

The Design, Synthesis and Assessment of Novel Haemagglutinin- Neuraminidase Inhibitors.

David James Tindal B. Biomed Sci. (Hons)

Submitted in fulfilment of the requirements of the degree of Doctor of Philosophy

Institute for Glycomics
Griffith University (Gold Coast Campus)

December 2006

Abstract

The human parainfluenza viruses (hPIV) are leading causes of respiratory disease in young children, the immunocompromised and the elderly. All subtypes can cause lower and upper respiratory disease. The surface of human parainfluenza virus and a number of other paramyxoviruses contains, amongst other biomolecules, a multifunctional glycoprotein known as haemagglutinin-neuraminidase (HN). The HN has three roles in viral pathogenesis; the binding to terminally bound sialic acid residues of cell oligosaccharides, thereby mediating viral attachment to the host cell; the promotion of viral and host cell fusion; and the cleavage of cell surface sialic acids to promote viral elution from host cells. The important role of HN in the viral life-cycle makes it an attractive target for therapeutic intervention. This thesis details an investigation into the design, synthesis and assessment of novel inhibitors of paramyxoviral haemagglutinin-neuraminidase.

Chapter 1 presents an introduction to the paramyxoviridae family. The role of HN in the life cycle of the paramyxoviruses hPIV and Newcastle Disease Virus and the characteristics of the glycoprotein are discussed. Included in this chapter is a discussion of the ligand for HN, the terminal glycoconjugate monosaccharide *N*-acetylneuraminic acid. Also discussed is the development of influenza virus sialidase inhibitors, as the processes used for their discovery were utilised within this project.

Chapter 2 describes the computational evaluation of the HN of Newcastle Disease Virus, a comparison with hPIV-3 HN, and the design of ligands to take advantage of potential interactions within the sialic acid binding site. The computational design process utilised *de novo* design on the C-4 position of 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid (Neu5Ac2en), a known sialidase inhibitor. The interactions within the C-4 pocket of HN are predominantly hydrophobic in nature, thus designed ligands incorporated a hydrophobic functionality at C-4 linked through either sulfur or oxygen. Also discussed is the *de novo* design of new templates principally using benzoic acid as the core structure.

Chapter 3 details the chemical synthesis of the ligands designed in Chapter 2 based on the Neu5Ac2en template. A series of 4-*S*-alkylated 4-thio-Neu5Ac2en derivatives and a series of 4-*O*-alkylated 4-ether-Neu5Ac2en derivatives were prepared, both series

starting from *N*-acetylneuraminic acid. Also discussed in this chapter is the observed epimerisation at C-4 of the C-4-thioether Neu5Ac2en derivatives upon deprotection mediated by sodium methoxide.

Chapter 4 describes the biological evaluation of the synthesised C-4-thioether and C-4-ether Neu5Ac2en derivatives as inhibitors of the sialidase activity of the HN's of human parainfluenza virus subtypes 1 and 3. Included in this chapter is the analysis of the X-ray crystallographic complex of one of the designed inhibitors, 4-*O*-benzyl-Neu5Ac2en and the HN of Newcastle Disease Virus. This chapter also draws together the final conclusions and remarks regarding the design, synthesis and assessment of novel inhibitors of haemagglutinin-neuraminidase.

Chapter 5 presents the experimental procedures that were utilised to gather the results described in Chapters 2 and 3. Computational input and parameter files to reproduce computational results discussed in Chapter 2 are presented in Appendix A. All computational data obtained for templates other than Neu5Ac2en is given in Appendix B. Representative NMR spectra of selected compounds and HPLC data of the deprotected C-4-thioether Neu5Ac2en derivatives to support the synthetic methods are presented in Appendices C and D respectively.

Table of Contents

Abstract	i
Statement of Originality	ix
Acknowledgements	xi
Communications and Publications	xiii
Abbreviations	xv
Table of Figures	xix
Table of Tables	xxiii
Nomenclature	xxv

CHAPTER 1 INTRODUCTION 1

1.1	Carbohydrates: A brief overview	1
1.1.1	Sialic acids: A general overview	2
1.2	Sialic acid recognising proteins.	3
1.2.1	Sialidases	4
1.3	An introduction to Paramyxoviruses	4
1.3.1	Human Parainfluenza Virus	5
1.3.1.1	Clinical information	6
1.3.2	Newcastle Disease Virus	7
1.3.2.1	Aetiology	7
1.3.3	The haemagglutinin-neuraminidase of the paramyxoviruses	8
1.3.3.1	Haemagglutinin-neuraminidase of NDV	10
1.3.3.2	Haemagglutinin-neuraminidase of hPIV-3	12
1.4	Current prevention and treatment of paramyxovirus infection	13
1.4.1	Vaccines	13
1.4.2	Therapeutic agents	14
1.4.2.1	Sialidase inhibitors.	14
1.4.2.2	Zanamivir mimetics	16
1.4.2.3	Inhibition of paramyxovirus haemagglutinin-neuraminidases	18
1.5	Aims and objectives	20

CHAPTER 2 COMPUTATIONAL ANALYSIS OF THE PARAMYXOVIRAL HN GLYCOPROTEIN 23

2.1	Introduction	23
2.1.1	Structure-based drug design	28
2.1.1.1	Influenza virus sialidase	28
2.2	Computational design	30

2.3	Computational analysis of NDV HN	31
2.3.1	Validation of NDV HN crystal structure	31
2.3.1.1	Procheck	32
2.3.1.2	Whatcheck	32
2.3.1.3	Hbplus	33
2.3.1.4	Correlation of results from Whatcheck and Hbplus	33
2.3.2	Initial investigation of NDV HN	33
2.3.2.1	The interactions of the active site of NDV HN with Neu5Ac2en.	33
2.3.2.2	Analysis of putative catalytic residue Asp 198.	35
2.3.2.3	DelPhi of unminimised NDV HN	36
2.3.2.4	C-4 binding pocket	37
2.3.2.5	PASS	38
2.3.2.6	GRID study of NDV HN crystal structure.	39
2.3.2.7	Pocket characterisation of NDV HN active site interactions	44
2.3.3	Validation of docking and scoring programs	45
2.3.3.1	AutoDock - Neu5Ac2en	45
2.3.3.2	SCORE - Neu5Ac2en	46
2.3.3.3	DOCK	47
2.3.4	Scoring of existing C-4 modified analogues of Neu5Ac2en.	48
2.3.4.1	Refinement of the initial modelling protocol	52
2.3.5	Scoring of existing C-5 modified analogues of Neu5Ac2en	56
2.3.6	AutoDock of variously linked sialyllactoses	57
2.3.7	<i>De novo</i> designed C-4 modified analogues of Neu5Ac2en.	58
2.3.7.1	Refinement of ligands to final forms.	61
2.3.7.2	Docking and scoring of refined <i>de novo</i> designed substituents – Initial examination of the Neu5Ac2en template	65
2.3.7.3	Improved computational examination of Neu5Ac2en template	70
2.3.8	Comparison of selected C-4- <i>S</i> and - <i>O</i> -linked Neu5Ac2en derivatives	76
2.3.8.1	Comparison of 4- <i>O</i> - and 4- <i>S</i> -benzyl-Neu5Ac2en computational structures.	76
2.3.8.2	Comparison of 4- <i>S</i> - and 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en computational structures.	77
2.3.8.3	Comparison of 4- <i>S</i> - and 4- <i>O</i> -ethyl-Neu5Ac2en computational structures.	78
2.3.8.4	Comparison of 4- <i>S</i> - and 4- <i>O</i> -hexyl-Neu5Ac2en computational structures.	79
2.3.8.5	Comparison of 4- <i>S</i> - and 4- <i>O</i> -decyl-Neu5Ac2en computational structures.	80
2.3.9	Scoring and docking results of templates other than Neu5Ac2en.	81
2.3.10	Docking and scoring of miscellaneous compounds	84
2.3.10.1	Modified C-4- <i>O</i> -Linked substituted Neu5Ac2en derivatives.	85
2.3.10.2	Neu5Gc2en and 5-glycolylamino derivative of 4- <i>S</i> -benzyl Neu5Ac2en	88
2.3.11	The development of potential new templates	89
2.3.11.1	LigBuilder design of alternative templates.	91
2.3.11.2	Di-substituted benzoic acid template	94
2.3.11.3	Final thoughts on the development of potential new templates	96

2.4	Homology Modelling	97
2.4.1.1	Preliminary analysis of homology models	98
2.4.2	Docking of Neu5Ac2en into hPIV-1, 2 and 3 HN homology models	101
2.4.2.1	Holzer <i>et al.</i> derivatives of Neu5Ac2en and hPIV-2 HN homology model	102
2.5	Computational analysis of hPIV-3 HN crystal structure.	103
2.5.1	Initial analysis of available hPIV-3 HN crystal structures.	103
2.5.1.1	Procheck	104
2.5.1.2	Secondary structure render	105
2.5.1.3	Electrostatics of hPIV-3 HN active site	105
2.5.1.4	C-4 binding pocket volume calculations	106
2.5.2	AutoDock of Neu5Ac2en and Zanamivir.	106
2.5.3	GRID analysis of hPIV-3 HN	107
2.5.4	Docking and scoring of 4- <i>O</i> -substituted-Neu5Ac2en derivatives and hPIV-3 HN	112
2.5.4.1	Examination of docking and scoring of hPIV-3 HN with crystal water residues within the C-4 binding pocket of the active site.	114
2.6	Overview and summation of the modelling procedure	115

CHAPTER 3 SYNTHESIS OF DESIGNED C-4 MODIFIED Neu5Ac2en DERIVATIVES AS POTENTIAL INHIBITORS OF PARAMYXOVIRAL HN. 119

3.1	Introduction	119
3.1.1	Synthesis of Neu5Ac2en	120
3.1.2	Synthesis of Neu5Ac2en derivatives	121
3.1.3	Approaches to Neu5Ac2en C-4 thioether and ether derivatives	123
3.2	C-4-Thioether Neu5Ac2en derivatives	123
3.2.1	Synthesis of C-4-Thioether Neu5Ac2en derivatives.	125
3.2.2	Synthesis of 2-methyl-(methyl 7,8,9-tri- <i>O</i> -acetyl-2,6-anhydro-3,5-dideoxy-D-glycero-D-talo-non-2-enonate)-[4,5-d]-2-oxazoline (56).	126
3.2.3	Synthesis of methyl 5-acetamido-7,8,9-tri- <i>O</i> -acetyl-4- <i>S</i> -acetyl-2,6-anhydro-3,5-dideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, (58).	129
3.2.4	Hydrazine acetate mediated C-4 thioether formation.	132
3.2.5	Deprotection of C-4-thioether derivatives of Neu5Ac2en	137
3.2.5.1	Deprotection of C-4-thioether Neu5Ac2en derivatives with potassium hydroxide.	139
3.3	Approaches to C-4-ether Neu5Ac2en derivatives	140
3.3.1	Synthesis of C-4-ether linked Neu5Ac2en derivatives	141
3.3.2	Preparation of key intermediate 8,9- <i>O</i> -isopropylidenated Neu5Ac2en1Me 52.	142
3.3.3	Coupling of alkyl/aralkyl halides with 8,9- <i>O</i> -isopropylidene-Neu5Ac2en1Me, 52.	145
3.3.3.1	Sodium hydride mediated alkylation of 52	145
3.3.3.2	Silver (I) oxide mediated alkylation of 52	146

3.3.3.3	Analysis of NaH and Ag ₂ O mediated alkylation reactions	147
3.3.4	Deprotection of 4- <i>O</i> -alkylated derivatives of Neu5Ac2en.	149
3.3.5	HPLC conditions for purification of C-4-thioether and C-4-ether Neu5Ac2en derivatives.	150
3.4	Summary and Conclusions	151
3.5	Future Work	152
3.5.1	Bis- <i>O</i> -benzylation of Neu5Ac2en1Me.	152
3.5.2	Alternative synthesis of 4-thiolacetyl Neu5Ac2en1Me derivative 58	153
3.5.3	Investigation of C-4 epimer formation from 4- <i>O</i> -benzyl-Neu5Ac2en1Me, 65-A .	154
3.5.4	Computational analysis of C-4-thioether and ether Neu5Ac2en derivatives.	155
CHAPTER 4 BIOLOGICAL EVALUATION OF LIGANDS DESIGNED TO INTERACT WITH PARAMYXOVIRAL HN.		157
4.1	Introduction	157
4.2	<i>In vitro</i> biological evaluation	158
4.2.1	Assay for inhibition of sialidase activity of hPIV-1 HN.	158
4.2.2	Assay for inhibition of sialidase activity of hPIV-3 HN.	160
4.2.3	Assay for inhibition of viral growth of rSV (hPIV-1 HN).	162
4.3	Comparisons between inhibitory data against hPIV sialidase activity and computational data (NDV HN) for C-4-ether and C-4-thioether derivatives.	163
4.3.1	Comparison of 4- <i>O</i> and 4- <i>S</i> -alkylated Neu5Ac2en derivatives	166
4.4	Analysis of the X-ray crystal structure of 4- <i>O</i> -benzyl-Neu5Ac2en/NDV HN complex.	166
4.4.1	Comparison of the X-ray crystal structure vs computationally docked 4- <i>O</i> -benzyl-Neu5Ac2en/NDV HN complexes.	170
4.5	Future Work	172
4.6	Summary and Conclusions	173
CHAPTER 5 EXPERIMENTAL		175
5.1	Modelling experimental	175
5.1.1	Structure validation	175
5.1.1.1	Optimisation of hydrogen-bond network of the protein.	175
5.1.2	Initial structure investigation	176
5.1.2.1	Ligplot identification of key interactions of Neu5Ac2en.	176
5.1.2.2	Delphi	176
5.1.2.3	PASS	176

5.1.2.4	Examination of the high and low pH conformations of NDV HN	177
5.1.2.5	GRID analysis of NDV HN active site	177
5.1.3	Homology models	177
5.1.4	<i>De novo</i> drug design using LigBuilder	178
5.1.4.1	Pocket	178
5.1.4.2	Grow or Link	178
5.1.4.3	Process	179
5.1.4.4	Further refinement of the ten best LigBuilder designed derivatives.	179
5.1.5	Initial docking and scoring method	181
5.1.6	Second docking and scoring method	183
5.1.7	<i>In Silico</i> Screening	187
5.1.7.1	DOCK	187
5.1.8	Miscellaneous programs	188
5.1.8.1	UNIX	188
5.1.8.2	InsightII	188
5.2	Synthesis experimental	188
5.2.1	General methods	188
5.2.2	Experimental methods	189
5.2.2.1	4- <i>S</i> -alkylated Neu5Ac2en derivatives	189
5.2.2.2	4- <i>O</i> -alkylated Neu5Ac2en derivatives	206
References		221
Appendix A – Computer Input and Parameter Files		229
Appendix B – Docking and Scoring Results of Alternative Templates		255
Appendix C – Selected ¹H NMR and COSY Spectra		263
Appendix D – HPLC data of C-4 thioether Neu5Ac2en Derivatives		273

Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself. This thesis is less than 100,000 words in length, excluding appendices and references.

David James Tindal

December 2006

Acknowledgments

Firstly, I would like to thank my principle supervisor Prof Mark von Itzstein for the opportunity to undertake my PhD at the Institute for Glycomics. Thankyou for the guidance and challenges you provided me during my candidature. I would also like to thank and acknowledge the support, patience, and boundless knowledge of my supervisors Dr Robin Thomson and Mr Jeffrey Dyason. Words cannot express how I have valued your support and care during my candidature. I thank you for all you have done. In many ways, I have changed over the course of my PhD and you have helped shape who I have become. It is much appreciated and I thankyou always.

I would also like to acknowledge the assistance of many of the researchers and administrative staff of the Institute for Glycomics. I would like to thank Dr Milton Kiefel and Dr Darren Grice for invaluable synthetic chemistry advice. I would not have made it to the end without it. I would like to thank Dr Jennifer Wilson for general and troubleshooting advice regarding the use of the NMR spectrometer. I would like to thank Prof Garry Taylor and his group in St Andrews, Scotland, for solving the crystal structure of 4-*O*-benzyl-Neu5Ac2en/NDV HN complex. I would also like to thank Prof Yasuo Suzuki and Prof Allen Portner and their respective groups for undertaking the biological testing of the compounds designed in this thesis. Thanks also to Dr Alex Szyzew for many little tips and tricks regarding the use and troubleshooting of the HPLC machines. I would like to thank Ms Fiona Crone for the management of the many administrative duties required. Thanks also to Mrs Sharon Ackerman for all the friendship, care and support that you have provided for me over the years. It will always be appreciated.

I would like to thank Dr Angela Liakatos for many enjoyable hours spent invariably laughing uncontrollably, you helped keep me sane. I would also like to thank you for many insightful comments made concerning both work and non-work related topics. The movie nights and conversations we had will stay with me for a long time to come. I would like to thank Greg Tredwell for his friendship, worldly knowledge, and his uncanny ability to cheer me up when the black clouds became too heavy; and for having an excellent knack at seeing through to the heart of an issue and opening me to many new experiences. Thankyou also to Patrick Collins for his friendship, support and many hilarious conversations. I have loved spending time with you over the years in uncontrollable laughter in the many situations that have arisen over the years. I have always appreciated that my PhD has allowed me to get to know the three of you over the years. I have many fantastic memories to take with me and look forward to creating new ones in the future. I would like to thank Dr Florence Chery for her many hours of help, and her friendship, I had many enjoyable times with you and will remember them fondly always. Thanks to Hannah Hartig for her friendship, care and worldly knowledge, I have learnt many things from

our many conversations. I would also like to thank Szilvia Papp and Ngahine Vines. You have hearts of gold and I have loved getting to know you over the years, in my good and bad moments you have been rocks in my life and I will always cherish the friendships that I have with each of you. Thanks to Joshua Rose, Kate Plant, Sarah Blyth (nee McAtamney) and Anthony Blyth for being kind, insightful and fun people. I value your friendship always and look forward to the many years of friendship in the future. Thanks to Andrea, Anna, Tareq, Chris, Lirio, Tessa, Stacey, Paul, Kristina, Marettta and Christian for all your help, support and all the good times we had. I would also like to acknowledge and thanks all the other people who have helped me over the duration of my candidature, you know who you are. Thankyou.

To my beautiful and dearly cherished family (Mum, Splinter, Teresa, Leisen, Josh, Chris, Peter and Kiera), I will never forget the constant support, love and care that you have bestowed upon me. I will always be eternally thankful that I was put on this earth to spend my life with the kindest, patient, and devoted Mum that I could ever hope for. This thesis is as much yours as it is mine, for without your endless support I would not have lasted to the end. I would also like to thank Ileana, Kerry, Angela and Christopher for their support and care. I look forward to the many years of friendship and love in the future; you are all beautiful and caring people. I am very thankful that the decisions I have made has brought you into my life. I would also like to thank and acknowledge the care and support provided by the extended Tindal and Blythman families.

Lastly, but not least, I would like to thank my fiancé Catherine Whelan. There are not enough words to express adequately how lucky I am to have found such a beautiful, kind, fascinating and endlessly supportive woman. You keep me grounded and sane (and I realise how difficult that may be). Without your support and care I would never have come this far. I look forward to spending many years with you and I appreciate everything you have done and will do in the future. Thankyou always and forever.

Communications and Publications

The following communications and publications have resulted from the work described in this thesis.

Communications

Tindal, D. J.; Dyason, J.C.; Thomson, R. J.; von Itzstein, M. (2006). "A GRID study of the multifunctional haemagglutinin-neuraminidase of human parainfluenza virus 3", *XXIIIrd International Carbohydrate Symposium*, Whistler, Canada, July 23-28, Abstract MON-PS.95.

Tindal, D. J.; Dyason, J.C.; Thomson, R. J.; Suzuki, T.; Suzuki, Y.; von Itzstein, M. (2005). "Design, synthesis and evaluation of inhibitors of paramyxoviral haemagglutinin-neuraminidase", *13th European Carbohydrate Symposium*, Bratislava, Slovakia, August 21-26, Abstract P157.

Tindal, D. J.; Dyason, J.C.; Thomson, R. J.; Suzuki, T.; Suzuki, Y.; von Itzstein, M. (2004). "The design, synthesis and evaluation of Paramyxoviridae haemagglutinin-neuraminidase inhibitors", *XXIInd International Carbohydrate Symposium*, Glasgow, United Kingdom, July 23-27, Abstract P290.

Tindal, D. J.; Dyason, J.C.; Thomson, R. J.; von Itzstein, M. (2002). "A GRID study of the haemagglutinin-neuraminidase of Newcastle Disease Virus" *International Conference on Sialoglycoscience and Other Forms of Glycosylation*, Gold Coast, Australia, July 14-17, Abstract PP18.

Tindal, D. J.; Dyason, J.C.; Thomson, R. J.; von Itzstein, M. (2002). "A GRID study of the haemagglutinin-neuraminidase of Newcastle Disease Virus", *XXIst International Carbohydrate Symposium*, Cairns, Australia, July 7-12, Abstract PP075.

Publications

Tindal, D. J.; Dyason, J.C.; Thomson, R. J.; von Itzstein, M. (2006). A comparative GRID study of the multifunctional Haemagglutinin-Neuraminidases of Newcastle Disease Virus and human parainfluenza virus type-3, manuscript in preparation.

Tindal, D. J.; Dyason, J.C.; Thomson, R. J.; Suzuki, T.; Suzuki, Y.; von Itzstein, M. (2007). Design, synthesis, and evaluation of C-4 substituted 2,3-anhydro-2-deoxy-*N*-acetylneuraminic acid derivatives as inhibitors of human paramyxoviruses, accepted for publication.

Ryan, C.; Zaitsev, V.; Tindal, D. J.; Dyason, J. C.; Thomson, R. J.; Alymova, I. V.; Portner, A.; von Itzstein, M.; Taylor, G. (2006). Structural analysis of a designed inhibitor complexed with the hemagglutinin-neuraminidase of Newcastle Disease Virus, *Glycoconjugate Journal*, 23, 135-141.

Abbreviations

General abbreviations

Å	ångström(s)
Ac	acetyl
Ac ₂ O	acetic anhydride
ACN	Acetonitrile
AcSH	Thiolacetic acid
ADT	AutoDock Tools
BCL	Biosym Command Language
BF ₃ •Et ₂ O	Boron trifluoride etherate
Bn	benzyl
br	broad (spectral)
°C	degree(s) Celsius
COSY	correlation spectroscopy
CVFF	Consistent Valence Force field
δ	chemical shift in parts per million
d	day(s), doublet (spectral)
dd	doublet of doublets (spectral)
ddd	doublet of doublet of doublets (spectral)
1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
DBU	1,8-diazobicyclo-[5.4.0]undec-7-ene
DCM	dichloromethane
DMF	<i>N,N</i> -dimethylformamide
DMTST	dimethyl(methylthio) sulfonium triflate
DNA	deoxyribonucleic acid
Et	ethyl
equiv	equivalent(s)
ESI	Electrospray Ionisation (in Mass Spectrometry)
F	Fusion protein
g	gram(s)
gaff	generalised amber force field
GUI	Graphical User Interface
h	hour(s)
HN	haemagglutinin-neuraminidase
hPIV	human parainfluenza virus
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry/spectrum

HSQC	heteronuclear single quantum correlation
Hz	hertz
IR	infrared
<i>J</i>	coupling constant
k	kilo
kcal/mol	kilocalorie per mole
λ	wavelength
L	litre
LRMS	low-resolution mass spectrometry/spectrum
μ	micro
m	multiplet (spectral), milli, metre
M	molar (moles per litre), mega
Me	methyl
MHz	Megahertz
min	minute(s)
mL	millilitre
MMI	Molecular Mechanics Interaction
MM-PBSA	Molecular Mechanical Poisson-Boltzmann Surface Area
mol	mole(s)
MS	mass spectrometry
MUN	4'-methylumbelliferyl
<i>m/z</i>	mass-to-charge ratio
n	nano
NA	Neuraminidase
NDV	Newcastle Disease Virus
NMR	nuclear magnetic resonance
o/n	overnight
Ph	phenyl
ppm	parts per million
Pr	propyl
q	quartet (spectral)
<i>R_f</i>	retention factor
rt	room temperature
s	second(s), singlet (spectral)
SARP	Sialic acid recognising protein
SGI	Silicon Graphics Incorporated
t	triplet (spectral)
Tf	trifluoromethanesulfonyl (triflyl)
THF	tetrahydrofuran
TLC	thin layer chromatography

TMSOTf	Trimethylsilyl trifluoromethanesulfonate
UV	ultraviolet

Amino acid abbreviations

Ala	alanine, A
Arg	arginine, R
Asn	asparagine, N
Asp	aspartic acid, D
Cys	cysteine, C
Gln	glutamine, Q
Glu	glutamic acid, E
Gly	glycine, G
His	histidine, H
Ile	isoleucine, I
Leu	leucine, L
Lys	lysine, K
Met	methionine, M
Phe	phenylalanine, F
Pro	proline, P
Ser	serine, S
Thr	threonine, T
Trp	tryptophan, W
Tyr	tyrosine, Y
Val	valine, V

AMBER amino acid abbreviations

CYX	cysteine involved in a disulfide linkage
HID	histidine with atom CD protonated
HIE	histidine with atom CE protonated
HIP	histidine with both atoms CD and CE protonated

Sialic acid abbreviations

KDN	3-deoxy-D-glycero-D-galacto-2-nonulosonic acid or 2-keto-3-deoxy-D-glycero-D-galacto-non-2-ulosonic acid.
Neu	5-amino-3,5-dideoxy-D-glycero-D-galacto-non-2-ulosonic acid or neuraminic acid.
Neu5Ac	5-acetamido-3,5-dideoxy-D-glycero-D-galacto-non-2-ulosonic acid or <i>N</i> -acetylneuraminic acid.
Neu5Ac2en	5-Acetamido-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-2-non-2-enoic acid or 2-deoxy-2,3-didehydro- <i>N</i> -acetylneuraminic acid

Neu5Ac1Me	Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-D- <i>glycero</i> -D- <i>galacto</i> -non-2-enonate
Neu5Gc	3,5-dideoxy-5-(2-hydroxyacetamido)-D- <i>glycero</i> -D- <i>galacto</i> -non-2-ulosonic acid or <i>N</i> -glycolylneuraminic acid.

List of Figures

Figure 1.1: Utilisation of carbohydrates as a receptor for adhesion to cells.....	2
Figure 1.2: Electron micrograph of hPIV-3 (Left) and NDV(Right) virions	5
Figure 1.3: Schematic representation of parainfluenza virus, its genome and surface glycoproteins. HN = Haemagglutinin-neuraminidase, F = Fusion protein.....	9
Figure 1.4: The structure of the NDV virus haemagglutinin-neuraminidase with bound inhibitor, Neu5Ac2en, 5.....	10
Figure 1.5: The structure of the hPIV-3 haemagglutinin-neuraminidase with bound inhibitor, Neu5Ac2en	12
Figure 2.1: The structure of an influenza virus sialidase monomer	29
Figure 2.2: Flow diagram of design process.	30
Figure 2.3: Ligplot of interactions of Neu5Ac2en with NDV HN.....	34
Figure 2.4: Superimposition of NDV HN and influenza virus N9 sialidase triarginyly clusters.	36
Figure 2.5: Initial DelPhi electrostatics calculation results of NDV HN	37
Figure 2.6: C-4 binding pocket of NDV HN with crystal waters CPK rendered.	38
Figure 2.7: Invaginations in the HN protein surface located by PASS	38
Figure 2.8: Area of interaction with the carboxylate (COO ⁻) probe, contoured at -17 kcal/mol.	40
Figure 2.9: Areas of interaction with the water (OH ₂) probe, contoured at -6 kcal/mol.....	41
Figure 2.10: Areas of interaction with the hydroxyl (OH) probe, contoured at -8 kcal/mol.	41
Figure 2.11: Areas of interaction with the methyl (C3) probe, contoured at -3 kcal/mol.....	42
Figure 2.12: Areas of interaction with the hydrophobic (DRY) probe, contoured at -0.2 kcal/mol.....	42
Figure 2.13: Areas of interaction with the guanidino (AMIDINE) probe, contoured at -12 kcal/mol. ...	43
Figure 2.14: Areas of interaction with the amine (N1=) probe, contoured at -6 kcal/mol.....	43
Figure 2.15: Area of interaction with the calcium ion (CA ⁺²) probe, contoured at -130 kcal/mol.....	44
Figure 2.16: Pocket analysis of NDV HN active site.....	45
Figure 2.17: Crystal bound (purple) and flexibly docked (green) Neu5Ac2en in NDV HN active site... 46	
Figure 2.18: SCORE classification of atom interactions of crystal bound Neu5Ac2en in NDV HN active site	46
Figure 2.19: AutoDock orientation of derivative 27.	54
Figure 2.20: SCORE result of derivative 23	55
Figure 2.21: Ligplot analysis of Neu5Ac2en derivative 23 in HDV HN.	56
Figure 2.22: The structure of the four LigBuilder seeds	59
Figure 2.23: An example of the process of structure refinement.	62
Figure 2.24: Structure of the final 14 substituents, with the point of attachment to the core template indicated.	63
Figure 2.25: Structures of the 12 templates used in ligand design.	64
Figure 2.26: Superimposition of the lowest energy structures of 4- <i>O</i> -benzyl-Neu5Ac2en (green) and 4- <i>S</i> - benzyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.....	77
Figure 2.27: Superimposition of the lowest energy structures of 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en (green) and 4- <i>S</i> -(2-phenyl)benzyl-Neu5Ac2en (purple) in the minimised active site of NDV HN..	78

Figure 2.28: Superimposition of the lowest energy structures of 4- <i>O</i> -ethyl-Neu5Ac2en (green) and 4- <i>S</i> -ethyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.	79
Figure 2.29: Superimposition of the lowest energy structure of 4- <i>O</i> -hexyl-Neu5Ac2en (green) and 4- <i>S</i> -hexyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.	80
Figure 2.30: Superimposition of the lowest energy structure of 4- <i>O</i> -decyl-Neu5Ac2en (green) and 4- <i>S</i> -decyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.....	81
Figure 2.31: Structures of C-8 or C-9-modified C-4-ether linked Neu5Ac2en derivatives.....	85
Figure 2.32: AutoDock orientation of 4,9-di- <i>O</i> -benzyl-Neu5Ac2en 29-P	86
Figure 2.33: AutoDock orientation of 8-amino-4- <i>O</i> -benzyl-8-deoxy-Neu5Ac2en 29-S	88
Figure 2.34: Structures of potential new alkyl and benzoic acid templates.....	90
Figure 2.35: AutoDock result for the alkyl template 42	91
Figure 2.36: AutoDock result for benzoic acid template 43	91
Figure 2.37: The 7 seed structures that were used in LigBuilder.....	92
Figure 2.38: The structure of 3-thiobenzyl-5-guanidino-benzoic acid 44	93
Figure 2.39: Docked conformation of the 3-thiobenzyl-5-guanidino-benzoic acid template 44	93
Figure 2.40: Structure of the 3,5-di-substituted benzoic acid template.	94
Figure 2.41: AutoDock (left) and SCORE (right) results for 3,5-di-thiobenzyl benzoic acid template 45	95
Figure 2.42: AutoDock (left) and SCORE (right) results for 3,5-di-benzyloxy benzoic acid template 46	96
Figure 2.43: Multiple HN Alignments.....	98
Figure 2.44: Molscript diagrams of the tertiary structure of hPIV-1, 2, & 3 homology models.....	99
Figure 2.45: Delphi analysis of the electrostatic potential of the active sites of the hPIV-1, 2, and 3 HN homology models with Neu5Ac2en docked in the active sites	100
Figure 2.46: Pocket analysis of the active sites of the hPIV-1, 2, and 3 HN homology models with Neu5Ac2en docked in the active sites.....	101
Figure 2.47: PASS analysis of the active sites of the hPIV-1, 2, and 3 HN homology models with Neu5Ac2en docked in the active sites.....	101
Figure 2.48: Secondary structure comparison between hPIV-3 HN crystal structure (left) and homology model (right)	105
Figure 2.49: The electrostatic potential of the active site of the crystal structure of hPIV-3 HN	106
Figure 2.50: Orientation of Neu5Ac2en (left) and zanamivir (right) in hPIV-3 HN	107
Figure 2.51: The carboxylate probe interaction site with hPIV-3 HN, contoured at -18 kcal/mol.	107
Figure 2.52: Calcium (CA2+) probe interaction site with hPIV-3 HN, contoured at -75.0 kcal/mol	108
Figure 2.53: Methyl (C3) probe interactions with hPIV-3 HN, contoured at -3 kcal/mol.	108
Figure 2.54: Hydrophobic (DRY) probe interaction sites with hPIV-3 HN, contoured at -0.8 kcal/mol.	109
Figure 2.55: Water (OH2) probe interactions with the active site of hPIV-3 HN, contoured at -8 kcal/mol.....	109
Figure 2.56: Aliphatic hydroxyl (O1) probe interaction sites with hPIV-3 HN, contoured at -8 kcal/mol.	110

Figure 2.57: Guanidino (AMIDINE) probe interactions with active site of hPIV-3 HN, contoured at – 10.0 kcal/mol.	110
Figure 2.58: Protonated Amino (N3+) probe interaction sites with hPIV-3 HN, contoured at –8 kcal/mol.	111
Figure 2.59: Ligplot diagram showing water interactions in hPIV-3 active site with Neu5Ac2en.	115
Figure 3.1: Structure of the nine 4- <i>S</i> -substituted Neu5Ac2en derivatives to be synthesised.	133
Figure 3.2: Structure of the C-4-thioether linked derivatives of Neu5Ac2en synthesised.	140
Figure 3.3: Structure of the five 4- <i>O</i> -substituted derivatives of Neu5Ac2en to be synthesised.	145
Figure 4.1: Rendered overview of the tertiary structure of the complex of NDV HN with 4- <i>O</i> -benzyl-Neu5Ac2en 29-A (white)	167
Figure 4.2: Electrostatics analysis of active site of NDV HN, with crystal 4- <i>O</i> -benzyl-Neu5Ac2en rendered.	168
Figure 4.3: Ligplot analysis of important interactions between 4- <i>O</i> -benzyl-Neu5Ac2en 29-A and NDV HN.	169
Figure 4.4: Close up of NDV active site complexed with 29-A	170
Figure 4.5: Superimposition of docked (purple) and crystal (green) bound 4- <i>O</i> -benzyl-Neu5Ac2en in NDV HN.	171
Figure 4.6: Superimposition of docked (purple) and crystal (green) bound 4- <i>O</i> -benzyl-Neu5Ac2en in the full active site of NDV HN.	172
Figure 5.1: Structures of the 12 templates used in ligand design.	180
Figure 5.2: Structures of the final fourteen substituents.	180

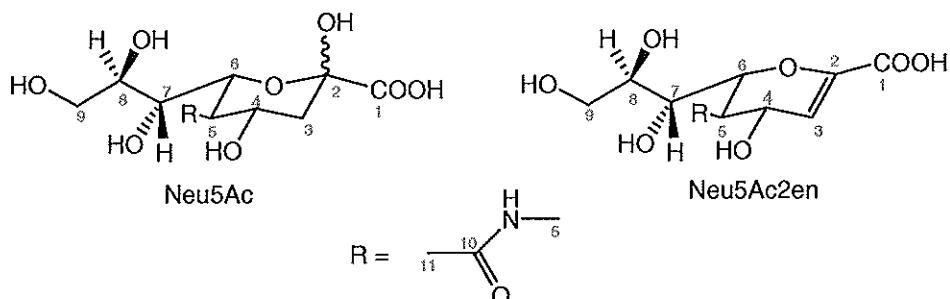
List of Tables

Table 2.1: Procheck results of NDV HN crystal structures 1E8T, 1E8U and 1E8V.....	32
Table 2.2: Residues with side-chains flipped 180°.....	33
Table 2.3: NDV HN and Neu5Ac2en H-bond interactions.	35
Table 2.4: AutoDock and SCORE results for C-4 modified Neu5Ac2en derivatives in NDV HN of ligands known to inhibit hPIV-2.....	49
Table 2.5: Flexible AutoDock and SCORE values for C-4 modified Neu5Ac2en derivatives of ligands known to inhibit influenza virus sialidase in NDV HN	51
Table 2.6: Docking and scoring results in NDV HN for the C-4 modified Neu5Ac2en derivative reported by Suzuki <i>et al.</i> as inhibitors of hPIV-1.....	53
Table 2.7: Scoring of flexibly docked C-5 modified analogues of Neu5Ac2en in NDV HN vs inhibition against NDV HN.	57
Table 2.8: <i>De novo</i> designed C-4 O/S-linked derivatives of Neu5Ac2en.	61
Table 2.9: Scoring results for Neu5Ac2en derivatives from AutoDock, MMI and SCORE.....	66
Table 2.10: Scoring results for Neu5Ac2en using updated modelling protocol (Section 2.3.7.3).	71
Table 2.11: Scoring results for Neu5Ac2en derivatives from AutoDock, MM-PBSA and X-Score.	72
Table 2.12: Comparative data for 4- <i>S</i> - and 4- <i>O</i> -benzyl-Neu5Ac2en	77
Table 2.13: Comparative data for 4- <i>S</i> - and 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en	78
Table 2.14: Comparative data for 4- <i>S</i> - and 4- <i>O</i> -ethyl-Neu5Ac2en.....	79
Table 2.15: Comparative data for 4- <i>S</i> - and 4- <i>O</i> -hexyl-Neu5Ac2en.....	80
Table 2.16: Comparative data for 4- <i>S</i> - and 4- <i>O</i> -decyl-Neu5Ac2en	81
Table 2.17: Energetic range of computational data (Autodock, MM-PBSA, X-Score) for all template derivatives.....	84
Table 2.18: Docking and scoring results for C-4 modified 9- <i>O</i> -benyl-Neu5Ac2en derivatives, 29-O, P, and Q.	86
Table 2.19: Docking and scoring results for C-4 modified 8-amino-8-deoxy-Neu5Ac2en derivatives, 29- R, S, and T.	87
Table 2.20: Scoring results of the di-substituted benzoic acid templates 45 and 46 with NDV HN.	95
Table 2.21: Flexible AutoDock results for Neu5Ac2en docked into the active sites of hPIV-1, 2 & 3 HN homology models.	102
Table 2.22: AutoDock, MMI and SCORE results for C-4 modified Neu5Ac2en derivatives in hPIV-2 HN homology model of ligands known to inhibit hPIV-2.....	103
Table 2.23: Superimposed active site residues of NDV and hPIV-3 HN.....	104
Table 2.24: Docking and scoring results of 4- <i>O</i> -substituted-Neu5Ac2en derivatives 29 in hPIV-3 HN.	113
Table 2.25: Hydrogen bonds formed in the docking of Neu5Ac2en to hPIV-3 HN.	115
Table 2.26: Benchmarks set by various programs for Neu5Ac2en in NDV HN.....	116
Table 3.1: Some of the modifications carried-out on Neu5Ac 1.	119
Table 3.2: Alternative reaction conditions for formation of 4- <i>S</i> -acetyl-Neu5Ac2en 1Me derivative 58 from oxazoline 56.	131

Table 3.3: Summary of coupling reaction conditions trialled for synthesis of 4- <i>S</i> -benzyl Neu5Ac2en1Me derivative 59-A from 58	134
Table 3.4: Yields of hydrazine acetate mediated coupling reactions of 4-SAc Neu5Ac2en derivative 58 with various alkyl/aralkyl halides.	136
Table 3.5: Selected data indicating successful coupling of alkyl/aralkyl halides with 58	137
Table 3.6: Summary of the elimination reactions examined for the synthesis of 49 from 47	144
Table 3.7: Yields of sodium hydride and silver (I) oxide alkylation reactions on 52	147
Table 3.8: Selected data for the 4- <i>O</i> -alkylated-Neu5Ac2en derivatives 65	149
Table 3.9: HPLC conditions for purification of C-4-thioether and C-4-ether Neu5Ac2en derivatives. .	151
Table 4.1: Compounds evaluated as inhibitors of paramyxoviral HN.	158
Table 4.2: Inhibition (IC ₅₀ , μM) of hPIV-1 sialidase activity by designed C-4-ether and C-4-thioether substituted Neu5Ac2en derivatives.	159
Table 4.3: Comparison between biological hPIV-3 sialidase activity inhibition data and computational binding energy (Section 2.3.7.3) for 4- <i>O</i> -substituted Neu5Ac2en derivatives.	161
Table 4.4: Inhibitory effect (EC ₅₀ , μM) of 4- <i>O</i> -linked Neu5Ac2en derivatives on rSV (hPIV-1 HN) growth <i>in vitro</i>	163
Table 4.5: Comparison between <i>in vitro</i> hPIV-1 sialidase activity inhibition data and computational data (Section 2.3.7.3) for 4- <i>O</i> -substituted Neu5Ac2en derivatives.	163
Table 4.6: Comparison between <i>in vitro</i> hPIV-1 sialidase activity inhibition data and computational data (Section 2.3.7.3) for 4- <i>S</i> -substituted Neu5Ac2en derivatives.	165

Nomenclature

Within this thesis, the carbon atoms of Neu5Ac/Neu5Ac2en are numbered from 1 to 11 as shown below.



Within this thesis, 4-thioether linked Neu5A2en derivatives are presented in the text as 4-*S*-alkyl-Neu5Ac2en and not 4-*S*-alkyl-4-thio-Neu5Ac2en for simplicity. For example: the derivative 4-*S*-benzyl-4-thio-Neu5Ac2en is named within this thesis as 4-*S*-benzyl-Neu5Ac2en.

1 Introduction

1.1 Carbohydrates: A brief overview

Carbohydrates are a class of compounds that is comprised of many varied molecules, classified by the broadly generalised formula $(\text{CH}_2\text{O})_n$ where $n \geq 3$; it is from this formula that their name is derived.¹ Originally the role of carbohydrates was thought to be two-fold; the production of energy derived from monosaccharides and secondly, structural scaffolding in the form of polysaccharides.² A third role has also been attributed to carbohydrates; the role of cell-cell recognition. Cell-cell recognition involves a pair of complementary structures (ligand and receptor) at the cell surface that bind together; a number of classes of biomolecules are involved in this process. Carbohydrates, proteins and nucleic acids to name a few serve in this function.³

The structure of most carbohydrates is complex. This complexity is due to the poly-hydroxylated nature of carbohydrates which can form a vast number of glycosidic linkages; this allows the carbohydrates to form complex multiply branched polysaccharides. The general structure of carbohydrates is well suited to the role of information transfer; for example: 35,560 unique tetrasaccharides may be formed from four different monosaccharides, however only 24 unique tetranucleotides may be formed from four different nucleotides.³ This structural diversity allows the carbohydrates to carry more information per unit weight than either the nucleic or amino acids. This is not to say that carbohydrates exclusively are used for information transfer and cell recognition. As an example, an important class of proteins known as lectins (carbohydrate-recognising proteins) are involved in cell recognition. Hormones which are generally protein based are also utilised by the body for information transfer, and the nucleic acids make the basis of cellular genetic material, which transfers genetic information from one generation to the next.³

For pathogenic bacteria and viruses to infect a host, they must be able to interact with at least one receptor on the target host cell otherwise the pathogen would be cleared from the body rapidly *via* the host's gastric, circulatory and/or immune systems.³ A number of bacterial and viral pathogens utilise carbohydrates found on the cell surface, as shown in Figure 1.1, for this purpose. Within the group of pathogens that recognise cell surface carbohydrates, some preferentially recognise a specific class of carbohydrate known as the sialic acids.

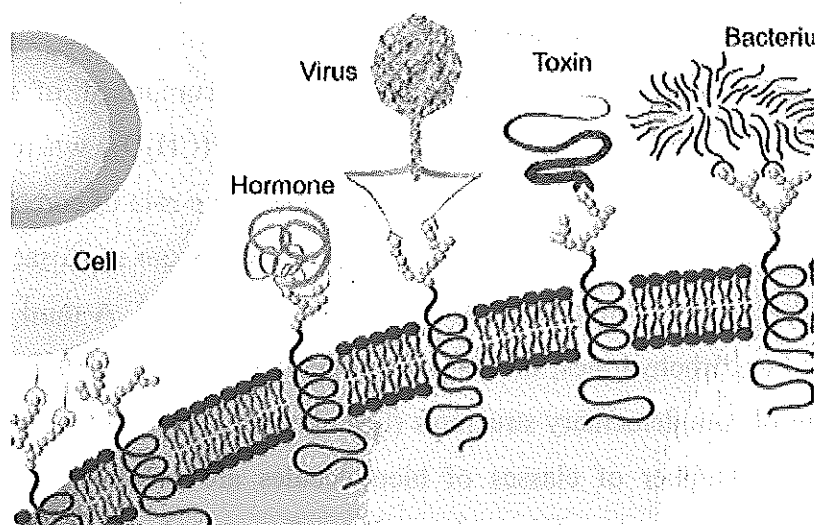
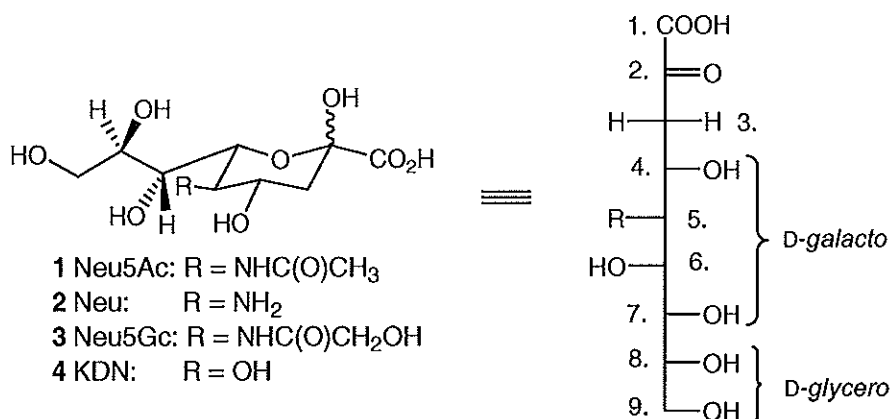


Figure 1.1: Utilisation of carbohydrates as a receptor for adhesion to cells.

1.1.1 Sialic acids: A general overview

Sialic acids are monosaccharides that contain a 9-carbon backbone with 5-acetamido-3,5-dideoxy-D-glycero-D-galacto-2-nonulosonic acid (**1**) as the core structure.⁴ The sialic acid family is large and diverse, with more than 50 naturally-occurring derivatives reported to date.⁵ The majority of the sialic acids are based on the 5-amino-5-deoxy parent that is known as neuraminic acid (Neu, **2**). The neuraminic acids are *N*-acylated to form either *N*-acetylneuraminic acid (Neu5Ac, **1**) or *N*-glycolylneuraminic acid (Neu5Gc, **3**),⁶ or the non-aminated 3-deoxy-D-glycero-D-galacto-2-nonulosonic acid (2-keto-3-deoxy-D-glycero-D-galacto-nononic acid, KDN, **4**).⁴ Most additional modification of the sialic acids typically involve *O*-acetylation at one or more hydroxyl groups on the carbohydrate ring or the glycerol side chain.⁷



Sialic acids have been found to have a widespread distribution throughout nature.⁵ They occur mainly as the terminal components of glycolipids and glycoproteins that are

found on the cell surface and play a major role in the chemical or biological activity of many glycoconjugates. Sialic acids are negatively charged at physiological pH and as such have a role in the charge of cellular membranes. Sialic acids have been shown to have roles as biological receptors for toxins, viruses, enzymes and bacteria.^{2,7}

Many derivatives of the sialic acid family have been synthesised, with many diverse modifications created on all carbons of the pyranose ring and the glycerol side-chain. A number of good reviews detailing the modifications performed have been recently published for the interested reader.^{4,8,9}

1.2 Sialic acid recognising proteins.

The receptors for sialic acids are commonly referred to as sialic acid recognising proteins (SARPs) and the proteins involved in pathogenic processes that utilise the recognition of sialic acids may provide attractive targets for drug design.² A number of SARPs have been identified.

In the mammals for example, a family of cell adhesion molecules consisting of three members have been identified known as selectins. The three members are known as E-selectin, L-selectin, and P-selectin.¹⁰ E-selectin is located on activated endothelia, L-selectin is located on leukocytes, and P-selectin is located on both activated platelets and endothelia. Another example of sialic acid recognising proteins is the Siglec family, of which there are 11 functional members and one Siglec-like molecule.⁵ Sialoadhesin (Siglec-1), located on specific macrophages; and CD22 (Siglec-2) located on B-cells of the immune system.¹⁰ Many Siglecs are involved in the human body's innate immunity.⁵ Complement Factor H, the first reported example of a SARP is also involved in the early stages of the innate immune system.⁵ SARPs have also been found in pathogenic microorganisms, influenza virus haemagglutinin is a well characterised sialic acid recognising glycoprotein. Influenza virus haemagglutinin is essential in the pathogenicity of influenza virus. It is involved in the initial attachment of the virus to the host cell, it is also involved in the penetration of the virus with the host cell by triggering fusion.¹¹ Another well characterised class of sialic acid recognising proteins are the sialidases which are discussed in the following section.

1.2.1 Sialidases

Sialidases are sometimes referred to as neuraminidases and the names are generally interchangeable, however a neuraminidase is more specifically targeted to the *N*-acetylneuraminic acids (C-5 acylamino derivatives).¹² They are exo-glycohydrolases that cleave the terminal α -ketosidically-linked sialic acids on cell surface glycoconjugates.¹²⁻¹⁴ Oligosaccharides with terminal KDN residues are resistant to *N*-acetylneuraminidases.¹⁵

The mammalian sialidases regulate the sialic acid profile of cells, and some cellular functions, such as apoptosis, cell adhesion and growth. They have also been shown to regulate red blood cell destruction and tumour metastasis.¹² Mammalian sialidases have been located membrane-bound, in the cell cytosol and in the extracellular compartment. Experimentally, they have been cloned from the cell lysosome, brain tissue,¹⁶ and more recently the human cell cytosol.^{17, 18}

Sialidases have been found to be involved in viral and bacterial pathogenesis. They mediate metabolism, adherence and infection. Not all sialidases are involved in pathogenicity, some bacteria utilise the enzyme for nutritional purposes, with terminal sialic acids cleaved from surface glycoconjugates being used as a carbon source.¹⁹ It has been suggested that the bacteria that possess sialidase activity have received the genes encoding the enzyme *via* horizontal gene transfer.¹⁶ A sialidase with a different function is located on the parasite *Trypanosoma cruzi*. It is a *trans*-sialidase, which is used by the parasite to transfer sialic acid from the host cells to the parasite. It also functions in the pathogenesis of Chagas' Disease.²⁰ Sialidases have not been found in plants. The only source of sialidase found amongst viruses is found within the myxovirus families; which includes human influenza virus, human parainfluenza virus, mumps, Sendai virus and Newcastle disease viruses amongst others.^{21, 22}

1.3 An introduction to Paramyxoviruses

The human parainfluenza viruses (hPIV) are members of the paramyxoviridae family of the genus paramyxovirus, as are the measles virus, Newcastle disease virus (NDV), mumps virus and respiratory syncytial virus.²³ Infection with hPIV usually causes upper respiratory tract symptoms, such as coughing and wheezing, however life threatening pneumonia can occur, and the fact that these viruses preferentially infect infants, is of major concern.²⁴ NDV is a highly infectious disease of birds which has

enormous impact, in terms of both economic impact and the number of infections/deaths per year.²⁵

The general structure of a paramyxovirus is a pleomorphic enveloped particle that is 150 to 300 nm in diameter.²⁶ The genome of paramyxoviruses are a single strand of non-segmented negative-sense RNA (approx 15kb), consisting of six or seven individual genes, which encode eight or nine different gene products, including fusion protein, nucleoprotein, large protein, matrix protein, phosphoprotein, and haemagglutinin-neraminidase (discussed in Section 1.3.3).²⁶ The nucleocapsid core is a filamentous, herringbone-like structure surrounded by a lipid envelope containing virus-specific glycoprotein spikes.²⁶

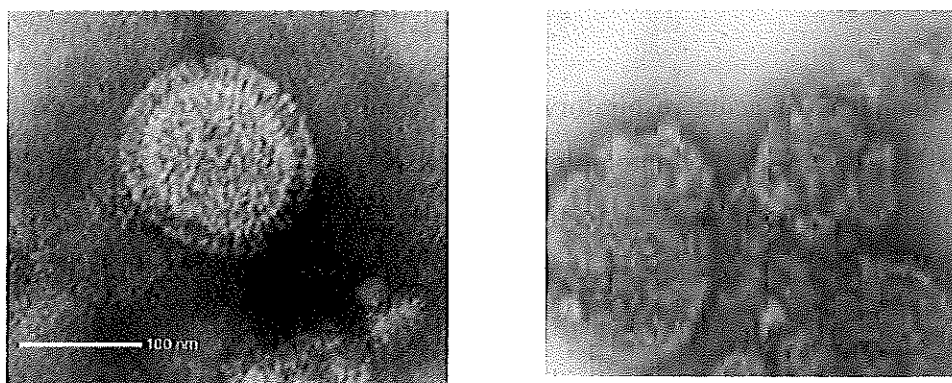


Figure 1.2: Electron micrograph of hPIV-3²⁷ (left) and NDV²⁸(right) virions

1.3.1 Human Parainfluenza Virus

There are four known subtypes of human parainfluenza viruses: 1, 2, 3 and 4. All four subtypes cause lower and upper respiratory tract infections in humans of all ages, but the most commonly infected with severe illness are children. The parainfluenza viruses are the second most important cause of lower respiratory disease in young children, second only to respiratory syncytial virus. The parainfluenza viruses can also reinfect older children and adults to produce upper respiratory disease.^{23, 24}

The first parainfluenza virus isolated from a human source was in 1956 and was designated “croup-associated” virus.²⁹ Another two were isolated in 1958 and the three known to exist were redesignated Para 1 (hPIV-1), Para 2 (hPIV-2) and Para 3 (hPIV-3).³⁰ Parainfluenza virus type 4 (Para 4, hPIV-4) was isolated in 1960.³¹

1.3.1.1 Clinical information

The transmission of the parainfluenza viruses is by large respiratory droplet spread or by direct person-to-person contact.²⁴ The viruses generally do not live well in the environment, but hPIV-1 virions have been collected from the air around an infected person, and up to 10% of hPIV-3 virions may be viable after 1 hour in an aerosol. hPIV has been demonstrated to infect animals such as hamsters, guinea pigs and ferrets.³² The incubation period of hPIV in human ranges from 3 to 6 days. A severe inflammatory response in the glottic tissues has been noted in a severe case of croup, and chemical mediators such as histamine may contribute to the observed inflammatory response. Viraemia has been reported with both primary viral infection and reinfection.²⁴

The diagnosis of acute respiratory disease is difficult but infections are usually diagnosed based on the symptomology and the known epidemiological pattern.²⁴ Laboratory diagnosis of hPIV infection is performed on respiratory secretions by visualisation due to cross-reactions occurring during serological tests. The collection of aspirations of washings of the upper respiratory tract can allow for visual identification of the virus in infected epithelial cells that have been shed.²⁴

The manifestations of primary parainfluenza infection are usually symptomatic and range from afebrile upper respiratory disease to severe, life threatening lower respiratory disease.^{24, 32} hPIV-1 causes croup or laryngotracheobronchitis in children, hPIV-2 also causes croup but serious disease states occur infrequently.^{24, 32} In rare cases, death has been reported with infection of the hPIV-1 and 2 virus. The clinical manifestations of hPIV-3 vary with the age of the affected patient. In infants less than 12 months old, bronchiolitis or pneumonia is found whereas infants from 6 to 18 months more commonly have croup. Older children generally present with tracheobronchitis and a severe infection generally results in pneumonia.²⁴ The understanding of the hPIV-4 subtype has been hampered by the lack of epidemiological data. hPIV-4 is found infrequently and the associated symptomology is generally of no consequence.³²

Maternal antibodies to the hPIV-1 and 2 viruses are generally accepted as the reason why these viruses do not cause serious disease in infants less than 4 months old.²⁴ The initial infection with hPIV-3 virus occurs early in life with 62% of infants

infected during the first year of life. This percentage is raised to 92% during the second year of life.²⁴ Children under the age of six are the most susceptible to parainfluenza virus infection, with males most frequently infected, and adolescent or adult infection is uncommon. After the age of six the sex difference disappears.²⁴

Epidemics of hPIV-1 generally occur in the autumn of every other year and hPIV-2 is also epidemic during the same period as hPIV-1 appearing to be as ubiquitous. hPIV-3 is generally considered endemic and occurs throughout the seasons.²⁴ No predictable pattern is detectable with respect to hPIV-3 occurrence although it has been noted that hPIV-3 infection can occur during the spring months when either hPIV-1 or 2 are active. hPIV-3 has also been found to be active during other viral epidemics. Having said this, hPIV-1, 2, and 3 virions can be isolated throughout the year but generally can be found during the periods specified above.²⁴

1.3.2 Newcastle Disease Virus

Newcastle disease virus (NDV), also known as avian paramyxovirus 1, is a member of the paramyxoviridae family, genus *Rubulavirus*. The first outbreaks were reported in poultry from Java, Indonesia and Newcastle-upon-Tyne in 1926.³³ Twenty-seven of the 50 orders of birds are reported²⁵ to be infected with NDV.³⁴ The virus also has a world wide distribution, causing severe losses to both governments and farmers each year.²⁵ An example of the impact of NDV is demonstrated by the fact that at least 90% of the infected poultry in Nepal is estimated to die each year as a direct result of NDV infection.³⁵ Humans exposed to NDV develop granular conjunctivitis, which can lead to some lasting impairment of vision. Other symptoms include headache, malaise, chills and lymphadenitis.³⁶

1.3.2.1 Aetiology

The virus is transmitted *via* ingestion or by the inhalation of infectious particles and three main pathotypes categorise isolates of NDV.²⁵

1. Lentogenic – Do not usually cause disease in birds.
2. Mesogenic – Viruses of intermediate virulence that cause respiratory disease.
3. Velogenic – Virulent viruses that cause high mortality split into two sub-categories.
 - a. viscerotropic – haemorrhagic lesions of the gastrointestinal tract
 - b. neurotropic – neurological disease

A velogenic viscerotropic infection is characterised by acute systemic illness with extensive necrosis of the spleen and intestinal lymphoid tissues. Infection by a velogenic neurotropic strain is characterised by central nervous system disease. Infection by the mesogenic or lentogenic strains causes no overt disease but may lead to decreased egg and meat production.³⁷ Surviving birds show impaired growth, poor food utilisation, reduced egg production, impaired eggshell formation, reduced fertility and hatchability of eggs. NDV is readily isolated during active infection, however the persistence of the virus has not been determined.²⁵ The clinical signs seen in a bird infected with NDV can vary widely depending on a number of factors such as the virus, age, host species, environmental stress, and infection with other organisms.³⁵

Analysis of birds that died in an outbreak of NDV in Canada in 1999 revealed that NDV and *Salmonella typhimurium* infected the brain and lung tissues. The infected birds showed symptoms of nervous system disease. The birds suffered from unilateral wing and leg paralysis. The strain which caused this outbreak was characterised as a mesogenic. It was suggested that pathogens such as *S. typhimurium* could influence the presentation of symptoms from the mesogenic strains of NDV.³⁸

Interestingly, the macrophages from chicken peripheral blood cultured *in vitro* with NDV show signs of apoptosis when inspected. The infected macrophages have also been found to support the growth and replication of NDV.³⁹

1.3.3 The haemagglutinin-neuraminidase of the paramyxoviruses

The entry of paramyxoviruses into the cells of the respiratory tract is mediated by two proteins; haemagglutinin-neuraminidase (HN) and fusion protein (F)^{40,41}, as shown in Figure 1.3. The haemagglutinin-neuraminidase protein is a type 2 glycoprotein. The transmembrane domain both anchors the protein and has a signal transduction function. Sequence analysis of the transmembrane region of NDV HN has shown that the region is the least conserved of the protein.⁴² Within the transmembrane region of NDV, three leucine residues in a heptad-repeat motif are found. This motif is found in most NDV isolates. Mutations of these residues have resulted in viruses with HN proteins with altered structure and/or function.⁴²

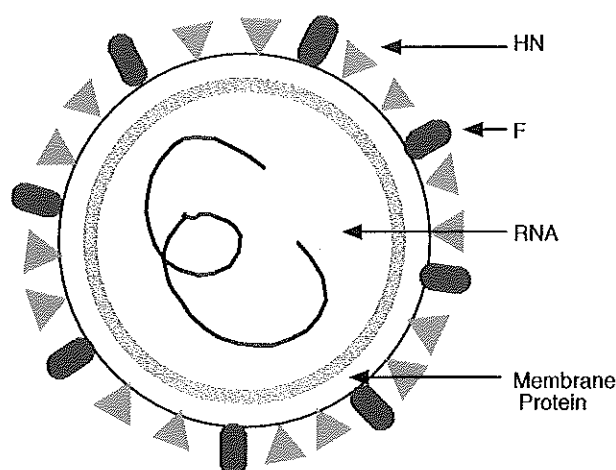


Figure 1.3: Schematic representation of parainfluenza virus, its genome and surface glycoproteins. HN = Haemagglutinin-neuraminidase, F = Fusion protein.

In humans and avians, HN binds to molecules on the target cells that contain Neu5Ac. The fusion protein mediates the actual process of the fusion of the virus with the target cell. The interaction of the HN glycoprotein is also required for the fusion of the virus.⁴⁰ The fusion protein is produced as a precursor and cleaved to its active form by a protease. In NDV it has been shown that the expression of the fusion protein and the HN protein to induce syncytium formation will only occur if the proteins are from homologous viruses.⁴³ It has been suggested that HN is involved in a virus-specific interaction with the fusion protein. Results imply that the fusion protein and HN contain a specific domain that allows the proteins to communicate with each other to promote fusion.⁴⁴ The removal of sufficient sialic acid residues (by the sialidase action of HN) to prevent cell fusion but still allowing the attachment of viral particles and cell infection lead to persistent infection.⁴⁰

The elucidation of dual roles of the haemagglutinin-neuraminidase glycoprotein has been the subject of research for many years. A single site model has been proposed, with the active site of the glycoprotein being flexible and undergoing a conformational change to induce either the receptor binding or receptor destroying activity.⁴⁵ A second sialic acid site has been identified on NDV HN that upon binding a sialic acid, is thought to drive the structural changes necessary for the HN protein to trigger the fusion protein, which in turn mediates the actual process of viral fusion.⁴⁶

The combined haemagglutinin-neuraminidase moiety performs the same roles as the separate molecules; haemagglutinin and sialidase found on the surface of influenza virus. Both the fusion and haemagglutinin-neuraminidase biomolecules have been

suggested to mediate viral fusion. The HN of hPIV consists of four identical subunits, arranged into dimers that are linked *via* disulfide bridges. The tetramer is linked to the viral coat *via* the N-terminal region of the protein.⁴¹

There is roughly 50 % sequence similarity between the different strains of parainfluenza haemagglutinin-neuraminidase, and some cysteine and glycine residues that are structurally important are conserved.²² Structural sequence alignment of the HN of both hPIV and NDV with influenza virus sialidase show that the important active site residues of influenza virus sialidase have analogues in HN.⁴⁷ The triarginyl cluster, putative catalytic aspartic acid, tyrosine residue which interacts the sugar ring of Neu5Ac, and glycerol side chain (2 interactions) in influenza virus sialidase are all present in parainfluenza virus HN.⁴⁷

1.3.3.1 Haemagglutinin-neuraminidase of NDV

The crystal structure of the multifunctional HN of NDV was solved to 2.5 Å in 2000.⁴⁸ The topology of the head of one monomer of the enzyme with bound inhibitor, 5-Acetamido-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-2-non-2-enoic acid (Neu5Ac2en), **5**, is shown in Figure 1.4. Two additional crystal structures were solved; one ligand-free and one with bound ligand Neu5Ac, **1**.⁴⁸

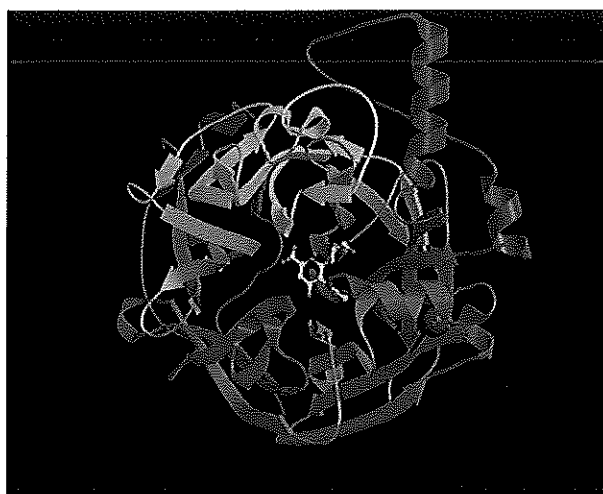
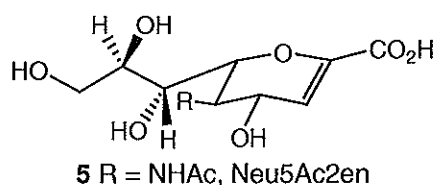


Figure 1.4: The structure of the NDV virus haemagglutinin-neuraminidase with bound inhibitor, Neu5Ac2en, **5**.^{49, 50}

The crystal structure of NDV HN shares many features with the other sialidases that have been crystallised.⁵¹⁻⁵³ The carboxyl group of the inhibitor **5** has been shown to interact with the Arg triad (Arg 174, Arg 416, and Arg 498) located in the active site. This interaction has been seen in many other bacterial and viral sialidases.⁵¹⁻⁵³ The hydroxyl groups of the glycerol side chain have been found to interact with three highly conserved residues: Glu 258, Tyr 317, and Tyr 262. The glycerol side-chain binding to three residues (two previously) has not been observed with other sialidases and may suggest that the glycerol side-chain hydroxyl groups may be a requirement of substrate recognition.⁴⁸ The residues Glu 258 and Tyr 262 are stabilised by a calcium ion and if the ion is removed enzymatic activity is lost.⁴⁸ The O-4 hydroxyl of Neu5Ac2en points into a large cavity lined with several invariant residues (based on alignment of Sendai, mumps, hPIV-1, 2 & 3, Simian virus 5, and NDV HN's) but does not have any interaction with the residues. The pocket is larger than the corresponding O-4 pocket seen in influenza virus sialidase.⁴⁸ The primary sequence alignment of NDV HN with influenza virus sialidase has yielded 7.1 to 10% similarity.⁵⁴

Sialidase activity of NDV HN

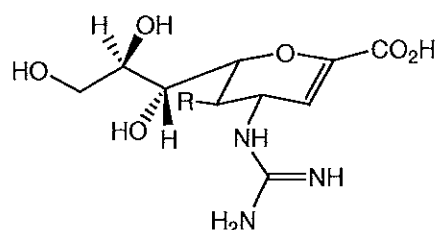
In the case of paramyxoviruses, most information available about the receptor determinants is regarding Sendai virus or Newcastle disease virus. Receptor determinants are *N*-acetylneuraminic acid $\alpha(2-3)$ linked to a galactose moiety, with disialo Neu5Ac $\alpha(2\rightarrow8)$ Neu5Ac $\alpha(2\rightarrow3)$ Gal groups being more efficiently recognised.²² Recognition of sialic acids is determined by several factors: the type of sialic acid, the linkage of the sialic acid, the oligosaccharide structure, the number of sialic acids molecules present and the arrangement of the sialic acid moieties.²²

The optimal pH of the NDV HN for sialidase activity is found within the range pH 5.1 – 5.3, most strains also have a second pH optimum; at pH 6.3.⁵⁵ The K_m of NDV HN catalysing the cleavage of sialic acid from *N*-acetylneuraminy1- $\alpha(2\rightarrow3)$ lactose is 1.06×10^{-3} M with a V_{max} of 6.70 μ mol of sialic acid released per minute. In comparison the K_m of influenza A virus sialidase catalysing the same reaction is 1.25×10^{-3} M and the V_{max} is 4.70 μ mol of sialic acid released per minute.⁵⁶ K_m is a measure of the substrate concentration required for effective catalysis to occur.¹ The fact the K_m of NDV HN is lower than that of influenza A virus sialidase means that NDV is more effective at catalysing the hydrolysis of the substrate. The higher the V_{max} of an enzyme results in more substrate that can be processed in a given time.¹ So with the above

results NDV requires lower concentrations of substrate for effective catalysis and can produce more product per unit time than influenza A virus sialidase.

1.3.3.2 Haemagglutinin-neuraminidase of hPIV-3

The crystal structure of the multifunctional HN of hPIV-3 was solved to 2.8 Å in 2004.⁵⁷ The topology of the head of the enzyme with bound inhibitor, Neu5Ac2en, **5**, is shown in Figure 1.5. Four additional structures were solved, two are ligand-free, one contains the ligand Neu5Ac, and another contains the inhibitor 4-deoxy-4-guanidino-Neu5Ac2en, (**6**).⁵⁷



6 R = NHAc



Figure 1.5: The structure of the hPIV-3 haemagglutinin-neuraminidase with bound inhibitor, Neu5Ac2en.^{49, 50}

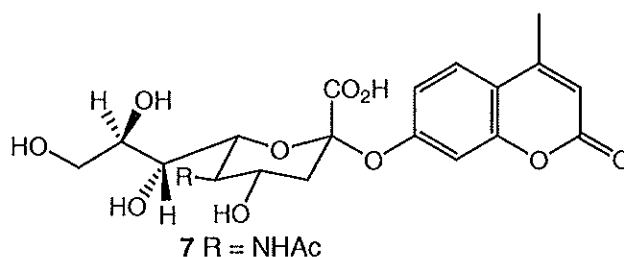
Sequence analysis of hPIV with influenza virus NA suggested that the tertiary fold of the globular head of hPIV3 HN would conform to a six-blade anti-parallel β -sheet propeller (analogous to that found in influenza virus sialidase).⁴⁷ The crystal structure of the hPIV3 HN monomer has indeed shown this to be the case. The sequence identity of the HN of hPIV-3 is 24% over the globular head domain in comparison to NDV HN.⁵⁷ The carboxyl group of the inhibitor Neu5Ac2en, **5**, has been shown to interact

with the Arg triad located in the active site and structural changes in the active site occur to accommodate the ligand.⁵⁷ The side-chain of Arg 192 is positioned deeper in the active site and creates the potential of four hydrogen bonds. The C-7 and C-9 hydroxyl groups of the glycerol side chain have been found to interact with Glu 276. The methyl group of the C-5 acetamido chain forms a hydrophobic interaction with Tyr 319. Analysis of bound Neu5Ac, **1**, shows the presence of a crystallographically observed water molecule beneath the sugar, and also shows that the C-4 OH group and the nitrogen of the C-5 acetamido group make interactions with the surface of the protein via hydrogen bonding with water molecules.⁵⁷ Upon binding of a ligand no significant structural changes are seen outside the active site.⁵⁷

Sialidase activity of hPIV HN

Recently, a study of the receptor determinants of hPIV-1 and hPIV-3 have shown that hPIV-1 HN preferentially binds to $\alpha(2-3)$ linked Neu5Ac and not $\alpha(2-6)$ linked Neu5Ac, in contrast hPIV-3 HN recognises $\alpha(2-3)$ linked and $\alpha(2-6)$ linked Neu5Ac.⁵⁸

hPIV-1 HN was expressed and characterised in 2001.⁵⁹ The optimal pH for recombinant purified hPIV-1 HN sialidase activity is 4.5 and the binding affinity for the fluorogenic substrate 4-methylumbelliferyl-*N*-acetylneuraminide, MUN, **7**, was found to be 234 μ M, in comparison the optimal pH for native HN obtained from the hPIV-1 lysate is 5 and the binding affinity of **7** was found to be 176 μ M.⁵⁹



1.4 Current prevention and treatment of paramyxovirus infection

1.4.1 Vaccines

Vaccines against the parainfluenza viruses have been developed in the form of monovalent, trivalent or multivalent formalin-inactivated, parenterally administered vaccines. All trialled vaccines failed to show protection even if the immune system was adequately activated.^{24, 60} Recently, a live attenuated chimeric recombinant parainfluenza virus vaccine has been shown to be protective in the experimental

infection of hamsters.⁶¹ The virus was produced *via* reverse genetics to produce a virus that encodes the internal proteins of hPIV-3 and the external proteins of hPIV-1. The vaccine candidate (known as rPIV3-1.cp45L) has been shown to induce complete resistance to hPIV-1 infection and confer partial resistance to hPIV-3 infection.⁶¹ The partial resistance is found to be short lived and disappears after four months.⁶² However, resistance induced by prior infection with hPIV-3 is found to remain high after four months. rPIV3-1.cp45L has been shown to be immunogenic to hPIV-1 in animals even if the animals are immune to hPIV-3, this study demonstrated that a sequential immunization schedule is feasible with rPIV3-1.cp45L.⁶²

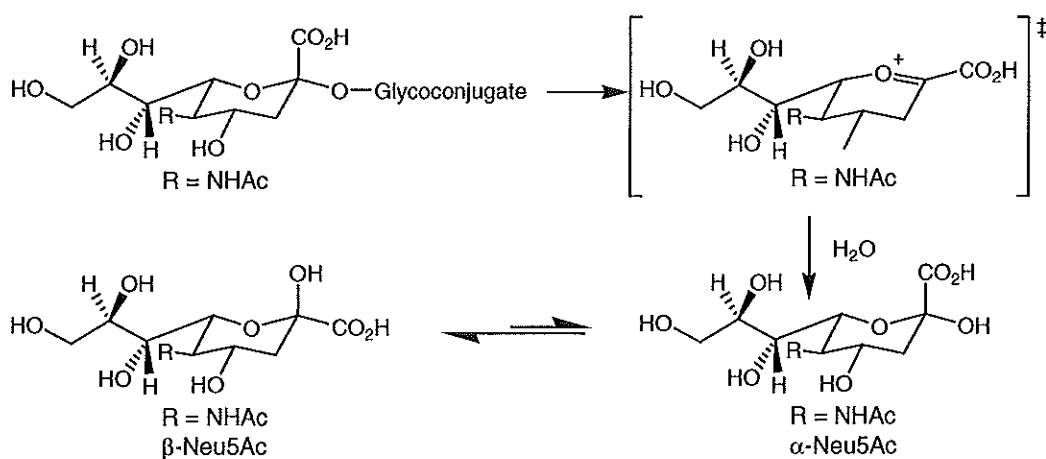
In the case of NDV, currently the control of NDV in poultry is aimed at vaccinating the bird population. Effective vaccines against NDV have been developed, the most widely used efficacious live-virus vaccines for prevention of NDV are the distribution of mildly virulent B1 and La Sota strains in food, water or by large droplet spread.²⁵

1.4.2 Therapeutic agents

An extensive search of the publicly available literature has shown that currently there is no drug on the market that can effectively treat paramyxovirus infections. The use of ribavirin has been limited due to expense and unproved therapeutic efficacy.⁶³ However recently, a number of drugs for the effective treatment of influenza virus infection have become available.^{14, 64} These drugs are known as sialidase inhibitors.

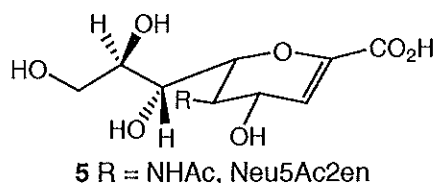
1.4.2.1 Sialidase inhibitors.

Many enzymes in order to cleave biological substrates induce a conformational change in the substrate to make it more like the transition state of the catalysed reaction. The active site residues are positioned so as to create the most efficient binding to the transition state. As such transition state analogues are one of the best for the design of therapeutic drugs to interact with an enzyme.⁶⁵ The reaction mechanism of the influenza virus sialidase showing the proposed⁶⁶ structure of the transition state (proceeding via S_N1 mechanism) is provided in Scheme 1.1 An alternate reaction mechanism utilising a mixed S_N1 - S_N2 mechanism that gives a covalently bound intermediate has also been proposed.⁶⁷

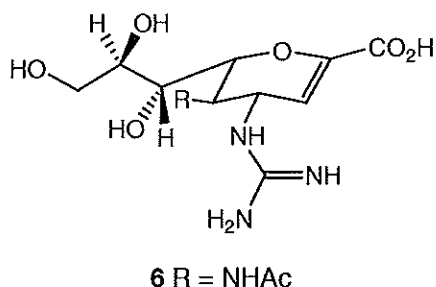


Scheme 1.1: Proposed S_N1 mechanism for influenza virus sialidase.⁶⁶ The putative proposed transition state structure is shown in the square brackets.

The first derivative of *N*-acetylneuraminic acid to be synthesised that potentially inhibited sialidases 5-acetamido-2,6-anhydro-3,5-dideoxy-D-glycero-D-galacto-2-non-2-enoic acid (Neu5Ac2en, **5**) was prepared in the mid-60's to inhibit sialidases. Neu5Ac2en is a transition state analog by virtue of it's flattened half-chair structure.¹³ This compound was found to have good inhibitory effect on a number of sialidases (with K_i range of 10^{-5} to 10^{-6} M), however it lacked specificity for any one sialidase.^{13, 68, 69} This compound has been used as a template to design new potent, selective inhibitors of the influenza virus sialidase.

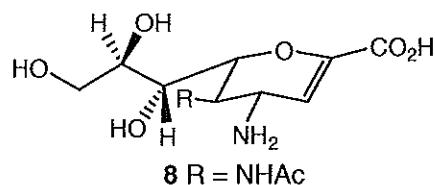


A potent and selective influenza virus sialidase inhibitor ($K_i \sim 10^{-11}$ M) was developed in the early 1990's. The compound is known as 4-deoxy-4-guanidino-Neu5Ac2en (**6**) or by it's tradename Zanamivir.^{13, 14, 70}



Zanamivir, **6**, was the result of a decade of collaborative research by a small number of Australian research groups. The solving of the influenza virus sialidase structure in 1983⁵¹ allowed the von Itzstein research group to begin looking for potential

inhibitors of the enzyme. The von Itzstein group's research centred on modifying the compound Neu5Ac2en, **5**, to develop a molecule that tightly binds to the enzyme's active site thus blocking activity. The result of their research was Zanamivir (**6**).^{13, 14} Modelling studies initially indicated that the incorporation of an amino functionality at C-4 of Neu5Ac2en should improve binding.⁷¹ 4-Amino-4-deoxy-Neu5Ac2en, **8**, showed a 100-fold improvement in potency over Neu5Ac2en, **5**.¹⁴



Extending the substitution at C-4 to the guanidino functionality gave the potent and selective inhibitor, Zanamivir (**6**).¹⁴ The specificity of **6** for influenza virus sialidases arises from the fact that two conserved glutamic acid residues (Glu 119 and Glu 227) are located in the O-4 pocket to which the guanidino group binds. These residues are lacking in bacterial sialidases.⁷² Indeed, in the case of *Vibrio cholerae* sialidase the O-4 pocket is too small to accommodate the size of the guanidino functionality.⁵³

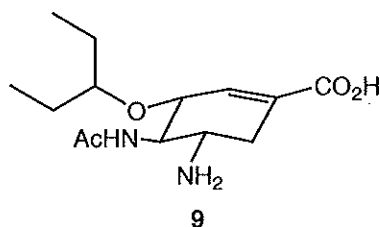
Zanamivir has been licensed in a number of countries for the treatment and prophylaxis of influenza A and B strains. The oral bioavailability of Zanamivir is low, however topically administering the drug by inhalation has solved the problem.^{13, 73} Twice daily inhalation of 10 mg of Zanamivir for five days has been found to be highly effective in the treatment of symptoms.^{74, 75} The plasma half-life of Zanamivir ranges from 2.5 to 5.1 hours, with the compound being excreted in the urine unchanged within 24 hours.^{74, 75}

1.4.2.2 Zanamivir mimetics

From the time of the discovery of Zanamivir, other research groups have been searching for other sialidase inhibitors as potential compounds for anti-influenza therapeutics. The result of this research has been mixed, with many compounds being synthesised but only a handful showing similar activity to that of Zanamivir.⁷⁶ Most compounds developed using a modified structure of Zanamivir have resulted in inhibitors that are weaker than the starting compound.¹³

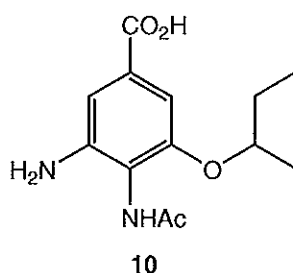
Zanamivir mimetics based on non-carbohydrate core structures have yielded another potent selective inhibitor of influenza virus sialidase. GS4071 or Oseltamivir

[4-acetamido-5-amino-3-(1-ethylpropoxy)-1-cyclohexene-1-carboxylate, **9**] is a carbocyclic mimetic of 4-amino-4-deoxy-Neu5Ac2en derivative, **8**, with potency equal to that of Zanamivir ($IC_{50} = 1 \times 10^{-9}$ M).^{13, 64, 73} Key alterations to the structure are the position of the double bond and the replacement of the glycerol side-chain with a hydrophobic 3-pentyl side-chain. The hydrophobic side-chain leads to an inhibitor more potent against influenza virus A than B.⁷⁷

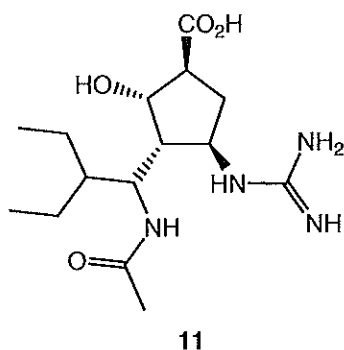


Oseltamivir, **9**, differs from Zanamivir, **6**, in that it can be administered orally as a prodrug, which is the ethyl-ester of Oseltamivir. Hepatic enzymes within the liver convert the prodrug to its active form.¹³ Twice daily administration for five days is the normal treatment regime, with eighty percent of the 75 mg oral dose converted into the active form.⁷⁴ The plasma half-life of Oseltamivir ranges from 6-10 hours. The major route for the elimination of Oseltamivir is excretion unchanged in the urine.^{70, 74}

In 1999 another inhibitor of influenza virus sialidase based on planar benzoic acid was reported. The inhibitor 4-(*N*-acetylamino)-5-amino-3-(3-pentyloxy)benzoic acid (**10**) was found to be highly selective for influenza A sialidase (IC_{50} 1 μ M) compared to influenza B sialidase (IC_{50} 500 μ M).⁷⁸

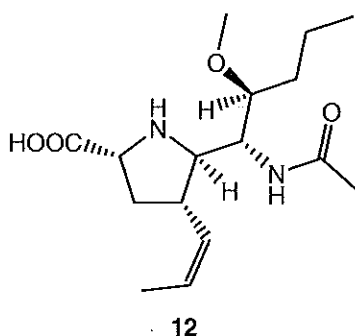


More recently Biocryst Pharmaceuticals, Inc reported the inhibitor BCX-1812 (3-(1-acetamido-2-ethylbutyl)-4-guanidino-2-hydroxycyclopentane-1-carboxylic acid, **11**).⁷⁹ In the mouse model, orally administered BCX-1812 was found to inhibit lethal concentrations of influenza virus at doses as low as 1 mg/kg/day, with a lack of toxicity. The compound was expected to enter phase III clinical trial in the northern winter of 2000-2001.⁷⁰ However, in 2002 BCX-1812 was removed from clinical trials due to lack of statistically significant reduction in the severity of influenza symptoms.⁸⁰



Interestingly the structure of 4-amino-4-deoxy-Neu5Ac2en, **8**, and the Zanamivir mimetics **9**, **10** & **11**, while different to Zanamivir in primary structure, all contain key elements of the Zanamivir template. All three mimetics mentioned above maintain the carboxylic acid and NHAc groups found in Zanamivir. Another point of note is that **10** and **11** also contain the hydrophobic side chain that is located on Oseltamivir. This side-chain binds in the same pocket as the glycerol side chain, but induces a change in the shape of the protein in influenza virus A sialidase to produce a hydrophobic pocket. This conformational change occurs to a lesser extent in influenza virus B sialidase leading to the difference in the activity against influenza virus A and B sialidase.⁷⁷

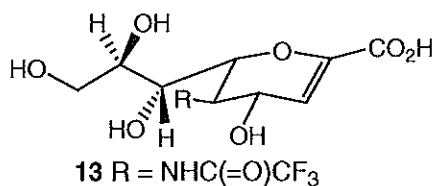
A highly potent novel pyrrolidine-based compound was developed by Abbott Laboratories in the US in 2002, 5-[(1*R*,2*S*)-1-(Acetyl-amino)-2-methoxy-2-methylpentyl]-4-[(1*Z*)-1-propenyl]-(4*S*,5*R*)-D-proline, (A-315675), **12**, has been shown to inhibit influenza virus A (N1, N2, & N9) and B sialidases in the order of 0.024 to 0.34 nM (K_i).⁸¹



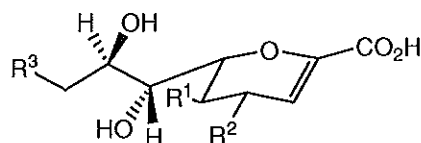
1.4.2.3 Inhibition of paramyxovirus haemagglutinin-neuraminidases

There is not a great deal of data available with regard to the inhibition of paramyxovirus HN by sialic acid analogues or derivatives. The data available relates to the inhibition of hPIV-2 by 4-substituted Neu5Ac2en derivatives (K_i of 8×10^{-4} M for 4-deoxy-4-guanidino-Neu5Ac2en, **6**, tested against hPIV-2).⁶⁹ Also 5-modified

Neu5Ac2en derivatives have been tested against NDV HN (K_i of 1.9×10^{-6} M for 2-deoxy-2,3-dehydro-*N*-trifluoroacetylneuraminic acid, **13**, tested against NDV).⁶⁸



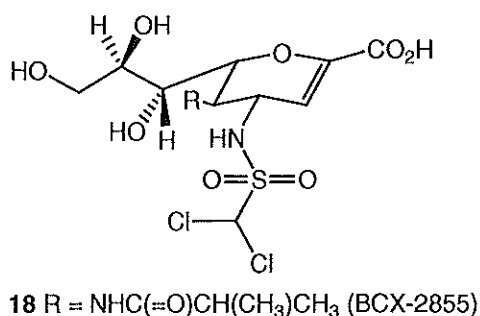
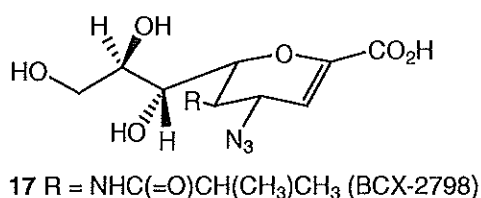
Both of these studies also tested the same derivatives against influenza virus sialidase for comparison.^{68, 69} The inhibition of 4-deoxy-4-guanidino-Neu5Ac2en, **6** against hPIV-3 has also been investigated and found to inhibitory activity in the order of 2.5×10^{-4} M (IC_{50}).^{48, 82} Suzuki *et al.* reported mono and di-substituted derivatives of Neu5Ac2en as inhibitors of hPIV-1 HN in 2001. Inhibition exhibited ranged from 1.0×10^{-5} M (IC_{50}) with 5-acetamido-4-*O*-thiocarbamoylmethyl-2,3-didehydro-2,3-dideoxy-D-glycero-D-galacto-2-nonulopyranosic acid, **14**, to 10×10^{-5} M (IC_{50}) with 5-acetamido-4-*O*-amidinomethyl-2,3-didehydro-2,3-dideoxy-D-glycero-D-galacto-2-nonulopyranosic acid, **15**.⁸³



- 14** R¹ = NHAc, R² = OCH₂C(=S)NH₂, R³ = OH
15 R¹ = NHAc, R² = OCH₂C(=NH)NH₂, R³ = OH
16 R¹ = NHAc, R² = OCH₂C(=S)NH₂, R³ = NHAc

The disubstituted 5,9-diacetamido-4-*O*-thiocarbamoylmethyl-2,3-didehydro-2,3-dideoxy-D-glycero-D-galacto-2-nonulopyranosic acid, **16**, inhibited hPIV-1 HN with an IC_{50} of 4.2×10^{-5} M.⁸³ Greengard *et al.* have reported that Zanamivir, **6**, Neu5Ac2en, **5**, and 4-amino-4-deoxy-Neu5Ac2en, **8**, inhibits cell fusion mediated by hPIV-3 HN.⁸² **6** was found to be most effective (IC_{50} 0.19 mM) compared to **5** (IC_{50} 3.2 mM) and **8** (IC_{50} 4.0 mM).⁸²

Alymova *et al.* recently (2004) reported two novel inhibitors of hPIV 1, 2 & 3 HN, BCX-2798 (4-azido-5-isobutyrylamino-2,3-didehydro-2,3,4,5-tetradeoxy-D-glycero-D-galacto-2-nonulopyranosic acid, **17**) and BCX-2855 (4-dichloromethanesulfonylamino-5-isobutyrylamino-2,3-didehydro-2,3,4,5-tetradeoxy-D-glycero-D-galacto-2-nonulopyranosic acid, **18**).⁸⁴



Compound **17** is found to inhibit the sialidase action of hPIV-1, 2, 3 HN with an IC₅₀ of 0.04, 1.6 & 20 μM and the haemagglutinin activity with an IC₅₀ of 0.1, 2.2 & 4.8 μM, **18** has an IC₅₀ of 1.2, 1.9 and 4.3 μM and 6.0, 2.7 & 2.0 μM respectively.⁸⁴ These two compounds represent the most potent inhibition of hPIV HN reported in the examined literature to date.

1.5 Aims and objectives

The aim of this project was to design novel, potent sialidase inhibitors of hPIV. At the commencement of this research project the crystal structure of hPIV HN was not available and therefore the crystal structure of the HN from NDV was probed *via* computational modelling as a model system for hPIV. The crystal structure of NDV theoretically can be used to model inhibitors for the treatment of hPIV, because there is good conservation of amino acid residues about the active site.⁴⁸ Subsequent to the design phase of this project, a crystal structure of hPIV-3 HN was published.⁵⁷ This allowed for the validation of the designed compounds and the confirmation of the hypothesis the NDV HN could be used in place of hPIV HN.

- Computer programs such as Insight II, AutoDock and GRID were used to characterise the potential interactions of the binding/active site of the HN glycoproteins. Once the interactions were characterised, potential ligands were *de novo* designed using Neu5Ac2en, **5**, as a seed template. The ligands that were designed were refined into suitable synthesisable structures and the ligands were re-scored and re-ranked. Other work performed involved the development of alternate starting templates (Chapter 2).
- The designed ligands were chemically synthesised (Chapter 3).
- The synthesised ligands were tested using screening assays against the HN's of hPIV-1 and hPIV-3. With successful binding of designed ligands to the HN

protein, a crystallography study on one ligand was performed to ascertain the ligand's conformation, which was subsequently used to validate computational docking experiments (Chapter 4).

2 Computational analysis of the paramyxoviral HN glycoprotein

2.1 Introduction

Computer-aided molecular modelling is a term that encompasses the calculation, visualisation and analysis of the molecular properties of the target of interest. Molecular modelling utilises force fields to explain the interactions between atoms, such as electrostatic and van der Waals interactions. Force fields also contain parameterisation to explain interatomic bond lengths, angles and torsions. These parameters allow the *in silico* construction of a model of a target (small or macromolecule). It was the discovery that similar molecular fragments were found to possess similar properties in varying environments that allowed for the creation of force fields to parameterise molecules.⁸⁵

The three dimensional structural elucidation of proteins has become an area of intense interest, with respect to the design of novel therapeutics, especially through the engagement of computational modelling. Structural elucidation can take many forms including X-ray crystallography, homology modelling, NMR spectroscopy experiments⁸⁶ and electron crystallography.⁸⁷ The 3D structure of a protein allows one to visually inspect the binding domain, and to design ligands that fit into the shape, hopefully with greater affinity than the natural substrate.⁸⁶

Generally, the first step taken in the computational characterisation of a protein is the validation of the structure. This is typically done by visualising the protein, using a visualisation program such as InsightII, and using programs such as Reduce,⁸⁸ Procheck,⁸⁹ Whatcheck,⁹⁰ or Hbplus⁹¹ designed to compare the experimental variables to those considered ideal. The program Reduce is used to evaluate and optimise the internal H-bond network of a protein by rotating the side-chains of the amino acids asparagine, glutamine and histidine. This is necessary due to the ambiguity in assigning these amino acid side chains in the X-ray crystal structure electron density. Procheck is utilised to assess how normal or unusual the geometry of the residues of the protein are.⁸⁹ The purpose of Whatcheck is to verify if the crystal structure proposed is indeed possible, given that torsion angles, interatomic distances etc. for a given pair of atoms are restricted to a finite range.⁹⁰ Hbplus checks (by measuring the distances and angles between H-bond donors and acceptors) the H-bond network of a protein.⁹¹ This

analysis will characterise the initial state of the structure and the suitability of the structure for computational modelling.

The next step in the *in silico* characterisation of a protein will normally involve the characterisation of the electrostatics of the protein, to determine potential areas of interaction. A program used to determine the electrostatics of a molecule is DelPhi.⁹² The electrostatic potential of a protein arises from the unequal distribution of electrons within atomic bonds; this is due to the withdrawing effect of electronegative elements. Ionised functional groups also play a substantial role in the charge distribution of a protein. This unequal distribution of electrons between the atoms of the protein is represented as a fractional charge on each atom of the protein. The electrostatic potential at a point on the protein is the force acting on a unit positive charge placed at that point on or surrounding the protein.⁹²

The areas of potential interaction of a protein are further characterised by the use of programs designed to probe the surface of the protein for areas of hydrogen bond donor or acceptor sites. Also probed for are areas of van der Waals (hydrophobic) interaction. A program designed to probe a protein in this fashion is GRID.⁹³ The Molecular Discovery suite of programs (GRID) is utilised to identify optimum areas of interaction of a receptor structure with a set of functional group probes. Since being first described in 1985,⁹³ the suite of programs have been used most notably in the successful design of the commercially available anti-influenza drug Relenza®.¹⁴ GRID predicts specific non-covalent interactions (steric and electronic) between a known protein structure and a probe.⁹⁴ Multiple sets of probes are available with GRID version 22.⁹⁴ The probes consist of traditional single atom probes containing one heavy atom, multi-atom probes containing more than one heavy atom, and special probes that include covalently-bound halogen atoms. The use of GRID to predict areas of interaction is highly advantageous in the derivatisation of a lead template in the drug discovery process. The use of this program in this work will allow for the focused analysis and discovery of potential inhibitors of the HN glycoprotein.

With the interactions of the protein characterised, the next step in the computational design process is to design potential ligands. This can be done in a number of ways, one of which is the *in silico* screening of large databases to obtain potential compounds that fit the sites of interaction. A program ideal for this method is DOCK.⁹⁵ DOCK

(version 4) is a molecular docking program that is used to explore the possible conformations of a ligand within a protein binding site. It assumes that the protein is rigid, the ligand on the other hand can be considered flexible (with the exception of rings). DOCK calculates the negative image (available conformational space) of the binding site with a series of binding site points. The ligand is then matched to these points (and optimised) using a process which generates multiple orientations of the ligand. Each conformation is given a score based upon the intermolecular interaction as defined by a molecular mechanics force field. The molecular docking algorithm is fast and allows the searching of databases of compounds.⁹⁶ *In silico* screening is especially useful if no ligand information is available during the initial computational modelling phase. In this study, however as a lead structure was known and available with the X-ray crystal structure, *in silico* screening was not utilised.

An alternate method of ligand design is fragment-based *de novo* design. *De novo* design involves the mapping of sites of potential interaction within the binding site, followed by the creation of ligands using fragments to form grow a core structure or by linking separate fragments together. One program that does this is LigBuilder.⁹⁷ LigBuilder is utilised in three steps: characterisation of the protein (Pocket), the ligand creation (Grow/Link), and processing of ligands for suitability (Process).

Pocket is utilised to characterise the interactions of the protein around a pre-docked ligand.⁹⁷ The site is probed with hydrogen bond donor and acceptor probes and a hydrophobic probe. The probing also identifies areas that cannot be built into, ie: areas too close to the protein surface. Once the active site has been characterised by Pocket, the next step is the creation of the ligands using either a grow or link strategy.⁹⁷ Both strategies utilise a fragment-based algorithm to build the ligands. Upon the formation of a bond between the user specified core fragment and the selected growing fragment, the newly formed bond is energy minimised by the rotation of the fragments to form the optimal interaction with the protein and to reduce intramolecular steric energy. The process of energy minimisation around the dihedral angles of the ligands allows for the flexibility of the ligand to be considered during the calculation. The Grow strategy “grows” from a user selected core fragment until the user specified criteria is satisfied. The Link strategy “links” multiple user selected core fragments with fragments from the library, which can be expanded with user defined fragments.⁹⁷

The output from Grow or Link needs to be processed by Process before it can be viewed or utilised in further calculations. Process has a parameter file that contains user specified criteria to select the most suitable ligands from the generated population of structures. These criteria are based on the Lipinski rules,⁹⁸ and as such take into account drug oral bioavailability. The library of suitable structures is characterised by the molecular formula, predicted octanol/water coefficient, and predicted binding score.⁹⁷ *De novo* ligand design is useful if a lead structure is known and derivatives of the lead are required, and as such was envisaged to be very useful in this particular study as a lead structure was already known.

A technique used to improve structures or structural interactions is the minimisation of structures. The minimisation of the protein is utilised to shift the atoms of a protein to the nearest local minimum, although this is not necessarily the global minimum of the system. The conformation generated may or may not be a conformation adopted at 0 K. Minimisation relieves the strain placed on residues within the structure (by reducing bad van der Waals interactions), and can help to relieve some errors found in the crystal structure.⁹⁹ Two programs are used to minimise proteins and protein/ligand assemblies: Discover⁹⁹ and Sander¹⁰⁰ (part of AMBER 7). The use of minimisation in this study will allow some protein flexibility in the calculation process once the ligands have been docked and thus should hopefully result in an increase in the energy of interaction of the protein ligand complex. Both Discover and Sander can also be used to perform molecular dynamics. Molecular dynamics provides a simulation of the motion of atoms at a given temperature and therefore sample the flexibility of molecules.⁹⁹ The energy of interaction of a ligand to a macromolecule can also be calculated using molecular mechanics.^{99, 100} Molecular dynamics is a computationally expensive procedure and it was envisaged that the technique would not be employed in a large way during this study.

Once the ligands have been designed they need to be ranked and characterised based on their ability to bind or dock to the protein structure of interest. Computational molecular docking involves finding the low-energy binding states between a ligand within an active site of the protein. This docking procedure requires accurate modelling of the force field parameters to explain the interactions of the systems, as well as an efficient search algorithm to examine the potential binding modes of the ligand.⁹⁶

Consensus scoring is utilised to obtain greater positive hits in the score of databases of molecules.¹⁰¹ Consensus scoring involves the use of a number of molecular docking programs to score the molecules of interest. The scores are then ranked and molecules that consistently rank in the top scores between a large proportion of programs are included in future modelling calculations/consideration in the design process of novel ligands.¹⁰¹ The use of multiple programs will potentially also reduce the number of false positives as individual programs are parameterised differently. The differences in parameterisation will test the properties of the molecules in different ways, such as the use of different force-fields to explain bond and torsion parameters, charges, and charge-charge or hydrophobic interactions. These differences will potentially find the false positives, if a molecule scores well in one program but does not score well in the others. There are many programs that dock or score or do both of which, AutoDock¹⁰² and X-Score¹⁰³ are examples.

AutoDock is a molecular docking and scoring program. That is, it is utilised to flexibly dock small molecules into the active site of a protein, and can also be used to score the static (rigid) conformation of a ligand. The program uses a genetic algorithm to reduce the time to search the conformational space within an active site.¹⁰² Upon small molecule binding the program can estimate the free energy of binding.¹⁰² The receptor is considered static throughout the calculation in AutoDock version 3.05, however the ligand can be considered either flexibly or statically (to score a previously obtained conformation).¹⁰² The root ring structure of the ligand is considered rigid throughout the calculation and thus careful thought about the ligand conformation before commencing the calculation is required. The ability of this program to dock flexibly and score statically is highly advantageous for this study within the consensus scoring routine as it will allow the flexible docking of the ligands and the rescoring of the ligands after energy minimisation. This will allow the examination of the protein-ligand complex energy of interaction both before and after the process to examine if the interactions are improved during in the energy minimisation process.

X-Score,¹⁰³ formally known as X-CSCORE,¹⁰⁴ is a program that utilises a non-force field empirical method to estimate the binding affinity of a protein-ligand complex. The interactions of each atom of a ligand with the protein are classified as good, bad, or neutral by X-Score. The classification is based on an atomic binding score. This was considered a good program to include in the consensus scoring routine

as it is non force field based and as such is not biased by the force-field parameterisation and will give a different perspective on the protein-ligand complex interaction.

2.1.1 Structure-based drug design

2.1.1.1 *Influenza virus sialidase*

The process of structure based drug design has been used to develop inhibitors of many classes of enzymes, the most pertinent example in relation to this project being influenza virus sialidase.¹⁴ The techniques used in the structure-based drug design of influenza virus sialidase inhibitors X-ray crystallography,⁵¹ various computational techniques,⁷¹ chemical synthesis^{105, 106} and biological evaluation. A brief discussion of influenza virus sialidase is presented here to highlight the information that could be obtained during a structure based drug design study.

Influenza virus sialidase has two postulated biological roles. The first is that it helps the invading virion through the mucus of the respiratory tract, by removing sialic acid from the mucins and sialoglycolipids, thereby facilitating access of the virion to the target epithelial cells. Secondly, it may assist the nascent virions to elute from the infected cell; it has been postulated that if the sialic acid residues found on the cell surface were not removed the nascent virions would clump together and would be easily cleared by the body's immune system.^{11, 107, 108}

The X-ray crystal structures of the sialidase from influenza A (N2) strains A/Tokyo/3/67 and A/RI/57/57 were solved to 2.9 Å by Varghese *et al.* in 1983.⁵¹ The X-ray crystal structure of influenza B sialidase from B/Beijing/1/87 solved to 2.2 Å, was determined in 1991.¹⁰⁹ Subsequently a large number of structures at the time of writing have been solved of both apo enzymes and complexes with inhibitors.

The sialidases of the influenza viruses are composed of four identical subunits, whose general shape is that of a mushroom (molecular weight approximately 240 kDa).¹⁰⁸ The tetramer of identical subunits is located at the top of a long thin stalk that is anchored into the viral coat by the N-terminus of the protein. Disulfide bridges link the subunits of the sialidase, and the mushroom head of the enzyme is located 60 Å from the viral coat.¹¹⁰ Each subunit of the sialidase head is composed of six four-stranded antiparallel β -sheets. Through the centre of each subunit lies a pseudo 6-fold

symmetry axis. The structure of a monomer of the influenza virus sialidase is shown in Figure 2.1.



Figure 2.1: The structure of an influenza virus sialidase monomer.^{49, 50}

Investigations of the viral sialidase show that up to 50% sequence variation in the amino acid sequence may be found between the different strains.¹¹⁰ However the sialidase has been found to have very similar three-dimensional structure and the amino acids that mediate active-site binding are highly conserved across all influenza A and B strains found.¹¹⁰ The active site is highly charged with a majority of residues at physiological pH carrying a charge. Two calcium-binding domains have been located and it is postulated that the calcium may have a role in the release of nascent virions from the cell by positively modulating the sialidase activity.¹¹⁰ The calcium ions are also believed to stabilise the symmetry of the sialidase heads.¹¹¹

The structure of *N*-acetylneuraminic acid when bound into the active site of influenza sialidase has been examined previously.¹¹² When a *N*-acetylneuraminic acid binds to the sialidase active site, the pyranose ring structure of the Neu5Ac residue is distorted from the normal ²C₅ glycoside chair conformation to bind in a pseudoboat conformation.¹¹⁰ As previously discussed in Section 1.4.2.1, the next step in the enzyme mechanism involves the donation of a proton from the solvent facilitating the formation of an endocyclic sialosyl cation transition-state intermediate. The final step involves the formation of *N*-acetylneuraminic acid and its release.¹¹⁰ An alternate reaction mechanism utilising a mixed S_N1-S_N2 mechanism that gives a covalently bound intermediate has also been proposed.⁶⁷

The binding of *N*-acetylneuraminic acid into the active site involves interactions with key residues in the active site. The carboxylate group is drawn into an equatorial conformation by interactions with the triarginyl cluster (Arg 118, 292 and 371). When the NH group of the acetamido side chain binds into the active site it is directed to the floor of the active site.¹¹² The methyl carbon at the terminus of the acetamido group is used for molecular recognition of the substrate and is located close to residues Trp 178 and Ile 222, these residues form a hydrophobic pocket.¹¹⁰ The hydroxyl groups of the glycerol side-chain hydrogen bond to one glutamic acid moiety, Glu 276, and through water interactions to Glu 277 and Glu 227.¹¹⁰ Analysis of the C-4 binding pocket of influenza virus sialidase with the program GRID⁹³ identified a favourable interaction with an amino probe.^{14, 71} When the 4-amino-4-deoxy-Neu5Ac2en derivative was tested, it was found to inhibit influenza virus sialidase at 4×10^{-8} M.⁷¹ Further visual analysis of the C-4 pocket revealed a pocket that was large enough to accommodate a guanidino functionality that would be able to form salt-bridges to two invariant glutamic acid residues: Glu 119, Glu 227. It was these glutamic acid residues that were exploited in the design of the first potent inhibitor of influenza virus sialidase, Zanamivir.^{14, 71}

2.2 Computational design

The design process for the development of inhibitors of NDV HN (and by extension hPIV HN) is presented in Figure 2.2.

