

**TOWARDS THE DEVELOPMENT AND DISCOVERY OF INHIBITORS FOR  
TRYPANOSOMA CRUZI TRANS-SIALIDASE**

by

Kyle Brian Robinson

B.Sc., McGill University, 2013

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

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES  
(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA  
(Vancouver)

April 2020

© Kyle Brian Robinson, 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Towards the development and discovery of inhibitors for Trypanosoma cruzi trans-sialidase

submitted by Kyle Robinson in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Chemistry

**Examining Committee:**

Prof. Stephen Withers, Chemistry and Biochemistry  
Supervisor

Prof. Martin Tanner, Chemistry  
Supervisory Committee Member

Prof. Glenn Sammis, Chemistry  
University Examiner

Prof. Michel Roberge, Biochemistry and Molecular Biology  
University Examiner

**Additional Supervisory Committee Members:**

Prof. Raymond Andersen, Chemistry and Earth Ocean & Atmospheric Sciences  
Supervisory Committee Member

Prof. Reinhard Jetter, Chemistry and Botany  
Supervisory Committee Member

## Abstract

The parasite *Trypanosoma cruzi* displays a trans-sialidase (TcTS) on its surface that is hypothesized to be a therapeutic target for Chagas disease. TcTS transfers sialic acid from the cells of infected hosts to the surface of the pathogenic parasite, masking it from immune recognition and enhancing cellular invasion. The design of TcTS inhibitors has been largely unsuccessful to date. Accordingly, this work aims to identify new TcTS inhibitors, both by modifying existing inhibitors and by screening natural product and peptide libraries to discover new chemical scaffolds for this purpose. First, analogues of the established mechanism-based inhibitor, difluorosialic acid (DFSA), were investigated in search of increased potency and selectivity for TcTS. The synthesis and kinetic analysis of nine C9 amide-linked DFSAs and seven *N*-acyl modified DFSAs was explored to this end. One candidate was identified that inhibited TcTS 10-fold better than the unmodified precursor. Further, TcTS showed a tolerance for the C5 functionalized inhibitors, a fact that can be leveraged – together with C9 substitution – for specificity versus human neuraminidases. Next, a library of ~1000 marine organism extracts was screened for TcTS inhibition, from which five hits were selected. Bioassay-guided isolation and structural determination of the active chemicals afforded two new natural product inhibitors of TcTS with  $IC_{50}$  values  $<25 \mu\text{M}$ . The most potent species belongs to the ircinialactam family and its structure-activity relationship was assessed by screening a library of synthetic ircinialactam-like compounds. These studies revealed a preference for chemical species bearing furan, tetrionic acid or glycinal lactam functional groups separated by unsaturated carbon chains 10-15 atoms in length. The eight best synthetic derivatives had a reversible, non-competitive mode of action with  $K_i < 5 \mu\text{M}$ . These species were further tested in a parasite growth inhibition assay affording inconclusive results. Finally, random nonstandard peptide integrated discovery (RaPID) screening identified two peptide sequences – bearing 4-5 hydroxylated non-proteinogenic amino acid residues – that were exogenously cyclized with disuccinimidyl glutarate. The best macrocyclic peptide inhibitor had a  $K_i = 1.4 \mu\text{M}$ . Together these studies identified a diverse series of new chemical inhibitors that can be further explored to assess the therapeutic relevance of TcTS.

## Lay Summary

Chagas is a neglected tropical disease that affects >10 million people worldwide and is caused by the pathogenic parasite *Trypanosoma cruzi*. This parasite produces an enzyme that steals specific sugar molecules from the body in order to camouflage itself from the human immune system. The goal of this work was to develop new chemical molecules that will selectively block this enzyme; two complementary approaches were taken to this end. First, I designed sugar-based inhibitors that can mimic the natural molecules but instead block the enzyme activity. Second, I searched for new chemical molecules from natural sources that could be used for this purpose. Both of these approaches yielded new best-in-class chemicals for inhibiting the enzyme activity. Finding new chemical inhibitors is the foundation drug discovery, as such these findings can be further investigated towards the development of new therapeutics for Chagas disease.

## Preface

Chapter 2 is based on work previously conducted in Prof. Stephen Withers lab by Dr. Sabrina Buchini and Yuan Yao at UBC. I carried out all steps in the synthesis of the DFSA except for the coupling step between **2.61** and seven commercial acyl chlorides. Dr. Hongming Chen performed this reaction, making compounds **2.62-2.69**. I performed kinetic experiments and analysis on all compounds.

The work in Chapter 3 was done in collaboration with Prof. Raymond Andersen's lab at UBC and Prof. Robert Capon's lab at the University of Queensland. Prof. Andersen's lab provided the library of natural product extracts for screening. I designed and executed the screening for TcTS inhibition. I performed the small-scale extraction of lead sponge extracts and the extraction of sponge sample PNG-03-395-2 under the supervision of Dr. David Williams. I synthesized **3.5** and performed all inhibition assays. Dr. Williams performed all steps during the isolation of RJA-00-287-2-A&B-D-CH<sub>2</sub>N<sub>2</sub>-3 from sample RJA-00-287-2. He also performed isocratic HPLC, NMR acquisition and structural identification for compounds **3.1-3.4**. Prof. Capon provided the synthetic ircinialactam library of 117 compounds, along with additional samples of compounds **3.10-3.16** and **3.18-3.21** and the set of 15 cacolide natural products **3.22-3.36**. A materials transfer agreement was signed between the University of Queensland and UBC. I performed all inhibition screening and kinetic analysis on the provided materials.

Chapter 4 was done in collaboration with Dr. Seino Jongkees in Prof. Hiroaki Suga's lab at the University of Tokyo. Dr. Jongkees performed all cyclic-peptide screening and synthesized peptides **4.1-4.9**. I synthesized modified cysteine derivative **4.24**, coupling agent **4.19**, linear peptides **4.20-4.21**; **4.29-4.32** and cyclic peptides **4.33-4.38**. I performed all enzyme kinetics on the linear and cyclic peptides.

Enzyme crystallization in Chapter 5 was based on previous reports by Dr. Alejandro Buschiazzo and cell-based assays were performed in collaboration with Prof. Fred Buckner's lab at the University of Washington. I was responsible for the recombinant expression and purification of TcTS and for the design and execution of all enzyme crystallization trials. Dr. Ranae Ranade performed parasite growth inhibition assays at the University of Washington.

# Table of Contents

<b>Abstract</b> .....	<b>iii</b>
<b>Lay Summary</b> .....	<b>iv</b>
<b>Preface</b> .....	<b>v</b>
<b>Table of Contents</b> .....	<b>vi</b>
<b>List of Tables</b> .....	<b>xiii</b>
<b>List of Figures</b> .....	<b>xv</b>
<b>List of Schemes</b> .....	<b>xix</b>
<b>List of Abbreviations</b> .....	<b>xx</b>
<b>Acknowledgements</b> .....	<b>xxvi</b>
<b>Dedication</b> .....	<b>xxvii</b>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1    Chagas disease.....	1
1.2 <i>Trypanosoma cruzi</i> .....	2
1.3 <i>Trypanosoma cruzi</i> trans-sialidase .....	5
1.4    Carbohydrate-based inhibitors.....	10
1.4.1    Transition state analogues .....	10
1.4.2    Substrate analogues.....	11
1.4.3    Mechanism-based inhibitors .....	15
1.5    Screening for enzyme inhibitors.....	18
1.5.1    Introduction .....	18
1.5.2    Screening for TcTS inhibitors .....	20
1.6    Aims of thesis .....	23
1.6.1    Design and synthesis of modified difluorosialic acid inactivators .....	23

1.6.2	Screening for new classes of TcTS inhibitors .....	24
<b>Chapter 2: Design of Mechanism-Based Inhibitors for TcTS .....</b>		<b>25</b>
2.1	Introduction .....	25
2.1.1	Case study: influenza neuraminidase .....	25
2.1.2	Structural modification of sialoside-mimics for enzyme specificity .....	28
2.1.2.1	Modification of DANA for hNEU specificity .....	28
2.1.2.2	Modification of difluorosialic acid glycerol chain for TcTS specificity .....	31
2.1.2.3	Modification of sialosides at C4 and C5 for TcTS specificity .....	36
2.2	Specific aims .....	39
2.3	Synthesis of 9-amido-difluorosialic acids .....	40
2.3.1	Azide substitution of 6-tosyl-ManNAc .....	42
2.3.2	Neu5Ac aldolase condensation .....	43
2.3.3	Acylation of 9-amino-difluorosialic acid .....	45
2.4	Synthesis of <i>N</i> -acyl-difluorosialic acids .....	48
2.4.1	Elimination of Neu5Ac .....	49
2.4.2	Electrophilic fluorination of Neu5Ac2en .....	50
2.4.3	Acylation of 5-amino-difluorosialic acid .....	50
2.5	Kinetic analysis .....	53
2.5.1	Enzyme expression and characterization .....	53
2.5.2	Inactivation kinetics: protocol .....	53
2.5.3	Inactivation kinetics: 9-amido-DFSAs .....	55
2.5.4	Inactivation kinetics: 5-amido-DFSAs .....	58
2.6	Conclusions .....	61
<b>Chapter 3: Marine Natural Product Screening and Bioassay Guided Isolation of Inhibitors for TcTS .....</b>		<b>63</b>

3.1	Introduction .....	63
3.1.1	Natural carbohydrate-based glycosidase inhibitors.....	65
3.1.2	Marine natural products .....	66
3.1.3	Screening natural product libraries for glycosidase inhibitors.....	70
3.1.4	Case study: screening natural product libraries for $\alpha$ -amylase inhibitors .....	71
3.1.5	Screening for TcTS inhibitors .....	73
3.2	Specific aims .....	73
3.3	Screen design and preparation of extract library .....	74
3.3.1	Screen design and validation.....	74
3.3.2	Preparation of extract library.....	76
3.4	Screening marine organism extracts for TcTS inhibition.....	77
3.4.1	Screening results and hit selection .....	77
3.4.2	Hit validation and TcTS selectivity.....	79
3.5	Isolation and structural determination of hits from marine organisms .....	81
3.5.1	Bioactivity-guided isolation of sample RJA-03-242.....	83
3.5.2	Bioactivity-guided isolation of sample PNG-03-395.....	84
3.5.3	Structural determination of sample(s) from PNG-03-395.....	88
3.5.4	Bioactivity-guided isolation of sample RJA-00-287.....	92
3.5.5	Structural determination of sample from RJA-00-287.....	95
3.6	Kinetic analysis and structure-activity relationship.....	99
3.6.1	Isolated natural products and structural derivatives .....	99
3.6.2	Screening ircinialactam library .....	101
3.6.2.1	Most inhibitory .....	103
3.6.2.2	Inhibitory .....	105
3.6.2.3	Mildly inhibitory and non-inhibitory.....	105

3.6.3	Screening cacolide library .....	106
3.6.4	Inhibition studies: ircinialactam library hits.....	107
3.7	Conclusions .....	110
<b>Chapter 4: High-Throughput Screening and Synthesis of Non-Natural Cyclic Peptides Inhibitors for TcTS .....</b>		<b>112</b>
4.1	Introduction .....	112
4.1.1	Peptide screening .....	112
4.1.2	Peptide-based therapeutics .....	115
4.1.3	Ribosomal peptide synthesis .....	118
4.1.4	Genetic code reprogramming .....	119
4.1.5	RaPID Screening .....	123
4.2	Specific aims .....	124
4.3	Screen 1: cyclization between <i>N</i> -chloroacetamide and cysteine .....	125
4.3.1	Kinetic analysis .....	127
4.3.2	Limitations of screening/selections .....	129
4.4	Screen 2: cyclization between disuccinimidyl glutarate and conserved amines.....	129
4.5	Peptide synthesis and macrocyclization .....	135
4.5.1	Synthesis of L-Cys* .....	135
4.5.2	Manual synthesis of non-proteinogenic peptide hits .....	141
4.5.3	Automated synthesis of proteinogenic peptide analogues .....	144
4.5.4	Deprotection and purification.....	146
4.5.5	Macrocyclization of linear peptides .....	146
4.6	Kinetic analysis .....	150
4.7	Conclusions .....	153
<b>Chapter 5: Towards Enzyme Crystallography and Cell-Based Studies .....</b>		<b>154</b>

5.1	Introduction .....	154
5.2	Specific aims .....	155
5.3	Crystallization of <i>T. cruzi</i> trans-sialidase .....	156
5.3.1	Enzyme preparation .....	157
5.3.2	Growth of seed crystals .....	158
5.3.3	Microseeding attempts .....	161
5.3.4	Screening for new conditions .....	162
5.4	Cell-based studies .....	163
5.4.1	<i>T. cruzi</i> growth inhibition assay .....	165
5.5	Conclusions .....	170
<b>Chapter 6: Concluding Remarks.....</b>		<b>171</b>
<b>Chapter 7: Experimental Procedures .....</b>		<b>177</b>
7.1	Synthesis of DFSAs.....	177
7.1.1	General materials .....	177
7.1.2	Synthesis of 9-amido-DFSA .....	178
7.1.3	General method for the synthesis of NHS-activated esters .....	185
7.1.4	General procedure for 9-amino-DFSA acylation .....	191
7.1.5	Synthesis of 5-amido-DFSA .....	202
7.1.6	General procedure for 5-amino-DFSA acylation .....	207
7.1.7	Enzyme expression.....	215
7.1.8	Enzyme kinetics .....	216
7.1.8.1	Measuring $k_i/K_i$ for 9-amido-difluorosialoside inactivators .....	216
7.1.8.2	Measuring $k_{obs}$ and half-life for 5-amido-difluorosialoside inactivators .....	216
7.2	Natural product screening and isolation for TcTS inhibition .....	217
7.2.1	Screening protocol .....	217

7.2.2	TcTS inhibition assay.....	218
7.2.3	$\beta$ -Galactosidase from <i>Aspergillus oryzae</i> inhibition assay.....	218
7.2.4	Human NEU2 inhibition assay.....	219
7.2.5	General isolation procedure .....	219
7.2.6	Ircinialactam screening .....	220
7.2.7	Enzyme kinetics .....	220
7.3	Cyclic-peptide inhibitors .....	222
7.3.1	Synthesis of DSG .....	222
7.3.2	Synthesis of L-Cys* .....	223
7.3.3	Peptide synthesis .....	227
7.3.4	Amino acid coupling with DSG .....	231
7.3.5	Peptide cyclization and purification.....	232
7.3.6	Quantifying peptide stocks.....	235
7.3.7	Enzyme kinetics .....	237
7.4	<i>Trypanosoma cruzi</i> growth inhibition assay.....	238
7.4.1	Standard growth inhibition assay procedure .....	238
7.4.2	Modified growth inhibition assay procedure.....	239
	<b>References.....</b>	<b>240</b>
	<b>Appendices.....</b>	<b>257</b>
	Appendix A : Enzyme kinetics.....	257
A.1	Michaelis-Menten kinetics .....	257
A.2	Inactivation kinetics (irreversible inhibitors) .....	259
A.3	Inhibition kinetics (reversible inhibitors).....	261
	Appendix B : Kinetic plots of TcTS inactivation by DFSA analogues .....	265
	Appendix C : Kinetic plots of TcTS inhibition by natural product and ircinialactam library hits .....	271

Appendix D : Kinetic plots of TcTS inhibition by peptide hits..... 278

## List of Tables

Table 2.1. Inhibition constants of DANA analogues versus vNEU and hNEU.....	26
Table 2.2. Inactivation and reactivation parameters of DFSA analogues versus vNEU (G70C H1N9).....	28
Table 2.3. Inactivation and reactivation parameters of C9 amide-linked DFSA analogues versus TcTS. .	32
Table 2.4. Structures and inactivation parameters of C9 triazole-linked DFSA analogues versus TcTS...	35
Table 2.5. Reaction yields from 9-amido-DFSA coupling reactions. ....	47
Table 2.6. Reaction yields from 5-amido-DFSA coupling reactions. ....	52
Table 2.7. Inactivation parameters of first-generation 9-amido-DFSAs versus TcTS.....	55
Table 2.8. Inactivation parameters of second-generation 9-amido-DFSAs versus TcTS.....	57
Table 2.9. Inactivation rate constants ( $k_{\text{obs}}$ ) and half-life ( $t_{1/2}$ ) of 5-amido-DFSAs (5 mM) versus TcTS. .	60
Table 3.1. $Z'$ -factor analysis and interpretation. ....	75
Table 3.2. Effect of natural product extract dilution on hit rate for TcTS inhibition.....	76
Table 3.3. Hit samples selected from the marine natural product extract screen.....	78
Table 3.4. Inhibition values for hit samples selected from the marine natural product extract screen. ....	79
Table 3.5. Percent inhibition of glycosidase activity by selected hits from natural product extract screening. ....	81
Table 3.6. Structural assignment of <b>3.2</b> . ....	90
Table 3.7. Structural assignment of <b>3.4</b> . ....	97
Table 3.8. Inhibition values of isolated natural products and their analogues.....	100
Table 3.9. Inhibition values for <i>most inhibitory</i> hits from ircinialactam library versus TcTS.....	104
Table 3.10. Inhibition values for ircinialactam library hits <b>3.10-3.21</b> . ....	109
Table 4.1. First-generation target sequences from RaPID screening.....	126
Table 4.2. Hydroxylated non-peptide based inhibitors of TcTS.....	130
Table 4.3. Second-generation target sequences from RaPID screening. ....	135
Table 4.4. Synthetic methods explored for synthesis of <b>4.24</b> . ....	139

Table 4.5. Coupling conditions for manual SPPS of <b>4.20</b> and <b>4.21</b> .....	144
Table 4.6. Aqueous cyclization conditions explored with linear peptides and DSG. ....	147
Table 4.7. Reaction conditions tested for-collidine mediated coupling between <b>4.31</b> and DSG.....	149
Table 4.8. Inhibition values for linear and cyclic peptides versus TcTS. ....	152
Table 5.1. Comparison of TcTS crystallization buffer components. ....	163
Table 5.2. Comparison of <i>in vitro</i> (IC <sub>50</sub> ) and <i>in vivo</i> (EC <sub>50</sub> ) values for ircinialactam analogues. ....	168
Table 7.1. Inhibition values ( $K_i$ ) and fitting parameter ( $\chi^2$ ) from GraFit 7.0 fitting software.....	222
Table 7.2. Occurrence (n) and calculated extinction coefficients ( $\epsilon_{\text{total}}$ ) for peptides by UV/Vis spectroscopy at A <sub>280</sub> nm. ....	236
Table 7.3. Calculated concentration of peptide stocks by UV/Vis spectroscopy at A <sub>280</sub> nm.....	237

## List of Figures

Figure 1.1. Global expansion and prevalence of Chagas disease. ....	1
Figure 1.2. <i>Trypanosoma cruzi</i> life cycle. ....	3
Figure 1.3. Approved drugs for Chagas disease. ....	3
Figure 1.4. Structure of Neu5Ac and enzymatic processing of sialoglycoconjugates. ....	5
Figure 1.5. Role of <i>T. cruzi</i> trans-sialidase in biological systems. ....	7
Figure 1.6. Active site interactions of Sia-Lac with TcTS D59A mutant. ....	8
Figure 1.7. Structure of select sialidase transition state analogues. ....	11
Figure 1.8. Carbohydrate-based inhibitors of <i>T. cruzi</i> trans-sialidase. ....	14
Figure 1.9. (a) General structure of glycosidase-targeting mechanism-based inhibitors (b) General mechanism for fluoroglycosyl fluorides (c) Structure of fluoroglycosyl fluorides <b>1.20</b> and <b>1.21</b> . ....	15
Figure 1.10. Mechanism based inhibitors of TcTS. Second-order rate constants for inactivation ( $k_i/K_i$ ) and reactivation ( $k_r^{lact}/K_d^{lact}$ ) with lactose for DFSA inactivators. ....	17
Figure 1.11. Lead inhibitors discovered by <i>T. cruzi</i> trans-sialidase screening approaches. ....	21
Figure 2.1. Substituted DANA derivatives as selective hNEU inhibitors. ....	30
Figure 2.2. Crystal structure of covalent 3-fluorosialosyl fluoride-enzyme complex with hNEU2 (left) and TcTS (right). ....	31
Figure 2.3. X-ray structure of covalent intermediate between <b>1.22</b> (BFN) and TcTS (1S0I in grey). ....	33
Figure 2.4. C7 modified DFSA analogues as inactivators of TcTS. ....	36
Figure 2.5. Biosynthesis of CMP-Neu5Ac in bacteria and mammals. ....	37
Figure 2.6. <i>N</i> -modified Neu5Ac analogues biosynthesized and assayed for effect on <i>T. cruzi</i> infection. ....	38
Figure 2.7. C5-modified DFSA analogues as inactivators of bacterial sialidases. ....	39
Figure 2.8. 9-azido-DFSA ( <b>2.26</b> ) structure and design features. ....	40
Figure 2.9. <sup>19</sup> F-NMR of Neu5Ac aldolase reaction progress after 6 hours. ....	44

Figure 2.10. Determination of inactivation rate constants ( $k_{obs}$ ) (left) and second-order inactivation rate constant ( $k_i/K_i$ ) (right) for <b>2.55</b> versus TcTS.....	54
Figure 2.11. Determination of inactivation rate of <b>2.68</b> versus TcTS.....	59
Figure 3.1. FDA Approved drugs 1981-2014. ....	64
Figure 3.2. Major classes of alkaloid-based glycosidase inhibitors.....	65
Figure 3.3. Marine natural products as FDA approved drugs.....	67
Figure 3.4. Structures and antitrypanosomal activity of select marine natural products. ....	68
Figure 3.5. Marine natural products as sialidase inhibitors. ....	69
Figure 3.6. Structural comparison of substrate and natural carbohydrate-based inhibitors of $\alpha$ -glucosidase (left), $\alpha$ -amylase (bottom) and neuraminidase (right). ....	70
Figure 3.7. Natural product inhibitors of human pancreatic amylase. ....	72
Figure 3.8. Data set from two plate-based screens of UBC marine natural product extract library. ....	77
Figure 3.9. Percent inhibition of $\beta$ Gal from <i>aspergillus oryzae</i> (top) or TcTS (bottom) by samples from plate 4 of the marine natural product extract library.....	80
Figure 3.10. Percent inhibition of TcTS activity for sponge extracts from aqueous (Aq) or organic (Org) phase from liquid-liquid phase extraction.....	83
Figure 3.11. Extraction scheme and bioassay summary for sample 03-242.....	84
Figure 3.12. Extraction scheme and bioassay summary for sample 03-395.....	85
Figure 3.13. Marine sponge PNG-03-395 prior to extraction.....	86
Figure 3.14. Extraction scheme and bioassay summary for sample 03-395-2.....	87
Figure 3.15. Structures of natural products identified from sponge 03-395. ....	89
Figure 3.16. Structures of select members of the adociasulfate family. ....	92
Figure 3.17. Extraction scheme and bioassay summary for sample 00-287.....	93
Figure 3.18. Extraction scheme and bioassay summary for sample 00-287-2.....	94
Figure 3.19. Structures of the ircinialactam family (A-I). ....	96

Figure 3.20. Structures of isolated natural products and analogues.....	99
Figure 3.21. Structures of select members of the sesterterpene tetronic acid family.....	101
Figure 3.22. General structure of the ircinialactam library compounds.....	102
Figure 3.23. Ircinialactam library screening results for TcTS inhibition.....	103
Figure 3.24. Structures of select <i>inhibitory</i> hits from ircinialactam library.....	105
Figure 3.25. Structure (top) and percent inhibition (bottom) of TcTS by cacolides <b>3.22-3.36</b> .....	107
Figure 3.26. Time-dependent inhibition assay with select ircinialactam library hits. ....	108
Figure 4.1. General summary of phage display screening. ....	113
Figure 4.2. Role and structure of puromycin in connecting peptide to encoding mRNA strand.....	115
Figure 4.3. Structure of cyclic peptide-based natural products.....	117
Figure 4.4. General summary of peptide translation.....	118
Figure 4.5. (a) Flexizyme catalyzed attachment of non-proteinogenic amino acid to tRNA-GAA (b) genetic code (left) and reprogramed NNU genetic code (right). ....	121
Figure 4.6. Post-translational modification of non-proteinogenic amino acids (a) <sup>ClAc</sup> Trp or (b) Vinylglycine (Vgl) following site specific incorporation with the FIT system. ....	122
Figure 4.7. General summary of RaPID screening. ....	123
Figure 4.8. First-generation library design and cyclization strategy for RaPID screening. ....	125
Figure 4.9. Percent inhibition of TcTS by cyclic peptides <b>4.1-4.9</b> . ....	127
Figure 4.10. Second-generation library design and cyclization strategies for RaPID screening. ....	133
Figure 4.11. Reaction progress (TLC) and reaction scheme for <i>t</i> -butyl ester formation during synthesis of <b>4.26</b> . ....	138
Figure 4.12. General summary of solid phase peptide synthesis (SPPS).....	142
Figure 4.13. Structure of linear peptides: <b>4.20</b> and ‘natural’ derivatives <b>4.29-4.30</b> (left); <b>4.21</b> and ‘natural’ derivatives <b>4.31-4.32</b> (right).....	145

Figure 4.14. Reaction scheme (top) and progress (bottom) of amino acids (Lys, AMB and Dopa) with DSG in the presence of varying organic bases.....	148
Figure 4.15. Percent inhibition of TcTS by peptides <b>4.20-4.21</b> ; <b>4.29-4.38</b> .....	151
Figure 5.1. SDS-PAGE of HisTrap (left) and MonoQ (right) column purification following <i>T. cruzi</i> trans-sialidase expression.....	157
Figure 5.2. SDS-PAGE of combined TcTS-containing fractions from each column purification (left). UV/Vis spectrum of the elution profile from SEC purification (right).....	158
Figure 5.3. Hanging drop vapour diffusion crystallization preparation (left) and incubation (right). ....	159
Figure 5.4. Light microscopy image of TcTS crystals (1:1 TcTS (15 mg/mL)/Buffer A). ....	159
Figure 5.5. Light microscopy images of TcTS crystals. (a) 2:1 TcTS (35 mg/mL)/Buffer A (b) 2:1 TcTS (20 mg/mL)/Buffer A (c) 1:2 TcTS (20 mg/mL)/Buffer A.....	160
Figure 5.6. Light microscopy images of TcTS crystals. (a) 3:0.5 TcTS (20 mg/mL)/Buffer A (b) 2:0.5 TcTS (20 mg/mL)/Buffer A (c) 1:2 TcTS (20 mg/mL)/Buffer A.....	161
Figure 5.7. Percent inhibition of <i>T. cruzi</i> growth and inhibitor structure for <b>3.15</b> (left) and benznidazole (right). ....	166
Figure 5.8. Percent inhibition of <i>T. cruzi</i> growth and inhibitor structures for (a) <b>3.10</b> (b) <b>3.12</b> (c) <b>3.13</b> (d) <b>3.14</b> (e) <b>3.15</b> (f) <b>3.16</b> . ....	167

## List of Schemes

Scheme 1.1. Catalytic mechanism of <i>T. cruzi</i> trans-sialidase. ....	9
Scheme 2.1. Synthesis of 9-amido-DFSAs. ....	41
Scheme 2.2. Enzymatic coupling of 3-fluoropyruvate and 6-azido-ManNAc. ....	43
Scheme 2.3. Synthesis of 3-oxabicyclo[3.2.1]octane-2,4-dione ( <b>2.46</b> ). ....	45
Scheme 2.4. Synthesis of 5-amido-DFSAs from Neu5Ac. ....	49
Scheme 3.1. Methylation of carboxylic acid group by reaction with diazomethane. ....	95
Scheme 3.2. Saponification of ircinialactam <b>3.4</b> . ....	98
Scheme 4.1. Biosynthetic preparation of Sia-Lac-MU ( <b>4.11</b> ). ....	128
Scheme 4.2. Comparison of TcTS assays using a (top) one-enzyme or (bottom) two-enzyme system. ..	128
Scheme 4.3. Post-translational modification of cysteine during RaPID screening. ....	134
Scheme 4.4. Proposed (a) and executed (b) synthesis of <b>4.24</b> . ....	136
Scheme 4.5. Cyclization of linear peptides with DSG and collidine. ....	150

## List of Abbreviations

[E]: Enzyme concentration

[E]<sub>0</sub>: Initial enzyme concentration

[I]: Inhibitor concentration

[S]: Substrate concentration

Abg: *Agrobacterium* sp.  $\beta$ -glucosidase

Ac<sub>2</sub>O: Acetic anhydride

AcOH: Acetic acid

AMB: 4-(Aminomethyl)benzoic acid

Boc: *tert*-Butyloxycarbonyl

BSA: Bovine serum albumin

CAZy: Carbohydrate active enzyme

CPRG: Chlorophenol red- $\beta$ -D-galactopyranoside

CMP-SA: Cytidine-5'-monophospho-*N*-acetylneuraminic acid

COSY: Correlation spectroscopy

CPdCA: Cyclopentylidicarboxylic acid

Cys\* (C\*): *S*-(3,5-Dihydroxybenzyl)-cysteine

DANA: 2-Deoxy-2,3-dehydro-*N*-acetylneuraminic acid

DAST: Diethylaminosulfur trifluoride

DBAA:  $\alpha,\alpha'$ -Dibromoadipic-bis-amide

DBU: 1,8-Diazabicyclo(5.4.0)undec-7-ene

DCM: Dichloromethane

DFSA: Difluorosialic acid

DIPEA: Diisopropylethylamine

DMAP: 4-Dimethylaminopyridine

DMF: Dimethylformamide

DMSO: Dimethyl sulfoxide

Dopa (d): 3,4-Dihydroxyphenylalanine

DSG: Disuccinimidyl glutarate

EC<sub>50</sub>: Half-maximal effective concentration

EDC: 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

EDTA: Ethylenediaminetetraacetic acid

ESI-TOF: Electrospray ionization – time of flight

EtOAc: Ethyl acetate

EtOH: Ethanol

FIT: Flexible *in vitro* translation

Fmoc: Fluorenylmethyloxycarbonyl group

Gal: Galactose

Gal- $\alpha$ -F:  $\alpha$ -D-Galactopyranosyl fluoride

GH: Glycoside hydrolase

Glc: Glucose

Glc- $\beta$ -MU: 4-Methylumbelliferyl  $\beta$ -D-glucopyranoside

GPI: Glycosylphosphatidylinositol

HBTU: 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HEPES: 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid

HMBC: Homonuclear multiple bond correlation

HOBt: Hydroxybenzotriazole

hNEU: Human neuraminidase

HPA: Human pancreatic amylase

HPLC: High performance liquid chromatography

HSQC: Heteronuclear single quantum coherence

HTS: High throughput screen

IC<sub>50</sub>: Half-maximal inhibitory concentration

IPTG: Isopropylthiogalactoside

K<sub>d</sub>: Dissociation constant

k<sub>i</sub>: Inactivation rate

K<sub>i</sub>: Inhibition constant

k<sub>i</sub>/K<sub>i</sub>: Second-order inactivation rate constant

K<sub>m</sub>: Michaelis-Menten constant

k<sub>obs</sub>: Observed pseudo-first order inactivation rate constant

k<sub>r</sub>: Reactivation rate

KIE: Kinetic isotope effect

Lac: Lactose

Lac-β-MU: 4-Methylumbelliferyl β-D-lactopyranoside

LacNAc: *N*-Acetyl-lactosamine

LB: Lysogeny broth

MALDI-TOF: Matrix-assisted laser desorption/ionization – time of flight

ManNAc: *N*-Acetylmannosamine

MbA: Montbretin A

MBI: Mechanism-based inhibitor

MeOH: Methanol

MS: Mass spectroscopy

Mtt: 4-Methyltrityl

MU: Methyl umbelliferone

NANA: *N*-Acetyl-neuraminic acid  
Neu5Ac: *N*-Acetyl-neuraminic acid  
NHS: *N*-Hydroxysuccinimide  
NMM: *N*-Methylmorpholine  
NMR: Nuclear magnetic resonance  
PDB: Protein data bank  
PEG: Polyethylene glycol  
PPh<sub>3</sub>: Triphenylphosphine  
Pyr: Pyridine  
QM: Quinone methide  
RaPID: Random nonstandard peptide integrated discovery  
ROESY: Rotating frame overhauser effect spectroscopy  
RPM: Rotations per minute  
RT: Room temperature  
SA: Sialic acid  
SAR: Structure-activity relationship  
SDS-PAGE: Sodium dodecyl sulfate - polyacrylamide gel electrophoresis  
Selectfluor: 1-Chloromethyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octane bis(tetrafluoroborate)  
Sia: Sialic acid  
SPPS: Solid phase peptide synthesis  
TBAF: Tetra-*n*-butylammonium fluoride  
tBu: *tert*-Butyl  
TBDPS: *tert*-Butyldiphenylsilyl  
TcTS: *Trypanosoma cruzi* trans-sialidase  
TEA: Triethylamine

Tf<sub>2</sub>O: Triflic anhydride  
TFA: Trifluoromethanesulfonic anhydride  
TFMU: 4-Trifluoromethylumbelliferone  
TFMU-SA: 4-Trifluoromethylumbelliferone sialoside  
THF: Tetrahydrofuran  
TLC: Thin layer chromatography  
TMS-N<sub>3</sub>: Trimethylsilyl azide  
TMSOTf: Trimethylsilyl trifluoromethanesulfonate  
Tris: Tris(hydroxymethyl)aminomethane  
UV/Vis: Ultraviolet/visible  
V<sub>max</sub>: Maximal velocity  
v<sub>o</sub>: Initial rate  
vNEU: Viral neuraminidase

**Standard Amino Acids:**

Ala (A): Alanine  
Arg (R): Arginine  
Asn (N): Asparagine  
Asp (D): Aspartic acid  
Cys (C): Cysteine  
Gln (Q): Glutamine  
Glu (E): Glutamic acid  
Gly (G): Glycine  
His (H): Histidine  
Ile (I): Isoleucine

Leu (L): Leucine

Lys (K): Lysine

Met (M): Methionine

Phe (F): Phenylalanine

Pro (P): Proline

Ser (S): Serine

Thr (T): Threonine

Trp (W): Tryptophan

Tyr (Y): Tyrosine

Val (V): Valine

## Acknowledgements

First and foremost I would like to thank my supervisor, Dr. Stephen Withers, for his unconditional support and guidance throughout my degree. I am forever grateful for his mentorship, patience and kindness in teaching me to become the scientist I am today.

Thank you to each and every member of the Withers group, past and present, who have created a fun and supportive community that has made coming to lab every day worthwhile. A special thanks to Andrés for being my mentor and my role model in science, to Spence for partnering in our joint business(ish) adventures, to Seyed and Phil for your absolute support and willingness to indulge my many questions, and to Hongming, Emily, Feng, Lyann and Miriam for your expertise and guidance. I am very grateful to the amazing staff in the Chemistry Shops and Services who make research in the department possible, a special thanks to Maria, Elena, Yun, Milan and David for your above and beyond support.

I would like to acknowledge ECOSCOPE, specifically Dr. Steven Hallam and Jennifer Bonderoff, for the funding and support. Thank you to the excellent scientists that I was fortunate to collaborate with: Dr. Seino Jongkees and Dr. Hiroshi Suga at the University of Tokyo; Dr. Robert Capon at the University of Queensland; Dr. Fred Buckner at the University of Washington; Dr. David Williams and Dr. Raymond Andersen at the University of British Columbia.

Thank you to Val, Rebecca, Chris and Danielle for being with me every step of the way, for the countless coffee walks, long lunches, occasional forays into exercise and many swift ones. I cannot articulate how much your unwavering support and friendship has meant to me. Thank you to Karen for being my rock, always being by my side no matter where in the world you are and for motivating me to be the best I can be.

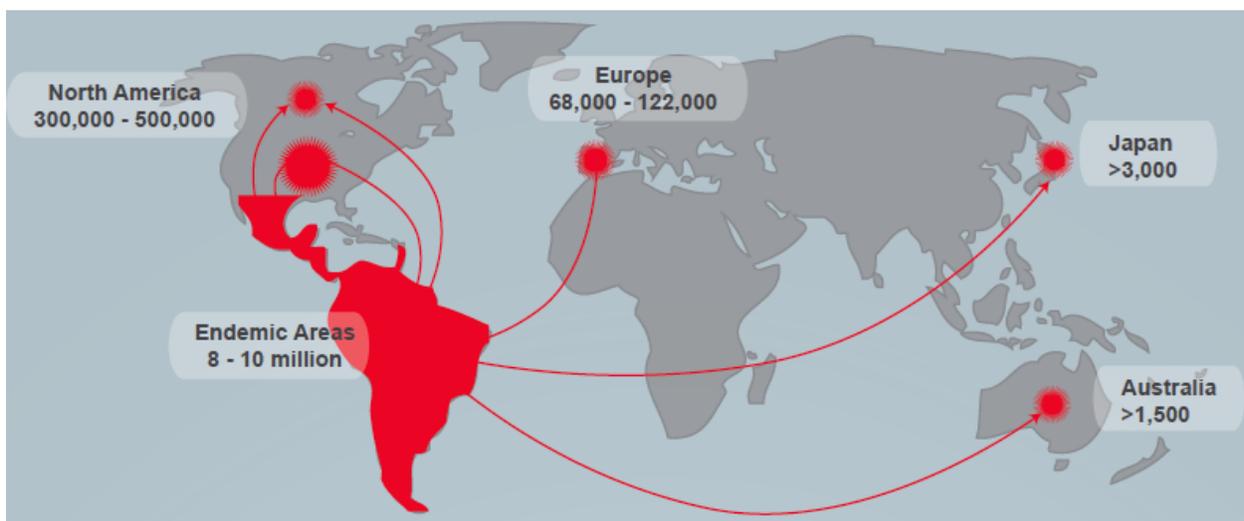
Finally, to my family who have been a source of constant support and motivation to pursue my passions and my dreams, thank you.

To my parents, Ken and Heather.

# Chapter 1: Introduction

## 1.1 Chagas disease

Chagas disease (*i.e.* American trypanosomiasis) is caused by a tropical parasite that is endemic to Latin America. The disease has expanded globally over the past decade and currently affects 8-10 million people worldwide, incurring an estimated annual cost >\$7 billion<sup>1</sup> (Figure 1.1). The World Health Organization (WHO) has classified Chagas on their list of *Neglected Tropical Diseases*, acknowledging that an increase in efforts and resources is required to combat this potentially life-threatening illness. A protozoan parasite – *Trypanosoma cruzi* – is the causal agent of the disease and is most often transmitted via infected blood-sucking insect carriers (*e.g.* triatomine bug) which reside in the cracked walls of homes throughout the impoverished areas of Latin America. Alternatively, the disease can be spread via blood transfusions, organ transplantations, or from mother to fetus.



**Figure 1.1. Global expansion and prevalence of Chagas disease.** Source: Martinez<sup>2</sup>

Following parasite infection, the disease manifests in two stages: the acute phase and the chronic phase. The first 12-18 week period is called the *acute phase*, when circulating parasite can be detected in the bloodstream and when the primary symptoms may be observable but are rarely detected. These are

typically unspecific fever-like symptoms, which make Chagas difficult to diagnose, although a small percentage of infected patients do present with epidermal lesions or eyelid swelling which can be suggestive of *T. cruzi* infection. If the disease goes untreated the *chronic phase* will ensue whereby the parasite buries into the cardiac and digestive tissues, where it remains for the duration of the host's lifetime. Clinical manifestations will become evident in ~40% of infected patients during this phase and may lead to severe organ damage or failure, with an associated mortality rate of 10-20%<sup>3</sup>.

## 1.2 *Trypanosoma cruzi*

*Trypanosoma cruzi* exists in four physiologically specialized forms throughout its life cycle (Figure 1.2): the infectious forms (1) metacyclic trypomastigote and (2) trypomastigote and the reproductive forms (3) amastigote and (4) epimastigote. In the hindgut of triatomine bugs, epimastigotes reproduce and transform into the infectious metacyclic trypomastigote. Following a blood meal the infected insect will defecate near the wound allowing the parasite to enter the new mammalian host bloodstream through broken skin or mucosal membranes<sup>4</sup>. The parasite will immediately invade nearby tissues and differentiate into amastigotes within the host cells where they will replicate by binary fission. Finally, the parasite transforms into the highly infectious trypomastigote and re-enters the bloodstream where it can invade new host cells to further propagate the infectious cycle.

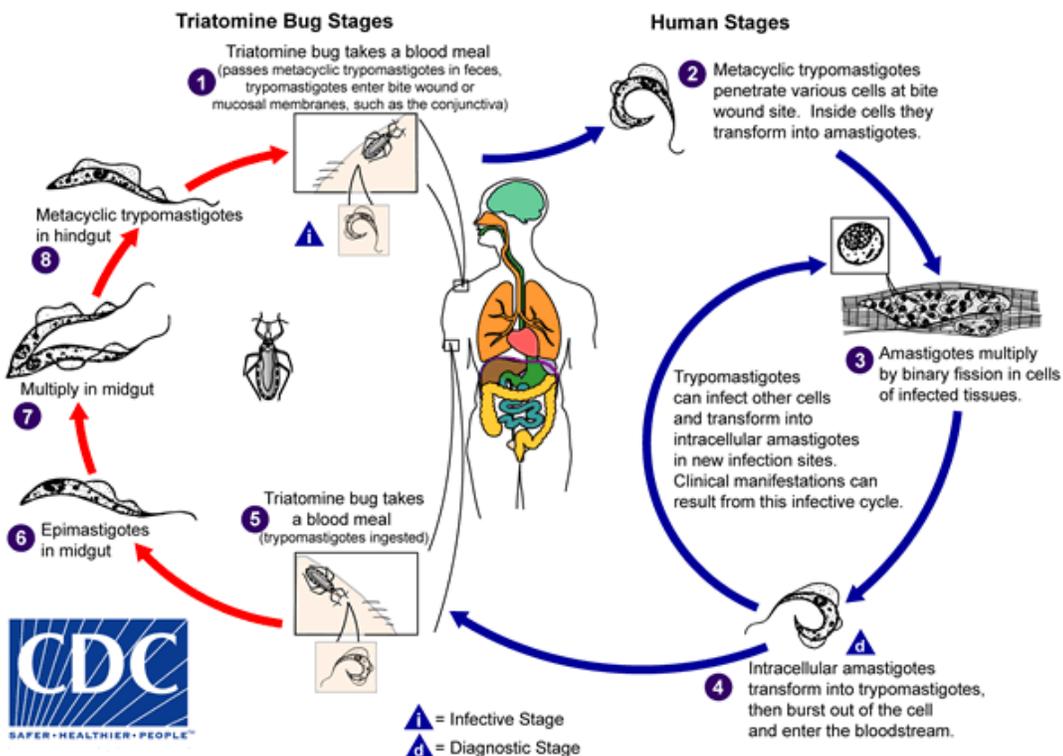
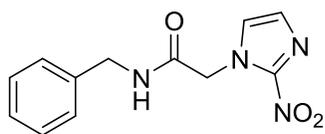
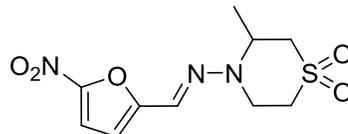


Figure 1.2. *Trypanosoma cruzi* life cycle. Source: www.cdc.gov<sup>5</sup>

Currently there are two approved treatments for Chagas disease<sup>6</sup>: benznidazole (Roche; **1.1**) and nifurtimox (Bayer; **1.2**) (Figure 1.3). Both are orally available therapeutics demonstrating modest acute phase antiparasitic efficacy and moderate to severe side-effects with prolonged usage.



Benznidazole  
**1.1**



Nifurtimox  
**1.2**

Figure 1.3. Approved drugs for Chagas disease.

It is generally accepted that broad-spectrum nitroheterocycle-containing antibiotic or antiparasitic compounds, such as **1.1** and **1.2**, act as prodrugs that generate reactive radical species upon enzyme-

mediated activation by nitroreductase within the pathogen<sup>7</sup>. Oxygen-insensitive (Type I) nitroreductases – such as the essential bacteria-like enzyme identified in the *T. cruzi* genome<sup>8</sup> – catalyze a series of two-electron reductions of the nitro-group to afford products that can preferentially damage the cellular and genetic material of the parasite<sup>9,10</sup>.

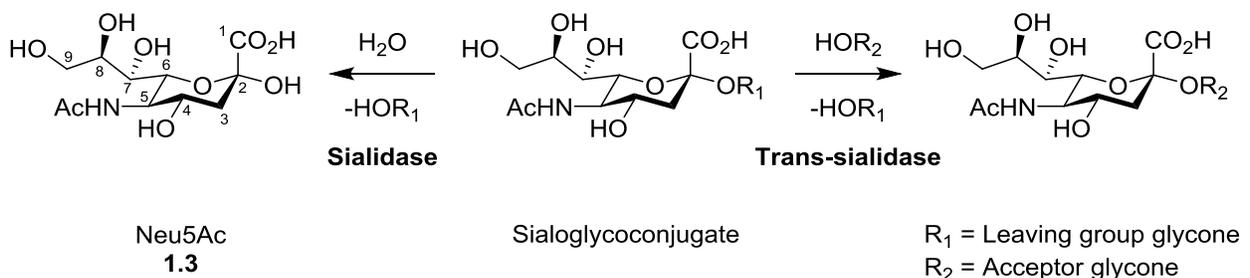
Benznidazole (**1.1**) is a nitroimidazole derivative that has reported antitrypanosomal efficacy of >60% during the acute phase of infection, making it the first line treatment for the disease<sup>11,12</sup>. It has been shown to interfere with *T. cruzi* RNA and protein synthesis<sup>13</sup> by generating double stranded breaks and nucleotide oxidation<sup>14</sup>. Side effects from the treatment are widely reported with regular use often leading to abdominal pain, nausea, vomiting and diarrhea. Nifurtimox (**1.2**), a nitrofurantoin derivative, demonstrates similar efficacy to benznidazole during the acute phase but has a more problematic range of side effects including gastrointestinal symptoms, anorexia, depression, rash and anxiety – making it the second line treatment for Chagas disease<sup>15</sup>. Due to the lengthy treatment periods, poor tolerance and lack of efficacy during the chronic phase, new therapies are required.

The limitations of on-market drugs and lack of specific, validated therapeutic targets has sparked research interests to this end. Several targets involved in essential parasite pathways have been suggested and lead inhibitors identified: K-11777 (*i.e.* *N*-methyl-piperazinyl-L-phenylalanyl-L-homophenylalanyl-vinyl sulfone) for inhibition of cruzipain<sup>16,17</sup>, the essential cysteine protease for *T. cruzi* proliferation; bisphosphonates as trypanocidal entities that act through inhibition of the parasite's farnesyl pyrophosphate synthase<sup>18</sup>; and posaconazole as an ergosterol biosynthesis inhibitor<sup>19</sup>. The latter compound completed clinical trials but ultimately failed to be more effective than current treatments<sup>20,21</sup>. Another specific target that has received considerable interest is the *T. cruzi* trans-sialidase (TcTS) which is anticipated to play an essential role in modulating the glycosylation of both the host cell and parasite surface during the acute phase of Chagas infection. However, to date, inhibitor development for this target has been challenging.

### 1.3 *Trypanosoma cruzi* trans-sialidase

The sialic acids are a family of structurally diverse 9-carbon monosaccharides<sup>22</sup>, the most prevalent of which is *N*-acetylneuraminic acid (*i.e.* Neu5Ac; NANA; **1.3**) (Figure 1.4). Neu5Ac is an  $\alpha$ -keto sugar bearing a glycerol side chain, an *N*-acetyl moiety at C5 and a carboxylate at C1 that exists as a charged species under physiological conditions. Sialic acids are most often found in higher eukaryotes and some microbial pathogens<sup>23</sup> and are typically displayed as terminal residues bound to  $\beta$ -galactopyranosyl moieties of *N*- and *O*-linked glycans<sup>24</sup>. Tens of millions of sialic acids are present on a single cell and can present apparent concentrations as high as 10 mM on the cell-surface glycocalyx<sup>25</sup>. The anionic character of these sugars and their high prevalence function to mediate cell-cell and cell-molecule interactions<sup>26</sup>.

Sialic acid-producing organisms also express enzymes that process these sugars: sialidases (*i.e.* neuraminidases) and trans-sialidases (Figure 1.4). In animals, these enzymes play an essential role by regulating the surface sialic acid levels<sup>27</sup>. Additionally, pathogens which cannot synthesize sialic acid *de novo*, such as viruses,<sup>28</sup> bacteria,<sup>29</sup> and protozoa,<sup>30</sup> have evolved mechanisms to exploit the sialylglycophenotype present on host cells.

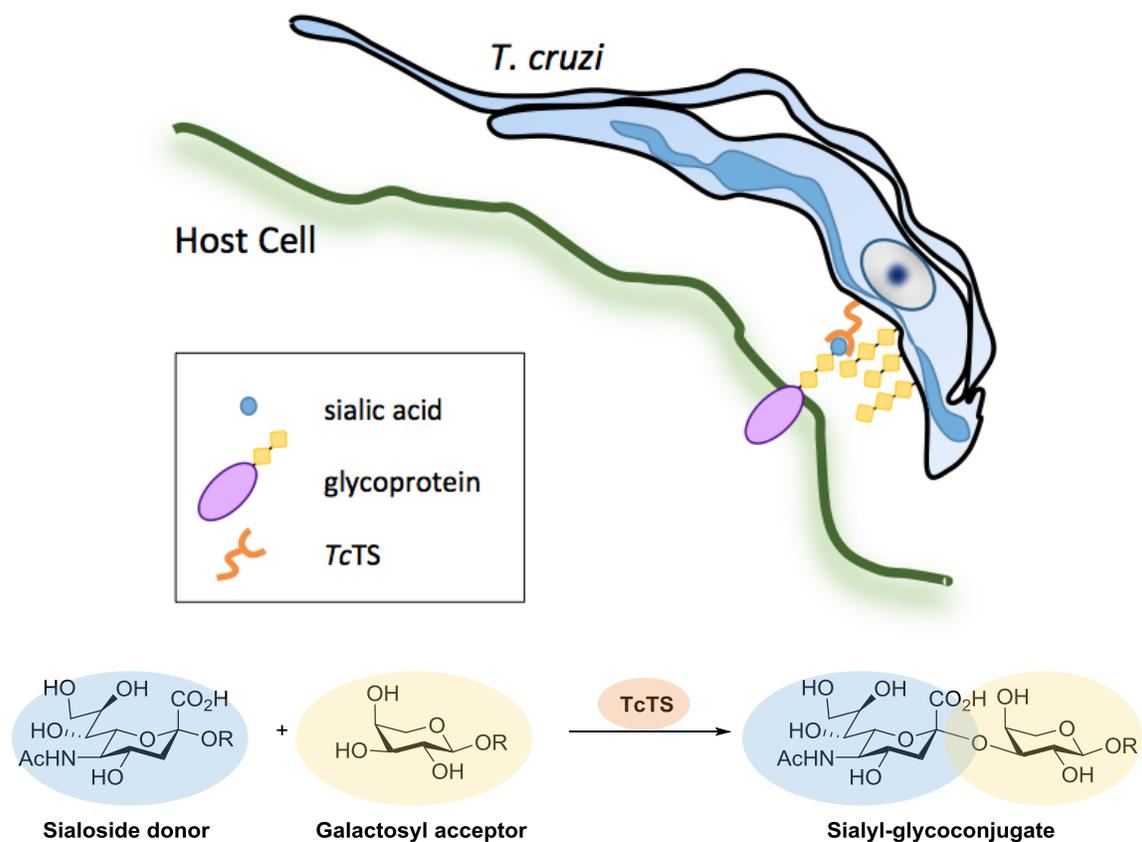


**Figure 1.4. Structure of Neu5Ac and enzymatic processing of sialoglycoconjugates.**

The protozoan parasite *T. cruzi* is one such example which has been reported to use its trans-sialylation activity to modulate host cell and parasite surface sialylation. TcTS is a GPI-anchored protein

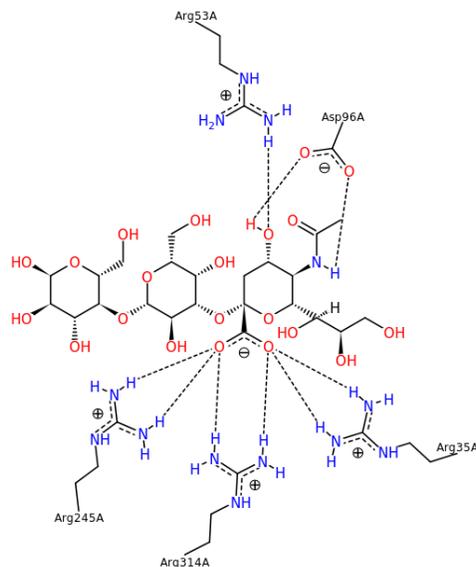
that is displayed on the parasite surface and/or excreted into the extracellular milieu. Since *T. cruzi* cannot produce sialic acid, it must scavenge the host sialyl-glycoconjugates using its trans-sialidase activity<sup>31</sup>. In Chagas-affected individuals, removal of sialic acid from the host surface and sialylation of the parasite surface are hypothesized to play two essential roles in the disease lifecycle: (1) sialylation of the parasite surface masks the pathogen from innate host immune recognition leading to reduced clearance and allowing for increased disease propagation<sup>32,33</sup> and (2) sialylation of trypomastigotes increases cell invasiveness<sup>34–36</sup>.

As such, TcTS is expected to play an important role during *T. cruzi* infection. Consequently, since its discovery by Pereira in the 1980s<sup>37</sup>, TcTS activity, structure, mechanism and function have been widely investigated. TcTS is a retaining glycosidase that catalyzes the transfer of  $\alpha$ -(2,3)-sialic acid from host glycoconjugates to the terminal  $\beta$ -galactopyranosyl moieties on the dense mucin-like coating of the parasite surface (Figure 1.5). TcTS belongs to GH 33 family in the sequence-based classification system of Carbohydrate Active enZymes (CAZy; [cazy.org](http://cazy.org)). The GH 33 family is composed of bacterial and eukaryotic exo- $\alpha$ -sialidases and trans-sialidases, known for their 6-fold beta-propeller structure and retaining mechanism of action in which the cleavage of the glycosidic bond occurs with a net retention of the anomeric stereochemistry. Unlike most retaining glycoside hydrolases, TcTS does not present a carboxylic nucleophile in the enzyme active site. Using fluoro-sugar inactivators, MS/MS analysis<sup>38</sup>, kinetic isotope effect (KIE) experiments<sup>39</sup> and protein crystallography<sup>40,41</sup>, researchers demonstrated that TcTS employs a double displacement mechanism by means of a nucleophilic tyrosine residue (Y342) which has since been shown to be highly conserved in exo-sialidases<sup>42</sup>.



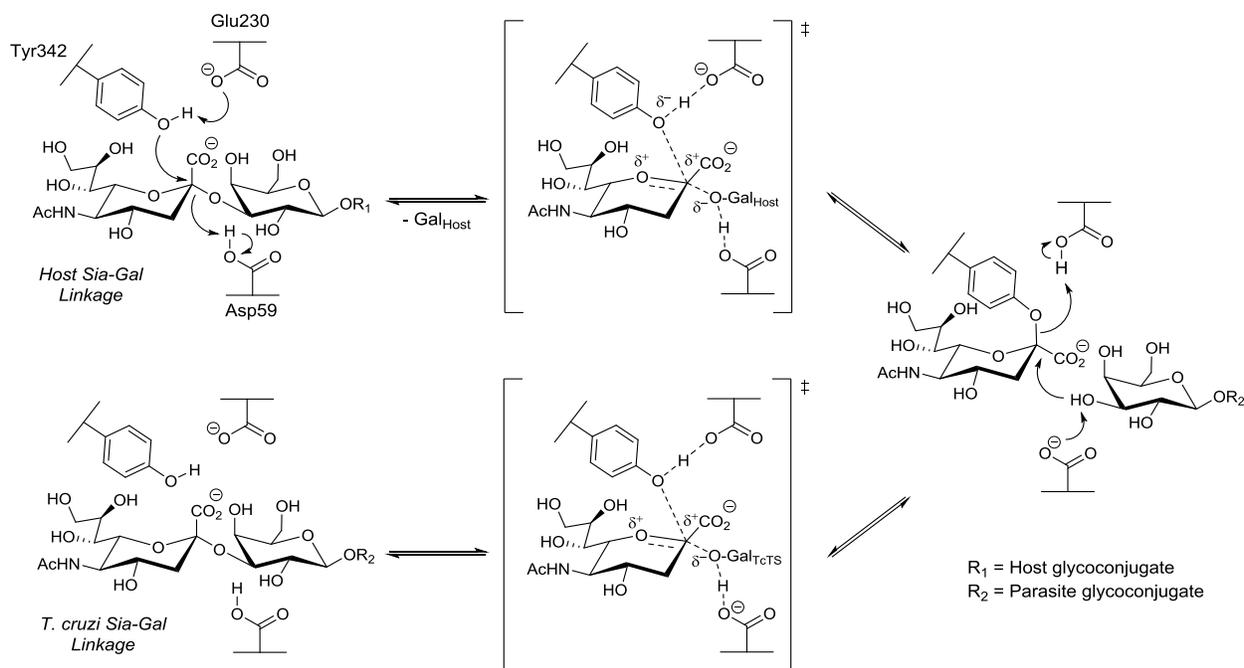
**Figure 1.5. Role of *T. cruzi* trans-sialidase in biological systems.**

Enzyme crystallography and mutagenesis studies have also identified D59 and E230 as essential residues in the catalytic cycle as well as several other key interactions in the donor and acceptor subsites (Figure 1.6): an arginine triad (R35, R245, R314) which interacts with the C1 carboxylate and a pair of hydrogen bond donors/acceptors (R53/D96) that interact with the C4 hydroxyl and C5 acetamide<sup>40</sup>. Additionally a pair of aromatic residues in the acceptor site (Y119/W312) are believed to confer flexibility of and control access to the active site cleft.



**Figure 1.6. Active site interactions of Sia-Lac with TcTS D59A mutant. Image from the RCSB PDB (rcsb.org) of PDB ID 1S0I<sup>41</sup>.**

The catalytic machinery of TcTS acts in concert to perform an enzymatic two-step ping-pong mechanism yielding the retention of anomeric stereochemistry<sup>43</sup> (Scheme 1.1). Specifically, E230 acts as a general base to facilitate nucleophilic attack by Y342 at the C2 centre of the terminal sialyl residue yielding a covalent sialyl-enzyme intermediate. The cleaved (a)glycone leaves the active site, allowing acceptor binding and D59 assisted nucleophilic attack on the sialyl-enzyme intermediate by the C3 hydroxyl group of the exogenous galactosyl moiety. Both the glycosylation and deglycosylation steps proceed via oxocarbenium-like transition states typical of the Koshland mechanism<sup>44</sup>. The nascent parasite-bound sialic acid has retained anomeric stereochemistry, while TcTS is again primed for the ensuing round of catalysis.



**Scheme 1.1. Catalytic mechanism of *T. cruzi* trans-sialidase.**

Effective therapeutics that bind this target could therefore reduce the ability of the parasite to invade host cells, expose their non-human antigens and allow for innate immune clearance of the pathogen. Attempts to validate this target through routine biological methods such as genetic knockout has been complicated by differential expression of TcTS during the *T. cruzi* life-cycle and the ~1400 copies of the TcTS gene found throughout the parasite genome<sup>34,45</sup>. Of those, an inactive mutant of TcTS – with a single amino acid substitution Y342H<sup>46</sup> – was found to occur frequently and acts as a carbohydrate-binding domain used to localize the parasite to the host cell surface. Consequently, the knockout of the enzyme's catalytic activity via small molecule modulation has been explored. Rational approaches – such as the design of carbohydrate-based inhibitors – have primarily been investigated to this end given the wealth of structural and mechanistic information available for TcTS.

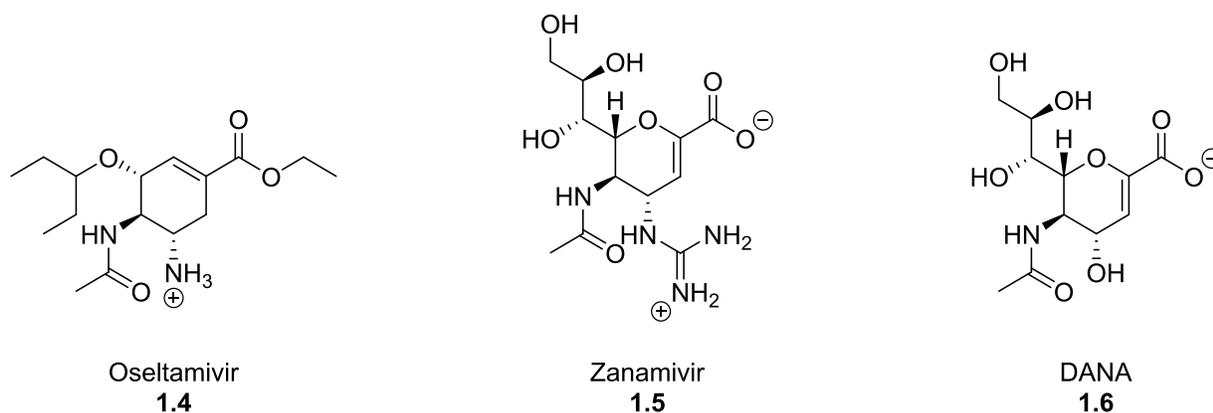
## 1.4 Carbohydrate-based inhibitors

Enzyme inhibitors can be generally divided into two categories: covalent and non-covalent. Non-covalent glycosidase inhibitors bind the active site reversibly and are typically used as the first-line approach in rational design towards well characterized enzyme targets. This process employs synthetically modified, catalytically inert, glycoside mimics that can retain endogenous active site binding interactions. In an ideal case these carbohydrate-based derivatives can be designed as substrate or transition-state mimics that can out compete the endogenous substrate for binding of essential enzyme residues. Alternatively, covalent glycoside inhibitors or suicide substrates bind irreversibly to the enzyme active site. Through appropriately placed reactive warheads these inhibitors covalently block the active site and render the target catalytically inactive. Each of these approaches has been extensively explored towards development of TcTS inhibitors.

### 1.4.1 Transition state analogues

The oxocarbenium ion-like transition state employed by retaining glycoside hydrolases – including TcTS – features partial positive character around the reaction center and increased planarity of the glycoside ring. Thus, transition-state analogues have been designed to mimic these features through the introduction of planarity via unsaturations and/or positively charged species (*e.g.* amines) onto or into the carbohydrate ring. This design strategy has been effectively employed against sialidase targets, highlighted by oseltamivir (*i.e.* Tamiflu; **1.4**) and zanamivir (*i.e.* Relenza; **1.5**), which target influenza neuraminidase and are the primary therapeutic agents for influenza infection (Figure 1.7). Though the structure of these compounds resemble features of the transition-state, Shidmoosavee et al. has rigorously demonstrated that Relenza instead mimics the enzyme-substrate interactions at the Michaelis complex – in the process describing the principle means of drug-resistance developed by the influenza virus<sup>47</sup>.

The broad scope sialidase inhibitor 2-deoxy-2,3-dehydro-*N*-acetyl-neuraminic acid (Neu2en5Ac; DANA; **1.6**) has shown micromolar inhibition against a variety of viral, bacterial and eukaryotic sialidases including influenza N2 neuraminidase ( $K_i = 4 \mu\text{M}$ ), *Trypanosoma rangeli* sialidase ( $K_i = 140 \mu\text{M}$ ) and *Salmonella typhimurium* sialidase ( $K_i = 380 \mu\text{M}$ ). Interestingly DANA exhibited negligible inhibitory activity when tested against TcTS ( $K_i = 12.3 \text{ mM}$ )<sup>48</sup>.



**Figure 1.7. Structure of select sialidase transition state analogues.**

#### 1.4.2 Substrate analogues

The TcTS active site consists of a donor binding site for sialic acid and an acceptor binding site for the  $\beta$ -galactopyranosyl unit of the parasite surface mucins; therefore, substrate analogues have been designed to harness interactions in one or both of these sites (Figure 1.8). Lactose or *N*-acetyl-lactosamine (LacNAc) is typically employed as a surrogate for the acceptor glycoside in laboratory settings. Lactitol (**1.7**) was investigated as an acceptor for TcTS trans-sialylation with a 3'-sialyllactose donor and it was discovered that **1.7** is a better acceptor ( $K_m = 0.26 \text{ mM}$ ) than LacNAc ( $K_m = 0.57 \text{ mM}$ ). In a competition experiment between **1.7** and LacNAc, **1.7** was shown to be preferentially sialylated, thus inhibiting trans-sialylation of LacNAc. *In vivo* testing showed that **1.7** inhibited parasite mucin sialylation, diminished *T. cruzi* infection by 20-30%<sup>49</sup> and prevented cellular apoptosis<sup>50</sup>. Modification of lactitol through addition

of linear and multi-arm polyethylene glycol (PEG) increased bioavailability<sup>51,52</sup>, while the oligosaccharide derivative penta-5ol (**1.8**) inhibited sialylation of LacNAc by synthetic mucin oligosaccharides with an  $IC_{50}$  value of 0.61 mM<sup>53</sup>. While interesting, the high concentrations needed for competition hindered therapeutic utility. More recently, a library of 46 C1 or C6 triazole-modified galactoside derivatives were synthesized using click chemistry in search of new interactions in the acceptor binding site. The best candidate (**1.9**) exhibited 37% inhibition of TcTS hydrolysis at 1 mM, while other members of the library (**1.10**) demonstrated trypanocidal activity in the hundreds of micromolar range (50% activity at 200  $\mu$ M). The authors attributed this observation to the cytotoxic nature of these derivatives<sup>54</sup>. More recently a library of ten aryl galactosides modified at C3 were synthesized, the best candidate – 4-methoxycarbonyl-2-nitrophenyl 3-*O*-carboxymethyl- $\beta$ -D-galactopyranoside (**1.11**) – exhibited 21% inhibition of TcTS enzymatic activity at 1 mM<sup>55</sup>. Thus far the efforts to design potent inhibitors targeting the TcTS acceptor site have been underwhelming.

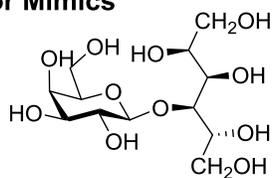
Donor mimics that emulate sialic acid binding modes have been similarly evaluated (Figure 1.8). Ferraro-Garcia and colleagues reported that the simple pyridoxal phosphate scaffold (**1.12**) demonstrated low millimolar inhibition against TcTS<sup>56</sup>. Subsequently, other minimalist hexose-like designs employing a six-atom cyclic core, bearing hydroxyl moieties, anionic charge (*e.g.* phosphate or carboxylate) and/or an acetamido group on the opposite side of the ring have been evaluated (**1.13**, **1.14**) and they exhibited comparable levels of enzyme inhibition<sup>57,58</sup>. An irreversible sialoside substrate NeuNAcFNP (**1.15**) was synthesized and designed such that the cleaved sialoside releases a highly reactive quinone methide (QM) aglycone<sup>59</sup>. The electrophilic QM product can indiscriminately react with surrounding nucleophiles resulting in covalent TcTS inactivation ( $IC_{50} = 0.57$  mM) as well as toxic off-target effects from non-specific reactivity with the cell. Comparison of this  $IC_{50}$  value to those of the competitive inhibitors is not appropriate since the covalent inhibition occurs in a time-dependent manner that is not accounted for in this parameter. *In vivo* these compounds demonstrated a 34% and 90% decrease in cellular parasite uptake

following pre-incubation of **1.15** with trypomastigotes at 1 mM and 10 mM respectively. However the non-specific reactivity of the covalent QM warhead makes these results difficult to assess.

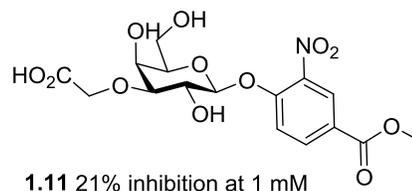
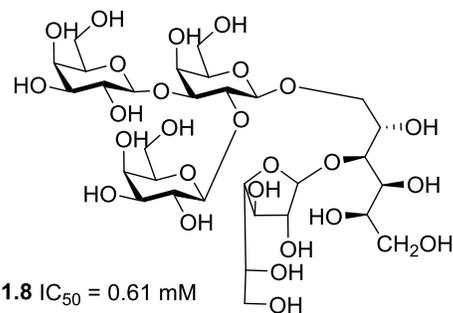
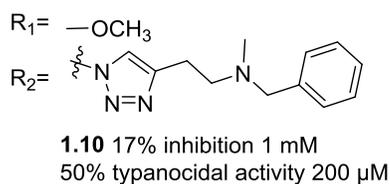
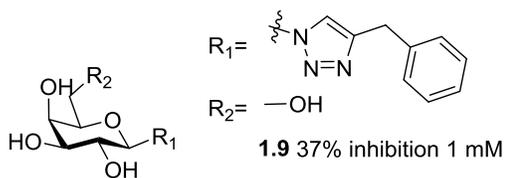
Carbohydrate analogues that are designed to harness interactions in both the donor and acceptor subsites generally result in better inhibitory efficiencies (Figure 1.8). In 2011 a series of sialoside-like C-glycosides bearing aryl aglycones was developed by Meinke and colleagues<sup>60</sup>. The best member of this series, **1.16**, bound with  $K_d = 160 \mu\text{M}$  and was hypothesized to interact with key acceptor residues Y119 and W312 based on the crystal structure of TcTS. In 2012 a series of triazole-linked sialyl-galactoside substrates were shown to inhibit with  $K_i$  values as low as  $260 \mu\text{M}$ <sup>61</sup> (**1.17**). Although some inhibitors targeting both donor and acceptor sites seem to provide enhanced enzyme binding, this is not universally applicable. A less successful report tested a set of pseudo-sialosides containing phosphonate-galactose linkages, the most promising of which (**1.18**) inhibited TcTS with low millimolar affinity ( $\text{IC}_{50} = 1.5 \text{ mM}$ )<sup>62</sup>. Further, the catalytically incompetent thioglycoside version of the ‘natural’ substrate, Sia-S- $\beta$ -Gal-*O*-octyl (**1.19**), did not exhibit observable inhibition at millimolar concentrations<sup>63</sup>. This was speculated to be a result of different conformational distributions relative to the *O*-linked version; similar to the arguments made for the differing affinities of cellobiose versus S-cellobiose with  $\beta$ -glucosidase from *Streptomyces* sp.<sup>64</sup>.

Libraries of modified donor and/or acceptor derivatives have yielded incremental increases in TcTS affinity culminating in **1.16** with  $K_d = 160 \mu\text{M}$ <sup>60</sup>. Evidence suggests a need for interactions with amino acids in both sites to achieve even moderate inhibitor affinity, although not all donor-acceptor designs have yielded the desired results. Overall, the potency of these derivatives remains orders of magnitudes away from the typical nanomolar values required for therapeutic relevance.

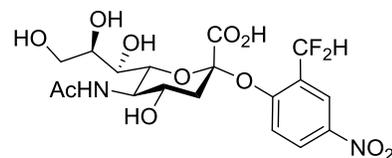
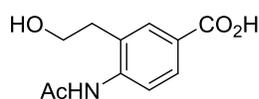
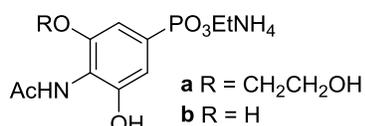
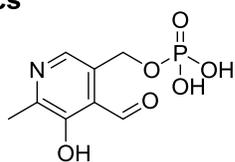
### Acceptor Mimics



1.7 20-32% decrease in cellular infection



### Donor Mimics



### Donor-Acceptor Mimics

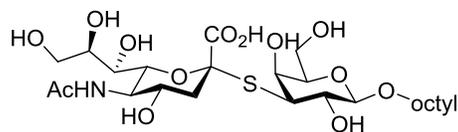
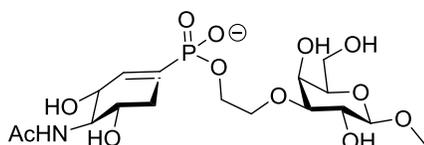
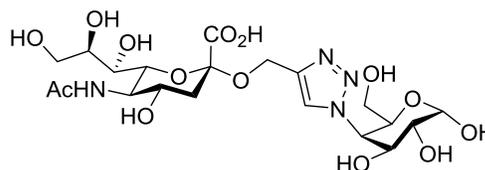
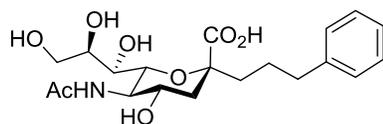
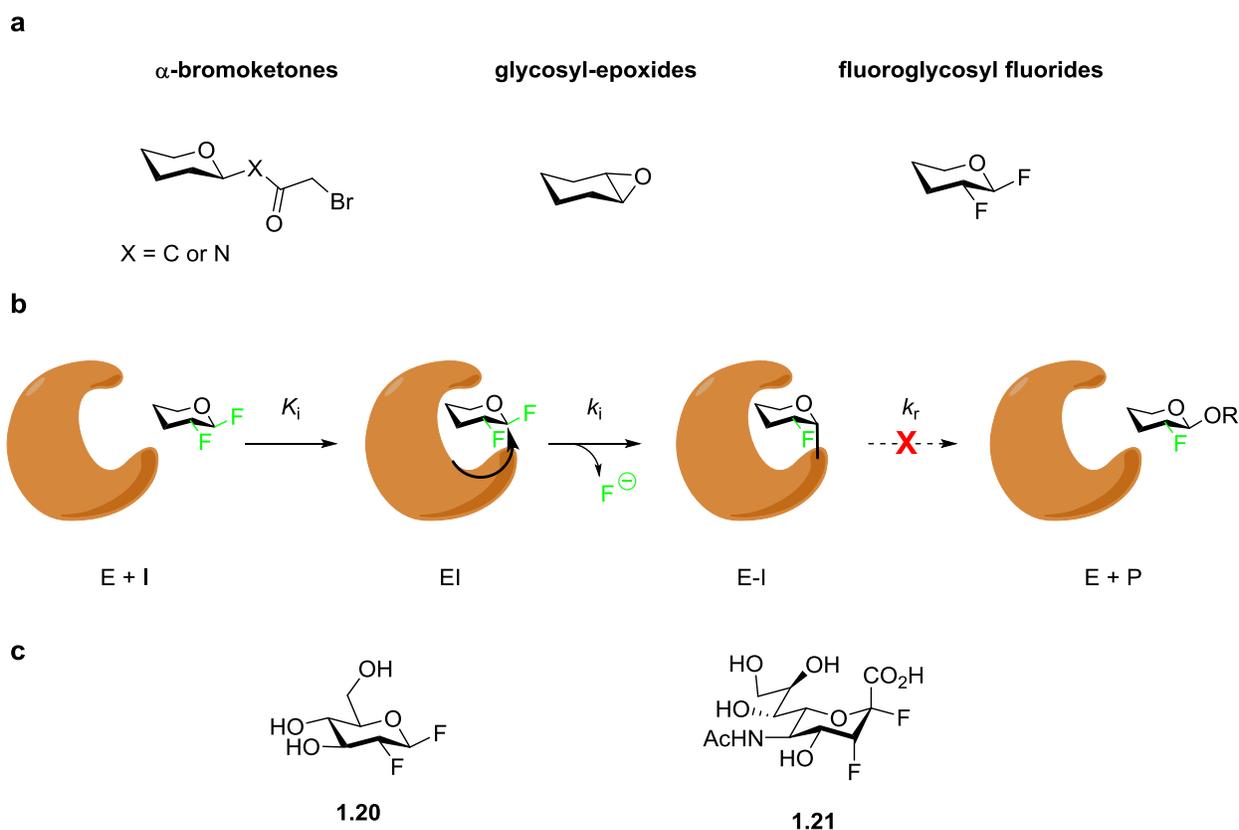


Figure 1.8. Carbohydrate-based inhibitors of *T. cruzi* trans-sialidase.

### 1.4.3 Mechanism-based inhibitors

Mechanism-based inhibitors (MBIs) are another class of rational carbohydrate-based inhibitors that have been explored. These substrate analogues utilize well-characterized enzyme mechanisms to covalently trap nucleophilic amino acids during the catalytic cycle rendering them inactive. Analogous to many competitive inhibitors, carbohydrate-based inactivators mimic the binding mode of the natural substrate but are designed such that they present an unnatural reactive group to the enzyme to induce their trapping mechanism. A few of the most common scaffolds used for inactivation of carbohydrate active enzymes include  $\alpha$ -bromoketone C-glycosides, glycosyl-epoxides and fluoroglycosyl fluorides<sup>65,66</sup> (Figure 1.9a).

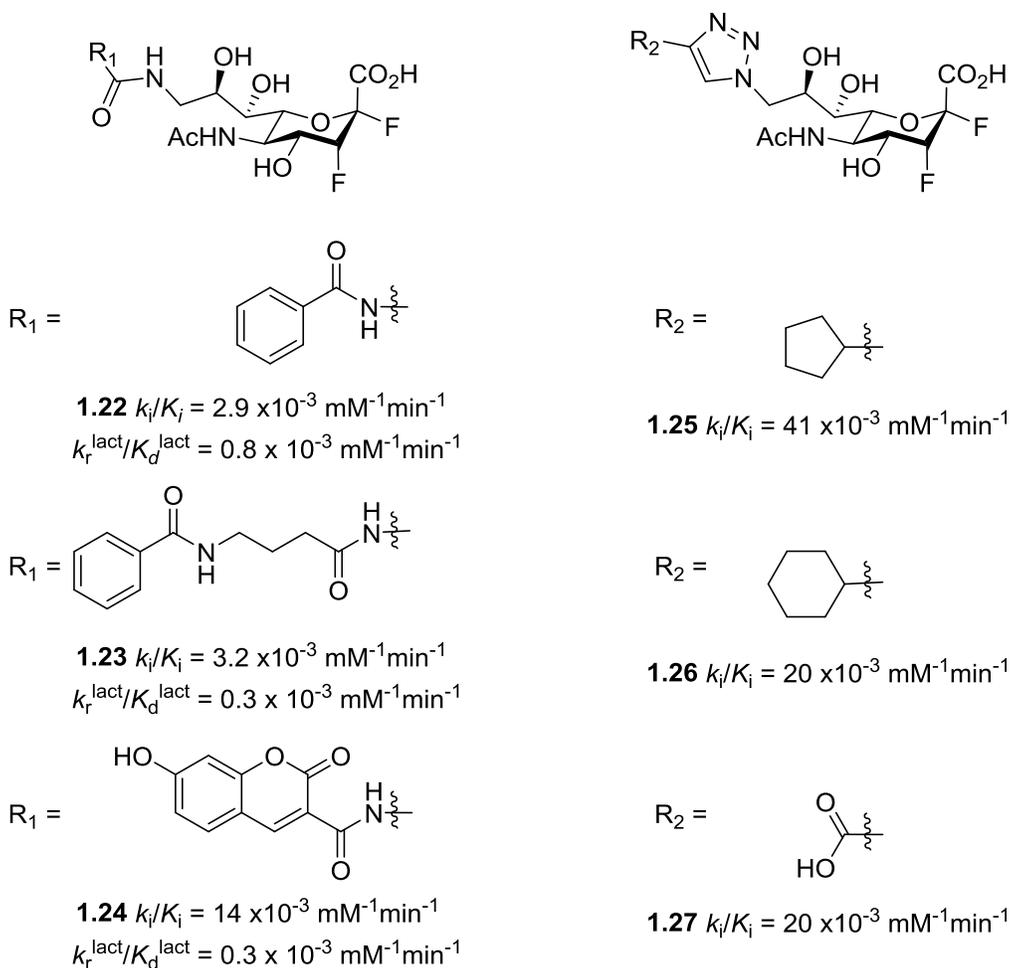


**Figure 1.9.** (a) General structure of glycosidase-targeting mechanism-based inhibitors (b) General mechanism for fluoroglycosyl fluorides (c) Structure of fluoroglycosyl fluorides 1.20 and 1.21.  $K_i$  is the inhibition constant,  $k_i$  is the inactivation rate,  $k_r$  is the turnover rate.

A successful method extensively employed by the Withers lab has been the use of fluoro-glycosyl fluorides (*e.g.* 2-deoxy-2-fluoro-, 5-fluoro- or 2-deoxy-2,2-difluoro-glycosides) as covalent inactivators of retaining glycoside hydrolases. These inactivators are designed such that a fluorine atom is chemically installed adjacent to the anomeric center of the glycoside target, while a second fluorine atom or aryl aglycone is installed at the anomeric centre. The seminal 1988 work by Withers showed that the modified glucoside substrate 2-fluoroglucosyl fluoride (**1.20**) could be partially processed by the GH 1  $\beta$ -glucosidase from *Agrobacterium* sp. (Abg) resulting in a covalently trapped fluoro-glucosyl enzyme species<sup>67</sup>. Abg employs the standard Koshland retaining mechanism which consists of an initial glycosylation event and a subsequent deglycosylation step to complete catalytic turnover. Substitution of the C2 hydroxyl moiety with fluorine inductively destabilizes the transition states of both catalytic steps slowing them substantially. The addition of a potent leaving group, such as fluoride, at the anomeric position sufficiently destabilizes the ground state of the substrate to facilitate the initial glycosylation event. However, the deglycosylation step (whose rate is independent of the aglycone) remains impaired, resulting in the inactivated 2-fluoroglucosyl enzyme species (Figure 1.9b). This strategy has been successfully translated into the sialic acid series to develop potent therapeutics<sup>68</sup> and titration agents<sup>69,70</sup> for influenza virus neuraminidase (Figure 1.9c).

3-Fluorosialosyl fluoride (*i.e.* DFSA; **1.21**) was first employed as an inactivator of TcTS to identify the catalytic nucleophile<sup>38</sup>. Subsequently DFSA has been examined as an inhibitory scaffold for therapeutic development. While DFSA indeed inactivated TcTS in the manner proposed, the inactivation efficacy was weak ( $k_i/K_i = 8 \times 10^{-3} \text{ mM}^{-1}\text{min}^{-1}$ ) and the fluorosialyl-enzyme complex turned over quickly in the absence of excess inactivator. Additionally DFSA does not exhibit any structural features that would confer specific TcTS binding in biological systems, thus simultaneous inactivation of human neuraminidases would be anticipated. In an attempt to increase its potency versus TcTS and its selectivity versus human neuraminidases Buchini and colleagues synthesized 9-amide-linked DFSA derivatives<sup>71</sup>

(**1.22-1.24**; Figure 1.10). Substitution at C9 by the addition of aryl amides yielded analogues with similar inactivation efficiencies to that of the parent DFSA compound, yet significantly reduced the rate of enzyme reactivation. Specifically the second-order rate constants for reactivation by lactose were decreased >1000-fold for compounds **1.22-1.24** ( $k_r^{\text{lact}}/K_d^{\text{lact}} < 1 \times 10^{-3} \text{ min}^{-1}\text{mM}^{-1}$ ) compared to the parent species **1.21** ( $k_r^{\text{lact}}/K_d^{\text{lact}} = 1330 \times 10^{-3} \text{ min}^{-1}\text{mM}^{-1}$ ). This effect has been attributed to a rearrangement of the inactivator's glycerol side chain into the TcTS acceptor site, where it interacts with Y119. This interaction – as demonstrated by crystallography – blocks entry to the active site by incoming acceptor molecules (*e.g.* lactose or water) leading to the longer-lived sialyl-enzyme complex.



**Figure 1.10. Mechanism based inhibitors of TcTS.** Second-order rate constants for inactivation ( $k_i/K_i$ ) and reactivation ( $k_r^{\text{lact}}/K_d^{\text{lact}}$ ) with lactose for DFSA inactivators.

In an attempt to replicate these interactions, while simultaneously searching for new ones, a diverse series of 9-triazole linked DFSA derivatives were synthesized using click chemistry and a central 9-azido-DFSA precursor<sup>72</sup>. Analysis of kinetic parameters for inactivation by a library of 12 substituted DFSAs showed that the 9-triazole(cyclopentyl)-DFSA (**1.25**) ( $k_i/K_i = 41 \times 10^{-3} \text{ min}^{-1} \text{ mM}^{-1}$ ) derivative inactivated TcTS three times faster than did the best 9-amido-DFSA derivative (**1.24**) and five times faster than DFSA (**1.21**). Overall, there was a preference for cyclic aliphatic derivatives (**1.25** and **1.26**) and some tolerance for anionic interactions (**1.27**), while polar, bulky, aromatic and cationic species exhibited decreased efficacy compared to **1.21**. Additionally, the conserved aryl triazole linkage did not decrease the reactivation rate in the same manner that was originally reported by Buchini. This can possibly be attributed to a lack of conformational flexibility of the planar triazole linkage, prohibiting the desired rearrangement into the acceptor binding pocket.

Overall, significant synthetic efforts have been invested into the development of rationally designed inhibitors of TcTS. The results have largely been underwhelming as the resultant inhibitors have yet to exhibit sufficient inhibition potency to be appropriately tested *in vivo*. Future inhibitor design should include development of compounds that simultaneously target interactions in both the donor and acceptor subsites (as the most potent inhibitor has) and continued investigations towards harnessing new active site interactions to increase binding efficiency and specificity. Additional efforts to find new inhibition scaffolds – such as screening approaches – could prove useful to this end.

## **1.5 Screening for enzyme inhibitors**

### **1.5.1 Introduction**

Rational inhibitor design is a logical starting point for information-rich targets in drug discovery and development processes. However these paths are typically synthetically arduous and low-throughput. When pursuing biological targets with no known small molecule modulators or when previous design attempts have been unsuccessful, high-throughput screening (HTS) approaches can yield new chemical

leads. HTS is a widely utilized technique that provides a way to quickly assess a large number of unique chemical species for a given biological function using automated technologies<sup>73</sup>. The scope and potential success of these screening approaches in a laboratory or industrial setting are largely dependent on two main factors: library selection and assay throughput.

The selection of an appropriate, high-quality library for the biological target is an essential aspect of the discovery process<sup>74</sup>. In practice, this is largely affected by a multitude of factors including laboratory goals and expertise, library access and availability, as well as assay throughput and overall cost. Common library types which may be used to discover new lead targets include curated synthetic commercial or homebrew compounds, characterized or crude extracts of natural products or peptide-based libraries. Additionally, for targets that have been characterized by crystallography, virtual screening and docking methods can be employed to generate theoretical lead compounds.

The nature of the library selected for screening will dictate the potential hits that can be obtained. For example, a pharmaceutical company that is searching for a lead chemical inhibitor for a new biological target may determine that in-house libraries of previously characterized compounds – with existing stability and toxicity profiles and favourable physicochemical properties – may best serve their purposes. On the other hand an academic laboratory that is searching for novel chemical entities to bind to an established biological target may have limited access to analogous commercial libraries and therefore chose to screen smaller diverse libraries of uncharacterized natural product extracts.

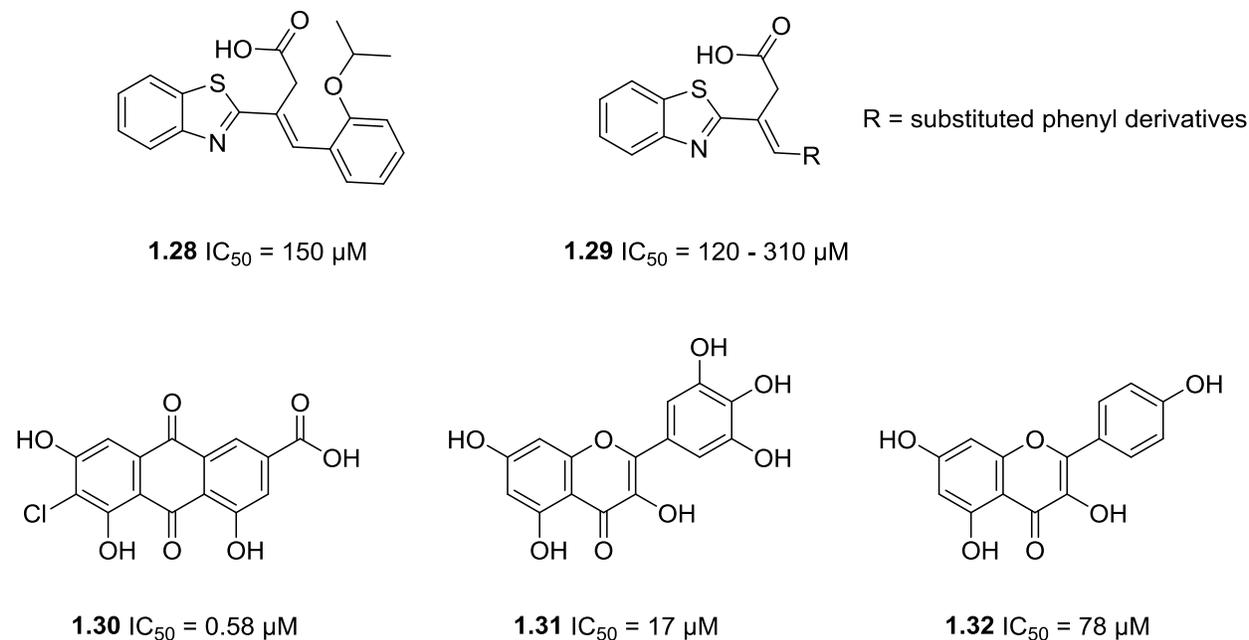
Once the desired library has been identified the development of an appropriate screening process is paramount. A central premise of the screen development process is that the biological assay must be relevant, reproducible, reliable and robust<sup>75</sup>. To this end each factor must be duly considered during assay design. To ensure the relevant biological assay will be scalable, plate-based (96-well or 384-well) or droplet-based systems are typically employed. Further, automated pipetting systems and high throughput detection methods (*e.g.* plate reader) assure standardization in the liquid handling and assay readout components. Finally, the reliability of the assay can be assessed by statistical analysis (*e.g.* Z-score<sup>76</sup>).

Optimization of these features can establish a robust assay that can be easily scaled to the library demands. Assay development is not only important for the initial discovery step but essential to the entire development process since the method used will guide selection and provide validation of lead compounds throughout.

While the selection of appropriate libraries and development of robust bioassays are the main strategies used to increase the likelihood of obtaining positive screening results, ultimately there is no way to guarantee success. Thus researchers must continue to investigate a wide variety of discovery methods and library sources in search of chemical leads for each unique biological target.

### **1.5.2 Screening for TcTS inhibitors**

Given the lack of success that was achieved by the standard approach (*i.e.* design and synthesis of carbohydrate-based inhibitors), discovery techniques such as chemical screening could provide an appealing alternative towards identifying new inhibitor scaffolds for the enzyme target. To date only two screens have been reported in search of TcTS inhibitors, both of which yielded lead compounds with superior inhibitory activity to those of the best known leads designed by rational methods (Figure 1.11).



**Figure 1.11. Lead inhibitors discovered by *T. cruzi* trans-sialidase screening approaches.**

In 2009, Neres and colleagues investigated the use of *in silico* screening to identify commercially available molecules as lead targets for TcTS binding<sup>77</sup>. By curating a database of ~2.5 million commercially available compounds using drug-likeness filters, removal of reactive groups and positive selection for compounds bearing anionic sidechains they chose 305,000 unique structures to interrogate by virtual screening. Utilizing the reported crystal structure of TcTS and selecting for both the sialic acid donor and acceptor binding pockets, virtual screening and visual inspection of docked leads yielded 23 diverse hits each bearing one negatively charged functional group for donor site binding and a hydrophobic group targeting the acceptor site. The hit molecules were purchased and experimentally evaluated as inhibitors of TcTS hydrolysis activity by fluorometric assay. The most promising compound **1.28** inhibited TcTS with a  $K_i = 0.15 \text{ mM}$  and its structure-activity relationship (SAR) was subsequently assessed. A set of eight derivatives (**1.29**) were experimentally evaluated, each exhibited mid-micromolar inhibitory activities ( $IC_{50} = 0.12\text{-}0.31 \text{ mM}$ ). Difficulties with co-crystallization led to inconclusive results

regarding the mode and site of molecular inhibition. Additional kinetic interrogation determined that these derivatives – bearing the 3-benzothiazol-2-yl-4-phenyl-but-3-enoic acid scaffold – modulate TcTS activity in a mixed or non-competitive manner and the authors suggest that multiple binding modes may be involved.

In 2010, Arioka and colleagues screened a library of 2283 commercially available natural products as inhibitors of TcTS hydrolysis<sup>78</sup>. Primary hits were selected by using a threshold of >40% inhibition activity at concentrations of 1  $\mu\text{M}$  which reduced the number of potential leads to 103 species. Secondary screening for  $\text{IC}_{50}$  activity <10  $\mu\text{M}$  and elimination of promiscuous inhibitors further focused the pool to 16 compounds. Analysis of structural complexity and application of Lipinski's constraints<sup>79</sup> led to the final selection of two lead compounds (**1.30**, **1.31**) which exhibited  $\text{IC}_{50}$  values of 0.58  $\mu\text{M}$  and 17  $\mu\text{M}$  respectively. SAR analysis of myricetin (**1.31**) derivatives revealed that the most heavily hydroxylated flavonoids exhibited the strongest inhibition, while the removal of functional groups from **1.30** led to a loss of inhibitor activity. Lead inhibitors were shown to specifically inhibit TcTS versus human neuraminidase NEU2 ( $\text{IC}_{50}^{\text{NEU2}}/\text{IC}_{50}^{\text{TcTS}} = 170$  for **1.30** and 8.8 for **1.31**). Mechanistic investigations of **1.30** and the myricetin analogue apigenin (**1.32**) with TcTS revealed reversible inhibition while Lineweaver-Burk plots indicated a non-competitive mode of action ( $K_i = 0.89 \mu\text{M}$  **1.30**;  $K_i = 60 \mu\text{M}$  **1.32**). The authors attribute the wide and shallow binding site of TcTS as a rationale for the non-competitive binding mode of the lead compounds. Further, **1.32** has been previously isolated from plant sources (*M. Kingii*) and was shown to elicit trypanocidal activity ( $\text{EC}_{50} = 0.68 \text{ mM}$ <sup>80</sup>) leading Arioka to speculate that these effects are caused by TcTS inhibition. However myricetin, apigenin and related compounds have also demonstrated significant cytotoxicity in *in vivo* studies<sup>81</sup> thus providing an alternative rationale for the observed antitrypanosomal activity. No follow-up studies have been reported that evaluate **1.30** in biological systems.

To date, two unique screening approaches have yielded TcTS inhibitor motifs which exhibit high-nanomolar to mid-micromolar activity. Each screen afforded molecules with better enzyme inhibition

than the most promising hits developed by any rational design approach. However cytotoxicity and non-competitive modes of action have limited the utility of the resultant compounds. Therefore additional screening efforts on new libraries may yield new promising scaffolds for TcTS inhibition.

## **1.6 Aims of thesis**

*Trypanosoma cruzi* trans-sialidase is hypothesized to be essentially involved in host cell invasion and immune response evasion during the acute phase of Chagas infection. Difficulties in target validation have arisen due to high gene copy number and low inhibitor efficacies. Therefore new small-molecule inhibitors with increased potency must be developed to provide the necessary chemical tools to evaluate TcTS as a therapeutic target.

### **1.6.1 Design and synthesis of modified difluorosialic acid inactivators**

Difluorosialic acids have been shown to inhibit *T. cruzi* trans-sialidase in a time-dependent fashion at millimolar concentrations. TcTS tolerates modification of sialoside derivatives at C9 as illustrated by previous members of the Withers lab who have achieved selectivity and improved kinetic parameters through synthetic modification at this position. Synthesis of a divergent library of C9 modified difluorosialic acids may further improve potency and selectivity for TcTS inhibition. Alternatively, modification of difluorosialic acid at previously unexplored sites may yield new beneficial interactions in the enzyme active site.

*Specific Aims:* Design and synthesize a library of 9-amido-difluorosialic acid derivatives and a library of 5-amido-difluorosialic acid derivatives. Characterize the inactivation efficacy of each compound by kinetic analysis and use lead compounds for further characterization.

### 1.6.2 Screening for new classes of TcTS inhibitors

Two screening efforts have unearthed the most potent TcTS inhibitors to date. Natural product extract and peptide-based libraries are a source of unexplored chemical diversity that have previously afforded potent glycosidase inhibitors. We believe these can be leveraged towards the discovery of new chemical scaffolds for TcTS inhibitor development.

*Specific Aims:* Design and validate a robust screening methodology for TcTS inhibition. Assay natural product extracts, perform bioassay-guided isolation of active components and structurally characterize hits. Kinetically investigate lead compounds, assess structure-activity relationships and evaluate lead TcTS inhibitors *in vivo* for activity versus *T. cruzi*. Utilize high-throughput screening technology to identify tight-binding cyclic peptide species for TcTS. Synthesize peptide hits and kinetically assess the lead compounds and their derivatives for inhibition activity versus TcTS.

