SYNTHESIS OF POLAR TERTIARY AMINE DENDRIMERS TERMINATED WITH GLUCOSAMINE

Michelle Watterson
B.A., California State University, Sacramento, 2005

THESIS

Submitted in partial satisfaction of the requirements for the degree of

MASTER OF SCIENCE

in

CHEMISTRY
(Biochemistry)

at

CALIFORNIA STATE UNIVERSITY, SACRAMENTO

FALL
2010
SYNTHESIS OF POLAR TERTIARY AMINE DENDRIMERS TERMINATED WITH GLUCOSAMINE

A Thesis

by

Michelle Watterson

Approved by:

__________________________, Committee Chair
Katherine McReynolds, Ph.D.

__________________________, Second Reader
Cynthia Kellen-Yuen, Ph.D.

__________________________, Third Reader
Claudia Lucero, Ph.D.

__________________________
Date
Student: Michelle Watterson

I certify that this student has met the requirements for format contained in the University format manual, and that this thesis is suitable for shelving in the Library and credit is to be awarded for the thesis.

__________________________, Graduate Coordinator
Brad Baker, Ph.D.                      Date

Department of Chemistry
Abstract

of

SYNTHESIS OF POLAR TERTIARY AMINE DENDRIMERS TERMINATED WITH GLUCOSAMINE

by

Michelle Watterson

This Thesis examines the synthesis of polar, tertiary amine dendrimers terminated with glucosamine. These glycodendrimers, once sulfated, could inhibit the binding of the HIV envelope glycoprotein, gp120, to host cells. Studies have shown that gp120 specifically binds to a variety of receptors, allowing HIV to infect a multitude of cell types. Upon further investigation, researchers have determined that gp120 uses sugar moieties on sphingolipids and sulfated polyanionic proteins as host cell receptors. Studies of analogs to these known gp120 receptors have shown that sulfated polyanionic sugars bind with a stronger affinity to gp120 than the host cell receptors themselves. Research has also shown that the infection process for HIV is multivalent in nature, which means multiple copies of gp120 bind to multiple copies of receptors simultaneously in order to gain access into the host cell. However, to mimic this process, multiple receptor analogs need to be clustered together on a scaffold.

Dendrimers are hyper-branched, three dimensional, globular macromolecules. They are comprised of three parts: 1) a core, 2) repeating units attached in a radial manner to the core, and 3) terminal functional groups attached to the outer surface. These macromolecules are synthesized either from the core out to the terminal groups,
known as divergent synthesis, or from the terminal groups inward toward the core, known as convergent synthesis. Due to the variability in synthesis, dendrimers can be any size and have any physical properties desired. Therefore, in this application, dendrimers allow gp120 receptor analogs to be positioned in a manner similar to what is found naturally on the surface of host cells. This work will show the successful synthesis of two core molecules [TEAₘ-(G1)₃, TEAₘ-G1-(DAP)₃], a glucosamine-terminated linker (G-AE-PA), as well as two glucosamine-terminated dendrimers [TEAₘ-G1-(GlcN)₃, TEAₘ-G1-(G-AE-PA)₃]. TEAₘ-(G1)₃ was synthesized with a 99.6% yield, using a Michael-like addition and a trifluoroacetic acid (TFA) deprotection. TEAₘ-G1-(DAP)₃ was synthesized in a similar manner, using two addition methods: methyl ester formation and amidation resulting in a 29.3% yield of product. G-AE-PA was synthesized with a 14.2% yield, using benzotriazol-1-yloxy-tris(dimethylamino) phosphonium hexafluorophosphate (BOP)-mediated amide coupling followed by a TFA deprotection. Both dendrimers, TEAₘ-G1-(GlcN)₃ and TEAₘ-G1-(G-AE-PA)₃ were synthesized with a 32.8% and 61.4% yield, respectively. This was accomplished using a combination of three steps: Michael-like addition, TFA deprotection and BOP-mediated amide-coupling. The successful synthesis and future sulfation of these dendrimers could lead to the creation of viable drugs which have the ability to inhibit the binding of gp120 onto host cells, thereby inhibiting infection by HIV.
ACKNOWLEDGMENTS

I would like to extend my heartfelt gratitude to a number of people who have made the completion of this Thesis possible. First, I must thank Dr. McReynolds, who not only served as my research advisor but also, challenged, encouraged and supported me throughout this process. I can honestly say, that this Thesis could not have been written without her. I would also like to thank my committee, Dr. Kellen-Yuen and Dr. Lucero who have patiently guided me through the writing process, never accepting anything less than my best efforts.

Next, I would like to thank all of the members of the McReynolds research group. The camaraderie we share and the readiness everyone has to help one another, when, where, and however they can, has made my experience in the Masters program a truly wonderful one. There are a few group members (Russ, Rachel and Alex), whose help and friendship have directly assisted me and my research. Russ and Rachel have sacrificed countless evenings and weekends to accompany me in the lab. Without their sacrifices, I never would have completed the lab work. Lastly, Alex graciously performed quite a bit of ‘grunt’ work in synthesizing starting materials, so that I could focus on the synthesis of new products. Without his help I would still be toiling in lab to this day.

As important as my research group has been during my time in the Masters program, I have to say I never could have succeeded without the love and support from my family. At times when I was discouraged and frustrated, they did their best to lift my
spirits and keep me focused. My husband, Billy, always knew what to say to get me back in the lab and keep my working. Just as Russ and Rachel sacrificed their evenings and weekends, so did my husband. Billy has put his life on hold, so that he could be there whenever I needed him, and I am forever thankful for that. I also have to thank mother, Vicki, who patiently listened to me go on and on in quite some depth about the many reactions I have attempted during my time in lab. She did this despite the fact that she probably didn’t understand even half of what I was talking about. My mom was always there to listen to the good or the bad and her support has been integral for my success.

Finally, I have to thank God, through whom everything is possible and without whom I never would have even attempted this great feat.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>List of Schemes</td>
<td>xi</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td><strong>1. BACKGROUND</strong></td>
<td>1</td>
</tr>
<tr>
<td>Introduction to HIV</td>
<td>1</td>
</tr>
<tr>
<td>The Mechanisms by which HIV Binds to a Host Cell</td>
<td>2</td>
</tr>
<tr>
<td>Overview HIV Replication</td>
<td>7</td>
</tr>
<tr>
<td>Current HIV Treatments</td>
<td>11</td>
</tr>
<tr>
<td>Drug Resistance</td>
<td>15</td>
</tr>
<tr>
<td>Drug Side Effects</td>
<td>18</td>
</tr>
<tr>
<td>Current Research for Treating HIV</td>
<td>19</td>
</tr>
<tr>
<td>Anionic Sulfated Compounds</td>
<td>26</td>
</tr>
<tr>
<td>Multivalency</td>
<td>27</td>
</tr>
<tr>
<td>Dendrimers</td>
<td>30</td>
</tr>
<tr>
<td>Proposal</td>
<td>31</td>
</tr>
<tr>
<td><strong>2. RESULTS AND DISCUSSION</strong></td>
<td>33</td>
</tr>
<tr>
<td>Compounds Synthesized</td>
<td>34</td>
</tr>
<tr>
<td>Future Work</td>
<td>54</td>
</tr>
<tr>
<td><strong>3. CONCLUSIONS</strong></td>
<td>58</td>
</tr>
<tr>
<td><strong>4. EXPERIMENTAL</strong></td>
<td>59</td>
</tr>
<tr>
<td>General Information</td>
<td>59</td>
</tr>
<tr>
<td>Compounds Synthesized</td>
<td>60</td>
</tr>
<tr>
<td><strong>Appendix A.</strong>  $^1$H NMR Spectra</td>
<td>67</td>
</tr>
<tr>
<td><strong>Appendix B.</strong>  $^{13}$C NMR Spectra</td>
<td>73</td>
</tr>
<tr>
<td>References</td>
<td>79</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Binding of HIV glycoprotein complex to host cell</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Fusion of viral and host cell membranes</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Gp120 binding domains</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>HIV structure and life cycle</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>The process of HIV integrase as it alters HIV DNA and integrates it into the host cell DNA</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>Structures of HIV drugs</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>Diagram of the HIV replication cycle signifying where the different classes of drugs interfere</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>(+)-Cananolide A</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>Microbicide SPL7013</td>
<td>24</td>
</tr>
<tr>
<td>10</td>
<td>Heparan sulfate (HS) and heparin sulfate</td>
<td>27</td>
</tr>
<tr>
<td>11</td>
<td>Multivalency</td>
<td>29</td>
</tr>
<tr>
<td>12</td>
<td>Structure of galactosyl ceramide sulfatide (SGalCer)</td>
<td>29</td>
</tr>
<tr>
<td>13</td>
<td>Dendrimer generations</td>
<td>30</td>
</tr>
<tr>
<td>14</td>
<td>Compound 1, $^1$H NMR</td>
<td>36</td>
</tr>
<tr>
<td>15</td>
<td>Compound 2, $^1$H NMR</td>
<td>37</td>
</tr>
<tr>
<td>16</td>
<td>Compound 3, $^1$H NMR</td>
<td>40</td>
</tr>
<tr>
<td>17</td>
<td>Glucosamine isomers</td>
<td>41</td>
</tr>
<tr>
<td>18</td>
<td>Compound 6, $^1$H NMR</td>
<td>45</td>
</tr>
<tr>
<td>19</td>
<td>Stick diagram of the diastereomeric proton splitting on Compound 6</td>
<td>45</td>
</tr>
</tbody>
</table>
20. Compound 7, $^1$H NMR .................................................................49
21. Compound 8, $^1$H NMR .................................................................53
22. Analysis of the splitting pattern of the multiplet in Compound 8..............53
# LIST OF SCHEMES

<table>
<thead>
<tr>
<th>No.</th>
<th>Scheme Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Synthesis of the dendrimer core</td>
<td>34</td>
</tr>
<tr>
<td>2.</td>
<td>Synthesis of generation 0 glycodendrimer</td>
<td>39</td>
</tr>
<tr>
<td>3.</td>
<td>Synthesis of sugar linker</td>
<td>43</td>
</tr>
<tr>
<td>4.</td>
<td>Synthesis of generation 0.5 glycodendrimer</td>
<td>47</td>
</tr>
<tr>
<td>5.</td>
<td>Synthesis of the amine core</td>
<td>51</td>
</tr>
<tr>
<td>6.</td>
<td>Branching of the amine core</td>
<td>54</td>
</tr>
<tr>
<td>7.</td>
<td>Carboxylic acid derivative of the branched core</td>
<td>55</td>
</tr>
<tr>
<td>8.</td>
<td>Synthesis of generation 1 glycodendrimer</td>
<td>55</td>
</tr>
<tr>
<td>9.</td>
<td>Synthesis of generation 1.5 glycodendrimer</td>
<td>56</td>
</tr>
</tbody>
</table>
Chapter 1

BACKGROUND

Introduction to HIV

HIV (human immunodeficiency virus) belongs to a class of viruses known as retroviruses that have RNA (ribonucleic acid) instead of DNA (deoxyribonucleic acid) for the genetic material. More specifically, HIV belongs to a subclass of retroviruses known as lentiviruses. Lentiviruses are characterized by the long period of time between the moment of infection and the onset of symptoms, hence being dubbed “slow” viruses. If HIV is left untreated, it will result in the onset of AIDS (acquired immune deficiency syndrome). The diagnosis of AIDS results when a patient has a combination of an infection from an opportunistic pathogen, such as pneumonia, and a T-cell count of less than 200 per milliliter of blood.¹

HIV has infected roughly 60 million people since the beginning of the epidemic in 1981, and has killed about 25 million worldwide. As of December 2009, there were 33.4 million people living with HIV; 2.7 million of whom were newly infected, and 2 million who died that year worldwide.² According to a recent estimate performed by the Centers for Disease Control and Prevention (CDC), there will be over 1 million people living with HIV within the United States by the end of this year. Of those infected, 21% of them will be unaware of their HIV status and will therefore potentially transmit the disease further. It is also estimated that by the end of this year, there will be 56,300 Americans who will have become newly infected, and 18,000 who will have died.³
The speed with which the virus mutates has made finding a cure most difficult. Also, the manner in which the virus spreads throughout the body has limited the possible avenues of treatment. To further understand the difficulties and limitations of developing new therapies, one must first take a look at how the virus infects host cells and replicates within them.