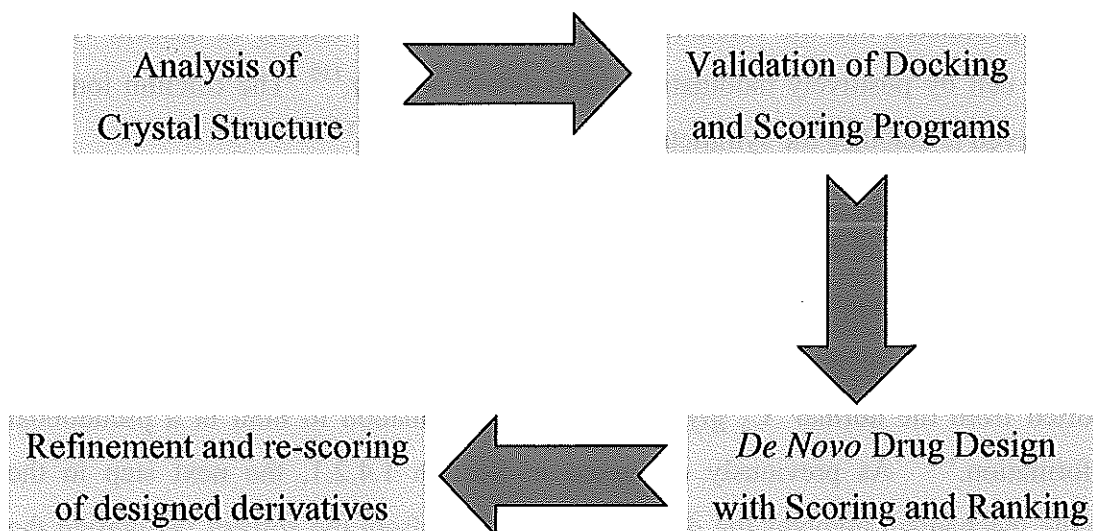


Figure 2.2: Flow diagram of design process.

The computational design process initially involved the analysis of the crystal structure. This involved the visualisation of the three dimensional structure of the enzyme using molecular visualisation software such as InsightII, and the analysis of the suitability of the structure for drug design using the programs Procheck and Whatcheck

(detailed in Section 2.3.1). The characterisation of key interaction sites and the electrostatics of the active site of the enzyme were performed using the programs GRID, PASS, Ligplot and DelPhi (detailed in Section 2.3.2). The next step in the computational process involved the validation of the various docking and scoring programs; the programs selected include DOCK, AutoDock, SCORE, X-Score, Discover (MMI), and AMBER (MMPBSA). The validation involved the benchmarking of the docking programs (AutoDock and DOCK, Section 2.3.3) with a known compound. In this case, 5-acetamido-2,6-anhydro-3,5-dideoxy-D-*glycero*-D-*galacto*-2-non-2-enoic acid, Neu5Ac2en (**5**) was selected. Docking programs (AutoDock, DOCK) were then analysed to examine whether they could reproduce the crystal structure, and the scoring programs (SCORE, X-Score, MMI and MMPBSA, Section 2.3.3 and Section 2.3.7.3.1) were analysed for a base (benchmark) score for comparison with the designed ligand scores. Following the validation of the various programs used, the *de novo* design process was performed using the program LigBuilder (detailed in Section 2.3.7). This involved a fragment-based construction process guided by active site interaction points. The ligands that were designed *via* the *de novo* process were then ranked, scored, and the top molecules were refined. The refined ligands were then docked, scored, and ranked (detailed in Section 2.3.7.2 and 2.3.7.3) which gave the final structures to be chemically synthesised (Chapter 3).

2.3 Computational analysis of NDV HN

2.3.1 Validation of NDV HN crystal structure

Three crystal structures of the globular head of NDV HN (Kansas Strain) have been solved.⁴⁸ A ligand-free structure at pH 4.6 (2.5 Å, PDB code: 1E8T), a complex with sialic acid at pH 4.6 (2.0 Å, PDB code: 1E8U), and a complex with Neu5Ac2en at pH 6.5 (2.8 Å, PDB code: 1E8V). Overall the topology of the globular head is a six-bladed β -propeller, similar to that of influenza virus sialidase. However, an additional α -helical domain not seen in influenza virus sialidase is present. Calcium is found co-ordinated in a pentagonal bipyrimidal arrangement with residues Asp 261 and Ser 264.⁴⁸ The crystal structure is organised into two monomer chains, A and B. Each chain is comprised of residues 124 to 570 (residues 1 - 123 comprise the stalk region which is cleaved before crystallisation⁴⁸), with 129 water residues complexed with chain A and 110 water residues complexed with chain B. Two glycosylation sites are present on each chain. The results presented in Chapter 2 are for the NDV HN crystal structure

complexed with Neu5Ac2en (1E8V) unless otherwise stated. This structure was chosen on the basis that the structure-based drug design would entail looking for an optimal transition state based inhibitor. As the structure of Neu5Ac2en is close to the proposed transition state for a sialidase, the structure 1E8V was chosen for computational drug design.

2.3.1.1 Procheck

The program Procheck⁸⁹ was used to assess the NDV HN X-ray crystal structures that were available for how each structure deviates from ideal values. Table 2.1 shows the Procheck results of all three available NDV HN crystal structures. The Ramachandran plot profile is less than ideal (favoured < 90%) for each of the structures, however the structure with bound inhibitor Neu5Ac2en **5** (1E8V) is marginally the best available structure.

Table 2.1: Procheck results of NDV HN crystal structures 1E8T, 1E8U and 1E8V.

	1E8T	1E8U	1E8V ^a
Ramachandran Plot			
Favoured	82.4%	83.5%	83.9%
Allowed	16.5%	15%	14.5%
Generously Allowed	0.5%	0.8%	0.8 %
Disfavoured	0.6%	0.4%	0.8%
Main-chain bond	100% within limits	100% within limits	100% within limits
Main-chain angles	98.2% within limits	98.1% within limits	98% within limits
Planar atoms	99% within plane	97.3% within plane	100% within plane

^a Complex with Neu5Ac2en

2.3.1.2 Whatcheck

The structure 1E8V was analysed using Whatcheck⁹⁰ to identify potential problems with the structure. Multiple Gln and a single Asn residues were found to be detrimental to the hydrogen-bond network of the NDV HN due to incorrect protonation states or incorrect side-chain orientation.¹¹³ The residues located were: Gln 157, Gln 204, Gln 280, Asn 298, Gln, 371, Gln 496. None of these residues are seen to make interaction with Neu5Ac2en **5**, or the active site.

Other problems found by Whatcheck with the protein structure were a high Matthew's coefficient, unusual bond and torsion angles, too tightly restrained omega angles, abnormal backbone torsion angles, abnormally short inter-atomic distances, and unusual backbone conformations.

2.3.1.3 Hbplus

Hbplus⁹¹ was used to calculate His, Gln and Asn residues that contribute detrimentally to the hydrogen-bond network. The residues found were: Gln 157, Asn 298, and Gln 496.

2.3.1.4 Correlation of results from Whatcheck and Hbplus

Of the Gln and Asn residues identified by Whatcheck, three correlate with the findings of Hbplus: Gln 157, Asn 298, and Gln 496. By reorienting the side-chains of these residues by 180°, theoretically the hydrogen bonding network of the protein should be improved.¹¹³ The torsions changed can be seen in Table 2.2. The three residues identified by Whatcheck but not by Hbplus, namely Gln 204, Gln 280 and Gln 371 were not flipped. These residues were visually inspected for distances and angles. The lack of consensus between Hbplus and Whatcheck is expected as both programs do not test for solvent accessibility. There His, Gln or Asn residues on the surface of the protein would be calculated by Hbplus as able to form H-bonds to solvent whereas Whatcheck only calculates protein mediated H-bonding. A consensus would be expected for any H-bonding within the protein.

Table 2.2: Residues with side-chains flipped 180°

Residue	Torsion	Original Angle	Flipped Angle
Gln 157	χ^3	103.94°	-76.06°
Asn 298	χ^2	101.51°	-78.49°
Gln 496	χ^3	-75.98°	104.02°

With the initial evaluation of crystal structure 1E8V finished, with minor improvements made, the next step was to evaluate interactions within the active site of NDV HN.

2.3.2 Initial investigation of NDV HN

2.3.2.1 The interactions of the active site of NDV HN with Neu5Ac2en.

The hydrogen bonds that are formed between the atoms of the active site and Neu5Ac2en can be seen in Table 2.3, and are indicated in Figure 2.3 (generated by Ligplot¹¹⁴) by dashed green lines. Hydrophobic interactions are indicated in Figure 2.3 by red arcs labelling active site residues and ligand atoms. Neu5Ac2en is seen to make

six hydrogen bonds between the carboxylate group and the triarginyl cluster (Arg 174, Arg 416, and Arg 498), similar to interactions found for Neu5Ac2en in other sialidase structures.^{48, 51-53} The three glycerol-side chain hydroxyl groups each make at least one hydrogen bond to the active site of NDV HN. Almost all carbon atoms of Neu5Ac2en with the exception of C-5 make hydrophobic interactions with numerous active site residues.

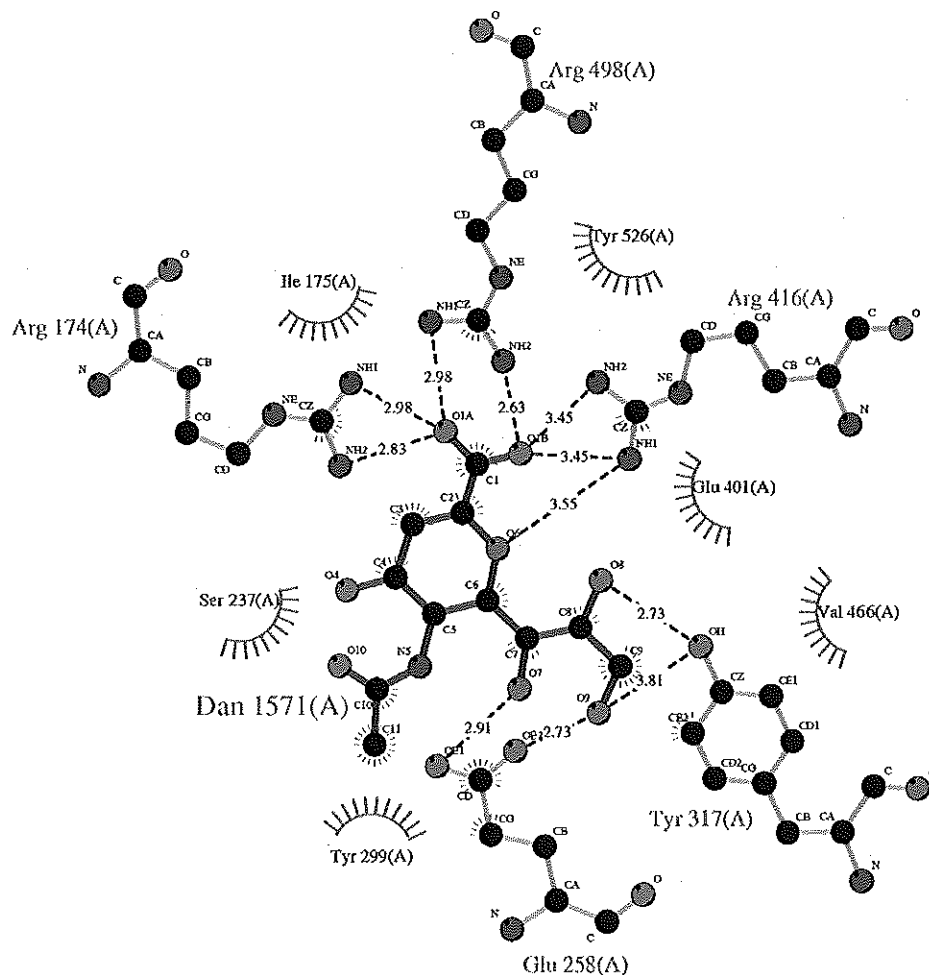


Figure 2.3: Ligplot of interactions of Neu5Ac2en with NDV HN.

Table 2.3: NDV HN and Neu5Ac2en H-bond interactions.

Protein Residue	Atom	Ligand Residue	Atom	Distance (Å)
Arg 174	NH1	DAN	O1A	2.98
Arg 174	NH2	DAN	O1A	2.83
Glu 258	OE1	DAN	O7	2.91
Glu 258	OE2	DAN	O9	2.73
Tyr 317	OH	DAN	O8	2.73
Arg 416	NH1	DAN	O1B	3.45
Arg 416	NH1	DAN	O6	3.55
Arg 416	NH2	DAN	O1B	3.45
Arg 498	NH1	DAN	O1A	2.98
Arg 498	NH2	DAN	O1B	2.63
Tyr 526	OH	DAN	O6	3.23

2.3.2.2 Analysis of putative catalytic residue Asp 198.

The side-chain conformation for the putative catalytic residue Asp 198 is different between the high and low pH structures of NDV HN (1E8T and 1E8V).⁴⁸ The side-chain of Asp 198 in the low pH structure (apo enzyme) is found to be oriented towards the bound Neu5Ac2en **5**, of the superimposed high pH structure in Figure 2.4. The measured distance between the side-chain atom OD2 and C-2 of Neu5Ac2en is 7.88 Å. By comparison the distance between the side-chain atom OD2 of high pH Asp 198 and C-2 of Neu5Ac2en is 16.31 Å. Superimposition of the triarginyl clusters of low and high pH HN and influenza virus N9 sialidase (Figure 2.4) allows for comparison between HN and N9 catalytic residues. The RMSD of these superimposed residues was 2.66 Å. The distance between the catalytic residue Asp 152 of N9 sialidase and bound Neu5Ac2en was measured at 3.81 Å (between OD2 and C-2). This is 4.07 Å less than the distance between Asp 198 in the low pH HN and Neu5Ac2en atoms. There is however the potential for water interaction between the low pH Asp 198 and Neu5Ac2en. These results suggest that the catalytic activity of Asp 198 in HN is mediated *via* a water molecule.

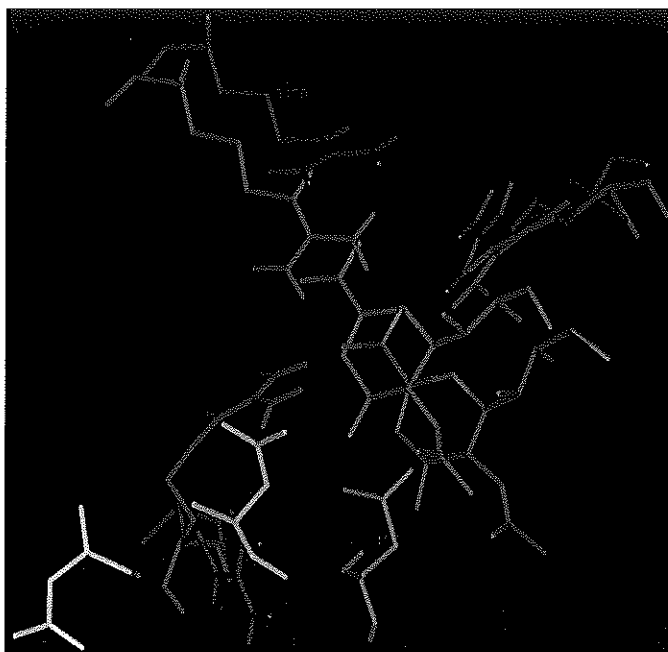


Figure 2.4: Superimposition of NDV HN and influenza virus N9 sialidase triarginyln clusters. The low pH HN structure is depicted in red, the high pH HN structure in yellow, and N9 sialidase in blue. The green Neu5Ac2en is from high pH HN, while the purple Neu5Ac2en is from N9 sialidase.

2.3.2.3 *DelPhi of unminimised NDV HN*

DelPhi was used to analyse the electrostatic potential of the active site of NDV HN, this can be seen in Figure 2.5. The active site of NDV HN is dominated around the carboxylate group of Neu5Ac2en **5** by a positive field potential. This positive potential around the triarginyln cluster is to be expected. The floor of the active site under the ring of **5** is mixed with large areas of neutral and negative potential. The glycerol binding pocket is lined by positive field potential on one side and negative on the other. This is separated by a very small area of neutral field potential. An area of neutral field potential is found around the end of the pocket, this is primarily contributed by Phe 364. The C-4 binding pocket is very large and has a predominantly positive potential, small neutral and negative areas are also found within the pocket. Electrostatic analysis suggested a number of modifications to the Neu5Ac2en, **5**, template; in the glycerol-side chain pocket a hydrophobic group at C-9 could potentially pick up an interaction with Phe 364, and in the C-4 pocket, negatively charged groups could provide interactions with a number of positively charged amino acids. Interactions with the C-4 binding pocket are detailed in the next section.

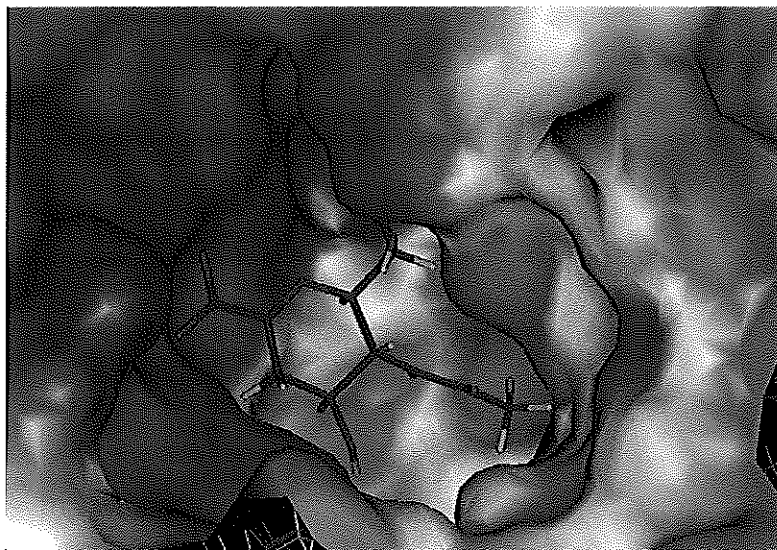


Figure 2.5: Initial DelPhi electrostatics calculation results of NDV HN. Blue: positive field potential, red: negative field potential, and white: neutral field potential.

2.3.2.4 C-4 binding pocket

The following amino acid residues are located around the C-4 binding pocket: Ser 237, Arg 235, Ile 175, Lys 236, Ser 177, His 199, Ile 192, Arg 174, Asp 198, Asn 190, Gln 204, Arg 197, Cys 238, His 189, Cys 251, and Cys 196. The following crystal water residues are located in the C-4 binding pocket: Z128H, Z55H. However, the temperature factors of Z128H suggest that it is not entirely stable (61.37). The crystal water residues can be seen in Figure 2.6. Also shown is the positive nature of the pocket, contributed to by at least six positively charged residues on the solvent accessible surface. The colouring of the surface is in the DelPhi electrostatic spectrum. The charges used for this calculation were from the CVFF force field with DelPhi radii applied, a solvent dielectric of 80 and solute dielectric of 4 were applied with no ionic parameters utilised.

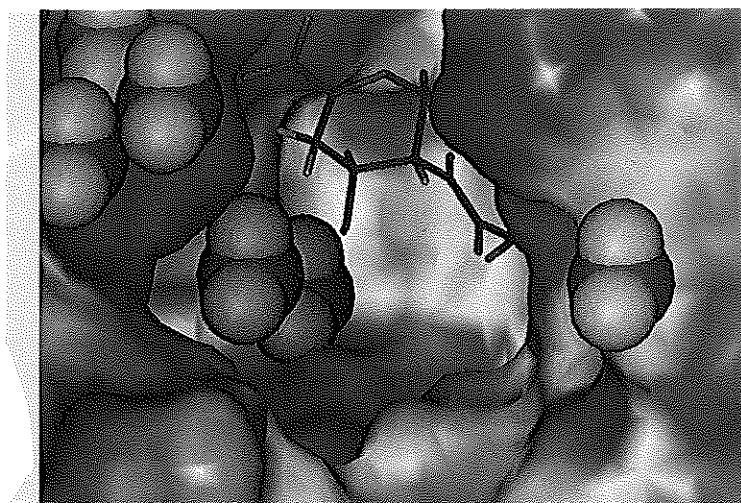


Figure 2.6: C-4 binding pocket of NDV HN with crystal waters CPK rendered.

2.3.2.5 PASS

The volume of the C-4 binding pocket of NDV HN was determined by the program PASS to be 612 Å². The C-4 pocket was the largest found on the protein. The glycerol side-chain binding pocket was also found to be large (though only approx half of C-4 binding pocket volume). Pockets identified by PASS in NDV HN can be seen in Figure 2.7, with the large red sphere indicating the C-4 binding pocket. The glycerol side-chain pocket is indicated by the large white sphere found above the red sphere. PASS highlights that the available space within the C-4 pocket is very large and thus should accept large C-4 modified Neu5Ac2en derivatives carrying bulky substituents at C-4.

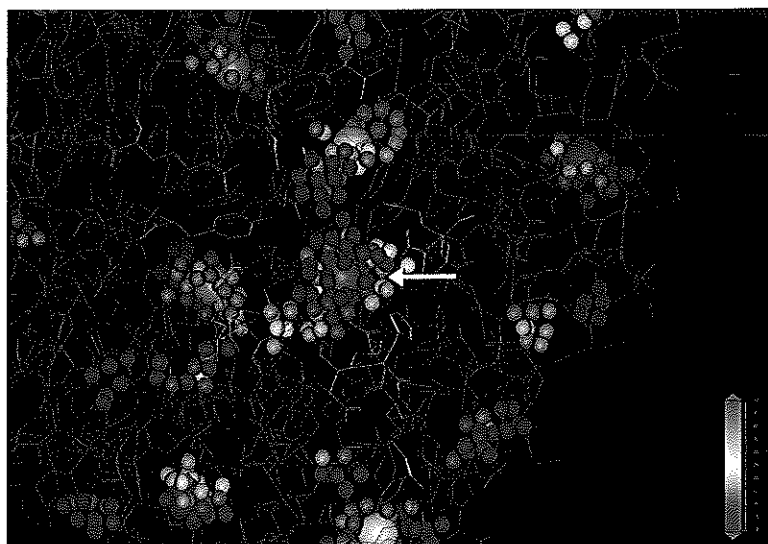


Figure 2.7: Invaginations in the HN protein surface located by PASS. The arrow indicates the C-4 binding pocket sphere.

2.3.2.6 GRID study of NDV HN crystal structure.

Initially the PDB file of the high pH structure of NDV HN (1E8V) was prepared for GRID calculations (see Section 2.1) by the program Greater (available with version 22 of GRID). The program Greater is a GUI to the other programs of the Molecular Discovery suite of programs.¹¹⁵

GRIN is utilised to create the input necessary for GRID. Two input files are required for the GRIN calculation. One is the target protein PDB file containing the XYZ co-ordinates of each atom of the protein. The other input for GRIN is a set of energy variables supplied by the program. The energy variables are based on the “extended” atom concept from the program CHARMM.⁹³ The “extended” atom concept consists of a single heavy atom with modified energy parameters to explain the corresponding heavy atom with ancillary hydrogens attached. Thus a methyl group is represented as a single carbon atom, modified to incorporate the energy parameters of the attached hydrogen atoms.⁹³ The primary purpose of the program GRIN is to merge the XYZ co-ordinates and the corresponding energy parameters to specify the correct interactions with the probes of interest for each atom of the target protein of interest. Hydrogen-bonding hydrogen atoms are added to the protein by GRIN if not in standard geometry.⁹³

The input files required for GRID are the output files from the program GRIN and the GRID input file. The setup of the GRID job requires information on the size of the grid required for the calculation, whether it surrounds the entire protein or whether it surrounds a particular region of the protein. Also at least one probe needs to be specified. The specification of the probe is necessary to utilise the correct energy parameters, as each probe has a different set of parameters.¹¹⁵ Four probe types are available with version 22 of GRID; traditional probes containing one heavy atom, multi atom containing more than one heavy atom, group probes such as glucose, and molecule probes such as methotrexate and special probes. During the calculation, water molecules, co-factors and chloride ions are generally not treated as part of the protein molecule. The main consideration in calculating the energy of interaction between the protein and the probes is the Lenard-Jones potential or the attractive/repulsive non-bonded interactions, with electrostatic force and hydrogen-bond contributions also calculated. The protein is considered to be a rigid molecule with tautomeric hydrogens by default. Different conformations of the target protein can be studied. The protein

can also be studied flexibly with the structure divided into a “core” region with flexible side chains. Upon interaction with the probe the side-chains alter their conformation.¹¹⁵

One significant area of interaction was found with the carboxylate probe as seen in Figure 2.8. The area corresponds to the carboxylate of Neu5Ac2en, with interaction occurring with the triarginyl cluster. Maximum interaction with the probe occurred at -25.13 kcal/mol.

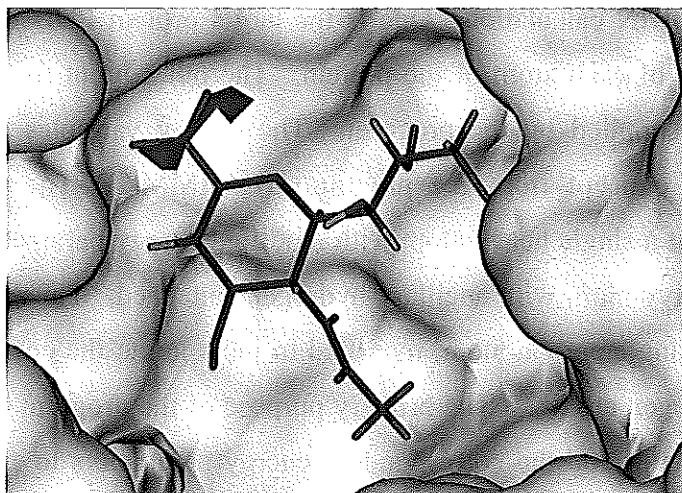


Figure 2.8: Area of interaction with the carboxylate (COO^-) probe, contoured at -17 kcal/mol.

Many areas of interaction were found with the water probe, as seen in Figure 2.9; areas of interaction were located along the glycerol side chain, and around the carboxylate group of Neu5Ac2en. Most significantly, the water probe predicted areas of interaction within the C-4 pocket of the active site corresponding to the areas occupied by crystal water residues. Maximum interaction with the water probe occurred at -15.22 kcal/mol.

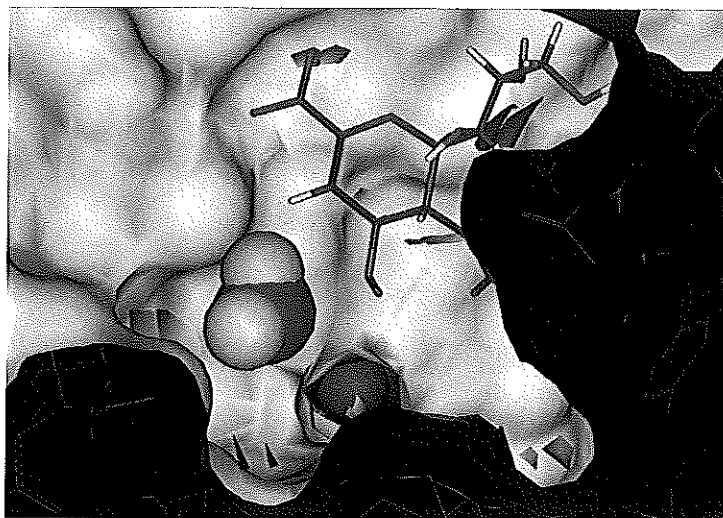


Figure 2.9: Areas of interaction with the water (OH₂) probe, contoured at -6 kcal/mol.

Many areas of interaction with the hydroxyl probe were located as seen in Figure 2.10. Interactions were found along the glycerol side-chain, in the region of the C-4 hydroxyl group, and near the carboxylate group. Maximum interaction with the hydroxyl probe occurred at -13.02 kcal/mol.

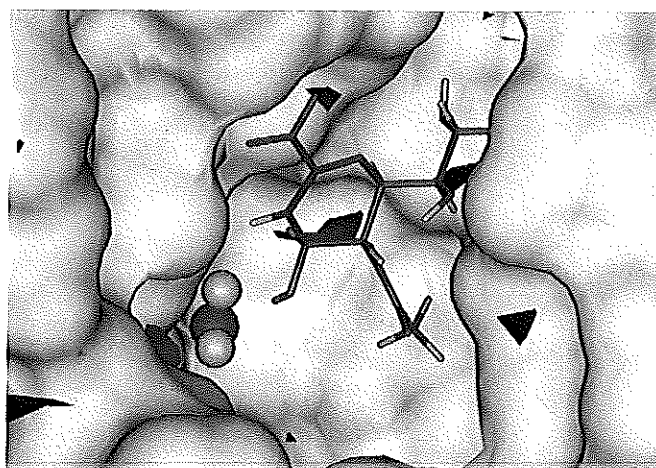


Figure 2.10: Areas of interaction with the hydroxyl (OH) probe, contoured at -8 kcal/mol.

Within the NDV active site three main areas of interaction were found with the methyl probe, as seen in Figure 2.11. Interestingly, the methyl probe was found to interact with a large region of the glycerol side chain pocket. Methyl interactions were also found within the C-5 binding pocket at the terminus of the NHAc group. Maximum interaction with the methyl probe occurred at -5.41 kcal/mol.

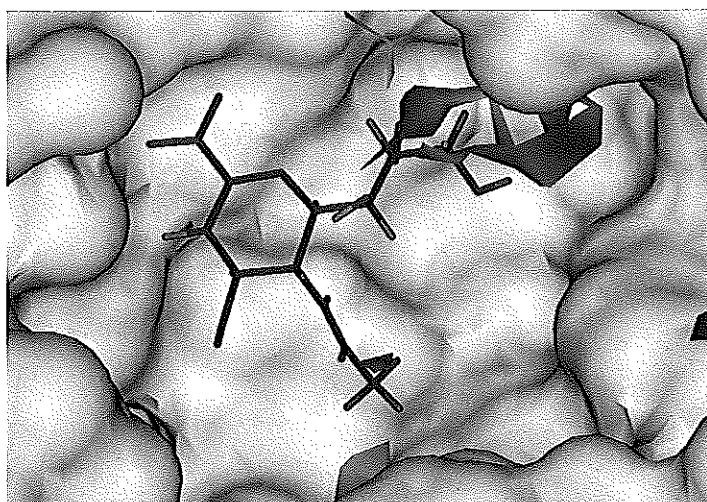


Figure 2.11: Areas of interaction with the methyl (C3) probe, contoured at -3 kcal/mol

Four areas of interaction with the hydrophobic probe were located within the NDV active site, as seen in Figure 2.12. Interaction regions were found at the terminus of the glycerol and NHAc side chains, and two areas of interaction were found within the C-4 binding pocket. Maximum interaction with the hydrophobic probe was -1.41 kcal/mol.

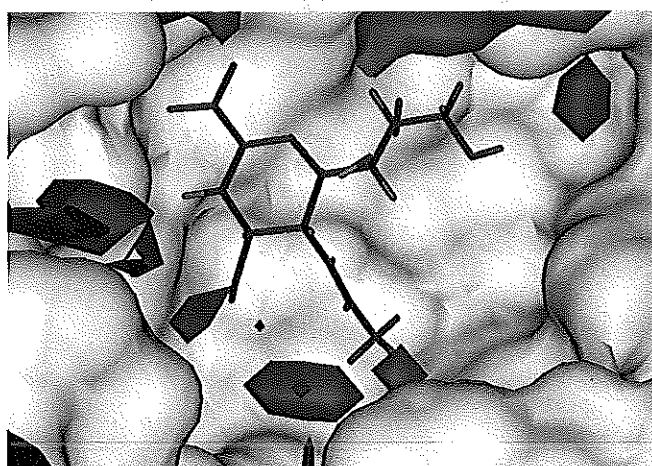


Figure 2.12: Areas of interaction with the hydrophobic (DRY) probe, contoured at -0.2 kcal/mol.

Two areas of interaction with the guanidino probe were located within the active site, as seen in Figure 2.13. The major area of interaction was within the glycerol side-chain binding pocket. An area is also located at the base of the C-4 binding pocket. Maximum interaction with the guanidino probe occurred at -32.84 kcal/mol. The aromatic version of the probe also scored a maximum interaction energy of -32.84 kcal/mol.

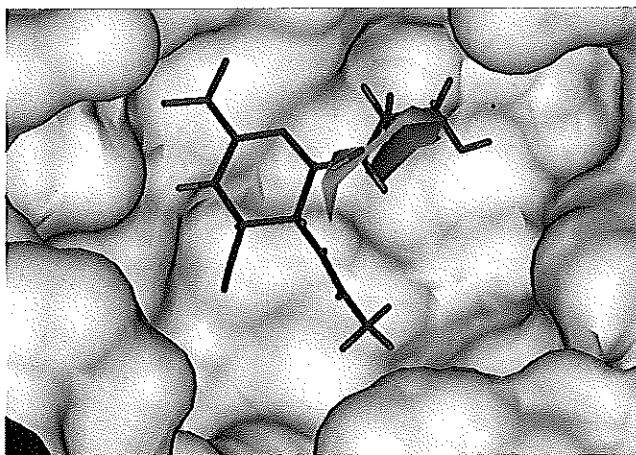


Figure 2.13: Areas of interaction with the guanidino (AMIDINE) probe, contoured at -12 kcal/mol.

Interactions with the amine probe were located within the glycerol-side chain pocket, and at the base of the C-4 and C-5 bonding pockets, as seen in Figure 2.14. The protonated form of the amine probe identified similar interactions. The maximum interaction with the amine probe occurred at -25.73 kcal/mol. The protonated amine probe identified maximum interactions at -25.57 kcal/mol.

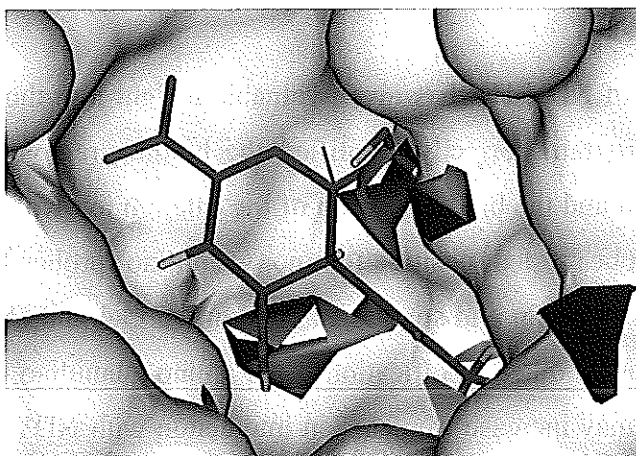


Figure 2.14: Areas of interaction with the amine (N1=) probe, contoured at -6 kcal/mol.

One significant area of interaction with the calcium ion probe was found, as seen in Figure 2.15. The calcium ion probe accurately located the area of maximum interaction in the position of the crystallographically observed calcium ion. Maximum interaction with the calcium ion probe occurred at -139.11 kcal/mol.

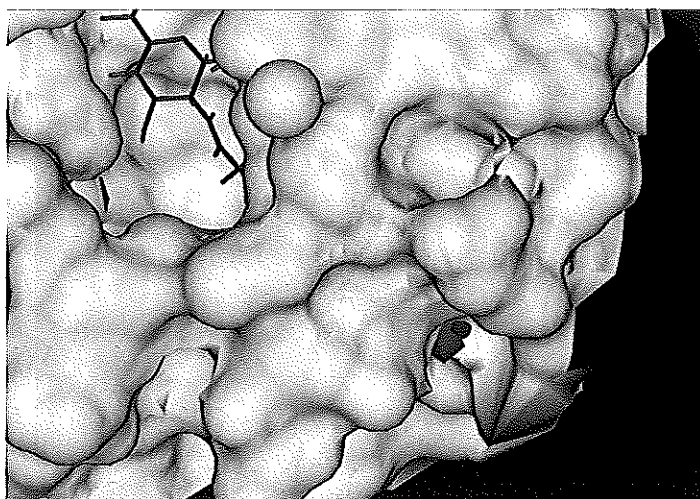


Figure 2.15: Area of interaction with the calcium ion (CA^{+2}) probe, contoured at -130 kcal/mol.

No competition between the structural water molecules and the probes tested (N^{3+} and DRY) was detected for NDV HN. Therefore the structural water residues observed in the C-4 pocket in NDV HN should be able to be displaced by ligand interactions with the active site. This observation coincides with the mid to high temperature factors of the crystal waters in the active site C-4 pocket, suggesting that the crystal waters are not entirely stable.

GRID analysis of NDV HN has given valuable insight into the interactions of the active site. The program correctly identified the binding region of the carboxylate group of Neu5Ac2en and the binding region of the calcium ion with the protein. In the analysis of other probes tested, positive (amino, guanidino) probes located interactions along the glycerol side-chain and small areas within the C-4 and C-5 binding pockets, suggesting that amongst other modifications, that amino/guanidino modification of the glycerol side chain may increase Neu5Ac2en interaction with NDV HN. Hydrophobic probes (C3 and DRY) found extensive interactions within the terminal end of the glycerol side-chain pocket due to interaction with Phe 364. The DRY probe in particular found extensive interactions within the C-4 and C-5 pockets suggesting that hydrophobic groups off C-4 may increase the binding affinity over the normal 4-hydroxy-Neu5Ac2en.

2.3.2.7 Pocket characterisation of NDV HN active site interactions

Pocket, part of LigBuilder, is a program used to characterise the interactions of a binding/active site. As shown in Figure 2.16, pocket characterisation of the active site identified a large area for hydrogen-bond formation (blue spheres) around the triarginyln

cluster, tyrosine residue 526, and into the C-4 binding pocket. A large area of H-bond donors (red spheres) is found adjacent to a H-bond donor region along the C-5 acetamido group and the C-7 hydroxyl group. Hydrophobic interactions (green spheres) are found at the terminus of the N-acetyl group binding pocket, at the end of the glycerol side-chain binding pocket (contributed by Phe 364), and within the C-4 binding pocket.

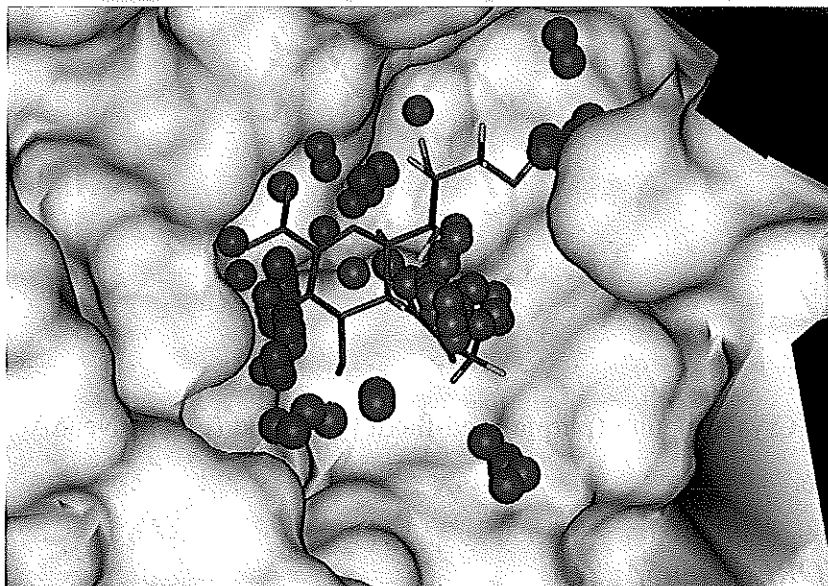


Figure 2.16: Pocket analysis of NDV HN active site. Blue – hydrogen bond acceptor site, red – hydrogen bond donor site, green – hydrophobic interaction.

Pocket analysis of the active site, while more generalised than GRID, has found many similar interactions to GRID around the carboxylate group, glycerol-side chain and most importantly, the hydrophobic interactions of the C-4 pocket. As Pocket analysis of the active site is used for the *de novo* design program LigBuilder, the final designed structures modified at C-4 will be hydrophobic in nature. However, before design of ligands, the next step in the process is to validate and benchmark the docking and scoring programs.

2.3.3 Validation of docking and scoring programs

2.3.3.1 AutoDock - Neu5Ac2en

Neu5Ac2en **5** was removed from crystal structure 1E8V and flexibly docked into the structure using AutoDock 3.05, as seen in Figure 2.17. The docked pose (green) of Neu5Ac2en is very similar to that of crystallographically bound (purple) Neu5Ac2en. A slight shift in the positioning of the Neu5Ac2en is seen. The docked energy of the

flexibly bound Neu5Ac2en was -10.06 kcal/mol. Static scoring of the crystal bound Neu5Ac2en for comparison (without flexible torsions) yielded a docked energy of -9.01 kcal/mol. Therefore the slightly different orientation of Neu5Ac2en in the docked form resulted in an increase of interaction of 1.05 kcal/mol.

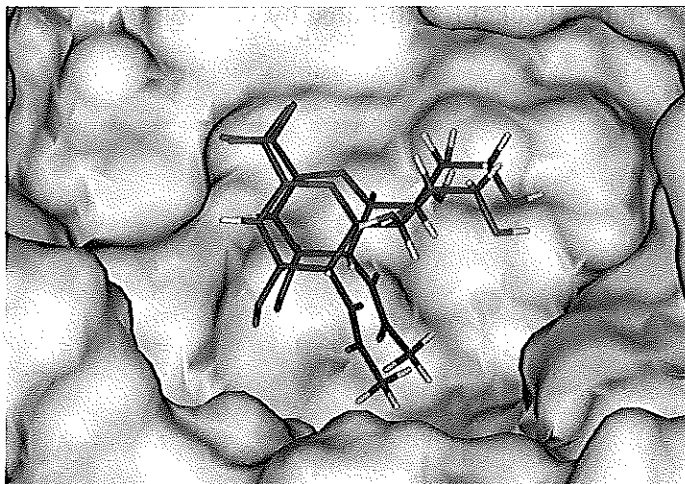


Figure 2.17: Crystal bound (purple) and flexibly docked (green) Neu5Ac2en in NDV HN active site.

2.3.3.2 *SCORE - Neu5Ac2en*

The SCORE algorithm uses a logarithmic scoring function to estimate the binding affinity of a small molecule to a protein. The binding affinity score of Neu5Ac2en (crystal-bound) was found to be 5.52. SCORE classifies atoms of the small molecule according to their interaction with the protein. The good (green), neutral (blue), and bad (red) interactions of Neu5Ac2en with the NDV HN active site can be seen in Figure 2.18.

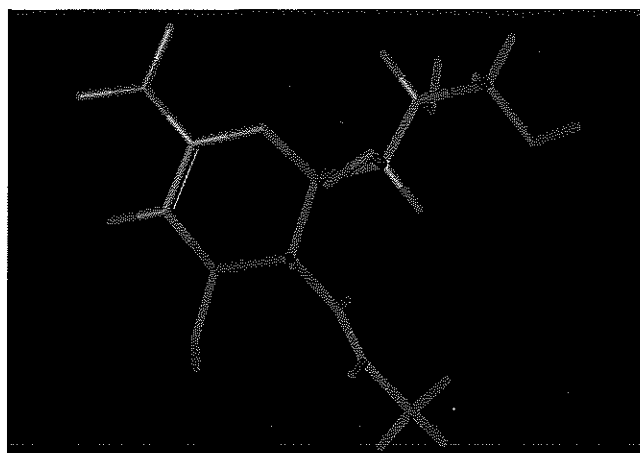


Figure 2.18: SCORE classification of atom interactions (good - green, neutral - blue, and bad - red) of crystal bound Neu5Ac2en in NDV HN active site.

According to SCORE, the C-2, C-3, C-7, C-8 and C-10 carbon atoms of crystal bound Neu5Ac2en cause bad interactions with the active site of NDV HN. The atoms are labelled bad (red) due to the presence of undesirable van der Waals contacts. The atoms O-1A, O-1B, O-6, N, O-7, O-8, and C-11 of Neu5Ac2en are the source of the good interactions with the protein. The good (green) interactions are caused by weak, medium or strong hydrogen bonds or hydrophobic interactions. All other atoms of the molecule are neutral (blue). The SCORE of docked Neu5Ac2en (docked using AutoDock) was found to increase to 6.06. This was due to the removal of some of the undesirable van der Waals contacts.

2.3.3.3 DOCK

DOCK 4⁹⁶ is a program that is used to dock a ligand to a protein and to search the available conformational space of the binding/active site. DOCK was unable to reproduce the crystal structure orientation of Neu5Ac2en with active site of NDV HN (1E8V). The methods tried include: no flexible ligands, no orientation of ligands, no minimisation of anchor or ligand, no binding spheres within the C-4 binding pocket, restriction of spheres (binding interactions) to the glycerol side chain and ring of Neu5Ac2en, manual matching, anchor searching, and random searching. The problem lay with the fact that as soon as any type of flexibility or orientation was introduced into the calculation, the molecule was flipped to position the glycerol side-chain of Neu5Ac2en in the C-4 pocket, while the carboxylate and the C-5 acetamido group interactions remained the same. This flipping of the orientation may have occurred because DOCK did not evaluate the crystal structure conformation of Neu5Ac2en as the lowest energy-state possible. The removal of the binding site interaction points (calculated by DOCK) from the C-4 pocket did not alleviate the problem. Due to this problem, the use of DOCK for *in silico* screening purposes would have to be treated carefully as the final result would be open to interpretation. Therefore, the proposed method chosen for ligand design in this work was *via de novo* methods.

The programs AutoDock and Score having been validated and benchmarked, the next step in the process was to develop the modelling protocol and score some of the known inhibitors of hPIV-2⁶⁹ and influenza virus sialidase.¹¹⁶

2.3.4 Scoring of existing C-4 modified analogues of Neu5Ac2en.

A number of inhibitors of hPIV-2⁶⁹ and influenza virus sialidase¹¹⁶ are known. It was of interest to examine whether these inhibitors would score well against the crystal structure of NDV HN. The results obtained could give some valuable insight into the potential interactions of C-4 modified Neu5Ac2en derivatives within the active site of NDV HN.

The modelling protocol developed for the scoring of each of the ligands was as follows.

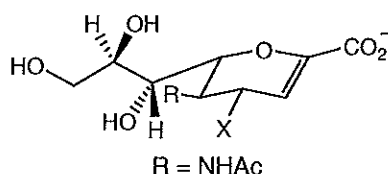
1. Each ligand was flexibly docked using AutoDock to place the ligand in an acceptable conformation within the protein structure for minimisation (if performed).
2. Each ligand (if to be minimised) was then minimised using Discover. The minimisation of the protein was performed using a dielectric constant of 80 (to simulate bulk solvent). The protein was initially fixed and gradually relaxed until the total complex was free to move. The input file for Discover is available in Appendix A.
3. Following the minimisation process or initial Autodock calculation each ligand was statically scored with SCORE and AutoDock.

A number of C-4 modified Neu5Ac2en derivatives prepared by Holzer *et al.*⁶⁹ to inhibit hPIV-2 were scored for interaction with NDV HN. These results are shown in Table 2.4. These derivatives were selected on the basis that this work is utilising the crystal structure of NDV HN as a model for hPIV. Therefore it was of interest to examine whether the biological results of these derivatives would correlate with the docked energy from NDV HN.

Each of the seven derivatives and Neu5Ac2en autodocked in a very similar orientation to the crystallographically bound Neu5Ac2en, however slight changes in the orientation of the glycerol side chain were observed. The Neu5Ac2en derivative that docked with the highest energy of interaction after minimisation was 4-(prop-2-enylamino) derivative (Ligand F, Table 2.4), with -14.70 kcal/mol in the minimised active site. The SCORE result of this ligand was also the second highest found at 6.24. In most cases an increase in the docked energy of the ligands was seen in the minimised

active site compared with the unminimised site. Similarly an increase in the binding affinity (SCORE result) of most of the ligands in the minimised site was also observed. The reported K_i for each ligand is the inhibition constant against hPIV-2 virus. Each of the ligands had been evaluated as inhibitors of hPIV-2; the inhibition constants (K_i) are given in Table 2.4. Only a weak correlation was found between the docked energy and experimental K_i . However more data points are required for conclusive proof that a correlation exists.

Table 2.4: AutoDock and SCORE results for C-4 modified Neu5Ac2en derivatives in NDV HN of ligands known to inhibit hPIV-2.⁶⁹



C-4-X Substituent	Flexible AutoDock (kcal/mol) (Unminimised)	SCORE (pK_d) (Unminimised HN)	Flexible AutoDock (kcal/mol) (Minimised HN)	SCORE (pK_d) (Minimised HN)	Experimental inhibition constant ^{69a}
A OH	-10.40	3.97	-12.16	4.97	$1 \times 10^{-5}M$
B NHC(=NH)NH ₂	-13.35 [2] ^b	3.41 [7]	-14.16 [2]	6.45 [2]	$8 \times 10^{-4}M$
C NH ₂	-12.13 [6]	4.19 [4]	-13.20 [6]	5.57 [5]	$3 \times 10^{-4}M$
D N(CH ₃) ₂	-12.46 [5]	4.47 [1]	-13.52 [4]	4.76 [6]	$1 \times 10^{-4}M$
E NHC(=O)CH ₃	-13.35 [2]	4.13 [5]	-13.28 [5]	6.88 [1]	$4 \times 10^{-5}M$
F NHCH ₂ CH=CH ₂	-13.95 [1]	4.33 [2]	-14.70 [1]	6.24 [3]	$2 \times 10^{-5}M$
G NH(OH)CH ₂ CH=CH ₂	-12.72 [3]	4.26 [3]	-12.16 [7]	4.14 [7]	$3 \times 10^{-5}M$
H NHCH ₂ CH ₂ OH	-12.70 [4]	3.70 [6]	-13.78 [3]	6.07 [4]	$1 \times 10^{-4}M$

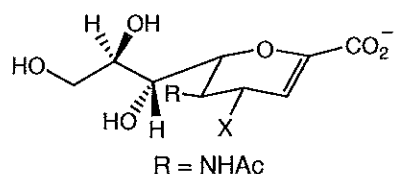
^a Inhibition data against hPIV-2 HN. ^b Numbers in square brackets indicate scoring rank.

The next set of ligands chosen for evaluation was a series of C-4 substituted Neu5Ac2en derivatives reported to inhibit influenza virus sialidase.¹¹⁶ Table 2.5 shows the scoring results of selected ligands from Taylor *et al.*¹¹⁶ against NDV HN (unminimised active site). The removal of some of the ligands detailed by Taylor *et al.*¹¹⁶ from these calculations was due to the fact that structures were included in the previous calculations involving C-4 modified analogues from Holzer *et al.*⁶⁹ (Table 2.4). The ligand found to dock with the greatest interaction energy was ligand P, Table 2.5, with -13.74 kcal/mol. An improvement of 2.57 kcal/mol over NeuAc2en is seen with

the addition of the C-4 substituent. The SCORE result of ligand P is 5.08, an increase of 0.58 over Neu5Ac2en. Using the SCORE algorithm, ligand S has a value 1.41 greater than the flexibly docked orientation of Neu5Ac2en (5.91 vs 4.50). However, the docked energy is 1.80 kcal/mol greater than Neu5Ac2en. The difference in docked energy could be explained by ligand S making less hydrogen bond contacts with the protein than ligand P. The examination of C-4 substituted Neu5Ac2en derivatives reported to inhibit influenza virus sialidase has not identified any compounds that differ significantly for Neu5Ac2en.

As can be seen in Tables 2.4 and 2.5 the AutoDock scores of the compounds are very similar and differ by a small amount (~1 to 2 kcal/mol). This difference is not expected to be significant and as such each of these compounds would be computationally expected to have similar affinity to the protein. It should also be noted that there is a small difference in the benchmark of Neu5Ac2en shown in Table 2.5 compared to that found in Table 2.4. This is due to the AutoDock protocol utilising 250,000 energy evaluations per docked pose generated. This setting may not allow for the complete sampling of conformational space available. However, this setting allows for the quick analysis of many ligands and it will allow for the analysis of many compounds in a timely fashion. After the identification of the potential compounds a more exhaustive search of the conformational space could be conducted on the smaller number of compounds.

Table 2.5: Flexible AutoDock and SCORE values for C-4 modified Neu5Ac2en derivatives of ligands known to inhibit influenza virus sialidase in NDV HN.¹¹⁶



Ligand No. ^a	C-4-X Substituent ^b	Flexible AutoDock (kcal/mol) (Unminimised)	SCORE (pK _d) (Unminimised)
A	OH	-11.17	4.50
B	NC(=NH ₂)NH ₂	-12.18 [6] ^c	4.14
C	NHCH ₃	-11.84 [13]	5.55
D	N(OH)CH ₂ CH=CH ₂	-12.01 [11]	4.87
E	NH ₂ (R)	-11.95 [12]	5.87
F	NHC(=O)NH ₂	-12.10 [8]	4.07
G	NHC(=O)CH ₂ NH ₂	-12.39 [5]	4.91
H	N(OH)CH ₂ CH ₂ OH	-12.09 [9]	5.03
I	NHC(=O)CH=CH ₂	-13.07 [2]	5.20
J	NC(=NH ₂)NH ₂ (R)	-12.40 [4]	5.09
K	=O	-11.51 [14]	5.39
L	=NOCH ₃	-11.51 [14]	4.58
M	=NNH ₂	-12.16 [7]	5.20
N	N(CH ₂ CH ₂ OH) ₂	-11.34 [16]	4.48
O	NHNNH ₂	-12.04 [10]	5.01
P	N(CH ₂ CH=CH ₂) ₂	-13.74 [1]	5.08
Q	N(CH ₃) ₃ ⁺	-10.99 [17]	5.38
R	N(CH ₂ CH=CH ₂)CH ₂ CH ₂ OH	-11.41 [15]	5.05
S	=NCH ₂ CH=CH ₂	-12.97 [3]	5.91

^a Ligand No. is derived from Taylor *et al.*¹¹⁶ ^b Stereochemistry at C-4 is *S* (equatorial substituent) unless otherwise indicated. ^c Numbers in square brackets indicate scoring rank.

2.3.4.1 Refinement of the initial modelling protocol

The refined modelling protocol was used to evaluate the binding to NDV HN of a series of C-4 modified Neu5Ac2en derivatives reported by Suzuki *et al.*⁸³ to inhibit hPIV-1 sialidase activity. The most potent inhibitor reported was the 4-*O*-thiocarbamoylmethyl derivative **21**. The purpose of refining the modelling protocol was to introduce the protocol of consensus scoring mentioned in Section 2.1. The refined initial modelling protocol developed for the scoring of each of the ligands was:

1. Each ligand was flexibly docked using AutoDock to place the ligand in an acceptable conformation within the protein structure for minimisation.
2. Each ligand was then minimised using Discover. The minimisation of the protein was performed using a dielectric constant of 80 (to simulate bulk solvent). The protein was initially fixed and gradually relaxed until the total complex was free to move. The input file for Discover is available in Appendix A.
3. Following the minimisation process each ligand was statically scored with MMI, SCORE and AutoDock.

The derivatives of Neu5Ac2en reported by Suzuki *et al.*⁸³ were docked and scored against NDV HN using the refined protocol, with the results given in Table 2.6. The interaction energies of each Neu5Ac2en derivative were determined by flexible AutoDock before minimisation. The top scoring pose was then minimised using Discover and the minimised complex was evaluated using AutoDock (statically), Molecular Mechanics Interaction (MMI) energy score, where the interaction energies of the ligands were determined with respect to the protein active site residues, and SCORE.

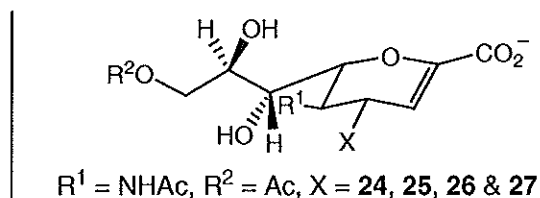
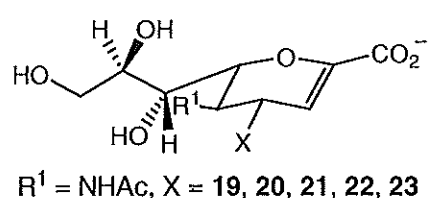


Table 2.6: Docking and scoring results in NDV HN for the C-4 modified Neu5Ac2en derivative reported by Suzuki *et al.*⁸³ as inhibitors of hPIV-1.

C-4-X-substituent ^a	AutoDock (Unminimised) kcal/mol	MMI (Minimised) kcal/mol	AutoDock (Minimised) kcal/mol	SCORE (Minimised) pK_d
OH	-10.06	-173.52	-12.42	6.06
	-14.66 [6] ^b	-185.93 [5]	-12.00 [3]	6.22 [4]
19				
	-14.19 [7]	-187.53 [4]	-5.31 [8]	6.50 [2]
20				
	-14.76 [5]	-191.18 [3]	-7.18 [5]	6.02 [7]
21				
	-13.72 [8]	-181.82 [6]	-3.44 [9]	6.09 [6]
22				
	-14.94 [4]	-202.77 [1]	-5.34 [7]	7.59 [1]
23				
	-15.20 [3]	-178.29 [7]	-8.67 [4]	5.67 [8]
24				
	-15.63 [1]	-195.41 [2]	-5.84 [6]	6.26 [3]
25				
	-15.47 [2]	-174.85 [8]	-13.00 [2]	5.58 [9]
26				
	-15.47 [2]	-174.34 [9]	-13.13 [1]	6.21 [5]
27				

^a denotes point of oxygen attachment. ^b Numbers in square brackets indicate scoring rank. ^c Derivative 23 is not from the literature.

The preliminary AutoDock (flexible AutoDock) score for each derivative of Neu5Ac2en in the unminimised active site ranged from -13.72 kcal/mol, with

derivative **22** (3.66 kcal/mol higher than Neu5Ac2en) to -15.63 kcal/mol, with derivative **25** (5.57 kcal/mol higher than Neu5Ac2en). After the minimisation procedure, AutoDock was also used as a scoring function (static AutoDock). No increase in the AutoDock score after minimisation was observed for any of the nine derivatives. The loss in the energy of interaction observed in the AutoDock score after the minimisation process ranged from a maximum gain of 0.71 kcal/mol with derivative **27**, to a maximum loss of 8.98 kcal/mol with derivative **22**. It is postulated that the lower energy of derivatives **20** to **25** could be due to the loss of favourable H-bond interactions. Figure 2.19 shows the docked orientation of derivative **27**, which showed minimal loss of interaction after minimisation.

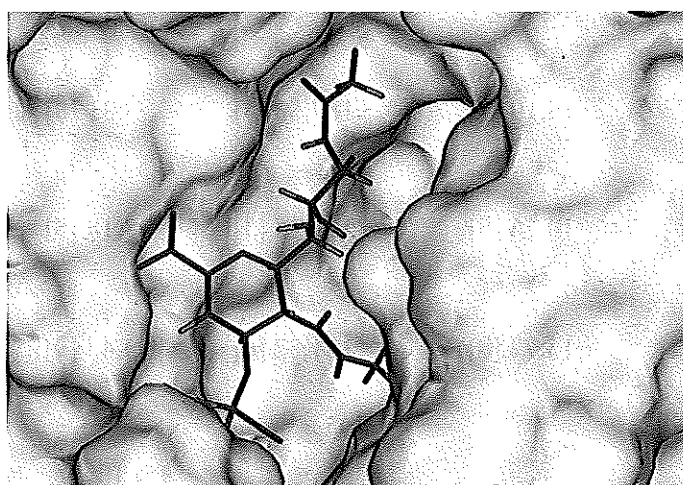


Figure 2.19: AutoDock orientation of derivative **27**.

The MMI energies of each derivative with the minimised NDV HN active site ranged from -174.34 to -202.77 kcal/mol (Table 2.6). Derivative **23** scored the highest (29.25 kcal/mol higher than Neu5Ac2en, -173.52kcal/mol) and derivative **27** scored the lowest (1.33 kcal/mol higher than Neu5Ac2en).

Using the SCORE algorithm, ligand was also ranked as the best binding ligand (Table 2.6). Figure 2.20 shows a diagrammatic representation of SCORE results for derivative **23**. The SCORE results for the C-4 derivative ranged from 5.58 with **26** to 7.59 with derivative **23**. The highest scoring derivative, **23**, showed an approximate increase of 100 times in binding affinity with respect to the lowest scoring derivative, **26**. The results range from approximately 3 times lower to approximately 33 times greater than the calculated binding affinity of Neu5Ac2en.

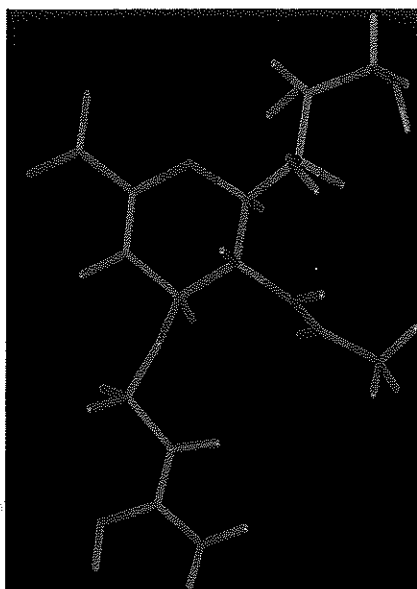


Figure 2.20: SCORE result of derivative **23** (Table 2.6). Colour coding: good interaction (green), neutral interaction (blue) and bad interaction (red).

The program Ligplot was used to more closely analyse potential interactions of ligands **23** within the active site of NDV HN. Hydrogen bonds formed between derivative **23**, and minimised the NDV active site are shown in Figure 2.21. All three arginine residues from the triarginyl cluster form hydrogen bonds with the carboxylate group of **23**. Three hydrogen bonds are formed between the C-7 hydroxyl group of the glycerol side-chain and a glutamic acid and an arginine residue. The glutamic acid residue also forms a hydrogen bond to the nitrogen of the C-5 acetamido group. One hydrophobic interaction (with Ile 192) is present within the C-4 binding pocket. Two hydrogen bonds are also formed, one from the C-4 oxygen (with a histidine residue) and the other between a nitrogen of the guanidino group of **23** and an aspartic acid residue.

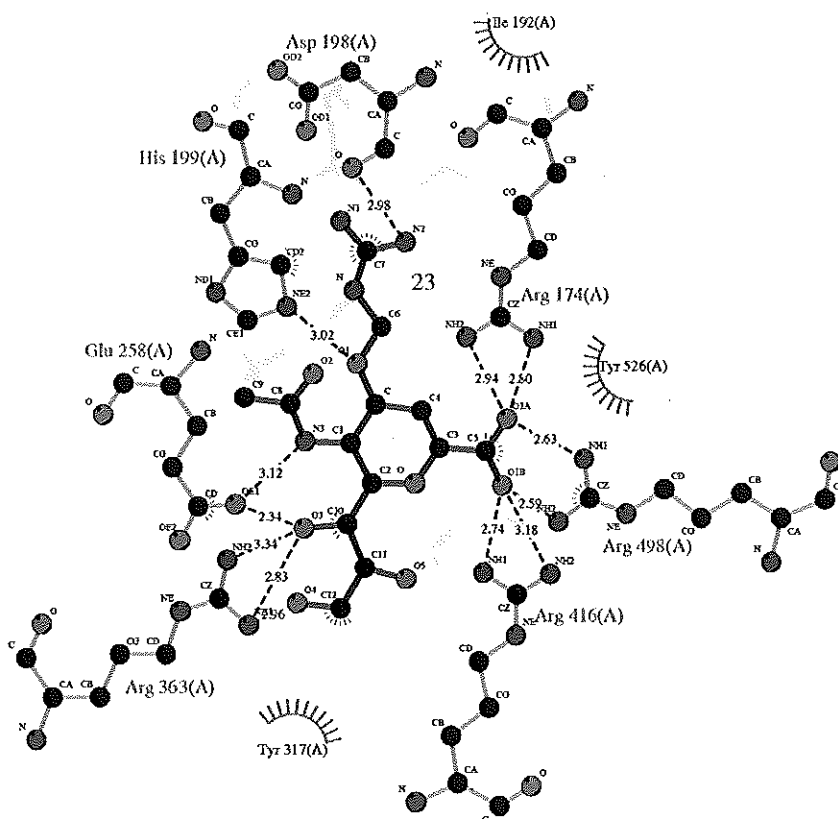


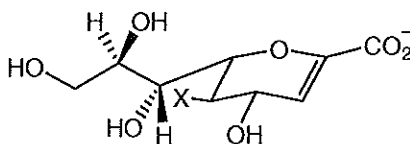
Figure 2.21: Ligplot analysis of Neu5Ac2en derivative **23** in NDV HN.

2.3.5 Scoring of existing C-5 modified analogues of Neu5Ac2en

The next set of ligands chosen for evaluation was a series of C-5 substituted Neu5Ac2en derivatives reported to inhibit NDV HN,⁶⁸ the results are shown in Table 2.7. The availability of NDV HN inhibition data would give an opportunity for a more direct validation of the computational modelling of NDV HN inhibitors. The highest scoring derivative was found to be the *N*-aminoacetyl derivative (Ligand I, Table 2.7) with -11.80 kcal/mol (1.4 kcal/mol greater than that of Neu5Ac2en with -10.40 kcal/mol). The lowest scoring derivative was the *N*-formyl derivative (Ligand B, Table 2.7) with -9.61 kcal/mol (0.79 kcal/mol lower than Neu5Ac2en). The difference between the lowest and highest scoring derivatives was 2.19 kcal/mol. The computational scoring of docked conformations of C-5 modified Neu5Ac2en derivatives against NDV HN did not show an appreciable difference in energy between the derivatives and the parent ligand Neu5Ac2en **5**. However, it is noted that biologically when tested against Newcastle Disease Virus sialidase activity, a marked difference is shown in the inhibitory data. The *N*-trifluoroacetyl derivative (Ligand E, the most active compound, ranking fifth computationally) is almost 10000 times more potent than the *N*-formyl derivative (Ligand B, ranking eighth), and is 100 times more

potent than the *N*-aminoacetyl derivative (Ligand I, ranking first computationally). Therefore, using this protocol, because of the minimal differences in docked energies no correlation between the calculated energy and the biological data could be observed. As discussed earlier (Section 2.3.4) the AutoDock scores of the compounds are very similar and are not expected to be significantly different from one another.

Table 2.7: Scoring of flexibly docked C-5 modified analogues of Neu5Ac2en in NDV HN vs inhibition against NDV HN.⁶⁸



C-5-X Substituent	Flexible AutoDock (unminimised) (kcal/mol)	IC ₅₀ against NDV(M)
NHAc	-10.06	
A NHC(=O)NH ₂	-10.72 [2] ^a	5.2 × 10 ⁻⁴
B NHC(=O)H	-9.61 [8]	1.1 × 10 ⁻³
C NHC(=O)CH ₂ F	-10.16 [7]	1.4 × 10 ⁻⁵
D NHC(=O)CHF ₂	-10.67 [3]	4.4 × 10 ⁻⁶
E NHC(=O)CF ₃	-10.52 [5]	1.9 × 10 ⁻⁶
F NHC(=O)CH ₂ Cl	-10.72 [2]	1.3 × 10 ⁻⁵
G NHC(=O)CH ₂ I	-10.46 [6]	1.8 × 10 ⁻⁵
H NHC(=O)CH ₂ CN	-10.58 [4]	1.7 × 10 ⁻⁵
I NHC(=O)CH ₂ NH ₂	-11.80 [1]	2.0 × 10 ⁻⁴

^aNumbers in square brackets indicate scoring rank.

2.3.6 AutoDock of variously linked sialyllactoses

The natural substrate for both binding and sialidase activities of NDV HN are cell surface sialylated glycoconjugates. Substrate specificity is reported²² to be for α(2,3)-linked sialic acids, with α(2,8)-α(2,3)-linked sialic acids being more efficiently recognised. The crystal structure of NDV HN was co-crystallised by Neu5Ac. It was of interest to therefore examine whether AutoDock could predict the binding orientation of the substrate α(2,3) sialyllactose and of alternately linked sialyllactoses. α(2,3)-, α(2,6)- sialyllactose, and α(2,8)-α(2,3)-disialyllactose were each modelled using

InsightII and the geometry was optimised in vacuum to give each carbohydrate residue in the low energy chair conformation. These structures were then docked using AutoDock into NDV HN with default parameterisation.

The binding energies for the 10 conformations generated by the docking of $\alpha(2,3)$ sialyllactose into the active site of NDV HN ranged from -3.70 kcal/mol to -7.24 kcal/mol. The carboxylate of the sialic acid residue of all conformations generated was seen to make contacts to all three residues within the triarginyl cluster. However the sialic acid residue was seen to be partially moved out of the active site in all 10 structures.

The binding energies for the 10 conformations generated by docking of $\alpha(2,6)$ sialyllactose into the active site of NDV HN was -11.38 kcal/mol to -6.88 kcal/mol. The sialic acid residue was found docked in a similar orientation to that seen previously for bound Neu5Ac2en (Figure 2.17). The $\alpha(2,6)$ linked lactose residue was seen to be stacked above the sialic acid residue and to stay within the boundaries of the active site for all conformations generated.

The binding energies for the 10 conformations generated by the docking of $\alpha(2,8)$ - $\alpha(2,3)$ -disialyllactose into the active site of NDV HN was 5.73 kcal/mol to -5.14 kcal/mol. The structure that came closest to predicting the positioning of the sialic acid residue in the active site had an interaction energy of -4.63 kcal/mol. In this case, the sialic acid residue makes only one interaction with the triarginyl cluster, with Arg 174. No other interactions within the active site were seen. This computational result is not expected to correlate with how the substrate would biologically bind.

The analysis of the results of the computational binding of sialyllactose shows that the ligands do bind, however the orientation is not as would be expected in the natural binding of the ligands to NDV HN.

2.3.7 *De novo* designed C-4 modified analogues of Neu5Ac2en.

With the characterisation of active site interactions and the validation and benchmarking of docking and scoring programs complete, the next step in this work is the *de novo* ligand design. The program selected for this was LigBuilder. LigBuilder utilises four programs to accomplish *de novo* ligand design, Pocket (already used, Section 2.3.2.7) to characterise the active site interactions, Grow and Link to

accomplish the actual building of the ligands *via* either a growing or linking strategy, and Process to process the obtained results to the users specifications. The strategy selected for the design of inhibitors of NDV HN was the Grow strategy, specifying the seed structure and the site of modification, with the seed being Neu5Ac2en and the site being C-4. A total of four seed structures were given, seen in Figure 2.22, with sites of modification marked by an asterix. The first two seed structures were selected to examine what LigBuilder would identify as a linker off the C-4 position of Neu5Ac2en. The second two structures were selected to give LigBuilder many potential sites to grow from.

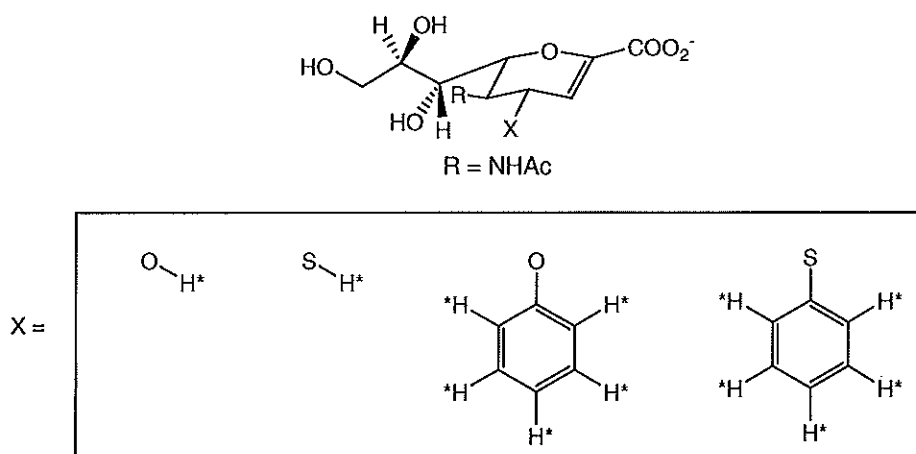


Figure 2.22: The structure of the four LigBuilder seeds. * indicates site of modification.

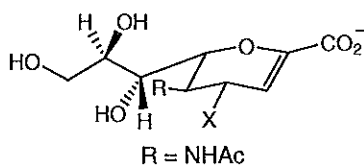
The LigBuilder procedure required the initial placement of the seed structure. This was accomplished by using AutoDock to dock the structures of the 4 seeds. The seed structures were then utilised by Grow to construct the molecules to user specified criteria (the input file is available in Appendix A). The fragment library used for the calculation is available with LigBuilder. A total of 20 generations were constructed and the results were processed to user specified criteria for a total of 90 molecules. Table 2.8 shows the structures of selected designed ligands. It can be seen that the structure of the designed substituents are predominately hydrophobic in nature (although many functional groups are available to Grow) as previously predicted by GRID (Section 2.3.2.6) and Pocket (Section 2.3.2.7).

During LigBuilder Grow calculations the oxygen/sulfur link between the ring of Neu5Ac2en and the attached alkyl or aralkyl chain was mutated to a carbon atom. This is due to a mutation algorithm present within Grow. The mutation algorithm mutates atoms as necessary to pick up interactions it perceives as favourable. At the end of the

Grow process the mutated atom was reverted back to the initial oxygen/sulfur atom to give compounds that could be readily synthesised. A total of 90 derivatives were designed and scored. The molecules presented in Table 2.8 represent the best in each category (4-*O*- or *S*- alkyl or aralkyl) of the *de novo* designed derivatives. SCORE, MMI and AutoDock were used to perform the initial static scoring of the *de novo* designed ligands. The static AutoDock results were highly variable, with the results ranging from expected values to in some cases in the hundreds of thousands. This was due to the inaccurate placement of the ligands. Visual inspection identified that some atoms of the ligands were placed through the protein surface. Such inaccurate placement attracts a heavy energy penalty in AutoDock. Flexible docking of the ligands alleviated this problem. The MMI energy for all ligands was lower than that found for Neu5Ac2en (-173.52 kcal). The reduction in interaction energy ranged from 32.7 to 98.7 %.

In all but three of the derivatives presented in Table 2.8, the final binding affinity of the flexibly docked ligands as measured by SCORE was lower than that of Neu5Ac2en (6.06, Section 2.3.3.2). The reduction in binding affinity may be due to an increase in the number of rotatable torsions in the C-4 side-chains. All presented derivatives increased in AutoDock binding energy compared to Neu5Ac2en (-10.06kcal/mol) with values ranging from 1.06 kcal/mol to 5.04 kcal/mol better than Neu5Ac2en.

Table 2.8: *De novo* designed C-4 O/S-linked derivatives of Neu5Ac2en.



X Substituent	MMI of designed ligands (kcal/mol)	SCORE of designed ligands (pK _a)	Static AutoDock of designed ligands (kcal/mol)	Flexible AutoDock of designed ligands (kcal/mol)	SCORE of Flexibly Docked Ligands (pK _a)
Neu5Ac2en-4-O-alkyl					
	-12.73 [6] ^a	4.07 [9]	28.93 [5]	-14.30 [2]	6.47 [2]
	-82.43 [3]	5.24 [7]	344.74 [7]	-11.15 [8]	4.79 [6]
Neu5Ac2en-4-O-aryl					
	-116.63 [1]	5.62 [6]	194.87 [6]	-11.12 [9]	5.69 [4]
Neu5Ac2en-4-S-alkyl					
	-2.21 [9]	6.31 [4]	35546 [8]	-11.71 [7]	6.55 [1]
	-2.30 [8]	4.97 [8]	122109.08 [9]	-12.76 [5]	6.45 [3]
Neu5Ac2en-4-S-aryl					
	-110.01 [2]	5.63 [5]	-1.64 [2]	-13.52 [4]	3.75 [9]
	-24.32 [5]	6.91 [1]	-4.34 [1]	-15.10 [1]	4.48 [8]
	-12.44 [7]	6.69 [2]	3.49 [3]	-13.58 [3]	5.52 [5]
	-35.67 [4]	6.36 [3]	24.48 [4]	-12.53 [6]	4.65 [7]

^aNumbers in square brackets indicate scoring rank.

2.3.7.1 Refinement of ligands to final forms.

Once the top ligands had been selected, the structures were analysed based on synthetic potential (ease of synthesis), and availability of synthetic materials. Most

molecules obtained from LigBuilder were modified in some fashion. Modifications included the addition of a methylene group between the *O* or *S* atom attached to C-4 of Neu5Ac2en and the aromatic carbon of a phenyl ring, and the removal of excessive flexible bonds deemed to be unnecessary and that could potentially cause an energy penalty upon binding due to the freezing of flexible bonds. An example of the refinement of the derivatives is shown in Figure 2.23 for the top scoring derivative in the AutoDock score: the 2-(5'-cyclohexyl-3'-ethyl-penta-1',4'-dienyl)-phenyl substituent was developed into the 2-phenylbenzyl derivative. Figure 2.23 illustrates the process of structural refinement showing the removal of the cyclohexyl group and the ethyl side chain, the addition of a CH₂ group between the sulfur and the aromatic ring attached to C-4, and the fusing and the unsaturation of the pentyl side chain into an additional aromatic ring. In this way, the derivative was refined to make it less flexible but to retain key structural and hydrophobic characteristics.

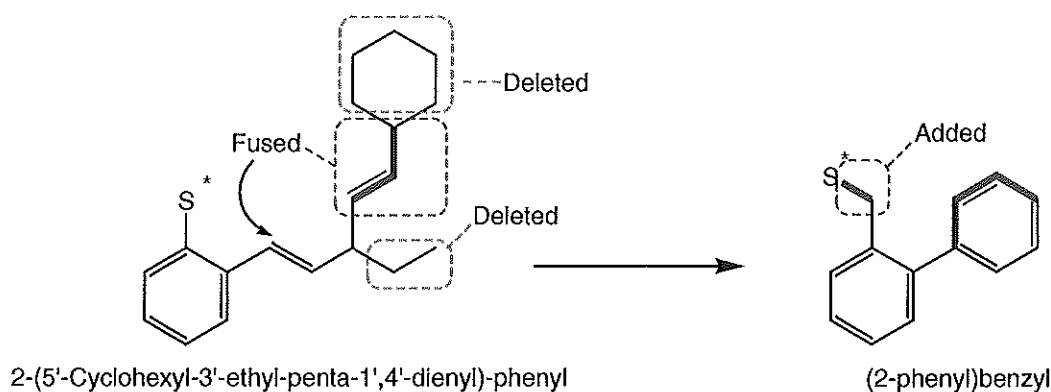


Figure 2.23: An example of the process of structure refinement.

Some of the additional molecules that were added to the set of substituents are extensions of existing substituents, such as the 2-phenylethyl substituent (B in Figure 2.24) being an extension of the benzyl derivative. Figure 2.24 shows the substituents chosen in their final form after refinement.

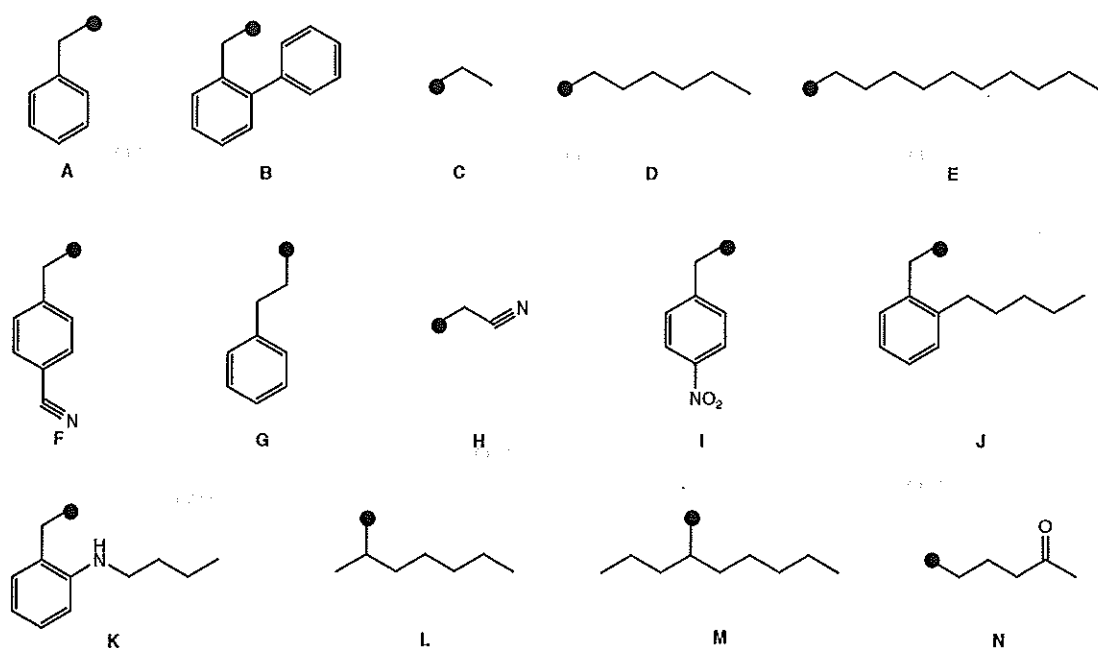


Figure 2.24: Structure of the final 14 substituents, with the point of attachment to the core template indicated.

The chosen fourteen substituents: benzyl, **A**, (2-phenyl)benzyl, **B**, ethyl, **C**, hexyl, **D**, decyl, **E**, *p*-(cyano)benzyl, **F**, 2-phenylethyl, **G**, cyanomethyl, **H**, *p*-(nitro)benzyl, **I**, (2-pentyl)benzyl, **J**, (2-butylamino)benzyl, **K**, 2-heptyl, **L**, 4-nonyl, **M**, and (4-oxo)pentyl, **N**; were added *via* an oxygen or sulfur ether-linkage to the C-4 position of 4-*S*-Neu5Ac2en **28**, 4-*O*-Neu5Ac2en, **29**, and ten other templates: 4-*S*-Gilead template, **30**, 4-*O*-Gilead template, **31**, 4-*S*-benzoic acid template, **32**, 4-*O*-benzoic acid template, **33**, 4-*S*-uronic acid template, **34**, 4-*O*-uronic acid template, **35**, 4-*S*-BCX-1812 template, **36**, 4-*O*-BCX-1812 template, **37**, 4-*S*-furanose template, **38** and 4-*O*-furanose template, **39**. The structures of all 12 core templates are shown in Figure 2.25.

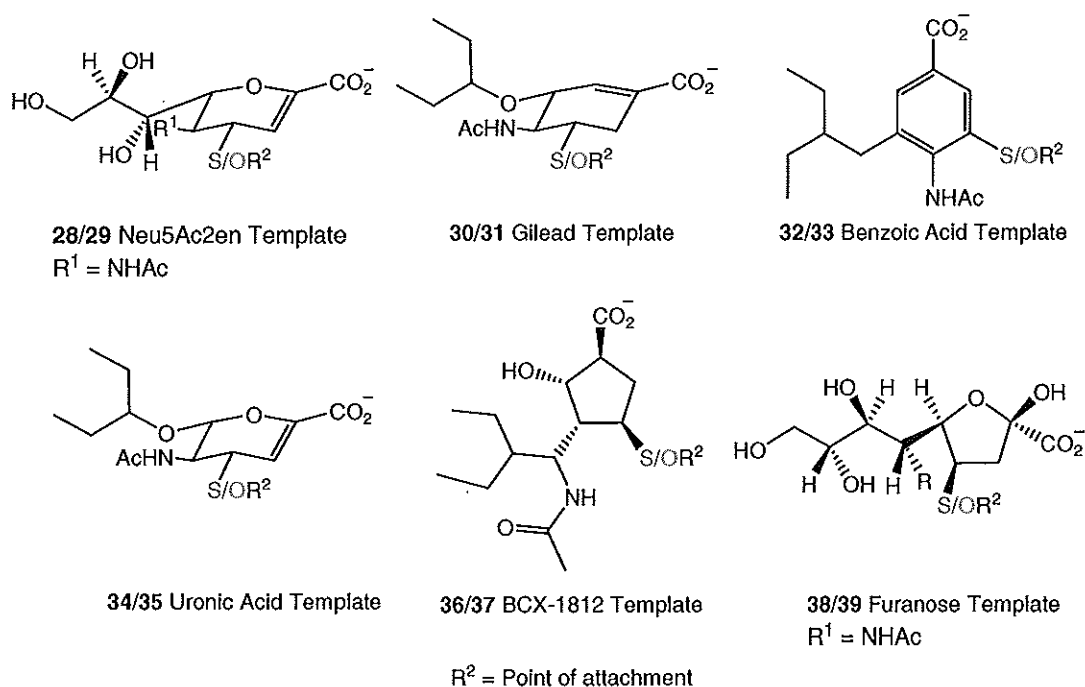


Figure 2.25: Structures of the 12 templates used in ligand design.

The alternative templates were selected on the basis of their having activity against influenza virus sialidase. Each of them has a unique difference from Neu5Ac2en which could shed some light on the particular substituents at a given place on the ring structure. In addition, using the different templates may shed some light on the type of ring structure needed for optimal binding, whether it is flexible or rigid or either a pyranose or furanose conformation.

Due to the refinement of the substituents designed by Grow to the final forms (Figure 2.24), scoring and docking calculations needed to be repeated. The initial modelling protocol used for the scoring of each of the ligands was:

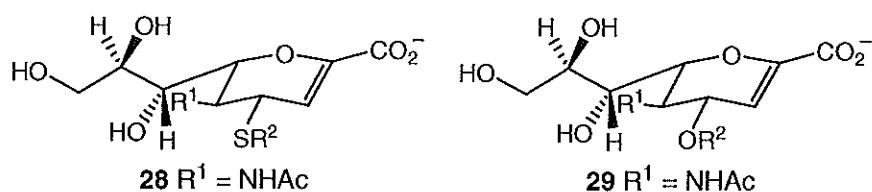
1. Each ligand was flexibly docked using AutoDock to place the ligand in an acceptable conformation within the protein structure for minimisation.
2. Each ligand was then minimised using Discover.
3. Following the minimisation process each ligand was statically scored with MMI, SCORE and AutoDock.

2.3.7.2 Docking and scoring of refined de novo designed substituents – *Initial examination of the Neu5Ac2en template*

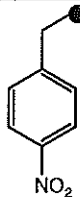
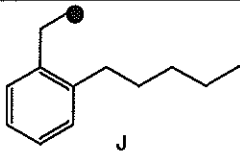
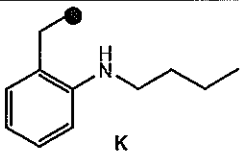
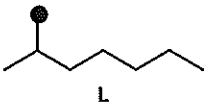
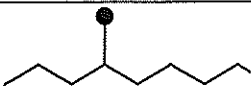
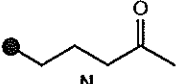
The first of the templates examined was the Neu5Ac2en template due to it being available in the crystal structure, and because previous calculations had utilised the Neu5Ac2en template. As stated earlier, the fourteen substituents (Figure 2.24) were linked through either a sulfur or oxygen atom to the C-4 position of Neu5Ac2en: the *S* linkage was used due to the availability of a suitable synthetic template,¹¹⁷ and the *O* linkage was modelled to compliment the *S*-linked data and was also accessible through a suitable synthetic template.¹¹⁸ The major differences between the two types of compounds conferred by the oxygen or sulfur atoms, is a longer bond length and smaller bond angle between sulfur and carbon,¹¹⁹ and the increased polarisability of the carbon-sulfur bond, due to the d shell orbitals and the reduced electronegativity of sulfur (compared to oxygen).¹¹⁹

Initially, a crude examination of the docking and scoring of the derivatives was used to identify which molecules were desirable, the full protocol can be found in Chapter 5. Briefly, the method consisted of a fast (250,000 energy evaluations) flexible AutoDock of the designed ligands to place each ligand in a suitable position for minimisation. Minimisation was then performed to remove bad contacts and optimise the hydrogen bond network. After minimisation the ligands were scored with the programs Discover (Molecular Mechanics Interaction energy), AutoDock (in Static Mode) and SCORE. Ligands were then ranked according to SCORE and consensus scored. In all calculations, CVFF charges were used. The results for the combination of the 14 templates and 2 templates (4-*S*- and 4-*O*- Neu5Ac2en) are given in Table 2.9

Table 2.9: Scoring results for Neu5Ac2en derivatives from AutoDock, MMI and SCORE.



R^2 Substituent	C-4 Link	AutoDock (Unminimised) kcal/mol	MMI (Minimised) kcal/mol	AutoDock (Minimised) kcal/mol	SCORE (Minimised) pK_d
 A	S	-16.69 [1] ^a	-189.35 [8]	-17.78 [2]	7.31 [11]
	O	-17.10 [1]	-254.58 [1]	-17.79 [1]	5.36 [14]
 B	S	-16.25 [2]	-191.77 [6]	-18.37 [1]	9.39 [2]
	O	-15.71 [4]	-200.80 [6]	-15.19 [5]	7.53 [6]
 C	S	-14.17 [8]	-177.67 [10]	-11.86 [11]	7.75 [8]
	O	-14.10 [10]	-193.33 [10]	-12.66 [9]	7.40 [7]
 D	S	-15.07 [6]	-176.81 [11]	-16.81 [5]	8.09 [7]
	O	-16.04 [3]	-194.52 [8]	-16.03 [4]	6.94 [10]
 E	S	-15.09 [5]	-171.93 [12]	-15.35 [9]	6.47 [12]
	O	-16.71 [2]	-180.59 [13]	-10.34 [13]	6.79 [11]
 F	S	-14.09 [9]	-187.10 [9]	-10.88 [12]	8.29 [6]
	O	-14.95 [6]	-200.10 [7]	-13.25 [7]	7.74 [5]
 G	S	-10.80 [14]	-198.02 [3]	-8.47 [13]	8.94 [3]
	O	-11.78 [13]	-211.83 [4]	-16.28 [3]	8.95 [2]
 H	S	-15.14 [4]	-190.75 [7]	-7.20 [14]	7.47 [10]
	O	-14.37 [9]	-194.405 [9]	-12.59 [10]	7.01 [9]

R ² Substituent	C-4 Link	AutoDock (Unminimised) kcal/mol	MMI (Minimised) kcal/mol	AutoDock (Minimised) kcal/mol	SCORE (Minimised) pK _d
	S	-13.87 [11]	-206.38 [1]	-17.10 [4]	10.09 [1]
	O	-14.66 [7]	-208.96 [3]	-14.11 [6]	9.15 [1]
 J	S	-13.54 [12]	-190.54 [5]	-12.49 [10]	7.59 [9]
	O	-11.67 [14]	-219.20 [2]	-17.66 [2]	8.25 [4]
 K	S	-14.05 [10]	-168.78 [13]	-15.76 [8]	5.73 [13]
	O	-12.11 [12]	-187.60 [11]	-13.22 [8]	7.40 [7]
 L	S	-16.21 [3]	-194.03 [4]	-16.33 [6]	8.79 [4]
	O	-15.24 [5]	-147.21 [14]	-11.91 [12]	7.27 [8]
 M	S	-13.15 [13]	-148.24 [14]	-15.89 [7]	5.45 [14]
	O	-13.17 [10]	-188.70 [11]	-12.20 [11]	8.79 [3]
 N	S	-14.70 [7]	-198.18 [2]	-17.41 [3]	8.64 [5]
	O	-14.65 [8]	-201.09 [5]	-9.03 [14]	7.40 [7]

^a Numbers in square brackets indicate scoring rank (Rank is within compound series).

2.3.7.2.1 C-4 Sulfur linked derivatives.

The initial flexible AutoDock scores in the unminimised protein ranged from -10.80 to -16.69 kcal/mol, compared to Neu5Ac2en at -10.06 kcal/mol (Section 2.3.3.1). Derivative **28-A** scored the highest (6.63kcal/mol higher than Neu5Ac2en), and derivative **28-M** scored the lowest (0.74kcal/mol higher than Neu5Ac2en). The derivatives were minimised and re-scored. The interaction energies of each derivative of Neu5Ac2en were determined by Discover. Using the Molecular Mechanics Interaction energy parameters, the interaction energy of the ligands were determined with respect to the protein active site residues. The interaction energies of each derivative with the minimised NDV active site ranged from -148.24 to -206.38 kcal/mol. The MMI interaction energy of crystal structure Neu5Ac2en with NDV HN

is -173.52 kcal/mol. Derivative **28-I** scored the highest (32.86 kcal/mol higher than Neu5Ac2en) and derivative **28-M** scored the lowest (25.28 kcal/mol lower than Neu5Ac2en).

After the minimisation procedure, AutoDock was used solely as a scoring function (no flexible torsions). In 9 out of the 14 derivatives, the static AutoDock energy of each compound increased after minimisation. The change in the AutoDock energy of interaction ranged from -7.94 to 3.23 kcal/mol after minimisation. Derivative **28-H** demonstrated the loss of 7.94 kcal/mol, whereas derivative **28-I** showed the greatest improvement of 3.23 kcal/mol. Neu5Ac2en compared with the pre-minimisation AutoDock score of -10.06 kcal/mol showed a 2.36 kcal/mol increase in the energy of interaction to -12.42 kcal/mol (Section 2.3.4.1, Table 2.6). The highest scoring derivative **28-B** scores 5.95 kcal/mol better than minimised Neu5Ac2en, and the lowest scoring derivative **28-H** scores 5.22 kcal/mol worse than minimised Neu5Ac2en.

The SCORE results ranged on a logarithmic scale from 5.45 with derivative **28-M** to 10.09 with derivative **28-I**. The difference in binding affinity between the highest and lowest scoring derivatives was approximately 43650 times better. The minimised crystal structure containing Neu5Ac2en has a SCORE result of 6.06 (Section 2.3.3.2). The results ranged between approximately 4 times less to approximately 10700 times greater than Neu5Ac2en.

Ligplot analyses of the interactions formed by the fourteen derivatives showed that interactions were similar to those of Neu5Ac2en. All three arginine residues from the triarginyl cluster form hydrogen bonds with the carboxylate group of each derivative with the exception of derivative **28-K**. In **28-K** the carboxylate group formed hydrogen bonds with two arginine residues of triarginyl cluster. Glycerol side chain interactions were found to vary between each of the designed derivatives and the seed structure Neu5Ac2en **5**; this could be due to the flexible nature of the glycerol side chain and the large size of the glycerol side-chain binding pocket. A minimum of one hydrogen bond was formed, with most of the derivatives forming two or more. One or more possible hydrophobic interactions (Ile 175, His 199, and Ile 192) were present within the C-4 binding pocket of most of the derivatives. These hydrophobic interactions represent the major difference between the crystal structure bound Neu5Ac2en and the designed derivatives.

2.3.7.2.2 C-4 Oxygen linked derivatives

The preliminary (flexibly docked) AutoDock score for each of the *O*-linked derivatives ranged from -11.67 to -17.10 kcal/mol. Derivative **29-A** scored the highest, 7.04 kcal/mol higher than Neu5Ac2en (-10.06 kcal/mol), and derivative **29-J** scored the lowest (1.61 kcal/mol higher than Neu5Ac2en). The interaction energies (MMI) of each of the *O*-linked derivatives of the Neu5Ac2en template with the minimised NDV active site ranged from -147.21 to -254.58 kcal/mol. Derivative **29-A** scored the highest (81.06 kcal/mol higher than Neu5Ac2en at -173.52 kcal/mol) and derivative **29-L** scored the lowest (26.31 kcal/mol lower than Neu5Ac2en).

After disabling the flexible torsions of the ligands, the static AutoDock energy of 4 of the 14 compounds increased after the minimisation process. The change in the AutoDock energy score ranged from -6.37 to 5.99 kcal/mol after minimisation. Derivative **29-E** demonstrated the loss of 6.37 kcal/mol, whereas derivative **29-J** showed the greatest improvement of 5.99 kcal/mol. Derivative **29-A** showed the greatest increase over the benchmark Neu5Ac2en (-12.42 kcal/mol) with an improvement of 5.37 kcal/mol. Derivative **29-N** shows the greatest loss (compared with Neu5Ac2en) with an interaction energy 3.39 kcal/mol less than Neu5Ac2en.

The SCORE results ranged from 5.36 with derivative **29-A** to 9.15 with derivative **29-G**; the highest scoring substituent is approximately 6200 times stronger than the lowest scoring derivative with respect to binding affinity. The results range between approximately 5 times lower to approximately 1200 times greater than the binding affinity of Neu5Ac2en (pK_d 6.06).

Ligplot analyses of the interactions formed by the fourteen derivatives showed that all three arginine residues from the triarginyl cluster formed hydrogen bonds with the carboxylate group of each derivative. Similarly to the sulfur-linked derivatives, glycerol side-chain interactions were found to vary between each of the derivatives and Neu5Ac2en. A minimum of one hydrogen bond was formed, with most derivatives forming two or more. One or more possible hydrophobic interactions (Ile 175, His 199, and Ile 192) were present within the C-4 binding pocket of most derivatives, with the exception of derivative **29-C** which was found to make no additional interactions from those formed by Neu5Ac2en.

2.3.7.3 *Improved computational examination of Neu5Ac2en template*

Upon examination of the modelling protocol and with the availability of new programs after the initial evaluation was complete, it was decided that the designed derivatives would undergo re-evaluation. This was to take advantage of parameterisation that AutoDock was designed to use. As AutoDock was going to be used to re-evaluate the ligands, it was decided that the conformational search parameters would be increased to evaluate predicted ligand binding more rigorously. The AMBER package was selected to perform the minimisation to take advantage of the gaff forcefield that is designed to be used with small ligands.

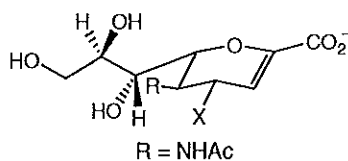
The new modelling protocol used for the scoring of each of the ligands was:

1. Each ligand was flexibly docked using AutoDock to place the ligand in an acceptable conformation within the protein structure for minimisation. This protocol was updated to use a higher number of energy evaluations and the Gasteiger charge set for the ligand. Previous calculations (Section 2.3.7.2) used the CVFF charge set.
2. Each ligand was then minimised using the AMBER 7 package.
3. Following the minimisation process each ligand was statically scored with MM-PBSA (MMI with the addition of solvation parameters), X-Score (updated SCORE) and AutoDock (optimised charge set).

2.3.7.3.1 Re-examination of Neu5Ac2en

The scoring of Neu5Ac2en against NDV HN was performed once again as new benchmarks were required for the new modelling protocol. The results obtained for Neu5Ac2en are shown in Table 2.10.

Table 2.10: Scoring results for Neu5Ac2en using updated modelling protocol (Section 2.3.7.3).



C-4-X-substituent	AutoDock (Unminimised) kcal/mol	MM-PBSA (Minimised) kcal/mol.	AutoDock (Minimised) kcal/mol	X-Score (Minimised) pK_d
OH	-12.01 ^a	-102.50	-6.36	5.12
	-10.06 ^b	MMI: -173.52	-12.42	SCORE: 6.06

^a Row of new benchmark scores using modelling protocol from Section 2.3.7.3. ^b Row of benchmark scores from initial modelling protocol (Section 2.3.4.1).

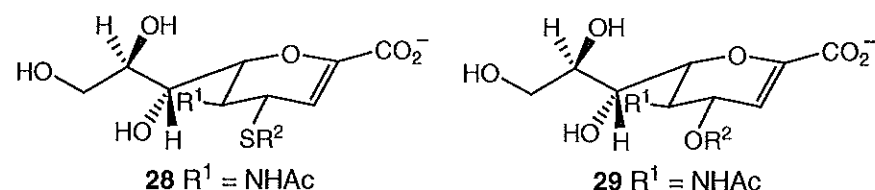
The flexible unminimised AutoDock score for Neu5Ac2en was found to be -12.01 kcal/mol, which is comparable to that found with unminimised score with the first method (-10.06 kcal/mol, Section 2.3.3.1). The minimised static AutoDock score however varies from the first protocol (Section 2.3.4.1) score by an appreciable amount (-6.36 vs -12.42 kcal/mol). The difference between the two scores could possibly be explained by the use of two different force fields during 3 stages of the second modelling protocol. The first protocol employed the same force field charges throughout the calculation process, therefore optimising and minimises using the same charges throughout the process. The second protocol however, optimises and evaluates different charges during each step of the calculation process. The MM-PBSA interaction energy was evaluated to -102.50 kcal/mol. This interaction energy represents both the molecular mechanics interaction energy and also a solvation energy as calculated by the program APBS. With the solvation energy being a positive quantity (62.23 kcal/mol), this accounts for most of the difference between the MMI score of -173.52 kcal/mol and the MM-PBSA score of -102.50. However, as both scoring protocols test different parameters involved in calculating binding energies, no useful comparison can be made between the two calculations. The X-Score score value was approximately 9 times lower than that found with the first method SCORE (5.12 vs 6.06).

2.3.7.3.2 Revised docking and scoring of C-4 substituted Neu5Ac2en derivatives.

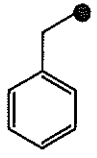
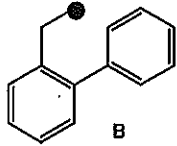

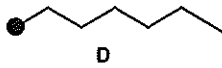
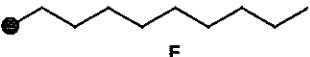
A more thorough examination of the docking and scoring of the Neu5Ac2en derivatives previously described (Section 2.3.7.2) was used to refine the initial data

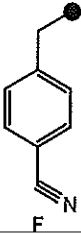
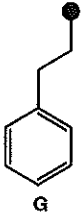
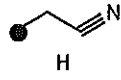
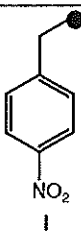
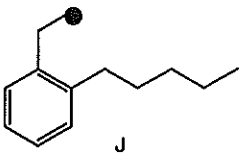
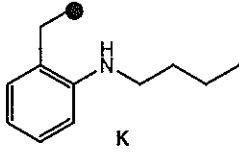
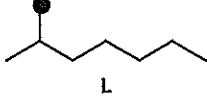
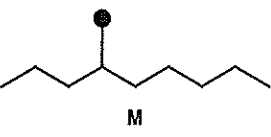
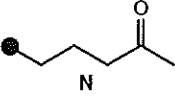
collected and to determine which molecules were of highest importance for synthesis. The full protocol can be found in Chapter 5. Briefly, the method consisted of a more extensive initial search (1,000,000 energy evaluations, previously 250,000 evaluations) with flexible AutoDock (using Kollman charges for protein and Gasteiger charges for ligands) of the designed ligands to place each ligand in a suitable position for minimisation. Minimisation with AMBER using the gaff force field for each ligand was performed to remove bad contacts and optimise the hydrogen bond network. After minimisation the ligands were scored with the programs AMBER/APBS (Molecular Mechanics Interaction and Solvation energies, MM-PBSA), AutoDock (in Static Mode with Kollman and Gasteiger charges) and X-Score. Ligands were then ranked according to score and consensus scored. The results for the 28 Neu5Ac2en derivatives are given in Table 2.11. Energy values and scores were compared to those for Neu5Ac2en given in Table 2.10.

Table 2.11: Scoring results for Neu5Ac2en derivatives from AutoDock, MM-PBSA and X-Score.



