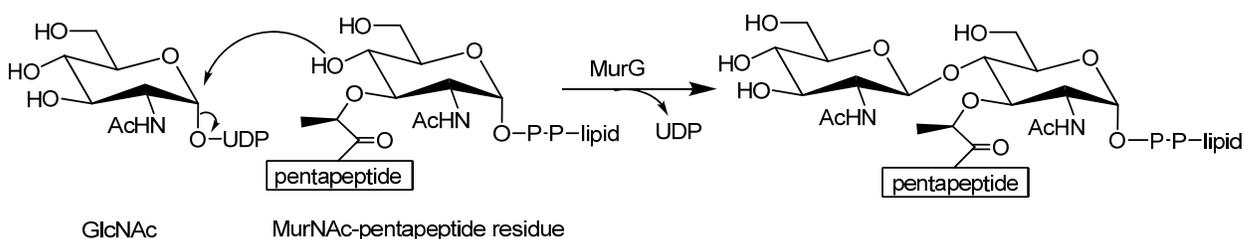


Figure 3.10 Glycosyl-transfer reaction catalyzed by MurG.

MurG is a member of the GT-B superfamily of enzymes (family 28) that are metal-independent inverting glycosyltransferases.^{159, 161} It catalyzes the transfer of GlcNAc from its UDP nucleoside onto peptide-attached *N*-acetylmuramic acid (MurNAc) to form the β -(1,4)-linked GlcNAc-MurNAc backbone in peptidoglycan (Figure 3.10). Previous studies on the

structure of the *E. coli* MurG•UDP-GlcNAc complex have revealed key conserved residues in the C-terminal half of the enzyme that are important for sugar nucleotide binding.¹⁶⁰ Although overall homologies between PseG and MurG are low (<15% identity), all of these key residues are present in the PseG sequence and they are well aligned with those in the MurG sequences (Figure 3.11). The GGS loop of the GT-B family members (residues 190-192 in *E. coli* MurG) provides the serine that interacts directly with the β -phosphate of the sugar nucleotide and is conserved as GGT in PseG (residues 165-167). The conserved Glu269 and Gln289 in *E. coli* MurG that are hydrogen-bonded to the ribose hydroxyls also present as Glu239 and Gln256 in *C. jejuni* PseG.

PseG	2	KVLFPSDSSSQIGFGHIKRDVLVAKQYSDVSFACLPL-----EGSLIDEIPYPVYELSS	55
		+++ + + GH+ L +A + L E L+ + + +	
MurG	7	RLMVMAGGTG----GHVFPGLAVAHHLMAQGWQVRWLG TADRMEADLVPKHGIEIDFIRI	62
PseG	56	ESIY-----ELINL I KEKPELLI I-DHYGISVDDEKLIKLETGVK	95
		+ + ++K K ++++ Y G+	
MurG	63	SGLRGKGIKALIAAPLRIFNAWRQARAIMKAYKPDVVLGMGGYVSGPGGLAAWS--LGIP	120
PseG	96	ILSFDDEIKPHHCDILLNVNAYAKASDYEGLVVPFKCEVRCGFSYALIREEFYQE-----A	150
		++ + + L A + G P V +R +	
MurG	121	VVLHEQNGIAGLTNKWLAKIATKVMQAFPGAFPNAEVV-----GNPVRTDVLALPLPQQR	175
PseG	151	KENREKKYDFFICMGGTD IKNLSLQIASELPKTKI--ISIATSSSNPMLKKLQKFAKL--	206
		RE + GG+ + Q ++ ++I S + + +++	
MurG	176	LAGREGPVRVVLVV-GGSQGARILNQTMPQVAAKLGDSVTIWHQSGKGSQQSVEQAYAEAG	234
PseG	207	HNNIRLFDIHENIAKLMNESNKL IISASS-LVMEALLLKANFKAICY--VKNQESTATW	262
		++ +++A ++ ++ + + V+E + + + Q A	
MurG	235	QPQHKVTEFIDDMAAAYAWADVVCVCRSGALTVSEIAAAGLPALFVFPQHKDRQQYWNALP	294
PseG	263	LAKKGYE	269
		L K G	
MurG	295	LEKAGAA	301

Figure 3.11 Alignment of *C. jejuni* PseG and *E. coli* MurG. The alignment was obtained from an iterative PSI-BLAST homology search using PseG as the query sequence (four iterations).¹⁶² Boxed residues are consistently conserved among MurG family members and have been implicated as important in the binding of UDP-sugars.¹⁶⁰

3.3.2.2 Over-expression of PseG and Activity Test

The gene *pseG* (Cj1312) was therefore cloned from the genomic DNA of *C. jejuni* NCTC 11168 and inserted into the pET 30 Xa/LIC vector. The resulting plasmid was transformed into *E. coli* BL21 (DE3) cells and PseG was over-expressed with an N-terminal His₆-tag. The soluble protein was purified by nickel affinity chromatography and was found to be >95% pure as analyzed by SDS-PAGE (Figure 3.12).



Figure 3.12 SDS-PAGE of purified PseG.

PseG was incubated with UDP-6-deoxy-AltdiNAc and the reaction was monitored by ESI-MS spectrometry and ³¹P NMR spectroscopy (Figure 3.13). It was shown that the disappearance of UDP-6-deoxy-AltdiNAc (ESI-MS (-) *m/z* 631(M-H⁺), ³¹P NMR δ -10.3, -12.4) was accompanied by the formation of free UDP (ESI-MS (-) *m/z* 403 (M-H⁺), ³¹P NMR δ -5.7, -9.8), and a free sugar whose mass was consistent with that of 6-deoxy-AltdiNAc (ESI-MS (+) *m/z* 269 (M+Na⁺)). The free sugar was isolated from all charged species by passage through ion-exchange columns and subsequent removal of solvent *in vacuo*. Its structure was confirmed by its ¹H NMR spectrum (Figure A.6), which was shown to be identical to that of a synthetic

standard of 6-deoxy-AltdiNAc. PseG was therefore confirmed to act as a UDP-6-deoxy-AltdiNAc hydrolase.

UDP-GlcNAc and UDP-6-deoxy-AltNAc4N, the immediate precursor of UDP-6-deoxy-AltdiNAc, were also briefly tested as potential alternative substrates. It was shown that PseG has no detectable activity with UDP-GlcNAc and extremely low level of activity (<1%) with UDP-6-deoxy-AltNAc4N, suggesting that PseG exhibits reasonably high substrate specificity.

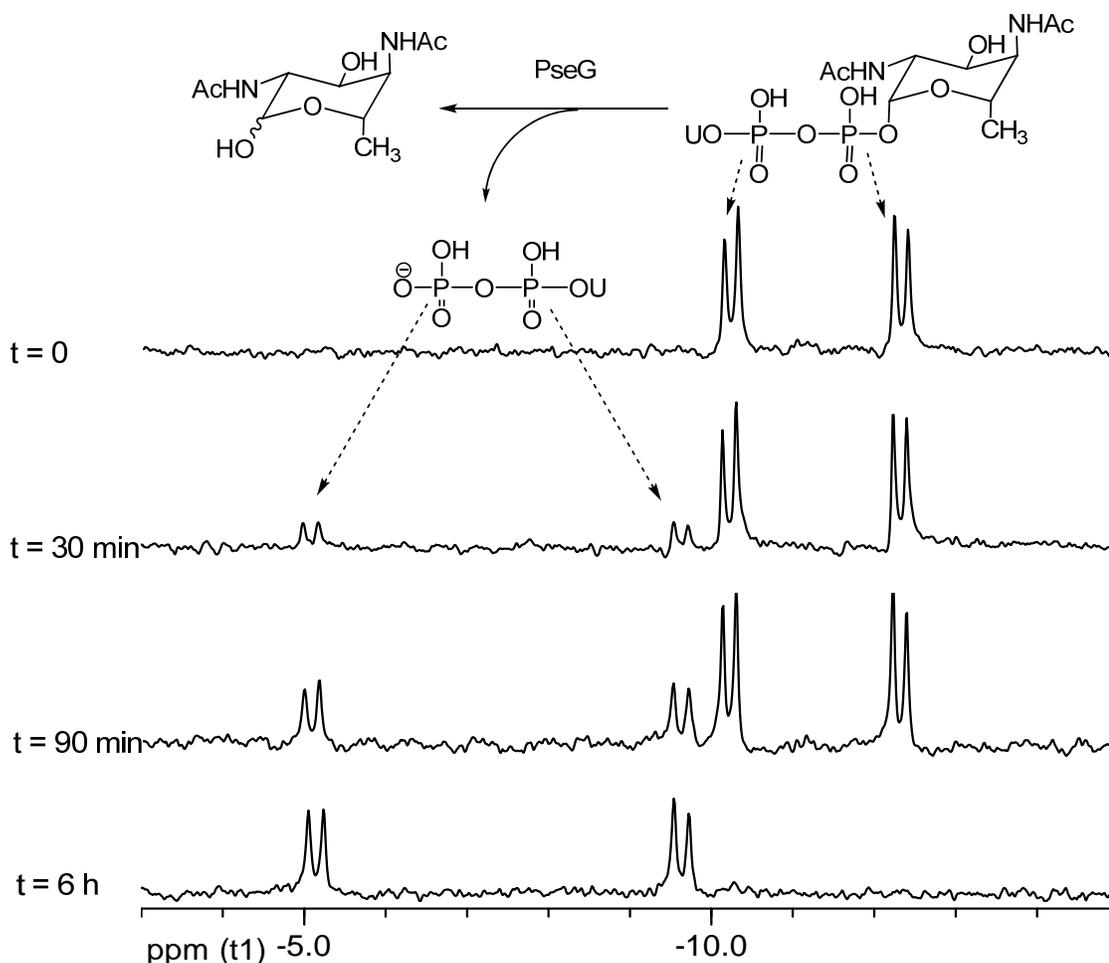


Figure 3.13 ³¹P NMR spectra monitoring the time course of the conversion of UDP-6-deoxy-AltdiNAc into UDP.

Although PseG was shown to be a UDP-6-deoxy-AltdiNAc-specific hydrolase, the fact that PseG is related to a family of glycosyltransferases raises a question: Is the hydrolase function the true role of PseG or is the enzyme indeed a transferase that exhibits hydrolase activity as an alternate reaction in the absence of an appropriate acceptor molecule? Several lines of evidence support the former notion. Firstly, The *pseG* gene (Cj1312) is located in the flagellar glycosylation gene cluster together with all other known and putative enzymes involved in CMP-pseudaminic acid biosynthesis, including Cj1317 encoding for Pse synthase (PseI) and Cj1311 encoding for a putative CMP-Pse synthetase (PseF).^{2, 80} Secondly, its homologue in *H. pylori*, FlmD or HP0326b, has been shown to be required for flagellin glycosylation.¹⁵³ These findings clearly demonstrate the involvement of PseG in the flagellar glycosylation process. Thirdly and most importantly, upon the completion of this project a report appeared showing that mutations of the PseG gene resulted in the loss of motility in the *C. jejuni* strain 81-176.⁶⁸ As we know, the only glycosyltransferase that would be required for flagellin modification with pseudaminic acid would be CMP-pseudaminyltransferase. The gene *pseF* (Cj1311) has been found to be homologous to that encoding for CMP-sialyltransferases and was already proposed to encode for the CMP-Pse synthetase (PseF) in pseudaminic biosynthesis. Moreover, since the structurally characterized CMP-sialyltransferase from *C. jejuni* was found to display a GT-A-like fold, it is highly unlikely that PseG, which is closely related to the GT-B family, would catalyze an analogous reaction.^{163, 164} Finally, other than its involvement in Pse biosynthesis, 6-deoxy-AltdiNAc has no known roles in *C. jejuni* biology. It is therefore very unlikely that a glycosyltransferase would exist that accepts the UDP-linked form of 6-deoxy-AltdiNAc as a substrate. To conclude, all evidence suggests that the true role of PseG is to act as a UDP-6-deoxy-AltdiNAc hydrolase in pseudaminic acid biosynthesis.

3.4 Kinetic Analysis of the UDP-6-deoxy-AltdiNAc Hydrolase Reaction

The kinetics of the hydrolysis reaction were monitored by a continuous coupled assay monitoring the release of UDP (Figure 3.14).³⁵ As PseG hydrolyzes UDP-6-deoxy-AltdiNAc, the released UDP is converted to uridine 5'-triphosphate (UTP) by the action of pyruvate kinase, which catalyzes a reversible phosphate transfer between pyruvate/UTP and PEP/UDP. The equilibrium of the reaction is driven towards the formation of pyruvate when a second coupling enzyme, lactate dehydrogenase (LDH), is added in the presence of nicotinamide adenine dinucleotide (NADH). As described previously, LDH reduces pyruvate and consumes NADH at the same time. The rate of reaction catalyzed by NeuB can therefore be followed by the disappearance of the NADH absorption at 340 nm ($\epsilon = 6\,220\text{ M}^{-1}\text{ cm}^{-1}$).

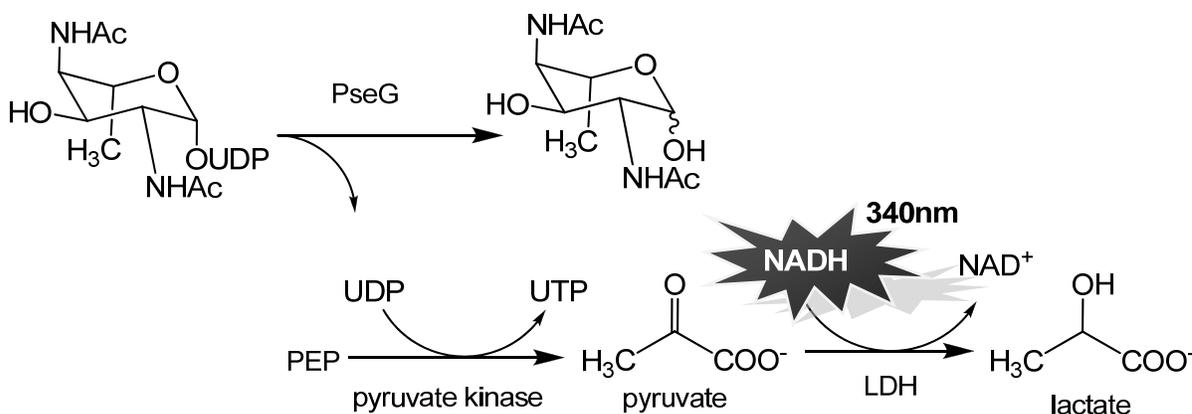


Figure 3.14 Continuous coupled assay monitoring the release of UDP.

The reaction was found to obey Michaelis-Menten kinetics with $k_{\text{cat}} = 26.6 \pm 0.6\text{ s}^{-1}$, $K_{\text{M}} = 174 \pm 11\text{ }\mu\text{M}$, and $k_{\text{cat}}/K_{\text{M}} = 1.5 \times 10^5\text{ M}^{-1}\text{ s}^{-1}$ (Figure 3.15). The reasonably large value of the specificity constant ($k_{\text{cat}}/K_{\text{M}}$) suggests that UDP-6-deoxy-AltdiNAc hydrolysis is the biologically relevant function of this enzyme. The measured kinetic constant (k_{cat}) for the hydrolysis reaction

is comparable or superior to those obtained for glycosyl transfer with related glycosyltransferases.^{165, 166} This further confirms the previously discussed notion that hydrolysis is the normal function of PseG rather than an alternative reaction catalyzed by a glycosyltransferase. In previous studies where glycosyltransferases have been observed to show a background hydrolysis reaction, the rate of hydrolysis is orders of magnitude slower than that of the normal glycosyl transfer reaction catalyzed by the same enzyme.¹⁶⁷⁻¹⁶⁹ A low hydrolase activity ensures that a glycosyltransferase does not needlessly hydrolyze its substrate in an unproductive fashion and one would expect that such an energy-wasting function would not have survived natural selection.

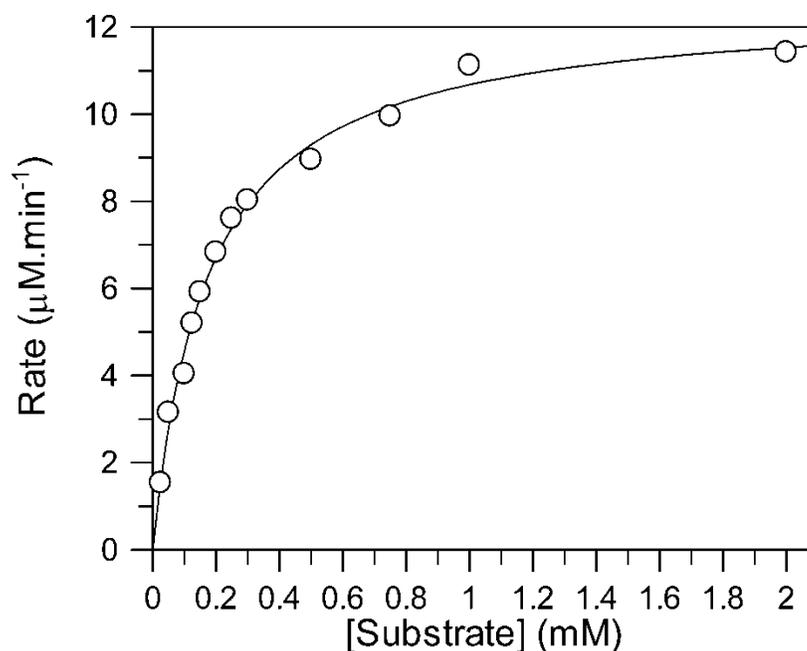


Figure 3.15 Enzyme kinetic plots of initial velocity vs. substrate concentration. The kinetic parameters as determined by fitting the data to Michealis-Menten kinetics are as follows: $k_{\text{cat}} = 26.6 \pm 0.6 \text{ s}^{-1}$, $K_{\text{M}} = 174 \pm 11 \text{ µM}$.

3.5 Metal Dependency of PseG

Many glycosyltransferases are metal-dependent enzymes.¹⁷⁰ Divalent metal ions bind to the nucleotide phosphate moiety and act as Lewis acids. GDP-mannose hydrolase, which is one of the few enzymes that have the same sole function as PseG, belongs to the Nudix superfamily of enzymes and is also metal-dependent.³⁸ However, some glycosyltransferases do not require a metal for catalysis. The fact that PseG is most closely related to the metal-independent GT-B superfamily of enzymes implies that a metal ion may not be involved in the catalysis. To test this notion, the effect of divalent metals on the rate of the PseG reaction was investigated. The continuous coupled assay described in the kinetics study (Section 3.4) cannot be employed because one of the coupling enzymes, pyruvate kinase, requires a metal for its activity.¹⁷¹ Instead, a qualitative ³¹P NMR spectroscopic assay was used to determine the metal requirements of PseG (Figure 3.16).

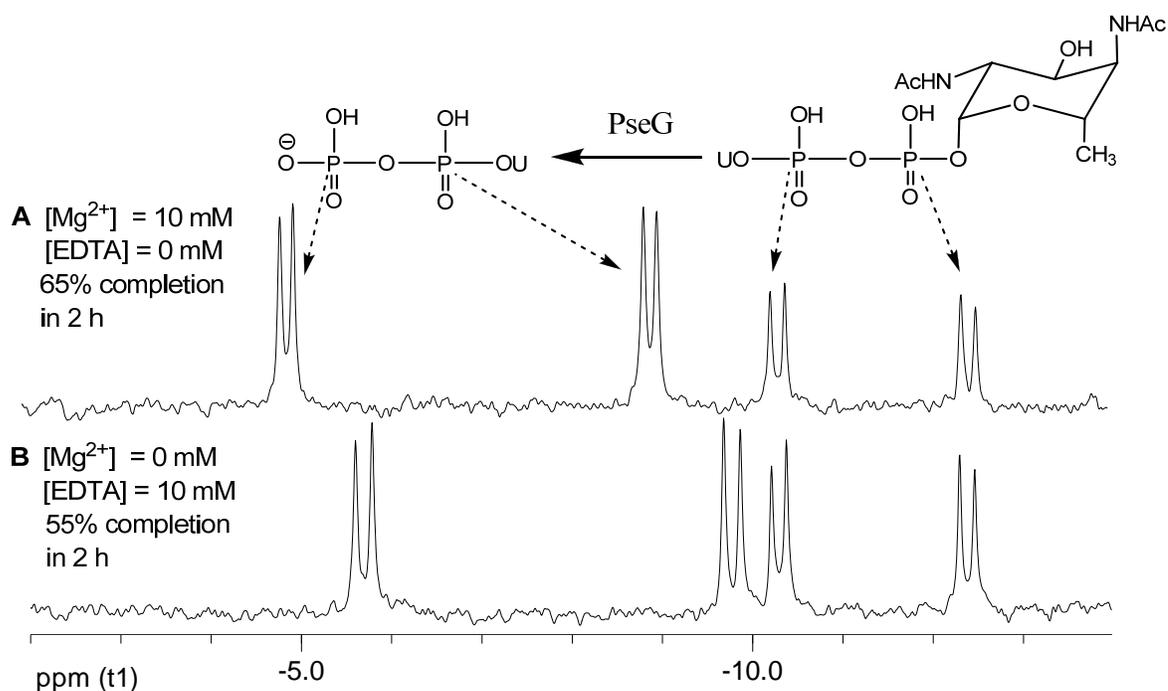


Figure 3.16 ³¹P NMR spectra showing the metal-independency of PseG.

Two identical samples of UDP-6-deoxy-AltdiNAc were incubated with PseG in the presence of either 10 mM Mg²⁺ or 10 mM EDTA. The extent of reaction was then determined by integration of the substrate (UDP-6-deoxy-AltdiNAc) and product (UDP) ³¹P NMR signals. The sample containing 10 mM Mg²⁺ was observed to reach 65% completion in two hours while the sample containing EDTA reached 55% completion during the same period of incubation. Since the amount of EDTA is sufficient to deplete all free and exchangeable divalent metal ions, the similarity between the values strongly suggests that exogenous metal ions are not required for the reaction catalyzed by PseG. Although it is possible that a non-exchangeable metal ion could be tightly bound at the active site, a crystallographic study published later in 2009 rules out this notion by showing the absence of any metal ion in the structures of both apo-PseG and the PseG-UDP complex.¹⁷² The slightly slower rate found in the reaction with 10 mM EDTA may reflect a weak inhibition caused by this anionic molecule. Similar inhibition is reported in a study on a metal-independent CAZy Family 6 glycosyltransferase, which exhibited a small progressive decline in activity upon the addition of EDTA.¹⁷³

3.6 Stereochemistry of the Hydrolase Reaction

As PseG has been identified as a masked glycosyltransferase, the first question raised concerning its action would be whether it is an inverting or a retaining enzyme. However, the experimental determination of the anomeric stereochemistry of the product is somewhat complicated by the fact that the monosaccharide undergoes rapid non-enzymatic mutarotation through the open chain form of the free aldehyde and reaches an equilibrium between the α and β anomers (Figure 3.17).

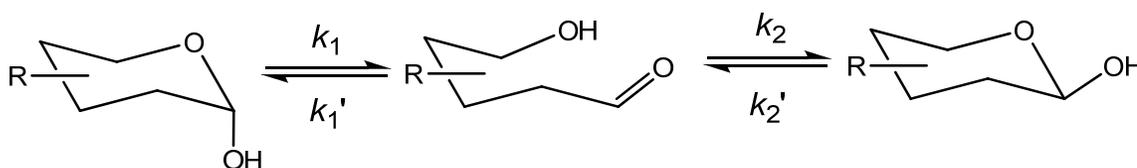


Figure 3.17 Mutarotation between anomers of a pyranose monosaccharide through the open chain aldehyde form.

Studies have shown that the rate constants for mutarotation increase as a linear function of the buffer concentration.¹⁷⁴ This means that mutarotation can be slowed down by decreased buffer concentration. Hence, a Tris-DCl buffer with low concentration (4 mM) was employed in the stereochemical analysis of the first formed product. The reaction was run in buffer prepared using D₂O with a neutral pH (pD 7.4), and was analyzed by ¹H NMR spectroscopy (Figure 3.18). A relatively large amount of PseG was added to the sample and ¹H NMR spectra were acquired at different time points. Immediately after the addition of the hydrolase (2 min), a major product was observed whose H-1 signal appeared as a doublet at 5.04 ppm with a coupling constant ($J_{1,2}$) of 2.8 Hz (Figure 3.18 b). As the reaction proceeded, a new ¹H signal appearing as a doublet at 5.26 ppm with a coupling constant ($J_{1,2}$) of 1.8 Hz started growing in as a result of the generation of the other anomer by non-enzymatic mutarotation (Figure 3.18 c and d). After extended

incubation times, the anomer corresponding to the ^1H signal at 5.26 ppm had become the major product (72 %), indicating that it is the more stable one of the two anomers.

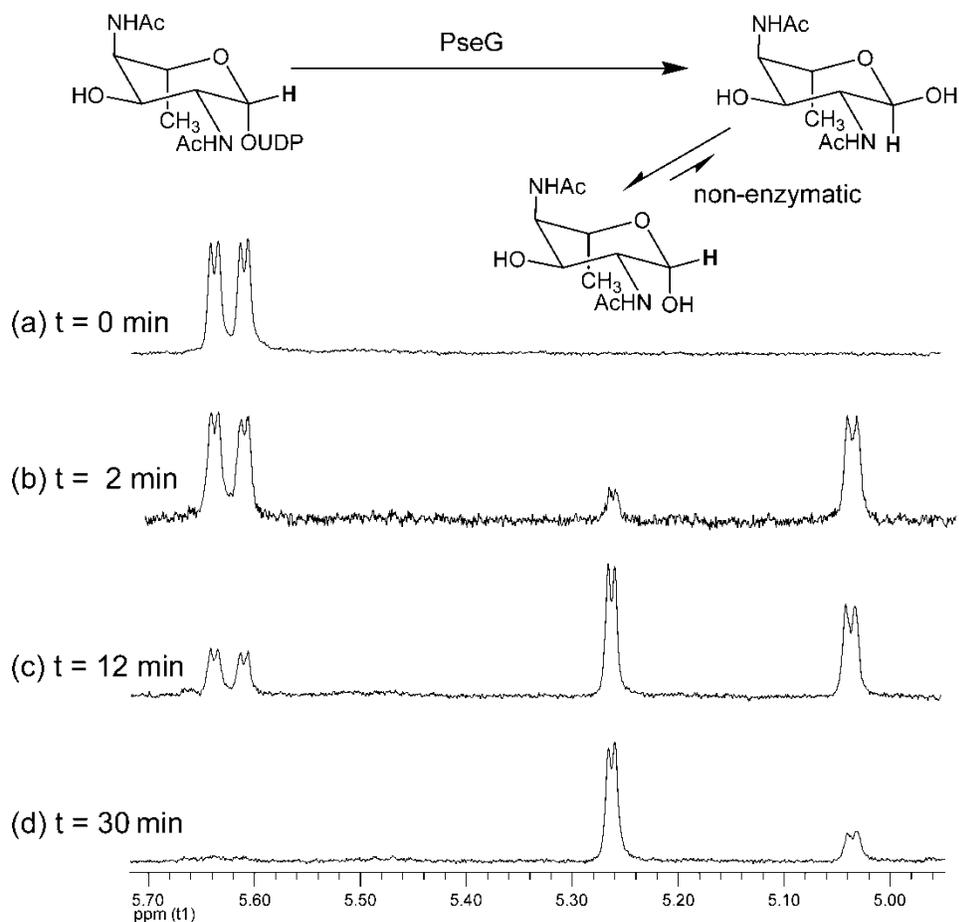


Figure 3.18 ^1H NMR spectra monitoring the time course of the conversion of UDP-6-deoxy-AltdiNAc into α -6-deoxy-AltdiNAc and β -6-deoxy-AltdiNAc. The regions of the spectra showing the anameric proton (H-1) of the 6-deoxy-AltdiNAc compounds are displayed.

