Applications of Chemical Biology in Drug Discovery and Systems Biology:
  Fragment-based Design of Histone Deacetylase Inhibitors
&
A Chemical Approach to Understanding Polysaccharide Biosynthesis and
Protein Glycosylation

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy
in the Graduate School of The Ohio State University

By

Robert L. Woodward Jr.

Graduate Program in Chemistry

The Ohio State University

2010

Dissertation Committee:

Peng George Wang, Advisor

Robert S. Coleman

Ross E. Dalbey
ABSTRACT

The development of the field of chemical biology has highlighted the utility of chemically synthesized small molecules for the targeting and study of various biological processes. Detailed in this work are a few examples of how this concept has come to fruition. First, the design and discovery of histone deacetylase inhibitors is considered. These inhibitors represent a relatively new class of chemotherapy agents, which target the aberrant hypoacetylation of core histone proteins, a process observed in cancer cells. A click chemistry based approach was employed to rapidly develop a small library of potential inhibitors. The inhibition of two histone deacetylase isoforms was then determined by in vitro enzymatic assays, after which the cytotoxicity was evaluated in the National Cancer Institute’s human cancer cell line screen. A lead compound was discovered that was equipotent with a benchmark compound in enzymatic studies and proved active across all cell lines in the National Cancer Institute screen. This click histone deacetylase inhibitor design has thus provided a new chemical scaffold that has not only revealed a lead compound, but one which is amendable to further structural modifications given the modular nature of this approach.

Secondly, a chemical approach to understanding bacterial polysaccharide biosynthesis is described. Polysaccharides constitute a major component of bacterial cell
surfaces and play critical roles in bacteria/host interactions. The biosynthesis of such molecules, however, has mainly been characterized through in vivo genetic studies, thus precluding discernment of the details of this pathway. Accordingly, a chemical approach is described which enabled reconstitution of the Escherichia coli O-polysaccharide biosynthetic pathway in vitro. Starting with chemically prepared N-Acetyl-D-galactosamine-diphospho-undecaprenyl, the E. coli O86 oligosaccharide repeating unit was assembled via sequential enzymatic glycosylation. Successful expression of the putative polymerase Wzy via a chaperone co-expression system then allowed demonstration of polymerization in vitro using this substrate. Analysis of additional substrates revealed a defined mode of recognition for Wzy towards the lipid moiety. Specific polysaccharide chain length modality was furthermore demonstrated to result from the action of Wzz. Collectively, polysaccharide biosynthesis was chemically reconstituted in vitro, providing a well-defined system for further underpinning molecular details of this biosynthetic pathway.

Finally, the pathways through which proteins are glycosylated in prokaryotes and eukaryotes are considered. Glycosylation represents one of the most common forms of co-/posttranslational modification that a protein can undergo. The importance of this process has been highlighted by its association with properties including protein stability and bacterial pathogenicity. Accordingly, to better understand eukaryotic glycosylation, an in vitro assay was established in collaboration with the Aebi group which has enabled the characterization of a protozoan oligosaccharyltransferase Stt3d. For prokaryotes, collaborative efforts with the Feldman group are enabling discernment of the substrate specificity of a surprisingly promiscuous bacterial oligosaccharyltransferase PgL.
Dedicated to my wife, Megan, for her love, understanding, prayers and support.
ACKNOWLEDGMENTS

The past four years of my life have been an incredible experience of growth, frustration, joy and just about every emotion one could imagine. As I come to the end of this time, there are several people who deserve special mention for the roles they played. First, I must obviously thank my beautiful wife, Megan Woodward, M.D. She was a rock during the trying times and a continual source of encouragement and strength. Thank you to my parents, Bob and Jane Woodward, as well as the rest of my family, for your support throughout my academic journey. Many thanks to Jie Shen who introduced to me to the art of organic synthesis in between his many naps and randomly placed sighs which always seemed to be directed at Crystal O’Neil. I also thank Crystal for providing a little bit of sanity (or insanity depending upon how you look at it) and rationality throughout the course of these past four years. The Wang group as a whole has also proven to be a great group of people, and thus deserves mention here. Finally, I want to thank Rob Coleman, my undergraduate research advisor, and Amy Hayes for shaping and preparing me as a scientist during my final two years of undergraduate study. It was because of their tutelage that the transition into graduate school proved quite easy.
VITA

August 21, 1983 ..................................................... Born in Gallipolis, OH

2006........................................................................ B.A., Chemistry; B.S., Biology
       The Ohio State University

2006 – 2010............................................................ Graduate Fellow
       The Ohio State University

PUBLICATIONS

Research Publications


FIELDS OF STUDY

Major Field: Chemistry
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiv</td>
</tr>
<tr>
<td>List of Schemes</td>
<td>xvii</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xx</td>
</tr>
<tr>
<td>Preface</td>
<td>xxv</td>
</tr>
</tbody>
</table>

Chapters

1. Fragment-based Design of Histone Deacetylase Inhibitors......... 1

1.1 Introduction ................................................................. 1

1.2 Relationship between HDACs and cancer.............................. 5

1.3 Current HDACi designs....................................................... 7

1.4 Fragment-based assembly of a HDACi library ...................... 11
1.4.1 Chemical synthesis of a HDACi library..........................11
1.4.2 Enzyme inhibition analysis ............................................15
1.4.3 Cell line cytotoxicity assays...........................................18
1.4.4 Further investigation of HDAC inhibition .............21
1.5 Conclusions ...................................................................................22
References ...........................................................................................23

2. A Chemical Approach to Understanding Bacterial Polysaccharide Biosynthesis.............................................................29

2.1 Introduction ...................................................................................29
2.2 Lipopolysaccharide: A Gram-negative bacterial cell surface polysaccharide .................................................................32
  2.2.1 Lipid A structure and function ...........................................33
  2.2.2 Core oligosaccharide structure and function..................37
  2.2.3 O-polysaccharide structure and function.......................38
2.3 Polysaccharide biosynthetic pathways .................................39
2.4 In vitro reconstitution of wzy-dependent polysaccharide biosynthesis .............................................................41
  2.4.1 Expression and purification of Wzy...............................41
  2.4.2 Synthesis of RU-PP-Und................................................43
  2.4.3 Polymerization assays ....................................................48
  2.4.4 Regeneration of modal distributions ..............................49
2.5 Analysis of Wzy substrate specificity .................................52
  2.5.1 Relevance of the pentasaccharide moiety ......................52
2.5.2 Relevance of the polyprenol moiety...............................54

2.5.2.1 Synthetic approaches to polyprenols ..............56

2.5.2.1.1 Solution phase synthetic methods ... 56

2.5.2.1.2 Solid phase synthetic methods ....... 65

2.5.2.2 Synthesis of a polyprenol library ..............67

2.5.2.3 Interpretation of polymerization data ..........74

2.6 Conclusions ...................................................................................77

References............................................................................................80