28 $R^1 = \text{NHAc}$ **29** $R^1 = \text{NHAc}$

R^2 Substituent	C-4 Link	AutoDock (Unminimised) kcal/mol	MM-PBSA (Minimised) kcal/mol.	AutoDock (Minimised) kcal/mol	X-Score (Minimised) pK_d
 A	S	-14.69 [11] ^a	-116.88 [6]	-13.74 [8]	6.32 [6]
	O	-15.24 [11]	-110.79 [3]	-13.26 [1]	6.26 [1]
 B	S	-15.85 [7]	-113.69 [8]	-14.27 [6]	6.69 [1]
	O	-16.71 [2]	-108.59 [4]	-11.97 [5]	5.94 [3]
 C	S	-12.82 [14]	-110.44 [10]	-12.04 [13]	5.60 [13]
	O	-15.59 [9]	-82.77 [12]	-10.03 [9]	5.32 [9]
 D	S	-15.24 [9]	-94.29 [13]	-13.29 [10]	5.81 [11]
	O	-15.73 [8]	-106.07 [5]	-11.93 [6]	5.54 [8]
 E	S	-17.31 [1]	-124.62 [4]	-15.86 [1]	6.10 [9]
	O	-18.12 [1]	-88.67 [11]	-11.70 [7]	4.97 [13]

R ² Substituent	C-4 Link	AutoDock (Unminimised) kcal/mol	MM-PBSA (Minimised) kcal/mol.	AutoDock (Minimised) kcal/mol	X-Score (Minimised) pK _d
	S	-13.96 [12]	-124.93 [3]	-13.26 [11]	6.18 [8]
	O	-13.66 [13]	-18.67 [14]	-12.16 [3]	5.94 [3]
	S	-16.05 [5]	-96.93 [12]	-14.25 [7]	6.29 [7]
	O	-16.68 [3]	-114.41 [2]	-9.00 [11]	5.72 [6]
	S	-13.86 [13]	-114.19 [7]	-11.39 [14]	5.32 [14]
	O	-13.02 [14]	-100.87 [6]	-10.50 [8]	5.08 [11]
	S	-15.36 [8]	-82.50 [14]	-12.07 [12]	6.09 [10]
	O	-16.36 [4]	-94.22 [8]	-12.04 [4]	5.89 [4]
	S	-16.46 [3]	-110.46 [9]	-15.07 [4]	6.38 [5]
	O	-15.91 [5]	-44.86 [13]	-5.92 [13]	5.04 [12]
	S	-17.06 [2]	-119.19 [5]	-15.75 [2]	6.58 [3]
	O	-15.90 [6]	-90.74 [10]	-12.76 [2]	6.19 [2]
	S	-15.93 [6]	-98.67 [11]	-14.91 [5]	6.43 [4]
	O	-15.89 [7]	-118.24 [1]	-12.76 [2]	5.80 [5]
	S	-16.18 [4]	-125.91 [2]	-15.42 [3]	6.66 [2]
	O	-15.39 [10]	-91.32 [9]	-8.93 [12]	5.70 [7]
	S	-14.75 [10]	-128.46 [1]	-13.44 [9]	5.80 [12]
	O	-15.18 [12]	-95.74 [7]	-9.48 [10]	5.15 [10]

^a Numbers in square brackets indicate scoring rank (Rank is within compound series).

2.3.7.3.3 C-4 Sulfur linked derivatives.

The flexible AutoDock scores in the unminimised protein ranged from -12.82 to -17.31 kcal/mol, compared to Neu5Ac2en at -12.01 kcal/mol. Derivative **28-E** scored the highest (5.30 kcal/mol higher than Neu5Ac2en), and derivative **28-C** scored the lowest (0.81 kcal/mol higher than Neu5Ac2en). The derivatives were minimised and re-scored.

The interaction energies of each derivative of Neu5Ac2en were determined by MM-PBSA. The interaction energy / solvation energy of the ligands were determined with respect to the protein active site residues. The interaction energies of each derivative with the minimised NDV active site ranged from -82.50 to -128.46 kcal/mol, compared to Neu5Ac2en itself with -102.50 kcal/mol. Derivative **28-N** scored the highest (25.96 kcal/mol higher than Neu5Ac2en) and derivative **28-I** scored the lowest (20.00 kcal/mol lower than Neu5Ac2en).

The static AutoDock energy of each compound ranged from -11.39 to -15.86 kcal/mol. Change in AutoDock energy of interaction ranged from 0.70 to 3.29 kcal/mol after minimisation. Derivative **28-F** demonstrated the small increase of 0.70 kcal/mol where as derivative **28-I** showed the greatest change of 3.29 kcal/mol. For Neu5Ac2en, a 5.65 kcal/mol decrease in AutoDock score was seen after minimisation (Table 2.10). The highest scoring derivative **28-I** scores 9.50 kcal/mol better than minimised Neu5Ac2en (-6.36 kcal/mol), and the lowest scoring derivative **28-F** scores 5.03 kcal/mol better than minimised Neu5Ac2en.

The X-Score results ranged on a logarithmic scale from 5.32 with derivative **28-H** to 6.69 with derivative **12-B**. The difference in binding affinity between the highest and lowest scoring derivatives was approximately 23 times. The results range between approximately 1.6 times greater to approximately 37 times greater than Neu5Ac2en (pK_d 5.12).

Ligplot analyses of the interactions formed by the fourteen derivatives were showed the interactions to be similar to that of Neu5Ac2en. The binding interactions shown are very similar to those seen with the previous docking protocol (Section 2.3.7.2.1).

2.3.7.3.4 C-4 Oxygen linked derivatives

The initial flexible AutoDock score for each of the *O*-linked derivatives ranged from –13.02 to –18.12 kcal/mol, a difference of 5.10 kcal/mol. Derivative **29-E** scored the highest (6.11 kcal/mol higher than Neu5Ac2en), compared to Neu5Ac2en at –12.01 kcal/mol, and derivative **29-H** scored the lowest (1.01 kcal/mol higher than Neu5Ac2en).

The MM-PBSA interaction energies of each of the *O*-linked derivatives of the Neu5Ac2en template with the minimised NDV active site ranged from –18.67 to –118.24 kcal/mol (compared to Neu5Ac2en –102.50 kcal/mol). Derivative **29-L** scored the highest (15.74 kcal/mol higher than Neu5Ac2en) and derivative **29-F** scored the lowest (83.83 kcal/mol lower than Neu5Ac2en).

The static AutoDock energy of each compound ranged between –5.92 to –13.26 kcal/mol. Change in AutoDock energy of interaction ranged from 1.50 to 9.99 kcal/mol after minimisation. Derivative **29-A** demonstrated the greatest change of 9.99 kcal/mol where as derivative **29-F** showed the least change of 1.50 kcal/mol. The highest scoring derivative **29-A** scores 6.90 kcal/mol better than minimised Neu5Ac2en, and the lowest scoring derivative **29-F** scores 0.44 kcal/mol worse than minimised Neu5Ac2en.

The X-Score results ranged from 4.97 with derivative **29-E**, to 6.26 with derivative **29-A**. The highest scoring substituent is approximately 19.5 times stronger than the lowest scoring derivative, with respect to binding affinity. The results range between approximately 0.7 times lower to approximately 1.14 times greater than the binding affinity of Neu5Ac2en (pK_d 5.12).

Ligplot analysis of the interactions formed by the fourteen derivatives found that all three arginine residues from the triarginyl cluster formed hydrogen bonds with the carboxylate group of half of the derivatives, the exceptions being derivatives **29-A**, **B**, **E**, **F**, **J**, **M**, and **N** which formed interactions with only two of the three arginine residues. Glycerol side chain interactions were found to vary between each of the derivatives and Neu5Ac2en. A minimum of one hydrogen bond was formed, with most derivatives forming two or more. One or more possible hydrophobic interactions (Ile 175, Ile 192, and Lys 236) are present within the C-4 binding pocket of most derivatives.

Consensus scoring of the designed ligands shows that in most cases the designed derivatives perform adequately (scoring higher than the parent compound, Neu5Ac2en) in two out of the three scoring functions. Exceptions were the *O*-linked derivatives **29-E**, **H** and **J** which each failed on the X-Score and MM-PBSA scoring functions. Interestingly each of the corresponding derivatives in the sulfur series **28-E**, **H** and **J** performed adequately.

2.3.8 Comparison of selected C-4-S- and -*O*-linked Neu5Ac2en derivatives

Before the comparative analysis of the computational structures is made, it is necessary to point out that each of the structures obtained is a single snap-shot of the potential interactions of each ligand with the NDV HN protein and as such does not represent the complete binding conformations available to the ligands. To elucidate such binding, molecular dynamics of each ligand would be required. Molecular dynamics will allow for both ligand and protein flexibility. This process is however, very computationally expensive. This said, the structures presented in the following analysis represent the best (based on binding energy) single binding conformation out of one million conformations analysed (Section 2.3.7.3).

2.3.8.1 Comparison of 4-*O*- and 4-*S*-benzyl-Neu5Ac2en computational structures.

In the two best structures of the 4-*O*- and 4-*S*-benzyl substituted Neu5Ac2en derivatives in NDV HN, shown overlaid in Figure 2.26, the benzyl ring of 4-*O*-benzyl-Neu5Ac2en appears to make better surface contact with the protein than the *S*-linked derivative; it lies flat against the protein surface as opposed to having an angular orientation of the *S*-linked ligand. Modelling shows the 4-benzyl substituent of both derivatives in similar areas in the large C-4 pocket. However, due to a shift in the interaction of the *S*-linked derivative and to the longer nature of the C-S bond, the flat alignment of the benzyl substituent of the *S*-linked derivative is not possible. However, it can be seen in the comparative data in Table 2.12 that the *S*-linked derivative in all scoring functions is predicted to bind better to the HN glycoprotein.

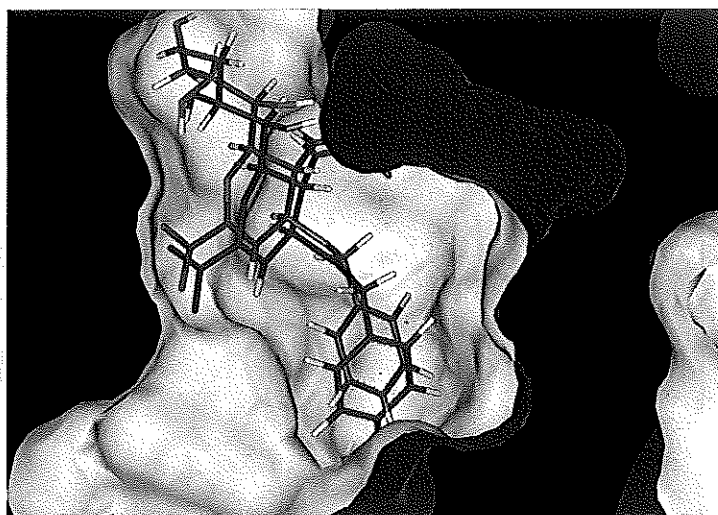


Figure 2.26: Superimposition of the lowest energy structures of 4-*O*-benzyl-Neu5Ac2en (green) and 4-*S*-benzyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.

Table 2.12: Comparative data for 4-*S*- and 4-*O*-benzyl-Neu5Ac2en.^a

Compound		MM-PBSA kcal/mol	AutoDock kcal/mol	X-Score pK _d
5	Neu5Ac2en	-102.50	-6.36	5.12
28-A	4- <i>S</i> -benzyl-Neu5Ac2en	-116.88 [1] ^a	-13.74 [1]	6.32 [1]
29-A	4- <i>O</i> -benzyl-Neu5Ac2en	-110.79 [2]	-13.26 [2]	6.26 [2]

^a Numbers in square brackets indicate relative ranking of computational score.

2.3.8.2 Comparison of 4-*S*- and 4-*O*-(2-phenyl)benzyl-Neu5Ac2en computational structures.

The comparison of computational structures of both lowest energy structures 4-*O*- and 4-*S*-(2-phenyl)benzyl Neu5Ac2en in NDV HN, is seen overlaid in Figure 2.27. Figure 2.27 shows that the α phenyl rings of both structures are positioned in the same area of the C-4 pocket, however the β phenyl rings are positioned perpendicular to one another. In the *O*-linked structure, the β phenyl ring is pointing out of the active site allowing the α phenyl ring to adopt a conformation that is similar to that seen in the 4-*O*-benzyl Neu5Ac2en structure (Figure 2.26). The β phenyl ring in the *S*-linked structure however is seen to point towards the C-5 binding pocket. Given the different positions and therefore different interactions of the biphenyl system, in particular the orientation of the β phenyl ring; of the two structures it is to be expected that the thioether linked derivative would score better in all scoring functions and indeed this is the case. The comparative computational scoring data can be seen in Table 2.13.

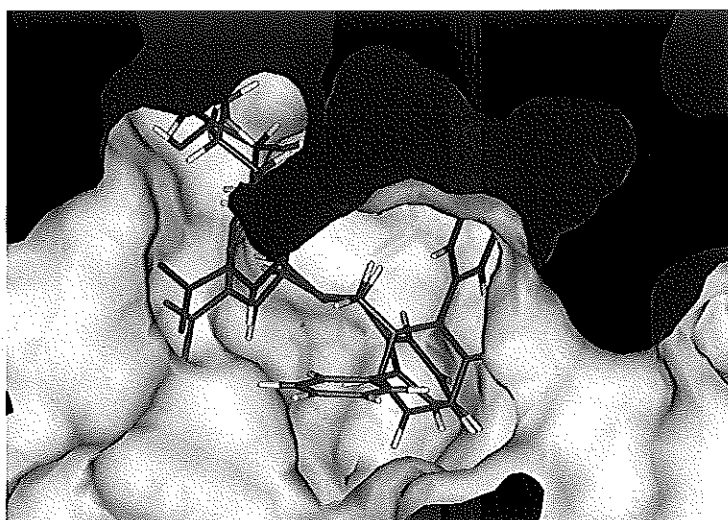


Figure 2.27: Superimposition of the lowest energy structures of 4-*O*-(2-phenyl)benzyl-Neu5Ac2en (green) and 4-*S*-(2-phenyl)benzyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.

Table 2.13: Comparative data for 4-*S*- and 4-*O*-(2-phenyl)benzyl-Neu5Ac2en.^a

Compound	MM-PBSA	AutoDock	X-Score
	kcal/mol	kcal/mol	pK _d
5 Neu5Ac2en	-102.50	-6.36	5.12
28-B 4- <i>S</i> -(2-phenyl)benzyl-Neu5Ac2en	-113.69 [1] ^a	-14.27 [1]	6.69 [1]
29-B 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en	-108.59 [2]	-11.97 [2]	5.94 [2]

^a Numbers in square brackets indicate relative ranking of computational score.

2.3.8.3 Comparison of 4-*S*- and 4-*O*-ethyl-Neu5Ac2en computational structures.

The superimposition of the lowest energy complexes of 4-*O*-ethyl-Neu5Ac2en and 4-*S*-ethyl-Neu5Ac2en is shown in Figure 2.28. *In silico*, the *S*-linked derivative is predicted to bind most favourably, with this trend being seen in all scoring functions (Table 2.14). There is a substantial difference in the positioning of the C-5 acetamido group of the *O*-linked derivative. This is usually an important interaction in Neu5Ac-based ligands binding to neuraminidases.^{48, 51, 53} The C-5 pocket in NDV HN is large, which would allow for the acetamido group to move, but may cause a reduction in overall interaction energy of the 4-*O*-ethyl derivative.

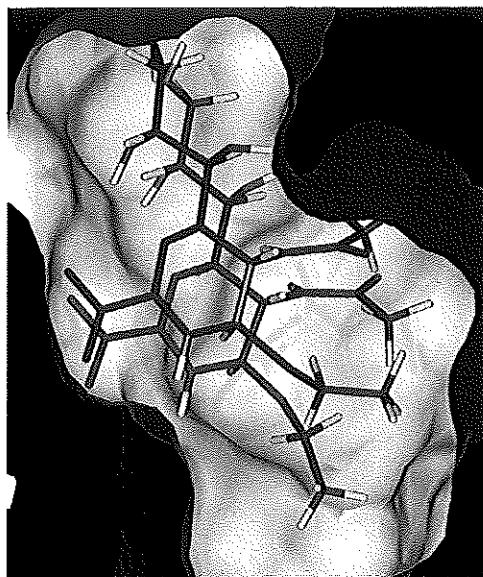


Figure 2.28: Superimposition of the lowest energy structures of 4-*O*-ethyl-Neu5Ac2en (green) and 4-*S*-ethyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.

Table 2.14: Comparative data for 4-*S*- and 4-*O*-ethyl-Neu5Ac2en.^a

Compound		MM-PBSA kcal/mol	AutoDock kcal/mol	X-Score pK _d
5	Neu5Ac2en	-102.50	-6.36	5.12
	28-C 4- <i>S</i> -ethyl-Neu5Ac2en	-110.44 [1] ^a	-12.04 [1]	5.60 [1]
	29-C 4- <i>O</i> -ethyl-Neu5Ac2en	-82.77 [2]	-10.03 [2]	5.32 [2]

^aNumbers in square brackets indicate relative ranking of computational score.

2.3.8.4 Comparison of 4-*S*- and 4-*O*-hexyl-Neu5Ac2en computational structures.

The superimposition of the lowest energy complexes of 4-*O*-hexyl-Neu5Ac2en and 4-*S*-hexyl-Neu5Ac2en can be seen in Figure 2.29. *In silico*, the *S*-linked derivative is predicted to bind more favourably than the *O*-linked derivative in two out of the three scoring functions, with the exception being MM-PBSA (Table 2.15). The difference in the interaction energy seen between 4-*O*- and 4-*S*-linked hexyl Neu5Ac2en may be due to slightly different binding especially in the positioning of the C-5 acetamido group, as discussed above in Section 2.3.8.3. The *S*-hexyl derivative may be able to bind in a more favourable conformation and therefore have more favourable interaction energy than the analogous *O*-linked derivative.

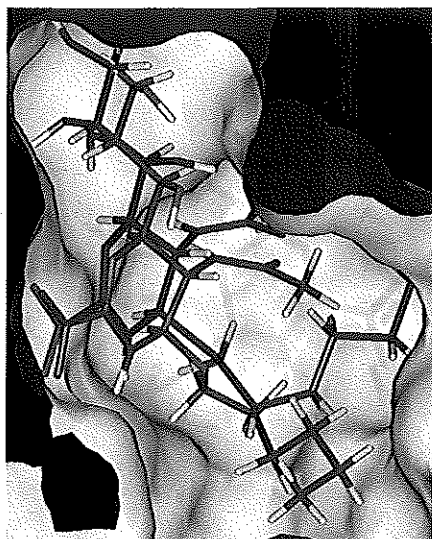


Figure 2.29: Superimposition of the lowest energy structure of 4-*O*-hexyl-Neu5Ac2en (green) and 4-*S*-hexyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.

Table 2.15: Comparative data for 4-*S*- and 4-*O*-hexyl-Neu5Ac2en.^a

Compound		MM-PBSA kcal/mol	AutoDock kcal/mol	X-Score pK _a
5	Neu5Ac2en	-102.50	-6.36	5.12
28-D	4- <i>S</i> -hexyl-Neu5Ac2en	-94.29 [2] ^a	-13.29 [1]	5.81 [1]
29-D	4- <i>O</i> -hexyl-Neu5Ac2en	-106.07 [1]	-11.93 [2]	5.54 [2]

^aNumbers in square brackets indicate relative ranking of computational score.

2.3.8.5 Comparison of 4-*S*- and 4-*O*-decyl-Neu5Ac2en computational structures.

The superimposition of the lowest energy complexes of 4-*O*-decyl-Neu5Ac2en and 4-*S*-decyl-Neu5Ac2en can be seen in Figure 2.30. *In silico*, the *S*-linked derivative is predicted to bind most favourably by consensus scoring as seen in Table 2.16. In the *O*-linked structure the decyl side chain is oriented over the main ring toward the glycerol side-chain binding pocket. The *S*-linked structure orients the decyl side-chain out of the active site, where it may pick up additional interactions, external to the active site, with the protein. The C-5 acetamido group is also noted to be oriented differently between the two derivatives as seen with other *O*-alkylated derivatives. These differences may account for the lower energy of interaction of the 4-*O*-decyl derivative.

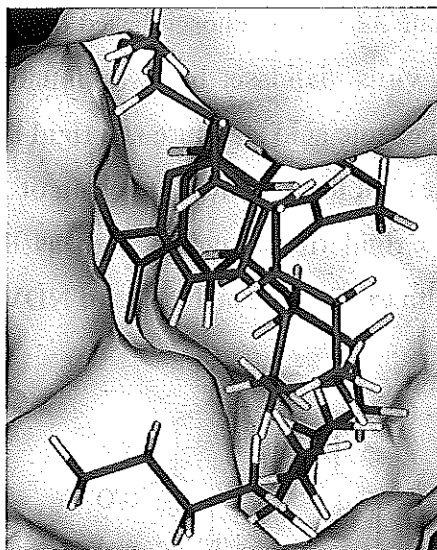


Figure 2.30: Superimposition of the lowest energy structure of 4-*O*-decyl-Neu5Ac2en (green) and 4-*S*-decyl-Neu5Ac2en (purple) in the minimised active site of NDV HN.

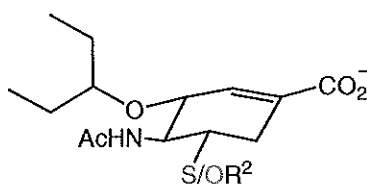
Table 2.16: Comparative data for 4-*S*- and 4-*O*-decyl-Neu5Ac2en.^a

Compound	MM-PBSA	AutoDock	X-Score
	kcal/mol	kcal/mol	pK _d
5 Neu5Ac2en	-102.50	-6.36	5.12
28-E 4- <i>S</i> -decyl-Neu5Ac2en	-124.62 [1] ^a	-15.86 [1]	6.10 [1]
29-E 4- <i>O</i> -decyl-Neu5Ac2en	-88.67 [2]	-11.70 [2]	4.97 [2]

^aNumbers in square brackets indicate relative ranking of computational score.

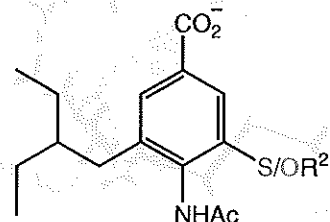
2.3.9 Scoring and docking results of templates other than Neu5Ac2en.

The first alternative template trialled was based on the Oseltamivir™ structure as designed by Gilead Sciences.⁶⁴ This template **30/31** the glycerol side chain of **5** is replaced with a hydrophobic 3-pentyl group and the ring oxygen of **5** is replaced with a methylene group. Using a carbocyclic structure also allowed re-positioning of the endocyclic double bond to more closely match the putative transition state of sialoside hydrolysis.⁶⁴ This template was used to examine whether a hydrophobic side-chain gives better binding than a glycerol side-chain, and also to examine whether the carbocycle is better than the carbohydrate ring structure of Neu5Ac2en **5**.



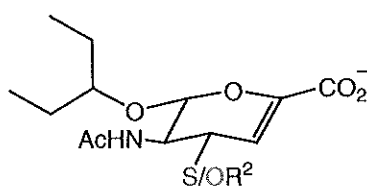
30/31 Gilead Template

The benzoic acid template **32/33** was chosen based on a template that was designed for inhibition of influenza virus sialidase, but was found to be highly specific for influenza A over B.⁷⁸ The benzoic acid structure is a rigid aromatic ring and also has a hydrophobic side-chain instead of the glycerol side-chain of Neu5Ac2en. This template was used to test whether a rigid ring structure such as the benzene ring will give derivatives that bind well with the flexible active site of NDV HN.



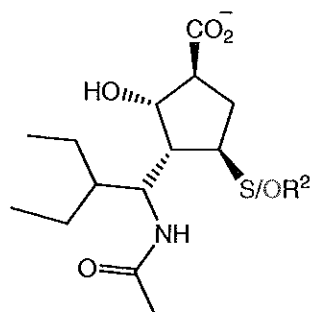
32/33 Benzoic Acid Template

The third template chosen was the uronic acid template **34/35**.¹²⁰ This template replaces the glycerol side chain of Neu5Ac2en with an *O*-linked hydrophobic 3-pentyl group. Therefore, this template was used to examine whether the *O*-linked hydrophobic side-chain is better than a glycerol side-chain.



34/35 Uronic Acid Template

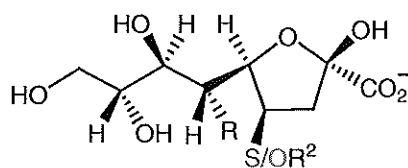
The fourth template chosen was the BCX-1812 template **36/37** as designed by BioCryst.⁷⁹ BCX-1812 was designed for the treatment of influenza. The structure of BCX-1812 is a cyclopentyl ring with a hydrophobic side-chain to mimic OseltamivirTM, and the carboxylic acid, and acetamido functionalisation to mimic the structure of Neu5Ac2en. This template was used to test whether a cyclopentyl ring with functionalisation could create derivatives that bind well with the active site of NDV HN.



36/37 BCX-1812 Template

The fifth of the templates trialled was the furanose template **38/39**.¹²¹ This template was trialled due to its five membered ring and also for its alternate placement of the acetamido group. This template should give further information on whether a five or

six membered ring is optimal for the placement of functionalisation within the active-site of NDV HN.



38/39 Furanose Template

$R^1 = \text{NHAc}$

The substituents A-N (Figure 2.24) were attached by a thioether or ether linkage to each of the five new templates *S*- & *O*-Gilead template **30/31**, *S*- & *O*-benzoic acid template **32/33**, *S*- & *O*- uronic acid template **34/35**, *S*- & *O*- BCX-1812 template **36/37**, and *S*- & *O*-furanose template **38/39**. Each of the compounds was built using InsightII. The compounds were then docked and scored using the second modelling protocol outlined in Section 2.3.7.3. Pertinent data concerning the discussion of these templates will be presented in main text (for comparative discussion of Neu5Ac2en-based derivatives) with all other docking and scoring data to be found in Appendix B.

Scoring data for all templates is given in Table 2.17 with the energy values given as a range for substituents A-N on each template. Computationally, no significant difference could be seen in the energy ranges obtained for each of the templates using each of the scoring programs, as seen in Table 2.17. These results however do not accurately assess the affect of each templates rigidity or flexibility. This is due to each compounds being treated as a static entity for three quarters of the protocol. If the affect of changing templates to have more rigidity or more flexibility is to be more accurately assessed, molecular dynamics with the averaging of energies obtained would have to be employed. Molecular dynamics simulations however, are computationally expensive given that the calculations would have to be performed on 168 compounds. This said, the results shown in Table 2.17 do show at least one important fact, that each of the templates does in fact successfully dock, minimise and score in the active site of NDV HN. This would suggest that the active site of NDV HN is large enough that the effect of each template is minimised. This could be due to the fact that each template could search the active site for a viable docking conformation, while still making the important active site residue interactions. The fact that each of the templates docked and scored, would suggest that any of the five templates is a viable alternative to the Neu5Ac2en template. The predicted activity of each template was at least comparable to the computational activity of Neu5Ac2en. However, with the chemistry of

Neu5Ac2en well established in the literature, it was decided that the Neu5Ac2en template was the most desirable template to pursue.

Table 2.17: Energetic range of computational data (Autodock, MM-PBSA, X-Score) for all template derivatives.^{a,b}

Flexible AutoDock (kcal/mol)		MM-PBSA (kcal/mol)	
Sulfur-Linked	Energy Range	Sulfur-Linked	Energy Range
Neu5Ac2en 28	-12.82 to -17.31	Neu5Ac2en 28	-82.50 to -128.46
Gilead 30	-13.39 to -16.60	Gilead 30	-81.09 to -110.22
Benzoic Acid 32	-12.59 to -16.91	Benzoic Acid 32	-78.19 to -108.47
Uronic Acid 34	-12.98 to -16.53	Uronic Acid 34	-69.61 to -112.87
BCX-1812 36	-10.83 to -16.06	BCX-1812 36	-79.13 to -119.24
Furanose 38	-12.72 to -16.80	Furanose 38	-100.27 to -137.93
Oxygen-Linked	Energy Range	Oxygen-Linked	Energy Range
Neu5Ac2en 29	-13.02 to -18.12	Neu5Ac2en 29	-18.67 to -118.24
Gilead 31	-11.44 to -17.08	Gilead 31	-80.00 to -112.28
Benzoic Acid 33	-12.18 to -16.71	Benzoic Acid 33	-73.52 to -104.87
Uronic Acid 35	-12.56 to -16.56	Uronic Acid 35	-73.27 to -107.19
BCX-1812 37	-10.47 to -15.53	BCX-1812 37	-88.66 to -190.38
Furanose 39	-10.65 to -15.68	Furanose 39	-90.54 to -137.81
Static AutoDock (kcal/mol)		X-Score (pK _d)	
Sulfur-Linked	Energy Range	Sulfur-Linked	Energy Range
Neu5Ac2en 28	-11.39 to -15.86	Neu5Ac2en 28	5.32 to 6.69
Gilead 30	-12.53 to -15.46	Gilead 30	5.75 to 7.26
Benzoic Acid 32	-11.73 to -15.64	Benzoic Acid 32	6.13 to 7.74
Uronic Acid 34	-12.48 to -16.40	Uronic Acid 34	5.79 to 7.54
BCX-1812 36	-10.28 to -15.78	BCX-1812 36	5.36 to 6.62
Furanose 38	-9.65 to -16.06	Furanose 38	4.88 to 6.68
Oxygen-Linked	Energy Range	Oxygen-Linked	Energy Range
Neu5Ac2en 29	-5.92 to -13.26	Neu5Ac2en 29	4.97 to 6.26
Gilead 31	-11.52 to -15.93	Gilead 31	5.74 to 7.22
Benzoic Acid 33	-11.63 to -15.94	Benzoic Acid 33	5.91 to 7.66
Uronic Acid 35	-12.04 to -16.11	Uronic Acid 35	5.61 to 7.38
BCX-1812 37	-9.65 to -14.62	BCX-1812 37	5.43 to 7.00
Furanose 39	-9.88 to -15.36	Furanose 39	5.12 to 6.91

^a Each template energy range is presented as the energy range between the lowest to highest docked energy / binding score for all derivatives A-N of that particular template.

^b Full data for each derivative can be seen in Appendix B.

2.3.10 Docking and scoring of miscellaneous compounds

After the establishment of the designed compounds, it was of interest to examine whether any new templates could be discovered. It was also of interest to examine other Neu5Ac2en derivatives with additional modifications to study whether they pick up additional favourable contacts with the active site of NDV HN. The scoring results in

Sections 2.3.10.1 and 2.3.10.2 have been calculated using the initial protocol presented in Section 2.3.4.1

2.3.10.1 Modified C-4-O-Linked substituted Neu5Ac2en derivatives.

Both the 4-*O*-benzyl **29-A** and 4-*O*-(2-phenyl)ethyl **29-G** Neu5Ac2en derivatives were substituted (as suggested from the GRID study, Section 2.3.2.6) at either C-9 (*O*-benzyl) or C-8 (amino group) as shown in Figure 2.31. Flexible AutoDock results for each of the derivatives in Figure 2.31 in the unminimised active site if NDV HN are shown in Table 2.18 (**29-O**, **P** and **Q**) and Table 2.19 (**29-R**, **S** and **T**) respectively.

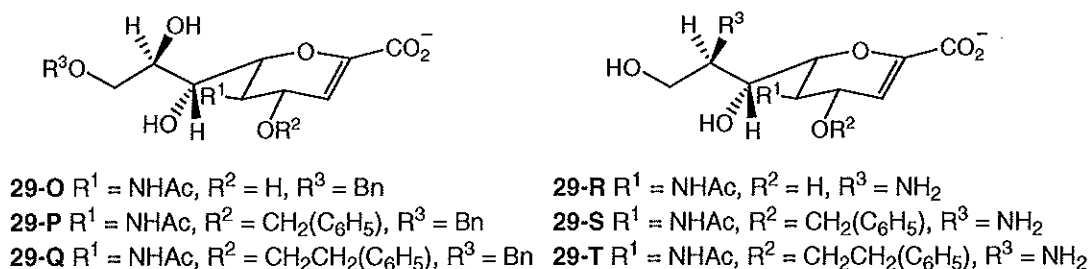
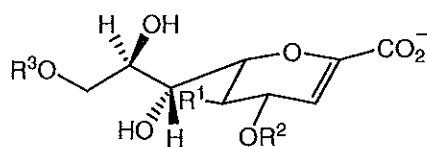


Figure 2.31: Structures of C-8 or C-9-modified C-4-ether linked Neu5Ac2en derivatives.

Table 2.18 shows the docking and scoring results for the C-9 modification to Neu5Ac2en and to C-4-ether linked derivatives **29-A** and **29-G**. The 9-*O*-benzyl derivative of Neu5Ac2en **29-O** scored 5.95 kcal/mol greater than Neu5Ac2en, this is slightly less than the 4-*O*-benzyl derivative **29-A** at -17.10 kcal/mol (7.04 kcal/mol increase from Neu5Ac2en). The 4,9-di-*O*-Bn derivative **29-P** scored 4.39 kcal/mol greater than 4-*O*-benzyl-Neu5Ac2en derivative **29-A**, the total increase over Neu5Ac2en is 11.43 kcal/mol. The 4-*O*-benzyl-9-*O*-(2-phenyl)ethyl derivative **29-Q** scored 8.61 kcal/mol greater than 4-*O*-(2-phenyl)ethyl-Neu5Ac2en derivative **29-G**, and 10.33 kcal/mol greater than Neu5Ac2en.

Table 2.18: Docking and scoring results for C-4 modified 9-*O*-benzyl-Neu5Ac2en derivatives, **29-O**, **P**, and **Q**.



29-O, **P**, and **Q** $R^1 = \text{NHAc}$

	R^2	R^3	AutoDock ^a (Unminimised) kcal/mol
29-O	H	benzyl	-16.01 [3]
29-P	benzyl	benzyl	-21.49 [1]
29-Q	(2-phenyl)ethyl	benzyl	-20.39 [2]

^a Without C-9 modification for reference 4-*O*-benzyl derivative **29-A** scored at -17.10 kcal/mol and 4-*O*-(2-phenyl)ethyl derivative **29-G** scored at -11.78 kcal/mol (AutoDock score). Neu5Ac2en **5** scored at -10.06 kcal/mol.

The docked orientation of 4,9-di-*O*-benzyl-Neu5Ac2en **29-P** can be seen in Figure 2.32. The carboxylate group is seen to make the expected interactions with the triarginyl cluster. Also seen in Figure 2.32 is the expected hydrophobic interaction of the C-5 acetamido methyl group with Tyr 299. The C-9 benzyl modification was designed to make interactions with Phe 364 located inside and at the end of the glycerol side-chain binding pocket. It is however observed forming an interaction with the outside the glycerol side-chain binding pocket.

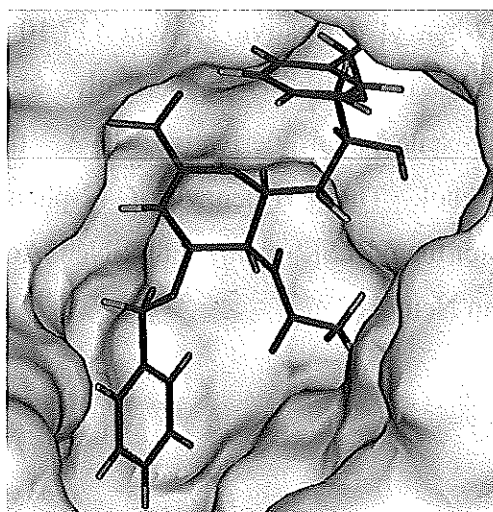
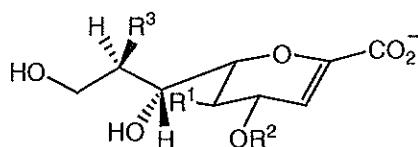


Figure 2.32: AutoDock orientation of 4,9-di-*O*-benzyl-Neu5Ac2en **29-P**.

Table 2.19 shows the docking and scoring results of the C-8 modification to Neu5Ac2en and to C-4-ether linked derivatives **29-A** and **29-G**. It can be seen that the

energy of interaction of the C-8 amino-8-deoxy derivative **29-R** was not significantly different to that of Neu5Ac2en (the energy is 2.78 kcal/mol greater than Neu5Ac2en). The increase in the energy of interaction seen in the AutoDock score ranged between 0.94 kcal/mol with the 8-amino-4-*O*-benzyl-8-deoxy derivative **29-A** (total increase over Neu5Ac2en of 7.98 kcal/mol) and 6.75 kcal/mol with the 8-amino-8-deoxy-4-*O*-(2-phenyl)ethyl derivative **29-T** (total increase over Neu5Ac2en of 8.47 kcal/mol).

Table 2.19: Docking and scoring results for C-4 modified 8-amino-8-deoxy-Neu5Ac2en derivatives, **29-R**, **S**, and **T**.



29-R $R^1 = \text{NHAc}$, $R^2 = \text{H}$, $R^3 = \text{NH}_2$

29-S $R^1 = \text{NHAc}$, $R^2 = \text{CH}_2(\text{C}_6\text{H}_5)$, $R^3 = \text{NH}_2$

29-T $R^1 = \text{NHAc}$, $R^2 = \text{CH}_2\text{CH}_2(\text{C}_6\text{H}_5)$, $R^3 = \text{NH}_2$

	R^2	R^3	AutoDock ^a (Unminimised) kcal/mol
29-R	H	NH ₂	-12.84 [3]
29-S	benzyl	NH ₂	-18.04 [2]
29-T	(2-phenyl)ethyl	NH ₂	-18.53 [1]

^a Without C-9 modification for reference 4-*O*-benzyl derivative **29-A** scored at -17.10 kcal/mol and 4-*O*-(2-phenyl)ethyl derivative **29-G** scored at -11.78 kcal/mol (AutoDock score). Neu5Ac2en **5** scored at -10.06 kcal/mol.

Figure 2.33 shows the AutoDock orientation of 8-amino-4-*O*-benzyl-8-deoxy-Neu5Ac2en, **29-S**. The interactions that the derivative was seen to make are similar to those of 4-*O*-benzyl-Neu5Ac2en **29-A**, with the carboxylate group making a strong interaction with the triarginyl cluster and the benzyl ring and the C-5 acetamido group making hydrophobic interactions in the C-4 and C-5 binding pockets respectively. The 8-amino functionality was seen to pick-up analogous interactions to those by the C-8 hydroxyl group of Neu5Ac2en. The lack of significantly improved energies and interactions coupled with the extra synthetic steps required to synthesise the C-8 amino derivatives, makes this modification of low priority for synthesis.

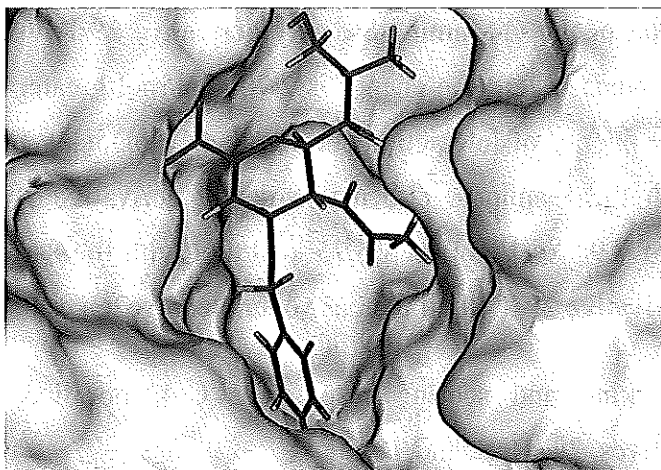
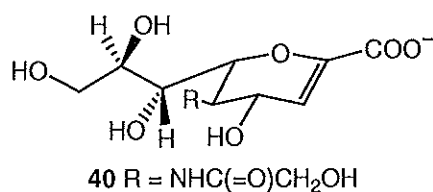


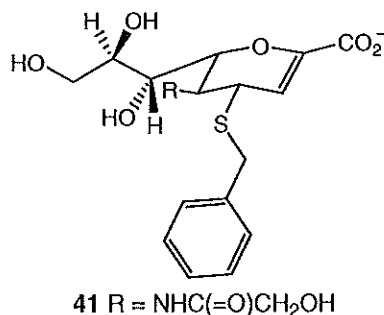
Figure 2.33: AutoDock orientation of 8-amino-4-*O*-benzyl-8-deoxy-Neu5Ac2en **29-S**.

2.3.10.2 *Neu5Gc2en* and 5-glycolylamino derivative of 4-*S*-benzyl *Neu5Ac2en*

Neu5Gc2en was selected to examine if a simple modification of Neu5Ac2en at the C-5 position would confer an increase in the interactions of derivatives. The initial modelling protocol (Section 2.3.4.1) was used to characterise the interactions of Neu5Gc2en. Flexible AutoDock of Neu5Gc2en, **40**, in NDV HN yielded an energy of interaction of -13.36 kcal/mol; this is 3.30 kcal/mol better than Neu5Ac2en (-10.06 kcal/mol). The structure was minimised and re-scored. The static AutoDock score for **40** was -2.87 kcal/mol (7.19 worse than Neu5Ac2en). Examination of the minimised conformation showed that the triarginyll cluster interactions had been diminished by the minimisation process and this is reflected in the static AutoDock score. The SCORE result for **40** was 6.62 (0.56 better or 3.6 times better than Neu5Ac2en) showing that the modification of the acetamido group of Neu5Ac2en to a glycolyl group ($\text{NHC}(=\text{O})\text{CH}_3$ to $\text{NHC}(=\text{O})\text{CH}_2\text{OH}$) should increase binding affinity. This is also shown by the Discover MMI binding energy of -193.64 kcal/mol, or 20.12 kcal/mol better than Neu5Ac2en. Ligplot analysis showed a change in glycerol side-chain interactions between Neu5Ac2en and Neu5Gc2en **40** and an additional hydrogen bond by the C-5 glycolylamino group.



The results of the scoring of Neu5Gc2en suggested that the addition of the 5-NHglycolyl group to C-4 modified Neu5Ac2en derivatives may cause an increase in favourable interactions. Therefore, the 5-glycolylamino derivative of 4-*S*-benzyl Neu5Ac2en **41**, was docked into NDV HN using the initial flexible AutoDock (unminimised, Section 2.3.4.1) protocol to examine whether an increase in docked energy would occur.



Flexible AutoDock of 4-*S*-benzyl-Neu5Ac2en, **28-A**, in NDV HN yielded an energy of interaction of -16.69 kcal/mol; this is 6.63 kcal/mol better than Neu5Ac2en (-10.06 kcal/mol). The flexible AutoDock score for 4-*S*-benzyl-Neu5Gc2en derivative, **41**, yielded an energy of interaction of -16.66 kcal/mol; this was 6.60 kcal/mol better than Neu5Ac2en. The difference in the interaction energies of **28-A** and **41** was 0.03 kcal/mol. It can be seen from the AutoDock scores obtained for both molecules that the substitution of a glycolylamino group for the C-5 acetamido group had no impact on the energy of interaction with NDV HN. Given the extra steps required to synthesise the glycolylamino derivative, this was not considered a target for this work.

2.3.11 The development of potential new templates

Two new templates shown in Figure 2.34 were designed in which functionality was minimised. The first template **42** included only the carboxylate group with an *S*-linked phenyl group positioned at a distance mimicking a C-4 substituent on a sialic acid.

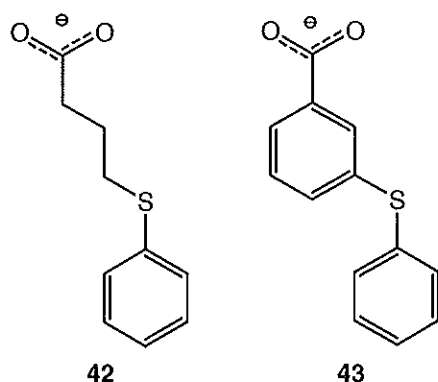


Figure 2.34: Structures of potential new alkyl and benzoic acid templates.

The second template **43** was based on a benzoic acid core structure. This structure was used to try to reduce the flexibility of the first template **42**. The more torsions that freeze upon binding can offset the free energy generated upon binding. This template also had the removal of the glycerol and acetamido side-chains.

Templates **42** and **43** were built using InsightII and the interaction energies of templates **42** and **43** were calculated using the initial AutoDock protocol from Section 2.3.4.1. The energies of the docked ten calculated poses obtained for the alkyl template **42** ranged from -10.39 kcal/mol to -8.82 kcal/mol, 0.33 kcal/mol better to -1.24 kcal/mol worse than Neu5Ac2en (-10.06 kcal/mol). Two distinct docked poses were generated upon binding, with the top scored structure bound with the benzyl group making interactions in the glycerol side-chain binding-pocket. Figure 2.35 shows the pose in which the benzyl group is bound in the C-4 pocket. This pose was the third ranked structure with a binding energy of -9.76 kcal/mol, only 0.3 kcal/mol worse than Neu5Ac2en. This template shows that it is possible to bind with a similar energy of interaction with a much smaller and significantly simplified compound than Neu5Ac2en.

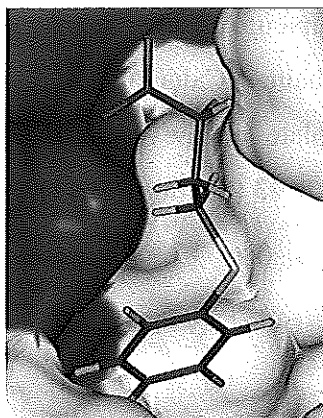


Figure 2.35: AutoDock result for the alkyl template **42**.

The 10 docked poses for the benzoic acid template **43** in NDV HN ranged from -11.67 kcal/mol to -10.26 kcal/mol (1.61 kcal/mol to 0.2 kcal/mol better than Neu5Ac2en). Two conformations were generated upon binding. Once again, the top scored structure bound with the benzyl group in the glycerol side-chain binding-pocket. Figure 2.36 shows the pose with the benzyl group bound in the C-4 pocket; this conformation was the fourth ranked structure with a binding energy of -10.51 kcal/mol, 0.45 kcal/mol better than Neu5Ac2en. It can be seen from the results between the alkyl **42** and benzoic acid **43** templates, that the rigidity of the latter template **43** is preferable to the flexible template **42**.

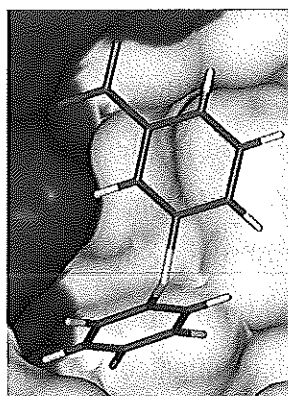


Figure 2.36: AutoDock result for benzoic acid template **43**.

2.3.11.1 *LigBuilder design of alternative templates.*

In a further search for alternative template structures, LigBuilder⁹⁷ was used to develop structures using both the Grow and Link strategies. The seven seed structures used, the type of calculation and the points of modification on the seed structures are shown in Figure 2.37. Three types of seed structures were chosen. For calculations 1, 4, and 5, an *S*-alkylated thiophenyl fragment was used which in essence represented the atoms off C-4 with an attached thiophenyl group. Calculation 2 used the Link program

with the key substituents of the Neu5Ac2en pyranose ring: the carboxylate, acetamido group and the glycerol side-chain. Calculations 3, 6, and 7 used a 2-thiophenyl substituted benzoic acids template as the seed, this was based on the previously investigated benzoic acid template **43** (Section 2.3.11).

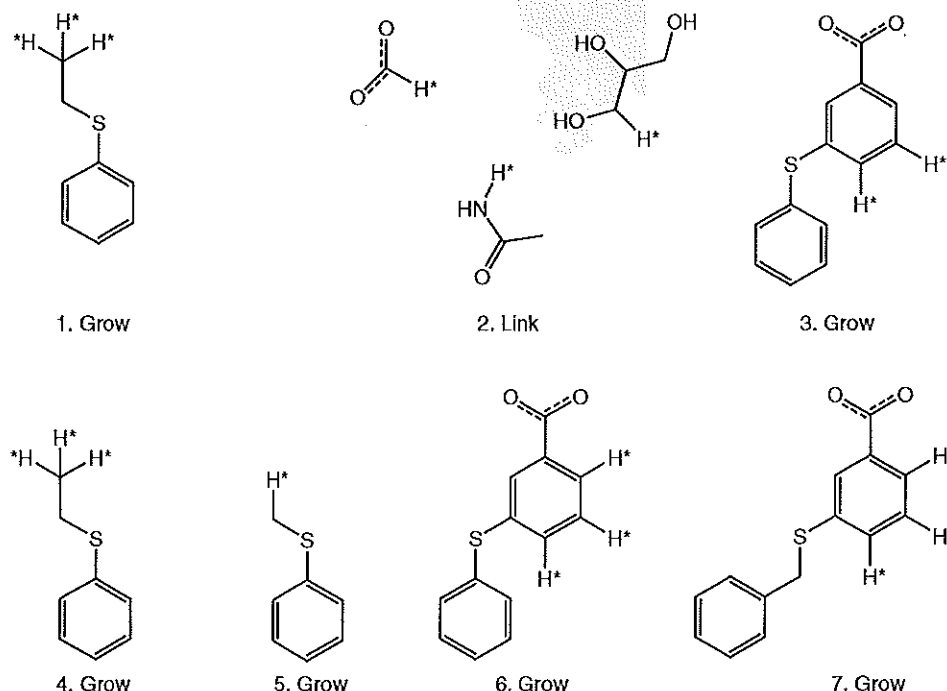


Figure 2.37: The 7 seed structures that were used in LigBuilder.⁹⁷ H* is the point of modification.

The first three calculations used the default fragment library available with LigBuilder. Most structures developed for the first calculation were highly conjugated aromatic systems. The second calculation, involving a linking procedure, did not connect all available fragments of the seed structures. The third calculation was performed to investigate possible improvements to the previously docked benzoic acid template **43** (Section 2.3.11). Both conjugated aromatic rings and hydrogen-bond donor groups were placed at the C-5 position.

Calculations 4-7 use a modified fragment library, library includes β -D-glucose, β -D-galactose, β -D-fructose, glycerol, Neu5Ac2en, Neu5Ac, and benzoic acid, Therefore, the previous calculations involving seed structures 1 to 3 were repeated to investigate whether selection of the new fragments in the calculations would occur, as either new core structures or modifications. It was found that none of the added carbohydrate fragments available for incorporation into the growing template were used, which may be due to the fact that carbohydrate fragments adversely affect the

calculation of Lipinski's "rule of 5" parameters taken into account by LigBuilder. The seventh calculation used a modified benzoic acid template in which the substituent phenyl ring was extended to a benzyl group. With the modified fragment library and the guanidinium and amidinium fragments enabled, it was found that the guanidinium fragment was placed on the position of the benzoic acid (C-5) analogous to the C-6 position of Neu5Ac2en. This finding led to the next template to be analysed using AutoDock. The structure of which can be seen in Figure 2.38.

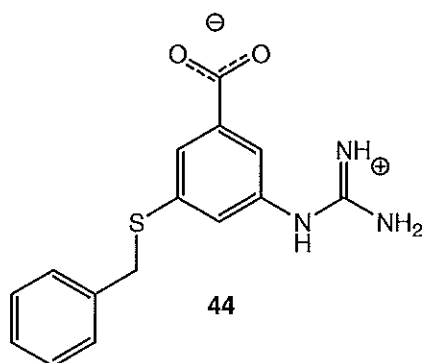


Figure 2.38: The structure of 3-thiobenzyl-5-guanidino-benzoic acid **44**.

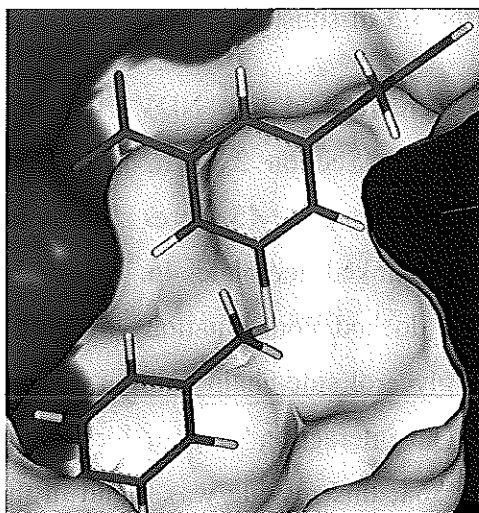


Figure 2.39: Docked conformation of the 3-thiobenzyl-5-guanidino-benzoic acid template **44**.

The docked poses for the 3-thiobenzyl-5-guanidino-benzoic acid template **44** ranged from -13.75 kcal/mol to -12.32 kcal/mol (3.69 kcal/mol to 2.26 kcal/mol better than Neu5Ac2en). As seen before, two alternative poses were generated upon binding, of which the top structure had the benzyl group bound in the glycerol side-chain binding pocket. This pose could be caused by the presence of Phe 364 in the glycerol side-chain pocket. Figure 2.39 shows the pose of **44** bound with the benzyl ring in the C-4 pocket.

This conformation was the second ranked structure with a binding energy of -13.16 kcal/mol, 3.1 kcal/mol better than Neu5Ac2en.

2.3.11.2 Di-substituted benzoic acid template

Following the observation of the second binding conformation of the alkyl **42** and benzoic acid templates **43**; in which the phenyl substituent is positioned in the glycerol side chain binding pocket (Section 2.3.11), another template was created. The template structure developed was a di-substituted benzoic acid in which the substituent benzyl groups at C-3 and C-5 could be either *S* or *O* linked.

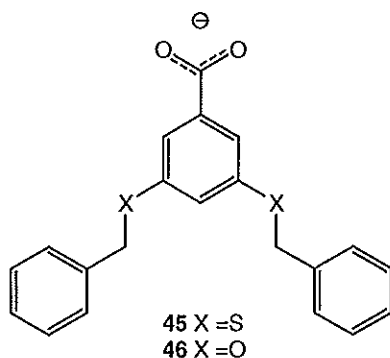


Figure 2.40: Structure of the 3,5-di-substituted benzoic acid template.

Structures **45** and **46** were docked into NDV HN and scored using the initial modelling protocol discussed in Section 2.3.4.1. The initial flexible unminimised AutoDock score for the *S*-linked template **45** scores slightly higher (-16.82 kcal/mol vs -16.30 kcal/mol) than the *O*-linked template **46**. Both templates are better than Neu5Ac2en (-10.06). The docked conformations of both derivatives showed one benzyl group bound in the glycerol side-chain binding pocket with the other bound in the C-5 binding pocket. The interaction of note formed in the glycerol side-chain binding pocket is with Phe 364; this residue was found in the previous GRID⁹³ study (Section 2.3.2.6) to give a hydrophobic character to the tail of the glycerol side-chain binding pocket. The interaction in the C-5 binding pocket is with Tyr 317, which is found to make interactions with the C-5 acetamido group of Neu5Ac2en. After minimisation, the interaction energies of the *S*-linked template **45**, was slightly higher (-174.40 kcal/mol) than Neu5Ac2en (-173.52 kcal/mol), whereas the *O*-linked template **46** was slightly lower (-172.39 kcal/mol) than Neu5Ac2en. In both the static AutoDock and SCORE scores the *S*-linked template **45** performs better (-16.74 kcal/mol and 8.04 respectively) than both Neu5Ac2en (-12.42 kcal/mol and 6.06 respectively). The *O*-linked template **46** (-15.93 kcal/mol and 5.42 respectively) scores better than

Neu5Ac2en in the static AutoDock scoring calculation and scores poorer in the SCORE calculation. In all scores the *S*-linked template scores better than the *O*-linked template. This may be due to the longer C-S bond lengths in template **45** which may allow slightly more flexibility to attain an optimal orientation within the active site of NDV HN.

Table 2.20: Scoring results of the di-substituted benzoic acid templates **45** and **46** with NDV HN.

3,5-X-substituent	Flexible AutoDock (kcal/mol)	MMI (kcal/mol)	Static AutoDock (kcal/mol)	SCORE (pK _d)
<i>S</i> 45	-16.82	-174.40	-16.74	8.04
<i>O</i> 46	-16.30	-172.39	-15.93	5.42

The main areas of interaction of the 3,5-di-thiobenzyl benzoic acid template **45** are indicated in Figure 2.41. In the AutoDock results, the dark blue area indicates the triarginyl cluster, the purple area indicates Tyr 317, and the dark red area indicates Phe 364 in the glycerol side-chain binding pocket. Also shown in Figure 2.41 are the SCORE results of the docked conformation of **45**. It can be seen that one area of bad interaction (red) is found on the 5-thiobenzyl CH₂ carbon atom, with most of the C-3 and C-5 benzyl rings forming good and neutral interactions with the protein.

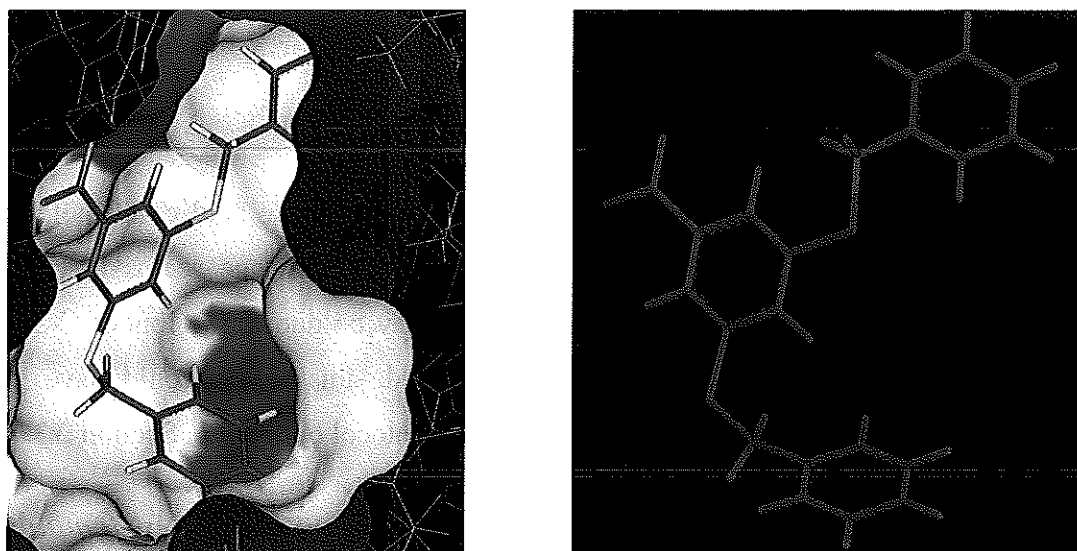


Figure 2.41: AutoDock (left) and SCORE (right) results for 3,5-di-thiobenzyl benzoic acid template **45**.

The main areas of interaction of the 3,5-di-benzyloxy benzoic acid template **46** are indicated in Figure 2.42. In the AutoDock results, the dark blue area indicates the triarginyl cluster, the purple area indicates Tyr 317 in the C-5 binding pocket, and the dark red area indicates Phe 364 in the glycerol side-chain binding pocket. Also shown in Figure 2.42 are the SCORE results of the docked conformation, it can be seen that two areas of bad interactions are found on the molecule. These interactions may be an artefact due to the docked orientation of the molecule; it can be seen in the docked structure that the benzoic acid ring docked close to vertical in the active site of NDV HN. This orientation will increase the likelihood of a “bump” (too close to the protein surface) with the protein.

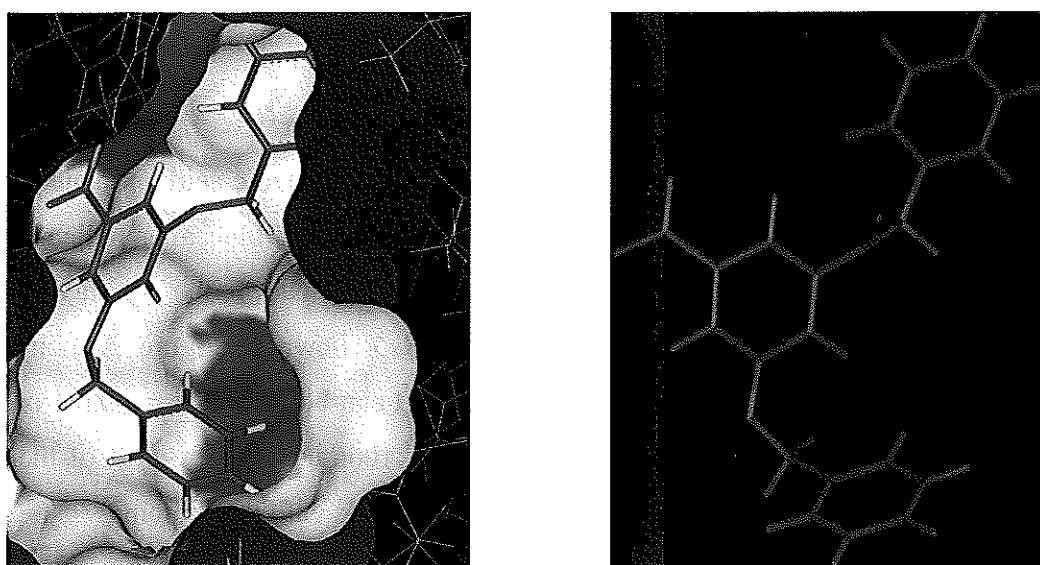


Figure 2.42: AutoDock (left) and SCORE (right) results for 3,5-di-benzyloxy benzoic acid template **46**.

This study demonstrates that a simply derivitised benzoic acid could potentially interact with NDV HN better than Neu5Ac2en. The *S*-linkage seems to be more favourable in these benzoic acid derivatives, this may be due to the longer C-S bond length. However, for firm conclusions to be made this study would have to be expanded and revised to use the refined modelling protocol outlined in Section 2.3.7.3.

2.3.11.3 *Final thoughts on the development of potential new templates*

However good the templates developed may score against NDV HN, it has been noted in the literature that modification of the glycerol side-chain interactions made by Neu5Ac2en leads to a ligand that does not bind as well to hPIV HN. This can be seen

in the work of Suzuki *et al.*,⁸³ where C-4 modified Neu5Ac2en derivatives and their corresponding C-4 and C-9 modified derivatives were evaluated as inhibitors of hPIV-1. The dual C-4/C-9 derivatives are found to be much less effective in inhibiting the sialidase activity of hPIV-1 HN than their singly C-4 modified counterparts.⁸³ This study would suggest that the modification of the C-9 position would result in a ligand that would have less affinity for hPIV HN.

2.4 Homology Modelling

In the present study towards the computational development of inhibitors of hPIV HN no crystal structure of hPIV HN was available. This necessitated the use of the crystal structure of the structurally and functionally related NDV HN for the development of inhibitors. It was of interest however, to see whether a homology model of hPIV HN could be constructed.

The first step in the homology modelling process was to generate the sequence alignments of NDV HN and the HN's from hPIV-1, 2 and 3. The alignments were generated by Clustal W (Internet based program),¹²² with the sequences obtained from the SWISSPROT database [accession codes: Q9Q2W5 – NDV HN crystal structure sequence; P16071 (HEMA_PI1HW) – hPIV-1 HN; P25465 (HEMA_PI2H) – hPIV-2 HN; P12561 (HEMA_PI3HA) – hPIV-3 HN]. Alignments were generated in two formats: alignment with numbers for sequence analysis (Figure 2.43), and alignment in a format suitable for the program Modeller. Analysis of key active site residues⁴⁸ from NDV HN showed that these residues were conserved across the alignments of hPIV-1, 2, and 3 HN's, as indicated by boxed residues (NDV HN numbering) in Figure 2.43. If additional structures of HN were present during the initial alignment phase, the resulting model quality would be improved, as the quality of the homology model obtained is directly related to the quality of the alignment. However, only the crystal structure of NDV HN (1E8T, 1E8U and 1E8V – with the same sequence) was available at the time of homology model creation, the models obtained are biased by the use of the crystal structure of NDV HN.

```

HEMA_PI1HW    MAEKGKTNSSYWSTTRNDNSTVNTYIDTPAGKTHIWLITATMHTILSFIMILCIDLII 60
HEMA_PI3HA    MEYWKHTNHGKDACNELGTSMATHGNKITNKITYILWTEILVLLSIIFILVLINSIKSEK 60
Q9Q2W5        -----MDRAVSQVALENDEREAKNTWRLIFRIALLLLTVVTLATSVASLVYSM 48
HEMA_PI2H     -----MEDYSHLSLKS---IPKRTCRIIFRTATILGICTLIVLCSSILHEII 44
               :         :         :         :         :
HEMA_PI1HW    KQDTCMAKNIMTVSSMNESAKTIKETITELIRQEVISRTINIQSSVQSGIPILLNKQSRD 120
HEMA_PI3HA    AHESLLQDVNNEFMEVTEKIQMASDNINDLIQSGVNTRLTLIQSHVQNYIPIISLTQQMSD 120
Q9Q2W5        GAS-----TPSDLVGIPTRISRAEEKITSALGSNQDVVDRIYKQVALESPLALLNTETI 103
HEMA_PI2H     HLD-----VSSGLMDSDDSQGGIIPILSLKSLIALANQILYNVAIIIPLKIDSJETVI 99
               :         :         :         :         :
HEMA_PI1HW    LTQLIEKSCNRQELAQICENTIAIHADGISPLDPHDFWRCPVGEPLLSNPNISLLPGP 180
HEMA_PI3HA    LRKFISEITIRNDNQEVPPQ--RITHDVGIKPLNPDDEFWRCTSGLEPSLMKTPKIRLMPGP 178
Q9Q2W5        MNAITSLSYQINGAANNHSGWGAPIHDPDFIGGIGKELIVDNASDVTSFYPS---AFQEH 160
HEMA_PI2H     YSALKDMHTGSMSTNCTPGNLLHDAAYINGENKFLVLKSYNGTPKYGP-----LLNIP 154
               :         :         :         :         :
               Arg 174
HEMA_PI1HW    SLLSGSTTISGCVRLPSLSIGDAIYAYSSNLIQGCADIGKSYQVLQLGYISLNSDMYFD 240
HEMA_PI3HA    GLLAMPTTVDGCVRTPSLVINDLIYAYTSNLIIRGCDIGKSYQVLQIGLITVNSDLVFD 238
Q9Q2W5        NFIPAPITGSGCTRIIPSFDMATHYCYTHNVILSGCRDHSQYLALGVLRITATGRIF 220
HEMA_PI2H     SFIPSATSPNGCTRIIPSFSLIKTHWCYTHNVILGDCLEDTTSNQYLAMGLIQSSAAAFPI 214
               :         :         :         :         :
               Glu 258
HEMA_PI1HW    LKPVISHTYDINDNRKSCSVIAAGTRGYQLCSLPTVNETTDSYSEGIEDLVFDILDLEK 300
HEMA_PI3HA    LNPRIHTFNINDNRKSCSLALLNTDVYQLCSTPKVDESDYASSGIEDIVLDIVNHDGS 298
Q9Q2W5        ESTLRSISLDDTQNRKSCSVSATPLGCDMLCSKVTETEEEDYNSAVPTLMAHGRIGFDGQ 280
HEMA_PI2H     FRTMKTIIYLSGGINRKSCSVTAIPGGCVLYCYVATRSKEDYATTDLAEELRLAFYYNDT 274
               :         :         :         :         :
               Tyr 317
HEMA_PI1HW    TKSHRYKNEDITFDHPFSAMYPSVSGSINKIENTLIFLGYGGLTTPLOGDTKCV----- 353
HEMA_PI3HA    ISTTREFKNNNISFDQPYAALYPSVGPPIYKGIIFLGYGGLEHPINENAICN----- 351
Q9Q2W5        YHEKDLDT--TFEEDWVANYPGVGGGSFIDGRVWFSVYGGLEKNSPSDTVQEGYVIYK 338
HEMA_PI2H     FIERVISLP--NTTGQWATINPAVGSYHGLFIFLPVYGGLEKGTSPYKQSSRYFIPK 332
               :         :         :         :         :
HEMA_PI1HW    --TNRCAN---VNQSVCDALKITWLKKRQVVNVLIRINNYLSDRPKIVVETIPITQNYL 408
HEMA_PI3HA    --TTCGPG---KTQRDCNQASHSPWFSDRRMVNSIIVVDKGLNSIPKLVVWTISMRQNYW 406
Q9Q2W5        RYNDTCPDEQDYQIRMAKSSYPKRGFGKRIQQAILSIKSVSTLGEDEPVLTVPPNTVTL 398
HEMA_PI2H     HPNITCAGKSSEQAAAAARSSYVIRYHSNRLQLSAVLICPLSDMHTARCNLVMMNNSQVMM 392
               :         :         :         :         :
               Arg 416
HEMA_PI1HW    GAEGRLKLGKKIYIYTESSGWHSHLQIGSLDINN-----PMTIKWAPHEVLSRPGNQD 462
HEMA_PI3HA    GSEGRLLLLGNKIYIYTESSTWHSKLGIGIITDYS---DIRIKWTWHNVLSRPGNNE 462
Q9Q2W5        GAEGRLITVGTSHFLYQSGSSYFSPALLYPMTVSNKT---ATLHSPYTFN-AFTREGSIP 454
HEMA_PI2H     GAEGRLYVIDNNLYYQSSSSWWSASLFYRINTDFSKGIPPIEBAQWVPSYQVPRPGVMP 452
               :         :         :         :         :
               Arg 498
HEMA_PI1HW    CNWYNRCPRECISGVYTDAYPLSPDAVNVAT-----TTLYANTSIVNPVIMYSNT 512
HEMA_PI3HA    CPWGHSCPDGCI TGVYTDAYPLNPTGSIVSS-----VILDSQKIRVNPVITYSTA 512
Q9Q2W5        CQASARCPNSCVTVGYTDPYPLIFYRNHTLRGVFG-----TMLDSEQARLNPAFAVFD 508
HEMA_PI2H     CNATSFPCANCITGVYADVWPLNDPEPTSQNALNPNYRFAGAFERNESNRTNPTFTYASA 512
               :         :         :         :         :
               Tyr 526
HEMA_PI1HW    SEIINMERLKNVQLEAAITTTSCITHFG--KGYCFHIVEINQTSINTLQPMLEKTSIPKI 570
HEMA_PI3HA    TERNELAIRNKTLSAGYITTTSCITHYN--KGYCFHIVEINHKSLDTEQPMLEKTEIPKS 570
Q9Q2W5        TSRSRITRVSSSSTKAAIYITSTCFKVVKNKTYCLSAEISNTLFGFEIRIVPLLEILKN 568
HEMA_PI2H     SALLNTTGFNNINHKAAIYITSSCTCFKNIGTQKTYCLIIEMGSSLLGEFQIIPFLRELIP- 571
               :         :         :         :         :
HEMA_PI1HW    CKITS---- 575
HEMA_PI3HA    CS----- 572
Q9Q2W5        DGVREARSG 577
HEMA_PI2H     -----
               :         :

```

Figure 2.43: Multiple HN Alignments,

Codes: (HEMA_PI1HW – hPIV-1 HN, HEMA_PI3HA – hPIV-3 HN, Q9Q2W5 – NDV HN, HEMA_PI2H – hPIV-2 HN.)

2.4.1.1 Preliminary analysis of homology models

Each hPIV-HN was homology modelled using the program Modeller. Each model was modelled using the crystal structure of NDV HN without flipped His, Gln, Asn residues. The tertiary structure of each model is shown in Figure 2.44. No errors

occurred during modelling of each structure, and Procheck⁸⁹ analysis, (Section 2.1) revealed that the models were comparable to the NDV HN crystal structure. Analysis using Whatcheck⁹⁰ (Section 2.1) of each model structure identified His, Gln, Asn residues for flipping, while analysis of the hydrogen bonding network showed many residues in the core of protein not satisfied. The hPIV-1 homology model was minimised in an aqueous environment in an attempt to fix the hydrogen-bonding network. Procheck analysis of the minimised structure showed an overall reduction of model quality. However, the same result is noted for the minimised structure of NDV HN. Whatcheck analysis shows that minimisation had little effect on improving the hydrogen bond network.

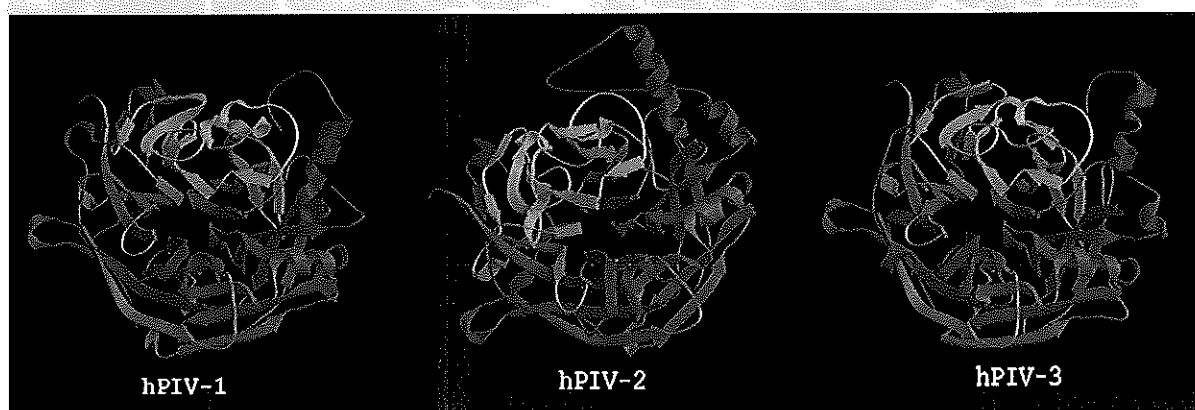


Figure 2.44: Molscript^{49, 50} diagrams of the tertiary structure of hPIV-1, 2, & 3 homology models.

2.4.1.1.1 Optimisation of H-bond network by flipping His, Gln and Asn residues.

The hPIV-1, 2 & 3 HN homology models were H-bond network optimised using the program Reduce.⁸⁸ Reduce was found to optimise the H-Bond network better than the program Whatcheck. Even though both programs can satisfactorily optimise the H-bond network, Reduce takes into account factors that Whatcheck does not, such as whether a flipped residue makes a bad contact with the protein surface.

2.4.1.1.2 Characterisation of homology model active sites with DelPhi, Pocket, and PASS.

The programs DelPhi, Pocket and PASS were used to evaluate the electrostatic potential, interactions, and size respectively of the active sites of the three generated homology models. Analysis of each of the homology models with DelPhi showed that the active sites are dominated around the triarginyl cluster by a positive field potential.

The glycerol binding pocket is lined by positive field potential on one side and negative on the other. This is separated by an area of neutral field potential. The C-4 binding pocket of each model is very large and has a predominantly positive potential, small neutral and negative areas are also found within the pocket. The electrostatic potential of each of the structures was seen to be very similar to that of NDV HN. The results can be seen in Figure 2.45.

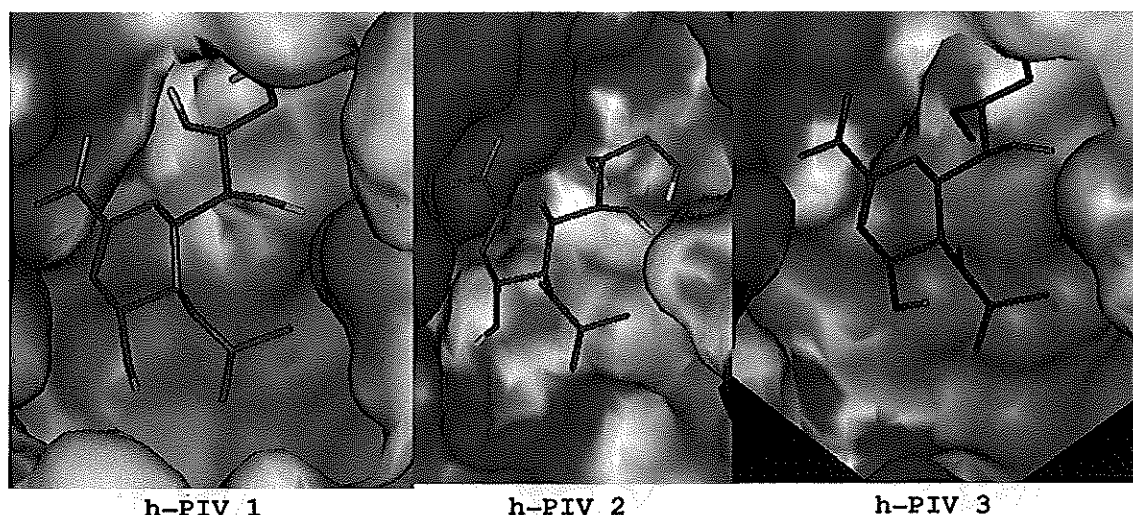


Figure 2.45: Delphi analysis of the electrostatic potential of the active sites of the hPIV-1, 2, and 3 HN homology models with Neu5Ac2en docked in the active sites (see Section 2.4.2). Blue: Positive potential, Red: Negative potential, White: Neutral potential.

Pocket analysis of the homology models showed that the active sites contain a large area (blue) for the hydrogen bond formation around the triarginyl cluster and into the C-4 binding pocket. A large area (red) of H-bond donors is found adjacent to the H-bond accepting region (blue). Hydrophobic interactions (green) are found at the terminus of the N-acetyl group binding pocket, along the pocket where the glycerol side-chain bonds and within the C-4 binding pocket. Figure 2.46 shows these interactions. All interactions found were similar to those found with the Pocket analysis of NDV HN; this is to be expected as NDV HN was used to model the structures.

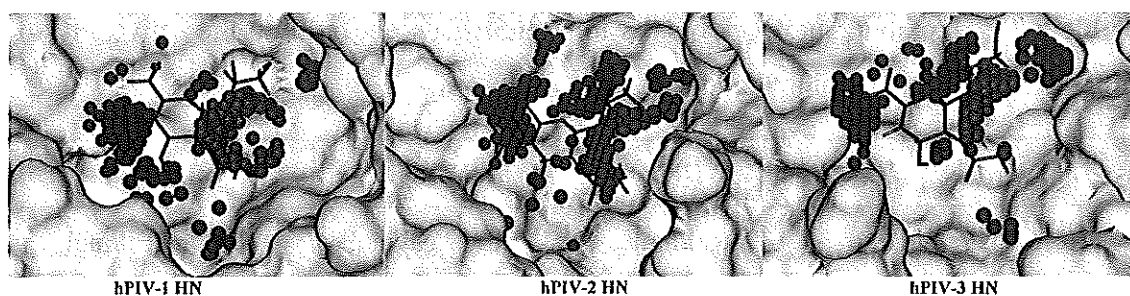


Figure 2.46: Pocket analysis of the active sites of the hPIV-1, 2, and 3 HN homology models with Neu5Ac2en docked in the active sites (see Section 2.4.2). Blue – hydrogen bond acceptor site, red – hydrogen bond donor site, green – hydrophobic interaction.

PASS was used to analyse the size of the active sites of each of the homology models. PASS analysis located a red sphere in the active site of each of the homology models. A red sphere is indicative of a large pocket, a white sphere is indicative of a medium sized pocket, and a blue sphere indicates a small pocket. PASS analyses are shown in Figure 2.47. Preliminary results indicated that the active site of each of the homology models is large. This is to be expected as the NDV HN crystal structure used to model the structures has a large active site.

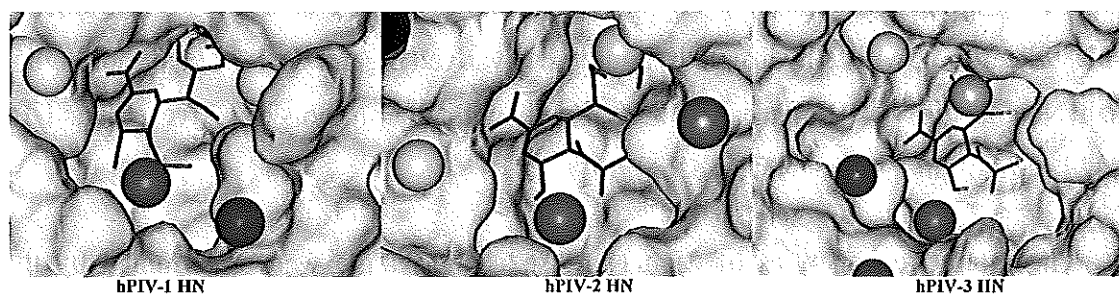


Figure 2.47: PASS analysis of the active sites of the hPIV-1, 2, and 3 HN homology models with Neu5Ac2en docked in the active sites (see Section 2.4.2).

2.4.2 Docking of Neu5Ac2en into hPIV-1, 2 and 3 HN homology models

The modelling protocol used to flexibly dock Neu5Ac2en to each model is discussed in Section 2.3.4.1. The flexible AutoDock interaction energies (Table 2.21) obtained for the docking of Neu5Ac2en to each of the homology models was very similar (a difference of -0.55 kcal/mol). They are also very similar to the energy of the flexible docking of Neu5Ac2en obtained previously for NDV HN (-10.06 kcal/mol, Section 2.3.3.1). These results are to be expected as high active site residue conservation occurs across all four proteins.

Table 2.21: Flexible AutoDock results for Neu5Ac2en docked into the active sites of hPIV-1, 2 & 3 HN homology models.

hPIV-1 HN	hPIV-2 HN	hPIV-3 HN	NDV HN
−10.16 kcal/mol	−10.42 kcal/mol	−10.71 kcal/mol	−10.06 kcal/mol

2.4.2.1 Holzer *et al.* derivatives of Neu5Ac2en and hPIV-2 HN homology model

Holzer *et al.* have reported the inhibition of the sialidase activity of hPIV-2 by a number of C-4 modified Neu5Ac2en derivatives.⁶⁹ It was of interest to see if a correlation could be observed between the energies for the docking these derivatives to the homology model of hPIV-2 HN and the experimental data. The compounds, presented in Table 2.22, were docked and scored against the homology model of hPIV-2 HN.

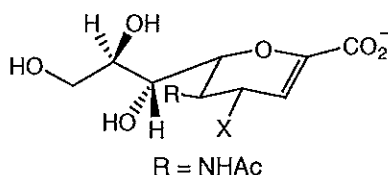
The flexible AutoDock score for each derivative of Neu5Ac2en ranged from −11.89 kcal/mol, with 2-hydroxyethylamino derivative (Ligand G, 1.83 kcal/mol higher than Neu5Ac2en: −10.06) to −10.48 kcal/mol, with the amino derivative (Ligand B, 0.42 kcal/mol higher than Neu5Ac2en).

The interaction energies of each Neu5Ac2en derivative were determined by the MMI protocol. The interaction energies of each derivative with the minimised NDV HN active site ranged from −176.35 to −198.82 kcal/mol. The C-4 prop-2-enylamino derivative (Ligand E, Table 2.22) scored the highest (25.30 kcal/mol higher than Neu5Ac2en: −173.52 kcal/mol) and the *N,N*-dimethylamino derivative (Ligand C, Table 2.22) scored the lowest (2.83 kcal/mol lower than Neu5Ac2en).

The SCORE results for each derivative range from the *N,N*-dimethylamino derivative (Ligand C, Table 2.22) scoring 4.88 to the C-4 prop-2-enylamino derivative (Ligand E, Table 2.22) scoring 6.82. The highest scoring derivative, C-4 prop-2-enylamino, showed an approximate increase of 90 times in binding affinity with respect to the lowest scoring derivative, *N,N*-dimethylamino. The results range from approximately 15 times less to approximately 5 times greater than the calculated binding affinity of Neu5Ac2en (pK_d 6.06, Section 2.3.3.2).

No increase in the static AutoDock score after minimisation was observed any of the nine derivatives. The change observed in the AutoDock score ranged from a maximum loss of 15.82 kcal/mol with the guanidino derivative (Ligand A, Table 2.22), to a minimum loss of 2.19 kcal/mol with the acetamido derivative (Ligand D, Table 2.22), after the minimisation process.

Table 2.22: AutoDock, MMI and SCORE results for C-4 modified Neu5Ac2en derivatives in hPIV-2 HN homology model of ligands known to inhibit hPIV-2.⁶⁹



C4-X Substituent	Flexible AutoDock (Unminimised) (kcal/mol)	MMI (Minimised) (kcal/mol)	Static AutoDock (Minimised) (kcal/mol)	SCORE (Minimised) (pK _d)	Experimental inhibition constant ⁶⁹
A NHC(=NH)NH ₂ ⁺	-11.71 [4] ^a	-183.97 [4]	4.11 [7]	5.56 [5]	1 x 10 ⁻⁵ M
B NH ₂ ⁺	-10.48 [7]	-181.27 [5]	2.37 [6]	6.34 [2]	8 x 10 ⁻⁴ M
C N(CH ₃) ₂	-11.15 [6]	-176.35 [7]	1.80 [5]	4.88 [7]	3 x 10 ⁻⁴ M
D NHC(=O)CH ₃	-11.82 [3]	-190.23 [3]	-9.63 [1]	5.67 [4]	1 x 10 ⁻⁴ M
E NHCH ₂ CH=CH ₂	-11.88 [2]	-198.82 [1]	-6.95 [2]	6.82 [1]	4 x 10 ⁻⁵ M
F NH(OH)CH ₂ CH=CH ₂	-11.43 [5]	-178.20 [6]	-2.41 [4]	5.44 [6]	2 x 10 ⁻⁵ M
G NHCH ₂ CH ₂ OH	-11.89 [1]	-191.59 [2]	-6.29 [3]	6.33 [3]	3 x 10 ⁻⁵ M

^a Numbers in square brackets indicate scoring rank.

Comparing the results from the various scoring methods with the experimentally determined inhibition data showed that no correlation between binding energy and inhibition had been found. This was not unexpected given that the modelling data is derived from a homology model, and as such there may be key differences between the actual structure of hPIV-2 HN and the homology model derived from NDV HN.

2.5 Computational analysis of hPIV-3 HN crystal structure.

2.5.1 Initial analysis of available hPIV-3 HN crystal structures.

Five crystal structures of hPIV-3 HN were recently released.⁵⁷ The structures available include: two ligand free structures (PDB accession code 1V2I: pH 6.5 structure at 2.2 Å and 1V3B: pH 7.5 structure at 2.0 Å), a Neu5Ac complex (1V3C: pH

7.5 structure at 2.3 Å), a Neu5Ac2en complex (1V3D: pH 7.5 structure at 2.28 Å) and a zanamivir complex (1V3E: pH 7.5 structure at 1.89 Å). Each crystal structure consists of 2 chains (a dimer), each containing residues 142 –572 of the HN protein. Each crystal structure also contains two calcium ions, N-acetylglucosamine, and either phosphate or sulphate ions. Superimposition of 1V2I (pH 6.5 apo structure) active site residues with NDV HN (pH 6.5) indicated an RMSD of 2.30 Å over all atoms of the residues. The RMSD is inflated by Tyrosine 530 which in hPIV-3 HN has the side chain flipped ~90°, to point the hydroxyl group away from the solvent exposed protein surface. The active site residues of hPIV-3 HN and the corresponding NDV HN residues are shown below in Table 2.23.

Table 2.23: Superimposed active site residues of NDV and hPIV-3 HN.

Amino Acid	hPIV-3 HN	NDV HN
Arginine	192	174
Arginine	424	416
Arginine	502	498
Tyrosine	530	526
Glutamic Acid	409	401
Glutamic Acid	549	547
Aspartic Acid	216	198

2.5.1.1 Procheck

Ramachandran plot analysis showed that core percentage was between 85.9 and 80.9% favoured, compared to 83.9% for NDV HN. Disallowed residues fell between 0.8 and 0.9%; NDV HN for comparison was 0.8%.

2.5.1.2 Secondary structure render

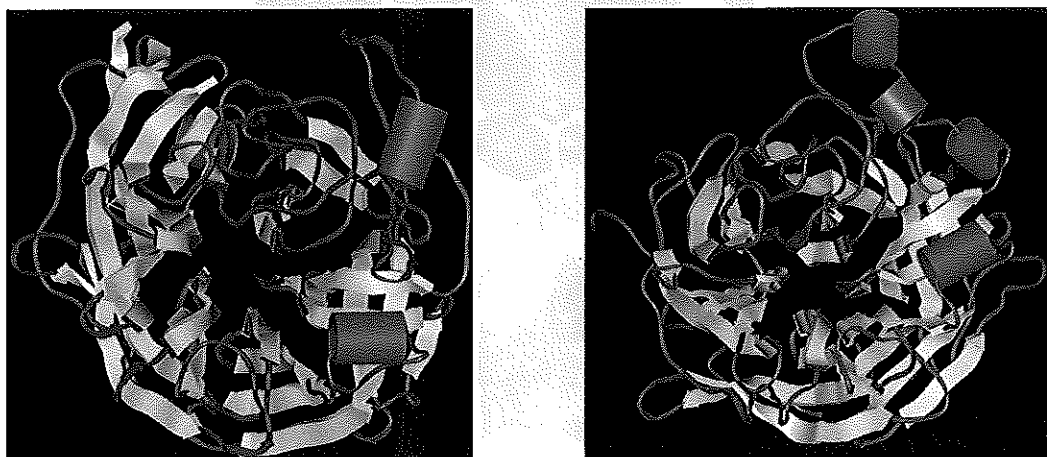


Figure 2.48: Secondary structure comparison between hPIV-3 HN crystal structure (PDB ID:1V3D) (left) and homology model (right).¹²³

Rendering the secondary structure of both the crystal structure 1V3D and the homology model of hPIV-3 HN (Section 2.4) showed similarities and differences between the two structures, as can be seen in Figure 2.48. The obvious similarity is the “classical” neuraminidase structure – pseudo six-fold symmetry β -sheet propeller. The structure of the homology model is noted, however, to be more disordered by the presence of more random coil domains than in the crystal structure. The α -helical domain of the homology model is also quite different to that of the crystal structure. It is as expected more similar to the NDV HN crystal structure used to generate the homology model.

2.5.1.3 Electrostatics of hPIV-3 HN active site

The electrostatic potential determined by DelPhi of the hPIV-3 HN crystal structure when compared to NDV HN was found to differ in a number of aspects. The positive potential around the triarginyl cluster is more localized around the carboxylate group of Neu5Ac2en than NDV HN, and the C-4 pocket is more positively charged than in NDV HN. The glycerol side-chain binding pocket is more negatively charged, whereas the pocket of NDV HN is found to be split down the middle lined by positive potential on one side and negative potential on the other.

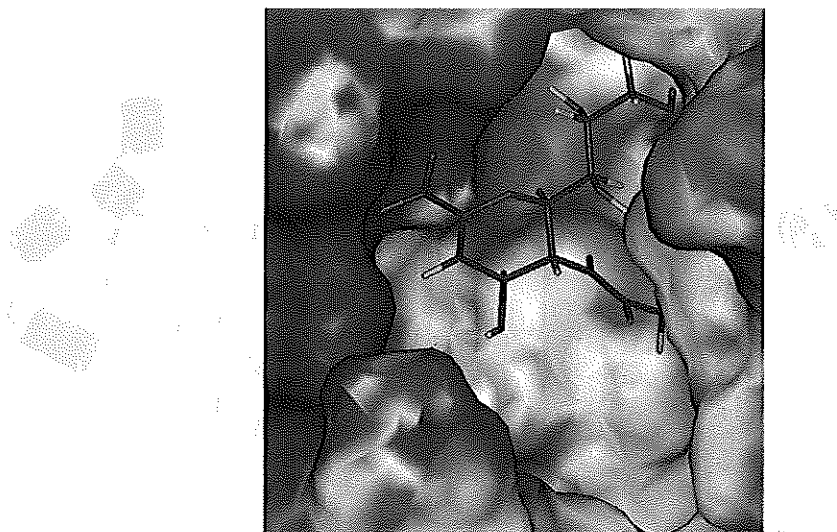


Figure 2.49: The electrostatic potential of the active site of the crystal structure of hPIV-3 HN (PDB ID: 1V3D).

2.5.1.4 C-4 binding pocket volume calculations

The volume of the C-4 pocket in the Neu5Ac2en bound crystal structure (PDB ID: 1V3D) as determined by PASS to be 1074 Å³, the volume of C-4 pocket in the zanamivir™ bound crystal structure was found to be not unexpectedly smaller at 706 Å³. For comparative purposes the volume of the C-4 pocket in the NDV HN crystal structure with Neu5Ac2en bound was found to be smaller than either hPIV-3 HN crystal structures at 612 Å³. These results show that the size of the binding pocket of hPIV-3 HN is much larger than that found in NDV HN.

2.5.2 AutoDock of Neu5Ac2en and Zanamivir.

Using AutoDock to score the crystal bound structure of Neu5Ac2en **5** in hPIV-3 HN (1V3D) yielded a static binding score of -7.27 kcal/mol. When flexibility of Neu5Ac2en was allowed (unminimised), a moderate increase to -9.72 kcal/mol was observed. However, it was found that a second conformation exists where the carboxylate group is found to bind into the glycerol side-chain binding pocket.

The crystal bound structure of zanamivir **6** in hPIV-3 HN (1V3E) yielded a static AutoDock binding score of -9.25 kcal/mol. When flexibility was introduced (unminimised), an increase in interaction energy to -12.76 kcal/mol was observed, with no alternate conformations observed. The docked orientation of the pose closest to that found in the crystal structure is shown in Figure 2.50.

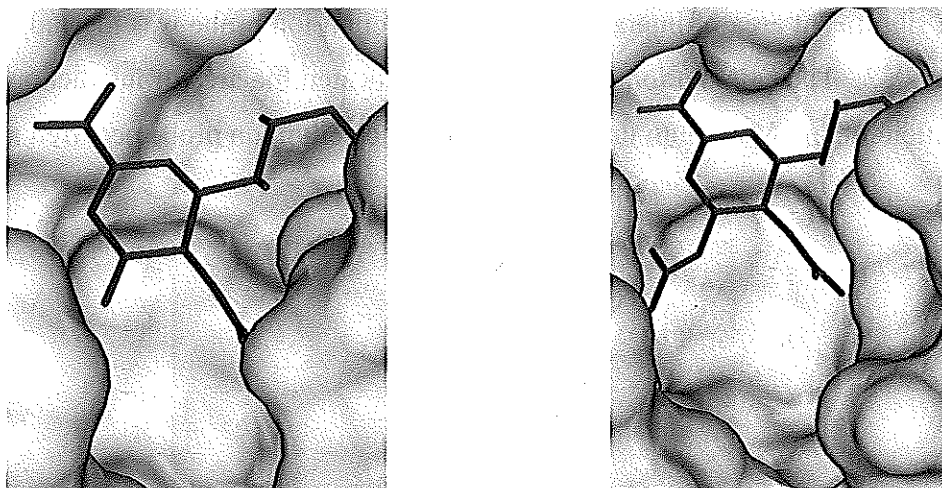


Figure 2.50: Orientation of Neu5Ac2en (left) and zanamivir (right) in hPIV-3 HN crystal structure. For Neu5Ac2en calculations, crystal structure 1V3D was used, in the case of zanamivir; structure 1V3E was used.

2.5.3 GRID analysis of hPIV-3 HN

GRID was utilised to identify optimum areas of interaction of the crystal structure of hPIV-3 HN (1V3D) with a series functional group probes. The structure was prepared using Greater and GRIN (Section 2.3.2.6).

One significant area of interaction in the active site of hPIV-3 HN was found with the carboxylate probe, seen in Figure 2.51. The area corresponds to the position of the carboxylate group of Neu5Ac2en, with interaction occurring with the triarginyl cluster. Maximum interaction with the probe occurred at -21.0 kcal/mol.

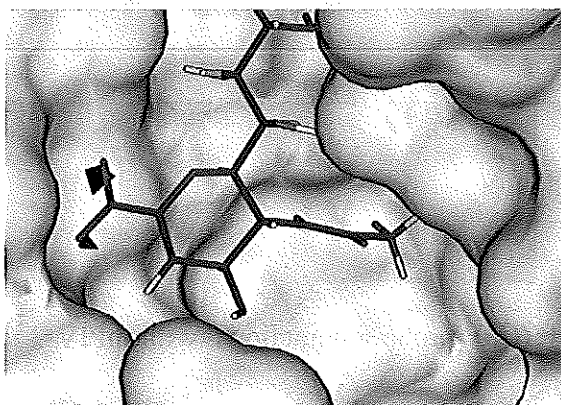


Figure 2.51: The carboxylate probe interaction site with hPIV-3 HN, contoured at -18 kcal/mol.

One significant area of interaction with the calcium ion probe was located, as seen in Figure 2.52. The calcium ion probe accurately located the area of interaction by the crystal calcium ion. Maximum interaction with the calcium ion probe occurred at -85.0 kcal/mol.

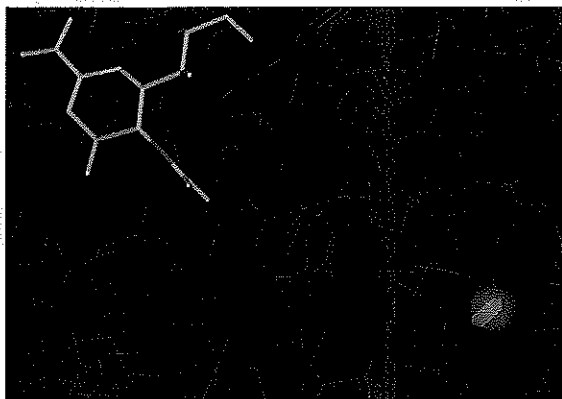


Figure 2.52: Calcium (Ca^{2+}) probe interaction site with hPIV-3 HN, contoured at -75.0 kcal/mol, crystal calcium is rendered in transparent white CPK, contour is coloured red.

Within the hPIV-3 active site three main areas of interaction were found with the methyl probe, seen in Figure 2.53. The methyl probe was found to interact with a large region of the glycerol side chain pocket, the C-5 binding pocket at the terminus of the NHAc group and the edge of the C-5 binding pocket. Maximum interaction with the methyl probe occurred at -4.0 kcal/mol.

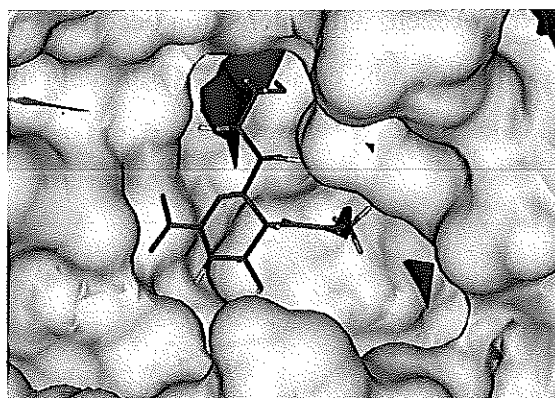


Figure 2.53: Methyl (C3) probe interactions with hPIV-3 HN, contoured at -3 kcal/mol.

Two areas of interaction with the hydrophobic probe were located within the active site of hPIV-3 HN, as seen in Figure 2.54. Interaction regions were found at the edge of the glycerol side-chain binding pocket, and within the C-5 binding pocket, stretching

into the C-4 binding pocket. Maximum interaction with the hydrophobic probe was -1.20 kcal/mol.

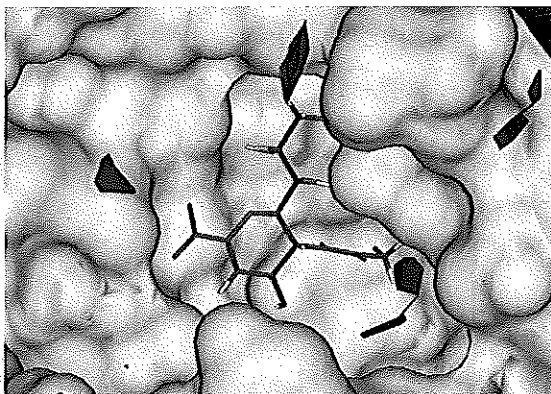


Figure 2.54: Hydrophobic (DRY) probe interaction sites with hPIV-3 HN, contoured at -0.8 kcal/mol.

Many areas of interaction were found with the water probe, as seen in Figure 2.55. Areas of interaction were located along the glycerol side chain, around the carboxylate group of Neu5Ac2en, and the region around the nitrogen of the C-5 acetamido group. The water probe predicted an area of interaction within the C-4 pocket of the active site corresponding to an area occupied by the crystal water residue 146 (rendered in CPK in Figure 2.55). Other water residues were not predicted. Maximum interaction with the water probe occurred at -10.0 kcal/mol.



Figure 2.55: Water (OH₂) probe interactions with the active site of hPIV-3 HN, contoured at -8 kcal/mol.

Four areas of interaction with the hydroxyl probe were located, as seen in Figure 2.56. Interactions were found along the glycerol side-chain, in the base of the C-4 binding pocket, in the area around the nitrogen of the C-5 acetamido group, and around

the carboxylate group of Neu5Ac2en. Maximum interaction with the hydroxyl probe occurred at -10.0 kcal/mol.

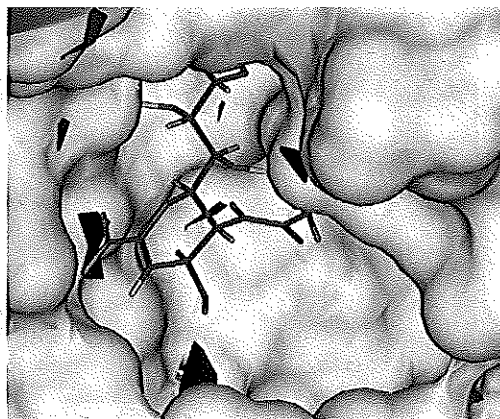


Figure 2.56: Aliphatic hydroxyl (O1) probe interaction sites with hPIV-3 HN, contoured at -8 kcal/mol.

Two areas of interaction with the guanidino probe were located within the active site, seen in Figure 2.57. Both areas of interaction are within the glycerol side-chain binding pocket. An area is located along the glycerol side chain and the other is found at the end of the pocket. Maximum interaction with the guanidino probe occurred at -12.50 kcal/mol.

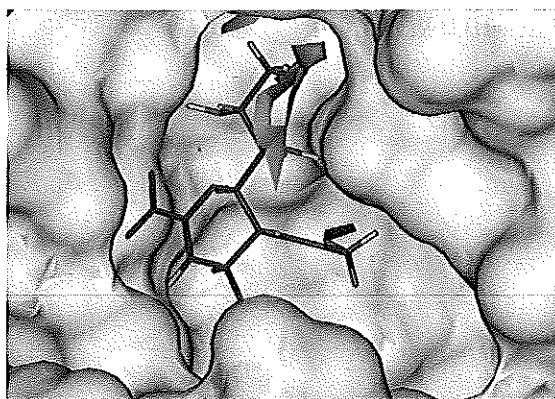


Figure 2.57: Guanidino (AMIDINE) probe interactions with active site of hPIV-3 HN, contoured at -10.0 kcal/mol.

Interactions with the protonated amino probe were located within the glycerol-side chain pocket, and multiple sites were found within and at the base of the C-4 and C-5 binding pockets, as seen in Figure 2.58. The maximum interaction with the protonated amino probe occurred at -10.0 kcal/mol.

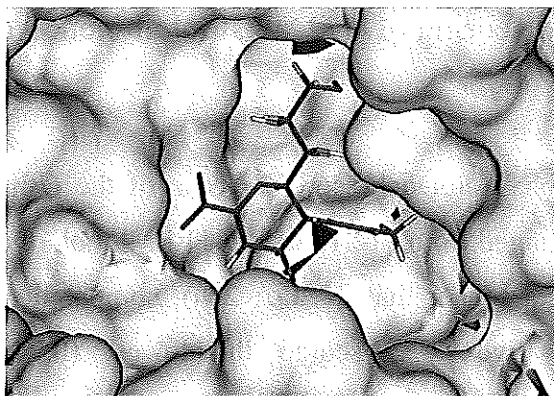


Figure 2.58: Protonated Amino (N3+) probe interaction sites with hPIV-3 HN, contoured at -8 kcal/mol.

No competition between structural water molecules and the probes tested (N3+ and DRY) was detected in the C-4 binding pocket hPIV-3 HN, therefore the water residues observed in the C-4 pocket in hPIV-3 HN should be displaced by ligand interactions with the active site. In the case of the N3+ probe, water would preferentially bind to the triarginyl cluster before the probe.