The conformational analysis of the resulting altrose derivatives is somewhat complicated by the fact that both of the α - and β -anomers may adopt either a $^4\text{C}_1$ or a $^1\text{C}_4$ conformation (Figure 3.19). Since the interconversion of these conformers will occur rapidly on the ^1H NMR time scale, an averaged signal will be observed. Nevertheless, only the $^4\text{C}_1$ conformation of the α -anomer would bear H-1 and H-2 protons positioned in a trans-diaxial relationship. Therefore

the anomer corresponding to the larger coupling constant ($J_{1,2}$ 2.8 Hz) at 5.04 ppm is expected to be the α -anomer that partially populates the 4C_1 conformation. The identity of the major product at equilibrium (1H NMR δ 5.26 ppm) as the β -anomer was confirmed by correlation spectroscopy (COSY) and nuclear Overhauser effect (NOE) experiments on a sample of the purified 6-deoxy-AltdiNAc. After the assignment of chemical shift using 1H COSY (Figure 3.20), irradiation of the H-1 signal at 5.26 ppm resulted in an enhancement of the H-5 signal (a doublet of quartets) at 3.94 ppm (Figure 3.21). This indicates the existence of a 1,3 diaxial relationship between H-1 and H-5, which would only be expected in the 1C_4 conformation of the β -anomer (Figure 3.19). Thus it can be concluded that PseG is an inverting hydrolase and generates the α -anomer as the first formed product.

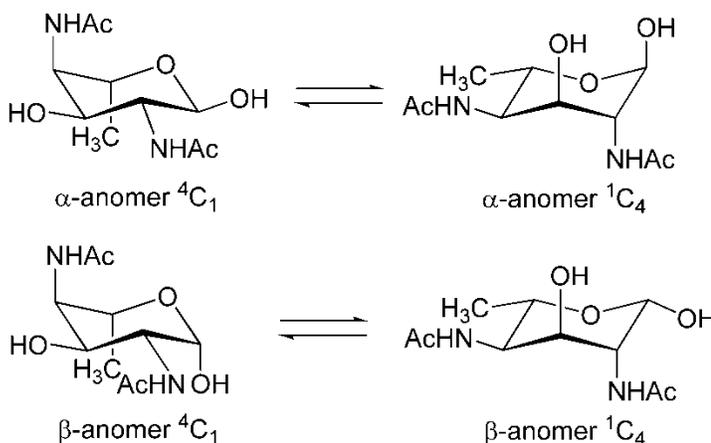


Figure 3.19 4C_1 and 1C_4 conformations of the α - and β -anomers of 6-deoxy AltdiNAc.

In addition to the stereochemical analysis, this experiment also shows that solvent-derived deuterium is not incorporated into the C-2 position during catalysis. This is clearly evidenced by the observed doublet in the H-1 signals of the product (Figure 3.18 d). Mass spectral analysis of the 6-deoxy-AltdiNAc generated in this experiment confirms that no (<5%) non-exchangeable deuterium is incorporated during catalysis.

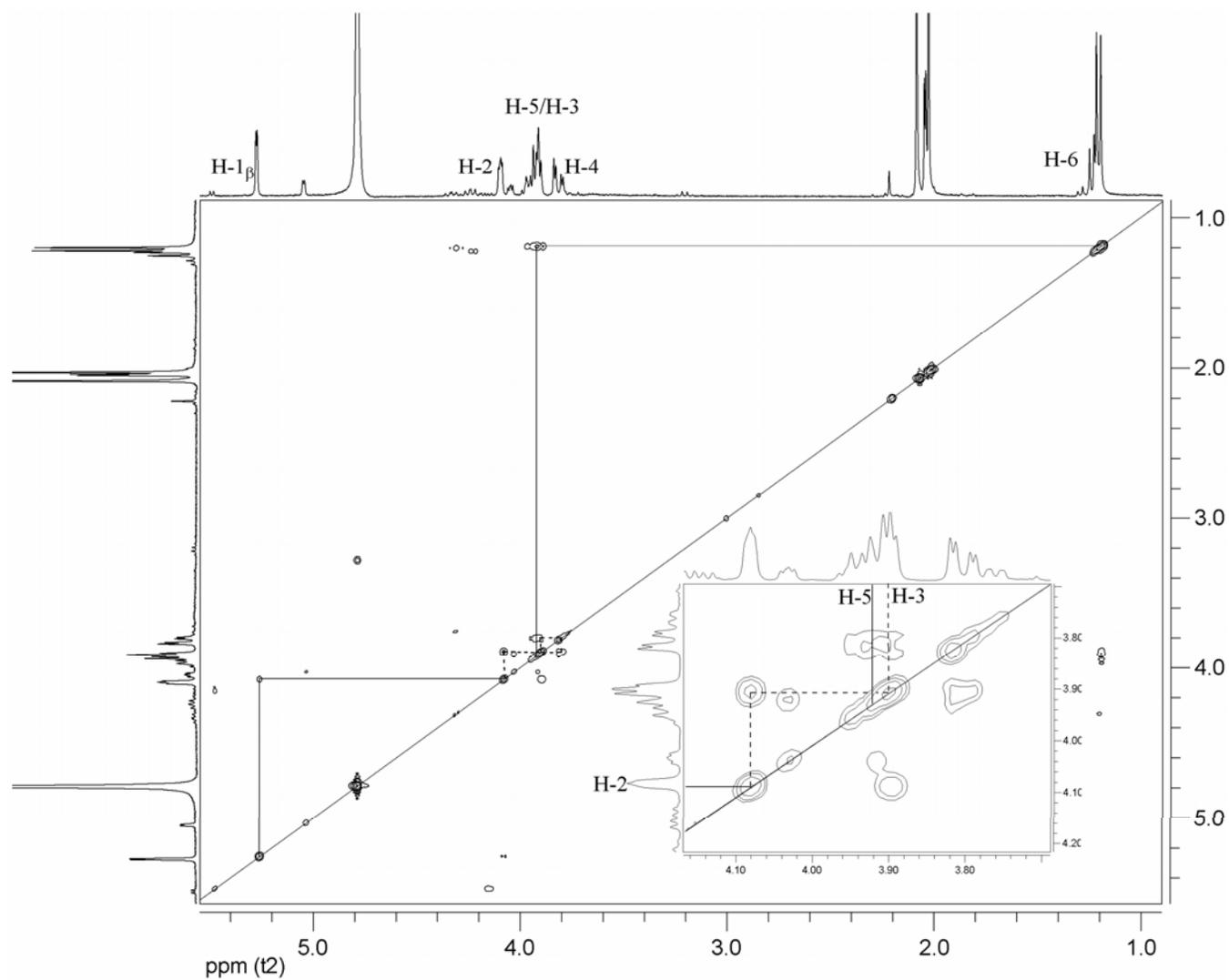


Figure 3.20 COSY spectrum of 6-deoxy-AldiNAc. Correlations between H-5 and H-6 as well as H-2 and H-1 of the β -anomer are shown with solid lines. The insert (dashed lines) shows the identification of H-3 via its coupling to H-2.

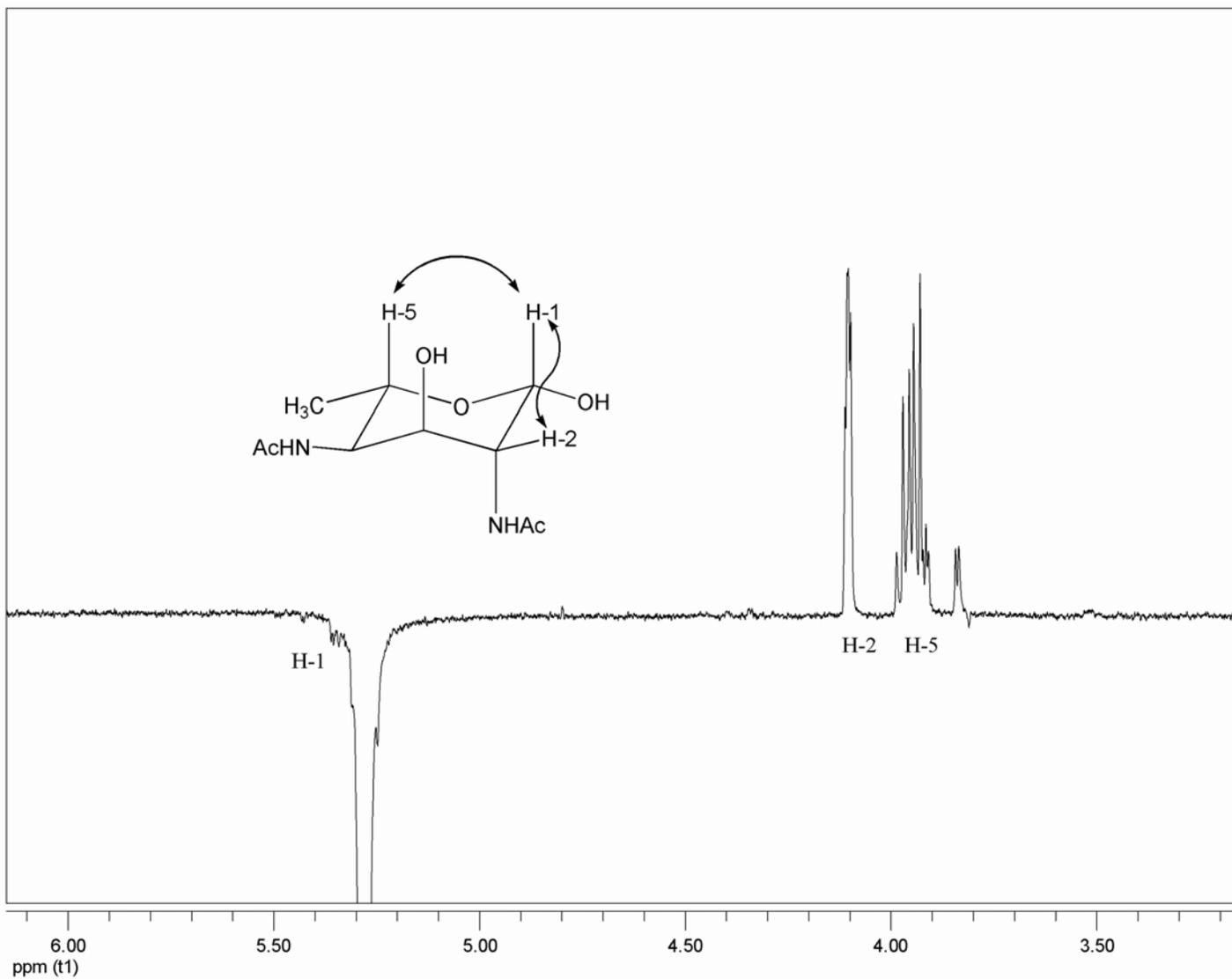


Figure 3.21 NOE difference spectrum of 6-deoxy-AltdiNAc with irradiation at 5.28 ppm (H-1 of the β -anomer).

3.7 Mechanistic Analysis of the Hydrolase Reaction

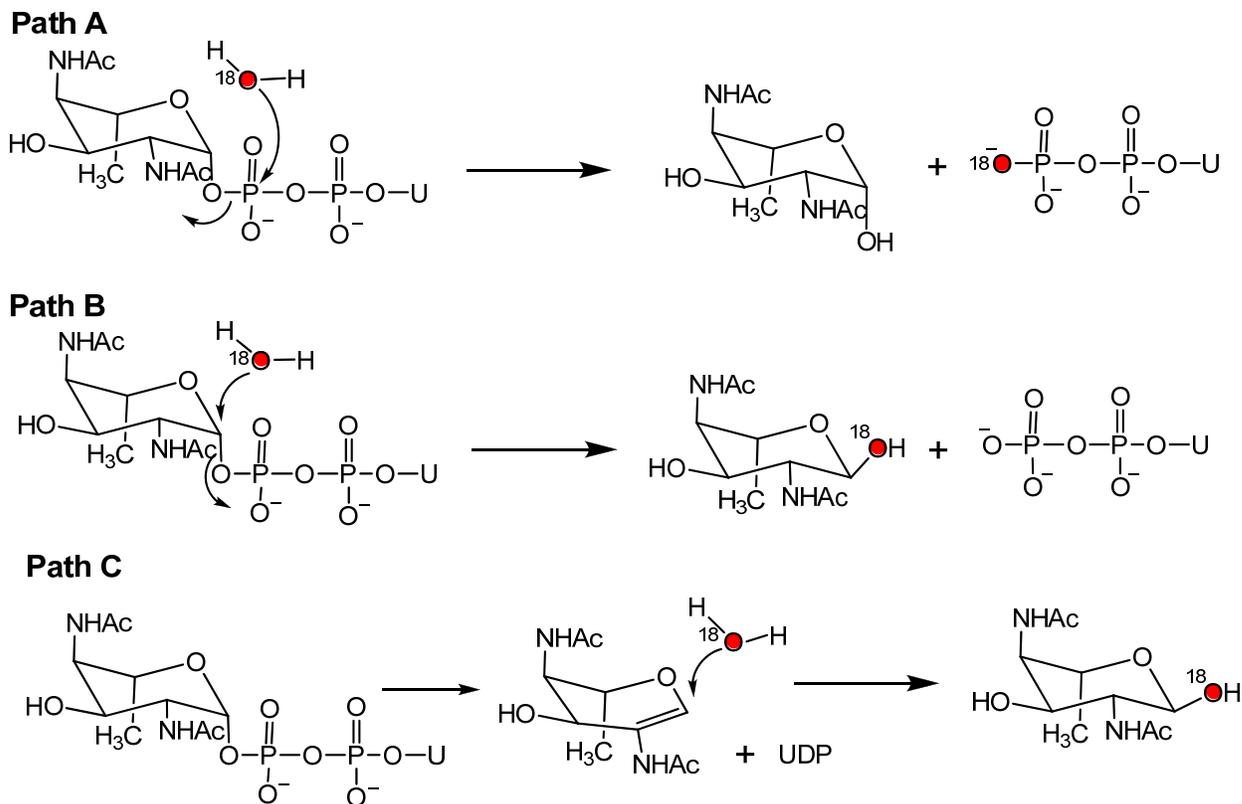


Figure 3.22 Potential mechanisms for the reaction catalyzed by UDP-6-deoxy AltDiNAc hydrolase (PseG). Path A) A P-O bond cleavage mechanism involving direct attack of water at phosphorus. Path B) A C-O bond cleavage mechanism involving direct attack of water at carbon. Path C) A C-O bond cleavage mechanism involving glycol formation. Isotopic labels indicate the consequence of carrying out the reaction in H_2^{18}O .

Three reasonable mechanisms could be employed by PseG during the hydrolysis reaction it catalyzes (Figure 3.22). Path A involves a P-O bond cleavage process in which water attacks the β -phosphorus of the sugar nucleotide and releases the hexose moiety. Alternatively, in a C-O bond cleavage mechanism (Path B), water could attack at the anomeric carbon and displace UDP. This strategy is often employed by glycosyltransferases¹⁶⁸ and is therefore a very possible

mechanism given that PseG is found to bear close resemblance to a glycosyltransferase. Path C, as the third possibility, also involves a C-O bond cleavage. In this case an *anti*-elimination of UDP would generate a glycal intermediate, which undergoes a subsequent *syn*-hydration to give the final product with an overall inversion of configuration at C-1. This resembles the mechanism employed by the hydrolyzing UDP-GlcNAc 2-epimerase, NeuC, which also shares a GT-B fold.^{34, 36, 80}

An ¹⁸O-incorporation experiment was carried out to determine whether PseG utilized a P-O or a C-O bond cleavage mechanism during catalysis. The reaction was run in a buffer prepared from 1:1 H₂¹⁸O:H₂¹⁶O and the products were immediately analyzed by mass spectrometry following the addition of a relatively large amount of enzyme. The observed mass for UDP that was generated in this reaction was identical to that observed in a control reaction conducted in 100% H₂¹⁶O (*m/z* 403, M-H⁺), showing no incorporation of ¹⁸O-label. On the other hand, the ESI-MS(+) signals of 6-deoxy-AltdiNAc showed a 1:1 ratio of *m/z* 269 (M+Na⁺) and *m/z* 271 (M+Na⁺+2), indicating that an ¹⁸O-label had been incorporated into the sugar moiety during catalysis. In order to confirm that the label had not been incorporated subsequent to hydrolysis, such as during the non-enzymatic hydration/dehydration of the free aldehyde, a control experiment was conducted by incubating a sample of 6-deoxy-AltdiNAc with PseG under identical conditions. Although a slow incorporation of ¹⁸O-label was observed over an extended incubation period, the rate of incorporation was several orders of magnitude too slow to account for the incorporation observed during the hydrolase reaction described above. These experiments showed that a solvent-derived oxygen atom is incorporated into the sugar moiety, and not into UDP, during the hydrolysis catalyzed by PseG, and therefore strongly supports the C-O bond cleavage mechanism in which water directly attacks the anomeric carbon to displace UDP. The

determination of a C-O bond cleavage, as well as the observation that the hydrolase is an inverting enzyme, rule out Path A.

Although it is difficult to conclusively distinguish between Paths B and C as they both involve C-O bond cleavage, several lines of evidence argue against the latter mechanism. Firstly, in Path C, the proton that was removed from C-2 of the substrate might exchange with bulk solvent during the lifetime of the glycal intermediate, which would result in solvent-derived deuterium incorporation during the hydration step if the reaction is carried out in D₂O. This type of exchange occurred with the hydrolyzing NeuC of sialic acid biosynthesis^{35, 36} and would likely be observed in the reaction catalyzed by PseG if Path C were at play. However, as mentioned in Section 3.6, no solvent-derived deuterium was incorporated into the C-2 position of 6-deoxy-AltdiNAc during catalysis. This observation is consistent with the mechanism shown in Path B, which involves a direct replacement at the anomeric position and no solvent-derived deuterium incorporation. Secondly, as NeuC catalyzes a hydrolysis involving the retention of stereochemistry at C-1, a related enzyme utilizing a similar mechanism would also be expected to have the same stereochemical outcome. Moreover, considering the significant structural differences between NeuC and PseG, it is unlikely for nature to develop a complex system, as a result of convergent evolution, just to exert the sole function of simple anomeric hydrolysis. Instead, PseG more closely resembles an inverting glycosyltransferase such as MurG, to which it shares the homology of key residues. The striking structural similarity of MurG and PseG has also been reported in the recent crystallographic study.¹⁷² It is thus reasonable to assume that UDP-6-deoxy-AltdiNAc hydrolase (PseG) employs a C-O bond cleavage mechanism involving the direct attack of water at C-1 (Path B).

3.8 Conclusions

The UDP-6-deoxy-AltdiNAc hydrolase of pseudaminic acid biosynthesis had never been identified prior to this study. The work in this chapter demonstrated a novel chemo-enzymatic synthesis of UDP-6-deoxy-AltdiNAc, as well as the first identification and characterization of UDP-6-deoxy-AltdiNAc hydrolase.

Instead of employing a lengthy chemical approach, UDP-6-deoxy-AltdiNAc was effectively synthesized by incubating UDP-GlcNAc with PseB and PseC and by a subsequent chemical acetylation. 6-Deoxy-AltdiNAc obtained from the hydrolysis of UDP-6-deoxy-AltdiNAc under mild acidic conditions was proven to be identical to that chemically synthesized in previous studies. This method provides a concise and useful way to synthesize substrate and substrate analogues for UDP-6-deoxy-AltdiNAc hydrolase and enzymes involved in subsequent biosynthetic steps.

In Section 3.3, two enzymes, encoded by the *Campylobacter* genes Cj1328 and Cj1312, respectively, were cloned, over-expressed and tested as potential candidates for UDP-6-deoxy-AltdiNAc hydrolase. The enzyme Cj1328 is homologous to UDP-GlcNAc 2-epimerase (NeuC) in the closely related sialic acid biosynthesis, and was suggested to be involved in pseudaminic acid biosynthesis by previous mass spectral analysis on metabolites. This enzyme was over-expressed as His₆-tagged and MBP-fusion constructs but showed no activity with UDP-6-deoxy-AltdiNAc. It was later proven to be involved in legionaminic acid biosynthesis in *Campylobacter spp.*. On the other hand, the other enzyme, Cj1312, is distantly related to a family of glycosyltransferases, and therefore could possibly serve as a hydrolase. Its homologue in *H. pylori*, FlmD (HP0326b), was previously shown to be involved in the flagellar glycosylation

process. In this study, we successfully determined that Cj1312 (or PseG) is the previously unidentified UDP-6-deoxy-AltdiNAc hydrolase. PseG exhibited zero or extremely low activity towards UDP-GlcNAc and UDP-6-deoxy-AltNAcN, implying good substrate specificity. This finding also strongly suggests that FlmD (HP0326b) plays the same role in *H. pylori*.

In following studies, PseG was characterized for the first time. The activity of this enzyme is independent of the presence of divalent metal ions, and the values of the catalytic constants were found to be $k_{\text{cat}} = 27 \text{ s}^{-1}$ and $K_{\text{M}} = 174 \text{ }\mu\text{M}$. The enzyme was shown to hydrolyze the substrate with an overall inversion of stereochemistry at C-1 using ^1H NMR analysis of the stereochemistry of the first formed product. This result, coupled with isotope incorporation studies, demonstrated that PseG utilizes a mechanism involving direct nucleophilic attack at the anomeric position and an accompanying C-O bond cleavage during catalysis.

Homology comparisons and results of the mechanistic study suggest that the closest ancestors to PseG are members of the metal-independent GT-B family of glycosyltransferases that include the peptidoglycan biosynthetic enzyme MurG. PseG, a masked glycosyltransferase, functions as a hydrolase by transferring UDP-6-deoxy-AltdiNAc to a water molecule. The closest precedent to this reaction is the one catalyzed by GDP-mannose hydrolase, which is also an inverting hydrolase that utilizes a C-O bond cleavage mechanism during catalysis.³⁸ However, GDP-mannose hydrolase is a member of the Nudix family of enzymes and employs a divalent cation that binds to the α - and β -phosphates of the substrate in a bidentate fashion.¹⁷⁵ Thus, UDP-6-deoxy-6-deoxy-AltdiNAc hydrolase (PseG) involved in the pseudaminic acid biosynthesis, appears to be one of the few examples of sugar nucleotide hydrolases that have evolved from the metal-independent GT-B family of glycosyltransferases.

3.9 Future Directions

The crystal structure of the PseG•UDP complex was published by Rangarajan *et al.* in 2009.¹⁷² Together with mutagenesis results, the roles of active site residues and conformations of the sugar moiety were suggested by software simulation and structural comparison with MurG. This work has provided helpful insights into the detailed mechanism of the action of PseG. However, information regarding the binding pocket of the sugar moiety is largely speculative. The attempts to co-crystallize PseG with 6-deoxy-AltdiNAc, UDP-GlcNAc and UDP-6-deoxy-AltdiNAc were unsuccessful, which perhaps is not too surprising. Firstly, the free sugar, 6-deoxy-AltdiNAc, likely does not have enough binding affinity and may require the presence of UDP in order to be crystallized with the enzyme. Secondly, as we have demonstrated earlier, PseG shows no detectable activity with UDP-GlcNAc. This implies there is a low binding affinity of UDP-GlcNAc towards the enzyme, as a result of the structural differences between UDP-GlcNAc and UDP-6-deoxy-AltdiNAc. Lastly, PseG could still be active under the crystallization conditions and capable of hydrolyzing UDP-6-deoxy-AltdiNAc. This is consistent with the fact that the co-crystallization of PseG and UDP-6-deoxy-AltdiNAc yielded a PseG•UDP complex.

In future studies, attempts could be made to co-crystallize PseG with 6-deoxy-AltdiNAc and UDP. It is also worthwhile trying to soak 6-deoxy-AltdiNAc or UDP-6-deoxy-AltdiNAc into the previously obtained crystal that had UDP and glycerol bound at the active site. Additionally, as UDP-GlcNAc conceivably has very low binding affinity towards PseG, and UDP-6-deoxy-AltdiNAc is likely hydrolyzed by the enzyme under crystallization conditions, UDP-6-deoxy-AltNac4N, the immediate precursor of UDP-6-deoxy-AltdiNAc, can be potentially co-crystallized with PseG (Figure 3.23). This compound was shown to be a very poor substrate for

PseG, and thus may survive the crystallization conditions. Furthermore, the substrate analogue **26** could be synthesized, tested as an inhibitor of PseG, and used in the crystallographic study (Figure 3.23). This compound would serve as an excellent structural mimic of UDP-6-deoxy-AltdiNAc but could not undergo hydrolysis due to the anomeric methylene group.

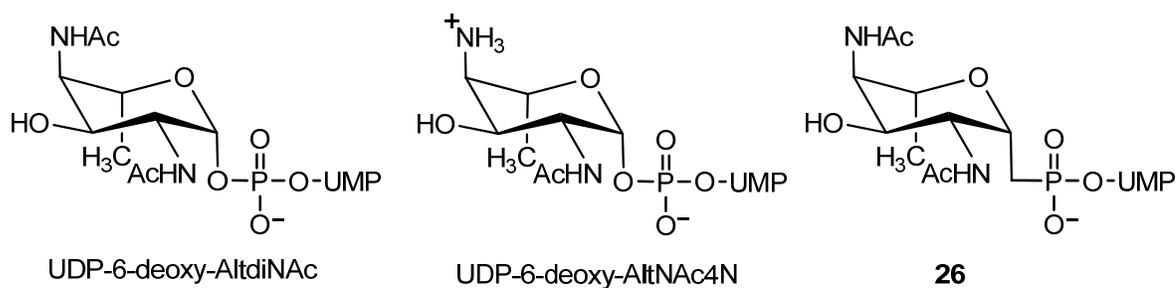


Figure 3.23 UDP-6-deoxy-AltdiNAc and potential candidates for co-crystallization with PseG.