3. In vitro Studies of a Protozoan Oligosaccharyltransferase...........89

3.1 Introduction ...................................................................................89

3.2 N-Glycosylation pathway overview ..............................................92

3.3 Biological significance of eukaryotic N-glycosylation.............96

3.4 Development of an in vitro assay for eukaryotic
oligosaccharide transfer .................................................................98

3.4.1 Design of the in vitro glycosylation system......... 99

3.4.2 In vitro reconstitution of oligosaccharide transfer ....102

3.5 Conclusions .................................................................................110

References .........................................................................................112

4. Analysis of the Substrate Specificity of a Bacterial
Oligosaccharyltransferase..............................................................116

4.1 Introduction ..................................................................................116

4.2 Bacterial protein glycosylation..................................................117
4.2.1 N-Glycosylation ............................................................117
4.2.2 O-Glycosylation ............................................................119
4.3 Substrate specificity of PgIL.........................................................121
  4.3.1 Relevance of the carbohydrate moiety ......................122
  4.3.2 Relevance of the polyprenol moiety .......................123
    4.3.2.1 Preliminary in vitro results .........................124
    4.3.2.2 Synthesis of RU-PP-Lipid analog library .....126
    4.3.2.3 Polyprenol specificity of PgIL ...................127
4.4 Conclusions ..................................................................................128
References ..........................................................................................129

5. Experimental Procedures.................................................................132
  5.1 HDACi chemical synthesis and compound characterization ......132
  5.2 Biological studies of HDACi......................................................141
  5.3 Sugar-PP-Polyprenol chemical synthesis and compound
      characterization .........................................................................142
  5.4 RU-PP-Polyprenol enzymatic synthesis and product
      characterization .........................................................................173
  5.5 Polymerization assays.............................................................178
  5.6 Chemical synthesis and characterization of the Stt3d substrate
      and associated precursors......................................................178
  5.7 Chemical synthesis and characterization of PgIL substrates and
      associated precursors...........................................................183
Bibliography..........................................................................................188

Appendix: HPLC data, NCI screening results, and NMR spectra ..........208
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Histone modifications and associated regulatory functions .....3</td>
</tr>
<tr>
<td>1.2</td>
<td>Activity of HDACi against HDAC1 and HDAC8 ...............17</td>
</tr>
<tr>
<td>1.3</td>
<td>Average inhibition values of a 10 μM HDACi treatment of the NCI cancer cell lines .........................................................19</td>
</tr>
<tr>
<td>2.1</td>
<td>General classification scheme for polysaccharides ............30</td>
</tr>
<tr>
<td>2.2</td>
<td>Comparison of natural and synthetic Lipid A lethality in rats ...........................................................................................36</td>
</tr>
<tr>
<td>2.3</td>
<td>Proposed library of polyprenols for Wzy substrate specificity study .................................................................55</td>
</tr>
<tr>
<td>2.4</td>
<td>Comparison of Wzy activity among RU-PP-Polyprenol analogs ....................................................................................76</td>
</tr>
<tr>
<td>3.1</td>
<td>Covalent modifications of proteins .....................................90</td>
</tr>
<tr>
<td>4.1</td>
<td>Proposed library of polyprenols for PglL substrate specificity study .................................................................126</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Packaging of DNA through complexation with histone proteins</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Regulation of transcription through reversible histone acetylation</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>Representative examples of current HDACi</td>
<td>7</td>
</tr>
<tr>
<td>1.4</td>
<td>Common structural motifs of current HDACi</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>Comparison of GI50 values from the NCI human cancer cell line screen for 1.14 and SAHA</td>
<td>20</td>
</tr>
<tr>
<td>1.6</td>
<td>Comparison of HDAC1 inhibition by SAHA and 1.14</td>
<td>21</td>
</tr>
<tr>
<td>2.1</td>
<td>Prevalence of polysaccharides in a) Gram-positive and b) Gram-negative bacteria</td>
<td>31</td>
</tr>
<tr>
<td>2.2</td>
<td>Structure of a representative lipopolysaccharide</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>Structure of the LPS Lipid A region</td>
<td>34</td>
</tr>
<tr>
<td>2.4</td>
<td>Proposed pathway for LPS initiated immune response</td>
<td>35</td>
</tr>
<tr>
<td>2.5</td>
<td>Structure of a representative Core oligosaccharide region</td>
<td>38</td>
</tr>
<tr>
<td>2.6</td>
<td>Structure of a representative O-polysaccharide region</td>
<td>39</td>
</tr>
<tr>
<td>2.7</td>
<td>wzy-dependent pathway of O-polysaccharide biosynthesis</td>
<td>40</td>
</tr>
<tr>
<td>2.8</td>
<td>Expression of polymerase Wzy</td>
<td>43</td>
</tr>
</tbody>
</table>
2.9 Analysis of Wzy activity using radiolabeled RU-PP-Und......49
2.10 Confirmation of the regulatory role of Wzz in polysaccharide biosynthesis.................................................51
2.11 Confirmation of the functional roles of Wzy and Wzz ..........52
2.12 Evaluation of RU specificity..............................................53
3.1 Diagrammatic representation of the five distinct types of sugar–peptide bonds that have currently been identified....91
3.2 The dolichol pathway of N-linked protein glycosylation in S. cerevisiae .................................................................93
3.3 The calnexin–calreticulin cycle.................................................95
3.4 Binding interactions among cells of the immune system necessary for the immune response..........................98
3.5 Demonstration of in vitro Stt3d-mediated peptide glycosylation ........................................................................102
3.6 Confirmation of Stt3d-mediated glycosylation.................103
3.7 Effect of Chitobiose-PP-MS-Pent concentration on Stt3d-mediated peptide glycosylation..........................104
3.8 Kinetics of Stt3d-mediated peptide glycosylation at different Chitobiose-PP-MS-Pent concentrations ...............105
3.9 Time course for Stt3d-mediated peptide glycosylation........106
3.10 Effect of pH on Stt3d-mediated peptide glycosylation ........107
3.11 Effect of divalent cation on Stt3d-mediated peptide glycosylation .................................................................108
3.12 Effect of free monosaccharides on Stt3d-mediated peptide glycosylation ..........................................................109
3.13 Effect of free UDP on Stt3d-mediated peptide glycosylation.................................................................109
3.14 Effect of temperature on Stt3d-mediated peptide glycosylation .......................................................... 110

4.1 Bacterial N-glycosylation in *C. jejuni* ........................................ 118

4.2 Protein glycosylation locus (*pgl*) gene cluster found in *C. jejuni* ........................................................................................................... 119