GRID analysis of the hPIV-3 HN crystal structure has identified key interactions in the active site. The program correctly identified the binding region of the carboxylate group of Neu5Ac2en and the binding region of the calcium ion to the protein. In the analysis of other probes tested, positive (amino, guanidino) probes located interactions along the glycerol side-chain and small areas within the C-4 and C-5 binding pockets, this result is similar to that found with the GRID study of NDV HN. The hydrophobic probes (C3 and DRY) found limited interactions within the C-4 pocket, this is a key difference to the results found in the study of NDV HN. The hydrophobic probe interactions found at the terminal end of the glycerol side-chain in NDV HN are also present in the structure of hPIV-HN. The C-4 hydrophobic derivatives of Neu5Ac2en designed based on the NDV HN crystal structure may not bind or may bind with decreased affinity to hPIV HN. To see if this could be predicted computationally the 4-*O*-substituted-Neu5Ac2en derivatives were docked into the active site hPIV-3 HN as described in Section 2.5.4.

2.5.4 Docking and scoring of 4-*O*-substituted-Neu5Ac2en derivatives and hPIV-3 HN.

The initial flexible AutoDock score for each of the *O*-linked derivatives docked against hPIV-3 HN ranged from -12.15 to -17.11 kcal/mol. Derivative **29-E** scored the highest (5.90 kcal/mol higher than Neu5Ac2en), compared to Neu5Ac2en at -11.21 kcal/mol, and derivative **29-C** scored the lowest (0.94 kcal/mol higher than Neu5Ac2en). After disabling the flexible torsions of the ligands to score the post minimisation interaction energy with AutoDock, the static AutoDock energy of all of the compounds decreased after the minimisation process. The interaction energy as scored by static AutoDock ranged from -9.55 to -14.10 kcal/mol after minimisation, this is compared to Neu5Ac2en which scored -9.36 kcal/mol. Derivative **29-E** post minimisation scored the highest score with -14.10 kcal/mol. The lowest scoring derivative was **29-I** with -9.55 kcal/mol.

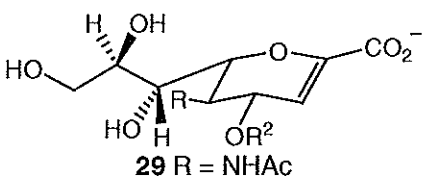
The MM-PBSA interaction energies of each of the *O*-linked derivatives of the Neu5Ac2en template with the minimised hPIV-3 HN active site ranged from -149.85 to -96.86 kcal/mol. Derivative **29-E** scored the highest (36.74 kcal/mol higher than Neu5Ac2en) and derivative **29-I** scored the lowest (16.25 kcal/mol lower than Neu5Ac2en). The MM-PBSA interaction energy of minimised Neu5Ac2en with NDV HN was -113.11 kcal/mol.

The X-Score results ranged from 5.21 with derivative **29-H** to 6.89 with derivative **29-B**; the highest scoring substituent is approximately 47.86 times stronger than the lowest scoring derivative, with respect to binding affinity. The results range between approximately 1.2 times lower to approximately 39.81 times greater than the binding affinity of Neu5Ac2en. The minimised crystal structure containing Neu5Ac2en has an X-Score result of 5.29. The top scoring molecule in the AutoDock energy scoring routine, **29-E**, shows a binding affinity 3.89 times greater than that of Neu5Ac2en.

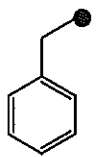
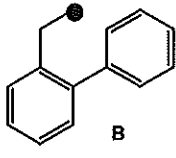

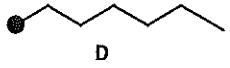
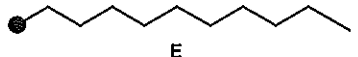
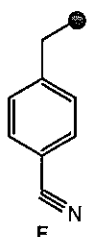
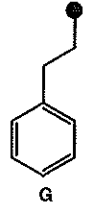
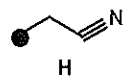
Consensus scoring of the derivatives scored against hPIV-3 HN showed that all derivatives scored better than Neu5Ac2en in two out three scoring functions and most scored better in three out of three; these are different results to those obtained from NDV HN. These results also contradict the hydrophobic GRID probe analysis performed in Section 2.5.3, where hydrophobic interaction areas were not extensively

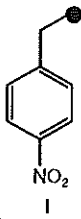
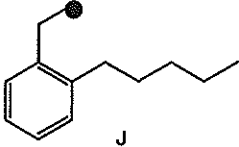
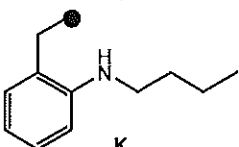
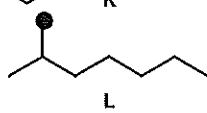
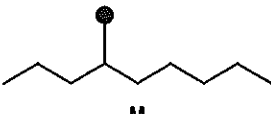
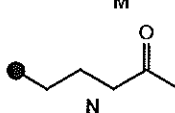
predicted for the C-4 binding pocket, but rather only a small area adjacent to the C-5 binding pocket.

Table 2.24: Docking and scoring results of 4-*O*-substituted-Neu5Ac2en derivatives **29** in hPIV-3 HN.^a



29 R = NHAc

R ² Substituent	Flexible AutoDock (Unminimised) (kcal/mol)	MM-PBSA (Minimised) (kcal/mol)	Static AutoDock (Minimised) (kcal/mol)	X-Score (Minimised) (pK _d)
H (Neu5Ac2en)	-11.21	-113.11	-9.36	5.29
	-14.13 [9] ^b	-110.75 [13]	-10.89 [11]	5.97 [6]
	-16.65 [2]	-128.05 [6]	-13.91 [2]	6.89 [1]
	-12.15 [14]	-112.48 [12]	-9.72 [13]	5.31 [11]
	-14.84 [5]	-136.19 [2]	-12.48 [7]	5.68 [9]
	-17.11 [1]	-149.85 [1]	-14.10 [1]	5.88 [8]
	-13.27 [11]	-126.57 [7]	-11.63 [9]	5.98 [5]
	-14.15 [8]	-131.12 [4]	-13.39 [5]	5.96 [7]
	-12.97 [12]	-118.35 [11]	-10.82 [12]	5.21 [12]

R ² Substituent	Flexible AutoDock (Unminimised) (kcal/mol)	MM-PBSA (Minimised) (kcal/mol)	Static AutoDock (Minimised) (kcal/mol)	X-Score (Minimised) (pK _d)
 I	-12.89 [13]	-96.86 [14]	-9.55 [14]	5.97 [6]
 J	-15.55 [4]	-119.80 [10]	-13.44 [4]	6.14 [4]
 K	-15.66 [3]	-120.05 [9]	-13.15 [6]	6.20 [3]
 L	-14.61 [7]	-133.73 [3]	-12.03 [8]	5.47 [10]
 M	-14.80 [6]	-130.56 [5]	-13.60 [3]	6.24 [2]
 N	-14.08 [10]	-121.90 [8]	-11.60 [10]	5.68 [9]

^a The data presented here has been generated by the second computational scoring protocol described in Section 2.3.7.3, ^b Numbers in square brackets indicate scoring rank.

2.5.4.1 Examination of docking and scoring of hPIV-3 HN with crystal water residues within the C-4 binding pocket of the active site.

It was of interest to examine the affect of crystal water residues on the binding of Neu5Ac2en to hPIV-3 HN, and to also examine the stability of crystal water residues within the active site. The known inhibitor Neu5Ac2en **5** was docked, using AutoDock (Section 2.3.4.1), into the active site (PDB ID: 1V3D) with and without the crystal water residues: 7, 137, 146, 185, 218, 237, 271, 274, 402, 664, 665, and 666. The AutoDock interaction energy without water residues was found to be -10.18 kcal/mol. When the water residues were included, the interaction energy is found to show a small increase to -11.66 kcal/mol (an increase of 1.48 kcal/mol). Interestingly, while these calculations were performed with flexibility allowed, the calculation where water residues were included showed no change from the crystal structure conformation. The

increased interaction energy is due to the formation additional hydrophobic interactions with Arg 424. The details of these bonds are found in Table 2.25 and Figure 2.59.

Table 2.25: Hydrogen bonds formed in the docking of Neu5Ac2en to hPIV-3 HN.

Atom	Water	Atom	Neu5Ac2en	Distance	Temperature factor of water
1	Residue	2		(Å)	residue
O	137	H	C-4 OH	2.85	High
H	271	O	C-4 OH	2.71	Moderate

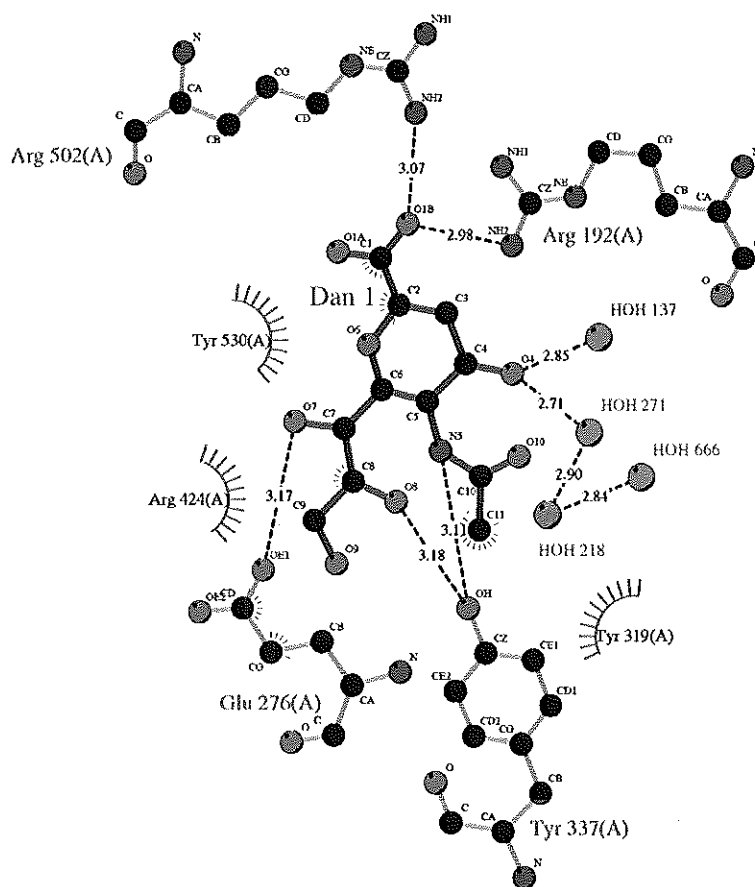


Figure 2.59: Ligplot diagram showing water interactions in hPIV-3 active site with Neu5Ac2en.

2.6 Overview and summation of the modelling procedure

The ultimate goal of the molecular modelling studies carried-out was to design inhibitors of hPIV HN. In the first instance, with structural data available at the time, this involved using the structure of the known sialidase inhibitor Neu5Ac2en **5** as a template, and the crystal structure of the paramyxoviral HN from NDV. Initial analyses and calculations set the benchmarks, against which subsequent calculations would be

measured. The first of these analyses consisted of defining the H-bond network within the NDV HN active site. After this, probing the active site with GRID and Pocket was performed, so to gain an insight as to what functional groups would be best accommodated by the active site architecture. With respect to the C-4 binding pocket, GRID and Pocket analyses demonstrated that hydrophobic interactions were favoured.

Following the probing of the active site, the scoring and docking programs were first used to score and dock crystal bound Neu5Ac2en. This had two purposes, firstly to evaluate whether the scoring functions and docking algorithms worked with this particular system, and secondly, to set benchmark scores. In the first instance, SCORE, AutoDock and Discover (MMI), X-Score, and AMBER (MM-PBSA) were found to perform adequately with this system. Dock 4 however could not successfully reproduce the position of Neu5Ac2en in the crystal structure of NDV HN. Benchmarks scores were successfully measured using the programs SCORE, AutoDock, MMI, X-Score, and MM-PBSA. A summary of selected benchmarks (Section 2.3.7.3) measured with Neu5Ac2en docked to NDV HN (either crystallographically or computationally) can be seen in Table 2.26.

Table 2.26: Benchmarks set by various programs^a for Neu5Ac2en in NDV HN.

Program Used	Calculated Benchmark
AutoDock (Flexible – preminimisation)	-12.01 kcal/mol
AutoDock (Static – postminimisation)	-102.50 kcal/mol
MM-PBSA (Postminimisation)	-6.36 kcal/mol
X-Score (Postminimisation)	pK _a 5.12

^a Benchmarks calculated using refined modelling protocol from Section 2.3.7.3.

Following the setting of the benchmark scores, for Neu5Ac2en the derivatives of the template Neu5Ac2en **5** were constructed using LigBuilder. LigBuilder constructed a total of 90 C-4 modified Neu5Ac2en derivatives, each of which was scored and docked against NDV HN. The resultant scores were compared to Neu5Ac2en benchmarks for each of the scoring programs. In most cases SCORE and AutoDock results were scattered around the benchmark results obtained for each program. Once the designed ligands had been scored and ranked (Section 2.3.7), the top 9 side-chains were examined and modified to make them more chemically accessible. Using the modified designed side-chain structures **A-N** (Section 2.3.7.1), the final fourteen

substituent structures were decided on. These modified side-chains were then added to the C-4 position of Neu5Ac2en by both thioether **28** and ether **29** linkages and the resulting Neu5Ac2en derivatives were minimised and scored. In most cases the modification of the structure increased the scores of the designed side-chains. The next phase in the development of potential inhibitors of paramyxoviral HN was to synthesise the designed compounds; this is discussed in Chapter 3.

3 Synthesis of designed C-4 modified Neu5Ac2en derivatives as potential inhibitors of paramyxoviral HN.

3.1 Introduction

Over the years the derivatisation of *N*-acetylneuraminic acid (Neu5Ac, **1**) has yielded modifications at all carbon atoms of the 9-carbon sugar (reviewed in references 8, 9, 124-128). The resulting modified sialic acids have been used to examine the activities of sialic acid recognizing proteins (SARPs). Synthetic strategies to derivitise *N*-acetylneuraminic acid generally utilise the sialic acid as a protected species with the C-1 and C-2 positions protected *via* esterification and glycosidation respectively. Various hydroxyl protecting groups such as acetate, benzoate, benzyl, silyl and isopropylidene groups have also been used during the synthesis of modified sialic acids. Examples of the modifications carried-out on *N*-acetylneuraminic acid **1**, are shown in Table 3.1.

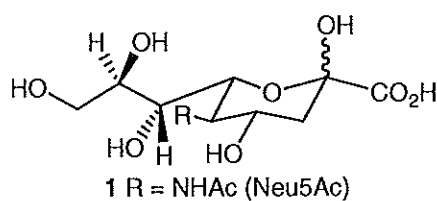
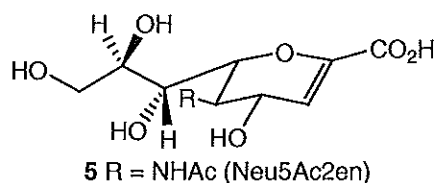


Table 3.1: Some of the modifications carried-out on Neu5Ac **1**.

Atom	Modifications carried-out
C-1	Carboxamide formation, ^{129, 130} reduction to hydroxymethyl, ^{129, 130} removal. ¹³¹
C-2	Dehydroxylation, ¹³² halogenation, ¹³³ elimination between C-2 and C-3. ¹³⁴⁻¹³⁶
C-3	Halogenation. ^{136, 137}
C-4	Epimerisation, ¹³⁸ oxidation, ¹³⁹⁻¹⁴¹ deoxygenation, ¹⁴² <i>O</i> -acylation, ¹⁴³ <i>O</i> -alkylation, ¹⁴⁴ introduction of <i>N</i> , ¹⁴⁵ <i>S</i> , ¹¹⁷ <i>C</i> -linked substituents, ¹⁴⁶ oxazoline formation with C-5 NHAc. ^{117, 147}
C-5	De- <i>N</i> -acetylation, ¹³⁰ alternate acylation. ¹³⁵
C-6	All sialic acids modified at C-6 have been synthesized from hexose sugars; amino, ¹⁴⁸ hydroxymethyl, ¹⁴⁹ methyl, ¹⁴⁹ and thio ¹⁵⁰ derivatives
C-7	Epimerisation, ^{139, 151} oxidation, ¹⁵² deoxygenation, ¹⁵³ epoxide formation, ^{151, 154, 155} introduction of azide, ¹⁵⁴ truncation at C-7 of glycerol side chain. ¹⁵⁶
C-8	Epimerisation, ^{139, 151} oxidation, ¹⁵² deoxygenation, ¹⁵³ epoxide formation, ^{151, 154, 155} introduction of azide, ¹⁵⁷ truncation at C-8 of glycerol side chain. ¹⁵⁶
C-9	Oxidation, ¹²⁴ deoxygenation, ¹⁵⁸⁻¹⁶⁰ <i>O</i> -acylation, ¹⁴⁴ <i>O</i> -alkylation, ¹⁴⁴ tosylation, ¹⁶¹ silylation, ¹⁵⁴ tritylation, ¹⁶² fluorination, ¹⁶² introduction of <i>N</i> -substituents. ¹⁶³

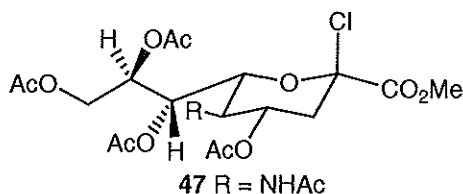
Of the modifications to *N*-acetylneuraminic acid presented in Table 3.1, the most relevant within the scope of this thesis is the modification of C-2/C-3 to the naturally

occurring 2-deoxy-2,3-didehydro Neu5Ac derivative; or as it is more commonly known, Neu5Ac2en **5**.

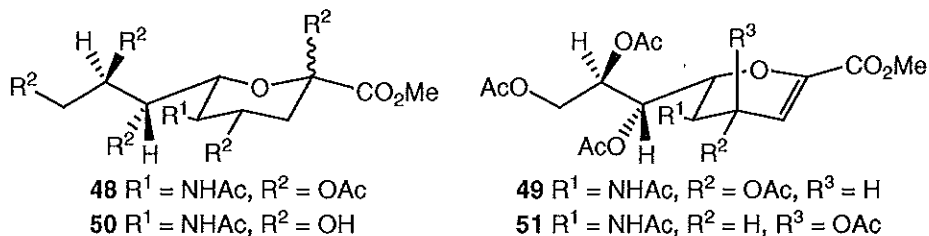


3.1.1 Synthesis of Neu5Ac2en

The synthesis of Neu5Ac2en, and of Neu5Ac2en derivatives, has been accomplished by a number of procedures. The chemical synthesis of Neu5Ac2en **5** from Neu5Ac **1** is usually accomplished by elimination of the β -chloride from methyl (5-acetamido-4,7,8,9-tetra-*O*-acetyl-2-chloro-3,5-dideoxy-D-glycero- β -D-galacto-2-nonulopyranosyl chloride)onate **47**. Esterification of the C-1 carboxyl group of Neu5Ac¹⁶⁴ is followed by treatment of the product with acetyl chloride to afford the per-*O*-acetylated β -chloride **47**.¹³³ Treatment of **47**, with a base such as pyridine,¹³⁴ triethylamine¹³⁵ or diazabicycloundecene (DBU),¹³⁶ and subsequent Zemplén deacetylation and de-esterification gives the final product, Neu5Ac2en **5**.



In an alternative syntheses of Neu5Ac2en, heating per-*O*-acetylated Neu5Ac1Me **48** in dioxane at 90°C,¹³⁵ or reaction of **48** with a catalytic amount of trimethylsilyltriflate at room temperature,¹⁶⁵ have been reported to give the per-*O*-acetylated derivative of Neu5Ac2en1Me, **49**. The reaction of Neu5Ac1Me **50** under acetolysis conditions (Ac₂O, cat. H₂SO₄, 50°C) affords a C-4 epimeric mixture (**49** and **51**) of acetylated Neu5Ac2en1Me.¹⁶⁶



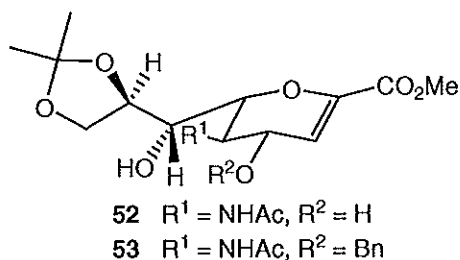
Neu5Ac2en has also been noted in the literature to form as a by-product in glycosidation reactions, and to be formed at basic pH from the natural nucleotide sugar

CMP-Neu5Ac.¹⁶⁷ It has also been observed to form during the action of influenza B sialidase on sialylactose.¹⁶⁸

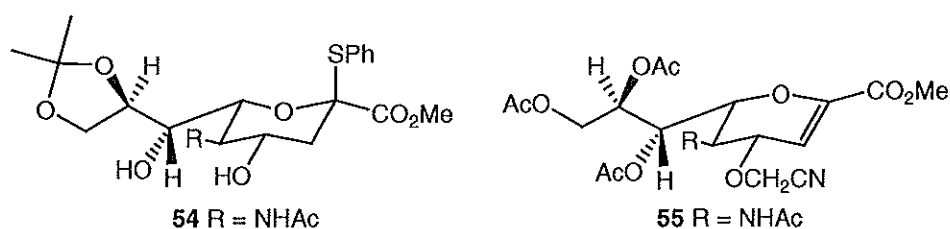
3.1.2 Synthesis of Neu5Ac2en derivatives

As with the derivatisation of Neu5Ac **1**, many derivatives of Neu5Ac2en, **5**, have been previously synthesised.^{8, 9, 117, 143-146} Within the scope of this thesis, the derivatisation of C-4 is of most interest. Functionalisation of the C-4 position has produced oxygen, nitrogen and sulfur linked derivatives of Neu5Ac2en.^{117, 143-146}

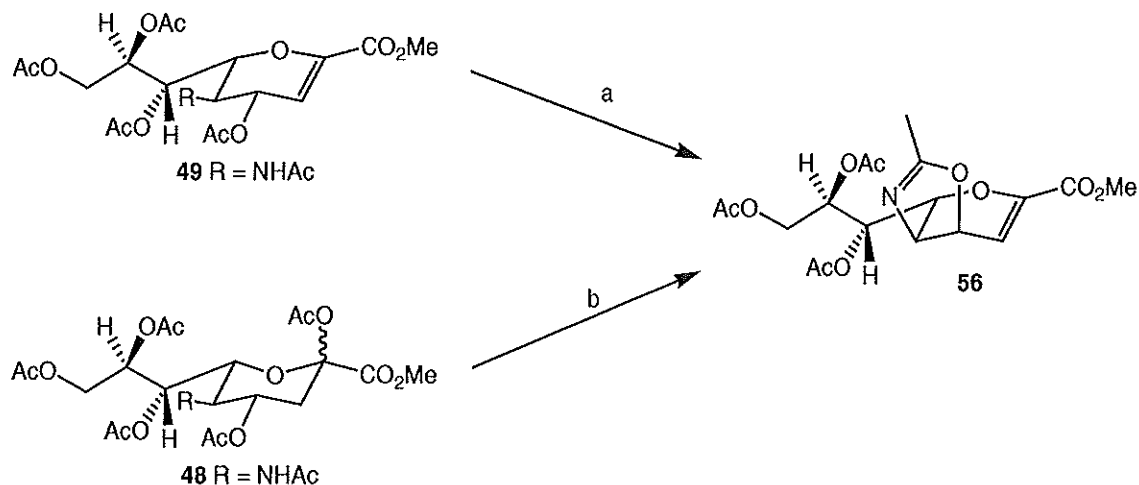
Access to C-4 oxygen linked derivatives of Neu5Ac2en **5** can be achieved *via* the 8,9-*O*-isopropylidenated Neu5Ac2en1Me derivative **52**. For example, reaction of **52** with benzyl bromide in the presence of sodium hydride has been reported to give 4-*O*-benzyl-8,9-*O*-isopropylidene-Neu5Ac2en1Me **53** in 83% yield. This appears to be the only reported *O*-alkylation of compound **52** in refereed publications at the time of writing.¹¹⁸ Conceivably, this reaction could be adapted to alkylation with any suitable alkyl or aryl halide.



In an alternative approach, C-4 oxygen linked derivatives have been produced *via* suitable modification at C-4 of Neu5Ac1Me derivatives and subsequent introduction of 2,3-unsaturation to give C-4 *O*-linked Neu5Ac2en derivatives. This approach has been demonstrated by Ikeda *et al.*¹⁴⁴ The cyanomethyl group has been introduced at C-4 of the 8,9-*O*-isopropylidenated thiophenyl glycoside of Neu5Ac1Me **54** *via* coupling with bromoacetonitrile in the presence of silver (I) oxide. De-*O*-isopropylidenation in 80% aqueous acetic acid, followed by acetylation using acetic anhydride in pyridine, gave suitable material for elimination. Elimination across C-2 and C-3 to give the desired per-*O*-acetylated 4-*O*-cyanomethyl-Neu5Ac2en derivative **55** was accomplished in two steps, by successive treatment with dimethyl(methylthio)sulfonium triflate (DMTST) and DBU. The isolated yield of **55** was 36% over 4 steps from **54**, with the yield limiting step in this reaction sequence being the dual silver (I) oxide coupling reaction/de-*O*-isopropylidenation step (47% yield).¹⁴⁴



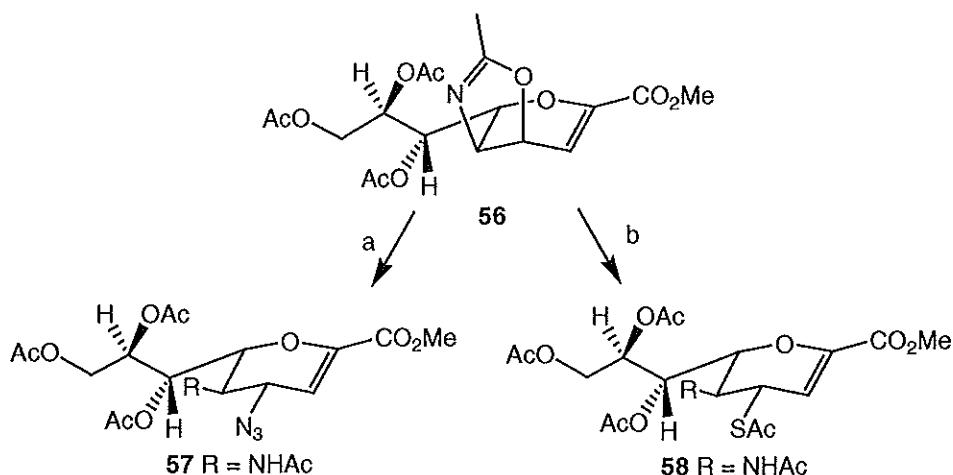
The derivatisation at C-4 of Neu5Ac2en with sulfur or nitrogen proceeds *via* similar synthetic routes. Treatment of per-*O*-acetylated Neu5Ac2en1Me **49** with a Lewis acid catalyst such as boron trifluoride etherate¹¹⁷ results in the formation of the 4,5-oxazoline **56** (Scheme 3.1a). Beginning from the saturated species, the treatment of per-*O*-acetylated Neu5Ac1Me **48** with trimethylsilyl trifluoromethanesulfonate (TMSOTf)¹⁴⁷ results in both formation of the 4,5-oxazoline and introduction of 2,3-unsaturation, giving **56** (Scheme 3.1b).



Scheme 3.1

Reagents & Conditions. a) $\text{BF}_3 \cdot \text{Et}_2\text{O}$, dry DCM, dry MeOH, 16 h, rt;¹¹⁷ b) TMSOTf, dry EtOAc, 2 h, 53°C.¹⁴⁷

The treatment of the oxazoline **56** with either lithium azide¹¹⁷ or thiolacetic acid¹¹⁷ gives the per-*O*-acetylated 4-azido-4-deoxy **57**, or 4-thiolacetate **58**, derivatives of Neu5Ac2en1Me, respectively, thus giving entry to both C-4-*N* or C-4-*S* linked Neu5Ac2en derivatives (Scheme 3.2). A range of 4-*N*-substituted Neu5Ac2en derivatives have been prepared from **57**.^{105, 106} However, the 4-thiolacetate derivative **58** is the only C-4 *S*-substituted Neu5Ac2en analogue reported in the literature to date.¹¹⁷

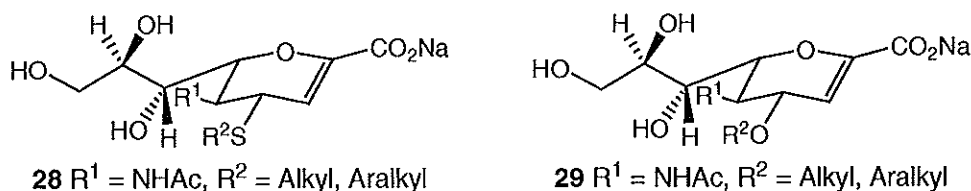


Scheme 3.2

Reagents & Conditions. a) LiN_3 , Dowex-50W \times 8 (H^+) resin, dry DMF, 16 h, 80°C ;¹¹⁷ b) thiolacetic acid, dry DMF, dry MeOH, 20 h, 83°C .¹¹⁷

3.1.3 Approaches to Neu5Ac2en C-4 thioether and ether derivatives

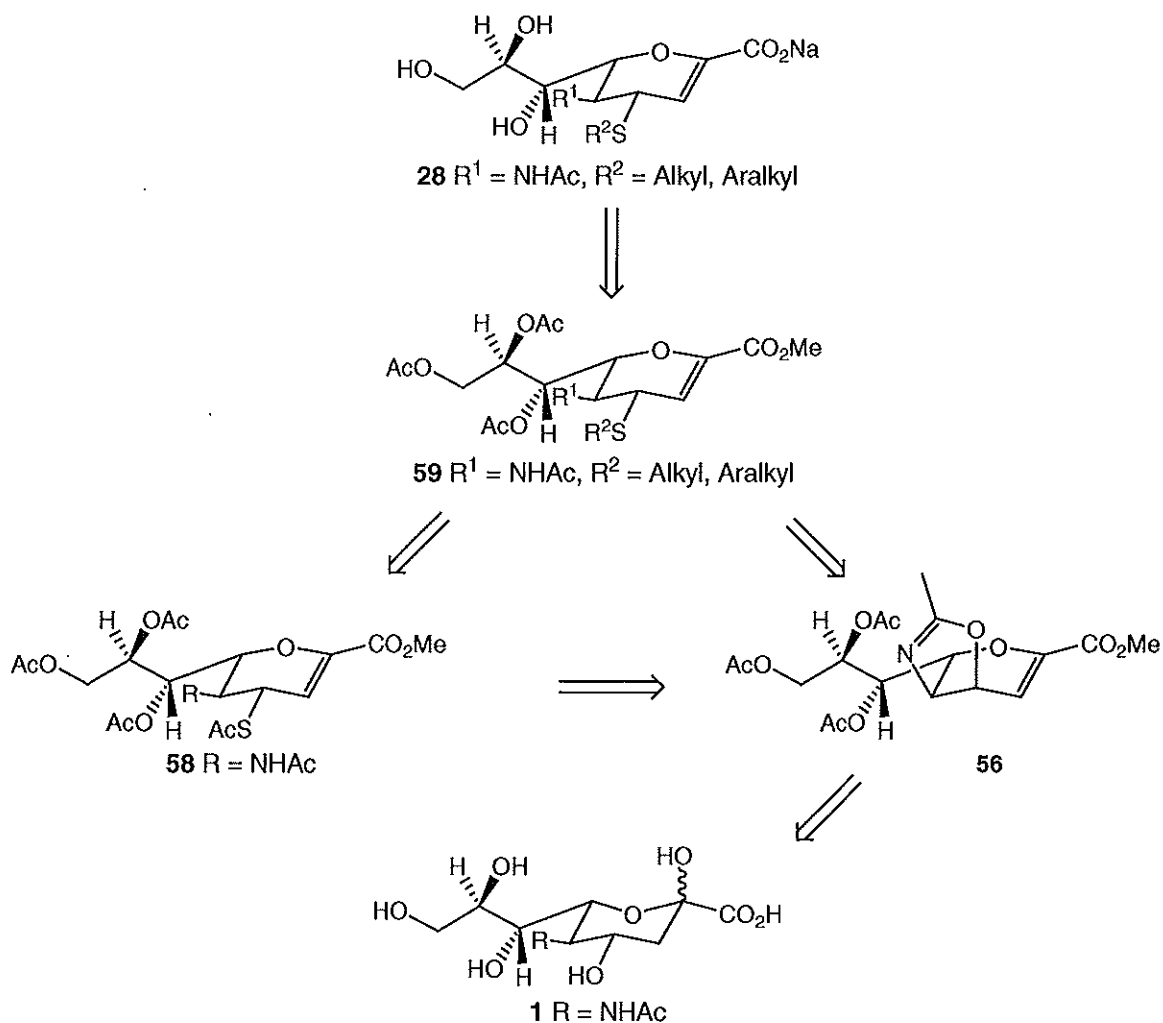
Modelling experiments (described in Chapter 2) suggested a number of modifications to Neu5Ac2en at the C-4 position that might increase the binding interactions of Neu5Ac2en to the HN glycoprotein active/binding site. These modifications involved either a thioether or ether linkage to a number of alkyl or aralkyl substituents. As already noted, there are no C-4-thioether derivatives of Neu5Ac2en reported in the literature. Thioether derivatives were therefore chosen for synthesis as interesting novel compounds with theoretically straight-forward synthesis from the known 4-thiolacetate Neu5Ac2en derivative **58**. A number of C-4-ether derivatives were also chosen to allow comparison of the activity between the two classes of compounds. The general structures of the C-4 thioether **28** and ether **29** derivatives are shown below. The strategies for preparation of the C-4-thioether derivatives are described in Section 3.2, and C-4-ether derivatives are described in Section 3.3.



3.2 C-4-Thioether Neu5Ac2en derivatives

A retro-synthetic approach to the formation of C-4-thioether Neu5Ac2en derivatives, **28**, is shown in Scheme 3.3. One method of thioether synthesis involves the attack of a thiolate ion on an alkyl bromide or triflate.¹⁶⁹ This method of thioether

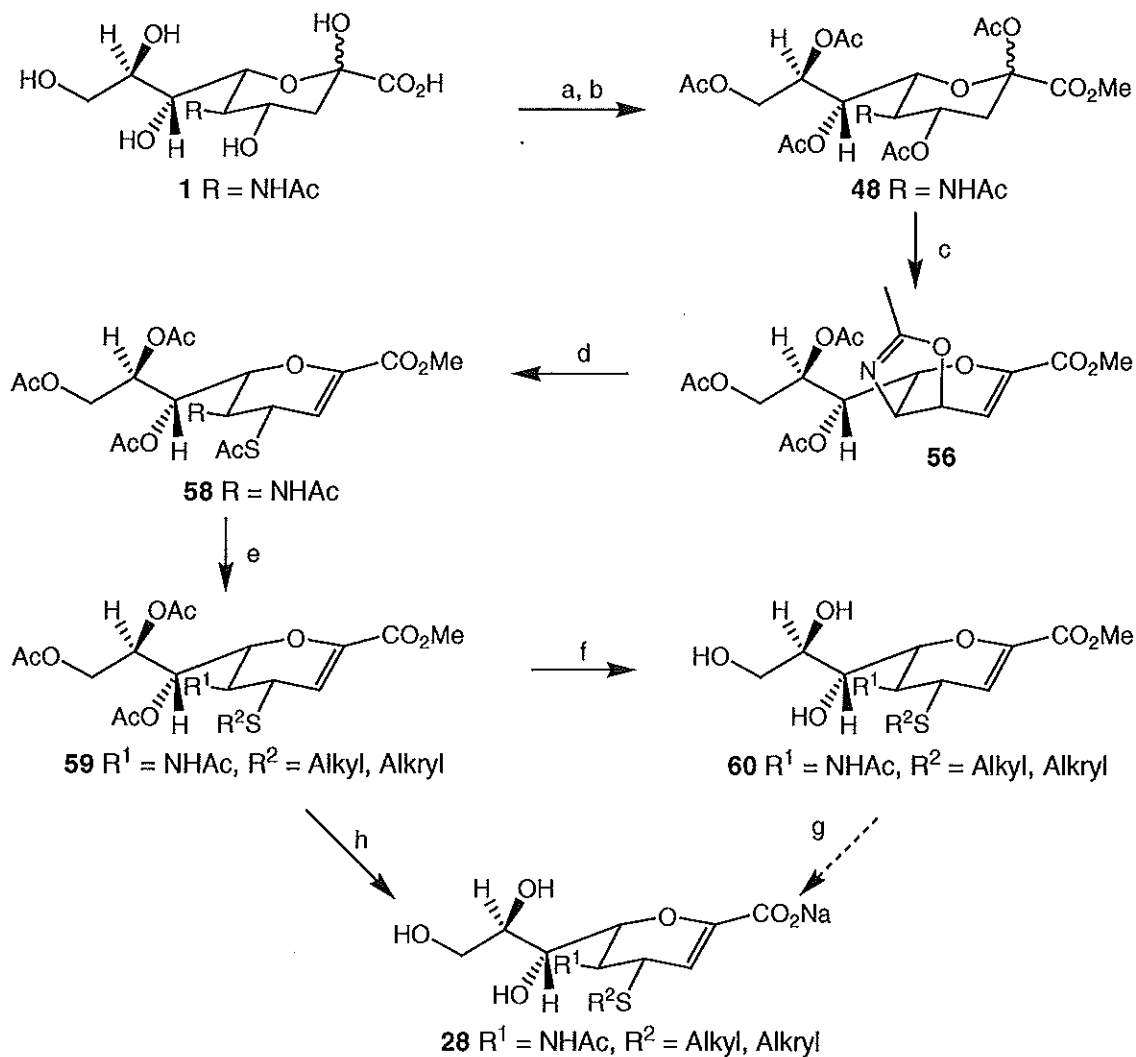
preparation has been successfully used for *S*-ether formation at both the anomeric position,¹⁶⁹⁻¹⁷¹ and at other positions (both primary and secondary) on carbohydrates.^{171, 172} Therefore, the starting material for the synthesis of Neu5Ac2en C-4-thioether derivatives would be the C-4 thiolacetate derivative **58**. Reaction of the C-4 thiolate ion generated from **58**, with appropriate alkyl halides would theoretically allow for the introduction of substituents suggested from the previous modelling experiments (described in Section 2.3.7.2). The synthesis of the key C-4 thiolacetate substituted Neu5Ac2en derivative **58** was reported in 1993.¹¹⁷ As described in Section 3.1.2, stereoselective introduction of the thiolacetate group at C-4 is achieved starting from the 4,5-oxazoline derivative **56**.^{117, 147} The oxazoline **56** has in turn been prepared in three steps from Neu5Ac **1**.¹⁴⁷ It might also be possible to synthesise the C-4-thioether Neu5Ac2en derivatives directly from the oxazoline **56** by thiol attack under acidic conditions.



Scheme 3.3: Retrosynthetic approaches to C-4-thioether derivatives of Neu5Ac2en.

3.2.1 Synthesis of C-4-Thioether Neu5Ac2en derivatives.

The synthetic approach taken to C-4-thioether Neu5Ac2en derivatives **28** is outlined in Scheme 3.4. The key intermediate in the synthesis of these derivatives, the 4-thiolacetylated derivative **58**,¹¹⁷ was prepared in four steps from Neu5Ac **1** via the per-*O*-acetylated methyl ester **48** and the oxazoline derivative **56**.



Scheme 3.4

Reagents & Conditions: a) Amberlite IR-120 (H⁺), dry MeOH, 2 d, rt (95%); b) pyridine, Ac₂O, 2 d, rt (98%); c) TMSOTf, dry EtOAc, 2 h, 53°C (77%); d) thiolacetic acid, dry DMF, 20 h, 83°C (85%); e) hydrazine acetate, R²X, dry DMF, dry NEt₃, 3 Å molecular sieves, 8 h, rt (17%-84%); f) 1 M NaOMe, dry MeOH, 1 h, 0°C (88%); g) 0.1 M KOH, MeOH, 1 h, rt, ii: Amberlite IR-120 (H⁺), iii: aq. NaOH to pH 7; h) i: 0.1M KOH, MeOH, 1 h, rt, ii: Amberlite IR-120 (H⁺), iii: aq. NaOH to pH 7 (40-72%).

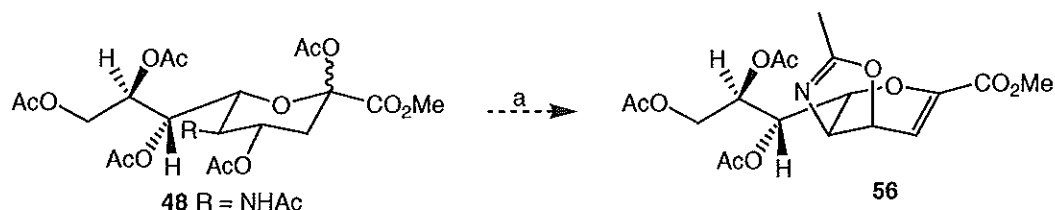
The first steps in the synthesis involved the esterification of the carboxylic acid group of Neu5Ac **1** followed by full *O*-acetylation. Esterification was accomplished by the treatment of Neu5Ac **1** with Amberlite IR-120 (H⁺) resin in dry MeOH, to afford the

methyl ester as a white solid in greater than 95% yield.¹⁶⁴ Per-*O*-acetylation of **1** was accomplished by treatment with pyridine in acetic anhydride at room temperature. Purification of the resultant brown solid yielded methyl 5-acetamido-2,4,7,8,9-penta-*O*-acetyl-3,5-dideoxy-D-*glycero*- α/β -D-*galacto*-non-2-ulopyranosonate (**48**) as a white foam in 98% yield, with the overall isolated yield (93%) of **48** from Neu5Ac **1** being comparable to that reported in the literature.¹⁷³

3.2.2 Synthesis of 2-methyl-(methyl 7,8,9-tri-*O*-acetyl-2,6-anhydro-3,5-dideoxy-D-*glycero*-D-*talo*-non-2-enonate)-[4,5-d]-2-oxazoline (**56**).

The Neu5Ac2en oxazoline, 2-methyl-(methyl 7,8,9-tri-*O*-acetyl-2,6-anhydro-3,5-dideoxy-D-*glycero*-D-*talo*-non-2-enonate)-[4,5-d]-2-oxazoline (**56**) has previously been prepared¹¹⁷ by the reaction of protected Neu5Ac2en, **49**, with boron trifluoride etherate (BF₃•Et₂O) at room temperature for 16 hours (96% yield) as shown in Scheme 3.1a, Section 3.1.2. It has also been prepared¹⁴⁷ by the reaction of fully protected Neu5Ac1Me, **48**, with trimethylsilyl trifluoromethanesulfonate (TMSOTf) (3 equivalents) in ethyl acetate at 52°C for 2.5 hours (61% yield) which both introduces the 2,3 double bond and forms the oxazoline (Scheme 3.1a, Section 3.1.2). Preparation of oxazoline **56** directly from Neu5Ac derivative **48** avoids the synthesis of the relatively unstable β -chloro-Neu5Ac1Me derivative **47** in the synthesis of Neu5Ac2en, and so was the method of choice in this work.

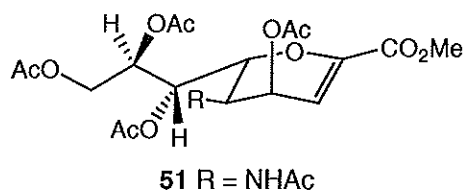
Given the reported synthesis of unsaturated oxazoline **56** directly from saturated per-*O*-acetylated Neu5Ac1Me **48** using TMSOTf, and the synthesis of the oxazoline from Neu5Ac2en derivative **49** using BF₃•Et₂O (Scheme 3.1), it was of interest to examine the effect of the alternate Lewis acid BF₃•Et₂O on protected Neu5Ac1Me **48**. The reaction conditions used for the BF₃•Et₂O reaction were as reported by von Itzstein *et al.*¹¹⁷ for the preparation of oxazoline **56** from protected Neu5Ac2en, **49**. The peracetylated methyl ester of Neu5Ac, **48**, was stirred with 10 equivalents of BF₃•Et₂O in a mixture of dry MeOH and dry DCM for 16 hours at ambient temperature (Scheme 3.5). After this time, the reaction was worked-up by addition of cold aqueous sodium hydrogen carbonate and extraction into EtOAc. As will be seen, the nature of the work-up procedure has a significant effect on the product isolated.



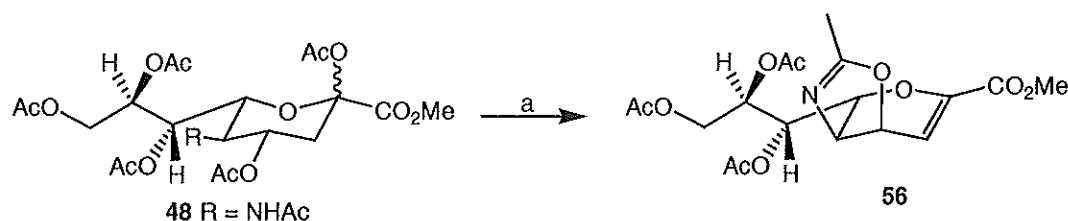
Scheme 3.5

Reagents & Conditions. a) $\text{BF}_3 \cdot \text{Et}_2\text{O}$, dry DCM, dry MeOH, 16 h, rt.

The reaction using $\text{BF}_3 \cdot \text{Et}_2\text{O}$ described above yielded two compounds, one in major yield (approximately 1:0.2 by ^1H NMR analysis). NMR analysis of the product mixture showed that the formation of the double bond had occurred in the major product as indicated by the resonance for an olefinic proton (H-3) at δ 6.08. However the coupling constant of H-3 (of the major product) was 5.7 Hz, not the expected coupling of 3.9 Hz previously reported¹¹⁷ for the oxazoline derivative **56**. The H-3 coupling constant of 5.7 Hz was more consistent with coupling to an equatorial proton at C-4 as in 4-*epi*-Neu4,5,7,8,9Ac₅en1Me **51**.¹⁷⁴ The minor product, that also showed an olefinic signal (δ 5.95, J 3.3 Hz), was tentatively identified as the desired product **56**.



As the first use of $\text{BF}_3 \cdot \text{Et}_2\text{O}$ failed to produce the desired product, the procedure of Chandler *et al.*¹⁴⁷ using TMSOTf was used to form the oxazoline **56** (Scheme 3.6). The fully protected Neu5Ac1Me derivative **48** was stirred in EtOAc at 30°C, TMSOTf was added and the temperature was raised to 52°C and the reaction was left to stir for 2.5 hours. Successful formation of oxazoline **56** was indicated in the ^1H NMR spectrum by the lack of the NH proton, loss of an acetate group methyl signal, and the olefinic proton signal at δ 6.37 with a coupling constant of 3.9 Hz (H-3). However the purity of the product and isolated yields of the oxazoline after work-up and chromatography (EtOAc) varied between 70 to 85% depending on the work-up procedure used. The work-up procedure of Chandler *et al.*¹⁴⁷ utilised ice-cold saturated sodium bicarbonate, followed by extraction into EtOAc. ^1H NMR analysis of the resultant product showed the formation of at least two compounds, with the oxazoline **56** as the major component (approximately 1:0.2 by ^1H NMR analysis). In an attempt to increase the isolated yield of **56**, a number of different work-up procedures were trialled.



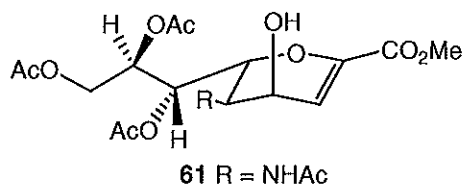
Scheme 3.6

Reagents & Conditions. a) TMSOTf, dry EtOAc, 2 h, 53°C (85%).

The second work-up procedure trialled was based on that used by Yu *et al.*¹⁷⁵ for the TMSOTf catalysed introduction of a 1,2-oxazoline into per-*O*-acetylated chitobiose. In place of saturated sodium bicarbonate, the reaction was neutralised with a slight excess of triethylamine (NEt₃), before evaporation to dryness. Isolated yield after work-up was comparable to that from the cold aqueous NaHCO₃ work-up procedure, with **56** still mixed with a second component. Dry NEt₃ was trialled next based on the hypothesis that the oxazoline when in solution is highly susceptible to acid attack, and water being a weak acid causes the formation of the unknown product. The reaction mixture was evaporated after neutralisation with dry NEt₃. The ¹H NMR of the product in this case showed pure oxazoline **56** with no other product visible. However experimental yield was in excess of 100%, indicative of the presence of triethyl ammonium salts. Two methods for salt extractions were then tried. In the first method, extraction into EtOAc and washing the EtOAc layer with distilled water resulted in partial conversion of the oxazoline to the unknown compound, as shown by ¹H NMR. A mild base extraction was trialled to counteract the effect of slightly acidic nature of water on the oxazoline. Mild base (0.1 M NaHCO₃, pH 8.2) was selected so as not to remove the *O*-acetate groups. Using this washing procedure, the triethyl ammonium salts were successfully removed, with minimal conversion of the oxazoline to the unknown product (below the level of ¹H NMR detection). After extraction of the ethyl acetate fraction with 0.1 M NaHCO₃ solution, the oxazoline **56** was obtained in 85% yield. The ¹H NMR data obtained for **56** corresponded to that reported in the literature.¹⁴⁷

Acid-catalysed hydrolysis of the oxazoline **56** would be expected to lead to the formation of 4-*epi*-hydroxy Neu5Ac2en1Me derivative **61**.¹⁷⁶ The formation of **61** may explain the presence of the unknown product seen in the work-up of the oxazoline reactions. The previously reported NMR data for **61**, recorded in CDCl₃,¹⁷⁷ could be used to check for the presence of **61** in the oxazoline reactions. NMR spectra of the

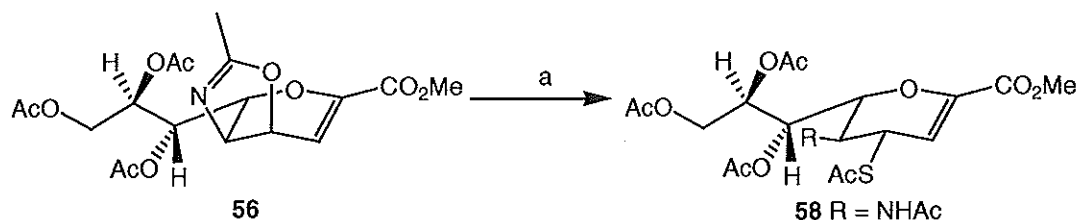
by-product from a work-up of an TMSOTf reaction, and the product from the initial $\text{BF}_3 \cdot \text{Et}_2\text{O}$ reaction, were compared with the spectrum of **61**.¹⁷⁷ This analysis suggested that the unknown product from the oxazoline aqueous base work-up and the major product from the $\text{BF}_3 \cdot \text{Et}_2\text{O}$ reaction are the same compound, **61**.



With a successful method for the clean isolation of the oxazoline **56**, the $\text{BF}_3 \cdot \text{Et}_2\text{O}$ reaction was tried again using the dry NEt_3 work-up procedure. Reaction conditions were not changed from those of the first attempt (EtOAc , 30-52°C). With the dry NEt_3 work-up, NMR analysis now showed the formation of the desired oxazoline **56**, with minimal conversion to 4-*epi*-hydroxy Neu5Ac2en1Me derivative, **61**. In summary, both TMSOTf and $\text{BF}_3 \cdot \text{Et}_2\text{O}$ could be used as Lewis acids to promote conversion of per-*O*-acetylated Neu5Ac1Me **48** to oxazoline **56**. Careful workup using dry NEt_3 for neutralization eliminated the formation of the by-product identified as the 4-*epi*-hydroxy derivative **61** and allowed the isolation of oxazoline **56** in 85% yield. Use of cold aqueous NaHCO_3 during work-up however led to formation of significant amounts of **61** (up to 85% of isolated product).

3.2.3 Synthesis of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,5-dideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, (**58**).

The introduction of the thiolacetate group at C-4 of Neu5Ac2en was carried out according to the procedure of von Itzstein *et al.*¹¹⁷ The C-4-thiolacetate derivative, **58**, was formed by treatment of oxazoline **56** with thiolacetic acid in anhydrous *N,N*-dimethylformamide at 83°C for 20 hours (Scheme 3.7). These reaction conditions resulted in the formation of the key intermediate, methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,5-dideoxy-4-thio-D-glycero-D-galacto-non-2-enonate (**58**) in 49% yield after purification. Successful formation of **58** was indicated by the presence of the characteristic *S*-acetyl methyl resonance at 2.35 ppm¹¹⁷ in the ^1H NMR spectrum. Equatorial placement of the thiolacetate group at C-4 was indicated by the $J_{3,4}$ coupling constant of 2.4 Hz (H-4: δ 4.37 ppm).¹¹⁷

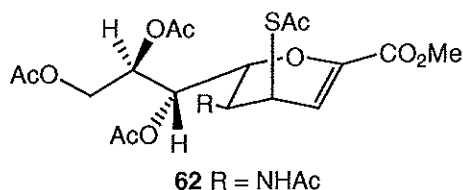


Scheme 3.7

Reagents & Conditions. a) Thiolacetic acid, dry DMF, 20 h, 83°C (49%).

On scale-up of the synthesis of **58** however, ^1H NMR analysis of the product showed two products with thiolacetate resonances in a ratio approx 1:0.3 as judged by integration of both the SAc and H-3 resonances in the ^1H NMR spectrum. A number of solvent systems were evaluated to try to separate and resolve the two thiolacetate derivatives; however no trialled solvent system could resolve the compounds. The similarity of the R_f values of the components meant they were very difficult to separate by column chromatography. The best purity of the desired product **58** was obtained using 20:1 DCM/MeOH as the column eluent, with the isolated yield at 49% with approximately 15% of the minor product (1:0.17 ratio by ^1H NMR).

The thiolacetate region of the NMR spectrum of the mixture isolated after chromatography showed the resonance for the thiolacetate methyl group of the major product **58** at 2.34 ppm, and that of the minor product at 2.36 ppm. To examine whether the minor resonance was residual thiolacetic acid, a drop of thiolacetic acid was added to a NMR sample. Subsequent analysis of this sample showed three thiolacetates, the previously identified major and minor thiolacetate methyl signals and the third thiolacetate methyl signal that resonated at 2.38 ppm, as expected for thiolacetic acid. The olefinic proton of the minor product (δ 6.05) had a coupling constant of 5.7 Hz, characteristic of a C-4 axially substituted derivative of Neu5Ac2en. From the ^1H NMR spectrum for the minor product, this species could be assigned as per-*O*-acetylated 4-*epi*-SAc Neu5Ac2en1Me **62**.



Under the standard reaction conditions for the formation of the C-4 thiolacetate derivative from oxazoline **56**, (dry DMF, 5 equivalents thiolacetic acid, 83°C, 20 h) approximately 15 to 23% of the 4-*epi*-SAc Neu5Ac2en derivative **62** was found to be

formed. The variance in the amount of **62** present may be due to partial removal by column chromatography. In an attempt to minimise the amount of epimer formed, a number of other reaction conditions were trialled, as summarised in Table 3.2.

Table 3.2: Alternative reaction conditions^a for formation of 4-*S*-acetyl-Neu5Ac2en1Me derivative **58** from oxazoline **56**.

Solvent	Molar equivalents	Product ratio ^b	Isolated combined
	Thiolacetic acid to 56	4- <i>eq</i> 58 :4- <i>epi</i> 62	yield of 58 and 62 ^c
DMF	5	1:0.17	49%
DMF	1.5	1:0.10	85%
Acetonitrile	1.5	1:0.17	47%
Thiolacetic acid	excess	1:0.24	89%

^a Reactions were carried out at 83°C for 20 hours.

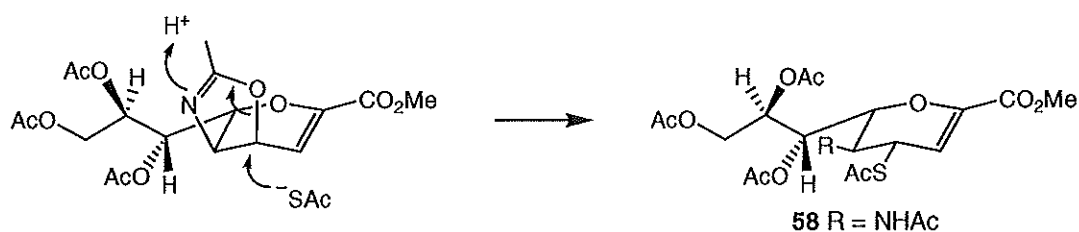
^b Epimer ratio determined from the ratio of H-3 signals in 300 MHz ¹H NMR spectrum of reaction product after the evaporation of reaction solvent.

^c After column chromatography.

The best reaction conditions involved the use of DMF as the reaction solvent with a reduced number of equivalents (1.5) of thiolacetic acid (Table 3.2, entry 2). These reaction conditions gave the best results in terms of an excellent isolated yield of product, and a reduced amount of 4-*epi*-SAc (down to 9% from 15%) compared to reaction in the presence of 5 equivalents of thiolacetic acid (Table 3.2, entry 1).

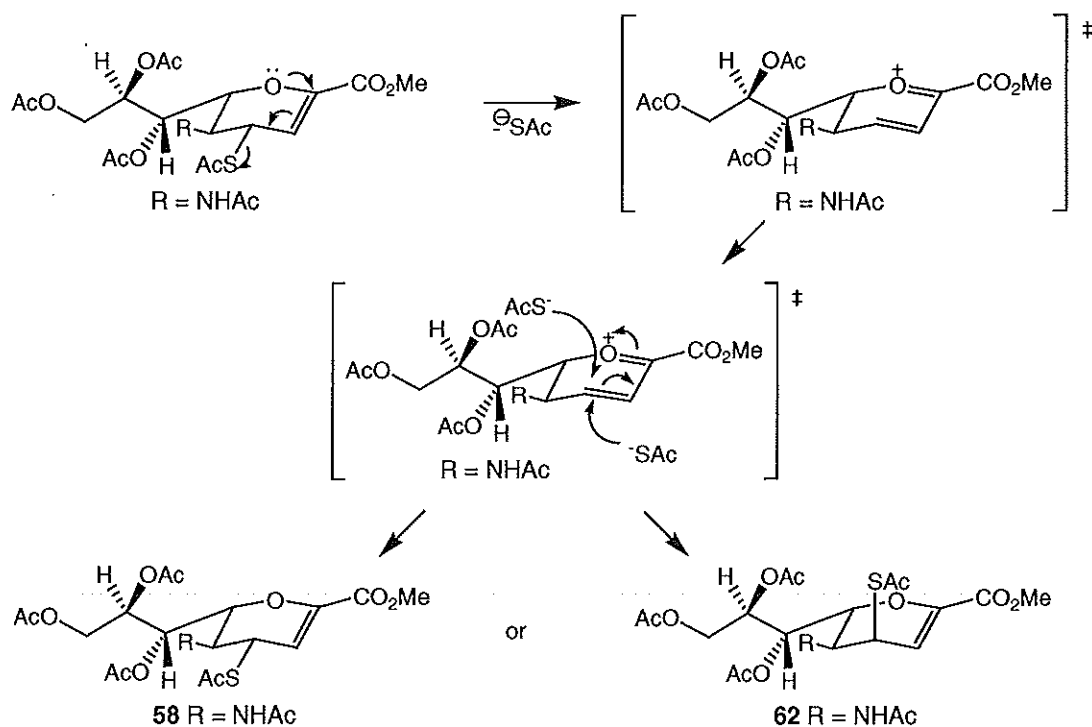
Alternate solvents for the reaction were investigated. DMF was replaced by acetonitrile as the reaction solvent (Table 3.2, entry 3) based on similarities in dielectric constant.¹⁷⁸ The replacement of DMF with acetonitrile and lowering the number of equivalents of thiolacetic acid gave a similar isolated yield and epimer ratio to the original reaction conditions. The use of neat thiolacetic acid as the reaction solvent (Table 3.2, entry 4) caused an increase in the formation of the 4-*epi*-SAc derivative **62** to approx 19%. However, the combined isolated yield of **58** and **62** also increased to 89%, although the resultant isolated product was slightly less pure than when using DMF as solvent.

The formation of 4-*eq*-SAc Neu5Ac2en1Me derivative **58** can occur by direct S_N2 attack of thiolacetic acid at C-4 of oxazoline **56** (Scheme 3.8).



Scheme 3.8

A postulated mechanism for the formation of 4-*epi*-SAc-Neu5Ac2en1Me **62** is shown in Scheme 3.9. The C-4 position of per-*O*-acetylated Neu5Ac2en1Me is known to undergo loss of the C-4 acetate under Lewis acidic conditions (Section 3.2.2). Under the acidic conditions of the thiolacetylation reaction, it may be possible that the loss of the C-4 thiolacetate could occur with assistance from a resonance structure in which the endocyclic oxygen electron lone pair creates a double bond to C-1. The resultant planar structure could then be attacked on either face by a further thiolate ion.



Scheme 3.9

For simplicity in future sections the chemical structures of C-4-thioether derivatives of Neu5Ac2en will be represented in the C-4 equatorial form.

3.2.4 Hydrazine acetate mediated C-4 thioether formation.

From the molecular modelling study (Section 2.3.7.3), nine 4-*S*-substituted Neu5Ac2en derivatives were selected to be prepared (Figure 3.1): 4-*S*-benzyl- **28-A**, 4-*S*-(2-phenyl)benzyl- **28-B**, 4-*S*-ethyl- **28-C**, 4-*S*-hexyl- **28-D**, 4-*S*-decyl- **28-E**, 4-*S*-*p*-

cyanobenzyl- **28-F**, 4-*S*-(2-phenyl)ethyl- **28-G**, 4-*S*-cyanomethyl- **28-H**, and 4-*S*-*p*-nitrobenzyl- **28-I** Neu5Ac2en.

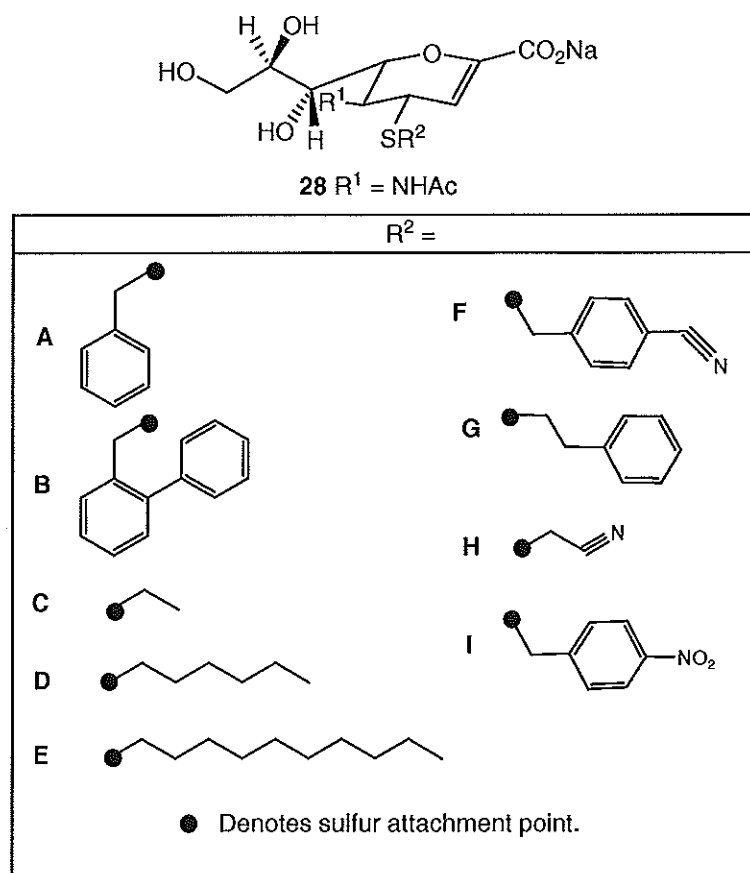
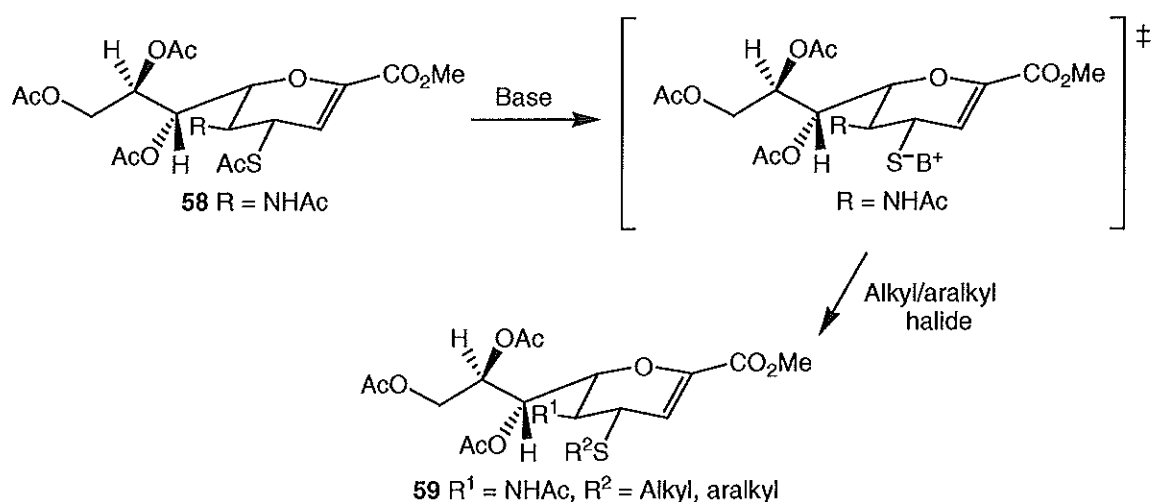


Figure 3.1: Structure of the nine 4-*S*-substituted Neu5Ac2en derivatives to be synthesised.

The derivatives shown in Figure 3.1 were prepared from 4-thiolacetyl Neu5Ac2en1Me derivative **58** in a two step one-pot synthesis involving the selective de-*S*-acetylation of the C-4 thiolacetate group, followed by the coupling of the required alkyl or aralkyl halide with the resultant thiolate ion (Scheme 3.10). De-*S*-acetylation of thiolacetate groups has been reported in the literature using hydrazine acetate,¹⁷¹ sodium methoxide,¹⁷⁹ or diethyl amine.¹⁷⁰ Sodium methoxide has been reported to have low selectivity for *S*-acetates over *O*-acetates.¹⁷⁹ While diethyl amine has been reported to be milder than both sodium methoxide and hydrazine acetate it has been found to be ineffective in the de-*S*-acetylation of non-anomeric thiolacetates.¹⁷² Hydrazine acetate has been used to de-*O*-acetylate the anomeric position on carbohydrates, and to de-*S*-acetylate anomeric¹⁷¹ or primary thioesters.¹⁷¹ Chemoselectivity (*S*Ac vs *O*Ac) at ambient temperatures over prolonged reaction times has also been observed.¹⁷¹ Thus hydrazine acetate was selected as the de-*S*-acetylating reagent.



Scheme 3.10

Coupling reactions were first attempted with benzyl bromide, using the benzyl group as a representative of the desired substituents to be introduced at C-4. Initial reaction conditions trialled were those reported in literature¹⁷¹ for primary thiolacetates. Accordingly the 4-SAc derivative **58** was reacted with 1 equivalent of hydrazine acetate in DMF at room temperature for 45 minutes, followed by the addition of 1.5 equivalents of benzyl bromide and 1 equivalent of NEt₃ and reaction for 3 hours at room temperature. ¹H NMR of the crude product isolated after evaporation of reaction solvent however, showed mostly recovered **58** indicating that the initial de-S-acetylation had not occurred to any appreciable extent.

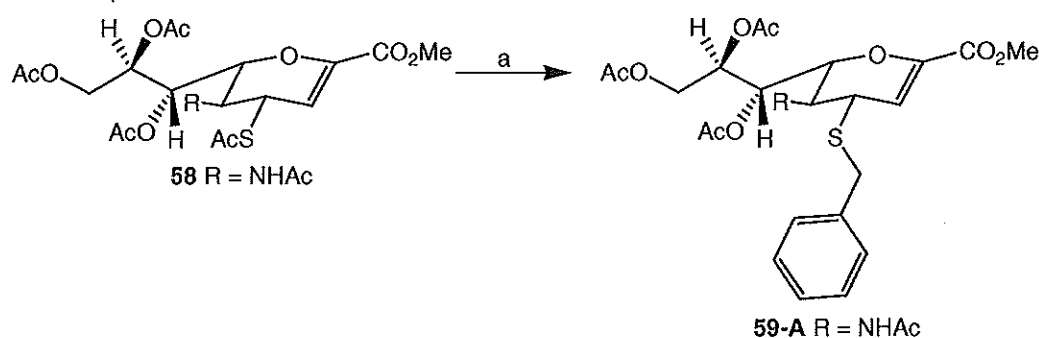
Modifications to the initial procedure to attempt to improve thiolate formation included the addition of 4 equivalents of hydrazine acetate, a longer reaction time (48 hours), and an increase in temperature (40°C) as summarised in Table 3.3. These reaction modifications did not produce the desired coupled material. It appeared that de-S-acetylation was not occurring as unreacted starting material was consistently observed.

Table 3.3: Summary of coupling reaction conditions trialled for synthesis of 4-S-benzyl Neu5Ac2en1Me derivative **59-A** from **58** (Scheme 3.11).

Hydrazine Acetate equiv to 58	BnBr equiv. to 58	Solvent	Temp	Time	Outcome
1 (DMF solution)	1.5 BnBr	DMF	rt	3 h	NR ^a
4 (DMF solution)	1.5 BnBr	DMF	40°C	2 d	NR
7 solid	1.5 BnBr	DMF	rt	8 h	68% 59-A

^a NR = No visible reaction by TLC

Formation of the desired C-4 thioether derivative using benzyl bromide as coupling partner was successfully achieved under the following conditions. The coupling partners were initially stirred with 3 Å molecular sieves in DMF for a minimum of 10 minutes. This was followed by the addition of 7 equivalents of solid hydrazine acetate (as opposed to a DMF solution) to the reaction mixture. These changes were made to minimise any water present in starting materials/reagents. The reaction was then left to stir at room temperature for 8 hours, after which ^1H NMR and TLC analysis indicated the successful formation of the 4-*S*-benzyl Neu5Ac2en1Me derivative **59-A** (Scheme 3.11). ^1H NMR analysis of the reaction product indicated a distinctive shift of the resonance for H-4, from δ 4.37 in thiolacetate derivative **58** to δ 3.47 for the successfully coupled 4-*S*-benzyl derivative **59-A**. The C-5 proton was also seen to shift from δ 4.20 in **58** to δ 4.02 in **59-A**. Successful formation of **59-A** was also confirmed by ESI LRMS in positive ion mode with the observation of an $(\text{M}+\text{Na})^+$ peak at 560 m/z for the product.



Scheme 3.11

Reagents & Conditions. a) 7 equiv. hydrazine acetate, dry DMF, 1.5 equiv. dry NEt_3 , 1.5 equiv BnBr , 3 Å molecular sieves, 8 h, rt, (68%).

While TLC analysis of the product isolated from the benzyl bromide coupling appeared to show only one compound, NMR analysis indicated two compounds (1:0.2) were present. The product produced in major yield had an olefinic proton (H-3) coupling of 2.7 Hz and the minor product had a coupling of 5.6 Hz. This is consistent with a mixture of the 4-equatorial and 4-*epi-S*-benzyl derivatives. NMR analysis of the C-4 thiolacetate starting material **58** established that it was composed of a 1:0.2 mixture of C-4 epimers; after the coupling reaction no detectable increase in the amount of the C-4 *epi* form was seen.

The following optimised coupling reaction conditions were used to synthesise the Neu5Ac2en C-4 thioether derivatives **59-A** to **59-I**: 7 equivalents hydrazine acetate, 1.5

equivalents NEt_3 , 1.5 equivalents alkyl/aralkyl halide, anhydrous DMF, 3 Å molecular sieves, and reaction for 8 hours at room temperature. The yields of the products isolated after chromatography are given in Table 3.4. In all cases good conversion of starting material **58** to the desired product was observed by TLC of the crude reaction product. However, the purification of the products derived from bromodecane (**59-E**), bromoacetonitrile (**59-H**) and 4-nitrobenzyl bromide (**59-I**) led to lower isolated yields than seen with the other coupling partners.

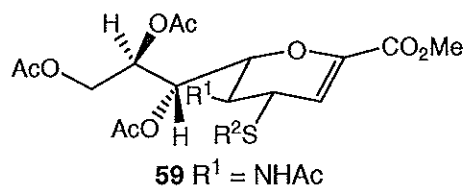
Table 3.4: Yields of hydrazine acetate mediated coupling reactions of 4-SAc Neu5Ac2en derivative **58** with various alkyl/aralkyl halides.

Alkyl halide	Product	Reaction yield ^a
benzyl bromide	59-A	68%
2-phenylbenzyl bromide	59-B	68%
iodoethane	59-C	61%
bromohexane	59-D	84%
bromodecane	59-E	39%
α -bromo- <i>p</i> -tolunitrile	59-F	80%
2-bromoethylbenzene	59-G	59%
bromoacetonitrile	59-H	17%
4-nitrobenzyl bromide	59-I	27%

^a Represents best isolated yield for coupling reaction.

Successful coupling of the alkyl/aralkyl halides was indicated in ^1H NMR analysis of the coupled products by characteristic changes in the chemical shifts of H-4 and H-5. Selected ^1H NMR data for the coupled products is given in Table 3.5. In all cases, H-4 is observed to shift upfield by $\sim 0.9 - 1$ ppm from the shift observed for the 4-SAc derivative **58** (H-4: 4.37 ppm). This is consistent with the loss of the deshielding acetate group on the sulfur atom. A smaller upfield shift was observed in the case of the cyanomethyl derivative **59-H**, which may be due to the triple bond of the cyano group deshielding the H-4 proton.¹⁴⁴ H-5 is also observed to shift upfield by between 0.2 and 0.4 ppm depending on the substituent attached, this may be due to a through space effect of substituents at O-4. No great difference is seen in the chemical shift of H-4 and H-5 between each of the synthesised derivatives. All compounds gave the expected mass for the $(\text{M}+\text{Na})^+$ ion in positive ion mode ESI LRMS.

Table 3.5: Selected data indicating successful coupling of alkyl/aralkyl halides with **58**.

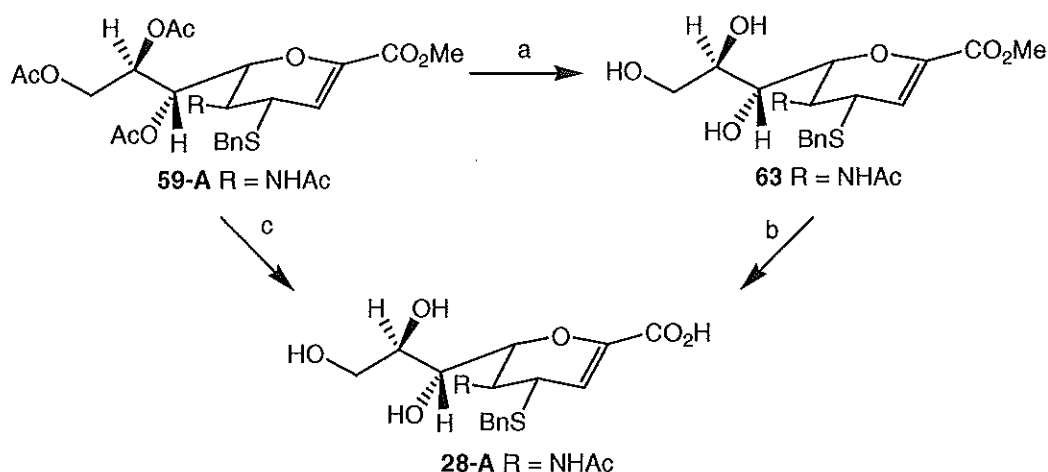


Compound (R^2)	Compound	R_f^a	H-4 ^c	H-5 [#]
acetyl	58	0.38	4.37	4.20
benzyl	59-A	0.62 ^b	3.47	4.02
2-phenylbenzyl	59-B	0.54	3.37	3.78
ethyl	59-C	0.46	3.47	3.99
hexyl	59-D	0.62	3.40	3.87
decyl	59-E	0.60	3.48	3.92
<i>p</i> -cyanobenzyl	59-F	0.48	3.38	4.00
2-phenylethyl	59-G	0.56	3.45	4.00
cyanomethyl	59-H	0.38	3.79	4.05
<i>p</i> -nitrobenzyl	59-I	0.49	3.41	4.04

^a TLC (2:1 EtOAc/Hex); ^b TLC (EtOAc, SM R_f = 0.51); ^c ¹H NMR 300 MHz.

3.2.5 Deprotection of C-4-thioether derivatives of Neu5Ac2en

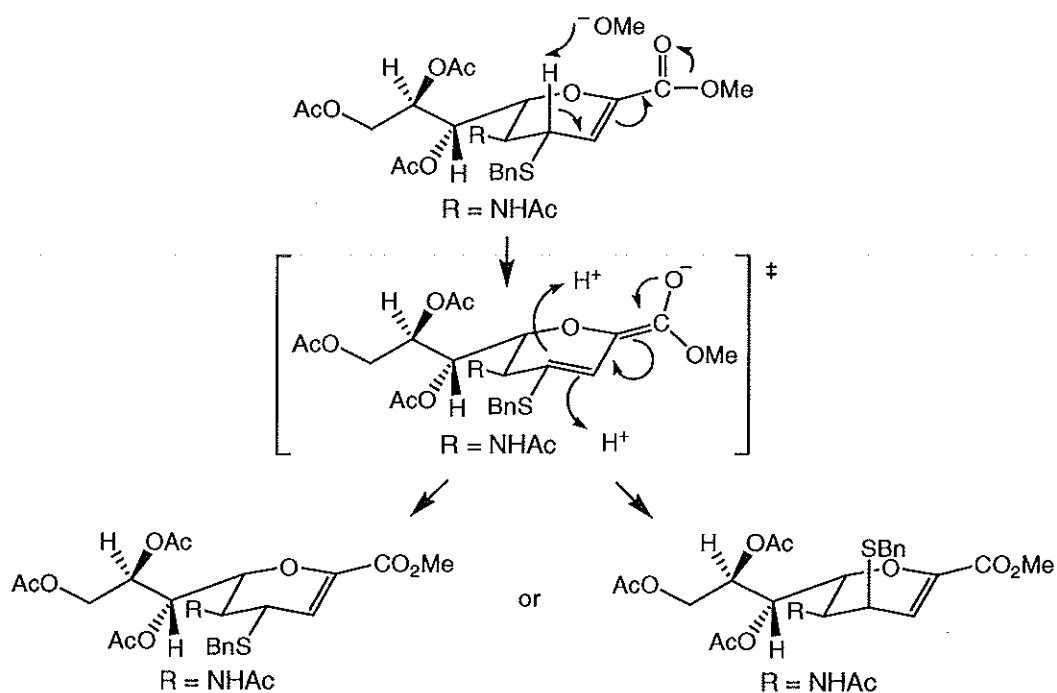
The deprotection strategy initially chosen for the C-4-thioether derivatives of Neu5Ac2en was a two step strategy; de-*O*-acetylation followed by de-esterification as shown in Scheme 3.12 (steps a and b), a standard deprotection strategy in sialic acid chemistry. Initially for the 4-*S*-benzyl derivative **59-A**, de-*O*-acetylation was performed using catalytic NaOMe in dry MeOH. When the de-*O*-acetylation reaction was conducted at room temperature, NMR analysis of the crude product, **63**, isolated after neutralisation and evaporation of the reaction solvent showed an increase in the proportion on the 4-*epi* derivative from 15% in the starting material to 39% in the product, based on integration of the H-3 resonance. To examine whether the temperature of the reaction affected the extent of epimerisation, the reaction was repeated at 0°C. The proportion of the 4-*epi* derivative in the product after reaction at 0°C was reduced (compared to reaction at room temperature) to 24% of the total product. At lower temperatures (-20°C and -10°C) the de-*O*-acetylation reaction did not proceed.



Scheme 3.12

Reagents & Conditions. a) cat. NaOMe, dry MeOH, 1 h, 0°C; b) 0.1 M KOH, MeOH, 1h, rt; c) 0.1 M KOH, MeOH, 1h, rt.

A mechanism proposed to explain the increased proportion of the 4-*epi* derivative in the 4-*S*-benzyl derivative **63** after reaction of **59-A** with NaOMe is given in Scheme 3.13. It is postulated that the reaction of **59-A** with NaOMe would be conducive to the deprotonation of C-4 and the resultant anion could be stabilised by a resonance structure through to the methyl ester. Re-protonation at C-4 could happen on either face of the ring and as such give rise to a greater proportion of the 4-*epi* derivative in the product than in the starting material.



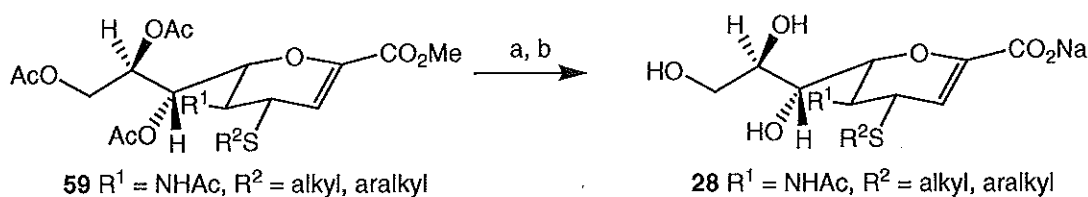
Scheme 3.13

Proposed reaction mechanism for the NaOMe mediated formation of 4-*S*-*epi*-benzyl-Neu5Ac2en1Me.

3.2.5.1 Deprotection of C-4-thioether Neu5Ac2en derivatives with potassium hydroxide.

A revised strategy for deprotection of the C-4-thioether Neu5Ac2en derivatives involving the removal of both the *O*-acetate and the methyl ester protecting groups with a 1:1 mixture of 0.1 M KOH and MeOH at 0°C was tried. Deprotection of the methyl ester under these conditions would place a formal negative charge at C-1 and therefore make less likely the formation of the resonance structure shown in Scheme 3.13. The revised reaction conditions were tried for the deprotection of the 4-*S*-benzyl-Neu5Ac2en1Me (**59-A**) derivative. ¹H NMR analysis of the fully deprotected 4-*S*-benzyl derivative **28-A** using these conditions showed that there was very little (if any) increase in the proportion of the 4-*epi* derivative compared to that seen in the starting material.

All C-4-thioether derivatives of Neu5Ac2en, with the exception of the 4-nitrobenzyl and 4-cyanomethyl derivatives **59-I** and **59-H** (due to poor coupling reaction yields) were deprotected using 0.1 M KOH and MeOH (Scheme 3.14). The deprotected materials were acidified to pH 3-4 with ion exchange resin, the resin was filtered off, and the solution was neutralised to pH 7.8 with NaOH solution to form the sodium salts. Yields of the desired fully deprotected products isolated after the formation of the sodium salt were greater than 90%.



Scheme 3.14

Reagents & Conditions. a) 0.1 M KOH, MeOH, 1h, rt; b) i: Amberlite IR-120 (H^+) to pH 3-4; ii: 1M NaOH to pH 7.8.

Successful deprotection was confirmed by ¹H NMR analysis which showed the loss of the *O*-acetate resonances ($3 \times 3\text{H}$ singlets, region of δ 2.00 to δ 2.30), and loss of the methyl ester resonance ($1 \times 3\text{H}$ singlet, $\sim \delta$ 3.70). In Figure 3.2, which shows the C-4-thioether Neu5Ac2en derivatives synthesised, the general structure is presented in equatorial form for simplicity, however in each case a C-4 epimeric mixture with 5 to 22% of the 4-*epi*-derivative (see Table 3.9, Section 3.3.5) was present. The deprotected

S-linked derivatives **28-A** to **28-G** were purified using semi-preparative reverse phase high performance liquid chromatography (HPLC). Details of the chromatography are shown in Table 3.9, Section 3.3.5 at the end of this chapter.

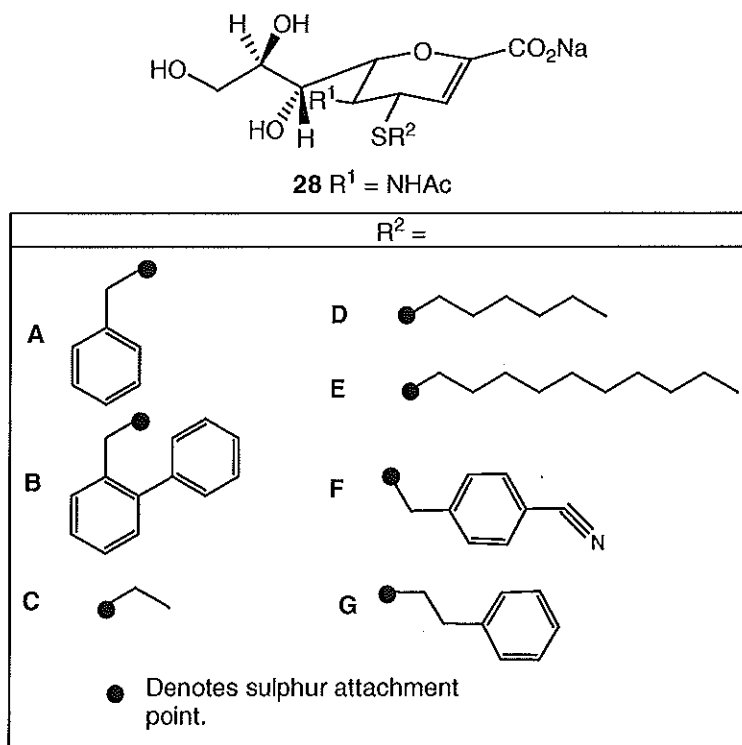
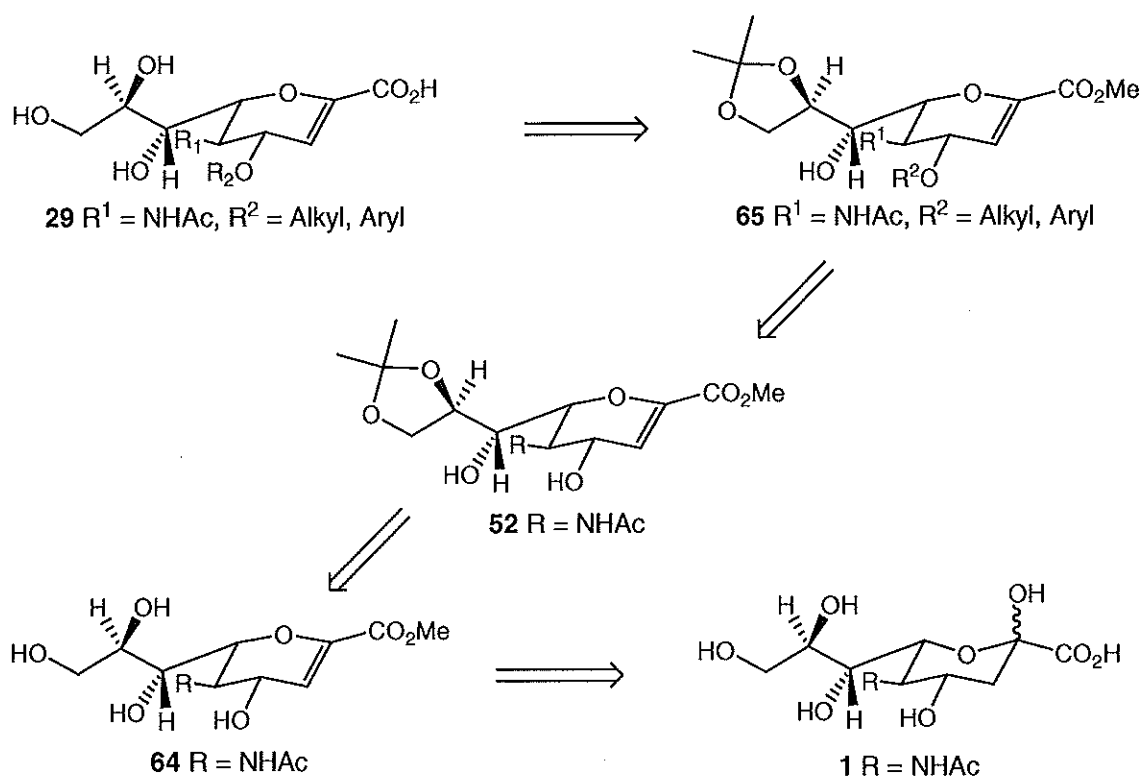


Figure 3.2: Structure of the C-4-thioether linked derivatives of Neu5Ac2en synthesised.

3.3 Approaches to C-4-ether Neu5Ac2en derivatives

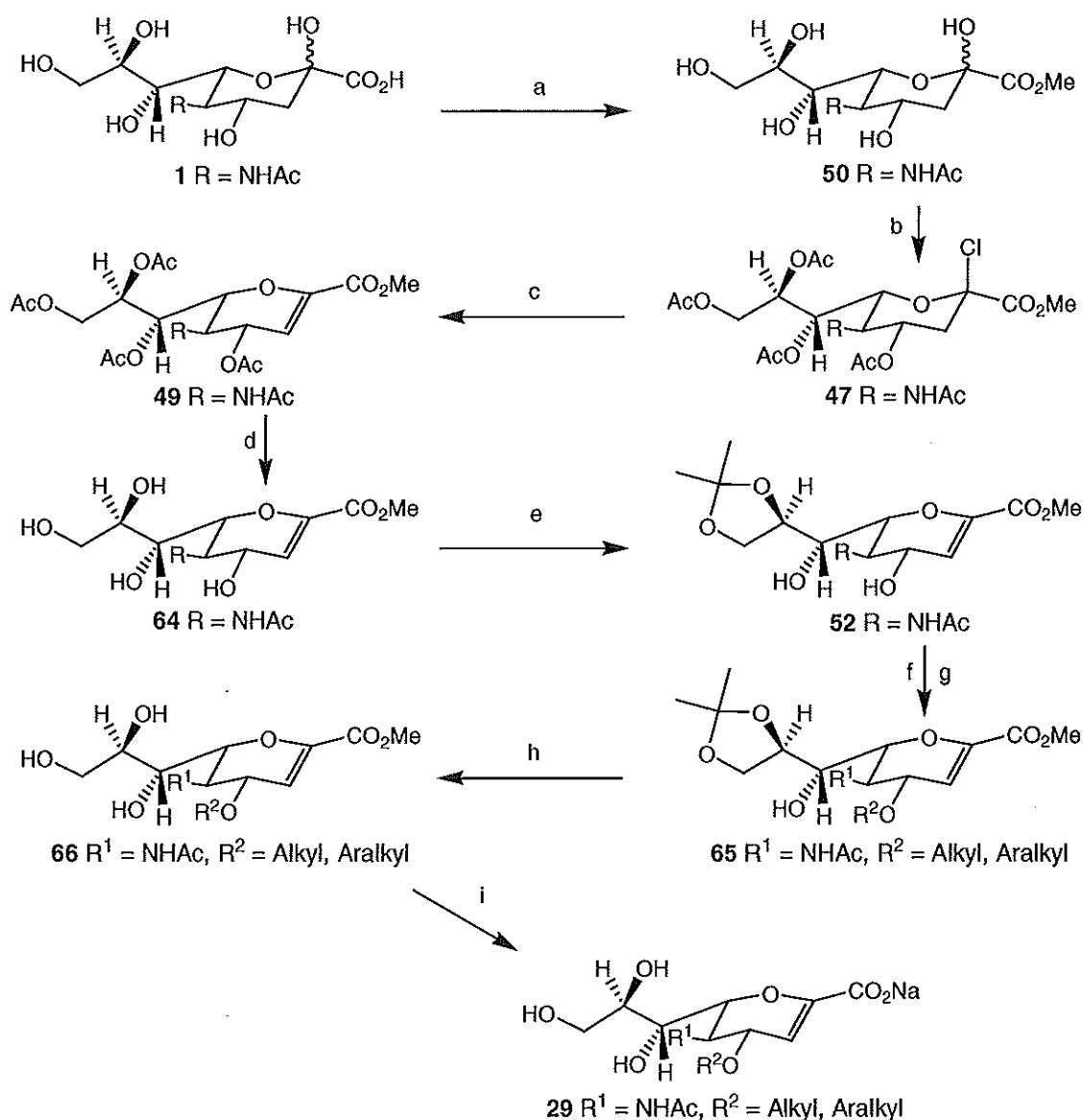
A retro-synthetic approach to the preparation of C-4-ether derivatives of Neu5Ac2en **29** is shown in Scheme 3.15. The alkylation of the C-4 hydroxyl group of a Neu5Ac2en derivative has been successfully performed in the presence of sodium hydride.¹¹⁸ The use of silver (I) oxide to alkylate at C-4 of saturated Neu5Ac derivatives has also been reported in the literature.^{144, 180} This reagent may also be of use in the alkylation of Neu5Ac2en derivatives. In either case, the coupling reaction has been performed on differentially protected Neu5Ac2en (NaH) or Neu5Ac (Ag_2O) derivatives. In the retro-synthesis of 4-*O*-alkylated Neu5Ac2en derivatives **29** presented in Scheme 3.15, the key intermediate proposed for both coupling strategies is methyl 5-acetamido-2,6-anhydro-3,5-dideoxy-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate (**52**). This compound has previously been synthesised from Neu5Ac2en1Me **64**¹¹⁸ which has in turn been synthesised from Neu5Ac, **1**, in four steps *via* the β -chloride of per-*O*-acetylated Neu5Ac1Me, **47**.^{134, 136}



Scheme 3.15: Retrosynthetic approaches to C-4-ether derivatives of Neu5Ac2en.

3.3.1 Synthesis of C-4-ether linked Neu5Ac2en derivatives

The approach taken to the preparation of the C-4-ether Neu5Ac2en derivatives is shown in Scheme 3.16. An advantage of the C-4-ether derivatives is that they can be prepared epimerically pure, in comparison to the C-4-thioether derivatives that contained ~5 to 22% of the C-4 epimer. Another advantage of synthesising the C-4-ether linked Neu5Ac2en derivatives is that they would allow assessment of the effect of the C-O or C-S linkage at C-4 on biological activity.



Scheme 3.16

Reagents & Conditions. a) dry MeOH, Amberlite IR-120 (H^+) resin, 1 d (> 95%); b) AcCl, 2 d, rt (~98%); c) DBU, dry DCM, 2.5 h, rt (74%); d) 1 M NaOMe. Dry MeOH, 1 h, 0°C (92%); e) 2,2-dimethoxypropane, Amberlite IR-120 (H^+) resin, dry acetone, 16 h, rt (100%); f) R^2X , dry DCM, Ag_2O , $n\text{-Bu}_4\text{NI}$, 3 Å molecular sieves, 2 d, rt (0-82%); g) R^2X , dry DMF, NaH, 8 h, rt, Ar (0-65%); h) 80% AcOH, 0.5 h, ~83°C (84-> 100%); i) i: 0.1M KOH, MeOH, 1 h, rt, ii: Amberlite IR-120 (H^+), iii: aq. NaOH to pH 7.8, (35-64%).

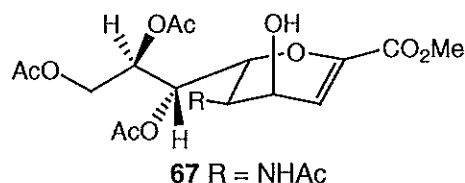
3.3.2 Preparation of key intermediate 8,9-*O*-isopropylidenated

Neu5Ac2en1Me **52**.

The key intermediate in Scheme 3.16 is methyl 5-acetamido-2,6-anhydro-3,5-dideoxy-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate (**52**). The preparation of **52** has been described in the literature¹¹⁸ in 5 steps from Neu5Ac, **1**. The methyl ester of Neu5Ac **1**, as described in Section 3.2.1, was synthesised in excellent

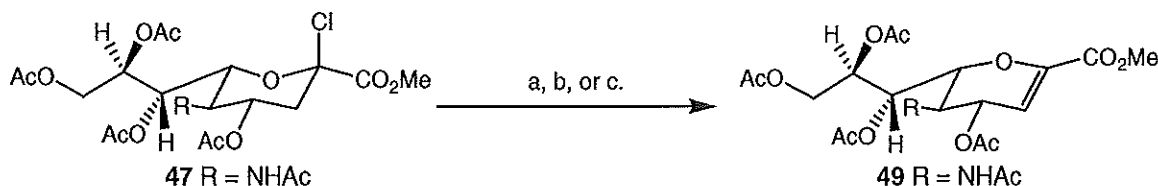
yield (> 95%) using standard conditions (MeOH, acidic resin, rt).¹⁶⁴ Chlorination and per-*O*-acetylation to give 2- β -chloro-Neu5Ac1Me **47** was performed by stirring Neu5Ac1Me **50** in neat AcCl at room temperature for 2 days.¹³³ Once complete, the reaction was worked-up by co-evaporation with toluene and the crude product was used without purification in the subsequent β -elimination step. A crude yield of ~98% was obtained.

The formation of per-*O*-acetylated Neu5Ac2en1Me **49** from the per-*O*-acetylated β -chloro-Neu5Ac1Me derivative **47** has been reported using a number of different bases to eliminate HCl and form the unsaturated product **49**.¹³⁴⁻¹³⁶ In this work, the elimination of HCl from the β -chloride was tried a number of different ways in an attempt to maximise the yield (Scheme 3.17, Table 3.6). The formation of the C2-C3 double bond by reacting **47** in dry pyridine at 50°C was reported by Sun *et al.*¹³⁴ (overall yield of **49** from **1**, 97%). Following this method, crude **47** was dissolved in pyridine and stirred at 50°C for 1 hour. After purification, ¹H NMR analysis showed the presence of an olefinic resonance (H-3) at 6.02 ppm, confirming the formation of the C2-C3 double bond. In our hands, the overall reaction yield for the conversion of Neu5Ac1Me **50** to per-*O*-acetylated Neu5Ac2en1Me **49** over two steps was ~55% (well below that reported¹³⁴ in the literature). In an effort to increase the reaction yield, an alternate method of elimination was used involving the reaction of the β -chloro derivative **47** with triethylamine in dioxane at 80°C for 30 minutes under Ar.¹³⁵ Overall yield for the conversion of methyl ester **50** to **49** over the two steps was 74%, which was an improvement over the pyridine-catalysed elimination method. A third method using DBU for the elimination, reported by Okamtoto *et al.*,¹³⁶ was also tried. In this method, per-*O*-acetylated β -chloro-Neu5Ac1Me **47** was reacted with DBU in dry DCM at room temperature for 2.5 hours under Ar. The overall yield of per-*O*-acetylated Neu5Ac2en1Me **50** from Neu5Ac1Me **49** after purification by chromatography was 74%.



On scale-up of the DBU reaction to multi-gram scale, an impurity was noted to form during the reaction that appeared to be (on the basis of an olefinic (H-3) coupling constant of 5.7 Hz and the loss of an acetate group) per-*O*-acetylated 4-*epi*-OH

Neu5Ac2en1Me **67**. Recrystallisation (EtOAc/hexanes, DCM/MeOH, hexanes/MeOH, DCM/acetone) was unable to separate the main product **49** from the impurity. However, careful chromatography on a smaller scale (< 1 g) (7:1 EtOAc/hexanes) enabled the isolation of fully protected Neu5Ac2en derivative **49** in clean form as confirmed by ^1H NMR spectroscopy. A summary of elimination reaction results can be seen in Table 3.6.



Scheme 3.17

Reagents & Conditions. a) dry pyridine, 1 h, 50°C, 55%; b) dry NEt_3 , dioxane, 0.5 h, 80°C, 74%; c) DBU, dry DCM, 2.5 h, rt, 74%.

Table 3.6: Summary of the elimination reactions examined for the synthesis of **49** from **47**.

Method of HCl elimination from 47			Yield of 49 (over two steps from 50)
pyridine	50°C	1.0 h	~55%
Et_3N in dioxane	80°C	0.5 h	~74%
DBU in DCM	rt	2.5 h	~74%

De-*O*-acetylation of tetra-*O*-acetyl-Neu5Ac2en1Me **49** was carried out using NaOMe in dry MeOH solution to give the desired compound Neu5Ac2en1Me **64** in excellent yield (92%). The key intermediate 8,9-*O*-isopropylidenated Neu5Ac2en1Me **52** has been prepared in the literature by the reaction of **64** with acetone in the presence of triflic acid.¹¹⁸ In this work, **52** was produced by suspending Neu5Ac2en1Me **64** in dry acetone and 2,2-dimethoxypropane at room temperature under an atmosphere of argon (Scheme 3.16). This was followed by the addition of Amberlite IR-120 (H^+) ion exchange resin and the reaction was left to stir at room temperature overnight. After the removal of the resin by filtration, and evaporation of the solvent, subsequent treatment with dry NEt_3 and resuspension of the resultant residue in DCM yielded the product as a white precipitate in quantitative yield. ^1H NMR analysis of the precipitated product indicated pure **52**; two characteristic isopropyl methyl resonances were present in the ^1H NMR spectrum at 1.35 and 1.39 ppm. The overall yield of the key isopropylidenated

Neu5Ac2en derivative **52** from Neu5Ac **1** (using DBU catalysed elimination) was ~65% over five steps.

3.3.3 Coupling of alkyl/aralkyl halides with 8,9-*O*-isopropylidene-Neu5Ac2en1Me, **52**.

The C-4-ether derivatives of Neu5Ac2en selected to be prepared based on the modelling studies from Section 2.3.7.3 of Chapter 2, were 4-*O*-benzyl- **29-A**, 4-*O*-(2-phenyl)benzyl- **29-B**, 4-*O*-ethyl- **29-C**, 4-*O*-hexyl- **29-D**, and 4-*O*-decyl- **29-E** Neu5Ac2en (Figure 3.3).

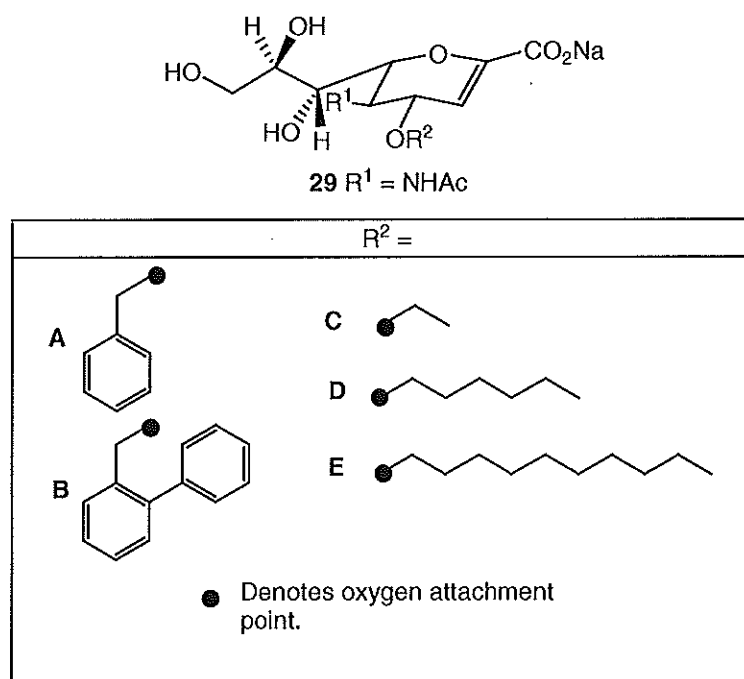
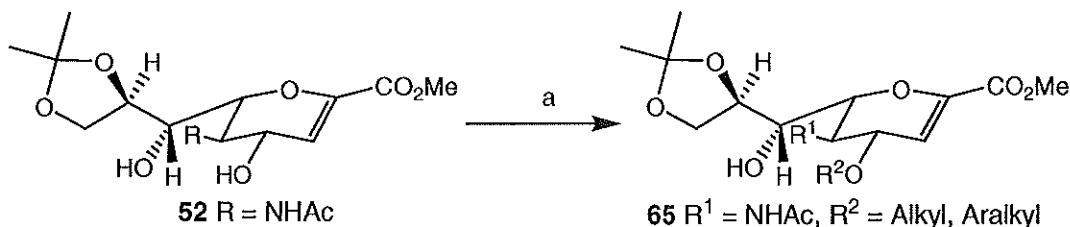


Figure 3.3: Structure of the five 4-*O*-substituted derivatives of Neu5Ac2en to be synthesised.