**The Mechanisms by which HIV Binds to a Host Cell**

The binding process of HIV to a host cell consists of three steps: 1) the HIV glycoprotein complex binds to a host cell receptor, 2) the variable loops on the glycoprotein complex bind to an additional co-receptor on the host cell and 3) the HIV transmembrane glycoprotein complex inserts into the host cell membrane, resulting in a fusion pore. The first step is HIV binding to the host cell through the trimeric envelope glycoprotein complex found on the surface of the virus (Figure 1). This complex consists of three copies of the envelope glycoprotein, gp120, which are attached non-covalently, but through electrostatic interactions, to three copies of the transmembrane glycoprotein gp41. This trimeric complex binds to host cells either in a CD4-dependent or a CD4-independent fashion, depending on the concentration of the CD4 receptor found on the host cell surface. Cells with an abundant amount of CD4, such as the A 3.01 T-cell line, are infected via the CD4-dependent mechanism. Cells with minimal, or no CD4 receptors present (such as monocytes, macrophages and microglial cells), are infected via the CD4-independent mechanism.
Figure 1. Binding of HIV glycoprotein complex to host cell. First, the gp120 binds to the host cell CD4. Next, the variable loops (V3 loops) located on the gp120 associate with either CCR5 or CXCR4. This allows gp41 to bind to the cell membrane and draw the viral membrane and host membrane together until the two fuse, at which point the gp120 is sloughed off, and the virus will inject the core/viral contents into the host cell.\(^7\)

**CD4-Dependant Mechanism**

The CD4-dependent mechanism for binding HIV to the host cell is initiated by gp120 binding to CD4, resulting in a conformational change in the gp120 molecule (Figure 1). This moves the first and second variable loops (V1 & V2) on gp120 and allows for the exposure of the basic residues on the third variable loop (V3) of the gp120
molecule. This third variable loop then binds to the anionic sulfated tyrosine residues on the co-receptor with 100-1000 times greater affinity than the initial gp120-CD4 binding event.\textsuperscript{4-6,8-11} The co-receptor that is bound is either the alpha-chemokine receptor (CXCR4), or the C-C motif chemokine receptor (CCR5), depending on which co-receptor is present on the host cell surface at the point of infection.\textsuperscript{8} Since CCR5 has a 100-fold stronger binding affinity to gp120 (K\textsubscript{D} = 4 nM) than CXCR4 (K\textsubscript{D} = 200 nM), HIV-1 preferentially binds to CCR5, and therefore will only bind to CXCR4 in the absence of CCR5.\textsuperscript{12-13} The co-receptor binding event causes another conformational change in gp120 which initiates the formation of a trimeric coiled-coil structure in the N-terminal heptad region (HR1) of gp41, resulting in fusion of the peptide into the host cell and the shedding of the gp120 which is the final step in HIV binding (Figure 2). The shedding of gp120 causes a conformational change in gp41 where the C-terminal heptad region (HR2) binds to the outside of the HR1 supercoil, forming a 6-helix bundle, which mechanically draws the host and viral membranes together until the two fuse.\textsuperscript{6,9}
Figure 2. Fusion of viral and host cell membranes. The 3-helical supercoil HR2 (represented in red) of the gp41 binds to the 3-helical supercoil HR1 (represented in blue), which creates a 6-helical bundle that draws and fuses the HIV and the host cell membranes.\textsuperscript{14}

\textbf{CD4-Independent Mechanism}

The CD4-independent mechanism for HIV binding to the host cell, suggests that HIV gp120 binds to cell surface heparan sulfate proteoglycans, more specifically syndecan transmembrane receptors that have distal heparan sulfate chains.\textsuperscript{5,15-16} The polyanionic sulfates associated with the heparan sulfate chains interact with four different domains on gp120. All of these domains have stretches of basic amino acids within them and three of the four domains are in close proximity to one another. The first domain is the N-terminus of the V2 loop, the second is at the very front of the V3 loop, the third is
at the C-terminal domain of gp120, and the fourth is on the CD4-induced bridging sheet (Figure 3). A few residues up from the second heparin sulfate-binding domain on gp120 is the CCR5 binding site. As a result, it is has been proposed that the binding of heparin sulfate on gp120 induces a similar conformational change in gp120 as CCR5. This allows gp41 to bind and fuse with the host cell membrane as described above. This interaction has a stronger binding affinity ($K_D=10 \text{ nM}$) than the gp120-CD4 interaction ($K_D=22 \text{ nM}$), which brings up interesting possibilities for future inhibition with other anionic sulfated compounds.

Figure 3. Gp120 binding domains. A) gp120 (light blue) bound to CD4, red. B) The structure of gp120, where the inner domain is red, the outer domain is yellow and the bridging sheet is blue. C) A closer look at gp120 (light blue), bound to CD4, (red) and co-receptor CCR5, (green) or co-receptor CXCR4, (orange).
Overview of HIV Replication

HIV replication is a complex series of events that follows the fusion of HIV to the host cell. Essentially, the virus injects its capsid into the host cell cytoplasm, where the viral components within the capsid are then released. At this point, these viral components reprogram the host cell to transcribe, then integrate viral DNA into the host cell genome, resulting in the synthesis and release of new HIV virions (Figure 4). However, before all of this can occur, the capsid must be uncoated to allow the viral components access to the host cell cytoplasm. This is achieved with assistance from a host cell protein, peptidyl-prolyl isomerase (Pin1). The viral capsid consists of multiple capsid protein (CA) isoforms interacting with one another to create a stable protein matrix. Pin1 phosphorylates the protein CA, which results in the disruption of protein-protein interaction within the matrix, therefore uncoating the capsid.\textsuperscript{23} Once released into the host cell cytoplasm, the viral components then recruit other host cell proteins to initiate replication. These viral elements consist of two copies of ribonucleic acid (RNA) and three viral enzymes: reverse transcriptase (RT), integrase and protease.\textsuperscript{24} Since humans carry their genetic information in double stranded DNA and HIV carries its genetic information in single stranded RNA, the virus requires a helper (RT) to translate its single stranded RNA into double stranded DNA.
Figure 4. HIV structure and life cycle. Life cycle steps: 1) Binding of HIV to the host cell. 2) Fusion of the HIV and host cell membranes followed by release of the viral components into the cell. 3) Uncoating of the viral capsid. 4) Production of viral DNA from viral RNA with RT. 5) Integration of viral DNA into the host DNA. 6) Production of viral RNA from the host-virus DNA hybrid. 7) Synthesis of long protein strands. 8) Cleavage of the long protein strands into individual viral components. 9) Packaging of these components into viral particles, followed by the budding of these viral particles off of the host cell.
Reverse transcriptase is a DNA polymerase enzyme with a ribonuclease H (RNase H) active site. RNase H is responsible for transcribing viral DNA from the viral RNA, and creating a RNA/DNA hybrid double helix. RNase H then degrades the attached RNA strand and transcribes a complementary viral DNA strand, thus forming a viral DNA double helix. Once the viral DNA complex is finished, viral integrase then cleaves nucleotides off of the 3’ end of the viral DNA, leaving it “sticky”. Next, integrase transports the modified DNA complex into the host cell nucleus, cleaves the host cell DNA to integrate the viral DNA, and allows for the formation a host-viral DNA hybrid (Figure 5). When the host cell is activated host cell proteins will transcribe the hybrid DNA into messenger RNA (mRNA). After post-transcriptional modification of the mRNA, it is transported outside of the nucleus via host cell proteins. Once in the cytoplasm, ribosomes make long continuous protein strands from the mRNA. Protease then cuts these large, inactive strands into the individual active viral proteins, such as the viral capsid, the viral envelope proteins, accessory proteins (Tat & Rev), and enzymes (RT, integrase and protease). The viral protein Tat activates the host cell, which extends the time the host cell spends in the growth phase. This in turn lengthens the amount of time in which the host cell transcribes DNA and increases the amount of viral RNA and protein produced. The viral protein Rev integrates into the host cell nucleus and transports unmodified viral genomic RNA outside of the nucleus. When all of the viral components have been remade by the host cell, they are then ready to be packaged into viral particles and released from the cell via a budding process (Figure 4). As all of the viral components use the host cell machinery to complete their tasks, treatment
methods are limited to ones that won't attack and thereby destroy normal healthy host
cells. This has lead to the development of 6 different classes of drugs used to treat HIV:
nucleoside RT inhibitors (NRTI), non-nucleoside RT inhibitors (NNRTI), protease 
inhibitors (PI), fusion inhibitors (FI), entry inhibitors (EI) and integrase inhibitors (INI).

Figure 5. The process of HIV integrase as it alters HIV DNA and integrates it into the 
host cell DNA. I) HIV DNA; II) Integrase cleaved off the 3’ ends of the viral DNA; III)
Viral DNA coordinates with the host cell DNA; IV) Integrase cleaves the host cell DNA 
and integrates the viral DNA into it; V) Hybrid viral/host DNA. 31-32
Current HIV Treatments

The ‘Overview of HIV Replication’ section discusses how HIV uses host cell machinery to replicate. This makes treating HIV a complicated problem since targeting host cell processes would harm healthy cells as well. The process of discovering anti-HIV drugs has been long and complex. The first antiretroviral treatment approved by the Food and Drug Administration (FDA) for the treatment of HIV was AZT, also known as Retrovir® (Figure 6, A). AZT was originally synthesized in 1964 as a possible anti-cancer agent, unfortunately it failed to adequately treat cancer and was shelved at that point. During a drug-screening for HIV 22 years later, AZT was found to inhibit HIV replication without harming host cells, which ultimately led to its approval by the FDA on March 19, 1987. At this point, the Nucleoside Reverse Transcriptase Inhibitor (NRTI) drug class was born.  

![Figure 6. Structures of HIV drugs. A) Retrovir®, B) Invirase®, C) Viramune®, D) Fuzeon®, E) Selzentry®, and F) Isentress®.][1]

---

[1]: #112x138 to 521x347
NRTIs are analogs of endogenous 2’-deoxynucleosides and nucleotides. However, in order for them to be active, they must be phosphorylated by the host cell kinases and phosphotransferases to their corresponding triphosphate form. Once phosphorylated, NRTIs are incorporated into the viral DNA facilitated by RT where they cause chain termination of the viral DNA. This chain termination is due to the lack of the terminal 3’ hydroxyl group needed for the following nucleotide to form the 3’-5’ phosphodiester bond required for making a DNA strand (Figure 7). As a whole, NRTIs interfere with the viral RT and impede the synthesis of the viral DNA, inhibiting further viral replication.