As mentioned in Section 3.3.1.2, Cj1328 or LegG was identified as a GDP/UDP-Bac2Ac4Ac 2-epimerase. It very likely employs a similar mechanism to that employed by the UDP-Bac2Ac4Ac 2-epimerase from *Legionella pneumophila*.¹⁷⁶ Similar strategies, such as determination of the anomeric stereochemistry of the first formed product, as well as isotope incorporation studies, could be used to characterize this enzyme (Figure 3.24).

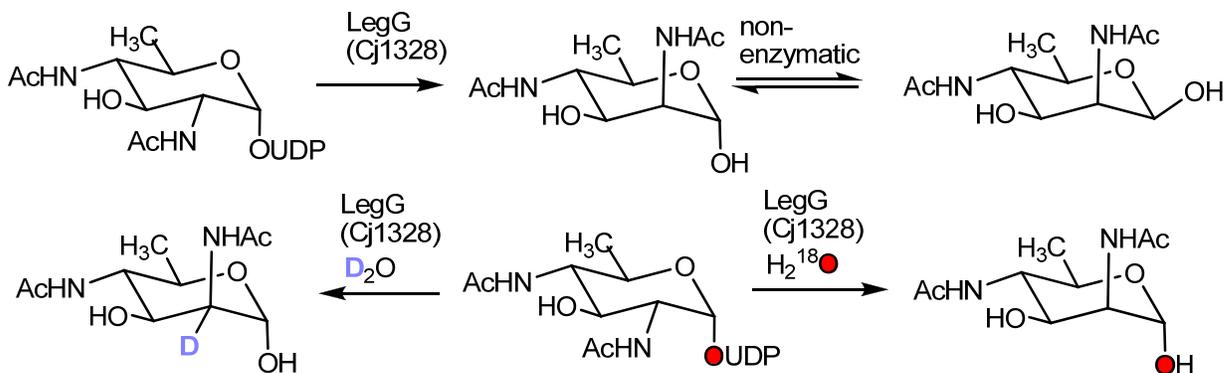
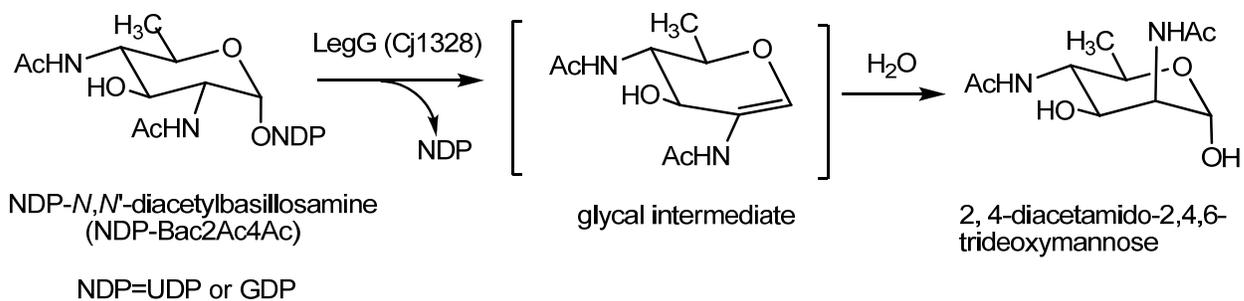


Figure 3.24 Proposed mechanism of LegG (Cj1328) and strategies to characterize it.

3.10 Experimental Procedures

3.10.1 Materials and General Methods

The previous materials and general methods section from Chapter 2 also applies in this study with the following additions. UDP-*N*-acetylglucosamine, lactate dehydrogenase (Type II from rabbit muscle) and pyruvate kinase (Type II from rabbit muscle) were purchased from Sigma-Aldrich. ¹⁸O-enriched H₂O (95%) was purchased from Cambridge Isotope Laboratories. pMAL™ Protein Fusion and Purification System, including Factor Xa and amylose resin, was purchased from New England Biolabs. pET 30 Vector Kit was purchased from Novagen. *Tag* DNA polymerase and dNTP mix were purchased from Invitrogen.

3.10.2 Chemoenzymatic Synthesis of UDP-6-deoxy-AltDiNAc

Samples of the crude lysates containing PseB and PseC (100 μL each) were added to 100 mL of triethanolamine-HCl buffer (50 mM, pH 8.0) containing 450 mg of UDP-*N*-acetylglucosamine disodium salt, 0.1 M monosodium glutamate, 0.25 mM pyridoxal 5'-phosphate (PLP), and 0.25 mM nicotine adenine dinucleotide (NAD⁺). This solution was incubated for 4 hr at 37 °C and the reaction progress was monitored by negative ESI-mass spectrometry. It was determined that > 90% of the UDP-GlcNAc (*m/z* 606, M-H⁺) was converted to UDP-6-deoxy-4-amino-2-diacetamido-L-altrose (UDP-6-deoxy-AltNAc4N, *m/z* 589, M-H⁺) during this time. Enzyme was then removed by centrifugation and centrifugal ultrafiltration. The resultant filtrate was loaded onto a 220 mL column of DEAE cellulose (DE 52, Whatman Inc.) and eluted with a linear gradient of 0.1 to 0.5 M triethylammonium bicarbonate buffer. The eluant was monitored at A₂₅₄ and UV-active fractions were analyzed by ESI mass spectrometry. Those containing UDP-6-deoxy-AltNAc4N were lyophilized to dryness.

The lyophilized UDP-6-deoxy-AltNAc4N (310 mg) was stirred with 0.5 g silver acetate and 0.8 mL acetic anhydride in 25 mL methanol at room temperature for 12 hrs. Negative ESI-MS showed that the starting material was completely converted to UDP-6-deoxy-AltdiNAc (m/z 631, $M-H^+$) during this time. After filtration and evaporation 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 H_2O and passed through a 20 mL column of Amberlite IR120 resin (sodium form, Aldrich). The eluant was re-lyophilized to give 173 mg of the UDP-6-deoxy-AltdiNAc disodium salt. 1H NMR(D_2O) δ 1.29 (d, 3H, $J_{5'',6''}$ 6.3 Hz, H-6''), 2.06 (s, 3H, CH_3), 2.09 (s, 3H, CH_3), 3.91 (dd, 1H, $J_{3'',4''}$ 3.1 Hz, $J_{4'',5''}$ 9.2 Hz, H-4''), 4.04 (dd, 1H, $J_{2'',3''}$ 4.6 Hz, $J_{3'',4''}$ 3.1 Hz, H-3''), 4.05 (dq, 1H, $J_{4'',5''}$ 9.2 Hz, $J_{5'',6''}$ 6.3 Hz, H-5''), 4.19 (dd, 1H, $J_{1'',2''}$ 2.2 Hz, $J_{2'',3''}$ 4.6 Hz, H-2''), 4.23 (m, 2H, H-5'), 4.30 (m, 1H, H-4'), 4.39 (m, 2H, H-2' and H-3'), 5.66 (dd, 1H, $J_{1'',2''}$ 2.2 Hz, $J_{H,P}$ 8.5 Hz, H-1''), 5.98 (m, 2H, H-1' and H-5), 8.00 (d, 1H, $J_{6',7'}$ 8.14 Hz, H-6). See Appendix Figure A.5 for detailed information. ^{31}P NMR(D_2O) δ -10.31 (d, $J_{P\alpha,P\beta}$ 20.0 Hz, $P\alpha$), -12.40 (d, $J_{P\alpha,P\beta}$ 20.0 Hz, $P\beta$). ESI-MS (-) m/z 631($M-H^+$).

3.10.3 Chemical Hydrolysis and Characterization of the Product

To a 1 mL solution containing 6 mg of UDP-6-deoxy-AltdiNAc was added 50 μL of concentrated HCl. The mixture was incubated at 37 $^{\circ}C$ for 10 min and then loaded on a 10 mL column of DE-52 anion exchange resin and eluted with distilled H_2O . The flow-through was collected, lyophilized and re-dissolved in D_2O . 1H -NMR spectroscopy showed that the mixture of 6-deoxy-2,4-diacetamido-L-altrose anomers produced was identical to that obtained from chemical synthesis (13,15). 1H (D_2O), δ 1.21 (d, 3H, $J_{5\beta,6\beta}$ 6.1 Hz, H-6 β), 1.24 (d, 3H, $J_{5\alpha,6\alpha}$ 6.5 Hz, H-6 α), 2.03 (s, 3H, β - CH_3), 2.04 (s, 3H, α - CH_3), 2.05 (s, 3H, α - CH_3), 2.09 (s, 3H, β - CH_3),

3.82 (dd, 1H, $J_{3,4}$ 2.9 Hz, $J_{4,5}$ 10.4 Hz, H-4), 3.91 (dd, 1H, $J_{2,3}$ 3.2 Hz, $J_{3,4}$ 2.9 Hz, H-3), 3.94 (dq, 1H, $J_{4,5}$ 10.1 Hz, $J_{5,6}$ 6.1 Hz, H-5), 4.05 (dd, 1H, $J_{1\alpha,2\alpha}$ 2.8 Hz, $J_{2\alpha,3\alpha}$ 4.3 Hz, H-2 $_{\alpha}$), 4.19 (dd, 1H, $J_{1\beta,2\beta}$ 1.8 Hz, $J_{2\beta,3\beta}$ 3.2 Hz, H-2 $_{\beta}$), 5.05 (d, 3H, $J_{1\alpha,2\alpha}$ 2.8 Hz, H-1 $_{\alpha}$), 5.28 (d, 3H, $J_{1\beta,2\beta}$ 1.8 Hz, H-1 $_{\beta}$). See Appendix Figure A.6 for detailed information. ESI-MS(+) m/z 269 (M+Na⁺).

3.10.4 Cloning, Over-expression and Purification of Enzymes.

3.10.4.1 Cloning of *neuC2* (Cj1328 or *legG*) and *pseG* (Cj1312)

The *neuC2* gene (Cj1328 or *legG*) was amplified from the MBP-*neuC2* construct plasmid by polymerase chain reaction (PCR). Oligonucleotide primers, including overhangs for ligation-independent cloning, were: 5'-GGTATTGAGGGTCGCGTGAGTAAA AGAAAAATTT-3' (sense) and 5'-AGAGGAGAGTTAGAGCCTTATAAATCGATGAAATT-3' (antisense). To a 200 μ L tube were added: 5 μ L of 10 \times PCR reaction buffer (Invitrogen), 1.5 μ L of 50 mM MgCl₂, 1 μ L of dNTP mix (PCR grade, Invitrogen), 125 ng of each primer, 100 ng of template DNA, 1.25 u *Tag* polymerase (Invitrogen) and distilled H₂O to make a total volume of 50 μ L. PCR was conducted on an iCycler Thermal Cycler (Bio-Rad) with the following cycles: 30 cycles of 1 min at 94 $^{\circ}$ C, 1 min at 50 $^{\circ}$ C, 90 s at 72 $^{\circ}$ C, and one cycle of 10 min at 72 $^{\circ}$ C. The PCR product was cloned into a pET-30 Xa/LIC vector (Novagen) according to the manufacturer's instructions. The resulting recombinant plasmid, which encodes a His₆-tag on the N-terminal of the target Cj1329 protein, was amplified in NovaBlue GigaSingles competent *E. coli* cells (Novagen).

The *pseG* gene (Cj1312) was amplified from *Campylobacter jejuni* (strain NCTC 11168) genomic DNA by PCR. Oligonucleotide primers, including overhangs for ligation-independent cloning, were: 5'-GGTATTGAGGGTCGCATGAAAGTGCTTTTTAGAAAG-3' (sense) and 5'-AGAGGAGAGTTAGAGCCTCAATACTTATACTCCA-3' (antisense). The conditions used to

clone *pseG* are identical to those described above for *neuC2*. The resulting recombinant plasmid, which encodes a His₆-tag on the N-terminal of the target *pseG* protein, was also amplified in NovaBlue GigaSingles competent *E. coli* cells (Novagen)

3.10.4.2 Over-expression and Purification of MBP-*neuC2* (Cj1312)

The MBP-*neuC2* (Cj1312) plasmid was transformed into *E. coli* BL21 (DE3) competent cells which were incubated in 10 mL Lucia-Bertani (LB) medium containing 2g/L glucose and 100 mg/L ampicillin at 37 °C/225 rpm for 10 h. The overnight culture was then poured into 500 mL of LB medium containing 2g/L glucose and 100 mg/L ampicillin and shaken at 37 °C/225 rpm until an OD₆₀₀ of 0.5 had been reached. The culture was allowed to continue to grow for 4 h after 72 mg/L of isopropyl β-D-thiogalactopyranoside (IPTG) was added. Cells were harvested by centrifuging at 4 000×g for 20 min, and then resuspended in 10 mL of phosphate buffer (20 mM, pH 8.0) 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 9 000×g for 30 min and passed through a 0.22 μm filter.

A column containing 10 mL of amylose resin (New England Biolabs) was washed with 20 mL of distilled H₂O and 80 mL of Tris-HCl buffer (20 mM, pH 7.4, containing 0.2 M NaCl and 1 mM of EDTA). The clarified lysate was loaded onto the column which was washed with 80 mL of the same buffer. The enzyme is eluted with 10 mM of maltose in the above described Tris-HCl buffer. Eluate fractions that showed absorbance at 280 nm were collected. These fractions were concentrated using Amicon Ultra Centricons before flash freezing with liquid N₂ in the presence of 10% glycerol and 2 mM DTT.

NeuC2 (Cj1312 or LegG) could be cleaved from the MBP-NeuC2 fusion protein (1 mg/mL) using Factor Xa (1 μ L, 200 μ g/mL) over a period of 8 h.

3.10.4.3 Over-expression and Purification of *neuC2* gene (Cj1328 or *legG*), *pseG*, *pseB* and *pseC*

The conditions used for over-expression and purification of His₆-tagged NeuC2 and His₆-tagged PseG are identical to those described previously for His₆-tagged NeuB in Section 2.7.3.1.

The preparation of the plasmids used in the over-expression of His₆-tagged PseB (Cj1293) and PseC (Cj1294) have been described previously.¹⁵⁴ The conditions used for over-expression of these proteins were also identical to those described above for His₆-tagged NeuB with the exception that the growth media used to prepare PseC contained 50 mg/L ampicillin in place of kanamycin. The clarified lysates obtained following lysis and filtering were used directly without further purification.

3.10.5 Enzyme Kinetics Using a Continuous Coupled Assay

Enzyme kinetics were measured using a continuous coupled assay for UDP formation, under conditions slightly modified from those described by Chou *et al.*⁸⁰ Each cuvette contained 50 mM NaH₂PO₄ buffer (pH 7.5), 10 mM MgCl₂, 2 mM PEP, 0.2 mM NADH, 20 units of lactate dehydrogenase, 18 units of pyruvate kinase, and the UDP-6-deoxy-AltdiNAc concentration was varied from 25-2000 μ M. The concentrations of stock UDP-sugar solutions were determined by measuring A_{262} ($\epsilon = 9\,890\text{ M}^{-1}\text{ cm}^{-1}$). Enzymatic reactions were initiated by addition of 20 μ L of enzyme solution (final concentration 7.9 nM). Rates were measured by monitoring the decrease A_{340} at 37 °C ($\epsilon = 6\,220\text{ M}^{-1}\text{ cm}^{-1}$). Kinetics parameters were determined by fitting initial velocities to the Michaelis-Menten equation using GraFit 4.0. The reported error

range was calculated from the non-linear least square fitting on results of kinetic trials carried out on the same day, with enzymes, substrates and/or inhibitors from the same batch. However, systematic errors can be occur due to factors such as use of different batches of enzymes and inhibitors, errors in concentration measurements, as well as temperature fluctuation. These typically lead to an additional 15% error when measurements were performed on different days. No detectable background release of UDP was observed in the absence of added PseG.

3.10.6 Metal Dependency Experiment

Two aliquots of a solution containing UDP-6-deoxy-AltdiNac disodium salt (6 mg per aliquot) in 50 mM Tris-HCl buffer (pH 7.4) were prepared. To one aliquot was added MgCl_2 and to the other was added EDTA tetrasodium salt (each at a 10 mM final conc. in a total volume of 990 μL). PseG (10 μL of a 33 μM stock solution) was added to each sample and the mixtures were incubated for 2 h at room temperature. The progress of the reactions was monitored using ^{31}P NMR spectroscopy with integration of the diphosphate signals.

3.10.7 Reaction Stereochemistry and Deuterium Incorporation

A glycerol stock solution of PseG (200 μL , 33 μM) was subjected to buffer-exchange with a 20 mM Tris-DCl/ D_2O buffer (pD 7.4, 200 μL final volume) using centrifugal ultrafiltration. This was added to a solution of UDP-6-deoxy-AltdiNac disodium salt (5 mg) dissolved in 800 μL D_2O (1 mL final volume) and ^1H NMR spectra were acquired at timed intervals.

In order to assign the signals of the anomeric products, a sample of UDP-6-deoxy-AltdiNac disodium salt was dissolved in D_2O and analyzed by NOE spectroscopy with irradiation at 5.28 ppm (H-1 of the major anomer at equilibrium). An NOE enhancement of the

H-5 signal confirmed that the β -anomer was the predominant species at equilibrium (Figure 3.21). Assignment of the H-5 signal was achieved using COSY spectroscopy via an observed correlation with the H-6 methyl signal (Figure 3.20).

3.10.8 ^{18}O Incorporation Experiment

A solution of 50 mM Tris-HCl buffer (pH 7.4, 400 μL) was prepared using 50% H_2^{16}O /50% H_2^{18}O and was divided into two equal aliquots. To one aliquot was added UDP-6-deoxy-AltdiNAc (5.0 mg) and to the other was added 6-deoxy-AltdiNAc (2.5 mg) and UDP (4.0 mg). PseG (10 μL of a 33 μM stock solution) was added to each sample and the mixtures were incubated at room temperature. 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 6-deoxy-AltdiNAc was deduced from the ratio of peaks at m/z 269 (^{16}O , $\text{M}+\text{Na}^+$) and m/z 271 (^{18}O , $\text{M}+\text{Na}^++2$).

Chapter 4 : The Engineering of Bacteria Bearing Azido-Pseudaminic Acid-Modified Flagella^d

^d A version of this chapter has been published, some figures are reproduced with permission from: Feng Liu, Annie J. Aubry, Ian C. Schoenhofen, Susan M. Logan, Martin E. Tanner, (2009) "The Engineering of Bacteria Bearing Azido-Pseudaminic Acid-Modified Flagella", *ChemBioChem* 10: 1317-1320. © 2009 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

4.1 Introduction

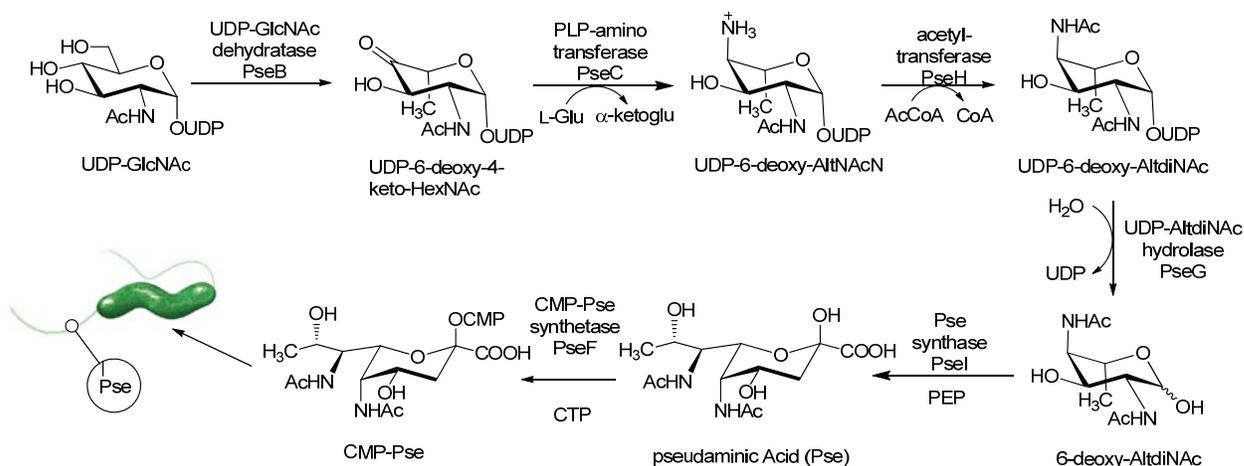


Figure 4.1 Biosynthesis of CMP-pseudaminic acid.

After the identification and characterization of UDP-6-deoxy-AldiNAc hydrolase (PseG), several papers were published on other enzymes involved in pseudaminic acid biosynthesis, including PseH and PseF.^{75, 76} These findings not only led to the full elucidation of the pseudaminic acid biosynthetic pathway in *C. jejuni* and *H. pylori* (Figure 4.1), but also made the manipulation of Pse-producing machinery possible. The fact that pseudaminic acid is displayed on the surface of flagella, and that we are able to synthesize all CMP-Pse precursors and their corresponding derivatives, inspired an idea: can we track Pse in its native environment, using an unnatural Pse precursor bearing a bio-orthogonal chemical reporter, to gain further insight into related biological processes at a cellular level? As was mentioned in Section 1.4, Bertozzi and coworkers had incorporated modified sialic acid onto cell surfaces by feeding ManNAc bearing bio-orthogonal chemical reporters to sialic acid-producing mammalian cells.^{95, 124, 177} These reporters can be subsequently modified using small reactive probes and visualized using a variety of detection methods. In this chapter, we will feed a mutant strain of *C. jejuni*, which

lacks one of the enzymes required for Pse biosynthesis and is therefore deficient in Pse production, with an azido-derivative of a Pse precursor (which is the product of the missing enzyme). Considering the positive results from the azido-labeling of cells using the closely related sialic acid biosynthetic system, an azido derivative will also likely be tolerated by the Pse production/flagella assembly machineries. The presence of azido-Pse on the flagella can then be detected using methods that employ Staudinger reagents.

The first goal of this project was to prepare a suitable precursor that might be taken up by *C. jejuni* cells and converted into an azido-labeled Pse derivative *in vivo*. Cell membranes mainly consist of a phospholipid bilayer that only allows small, non-polar molecules to permeate by passive diffusion.¹⁷⁸ The phospholipid bilayer is impermeable to most water soluble (hydrophilic) molecules, especially charged species. These hydrophilic/polar molecules have to be transported into the cells by specific active transport systems that require energy. The uptake of neutral sugars into bacterial cells occurs in several different ways including¹⁷⁸: a) phosphate transport systems (PTSs) that phosphorylate sugars as they are transported across the membrane, b) ATP-binding cassette (ABC) importers that contain binding proteins exterior to the cell membrane and c) chemiosmotic systems energized by H⁺ or Na⁺ cotransport. Charged species, such as UDP-sugar nucleotides and Pse, are very unlikely to pass through the phospholipid bilayer by diffusion or to be tolerated by the neutral sugar-specific active transport systems. Therefore, the only uncharged compound in the Pse biosynthetic pathway, 6-deoxy-AltdiNAc, became our target for labeling with bio-orthogonal reporters.

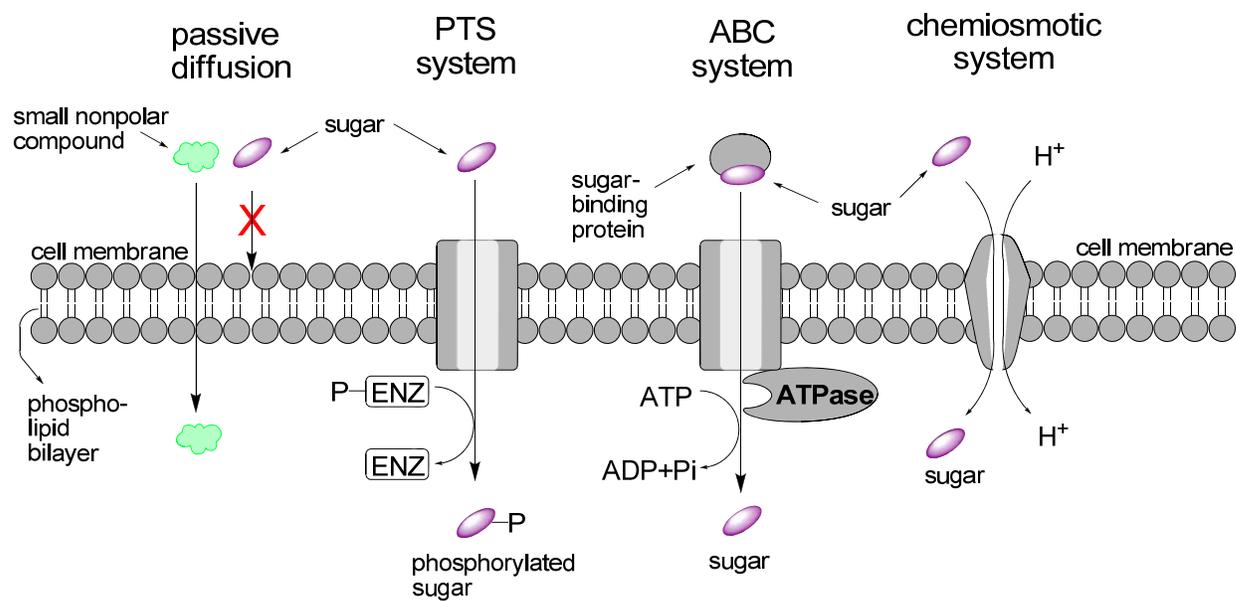


Figure 4.2 Potential mechanisms for the transport of sugar across the cell membrane.