## Chapter 2: Design of Mechanism-Based Inhibitors for TcTS

### 2.1 Introduction

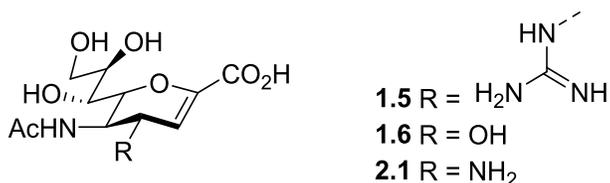
Substrate-based design of synthetic glycoside mimics is an established method to develop inhibitors of carbohydrate active enzymes. Classically, reversible inhibitors have been most sought after as they can offer tight-binding and non-permanent modulation of the target enzyme activity. Contrarily irreversible covalent bond forming inhibitors have generally been met with skepticism in the pharmaceutical community – due to potential safety concerns regarding compound specificity and the potential immunogenicity of resultant adducts<sup>82</sup> – despite essential medicines such as aspirin and penicillin having used covalent mechanisms to elicit their effects throughout the past century. Specific, irreversible, active-site targeting inactivators have some appealing advantages and have managed to recently gain cautious acceptance<sup>83</sup>. They employ a transient or permanent covalent bond-forming mode of action that can evoke strong target affinity and reduced biological clearance leading to prolonged therapeutic effects. Further, specific inhibitors presenting reactive chemical warheads will only inactivate enzymes bearing the required catalytic machinery to process them. Thus if the reactive species can be appropriately targeted and/or the turnover of the inactivated enzyme modulated, the risk of potential permanent off-target effects can be minimized. To this end, researchers are carefully characterizing and controlling inhibition specificity, potency as well as the lifetimes of the potential therapeutics in an attempt to quell these concerns. These elements were duly considered during the development of chemical therapeutics for influenza neuraminidase and serves as an ideal case study.

#### 2.1.1 Case study: influenza neuraminidase

Influenza neuraminidase (vNEU) plays an essential role during flu infection through its catalytic cleavage of the terminal sialic acid residues from the surface glycans. This process releases nascent virions from infected host cells and propagates influenza infection. Thus inhibition of this target was

hypothesized to disrupt virus release from the parent cell and was proposed as a therapeutic target for influenza infection. The transition-state analogue inhibitor DANA (**1.6**) and related derivatives inhibit vNEU with measured IC<sub>50</sub> values in the range of 1-10 μM. However *in vivo* studies failed to demonstrate any effects on animal influenza models leading to the temporary invalidation of vNEU as a relevant target<sup>84</sup>. Insufficient potency and a lack of specificity for vNEU versus human neuraminidase (hNEU) necessitated the chemical introduction of functional groups as selectivity factors. Upon comparison and analysis of the crystal structure complexes between vNEU bound to Neu5Ac and DANA respectively, an anionic pocket surrounding C4 was identified in vNEU. Von Itzstein and colleagues functionalized the established sialidase inhibitor at this position with positively charged substituents and demonstrated a dramatic increase in potency and selectivity for vNEU<sup>85,86</sup> (Table 2.1).

**Table 2.1. Inhibition constants of DANA analogues versus vNEU and hNEU<sup>85</sup>.**



Compound	K <sub>i</sub> (μM) vNEU (N2 A/Tokyo/3/67)	K <sub>i</sub> (μM) hNEU (lysosomal)
DANA ( <b>1.6</b> )	10	12
4-amino-DANA ( <b>2.1</b> )	0.5	9300
4-guanidino-DANA ( <b>1.5</b> )	0.002	1000

Substitution of the C4 hydroxyl for an amino functionality (**2.1**) increased potency by two orders of magnitude while addition of a C4 guanidino group (**1.5**) was so effective that this analogue has since been developed as a commercial therapeutic for influenza infection (*i.e.* Zanamivir) which elicits its effects through reversible target inhibition. Installing a large cationic species at C4 led to beneficial salt bridge interactions with surrounding carboxylate residues in the vNEU active site (D151, E227), while

simultaneously disfavoured binding with hNEU – likely due to unfavourable stereoelectronic interactions. This selectivity factor has become a standard feature in the effective development of potent influenza virus inhibitors. Addition of the established C4 guanidine to the mechanism-based inactivator 3-fluorosialosyl fluoride (**1.21**) has translated into the development of exciting new covalent inhibitors of vNEU<sup>87</sup>. Although dramatic enhancements in second-order inactivation rate constants ( $k_i/K_i$ ) were not observed following the addition of C4 guanidine the turnover of the covalent intermediate was significantly impeded (Table 2.2). Mechanistic investigations into the importance of the stereochemistry of the C3 fluorine substitution<sup>88</sup> and the stereoelectronic effects of C4 substitutions<sup>89</sup> have provided valuable insights into how the lifetime of the covalent fluorosialosyl-enzyme species can be modulated. For example, comparison of the 2eq3axDFSA (**2.2**, **2.3**) to the 2eq3eqDFSA (**2.4**, **2.5**) revealed that inactivators bearing 3-axial fluorine stereochemistry exhibited slower inactivation than did their 3-equatorial counterparts. However, the lifetime of the inactivated covalent intermediate is up to 50-times longer for the 2eq3axDFSA species. Thus by introducing selectivity versus hNEU through chemical modification of C4 and by controlling inactivation and reactivation rates via C3 fluorine stereochemistry, potent and targeted small-molecule covalent therapeutics have been developed for the influenza neuraminidase.

**Table 2.2. Inactivation and reactivation parameters of DFSA analogues versus vNEU (G70C H1N9)<sup>68</sup>.**

Compound (2eq3ax)	$k_i/K_i$ (mM <sup>-1</sup> min <sup>-1</sup> ) vNEU	$t_{1/2}$ reactivation (min)	Compound (2eq3eq)	$k_i/K_i$ (mM <sup>-1</sup> min <sup>-1</sup> ) vNEU	$t_{1/2}$ reactivation (min)
DFSA ( <b>1.21</b> )	196	<1	DFSA	-	-
4-am-DFSA ( <b>2.2</b> )	74	2300	4-am-DFSA ( <b>2.4</b> )	4332	49
4-gua-DFSA ( <b>2.3</b> )	246	>6900	4-gua-DFSA ( <b>2.5</b> )	3879	363

## 2.1.2 Structural modification of sialoside-mimics for enzyme specificity

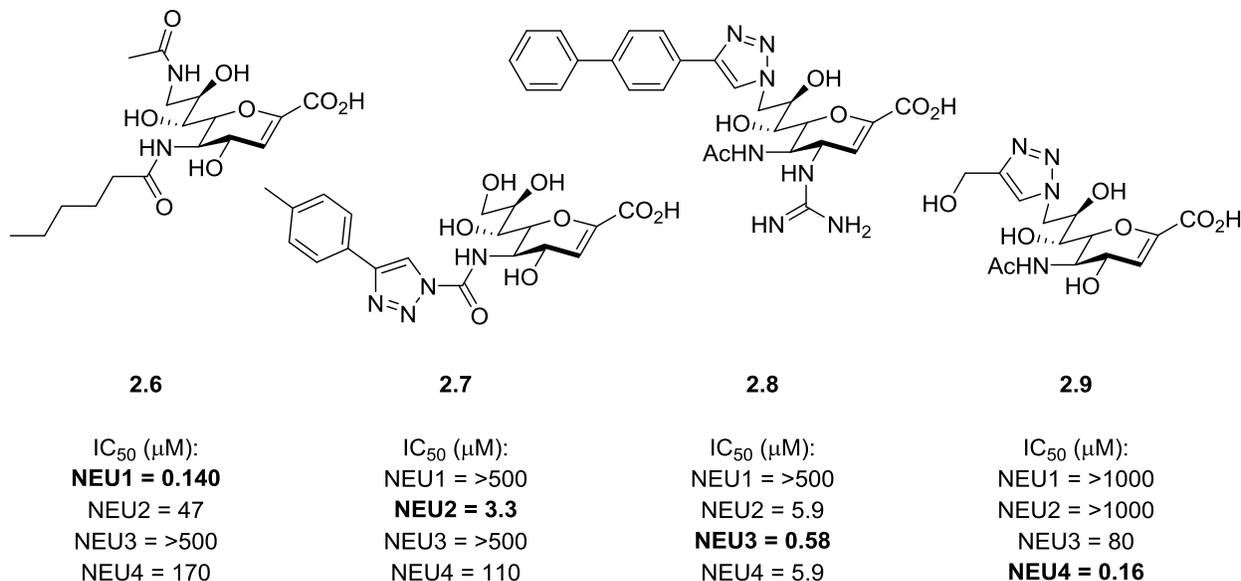
### 2.1.2.1 Modification of DANA for hNEU specificity

Similar efforts have been made to develop selective *T. cruzi* trans-sialidase (TcTS) inhibitors. As detailed in Chapter 1, potent TcTS inhibitors have not yet been identified; however, millimolar concentrations of the mechanism-based inhibitor **1.21** has been shown to ablate TcTS activity. Though kinetic analysis of this process determined that **1.21** inactivates TcTS ( $k_i/K_i = 8 \times 10^{-3} \text{ mM}^{-1} \text{ min}^{-1}$ ) some 180-fold slower than hNEU2 ( $k_i/K_i = 1440 \times 10^{-3} \text{ mM}^{-1} \text{ min}^{-1}$ ). Consequently, if this inhibitor scaffold is going to be useful for TcTS the selectivity will have to be dramatically inverted through chemical modification of the reagent.

Comparison of the known structural information regarding TcTS provides suggestive insight as to which positions of sialic acid are most amenable to substitution. Buschiazzo and colleagues first reported the successful crystallization of TcTS in 2002 – as well as an additional 14 structures of TcTS or its active site mutants to the protein data bank (PDB) over the following decade. These include enzyme structures in complex with either donor, acceptor, inhibitor or complete substrate molecules. These structures have

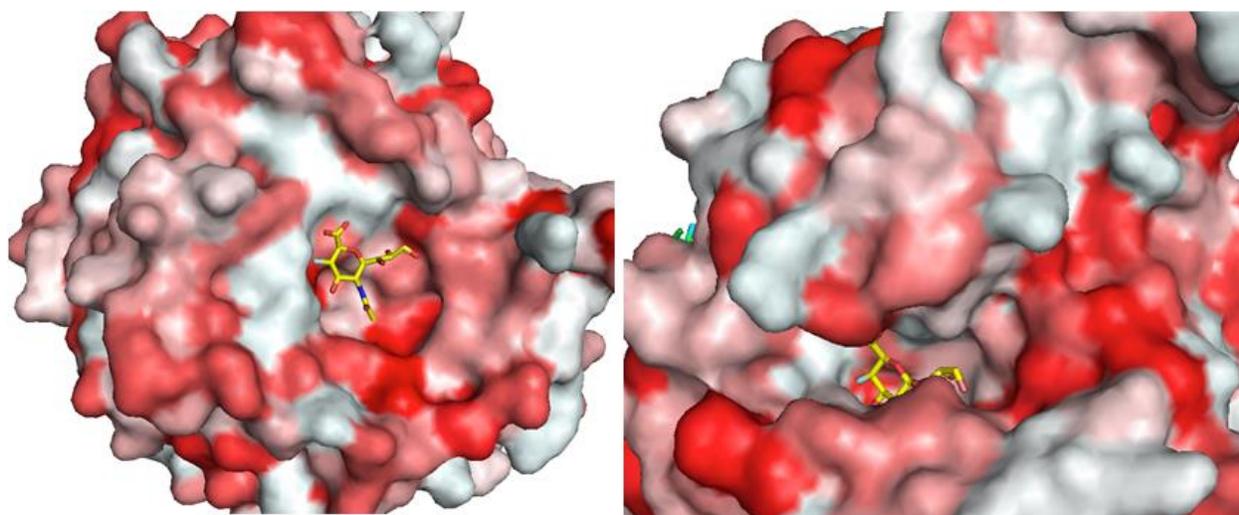
played an essential role in elucidating the mechanism of trans-sialidase activity, confirming the debated catalytic nucleophile residue<sup>41</sup> and understanding the dynamic interaction landscape of the TcTS active site<sup>40</sup>. Analysis of the key interactions with the native donor and acceptor glycosides has directed researchers towards the most promising sites of substrate-based inhibitor functionalization.

The substituents chosen to derivatize DFSA inactivators should not only enhance binding affinity with the target enzyme but simultaneously disfavour binding towards the endogenous neuraminidases. Four human neuraminidases have been identified (hNEU1-4) and their unique structural, catalytic and biological properties investigated. The subcellular localization of each isozyme can vary<sup>90</sup> but it has been reported that hNEU1 is primarily found in the lysosome or the plasma membrane, hNEU2 exists in the cytosol, hNEU3 is located in the plasma membrane and hNEU4 has been identified in the lysosome and endoplasmic reticulum<sup>91</sup>. The individual roles of each enzyme continue to be a subject of study yet it is accepted that together these hNEUs modulate the sialylation of cellular glycan structures and are implicated in the regulation of a range of human pathologies<sup>92</sup>. Research efforts have been largely dedicated to a better understanding of the specific biological roles these enzymes play while less information had been revealed about hNEU structural features or substrate/inhibitor binding preferences. However over the last decade the Cairo lab has made significant advances in this area, largely through the functionalization of DANA<sup>93</sup> to yield specific inhibitors of NEU1<sup>94</sup> (**2.6**), NEU2 (**2.7**), NEU3<sup>95,96</sup> (**2.8**) and NEU4<sup>97</sup> (**2.9**), shedding light on the isoform-specific binding preferences of the human neuraminidases (Figure 2.1).



**Figure 2.1. Substituted DANA derivatives as selective hNEU inhibitors**<sup>94,96,97</sup>.

hNEU2 has typically served as the primary model for human neuraminidases since its crystal structure has been elucidated<sup>98</sup> and the robust expression of the recombinant enzyme established<sup>99</sup>. By contrast, the membrane-associated isoforms hNEU1, hNEU3 and hNEU4 have generally proved more challenging to access. hNEU2 shares many of the structural features common to GH 33 family sialidases (such as TcTS), including the 6 bladed  $\beta$ -propeller fold, retaining mechanism, catalytic machinery and active site interactions with the C1 carboxylate and the C5 acetamide sidechain of sialic acid. The distinguishing structural difference between the trypanosomal trans-sialidase and the endogenous neuraminidase is the presence of three residues (E111, Y179 and Y181) in NEU2 that form a network of hydrogen bonding interactions with the hydroxyl groups of the Neu5Ac glycerol moiety<sup>98</sup>. Comparison to the TcTS crystal structure revealed that the trans-sialidase has a larger, hydrophobic pocket in the region surrounding the glycerol chain, thus indicating a potential subsite that could be leveraged to confer specificity for TcTS<sup>71</sup> (Figure 2.2).



**Figure 2.2. Crystal structure of covalent 3-fluorosialosyl fluoride-enzyme complex with hNEU2 (left) and TcTS (right).**

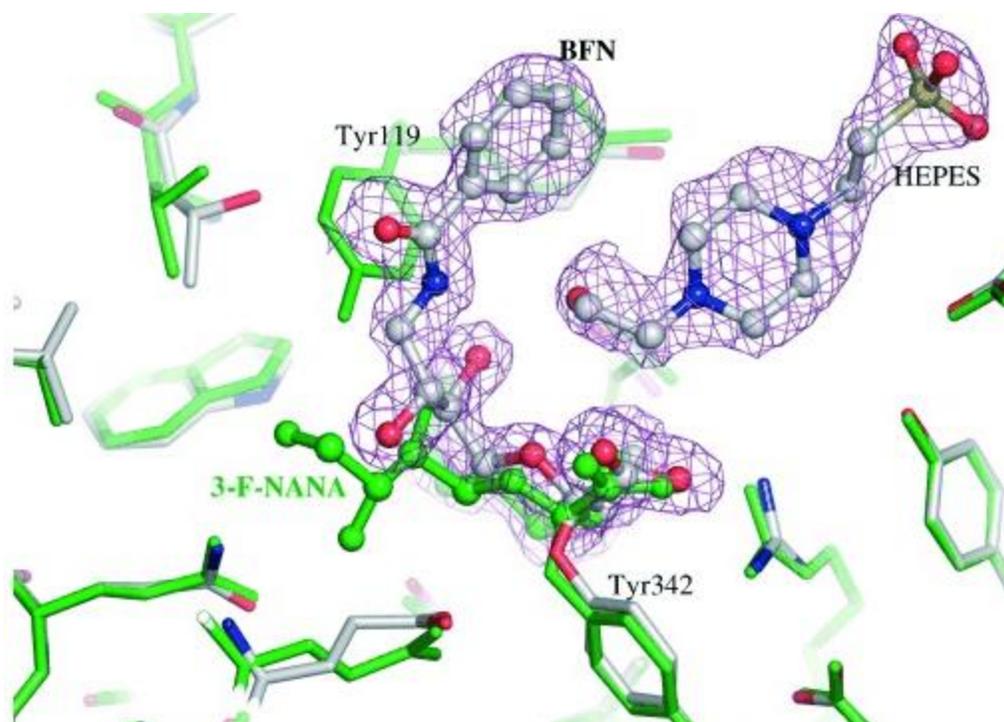
### 2.1.2.2 Modification of difluorosialic acid glycerol chain for TcTS specificity

Substrate specificity assays conducted with hNEU2 yielded results that are in line with structural expectations, whereby the enzyme demonstrates significantly reduced catalytic activity towards C9 *O*-methyl- or fluoro- or deoxy- modified sialosides<sup>100</sup>. By contrast processing of the C9-substituted sialosides by bacterial sialidases was largely unaffected. This feature has been explored against TcTS by Buchini et al. who synthesized a library of six 3-fluorosialosyl fluoride inhibitors functionalized at the C9 position<sup>71</sup>. An amine handle was installed in place of a hydroxyl group at this position and was used to create a divergent library of amide-linked DFSA derivatives. Kinetic evaluation of the resultant library indicated that the best C9 modification (**1.24**) marginally increased inactivation rates – with  $k_i/K_i$  1.2-fold higher than its DFSA parent compound (Table 2.3) – a far cry from the 100-fold increases observed in the vNEU case study. However, they determined that C9 amide-linked aryl substitutions led to >1000-fold decrease in sialyl-enzyme turnover rates in the presence of millimolar concentrations of lactose, relative to **1.21**. Their studies also showed that C9-substituted DFSAs inactivate hNEU2 150-fold slower the unsubstituted parent compound.

**Table 2.3. Inactivation and reactivation parameters of C9 amide-linked DFSA analogues versus TcTS<sup>101</sup>.**

<b>Compound</b> (2eq3ax)	$k_i/K_i \times 10^{-3}$ (mM <sup>-1</sup> min <sup>-1</sup> )	$k_r^{\text{Lact}} \times 10^{-3}$ (min <sup>-1</sup> )	$(k_r/K_i)^{\text{Lact}} \times 10^{-3}$ (mM <sup>-1</sup> min <sup>-1</sup> )
<b>DFSA (1.21)</b>	8	120	1330
<b>9-Bz-DFSA (1.22)</b>	2.9	N.D.	0.8
<b>9-MU-DFSA (1.24)</b>	14	N.D.	0.3

Structural analysis of the covalent complex between 9-benzoyl-3-fluoro-*N*-acetylneuraminic acid (**1.22**; BFN) and TcTS revealed that the modified glycerol side chain had rearranged to form interactions between the aryl substituent and Y119 in the acceptor binding site, prohibiting access of water or lactose thus decreasing the rate of turnover (Figure 2.3). Interestingly the glycerol chain does not occupy the hypothesized active-site pocket surrounding C9 (as observed for the unsubstituted 3F-NANA complex) but instead leans into the vacant acceptor site. While this approach yielded minimal change to the rates of TcTS inactivation, 9-amide linked aryl DFSA derivatives achieved both improved inactivator specificity versus human neuraminidase and significantly reduced the turnover rate of the inactivated enzyme species.



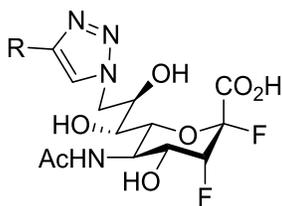
**Figure 2.3. X-ray structure of covalent intermediate between 1.22 (BFN) and TcTS (1S0I in grey). Structure is superimposed on the covalent intermediate of 1.21 (3F-NANA) and TcTS (2AH2 in green). Source: Buchini<sup>101</sup>**

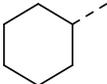
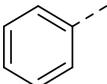
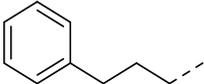
Subsequent work in the Withers lab reported the synthesis of 9-azido-DFSA, which was then derivatized using click chemistry to produce a library of twelve C9 triazole-linked derivatives. This aim of the library design was such that the conserved triazole linker would replicate the aromatic interactions observed between **1.24** and Y119 while the range of chemical functionalities employed would assess the surrounding active site residues for new binding interactions. However, kinetic analysis of **1.25** revealed that the reactivation rate of the sialyl-enzyme intermediate in the presence of lactose was similar to that of the unsubstituted derivative **1.21**. These studies suggest that the aforementioned glycerol chain rearrangement and resultant stabilization of the covalent intermediate does not occur in this case. The steric bulk and rigidity of the conserved triazole group relative to the amide linkage may have impeded glycerol chain rearrangement, or the stereoelectronic properties of the library derivatives may not be appropriate to harness analogous binding interactions with acceptor site amino acids. It was hypothesized

that the glycerol chain and appended triazole functionality may instead occupy the originally targeted C9 pocket distal to the acceptor binding site.

Kinetic evaluation of the DFSA library revealed previously unknown binding preferences in the enzyme active site, where the best derivative **1.25** had a second-order rate of inactivation that is five-fold higher than that of **1.21** (Table 2.4). C9 triazole-linked DFSAs bearing bulky aliphatic groups such as cyclopentyl or cyclohexyl groups, or anionic substitution had significantly higher inactivation rates than DFSA; while, hydrophilic functionality (*e.g.* hydroxyl groups), aromatic substituents or positively charged amines exhibited similar or diminished inactivation rates relative to that of the parent species. The disfavoured binding of cationic substituents is consistent with the reported lack of inhibition by 9-amino-DFSA<sup>71</sup>. Overall, these studies demonstrated that inactivation parameters for substituted DFSA derivatives can be differentially affected by functional group and linker selection. For example, relative to **1.21**, DFSAs with C9-triazole linked cyclic aliphatic groups had increased rates of TcTS inactivation and unchanged turnover rates while DFSAs with C9-amide linked aryl substitutions had markedly decreased reactivation rates, yet maintained comparable inactivation rate constants.

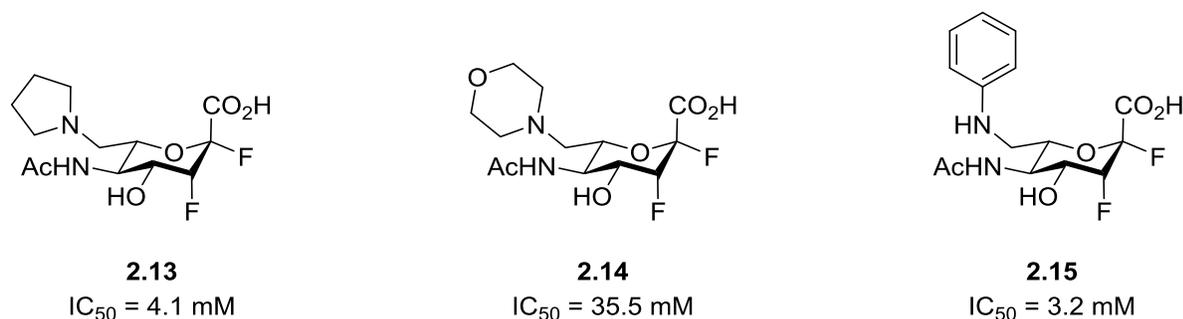
**Table 2.4. Structures and inactivation parameters of C9 triazole-linked DFSA analogues versus TcTS<sup>72</sup>.**



Name	R	$k_i/K_i \times 10^{-3}$ (mM <sup>-1</sup> min <sup>-1</sup> )	$k_i \times 10^{-3}$ (min <sup>-1</sup> )	$K_i$ (mM)
DFSA (1.21)	-	8	-	-
1.25		41	31	0.8
1.26		20	21	1.0
1.27		21	30	1.5
2.10		7	16	2.3
2.11		4	17	4.5
2.12	H <sub>2</sub> N <sup>+</sup>	5	44	8.4

Modification of the sialic acid glycerol chain has not been limited to the C9 position. A series of three truncated DFSA-like inactivators were synthesized – bearing cyclic functionality from a central C8-amine linkage (Figure 2.4) – and their inhibition kinetics assayed<sup>102</sup>. Transfer of sialic acid between Sia-Lac and <sup>14</sup>C-labelled lactose was measured by pre-incubating **2.13-2.15** with TcTS at a range of concentrations for 30 minutes prior to quantifying the residual TcTS activity. These experiments yielded IC<sub>50</sub> values between 3.2-35.5 mM. Despite the high inhibitory concentrations, these species were assessed

in cell-based assays at concentrations of 0.1 mM and 5 mM. Unsurprisingly, they found no observable changes to parasite infectivity at 0.1 mM, while trials at 5 mM showed cytotoxic effects to control cells.

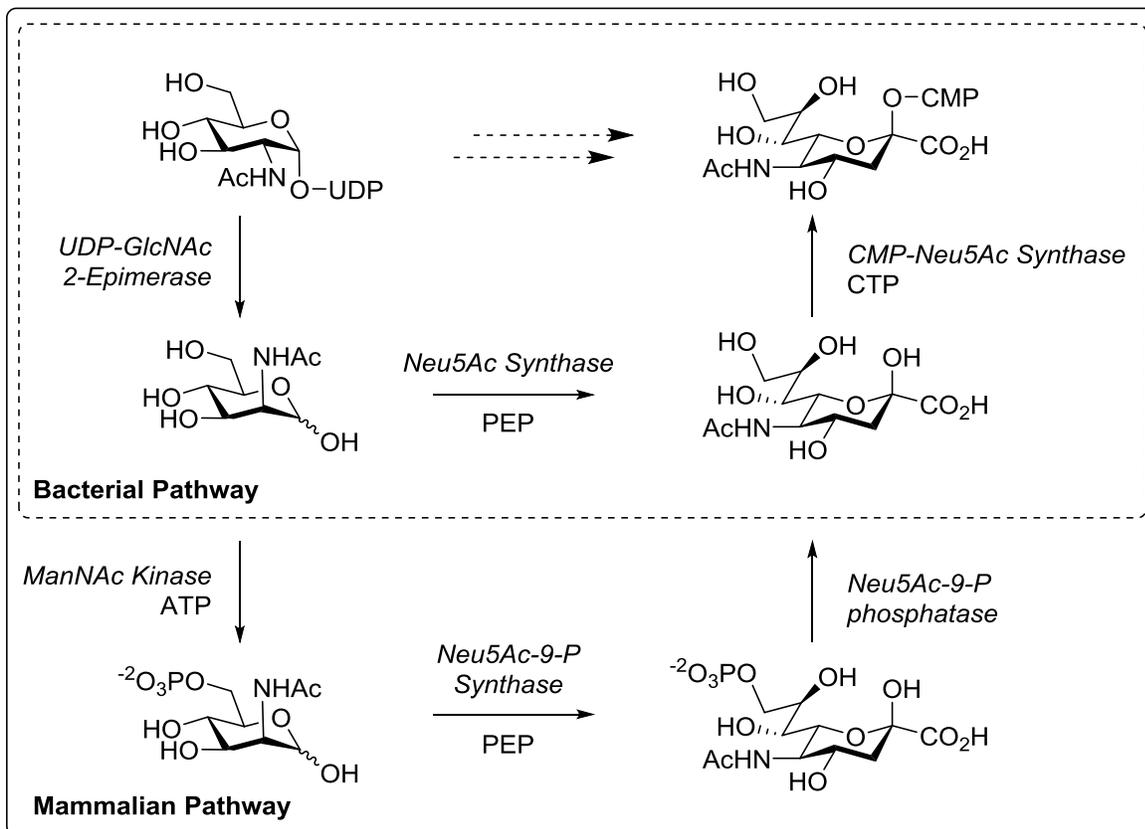


**Figure 2.4.** C7 modified DFSA analogues as inactivators of TcTS<sup>102</sup>.

### 2.1.2.3 Modification of sialosides at C4 and C5 for TcTS specificity

Although the glycerol chain of sialic acid was identified as the most promising modification site for TcTS selectivity, alternative positions of the sialic acid backbone have been examined. Given the success obtained by C4 substituted sialic acid mimics as selective vNEU inhibitors, these compounds have been investigated. Primarily, potent vNEU inhibitors were assessed for TcTS inhibition; however the positive charge provided by the amine or guanidine substitution – which is essential for vNEU specificity – cannot be accommodated by TcTS. As a result, vNEU therapeutics such as zanamivir and oseltamivir-carboxylate do not inhibit TcTS<sup>103</sup>. Crystal structures of DANA in complex with TcTS reveal an H-bonding network between the C4 hydroxyl and neighbouring arginine and carboxylate residues (R53/D96); disrupting these interactions may therefore lead to deleterious effects on binding.

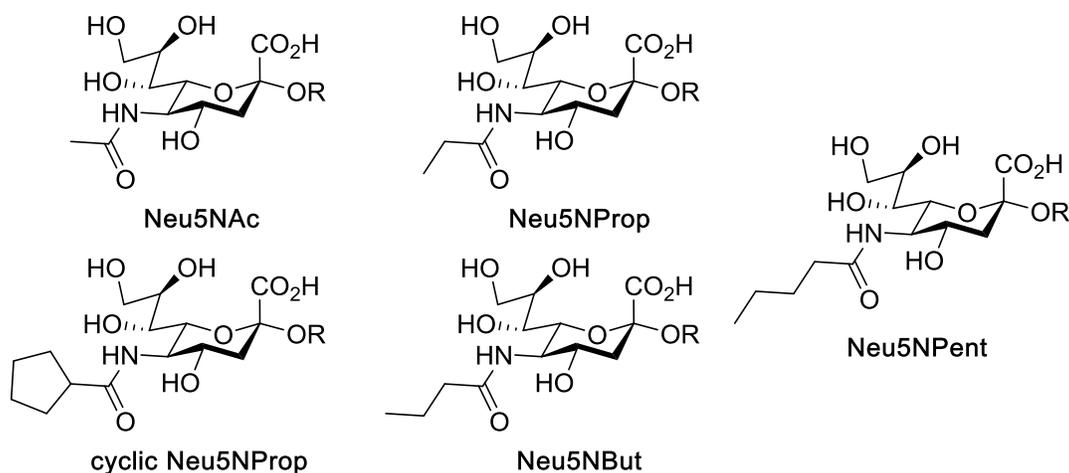
Members of the sialidase superfamily bear an essential carboxylate residue near C5 that forms H-bonding interactions with the *N*-acetyl group and appropriately orients the sialic acid donor for catalysis. It would therefore seem unwise to functionalize the C5 position since this might disturb the bonding network. However, it has been hypothesized that extension or modification of the *N*-acetyl group ‘tail’ may be tolerated if the H-bonding network with the carbonyl group is retained. To this end researchers have investigated the synthesis and bioactivity of such *N*-acetyl modified sialosides.



**Figure 2.5. Biosynthesis of CMP-Neu5Ac in bacteria and mammals.**

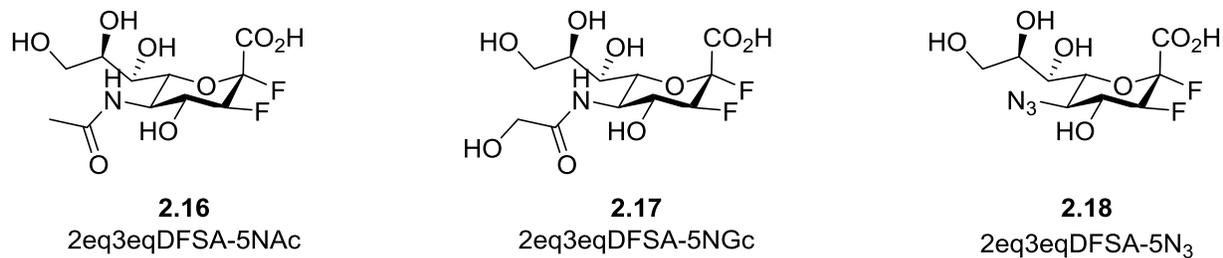
The biosynthesis of C5 modified sialic acid derivatives from *N*-substituted D-mannosamine precursors (Figure 2.5) and subsequent incorporation into surface displayed glycoconjugates in cellular systems was demonstrated by Kayser in 1992<sup>104</sup>. Several reports have subsequently assessed the ability of such *N*-substituted sialic acids to modulate interactions with cell surface receptors<sup>105–107</sup>. For example, Keppler and colleagues found that by displaying different *N*-modified sialic acids on the cell surface, binding interactions with sialic-acid dependent receptors of polyoma viruses could be significantly inhibited (95%) or enhanced (7-fold) relative to Neu5Ac depending on the individual modification<sup>105</sup>. More recently, a 2011 study by Lieke and colleagues assessed the relationship between *Trypanosoma cruzi* infection and the structure of *N*-modified sialylglycoconjugates on the surface of host cells (86-HG-39)<sup>108</sup>. Five distinct *N*-acyl sialic acids were synthesized from their *N*-modified mannosamine precursors

(Figure 2.6). The number of parasite-infected cells was reduced up to 60% (Neu5NProp) and infectivity decreased with increasing chain length. Together these studies highlight that the substituted *N*-acetyl moieties can differentially modulate the interactions between sialic acid and sialic acid receptors. Thus, modifying the *N*-acetyl group of sialic acid inhibitors may provide an alternative means of incurring selectivity for TcTS.



**Figure 2.6.** *N*-modified Neu5Ac analogues biosynthesized and assayed for effect on *T. cruzi* infection<sup>108</sup>.

A set of chemoenzymatically synthesized C5 or C9 modified DFSA analogues were very recently reported by Li and colleagues and the kinetic parameters for inhibition of eight bacterial sialidases determined<sup>109</sup>. Three C5 modified 3-fluorosialyl fluorides were assessed (**2.16-2.18**; Figure 2.7) and each species exhibited a unique profile of inhibition activity and selectivity versus the panel of sialidases. However, in each case the measured IC<sub>50</sub> value (determined after 30 mins of incubation between DFSAs and sialidase) for the natural DFSA **2.16** was always lower (more potent) than modified derivatives **2.17** and **2.18**. By assessing a larger, more diverse set of *N*-acyl modified DFSAs the authors may have found inactivators that offer improved selectivity for their panel of bacterial sialidases. Nevertheless, these species represent a new site of functionalization that can be interrogated for TcTS-targeted inactivator development.



**Figure 2.7.** C5-modified DFSA analogues as inactivators of bacterial sialidases<sup>109</sup>.

## 2.2 Specific aims

The overall goal of this chapter is to generate synthetically modified libraries of a known sialidase inactivator – 3-fluorosialosyl fluoride – in such a manner that its potency and selectivity towards TcTS may be improved. To this end, our first aim is to extend the scope of C9-modified 3-fluorosialosyl fluorides by building on the known linkages and chemical functionalities previously employed by the best TcTS inactivators. Our second aim is to synthesize a series *N*-acyl modified 3-fluorosialosyl fluorides. Inactivation kinetics will then be measured versus TcTS for each of the C9 or C5 modified difluorosialic acids.

## Results and Discussion

### 2.3 Synthesis of 9-amido-difluorosialic acids

To access a synthetic library of C9 modified inactivators, 9-azido-2eq,3ax-DFSA (**2.26**) was chosen as the parent compound (Figure 2.8). This scaffold bears (1) an appropriate handle at C9 for divergent library generation (2) the desired stereochemistry of the fluorine atoms and (3) maintains the necessary elements of the native substrate for optimal binding interactions within the enzyme active site.

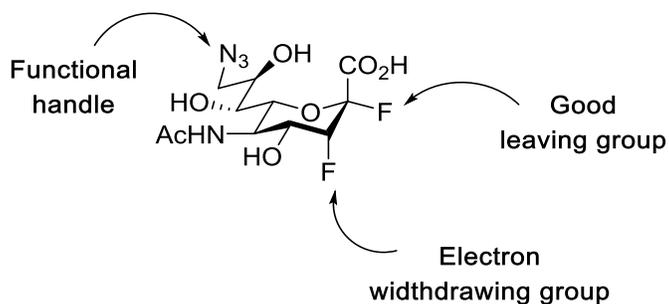
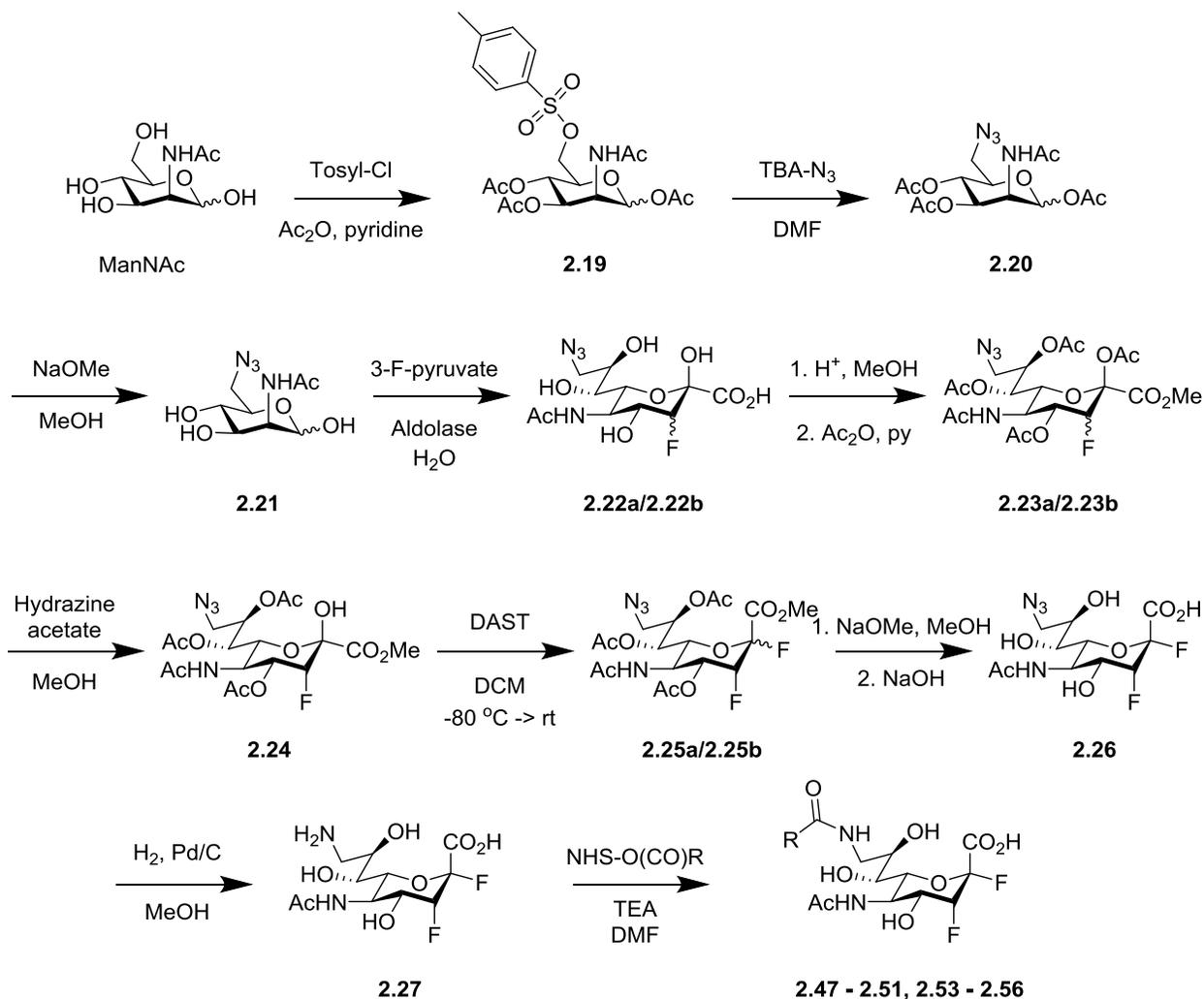


Figure 2.8. 9-azido-DFSA (**2.26**) structure and design features.

The azide substituent at C9 can be easily reduced to the corresponding amine and both the azide and the amine offer an appealing range of coupling techniques that can be used to generate libraries of modified DFSA derivatives. The addition of a fluorine atom with axial stereochemistry at the C3 position generates a longer-lived covalent sialyl enzyme intermediate of TcTS relative to the equatorial isomer, thus prolonging enzyme trapping. Amide coupling was selected as the method of choice for our C9-modified library, which was carried out between the primary amine of our parent compound and a series of activated esters to generate 9-amido-linked DFSA inactivators.

Synthesis of 9-azido-DFSA (**2.26**) can be achieved by two possible routes. The first is a purely chemical approach starting from Neu5Ac, whereby the azide moiety is first installed through substitution, then elimination across C2/C3 affords 9-azido-DANA, and finally sequential fluorination steps affords the 9-azido-DFSA product. The second is a chemoenzymatic approach from ManNAc, whereby 6-azido-

ManNAc is first prepared, then reacted with fluoropyruvate in an enzyme-catalyzed condensation followed by chemical installation of the anomeric fluoride to yield 9-azido-DFSA. This product is then reduced to afford **2.27** in both cases. The chemoenzymatic route is preferred in our case since the axial fluorine isomer of 3-fluorosialic acid can be preferentially installed by Neu5Ac aldolase whereas the chemical fluorination step affords a 1:1 mixture of axial and equatorial fluorine isomers. Therefore, this route was employed for the synthesis of **2.27** and is largely based on reported protocols<sup>71</sup>, with procedural changes made where necessary.



**Scheme 2.1. Synthesis of 9-amido-DFSAs.**

Briefly, the synthesis (Scheme 2.1) begins by activation of the C6 primary alcohol of ManNAc through reaction with tosyl chloride to afford 6-tosyl-ManNAc (**2.19**). Next, the remaining hydroxyl moieties were acetylated and the activated alcohol substituted by azide to yield per-*O*-acetylated 6-azido-ManNAc (**2.20**). Zemplen deacetylation yields the requisite starting material for enzymatic condensation with 3-fluoropyruvate to form 3-fluoro-Neu5Ac (**2.22**) as a mixture of stereoisomers (5:1 ax/eq). Esterification of the C1 carboxylic acid and acetylation of the remaining hydroxyls affords a mixture of 3-fluoro stereoisomers that can be separated by silica gel column chromatography (**2.23**). Selective deacetylation of the anomeric alcohol and nucleophilic substitution using diethylaminosulfur trifluoride (DAST) yields the difluorinated per-*O*-acetylated sialoside as a mixture of C2 stereoisomers (**2.25**; 3:1 eq/ax). 9-amino-2eq,3ax-DFSA is finally achieved by sequential deacetylation, saponification of the C1 methyl ester and palladium catalyzed reduction of the C9 azide moiety. This synthetic route was generally robust, although two steps (described below) in particular required investigation and optimization to achieve reasonable yields in my hands.

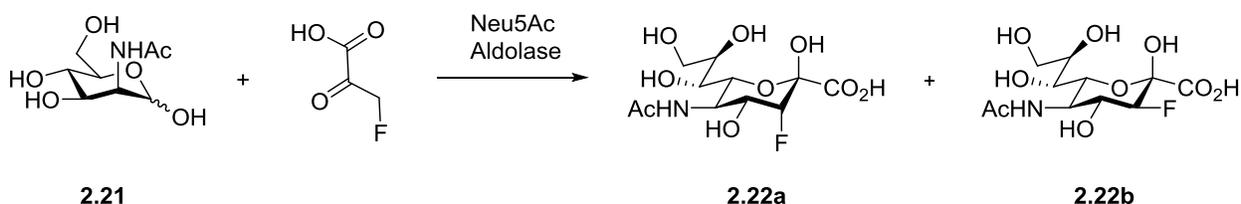
### 2.3.1 Azide substitution of 6-tosyl-ManNAc

Towards the synthesis of per-*O*Ac-6-azido-ManNAc (**2.20**), Buchini et al. reported the substitution of per-*O*Ac-6-tosyl-ManNAc (**2.19**) using trimethylsilyl azide (TMS-N<sub>3</sub>) and tetrabutylammonium fluoride (TBAF) in refluxing acetonitrile to afford the desired product in 73% yield. However, the hygroscopic nature of these reagents and the sensitivity of the reaction to water led to highly variable yields and was found to be unreliable for large-scale synthesis. A highly efficient substitution protocol<sup>110</sup> (>90% yield) was reported using sodium azide (NaN<sub>3</sub>) in warm DMF (80 °C), yet these results were irreproducible on multiple attempts, failing to afford products in >40% yield. Reactivity was notably improved (60-70% yield) through the addition of excess amounts of coordinating ligand (18-crown-6), yet the apparent amphiphilic nature of the resultant coordination complex led to separation issues during reaction workup. We speculated that the increased yield in the presence of 18-

crowd-6 is likely due to the coordination of the sodium cation, leading to increased availability of the nucleophilic azide anion in solution. Therefore, the use of a larger, solvent-compatible counter-ion may increase the reactivity of the azide reagent. To this end, tetrabutylammonium azide (TBA-N<sub>3</sub>) was freshly prepared and reaction with **2.19** afforded the desired substitution product in >80% yield. These conditions seemed to be more robust and efficient than previous protocols and were used during up-scale synthetic procedures.

### 2.3.2 Neu5Ac aldolase condensation

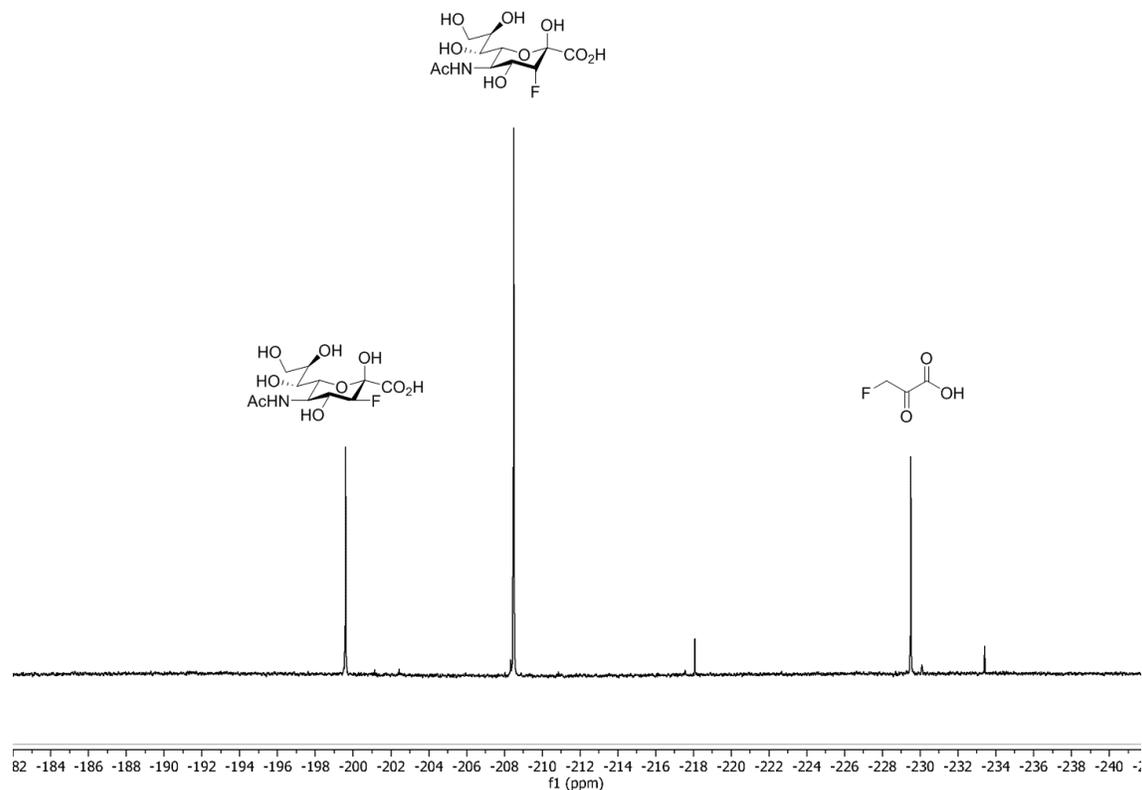
In the biosynthetic pathway of cytidine-5'-monophospho-*N*-acetylneuraminic acid (CMP-SA), ManNAc and pyruvate are coupled in an enzyme catalyzed process to afford Neu5Ac. Neu5Ac aldolase (*i.e.* sialic acid synthase; EC:4.1.3.3) is the class I aldolase responsible for this transformation. Fortunately, this enzyme exhibits robust activity and tolerates some modifications to both starting materials in the reaction process, although typically at lower rates. Thus our synthesis employs recombinant Neu5Ac aldolase (Carbosynth) to catalyze the condensation of 3-fluoropyruvate and 6-azido-ManNAc (**2.21**) to afford 9-azido-3FSA (**2.22**) (Scheme 2.2).



**Scheme 2.2. Enzymatic coupling of 3-fluoropyruvate and 6-azido-ManNAc.**

This reaction yields a mixture of C3 fluorine isomers, the ratio of which can be altered by controlling the reaction conditions<sup>71,111</sup>. The axial isomer is the first formed product but upon equilibration of the reaction with excess 3-fluoropyruvate a 1:1 mixture ensues. Thus to optimize for the desired C3 stereochemistry 0.3 equivalents of pyruvate were added in four portions (1.2 eq total) over ~3

days. Use of  $^{19}\text{F}$ -NMR to monitor reaction progress allows us to track the relative ratio of starting material consumed to products formed (Figure 2.9).



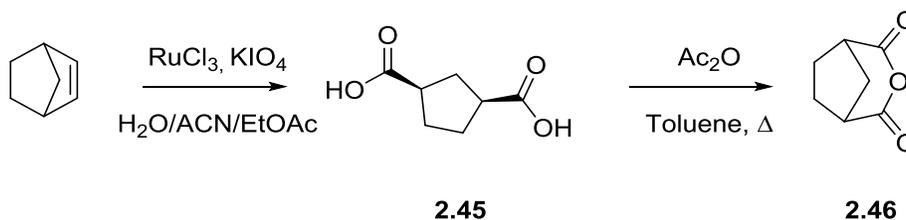
**Figure 2.9.**  $^{19}\text{F}$ -NMR of Neu5Ac aldolase reaction progress after 6 hours.

This information is crucial as it informs the timing of subsequent fluoropyruvate addition and allows the ratio of formed products to be assessed as the reaction progresses. We observed that (1) timing of starting material addition and (2) freshness of purchased aldolase greatly affected stereochemical outcomes. To investigate the latter point, a colleague in our lab (Dr. Lars Baumann) dissolved and analyzed the fresh and stored batches of lyophilized enzyme by SDS-PAGE and discovered many contaminating bands in both samples. The recombinant enzyme was subsequently expressed in *E. coli* and purified in house; I then assessed the reaction performance with 3-fluoropyruvate. Stereochemical specificity was 5:1 ax/eq, an equivalent ratio to that obtained for the lyophilized commercial enzyme.

Following reaction completion the enzyme can be removed with a spin filter (Amicon 30 kDa) or the complete reaction mixture can be flash frozen (-78 °C), lyophilized and the products dissolved in methanol. Separation of the fluorine isomers of the highly polar products was challenging, therefore protecting groups were first installed on the C1 carboxylate and hydroxyl moieties. Iterative rounds of silica gel column chromatography or preparative scale HPLC were employed to isolate **2.23a** and **2.23b** respectively.

### 2.3.3 Acylation of 9-amino-difluorosialic acid

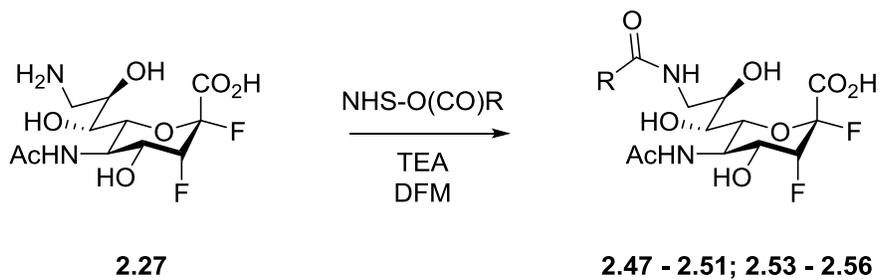
Generation of the desired 9-amide linked library of 3-fluorosialyl fluorides was achieved by coupling of the primary amine of 9-amino-DFSA with a series of *N*-hydroxysuccinimidyl (NHS) esters. These activated esters were stable to purification and storage while also providing selective reactivity towards the desired amine functionality in the chosen reaction conditions. Eight of the desired esters were prepared from the commercially available acids and *N*-hydroxysuccinimide using standard 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) coupling procedures. However the attempted activation of cyclopentane-1,3-dicarboxylic acid (**2.45**) using this strategy was unsuccessful. In this case **2.45** (prepared from oxidative ring opening of norbornene) was cyclized using acetic anhydride to form the bicyclic anhydride product 3-oxabicyclo[3.2.1]octane-2,4-dione (**2.46**) (Scheme 2.3).



**Scheme 2.3. Synthesis of 3-oxabicyclo[3.2.1]octane-2,4-dione (2.46).**

Coupling of the activated NHS-esters or bicyclic anhydride reagent and 9-amino-DFSA was accomplished following literature procedures in DMF with excess triethylamine affording the 9-amido-DFSA products in 46-70% yield after purification (Table 2.5).

**Table 2.5. Reaction yields from 9-amido-DFSA coupling reactions.**



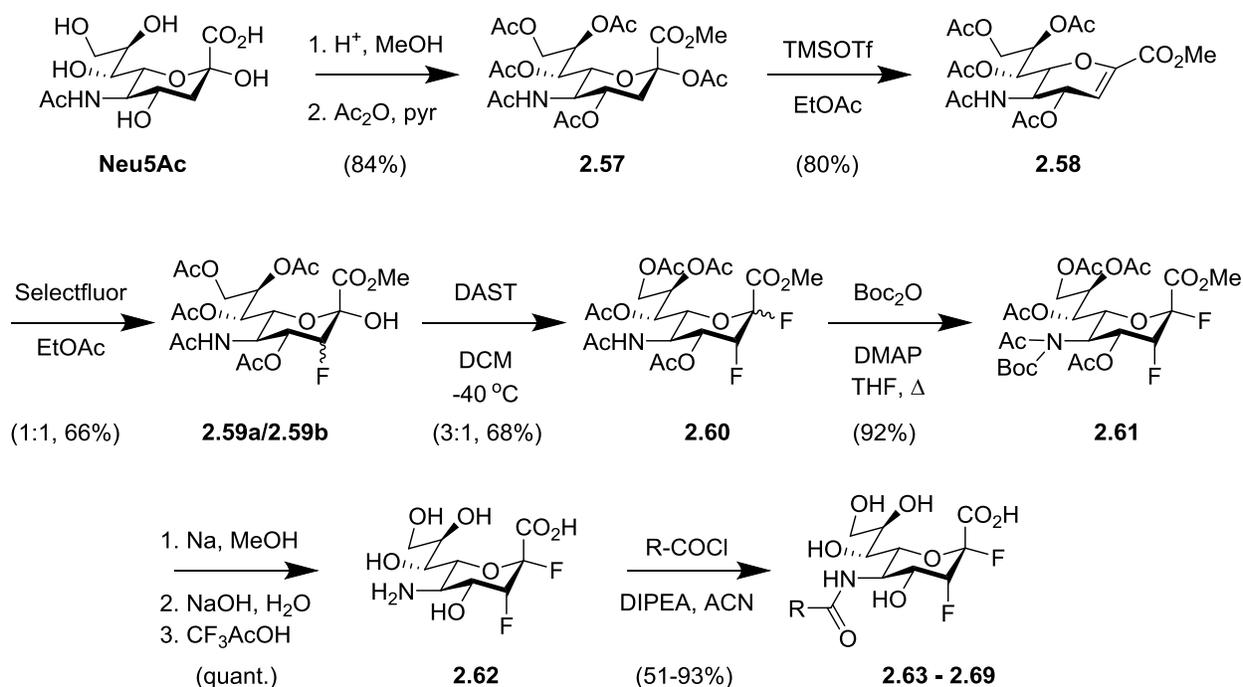
Name	R =	Isolated Yield (%)*
9-amido(Bz)DFSA ( <b>2.47</b> )		49
9-amido(CH)DFSA ( <b>2.48</b> )		46
9-amido(CP)DFSA ( <b>2.49</b> )		48
9-amido(THF)DFSA ( <b>2.50</b> )		65
9-amido(Furan)DFSA ( <b>2.51</b> )		61
9-amido(Pro)DFSA ( <b>2.53</b> )		70
9-amido(Pyr)DFSA ( <b>2.54</b> )		68
9-amido(CPAA)DFSA ( <b>2.55</b> )		51
9-amido(CPCA)DFSA ( <b>2.56</b> )		68

\*General reaction conditions employ 1.5 eq. activated ester and 2 eq. triethylamine; final products purified by silica gel chromatography (See Chapter 7.1.4 for details).

## 2.4 Synthesis of *N*-acyl-difluorosialic acids

In order to generate the library of *N*-acyl modified DFSA derivatives, 5-amino-2eq,3ax-DFSA (**2.62**) was chosen as the target parent compound. This species maintains the same fluorine atom stereochemistry as the aforementioned libraries and bears the necessary amine at C5 for chemical derivatization. There are two potential synthetic routes for 5-amino-DFSA that are analogous to those considered for 9-amino-DFSA. However in this case the chemoenzymatic method is more synthetically arduous despite the preferential means of C3 fluorine installation. Therefore the synthetic method – starting from Neu5Ac – was employed to achieve the product in nine steps.

First, esterification and per-*O*-acetylation of Neu5Ac followed by acid-catalyzed elimination of the anomeric acetate yields the protected sialidase inhibitor DANA (**2.58**). The fluorine atom at C3 is installed by reaction with the electrophilic fluorination reagent Selectfluor, yielding the hemiketal (**2.59**) as a 1:1 mixture of fluorine stereoisomers. Reaction with the nucleophilic fluorine source DAST affords the anomeric fluoride **2.60** in a mixture of anomers. To enable the removal of the *N*-acetyl group at C5, 4-dimethylaminopyridine (DMAP)-mediated addition of a *tert*-butyloxycarbonyl (Boc) ester to the acetamide was enlisted to yield the peracetylated NAc(Boc)-DFSA methyl ester (**2.61**). Sequential global deacetylation, saponification of the methyl ester and removal of the Boc ester with trifluoroacetic acid (TFA) afforded the target compound **2.62** (Scheme 2.4).



**Scheme 2.4. Synthesis of 5-amido-DFSAs from Neu5Ac.**

### 2.4.1 Elimination of Neu5Ac

Neu5Ac2en (*i.e.* DANA) is a common inhibitory scaffold for sialidases and its synthesis has been described many times. Yet the most efficient method for the formation of the C2/C3 unsaturation remains contested. The most common route involves the formation of the anomeric chloride species and subsequent elimination with an organic base (typically DBU). However, formation of the glycosyl chloride can be variably efficient and competing intermolecular and intramolecular substitution reactions can yield deleterious side products. An alternative method involves the activation of the anomeric acetate via addition of a Lewis acid (*e.g.* trimethylsilyl trifluoromethanesulfonate (TMSOTf)) and elimination under cooling conditions. A major reaction that is observed at warmer temperatures is the intramolecular substitution of the C4 acetate by the neighbouring *N*-acetyl group to afford the oxazoline Neu5Ac2en derivative. In fact this species is a key intermediate in the synthesis of C4 guanidine substituted DFSA or zanamivir, but is not wanted in this case. To deter the unwanted reactivity I employed dropwise addition

of TMSOTf to the solution of protected Neu5Ac (**2.57**) in ethyl acetate at 0 °C, which afforded the desired product **2.58** in 80% yield after purification by silica gel column chromatography.

#### 2.4.2 Electrophilic fluorination of Neu5Ac2en

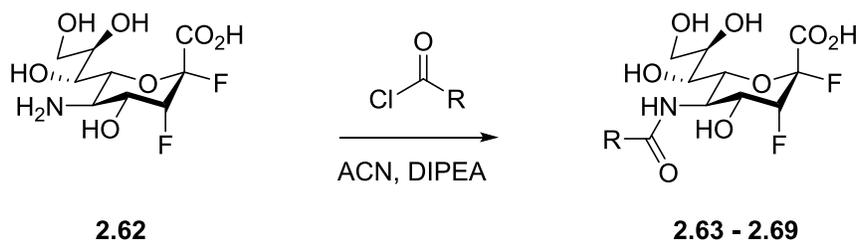
Selectfluor has been the most commonly employed electrophilic fluorinating agent of its kind and is produced on the multi-ton scale for industrial purposes<sup>112</sup>. However, the reported fluorination reaction conditions with **2.58** require 7 days of stirring at room temperature in nitromethane to obtain ~70% yield<sup>87</sup>. In principle, the reaction time could be shortened by increasing the reaction temperature but heating nitromethane presents significant safety concerns. Therefore others have explored alternative solvents for similar reactions<sup>113,114</sup>. I followed suit, aiming to find a more amenable solvent that could afford comparable yields while reducing the reaction time. I found that four equivalents of Selectfluor in refluxing EtOAc afforded spot-to-spot conversion on TLC over the course of 48 hours. However, upon scale-up the isolated yield of **2.59** was 66% as a 1:1 mixture of isomers despite the promising test scale attempts. The partial aqueous solubility of **2.59** may have contributed to substantial product loss during liquid-liquid extraction. Regardless, similar yields were achieved to those reported in nitromethane with a reduction in reaction time by five days.

#### 2.4.3 Acylation of 5-amino-difluorosialic acid

While the formation of 9-amido-DFSA derivatives was achieved by reaction of **2.27** with freshly prepared NHS-esters, 5-amino-DFSA requires more reactive acylation reagents to form the targeted amido species. Therefore commercial acyl chlorides were purchased and employed as coupling partners. To date, only one protocol for a coupling involving 5-amino-DFSA has been reported<sup>109</sup> and they used acetoxyacetyl chloride in acetonitrile/water with sodium bicarbonate to yield their target Neu5Gc. These general reaction conditions were replicated by Dr. Hongming Chen and successfully used to afford **2.63** upon coupling with benzoyl chloride. However the same conditions failed to produce product upon

reaction with propionyl chloride. Therefore, alternate coupling conditions were explored. Dr. Chen reacted **2.62** with acyl chlorides in acetonitrile and two equivalents of *N,N*-diisopropylethylamine (DIPEA) to afford the desired 5-amido-DFSA products. After purification by silica gel chromatography the products were obtained in 51% to 93% yield (Table 2.6).

**Table 2.6. Reaction yields from 5-amido-DFSA coupling reactions.**



Name	R =	Isolated Yield (%)*
5-NHAc(Bz)DFSA ( <b>2.63</b> )		81**
5-NHAc( <i>n</i> Prop)DFSA ( <b>2.64</b> )		89
5-NH( <i>n</i> But)DFSA ( <b>2.65</b> )		93
5-NH( <i>n</i> Pent)DFSA ( <b>2.66</b> )		51
5-NH( <i>i</i> Pent)DFSA ( <b>2.67</b> )		72
5-NH( <i>c</i> Pent)DFSA ( <b>2.68</b> )		58
5-NH( <i>n</i> Hex)DFSA ( <b>2.69</b> )		75

\*General reaction conditions employ 2 eq. acyl chloride and 6 eq. *N,N*-diisopropylethylamine; final products purified by silica gel chromatography (See Chapter 7.1.6 for details). \*\* Formed using alternate reaction conditions: 5 eq. benzoyl chloride, 1:1 ACN/H<sub>2</sub>O and sodium bicarbonate<sup>109</sup>.

## 2.5 Kinetic analysis

### 2.5.1 Enzyme expression and characterization

TcTS was expressed and purified according to literature protocols with minor adjustments<sup>115</sup>. Recombinant TcTS was expressed in *E. coli* BL21(DE3) cells and purified by successive HisTrap and MonoQ anion exchange column chromatography (see Chapter 7.1.7 for details). 4-Trifluoromethylumbelliferyl sialoside (TFMU-SA) was used as a chromogenic substrate, hydrolysis by TcTS was monitored at 385 nm using a UV/Vis spectrophotometer. All assays were conducted in standard TcTS buffer (20 mM Tris, 30 mM NaCl, 0.1% BSA, pH = 7.5). The concentration of TFMU-SA was varied and the reaction rates ( $v$ ) determined by linear regression. Plotting the initial rates versus concentration of substrate and fitting to the Michaelis-Menten equation (Equation 2.1) affords kinetic parameters  $K_m$  and  $V_{max}$  (which can be used to calculate  $k_{cat}$ ; see Appendix A); values obtained were in accordance with literature<sup>43</sup>.

#### Equation 2.1. Michaelis-Menten equation.

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

### 2.5.2 Inactivation kinetics: protocol

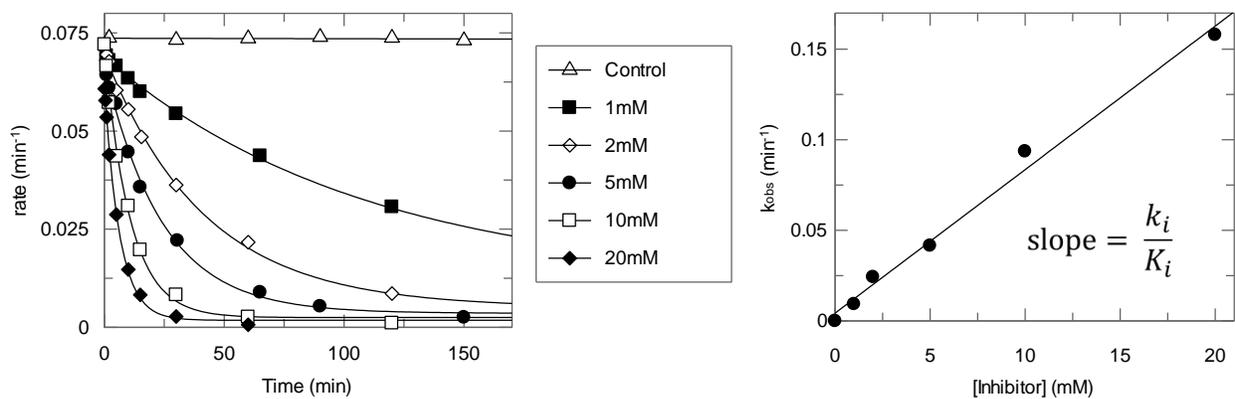
Following reported procedures<sup>71</sup>, TcTS was incubated with varying concentrations of 3-fluorosialosyl fluoride at 25 °C for 2-3 hours. For each concentration of the individual inactivator nine time points were measured in an effort to capture the complete TcTS decay curve. As a control, TcTS was incubated without inactivator and the enzyme activity was measured over the length of the experiment: no significant change in activity was observed. Aliquots of the incubating inactivator/enzyme mixture were diluted 10-fold into quartz cuvettes containing excess buffered TFMU-SA (0.5 mM) at 25 °C. Initial rates ( $v_0$ ) of substrate cleavage were measured and plotted versus incubation time for each inactivator concentration (Figure 2.10, left panel). Resultant data was then fit to a first-order exponential decay

equation (Equation 2.2) using data analysis software GraFit 7.0 (Erithacus) to obtain the rate constants  $k_{\text{obs}}$  that were then plotted versus inhibitor concentration (Figure 2.10, right panel).

**Equation 2.2. Exponential decay equation with offset.**

$$v = v_0 e^{k_{\text{obs}} t} + \text{offset}$$

Ideally the plot of  $k_{\text{obs}}$  versus inactivator concentration should yield a Michaelis-Menten like curve from which the individual inhibition parameters  $k_i$  and  $K_i$  (analogous to  $K_m$  and  $V_{\text{max}}$ ) can be extracted. However, as was observed for previous DFSA inactivators measured against TcTS, saturation cannot be achieved thus only the initial linear portion of the curve can be obtained (Figure 2.10). Higher concentrations of inactivator would be required to determine the complete shape of the inactivation curve, but the practicalities of measuring the fast rates of decay at those concentrations makes this infeasible. While being unable to parse the individual parameters, the second-order rate constant of inactivation can be determined ( $k_i/K_i$ ) from the slope of the linear plot. The second-order rate constant is used as the basis for comparison between DFSA inactivators: the greater the value, the better the inhibitor.



**Figure 2.10. Determination of inactivation rate constants ( $k_{\text{obs}}$ ) (left) and second-order inactivation rate constant ( $k_i/K_i$ ) (right) for 2.55 versus TcTS.**

### 2.5.3 Inactivation kinetics: 9-amido-DFSAs

The first three inactivators synthesized (**2.47-2.49**) were designed based on previously reported results. A reference compound, **2.47**, was selected for which the kinetic parameters for inactivation were previously reported. Compounds **2.48** and **2.49** are the amide-linked analogues of the best functionalities reported in Yao's triazole-linked library. Employing the methods described above, second-order rate constants of inactivation were obtained for each derivative (Table 2.7).

**Table 2.7. Inactivation parameters of first-generation 9-amido-DFSAs versus TcTS.**

Name	Structure	$k_i/K_i \times 10^{-3} \text{ (mM}^{-1}\text{min}^{-1}\text{)}$
<b>2.47</b>		$5.2 \pm 1.1$
<b>2.48</b>		$7.2 \pm 0.4$
<b>2.49</b>		$81.6 \pm 6.0$

The second-order rate constant measured for **2.47** ( $k_i/K_i = 5.2 \times 10^{-3} \text{ mM}^{-1} \text{ min}^{-1}$ ) was comparable to the value reported by Buchini ( $k_i/K_i = 2.9 \times 10^{-3} \text{ mM}^{-1} \text{ min}^{-1}$ ). Meanwhile some parallels can be noted between the results obtained for **2.48** and **2.49** and the corresponding triazole derivatives. The cyclopentyl amide-linked species **2.49** was the most potent DFSA inactivator of TcTS to date, just as the cyclopentyl triazole-linked DFSA was the best of Yao's inactivators. Although the same relationship was not

observed between the triazole- and amide-linked cyclohexyl derivatives, since **2.48** had  $k_i/K_i = 7.2 \times 10^{-3} \text{ mM}^{-1} \text{ min}^{-1}$ , which is essentially the same as unsubstituted DFSA. The 5-membered ring structure at C9 seems to be essential for optimal inactivation of TcTS, where increasing or decreasing ring size led decreased inactivation rates with amide- and/or triazole-linked inactivators. While it may seem that these interactions with TcTS are likely hydrophobic in nature, we sought to investigate whether additional binding interactions can be harnessed (*e.g.* H-bonding, ion pairs,  $\pi$ - $\pi$  stacking) while maintaining the cyclopentyl core structure. A second-generation library was therefore synthesized bearing aromatic or aliphatic five-membered rings containing potential H-bond acceptors and donors or anionic functionality and their kinetic parameters for TcTS inactivation measured (Table 2.8; See Appendix B for all plots).

**Table 2.8. Inactivation parameters of second-generation 9-amido-DFSAs versus TcTS.**

Name	Structure	$k_i/K_i \times 10^{-3} \text{ (mM}^{-1} \text{ min}^{-1}\text{)}$
2.50		$2.2 \pm 0.4$
2.51		$6.0 \pm 0.4$
2.53		$5.0 \pm 0.1$
2.54		$17.4 \pm 3.0$
2.55		$7.9 \pm 0.4$
2.56		$31.8 \pm 4.1$

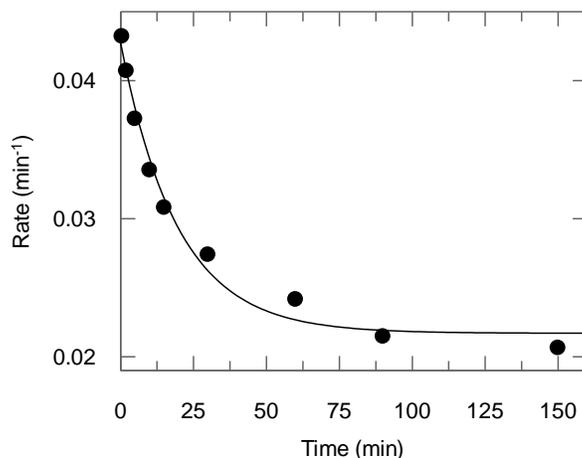
Disappointingly, the heterocycles (tetrahydrofuran, furan, proline and pyrrole) had inactivation parameters that were similar to or slower than that of the unsubstituted derivatives, while surprisingly the aliphatic cyclopentyl acetamide species **2.55** exhibited similarly poor results. It is possible that a local maximum of binding has been found, especially since either elongation or expansion of the cyclopentyl derivative by a single carbon led to ~10-fold decreases in rate constants for TcTS inactivation. If this is the case, hydrophilic atoms of the heterocyclic functional groups may have disrupted the proposed hydrophobic interactions of the 5-membered aliphatic ring, leading to the poor inactivation results we obtained.

However, the second best results were obtained with the anionic derivative **2.56**, which had a measured  $k_i/K_i$  of  $31.8 \times 10^{-3} \text{ mM}^{-1} \text{ min}^{-1}$ . While this result may be contrary to the above hypothesis, it is consistent with the above-average inactivation constant measured on the triazole-linked carboxylate species **1.27** ( $k_i/K_i = 20 \times 10^{-3} \text{ mM}^{-1} \text{ min}^{-1}$ ). Further, C9 functionalized DFSAs with cationic charge (**2.12**, **2.53**) have second-order rate constants  $\leq 5 \times 10^{-3} \text{ mM}^{-1} \text{ min}^{-1}$  and in the case of 9-amino-DFSA (**2.27**) have no detectable inactivation at all. The presence of a stabilizing ionic interaction with the carboxylate groups and a destabilizing interaction with the amino groups could suggest the presence of a positively charged amino acid in the glycerol binding pocket. If so, a salt bridge interaction could be leveraged for increased selectivity and potency towards TcTS (as was achieved for vNEU with a C4 guanidino group). Yet without additional structural information about the binding mode of these compounds any conclusions regarding the basis for the contrasting inactivation parameters between cationically and anionically charged DFSA derivatives is purely speculative. Additional derivatization at the C9 position of DFSA focusing on anionic substituents could be explored to further investigate this matter.

#### **2.5.4 Inactivation kinetics: 5-amido-DFSAs**

Literature studies demonstrated that *N*-modified sialylglycoconjugates bearing elongated aliphatic chains could differentially regulate the interactions with sialic acid receptors in cell-based studies.

Therefore, incorporation of such modification into DFSA inactivators may differentially affect inactivator potency and selectivity towards DFSA. To this end, seven *N*-modified DFSA derivatives were synthesized. Elongation of the linear *N*-acyl chain was examined as well as the tolerance for branched and cyclic pentane derivatives. To assess the relative inactivation of compounds **2.63-2.69** a comparative assay was employed. A single concentration (5 mM) of each derivative was incubated with TcTS and the time-dependent inactivation assayed employed as described for the 9-amido linked DFSAs (Figure 2.11).



**Figure 2.11. Determination of inactivation rate of 2.68 versus TcTS.**

Each curve was fit to a first-order decay equation (Equation 2.2) to obtain a single pseudo first-order rate constant ( $k_{\text{obs}}$ ) for each derivative (Table 2.9). The half-life ( $t_{1/2}$ ) of the enzyme in the presence of each inactivator was also calculated (Equation 2.3).

**Equation 2.3. Half-life ( $t_{1/2}$ ) of exponential decay ( $k_{\text{obs}}$ ).**

$$t_{1/2} = \frac{0.693}{k_{\text{obs}}}$$

**Table 2.9. Inactivation rate constants ( $k_{\text{obs}}$ ) and half-life ( $t_{1/2}$ ) of 5-amido-DFSAs (5 mM) versus TcTS.**

Name	Structure	$k_{\text{obs}}$ ( $\times 10^{-3} \text{ min}^{-1}$ )*	$t_{1/2}$ (min)
1.21		$50^{101}$	14
2.63		$47.7 \pm 8.8$	14.5
2.64		$43.3 \pm 5.2$	16.0
2.65		$70.1 \pm 12.6$	9.89
2.66		$40.1 \pm 5.1$	17.3
2.67		$36.6 \pm 9.6$	18.9
2.68		$51.3 \pm 6.9$	13.5

Name	Structure	$k_{\text{obs}}$ ( $\times 10^{-3} \text{ min}^{-1}$ )*	$t_{1/2}$ (min)
2.69		$25.0 \pm 5.5$	27.7

\* All pseudo-first order rate constants ( $k_{\text{obs}}$ ) determined by incubation with 5 mM of DFSA analogues.

All seven species inactivated TcTS at similar rates ( $0.025$ - $0.070 \text{ min}^{-1}$ ); although the butyryl (**2.65**) and cyclopentanoyl (**2.68**) derivatives were the fastest inactivators and the *n*-hexanoyl was the slowest. Importantly these results are comparable to that of the *N*-acetyl species ( $k_{\text{obs}} = 0.050 \text{ min}^{-1}$ ), which means that *N*-acyl functionalization – with aliphatic groups of six or less carbons – is generally accepted in the TcTS active site. Although target affinity does not appear to be markedly increased, selectivity versus human neuraminidases that do not tolerate *N*-acetyl modification can likely be conferred. For example, selectivity for hNEU1 and hNEU2 can be achieved versus hNEU3 and hNEU4 by addition of short, branched aliphatic modifications on the *N*-acyl group of DANA<sup>94</sup>. Similarly, selectivity for hNEU2 can be achieved with modified *N*-acyl groups bearing large triazole-linked derivatives<sup>97</sup>. Conversely, addition of C9 substitution such as triazole-linked biphenyl or amide-linked aliphatic chains can significantly disfavour inhibition of hNEU1 or hNEU2<sup>96</sup>. Therefore using our acquired knowledge about TcTS and the literature information regarding hNEUs, the appropriate selection and combination of such modifications could afford DFSAs with selectivity for TcTS over endogenous neuraminidases.

## 2.6 Conclusions

In an attempt to harness new active site interactions to enhance binding with TcTS two sites of modification were selected for chemical functionalization of the known sialidase inactivator difluorosialic

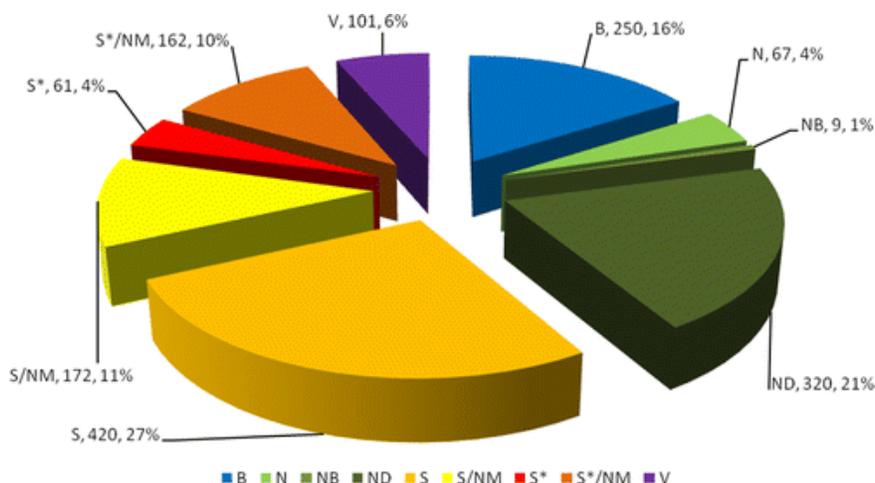
acid. Inspired by previous studies, DFSA was derivatized by amide coupling at C9 and C5 to yield libraries of nine and seven members respectively. The C9 amide-linked cyclopentyl analogue **2.49** had best-in-class inactivation of TcTS, yet highly related 5- or 6- membered cyclic analogues failed to achieve similar results. The only anionic derivative **2.56** in the library had above average inhibition of TcTS, which is consistent with previous reported results for negatively charged 9-triazole linked DFSA derivatives. A series of linear, branched or cyclic *N*-acyl derivatives of DFSA were synthesized and evaluated at a single concentration versus TcTS. While the results indicated modest selectivity for the butyryl and cyclopentanoyl modifications, all species exhibited comparable inactivation rates to the natural *N*-acetyl species. Although new beneficial binding interactions were not discovered, a notable tolerance by TcTS to *N*-acyl modifications of DFSA inactivators was observed.

Overall, we sought to increase potency and selectivity of the DFSA analogues for TcTS. Our best DFSA analogue **2.49** had a two-fold increase in inactivation rate constant versus TcTS relative to the previous champion (**1.25**) and a 10-fold increase over the unsubstituted DFSA derivative (**1.21**). Future attempts to increase DFSA potency should focus on expanding the scope of anionic derivatives around C9 to harness the suggested ionic interactions. Alternatively, the synthesis of 3-fluorosialosides with larger activated aglycones, that can simultaneously harness interactions in the acceptor site, may enhance binding with TcTS. Additionally, we discovered that TcTS is tolerant to *N*-acyl modification of DFSAs. Based on published reports that have investigated hNEU isoform tolerance to inhibition by substituted DANA analogues, we propose that simultaneous modification of DFSA at the C9 and C5 position may yield inactivators that are specific for TcTS relative to endogenous neuraminidase. Further investigations to this end are warranted.

## Chapter 3: Marine Natural Product Screening and Bioassay Guided Isolation of Inhibitors for TcTS

### 3.1 Introduction

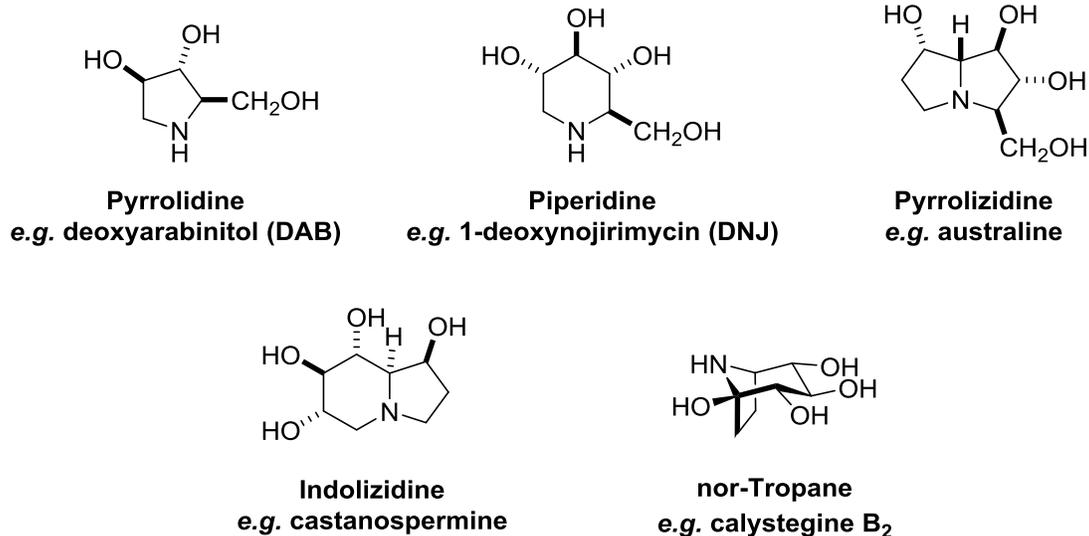
Decades of research on TcTS inhibition has focused on the rational design and development of substrate-based inhibitors. However these approaches have failed to achieve any compounds with inhibition constants  $<100 \mu\text{M}$ . Therefore, exploratory methods such as functional screening are necessary to find new chemical scaffolds that can elicit the desired enzyme inhibition. Libraries of isolated natural products or environmental extracts serve as a source of vast chemical diversity that can be mined for such a purpose. Bioactive materials isolated from plants and/or microorganisms are considered the foundation of modern medicine, since many of the earliest drugs were directly isolated from such sources (*e.g.* morphine, penicillin) or synthesized using natural product structures as inspiration (*e.g.* aspirin). This class of molecules remain highly important today. A comprehensive report of the commercial therapeutics developed between 1981 and 2015 revealed that ~42% of the 1562 chemical entities approved for medical use are from natural sources<sup>116</sup>. As shown in Figure 3.1 biologicals (*e.g.* large peptides; B, 16%), unmodified natural products (N, 4%) and natural product derivatives (ND, 21%) make up the major categories of these compounds. Furthermore of the 815 synthetic or semi-synthetic therapeutics ~50% employ a pharmacophore of natural origin. Cancer chemotherapies are an exemplary case study, where only 17% of the approved small-molecule drugs are classified as synthetic, meaning 83% of approved anti-cancer drugs are natural products or derivatives/mimics thereof.



**Figure 3.1. FDA Approved drugs 1981-2014. B: Biological, N: Natural product, NB: Natural botanical, ND: Natural product derivative, S: Synthetic, S\*: Semi-synthetic, V: Vaccine, /NM: Natural product mimic.**  
Source: Newman<sup>116</sup>.

A long-standing survival approach for immotile and vulnerable prey has been to employ combinations of bioactive metabolites as their primary defense system. On land, plants are an abundant, immobile, carbohydrate-rich source of nutrition; these complex glycoside polymers are thus a primary food source for organisms that express the necessary machinery to digest them. Therefore some plants use chemicals that can inhibit enzymatic degradation or elicit toxic effects to guard against predators. In the ocean, marine invertebrates (*e.g.* sponges) suffer from analogous defense limitations, making them an easy target for predation. They too have evolved a chemical-based defense system whereby bioactive metabolites, or chemicals sequestered from associated bacteria, are excreted against foreign attackers. The characterization of these molecules and the organisms that produce them has been the basis of many academic and commercial investigations alike.

### 3.1.1 Natural carbohydrate-based glycosidase inhibitors



**Figure 3.2. Major classes of alkaloid-based glycosidase inhibitors.**

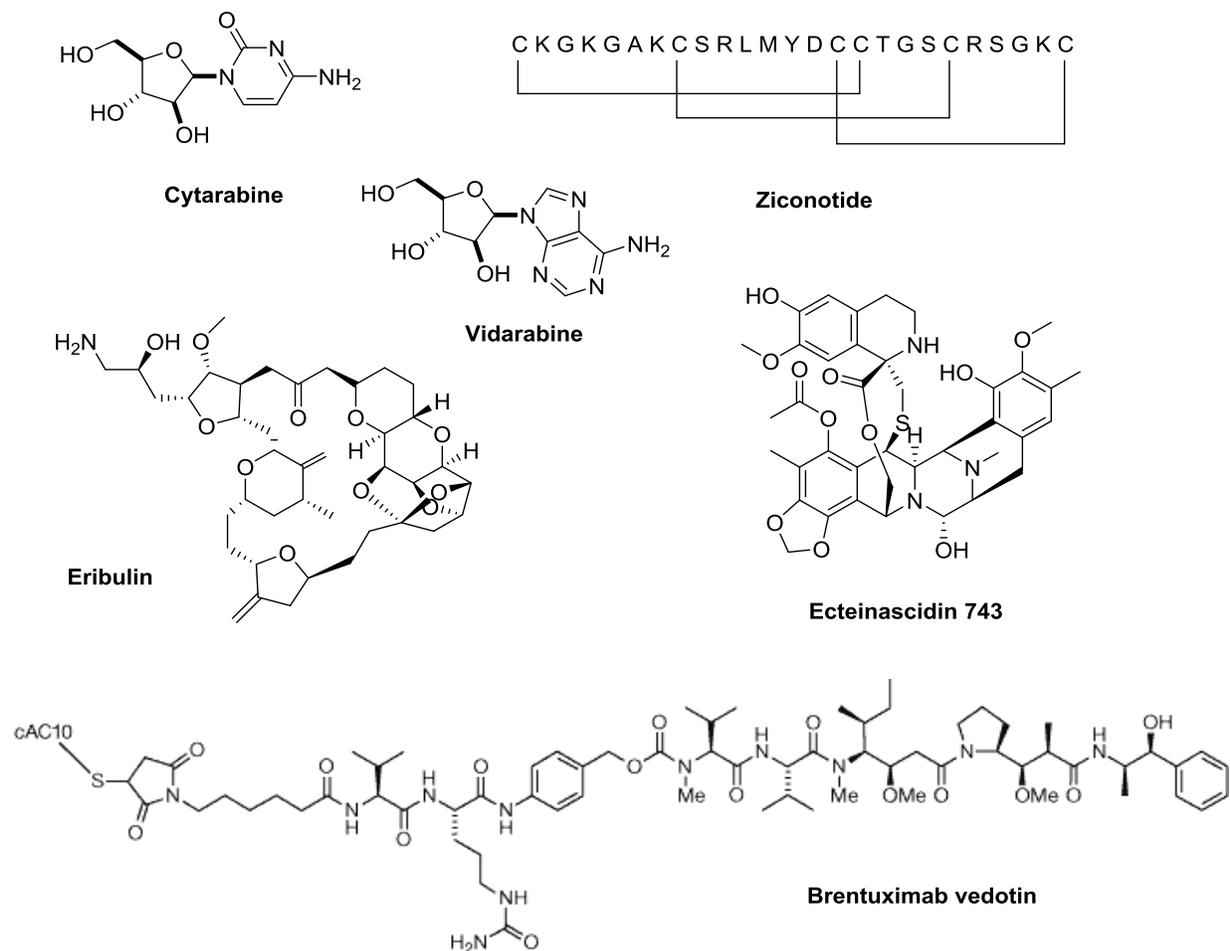
Among the chemicals identified from terrestrial sources, glycosidase inhibitors have been particularly prevalent. In plants and associated microbes, polyhydroxylated nitrogen-containing organic compounds – generally termed alkaloids (or aza/imino sugars) – have been routinely isolated and are grouped into five major categories based on ring size and structure (Figure 3.2). These nitrogenous heterocycles are convincing carbohydrate mimics as evidenced by their ability to both form strong interactions with sugar receptors in biological systems and act as potent glycosidase inhibitors<sup>117</sup>. For example, deoxynojirimycins (DNJs) – a subclass of polyhydroxylated piperidine alkaloids – are potent inhibitors of  $\alpha$ -glucosidase activity with, typically, micro- to nanomolar inhibition constants. The breakdown of starch polymers requires the enzymatic cleavage of  $\alpha(1,4)$ -glucoside linkages by such glucosidases; therefore the release of DNJs in starch-rich environments can disrupt the digestion of the glucose polymer. Plants leverage this activity by employing combinations of alkaloid metabolites to elicit broad scope enzyme inhibition as protection against predation. Medically, DNJs have gained recognition for their therapeutic potential in a range of diseases including cancer, viral infections<sup>118</sup> and diabetes<sup>119</sup>.

Similar glycoside-based natural product inhibitors have been identified for many carbohydrate active enzymes, including sialidases. The transition-state analogue Neu5Ac2en (*i.e.* DANA; discussed in Chapter 1 and Chapter 2) is one such example. Although DANA was first described in 1969 by Meindl and Tuppy<sup>120</sup> – following its synthetic preparation from tetra-*O*-acetyl-2-chloro-2-deoxy-Neu5Ac – it was discovered in 1975 that DANA was present in the urine, saliva and blood serum of a sialuria-affected patient, thus indicating a biological origin of the neuraminidase inhibitor<sup>121</sup>. It has since been shown that some neuraminidases can perform an elimination reaction (rather than hydrolysis) on the covalent sialosyl-enzyme intermediate during catalysis to produce Neu5Ac2en in biological systems<sup>122</sup>. Although DANA exerts potent inhibition against many members of the sialidase superfamily it requires millimolar concentrations of material to affect TcTS activity.

### 3.1.2 Marine natural products

Terrestrial organisms have been the primary source of natural product research and for centuries, extracts of this nature have been heavily investigated for therapeutic development. Yet, oceans cover >70% of the earth's surface and to date comprise >230,000 unique organisms<sup>123</sup>, and experts predict the actual number to surpass 1,000,000 species<sup>124</sup>. Despite this, serious investigations into marine natural products were only spawned in the 1950s following Bergmann's discovery of non-ribose nucleosides – spongothymidine and spongouridine – from the marine sponge *Cryptotethia crypta*<sup>125</sup>. That report birthed new fields of study – focused on marine toxins, marine natural product chemistry and marine chemical ecology – which have since been heavily investigated<sup>126</sup>. Marine sponges and the associated microorganisms are now recognized to be among the richest environmental sources of pharmacologically-active chemicals<sup>127</sup>. Through 2016 an estimated 28,500 marine natural products have been identified, with anticancer and cytotoxic properties accounting for the majority of the reported biological activities<sup>128</sup>. The most promising of these chemicals are currently in clinical trials, while six marine natural products (or their close derivatives) have already been approved by the FDA for therapeutic use (Figure 3.3):

cytarabine (*ara-C*; anticancer; 1969), vidarabine (*ara-A*; antiviral, 1976), ziconotide (Prialt; pain control, 2004), ecteinascidin 743 (Trabectedin; anticancer, 2015), eribulin mesylate (Halaven; anticancer, 2010) and brentuximab vedotin (Adcetris; anticancer, 2011).

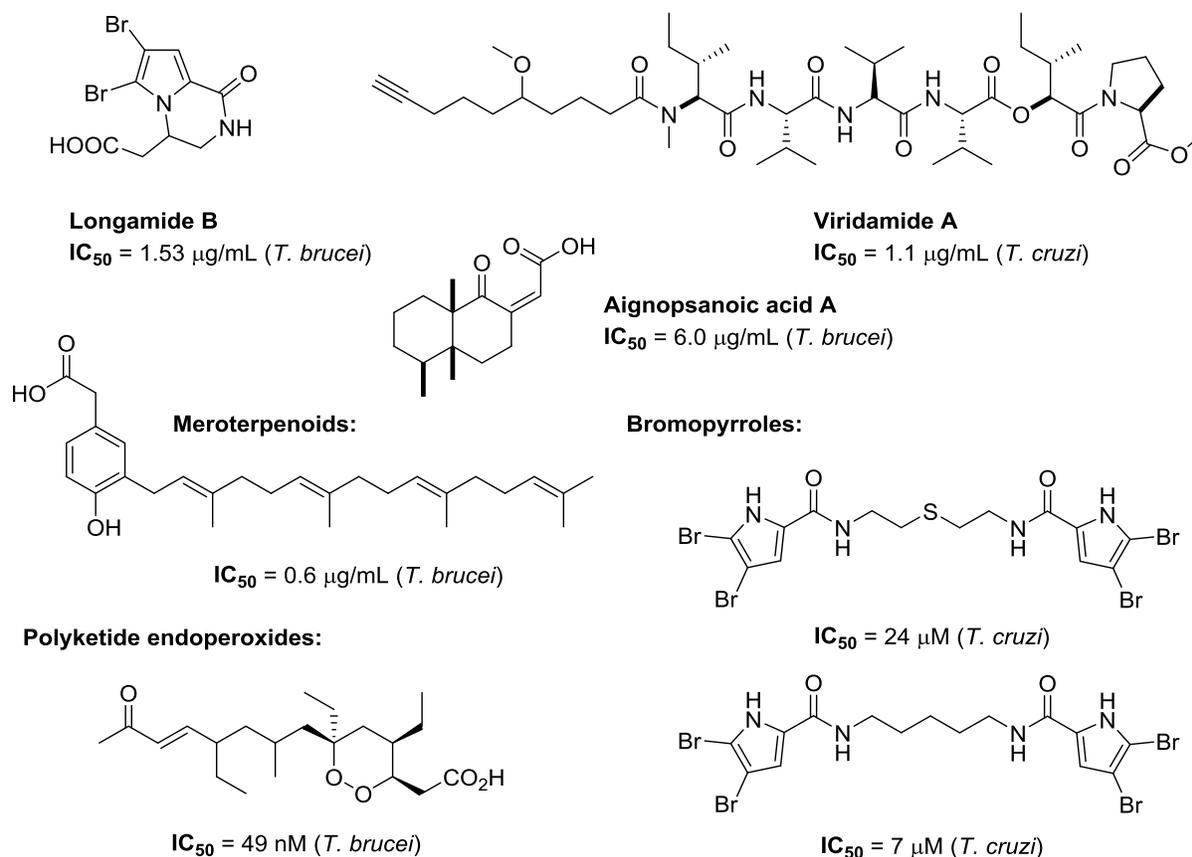


**Figure 3.3.** Marine natural products as FDA approved drugs.

Bioactive molecules from marine invertebrates have also been shown to elicit strong antiparasitic activity or to be potent enzyme inhibitors, both of which are of particular interest to our research.

Antitrypanosomal activity has been reported against *T. cruzi* and *T. brucei* by a handful of these natural product species such as alkaloids, terpenes, and endoperoxide derivatives<sup>129</sup> (Figure 3.4). Further, two bromopyrrole alkaloids were identified recently from the sponge *Tedania brasiliensis* that had specific

antiparasitic activity versus *T. cruzi* with effective concentrations of  $EC_{50} = 7 \mu\text{M}$  and  $24 \mu\text{M}$  respectively<sup>130</sup>. While marine natural products with antiparasitic activity against *T. cruzi* have been identified the mechanism of action often remains unknown which can significantly hamper the therapeutic utility of these molecules. Therefore researchers tend to direct their efforts towards inhibitors of defined, therapeutic relevant molecular targets.