Figure 7. Diagram of the HIV replication cycle signifying where the different classes of drugs interfere.
The next class of drugs established for HIV treatment were the Protease Inhibitors (PIs), with the FDA approval of Invirase® (Figure 6, B) on December 6, 1995. PIs competitively bind to protease and prevent it from binding to the long, continuous viral protein strands that are formed in the host cell. This inhibits any cleavage of the protein strand and ultimately inhibits the production of active viral particles (Figure 7). Next came the FDA approval of Viramune® (Figure 6, C), the first Non-Nucleoside Reverse Transcriptase Inhibitor (NNRTI), just six months after the first PI approval. NNRTIs mimic deoxynucleotides and bind to the hydrophobic pocket, which is ~10Å away from the catalytic site on RT. This induces a conformational change in RT that decreases the viral enzyme affinity for the real deoxynucleotides (Figure 7), which ultimately prevents the synthesis of viral DNA and any further viral replication, typical of NRTIs.

Shortly after the approval of the first NNRTI, Viramune®, a combination therapy was implemented for treatment of HIV. This is now known as Highly Active Antiretroviral Therapy (HAART), and is currently considered the standard of care for HIV patients. HAART is a combination of two NRTIs and a third drug from a different class, which has been determined to be the most effective drug class combination to treat HIV. The result is a marked decrease in mortality rates, which has reduced the rate of AIDS-related deaths by 86%. The increase in life expectancy is a result of targeting different steps in the replication cycle of the virus, thereby ensuring that viral replication is significantly decreased which in turn decreases the concentration of virus in HIV patients. The next course of action was the simplification of treatment regimens. The first fixed dose combination (FDC) drug, Combivir®, was approved by the FDA on
September 27, 1997. It combines two NRTIs in a tablet to be taken orally twice daily. This decreases the number of pills taken within a complete regimen and in turn increases the patient compliance to said regimen which will ultimately decrease drug resistance.\textsuperscript{33}

In 2006, Atripla\textregistered was the first FDC drug recognized as possessing a complete regimen of medications (Emtriva\textregistered, Viread\textregistered & Sustiva\textregistered) in one pill to treat HIV.\textsuperscript{33,35} At the same time that treatment options were being simplified, it was discovered that Ritonavir\textregistered boosts the efficiency of other PIs. This is because Ritonavir\textregistered is a cytochrome P450 3A4 inhibitor, which decreases the metabolism of PIs within the host cell. This conversely increases the plasma concentrations of the PIs, as well as their half-lives.\textsuperscript{40,44}

The fourth class of HIV drugs are the fusion inhibitors (FI), which were established with the creation of Fuzeon\textregistered\textsuperscript{36} (Figure 6, D), approved by the FDA in 2003. Fuzeon\textregistered interferes with the ability of HIV to enter into a host cell by binding to the viral transmembrane protein gp41, which is responsible for fusing the virus with the host cell. This is accomplished because Fuzeon\textregistered is a synthetic peptide mimic of the amino acid sequence 127-162 on HR2 of gp41. After gp41 has been inserted into the host cell membrane, the gp41 HR domains are exposed and Fuzeon\textregistered binds to HR1 of gp41, preventing both the formation of the 6-helix bundle of gp41 and the subsequent fusion of the two cells (Figure 7).\textsuperscript{35,39,45-46}

In August of 2007, the FDA approved the first Entry Inhibitor (EI), Selzentry\textsuperscript{®} (Figure 6, E) and the fifth drug class of HIV. Selzentry\textsuperscript{®} is a CCR5 co-receptor antagonist that binds to the host cell co-receptor CCR5 and blocks gp120 from binding to
CCR5 and consequently infecting the host cell.\textsuperscript{46-47} Finally, the last drug class established to treat HIV was the Integrase inhibitors (INIs), represented by Isentress\textsuperscript{®} (Figure 6, F), which was approved by the FDA in October of 2007. Isentress\textsuperscript{®} targets integrase (IN) after it has processed the 3’ end of the viral DNA. The 3’ processing induces a conformational change in the IN which allows Isentress\textsuperscript{®} to bind to the active site on IN. Once bound, Isentress\textsuperscript{®} chelates the metal ions coordinated by the catalytic triad on IN, thus blocking the covalent integration of viral DNA into host genome.\textsuperscript{39,48} This in turn eliminates the production of viral RNA and viral proteins necessary for viral replication.

In summary, there are currently 6 classes of drugs developed to treat HIV: NRTI, PI, NNRTI, FI, EI and INI. Within those drug classes reside a total of 32 drugs, 6 of which are classified as fixed dose combinations (FDCs).\textsuperscript{33} Unfortunately, despite the plethora of current HIV medications, HIV patients are slowly becoming resistant to these drugs.

**Drug Resistance**

Every treatment for HIV has a limit regarding how effective it can be. Unfortunately, this is due to the resistance HIV builds against each drug to which it is exposed. Drug resistance is defined as an increase of antiretroviral (ARV) drug concentrations needed to suppress HIV replication relative to non-resistant viruses.\textsuperscript{49} This resistance is primarily due to the high rate of HIV mutation, \( \sim 10^{-4} \) mutations per nucleotide (1 or 2 mutations per genome) per viral replication cycle. Since HIV rapidly
replicates, producing $\sim 10^9$ new virons daily, the end result is a large number of genetically related, but distinct viral variants.\textsuperscript{49-50} To compound the issue, the molecular structures of drugs within the same class are similar, and have the same target sites. This means that when HIV becomes resistant to a particular drug, it also becomes resistant to the majority of drugs within the same drug class.\textsuperscript{49}

To understand drug resistance one must explore the different types of mutations that lead to the resistance of the differing drug classes. For the NRTIs, mutations occur in the catalytic binding site of the RT, thus making it too sterically hindered for the NRTIs to bind. Also, mutations of the RT lead to the cleavage of the already bound NRTI from viral DNA, allowing the DNA elongation to occur.\textsuperscript{38} As for the PIs, the mutations to the viral protease are stepwise in fashion. Initially, mutations occur within the protease substrate-binding cleft and decrease the ability of the PIs to bind. Next, mutations occur within the catalytic site of the protease and result in an overall enlargement. This enlargement allows the long viral protein strands to gain access and be cleaved into individual viral proteins.\textsuperscript{38} The mutations affecting NNRTIs are simple: they occur within the RT hydrophobic binding pocket and prevent NNRTIs from binding to RT.\textsuperscript{41}

The remaining three drug classes (FIs, EIs & INIs), all contain only one drug per drug class. Therefore, the mutations to HIV are drug specific and may or may not represent a resistance to the entire drug class. Because of this, the HIV resistance for the remaining drug classes is discussed on a drug-by-drug basis. The first drug resistance discussed is Fuzeon®, which results from mutations occurring at codons 36-38 on HR1
of gp41. These changes create an inability of Fuzeon® to bind to the HR1 domain of gp41. The normal formation of the 6-helical bundle and subsequent fusion of HIV to the host cell is then achieved. As for the resistance to Selzentry®, two main mutations in HIV occur: 1) a change in the tropism of HIV from R5 to X4. This means that while HIV initially requires co-receptor CCR5 to bind to the host cell, once mutated, it will require the alternative co-receptor CXCR4 to bind to the host cell; 2) mutations to gp120 that allow it to bind to CCR5 even though CCR5 is bound by Selzentry®, which was designed to inhibit the binding of gp120. Finally, drug resistance to Isentress® is the result of simple mutations. The catalytic core domain of IN is mutated close enough to the binding site of Isentress® to disrupt the binding interaction of Isentress® to IN, allowing the viral DNA to continue to integrate into the host genome. In general, these mutations affect the sites where the drugs (drug classes) interact with HIV, making the drugs ineffective. At the rate at which HIV can mutate and replicate, these resistances can add up quickly and make treatment problematic. For example, a case study of 132,500 viremic patients reported in 2004 showed that 76.3% of the patients were found to be resistant to at least one class of drug, 71.4% were resistant to NRTIs, 40.5% were resistant to PIs, 25.2% were resistant to NNRTIs and 13% were resistant to all three types. As the mutations and resistance to treatment increases, the available therapeutic arsenal decreases. This in turn leads to more complex treatment options that are more expensive and generally have a greater number of side effects associated with them, many of which are detrimental or intolerable to the patient.
Drug Side Effects

In addition to drug resistance presenting an impediment for current HIV/AIDS treatment, there are also many side effects associated with these treatments. In a case study in 2003, 109 people were asked to rate a list of 38 common side effects associated with these treatments and to provide insight into which were the worst to endure. The most common symptoms were fatigue, stiff/painful joints, aching muscles, diarrhea, depression and peripheral neuropathy (pain in the arms and legs). The five worst symptoms to experience, as reported by the study, were peripheral neuropathy, lipodystrophy (an abnormal gain or loss of fat in certain areas of the body), sleep difficulties, fatigue and stiff/painful joints. It was concluded that three of the worst symptoms experienced were also the most commonly experienced.

More recently in 2009, the side effects of current HIV treatments were reassessed and correlated to drug classes. For patients on NRTIs and NNRTIs common side effects were lipodystrophy, hepatitis and pancreatitis. For particular drugs within these two drug classes, a hypersensitivity response can occur, which is a multisystem response that can lead to death. It has been shown that this particular response is strongly associated with particular HLA haplotypes (alleles on linked genes), depending on the drug in question. For patients on PIs, common side effects were hyperlipidemia (increase in LDL and triglycerides), lipodystrophy and type II diabetes. For Fuzeon®, common side effects included injection-site reactions (rashes, cysts) and a bad taste in the mouth, while for Selzentry® it was hepatitis. Lastly, for Isentress® the common side effects were dizziness, headache, fatigue, insomnia and diarrhea. The severity and frequency
of side effects, in addition to increasing resistance to current treatments, has created a need for new approaches in treating HIV/AIDS.

**Current Research for Treating HIV**

As mentioned in the previous section, HIV resistance to current medications is a significant problem. This has lead to vigorous research in drug design for more effective and tolerable HIV treatments. Currently there are 4 NRTIs, 6 NNRTIs, 1 FI, 7 EIs and 1 INI in phase II or III clinical trials. Interestingly, one of the NNRTIs under study, (+)-Calanolide A (Figure 8), is the only natural product to have entered clinical trials for HIV. (+)-Calanolide A is isolated from a tropical rain forest plant species *Calophyllum lanigerum*, and has been shown to have anti-HIV activity. Although the mechanism with which (+)-Calanolide A inhibits HIV is still not fully understood, it is suspected that (+)-Calanolide A binds to one or more sites on RT. This type of RT inhibition places (+)-Calanolide A in the drug class NNRTI.