Considering the stability and biocompatibility of organic azides, we decided to introduce an azido functionality at the C-4 acetyl group of 6-deoxy-AltNac. Therefore, 2-acetamido-4-azidoacetamido-2,4,6-trideoxy-L-altrose (6-deoxy-AltNac4NAz), **27**, became our first synthetic target (Figure 4.3). This neutral sugar is likely to be transported into the *C. jejuni* cell by the active sugar transport systems and utilized by the Pse production machinery. If the uptake of this neutral, unprotected sugar turns out to be insufficient for Pse production, a peracetylated version of this sugar (**28**) can be employed for better uptake. It had been shown that acetylation greatly increases membrane permeability of monosaccharides¹²³, which enables the uptake of sugars by cells via passive diffusion. Once transported into the cell, the *O*-acetyl groups on the sugar would be readily cleaved by cytosolic esterases.¹⁷⁷ This could result in a 200-fold increase in the sugar uptake as compared to treatment with unprotected sugar.¹²³

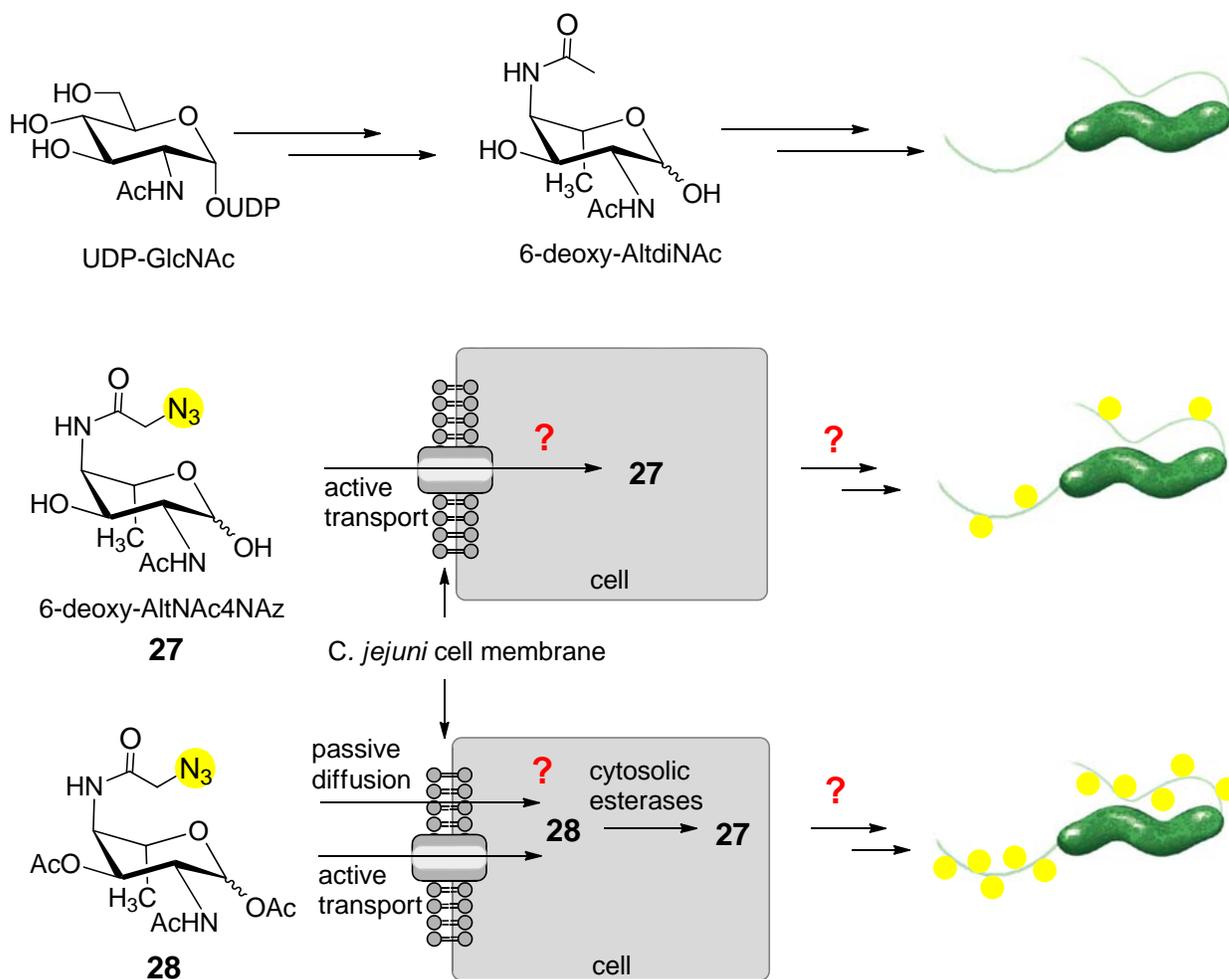


Figure 4.3 6-Deoxy-AltNac4NAz (27) and compound 28 as potential targets to introduce an azido functionality onto the *C. jejuni* flagella.

Once synthesized, the modified Pse precursor 6-deoxy-AltNac4NAz would then be subjected to *in vitro* studies to test if it can be converted to azido-pseudaminic acid (Az-Pse) by pseudaminic acid synthase (PseI). Ultimately, motility tests using semi-solid agar, as well as streptavidin-HRP detection of biotin-Staudinger reagent-treated flagella could be used to demonstrate that azido-pseudaminic acid was incorporated into the flagellin proteins of functional flagella.

4.2 Synthesis of Azido Sugars and Labeling Reagents

4.2.1 Chemo-enzymatic Synthesis of Azido-altrose

A chemo-enzymatic synthetic strategy similar to that utilized in the synthesis of 6-deoxy-AltDiNAc was employed. UDP-GlcNAc was first treated with PseB and PseC in the presence of pyridoxal 5'-phosphate (PLP) and glutamate to give UDP-6-deoxy-AltNAc4N. This compound was purified by ion-exchange chromatography, and the C-4 amino group was then chemically acetylated with chloroacetic anhydride to give UDP-2-acetamido-4-chloroacetamido-2,4,6-trideoxy- β -L-altrose (UDP-6-deoxy-AltNAc4NAcCl), **29**. The UDP moiety was subsequently removed using a mild enzymatic hydrolysis catalyzed by the UDP-6-deoxy-AltDiNAc hydrolase (PseG), and 2-acetamido-4-chloroacetamido-2,4,6-trideoxy-L-altrose (6-deoxy-AltNAc4NAcCl, **30**) was obtained after purification by ion-exchange chromatography. Treatment of compound **30** with sodium azide in acetone/H₂O¹⁷⁹ generated the desired precursor 6-deoxy-AltNAc4NAz, **27**.

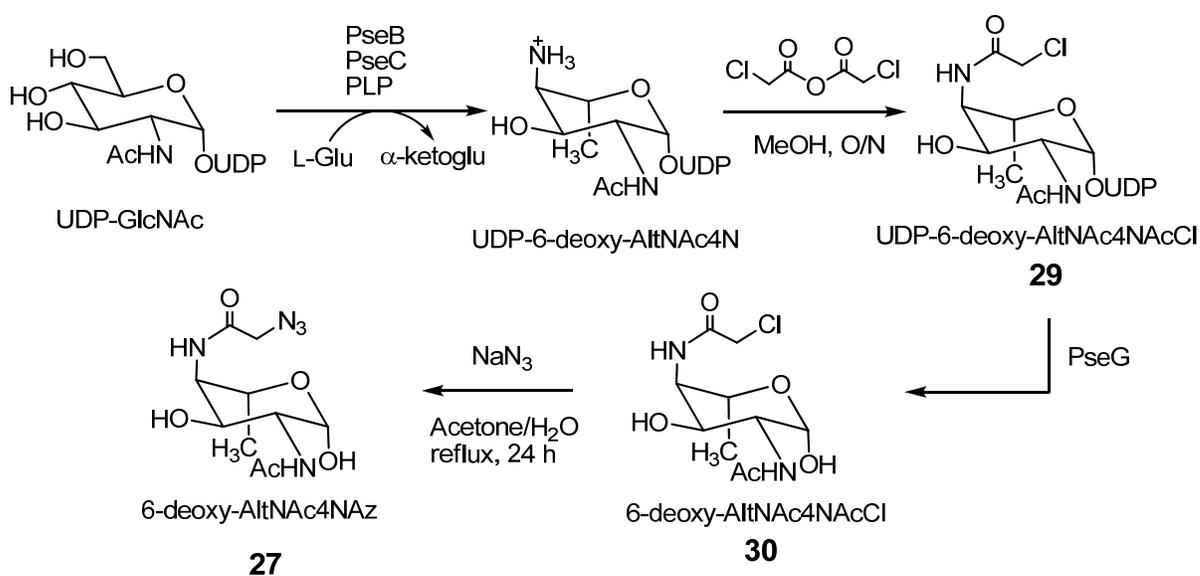


Figure 4.4 Chemo-enzymatic synthesis of 6-deoxy-AltNAc4NAz (**27**).

4.2.2 Syntheses of Staudinger Reagent 31

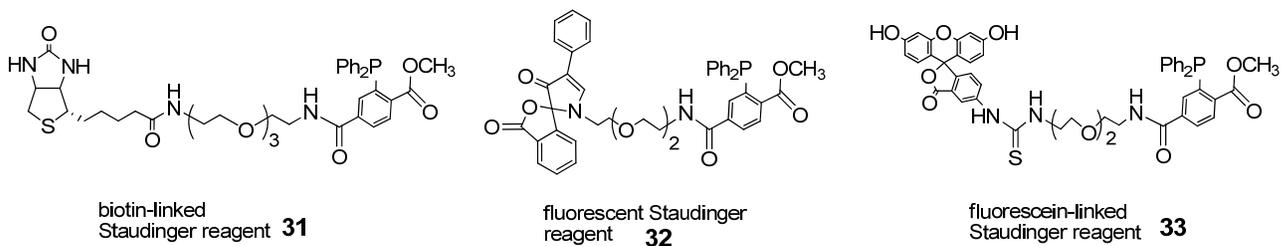


Figure 4.5 Staudinger reagents.

We decided to probe the azido functionality using the Staudinger ligation because of its mild and non-toxic conditions, its satisfactory rate, and the ease with which one can synthesize the required reagents. The three Staudinger reagents shown in Figure 4.5 all contain a 2-(diphenylphosphino) terephthalic acid 1-methyl ester motif, for reaction with the cell surface azido functionality via the Staudinger ligation, as well as a (PEO)_n linker to enhance the limited water solubility of the phosphine.¹²⁴ Compound **31** bears a biotin moiety that may be detected by a variety of avidin-containing labeling reagents. Compound **32** and **33** are fluorescent and can be directly visualized by fluorescence spectroscopy.

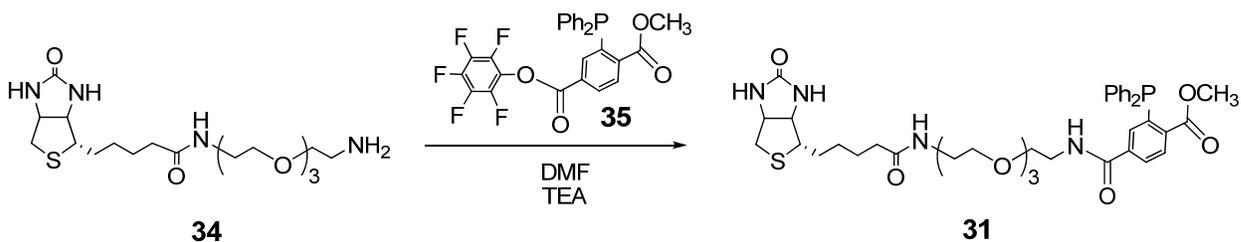


Figure 4.6 Synthesis of biotin-Staudinger reagent **31**.

Compound **31** was synthesized in a one-step reaction, which was slightly modified from that described previously by Saxon *et al.*¹²⁴, by treating (+)-biotin-(PEO)₄-amine (**34**) with commercially available 2-(diphenylphosphino) terephthalic acid 1-methyl 4-pentafluorophenyl

diester (**35**). It was purified by silica gel column chromatography and was characterized by ^1H NMR spectroscopy and ESI-mass spectrometry (m/z 787 ($\text{M}+\text{Na}^+$)). Compound **31** is soluble in phosphate buffered saline (PBS), and was found to be reasonably stable to oxidation of the phosphine group under appropriate storage conditions.

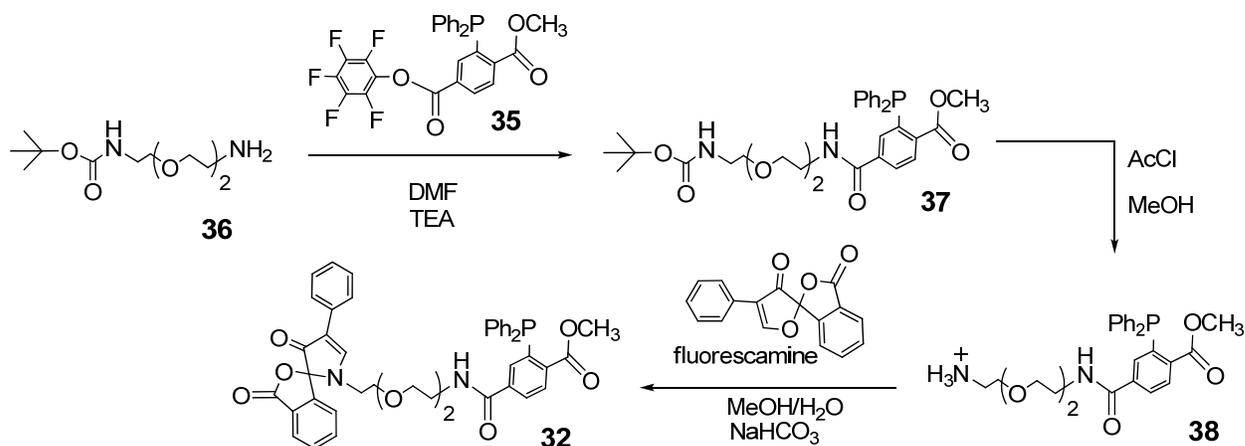


Figure 4.7 Synthesis of fluorescent Staudinger reagent **32** using fluorescamine.

The detection of an azido functionality by modification with a fluorescent Staudinger reagent and the use of fluorescence spectroscopy would be fast and convenient. Our first fluorescent Staudinger reagent, **32**, was synthesized by taking advantage of the reactivity of fluorescamine towards primary amino groups. Fluorescamine has been widely used in assays of amino acids, peptides, proteins and other primary amines in the picomolar range.¹⁸⁰ It is non-fluorescent but readily reacts with primary amines to yield highly fluorescent products ($\lambda_{\text{ex}} = 390$ nm, $\lambda_{\text{em}} = 475$ nm). To synthesize the fluorescent Staudinger reagent **32**, mono-(Boc)-(PEO)₃-diamine (**36**) was first reacted with phosphine **35** to yield compound **37**, which was subjected to methanolysis under acidic conditions to expose the terminal amino group. The resultant amine **38** was then treated with fluorescamine to give the fluorescent Staudinger reagent **32**. Unfortunately,

phosphine **32** was found to be liable to air oxidation and >50% of the material was oxidized to the corresponding phosphine oxide after purification by silica gel column chromatography.

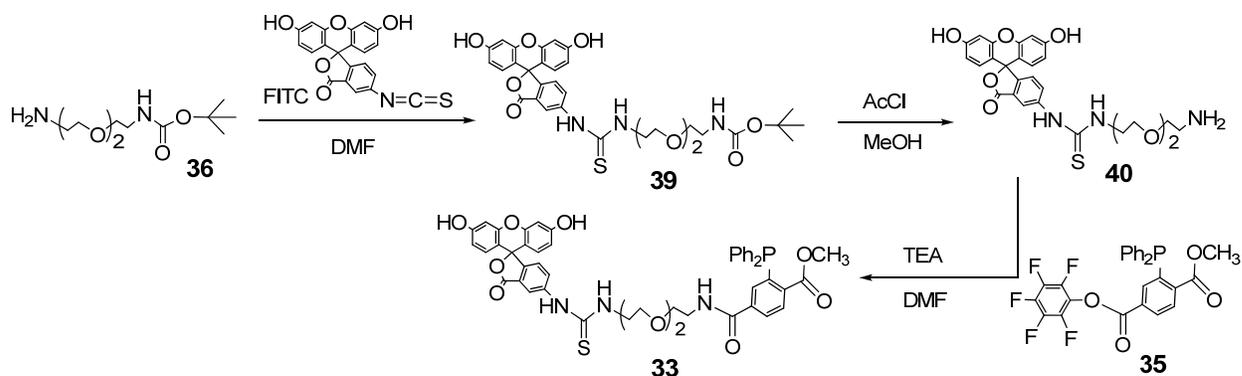


Figure 4.8 Synthesis of fluorescein-linked Staudinger reagent **33**.

Our second fluorescent Staudinger reagent, **33**, bears a fluorescein moiety in its structure. Fluorescein ($\lambda_{\text{ex}} = 492 \text{ nm}$, $\lambda_{\text{em}} = 518 \text{ nm}$) displays much stronger fluorescence than the ligation product of fluorescamine and an amine. Therefore it is expected that azido-bearing molecules labeled with reagent **33** would be easier to detect than those labeled by reagent **32**. A strategy similar to that utilized in the synthesis of reagent **32** was employed to generate fluorescein-linked Staudinger reagent **33**. Mono-(Boc)-(PEO)₃-diamine (**36**) was first treated with fluorescein isothiocyanate (FITC) to give compound **39**. Compound **40** was prepared by the deprotection of the amino-Boc group and then coupled with phosphine **35** to give the desired product **33**. However, the fluorescein-linked Staudinger reagent **33** was also found to be liable to oxidation. Hence, in the following studies, the relatively stable biotin-linked Staudinger reagent **31** was used as the major probing reagent for the detection of the azido functionality.

4.3 *In vitro* Tests

4.3.1 *In vitro* Conversion of 6-Deoxy-AltNAc4NAz to Az-Pse

It was important to demonstrate that compound **27** serves as an alternative substrate for pseudaminic acid synthase (PseI) before using it in *in vivo* studies. This was examined by incubating compound **27** with phosphoenolpyruvate (PEP), PseI, and Mn^{2+} under *in vitro* conditions (Figure 4.9). The clean production of azido-pseudaminic acid (Az-Pse, **41**) was observed. Az-Pse (**41**) was then fully characterized by 1H NMR spectroscopy (see Appendix Figure A.10) and ESI-mass spectrometry.

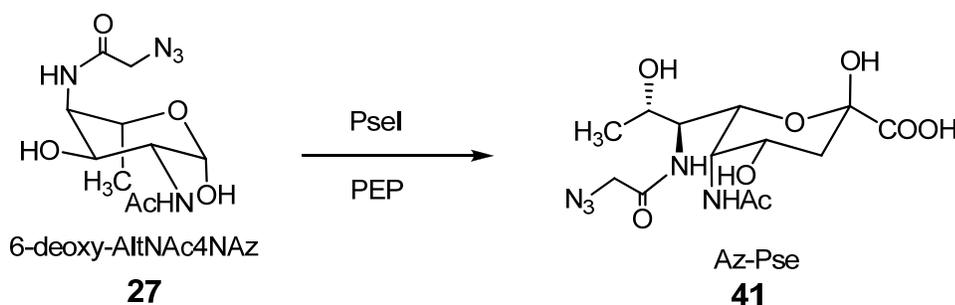


Figure 4.9 *In vitro* conversion of 6-deoxy-AltNAc4NAz (**27**) to Az-Pse (**41**).

4.3.2 *In vitro* Labeling of Az-Pse with Biotin-Staudinger Reagent (**31**)

The conditions required for the Staudinger ligation were briefly examined by incubating Az-Pse (**41**) with biotin-linked Staudinger reagent **31** in MeOH/H₂O or PBS buffer at room temperature. Rapid and clean formation of ligation product **42** was detected as shown by ESI-mass spectrometry and 1H NMR spectroscopy.

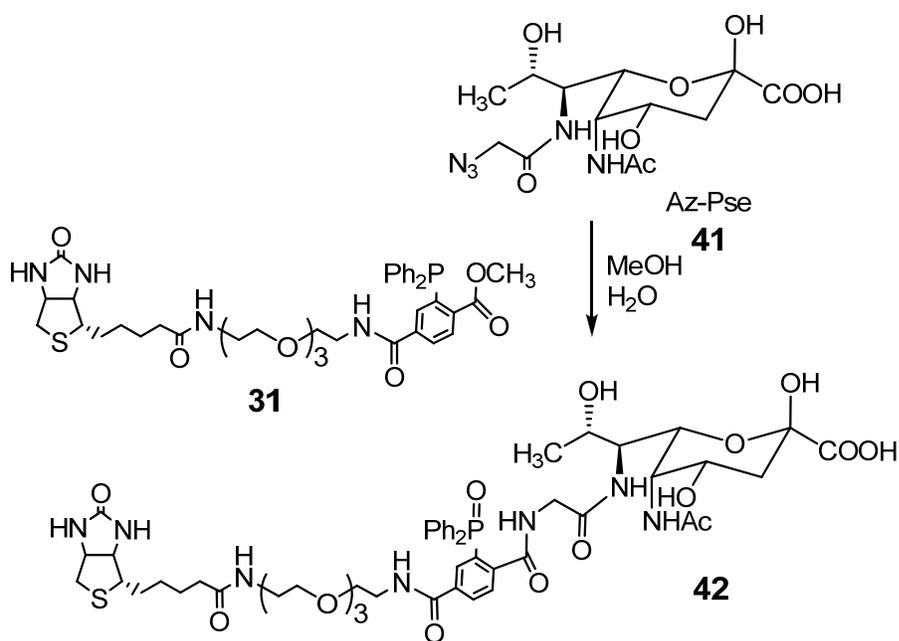


Figure 4.10 *In vitro* labeling of Az-Pse with biotin-Staudinger reagent **31**.

4.4 Motility Test^e

Propelled by its flagella, wild type *C. jejuni* 81-176 is motile and produces a spreading diffuse pattern when grown in Mueller-Hinton (MH) semi-solid motility agar¹⁸¹ (Figure 4.11A). A mutant strain *C. jejuni* 81-176 *pseG::cat*, which is unable to produce the UDP-6-deoxy-AldiNAc hydrolase PseG and is therefore deficient in the production of Pse, was employed to test for the assembly of functional flagella using a motility test. This strain is unable to assemble flagella and is completely non-motile as evidenced by the small, sharply delineated colony restricted to the centre of inoculation within the motility agar plate (Figure 4.11B).⁶⁸ However, if the azido-labeled Pse precursor, **27**, is transported into *C. jejuni* cells from the media, it should

^e This work was conducted in the Institute for Biological Sciences, NRC Canada.

be converted into azido-Pse, **41**. Thus, the mutant strain *C. jejuni* 81-176 *pseG::cat* may be able to assemble azido-labeled flagella when grown in the presence of 6-deoxy-AltNac4NAz (**27**).

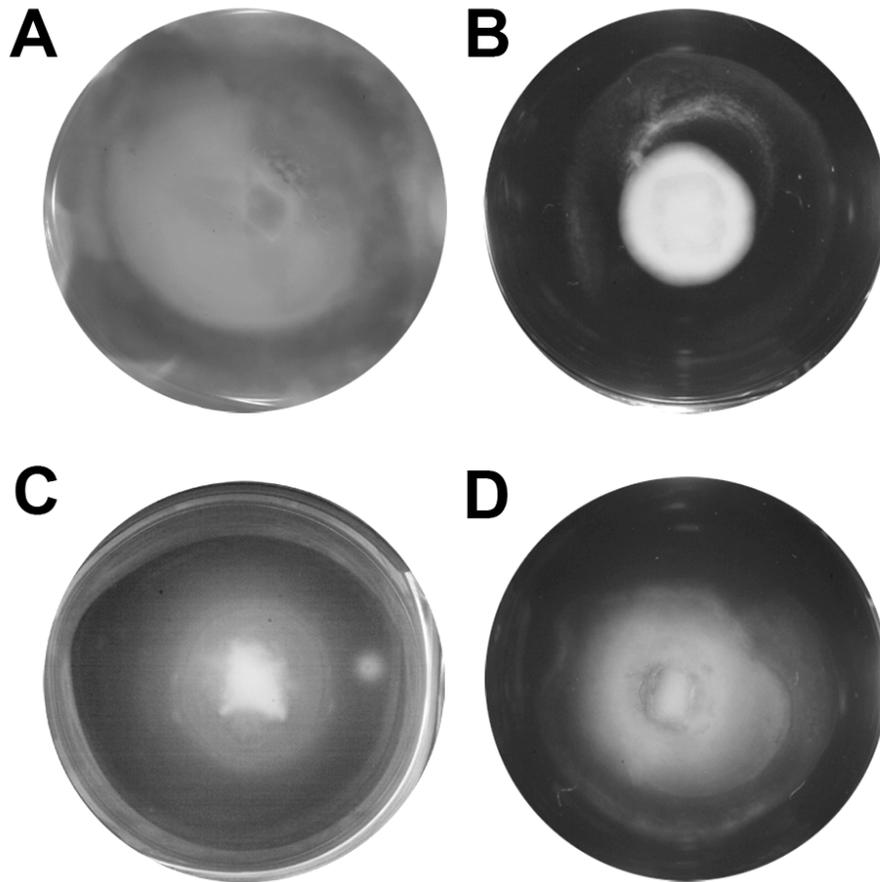


Figure 4.11 Motility of *C. jejuni* 81-176 and *C. jejuni* 81-176 *pseG::cat*. **A.** *C. jejuni* 81-176 in MH agar **B.** *C. jejuni* 81-176 *pseG::cat* in MH agar **C.** *C. jejuni* 81-176 *pseG::cat* in MH agar containing 6-deoxy-AltdiNac **D.** *C. jejuni* 81-176 *pseG::cat* in MH agar containing 6-deoxy-AltNac4NAz **27**.

The uptake of the Pse precursor from the media by *C. jejuni* cells was first tested with the natural precursor, 6-deoxy-AltdiNac. When 6-deoxy-AltdiNac (8 mM final concentration) is included in the motility agar, the motility of *C. jejuni* 81-176 *pseG::cat* is restored as evidenced by a spreading diffuse growth pattern (Figure 4.11C) similar to that shown by the wild type strain (Figure 4.11A). Although the motility was not restored to wild type levels, this study still

indicates that the Pse precursor is taken up directly from the medium, and that the uptake level is sufficient to generate functional flagella.

The same motility assay using the azido-labeled precursor, compound **27**, was employed to test the feasibility of generating azido-labeled bacteria in this manner. As is shown in Figure 4.11D, *C. jejuni* 81-176 *pseG::cat* cells that were grown on compound **27**-containing motility agar, also displayed a spreading diffuse growth pattern that is similar to that found for motile wild type *C. jejuni* 81-176. This clearly demonstrates that 6-deoxy-AltNAc4NAz (**27**) is taken up by *C. jejuni* cells and serves to generate functional flagella that confer motility. Additionally, this finding indicates that the compound is well accepted by the enzymes required to synthesize (PseI) and activate (PseF) pseudaminic acid, as well as the unidentified enzyme that transfers Pse to the flagellin monomer and the machinery that assembles the flagellin monomers into functional flagella.