**Figure 3.4. Structures and antitrypanosomal activity of select marine natural products.**

A comprehensive compilation of such compounds was assembled in 2012 and describes a diverse set of >370 metabolites isolated from marine invertebrate extracts that exhibit inhibition against a wide range of enzyme classes: oxidoreductases, transferases, hydrolases, lyases and isomerases<sup>131</sup>. Among the hydrolase-inhibiting species, both glycoside and non-glycoside-based compounds have been identified that inhibit carbohydrate active enzymes (*e.g.* chitinase,  $\alpha$ -glucosidase and sialidase). Specifically,

inhibitors of viral neuraminidase as well as bacterial sialidases from *Clostridium perfringens*, *Vibraea chloae* and *Staphylococcus typhimurium* have been isolated that exhibit high micromolar to mid nanomolar potencies (Figure 3.5). These compounds highlight the rich diversity of bioactive chemicals employed by marine organisms and the largely untapped therapeutic potential they encompass. The discovery of chemicals from sponge extracts that exert specific and potent enzyme inhibition makes organisms of this type an attractive area of study for new natural product motifs and for discovery of enzyme inhibitors.

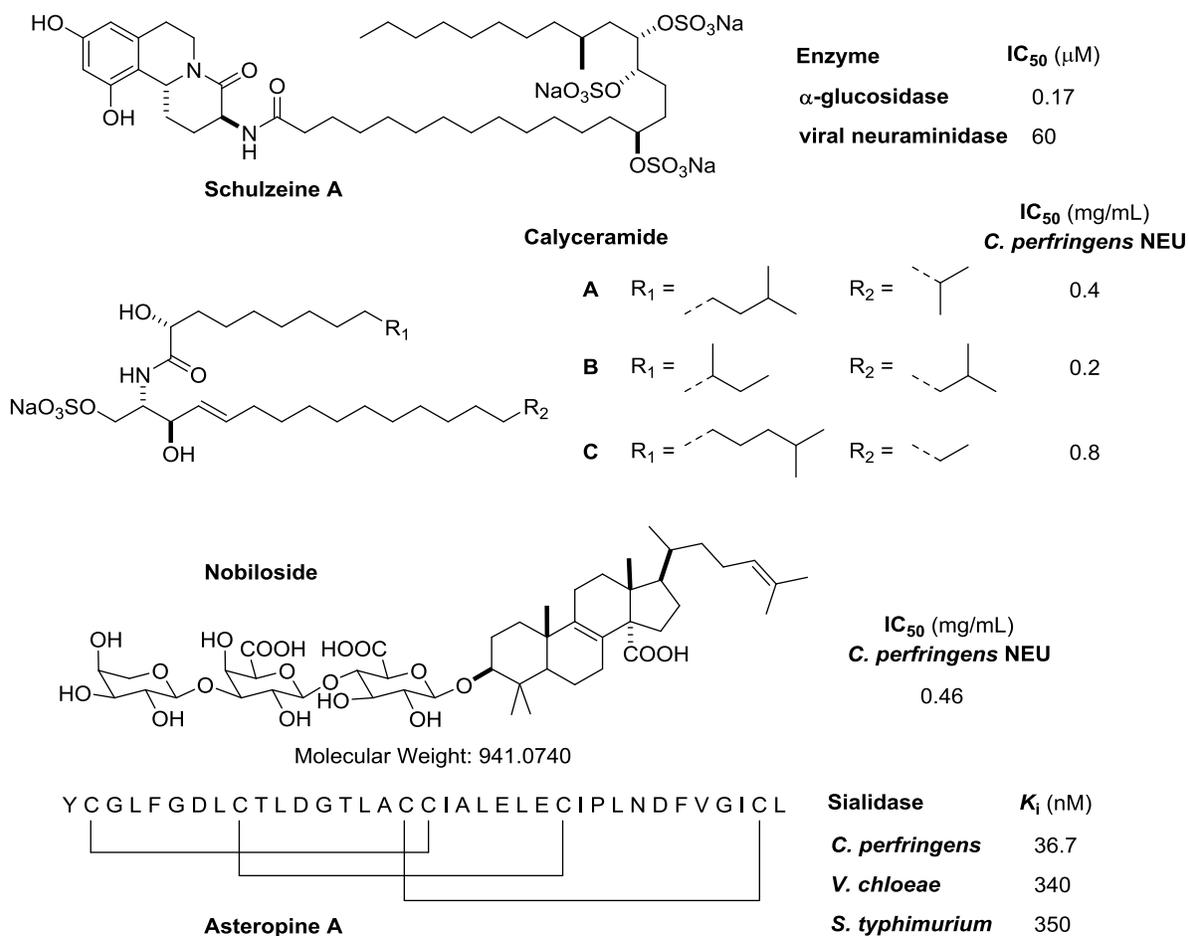


Figure 3.5. Marine natural products as sialidase inhibitors.

### 3.1.3 Screening natural product libraries for glycosidase inhibitors

A glaring need to find new classes of TcTS inhibitors has been clearly demonstrated. While many bioactive inhibitors exhibit structural features that suggest an evolutionary development for the role (*e.g.* DNJ, DANA, Acarbose; Figure 3.6), this relationship is not required. Compounds bearing no structural resemblance to endogenous substrates, nor showing any indication of evolved functionality against the target enzyme, can also exhibit the desired activity. Thus screening natural product libraries, regardless of source, presents the possibility of finding new chemical scaffolds for enzyme inhibition. An illustrative case study below describes the successful use of screening environmental libraries to find new glycoside and non-glycoside based therapeutic leads that target the ever-growing diabetes and obesity epidemic.

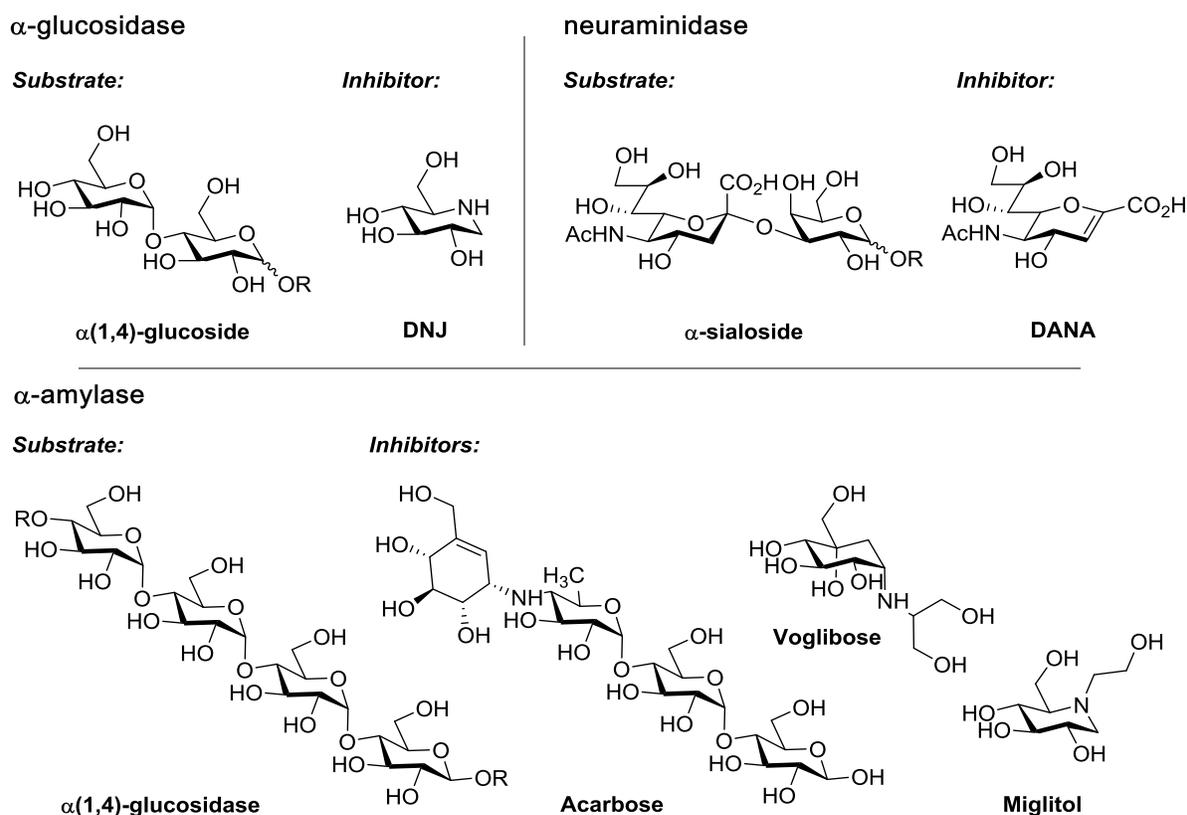
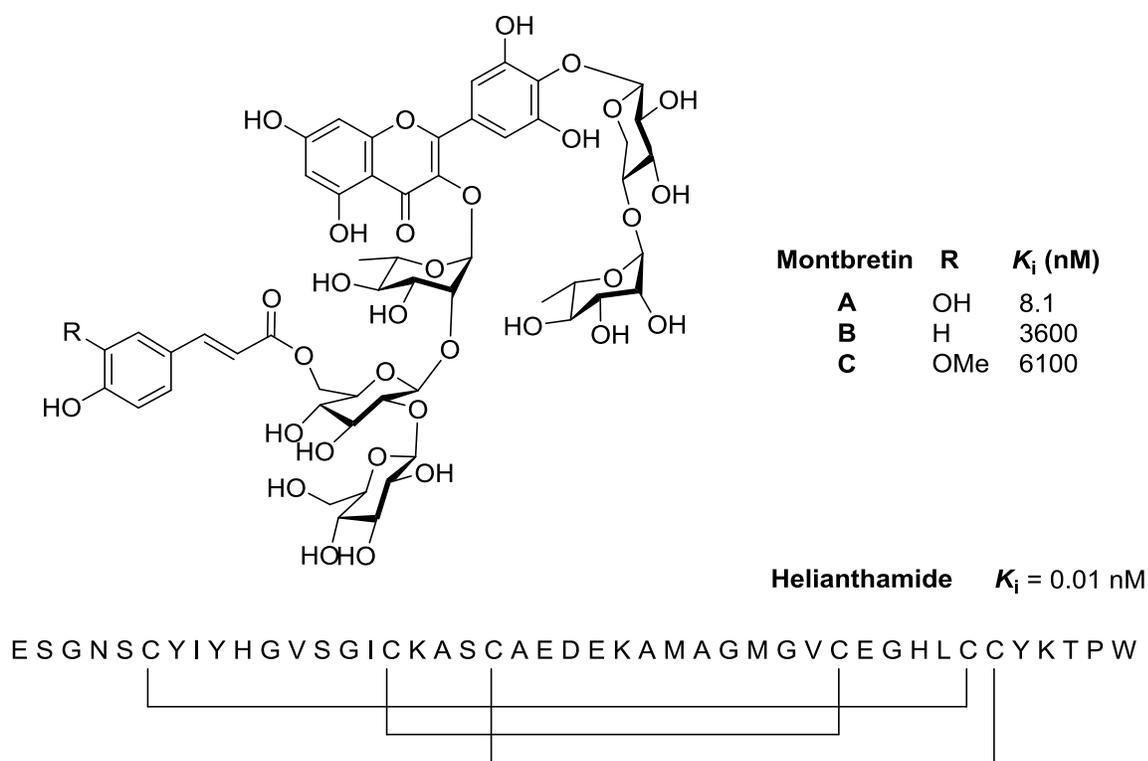


Figure 3.6. Structural comparison of substrate and natural carbohydrate-based inhibitors of  $\alpha$ -glucosidase (left),  $\alpha$ -amylase (bottom) and neuraminidase (right).

### 3.1.4 Case study: screening natural product libraries for $\alpha$ -amylase inhibitors

Type 2 diabetes mellitus is caused by insufficient insulin production in the body, leading to individuals who cannot endogenously manage their blood sugar levels. Therefore diabetics must closely monitor and control their blood sugar levels with medical treatments as well as lifestyle and dietary changes. Three natural product-based drugs have been developed as oral therapeutics to this end (*i.e.* acarbose, miglitol, voglibose) all of which exert their function by inhibiting  $\alpha$ -glucosidase activity to slow the digestion and absorbance of dietary carbohydrates. Ubiquitous acceptance of  $\alpha$ -glucosidase inhibitors as a therapy for diabetes has been hampered by unpleasant side effects associated with these drugs leading to poor patient compliance<sup>132,133</sup>. This has sparked the investigation into alternative therapeutic targets in the starch breakdown pathway such as human pancreatic amylase (HPA). HPA acts on dietary starch upstream of the  $\alpha$ -glucosidase breakdown which is hypothesized to reduce the gastrointestinal side-effects associated with existing treatments. Inhibitor development towards HPA had been largely unsuccessful until the recent discoveries of potent and selective natural product inhibitors by the Withers lab from plant and marine sources (Figure 3.7).



**Figure 3.7. Natural product inhibitors of human pancreatic amylase.**

Screening for inhibitors of HPA catalysis against a library of 30,000 natural product extracts (National Cancer Institute, U.S.) from terrestrial and marine sources afforded a family of glycosylated flavonol derivatives, Montbretins A-C (MbX). The most active of the isolated species was MbA which inhibited the target enzyme with a high affinity<sup>134</sup> ( $K_i = 9$  nM). Moreover a small peptidic inhibitor – Helianthamide – was separately discovered from a library of 10,000 marine natural product extracts (assembled by the Andersen lab) and demonstrated sub-nanomolar inhibition ( $K_i = 10$  pM) of the HPA target<sup>135</sup>. Helianthamide was identified after bioactivity-guided isolation from the Caribbean sea anemone *Stichodactyla helianthus*. These two novel  $\alpha$ -amylase inhibitors have since garnered considerable commercial<sup>136</sup> and academic interest<sup>137–139</sup>. These findings highlight the power of screening terrestrial or marine-based libraries of natural product extracts for inhibitors of elusive glycosidase targets.

### 3.1.5 Screening for TcTS inhibitors

As detailed in Chapter 1, two screening approaches have been reported in search of TcTS inhibitors. One of which was a commercial library of ~2300 unique natural products was screened that yielded hits with low to sub micromolar inhibition of TcTS. The two most promising classes of identified inhibitors – flavanols and anthraquinones –exerted non-competitive inhibition and latent cytotoxicity during *in vivo* studies. Both reported screening attempts led to the discovery of new inhibitor scaffolds with better or comparable affinity to the best glycoside mimetics developed by rational design, thus providing the motivation for additional screening efforts to be pursued in this vein.

Given the aforementioned success of screening natural product libraries for potent glycosidase inhibitors and the previous success in screening for TcTS inhibitors, we decided to pursue a screen of our own. Collaboration with the research group of Dr. Raymond Andersen afforded us access to both the UBC Marine Natural Products Extract Library and the expertise their lab provides in bioassay guided natural product isolation and characterization.

## 3.2 Specific aims

The primary goal of this chapter is to identify new chemical scaffolds that inhibit TcTS activity. We aim to screen the UBC Marine Natural Product library to find sponge extracts that exhibit the desired activity. We will then use bioassay-guided separation techniques to isolate the active species and structural analysis methods to identify the compounds responsible for the activity. Subsequently, kinetic studies and structure-activity relationships will be employed to characterize the lead compounds. From these results, we can select the best natural products or their derivatives for further analysis and development.

## Results and Discussion

### 3.3 Screen design and preparation of extract library

#### 3.3.1 Screen design and validation

Our first step was to develop and validate an appropriate bioassay for TcTS inhibition that is amenable to automated screening. While TcTS preferentially transfers sialic acid to a  $\beta$ -galactosyl acceptor (*e.g.* lactose) it can also act as a sialidase by hydrolyzing sialic acid from sialylglycoconjugates. The same active site has been shown to be responsible for both catalytic activities<sup>140</sup>, thus TcTS activity has typically been quantified by either a trans-sialylation assay using radiolabeled substrate or by monitoring the cleavage of chromophore-generating sialoside substrates. The regulatory burden of radioactive methods as well as the technical ease of plate-based screening with UV/Vis detection has significantly popularized the latter method. Activated sialoside derivatives bearing fluorogenic leaving groups such as methyl umbelliferone (MU) or trifluoromethyl umbelliferone (TFMU) have been most often used to this end.

Our plate-based screening method was developed using the more reactive TFMU-sialoside substrate, while continuous release of the cleaved chromophore was monitored by UV/Vis spectroscopy at  $\lambda = 405$  nm. Although UV/Vis spectroscopy is generally less sensitive than fluorescence detection, the ability to directly monitor the initial rates of substrate cleavage was more appealing than the alternative stopped-assay approach that is necessarily employed in fluorescence-based techniques<sup>141</sup>. Thus the general screening protocol was designed such that 96-well plates would be filled with buffered enzyme solution and incubated with DMSO-solubilized natural product extracts for 10 minutes. The assay would then be initiated by addition of substrate solution and the rate of chromophore release monitored by UV/Vis spectroscopy. Linear regression of the initial rates would afford the initial rates of substrate cleavage by TcTS. The percent inhibition by each extract sample was determined by comparison to control activity.

The proposed screening method was assessed by Z-factor analysis to ensure the reproducibility of the devised assay. To this end, three reduced volume 96-well plates (Costar) were assayed containing alternating rows of positive and negative controls. The positive wells contained enzyme (TcTS), buffer and DMSO, while the negative wells contained only buffer and DMSO. Normally the enzyme and a potent enzyme inhibitor would be used in the negative control rather than complete omission of the enzyme; however, in this case no such compound is available. The assay was initiated by addition of TFMU-SA and the rate of chromophore release was recorded for 15 minutes. Zhang and colleagues have developed and extensively described standard methods for assessing the quality of an assay during development. In our case, the Z'-factor value can be calculated (Equation 3.1) and interpreted to determine the suitability of our assay for our intended purpose<sup>76</sup> (Table 3.1). A Z'-factor of 0.7 was determined from the measured values of the control plates for TcTS activity, thus indicating the employed assay conditions are of 'excellent' quality.

**Equation 3.1. Z'-factor.**

$$Z' = 1 - \frac{3SD_{\text{positive}} + 3SD_{\text{negative}}}{\text{mean}_{\text{positive}} - \text{mean}_{\text{negative}}}$$

**Table 3.1. Z'-factor analysis and interpretation<sup>76</sup>.**

<b>Value</b>	<b>Assay Structure</b>	<b>Meaning</b>
$Z' = 1$	Dynamic range $\rightarrow$ infinity	Perfect assay
$1 > Z' \geq 0.5$	Large separation band	Excellent assay
$0.5 > Z' > 0$	Small separation band	Double assay
$Z' = 0$	No separation band	Yes/No assay
$Z' < 0$	No separation band	Unusable

### 3.3.2 Preparation of extract library

The UBC Marine Natural Products Extract Library was assembled from marine invertebrates collected by the Andersen lab and has been curated over the past ~30 years. The library subset provided for screening in this report consists of methanolic extracts of 986 such organisms distributed across twelve 96-well plates. The master plates of these extracts (containing an estimated 5 mg/well) were solubilized in DMSO (100  $\mu$ L) and subsequently replicated in 20-, 400- and 4000-fold dilutions. Two of the twelve plates (plates 9 and 10) were assayed for TcTS inhibition at each dilution to gauge the appropriate concentration for use in library screening (Table 3.2).

**Table 3.2. Effect of natural product extract dilution on hit rate for TcTS inhibition.**

<b>Plate 9</b>			<b>Plate 10</b>		
<b>Extract Dilution</b>	<b>Hits* (#)</b>	<b>Hit Rate (%)</b>	<b>Extract Dilution</b>	<b>Hits* (#)</b>	<b>Hit Rate (%)</b>
<b>10</b>	67	84	<b>10</b>	65	81
<b>200</b>	59	74	<b>200</b>	50	63
<b>4,000</b>	13	16	<b>4,000</b>	8	10
<b>40,000</b>	1	1.3	<b>40,000</b>	1	1.3

\*Hits are defined as wells with >75% inhibition of TcTS relative to controls.

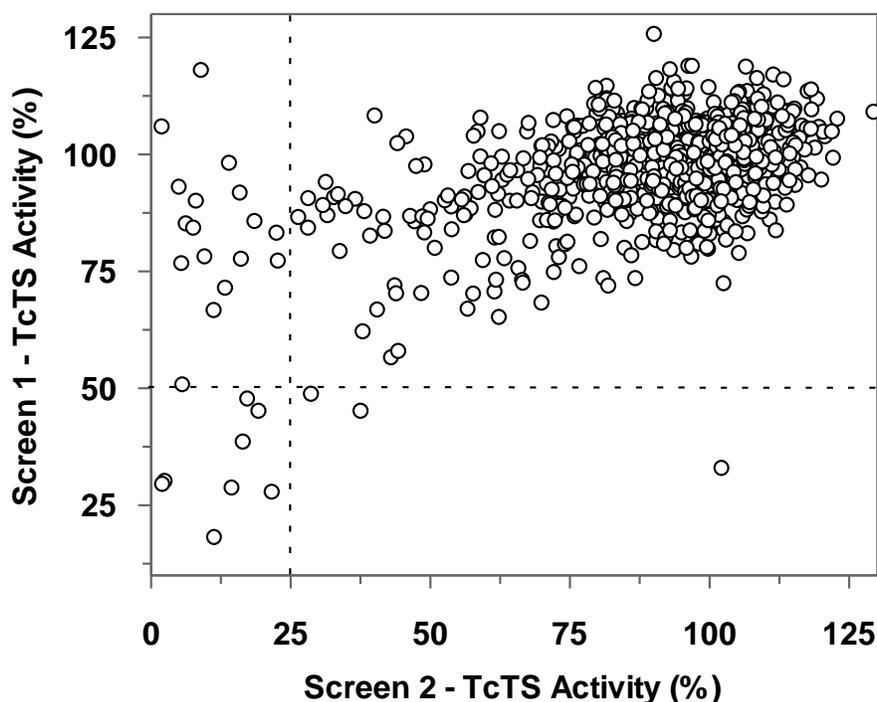
Our hit threshold was defined at >75% inhibition versus average activity of TcTS control wells. The highest concentrations of the extract library afforded near complete suppression of TcTS activity (hit rate >80%), whereas 4,000- and 40,000-fold dilution of the natural product extracts afforded hit rates of  $\leq$ 16% per plate. Extract dilution by 40,000-fold yielded 2 total hits (1.3% hit rate). Projected across the complete library this would yield ~1 hit per plate (~12 total hits). A hit rate of ~1% should allow for the

most active extracts to be selected while providing a manageable number of samples for secondary analysis. Therefore, the extract dilution of 40,000-fold was selected for primary screening.

### 3.4 Screening marine organism extracts for TcTS inhibition

#### 3.4.1 Screening results and hit selection

With an appropriate library and a robust assay in-hand 986 unique marine sponge extracts were screened for inhibition of TcTS activity. Specifically, buffered TcTS (55 nM, 70 $\mu$ L) was incubated with 4,000-fold diluted DMSO-solubilized extract solution (10  $\mu$ L) for 10 minutes and the residual TcTS activity was monitored by UV/Vis spectroscopy ( $A_{405}$ ) after addition of buffered substrate solution (TFMU-SA; 0.5 mM; 20  $\mu$ L). The primary screen was performed twice on a Beckman Coulter Biomek FX Laboratory Automation Workstation (Fullerton, CA, USA) integrated with a Beckman Coulter DTX880 plate reader with UV/Vis capability.



**Figure 3.8.** Data set from two plate-based screens of UBC marine natural product extract library. Dashed lines indicate hit threshold for each screen, where hit samples selected for further testing are near the origin. Data is presented as percent residual TcTS activity relative to uninhibited controls.

The first library screen afforded 25 hits (>75% inhibition) corresponding to a hit rate of 2.5%. However, analysis of the transfer/mixing settings employed by the automated liquid handler during library dilution raised concerns that the ‘concentration’ of the 4,000-fold dilution library may be inflated, yielding higher than expected concentrations of extracts. Thus a second 4,000-fold dilution library was freshly prepared (after altering the transfer method) and a second round of screening was performed. In this case a single sample had >75% inhibition of TcTS activity, while 11 samples had >50% inhibition (1.1% hit rate). The percentage of residual TcTS activity – relative to uninhibited controls – in each well were plotted against each other for both screens to afford a 2D plot (Figure 3.8). Hits from eight wells were selected based on their performance in both screens (Table 3.3; >75% inhibition in Screen 1 and >50% inhibition in Screen 2), thereby decreasing the likelihood of choosing false positives.

**Table 3.3. Hit samples selected from the marine natural product extract screen.**

<b>Hit (Plate/Row/Col)</b>	<b>% TcTS Inhibition (Screen 1)*</b>	<b>% TcTS Inhibition (Screen 2)**</b>
<b>3H3</b>	81	55
<b>4A8</b>	89	82
<b>4D7</b>	97	70
<b>4D8</b>	78	72
<b>4H5</b>	98	71
<b>6C11</b>	83	52
<b>9C6</b>	85	71
<b>10E10</b>	83	62

\*Selection threshold >75% enzyme inhibition; \*\*Selection threshold >50% enzyme inhibition

### 3.4.2 Hit validation and TcTS selectivity

In order to confirm the activity of the selected extracts and to further refine the number of hits prior to bioassay guided isolation, secondary assessments were performed. To confirm the observed inhibition and to assess whether the mode of inhibition is exerted in a typical concentration dependent fashion, a dilution series (5 mg/mL to 0.5 µg/mL) of each hit was assayed for TcTS inhibition. All extracts demonstrated a sigmoidal concentration dependence and estimated IC<sub>50</sub> values were determined (Table 3.4). To estimate these values, two assumptions were made: (1) the amount of initial sponge extract in all wells was 5 mg and (2) that the abundance of inhibitor was consistent in each well. While neither of these assumptions are likely to be correct, they provide a practical way to standardize and compare the observed results. Given these assumptions, and since each of the resultant IC<sub>50</sub> values were of the same order of magnitude, the eight hits were considered equally promising at this point.

**Table 3.4. Inhibition values for hit samples selected from the marine natural product extract screen.**

Hit (Plate/Row/Col)	IC <sub>50</sub> (µg/mL)*	Hit (Plate/Row/Col)	IC <sub>50</sub> (µg/mL)*
<b>3H3</b>	46 ± 3	<b>4H5</b>	30 ± 1
<b>4A8</b>	35 ± 3	<b>6C11</b>	31 ± 2
<b>4D7</b>	25 ± 2	<b>9C6</b>	12 ± 1
<b>4D8</b>	36 ± 2	<b>10E10</b>	26 ± 2

\*Calculated based on assumption that initial mass of solid crude extract was 5 mg/well.

Next we wanted to assess the selectivity of each extract for TcTS. It is possible that the lead extracts may contain chemical species that exhibit nonspecific inhibition towards TcTS (*e.g.* pan-assay interference compounds), or they could contain a molecule that has promiscuous glycosidase and/or sialidase inhibition. Thus, two additional enzymes were chosen to help narrow down the selected hits: cytosolic human neuraminidase (hNEU2) and β-galactosidase from *Aspergillus oryzae* (βGal). hNEU2

was selected as a representative for endogenous human neuraminidases (which we do not want to inhibit), while  $\beta$ -galactosidase was chosen to control for nonspecific inhibitory interactions.

Assay conditions were developed for each of the enzymes.  $\beta$ -Galactosidase was assessed phosphate buffer (200 mM pH = 6.5) and its activity monitored by UV/Vis spectroscopy while employing 4-methylumbelliferyl  $\beta$ -D-galactopyranoside (MU-Gal; 0.5 mM) as a substrate. hNEU2 was tested using a fluorescence assay in sodium acetate buffer (100 mM; pH = 5.5) with TFMU-SA (0.8 mM). Plate 4 of the extract library contained the most samples with TcTS inhibition and was selected to be assayed for inhibition of  $\beta$ Gal activity (Figure 3.9). Satisfyingly, none of the hit wells versus TcTS showed any inhibition of  $\beta$ Gal activity at the same concentration. Some hit wells for  $\beta$ Gal had mild inhibition of TcTS activity.

beta-Gal	Ctrl (+)	2	3	4	5	6	7	8	9	10	11	Ctrl (-)
A	-4	-7	-20	-4	-10	6	-11	-7	-14	-3	-11	100
B	-5	-9	27	-9	-10	-20	-15	-7	-5	-8	-9	99
C	0	-8	-7	-8	-13	-10	-6	-4	-1	-8	-18	90
D	-6	-3	-17	-14	-8	-1	-1	-7	-5	-6	-13	94
E	-6	13	-8	-6	42	-10	-4	-6	-10	38	-8	94
F	3	-3	-5	-5	100	-8	-4	2	-5	46	-5	100
G	-3	-2	-15	-4	-3	-5	-11	17	-9	-7	-2	95
H	-11	-9	-4	-12	-1	-2	-6	-7	-8	-9	-1	96
TCTS	Ctrl (+)	2	3	4	5	6	7	8	9	10	11	Ctrl (-)
A	-21	7	-15	-9	12	-29	-22	89	-20	16	-15	100
B	-26	-11	-38	-3	5	-9	-14	24	-10	-11	-5	99
C	6	12	46	14	15	12	50	22	20	38	18	99
D	7	20	8	12	21	18	97	78	19	51	20	98
E	9	4	29	20	9	8	0	19	23	4	59	100
F	4	25	34	11	67	4	10	1	3	-4	14	99
G	9	44	22	6	19	13	8	6	33	18	12	96
H	1	7	17	23	98	40	13	24	71	30	98	100

**Figure 3.9. Percent inhibition of  $\beta$ Gal from *aspergillus oryzae* (top) or TcTS (bottom) by samples from plate 4 of the marine natural product extract library.** Extracts were diluted 40,000-fold; the controls represent measurements conducted without inhibitor (+) or without enzyme (-). Colour intensity scales with percent inhibition.

The selectivity of our lead extracts was assessed by testing for inhibition of hNEU2,  $\beta$ Gal and TcTS activity at concentrations 40-fold higher than the original screening conditions (Table 3.5). While all eight species demonstrated complete inhibition of TcTS activity, three extracts also demonstrated  $\geq 20\%$  inhibition of  $\beta$ Gal activity (4A8, 4D7, 4H5) and were eliminated from further investigation. hNEU2 activity was not inhibited more than 11% at this concentration, which boded well for the potential selectivity towards TcTS by the lead extracts.

**Table 3.5. Percent inhibition of glycosidase activity by selected hits from natural product extract screening.**

<b>Hit (Plate/Row/Col)*</b>	<b>hNEU2 Inhibition (%)</b>	<b><math>\beta</math>Gal Inhibition (%)</b>	<b>TcTS Inhibition (%)</b>
<b>3H3</b>	-0.6	1.5	100
<b>4A8</b>	1.6	20	100
<b>4D7</b>	-6.3	22	96
<b>4D8</b>	8.3	6.7	99
<b>4H5</b>	3.3	23	100
<b>6C11</b>	5.4	9.3	99
<b>9C6</b>	11	-2.6	97
<b>10E10</b>	6.2	4.7	99

\*Hit extract samples were assessed at 1,000-fold dilution.

### **3.5 Isolation and structural determination of hits from marine organisms**

Primary and secondary screening had narrowed the initial 986 sample library to five lead candidates. At this point, bioactivity-guided (*i.e.* TcTS inhibition) isolation of the active components from the marine organisms was pursued in collaboration with the Andersen lab. First, I was provided with a few hundred milligrams of each of the methanolic sponge extracts (3H3, 4D8, 6C11, 9C6, 10E10) to perform trial scale investigations. The samples were subjected to iterative orthogonal partitioning and subsequent assessment for TcTS inhibition. This process was conducted until either the bioactive chemical was isolated or there was an insufficient amount of material to characterize.

The general separation process began with liquid-liquid partitioning of the methanolic extract between an aqueous and organic phase (ethyl acetate or *n*-butanol). The organic component would then be fractionated by size exclusion chromatography using Sephadex LH-20 resin. Selected fractions would be subsequently partitioned by reverse phase chromatography over C18 silica gel and finally purified by

HPLC. If the results from the trial scale isolations were promising, additional sponge material was extracted following the same protocols. Following each stage of purification the solvent was evaporated and the extract material was assayed for TcTS inhibition in an analogous manner to the assay employed during screening. Fractionated extracts were dissolved in DMSO to a stock concentration of 1 mg/mL, diluted 10-, 100- and 1000-fold, and each was incubated with TcTS. The residual enzymatic activity was monitored by UV/Vis spectroscopy and compared versus controls. The materials that showed the highest levels of TcTS inhibition at the lowest concentration of extract were selected for additional separation at each step.

First, each of the methanolic extracts were separated by liquid-liquid extraction and assessed for enzyme inhibition (Figure 3.10). The most potent inhibition was observed in the ethyl acetate extracts from samples 03-242 (4D8), 03-395 (9C6) and 00-287 (6C11), which completely blocked TcTS activity at 0.1 mg/mL while reducing activity by ~50% at 0.01 mg/mL. Samples 97-199 (10E10) and 11-312 (3H3) were less active and the majority of the material was partitioned into the aqueous layer. Thus insufficient quantities of material and poor TcTS inhibition eliminated these samples from further investigation. The fractionation and bioactivity guided isolation on the three lead samples are described below.

**Sponge Extract**

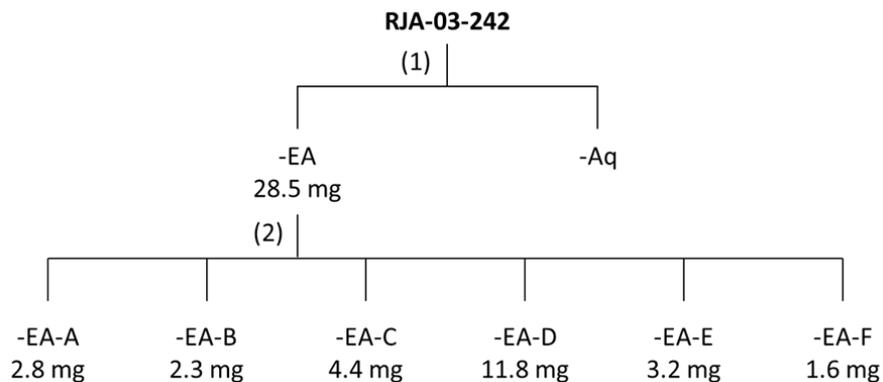
|  
└─ Liquid-liquid extraction ─┘  
├─ Organic Layer (Org)      └─ Aqueous Layer (Aq)

Sponge	Phase	0.1 mg/mL	0.01 mg/mL	0.001 mg/mL
97-199	Aq	99	26	9
	Org	83	18	4
03-242	Aq	100	42	14
	Org	100	54	8
00-287	Aq	38	9	2
	Org	100	57	17
11-312	Aq	10	1	-3
	Org	100	27	9
03-395	Aq	99	24	6
	Org	100	43	11

**Figure 3.10. Percent inhibition of TcTS activity for sponge extracts from aqueous (Aq) or organic (Org) phase from liquid-liquid phase extraction.**

### 3.5.1 Bioactivity-guided isolation of sample RJA-03-242

Sample RJA-03-242 was partitioned into six fractions (A-F) by size exclusion chromatography and tested for TcTS inhibition (Figure 3.11). Fraction C had increased inhibition relative to the ethyl acetate extract, yet only 4.4 mg of material remained at this point. <sup>1</sup>H NMR analysis of Fraction C indicated a mixture of compounds (data not shown), while a lack of available sponge material halted additional extraction efforts on this sample.

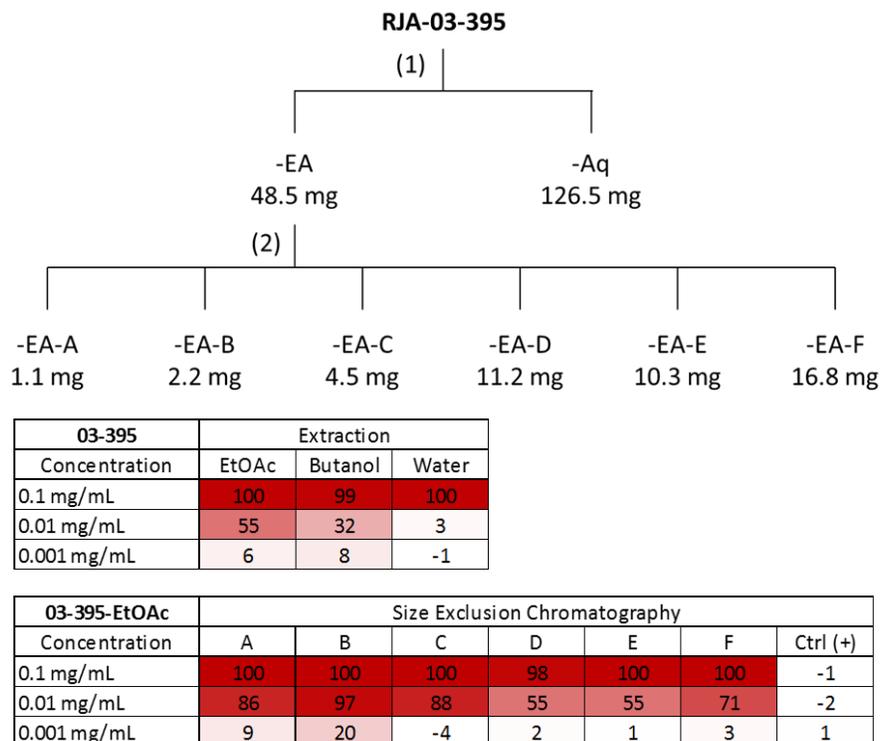


03-242-EtOAc	Size Exclusion Chromatography						
Concentration	A	B	C	D	E	F	Ctrl (+)
0.1 mg/mL	100	100	100	100	93	99	1
0.01mg/mL	26	21	73	58	6	38	1
0.001mg/mL	1	0	0	5	-7	1	-5

**Figure 3.11. Extraction scheme and bioassay summary for sample 03-242. (1) Liquid-liquid extraction (2) Size-exclusion chromatography.**

### 3.5.2 Bioactivity-guided isolation of sample PNG-03-395

The hit PNG-03-395 is the methanolic extract of a marine sponge that was collected by the Andersen lab off the coast of Papua New Guinea. This sample demonstrated strong inhibition in the selection assay versus TcTS and less than 10% activity versus other glycosidases tested. Initial attempts to isolate the bioactive chemical(s) responsible for the observed activity were pursued following the general protocol described above (Figure 3.12).



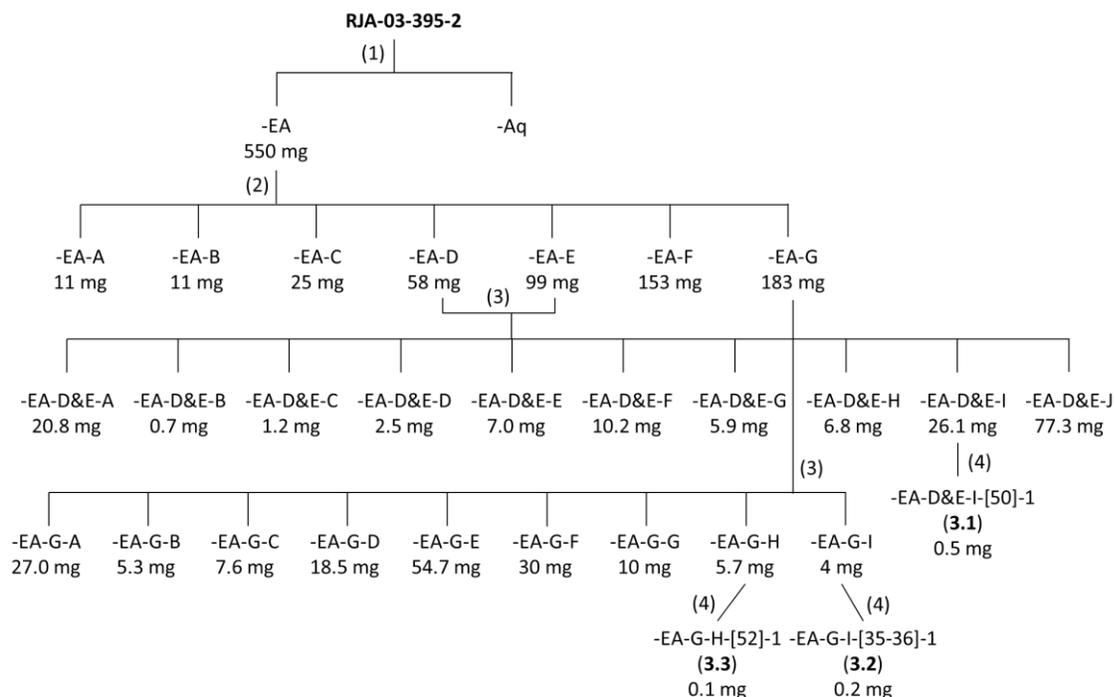
**Figure 3.12. Extraction scheme and bioassay summary for sample 03-395. (1) Liquid-liquid extraction (2) Size-exclusion chromatography.**

The enzyme inhibition assays – following liquid-liquid extraction between ethyl acetate/water and *n*-butanol/water – revealed that the ethyl acetate layer had higher inhibition (55%) at 0.01 mg/mL than both the butanol (32%) and aqueous (3%) components. The ethyl acetate extract was then fractionated by gel filtration over Sephadex LH-20 resin into six groups (A-F). Fractions A, B, C and F contained 1.6 mg, 2.2 mg, 5.8 mg and 15 mg respectively and each had enriched inhibition relative to the starting material. <sup>1</sup>H NMR spectra were collected for each sample to assess purity and structural similarities between the fractions. However, low quantities and low purity of the most active fraction (B) led us to determine that additional material would be necessary to proceed further. To this end the Andersen lab provided us with the remainder of the available sponge sample 03-395 (200 g).



**Figure 3.13. Marine sponge PNG-03-395 prior to extraction.**

The frozen organism was cut into small pieces (Figure 3.13) and exhaustively extracted with methanol (3 x 24 h). The combined extracts were filtered and evaporated to afford a dark solid. The sample was partitioned analogously to the initial extracts until three bioactive species were isolated (Figure 3.14). The methanolic extract was solubilized and partitioned between ethyl acetate/water. Size exclusion chromatography of 03-395-2-EA afforded seven fractions (A-G) which were assayed against TcTS. Fractions D, E and G exhibited the most potent inhibition (complete inhibition at 0.01 mg/mL) and fractions D and E were combined based on their similar retention times and inhibition profiles. Lead extracts (03-395-2-EA-D&E and 03-395-2-EA-G) were then eluted over reverse phase C18 silica gel column using a stepwise gradient of water/methanol from 100% water to 100% methanol and collected in nine fractions (A-I). The inhibition assay indicated that fractions G-I were the most potent in both cases. Each of the six samples was then partitioned by semi-preparative C18 column HPLC in an attempt to isolate the most potent bioactive components from this marine organism.



03-395-EtOAc	Size Exclusion Chromatography							
Concentration	A	B	C	D	E	F	G	Ctrl (+)
0.1 mg/mL	91	99	96	100	100	100	100	2
0.01 mg/mL	17	39	54	97	97	45	100	4
0.001 mg/mL	1	4	5	16	15	5	25	-3
0.1 µg/mL	-3	0	4	-3	2	3	18	0

03-395-EtOAc-D&E	Reverse Phase Chromatography									
Concentration	A	B	C	D	E	F	G	H	I	Ctrl (+)
0.1 mg/mL	100	95	99	100	100	100	100	100	100	-1
0.01 mg/mL	35	7	13	17	21	42	73	71	86	1
0.001 mg/mL	-6	0	-14	-13	-4	-4	-5	-2	-7	-2
0.1 µg/mL	-2	-14	0	-12	-11	-10	-12	-6	0	-3

03-395-EtOAc-G	Reverse Phase Chromatography									
Concentration	A	B	C	D	E	F	G	H	I	Ctrl (+)
0.1 mg/mL	70	72	100	100	100	100	100	100	100	-3
0.01 mg/mL	21	10	19	36	31	65	91	91	68	2
0.001 mg/mL	12	2	10	3	2	-7	9	1	3	-1
0.1 µg/mL	-4	-10	0	0	3	-10	13	-9	-7	3

**Figure 3.14. Extraction scheme and bioassay summary for sample 03-395-2. (1) Liquid-liquid extraction (2) Size-exclusion chromatography (3) Reverse-phase chromatography (4) HPLC.**

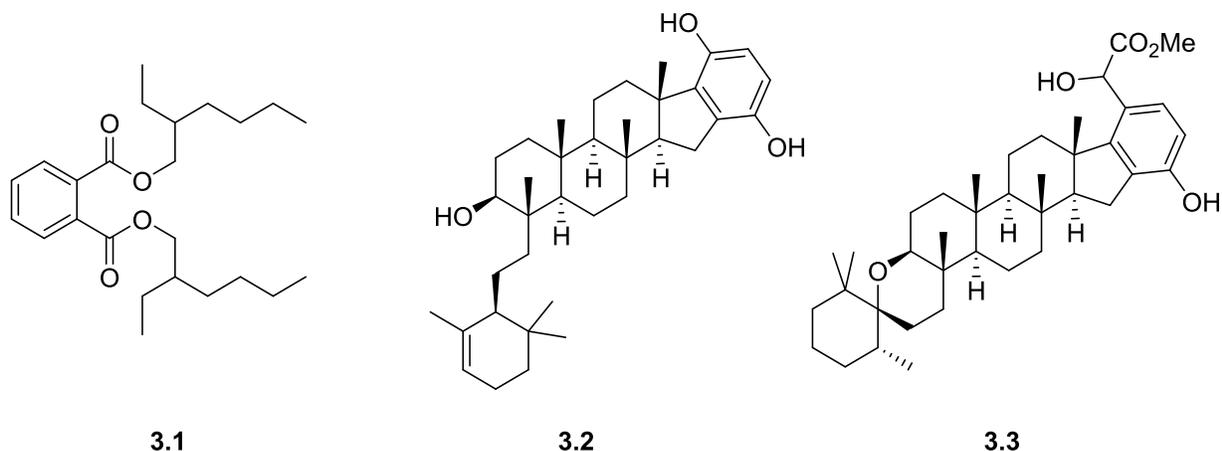
The appropriate HPLC elution gradients were first investigated on the analytical column (Agilent Poroshell 120) and then fine-tuned on the semi-preparative C18 column using sample 03-395-2-EA-D&E-I. Observable peaks at  $\lambda = 214$  nm were well spaced using a gradient from 70-95% acetonitrile over 40 minutes of elution. This method was employed on each of the six samples selected from reverse phase

partitioning. Fractions were collected in one minute increments, each was dried and dissolved in methanol/water and tested for TcTS inhibition. Different than the previous assays, fractions were diluted with the same volume of solvent (rather than adjusted to the same concentration), thus both the abundance and potency of the active compound plays a role in this assessment. We selected the 5-10 fractions that had the highest inhibition from each run and subsequently assessed them by  $^1\text{H}$  NMR for structural information. Most of the 47 samples had insufficient signal to noise ratio and/or purity for reliable structural determination, yet four lead candidates (03-395-2-EA-D&E-I[50]; 03-395-2-EA-G-G[52]; 03-395-2-EA-G-I[35-36]; 03-395-2-EA-G-G[36]) were chosen for additional analysis. Selection was based on abundance and purity of the sample – both of which are necessary to elucidate the structure and characterize enzyme inhibition.

### 3.5.3 Structural determination of sample(s) from PNG-03-395

The  $^1\text{H}$  NMR spectra of samples 03-395-2-EA-D&E-I[50] and 03-395-2-EA-G-G[36] were relatively simple (few distinct peaks) and displayed strong signal intensity. The relatively high concentration of material, the structural simplicity of the molecule and the modest inhibition observed for these fractions led us to believe they may not be of significant interest. Further, Dr. David Williams suggested that the  $^1\text{H}$  NMR fingerprint was likely indicative of the common class of plasticizers: phthalate esters. To confirm this, additional NMR experiments ( $^{13}\text{C}$ , COSY, [ $^1\text{H}$ ,  $^{13}\text{C}$ ] HSQC, HMBC, ROESY and [ $^1\text{H}$ ,  $^{15}\text{N}$ ] HSQC) were performed. Structural elucidation of 03-395-2-EA-D&E-I[50] and comparison to literature reports revealed the extracted compound to be bis(2-ethylhexyl) phthalate (DEHP; **3.1**). DEHP has been shown to be harmful to human health and the environment<sup>142</sup>; nevertheless it is the most common phthalate-based plasticizer in the world (3 billion kg produced per year<sup>143</sup>) and can be leached into liquids which contact these materials. Therefore **3.1** may have been present in our studies as a result of leaching from the materials used during the isolation process or extracted as a marine pollutant that was sequestered by the sponge. Similarities between the  $^1\text{H}$  NMR spectra of **3.1** and 03-

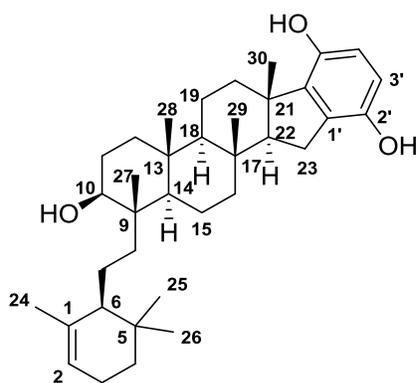
395-2-EA-G-G[36] led us to infer that it likely belongs to the phthalate family, yet the precise structure was not determined.



**Figure 3.15. Structures of natural products identified from sponge 03-395.**

In the Andersen lab, samples 03-395-2-EA-B-G-I[35-36] and 03-395-2-EA-B-G-G[52] were purified by a second round of HPLC – using an isocratic gradient – affording 0.2 mg and 0.1 mg of material respectively. A comprehensive set of NMR experiments (<sup>1</sup>H, <sup>13</sup>C, COSY, [<sup>1</sup>H, <sup>13</sup>C] HSQC, HMBC, ROESY and [<sup>1</sup>H, <sup>15</sup>N] HSQC) were collected for each sample and their structures proposed by Dr. Williams (Figure 3.15): **3.2** (from 03-395-2-EA-B-G-I[35-36]) and **3.3** (from 03-395-2-EA-B-G-G[52]). The structure of **3.2** was supported by mass spectrometry and by comparison to NMR data for related literature molecules. With the structure and NMR data set in hand I retroactively assigned each of the <sup>1</sup>H and <sup>13</sup>C signals (Table 3.6). The signal strength for **3.3** was insufficient to conclusively determine the structure, although a preliminary structure was suggested.

**Table 3.6. Structural assignment of 3.2.**



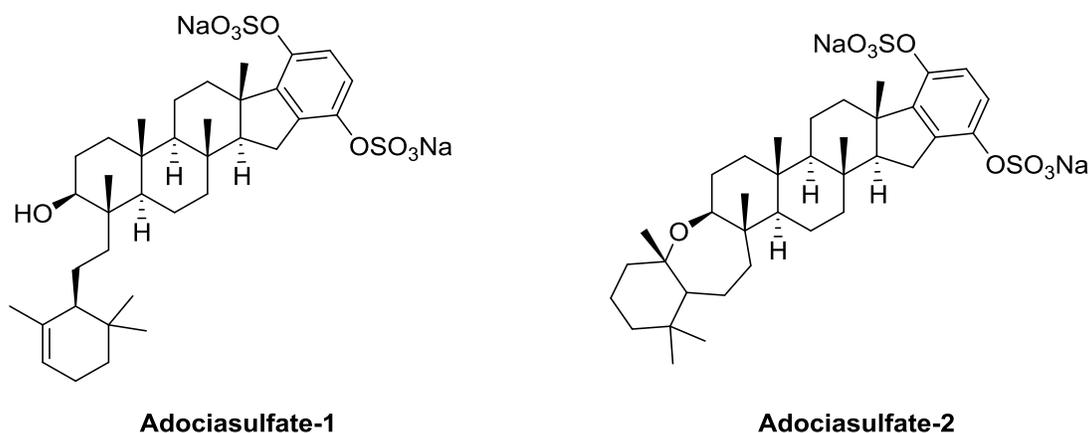
**3.2**

#	$\delta_C$		$\delta_H$ (mult., $J$ (Hz))	COSY	HMBC ( $^1H - ^{13}C$ )
1	137.2	C			
2	120.0	CH	5.37 (s)	3	
3	23.4	CH <sub>2</sub>	1.99 (m) 1.82 (q, 1.9)	2, 3, 4	1, 2, 6
4	31.8	CH <sub>2</sub>	1.59 (ddd, 13.2, 10.5, 6.9) 1.14 (dd, 13.0, 5.9)	3, 4	
5	32.9	C			
6	50.2	CH	1.41*		
7	24.6	CH <sub>2</sub>	1.50* 1.20 (m)	7, 8	
8	38.6	CH <sub>2</sub>	1.80* 1.37*	7, 8	
9	41.6	C			
10	72.5	CH	3.39 (m)	11	-
11	28.1	CH <sub>2</sub>	1.45* 0.69*	11, 12	
12	38.5	CH <sub>2</sub>	1.47* 0.66*	11, 12	
13	37.2	C			
14	50.0	CH	0.82 (s)	15	13, 15
15	18.7	CH <sub>2</sub>	1.45* 1.37*	14, 16	
16	41.4	CH <sub>2</sub>	1.41* 0.83*	15	
17	37.1	C			
18	61.4	CH	0.57 (dd, 12.3, 2.7)	19	13, 17

#	$\delta_C$		$\delta_H$ (mult., J (Hz))	COSY	HMBC ( $^1H$ - $^{13}C$ )
19	17.8	CH <sub>2</sub>	1.41* 1.35*	18, 20	-
20	37.9	CH <sub>2</sub>	2.37 (dt, 12.4, 3.3) 1.68*	19	-
21	47.0	C			
22	63.6	CH	1.31 (dd, 12.2, 6.4)	23	16, 17, 18, 20, 21, 23, 29, 30
23	24.9	CH <sub>2</sub>	2.24 (dd, 14.1, 6.5) 2.18 (m)	22	21, 22, 1', 2', 6'
24	23.5	CH <sub>3</sub>	1.81 (s)		1, 2, 6
25	27.6	CH <sub>3</sub>	0.98 (s)		4, 5, 6
26	28.1	CH <sub>3</sub>	1.09 (s)		4, 5, 6
27	17.8	CH <sub>3</sub>	0.81 (s)		8, 9, 10, 14
28	16.8	CH <sub>3</sub>	0.76 (s)		12, 13, 14, 18
29	17.3	CH <sub>3</sub>	0.85 (s)		16, 17, 18, 22
30	20.7	CH <sub>3</sub>	1.09 (s)		20, 21, 22, 6'
1'	129.4	C			
2'	146.8	C			
3'	113.5	CH	6.33 (d, 8.5)	4'	1', 2', 5'
4'	114.9	CH	6.16 (d, 8.4)	3'	2', 5', 6'
5'	145.7	C			
6'	140.3	C			
		OH	3.62 (s)		1', 2', 3'
		OH	3.66 (s)		4', 5', 6'
		OH	4.09 (s)		

HSQC and HMBC experiments supported the assignment of  $^{13}C$  signals; \* indicates overlapping signals.

Neither of these compounds have been previously reported in the literature, yet their structures are closely related to the family of adociasulfate natural products (Figure 3.16). This class of compounds was first extracted from the Palauan sponge *Haliclona* (i.e. *Adocia*) sp. and has demonstrated inhibition of both microtubule transport by motor protein kinesins<sup>144</sup> and proton pump activity in hen bone-derived membrane vesicles<sup>145</sup>.

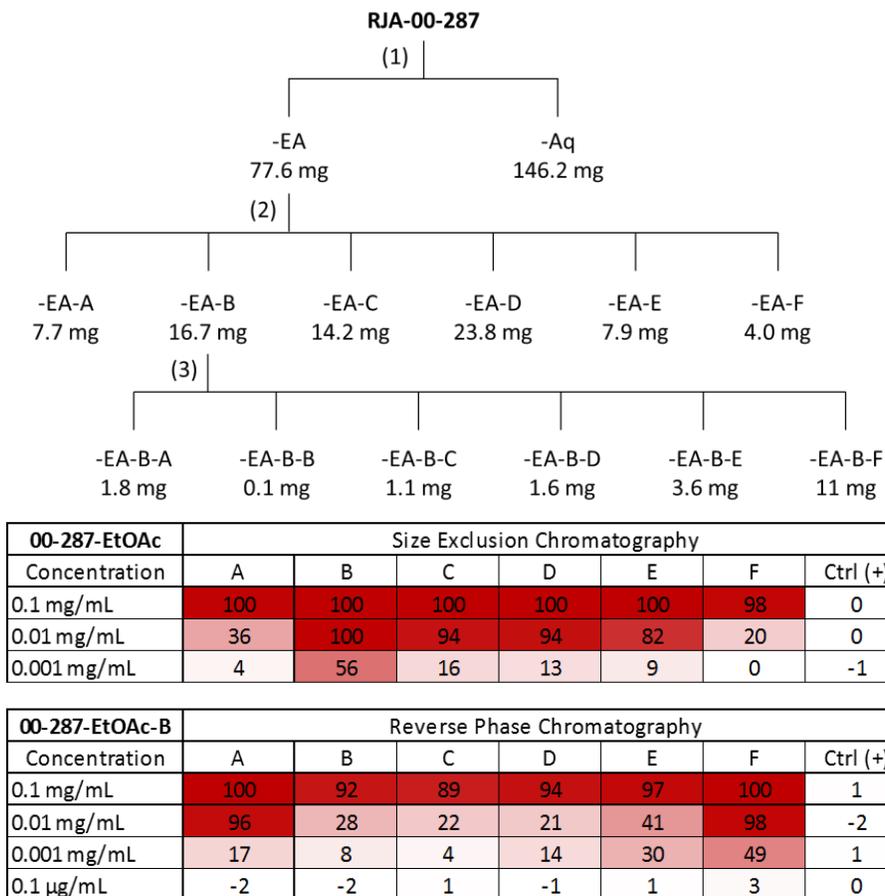


**Figure 3.16. Structures of select members of the adociasulfate family.**

### 3.5.4 Bioactivity-guided isolation of sample RJA-00-287

RJA-00-287 is the methanolic extract of a marine sponge that was collected by the Andersen lab in the Caribbean Sea near the shores of Dominica. This sample exhibited selective and potent inhibition of TcTS in the initial selection assays. Initial extraction procedures were investigated (Figure 3.17) in the same manner as for sample PNG-03-395.

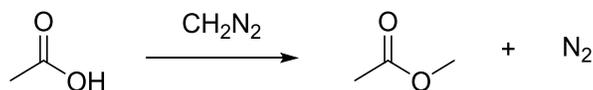
First, the methanolic extract was partitioned between ethyl acetate/water. The enzyme inhibition assay indicated that all inhibition activity was separated into the organic layer. The ethyl acetate extract was then fractionated by size exclusion chromatography into eight sections (A-G). Fraction B was significantly enriched in the bioactive species, demonstrating complete inhibition of TcTS at 0.01 mg/mL and 55% inhibition at 0.001 mg/mL. Sample 00-287-EA-B was further divided into six fractions (A-F) by reverse phase chromatography (C18 SepPak) via stepwise elution from water to methanol. The enzyme inhibition assay revealed no further enrichment at this stage since most of the sample eluted in the final step (100% methanol).



**Figure 3.17. Extraction scheme and bioassay summary for sample 00-287. (1) Liquid-liquid extraction (2) Size-exclusion chromatography (3) Reverse-phase chromatography.**

Next, the sample 00-287-B-F was fractionated by C18 semi-preparative HPLC and the absorbance was monitored at four different wavelengths ( $\lambda = 214 \text{ nm}, 240 \text{ nm}, 280 \text{ nm}, 300 \text{ nm}$ ), yet none of the traces showed any peaks throughout the elution. Each fraction was evaporated and dissolved in equal amounts of solvent and tested for TcTS inhibition. The bioactivity assay indicated that enzyme inhibition was spread across many consecutive fractions, leading us to speculate that the active component streaks on the C18 column. Those fractions were combined and later mixed with additional material (00-287-2-A&B-D) extracted during the scale-up procedure.



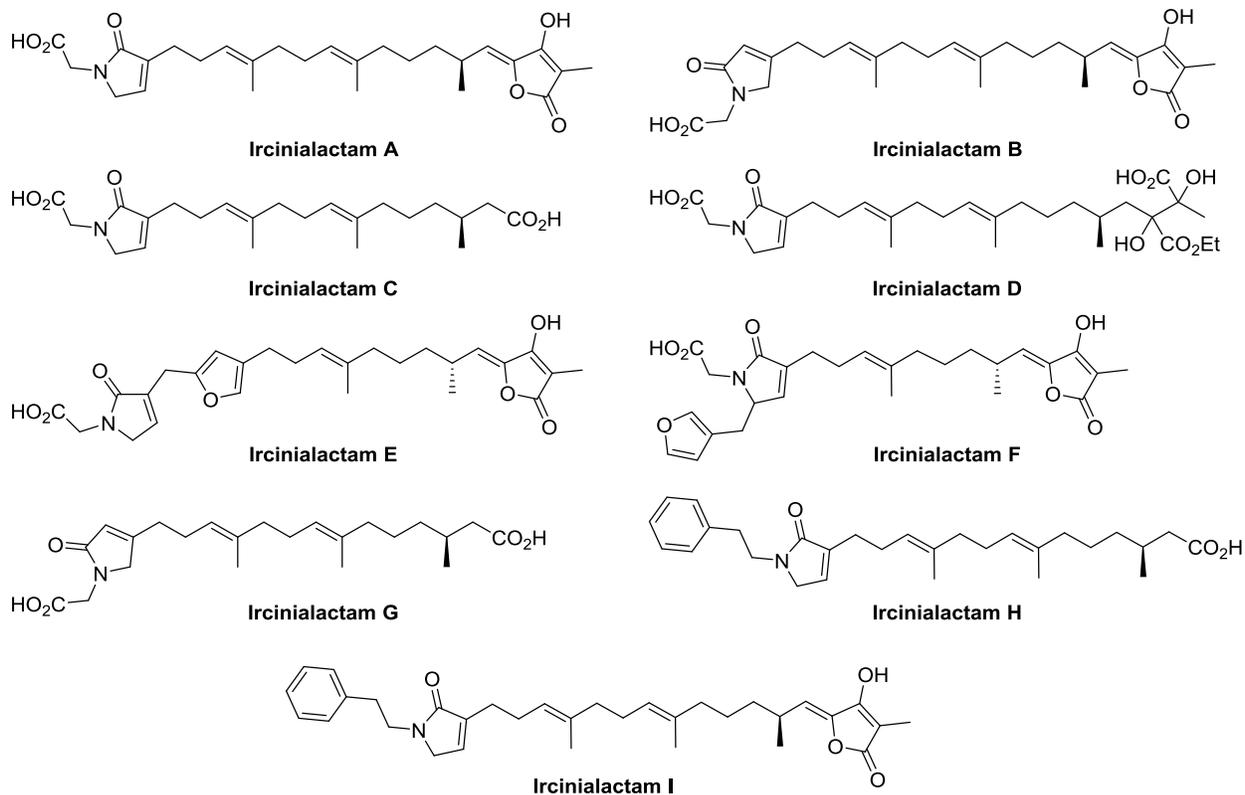


**Scheme 3.1. Methylation of carboxylic acid group by reaction with diazomethane.**

To this end, Dr. Williams treated sample 00-287-2-EA-A&B-D with fresh diazomethane and subsequently purified the products by HPLC. This time the absorbance trace indicated three distinct peaks being eluted, the major peak (00-287-2-EA-A&B-D-CH<sub>2</sub>N<sub>2</sub>-3) eluted after 55 minutes while two minor peaks (00-287-2-EA-A&B-D-CH<sub>2</sub>N<sub>2</sub>-1 and 00-287-2-EA-A&B-D-CH<sub>2</sub>N<sub>2</sub>-2) eluted earlier.

### 3.5.5 Structural determination of sample from RJA-00-287

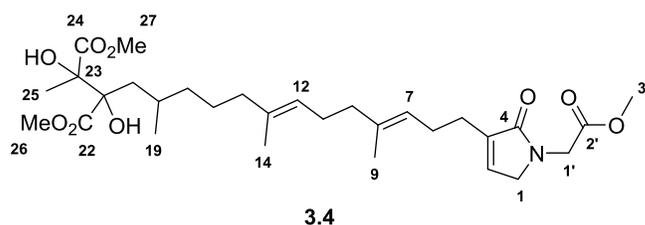
The largest isolated peak (00-287-2-EA-A&B-D-CH<sub>2</sub>N<sub>2</sub>-3) was dried to afford 0.4 mg of solid, which was subjected to NMR spectroscopic analysis (<sup>1</sup>H, <sup>13</sup>C, COSY, [<sup>1</sup>H, <sup>13</sup>C] HSQC, HMBC, ROESY and [<sup>1</sup>H, <sup>15</sup>N] HSQC) and mass spectrometry. Dr. Williams identified the material as **3.4**, a member of the ircinialactam family of natural products (Figure 3.19). Central to this family is the presence of a glycinolactam moiety, yet distally there is variation between metabolically related functional groups. Tetronic acid, the moiety opposite the lactam in ircinialactam A, is most common. The gamma lactone species has a pK<sub>a</sub> ~ 4 and is thus negatively charged in aqueous environments. If the tetronic acid is hydrolyzed, it can undergo oxidation processes to afford a variety of carboxylate bearing substituents (*e.g.* Ircinialactam C, D, G).



**Figure 3.19. Structures of the ircinialactam family (A-I).**

I retroactively assigned the  $^1\text{H}$  and  $^{13}\text{C}$  NMR resonances for **3.4** as shown in Table 3.7. Samples 00-287-2-EA-A&B-D- $\text{CH}_2\text{N}_2$ -1 and 00-287-2-EA-A&B-D- $\text{CH}_2\text{N}_2$ -2 were obtained in amounts  $<0.1$  mg thus hampering our ability to obtain quality NMR data. However, mass spectrometry results suggested that the species are partially esterified versions of **3.4**.

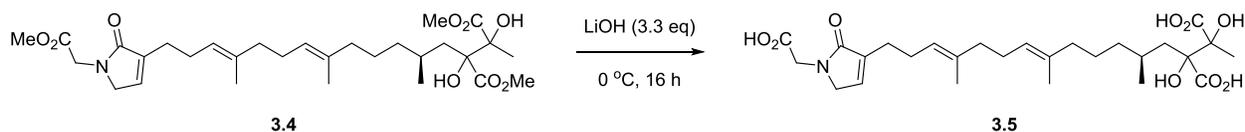
**Table 3.7. Structural assignment of 3.4.**



#	$\delta_c$		$\delta_H$ (mult., $J$ (Hz))	COSY	HMBC ( $^1H - ^{13}C$ )
<b>1</b>	51.4	CH <sub>2</sub>	3.98 (q, 2.0)	2	2,3
<b>2</b>	135.7	CH	5.71 (t, 1.6)	1	1,4
<b>3</b>	139.3	C			
<b>4</b>	172	C=O			
<b>5</b>	26.2	CH <sub>2</sub>	2.34 (t, 7.5)	6	2, 3, 4, 6
<b>6</b>	26	CH <sub>2</sub>	2.26 (q, 7.3)	5, 7	3, 5, 7, 8
<b>7</b>	123.5	CH	5.15 (m)	6	5, 9, 10
<b>8</b>	136.1	C			
<b>9</b>	16.2	CH <sub>3</sub>	1.60 (s)		7, 8, 10
<b>10</b>	39.9	CH <sub>2</sub>	1.98*	11	7, 8, 9, 11
<b>11</b>	26.8	CH <sub>2</sub>	2.07 (q, 7.6)	10, 12	10, 12, 13
<b>12</b>	124.3	CH	5.08 (m)	11	11, 14, 15
<b>13</b>	135.1	C			
<b>14</b>	16	CH <sub>3</sub>	1.57 (s)		12, 13, 15
<b>15</b>	39.9	CH <sub>2</sub>	1.90 (m)	16	12, 13, 14, 16
<b>16</b>		CH <sub>2</sub>	1.33*	15, 17	17
<b>17</b>	37.6	CH <sub>2</sub>	1.14 1.05	15, 17, 18	
<b>18</b>	29.7	CH	1.34*	17, 19, 20	
<b>19</b>	21.7	CH <sub>3</sub>	0.95 (d, 6.6)	18	17, 18, 20
<b>20</b>	39.1	CH <sub>2</sub>	1.96* 1.78 (dd, 14.2, 7.5)	18, 20	18, 19, 21, 22
<b>21</b>	82.4	C			
<b>22</b>	175.1	C=O			
<b>23</b>	80.1	C			
<b>24</b>	175.1	C=O			
<b>25</b>	20.5	CH <sub>3</sub>	1.51 (s)		21, 23, 24
<b>26</b>	53.1	OCH <sub>3</sub>	3.82 (s)		
<b>27</b>	53.1	OCH <sub>3</sub>	3.80 (s)		
<b>1'</b>	43.7	CH <sub>2</sub>	4.26 (s)		1, 2', 4
<b>2'</b>	169.9	C=O			
<b>3'</b>	52.2	OCH <sub>3</sub>	3.74 (s)		2'

HSQC and HMBC experiments supported the assignment of  $^{13}C$  signals; \* indicates overlapping signals.

Although the molecule **3.4** was identified by NMR it is not the original bioactive species that we were tracking throughout the isolation process. Removing the afore installed modification would typically afford the target compound; however, the presence of not one but three methyl ester moieties on **3.4** significantly complicated our ability to determine the identity of the natural product. It is plausible that the natural species is the triacid version of **3.4** (*i.e.* no methyl esters present), yet it is also possible that a combination of acid and ester moieties were initially present – as is the case for ircinialactam D. Conclusively determining which of the six partially esterified isomers may constitute the ‘natural’ species is not possible with the small amount of material in hand (~0.4 mg). Regardless, we decided that assessing the inhibitory activity of both **3.4** and the triacid (**3.5**) would be useful in determining the role of the anionic moieties on the inhibitory activity of the ircinialactam derivatives. Therefore we removed all three methyl ester moieties to this end (Scheme 3.2).



**Scheme 3.2. Saponification of ircinialactam 3.4.**

Aqueous lithium hydroxide (50 mM, 35  $\mu\text{L}$ , 3.3 equivalents) was added to a solution of **3.4** (0.3 mg) dissolved in THF/ $\text{H}_2\text{O}$  (7:3, 300  $\mu\text{L}$ ). The reaction was stirred overnight at 0  $^{\circ}\text{C}$  and after 16 hours the solvent was evaporated. The reaction progress was monitored by mass spectrometry, which at completion indicated the presence of **3.5** and the absence of both **3.4** and the partially saponified derivatives.

### 3.6 Kinetic analysis and structure-activity relationship

#### 3.6.1 Isolated natural products and structural derivatives

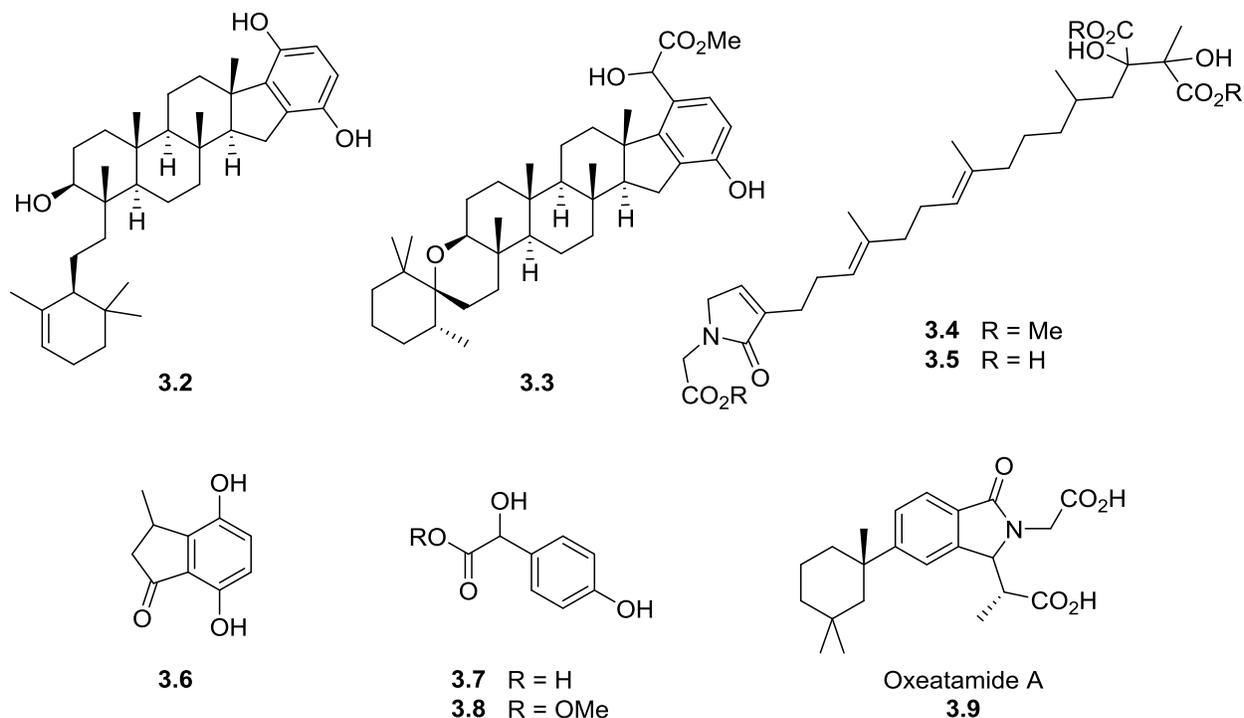


Figure 3.20. Structures of isolated natural products and analogues.

Five structures extracted from two marine sponges have been identified as potential TcTS inhibitors through bioactivity guided isolation. Additionally, molecules with structures similar to components of each natural product (**3.6-3.9**) were synthesized or obtained from commercial or laboratory sources (Figure 3.20). Each was assessed for TcTS inhibition (Table 3.8, see Appendix C for plots).

**Table 3.8. Inhibition values of isolated natural products and their analogues.**

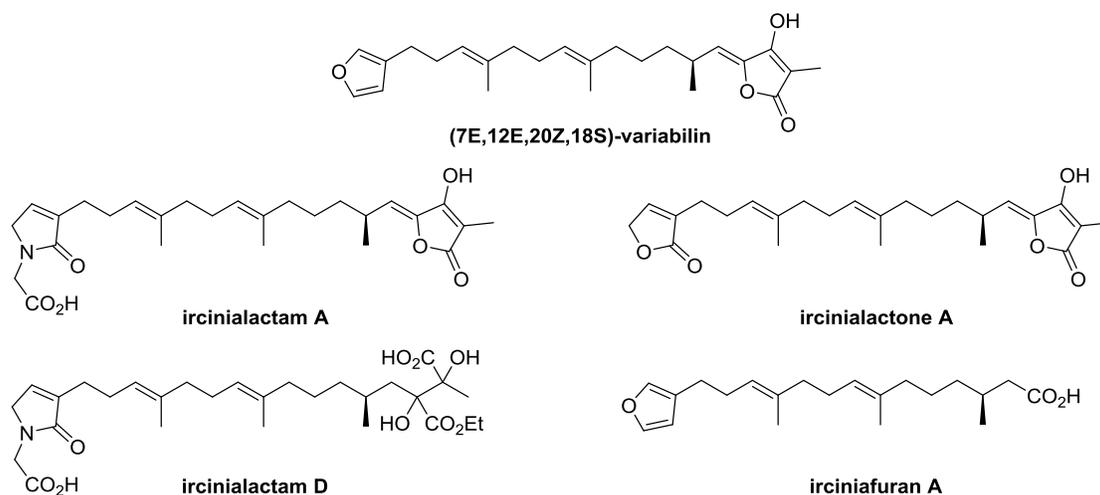
Compound	IC <sub>50</sub> (μM)	Compound	IC <sub>50</sub> (μM)
<b>3.1</b>	283 ± 46	<b>3.6</b>	213 ± 23
<b>3.2</b>	23.7 ± 0.7	<b>3.7</b>	542 ± 49
<b>3.3</b>	51.3 ± 2.1	<b>3.8</b>	507 ± 42
<b>3.4</b>	17.2 ± 1.5	<b>3.9</b>	516 ± 14
<b>3.5</b>	12.4 ± 0.7		

Of all the natural products and natural product derivatives tested, isolated species **3.2**, **3.4** and **3.5** have the most potent inhibition versus TcTS with IC<sub>50</sub> values <25 μM. Adociasulfates **3.2** and **3.3** are 10-fold better than their respective structural derivatives **3.6**, **3.7** and **3.8**. The plasticizer DEHP (**3.1**) was much less potent than the other isolated natural products (IC<sub>50</sub> = 283 ± 46 μM). Its isolation can likely be attributed to its relative abundance in the extract sample rather than its inhibition potency versus TcTS.

We hypothesized that a carboxylate moiety of the ircinialactam derivative may contribute significantly to the observed TcTS inhibition in the isolation assays, since the natural substrate bears an essential C1 carboxylate. However, comparison of the IC<sub>50</sub> values for **3.4** (IC<sub>50</sub> = 17.2 ± 1.5 μM) and **3.5** (IC<sub>50</sub> = 12.4 ± 0.7 μM) indicates that the triacid derivative is only a slightly better inhibitor (1.4-fold) than the triester. This result implies the charge state of the product may not be essential for inhibition.

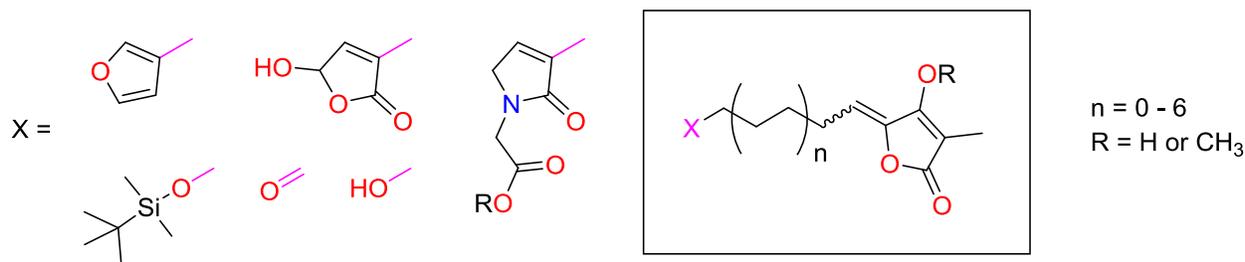
We also hypothesized that the glycinal lactam moiety has structural features (*e.g.* planar, charged, cyclic) which loosely resemble those of common transition-state mimics that can be potent glycosidase inhibitors. Therefore another glycinal lactam-bearing natural product, Oxeatamide A (**3.9**) (provided by the Andersen lab), was tested as a TcTS inhibitor. This compound had an IC<sub>50</sub> >500 μM, at least 50-fold worse than the ircinialactams, suggesting that the glycinal lactam moiety is not the principal source of the TcTS inhibition exhibited by **3.4** and **3.5**. No structural similarities are apparent between the icinialactam and adociasulfate species isolated in this process. Thus two unique structural scaffolds – with low micromolar IC<sub>50</sub> values – have been discovered for TcTS inhibition.

### 3.6.2 Screening ircinialactam library



**Figure 3.21. Structures of select members of the sesterterpene tetronic acid family.**

In 2010, Balansa and colleagues reported the isolation of the natural product variabilin and its rare glycinyl lactam-containing metabolites, ircinialactams, from three sponges of the Family Irciniidae<sup>146</sup>. These were discovered while screening a library of >2500 extracts of marine invertebrates for modulators of the glycine receptor (GlyR) chloride channels. Several of the isolated natural products – belonging to the common biosynthetic family called sesterterpene tetronic acids – exhibited promising biological activity for their target protein. The Capon group has continued to isolate new members of this family from Australian Irciniidae sponges further highlighting the structural diversity of the metabolites (Figure 3.21) and helping to elucidate the potential biosynthetic relationship between the  $\beta$ -furan, lactone and glycinyl lactam moieties<sup>147</sup>. Our natural product screening results showed that the glycinyl lactam containing species **3.5** – which closely resemble the structure of ircinialactam D – exhibited the strongest inhibition of TcTS. We therefore initiated a collaboration with the Capon lab at the University of Queensland and they courteously provided a library of synthetic variabilin-like derivatives (termed the ircinialactam library) designed to mimic the metabolic diversity of the sesterterpene tetronic acid family.



**Figure 3.22. General structure of the ircinialactam library compounds.**

The proprietary library of 117 compounds was based on the structure of variabilin and was designed to assess structural components such as linker length, double bond regiochemistry and stereochemistry, furan derivatization, oxidation, and charge state of the conserved tetronic acid core (Figure 3.22). Linker length was varied between 3 and 15 carbons, linkers contained 0-2 double bonds in either the E or Z conformation, while some linkers were functionalized to contain methyl or hydroxyl moieties. Distal to the conserved tetronic acid, functional groups were varied between glycinolactam, aldehyde, furan, alcohol or tetronic acid moieties. Finally, the molecules were either anionic or neutral in charge depending on the methylation state of the tetronic acid and/or glycinolactam components.

The library was provided as two 96-well plates containing approximately 0.1 mg of dried synthetic sample per well. Sample plates were suspended in DMSO (100  $\mu\text{L}$ /well), replicated using a 10-fold dilution and assayed for TcTS inhibition at 0.1 mg/mL and 0.01 mg/mL using the same conditions employed when screening the UBC Marine Natural Product Extract library. The inhibition levels at 0.1 mg/mL were high across the entire library (58% average) while the average inhibition at 0.01 mg/mL was 25%. Assessing the level of TcTS inhibition at both concentrations allowed us to bin the library components into four sections for analysis (most inhibitory, inhibitory, mildly inhibitory, non-inhibitory). The results revealed some trends regarding the structural features best suited for TcTS inhibition and allowed us to select the most potent derivatives for secondary analysis.

	1	2	3	4	5	6	7	8	9	10	11	12
P1A												
P1B												
P1C												
P1D												
P1E												
P1F												
P1G												
P1H												
P2A												
P2B												

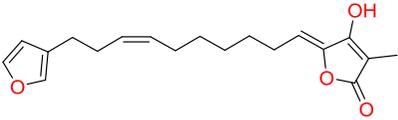
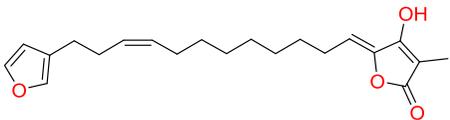
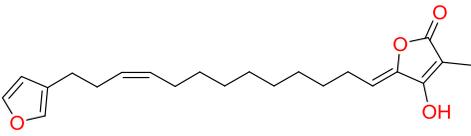
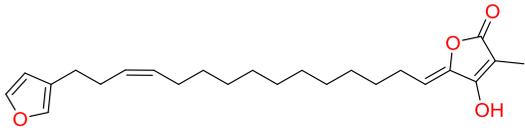
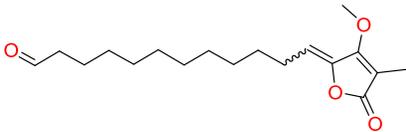
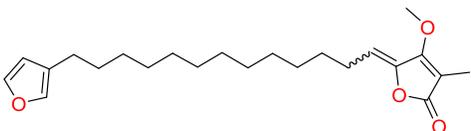
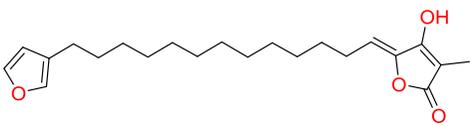
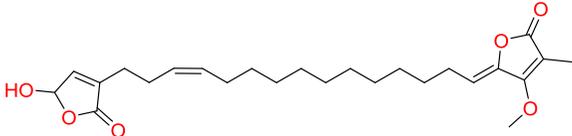
Most Inhibitory   
 Inhibitory   
 Mildly Inhibitory   
 Non-Inhibitory

Figure 3.23. Ircinialactam library screening results for TcTS inhibition.

### 3.6.2.1 Most inhibitory

The *most inhibitory* category was selected based on >95% inhibition at 0.1 mg/mL and >85% inhibition at 0.01 mg/mL. Eight members of the library belonged to this section, each of which had long linkers of 10-15 carbons (average =  $12.8 \pm 1.8$  carbons) and most bore furan as the distal functionality (6/8 members). Neither charge state (5 OH/3 OMe) nor linker saturation (5 with/3 without) were conserved across the hits, implying that these features are not essential for enzyme inhibition. Interestingly, none of the glycinal lactam-containing members of this library were selected in this category. A dilution series of each hit well was measured and an  $IC_{50}$  of each compound was determined (Table 3.9). The  $IC_{50}$  value was calculated under the assumption that each well of the ircinialactam library contains 0.1 mg of sample.

**Table 3.9. Inhibition values for *most inhibitory* hits from ircinialactam library versus TcTS.**

Compound	Structure	IC <sub>50</sub> * (μg/mL)
P1D8 (3.10)		6.0 ± 0.8
P1D10 (3.11)		4.3 ± 0.5
P1D11 (3.12)		8.5 ± 0.9
P1D12 (3.13)		4.7 ± 0.4
P1E9 (3.14)		8.6 ± 1.0
P1G2 (3.15)		5.5 ± 0.6
P1G3 (3.16)		6.0 ± 0.7
P2A7 (3.17)		5.7 ± 0.7

\*Concentration was determined under the assumption that each well contains 0.1 mg of material.

### 3.6.2.2 Inhibitory

The *inhibitory* category was selected as the second tier for compounds that showed >90% inhibition at 0.1 mg/mL and >50% inhibition at 0.01 mg/mL. Twelve members of the library fulfilled these requirements, many of which closely resembled the molecules selected in the *most inhibitory* category. However they have small structural differences such as shorter linker length (3-11 carbons; average =  $9.4 \pm 2.8$  carbons) and a more diverse assortment of distal functionality: 4 furans, 3 aldehydes, 3 tetronic acids and 2 glycinal lactams (Figure 3.24).

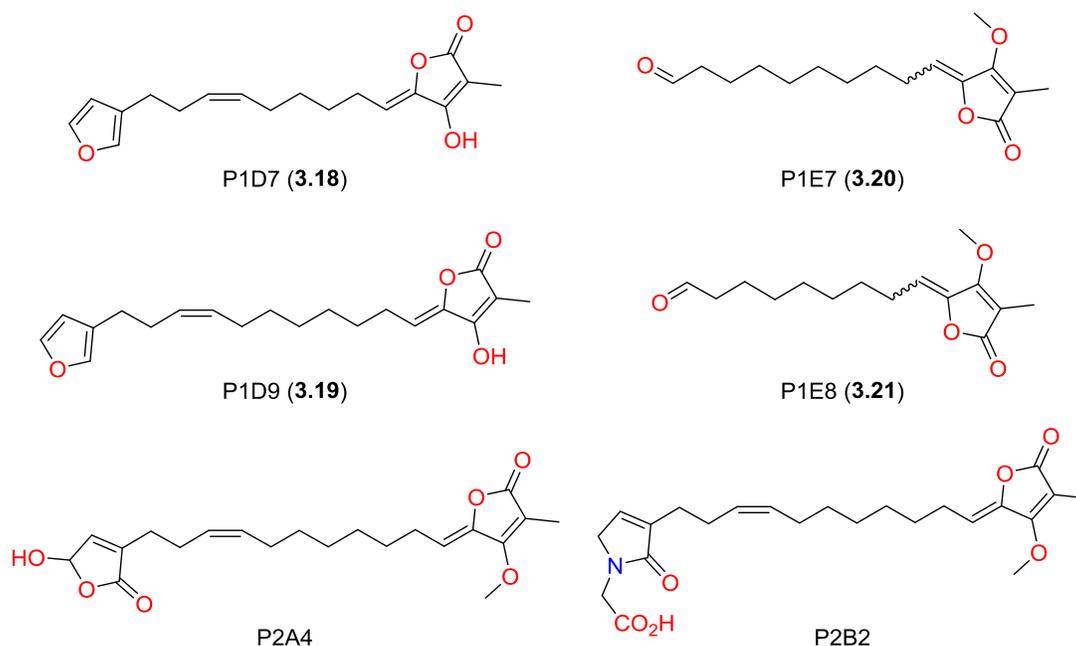


Figure 3.24. Structures of select *inhibitory* hits from ircinialactam library.

### 3.6.2.3 Mildly inhibitory and non-inhibitory

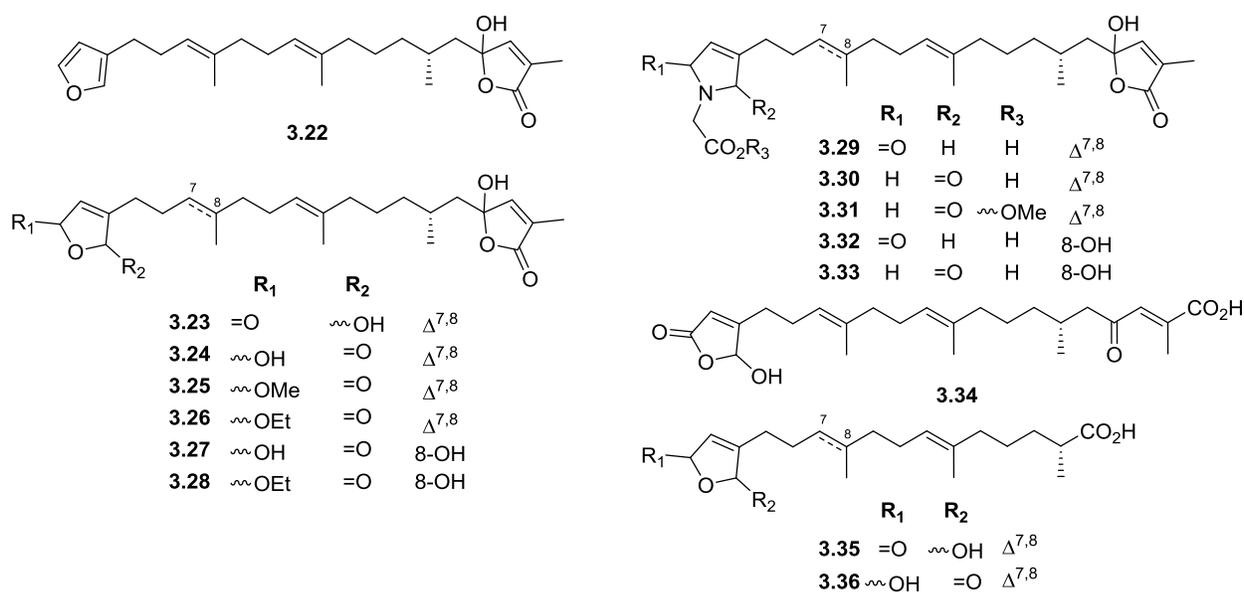
The *mildly inhibitory* category is the largest and accounts for >50% of the library members. Selection threshold was set at >20% activity at 0.1 mg/mL. This criteria was chosen to distinguish the compounds with low inhibition from the completely inactive members of the library. The chemical diversity of the substituents is high and qualitative observations suggest that shorter linkers (average  $7.9 \pm$

3.2 carbons) and distal *tert*-butyldiphenylsilyl (TBDPS) protected alcohols (15/15 members) are more prevalent than in the previous two categories.

Lastly, the *non-inhibitory* category comprised all species that demonstrated <20% inhibition of TcTS at the highest tested concentration (0.1 mg/mL). These 29 compounds provide important information regarding which scaffolds in the library of highly related structures did not exhibit any inhibition in the enzyme assay. Interestingly, many of the individual components (*e.g.* furan, glycinolactam, tetronic acid) that did not contain two chemical groups bridged by a linker were members of this category. Further, short linkers (average  $5.8 \pm 2.1$  carbons) and tetronic acids bearing distal hydroxyl functionalization (10/15 members) were highly represented. This category also included many glycinolactam-linked derivatives with short linkers (14 compounds), supporting the notion that the glycinolactam component of **3.4** and **3.5** is not solely responsible for TcTS inhibition.

### 3.6.3 Screening cacolide library

In addition to the synthetic ircinialactam library, the Capon lab also provided a set of fifteen sesterpene butenolides (*i.e.* cacolides) which they had isolated from the Australian marine sponge *Cacospongia* sp. in 2018<sup>148</sup> (Figure 3.25). This library was provided as dried samples (0.1 mg each) that were reconstituted in DMSO (100  $\mu$ L), diluted 10- and 100-fold, and analogously tested for TcTS inhibition. While all samples exhibited >50% inhibition of TcTS at 0.1 mg/mL two compounds – the furan containing cacolide A (**3.22**) and the glycinolactam bearing cacolide I (**3.30**) – abolished TcTS activity at 0.1 mg/mL and had >75% and >40% inhibition at 0.01 and 0.001 mg/mL respectively. The 42% and 51% inhibition of TcTS by **3.22** and **3.30** respectively at 0.001 mg/mL is the most of the inhibition we have measured under these conditions. Crude IC<sub>50</sub> estimates from the three data points affords values between 1-4  $\mu$ M, however more robust kinetic analysis is required. While these derivatives rank among the most potent compounds tested for TcTS inhibition, a lack of available material prevented further investigations at this time.



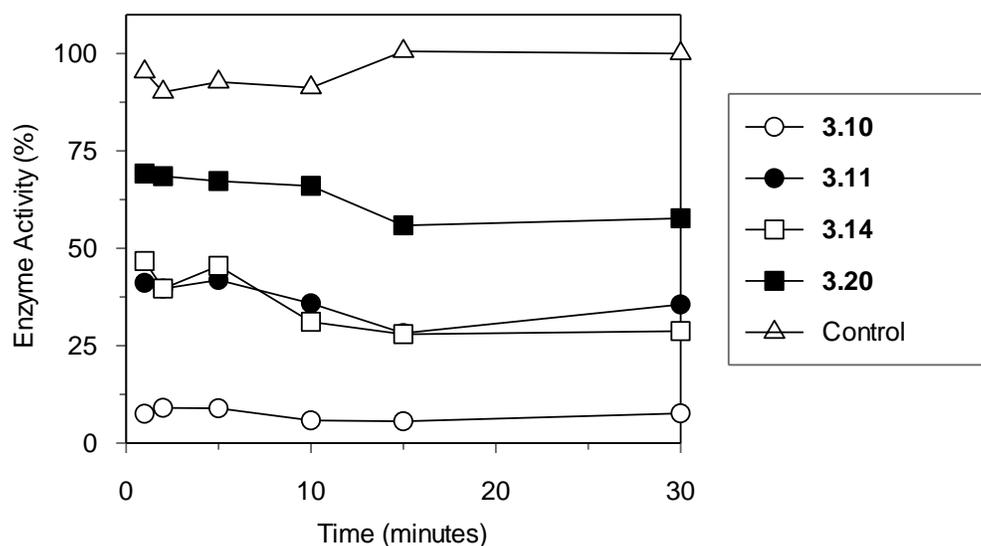
Conc. (mg/mL)	3.22	3.23	3.24	3.25	3.26	3.27	3.28	3.29
0.1	100	77	77	94	91	50	85	72
0.01	76	24	25	26	21	6	14	23
0.001	42	22	-5	-4	-6	-12	-9	31
Conc. (mg/mL)	3.30	3.31	3.32	3.33	3.34	3.35	3.36	Ctrl (+)
0.1	100	99	61	63	96	74	97	-1
0.01	90	48	3	2	21	8	22	4
0.001	51	6	-5	-6	-5	-2	3	-4

Figure 3.25. Structure (top) and percent inhibition (bottom) of TcTS by cacolides 3.22-3.36.

### 3.6.4 Inhibition studies: ircinialactam library hits

In order to further characterize the inhibition of the best ircinialactam derivatives, the Capon lab provided 1-3 mg of additional material for eleven compounds (**3.10-3.16** and **3.18-3.21**). The chemical structure of the synthetic ircinialactam derivatives contain several electrophilic groups (*e.g.* aldehyde, alkene, Michael acceptor) which may be prone to chemical reaction with enzyme residues. Therefore, representative species bearing furan (**3.10**, **3.11**) or aldehyde (**3.14**, **3.20**) moieties were tested for time-dependent inhibition versus TcTS (Figure 3.26). The synthetic ircinialactam inhibitors (25 μM) were incubated with TcTS at 30 °C for 30 minutes. Aliquots were sampled at six time points and were diluted

10-fold into a solution of buffered TFMU-SA. The substrate cleavage was monitored at  $\lambda = 385$  nm using the Cary4000 UV/Vis spectrophotometer. Control samples were incubated with an equivalent volume of DMSO and the percent inhibition for each sample was calculated relative to the average control value. Each of the four inhibitors exhibited relatively constant levels of enzyme inhibition over time, which indicated that a covalent time-dependent mechanism of action is not employed.