Despite the promising results from (+)-Calanolide A, the majority of known biologically active natural compounds lack anti-HIV activity. This make (+)-Calanolide A and phloroglucinols quite unique. Phloroglucinols are found in certain plant species, such as *Eucalyptus* and *Aspidium*, just to name a few and have been found to inhibit HIV. There are current computational studies exploring 21 phloroglucinols and how they theoretically interact with the three HIV enzymes; RT, Protease and IN. These studies have shown that 3 of these 21 phloroglucinols have good binding inhibition of IN, 2 derivatives have good binding inhibition of protease and 1 derivative inhibits both.
These findings have opened the door to a whole new treatment option for HIV utilizing natural products.

Figure 8. (+)-Calanolide A.$^{37}$

Another drug target under study for HIV inhibition is the enzyme RNase H. The function of RNase H with respect to RT has been known for some time. Ever since the discovery of RNase H there has been a lot of effort directed toward finding an inhibitor for this enzyme. It is suspected that there are two magnesium ions in the active site of RNase H, which are believed to be involved in the mechanism of viral RNA phosphodiester bond hydrolysis. Therefore, current research is focusing on the synthesis of molecules that will bind to the magnesium ions and prevent the hydrolysis of viral RNA, which will in turn halt viral replication.$^{57}$ Drug synthesis addressing new HIV targets, such as RNase H, is one area of current HIV research.

Another area of HIV research is the pathogenesis of the virus. New discoveries are being made regularly about the infection process of HIV, leading to new approaches for treating HIV. A few of the recent discoveries regarding HIV include the discovery of
the mechanism for the uncoating of the HIV capsid. Also, it was discovered that integrase uses a co-factor, lens epithelial derived growth factor/75 kDa protein (LEDGF/p75), to integrate viral DNA into the host genome. Additionally, the functions of the viral proteins (Tat, Rev, Vpr, Vpu and Vif) have been determined.\textsuperscript{58-59} By understanding these new aspects of the HIV infection process, researchers can study how to manipulate and possibly inhibit them. Therefore, active research is underway in these new areas of HIV pathogenesis in hopes of discovering new drug classes.

Another avenue of HIV research, outside of drug design and pathogenesis, is gene therapy, which is being explored as a way of inhibiting HIV replication in host cells. It has been known for some time that \textasciitilde 1\% of the Caucasian population has a homozygous deletion of the 32\textsuperscript{nd} base pair on the CCR5 gene and are naturally resistant to HIV infection. The result of this deletion is the absence of CCR5 on host cells, which doesn’t appear to affect the function of these cells.\textsuperscript{60} Ever since the discovery of this CCR5 mutation, it has been a goal to extend this natural resistance to those infected with HIV. In 2007, an HIV positive patient with acute myeloid leukemia underwent stem cell transplantation with cells from an HLA-identical donor who was homozygous for the 32-base-pair deletion on the CCR5 gene. At the time of transplantation the patient terminated all forms of HIV treatment. After 27 months, the patient was in complete remission of the leukemia and had no detectable levels of HIV.\textsuperscript{61-62} Unfortunately, the odds of finding a stem cell donor who also is homozygous for the delta 32 alleles are close to negligible, making this treatment option impractical, if not impossible. However, Anderson, J. \textit{et al.}\textsuperscript{60} have successfully created a triple combination anti-HIV lentiviral
vector that addresses this concept. This lentiviral vector targets CD34+ hematopoietic progenitor cells, with the goal of transducing the modified genetic information into these cells so that they will transform into cells that exhibit the desired genetic traits. This is possible due to the fact that CD34+ hematopoietic progenitor cells differentiate into either a common lymphoid progenitor cell or a common myeloid progenitor cell. The common lymphoid progenitor cell then further differentiates into either a natural killer T cell, a helper T cell, or a B cell. The common myeloid progenitor cell can also transform into either a monocyte, neutrophil, red blood cell or a platelet. Therefore, by engineering a CD34+ hematopoietic progenitor cell to be HIV-1 resistant, all of the future cells that differentiate from it will also be resistant or 'protected' from HIV-1.

The lentiviral vector is coded for three different types of gene expression: TRIM5α human/rhesus isoform, CCR5 short hairpin RNA (shRNA) and a TAR decoy. The protein TRIM5α has been shown to inhibit the HIV capsid from uncoating in the host cell in rhesus monkeys, however, the human isoform of TRIM5α does not. Upon further research, it was discovered that the rhesus monkey isoform of TRIM5α contains mutations that prevents the uncoating of the HIV capsid. So by inserting the key amino acid sequence from the rhesus isoform into the human isoform, a new TRIM5α human/rhesus isoform was created. This new human/rhesus isoform maintains the effectiveness of the rhesus isoform against HIV and is tolerated by humans. The next genetic modification on the lentiviral vector is the shRNA, which is an interfering RNA that silences the gene expression of CCR5 on host cells. As mentioned previously, 1% of the Caucasian population has this abnormality naturally, while still exhibiting normal
immune and cellular responses. Therefore, it is expected that this genetic modification to host cells will result in little to no adverse side effects. For the TAR decoy, a genetic modification will be made to code for a protein that will bind to the viral Tat protein, inhibiting the function of the Tat protein. So far, the lentiviral vector has been successfully transduced into CD34+ hematopoietic progenitor cells, in vitro, and the safety, toxicity and efficacy of the vector are currently under study. Assuming the lentiviral vector proves to be safe and effective in vitro, it will enter clinical trials.60

As of yet, the areas of HIV research that have been discussed have involved inhibiting the different modes of action of the virus post infection. These modes of action include replication and further host cell infection, which aren’t the only areas being researched. There is an area of study that is aimed at the prevention of HIV infection, the study of topical microbicides, which has received little attention. Microbicides are molecules that attack cellular or viral targets and prevent these targets from either infecting host cells, or from replicating. To be considered a successful microbicide, the molecule under study must have several key traits: it should be biologically effective, inexpensive, colorless, odorless, tasteless, easy to use and won’t interfere with sexual pleasure.65 There are currently several types of microbicides being studied for safety and efficacy which cover a variety of classes: nonspecific surfactants/detergents, acid buffering agents, moderately specific macromolecular anionic polymers and HIV drugs. Within the macromolecular anionic polymer class, there is a new molecule (SPL7013, Figure 9) that has been developed that shows anti-HIV activity against both CCR5 and CXCR4 strains of HIV. SPL7013 is a fourth generation dendrimer made up of L-lysine
branches, with surface naphthalene disulfonic acid groups and has completed phase I clinical trials.\textsuperscript{66}

\textbf{Figure 9.} Microbicide SPL7013. Lysine residues are highlighted in red, green, purple, brown and black, and the surface naphthalene disulfonic acid residues are in blue. Each color differentiation marks a branching point and therefore a different generation. Red represents the core or generation 0, green is generation 1, purple is generation 2, brown is generation 3 and black is generation 4.\textsuperscript{65-66}
Unfortunately, until a microbicide is approved by the FDA, there aren’t any treatment options to prevent transmission of HIV to host cells. As a result of the ongoing research into the pathogenesis of the active HIV virus, it has been discovered that HIV can infect a host cell and lay dormant for an extended period of time. These host cells infected with dormant HIV viruses are known as latent reservoirs. These latent reservoirs appear to be normal healthy cells due to the lack of HIV replication and budding of new viron. Therefore, these cells won’t be targeted by any of the current treatment options or even by the host immune system. This provides HIV with a resting place for an extended period of time until some event causes reactivation of the virus. Current research is focusing on treatment methods to activate these reservoirs. Once activated, these latent cells will no longer appear healthy and can be targeted by current HIV drugs as well as the host immune system.

In this section current research for treating HIV infection and replication has been discussed. Overall, the most promising area of research is topical microbicides, which explores the prevention of HIV transmission. Until these microbicides are made available for use, people are still being exposed to and becoming infected with HIV. Though there are drugs available to inhibit HIV replication and decrease infection rates, the result is still the habitation of active HIV viron within the host cells. This habitation causes the immune system of the host to destroy these infected cells. Since host cells are destroyed either by HIV or by the host immune system, it appears that prevention of the initial HIV binding to the host cell receptors would be an ideal form of HIV inhibition. One possible avenue to address this is the use of polyanionic sulfated compounds.
Anionic Sulfated Compounds

Anionic sulfated compounds, such as heparan sulfate (HS, Figure 10), act as cell surface receptors for HIV. Since HS binds more strongly to gp120 ($k_D = 10$ nM) than gp120 binds to the CD4 receptor ($k_D = 22$ nM), it is believed that anionic sulfated compounds have the possibility of inhibiting the binding of HIV to host cells.\textsuperscript{19-20} There have been a number of studies that have already explored the binding of anionic sulfated compounds to gp120. In studies by both Mondor, \textit{et. al.}\textsuperscript{67} and Saphire, \textit{et. al.}\textsuperscript{5} it was shown that the anionic charges of heparan sulfate bind to the basic residues on the gp120 molecule and inhibit HIV from binding to the host cell.\textsuperscript{5,67} These gp120 residues have since become a new target for HIV inhibition, and therefore, other anionic sulfated compounds have also been studied to inhibit gp120 binding to host cells. In these studies, the binding affinity to the monomeric gp120 complex is usually lower than to the oligomeric gp120. This low affinity is typical of carbohydrate-protein binding interactions and presents a complicated problem. A problem that can be resolved by multivalency.
Multivalency

Multivalency is nature’s answer to poor binding affinity. There are many events in nature that utilize carbohydrate-protein interactions, such as infection by most pathogens. These pathogens have protein receptors that are used to bind to surface carbohydrates on host cells.\(^\text{10}\) The affinity of carbohydrate-protein binding normally is millimolar at best, however, nature has overcome this by clustering carbohydrate ligands together to yield a ‘multivalent’ effect. The multivalent effect is the simultaneous attachment of two or more binding sites within one entity, to multiple receptor sites on another (Figure 11).\(^{69}\) This multiplicity of binding increases the strength of the binding interaction exponentially.\(^{70}\) An example of this principle would be illustrated by noting that the monovalent binding coefficient of a tetra-amine ligand to DNA is in the micromolar concentration range, while the multivalent binding coefficient of the same
tetra-amine ligand is in the nanomolar concentration range.71 Further examples of multivalent binding strengths were shown by the studies of Kensinger, et. al.10 This work explores the use of different sulfated anions in a multivalent nature to inhibit gp120. The most effective inhibitor from this study was galatosyl ceramide sulfatide (SGalCer, Figure 12), bound to a poly(propylene imine) dendrimer (Figure 13). SGalCer is a compound that contains a sugar residue with an anionic sulfate. Like heparan sulfate, SGalCer has a stronger binding affinity for gp120 (K_D=1.9nM, multivalent, 64-mer) than gp120 binding to the receptor CD4 (K_D=22nM, multivalent).12,72 This results in the preferred binding of SGalCer to gp120 despite the presence of CD4 on the host cell, thereby inhibiting HIV from binding and infecting host cells. More recently, binding studies of heparin sulfate (Figure 10), a more sulfated form of HS, have shown that it has an even stronger binding affinity to gp120 (K_D=0.6nM), when compared to SGalCer.73 Though the binding strength of heparin sulfate would lead to a promising drug target against gp120, it is an anticoagulant and the concentration needed to effectively inhibit gp120 would result in the patient bleeding to death. Despite these unfortunate results, the binding affinity of heparin sulfate supports the belief of polyanion sulfated sugars as potential gp120 inhibitors. Since gp120 binds to host cells in a multivalent fashion, it is believed that future drug targets should compensate for this. To further investigate multivalency and its multiple applications in drug design, 'chemical scaffolds' such as dendrimers are being investigated.
Figure 11. Multivalency. An example of the multivalent effect where the simultaneous attachment of two or more binding sites within one entity, bind to multiple receptor sites on another.\textsuperscript{74}