4.3 Bacterial O-glycosylation in *N. meningitidis* strain C311#3 ................................................................................................................... 121

4.4 Carbohydrate structures transferred by PglL ....................... 123

4.5 Preliminary investigation of PglL polyprenol specificity ...... 125

4.6 Polyprenol substrate specificity screen for PglL................. 128
# LIST OF SCHEMES

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Phase I metabolism of SAHA to form metabolite 4-Anilino-4-oxobutanoic acid (1.6)</td>
<td>9</td>
</tr>
<tr>
<td>1.2</td>
<td>Phase II transformation of SAHA to form glucuronylated metabolite 1.5</td>
<td>9</td>
</tr>
<tr>
<td>1.3</td>
<td>Design of clicked HDACi library</td>
<td>12</td>
</tr>
<tr>
<td>1.4</td>
<td>Preparation of the azide library</td>
<td>13</td>
</tr>
<tr>
<td>1.5</td>
<td>Synthesis of the alkyne precursor 1.23</td>
<td>14</td>
</tr>
<tr>
<td>1.6</td>
<td>Synthesis of the alkyne precursor 1.22</td>
<td>14</td>
</tr>
<tr>
<td>1.7</td>
<td>Assembly of HDACi library via click chemistry</td>
<td>15</td>
</tr>
<tr>
<td>2.1</td>
<td>Synthesis of peracetylated GalNAc-1-phosphate (2.5)</td>
<td>44</td>
</tr>
<tr>
<td>2.2</td>
<td>Synthesis of Undecaprenyl-phosphate (2.8)</td>
<td>45</td>
</tr>
<tr>
<td>2.3</td>
<td>Mechanism of formation of the coupled adduct peracetylated GalNAc-PP-Und (2.14)</td>
<td>46</td>
</tr>
<tr>
<td>2.4</td>
<td>Synthesis of GalNAc-PP-Und (2.15)</td>
<td>47</td>
</tr>
<tr>
<td>2.5</td>
<td>Enzymatic synthesis of RU-PP-Und (2.19)</td>
<td>48</td>
</tr>
<tr>
<td>2.6</td>
<td>Stereoselective formation of trisubstituted cis-olefins via Wittig olefination</td>
<td>56</td>
</tr>
<tr>
<td>2.7</td>
<td>Sato method for the assembly of trisubstituted cis-olefin polyenes</td>
<td>57</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>2.8</td>
<td>Preparation of the key bifunctional diene for the Sato method</td>
<td>57</td>
</tr>
<tr>
<td>2.9</td>
<td>Diversification of Sato’s key diene intermediate</td>
<td>58</td>
</tr>
<tr>
<td>2.10</td>
<td>Polyprenol formation via the Sato method</td>
<td>59</td>
</tr>
<tr>
<td>2.11</td>
<td>Moiseenkov sulfonamide approach to polyprenol synthesis</td>
<td>60</td>
</tr>
<tr>
<td>2.12</td>
<td>Moiseenkov imine approach to polyprenol synthesis</td>
<td>61</td>
</tr>
<tr>
<td>2.13</td>
<td>Gibbs synthesis of farnesol stereoisomers</td>
<td>63</td>
</tr>
<tr>
<td>2.14</td>
<td>Preparation of Nerylacetone (2.72) as a common starting material for the Weimer polyprenol synthetic strategy</td>
<td>64</td>
</tr>
<tr>
<td>2.15</td>
<td>Weimer synthesis of farnesol stereoisomers</td>
<td>65</td>
</tr>
<tr>
<td>2.16</td>
<td>Solid phase adaptation of the Sato approach to polyprenol synthesis</td>
<td>66</td>
</tr>
<tr>
<td>2.17</td>
<td>Synthesis of the bifunctional diene building block 2.39</td>
<td>68</td>
</tr>
<tr>
<td>2.18</td>
<td>Preparation of the sulfone derivative of 2.39</td>
<td>69</td>
</tr>
<tr>
<td>2.19</td>
<td>Preparation of the chloride derivative of 2.39</td>
<td>69</td>
</tr>
<tr>
<td>2.20</td>
<td>Synthesis and elaboration of the cis-tetraene 2.43</td>
<td>70</td>
</tr>
<tr>
<td>2.21</td>
<td>Completion of the Heptaprenol (2.47) synthesis</td>
<td>71</td>
</tr>
<tr>
<td>2.22</td>
<td>Synthesis of Pentaprenol (2.92)</td>
<td>72</td>
</tr>
<tr>
<td>2.23</td>
<td>Synthesis of cis-Pentaprenol (2.95)</td>
<td>72</td>
</tr>
<tr>
<td>2.24</td>
<td>Synthesis of the α-saturated diene building block 2.101</td>
<td>73</td>
</tr>
<tr>
<td>2.25</td>
<td>Preparation of coupling partner 2.103 for the MS-Pentaprenol synthesis</td>
<td>74</td>
</tr>
<tr>
<td>2.26</td>
<td>Synthesis of MS-Pentaprenol (2.105)</td>
<td>74</td>
</tr>
<tr>
<td>3.1</td>
<td>Synthesis of the OTase substrate Chitobiose-PP-MS-Pent (3.4)</td>
<td>101</td>
</tr>
</tbody>
</table>
4.1  Synthesis of cis-Farnesol (4.8) .............................................. 127
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>ABC</td>
<td>ATP-binding cassette</td>
</tr>
<tr>
<td>Ac</td>
<td>acetyl</td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine diphosphate</td>
</tr>
<tr>
<td>Ar</td>
<td>aryl</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>br</td>
<td>broad (IR and NMR)</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>Bac</td>
<td>bacillosamine</td>
</tr>
<tr>
<td>Bn</td>
<td>benzyl</td>
</tr>
<tr>
<td>n-Bu</td>
<td>normal-butyl</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celsius</td>
</tr>
<tr>
<td>calcd</td>
<td>calculated</td>
</tr>
<tr>
<td>CDI</td>
<td>1,1’-carbonyldiimidazole</td>
</tr>
<tr>
<td>CNX</td>
<td>calnexin</td>
</tr>
<tr>
<td>CRT</td>
<td>calreticulin</td>
</tr>
<tr>
<td>δ</td>
<td>chemical shift in parts per million downfield from tetramethylsilane</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>d</td>
<td>doublet (spectra); day(s)</td>
</tr>
<tr>
<td>DATDH</td>
<td>2,4-diacetamido-2,4,6-trideoxyhexose</td>
</tr>
<tr>
<td>DC-SIGN</td>
<td>dendritic cell-specific ICAM-3-grabbing non-integrin</td>
</tr>
<tr>
<td>DIBAL-H</td>
<td>diisobutylaluminum hydride</td>
</tr>
<tr>
<td>DMF</td>
<td>$N,N$-dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>eq.</td>
<td>equivalent</td>
</tr>
<tr>
<td>Et</td>
<td>ethyl</td>
</tr>
<tr>
<td>Farn</td>
<td>farnesol</td>
</tr>
<tr>
<td>FOA</td>
<td>5-fluoroorotic acid</td>
</tr>
<tr>
<td>Fuc</td>
<td>fucose</td>
</tr>
<tr>
<td>FucNAc</td>
<td>$N$-acetylfucosamine</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>gamma</td>
</tr>
<tr>
<td>g</td>
<td>gram(s)</td>
</tr>
<tr>
<td>Gal</td>
<td>galactose</td>
</tr>
<tr>
<td>GalNAc</td>
<td>$N$-acetylgalactosamine</td>
</tr>
<tr>
<td>GI$_{50}$</td>
<td>concentration at which 50% growth inhibition is observed</td>
</tr>
<tr>
<td>Glc</td>
<td>glucose</td>
</tr>
<tr>
<td>GlcA</td>
<td>glucuronic acid</td>
</tr>
<tr>
<td>GlcNAc</td>
<td>$N$-acetylglucosamine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>h</td>
<td>hour(s)</td>
</tr>
<tr>
<td>HA</td>
<td>hydroxamic acid</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyltransferase</td>
</tr>
<tr>
<td>HDAC</td>
<td>histone deacetylase</td>
</tr>
<tr>
<td>Hept</td>
<td>heptaprenol</td>
</tr>
<tr>
<td>HMPA</td>
<td>hexamethylphosphoramide</td>
</tr>
<tr>
<td>ICAM-3</td>
<td>intracellular adhesion molecule 3</td>
</tr>
<tr>
<td>IL1-β</td>
<td>interleukin 1, beta</td>
</tr>
<tr>
<td>Im</td>
<td>imidazole</td>
</tr>
<tr>
<td>iPr</td>
<td>iso-propyl</td>
</tr>
<tr>
<td>J</td>
<td>coupling constant in Hz (NMR)</td>
</tr>
<tr>
<td>Kdo</td>
<td>3-deoxy-D-manno-oct-2-ulosonic acid</td>
</tr>
<tr>
<td>K_m</td>
<td>Michaelis constant</td>
</tr>
<tr>
<td>L</td>
<td>liter(s)</td>
</tr>
<tr>
<td>LBP</td>
<td>serum LPS-binding protein</td>
</tr>
<tr>
<td>LDA</td>
<td>lithium diisopropylamide</td>
</tr>
<tr>
<td>m</td>
<td>milli; multiplet (NMR)</td>
</tr>
<tr>
<td>μ</td>
<td>micro</td>
</tr>
<tr>
<td>M</td>
<td>moles per liter</td>
</tr>
<tr>
<td>Man</td>
<td>mannose</td>
</tr>
<tr>
<td>mCPBA</td>
<td>meta-chloroperoxybenzoic acid</td>
</tr>
<tr>
<td>Me</td>
<td>methyl</td>
</tr>
<tr>
<td>MEM</td>
<td>methoxyethoxymethyl</td>
</tr>
</tbody>
</table>
MHz  megahertz