**Figure 3.26. Time-dependent inhibition assay with select ircinialactam library hits.**

Next, the  $IC_{50}$  values of each of the 11 hits were measured using two-fold serial dilutions of inhibitor from  $100 \mu\text{M}$  to  $10 \text{ nM}$ . Ten of the samples had values between than  $1.8 \mu\text{M}$  and  $4.9 \mu\text{M}$  while **3.20** was the least potent ( $IC_{50} = 7.5 \mu\text{M}$ ). The mode of inhibition and the inhibition constant ( $K_i$ ) were then determined for each ircinialactam analogue (Table 3.10). The inhibitors were incubated with TcTS at six concentrations between  $0.1 \mu\text{M}$  and  $10 \mu\text{M}$ , and the residual enzyme activity was determined at three concentrations of TFMU-SA substrate ( $0.25$ ,  $0.5$ ,  $1 \text{ mM}$ ). Results were separately fit to three equations describing either competitive, non-competitive or uncompetitive inhibition using data analysis software GraFit 7.0 (Erithacus). Qualitative (visual inspection) and quantitative means (reduced  $\chi^2$ ) of assessment

for each fit revealed that the compounds do not employ a competitive mode of action. Fitting the kinetic data to non-competitive or uncompetitive inhibition equations afforded reduced  $\chi^2$  values that were significantly lower than when using the competitive fit in all cases. The quality of fit and resultant  $K_i$  values were similar between non-competitive and uncompetitive models but non-competitive fits had lower  $\chi^2$  values. The  $K_i$  values calculated using a non-competitive fit (Equation 3.2) are listed in Table 3.10, many of which are comparable to the previously measured  $IC_{50}$  values. Our findings are similar to those reported by Arioka and colleagues<sup>78</sup>, where synthetic derivatives of natural products isolated from screening efforts afforded low micromolar inhibitors of TcTS that exhibit a non-competitive mode of action.

**Equation 3.2. Non-competitive inhibition.**

$$v = \frac{V_{\max} [S] \frac{1}{1 + [I]/K_i}}{K_m + [S]}$$

**Table 3.10. Inhibition values for ircinialactam library hits 3.10-3.21.**

<b>Compound</b>	<b>IC<sub>50</sub> (μM)</b>	<b>K<sub>i</sub> (μM)</b>	<b>Compound</b>	<b>IC<sub>50</sub> (μM)</b>	<b>K<sub>i</sub> (μM)</b>
<b>3.10</b> (1D8)	1.8 ± 0.1	0.8 ± 0.2	<b>3.16</b> (1G3)	3.1 ± 0.1	6.0 ± 1.1
<b>3.11</b> (1D10)	4.5 ± 0.2	4.7 ± 1.0	<b>3.17</b> (2A7)	N.D.	N.D.
<b>3.12</b> (1D11)	2.0 ± 0.1	2.8 ± 0.9	<b>3.18</b> (1D7)	2.0 ± 0.1	2.2 ± 0.3
<b>3.13</b> (1D12)	1.8 ± 0.1	3.8 ± 0.7	<b>3.19</b> (1D9)	2.1 ± 0.1	2.1 ± 0.4
<b>3.14</b> (1E9)	3.1 ± 0.1	4.6 ± 0.7	<b>3.20</b> (1E7)	7.5 ± 0.1	7.1 ± 1.0
<b>3.15</b> (1G2)	2.1 ± 0.1	4.5 ± 1.9	<b>3.21</b> (1E8)	4.9 ± 0.1	6.5 ± 0.9

**N.D. = Not determined.**

### 3.7 Conclusions

The primary goal of this chapter was to find new natural product molecules exhibiting potent TcTS inhibition. To this end we designed a functional screen for enzyme inhibition and collaborated with the Andersen lab to assay the UBC Marine Natural Product Extract library. Eight hits were selected for secondary analysis where they were assessed for inhibition potency and selectivity versus  $\beta$ -galactosidase and cytosolic human neuraminidase. Bioactivity guided isolation was performed with Dr. Williams on the lead sponges, affording four isolated molecules from two organisms. Structural elucidation revealed that the two most active compounds, **3.2** ( $IC_{50} = 23.7 \mu\text{M}$ ) and **3.5** ( $IC_{50} = 12.4 \mu\text{M}$ ), are novel in structure and belong to the adociasulfate and ircinialactam families respectively.

By collaborating with the Capon lab we screened a library of 117 synthetic sesterterpene tetronic acid mimics. The library contained highly related derivatives that systematically differed in linker length and structure, charge state and functionality distal to the tetronic acid core. Analysis of the best and worst inhibitors revealed that anionic or neutral tetronic acids, bearing a furan or aldehyde functionality – separated by a saturated or unsaturated linker of >10 carbons – performed better than related structures in the library. While 42 glycinal lactam-containing structures were included in the screen, they typically demonstrated worse inhibition than their furan analogues. The best inhibitors were shown to exhibit non-covalent non-competitive inhibition of the enzyme target with inhibition constants <5  $\mu\text{M}$ . The best inhibitors from a series of fifteen cacolide natural products isolated by the Capon lab demonstrated similar potency to the most inhibitory ircinialactam derivatives from the synthetic library and may warrant further investigation. However, additional materials would be required to further investigate the inhibition parameters and biological activity of cacolide species.

Kinetic results indicate that sesterterpene butenolides, sesterterpene tetronic acids and their metabolites exhibit strong inhibition of TcTS activity. The most potent sponge compounds (**3.4**, **3.5**), ircinialactams analogues (**3.10**, **3.13**) and cacolides (**3.22**, **3.30**) all bear two functional groups (*e.g.* furan, lactam, dicarboxylate, tetronic acid, butenolide) separated by unsaturated linkers of 10-15 carbons. If

additional research efforts were dedicated to synthesizing a combinatorial library of these features, sub-micromolar inhibitors of TcTS could likely be identified. However, significant investment of time and resources into improving inhibitor potency may necessarily be the top priority. Instead, efforts to better understand the site and mode of binding/inhibition employed by these non-competitive species via kinetic and structural analysis (*e.g.* enzyme crystallography) may be more important at this stage. Additionally, using the low micromolar inhibitors we have already identified – such as the synthetic ircinialactam analogues – to assess TcTS as a viable therapeutic target in cell-based systems should also be prioritized. Initial research efforts towards these goals have been explored in Chapter 5.

## Chapter 4: High-Throughput Screening and Synthesis of Non-Natural Cyclic Peptide Inhibitors for TcTS

### 4.1 Introduction

As described in previous chapters, rational design of carbohydrate-based inhibitors for TcTS has been fruitless to date and screening approaches have uncovered a few new inhibitor scaffolds. Screening studies have evaluated commercial small molecule or natural product libraries, no such investigations on peptide-based inhibitors have been reported versus TcTS. Given the advances in peptide screening technology and the success of the Withers and Suga labs in finding nanomolar inhibitors of glycosidase targets, we too decided to interrogate macrocyclic peptides as potential TcTS inhibitors.

#### 4.1.1 Peptide screening

Chemical diversity is an essential variable to consider when selecting a library for screening. Combinatorial synthetic libraries have enabled access to vast arrays of unique compounds, which will surely continue to increase with technological advances. Such libraries scale by the number of discrete building blocks and the number of distinct connections between building blocks. Peptide-based assemblies can also be viewed as combinatorial libraries; where the length of the peptide chain is  $X$ , and the number of available building blocks for a given node is  $N$ , the potential combinations can thus be represented as  $N^X$ . Therefore, in the case of peptides where any of the 20 proteinogenic amino acids are available, the potential diversity can be represented as  $20^X$ . If non-standard amino acids are also available, diversity will further increase. To harness the power of these massive libraries one must be able to both screen and deconvolute the highly diverse peptide libraries. Technologies such as phage display and mRNA display have demonstrated this capacity for library sizes on the order of  $10^9$  and  $10^{14}$  respectively<sup>149,150</sup>.

Phage display is a powerful technology that has allowed researchers to access diverse libraries of polypeptides or antibodies. In 2018 the inventors Smith and Winter were awarded the Nobel prize in chemistry for its profound impact on the scientific community. The basis of this technology is the ability to insert a gene of interest – which encodes a polypeptide – into the coat protein gene of bacteriophage (the virus that infects bacteria). The encoded sequence will then be ‘displayed’ on the surface of the bacteriophage; in this way, a physical linkage between the phenotype (*i.e.* displayed peptide) and the genotype (*i.e.* DNA which encodes it) is established. Immobilized targets can then be screened *in vitro* by these large libraries of phage displayed peptides and the resultant hits can be deconvoluted by sequencing of the DNA tags (Figure 4.1).

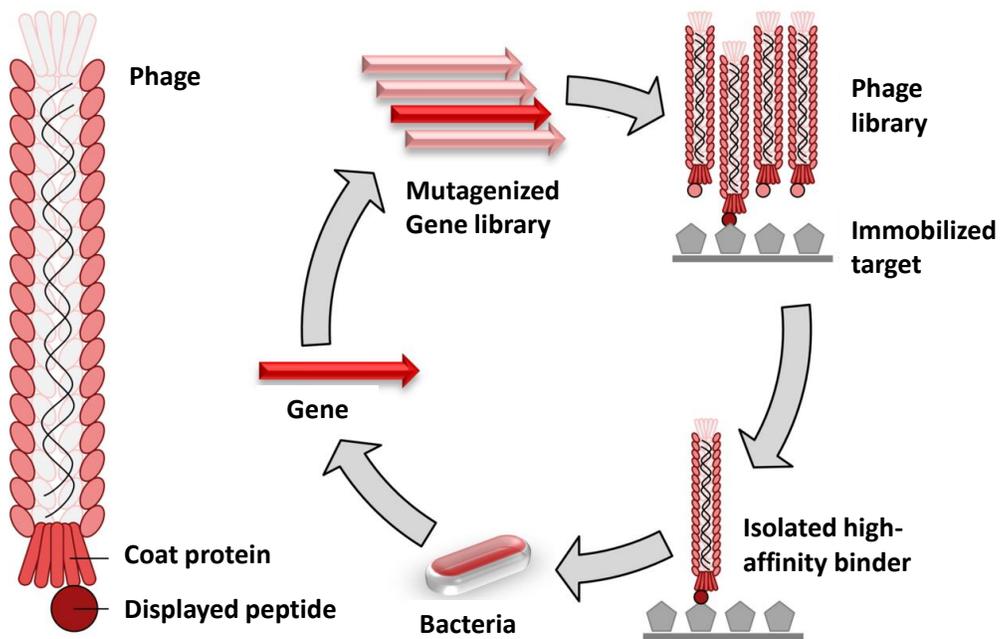
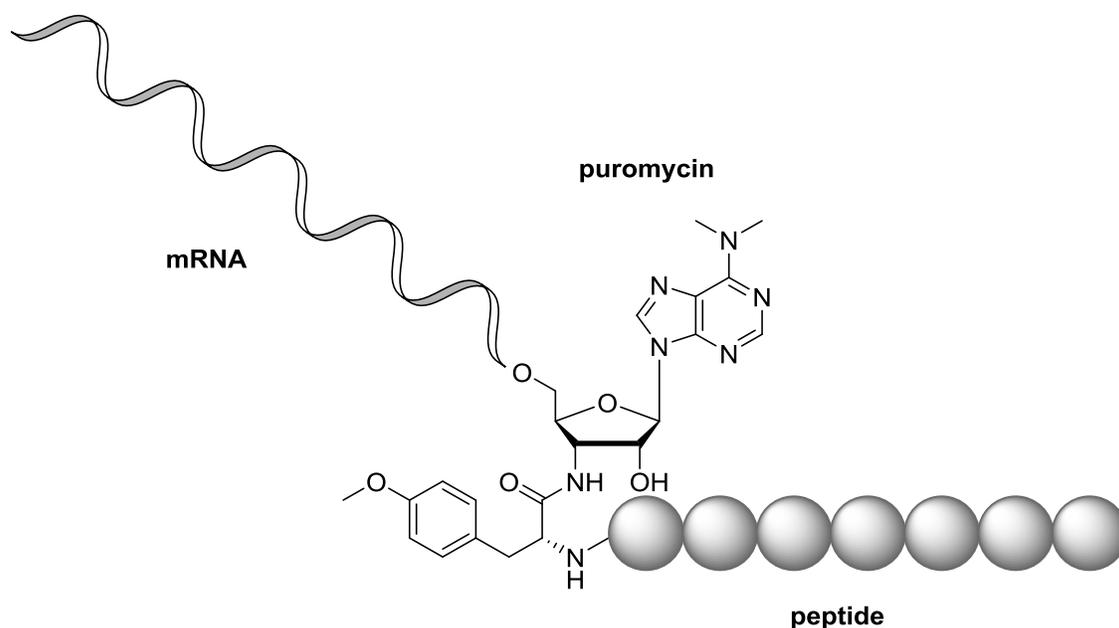


Figure 4.1. General summary of phage display screening. Source: Shafee<sup>151</sup>

Phage display was first described in 1985 using *E. coli* specific M13 bacteriophage<sup>152</sup>. Many systems have since been improved for niche use, but this seminal system is still predominantly employed today. Factors such as vector display, coat engineering and library construction have further been optimized to yield a robust and improved methodology<sup>153</sup>. Although generally effective, phage display

suffers from two major issues. First, unique library construction is limited to  $\sim 10^9$  independent members due to the bacterial transfection efficiency<sup>154</sup>. The second limitation is that the bacteriophage ( $\sim 1.6 \times 10^7$  Da) is significantly larger than the peptides ( $< 10,000$  Da) displayed on the surface. This feature can lead to significant interference between the target and the polypeptide or to undesirable complexation between the phage surface and the peptide, rendering the sequence inaccessible to target interaction<sup>150</sup>. The former issue was resolved following the invention of ribosome display<sup>155,156</sup> and mRNA display technology. Ribosome display is a completely *in vitro* method which allows library diversity to be improved up to  $10^{12}$ . The peptide-mRNA link is established by stalling the ribosome complex during translation; however, this method requires experimental conditions which are tailored such that the ribosome-mRNA-peptide complex remains intact. Further, this display method is still plagued by the size discrepancy issues (Ribosome:  $\sim 2,000,000$  Da) described for phage-based systems. Roberts and Szostak reported an elegant solution whereby a covalent linkage was established between the nascent peptide and the encoding mRNA strand through a puromycin linker that is installed at the 3' end of the mRNA fragment (Figure 4.2). Puromycin shares structural similarities between nucleic and amino acids thus making it ideally equipped for the task. Together with Wilson and Keefe they reported the *in vitro* selection method (termed mRNA display) for preparation and screening of peptide libraries ( $10^{13}$ ) with greater complexity than can be generated with phage display<sup>157</sup>.



**Figure 4.2. Role and structure of puromycin in connecting peptide to encoding mRNA strand.**

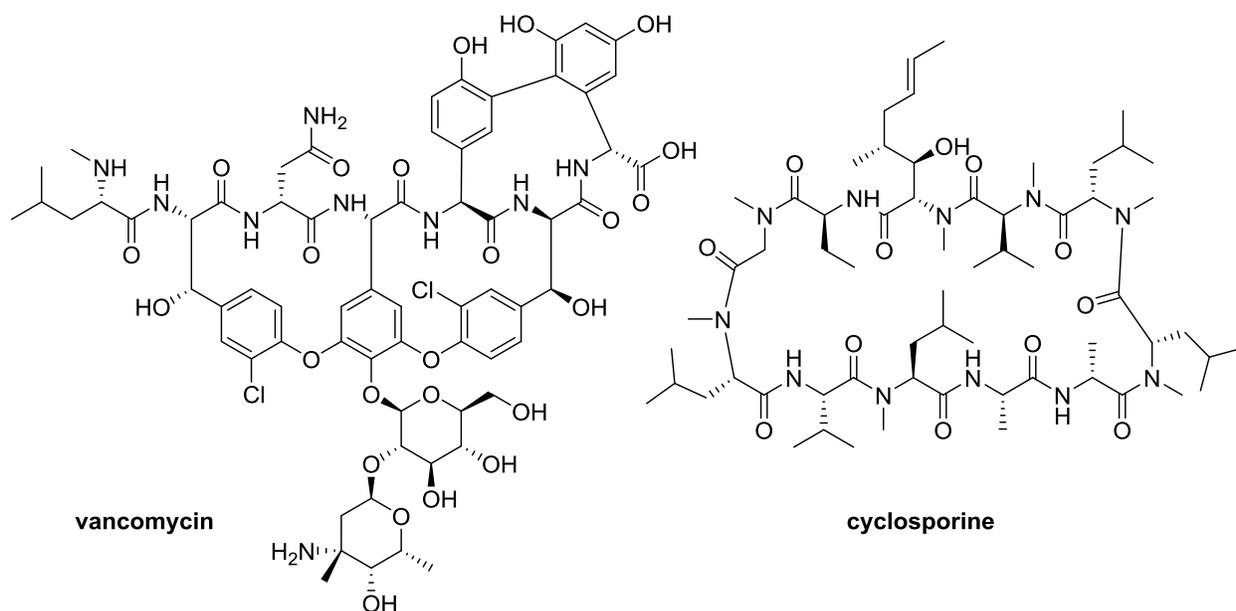
Immediately, the fruits of this technology were realized, as screening experiments using mRNA display elucidated peptides that exhibited low dissociation constants (*i.e.* high binding affinity, <10 nM) against the protein target streptavidin. These results were particularly illustrative because an analogous screening attempt using phage display methods – targeting the exact same protein – failed to uncover equally strong binding peptide sequences<sup>157</sup>. The superior results obtained by mRNA display screening can potentially be attributed to several factors including a larger initial library size ( $10^4$ -fold larger), longer peptide lengths screened (88 vs 38 amino acids) and/or the avoidance of interactions between the support (phage) and peptide binding. These enormous peptide-based libraries have been screened in search of peptides with high binding affinity for therapeutically relevant protein targets.

#### **4.1.2 Peptide-based therapeutics**

Peptides play an essential role in human physiology through their multitude of functions as hormones, neurotransmitters, growth factors, ion channel ligands or anti-infectives. In most cases these

peptides can perform their role by exerting selective and efficient binding as signaling molecules for specific cellular receptors. The specific nature of their interactions and the intrinsic pharmacological profile of peptides species (*i.e.* safe, predictable metabolism, well tolerated) make them attractive starting points for therapeutic development<sup>158</sup>. Peptides are also easily accessible using standard synthetic protocols, which makes them even more attractive as drug candidates. The size and potential complexity of peptide-based therapeutics set them apart from common small-molecule drugs since they can target large binding areas beyond the active site<sup>159</sup>. However, peptides also have some potential drawbacks as therapeutics, such as their instability to chemical hydrolysis and oxidation, short biological half-lives (from rapid enzymatic degradation), low membrane permeability and their tendency to aggregate in aqueous environments. Researchers and organisms alike have made some headway in addressing these pitfalls.

Peptide-based natural products attracted considerable interest after the discovery of the antibiotic vancomycin in the 1950s and the immunosuppressive agent cyclosporine (Figure 4.3) in the 1970s. The latter was quickly granted approval for clinical use in 1983. The lipophilic undecamer contains a cyclic peptide core, frequent *N*-methylation, a D- amino acid, hydrophobic side chains and an unnatural functional group.

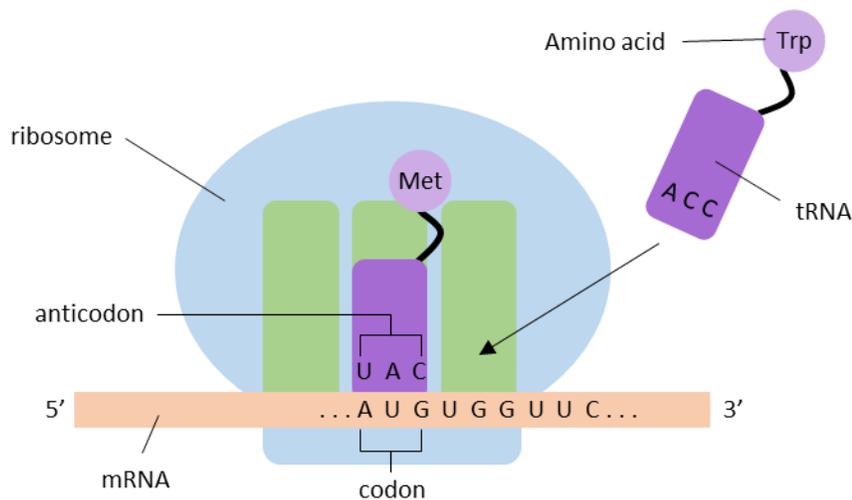


**Figure 4.3. Structure of cyclic peptide-based natural products.**

These features mirror the primary strategies that have been employed by chemists to combat enzymatic degradation and increase circulation lifetimes in other peptides. The introduction of non-proteinogenic amino acids and the addition of secondary or tertiary structural elements via macrocyclization can simultaneously enhance binding affinity while disfavoring peptidase recognition and associated cleavage activity<sup>160</sup>. Therefore the rational incorporation of these elements into therapeutically relevant peptides can enhance its physicochemical properties. Practically this requires the systematic evaluation of the peptide sequence to determine a structure-activity relationship (SAR) with the target receptor of interest. From this information, sites for chemical modification (*e.g.* *N*-methylation), substitution or cyclization can be determined. New peptide derivatives bearing these elements can be synthesized and assessed in an iterative process until the optimal properties have been developed<sup>161</sup>. On a case-by-case basis, SAR-informed rational design has afforded bioactive peptides with improved physicochemical properties in an effective but highly cumbersome process. Researchers have therefore sought out methods to include non-proteinogenic amino acids in higher throughput processes.

### 4.1.3 Ribosomal peptide synthesis

Methods to employ non-proteinogenic amino acids into peptide expression systems have therefore been devised to synthesize therapeutically interesting sequences. However to understand the way in which this has been done we must first explore how proteinogenic peptides are synthesized in living systems. The central dogma of molecular biology states that the information flows from genes to proteins in a central two-step pathway: DNA → RNA → protein. The first step is transcription, whereby genetic DNA is copied to make a messenger RNA (mRNA) sequence. In the second step, mRNA is translated using cellular machinery to express the encoded peptide sequence. This is a highly complex process, yet below I provide a non-comprehensive summary outlining only the fundamental components of ribosomal protein synthesis (Figure 4.4).



**Figure 4.4. General summary of peptide translation.**

First let's examine the participants. mRNA is the single strand of ribonucleic acids that carries the genetic information transcribed from DNA. The strand is made of four building blocks (Uridine, Cytidine, Guanine, Adenine) which can be read in triplets (*i.e.* groups of three) called codons. Each triplet combination encodes a specific amino acid, although there are more combinations than there are unique amino acids, which is called the genetic-code redundancy. The ribosome is the 'factory' where the mRNA

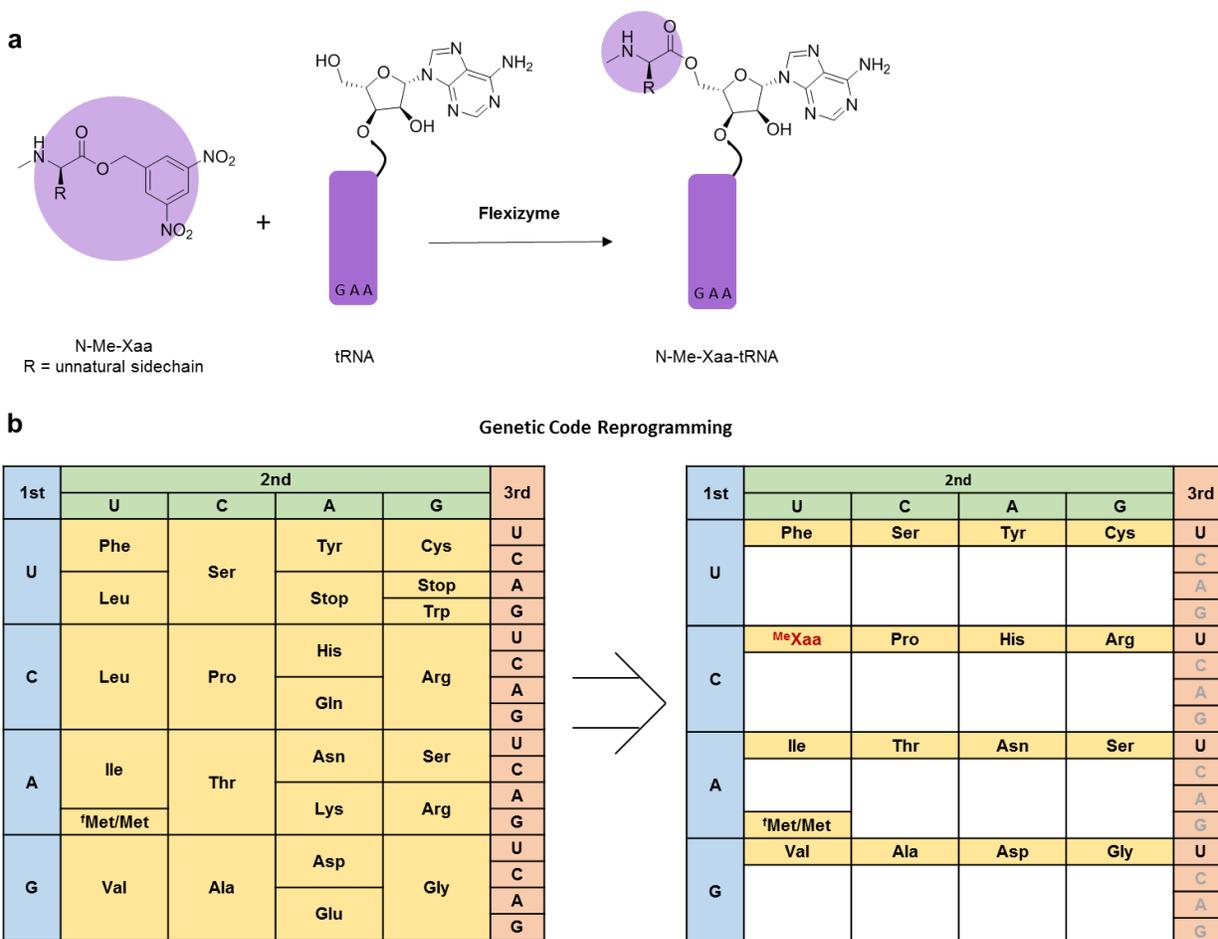
is decoded to afford the encoded amino acid chain. Transfer RNA (tRNA) is the molecule that bears both an anticodon and an amino acid, which makes it the physical link between mRNA and amino acid sequence. The anticodon is a trinucleotide sequence that is complementary to the corresponding codon. Finally, amino acid-tRNA-synthases (aatRS) are the enzymes responsible for linking the specific amino acids to the tRNA with the encoding anticodon. The entire process can now be described as the iterative translation of codons on the ribosomally-mounted mRNA strand by amino acid-charged tRNAs to sequentially transfer the encoded amino acid residues to a polypeptide chain being formed. Researchers have developed ways in which to modify (or ‘reprogram’) the genetic code to install non-proteinogenic amino acids into oligopeptides and proteins.

#### **4.1.4 Genetic code reprogramming**

The translation process is extremely efficient at producing polypeptides/proteins using the 20 amino acids for which it encodes. Researchers have even managed to lift the necessary machinery from cellular systems to simulate ribosomal expression *in vitro*. However, peptides are inherently limited in their chemical scope and their utility for therapeutic purposes. Therefore researchers have attempted to reprogram the genetic code to incorporate non-proteinogenic amino acids into translated peptide sequences. Engineering efforts (*e.g.* directed evolution) have successfully developed aatRSs that can charge non-proteinogenic amino acids onto tRNA<sup>162</sup>. However the chemical structure of the modified residue must closely resemble that of the proteinogenic amino acid, limiting the overall scope of this method. Alternatively, non-proteinogenic amino acids have been incorporated into protein sequences using mixed chemical/biochemical synthetic protocols<sup>163</sup> however the very arduous nature of these methods make them less appealing than ribosomal expression systems.

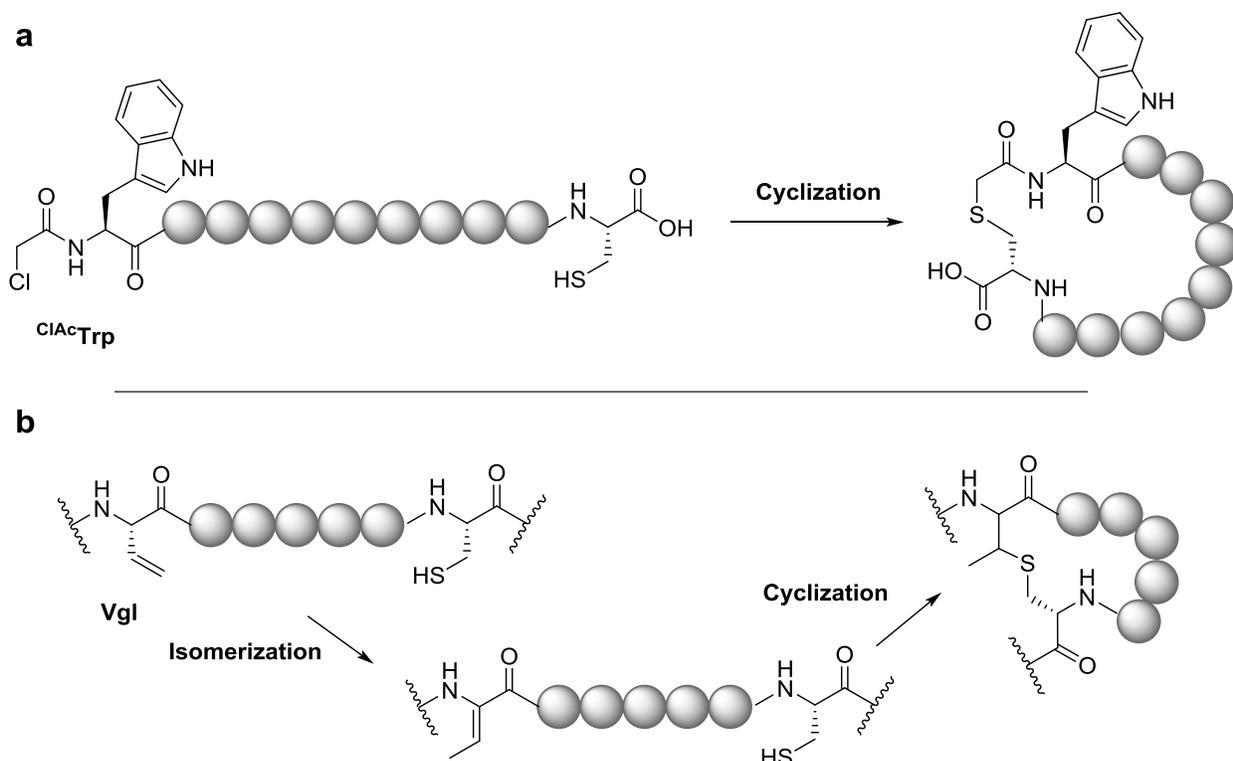
The Suga lab has developed ribozymes (catalytic RNA molecules) that are capable of charging tRNAs with a range of non-proteinogenic amino acids<sup>164</sup> (Figure 4.5a). The ribozymes (also termed Flexizymes) recognize the 3'-end of the tRNA acceptor and a benzylic leaving group installed on the

amino acid donor. Therefore since the Flexizyme recognizes the leaving group, rather than the amino acid structure, a wide variety of aminoacyl benzyl esters can be employed to construct new non-canonical aminoacyl-tRNA complexes. The ability to control the identity of the amino acid that will be charged onto a specific tRNA allows researchers to now modify the genetic code (Figure 4.5b). The combination of Flexizyme-catalyzed tRNA acylation and cell-free translation systems enables the facile synthesis of polypeptides bearing non-proteinogenic amino acids<sup>165</sup> and has since been termed the flexible *in vitro* translation (FIT) system. This *in vitro* translation method requires two groups of starting materials: (1) reconstituted *E. coli* translation components and (2) non-standard amino acids charged onto tRNA complexes. By appropriately omitting components from (1) corresponding to a specific codon box then the chosen additives from (2) will replace it<sup>166</sup>. For example in a NNU system (where N is any ribonucleotide), 15 proteinogenic amino acids can typically be accessed. However if leucine is omitted from the translation system and a pre-charged tRNA – bearing both an non-proteinogenic amino acid and the anticodon for CUU – was added, the non-proteinogenic residue would replace any encoded Leu in the resultant polypeptide.



**Figure 4.5. (a) Flexizyme catalyzed attachment of non-proteinogenic amino acid to tRNA-GAA (b) genetic code (left) and reprogrammed NNU genetic code (right) where CUU has been exchanged for N-Me-Xaa.**

Flexizyme technology has been extensively used in the Suga lab to install a multitude of functional groups to modulate a variety of peptide features including peptide stability (via macrocyclization or addition of non-proteinogenic amino acids). The FIT system has been used to reprogram the initiation step of translation by substituting the highly conserved methionine initiation codon for amino acids bearing *N*-terminal  $\alpha$ -chloro-acetamide moieties. By employing this substitution Goto and colleagues demonstrated that the linear peptides could be readily cyclized between an encoded cysteine residue and the *N*-terminal functional group (<sup>ClAc</sup>Trp) to afford a new physiologically stable bond (Figure 4.6a)<sup>167</sup>.



**Figure 4.6. Post-translational modification of non-proteinogenic amino acids (a)  $^{ClAc}Trp$  or (b) Vinylglycine (Vgl) following site specific incorporation with the FIT system.**

Similarly, D-amino acids can have been implemented (in place of the typical L-amino acids) at the initiation site. This is particularly appealing for therapeutic peptide development as those bearing amino acids with non-proteinogenic stereochemistry confer resistance to proteolytic degradation<sup>168</sup>. Goto has further demonstrated the site-specific incorporation of an non-proteinogenic amino acid vinylglycine (Vgl) using the FIT system<sup>169</sup>. This enabled the synthesis of a naturally occurring methylanthionine-containing peptide through post-translational isomerization and subsequent Michael addition to yield the cyclic peptide of interest (Figure 4.6b). The FIT system has since been employed in combination with their mRNA display screening platforms to enable access to libraries of natural product-like peptide macrocycles.

#### 4.1.5 RaPID Screening

The Suga lab has enhanced mRNA display methods by integrating their FIT system into the screening method in order to incorporate non-proteinogenic amino acids and macrocyclization. This combination has been termed the random non-standard peptide integrated discovery (RaPID) system (Figure 4.7)<sup>170,171</sup>.

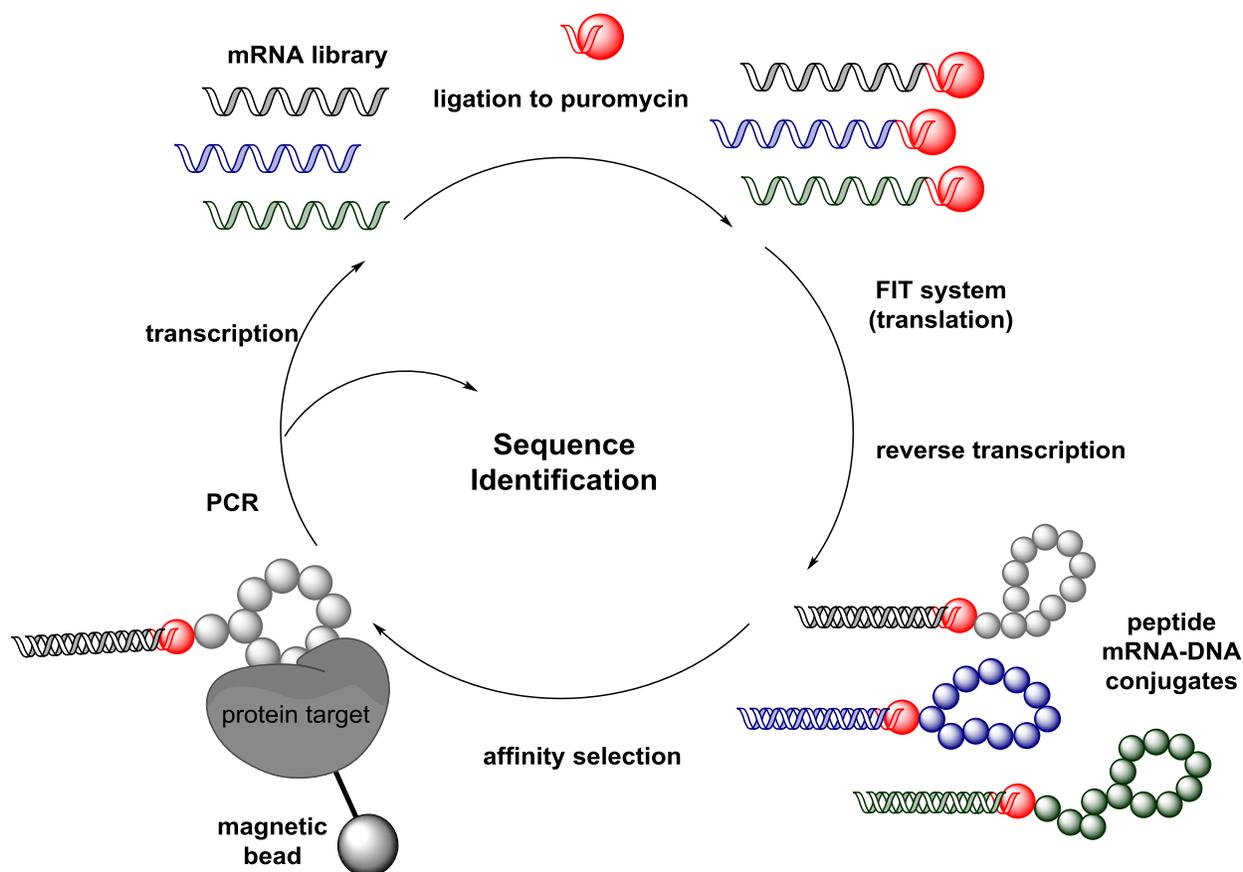


Figure 4.7. General summary of RaPID screening.

Their lab has subsequently reported the discovery of new cyclic non-natural peptide inhibitors for a wide variety of protein targets. Potent inhibitors of ubiquitin ligase<sup>170</sup>, isoform selective inhibitors of Akt2<sup>172</sup>, allosteric inhibitors of Semaphorin 4D Receptor Plexin B1<sup>173</sup>, selective inhibitors of histone demethylases<sup>174</sup> and the inhibition of cellular entry for the Hepatitis B virus<sup>175</sup> highlight some of the key discoveries enabled by the RaPID screening system.

In 2017, Jongkees reported the discovery of potent and selective inhibitors of a human glycosidase by means of the RaPID system<sup>176</sup>. A cyclic nonapeptide inhibitor of human pancreatic amylase (HPA) was determined to exert potent ( $K_i = 7$  nM) and selective inhibition of HPA over other glycosidases tested. Structural studies demonstrated that the target peptide formed a compact tertiary structure whereby two phenolic groups are involved in the tight-binding inhibition of HPA. These results are highly consistent with observations from non-protein based inhibitors of the same target. Thus, the discovery of new peptide-based structures exhibiting selective and potent inhibition towards a glycosidase target makes RaPID screening a very appealing approach to pursue in search of new scaffolds for TcTS inhibition.

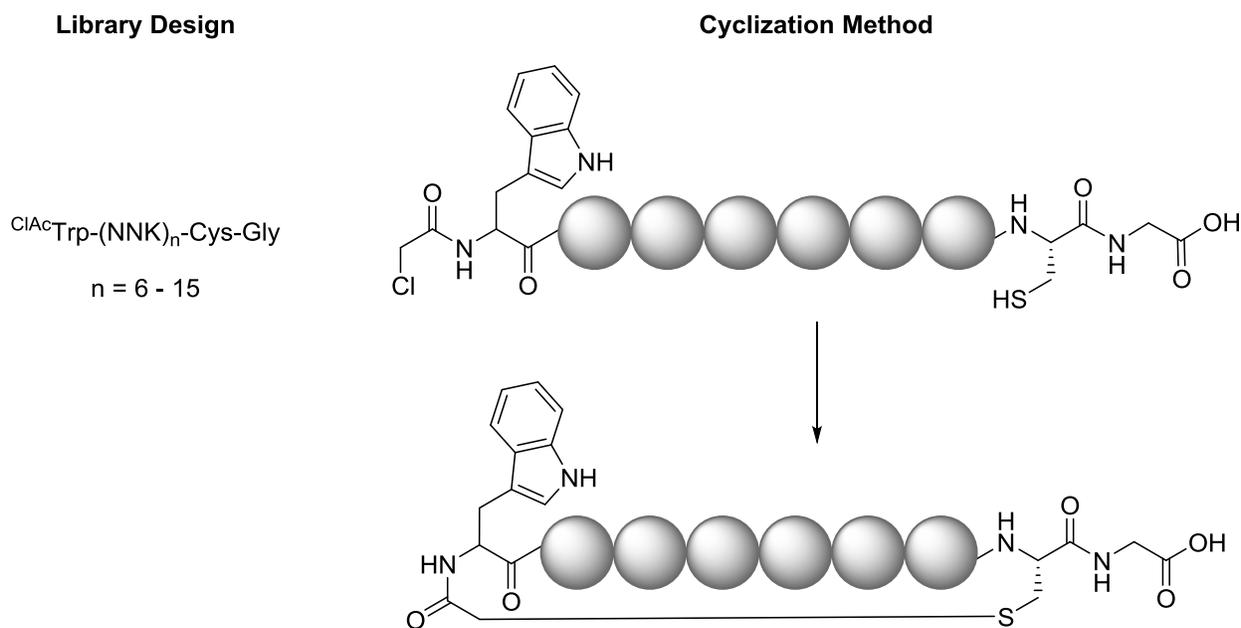
## 4.2 Specific aims

The primary aim of this chapter is to collaborate with Dr. Seino Jongkees in the Suga lab to discover potent peptide-based inhibitors of TcTS. Therefore the first goal is for the Suga lab to employ their RaPID screening system to discover macrocyclic peptide sequences that demonstrate high-affinity binding to our enzyme target. The second goal is to synthesize and subsequently cyclize the selected peptide-hits by solid phase peptide synthesis and chemical macrocyclization. Finally, we will measure the inhibition parameters ( $IC_{50}$  and  $K_i$ ) of the lead peptides and their derivatives.

## Results and Discussion

### 4.3 Screen 1: cyclization between *N*-chloroacetamide and cysteine

In the Suga lab, Dr. Jongkees performed the first round of RaPID screening against TcTS. He employed a peptide scaffold that was designed such that a variable region of 6 to 15 amino acids was flanked by conserved *C*- and *N*-terminal residues used for intramolecular macrocyclization (sequence (ClAc-Trp)-X<sub>6-15</sub>-CGSGSGS). Following translation of the encoded peptides during screening, cyclization was initiated between conserved *N*-chloroacetamidyl D- or L-tryptophan residues and L-cysteine (Figure 4.8).



**Figure 4.8.** First-generation library design and cyclization strategy for RaPID screening.

After eight rounds of selection, eight of the most prevalent peptides were selected for synthesis (Table 4.1). Oligomers **4.1-4.6** were chosen as the most abundant D- and L-initiated sequences from the eighth round of selection, **4.7** was from the seventh round while **4.8** and **4.9** were selected from the fourth round. All sequences contained a high degree of positive charge in the variable region as evidenced by the

3-6 cationic residues present in each peptide. Positive selection was carried out against the HisTrap column bearing resin-immobilized TcTS, while negative selection was conducted against the HisTrap resin. The ratio of positive/negative was represented as percent recovery and used as an indicator of binding affinity for TcTS.

**Table 4.1. First-generation target sequences from RaPID screening.**

Code	Name	Sequence*	Recovery (%)
4.1	TcTS2L8.1	ClAcW-I-R-A-L-W-I-N-H-R-A-L-Q-I-H-C-G	7.7
4.2	TcTS2L8.2	ClAcW-L-T-A-I-L-V-R-G-H-R-R-H-L-T-L-C-G	11.0
4.3	TcTS2L8.3	ClAcW-I-R-S-L-M-I-N-G-H-R-K-H-V-L-C-G	6.1
4.4	TcTS2L8.4	ClAcW-V-W-L-H-Y-P-R-R-L-L-W-C-G	10.7
4.5	TcTS2L8.5	ClAcW-V-Q-F-L-W-V-G-H-H-R-K-A-V-S-C-G	11.1
4.6	TcTS2D8.1	ClAcw-I-V-G-I-W-V-Q-H-N-I-R-H-V-R-G-C-G	4.9
4.7	TcTS2D7.2	ClAcw-K-F-T-S-L-V-V-K-H-H-R-R-L-I-A-C-G	3.1
4.8	TcTS2D4.12	ClAcw-L-R-R-K-C-T-V-R-S-V-D-T-P-H	2.0
4.9	TcTS2D4cons	ClAcw-L-R-R-K-C-T	-

\*L-Amino acids represented by standard one letter codes; ClAc = *N*-chloroacetamide; w = D-tryptophan.

Using standard amino acids building blocks, L- and D- *N*-chloroacetamidyl-tryptophan and an Fmoc/Boc protection scheme, Dr. Jongkees employed automated solid phase peptide synthesis (SPPS) to construct the linear target peptides. Following synthesis, peptides were cleaved from the resin support, their sidechain protection removed, and intramolecular cyclization initiated by the addition of triethylamine (TEA). Cyclized products were purified by HPLC, characterized by MALDI-TOF and sent to me for kinetic analysis.

### 4.3.1 Kinetic analysis

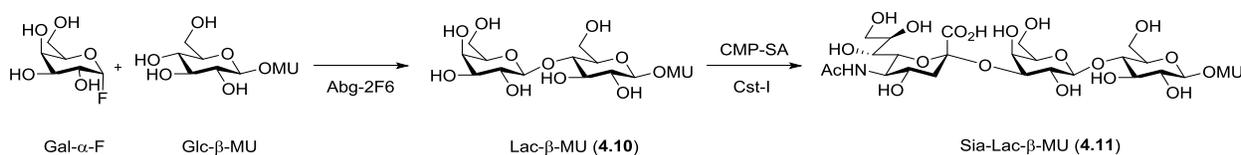
The hit peptides were selected based-on their binding affinities for TcTS, therefore we needed to determine the ability of the peptide hits to modulate TcTS catalysis. To this end, inhibitors were incubated with buffered enzyme at a range of concentrations (0.1 nM - 25  $\mu$ M), the substrate TFMU-SA was added and the initial rates of TcTS hydrolysis monitored by UV/Vis ( $\lambda = 405$  nm). Surprisingly, when these rates were compared to the rates of substrate cleavage in the absence of cyclic peptide, no inhibition was observed. Moreover, TcTS catalysis was enhanced by ~40% versus controls in some cases (Figure 4.9).

Conc ( $\mu$ M)	(+)	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	(+)	(-)
25	0	-9	-15	-18	-38	-23	-	-	-	-	-	100
10	-1	-33	-46	-27	-42	-32	-22	-17	-10	-19	-3	95
1	1	-20	-23	-22	-17	-18	-28	-19	-16	-13	-7	100
0.1	5	-10	-22	-13	-17	-16	-15	-20	-12	-9	-5	100
0.01	2	-9	-11	-8	-11	-10	-7	-14	-9	-5	-5	100
0.001	1	-7	-5	0	-5	-6	-2	-7	-6	-6	-1	100
0.0001	5	-4	-5	-1	-3	-5	-4	-7	-2	-5	-3	99
(+)	5	-2	1	0	3	2	4	2	-4	1	1	98

**Figure 4.9. Percent inhibition of TcTS by cyclic peptides 4.1-4.9.** Heatmap codes for strong inhibition in red, enzyme activation in blue and no change versus control wells in white. Positive wells (+) contain DMSO in place of peptides and negative wells (-) do not contain TcTS.

Therefore, we determined that the selected peptide hits do not inhibit TcTS hydrolysis of TFMU-SA at concentrations up to 10-25  $\mu$ M. To alleviate concerns that the chosen substrate may not be appropriate to mimic binding of the native substrate (in both the donor and acceptor site of TcTS), a trisaccharide substrate bearing a similar coumarin leaving group was synthesized. Sia-Lac- $\beta$ -MU (**4.11**) should harness specific interactions in the lactose binding pocket that are possibly not recruited by TFMU-SA, meaning that if the cyclic peptides interact with TcTS in that region, modulation of hydrolytic activity may only be detected with the larger substrate. Therefore **4.11** was synthesized in two enzymatic steps from the appropriate monosaccharide derivatives (Scheme 4.1). First, the glycosynthase Abg-2F6 catalyzed the assembly of Gal- $\alpha$ -F and Glc- $\beta$ -MU to afford Lac- $\beta$ -MU<sup>177</sup> (**4.10**). Then, using the sialyl-

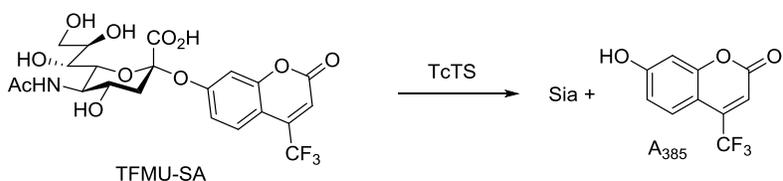
transferase Cst-I the product **4.10** was coupled with CMP-SA and purified to afford the trisaccharide substrate **4.11** in 89% yield<sup>178</sup>.



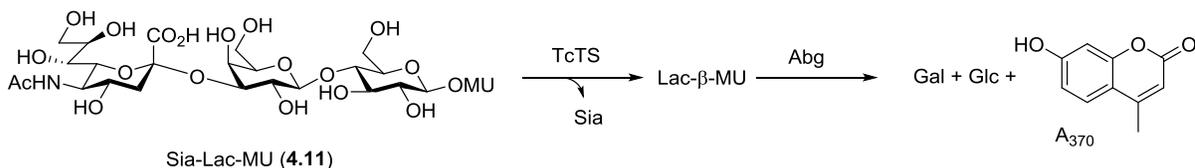
**Scheme 4.1. Biosynthetic preparation of Sia-Lac-MU (4.11).**

While TcTS catalysis is directly responsible for the release of the chromophore utilized in the TFMU-SA based assays, a coupled enzyme system must be used for Sia-Lac-MU assays in order to hydrolyze all glycosidic linkages to release the chromophore for UV detection.  $\beta$ -Glucosidase from *Agrobacterium faecalis* (Abg) is a robust enzyme in the GH1 family that can efficiently catalyze both the hydrolysis of the galactosyl moiety of Lac- $\beta$ -MU and the resultant Glc- $\beta$ -MU product to afford the corresponding monosaccharide products and the MU chromophore<sup>179</sup> (Scheme 4.2). Importantly Abg is catalytically inactive on the starting material Sia-Lac-MU.

#### TcTS Direct Assay



#### TcTS Coupled Assay



**Scheme 4.2. Comparison of TcTS assays using a (top) one-enzyme or (bottom) two-enzyme system.**

Therefore, this system was employed to assay the inhibitory activity of the cyclic peptides **4.1-4.9**. Sia-Lac-MU was added to a pre-incubated solution of TcTS, Abg and varying concentrations of peptide and the resultant rate of enzymatic cleavage monitored. Again, upon comparison of the resultant rates to the control activity, minimal differences were detected. Therefore, it was concluded that the cyclic peptides do not significantly inhibit TcTS activity at concentrations up to 10  $\mu$ M.

#### **4.3.2 Limitations of screening/selections**

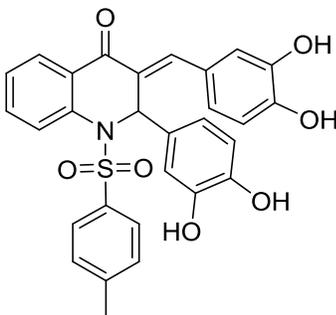
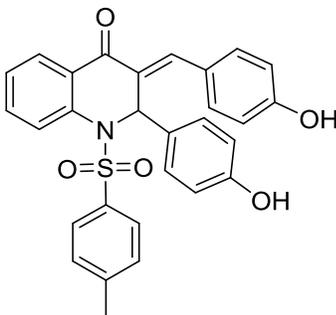
Based on the selection method employed in RaPID screening it is plausible to conclude that the chosen peptides do not bind TcTS in a position that elicits inhibitory effects, instead interacting at a remote site that does not affect catalytic activity. Alternatively, the low recovery values obtained for the lead sequences during selection may indicate a lack of significant sequence enrichment; thus suggesting that the cyclic peptides generated in this library may not exhibit high-affinity binding for TcTS. In either case, it is pertinent to acknowledge the potential selection pitfalls of the given screening method. This example highlights why functional screening, or the addition of a functional assessment step to a screening protocol may be desirable when targeting activity-modulating compounds.

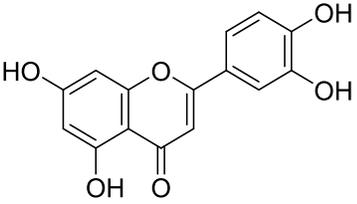
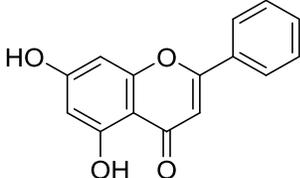
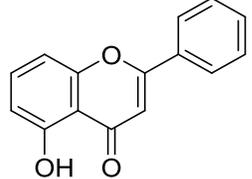
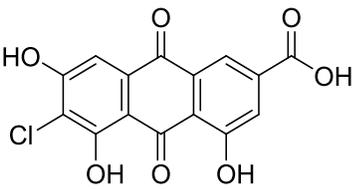
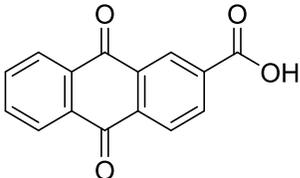
#### **4.4 Screen 2: cyclization between disuccinimidyl glutarate and conserved amines**

The first-generation library employed for RaPID screening encoded for *N*-chloroacetamidyl-tryptophan initiated intramolecularly cyclized peptides containing a variable region of 6-15 proteinogenic amino acids and failed to yield peptide hits that demonstrated any inhibitory activity towards the target enzyme. Therefore a second set of libraries was assembled by Dr. Jongkees which included additional non-proteinogenic amino acids and two alternative cyclization methods in search of better affinity for TcTS. Previously reported non-glycosidic inhibitors of TcTS have commonly included hydroxylated phenolic compounds, often bearing catechol or resorcinol-like substitutions. These compounds are

frequently better inhibitors than their non-substituted analogues<sup>77,78,180</sup>. A few exemplary cases illustrate this point, such as the comparison of the reported inhibition of hydroxylated quinolinone, flavone or anthraquinone derived compounds to their dehydroxylated relatives (**4.12-4.18**; Table 4.2). In the case of luteolin (**4.14**), its inhibitory potency is decreased 15-fold by removal of the hydroxyl moieties on the B-ring catechol (**4.15**) and an additional >3-fold activity loss is observed when the A-ring resorcinol moiety is disrupted (**4.16**). In the case of the 6-chloro-4,5,7-trihydroxy-anthraquinone-2-carboxylic acid (**4.17**), removal of the three hydroxyl groups and the chlorine atom leads to a >700-fold loss of inhibitory activity (**4.18**). Thus, based on the observed activity enhancements from polyhydroxylated phenolic species, specifically catechol and resorcinol-like substitutions, these features were incorporated into the design of the second library for assessment using the RaPID screening protocol.

**Table 4.2. Hydroxylated non-peptide based inhibitors of TcTS.**

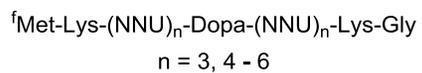
Compound Name	Structure	IC <sub>50</sub>
2,3-Dihydro-2-(3',4'-dihydroxyphenyl)-3-(3'',4''-dihydroxybenzylidene)-1-(p-toluenesulfonyl-4-(1H)-quinolinone ( <b>4.12</b> )		0.6 μM <sup>180</sup>
2,3-Dihydro-2-(4'-hydroxyphenyl)-3-(4''-hydroxybenzylidene)-1-(p-toluenesulfonyl-4-(1H)-quinolinone ( <b>4.13</b> )		30 μM <sup>180</sup>

Compound Name	Structure	IC <sub>50</sub>
Luteolin (4.14)		22 μM <sup>78</sup>
Chrysin (4.15)		340 μM <sup>78</sup>
5-Hydroxyflavone (4.16)		>1000 μM <sup>78</sup>
6-Chloro-4,5,7-trihydroxy-anthraquinone-2-carboxylic acid (4.17)		0.58 μM <sup>78</sup>
Anthraquinone-2-carboxylic acid (4.18)		410 μM <sup>78</sup>

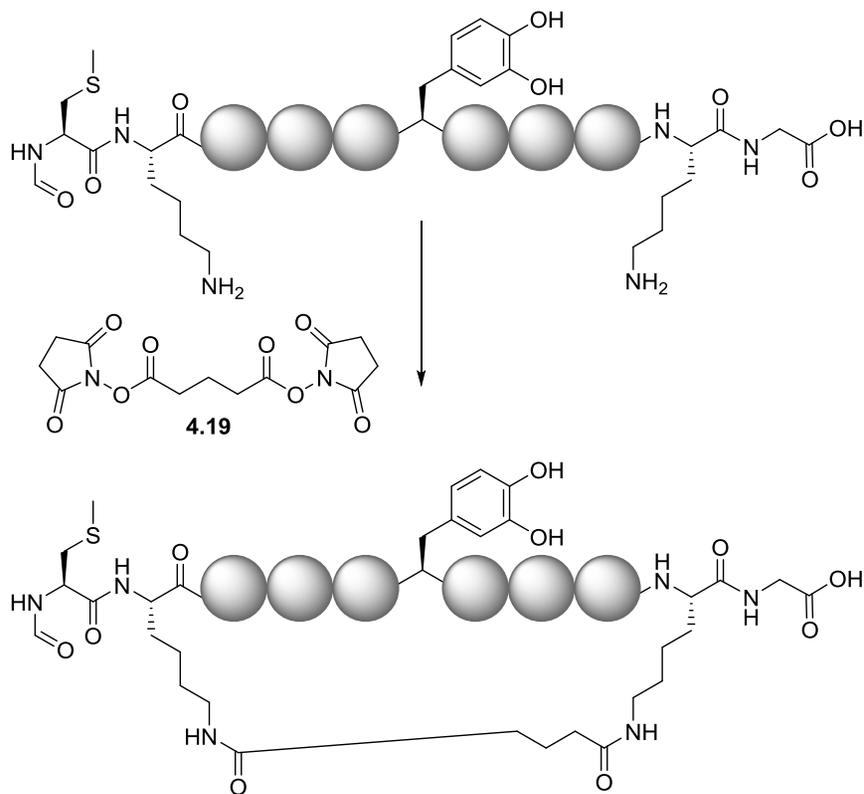
In the Suga lab, three second-generation libraries were constructed. The first library employed the same scaffold as that used in the first-generation screening (Figure 4.8) except that the *C*-terminal tryptophan residue has been replaced by tyrosine. The second library presents a new motif whereby two variable regions of 3-6 amino acids in length flank a conserved cysteine residue. Additionally, a pair of amine-bearing residues are conserved at either end of the sequence in the form of a *C*-terminal lysine residue and an *N*-terminal 4-aminomethylbenzoic acid. The third library closely resembles the second

with minor modifications: the *N*-terminal residues have been replaced by formyl-methionine (<sup>f</sup>Met) and lysine, while the conserved cysteine residue has been substituted by L-dopa. In both the second and third library, the pair of nitrogenous residues at opposing termini were post-translationally cyclized using an exogenous stapling agent, disuccinimidyl glutarate (DSG, **4.19**), to form the desired cyclic structures (Figure 4.10).

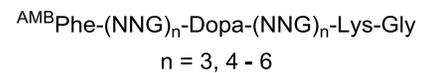
### Library Design



### Cyclization Method



### Library Design



### Cyclization Method

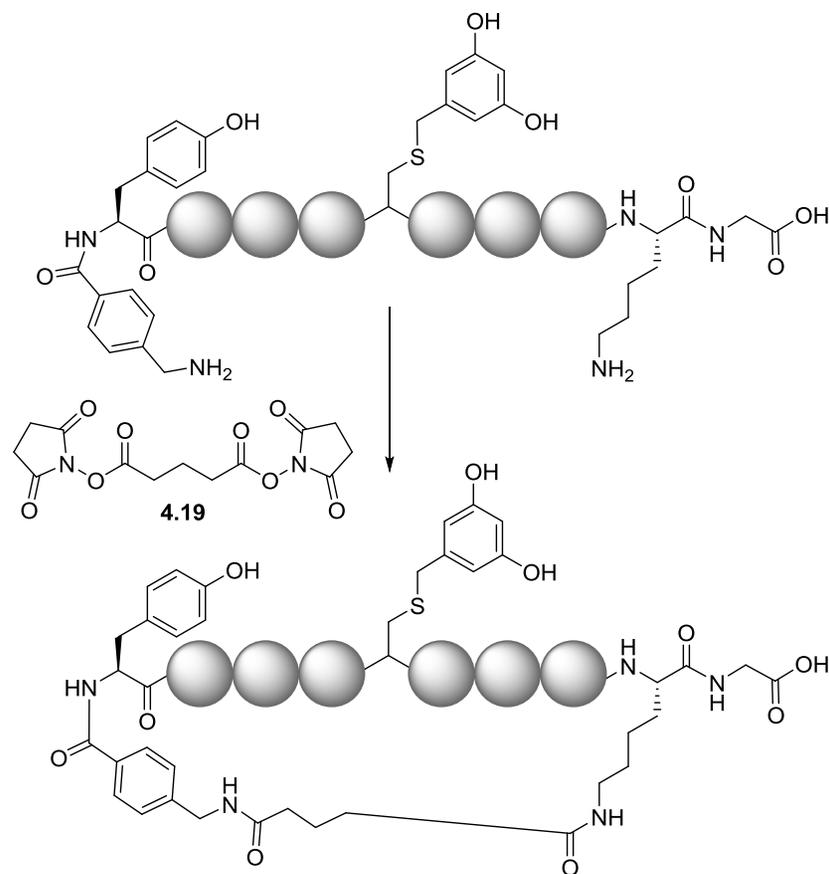
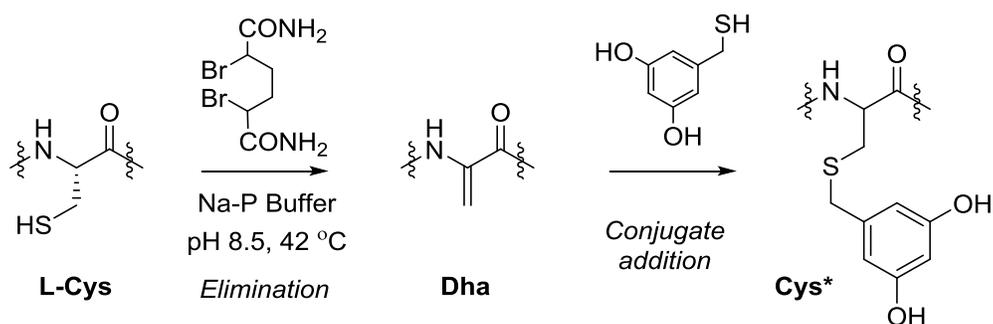


Figure 4.10. Second-generation library design and cyclization strategies for RaPID screening.

The variable regions of the peptide sequences were randomized using either NNK (19 amino acids), NNG (13 amino acids and L-dopa by reprogramming) or NNU (16 amino acids including Cys) random codons during FIT system translation. Additionally, the fixed L-cysteine residue was modified post-translationally for DSG cyclized libraries to yield a new thioether side chain (Cys\*) bearing a resorcinol-like hydroxylation pattern. This was installed in two steps by Dr. Jongkees: first, activation of cysteine with  $\alpha,\alpha'$ -dibromoadipic-bis-amide (DBAA) and elimination of the sulfur moiety in buffer (pH = 8.5) afforded dehydroalanine (Dha); second, conjugate addition to Dha with the 3,5-dihydroxybenzyl thiol yielded the desired thioether product Cys\* (Scheme 4.3).



**Scheme 4.3. Post-translational modification of cysteine during RaPID screening<sup>176,181</sup>.**

The recovery percentages for the most abundant peptides are reported in Table 4.3 after 5-6 rounds of RaPID screening (calculated in the same manner as for ‘Screen 1’). These values are significantly higher than those observed in the first-generation screen after 7-8 rounds of selection.

**Table 4.3. Second-generation target sequences from RaPID screening.**

Code	Name	Sequence*	Recovery (%)
<b>4.20</b>	I4-6TcTS6.4	AMB-F-H-d-S-R-d-C*-S-d-H-d-W-K-G	52
<b>4.21</b>	I4-6TcTS6.13	AMB-F-T-R-d-d-C*-V-d-W-W-K-G	52
<b>4.22</b>	I3TcTS6.1	AMB-F-Q-W-Y-W-K-V-W-K-Q-C*-G	26
<b>4.23</b>	I3TcTS6.2	AMB-F-Q-d-H-C*-H-d-H-K-G	35

\*L-Amino acids represented by standard one letter codes; AMB = 4-(aminomethyl)benzoate; d = L-dopa; C\* = S-(3,5-dihydroxybenzyl)cysteine.

Two sequences encoding for cyclic peptides **4.20** and **4.21** appeared to be the most promising candidates with percent recovery values >50, thus were selected for further investigation. Interestingly, the selected peptides shared some structural similarity, each sequence incorporates a large percentage of aromatic amino acids, especially the catechol-bearing residue L-dopa (3/8 and 4/10 respectively).

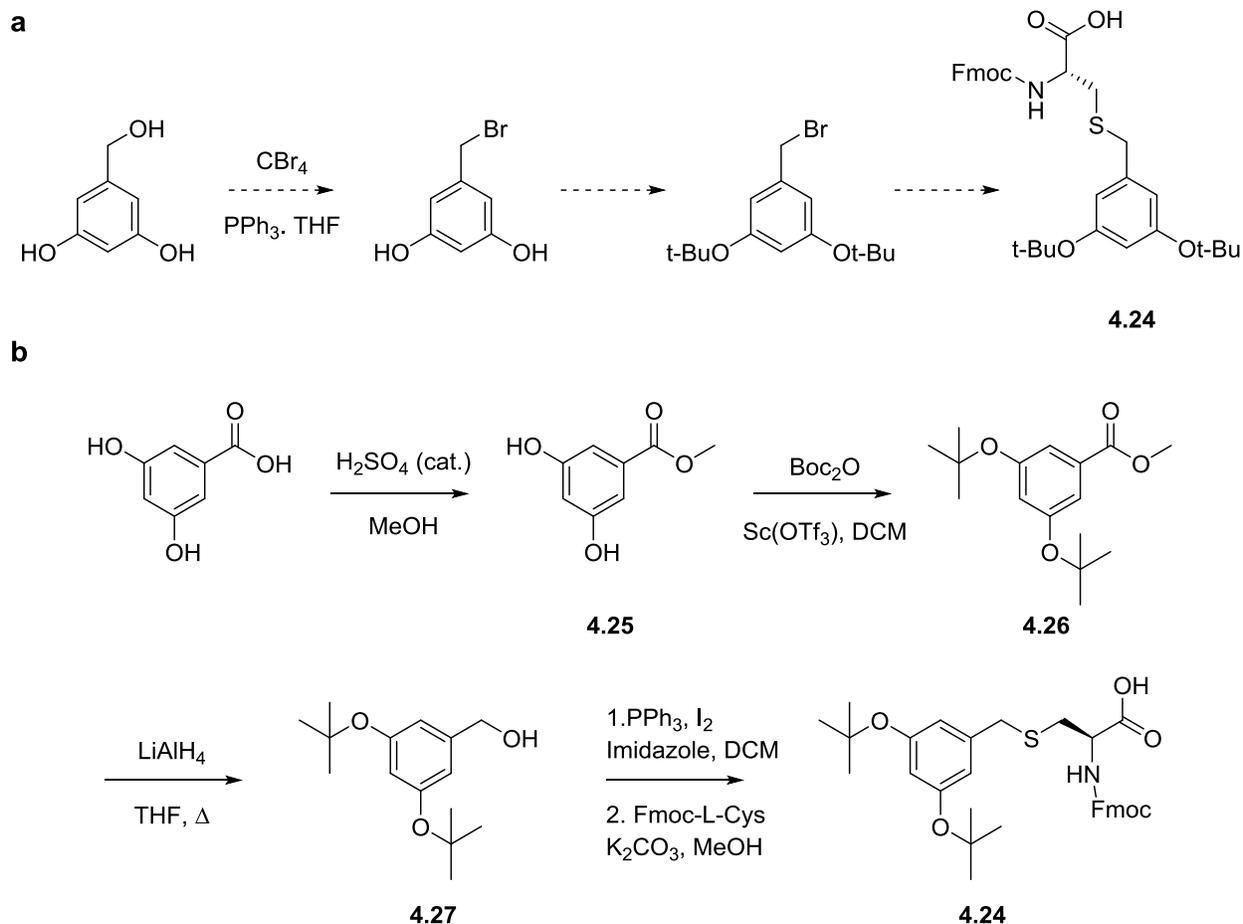
## 4.5 Peptide synthesis and macrocyclization

### 4.5.1 Synthesis of L-Cys\*

The post-translational modification of cysteine employed during screening yielded a mixture of cysteine isomers (L- and D-) *in situ*. In order to establish if a given isomer is preferred for enzyme binding, access to the stereochemically defined cysteine derivatives was required prior to peptide synthesis. To this end two synthetic routes for the desired cysteine analogue (L-Cys\*) were explored.

*Tert*-butyl protection of the tyrosine sidechain is most often employed during Fmoc/Boc SPPS, thus the same scheme was selected for resorcinol protection. Possible synthetic routes to the target compound Fmoc-L-Cys\*(*t*Bu)-OH (**4.24**) were considered. The most direct synthetic route achieves the target in three steps: (1) bromination of 5-(hydroxymethyl)resorcinol followed by (2) *tert*-butyl protection of the resorcinol alcohols and (3) coupling with Fmoc-L-Cys-OH to yield **4.24** (Scheme 4.4a). However,

this method was quickly foiled by the instability of the 5-(bromomethyl)resorcinol product, thus an alternative route was investigated.



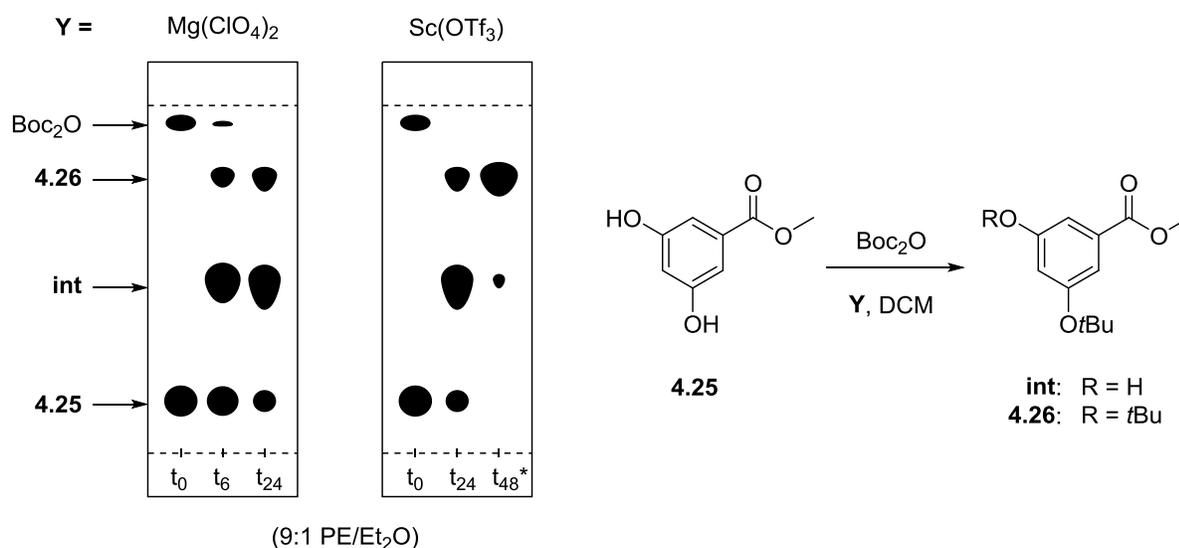
**Scheme 4.4. Proposed (a) and executed (b) synthesis of 4.24.**

The second approach began by installing the *t*-butyl ether protection scheme on the resorcinol-like hydroxyl groups prior to activating the benzyl alcohol. From methyl-3,5-dihydroxybenzoate the desired *t*-butyl ethers can be formed. The product can be reduced to yield the benzylic alcohol which can in turn be activated for coupling with Fmoc-Cys-OH using a myriad of standard methods (Scheme 4.4b).

Preparation of the starting material **4.25** was accomplished in 88% yield via acid-catalyzed esterification of 3,5-dihydroxybenzoic acid in methanol. *Tert*-butyl protection of the L-tyrosine side chain is classically achieved through reaction of isobutylene with an acid catalyst to generate the reactive

carbocationic intermediate *in situ*<sup>182</sup>. Although the reported *t*-butyl ether formation proceeds efficiently we decided to explore alternative chemical installation methods due to a lack of access to isobutylene gas and to avoid the competing Friedel-Crafts alkylation pathway<sup>183</sup>.

In 2005, Bartoli and colleagues described the mild and efficient formation of *t*-butyl esters using *t*-butyl dicarbonate (Boc<sub>2</sub>O) and magnesium perchlorate<sup>184</sup>. Their reports indicate that a wide range of aliphatic or aromatic alcohols react in 71-95% yield to form the desired ether products via the concerted cyclic decomposition of the mixed dicarbonate intermediate. Following the reported conditions (2.3:1 Boc<sub>2</sub>O/alcohol, 10% Mg(ClO<sub>4</sub>)<sub>2</sub>, DCM, 40 °C), I was able to obtain the desired product in ~15% yield. However, TLC of the reaction clearly showed three spots corresponding to the starting material (**4.25**), the mono-protected intermediate species (int) and the desired product (**4.26**; Figure 4.11). Attempts to increase product formation by lengthening reaction time and increasing catalyst loading or *t*-butyl dicarbonate equivalents failed to generate significant improvements to reaction yield; all the while TLC continued to indicate that the majority of the reaction material had ‘stalled’ at the intermediate species. Batch-wise addition of *t*-butyl dicarbonate over the course of the reaction, as opposed to a single addition of the reagent, seemed to provide modest improvements. This can likely be attributed to the competing Mg(ClO<sub>4</sub>)<sub>2</sub>-catalyzed degradation of the *t*-butyl dicarbonate into CO<sub>2</sub>, *t*-BuOH and isobutylene over the lengthened reaction times in refluxing conditions.

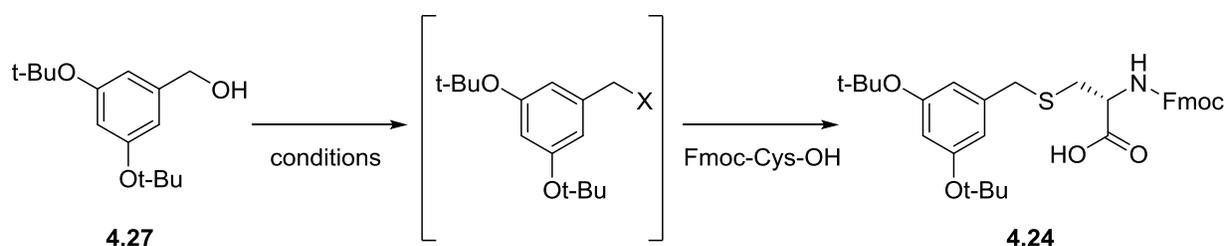


**Figure 4.11. Reaction progress (TLC) and reaction scheme for *t*-butyl ester formation during synthesis of 4.26.**

Alternative catalysts have also been surveyed for *t*-butyl ester formation and their catalytic properties assessed<sup>185</sup>. It was determined that for phenolic alcohols scandium triflate (Sc(OTf)<sub>3</sub>) catalysis yielded the product in higher yield than the Mg(ClO<sub>4</sub>)<sub>2</sub> catalyzed reactions at ambient temperatures and with a 2-fold decrease of catalyst loading. These procedures were especially appealing as ambient temperatures may slow the degradation of the anhydride reagent and positively affect reaction yield. However, substitution of the acid catalyst did not drive the reaction to completion. One hypothesis pertaining to the ‘stalled’ reaction progress was that a buildup of *t*-BuOH over the course of the reaction may compete with the hydroxyl moiety of the resorcinol intermediate for reactivity towards Boc<sub>2</sub>O, leading to a state of equilibrium being established. To test this theory the ‘stalled’ reaction – after 24 hours of stirring using the conditions above – was evaporated to dryness in order to remove all residual solvent, Boc<sub>2</sub>O and *t*-BuOH. Fresh aliquots of DCM, ScOTf<sub>3</sub> and Boc<sub>2</sub>O were added to the reaction flask and stirred overnight. After 24 hours TLC indicated that the reaction had converted nearly all starting material and intermediate to the desired di-*t*-butyl-protected product (Figure 4.12). Subsequent workup and purification by silica gel chromatography afforded **4.26** in 78% yield.

Upon successful *t*-butyl ether protection of the hydroxyl groups the benzyl ester was reduced by lithium aluminum hydride in anhydrous THF to afford the benzyl alcohol **4.27** in 96% yield<sup>186</sup>. With the appropriately protected 5-(hydroxymethyl)resorcinol in hand, the final coupling step with Fmoc-L-Cys-OH was investigated. To leverage the nucleophilic nature of the cysteine sidechain a survey of the different approaches to activate benzyl alcohol towards substitution was conducted (Table 4.4).

**Table 4.4. Synthetic methods explored for synthesis of 4.24.**



Entry	Method	Leaving Group	Conditions	Yield	Comments
1	Bromination	R-Br	PBr <sub>3</sub> (1.2 eq), DCM	None	Boc removal
2	Bromination	R-Br	PBr <sub>3</sub> (1.5 eq), Imidazole (3 eq), DCM	None	
3	Sulfonate activation	R-OTf	Tf <sub>2</sub> O, py	None	Solvent reactivity
4	Sulfonate activation <sup>187</sup>	R-OMs	MsCl (1.4 eq), TEA (2 eq), DCM, -40 °C	None	Side reactivity
5	Sulfonate activation	R-OMs	MsCl (1.2 eq), DIPEA (1.2 eq), DCM, -40 C	Low	
6	Mitsunobu Reaction	R-OPPh <sub>3</sub>	PPh <sub>3</sub> (2 eq), DIAD (2 eq), Cys (1.2 eq)	None	DIAD reactivity
7	Appel Reaction	R-I	PPh <sub>3</sub> (2 eq), Iodine (2 eq), Imidazole (2.2 eq), DCM	90%	

L-Cysteine can be readily alkylated in the presence of benzyl bromide<sup>188</sup>, thus bromination of **4.27** was attempted with phosphorus tribromide (PBr<sub>3</sub>); however, no product was achieved, instead the acid-labile protecting group was removed during the course of the reaction. My attempts to mediate this effect with imidazole failed to improve the reaction success. Sulfonyl ester activation of aryl or aliphatic alcohols – synthesized from the corresponding sulfonyl chloride or anhydride – is a versatile approach often employed in substitution reactions. The excellent leaving group ability of the common mesylate, tosylate or triflate species allows for facile replacement of hydroxyl groups by good nucleophiles. However it was evident that upon formation of the activated sulfonate species substitution immediately occurs *in situ* by any nucleophilic species (*e.g.* solvent or base employed in the reaction) to afford undesired products. By selecting the least activated of the three sulfonate species and low temperatures, reaction with mesyl chloride and diisopropylamine at -40 °C for one hour quantifiable amounts of the desired product could be obtained. However, the low yield, delicate nature of the reaction and the instability of the product led to a continued search for activation conditions.

In response to the unsuccessful attempts to isolate the activated form of the benzylic alcohol, Mitsunobu reaction conditions were explored. These protocols employ a two-step one-pot synthesis of the desired product by step-wise addition of reagents. This reaction leverages the reactivity of the oxyphosphonium intermediate *in situ* to generate the thioether product in a single reaction. However the azodicarboxylate reagent (DIAD) proved too nucleophilic for this approach, as the only isolated species after addition of Fmoc-L-Cys-OH was the substitution product of the activated 3,5-di-*t*-butyl-benzyl alcohol with DIAD.

In search of a nucleophile-free non-acidic method to activate **4.27** for cysteine alkylation, the Appel reaction was investigated. These conditions (Table 4.4, entry 7) induce benzyl iodide formation through the analogous oxyphosphonium ion intermediate employed in the Mitsunobu reaction. In the absence of competing nucleophiles iodide readily attacks the reactive intermediate to obtain the benzyl iodide target **4.28** in 90% yield. The iodinated product was found to be surprisingly stable as it was

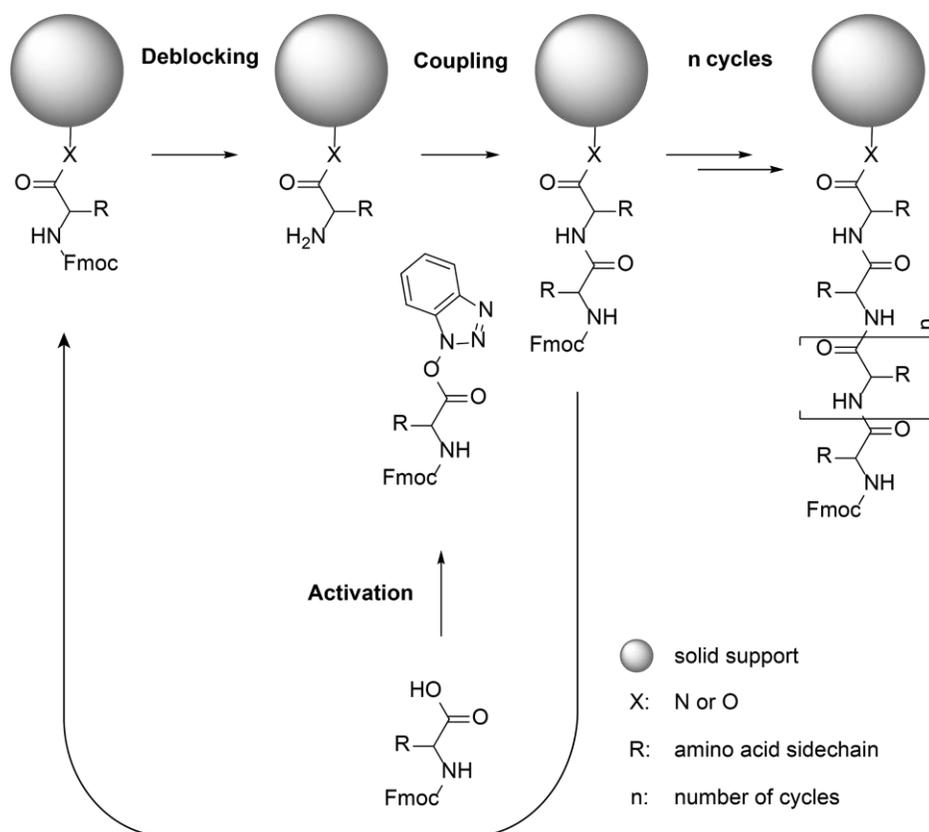
amenable to purification by silica gel column chromatography. Upon cooling the oily product crystallized to afford an off-white solid that has remained intact when subject to long-term storage at -20 °C.

Alkylation of cysteine with excess 3,5-di-*t*-butyl-benzyl iodide was finally achieved in the presence of potassium carbonate after stirring for one hour in methanol. Alternative organic base options were explored but undesired Fmoc removal was observed due to the base-labile nature of the protection group. The L-Cys\* product was purified by silica gel chromatography to afford **4.24** in 67% yield.

#### **4.5.2 Manual synthesis of non-proteinogenic peptide hits**

With all of the necessary starting materials in hand, synthesis of the target peptides **4.20** and **4.21** was pursued. Standard Fmoc/Boc solid phase peptide synthesis (SPPS) techniques were employed for the manual assembly of the linear peptide targets<sup>189</sup>. This process consists of three sequential steps that can be iteratively repeated until the desired peptide has been synthesized (Figure 4.12). A fourth step *capping* can be employed to terminate the elongation of incomplete peptides by reacting residual free amines after coupling with acetic anhydride, however this step was not carried out in any synthesis described herein.

- (1) *Deblocking* of the *N*-terminal amine is achieved by base activated elimination of the Fmoc protecting group.
- (2) *Activation* of the carboxylate moiety of the reacting amino acid is commonly achieved by addition of hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) to afford the activated HOBt-ester of the protected amino acid.
- (3) *Coupling* of the deblocked resin-bound amine with an excess of the activated ester of the impending amino acid is achieved by mixing. The presence of free amine, and thus reaction progress can be monitored using the Kaiser test.



**Figure 4.12. General summary of solid phase peptide synthesis (SPPS).**

The solid phase resin Tentagel S PHB-Gly-Fmoc was employed on a 25  $\mu\text{mol}$  scale as the anchor for the manual synthesis of **4.20** and **4.21** in the *C*- to *N*-terminus direction. The selected resin is composed of polyethylene glycol-grafted polystyrene support and presents a less than average loading capacity (0.22 mmol/g). Together these features increase resin swelling and disfavour interactions between neighbouring peptides or between the peptide and the solid support. Further, this resin is pre-charged with a *C*-terminal glycine residue bound via ester linkage with *p*-hydroxybenzyl alcohol. *Deblocking* of the *N*-terminal resin-bound amine was achieved through two rounds of 5-10 minute reactions with piperidine/DMF (1:4). *Activation* of the incoming protected amino acid was achieved by mixing for 10 minutes with equimolar amounts of HBTU and *N*-methylmorpholine (NMM) to afford the corresponding HOBt activated esters. *Coupling* procedures provided the most experimental variation by

residue, as the equivalents of amino acid, reaction length and number of coupling rounds were adjusted to best accommodate each coupling step (Table 4.5). The standard coupling procedures employed 5 equivalents of activated amino acid and reaction times between 20-30 minutes. In the case of difficult coupling steps the reaction times were lengthened to a maximum of 120 minutes and if completion was not observed a second round of coupling with an additional 5 equivalents of amino acid was conducted. Coupling efficiencies decreased with increasing peptide length (as is well documented<sup>190</sup>) and when introducing sequential amino acids bearing large side chains (*e.g.* tryptophan). For reactions involving the non-proteinogenic amino acids, 1.5 equivalents of amino acid was used to conserve materials. To compensate, reaction times were lengthened to up to 120 minutes and one to three rounds of couplings were performed. Peptide synthesis was completed following the addition of the *N*-terminal residue 4-(*boc*-aminomethyl)benzoic acid (*N*-*boc*-AMB) for each sequence. The Kaiser test was used to assess the completion of the coupling reaction (by detecting if any unreacted amines remain in solution) after each step. If residual amines were detected, the coupling reactions were repeated.

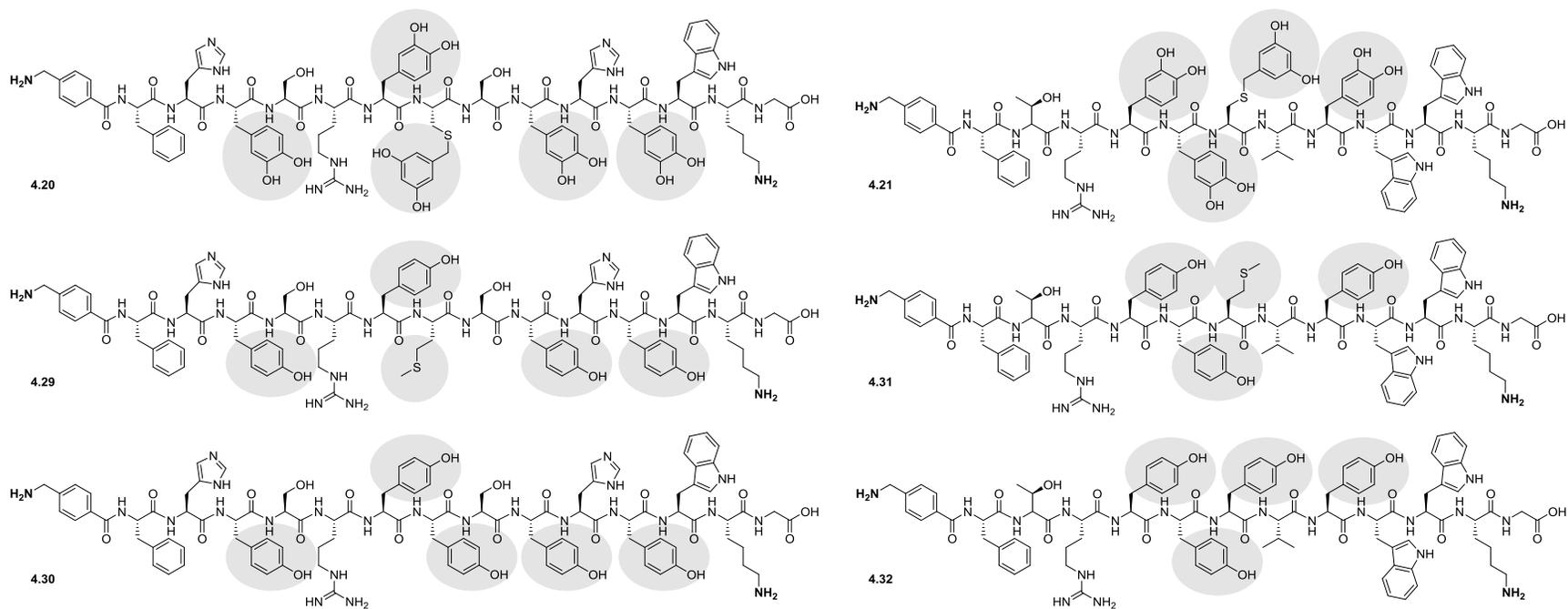
**Table 4.5. Coupling conditions for manual SPPS of 4.20 and 4.21.**

Peptide 4.20			Peptide 4.21		
Residue	eq. (rounds)	Time (min)	Residue	eq. (rounds)	Time (min)
<b>K</b>	5 (2)	30, 30	<b>K</b>	5 (1)	30
<b>W</b>	5 (1)	30	<b>W</b>	5 (1)	30
<b>d</b>	1.5 (1)	60	<b>W</b>	5 (3)	30, 60, 120
<b>H</b>	5 (1)	30	<b>d</b>	1.5 (1)	60
<b>d</b>	1.5 (1)	60	<b>V</b>	5 (1)	30
<b>S</b>	5 (1)	45	<b>C*</b>	1.5 (1)	60
<b>C*</b>	1.5 (1)	90	<b>d</b>	1.5 (3)	45, 60, 90
<b>d</b>	1.5 (3)	45, 60, 90	<b>d</b>	1.5 (2)	60, 90
<b>R</b>	5 (1)	45	<b>R</b>	5 (2)	60, 120
<b>S</b>	5 (3)	45, 60, 120	<b>T</b>	5 (2)	60, 120
<b>d</b>	1.5 (3)	45, 60, 120	<b>F</b>	5 (2)	60, 120
<b>H</b>	5 (3)	45, 60, 120	<b>AMB</b>	5 (2)	60, 120
<b>F</b>	5 (3)	45, 60, 120			
<b>AMB</b>	5 (2)	45, 60			

\*L-Amino acids represented by standard one letter codes; AMB = 4-(aminomethyl)benzoate; d = L-dopa; C\* = S-(3,5-dihydroxybenzyl)-L-cysteine.

### 4.5.3 Automated synthesis of proteinogenic peptide analogues

In order to assess the impact of adding non-proteinogenic polyhydroxylated phenolic residues to the screening library, a set of peptides bearing only proteinogenic amino acids was envisioned (Figure 4.13). For each hit peptide (**4.20** and **4.21**) two ‘natural’ derivatives were synthesized in which L-Dopa was substituted by L-Tyr and the modified L-Cys\* residue was replaced by L-Met or L-Tyr respectively. Methionine was chosen to emulate the thioether functionality of L-Cys\* while tyrosine was the most similar proteinogenic amino acid substitute for the resorcinol-like modification.



**Figure 4.13. Structure of linear peptides: 4.20 and ‘natural’ derivatives 4.29-4.30 (left); 4.21 and ‘natural’ derivatives 4.31-4.32 (right).**

Linear peptides **4.29-4.32** were synthesized on 5  $\mu$ mol scale/well using the MultipPep automated parallel SPPS instrument (Intravis AG, Bioanalytical Instruments). Each well was loaded with Tentagel S PHB-Gly-Fmoc resin, deblocking was accomplished with piperidine and HOBt-activated amino acid (5 eq.) couplings were performed twice per residue for 10-20 minutes respectively. Iterative coupling steps yielded the desired resin-charged linear peptides.

#### **4.5.4 Deprotection and purification**

Cleavage of the linear peptides from the resin support and removal of the acid-labile side chain protecting groups employed in this synthesis can be achieved in acidic conditions. For sequences containing Cys, Met, Trp, Arg or His, cleavage cocktails that include carbocation scavengers such as triisopropylsilane (TIPS) are recommended<sup>191</sup>. As such Reagent B (88:5:5:2 trifluoroacetic acid (TFA)/phenol/water/TIPS) was freshly prepared and reacted with peptides **4.20-4.21** and **4.29-4.32** respectively to achieve global deprotection and resin cleavage. Peptide solutions were precipitated by 10-fold dilution in ice-cold diethyl ether; the precipitate was pelleted and washed 3-5 times with fresh ether. After drying, the white solids were dissolved in acetonitrile/water/TFA (1:1:0.002) solution and purified by semi-preparative C18 reverse phase HPLC.

#### **4.5.5 Macrocyclization of linear peptides**

In order to cyclize the purified linear peptides, intermolecular cross-linking with the disuccinimidyl glutarate (DSG) was required. The activated linker was synthesized from *N*-hydroxysuccinimide and glutaryl chloride; recrystallization from ether afforded **4.19** in 97% yield. Attempts to replicate the literature cyclization protocol<sup>192</sup> upon which the screening method was based failed to produce significant amounts of the target products. Poor solubility of the linear peptides in the reported conditions (9:1 50 mM HEPES, pH = 8/DMF) made the target concentration of 1 mM unachievable. Reaction parameters were arrayed (Table 4.6) in an attempt to obtain optimal conditions for

the target peptides, but only qualitative amounts of the peptide were obtained in the best case scenario. Complications arising from peptide solubility, DSG instability in aqueous conditions and undesired side reactivity led me to explore alternative approaches.

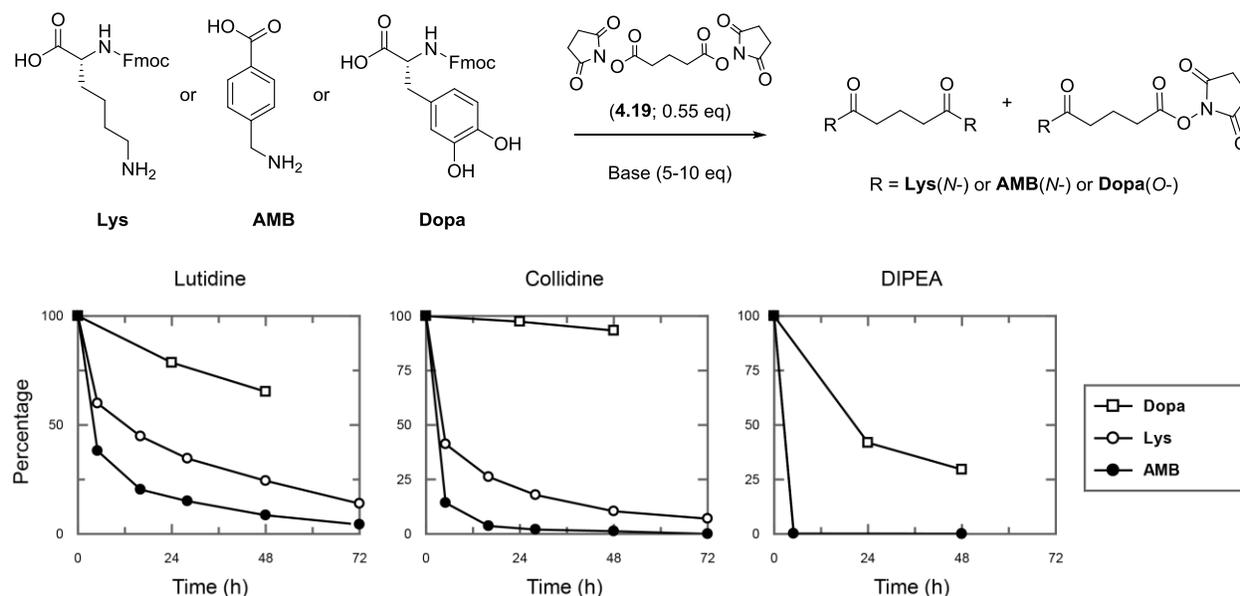
**Table 4.6. Aqueous cyclization conditions explored with linear peptides and DSG.**

Entry	Peptide	Conc. ( $\mu\text{M}$ )	DSG ( $\mu\text{M}$ )	Vol. (mL)	Yield/Comment
1	<b>4.31</b>	150	300	1	Product detected by MALDI
2	<b>4.31</b>	150	750	1	Product detected by MALDI
3	<b>4.31</b>	150	1500	1	SM consumed, no product
4	<b>4.31</b>	150	3000	1	SM consumed, no product
5	<b>4.29</b>	150	300	10	Product observed on HPLC; <10% yield
6	<b>4.31</b>	150	300	10	Product observed on HPLC; <10% yield

One tactic that can both eliminate DSG degradation and limit potential intermolecular polymerization is on-resin coupling. Orthogonal protection of the reactive amine functional groups is required such that selective deblocking and cyclization can occur prior to global deprotection. To this end 4-(Fmoc-aminomethyl)benzoic acid and Fmoc-Lys(Mtt)-OH were employed in place of their Boc protected versions. Automated SPPS was employed to assemble the orthogonally protected derivatives of **4.29-4.32**, the *N*-terminal Fmoc was removed using standard deblocking protocols and the *N*- $\epsilon$ -4-methyltrityl (Mtt) protection was removed with a solution of 1% TFA in DCM<sup>193</sup>. Subsequent reaction of the resin-bound linear peptides with a solution of DSG and diisopropylethylamine (DIPEA) in DMF was initiated and the reaction mixed for 2 hours. Following resin-cleavage and global deprotection, the peptides were analyzed by HPLC. The resultant traces were littered with peaks and did not appear to indicate the synthesis of unique peptide products. Additionally, MALDI-TOF analysis was conducted on all samples collected throughout the entire elution, none of which corresponded to the expected mass of

the desired products. Therefore we determined that the inability to appropriately monitor reaction progress and the inefficient synthesis of the linear peptides made this approach untenable.