Figure 12. Structure of galactosyl ceramide sulfatide (SGalCer).\textsuperscript{72}
Figure 13. Dendrimer generations. Multiple generations of a poly(propylene imine) dendrimer, where the terminal group SGalCer would be placed at the end of generation 5.10

**Dendrimers**

Dendrimers are a remarkable class of polymers, distinguished by their three-dimensional globular shape and extensive branching.75 Dendrimers are either constructed by divergent synthesis, involving building the molecule outward from the core, or by convergent synthesis, where blocks of the molecule are first constructed, then attached to
the core. Either method of synthesis creates branching points known as generations, where the number of branches increase by a factor of 2 with each increasing generation, resulting in the globular macromolecule (Figure 13).^75

A variety of surface functional groups may be placed at the terminal ends of the dendrimer branches, allowing dendrimers to have a wide range of physical properties and desirable traits. This is apparent in the scenario where the terminal groups are poly anionic sulfated sugars, dendrimers would allow gp120 to have access to these sugars while keeping the sugars relatively small. This would diminish any possibility of these sugars becoming anti-coagulants, such as heparin sulfate. Dendrimers are therefore ideal for the development of a variety of different HIV inhibitors due to their compact globular shape, controllable size, and variable functionality.^72,75

**Proposal**

The discussion presented (*vida supra*) has shown that there is a great need for new approaches in the treatment of HIV. Of particular interest is the approach which seeks to block binding of host cell receptors to gp120 competitively. By inhibiting the binding of gp120 to host cells, not only will the infection of HIV be prevented, but the host cell will remain unaffected by the virus. Because of this, it is believed that entry inhibitors would be ideal for treating HIV. As gp120 has a high affinity for anionic sulfated sugar compounds, it is believed that sulfated sugars attached to a dendrimer scaffold could prevent or decrease the binding of gp120 to host cells. Therefore, our research group is
exploring the effects of a variety of sulfated sugars as possible gp120 inhibitors. The sugar assigned to this body of work was glucosamine, due to it’s extensive use in the human body. Unfortunately, sugars are readily metabolized, which would negate any possibility for administering these drug targets orally. This leaves only two other options, either these drug targets could be used as a microbicide in a topical gel, or be administered intravenously. Therefore, these drug targets and the dendrimer scaffolds used to make them, will need to be water soluble. In order to achieve this, tertiary amines such as triethanolamine, which are both polar (water soluble) and readily available, will be used to build the dendrimers in this study. In summary, it is hypothesized that polar tertiary amine dendrimers terminated with glucosamine, once sulfated, could inhibit the binding of HIV envelope glycoprotein, gp120 to host cells.
Chapter 2

RESULTS AND DISCUSSION

The goal of the work presented in this thesis was to synthesize polar tertiary amine dendrimers terminated with glucosamine. These glycodendrimers, once sulfated, can then be used to study the possible inhibition of binding of the HIV envelope glycoprotein, gp120, to host cells. Studies have shown that gp120 uses sugar moieties on sphingolipids and sulfated, polyanionic proteins as host cell receptors.\textsuperscript{5,16-18} Further studies into this binding event have shown that sulfated, polyanionic sugars bind with a stronger affinity to gp120 than the host cell receptors themselves.\textsuperscript{5,10,12,19-20,67,72} It was also shown that the infection process for HIV requires multiple copies of gp120 to bind to multiple copies of receptors simultaneously.\textsuperscript{4-6,8-11} In order to mimic this process, dendrimers were determined to be an ideal scaffold for the sulfated polyanionic sugars. This determination was made based on the fact that dendrimers are hyper-branched, three dimensional, globular macromolecule scaffolds that can be manipulated to give any desired size and/or physical properties.\textsuperscript{75} Therefore, based on these previous studies and current investigations, this work will show the successful synthesis of two cores [TEA\textsubscript{m}-(G1)\textsubscript{3}, TEA\textsubscript{m}-G1-(DAP)\textsubscript{3}], one sugar-linker (G-AE-PA) and two glycodendrimers [TEA\textsubscript{m}-G1-(GlcN)\textsubscript{3}, TEA-G1-(G-AE-PA)\textsubscript{3}], with the aim of inhibiting gp120 from binding to host cells.
Compounds Synthesized

$\text{TEA}_m$-$\text{G1}_3$ (2)

Compound 1 ($\text{t}$-butyl protected $\text{TEA}_m$-$\text{G1}_3$), was synthesized using a Michael-like addition of $\text{t}$-butyl acrylate to triethanolamine (Scheme 1). Thin layer chromatography (TLC) was unable to elucidate any information regarding this reaction, so the initial reaction time was determined from the work done by Seitz & Kunz.\textsuperscript{76}

![Diagram](image)

**Scheme 1.** Synthesis of the dendrimer core. Synthesis of Compound 1 using a Michael-like addition and subsequent deprotection with TFA to produce Compound 2.

Since the reaction time for the first reaction shown in Scheme 1 is based primarily off of the work done by Seitz and Kunz, it is advantageous to look at their Michael-like
addition reaction to understand how it was adapted for this purpose. Their reaction conditions included stirring for 20 hours, to add methyl acrylate in a 1,4-fashion to one end of triethyleneglycol, hence 1 reaction site. Since the goal of the Michael-like addition in this body of work was to add t-butyl acrylate to each of the 3 ends of triethanolamine (3 reaction sites), the reaction time of Seitz & Kunz was tripled (to ~3 days). This reaction was subjected to time trials ranging from 24 hours to 4 days, in an attempt to optimize the reaction. These studies showed that the maximum yield (36.3%) was attained after the reaction ran for 3 days. At this point, it was concluded that ~1 day per reaction site was the optimal reaction condition for this step.

Once the reaction was complete, it was then neutralized with hydrochloric acid and purified by extraction. The $^1$H NMR of the pure product, Compound 1 (Figure 14), shows the successful 1,4-addition of t-butyl acrylate onto triethanolamine. The singlet at 1.39 ppm, indicates an alkane, which is typical for a t-butyl group. As a result, that peak was integrated for 27 protons, in correlation to the $t$-butyl group on Compound 1. The resultant integrations on the remaining 4 triplets equate to 6 protons each, which in turn correspond to the remaining ($-\text{CH}_2$-) groups on Compound 1. The triplet at 2.42 ppm, indicates a ($-\text{CH}_2$-) group adjacent to a carbonyl and split by 2 protons, labeled ‘4’ on the structure in (Figure 14). The coupling constant of triplet ‘4’ ($J = 6.5$ Hz), is identical to the coupling constant of the triplet at 3.60 ppm. This triplet indicates a ($-\text{CH}_2$-) group adjacent to an oxygen, also split by 2 protons. Thus, this triplet represents the ($-\text{CH}_2$-) group ‘3’. The triplet at 2.70 ppm, indicates a ($-\text{CH}_2$-) group adjacent to an amine, split by 2 protons, labeled ‘1’. The coupling constant of triplet ‘1’ ($J = 6.0$ Hz), matches the
coupling constant of the final triplet at 3.45 ppm. This triplet also indicates a (-CH\textsubscript{2}-) group adjacent to an oxygen, split by 2 protons, representing the (-CH\textsubscript{2}-) group ‘2’. The lack of additional peaks and the fact that the spectrum integrated for the exact number of (-CH\textsubscript{2}-) protons in the structure, imply purity of Compound 1. This was further confirmed via \textsuperscript{13}C NMR (Appendix B).

\textbf{Figure 14.} Compound 1, \textsuperscript{1}H NMR, 300 MHz, CDCl\textsubscript{3}. An expanded version is located in Appendix A.

Compound 1 was next subjected to a trifluoroacetic acid (TFA) deprotection, using anisole as a cation scavenger (Scheme 1). All side products were volatile and evaporated during concentration \textit{in vacuo}, resulting in a 99.6\% yield of Compound 2. The \textsuperscript{1}H NMR of Compound 2 (Figure 15), shows the successful deprotection of Compound 1 and
purification of Compound 2. The absence of the singlet at 1.39 ppm proves the successful removal of the \( t \)-butyl group. The 4 triplets that correspond to \((-\text{CH}_2-)\) groups ‘1-4’ are still present and the lack of any additional peaks suggest purity, which was later confirmed via \(^{13}\text{C} \text{NMR} \) (Appendix B) and fast atom bombardment mass spectroscopy (FAB MS).

![Figure 15](image)

**Figure 15.** Compound 2, \(^1\text{H} \text{NMR}, 300 \text{ MHz, D}_2\text{O.} \) An expanded version is located in Appendix A.

**TEA\textsubscript{m-G1(GlcN)}\textsubscript{3} (3)**

Compound 3 was synthesized by coupling glucosamine to Compound 2 via benzotriazol-1-yloxy-tris(dimethylamino)phosponium hexafluorophosphate (BOP) mediated amide-coupling reaction (Scheme 2). After the reaction was complete, as
determined by TLC, it was subjected to flash chromatography using 6:4:1 chloroform:methanol:water. TLC showed that the starting material had eluted after ~750 mL had been collected, however the product remained on the column. In order to elute the product, the solvent system was transitioned to methanol and an additional ~750 mL were collected. TLC showed product had just begun to elute, so in attempt to expedite the process, the column was then transitioned to “wet” methanol (~20% water, ~80% methanol) and an additional ~250 mL of solution were needed to complete the elution of the product. Unfortunately, $^1$H NMR showed diisopropylethylamine ((iPr)$_2$EtN) had co-eluted with Compound 3. The sample was therefore subjected to dialysis, using 500 molecular weight cut off (MWCO) tubing for 3 hours, in another attempt to purify Compound 3. Disappointingly, $^1$H NMR showed (iPr)$_2$EtN was still present, although diminished. Despite an additional attempt via dialysis to purify Compound 3 further, the amount of (iPr)$_2$EtN remained constant. This was evident due to the conservation of mass and was confirmed by $^1$H NMR. Finally, the sample was subjected to reverse phase high performance liquid chromatography (RP-HPLC) and Compound 3 was successfully purified with a 32.8% yield.

The $^1$H NMR of pure product, Compound 3 (Figure 16), shows the successful coupling of glucosamine onto Compound 2, and the subsequent removal of the (iPr)$_2$EtN impurity of Compound 3. The multiplet spanning from 2.64-2.65 ppm indicates a (-CH$_2$-) group adjacent to a carbonyl. The peak was integrated for 6 protons, and was assigned as the (-CH$_2$-) group ‘2’ on Compound 3 (Figure 16). The multiplets spanning 3.2-3.59 ppm and 3.72-3.95 ppm indicate one or more (-CH-) groups adjacent to an oxygen or a nitrogen. These multiplets have a combined integration of 36 protons, which represent the group of protons ‘1’ and ‘3’ on Compound 3. The doublets at 4.76 and 5.23 ppm are characteristic of anomeric protons.
Figure 16. Compound 3, $^1$H NMR, 500 MHz, $D_2$O. An expanded version is located in Appendix A.