min  minute(s)

mol  mole(s)

mRNA  messenger ribonucleic acid

Ms  methanesulfonyl

MS  mass spectrometry

MS-Pent  monosaturated-pentaprenol

MurNAc  N-acetylmuramic acid

m/z  mass to charge ratio (MS)

NBS  N-bromosuccinimide

NCI  National Cancer Institute

nm  nanometer

NMR  nuclear magnetic resonance

NTA  nitrilotriacetic acid

OPS  O-polysaccharide

OTase  oligosaccharyltransferase

p  para

Pent  Pentaprenol

pgl  protein glycosylation locus

Ph  phenyl

PMB  p-methoxybenzyl

PP  pyrophosphate

ppm  parts per million
q  quartet (NMR)
Rf  retention factor
rt  room temperature
RU  repeating unit
s  singlet (NMR); second(s)
SAHA  suberoylanilide hydroxamic acid
SDS-PAGE  sodium dodecyl sulfate polyacrylamide gel electrophoresis
Solan  solanesol
t  triplet (NMR)
TBAB  tetrabutylammonium bromide
TBAF  tetrabutylammonium fluoride
TBDPS  t-butyldiphenylsilyl
THF  tetrahydrofuran
THP  tetrahydropyranyl
TLC  thin layer chromatography
TLR4  toll-like receptor 4
TMS  trimethylsilyl
TNF-α  tumor necrosis factor-alpha
TSA  trichostatin A
TsOH  tosyllic acid
UDP  uridine diphosphate
Und  undecaprenol
Researchers who have treated chemistry and biology as two distinct fields have been responsible for various scientific advances. For example, synthetic organic chemists such as R. B. Woodward have provided artificial routes to highly complex molecules including strychnine and chlorophyll. Biologists including Oswald Avery and Frederick Griffith, in contrast, have elucidated the complexities of natural biological processes such as determining that DNA functions as the hereditary material. While these discoveries have helped to considerably advance the fields of chemistry and biology, a more recent mindset has been to explore the interface between these disciplines. Such exploration at the chemistry-biology interface has given rise to a new field, that of chemical biology, which will serve as the central theme in this dissertation.

Chemical biology, in simplest terms, deals with the application of chemistry to the study of biological systems. Most commonly, this application involves the use of chemically synthesized small molecules to probe various biological processes. The advantage of utilizing synthetic molecules is that chemists can readily modify the natural skeleton of a particular biomolecule to afford various derivatives. How these modifications influence the different processes can then be studied, bringing to light important information such as structure-activity relationships and inhibitory potential.
Detailed in this dissertation are a few examples of this approach including the evaluation of a library of chemically prepared small molecules for histone deacetylase inhibition and the synthesis of naturally occurring substrates and associated derivatives to better understand polysaccharide biosynthesis and protein glycosylation. However, as will be seen, whether it is the development of an antitumor agent or the probing of a biochemical pathway, a common theme exists. That theme is the power of uniting the fields of chemistry and biology such that 1) the process of drug discovery is greatly facilitated and 2) critical details of biological pathways are revealed.
CHAPTER 1

FRAGMENT-BASED DESIGN OF HISTONE DEACETYLASE INHIBITORS

1.1 Introduction

Early research in the field of cancer development suggested that the onset of this disease rested primarily with direct alterations of the genome. Specifically, mutations and deletions leading to the activation of proto-oncogenes or the inactivation of tumor-suppressor genes were held to be the major initiators of cancer.\(^1\) More recently, however, a much broader perspective has developed in which causes outside of direct changes to the DNA, such as variations from typical gene regulation patterns, have garnered increasing attention. One such example of altered expression regulation occurs at the transcriptional level, specifically through regulation of the structural state of chromatin.\(^2\)

Chromatin can be generally defined as the complex formed between DNA and proteins. At the most basic level, such complexation causes chromatin to adopt the appearance of “beads on a string” in which 147 bp of DNA (the string) is wrapped around an octamer of the four core histone proteins H3, H4, H2A and H2B (the beads) to form
nucleosomes (Figure 1.1). Through the help of additional proteins such as histone H1, these “beaded strings” subsequently form higher-order packing structures including the 30 nm chromatin fiber and the more familiar chromosome.