As a result of the observed instability of DSG in aqueous solution and failures with on-resin cyclization, alternative off-resin cyclization procedures employing organic solvents, a soluble base and purified linear peptides were explored. In this case, the coupling conditions, including the choice of organic base, were first optimized between DSG and the individual amino acids of interest. Three such residues were selected: target amine-bearing residues (1) Fmoc-Lys-OH and (2) 4-(aminomethyl)benzoic acid (AMB) as well as (3) Fmoc-Dopa-OH, since its catechol functionality may present an alternative site for base mediated reaction with DSG. Three organic bases were selected for assessment: (1) DIPEA ( $pK_a = 10.75$ ), (2) 2,4,6-collidine ( $pK_a = 7.43$ ), (3) 2,6-lutidine ( $pK_a = 6.6$ ). Nine unique analytical scale reactions were conducted whereby each amino acid was treated with DSG in the presence of each selected base for 48-72 hours. Aliquots were taken periodically, quenched with TFA and analyzed by HPLC-MS. Signals corresponding to starting materials, intermediates and products were integrated and the ratio of products calculated. Starting material consumption over the reaction progress was plotted (Figure 4.14).



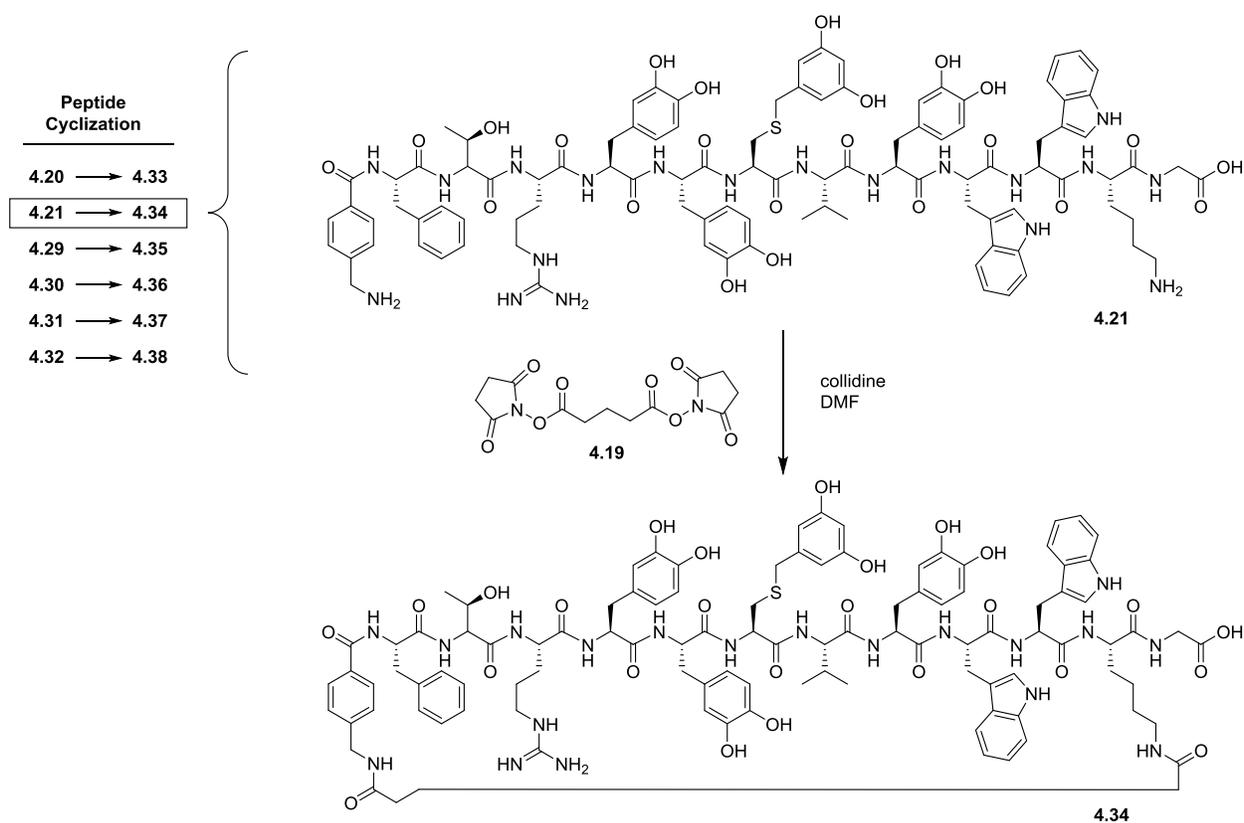
**Figure 4.14.** Reaction scheme (top) and progress (bottom) of amino acids (Lys, AMB and Dopa) with DSG in the presence of varying organic bases.

Coupling of both Lys and AMB respectively with DSG proceeds very efficiently in the presence of DIPEA, each showing complete consumption of starting material in less than 5 hours. However DSG also reacts efficiently with Dopa making it a poor choice for use with the linear peptides. Practically, solubility issues were also encountered at millimolar concentrations in DMF following deprotonation of the many catechol residues on the linear peptides. On the other hand, the reactions containing either lutidine or collidine also showed reactivity between DSG and Lys or AMB; where the coupling reactions approach completion after 24-72 hours. More importantly collidine does not facilitate the reaction of Dopa with DSG; while in the presence of lutidine ~30% of the Dopa starting material is consumed after 72 hours. Thus it was concluded that collidine offers the optimal base reactivity and selectivity with the amino acids of interest, providing a basis from which to optimize the peptide cyclization protocol. The concentration of peptide **4.31** was thus varied between 1 mM and 10 mM and that concentration which both maximized reaction speed and minimized by-product formation was selected (Table 4.7).

**Table 4.7. Reaction conditions tested for-collidine mediated coupling between 4.31 and DSG.**

Entry	4.31 (mM)	DSG (mM)	Collidine (mM)	Yield/Comment
1	10	11	100	Bimolecular reaction product
2	2	2.2	20	Selected conditions
3	1	1.1	10	Incomplete after 5 days

Optimized reaction conditions of 2 mM peptide, 2.2 mM DSG and 10 equivalents of collidine (Table 4.7; Entry 2) were employed to cyclize the six linear peptides to afford the cyclic peptide products **4.33-4.38** (Scheme 4.5). Purification by HPLC afforded the cyclic peptides as white solids, which were subsequently characterized by MALDI-TOF and high-resolution mass spectrometry (Waters Q-TOF spectrometer). Solutions of each linear and cyclic peptide (10 mM) were prepared in DMSO and their inhibition behaviour assessed.



**Scheme 4.5. Cyclization of linear peptides with DSG and collidine.**

## 4.6 Kinetic analysis

With the target peptides in hand the inhibition of TcTS by the linear and cyclic versions was assessed. Enzyme activity was assayed in a 96-well plate format and the rate of TFMU-SA cleavage by UV/Vis was monitored in the presence of varying peptide concentrations from 1-200  $\mu\text{M}$  (Figure 4.15).

Conc ( $\mu\text{M}$ )	Ctrl	4.20	4.21	4.29	4.30	4.31	4.32	Ctrl
200	1	35	2	19	11	-7	0	10.1
100	8	10	3	20	1	-3	-10	1.9
50	2	-8	-16	-3	-6	-24	-28	5.8
20	-3	-12	-20	-12	-9	-22	-26	-1.7
10	-8	-3	-10	-9	-6	-12	-10	-9.7
1	-1	-5	-6	-12	-8	-14	-11	-7.1
Conc ( $\mu\text{M}$ )	Ctrl	4.33	4.34	4.35	4.36	4.37	4.38	Ctrl
200	5	98	49	50	54	75	76	-6.6
100	1	88	33	28	27	74	67	-0.9
50	6	80	19	13	12	24	26	-11.1
20	-5	93	16	-5	5	6	6	-6.6
10	0	83	6	-13	-8	-3	-8	-1.0
1	-6	3	1	-9	-6	-7	-5	1.3

**Figure 4.15. Percent inhibition of TcTS by peptides 4.20-4.21; 4.29-4.38.**

Using the data analysis software GraFit 7.0 (Erithacus) the half-maximal inhibition ( $\text{IC}_{50}$ ) was determined for all cyclic peptides (all plots can be found in Appendix D). Cyclic peptides were insoluble above  $200 \mu\text{M}$  in the assay conditions limiting my ability to measure  $\text{IC}_{50}$  values for the more poorly inhibitory peptides. The most potent species (**4.33**) required that the range of concentrations be altered to appropriately characterize the inhibition. Thus the inhibition assay was repeated using the Cary4000 UV/Vis spectrometer at a lower range of peptide concentrations ( $0.1\text{-}50 \mu\text{M}$ ) and the  $\text{IC}_{50}$  determined ( $\text{IC}_{50} = 2.6 \pm 0.1 \mu\text{M}$ ). Additionally, both the substrate and inhibitor concentrations were varied and the data fit to a Dixon plot to obtain the inhibition constant  $K_i = 1.4 \pm 0.3 \mu\text{M}$  for **4.33** versus TcTS (Table 4.8).

**Table 4.8. Inhibition values for linear and cyclic peptides versus TcTS.**

Linear Peptide	IC <sub>50</sub> (μM)	Cyclic Peptide	IC <sub>50</sub> (μM)	K <sub>i</sub> (μM)
<b>4.20</b>	242 ± 29	<b>4.33</b>	2.6 ± 0.1	1.4 ± 0.3
<b>4.21</b>	>500	<b>4.34</b>	4.3 ± 0.7*	-
<b>4.29</b>	>500	<b>4.35</b>	167 ± 16	-
<b>4.30</b>	>500	<b>4.36</b>	169 ± 15	-
<b>4.31</b>	>500	<b>4.37</b>	70 ± 11	-
<b>4.32</b>	>500	<b>4.38</b>	74 ± 8	-

\*The concentration of **4.34** stock concentration, and thus IC<sub>50</sub> value, was adjusted after kinetics were performed based on UV/Vis analysis (See Chapter 7.3.6 for details).

By comparing the measured IC<sub>50</sub> values in Table 4.8 a few trends can be immediately observed. First, in all cases the cyclic peptide is a better inhibitor than the linear precursor. This may indicate that secondary structural features are induced by macrocyclization and/or it may reflect the entropic cost of linear chain rearrangement to bind in the same fashion as their cyclic relative. Further, the only linear peptide with measurable inhibitory activity versus TcTS was **4.20** and even so the activity was two orders of magnitude worse than that of the cyclized derivative. Another observation is that the peptide hits from the RaPID screen **4.20** and **4.21** are better inhibitors than their respective ‘natural’ analogues. This observation confirms our initial hypothesis that increased aryl hydroxylation leads to better inhibitors of TcTS. Finally, the comparison of the inhibition afforded by both pairs of ‘natural’ analogues **4.35/4.36** and **4.37/4.38** suggests that neither L-Tyr nor L-Met – substituted in place of L-Cys\* – contribute significantly to enzyme inhibition.

## 4.7 Conclusions

In search of new inhibitor scaffolds for TcTS, peptide-based libraries were assessed by RaPID screening. Three unique cyclization scaffolds, L- and D- amino acid initiation and hydroxylated non-proteinogenic amino acids were included in the screened peptide libraries. None of the most prevalent sequences from the library of macrocyclized peptides that were derivatised via intramolecular cysteine-based methods were inhibitors of TcTS at concentrations up to 25  $\mu\text{M}$ . Second-generation libraries containing non-proteinogenic amino acids were cyclized with an exogenous stapling agent and yielded more exciting screening results. Synthesis of requisite amino acids followed by manual SPPS was employed to produce the most promising linear peptides. The cross-linking reaction with disuccinimidyl glutarate was optimized to yield the desired cyclic peptides, and kinetic analysis revealed low micromolar inhibition activity for these new peptide-based scaffolds. Meanwhile, ‘natural’ analogues of the lead cyclic peptides – bearing the closest proteinogenic amino acids in place of the non-proteinogenic residues – proved to be mid to high micromolar inhibitors of TcTS. Further, the linear versions of the peptides showed no significant inhibition at concentrations up to 200  $\mu\text{M}$ . Thus, the incorporation of non-proteinogenic amino acids into DSG-stapled cyclic peptides and subsequent screening using the RaPID system has elucidated two new peptide motifs that inhibit TcTS at the same level as the best-known compounds to date. Additional structural and kinetic studies with peptide derivatives will be necessary to reveal deeper structure-activity relationships.

## Chapter 5: Towards Enzyme Crystallography and Cell-Based Studies

### 5.1 Introduction

In the previous three chapters we identified new inhibitors of our enzyme target through both rational design and screening approaches. These methods fulfilled the first goal of our drug discovery and development process: finding new ‘lead’ compounds with promising *in vitro* inhibition of TcTS activity. Once these candidates were identified we shifted our focus towards two new goals (1) determining the interactions that these compounds form with TcTS and (2) assessing how TcTS inhibition affects *T. cruzi* in cell-based systems.

Structural assessments – such as enzyme crystallography and/or structure-activity relationship (SAR) studies – are typically employed to assess the manner in which the observed inhibition is exerted. This information is crucial as it illustrates the mode and site of interaction between the chemical lead and the enzyme target. Further, these studies can provide valuable information about which molecular interactions are essential for binding and conversely which parts of the compound can be altered. Then, iterative derivatization – focused on increasing potency, selectivity or improving pharmacological properties – can be achieved by chemical synthesis until the characteristics of the lead compound have been sufficiently enhanced. Additionally, virtual screening of related compounds could be investigated based on the obtained structural information.

In parallel, assessment of the lead candidates in cell-based assays can determine the functional activity of the lead molecules in biologically relevant environments. Successful *in vitro* enzyme inhibition does not guarantee comparable performance in cell-based systems, as many biological factors such as off-target effects, cell-permeability and compound stability may lead to complications. However, if the candidates do exhibit promising cellular activity, further trials in animal and then human systems (*i.e.* clinical trials) are conducted en route to commercial development of the therapeutic compound. If not, the above development processes can be refined and repeated, or tabled in search of alternative approaches.

## 5.2 Specific aims

Chapters 2, 3 and 4 described the orthogonal design and discovery approaches employed to obtain new inhibitors of *Trypanosoma cruzi* trans-sialidase. Each of the methods afforded near best-in-class inhibitors of TcTS yet much remains unknown regarding the structure-activity relationship for any of these compounds. Thus our first aim is to reproduce the crystallization protocols of the target enzyme and subsequently obtain complexed structures with our lead molecules. Furthermore, the biological impact of TcTS inhibition remains elusive. Therefore our second aim is to evaluate our most promising synthetic natural product analogues in cell-based assays using established *in vivo* protocols through collaboration with Dr. Buckner's group at the University of Washington.

## Results and Discussion

### 5.3 Crystallization of *T. cruzi* trans-sialidase

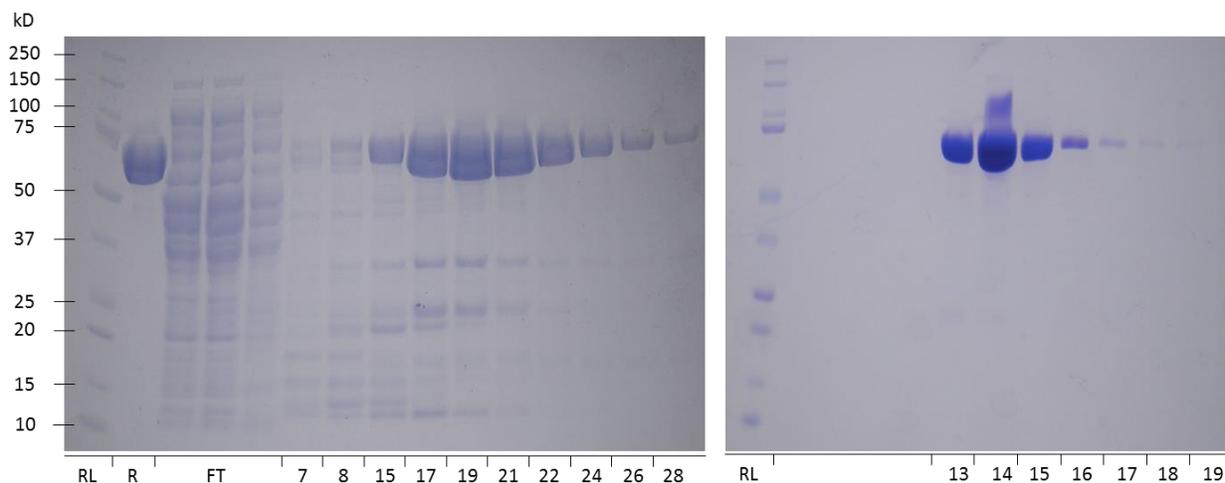
Enzyme crystallography is an invaluable tool in the drug development process. Structure-activity information, as it pertains to our synthetic DFSA libraries would provide insight on the orientation of, and the interactions formed by, the 5- and 9-amido functional groups. Further, co-crystallization with our lead compounds from natural product and peptide screening would illuminate the currently unknown binding site and interactions formed between the inhibitors and the enzyme. With such information in hand the determination of minimum inhibitory structures and the optimization of binding interactions could readily be explored.

Fourteen unique crystal structures of TcTS have been reported to the Protein Data Bank (PDB<sup>194</sup>; <http://www.rcsb.org/pdb>) from five articles published between 2002-2012, each of which were solved by Dr. Buschiazzo. The first reported structures were the monoclinic, orthorhombic and triclinic crystal systems<sup>40</sup>, followed by inhibitor- and substrate-bound co-crystals with TcTS or its catalytically inactive mutants<sup>41,101,195,196</sup> (D59A or Y342H). The initial report by Buschiazzo and colleagues describes the arduous, systematic approach used to determine that seven surface mutations and a two-step seeding/streaking protocol are required to obtain useable enzyme crystals. Specifically the reported procedure requires the growth of seed crystals in 'Buffer A' (2 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM HEPES, 2% PEG-400, pH = 8, 2 mM DANA) that are then fragmented and microseeds of these crystals are streaked into pre-equilibrated drops of TcTS in 'Buffer B' (100 mM Tris-HCl, 10% PEG-4000, pH = 7.5, 5% isopropanol, 2 mM DANA). The authors detailed that crystallization in the presence of millimolar amounts of Neu5Ac or DANA is necessary to obtain the monoclinic form of the enzyme. This is important because the triclinic and orthorhombic crystals were found to crack immediately upon soaking with substrates or inhibitors, leading the authors to conclude that a conformational change upon substrate binding disrupts the crystal lattice in these systems. This non-trivial crystallization protocol has eluded

any published replication to date. My attempts to crystallize TcTS following these protocols were also unsuccessful, yet efforts to this end and proposed future directions are discussed below.

### 5.3.1 Enzyme preparation

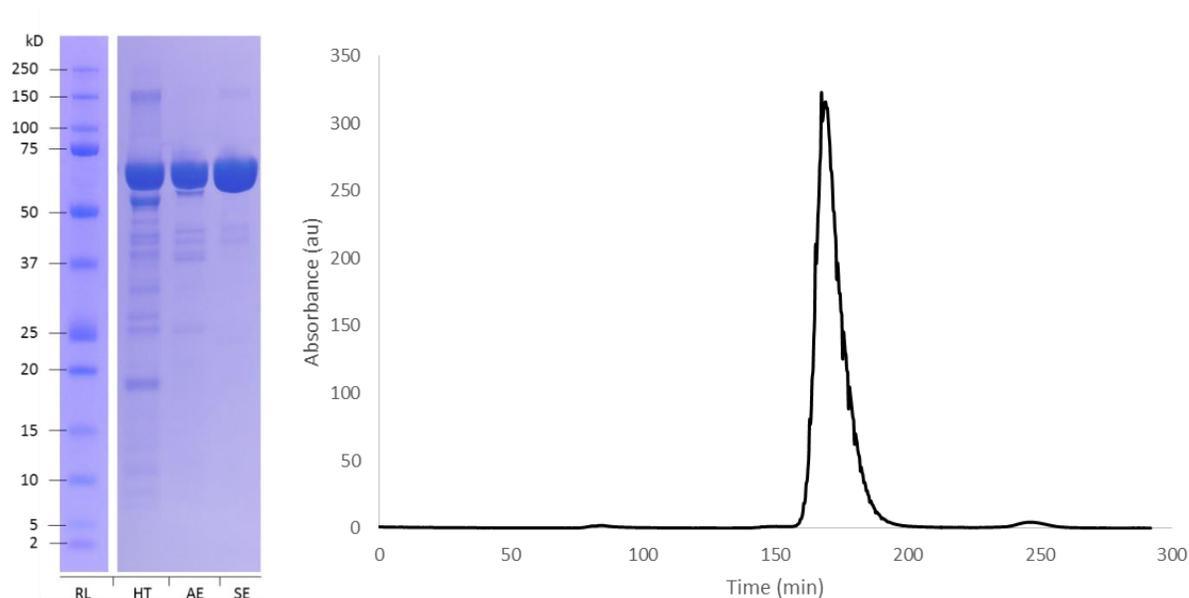
Recombinant TcTS, bearing the seven aforementioned surface mutations, was freshly prepared (within 1-2 weeks) prior to each round of crystallization trials. TcTS was expressed in BL21(DE3) competent *E. coli* cells (see Chapter 7.1.7 for details) and initially purified by sequential HisTrap (nickel-charged immobilized metal affinity) and MonoQ (anion exchange) chromatography (Figure 5.1).



**Figure 5.1. SDS-PAGE of HisTrap (left) and MonoQ (right) column purification following *T. cruzi* trans-sialidase expression.** RL: reference ladder; R: TcTS standard; FT: column flow-through; numbers indicate fractions from column purification.

Following some unsuccessful attempts to replicate the seed crystals described in the literature (discussed below) we examined the enzyme stock by SDS-PAGE. Analysis of the combined fractions (after anion exchange chromatography) indicated that a small contaminating band was present directly beneath the TcTS spot (Figure 5.2; AE). Thus, an additional purification step was employed: size exclusion chromatography. Superdex 200 resin (GE) was loaded with TcTS and eluted over 5 hours (Figure 5.2). Subsequent SDS-PAGE indicated the successful removal of the previous impurity (Figure

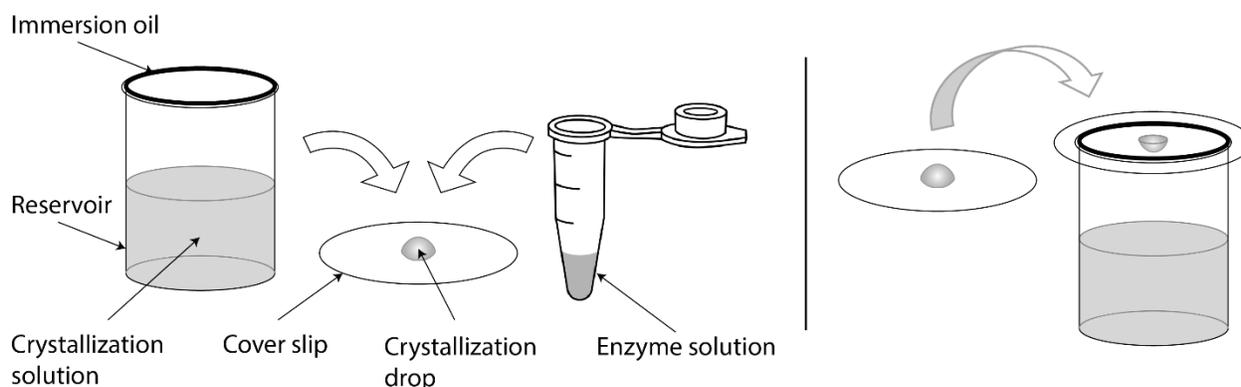
5.2; SE). TcTS was exchanged into storage buffer (20 mM Tris, 30 mM NaCl, pH = 7.5) and concentrated to 10-35 mg/mL prior to storage at 4 °C.



**Figure 5.2. SDS-PAGE of combined TcTS-containing fractions from each column purification (left). UV/Vis spectrum of the elution profile from SEC purification (right).** HT: HisTrap; AE: anion exchange; SE: size exclusion; RL: reference ladder. Reference ladder was developed on the same SDS-PAGE but repositioned for viewing clarity.

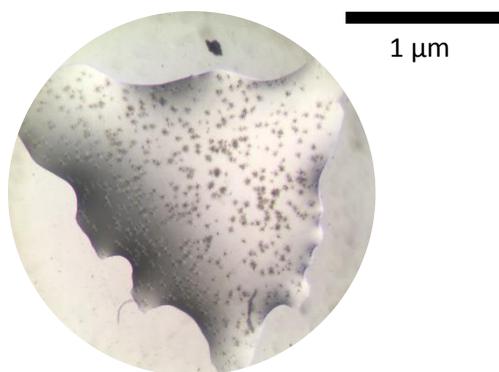
### 5.3.2 Growth of seed crystals

Crystallization conditions were tested using the hanging drop vapour diffusion technique (Figure 5.3). In this method, crystallization drops – made of mixed crystallization buffer and enzyme solution – are suspended on a glass cover slip above a reservoir of the crystallization buffer. Drops are left to incubate for hours to days as the water vapour diffuses between the drop and reservoir until equilibrium occurs. The slow change in the drop solute concentration facilitates the gradual formation of protein crystals. The use of a Carl Zeiss JENA Technical 2 stereo microscope enabled the visual inspection of each hanging drop.



**Figure 5.3. Hanging drop vapour diffusion crystallization preparation (left) and incubation (right).**

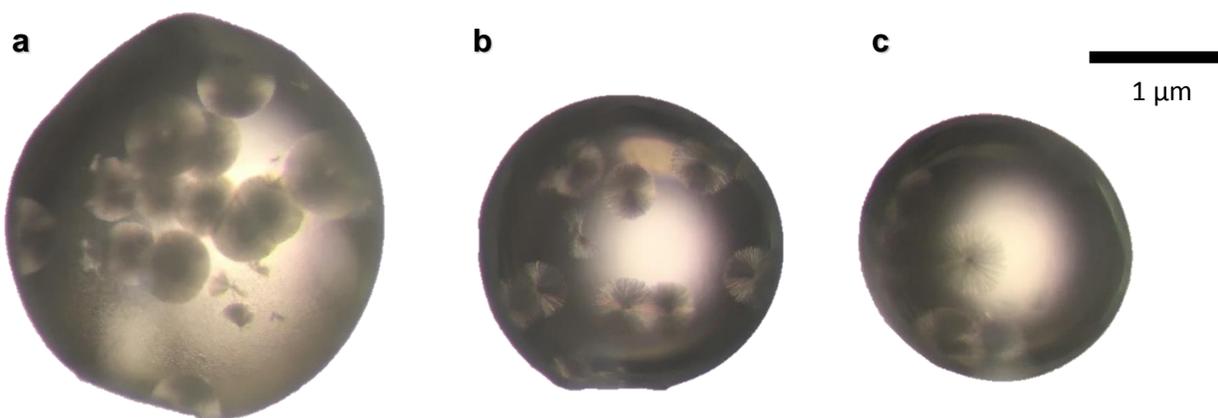
The first step of our process was to investigate the reproducibility of the literature conditions by systematically varying the concentration of the Buffer A components and assessing the effects on seed crystal formation. Each component was varied from 0.5- to 2-fold of the reported concentrations: salt solution (ammonium sulfate;  $(\text{NH}_4)_2\text{SO}_4$ ) from 1.5-2.5 M, polyethylene glycol 400 (PEG-400) from 1-4% and buffer (HEPES, pH = 8) from 50-200 mM. For the initial trial, TcTS was prepared following only HisTrap and MonoQ column purification (15 mg/mL), incubated with DANA (2 mM) and diluted 1:1 in crystallization buffer. A few observations were immediately evident:  $\geq 2.2$  M ammonium sulfate afforded immediate precipitation of TcTS, while  $\leq 1.8$  M precipitant failed to yield solid after 1-7 days. The replicated literature conditions (and the conditions with 4% PEG-400) afforded small needle-like clusters (Figure 5.4).



**Figure 5.4. Light microscopy image of TcTS crystals (1:1 TcTS (15 mg/mL)/Buffer A).**

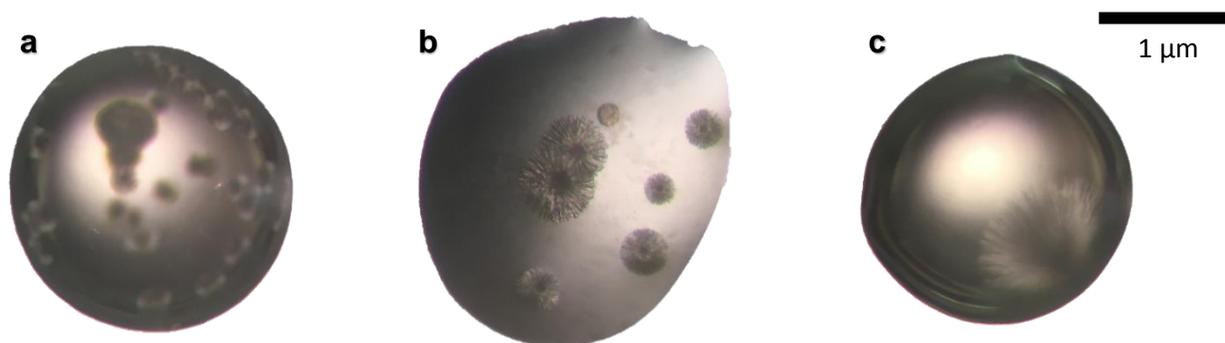
The effect of PEG-400 was further assessed by varying the concentration from 1-4% in the presence of two different enzyme/buffer ratios (1.5:1; 2:1). Similar, albeit larger, clusters of needle-like crystals were observed at both 2% and 4% PEG-400, so the literature ‘Buffer A’ conditions were employed for the remaining trials. The needle-like shape of the crystals contradict the plate-shaped crystals reported by Buschiazzo but nonetheless provided a starting point for further trials.

The next parameters investigated were the stock enzyme concentration and the dilution ratio with crystallization buffer. Additionally, size exclusion purification was added to all enzyme preparation protocols at this point. TcTS stock concentrations were varied from 12-35 mg/mL and the ratio of enzyme/buffer per drop investigated (2:1-1:2). Hanging drops of the selected conditions were employed and the drops assessed by visual inspection (Figure 5.5). Based on the quantity, shape and thickness of the resulting crystals some conclusions were drawn. First, 35 mg/mL was too concentrated and led to many small disc-shaped clusters of thin needles which grew within an hour of incubation. Second, 20 mg/mL drops generally yielded distinct clusters of needles which displayed a range of thicknesses depending on dilution ratios: 2  $\mu$ L of enzyme stock to 1  $\mu$ L of buffer yielded thicker needles, while 1  $\mu$ L enzyme and 2  $\mu$ L of buffer afforded thinner needles (Fig. 5.5).



**Figure 5.5. Light microscopy images of TcTS crystals.** (a) 2:1 TcTS (35 mg/mL)/Buffer A (b) 2:1 TcTS (20 mg/mL)/Buffer A (c) 1:2 TcTS (20 mg/mL)/Buffer A.

Initial findings indicated that larger crystalline needles resulted from enzyme stock concentrations of ~20 mg/mL and high initial ratios of enzyme to precipitant. Another set of crystallization conditions was explored – using finer gradients (3:0.5-1:2) – to find the best ratio of enzyme solution to precipitate for the 20 mg/mL stock. Similar trends were observed as before (Figure 5.6), and the optimal conditions to obtain clusters of thick needle like crystals were determined to be 20 mg/mL TcTS in a 2:0.5  $\mu$ L ratio in the reported Buffer A conditions after a 24-48 hour growth period.



**Figure 5.6. Light microscopy images of TcTS crystals.** (a) 3:0.5 TcTS (20 mg/mL)/Buffer A (b) 2:0.5 TcTS (20 mg/mL)/Buffer A (c) 1:2 TcTS (20 mg/mL)/Buffer A.

### 5.3.3 Microseeding attempts

In parallel, the conditions for Buffer B were similarly assessed. Following each round of seed crystal formation described above, the most promising crystals were either rinsed or directly fragmented for microseeding. Using a cat whisker, the resultant TcTS seeds were streaked into saturated solutions of enzyme that had been pre-incubated ‘Buffer B’ (1:1 ratio), yet these drops consistently failed to afford any solids after 1-14 days. Once the Buffer A solution had been optimized, crystals grown under these conditions were used to assess Buffer B components in a more systematic fashion. Isopropyl alcohol was added in 0-5% amounts, PEG-4000 concentration was varied from 8-12% and even substituted for PEG-3350 which was analogously evaluated. The pH was similarly varied (pH = 7, 7.5, 8) and the pre-incubation time before streaking assessed between 0-4 hours. None of the adjustments to the components or protocols made any discernible difference since all drops remained clear after weeks of incubation.

Considering the important role that stock enzyme concentration played on the crystals obtained from Buffer A, we similarly investigated its impact in Buffer B. However, the only conclusion we could draw was that at concentrations >25 mg/mL TcTS first oils out and then slowly precipitates upon mixing with crystallization buffer (1:1), therefore indicating an upper limit for enzyme concentration. Additional factors such as concentration of seeds used for streaking, method of streaking (micropipette vs. whisker) and ratio of enzyme to buffer were evaluated to no avail. Finally, attempts to crystallize TcTS directly in Buffer B – as was first reported by Buschiazzo in 2002 – were also unsuccessful. Difficulties with the microseeding and streaking method spurred a small screening attempt for new crystallization conditions.

#### **5.3.4 Screening for new conditions**

Considering that the established crystallization conditions for TcTS were developed more than 15 years ago and that enzyme crystallization screening has been significantly developed over that time, an initial screening attempt was conducted in search of new crystallization conditions. A 96-well two-drop crystallization plate (Swissci) was filled with 100  $\mu$ L/well of Hampton Crystal Screen HT solution – which was kindly provided by Dr. Katherine Ryan. TcTS was pre-incubated with DANA (2 mM) and sitting drops of the enzyme solution (20 mg/mL and 10 mg/mL stocks) were assessed in a 1:1 enzyme/buffer ratio. The plates were sealed with clear tape and the solutions were stored at room temperature for 14 days. Visual inspection of each well was performed after 1, 2, 5 and 14 days of incubation using a light microscope (Nikon SMZ1000 Zoom Stereo Microscope). After two weeks of incubation, two wells (G6 and H1) afforded small crystalline solids as well as some initial proteinaceous precipitate. The corresponding buffer compositions of G6 (100 mM HEPES, pH = 7.5, 10% PEG-6000, 5% 2-methyl-2,4-pentanediol) and H1 (100 mM HEPES, pH = 7.5, 10% PEG-8000, 8% ethylene glycol) are similar to each other and to elements of the Buffer A and Buffer B solutions used in the reported TcTS crystallization procedures (Table 5.1). Given these similarities the use of a different buffer (*e.g.* HEPES in place of Tris-HCl), larger PEG derivatives (*e.g.* PEG-6000 or PEG-8000) or alternative precipitant

additives (e.g. 2-methyl-2,4-pentanediol or ethylene glycol) in Buffer B may provide a promising starting point for future crystallization attempts – either directly with TcTS solution or through microseeding techniques with enzyme crystals from the optimized Buffer A conditions. Due to time constraints no further developments on this front have been explored.

**Table 5.1. Comparison of TcTS crystallization buffer components.**

<b>Components</b>	<b>Buffer A</b>	<b>Buffer B</b>	<b>G6 Solution</b>	<b>H1 Solution</b>
<b>Buffer</b>	100 mM HEPES	100 mM Tris-HCl	100 mM HEPES	100 mM HEPES
<b>pH</b>	8	7.5	7.5	7.5
<b>PEG</b>	2% PEG-400	10% PEG-4000	10% PEG-6000	10% PEG-8000
<b>Salt/precipitant</b>	2 M (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	5% isopropanol	5% 2-methyl-2,4-pentanediol	8% ethylene glycol

#### 5.4 Cell-based studies

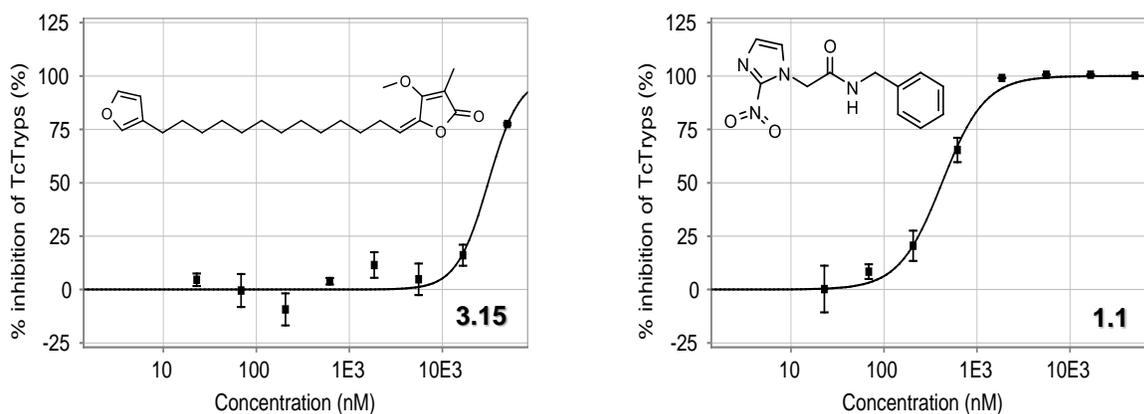
The biological experiments typically employed to assess potential Chagas disease therapeutics are based on the ability of a chemical lead to kill or affect growth of *Trypanosoma cruzi*. In this way the activity of new trypanocidal species can be compared directly to that of the antiparasitic drug benznidazole. In cell viability assays parasites are incubated in media with varying concentration of the potential therapeutic agent for ~24 hours prior to counting the number of living cells<sup>197</sup> (e.g. resazurin assay). A positive response in such an assay indicates an ability of the chemical compound to kill the trypanosome (i.e. trypanocidal activity). Whereas growth inhibition assays require the incubation of *T. cruzi* – in the infectious trypomastigote form – with mammalian cells and the chemical lead for several days prior to determining the number of parasites present at the end of the experiment (via cell counting or UV/Vis methods). Decreased growth relative to controls over the time of the experiment can either be attributed to antitrypanosomal activity of the agent or its ability to disturb the parasite’s reproduction cycle (since *T. cruzi* requires access to the cellular machinery of the mammalian cells for effective

reproduction). In practice both assays are commonly used to quantify (in the form of an EC<sub>50</sub> value) the trypanocidal activity of a chemical compound. While many agents that are anti-parasitic towards *T. cruzi* have been reported, most suffer from one or all of the following significant shortcomings: lack of potency relative to benznidazole in lab or clinical settings, poor selectivity for trypanosomes and/or an unknown mechanism of action. As such the door remains open for new classes of Chagas therapeutics to be discovered.

Having amassed a collection of new inhibitors with low micromolar inhibition constants versus recombinant TcTS we aimed to use an *in vivo* system to assess the biological effects of enzyme inhibition on the parasite species *Trypanosoma cruzi*. The hypothesized outcomes of TcTS inhibition are two-fold: (1) cellular invasion by non-sialylated parasite would be decreased and (2) clearance of the parasite by host innate immune response would be increased. We believe that in cell-based systems the assessment of cellular invasion should be more accessible than attempting to simulate the complex mammalian immune response system. In 1996 Buckner et al. described a drug screening method that used *T. cruzi* cells which were genetically modified to express the  $\beta$ -galactosidase gene from *E. coli*<sup>198</sup>. The transfected enzyme activity enabled the quantitation of the parasite cells by monitoring the enzymatic cleavage and resultant colorimetric response of chlorophenyl red  $\beta$ -D-galactopyranoside (CPRG). In this method parasite cells are incubated with murine fibroblast cells and varying concentrations of potential drug species for 6 days. By employing the galactosidase substrate to quantify the end-point parasite levels after this time the concentration-dependent effects of the drug species on parasite growth can be easily assessed. The derived EC<sub>50</sub> curve represents the ability of the compound of interest to disrupt growth of *T. cruzi* in any manner possible, including but not limited to intra or extracellular trypanocidal activity, inhibition of parasite reproduction or prevention of parasite invasion. This method has become a routine assessment tool for antitrypanosomal drug candidates and has been further advanced by adaptation to HTS format<sup>199</sup>. Therefore, we hypothesized that *in vivo* inhibition of TcTS would slow parasite surface sialylation, reduce cell invasion and lead to slower *T. cruzi* growth rates.

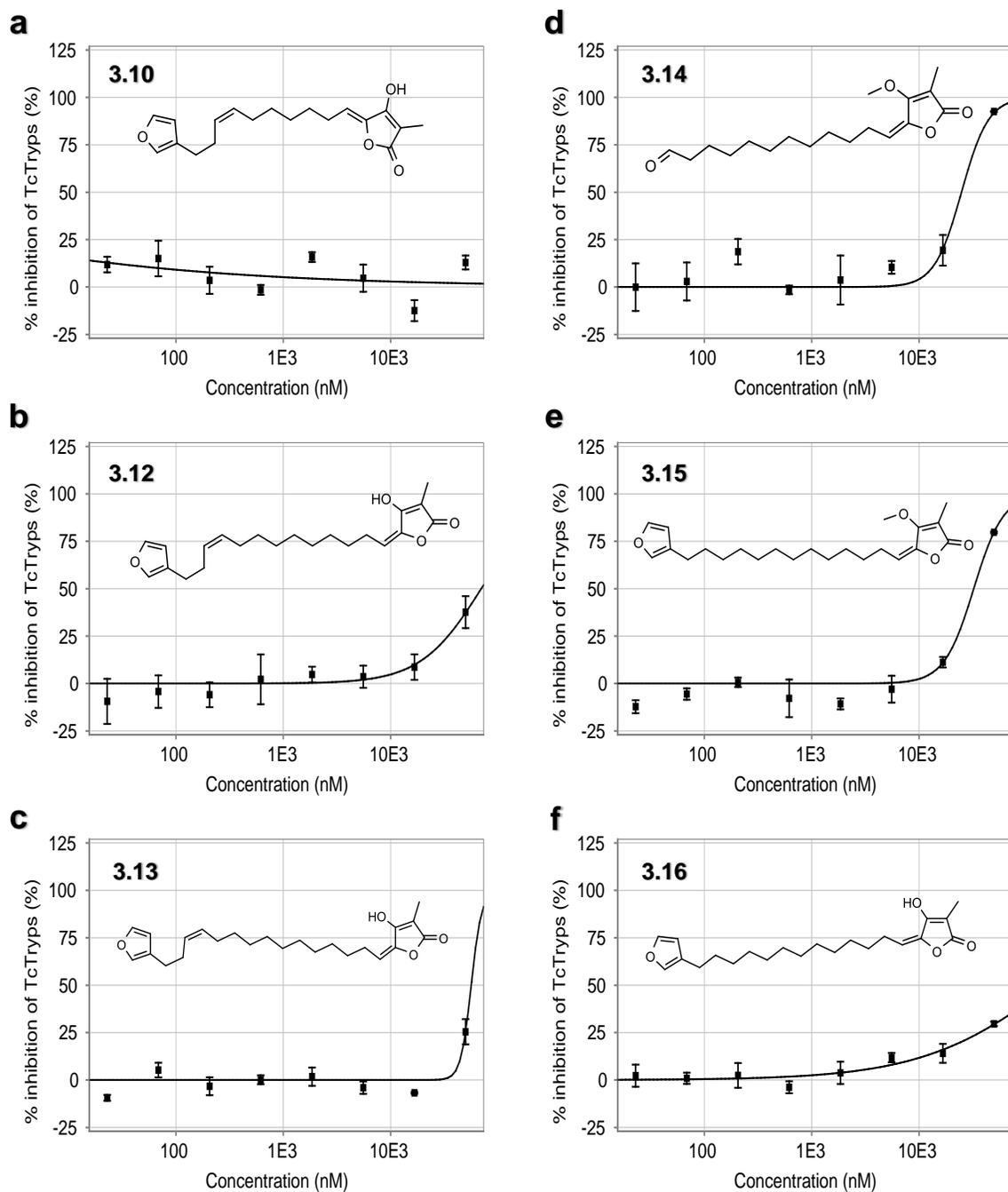
#### 5.4.1 *T. cruzi* growth inhibition assay

In collaboration with the Buckner lab, we selected six of the synthetic natural product analogues from Chapter 3 (**3.10**, **3.12-3.16**) to be assessed using their standard protocol<sup>200</sup> (See Chapter 7.4.1 for details). Specifically, this method begins with the addition of *T. cruzi* trypomastigotes ( $\sim 1 \times 10^4$  cells) to a monolayer of murine 3T3 fibroblast cells ( $\sim 2 \times 10^3$  cells) which are incubated together for 24 hours. Next, a series of drug concentration are added to each culture (triplicate serial dilutions) and the mixture is incubated for 6 days. Lysis of cellular material and colorometric response from  $\beta$ -galactosidase cleavage of the CPRG substrate affords the end-point parasite levels. Three control conditions were also assayed: (1) background (3T3 cells, trypomastigotes, 50  $\mu$ M benznidazole), (2) high control (3T3 cells, trypomastigotes), (3) viability control (3T3 cells, 50  $\mu$ M test compounds). In the presence of the ircinialactam analogues parasite growth was only impacted at the highest inhibitor concentrations tested (50  $\mu$ M) for select species, while most compounds failed to disrupt parasite growth at concentrations up to 50  $\mu$ M. The antitrypanosomal drug benznidazole was also measured and elicited potent assay activity  $EC_{50} = 420$  nM (Figure 5.7). We found the assay results to be surprising since there does not appear to be a correlation between the *in vitro*  $IC_{50}$  values and the *in vivo*  $EC_{50}$  results for the ircinialactam analogues (Table 5.2). The compounds with the most potent  $IC_{50}$  values (**3.10** and **3.13**) were among the worst of the compounds tested *in vivo* ( $EC_{50} = >50$   $\mu$ M). On the other hand, the most potent compounds *in vivo* **3.14** ( $EC_{50} = 48.1$   $\mu$ M) and **3.15** ( $EC_{50} = 31.1$   $\mu$ M) have mid-range  $IC_{50}$  values *in vitro* relative to the other tested products. Further, control assays did not show evidence of disruption to 3T3 cell viability by any of the test compounds.



**Figure 5.7. Percent inhibition of *T. cruzi* growth and inhibitor structure for 3.15 (left) and benznidazole (right).**

We feared that allowing uninhibited trans-sialylation of the parasite surface to occur during the 24 hour pre-incubation period between the parasite and the fibroblast cells prior to irlincinalactam addition may have adversely affected the experimental results. Therefore, a second assay was performed where the standard Buckner lab protocol was adjusted such that the 24 hour pre-incubation period was eliminated. Instead, the test compounds were first added to fibroblast-containing wells followed immediately by the parasite cells. The mixture was again incubated for six days, the parasite levels were assessed and the  $EC_{50}$  values determined (Figure 5.8).



**Figure 5.8.** Percent inhibition of *T. cruzi* growth and inhibitor structures for (a) 3.10 (b) 3.12 (c) 3.13 (d) 3.14 (e) 3.15 (f) 3.16. Assay conditions were adjusted to eliminate pre-incubation period of parasite and fibroblast cells.

However, the altered protocol minimally affected the experimental results as the same two ircinialactam analogues (3.14 and 3.15) were the only compounds that decreased parasite growth. In the

case of **3.14** the measured EC<sub>50</sub> value was decreased by ~50%, while the EC<sub>50</sub> remained unchanged for **3.15** (Table 5.2). Unfortunately, the most potent ircinialactam inhibitor of *T. cruzi* growth (EC<sub>50</sub> = 24.9 μM) is still >40-fold worse than the benznidazole standard.

**Table 5.2. Comparison of *in vitro* (IC<sub>50</sub>) and *in vivo* (EC<sub>50</sub>) values for ircinialactam analogues.**

Entry	Name	EC <sub>50</sub> <sup>a</sup> (μM)	EC <sub>50</sub> <sup>b</sup> (μM)	IC <sub>50</sub> (μM)
1	Benznidazole ( <b>1.1</b> )	0.42	0.57	-
2	<b>3.10</b>	>50	>50	1.8 ± 0.1
3	<b>3.12</b>	>50	>50	2.0 ± 0.1
4	<b>3.13</b>	>50	>50	1.8 ± 0.1
5	<b>3.14</b>	48.1	24.9	3.1 ± 0.1
6	<b>3.15</b>	31.1	32.4	2.1 ± 0.1
7	<b>3.16</b>	>50	>50	3.1 ± 0.1

<sup>a</sup> with pre-incubation period of *T. cruzi* and 3T3 cells.

<sup>b</sup> without pre-incubation period of *T. cruzi* and 3T3 cells.

Due to inconsistent *in vivo* and *in vitro* results and a lack of evidence indicating how **3.14** and **3.15** elicit their observed effects, conclusions towards our initial hypothesis remain elusive. We can say that two ircinialactam analogues with known IC<sub>50</sub> <5 μM disrupted *T. cruzi* growth with EC<sub>50</sub> values determined to be <50 μM. We can speculate that the decreased parasite levels observed in this assay are caused by inhibition of TcTS mediated trans-sialylation of the parasite cell surface and subsequent interference with cellular invasion. However, it is equally valid that four ircinialactam analogues with known IC<sub>50</sub> <5 μM failed to exhibit any effects on the observed parasite levels under these assay conditions. One structural feature – *O*-methylation of the tetronic acid moiety – is shared between the two active molecules and is absent in the other four species tested. It is possible that the anionic character of the tetronic acid bearing compound may differentially influence the biological activity relative to the

neutral, *O*-methylated species. Comparison of the cellular activity between **3.15** ( $EC_{50} = 32.4 \mu\text{M}$ ) and **3.16** ( $EC_{50} > 50 \mu\text{M}$ ) supports this theory since *O*-methylation is the only distinguishing feature between the otherwise structurally identical compounds. Therefore developing a better understanding about the role of tetronic acid methylation in the cellular system may be the key to deciphering the above results.

One potential explanation is that ircinialactams **3.14** and **3.15** both exhibit non-productive TcTS inhibition and the observed trypanocidal activity occurs through an alternative mode of action. One assay which could be used to test this hypothesis is a parasite viability assay. In this case, *T. cruzi* is incubated overnight in the presence of varying concentrations of ircinialactam and the cell viability determined through standard means (*e.g.* resazurin assay). It could be useful to perform this assay on both the epimastigote and trypomastigote forms of the parasite since only the trypomastigote displays TcTS on the cell surface. Results from this comparison could provide information regarding the trypanocidal activity of these compounds and the relationship of this activity to interactions with the parasite trans-sialidase.

Additionally, the appropriateness of the selected assay may be critiqued. The parasite growth assay employed by the Buckner lab is most often used to determine the assumed trypanocidal activity of potential drug candidates that act in a similar manner to benznidazole. Unlike those compounds, the hypothesized mechanism of action for our inhibitors does not lead directly to parasite death. Instead the inhibition of parasite surface TcTS should slow/prevent surface sialylation and cellular invasion. In more complex systems this would in turn allow for more rapid clearance (*i.e.* trypanocidal activity) to be carried out by host immune cells. In order to confirm our hypothesis in the growth inhibition assay, we are hoping that slowing parasite cell invasion will affect the final parasite concentration after six days of growth. However it is not clear if these experimental conditions are on the appropriate time-scale to observe the hypothesized effects. By measuring the number of parasites only at the end of the six day experiment, any potential effects on the parasite growth curve caused by the ircinialactams before the end-point are not detected. Therefore the development of new assay protocols is required to more rigorously assess the effects of TcTS inhibition on parasite invasivity. Moreover, animal models will

likely be required to appropriately assess the effect of TcTS inhibition on *T. cruzi* clearance rates by immune response systems.

## 5.5 Conclusions

Following the development and kinetic characterization of new carbohydrate-, natural product- and peptide-based inhibitors of TcTS using *in vitro* studies we aimed to further assess the mode of inhibition and biological relevance of these compounds. To better understand the interactions between the inhibitors and the enzyme target, structural studies were attempted. Systematic optimization of the protocols described by Buschiazzo led to modified enzyme purification and crystallization conditions that afforded thick needle-based crystal clusters. Unfortunately, attempts to replicate the streaking of TcTS seed crystals into the literature growth conditions failed to generate any usable crystals. Two new crystallization buffers were identified from the 96-well screen of the Hampton Crystal Screen HT solutions. These conditions afforded small TcTS crystals after two weeks and can possibly be optimized to obtain useable TcTS crystals for structural analysis and subsequent soaking experiments with inhibitor species.

To assess the biological relevance of TcTS inhibition we collaborated with the Buckner lab at the University of Washington to test six synthetic ircinialactam inhibitors in their parasite growth inhibition assay. Results of the two assays – conducted with and without parasite/3T3 cell pre-incubation – indicated that two of the six compounds have EC<sub>50</sub> values <50 μM while four of the compounds did not affect parasite growth up to 50 μM. The two ‘active’ ircinialactams share a structural similarity – *O*-methylation of the conserved tetronic acid moiety – which may play a central role in the inhibition of *T. cruzi* growth and warrants additional investigation. Thus conclusions regarding the effect of TcTS inhibition on parasite invasivity and the therapeutic relevancy of TcTS remain unclear.

## Chapter 6: Concluding Remarks

*Trypanosoma cruzi* trans-sialidase catalyzes the transfer of terminal sialic acid residues from human sialoglycoconjugates to the mucin-like coating on the parasite surface. Through this action, *T. cruzi* is effectively masked from host immune recognition, thus allowing the parasite to navigate the bloodstream incognito while propagating its infectious life cycle. Further, parasite surface sialylation and the associated desialylation of the host glycans is believed to enhance parasite invasion – an essential step in *T. cruzi* replication. Eliminating the enzymatic trans-sialidase activity would prevent the transfer of sialic acid from host to parasite surface. Consequently, parasite invasivity would be hindered leaving the vulnerable ‘unmasked’ *T. cruzi* to be cleared by the host’s immune system. As such, TcTS has garnered interest as a potential therapeutic target for small molecule inhibitor development. To the benefit of researchers in this field, TcTS structure, activity and mechanism have been thoroughly investigated. These studies have highlighted many of the essential amino acids in the enzyme active site that form binding interactions with the sialyl donor (R45, R245, R314, D96) and the lactosyl acceptor (Y119, W312) during catalysis. Moreover the two-step double-displacement mechanism has been solved and the catalytic residues (Y342, D59, E230) responsible for enzymatic activity identified. This level of insight is nearly ideal for the rational design of substrate-based inhibitors.

Despite the bevy of available structural information the design of carbohydrate-based inhibitors for TcTS has been largely unsuccessful, failing to eclipse inhibition constants  $<100 \mu\text{M}$ . Simplified sialoside mimics bearing anionic charge and/or *N*-acetyl moieties (*e.g.* pyridine, benzoic acid and cyclohexene phosphonate derivatives) were designed to target the key interactions in the sialic acid binding pocket; however  $K_i$  values for these species were in the millimolar range. Likewise, lactoside-based analogues were designed to interact with acceptor site amino acids, yet exhibited similarly high inhibition constants. The most successful synthetic mimic was designed to target amino acids in both the donor and acceptor sites. This was achieved using the *C*-glycoside version of a sialic acid bearing a benzyl aglycone (**1.16**). Even so, the binding constant was  $K_d = 120 \mu\text{M}$ , a far-cry from the nanomolar

range typically required for therapeutic relevance. The difficulties encountered with carbohydrate-based inhibitors for TcTS are epitomized by the transition-state analogue DANA. This prototypical sialidase inhibitor has an inhibition constant  $>10$  mM for TcTS which is in stark contrast to the low to sub micromolar inhibition typically exerted by DANA against bacterial and viral sialidases.

Mechanism-based covalent inactivators were first used to identify the catalytic nucleophile of TcTS and have further been explored as an inhibitor scaffold. Synthetically modified versions of difluorosialic acid (DFSA) have demonstrated exceptional potency and selectivity versus influenza neuraminidase (vNEU); yet, like DANA the potency of the parent DFSA is orders of magnitude (25,000-fold) higher for vNEU compared to TcTS. To close this gap researchers have sought after an appropriate modification to the DFSA backbone that can be leveraged to improve TcTS binding – in an analogous manner to the C4 guanidine moiety found for vNEU. The covalent nature of these inhibitors potentially affords a higher tolerance with regard to inhibitor potency (relative to reversible species) since the ‘permanent’ inhibition accumulates over time. In this case a greater emphasis on target selectivity may be considered. To this end an appropriate chemical modification (*i.e.* selectivity factor) to the inactivator scaffold must again be determined to incur the desired selectivity for TcTS.

Previous kinetic studies had evaluated C9-modified DFSA derivatives as inactivators for TcTS. They determined that TcTS tolerates modification of the glycerol chain and this position was leveraged to increase inhibitor potency, slow turnover of the covalent intermediate and to confer selectivity versus human neuraminidases. We used the insights gained from previous trials to design a series of 9-amido linked DFSA inactivators and determined their kinetic parameters versus TcTS. The most potent of which – the C9 amide-linked cyclopentyl DFSA (**2.49**) – afforded a 10-fold increase in second order rate constant ( $k_i/K_i$ ) versus the parent DFSA (**1.21**). However attempts to further enhance binding by modifying five-membered ring (*e.g.* aromaticity, heteroatoms or elongation) led to deleterious effects on enzyme inhibition. Yet the addition of a negatively charged functional group to the cyclopentyl ring – in the form of a carboxylate moiety – led to 3.5-fold improvement over unsubstituted DFSA. A second

series of *N*-modified DFSA derivatives was similarly synthesized bearing *N*-acyl groups with linear, branched and cyclic alkyl chain functionality. The rate of TcTS inactivation at 5 mM was measured for each species and the corresponding half-lives ( $t_{1/2}$ ) determined. All seven species demonstrated comparable inactivation to each other and the *N*-acetyl DFSA control. While improved binding affinity was not observed, we learned that TcTS is tolerant to substitution at this position. We were able to demonstrate that TcTS inactivation can be increased by C9 modification DFSAs while *N*-acyl functionalization is tolerated in the TcTS active site. Future research efforts should therefore focus on leveraging this information to achieve selectivity against human neuraminidases and continue to search for functional groups that can further enhance DFSA potency towards TcTS.

Towards the development of selective inhibitors, the synthesis of C9 and C5 disubstituted DFSA derivatives should be explored. The Withers lab has amassed much information about TcTS binding preferences through the use of modified DFSA analogues. Meanwhile the Cairo lab has synthesized libraries of C4, C5 and C9 modified DANA analogues and have determined that each human neuraminidase isoform has its own distinct binding preferences. By comparing the results from each lab, a targeted library of disubstituted DFSAs can be designed as promising candidates for future specificity studies. Alternatively, new C9 linkers (*e.g.* sulfonamides or phosphoramidates) or a more diverse set of functional groups could be appended to DFSA in search of binding interactions that have yet to be discovered. Ideally these synthetic efforts could be guided by structural information obtained from co-crystallization of existing DFSA derivatives to reveal important binding interactions which may be responsible for the improved inactivation measured for select C9 modified DFSAs (**2.49**, **2.56**). Alternatively, efforts to improve binding affinity for 3-fluorosialosides could be explored by surveying a set of larger activated aglycones – in place of the anomeric fluoride – that could simultaneously harness interactions in the lactose binding pocket. This strategy is reminiscent of the one employed to successfully afford the most potent reversible carbohydrate-based inhibitors of TcTS.

The lack of potent inhibitors obtained from decades of carbohydrate-based design efforts has pushed some researchers to explore screening-based approaches in search of new chemical inhibitor scaffolds. Virtual screening of commercial chemical libraries and subsequent kinetic characterization revealed a series of inhibitors with  $IC_{50}$  values in the mid-micromolar range. Separately, screening a commercial natural product library afforded two hydroxylated inhibitor scaffolds – flavanols and anthraquinones – which had low- and sub-micromolar inhibition constants versus TcTS. Similarly, hydroxylated sulfonamide chalcones were found to competitively inhibit TcTS in the same range. These findings revealed that non-carbohydrate based TcTS inhibitors were by 2-3 orders of magnitude more potent than the best designed carbohydrate-based species. However, concerns regarding the known cytotoxicity of flavonoids hampered further therapeutic studies on these particular derivatives. We therefore pursued two collaborative screening approaches in order to uncover new inhibitors of TcTS from natural product or peptide-based libraries.

Screening the UBC Marine Natural Product Extract Library and subsequent bioactivity guided isolation (with Dr. Raymond Andersen and Dr. David Williams) afforded two lead structures **3.2** and **3.4** with  $IC_{50}$  values  $<25 \mu\text{M}$ . Structure-activity relationships (SAR) were investigated on a set of synthetic ircinialactam-like species related to **3.4** and handful of recently isolated cacolide natural products (both supplied by Dr. Robert Capon). These studies revealed a general motif – two functional groups (furan, lactam, dicarboxylate, tetrionic acid or butenolide) separated by unsaturated linkers of 10-15 carbons – that is shared by the most potent structures. Kinetic studies found that the best eight derivatives inhibited TcTS in a non-competitive manner with  $K_i$  values between  $0.8 \mu\text{M}$  and  $5 \mu\text{M}$ . These findings highlight a new natural product-based scaffold that can be differentially functionalized to achieve near best-in-class inhibition of TcTS. Dedicated synthetic efforts to this end could surely optimize this relationship for improved inhibitor potency.

The Suga lab (specifically Dr. Seino Jongkees) employed their Random nonstandard Peptide Integrated Discovery (RaPID) method to identify macrocyclic-peptides candidates for TcTS inhibition.

The two most promising hits **4.33** ( $K_i = 1.4 \mu\text{M}$ ) and **4.34** ( $\text{IC}_{50} = 4.3 \mu\text{M}$ ) inhibited TcTS hydrolysis with a similar potency to the best natural product-based species. These peptidaceous inhibitors bear many hydroxylated non-proteinogenic amino acids, which are comparable to the structural features employed by literature natural product inhibitors of similar potency. Initial SAR studies revealed that substitution of the hydroxylated non-proteinogenic amino acids for their ‘natural’ derivatives reduced TcTS inhibition by at least 10-fold. Further, the enzyme inhibition for the linear analogues of **4.33** and **4.34** were >100-fold worse than their cyclic relatives. Therefore we concluded that both the molecular interactions formed by the hydroxylated non-proteinogenic amino acids and the secondary structural features of the cyclic peptides are essential for the observed TcTS inhibition. These results add to the growing body of non-carbohydrate based species which exhibiting low micromolar inhibition of TcTS.

Priorities in TcTS research have shifted now that a collection adequately potent inhibitors of TcTS have been discovered. Specifically, two important questions come to mind: (1) how do these non-carbohydrate based inhibitors interact with TcTS and (2) how does TcTS inhibition actually affect *T. cruzi* in biological environments. To this end structural studies and cell-based investigations must be prioritized. Previous structural insights have highlighted the important active site residues for the binding of natural sialosyl or lactosyl based compounds, yet no such information has been attained with regards to the more potent non-carbohydrate based inhibitors. The non-competitive inhibition of many of these species and the synthetically arduous nature of SAR studies makes co-crystallization with TcTS a top priority – both to understand the mode/site of inhibition and to inform future inhibitor development. Our attempts to replicate TcTS crystallization did not afford useable crystals; however, the reported progress in seed crystal optimization and from screening efforts towards new crystallization buffers has us hopeful for future success.

On the other hand, the biological effects of TcTS inhibition have remained elusive, largely due to the lack of sufficiently potent inhibitors for use in cell-based systems. While a couple of non-carbohydrate based inhibitors with suitable potency have been more recently identified, none of which have been

evaluated in such systems. Our first attempt to investigate the effects of TcTS inhibition on cell-invasion by *T. cruzi* (with Dr. Fred Buckner) led to mixed results. We found that two synthetic ircinialactams **3.14** and **3.15** inhibited trypomastigote growth over 6 days with half maximal effective concentrations  $<50 \mu\text{M}$  ( $\text{EC}_{50} = 25 \mu\text{M}$  and  $\text{EC}_{50} = 32 \mu\text{M}$  respectively). Yet four structurally related compounds failed to exhibit any inhibition at similar concentrations despite equally potent  $\text{IC}_{50}$  values measured against recombinant TcTS. We posited that charge state of the ircinialactams may have influenced the biological activity and warrants further investigation. Additionally, the bioassay we employed was developed to assess the potency of antiparasitic therapeutics like benznidazole, and thus may not be suitable for our purposes. Therefore additional cell-based experiments must be designed and executed to convincingly understand the role of TcTS inhibition on *T. cruzi* invasivity.

The overall goal of this thesis was to identify new chemical inhibitors of TcTS as tools that could be used to evaluate the enzyme as a therapeutic target in biological systems. The projects presented illustrate three different yet promising approaches for the development of such inhibitors. The synthesis of rational DFSA based inactivators afforded the new gold-standard for covalent inhibitors of TcTS ( $k_i/K_i = 81 \times 10^{-3} \text{ min}^{-1} \text{ mM}^{-1}$ ) and illustrated that inactivation potency can be enhanced 10-fold by targeted modification of DFSAs. Further, the combined kinetic information acquired from the C9 and C5 modified DFSA libraries spawns a promising direction for the future synthesis of disubstituted DFSA as selective inactivators of TcTS versus human neuraminidases. Meanwhile the potent non-carbohydrate based inhibitors of TcTS that we revealed through marine natural product and peptide-based screening were of similar potency to the best known reversible inhibitors; moreover, our results highlight the power of these screening techniques for the discovery of new inhibitors of difficult enzyme targets. These new chemical scaffolds form the basis for future synthetic studies which can further enhance the potency of these reversible TcTS inhibitors. Finally, these species can now be leveraged as chemical tools in biological systems to conclusively assess the therapeutic importance of TcTS for Chagas disease.

## Chapter 7: Experimental Procedures

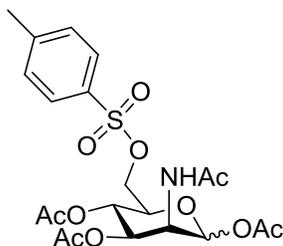
### 7.1 Synthesis of DFSAs

#### 7.1.1 General materials

All reagents were obtained from commercial suppliers (*i.e.* Sigma-Aldrich, Carbosynth, Alfa Aesar, TCI) and were used without further purification unless otherwise stated. Column chromatography was performed using silica gel (230-400 mesh), TLC was performed on pre-coated silica plates (60 F<sub>254</sub> aluminium sheets). TLC was visualized by UV light and/or by staining; typical stain solutions were ammonium molybdate (10%) in 2 M H<sub>2</sub>SO<sub>4</sub> or ninhydrin (1%) in 2:1 *n*-BuOH/EtOH followed by heating. Anhydrous solvents were distilled under a nitrogen atmosphere; pyridine, DCM and ACN over CaH<sub>2</sub>, THF over Na, and MeOH over Mg. DMF was dried over molecular sieves (3 Å). Moisture-sensitive reactions were carried out under an atmosphere of nitrogen or argon. NMR samples were recorded in CDCl<sub>3</sub>, DMSO-d<sub>6</sub>, MeOD or D<sub>2</sub>O and spectra were collected on Bruker AV-300, AV-400 or AV-600 instruments. All chemical shifts are reported in ppm ( $\delta$ ) relative to solvent standard. Low resolution mass spectrometry (LRMS) was performed on a Waters ZQ mass spectrometer equipped with ESCI ion source. High resolution mass spectrometry was performed by the University of British Columbia mass spectrometry facility using a Waters/Micromass LCT instrument with time of flight detection and electrospray ionization.

Compounds **2.19-2.27**, **2.47** and **2.57-2.59** have been previously synthesized and characterized by Buchini et al. <sup>1</sup>H-NMR, <sup>19</sup>F-NMR and ESI-MS data presented below match reported values.<sup>88,101</sup>

### 7.1.2 Synthesis of 9-amido-DFSA

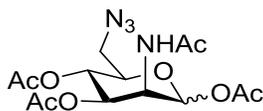


#### 2-Acetamido-1,3,4-tri-O-acetyl-2-deoxy-6-O-tosyl-D-mannopyranose (2.19)

Recrystallized tosyl chloride (20.74 g, 109 mmol) was dissolved in pyridine (100 mL) and added dropwise over 75 minutes to a solution of *N*-acetylmannosamine (10 g, 45.2 mmol) in pyridine (140 mL) while stirring at 0 °C. After 5 hours acetic anhydride (80 mL) was added, warmed to room temperature and stirred overnight. Solvent was evaporated, after which EtOAc (100 mL) was added. Organic phase was washed with 1 M HCl, NaHCO<sub>3</sub> solution and brine, dried over MgSO<sub>4</sub>, filtered and solvent evaporated. Purification by silica gel column chromatography (1:1 to 5:1 EtOAc/Hexanes) yielded partially purified product (contaminated with ditosylated by-product). A second silica gel column was performed (4:1 to 2:1 DCM/Acetone) to yield 18.2 g of product (80% yield).

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.76 (d, *J* = 8.3 Hz, 2H, H-2', H-6' Ar), 7.34 (d, *J* = 7.9 Hz, 2H, H-3', H-5' Ar), 6.27 (d, *J*<sub>NH,2</sub> = 9.3 Hz, 1H, NH), 5.98 (d, *J*<sub>1,2</sub> = 1.9 Hz, 1H, H-1), 5.38 – 5.15 (m, 2H, H-3, H-4), 4.60 (ddd, *J*<sub>2,NH</sub> = 9.4 Hz, *J*<sub>2,3</sub> = 3.9 Hz, *J*<sub>2,1</sub> = 1.9 Hz, 1H, H-2), 4.24 (dd, *J*<sub>5,6</sub> = 11.3 Hz, *J*<sub>5,4</sub> = 2.1 Hz, 1H, H-5), 4.09 – 3.98 (m, 2H, H-6), 2.43 (s, 3H, *Me*-Ar), 2.12 (s, 3H, NHAc), 2.03 (s, 3H, OAc), 1.99 (s, 3H, OAc), 1.98 (s, 3H, OAc).

ESI-LRMS: Calc'd for [C<sub>21</sub>H<sub>27</sub>NO<sub>11</sub>S + Na]<sup>+</sup>: 524.1197; found: 524.2.

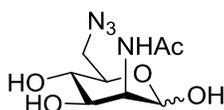


#### 2-Acetamido-1,3,4-tri-O-acetyl-6-azido-2,6-dideoxy-D-mannopyranose (2.20)

Tetrabutylammonium azide (TBA-N<sub>3</sub>) was freshly prepared from tetrabutylammonium hydroxide and sodium azide<sup>201</sup>. TBA-N<sub>3</sub> (6.9 g, 24.3 mmol) was added to a solution of **2.19** (6.0 g, 12.0 mmol) in DMF (80 mL; dried over MS) and heated to 70 °C. After consumption of starting material (~12 hours), the solvent was evaporated and the dark oil dissolved in EtOAc (150 mL). The organic phase was washed with H<sub>2</sub>O (2 x 50 mL) and brine (2 x 50 mL), the combined water extracts were back-extracted with EtOAc (2 x 75 mL) and washed with brine (75 mL). The combined organic layers were dried over MgSO<sub>4</sub> and evaporated to yield an off-white foam. Silica gel column chromatography (2:1 to 3:1 EtOAc/Hexanes) afforded 3.54 g (81% yield) of product as a white foam.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 6.05 (d, *J*<sub>1,2</sub> = 1.9 Hz, 1H, H-1), 5.83 (d, *J*<sub>NH,2</sub> = 9.5 Hz, 1H, NH), 5.33 (dd, *J*<sub>3,4</sub> = 10.1 Hz, *J*<sub>3,2</sub> = 4.5 Hz, 1H, H-3), 5.19 (t, *J*<sub>4,3</sub> = 10.0 Hz, *J*<sub>4,5</sub> = 10.0 Hz, 1H, H-4), 4.63 (ddd, *J*<sub>2,NH</sub> = 9.2 Hz, *J*<sub>2,3</sub> = 4.4 Hz, *J*<sub>2,1</sub> = 1.9 Hz, 1H, H-2), 3.97 (ddd, *J*<sub>5,4</sub> = 9.8 Hz, *J*<sub>5,6b</sub> = 4.8 Hz, *J*<sub>5,6a</sub> = 2.8 Hz, 1H, H-5), 3.45 (dd, *J*<sub>6a,6b</sub> = 13.5 Hz, *J*<sub>6a,5</sub> = 2.8 Hz, 1H, H-6a), 3.31 (dd, *J*<sub>6b,6a</sub> = 13.5 Hz, *J*<sub>6b,5</sub> = 4.8 Hz, 1H, H-6b), 2.17 (s, 3H, NHAc), 2.07 (s, 6H, 2 x OAc), 2.00 (s, 3H, OAc).

ESI-LRMS: Calc'd for [C<sub>14</sub>H<sub>20</sub>N<sub>4</sub>O<sub>8</sub> + Na]<sup>+</sup>: 395.1173; found: 395.2.

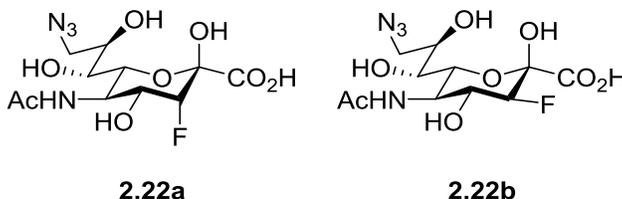


### **2-Acetamido-6-azido-2,6-dideoxy- $\alpha$ -D-mannopyranose (2.21)**

Na<sub>(s)</sub> was dissolved in methanol (0.34 g in 74 mL, 0.2 M) and added slowly to a cooled solution of **2.20** (3.2 g, 8.6 mmol) in anhydrous methanol (125 mL). The reaction mixture was stirred at 0 °C until completion (90 mins) as determined by TLC (5:1 DCM/MeOH). Reaction was quenched by addition of Amberlite IR120-H resin to the stirred solution until the pH was neutral. Filtration of the resin evaporation of solvent afforded 2.08 g (98% yield) of product as a white solid. Product was used without purification for subsequent step.

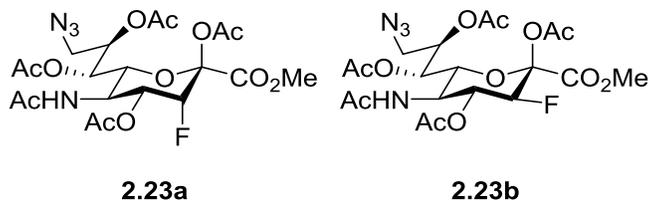
$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  5.05 (d,  $J_{1,2} = 1.6$  Hz, 1H, H-1), 4.25 (dd,  $J_{2,3} = 4.9$  Hz,  $J_{2,1} = 1.6$  Hz, 1H, H-2), 4.00 (dd,  $J_{3,4} = 9.6$  Hz,  $J_{3,2} = 4.8$  Hz, 1H, H-3), 3.93 (ddd,  $J_{5,4} = 9.6$  Hz,  $J_{5,6} = 6.4$  Hz,  $J_{5,6} = 2.9$  Hz, 1H, H-5), 3.58 – 3.46 (m, 3H, H-4, H-6), 2.03 (s, 3H,  $\text{NHAc}$ ).

ESI-LRMS: Calc'd for  $[\text{C}_8\text{H}_{14}\text{N}_4\text{O}_5 + \text{Na}]^+$ : 269.0856; found: 269.3.



**5-Acetamido-9-azido-3,5,9-trideoxy-3-fluoro-D-erythro- $\alpha$ -L-manno-non-2-ulopyranosonic acid (2.22a) and gluco-non-2-ulopyranosonic acid (2.22b)**

Neu5Ac aldolase (80 mg, Carbosynth) was added to a solution of **2.21** (2.0 g, 8.1 mmol) and 3-fluoropyruvate (0.3 g, 2.3 mmol, 0.28 eq) in  $\text{H}_2\text{O}$  (50 mL). Reaction progress was monitored by  $^{19}\text{F}$  NMR, aliquots of 3-fluoropyruvate were added after 10 h (0.3 g, 0.28 eq), 28 h (0.3 g, 0.28 eq) and 55 h (0.3 g, 0.28 eq) respectively. Reaction was quenched by flash freezing on dry ice after 74 hours and lyophilized to yield 3.2 g of crude off-white solid (5:1 **2.22a/2.22b**). Solid was used directly in next reaction.



**Methyl 5-acetamido-2,4,7,8-tetra-O-acetyl-9-azido-3,5,9-trideoxy-3-fluoro-D-erythro- $\alpha$ -L-manno-non-2-ulopyranosonate (2.23a) and gluco-non-2-ulopyranosonate (2.23b)**

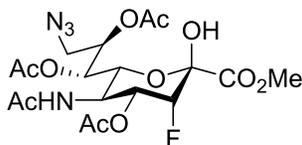
3.15 g of crude **2.22** was dissolved in anhydrous MeOH (300 mL) and stirred with Amberlite IR120-H resin (12.8 g) overnight. The cloudy reaction mixture was filtered over Celite, washed with MeOH and

solvent removed to yield 3.02 g crude solid. TLC (9:1 DCM/MeOH) showed two overlapping spots and ESI-LRMS confirmed m/z of product (ESI-LRMS: Calc'd for  $[C_{12}H_{19}FN_4O_8 + Na]^+$ : 389.1079; found: 389.3), so the crude material was carried forward without further purification.  $Ac_2O$  (35 mL) was added to a solution of the dissolved solid (2.0 g) in pyridine (15 mL) and the reaction mixture was stirred overnight. Solvent volume was reduced by rotary evaporation and co-evaporated thrice with toluene to yield a clear oil. Products were isolated by silica gel column chromatography (4:1 to 6:1 EtOAc/Hexanes) to afford 2.23 g of **2.23a** (76% yield, rf = 0.6, EtOAc) and 0.47 g **2.23b** (16% yield, rf = 0.5, EtOAc).

$^1H$  NMR (300 MHz,  $CDCl_3$ )  $\delta$  5.56 – 5.41 (m, 2H, H-4, *NH*), 5.39 (dd,  $J_{7,8} = 3.5$  Hz,  $J_{7,6} = 2.0$  Hz, 1H, H-7), 4.96 (dt,  $J_{8,9b} = 7.8$  Hz,  $J_{8,7} = 3.0$  Hz,  $J_{8,9a} = 3.0$  Hz, 1H, H-8), 4.95 (dd,  $J_{3,F3} = 49.0$  Hz,  $J_{3,4} = 2.5$  Hz, 1H, H-3), 4.30 (q,  $J_{5,4} = 10.2$  Hz,  $J_{5,6} = 10.2$  Hz,  $J_{5,NH} = 10.2$  Hz, 1H, H-5), 4.20 (dd,  $J_{6,5} = 10.6$  Hz,  $J_{6,7} = 2.1$  Hz, 1H, H-6), 3.92 (dd,  $J_{9a,9b} = 13.7$  Hz,  $J_{9a,8} = 2.7$  Hz, 1H, H-9a), 3.88 (s, 3H, *OMe*), 3.45 (dd,  $J_{9b,9a} = 13.6$  Hz,  $J_{9b,8} = 7.8$  Hz, 1H, H-9b), 2.21 (s, 3H, *Ac*), 2.21 (s, 3H, *Ac*), 2.13 (s, 3H, *Ac*), 2.11 (s, 3H, *Ac*), 1.93 (s, 3H, *NHAc*).

$^{19}F$  NMR (282 MHz,  $CDCl_3$ )  $\delta$  -209.21 (dd,  $J_{F3,H3} = 49.1$  Hz,  $J_{F3,H4} = 27.8$  Hz, F-3 ax).

ESI-LRMS: Calc'd for  $[C_{20}H_{27}FN_4O_{12} + Na]^+$ : 557.1502; found: 557.3.



**Methyl 5-acetamido-4,7,8-tri-*O*-acetyl-9-azido-3,5,9-trideoxy-3-fluoro-*D*-erythro- $\alpha$ -*L*-manno-non-2-ulopyranosonate (2.24)**

Hydrazine acetate (0.95 g, 10.3 mmol) was dissolved in MeOH (20 mL) and added slowly to a cooled solution of **2.23a** (1.30 g, 2.43 mmol) in 10 mL DCM, and the reaction mixture stirred for 6 hours at 0 °C. The solvent was evaporated, solid was dissolved in EtOAc (100 mL) and washed with 1 M HCl (50 mL),  $NaHCO_3$  (50 mL),  $H_2O$  (50 mL). Organic layer was dried over anhydrous  $MgSO_4$  and evaporated to yield

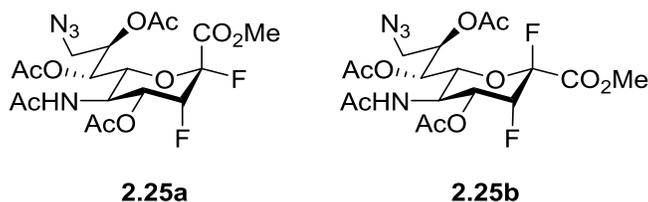
crude white solid. Purification by silica gel column chromatography (4:1 to 6:1 EtOAc/Hexanes) afforded 0.76 g (63% yield) of the product **2.24** as a white foam.

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  6.11 (d,  $J_{\text{NH},5} = 9.4$  Hz, 1H, NH), 5.80 (s, 1H, OH), 5.54 (ddd,  $J_{4,3\text{F}} = 28.1$  Hz,  $J_{4,5} = 10.9$  Hz,  $J_{4,3} = 2.4$  Hz, 1H, H-4), 5.37 (t,  $J_{7,6} = 2.6$  Hz,  $J_{7,8} = 2.6$  Hz, 1H, H-7), 5.14 (dt,  $J_{8,9\text{b}} = 9.1$  Hz,  $J_{8,7} = 2.6$  Hz,  $J_{8,9\text{a}} = 2.6$  Hz, 1H, H-8), 4.98 (dd,  $J_{3,3\text{F}} = 49.6$  Hz,  $J_{3,4} = 2.3$  Hz, 1H, H-3), 4.49 (dd,  $J_{6,5} = 10.6$  Hz,  $J_{6,7} = 2.0$  Hz, 1H, H-6), 4.21 (q,  $J_{5,\text{NH}} = 10.4$  Hz,  $J_{5,4} = 10.4$  Hz,  $J_{5,6} = 10.4$  Hz, 1H, H-5), 3.95 (dd,  $J_{9\text{a},9\text{b}} = 13.7$  Hz,  $J_{9\text{a},8} = 2.2$  Hz, 1H, H-9a), 3.86 (s, 3H, OMe), 3.44 (dd,  $J_{9\text{a},9\text{a}} = 13.6$  Hz,  $J_{9\text{a},8} = 9.0$  Hz, 1H, H-9b), 2.18 (s, 3H, OAc), 2.13 (s, 3H, OAc), 2.09 (s, 3H, OAc), 1.93 (s, 3H, NHAc).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.1 (NH-CO), 171.0 (2C, CO), 170.8 (CO), 167.6 (C-1), 94.5 (d,  $J_{\text{C}2,\text{F}3} = 25.3$  Hz, C-2), 86.8 (d,  $J_{\text{C}3,\text{F}3} = 184.8$  Hz, C-3), 75.2 (C-8), 71.0 (C-6), 69.1 (d,  $J_{\text{C}4,\text{F}3} = 17.5$  Hz, C-4), 68.9 (C-7), 53.5 (OMe), 50.7 (C-9), 45.7 (C-5), 29.8, 23.2 (NHAc), 21.4 (OAc), 20.9 (OAc), 20.9 (OAc).

$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -206.17 (s, F-3 ax).

ESI-LRMS: Calc'd for  $[\text{C}_{18}\text{H}_{25}\text{FN}_4\text{O}_{11} + \text{Na}]^+$ : 515.1396; found: 515.2.



**Methyl 5-acetamido-4,7,8-tri-O-acetyl-9-azido-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulopyranosylonate fluoride (2.25a) and - $\alpha$ -L-manno-non-2-ulopyranosylonate fluoride (2.25b)**

Diethylaminosulfur trifluoride (DAST, 70  $\mu\text{L}$ , 0.53 mmol) was added dropwise via glass syringe over 15 minutes to a solution of **2.24** (180 mg, 0.36 mmol) in DCM (7 mL) and stirred at  $-30$   $^\circ\text{C}$  for 60 minutes. TLC conditions to monitor the reaction by separating the starting material and product were determined (4:1 DCM/Acetone); however, separate conditions (4:1 EtOAc/Hexanes) were needed to isolate the two

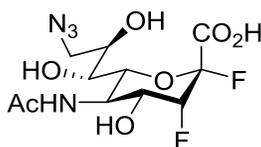
stereoisomers formed during the reaction. Methanol (100  $\mu$ L) was added to quench the excess DAST, the solvent was removed and solid dissolved in DCM. Organic phase was washed with water and brine, the aqueous wash was extracted with DCM and washed with brine. Combined organic layer was dried over  $MgSO_4$  and the solvent evaporated. Silica gel column chromatography (2:1 to 3:1 EtOAc/Hexanes) afforded 102 mg (56% yield) of **2.25a** and 50 mg (28% yield) of **2.25b**.

$^1H$  NMR (300 MHz,  $CDCl_3$ )  $\delta$  5.54 (d,  $J_{NH,5} = 8.9$  Hz, 1H, NH), 5.49 – 4.99 (m, 4H, H-4, H-7, H-8, H-3), 4.29 (d,  $J_{6,5} = 10.8$  Hz, 1H, H-6), 4.18 (q,  $J_{5,4} = 10.0$  Hz,  $J_{5,6} = 10.0$  Hz,  $J_{5,NH} = 10.0$  Hz, 1H, H-5), 3.91 (s, 3H, OMe), 3.68 (dd,  $J_{9a,9b} = 13.6$  Hz,  $J_{9a,98} = 2.8$  Hz, 1H), 3.34 (dd,  $J_{9b,9a} = 13.6$  Hz,  $J_{9b,8} = 6.1$  Hz, 1H, H-9b), 2.18 (s, 3H, OAc), 2.13 (s, 3H, OAc), 2.12 (s, 3H, OAc), 1.93 (s, 3H, NHAc).

$^{13}C$  NMR (101 MHz,  $CDCl_3$ )  $\delta$  170.9, 170.6, 170.6, 170.4, 164.5 (dd,  $J_{C1,F2} = 29.6$  Hz,  $J_{C1,F3} = 3.5$  Hz, C-1), 104.8 (dd,  $J_{C2,F2} = 225.9$  Hz,  $J_{C2,F3} = 17.2$  Hz, C-2), 85.6 (dd,  $J_{C3,F3} = 194.0$  Hz,  $J_{C3,F3} = 19.4$  Hz, C-3), 72.8 (C-6), 71.2 (C-8), 68.9 (dd,  $J_{C4,F3} = 16.1$  Hz,  $J_{C4,F2} = 4.8$  Hz, C-4), 67.8 (C-7), 54.0 (OMe), 50.5 (C-9), 45.3 (C-5), 23.3 (NHAc), 21.0 (OAc), 20.8 (OAc), 20.7 (OAc).

$^{19}F$  NMR (282 MHz,  $CDCl_3$ )  $\delta$  -123.56 (d,  $J_{F2,F3} = 11.5$  Hz, F-2 eq), -217.06 (ddd,  $J_{F3,H3} = 51.3$  Hz,  $J_{F3,H4} = 25.9$  Hz,  $J_{F3,F2} = 11.3$  Hz, F-3 ax).

ESI-LRMS: Calc'd for  $[C_{18}H_{24}F_2N_4O_{10} + Na]^+$ : 517.1352; found: 517.2.



### 5-Acetamido-9-azido-2,3,5,9-tetraoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulopyranosylonic fluoride (**2.26**)

NaOMe in methanol was prepared by addition of solid sodium (60 mg) to 10 mL of anhydrous MeOH. Solution was added slowly to a cooled solution of **2.25a** (0.67 g, mmol) in MeOH (30 mL) and stirred for 3 hours at 0  $^{\circ}C$  until all acetates were removed ( $R_f = 0.3$ , EtOAc). 1 M NaOH (1 mL) was slowly added

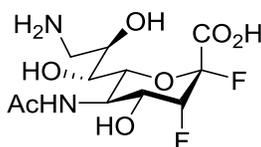
to the cooled reaction and stirred for an additional 30 minutes. Reaction was quenched by addition of Amberlite IR120-H<sup>+</sup> resin, the solution was filtered and the solvent evaporated, dissolved in H<sub>2</sub>O and lyophilized to yield 0.48 g (quantitative) of **2.26** as a white solid.

<sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O) δ 5.20 (dt,  $J_{3,F3} = 51.3$  Hz,  $J_{3,F2} = 2.6$  Hz,  $J_{3,4} = 2.6$  Hz, 1H, H-3), 4.36 – 4.07 (m, 2H, H-4, H-5), 4.00 (ddd,  $J_{8,7} = 9.0$  Hz,  $J_{8,9b} = 5.9$  Hz,  $J_{8,9a} = 2.8$  Hz, 1H, H-8), 3.86 (dd,  $J_{6,5} = 10.4$  Hz,  $J_{6,7} = 1.4$  Hz, 1H, H-6), 3.63 (dd,  $J_{9a,9b} = 13.2$  Hz,  $J_{9a,8} = 2.8$  Hz, 1H, H-9a), 3.56 (dt,  $J_{7,8} = 9.3$  Hz,  $J_{7,6} = 1.7$  Hz, 1H, H-7), 3.46 (dd,  $J_{9b,9a} = 13.2$  Hz,  $J_{9b,8} = 5.9$  Hz, 1H, H-9b), 2.04 (s, 3H, NHAc).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) δ 175.2 (NH CO), 168.3 (dd,  $J_{C1,F2} = 28.7$  Hz,  $J_{C1,F3} = 4.0$  Hz, C-1), 106.2 (dd,  $J_{C2,F2} = 219.6$  Hz,  $J_{C2,F3} = 15.5$  Hz, C-2), 88.5 (dd,  $J_{C3,F3} = 184.8$  Hz,  $J_{C3,F2} = 18.2$  Hz, C-3), 72.8 (d,  $J_{C6,F2} = 4.0$  Hz, C-6), 69.2 (C-7), 69.0 (dd,  $J_{C4,F3} = 17.9$  Hz,  $J_{C4,F2} = 5.8$  Hz, C-4), 68.5 (C-8), 53.9 (C-9), 47.0 (d,  $J_{C5,F3} = 3.5$  Hz, C-5), 22.3 (NHAc).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O) δ -122.17 (d,  $J_{F2,F3} = 11.1$  Hz, F-2 eq), -218.77 (ddd,  $J_{F3,H3} = 48.9$  Hz,  $J_{F3,H4} = 28.5$  Hz,  $J_{F3,F2} = 10.9$  Hz, F-3 ax).

ESI-LRMS: Calc'd for [C<sub>11</sub>H<sub>16</sub>F<sub>2</sub>N<sub>4</sub>O<sub>7</sub> + Na]<sup>+</sup>: 377.0878; found: 377.1.



### **5-Acetamido-9-amino-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro-β-L-manno-non-2-ulopyranosylonic fluoride (2.27)**

Pd/C (32.4 mg, 10% w/w) was added to **2.26** (320 mg, 0.90 mmol) and dissolved in MeOH (35 mL). H<sub>2</sub>(g) was gently bubbled through the solution for 15 minutes, then the reaction was stirred for 6 hours. The dark reaction mixture was filtered over Celite and washed with MeOH and solvent was evaporated. TLC (5:2:1 EtOAc/MeOH/H<sub>2</sub>O) showed a single spot which appeared red after staining with ninhydrin and

charring. The product was dissolved in H<sub>2</sub>O and lyophilized to afford 265 mg (90% yield) of **2.27** as a white solid.

<sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O) δ 5.19 (dt,  $J_{3,F3} = 51.7$  Hz,  $J_{3,F2} = 2.6$  Hz,  $J_{3,4} = 2.6$  Hz, 1H, H-3), 4.29 – 4.07 (m, 2H, H-4, H-5), 4.03 (ddd,  $J_{8,9a} = 9.5$  Hz,  $J_{8,7} = 8.5$  Hz,  $J_{8,9a} = 3.3$  Hz, 1H, H-8), 3.77 (d,  $J_{6,5} = 9.9$  Hz, 1H, H-6), 3.54 (dt,  $J_{7,8} = 8.5$  Hz,  $J_{7,6} = 1.7$  Hz, 1H, H-7), 3.40 (dd,  $J_{9a,9b} = 13.7$  Hz,  $J_{9a,8} = 3.8$  Hz, 1H, H-9a), 2.98 (dd,  $J_{9b,9a} = 13.1$  Hz,  $J_{9b,8} = 9.5$  Hz, 1H, H-9b), 2.02 (s, 3H, NHAc).

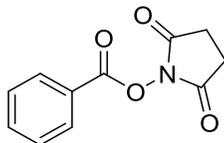
<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) δ 175.1 (NH CO), 169.4 (dd,  $J_{C1,F2} = 26.6$  Hz,  $J_{C1,F3} = 3.4$  Hz, C-1), 106.6 (dd,  $J_{C2,F2} = 220.0$  Hz,  $J_{C2,F3} = 15.0$  Hz, C-2), 88.9 (dd,  $J_{C3,F3} = 184.1$  Hz,  $J_{C3,F2} = 18.5$  Hz, C-3), 72.4 (d,  $J_{C6,F2} = 4.4$  Hz, C-6), 69.7 (C-7/C-8), 69.0 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 6.2$  Hz, C-4), 67.1 (C-7/C-8), 46.8 (d,  $J_{C5,F3} = 3.4$  Hz, C-5), 42.3 (C-9), 22.0 (NHAc).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O) δ -121.65 (d,  $J_{F2,F3} = 12.1$  Hz, F-2 eq), -218.34 (ddd,  $J_{F3,H3} = 51.6$  Hz,  $J_{F3,H4} = 27.8$  Hz,  $J_{F3,F2} = 11.9$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>11</sub>H<sub>18</sub>F<sub>2</sub>N<sub>2</sub>O<sub>7</sub> + H]<sup>+</sup>: 329.1161; found: 329.1164.

### 7.1.3 General method for the synthesis of NHS-activated esters

Carboxylic acid starting materials (1 g), *N*-hydroxysuccinimide (NHS; 1.1 eq) and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC, 1.1 eq) were dissolved in DCM (15 mL) and were reacted at room temperature overnight. Reaction was diluted with DCM, washed with H<sub>2</sub>O and brine, the organic phase was dried over MgSO<sub>4</sub> and the solvent was evaporated. Purification by silica gel column chromatography (DCM/MeOH) and/or crystallization (EtOAc/hexanes) yielded the products **2.37-2.44** as white solids.



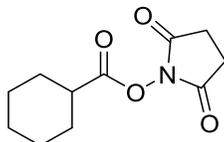
### 2,5-Dioxopyrrolidin-1-yl benzoate (2.37)

Benzoic acid (1.0 g, 8.2 mmol), EDC (1.72 g, 1.1 eq) and *N*-hydroxysuccinimide (1.04 g, 1.1 eq) were coupled following the *General method for the synthesis of activated esters*. Silica gel column chromatography (9:1 DCM/MeOH) afforded 0.90 g (50% yield) of **2.37** as a white solid.

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.17 – 8.08 (m, 2H; H-2, H-6), 7.66 (td,  $J = 7.4, 1.4$  Hz, 1H; H-4), 7.50 (t,  $J = 7.8$  Hz, 2H; H-3, H-5), 2.88 (s, 4H;  $\text{CH}_2$ ).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  169.5 (CO succinimide), 162.0 (CO Bz), 135.0 (C-4), 130.6 (2C, C-3, C-5), 129.0 (2C, C-2, C-6), 125.2 (C-1), 25.7 (2 x  $\text{CH}_2$  succinimide).

ESI-LRMS: Calc'd for  $[\text{C}_{11}\text{H}_9\text{NO}_4 + \text{Na}]^+$ : 242.1852; found: 242.3.



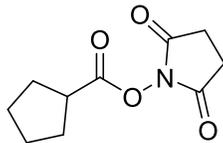
### 2,5-Dioxopyrrolidin-1-yl cyclohexanecarboxylate (2.38)

Cyclohexanecarboxylic acid (1.0 g, 7.8 mmol), EDC (1.64, 1.1 eq) and *N*-hydroxysuccinimide (1.0 g, 1.1 eq) were coupled following the *General method for the synthesis of activated esters*. Silica gel column chromatography (9:1 DCM/MeOH) afforded 0.85 g (48% yield) of **2.38** as a white solid.