3.3.3.1 Sodium hydride mediated alkylation of **52**

The 4-*O*-alkylation of 8,9-*O*-isopropylidene-Neu5Ac2en1Me **52** was first tried using the method reported by Castro-Palomino *et al.*¹¹⁸ with sodium hydride as coupling mediator, as shown in Scheme 3.18. The alkyl/aralkyl halides used and reaction yields are indicated in Table 3.7, Section 3.3.3.3. The NaH mediated coupling involved stirring 8,9-*O*-isopropylidene-Neu5Ac2en1Me, **52**, with 1.5 equivalents of the coupling partner in dry DMF (Scheme 3.18). Sodium hydride was added at 0°C and the reaction was left to stir until TLC (EtOAc) indicated reaction completion (typically 8 hours). The reaction mixture was then quenched with 0.1 mL of dry MeOH and, after a work-up consisting of evaporation of DMF and aqueous extraction, the crude product was

chromatographed using 5:1 EtOAc/hexanes as eluent to give the desired product generally as an off-white foam. The yields for these reactions, which ranged from 0 to 63% depending on the alkyl halide used, are presented in Table 3.7. Due to problems involving lack of reproducibility of the reaction outcome (discussed in Section 3.3.3.3), the use of silver (I) oxide as the reaction mediator was also investigated.

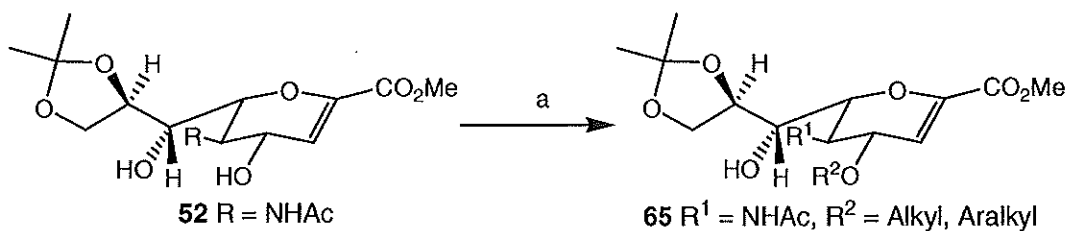


Scheme 3.18

Reagents & Conditions. a) Alkyl/aralkyl halide, NaH, dry DMF, 8 h, rt, Ar.

3.3.3.2 Silver (I) oxide mediated alkylation of **52**

The 4-*O*-alkylation of 8,9-*O*-isopropylidene-Neu5Ac2en1Me **52** was also tried using silver (I) oxide to mediate the coupling procedure (Scheme 3.19).



Scheme 3.19

Reagents & Conditions. a) Alkyl/aralkyl halide, Ag₂O, *n*-Bu₄NI, dry DCM 3 Å molecular sieves, 2 d, Ar, rt.

The method was performed as reported by Ikeda *et al.*¹⁴⁴ with minor modification for the 4-*O*-alkylation of a Neu5Ac thiophenyl glycoside **54** (see Section 3.1.2). In these reactions, 8,9-*O*-isopropylidene-Neu5Ac2en1Me **52** and 3 molar equivalents of alkyl/aralkyl halide were pre-dried by stirring over 3 Å molecular sieves in dry DCM for 2 hours at room temperature. Subsequently, 3 molar equivalents of silver oxide, and 1 equivalent of tetrabutyl ammonium iodide (*n*-Bu₄NI) were added. The reaction was protected from light and stirred at room temperature for two days. TLC (EtOAc) analysis after the 2 days had elapsed indicated that all starting material had been consumed and converted to a single major product. The reaction was filtered through Celite[®] and the reaction mixture was evaporated to dryness. Column chromatography

on the crude product was performed using EtOAc as the eluent and reaction yields (ranging from 0 to 82%), and the aryl/alkyl halides, used are presented in Table 3.7.

3.3.3.3 Analysis of NaH and Ag₂O mediated alkylation reactions

The reaction yields of the sodium hydride and silver (I) oxide mediated alkylation reactions are shown in Table 3.7. As can be seen, all alkyl halides with the exception of alkylation with 2-phenylethyl bromide were successfully coupled with the starting material **52**. The alkylation seen was specific for C-4, as no ¹H NMR spectrum showed the presence of C-7 alkylation or 4,7-di-*O*-alkylation. Yields when using NaH were more uniform across the different alkyl/aralkyl halides than when using Ag₂O.

Table 3.7: Yields of sodium hydride and silver (I) oxide alkylation reactions on **52**.

Coupling Partner	Product	Reaction yield ^a (NaH)	Reaction yield ^a (Ag ₂ O)
benzyl bromide	65-A	65%	62%
2-phenylbenzyl bromide	65-B	61%	82%
iodoethane	65-C	62%	40%
bromohexane	65-D	61%	34%
bromodecane	65-E	33%	11%
2-phenylethyl bromide	65-G	0%	0%

^a Isolated yields after chromatography. Represents best yield for alkylation reaction.

Yields for the alkylation reactions with sodium hydride ranged from 33% with bromodecane (**65-E**) to 65% with benzyl bromide (**65-A**). No 4-*O*-alkylated derivative was conclusively produced in the reaction with 2-phenylethyl bromide. An explanation for the lack of reactivity of 2-phenylethyl bromide under these conditions could be dehydrobromination to the form 2-phenylethene, thereby removing the coupling partner from the reaction.

The coupling of benzyl bromide was successful, with isolated yields in the region of 65%. This is somewhat below the 83% yield reported¹¹⁸ for this reaction. Although reasonable yields with NaH could be obtained, this reaction was found to be quite variable. The bromohexane coupling reaction was particularly variable in its outcome; initially the reaction was successful, however subsequent attempts led to recovery of partially deprotected starting material (either de-*O*-isopropylidenated or de-esterified during the workup procedure).

The coupling reaction using bromohexane was further investigated in particular to evaluate reaction pH at all steps in order to determine if the removal of the protecting

groups (methyl ester and/or isopropylidene) was the cause of the reduced reaction yield. The reaction was performed under Ar using the following conditions, 1.3 equivalents of alkyl halide, dry DMF, 1.3 equivalents of NaH, 3 Å molecular sieves, 5 h, rt. At completion of the reaction (as judged by TLC) the pH was tested and found to be 8.5; after subsequent addition of 0.1 mL of dry MeOH, the pH was once again tested and found to be 9.0. The reaction mixture was diluted with chloroform, filtered and concentrated under vacuum. The residue was then taken up in DCM, washed with ice-cold 0.1 M HCl, H₂O, sat. NaCl, dried (Na₂SO₄), filtered and concentrated to dryness. NMR analysis of the crude product showed the desired product and indicated that the methyl ester and the isopropylidene group were not removed during the normal work-up.

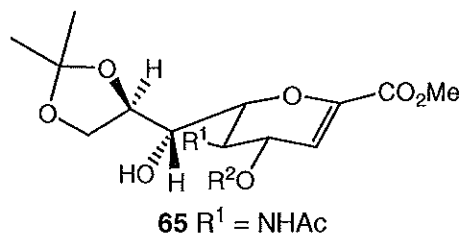
The purification of the successfully coupled 8,9-*O*-isopropylidene-4-*O*-hexyl-Neu5Ac2en1Me (**65-D**) however was problematic. It is postulated that the low isolated yields of some of the reactions could be due to de-*O*-isopropylidenation on the silica column (due to the slightly acidic nature of silica). This however could be countered by the addition of NEt₃ to the column eluent. As a result of this study, the optimised work-up for the NaH mediated alkylation of 8,9-*O*-isopropylidene-Neu5Ac2en1Me, **52**, involved the addition of 0.1 mL dry MeOH, filtration and evaporation of reaction solvent and column chromatography using 5:1 EtOAc/Hex to which a small amount (~1 mL) of NEt₃ had been added.

The coupling reactions performed using silver (I) oxide, with the exception of the 2-phenylethyl bromide reaction, were successful to varying degrees. The least successful was the bromodecane reaction at 11% yield, in comparison the most successful reaction 2-phenylbenzyl bromide at 82% yield. Once again there was no reaction with the 2-phenylethyl bromide coupling partner; it may be that the elimination of bromide from the coupling partner also occurs in the presence of Ag₂O.

Products from the 4-*O*-alkylation reactions were characterised by ¹H NMR and MS analysis. Selected ¹H NMR data for the 4-*O*-alkylated Neu5Ac2en derivatives is given in Table 3.8. In the ¹H NMR of the 4-*O*-alkylated derivatives, H-4 is generally observed to shift upfield by ~0.4 ppm from the shift observed for the starting 8,9-*O*-isopropylidene-Neu5Ac2en1Me derivative **52** (H-4: 4.46 ppm). This is consistent with the addition of a shielding group on the oxygen atom. H-5 is observed to shift downfield field by between ~0.2 to 0.4 ppm depending on the substituent attached, this

may be due to through space interactions. No great difference is seen in the shift of H-4 and H-5 between each of the synthesised derivatives. All compounds gave the expected mass for the (M-H)⁻ ion in negative ion mode ESI LRMS.

Table 3.8: Selected data for the 4-*O*-alkylated-Neu5Ac2en derivatives **65**.

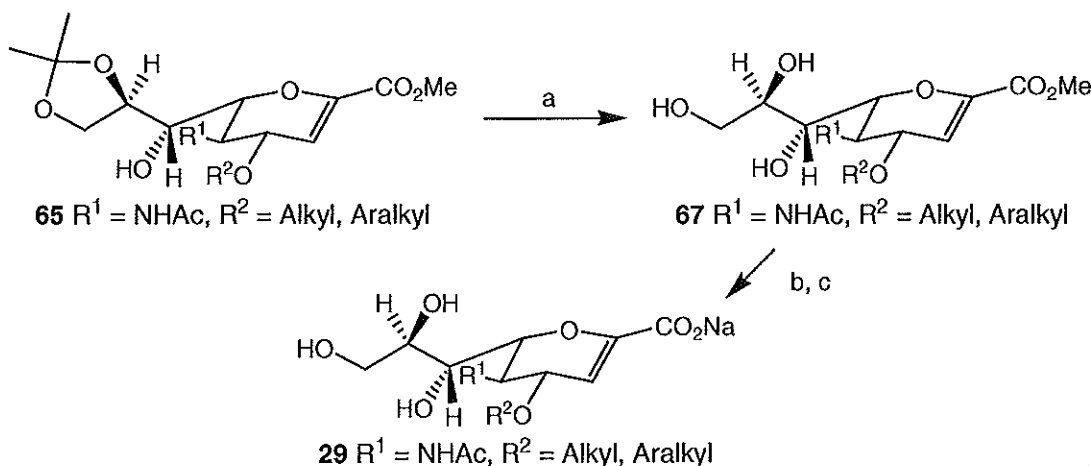


R ²	Compound Number	R _f ^a	H-4 ^b	H-5 ^b
H	52	0.46	4.46	3.97
benzyl	65-A	0.60	4.11	4.18
2-phenylbenzyl	65-B	0.75	4.10	4.15
ethyl	65-C	0.51	4.03	4.16
hexyl	65-D	0.71	4.01	4.13
decyl	65-E	0.75	4.04	4.36

^a TLC (EtOAc), ^b ¹H NMR 300 MHz.

3.3.4 Deprotection of 4-*O*-alkylated derivatives of Neu5Ac2en.

The deprotection of the 4-*O*-alkylated derivatives of Neu5Ac2en **65** was performed in a two-step one-pot reaction as shown in Scheme 3.20.



Scheme 3.20

Reagents & Conditions. a:) 80% AcOH, 0.5 h, ~83°C; b) 0.1 M KOH, MeOH, 1 h, rt; c) i: Amberlite IR-120 (H⁺), ii: aq. NaOH to pH 7.8, (35-64% over two steps).

The deprotection steps involved the initial removal of the isopropylidene group protecting the C-8 and C-9 hydroxyl groups followed by the de-esterification of the C-1 carboxylic acid. De-isopropylidination was carried-out by the use of 80% AcOH at 83°C for 1 hr. After evaporation of the AcOH, de-esterification was carried out by addition of 1:1 0.1 M KOH and MeOH. The deprotected derivatives were acidified to pH 3-4 with Amberlite IR-120 (H⁺) resin. The solution was then filtered to remove the resin and concentrated to dryness. The residue was redissolved in H₂O and neutralised to pH 7.8 by the addition of 1M NaOH to form the sodium salt in greater than 100% yield after lyophilisation. The greater than 100% yield is caused by the neutralisation process where a salt residue is formed due to the pH change.

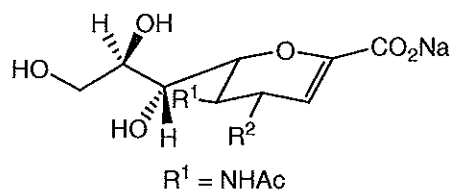
Successful deprotection was indicated in ¹H NMR spectroscopic analysis by the loss of the isopropylidene resonance (2 × 3H singlets, region of δ 1.25 to δ 1.35), and loss of the methyl ester resonance (1 × 3H singlet, δ ~3.70). ¹H NMR spectroscopic analysis also showed no epimerisation of the deprotected compounds was apparent. The deprotected 4-*O*-alkylated derivatives **29-A** to **29-E** were finally purified using semi-preparative reverse phase high performance liquid chromatography as discussed in Section 3.3.5. The structures of the C-4-ether linked Neu5Ac2en derivatives synthesised are shown in Figure 3.3 in Section 3.3.3.

3.3.5 HPLC conditions for purification of C-4-thioether and C-4-ether Neu5Ac2en derivatives.

The fully deprotected final C-4-thioether and C-4-ether Neu5Ac2en derivatives were purified using isocratic normal phase HLPC on a semi-preparative C18 column. The conditions for each compound, the elution time off the column at a flow rate of 4 mL/min and the epimeric ratio (in the case of C-4-thioether derivatives) are shown in Table 3.9.

A number of conditions were evaluated in an attempt to separate the epimers of the C-4-thioether derivatives. However, in no case were the HPLC conditions sufficient to remove the C-4-*epi* derivatives from the desired equatorially *S*-substituted derivatives. Therefore each of the C-4-thioether derivatives was sent for biological evaluation with the presence of the C-4-*epi* epimer in the ratios shown in Table 3.9.

Table 3.9: HPLC conditions for purification of C-4-thioether and C-4-ether Neu5Ac2en derivatives.



R^2	Isocratic ACN %	Elution time (Min)	Epimeric mixture ^a (eq:epi)
<i>S</i> -benzyl 28-A	2%	18.2	1:0.1
<i>O</i> -benzyl 29-A	5%	8.6	-
<i>S</i> -(2-phenyl)benzyl 28-B	20%	10.3	1:0.2
<i>O</i> -(2-phenyl)benzyl 29-B	20%	7.5	-
<i>S</i> -ethyl 28-C	3%	5.1	1:0.2
<i>O</i> -ethyl 29-C	1%	4.8	-
<i>S</i> -hexyl 28-D	10%	15.7	1:0.3
<i>O</i> -hexyl 29-D	15%	10.9	-
<i>S</i> -decyl 28-E	25%	13.6	1:0.3
<i>O</i> -decyl 29-E	25%	13.1	-
<i>S</i> -(<i>p</i> -cyanobenzyl) 28-F	5%	11.2	1:0.1
<i>S</i> -(2-phenyl)ethyl 29-G	10%	8.8	1:0.3

^a C-4 thioether derivatives, epimeric ratio was determined by 300 MHz ¹H NMR.

3.4 Summary and Conclusions

In conclusion, the synthesis of novel C-4-thioether **28-A** to **28-G** (Figure 3.1) and C-4-ether **29-A** to **29-E** (Figure 3.3) Neu5Ac2en derivatives has been achieved. The synthesis of C-4 thioether derivatives was achieved in 6 steps from Neu5Ac **1**. The key intermediate of the C-4-thioether synthesis, per-*O*-acetylated 4-thiolacetyl-Neu5Ac2en1Me **58**, was prepared *via* the Neu5Ac2en oxazoline **56**, which also resulted in the formation of the 4-*epi*-thiolacetyl-Neu5Ac2en1Me derivative (~14% of total product). Modification of the literature¹¹⁷ procedure for this reaction by reducing the number of equivalents of thiolacetic acid from 5 to 1.5 minimised the formation of the epimer. With the synthesis of the key intermediate **58**, 4-*S*-alkylation was then achieved by base-mediated thiolate formation and reaction with alkyl/aralkyl halides. The *S*-alkylations proceeded in isolated yields ranging from 17% to 84% and no further C-4 epimerisation was noted during the alkylation reaction. Initially, de-*O*-acetylation of per-*O*-acetylated 4-*S*-benzyl-Neu5Ac2en1Me derivative **59-A** was performed using NaOMe in dry MeOH. These reaction conditions however, caused an increase in the proportion of the C-4 epimer. By performing a dual de-*O*-acetylation/de-esterification

in 1:1 0.1M KOH/MeOH the increase of the proportion of the epimer was reduced. After neutralisation the seven C-4-thioether derivatives were purified by HPLC to give the target compounds **28-A** to **28-G** as a variable mixture of 4-*S*-epi and 4-*S*-equatorial substituted product.

The synthesis of C-4-ether derivatives was achieved in 8 steps from Neu5Ac **1**. The key intermediate of the synthesis, 8,9-*O*-isopropylidene-Neu5Ac2en1Me **52**, was prepared through a sequence of reactions involving the β -elimination of a C-2 halide from Neu5Ac1Me derivative, **47**, followed by de-*O*-acetylation and 8,9-*O*-isopropylidenation. Initially, selective 4-*O*-alkylation was mediated with NaH. However the yields obtained for the alkylation reactions were highly variable upon repetition of the reaction. 4-*O*-Alkylation was then achieved by the use of silver (I) oxide; this reaction was slower than the NaH reaction however reaction yield was not as variable. The isolated yields of the *O*-alkylation reactions ranged from 11% to 82%. Deprotection involving de-*O*-isopropylidenation and de-esterification were achieved using standard conditions without difficulty. Finally, the C-4-ether Neu5Ac2en derivatives were purified by HPLC to give the target compounds **29-A** to **29-E**.

Ultimately, the C-4-substituted Neu5Ac2en derivatives that were synthesised were: *S* and *O*-benzyl- **28/29-A**, *S* and *O*-2-phenylbenzyl- **28/29-B**, *S* and *O*-ethyl- **28/29-C**, *S* and *O*-hexyl- **28/29-D**, *S* and *O*-decyl- **28/29-E**, *S*-*p*-(cyano)benzyl- **28-F**, and *S*-(2-phenyl)ethyl- **28-G** derivatives.

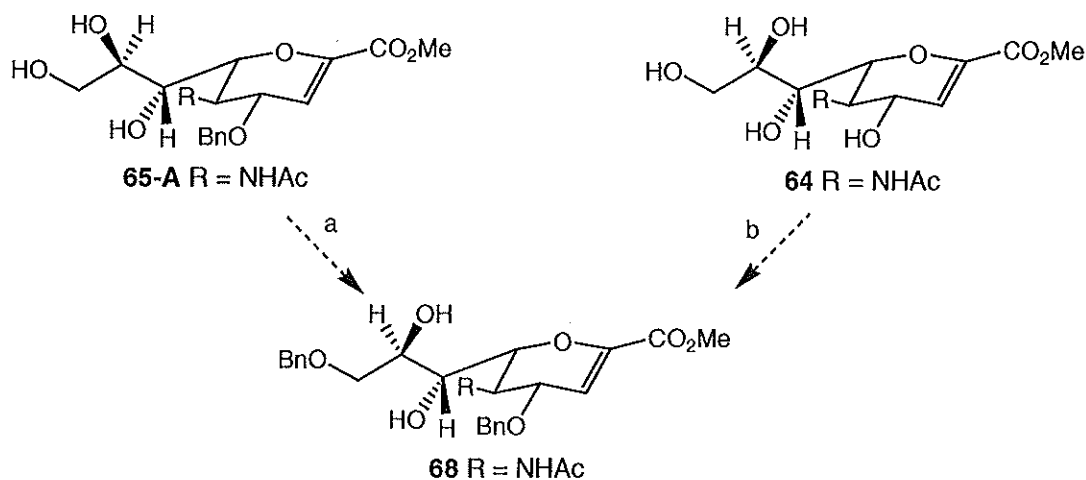
3.5 Future Work

A number of areas of research have been identified throughout this work that could be done in the area of synthesis, and in the analysis of synthetic reactions; these are outlined below in Sections 3.5.1 to 3.5.4.

3.5.1 Bis-*O*-benzylation of Neu5Ac2en1Me.

Modelling studies of NDV HN (Section 2.3.2.6) predicted an area of potential hydrophobic interaction in the glycerol-side chain pocket. The interaction could be potentially exploited by a benzyl group substituted on the C-9 hydroxyl. The preparation of such a derivative could be accomplished *via* two methods (Scheme 3.21); either C-9 derivatisation after the 4-*O*-benzyl derivative **65-A** has been prepared

(Scheme 3.21, step a) or benzylation of both the C-9 and C-4 hydroxyl groups of Neu5Ac2en1Me **64** simultaneously with 2 equivalents of benzyl bromide (Scheme 3.21, step b). Either of the two methods could afford the 4,9-di-*O*-benzyl-Neu5Ac2en1Me derivative **68**.

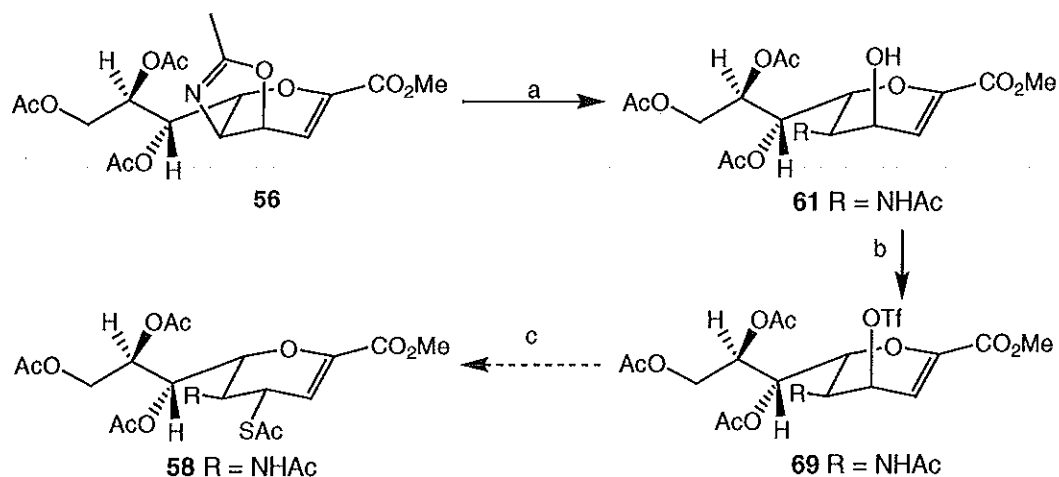


Scheme 3.21

Reagents & Conditions. a) 1.5 equiv. BnBr, Ag₂O, n-Bu₄NI, dry DCM, 2 d, rt; b) 2 equiv. BnBr, NaH, dry DMF, 0.5 h, rt.

3.5.2 Alternative synthesis of 4-thiolacetyl Neu5Ac2en1Me derivative **58**

An alternative synthetic approach that could lead to the preparation of epimerically pure 4-thiolacetylated Neu5Ac2en is required. One potential approach is that outlined in Scheme 3.22.



Scheme 3.22

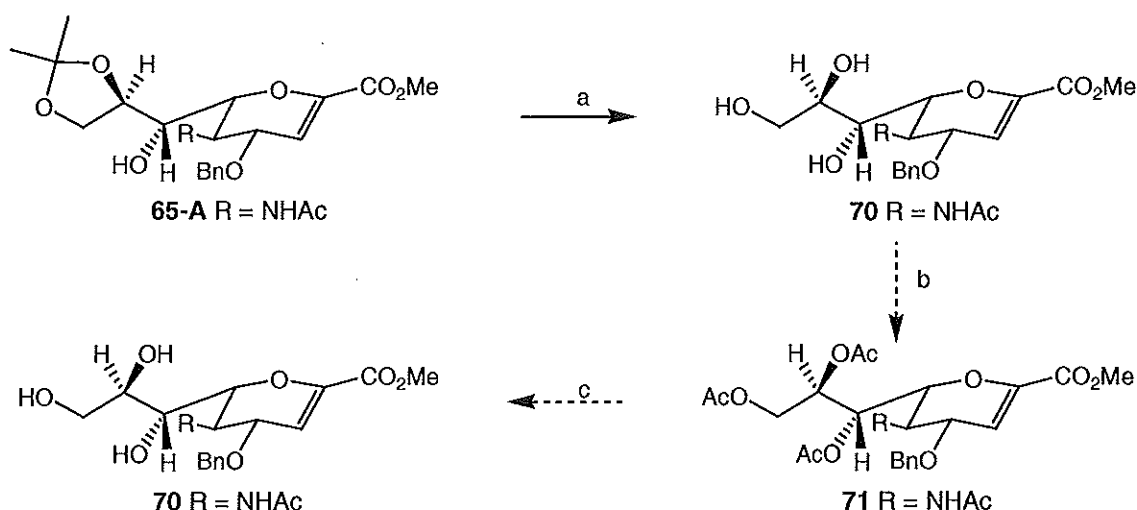
Reagents & Conditions. a) TFA, THF, H₂O, 4 h, rt; b) Tf₂O, DCM, pyridine, 0°C; c) KSAc, dry DMF, 70°C, 2 d.

The acid-catalysed opening of the Neu5Ac2en oxazoline **56** to the 4-*epi*-OH analogue of 7,8,9-tri-*O*-acetylated-Neu5Ac2en1Me, **61**, has previously been reported in the literature by Schreiner *et al.*¹⁷⁶ The reaction was carried-out and yielded 82% of **61** after chromatographic purification. 4-*Epi*-hydroxy derivative **61** was then reacted using a standard triflation procedure (triflic anhydride in DCM/pyridine) to give the 4-*epi* triflate **69** after evaporation of the reaction solvent. The triflated Neu5Ac2en derivative could then potentially be reacted in a nucleophilic displacement reaction with potassium thiolacetate in dry DMF, to theoretically give the equatorially-substituted 4-thiolacetyl derivative of Neu5Ac2en, **58**.

3.5.3 Investigation of C-4 epimer formation from 4-*O*-benzyl-Neu5Ac2en1Me, **65-A**.

As discussed earlier (Section 3.2.5) during the de-*O*-acetylation of 7,8,9-tri-*O*-acetyl-4-*S*-benzyl-Neu5Ac2en1Me **59-A** with NaOMe at room temperature or 0°C the proportion of the 4-*epi* form of the derivative was increased. It would be of interest to investigate whether such epimerisation at C-4 could occur during the NaOMe mediated deprotection of the analogous C-4-ether linked derivative.

A sample of 4-*O*-benzyl-8,9-*O*-isopropylidene-Neu5Ac2en1Me **65-A** could be de-*O*-isopropylidenated using 80% AcOH (Section 3.3.4), and the deprotected compound **70** could then be per-*O*-acetylated to give 7,8,9-tri-*O*-acetyl-4-*O*-benzyl-Neu5Ac2en1Me **71** (Scheme 3.23). De-*O*-acetylation of **71** would then be carried-out using the same conditions (1 M NaOMe, MeOH, rt, 1 h) under which epimerisation was observed for the corresponding 4-*S*-benzyl derivative **59-A**. NMR analysis of the product would establish whether two olefinic (H-3) signals were present. This could demonstrate whether epimerisation of 7,8,9-tri-*O*-acetyl-4-*O*-benzyl-Neu5Ac2en1Me **71** occurs under NaOMe mediated deprotection conditions.



Scheme 3.23

Reagents & Conditions. a.) 80% AcOH, 0.5 h, ~83°C; b) pyridine, Ac₂O, o/n, rt; c) 1 M NaOMe, dry MeOH, Ar, 1 h, rt.

3.5.4 Computational analysis of C-4-thioether and ether Neu5Ac2en derivatives.

It would be of interest to examine the conformational stability of C-4-thioether and C-4-ether linked Neu5Ac2en derivatives with regard to the epimerisation of the thioether linked derivatives. This could be accomplished by computational means using programs such as MOPAC.¹⁸¹ The analysis would involve mapping the potential energy surface of the *epi* and equatorial forms of the C-4-ether linked Neu5Ac2en derivatives, mapping the partial charges of each atom in each molecule and measuring the heat of formation. The purpose of these calculations would be to evaluate the C-4 proton to examine whether it could be removed in a basic reaction environment, and whether once formed is the *epi* form more stable than the equatorial form. Considerations in performing the calculation include the starting conformation of each molecule as the calculation is conformation dependant. It would also be of interest to extend the calculation to include the generation of multiple conformers to examine the effect of geometry on the calculation. Subsequent averaging of the partial charges of the multiple conformers may give a better picture of the electrostatic potential and thus conformational stability. Using MOPAC to examine the reaction by transition state analysis may also give a better understanding as to why epimerisation occurs during NaOMe mediated deprotection of C-4-thioether derivatives of Neu5Ac2en but not the C-4-ether derivatives.

4 Biological evaluation of ligands designed to interact with paramyxoviral HN.

4.1 Introduction

The compounds prepared in this work (Chapter 3) had been designed (Chapter 2) to interact with the active site of paramyxoviral HN's. Two biological assays were available to us to assess the interactions of these compounds with paramyxoviral HN's; an assay to assess inhibition of the sialidase activity of the HN protein and an assay to assess inhibition of paramyxoviral growth. The neuraminidase inhibition assay could be carried-out with human parainfluenza viruses 1 and 3,⁸³ while the viral growth inhibition assay used a recombinant Sendai virus in which the natural HN had been replaced with the HN from hPIV-1.⁸⁴ Results from these studies, carried out in the laboratories of Prof. Yasuo Suzuki (Shizuoka, Japan), and Prof. Allen Portner (Memphis, USA), respectively, presented herein are gratefully acknowledged. Unfortunately the error values for these assays had not been communicated to us at the time of writing this thesis. So it is not possible to clearly state what is an experimentally significant difference in inhibition values. Therefore it is not possible to make firm comparisons between the biological and computational activity of the thioether and ether linked derivatives. However, in this case it is still of interest to examine if a relationship between the biological and computational activity of the thioether and ether linked derivatives exists. The synthesised compounds that were evaluated for biological activity are listed in Table 4.1. The known neuraminidase inhibitors Neu5Ac2en **5** and 4-deoxy-4-guanidino-Neu5Ac2en **6** were also evaluated as benchmark compounds. Structure determination of one of the designed and synthesised compounds 4-*O*-benzyl-Neu5Ac2en **29-A** in complex with NDV HN was carried out in the lab of Prof. Garry Taylor (St. Andrews, Scotland).

Table 4.1: Compounds evaluated as inhibitors of paramyxoviral HN.

Compound	NA Activity		Inhibition of viral growth
	hPIV-1 HN	hPIV-3 HN	rHN ^a
28-A 4- <i>S</i> -benzyl-Neu5Ac2en	✓	-	-
28-B 4- <i>S</i> -(2-phenyl)benzyl-Neu5Ac2en	✓	-	-
28-C 4- <i>S</i> -ethyl-Neu5Ac2en	✓	-	-
28-D 4- <i>S</i> -hexyl-Neu5Ac2en	✓	-	-
28-E 4- <i>S</i> -decyl-Neu5Ac2en	✓	-	-
28-F 4- <i>S</i> -(<i>p</i> -cyanobenzyl)-Neu5Ac2en	✓	-	-
28-G 4- <i>S</i> -(2-phenyl)ethyl-Neu5Ac2en	✓	-	-
29-A 4- <i>O</i> -benzyl-Neu5Ac2en	✓	✓	✓
29-B 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en	✓	✓	✓
29-C 4- <i>O</i> -ethyl-Neu5Ac2en	✓	-	-
29-D 4- <i>O</i> -hexyl-Neu5Ac2en	✓	-	-
29-E 4- <i>O</i> -decyl-Neu5Ac2en	✓	✓	✓

^a Recombinant Sendai virus with hPIV-1 HN

4.2 *In vitro* biological evaluation

4.2.1 Assay for inhibition of sialidase activity of hPIV-1 HN.

Evaluation of inhibition of the sialidase activity of hPIV-1 and 3 HN was carried out in a whole virus assay using the fluorogenic substrate 4-methylumbelliferyl-*N*-acetylneuraminide (MUN, **7**) (Scheme 4.1). The assay⁸³ involved the incubation of the virus with inhibitor for 60 minutes, followed by incubation with MUN for 30 mins. Inhibition of sialidase activity was assessed by measuring the fluorescence of each experiment. Inhibition of sialidase activity results in a decrease in fluorescence (released free methylumbelliferone **72**) relative to the control (no inhibitor).

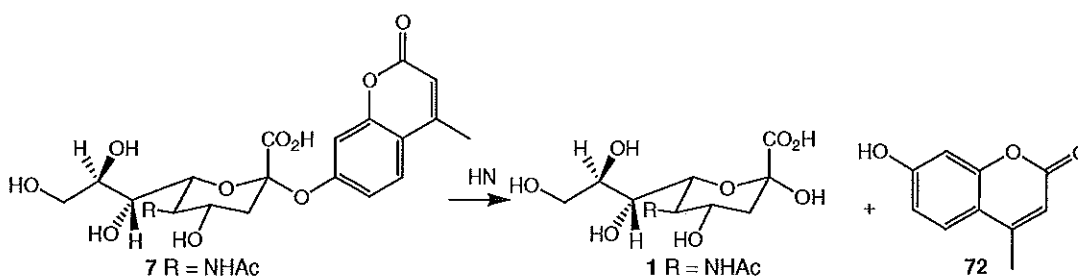
**Scheme 4.1**

Table 4.2 shows the biological data (IC₅₀ values) obtained for the inhibition of hPIV-1 sialidase activity by designed ligands, with numbers in square brackets designating rank from most active [1] to least active [12].

Table 4.2: Inhibition (IC₅₀, μM) of hPIV-1 sialidase activity by designed C-4-ether and C-4-thioether substituted Neu5Ac2en derivatives.^a

Compound	IC ₅₀ (μM)	Compound	IC ₅₀ (μM)
5 Neu5Ac2en	189		
29-A 4- <i>O</i> -benzyl-Neu5Ac2en	34 [2] ^b	28-A 4- <i>S</i> -benzyl-Neu5Ac2en	129 [9]
29-B 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en	61 [4]	28-B 4- <i>S</i> -(2-phenyl)benzyl-Neu5Ac2en	82 [6]
29-C 4- <i>O</i> -ethyl-Neu5Ac2en	72 [5]	28-C 4- <i>S</i> -ethyl-Neu5Ac2en	21 [1]
29-D 4- <i>O</i> -hexyl-Neu5Ac2en	145 [11]	28-D 4- <i>S</i> -hexyl-Neu5Ac2en	115 [8]
29-E 4- <i>O</i> -decyl-Neu5Ac2en	114 [7]	28-E 4- <i>S</i> -decyl-Neu5Ac2en	585 [12]
		28-F 4- <i>S</i> - <i>p</i> -(cyanobenzyl)-Neu5Ac2en	135 [10]
		28-G 4- <i>S</i> -(2-phenyl)ethyl-Neu5Ac2en	38 [3]

^a Assays were performed in duplicate. Results have been rounded to the nearest whole number

^b Numbers in square brackets indicate relative ranking of inhibitory potency.

The benchmark for hPIV-1 HN sialidase activity inhibition was established with Neu5Ac2en **5** which had an IC₅₀ of 189 μM. All of the compounds tested, with the exception of 4-*S*-decyl-Neu5Ac2en **28-E**, showed greater inhibitory activity than Neu5Ac2en **5**. The most active compound analysed was 4-*S*-ethyl-Neu5Ac2en **28-C** with an IC₅₀ measured at 21 μM. This gain represents an almost 10 fold increase in activity over the benchmark inhibitor. 4-Deoxy-4-guanidino-Neu5Ac2en **6** was sent to be tested as a benchmark against hPIV-1 NA activity, however no activity was seen.

From the data in Table 4.2 there is no clear trend observed for the 4-*O*-alkylated vs the 4-*S*-alkylated derivatives in regard to relative potency of inhibition. Amongst the 4-*O* and 4-*S* I-alkylated derivatives, the 4-*S*-ethyl (**28-C**), 4-*O*-decyl (**29-E**) derivatives appeared to be more active than their corresponding *O*- or *S*- counterparts. However for the benzyl and (2-phenyl)benzyl derivatives, the *O*-derivative was either more active (**29-A**) or of comparable activity (**29-B**) to the analogous *S*-derivative. It should be noted that the 4-*S*-alkylated derivatives are mixtures of epimers at C-4 with the 4-*epi* derivative making up approximately 5 to 22% of the mixture. The 4-*S-epi*-substituted compounds were not anticipated to bind to the HN active site and this therefore decreases the amount of compound available to inhibit the sialidase activity of hPIV-1.

As noted in the preceding paragraph each of the 4-*S*-alkylated Neu5Ac2en samples contains a different proportion of the 4-*epi* isomer. Due to the variant amount of *epi* compound present in each sample, comparisons and conclusions drawn between each of the thioether derivatives is made lightly. The biological data shows that the best *S*-linked compound is a short chain alkyl derivative 4-*S*-ethyl derivative **28-C** (IC₅₀ 21 μM). This is followed by a phenyl ring extended from the linker by an ethyl group **28-G** (IC₅₀ 38 μM), the most active form of the 4-*S*-aralkyl derivatives. In the case of the *S*-linked benzyl and (2-phenyl)benzyl derivatives two aromatic rings are better than one. The inhibition activity obtained for the alkyl derivatives follows a clear trend based on length, from shortest being the most active to the longest being the least active.

For the 4-*O*-alkylated Neu5Ac2en derivatives, the biological data suggests that the best *O*-linked compound contains a single aromatic ring (4-*O*-benzyl derivative **29-A**, IC₅₀ 34 μM), followed by two aromatic rings (4-*O*-(2-phenyl)benzyl derivative **29-B**, IC₅₀ 61 μM) in contrast to the *S*-linked derivatives. The data obtained for the three *n*-alkyl derivatives does not follow a trend based on length from smallest to largest. It would be anticipated that the short alkyl chains would be better accommodated within the active site than the long alkyl chains, although the shortest chain may not benefit from interactions in the active site pocket.

4.2.2 Assay for inhibition of sialidase activity of hPIV-3 HN.

The compounds **29-A**, **29-B** and **29-E** were assayed *in vitro* for inhibition of hPIV-3 sialidase activity in a whole virus fluorogenic sialidase assay as described in Section 4.2.1. While all the 4-*O*- and 4-*S*-alkylated Neu5Ac2en derivatives prepared in this work have been sent for evaluation against hPIV-3, not all compound data has been received from Japan at the time of writing this thesis. It is anticipated that the additional data will be published in the future.

Table 4.3 (Column 2) shows the biological data obtained on the inhibition of hPIV-3 NA by designed ligands. The numbers in brackets designate rank in activity or binding energy. Also shown in Table 4.3 (columns 3-5) is computational binding data (discussed in Section 4.3).

Table 4.3: Comparison between biological hPIV-3 sialidase activity inhibition data and computational binding energy (Section 2.3.7.3) for 4-*O*-substituted Neu5Ac2en derivatives.^a

Compound	IC ₅₀ (μ M)	MM-PBSA kcal/mol	AutoDock kcal/mol	X-Score pK _d
5 Neu5Ac2en	24 ^b	-102.5 ^c	-6.4 ^c	5.1 ^c
6 4-deoxy-4-guanidino-Neu5Ac2en	25			
29-A 4- <i>O</i> -benzyl-Neu5Ac2en	10 [1]	-110.8 [1]	-13.3 [1]	6.3 [1]
29-B 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en	30 [2]	-108.6 [2]	-12.0 [2]	5.9 [2]
29-E 4- <i>O</i> -decyl-Neu5Ac2en	10 [1]	-88.7 [3]	-11.7 [3]	5.0 [3]

^a Assays were performed in duplicate. Results have been rounded to the nearest whole number

^b Numbers in brackets indicate relative ranking of inhibitory potency.

^c Numbers in brackets indicate lowest binding energy/ highest X-Score.

The benchmark for the inhibition of hPIV-3 sialidase activity was established with Neu5Ac2en **5** which had an IC₅₀ of 24 μ M. The 4-*O*-benzyl- and 4-*O*-decyl-Neu5Ac2en derivatives showed slightly greater inhibitory activity than Neu5Ac2en **5** with IC₅₀ values measured at 10 μ M. In this biological assay, there was no apparent difference in the activity of the long alkyl chain, 4-*O*-decyl-Neu5Ac2en derivative **29-E**, on the inhibition of hPIV-3 HN NA over that shown with the single aromatic ring derivative **29-A**. The aralkyl derivative 4-*O*-(2-phenyl)benzyl-Neu5Ac2en **29-B** was seen to have marginally less activity than the benchmark compound.

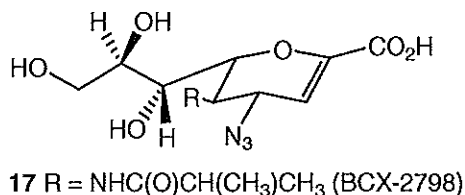
The potent influenza virus sialidase inhibitor 4-deoxy-4-guanidino-Neu5Ac2en **6** was also evaluated for inhibition of hPIV-3 sialidase activity. The IC₅₀ was measured at 25 μ M, making it equal in activity to Neu5Ac2en. Interestingly, in the literature 4-deoxy-4-guanidino-Neu5Ac2en **6** was reported to inhibit hPIV-3 HN sialidase activity at an IC₅₀ of 190 μ M,⁸² with Neu5Ac2en inhibiting at an IC₅₀ of 2.1 mM.⁸² However, the nature of the assay which used sonicated hPIV-3 preparations was quite different to the assay used in this work, so the two sets of results cannot be directly compared.

The designed Neu5Ac2en derivatives show in general a modest increase in inhibitory activity against hPIV-1 and hPIV-3 HN sialidase activity. The sialidase inhibition results obtained to date do show that the use of a hydrophobic functionality at C-4 of Neu5Ac2en is sound in the development of inhibitors of paramyxoviral HN. Structural differences between the HN of hPIV-1 and hPIV-3 could account for the

differences in levels of NA activity inhibition seen between the two paramyxoviruses. However the lack of crystallographic data at this time for hPIV-1 HN means that we do not yet know what differences there are between the two binding/active sites. The inhibition demonstrated by the designed compounds does validate the working hypothesis of this thesis, that NDV HN could be used as a model of parainfluenza HN.

4.2.3 Assay for inhibition of viral growth of rSV (hPIV-1 HN).

Three of the designed 4-*O*-alkylated Neu5Ac2en derivatives **29-A**, **29-B**, and **29-E** were evaluated for *in vitro* inhibition of viral replication of a recombinant Sendai Virus containing hPIV-1 HN.⁸⁴ This study was of interest due to the availability of the inhibitor BCX-2798, **17**, as a benchmark, which would allow for comparison between the best reported⁸⁴ inhibitor to date and some of the compounds produced in this study. Inhibition effects of the ligands on virus growth were determined in a plaque reduction assay. LLC-MK2 cells were pre-treated with different concentrations (0.001 to 10 μ M) of each compound for 1 hour, after which virus was added. Infection of cells by virus was visually assessed after incubation for 72 hours.



The results for the designed 4-*O*-linked Neu5Ac2en derivatives in the plaque reduction assay can be seen in Table 4.4. Unfortunately, none of the compounds showed inhibition of viral growth at or below 10 μ M concentration. A similar result was obtained for the benchmark Neu5Ac2en **5**. This result is consistent with the IC₅₀ values found for the inhibition of hPIV-1 NA activity (Section 4.2.1). The inhibitory effect of BCX-2798 was measured in the assay at 0.01 to 0.1 μ M. The reported⁸⁴ inhibitory activity of BCX-2798 in a similar assay is 0.7 μ M. In a test of the toxicity of the designed compounds, no differences between treated uninfected cells and untreated uninfected cells were found visually at all tested concentrations. The reason for the significant difference in activity between the Neu5Ac2en derivatives and BCX-2798 **17** is uncertain but interaction of **17** with hPIV-1 HN may have contribution from the alternate acyl group on the C-5 position.

Table 4.4: Inhibitory effect (EC_{50} , μM) of 4-*O*-linked Neu5Ac2en derivatives on rSV (hPIV-1 HN) growth *in vitro*.

Compound	EC_{50} (μM)
5 Neu5Ac2en	≥ 10
6 4-deoxy-4-guanidino-Neu5Ac2en	≥ 10
17 BCX-2798	0.01 - 0.1
29-A 4- <i>O</i> -benzyl-Neu5Ac2en	≥ 10
29-B 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en	≥ 10
29-E 4- <i>O</i> -decyl-Neu5Ac2en	≥ 10

4.3 Comparisons between inhibitory data against hPIV sialidase activity and computational data (NDV HN) for C-4-ether and C-4-thioether derivatives.

Comparisons between the observed hPIV-1 sialidase activity inhibition data of 4-*O*-alkylated Neu5Ac2en derivatives (Section 4.2.1) and the predicted relative strengths of binding to NDV HN derived from computational data (Chapter 2, Section 2.3.7.3) can be seen in Table 4.5. Comparisons are made to the data derived using minimised NDV HN due to the more rigorous computational analysis used in Section 2.3.7.3 compared to Section 2.3.7.2 in the design phase. The five designed compounds are ranked from best [1], (greatest inhibition, best binding energy, highest X-Score value), to the worst [5].

Table 4.5: Comparison between *in vitro* hPIV-1 sialidase activity inhibition data and computational data (Section 2.3.7.3) for 4-*O*-substituted Neu5Ac2en derivatives.^a

Compound	IC_{50} (μM)	MM-PBSA kcal/mol	AutoDock kcal/mol	X-Score pK_d
5 Neu5Ac2en	189	-102.5	-6.4	5.1
29-A 4- <i>O</i> -benzyl-Neu5Ac2en	34 [1]	-110.8 [1]	-13.3 [1]	6.3 [1]
29-B 4- <i>O</i> -(2-phenyl)benzyl-Neu5Ac2en	61 [2]	-108.6 [2]	-12.0 [2]	5.9 [2]
29-C 4- <i>O</i> -ethyl-Neu5Ac2en	72 [3]	-82.8 [5]	-10.0 [5]	5.3 [4]
29-D 4- <i>O</i> -hexyl-Neu5Ac2en	145 [5]	-106.1 [3]	-11.9 [3]	5.5 [3]
29-E 4- <i>O</i> -decyl-Neu5Ac2en	114 [4]	-88.67 [4]	-11.7 [4]	5.0 [5]

^a Blue represents inhibition or a score better than the benchmark; red represents a score worse than the benchmark. Numbers in square brackets indicate relative ranking of inhibitors potency or computational score.

The hPIV-1 sialidase activity inhibition data for the 4-*O*-benzyl- **29-A** and 4-*O*-(2-phenyl)benzyl- **29-B** Neu5Ac2en derivatives is in line with the energy values and pK_d values (after minimisation) calculated for these compounds with NDV HN, which predicted that the derivative with the single aromatic ring would bind better to HN than the “biphenyl” derivative. The biological data was also in line with the computational data that suggested that the activity of the aralkyl derivatives would be greater than that of the alkyl derivatives. The biological data obtained for the alkyl derivatives (4-*O*-ethyl, 4-*O*-hexyl, 4-*O*-decyl) does not follow any discernable trend (size, chain length). Computationally, at least two derivatives scored worse than the benchmark Neu5Ac2en, however this was not reflected in the biological data. While computational analysis suggested that the 4-*O*-hexyl derivative **29-D** would bind better than the 4-*O*-ethyl derivative **29-C** to NDV HN, the 4-*O*-ethyl derivative **29-C** was a better inhibitor of hPIV-1 sialidase activity. This could be due to the ethyl derivative lacking a number of rotatable bonds which is more favourable upon binding as it will not attract an entropic penalty (caused by freezing single bonds in a bound state). The computational data for the 4-*O*-ethyl-Neu5Ac2en derivative **29-C** in two out of three scoring functions performs better than the benchmark.

The data presented in Table 4.5 demonstrates that the use of the NDV HN X-ray crystal structure as a model for hPIV-1 HN was a valid premise for 4-*O*-substituted Neu5Ac2en derivatives. Each compound was designed and computationally evaluated against the NDV HN crystal structure. The subsequent *in vitro* sialidase inhibitory activity against hPIV-1 demonstrates that the C-4 binding domain of hPIV-1 HN can accommodate hydrophobic groups, in particular the large 2-phenylbenzyl group. This also suggests that the active site of hPIV-1 HN could be hydrophobic in nature.

Table 4.6 shows the trend comparison between the biological data obtained for inhibition of hPIV-1 sialidase activity by the 4-*S*-alkylated Neu5Ac2en inhibitors and the associated computational data (Section 2.3.7.3). It should be noted that in comparison to computational data for hPIV-3 (Section 2.5.4.1) and the 4-*O*-substituted Neu5Ac2en derivatives against hPIV-1, computational data for the 4-*S*-substituted derivatives with hPIV-1 shows poor consensus across the three scoring methods used. This makes trend comparison of biological and computational data difficult. The seven

designed compounds are ranked from best [1], greatest inhibition, lowest binding energy, highest X-Score value, to the worst [7].

Table 4.6: Comparison between *in vitro* hPIV-1 sialidase activity inhibition data and computational data (Section 2.3.7.3) for 4-*S*-substituted Neu5Ac2en derivatives.^a

Compound		IC ₅₀ (μ M)	MM- PBSA kcal/mol	AutoDock kcal/mol	X-Score pK _d
5	Neu5Ac2en	189.0	-102.5	-6.4	5.1
28-A	4- <i>S</i> -benzyl-Neu5Ac2en	129 [5]	-116.9 [3]	-13.7 [4]	6.3 [2]
28-B	4- <i>S</i> -(2-phenyl)benzyl-Neu5Ac2en	82 [3]	-113.7 [4]	-14.3 [2]	6.7 [1]
28-C	4- <i>S</i> -ethyl-Neu5Ac2en	21 [1]	-110.4 [5]	-12.0 [7]	5.6 [7]
28-D	4- <i>S</i> -hexyl-Neu5Ac2en	115 [4]	-94.3 [7]	-13.3 [5]	5.8 [6]
28-E	4- <i>S</i> -decyl-Neu5Ac2en	585 [7]	-124.6 [2]	-15.9 [1]	6.1 [5]
28-F	4- <i>S</i> - <i>p</i> -(cyanobenzyl)-Neu5Ac2en	135 [6]	-124.9 [1]	-13.3 [6]	6.2 [4]
28-G	4- <i>S</i> -(2-phenyl)ethyl-Neu5Ac2en	38 [2]	-96.9 [6]	-14.3 [3]	6.3 [3]

^a Blue represents inhibition or a score better than the benchmark; red represents a score worse than the benchmark. Numbers in square brackets indicate relative ranking of inhibitors potency or computational score.

Overall, trends for the inhibition of hPIV-1 sialidase activity by the C-4-thioether linked derivatives did not match well with the predicted strength of binding to NDV HN given by all the scoring protocols. With the presence of a variable amount of the 4-*epi* form in all C-4-thioether derivatives, the biological results may not actually reflect actual activity and therefore this will impact on matching with scoring data.

A comparison between hPIV-3 sialidase activity of 4-*O*-alkylated Neu5Ac2en derivatives **29-A**, **29-B** and **29-E** and computational estimation of the binding energies of these compounds to the HN of NDV using three scoring functions is given in Table 4.3. In each case this computational data suggested that the most effectively bound compound would be the 4-*O*-benzyl derivative **29-A**, followed by the 4-*O*-(2-phenyl)benzyl derivative **29-B**, both of which would bind more effectively than the parent 4-hydroxy compound Neu5Ac2en **5**. The binding of the 4-*O*-decyl derivative **29-E**, however, was predicted to be less than that of the benchmark in two of the scoring functions (Table 4.3, Columns 3 and 5). The trend in the estimated binding energies is mirrored in the biological data for inhibition of hPIV-3 HN sialidase activity by **29-A** and **29-B**. It is expected that with such a small dataset that drawing comparisons

between biological and computational data is going to be difficult at best and not really indicative of a relationship between the computationally predicted and biological results.

The designed ligands were assessed for their binding to NDV HN using a number of scoring functions. Each of the scoring functions used involves the solving of parameters that could describe the binding of a ligand to a macromolecule. However each calculation is different from the others in key ways. The MM-PBSA includes a solvation calculation in addition to a molecular interaction energy calculation. This is the key difference between MM-PBSA and AutoDock, which considers only the molecular interaction energy calculation. The X-Score binding function is empirical in nature and force-field independent. The use of these three different binding functions was to try to establish a consensus between the functions and the biological inhibitory data. As particularly seen in Table 4.6 for the most part this consensus was not apparent. This could suggest that there are real differences in the computationally predicted binding to NDV HN and the actual biological binding of these molecules. Or it could suggest that new scoring protocols are required.

4.3.1 Comparison of 4-*O* and 4-*S*-alkylated Neu5Ac2en derivatives

As discussed above (Section 4.2.1), there are differences in inhibition against hPIV-1 sialidase activity between the 4-*O*- and 4-*S*-substituted Neu5Ac2en derivatives with the same substituent. Computationally, the derivatives are discussed in Section 2.3.8. It was generally predicted computationally that the thioether derivatives would be more active than their ether linked counterparts, however it has been seen that the ether derivatives are generally the more biologically active form, with two out of five of the ether linked compounds clearly out performing the analogous thioether linked derivatives.

4.4 Analysis of the X-ray crystal structure of 4-*O*-benzyl-Neu5Ac2en/NDV HN complex.

The X-ray crystal structure of 4-*O*-benzyl-Neu5Ac2en **29-A** in complex with NDV HN was undertaken in the laboratory of Prof. Garry Taylor (St. Andrews, Scotland). Crystals of NDV HN (globular ectodomain, Residues 124-370) in complex with **29-A** were prepared by the soaking method.^{182, 183} The crystal structure of the complex was solved to 2.7 Å.¹⁸³ The HN-ligand complex was prepared using crystals of NDV HN formed at pH 4.6, in contrast to the crystal structure previously solved with Neu5Ac2en

5 which was prepared by co-crystallisation at pH 6.5.⁴⁸ The topology of the head of a monomer of the enzyme with bound inhibitor 4-*O*-benzyl-Neu5Ac2en is shown in Figure 4.1. It can be seen that the globular head of NDV HN still⁴⁸ conforms to the six-blade anti-parallel β -sheet propeller shown previously in Figure 1.4 for the NDV HN-Neu5Ac2en complex. Also present in the structure with **29-A** are the three alpha helical domains (with one of these domains binding a calcium ion) that were observed in the original X-ray crystal structure.⁴⁸ This section describes the analysis of the X-ray crystal structure of the 4-*O*-benzyl-Neu5Ac2en/NDV-HN complex. Subsequent to this study being performed, Ryan *et al.*¹⁸³ published an analysis of the same structure with similar conclusions to our study being found.



Figure 4.1: Rendered overview of the tertiary structure of the complex of NDV HN with 4-*O*-benzyl-Neu5Ac2en **29-A** (white).^{49, 50}

An electrostatics analysis of the active site of the 4-*O*-benzyl-Neu5Ac2en **29-A**/NDV HN complex is shown in Figure 4.2. The active site of NDV HN is dominated around the carboxylate group of the ligand by a positive field potential. The floor of the active site is mixed, with a large area of negative potential. The C-5 binding pocket is predominantly neutral in character, which would facilitate the binding of the methyl functionality of the C-5 acetamido group. The glycerol side-chain binding pocket is lined by positive field potential on one side and negative on the other (not seen in Figure 4.2), with these areas separated by a small area of neutral field potential.

These results are no different to those found upon Neu5Ac2en binding NDV HN. The very large C-4 binding pocket [calculated volume of C-4 pocket of crystal structure of NDV HN with bound **29-A** (with **29-A** removed); 721 Å³, volume of C-4 pocket calculated from NDV/Neu5Ac2en crystal structure (Section 2.3.2.5): 612 Å³] has a predominantly weak positive potential, however neutral and negative areas are also found within the pocket. The 4-*O*-benzyl group is seen to make interaction with an area of neutral field potential.

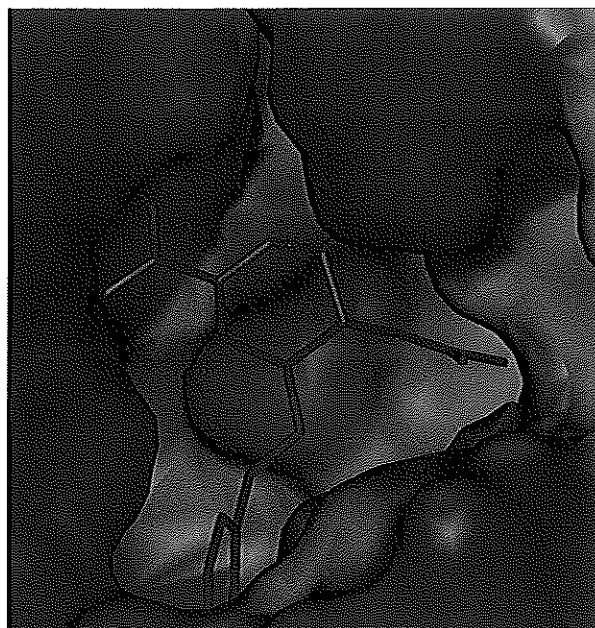


Figure 4.2: Electrostatics analysis of active site of NDV HN, with crystal 4-*O*-benzyl-Neu5Ac2en rendered. Blue is an area of positive electrostatic potential, white is neutral and red is negative potential.

A Ligplot analysis of the important interactions between the inhibitor 4-*O*-benzyl-Neu5Ac2en **29-A** and NDV HN is shown in Figure 4.3. As anticipated, the carboxylate functionality of 4-*O*-benzyl-Neu5Ac2en is positioned to make six H-bonding interactions with the arginine residues (Arg 174, Arg 416 and Arg 498) of the triarginyl cluster. The glycerol side-chain of **29-A** can make both hydrogen bonding, and non-bonding interactions with five residues in this area of the active site, although not all are identified by Ligplot (Glu 258, Tyr 262, Arg 416, Ser 418 and Tyr 317). Three water mediated interactions also are noted between the nitrogen of the C-5 acetamido group and Ser 237, Tyr 317 and Glu 401, while the acetamido methyl group makes non-bonding hydrophobic interactions with Tyr 299. These interactions with the carboxylate, glycerol side chain and acetamido methyl group of **29-A** are the same as those previously reported in the crystal structure of NDV HN complexed with

Neu5Ac2en.⁴⁸ The new 4-*O*-benzyl group is positioned to make multiple hydrophobic interactions within the C-4 pocket with residues Ile 175, Ile 192, Cys 196, Arg 197, and Asp 198. The side chain of Ile 175 in the NDV HN/29-A complex is noted to have an alternate conformation compared to that in the NDV/Neu5Ac2en **5** complex, in order to make interaction with the 4-*O*-benzyl group. The catalytic residue Asp 198, previously noted to shift position in different pH forms, is noted to take the same position as the pH 4.6 NDV HN/Neu5Ac **1** complex (1E8U), it is oriented towards **29-A**. This is to be expected as the pH of the NDV HN/29-A complex is also 4.6.

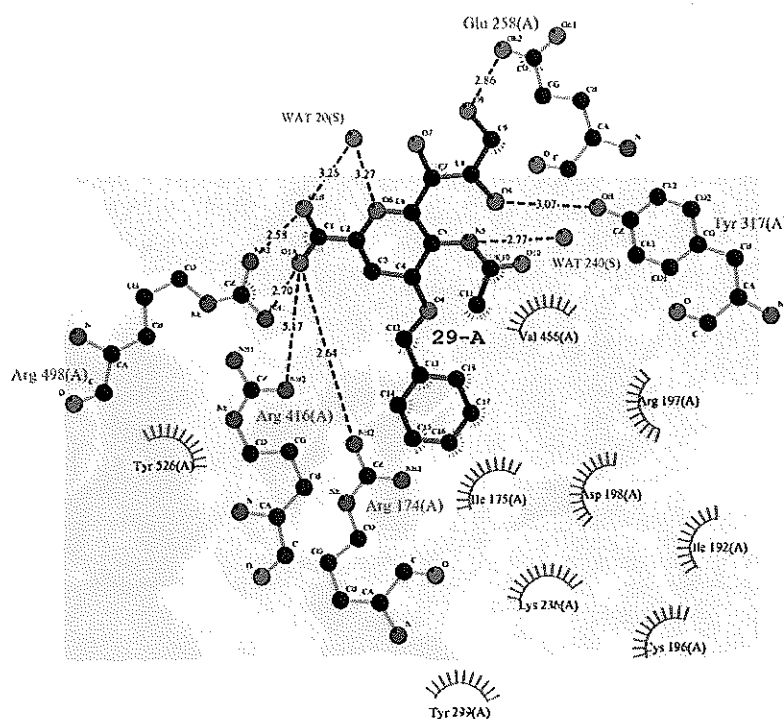


Figure 4.3: Ligplot analysis of important interactions between 4-*O*-benzyl-Neu5Ac2en **29-A** and NDV HN. H-bonding interactions are shown as dotted green bonds, with the number on the green bond being the measured distance of the H-bond (in Å).

A close-up of the active site of the crystal structure with both the ligand and crystal water residues rendered is shown in Figure 4.4. One water of crystallisation can be seen at the bottom of the deep C-4 pocket. The reported crystal structure of NDV HN in complex with Neu5Ac2en showed two waters of crystallisation within the C-4 pocket (Figure 2.6, reproduced here with modification for clarity). It can be seen that the two crystal water residues previously found to bind within the C-4 pocket are not present in the crystal structure with 4-*O*-benzyl-Neu5Ac2en bound. This supports GRID competition calculations (Section 2.3.2.6) that suggested that these water molecules would be displaced by a hydrophobic functionality.

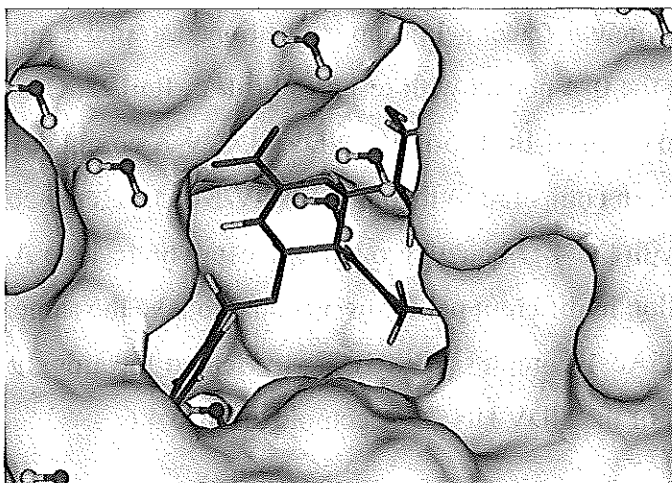


Figure 4.4: Close up of NDV active site complexed with **29-A**. Crystal waters around the active site are included and rendered in ball-and-stick representation.

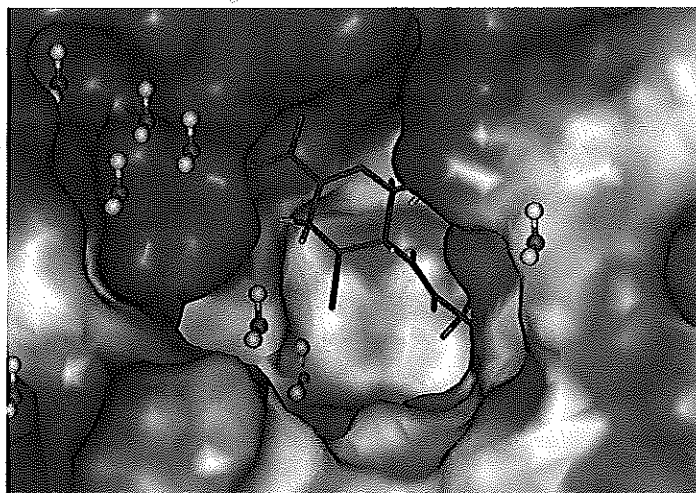


Figure 2.6: C-4 Binding pocket of NDV HN with Neu5Ac2en **5** bound and crystal waters ball-and-stick rendered.

4.4.1 Comparison of the X-ray crystal structure vs computationally docked 4-*O*-benzyl-Neu5Ac2en/NDV HN complexes.

The availability of an X-ray crystal structure of NDV HN complexed with **29-A** allows for a comparative analysis between the X-ray crystal structure orientation and the predicted docked orientation of ligand **29-A**. It also allows for the comparison of the predicted binding energy with the computational scoring of the X-ray crystal structure orientation. It also allows the analysis of the impact of the presence of crystal water residues in the AutoDock calculations.

Superimposition of the docked pose of **29-A** in NDV HN and the NDV HN X-ray crystal structure orientation of 4-*O*-benzyl-Neu5Ac2en **29-A** is shown in Figure 4.5,

where the top of the active site has been removed for clarity. The docked pose of **29-A** was obtained from the second improved scoring protocol (Section 2.3.7.3). Superimposition of the X-ray crystal structures was performed by superimposition of the backbone of each protein. Both structures presented in Figure 4.5 represent a "snapshot" of the binding of the ligand to the active site of NDV HN. It can be seen in Figure 4.5 that the crystal pose of **29-A** is slightly different to that of the docked pose. The side-chains of the crystal and docked poses of the ligand are observed to bind to the appropriate active site pockets although there are differences in exact orientation of the glycerol side chain. The superimposed structures in the full active site can be seen in Figure 4.6. The placement of the benzyl ring in the crystal structure form of the ligand is actually similar to the computational structure generated for the docking of 4-*S*-benzyl-Neu5Ac2en (Section 2.3.7.2.1).

The major difference between the docked and crystal structures of 4-*O*-benzyl-Neu5Ac2en is the placement of the benzyl ring in the large C-4 pocket which causes a shift in the positioning of the other side-chains. The 4-*O*-benzyl ring of **29-A** makes interactions with Ile 175 in both the docked and crystal poses, however due to the shift of Ile 175 in the NDV HN/**29-A** crystal structure, the crystal pose is positioned within the C-4 pocket of the active site near the edge as opposed to the docked pose which lies flat on the bottom of the C-4 pocket.

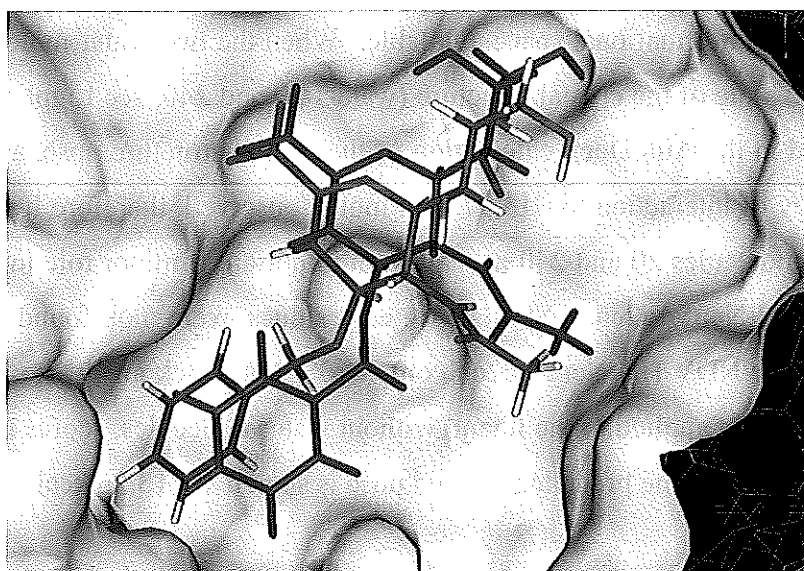


Figure 4.5: Superimposition of docked (purple) and crystal (green) bound 4-*O*-benzyl-Neu5Ac2en in NDV HN; the top of the active site has been removed for clarity.

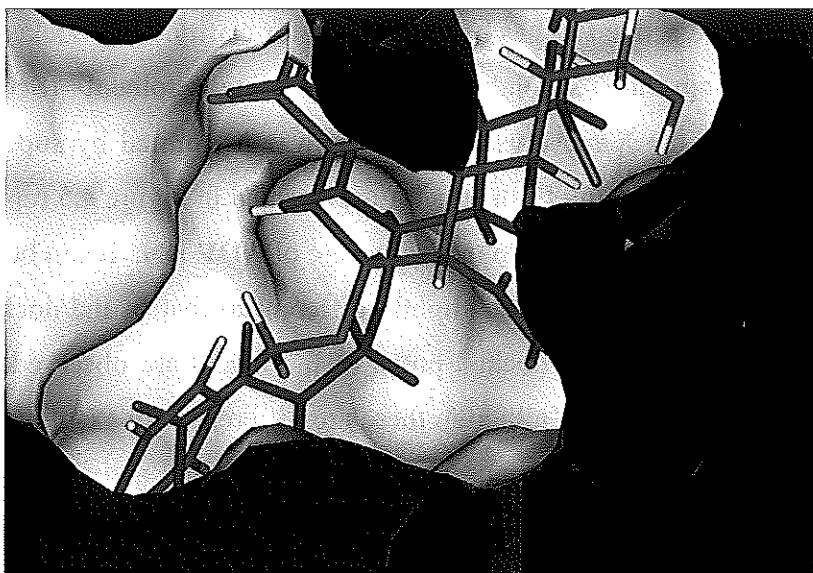


Figure 4.6: Superimposition of docked (purple) and crystal (green) bound 4-*O*-benzyl-Neu5Ac2en in the full active site of NDV HN.

AutoDock analysis (by the first method, Section 2.3.7.2) statically scored the crystal structure conformation of 4-*O*-benzyl-Neu5Ac2en without bound water molecules at -12.37 kcal/mol; the previously docked 4-*O*-benzyl-Neu5Ac2en (without added water molecules) was statically scored at -17.10 kcal/mol (Section 2.3.7.2). The difference in interaction energy is postulated to be a combination of factors involving the conformational shift of Ile 175 and the presence of water during crystallisation (which would guide the ligand into its final pose), whereas the docking calculation has all active site conformational space available. When the two water molecules observed in the binding site of the crystal structure (waters 20 and 240, see Ligplot in Figure 4.3) are included in the AutoDock calculation of the crystal structure pose, a slight increase in the energy of interaction to -13.92 kcal/mol is observed, therefore ligand interaction with water molecules 20 and 240 may play an important if minor role in the interaction of 4-*O*-benzyl-Neu5Ac2en. Similar water residues were not found in the crystal structure of the NDV HN/Neu5Ac2en complex.⁴⁸ AutoDock analysis, using ADT with Gasteiger charges applied to the ligand without water molecules, scores the crystal pose at -12.15 kcal/mol. The predicted docked pose is scored at -13.26 kcal/mol.

4.5 Future Work

The first round of computational design to develop ligands for paramyxoviral HN that may act as inhibitors of HN function has been completed with the successful design (Section 2.3.7) of a number of ligands which are found inhibit hPIV-1 sialidase activity

better than the benchmark ligand, Neu5Ac2en (Section 4.2.1). Future work could entail the design of enhancements to the more active ligands in the current series to increase inhibitor potency. The current study has touched on this point with the proposal for substitution at the C-9 position of Neu5Ac2en (Section 2.3.10.1), however the secondary modification of Neu5Ac2en was only a preliminary study and has great scope to be expanded.

The design work using alternate templates such as cyclopentane (BCX-1812) and cyclohexane (Tamiflu™) modified with hydrophobic substitutions could be taken into the synthetic phase of development and biological testing. These templates and others were computationally predicted to bind to NDV HN with affinity comparable to the Neu5Ac2en based derivatives (detailed in Section 0). Additionally new templates could be located *via* pharmacophore or database searching.

In the area of biological testing, it would be of great interest to test the designed ligands against the HN of NDV, as the crystal structure of NDV HN was used to develop and refine the ligands. This would allow a clearer evaluation of the ability of the scoring methods used to predict biological activity. The testing of the full range of ligands against the HN of hPIV-3 is also of interest, to examine the potency of each ligand against hPIV-3, and to examine whether there is a marked difference in a compound's activity between hPIV-1 and 3. In addition the effect of these ligands on the hemagglutination activity of the HN, as explored by the Moscona group⁸² and others⁸⁴ would also be of interest.

4.6 Summary and Conclusions

This study began with the aim of the development of novel inhibitors of the HN activity of hPIV. This was to be accomplished *via* three processes. Firstly, the computational design of potential ligands of the HN glycoprotein using the X-ray crystal structure of the structurally related NDV HN. Secondly, the chemical synthesis of the designed ligands. Finally, the biological assessment of the synthesised ligands against the HN of hPIV. The computational design process identified 168 potential candidates using six individual templates, each utilising two different linkages between the template and an alkyl/aralkyl substituent positioned to fill the C-4 pocket of HN. The template selected for the synthesis of potential inhibitors was the Neu5Ac2en

template, due to the well established biological inhibition of the target protein by Neu5Ac2en **5** and the relative ease of synthetic access to these compounds.

Of the twenty-eight possible molecules (14 *O*-linked and 14 *S*-linked) designed, sixteen molecules were selected for synthesis. Five *O*-linked and seven *S*-linked derivatives of Neu5Ac2en were successfully synthesised and purified. The purified compounds were evaluated for inhibition of the sialidase activity of hPIV-1 and hPIV-3 HN, with results obtained showing that of the twelve compounds evaluated, eleven of them were more potent than the base template, Neu5Ac2en against hPIV-1. Two out of three compounds showed improved inhibition of hPIV-3 sialidase activity over Neu5Ac2en. One of the most potent *O*-linked derivatives, 4-*O*-benzyl-Neu5Ac2en **29-A** was crystallised with the HN of NDV. Analysis of the structure obtained, and molecular modelling studies, show that the ligand binds to the active site in a similar conformation to that predicted computationally.

The most active of the Neu5Ac2en derivatives modified by hydrophobic C-4 substituents against hPIV-1 sialidase activity was 4-*S*-ethyl-Neu5Ac2en which inhibits at 21 μ M; two compounds, 4-*O*-benzyl- and 4-*O*-decyl-Neu5Ac2en showed activity against hPIV-3 HN at 10 μ M. These compounds are a good starting point for the further refinement and development of inhibitors of the multifunctional haemagglutinin-neuraminidase of the paramyxoviruses. In conclusion the use of structure-based drug design against hPIV HN using NDV HN as a model has yielded twelve C-4 substituted Neu5Ac2en derivatives, eleven of which are more active than the starting template. This inhibition activity demonstrates that computational design using NDV HN as a model for hPIV HN could be a valuable tool in the development of inhibitors of paramyxoviral diseases.

5 Experimental

5.1 Modelling experimental

5.1.1 Structure validation

Prior to molecular modelling the PDB files to be used were checked for errors by the following programs and shell scripts: Procheck 3.5.3⁸⁹, Whatcheck 4.99g⁹⁰, and Hbplus 3.06.⁹¹ Procheck was run by specifying the PDB file to be analysed and the crystallographic resolution of the PDB file, specified in the header of the PDB file. Procheck is utilised to assess how normal or unusual the geometry of the residues of protein.⁸⁹

5.1.1.1 Optimisation of hydrogen-bond network of the protein.

Whatcheck was executed by specifying the PDB file. The purpose of Whatcheck is to verify if the crystal structure proposed, given that torsion angles, interatomic distances etc. for a given pair of atoms are restricted to a finite range.⁹⁰ Hbplus was run by specifying the PDB file and the -X flag. Hbplus checks the H-bond network of a protein.⁹¹ The His, Gln, and Asp residues identified to be flipped were manually rotated 180° in InsightII.¹²³ The structure generated was used in the first method for ligand design.

REDUCE 2.21⁸⁸ was used to automatically optimise the hydrogen bond network *via* flipping His, Gln, and Asn residues. REDUCE was executed at the UNIX command line with the build switch enabled. Calculation times varied from half an hour to multiple hours. The optimised structures were viewed in Kinemage format. The Kinemage files were produced by the Perl script flipkin. Flipkin was executed on the Reduced structure twice, once for Asn & Gln residues and once for His residues. The Kinemage files were viewed in the program Mage 5.93. Each flipped residue was viewed in flipped and non-flipped conformation to ascertain whether REDUCE was accurate with all changes made by REDUCE accepted. The structure generated was used in the second method for ligand design.

5.1.2 Initial structure investigation

5.1.2.1 Ligplot identification of key interactions of Neu5Ac2en.

The key interactions between the bound ligand, Neu5Ac2en, and HN were identified by Ligplot 4.¹¹⁴ Ligplot takes the protein-ligand complex and both the ligand residues and interacting protein residues are “unrolled” about its rotatable bond to a 2D representation. The H-bond interactions are determined by Hbplus. Before Hbplus can calculate the H-bond the ligand needs to be processed by HBADD, otherwise Hbplus will not recognise the ligand residues H-bond acceptor and donor sites.¹¹⁴

5.1.2.2 Delphi

The protein was prepared for DelPhi 3.00⁹² calculations by the Builder module of InsightII 2000. Hydrogens were added to the protein at pH 7 and the potential and charges were set using the CVFF force-field. The conditions of the calculation were setup within the DelPhi module. The grid consisted of 33 points per axis (unminimised structure) or 65 points per axis (minimised structure) and a solvent dielectric constant of 80.0. The solute radius was set to DelPhi defaults and the solvent was set to zero ionic strength. After the conditions were set up, the DelPhi calculation was performed. The calculation took approximately 5 minutes. The results were viewed by importing the DelPhi grid and colouring the surfaced active site by the DelPhi spectrum. DelPhi is used to quantify the electrostatics of a protein. The electrostatic potential of a protein arises from the unequal distribution of electrons within bonds, due to the withdrawing effect of electronegative elements and ionised functional groups. The unequal distribution of electrons, hence charge, is represented as a fractional point charge on each atom. The electrostatic potential at a point is the force acting on a unit positive charge at a point on the protein.⁹²

5.1.2.3 PASS

The volume of the proteins active site was determined using PASS 1.0.¹⁸⁴ PASS was executed at the UNIX prompt, specifying the protein file and the InsightII flag (-InsightII). The results were viewed in InsightII by sourcing the BCL macro generated by PASS. PASS utilises geometry to characterise regions of buried volumes within the protein. The identification of such volumes may lead to the identification of the active site of the protein.¹⁸⁴

5.1.2.4 Examination of the high and low pH conformations of NDV HN

Using InsightII the NDV HN protein residue Asp198 was visually inspected in both low (4.6) pH and high (6.5) pH structure. The PDB code for the low pH structures is 1E8T, the high pH structures PDB code is 1E8V. The proteins were superimposed by the backbone, and the Asp residue was visually inspected for the conformation of the side-chain and the distance from Neu5Ac2en atom C2. The triarginyll clusters of both HN proteins and influenza N9 virus sialidase (PDB code: 1NNB) were superimposed for comparison. The distance between the low pH HN, N9 sialidase residue Asp 152 and the crystal Neu5Ac2en bound in each protein was measured.

5.1.2.5 GRID analysis of NDV HN active site

Active site extents were defined by InsightII, the extents of the active site are the values required for the cage within GRID 22.⁹³ For analysis of the protein by GRID a PDB file is required. Firstly the program GRIN is executed, and the PDB file is made suitable for computational analysis by GRID. Within the program GRID the cage defining the active site and the probe used to analyse the active site is selected. A grid spacing of 1 Å was used within the grid centred about the active site, with a 5 Å border surrounding the grid. A total of 47652 grid points were calculated. A dielectric constant of 80 was used to simulate an aqueous environment. Within the protein a dielectric constant of 4 was used to simulate the interior of the protein. The GRID calculation takes about 10 minutes on a SGI Octane system. The program GVIEW, available with version 22 of GRID views the output of the GRID calculation. GVIEW is utilised to view active site/probe interactions and to obtain a suitable contouring level for InsightII. The program gridi is utilised to generate suitable output for InsightII. In Insight II, the grid generated by gridi is read in with reference to the protein analysed. The probe is contoured with respect to the active site by the contour feature within InsightII, using the value established from GVIEW. The input files for GRIN and GRID can be seen in Appendix A.

5.1.3 Homology models

CLUSTAL W, a web-based (<http://www.cbi.ac.uk/clustalw>) alignment engine was utilised to produce the sequence alignment necessary for Modeller 6. The primary sequences of NDV HN, hPIV-1, hPIV-2 and hPIV-3 HN's were obtained from SwissProt, SwissProt accession codes: Q9Q2W5 (NDV HN), P16071 (hPIV-1 HN), P25465 (hPIV-2 HN), and P12561 (hPIV-3 HN). Sequence alignment for Modeller¹⁸⁵

was output in PIR format from CLUSTAL W.¹²² Alignment was modified for input into Modeller (modification can be seen in Appendix A). The control file was created, with all parameters left as default (seen in Appendix A) and Modeller was executed at the system prompt, the calculation takes about half an hour. Preliminary analysis of models was done with Procheck and Whatcheck.

5.1.4 *De novo* drug design using LigBuilder

LigBuilder 1.2⁹⁷ is a program used for the *de novo* generation of ligands in a rigid receptor pocket. Three steps are involved in the process: 1. The analysis of the pocket of interest using the program Pocket, which characterises the pocket based on hydrogen bond donor or acceptor sites and hydrophobic interactions. 2. To grow or link fragments of molecules together to suit the interactions found within the pocket using either the programs Grow or Link. This is accomplished using a starting seed structure and a fragment library available with LigBuilder (this library can be user expanded). 3. To process the available grown or linked ligands using the program Process and user defined parameters.

5.1.4.1 *Pocket*

The minimised NDV HN protein was used for the following generation of C-4 modified *O* and *S*-linked analogues of Neu5Ac2en. Water residues were removed from the PDB file using Insight II and the protein was saved as a PDB file, while Neu5Ac2en was saved as a Mol2 file, the potential and charges were set using the CVFF force-field. The carboxylate oxygen atoms potentials were incorrectly recorded in the Mol2 file. As such they were changed from the incorrect (O.2) to the correct (O.co2). The input parameter file can be viewed in Appendix A. The results of Pocket were viewed in InsightII; the active site interaction PDB file was split into a series of PDB files containing no more than 50 atoms (using a text editor) to facilitate easier incorporation into InsightII. The Pocket output PDB file contains nitrogen, carbon and oxygen atoms that were coloured red, blue and white to represent a H-bond donor site, H-bond acceptor site and a hydrophobic interaction site respectively.⁹⁷

5.1.4.2 *Grow or Link*

Once the active site had been characterised by Pocket, the next LigBuilder program Grow was used to grow the C-4 modified analogues of Neu5Ac2en. The seed structure Neu5Ac2en was extracted from the crystal structure of NDV and prepared for

LigBuilder by InsightII. The seed structure from the crystal structure in the active site was output as a Mol2 file for input into LigBuilder. Using a text editor the Mol2 file was opened and the C-4 hydrogen on the oxygen or sulfur atom was re-labelled H.spc (to specify 'grow from here'), and the grow parameter input file was created.⁹⁷ The parameter file for Grow and Link can be viewed in Appendix A.

5.1.4.3 Process

The output from Grow or Link needs to be processed by Process before it can be viewed or utilised in further calculations. Process has a parameter file that is created by a text editor; an example parameter file can be seen in Appendix A. Once Process was run the Mol2 files that were generated were examined in the active site of NDV HN by InsightII. An INDEX file was also generated by Process, listing the molecular formula, predicted octanol/water coefficient, and predicted binding score.⁹⁷

5.1.4.4 Further refinement of the ten best LigBuilder designed derivatives.

Twelve base templates were used for further modelling calculations: 4-*S*-Neu5Ac2en **28**, 4-*O*-Neu5Ac2en, **29**, and ten other templates: 4-*S*-Gilead template, **30**, 4-*O*-Gilead template, **31**, 4-*S*-benzoic acid template, **32**, 4-*O*-benzoic acid template, **33**, 4-*S*-uronic acid template, **34**, 4-*O*-uronic acid template, **35**, 4-*S*-BCX-1812 template, **36**, 4-*O*-BCX-1812 template, **37**, 4-*S*-furanose template, **38** and 4-*O*-furanose template, **39**. The structures of these templates are shown in Figure 5.1. The templates were modelled in InsightII¹²³, using the Builder Module. The pyranose ring templates were modelled from Neu5Ac2en, the furanose ring templates were modelled from ribose.

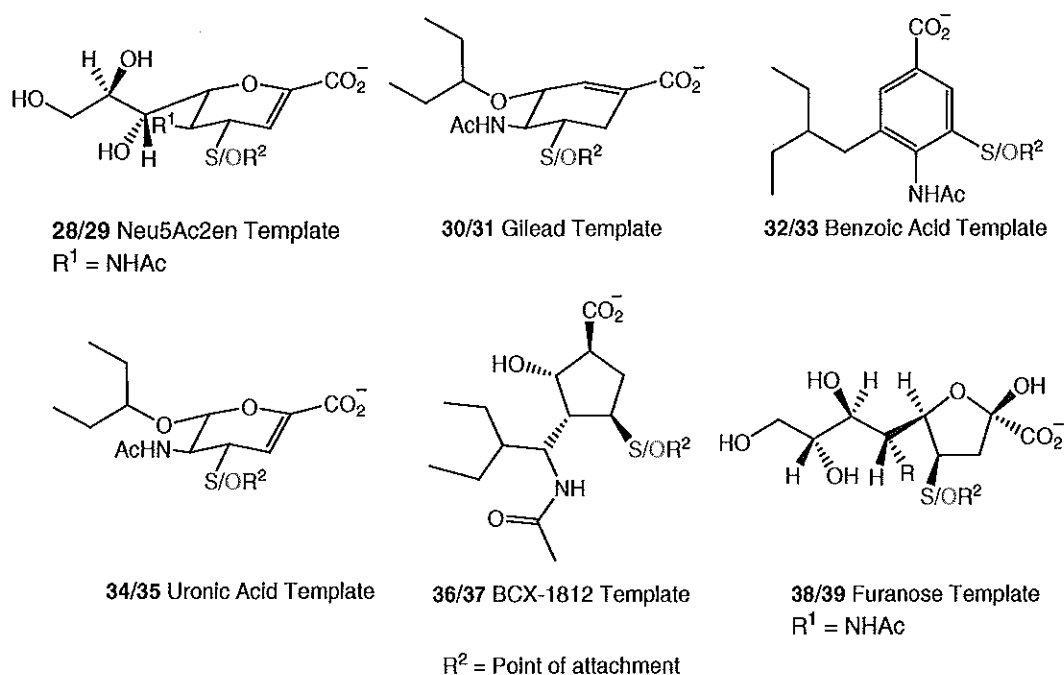


Figure 5.1: Structures of the 12 templates used in ligand design.

The ten best LigBuilder⁹⁷ designed substituents were modified to make them more chemically accessible. Subsequently after additional analysis, fourteen structures were developed from the available ten. The final structures of the 14 substituents selected for modelling are shown in Figure 5.2.

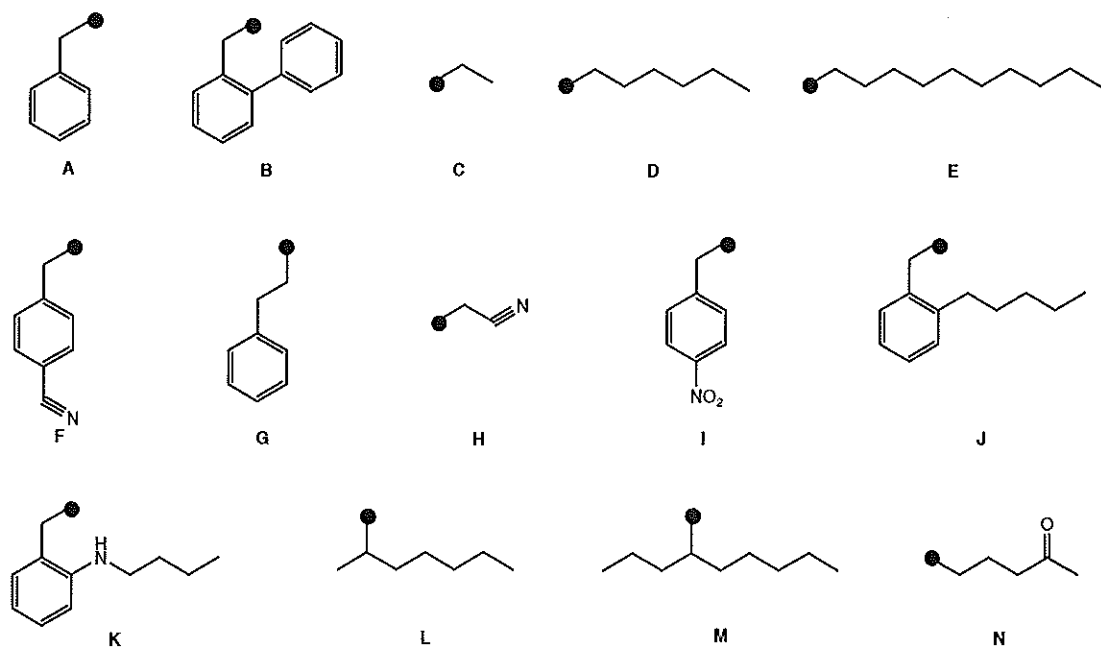


Figure 5.2: Structures of the final fourteen substituents.

Initially, substituents A – I were modelled on the C-4 position or equivalent on each of the 6 templates via a sulfur linkage, the Builder module was utilised to create the structures. These 54 structures were used in the first modelling study. With the

development of the second method, the Builder module was used to connect each of the fourteen substituents to the C-4 or equivalent position of the 6 templates via either a sulfur or oxygen linkage, to yield a total of 168 derivatives in the second method developed.

5.1.5 Initial docking and scoring method

The initial docking and scoring method was used to investigate the S-linked templates. The method consisted of a consensus score between Molecular Mechanics Interaction (MMI), AutoDock 3.05 and SCORE 2.01. AutoDock can be utilised as a docking or scoring program. Within this thesis the docking function is defined as flexible AutoDock and the scoring function is defined as static AutoDock (no flexible torsions of the ligand allowed).

Initially the ligands are positioned in the active site by AutoDock.¹⁰² AutoDock is a molecular docking program. That is, it is utilised to dock small molecules into the active site of a macromolecule, such as a protein. The program uses a genetic algorithm to reduce the time to search the conformational space within an active site. The input files for AutoDock were prepared by InsightII. Hydrogens were added to the protein and ligand at pH 7. The potential and charges were set for both the ligand and protein using the CVFF force-field. If a different ligand is required in the active site, the new ligand was superimposed over the existing ligand. The protein was saved in the Biosym archive format. The ligand was also saved in the Biosym archive format, but transformed with respect to the protein. In the UNIX shell, the resulting Biosym format files were converted to AutoDock pdbq files by the cartopdbq script included with AutoDock. The protein was solvated by the addsol script. The bonds of the ligand were defined by the pdbqtobnd script. The movable torsions of the ligand were defined by program Autotors. The ring of Neu5Ac2en was the cycle of root atoms. The other single bonds were allowed to rotate around the bond axis by Autodock with the exception of the bond between the carboxyl carbon and the ring carbon (C1 and C2). The AutoGrid and Autodock parameter files were created by mkgpf3 and mkdpf3 respectively and can be viewed in Appendix A. The parameters used were as default with the exception of the docking of sialyllactose; the grid spacing was increased to 0.475 from 0.375 and the number of points per axis was increased to 100 from 60. The scoring grid was created by AutoGrid, followed by the dock procedure by Autodock. The results were split into individual Mol2 files by the script bb, and viewed using InsightII.¹⁰² On a SGI Octane processor the AutoGrid calculation takes approximately

10 minutes and the Autodock calculation also takes approximately 10 minutes. After the initial placement the ligands are minimised in the active site by Discover 2.98.

The minimisation of the protein is utilised to shift the atoms of a protein to the nearest local minimum, not necessarily the global minimum. Minimisation relieves the strain placed on residues within the structure, and can help to relieve some errors found in the crystal structure.⁹⁹ The complex generated by flexible AutoDock analysis, was used for minimisation. Hydrogens were added to the protein and ligand residues at pH 7. The residue name of the ligand was set via the Biopolymer module within InsightII. The charge groups of the ligands were broken into smaller charge groups to satisfy Discover parameter requirements. All potentials and charges were set, using the consistent valence force-field (CVFF). The protein and the ligand molecules were joined into an assembly, by association. The ligand was added first, followed by the protein. The assembly was saved in Biosym car format. The input file can be viewed in Appendix A. Discover was run as a standalone application with the CVFF force-field, nice number of 19, on a single processor. Energy minimisation by both a distance dependant and a constant dielectric (80, to simulate a water environment) field were performed. The final structure generated by energy minimisation was compared to the original starting structure by superimposition by the protein backbone in Insight II.⁹⁹ On a SGI Octane processor the calculations took approximately 4.5 hours.

Once minimisation was performed the ligands minimised structure was statically scored used AutoDock, MMI and SCORE. The preparation of the ligands for static scoring with AutoDock was performed with InsightII. The assembly was disassociated and the potential and charges were set for both the ligand and protein using the CVFF force-field. The protein was saved in the Biosym archive format. The ligand was also saved in the Biosym archive format, but transformed with respect to the protein. In the UNIX shell, the resulting Biosym format files were converted to AutoDock pdbq files by the cartopdbq script included with AutoDock. The protein was solvated by the addsol script. The AutoGrid and Autodock parameter files were created by mkgpf3 and mkdpf3 respectively. The scoring grid was created by AutoGrid, followed by the dock procedure by Autodock. Autodock was run in command mode with the `-c` flag. Once the Autodock prompt was ready the command `epdb "ligand name"`, followed by the stop command to evaluate the static docked energy.¹⁰² The AutoGrid calculation takes about 10 minutes and the Autodock takes about 1 minute on a SGI Octane processor.

MMI was executed using the output file from the minimisation (the .cor file). The MMI scoring routine is performed within Discover and as such has an input file, which can be seen in Appendix A. The job takes about 2 minutes on an SGI Octane processor.

SCORE⁸⁶ is a program that utilises a non-forcefield empirical method to estimate the binding affinity of a protein-ligand complex. The interactions of each atom of a ligand with the protein were classified as good, bad or neutral by SCORE. The classification is based on the atomic binding score. A score greater than 0.1 is defined as a good interaction, less than -0.1 is defined as a bad interaction, while between -0.1 and 0.1 is classified as a neutral interaction.⁸⁶ The input files for SCORE were prepared using InsightII. The minimised complex was disassociated in InsightII and the minimised protein was exported as a PDB file and the minimised ligand was exported as a SYBYL Mol2 file. The Tripos atom types of the carboxyl oxygen atoms of the ligands were edited from O.2 to O.co2. The program was executed at the UNIX prompt specifying the protein and ligand. The calculation took approximately 15 seconds. The structures generated by SCORE were coloured by interactions: Green for good, Blue for neutral and Red for bad. The accompanying log file generated by SCORE contains the numerical results.⁸⁶

The method outlined above was used for the initial 54 derivatives constructed, it was also used to score the previously designed C4 derivatives in both NDV HN and the Holzer *et al.*⁶⁹ derivatives into the hPIV-2 HN homology model. The initial flexible AutoDock method was used for the docking of Neu5Ac2en into the hPIV HN homology models, it was also used for the initial validation of the AutoDock program, and it was also used dock sialyllactose with NDV HN.

5.1.6 Second docking and scoring method

The second docking and scoring method was used to investigate the *S*-linked and *O*-linked templates. The new method was designed with a few points in mind. Firstly, the use of AutoDock 3.05 would involve the use of Gasteiger and Kollman charges, this is because the program was designed and validated using these charge sets. In addition to the new charge sets, the number of energy evaluations was increased from 250,000 to 1,000,000, this allows for more conformational searching to take place and as such generate ten structures that are tighter in terms of final docked energy. The second point was the use of AMBER 7, the AMBER force-field also include the generalised

amber force-field or gaff. This force-field is designed for small ligands and can be customised for individual ligands, as such this force-field is better suited to parameterise the ligands designed as they are either carbohydrate based or carbohydrate mimetics, the CVFF by comparison was developed with DNA or amino acids in mind. The third point is the redesign of the MMI score, originally this score calculated the pair-wise atomic interactions of the ligand with the active site. The MM-PBSA (Amber 7 and ABPS 0.3.0) score calculates this as well as the contribution of energy of solvation. The fourth point involves the use of X-Score 1.2 (to replace SCORE), X-Score is a better parameterised (in the case of sulfur linkages) non-empirical force-field than the one implemented in SCORE. So the second method consists of a consensus score created between the programs MM-PBSA, AutoDock¹⁰² and X-Score. Globally the method is the same, with initial placement of the ligands followed by minimisation and finished with consensus scoring.

Initially the ligands are positioned in the active site by AutoDock.¹⁰² The input files for AutoDock¹⁰² were initially prepared by InsightII and finalised with AutoDock Tools 1.1.1 (ADT).¹⁸⁶ Within ADT, Gasteiger charges were applied to the ligand and the non-polar hydrogens (with charges) were merged into the attached parent atoms. Polar hydrogens were left as is. The ligands torsion tree was automatically assigned within ADT (the root atoms were manually checked for suitability, and amide bond was checked to be non-rotatable), with the resultant pdbq file automatically saved. The protein had United Kollman charges applied and then the non-polar hydrogens were merged. Solvation parameters were added to the protein and the protein was saved in AutoDock PDBQS format. The AutoGrid and Autodock parameter files were created by the respective menus and can be viewed in Appendix A. The parameters used were as default with the exception of the number of energy evaluations in the genetic algorithm which were increased to 1,000,000. The scoring grid was created by AutoGrid, followed by the dock procedure by Autodock, once again executed via appropriate menus. The results were split into individual Mol2 files by the script bb (supplied with AutoDock 3.05), and viewed using InsightII,¹⁰² or alternatively the results can be viewed with ADT. On a SGI Octane processor the AutoGrid calculation takes approximately 5 minutes and the Autodock calculation also takes approximately 30 minutes. After the initial placement the ligands are minimised in the active site by AMBER (version 7).¹⁰⁰

The minimisation of the ligands complexed with the protein was performed with AMBER. Initially the ligand in prepared in a SYBYL Mol2 file, Antechamber (part of AMBER) was run on the ligands to prepare AMBER Prep files and PDB files of the ligands. Antechamber was executed to perform MOPAC (version 5.09mn) charge calculations on the ligands with the total charge of the system being -1. The resultant AMBER Prep file was checked by the utility Parmchk to generate a gaff force-field modification file. The protein file was prepared by Reduce, the disulfide-linked cysteine residues were noted in addition to the protonation states of the various histidine residues. In the output PDB file from Reduce the disulfide-linked cysteine residues were renamed CYX (from CYS) and the protonated histidine residues were renamed either HIS, HID, HIE, or HIP, depending on protonation state. In the utility Leap (xleap or tleap) the gaff force-field was loaded, in addition of the AMBER 99 force-field and the gaff force-field modification file (generated by Parmchk). The Prep and PDB files of the ligand were loaded and checked, the protein PDB file was also loaded and checked, the disulfide linkages were also defined. An AMBER topology (.top) and co-ordinate (.crd) file were output for each of the ligands and protein files. In addition to these the ligand and protein were combined into a non-covalently associated complex and also output as a topology and co-ordinate file. The minimisation process was started at the UNIX prompt by the use of sander (part of AMBER). The minimisation of one complex by sander on a SGI Octane processor takes between 1 and 4 hours.

Once minimisation was performed the ligands minimised structure was statically scored used AutoDock, MM-PBSA (AMBER and APBS) and X-Score. The preparation of the ligands for static scoring with AutoDock was performed with the utilities ptraj (part of AMBER), and antechamber. Ptraj was used to obtain a PDB file from the output AMBER restart file, this file after some modification (AMBER atom names for hydrogens) was read into InsightII, and output as a PDB file (fixes Brookhaven file conventions), this file was suitable for reading into ADT. The script for ptraj can be found in Appendix A. Antechamber was used to convert the AMBER Restart file into a SYBYL mol2 file of the ligand using the initial topology file output at the start of the minimisation preparation. This mol2 file requires modification to make it readable by ADT. Modifications included the changing of AMBER gaff atom types into Tripos standard atom types, the redefinition of the carboxyl bond type, and the removal of all charge records. Within ADT, Gasteiger charges were applied to the ligand and the non-polar hydrogens (with charges) were merged into the attached parent

atoms. Polar hydrogens were left as is. The ligands torsion tree was automatically assigned within ADT and all rotatable torsions were removed from the ligand, with the resultant pdbq file saved. The protein had United Kollman charges applied and then the non-polar hydrogens were merged. Solvation parameters were added to the protein and the protein was saved in AutoDock PDBQS format. The AutoGrid and Autodock parameter files were created by the respective menus and can be viewed in Appendix A. The scoring grid was created by AutoGrid, followed by the dock procedure by Autodock. Autodock was run in command mode at the system prompt with the `-c` flag. Once the Autodock prompt was ready the command `epdb "ligand name"`, followed by the stop command to evaluate the static docked energy.¹⁰² The AutoGrid calculation takes about 5 minutes and the Autodock takes about 1 minute on a SGI Octane processor.

The MM-PBSA calculation calculates the free energies of binding or it is used to evaluate the absolute free energies of molecules in solution.¹⁰⁰ The force-field modification (from minimisation protocol), the generalised amber force-field, AMBER99 force-field and ligand PDB/Prep and protein PDB files (also from minimisation protocol) were processed by Xleap to generate AMBER topology, co-ordinate and pdb files for ligand, protein and non covalent complex. The program ptraj (available with AMBER 7) was used to convert minimised co-ordinates to mdcrd format. The input file for MM-PBSA was created (seen in Appendix A) and executed using `mm_pbsa.pl` script available with AMBER 7 to make input files for MM-PBSA calculation. The molecular mechanics (MM) part of the calculation uses sander to compute the MM interaction energy. Sander was used on the ligand, protein and complex separately. The second part of the calculation uses APBS to calculate the solvation energy with respect to accessible surface area. The calculation requires the creation of a pdb file with charge and radius terms, the program ptraj was used to create the required files for the ligand, protein and complex. APBS was executed using the `apbs` input file (`bind_energy.in`, available in Appendix A) to perform the calculation. The results were processed using a perl script (`array2.perl`, available in Appendix A, perl version 5.004_05), to calculate and collate the MM and APBS results into the final MM-PBSA score.

X-Score is a program that utilises a non-forcefield empirical method to estimate the binding affinity of a protein-ligand complex. The input files for X-Score were prepared

using. The ligand was docked in the active site of the protein by manual superimposition over Neu5Ac2en. The protein was exported as a PDB file and the ligand was exported as a SYBYL Mol2 file. The Tripos atom types of the carboxyl oxygen atoms of the ligands were edited from O.2 to O.co2. The program was executed at the UNIX prompt specifying the protein and ligand. The calculation took approximately 30 seconds. The structures generated by X-Score were coloured by interactions: Green for good, Blue for neutral and Red for bad. The accompanying log file generated by X-Score contains the numerical results.⁸⁶

The method outlined above was used for the complete set of 168 derivatives constructed.

5.1.7 *In Silico* Screening

5.1.7.1 DOCK

The protein was prepared for calculations by removing the ligand and water residues and defining a 10 Å subset around the active site. The active site subset was undisplayed, and the remaining protein structure was saved as exclude.pdb (the file name is required by DOCK 4⁹⁶). The ligand was saved as a Mol2 file. The active site was surfaced using the program DOCK_interface which uses the QCPE Connolly surface program. Sphgen was used to generate nodes of protein-ligand interaction. The spheres clusters generated by sphgen were processed to a PDB file by the program showsphere. The PDB file was viewed using InsightII for correct spheres, any erroneous spheres were removed from the sphere file generated by sphgen and showsphere was run again until the PDB file contained the spheres that defined the ligand in the active site. A box defining the spheres was generated using showbox. Once the active site had been characterised the electrostatic potential was analysed using Grid (available with DOCK 4). Grid was run inter-actively to generate the input file. The input file can be seen in Appendix A. The ligand was docked into the characterised active site by DOCK. DOCK was run inter-actively to generate the input file.⁹⁶ The input file can be viewed in Appendix A. The results were the viewed in InsightII.

5.1.8 Miscellaneous programs

5.1.8.1 UNIX

The version of UNIX used on the SGI Octane system for all calculations was IRIX64 6.5.

5.1.8.2 *InsightII*

All visualisation was performed by *InsightII*,¹²³ and the contained Builder, Biopolymer and DelPhi modules.