Figure 1.1. Packaging of DNA through complexation with histone proteins.
As alluded to above, these condensed structures help regulate gene expression. More specifically, such regulation is brought about through the posttranslational modification of the histone component of chromatin. Posttranslational modifications occur in a variety of forms (Table 1.1) and have primarily been associated with the unstructured N-terminal “tails” of the core histones.\textsuperscript{2} Within the past five years, however, over 30 modifications have also been found to occur on the more structured regions of the nucleosome.\textsuperscript{4} Nevertheless, it is the “tail-modifications,” in particular the example of acetylation, which have been of significant interest to the cancer research community.

<table>
<thead>
<tr>
<th>Posttranslational Modifications</th>
<th>Functions Regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylation</td>
<td>Transcription, Repair, Replication, Condensation</td>
</tr>
<tr>
<td>ADP Ribosylation</td>
<td>Transcription</td>
</tr>
<tr>
<td>Deimination</td>
<td>Transcription</td>
</tr>
<tr>
<td>Methylation</td>
<td>Transcription, Repair</td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>Transcription, Repair, Condensation</td>
</tr>
<tr>
<td>Proline Isomerization</td>
<td>Transcription</td>
</tr>
<tr>
<td>Sumoylation</td>
<td>Transcription</td>
</tr>
<tr>
<td>Ubiquitylation</td>
<td>Transcription, Repair</td>
</tr>
</tbody>
</table>

\textbf{Table 1.1.} Histone modifications and associated regulatory functions.\textsuperscript{2}
The reversible acetylation of N-terminal Lys residues helps control the remodeling status of chromatin via charge-charge interactions between negatively charged DNA and neutral or positively charged histones.\textsuperscript{5-7} As shown below, the removal and addition of these acetyl groups is catalyzed by the enzymes histone deacetylase (HDAC) and histone acetyltransferase (HAT), respectively (Figure 1.2). HAT catalyzed addition of an acetyl group removes the positive charge which would normally be present on a lysine residue. The histones thus lose their affinity for the negatively charged backbone of DNA, allowing the DNA to bind much more loosely. This loose form renders the DNA more susceptible to expression as the cellular machinery can now access genes. Upon removal of an acetyl group by HDAC, however, the positive charge returns to the histone Lys residues, forming a tighter complex between the DNA and histones. In contrast to the loose form, this tight form is much less accessible to gene expression machinery.
1.2 Relationship between HDACs and cancer

In the past ten years, studies on the above process of reversible histone acetylation have attracted an increasing amount of attention. This popularity has largely been driven by observations that aberrant hypoacetylation of histones frequently occurs in tumor cells. As described above, the absence of acetyl groups causes DNA to be present in a tightly associated complex with the histones, thus limiting the opportunity for gene expression. Given that some of these genes are tumor-suppressor genes, the observed hypoacetylation is then effectively silencing the means by which cells would normally regulate growth.\(^8\) Accordingly, clinical and biological studies have shown that inhibition of HDACs can selectively inhibit cancer cell growth.\(^9\) It has also been revealed
that HDACs have a panel of non-histone targets, many of which are involved in
tumorigenesis.\textsuperscript{10-18} All of these features therefore establish HDAC as an attractive target in cancer chemotherapy.\textsuperscript{1,19-21}

Thus far, eighteen human HDAC subtypes have been identified and divided into four general classes with respect to their homology to yeast HDACs.\textsuperscript{1} Class I (yeast transcriptional regulator RPD3) and II (yeast Hda 1) HDACs both require a zinc ion to mediate deacetylation. However, the localization of these isoforms is different in that Class I HDACs are found exclusively in the nucleus, whereas Class II HDACs are found in both the nucleus and cytoplasm. These two classes are furthermore mechanistically different from the Class III (SIR2 yeast family) HDACs which are NAD\textsuperscript{+} dependent. While HDAC 11 possesses similar catalytic activity to Class I and II HDACs in terms of requiring a zinc ion, it actually serves as the sole member of Class IV due to low overall sequence similarity with Classes I and II.

In the case of the aforementioned Zn(II) dependant HDAC isozymes, the active site structure was first revealed by a homolog of the hyperthermophilic bacterium Aquifex aeolicus in 1999.\textsuperscript{22} In general, the active site is a tube-like pocket that accommodates the side chain of client Lys residues. Located at the bottom of this pocket is the Zn(II) cofactor, which acts as a Lewis acid to catalyze hydrolysis of the acetylated Lys side chain. The opening region, or “cap” region, of the active site consists of multiple loops and is thus highly malleable.
1.3 Current HDACi designs

From this elucidation of the HDAC active site structure, progress on HDAC inhibitor (HDACi) design has been greatly facilitated (Figure 1.3).\textsuperscript{1,23-30} In most cases, the reported inhibitor is composed of the following three moieties: a Zn(II) binding moiety, a spacer moiety and a “cap” moiety (Figure 1.4).\textsuperscript{31-35} Of the demonstrated Zn(II) binding moieties, the hydroxamic acid (HA) functionality is the most common. For example, the HA moiety exists in the naturally occurring compound \textbf{TSA} as well as in the clinically approved drug \textbf{SAHA}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Figure 1.3. Representative examples of current HDACi.}
\end{figure}
Figure 1.4. Common structural motifs of current HDACi (Blue = Cap; Orange = Spacer; Pink = Zn(II) binder)

In spite of the strong affinity of HA for Zn(II), the *in vivo* efficacy of HA-containing HDACi is impaired by the short clearance time of these agents.\(^{36-38}\) For example, in the case of **SAHA**, two major pathways appear responsible for this clearance.\(^{39-42}\) The first pathway incorporates what are known as phase I transformations, or rather reactions which either introduce or reveal a new functional group within the target molecule.\(^{43}\) **SAHA** undergoes a series of these transformations, first being converted into the corresponding amide through hydrolytic cleavage of the HA functional group (*Scheme 1.1*). Subsequent deamination forms carboxylic acid derivative 1.5 which then twice undergoes β-oxidation to afford the final metabolite (1.6).
Scheme 1.1. Phase I metabolism of SAHA to form metabolite 4-Anilino-4-oxobutanoic acid (1.6).