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  2.82 (d,  $J = 1.8$  Hz, 4H;  $\text{CH}_2$  succinimide), 2.67 (tt,  $J = 10.9, 3.7$  Hz, 1H; H-1), 2.04 (dt,  $J = 12.8, 3.7$  Hz, 2H), 1.88 – 1.72 (m, 2H), 1.62 (tdd,  $J = 16.0, 8.7, 3.1$  Hz, 4H), 1.45 – 1.17 (m, 2H).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  171.0 (CO), 169.4 (CO succinimide), 40.6, 28.9, 25.8, 25.6, 25.1.

ESI-LRMS: Calc'd for  $[\text{C}_{11}\text{H}_{15}\text{NO}_4 + \text{Na}]^+$ : 248.0894; found: 248.3.



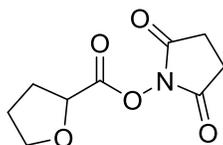
### 2,5-Dioxopyrrolidin-1-yl cyclopentanecarboxylate (**2.39**)

Cyclopentanecarboxylic acid (1.0 g, 8.8 mmol), EDC (1.86, 1.1 eq) and *N*-hydroxysuccinimide (1.11 g, 9.6 mmol, 1.1 eq) were coupled following the *General method for the synthesis of activated esters*. Silica gel column chromatography (9:1 DCM/MeOH) afforded 0.88 g (47% yield) of **2.39** as a white solid.

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.03 (t,  $J = 7.9$  Hz, 1H; H-1), 2.80 (s, 4H;  $\text{CH}_2$  succinimide), 1.99 (dq,  $J = 13.5, 7.0$  Hz, 4H), 1.85 – 1.53 (m, 4H).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  172.0 (CO), 169.4 (CO succinimide), 40.7 (C-1), 30.2 (C-2, C-5), 26.0 (C-3, C-4), 25.7 ( $\text{CH}_2$  succinimide).

ESI-LRMS: Calc'd for  $[\text{C}_{10}\text{H}_{13}\text{NO}_4 + \text{Na}]^+$ : 234.0437; found: 234.3.



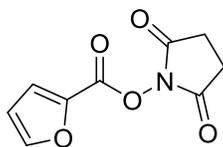
### 2,5-Dioxopyrrolidin-1-yl tetrahydrofuran-2-carboxylate (**2.40**)

Tetrahydrofuran-2-carboxylic acid (1.0 g, 8.6 mmol), EDC (1.74 g, 1.05 eq) and *N*-hydroxysuccinimide (1.04 g, 1.05 eq) were coupled following the *General method for the synthesis of activated esters*. Silica gel column chromatography (9:1 DCM/MeOH) afforded 0.93 g (51% yield) of **2.40** as a white solid.

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  4.80 (dd,  $J = 8.4, 5.1$  Hz, 1H), 4.12 – 3.91 (m, 2H), 2.84 (s, 4H,  $\text{CH}_2$  succinimide), 2.48 – 2.23 (m, 2H), 2.15 – 1.88 (m, 2H).

$^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  169.1 (CO), 168.9 (CO), 74.9 (C-1), 69.9 (C-4), 30.8 (C-2), 25.6 ( $\text{CH}_2$  succinimide), 25.1 (C-3).

ESI-LRMS: Calc'd for  $[C_9H_{11}NO_5 + Na]^+$ : 236.0529; found: 236.4.



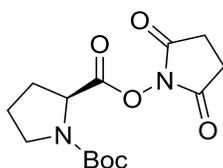
### 2,5-Dioxopyrrolidin-1-yl furan-2-carboxylate (**2.41**)

2-Furoic acid (1.0 g, 8.9 mmol), EDC (1.80 g, 1.1 eq) and *N*-hydroxysuccinimide (1.10 g, 1.1 eq) were coupled following the *General method for the synthesis of activated esters*. Silica gel column chromatography (9:1 DCM/MeOH) afforded 0.91 g (49% yield) of **2.41** as a white solid.

$^1H$  NMR (300 MHz,  $CDCl_3$ )  $\delta$  7.73 (d,  $J = 1.7$  Hz, 1H; H-4), 7.53 – 7.43 (m, 1H; H-2), 6.62 (dd,  $J = 3.7$ , 1.7 Hz, 1H; H-3), 2.89 (s, 4H;  $CH_2$  succinimide).

$^{13}C$  NMR (75 MHz,  $CDCl_3$ )  $\delta$  169.3 (CO succinimide), 153.6 (CO) 148.9 (C-4), 139.7 (C-1), 122.3 (C-2), 112.7 (C-3), 25.7 ( $CH_2$  succinimide).

ESI-LRMS: Calc'd for  $[C_9H_7NO_5 + Na]^+$ : 232.0216; found: 232.3.



### 1-(*tert*-Butyl) 2-(2,5-dioxopyrrolidin-1-yl) pyrrolidine-1,2-dicarboxylate (**2.42**)

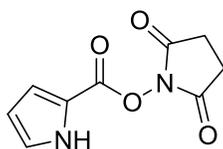
L-Proline (1.08 g, 9.4 mmol) and boc anhydride (2.15 g, 9.9 mmol) were dissolved in DCM (15 mL) and triethylamine (1.3 mL, 9.3 mmol) was added. The reaction was stirred for 2 hours until completion, reaction was quenched with acetic acid and the solvent evaporated.

Product, EDC (2.15 g, 1.2 eq) and *N*-hydroxysuccinimide (1.30 g, 1.2 eq) were coupled following the *General method for the synthesis of activated esters*.

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.54 (dd,  $J = 8.8, 3.7$  Hz, 1H), 3.60 (ddd,  $J = 10.4, 7.8, 4.5$  Hz, 1H), 3.46 (dt,  $J = 10.5, 7.5$  Hz, 1H), 2.93 – 2.74 (m, 4H), 2.47 – 2.24 (m, 2H), 2.10 – 1.88 (m, 2H), 1.47 (s, 9H,  $\text{C}(\text{CH}_3)_3$ ).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  168.9 (CO succinimide), 168.8 (CO Boc), 153.6 (CO), 81.3 (C Boc), 57.3 (C-1), 46.4 (C-4), 31.5 (C-2), 28.2 ( $\text{CH}_3$  Boc), 25.7 ( $\text{CH}_2$  succinimide), 23.6 (C-3).

ESI-LRMS: Calc'd for  $[\text{C}_{14}\text{H}_{20}\text{N}_2\text{O}_6 + \text{Na}]^+$ : 335.1213; found: 335.4.



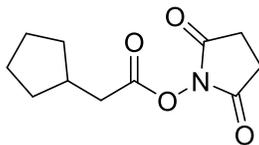
#### 2,5-Dioxopyrrolidin-1-yl 1H-pyrrole-2-carboxylate (2.43)

Pyrrole-2-carboxylic acid (0.5 g, 4.5 mmol), EDC (0.95 g, 1.1 eq) and *N*-hydroxysuccinimide (0.57 g, 1.1 eq) were coupled following the *General method for the synthesis of activated esters*. Silica gel column chromatography (9:1 DCM/MeOH) afforded 0.45 g (48% yield) of **2.43** as a white solid.

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  9.54 (s, 1H; NH), 7.17 (s, 1H; H-4), 7.09 (q,  $J = 1.4$  Hz, 1H; H-2), 6.39 – 6.26 (m, 1H; H-3), 2.87 (s, 4H;  $\text{CH}_2$  succinimide).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  169.9 (CO succinimide), 155.9 (CO), 126.3 (C-1), 119.6 (C-4), 116.8 (C-2), 111.7 (C-3), 25.8 ( $\text{CH}_2$  succinimide).

ESI-LRMS: Calc'd for  $[\text{C}_9\text{H}_8\text{N}_2\text{O}_4 + \text{Na}]^+$ : 231.0376; found: 231.3.



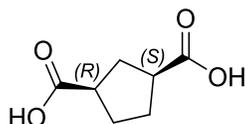
#### 2,5-Dioxopyrrolidin-1-yl 2-cyclopentylacetate (2.44)

Cyclopentaneacetic acid (1.0 g, 7.8 mmol), EDC (1.64 g, 1.1 eq) and *N*-hydroxysuccinimide (0.96 g, 1.1 eq) were coupled following the *General method for the synthesis of activated esters*. Silica gel column chromatography (9:1 DCM/MeOH) afforded 0.82 g (52% yield) of **2.44** as a white solid.

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 2.78 (d, *J* = 2.9 Hz, 4H; CH<sub>2</sub> succinimide), 2.56 (d, *J* = 7.4 Hz, 2H), 2.27 (p, *J* = 7.8 Hz, 1H), 1.95 – 1.82 (m, 2H), 1.70 – 1.44 (m, 4H), 1.29 – 1.14 (m, 2H).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 169.3 (CO succinimide), 168.1 (CO), 36.8, 36.3, 32.3, 25.6 (CH<sub>2</sub> succinimide), 24.9.

ESI-LRMS: Calc'd for [C<sub>11</sub>H<sub>15</sub>NO<sub>4</sub> + Na]<sup>+</sup>: 248.0990; found: 248.1.



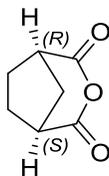
#### (1R,3S)-Cyclopentane-1,3-dicarboxylic acid (**2.45**)

Norbornene (1 g, 10.6 mmol), ruthenium trichloride (48 mg, 2.2 mol %) and potassium periodate (1 g, 4.1 eq) were dissolved in 3:2:2 H<sub>2</sub>O/ACN/EtOAc (70 mL) and stirred vigorously for 4 hours at room temperature. The reaction was extracted with ethyl acetate (3 x 50 mL), washed with brine (2 x 50 mL) and dried over MgSO<sub>4</sub>. Solvent was evaporated and the solid was recrystallized from chloroform to afford 1.2 g (71% yield) of **2.45**<sup>202</sup>.

<sup>1</sup>H NMR (300 MHz, DMSO-d<sub>6</sub>) δ 12.07 (s, 2H, COOH), 2.69 (ddq, *J* = 11.0, 5.9, 2.8 Hz, 2H, CH), 2.10 (dt, *J* = 12.8, 8.1 Hz, 1H), 2.00 – 1.63 (m, 5H).

<sup>13</sup>C NMR (75 MHz, DMSO-d<sub>6</sub>) δ 176.4 (2C, COOH), 43.2 (2C, CH), 32.8 (CH<sub>2</sub>), 28.9 (2C, CH<sub>2</sub>).

ESI-LRMS: Calc'd for [C<sub>7</sub>H<sub>10</sub>O<sub>4</sub> + Na]<sup>+</sup>: 181.0471; found: 181.4.



### 3-oxabicyclo[3.2.1]octane-2,4-dione (2.46)

Acetic anhydride (1 mL) was added to a solution of cyclopentane-1,3-dicarboxylic acid (600 mg, 1.2 mmol) stirring in anhydrous toluene (7 mL). The reaction was refluxed for 24 hours (115 °C), the solvent was evaporated and **2.46** was recrystallized in two crops (THF/Hexanes) to afford 0.46 g (85% yield) of off-white crystals.

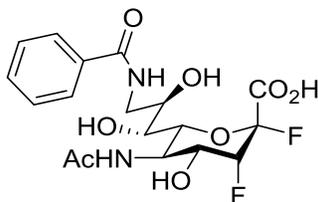
<sup>1</sup>H NMR (400 MHz, Acetone-d<sub>6</sub>) δ 3.23 (tdd, *J* = 4.1, 2.7, 0.9 Hz, 2H), 2.38 (d, *J* = 12.6 Hz, 1H), 2.22 (ddd, *J* = 8.3, 3.5, 2.1 Hz, 2H), 1.99 (dd, *J* = 8.0, 1.3 Hz, 2H), 1.84 (dt, *J* = 12.6, 4.1 Hz, 1H).

<sup>13</sup>C NMR (101 MHz, Acetone-d<sub>6</sub>) δ 170.2 (CO), 42.0 (2C, C-1, C-5), 30.8 (C-8), 26.2 (2C, C-6, C-7).

ESI-LRMS: Calc'd for [C<sub>7</sub>H<sub>8</sub>O<sub>3</sub> + Na]<sup>+</sup>: 163.0365; found: 163.3.

#### 7.1.4 General procedure for 9-amino-DFSA acylation

Activated ester (**2.37-2.44**, 1.5 eq) or anhydride (**2.46**, 1.3 eq) was added to a solution of **2.27** (0.1 mmol) in DMF (1 mL). Triethylamine (TEA, 30 μL, 2 eq) was added dropwise and the reaction stirred for 2-16 hours. Solvent was evaporated and the product purified by silica gel column chromatography (EtOAc/MeOH/H<sub>2</sub>O). Solid was dissolved in H<sub>2</sub>O, filtered (0.22 micron syringe filter) and lyophilized to afford a white solid product (46-70% yield).



**5-Acetamido-9-benzamido-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro-β-L-manno-non-2-  
ulopyranosylonic fluoride (2.47)**

**2.27** (30.6 mg, 0.093 mmol) and **2.37** (32.0 mg, 0.146 mmol) were dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30 μL) was added dropwise to the mixture and the reaction mixture was stirred for 3 hours at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (EtOAc/MeOH/H<sub>2</sub>O 17:2:1 – 10:2:1). Product was dissolved in H<sub>2</sub>O filtered through 0.22 μm Millex filter unit (0.22 μm) and lyophilized to afford 19.7 mg (49% yield) of **2.47** as a white solid.

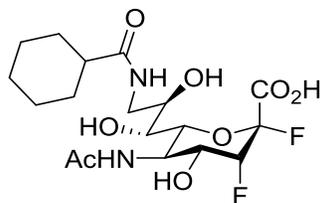
R<sub>f</sub> = 0.3 (7:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O) δ 7.77 (d, *J* = 7.0, 2H, H-ortho), 7.62 (t, *J* = 7.4 Hz, 1H, H-para), 7.53 (t, *J* = 7.3 Hz, 2H, H-meta), 5.21 (dt, *J*<sub>3,F3</sub> = 51.5 Hz, *J*<sub>3,F2</sub> = 2.7 Hz, *J*<sub>3,4</sub> = 2.7 Hz, 1H, H-3), 4.31 (td, *J*<sub>5,4</sub> = 10.5 Hz, *J*<sub>5,6</sub> = 10.5 Hz, *J*<sub>5,3F</sub> = 1.5 Hz, 1H, H-5), 4.18 (ddt, *J*<sub>4,3F</sub> = 28.2 Hz, *J*<sub>4,5</sub> = 10.7 Hz, *J*<sub>4,F2</sub> = 2.0 Hz, *J*<sub>4,3</sub> = 2.0 Hz, 1H, H-4), 4.04 (ddd, *J*<sub>8,7</sub> = 9.1 Hz, *J*<sub>8,9b</sub> = 7.2 Hz, *J*<sub>8,9a</sub> = 3.0 Hz, 1H, H-8), 3.85 (d, *J*<sub>6,5</sub> = 10.3 Hz, 1H, H-6), 3.76 (dd, *J*<sub>9a,9b</sub> = 14.1 Hz, *J*<sub>9a,8</sub> = 3.1 Hz, 1H, H-9a), 3.63 – 3.56 (m, 1H, H-9b), 3.53 (d, *J*<sub>7,8</sub> = 10.0 Hz, 1H, H-7), 2.00 (s, 3H, NHAc).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) 175.1 (CO), 171.6 (CO), 169.7 (dd, *J*<sub>C1,F2</sub> = 26.3 Hz, *J*<sub>C1,F3</sub> = 3.5 Hz, C-1), 133.7 (Bz), 132.2, 128.9 (2C, Bz), 127.2 (2C, Bz), 106.8 (dd, *J*<sub>C2,F2</sub> = 221.2 Hz, *J*<sub>C2,F3</sub> = 15.2 Hz, C-2), 89.1 (dd, *J*<sub>C3,F3</sub> = 183.8 Hz, *J*<sub>C3,F2</sub> = 18.5 Hz, C-3), 72.6 (d, *J*<sub>C6,F2</sub> = 4.3 Hz, C-6), 69.3 (d, *J*<sub>C4,F3</sub> = 18.2 Hz, C-4), 69.3, 69.1 (C-7, C-8), 47.1 (d, *J*<sub>C5,F3</sub> = 3.3 Hz, C-5), 43.1 (C-9), 22.1 (NHAc).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O) δ -121.48 (d, *J*<sub>F2,F3</sub> = 11.6 Hz, F-2 eq), -218.18 (d, *J*<sub>F3,F2</sub> = 11.6 Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>18</sub>H<sub>22</sub>F<sub>2</sub>N<sub>2</sub>O<sub>8</sub> + H]<sup>+</sup>: 433.1422; found: 433.1430.



**5-Acetamido-9-cyclohexanecarboxamido-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulopyranosylonic fluoride (2.48)**

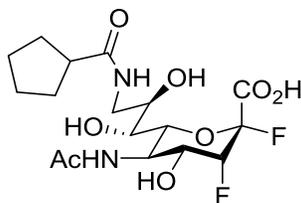
**2.27** (31.7 mg, 0.097 mmol) and **2.38** (32.7 mg, 0.145 mmol) was dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30  $\mu$ L) was added dropwise to the mixture and the reaction was stirred for 3 hours at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (EtOAc/MeOH/H<sub>2</sub>O 17:2:1 – 10:2:1). Product was dissolved in H<sub>2</sub>O filtered through 0.22  $\mu$ m Millex filter unit and lyophilized to afford 19.5 mg (46% yield) of **2.48** as a white solid. R<sub>f</sub> = 0.3 (7:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  5.20 (dt,  $J_{3,F3} = 51.5$  Hz,  $J_{3,F2} = 2.6$  Hz,  $J_{3,4} = 2.6$  Hz, 1H, H-3), 4.29 (t,  $J_{5,4} = 10.5$  Hz,  $J_{5,6} = 10.5$  Hz, 1H, H-5), 4.17 (dd,  $J_{4,F3} = 28.4$ ,  $J_{4,5} = 10.6$  Hz, 1H, H-4), 3.92 (ddd,  $J_{8,9b} = 9.3$  Hz,  $J_{8,7} = 6.2$  Hz,  $J_{8,9a} = 3.2$  Hz, 1H, H-8), 3.81 (d,  $J_{6,5} = 10.0$  Hz, 1H, H-6), 3.50 (dd,  $J_{9a,9b} = 14.2$  Hz,  $J_{9a,8} = 3.2$  Hz, 1H, H-9a), 3.46 – 3.37 (m, 2H, H-7, H-9b), 2.29 (td,  $J = 11.5, 3.3$  Hz, 1H), 2.04 (s, 3H, NHAc), 1.81 – 1.76 (m, 5H), 1.68 (d,  $J = 11.9$  Hz, 1H), 1.47 – 1.12 (m, 4H).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  181.1 (CO), 175.1 (CO), 169.8 (dd,  $J_{C1,F2} = 26.5$  Hz,  $J_{C1,F3} = 3.5$  Hz, C-1), 106.9 (dd,  $J_{C2,F2} = 219.5$  Hz,  $J_{C2,F3} = 15.1$  Hz, C-2), 89.2 (dd,  $J_{C3,F3} = 184.0$  Hz,  $J_{C3,F2} = 18.8$  Hz, C-3), 72.7 (d,  $J_{C6,F2} = 4.5$  Hz, C-6), 69.4 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 6.1$  Hz, C-4), 69.1, 68.9 (C-7, C-8), 47.2 (d,  $J_{C5,F3} = 3.3$  Hz, C-5), 45.4, 42.2 (C-9), 29.7, 29.3, 25.5, 25.4, 25.4, 22.3 (NHAc).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O)  $\delta$  -121.54 (d,  $J_{F2,F3} = 11.9$  Hz, F-2 eq), -218.24 (d,  $J_{F3,F2} = 11.6$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>18</sub>H<sub>28</sub>F<sub>2</sub>N<sub>2</sub>O<sub>8</sub> + H]<sup>+</sup>: 439.1892; found: 439.1895.



**5-Acetamido-9-cyclopentanecarboxamido-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-uloopyranosylonic fluoride (2.49)**

**2.27** (32.1 mg, 0.098 mmol) and **2.39** (33.1 mg, 0.157 mmol) were dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30  $\mu$ L) was added dropwise and the reaction mixture was stirred for 2 hours at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (EtOAc/MeOH/H<sub>2</sub>O 17:2:1 – 10:2:1). Product was dissolved in H<sub>2</sub>O filtered through 0.22  $\mu$ m Millex filter unit and lyophilized to afford 19.7 mg (48% yield) of **2.49** as a white solid.

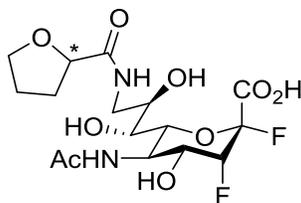
R<sub>f</sub> = 0.3 (7:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O)  $\delta$  5.20 (dt,  $J_{3,F3} = 51.6$  Hz,  $J_{3,F3} = 2.6$  Hz, 1H, H-3), 4.29 (t,  $J_{5,4} = 10.0$  Hz,  $J_{5,6} = 10.0$  Hz, 1H, H-5), 4.16 (ddt,  $J_{4,F3} = 28.2$  Hz,  $J_{4,5} = 10.7$  Hz,  $J_{4,3} = 2.0$  Hz, 1H, H-4), 3.91 (ddd,  $J_{8,9b} = 9.3$  Hz,  $J_{8,7} = 6.3$  Hz,  $J_{8,9a} = 3.1$  Hz, 1H, H-8), 3.81 (d,  $J_{6,5} = 10.4$  Hz, 1H, H-6), 3.52 (dd,  $J_{9a,9b} = 14.2$  Hz,  $J_{9a,8} = 3.1$  Hz, 1H, H-9a), 3.47 – 3.35 (m, 2H, H-7, H-9b), 2.82 – 2.61 (m, 1H), 2.04 (s, 3H, NHAc), 1.97 – 1.78 (m, 2H), 1.77 – 1.50 (m, 6H).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) 181.0 (CO), 175.0 (CO), 169.7 (dd,  $J_{C1,F2} = 26.5$  Hz,  $J_{C1,F3} = 3.7$  Hz), 106.8 (dd,  $J_{C2,F2} = 219.6$  Hz,  $J_{C2,F3} = 14.6$  Hz, C-2), 89.1 (dd,  $J_{C3,F3} = 184.0$  Hz,  $J_{C3,F2} = 18.6$  Hz, C-3), 72.5 (d,  $J_{C6,F2} = 4.5$  Hz, C-6), 69.3 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 6.1$  Hz, C-4), 69.0, 68.9 (C-7, C-8), 47.0 (d,  $J_{C5,F3} = 3.3$  Hz, C-5), 45.4, 42.3 (C-9), 30.5, 30.1, 25.7, 25.7, 22.2 (NHAc).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O)  $\delta$  -121.49 (d,  $J_{F2,F3} = 12.1$  Hz, F-2 eq), -218.19 (ddd,  $J_{F3,H3} = 51.6$  Hz,  $J_{F3,H4} = 28.3$ ,  $J_{F3,F2} = 11.7$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>17</sub>H<sub>26</sub>F<sub>2</sub>N<sub>2</sub>O<sub>8</sub> + H]<sup>+</sup>: 425.1735; found: 425.1740.



**5-Acetamido-9-[2'-tetrahydrofurancarboxamido]-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulopyranosylonic fluoride (2.50)**

**2.27** (29.8 mg, 0.091 mmol) and **2.40** (29.2 mg, 0.137 mmol) were dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30  $\mu$ L) was added dropwise and the reaction mixture was stirred for 2 hours at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (EtOAc/MeOH/H<sub>2</sub>O 17:2:1 – 10:2:1). Product was dissolved in H<sub>2</sub>O filtered through 0.22  $\mu$ m Millex filter unit and lyophilized to afford 25.2 mg (65% yield) of **2.50** as a white solid. \* **2.50** is a 1:1 mixture of diastereomers as determined by the <sup>13</sup>C NMR shifts.

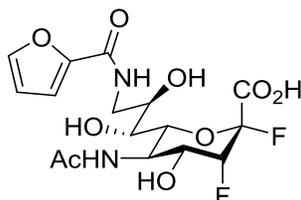
R<sub>f</sub> = 0.3 (5:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  5.09 (dt,  $J_{3,F3} = 51.5$  Hz,  $J_{3,F2} = 1.9$  Hz,  $J_{3,4} = 1.9$  Hz, 1H, H-3), 4.33 (ddd,  $J_{8,7} = 8.8$  Hz,  $J_{8,9b} = 5.0$  Hz,  $J_{8,9a} = 2.0$  Hz, 1H, H-8), 4.17 (t,  $J_{5,4} = 10.2$  Hz,  $J_{5,6} = 10.2$  Hz, 1H, H-5), 4.05 (dd,  $J_{4,F3} = 28.9$  Hz,  $J_{4,5} = 10.3$  Hz, 1H, H-4), 3.96 – 3.87 (m, 1H, THF), 3.82 (ddq,  $J = 9.9, 5.8, 3.3$  Hz, 2H, THF), 3.70 (d,  $J_{6,5} = 10.4$  Hz, 1H, H-6), 3.46 (ddd,  $J = 26.1$  Hz,  $J_{9a,9b} = 14.1$  Hz,  $J_{9a,8} = 3.2$  Hz, 1H, H-9a), 3.39 – 3.23 (m, 2H, H-7, H-9b), 2.27 – 2.16 (m, 1H, THF), 1.93 (s, 3H, NHAc), 1.90 – 1.73 (m, 3H, THF).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  176.6, 176.5 (2C, CO), 175.0, 174.9 (2C, CO), 169.5 (dd,  $J_{C1,F2} = 26.7$  Hz,  $J_{C1,F3} = 3.9$  Hz, C-1), 106.7 (dd,  $J_{C2,F2} = 219.6$  Hz,  $J_{C2,F3} = 14.8$  Hz, C-2), 89.0 (dd,  $J_{C3,F3} = 183.9$  Hz,  $J_{C3,F2} = 18.5$  Hz, C-3), 77.7, 77.7 (2C, C-1'), 72.5 (q,  $J = 3.0$  Hz, 2C, C-6), 69.8 (C-7), 69.3 – 68.9 (m, C-4), 69.3, 69.0 (2C, C-8), 68.8, 68.4 (2C, C-4') 46.9 (t,  $J = 3.5$  Hz, 2C, C-5), 42.0, 41.8 (2C, C-9), 30.4, 30.3 (2C, C-2'), 24.9, 24.9 (2C, C-3'), 22.1 (NHAc).

$^{19}\text{F}$  NMR (282 MHz,  $\text{D}_2\text{O}$ )  $\delta$  -121.50 (dd,  $J_{F_2,F_3} = 12.0, 5.3$  Hz, F-2 eq), -218.22 (dd,  $J_{F_2,F_3} = 11.5, 7.4$  Hz, F-3 ax).

ESI-HRMS: Calc'd for  $[\text{C}_{16}\text{H}_{24}\text{F}_2\text{N}_2\text{O}_9 + \text{H}]^+$ : 427.1528; found: 427.1527.



**5-Acetamido-9-[2'-furancarboxamido]-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulo-pyranosylonic fluoride (2.51)**

**2.27** (30.5 mg, 0.090 mmol) and **2.41** (29.3 mg, 0.140 mmol) were dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30  $\mu\text{L}$ ) was added dropwise and the reaction mixture was stirred for 2 hours at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (EtOAc/MeOH/ $\text{H}_2\text{O}$  17:2:1 – 10:2:1). Product was dissolved in  $\text{H}_2\text{O}$  filtered through 0.22  $\mu\text{m}$  Millix filter unit and lyophilized to afford X mg (61% yield) of **2.51** as a white solid.

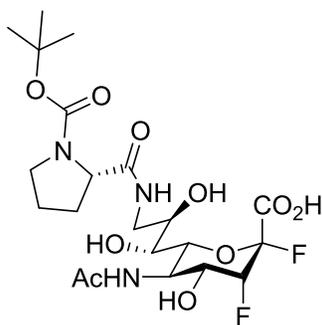
Rf = 0.3 (7:2:1 EtOAc/MeOH/ $\text{H}_2\text{O}$ )

$^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  7.67 (dd,  $J = 1.8, 0.8$  Hz, 1H, H-4'), 7.16 (dd,  $J = 3.6, 0.8$  Hz, 1H, H-2'), 6.62 (dd,  $J = 3.6, 1.8$  Hz, 1H, H-3'), 5.19 (dt,  $J_{3,F_3} = 51.5$  Hz,  $J_{3,F_2} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.29 (t,  $J_{5,4} = 10.2$  Hz,  $J_{5,6} = 10.2$  Hz, 1H, H-5), 4.16 (ddt,  $J_{4,F_3} = 28.3$  Hz,  $J_{4,5} = 10.7$  Hz,  $J_{4,3} = 2.0$  Hz, 1H, H-4), 4.00 (ddd,  $J_{8,9b} = 9.1$  Hz,  $J_{8,7} = 7.4$  Hz,  $J_{8,9a} = 3.1$  Hz, 1H, H-8), 3.83 (d,  $J_{6,5} = 10.4$  Hz, 1H, H-6), 3.73 (dd,  $J_{9a,9b} = 14.2$  Hz,  $J_{9a,8} = 3.1$  Hz, 1H, H-9a), 3.57 – 3.47 (m, 2H, H-7, H-9b), 2.00 (s, 3H, NHAc).

$^{13}\text{C}$  NMR (101 MHz,  $\text{D}_2\text{O}$ )  $\delta$  175.0 (CO), 169.5 (dd,  $J_{C_1,F_2} = 26.6$  Hz,  $J_{C_1,F_3} = 3.4$  Hz, C-1), 161.1 (CO), 146.3 (C-1'), 146.0 (C-4'), 115.2, 112.1 (C-2', C-3'), 106.7 (dd,  $J_{C_2,F_2} = 219.7$  Hz,  $J_{C_1,F_3} = 14.9$  Hz, C-2), 89.0 (dd,  $J_{C_3,F_3} = 183.9$  Hz,  $J_{C_3,F_2} = 18.5$  Hz, C-3), 72.5 (d,  $J_{C_6,F_2} = 4.4$  Hz, C-6), 69.3 – 68.9 (m, C-4), 69.2, 69.0 (C-7, C-8), 46.9 (d,  $J_{C_5,F_3} = 3.4$  Hz, C-5), 42.2 (C-9), 22.0 (NHAc).

$^{19}\text{F}$  NMR (282 MHz,  $\text{D}_2\text{O}$ )  $\delta$  -121.48 (d,  $J_{F_2,F_3} = 11.8$  Hz, F-2 eq), -218.16 (ddd,  $J_{F_3,H_3} = 51.3$  Hz,  $J_{F_3,H_4} = 28.1$  Hz,  $J_{F_3,F_2} = 11.7$  Hz, F-3 ax).

ESI-HRMS: Calc'd for  $[\text{C}_{16}\text{H}_{20}\text{F}_2\text{N}_2\text{O}_9 + \text{H}]^+$ : 423.1215; found: 423.1217.



**5-Acetamido-9-[N-boc-2'-pyrrolidamido]-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-uloopyranosylonic fluoride (2.52)**

**2.27** (30.1 mg, 0.092 mmol) and **2.42** (43.2 mg, 0.138 mmol) were dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30  $\mu\text{L}$ ) was added dropwise and the reaction mixture was stirred for 12 hours at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (EtOAc/MeOH/ $\text{H}_2\text{O}$  17:2:1 – 10:2:1). Product was dissolved in  $\text{H}_2\text{O}$  filtered through 0.22  $\mu\text{m}$  Millex filter unit and lyophilized to afford 35 mg (72% yield) of **2.52** as a white solid.

Interpretation of the NMR spectra for **2.52** was complicated by the presence of rotamers after Boc addition.

Rf = 0.5 (5:2:1 EtOAc/MeOH/ $\text{H}_2\text{O}$ )

$^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  5.11 (dt,  $J_{3,3F} = 51.5$  Hz,  $J_{3,2F} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.22 – 4.08 (m, 2H, H-5, H-7), 4.02 (ddt,  $J_{4,3F} = 28.3$  Hz,  $J_{4,5} = 10.6$  Hz,  $J_{4,3} = 2.0$  Hz, 1H, H-4), 3.83 (ddd,  $J_{8,9b} = 9.6$  Hz,  $J_{8,7} = 7.4$  Hz,  $J_{8,9a} = 2.8$  Hz, 1H, H-8), 3.70 (d,  $J_{6,5} = 10.9$  Hz, 1H, H-6), 3.46 (dd,  $J_{9a,9b} = 14.1$  Hz,  $J_{9a,8}$

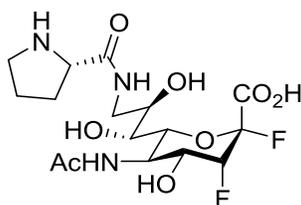
= 2.9 Hz, 1H, H-9a), 3.41 – 3.26 (m, 4H, H-7, H-9b, CH<sub>2</sub> H-4'), 2.18 (tt,  $J = 9.6, 6.4$  Hz, 1H, H-1'), 1.94 (s, 3H, *OMe*), 1.87 – 1.72 (m, 3H), 1.35 (s, 2H), 1.29 (s, 7H).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  176.5, 175.0, 156.0, 90.1, 89.9, 88.3, 88.1, 82.0, 72.6, 69.4, 69.3, 69.2, 69.1, 60.8, 60.5, 47.0, 46.9, 42.3, 31.3, 27.8, 27.7, 23.6, 22.2.

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  176.5 (CO), 175.0 (CO), 156.0 (NH CO), 89.11 (dd,  $J = 183.7, 18.8$  Hz, C-3), 82.0, 72.6, 69.5 – 69.0 (m, C-4), 69.2, 69.1, 60.8, 47.0, 46.9, 42.4, 31.3, 27.8 (Ac), 27.7 (Ac), 23.6 C(CH<sub>3</sub>)<sub>3</sub>, 22.2 (Ac).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O)  $\delta$  -121.62 (dd,  $J = 70.6, 11.8$  Hz, F-2), -218.28 (ddd,  $J = 51.3, 28.3, 11.5$  Hz, F-3).

ESI-LRMS: Calc'd for [C<sub>21</sub>H<sub>33</sub>F<sub>2</sub>N<sub>3</sub>O<sub>10</sub> + Na]<sup>+</sup>: 548.2032; found: 548.3.



**5-Acetamido-9-[2'-pyrrolidamido]-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulopyranosylonic fluoride (2.53)**

**2.52** was dissolved in trifluoroacetic acid (1 mL) and the reaction mixture was stirred for 2 hours at room temperature. Reaction was quenched with acetic acid and the solvent evaporated. The product was dissolved in H<sub>2</sub>O filtered through 0.22  $\mu$ m Millex filter unit and lyophilized to afford 27.4 mg (70% two-step yield) of **2.53** as a white solid.

R<sub>f</sub> = 0.3 (5:2:1 EtOAc/MeOH/H<sub>2</sub>O)

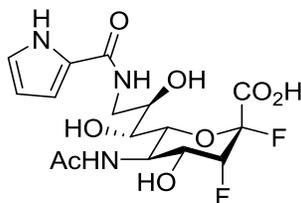
<sup>1</sup>H NMR (600 MHz, D<sub>2</sub>O)  $\delta$  5.21 (dt,  $J_{3,F3} = 51.1$  Hz,  $J_{3,F2} = 2.6$  Hz,  $J_{3,4} = 2.6$  Hz, 1H, H-3), 4.39 (dd,  $J = 8.6, 6.2$  Hz, 1H, proline), 4.27 (t,  $J_{5,4} = 10.5$  Hz,  $J_{5,6} = 10.5$  Hz, 1H, H-5), 4.18 (ddt,  $J_{4,F3} = 28.3$  Hz,  $J_{4,5} = 10.7$ ,  $J_{4,3} = 1.7$  Hz, 1H, H-4), 3.90 (ddd,  $J_{8,9b} = 9.7$  Hz,  $J_{8,7} = 7.2$  Hz,  $J_{8,9a} = 3.0$  Hz, 1H, H-8), 3.85 (d,  $J_{6,5}$

= 10.5 Hz, 1H, H-6), 3.61 (dd,  $J_{9a,9b} = 14.2$  Hz,  $J_{9a,8} = 3.1$  Hz, 1H, H-9a), 3.50 – 3.34 (m, 4H, H-7, H-9b, proline 2H), 2.54 – 2.39 (m, 1H, proline), 2.12 – 1.99 (m, 3H, proline), 2.03 (s, 3H, NHAc).

$^{13}\text{C}$  NMR (151 MHz,  $\text{D}_2\text{O}$ ) 174.2 (CO), 169.1 (CO), 167.8 (dd,  $J_{\text{C1},\text{F2}} = 27.5$  Hz,  $J_{\text{C1},\text{F3}} = 2.9$  Hz, C-1), 105.5 (dd,  $J_{\text{C2},\text{F2}} = 220.1$ ,  $J_{\text{C2},\text{F3}} = 15.0$  Hz, C-2), 87.7 (dd,  $J_{\text{C3},\text{F3}} = 184.3$  Hz,  $J_{\text{C3},\text{F2}} = 18.3$  Hz, C-3), 71.7 (d,  $J_{\text{C6},\text{F2}} = 4.0$  Hz, C-6), 68.2 (C-7), 68.0 (dd,  $J_{\text{C4},\text{F3}} = 17.9$  Hz,  $J_{\text{C4},\text{F2}} = 5.7$  Hz, C-4), 67.7 (C-8), 59.1, 46.0 (d,  $J_{\text{C5},\text{F3}} = 3.5$  Hz, C-5), 45.7, 41.9 (C-9), 29.1, 23.0, 21.3 (NHAc).

$^{19}\text{F}$  NMR (377 MHz,  $\text{D}_2\text{O}$ )  $\delta$  -121.24 (d,  $J_{\text{F2},\text{F3}} = 12.0$  Hz, F-2 eq), -217.94 (ddd,  $J_{\text{F3},\text{H3}} = 51.2$  Hz,  $J_{\text{F3},\text{H4}} = 27.9$  Hz,  $J_{\text{F3},\text{F2}} = 11.6$  Hz, F-3 ax).

ESI-HRMS: Calc'd for  $[\text{C}_{16}\text{H}_{25}\text{F}_2\text{N}_3\text{O}_8 + \text{Na}]^+$ : 448.1508; found: 448.1513.



#### 5-Acetamido-9-[2'-pyrrolecarboxamido]-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulopyranosylonic fluoride (**2.54**)

**2.27** (29.6 mg, 0.090 mmol) and **2.43** (29.5 mg, 0.139 mmol) was dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30  $\mu\text{L}$ ) was added dropwise and the reaction mixture was stirred overnight at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (EtOAc/MeOH/ $\text{H}_2\text{O}$  17:2:1 – 10:2:1). Product was dissolved in  $\text{H}_2\text{O}$  filtered through 0.22  $\mu\text{m}$  Millex filter unit and lyophilized to afford 25.7 mg (68% yield) of **2.54** as a white solid.

Rf = 0.3 (7:2:1 EtOAc/MeOH/ $\text{H}_2\text{O}$ )

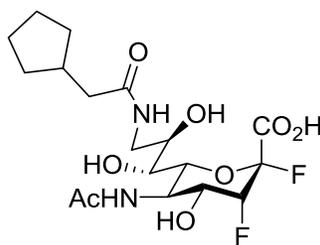
$^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  7.09 (dd,  $J = 2.6, 1.4$  Hz, 1H, H-4'), 6.88 (dd,  $J = 3.7, 1.4$  Hz, 1H, H-2'), 6.32 (dd,  $J = 3.8, 2.6$  Hz, 1H, H-3'), 5.22 (dt,  $J_{3,\text{F3}} = 51.4$  Hz,  $J_{3,\text{F2}} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.31 (td,  $J_{5,4} = 10.6$  Hz,  $J_{5,6} = 10.6$  Hz,  $J_{5,\text{F3}} = 1.3$  Hz, 1H, H-5), 4.18 (ddt,  $J_{4,\text{F3}} = 28.4$  Hz,  $J_{4,5} = 10.9$  Hz,  $J_{4,\text{F2}} =$

2.1 Hz,  $J_{4,3} = 2.1$  Hz, 1H, C-4), 4.00 (ddd,  $J_{8,9b} = 9.6$  Hz,  $J_{8,7} = 7.0$  Hz,  $J_{8,9a} = 3.0$  Hz, 1H, H-8), 3.85 (dd,  $J_{6,5} = 10.4$  Hz,  $J_{6,4} = 1.4$  Hz, 1H, H-6), 3.71 (dd,  $J_{9a,9b} = 14.2$  Hz,  $J_{9a,8} = 3.1$  Hz, 1H, H-9a), 3.61 – 3.46 (m, 2H, H-7, H-9b), 1.99 (s, 3H, NHAc).

$^{13}\text{C}$  NMR (101 MHz,  $\text{D}_2\text{O}$ )  $\delta$  175.1 (CO), 169.7 (dd,  $J_{\text{C}1,\text{F}2} = 26.6$  Hz,  $J_{\text{C}1,\text{F}3} = 3.5$  Hz, C-1), 163.7 (CO), 124.8, 123.3, 111.6, 109.7, 106.8 (dd,  $J_{\text{C}2,\text{F}2} = 219.2$  Hz,  $J_{\text{C}2,\text{F}3} = 14.5$  Hz, C-2), 89.1 (dd,  $J_{\text{C}3,\text{F}3} = 183.7$  Hz,  $J_{\text{C}3,\text{F}2} = 18.5$  Hz, C-3), 72.6 (d,  $J_{\text{C}6,\text{F}2} = 4.4$  Hz, C-6), 69.3 (d,  $J_{\text{C}4,\text{F}3} = 18.2$  Hz, C-4), 69.3, 69.1 (C-7,C-8), 47.0 (d,  $J_{\text{C}5,\text{F}3} = 3.4$  Hz, C-5), 42.3 (C-9), 22.1 (NHAc).

$^{19}\text{F}$  NMR (282 MHz,  $\text{D}_2\text{O}$ )  $\delta$  -121.45 (d,  $J_{\text{F}2,\text{F}3} = 11.7$  Hz, F-2 eq), -218.14 (ddd,  $J_{\text{F}3,\text{H}3} = 51.7$  Hz,  $J_{\text{F}3,\text{H}4} = 28.1$  Hz,  $J_{\text{F}3,\text{F}2} = 11.9$  Hz, F-3 ax).

ESI-HRMS: Calc'd for  $[\text{C}_{16}\text{H}_{21}\text{F}_2\text{N}_3\text{O}_8 + \text{H}]^+$ : 422.1375; found: 422.1379.



**5-Acetamido-9-cyclopentylacetamido-2,3,5,9-tetra-deoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulopyranosylonic fluoride (2.55)**

**2.27** (29.5 mg, 0.090 mmol) and **2.44** (32.3 mg, 0.144 mmol) was dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30  $\mu\text{L}$ ) was added dropwise and the reaction mixture was stirred for 2 hours at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (15:2:1 EtOAc/MeOH/ $\text{H}_2\text{O}$ ). Product was dissolved in  $\text{H}_2\text{O}$  filtered through 0.22  $\mu\text{m}$  Millex filter unit and lyophilized to afford 20.0 mg (51% yield) of **2.55** as a white solid.

$R_f = 0.3$  (7:2:1 EtOAc/MeOH/ $\text{H}_2\text{O}$ )

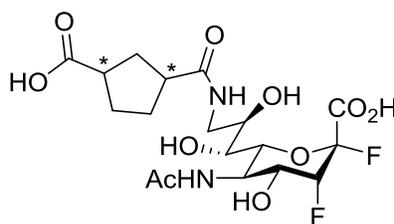
$^1\text{H}$  NMR (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  5.20 (dt,  $J_{3,\text{F}3} = 51.5$  Hz,  $J_{3,\text{F}2} = 2.6$  Hz,  $J_{3,4} = 2.6$  Hz, 1H, H-3), 4.29 (t,  $J_{5,4} = 10.5$  Hz,  $J_{5,6} = 10.5$  Hz, 1H, H-5), 4.15 (ddt,  $J_{4,\text{F}3} = 28.4$  Hz,  $J_{4,5} = 10.6$ ,  $J_{4,3} = 1.9$  Hz, 1H, H-4), 3.92 (ddd,

$J_{8,9b} = 9.4$  Hz,  $J_{8,7} = 6.4$ ,  $J_{8,9a} = 3.0$  Hz, 1H, H-8), 3.81 (d,  $J_{6,5} = 10.5$  Hz, 1H, H-6), 3.52 (dd,  $J_{9a,9b} = 14.2$ ,  $J_{9a,8} = 3.1$  Hz, 1H, H-9a), 3.48 – 3.34 (m, 2H, H-7, H-9b), 2.27 (d,  $J = 8.2$  Hz, 2H), 2.24 – 2.09 (m, 1H), 2.04 (s, 3H, NHAc), 1.82 – 1.69 (m, 2H), 1.67 – 1.48 (m, 4H), 1.25 – 1.09 (m, 2H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{D}_2\text{O}$ )  $\delta$  177.5 (CO), 175.0 (CO), 169.7 (dd,  $J_{C1,F2} = 26.5$  Hz,  $J_{C1,F3} = 3.3$  Hz, C-1), 106.8 (dd,  $J_{C2,F2} = 219.6$  Hz,  $J_{C2,F3} = 14.9$  Hz, C-2), 89.1 (dd,  $J_{C3,F3} = 183.9$  Hz,  $J_{C3,F2} = 18.7$  Hz, C-3), 72.6 (d,  $J_{C6,F2} = 4.5$  Hz, C-6), 69.3 (dd,  $J_{C4,F3} = 17.9$  Hz,  $J_{C4,F2} = 6.2$  Hz, C-4), 69.0, 68.9 (C-7, C-8), 47.02 (d,  $J_{C5,F3} = 3.4$  Hz, C-5), 42.2 (C-9), 41.9, 37.2, 31.9, 31.8, 24.6, 22.2 (NHAc).

$^{19}\text{F}$  NMR (377 MHz,  $\text{D}_2\text{O}$ )  $\delta$  -120.86 (d,  $J_{F2,F3} = 11.8$  Hz, F-2 eq), -217.56 (ddd,  $J_{F3,H3} = 51.6$  Hz,  $J_{F3,H4} = 28.5$  Hz,  $J_{F3,F2} = 11.7$  Hz, F-3 ax).

ESI-HRMS: Calc'd for  $[\text{C}_{18}\text{H}_{28}\text{F}_2\text{N}_2\text{O}_8 + \text{H}]^+$ : 439.1892; found: 439.1894.



**5-Acetamido-9-[3'-carboxy-cyclopentamido]-2,3,5,9-tetraoxy-3-fluoro-D-erythro- $\beta$ -L-manno-non-2-ulopyranosylonic fluoride (2.56)**

**2.27** (31.0 mg, 0.095 mmol) and **2.46** (16.5 mg, 0.118 mmol) was dissolved in dimethylformamide (DMF, 1 mL). Triethylamine (30  $\mu\text{L}$ ) was added dropwise and the reaction mixture was stirred for 2 hours at room temperature. DMF was co-evaporated with toluene and the product was purified by column chromatography (7:2:1 EtOAc/MeOH/ $\text{H}_2\text{O}$ ). Product was dissolved in  $\text{H}_2\text{O}$  filtered through 0.22  $\mu\text{m}$  Millex filter unit and lyophilized to afford 30.2 mg (68% yield) of **2.56** as a white solid. \* **2.56** is a 1:1 mixture of diastereomers ( $R,S$  or  $S,R$ ) as determined by the  $^{13}\text{C}$  NMR shifts.

Rf = 0.3 (7:2:1 EtOAc/MeOH/ $\text{H}_2\text{O}$ )

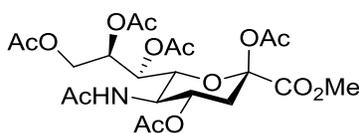
$^1\text{H}$  NMR (300 MHz,  $\text{D}_2\text{O}$ )  $\delta$  5.13 (dt,  $J_{3,F3} = 51.5$  Hz,  $J_{3,F2} = 2.6$  Hz,  $J_{3,4} = 2.6$  Hz, 1H, H-3), 4.21 (t,  $J_{5,4} = 9.9$  Hz,  $J_{5,6} = 9.9$  Hz, 1H, H-5), 4.08 (ddt,  $J_{4,F3} = 28.2$  Hz,  $J_{4,5} = 10.4$  Hz,  $J_{4,3} = 1.8$  Hz, 1H, H-4), 3.84 (ddd,  $J_{8,9b} = 9.5$  Hz,  $J_{8,7} = 6.6$  Hz,  $J_{8,9a} = 3.1$  Hz, 1H, H-8), 3.74 (d,  $J_{6,5} = 10.0$  Hz, 1H, H-6), 3.47 (dt,  $J_{9a,9b} = 14.2$  Hz,  $J_{9a,8} = 3.1$  Hz, 1H, H-9a), 3.41 – 3.29 (m, 2H, H-7, H-9b), 2.79 (ddt,  $J = 15.3, 9.9, 7.7$  Hz, 2H), 2.16 (dtd,  $J = 11.4, 7.6, 3.5$  Hz, 1H), 1.96 (s, 3H, NHAc), 1.93 – 1.65 (m, 5H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{D}_2\text{O}$ ) 181.1 (CO), 179.0 (CO), 174.9 (CO), 169.5 (dd,  $J_{C1,F2} = 26.5$  Hz,  $J_{C1,F3} = 3.4$  Hz, C-1), 106.7 (dd,  $J_{C2,F2} = 219.9$  Hz,  $J_{C2,F3} = 15.2$  Hz, C-2), 89.0 (dd,  $J_{C3,F3} = 183.6$  Hz,  $J_{C3,F2} = 18.5$  Hz, C-3), 72.4 (d,  $J_{C6,F2} = 4.4$  Hz, C-6), 69.1 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 5.9$  Hz, C-4), 68.8, 68.8 (C-7, C-8), 46.9 (d,  $J_{C5,F3} = 3.4$  Hz, C-5), 45.2, 45.2, 44.2, 44.2 (4C, C-1', C-3'), 42.2 (C-9), 33.9, 33.6 (2C, C-2'), 29.5, 29.2, 29.0, 29.0 (4C, C-4', C-5'), 22.0 (NHAc).

$^{19}\text{F}$  NMR (377 MHz,  $\text{D}_2\text{O}$ )  $\delta$  -120.89 (d,  $J_{F2,F3} = 11.7$  Hz, F-2 eq), -217.59 (ddd,  $J_{F3,H3} = 51.6$  Hz,  $J_{F3,H4} = 28.5$  Hz,  $J_{F3,F2} = 11.7$  Hz, F-3 ax).

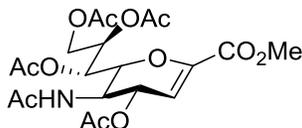
ESI-HRMS: Calc'd for  $[\text{C}_{18}\text{H}_{26}\text{F}_2\text{N}_2\text{O}_{10} + \text{H}]^+$ : 469.1634; found: 469.1636.

### 7.1.5 Synthesis of 5-amido-DFSA



#### Methyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-*D*-glycero- $\beta$ -*D*-galacto- non-2-ulopyranosonate (2.57)

Neu5Ac (10 g) was dissolved in methanol (500 mL) and Amberlite IR-120H ion exchange resin was added. After stirring overnight, the solution was filtered, the resin washed with methanol and solvent evaporated to afford a white foam. Solid was dissolved in pyridine (35 mL) to which acetic anhydride (70 mL) was added and the reaction stirred for 16 hours. Solvents were evaporated and the product extracted with EtOAc. Organic phase was washed with 1 M HCl,  $\text{NaHCO}_3$  and brine then dried over  $\text{MgSO}_4$ . Solvent was removed to afford 14.5 g of **2.57** as a white foam (84% yield).



**Methyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-2,6-anhydro-D-glycero-D-galacto-non-2-enonate (2.58)**

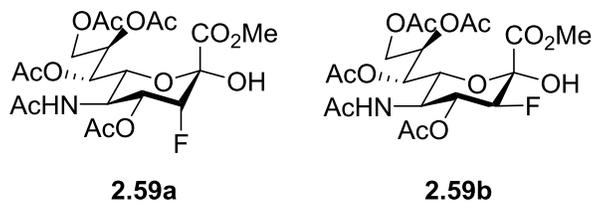
trimethylsilyl trifluoromethanesulfonate (9.5 mL, 52.5 mmol) was added dropwise to 2,4,7,8,9-penta-*O*-acetyl-*N*-acetylneuraminic acid methyl ester (**2.57**; 11.5 g, 21.6 mmol) stirring in EtOAc (125 mL) at 0 °C. The mixture was warmed to room temperature and stirred for 4 hours then quenched with triethylamine (18 mL) and H<sub>2</sub>O (150 mL). Aqueous layer was separated and extracted with EtOAc (2 x 150 mL). Combined organic layers were washed with brine (2 x 150 mL) and dried over MgSO<sub>4</sub>. Solvent was removed and silica gel column chromatography (9:1 to 3:1 DCM/acetone) of the mixture afforded 8.2 g of **2.58** (80% yield) as a white foam.

R<sub>f</sub> = 0.3 (EtOAc)

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.15 (d, *J*<sub>NH,5</sub> = 8.7 Hz, 1H, NH), 5.93 (d, *J*<sub>3,4</sub> = 3.0 Hz, 1H, H-3), 5.51 – 5.42 (m, 2H, H-4, H-7), 5.30 (ddd, *J*<sub>8,9b</sub> = 7.5 Hz, *J*<sub>8,7</sub> = 4.5 Hz, *J*<sub>8,9a</sub> = 3.1 Hz, 1H, H-8), 4.60 (dd, *J*<sub>9a,9b</sub> = 12.3 Hz, *J*<sub>9a,8</sub> = 3.1 Hz, 1H, H-9a), 4.40 – 4.32 (m, 2H, H-5, H-6), 4.14 (dd, *J*<sub>9b,9a</sub> = 12.3 Hz, *J*<sub>9b,8</sub> = 7.3 Hz, 1H, H-9b), 3.75 (s, 3H, OMe), 2.07 (s, 3H, OAc), 2.02 (s, 3H, OAc), 2.01 (s, 3H, OAc), 2.00 (s, 3H, OAc), 1.87 (s, 3H, NHAc).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.9 (CO), 170.7 (CO), 170.3 (CO), 170.2 (CO), 170.2 (CO), 161.7 (NHAc), 145.1 (C-2), 108.1 (C-3), 76.7 (C-6), 70.9 (C-8), 68.2 (C-7), 67.7 (C-4), 62.1 (C-9), 52.6 (OMe), 46.5 (C-5), 23.1 (NHAc), 20.9 (OAc), 20.9 (OAc), 20.8 (OAc), 20.8 (OAc).

ESI-LRMS: Calc'd for [C<sub>20</sub>H<sub>27</sub>NO<sub>12</sub> + Na]<sup>+</sup>: 496.1425; found: 496.3.



**Methyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-3-fluoro-*D*-erythro- $\alpha$ -*L*-manno- non-2-ulopyranosonate (2.59a) and gluco-non-2-ulopyranosonate (2.59b)**

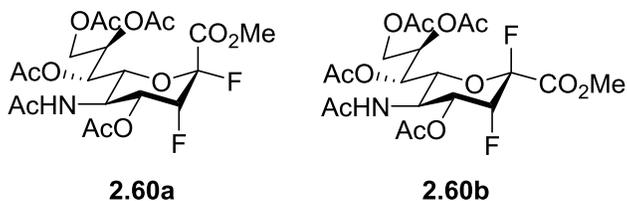
Selectfluor (15.73 g, 0.044 mol) and **2.58** (5.25 g, 0.011 mol) were dissolved in a 9:1 mixture of acetonitrile and water (150 mL) and the solution was stirred at 60 °C for 5 hours. EtOAc (125 mL) was added to the slurry and washed with NaHCO<sub>3</sub> (2 x 50 mL). Aqueous solution was extracted with EtOAc (2 x 75 mL), combined organic layers washed with brine (2 x 100 mL) and dried over MgSO<sub>4</sub>. Products were purified by silica gel column chromatography (5:1 to 3:1 DCM/acetone) to afford 1.95 g (35% yield) **2.59a** and 1.75 g (31% yield) **2.59b**.

R<sub>f</sub> = 0.4 (3:1 DCM/acetone)

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  8.14 (s, 1H, OH), 6.06 (d,  $J_{NH,5}$  = 8.4 Hz, 1H, NH), 5.71 (ddd,  $J_{4,F3}$  = 28.1 Hz,  $J_{4,5}$  = 10.8 Hz,  $J_{4,3}$  = 2.6 Hz, 1H, H-4), 5.37 (dd,  $J_{7,8}$  = 4.3 Hz,  $J_{7,6}$  = 1.8 Hz, 1H, H-7), 5.33 – 5.08 (m, 2H, H-3, H-8), 4.78 (dd,  $J_{9a,9b}$  = 12.5 Hz,  $J_{9a,8}$  = 2.1 Hz, 1H, H-9a), 4.46 (d,  $J_{6,5}$  = 10.2 Hz, 1H, H-6), 4.23 (dd,  $J_{9b,9a}$  = 12.5 Hz,  $J_{9b,8}$  = 7.5 Hz, 1H, H-9b), 3.89 – 3.73 (m, 4H, OMe, H-5), 2.18 (s, 3H, OAc), 2.09 (s, 3H, OAc), 2.04 (s, 3H, OAc), 2.03 (s, 3H, OAc), 1.95 (s, 3H, NHAc).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)  $\delta$  -206.37 (dd,  $J_{F3,H3}$  = 47.5 Hz,  $J_{F3,H4}$  = 27.0 Hz, F-3 ax).

ESI-LRMS: Calc'd for [C<sub>20</sub>H<sub>28</sub>FNO<sub>13</sub> + Na]<sup>+</sup>: 532.1437; found: 532.3.



**Methyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-2,3,5-trideoxy-3-fluoro-*D*-erythro- $\beta$ -*L*-manno-non-2-ulopyranosylonate fluoride (2.60a) and - $\alpha$ -*L*-manno-non-2-ulopyranosylonate fluoride (2.60b)**

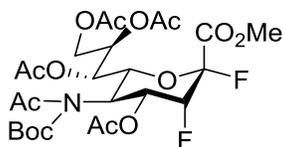
**2.59a** (0.50 g, 0.98 mmol) was dissolved in 20 mL of DCM and cooled to -40 °C. Diethylaminosulfur trifluoride (DAST, 145  $\mu$ L, 1.1 eq) was added dropwise over 1 hour. Reaction was left to stir for 30 mins, gradually warming to -10 °C. Reaction was quenched with methanol, diluted with DCM and washed with NaHCO<sub>3</sub> and brine. Organic phase was dried over MgSO<sub>4</sub>, solvent evaporated and silica gel column chromatography (EtOAc) isolated a mixture of **2.60a** and **2.60b**. Semi-preparative silica gel HPLC (5:1 EtOAc/hexanes) separated the isomers to afford 257 mg of **2.60a** (51% yield) and 86 mg of **2.60b** (17% yield) as white foams.

R<sub>f</sub> = 0.4 (EtOAc/hexanes)

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  5.51 – 5.23 (m, 4H, NH, H-4, H-7, H-8), 5.13 (dt,  $J_{3,F3} = 50.8$  Hz,  $J_{3,F2} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.30 (dd,  $J_{6,5} = 10.8$  Hz,  $J_{6,7} = 1.7$  Hz, 1H, H-6), 4.15 (q,  $J_{5,4} = 10.2$  Hz,  $J_{5,6} = 10.2$  Hz,  $J_{5,NH} = 10.2$  Hz, 1H, H-5), 3.92 (s, 3H, OMe), 3.68 (dd,  $J_{9a,9b} = 13.6$  Hz,  $J_{9a,8} = 2.8$  Hz, 1H, H-9a), 3.34 (dd,  $J_{9b,9a} = 13.6$  Hz,  $J_{9b,8} = 6.0$  Hz, 1H, H-9b), 2.19 (s, 3H, OAc), 2.14 (s, 3H, OAc), 2.12 (s, 3H, OAc), 1.93 (s, 3H, OAc), 1.62 (s, 3H, NHAc).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)  $\delta$  -123.68 (d,  $J_{F2,F3} = 11.5$  Hz, F-2 eq), -217.25 (d,  $J_{F3,F2} = 11.3$  Hz, F-3 ax).

ESI-LRMS: Calc'd for [C<sub>20</sub>H<sub>27</sub>F<sub>2</sub>NO<sub>12</sub> + Na]<sup>+</sup>: 534.1393; found: 534.3.



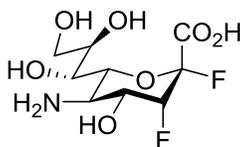
**Methyl 5-(*N*-(tert-butoxycarbonyl)acetamido)-4,7,8,9-tetra-*O*-acetyl-2,3,5-trideoxy-3-fluoro-*D*-erythro- $\beta$ -*L*-manno-non-2-ulopyranosylonate fluoride (2.61)**

4-Dimethylaminopyridine (DMAP, 10 mg, 0.082 mmol) was added to a solution of di-*tert*-butyl dicarbonate (550 mg, 2.52 mmol) and **2.60a** (257 mg, 0.50 mmol) in anhydrous THF (1 mL). The reaction was stirred overnight at room temperature, following completion the solution was diluted with DCM and washed with NaHCO<sub>3</sub> and brine. Organic layer was dried over MgSO<sub>4</sub>, evaporated and product isolated by silica gel column chromatography (7:1 to 5:1 Hexanes/EtOAc) to afford 280 mg (92% yield) of **2.61** as a clear oil. NMR analysis was complicated by the presence of rotamers following Boc addition. Given the presence of a single spot on TLC (R<sub>f</sub> = 0.4; 9:1 Hexanes/EtOAc) product was used directly in the following steps.

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  5.91 (ddd, *J* = 16.2, 9.4, 6.6 Hz, 1H), 5.42 – 5.21 (m, 2H), 5.20 – 5.03 (m, 2H), 4.89 – 4.57 (m, 1H), 4.31 (dd, *J* = 12.5, 2.6 Hz, 1H), 4.14 – 3.97 (m, 1H), 3.87 (d, *J* = 7.3 Hz, 3H), 2.37 (d, *J* = 6.1 Hz, 3H), 2.12 – 1.89 (m, 9H), 1.59 (d, *J* = 13.4 Hz, 5H), 1.54 – 1.42 (m, 3H).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)  $\delta$  -114.38 (t, *J* = 12.5 Hz), -197.37 (ddd, *J* = 47.4, 16.3, 12.7 Hz).

ESI-LRMS: Calc'd for [C<sub>25</sub>H<sub>35</sub>F<sub>2</sub>NO<sub>14</sub> + Na]<sup>+</sup>: 634.1918; found: 634.3.



**5-Amino-2,3,5-trideoxy-2,3-difluoro-*D*-erythro- $\beta$ -*L*-manno-2-nonulopyranosonic acid (2.62)**

Sodium (43 mg) was dissolved in MeOH (6 mL) and 4 mL was added slowly to a cooled solution of **2.61** (280 mg, 0.46 mmol) in MeOH (15 mL). The reaction mixture was stirred at 0 °C for 2 hours until TLC

showed complete removal of acetates (9:2:1 EtOAc/MeOH/H<sub>2</sub>O; R<sub>f</sub> = 0.8). 0.1 M NaOH (2 mL) was added slowly and the reaction was stirred for 1 hour. The solution was neutralized with Amberlite IR120 H resin, filtered and evaporated to yield 140 mg of white solid. Solid was dissolved in neat trifluoroacetic acid (6 mL) and stirred for 2 hours, the solvent was evaporated and the product dissolved in H<sub>2</sub>O, filtered (0.22 micron Millex syringe filter) and lyophilized to afford 130 mg (quantitative) of **2.62** as a white solid.

R<sub>f</sub> = 0.3 (5:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 5.20 (dt,  $J_{3,F3} = 50.5$  Hz,  $J_{3,F2} = 2.4$  Hz,  $J_{3,4} = 2.4$  Hz, 1H, H-3), 4.35 (ddt,  $J_{4,F3} = 27.5$  Hz,  $J_{4,5} = 10.7$  Hz,  $J_{4,3} = 1.9$  Hz, 1H, H-4), 4.17 (d,  $J_{6,5} = 10.4$  Hz, 1H, H-6), 3.88 (ddd,  $J_{8,7} = 8.5$  Hz,  $J_{8,9b} = 5.5$  Hz,  $J_{8,9a} = 2.5$  Hz, 1H, H-8), 3.82 (dd,  $J_{9a,9b} = 12.1$  Hz,  $J_{9a,8} = 2.6$  Hz, 1H, H-9a), 3.74 (d,  $J_{7,8} = 8.9$  Hz, 1H, H-7), 3.70 – 3.61 (m, 2H, H-5, H-9b).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) δ 168.1 (dd,  $J_{C1,F2} = 28.0$  Hz,  $J_{C1,F3} = 3.7$  Hz, C-1), 106.0 (dd,  $J_{C2,F2} = 222.3$  Hz,  $J_{C2,F3} = 15.4$  Hz, C-2), 88.0 (dd,  $J_{C3,F3} = 186.1$  Hz,  $J_{C3,F2} = 18.3$  Hz, C-3), 71.7 (d,  $J_{C6,F2} = 4.6$  Hz, C-6), 70.2, 67.6 (C-7, C-8), 67.6 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 5.4$  Hz, C-4), 62.8 (C-9), 47.7 (d,  $J_{C5,F3} = 3.8$  Hz, C-5).

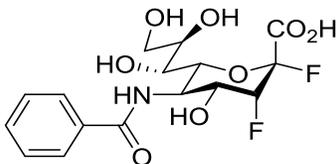
<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O) δ -75.95 (s, 3F, CF<sub>3</sub>CO<sub>2</sub>H), -122.40 (dd,  $J_{F2,F3} = 11.3$  Hz,  $J_{F2,H3} = 2.3$  Hz, F-2 eq), -218.94 (ddd,  $J_{F3,H3} = 50.5$  Hz,  $J_{F3,H4} = 27.6$  Hz,  $J_{F3,F2} = 11.3$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>9</sub>H<sub>15</sub>F<sub>2</sub>NO<sub>7</sub> + H]<sup>+</sup>: 288.0896; found: 288.0894.

### 7.1.6 General procedure for 5-amino-DFSA acylation

Commercial acyl chlorides (2 eq.) were added dropwise to a solution of **2.62** (0.035 mmol) and diisopropylethylamine (DIPEA, 35 μL, 6 eq.) in anhydrous acetonitrile (1 mL). The reaction mixture was stirred for 2 hours, quenched with water (0.5 mL), neutralized with Amberlite IR120-H<sup>+</sup> resin, filtered and the solvent evaporated. Silica gel column chromatography (EtOAc/MeOH/H<sub>2</sub>O) afforded white solids

which were dissolved in H<sub>2</sub>O, filtered (0.22 μm Millex syringe filter) and lyophilized to afford the respective 5-amido-DFSA products **2.63-2.69** (51-93% yield).



**5-Benzamido-2,3,5-trideoxy-2,3-difluoro-D-erythro-β-L-manno-2-nonulopyranosonic acid (2.63)**

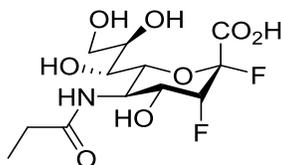
**2.62** (11 mg, 0.038 mmol) and NaHCO<sub>3</sub> (40 mg) was dissolved in a 1:1 acetonitrile/water (0.4 mL) solution, cooled to 0 °C for 15 minutes and benzoyl chloride (30 μL in 80 μL ACN, 5 eq.) was added dropwise and stirred for 3 hours. After neutralization with acetic acid, solvent was removed and partially purified by C18 Sep-Pak. Silica gel column chromatography (9:2:1 EtOAc/MeOH/H<sub>2</sub>O) and lyophilization afforded 12 mg (81% yield) of **2.63** as a white solid.

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ 7.81 (dd, *J* = 8.4, 1.4 Hz, 2H), 7.70 – 7.62 (m, 1H), 7.55 (dd, *J* = 8.4, 7.1 Hz, 2H), 5.29 (dt, *J*<sub>3,F3</sub> = 51.5 Hz, *J*<sub>3,F2</sub> = 2.7 Hz, *J*<sub>3,4</sub> = 2.7 Hz, 1H, H-3), 4.56 (t, *J*<sub>5,4</sub> = 10.6 Hz, *J*<sub>5,6</sub> = 10.6 Hz, 1H, H-5), 4.35 (ddt, *J*<sub>4,F3</sub> = 28.4 Hz, *J*<sub>4,5</sub> = 10.6 Hz, *J*<sub>4,3</sub> = 1.6 Hz, 1H, H-4), 4.00 (d, *J*<sub>6,5</sub> = 10.3 Hz, 1H, H-6), 3.96 – 3.85 (m, 2H, H-8, H-9a), 3.68 – 3.60 (m, 2H, H-7, H9b).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) δ 172.1 (NH CO), 169.8 (dd, *J*<sub>C1,F2</sub> = 26.6 Hz, *J*<sub>C1,F3</sub> = 3.5 Hz, C-1), 133.5 (C-1'), 132.7 (C-4'), 129.0 (2C, C-3', C-5'), 127.4 (2C, C-2', C-6'), 107.0 (dd, *J*<sub>C2,F2</sub> = 219.6 Hz, *J*<sub>C2,F3</sub> = 14.7 Hz, C-2), 89.4 (dd, *J*<sub>C3,F3</sub> = 183.7 Hz, *J*<sub>C3,F2</sub> = 18.5 Hz, C-3), 72.9 (d, *J*<sub>C6,F2</sub> = 4.2 Hz, H-6), 70.9 (C-7/C-8), 69.3 (dd, *J*<sub>C4,F3</sub> = 17.8 Hz, *J*<sub>C4,F2</sub> = 6.0 Hz, C-4), 68.3 (C-7/C-8), 63.2 (C-9), 47.7 (d, *J*<sub>C5,F3</sub> = 3.4 Hz, C-5).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O) δ -121.52 (d, *J*<sub>F2,F3</sub> = 11.7 Hz, F-2 eq), -218.27 (d, *J*<sub>F3,F2</sub> = 11.5 Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>16</sub>H<sub>19</sub>F<sub>2</sub>NO<sub>8</sub> + H]<sup>+</sup>: 392.1157; found: 392.1158.



**2,3,5-Trideoxy-2,3-difluoro-5-propionamido-D-erythro- $\beta$ -L-manno-2-nonulopyranosonic acid (2.64)**

*N,N*-Diisopropylethylamine (31  $\mu$ L) was added to a solution of **2.62** (14 mg, 0.049 mmol) in acetonitrile (1 mL). Propionyl chloride (6  $\mu$ L in 0.1 mL ACN, 2 eq.) was added dropwise, the reaction stirred for 2 hours at room temperature and then quenched by addition of water (0.5 mL) and neutralized with Amberlite IR120-H resin. Solvent was removed, purification by silica gel column chromatography (9:2:1 EtOAc/MeOH/H<sub>2</sub>O) and lyophilization afforded 15 mg (89% yield) of **2.64** as a white solid.

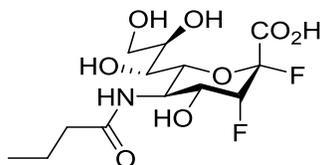
R<sub>f</sub> = 0.2 (7:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  5.23 (dt,  $J_{3,F3} = 51.5$  Hz,  $J_{3,F2} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.29 (td,  $J_{5,4} = 10.5$  Hz,  $J_{5,6} = 10.5$  Hz,  $J_{5,F3} = 1.4$  Hz, 1H, H-5), 4.17 (ddt,  $J_{4,F3} = 28.1$  Hz,  $J_{4,5} = 10.5$ ,  $J_{4,3} = 2.0$  Hz, 1H, H-4), 3.93 – 3.85 (m, 2H, H-8, H-9a), 3.82 (dd,  $J_{6,5} = 10.5$  Hz,  $J_{6,7} = 1.4$  Hz, 1H, H-6), 3.71 – 3.60 (m, 1H, H-9b), 3.55 (dt,  $J_{7,8} = 9.2$  Hz,  $J_{7,6} = 1.7$  Hz, 1H, H-7), 2.32 (q,  $J = 7.7$  Hz, 2H, CH<sub>2</sub>), 1.13 (t,  $J = 7.7$  Hz, 3H, CH<sub>3</sub>).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  179.3 (NH CO), 169.8 (dd,  $J_{C1,F2} = 26.5$  Hz,  $J_{C1,F3} = 3.7$  Hz, C-1), 106.9 (dd,  $J_{C2,F2} = 219.4$  Hz,  $J_{C2,F3} = 14.6$  Hz, C-2), 89.3 (dd,  $J_{C3,F3} = 183.6$  Hz,  $J_{C3,F2} = 18.5$  Hz, C-3), 72.9 (d,  $J_{C6,F2} = 4.3$  Hz, C-6), 70.8 (C-7/C-8), 69.3 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 6.1$  Hz, C-4), 68.1 (C-8/C-7), 63.1 (C-9), 47.0 (d,  $J_{C5,F3} = 3.4$  Hz, C-5), 29.4 (CH<sub>2</sub>), 9.7 (CH<sub>3</sub>).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O)  $\delta$  -121.52 (d,  $J_{F2,F3} = 12.0$  Hz, F-2 eq), -218.26 (d,  $J_{F3,F2} = 11.7$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>12</sub>H<sub>19</sub>F<sub>2</sub>NO<sub>8</sub> + H]<sup>+</sup>: 344.1157; found: 344.1158.



**5-Butyramido-2,3,5-trideoxy-2,3-difluoro-D-erythro- $\beta$ -L-manno-2-nonulopyranosonic acid (2.65)**

*N,N*-Diisopropylethylamine (15  $\mu$ L) was added to a solution of **2.62** (7 mg, 0.024 mmol) in acetonitrile (0.5 mL). Butyryl chloride (5  $\mu$ L in 0.1 mL ACN, 2 eq.) was added dropwise, the reaction stirred for 2 hours at room temperature and was quenched by addition of water (0.5 mL) and neutralized with Amberlite IR120-H resin. Solvent was removed, purification by silica gel column chromatography (9:2:1 EtOAc/MeOH/H<sub>2</sub>O) and lyophilization afforded 8 mg (93% yield) of **2.65** as a white solid.

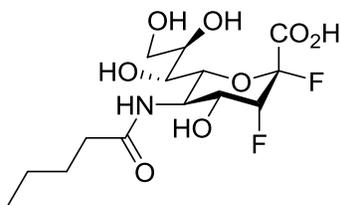
R<sub>f</sub> = 0.2 (7:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  5.23 (dt,  $J_{3,F3} = 51.5$  Hz,  $J_{3,F2} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.30 (t,  $J_{5,4} = 10.5$  Hz,  $J_{5,6} = 10.5$  Hz, 1H, H-5), 4.18 (ddt,  $J_{4,F3} = 28.5$  Hz,  $J_{4,5} = 10.9$  Hz,  $J_{4,3} = 2.0$  Hz, 1H, H-4), 3.93 – 3.85 (m, 2H, H-8, H-9a), 3.82 (d,  $J_{6,5} = 10.6$  Hz, 1H, H-6), 3.69 – 3.60 (m, 1H, H-9b), 3.57 (dt,  $J_{7,8} = 9.0$  Hz,  $J_{7,6} = 1.7$  Hz, 1H, H-7), 2.29 (t,  $J = 7.3$  Hz, 2H, CH<sub>2</sub>), 1.64 (h,  $J = 7.4$  Hz, 2H, CH<sub>2</sub>), 0.93 (t,  $J = 7.4$  Hz, 3H, CH<sub>3</sub>).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  178.2 (NH CO), 169.6 (dd,  $J_{C1,F2} = 26.7$  Hz,  $J_{C1,F3} = 3.6$  Hz, C-1), 106.7 (dd,  $J_{C2,F2} = 219.4$  Hz,  $J_{C2,F3} = 14.7$  Hz, C-2), 89.1 (dd,  $J_{C3,F3} = 183.5$  Hz,  $J_{C3,F2} = 18.5$  Hz, C-3), 72.6 (d,  $J_{C6,F2} = 4.3$  Hz, C-6), 70.6 (C-7), 69.1 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 6.0$  Hz, C-4), 68.0 (C-8), 62.9 (C-9), 46.8 (d,  $J_{C5,F3} = 3.4$  Hz, C-5), 37.9 (CH<sub>2</sub>), 19.0 (CH<sub>2</sub>), 12.8 (CH<sub>3</sub>).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O)  $\delta$  -121.49 (d,  $J_{F2,F3} = 11.6$  Hz, F-2 eq), -218.22 (d,  $J_{F3,F2} = 11.8$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>13</sub>H<sub>21</sub>F<sub>2</sub>NO<sub>8</sub> + H]<sup>+</sup>: 358.1313; found: 358.1318.



**2,3,5-Trideoxy-2,3-difluoro-5-pentanamido-D-erythro- $\beta$ -L-manno-2-nonulopyranosonic acid (2.66)**

*N,N*-Diisopropylethylamine (25  $\mu$ L) was added to a solution of **2.62** (12 mg, 0.042 mmol) in acetonitrile (1 mL). Valeroyl chloride (10  $\mu$ L in 0.1 mL ACN, 2 eq.) was added dropwise, the reaction stirred for 2 hours at room temperature and was quenched by addition of water (0.5 mL) and neutralized with Amberlite IR120-H resin. Solvent was removed, purification by silica gel column chromatography (12:2:1 EtOAc/MeOH/H<sub>2</sub>O) and lyophilization afforded 8 mg (51% yield) of **2.66** as a white solid.

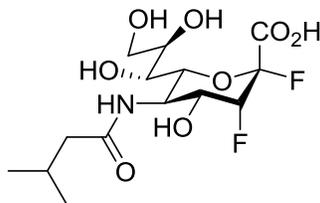
R<sub>f</sub> = 0.4 (7:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  5.22 (dt,  $J_{3,F3} = 51.4$  Hz,  $J_{3,F2} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.29 (t,  $J_{5,4} = 10.2$  Hz,  $J_{5,6} = 10.2$  Hz, 1H, H-5), 4.16 (ddt,  $J_{4,F3} = 28.4$  Hz,  $J_{4,5} = 10.7$  Hz,  $J_{4,3} = 2.0$  Hz, 1H, H-4), 3.94 – 3.84 (m, 2H, H-8, H-9a), 3.81 (d,  $J_{6,5} = 10.5$  Hz, 1H, H-6), 3.68 – 3.59 (m, 1H, H-9b), 3.56 (dt,  $J_{7,8} = 9.2$  Hz,  $J_{7,6} = 1.8$  Hz, 1H, H-7), 2.31 (t,  $J = 7.4$  Hz, 2H), 1.59 (p,  $J = 7.4$  Hz, 2H), 1.32 (dt,  $J = 16.3, 7.4$  Hz, 2H), 0.90 (t,  $J = 7.3$  Hz, 3H).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  178.6 (NH CO), 169.8 (dd,  $J_{C1,F2} = 26.7$  Hz,  $J_{C1,F3} = 3.3$  Hz, C-1), 106.9 (dd,  $J_{C2,F2} = 219.4$  Hz,  $J_{C2,F3} = 14.7$  Hz, C-2), 89.3 (dd,  $J_{C3,F3} = 183.6$  Hz,  $J_{C3,F2} = 18.4$  Hz, C-3), 72.9 (d,  $J_{C6,F2} = 4.4$  Hz, C-6), 70.8 (C-8), 69.3 (dd,  $J_{C4,F3} = 18.1$  Hz,  $J_{C4,F2} = 6.0$  Hz, C-4), 68.2 (C-7), 63.1 (C-9), 46.9 (d,  $J_{C5,F3} = 3.3$  Hz, C-5), 35.9 (CH<sub>2</sub>), 27.7 (CH<sub>2</sub>), 21.8 (CH<sub>2</sub>), 13.2 (CH<sub>3</sub>).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O)  $\delta$  -121.45 (d,  $J_{F2,F3} = 11.7$  Hz, F-2 eq), -218.16 (ddd,  $J_{F3,H3} = 51.6$  Hz,  $J_{F3,H4} = 27.8$  Hz,  $J_{F3,F2} = 11.9$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>14</sub>H<sub>23</sub>F<sub>2</sub>NO<sub>8</sub> + H]<sup>+</sup>: 372.1470; found: 372.1471.



**2,3,5-Trideoxy-2,3-difluoro-5-(3'-methylbutanamido)-D-erythro- $\beta$ -L-manno-2-nonulopyranosonic acid (2.67)**

*N,N*-Diisopropylethylamine (30  $\mu$ L) was added to a solution of **2.62** (14 mg, 0.049 mmol) in acetonitrile (1 mL). Valeroyl chloride (12  $\mu$ L in 0.1 mL ACN, 2 eq.) was added dropwise, the reaction stirred for 2 hours at room temperature and then quenched by addition of water (0.5 mL) and neutralized with Amberlite IR120-H resin. Solvent was removed: purification by silica gel column chromatography (10:2:1 EtOAc/MeOH/H<sub>2</sub>O) and lyophilization afforded 13 mg (72% yield) of **2.67** as a white solid.

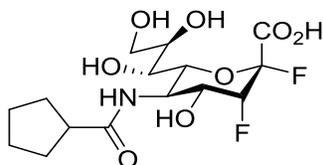
R<sub>f</sub> = 0.4 (7:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  5.23 (dt,  $J_{3,F3} = 51.5$  Hz,  $J_{3,F2} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.30 (t,  $J_{5,4} = 10.5$  Hz,  $J_{5,6} = 10.5$  Hz, 1H, H-5), 4.17 (ddt,  $J_{4,F3} = 28.5$  Hz,  $J_{4,5} = 10.9$  Hz,  $J_{4,3} = 2.1$  Hz, 1H, H-4), 3.96 – 3.85 (m, 2H, H-8, H-9a), 3.82 (d,  $J_{6,5} = 10.5$  Hz, 1H, H-6), 3.64 (dd,  $J_{9b,9a} = 12.0$  Hz,  $J_{9b,8} = 6.3$  Hz, 1H, H-9b), 3.59 (dt,  $J_{7,8} = 9.1$  Hz,  $J_{7,6} = 1.8$  Hz, 1H, H-7), 2.19 (d,  $J = 8.0$  Hz, 2H, CH<sub>2</sub>), 2.02 (dt,  $J = 13.7, 6.9$  Hz, 1H, CH), 0.95 (dd,  $J = 6.6, 3.0$  Hz, 6H, 2 x CH<sub>3</sub>).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  177.8 (NH CO), 169.8 (dd,  $J_{C1,F2} = 26.5$  Hz,  $J_{C1,F3} = 3.6$  Hz, C-1), 107.0 (dd,  $J_{C2,F2} = 219.5$  Hz,  $J_{C2,F3} = 14.6$  Hz, C-2), 89.4 (dd,  $J_{C3,F3} = 183.5$  Hz,  $J_{C3,F2} = 18.5$  Hz, C-3), 72.9 (d,  $J_{C6,F2} = 4.4$  Hz, C-6), 70.8 (C-7/C-8), 69.3 (dd,  $J_{C4,F3} = 18.1$  Hz,  $J_{C4,F2} = 6.1$  Hz, C-4), 68.3 (C-7/C-8), 63.2 (C-9), 47.0 (d,  $J_{5,F3} = 3.4$  Hz, C-5), 45.4 (CH<sub>2</sub>), 26.4 (CH), 21.8 (CH<sub>3</sub>), 21.8 (CH<sub>3</sub>).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O)  $\delta$  -121.49 (d,  $J_{F2,F3} = 11.6$  Hz, F-2 eq), -218.20 (d,  $J_{F3,F2} = 11.6$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>14</sub>H<sub>23</sub>F<sub>2</sub>NO<sub>8</sub> + H]<sup>+</sup>: 372.1470; found: 372.1472.



**5-(Cyclopentanecarboxamido)-2,3,5-trideoxy-2,3-difluoro-D-erythro- $\beta$ -L-manno-2-nonulopyranosonic acid (2.68)**

*N,N*-Diisopropylethylamine (25  $\mu$ L) was added to a solution of **2.62** (13 mg, 0.045 mmol) in acetonitrile (1 mL). Cyclopentanoyl chloride (11  $\mu$ L in 0.1 mL ACN, 2 eq.) was added dropwise, the reaction mixture stirred for 2 hours at room temperature and then quenched by addition of water (0.5 mL) and neutralized with Amberlite IR120-H resin. Solvent was removed; purification by silica gel column chromatography (11:2:1 EtOAc/MeOH/H<sub>2</sub>O) and lyophilization afforded 10 mg (58% yield) of **2.68** as a white solid.

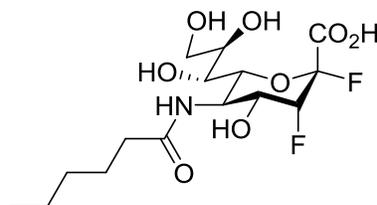
R<sub>f</sub> = 0.3 (7:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  5.22 (dt,  $J_{3,F3} = 51.5$  Hz,  $J_{3,F2} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-5), 4.28 (td,  $J_{5,4} = 10.5$  Hz,  $J_{5,6} = 10.5$  Hz,  $J_{5,F3} = 1.5$  Hz, 1H, H-5), 4.17 (ddt,  $J_{4,F3} = 28.2$  Hz,  $J_{4,5} = 10.6$  Hz,  $J_{4,3} = 2.0$  Hz, 1H, H-4), 3.92 – 3.84 (m, 2H, H-8, H-9a), 3.81 (dd,  $J_{6,5} = 10.6$  Hz,  $J_{6,7} = 1.4$  Hz, 1H, H-6), 3.67 – 3.58 (m, 1H, H-9b), 3.52 (dt,  $J_{7,8} = 9.3$  Hz,  $J_{7,6} = 1.7$  Hz, 1H, H-7), 2.73 (dd,  $J = 10.3, 4.9$  Hz, 1H, CH), 1.90 (dddd,  $J = 11.9, 9.7, 7.0, 4.1$  Hz, 2H), 1.76 – 1.52 (m, 6H).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  181.4 (NH CO), 169.6 (dd,  $J_{C1,F2} = 26.5$  Hz,  $J_{C1,F3} = 3.6$  Hz, C-1), 106.7 (dd,  $J_{C2,F2} = 219.5$  Hz,  $J_{C2,F3} = 14.6$  Hz, C-2), 89.1 (dd,  $J_{C3,F3} = 183.5$  Hz,  $J_{C3,F2} = 18.5$  Hz, C-3), 72.7 (d,  $J_{C6,F2} = 4.2$  Hz, C-6), 70.5 (C-7/C-8), 69.0 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 6.0$  Hz, C-4), 69.0 (C-7/C-8), 62.9 (C-9), 46.8 (d,  $J_{C5,F3} = 3.4$  Hz, C-5), 45.3 (CH), 30.4 (CH<sub>2</sub>), 29.9 (CH<sub>2</sub>), 25.6 (CH<sub>2</sub>), 25.6 (CH<sub>2</sub>).

<sup>19</sup>F NMR (377 MHz, D<sub>2</sub>O)  $\delta$  -120.85 (dd,  $J_{F2,F3} = 12.0$  Hz,  $J_{F2,H3} = 2.4$  Hz, F-2 eq), -217.59 (dddd,  $J_{F3,H3} = 51.5$  Hz,  $J_{F3,H4} = 28.2$  Hz,  $J_{F3,F2} = 11.7$  Hz,  $J_{F3,H5} = 1.6$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>15</sub>H<sub>23</sub>F<sub>2</sub>NO<sub>8</sub> + H]<sup>+</sup>: 384.1470; found: 384.1475.



**2,3,5-Trideoxy-2,3-difluoro-5-hexanamido-D-erythro- $\beta$ -L-manno-2-nonulopyranosonic acid (2.69)**

*N,N*-Diisopropylethylamine (27  $\mu$ L) was added to a solution of **2.62** (13 mg, 0.045 mmol) in acetonitrile (1 mL). Hexanoyl chloride (12  $\mu$ L in 0.1 mL ACN, 2 eq.) was added dropwise, the reaction mixture stirred for 2 hours at room temperature and then quenched by addition of water (0.5 mL) and neutralized with Amberlite IR120-H resin. Solvent was removed: purification by silica gel column chromatography (12:2:1 EtOAc/MeOH/H<sub>2</sub>O) and lyophilization afforded 13 mg (75% yield) of **2.69** as a white solid.

R<sub>f</sub> = 0.3 (9:2:1 EtOAc/MeOH/H<sub>2</sub>O)

<sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  5.21 (dt,  $J_{3,F3} = 51.5$  Hz,  $J_{3,F2} = 2.7$  Hz,  $J_{3,4} = 2.7$  Hz, 1H, H-3), 4.28 (t,  $J_{5,4} = 10.2$  Hz,  $J_{5,6} = 10.2$  Hz, 1H, H-5), 4.16 (ddt,  $J_{4,F3} = 28.3$  Hz,  $J_{4,5} = 10.7$ ,  $J_{4,3} = 2.0$  Hz,  $J_{4,F2} = 2.0$  Hz, 1H, H-4), 3.93 – 3.84 (m, 2H, H-8, H-9a), 3.80 (d,  $J_{6,5} = 10.1$  Hz, 1H, H-6), 3.68 – 3.58 (m, 1H, H-9b), 3.55 (dt,  $J_{7,8} = 9.2$  Hz,  $J_{7,6} = 1.8$  Hz, 1H, H-7), 2.30 (t,  $J = 7.4$  Hz, 2H, CH<sub>2</sub>), 1.61 (p,  $J = 7.3$  Hz, 2H, CH<sub>2</sub>), 1.40 – 1.18 (m, 4H, 2 x CH<sub>2</sub>), 0.88 (t,  $J = 6.8$  Hz, 3H, CH<sub>3</sub>).

<sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  178.6 (NH CO), 169.8 (dd,  $J_{C1,F2} = 26.5$  Hz,  $J_{C1,F3} = 3.6$  Hz, C-1), 106.9 (dd,  $J_{C2,F2} = 219.5$  Hz,  $J_{C2,F3} = 14.7$  Hz, C-2), 89.3 (dd,  $J_{C3,F3} = 183.5$  Hz,  $J_{C3,F2} = 18.5$  Hz, C-3), 72.9 (d,  $J_{C6,F2} = 4.3$  Hz, C-6), 70.8 (C-7/C-8), 69.3 (dd,  $J_{C4,F3} = 18.0$  Hz,  $J_{C4,F2} = 6.0$  Hz), 68.2 (C-7/C-8), 63.2 (C-9), 47.0 (d,  $J_{C5,F3} = 3.4$  Hz, C-5), 36.1 (CH<sub>2</sub>), 30.7 (CH<sub>2</sub>), 25.3 (CH<sub>2</sub>), 21.9 (CH<sub>2</sub>), 13.4 (CH<sub>3</sub>).

<sup>19</sup>F NMR (282 MHz, D<sub>2</sub>O)  $\delta$  -121.54 (d,  $J_{F2,F3} = 11.9$  Hz, F-2 eq), -218.23 (d,  $J_{F3,F2} = 12.0$  Hz, F-3 ax).

ESI-HRMS: Calc'd for [C<sub>15</sub>H<sub>25</sub>F<sub>2</sub>NO<sub>8</sub> + H]<sup>+</sup>: 386.1626; found: 386.1626.

### 7.1.7 Enzyme expression

Wild type TcTS was expressed in BL21 (DE3) competent *E. coli* cells and purified following reported procedures with minor modifications<sup>115</sup>. A single colony of cells was inoculated into a small Erlenmeyer flask containing 65 mL lysogeny broth (LB) with carbenicillin (0.1 mg/mL) and then incubated overnight at 25 °C with constant shaking (220 RPM). From the overnight growth, ~5 mL of culture was used to inoculate each of four 1 L flasks containing 500 mL/flask of autoclaved LB and 0.1 mg/mL carbenicillin. The 2 L of cell culture were grown for ~4 hours at 37 °C with constant shaking at 220 RPM (until Abs<sub>600</sub> ~ 0.6). Isopropylthiogalactoside (IPTG) was then added (0.5 mM) to each flask to induce enzyme expression. Cells were incubated at 30 °C with constant agitation (220 RPM) overnight. Cells were harvested by centrifugation (30 min, 5000 RPM), the supernatant was removed and the cells resuspended in binding buffer (20 mM Tris, 0.5 M NaCl, 5 mM imidazole, pH = 8). One tablet of EDTA-free protease inhibitor cocktail (Roche) and benzonase (0.5 µL) were added to the resuspended solution. Cells were lysed by three successive rounds of homogenization using the French cell press. Cell extracts were clarified by centrifugation (30 min, 15,000 RPM) and subsequently loaded onto a 5 mL HisTrap Ni-NTA affinity column (GE Healthcare). The column was rinsed with binding buffer to remove any unbound protein. Target protein was eluted using a linear gradient of imidazole from 5 mM to 0.5 M in 20 mM Tris-HCl, 0.5 M NaCl, pH = 8 with an AKTA purification system and UV detection at A<sub>280</sub> nm. The proteinaceous fractions were analyzed by SDS-PAGE and those containing TcTS pooled. The enzyme was buffer exchanged into 20 mM Tris-HCl, pH = 8, by 30 kDa spin filter (Amicon) centrifugation (3 x 15 min at 5,000 RPM). TcTS was then eluted with a linear NaCl elution gradient from 0-1 M NaCl over a 5 mL MonoQ column using the AKTA purification system and UV detection at A<sub>280</sub> nm. Proteinaceous fractions were again analyzed by SDS-PAGE, fractions containing TcTS were combined and buffer exchanged into 20 mM Tris/30 mM NaCl, pH = 8, using a 30 kDa spin filter (Amicon) by centrifugation, concentrated to ~10 mg/mL and stored at 4 °C.

### 7.1.8 Enzyme kinetics

All experiments were carried out in standard TcTS buffer (20 mM Tris, 30 mM NaCl, pH = 7.5, 0.1% BSA). Quartz cuvettes with a path length of 1 cm were used in a Cary 4000 UV/Vis spectrophotometer. All kinetic data was analyzed using the data analysis software GraFit 7.0 (Erithacus).

#### 7.1.8.1 Measuring $k_i/K_i$ for 9-amido-difluorosialoside inactivators

Time-dependent inactivations were performed by pre-incubating buffered TcTS (0.5  $\mu$ M, 180  $\mu$ L) in 0.6 mL Eppendorf tubes in a 25 °C water bath, followed by the addition of varying concentrations (from 0.1 mM to 20 mM) of 9-amido-DFSA inactivators (20  $\mu$ L) to each tube. Residual enzyme activity was determined over the course of the 120-180 min incubation period by the diluting nine aliquots (20  $\mu$ L) of the inactivation mixture (over the total time of the inactivation) 10-fold into quartz cuvettes charged with buffered substrate solution (180  $\mu$ L, 0.5 mM): 4-trifluoromethylumbelliferyl- $\beta$ -D-sialoside (TFMU-SA). Residual TcTS activity was determined by measuring the initial linear increase in absorbance at  $\lambda = 385$  nm for 5-10 minutes. The initial rates ( $v_o$ ), were plotted against the incubation time to obtain time-dependent decay curves were obtained for each concentration of inactivator. By fitting to a pseudo first-order exponential decay equation,  $v = v_o e^{(k_{obs})t} + \text{offset}$ , the rate ( $k_{obs}$ ) could be obtained for each curve. An equation with an offset was used because the rates did not decay to zero. The observed rate ( $k_{obs}$ ) was then plotted against inactivator concentration to a Michaelis-Menten like plot from which the inhibition parameters could be obtained. However, in each case  $[I] \ll K_i$ , thus only the second-order rate constant ( $k_i/K_i$ ) was determined by fitting the data to the equation:  $k_{obs} = k_i[I]/K_i$ .

#### 7.1.8.2 Measuring $k_{obs}$ and half-life for 5-amido-difluorosialoside inactivators

Time-dependent inactivation curves were measured for 5-amido-DFSA inactivators versus TcTS activity in an analogous manner to those described in section 7.1.8.1. In this case, TcTS (0.5  $\mu$ M, 180  $\mu$ L) was incubated with 5-amido-DFSAs at a single concentration (5 mM) for 120-180 minutes. Nine aliquots

(20  $\mu\text{L}$ ) of enzyme-inactivator solution was diluted 10-fold over the course the inactivation into quartz cuvettes charged with buffered TFMU-SA solution (180  $\mu\text{L}$ , 0.5 mM). Residual TcTS activity was determined by measuring the initial linear increase in absorbance at  $\lambda = 385 \text{ nm}$ , corresponding to TFMU-SA cleavage, for 5-10 minutes. The initial rates ( $v_0$ ), were plotted against the incubation time to obtain time-dependent decay curves for each inactivator. By fitting to a pseudo first-order exponential decay equation,  $v = v_0 e^{(k_{\text{obs}})t} + \text{offset}$ , the inactivation rate ( $k_{\text{obs}}$ ) at 5 mM could be obtained for each curve. The half-life ( $t_{1/2}$ ) of TcTS in the presence of 5 mM inactivator was calculated using the equation  $t_{1/2} = 0.693/k_{\text{obs}}$ .