5.2 Synthesis experimental

5.2.1 General methods

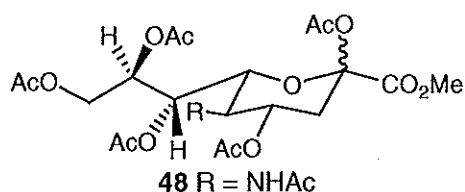
¹H and ¹³C spectra were recorded using a Brüker Avance 300 MHz spectrometer. Chemical shifts are given in ppm relative to the solvent used (CDCl₃: 7.26, 77.0, CD₃OD: 3.36, 37.0, D₂O: 4.64, for ¹H and ¹³C respectively). Multiplicities are denoted as s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublets), ddd (doublet of doublet of doublets), br (broad) and app (apparent). Chemical shifts indicated with a multiplication (×) descriptor correspond to more than one signal of the same functionality. Two-dimensional ¹H-¹H-COSY and ¹H-¹³C- HSQC NMR spectra were obtained where necessary, in order to assist with spectral assignment. Reactions were monitored by TLC (Merck silica gel plates GF₂₅₄, cat. #1.05554) and products were generally purified by flash chromatography using Merck silica gel 60 (0.040-0.063mm, cat. #1.09385). TLC's were typically visualized using 5% H₂SO₄ in EtOH followed by heating to 200°C. Elemental analyses were carried out by the microanalysis service of the Department of Chemistry at the University of Queensland and were recorded on a Carlo Erba Elemental Microanalyser, model 1106. Electrospray ionization (ESI) low resolution mass spectra (LRMS) were recorded on a Brüker Daltonics® Esquire 3000 Ion-Trap LC MS, using the positive or negative mode (as indicated) with samples introduced at 180µL/h. High resolution mass spectrometry (HRMS) were carried out by the mass spectral service of the Department of Chemistry at the University of Queensland and were recorded on a Finnigan MAT 900 XL Trap with a Finnigan API III sprayer. Infrared (IR) spectra were recorded on a Brüker Optik Vector 22 instrument using OPUS (version 3.01) software. IR spectra were obtained in neat CHCl₃ solutions between KBr plates. Deprotected products were generally purified by high performance

liquid chromatography (HPLC) using a Phenomenex C₁₈ reverse phase column (Semi-preparative) unless otherwise stated. All commercial solvents (MeOH, EtOAc, hexanes, CHCl₃, CH₂Cl₂) were distilled prior to use. Dried solvents were distilled under N₂ according to Perrin and Armarego.¹⁸⁷ High resolution mass spectroscopy of some of the final compounds indicated that a high amount of potassium ions were present (as indicated in the mass spectrum). It is postulated that they are present due to incomplete neutralisation of the final compounds.

5.2.2 Experimental methods

5.2.2.1 4-*S*-alkylated Neu5Ac2en derivatives

Methyl 5-acetamido-2,4,7,8,9-penta-*O*-acetyl-3,5-dideoxy-D-glycero- α/β -D-galactono-2-ulopyranosonate, 48.



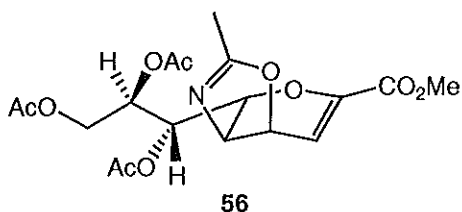
A mixture of Neu5Ac (**1**, 10 g, 30.6 mmol) and Amberlite IR-120(H⁺) resin (5 g) in dry MeOH (600 mL) was stirred at room temperature for 2 days. Reaction completion was determined by TLC (15:4:0.5 *i*PrOH:H₂O:AcOH, *R*_f = 0.5). The reaction mixture was filtered and concentrated under reduced pressure to yield the methyl ester as a pale brown solid (9.1g, 92%). A mixture of the methyl ester **50** (8.0 g, 26.7 mmol), and acetic anhydride (25 mL) in pyridine (40 mL) was stirred at room temperature for 4 days. Reaction completion was determined by TLC (EtOAc, *R*_f 0.33). The reaction mixture was concentrated under reduced pressure to yield a brown oily residue. Column chromatography (EtOAc) yielded a mixture of the α and β anomers, approximately 1:1 by ¹H NMR (13.5g, 98%). ¹H NMR data is in accordance with that reported in the literature.¹⁷³

Methyl 5-acetamido-2,4,7,8,9-penta-*O*-acetyl-3,5-dideoxy-D-glycero- α -D-galactono-2-ulopyranosonate. ¹H NMR (300 MHz, CDCl₃): δ 1.84 (3H, s, NAc), 1.98 (6H, s, 2 \times OAc), 1.94 (1H, app t, *J*_{3a,3e} 13.2, *J*_{3a,4} 13.2, H-3a), 2.04, 2.05, 2.08 (9H, 3 \times s, 3 \times OAc), 2.52 (1H, dd, *J*_{3e,3a} 13.2, *J*_{3e,4} 4.8 Hz, H-3e), 3.70 (3H, s, CO₂Me), 4.03 (1H, dd, *J*_{9,8} 6.3, *J*_{9,9'} 12.6 Hz, H-9), 4.12 (1H, ddd, *J*_{5,4} = *J*_{5,NH} = *J*_{5,6} = 10.2 Hz, H-5), 4.33 (1H, dd, *J*_{9',8} 2.7, *J*_{9',9} 12.6 Hz, H-9'), 4.65 (1H, dd, *J*_{6,5} 10.2, *J*_{6,7} 2.4 Hz, H-6),

4.97 (1H, ddd, $J_{4,3a}$ 13.2, $J_{4,3e}$ 4.8, $J_{4,5}$ 10.2 Hz, H-4), 5.14 (1H, ddd, $J_{8,7}$ 6.3, $J_{8,9}$ 6.3, $J_{8,9'}$ 2.7 Hz, H-8), 5.33 (1H, dd, $J_{7,6}$ 2.4, $J_{7,8}$ 6.3 Hz, H-7), 5.96 (1H, br d, $J_{NH,5}$ 10.2 Hz, NH)

Methyl 5-acetamido-2,4,7,8,9-penta-*O*-acetyl-3,5-dideoxy-D-glycero- β -D-galacto-non-2-ulopyranosonate ^1H NMR (300 MHz, CDCl_3): δ 1.82 (3H, s, NAc), 1.96, 1.97, 1.99, (9H, 3 \times s, 3 \times OAc), 2.02 (1H, dd, $J_{3a,3e}$ 13.2, $J_{3a,4}$ 12.9, H-3a), 2.07 (6H, s, 2 \times OAc), 2.47 (1H, dd, $J_{3e,3a}$ 13.2, $J_{3e,4}$ 5.1 Hz, H-3e), 3.70 (3H, s, CO_2Me), 4.01-4.10 (3H, m, H-5, H-6, H-9), 4.43 (1H, dd, $J_{9',8}$ 2.7, $J_{9',9}$ 12.6 Hz, H-9'), 5.00 (1H, ddd, $J_{8,7}$ 5.1, $J_{8,9}$ 6.9, $J_{8,9'}$ 2.7 Hz, H-8), 5.18 (1H, ddd, $J_{4,3a}$ 11.4, $J_{4,3e}$ 5.1, $J_{4,5}$ 10.2 Hz, H-4), 5.30 (1H, dd, $J_{7,6}$ 1.8, $J_{7,8}$ 5.1 Hz, H-7), 5.39 (1H, br d, $J_{NH,5}$ 9.6 Hz, NH).

2-Methyl-(methyl 7,8,9-tri-*O*-acetyl-2,6-anhydro-3,5-dideoxy-D-glycero-D-talo-non-2-enonate)-[4,5-d]-2-oxazoline, **56.**



Method 1: Reaction with $\text{BF}_3 \cdot \text{Et}_2\text{O}$

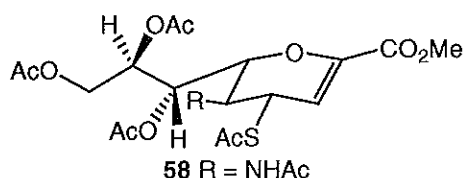
At room temperature a solution of methyl 5-acetamido-2,4,7,8,9-penta-*O*-acetyl-3,5-dideoxy-D-glycero- α/β -D-galacto-non-2-ulopyranosonate, **48**, (200 mg, 0.39 mmol) in dry DCM (3 mL) and dry MeOH (16 μL) and $\text{BF}_3 \cdot \text{Et}_2\text{O}$ (0.5 mL, 3.98 mmol) was stirred for 16 hours under N_2 . TLC (EtOAc, R_f 0.42) showed the development of a UV active spot. The reaction mixture was cooled in ice and dry NEt_3 (1 mL) was added and the reaction stirred for 10 mins. The neutralised reaction mixture was concentrated to dryness and then taken up in EtOAc (5 mL). The organic phase was extracted with 0.1M NaHCO_3 (3 \times 5 mL), and evaporated under reduced pressure to yield the crude oxazoline **56** as a slightly brown foam (131 mg, 82%).

Method 2: Reaction with TMSOTf

According to the method of Chandler *et al.*¹⁴⁷ a stirred solution of **48** (200 mg, 0.39 mmol) in EtOAc (3 mL) at 30°C under N_2 was added TMSOTf (210 μL , 1.16 mmol) dropwise. Over 20 minutes the temperature was raised to 52°C, then the reaction was left to stir for 2.5 hours. TLC (EtOAc, R_f 0.42) at 2.5 hrs showed the reaction was complete. The reaction mixture was then cooled in ice and 600 μL of dry NEt_3 was added to the reaction, which was then stirred for 10 minutes. The reaction mixture was

concentrated under reduced pressure. The brown residue was taken-up in 5 mL EtOAc and extracted with (3 × 5 mL) 0.1M NaHCO₃. The organic layer was concentrated under reduced pressure, dissolved in 1 mL DCM and re-concentrated to afford the title compound **56** as a off-white foam (137 mg, 85%). ¹H NMR data is in accordance with that reported in the literature.¹⁴⁷ ¹H NMR (300 MHz, CDCl₃): δ 1.99 (3H, s, CH₃), 2.01, 2.05, 2.14 (9H, 3 × s, 3 × OAc), 3.43 (1H, dd, *J*_{6,5} 9.9, *J*_{6,7} 2.7 Hz, H-6), 3.72 (3H, s, CO₂Me), 3.95 (1H, t, *J*_{5,4} 9.0, *J*_{5,6} 9.0 Hz, H-5), 4.21 (1H, app q, *J*_{9,8} 6.3, *J*_{9,9'} 12.6 Hz, H-9), 4.59 (1H, dd, *J*_{9,8} 2.4, *J*_{9,9'} 12.6 Hz, H-9'), 4.82 (1H, dd, *J*_{4,3} 3.9, *J*_{4,5} 9.0 Hz, H-4), 5.43 (1H, ddd, *J*_{8,7} 6.0, *J*_{8,9} 6.3, *J*_{8,9'} 2.4 Hz, H-8), 5.62 (1H, dd, *J*_{7,6} 2.7, *J*_{7,8} 6.0 Hz, H-7), 6.37 (1H, d, *J*_{3,4} 3.9 Hz, H-3).

Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,5-dideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **58.**



Method 1: According to von Itzstein *et al.*¹¹⁷ a solution of 2-methyl-(methyl 7,8,9-tri-*O*-acetyl-2,6-anhydro-3,5-dideoxy-D-glycero-D-talo-non-2-enonate)-[4,5-d]-2-oxazoline, **56**, (112 mg, 0.27

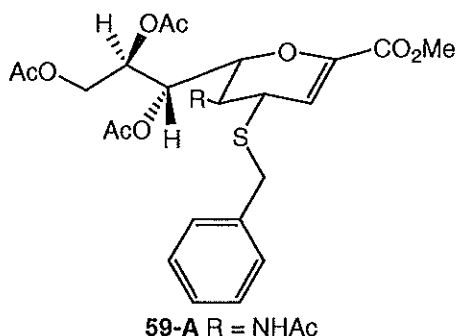
mmol) and thiolacetic acid (100 μL, 1.39 mmol) in dry DMF (2 mL) was stirred for 20 hours at 83°C under N₂. After this time, TLC (EtOAc, *R*_f 0.38) showed complete consumption of starting material (*R*_f 0.46). The reaction mixture was evaporated and the residue triturated with ether (2 × 10 mL). The resultant viscous yellow liquid was purified by column chromatography (20:1 DCM/MeOH) to afford a pale yellow solid as a 1 (product): 0.17 (*epi* SAc) mixture (65 mg, 49%). The two epimers were inseparable by column chromatography. The ¹H NMR spectrum of the major component was in agreement with that reported in the literature.¹¹⁷

Method 2: A solution of **56**, (2 g, 4.84 mmol) and thiolacetic acid (518 μL, 7.26 mmol) in dry DMF (30 mL) was stirred for 20 hours at 83°C under N₂. After this time, TLC (EtOAc, *R*_f 0.38) showed complete consumption of starting material (*R*_f 0.46). The reaction mixture was evaporated and the residue triturated with ether (2 × 10 mL). The resultant viscous yellow liquid was purified by column chromatography (20:1 DCM/MeOH) to afford a pale yellow solid as a 1 (product): 0.10 (*epi* SAc) mixture (2 g, 85%). The two epimers were inseparable by column chromatography.

Method 3: A solution of **56** (109 mg, 0.22 mmol) and thiolacetic acid (25 μ L, 0.34 mmol) in dry acetonitrile (2 mL) was stirred for 20 hours at 83°C. After this time, TLC (EtOAc, R_f 0.38) showed complete consumption of starting material (R_f 0.46). The reaction mixture was evaporated and the residue triturated with ether (2 \times 10 mL). The resultant viscous yellow liquid afforded the product as a 1 (product): 0.17 (*epi* SAc) mixture (61 mg, 47%). The two epimers were inseparable by column chromatography.

Method 4: A solution of **56** (108 mg, 0.22 mmol) in neat thiolacetic acid (3 mL) was stirred for 20 hours at 83°C. After this time, TLC (EtOAc, R_f 0.38) showed complete consumption of starting material (R_f 0.46). The reaction mixture was evaporated and the residue triturated with ether (2 \times 10 mL). The resultant viscous yellow oil afforded the product as a 1 (product): 0.24 (*epi* SAc) mixture (114 mg, 89%). The two epimers were inseparable by column chromatography. NMR data of major component: ^1H NMR (300 MHz, CDCl_3): δ 1.89 (3H, s, NAc), 2.05, 2.06, 2.10 (9H, 3 \times s, 3 \times OAc), 2.35 (3H, s, SAc), 3.70 (3H, s, CO_2Me), 4.15 (1H, app q, $J_{9,8}$ 2.7, $J_{9,9}$ 12.3 Hz, H-9), 4.20 – 4.37 (3H, m, H-4, H-5, H-6), 4.66 (1H, dd, $J_{9,8}$ 7.2, $J_{9,9}$ 12.3 Hz, H-9'), 5.30 (1H, ddd, $J_{8,7}$ 5.1, $J_{8,9}$ 7.2, $J_{8,9}$ 2.7 Hz, H-8), 5.48 (1H, dd, $J_{7,6}$ 1.5, $J_{7,8}$ 5.1 Hz, H-7), 5.57 (1H, br d, $J_{\text{NH},5}$ 8.7 Hz, NH), 5.94 (1H, d, $J_{3,4}$ 2.4 Hz, H-3). Partial ^1H NMR spectrum of minor component: ^1H NMR (300 MHz, CDCl_3): δ 1.83 (3H, s, NAc), 2.01, 2.02, 2.05 (9H, 3 \times s, 3 \times OAc), 2.36 (3H, s, SAc), 3.77 (3H, s, CO_2Me), 4.46 (1H, t, $J_{4,3}$ 5.7, $J_{4,5}$ 5.1, H-4), 6.05 (1H, d, $J_{3,4}$ 5.7, H-3).

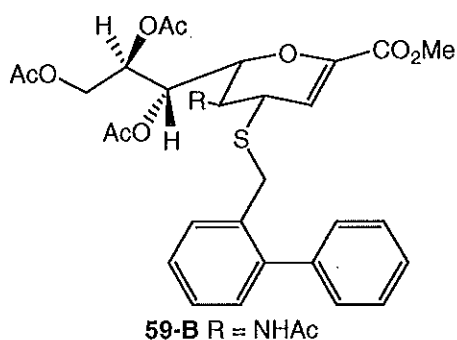
Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-4-*S*-benzyl-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **59-A.**



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,4,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **58**, (200 mg, 0.41 mmol) as a 1:0.18 epimeric mixture in dry DMF (8 mL) with 3 Å molecular sieves under Ar was added benzyl bromide (80 μ L, 0.67 mmol) and dry triethylamine (60 μ L, 0.43 mmol). The mixture was left to stir for 10 minutes, following this hydrazine acetate (280 mg, 3.04 mmol) was added and the

reaction mixture was left to stir at room temperature for 8 hours. After this time, TLC (EtOAc, R_f 0.62) showed complete conversion of starting material (R_f 0.35). EtOAc (10 mL) and water (10 mL) were added, the organic phase was washed with water and sat. NaCl (10 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was subject to column chromatography on silica (EtOAc) to give the purified title compound **59-A** as a 1:0.18 epimeric mixture (by ^1H NMR spectroscopy) as a pale yellow foam (150 mg, 68%). The two epimers were inseparable by column chromatography. ν_{max} 1266, 1421, 1652, 1740, 2359, 2989, 3056, 3426 cm^{-1} . NMR data of major component: ^1H NMR (300 MHz, CDCl_3): δ 1.94 (3H, s, NAc), 2.04, 2.05, 2.11 (9H, $9 \times$ s, $3 \times$ OAc), 3.47 (1H, dd, $J_{4,3}$ 3.0, $J_{4,5}$ 9.0 Hz, H-4), 3.74 (1H, d, J_{gem} 12.6 Hz, PhCH_2), 3.77 (3H, s, CO_2Me), 3.82 (1H, d, J_{gem} 13.2 Hz, PhCH_2), 4.02 (1H, app q, $J_{5,4}$ 9.0, $J_{5,\text{NH}}$ 9.3, $J_{5,6}$ 9.0 Hz, H-5), 4.17 (1H, dd, $J_{9,8}$ 6.9, $J_{9,9'}$ 12.3 Hz, H-9), 4.27 (1H, dd, $J_{6,5}$ 9.0, $J_{6,7}$ 3.0 Hz, H-6), 4.62 (1H, dd, $J_{9',8}$ 3.0, $J_{9',9}$ 12.3 Hz, H-9'), 5.34 (1H, ddd, $J_{8,7}$ 4.8, $J_{8,9}$ 6.9, $J_{8,9'}$ 3.0 Hz, H-8), 5.43 (1H, br d, $J_{\text{NH},5}$ 9.3 Hz, NH), 5.48 (1H, dd, $J_{7,6}$ 2.7, $J_{7,8}$ 4.8 Hz, H-7), 6.07 (1H, d, $J_{3,4}$ 3.0 Hz, H-3), 7.29-7.33 (5H, m, ArH). ^{13}C NMR (75.5 MHz, CDCl_3): δ 21.9, 20.8, 20.9 ($3 \times \text{OCOCH}_3$), 23.2 (NHCOCH_3), 33.8 (Ph-CH_2), 46.6 (C-5), 43.4 (C-4), 52.4 (COOCH_3), 62.1 (C-9), 68.1 (C-7), 71.1 (C-8), 76.7 (C-6), 112.7 (C-3), 128.5, 128.7, 128.8, 128.9, 129.1, 129.3 (Ph), 169.7 (NHCOCH_3), 170.1, 170.2, 170.3 ($3 \times \text{OCOCH}_3$). $^*\text{COOCH}_3$ and C-2 were not observed. LRMS m/z 560 ($[\text{M}+\text{Na}]^+$, 100%). HRMS calcd for $\text{C}_{25}\text{H}_{31}\text{NO}_{10}\text{S}$ ($\text{M}+\text{Na}$) 560.1561, found 560.1581. Partial ^1H NMR spectrum of minor component: ^1H NMR (300 MHz, CDCl_3): δ 3.44 (1H, t, $J_{4,3}$ 5.7, $J_{4,5}$ 5.4, H-4), 6.03 (1H, d, $J_{3,4}$ 5.7, H-3).

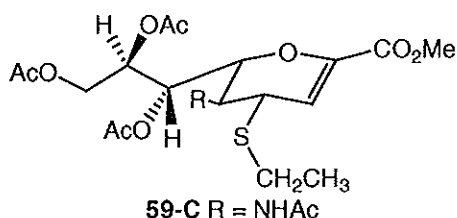
Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-2,3,5-trideoxy-4-*S*-(2-phenylbenzyl)-4-thio-D-glycero-D-galacto-non-2-enonate, 59-B.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,4,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **58**, (200 mg, 0.41 mmol) as a 1:0.16 epimeric mixture in dry DMF (8 mL) with 3 Å molecular sieves under Ar was added 2-phenylbenzyl bromide (120 μL , 0.66 mmol) and dry triethylamine (60 μL , 0.43 mmol). The mixture was left to stir for 10 minutes, following this hydrazine acetate (280 mg, 3.04 mmol) was added and

the reaction mixture was left to stir at room temperature for 8 hours. After this time, TLC (2:1 EtOAc/hexanes, R_f 0.54) showed complete conversion of starting material (R_f 0.28). EtOAc (10 mL) and water (10 mL) were added, the organic phase was washed with water and sat. NaCl (10 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was subject to column chromatography on silica (2:1 EtOAc/hexanes) to give the purified title compound **59-B** as a 1:0.16 epimeric mixture (by ^1H NMR spectroscopy) as a dark yellow foam (165 mg, 68%). The two epimers were inseparable by column chromatography. ν_{max} 1266, 1422, 1743, 2360, 2988, 3055 cm^{-1} . NMR data of major component: ^1H NMR (300 MHz, CDCl_3): δ 1.83 (3H, s, NAc), 1.96 (9H, s, $3 \times \text{OAc}$), 3.37 (1H, dd, $J_{4,3}$ 2.7, $J_{4,5}$ 9.3 Hz, H-4), 3.65 (2H, m, PhCH_2), 3.68 (3H, s, CO_2Me), 3.78 (1H, app q, $J_{5,4}$ 9.0 $J_{5,\text{NH}}$ 9.3, $J_{5,6}$ 9.0 Hz, H-5), 4.09 (1H, dd, $J_{9,8}$ 7.2, $J_{9,9'}$ 12.3 Hz, H-9), 4.18 (1H, dd, $J_{6,5}$ 9.6 $J_{6,7}$ 2.4 Hz, H-6), 4.59 (1H, dd, $J_{9',8}$ 2.7, $J_{9',9}$ 12.3 Hz, H-9'), 5.21 (1H, ddd, $J_{8,7}$ 4.8, $J_{8,9}$ 7.2, $J_{8,9'}$ 2.7 Hz, H-8), 5.35 (1H, dd, $J_{7,6}$ 2.4, $J_{7,8}$ 4.8 Hz, H-7), 5.50 (1H, br d, $J_{\text{NH},5}$ 9.3 Hz, NH), 5.78 (1H, d, $J_{3,4}$ 2.7 Hz, H-3), 7.12-7.48 (9H, m, $2 \times \text{ArH}$). ^{13}C NMR (75.5 MHz, CDCl_3): δ 20.8, 20.9, 21.0 ($3 \times \text{OCOCH}_3$), 23.3 (NHCOCH_3), 31.6 (Ph-CH_2), 36.8 (COOCH_3), 43.6 (C-4), 47.4 (C-5), 52.3 (COOCH_3), 62.1 (C-9), 68.1 (C-7), 71.2 (C-8), 76.9 (C-6), 112.5 (C-3), 127.6-130.5 ($2 \times \text{Ph}$), 169.6 (NHCOCH_3), 170.1, 170.2, 170.3 ($3 \times \text{OCOCH}_3$). $^*\text{COOCH}_3$ and C-2 were not observed. LRMS m/z 636 ($[\text{M}+\text{Na}]^+$, 100%). HRMS calcd for $\text{C}_{31}\text{H}_{35}\text{NO}_{10}\text{S}$ ($\text{M}+\text{Na}$) 636.1874, found 636.1875. Partial ^1H NMR spectrum of minor component: ^1H NMR (300 MHz, CDCl_3): δ 3.25 (1H, t, $J_{4,3}$ 5.7, $J_{4,5}$ 5.1, H-4), 5.91 (1H, d, $J_{3,4}$ 5.7, H-3).

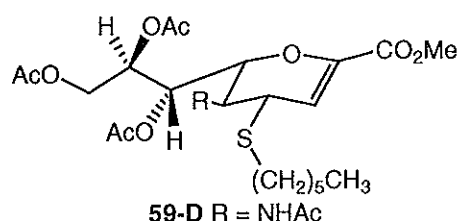
Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-2,3,5-trideoxy-4-*S*-ethyl-4-thio-*D*-glycero-*D*-galacto-non-2-enonate, 59-C.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,4,5-trideoxy-4-thio-*D*-glycero-*D*-galacto-non-2-enonate, **58**, (200 mg, 0.41 mmol) as a 1:0.14 epimeric mixture in dry DMF (8 mL) with 3 Å molecular sieves under Ar was added iodoethane (48 μL , 0.60 mmol) and dry triethylamine (60 μL , 0.43 mmol). The mixture was left to stir for 10 minutes, following this hydrazine acetate (280 mg, 3.04 mmol) was added and the reaction mixture was left to stir at room temperature for 8 hours.

After this time, TLC (2:1 EtOAc/hexanes, R_f 0.46) showed complete conversion of starting material (R_f 0.28). EtOAc (10 mL) and water (10 mL) were added, the organic phase was washed with water and sat. NaCl (10 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was subject to column chromatography on silica (2:1 EtOAc/hexanes) to give the purified title compound **59-C** as a 1:0.14 epimeric mixture (by ^1H NMR spectroscopy) as a pale yellow foam (118 mg, 61%). The two epimers were inseparable by column chromatography. ν_{max} 1266, 1422, 1439, 1513, 1650, 1742, 2307, 2362, 2988, 3056, 3427 cm^{-1} . NMR data of major component: ^1H NMR (300 MHz, CDCl_3): δ 1.16 (3H, t, CH_3), 1.91 (3H, s, NAc), 1.99 (6H, s, $2 \times \text{OAc}$), 2.06 (3H, s, OAc), 2.42-2.62 (2H, m, SCH_2), 3.47 (1H, dd, $J_{4,3}$ 3.0, $J_{4,5}$ 9.3 Hz, H-4), 3.72 (3H, s, CO_2Me), 3.99 (1H, app q, $J_{5,4}$ 9.3, $J_{5,\text{NH}}$ 9.3, $J_{5,6}$ 9.3 Hz, H-5), 4.10 (1H, dd, $J_{9,8}$ 7.2, $J_{9,9'}$ 12.6 Hz, H-9), 4.22 (1H, dd, $J_{6,5}$ 9.3, $J_{6,7}$ 3 Hz, H-6), 4.62 (1H, dd, $J_{9',8}$ 3.0, $J_{9',9}$ 12.3 Hz, H-9'), 5.26 (1H, ddd, $J_{8,7}$ 4.8, $J_{8,9}$ 7.2, $J_{8,9'}$ 3.0 Hz, H-8), 5.45 (1H, dd, $J_{7,6}$ 2.7, $J_{7,8}$ 4.5 Hz, H-7), 5.90 (1H, br d, $J_{\text{NH},5}$ 9.3 Hz, NH), 6.04 (1H, d, $J_{3,4}$ 3.0 Hz, H-3). ^{13}C NMR (75.5 MHz, CDCl_3): δ 14.4 (SCH_2CH_3), 20.9 (OCOCH_3), 23.4 (NHCOCH_3), 23.3 (SCH_2), 43.2 (C-4), 47.3 (C-5), 52.4 (COOCH_3), 62.1 (C-9), 68.1 (C-7), 71.3 (C-8), 77.1 (C-6), 112.9 (C-3), 169.9 (NHCOCH_3), 170.2, 170.3, 170.5 ($3 \times \text{OCOCH}_3$). $^*\text{COOCH}_3$ and C-2 were not observed. LRMS m/z 498 ($[\text{M}+\text{Na}]^+$, 100%). HRMS calcd for $\text{C}_{20}\text{H}_{29}\text{NO}_{10}\text{S}$ (M+K) 514.1149, found 514.1142. Partial ^1H NMR spectrum of minor component: ^1H NMR (300 MHz, CDCl_3): δ 3.43 (1H, t, $J_{4,3}$ 5.7, $J_{4,5}$ 5.4, H-4), 6.11 (1H, d, $J_{3,4}$ 5.7, H-3).

Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-2,3,5-trideoxy-4-*S*-hexyl-4-thio-D-glycero-D-galacto-non-2-enonate, 59-D.

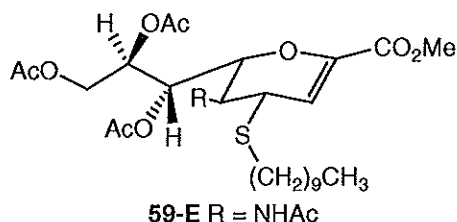


To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,4,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **58**, (200 mg, 0.41 mmol) as a 1:0.18

epimeric mixture in dry DMF (8 mL) with 3 Å molecular sieves under Ar was added bromohexane (84 μL , 0.60 mmol) and dry triethylamine (60 μL , 0.43 mmol). The mixture was left to stir for 10 minutes, following this hydrazine acetate (280 mg, 3.04 mmol) was added and the reaction mixture was left to stir at room temperature for 8 hours. After this time, TLC (2:1 EtOAc/hexanes, R_f 0.62) showed complete conversion

of starting material (R_f 0.28). EtOAc (10 mL) and water (10 mL) were added, the organic phase was washed with water and sat. NaCl (10 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was subject to column chromatography on silica (2:1 EtOAc/hexanes) to give the purified title compound **59-D** as a 1:0.18 epimeric mixture (by ^1H NMR spectroscopy) as a pale yellow foam (182 mg, 84%). The two epimers were inseparable by column chromatography. ν_{max} 1266, 1422, 1509, 1650, 1742, 2307, 3055, 3440 cm^{-1} . NMR data of major component: ^1H NMR (300 MHz, CDCl_3): δ 0.75 (3H, br. app. t, CH_3), 1.09-1.26 (6H, s, $3 \times \text{CH}_2$), 1.35-1.44 (2H, m, SCH_2CH_2), 1.83 (3H, s, NAc), 1.90 (6H, s, $2 \times \text{OAc}$), 1.98 (3H, s, OAc), 2.31-2.48 (2H, m, SCH_2) 3.40 (1H, dd, $J_{4,3}$ 2.7, $J_{4,5}$ 9.0 Hz, H-4), 3.65 (3H, s, CO_2Me), 3.87 (1H, app q, $J_{5,4}$ 9.3, $J_{5,\text{NH}}$ 9.3, $J_{5,6}$ 9.3 Hz, H-5), 4.03 (1H, dd, $J_{9,8}$ 7.2, $J_{9,9'}$ 12.6 Hz, H-9), 4.14 (1H, dd, $J_{6,5}$ 9.3, $J_{6,7}$ 2.7 Hz, H-6), 4.53 (1H, dd, $J_{9,8}$ 3.0, $J_{9,9'}$ 12.3 Hz, H-9'), 5.19 (1H, ddd, $J_{8,7}$ 5.1, $J_{8,9}$ 7.5, $J_{8,9'}$ 3.0 Hz, H-8), 5.37 (1H, dd, $J_{7,6}$ 2.7, $J_{7,8}$ 4.8 Hz, H-7), 5.92 (1H, br d, $J_{\text{NH},5}$ 9.0 Hz, NH), 5.97 (1H, d, $J_{3,4}$ 3.0 Hz, H-3). ^{13}C NMR (75.5 MHz, CDCl_3): δ 14.0 (CH_3), 20.7, 20.8, 20.9 ($3 \times \text{OCOCH}_3$), 23.2 (NHCOCH_3), 28.6, 29.3, 29.6 ($3 \times \text{CH}_2$), 30.1 (SCH_2CH_2), 31.3 (SCH_2CH_2), 43.3 (C-4), 47.3 (C-5), 52.3 (COOCH_3), 62.1 (C-9), 68.1 (C-7), 71.0 (C-8), 77.1 (C-6), 113.0 (C-3), 169.8 (NHCOCH_3), 170.1, 170.2, 170.3 ($3 \times \text{OCOCH}_3$). $^*\text{COOCH}_3$ and C-2 were not observed. LRMS m/z 554 ($[\text{M}+\text{Na}]^+$, 100%). HRMS calcd for $\text{C}_{24}\text{H}_{37}\text{NO}_{10}\text{S}$ ($\text{M}+\text{Na}$) 554.2036, found 554.2043. Partial ^1H NMR spectrum of minor component: ^1H NMR (300 MHz, CDCl_3): δ 3.36 (1H, t, $J_{4,3}$ 5.7, $J_{4,5}$ 5.4, H-4), 6.03 (1H, d, $J_{3,4}$ 5.7, H-3).

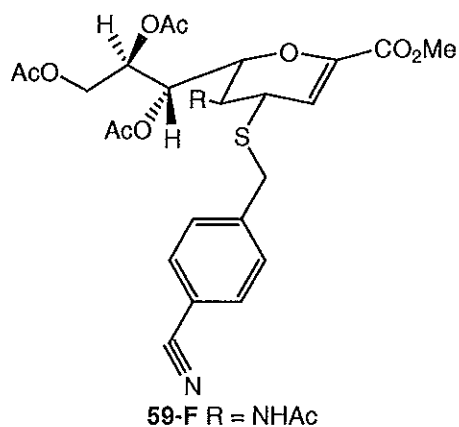
Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-4-*S*-decyl-2,3,5-trideoxy-4-thio-*D*-glycero-*D*-galacto-non-2-enonate, **59-E.**



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,4,5-trideoxy-4-thio-*D*-glycero-*D*-galacto-non-2-enonate, **58**, (200 mg, 0.41 mmol) as a 1:0.31 epimeric mixture in dry DMF (8 mL) with 3 Å molecular sieves under Ar was added bromodecane (124 μL , 0.60 mmol) and dry triethylamine (60 μL , 0.43 mmol). The mixture was left to stir for 10 minutes, following this hydrazine acetate (280 mg, 3.04 mmol) was added and the reaction mixture was left to stir at room temperature for 8 hours. After this time,

TLC (2:1 EtOAc/hexanes, R_f 0.6) showed complete conversion of starting material (R_f 0.28). EtOAc (10 mL) and water (10 mL) were added, the organic phase was washed with water and sat. NaCl (10 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was subject to column chromatography on silica (2:1 EtOAc/hexanes) to give the purified title compound **59-E** as a 1:0.31 epimeric mixture (by ^1H NMR spectroscopy) as a pale yellow foam (95 mg, 39%). The two epimers were inseparable by column chromatography. ν_{max} 1265, 1422, 1644, 2059, 2411, 3448 cm^{-1} . NMR data of major component: ^1H NMR (300 MHz, CDCl_3): δ 0.80 (3H, br app t, CH_3), 1.18 (14H, s, $7 \times \text{CH}_2$), 1.42-1.61 (2H, m, SCH_2CH_2), 1.91 (3H, s, NAc), 1.99 (6H, s, $2 \times \text{OAc}$), 2.06 (3H, s, OAc), 2.38-2.67 (2H, m, SCH_2), 3.48 (1H, dd, $J_{4,3}$ 3.0, $J_{4,5}$ 9.3 Hz, H-4), 3.72 (3H, s, CO_2Me), 3.92 (1H, app q, $J_{5,4}$ 9.3 $J_{5,\text{NH}}$ 9.3, $J_{5,6}$ 9.3 Hz, H-5), 4.13 (1H, dd, $J_{9,8}$ 7.2, $J_{9,9'}$ 12.3 Hz, H-9), 4.23 (1H, dd, $J_{6,5}$ 9.3, $J_{6,7}$ 3.0 Hz, H-6), 4.64 (1H, dd, $J_{9',8}$ 3.0, $J_{9',9}$ 12.6 Hz, H-9'), 5.27 (1H, ddd, $J_{8,7}$ 4.8, $J_{8,9}$ 7.5, $J_{8,9'}$ 2.7 Hz, H-8), 5.45 (1H, dd, $J_{7,6}$ 2.7, $J_{7,8}$ 4.8 Hz, H-7), 5.80 (1H, br d, $J_{\text{NH},5}$ 9.3 Hz, NH), 6.05 (1H, d, $J_{3,4}$ 3.0 Hz, H-3). ^{13}C NMR (75.5 MHz, CDCl_3): δ 14.2 (CH_3), 21.1 (OCOCH_3), 23.4 (NHCOCH_3), 29.4 (SCH_2), 29.4 (SCH_2CH_2), 29.0, 29.1, 29.2, 29.3, 29.4, 29.5, 29.6 ($6 \times \text{CH}_2$), 43.0 (C-4), 47.5 (C-5), 52.4 (COOCH_3), 62.1 (C-9), 68.1 (C-7), 71.0 (C-8), 77.0 (C-6), 113.0 (C-3), 169.6 (NHCOCH_3), 170.1, 170.2, 170.3 ($3 \times \text{OCOCH}_3$). $^*\text{COOCH}_3$ and C-2 were not observed. LRMS m/z 610 ($[\text{M}+\text{Na}]^+$, 100%). HRMS calcd for $\text{C}_{28}\text{H}_{45}\text{NO}_{10}\text{S}$ ($\text{M}+\text{Na}$) 610.2662, found 610.2649. Partial ^1H NMR spectrum of minor component: ^1H NMR (300 MHz, CDCl_3): δ 3.40 (1H, t, $J_{4,3}$ 5.7, $J_{4,5}$ 5.1, H-4), 6.11 (1H, d, $J_{3,4}$ 5.7, H-3).

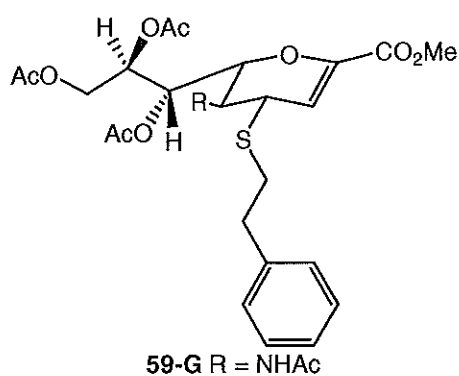
Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-4-*S*-(*p*-cyanobenzyl)-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, 59-F.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,4,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **58**, (200 mg, 0.41 mmol) as a 1:0.18 epimeric mixture in dry DMF (8 mL) with 3 Å molecular sieves under Ar was added α -bromo-*p*-tolunitrile (120 mg, 0.61 mmol) and dry triethylamine (60 μL , 0.432 mmol). The mixture

was left to stir for 10 minutes, following this hydrazine acetate (280 mg, 3.04 mmol) was added and the reaction mixture was left to stir at room temperature for 8 hours. After this time, TLC (2:1 EtOAc/hexanes, R_f 0.48) showed complete conversion of starting material (R_f 0.28). EtOAc (10 mL) and water (10 mL) were added, the organic phase was washed with water and sat. NaCl (10 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was subject to column chromatography on silica (2:1 EtOAc/hexanes) to give the purified title compound **59-F** as a 1:0.18 epimeric mixture (by ^1H NMR spectroscopy) as a pale yellow foam (168 mg, 80%). The two epimers were inseparable by column chromatography. ν_{max} 1266, 1422, 1747, 2306, 2412, 2988, 3055 cm^{-1} . NMR data of major component: ^1H NMR (300 MHz, CDCl_3): δ 1.90 (3H, s, NAc), 1.96, 1.97, 2.05 (9H, 3 \times s, 3 OAc), 3.38 (1H, dd, $J_{4,3}$ 3.0, $J_{4,5}$ 9.3 Hz, H-4), 3.70 (3H, s, CO_2Me), 3.76 (2H, m, PhCH_2), 4.00 (3H, m, H-5, H-6, H-9), 4.62 (1H, dd, $J_{9',8}$ 2.7, $J_{9',9}$ 12.3 Hz, H-9'), 5.25 (1H, ddd, $J_{8,7}$ 4.5, $J_{8,9}$ 7.2, $J_{8,9'}$ 2.7 Hz, H-8), 5.47 (1H, dd, $J_{7,6}$ 2.4, $J_{7,8}$ 4.5 Hz, H-7), 5.92 (1H, d, $J_{3,4}$ 3.0 Hz, H-3), 6.24 (1H, br d, $J_{\text{NH},5}$ 9.0 Hz, NH), 7.38 (2H, d, $J_{\text{ArH}-o}$, $J_{\text{ArH}-m}$ 8.7 Hz, ArH-*o*), 7.56 (2H, d, $J_{\text{ArH}-m}$, $J_{\text{ArH}-o}$ 8.7 Hz, ArH-*m*). ^{13}C NMR (75.5 MHz, CDCl_3): δ 20.7, 20.8, 21.0 (3 \times OCOCH_3), 23.1 (NHCOCH_3), 33.1 (Ph-CH_2), 44.1 (C-4), 45.9 (C-5), 52.4 (COOCH_3), 62.1 (C-9), 68.1 (C-7), 71.2 (C-8), 77.4 (C-6), 112.2 (C-3), 128.3, 128.4, 128.8, 129.4, 129.5, 129.6 (Ph), 169.8 (NHCOCH_3), 170.0, 170.1, 170.3 (3 \times OCOCH_3). $^*\text{COOCH}_3$ and C-2 were not observed. LRMS m/z 585 ($[\text{M}+\text{Na}]^+$, 100%). HRMS calcd for $\text{C}_{26}\text{H}_{30}\text{N}_2\text{O}_{10}\text{S}$ ($\text{M}+\text{Na}$) 585.1519, found 585.1522. Partial ^1H NMR spectrum of minor component: ^1H NMR (300 MHz, CDCl_3): δ 3.43 (1H, t, $J_{4,3}$ 5.7, $J_{4,5}$ 5.4, H-4), 5.89 (1H, d, $J_{3,4}$ 5.7, H-3).

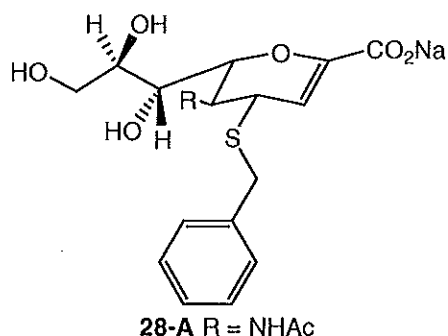
Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-2,3,5-trideoxy-4-*S*-(2-phenylethyl)-4-thio-D-glycero-D-galacto-non-2-enonate, 59-G.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-4-*S*-acetyl-2,6-anhydro-3,4,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **58**, (200 mg, 0.41 mmol) as a 1:0.22 epimeric mixture in dry DMF (8 mL) with 3 Å molecular sieves under Ar was added 2-bromoethylbenzene (80 μL , 0.58 mmol) and dry triethylamine (60 μL , 0.43

mmol). The mixture was left to stir for 10 minutes, following this hydrazine acetate (280 mg, 3.04 mmol) was added and the reaction mixture was left to stir at room temperature for 8 hours. After this time, TLC (2:1 EtOAc/hexanes, R_f 0.56) showed complete conversion of starting material (R_f 0.28). EtOAc (10 mL) and water (10 mL) were added, the organic phase was washed with water and sat. NaCl (10 mL), dried (Na_2SO_4), filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was subject to column chromatography on silica (2:1 EtOAc/hexanes) to give the purified title compound **59-G** as a 1:0.22 epimeric mixture (by ^1H NMR spectroscopy) as a pale yellow foam (133 mg, 59%). The two epimers were inseparable by column chromatography. ν_{max} 1266, 1371, 1422, 1679, 1740, 2306, 3056 cm^{-1} . NMR data of major component: ^1H NMR (300 MHz, CDCl_3) δ 1.88 (3H, s, NAc), 1.97 (6H, s, 2 \times OAc), 2.04 (3H, s, OAc), 2.68 – 2.90 (4H, m, SCH_2CH_2), 3.45 (1H, dd, $J_{4,3}$ 2.7, $J_{4,5}$ 9.3 Hz, H-4), 3.70 (3H, s, CO_2Me), 4.00 (1H, q, $J_{5,4}$ 9.3, $J_{5,\text{NH}}$ 9.3, $J_{5,6}$ 9.3 Hz, H-5), 4.10 (1H, dd, $J_{9,8}$ 7.2, $J_{9,9'}$ 12.3 Hz, H-9), 4.19 (1H, dd, $J_{6,5}$ 9.3, $J_{6,7}$ 2.7 Hz, H-6), 4.62 (1H, dd, $J_{9',8}$ 2.7, $J_{9',9}$ 12.3 Hz, H-9'), 5.27 (1H, ddd, $J_{8,7}$ 4.8, $J_{8,9}$ 7.5, $J_{8,9'}$ 3.0 Hz, H-8), 5.44 (1H, dd, $J_{7,6}$ 2.7, $J_{7,8}$ 4.8 Hz, H-7), 5.92 (1H, br d, $J_{\text{NH},5}$ 9.3 Hz, NH), 6.01 (1H, d, $J_{3,4}$ 2.7 Hz, H-3), 7.10–7.27 (5H, m, ArH). ^{13}C NMR (75.5 MHz, CDCl_3): δ 20.7, 20.8, 21.1 (3 \times OCOCH_3), 23.3 (NHCOCH_3), 30.5 ($\text{Ph-CH}_2\text{CH}_2$), 35.6 ($\text{Ph-CH}_2\text{CH}_2$), 43.4 (C-4), 47.2 (C-5), 52.4 (COOCH_3), 62.1 (C-9), 68.2 (C-7), 71.1 (C-8), 77.1 (C-6), 112.7 (C-3), 128.5 (Ph), 169.7 (NHCOCH_3), 170.0, 170.1, 170.2 (3 \times OCOCH_3). $^*\text{COOCH}_3$ and C-2 were not observed. LRMS m/z 574 ($[\text{M}+\text{Na}]^+$, 100%). HRMS calcd for $\text{C}_{26}\text{H}_{33}\text{NO}_{10}\text{S}$ ($\text{M}+\text{Na}$) 574.1723, found 574.1718. Partial ^1H NMR spectrum of minor component: ^1H NMR (300 MHz, CDCl_3): δ 3.39 (1H, t, $J_{4,3}$ 5.7, $J_{4,5}$ 5.4, H-4), 6.03 (1H, d, $J_{3,4}$ 5.7, H-3).

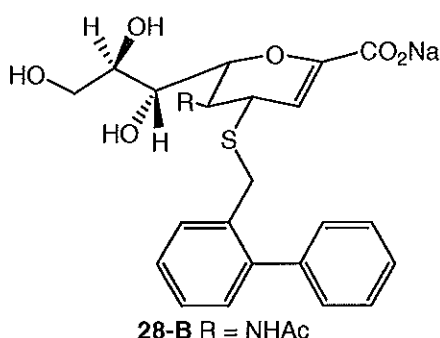
Sodium (5-acetamido-2,6-anhydro-4-S-benzyl-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-en)onate, 28-A.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-4-*S*-benzyl-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **59-A**, (100 mg, 0.19 mmol) in MeOH (5 mL) was added 0.1 M KOH (5 mL). The reaction mixture was left to stir at room temperature for 2 hours. TLC (7:2:1 EtOAc/MeOH/ H_2O , R_f 0.32) showed complete conversion of starting material (R_f 0.70).

The reaction mixture was acidified using Amberlite IR-120 (H^+) resin to pH 3 - 4, filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was resuspended in water and neutralised to pH 7.8 with aqueous NaOH and concentrated under reduced pressure to afford a dark yellow mass, 75 mg, > 100%. The crude product was purified by reverse-phase HPLC using isocratic 2% acetonitrile/98% H_2O , elution time 18.2 mins, and lyophilised to give **28-A** as a pale yellow fine powder (35 mg, 47%). The two epimers were inseparable by HPLC conditions (approx ratio: 1 eq:0.07 *epi*). NMR data of major component: 1H NMR (300 MHz, D_2O): δ 2.02 (3H, s, NAc), 3.29 - 3.47 (3H, m, H-4, H-7, H-8), 3.64 - 3.76 (4H, m, H-9, H-9', PhCH₂), 3.82 (1H, app t, $J_{5,4}$ 10.2, $J_{5,6}$ 10.2 Hz, H-5), 3.87 (1H, app d, $J_{6,5}$ 10.2 Hz, H-6), 5.62 (1H, d, $J_{3,4}$ 2.4 Hz, H-3), 7.07 - 7.18 (5H, m, ArH). ^{13}C NMR (75.5 MHz, D_2O): δ 21.9 (NHCOCH₃), 32.9 (PhCH₂), 43.3 (C-4), 46.7 (C-5), 63.0 (C-9), 68.3 (C-7), 70.0 (C-6), 76.0 (C-8), 107.7 (C-3), 127.3, 128.7, 129.1 (Ar \times 5), 174.2 (CO₂Na). *C-2 and NHCOCH₃ were not observed. LRMS m/z 420 ($[M+H]^+$, 100%). HRMS calcd for C₁₈H₂₂NNaO₇S (M+H) 420.1087, found 420.1079. Partial 1H NMR spectrum of minor component: δ 1H NMR (300 MHz, D_2O): δ 3.45 (3H, m, H-4, H-7, H-8), 5.60 (1H, d, $J_{3,4}$ 6.0, H-3).

Sodium (5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*S*-(2-phenylbenzyl)-4-thio-D-glycero-D-galacto-non-2-en)onate, 28-B.

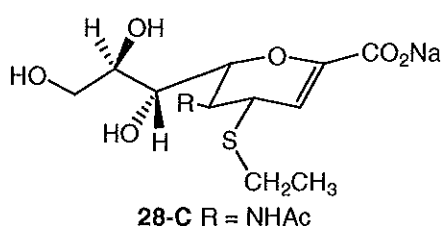


To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-2,3,5-trideoxy-4-*S*-(2-phenyl)benzyl-4-thio-D-glycero-D-galacto-non-2-enonate, **59-B**, (100 mg, 0.16 mmol) in MeOH (5 mL) was added 0.1 M KOH (5 mL). The reaction mixture was left to stir at room temperature for 2

hours. TLC (7:2:1 EtOAc/MeOH/ H_2O , R_f 0.34) showed complete conversion of starting material (R_f 0.75). The reaction mixture was acidified using Amberlite IR-120 (H^+) resin to pH 3 - 4, filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was resuspended in water and neutralised to pH 7.8 with aqueous NaOH and concentrated under reduced pressure to afford a dark yellow mass, 90 mg, > 100%. The crude product was purified by reverse-phase HPLC using isocratic 20% acetonitrile/80% H_2O , elution time 10.3 mins, and lyophilised to give **28-B** as a pale yellow fine powder (56 mg, 73%). The two epimers were inseparable by

HPLC conditions (approx ratio: 1 eq:0.15 *epi*). NMR data of major component: ^1H NMR (300 MHz, CD_3OD): δ 2.02 (3H, s, NAc), 3.51 - 3.70 (3H, m, H-4, H-7, H-9), 3.78 - 3.92 (4H, m, H-8, H-9', PhCH_2), 4.05 (1H, app t, $J_{5,4}$ 10.5, $J_{5,6}$ 10.5 Hz, H-5), 4.10 (1H, app d, $J_{6,5}$ 10.5 Hz, H-6), 5.78 (1H, d, $J_{3,4}$ 2.4 Hz, H-3), 7.23 - 7.50 (9H, m, ArH). ^{13}C NMR (75.5 MHz, CD_3OD): δ 22.7 (NHCOCH_3), 32.6 (PhCH_2), 44.3 (C-4), 45.7 (C-5), 65.0 (C-9), 70.4 (C-7), 71.2 (C-6), 78.0 (C-8), 107.0 (C-3), 128.2, 128.3, 128.6, 129.3, 129.4, 130.2, 130.4, 131.1, 131.9 (Ar \times 9), 174.1 (CO_2Na). *C-2 and NHCOCH_3 were not observed. LRMS m/z 472 ($[\text{M}-\text{Na}]^-$, 100%). HRMS calcd for $\text{C}_{23}\text{H}_{24}\text{KNO}_7\text{S}$ (M+H) 498.0989, found 498.1410. Partial ^1H NMR spectrum of minor component: δ ^1H NMR (300 MHz, CD_3OD): δ 3.61 (3H, m, H-4, H-7, H-9), 5.76 (1H, d, $J_{3,4}$ 6.0, H-3).

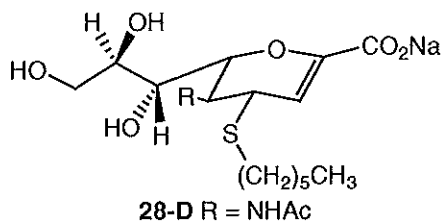
Sodium (5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-S-ethyl-4-thio-D-glycero-D-galacto-non-2-en)onate, 28-C.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-2,3,5-trideoxy-4-*S*-ethyl-4-thio-D-glycero-D-galacto-non-2-enonate, **59-C**, (100 mg, 0.21 mmol) in MeOH (5 mL) was added 0.1 M KOH (5 mL). The reaction mixture was left to stir at room temperature for 2 hours. TLC (7:2:1 EtOAc/MeOH/ H_2O , R_f 0.11) showed complete conversion of starting material (R_f 0.74). The reaction mixture was acidified using Amberlite IR-120 (H^+) resin to pH 3 - 4, filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was resuspended in water and neutralised to pH 7.8 with aqueous NaOH and concentrated under reduced pressure to afford a dark yellow mass, 82 mg, > 100%. The crude product was purified by reverse-phase HPLC using isocratic 3% acetonitrile/97% H_2O , elution time 5.1 mins, and lyophilised to give **28-C** as a pale yellow fine powder (65 mg, 92%). The two epimers were inseparable by HPLC conditions (approx ratio: 1 eq:0.16 *epi*). NMR data of major component: ^1H NMR (300 MHz, D_2O): δ 1.12 (3H, t, $J_{\text{CH}_3, \text{SCH}_2}$ 7.2, Hz, CH_3), 1.96 (3H, s, NAc), 2.55 (2H, t, $J_{\text{SCH}_2, \text{CH}_3}$ 7.2 Hz, SCH_2), 3.51 - 3.56 (3H, m, H-4, H-7, H-9), 3.72 (1H, dd, $J_{9',8}$ 2.7, $J_{9',9}$ 12 Hz, H-9'), 3.85 (1H, ddd, $J_{8,7}$ 3.6, $J_{8,9}$ 6.3, $J_{8,9'}$ 2.7 Hz, H-8), 4.08 (2H, m, H-5, H-6), 5.73 (1H, d, $J_{3,4}$ 2.7 Hz, H-3). ^{13}C NMR (75.5 MHz, D_2O): δ 13.9 (CH_3), 22.0 (NHCOCH_3), 22.9 (SCH_2), 43.2 (C-4), 47.3 (C-5), 63.1 (C-9), 68.4 (C-7), 69.8 (C-6),

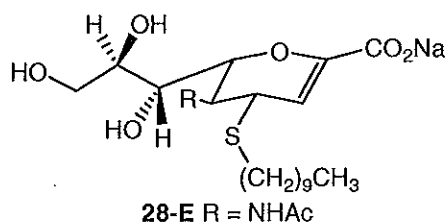
76.1 (C-8), 107.9 (C-3), 174.2 (CO₂Na). *C-2 and NHCOCH₃ were not observed. LRMS *m/z* 334 ([M-Na]⁺, 100%). HRMS calcd for C₁₃H₂₀NNaO₇S (M+H) 358.0936, found 358.0937. Partial ¹H NMR spectrum of minor component: δ ¹H NMR (300 MHz, D₂O): δ 3.64 (1H, t, *J*_{4,3} 6.0, *J*_{4,5} 5.1, H-4), 5.81 (1H, d, *J*_{3,4} 6.0, H-3).

Sodium (5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*S*-hexyl-4-thio-D-glycero-D-galacto-non-2-en)onate, 28-D.



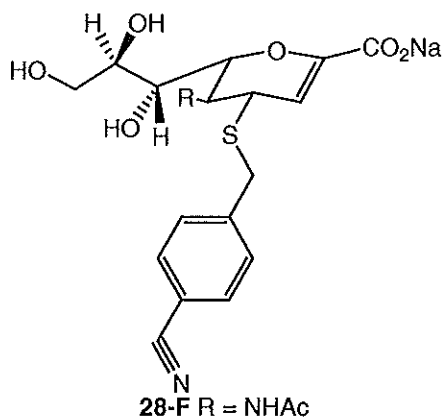
To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-2,3,5-trideoxy-4-*S*-hexyl-4-thio-D-glycero-D-galacto-non-2-enonate, **59-D**, (100 mg, 0.19 mmol) in MeOH (5 mL) was added 0.1 M KOH (5 mL). The reaction mixture was left to stir at room temperature for 2 hours. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.16) showed complete conversion of starting material (*R_f* 0.70). The reaction mixture was acidified using Amberlite IR-120 (H⁺) resin to pH 3 - 4, filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was resuspended in water and neutralised to pH 7.8 with aqueous NaOH and concentrated under reduced pressure to afford a dark yellow mass, 86 mg, > 100%. The crude product was purified by reverse-phase HPLC using isocratic 10% acetonitrile/90% H₂O, elution time 15.7 mins, and lyophilised to give **28-D** as a pale yellow fine powder (38 mg, 52%). The two epimers were inseparable by HPLC conditions (approx ratio: 1 eq:0.29 *epi*). NMR data of major component: ¹H NMR (300 MHz, CD₃OD): δ 0.93 (3H, t, *J* CH₃,CH₂ 6.6 Hz, CH₃), 1.33 (6H, s, CH₂), 1.56 (2H, t, *J* SCH₂CH₂,SCH₂ 7.2 Hz, SCH₂CH₂), 2.04 (3H, s, NAc), 2.59 (2H, t, *J* SCH₂,SCH₂CH₂ 7.2 Hz, SCH₂), 3.51, (1H, dd, *J*_{4,3} 2.1, *J*_{4,5} 9.3 Hz, H-4), 3.62 - 3.72 (2H, m, H-7 & 9), 3.81 - 3.92 (2H, m, H-8, H-9'), 4.05 (1H, app t, *J*_{5,4} 10.5, *J*_{5,6} 10.5 Hz, H-5), 4.11 (1H, app d, *J*_{6,5} 10.5 Hz, H-6), 5.80 (1H, d, *J*_{3,4} 2.4 Hz, H-3). ¹³C NMR (75.5 MHz, CD₃OD): δ 14.4 (CH₃), 22.7 (NHCOCH₃), 23.7 (CH₂), 29.7 (SCH₂), 29.8 (SCH₂CH₂), 30.8, 31.3 (2 × CH₂), 43.4 (C-4), 44.7 (C-5), 65.0 (C-9), 70.4 (C-7), 71.3 (C-6), 78.0 (C-8), 107.8 (C-3), 174.2 (CO₂Na). *C-2 and NHCOCH₃ were not observed. LRMS *m/z* 414 ([M+H]⁺, 100%). HRMS calcd for C₁₇H₂₈NNaO₇S (M+Na) 436.1382, found 436.1375. Partial ¹H NMR spectrum of minor component: δ ¹H NMR (300 MHz, CD₃OD): δ 3.40 (3H, m, H-4, H-7, H-9), 5.56 (1H, d, *J*_{3,4} 6.0, H-3).

Sodium (5-acetamido-2,6-anhydro-4-*S*-decyl-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-en)onate, 28-E.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-4-*S*-decyl-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **59-E**, (100 mg, 0.17 mmol) in MeOH (5 mL) was added 0.1 M KOH (5 mL). The reaction mixture was left to stir at room temperature for 2 hours. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.13) showed complete conversion of starting material (*R_f* 0.65). The reaction mixture was acidified using Amberlite IR-120 (H⁺) resin to pH 3 - 4, filtered and concentrated under reduced pressure to afford a dark yellow mass, 95 mg, > 100%. The crude product was resuspended in water and neutralised to pH 7.8 with aqueous NaOH and concentrated under reduced pressure to afford a dark yellow mass. The crude product was purified by reverse-phase HPLC using isocratic 25% acetonitrile/75% H₂O, elution time 13.6 mins, and lyophilised to give **28-E** as a pale yellow fine powder (56 mg, 73%). The two epimers were inseparable by HPLC conditions (approx ratio: 1 eq:0.25 *epi*). NMR data of major component: ¹H NMR (300 MHz, CD₃OD): δ 0.92 (3H, t, *J* CH₃,CH₂ 6.6 Hz, CH₃), 1.31 (6H, s, CH₂), 1.58 (2H, t, *J* SCH₂CH₂,SCH₂ 7.5 Hz, SCH₂CH₂), 2.04 (3H, s, NAc), 2.59 (2H, t, *J* SCH₂,SCH₂CH₂ 7.5 Hz, SCH₂), 3.50, (1H, dd, *J*_{4,3} 2.1, *J*_{4,5} 9.6 Hz, H-4), 3.61 - 3.71 (2H, m, H-7, H-9), 3.81 - 3.81 (2H, m, H-8, H-9'), 4.04 (1H, app t, *J*_{5,4} 10.5, *J*_{5,6} 10.5 Hz, H-5), 4.10 (1H, app d, *J*_{6,5} 10.5 Hz, H-6), 5.95 (1H, d, *J*_{3,4} 2.4 Hz, H-3). ¹³C NMR (75.5 MHz, CD₃OD): δ 14.5 (CH₃), 22.7 (NHCOCH₃), 23.8 (CH₂), 29.7, 30.2 (2 × CH₂), 30.4 (SCH₂), 30.5 (SCH₂CH₂), 30.7, 30.9 (5 × CH₂), 43.5 (C-4), 44.8 (C-5), 65.0 (C-9), 70.4 (C-7), 71.4 (C-6), 78.0 (C-8), 107.8 (C-3), 174.2 (CO₂Na). *C-2 and NHCOCH₃ were not observed. LRMS *m/z* 446 ([M-Na]⁺, 100%). HRMS calcd for C₂₁H₃₇KNO₇S (M+H) 486.1928, found 486.1925. Partial ¹H NMR spectrum of minor component: δ ¹H NMR (300 MHz, CD₃OD): δ 3.70 (3H, m, H-4, H-7, H-9), 5.86 (1H, d, *J*_{3,4} 6.0, H-3).

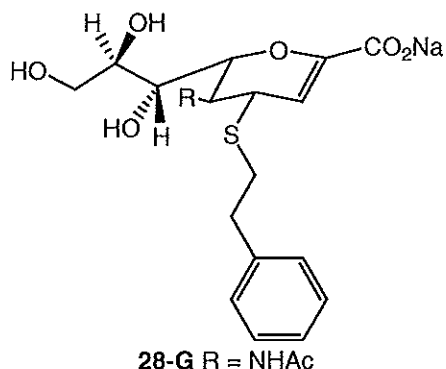
Sodium (5-acetamido-2,6-anhydro-4-*S*-(*p*-cyanobenzyl)-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-en)onate, 28-F.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-4-*S*-(*p*-cyanobenzyl)-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, **59-F**, (100 mg, 0.18 mmol) in MeOH (5 mL) was added 0.1 M KOH (5 mL). The reaction mixture was left to stir at room temperature for 2 hours. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.27) showed complete conversion of starting material (*R_f*

0.75). The reaction mixture was acidified using Amberlite IR-120 (H⁺) resin to pH 3 - 4, filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was resuspended in water and neutralised to pH 7.8 with aqueous NaOH and concentrated under reduced pressure to afford a dark yellow mass, 82 mg, > 100%. The crude product was purified by reverse-phase HPLC using isocratic 5% acetonitrile/95% H₂O, elution time 11.2 mins, and lyophilised to give **28-F** as a pale yellow fine powder (30 mg, 40%). The two epimers were inseparable by HPLC conditions (approx ratio: 1 eq:0.06 *epi*). NMR data of major component: ¹H NMR (300 MHz, D₂O): δ 1.86 (3H, s, NAc), 3.35 - 3.51 (3H, m, H-4, H-7, H-9), 3.72 (1H, dd, *J*_{9',8} 2.7, *J*_{9',9} 11.4 Hz, H-9'), 3.77 - 3.81 (3H, m, H-8, PhCH₂), 3.86 (2H, m, H-5, H-6), 5.62 (1H, d, *J*_{3,4} 2.4 Hz, H-3), 7.38 (2H, d, *J*_{ArH-o, ArH-m} 8.1 Hz, ArH-*o*), 7.60 (2H, d, *J*_{ArH-m, ArH-o} 8.1 Hz, ArH-*m*). ¹³C NMR (75.5 MHz, D₂O): δ 21.8 (NHCOCH₃), 32.6 (PhCH₂), 43.6 (C-4), 46.5 (C-5), 63.0 (C-9), 68.3 (C-7), 69.6 (C-8), 75.9 (C-6), 107.6 (C-3), 109.5 (PhCN), 129.8 (2 × Ph-*o*), 132.6 (2 × Ph-*m*), 144.7, 148.3 (2 × Ar), 174.2 (CO₂Na). *C-2 and NHCOCH₃ were not observed. LRMS *m/z* 421 ([M-Na]⁺, 100%). HRMS calcd for C₁₉H₂₁N₂NaO₇S (M+H) 445.1045, found 445.1047. Partial ¹H NMR spectrum of minor component: δ ¹H NMR (300 MHz, D₂O): δ 3.45 (3H, m, H-4, H-7, H-9), 5.66 (1H, d, *J*_{3,4} 6.0, H-3).

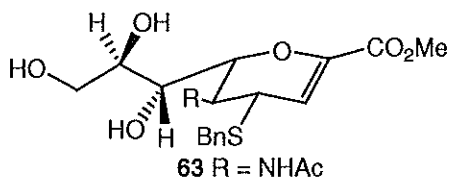
Sodium (5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*S*-(2-phenylethyl)-4-thio-D-glycero-D-galacto-non-2-en)onate, 28-G.



To a stirred solution of methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-2,3,5-trideoxy-4-*S*-(2-phenylethyl)-4-thio-D-glycero-D-galacto-non-2-enonate, **59-G**, (100 mg, 0.180 mmol) in MeOH (5 mL) was added 0.1 M KOH (5 mL). The reaction mixture was left to stir at room temperature for 2 hours. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.40)

showed complete conversion of starting material (*R_f* 0.67). The reaction mixture was acidified using Amberlite IR-120 (H⁺) resin to pH 3 - 4, filtered and concentrated under reduced pressure to afford a dark yellow mass. The crude product was resuspended in water and neutralised to pH 7.8 with aqueous NaOH and concentrated under reduced pressure to afford a dark yellow mass 118 mg, > 100%. The crude product was purified by reverse-phase HPLC using isocratic 10% acetonitrile/90% H₂O, elution time 8.8 mins, and lyophilised to give **28-G** as a pale yellow fine powder (35 mg, 44%). The two epimers were inseparable by HPLC conditions (approx ratio: 1 eq:0.29 *epi*). NMR data of major component: ¹H NMR (300 MHz, D₂O): δ 1.88 (3H, s, NAc), 2.78 (4H, s, SCH₂CH₂), 3.43 - 3.52 (3H, m, H-4, H-7, H-9), 3.73 (1H, dd, *J*_{9',8} 2.4, *J*_{9',9} 12.6 Hz, H-9'), 3.78 (1H, ddd, *J*_{8,7} 3.9, *J*_{8,9} 6.3, *J*_{8,9'} 2.4 Hz, H-8), 3.92 - 4.01 (2H, m, H-5, H-6), 5.64 (1H, d, *J*_{3,4} 2.4 Hz, H-3), 7.10 - 7.30 (5H, m, ArH). ¹³C NMR (75.5 MHz, D₂O): δ 21.9 (NHCOCH₃), 29.7 (SCH₂), 34.7 (SCH₂CH₂), 43.3 (C-4), 47.5 (C-5), 63.0 (C-9), 68.4 (C-7), 69.7 (C-8), 76.0 (C-6), 107.8 (C-3), 126.5, 128.6, 128.7 (5 × ArH). *CO₂Na, C-2, and NHCOCH₃ were not observed. LRMS *m/z* 434 ([M+H]⁺, 100%). HRMS calcd for C₂₃H₂₅NO₇S (M+Na) 482.1244, found 482.1251. Partial ¹H NMR spectrum of minor component: δ ¹H NMR (300 MHz, D₂O): δ 3.47 (3H, m, H-4, H-7, H-9), 5.67 (1H, d, *J*_{3,4} 6.0, H-3).

Methyl 5-acetamido-2,6-anhydro-4-*S*-benzyl-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, 63.

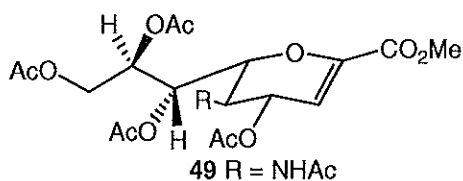


Methyl 5-acetamido-7,8,9-tri-*O*-acetyl-2,6-anhydro-4-*S*-benzyl-2,3,5-trideoxy-4-thio-D-glycero-D-galacto-non-2-enonate, 59-A (56 mg, 0.10 mmol) was stirred in a solution of 5 mL dry

MeOH and 66 μ L 1M NaOMe (66 μ mol) under N₂ at room temperature for 2 hours. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.55) showed complete consumption of starting material. The reaction mixture was neutralised using Amberlite IR-120 (H⁺) resin, filtered and concentrated under reduced pressure to afford the title compound **63** as a 1:0.64 epimeric mixture as a white foam without further purification, yield 38 mg, 88%. Partial NMR data of major component: ¹H NMR (300 MHz, CD₃OD): δ 2.06 (3H, s, NAc), 3.55 (1H, dd, *J*_{4,3} 2.4, *J*_{4,5} 9.6 Hz, H-4), 3.75 (3H, s, CO₂Me), 4.05 (1H, d, *J*_{gem} 10.8 Hz, PhCH₂), 4.29 (1H, d, *J*_{gem} 10.8 Hz, PhCH₂), 6.43 (1H, d, *J*_{3,4} 2.7 Hz, H-3), 7.18-7.33 (5H, m, ArH). Partial NMR data of minor component: ¹H NMR (300 MHz, CD₃OD): δ 2.03 (3H, s, NAc), 3.76 (3H, s, CO₂Me), 6.41 (1H, d, *J*_{3,4} 6.0 Hz, H-3).

5.2.2.2 4-*O*-alkylated Neu5Ac2en derivatives

Methyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-2,3,5-trideoxy-2,3-dehydro-D-glycero-D-galacto-non-2-ulosonate, 49.



To glycosyl chloride: A mixture of Neu5Ac, **1**, (5 g, 14.48 mmol) and Amberlite IR-120 (H⁺) resin (2.5 g) in dry MeOH (300 mL) was stirred at room temperature for 1 day. The reaction mixture was

filtered and concentrated under reduced pressure to yield a white solid (4.43g, 95%). The crude methyl 5-acetamido-3,5-dideoxy-D-glycero-D-galacto-non-2-ulosonate, **50** (2.5 g, approx 7.80 mmol) was stirred in acetyl chloride (20 mL) for 2 days at room temperature in a sealed glass-stoppered round bottom flask. Reaction completion was determined by TLC (EtOAc, *R_f* 0.7). The reaction mixture was concentrated under reduced pressure and azeotroped with toluene (3 \times 50 mL) to yield an off white foam.

Method 1: Elimination using pyridine.

The crude chloride **47** was then taken up in dry pyridine (10 mL) and stirred at 50°C for 1 hour under Ar. The reaction mixture was concentrated under reduced pressure and azeotroped with toluene (3 × 50 mL) to yield a brown oil, column chromatography (EtOAc) gave the title product **49** as a off-white foam (3.90 g, 55%). The ¹H NMR spectrum was in agreement with that reported in the literature.¹³⁶

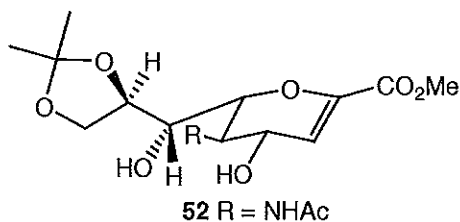
Method 2: Elimination using triethylamine.

The crude chloride **47** was then taken up in dioxane (30 mL, 0.36 mmol), to which triethylamine (3.0 mL, 21.46 mmol) was added dropwise, the reaction was then heated to 80°C and left to stir for half an hour under Ar. The reaction mixture was concentrated under reduced pressure and azeotroped with toluene (3 × 50 mL) to yield a brown oil, column chromatography (EtOAc) gave the title product **49** as a brown foam (5.25 g, 74%). The ¹H NMR spectrum was in agreement with that reported in the literature.¹³⁶

Method 3: Elimination using DBU.

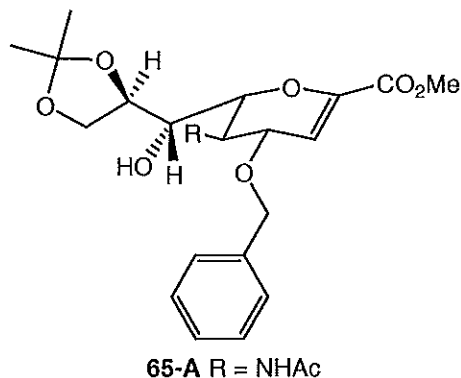
The crude chloride, **47** was then taken up in dry DCM (10 mL), to which DBU (176 μL, 1.17 mmol) was added. The reaction was left to stir for 2 and a half hours at room temperature under Ar. TLC (EtOAc, *R_f* 0.31) indicated reaction completion. Reaction mixture was evaporated to dryness, taken up in chloroform and washed with sat. aq. NH₄Cl, H₂O and dried with sat. aq. NaCl and Na₂SO₄, the organic layer was evaporated under reduced pressure, and chromatographed (7:1 EtOAc/hexanes) to give the title compound **49** as an off-white foam (5.26 g, 74%). The ¹H NMR spectrum was in agreement with that reported in the literature.¹³⁶ ¹H NMR (300 MHz, CDCl₃): δ 1.94 (3H, s, NAc), 2.05, 2.06, 2.08, 2.12 (12H, 4 × s, 4 × OAc), 3.80 (3H, s, CO₂Me), 4.19 (1H, dd, *J*_{9,8} 6.9, *J*_{9,9'} 12.3 Hz, H-9), 4.37-4.40 (2H, m, H-5, H-6), 4.57 (1H, dd, *J*_{9',8} 3.3, *J*_{9',9} 12.6 Hz, H-9'), 5.37 (1H, ddd, *J*_{8,7} 4.8, *J*_{8,9} 6.9, *J*_{8,9'} 2.7 Hz, H-8), 5.45-5.51 (2H, m, H-7, H-4), 6.02 (1H, d, *J*_{3,4} 3.3 Hz, H-3). LRMS *m/z* 496 ([M+Na]⁺, 100%).

Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate, 52.



To a stirring solution of methyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-2,3,5-trideoxy-2,3-dehydro-D-glycero-D-galacto-non-2-ulosonate, **49** (1 g, 2.31 mmol) in 10 mL dry MeOH under Ar at 0°C was added 10 mL 1M NaOMe. The reaction was left for 1 hour. After this time, TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.4) showed complete consumption of starting material (*R_f* 0.62). The reaction mixture was neutralised using Amberlite IR-120 (H⁺) resin, filtered and concentrated under reduced pressure to afford the title compound **64** as a white foam (594 mg, 92%). A suspension of crude methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-D-glycero-D-galacto-non-2-enonate, **64**, (594 mg, 1.95 mmol) was stirred in a mixture of 2,2-dimethoxypropane (1.16 mL, 9.47 mmol), dry acetone (25 mL) and dry Amberlite IR-120 (H⁺) resin (580 mg) at room temperature under Ar until solution was clear (o/n). TLC (8.5:1.5 EtOAc/MeOH, *R_f* 0.26), showed complete consumption of starting material (*R_f* 0.1). Filtration and evaporation of solvent under reduced pressure gave the title compound **52** as a white foam (730 mg, 100%). The ¹H NMR spectrum was in agreement with that reported in the literature.¹⁸⁸ ¹H NMR (300 MHz, CD₃OD): δ 1.35, 1.39 (6H, 2 × s, 2 × CCH₃), 2.07 (3H, s, NAc), 3.60 (1H, dd, *J*_{9,8} 2.1 *J*_{9,9'} 8.7 Hz, H-9), 3.80 (3H, s, CO₂Me), 3.97 (1H, app q, *J*_{5,4} 8.7, *J*_{5,6} 10.8 Hz, H-5), 4.04 (1H, dd, *J*_{9,8} 2.1, *J*_{9,9'} 8.7 Hz, H-9'), 4.09 (1H, dd, *J*_{6,5} 10.8, *J*_{6,7} 1.2 Hz, H-6), 4.17 (1H, dd, *J*_{7,6} 1.2, *J*_{7,8} 6.3 Hz, H-7), 4.34 (1H, ddd, *J*_{8,7} 6.3, *J*_{8,9} 2.1, *J*_{8,9'} 5.7 Hz, H-8), 4.46 (1H, dd, *J*_{4,3} 2.4, *J*_{4,5} 8.7 Hz, H-4), 5.97 (1H, d, *J*_{3,4} 2.4 Hz, H-3). ¹³C NMR (75.5 MHz, CD₃OD): δ 22.7 (NHCOCH₃), 25.6 (CH₃CCH₃'), 27.2 (CH₃CCH₃'), 52.0 (C-5), 52.8 (CO₂CH₃), 67.7 (C-4), 68.0 (C-9), 71.0 (C-7), 76.1 (C-8), 79.2 (C-6), 113.6 (C-3). *CO₂CH₃, C-2, NHCOCH₃ and CH₃CCH₃ were not observed. LRMS *m/z* 344 ([M-H]⁺, 100%).

Methyl 5-acetamido-2,6-anhydro-4-*O*-benzyl-2,3,5-trideoxy-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate, 65-A.



Method 1: Alkylation with NaH.

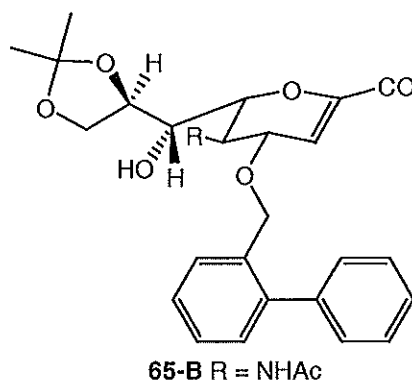
According to the method of Castro-Palomino *et al.*,¹¹⁸ a stirred solution of methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate, **52**, (300 mg, 0.87 mmol), dry DMF (3 mL) and benzyl bromide (135 μ L, 1.13 mmol) at 0°C under Ar, was added NaH (27 mg, 1.17 mmol) and the reaction was left to stir at 0°C for half an hour. TLC (EtOAc, R_f 0.60) showed complete consumption of starting material (R_f 0.46). Dry MeOH (0.1 mL) was added to quench reaction and the solvent was evaporated to dryness under reduced pressure. The resultant white solid was recrystallised from boiling EtOAc to give the title compound **65-A** as fine white crystals, yield 245 mg, 65%. The ^1H NMR spectrum was in agreement with that reported in the literature.¹¹⁸

Method 2: Alkylation with Ag₂O.

According to the method of Ikeda *et al.*,¹⁴⁴ **52** (100 mg, 0.29 mmol) was stirred in dry DCM (10 mL) with 3 Å molecular sieves and benzyl bromide (45 μ L, 0.65 mmol) at room temperature under Ar in a foil covered flask for 2 hrs. Ag₂O (200 mg, 0.86 mmol) and *n*-Bu₄NI (106 mg, 0.29 mmol) was added and the reaction was left to stir at room temperature for 2 days. TLC (EtOAc, R_f 0.51) showed complete consumption of starting material (R_f 0.46). The reaction mixture was filtered through Celite® and the solvent was evaporated to dryness under reduced pressure. The resultant dark brown amorphous mass was subject to column chromatography (5:1 EtOAc/hexanes) to give the title compound **65-A** as an off-white foam (79 mg, 62). The ^1H NMR spectrum was in agreement with that reported in the literature.¹¹⁸ ^1H NMR (300 MHz, CDCl₃): δ 1.27, 1.30 (6H, 2 \times s, 2 \times CCH₃), 1.92 (3H, s, NAc), 3.45 (1H, qd, $J_{7,6}$ 0.9, $J_{7,8}$ 8.1 $J_{7,\text{OH}}$ 3.9 Hz, H-7), 3.73 (3H, s, CO₂Me), 3.98 (2H, m, H-6, H-9), 4.11 (2H, m, H-4, H-9'), 4.18 (1H, m, H-5), 4.26 (1H, m, H-8), 4.45 (1H, d, J_{gem} 12.3 Hz, PhCH₂), 4.48 (1H, d, $J_{\text{OH},7}$ 3.9 Hz, 7-OH), 4.64 (1H, d, J_{gem} 12.3 Hz, PhCH₂), 5.23 (1H, br d, $J_{\text{NH},5}$ 7.2 Hz, NH), 6.02 (1H, d, $J_{3,4}$ 2.7 Hz, H-3), 7.30 (5H, m, Ph). ^{13}C NMR (75.5 MHz, CDCl₃): δ 23.2 (NHCOCH₃), 25.2 (CH₃CCH₃'), 27.0 (CH₃CCH₃'), 48.3 (C-4), 48.3 (COOCH₃), 67.3

(C-9), 69.9 (Ph-CH₂), 70.8 (C-7), 72.0 (C-5), 74.2 (C-8), 77.2 (C-6), 107.4 (C-3), 128.2 – 128.7 (Ph). * CO_2CH_3 , C-2, NHCOCH_3 and CH_3CCH_3 were not observed. LRMS m/z 434 ($[\text{M}-\text{H}]^-$, 100%). HRMS calcd for $\text{C}_{22}\text{H}_{29}\text{NO}_8$ ($\text{M}+\text{Na}$) 458.1791, found 458.1784.

Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-*O*-isopropylidene-4-*O*-(2-phenylbenzyl)-D-glycero-D-galacto-non-2-enonate, 65-B.



Method 1: Alkylation with NaH.

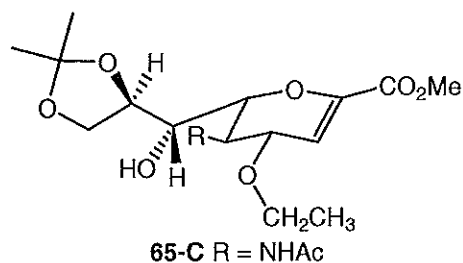
To a stirred solution of a stirred solution of methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate, **52**, (300 mg, 0.87 mmol), dry DMF (3 mL) and 2-phenylbenzyl bromide (202 μL , 1.11 mmol) at 0°C under Ar, was added NaH (27 mg, 1.13 mmol) and the reaction was left to stir at 0°C for 4 hours. TLC (EtOAc, R_f 0.75) showed complete consumption of starting material (R_f 0.46). Dry MeOH (0.1 mL) was added to quench reaction and the solvent was evaporated to dryness under reduced pressure. Flash chromatography (5:1 EtOAc/hexanes) was performed to give the title product **65-B** as an off-white amorphous solid, with a yield of 176 mg, 61%.

Method 2: Alkylation with Ag₂O.

52 (200 mg, 0.58 mmol) was stirred in dry DCM (20 mL) with 3 Å molecular sieves and 2-phenylbenzyl bromide (135 μL , 0.74 mmol) at room temperature under Ar in a foil covered flask for 2 hrs. Ag₂O (400 mg, 1.72 mmol) and *n*-Bu₄NI (212 mg, 0.58 mmol) was added and the reaction was left to stir at room temperature for 2 days. TLC (EtOAc, R_f 0.75) showed complete consumption of starting material (R_f 0.46). The reaction mixture was filtered through Celite® and the solvent was evaporated to dryness under reduced pressure. The resultant dark brown amorphous mass was subject to column chromatography (5:1 EtOAc/hexanes) to give the title compound **65-B** as an off-white foam (157 mg, 82%). ν_{max} 1266, 1422, 1651, 2306, 2988, 3055, 3447 cm^{-1} . ¹H NMR (300 MHz, CDCl₃): δ 1.33, 1.37 (6H, 2 \times s, 2 \times CCH₃), 1.96 (3H, s, NAc), 3.48 (1H, qd, $J_{7,6}$ 0.9, $J_{7,8}$ 8.1, $J_{7,\text{OH}}$ 4.5 Hz, H-7), 3.76 (3H, s, CO₂Me), 3.97-4.23 (5H,

m, H-4, H-5, H-6, H-9, H-9'), 4.32 (1H, m, H-8), 4.47 (1H, d, J_{gem} 11.4 Hz, PhCH₂), 4.56 (1H, d, J_{gem} 11.4 Hz, PhCH₂), 4.57 (1H, d, $J_{\text{OH},7}$ 3.9 Hz, 7-OH), 5.57 (1H, br d, $J_{\text{NH},5}$ 6.6 Hz, NH), 5.90 (1H, d, $J_{3,4}$ 2.4 Hz, H-3), 7.30-7.50 (9H, m, Ph). ¹³C NMR (75.5 MHz, CDCl₃): δ 23.2 (NHCOCH₃), 26.9 (CH₃CCH₃'), 27.0 (CH₃CCH₃'), 48.5 (C-4), 52.4 (COOCH₃), 67.2 (C-9), 68.0 (Ph-CH₂), 70.5 (C-7), 72.6 (C-5), 74.1 (C-8), 77.1 (C-6), 107.8 (C-3), 127.6-130.3 (Ph). LRMS m/z 510 ([M-H]⁻, 100%). * CO₂CH₃, C-2, NHCOCH₃ and CH₃CCH₃ were not observed. HRMS calcd for C₂₈H₃₃NO₈ (M+Na) 534.2104, found 534.2111.

Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-O-ethyl-8,9-O-isopropylidene-D-glycero-D-galacto-non-2-enonate, 65-C.



Method 1: Alkylation with NaH.

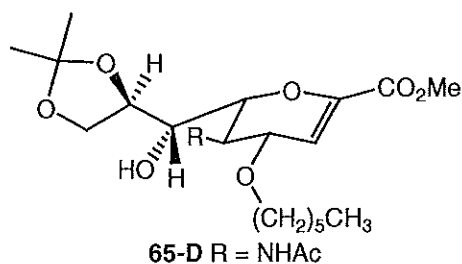
To a stirred solution of a stirred solution of methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-O-isopropylidene-D-glycero-D-galacto-non-2-enonate, **52**, (300 mg, 0.87 mmol), dry DMF (3 mL) and iodoethane (81 μ L, 1.01 mmol) at 0°C under Ar, was added NaH (27 mg, 1.13 mmol) and the reaction was left to stir at 0°C for 4 hours. TLC (EtOAc, R_f 0.51) showed complete consumption of starting material (R_f 0.46). Dry MeOH (0.1 mL) was added to quench reaction and the solvent was evaporated to dryness under reduced pressure. Flash chromatography (5:1 EtOAc/hexanes) was performed to give the title product **65-C** as an off-white amorphous solid, with a yield of 200 mg, 62%.

Method 2: Alkylation with Ag₂O.

52 (300 mg, 0.87 mmol) was stirred in dry DCM (20 mL) with 3 Å molecular sieves and iodoethane (81 μ L, 1.01 mmol) at room temperature under Ar in a foil covered flask for 2 hrs. Ag₂O (600 mg, 2.58 mmol) and *n*-Bu₄NI (319 mg, 0.87 mmol) was added and the reaction was left to stir at room temperature for 2 days. TLC (EtOAc, R_f 0.51) showed complete consumption of starting material (R_f 0.46). The reaction mixture was filtered through Celite[®] and the solvent was evaporated to dryness under reduced pressure. The resultant dark brown amorphous mass was subject to column chromatography (5:1 EtOAc/hexanes) to give the title compound **65-C** as an off-white foam (130 mg, 40%). ν_{max} 1655, 1729, 3020, 3455 cm⁻¹. ¹H NMR (300 MHz,

CDCl₃): δ 1.15 (3H, t, $J_{\text{CH}_3\text{OCH}_2}$ 6.9 Hz, CH₃), 1.29, 1.33 (6H, 2 \times s, 2 \times CCH₃), 1.99 (3H, s, NAc), 3.46-3.62 (3H, m, H-7 & OCH₂), 3.73 (3H, s, CO₂Me), 3.97-4.29 (6H, m, H-4, H-5, H-6, H-8, H-9, H-9'), 4.47 (1H, d, $J_{\text{OH},7}$ 3.6 Hz, 7-OH), 5.60 (1H, br d, $J_{\text{NH},5}$ 7.2 Hz, NH), 5.98 (1H, d, $J_{3,4}$ 3.3 Hz, H-3). ¹³C NMR (75.5 MHz, CDCl₃): δ 15.4 (CH₃), 23.2 (NHCOCH₃), 25.3 (CH₃CCH₃'), 27.3 (CH₃CCH₃'), 48.0 (C-4), 52.4 (COOCH₃), 67.2 (C-9), 67.6 (O-CH₂), 71.3 (C-7), 72.3 (C-5), 74.4 (C-8), 77.1 (C-6), 107.6 (C-3). LRMS m/z 372 ([M-H], 100%). * CO₂CH₃, C-2, NHCOCH₃ and CH₃CCH₃ were not observed. HRMS calcd for C₁₇H₂₇NO₈ (M+Na) 396.1634, found 396.1637. Found C, 53.53; H, 7.29; N, 3.55. C₁₇H₂₇NO₈·0.5 H₂O requires C, 53.39; H, 7.38; N, 3.66.

Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*O*-hexyl-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate, 65-D.



Method 1: Alkylation with NaH.

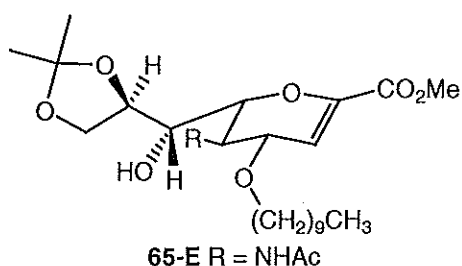
To a stirred solution of a stirred solution methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate, **52**, (300 mg, 0.87 mmol), dry DMF (3 mL) and bromohexane (141 μ L, 1.01 mmol) at 0°C under Ar, was added NaH (27 mg, 1.13 mmol) and the reaction was left to stir at 0°C for 4 hours. TLC (EtOAc, R_f 0.71) showed complete consumption of starting material (R_f 0.46). Dry MeOH (0.1 mL) was added to quench reaction and the solvent was evaporated to dryness under reduced pressure. The resultant white solid gave the title compound **65-D**, with a purified (5:1 EtOAc/hexanes flash chromatography) reaction yield of 226 mg, 61%.

Method 2: Alkylation with Ag₂O.

52 (300 mg, 0.87 mmol) was stirred in dry DCM (25 mL) with 3 Å molecular sieves and bromohexane (142 μ L, 1.01 mmol) at room temperature under Ar in a foil covered flask for 2 hrs. Ag₂O (600 mg, 2.58 mmol) and *n*-Bu₄NI (319 mg, 0.87 mmol) was added and the reaction was left to stir at room temperature for 2 days. TLC (2:1 EtOAc/hexanes, R_f 0.56) showed almost complete consumption of starting material (R_f 0.34). The reaction mixture was filtered through Celite® and the solvent was evaporated to dryness under reduced pressure. The resultant dark brown amorphous mass was

subject to column chromatography (5:1 EtOAc/hexanes) to give the title compound **65-D** as an off-white foam (126 mg, 34%). ν_{\max} 1265, 1373, 1422, 1438, 1654, 2305, 3056, 3420 cm^{-1} . ^1H NMR (300 MHz, CDCl_3): δ 0.78 (3H, t, $J_{\text{CH}_3, \text{OCH}_2}$ 6.9 Hz, CH_3), 1.18 (6H, s, CH_2), 1.26, 1.30 (6H, $2 \times$ s, $2 \times \text{CCH}_3$), 1.50 (2H, s, OCH_2CH_2), 1.96 (3H, s, NAc), 3.36-3.53 (3H, m, H-7 & OCH_2), 3.69, (3H, s, CO_2Me), 3.94-4.07 (4H, m, H-4, H-6, H-9, H-9'), 4.13 (1H, app t, $J_{5,4}$ 6.9, $J_{5,6}$ 6.9 Hz, H-5), 4.18-4.25 (1H, m, H-8), 4.45 (1H, d, $J_{\text{OH},7}$ 3.6 Hz, 7-OH), 5.48 (1H, br d, $J_{\text{NH},5}$ 7.2 Hz, NH), 5.94 (1H, d, $J_{3,4}$ 3.3 Hz, H-3). ^{13}C NMR (75.5 MHz, CDCl_3): δ 14.0 (CH_3), 22.6 (CH_2), 23.2 (NHCOCH_3), 25.3 (CH_3CCH_3 '), 25.7 (CH_2), 27.0 (CH_3CCH_3 '), 29.7 (CH_2), 31.5 ($\text{O-CH}_2\text{CH}_2$), 48.0 (C-4), 52.4 (COOCH_3), 67.2 (C-9), 68.2 (O-CH_2), 71.3 (C-7), 72.6 (C-5), 74.4 (C-8), 77.0 (C-6), 107.6 (C-3). $^*\text{CO}_2\text{CH}_3$, C-2, NHCOCH_3 and CH_3CCH_3 were not observed. LRMS m/z 429 ($[\text{M-H}]^-$, 100%). HRMS calcd for $\text{C}_{21}\text{H}_{35}\text{NO}_8$ ($\text{M}+\text{Na}$) 452.2260, found 452.2258.

Methyl 5-acetamido-2,6-anhydro-4-O-decyl-2,3,5-trideoxy-8,9-O-isopropylidene-D-glycero-D-galacto-non-2-enonate, 65-E.



Method 1: Alkylation with NaH.

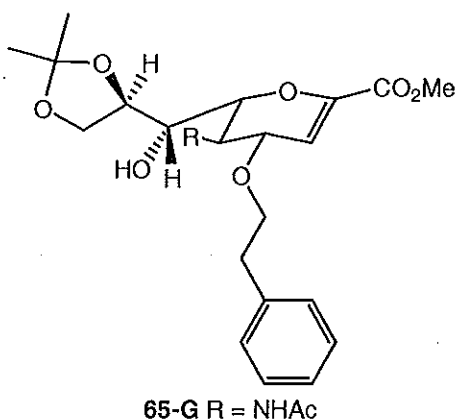
To a stirred solution of a stirred solution of methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-O-isopropylidene-D-glycero-D-galacto-non-2-enonate, **52**, (300 mg, 0.87 mmol), dry DMF (3 mL) and bromodecane (209 μL , 1.01 mmol) at 0°C under Ar, was added NaH (27 mg, 1.13 mmol) and the reaction was left to stir at 0°C for 4 hours. TLC (EtOAc, R_f 0.75) showed complete consumption of starting material (R_f 0.46). Dry MeOH (0.1 mL) was added to quench reaction and the solvent was evaporated to dryness under reduced pressure. The resultant white solid gave the title compound **65-E**, with a purified (5:1 EtOAc/hexanes flash chromatography) reaction yield of 140 mg, 33%.

Method 2: Alkylation with Ag_2O .

52 (300 mg, 0.87 mmol) was stirred in dry DCM (25 mL) with 3 Å molecular sieves and bromodecane (210 μL , 1.01 mmol) at room temperature under Ar in a foil covered flask for 2 hrs. Ag_2O (600 mg, 2.58 mmol) and $n\text{-Bu}_4\text{NI}$ (319 mg, 0.87 mmol) was added and the reaction was left to stir at room temperature for 2 days. TLC (8.5:1.5

EtOAc/MeOH, R_f 0.68) showed complete consumption of starting material (R_f 0.26). The reaction mixture was filtered through Celite[®] and the solvent was evaporated to dryness under reduced pressure. The resultant dark brown amorphous mass was subject to column chromatography (20:1 DCM/MeOH) to give the title compound **65-E** as an off-white foam (45 mg, 11%). ν_{\max} 1266, 1656, 1735, 2305, 3055, 3425 cm^{-1} . ^1H NMR (300 MHz, CDCl_3): δ 0.87 (3H, t, $J_{\text{CH}_3, \text{CH}_2}$ 6.6 Hz, CH_3), 1.26 (14H, s, $7 \times \text{CH}_2$), 1.36, 1.40 (6H, $2 \times$ s, $2 \times \text{CCH}_3$), 1.58, (2H, s, OCH_2CH_2), 2.05 (3H, s, NAc), 3.45-3.61 (3H, m, H-7, OCH_2), 3.80 (3H, s, CO_2Me), 4.04-4.36 (6H, m, H-4, H-5, H-6, H-8, H-9, H-9'), 4.53 (1H, d, $J_{\text{OH}, 7}$ 3.3 Hz, 7-OH), 5.55 (1H, br d, $J_{\text{NH}, 5}$ 7.5 Hz, NH), 6.04 (1H, d, $J_{3,4}$ 3.3 Hz, H-3). ^{13}C NMR (75.5 MHz, CDCl_3): δ 14.1 (CH_3), 22.7 (CH_2), 23.2 (NHCOCH_3), 25.3 (CH_3CCH_3 '), 26.0 (CH_2), 27.0 (CH_3CCH_3 '), 29.5 (CH_2), 31.9 ($\text{O-CH}_2\text{CH}_2$), 48.0 (C-4), 52.4 (COOCH_3), 67.3 (C-9), 68.2 (O-CH_2), 71.4 (C-7), 72.5 (C-5), 74.3 (C-6), 77.0 (C-8), 107.4 (C-3). $^*\text{CO}_2\text{CH}_3$, C-2, NHCOCH_3 and CH_3CCH_3 were not observed. LRMS m/z 484 ($[\text{M-H}]^-$, 100%). Found C, 61.71; H, 9.10; N, 2.69. $\text{C}_{25}\text{H}_{43}\text{NO}_8$ requires C, 61.83; H, 8.93; N, 2.88.

Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-O-ethylbenzyl-8,9-O-isopropylidene-D-glycero-D-galacto-non-2-enonate, 65-G.



Method 1: Alkylation with NaH.

To a stirred solution of methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-O-isopropylidene-D-glycero-D-galacto-non-2-enonate, **52**, (300 mg, 0.87 mmol) was stirred in a solution of dry DMF (3 mL) and 2-phenylethyl bromide (135 μL , 0.99 mmol) at 0°C under Ar. NaH (27 mg, 1.13 mmol) was added and the reaction was left to stir at 0°C

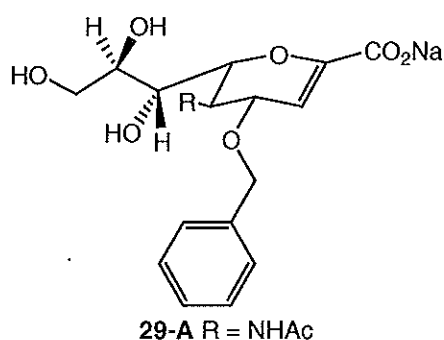
for 4 hours. TLC (EtOAc) showed only unreacted starting material (R_f 0.46). 0.1 mL of dry MeOH was added to quench reaction and the solvent was evaporated to dryness under reduced pressure. Analysis of the residue showed unreacted starting material.

Method 2: Alkylation with Ag_2O .

52 (300 mg, 0.87 mmol) was stirred in dry DCM (20 mL) with 3 Å molecular sieves and 2-phenylethyl bromide (135 μL , 0.99 mmol) at room temperature under Ar in

a foil covered flask for 2 hrs. Ag_2O (600 mg, 2.58 mmol) and $n\text{-Bu}_4\text{NI}$ (319 mg, 0.87 mmol) was added and the reaction was left to stir at room temperature for 2 days. TLC (EtOAc) showed unreacted starting material (R_f 0.46). Reaction was filtered through Celite[®] and the solvent was evaporated to dryness under reduced pressure. The resultant dark brown amorphous mass gave unreacted starting material.

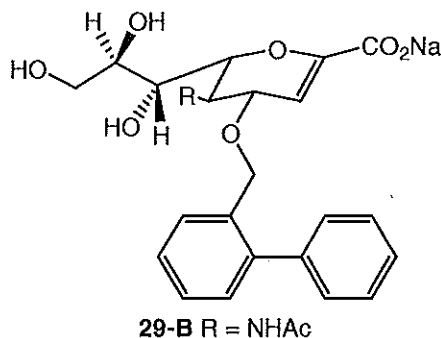
Sodium (5-acetamido-2,6-anhydro-4-*O*-benzyl-2,3,5-trideoxy-D-glycero-D-galacto-non-2-en)onate, 29-A.



Methyl 5-acetamido-2,6-anhydro-4-*O*-benzyl-2,3,5-trideoxy-8,9-*O*-isopropylidene-D-glycero-D-galacto-non-2-enonate, **65-A**, (200 mg, 0.46 mmol) was stirred in 80% AcOH at 85°C for half an hour. TLC (EtOAc, R_f 0 - 0.4) showed the formation of a broad lower R_f UV active spot. Evaporation of the solvent

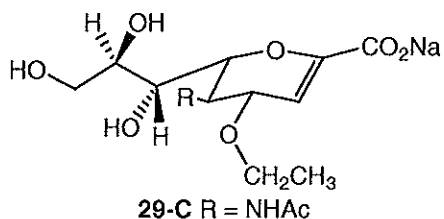
gave the crude product, methyl 5-acetamido-2,6-anhydro-4-*O*-benzyl-2,3,5-trideoxy-D-glycero-D-galacto-non-2-enonate, as an opaque foam (152 mg, 84%). The crude product (95 mg, approx 0.24 mmol) was stirred at 0°C in a solution of MeOH (4 mL) and 0.1 M KOH (4 mL) for 1 hr. TLC (7:2:1 EtOAc/MeOH/H₂O, R_f 0.25) showed the complete consumption of starting material (R_f 0.58). The reaction mixture was acidified to pH 3-4 with Amberlite IR-120 (H⁺) resin, filtered and the solvent was evaporated. The residue was taken up in H₂O and neutralised to pH 7.8 with aqueous NaOH and the solvent removed under vacuum, to give the title compound **29-A** as an off-white fine powder, 93 mg, > 100%. Reverse-phase HPLC (5% isocratic acetonitrile/95% H₂O, elution time: 8.6 min) yielded 40 mg, 44%. ¹H NMR (300 MHz, D₂O): δ 1.80 (3H, s, NAc), 3.42 (1H, d, $J_{7,8}$ 9.0 Hz, H-7), 3.79 (1H, dd, $J_{9,8}$ 6.3, $J_{9,9'}$ 11.7 Hz, H-9), 3.71 - 3.81 (2H, m, H-8, H-9'), 4.05 (2H, m, H-5, H-6), 4.26 (1H, dd, $J_{4,3}$ 2.4, $J_{4,5}$ 6.6 Hz, H-4), 4.43 (1H, d, J_{gem} 11.7 Hz, PhCH₂), 4.59 (1H, d, J_{gem} 11.7 Hz, PhCH₂), 5.73 (1H, d, $J_{3,4}$ 2.1 Hz, H-3), 7.30 (5H, m, Ph). ¹³C NMR (75.5 MHz, D₂O): δ 22.0 (NHCOCH₃), 47.3 (C-6), 63.0 (C-9), 68.1 (C-7), 69.7 (C-8), 70.2 (Ph-CH₂), 74.31 (C-4), 75.3 (C-5), 105.0 (C-3), 128.4 - 128.7 (Ph). *CO₂Na, C-2, and NHCOCH₃ were not observed. LRMS m/z 380 ([M-Na]⁺, 100%). HRMS calcd for C₁₈H₂₂NNaO₈ (M+H) 404.1321, found 404.1321. Found C, 49.06; H, 5.70; N, 3.08. C₁₈H₂₂NNaO₈.2H₂O requires C, 49.20; H, 5.96; N, 3.19.

Sodium (5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*O*-(2-phenylbenzyl)-D-glycero-D-galacto-non-2-en)onate, 29-B.