The second metabolic pathway involves a phase II transformation. Such biotransformations involve conjugation of highly polar molecules to drugs so as to cause inactivation and promote urine/bile excretion. For SAHA, UDP-Glucuronosyltrasferase conjugates glucuronic acid to the HA moiety (Scheme 1.2).

Scheme 1.2. Phase II transformation of SAHA to form glucuronylated metabolite 1.5.
Given these clearance issues, non-HA HDACi design has been explored.\textsuperscript{44,45} These non-HA moieties include electrophilic ketones (1.1), \textit{O}-aminoanilides (\textbf{MS-275}), thiols (or the pro-drug forms) (\textbf{FK-228}, 1.2), and a variety of others. However, in terms of binding affinity, none are as strong as the HA moiety.

As for the spacer moiety, proper length is the key factor in successful HDAC inhibition. For example, six CH\textsubscript{2} units are preferred in the case of \textbf{SAHA} and its analogs, whereas chiral center and HA-containing HDACi (\textbf{CHAPs}, 1.3) prefer five CH\textsubscript{2} units between the HA and chiral center.\textsuperscript{46-48} Similarly, chiral center and thiol-containing HDACi (1.2) prefer five CH\textsubscript{2} units in the spacer moiety. In addition to these length requirements, introduction of an unsaturated moiety within the spacer region has also proved to be a valuable feature in enhancing potency (\textbf{TSA, MS275, (S)-HDAC-42}, 1.4). Such value appears to arise from \(\pi-\pi\) stacking interactions between the unsaturated moiety and two conserved Phe residues that occur near the midpoint of the tube-like pocket in the eleven Zn(II) dependant human HDACs.\textsuperscript{49}

In regards to the “cap” moiety, a hydrophobic substituent, especially an aromatic substituent, is frequently incorporated. It should be noted that structure modification in the “cap” region is a key strategy in current HDACi design, since topological differences are observed in the corresponding “cap” regions of HDAC isozymes.
1.4 Fragment-based assembly of a HDACi library

1.4.1 Chemical synthesis of a HDACi library

To further facilitate inhibitor design, a scaffold which is readily amendable to structure diversification was highly desired. A fragment-based approach was thus envisioned in which two libraries of small molecules would be assembled in all possible combinations to afford a larger library of potential inhibitors. Subsequent enzymatic and cell cytotoxicity screening would then be employed to identify hits which would be further developed structurally and evaluated in an in vivo context.

Accordingly, a click chemistry approach was adopted in which the Huisgen 1,3-dipolar cycloaddition of azides and alkynes was to serve as the key transformation. This reaction has been widely applied in chemical synthesis in recent years in efforts to develop new drugs.\textsuperscript{50-54} It is now recognized as the premier example of the click reaction due to the convenient availability and high stability of the alkyne and azide precursors, as well as the mild reaction conditions and nearly quantitative yields.

To employ this strategy, the aforementioned small molecule libraries were designed, one of which was composed of azide-bearing precursors corresponding to the “cap” moiety of HDACi (Scheme 1.3). The other library consisted of precursors containing the zinc chelating functionality and an alkyne group. From these libraries, the desired “clicked” products were to be produced. (Preparation of 1.8-1.14 was previously
discussed by Jie Shen (Doctoral Dissertation). These results will only be summarized here. Appropriate experimental and spectral data can be found in the original work).

![Scheme 1.3. Design of clicked HDACi library.](image)

Preparation of the azide library was initiated from the corresponding halides (Scheme 1.4). Specifically, 1.31-1.37 were stirred overnight in a 0.5 M NaN₃ DMSO solution. Addition of tetraethylammonium bromide was necessary in the case of 1.33 and 1.34 to catalyze product formation. Furthermore, generation of the corresponding azide from 1.35 required a lower temperature (0 °C). With the exception of 1.28, all of the products were of suitable purity to be used in the next reaction following work-up.
Scheme 1.4. Preparation of the azide library.

Synthesis of the alkyne precursor \textbf{1.23} commenced with PMB-protection of \(N\)-hydroxyphthalimide (\textbf{1.38}) to form \textbf{1.39} (Scheme 1.5).\textsuperscript{56} Subsequent treatment with hydrazine monohydrate released the PMB-protected hydroxylamine which was then precipitated as the hydrogen chloride salt (\textbf{1.40}). Initial conversion of the acid \textbf{1.41} to the corresponding acyl chloride was followed by addition of the PMB-protected hydroxylamine, which was pretreated with Hünig’s base. This sequence yielded the desired alkyne precursor \textbf{1.23}.

In contrast, synthesis of the alkyne precursor \textbf{1.22} (Scheme 1.6) was initiated through conversion of the propiolic acid \textbf{1.42} to the trans-iodopropenoic acid \textbf{1.43}.	extsuperscript{57} Elaboration of \textbf{1.43} to the desired eneyne \textbf{1.46} made use of a route previously described for preparation of the corresponding (\(Z\))-isomer.\textsuperscript{58} This route entailed initial protection of the acid \textbf{1.43} with MEMCl to furnish ester \textbf{1.44}. \textbf{1.44} and TMS-protected acetylene were
then coupled via a Sonogashira reaction to afford 1.45. Deprotection of 1.45 with HCl finally yielded the desired enyne 1.46. This acid then underwent amidation following the pathway described above (1.46 replaces 1.41).

Scheme 1.5. Synthesis of the alkyne precursor 1.23.

Scheme 1.6. Synthesis of the alkyne precursor 1.22.

Following preparation of these two libraries, a series of click reactions between alkyne precursors 1.22 and 1.23 and the seven azide-containing precursors 1.24-1.30 were performed to afford intermediates 1.47-1.53 and 1.54-1.60 (Scheme 1.7). In this combinatorial step, the alkyne precursors were first desilylated with CsF. Following
work-up, the crude intermediates were treated with 1.24-1.30 in the presence of a Cu(I) catalyst to give 1.47-1.53 and 1.54-1.60. The purified products were deprotected using TFA to yield the click HDACi 1.8-1.14 and 1.15-1.21. These final target compounds were purified by C-18 reverse phase column chromatography due to the presence of an unidentified red impurity which appeared upon addition to normal phase silica gel.

Scheme 1.7. Assembly of HDACi library via click chemistry.