It should be noted that the anomeric proton on an $\alpha$-pyranoside, such as glucosamine (GlcN), is always more deshielded than the anomeric proton on a $\beta$-pyranoside. This is because the $\alpha$-proton is in the equatorial position (Figure 17), which is closer to the ring oxygen than the axial position of the $\beta$-proton. This close proximity allows the ring oxygen to withdraw more electron density from the $\alpha$-proton than the $\beta$-proton, resulting in the $\alpha$-proton being more deshielded. Also, it should be mentioned that there are 3 GlcN and therefore 3 anomeric protons on Compound 3. The integration of $\alpha$ and $\beta$ protons is $\sim$1.5, which indicates an equal mixture of $\alpha$ and $\beta$-GlcN in the
sample. It is most likely a mixture of 4 possible dendrimers, Compound 2 with: 3 $\alpha$-GlcN; 3 $\beta$-GlcN; 2 $\alpha$-GlcN and 1 $\beta$-GlcN or 2 $\beta$-GlcN and 1 $\alpha$-GlcN. The result is the same amount of $\alpha$ vs. $\beta$-GlcN. Therefore, if the glucosamine is split 50:50 in the sample, it makes sense that this ratio would be represented in the $^1$H NMR, by dividing the anomeric proton integration (3 protons), in half (1.5 protons per isomer). Since there weren’t any auxiliary peaks and the spectrum integrated for the exact number of (-CH-) protons in the structure, Compound 3 is assumed to be pure. The purity was further confirmed via $^{13}$C NMR (Appendix B) and electrospray ionization-fourier transform mass spectroscopy (ESI-FT MS).

**Figure 17.** Glucosamine isomers, where the equatorial proton is in red and the axial proton is in blue.
**G-AE-PA (6)**

Compound 6 was synthesized using BOP-mediated amide coupling of glucosamine to Compound 4, which was synthesized previously by Doug Glick, a former group member. The newly synthesized Compound 5 was then subjected to a TFA deprotection, resulting in Compound 6 (*Scheme 3*). Purification of Compound 5 was attempted prior to the deprotection step. This proved to be difficult and ultimately was unsuccessful. First, the crude mixture containing Compound 5 was subjected to flash chromatography using 8:1 chloroform: methanol. Collection of ~1.7 L resulted in what appeared to be the complete elution of starting materials from the column, as determined by TLC. Next, the column was transitioned to 6:4:1 chloroform:methanol:water to elute Compound 5. Collection of ~500 mL appeared to elute Compound 5 from the column, again determined by TLC. Despite multiple purification attempts, $^1$H NMR showed (iPr)$_2$EtN had co-eluted off with Compound 5. The crude mixture of Compound 5 was subjected to a TFA deprotection to produce Compound 6 which was purified from (iPr)$_2$EtN via RP-HPLC, resulting in an overall 14.2% yield for the two-step process.
**Scheme 3.** Synthesis of the sugar linker. Synthesis of Compound 6 via BOP-mediated amide coupling followed by a TFA deprotection.

The $^1$H NMR of pure product, Compound 6 (*Figure 18*), shows the successful coupling of glucosamine onto Compound 4. The quartet at 2.60 ppm indicates a (-CH$_2$-) group adjacent to a carbonyl. This the peak was integrated for 2 protons, which correspond to the (-CH$_2$-) group ‘3’. The integrations on the remaining peaks equate to 13 protons, which in turn correspond to the remaining (-CH-) groups on Compound 6. The quartet at 2.6 ppm is due to the diastereotopic protons on carbon 3, which result from the nearby chiral center on the glucosamine. Since the protons on carbon 3 are non-equivalent, the splitting from the two neighboring protons on the adjacent carbon, should result in two sets triplet peaks. However, since a quartet is the result of these splitting events, it is believed that the triplets merge together to create the quartet (*Figure 19*). The triplet at 3.19 ppm indicates a (-CH$_2$-) group adjacent to a nitrogen and is labeled ‘1’. The multiplet spanning from 3.44-3.55 ppm indicates a (-CH-) group adjacent to oxygen. Since this multiplet only integrated for 2 protons, it must correspond to the (-CH-) group.
‘4’. The multiplet spanning from 3.70-3.94 ppm also indicates (-CH-) groups adjacent to an oxygen and corresponds to the (-CH-) groups ‘1’ and ‘4’. As discussed previously, doublets between 4-5 ppm are characteristic of anomeric protons and the α proton is always more deshielded than the β proton. Based on this information it can be determined that the doublets 4.70 and 5.20 ppm represent the 2 isoforms of GlcN (α & β), where α-GlcN is the 5.20 ppm doublet and the β-GlcN is the 4.70 ppm doublet. The absence of additional peaks and the fact that the spectrum integrated for the exact number of (-CH-) protons in the structure, suggest the purity of Compound 6. This was further confirmed via 13C NMR (Appendix B). Mass spectroscopy was not performed on this compound. Since Compound 6 was used in the synthesis of Compound 7, which was confirmed via mass spectrometry, it can be concluded indirectly that Compound 6 was successfully synthesized.
Figure 18. Compound 6, $^1$H NMR, 300 MHz, D$_2$O. An extended version is located in Appendix A.

Figure 19. Stick diagram of the diastereomeric proton splitting on Compound 6. This shows the two triplet peaks merging to form the quartet seen at 2.60 ppm in Figure 18.
It should be mentioned that (iPr)$_2$EtN appears to associate strongly with Compounds 3 & 5. It is hypothesized that (iPr)$_2$EtN is hydrogen bonding to the amide on Compounds 3 & 5. This hypothesis is supported by the fact that once (iPr)$_2$EtN is protonated it can then be separated from either product. Consequently, the BOP-mediated amide coupling reaction, used to synthesize Compound 5, was repeated using triethylamine as an alternative base to optimize the reaction. Unfortunately, the glucosamine never dissolved, and the amide coupling failed to occur. Therefore, it is believed that the strength of the base used for BOP-mediated amide coupling drives this reaction. The base must deprotonate the glucosamine, creating a charged species that then allows dimethylformamide (DMF) to solvate glucosamine. Once the glucosamine is in solution, it can readily undergo amide coupling. Though the pKa values of (iPr)$_2$EtN (pKa = 11) and triethylamine (pKa=10.78) are close to one another, perhaps glucosamine requires a base at or above pKa = 11 to deprotonate the amine. This would be an interesting study for the future.

**TEA$_m$-G1-(G-AE-PA)$_3$ (7)**

Compound 7 was synthesized by performing a BOP-mediated amide coupling of Compound 2 to Compound 6 (Scheme 4). It should be mentioned that the base used in this reaction was triethylamine (TEA) instead of (iPr)$_2$EtN. This was done in an attempt to simplify purification, which was successful. After the reaction was complete, as determined by TLC, the reaction mixture was subjected to dialysis using 500 MWCO tubing. This resulted in the removal of the majority of starting material. To finish the
purification of Compound 7, the dialyzed reaction mixture was subjected to Fast Pace Liquid Chromatography (FPLC), yielding pure Compound 7 in 61.4%.


The $^1$H NMR of pure product, Compound 7 (Figure 20), shows the successful coupling of Compound 2 onto Compound 6. The multiplet spanning 2.55-2.64 ppm indicates two (-CH$_2$-) groups adjacent to a carbonyl, which integrated for 12 protons, and corresponds to ‘3’ & ‘6’. The remaining integrations equate to 66 protons, which in turn corresponds to the remaining (-CH-) groups on Compound 7, except for the anomeric β-protons, which are obscured by the broadening of the HOD peak. The multiplet spanning 2.81-2.83 ppm indicates a (-CH$_2$-) group adjacent to a nitrogen. This multiplet corresponds to the (-CH$_2$-) group ‘1’. The multiplet spanning from 3.38-3.43 ppm, also indicates a (-CH$_2$-) group adjacent to a nitrogen. Since this multiplet is more downfield than the other amine multiplet, it should correspond to (-CH$_2$-) group ‘4’. The ‘4’ (-CH$_2$-) group has more electronegative atoms within a 3 atom radius than the other (-CH$_2$-)
group adjacent to a nitrogen ‘1’, thus making ‘4’ more deshielded. This leaves the final set of multiplets spanning 3.49-3.96 ppm to correspond to the (-CH-) and (-CH\_2-) groups of ‘2’, ‘5’ and ‘7’. As mentioned previously, doublets between 4-5 ppm are characteristic of anomeric protons and the α proton is always more deshielded than the β proton. Using this information, it was determined that the doublet at 5.23 ppm represents the α-proton. Mentioned before, the β-protons are obscured by the HOD peak and it can be assumed that since there is only 1 α-proton, from integration, the remaining 2 anomeric protons are β-protons. This would indicate that more of the β isoform of Compound 6 was coupled to Compound 2 than the α isoform. There could be a correlation between this result and the fact that TEA was used instead of (iPr)\_2EtN. However, more studies would have to be performed to determine a true correlation. The lack of additional peaks implies that Compound 7 is pure. This was also confirmed via \(^{13}\)C NMR (Appendix B) and matrix assisted laser desorption ionization-time of flight mass spectroscopy (MALDI-TOF MS).
Figure 20. Compound 7, $^1$H NMR, 500 MHz, D$_2$O. An extended version is located in Appendix A.

Since both glycodendrimers have been discussed at this point, a determination of which generation each one corresponds to, is required. This is because generations are an integral part of dendrimer nomenclature. In order for a dendrimer to be a whole number generation there must branching within the molecule. At each branch point, a new generation is formed. For that reason, it should be noted that both of the dendrimers synthesized are less than a complete generation because they both are lacking any branching. As for the assignment of generation numbers, the size and structure of each of the dendrimers has to be considered. Assessing the structure of the dendrimers it was observed that they both share the same core (TEA$_m$-(G1)$_3$) and
terminal group (GlcN), which suggests that they are within the same family of dendrimer. This indicates that the generations of these dendrimers are structurally related to one another. Next, the size of these dendrimers needs to be considered, TEAₘ-G1-(GlcN)₃ is the smallest glycodendrimer possible within this family. This results in the lowest designation of dendrimer possible, generation 0. As for the glycodendrimer TEAₘ-G1-(G-AE-PA)₃, it is larger than TEAₘ-G1-(GlcN)₃, however it still isn’t branched, so it is given a designation in between generation 0 and 1, resulting in the 0.5 generation designation. Naturally, the next step would be the synthesis of a generation 1 dendrimer within this family. In order to accomplish this, the core TEAₘ-(G1)₃ will need to be branched. Work in this direction was performed, with the synthesis of the second core TEAₘ-G1-(DAP)₃, Compound 8.

**TEAₘ-G1-(DAP)₃ (8)**

Compound 8 was synthesized using a two step, one pot procedure. First, Compound 2 was converted into the corresponding methyl ester via Fischer esterification. Then, the newly formed methyl ester underwent an aminolysis, resulting in the conversion of the ester to an amide (Scheme 5). Purification of Compound 8 was conducted first with dialysis using 100 MWCO tubing for 3 hours to remove any salts from this reaction. Next, RP-HPLC was used to finish the purification of Compound 8, resulting in a 29.3% yield.
Scheme 5. Synthesis of the amine core. Synthesis of Compound 8, using a one pot, two step procedure combining methyl ester formation and aminolysis.