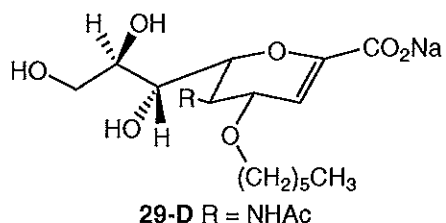
Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-8,9-*O*-isopropylidene-4-*O*-(2-phenylbenzyl)-D-glycero-D-galacto-non-2-enonate, **65-B**, (344mg, 0.67 mmol) was stirred in 80% AcOH at 85°C for half an hour. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.65) showed the formation of a broad lower *R_f* UV active spot. Evaporation of the solvent gave the crude product as an opaque foam (334 mg, > 100%). The crude product, methyl 5-acetamido-2,6-anhydro-4-*O*-(2-phenylbenzyl)-2,3,5-trideoxy-D-glycero-D-galacto-non-2-enonate, (334 mg, approx 0.73 mmol) was stirred at 0°C in a solution of MeOH (4 mL) and 0.1 M KOH (4 mL) for 1 hr. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.42) showed the formation of a broad lower *R_f* UV active spot. The reaction mixture was acidified to pH 3-4 with Amberlite IR-120 (H⁺) resin, filtered and the solvent was evaporated. The residue was taken up in H₂O and neutralised to pH 7.8 with aqueous NaOH and the solvent removed under vacuum, to give the title compound **29-B** as an off-white amorphous mass, 442 mg, > 100%. Reverse-phase HPLC (150 mg loaded onto the column) (20% isocratic acetonitrile/80% H₂O, elution time: 7.5 min) yielded 90 mg, 64%. ¹H NMR (300 MHz, D₂O): δ 1.75 (3H, s, NAc), 3.37 (1H, d, *J*_{7,8} 9.3 Hz, H-7), 3.46 (1H, dd, *J*_{9,8} 6.3, *J*_{9,9'} 12.0 Hz, H-9), 3.69 - 3.79 (2H, m, H-8, H-9'), 3.91-4.02 (2H, m, H-5, H-6), 4.14 (1H, dd, *J*_{4,3} 1.5, *J*_{4,5} 8.1 Hz, H-4), 4.38 (1H, d, *J*_{gem} 11.1 Hz, PhCH₂), 4.48 (1H, d, *J*_{gem} 11.1 Hz, PhCH₂), 5.49 (1H, d, *J*_{3,4} 2.1 Hz, H-3), 7.20-7.41 (9H, m, Ph). ¹³C NMR (75.5 MHz, D₂O): δ 22.0 (NHCOCH₃), 47.6 (C-6), 63.0 (C-9), 68.1 (C-7), 68.5 (Ph-CH₂), 69.6 (C-8), 74.4 (C-4), 75.3 (C-5), 104.8 (C-3), 127.5-130.1 (Ph). *C=O₂Na, C-2, and NHCOCH₃ were not observed. LRMS *m/z* 456 ([M-Na]⁺, 100%). Found C, 53.24; H, 5.59; N, 2.58. C₂₄H₂₆KNO₈•2.5 H₂O requires C, 53.32; H, 5.78; N, 2.59.

Sodium (5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*O*-ethyl-D-glycero-D-galactonon-2-en)onate, 29-C.



Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*O*-ethyl-8,9-*O*-isopropylidene-D-glycero-D-galactonon-2-enonate, **65-C**, (397 mg, 1.06 mmol) was stirred in 80% AcOH at 85°C for half an hour. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0.4) showed the formation of a broad lower *R_f* UV active spot. Evaporation of the solvent gave the title compound as an off-white solid mass, 307 mg, 87%. 145 mg (approx 0.44 mmol) of the crude methyl 5-acetamido-2,6-anhydro-4-*O*-ethyl-2,3,5-trideoxy-D-glycero-D-galactonon-2-enonate was stirred at 0°C in a solution of MeOH (4 mL) and 0.1 M KOH (4 mL) for 1 hr. TLC (7:2:1 EtOAc/MeOH/H₂O, *R_f* 0 - 0.2) showed the formation of a broad lower *R_f* UV active spot. The reaction mixture was acidified to pH 3-4 with Amberlite IR-120 (H⁺) resin, filtered and the solvent was evaporated. The residue was taken up in H₂O and neutralised to pH 7.8 with aqueous NaOH and the solvent removed under vacuum, to give the title compound **29-C** as an off-white amorphous mass, 234 mg, > 100%. Reverse-phase HPLC (1% isocratic acetonitrile/99% H₂O, elution time: 4.8 min) yielded 72 mg, 52%. ¹H NMR (300 MHz, D₂O): δ 1.03 (3H, t, *J*_{CH₃,OCH₂} 6.9 Hz, CH₃), 1.91 (3H, s, NAc), 3.44-3.62 (4H, m, H-7, H-9, O-CH₂), 3.73 (1H, dd, *J*_{9,8} 2.7, *J*_{9,9'} 11.4 Hz, H-9'), 3.80 (1H, ddd, *J*_{8,7} 3.9, *J*_{8,9} 6.3, *J*_{8,9'} 2.4 Hz, H-8), 4.03 (1H, dd, *J*_{5,4} 8.1, *J*_{5,6} 10.5 Hz, H-5), 4.10 (1H, dd, *J*_{6,5} 10.5, *J*_{6,7} 1.2 Hz, H-6), 4.22 (1H, dd, *J*_{4,3} 2.1, *J*_{4,5} 9.3 Hz, H-4), 5.64 (1H, d, *J*_{3,4} 2.4 Hz, H-3). ¹³C NMR (75.5 MHz, D₂O): δ 14.4 (CH₃), 21.9 (NHCOCH₃), 47.1 (C-6), 63.0 (C-9), 64.1 (O-CH₂), 68.1 (C-7), 69.6 (C-8), 74.7 (C-4), 75.3 (C-5), 105.3 (C-3), 169.4 (NHCOCH₃), 174.3 (CO₂Na). *C-2 was not observed. LRMS *m/z* 318 ([M-Na]⁺, 100%).

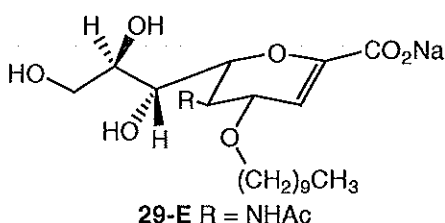
Sodium (5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*O*-hexyl-D-glycero-D-galactonon-2-en)onate, 29-D.



Methyl 5-acetamido-2,6-anhydro-2,3,5-trideoxy-4-*O*-hexyl-8,9-*O*-isopropylidene-D-glycero-D-galactonon-2-enonate, **65-D**, (72 mg, 0.17 mmol) was stirred in 80% AcOH at 85°C for half an hour. TLC

(7:2:1 EtOAc/MeOH/H₂O, R_f 0.42) showed the formation of a broad lower R_f UV active spot. Evaporation of the solvent gave an off-white solid mass, 93 mg, > 100%. The crude product methyl 5-acetamido-2,6-anhydro-4-*O*-hexyl-2,3,5-trideoxy-D-*glycero*-D-*galacto*-non-2-enonate (74 mg, approx 0.19 mmol) was stirred at 0°C in a solution of 1:1 MeOH (4 mL) /0.1 M KOH (4 mL) for 1 hr. TLC (7:2:1 EtOAc/MeOH/H₂O, R_f 0.02) showed the formation of a broad lower R_f UV active spot. The reaction mixture was acidified to pH 3-4 with Amberlite IR-120 (H⁺) resin, filtered and the solvent was evaporated. The residue was taken up in H₂O and neutralised to pH 7.8 with aqueous NaOH and the solvent removed under vacuum, to give the title compound **29-D** as a yellow amorphous mass, 140 mg, > 100%. Reverse-phase HPLC (15% isocratic acetonitrile/85% H₂O, elution time: 10.9 min) yielded 25 mg, 35%. ¹H NMR (300 MHz, CD₃OD): δ 0.63 (3H, t, $J_{CH_3CH_2}$ 6.6 Hz, CH₃), 1.03 (6H, s, 3 × CH₂), 1.26 (2H, t, $J_{OCH_2CH_2OCH_2}$ 7.2 Hz, OCH₂CH₂), 1.74 (3H, s, NAc), 2.29 (2H, t, $J_{OCH_2OCH_2CH_2}$ 7.2 Hz, OCH₂), 3.32-3.42 (2H, m, H-7, H-9), 3.51 – 3.62 (2H, m, H-8, H-9'), 3.75 (1H, dd, $J_{5,4}$ 9.3, $J_{5,6}$ 10.5 Hz, H-5), 3.81 (1H, app d, $J_{6,5}$ 10.5 Hz, H-6), 4.23 (1H, dd, $J_{4,3}$ 2.4, $J_{4,5}$ 9.3 Hz, H-4), 5.50 (1H, d, $J_{3,4}$ 2.4 Hz, H-3). ¹³C NMR (75.5 MHz, CD₃OD): δ 14.4 (CH₃), 22.7 (NHCOCH₃), 23.7 (CH₂), 29.7 (O-CH₂), 29.8 (OCH₂CH₂), 30.8, 31.25 (2 × CH₂), 43.4 (C-4), 44.7 (C-5), 65.0 (C-9), 70.4 (C-7), 71.3 (C-6), 78.0 (C-8), 107.8 (C-3), 174.2 (CO₂Na). *C-2 and NHCOCH₃ were not observed. LRMS m/z 374 ([M-Na]⁺, 100%).

Sodium (5-acetamido-2,6-anhydro-4-*O*-decyl-2,3,5-trideoxy-D-*glycero*-D-*galacto*-non-2-en)onate, 29-E.



Methyl 5-acetamido-2,6-anhydro-4-*O*-decyl-2,3,5-trideoxy-8,9-*O*-isopropylidene-D-*glycero*-D-*galacto*-non-2-enonate, **65-E**, (155 mg, 0.32 mmol) was stirred in 80% AcOH at 85°C for half an hour.

TLC (7:2:1 EtOAc/MeOH/H₂O, R_f 0.4) showed the formation of a broad lower R_f UV active spot. Evaporation of the solvent gave the crude product as an opaque foam (120 mg, 84%). The crude product methyl 5-acetamido-2,6-anhydro-4-*O*-decyl-2,3,5-trideoxy-D-*glycero*-D-*galacto*-non-2-enonate (120 mg, 0.27 mmol) was stirred at 0°C in a solution of 1:1 MeOH (4 mL) /0.1 M KOH (4 mL) for 1 hr. TLC (7:2:1 EtOAc/MeOH/H₂O, R_f 0) showed the formation of a lower R_f UV active spot. The reaction mixture was acidified to pH 3-4 with Amberlite IR-120 (H⁺) resin, filtered and

the solvent was evaporated. The residue was taken up in H₂O and neutralised to pH 7.8 with aqueous NaOH and the solvent removed under vacuum, to give the title compound **29-E** as an off-white fine powder, 160 mg, > 100%. Reverse-phase HPLC (25% isocratic acetonitrile/75% H₂O, elution time: 13.1 min, 40 mg loaded onto the column) yielded 21 mg, 56%. ¹H NMR (300 MHz, CD₃OD): δ 0.81 (3H, t, *J*_{CH₃,CH₂} 6.6 Hz, CH₃), 1.19 (14H, s, 7 × CH₂), 1.44 (2H, m, OCH₂CH₂), 1.90 (3H, s, NAc), 3.32-3.44 (3H, m, H-7, O-CH₂), 3.51-3.57 (2H, m, H-9 & O-CH₂'), 3.69 - 3.79 (2H, m, H-8, H-9'), 4.02-4.13 (3H, m, H-4, H-5, H-6), 5.68 (1H, d, *J*_{3,4} 2.1 Hz, H-3). ¹³C NMR (75.5 MHz, CD₃OD): δ 14.4 (CH₃), 22.8 (NHCOCH₃), 23.8, 27.3, 30.5, 30.6, 30.7, 30.8 (6 × CH₂), 31.9 (O-CH₂CH₂), 33.1 (CH₂), 49.0 (C-6), 64.9 (C-9), 69.3 (O-CH₂), 70.2 (C-7), 71.3 (C-8), 76.3 (C-4), 77.2 (C-5), 105.1 (C-3). *CO₂Na, C-2, and NHCOCH₃ were not observed. LRMS *m/z* 430 ([M-Na]⁺, 100%). HRMS calculated for C₂₁H₃₆NNaO₈ [M+H] 454.2417, found 454.2418.

References

1. C. K. Mathews and K. E. Van Holde, *Biochemistry*. Second ed.; The Benjamin/Cummings Publishing Company, Inc: Sydney, 1996.
2. M. von Itzstein and B. Smalec, *Today's Life Science* **1994**, 6, (4), 22-24.
3. N. Sharon and H. Lis, Carbohydrates in cell recognition. In *Scientific American*, 1993; January, pp 74-81.
4. M. J. Kiefel and M. von Itzstein, *Chemical Review* **2002**, 102, 471-490.
5. T. Angata and A. Varki, *Chemical Review* **2002**, 102, 439-469.
6. M. von Itzstein and R. J. Thomson, *Current Medicinal Chemistry* **1997**, 4, 185-210.
7. R. Schauer and J. P. Kamerling, Chemistry, biochemistry and biology of sialic acids. In *Glycoproteins II*, ed.; J. Montreuil, J. F. G. Vliegenthart and H. Schaechter, Ed. Elsevier science B.V.: New york, 1997; Vol. pp 243-402.
8. M. von Itzstein and M. J. Kiefel, Sialic acid analogues as potential antimicrobial agents. In *Carbohydrates in drug design*, Z. J. Wiczak and K. A. Nieforth, Ed. Marcel dekker, Inc: New York, 1997; Vol. pp 39-82.
9. M. von Itzstein and R. J. Thomson, *Topics in Current Chemistry* **1997**, 186, 119-170.
10. R. Schauer, S. Kelm, G. Reuter, P. Roggentin and L. Shaw, Biochemistry and role of sialic acids. In *Biology of the Sialic Acids*, A. Rosenberg, Ed. Plenum Press: London, 1995; pp 7-67.
11. H.-D. Klenk and R. Rott, *Advances in Virus Research* **1988**, 34, 247-281.
12. M. D. Dowle and P. D. Howes, *Expert Opinion on Therapeutic Patents* **1998**, 8, (11), 1461-1478.
13. M. J. Kiefel and M. von Itzstein, *Progress in Medicinal Chemistry* **1999**, 36, 1-28.
14. M. von Itzstein, W.-Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. V. Phan, M. L. Smythe, H. F. White, S. W. Oliver, P. M. Colman, J. N. Varghese, D. M. Ryan, J. M. Woods, R. C. Bethell, V. J. Hotham, J. M. Cameron and C. R. Penn, *Nature* **1993**, 363, (6428), 418-423.
15. D. Nadano, M. Iwasaki, S. Endo, K. Kitajima, S. Inoue and Y. Inoue, *The Journal of Biological Chemistry* **1986**, 261, (25), 11550-11557.
16. G. L. Taylor, S. J. Crennell, C. Thompson and M. Chuenkova, Sialidases. In *Carbohydrates in chemistry and biology*, B. Ernst, G. W. Hart and P. Sinay, Ed. Wiley-VCH: Weinhiem, 2000; Vol. 3, pp 485-494.
17. E. Monti, A. Preti, E. Rossi, A. Ballabio and G. Borsani, *Genomics* **1999**, 57, (1), 137-143.
18. L. M. G. Chavas, C. Tringali, P. Fusi, B. Venerando, G. Tettamanti, R. Kato, E. Monti and S. Wakatsuki, *The Journal of Biological Chemistry* **2005**, 280, (1), 469-475.
19. T. Corfield, *Glycobiology* **1992**, 2, (6), 209-521.
20. F. Villalta, C. M. Smith, A. Ruiz-Ruano and M. F. Lima, *Febs Letters* **2001**, 505, 383-388.
21. R. Drzeniek, *Current Topics in Microbiology and Immunology* **1972**, 59, 35-74.
22. G. Herrler, J. Hausmann and H.-D. Klenk, Sialic acid as receptor determinant of ortho- and paramyxoviruses. In *Biology of the Sialic Acids*, A. Rosenberg, Ed. Plenum Press: London, 1995; pp 315-336.
23. D. S. Leland, *Clinical Virology*. ed.; W.B. Saunders Company: Sydney, 1996.

24. W. P. Glezen, F. A. Loda and F. W. Denny, Parainfluenza viruses. In *Viral infections of humans: Epidemiology and control*, Third ed.; A. S. Evans, Ed. Plenum Medical Book Company: New York, 1991; pp 493-507.
25. B. S. Seal, D. J. King and H. S. Sellers, *Developmental and Comparative Immunology* **2000**, 24, 257-268.
26. R. Vainionpaa and T. Hyypia, *Clinical Microbiology Reviews* **1994**, 7, (2), 265-275.
27. Unknown, Human Parainfluenza Virus - 3. In [texte_alt_jlevir00016_gr2.jpg](http://www.john-libbey-eurotext.fr/fr/revues/bio_rech/vir/c-docs/00/04/09/48/texte_alt_jlevir00016_gr2.jpg), http://www.john-libbey-eurotext.fr/fr/revues/bio_rech/vir/c-docs/00/04/09/48/texte_alt_jlevir00016_gr2.jpg; 2005; 24.4 kB,.
28. Unknown, Newcastle Disease Virus. In [newcastle.jpg](http://aragriculture.org/News/avain/2001/summer2001/newcastle.jpg), <http://aragriculture.org/News/avain/2001/summer2001/newcastle.jpg>; 2001; 1.44 kB.
29. R. M. Chanock, *Journal of Experimental Medicine* **1956**, 104, 555-576.
30. R. M. Chanock, R. H. Parrott, K. Cook, B. E. Andrews, J. A. Bell, T. Riechelderfer, A. Z. Kapikian, F. M. Mastrota and R. J. Huebner, *The New England Journal of Medicine* **1958**, 258, 207-213.
31. K. M. Johnson, R. M. Chanock, M. K. Cook and R. J. Huebner, *American Journal of Hygiene* **1960**, 71, 81-92.
32. K. J. Henrickson, *Clinical Microbiology Reviews* **2003**, April 2003, 242-264.
33. T. M. Doyle, *Journal of Comparative Pathology and Therapeutics* **1927**, 40, 144-169.
34. E. F. Kaleta and Baldauf, Newcastle disease in free-living and pet birds. In *Newcastle Disease*, D. J. Alexander, Ed. Kluwer: Boston, 1988; pp 197-256.
35. D. J. Alexander, *Revue Scientifique et Technique de L Office International des Epizooties* **2001**, 19, (2), 443-432.
36. G. E. Gallili and D. BenNathan, *Biotechnology Advances* **1998**, 16, (2), 343-366.
37. C. Brown, D. J. King and B. S. Seal, *Veterinary Pathology* **1999**, 36, (2), 125-132.
38. A. Clavijo, Y. Robinson and J. Lopez, *Avian Diseases* **2001**, 45, (1), 245-250.
39. K. M. Lam, *Journal of Comparative Pathology* **1996**, 115, (3), 253-263.
40. K. Huberman, R. W. Peluso and A. Moscona, *Virology* **1995**, 214, 294-300.
41. G. L. Taylor, *Current Opinion in Structural Biology* **1996**, 6, 830-837.
42. L. McGinnes, T. Sergel and T. Morrison, *Virology* **1993**, 196, (1), 101-110.
43. R. T. Deng, Z. Y. Wang, P. J. Mahon, M. Marinello, A. Mirza and R. M. Iorio, *Virology* **1999**, 253, (1), 43-54.
44. R. T. Deng, Z. Y. Wang, R. L. Glickman and R. M. Iorio, *Virology* **1994**, 204, (1), 17-26.
45. H. Connaris, T. Takimoto, R. Russell, S. Crennell, I. Moustafa, A. Portner and G. Taylor, *Journal of Virology* **2002**, 76, (4), 1816-1824.
46. V. Zaitsev, M. von Itzstein, D. Groves, M. J. Kiefel, T. Takimoto, A. Portner and G. Taylor, *Journal of Virology* **2004**, 78, (7), 3733-3741.
47. P. M. Colman, P. A. Hoyne and M. C. Lawrence, *Journal of Virology* **1993**, 67, (6), 2972-2980.
48. S. J. Crennell, T. Takimoto, A. Portner and G. L. Taylor, *Nature Structural Biology* **2000**, 7, (11), 1068-1074.
49. P. J. Kraulis, *Journal of Applied Crystallography* **1991**, 24, 946-950.
50. E. A. Merritt and M. E. P. Murphy, *Acta Crystallographica* **1994**, D50, 869-873.
51. J. N. Varghese, W. G. Laver and P. M. Colman, *Nature* **1983**, 303, 35-40.
52. S. J. Crennell, E. F. Garman, W. G. Laver, E. R. Vimr and G. L. Taylor, *Proceedings of the National Academy of Science USA* **1993**, 90, 9852-9856.

53. S. J. Crennell, E. Garman, W. G. Laver, E. Vimr and G. L. Taylor, *Structure* **1994**, 15, 535-544.
54. J. P. M. Langedijk, F. J. Daus and J. T. VanOirschot, *Journal of Virology* **1997**, 71, (8), 6155-6167.
55. D. J. Alexander, P. Reeve and W. H. Allan, *Microbios* **1970**, 2, (6), 155-165.
56. A. P. Corfield, R. W. Veh, M. Wember, J.-C. Michalski and R. Schauer, *Biochemical Journal* **1981**, 197, 293-299.
57. M. C. Lawrence, N. A. Borg, V. A. Streltsov, P. A. Pilling, V. Chandana Epa, J. N. Varghese, J. L. McKimm-Breschkin and P. M. Colman, *Journal of Molecular Biology* **2004**, 335, 1343-1357.
58. T. Suzuki, A. Portner, R. A. Scroggs, M. Uchikawa, N. Koyama, K. Matsuo, Y. Suzuki and T. Takimoto, *Journal of Virology* **2001**, 75, (10), 4604-4613.
59. Z. M. Wang, L. L. Tong, D. Grant and T. Cihlar, *Journal of Virological Methods* **2001**, 98, 53-61.
60. J. Treanor and A. Falsey, *Antiviral Research* **1999**, 44, 79-102.
61. T. Tao, M. H. Skiadopoulos, A. P. Durbin, F. Davoodi, P. L. Collins and B. R. Murphy, *Vaccine* **1999**, 17, 1100-1108.
62. T. Tao, F. Davoodi, C. J. Cho, M. H. Skiadopoulos, A. P. Durbin, P. L. Collins and B. R. Murphy, *Vaccine* **2000**, 18, 1359-1366.
63. WHO., State of the art new vaccines research & development. In *Initiative for vaccine research*, WHO: Geneva, 2003; pp 1-75.
64. C. H. Kim, W. Lew, M. A. Williams, H. Liu, L. Zhang, S. Swaminathan, N. Bischofberger, M. S. Chen, D. B. Mendel, C. Y. Tai, W. G. Laver and R. C. Stevens, *Journal of the American Chemical Society* **1997**, 119, 681-690.
65. M. J. Bamford, *Journal of Enzyme Inhibition* **1995**, 10, 1-16.
66. A. K. J. Chong, M. S. Pegg, N. R. Taylor and M. von Itzstein, *European Journal of Biochemistry* **1992**, 207, 335-343.
67. A. G. Watts, P. Oppezso, S. G. Withers, P. M. Alzari and A. Buschiazso, *The Journal of Biological Chemistry* **2006**, 281, (7), 4149-4155.
68. P. Meindl, G. Bodo, P. Palese, J. L. Schulman and H. Tuppy, *Virology* **1974**, 58, 457-463.
69. C. T. Holzer, M. von Itzstein, B. Jin, M. S. Pegg, W. P. Stewart and W.-Y. Wu, *Glycoconjugate Journal* **1993**, 10, 40-44.
70. L. S. Gonzalez III, *Formulary* **2000**, 35, 812-831.
71. M. von Itzstein, J. C. Dyason, S. W. Oliver, H. F. White, W.-Y. Wu, G. B. Kok and M. S. Pegg, *Journal of Medicinal Chemistry* **1996**, 39, 388-391.
72. J. N. Varghese, V. C. Epa and P. M. Colman, *Protein Science* **1995**, 4, 1081-1087.
73. N. J. Montalto, K. D. Gum and J. V. Ashley, *American Family Physician* **2000**, 62, (11), 2467-2476.
74. L. V. Gubareva, L. Kaiser and F. G. Hayden, *The Lancet* **2000**, 355, (9206), 827-835.
75. P. Wutzler and G. Vogel, *Infection* **2000**, 28, (5), 261-266.
76. J. C. Dyason and M. von Itzstein, *Australian Journal Of Chemistry* **2001**, 54, 663-670.
77. W. Lew, X. Chen and C. H. Kim, *Current Medicinal Chemistry* **2000**, 7, (6), 663-672.
78. V. R. Atigadda, W. J. Brouillette, F. Duarte, Y. S. Babu, S. Bantia, P. Chand, N. Chu, J. A. Montgomery, D. A. Walsh, E. Sudbeck, J. Finley, G. M. Air, M. Luo and W. G. Laver, *Bioorganic & Medicinal Chemistry* **1999**, 7, 2487-2497.
79. Y. S. Babu, P. Chand, S. Bantia, P. Kotian, A. Dehghani, Y. El-Kattan, T.-H. Lin, T. L. Hutchison, A. J. Elliott, C. D. Parker, S. L. Ananth, L. L. Horn, W. G.

- Laver and J. A. Montgomery, *Journal of Medicinal Chemistry* **2000**, 43, 3482-3486.
80. Biocryst Pharmaceuticals Inc. announces preliminary Phase III trial results for Influenza neuraminidase inhibitor, Peramivir.
<http://www.shareholder.com/biocryst/news/> (26/04/2005).
 81. W. M. Kati, D. Montgomery, R. Carrick, L. Gubareva, C. Maring, K. McDaniel, K. Steffy, A. Molla, F. Hayden, D. Kempf and W. Kohlbrenner, *Antimicrobial Agents and Chemotherapy* **2002**, 46, (4), 1014-1021.
 82. O. Greengard, N. Poltoratskaia, E. Leikina, J. Zimmerberg and A. Moscona, *Journal of Virology* **2000**, 74, (23), 11108-11114.
 83. T. Suzuki, K. Ikeda, N. Koyama, C. Hosokawa, T. Kogure, T. Takahashi, K. Hidari, D. Miyamoto, K. Tanaka and Y. Suzuki, *Glycoconjugate Journal* **2001**, 18, (4), 331-337.
 84. I. V. Alymova, G. Taylor, T. Takimoto, T.-H. Lin, P. Chand, Y. S. Babu, C. Li, X. Xiong and A. Portner, *Antimicrobial Agents and Chemotherapy* **2004**, 48, (5), 1495-1502.
 85. P. Grootenhuis, A. van Boeckel and C. Haasnoot, *Tibtech* **1994**, 12, 9-14.
 86. R. Wang, L. Liu, L. Lai and Y. Tang, *Journal of Molecular Modelling* **1998**, 4, 379-394.
 87. M. J. Ellis and H. Hebert, *Micron* **2001**, 32, (5), 541-550.
 88. J. M. Word, S. C. Lovell, J. S. Richardson and D. C. Richardson, *Journal of Molecular Biology* **1999**, 285, (4), 1735-1747.
 89. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *Journal of Applied Crystallography* **1993**, 26, 283-291.
 90. R. W. W. Hoof, G. Vriend and E. E. Abola, *Nature* **1996**, 381, 272-272.
 91. I. K. McDonald and J. M. Thornton, *Journal of Molecular Biology* **1994**, 238, 777-793.
 92. M. Gilson and B. Honig, *Nature* **1987**, 330, 84.
 93. P. J. Goodford, *Journal of Medicinal Chemistry* **1985**, 28, 849-857.
 94. Molecular Discovery Ltd. *GRID 22*; Molecular Discovery Ltd.: 2004.
 95. T. J. A. Ewing and I. D. Kuntz, *Journal of Computational Chemistry* **1997**, 18, (9), 1175-1189.
 96. T. J. A. Ewing, S. Makino, A. G. Skillman and I. D. Kuntz, *Journal of Computer-Aided Molecular Design* **2001**, 15, 411-428.
 97. R. Wang, Y. Gao and L. Lai, *Journal of Molecular Modelling* **2000**, 6, 498-516.
 98. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Advanced Drug Delivery Reviews* **2001**, 46, (1-3), 3-26.
 99. Accelrys *Discover 2.9.7*, Accelrys Inc.: San Diego, 2001.
 100. D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, J. Wang, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R. J. Radmer, Y. Duan, J. Pitner, G. L. Massova, G. L. Siebel, U. C. Singh, P. K. Weiner and P. A. Kollman *AMBER 7*, 7; University Of California: San Francisco, 2002.
 101. P. S. Charifson, J. J. Corkery, M. A. Murcko and W. P. Walters, *Journal of Medicinal Chemistry* **1999**, 42, 5100-5109.
 102. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *Journal of Computational Chemistry* **1998**, 19, (14), 1639-1662.
 103. R. Wang, Y. Lu and S. Wang, *Journal of Medicinal Chemistry* **2003**, 46, 2287-2303.
 104. R. Wang, L. Lai and S. Wang, *Journal of Computer Aided Molecular Design* **2002**, 16, 11-26.

105. L. M. von Itzstein, W. Y. Wu, V. Phan Tho, B. Danylec and B. Jin Preparation of derivatives and analogs of 2-deoxy-2,3-didehydro-*N*-acetylneuraminic acid as antiviral agents. 9116320, 1991.
106. M. von Itzstein, W.-Y. Wu and B. Jin, *Carbohydrate Research* **1994**, 259, (2), 301-305.
107. M. von Itzstein, J. G. Barry and A. K. J. Chong, *Current Opinion on Therapeutic Patents* **1993**, 3, (12), 1755-1762.
108. M. von Itzstein and J. C. Dyason, *Sialidase inhibitors as anti-influenza drugs*. ed.; JAI Press Inc: Stamford, 1999.
109. W. P. Burmeister, R. W. H. Ruigrok and S. Cusack, *The EMBO Journal* **1992**, 11, (1), 49-56.
110. N. R. Taylor and M. von Itzstein, *Journal of Medicinal Chemistry* **1994**, 37, (5), 616-624.
111. R. A. Lamb and R. M. Krug, Orthomyxoviridae: The viruses and their replication. In *Fundamental Virology*, Third ed.; B. N. Fields, D. M. Knipe, P. M. Howely, R. M. Chanock, J. L. Melnick, T. P. Monath, B. Roizman and S. E. Straus, Ed. Lippincott-Raven: New York, 1996; pp 605-648.
112. J. N. Varghese, J. L. McKimm-Breschkin, J. B. Caldwell, A. A. Kortt and P. M. Colman, *Proteins: Structure, Function, and Genetics* **1992**, 14, 327-332.
113. J. E. Nielsen, K. V. Andersen, B. Honig, R. W. W. Hooft, G. Klebe, G. Vriend and R. C. Wade, *Protein Engineering* **1999**, 12, (8), 657-662.
114. A. C. Wallace, R. A. Laskowski and J. M. Thornton, *Protein Engineering* **1995**, 8, 127-134.
115. P. J. Goodford *GRID User Manual*, 19; Molecular Discovery Ltd.: 2001.
116. N. R. Taylor and M. von Itzstein, *Journal of Computer-Aided Molecular Design* **1996**, 10, 233-246.
117. M. von Itzstein, B. Jin, W.-Y. Wu and M. Chandler, *Carbohydrate Research* **1993**, 244, 181-185.
118. J. C. Castro-Palomino, B. Simon, O. Speer, M. Leist and R. R. Schmidt, *Chemistry a European Journal* **2001**, 7, (10), 2178-2184.
119. C. Press, *CRC Handbook of Chemistry and Physics*. 60 ed.; CRC Press: Florida, 1980.
120. P. Florio, R. J. Thomson, A. Alafaci, S. Abo and I. M. von, *Bioorganic and Medicinal Chemistry Letters* **1999**, 9, (14), 2065-2068.
121. T. Yamamoto, H. Kumazawa, K. Inami, T. Teshima and T. Shiba, *Tetrahedron Letters* **1992**, 33, (39), 5791-5794.
122. D. G. Higgins, J. D. Thompson and T. J. Gibson, *Nucleic Acids Research* **1994**, 22, 4673-4680.
123. *Accelrys InsightII 2000*, Accelrys Inc.: San Diego, 2000.
124. L. Holmquist, *FOA Reports* **1975**, 9, (3), 1-20.
125. H. Tuppy and A. Gottschalk, The structure of sialic acids and their quantitation. In *Glycoproteins. Their composition, structure and function.*, A. Gottschalk, Ed. Elsevier: Amsterdam, 1972; pp 403-449.
126. J. F. G. Vliegthart and J. P. Kamerling, Synthesis of sialic acids and sialic acid derivatives. In *Sialic Acids: Chemistry, Metabolism and Function. Cell Biology Monographs*, R. Schauer, Ed. Springer Verlag: Wien, New York, 1982; Vol. 10, pp 59-76.
127. E. Zbiral, Synthesis of sialic acid analogs and their behaviour towards the enzymes of sialic acid metabolism and hemagglutini X-31 of influenza A-virus. In *Carbohydrates- Synthetic methods and applications in medicinal chemistry.*, H. Ogura, A. Hasegawa and T. Suami, Ed. VCH: Weinheim, 1992; pp 304-339.

128. K. Furuhata, *Trends in Glycoscience and Glycotechnology* **2004**, 16, (89), 143-169.
129. R. Brossmer and L. Homquist, *Hoppe-Seyler 's Z. Physiol. Chem.* **1971**, 352, 1715-1719.
130. R. Roy and C. A. LaFerrière, *Canadian Journal of Chemistry* **1990**, 68, 2045-2054.
131. K. Wallimann and A. Vasella, *Helvetica Chimica Acta* **1990**, 73, 1359-1372.
132. W. Schmid, R. Christian and E. Zbiral, *Tetrahedron Letters* **1988**, 29, (30), 3643-3646.
133. H. Ogura and K. Furuhata, *Carbohydrate Research* **1986**, 158, 37-51.
134. X.-L. Sun, Y. Kanie, C.-T. Guo, O. Kanie, Y. Suzuki and C. H. Wong, *European Journal of Organic Chemistry* **2000**, 2643-2653.
135. P. Meindl and H. Tuppy, *Monatshefte Für Chemie* **1969**, 100, 1296-1306.
136. K. Okamoto, T. Kondo and T. Goto, *Bulletin of the Chemical Society of Japan*. **1987**, 60, 631-636.
137. R. Gantt, S. Millner and S. B. Binkley, *Biochemistry* **1964**, 3, (12), 1952-1960.
138. F. Baumberger and A. Vasella, *Helvetica Chimica Acta* **1986**, 69, 1205-1215.
139. N. Hartmann and E. Zbiral, *Monatshefte Für Chemie* **1989**, 120, (10), 899-906.
140. M. Hartmann, R. Christian and E. Zbiral, *Liebigs Annalen der Chemie* **1990**, 83-91.
141. D. R. Groves, J. C. Wilson and M. von Itzstein, *Australian Journal Of Chemistry* **1995**, 48, 1217-1220.
142. G. Estenne, A. Saroli and A. Doutheau, *Journal of Carbohydrate Chemistry* **1991**, 10, (2), 181-195.
143. E. Schreiner, R. Christian and E. Zbiral, *Liebigs Annalen der Chemie* **1990**, 93-97.
144. K. Ikeda, K. Sano, M. Ito, M. Saito, K. Hidari, T. Suzuki, Y. Suzuki and K. Tanaka, *Carbohydrate Research* **2001**, 330, (1), 31-41.
145. E. Zbiral, E. Schreiner and R. Christian, *Carbohydrate Research* **1989**, 194, C15-C18.
146. M. Hartmann, R. Christian and E. Zbiral, *Monatshefte Für Chemie* **1991**, 122, 111-125.
147. M. Chandler, M. J. Bamford, R. Conroy, B. Lamont, B. Patel, V. K. Patel, I. P. Steeples, R. Storer, N. G. Weir, M. Wright and C. Williamson, *Journal of the Chemical Society Perkin Transactions 1* **1995**, 9, 1173-1180.
148. F. Baumberger, A. Vasella and R. Schauer, *Helvetica Chimica Acta* **1988**, 71, 429-445.
149. A. Vasella and R. Wyler, *Helvetica Chimica Acta* **1990**, 73, 1742-1763.
150. H. Mack and R. Brossmer, *Tetrahedron Letters* **1987**, 28, (2), 191-194.
151. M. Salunkhe, M. Hartmann, W. Schmid and E. Zbiral, *Liebigs Annalen der Chemie* **1988**, 187-189.
152. E. Zbiral, H. H. Brandstetter, R. Christian and R. Schauer, *Liebigs Annalen der Chemie* **1987**, 781-786.
153. E. Zbiral, H. H. Brandstetter and E. P. Schreiner, *Monatshefte für Chemie* **1988**, 119, 127-141.
154. H. H. Brandstetter, E. Zbiral and G. Schulz, *Liebigs Annalen der Chemie* **1982**, 1-13.
155. H. H. Brandstetter and E. Zbiral, *Liebigs Annalen der Chemie* **1983**, 12, 2055-2065.
156. R. L. McLean, M. Suttajit, J. Beidler and R. J. Winzler, *Journal of Biological Chemistry* **1971**, 246, (3), 803-809.
157. D. C. M. Kong and M. von Itzstein, *Tetrahedron Letters* **1995**, 36, (6), 957-960.

158. K. Furuhashi and H. Ogura, *Chemical and Pharmaceutical Bulletin* **1989**, 37, (8), 2037-2040.
159. B. Reinhard and H. Faillard, *Liebigs Annalen der Chemie* **1994**, (2), 193-203.
160. S. Sepulveda-Boza and U. Stather, *Boletin de la Sociedad Chilena de Quimica* **1994**, 39, 299-303.
161. R. Isecke and R. Brossmer, *Tetrahedron* **1994**, 50, (25), 7445-7460.
162. M. Sharma, C. R. Petrie 3rd and W. Korytnyk, *Carbohydrate Research* **1988**, 175, 25-34.
163. H. J. r. Gross, A. B nsch, J. C. Paulson and R. Brossmer, *European Journal of Biochemistry* **1987**, 168, 595-602.
164. R. Kuhn, P. Lutz and D. L. MacDonald, *Chemische Berichte* **1966**, 99, (2), 611-617.
165. A. Claesson and K. Luthman, *Acta Chemica Scandinavia* **1982**, B36, 719-731.
166. V. Kumar, J. Kessler, M. E. Scott, B. H. TPatwardhan, S. W. Tanenbaum and M. Flashner, *Carbohydrate Research* **1981**, 94, 123-130.
167. J. M. Beau, R. Schauer, J. Haverkamp, J. P. Kamerling, L. Dorland and J. F. Vliegthart, *European Journal of Biochemistry* **1984**, 140, 203-208.
168. W. P. Burmeister, B. Henrissat, C. Bosso, S. Cusack and R. W. H. Ruigrok, *Structure* **1993**, 1, (1), 19-26.
169. H. Driguez, *Topics in Current Chemistry* **1997**, 187, 85-116.
170. S. Bennett, M. von Itzstein and M. J. Kiefel, *Carbohydrate Research* **1994**, 259, (2), 293-299.
171. W. K. C. Park, S. J. Meunier, D. Zanini and R. Roy, *Carbohydrate Letters* **1995**, 1, 179-184.
172. A. Fazli, S. J. Bradley, M. J. Kiefel, C. Jolly, I. H. Holmes and M. von Itzstein, *Journal of Medicinal Chemistry* **2001**, 44, (20), 3292-3301.
173. A. Marra and P. Sinay, *Carbohydrate Research* **1989**, 190, 317-322.
174. G. B. Kok, A. K. Norton and M. von Itzstein, *Synthesis Stuttgart* **1997**, 10, 1185&.
175. B. Yu, Q. Q. Ouyang, C. Li and Y. Z. Hui, *Journal of Carbohydrate Chemistry* **1996**, 15, (3), 297-302.
176. E. Schreiner, E. Zbiral, R. G. Kleincidam and R. Schauer, *Liebigs Annalen der Chemie* **1991**, 129-134.
177. G. B. Kok, D. Groves and M. von Itzstein, *Journal of the Chemical Society Perkin Transactions I* **1999**, 15, 2109-2115.
178. A. J. Gordon and R. A. Ford, *The chemist's companion: a handbook of practical data, techniques, and references*. ed.; John Wiley & Sons: Brisbane, 1972.
179. A. Hasegawa, K. Nakamura and M. Kiso, *Journal of Carbohydrate Chemistry* **1986**, 5, (1), 11-19.
180. P. Chand, Y. S. Babu, S. R. Rowland and T.-H. Lin Preparation of neuraminic acids and analogs useful for inhibiting paramyxovirus neuraminidase. In *PCT International Applications*. 20020308, 2002.
181. J. J. P. Stewart, I. Rossi, W.-P. Hu, G. C. Lynch, Y.-P. Liu, Y.-Y. Chuang, J. Li, C. J. Cramer, P. L. Fast and D. G. Truhlar *MOPAC*, 5.09mn; University Of Minnesota, Minneapolis: 1999.
182. T. Takimoto, G. L. Taylor, S. J. Crennell, R. A. Scroggs and A. Portner, *Virology* **2000**, 270, 208-214.
183. C. Ryan, V. Zaitsev, D. J. Tindal, J. C. Dyason, R. J. Thomson, I. V. Alymova, A. Portner, M. von Itzstein and G. Taylor, *Glycoconjugate Journal* **2006**, 23, 135-141.
184. G. P. Brady and P. F. W. Stouten, *Journal of Computer Aided Molecular Design* **2000**, 14, (4), 383-401.

185. A. Sali, L. Potterton, F. Yuan, H. van Vlijmen and M. Karplus, *Proteins Structure Function and Genetics* **1995**, 23, 318-326.
186. M. F. Sanner, *Journal of Molecular Graphics and Modelling* **1999**, 17, 57-61.
187. D. D. Perrin and W. L. F. Armarego, *Purification of Laboratory Chemicals*. 3rd ed.; Pergamon Press: Oxford, 1988.
188. J. C. Castro-Palomino, Y. E. Tscetkov, R. Schneider and R. R. Schmidt, *Tetrahedron Letters* **1997**, 38, (39), 6837-6840.

Appendix A: Computer Input and Parameter Files

Discover

Minimisation input file for distance dependant dielectric minimisation

```
! INPUT FILE FOR DISCOVER GENERATED BY INSIGHT
!
!
! overlap = 0.01
!
! begin simulation
! * add-automatic bond torsion valence out-of-plane
! reduce
!
! set dielectric = 1.0*r
! cutoff = 15.0
! cutdis = 11.5
! swtdis = 1.5
! nrneib = 20
! igrpck = 0
!
!
!
! fixed atom list generation; brief
! * add heavy atoms molecules 1 to 131
!
! minimize using steepest descents
! * until the maximum derivative is less than 10.0
! * with no cross terms and no morse functions
!
!
!
! fixed atom list generation; brief
! * clear
!
! archive as file number 1
!
! fixed atom list generation; brief
! * add main molecule 2
!
!
!
! minimize using steepest descents
! * until the maximum derivative is less than 10.0
! * with no cross terms and no morse functions
!
!
!
! fixed atom list generation; brief
! * clear
!
! archive as file number 2
!
! tethered atoms list generation
! * add main
! * molecule 2
!
!
! template force until the maximum derivative is less than 1.0
! * using steepest descents
```

```

* with no morse functions and
* no cross terms
* and a forcing constant of 500.0 kcal/angstrom
!
archive as file number 3
!
template force until the maximum derivative is less than 1.0
* using steepest descents
* with no morse functions and
* no cross terms
* and a forcing constant of 100.0 kcal/angstrom
!
archive as file number 4
!
template force until the maximum derivative is less than 1.0
* using steepest descents
* with no morse functions and
* no cross terms
* and a forcing constant of 50.0 kcal/angstrom
!
archive as file number 5
!
tethered atoms list generation
* clear
!
archive as file number 6
!
minimize using steepest descents
* until the maximum derivative is less than 5.0 kcal/angstrom
* with no cross terms and no morse functions
!
archive as file number 7
!
minimize using conjugate gradients
* until the maximum derivative is less than 0.1 kcal/angstrom
* with no cross terms and no morse functions
!
!
end

```

Minimisation input file for constant dielectric minimisation

```

! INPUT FILE FOR DISCOVER GENERATED BY INSIGHT
!
!
overlap = 0.01
!
begin simulation
* add-automatic bond torsion valence out-of-plane
reduce
!
set dielectric = 80
cutoff = 15.0
cutdis = 11.5
swtdis = 1.5
nrneib = 20
!

```

```

!
!
    fixed atom list generation; brief
    * add heavy atoms molecules 1 to 131
!
    minimize using steepest descents
    * until the maximum derivative is less than 10.0
    * with no cross terms and no morse functions
!
!
    fixed atom list generation; brief
    * clear
!
    archive as file number 1
!
    fixed atom list generation; brief
    * add main molecule 2
!
!
    minimize using steepest descents
    * until the maximum derivative is less than 10.0
    * with no cross terms and no morse functions
!
!
    fixed atom list generation; brief
    * clear
!
    archive as file number 2
!
    tethered atoms list generation
    * add main
    * molecule 2
!
    template force until the maximum derivative is less than 1.0
    * using steepest descents
    * with no morse functions and
    * no cross terms
    * and a forcing constant of 500.0 kcal/angstrom
!
    archive as file number 3
!
    template force until the maximum derivative is less than 1.0
    * using steepest descents
    * with no morse functions and
    * no cross terms
    * and a forcing constant of 100.0 kcal/angstrom
!
    archive as file number 4
!
    template force until the maximum derivative is less than 1.0
    * using steepest descents
    * with no morse functions and
    * no cross terms
    * and a forcing constant of 50.0 kcal/angstrom
!
    archive as file number 5
!

```

```

    tethered atoms list generation
    * clear
!
    archive as file number 6
!
    minimize using steepest descents
    * until the maximum derivative is less than 5.0 kcal/angstrom
    * with no cross terms and no morse functions
!
    archive as file number 7
!
    minimize using conjugate gradients
    * until the maximum derivative is less than 0.1 kcal/angstrom
    * with no cross terms and no morse functions
!
!
    end

```

Conformational search of NDV HN active site using Neu5Ac2en

```

! INPUT FILE FOR DISCOVER GENERATED BY INSIGHT
!
!
    overlap = 0.01
    cutoff = 15.0
    cutdis = 11.5
    swtdis = 1.5
    dseed = 666

    begin simulation
    * add-automatic bond torsion valence out-of-plane
    reduce
!
    set dielectric = 1.000*r
!
!
!
    Fixed atom list generation; brief
    * add all
    * molecule 2
    * delete all
    * molecule 2 residue ILE A163
    * delete all
    * molecule 2 residue PRO A166
    * delete all
    * molecule 2 residue SER A170 to PHE A178
    * delete all
    * molecule 2 residue TYR A187 to LEU A206
    * delete all
    * molecule 2 residue LEU A229 to VAL A240
    * delete all
    * molecule 2 residue ASP- A248 to GLY A273
    * delete all
    * molecule 2 residue TYR A281
    * delete all
    * molecule 2 residue GLU- A283
    * delete all

```


- * molecule 2 residue ASP- A285 to LEU A286
- * delete all
- * molecule 2 residue VAL A288
- * delete all
- * molecule 2 residue PHE A292
- * delete all
- * molecule 2 residue ASP- A294 to GLY A305
- * delete all
- * molecule 2 residue PHE A314 to PRO A322
- * delete all
- * molecule 2 residue ASP- A327
- * delete all
- * molecule 2 residue GLN A330
- * delete all
- * molecule 2 residue ALA A355
- * delete all
- * molecule 2 residue TYR A337
- * delete all
- * molecule 2 residue ALA A355 to LEU A374
- * delete all
- * molecule 2 residue ILE A376
- * delete all
- * molecule 2 residue LEU A388 to PRO A391
- * delete all
- * molecule 2 residue ASN A393
- * delete all
- * molecule 2 residue VAL A395 to LEU A405
- * delete all
- * molecule 2 residue LEU A413 to TYR A427
- * delete all
- * molecule 2 residue LEU A438
- * delete all
- * molecule 2 residue PHE A447 to ALA A457
- * delete all
- * molecule 2 residue ALA A459
- * delete all
- * molecule 2 residue PRO A462 to PRO A473
- * delete all
- * molecule 2 residue THR A490 to PRO A501
- * delete all
- * molecule 2 residue THR A522 to SER A529
- * delete all
- * molecule 2 residue ILE A545 to ASN A550
- * delete all
- * molecule 2 residue PHE A553 to VAL A559
- * add all
- * molecule 3 to 131
- * delete all
- * molecule 18 residue HOH Z16H
- * delete all
- * molecule 20 residue HOH Z18H
- * delete all
- * molecule 30 residue HOH Z28H
- * delete all
- * molecule 31 residue HOH Z29H
- * delete all
- * molecule 33 residue HOH Z31H

```

* delete all
* molecule 54 residue HOH Z52H
* delete all
* molecule 55 residue HOH Z53H
* delete all
* molecule 56 residue HOH Z54H
* delete all
* molecule 57 residue HOH Z55H
* delete all
* molecule 60 residue HOH Z58H
* delete all
* molecule 72 residue HOH Z70H
* delete all
* molecule 74 residue HOH Z72H
* delete all
* molecule 80 residue HOH Z78H
* delete all
* molecule 84 residue HOH Z82H
* delete all
* molecule 88 residue HOH Z86H
* delete all
* molecule 89 residue HOH Z87H
* delete all
* molecule 90 residue HOH Z88H
* delete all
* molecule 91 residue HOH Z89H
* delete all
* molecule 92 residue HOH Z90H
* delete all
* molecule 98 residue HOH Z96H
* delete all
* molecule 100 residue HOH Z98H
* delete all
* molecule 106 residue HOH Z104H
* delete all
* molecule 107 residue HOH Z105H
* delete all
* molecule 112 residue HOH Z110H
* delete all
* molecule 113 residue HOH Z111H
* delete all
* molecule 123 residue HOH Z121H
* delete all
* molecule 124 residue HOH Z122H
* delete all
* molecule 125 residue HOH Z123H
* delete all
* molecule 126 residue HOH Z124H
* delete all
* molecule 130 residue HOH Z128H
* delete all
* molecule 131 residue HOH Z129H

```

!

Minimize

```

* no cross terms
* no morse
* using steep descents

```

```

* until the maximum derivative is less than 1.000000000 kcal/A
!
  Minimize
  * no cross terms
  * no morse
  * using conjugate gradients
  * until the maximum derivative is less than 0.01 kcal/A
!
  archive as file number 1
!
  ifile = 2
!
lop1 retrieve file number 1
!
  iloop = 0
!
  initialize dynamics
  * for 2000 iterations
  * at 350.000 K
  * steps of 1.000 fs
  * no cross terms
  * no morse
!
lop2 resume dynamics
  * for 2000 iterations
  * at 350.000 K
  * steps of 2.000 fs
  * no cross terms
  * no morse
!
  iloop = iloop + 1
!
  minimize
  * no cross terms
  * no morse
  * using steepest descents
  * until the maximum derivative is less than 1 kcal/A
!
  minimize
  * no cross terms
  * no morse
  * using conjugate gradients
  * until the maximum derivative is less than 0.01 kcal/A
!
  archive as file number ifile
  ifile = ifile + 1
  if iloop.eq.1 then lop2
  if ifile.le.14 then lop1
!
end

```

Discover – Interaction Energies

```

!
!
begin simulation
* add-automatic bond torsion valence out-of-plane

```

```

!
  reduce output
!
  set dielectric = 1.0*r
!
  print residue-residue interactions
  * add molecule 1
  * add molecule 2
!
  end

```

DOCK

GRID parameter file

```

compute_grids yes
grid_spacing 0.3
output_molecule no
contact_score yes
contact_cutoff_distance 4.5
chemical_score yes
energy_score yes
energy_cutoff_distance 10
atom_model a
attractive_exponent 6
repulsive_exponent 12
distance_dielectric yes
dielectric_factor 4
bump_filter yes
bump_overlap 0.75
receptor_file ../1_struc/pro1/pro1.mol2
box_file box1.pdb
vdw_definition_file /programs/dock/4.0.1/parameter/vdw.defn
chemical_definition_file /programs/dock/4.0.1/parameter/chem.defn
score_grid_prefix grid1

```

Static DOCK parameter file

```

flexible_ligand no
orient_ligand no
score_ligand yes
minimize_ligand no
multiple_ligands no
intermolecular_score yes
gridded_score yes
grid_version 4
contact_score no
chemical_score no
energy_score yes
atom_model a
vdw_scale 1
electrostatic_scale 1
ligand_atom_file lig1.mol2
score_grid_prefix ../3_grid/grid1
vdw_definition_file /programs/dock/4.0.1/parameter/vdw.defn
ligand_energy_file lig1_nrg.mol2

```

Flexible DOCK parameter file

flexible_ligand	yes
orient_ligand	yes
score_ligand	yes
minimize_ligand	yes
multiple_ligands	no
random_seed	0
anchor_search	no
torsion_drive	yes
clash_overlap	0.5
conformation_cutoff_factor	5
torsion_minimize	yes
match_receptor_sites	yes
random_search	no
ligand_centers	no
automated_matching	yes
maximum_orientations	500
write_configurations	yes
write_configuration_total	100
intramolecular_score	yes
intermolecular_score	yes
gridded_score	yes
grid_version	4
bump_filter	yes
bump_maximum	2
contact_score	no
chemical_score	no
energy_score	yes
energy_cutoff_distance	10
distance_dielectric	yes
dielectric_factor	4
attractive_exponent	6
repulsive_exponent	12
atom_model	a
vdw_scale	1
electrostatic_scale	1
energy_minimize	yes
initial_translation	1
initial_rotation	0.1
initial_torsion	10
maximum_iterations	100
energy_convergence	0.1
maximum_cycles	1
ligand_atom_file	dana.mol2
receptor_site_file	../1_struc/hn1.sph
score_grid_prefix	../3_grid/grid
vdw_definition_file	/programs/dock/4.0.1/parameter/vdw.defn
flex_definition_file	/programs/dock/4.0.1/parameter/flex.defn
flex_drive_file	/programs/dock/4.0.1/parameter/flex_drive.tbl
ligand_energy_file	dana1_nrg.mol2

Ligbuilder

Pocket

```
#
# input files
#
RECEPTOR_FILE          /usr/compchem/david/book0032/page008/minhn.pdb
LIGAND_FILE              /usr/compchem/david/book0032/page008/dana.mol2
PARAMETER_DIRECTORY     /programs/ligbuilder/LigBuilderv1.2/parameter/
#
# output files
#
POCKET_ATOM_FILE        /usr/compchem/david/book0032/page008/pocket/minhnatom.txt
POCKET_GRID_FILE        /usr/compchem/david/book0032/page008/pocket/minhngrid.txt
#
# key interaction sites and pharmacophore
#
KEY_SITE_FILE           /usr/compchem/david/book0032/page008/pocket/minhnsite.pdb
PHARMACOPHORE_PDB_FILE  /usr/compchem/david/book0032/page008/pocket/minhnpharm.pdb
PHARMACOPHORE_TXT_FILE  /usr/compchem/david/book0032/page008/pocket/minhnpharm.txt
MINIMAL_FEATURE_DISTANCE 3.00
MAXIMAL_FEATURE_DISTANCE 8.00
#
#
```

Grow

```
#
# input files
#
SEED_LIGAND_FILE /usr/compchem/david/book0032/page008/grow/seed2.mol2
POCKET_ATOM_FILE /usr/compchem/david/book0032/page008/pocket/minhnatom.txt
POCKET_GRID_FILE /usr/compchem/david/book0032/page008/pocket/minhngrid.txt
#
# force field directory
#
PARAMETER_DIRECTORY /programs/ligbuilder/LigBuilderv1.2/parameter/
#
# fragment libraries
#
BUILDING_BLOCK_LIBRARY /programs/ligbuilder/LigBuilderv1.2/fragment.mdb/
FORBIDDEN_STRUCTURE_LIBRARY /programs/ligbuilder/LigBuilderv1.2/forbidden.mdb/
TOXIC_STRUCTURE_LIBRARY /programs/ligbuilder/LigBuilderv1.2/toxicity.mdb/
#
# structural construction parameters
#
GROWING_PROBABILITY          1.00
LINKING_PROBABILITY          0.50
MUTATION_PROBABILITY         0.50
#
# chemical viability rules
#
```

```

APPLY_CHEMICAL_RULES                YES
APPLY_FORBIDDEN_STRUCTURE_CHECK      YES
APPLY_TOXIC_STRUCTURE_CHECK          YES
MAXIMAL_MOLECULAR_WEIGHT             600
MINIMAL_MOLECULAR_WEIGHT             300
MAXIMAL_LOGP                         6.00
MINIMAL_LOGP                         3.00
MAXIMAL_HB_DONOR_ATOM                6
MINIMAL_HB_DONOR_ATOM                2
MAXIMAL_HB_ACCEPTOR_ATOM             6
MINIMAL_HB_ACCEPTOR_ATOM             2
MAXIMAL_PKD                         10.00
MINIMAL_PKD                         5.00
#
# genetic algorithm parameters
#
NUMBER_OF_GENERATION                 20
NUMBER_OF_POPULATION                 3000
NUMBER_OF_PARENTS                    200
ELITISM_RATIO                       0.10
SIMILARITY_CUTOFF                   0.90
#
# output files
#
POPULATION_RECORD_FILE
    /usr/compchem/david/book0032/page008/grow/population2.lig
LIGAND_COLLECTION_FILE
    /usr/compchem/david/book0032/page008/grow/ligands2.lig
#

```

Process

```

#
# input files
#
LIGAND_COLLECTION_FILE
    /usr/compchem/david/book0032/page008/grow/population.lig
#
# chemical rules
#
MAXIMAL_MOLECULAR_WEIGHT             600
MINIMAL_MOLECULAR_WEIGHT             300
MAXIMAL_LOGP                         6.00
MINIMAL_LOGP                         3.00
MAXIMAL_PKD                         10.00
MINIMAL_PKD                         5.00
#
# similarity threshold
#
SIMILARITY_CUTOFF                   0.90
#
# output files
#
NUMBER_OF_OUTPUT_MOLECULES           200
OUTPUT_DIRECTORY                    /usr/compchem/david/book0032/page008/process/c4o
#

```

GRID Input Parameter Files

GRIN parameter file

```
:  
:  
:      MOLECULAR DISCOVERY LIMITED  
:      *****  
:  
:      Command File for Programme GRIN  
:  
:  
: Assign Channel Numbers and Output File Names:  
: -----  
:  
: LOUT    6  
: LOUT pro.lout  
: KOUT    20  
: KOUT pro.kout  
:  
: Assign Channel Numbers and Input File Names:  
: -----  
:  
: INKO    11  
: INKO pro.pdb  
: INAT    10  
: INAT /programs/grid19/grub.dat  
:  
: Provide Control Parameters:  
: -----  
:  
: ALHY    0.040  
: NEHY    1  
: IHVA    0  
: LEVL    3  
: MOVE    1  
: QQHY    0.000  
: VDHY    0.600  
: IEND  
:  
:Record the name of the executable Programme and the name of  
:the Default Directory as comment lines starting with a colon:  
:  
:PROG: /programs/grid21/grin  
:DEFA:
```

GRID parameter file

```
:  
:  
:      MOLECULAR DISCOVERY LIMITED  
:      *****  
:  
:      Command File for Programme GRID  
:  
:  
:  
:  
:
```



```

: Assign Channel Numbers and Output File names:
: -----
:
LONT    6
LONT pro.lont
KONT    20
KONT pro.kont
:
: Assign Channel Numbers and Input File names:
: -----
:
INPT    10
INPT pro.kout
:
: Provide Control Parameters:
: -----
:
CLER     5.000
DEEP     5.000
DPRO     4.000
DWAT    80.000
EACH     5.000
EMAX     5.000
FARH     5.000
FARR     8.000
KWIK     0
LEAU     0
LENG     30
LEVL     1
LIST     1
MOVE     0
NETA     0
NPLA     1
NUMB     1
VALU     0.000
OH2
TOPX    47.430
TOPY    61.780
TOPZ    13.560
BOTX    15.880
BOTY    25.260
BOTZ   -23.140
IEND
OH2
    0  1
:
:Record the name of the executable Programme and the name of
:the Default Directory as comment lines starting with a colon:
:
:PROG: /programs/grid21/grid
:DEFA:

```

AutoDock

AutoGrid GPF File

```
receptor hn.pdbqs          # macromolecule
gridfld hn.maps.fld       # grid_data_file
npts 60 60 60             # num.grid points in xyz
spacing 0.375             # spacing(A)
gridcenter 33.38 45.07 -4.44 # xyz-coordinates or auto
types CAONSH              # atom type names
smooth 0.5                # store minimum energy w/in rad(A)
map hn.C.map              # atom-specific affinity map
nbp_r_eps 4.00 0.0222750 12 6 # C-C lj
nbp_r_eps 3.75 0.0230026 12 6 # C-N lj
nbp_r_eps 3.60 0.0257202 12 6 # C-O lj
nbp_r_eps 4.00 0.0257202 12 6 # C-S lj
nbp_r_eps 3.00 0.0081378 12 6 # C-H lj
nbp_r_eps 3.00 0.0081378 12 6 # C-H lj
nbp_r_eps 3.00 0.0081378 12 6 # C-H lj
sol_par 12.77 0.6844      # C atomic fragmental volume, solvation parameters
constant 0.000           # C grid map constant energy
map hn.A.map              # atom-specific affinity map
nbp_r_eps 4.00 0.0222750 12 6 # A-C lj
nbp_r_eps 3.75 0.0230026 12 6 # A-N lj
nbp_r_eps 3.60 0.0257202 12 6 # A-O lj
nbp_r_eps 4.00 0.0257202 12 6 # A-S lj
nbp_r_eps 3.00 0.0081378 12 6 # A-H lj
nbp_r_eps 3.00 0.0081378 12 6 # A-H lj
nbp_r_eps 3.00 0.0081378 12 6 # A-H lj
sol_par 10.80 0.1027     # A atomic fragmental volume, solvation parameters
constant 0.000           # A grid map constant energy
map hn.O.map              # atom-specific affinity map
nbp_r_eps 3.60 0.0257202 12 6 # O-C lj
nbp_r_eps 3.35 0.0265667 12 6 # O-N lj
nbp_r_eps 3.20 0.0297000 12 6 # O-O lj
nbp_r_eps 3.60 0.0297000 12 6 # O-S lj
nbp_r_eps 1.90 0.3280000 12 10 # O-H hb
nbp_r_eps 1.90 0.3280000 12 10 # O-H hb
nbp_r_eps 1.90 0.3280000 12 10 # O-H hb
sol_par 0.00 0.0000      # O atomic fragmental volume, solvation parameters
constant 0.236           # O grid map constant energy
map hn.N.map              # atom-specific affinity map
nbp_r_eps 3.75 0.0230026 12 6 # N-C lj
nbp_r_eps 3.50 0.0237600 12 6 # N-N lj
nbp_r_eps 3.35 0.0265667 12 6 # N-O lj
nbp_r_eps 3.75 0.0265667 12 6 # N-S lj
nbp_r_eps 1.90 0.3280000 12 10 # N-H hb
nbp_r_eps 1.90 0.3280000 12 10 # N-H hb
nbp_r_eps 1.90 0.3280000 12 10 # N-H hb
sol_par 0.00 0.0000      # N atomic fragmental volume, solvation parameters
constant 0.000           # N grid map constant energy
map hn.S.map              # atom-specific affinity map
nbp_r_eps 4.00 0.0257202 12 6 # S-C lj
nbp_r_eps 3.75 0.0265667 12 6 # S-N lj
nbp_r_eps 3.60 0.0297000 12 6 # S-O lj
nbp_r_eps 4.00 0.0297000 12 6 # S-S lj
nbp_r_eps 2.50 0.0656000 12 10 # S-H hb
```

```

nbp_r_eps 2.50 0.0656000 12 10 # S-H hb
nbp_r_eps 2.50 0.0656000 12 10 # S-H hb
sol_par 0.000 0.000 #S atomic fragmental volume, solvation parameters
constant 0.000 #S grid map constant energy
map hn.H.map # atom-specific affinity map
nbp_r_eps 3.00 0.0081378 12 6 # H-C lj
nbp_r_eps 1.90 0.3280000 12 10 # H-N hb
nbp_r_eps 1.90 0.3280000 12 10 # H-O hb
nbp_r_eps 2.50 0.0656000 12 10 # H-S hb
nbp_r_eps 2.00 0.0029700 12 6 # H-H lj
nbp_r_eps 2.00 0.0029700 12 6 # H-H lj
nbp_r_eps 2.00 0.0029700 12 6 # H-H lj
sol_par 0.00 0.0000 # H atomic fragmental volume, solvation parameters
constant 0.118 # H grid map constant energy
elecmap hn.e.map # electrostatic potential map
dielectric -0.1146 # <0, distance-dep.diel;>0, constant
#

```

AutoDock DPF File

```

seed time pid # for random number generator
types CANOH # atom type names
fld hn.maps.fld # grid data file
map hn.C.map # C-atomic affinity map file
map hn.A.map # A-atomic affinity map file
map hn.N.map # N-atomic affinity map file
map hn.O.map # O-atomic affinity map file
map hn.H.map # H-atomic affinity map file
map hn.e.map # electrostatics map file

move mol1_lig.out.pdbq # small molecule file
about 33.356 41.951 -3.463 # small molecule center

# Initial Translation, Quaternion and Torsions
tran0 random # initial coordinates/A or "random"
quat0 random # initial quaternion or "random"
ndihe 12 # number of initial torsions
dihe0 random # initial torsions

torsdof 12 0.3113 # num. non-Hydrogen torsional DOF & coeff.
#ligand_is_not_inhibitor # uncomment if small molecule is substrate or T.S.

# Initial Translation, Quaternion and Torsion Step Sizes and Reduction Factors
tstep 2.0 # translation step/A
qstep 50.0 # quaternion step/deg
dstep 50.0 # torsion step/deg
trnrf 1. # trans reduction factor/per cycle
quarf 1. # quat reduction factor/per cycle
dihrf 1. # tors reduction factor/per cycle

# Hard Torsion Constraints
#hardtorcon 1 -180. 30. # constrain torsion, num., angle(deg), range(deg)

# Internal Non-Bonded Parameters
intnbp_r_eps 4.00 0.0222750 12 6 #C-C lj
intnbp_r_eps 4.00 0.0222750 12 6 #C-A lj
intnbp_r_eps 3.75 0.0230026 12 6 #C-N lj

```

```

intnbp_r_eps 3.60 0.0257202 12 6 #C-O lj
intnbp_r_eps 3.00 0.0081378 12 6 #C-H lj
intnbp_r_eps 0.00 0.0000000 12 6 #A-A lj
intnbp_r_eps 3.75 0.0230026 12 6 #A-N lj
intnbp_r_eps 3.60 0.0257202 12 6 #A-O lj
intnbp_r_eps 3.00 0.0081378 12 6 #A-H lj
intnbp_r_eps 3.50 0.0237600 12 6 #N-N lj
intnbp_r_eps 3.35 0.0265667 12 6 #N-O lj
intnbp_r_eps 2.75 0.0084051 12 6 #N-H lj
intnbp_r_eps 3.20 0.0297000 12 6 #O-O lj
intnbp_r_eps 1.90 0.3280000 12 10 #O-H hb
intnbp_r_eps 2.00 0.0029700 12 6 #H-H lj

#intelec 0.1146      # calculate internal electrostatic energy

# Simulated Annealing Parameters
#rt0 616.           # SA: initial RT
#trtf 0.95          # SA: RT reduction factor/per cycle
#linear_schedule    # SA: do not use geometric cooling
#runs 10            # SA: number of runs
#cycles 50          # SA: cycles
#accs 100           # SA: steps accepted
#rejs 100           # SA: steps rejected
#select m           # SA: minimum or last

# Trajectory Parameters (Simulated Annealing Only)
#trjfrq 100         # trajectory frequency
#trjbeg 1           # start trj output at cycle
#trjend 50          # end trj output at cycle
#trjout mol11_lig.out.trj # trajectory file
#trjsel E           # A=acc only;E=either acc or rej

#watch mol11_lig.out.watch.pdb # real-time monitoring file

outlev 1            # diagnostic output level

# Docked Conformation Clustering Parameters for "analysis" command
rmstol 1.0          # cluster tolerance (Angstroms)
rmsref mol11_lig.out.pdbq # reference structure file for RMS calc.
#rmsnosym           # do no symmetry checking in RMS calc.
write_all           # write all conformations in a cluster

extnrg 1000.        # external grid energy
e0max 0.10000       # max. allowable initial energy, max. num. retries

# Genetic Algorithm (GA) and Lamarckian Genetic Algorithm Parameters (LGA)
ga_pop_size 50      # number of individuals in population
ga_num_evals 1000000 # maximum number of energy evaluations
ga_num_generations 27000 # maximum number of generations
ga_elitism 1        # num. of top individuals that automatically survive
ga_mutation_rate 0.02 # rate of gene mutation
ga_crossover_rate 0.80 # rate of crossover
ga_window_size 10   # num. of generations for picking worst individual
ga_cauchy_alpha 0   # ~mean of Cauchy distribution for gene mutation
ga_cauchy_beta 1    # ~variance of Cauchy distribution for gene mutation
set_ga              # set the above parameters for GA or LGA

```

```

# Local Search (Solis & Wets) Parameters (for LS alone and for LGA)
sw_max_its 300      # number of iterations of Solis & Wets local search
sw_max_succ 4      # number of consecutive successes before changing rho
sw_max_fail 4      # number of consecutive failures before changing rho
sw_rho 1.0         # size of local search space to sample
sw_lb_rho 0.01     # lower bound on rho
ls_search_freq 0.06 # probability of performing local search on an indiv.
set_psw1          # set the above pseudo-Solis & Wets parameters

# Perform Dockings
#do_local_only 50    # do only local search
#do_global_only 10 # do only global search (traditional GA)

#simanneal          # do as many SA runs as set by the "runs" command above

ga_run 10           # do this many GA or LGA runs

# Perform Cluster Analysis
analysis           # do cluster analysis on results
# dpf3gen.awk 3.0.5 #

```

Modeller

Modeller Alignment File

```

>P1;P_1V3D
structureX:P_1V3D : 142:A: 572:A:unknown:unknown:-1:-1
ITHDVGIKPLNPDDFWRCTSGLPMLKTPKIRLMPGPGLLAMPTTVDGCIRTPSLVINDLI
YAYTSNLITRGCDIGKSYQVLQIGIITVNSDLVPDLNPRISHTFNINDNRKSCSLALLNT
DVYQLCSTPKVDERSDYASPGIEDIVLDIVNYDGSISTTRFKNNNISFDQPYAALYPSVG
PGIYYKGKIIFLGYGGL-----EHPINENVICNTTGCPGKTQRDCNQASHSPWFSDRR
MVNSIIVVDKGLNSIPKLKVTISMQRQNYWGSEGRLLLLGNKIYIYTRSTSWHSLQLG
IIDITDYSIRIKWTWHNVLSRPGNNECPWGHSCPDGCITGVYTDAYPL----NPTGSIVSS
VILDSQKSRVNPVITYSTATERVNELAILNRTLSTAGYTTTSC--ITHYNGGYCFHIVEINHK
SLNTLQPMFLFKTEIPKSCS*
>P1;P_1E8V
structureX:P_1E8V : 124:A: 570:A:unknown:unknown:-1:-1
GAPIHDPDFIGGIGKELIVDNASDVTSFYPSAFQEHLNFIPAPTTGSGCTRIPSFDMSATH
YCYTHNVILSGCRDHSLSHQYLALGVLRRTATGRIFFTLRSISLDDTQNRKSCSVSATP
LGCDMLCSKVTEETEEEDYNSA--VPTLMAHGRLGFDGQYHEKDLDVTLFEDWVANYP
GVGGGSFIDGRVWFSVYGGGLKPNSPSDTVQEGKYVIYKRYNDTCPDEQDYQIRMAKSS
YKPGRFGGKRIQQAILSIKVSTSLGEDPVLTVPPNTVTLMGAEGRLTVGTSHFLYQRGS
SYFSPALLYPMTVSNKTATLHSPYTFNAFTRPGSIPCQASARCPNSCVTGVYTDYPLIF
YRNHTLRGVFGTMLDSEQARLNPASAVFDSTSRSRITRVSSSSTKAA YTTSTCFKVVKT
NKTYCLSAEISNTLFGFRIVPLLVEILKNDG*

```

MODEL.TOP FILE

```

# Homology modelling by the MODELLER TOP routine 'model'.

INCLUDE          # Include the predefined TOP routines

SET OUTPUT_CONTROL = 1 1 1 1 1  # uncomment to produce a large log file
SET ALNFILE = 'para1_alignment2.ali'  # alignment filename
SET KNOWN = '1e8v'          # codes of the templates
SET SEQUENCE = 'par1'       # code of the target

```

```

SET ATOM_FILES_DIRECTORY = '/usr/compchem/david/sequences/test1'
SET STARTING_MODEL= 1      # index of the first model
SET ENDING_MODEL = 1       # index of the last model
                           # (determines how many models to calculate)

```

```

CALL ROUTINE = 'model'      # do homology modelling

```

Amber

Leap Input File

```

source leaprc.gaff
mods = loadAmberParams frcmod1
loadAmberPrep "mol1_lig1.prep"
lig = loadPDB "mol1_lig1.pdb"
check lig
hnh = loadPDB "hnhfix.pdb"
bond hnh.172.SG hnh.196.SG
bond hnh.186.SG hnh.247.SG
bond hnh.238.SG hnh.251.SG
bond hnh.344.SG hnh.461.SG
bond hnh.455.SG hnh.465.SG
bond hnh.531.SG hnh.542.SG
check hnh
tot = combine {lig hnh}
saveAmberParm tot tot1.top tot1.crd
savePdb tot tot1.pdb
quit

```

PTRAJ

```

trajin min1.rstst
trajout min1.mdcrd
go

```

MMGBSA

```

#
# Input parameters for mm_pbsa.pl
# This example just generates snapshots from a trajectory file
#
# Holger Gohlke
# 08.01.2002
#
#####
@GENERAL
#
# General parameters
# 0: means NO; >0: means YES
#
# mm_pbsa allows to calculate (absolute) free energies for one molecular
# species or a free energy difference according to Receptor + Ligand = Complex,
#  $\Delta G = G(\text{Complex}) - G(\text{Receptor}) - G(\text{Ligand})$ .
#
# PREFIX - To the prefix, "{_com, _rec, _lig}.crd.Number" is added during
# generation of snapshots as well as during mm_pbsa calculations.
# PATH - Specifies the location where to store or get snapshots.

```

```

# COMPLEX - Set to 1 if free energy difference is calculated.
# RECEPTOR - Set to 1 if either (absolute) free energy or free energy difference
#       are calculated.
# LIGAND - Set to 1 if free energy difference is calculated.
#
# COMPT - Name of the parmtop file for the complex (not necessary for option GC).
# RECPT - Name of the parmtop file for the receptor (not necessary for option GC).
# LIGPT - Name of the parmtop file for the ligand (not necessary for option GC).
#
# GC - Snapshots are generated from trajectories (see below).
# AS - Residues are mutated during generation of snapshots from trajectories.
# DC - Decompose the free energies in single contributions (only works with MM and GB
yet).
#
# MM - Calculation of gasphase energies using sander.
# GB - Calculation of desolvation free energies using the GB models in sander (see below).
# PB - Calculation of desolvation free energies using delphi (see below).
# MS - Calculation of nonpolar contributions to desolvation using molsurf (see below).
#     If MS == 0, nonpolar contributions are calculated with the LCPO method in sander.
# NM - Calculation of entropies with nmode.
#
PREFIX      lig1
PATH        ./
COMPLEX      1
RECEPTOR   1
LIGAND       1
#
GC           1
AS           0
DC           0
#
MM           0
GB           0
PB           0
MS           0
#
NM           0
#
#####
@MAKECRD
#
# The following parameters are passed to make_crd_hg, which extracts snapshots
# from trajectory files.
#
# BOX - "YES" means that periodic boundary conditions were used during MD
#       simulation and that box information has been printed in the trajectory
#       files; "NO" means opposite.
# NTOTAL - Total number of atoms per snapshot printed in the trajectory file
#           (including water, ions, ...).
# NSTART - Start structure extraction from NSTART snapshot.
# NSTOP - Stop structure extraction at NSTOP snapshot.
# NFREQ - Every NFREQ structure will be extracted from the trajectory.
#
# NUMBER_LIG_GROUPS - Number of subsequent LSTART/LSTOP combinations to
extract
#                   atoms belonging to the ligand.
# LSTART - Number of first ligand atom in the trajectory entry.

```

```

# LSTOP - Number of last ligand atom in the trajectory entry.
# NUMBER_REC_GROUPS - Number of subsequent RSTART/RSTOP combinations to
extract
#          atoms belonging to the receptor.
# RSTART - Number of first receptor atom in the trajectory entry.
# RSTOP - Number of last receptor atom in the trajectory entry.
# Note: If only one molecular species is to be extracted, use only the receptor
# parameters (NUMBER_REC_GROUPS, RSTART, RSTOP).
#
BOX          NO
NTOTAL       6855
NSTART       1
NSTOP        1
NFREQ        1
#
NUMBER_LIG_GROUPS 1
LSTART       1
LSTOP        57
NUMBER_REC_GROUPS 1
RSTART       58
RSTOP        6855
#
#####
@TRAJECTORY
#
# Trajectory names
#
# The following trajectories are used to extract snapshots with the make_crd_hg
# program.
# Each trajectory name must be preceded by the TRAJECTORY card.
# Subsequent trajectories are considered together, trajectories may be in ascii
# as well as in .gz format.
# To be able to identify the title line, it must be identical in all files.
#
TRAJECTORY    ./min1.mdcrd
#TRAJECTORY    ./md_traj_short_01.mdcrd
#TRAJECTORY    ./md_traj_short_02.mdcrd
#
#####

MMPBSA

# data in hn
# Full binding energy calc as per the FKBP example with DMSO.

read
  mol pqr com1.pqr.1
  mol pqr lig1.pqr.1
  mol pqr rec1.pqr.1
end

# Complex - Solvated state (Focusing)
elec name complex-solv-course
  mg-manual
  dime 65 65 65
  nlev 4
  grid 1.50 1.50 1.50

```



```

gcent mol 1
mol 1
lpbe
bcfl mdh
ion 1 0.000 2.0
ion -1 0.000 2.0
pdie 2.0
sdie 78.0
chgm spl0
srfin smol
srad 0.0
swin 0.3
temp 300.0
gamma 0.105
calcenergy total
calcforce no
end

```

```

elec name complex-solv-fine
mg-manual
dime 65 65 65
nlev 4
grid 0.225 0.225 0.225
gcent mol 2
mol 1
lpbe
bcfl focus
ion 1 0.000 2.0
ion -1 0.000 2.0
pdie 2.0
sdie 78.0
chgm spl0
srfin smol
srad 0.0
swin 0.3
temp 300.0
gamma 0.105
calcenergy total
calcforce no
end

```

Complex - Reference state (Focusing)

```

elec name complex-ref-course
mg-manual
dime 65 65 65
nlev 4
grid 1.50 1.50 1.50
gcent mol 1
mol 1
lpbe
bcfl mdh
ion 1 0.000 2.0
ion -1 0.000 2.0
pdie 2.0
sdie 2.0
chgm spl0
srfin smol

```

```

    srad 0.0
    swin 0.3
    temp 300.0
    gamma 0.105
    calcenergy total
    calcforce no
end

```

```

elec name complex-ref-fine
  mg-manual
  dime 65 65 65
  nlev 4
  grid 0.225 0.225 0.225
  gcent mol 2
  mol 1
  lpbe
  bcfl focus
  ion 1 0.000 2.0
  ion -1 0.000 2.0
  pdie 2.0
  sdie 2.0
  chgm spl0
  srfm smol
  srad 0.0
  swin 0.3
  temp 300.0
  gamma 0.105
  calcenergy total
  calcforce no
end

```

Ligand - Solvated state (Focusing)

```

elec name ligand-solv-course
  mg-manual
  dime 65 65 65
  nlev 4
  grid 1.50 1.50 1.50
  gcent mol 2
  mol 2
  lpbe
  bcfl mdh
  ion 1 0.000 2.0
  ion -1 0.000 2.0
  pdie 2.0
  sdie 78.0
  chgm spl0
  srfm smol
  srad 0.0
  swin 0.3
  temp 300.0
  gamma 0.105
  calcenergy total
  calcforce no
end

```

```

elec name ligand-solv-fine
  mg-manual

```

```

dime 65 65 65
nlev 4
grid 0.225 0.225 0.225
gcent mol 2
mol 2
lpbe
bcfl focus
ion 1 0.000 2.0
ion -1 0.000 2.0
pdie 2.0
sdie 78.0
chgm spl0
srfin smol
sradi 0.0
swin 0.3
temp 300.0
gamma 0.105
calcenergy total
calcforce no
end

```

Ligand - Reference state (Focusing)

```

elec name ligand-ref-course
mg-manual
dime 65 65 65
nlev 4
grid 1.50 1.50 1.50
gcent mol 2
mol 2
lpbe
bcfl mdh
ion 1 0.000 2.0
ion -1 0.000 2.0
pdie 2.0
sdie 2.0
chgm spl0
srfin smol
sradi 0.0
swin 0.3
temp 300.0
gamma 0.105
calcenergy total
calcforce no
end

```

```

elec name ligand-ref-fine
mg-manual
dime 65 65 65
nlev 4
grid 0.225 0.225 0.225
gcent mol 2
mol 2
lpbe
bcfl focus
ion 1 0.000 2.0
ion -1 0.000 2.0
pdie 2.0

```

```

sdie 2.0
chgm spl0
srfm smol
srad 0.0
swin 0.3
temp 300.0
gamma 0.105
calcenergy total
calcforce no
end

```

Protein - Solvated state (Focusing)

```

elec name protein-solv-course
mg-manual
dime 65 65 65
nlev 4
grid 1.50 1.50 1.50
gcent mol 3
mol 3
lpbe
bcfl mdh
ion 1 0.000 2.0
ion -1 0.000 2.0
pdie 2.0
sdie 78.0
chgm spl0
srfm smol
srad 0.0
swin 0.3
temp 300.0
gamma 0.105
calcenergy total
calcforce no
end

```

```

elec name protein-solv-fine
mg-manual
dime 65 65 65
nlev 4
grid 0.225 0.225 0.225
gcent mol 2
mol 3
lpbe
bcfl focus
ion 1 0.000 2.0
ion -1 0.000 2.0
pdie 2.0
sdie 78.0
chgm spl0
srfm smol
srad 0.0
swin 0.3
temp 300.0
gamma 0.105
calcenergy total
calcforce no
end

```

Protein - Reference state (Focusing)

elec name protein-ref-course

mg-manual

dime 65 65 65

nlev 4

grid 1.50 1.50 1.50

gcent mol 3

mol 3

lpbe

bcfl mdh

ion 1 0.000 2.0

ion -1 0.000 2.0

pdie 2.0

sdie 2.0

chgm spl0

srfm smol

srad 0.0

swin 0.3

temp 300.0

gamma 0.105

calcenergy total

calcforce no

end

elec name protein-ref-fine

mg-manual

dime 65 65 65

nlev 4

grid 0.225 0.225 0.225

gcent mol 2

mol 3

lpbe

bcfl focus

ion 1 0.000 2.0

ion -1 0.000 2.0

pdie 2.0

sdie 2.0

chgm spl0

srfm smol

srad 0.0

swin 0.3

temp 300.0

gamma 0.105

calcenergy total

calcforce no

end

Change in Solvation energy on binding.

print energy complex-solv-fine - complex-ref-fine - ligand-solv-fine + ligand-ref-fine - protein-solv-fine + protein-ref-fine end

quit

Appendix B: Docking and Scoring Results of Alternative Templates

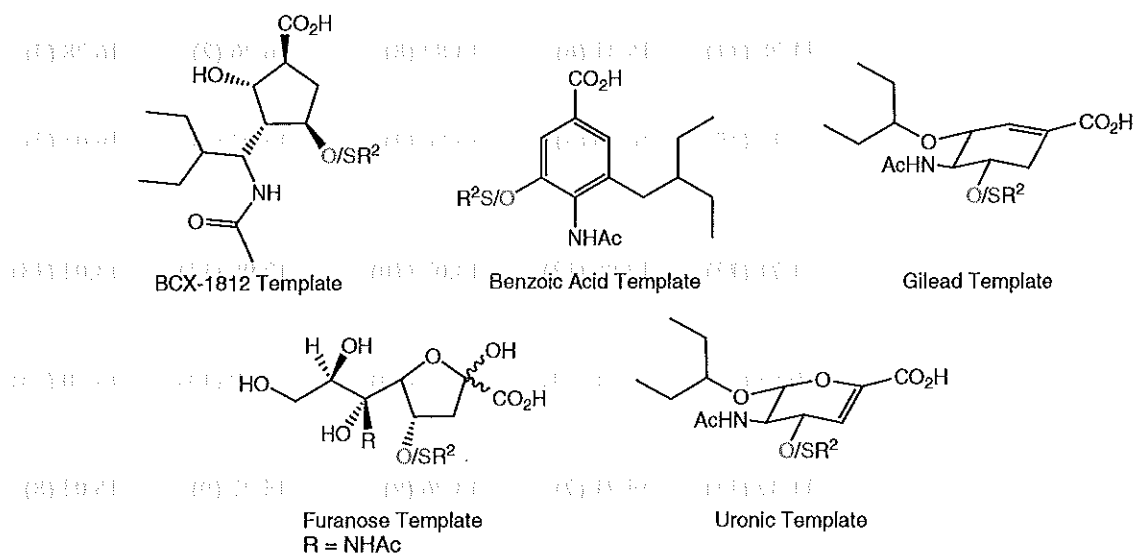


Table B.1: Flexible Autodock score for *O* and *S* linked derivatives.

R ² Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 A	-12.11 (10)	-14.37 (11)	-14.13 (5)	-13.74 (10)	-14.06 (12)
	-12.80 (8)	-14.67 (11)	-13.93 (8)	-13.70 (11)	-14.40 (12)
 B	-15.53 (1)	-16.71 (1)	-15.30 (4)	-15.80 (3)	-16.36 (3)
	-14.77 (2)	-16.91 (1)	-15.45 (5)	-14.46 (8)	-15.86 (5)
 C	-10.47 (14)	-12.18 (14)	-12.20 (13)	-11.44 (14)	-12.56 (14)
	-11.05 (12)	-12.59 (14)	-12.80 (12)	-12.57 (12)	-12.98 (14)
 D	-12.57 (9)	-14.68 (8)	-12.48 (12)	-14.72 (7)	-14.90 (9)
	-13.48 (5)	15.02 (9)	-13.91 (9)	14.78 (6)	-15.21 (8)
 E	-14.14 (4)	-16.15 (3)	-15.68 (1)	-17.08 (1)	-14.89 (10)
	-16.06 (1)	-16.85 (2)	13.29 (11)	-16.60 (1)	-16.06 (3)
 F	-13.43 (6)	-14.47 (10)	-14.11 (6)	-14.63 (8)	-15.19 (6)
	-13.45 (6)	-14.99 (10)	-14.79 (6)	-14.44 (9)	-14.95 (10)

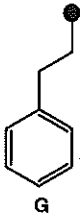
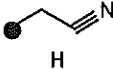
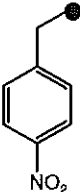
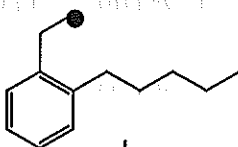
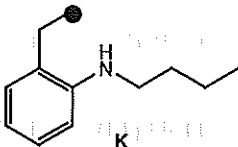
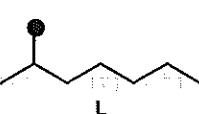
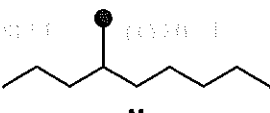
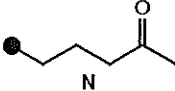
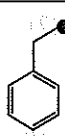
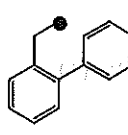

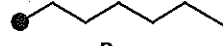
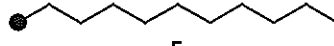
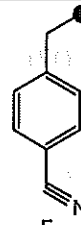
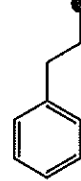

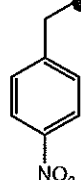
R2 Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 G	-11.78 (11)	-15.31 (6)	-13.87 (8)	-16.26 (2)	-16.28 (4)
	-12.17 (10)	-16.00 (5)	-15.86 (4)	-15.03 (4)	-16.03 (4)
 H	-11.71 (12)	-12.85 (13)	-13.07 (10)	-12.98 (12)	-13.04 (13)
	-10.83 (13)	-13.41 (13)	-13.68 (10)	-13.39 (13)	-13.50 (13)
 I	-11.37 (13)	-14.71 (7)	-13.26 (9)	-14.27 (9)	-15.04 (8)
	-11.62 (11)	-15.29 (8)	-12.72 (13)	-14.44 (9)	-14.77 (11)
 J	-14.25 (3)	-16.12 (4)	-15.49 (2)	-12.84 (13)	-15.59 (5)
	-13.27 (7)	-15.74 (6)	-15.86 (3)	-14.71 (7)	-15.38 (7)
 K	-13.52 (5)	-16.54 (2)	-14.10 (7)	-15.17 (5)	-16.56 (1)
	-13.72 (4)	-16.79 (3)	-16.80 (1)	-15.01 (5)	-16.16 (2)
 L	-13.09 (7)	-14.65 (9)	-10.65 (14)	-14.93 (6)	-15.09 (7)
	-12.52 (9)	-15.45 (7)	-14.23 (7)	-15.04 (3)	-15.61 (6)
 M	-14.56 (2)	-15.98 (5)	-15.45 (3)	-15.57 (4)	-16.52 (2)
	-13.93 (3)	-16.49 (4)	-16.01 (2)	-15.30 (2)	-16.53 (1)
 N	-12.82 (8)	-14.25 (12)	-12.93 (11)	-13.91 (11)	-14.31 (11)
	-12.17 (10)	-14.32 (12)	-14.23 (7)	-14.13 (10)	-14.96 (9)

Table B.2: Static Autodock Scores for *O* and *S* linked derivatives.

R ² Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 A	-12.12 (11)	-13.40 (10)	-12.81 (7)	-13.93 (11)	-13.72 (10)
	-12.46 (9)	-13.87 (11)	-13.04 (9)	-13.44 (11)	-13.67 (12)
 B	-14.62 (1)	-15.51 (2)	-14.05 (3)	-15.20 (5)	-15.43 (3)
	-13.94 (3)	-15.64 (1)	-13.49 (5)	-14.90 (4)	-16.40 (1)
 C	-9.65 (14)	-11.63 (14)	-10.88 (12)	-11.52 (14)	-12.04 (13)
	-10.28 (14)	-11.73 (14)	-11.89 (12)	-12.53 (14)	-12.48 (14)
 D	-13.32 (3)	-13.91 (9)	-11.14 (10)	-14.01 (10)	-14.00 (9)
	-13.23 (5)	-14.03 (10)	-13.02 (10)	-13.91 (9)	-14.52 (10)
 E	-13.71 (2)	-15.94 (1)	-15.36 (1)	-15.73 (4)	-16.11 (1)
	-15.78 (1)	-14.67 (5)	-12.93 (11)	-15.29 (2)	-15.74 (3)
 F	-12.43 (6)	-13.93 (8)	-14.01 (4)	-14.06 (9)	-14.21 (7)
	-13.28 (4)	-14.54 (9)	-13.51 (4)	-14.55 (6)	-14.66 (9)
 G	-12.23 (10)	-14.13 (7)	-10.81 (13)	-14.72 (6)	-14.71 (5)
	-12.83 (6)	-15.19 (5)	-13.25 (8)	-14.68 (5)	-14.70 (5)
 H	-10.47 (13)	-11.94 (13)	-10.99 (11)	-12.14 (6)	-12.30 (12)
	-10.29 (13)	-12.40 (13)	-11.26 (13)	-12.57 (13)	-12.58 (13)
 I	-12.42 (7)	-13.13 (11)	-12.24 (8)	-14.55 (7)	-14.39 (6)
	-14.63 (2)	-14.70 (7)	-9.65 (14)	-13.92 (8)	-14.78 (6)

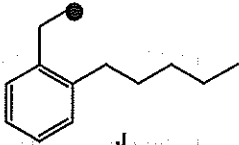
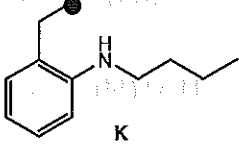
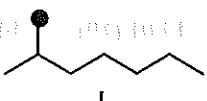
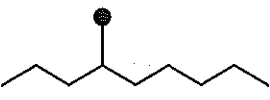
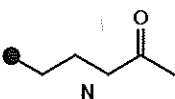

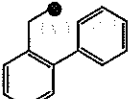

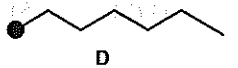
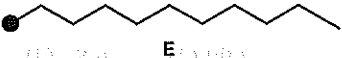
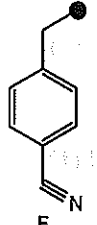
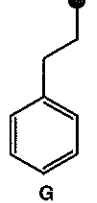
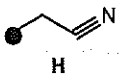
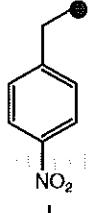
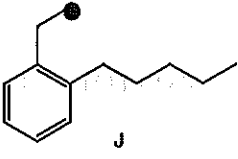
R ² Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 J	-12.75 (5)	-15.07 (4)	-13.58 (5)	-15.85 (2)	-15.10 (4)
	-11.94 (11)	-15.42 (4)	-15.43 (2)	-13.73 (10)	-14.79 (5)
 K	-12.41 (8)	-15.46 (3)	-13.18 (6)	-15.93 (1)	-15.43 (3)
	-12.60 (7)	-15.59 (2)	-16.06 (1)	-15.46 (1)	-15.32 (4)
 L	-12.24 (9)	-14.23 (6)	-9.88 (14)	-14.53 (8)	-14.17 (8)
	-12.00 (10)	-14.96 (6)	-13.46 (6)	-14.12 (7)	-14.77 (7)
 M	-12.84 (4)	-14.97 (5)	-14.17 (2)	-15.75 (3)	-15.49 (2)
	-12.47 (8)	-15.44 (3)	-14.26 (3)	-15.05 (3)	-15.83 (2)
 N	-11.96 (12)	-13.05 (12)	-12.14 (9)	-13.26 (12)	-12.98 (11)
	-11.34 (12)	-13.20 (12)	-13.37 (7)	-13.31 (12)	-14.10 (11)

Table B.3: X-score results for *O* and *S* linked derivatives.

R ² Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 A	6.42 (2)	6.75 (8)	5.81 (7)	6.39 (11)	6.53 (7)
	6.46 (4)	6.89 (9)	6.10 (6)	6.60 (9)	6.53 (8)
 B	7.00 (1)	7.66 (1)	6.91 (1)	7.22 (1)	7.38 (1)
	6.62 (1)	7.74 (1)	6.33 (3)	7.18 (2)	7.54 (1)

R2 Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 C	5.53 (12)	6.11 (11)	5.12 (13)	5.89 (13)	5.81 (13)
	5.70 (12)	6.17 (13)	5.46 (11)	6.12 (13)	5.91 (11)
 D	6.26 (4)	6.47 (9)	5.26 (12)	6.57 (10)	6.30 (11)
	6.20 (6)	6.77 (10)	5.93 (9)	6.47 (11)	6.42 (9)
 E	5.69 (11)	6.87 (6)	6.26 (3)	6.62 (9)	6.52 (8)
	6.28 (5)	6.41 (11)	4.88 (14)	6.85 (6)	6.42 (9)
 F	5.94 (8)	6.92 (4)	6.24 (4)	6.69 (7)	6.44 (10)
	6.49 (3)	7.09 (5)	6.29 (4)	6.96 (5)	6.78 (5)
 G	6.09 (7)	6.90 (5)	5.74 (8)	6.72 (6)	6.81 (4)
	6.07 (8)	7.08 (6)	6.13 (5)	6.99 (3)	6.79 (4)
 H	5.43 (13)	5.91 (12)	5.12 (13)	5.74 (14)	5.61 (14)
	5.36 (14)	6.13 (14)	5.12 (13)	5.75 (14)	5.79 (12)
 I	6.23 (5)	6.75 (8)	5.38 (11)	6.91 (4)	6.75 (5)
	6.61 (2)	7.04 (7)	5.28 (12)	6.98 (4)	6.88 (3)
 J	6.14 (6)	6.87 (6)	6.20 (5)	7.02 (3)	6.65 (6)
	6.02 (9)	7.16 (3)	6.45 (2)	6.56 (10)	6.71 (6)

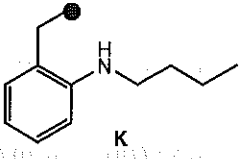
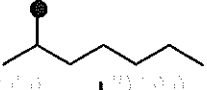
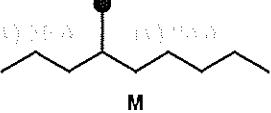
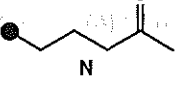
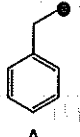
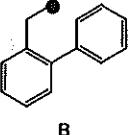

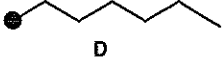
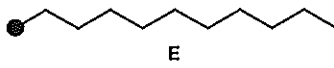
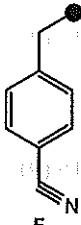
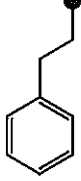
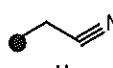
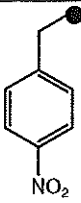
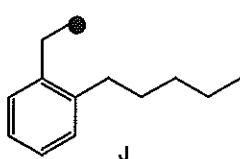
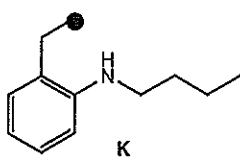
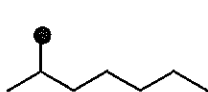
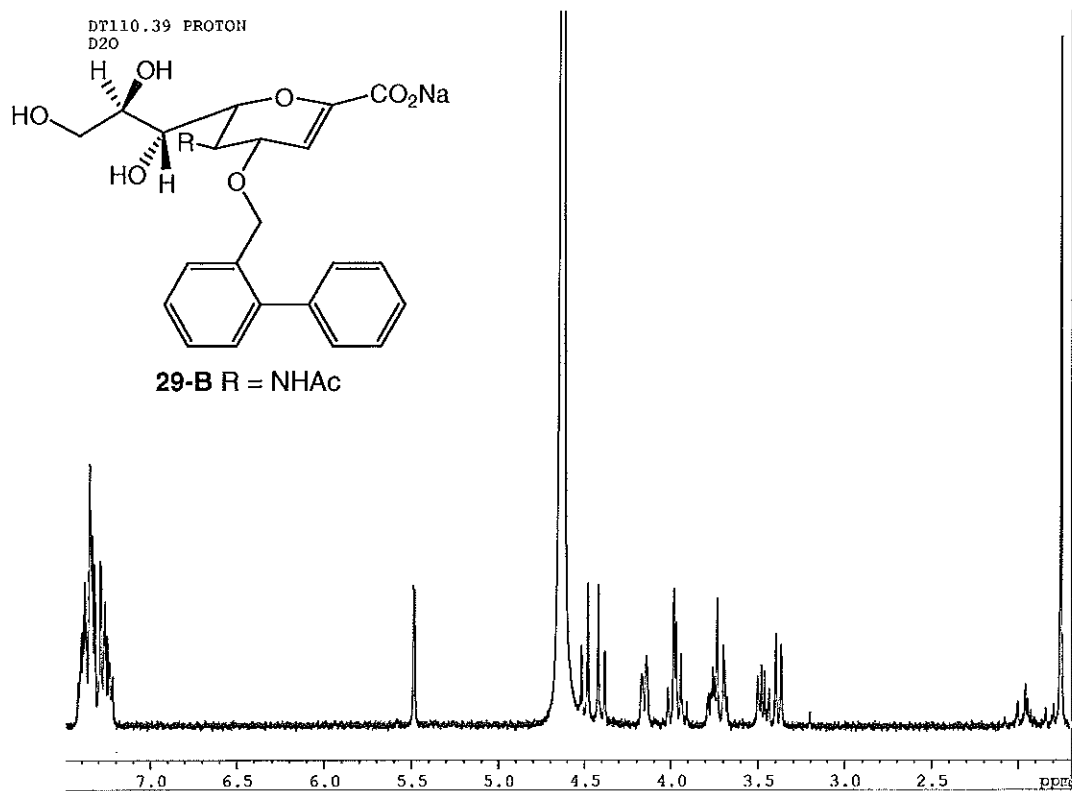
R2 Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 K	5.91 (9)	7.16 (2)	6.44 (2)	6.87 (5)	6.99 (2)
	5.95 (10)	7.20 (2)	6.68 (1)	7.26 (1)	7.03 (2)
 L	5.94 (8)	6.79 (7)	5.42 (9)	6.63 (8)	6.51 (9)
	5.94 (11)	7.00 (8)	6.05 (7)	6.75 (8)	6.66 (7)
 M	6.34 (3)	7.02 (3)	6.09 (6)	7.09 (2)	6.85 (3)
	6.13 (7)	7.13 (4)	5.95 (8)	6.96 (5)	6.88 (3)
 N	5.78 (10)	6.32 (10)	5.41 (10)	6.04 (12)	6.10 (12)
	5.47 (13)	6.26 (12)	5.74 (10)	6.32 (12)	6.09 (10)

Table B.4: MMPBSA Scores for *O* and *S* linked derivatives.