4.5 Probing the Azido Functionality on the Flagella^f

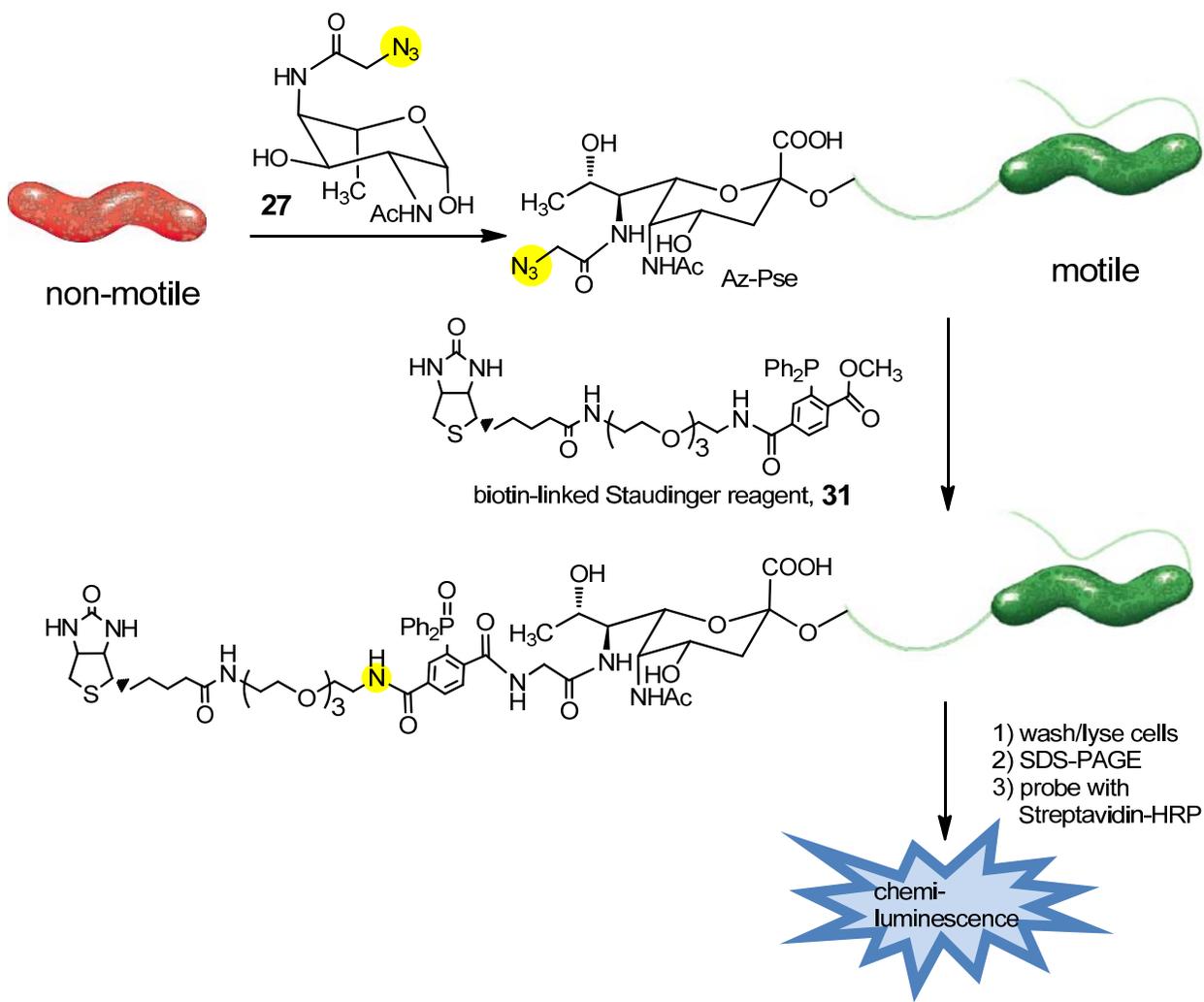


Figure 4.12 Generation of azido-pseudaminic acid-bearing bacteria and probing with biotin-linked Staudinger reagent, **31**.

In order to demonstrate that the azido functionality survived the *in vivo* conditions, and that Az-Pse was still amenable to chemical modification, the bacterial cell surface was probed with the biotin-linked Staudinger reagent, **31** (Figure 4.12).^{107, 110, 124, 125} After overnight growth on an MH agar plate containing 6-deoxy-AltNAc4NAz, **27**, (8 mM final concentration), *C. jejuni*

^f This work was conducted in the Institute for Biological Sciences, NRC Canada.

81-176 *pseG::cat* and *C. jejuni* 81-176 cells were suspended in PBS buffer and treated with biotin-linked Staudinger reagent **31** (2 mM final concentration) for 2 hours at room temperature. The cells were then harvested by centrifugation and lysed. The whole cell lysate was applied to an SDS-PAGE gel and the biotinylated proteins were detected via Western blotting using streptavidin-linked horseradish peroxidase. Streptavidin is a tetrameric protein that has an extraordinarily high affinity for biotin through non-covalent interactions. Horseradish peroxidase (HRP) oxidizes luminol to the triplet state dianion of 3-aminophthalate, which converts into its singlet state via intersystem crossing. The singlet state dianion of 3-aminophthalate relaxes to the ground state and emits energy in the form of blue light ($\lambda_{em} = 428$ nm). HRP is widely used as a reporter enzyme since it is able to amplify weak signals and increase the detectability of a target molecule.

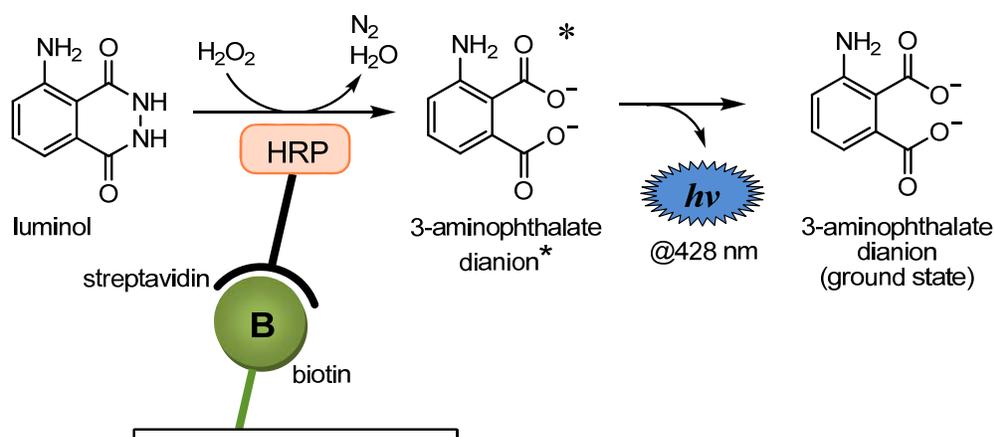


Figure 4.13 Chemiluminescence detection by streptavidin-HRP.

Using the Western blotting technique, a biotinylated protein was detected in the whole cell lysates of both mutant *C. jejuni* 81-176 *pseG::cat* and wild type *C. jejuni* 81-176 cells grown in the presence of 6-deoxy-AltNAc4NAz, **27** (Figure 4.14A, lane 1 and lane 4). No streptavidin

reactive bands were detected in control experiments where these strains were grown in the absence of compound **27** (Figure 4.14A lanes 2 and lane 3). This implies that these streptavidin reactive bands correspond to a biotin-containing protein.

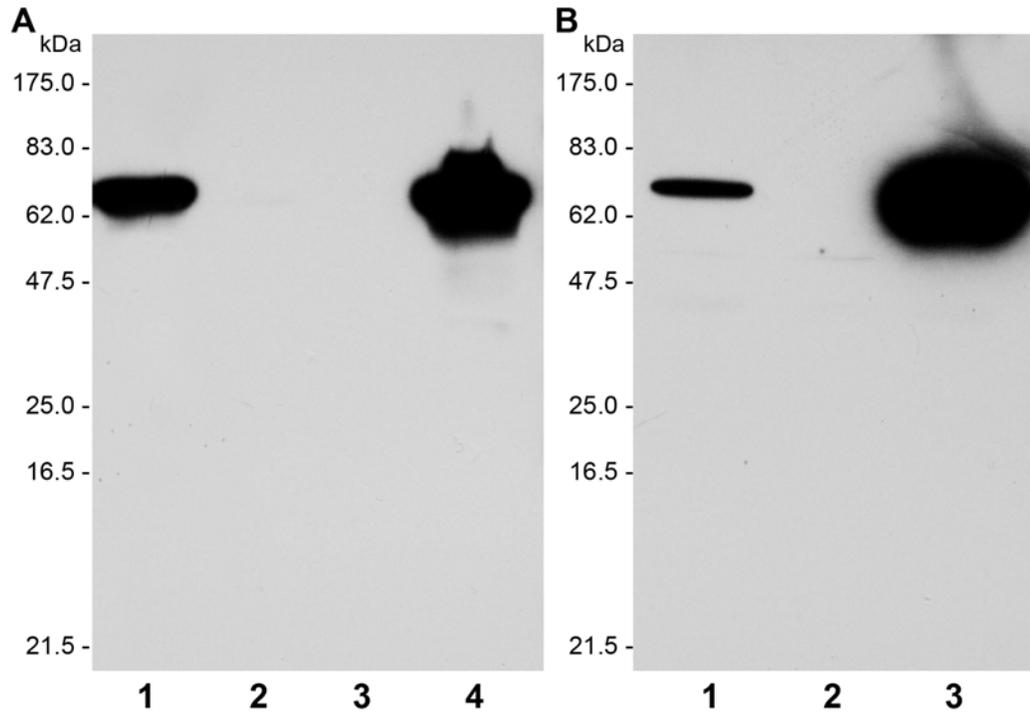


Figure 4.14 Incorporation of Az-Pse into the flagellin glycoprotein. A. Streptavidin-HRP detection of biotinylated flagellin in whole cell lysates. 1) *C. jejuni* 81-176 *pseG::cat* on MH medium containing 6-deoxy-AltNAc4NAz **27**, 2) *C. jejuni* 81-176 *pseG::cat* on MH medium, 3) *C. jejuni* 81-176 on MH medium, 4) *C. jejuni* 81-176 on MH medium containing 6-deoxy-AltNAc4NAz **27**. **B.** Flagellin detection in whole cell lysates using antiflagellin single domain antibody. 1) *C. jejuni* 81-176 *pseG::cat* on MH medium containing 6-deoxy-AltNAc4NAz **27**, 2) *C. jejuni* 81-176 *pseG::cat* on MH medium, Lane 3. *C. jejuni* 81-176 on MH medium.

To confirm that this biotin-containing protein was flagellin, the whole cell lysates were subjected to Western blot analysis probed with an antiflagellin antibody (Figure 4.14B lane 1-3).

The anti-flagellin antibody reactive bands in each strain were shown to have an identical migration pattern to the streptavidin reactive bands in Figure 4.14A.

The above experiments demonstrated that the azido functionality was still intact and was displayed on the surface of the bacterial flagella. The fact that no other streptavidin reactive proteins were found in whole cell lysates of both wild type *C. jejuni* 81-176 and mutant *C. jejuni* 81-176 *pseG::cat* grown in the presence of 6-deoxy-AltNAc4NAz (**27**) indicates that the flagellin proteins are the only proteins that are glycosylated with pseudaminic acid (Figure 4.14A, lane 1 and lane 4). The biotin-containing flagellin band of the wild type *C. jejuni* 81-176 was found to be much thicker than that of the mutant *C. jejuni* 81-176 *pseG::cat*. Since approximately equal amounts of total protein were loaded in each lane of the western blot (Figure 4.15, lane 1-4), this difference reflects the relatively larger amount of flagellin protein produced by the wild type cells when sufficient endogenous pseudaminic acid is available. The incorporation of Az-Pse in wild type *C. jejuni* 81-176 flagellin proteins clearly demonstrates that 6-deoxy-AltNAc4NAz (**27**) serves as a good substrate that competes readily with the natural substrate 6-deoxy-AltdiNAc, and that azido-derivatives are well tolerated by the enzymes involved in the CMP-Pse biosynthesis.

Attempts were also made to directly visualize the intact azido-bacteria using fluorescent Staudinger reagents **32** and **33**. However, no fluorescence-labeling trials have yet been successful, probably due to the fact that the phosphine groups in these reagents are very liable to oxidation. Visualization of the azido tag using different reagents is currently underway.

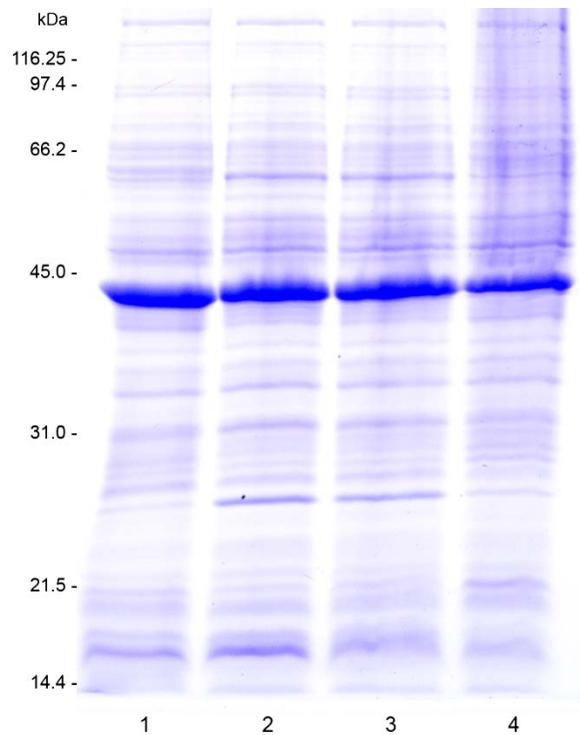


Figure 4.15 Coomassie blue-stained SDS-PAGE gel of samples used in Western blot. **1)** *C. jejuni* 81-176 grown on MH medium, **2)** *C. jejuni* 81-176 *pseG::cat* grown on MH medium, **3)** *C. jejuni* 81-176 *pseG::cat* grown on MH medium containing 6-deoxy AltNAc4NAz (**27**), **4)** *C. jejuni* 81-176 grown on MH medium containing 6-deoxy AltNAc4NAz. The major protein at 45 kDa is the major outer membrane protein (MOMP).

4.6 Conclusions

Pseudaminic acid is found as the major *O*-linked posttranslational modification of the flagellin proteins in pathogenic bacteria *H. pylori*, the causative agent of duodenal/ gastric ulcers and gastric cancer, as well as *C. jejuni*, a major cause of food poisoning in humans. It would be of particular interest to track Pse in its native environment to gain further insight into processes such as flagella assembly. In this chapter we described the first labeling of flagella with azido-pseudaminic acid using the *C. jejuni* Pse production and flagella assembly machineries.

6-Deoxy-AltNAc4NAz (**27**), an azido-derivative of the neutral Pse precursor 6-deoxy-AltNAc, was synthesized chemo-enzymatically. This compound was then fed to a mutant strain *C. jejuni* 81-176 *pseG::cat* which is unable to express PseG required for Pse biosynthesis. *C. jejuni* 81-176 *pseG::cat* is deficient in Pse production and is therefore rendered aflagellate and non-motile. However, when grown in media containing 6-deoxy-AltNAc4NAz (**27**), the Pse/flagella production and cell motility were restored as shown by a spreading diffuse growth pattern on a motility agar, which is typical of motile strains. This indicates that neutral sugars can be directly taken up by *C. jejuni* cells from the media, and that the azido-derivatives are well tolerated by the Pse production and flagella assembly machineries.

Western blots on whole cell lysates using a synthetic biotin-linked Staudinger reagent coupled with streptavidin-HRP detection confirmed that the azido functionality was still intact and was able to react with the Staudinger reagent. Parallel western blots probed with an anti-flagellin antibody indicated that only flagellin proteins were glycosylated with pseudaminic acid (Pse). The fact that the azido functionality was also incorporated into the flagella of wild type *C. jejuni* 81-176 upon incubation with media containing 6-deoxy-AltNAc4NAz (**27**) implies

that the azido derivatives, including 6-deoxy-AltNAc4NAz (**27**), Az-Pse (**41**) and the corresponding azido-derivative of CMP-Pse, serve as good substrates for the enzymes involved in Pse and flagella production, and readily compete with the natural substrates.

This chapter demonstrated the *in vivo* labeling of bacteria with bio-orthogonal azido tags using the pseudaminic biosynthetic machinery. The presence of the azido-pseudaminic acid on the surface of the flagella provides a bio-orthogonal chemical handle that not only can be used to modify the flagellar proteins, but also opens up endless possibilities for further studies.

4.7 Future Directions

The finding that wild type *C. jejuni* had been shown to take up 6-deoxy-AltNAc4NAz (**27**) and produce azido-labeled flagella at satisfactory levels suggests that wild type *C. jejuni* and *H. pylori* can be directly used in subsequent studies instead of the mutant strains. An acetylated version of 6-deoxy-AltNAc4NAz, compound **28**, can also be tested for higher levels of cellular uptake.

It would be fast and easy to track Pse and the corresponding bacteria using fluorescent probes. However, the propensity for oxidation of the fluorescent Staudinger reagents made them unsuitable for azide-labeling. Other fluorescent chemical reporters, such as cyclooctyne-linked derivatives, could be used to label the azido-functionality by using click chemistry that is independent of cytotoxic copper. Visualization of the azido-bearing flagella with gold nano particles is also currently underway.

The ability to incorporate reactive chemical handles selectively onto the flagella of pathogenic bacteria provides a useful tool for further studies. Such tags could be used for labeling of cells to be used in *in vivo* animal models of infection, biophysical studies of bacterial motility or in studies probing flagella assembly and surface accessibility of glycan modifications on flagellin proteins. Azido-labeled bacteria can also be designed as drug delivery vehicles. Engineered non-pathogenic bacteria might transport drug precursors that are linked to their flagella to the target tissue. Additionally, as flagellin proteins are strongly immunogenic, they can be appended with a variety of compounds to develop novel conjugate vaccines for small molecules.

4.8 Experimental Procedures

4.8.1 Materials and General Methods

The previous materials and general methods from Chapter 2 and 3 also apply in this study with the following additions. (+)-Biotin-(PEO)₄-amine (**34**) and mono-(Boc)-(PEO)₃-diamine (**36**) were purchased from Molecular Biosciences. Streptavidin-HRP was purchased from Cedarlane Labs (Canada). Pierce In-Gel Chemiluminescence Detection Kit was purchased from Thermo Scientific (USA).

4.8.2 Over-expression and Purification of His₆-Tagged PseB, PseC, PseG and PseI

PseB, PseC and PseG are over-expressed as previously described in Chapter 3.

The conditions used for over-expression of PseI were identical to those described for PseG with the exception that the crude cell lysate was used directly without purification by Ni-affinity chromatography.

4.8.3 Chemical Syntheses

4.8.3.1 Chemoenzymatic synthesis of UDP-2-acetamido-4-chloroacetamido-2,4,6-trideoxy-L-altrose (UDP-6-deoxy-AltNAc4NAcCl, **29**)

UDP-4-amino-2-acetamido-2,4,6-trideoxy-L-altrose (UDP-6-deoxy-AltNAc4N) was synthesized as previously described in Chapter 3. UDP-6-deoxy-AltNAc4NAcCl was synthesized using a slightly modified method from that used to generate UDP-6-deoxy-AltNAc. UDP-6-deoxy-AltNAc4N (240 mg, ditriethylammonium salt) was stirred with triethylamine (887 μ L) and chloroacetic anhydride (1.08 g) in methanol (25 mL) at room temperature for 72 h. ESI-

MS(-) showed that the starting material was completely converted to UDP-6-deoxy-AltNAc4NAcCl, **29**, (m/z 665 (M-H⁺)) during this time. After filtration and evaporation of the solvent under reduced pressure, the product was loaded onto a DE-52 anion exchange column and eluted with a linear gradient of triethylammonium bicarbonate buffer as described above. Fractions containing the product were lyophilized to dryness and passed through 5 mL of DOWEX 50WX8-200 resin (sodium form). The yield of UDP-6-deoxy-AltNAc4NAcCl (**29**) was determined to be 160 mg (disodium salt) by UV absorbance at 262 nm ($\epsilon_{262} = 9890 \text{ M}^{-1}$). The product was lyophilized twice to obtain UDP-6-deoxy-AltNAc4NAcCl (**29**) as a white solid. ¹H NMR(D₂O, pD = 6.4, 400 MHz) δ 1.31 (d, 2H, $J_{5'',6''}$ 6.2 Hz, H-6''), 2.10 (s, 3H, CH₃), 3.96 (dd, 1H, $J_{3'',4''}$ 3.0 Hz, $J_{4'',5''}$ 9.4 Hz, H-4''), 4.06 (dd, 1H, $J_{2'',3''}$ 4.6 Hz, $J_{3'',4''}$ 3.0 Hz, H-3''), 4.09 (dq, 1H, $J_{4'',5''}$ 9.4 Hz, $J_{5'',6''}$ 6.2 Hz, H-5''), 4.20 (s, 2H, Cl-CH₂), 4.21 (dd, 1H, $J_{1'',2''}$ 2.2 Hz, $J_{2'',3''}$ 4.6 Hz, H-2''), 4.27 (m, 2H, H-5'), 4.32 (m, 1H, H-4'), 4.39 (m, 1H, H-3'), 4.41 (m, 1H, H-2'), 5.69 (dd, 1H, $J_{1'',2''}$ 2.2 Hz, $J_{H,p}$ 8.4 Hz, H-1''), 6.01 (m, 2H, H-1' and H-5), 8.02 (d, 1H, $J_{5,6}$ 8.1 Hz, H-6). See Appendix Figure A.7 for detailed information. ESI-MS(-) m/z 665 (M-H⁺).

4.8.3.2 Enzymatic synthesis of 2-acetamido-4-chloroacetamido-2,4,6-trideoxy-L-altrose (6-deoxy-AltNAc4NAcCl, **30**)

A solution of UDP-6-deoxy-AltNAc4NAcCl (**29**, disodium salt, 53 mg) in sodium phosphate buffer (1 mM, 5 mL, pH = 7.5) containing PseG (2 mg) was incubated at 37 °C for 1 hr. PseG was removed by centrifugal ultrafiltration. The resulting filtrate was then passed through BIO-RAD AG 1-X8 resin (10 mL) and DOWEX 50WX8-200 resin (5 mL) eluting with distilled water. The eluant (100 mL) was collected and 6-deoxy-AltNAc4NAcCl (**30**, 16mg) was obtained as a white solid after the removal of solvent *in vacuo*. ¹H NMR (D₂O, pD = 7.2, 400 MHz), δ 1.24 (d, 3H, $J_{5\beta,6\beta}$ 6.1 Hz, H-6 β), 1.27 (d, 3H, $J_{5\alpha,6\alpha}$ 6.5 Hz, H-6 α), 2.06 (s, 3H, α -

CH₃), 2.11 (s, 3H, β-CH₃), 3.89 (dd, 1H, $J_{3,4}$ 2.8 Hz, $J_{4,5}$ 10.4 Hz, H-4), 3.98 (dd, 1H, $J_{2,3}$ 3.2 Hz, $J_{3,4}$ 2.8 Hz, H-3), 4.01 (dq, 1H, $J_{4,5}$ 10.4 Hz, $J_{5,6}$ 6.1 Hz, H-5), 4.08 (dd, 1H, $J_{1\alpha,2\alpha}$ 2.8 Hz, $J_{2\alpha,3\alpha}$ 4.3 Hz, H-2_α), 4.13 (dd, 1H, $J_{1\beta,2\beta}$ 1.8 Hz, $J_{2\beta,3\beta}$ 3.2 Hz, H-2_β), 4.18 (s, 2H, β-Cl-CH₂), 4.20 (s, 2H, α-Cl-CH₂), 5.08 (d, 3H, $J_{1\alpha,2\alpha}$ 2.8 Hz, H-1_α), 5.31 (d, 3H, $J_{1\beta,2\beta}$ 1.8 Hz, H-1_β). See Appendix Figure A.8 for detailed information. ESI-MS(+) m/z 303 (M+Na⁺).

4.8.3.3 Synthesis of 2-acetamido-4-azidoacetamido-2,4,6-trideoxy-L-altrose (6-deoxy-AltNAc4NAz, **27**)

6-Deoxy-AltNAc4NAcCl (**30**, 15 mg) was refluxed with sodium azide (17 mg) in acetone/water (15 mL/7 mL) for 24 h. After the removal of acetone at reduced pressure, the reaction mixture was passed through BIO-RAD AG 1-X8 and DOWEX 50WX8-200 resin as described above. This procedure yielded 14 mg of 6-deoxy-AltNAc4NAz (**27**). ¹H NMR (D₂O, pD = 7.2, 400 MHz), δ 1.24 (d, 3H, $J_{5\beta,6\beta}$ 6.0 Hz, H-6_β), 1.27 (d, 3H, $J_{5\alpha,6\alpha}$ 6.5 Hz, H-6_α), 2.07 (s, 3H, α-CH₃), 2.11 (s, 3H, β-CH₃), 3.91 (dd, 1H, $J_{3,4}$ 3.1 Hz, $J_{4,5}$ 10.2 Hz, H-4), 3.94 (dd, 1H, $J_{2,3}$ 3.4 Hz, $J_{3,4}$ 3.1 Hz, H-3), 3.98 (dq, 1H, $J_{4,5}$ 10.2 Hz, $J_{5\beta,6\beta}$ 6.0 Hz, H-5), 4.07 (dd, 1H, $J_{1\alpha,2\alpha}$ 2.5 Hz, $J_{2\alpha,3\alpha}$ 3.3 Hz, H-2_α), 4.08 (s, 2H, β-Az-CH₂), 4.10 (s, 2H, α-Az-CH₂), 4.14 (dd, 1H, $J_{1\beta,2\beta}$ 1.9 Hz, $J_{2\beta,3\beta}$ 3.4 Hz, H-2_β), 5.08 (d, 3H, $J_{1\alpha,2\alpha}$ 2.5 Hz, H-1_α), 5.31 (d, 3H, $J_{1\beta,2\beta}$ 1.9 Hz, H-1_β). See Appendix Figure A.9 detailed information. ¹³C NMR (D₂O, pD=7.2, 100MHz), δ 17.40, 22.13, 50.94, 51.94, 53.69, 68.22, 69.48, 90.75, 170.49, 175.01. HRMS calculated for C₁₀H₁₇N₅O₅Na (M+Na⁺) 310.1127, found 310.1120.