The $^1$H NMR of the pure product, Compound 8 (Figure 21), displays the successful addition of 1,3-diaminopropane onto Compound 2. The multiplet spanning 1.86-1.96 ppm represents an alkane group that is the most shielded group on Compound 8. Since there is only one (-CH$_2$-) group on this molecule that isn’t adjacent an electron withdrawing group, ‘6’ clearly corresponds to this multiplet. However, this multiplet appears to be a quintet, yet it lacks a consistent coupling constant which is typical of a true quintet. What is most likely occurring is the overlaying of two triplets with differing coupling constants, which results in the quintet (Figure 22), which has two distinct coupling constants.

The triplet at 2.58 ppm (Figure 21), indicates a (-CH$_2$-) group adjacent to a carbonyl, with a coupling constant of $J = 5.9$ Hz. This triplet must correspond to the (-CH$_2$-) group ‘4’ since there is only one (-CH$_2$-) group on this molecule that is near a carbonyl. Consequently, that peak was integrated for 6 protons, and the remaining peaks also integrate for 6 protons each, which is what would be expected for Compound 8. The following three triplets at 3.04, 3.33 and 3.51 ppm, are characteristic of (-CH$_2$-) groups adjacent to a nitrogen. The (-CH$_2$-) group adjacent to a nitrogen that has the fewest
electronegative withdrawing groups within 3 atoms, is ‘7’, which results in the triplet located at 3.04 ppm, with a coupling constant of $J = 7.6$ Hz. The (-CH$_2$-) group ‘adjacent a nitrogen’, surrounded by the most electronegative atoms is ‘1’, which results in the most deshielded, triplet located at 3.51 ppm with a coupling constant of $J = 4.8$ Hz. The final triplet in the ‘adjacent a nitrogen’ range at 3.33 ppm represents the (-CH$_2$-) group ‘5’ which has a coupling constant of $J = 6.8$ Hz. The 2 final triplets at 3.82 and 3.85 ppm, are indicative of (-CH$_2$-) groups adjacent to an oxygen. The triplet at 3.82 ppm has a coupling constant of $J = 5.9$ Hz, which suggests that this (-CH$_2$-) group is being split by the (-CH$_2$-) group ‘4’. This in turn means that this triplet represents the (-CH$_2$-) group ‘3’. The final triplet at 3.85 ppm has a coupling constant of $J = 4.8$ Hz, which implies that this (-CH$_2$-) group is being split by the (-CH$_2$-) group ‘1’. This then means that this triplet represents the (-CH$_2$-) group ‘2’. Since there aren’t any additional peaks and the spectrum integrated for the exact number of (-CH$_2$-) protons in the structure, this indicates that Compound 8 is pure. $^{13}$C NMR (Appendix B) and electrospray ionization-fourier transform mass spectroscopy (ESI-FT MS) also confirmed purity of Compound 8.
Figure 21. Compound 8, $^1$H NMR, 300 MHz, D$_2$O. An extended version is located in Appendix A.

Figure 22. Analysis of the splitting pattern of the multiplet in Compound 8.
Future Work

The goal of future work is to functionalize the terminal amines on TEA<sub>m</sub>-G1-(DAP)<sub>3</sub>, which would result in the doubling of terminal groups (branching). Ideally, this would be done in a manner that results in terminal carboxylic acids (Schemes 6 & 7). This is desirable because the new branched core would be primed for amide coupling of either GlcN or G-AE-PA, which would give generations 1 and 1.5 dendrimers respectively (Schemes 8 & 9).

Scheme 6. Branching of the amine core. Future functionalization of TEA<sub>m</sub>-G1-(DAP) (8).

Scheme 8. Synthesis of generation 1 glycodendrimer. Future amide coupling of Glucosamine onto the branched core, resulting in a Generation 1 glycodendrimer.
Scheme 9. Synthesis of generation 1.5 glycodendrimer. Future amide coupling of Compound 6 onto the branched core, resulting in a Generation 1.5 glycodendrimer. 81
In summary, this body of work has shown the successful synthesis of two “core” molecules [TEA\textsubscript{m}-(G1)\textsubscript{3} and TEA\textsubscript{m}-G1-(DAP)\textsubscript{3}] , a “linker” molecule (G-AE-PA), and two glycodendrimers, generations 0 and 0.5 [TEA\textsubscript{m}-G1(GlcN)\textsubscript{3} and TEA\textsubscript{m}-G1-(G-AE-PA)\textsubscript{3}]. As hypothesized (\textit{vida supra}), once these dendrimers are sulfated they could inhibit gp120 from binding to host cells \textit{in vitro}. For that reason, future work will focus on the sulfation of these two dendrimers and the subsequent inhibition studies of gp120.
This Thesis examines the synthesis of polar, tertiary amine dendrimers terminated with glucosamine. These glycodendrimers, once sulfated, could inhibit the binding of HIV envelope glycoprotein, gp120 to host cells. Previous studies have shown that sulfated, polyanionic sugars bind to gp120 on HIV more strongly than to host cell receptors. It has also been shown that gp120 binds to host cells in a multivalent fashion and that dendrimers are an ideal scaffold for the exploration of this interaction. In this study, two glycodendrimers were synthesized convergently from simple starting materials, primarily using benzotriazol-1-yl-oxy-tris(dimethylamino)phosphonium hexafluorophosphat (BOP) amide coupling chemistry. A common base used for BOP amide coupling is diisopropylethylamine (iPr)$_2$EtN, which appeared to associate via hydrogen bonding with the compounds synthesized. One way around this would be to substitute triethylamine (TEA) for (iPr)$_2$EtN. These studies showed that TEA worked well for deprotonating G-AE-PA, but failed to deprotonate glucosamine. Therefore, TEA also has some drawbacks, so if (iPr)$_2$EtN must be used then an alternative resolution is to protonate the (iPr)$_2$EtN then subject it to RP-HPLC for purification.

Future work should explore other non-nucleophilic bases useful for BOP amide coupling chemistry to address the drawbacks of TEA and (iPr)$_2$EtN. Also, since the only glycodendrimers synthesized in this study were both less than a complete generation (0 & 0.5), future work should also focus on synthesizing larger generations within this family of dendrimers so that they too can be sulfated and studied for gp120 inhibition.
Chapter 4

EXPERIMENTAL

General Information

**Abbreviations.** Benzotriazol-1-yl oxy-tris(dimethylamino)phosphonium hexafluorophosphate, (BOP); Trifluoroacetic acid, (TFA); Glucosamine, (GlcN); Molecular weight cut off, (MWCO); Reverse phase high performance liquid chromatography, (RP-HPLC); Fast pace liquid chromatography, (FPLC).

**Characterization.** NMR spectra were recorded on Bruker Avance 300 or Avance III 500 MHz instruments. FAB$^+$ MS was performed at the University of Arizona Mass Spectrometry Facility. ESI FTMS$^+$ was performed at the Ohio State University Mass Spectrometry and Proteomics Facility. MALDI-TOF MS was performed at the University of the Pacific Mass Spectrometry Facility. RP-HPLC was performed on a Hewlett Packard Series 1050 system, using an Alltech Prevail C18 5µ column, 250 mm in length and with a 10 mm inner diameter, unless otherwise noted.

**Materials.** Triethanolamine was purchased from JT Baker. $t$-Butyl acrylate and 1,3-diaminopropane were purchased from Acros. Glucosamine was purchased from Sigma-Aldrich. BOP was purchased from Nova Biochem. All other reagents and solvents were of analytical grade, were obtained from local suppliers and were used as obtained. All
dialysis tubing used was Spectra/Por (R) Biotech Cellulose Ester membrane and was purchased from Spectrum.

**Nomenclature.** A triethanolamine core (TEAₘ) with one generation of propionic acid (G1) and 1,3-diaminopropane (DAP) is abbreviated TEAₘ-G1-(DAP)₃. The suffix indicates the quantity of terminal groups. An alternative terminal group is Glucosamine-3-(2-amino-ethoxy)propionaldehyde (G-AE-PA).

**Spectra.** ¹H NMR spectra for each of the synthesized compounds are located in Appendix A and ¹³C NMR spectra are located in Appendix B.

**Compounds Synthesized**

**Preparation of t-butyl protected TEAₘ-(G1)₃ (1).**⁷⁶,⁸² Triethanolamine (1.08 g, 7.21 mmol), sodium hydroxide (1.11 g, 27.8 mmol) in water (2.8 mL), and t-butyl acrylate (3.58 g, 27.9 mmol) were combined in a round-bottomed flask (50 mL) and sealed with a septum. After 3 days of stirring at room temperature, the reaction was neutralized using 6M hydrochloric acid. The reaction was diluted with ~50 mL of chloroform and washed with ~50 mL of water. The organic layer was then dried over magnesium sulfate, filtered, concentrated *in vacuo* and lyophilized to afford the product as a yellow oil (1.412 g, 36.3%): ¹H NMR (300 MHz, CDCl₃) δ 1.39 (s, 27H), 2.42 (t, 6H, J = 6.5 Hz), 2.70 (t,
6H, J = 6.0 Hz), 3.45 (t, 6H, \( J = 6.0 \) Hz), 3.60 (t, 6H, \( J = 6.5 \) Hz), \(^{13}\)C NMR (300 MHz, CDCl\(_3\)) \( \delta \) 27.81, 36.02, 54.20, 66.40, 69.23, 80.04, 170.55.

**TEA\(_m\)-(G1)\(_3\) (2).** \(^{78-80}\) Compound 1 (0.110 g, 0.0200 mmol), trifluoroacetic acid (0.084 g, 0.740 mmol), and a catalytic amount (~1 drop) of anisole dissolved in dichloromethane (5 mL) were combined in a round-bottomed flask (50 mL) and sealed with a septum. After 1 hour of stirring at room temperature, the reaction was concentrated in vacuo, diluted with ~50 mL of chloroform and washed with ~50 mL of water. Lyophilization of the water layer gave a dark yellow oil (0.071 g, 99.6%): \(^1\)H NMR (300 MHz, D\(_2\)O) \( \delta \) 2.16 (t, 6H, \( J = 5.8 \) Hz), 3.01 (t, 6H, \( J = 4.2 \) Hz), 3.27 (t, 6H, \( J = 5.8 \) Hz), 3.33 (t, 6H, \( J = 4.2 \) Hz); \(^{13}\)C NMR (300 MHz, D\(_2\)O, internal CH\(_3\)OH standard) \( \delta \) 33.40, 52.25, 62.95, 65.37, 175.34. FAB\(^+\) MS. [M + H]\(^+\) (C\(_{15}\)H\(_{28}\)NO\(_9\)) Calcd: \( m/\varepsilon = 366.38860 \), found: \( m/\varepsilon = 366.17644 \).