## 7.2 Natural product screening and isolation for TcTS inhibition

### 7.2.1 Screening protocol

Marine sponge extract library was obtained courtesy of the Andersen lab at the University of British Columbia. The library of 986 crude methanolic sponge extracts provided as dried samples distributed across 12 x 96-well plates. Samples were solubilized in DMSO (100  $\mu\text{L}/\text{well}$ ) and library was replicated by 20-, 400- and 4000-fold dilutions in DMSO using the Beckman Coulter Biomek FX Laboratory Automation Workstation. The 96-channel pipetting instrument first charged the 96-well reduced area plates (Costar) with a buffered solution of TcTS (70  $\mu\text{L}$ , 56 nM, 20 mM Tris, 30 mM NaCl, 0.1% BSA, pH = 7.5), then added DMSO-solubilized inhibitor solution (10  $\mu\text{L}$ ). After a 10 minute incubation period, the pipetting instrument added 20  $\mu\text{L}$  of buffered substrate solution, TFMU-SA (0.5 mM). Initial rates of substrate cleavage were monitored at  $A_{405 \text{ nm}}$  using UV/Vis detection on the Beckman Coulter DTX880 plate reader over the course of 15 minutes. Each 96-well plate included at least two columns of positive control wells (10% DMSO/extract-free) and typically included one column of negative control wells (enzyme-free). Linear regression was performed on the data obtained from each well and the resultant slope indicated the initial rate ( $v_0$ ) of substrate cleavage. Samples that displayed a negative slope were represented with rate of zero. On a per plate basis, the difference between the average

initial rates in positive and negative controls was obtained. Background-subtracted initial rates from each sample well were then compared to the positive controls to determine the percent TcTS inhibition (relative to the average non-inhibited control). Visual representation of these plates are presented in the text, where values are shown as percent inhibition and are coloured for visualization from white (no inhibition) to red (strong inhibition).

### 7.2.2 TcTS inhibition assay

Recombinant TcTS was diluted in buffer (20 mM Tris-HCl, 30 mM NaCl, 0.1% BSA, pH = 7.5) and aliquoted (70  $\mu$ L/well) into 96-well reduced volume clear plate (Costar). DMSO (10  $\mu$ L) or DMSO-solubilized samples/extracts (10  $\mu$ L) were added to each sample using an 8-channel or 12-channel multi-channel pipette. Samples/extracts were assayed in duplicate following serial 2-fold or 10-fold dilutions. The assay was initiated by addition of TFMU-SA (20  $\mu$ L; 0.5 mM), and the initial rates of substrate cleavage were monitored by UV/Vis at  $A_{405}$  nm on the Beckman Coulter DTX880 plate reader for 15 minutes. Background-subtracted initial rates from each sample were compared to the positive controls to determine the percentage TcTS inhibition at each concentration. When appropriate, half-maximal inhibitor concentrations ( $IC_{50}$ ) were determined by fitting to the equation:  $y = \frac{range}{1 + (\frac{x}{IC_{50}})^s}$ . Where *range* is the uninhibited value, *s* is the slope factor and *y* falls with increasing *x*.

### 7.2.3 $\beta$ -Galactosidase from *Aspergillus oryzae* inhibition assay

$\beta$ -Galactosidase ( $\beta$ Gal) from *Aspergillus oryzae* was obtained from Sigma and used without purification. Buffer (200 mM sodium phosphate, pH = 6.5) and enzyme substrate (Gal- $\beta$ -MU; 500  $\mu$ M), were selected from literature examples. Buffered  $\beta$ Gal (70  $\mu$ L) was incubated with DMSO (10  $\mu$ L) or DMSO-solubilized extracts (10  $\mu$ L) for 5-10 minutes in a 96-well reduced volume clear plate. Buffered Gal- $\beta$ -MU was added (20  $\mu$ L) and MU release was monitored by UV/Vis spectroscopy at  $A_{370}$  and the

initial rates determined by linear regression. Percent enzyme activity was calculated relative to the initial rate of uninhibited (DMSO-containing) controls.

#### **7.2.4 Human NEU2 inhibition assay**

The purified recombinant enzyme – cytosolic human neuraminidase 2 (hNEU2) – was obtained courtesy of the Cairo lab at the University of Alberta. Standard hNEU2 assay conditions require low pH buffer (0.1 M NaOAc, pH = 5.5) which precludes the effective use of UV/Vis detection with the TFMU-SA substrate, thus a fluorescence assay was instead employed. Buffered hNEU2 (180  $\mu$ L) was added to a 0.6 mL Eppendorf tube and incubated with DMSO-solubilized inhibitor (20  $\mu$ L) at 30 °C in a water bath for 5-10 minutes. Fluorogenic substrate TFMU-SA (20  $\mu$ L; 0.8 mM) was added and the reaction aliquoted (25  $\mu$ L) every 2 minutes into quenching buffer (375  $\mu$ L; 200 mM glycine buffer, pH = 8.5) and the fluorescence measured in clear plastic cuvettes on the Cary Eclipse fluorescence spectrophotometer ( $\lambda_{\text{ex}}$  = 370 nm;  $\lambda_{\text{em}}$  = 450 nm). Linear regression of the stopped assay time-points affords a linear initial rate that can be compared to control (DMSO only/extract-free) TcTS activity to determine the percent inhibition.

#### **7.2.5 General isolation procedure**

Methanolic extracts of sponge hits (03-395, 11-312, 00-287, 03-242, 97-191) were obtained courtesy of the Andersen lab at the University of British Columbia. Solid samples (100-300 mg) were solubilized in EtOAc and partitioned with H<sub>2</sub>O. Aqueous samples were in some cases additionally extracted with *n*-BuOH. Extract samples from liquid-liquid extraction were dried and assayed versus TcTS at concentrations from 0.1-0.001 mg/mL. Dried extracts from the organic phase were respectively dissolved in 4:1 MeOH/DCM and partitioned via gel filtration chromatography (Sephadex LH-20). Samples were eluted with 4:1 MeOH/DCM by gravity and fractions were collected in 8 mL glass tubes using an automated fraction collector. Fractions were assessed by TLC (4:1 DCM/MeOH) and those with similar polarity products (as visualized by UV or by staining with 6% (w/v) vanillin/1% (v/v) H<sub>2</sub>SO<sub>4</sub> in

EtOH and charring) were pooled. Pooled fractions were dried, weighed, redissolved in DMSO and assayed for TcTS inhibition at concentrations of 0.1, 0.01, and 0.001 mg/mL. Dried samples were then suspended in a minimal volume of MeOH and further partitioned by stepwise elution of MeOH/H<sub>2</sub>O on reverse phase chromatography (Sep-Pak tC18) from 0-100% MeOH. Resultant fractions were dried, weighed, redissolved in DMSO and tested for TcTS inhibition at concentrations of 0.1, 0.01, 0.001 and 0.0001 mg/mL. Fractions that were showed TcTS inhibition, with sufficient material (>5 mg), were partitioned by semi-preparative C18 HPLC (65-80% gradient of ACN/H<sub>2</sub>O over 30 min at 4 mL/min) and fractions collected each minute. HPLC fractions were dried, dissolved in MeOH/H<sub>2</sub>O (1 mL) and assayed for TcTS inhibition. Active samples were analyzed by NMR spectroscopy and mass spectrometry.

#### **7.2.6 Ircinialactam screening**

Two 96-well plates containing a total of 117 synthetic ircinialactam-like compounds were provided courtesy of the Capon lab at the University of Queensland. Samples were sent as dried aliquots of ~0.1 mg of compound per well. Each well was solubilized in DMSO (100 µL/well) and two dilution plates were prepared (10- and 100-fold). Each plate was assayed for inhibitory activity following the *TcTS inhibition assay* protocol (see Chapter 7.2.2).

#### **7.2.7 Enzyme kinetics**

**IC<sub>50</sub> determination:** Eleven compounds (**3.10-3.16; 3.18-3.21**) identified from ircinialactam library were obtained as pre-weighed solids in 1.3-2.2 mg quantities, courtesy of the Capon lab at the University of Queensland. Concentration-dependent inhibition of TcTS activity was measured at concentrations of ircinialactams ranging from 0.01 µM to 100 µM following the *TcTS inhibition assay* protocol (see Chapter 7.2.2).

**Time-dependent inhibition assay:** Four hit compounds were assayed for time-dependent inhibition.

TcTS was incubated with 25  $\mu\text{M}$  of **3.10**, **3.11**, **3.14** and **3.20** respectively in buffer (20 mM Tris, 30 mM NaCl, pH = 7.4) at 25  $^{\circ}\text{C}$  for 30 minutes. Aliquots of the mixtures were collected throughout the incubation (1, 2, 5, 10, 15, 30 minutes) and diluted 10-fold into quartz cuvettes charged with buffered TFMU-SA (500  $\mu\text{M}$ ) solution. Initial rate of TFMU-SA cleavage was monitored in the Cary 4000 UV/Vis spectrophotometer  $A_{385\text{ nm}}$  for 10-15 minutes. The percent residual active, relative to uninhibited control, was plotted versus time for each sample.

**$K_i$  determination:** The inhibition constant and mode of inhibition for each of the eleven ircinialactam species (**3.10-3.16**; **3.18-3.21**) was determined. Clear 96-well reduced volume plates (Costar) were charged with 70  $\mu\text{L}$  of buffered TcTS solution (55 nM; 20 mM Tris, 30 mM NaCl, pH = 7.5), to which 10  $\mu\text{L}$  of DMSO or DMSO-dissolved ircinialactams were added (0.1-25  $\mu\text{M}$ ). Each concentration of inhibitor was assayed at three concentrations of TFMU-SA (0.25, 0.5, 1 mM) and the initial rate of enzymatic hydrolysis measured by UV/Vis at 385 nm for 15 minutes. Initial rates, inhibitor concentration and substrate concentrations were input into data analysis software GraFit 7.0 (Erithacus). Data was fit to equations describing competitive, non-competitive and uncompetitive inhibition (See Appendix A) and the fitting parameter, reduced  $\chi^2$ , was used to select the best fit (Table 7.1). When the fit parameter reduced  $\chi^2 > 100$ , the inhibition constant  $K_i$  was said to be not determined (N.D.), since the fit to the corresponding equation was unsatisfactory.

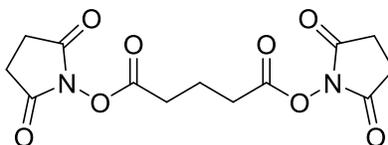
**Table 7.1. Inhibition values ( $K_i$ ) and fitting parameter ( $\chi^2$ ) from GraFit 7.0 fitting software.**

Compound	Competitive $K_i$ $\mu\text{M}$ ( $\chi^2$ )	Non-competitive $K_i$ $\mu\text{M}$ ( $\chi^2$ )	Uncompetitive $K_i$ $\mu\text{M}$ ( $\chi^2$ )	Mechanism*
3.10	N.D. (>100)	$0.8 \pm 0.2$ (5)	$0.9 \pm 0.1$ (26)	non-competitive
3.11	N.D. (>100)	$4.7 \pm 1.0$ (2)	$3.2 \pm 1.1$ (17)	non-competitive
3.12	N.D. (>100)	$2.8 \pm 0.9$ (8)	$2.3 \pm 0.5$ (35)	non-competitive
3.13	N.D. (>100)	$3.8 \pm 0.7$ (12)	$2.7 \pm 0.8$ (71)	non-competitive
3.14	N.D. (>100)	$4.6 \pm 0.7$ (7)	$3.6 \pm 0.7$ (33)	non-competitive
3.15	N.D. (>100)	$4.5 \pm 1.9$ (7)	$3.4 \pm 0.8$ (39)	non-competitive
3.16	N.D. (>100)	$6.0 \pm 1.1$ (6)	$4.6 \pm 0.8$ (47)	non-competitive
3.18	N.D. (>100)	$2.2 \pm 0.3$ (4)	$1.8 \pm 0.3$ (24)	non-competitive
3.19	N.D. (>100)	$2.1 \pm 0.4$ (2)	$1.7 \pm 0.3$ (21)	non-competitive
3.20	N.D. (>100)	$7.1 \pm 1.0$ (6)	$5.9 \pm 0.8$ (41)	non-competitive
3.21	N.D. (>100)	$6.5 \pm 0.9$ (3)	$5.3 \pm 1.1$ (22)	non-competitive

N.D. = not determined; \*inhibitor mechanism selected based on best fit (*i.e.* lowest reduced  $\chi^2$  value).

### 7.3 Cyclic-peptide inhibitors

#### 7.3.1 Synthesis of DSG



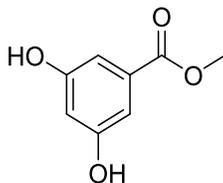
#### Disuccinimidyl glutarate<sup>203</sup> (4.19)

To a cooled (0 °C) solution of *N*-hydroxysuccinimide (2.14 g, 17.3 mmol) in dry THF (50 mL), triethylamine (2.4 mL, 17.3 mmol) was added followed by dropwise addition of glutaryl chloride (1.0 mL, 7.9 mmol). The solution was warmed gradually to room temperature and stirred for 2 hours. Solvent was evaporated, the solid was dissolved in DCM and washed with water. Organic solution was dried over  $\text{MgSO}_4$ , the solvent was evaporated and the solid was recrystallized from isopropanol to yield 2.47 g (96% yield) of a crystalline off-white solid.

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  2.83 (s, 8H), 2.79 (t,  $J = 7.3$  Hz, 4H), 2.19 (p,  $J = 7.3$  Hz, 2H).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  169.1 (4C, CO succinimide), 167.8 (2C, CO), 29.7 (2C, CH<sub>2</sub>), 25.7 (4C, CH<sub>2</sub>), 19.7 (CH<sub>2</sub>).

### 7.3.2 Synthesis of L-Cys\*



#### Methyl 3,5-dihydroxybenzoate<sup>204</sup> (4.25)

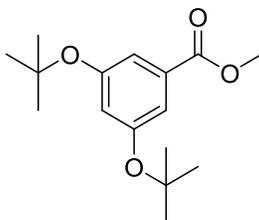
3,5-Dihydroxybenzoic acid (5.7 g, 37 mmol) was suspended in MeOH (50 mL), to which sulfuric acid (5 mL) was added dropwise and the reaction stirred for 4 hours at 60 °C. Solvent was evaporated, the resultant solid was dissolved in fresh EtOAc (100 mL) and washed with H<sub>2</sub>O (2 x 50 mL), NaHCO<sub>3</sub> (2 x 50 mL) and brine (2 x 50 mL). The aqueous portion was again extracted with EtOAc (2 x 50 mL) and washed with brine (50 mL), the combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>. Solvent was removed to yield 5.5 g (88% yield) of **4.25** as a white solid.

R<sub>f</sub> = 0.5 (1:1 Toluene/EtOAc + AcOH)

<sup>1</sup>H NMR (400 MHz, Acetone-d<sub>6</sub>) δ 8.55 (s, 2H, 2 x OH), 7.01 (d, *J* = 2.3 Hz, 2H, H-2, H-6), 6.59 (t, *J* = 2.3 Hz, 1H; H-4), 3.82 (s, 3H; OMe).

<sup>13</sup>C NMR (101 MHz, Acetone-d<sub>6</sub>) δ 167.2 (CO), 159.4 (2C, C-3, C-5), 133.1 (C-1), 108.6 (2C, C-2, C-6), 108.0 (C-4), 52.2 (OMe).

ESI-LRMS: Calc'd for [C<sub>8</sub>H<sub>8</sub>O<sub>4</sub> - H]<sup>-</sup>: 167.0344; found: 167.1.



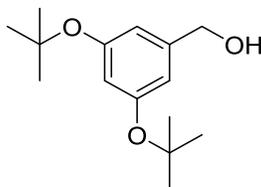
#### Methyl 3,5-di-*tert*-butoxybenzoate<sup>185</sup> (4.26)

**4.25** (3.1 g, 18.5 mmol) and scandium(III) triflate (0.5 g, 1.0 mmol) were suspended in 3:1 DCM/ACN (40 mL) and stirred vigorously. Three aliquots of *t*-butyl dicarbonate (each aliquot: 17.5 g, 80 mmol, 4.5 eq.) were added over 24 hours of stirring at room temperature. Solvents were evaporated, recharged with DCM (30 mL), *t*-butyl dicarbonate (17.5 g, 4.5 eq.) and scandium(III) triflate (1.0 mmol) and stirred for 24 hours. Reaction was diluted with water, extracted with DCM, washed with brine and the separated organic layer was dried over MgSO<sub>4</sub>. Solvent was evaporated and the product purified by silica gel column chromatography (19:1 PE/Et<sub>2</sub>O) afforded 4.04 g (78% yield) of **4.26** as a clear oil.

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.40 (d, *J* = 2.2 Hz, 2H, H-2, H-6), 6.83 (t, *J* = 2.2 Hz, 1H, H-4), 3.88 (s, 3H, *OMe*), 1.34 (s, 18H, 2 x *t*Bu).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 166.9 (CO), 156.0 (2C, C-3, C-5), 131.0 (C-1), 125.0 (C-4), 120.4 (2C, C-2, C-6), 79.4 (2C, C(CH<sub>3</sub>)<sub>3</sub>), 52.3 (*OMe*), 29.0 (6C, C(CH<sub>3</sub>)<sub>3</sub>).

ESI-LRMS: Calc'd for [C<sub>16</sub>H<sub>24</sub>O<sub>4</sub> + Na]<sup>+</sup>: 303.3532; found: 303.2.



**(3,5-Di-*tert*-butoxyphenyl)methanol**<sup>205</sup> (**4.27**)

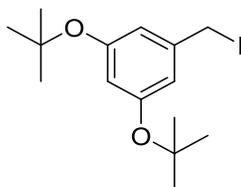
A dry round-bottom flask was charged with LiAlH<sub>4</sub> (1.4 g, 37 mmol) and freshly distilled THF (50 mL), which was stirred at 0 °C. **4.26** (4.04 g, 14.4 mmol) was dissolved in THF (10 mL) and added slowly to the stirring suspension. The mixture was stirred vigorously for 3 hours at 70 °C. Upon completion of the reaction, the mixture was cooled to 0 °C and excess LiAlH<sub>4</sub> was quenched by sequential dropwise addition of EtOAc, EtOH, H<sub>2</sub>O and HCl (1 M) until the solution was acidic. All solvents were evaporated, fresh EtOAc (50 mL) was added, stirred vigorously and the resultant suspension was filtered and washed with EtOAc. Filtrate was evaporated and purification by silica gel column chromatography (5:1

DCM/acetone) afforded 3.5 g (96% yield) of **4.27** as a clear oil, which formed a white solid upon freezing.

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.60 (d,  $J = 2.2$  Hz, 2H, H-2, H-6), 6.45 (t,  $J = 2.2$  Hz, 1H, H-4), 4.41 (s, 2H,  $\text{CH}_2$ ), 1.23 (s, 18H, *t*Bu).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  155.4 (2C, C-3, C-5), 142.2 (C-1), 118.9 (C-4), 117.7 (2C, C-2, C-6), 78.6 (2C,  $\text{C}(\text{CH}_3)_3$ ), 64.4 ( $\text{CH}_2$ ), 28.8 (6C,  $\text{C}(\text{CH}_3)_3$ ).

ESI-LRMS: Calc'd for  $[\text{C}_{15}\text{H}_{24}\text{O}_3 + \text{Na}]^+$ : 275.3432; found: 275.3.



### 1,3-Di-tert-butoxy-5-(iodomethyl)benzene (**4.28**)

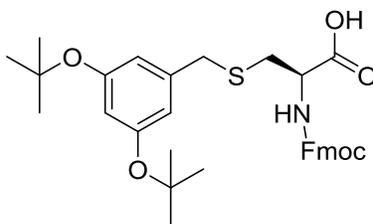
Iodine (0.5 g, 2.0 mmol) was dissolved in anhydrous DCM (60 mL) and the purple solution was stirred for 5 minutes. Triphenylphosphine (0.53 g, 2.0 mmol) and imidazole (0.15 g, 2.2 mmol) were added, the solution turned yellow and stirring was continued for 15 minutes. **4.27** (0.25 g, 1 mmol) was dissolved in DCM (30 mL) and added slowly to the reaction mixture, which was then stirred for 1 hour. Reaction was diluted with DCM, washed with a saturated solution of  $\text{Na}_2\text{S}_2\text{O}_3$  (2 x 100 mL) and brine (2 x 100 mL).

The organic layer was dried over  $\text{MgSO}_4$  and flash column chromatography on silica gel (14:1

PE/EtOAc) afforded 360 mg (90% yield) of **4.28** as a faint yellow oil that solidified upon freezing.

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.74 (d,  $J = 2.1$  Hz, 2H, H-2, H-6), 6.53 (t,  $J = 2.1$  Hz, 1H, H-4), 4.36 (s, 2H,  $\text{CH}_2$ ), 1.34 (s, 18H, *t*Bu).

$^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  156.0 (2C, C-3, C-5), 139.9 (C-1), 119.9 (2C, C-2, C-6), 119.7 (C-4), 79.0 (2C,  $\text{C}(\text{CH}_3)_3$ ), 29.0 (6C,  $\text{C}(\text{CH}_3)_3$ ), 5.7 ( $\text{CH}_2$ ).



***N*-(((9H-Fluoren-9-yl)methoxy)carbonyl)-*S*-(3,5-di-*tert*-butoxybenzyl)-*L*-cysteine (*L*-Cys\*; **4.24**)**

**4.28** (0.65 g, 1.8 mmol) and Fmoc-*L*-Cys-OH (0.49 g, 1.43 mmol) were dissolved in MeOH (30 mL).

K<sub>2</sub>CO<sub>3</sub> (0.4 g, 2.8 mmol, 2 eq.) was added and the reaction stirred for 1 hour until completion. Solution was filtered, neutralized with acetic acid and evaporated to yield an off-white foam. Solid was suspended in EtOAc, centrifuged (15 min, 4000 rpm) and the supernatant was purified by silica gel column chromatography (29:1 to 9:1 DCM/MeOH) to afford 0.56 g (68% yield) of **4.24** as a pale yellow solid.

R<sub>f</sub> = 0.5 (19:1 DCM/MeOH)

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.75 (d, *J* = 7.5 Hz, 2H, Fmoc), 7.65 – 7.57 (m, 2H, Fmoc), 7.44 – 7.34 (m, 2H, Fmoc), 7.31 (d, *J* = 7.6 Hz, 3H, Fmoc), 6.75 – 6.68 (m, 2H, *ortho*Bz), 6.57 (d, *J* = 2.4 Hz, 1H, *para*Bz), 5.78 (s, 1H, NH), 4.57 (s, 1H, CH Cys), 4.46 – 4.34 (m, 2H, CH<sub>2</sub> Fmoc), 4.24 (t, *J* = 7.3 Hz, 1H, CH Fmoc), 3.65 (s, 2H, CH<sub>2</sub> Bz), 2.98 (dd, *J* = 13.8, 4.7 Hz, 1H, CH<sub>2a</sub> Cys), 2.86 (dd, *J* = 14.3, 6.1 Hz, 1H, CH<sub>2b</sub> Cys), 1.33 (s, 18H, 2x*t*Bu).

<sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 175.1 (CO<sub>2</sub>H), 156.3 (NHCO), 155.8 (2C, *meta*Bz), 144.0 (Fmoc), 143.8 (Fmoc), 141.4 (2C, Fmoc), 138.5 (C1'Bz), 127.8 (2C, Fmoc), 127.2 (2C, Fmoc), 125.3 (2C, Fmoc), 120.3 (2C, Fmoc), 120.1 (2C, *ortho*Bz), 119.1 (*para*Bz), 79.1 (2C, 2xC(CH<sub>3</sub>)<sub>3</sub>), 67.5 (CH<sub>2</sub> Fmoc), 53.8 (CH Cys), 47.2 (CH Fmoc), 36.6 (CH<sub>2</sub> Bz), 33.3 (CH<sub>2</sub> Cys), 29.0 (6C, 2xC(CH<sub>3</sub>)<sub>3</sub>).

HRMS: Calc'd for [C<sub>33</sub>H<sub>39</sub>NO<sub>6</sub>S + Na]<sup>+</sup>: 600.2390; found: 600.2392.

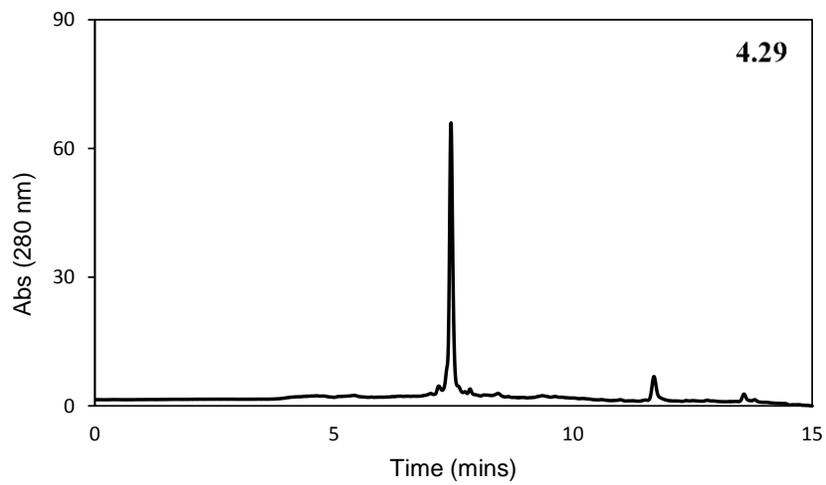
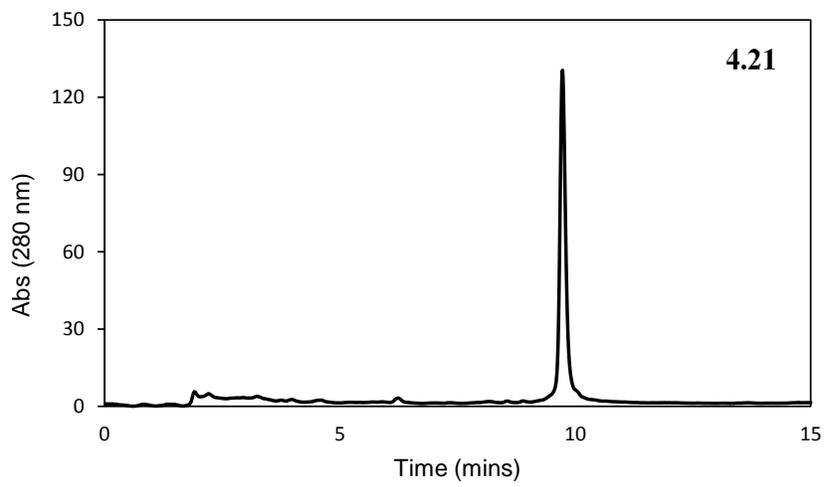
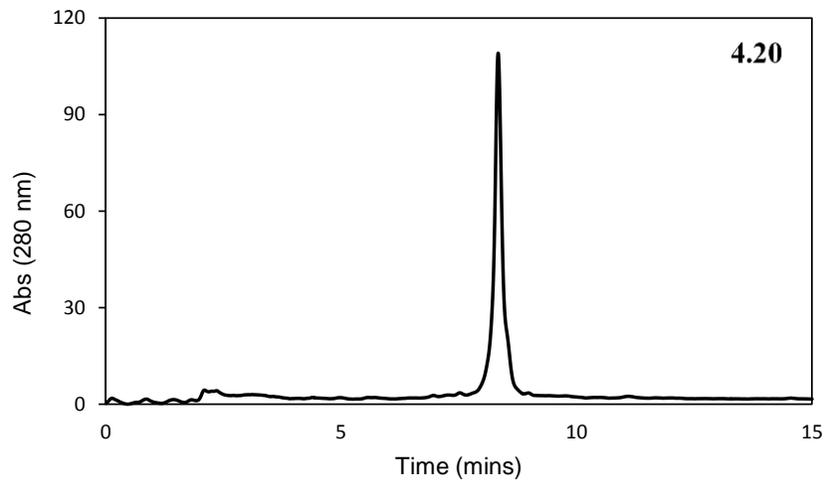
### 7.3.3 Peptide synthesis

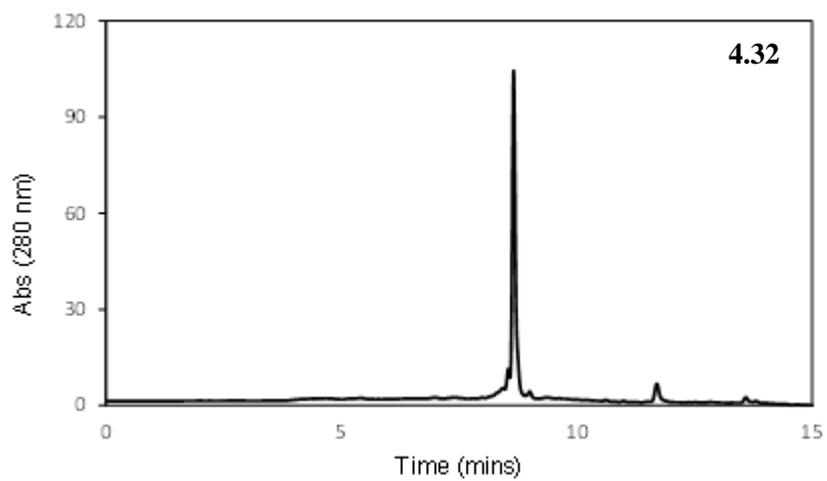
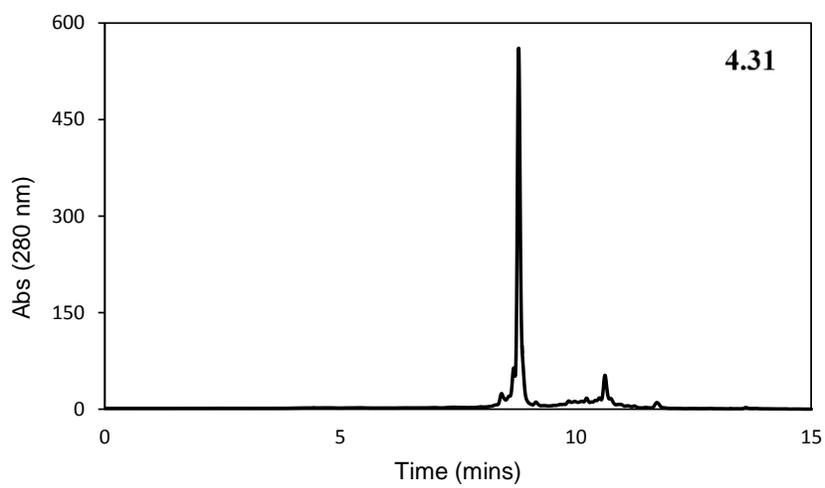
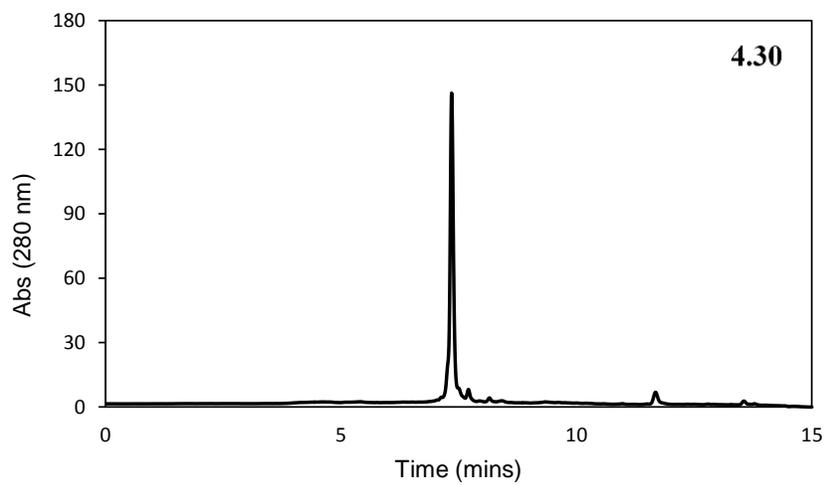
**Manual solid phase peptide synthesis:** Linear peptides **4.20** and **4.21** were synthesized on 25  $\mu\text{mol}$  scale by manual solid phase peptide synthesis. Tentagel S PHB-Gly-Fmoc (Sigma-Aldrich) resin was charged in a 10 mL polypropylene reaction vessel and soaked in DMF for 15 minutes. *Deblocking:* the solvent was drained and piperidine/DMF (1:4) was added for 2 rounds of 5-10 minutes. Resin was rinsed twice with excess DMF. *Activation:* Fmoc-protected amino acids were mixed with equimolar amounts of HBTU and *N*-methylmorpholine (NMM) for 10 minutes in an Eppendorf tube. *Coupling:* the solution of activated amino acid (1.5-5 eq.) was added to the resin and reacted for 30-120 mins. After coupling the resin was washed twice with DMF and a few beads were transferred to a 0.6 mL Eppendorf tube. *Kaiser test:* to these resin beads 2-5 drops of ninhydrin (5% w/v in EtOH), phenol (80% w/v in EtOH) and KCN (20  $\mu\text{M}$  in 50:1 pyridine/H<sub>2</sub>O) were added and the solution heated at 100 °C for 5 minutes. Colour observations (blue = positive; yellow/clear = negative) informed if free amine was present and thus whether additional rounds of coupling were needed/performed. Once each coupling step was complete the cycle (deblocking, activation, coupling) was repeated for each amino acid until peptide synthesis was finished. Peptide resin was washed three times with DMF, three times with DCM and the solvent evaporated overnight on high-vacuum pump. Peptide-loaded resins were stored at 4 °C until needed.

**Automated solid phase peptide synthesis:** Linear peptides **4.29-4.32** were synthesized on 5  $\mu\text{mol}$ /well scale using the MultipPep (Intravis AG, Bioanalytical Instruments) automated parallel solid phase peptide synthesis (SPPS) instrument. Wells were loaded in quadruplicate for each peptide target with Tentagel S PHB-Gly-Fmoc resin. Beads were soaked in DMF for 15 minutes prior to synthesis. Automated protocols used piperidine/DMF (1:4) in double deblocking steps for 5 minutes each followed by washing with excess DMF. Fmoc-amino acids were separately reacted with HBTU and NMM for 15 minutes immediately prior to coupling. Coupling with activated amino acid solutions (5 eq.) were performed twice

per residue for 10 and 20 minutes respectively. Iterative SPPS cycles were performed until all amino acids were added. Resin-loaded peptides were subject to deprotection/cleavage upon completion.

**General deprotection procedure:** To peptide-loaded resin, freshly prepared cleavage cocktail B (88:5:5:2 TFA/H<sub>2</sub>O/phenol/TIPS) was added and mixed for 90-120 minutes. Cleavage cocktail was diluted 10-fold into ice-cold anhydrous ether, yielding white precipitate. Solid was collected by centrifugation (10 min at 5,000 rpm), washed 2-3 times with fresh, cooled anhydrous ether and the solvent evaporated under a stream of nitrogen and then overnight on a high-vacuum pump. The crude peptides were dissolved in DMSO and injected onto semi-preparative C18 reverse phase high performance liquid chromatography (HPLC) for purification. Samples were fractionated based on their strong absorbance at 214 nm and 280 nm and the resultant fractions were assessed by MALDI-TOF analysis to identify the peptide species. Product-containing fractions were concentrated with the SpeedVac and then lyophilized to yield the peptidic products as white solids. HPLC-purified peptides were injected on the Q-TOF LC/MS system to obtain the HRMS characterization.





4.20: Calc'd for  $[C_{106}H_{126}N_{24}O_{28}S + 3H]^{+3}$ : 739.3043; found: 739.3029.

4.21: Calc'd for  $[C_{99}H_{119}N_{19}O_{23}S + 3H]^{+3}$ : 658.9560; found: 658.9577.

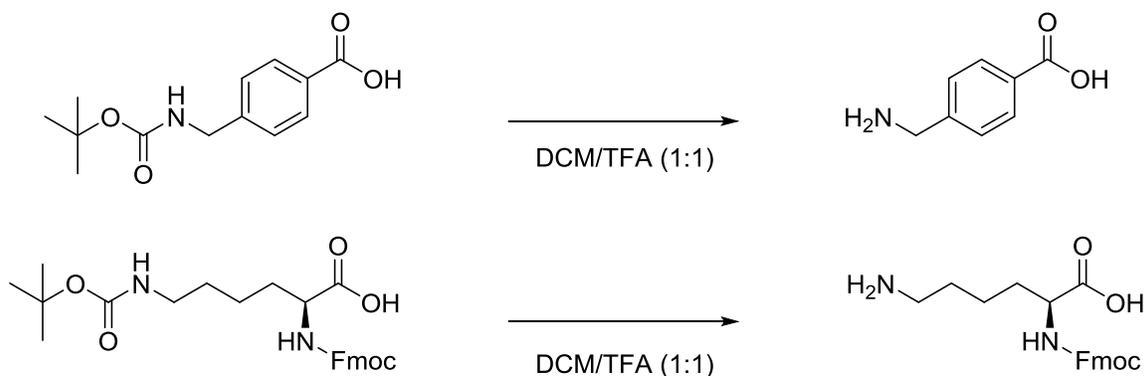
4.29: Calc'd for  $[C_{101}H_{124}N_{24}O_{22}S + 3H]^{+3}$ : 686.6426; found: 686.6471.

4.30: Calc'd for  $[C_{105}H_{124}N_{24}O_{23} + 3H]^{+3}$ : 697.3168; found: 697.3206.

4.31: Calc'd for  $[C_{94}H_{117}N_{19}O_{18}S + 3H]^{+3}$ : 611.6260; found: 611.6264.

4.32: Calc'd for  $[C_{98}H_{117}N_{19}O_{19} + 3H]^{+3}$ : 622.3002; found: 622.3042.

### 7.3.4 Amino acid coupling with DSG



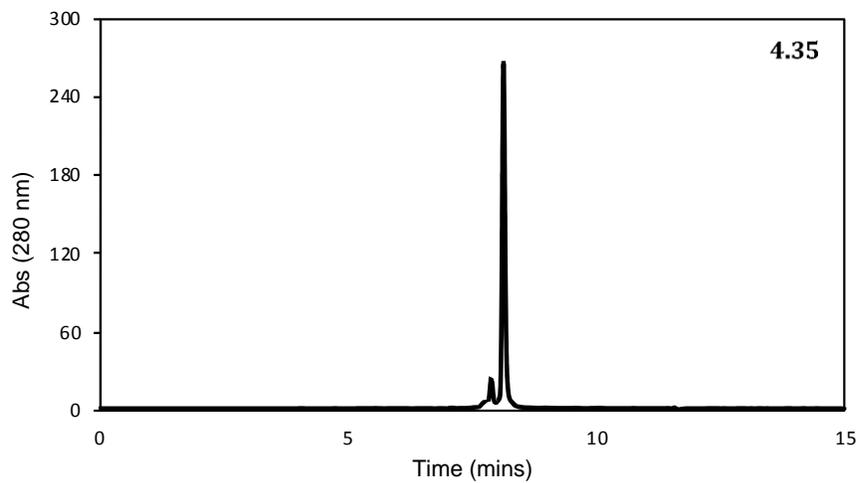
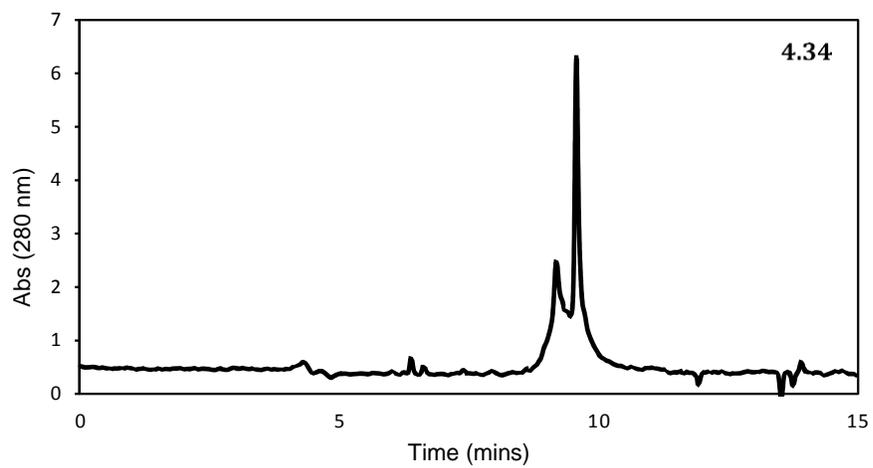
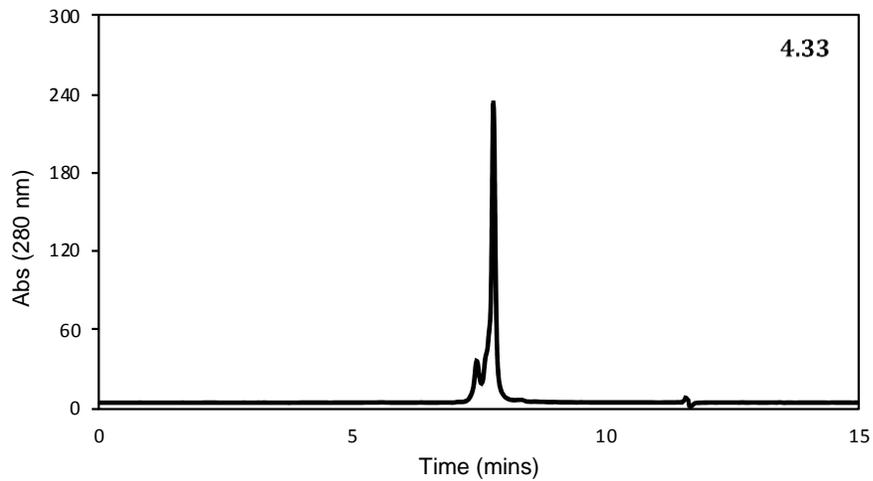
**General procedure for *N*-Boc deprotection:** *N*-Fmoc-amino acids bearing *N*-Boc protection (Fmoc-Lys(Boc)-OH or 4-(*N*-boc-aminomethyl)benzoic acid) were dissolved in 1:1 DCM/TFA and stirred for 2 hours at room temperature. Solvent was evaporated, product was extracted with diethyl ether and washed with water to yield the desired product in quantitative yield.  $^1H$  and  $^{13}C$  NMR signals matched reported values<sup>206</sup>.

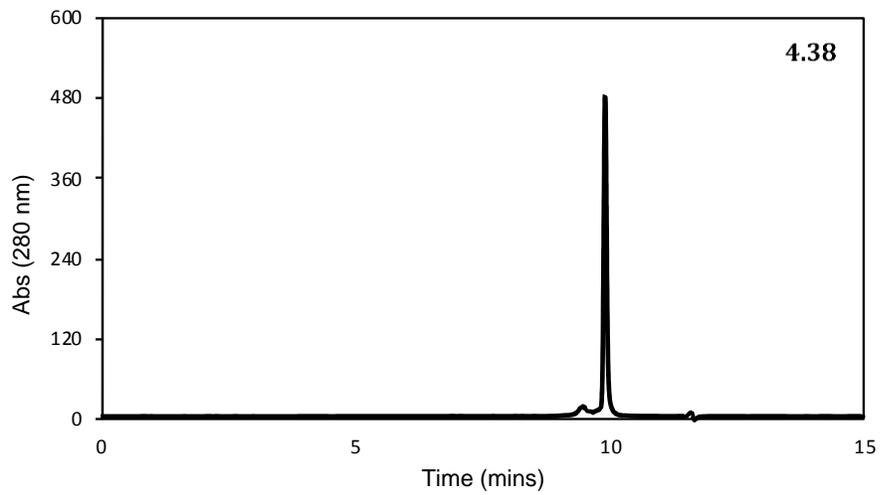
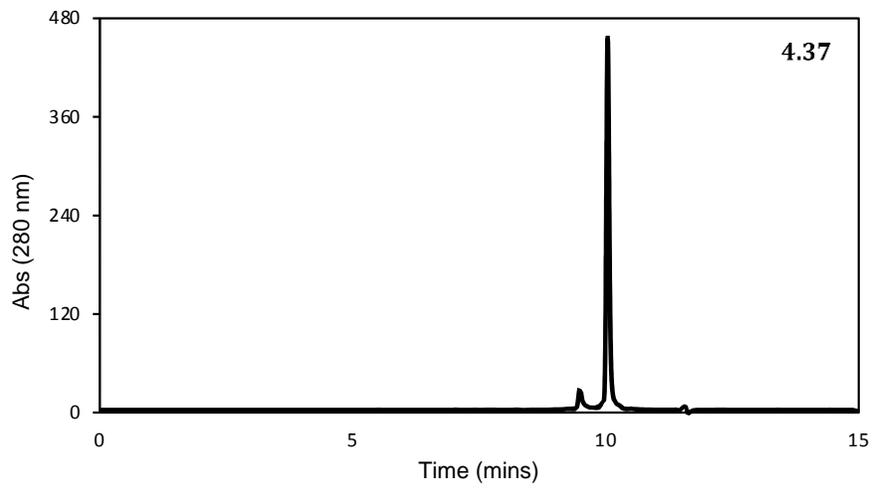
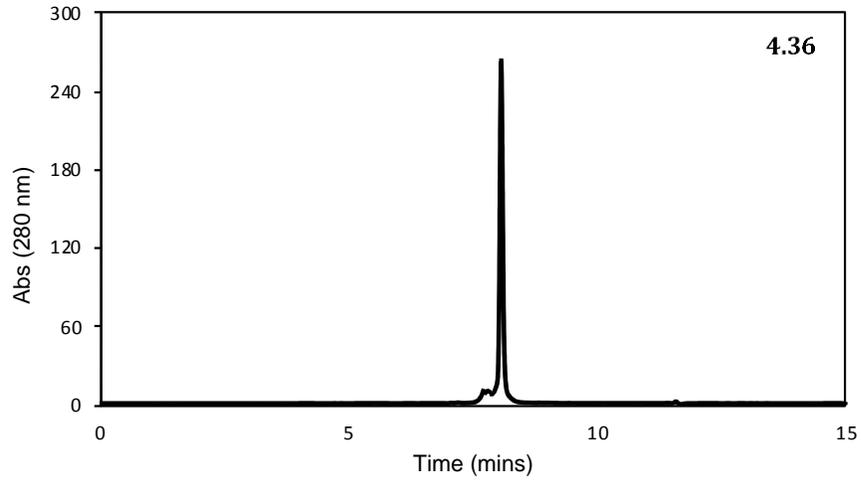
**Amino acid reactivity with disuccinimidyl glutarate:** 4-(Aminomethyl)benzoic acid, Fmoc-Lys-OH and Fmoc-Dopa-OH were each dissolved in DMF to a concentration of 10 mM and aliquoted into three in glass scintillation vials each. A solution of disuccinimidyl glutarate (0.55 eq.) in DMF was added to each

of the nine vials and the reactions were initiated by addition of base. Three organic bases 2,6-lutidine, 2,4,6-collidine or diisopropylethylamine (DIPEA) were added in 10 eq., 10 eq. and 5 eq. amounts respectively. Reaction mixtures were stirred at room temperature and aliquots of the reaction mixtures were obtained between 0 and 72 hours. Aliquots were immediately quenched with TFA and flash frozen at -78 °C. Upon completion of the experiment all aliquots were thawed and injected onto the Ultra High Performance Liquid Chromatography (uHPLC) Mass Spectrometer (MS) instrument for analysis. UV/Vis traces at  $A_{280}$  nm (Lys, Dopa) and  $A_{240}$  nm (AMB) revealed peaks corresponding to the expected m/z for starting material, reaction intermediates and products. Each peak was integrated and the relatively amount of each species plotted to examine the base-assisted reactivity between amino acids and DSG.

### 7.3.5 Peptide cyclization and purification

**General procedure:** Disuccinimidyl glutarate (2 mM; 1.1 eq.) and linear peptides (2 mM) were dissolved in DMF and collidine (20 mM; 10 eq.) was added to initiate the reaction. After 72 hours the reaction was quenched with trifluoroacetic acid and purified using semi-preparative C18 reverse-phase high performance liquid chromatography (HPLC). Fractions were collected by peak absorbance at  $A_{280}$ , MALDI-TOF analysis confirmed the presence of desired peptide in selected fractions. Samples were pooled from successive runs. The organic solvents were removed by SpeedVac, aqueous suspensions were then lyophilized to yield the products as white solids. Product solutions (1  $\mu$ M) were injected on the Q-TOF LC/MS system to obtain the HRMS characterization.





### High-resolution mass spectrometry of cyclic peptides:

**4.33:** Calc'd for  $[C_{111}H_{130}N_{24}O_{30}S + 3H]^{+3}$ : 771.3113; found: 771.3100.

**4.34:** Calc'd for  $[C_{104}H_{123}N_{19}O_{25}S + 3H]^{+3}$ : 690.9631; found: 690.9614.

**4.35:** Calc'd for  $[C_{106}H_{128}N_{24}O_{24}S + 3H]^{+3}$ : 718.6496; found: 718.6474.

**4.36:** Calc'd for  $[C_{110}H_{128}N_{24}O_{25} + 3H]^{+3}$ : 729.3239; found: 729.3251.

**4.37:** Calc'd for  $[C_{99}H_{121}N_{19}O_{20}S + 2H]^{+2}$ : 964.9456; found: 964.9420.

**4.38:** Calc'd for  $[C_{103}H_{121}N_{19}O_{21} + 2H]^{+2}$ : 980.9570; found: 980.9563.

### 7.3.6 Quantifying peptide stocks

Linear and cyclic peptides were weighed by difference on the analytical balance and dissolved in DMSO to a concentration of 10 mM. The stock solutions were diluted 10-fold in DMSO and then again diluted 10-fold in H<sub>2</sub>O to a concentration of 100 μM. Solutions were loaded onto the Take3 micro-volume plate in triplicate (2 μL/well) and the absorbance at A<sub>280</sub> nm measured using the Synergy HTX multi-mode plate reader (BioTek). Absorbance values were corrected by the average background A<sub>280</sub> nm value of a 10% aqueous solution of DMSO. The concentrations (c) were determined as a function of UV absorbance (A) at λ = 280, the extinction coefficient (ε) and the path length of the cell (l = 0.5 mm) by employing the Beer-Lambert law (Equation 7.1).

#### Equation 7.1. Beers-Lambert law.

$$A = \epsilon l c$$

The extinction coefficient of each peptide was determined by calculation (Equation 7.2) using a combination of literature values for extinction coefficients at A<sub>280</sub> nm for amino acids (Tyr = 1490 M<sup>-1</sup>cm<sup>-1</sup>, Trp = 5500 M<sup>-1</sup>cm<sup>-1</sup>, Dopa = M<sup>-1</sup>cm<sup>-1</sup>) and chemically similar moieties (resorcinol = 1620 M<sup>-1</sup>cm<sup>-1</sup> as a surrogate for Cys\*).

**Equation 7.2. Determination of peptide extinction coefficient.**

$$\varepsilon_{total} = \varepsilon_{Trp}(n_{Trp}) + \varepsilon_{Tyr}(n_{Tyr}) + \varepsilon_{AMB}(n_{AMB}) + \varepsilon_{Dopa}(n_{Dopa}) + \varepsilon_{Cys*}(n_{Cys*})$$

The extinction coefficient value can be estimated for each sequence from the extinction coefficient of each of the constituent amino acid residues that have significant absorbance at  $\lambda = 280$  nm (Equation 7.2). The cumulative sum of the occurrence (n) and extinction coefficient ( $\varepsilon$ ) of each contributing amino acid yields the approximated value ( $\varepsilon_{total}$ ) for each peptide (Table 7.2).

**Table 7.2. Occurrence (n) and calculated extinction coefficients ( $\varepsilon_{total}$ ) for peptides by UV/Vis spectroscopy at  $\lambda_{280}$  nm.**

Peptide	n <sub>Dopa</sub>	n <sub>Tyr</sub>	n <sub>Trp</sub>	n <sub>Cys*</sub>	$\varepsilon_{total}$ (M <sup>-1</sup> cm <sup>-1</sup> )
<b>4.20/4.33</b>	4	-	1	1	20,510
<b>4.21/4.34</b>	3	-	2	1	17,640
<b>4.29/4.35</b>	-	4	1	-	11,460
<b>4.30/4.36</b>	-	5	1	-	12,950
<b>4.31/4.37</b>	-	3	2	-	15,470
<b>4.32/4.38</b>	-	4	2	-	16,960

Concentration values as determined by UV/Vis were lower than the concentrations determined by weight (Table 7.3), yet values were generally of the same order of magnitude. Therefore, the UV/Vis assay was primarily used to confirm the approximate concentration (10 mM) of weighed samples. However, for **4.34**, concentration determination by UV/Vis spectroscopy indicated a significantly (50-fold) lower value than the one calculated by weight. The stock concentration of **4.34** was therefore adjusted from 10 mM to 0.2 mM for kinetic parameter determination.

**Table 7.3. Calculated concentration of peptide stocks by UV/Vis spectroscopy at A<sub>280</sub> nm.**

<b>Linear Peptide*</b>	<b>Concentration (A<sub>280</sub>)</b>	<b>Cyclic Peptide*</b>	<b>Concentration (A<sub>280</sub>)</b>
<b>4.20</b>	6.6 mM	<b>4.33</b>	4.4 mM
<b>4.21</b>	5.2 mM	<b>4.34</b>	0.2 mM
<b>4.29</b>	5.8 mM	<b>4.35</b>	4.9 mM
<b>4.30</b>	6.0 mM	<b>4.36</b>	3.9 mM
<b>4.31</b>	5.1 mM	<b>4.37</b>	7.2 mM
<b>4.32</b>	6.1 mM	<b>4.38</b>	6.6 mM

\*All peptide stocks initially determined to be 10 mM by weight.

### 7.3.7 Enzyme kinetics

**Direct assay:** 96-Well reduced volume plates (Costar) were charged with a buffered solution of TcTS (44 nM; 20 mM Tris, 30 mM NaCl, 0.1% BSA, pH = 7.4) and incubated at 25 °C. Peptides were dissolved in DMSO and were subsequently added to the enzyme solution (10 µL/well) and incubated for 5-10 minutes. Assay was initiated by addition of TFMU-SA (0.5 mM) and the initial rates were monitored in the plate reader by UV/Vis at A<sub>405</sub> nm for 15 minutes. Assay plates contained eight wells of both positive controls (containing equal volumes of DMSO in place of peptide species) and negative controls (absence of enzyme). Initial rates for each well were determined by linear regression and the percent residual TcTS activity determined by comparison to the average of the background subtracted positive control samples.

**Coupled enzyme assay development:** The two-enzyme (TcTS/Abg) coupled assay was designed by first selecting a concentration of Abg (100 nM) that cleaved Lac-MU at an appropriate initial rate (~0.1 ΔAbs/min). Next, TcTS was varied at lower concentration (from 10-60 nM). Initial rate of Sia-Lac-MU cleavage was monitored, and rates increased proportional to TcTS concentration across the range. Sia-Lac-MU was then added in varying concentrations from 10-1000 µM to cuvettes preincubated with TcTS (40 nM) and Abg (100 nM) at 25 °C. Initial rates were plotted against concentration of substrate and the Michaelis constant determined  $K_m = 145 \mu\text{M}$ . Additionally, Abg (100 nM) was incubated with Sia-Lac-MU (100 µM) for 10 minutes, no observable hydrolysis of MU (A<sub>370</sub> nm) was detected.

**Coupled enzyme inhibition assay:** Quartz cuvettes were charged with a buffered solution (50 mM HEPES, pH 7.45) of TcTS (40 nM) and Abg (100 nM) at 25 °C. Inhibitor solution (100 μM in DMSO) was diluted 10-fold into buffered enzyme solution and incubated for 5-10 minutes. Assay was initiated by addition of Sia-Lac-MU (500 μM) and the initial rate of cleavage was monitored at  $A_{370}$  nm using the Cary4000 UV/Vis spectrophotometer. In parallel, control cuvettes were incubated with 10% DMSO (20 μL) and the initial rates determined. Percent TcTS activity in the presence of the inhibitors was calculated as the ratio of the initial rate of sample well over the initial rate of the positive control.

**$K_i$  determination:** The inhibition constant ( $K_i$ ) for **4.33** was determined. Quartz cuvettes charged with buffered TcTS (45 nM; 20 mM Tris, 30 mM NaCl, pH = 7.4) and were incubated with **4.33** at a range of concentrations (0.1-5 μM). Enzyme inhibition was assayed at three concentrations of TFMU-SA substrate (0.25, 0.5, 1 mM) and the initial rate of enzymatic hydrolysis measured by UV/Vis at 385 nm for 10 minutes. Initial rates, inhibitor concentration and substrate concentrations were input into data analysis software GraFit 7.0 (Eritacus) and fit to a Dixon plot to determine the  $K_i$  value.

## 7.4 *Trypanosoma cruzi* growth inhibition assay

### 7.4.1 Standard growth inhibition assay procedure

Compounds **3.10**, **3.12-3.16** were screened against *Trypanosoma cruzi* (Tulahuen strain, DTU VI) expressing β-galactosidase using the growth inhibition assay described by the Buckner lab<sup>198,200</sup>. Briefly, assays were conducted at 37 °C and 5% CO<sub>2</sub> using RPMI-1640 (without phenol red) with 10% heat inactivated FBS, penicillin (100 u/mL) and streptomycin (100 μg/mL). Murine 3T3 fibroblast cells (~2 x 10<sup>3</sup>/well) were allowed to attach to 96-well tissues culture plates for 24 hours. Trypomastigotes (~1 x 10<sup>4</sup>/well) were added and incubated with fibroblast cells for 24 hours, then test compounds were added (triplicate serial dilutions) and the plates were incubated for 6 days. Parasite levels were quantified after

lysis using CRBG (Roche) as described<sup>207</sup>. Percent inhibition was determined by comparison to high control (3T3 cells, trypomastigotes) after subtraction of the background (3T3 cells, trypomastigotes, 50  $\mu$ M benznidazole). EC<sub>50</sub> values were calculated by non-linear regression using software by the Collaborative Drug Database (Burlingame, CA. [www.collaboratedrug.com](http://www.collaboratedrug.com)).

#### **7.4.2 Modified growth inhibition assay procedure**

Compounds **3.10**, **3.12-3.16** were assessed using the *Trypanosoma cruzi* growth inhibition assay as described above with a single modification. The 24-hour incubation period between *T. cruzi* and 3T3 cells (as described in 7.4.1) was removed. Instead, test compounds were added to 3T3 cells followed by addition of parasite cells.

## References

- (1) Lee, B. Y.; Bacon, K. M.; Bottazzi, M. E.; Hotez, P. J. Global Economic Burden of Chagas Disease: A Computational Simulation Model. *Lancet Infect. Dis.* **2013**.
- (2) Martinez, F.; Perna, E.; Perrone, S. V.; Liprandi, A. S. Chagas Disease and Heart Failure: An Expanding Issue Worldwide. *Eur. Cardiol. Rev.* **2019**, *14* (2), 82–88. <https://doi.org/10.15420/ecr.2018.30.2>.
- (3) Albajar-Vinãs, P.; Dias, J. C. P. Advancing the Treatment for Chagas' Disease. *N. Engl. J. Med.* **2014**, *370* (20), 1942–1943. <https://doi.org/10.1056/NEJMe1403689>.
- (4) Pereira, K. S.; Schmidt, F. L.; Guaraldo, A. M. A.; Franco, R. M. B.; Dias, V. L.; Passos, L. A. C. Chagas' Disease as a Foodborne Illness. *J. Food Prot.* **2009**, *72* (2), 441–446. <https://doi.org/10.4315/0362-028x-72.2.441>.
- (5) CDC - Chagas Disease - Biology <https://www.cdc.gov/parasites/chagas/biology.html> (accessed Jan 29, 2020).
- (6) Melton, L. Spanish Group Defeats Shkreli in Chagas Voucher Race. *Nat. Biotechnol.* **2017**, *35* (11), 1003–1003. <https://doi.org/10.1038/nbt1117-1003>.
- (7) Boiani, M.; Piacenza, L.; Hernández, P.; Boiani, L.; Cerecetto, H.; González, M.; Denicola, A. Mode of Action of Nifurtimox and N-Oxide-Containing Heterocycles against *Trypanosoma Cruzi*: Is Oxidative Stress Involved? *Biochem. Pharmacol.* **2010**, *79* (12), 1736–1745. <https://doi.org/10.1016/j.bcp.2010.02.009>.
- (8) Wilkinson, S. R.; Taylor, M. C.; Horn, D.; Kelly, J. M.; Cheeseman, I. A Mechanism for Cross-Resistance to Nifurtimox and Benznidazole in Trypanosomes. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (13), 5022–5027. <https://doi.org/10.1073/pnas.0711014105>.
- (9) McCalla, D. R.; Reuvers, A.; Kaiser, C. Breakage of Bacterial DNA by Nitrofurans Derivatives. *Cancer Res.* **1971**, *31* (12), 2184–2188.
- (10) Streeter, A. J.; Hoener, B. Evidence for the Involvement of a Nitrenium Ion in the Covalent Binding of Nitrofurazone to DNA. *Pharm. Res. An Off. J. Am. Assoc. Pharm. Sci.* **1988**, *5* (7), 434–436. <https://doi.org/10.1023/A:1015988401601>.
- (11) Sgambatti de Andrade, A. L. S.; Zicker, F.; de Oliveira, R. M.; Almeida e Silva, S.; Luquetti, A.; Travassos, L. R.; Almeida, I. C.; de Andrade, S. S.; Guimarães de Andrade, J.; Martelli, C. M. Randomised Trial of Efficacy of Benznidazole in Treatment of Early *Trypanosoma Cruzi* Infection. *Lancet* **1996**, *348* (9039), 1407–1413. [https://doi.org/10.1016/S0140-6736\(96\)04128-1](https://doi.org/10.1016/S0140-6736(96)04128-1).
- (12) Estani, S. S.; Segura, E. L.; Ruiz, A. M.; Velazquez, E.; Porcel, B. M.; Yampotis, C. Efficacy of Chemotherapy with Benznidazole in Children in the Indeterminate Phase of Chagas' Disease. *Am. J. Trop. Med. Hyg.* **1998**, *59* (4), 526–529. <https://doi.org/10.4269/ajtmh.1998.59.526>.
- (13) Polak, A.; Richle, R. Mode of Action of the 2-Nitroimidazole Derivative Benznidazole. *Ann. Trop. Med. Parasitol.* **1978**, *72* (1), 45–54. <https://doi.org/10.1080/00034983.1978.11719278>.

- (14) Rajão, M. A.; Furtado, C.; Alves, C. L.; Passos-Silva, D. G.; De Moura, M. B.; Schamber-Reis, B. L.; Kunrath-Lima, M.; Zuma, A. A.; Vieira-da-Rocha, J. P.; Garcia, J. B. F.; Mendes, I. C.; Pena, S. D. J.; Macedo, A. M.; Franco, G. R.; De Souza-Pinto, N. C.; De Medeiros, M. H. G.; Cruz, A. K.; Motta, M. C. M.; Teixeira, S. M. R.; Machado, C. R. Unveiling Benznidazole's Mechanism of Action through Overexpression of DNA Repair Proteins in *Trypanosoma Cruzi*. *Environ. Mol. Mutagen.* **2014**, *55* (4), 309–321. <https://doi.org/10.1002/em.21839>.
- (15) Jackson, Y.; Alirol, E.; Getaz, L.; Wolff, H.; Combescure, C.; Chappuis, F. Tolerance and Safety of Nifurtimox in Patients with Chronic Chagas Disease. *Clin. Infect. Dis.* **2010**, *51* (10), 69–75. <https://doi.org/10.1086/656917>.
- (16) Barr, S. C.; Warner, K. L.; Kornreic, B. G.; Piscitelli, J.; Wolfe, A.; Benet, L.; McKerrow, J. H. A Cysteine Protease Inhibitor Protects Dogs from Cardiac Damage during Infection by *Trypanosoma Cruzi*. *Antimicrob. Agents Chemother.* **2005**, *49* (12), 5160–5161. <https://doi.org/10.1128/AAC.49.12.5160-5161.2005>.
- (17) Engel, J. C.; Doyle, P. S.; Hsieh, I.; McKerrow, J. H. Cysteine Protease Inhibitors Cure an Experimental *Trypanosoma Cruzi* Infection. *J. Exp. Med.* **1998**, *188* (4), 725–734. <https://doi.org/10.1084/jem.188.4.725>.
- (18) Montalvetti, A.; Bailey, B. N.; Martin, M. B.; Severin, G. W.; Oldfield, E.; Docampo, R. Bisphosphonates Are Potent Inhibitors of *Trypanosoma Cruzi* Farnesyl Pyrophosphate Synthase. *J. Biol. Chem.* **2001**, *276* (36), 33930–33937. <https://doi.org/10.1074/jbc.M103950200>.
- (19) Urbina, J. A. Ergosterol Biosynthesis and Drug Development for Chagas Disease. *Mem. Inst. Oswaldo Cruz* **2009**, *104*, 311–318. <https://doi.org/10.1590/s0074-02762009000900041>.
- (20) Urbina, J. A. Specific Chemotherapy of Chagas Disease: Relevance, Current Limitations and New Approaches. *Acta Trop.* **2010**, *115* (1–2), 55–68. <https://doi.org/10.1016/j.actatropica.2009.10.023>.
- (21) Molina, I.; Gómez I Prat, J.; Salvador, F.; Treviño, B.; Sulleiro, E.; Serre, N.; Pou, D.; Roure, S.; Cabezos, J.; Valerio, L.; Blanco-Grau, A.; Sánchez-Montalva, A.; Vidal, X.; Pahissa, A. Randomized Trial of Posaconazole and Benznidazole for Chronic Chagas' Disease. *N. Engl. J. Med.* **2014**, *370* (20), 1899–1908. <https://doi.org/10.1056/NEJMoa1313122>.
- (22) Varki, A.; Schauer, R. Chapter 14: Sialic Acids. In *Essentials of Glycobiology - second edition*; 2014; pp 1–17.
- (23) Vimr, E.; Lichtensteiger, C. To Sialylate, or Not to Sialylate: That Is the Question. *Trends in Microbiology*. June 1, 2002, pp 254–257. [https://doi.org/10.1016/S0966-842X\(02\)02361-2](https://doi.org/10.1016/S0966-842X(02)02361-2).
- (24) Chen, X.; Varki, A. Advances in the Biology and Chemistry of Sialic Acids. *ACS Chemical Biology*. February 19, 2010, pp 163–176. <https://doi.org/10.1021/cb900266r>.
- (25) Varki, A.; Gagneux, P. Multifarious Roles of Sialic Acids in Immunity. *Ann. N. Y. Acad. Sci.* **2012**, *1253* (1), 16–36. <https://doi.org/10.1111/j.1749-6632.2012.06517.x>.
- (26) Varki, A. Sialic Acids as Ligands in Recognition Phenomena. *FASEB Journal*. FASEB 1997, pp 248–255. <https://doi.org/10.1096/fasebj.11.4.9068613>.

- (27) Schauer, R. Achievements and Challenges of Sialic Acid Research. *Glycoconjugate Journal*. 2000, pp 485–499. <https://doi.org/10.1023/A:1011062223612>.
- (28) Neu, U.; Bauer, J.; Stehle, T. Viruses and Sialic Acids: Rules of Engagement. *Current Opinion in Structural Biology*. October 2011, pp 610–618. <https://doi.org/10.1016/j.sbi.2011.08.009>.
- (29) Chang, Y. C.; Nizet, V. The Interplay between Siglecs and Sialylated Pathogens. *Glycobiology*. Oxford University Press 2014, pp 818–825. <https://doi.org/10.1093/glycob/cwu067>.
- (30) Nishikawa, Y.; Ogiso, A.; Kameyama, K.; Nishimura, M.; Xuan, X.; Ikehara, Y. A2-3 Sialic Acid Glycoconjugate Loss and Its Effect on Infection with Toxoplasma Parasites. *Exp. Parasitol.* **2013**, *135* (3), 479–485. <https://doi.org/10.1016/j.exppara.2013.08.009>.
- (31) Previato, J.; Andrade, A. F. B.; Pessolani, M. C. V.; Mendonça-Previato, L. Incorporation of Sialic Acid into Trypanosoma Cruzi Macromolecules. A Proposal for a New Metabolic Route. *Mol. Biochem. Parasitol.* **1985**, *16* (1), 85–96. [https://doi.org/10.1016/0166-6851\(85\)90051-9](https://doi.org/10.1016/0166-6851(85)90051-9).
- (32) Tribulatti, M. V.; Mucci, J.; Van Rooijen, N.; Leguizamón, M. S.; Campetella, O. The Trans-Sialidase from Trypanosoma Cruzi Induces Thrombocytopenia during Acute Chagas' Disease by Reducing the Platelet Sialic Acid Contents. *Infect. Immun.* **2005**, *73* (1), 201–207. <https://doi.org/10.1128/IAI.73.1.201-207.2005>.
- (33) Leguizamón, M. S.; Mocetti, E.; García Rivello, H.; Argibay, P.; Campetella, O. Trans-Sialidase from Trypanosoma Cruzi Induces Apoptosis in Cells from the Immune System In Vivo. *J. Infect. Dis.* **1999**, *180* (4), 1398–1402. <https://doi.org/10.1086/315001>.
- (34) Schenkman, S.; Eichinger, D.; Pereira, M. E. A.; Nussenzweig, V. Structural and Functional Properties of Trypanosoma Trans-Sialidase. *Annu. Rev. Microbiol.* **1994**, *48* (1), 499–523. <https://doi.org/10.1146/annurev.mi.48.100194.002435>.
- (35) Docampo, R.; Moreno, S. N. J. The Role of Ca<sup>2+</sup> in the Process of Cell Invasion by Intracellular Parasites. *Parasitology Today*. Elsevier Ltd 1996, pp 61–65. [https://doi.org/10.1016/0169-4758\(96\)80656-9](https://doi.org/10.1016/0169-4758(96)80656-9).
- (36) Yoshida, N.; Dorta, M. L.; Ferreira, A. T.; Oshiro, M. E.; Mortara, R. A.; Acosta-Serrano, A.; Favoreto Júnior, S. Removal of Sialic Acid from Mucin-like Surface Molecules of Trypanosoma Cruzi Metacyclic Trypomastigotes Enhances Parasite-Host Cell Interaction. *Mol. Biochem. Parasitol.* **1997**, *84* (1), 57–67. [https://doi.org/10.1016/s0166-6851\(96\)02783-1](https://doi.org/10.1016/s0166-6851(96)02783-1).
- (37) Pereira, M. E. A. A Developmentally Regulated Neuraminidase Activity in Trypanosoma Cruzi. *Science* **1983**, *219* (4591), 1444–1446. <https://doi.org/10.1126/science.6338592>.
- (38) Watts, A. G.; Damager, I.; Amaya, M. L.; Buschiazzi, A.; Alzari, P.; Frasch, A. C.; Withers, S. G. Trypanosoma Cruzi Trans-Sialidase Operates through a Covalent Sialyl-Enzyme Intermediate: Tyrosine Is the Catalytic Nucleophile. *J. Am. Chem. Soc.* **2003**, *125* (25), 7532–7533. <https://doi.org/10.1021/ja0344967>.
- (39) Yang, J.; Schenkman, S.; Horenstein, B. A. Primary <sup>13</sup>C and β-Secondary <sup>2</sup>H KIEs for Trans-Sialidase. A Snapshot of Nucleophilic Participation during Catalysis. *Biochemistry* **2000**, *39*, 5902. <https://doi.org/10.1021/BI000061+>.

- (40) Buschiazzo, A.; Amaya, M. F.; Cremona, M. L.; Frasch, A. C.; Alzari, P. M. The Crystal Structure and Mode of Action of Trans-Sialidase, a Key Enzyme in Trypanosoma Cruzi Pathogenesis. *Mol. Cell* **2002**, *10* (4), 757–768. [https://doi.org/10.1016/S1097-2765\(02\)00680-9](https://doi.org/10.1016/S1097-2765(02)00680-9).
- (41) Amaya, M. F.; Watts, A. G.; Damager, I.; Wehenkel, A.; Nguyen, T.; Buschiazzo, A.; Paris, G.; Frasch, A. C.; Withers, S. G.; Alzari, P. M. Structural Insights into the Catalytic Mechanism of Trypanosoma Cruzi Trans-Sialidase. *Structure* **2004**, *12* (5), 775–784. <https://doi.org/10.1016/j.str.2004.02.036>.
- (42) Vocadlo, D. J.; Davies, G. J. Mechanistic Insights into Glycosidase Chemistry. *Current Opinion in Chemical Biology*. October 2008, pp 539–555. <https://doi.org/10.1016/j.cbpa.2008.05.010>.
- (43) Damager, I.; Buchini, S.; Amaya, M. F.; Buschiazzo, A.; Alzari, P.; Frasch, A. C.; Watts, A.; Withers, S. G. Kinetic and Mechanistic Analysis of Trypanosoma Cruzi Trans-Sialidase Reveals a Classical Ping-Pong Mechanism with Acid/Base Catalysis. *Biochemistry* **2008**, *47* (11), 3507–3512. <https://doi.org/10.1021/bi7024832>.
- (44) Koshland, D. E. Stereochemistry And The Mechanism Of Enzymatic Reactions. *Biol. Rev.* **1953**, *28* (4), 416–436. <https://doi.org/10.1111/j.1469-185X.1953.tb01386.x>.
- (45) Arner, E.; Kindlund, E.; Nilsson, D.; Farzana, F.; Ferella, M.; Tammi, M. T.; Andersson, B. Database of Trypanosoma Cruzi Repeated Genes: 20 000 Additional Gene Variants. *BMC Genomics* **2007**, *8*. <https://doi.org/10.1186/1471-2164-8-391>.
- (46) Cremona, M. L.; Sánchez, D. O.; Frasch, A. C. C.; Campetella, O. A Single Tyrosine Differentiates Active and Inactive Trypanosoma Cruzi Trans-Sialidases. *Gene* **1995**, *160* (1), 123–128. [https://doi.org/10.1016/0378-1119\(95\)00175-6](https://doi.org/10.1016/0378-1119(95)00175-6).
- (47) Shidmoosavee, F. S.; Watson, J. N.; Bennet, A. J. Chemical Insight into the Emergence of Influenza Virus Strains That Are Resistant to Relenza. *J. Am. Chem. Soc.* **2013**, *135* (36), 13254–13257. <https://doi.org/10.1021/ja405916q>.
- (48) Amaya, M. F.; Buschiazzo, A.; Nguyen, T.; Alzari, P. M. The High Resolution Structures of Free and Inhibitor-Bound Trypanosoma Rangeli Sialidase and Its Comparison with T. Cruzi Trans-Sialidase. *J. Mol. Biol.* **2003**, *325* (4), 773–784. [https://doi.org/10.1016/S0022-2836\(02\)01306-2](https://doi.org/10.1016/S0022-2836(02)01306-2).
- (49) Agustí, R.; París, G.; Ratier, L.; Frasch, A. C. C.; de Lederkremer, R. M. Lactose Derivatives Are Inhibitors of Trypanosoma Cruzi Trans-Sialidase Activity toward Conventional Substrates in Vitro and in Vivo. *Glycobiology* **2004**, *14* (7), 659–670. <https://doi.org/10.1093/glycob/cwh079>.
- (50) Mucci, J.; Risso, M. G.; Leguizamón, M. S.; Frasch, A. C. C.; Campetella, O. The Trans-Sialidase from Trypanosoma Cruzi Triggers Apoptosis by Target Cell Sialylation. *Cell. Microbiol.* **2006**, *8* (7), 1086–1095. <https://doi.org/10.1111/j.1462-5822.2006.00689.x>.
- (51) Giorgi, M. E.; Ratier, L.; Agustí, R.; Frasch, A. C. C.; De Lederkremer, R. M. Synthesis of PEGylated Lactose Analogs for Inhibition Studies on T.Cruzi Trans-Sialidase. *Glycoconj. J.* **2010**, *27* (5), 549–559. <https://doi.org/10.1007/s10719-010-9300-7>.
- (52) Giorgi, M. E.; Ratier, L.; Agustí, R.; Frasch, A. C. C.; De Lederkremer, R. M. Improved Bioavailability of Inhibitors of Trypanosoma Cruzi Trans-Sialidase: PEGylation of Lactose

- Analogs with Multiarm Polyethyleneglycol. *Glycobiology* **2012**, *22* (10), 1363–1373. <https://doi.org/10.1093/glycob/cws091>.
- (53) Agustí, R.; Giorgi, M. E.; Mendoza, V. M.; Gallo-Rodriguez, C.; de Lederkremer, R. M. Comparative Rates of Sialylation by Recombinant Trans-Sialidase and Inhibitor Properties of Synthetic Oligosaccharides from Trypanosoma Cruzi Mucins-Containing Galactofuranose and Galactopyranose. *Bioorganic Med. Chem.* **2007**, *15* (7), 2611–2616. <https://doi.org/10.1016/j.bmc.2007.01.045>.
- (54) Carvalho, I.; Andrade, P.; Campo, V. L.; Guedes, P. M. M.; Sesti-Costa, R.; Silva, J. S.; Schenkman, S.; Dedola, S.; Hill, L.; Rejzek, M.; Nepogodiev, S. A.; Field, R. A. “Click Chemistry” Synthesis of a Library of 1,2,3-Triazole-Substituted Galactose Derivatives and Their Evaluation against Trypanosoma Cruzi and Its Cell Surface Trans-Sialidase. *Bioorganic Med. Chem.* **2010**, *18* (7), 2412–2427. <https://doi.org/10.1016/j.bmc.2010.02.053>.
- (55) Silva, B. L.; S. Filho, J. D.; Andrade, P.; Carvalho, I.; Alves, R. J. Design, Synthesis and Enzymatic Evaluation of 3-O-Substituted Aryl  $\beta$ -d-Galactopyranosides as Inhibitors of Trypanosoma Cruzi Trans-Sialidase. *Bioorg. Med. Chem. Lett.* **2014**, *24* (18), 4529–4532. <https://doi.org/10.1016/j.bmcl.2014.07.088>.
- (56) Freire-de-Lima, L.; Oliveira, I. a.; Neves, J. L.; Penha, L. L.; Alisson-Silva, F.; Dias, W. B.; Todeschini, A. R. Sialic Acid: A Sweet Swing between Mammalian Host and Trypanosoma Cruzi. *Front. Immunol.* **2012**, *3*, 1–12. <https://doi.org/10.3389/fimmu.2012.00356>.
- (57) Streicher, H.; Busse, H. Building a Successful Structural Motif into Sialylmimetics - Cyclohexenephosphonate Monoesters as Pseudo-Sialosides with Promising Inhibitory Properties. *Bioorganic Med. Chem.* **2006**, *14* (4), 1047–1057. <https://doi.org/10.1016/j.bmc.2005.09.025>.
- (58) Neres, J.; Bonnet, P.; Edwards, P. N.; Kotian, P. L.; Buschiazzi, A.; Alzari, P. M.; Bryce, R. A.; Douglas, K. T. Benzoic Acid and Pyridine Derivatives as Inhibitors of Trypanosoma Cruzi Trans-Sialidase. *Bioorganic Med. Chem.* **2007**, *15* (5), 2106–2119. <https://doi.org/10.1016/j.bmc.2006.12.024>.
- (59) Carvalho, S. T.; Sola-Penna, M.; Oliveira, I. A.; Pita, S.; Gonçalves, A. S.; Neves, B. C.; Sousa, F. R.; Freire-de-Lima, L.; Kuroguchi, M.; Hinou, H.; Nishimura, S. I.; Mendonça-Previato, L.; Previato, J. O.; Todeschini, A. R. A New Class of Mechanism-Based Inhibitors for Trypanosoma Cruzi Transsialidase and Their Influence on Parasite Virulence. *Glycobiology* **2010**, *20* (8), 1034–1045. <https://doi.org/10.1093/glycob/cwq065>.
- (60) Meinke, S.; Schroven, A.; Thiem, J. Sialic Acid C-Glycosides with Aromatic Residues: Investigating Enzyme Binding and Inhibition of Trypanosoma Cruzi Trans-Sialidase. *Org. Biomol. Chem.* **2011**, *9* (12), 4487–4497. <https://doi.org/10.1039/c0ob01176b>.
- (61) Campo, V. L.; Sesti-Costa, R.; Carneiro, Z. A.; Silva, J. S.; Schenkman, S.; Carvalho, I. Design, Synthesis and the Effect of 1,2,3-Triazole Sialylmimetic Neoglycoconjugates on Trypanosoma Cruzi and Its Cell Surface Trans-Sialidase. *Bioorganic Med. Chem.* **2012**, *20* (1), 145–156. <https://doi.org/10.1016/j.bmc.2011.11.022>.
- (62) Busse, H.; Hakoda, M.; Stanley, M.; Streicher, H. Galactose-Phosphonates as Mimetics of the Sialyltransfer by Trypanosomal Sialidases. *J. Carbohydr. Chem.* **2007**, *26* (3), 159–194.

<https://doi.org/10.1080/07328300701351532>.

- (63) Harrison, J. A.; Kartha, K. P. R.; Fournier, E. J. L.; Lowary, T. L.; Malet, C.; Nilsson, U. J.; Hindsgaul, O.; Schenkman, S.; Naismith, J. H.; Field, R. A. Probing the Acceptor Substrate Binding Site of Trypanosoma Cruzi Trans-Sialidase with Systematically Modified Substrates and Glycoside Libraries. *Org. Biomol. Chem.* **2011**, *9* (5), 1653–1660. <https://doi.org/10.1039/c0ob00826e>.
- (64) Montero, E.; Vallmitjana, M.; Pérez-Pons, J. A.; Querol, E.; Jiménez-Barbero, J.; Cañada, F. J. NMR Studies of the Conformation of Thiocellobiose Bound to a  $\beta$ -Glucosidase from Streptomyces Sp. *FEBS Lett.* **1998**, *421* (3), 243–248. [https://doi.org/10.1016/S0014-5793\(97\)01571-8](https://doi.org/10.1016/S0014-5793(97)01571-8).
- (65) Howard, S.; Withers, S. G. Bromoketone C-Glycosides, a New Class of  $\beta$ -Glucanase Inactivators. *J. Am. Chem. Soc.* **1998**, *120* (40), 10326–10331. <https://doi.org/10.1021/ja981580r>.
- (66) Rempel, B. P.; Withers, S. G. Covalent Inhibitors of Glycosidases and Their Applications in Biochemistry and Biology. *Glycobiology*. 2008, pp 570–586. <https://doi.org/10.1093/glycob/cwn041>.
- (67) Withers, S. G.; Rupitz, K.; Street, I. P. 2-Deoxy-2-Fluoro-D-Glycosyl Fluorides. *J. Biol. Chem.* **1988**, *263* (17), 17–20.
- (68) Kim, J.-H.; Resende, R.; Wennekes, T.; Chen, H.-M.; NBance, N.; Buchini, S.; Watts, A. G.; Pilling, P.; Streltsov, V. A.; Petric, M.; Liggins, R.; Barrett, S.; McKimm-Breschkin, J. L.; Niikura, M.; Withers, S. G. Mechanism-Based Covalent Neuraminidase. *Science* **2013**, *340* (April), 71–75. <https://doi.org/10.1126/science.1232552>.
- (69) Gao, Z.; Niikura, M.; Withers, S. G. Ultrasensitive Fluorogenic Reagents for Neuraminidase Titration. *Angew. Chemie - Int. Ed.* **2017**, *56* (22), 6112–6116. <https://doi.org/10.1002/anie.201610544>.
- (70) Gao, Z.; Robinson, K.; Skowronski, D. M.; De Serres, G.; Withers, S. G. Quantification of the Total Neuraminidase Content of Recent Commercially-Available Influenza Vaccines: Introducing a Neuraminidase Titration Reagent. *Vaccine* **2020**, *38* (4), 715–718. <https://doi.org/10.1016/j.vaccine.2019.11.041>.
- (71) Buchini, S.; Buschiazzo, A.; Withers, S. G. A New Generation of Specific Trypanosoma Cruzi Trans-Sialidase Inhibitors. *Angew. Chemie Int. Ed.* **2008**, *47* (14), 2700–2703. <https://doi.org/10.1002/anie.200705435>.
- (72) Yao, Y. Towards Specific Inactivators of Trypanosoma Cruzi Trans-Sialidase, 2013.
- (73) Inglese, J.; Johnson, R. L.; Simeonov, A.; Xia, M.; Zheng, W.; Austin, C. P.; Auld, D. S. High-Throughput Screening Assays for the Identification of Chemical Probes. *Nat. Chem. Biol.* **2007**, *3* (8), 466–479. <https://doi.org/10.1038/nchembio.2007.17>.
- (74) Dandapani, S.; Rosse, G.; Southall, N.; Salvino, J. M.; Thomas, C. J. Selecting, Acquiring, and Using Small Molecule Libraries for High-Throughput Screening. In *Current Protocols in Chemical Biology*; John Wiley & Sons, Inc., 2012.

<https://doi.org/10.1002/9780470559277.ch110252>.

- (75) Janzen, W. P. Screening Technologies for Small Molecule Discovery: The State of the Art. *Chem. Biol.* **2014**, *21* (9), 1162–1170. <https://doi.org/10.1016/j.chembiol.2014.07.015>.
- (76) Zhang, J.-H. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of Biomolecular Screening*. 1999, pp 67–73. <https://doi.org/10.1177/108705719900400206>.
- (77) Neres, J.; Brewer, M. L.; Ratier, L.; Botti, H.; Buschiazzo, A.; Edwards, P. N.; Mortenson, P. N.; Charlton, M. H.; Alzari, P. M.; Frasch, A. C.; Bryce, R. a.; Douglas, K. T. Discovery of Novel Inhibitors of Trypanosoma Cruzi Trans-Sialidase from in Silico Screening. *Bioorganic Med. Chem. Lett.* **2009**, *19* (3), 589–596. <https://doi.org/10.1016/j.bmcl.2008.12.065>.
- (78) Arioka, S.; Sakagami, M.; Uematsu, R.; Yamaguchi, H.; Togame, H.; Takemoto, H.; Hinou, H.; Nishimura, S. I. Potent Inhibitor Scaffold against Trypanosoma Cruzi Trans-Sialidase. *Bioorganic Med. Chem.* **2010**, *18* (4), 1633–1640. <https://doi.org/10.1016/j.bmc.2009.12.062>.
- (79) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Deliv. Rev.* **2001**, *46* (1–3), 3–26. [https://doi.org/10.1016/S0169-409X\(00\)00129-0](https://doi.org/10.1016/S0169-409X(00)00129-0).
- (80) Schinor, E. C.; Salvador, M. J.; Ito, I. Y.; De Albuquerque, S.; Dias, D. A. Trypanocidal and Antimicrobial Activities of Moquinia Kingii. *Phytomedicine* **2004**, *11* (2–3), 224–229. <https://doi.org/10.1078/0944-7113-00342>.
- (81) Kuntz, S.; Wenzel, U.; Daniel, H. Comparative Analysis of the Effects of Flavonoids on Proliferation, Cytotoxicity, and Apoptosis in Human Colon Cancer Cell Lines. *Eur. J. Nutr.* **1999**, *38* (3), 133–142. <https://doi.org/10.1007/s003940050054>.
- (82) Johnson, D. S.; Weerapana, E.; Cravatt, B. F. Strategies for Discovering and Derisking Covalent, Irreversible Enzyme Inhibitors. *Future Medicinal Chemistry*. June 2010, pp 949–964. <https://doi.org/10.4155/fmc.10.21>.
- (83) Ghosh, A. K.; Samanta, I.; Mondal, A.; Liu, W. R. Covalent Inhibition in Drug Discovery. *ChemMedChem* **2019**, *14* (9), 889–906. <https://doi.org/10.1002/cmdc.201900107>.
- (84) Meindl, P.; Bodo, G.; Palese, P.; Schulman, J.; Tuppy, H. Inhibition of Neuraminidase Activity by Derivatives of 2-Deoxy-2,3-Dehydro-N-Acetylneuraminic Acid. *Virology* **1974**, *58* (2), 457–463. [https://doi.org/10.1016/0042-6822\(74\)90080-4](https://doi.org/10.1016/0042-6822(74)90080-4).
- (85) Von Itzstein, M.; Wu, W. Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Jin, B.; Phan, T. Van; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, P. M.; Varghese, J. N.; Ryan, D. M.; Woods, J. M.; Bethell, R. C.; Hotham, V. J.; Cameron, J. M.; Penn, C. R. Rational Design of Potent Sialidase-Based Inhibitors of Influenza Virus Replication. *Nature* **1993**, *363* (6428), 418–423. <https://doi.org/10.1038/363418a0>.
- (86) von Itzstein, M.; Wu, W. Y.; Jin, B. The Synthesis of 2,3-Didehydro-2,4-Dideoxy-4-Guanidinyln-Acetylneuraminic Acid: A Potent Influenza Virus Sialidase Inhibitor. *Carbohydr. Res.* **1994**, *259* (2), 301–305. [https://doi.org/10.1016/0008-6215\(94\)84065-2](https://doi.org/10.1016/0008-6215(94)84065-2).

- (87) Kim, J.-H.; Resende, R.; Wennekes, T.; Chen, H.-M.; Bance, N.; Buchini, S.; Watts, a. G.; Pilling, P.; Streltsov, V. a.; Petric, M.; Liggins, R.; Barrett, S.; McKimm-Breschkin, J. L.; Niikura, M.; Withers, S. G. Mechanism-Based Covalent Neuraminidase Inhibitors with Broad Spectrum Influenza Antiviral Activity. *Science* **2013**, *71* (340), 71–75. <https://doi.org/10.1126/science.1232552>.
- (88) Buchini, S.; Gallat, F. X.; Greig, I. R.; Kim, J. H.; Wakatsuki, S.; Chavas, L. M. G.; Withers, S. G. Tuning Mechanism-Based Inactivators of Neuraminidases: Mechanistic and Structural Insights. *Angew. Chemie Int. Ed.* **2014**, *53* (13), 3382–3386. <https://doi.org/10.1002/anie.201309675>.
- (89) Weck, S.; Robinson, K.; Smith, M. R.; Withers, S. G. Understanding Viral Neuraminidase Inhibition by Substituted Difluorosialic Acids. *Chem. Commun.* **2015**, *51* (14), 2933–2935. <https://doi.org/10.1039/c4cc08256g>.
- (90) Miyagi, T.; Wada, T.; Yamaguchi, K.; Shiozaki, K.; Sato, I.; Kakugawa, Y.; Yamanami, H.; Fujiya, T. Human Sialidase as a Cancer Marker. *Proteomics* **2008**, *8* (16), 3303–3311. <https://doi.org/10.1002/pmic.200800248>.
- (91) Miyagi, T.; Yamaguchi, K. Mammalian Sialidases: Physiological and Pathological Roles in Cellular Functions. *Glycobiology*. July 2012, pp 880–896. <https://doi.org/10.1093/glycob/cws057>.
- (92) Cairo, C. W. Inhibitors of the Human Neuraminidase Enzymes. *Medchemcomm* **2014**, *5* (8), 1067–1074. <https://doi.org/10.1039/c4md00089g>.
- (93) Zhang, Y.; Albohy, A.; Zou, Y.; Smutova, V.; Pshezhetsky, A. V.; Cairo, C. W. Identification of Selective Inhibitors for Human Neuraminidase Isoenzymes Using C4,C7-Modified 2-Deoxy-2,3-Didehydro-N-Acetylneuraminic Acid (DANA) Analogues. *J. Med. Chem.* **2013**, *56* (7), 2948–2958. <https://doi.org/10.1021/jm301892f>.
- (94) Guo, T.; Héon-Roberts, R.; Zou, C.; Zheng, R.; Pshezhetsky, A. V.; Cairo, C. W. Selective Inhibitors of Human Neuraminidase 1 (NEU1). *J. Med. Chem.* **2018**, *61* (24), 11261–11279. <https://doi.org/10.1021/acs.jmedchem.8b01411>.
- (95) Zou, Y.; Albohy, A.; Sandbhor, M.; Cairo, C. W. Inhibition of Human Neuraminidase 3 (NEU3) by C9-Triazole Derivatives of 2,3-Didehydro-N-Acetyl-Neuraminic Acid. *Bioorganic Med. Chem. Lett.* **2010**, *20* (24), 7529–7533. <https://doi.org/10.1016/j.bmcl.2010.09.111>.
- (96) Guo, T.; Dätwyler, P.; Demina, E.; Richards, M. R.; Ge, P.; Zou, C.; Zheng, R.; Fougerat, A.; Pshezhetsky, A. V.; Ernst, B.; Cairo, C. W. Selective Inhibitors of Human Neuraminidase 3. *J. Med. Chem.* **2018**, *61* (5), 1990–2008. <https://doi.org/10.1021/acs.jmedchem.7b01574>.
- (97) Albohy, A.; Zhang, Y.; Smutova, V.; Pshezhetsky, A. V.; Cairo, C. W. Identification of Selective Nanomolar Inhibitors of the Human Neuraminidase, NEU4. *ACS Med. Chem. Lett.* **2013**, *4* (6), 532–537. <https://doi.org/10.1021/ml400080t>.
- (98) Chavas, L. M. G.; Tringali, C.; Fusi, P.; Venerando, B.; Tettamanti, G.; Kato, R.; Monti, E.; Wakatsuki, S. Crystal Structure of the Human Cytosolic Sialidase Neu2: Evidence for the Dynamic Nature of Substrate Recognition. *J. Biol. Chem.* **2005**, *280* (1), 469–475. <https://doi.org/10.1074/jbc.M411506200>.

- (99) Monti, E.; Preti, A.; Nesti, C.; Ballabio, A. Expression of a Novel Human Sialidase Encoded by the NEU2 Gene. *Glycobiology* **1999**, *9* (12), 1313–1321. <https://doi.org/10.1093/glycob/9.12.1313>.
- (100) Khedri, Z.; Muthana, M. M.; Li, Y.; Muthana, S. M.; Yu, H.; Cao, H.; Chen, X. Probe Sialidase Substrate Specificity Using Chemoenzymatically Synthesized Sialosides Containing C9-Modified Sialic Acid. *Chem. Commun.* **2012**, *48* (27), 3357–3359. <https://doi.org/10.1039/c2cc17393j>.
- (101) Buchini, S.; Buschiazzo, A.; Withers, S. G. A New Generation of Specific Trypanosoma Cruzi Trans-Sialidase Inhibitors. *Angew. Chemie Int. Ed.* **2008**, *47* (14), 2700–2703. <https://doi.org/10.1002/anie.200705435>.
- (102) Resende, R. Synthesis of Novel Sialidase Inhibitors to Target Influenza A Virus and Chagas' Disease, 2010.
- (103) Freire-de-Lima, L.; Oliveira, I. A.; Neves, J. L.; Penha, L. L.; Alisson-Silva, F.; Dias, W. B.; Todeschini, A. R. Sialic Acid: A Sweet Swing between Mammalian Host and Trypanosoma Cruzi. *Front. Immunol.* **2012**, *3* (NOV). <https://doi.org/10.3389/fimmu.2012.00356>.
- (104) Kayser, H.; Zeitler, R.; Kannicht, C.; Grunow, D.; Nuck, R.; Reutter, W. Biosynthesis of a Nonphysiological Sialic Acid in Different Rat Organs, Using N-Propanoyl-D-Hexosamines as Precursors. *J. Biol. Chem.* **1992**, *267* (24), 16934–16938.
- (105) Keppler, O. T.; Stehling, P.; Herrmann, M.; Kayser, H.; Grunow, D.; Reutter, W.; Pawlita, M. Biosynthetic Modulation of Sialic Acid-Dependent Virus-Receptor Interactions of Two Primate Polyoma Viruses. *J. Biol. Chem.* **1995**, *270* (3), 1308–1314. <https://doi.org/10.1074/jbc.270.3.1308>.
- (106) Keppler, O. T.; Herrmann, M.; Von Der Lieth, C. W.; Stehling, P.; Reutter, W.; Pawlita, M. Elongation of the N-Acyl Side Chain of Sialic Acids in MDCK II Cells Inhibits Influenza A Virus Infection. *Biochem. Biophys. Res. Commun.* **1998**, *253* (2), 437–442. <https://doi.org/10.1006/bbrc.1998.9650>.
- (107) Madhunapantula, S. V.; Achur, R. N.; Bhavanandan, V. P.; Gowda, D. C. The Effect of Substitution of the N-Acetyl Groups of N-Acetylgalactosamine Residues in Chondroitin Sulfate on Its Degradation by Chondroitinase ABC. *Glycoconj. J.* **2007**, *24* (8), 465–473. <https://doi.org/10.1007/s10719-007-9039-y>.
- (108) Lieke, T.; Gröbe, D.; Blanchard, V.; Grunow, D.; Tauber, R.; Zimmermann-Kordmann, M.; Jacobs, T.; Reutter, W. Invasion of Trypanosoma Cruzi into Host Cells Is Impaired by N-Propionylmannosamine and Other N-Acylmannosamines. *Glycoconj. J.* **2011**, *28* (1), 31–37. <https://doi.org/10.1007/s10719-010-9321-2>.
- (109) Li, W.; Santra, A.; Yu, H.; Slack, T. J.; Muthana, M. M.; Shi, D.; Liu, Y.; Chen, X. 9-Azido-9-Deoxy-2,3-Difluorosialic Acid as a Subnanomolar Inhibitor against Bacterial Sialidases. *J. Org. Chem.* **2019**, *84* (11), 6697–6708. <https://doi.org/10.1021/acs.joc.9b00385>.
- (110) Braun, P.; Nägele, B.; Wittmann, V.; Drescher, M. Mechanism of Multivalent Carbohydrate-Protein Interactions Studied by EPR Spectroscopy. *Angew. Chemie Int. Ed.* **2011**, *50* (36), 8428–8431. <https://doi.org/10.1002/anie.201101074>.