4.8.3.4 Enzymatic synthesis of 5-acetamido-7-azidoacetamido-3,5,7,9-tetra-deoxy-L-glycero-L-manno-nonulosonic acid (Az-Pse,41)

6-Deoxy-AltNAc4NAz (**27**, 10 mg) was added to sodium phosphate buffer (15 mM, 6 mL, pH = 8) containing the crude lysate of PseI from 150 mL of *E. coli* cell culture, phospho(enol)pyruvic acid monosodium salt hydrate (PEP, 16 mg), MgCl₂ (1.2 mg) and Na₂CO₃ (6.5 mg) and was incubated at 37 °C for 2.5 hrs. After removal of PseI by centrifugal ultrafiltration, the filtrate was loaded onto a column containing BIO-RAD AG 1-X8 resin (10 mL) and purified by stepwise elution with 0~1 M formic acid as described by Chou et al.⁸⁰ ¹H NMR (D₂O, pD = 8.2, 400 MHz), δ 1.11 (d, 3 H, *J*_{8,9} 6.5 Hz, H-9), 1.78 (dd, 1 H, *J*_{3ax,4} 12.9 Hz, *J*_{3ax,3eq} 13.2 Hz, H-3ax), 1.94 (dd, 1 H, *J*_{3eq,3ax} 13.2, *J*_{3eq,4} 4.9 Hz, H-3eq), 2.01 (s, 3H, CH₃), 3.96 (d, 1 H, *J*_{H,H} 16.4 Hz, Az-CH₂), 4.03 (d, 1 H, *J*_{H,H} 16.4 Hz, Az-CH₂), 4.10 (dd, 1 H, *J*_{6,7} 10.5 Hz, H-6), 4.15 (m, 1H, H-4), 4.16 (dq, 1H, *J*_{7,8} 3.4 Hz, *J*_{8,9} 6.5 Hz, H-8), 4.22 (dd, 1 H, *J*_{6,7} 10.5 Hz, *J*_{7,8} 3.4 Hz, H-7), 4.25 (dd, 1 H, H-5). See Appendix Figure A.10 for detailed information. ESI-MS(+) *m/z* 398 (M+Na⁺), ESI-MS(-) *m/z* 374 (M+Na⁺).

4.8.3.5 Synthesis of biotin-linked Staudinger reagent 31

Biotin-linked Staudinger reagent **31** was synthesized using a slightly modified method from that described by Saxon and Bertozzi.¹²⁴ Commercially available 2-(diphenylphosphino) terephthalic acid 1-methyl 4-pentafluorophenyl diester (26 mg, 0.049 mmol) was stirred with (+)-biotin-(PEO)₄-amine (14.3 mg, 0.049 mmol) and triethylamine (43 μL, 0.245 mmol) in 2 mL of dimethylformamide (DMF) under argon at room temperature for 1 hr. After the removal of DMF *in vacuo*, the product was purified by silica gel chromatography eluting with 1:9 MeOH:DCM (R_f = 0.2). Compound **31** was yielded as a white solid (31 mg, 79%). ¹H NMR (MeOD, 400 MHz), δ 1.41 (dt, 2 H, *J* = 7.5 Hz, 15.0 Hz), 1.53-1.76 (m, 4 H), 2.18 (t, 2 H, *J* =

7.4 Hz), 2.69 (dd, 1 H, $J = 12.7$ Hz), 2.90 (dd, 1 H, $J = 5.0$ Hz, 12.7 Hz), 3.17 (m, 1 H), 3.32 (m, 4 H), 3.48 (m, 4 H), 3.55-3.63 (m, 4H), 3.69 (s, 3 H, O-CH₃), 4.27 (dd, 1 H, $J = 5.5$ Hz, 7.8 Hz), 4.46 (dd, 1 H, $J = 4.9$ Hz, 7.8 Hz), 7.25-7.37 (m, 10 H, P-Ph), 7.46 (dd, 1 H, $J = 1.6$ Hz, 3.8 Hz), 7.80 (dd, 1 H, $J = 1.7$ Hz, 8.1 Hz), 8.05 (dd, 1 H, $J = 3.5$ Hz, 8.1 Hz). ESI-MS(+) m/z 787 (M+Na⁺).

4.8.3.6 Synthesis of fluorescent Staudinger reagent **32**

A solution of mono-(Boc)-(PEO)₃-diamine (**36**, 8 mg, 0.03 mmol), phosphine **35** (17 mg, 0.03 mmol) and triethylamine (TEA, 6 μ L, 0.03 mmol) in dimethylformamide (DMF, 2 mL) was stirred under argon at room temperature for 1 h. After the removal of DMF *in vacuo*, the resultant product **37** was re-dissolved in methanol (25 mL), to which was added acetyl chloride (AcCl, 50 μ L). The reaction mixture was stirred under argon at room temperature for 16 h. The solvent was removed *in vacuo* after the reaction had been shown to be completed by ESI-MS(+). The resulting compound **38** (14 mg, 0.026 mmol), as well as fluorescamine (7.5 mg, 0.026 mmol), were dissolved in a mixture of MeOH:0.01 M NaHCO₃ (1 mL: 3 mL). The reaction mixture was stirred under argon at room temperature for 30 min before the removal of solvent *in vacuo*. The resultant fluorescent Staudinger reagent **32** was purified by silica gel column chromatography (MeOH:DCM = 1:9, $R_f = 0.233$) to give a bright yellow solid (15 mg, 74%) that gives off a light blue fluorescence upon excitation with UV light (365 nm). ¹H NMR (CDCl₃, 400 MHz), δ 3.46-3.77 (m, 12H, N-CH₂- and CH₂-O), 3.71 (s, 3H, O-CH₃), 6.96-8.00 (m, 23H, aromatic and enamine H), 8.59 (s, 1H, NH). ESI-MS(+) m/z 795 (M+Na⁺).

4.8.3.7 Synthesis of fluorescein-linked Staudinger reagent **33**

Synthesis of fluorescein-linked (PEO)₃-amine (38) – To a DMF (2 mL) solution containing mono-(Boc)-(PEO)₃-diamine (**36**, 49.7 mg, 0.2 mmol) and fluorescein isothiocyanate (FITC, 77.5 mg, 0.2 mmol) was added triethylamine bicarbonate solution (1 M, 1 mL) and the mixture was stirred at room temperature for 1 h. After the removal of DMF and water *in vacuo*, the resultant product **39** was re-dissolved in MeOH (25 mL), to which was added acetyl chloride (AcCl, 100 μ L). The reaction mixture was stirred under argon at room temperature for 16 h. The solvent was removed *in vacuo* after the reaction had been shown to be completed by ESI-MS(+). After purification by preparative TLC (MeOH:DCM=1:9, R_f=0 (baseline)), the amine hydrochloride salt **38** was produced as a orange solid (93.6 mg, 81%). ¹H NMR (D₂O, pD = 6.4, 400MHz), δ 3.12 (t, 2 H, $J = 4.6$ Hz, CH₂-NH₃⁺), 3.68-3.80 (m, 10 H, CH₂-O and CH₂-N-C=S), 7.17 (dd, 2 H, $J = 2.2$ Hz, 9.2 Hz), 7.32 (d, 2 H, $J = 2.2$ Hz), 7.37 (d, 1 H, $J = 8.3$ Hz), 7.59 (d, 2 H, $J = 9.2$ Hz), 6.96-8.00 (m, 23H, aromatic and enamine H), 8.11 (dd, 1 H, $J = 1.6$ Hz, 8.1 Hz), 8.58 (d, 1 H, $J = 1.6$ Hz). ESI-MS(+) m/z 538 (M-Cl⁻)

*Synthesis of fluorescein-linked Staudinger reagent **33***- A DMF solution (3 mL) containing compound **38** (23 mg, 0.04 mmol), phosphine **35** (7.5 mg, 0.026 mmol) and triethylamine (TEA, 6 μ L, 0.04 mmol) was stirred under argon at room temperature for 30 min before the removal of solvent *in vacuo*. The resultant fluorescein-linked Staudinger reagent **33** was purified by silica gel column chromatography (MeOH:DCM=1:9, R_f=0.25) to give an orange solid (31 mg, 85%) that gives off a strong green fluorescence upon excitation with UV light (365 nm). ¹H NMR (D₂O, pD = 7.4, 400 MHz), δ 3.50 (s, 3 H, O-CH₃), 3.52 (t, 2 H, $J = 5.3$ Hz, CH₂-NH-C=O), 3.62-3.78 (m, 10 H, CH₂-O and CH₂-N-C=S), 6.57 (dd, 2 H, $J = 2.3$ Hz, 8.9

Hz), 6.68 (d, 2 H, $J = 2.5$ Hz), 6.85 (d, 1 H, $J = 9.0$ Hz), 7.13 (d, 2 H, $J = 8.3$ Hz), 7.50-7.68 (m, 10 H, P-Ph), 7.73 (m, 1 H), 7.95-8.02 (m, 2 H), 8.09-8.13 (m, 2 H). ESI-MS(+) m/z 906 (M+Na⁺).

4.8.3.8 Synthesis of Biotin-linked Pse (**42**)

Az-Pse (**41**, 3 mg, 8 μ mol) and a solution of biotin-linked Staudinger reagent **31** (6 mg, 8 μ mol) in MeOH (1 mL) was added to PBS buffer (3 mL) and the mixture was stirred under argon at room temperature. The progress of the reaction was monitored by TLC. It was shown that the reaction was completed in 2 h as reagent **31** was consumed (MeOH:DCM=1:1, R_f =0.76) and a new UV-active spot corresponding to biotin-linked Pse (**42**) appeared (MeOH:DCM=1:1, R_f =0.18). Compound **42** was purified by preparative TLC (MeOH:DCM=1:1, R_f =0.18) and the structure was confirmed by ¹H NMR spectroscopy and ESI mass spectrometry. ¹H NMR (D₂O, $pD = 7.6$, 400 MHz), δ 1.07 (d, 3 H, $J = 6.5$ Hz), 1.30 (dt, 2 H, $J = 7.0$ Hz, 14.4 Hz), 1.45-1.65 (m, 4 H), 1.88 (dd, 1 H, $J = 4.9$ Hz), 1.99 (dd, 1 H), 2.00 (s, 3 H, Ac), 2.17 (t, 2 H, $J = 7.3$ Hz), 2.70 (dd, 1 H, $J = 13.0$ Hz), 2.89 (dd, 1 H, $J = 4.9$ Hz, 13.0 Hz), 3.17 (m, 1 H), 3.30 (dd, 2 H, $J = 5.2$ Hz), 3.44 (dd, 2 H, $J = 6.8$ Hz, 11.4 Hz), 3.52-3.72 (m, 14 H), 3.66 (s, 3 H, O-CH₃), 4.07-4.15 (m, 4 H), 4.20 (dd, 2 H, $J = 3.8$ Hz), 4.28 (dd, 1 H, $J = 4.5$ Hz, 8.0 Hz), 4.51 (dd, 1 H, $J = 4.7$ Hz, 7.9 Hz), 7.59-7.75 (m, 10 H, P-Ph), 7.87 (dd, 1 H, $J = 3.7$ Hz, 7.9 Hz), 7.96 (dd, 1 H, $J = 1.6$ Hz, 13.6 Hz), 8.13 (m, 1 H). ESI-MS(+) m/z 787 (M+Na⁺).

4.8.4 Motility Assays

Motility assays were conducted at the Institute for Biological Sciences, National Research Council (NRC) of Canada (Ottawa, ON). Strains were grown overnight on MH agar at 37°C under microaerophilic conditions. Motility stab assays were performed in four-well tissue

culture plates (Nunc, Fisher Scientific, Ottawa, Canada). Test wells contained 0.4% semi-solid MH agar (1 mL) or 0.4% semi-solid MH agar (1 mL) containing either 6-deoxy-AltDiNAc or 6-deoxy-AltNAc4NAz, **27** (8 mM final concentration). A loopful of an overnight culture was stab inoculated into the centre of the agar in each well and then incubated for 24 h at 37°C under microaerophilic conditions. The growth pattern was observed and recorded.

4.8.5 Cell surface Tagging of Az-Pse with Staudinger Reagent, 31

Cell surface tagging was conducted in Institute for Biological Sciences, NRC Canada. *C. jejuni* 81–176 and *C. jejuni* 81–176 *pseG::cat* were inoculated onto MH agar or agar containing 6-deoxy-AltNAc4NAz (2 mM), **27** at 8 mM final concentration in 5 cm Petri plates (Nunc, Fisher Scientific). Following overnight growth at 37 °C, cells were harvested from the surface of the agar by using a loop and resuspended in PBS that contained Staudinger reagent **31** (2 mM final concentration) and reacted for 2 h at room temperature. Bacterial cells were then collected by centrifugation, washed twice in PBS and lysed in SDS-PAGE solubilisation buffer by heating to 100 °C for 10 min.

4.8.6 Streptavidin HRP Detection of Biotinylated Az-Pse

Streptavidin HRP detection was also conducted in Institute for Biological Sciences, NRC Canada. Following SDS-PAGE of whole-cell lysates on 12.5% acrylamide gels as described by Laemmli, proteins were transferred to polyvinylidene difluoride (PVDF) membrane (Biorad, Mississauga, ON, Canada) as described by Towbin et al.¹⁸² Western blotting was performed by using Streptavidin-HRP (Cedar Lane, Hornby, ON, Canada) and developed with a chemiluminescence kit (Pierce, Rockford, IL, USA) according to the manufacturer's directions. The theoretical masses of the unglycosylated proteins FlaA and FlaB are 59543 and 59704 Da, respectively.

Bibliography

1. Angata, T.; Varki, A., *Chem. Rev.* **2002**, *102* (2), 439-470.
2. Szymanski, C. M.; Wren, B. W., *Nat Rev Micro* **2005**, *3* (3), 225-237.
3. Guerry, P., *Trends Microbiol.* **2007**, *15* (10), 456-461.
4. Logan, S. M., *Microbiology* **2006**, *152* (5), 1249-1262.
5. Raetz, C. R., *J. Bacteriol.* **1993**, *175* (18), 5745-5753.
6. Herrmann, K. M.; Weaver, L. M., *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1999**, *50* (1), 473-503.
7. Albert-Weissenberger, C.; Cazalet, C.; Buchrieser, C., *Cell. Mol. Life Sci.* **2007**, *64* (4), 432-448.
8. Fields, B. S.; Benson, R. F.; Besser, R. E., *Clin. Microbiol. Rev.* **2002**, *15* (3), 506-526.
9. Swanson, M. S.; Hammer, B. K., *Annu. Rev. Microbiol.* **2000**, *54* (1), 567-613.
10. Knirel, Y. A.; Rietschel, E. T.; Marre, R.; Zähringer, U., *Eur. J. Biochem.* **1994**, *221* (1), 239-245.
11. Kooistra, O.; Lüneberg, E.; Knirel, Y. A.; Frosch, M.; Zähringer, U., *Eur. J. Biochem.* **2002**, *269* (2), 560-572.
12. McNally, D. J.; Aubry, A. J.; Hui, J. P. M.; Khieu, N. H.; Whitfield, D.; Ewing, C. P.; Guerry, P.; Brisson, J.-R.; Logan, S. M.; Soo, E. C., *J. Biol. Chem.* **2007**, *282* (19), 14463-14475.
13. Faillard, H., *Trends Biochem. Sci.* **1989**, *14* (6), 237-241.
14. Gottschalk, A., *Nature* **1955**, *176* (4488), 881-882.
15. Kuhn, R.; Brossmer, R., *Angew. Chem.* **1962**, *74* (7), 252-253.
16. Blix, F. G.; Gottschalk, A.; Klenk, E., *Nature* **1957**, *179* (4569), 1088-1088.
17. Traving, C.; Schauer, R., *Cell. Mol. Life Sci.* **1998**, *54* (12), 1330-1349.
18. Cunningham, B. A.; Hoffman, S.; Rutishauser, U.; Hemperly, J. J.; Edelman, G. M., *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80* (10), 3116-3120.

19. Kleene, R.; Schachner, M., *Nat. Rev. Neurosci.* **2004**, 5 (3), 195-208.
20. Varki, A., *FASEB J.* **1997**, 11 (4), 248-255.
21. Lasky, L. A., *Annu. Rev. Biochem.* **1995**, 64 (1), 113-140.
22. Vestwever, D.; Blanks, J. E., *Physiol. Rev.* **1999**, 79 (1), 181-213.
23. Kim, Y. J.; Borsig, L.; Varki, N. M.; Varki, A., *Proc. Nat. Acad. Sci. U.S.A.* **1998**, 95 (16), 9325-9330.
24. Ugorski, M.; Laskowska, A., *Acta Biochim. Pol.* **2002**, 49 (2), 303-310.
25. Fukuda, M., *Cancer. Res.* **1996**, 56 (10), 2237-2244.
26. Kazatchkine, M. D.; Fearon, D. T.; Austen, K. F., *J. Immunol.* **1979**, 122 (1), 75-81.
27. Shu Li, X.; Min, Z.; Fang Cheng, C.; Bing, D.; Xiao Ping, Z., *J. Neuroimmunol.* **2006**, 174 (1), 126-132.
28. Vogel, U.; Hammerschmidt, S.; Frosch, M., *Med. Microbiol. Immunol.* **1996**, 185 (2), 81-87.
29. Preston, A.; Mandrell, R. E.; Gibson, B. W.; Apicella, M. A., *Crit. Rev. Microbiol.* **1996**, 22 (3), 139-180.
30. Robbins, J. B.; McCracken, G. H., Jr.; Gotschlich, E. C.; Orskov, F.; Orskov, I.; Hanson, L. A., *N. Engl. J. Med.* **1974**, 290, 1216.
31. Timmis, K. N.; Boulnois, G. J.; Bitter-Suermann, D.; Cabello, F. C., *Curr. Top. Microbiol. Immunol.* **1985**, 118, 197.
32. Moxon, E. R.; Kroll, J. S., *Curr. Top. Microbiol. Immunol.* **1990**, 150, 65.
33. Toshiyuki Hamada; Yoko Ito; Takamasa Abe; Fumiaki Hayashi; Peter Güntert; Makoto Inoue; Takanori Kigawa; Takaho Terada; Mikako Shirouzu; Mayumi Yoshida; Akiko Tanaka; Sumio Sugano; Shigeyuki Yokoyama; Hiroshi Hirota, *Protein Sci.* **2006**, 15 (5), 1010-1016.
34. Campbell, R. E.; Mosimann, S. C.; Tanner, M. E.; Strynadka, N. C. J., *Biochemistry* **2000**, 39 (49), 14993-15001.
35. Chou, W. K.; Hinderlich, S.; Reutter, W.; Tanner, M. E., *J. Am. Chem. Soc.* **2003**, 125 (9), 2455-2461.
36. Murkin, A. S.; Chou, W. K.; Wakarchuk, W. W.; Tanner, M. E., *Biochemistry* **2004**, 43 (44), 14290-14298.

37. Sommar, K. M.; Ellis, D. B., *Biochimica et Biophysica Acta (BBA) - Enzymology* **1972**, 268 (2), 590-595.
38. Legler, P. M.; Massiah, M. A.; Bessman, M. J.; Mildvan, A. S., *Biochemistry* **2000**, 39 (29), 8603-8608.
39. Glaser, L., *Biochim. Biophys. Acta* **1960**, 41 (3), 534-536.
40. Comb, D. G.; Roseman, S., *Biochim. Biophys. Acta* **1958**, 29 (3), 653-654.
41. Morgan, P. M.; Sala, R. F.; Tanner, M. E., *J. Am. Chem. Soc.* **1997**, 119 (43), 10269-10277.
42. Gunawan, J.; Simard, D.; Gilbert, M.; Lovering, A. L.; Wakarchuk, W. W.; Tanner, M. E.; Strynadka, N. C. J., *J. Biol. Chem.* **2005**, 280 (5), 3555-3563.
43. Baardsnes, J.; Davies, P. L., *Trends Biochem. Sci.* **2001**, 26 (8), 468-469.
44. Hamada, T.; Ito, Y.; Abe, T.; Hayashi, F.; Güntert, P.; Inoue, M.; Kigawa, T.; Terada, T.; Shirouzu, M.; Yoshida, M.; Tanaka, A.; Sugano, S.; Yokoyama, S.; Hirota, H., *Protein Sci.* **2006**, 15 (5), 1010-1016.
45. DeLeo, A. B.; Sprinson, D. B., *Biochem. Biophys. Res. Commun.* **1968**, 32 (5), 873-877.
46. Hedstrom, L.; Abeles, R., *Biochem. Biophys. Res. Commun.* **1988**, 157 (2), 816-820.
47. Sundaram, A. K.; Pitts, L.; Muhammad, K.; Wu, J.; Betenbaugh, M.; Woodard, R. W.; Vann, W. F., *Biochem. J.* **2004**, 383 (1), 83-89.
48. Du, S.; Faiger, H.; Belakhov, V.; Baasov, T., *Bioorg. Med. Chem.* **1999**, 7 (12), 2671-2682.
49. Du, S.; Tsipori, H.; Baasov, T., *Bioorg. Med. Chem. Lett.* **1997**, 7 (19), 2469-2472.
50. Asojo, O.; Friedman, J.; Adir, N.; Belakhov, V.; Shoham, Y.; Baasov, T., *Biochemistry* **2001**, 40 (21), 6326-6334.
51. Wang, J.; Duewel, H. S.; Woodard, R. W.; Gatti, D. L., *Biochemistry* **2001**, 40 (51), 15676-15683.
52. Kaustov, L.; Kababya, S.; Belakhov, V.; Baasov, T.; Shoham, Y.; Schmidt, A., *J. Am. Chem. Soc.* **2003**, 125 (15), 4662-4669.
53. Belakhov, V.; Dovgolevsky, E.; Rabkin, E.; Shulami, S.; Shoham, Y.; Baasov, T., *Carbohydr. Res.* **2004**, 339 (2), 385-392.
54. Li, Z.; Sau, A. K.; Shen, S.; Whitehouse, C.; Baasov, T.; Anderson, K. S., *J. Am. Chem. Soc.* **2003**, 125 (33), 9938-9939.

55. Grison, C.; Petek, S.; Finance, C.; Coutrot, P., *Carbohydr. Res.* **2005**, *340* (4), 529-537.
56. Knirel, Y. A.; Vinogradov, E. V.; L'Vov, V. L.; Kocharova, N. A.; Shashkov, A. S.; Dmitriev, B. A.; Kochetkov, N. K., *Carbohydr. Res.* **1984**, *133* (2), C5-C8.
57. Kiss, E.; Kereszt, A.; Barta, F.; Stephens, S.; Reuhs, B. L.; Kondorosi, A.; Putnoky, P., *Mol. Plant-Microbe Interact.* **2001**, *14* (12), 1395-1403.
58. Castric, P.; Cassels, F. J.; Carlson, R. W., *J. Biol. Chem.* **2001**, *276* (28), 26479-26485.
59. Thibault, P.; Logan, S. M.; Kelly, J. F.; Brisson, J.-R.; Ewing, C. P.; Trust, T. J.; Guerry, P., *J. Biol. Chem.* **2001**, *276* (37), 34862-34870.
60. Schirm, M.; Soo, E. C.; Aubry, A. J.; Austin, J.; Thibault, P.; Logan, S. M., *Mol. Microbiol.* **2003**, *48* (6), 1579-1592.
61. Logan, S. M.; Trust, T. J.; Guerry, P., *J. Bacteriol.* **1989**, *171* (6), 3031-3038.
62. Doig, P.; Trust, T. J., *Infect. Immun.* **1994**, *62* (10), 4526-4533.
63. Linton, D.; Karlyshev, A. V.; Hitchen, P. G.; Morris, H. R.; Dell, A.; Gregson, N. A.; Wren, B. W., *Mol. Microbiol.* **2000**, *35* (5), 1120-1134.
64. Yuki, N., *Curr. Opin. Immunol.* **2005**, *17* (6), 577-582.
65. Logan, S. M.; Kelly, J. F.; Thibault, P.; Ewing, C. P.; Guerry, P., *Mol. Microbiol.* **2002**, *46* (2), 587-597.
66. Amsterdam, K. v.; Vliet, A. H. M. v.; Kusters, J. G.; Ende, A. v. d., *FEMS Microbiol. Rev.* **2006**, *30* (1), 131-156.
67. van Amsterdam, K.; van Vliet, A. H. M.; Kusters, J. G.; van der Ende, A., *FEMS Microbiol. Rev.* **2006**, *30* (1), 131-156.
68. Guerry, P.; Ewing, C. P.; Schirm, M.; Lorenzo, M.; Kelly, J.; Pattarini, D.; Majam, G.; Thibault, P.; Logan, S., *Mol. Microbiol.* **2006**, *60* (2), 299-311.
69. Guerry, P.; Szymanski, C. M.; Prendergast, M. M.; Hickey, T. E.; Ewing, C. P.; Pattarini, D. L.; Moran, A. P., *Infect. Immun.* **2002**, *70* (2), 787-793.
70. Misawa, N.; Blaser, M. J., *Infect. Immun.* **2000**, *68* (11), 6168-6175.
71. Sherlock, O.; Vejborg, R. M.; Klemm, P., *Infect. Immun.* **2005**, *73* (4), 1954-1963.
72. Skurnik, M.; Bolin, I.; Heikkinen, H.; Piha, S.; Wolf-Watz, H., *J. Bacteriol.* **1984**, *158* (3), 1033-1036.
73. Parkhill, J.; Wren, B. W.; Mungall, K.; Ketley, J. M.; Churcher, C.; Basham, D.; Chillingworth, T.; Davies, R. M.; Feltwell, T.; Holroyd, S.; Jagels, K.; Karlyshev, A. V.;

- Moule, S.; Pallen, M. J.; Penn, C. W.; Quail, M. A.; Rajandream, M. A.; Rutherford, K. M.; van Vliet, A. H. M.; Whitehead, S.; Barrell, B. G., *Nature* **2000**, *403* (6770), 665-668.
74. Goon, S.; Kelly, J. F.; Logan, S. M.; Ewing, C. P.; Guerry, P., *Mol. Microbiol.* **2003**, *50* (2), 659-671.
75. McNally, D. J.; Hui, J. P. M.; Aubry, A. J.; Mui, K. K. K.; Guerry, P.; Brisson, J.-R.; Logan, S. M.; Soo, E. C., *J. Biol. Chem.* **2006**, *281* (27), 18489-18498.
76. Schoenhofen, I. C.; McNally, D. J.; Brisson, J.-R.; Logan, S. M., *Glycobiology* **2006**, *16* (9), 8C-14.
77. Creuzenet, C., *FEBS Lett.* **2004**, *559* (1-3), 136-140.
78. Obhi, R. K.; Creuzenet, C., *J. Biol. Chem.* **2005**, *280* (21), 20902-20908.
79. Schoenhofen, I. C.; Lunin, V. V.; Julien, J.-P.; Li, Y.; Ajamian, E.; Matte, A.; Cygler, M.; Brisson, J.-R.; Aubry, A.; Logan, S. M.; Bhatia, S.; Wakarchuk, W. W.; Young, N. M., *J. Biol. Chem.* **2006**, *281* (13), 8907-8916.
80. Chou, W. K.; Dick, S.; Wakarchuk, W. W.; Tanner, M. E., *J. Biol. Chem.* **2005**, *280* (43), 35922-35928.
81. Tsien, R. Y., *Annu. Rev. Biochem.* **1998**, *67* (1), 509-544.
82. Wang, Y.; Shyy, J. Y.-J.; Chien, S., *Annu. Rev. Biomed. Eng.* **2008**, *10* (1), 1-38.
83. Dunder, M.; Hoffmann-Rohrer, U.; Hu, Q.; Grummt, I.; Rothblum, L. I.; Phair, R. D.; Misteli, T., *Science* **2002**, *298* (5598), 1623-1626.
84. Hadjantonakis, A.-K.; Dickinson, M. E.; Fraser, S. E.; Papaioannou, V. E., *Nat. Rev. Genet.* **2003**, *4* (8), 613-625.
85. Presley, J. F.; Cole, N. B.; Schroer, T. A.; Hirschberg, K.; Zaal, K. J. M.; Lippincott-Schwartz, J., *Nature* **1997**, *389* (6646), 81-85.
86. Weissleder, R.; Ntziachristos, V., *Nat. Med.* **2003**, *9* (1), 123-128.
87. Miyawaki, A.; Tsien, R. Y.; Jeremy Thorner, S. D. E.; John, N. A., Monitoring protein conformations and interactions by fluorescence resonance energy transfer between mutants of green fluorescent protein. In *Methods in Enzymology*, Academic Press: 2000; Vol. Volume 327, pp 472-500.
88. Truong, K.; Sawano, A.; Mizuno, H.; Hama, H.; Tong, K. I.; Mal, T. K.; Miyawaki, A.; Ikura, M., *Nat. Struct. Mol. Biol.* **2001**, *8* (12), 1069-1073.
89. Luo, K. Q.; Yu, V. C.; Pu, Y.; Chang, D. C., *Biochem. Biophys. Res. Commun.* **2001**, *283* (5), 1054-1060.