**TEA\(_m\)-G1-(GlcN)\(_3\) (3).** \(^{81}\) Compound 2 (0.120 g, 0.340 mmol), diisopropylethylamine (0.430 g, 3.33 mmol) and BOP (0.550 g, 1.23 mmol) in DMF (3.25 mL) were combined in a round-bottomed flask (25 mL), sealed with a septum and placed under nitrogen. After 30 minutes of stirring at room temperature, glucosamine (0.210 g, 0.990 mmol) slurried in 3 mL DMF was added. After 6 days of stirring under nitrogen at room temperature, the reaction was concentrated in vacuo, and purified by flash column chromatography (6:4:1 chloroform:methanol:water, to ~2:8 water:methanol). The fractions were condensed in vacuo and dialyzed using 500 MWCO dialysis tubing in
water for 3 hours. The dialyzed product was lyophilized, then subjected to RP-HPLC for purification, using a linear gradient. All eluents contained 0.1% TFA (v/v). The starting concentrations of eluents were 90% water and 10% acetonitrile. After 40 minutes, the ending concentrations were 30% water and 70% acetonitrile. Lyophilization of the peak at 7.1 min. gave an off-white solid (0.093 g, 32.8%): $^1$H NMR (500 MHz, D$_2$O) $\delta$ 2.64-2.65 (m, 6H), 3.32-3.59 (m, 10H), 3.72-3.95 (m, 26H), 4.76 (d, 1.5H$_\beta$, $J = 7.0$ Hz), 5.23 (d, 1.5H$_\alpha$, $J = 3.5$ Hz); $^{13}$C NMR (500 MHz, D$_2$O, internal CH$_3$OH standard) $\delta$ 35.94, 36.26, 46.86, 53.37, 54.27, 56.89, 60.80, 60.94, 64.72, 65.31, 66.91, 70.16, 70.37, 70.81, 71.72, 74.01, 76.12, 91.05, 95.14, 174.39, 174.57. ESI FTMS$^+$ [M+Na]$^+$

(C$_{33}$H$_{60}$N$_4$O$_{21}$Na) Calcd: $m/z = 871.3642$, found: $m/z = 871.3667$.

$t$-Butyl carbamate protected G-AE-PA (5).$^{81}$ Compound 4 (0.110 g, 0.410 mmol), diisopropylethylamine (0.310 g, 2.40 mmol) and BOP (0.560 g, 1.30 mmol) in DMF (2.0 mL), were combined in a round-bottomed flask (50 mL), sealed with a septum and placed under nitrogen. After 30 minutes of stirring at room temperature, glucosamine (0.110 g, 0.490 mmol) slurred in 1.5 mL DMF was added. After 4 days of stirring under nitrogen at room temperature the reaction was concentrated in vacuo, and purified by flash column chromatography (8:1 chloroform:methanol). When the BOP and diisopropylethylamine finished eluting off of the flash column, as visualized by TLC, the flash conditions were transitioned to 6:4:1 chloroform:methanol:water. The fractions determined by TLC to contain crude product were concentrated in vacuo, giving a light yellow oil (0.055 g),
which was analyzed using ESI FTMS\(^+\) \([M+Na]^+\) (C\(_{11}\)H\(_{22}\)N\(_2\)O\(_7\)Na) Calcd: \(m/z = 417.1843\), found: \(m/z = 417.1850\).

**G-AE-PA (6).**\(^{78-80}\) Crude Compound 5 (0.055 g, 0.140 mmol), TFA (15.0 g, 130 mmol), a catalytic amount (~1 drop) of anisole and dichloromethane (10 mL), were combined in a round-bottomed flask (50 mL) and sealed with a septum. The reaction was then stirred for 1 hour at room temperature, concentrated in vacuo, diluted with ~50 mL of chloroform and washed with ~50 mL water. The water layer was lyophilized, then dialyzed with 100 MWCO dialysis tubing for 3 hours. The dialyzed product was lyophilized again, then subjected to RP-HPLC for purification using a linear gradient and an Alltech Prevail C18 5\(\mu\) column, 250 mm in length and with a 22 mm inner diameter. All eluents contained 0.1% TFA (v/v). The starting concentrations of eluents were 90% water and 10% acetonitrile. After 40 minutes, the ending concentrations were 78.7% water and 21.3% acetonitrile. Lyophilization of the peak at 12.1 min. gave a white solid (0.017 g, 14.2%, from Compound 4, two steps): \(^1\)H NMR (300 MHz, D\(_2\)O) \(\delta\) 2.60 (q, 2H, \(J = 6.0\) Hz), 3.19 (t, 2H, \(J = 5.0\) Hz), 3.44-3.55 (m, 2H), 3.70-3.94 (m, 8H), 4.70 (d, 0.5H, \(J = 8.3\) Hz), 5.20 (d, 0.5H, \(J = 3.5\) Hz); \(^{13}\)C NMR (300 MHz, D\(_2\)O, internal CH\(_3\)OH standard): \(\delta\) 35.95, 36.31, 39.19, 54.20, 56.79, 60.74, 66.27, 66.77, 70.09, 70.32, 70.83, 71.71, 74.03, 76.14, 91.02, 95.11, 174.64, 174.86.

**TEA\(_m\)-G1-(G-AE-PA)\(_3\) (7).**\(^{81}\) Compound 2 (0.004 g, 0.0100 mmol), triethylamine (0.020 g, 0.190 mmol) and BOP (0.040 g, 0.0800 mmol) in DMF (0.5 mL), were
combined in a round-bottomed flask (25 mL) and sealed with a septum and placed under nitrogen. After 30 minutes of stirring at room temperature, Compound 6 (0.0200 g, 0.0500 mmol) slurried in 0.5 mL DMF was added. After 6 days of stirring under nitrogen at room temperature, the reaction was concentrated in vacuo and dialyzed using 500 MWCO dialysis tubing in water for 7 days. The dialyzed product was lyophilized, then subjected to FPLC using a 2.5 cm x 120 cm column and Bio-Rad P-10 size exclusion gel. The column was eluted with 0.03M sodium bicarbonate. Lyophilization of the corresponding peaks gave a yellow solid (9.0 mg, 61.4%): ¹H NMR (500 MHz, D₂O) δ 2.55-2.64 (m, 12H), 2.81-2.83 (m, 6H), 3.38-3.43 (m, 6H), 3.49-3.54 (m, 3H), 3.63-3.66 (m, 13H), 3.72-3.96 (m, 26H), 5.23 (d, 1H, J = 3.5 Hz); ¹³C NMR (500 MHz, D₂O, internal CH₃OH standard) δ 36.04, 36.19, 36.42, 39.14, 53.07, 54.19, 56.88, 60.76, 60.91, 66.67, 66.81, 68.11, 68.83, 70.06, 70.27, 70.84, 71.71, 73.98, 76.10, 91.01, 95.08, 174.29, 174.50. MALDI-TOF MS [M+Na]⁺ (C₄₈H₈₇N₇O₂₇Na) Calcd: m/z = 1216.5, found: m/z = 1216.0.

**TEAₙ-G1-(DAP)₃ (8).**⁸³⁻⁸⁴ Compound 2 (0.110 g, 0.286 mmol), sulfuric acid (0.020 mL, 0.375 mmol) and methanol (5.60 mL, 138 mmol) were mixed together in a round-bottomed flask (50 mL) and sealed with a septum. After 3 days of stirring at room temperature, magnesium hydroxide (0.022 g, 0.377 mmol) and 1,3-diaminopropane (6.00 mL, 71.0 mmol) were added to the reaction. After 8 days of stirring at room temperature, the reaction was filtered and dialyzed using 100 MWCO tubing for 3 hours. The dialyzed product was then lyophilized and subjected to RP-HPLC for purification, using a linear
gradient. All eluents contained 0.1% TFA (v/v). The starting concentrations of eluents were 90% water and 10% acetonitrile. After 20 minutes, the ending concentrations were 60% water and 40% acetonitrile. Lyophilization of the peak at 13.6 min. gave a white solid (0.045 g, 29.3%): \( ^1\)H NMR (300 MHz, D\(_2\)O) \( \delta \) 1.86-1.96 (m, 6H), 2.58 (t, 6H, \( J = 5.9 \) Hz), 3.04 (t, 6H, \( J = 7.6 \) Hz), 3.33 (t, 6H, \( J = 6.8 \) Hz), 3.51 (t, 6H, \( J = 4.8 \) Hz), 3.82 (t, 6H, \( J = 5.9 \) Hz), 3.85 (t, 6H, \( J = 4.8 \) Hz); \( ^{13}\)C NMR (300 MHz, D\(_2\)O, internal CH\(_3\)OH standard) \( \delta \) 26.85, 35.94, 36.26, 37.16, 53.39, 64.30, 66.98, 174.36. ESI FTMS\(^+\) (performed at the University of Arizona Mass Spectrometry Facility) \([M+H]\)^{+} (C\(_{24}\)H\(_{52}\)N\(_7\)O\(_6\)) Calcd: \( m/z = 534.3974 \). Found: \( m/z = 534.3966 \).
APPENDIX A

$^1$H NMR Spectra

t-butyl protected TEA$_{m^-}$(G1)$_3$ (Compound 1), $^1$H NMR, 300 MHz, CDCl$_3$. 
TEAₘ₋(G1)₃ (Compound 2). $^1$H NMR, 300 MHz, D₂O.
TEA$_m$-G1-(GlcN)$_3$ (Compound 3), $^1$H NMR, 500 MHz, D$_2$O.
G-AE-PA (Compound 6), $^1$H NMR, 300 MHz, D$_2$O.
TEA$_m$-G1-(G-AE-PA)$_3$ (Compound 7), $^1$H NMR, 500 MHz, D$_2$O.
TEAm-G1-(DAP)₃ (Compound 8), $^1$H NMR, 300 MHz, D₂O.
APPENDIX B

$^{13}$C NMR Spectra

t-butyl protected TEA$_{m}$-(G1)$_3$ (Compound 1). $^{13}$C NMR 300 MHz, CDCl$_3$. 
TEA$_m$-(G1)$_3$ (Compound 2). $^{13}$C NMR, 300 MHz, D$_2$O.
TEAm$_{-}$G1-(GlcN)$_3$ (Compound 3), $^{13}$C NMR, 500 MHz, D$_2$O.
G-AE-PA (Compound 6), $^1$H NMR, 300 MHz, D$_2$O.
TEAm-G1-(G-AE-PA)$_3$ (Compound 7), $^1$H NMR, 500 MHz, D$_2$O.
TEAm-G1-(DAP)₃ (Compound 8), ¹³C NMR, 300 MHz, D₂O.
REFERENCES


(2) www.unaidstoday.org UNAIDS Annual Report 2009: Uniting the world against AIDS; UNAIDS.


(33) [http://www.fda.gov/ForConsumers/ByAudience/ForPatientAdvocates](http://www.fda.gov/ForConsumers/ByAudience/ForPatientAdvocates) Antiretroviral drugs used in the treatment of HIV infection, October 8, 2010.

(34) Broder, S. *Antiviral Research* 2010, 85, 1.


(41) de Bethune, M. *Antiviral Research* 2010, 85, 75.


(44) Hughes, B. *Nature Reviews* 2009, 8, 439.

(45) Tilton, J.; Doms, R. *Antiviral Research* 2010, 85, 91.


(47) Perry, C. *Drugs* 2010, 70, 1189.

(48) Croxtall, J.; Scott, L. *Drugs* 2010, 70, 631.

(49) Paredes, R.; Clotet, B. *Antiviral Research* 2010, 85, 245.


(74) [http://www.york.ac.uk/media/chemistry/aboutstaff/davidsmith/dks4.jpg](http://www.york.ac.uk/media/chemistry/aboutstaff/davidsmith/dks4.jpg) David Smith, October 15, 2010.