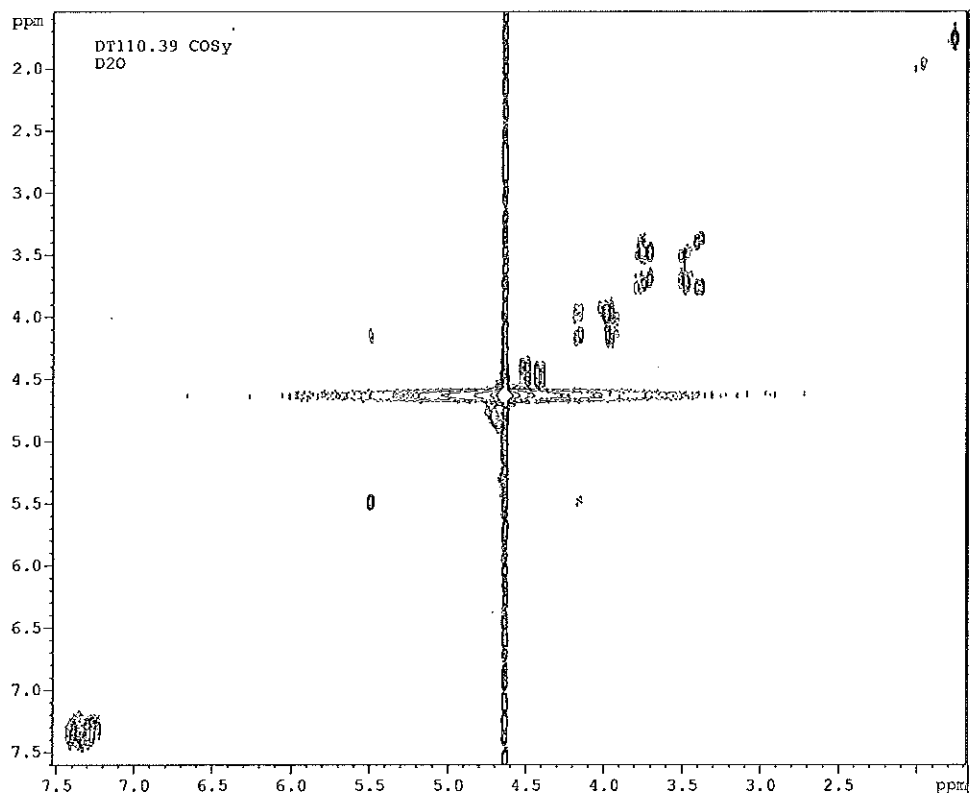
R ² Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 A	-91.23 (11)	-86.29 (12)	-122.97 (8)	-89.48 (9)	-82.55 (12)
	-108.85 (7)	-90.54 (10)	-109.64 (7)	-91.72 (8)	-97.15 (5)
 B	-108.23 (5)	-108.47 (1)	-112.65 (11)	-107.32 (3)	-105.44 (3)
	-109.27 (6)	-92.11 (7)	-117.64 (5)	-86.13 (11)	-94.99 (7)
 C	-90.40 (12)	-90.73 (7)	-123.70 (7)	-88.99 (11)	-91.73 (9)
	-88.66 (14)	-90.08 (11)	-109.87 (6)	-85.91 (12)	-89.19 (11)
 D	-101.74 (7)	-85.05 (13)	-124.16 (6)	-87.95 (12)	-95.59 (7)
	-101.45 (11)	-95.51 (5)	-93.48 (13)	-80.00 (14)	-95.20 (6)

R ² Substituent	BCX (kcal/mol)	BENZOIC (kcal/mol)	FURANOSE (kcal/mol)	GILEAD (kcal/mol)	URONIC (kcal/mol)
 E	-119.24 (1)	-78.19 (14)	-102.47 (12)	-99.80 (5)	-84.46 (11)
	-117.12 (3)	-104.24 (2)	-133.65 (2)	-97.79 (5)	-73.27 (14)
 F	-93.81 (9)	-90.75 (6)	-101.96 (13)	-106.02 (4)	-84.83 (10)
	-104.25 (10)	-79.91 (12)	-101.71 (11)	-89.69 (10)	-89.31 (10)
 G	-119.00 (2)	-100.22 (3)	-126.06 (4)	-99.50 (6)	-104.35 (4)
	-106.39 (8)	-104.87 (1)	-137.81 (1)	-100.64 (4)	-92.97 (8)
 H	-111.25 (3)	-89.18 (10)	-100.27 (14)	-81.09 (14)	-95.31 (8)
	-95.89 (13)	-92.02 (8)	-102.01 (10)	-85.61 (13)	-81.12 (12)
 I	-103.35 (6)	-90.35 (8)	-115.31 (10)	-89.40 (10)	-80.86 (13)
	-190.38 (1)	-74.78 (13)	-97.61 (12)	-101.34 (3)	-101.23 (2)
 J	-109.57 (4)	-87.69 (11)	-134.75 (2)	-81.78 (13)	-112.87 (1)
	-105.79 (9)	-97.39 (3)	-109.03 (8)	-93.86 (7)	-107.19 (1)
 K	-101.16 (8)	-103.17 (2)	-130.94 (3)	-109.74 (2)	-106.11 (2)
	-100.36 (12)	-95.40 (6)	-103.75 (9)	-102.10 (2)	-90.37 (9)
 L	-93.75 (10)	-98.36 (4)	-117.67 (9)	-91.67 (8)	-69.61 (14)
	-111.36 (4)	-97.27 (4)	-90.54 (14)	-95.48 (6)	-99.49 (3)

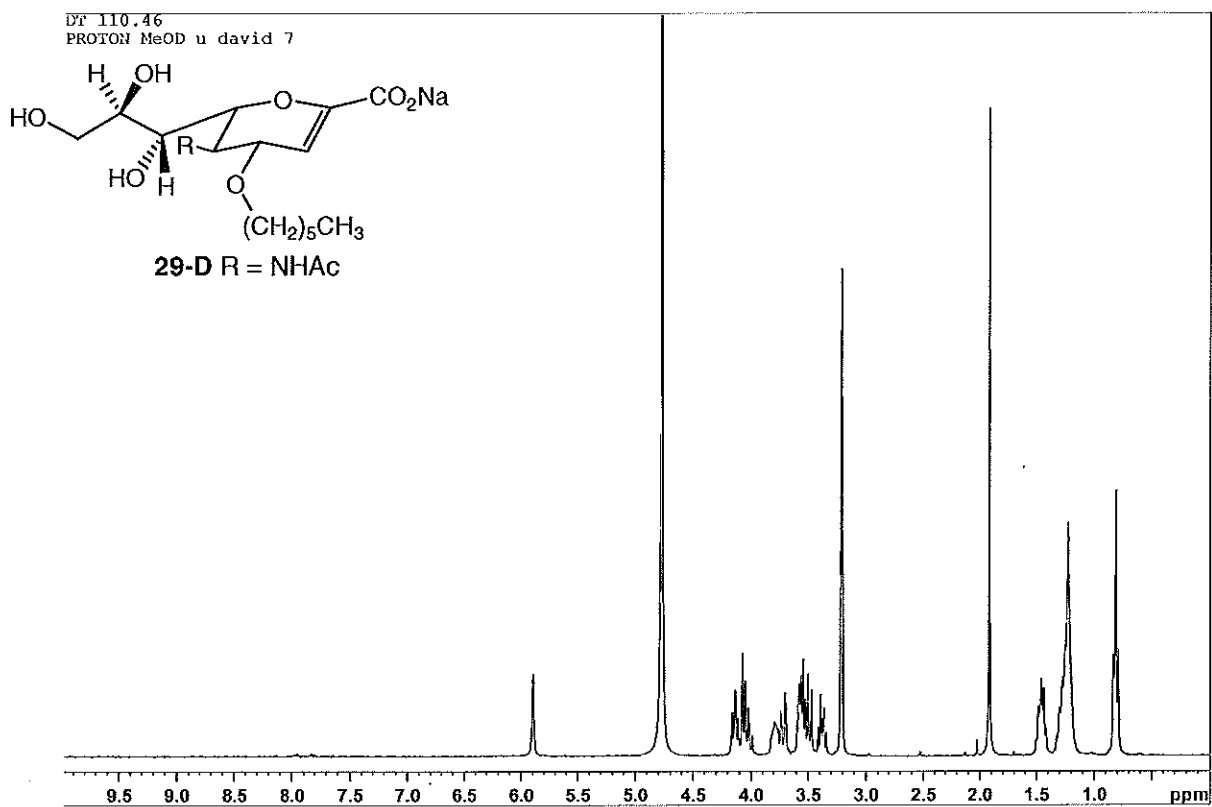
Appendix C: Selected ^1H NMR and COSY spectra



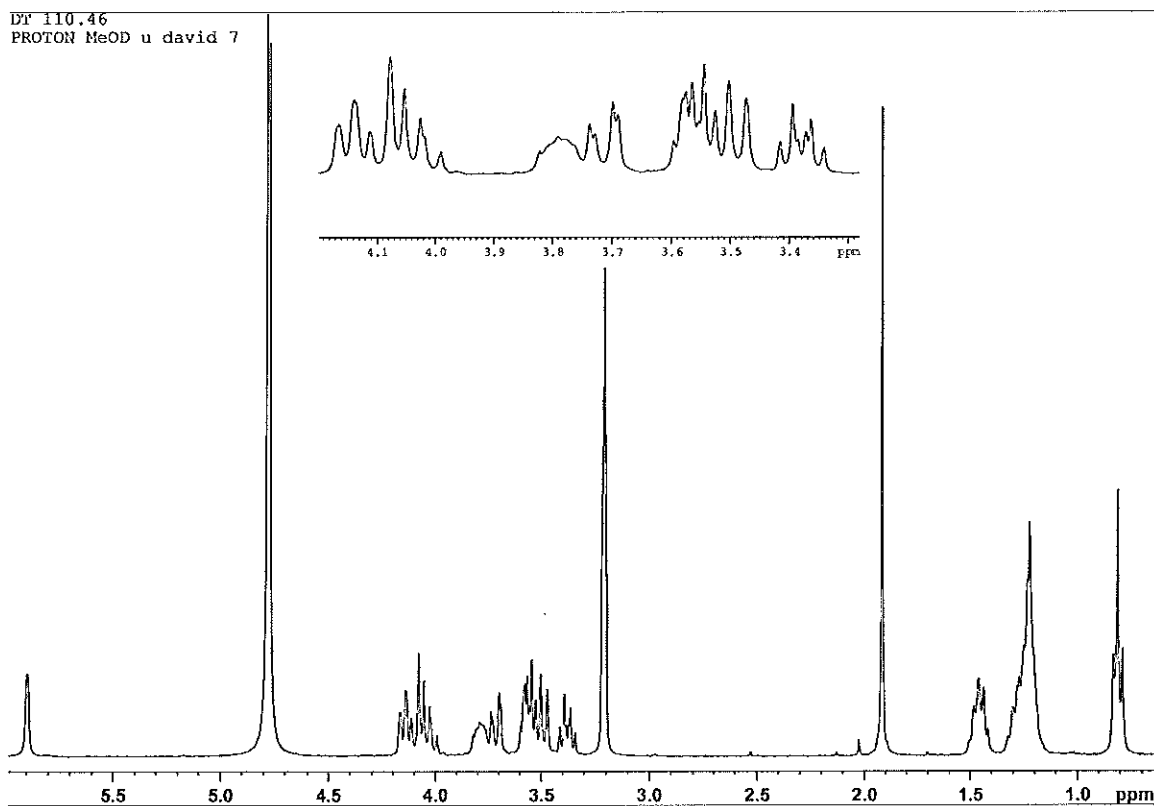
^1H NMR of compound **29-B** (300MHz, D_2O)



^1H - ^1H COSY of compound **29-B** (300MHz, D_2O)

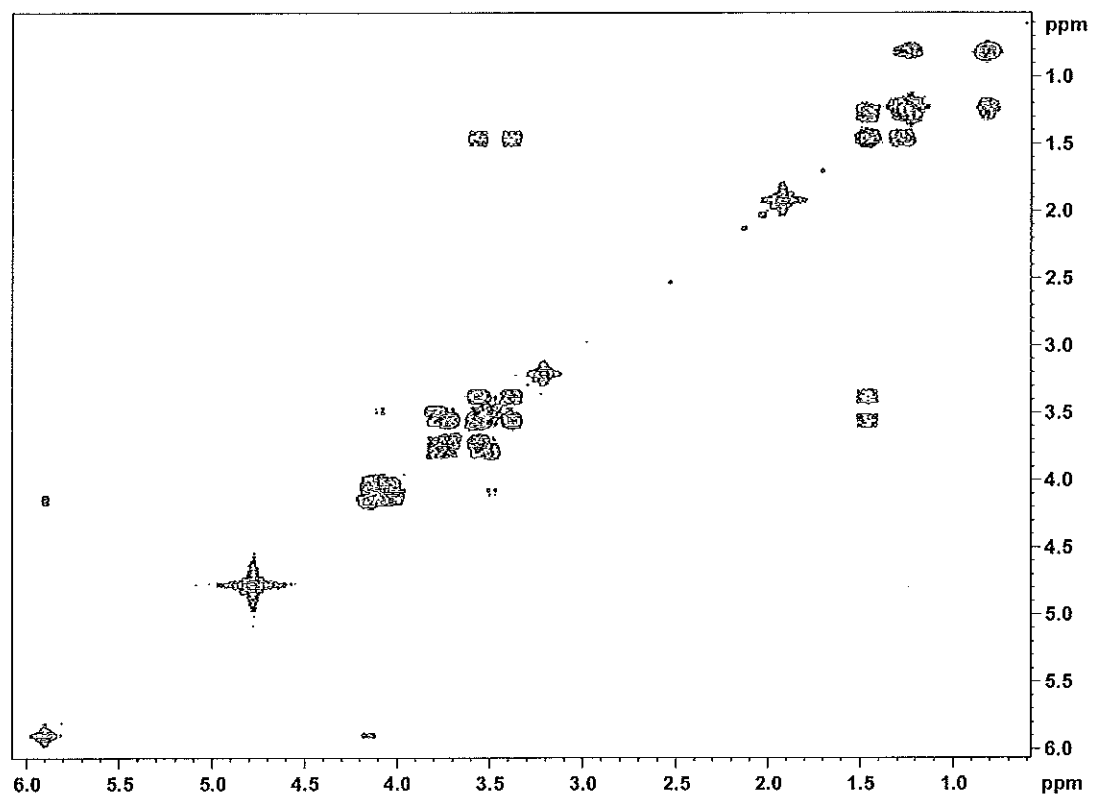


^1H NMR of compound **29-D** (300MHz, CD_3OD)

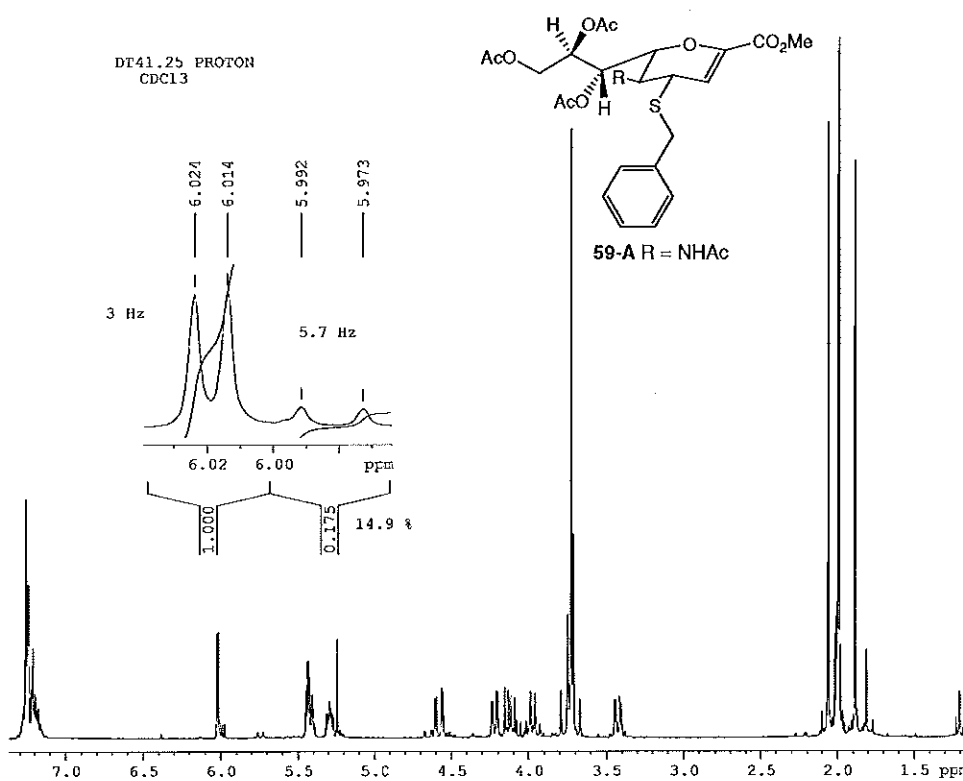


Expansion of ^1H NMR of compound **29-D** (300MHz, D_2O)

DT 110.46
PROTON MeOD u david 7

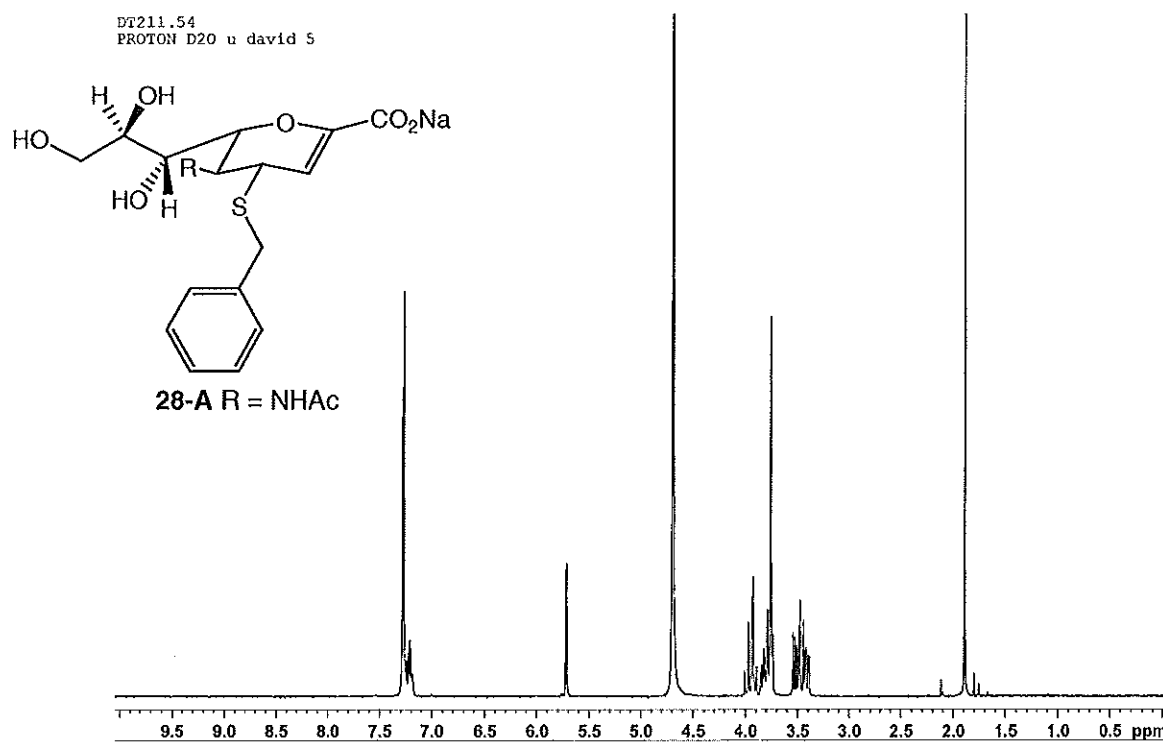


^1H - ^1H COSY of compound **29-D** (300MHz, D_2O)



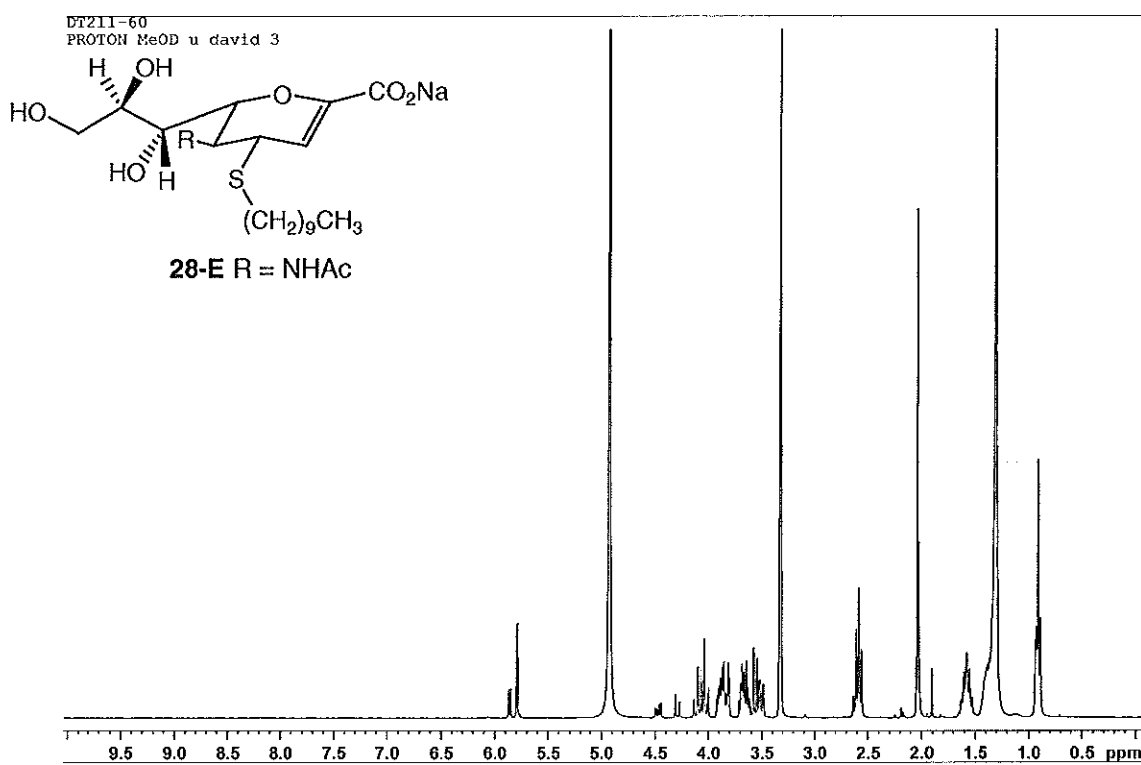
^1H NMR of compound **59-A** (300MHz, CDCl_3)

DT211-54
PROTON D2O u david 5



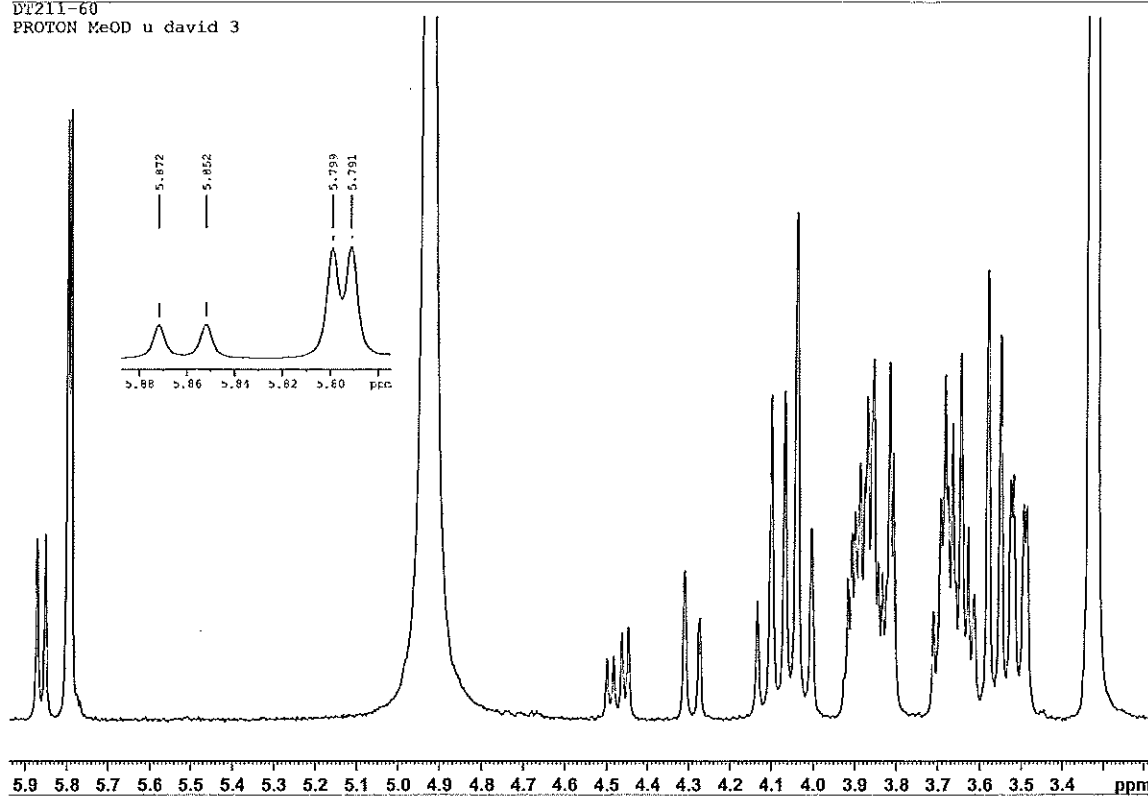
^1H NMR of compound **28-A** (300MHz, D_2O)

DT211-60
PROTON MeOD u david 3



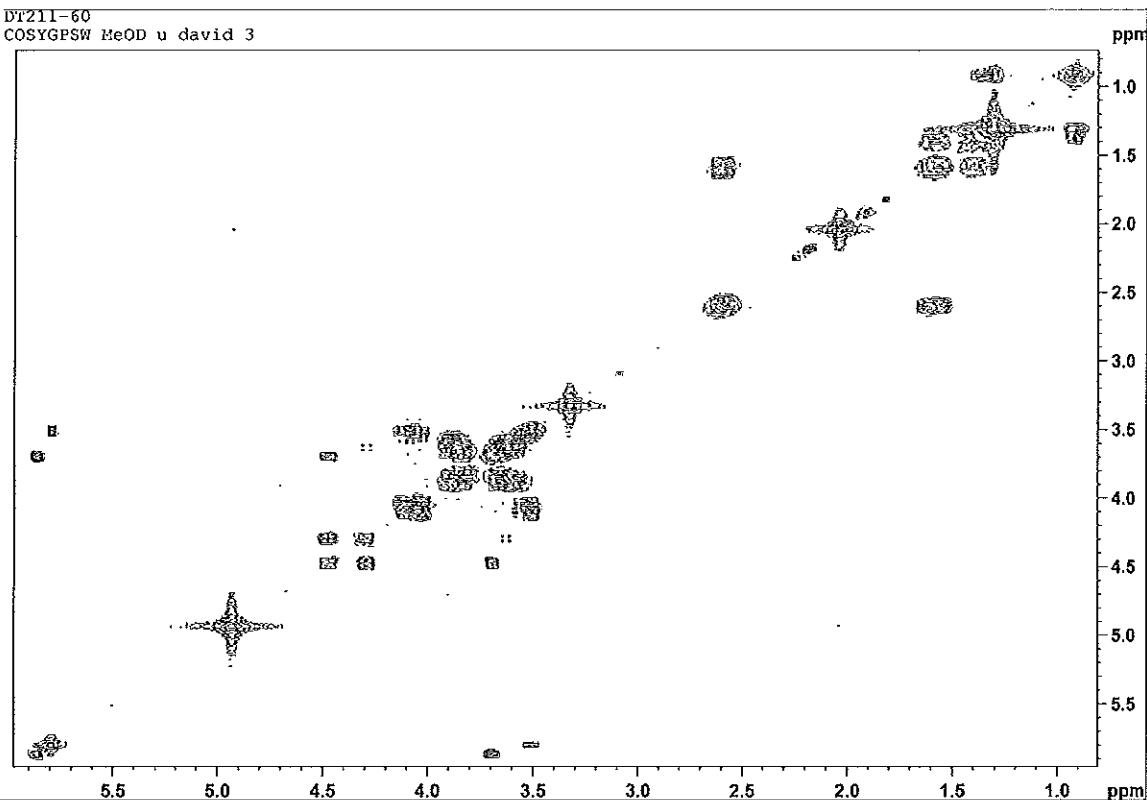
^1H NMR of compound **28-E** (300MHz, D_2O)

DT211-60
 PROTON MeOD u david 3



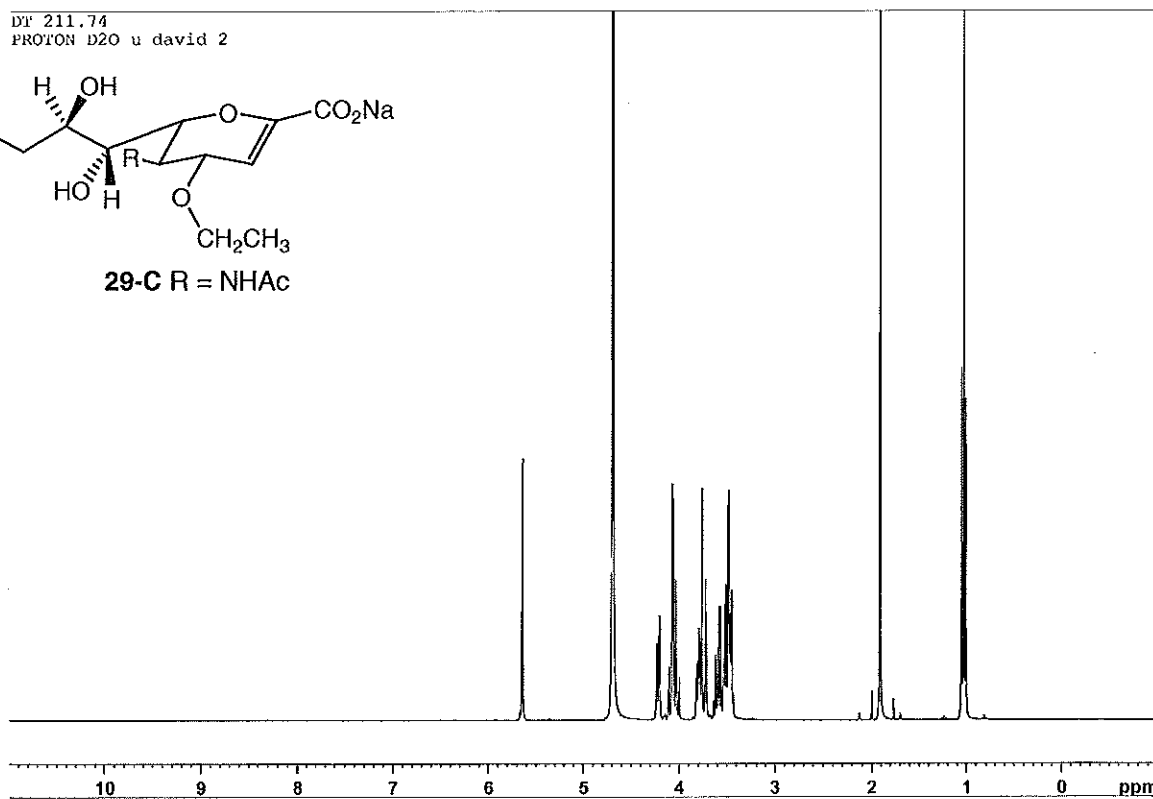
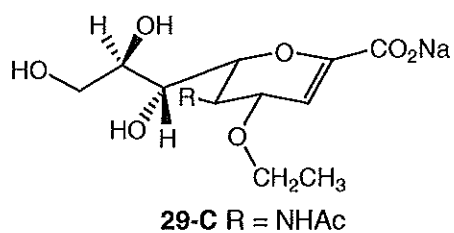
Expansion of ^1H NMR of compound **28-E** (300MHz, D_2O)

DT211-60
 COSYGPSW MeOD u david 3



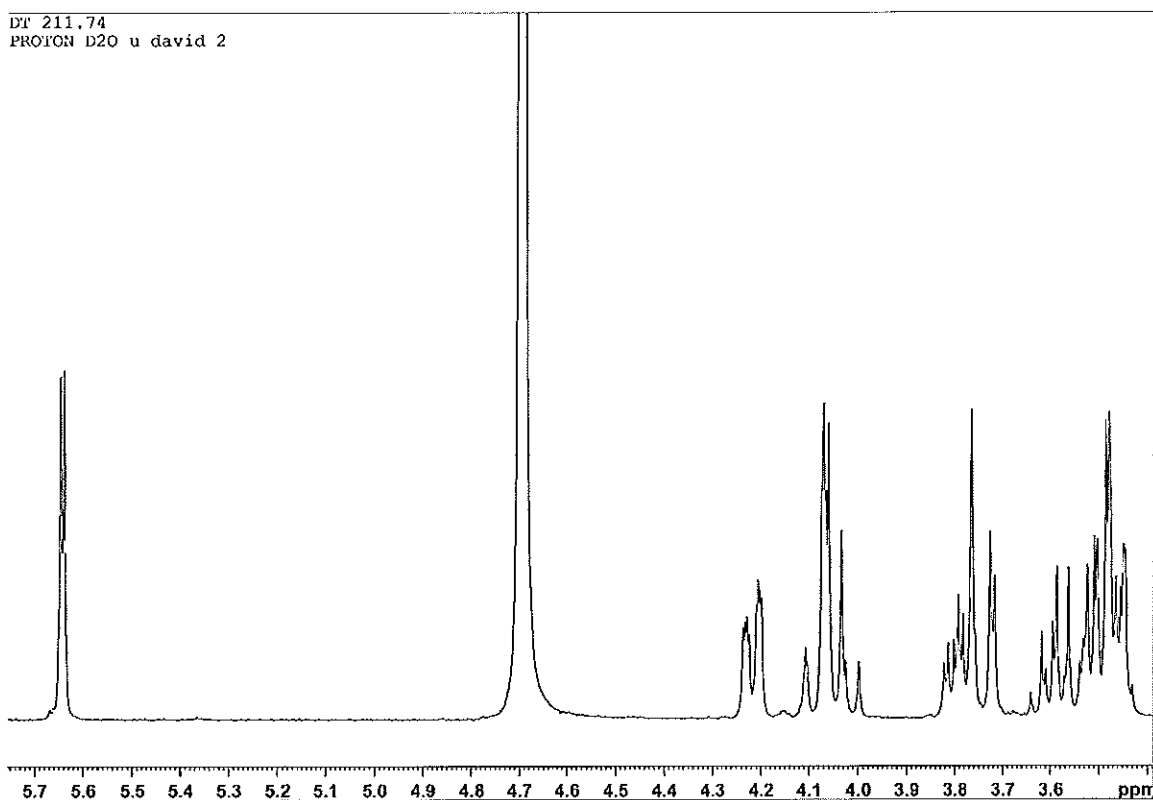
^1H - ^1H COSY of compound **28-E** (300MHz, D_2O)

DT 211.74
PROTON D2O u david 2



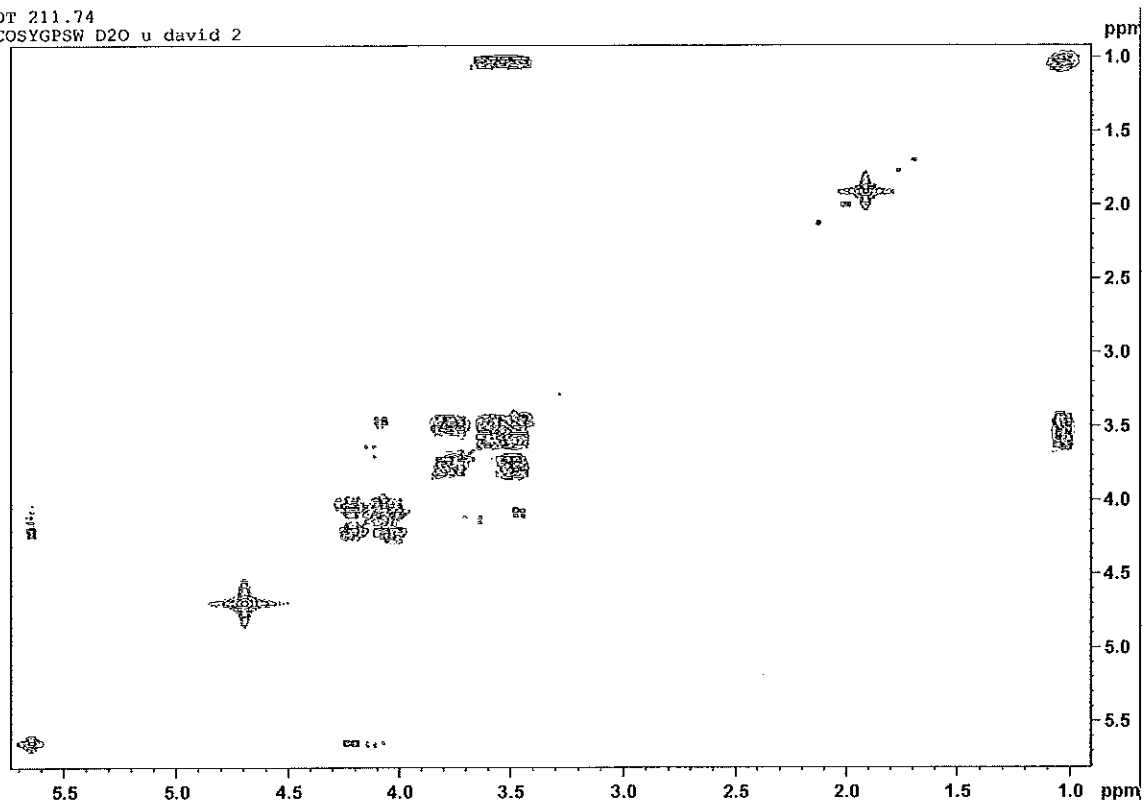
^1H NMR of compound **29-C** (300MHz, D_2O)

DT 211.74
PROTON D2O u david 2



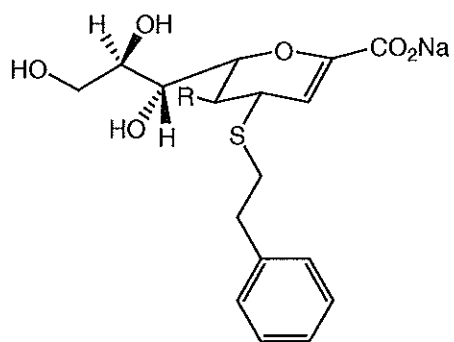
Expansion of ^1H NMR of compound **29-C** (300MHz, D_2O)

DT 211.74
COSYGPBW D2O u david 2

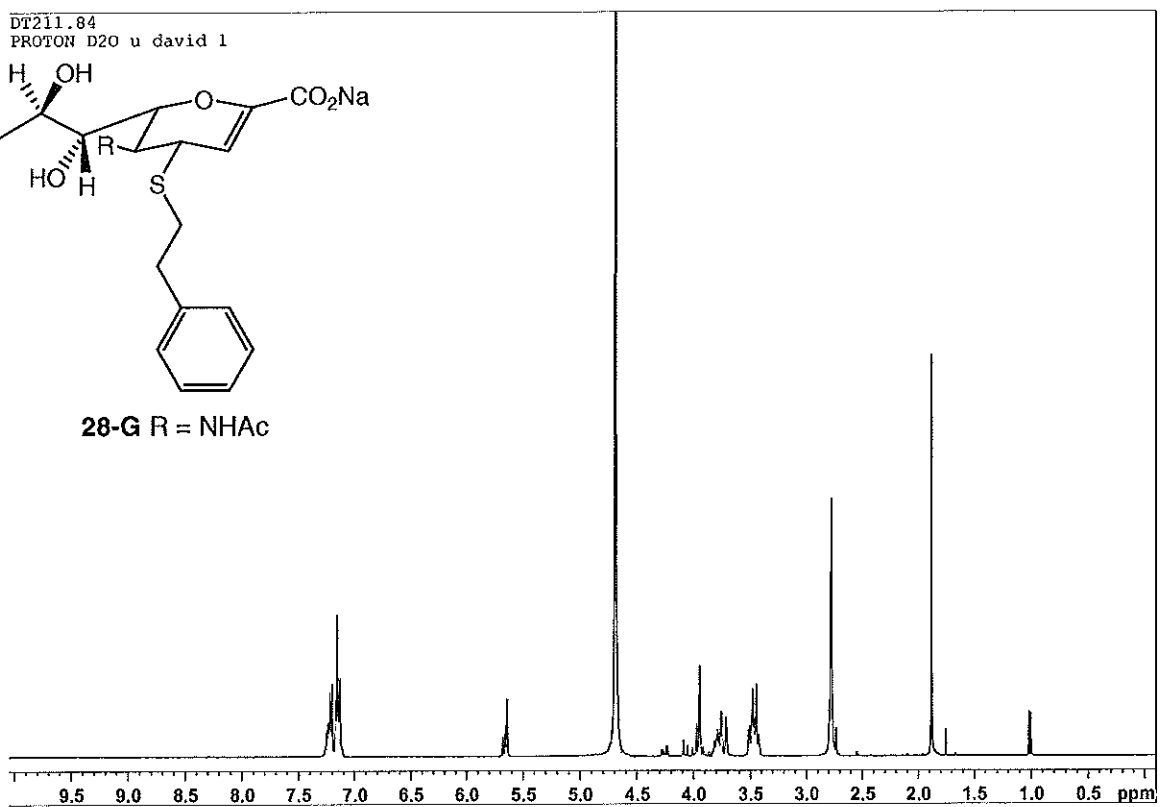


^1H - ^1H COSY of compound **29-C** (300MHz, D_2O)

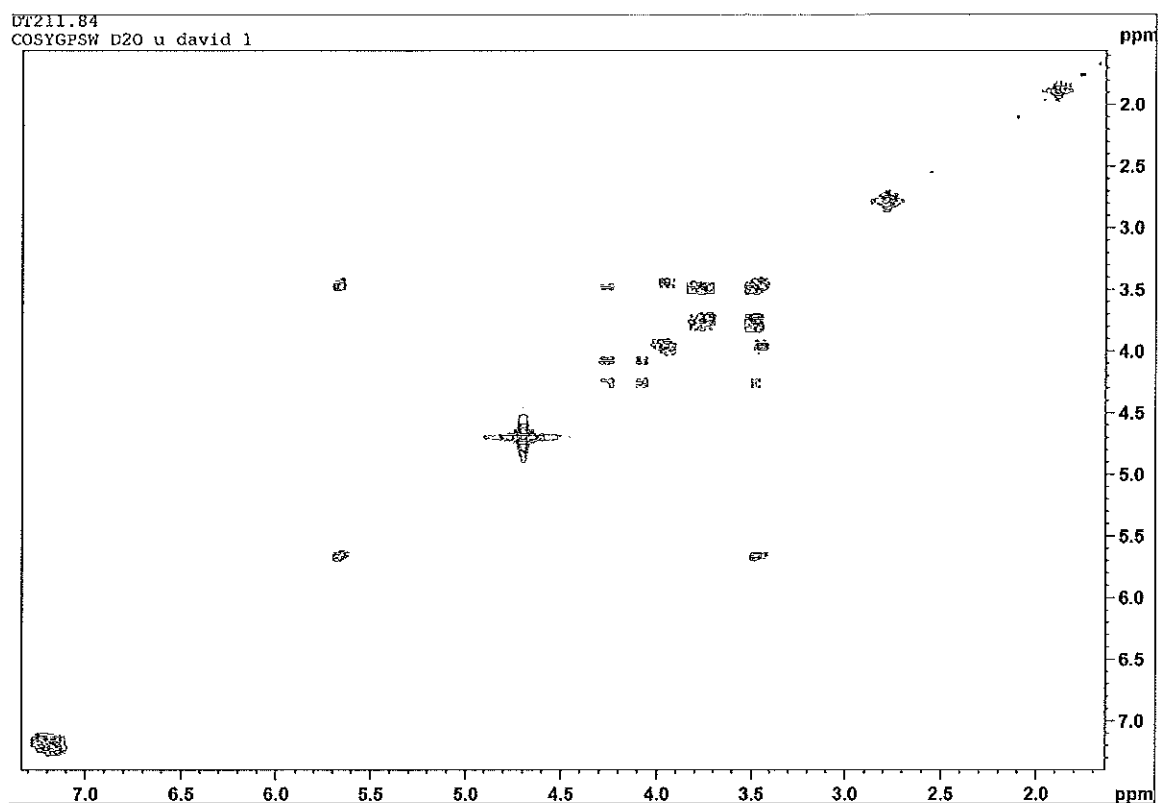
DT211.84
PROTON D2O u david 1



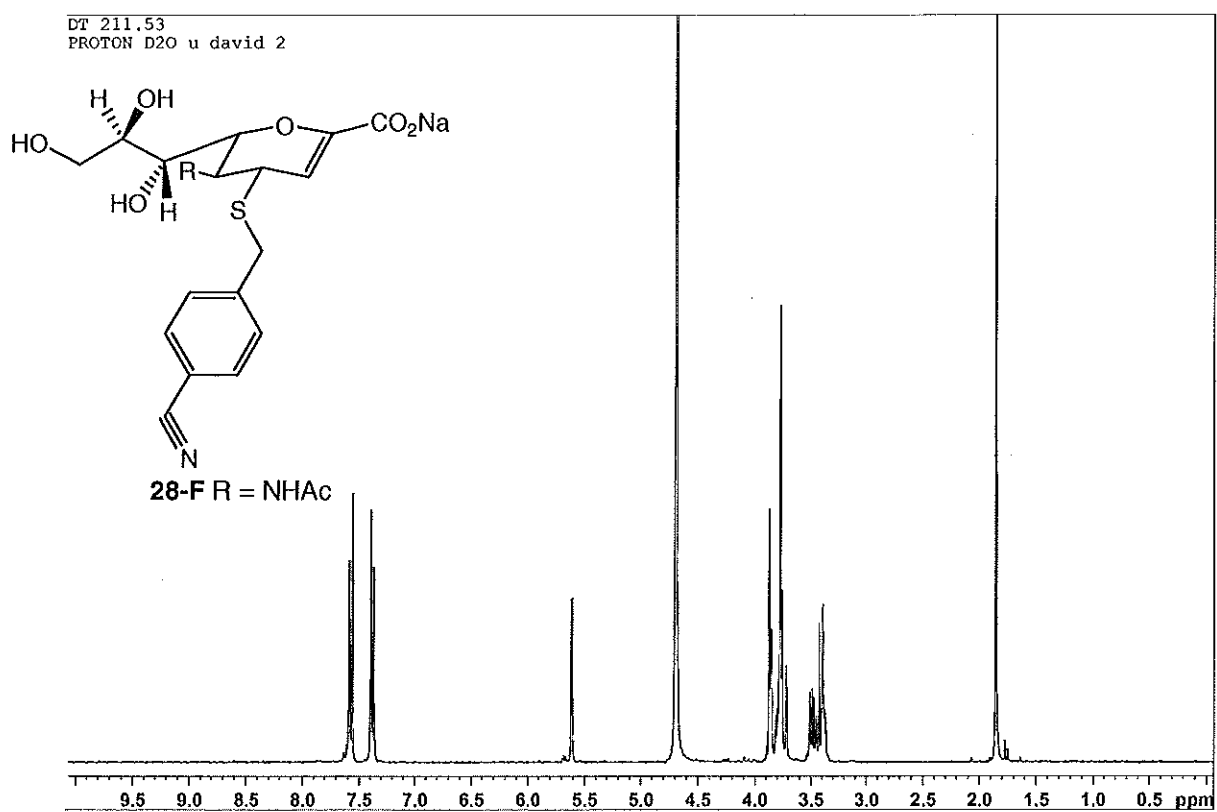
28-G R = NHAc



^1H NMR of compound **28-G** (300MHz, D_2O)

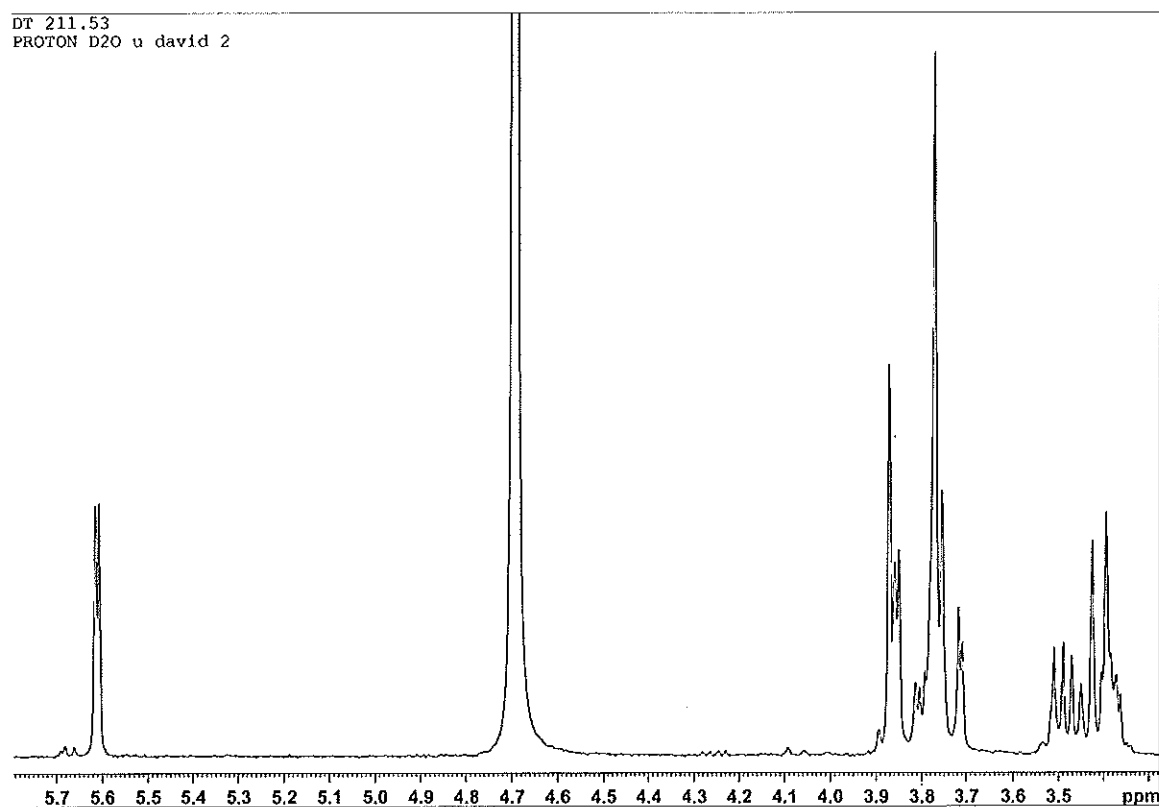


^1H - ^1H COSY of compound **28-G** (300MHz, D_2O)



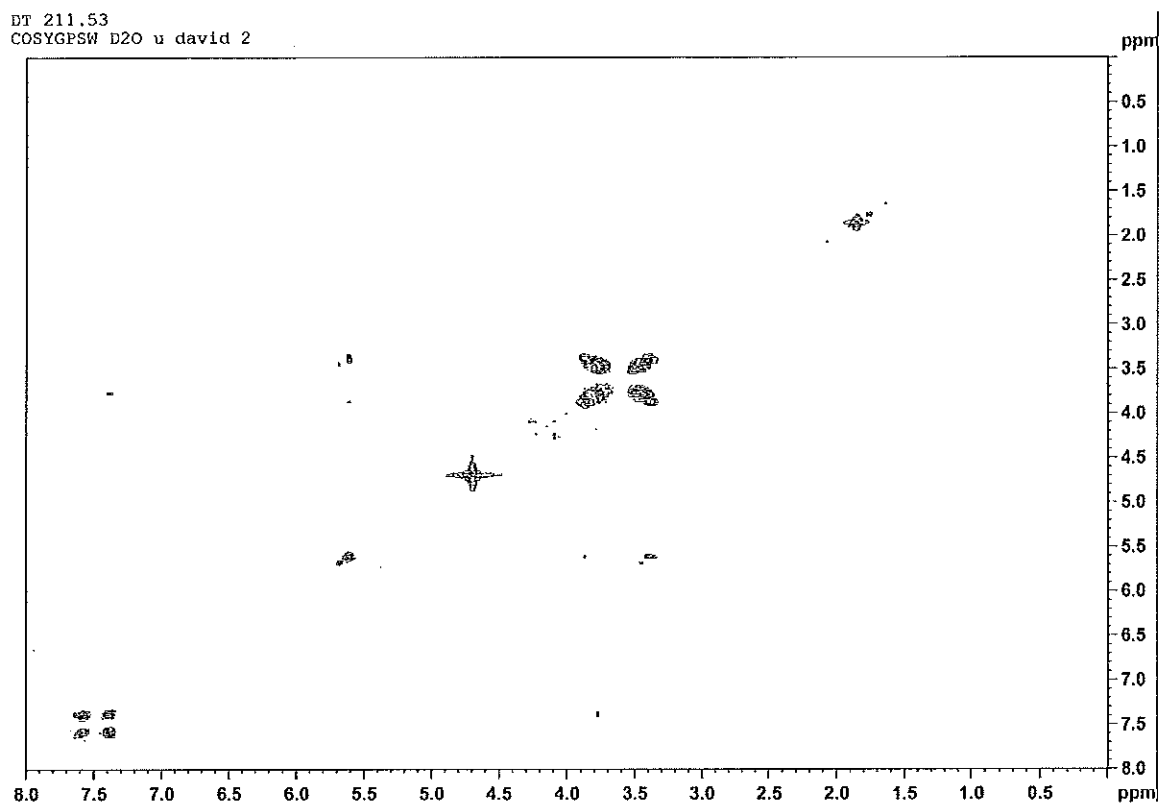
^1H NMR of compound **28-F** (300MHz, D_2O)

DT 211.53
 PROTON D2O u david 2



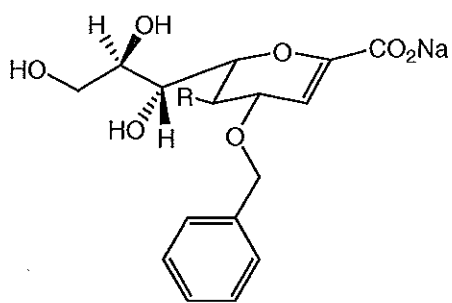
Expansion of ^1H NMR of compound **28-F** (300MHz, D_2O)

DT 211.53
 COSYGPSW D2O u david 2

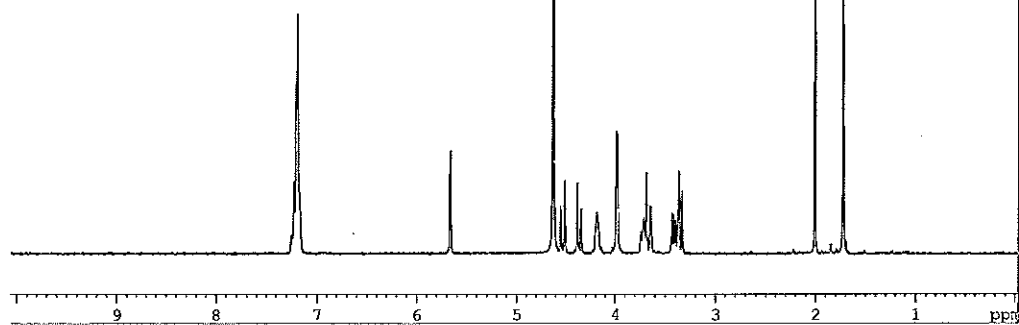


^1H - ^1H COSY of compound **28-F** (300MHz, D_2O)

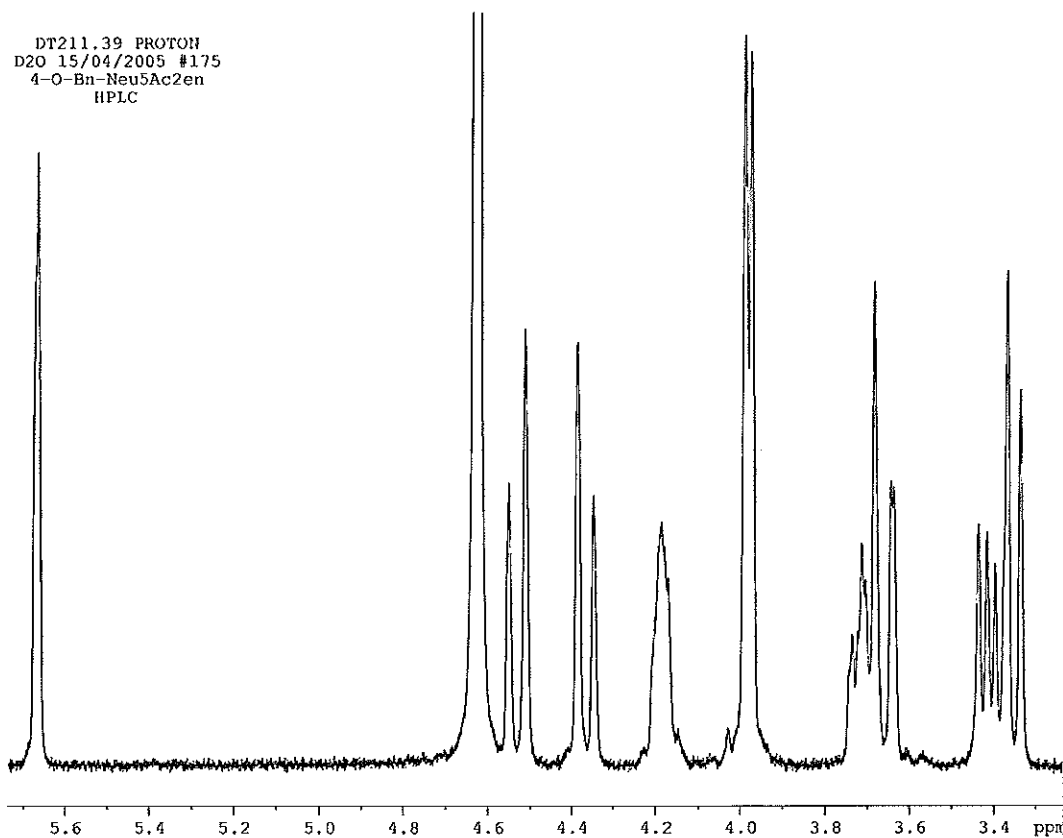
DT211.39 PROTON
D2O 15/04/2005 #175
4-O-Bn-Neu5Ac2en
HPLC



29-A R = NHAc



^1H NMR of compound **29-A** (300MHz, D_2O)



Expansion of ^1H NMR of compound **29-A** (300MHz, D_2O)

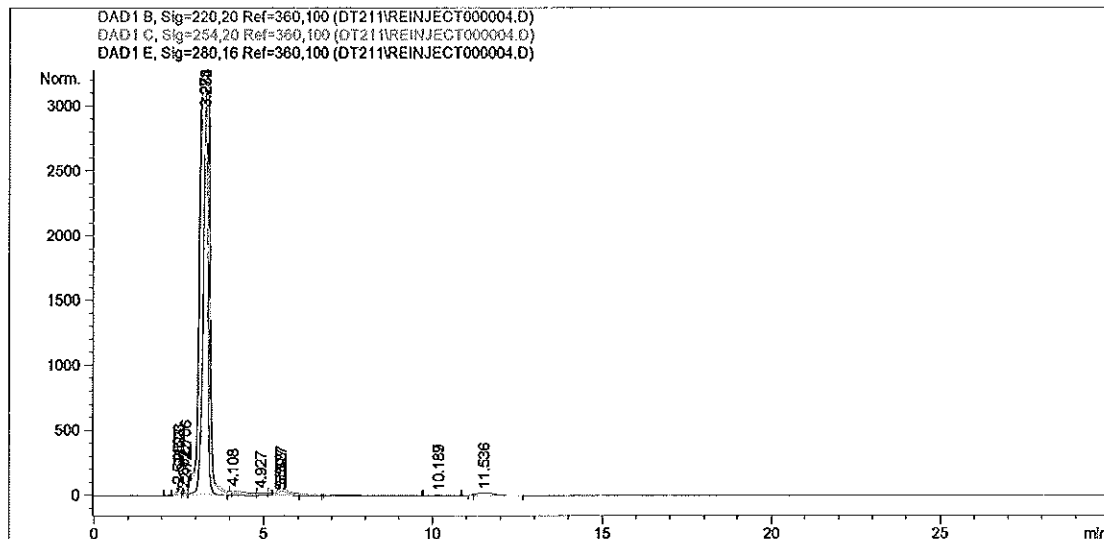
Appendix D: HPLC data of C-4 thioether Neu5Ac2en derivatives

The codes used for the sample names in the following HPLC chromatograms are as follows: DT211.53 – 4-*S*-*p*-(cyano)benzyl-Neu5Ac2en; DT211.54 – 4-*S*-benzyl-Neu5Ac2en; DT211.55 – 4-*S*-(2-phenyl)benzyl-Neu5Ac2en; DT211.60 – 4-*S*-decyl-Neu5Ac2en; DT211.74 – 4-*S*-ethyl-Neu5Ac2en; DT211.82 – 4-*S*-hexyl-Neu5Ac2en; DT211.84 – 4-*S*-(2-phenyl)ethyl-Neu5Ac2en. The sample name in each scanned picture is found on the second from the top line.

=====

Injection Date	: 5/30/2005 5:34:22 PM	Seq. Line	: 4
Sample Name	: DT211.82	Location	: Vial 23
Acq. Operator	: David	Inj	: 1
Acq. Instrument	: Glycomics-HPLC1	Inj Volume	: 20 µl
Method	: C:\HPCHEM\1\METHODS\5ACN.M		
Last changed	: 5/30/2005 3:31:21 PM by David		

=====



Area Percent Report

Sorted By : Signal
Multiplier : 1.0000
Dilution : 1.0000
Use Multiplier & Dilution Factor with ISTDs

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.523	BV	0.0945	1574.16101	228.27893	2.0526
2	2.706	VV	0.1115	1887.83167	251.81001	2.4616
3	3.274	VV	0.2800	7.01236e4	3122.21851	91.4349
4	4.108	VV	0.4279	1139.69751	33.58434	1.4861
5	4.927	VV	0.2579	278.83542	14.79420	0.3636
6	5.527	VB	0.4372	1608.01233	52.22953	2.0967
7	10.187	BB	0.4842	80.27020	2.81880	0.1047

Totals : 7.66924e4 3705.73431

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.523	BV	0.0800	997.70728	176.27391	1.6334
2	2.722	VV	0.0776	362.22836	66.44983	0.5930
3	3.278	VB	0.3312	5.91792e4	2910.39502	96.8832
4	5.527	BB	0.3402	457.89520	20.37652	0.7496
5	10.189	BB	0.5119	85.98646	2.56008	0.1408

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000004.D
Sample Name: DT211.82

```
=====
Injection Date   : 5/30/2005 5:34:22 PM      Seq. Line :    4
Sample Name      : DT211.82                  Location  : Vial 23
Acq. Operator    : David                      Inj       :    1
Acq. Instrument  : Glycomics-HPLC1            Inj Volume: 20 µl
Method           : C:\HPCHEM\1\METHODS\5ACN.M
Last changed     : 5/30/2005 3:31:21 PM by David
=====
```

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
Totals :				6.10830e4	3176.05536	

Signal 3: DAD1 E, Sig=280,16 Ref=360,100

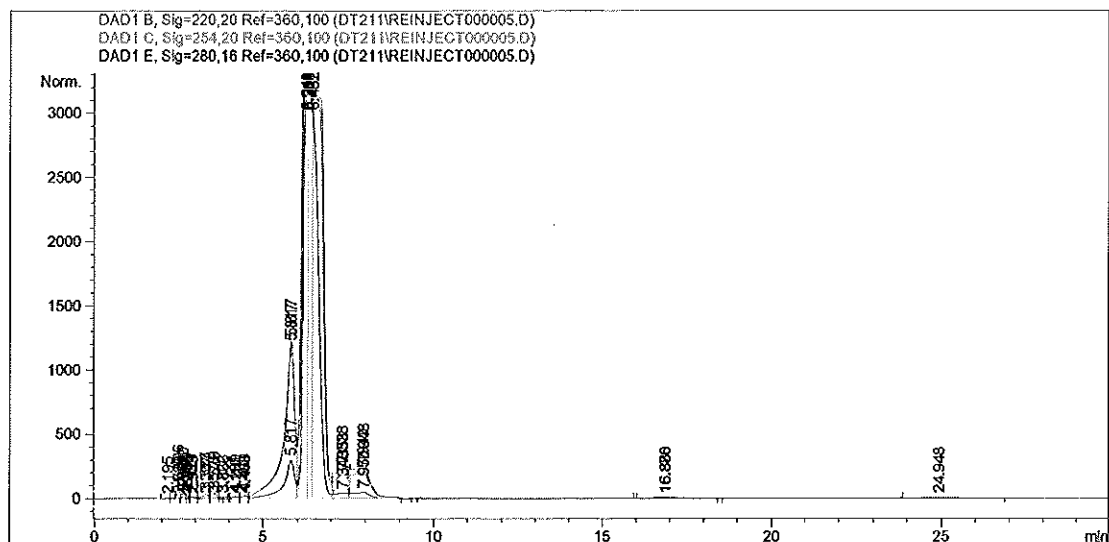
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.511	BV	0.0807	125.59361	23.37514	0.8681
2	2.662	VV	0.1123	82.24754	9.57031	0.5685
3	3.282	VB	0.1480	1.38430e4	1416.07654	95.6789
4	5.518	BB	0.3046	28.41275	1.43766	0.1964
5	11.536	BB	0.6506	388.92636	9.40309	2.6881

Totals : 1.44682e4 1459.86275

*** End of Report ***

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000005.D
Sample Name: DT211.53

```
=====
Injection Date   : 5/30/2005 6:06:02 PM      Seq. Line   :    5
Sample Name     : DT211.53                  Location    : Vial 24
Acq. Operator   : David                     Inj         :    1
Acq. Instrument : Glycomics-HPLC1           Inj Volume  : 20 µl
Method          : C:\HPCHEM\1\METHODS\5ACN.M
Last changed    : 5/30/2005 3:31:21 PM by David
=====
```



Area Percent Report

```
=====
Sorted By       :      Signal
Multiplier      :      1.0000
Dilution        :      1.0000
Use Multiplier & Dilution Factor with ISTDs
=====
```

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.195	BV	0.1102	28.17070	3.41555	0.0153
2	2.506	VV	0.0677	429.93912	93.47238	0.2335
3	2.650	VV	0.0896	464.80905	75.68451	0.2524
4	2.756	VV	0.0703	127.02332	25.43535	0.0690
5	2.929	VV	0.1771	183.27060	14.73258	0.0995
6	3.323	VV	0.1807	394.20981	29.73436	0.2141
7	3.579	VV	0.1665	424.21616	37.91571	0.2304
8	3.866	VV	0.2401	217.98575	12.45908	0.1184
9	4.189	VV	0.2389	216.07138	13.93819	0.1173
10	4.479	VV	0.2020	268.15753	20.32139	0.1456
11	5.817	VV	0.2919	2.57038e4	1190.22729	13.9593
12	6.248	VV	0.1651	4.24941e4	3155.56104	23.0779
13	6.315	VV	0.0835	2.04145e4	3154.84790	11.0868
14	6.452	VV	0.2897	7.55661e4	3151.80688	41.0388
15	7.338	VV	0.4200	6644.32568	234.91617	3.6084
16	7.948	VB	0.5064	9374.96680	252.93994	5.0914
17	16.880	BB	0.9939	675.42480	10.59285	0.3668
18	24.948	BB	1.0876	506.36978	5.48220	0.2750

Totals : 1.84133e5 1.14835e4

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000005.D
Sample Name: DT211.53

```
=====
Injection Date   : 5/30/2005 6:06:02 PM      Seq. Line :    5
Sample Name      : DT211.53                  Location  : Vial 24
Acq. Operator    : David                      Inj       :    1
Acq. Instrument  : Glycomics-HPLC1            Inj Volume: 20 µl
Method           : C:\HPCHEM\1\METHODS\5ACN.M
Last changed     : 5/30/2005 3:31:21 PM by David
=====
```

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.520	BV	0.0574	17.37200	4.46306	0.0106
2	2.652	VV	0.0741	256.73694	53.40440	0.1562
3	2.758	VV	0.0590	48.72321	12.09627	0.0296
4	2.948	VV	0.1439	95.55136	9.62146	0.0581
5	3.317	VV	0.1627	339.95483	29.01978	0.2069
6	3.576	VV	0.1593	379.78568	35.94292	0.2311
7	3.822	VV	0.2386	94.14611	6.08324	0.0573
8	4.201	VV	0.2094	154.11154	11.27696	0.0938
9	4.466	VV	0.1820	218.11807	17.91978	0.1327
10	5.817	VV	0.2931	2.52656e4	1164.40686	15.3745
11	6.329	VV	0.2679	6.53786e4	3004.42236	39.7839
12	6.452	VV	0.2481	5.95401e4	2992.68140	36.2311
13	7.358	VV	0.3614	4512.30762	175.85524	2.7458
14	7.942	VB	0.5357	7658.18896	193.51860	4.6601
15	16.876	BB	0.9125	375.09821	5.90958	0.2283

Totals : 1.64334e5 7716.62194

Signal 3: DAD1 E, Sig=280,16 Ref=360,100

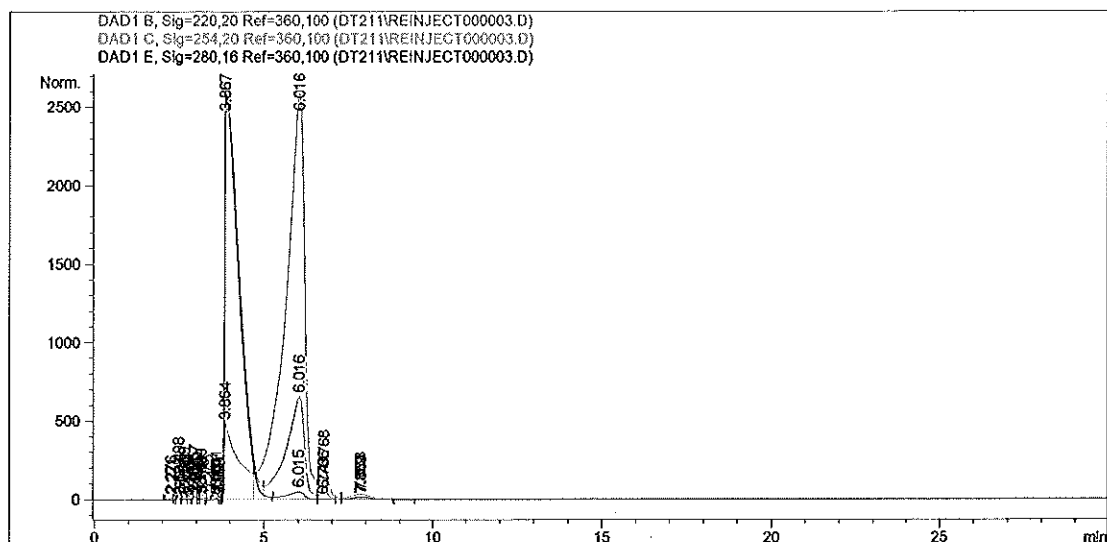
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.520	BV	0.0581	12.58144	3.18310	0.0145
2	2.641	VV	0.0731	65.22346	12.88001	0.0754
3	2.763	VV	0.0595	13.12652	3.22463	0.0152
4	2.933	VV	0.1434	13.85690	1.50582	0.0160
5	3.337	VV	0.1507	32.21320	2.96580	0.0372
6	3.578	VB	0.1580	42.05927	4.02097	0.0486
7	4.197	BV	0.1797	20.24778	1.85027	0.0234
8	4.463	VV	0.1576	31.66262	3.03788	0.0366
9	5.817	VV	0.2903	5365.89844	250.06776	6.2024
10	6.281	VV	0.4898	7.86032e4	2652.09619	90.8573
11	7.348	VV	0.3541	859.69653	34.33578	0.9937
12	7.950	VB	0.5174	1453.08252	38.22967	1.6796

Totals : 8.65129e4 3007.39789

*** End of Report ***

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000003.D
Sample Name: DT211.54

```
=====
Injection Date   : 5/30/2005 5:02:43 PM      Seq. Line :    3
Sample Name     : DT211.54                  Location  : Vial 22
Acq. Operator   : David                     Inj       :    1
Acq. Instrument : Glycomics-HPLC1           Inj Volume: 20 µl
Method          : C:\HPCHEM\1\METHODS\5ACN.M
Last changed    : 5/30/2005 3:31:21 PM by David
=====
```



Area Percent Report

```
=====
Sorted By      :      Signal
Multiplier     :      1.0000
Dilution       :      1.0000
Use Multiplier & Dilution Factor with ISTDs
=====
```

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.276	BV	0.1008	54.48933	7.32608	0.0858
2	2.498	VV	0.0564	244.58868	67.37842	0.3850
3	2.659	VV	0.0705	182.19691	36.31833	0.2868
4	2.781	VV	0.0638	88.01978	19.80893	0.1385
5	2.957	VV	0.1136	342.78317	43.63585	0.5396
6	3.040	VV	0.0514	68.33237	19.25927	0.1076
7	3.149	VV	0.0721	156.97586	30.48677	0.2471
8	3.462	VV	0.1010	77.82953	10.68656	0.1225
9	3.601	VV	0.1065	183.64082	26.60490	0.2891
10	3.691	VV	0.0702	37.43563	7.25073	0.0589
11	3.864	VV	0.3747	7507.89014	248.32466	11.8177
12	6.016	VV	0.5395	5.21436e4	1335.40759	82.0760
13	6.768	VB	0.3232	1875.17932	88.48716	2.9516
14	7.858	BB	0.5841	567.91199	14.98731	0.8939

Totals : 6.35309e4 1955.96258

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000003.D
Sample Name: DT211.54

```
=====
Injection Date   : 5/30/2005 5:02:43 PM      Seq. Line :    3
Sample Name      : DT211.54                  Location  : Vial 22
Acq. Operator    : David                      Inj       :    1
Acq. Instrument  : Glycomics-HPLC1           Inj Volume: 20 µl
Method           : C:\HPCHEM\1\METHODS\5ACN.M
Last changed     : 5/30/2005 3:31:21 PM by David
=====
```

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.271	BV	0.1112	31.07165	4.06267	0.0296
2	2.507	VV	0.0568	21.21320	5.51812	0.0202
3	2.661	VV	0.0642	93.23981	20.82556	0.0888
4	2.776	VV	0.0545	26.59566	7.29567	0.0253
5	2.957	VV	0.1083	155.55569	21.02205	0.1482
6	3.041	VV	0.0504	33.29402	9.62630	0.0317
7	3.149	VB	0.0614	67.39148	16.57545	0.0642
8	3.601	BV	0.1149	70.94853	9.52801	0.0676
9	3.698	VV	0.0626	11.19408	2.58175	0.0107
10	3.867	VV	0.4549	7.82523e4	2586.64307	74.5451
11	6.016	VV	0.5289	2.48819e4	652.43707	23.7031
12	6.766	VB	0.3274	949.34027	44.05315	0.9044
13	7.859	BB	0.6129	378.98303	9.35059	0.3610

Totals : 1.04973e5 3389.51945

Signal 3: DAD1 E, Sig=280,16 Ref=360,100

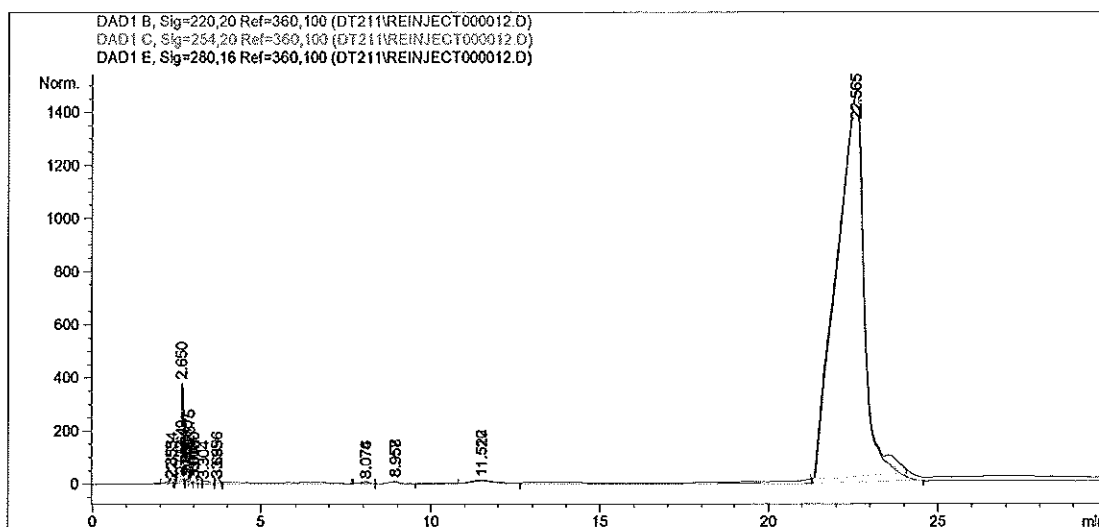
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.508	BV	0.0513	13.90108	4.12383	0.0179
2	2.631	VV	0.0899	37.64062	6.46332	0.0484
3	2.775	VB	0.0717	8.21038	1.60542	0.0105
4	2.976	BV	0.0837	8.96288	1.50061	0.0115
5	3.160	VB	0.0594	16.05606	4.12586	0.0206
6	3.601	BV	0.1017	37.70230	5.97072	0.0484
7	3.699	VV	0.0560	5.15229	1.30623	6.619e-3
8	3.867	VB	0.4410	7.50677e4	2568.13818	96.4310
9	6.015	BV	0.5310	1849.23853	48.26309	2.3755
10	6.743	VV	0.3374	89.00774	3.97276	0.1143
11	7.863	VB	0.5851	712.42413	19.01387	0.9152

Totals : 7.78460e4 2664.48390

*** End of Report ***

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000012.D
Sample Name: DT211.55

```
=====
Injection Date   : 5/30/2005 10:02:25 PM      Seq. Line : 12
Sample Name     : DT211.55                  Location  : Vial 28
Acq. Operator   : David                      Inj       : 1
Acq. Instrument : Glycomics-HPLC1            Inj Volume: 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\20ACN.M
Last changed    : 5/30/2005 3:32:20 PM by David
Analysis Method : C:\HPCHEM\1\METHODS\5ACN.M
Last changed    : 5/30/2005 3:31:21 PM by David
=====
```



=====
Area Percent Report
=====

Sorted By : Signal
Multiplier : 1.0000
Dilution : 1.0000
Use Multiplier & Dilution Factor with ISTDs

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.353	BV	0.1252	62.67691	6.46009	0.0738
2	2.649	VV	0.0722	415.22372	83.30647	0.4887
3	2.800	VV	0.0690	84.93230	18.69340	0.1000
4	2.872	VV	0.0687	131.06685	26.95637	0.1543
5	3.018	VV	0.1124	66.37849	7.71540	0.0781
6	3.304	VV	0.1531	32.93591	2.80305	0.0388
7	3.695	VB	0.0736	30.47073	6.17062	0.0359
8	8.074	BV	0.2652	50.43297	3.04505	0.0594
9	8.957	VB	0.3229	194.03682	8.95294	0.2284
10	11.520	BB	0.6495	485.14371	11.24204	0.5710
11	22.565	BB	0.8221	8.34122e4	1465.25513	98.1718

Totals : 8.49655e4 1640.60056

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000012.D
Sample Name: DT211.55

```
=====
Injection Date   : 5/30/2005 10:02:25 PM      Seq. Line : 12
Sample Name      : DT211.55                  Location  : Vial 28
Acq. Operator    : David                     Inj       : 1
Acq. Instrument  : Glycomics-HPLC1           Inj Volume: 20 µl
Acq. Method      : C:\HPCHEM\1\METHODS\20ACN.M
Last changed     : 5/30/2005 3:32:20 PM by David
Analysis Method  : C:\HPCHEM\1\METHODS\5ACN.M
Last changed     : 5/30/2005 3:31:21 PM by David
=====
```

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.353	BB	0.0682	10.59450	2.19956	0.0179
2	2.661	BV	0.0768	66.74113	12.02306	0.1130
3	2.805	VV	0.0538	27.48238	7.66758	0.0465
4	2.875	VV	0.0607	38.16763	9.14929	0.0646
5	3.018	VB	0.0633	9.20163	2.09277	0.0156
6	3.695	BB	0.0707	20.93880	4.30606	0.0355
7	8.076	BV	0.2648	63.39040	3.75867	0.1073
8	8.958	VB	0.3268	157.08046	7.19240	0.2660
9	11.512	BB	0.6475	423.01569	9.87940	0.7163
10	22.565	BB	0.8285	5.82401e4	1019.41705	98.6172

Totals : 5.90567e4 1077.68584

Signal 3: DAD1 E, Sig=280,16 Ref=360,100

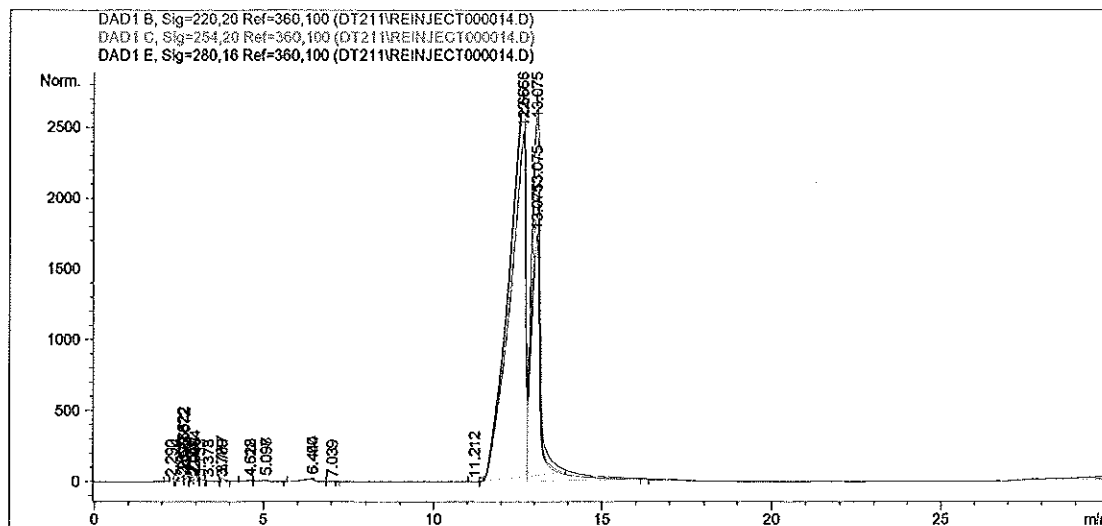
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.354	BB	0.0642	5.30922	1.18765	0.1463
2	2.650	BV	0.0847	83.45076	15.52533	2.3002
3	2.807	VV	0.0574	16.71251	4.29769	0.4606
4	2.875	VV	0.0629	20.30084	4.65496	0.5596
5	3.020	VB	0.0672	5.32583	1.12611	0.1468
6	3.696	BB	0.0650	6.39816	1.46458	0.1764
7	22.565	BB	0.7967	3490.55933	60.89958	96.2102

Totals : 3628.05664 89.15591

*** End of Report ***

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000014.D
Sample Name: DT211.60

```
=====
Injection Date   : 5/30/2005 11:10:37 PM      Seq. Line :   14
Sample Name     : DT211.60                  Location  : Vial 29
Acq. Operator   : David                      Inj       :    1
Acq. Instrument : Glycomics-HPLC1            Inj Volume: 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\25ACN.M
Last changed    : 5/30/2005 3:32:36 PM by David
Analysis Method : C:\HPCHEM\1\METHODS\5ACN.M
Last changed    : 5/30/2005 3:31:21 PM by David
=====
```



Area Percent Report

```
=====
Sorted By      : Signal
Multiplier     : 1.0000
Dilution       : 1.0000
Use Multiplier & Dilution Factor with ISTDs
=====
```

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.290	BV	0.1077	26.17551	3.25758	0.0196
2	2.623	VV	0.0569	192.14078	47.82999	0.1440
3	2.665	VV	0.0496	192.22287	56.62530	0.1441
4	2.886	VV	0.0779	69.40511	13.07369	0.0520
5	2.968	VV	0.0696	104.23949	21.10854	0.0781
6	3.132	VV	0.1031	33.18549	4.24838	0.0249
7	3.378	VV	0.1955	43.61155	2.96644	0.0327
8	3.780	VB	0.0899	41.35683	6.35570	0.0310
9	4.622	BV	0.2025	48.64078	3.77009	0.0365
10	5.098	VB	0.2796	229.18282	11.74836	0.1718
11	6.450	BV	0.3636	485.17548	20.69070	0.3636
12	7.039	VB	0.1219	8.93966	1.11111	6.700e-3
13	11.212	BV	0.1496	14.74729	1.48869	0.0111
14	12.656	VV	0.5083	9.77165e4	2751.83838	73.2371
15	13.075	VB	0.3108	3.42193e4	1744.72937	25.6469

Totals : 1.33425e5 4690.84232

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000014.D
Sample Name: DT211.60

```
=====
Injection Date   : 5/30/2005 11:10:37 PM      Seq. Line :   14
Sample Name     : DT211.60                  Location  : Vial 29
Acq. Operator   : David                      Inj       :    1
Acq. Instrument : Glycomics-HPLC1            Inj Volume: 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\25ACN.M
Last changed    : 5/30/2005 3:32:36 PM by David
Analysis Method : C:\HPCHEM\1\METHODS\5ACN.M
Last changed    : 5/30/2005 3:31:21 PM by David
=====
```

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.292	BB	0.0773	6.80936	1.25502	7.033e-3
2	2.620	BV	0.0541	30.80171	8.14771	0.0318
3	2.675	VB	0.0536	58.70607	16.46337	0.0606
4	2.976	BV	0.0858	45.44607	7.18753	0.0469
5	3.133	VB	0.0963	14.91393	2.11400	0.0154
6	3.779	BB	0.0890	26.98674	4.19515	0.0279
7	4.618	BV	0.2080	54.96881	4.05786	0.0568
8	5.097	VB	0.2686	176.84933	9.27026	0.1827
9	6.444	BB	0.3894	427.98938	17.37193	0.4421
10	12.656	BV	0.5002	6.77887e4	1966.54968	70.0170
11	13.075	VB	0.3146	2.81853e4	1438.35437	29.1118

Totals : 9.68175e4 3474.96688

Signal 3: DAD1 E, Sig=280,16 Ref=360,100

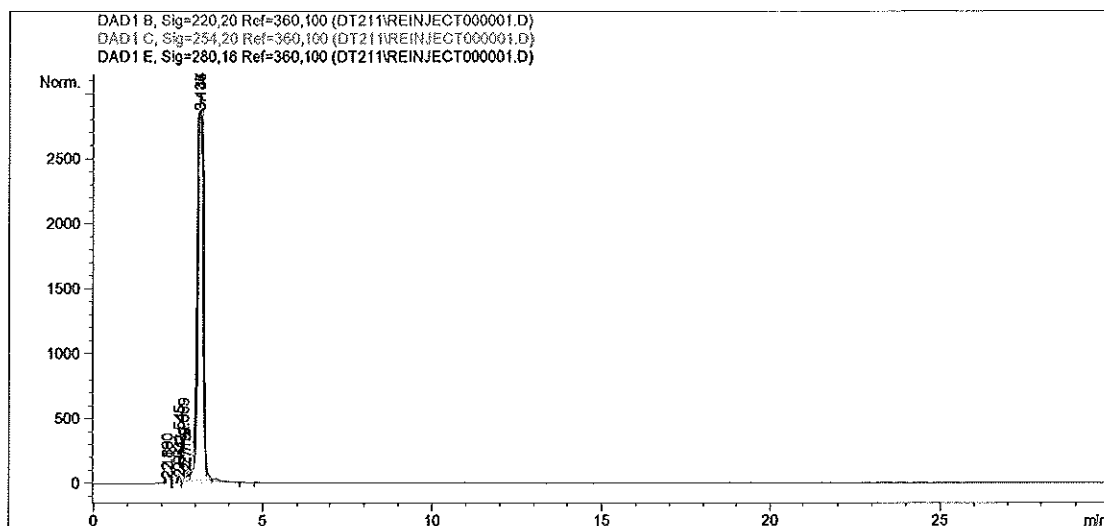
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.622	BV	0.0520	38.32864	11.17493	0.5884
2	2.672	VV	0.0585	44.19694	11.10049	0.6784
3	2.974	VV	0.0959	20.94188	2.91174	0.3215
4	3.777	BB	0.0726	6.70520	1.33541	0.1029
5	12.656	BV	0.4623	4039.12500	122.63206	62.0019
6	13.075	VB	0.2905	2365.21558	136.27567	36.3069

Totals : 6514.51324 285.43030

*** End of Report ***

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000001.D
Sample Name: DT211.52

```
=====
Injection Date   : 5/30/2005 3:54:24 PM      Seq. Line :    1
Sample Name     : DT211.52                  Location  : Vial 21
Acq. Operator   : David                     Inj       :    1
Acq. Instrument : Glycomics-HPLC1           Inj Volume: 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\O-ET-REINJECT.M
Last changed    : 5/30/2005 3:30:44 PM by David
Analysis Method : C:\HPCHEM\1\METHODS\5ACN.M
Last changed    : 5/30/2005 3:31:21 PM by David
=====
```



=====
Area Percent Report
=====

Sorted By : Signal
Multiplier : 1.0000
Dilution : 1.0000
Use Multiplier & Dilution Factor with ISTDs

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.186	BV	0.0674	6.71077	1.46498	0.0139
2	2.531	VV	0.0538	216.44858	60.39103	0.4481
3	2.719	VV	0.0848	618.78271	101.92947	1.2810
4	3.144	VB	0.2242	4.74614e4	3006.40332	98.2570

Totals : 4.83034e4 3170.18881

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.189	BB	0.0726	8.82101	1.69767	0.0276
2	2.541	BV	0.0536	38.85307	10.88609	0.1216
3	2.719	VV	0.0798	297.71637	54.45380	0.9316
4	3.137	VB	0.2124	3.16127e4	2451.46851	98.9192

Totals : 3.19581e4 2518.50606

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000001.D
Sample Name: DT211.52

```
=====
Injection Date   : 5/30/2005 3:54:24 PM      Seq. Line :    1
Sample Name      : DT211.52                  Location  : Vial 21
Acq. Operator    : David                     Inj       :    1
Acq. Instrument  : Glycomics-HPLC1           Inj Volume: 20 µl
Acq. Method      : C:\HPCHEM\1\METHODS\O-ET-REINJECT.M
Last changed     : 5/30/2005 3:30:44 PM by David
Analysis Method  : C:\HPCHEM\1\METHODS\5ACN.M
Last changed     : 5/30/2005 3:31:21 PM by David
=====
```

Signal 3: DAD1 E, Sig=280,16 Ref=360,100

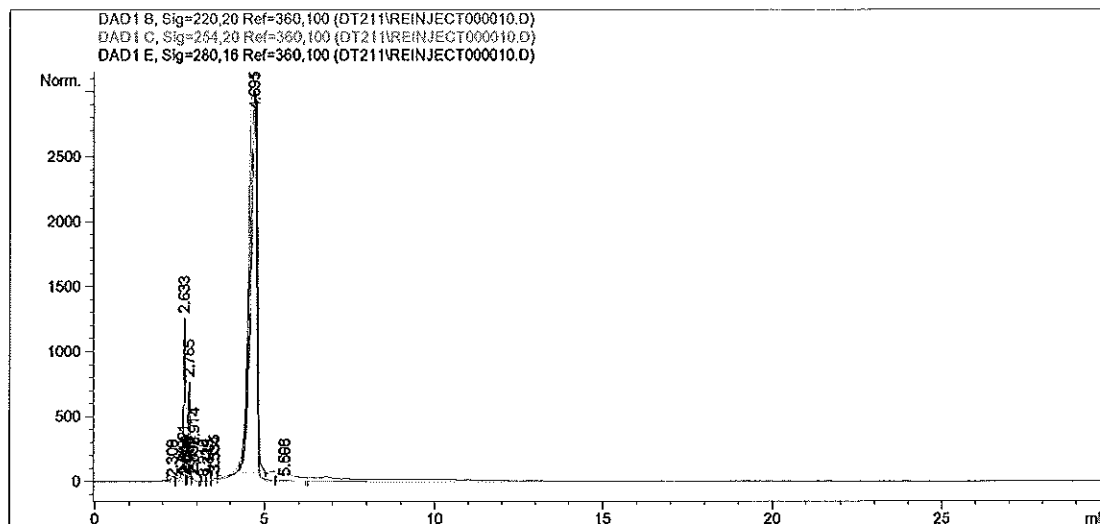
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.190	BB	0.0762	6.95844	1.26457	1.0495
2	2.545	BV	0.0556	20.87447	5.85127	3.1484
3	2.669	VB	0.0878	46.22507	7.11859	6.9719
4	3.136	BB	0.1344	588.96063	65.84416	88.8302

Totals : 663.01861 80.07859

*** End of Report ***

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000010.D
Sample Name: DT211.75

```
=====
Injection Date   : 5/30/2005 8:54:12 PM      Seq. Line :   10
Sample Name     : DT211.75                  Location  : Vial 27
Acq. Operator   : David                     Inj       :    1
Acq. Instrument : Glycomics-HPLC1            Inj Volume: 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\15ACN.M
Last changed    : 5/30/2005 3:32:05 PM by David
Analysis Method : C:\HPCHEM\1\METHODS\5ACN.M
Last changed    : 5/30/2005 3:31:21 PM by David
=====
```



=====
Area Percent Report
=====

Sorted By : Signal
Multiplier : 1.0000
Dilution : 1.0000
Use Multiplier & Dilution Factor with ISTDs

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.308	BV	0.0801	24.81610	4.25033	0.0419
2	2.621	VV	0.0692	592.57318	120.91873	1.0008
3	2.690	VV	0.0385	75.13650	28.70242	0.1269
4	2.767	VV	0.0688	253.79358	52.15182	0.4286
5	2.894	VB	0.1109	206.90065	25.98501	0.3494
6	3.214	BV	0.1009	48.21884	6.47073	0.0814
7	3.329	VV	0.1167	35.53308	4.67218	0.0600
8	3.535	VV	0.0999	243.89713	35.61075	0.4119
9	4.691	VV	0.2594	5.75292e4	3004.74390	97.1625
10	5.598	VB	0.4011	199.17188	7.42193	0.3364

Totals : 5.92092e4 3290.92780

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Sample Name: DT211.75

Injection Date	: 5/30/2005 8:54:12 PM	Seq. Line	: 10
Sample Name	: DT211.75	Location	: Vial 27
Acq. Operator	: David	Inj	: 1
Acq. Instrument	: Glycomics-HPLC1	Inj Volume	: 20 µl
Acq. Method	: C:\HPCHEM\1\METHODS\15ACN.M		
Last changed	: 5/30/2005 3:32:05 PM by David		
Analysis Method	: C:\HPCHEM\1\METHODS\5ACN.M		
Last changed	: 5/30/2005 3:31:21 PM by David		

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.309	BV	0.0713	7.62329	1.49993	0.0216
2	2.634	VV	0.0777	70.16982	12.45682	0.1991
3	2.694	VV	0.0371	19.68122	7.88784	0.0558
4	2.768	VV	0.0608	93.40009	22.30805	0.2650
5	2.892	VB	0.1009	81.09901	11.40949	0.2301
6	3.218	BV	0.0830	14.04432	2.30825	0.0398
7	3.345	VV	0.1047	9.34005	1.34955	0.0265
8	3.534	VV	0.0976	124.49924	18.71817	0.3533
9	4.695	VB	0.2506	3.46940e4	2193.18823	98.4411
10	5.606	BB	0.4036	129.55333	4.91436	0.3676

Totals : 3.52434e4 2276.04068

Signal 3: DAD1 E, Sig=280,16 Ref=360,100

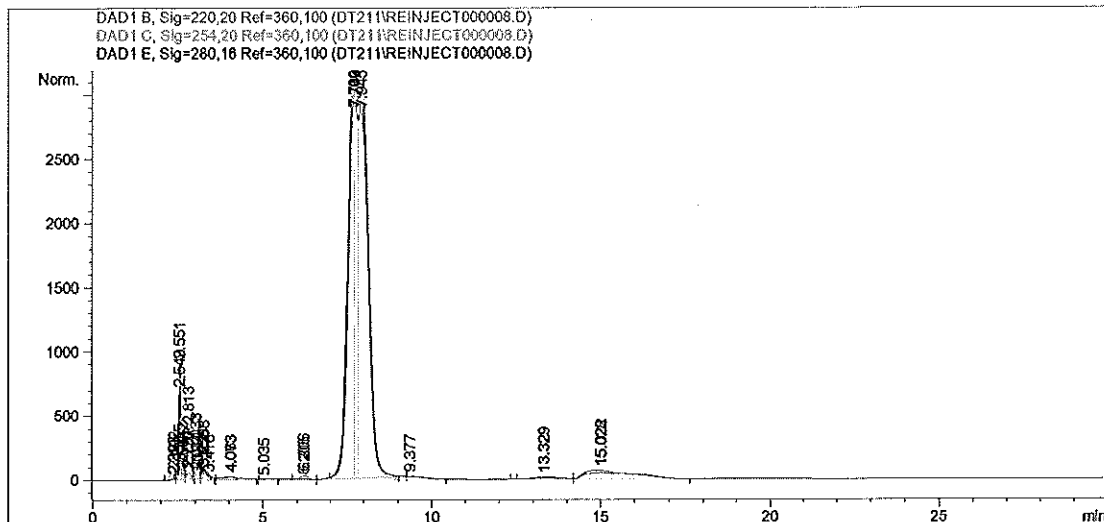
Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.633	BV	0.0711	81.93826	16.74919	13.1759
2	2.765	VV	0.0608	41.65164	9.96741	6.6977
3	2.914	VB	0.0874	17.03515	2.70767	2.7393
4	4.695	BB	0.1701	481.25620	40.08751	77.3872

Totals :	621.88125	69.51178
----------	-----------	----------

*** End of Report ***

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000008.D
Sample Name: DT211.82

```
=====
Injection Date   : 5/30/2005 7:45:58 PM          Seq. Line :    8
Sample Name      : DT211.82                      Location  : Vial 26
Acq. Operator    : David                          Inj       :    1
Acq. Instrument  : Glycomics-HPLC1                Inj Volume: 20 µl
Acq. Method      : C:\HPCHEM\1\METHODS\10ACN.M
Last changed     : 5/30/2005 3:31:34 PM by David
Analysis Method  : C:\HPCHEM\1\METHODS\5ACN.M
Last changed     : 5/30/2005 3:31:21 PM by David
=====
```



Area Percent Report

```
Sorted By      : Signal
Multiplier     : 1.0000
Dilution      : 1.0000
Use Multiplier & Dilution Factor with ISTDs
```

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.393	BV	0.0679	95.37785	19.89544	0.0677
2	2.549	VV	0.0556	2585.03418	691.96381	1.8353
3	2.642	VV	0.0610	526.00525	125.29282	0.3734
4	2.770	VV	0.1254	546.36371	57.21726	0.3879
5	3.074	VV	0.1426	497.80679	45.87096	0.3534
6	3.258	VV	0.1682	997.25726	75.46400	0.7080
7	4.063	VB	0.3969	712.25500	24.96196	0.5057
8	6.206	BV	0.2805	589.45618	32.99024	0.4185
9	7.713	VV	0.2313	4.56578e4	3039.49243	32.4149
10	7.746	VV	0.3095	7.96107e4	3040.00879	56.5199
11	9.377	VB	0.4601	973.92426	26.42100	0.6914
12	13.329	BV	0.8431	1002.44086	17.79763	0.7117
13	15.022	VB	1.2213	7059.85840	70.92416	5.0122

Totals : 1.40854e5 7268.30051

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000008.D
Sample Name: DT211.82

```
=====
Injection Date   : 5/30/2005 7:45:58 PM          Seq. Line :    8
Sample Name      : DT211.82                      Location  : Vial 26
Acq. Operator    : David                        Inj       :    1
Acq. Instrument  : Glycomics-HPLC1              Inj Volume: 20 µl
Acq. Method      : C:\HPCHEM\1\METHODS\10ACN.M
Last changed     : 5/30/2005 3:31:34 PM by David
Analysis Method  : C:\HPCHEM\1\METHODS\5ACN.M
Last changed     : 5/30/2005 3:31:21 PM by David
=====
```

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.393	BV	0.0602	34.83666	8.44006	0.0327
2	2.555	VV	0.0704	456.13901	94.40737	0.4284
3	2.645	VV	0.0599	216.45300	52.75651	0.2033
4	2.797	VV	0.1303	465.46948	53.07554	0.4371
5	3.087	VV	0.1224	103.01102	11.50624	0.0967
6	3.255	VV	0.1002	412.70419	54.56799	0.3876
7	3.418	VB	0.0895	148.76250	24.97529	0.1397
8	4.073	BB	0.3874	493.37558	21.36809	0.4633
9	5.035	BB	0.2582	25.44212	1.37281	0.0239
10	6.205	BV	0.2753	164.84569	9.45897	0.1548
11	7.739	VB	0.5443	9.81208e4	2508.92969	92.1447
12	13.329	BV	0.7798	638.98395	11.98891	0.6001
13	15.028	VB	1.2165	5204.68604	52.68541	4.8877

Totals : 1.06486e5 2905.53289

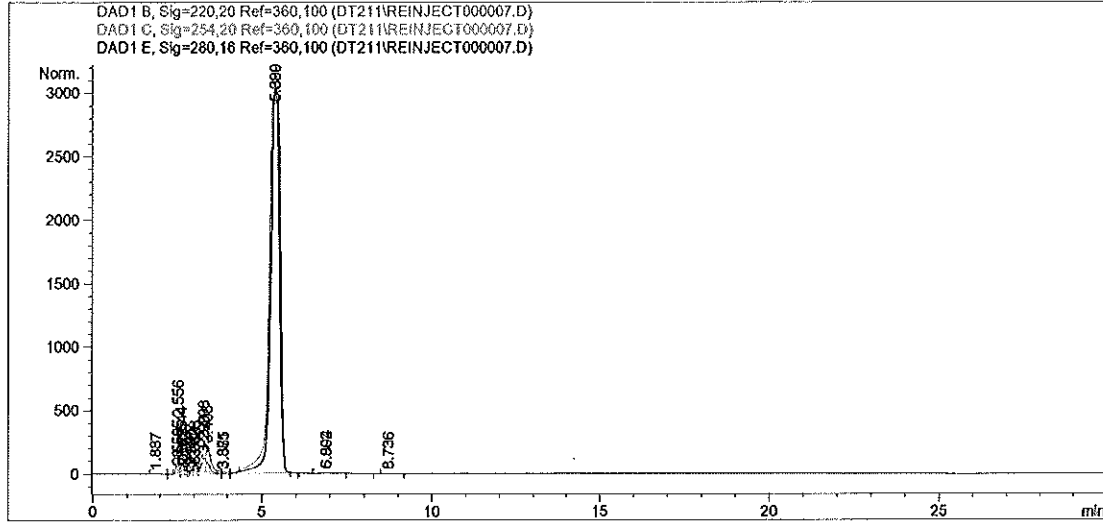
Signal 3: DAD1 E, Sig=280,16 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.402	BV	0.0496	12.12990	3.79863	0.1169
2	2.551	VV	0.0894	449.81766	69.60860	4.3357
3	2.813	VV	0.1461	302.88260	30.43967	2.9195
4	3.033	VV	0.1115	113.34844	14.44802	1.0926
5	3.258	VB	0.1035	72.53064	9.67331	0.6991
6	7.700	BV	0.3022	4669.55371	224.90622	45.0093
7	7.945	VB	0.3108	4754.37109	232.29912	45.8269

Totals : 1.03746e4 585.17356

*** End of Report ***

```
=====
Injection Date   : 5/30/2005 7:14:16 PM      Seq. Line :    7
Sample Name     : DT211.84                  Location  : Vial 25
Acq. Operator   : David                     Inj       :    1
Acq. Instrument : Glycomics-HPLC1           Inj Volume: 20 µl
Acq. Method     : C:\HPCHEM\1\METHODS\10ACN.M
Last changed    : 5/30/2005 3:31:34 PM by David
Analysis Method : C:\HPCHEM\1\METHODS\5ACN.M
Last changed    : 5/30/2005 3:31:21 PM by David
=====
```



=====
Area Percent Report
=====

```
Sorted By      :      Signal
Multiplier     :      1.0000
Dilution       :      1.0000
Use Multiplier & Dilution Factor with ISTDs
```

Signal 1: DAD1 B, Sig=220,20 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	1.887	BV	0.2437	17.04161	1.10618	0.0221
2	2.555	VV	0.0612	410.88754	97.45527	0.5340
3	2.655	VV	0.0892	577.85327	89.61031	0.7509
4	2.925	VV	0.1234	321.63779	37.74090	0.4180
5	3.034	VV	0.0939	282.50082	43.42561	0.3671
6	3.299	VV	0.2125	2102.26465	130.13390	2.7319
7	3.871	VV	0.1420	53.04193	4.99117	0.0689
8	5.389	VB	0.3817	7.30823e4	3070.32568	94.9708
9	6.894	BB	0.2850	64.21659	3.55077	0.0834
10	8.736	BB	0.3011	40.65155	2.01756	0.0528

Totals : 7.69524e4 3480.35735

Signal 2: DAD1 C, Sig=254,20 Ref=360,100

Data File C:\HPCHEM\1\DATA\DT211\REINJECT000007.D
Sample Name: DT211.84

```
=====
Injection Date   : 5/30/2005 7:14:16 PM          Seq. Line :    7
Sample Name      : DT211.84                      Location  : Vial 25
Acq. Operator    : David                          Inj       :    1
Acq. Instrument  : Glycomics-HPLC1                Inj Volume: 20 µl
Acq. Method      : C:\HPCHEM\1\METHODS\10ACN.M
Last changed     : 5/30/2005 3:31:34 PM by David
Analysis Method  : C:\HPCHEM\1\METHODS\5ACN.M
Last changed     : 5/30/2005 3:31:21 PM by David
=====
```

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.558	BV	0.0608	113.28417	27.06050	0.2184
2	2.673	VV	0.0783	169.90862	29.00428	0.3276
3	2.737	VV	0.0627	103.01408	24.68801	0.1986
4	2.927	VV	0.1084	148.46213	19.59868	0.2863
5	3.030	VV	0.1169	274.14929	33.01569	0.5286
6	3.299	VV	0.1958	1634.71472	111.04140	3.1520
7	3.885	VV	0.1358	25.08338	2.52657	0.0484
8	5.390	VB	0.3031	4.93197e4	2556.13965	95.0951
9	6.902	BB	0.3756	75.21820	2.89089	0.1450

Totals : 5.18636e4 2805.96569

Signal 3: DAD1 E, Sig=280,16 Ref=360,100

Peak #	RetTime [min]	Type	Width [min]	Area [mAU*s]	Height [mAU]	Area %
1	2.556	BV	0.0641	146.31795	32.73884	3.1176
2	2.644	VV	0.0724	74.88522	15.47677	1.5956
3	2.739	VB	0.0645	22.32774	4.95807	0.4757
4	3.028	BV	0.1261	58.55417	6.56451	1.2476
5	3.298	VV	0.1129	148.39014	19.04381	3.1617
6	3.406	VB	0.1345	142.00948	15.55700	3.0258
7	5.389	BB	0.2519	4100.84863	246.74292	87.3760

Totals : 4693.33332 341.08192

*** End of Report ***