- (111) Beliczey, J.; Kragl, U.; Wandrey, C.; Hamacher, K.; Coenen, H.; Tierling, T. Methods for Making Fluorinated Sugars Having a Side Chain and Use Thereof. US6355453B1, 2002.
- (112) Yang, L.; Dong, T.; Revankar, H. M.; Zhang, C. P. Recent Progress on Fluorination in Aqueous Media. *Green Chemistry*. Royal Society of Chemistry 2017, pp 3951–3992. <https://doi.org/10.1039/c7gc01566f>.
- (113) Burkart, M. D.; Zhang, Z.; Hung, S. C.; Wong, C. H. A New Method for the Synthesis of Fluoro-Carbohydrates and Glycosides Using Selectfluor. *J. Am. Chem. Soc.* **1997**, *119* (49), 11743–11746. <https://doi.org/10.1021/ja9723904>.
- (114) Francisco, C. G.; González, C. C.; Kennedy, A. R.; Paz, N. R.; Suárez, E. Fragmentation of Carbohydrate Anomeric Alkoxyl Radicals: New Synthesis of Chiral 1-Fluoro-1-Halo-1-Iodoalditols. *Chemistry* **2008**, *14* (22), 6704–6712. <https://doi.org/10.1002/chem.200800734>.
- (115) Paris, G.; Cremona, M. L.; Amaya, M. F.; Buschiazzo, A.; Giambiagi, S.; Frasci, A. C. C.; Alzari, P. M. Probing Molecular Function of Trypanosomal Sialidases: Single Point Mutations Can Change Substrate Specificity and Increase Hydrolytic Activity. *Glycobiology* **2001**, *11* (4), 305–311. <https://doi.org/10.1093/glycob/11.4.305>.
- (116) Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products*. American Chemical Society March 25, 2016, pp 629–661. <https://doi.org/10.1021/acs.jnatprod.5b01055>.
- (117) Watson, A. A.; Fleet, G. W. J.; Asano, N.; Molyneux, R. J.; Nash, R. J. Polyhydroxylated Alkaloids - Natural Occurrence and Therapeutic Applications. *Phytochemistry* **2001**, *56* (3), 265–295. [https://doi.org/10.1016/S0031-9422\(00\)00451-9](https://doi.org/10.1016/S0031-9422(00)00451-9).
- (118) Liang, P. H.; Cheng, W. C.; Lee, Y. L.; Yu, H. P.; Wu, Y. T.; Lin, Y. L.; Wong, C. H. Novel Five-Membered Iminocyclitol Derivatives as Selective and Potent Glycosidase Inhibitors: New Structures for Antivirals and Osteoarthritis. *ChemBioChem* **2006**, *7* (1), 165–173. <https://doi.org/10.1002/cbic.200500321>.
- (119) Li, Y. G.; Ji, D. F.; Zhong, S.; Lin, T. B.; Lv, Z. Q.; Hu, G. Y.; Wang, X. 1-Deoxynojirimycin Inhibits Glucose Absorption and Accelerates Glucose Metabolism in Streptozotocin-Induced Diabetic Mice. *Sci. Rep.* **2013**, *3*. <https://doi.org/10.1038/srep01377>.
- (120) Meindl, P.; Tuppy, H. Über 2-Deoxy-2,3-Dehydro-Sialinsäuren, 1. Mitt.: Synthese Und Eigenschaften von 2-Deoxy-2,3-Dehydro-N-Acylneuraminsäuren Und Deren Methylestern. *Monatshefte für Chemie* **1969**, *100* (4), 1295–1306. <https://doi.org/10.1007/BF00903465>.
- (121) Kamerling, J. P.; Vliegthart, J. F. G.; Schauer, R.; Strecker, G.; Moötreuil, J. Isolation and Identification of 2-Deoxy-2,3-dehydro-N-Acetylneuraminic Acid from the Urine of a Patient with Sialuria. *Eur. J. Biochem.* **1975**, *56* (1), 253–258. <https://doi.org/10.1111/j.1432-1033.1975.tb02228.x>.
- (122) Owen, C. D.; Lukacik, P.; Potter, J. A.; Sleator, O.; Taylor, G. L.; Walsh, M. A. Streptococcus Pneumoniae NanC: Structural Insights into the Specificity and Mechanism of a Sialidase That Produces a Sialidase Inhibitor. *J. Biol. Chem.* **2015**, *290* (46), 27736–27748. <https://doi.org/10.1074/jbc.M115.673632>.

- (123) WoRMS Editorial Board. No Title <http://www.marinespecies.org>. <https://doi.org/10.14284/170>.
- (124) Bouchet, P.; Duarte, C. M. *The Exploration of Marine Biodiversity Scientific and Technological Challenges*; BBVA Foundation: Bilbao, Spain, 2006.
- (125) Bergmann, W.; Feeney, R. J. Contributions to the Study of Marine Products. XXXII. the Nucleosides of Sponges. I. *J. Org. Chem.* **1951**, *16* (6), 981–987. <https://doi.org/10.1021/jo01146a023>.
- (126) Faulkner, D. J. Highlights of Marine Natural Products Chemistry (1972-1999). *Natural Product Reports*. February 2000, pp 1–6. <https://doi.org/10.1039/a909113k>.
- (127) Laport, M.; Santos, O.; Muricy, G. Marine Sponges: Potential Sources of New Antimicrobial Drugs. *Curr. Pharm. Biotechnol.* **2009**, *10* (1), 86–105. <https://doi.org/10.2174/138920109787048625>.
- (128) Jiménez, C. Marine Natural Products in Medicinal Chemistry. *ACS Med. Chem. Lett.* **2018**, *9* (10), 959–961. <https://doi.org/10.1021/acsmmedchemlett.8b00368>.
- (129) Fattorusso, E.; Tagliatalata-Scafati, O. Marine Natural Products Active against Protozoan Parasites. In *Handbook of Marine Natural Products*; Springer Netherlands, 2012; pp 1075–1110. [https://doi.org/10.1007/978-90-481-3834-0\\_21](https://doi.org/10.1007/978-90-481-3834-0_21).
- (130) Parra, L. L. L.; Bertonha, A. F.; Severo, I. R. M.; Aguiar, A. C. C.; de Souza, G. E.; Oliva, G.; Guido, R. V. C.; Grazzia, N.; Costa, T. R.; Miguel, D. C.; Gadelha, F. R.; Ferreira, A. G.; Hajdu, E.; Romo, D.; Berlinck, R. G. S. Isolation, Derivative Synthesis, and Structure–Activity Relationships of Antiparasitic Bromopyrrole Alkaloids from the Marine Sponge *Tedania Brasiliensis*. *J. Nat. Prod.* **2018**, *81* (1), 188–202. <https://doi.org/10.1021/acs.jnatprod.7b00876>.
- (131) Nakao, Y.; Fusetani, N. Enzyme Inhibitors from Marine Invertebrates. In *Handbook of Marine Natural Products*; Fattorusso, E., Gerwick, W. H., Tagliatalata-Scafati, O., Eds.; Springer Netherlands: Dordrecht, 2012; pp 1145–1229. [https://doi.org/10.1007/978-90-481-3834-0\\_23](https://doi.org/10.1007/978-90-481-3834-0_23).
- (132) He, K.; Shi, J. C.; Mao, X. M. Safety and Efficacy of Acarbose in the Treatment of Diabetes in Chinese Patients. *Therapeutics and Clinical Risk Management*. Dove Medical Press Ltd. 2014, pp 505–511. <https://doi.org/10.2147/TCRM.S50362>.
- (133) Rosak, C.; Mertes, G. Critical Evaluation of the Role of Acarbose in the Treatment of Diabetes: Patient Considerations. *Diabetes. Metab. Syndr. Obes.* **2012**, *5*, 357–367. <https://doi.org/10.2147/DMSO.S28340>.
- (134) Tarling, C. A.; Woods, K.; Zhang, R.; Brastianos, H. C.; Brayer, G. D.; Andersen, R. J.; Withers, S. G. The Search for Novel Human Pancreatic Alpha-Amylase Inhibitors: High-Throughput Screening of Terrestrial and Marine Natural Product Extracts. *ChemBioChem* **2008**, *9* (3), 433–438. <https://doi.org/10.1002/cbic.200700470> [doi].
- (135) Tysoe, C.; Williams, L. K.; Keyzers, R.; Nguyen, N. T.; Tarling, C.; Wicki, J.; Goddard-Borger, E. D.; Aguda, A. H.; Perry, S.; Foster, L. J.; Andersen, R. J.; Brayer, G. D.; Withers, S. G. Potent Human  $\alpha$ -Amylase Inhibition by the  $\beta$ -Defensin-like Protein Helianthamide. *ACS Cent. Sci.* **2016**, *2* (3), 154–161. <https://doi.org/10.1021/acscentsci.5b00399>.

- (136) Withers, S.; Tarling, A. C.; Andersen, R.; Brayer, G. D.; Woods, K. Alpha-Amylase Inhibitors: The Montbretins and Uses Thereof. US8431541B2, 2008.
- (137) Williams, L. K.; Zhang, X.; Caner, S.; Tysoe, C.; Nguyen, N. T.; Wicki, J.; Williams, D. E.; Coleman, J.; McNeill, J. H.; Yuen, V. G.; Andersen, R. J.; Withers, S. G.; Brayer, G. D. The Amylase Inhibitor Montbretin A Reveals a New Glycosidase Inhibition Motif. *Nat. Chem. Biol.* **2015**, No. 11, 691–696.
- (138) Yuen, V. G.; Coleman, J.; Withers, S. G.; Andersen, R. J.; Brayer, G. D.; Mustafa, S.; McNeill, J. H. Glucose Lowering Effect of Montbretin A in Zucker Diabetic Fatty Rats. *Mol. Cell. Biochem.* **2016**, *411* (1–2), 373–381. <https://doi.org/10.1007/s11010-015-2599-4>.
- (139) Irmisch, S.; Jancsik, S.; Yuen, M. M. S.; Madilao, L. L.; Bohlmann, J. Biosynthesis of the Anti-diabetic Metabolite Montbretin A: Glucosylation of the Central Intermediate Mini-MbA. *Plant J.* **2019**, *100* (5), 879–891. <https://doi.org/10.1111/tbj.14493>.
- (140) Colman, P. M.; Smith, B. J. The Trypanosomal Trans-Sialidase: Two Catalytic Functions Associated with One Catalytic Site. *Structure* **2002**, *10* (11), 1466–1468. [https://doi.org/10.1016/s0969-2126\(02\)00888-2](https://doi.org/10.1016/s0969-2126(02)00888-2).
- (141) Palmier, M. O.; Van Doren, S. R. Rapid Determination of Enzyme Kinetics from Fluorescence: Overcoming the Inner Filter Effect. *Anal. Biochem.* **2007**, *371* (1), 43–51. <https://doi.org/10.1016/j.ab.2007.07.008>.
- (142) Rowdhwal, S. S. S.; Chen, J. Toxic Effects of Di-2-Ethylhexyl Phthalate: An Overview. *Biomed Res. Int.* **2018**, *2018*, 1750368. <https://doi.org/10.1155/2018/1750368>.
- (143) Lorz, P. M.; Towae, F. K.; Enke, W.; Jäckh, R.; Bhargava, N.; Hillesheim, W. Phthalic Acid and Derivatives. In *Ullmann's Encyclopedia of Industrial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2007.
- (144) Blackburn, C. L.; Hopmann, C.; Sakowicz, R.; Berdelis, M. S.; Goldstein, L. S. B.; Faulkner, D. J. Adociasulfates 1–6, Inhibitors of Kinesin Motor Proteins from the Sponge *Haliclona* (Aka *Adocia*) Sp. *J. Org. Chem.* **1999**, *64* (15), 5565–5570. <https://doi.org/10.1021/jo9824448>.
- (145) Kalaitzis, J. A.; De Almeida Leone, P.; Harris, L.; Butler, M. S.; Ngo, A.; Hooper, J. N. A.; Quinn, R. J. Adociasulfates 1, 7, and 8: New Bioactive Hexaprenoid Hydroquinones from the Marine Sponge *Adocia* Sp. *J. Org. Chem.* **1999**, *64* (15), 5571–5574. <https://doi.org/10.1021/jo990404d>.
- (146) Balansa, W.; Islam, R.; Fontaine, F.; Piggott, A. M.; Zhang, H.; Webb, T. I.; Gilbert, D. F.; Lynch, J. W.; Capon, R. J. Ircinialactams: Subunit-Selective Glycine Receptor Modulators from Australian Sponges of the Family Irciniidae. *Bioorganic Med. Chem.* **2010**, *18* (8), 2912–2919. <https://doi.org/10.1016/j.bmc.2010.03.002>.
- (147) Prasad, P.; Zhang, A.; Salim, A. A.; Capon, R. J. Pursuing Sesterterpene Lactams in Australian Irciniidae Sponges. *Fitoterapia* **2018**, *126*, 83–89. <https://doi.org/10.1016/j.fitote.2017.09.003>.
- (148) Khushi, S.; Nahar, L.; Salim, A. A.; Capon, R. J. Cacolides: Sesterterpene Butenolides from a Southern Australian Marine Sponge, *Cacospongia* Sp. *Mar. Drugs* **2018**, *16* (11), 1–17. <https://doi.org/10.3390/md16110456>.

- (149) Sidhu, S. S. Phage Display in Pharmaceutical Biotechnology. *Current Opinion in Biotechnology*. Elsevier Ltd 2000, pp 610–616. [https://doi.org/10.1016/S0958-1669\(00\)00152-X](https://doi.org/10.1016/S0958-1669(00)00152-X).
- (150) Gold, L. MRNA Display: Diversity Matters during in Vitro Selection. *Proceedings of the National Academy of Sciences of the United States of America*. April 24, 2001, pp 4825–4826. <https://doi.org/10.1073/pnas.091101698>.
- (151) Shafee, T. Phage display - Wikipedia [https://en.wikipedia.org/wiki/Phage\\_display#/media/File:Phage\\_display.png](https://en.wikipedia.org/wiki/Phage_display#/media/File:Phage_display.png) (accessed Jan 29, 2020).
- (152) Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science* **1985**, 228 (4705), 1315–1317. <https://doi.org/10.1126/science.4001944>.
- (153) Sidhu, S. S.; Lowman, H. B.; Cunningham, B. C.; Wells, J. A. Phage Display for Selection of Novel Binding Peptides. *Methods Enzymol.* **2000**, 328, 333–363. [https://doi.org/10.1016/s0076-6879\(00\)28406-1](https://doi.org/10.1016/s0076-6879(00)28406-1).
- (154) Roberts, R. W.; Szostak, J. W. RNA-Peptide Fusions for the in Vitro Selection of Peptides and Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, 94 (23), 12297–12302. <https://doi.org/10.1073/pnas.94.23.12297>.
- (155) Mattheakis, L. C.; Dias, J. M.; Dower, W. J. Cell-Free Synthesis of Peptide Libraries Displayed on Polysomes. *Methods Enzymol.* **1996**, 267, 195–207. [https://doi.org/10.1016/s0076-6879\(96\)67013-x](https://doi.org/10.1016/s0076-6879(96)67013-x).
- (156) Hanes, J.; Plückthun, A. In Vitro Selection and Evolution of Functional Proteins by Using Ribosome Display. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, 94 (10), 4937–4942. <https://doi.org/10.1073/pnas.94.10.4937>.
- (157) Wilson, D. S.; Keefe, A. D.; Szostak, J. W. The Use of MRNA Display to Select High-Affinity Protein-Binding Peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98 (7), 3750–3755. <https://doi.org/10.1073/pnas.061028198>.
- (158) Lau, J. L.; Dunn, M. K. Therapeutic Peptides: Historical Perspectives, Current Development Trends, and Future Directions. *Bioorganic Med. Chem.* **2018**, 26 (10), 2700–2707. <https://doi.org/10.1016/j.bmc.2017.06.052>.
- (159) Bockus, A. T.; Lokey, R. S. *Practical Medicinal Chemistry with Macrocycles: Design, Synthesis, and Case Studies*; Marsault, E., Peterson, M. L., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2017.
- (160) Tsomaia, N. Peptide Therapeutics: Targeting the Undruggable Space. *Eur. J. Med. Chem.* **2015**, 94, 459–470. <https://doi.org/10.1016/j.ejmech.2015.01.014>.
- (161) Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discovery Today*. Elsevier Ltd 2015, pp 122–128. <https://doi.org/10.1016/j.drudis.2014.10.003>.
- (162) Josephson, K.; Hartman, M. C. T.; Szostak, J. W. Ribosomal Synthesis of Unnatural Peptides. *J.*

- Am. Chem. Soc.* **2005**, *127* (33), 11727–11735. <https://doi.org/10.1021/ja0515809>.
- (163) Kötzler, M. P.; Robinson, K.; Chen, H. M.; Okon, M.; McIntosh, L. P.; Withers, S. G. Modulating the Nucleophile of a Glycoside Hydrolase through Site-Specific Incorporation of Fluoroglutamic Acids. *J. Am. Chem. Soc.* **2018**, *140* (26), 8268–8276.
- (164) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. A Highly Flexible TRNA Acylation Method for Non-Natural Polypeptide Synthesis. *Nat. Methods* **2006**, *3* (5), 357–359. <https://doi.org/10.1038/nmeth877>.
- (165) Ohuchi, M.; Murakami, H.; Suga, H. The Flexizyme System: A Highly Flexible TRNA Aminoacylation Tool for the Translation Apparatus. *Current Opinion in Chemical Biology*. October 2007, pp 537–542. <https://doi.org/10.1016/j.cbpa.2007.08.011>.
- (166) Hipolito, C. J.; Suga, H. Ribosomal Production and in Vitro Selection of Natural Product-like Peptidomimetics: The FIT and RaPID Systems. *Curr. Opin. Chem. Biol.* **2012**, *16* (1–2), 196–203. <https://doi.org/10.1016/j.cbpa.2012.02.014>.
- (167) Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H. Reprogramming the Translation Initiation for the Synthesis of Physiologically Stable Cyclic Peptides. *ACS Chem. Biol.* **2008**, *3* (2), 120–129. <https://doi.org/10.1021/cb700233t>.
- (168) Tugyi, R.; Uray, K.; Iván, D.; Fellingner, E.; Perkins, A.; Hudecz, F. Partial D-Amino Acid Substitution: Improved Enzymatic Stability and Preserved Ab Recognition of a MUC2 Epitope Peptide. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (2), 413–418. <https://doi.org/10.1073/pnas.0407677102>.
- (169) Goto, Y.; Iwasaki, K.; Torikai, K.; Murakami, H.; Suga, H. Ribosomal Synthesis of Dehydrobutyrine- and Methyllanthionine-Containing Peptides. *Chem. Commun.* **2009**, No. 23, 3419–3421.
- (170) Yamagishi, Y.; Shoji, I.; Miyagawa, S.; Kawakami, T.; Katoh, T.; Goto, Y.; Suga, H. Natural Product-like Macrocyclic N-Methyl-Peptide Inhibitors against a Ubiquitin Ligase Uncovered from a Ribosome-Expressed de Novo Library. *Chem. Biol.* **2011**, *18* (12), 1562–1570. <https://doi.org/10.1016/j.chembiol.2011.09.013>.
- (171) Passioura, T.; Katoh, T.; Goto, Y.; Suga, H. Selection-Based Discovery of Druglike Macrocyclic Peptides. *Annu. Rev. Biochem.* **2014**, *83*, 727–752. <https://doi.org/10.1146/annurev-biochem-060713-035456>.
- (172) Hayashi, Y.; Morimoto, J.; Suga, H. In Vitro Selection of Anti-Akt2 Thioether-Macrocyclic Peptides Leading to Isoform-Selective Inhibitors. *ACS Chem. Biol.* **2012**, *7* (3), 607–613. <https://doi.org/10.1021/cb200388k>.
- (173) Matsunaga, Y.; Bashiruddin, N. K.; Kitago, Y.; Takagi, J.; Correspondence, H. S.; Suga, H. Allosteric Inhibition of a Semaphorin 4D Receptor Plexin B1 by a High-Affinity Macrocyclic Peptide. *Cell Chem. Biol.* **2016**, *23*, 1341–1350. <https://doi.org/10.1016/j.chembiol.2016.09.015>.
- (174) Kawamura, A.; Münzel, M.; Kojima, T.; Yapp, C.; Bhushan, B.; Goto, Y.; Tumber, A.; Katoh, T.; King, O. N. F.; Passioura, T.; Walport, L. J.; Hatch, S. B.; Madden, S.; Müller, S.; Brennan, P. E.;

- Chowdhury, R.; Hopkinson, R. J.; Suga, H.; Schofield, C. J. Highly Selective Inhibition of Histone Demethylases by de Novo Macrocyclic Peptides. *Nat. Commun.* **2017**, *8*.  
<https://doi.org/10.1038/ncomms14773>.
- (175) Passioura, T.; Liu, W.; Dunkelmann, D.; Higuchi, T.; Suga, H. Display Selection of Exotic Macrocyclic Peptides Expressed under a Radically Reprogrammed 23 Amino Acid Genetic Code. *J. Am. Chem. Soc.* **2018**, *140* (37), 11551–11555. <https://doi.org/10.1021/jacs.8b03367>.
- (176) Jongkees, S. A. K.; Caner, S.; Tysoe, C.; Brayer, G. D.; Withers, S. G.; Suga, H. Rapid Discovery of Potent and Selective Glycosidase-Inhibiting De Novo Peptides. *Cell Chem. Biol.* **2017**, *24* (3), 381–390. <https://doi.org/10.1016/j.chembiol.2017.02.001>.
- (177) Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G. Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis. *J. Am. Chem. Soc.* **1998**, *120* (22), 5583–5584. <https://doi.org/10.1021/ja980833d>.
- (178) Gilbert, M.; Brisson, J. R.; Karwaski, M. F.; Michniewicz, J.; Cunningham, A. M.; Wu, Y.; Young, N. M.; Wakarchuk, W. W. Biosynthesis of Ganglioside Mimics in *Campylobacter* Jejuni OH4384. Identification of the Glycosyltransferase Genes, Enzymatic Synthesis of Model Compounds, and Characterization of Nanomole Amounts by 600-MHz <sup>1</sup>H and <sup>13</sup>C NMR Analysis. *J. Biol. Chem.* **2000**, *275* (6), 3896–3906. <https://doi.org/10.1074/jbc.275.6.3896>.
- (179) Kempton, J. B.; Withers, S. G. Mechanism of *Agrobacterium*  $\beta$ -Glucosidase: Kinetic Studies. *Biochemistry* **1992**, *31* (41), 9961–9969. <https://doi.org/10.1021/bi00156a015>.
- (180) Kim, J. H.; Ryu, H. W.; Shim, J. H.; Park, K. H.; Withers, S. G. Development of New and Selective *Trypanosoma Cruzi* Trans-Sialidase Inhibitors from Sulfonamide Chalcones and Their Derivatives. *ChemBioChem* **2009**, *10* (15), 2475–2479. <https://doi.org/10.1002/cbic.200900108>.
- (181) Jongkees, S. A. K.; Umemoto, S.; Suga, H. Linker-Free Incorporation of Carbohydrates into in Vitro Displayed Macrocyclic Peptides. *Chem. Sci.* **2017**, *8* (2), 1474–1481. <https://doi.org/10.1039/C6SC04381J>.
- (182) Adamson, J. G.; Blaskovich, M. A.; Groenevelt, H.; Lajoie, G. A. Simple and Convenient Synthesis of Tert-Butyl Ethers of Fmoc-Serine, Fmoc-Threonine, and Fmoc-Tyrosine. *J. Org. Chem.* **1991**, *56* (10), 3447–3449. <https://doi.org/10.1021/jo00010a050>.
- (183) Zhang, K.; Zhang, H.; Xu, G.; Xiang, S.; Xu, D.; Liu, S.; Li, H. Alkylation of Phenol with Tert-Butyl Alcohol Catalyzed by Large Pore Zeolites. *Appl. Catal. A Gen.* **2001**, *207* (1–2), 183–190. [https://doi.org/10.1016/S0926-860X\(00\)00663-3](https://doi.org/10.1016/S0926-860X(00)00663-3).
- (184) Bartoli, G.; Bosco, M.; Locatelli, M.; Marcantoni, E.; Melchiorre, P.; Sambri, L. Unusual and Unexpected Reactivity of *t*-Butyl Dicarboxylate (Boc<sub>2</sub>O) with Alcohols in the Presence of Magnesium Perchlorate. A New and General Route to *t*-Butyl Ethers. *Org. Lett.* **2005**, *7* (3), 427–430. <https://doi.org/10.1021/ol047704o>.
- (185) Bartoli, G.; Bosco, M.; Carlone, A.; Dalpozzo, R.; Locatelli, M.; Melchiorre, P.; Sambri, L. Alcohols and Di-*Tert*-Butyl Dicarboxylate: How the Nature of the Lewis Acid Catalyst May Address the Reaction to the Synthesis of *Tert*-Butyl Ethers. *J. Org. Chem.* **2006**, *71* (26), 9580–9588. <https://doi.org/10.1021/jo061402d>.

- (186) Hirao, A.; Imai, T.; Watanabe, K.; Hayashi, M.; Sugiyama, K. Living Anionic Polymerization of 4-( $\alpha$ -Alkylvinyl)Styrene Derivatives. *Monatshefte für Chemie* **2006**, *137*, 855–867. <https://doi.org/10.1007/s00706-006-0493-1>.
- (187) De O Freitas, L. B.; Borgati, T. F.; De Freitas, R. P.; Ruiz, A. L. T. G.; Marchetti, G. M.; De Carvalho, J. E.; Da Cunha, E. F. F.; Ramalho, T. C.; Alves, R. B. Synthesis and Antiproliferative Activity of 8-Hydroxyquinoline Derivatives Containing a 1,2,3-Triazole Moiety. *Eur. J. Med. Chem.* **2014**, *84*, 595–604. <https://doi.org/10.1016/j.ejmech.2014.07.061>.
- (188) Dymicky, M.; Byler, D. M. Synthesis of Ethyl N-Carbobenzoxytyrosyl-S-Benzylcysteinylglycinate. *Org. Prep. Proced. Int.* **1991**, *23* (1), 93–101. <https://doi.org/10.1080/00304949109458288>.
- (189) Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis : A Practical Approach*; Oxford University Press, 2000.
- (190) Van Woerkem, W. J.; Van Nispen, J. W. Difficult Couplings in Stepwise Solid Phase Peptide Synthesis: Predictable or Just a Guess? *Int. J. Pept. Protein Res.* **2009**, *38* (2), 103–113. <https://doi.org/10.1111/j.1399-3011.1991.tb01417.x>.
- (191) Pearson, D. A.; Blanchette, M.; Baker, M. Lou; Guindon, C. A. Trialkylsilanes as Scavengers for the Trifluoroacetic Acid Deblocking of Protecting Groups in Peptide Synthesis. *Tetrahedron Lett.* **1989**, *30* (21), 2739–2742. [https://doi.org/10.1016/S0040-4039\(00\)99113-5](https://doi.org/10.1016/S0040-4039(00)99113-5).
- (192) Millward, S. W.; Fiacco, S.; Austin, R. J.; Roberts, R. W. Design of Cyclic Peptides That Bind Protein Surfaces with Antibody-like Affinity. *ACS Chem. Biol.* **2007**, *2* (9), 625–634. <https://doi.org/10.1021/cb7001126>.
- (193) Isidro-Llobet, A.; Álvarez, M.; Albericio, F. Amino Acid-Protecting Groups. *Chem. Rev.* **2009**, *109* (6), 2455–2504. <https://doi.org/10.1021/cr800323s>.
- (194) Berman, H. M. The Protein Data Bank / Biopython. *Presentation* **2000**, *28* (1), 235–242. <https://doi.org/10.1093/nar/28.1.235>.
- (195) Oppezzo, P.; Obal, G.; Baraibar, M. A.; Pritsch, O.; Alzari, P. M.; Buschiazzo, A. Crystal Structure of an Enzymatically Inactive Trans-Sialidase-like Lectin from Trypanosoma Cruzi: The Carbohydrate Binding Mechanism Involves Residual Sialidase Activity. *Biochim. Biophys. Acta - Proteins Proteomics* **2011**, *1814* (9), 1154–1161. <https://doi.org/10.1016/j.bbapap.2011.04.012>.
- (196) Buschiazzo, A.; Muiá, R.; Larrieux, N.; Pitcovsky, T.; Mucci, J.; Campetella, O. Trypanosoma Cruzi Trans-Sialidase in Complex with a Neutralizing Antibody: Structure/Function Studies towards the Rational Design of Inhibitors. *PLoS Pathog.* **2012**, *8* (1). <https://doi.org/10.1371/journal.ppat.1002474>.
- (197) Martins, L. F.; Mesquita, J. T.; Pinto, E. G.; Costa-Silva, T. A.; Borborema, S. E. T.; Galisteo Junior, A. J.; Neves, B. J.; Andrade, C. H.; Shuhaib, Z. Al; Bennett, E. L.; Black, G. P.; Harper, P. M.; Evans, D. M.; Fituri, H. S.; Leyland, J. P.; Martin, C.; Roberts, T. D.; Thornhill, A. J.; Vale, S. A.; Howard-Jones, A.; Thomas, D. A.; Williams, H. L.; Overman, L. E.; Berlinck, R. G. S.; Murphy, P. J.; Tempone, A. G. Analogues of Marine Guanidine Alkaloids Are *in Vitro* Effective against *Trypanosoma Cruzi* and Selectively Eliminate *Leishmania* ( *L.* ) *Infantum* Intracellular

- Amastigotes. *J. Nat. Prod.* **2016**, 79 (9), 2202–2210. <https://doi.org/10.1021/acs.jnatprod.6b00256>.
- (198) Buckner, F. S.; Verlinde, C. L. M. J.; La Flamme, A. C.; Van Voorhis, W. C. Efficient Technique for Screening Drugs for Activity against *Trypanosoma Cruzi* Using Parasites Expressing  $\beta$ -Galactosidase. *Antimicrob. Agents Chemother.* **1996**, 40 (11), 2592–2597. <https://doi.org/10.1128/aac.40.11.2592>.
- (199) Chatelain, E. Chagas Disease Drug Discovery: Toward a New Era. *Journal of Biomolecular Screening*. SAGE Publications Inc. January 30, 2015, pp 22–35. <https://doi.org/10.1177/1087057114550585>.
- (200) Silva, D. G.; Gillespie, J. R.; Ranade, R. M.; Herbst, Z. M.; Nguyen, U. T. T.; Buckner, F. S.; Montanari, C. A.; Gelb, M. H. New Class of Antitrypanosomal Agents Based on Imidazopyridines. *ACS Med. Chem. Lett.* **2017**, 8 (7), 766–770. <https://doi.org/10.1021/acsmedchemlett.7b00202>.
- (201) Moss, R. A.; Terpinski, J.; Cox, D. P.; Denney, D. Z.; Krogh-Jespersen, K. Azide and Fluoride Exchange Reactions of Halodiazirines. *J. Am. Chem. Soc.* **1985**, 107 (9), 2743–2748. <https://doi.org/10.1021/ja00295a029>.
- (202) Zimmermann, F.; Meux, E.; Mieloszynski, J. L.; Lecuire, J. M.; Oget, N. Ruthenium Catalysed Oxidation without CCl<sub>4</sub> of Oleic Acid, Other Monoenic Fatty Acids and Alkenes. *Tetrahedron Lett.* **2005**, 46 (18), 3201–3203. <https://doi.org/10.1016/j.tetlet.2005.03.052>.
- (203) Van Dongen, S. F. M.; Janvore, J.; Van Berkel, S. S.; Marie, E.; Piel, M.; Tribet, C. Reactive Protein-Repellent Surfaces for the Straightforward Attachment of Small Molecules up to Whole Cells. *Chem. Sci.* **2012**, 3 (10), 3000–3006. <https://doi.org/10.1039/c2sc20652h>.
- (204) Srinivas, B. T. V.; Maadhur, A. R.; Bojja, S. Total Synthesis of Racemic, Natural (+) and Unnatural (-) Scorzocreticin. *Tetrahedron* **2014**, 70 (43), 8161–8167. <https://doi.org/10.1016/j.tet.2014.07.107>.
- (205) Hirao, A.; Higashihara, T.; Nagura, M.; Sakurai, T. Successive Synthesis of Well-Defined Many Arm Star-Branched Polymers by an Iterative Methodology Using a Specially Designed 1,1-Diphenylethylene. *Macromolecules* **2006**, 39 (18), 6081–6091. <https://doi.org/10.1021/ma061075b>.
- (206) Vilà, S.; Badosa, E.; Montesinos, E.; Feliu, L.; Planas, M. Solid-Phase Synthesis of Peptide Conjugates Derived from the Antimicrobial Cyclic Decapeptide BPC194. *European J. Org. Chem.* **2015**, 2015 (5), 1117–1129. <https://doi.org/10.1002/ejoc.201403344>.
- (207) Planer, J. D.; Hulverson, M. A.; Arif, J. A.; Ranade, R. M.; Don, R.; Buckner, F. S. Synergy Testing of FDA-Approved Drugs Identifies Potent Drug Combinations against *Trypanosoma Cruzi*. *PLoS Negl. Trop. Dis.* **2014**, 8 (7), 2977. <https://doi.org/10.1371/journal.pntd.0002977>.

## Appendices

### Appendix A : Enzyme kinetics

#### A.1 Michaelis-Menten kinetics

The model of an enzyme (E) catalyzed conversion of a single substrate (S), via the reversible formation of an enzyme-substrate complex (ES), to a single product (P) was described by Michaelis and Menten and is shown below:



Where  $k_1$  (substrate binding rate constant),  $k_{-1}$  (substrate dissociation rate constant) and  $k_2$  (catalysis rate constant) are the rate constants describing this relationship. Substrate binding to free enzyme and dissociation from enzyme-substrate complex are in rapid equilibrium, thus it is typically assumed that the catalysis rate constant  $k_2$  is rate limiting. In this case the concentration of intermediate ES is assumed to be constant, known as the steady-state approximation, which implies that the intermediate ES is consumed as quickly as it is generated. These assumptions greatly simplify the rate law determination:

$$\frac{d[\text{ES}]}{dt} = k_1[\text{E}]_o[\text{S}] - k_{-1}[\text{ES}] - k_2[\text{ES}] = 0$$

Rearranged as:

$$[\text{ES}] = \frac{k_1[\text{E}]_o[\text{S}]}{(k_{-1} + k_2)}$$

since,

$$[\text{E}]_o = [\text{E}]_{\text{total}} - [\text{ES}]$$

and  $K_m$  is defined as:

$$K_m = \frac{(k_{-1} + k_2)}{k_1}$$

then,

$$[ES] = ([E]_{\text{total}} - [ES])[S]/K_m$$

$$[ES] = \frac{[E]_{\text{total}}[S]}{K_m + [S]}$$

Therefore the overall rate ( $v$ ) can be described by:

$$v = \frac{k_2[E]_{\text{total}}[S]}{K_m + [S]} \quad \text{or} \quad v = \frac{V_{\text{max}}[S]}{K_m + [S]}$$

This is termed the Michaelis-Menten equation, which in most cases, describes the relationship between the rate of enzyme catalysis and substrate concentration. Two important parameters are described by this equation,  $k_2$  (better known as  $k_{\text{cat}}$  and herein referred to as such) and  $K_m$ . The Michaelis constant,  $K_m$ , is the concentration of substrate when the reaction velocity is half of the maximal velocity, and is typically considered a measure of the binding affinity of a particular substrate with a given enzyme. The turnover number,  $k_{\text{cat}}$ , is the number of substrate molecules each enzyme site converts per unit time. If the total enzyme concentration is known it can be used together with  $k_{\text{cat}}$  to calculate the maximal velocity ( $V_{\text{max}}$ ):

$$V_{\text{max}} = k_{\text{cat}}[E]$$

$V_{\text{max}}$  represents the maximum kinetic rate for a fixed amount of enzyme.

where  $[S] \gg K_m$ ,

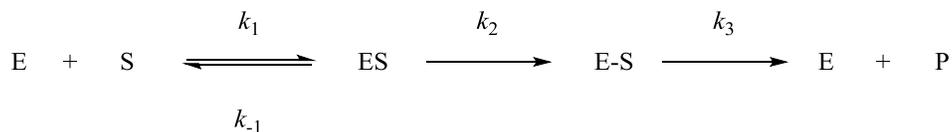
$$v = \frac{k_{\text{cat}}[E]_{\text{total}}[S]}{K_m + [S]} \sim \frac{k_{\text{cat}}[E]_{\text{total}}[S]}{[S]} = V_{\text{max}}$$

where  $[S] \ll K_m$ ,

$$v = \frac{k_{\text{cat}}[E]_{\text{total}}[S]}{K_m + [S]} \sim \frac{k_{\text{cat}}[E]_{\text{total}}[S]}{K_m} = \frac{k_{\text{cat}}}{K_m} [E][S]$$

In the second case the rate is directly proportional to substrate concentration and becomes a first-order reaction. Thus the constant  $k_{\text{cat}}/K_m$  can be used to compare the relative efficiency of catalysis for a given enzyme with different substrates and is referred to as the catalytic efficiency or the specificity constant.

In some cases, such as that described in this regarding TcTS, and additional step must be considered in the model:



Where  $k_2$  is now the chemistry rate constant, formation of the covalent intermediate (E-S), and  $k_3$  is the turnover rate constant, describing the cleavage of the covalent intermediate and release of product (P).

The catalysis constant is now described by a combination of  $k_2$  and  $k_3$  as below:

$$\frac{1}{k_{cat}} = \frac{1}{k_2} + \frac{1}{k_3}$$

$$k_{cat} = \frac{k_2 k_3}{(k_2 + k_3)}$$

If turnover is rate-limiting then  $k_2 \gg k_3$

$$k_{cat} = \frac{k_2 k_3}{(k_2 + k_3)} \sim \frac{k_2 k_3}{k_2} = k_3$$

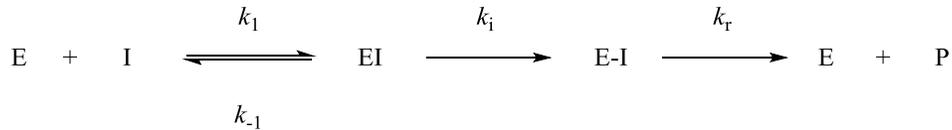
or if chemistry is rate-limiting  $k_3 \gg k_2$

$$k_{cat} = \frac{k_2 k_3}{(k_2 + k_3)} \sim \frac{k_2 k_3}{k_3} = k_2$$

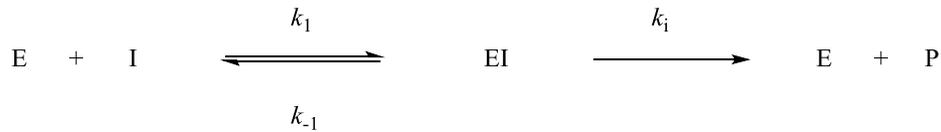
Therefore the rate-limiting step (glycosylation or deglycosylation) will correspond to the overall rates measured by experimental means with a selected substrate.

## A.2 Inactivation kinetics (irreversible inhibitors)

The mechanism-based inhibitors developed in this thesis act through irreversible inhibition by way of covalent bond formation with the enzyme. This relationship between inhibitor (I), free enzyme (E), enzyme-inhibitor complex (EI), covalent intermediate (E-I) and product (P) can be described as:



Where the rate constants are  $k_1$  (inhibitor binding rate constant),  $k_{-1}$  (inhibitor dissociation rate constant),  $k_i$  (inactivation rate constant) and  $k_r$  (reactivation rate constant). Typically, the inactivation rate constant is much greater than the reactivation rate constant ( $k_i \gg k_r$ ) and the relationship can be simplified as:



In most cases, the concentration of inhibitor is much greater than enzyme concentration ( $[\text{I}] \gg [\text{E}]$ ), and the overall rate ( $v$ ) of this relationship can be described analogously to that of the Michaelis-Menten equation:

$$v = \frac{k_i[\text{E}][\text{I}]}{K_i + [\text{I}]}$$

If the observed rate  $k_{\text{obs}}$  is defined as:

$$k_{\text{obs}} = \frac{k_i[\text{I}]}{K_i + [\text{I}]}$$

Then,

$$v = \frac{k_i[\text{E}][\text{I}]}{K_i + [\text{I}]} = k_{\text{obs}}[\text{E}]$$

Since  $k_{\text{obs}}$  is dependent on the inactivation rate constant, the apparent dissociation constant between the enzyme and inactivator,  $K_i$ , and the inactivator concentration, the relationship can be considered a pseudo-first order reaction. Plotting  $k_{\text{obs}}$  versus a range of inactivator concentrations allows for values of  $k_i$  and  $K_i$  to be determined. In the case where the dissociation constant is much greater than the experimentally achievable inactivator concentration ( $K_i \gg [\text{I}]$ ) then:

$$k_{\text{obs}} = \frac{k_i}{K_i} [I]$$

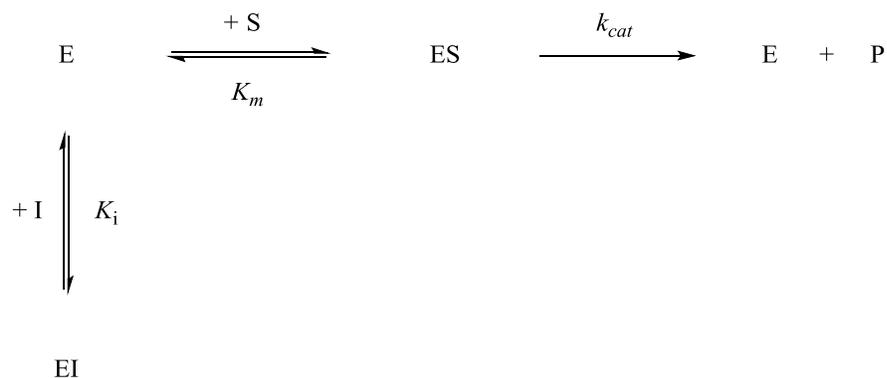
The second-order rate of inactivation  $k_i/K_i$  can instead be determined from this relationship.

### A.3 Inhibition kinetics (reversible inhibitors)

Reversible inhibition occurs via non-covalent interactions between an inhibitor and an enzyme species. This interaction can proceed through different modes of action, the kinetic schemes and equations describing them for three of the most common types – competitive, non-competitive and uncompetitive inhibition – are briefly described below. In each case the Michaelis-Menten scheme (described above) serves as the starting point for these analyses.

#### Competitive inhibition

Competitive inhibitors can only bind to the free enzyme species (E) and will reversibly form a catalytically inactive enzyme-inhibitor complex (EI). These inhibitors thus directly compete with substrate (S) for binding of the enzyme active site. The scheme for this interaction is:



where  $K_i = \frac{[\text{EI}]}{[\text{E}][\text{I}]}$  and  $K_m = \frac{[\text{ES}]}{[\text{E}][\text{S}]}$ .

The expression for the reaction rate ( $v$ ) can thus be described as:

$$v = \frac{V_{\text{max}}[\text{S}]}{[\text{S}] + K_m \left(1 + \frac{[\text{I}]}{K_i}\right)}$$

Otherwise shown as:

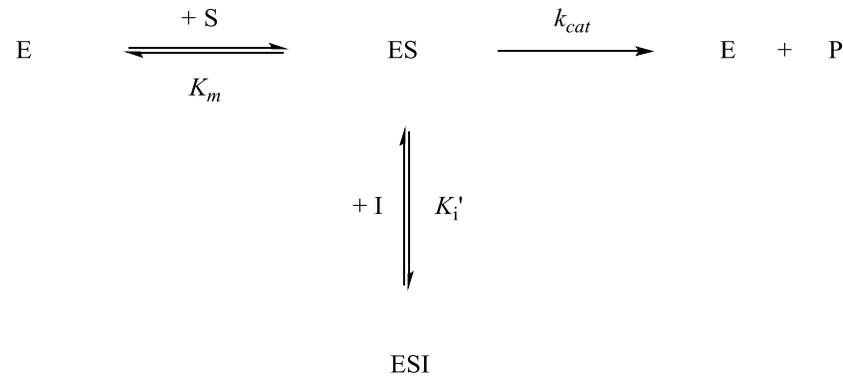
$$v = \frac{V_{max}[S]}{[S] + K_m^{app}}$$

where  $K_m^{app} = K_m \left(1 + \frac{[I]}{K_i}\right)$ .

Comparison of this expression to the Michaelis-Menten equation shows that competitive inhibitors will interfere with substrate binding – as shown by  $K_m^{app}$  – and that they do not affect  $V_{max}$  – since they cannot bind the enzyme-substrate (ES) complex. Practically, this means that by increasing substrate concentration,  $V_{max}$  can be achieved.

### Uncompetitive inhibition

Uncompetitive inhibitors can only bind to the enzyme-substrate complex (ES) leading to the catalytically incompetent enzyme-substrate-inhibitor complex ESI. This implies that the binding site for the inhibitor species only becomes accessible after the enzyme (E) has bound the substrate (S). Therefore, unlike for competitive inhibition, uncompetitive inhibition cannot be reversed by increasing substrate concentration. The scheme describing this relationship is:



where  $K_i' = \frac{[ESI]}{[ES][I]}$  and  $K_m = \frac{[ES]}{[E][S]}$ .

The expression for the reaction rate ( $v$ ) is described as:

$$v = \frac{V_{max}[S]}{[S] \left( 1 + \frac{[I]}{K_i'} \right) + K_m}$$

Otherwise shown as:

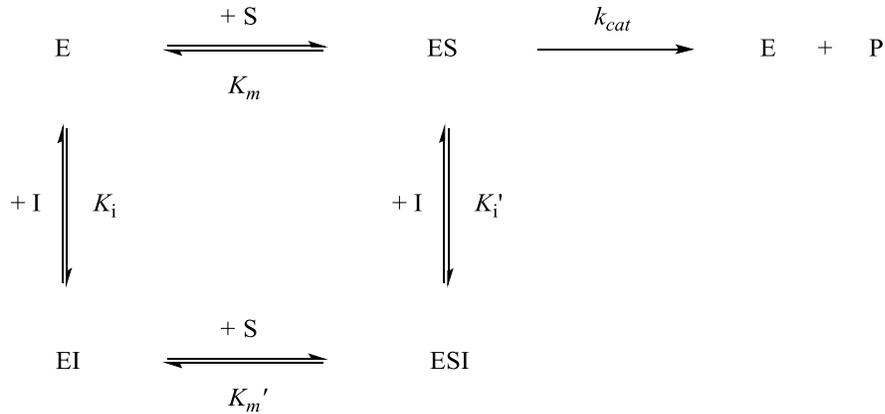
$$v = \frac{V_{max}^{app}[S]}{[S] + K_m^{app}}$$

where  $V_{max}^{app} = V_{max} \left( \frac{1}{1 + \frac{[I]}{K_i'}} \right)$  and  $K_m^{app} = K_m \left( \frac{1}{1 + \frac{[I]}{K_i'}} \right)$ .

The binding of inhibitor by the ES complex leads to a reduction in effective concentration of ES complex, thus resulting in a decrease to the maximum enzyme activity (as shown by  $V_{max}^{app}$ ). Further, the effective removal of ES complex by binding to inhibitor (forming ESI) leads to increased binding affinity for the substrate (S), thereby decreasing the value of the binding constant (as shown by  $K_m^{app}$ ).

### Non-competitive inhibition

Non-competitive inhibitors can bind with either the free enzyme (E) or the enzyme-substrate complex (ES) to afford the catalytically inactive enzyme-substrate-inhibitor species ESI. Typically these inhibitors bind at an allosteric site, away from the site of substrate binding. The scheme for this relationship is:



where  $K_i = \frac{[\text{EI}]}{[\text{E}][\text{I}]}$ ;  $K_m = \frac{[\text{ES}]}{[\text{E}][\text{S}]}$ ;  $K_i' = \frac{[\text{ESI}]}{[\text{ES}][\text{I}]}$ ;  $K_m' = \frac{[\text{ESI}]}{[\text{EI}][\text{S}]}$ .

We assume that the binding constant of the inhibitor with the free enzyme (E) or enzyme-substrate complex (ES) is equivalent ( $K_i = K_i'$ ) and that substrate binding to E or the enzyme-inhibitor complex is constant ( $K_m = K_m'$ ). Then the rate ( $v$ ) expression for this relationship is described as:

$$v = \frac{V_{max}[S]}{([S] + K_m) \left(1 + \frac{[I]}{K_i}\right)}$$

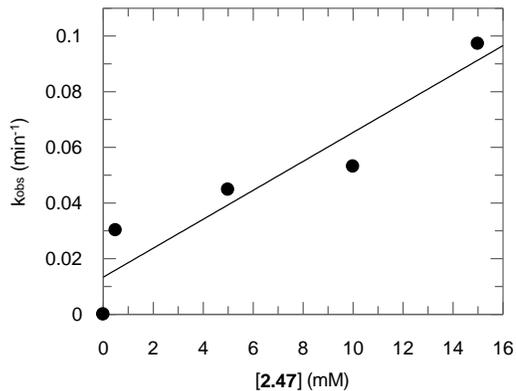
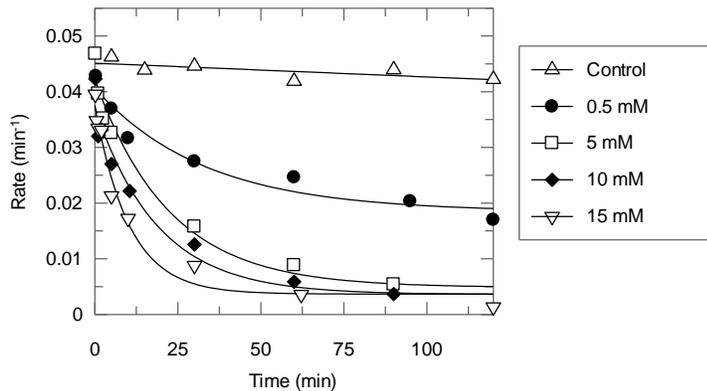
Otherwise shown as:

$$v = \frac{V_{max}^{app}[S]}{([S] + K_m)}$$

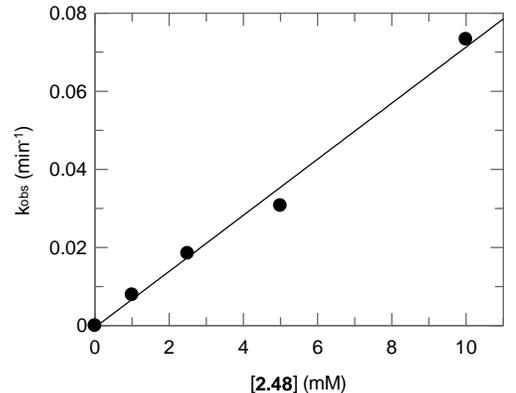
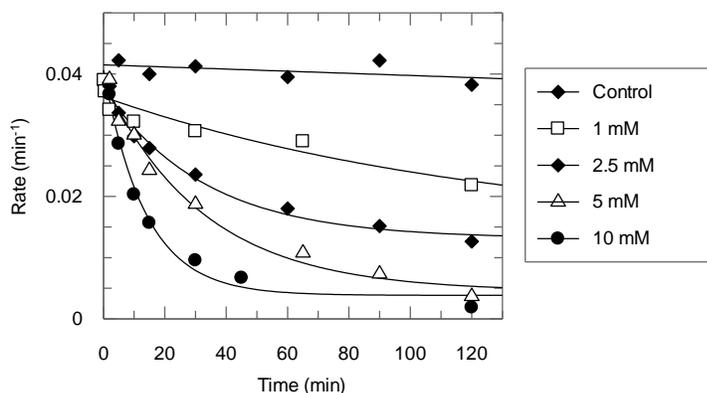
where  $V_{max}^{app} = V_{max} \left( \frac{1}{1 + \frac{[I]}{K_i}} \right)$ .

In this case, binding of the inhibitor (I) by the enzyme impedes the reaction efficiency thereby reducing the maximum activity of the enzyme (as shown by  $V_{max}^{app}$ ), yet does not affect binding of the substrate. Therefore the extent of inhibition is dependent on the concentration of inhibitor (I), while binding affinity of the substrate remains unchanged.

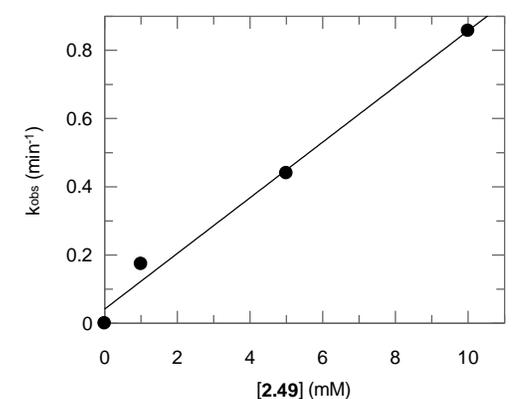
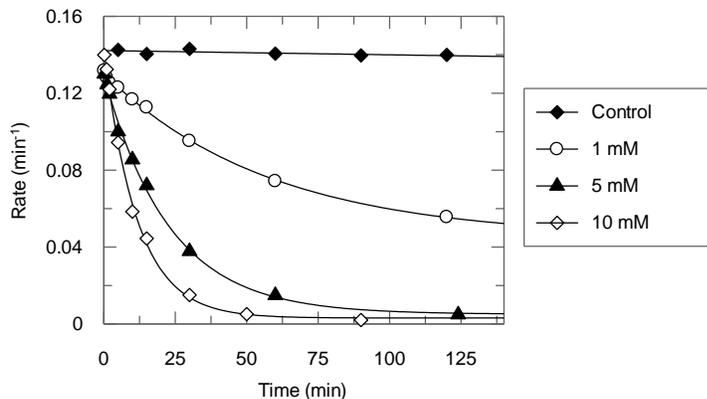
**Appendix B : Kinetic plots of TcTS inactivation by DFSA analogues**



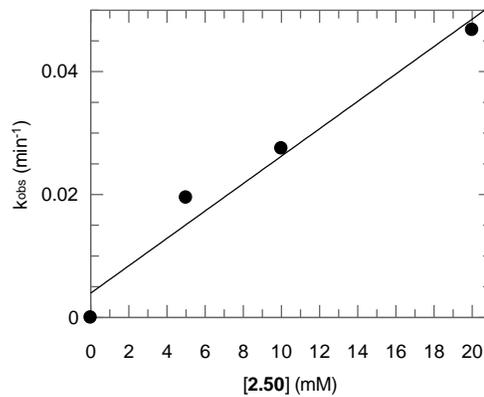
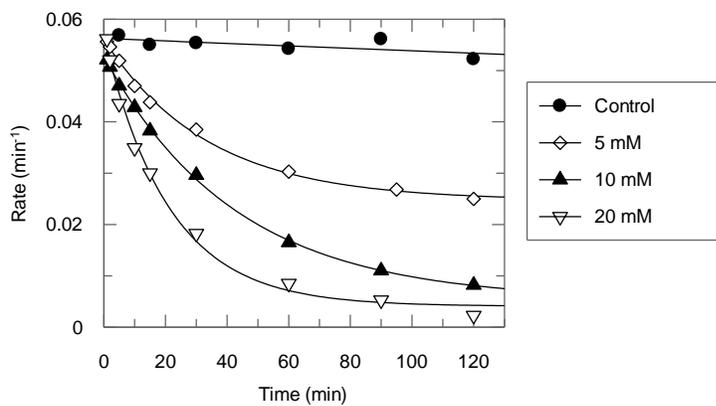
**Appendix Figure B.1. Inactivation plot 2.47 versus TcTS.**



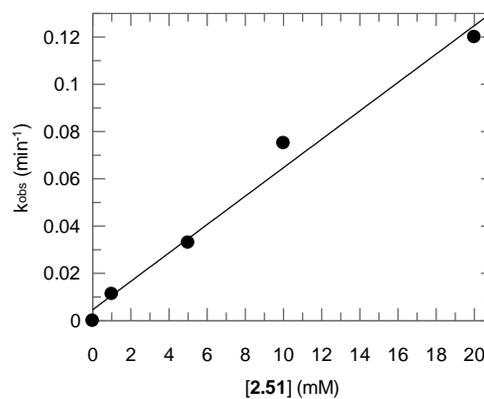
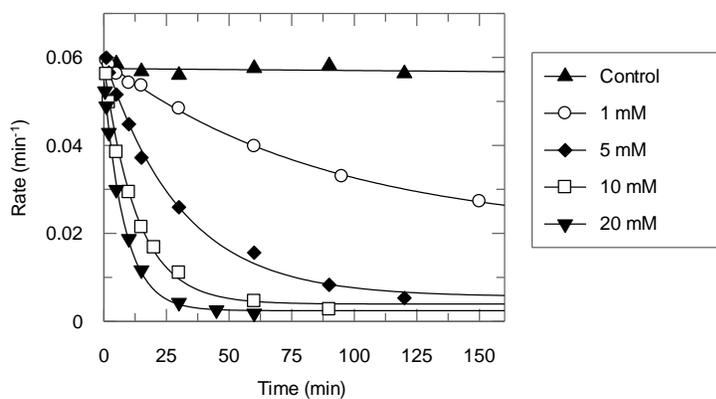
**Appendix Figure B.2. Inactivation plot 2.48 versus TcTS.**



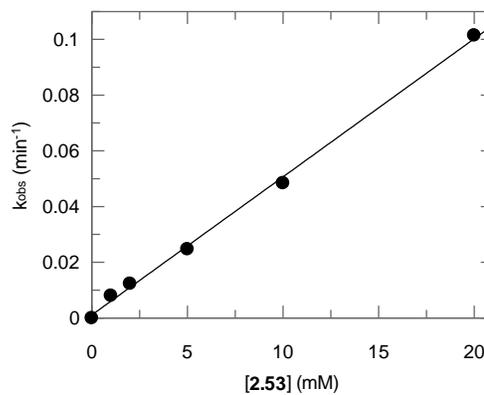
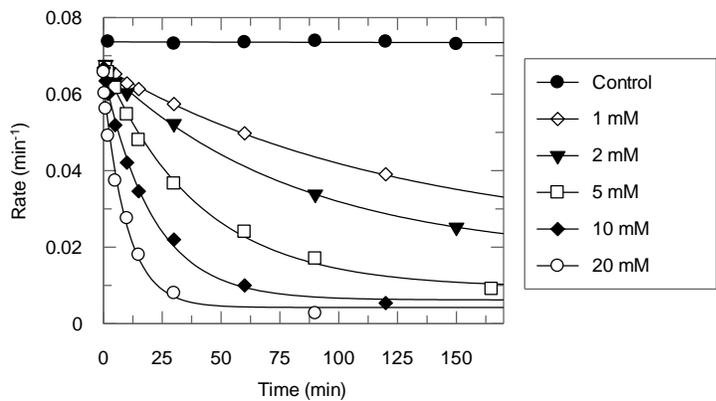
**Appendix Figure B.3. Inactivation plot 2.49 versus TcTS.**



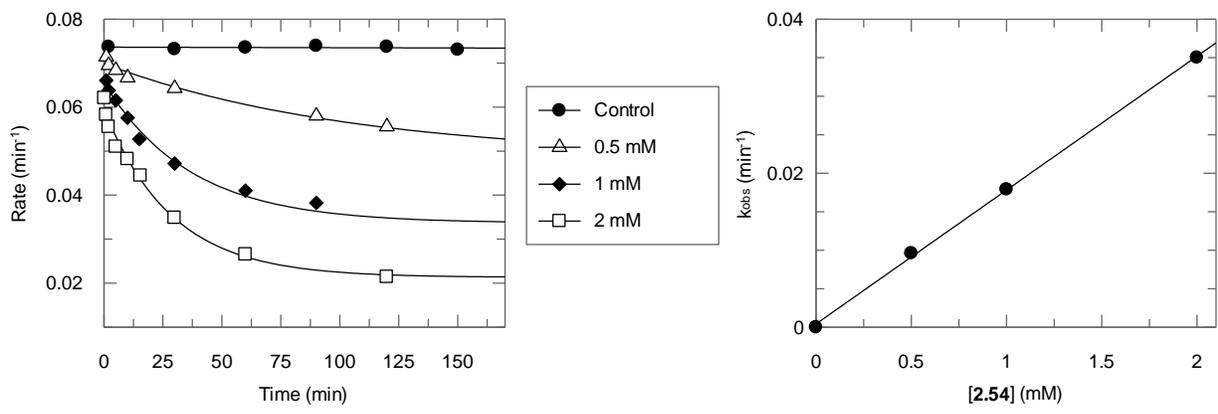
Appendix Figure B.4. Inactivation plot 2.50 versus TcTS.



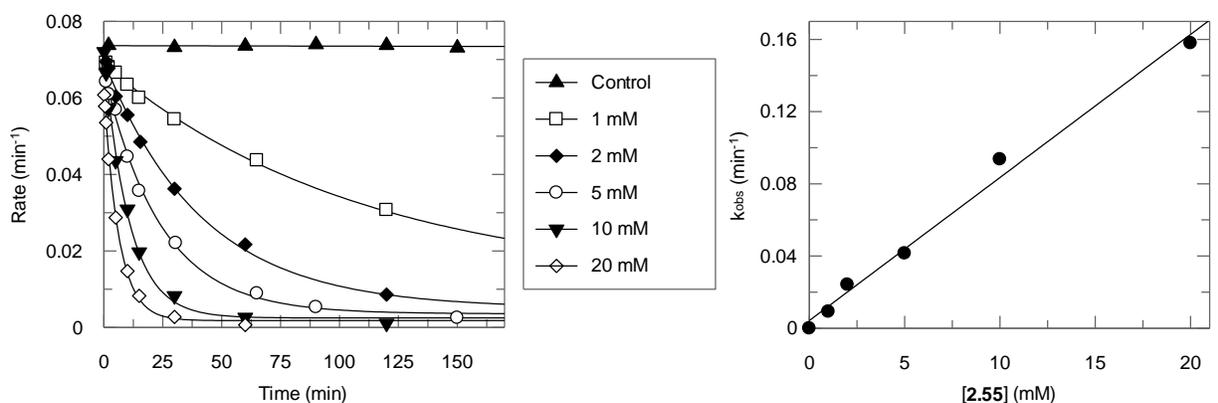
Appendix Figure B.5. Inactivation plot 2.51 versus TcTS.



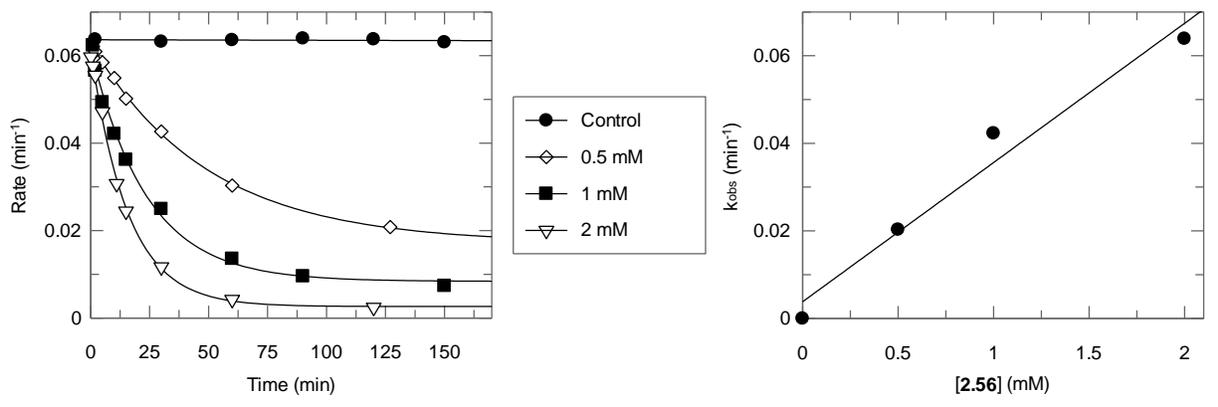
Appendix Figure B.6. Inactivation plot 2.53 versus TcTS.



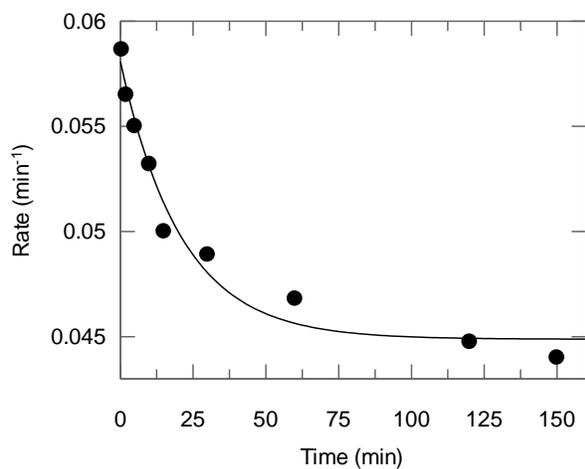
Appendix Figure B.7. Inactivation plot 2.54 versus TcTS.



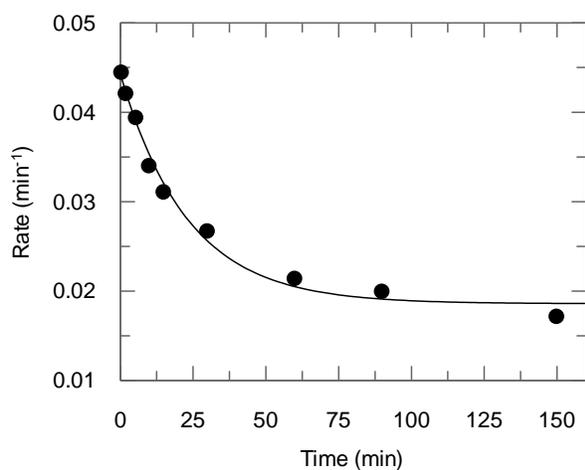
Appendix Figure B.8. Inactivation plot 2.55 versus TcTS.



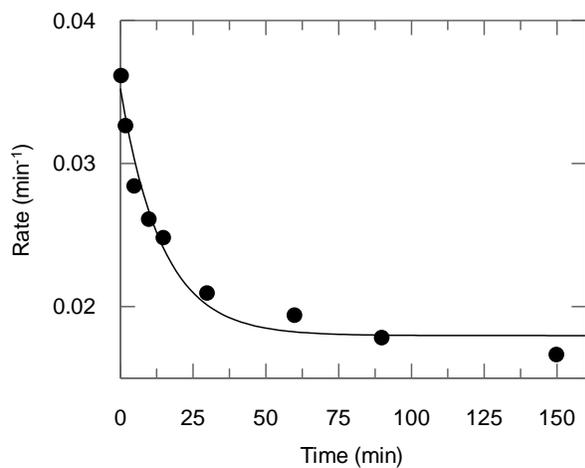
Appendix Figure B.9. Inactivation plot 2.56 versus TcTS.



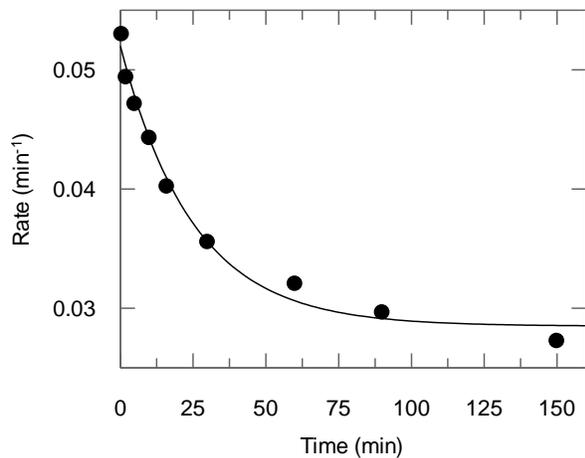
**Appendix Figure B.10. Inactivation plot 2.63 (5 mM) versus TcTS.**



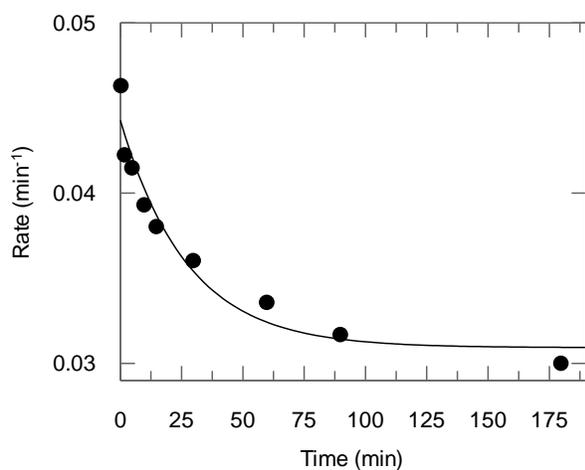
**Appendix Figure B.11. Inactivation plot 2.64 (5 mM) versus TcTS.**



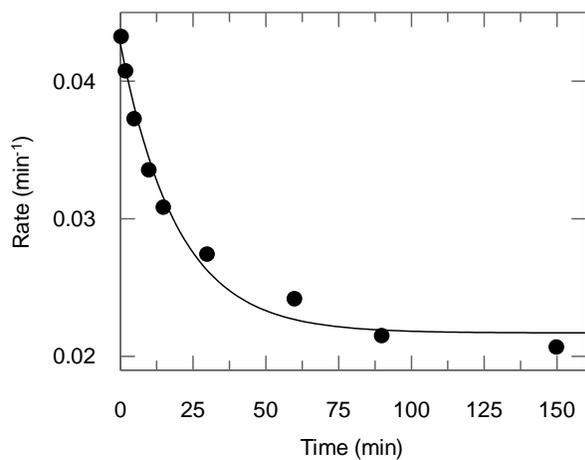
**Appendix Figure B.12. Inactivation plot 2.65 (5 mM) versus TcTS.**



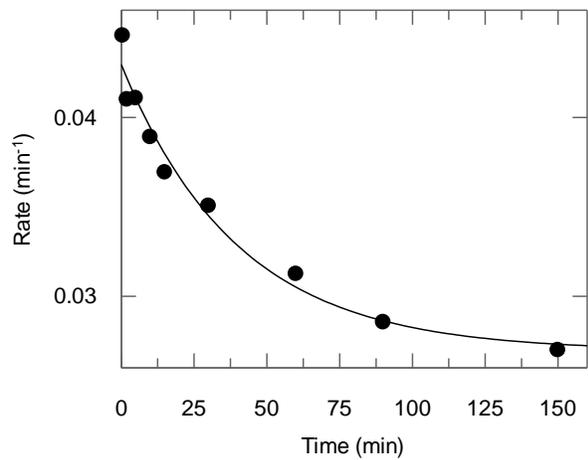
**Appendix Figure B.13. Inactivation plot 2.66 (5 mM) versus TcTS.**



**Appendix Figure B.14. Inactivation plot 2.67 (5 mM) versus TcTS.**

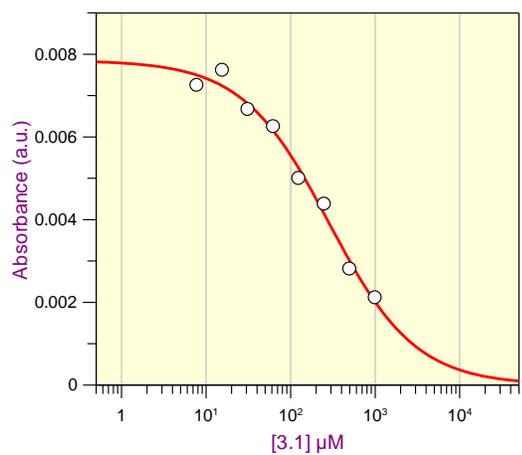


**Appendix Figure B.15. Inactivation plot 2.68 (5 mM) versus TcTS.**

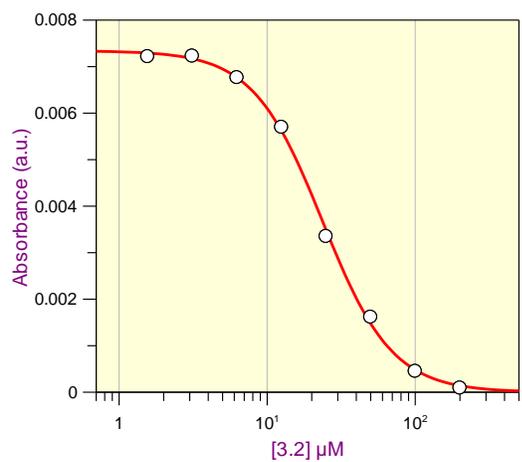


**Appendix Figure B.16. Inactivation plot 2.69 (5 mM) versus TcTS.**

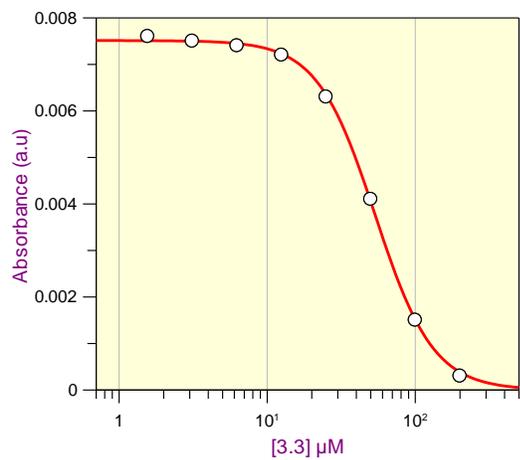
## Appendix C : Kinetic plots of TcTS inhibition by natural product and ircinialactam library hits



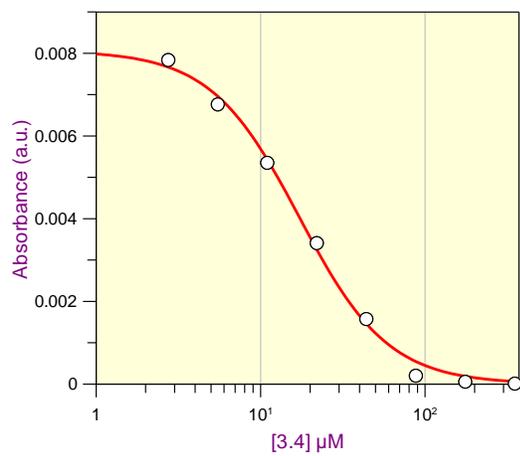
Appendix Figure C.1. 3.1 versus TcTS  $\text{IC}_{50}$ .



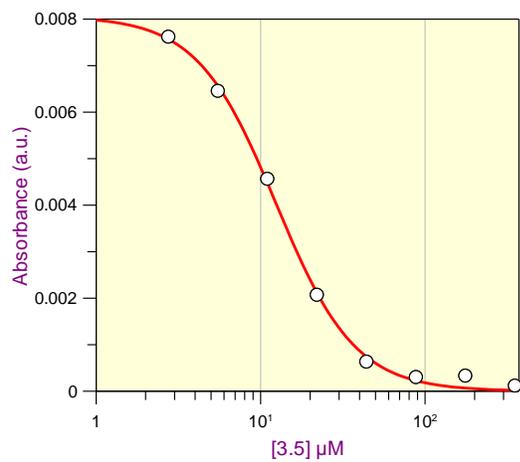
Appendix Figure C.2. 3.2 versus TcTS  $\text{IC}_{50}$ .



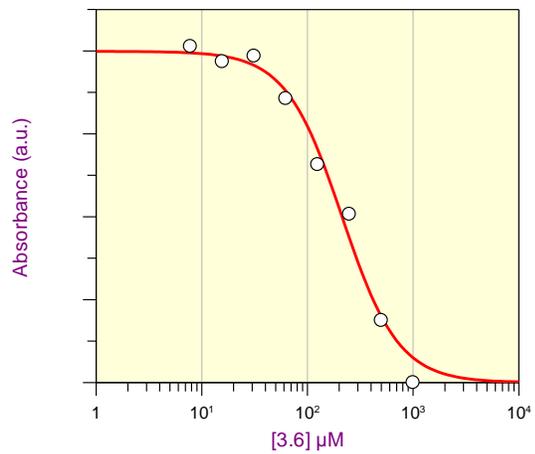
**Appendix Figure C.3. 3.3 versus TcTS  $\text{IC}_{50}$ .**



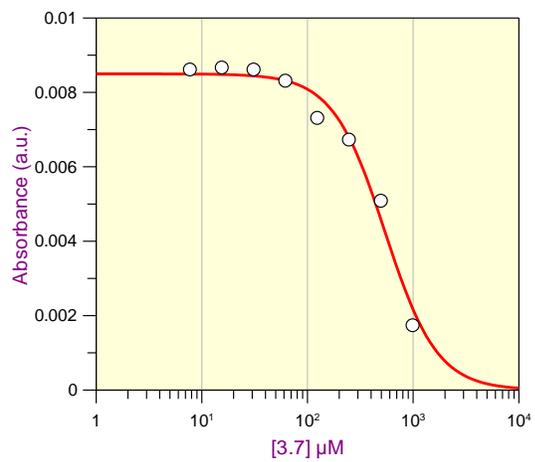
**Appendix Figure C.4. 3.4 versus TcTS  $\text{IC}_{50}$ .**



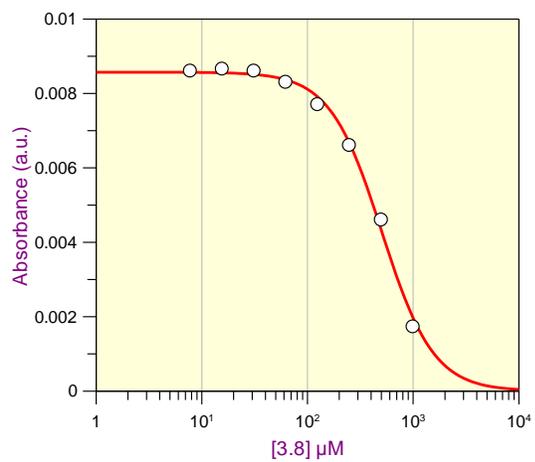
**Appendix Figure C.5. 3.5 versus TcTS  $\text{IC}_{50}$ .**



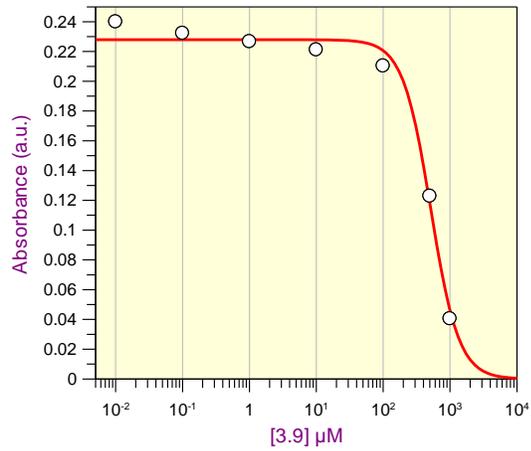
**Appendix Figure C.6. 3.6 versus TcTS  $\text{IC}_{50}$ .**



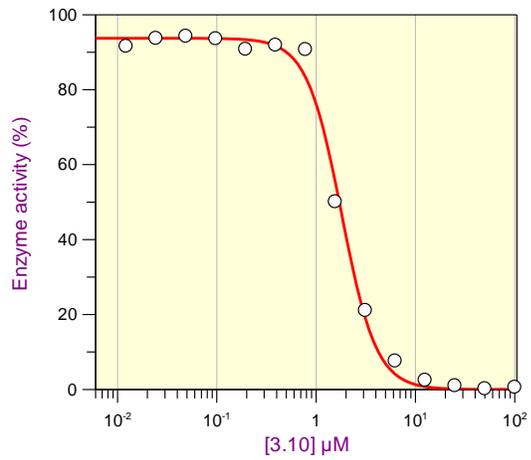
**Appendix Figure C.7. 3.7 versus TcTS  $\text{IC}_{50}$ .**



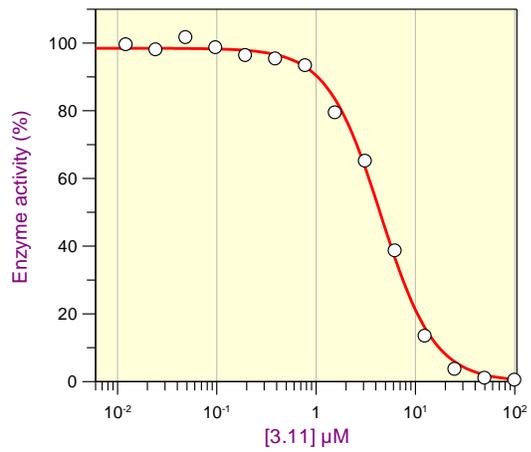
**Appendix Figure C.8. 3.8 versus TcTS  $\text{IC}_{50}$ .**



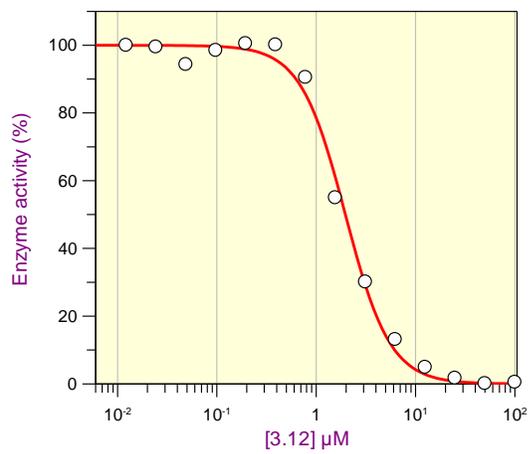
**Appendix Figure C.9. 3.9 versus TcTS IC<sub>50</sub>.**



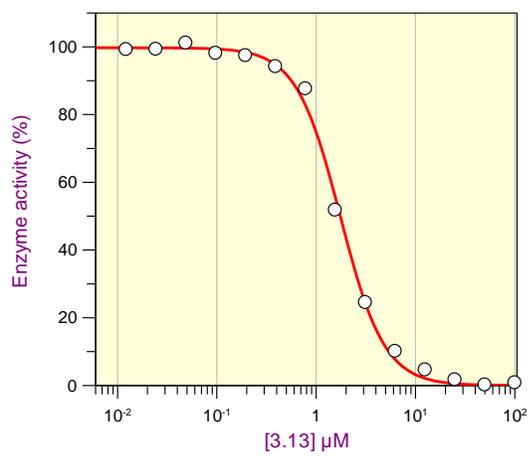
**Appendix Figure C.10. 3.10 versus TcTS IC<sub>50</sub>.**



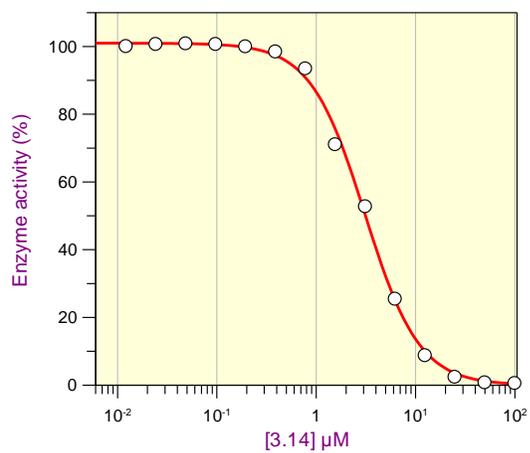
**Appendix Figure C.11. 3.11 versus TcTS IC<sub>50</sub>.**



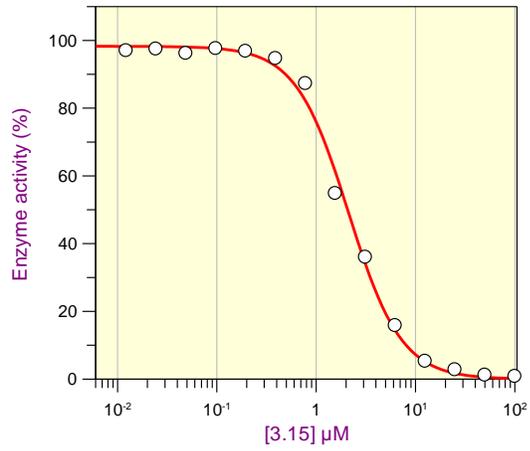
Appendix Figure C.12. 3.12 versus TcTS IC<sub>50</sub>.



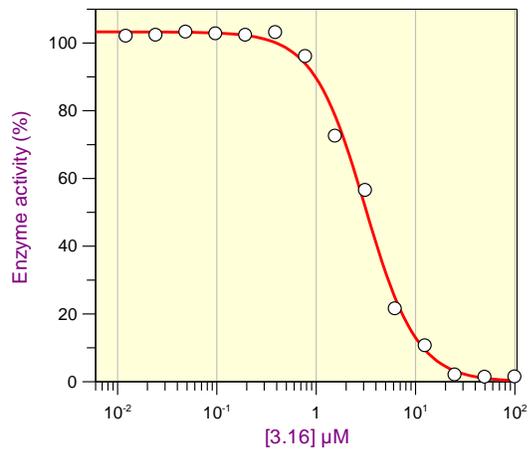
Appendix Figure C.13. 3.13 versus TcTS IC<sub>50</sub>.



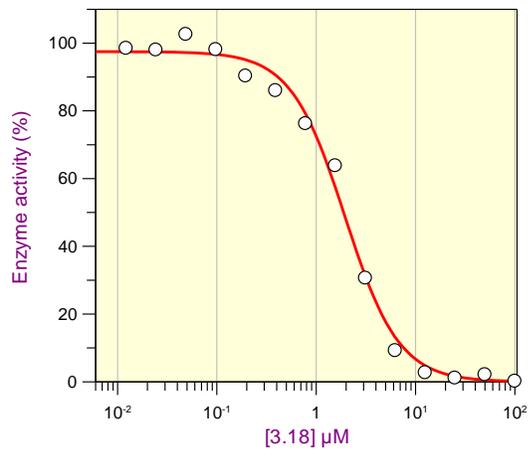
Appendix Figure C.14. 3.14 versus TcTS IC<sub>50</sub>.



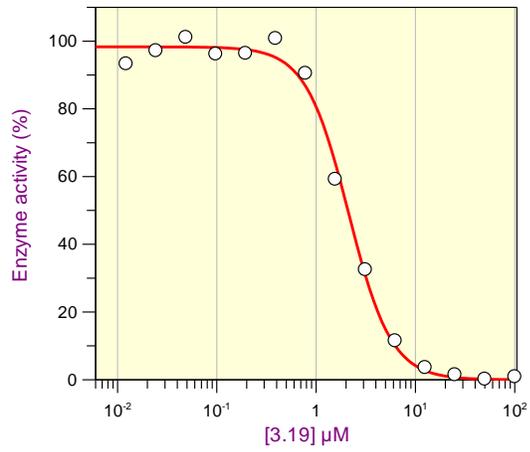
**Appendix Figure C.15. 3.15 versus TcTS IC<sub>50</sub>.**



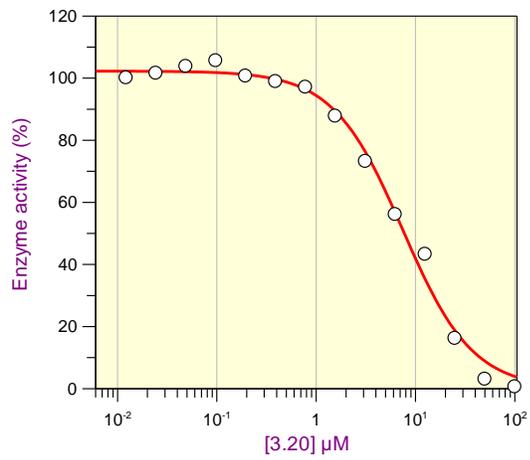
**Appendix Figure C.16. 3.16 versus TcTS IC<sub>50</sub>.**



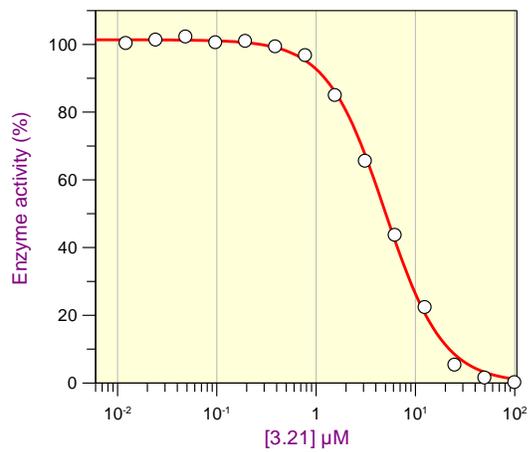
**Appendix Figure C.17. 3.18 versus TcTS IC<sub>50</sub>.**



**Appendix Figure C.18. 3.19 versus TcTS IC<sub>50</sub>.**

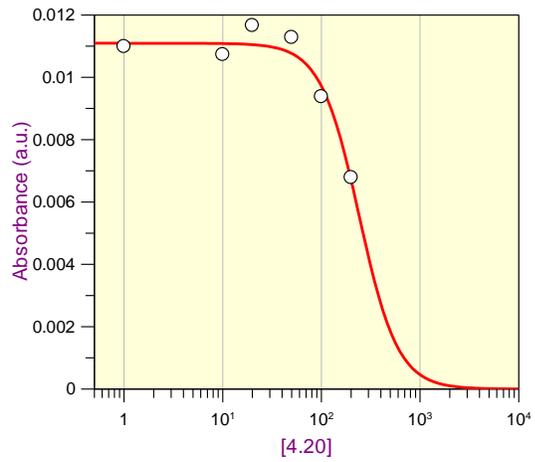


**Appendix Figure C.19. 3.20 versus TcTS IC<sub>50</sub>.**

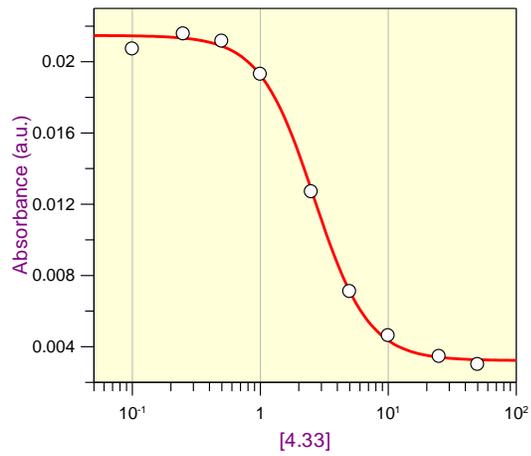


**Appendix Figure C.20. 3.21 versus TcTS IC<sub>50</sub>.**

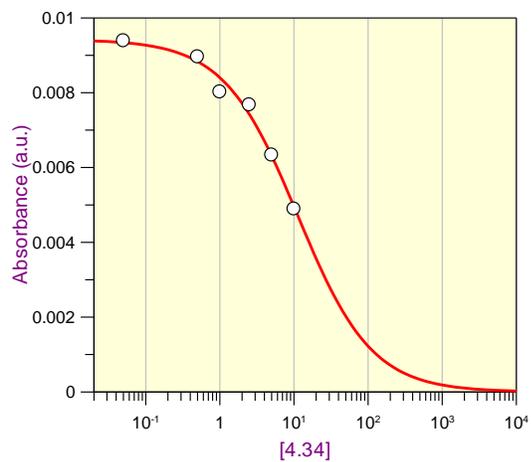
## Appendix D : Kinetic plots of TcTS inhibition by peptide hits



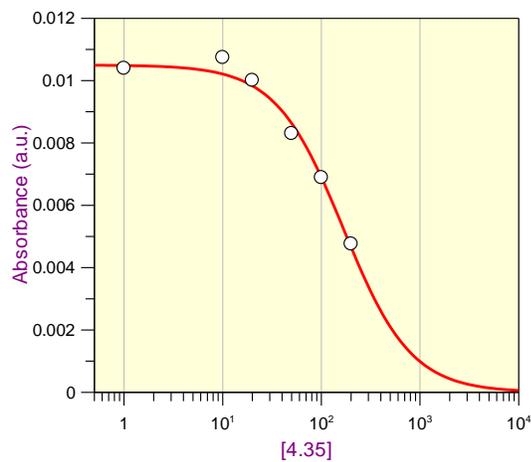
Appendix Figure D.1. 4.20 versus TcTS  $IC_{50}$ .



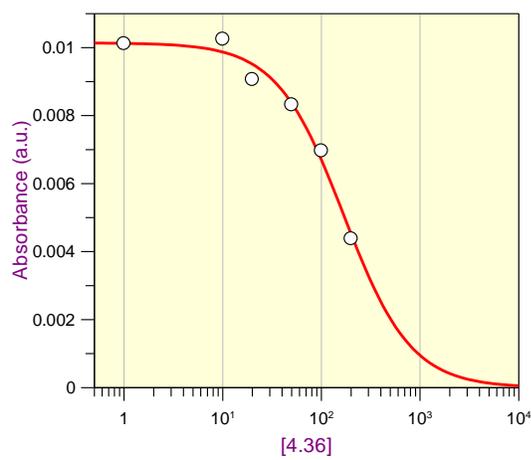
Appendix Figure D.2. 4.33 versus TcTS  $IC_{50}$ .



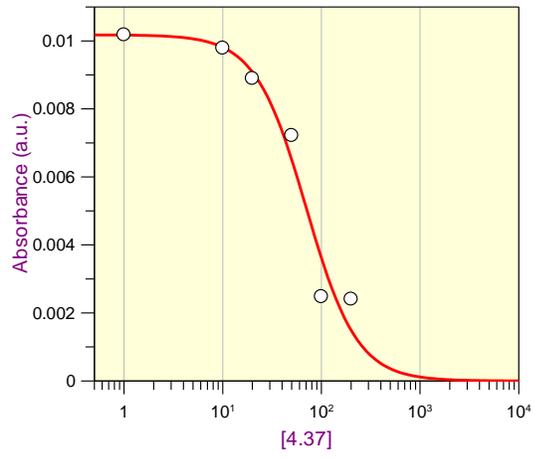
Appendix Figure D.3. 4.34 versus TcTS IC<sub>50</sub>.



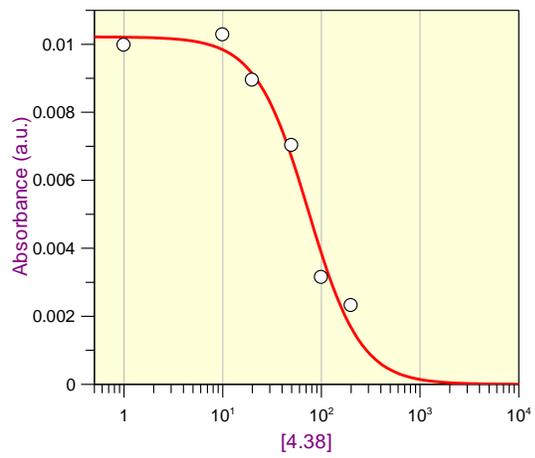
Appendix Figure D.4. 4.35 versus TcTS IC<sub>50</sub>.



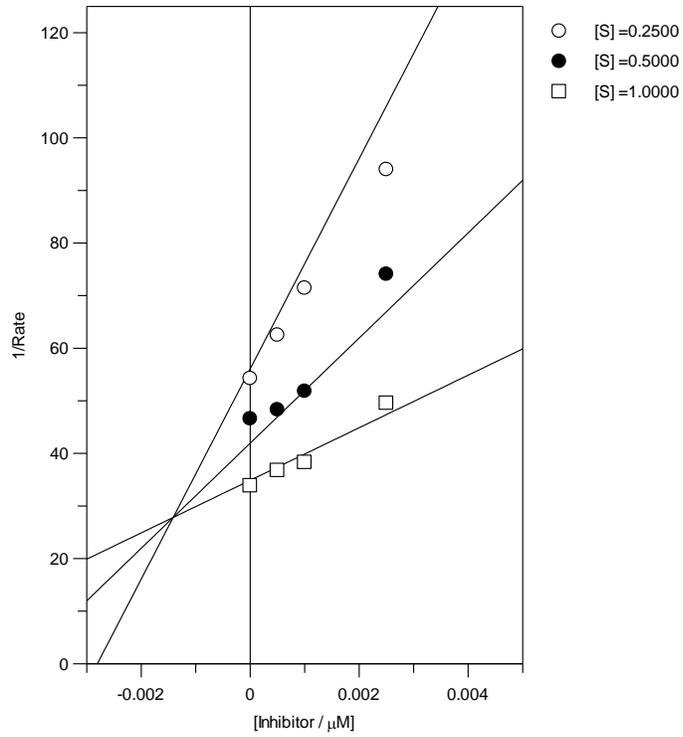
Appendix Figure D.5. 4.36 versus TcTS IC<sub>50</sub>.



**Appendix Figure D.6. 4.37 versus TcTS IC<sub>50</sub>.**



**Appendix Figure D.7. 4.38 versus TcTS IC<sub>50</sub>.**



**Appendix Figure D.8. 4.33 versus TcTS Dixon Plot.**