90. Mahajan, N. P.; Corinne Harrison-Shostak, D.; Michaux, J.; Herman, B., *Chem. Biol.* **1999**, *6* (6), 401-409.
91. Hudson, P. J.; Souriau, C., *Nat. Med.* **2003**, *9* (1), 129-134.
92. Mehren, M. v.; Adams, G. P.; Weiner, L. M., *Annu. Rev. Med.* **2003**, *54* (1), 343-369.
93. Sidhu, S. S.; Fellouse, F. A., *Nat. Chem. Biol.* **2006**, *2* (12), 682-688.
94. Prescher, J. A.; Bertozzi, C. R., *Nat. Chem. Biol.* **2005**, *1* (1), 13-21.
95. Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R., *Science* **1997**, *276* (5315), 1125-1128.
96. Yarema, K. J.; Mahal, L. K.; Bruehl, R. E.; Rodriguez, E. C.; Bertozzi, C. R., *J. Biol. Chem.* **1998**, *273* (47), 31168-31179.
97. Sadamoto, R.; Niikura, K.; Ueda, T.; Monde, K.; Fukuhara, N.; Nishimura, S.-I., *J. Am. Chem. Soc.* **2004**, *126* (12), 3755-3761.
98. Zhang, Z.; Smith, B. A. C.; Wang, L.; Brock, A.; Cho, C.; Schultz, P. G., *Biochemistry* **2003**, *42* (22), 6735-6746.
99. Chen, I.; Howarth, M.; Lin, W.; Ting, A. Y., *Nat. Meth.* **2005**, *2* (2), 99-104.
100. Rush, J. S.; Bertozzi, C. R., *J. Am. Chem. Soc.* **2008**, *130* (37), 12240-12241.
101. Berteau, O.; Guillot, A.; Benjdia, A.; Rabot, S., *J. Biol. Chem.* **2006**, *281* (32), 22464-22470.
102. Griffin, A. B.; Adams, S. R.; Jones, J.; Tsien, R. Y., Fluorescent labeling of recombinant proteins in living cells with FIAsh. In *Methods in Enzymology*, Academic Press: 2000; Vol. Volume 327, pp 565-578.
103. Griffin, A. B.; Adams, S. R.; Tsien, R. Y., *Science* **1998**, *281* (5374), 269-272.
104. Adams, S. R.; Campbell, R. E.; Gross, L. A.; Martin, B. R.; Walkup, G. K.; Yao, Y.; Llopis, J.; Tsien, R. Y., *J. Am. Chem. Soc.* **2002**, *124* (21), 6063-6076.
105. Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R., *ACS Chem. Biol.* **2006**, *1* (10), 644-648.
106. Saxon, E.; Armstrong, J. I.; Bertozzi, C. R., *Org. Lett.* **2000**, *2* (14), 2141-2143.
107. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R., *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (1), 19-24.
108. Link, A. J.; Tirrell, D. A., *J. Am. Chem. Soc.* **2003**, *125* (37), 11164-11165.

109. Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G., *J. Am. Chem. Soc.* **2002**, *124* (31), 9026-9027.
110. Tsao, M.-L.; Tian, F.; Schultz, P. G., *ChemBioChem* **2005**, *6* (12), 2147-2149.
111. Kho, Y.; Kim, S. C.; Jiang, C.; Barma, D.; Kwon, S. W.; Cheng, J.; Jaunbergs, J.; Weinbaum, C.; Tamanoi, F.; Falck, J.; Zhao, Y., *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (34), 12479-12484.
112. Weller, R. L.; Rajski, S. R., *Org. Lett.* **2005**, *7* (11), 2141-2144.
113. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., *Angew. Chem., Int. Ed.* **2001**, *40* (11), 2004-2021.
114. Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G., *J. Am. Chem. Soc.* **2003**, *125* (11), 3192-3193.
115. Link, A. J.; Vink, M. K. S.; Tirrell, D. A., *J. Am. Chem. Soc.* **2004**, *126* (34), 10598-10602.
116. Seo, T. S.; Bai, X.; Ruparel, H.; Li, Z.; Turro, N. J.; Ju, J., *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (15), 5488-5493.
117. Speers, A. E.; Cravatt, B. F., *ChemBioChem* **2004**, *5* (1), 41-47.
118. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R., *J. Am. Chem. Soc.* **2004**, *126* (46), 15046-15047.
119. Baskin, J. M.; Bertozzi, C. R., *QSAR Comb. Sci.* **2007**, *26* (11-12), 1211-1219.
120. Speers, A. E.; Cravatt, B. F., *Chem. Biol.* **2004**, *11* (4), 535-546.
121. Staudinger, H.; Meyer, J., *Helv. Chim. Acta* **1919**, *2* (1), 635-646.
122. Lin, F. L.; Hoyt, H. M.; van Halbeek, H.; Bergman, R. G.; Bertozzi, C. R., *J. Am. Chem. Soc.* **2005**, *127* (8), 2686-2695.
123. Sarkar, A. K.; Fritz, T. A.; Taylor, W. H.; Esko, J. D., *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (8), 3323-3327.
124. Saxon, E.; Bertozzi, C. R., *Science* **2000**, *287* (5460), 2007-2010.
125. Vocadlo, D. J.; Hang, H. C.; Kim, E.-J.; Hanover, J. A.; Bertozzi, C. R., *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (16), 9116-9121.
126. de Boer, H. A.; Comstock, L. J.; Vasser, M., *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80* (1), 21-25.
127. Franke, F. P.; Kapuscinski, M.; Lyndon, P., *Carbohydr. Res.* **1985**, *143*, 69-76.

128. Weitz, D. J.; Bednarski, M. D., *J. Org. Chem.* **1989**, *54* (20), 4957-4959.
129. Severin, T.; Lerche, H., *Synthesis* **1982**, 1982 (04), 305-307.
130. Ryu, I.; Kuriyama, H.; Miyazato, H.; Minakata, S.; Komatsu, M.; Yoon, J.-Y.; Kim, S., *Bull. Chem. Soc. Jpn.* **2004**, *77* (7), 1407-1408.
131. Li, C. J.; Chan, T. H., *Tetrahedron Lett.* **1991**, *32* (48), 7017-7020.
132. Einhorn, C.; Luche, J.-L., *J. Organomet. Chem.* **1987**, *322* (2), 177-183.
133. Nokami, J.; Otera, J.; Sudo, T.; Okawara, R., *Organometallics* **2002**, *2* (1), 191-193.
134. Miyamoto, H.; Daikawa, N.; Tanaka, K., *Tetrahedron Lett.* **2003**, *44* (36), 6963-6964.
135. Chan, T.-H.; Lee, M.-C., *J. Org. Chem.* **1995**, *60* (13), 4228-4232.
136. Chan, T.-H.; Li, C.-J., *J. Chem. Soc., Chem. Commun.* **1992**, (10), 747-748.
137. Gordon, D. M.; Whitesides, G. M., *J. Org. Chem.* **1993**, *58* (27), 7937-7938.
138. Choi, S.-K.; Lee, S.; Whitesides, G. M., *J. Org. Chem.* **1996**, *61* (25), 8739-8745.
139. Paquette, L. A.; Mitzel, T. M.; Isaac, M. B.; Crasto, C. F.; Schomer, W. W., *J. Org. Chem.* **1997**, *62* (25), 8960-8960.
140. Vorwerk, S.; Vasella, A., *Angew. Chem., Int. Ed.* **1998**, *37* (12), 1732-1734.
141. Luo, Y.; Zechel, D. L., *Can. J. Chem.* **2006**, *84*, 743-747.
142. Webb, M. R., *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (11), 4884-4887.
143. Tovar-Méndez, A.; Rodríguez-Sotres, R.; López-Valentín, D. M.; Muñoz-Clares, R. A., *Biochem. J.* **1998**, *332* (3), 633-642.
144. Tovar-Méndez, A.; Muñoz-Clares, R. A., *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **2001**, *1546* (1), 242-252.
145. An, M.; Maitra, U.; Neidlein, U.; Bartlett, P. A., *J. Am. Chem. Soc.* **2003**, *125* (42), 12759-12767.
146. Bradford, M. M., *Anal. Biochem.* **1976**, *72*, 248-254.
147. Laemmli, U. K., *Nature* **1970**, *227*, 680-685.
148. Leslie, A. G. W., *Joint CCP4+ESF-EAMCB Newsletter on Protein Crystallography* **1992**, 26.

149. Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2003**, *59* (7), 1131-1137.
150. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53* (3), 240-255.
151. Schüttelkopf, A. W.; van Aalten, D. M. F., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (8), 1355-1363.
152. Liav, A.; Sharon, N., *Carbohydr. Res.* **1973**, *30* (1), 109-126.
153. Josenhans, C.; Vossebein, L.; Friedrich, S.; Suerbaum, S., *FEMS Microbiol. Lett.* **2002**, *210* (2), 165-172.
154. Schoenhofen, I. C.; McNally, D. J.; Vinogradov, E.; Whitfield, D.; Young, N. M.; Dick, S.; Wakarchuk, W. W.; Brisson, J.-R.; Logan, S. M., *J. Biol. Chem.* **2006**, *281* (2), 723-732.
155. Fox, J. D.; Kapust, R. B.; Waugh, D. S., *Protein Sci.* **2001**, *10* (3), 622-630.
156. Kapust, R. B.; Waugh, D. S., *Protein Sci.* **1999**, *8* (8), 1668-1674.
157. Richarme, G.; Caldas, T. D., *J. Biol. Chem.* **1997**, *272* (25), 15607-15612.
158. Schoenhofen, I. C.; Vinogradov, E.; Whitfield, D. M.; Brisson, J.-R.; Logan, S. M., *Glycobiology* **2009**, *19* (7), 715-725.
159. Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B., *J. Mol. Biol.* **2003**, *328* (2), 307-317.
160. Hu, Y.; Chen, L.; Ha, S.; Gross, B.; Falcone, B.; Walker, D.; Mokhtarzadeh, M.; Walker, S., *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (3), 845-849.
161. Wrabl, J. O.; Grishin, N. V., *J. Mol. Biol.* **2001**, *314* (3), 365-374.
162. Altschul, S.; Madden, T.; Schaffer, A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D., *Nucl. Acids Res.* **1997**, *25* (17), 3389-3402.
163. Breton, C.; Snajdrova, L.; Jeanneau, C.; Koca, J.; Imberty, A., *Glycobiology* **2006**, *16* (2), 29R-37.
164. Chiu, C. P. C.; Watts, A. G.; Lairson, L. L.; Gilbert, M.; Lim, D.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C. J., *Nat. Struct. Mol. Biol.* **2004**, *11* (2), 163-170.
165. Chen, L.; Men, H.; Ha, S.; Ye, X.-Y.; Brunner, L.; Hu, Y.; Walker, S., *Biochemistry* **2002**, *41* (21), 6824-6833.
166. Losey, H. C.; Peczuh, M. W.; Chen, Z.; Eggert, U. S.; Dong, S. D.; Pelczar, I.; Kahne, D.; Walsh, C. T., *Biochemistry* **2001**, *40* (15), 4745-4755.

167. Ciesla, W. P.; Bobak, D. A., *J. Biol. Chem.* **1998**, 273 (26), 16021-16026.
168. Lairson, L. L.; Withers, S. G., *Chem. Commun. (Cambridge, U. K.)* **2004**, (20), 2243 - 2248.
169. Zhang, Y.; Wang, P. G.; Brew, K., *J. Biol. Chem.* **2001**, 276 (15), 11567-11574.
170. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G., *Annu. Rev. Biochem.* **2008**, 77 (1), 521-555.
171. Kachmar, J. F.; Boyer, P. D., *J. Biol. Chem.* **1953**, 200 (2), 669-682.
172. Rangarajan, E. S.; Proteau, A.; Cui, Q.; Logan, S. M.; Potetinova, Z.; Whitfield, D.; Purisima, E. O.; Cygler, M.; Matte, A.; Sulea, T.; Schoenhofen, I. C., *J. Biol. Chem.* **2009**, 284 (31), 20989-21000.
173. Tumbale, P.; Brew, K., *J. Biol. Chem.* **2009**, 284 (37), 25126-25134.
174. Los, J. M.; Simpson, L. B.; Wiesner, K., *J. Am. Chem. Soc.* **1956**, 78 (8), 1564-1568.
175. Mildvan, A. S.; Xia, Z.; Azurmendi, H. F.; Saraswat, V.; Legler, P. M.; Massiah, M. A.; Gabelli, S. B.; Bianchet, M. A.; Kang, L. W.; Amzel, L. M., *Arch. Biochem. Biophys.* **2005**, 433 (1), 129-143.
176. Glaze, P. A.; Watson, D. C.; Young, N. M.; Tanner, M. E., *Biochemistry* **2008**, 47 (10), 3272-3282.
177. Saxon, E.; Luchansky, S. J.; Hang, H. C.; Yu, C.; Lee, S. C.; Bertozzi, C. R., *J. Am. Chem. Soc.* **2002**, 124 (50), 14893-14902.
178. Barton, L. L., *Structural and Functional Relationships in Prokaryotes*. Springer: 2005; p 468.
179. Aich, U.; Loganathan, D., *J. Carbohydr. Chem.* **2005**, 24 (1), 1 - 12.
180. Udenfriend, S.; Stein, S.; Bohlen, P.; Dairman, W.; Leimgruber, W.; Weigele, M., *Science* **1972**, 178 (4063), 871-872.
181. Tittsler, R. P.; Sandholzer, L. A., *J. Bacteriol.* **1936**, 31 (6), 575-580.
182. Towbin, H.; Staehelin, T.; Gordon, J., *Proc. Natl. Acad. Sci. U. S. A.* **1979**, 76 (9), 4350-4354.

Appendix

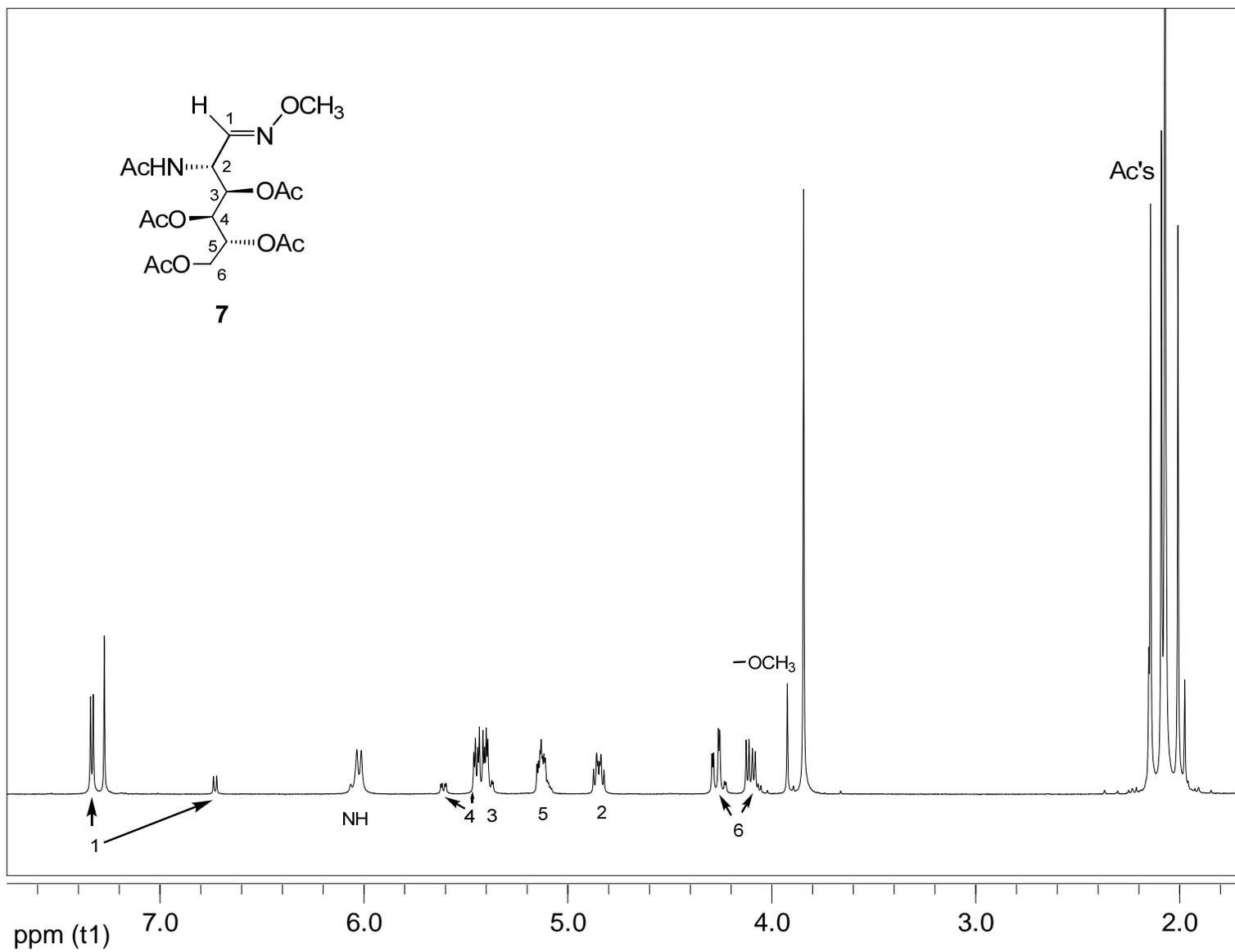


Figure A.1 ¹H NMR spectrum of 2-acetamido-2-deoxy-3,4,5,6-tetra-O-acetyl-D-mannose O-methyloxime (7).

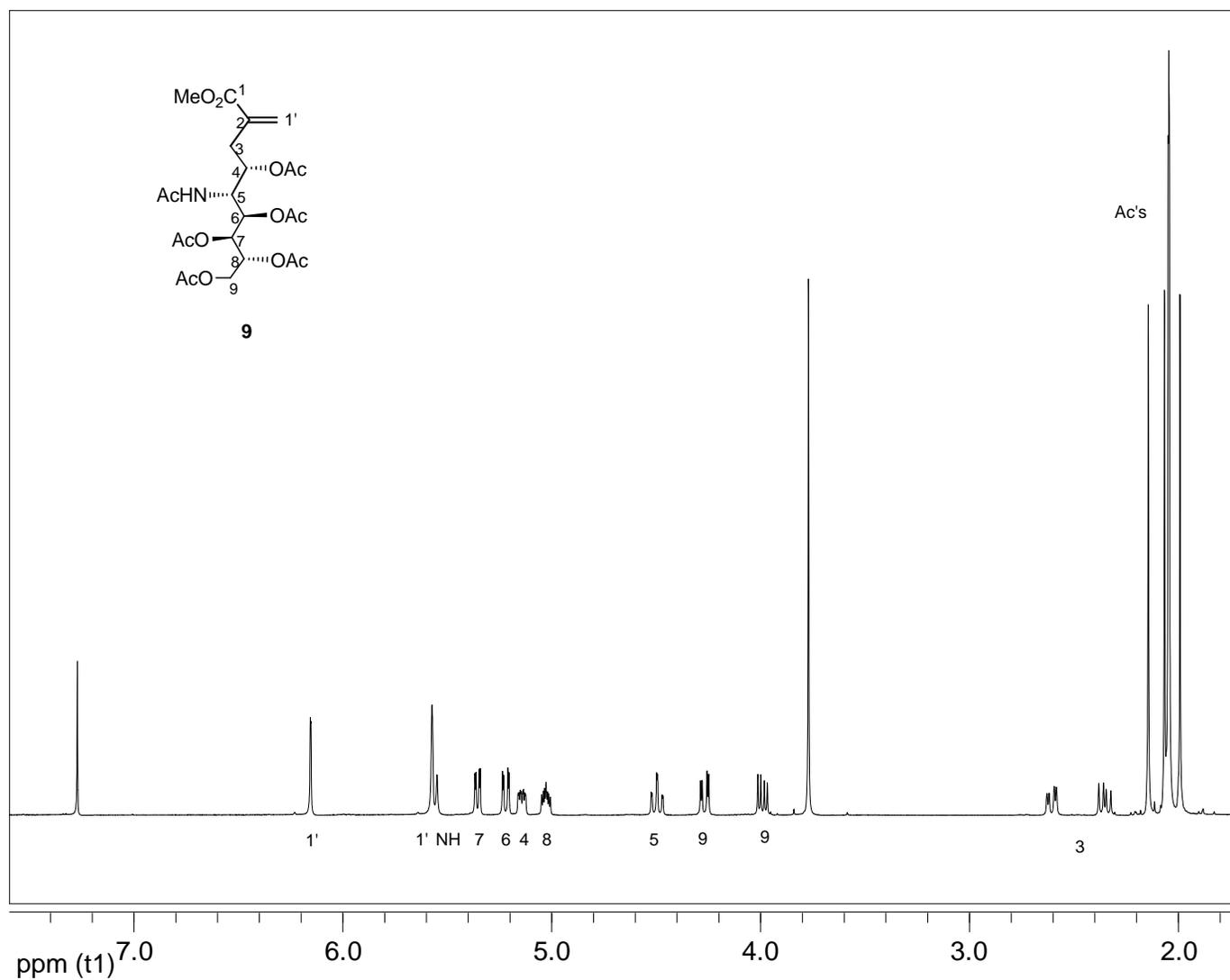


Figure A. 2 ¹H NMR spectrum of methyl 5-acetamido-3,5-dideoxy-2-methylidene-4,6,7,8,9-penta-*O*-acetyl-*D*-glycero-*D*-galacto-2-nonulosonate (9).

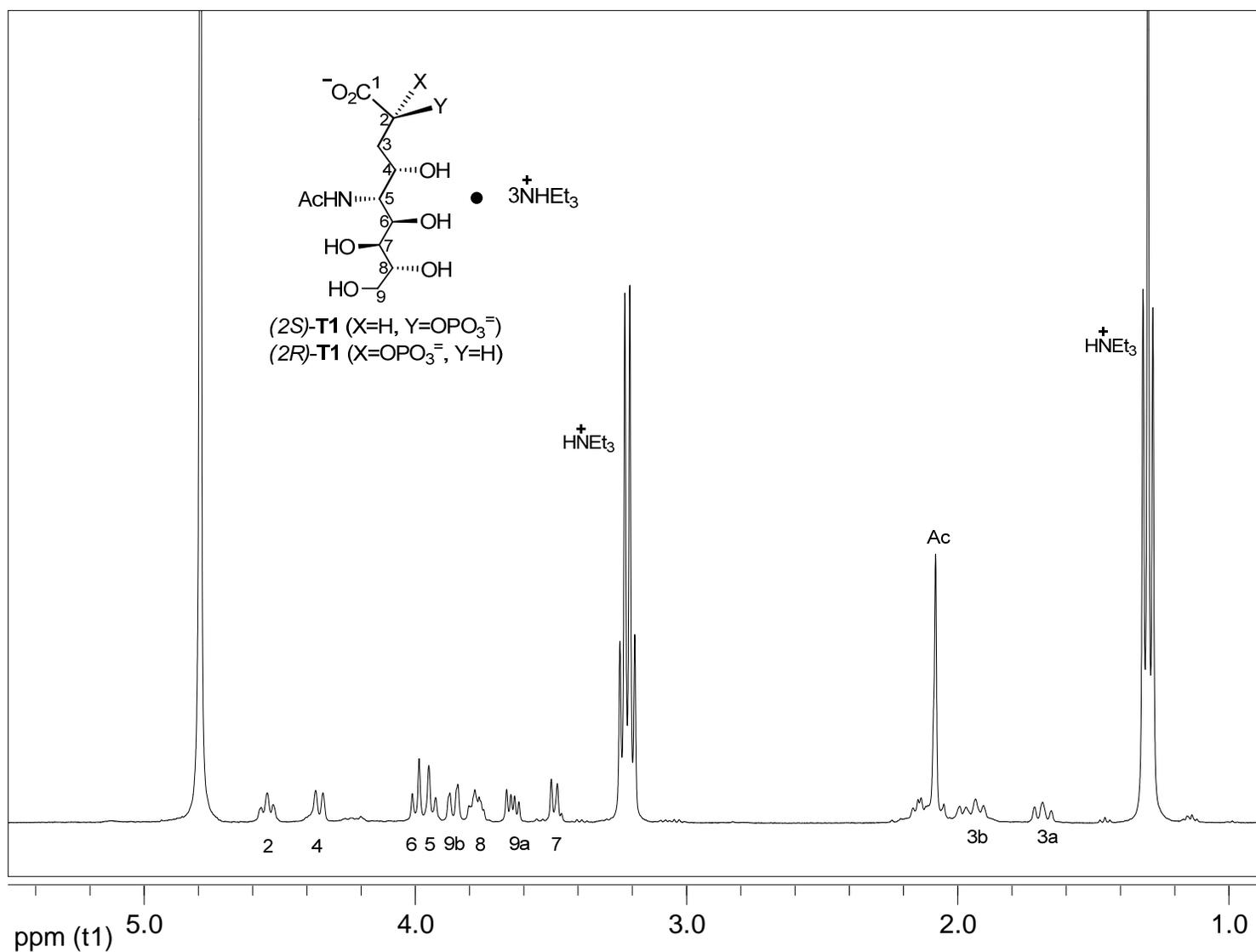


Figure A.3 ¹H NMR spectrum of 5-acetamido-4,6,7,8,9-pentahydroxy-2-phosphorylnonanoic acid ditriethylammonium salt (T1).