1.4.2 Enzyme inhibition analysis

In order to assess the biological activity of these fourteen compounds, they were first tested for inhibition of HDAC. Specifically, HDAC1 and HDAC8 were chosen for
the preliminary screen. Selection of these two isozymes stemmed from the previously reported correlation between HDAC1 inhibition and cancer cell growth inhibition, as well as the existence of preferences among several HDAC inhibitors for HDAC1 or HDAC8. Accordingly, preliminary screening at a concentration of 0.50 μM (Table 1.2) was conducted. This screening revealed that one of the compounds, 1.14, possessed an inhibitory potency towards HDAC1 that was quite similar to that of SAHA. In addition, several other inhibitors (1.8-1.13, 1.15, 1.16, 1.21) possessed activity against HDAC1, albeit to a lesser extent than either SAHA or 1.14. Of particular importance in this regard is 1.21, given that it showed only 10 ± 1% inhibition at the concentration tested. This is notable due to the fact that it and 1.14, which demonstrated 75 ± 5% inhibition, only differ in the length of the linker region. This suggests that a longer linker length considerably aids HDAC1 inhibition with this particular class of inhibitors.
<table>
<thead>
<tr>
<th></th>
<th>% Inhibition of HDAC1</th>
<th>% Inhibition of HDAC8</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAHA</td>
<td>72 ± 2</td>
<td>NS</td>
</tr>
<tr>
<td>1.8</td>
<td>31 ± 4</td>
<td>12 ± 6</td>
</tr>
<tr>
<td>1.9</td>
<td>10 ± 3</td>
<td>NS</td>
</tr>
<tr>
<td>1.10</td>
<td>8 ± 2</td>
<td>NS</td>
</tr>
<tr>
<td>1.11</td>
<td>13 ± 4</td>
<td>NS</td>
</tr>
<tr>
<td>1.12</td>
<td>8 ± 2</td>
<td>NS</td>
</tr>
<tr>
<td>1.13</td>
<td>15 ± 3</td>
<td>NS</td>
</tr>
<tr>
<td>1.14</td>
<td>75 ± 5</td>
<td>NS</td>
</tr>
<tr>
<td>1.15</td>
<td>18 ± 3</td>
<td>12 ± 2</td>
</tr>
<tr>
<td>1.16</td>
<td>9 ± 1</td>
<td>12 ± 2</td>
</tr>
<tr>
<td>1.17</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>1.18</td>
<td>NS</td>
<td>26 ± 3</td>
</tr>
<tr>
<td>1.19</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>1.20</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>1.21</td>
<td>10 ± 1</td>
<td>13 ± 2</td>
</tr>
</tbody>
</table>

**Table 1.2.** Activity of HDACi against HDAC1 and HDAC8. The values given represent the percentage by which the HDAC activity was inhibited by a 0.50 μM HDACi treatment. All of the assays were repeated at least three times. When the inhibition capacity was less than 5%, NS (meaning no significant inhibition) was assigned.

While the above ten compounds showed inhibition against HDAC1, only five of the fourteen (1.8, 1.15, 1.16, 1.18, 1.21) demonstrated any significant inhibition of HDAC8. Of these five, four also possessed activity against HDAC1 with 1.18 being the exception. Nonetheless, none of the inhibitors demonstrated as significant HDAC8 inhibition as that seen with the inhibition of HDAC1 by **SAHA** and 1.14.
1.4.3 Cell line cytotoxicity assays

Before better characterizing the activity of these compounds against HDAC, a list of the fourteen inhibitors was submitted to the National Cancer Institute (NCI) for acceptance into the human cancer cell line screening program which is provided at no cost. Of the fourteen, only one, $1.14^*$, ultimately advanced to the final five-dose screen. This selection for advancement stemmed from the results of the single-dose preliminary screen in which the percent inhibition of cell line growth was determined through comparison to untreated samples. As shown in Table 1.3, $1.14^*$ was the only compound in this preliminary screen to show good inhibition across the various cancer cell lines.

In the five-dose screen, $1.14^*$ proved to be active against each of the cancer cell lines, albeit to varying degrees (Figure 1.5 with SAHA as a benchmark). The differences among the cell lines, however, were generally within or close to one order of magnitude. For example, the GI$_{50}$ values obtained against the various leukemia cell lines all fell within the range of 0.249 - 2.99 μM. This same trend appeared among the other cancer types for which the ranges were as follows: non-small cell lung 0.199 - 2.26 μM, colon 0.347 – 2.07 μM, CNS 0.318 – 1.53 μM, melanoma 0.271 – 1.08 μM, ovarian 0.187 – 3.21 μM, renal 0.0100 – 2.38 μM, prostate 0.580 – 1.30 μM, and breast 0.180 – 3.92 μM. As can be seen, the lower end of the range for the renal cancer cell lines is noticeably lower than those of the other cancer types. Specifically, the RXF 393 cell line possessed a GI$_{50}$ < 10 nM. This represents the lowest GI$_{50}$ obtained across the
60 cell lines; however, this low value does not appear to be specific to this type of cancer, but rather cell line specific as the next most sensitive renal cell line (TK-10) possesses a GI50 value of 0.495 μM.

<table>
<thead>
<tr>
<th></th>
<th>SAHA</th>
<th>1.09</th>
<th>1.12</th>
<th>1.13</th>
<th>1.14</th>
<th>1.19</th>
<th>1.21</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td>0 %</td>
<td>45 %</td>
<td>53 %</td>
<td>68 %</td>
<td>1 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon</td>
<td>-10 %</td>
<td>51 %</td>
<td>78 %</td>
<td>83 %</td>
<td>-25 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>-3 %</td>
<td>40 %</td>
<td>78 %</td>
<td>71 %</td>
<td>8 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ovarian</td>
<td>-4 %</td>
<td>48 %</td>
<td>65 %</td>
<td>74 %</td>
<td>1 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leukemia</td>
<td>0 %</td>
<td>7 %</td>
<td>104 %</td>
<td>84 %</td>
<td>-6 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal</td>
<td>-23 %</td>
<td>46 %</td>
<td>65 %</td>
<td>90 %</td>
<td>-20 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melanoma</td>
<td>-18 %</td>
<td>33 %</td>
<td>60 %</td>
<td>69 %</td>
<td>-43 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>6 %</td>
<td>36 %</td>
<td>66 %</td>
<td>84 %</td>
<td>-9 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNS</td>
<td>5 %</td>
<td>58 %</td>
<td>67 %</td>
<td>89 %</td>
<td>-18 %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>-7 %</td>
<td>42 %</td>
<td>70 %</td>
<td>78 %</td>
<td>-12 %</td>
<td>101 %</td>
<td>103 %</td>
</tr>
</tbody>
</table>