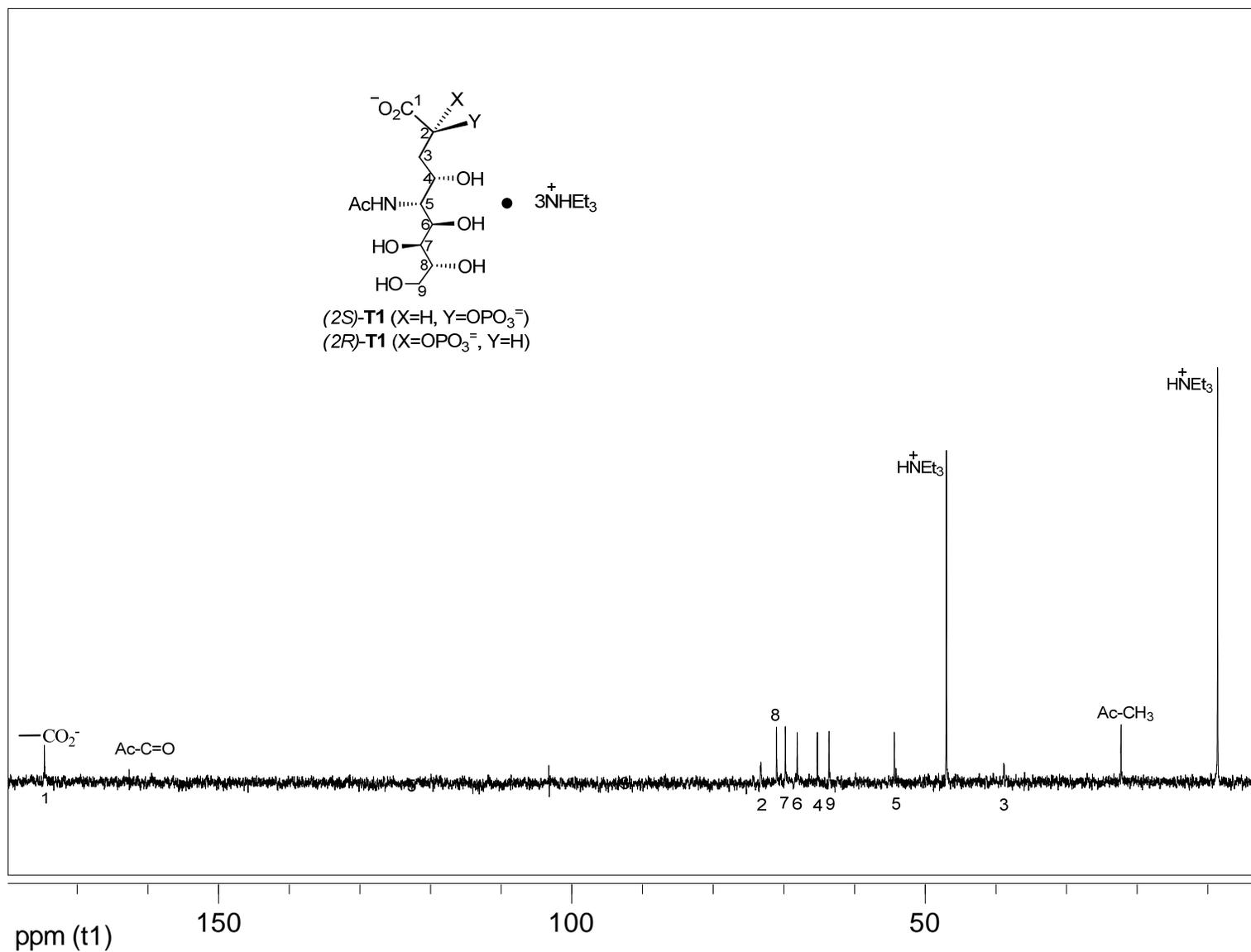


Figure A.4 ¹³C NMR spectrum of 5-acetamido-4,6,7,8,9-pentahydroxy-2-phosphorylnonanoic acid ditriethylammonium salt (T1).

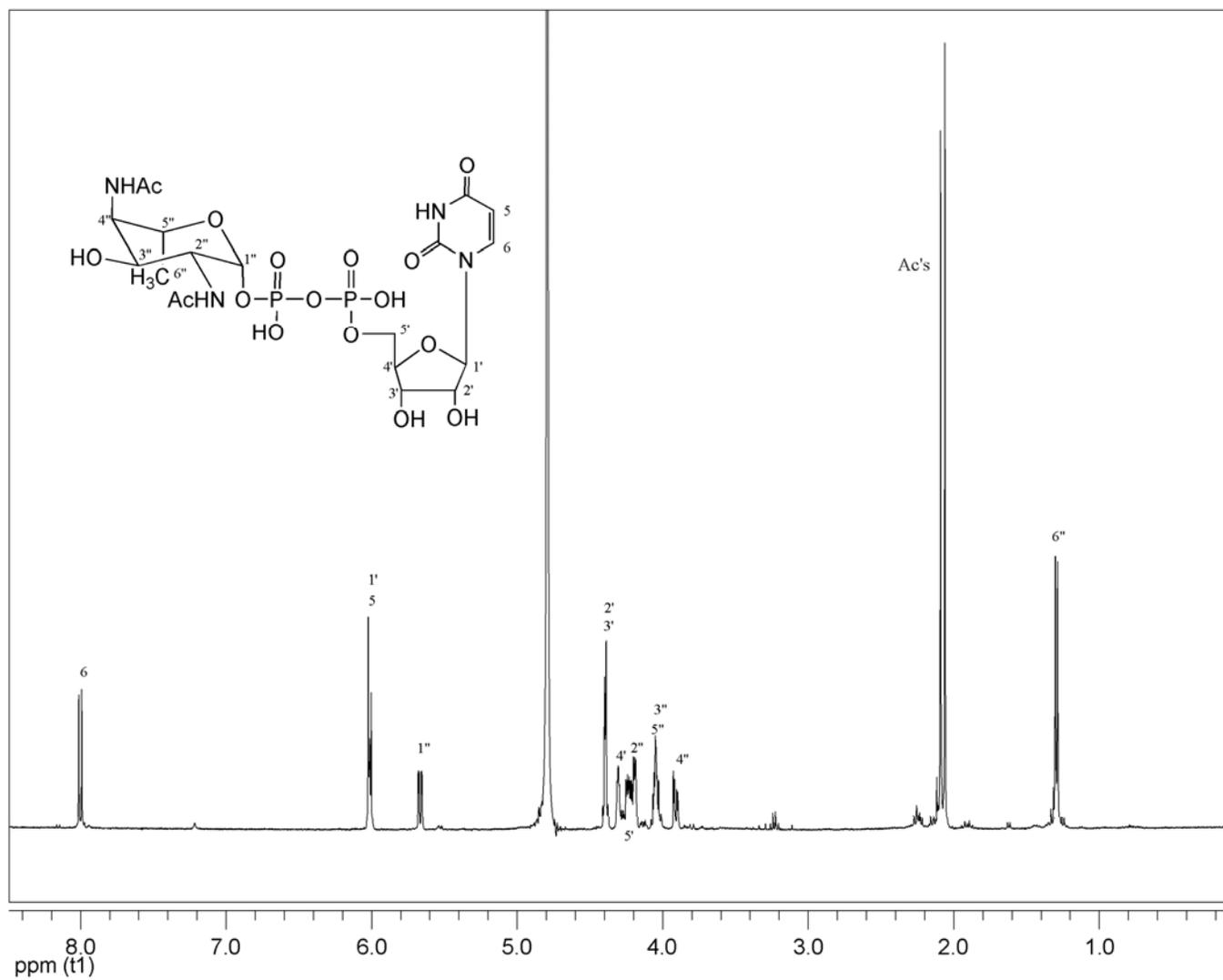


Figure A.5 ¹H NMR spectrum of UDP-6-deoxy-AltdiNAc.

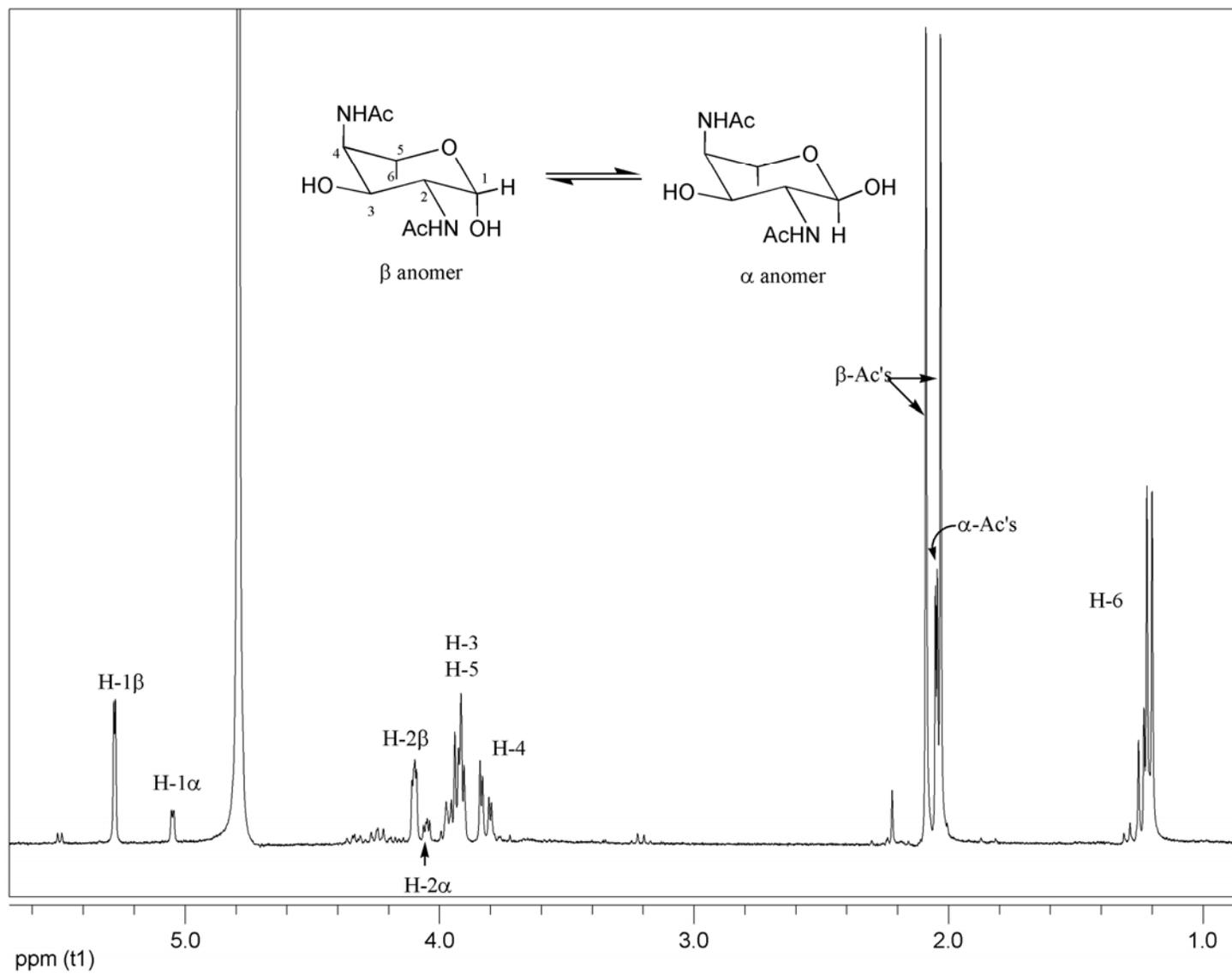


Figure A.6 ^1H NMR spectrum of 6-deoxy-AltdiNAc.

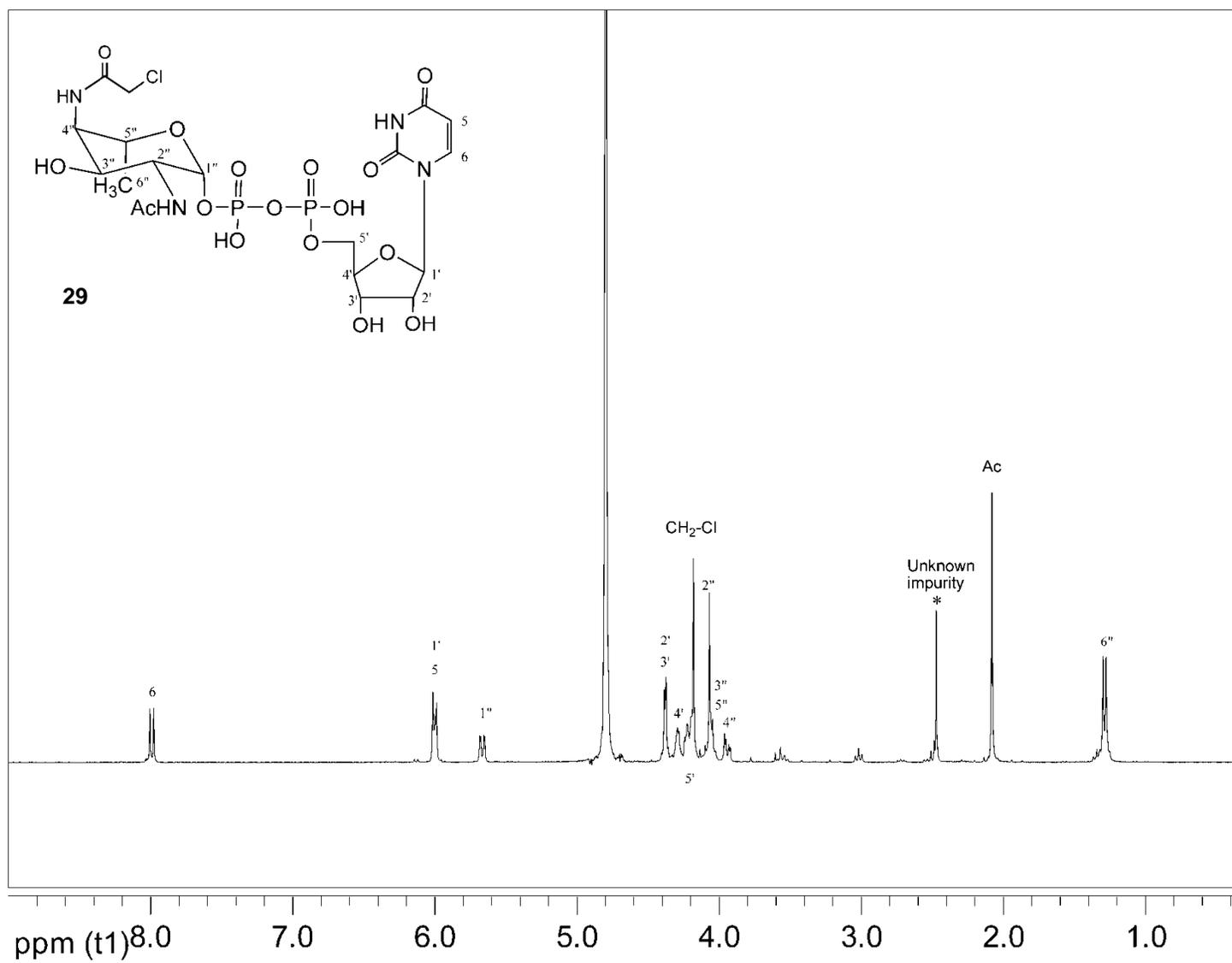


Figure A.7 ¹H NMR spectrum of UDP-6-deoxy-AltNAc4NAcCl (29).

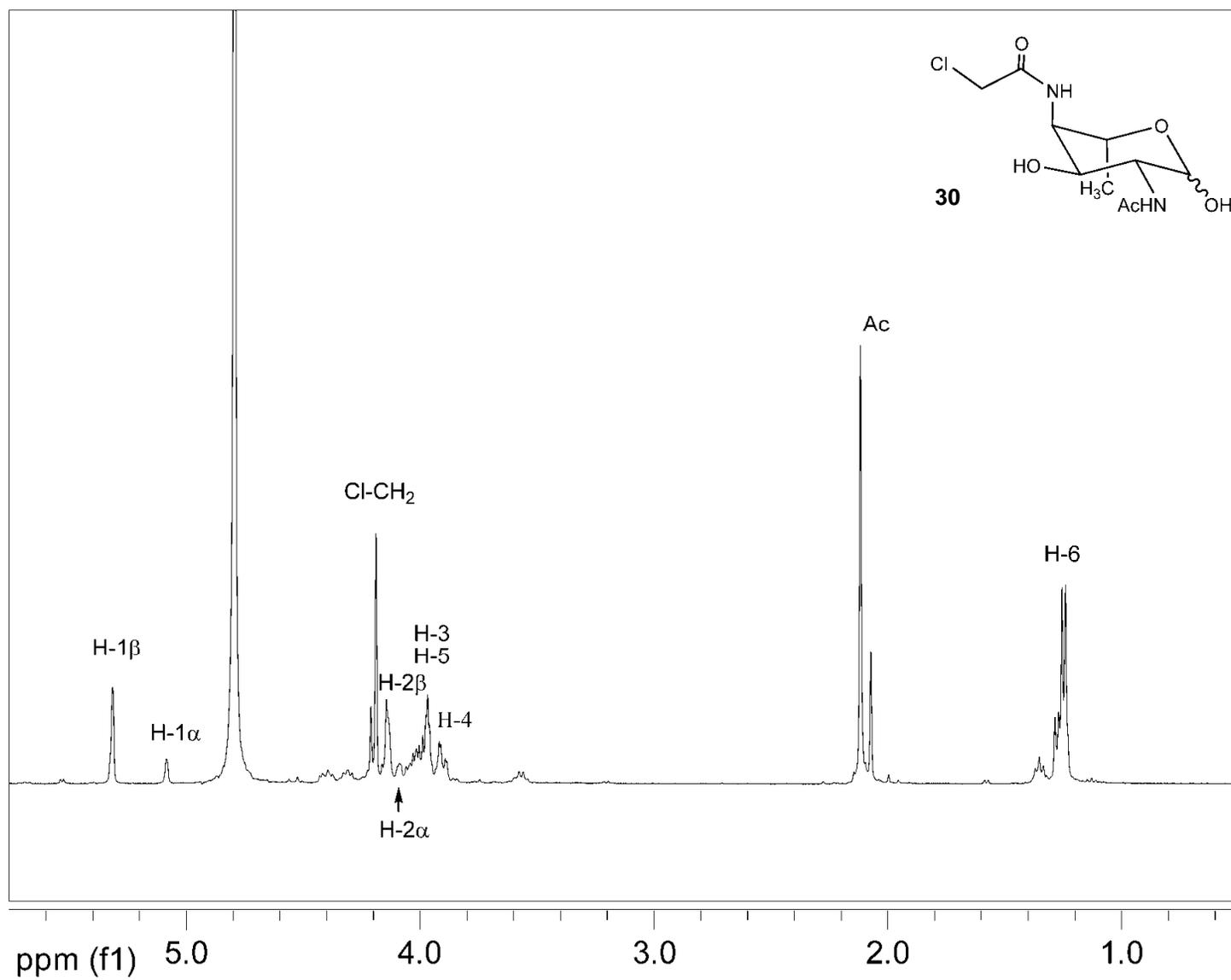


Figure A.8 ^1H NMR spectrum of 6-deoxy-AltNac4NAcCl (30).

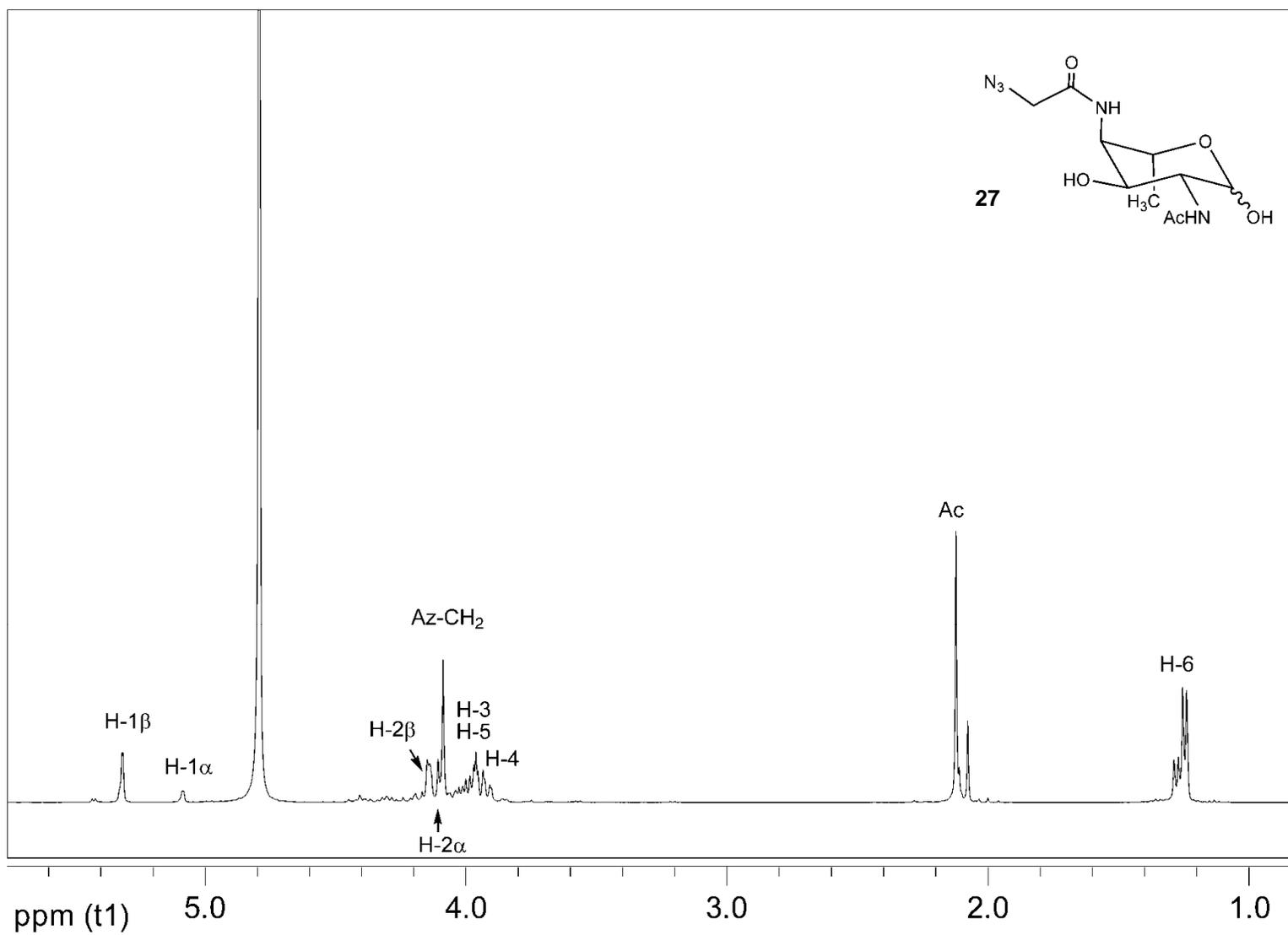


Figure A.9 ^1H NMR spectrum of 6-deoxy-AltdiNAc (27).

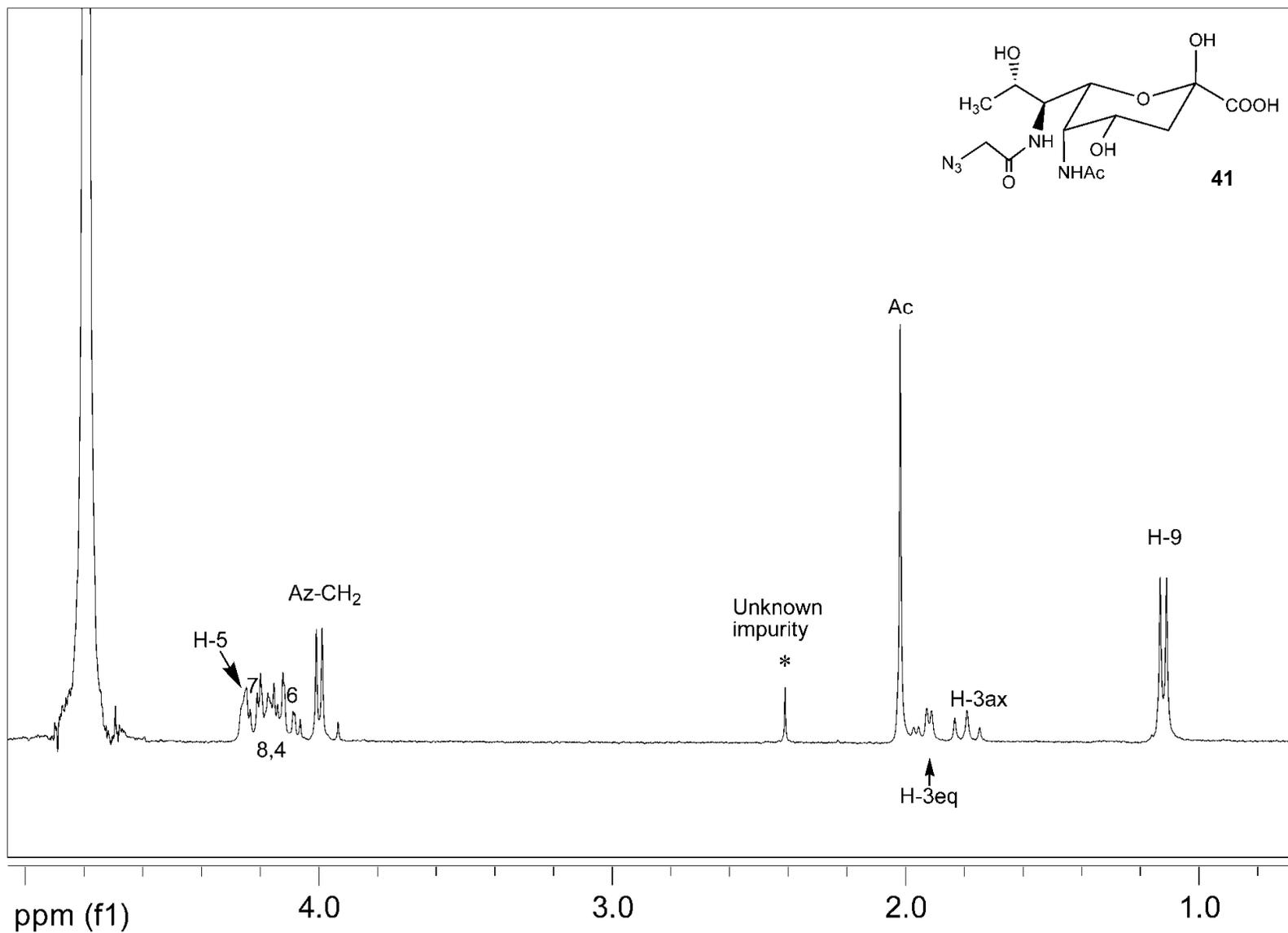


Figure A.10 ¹H NMR spectrum of Az-Pse (41).