Table 1.3. Average inhibition values of a 10 μM HDACi treatment of the NCI cancer cell lines. a,b (a) The data listed were the average percentage by which the 10 μM HDACi treatment inhibited the growth of cancer cell lines. Without the HDACi treatment, after a 48 hour incubation, the cell number corresponds to 100%. The cell number at the initial moment corresponds to 0%. (b) A negative percentage indicates that the HDACi at 10 μM decreases the cell number in comparison to the number seen at the initial moment. (c) NSI means no significant inhibition of cell growth (> 95 %).
Figure 1.5. Comparison of GI_{50} values from the NCI human cancer cell line screen for 1.14 and SAHA. Tested doses included 1x10^{-4}, 1x10^{-5}, 1x10^{-6}, 1x10^{-7} and 1x10^{-8} M. Activity against a) leukemia and non-small cell lung cancer cell lines, b) colon, central nervous system (CNS) and prostate cancer cell lines, c) melanoma and ovarian cancer cell lines, and d) renal and breast cancer cell line.
1.4.4. Further investigation of HDAC inhibition

Given the results of the NCI screen and the fact that 1.14 did not show any significant inhibition of HDAC8 at the concentration used in the preliminary screen, a more detailed analysis of HDAC1 inhibition was performed (Figure 1.6). It was found that 1.14 possesses an IC$_{50}$ value against HDAC1 (104 ± 30 nM) that is comparable to that found with SAHA (140 ± 65 nM). Thus, taken together, the results of the enzyme and cell-based assays provide support to the proposed correlation between the cytotoxicity of HDAC inhibitors and the ability of these agents to inhibit HDAC1. This notion is further supported by the NCI preliminary cell inhibition studies on the six candidates which were originally selected (Table 1.3). Those compounds which demonstrated a high degree of HDAC1 inhibition (SAHA and 1.14) also exhibited high levels of potency in the cell-based assay. In contrast, the weaker inhibitors of HDAC1 also generally exhibited reduced potency against cancer cells.

![Figure 1.6. Comparison of HDAC1 inhibition by SAHA and 1.14 (5g in this figure).](image-url)
1.5 Conclusions

The observed relationship between HDAC activity and cancer has triggered considerable interest into the targeting of these enzymes. Accordingly, a drug design strategy which enables rapid generation of potential inhibitors is highly desirable. The above click chemistry approach offers one such strategy. This fact was highlighted by the discovery of lead compound 1.14, which was found to inhibit HDAC1 with an IC$_{50}$ = 104 ± 30 nM while demonstrating no significant inhibition of HDAC8 in preliminary testing. This level of inhibition was found to be comparable to that exhibited by the clinically approved drug SAHA (140 ± 65 nM). Furthermore, 1.14 was also found to be active across all cell lines in the NCI human cancer cell line screen with GI$_{50}$ values ranging from 10 nM to 3.92 μM.

In summary, a click chemistry approach to HDACi design has been developed which has not only led to a potent lead compound, but more importantly, one which is readily amendable to further derivatization. For example, functionalization of the aromatic ring, replacement of the aromatic ring with a heteroaromatic species, or introduction of a chiral center in the “cap” moiety can all occur quite readily by incorporating these features into the azide-containing fragment. This will allow modifications to occur prior to the combinatorial step, thus maintaining the modular nature of this approach. A large number of new candidate molecules can therefore be rapidly prepared via modification of the azide library.


CHAPTER 2

A CHEMICAL APPROACH TO UNDERSTANDING BACTERIAL POLYSACCHARIDE BIOSYNTHESIS

2.1 Introduction

Polysaccharides constitute one of the major classes of macromolecules in living organisms. These molecules are ubiquitous in nature and play essential roles in various biological processes including embryonic development, energy storage, signal transduction, cell-cell interactions, and stimulation of the immune response.\textsuperscript{1-3} Structurally, polysaccharides are classified into one of the following two groups: Homopolysaccharides and Heteropolysaccharides (Table 2.1). Homopolysaccharides, which are composed of a single type of monosaccharide residue, primarily perform structural and energy storage roles.\textsuperscript{4-7} Examples include glycogen, starch, cellulose and chitin. In contrast, heteropolysaccharides exhibit a much more diverse set of functions, a fact which likely arises from the increased structural complexities seen with this type of polysaccharide. Specifically, heteropolysaccharides serve as essential components of cell walls and extracellular matrices (peptidoglycan, glycosaminoglycan), major pathogenic factors in bacteria (O-polysaccharides (OPSs), capsular polysaccharides, exopolysaccharides), and factors of cellular development and cell-cell interactions.\textsuperscript{8-17}
<table>
<thead>
<tr>
<th>Name</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Homopolysaccharides</strong></td>
<td></td>
</tr>
<tr>
<td>Glycogen/Starch</td>
<td>-(Glc-α1,4-Glc)$_n$-</td>
</tr>
<tr>
<td>Cellulose</td>
<td>-(Glc-β1,4-Glc)$_n$-</td>
</tr>
<tr>
<td>Chitin</td>
<td>-(GlcNac-β1,4-GlcNac)$_n$-</td>
</tr>
<tr>
<td><strong>Heteropolysaccharides</strong></td>
<td></td>
</tr>
<tr>
<td>Peptidoglycan</td>
<td>-(GlcNac-β1,4-MurNac)$_n$-</td>
</tr>
<tr>
<td>Glycosaminoglycan</td>
<td>-(GlcNac-β1,4-Glc)$_n$-</td>
</tr>
<tr>
<td></td>
<td>-(GalNac-β1,4-Glc)$_n$-</td>
</tr>
<tr>
<td></td>
<td>-(GlcNac-β1,3-Gal)$_n$-</td>
</tr>
<tr>
<td>O-polysaccharides</td>
<td>-(4Fuc-α1,2-Gal-β1,3-GalNac-α1,3-GalNac)$_n$-</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>α-Gal</td>
</tr>
<tr>
<td>Capsular polysaccharides</td>
<td>-(3Glc-β1,6-GlcNac-β1,3-Gal)$_n$-</td>
</tr>
<tr>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>β-Gal</td>
</tr>
<tr>
<td>Exopolysaccharides</td>
<td>-(Glc-β1,4-Glc)$_n$-</td>
</tr>
<tr>
<td></td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2.1. General classification scheme for polysaccharides.
Given the ability of these complex polysaccharides to serve in such diverse roles, it is not surprising that bacteria have exploited these molecules for many processes critical to their survival. For example, both Gram-positive and Gram-negative bacterial cell walls are adorned with several types of heteropolysaccharides (Figure 2.1). In the case of Gram-negative bacteria, the surface polysaccharides primarily expressed at 37 °C are the capsular polysaccharides and, more importantly, the OPSs which will now be considered in greater detail.19

2.2 Lipopolysaccharide: A Gram-negative bacterial cell surface polysaccharide

As described above, OPSs represent one of the major types of polysaccharides expressed on Gram-negative bacterial cell surfaces. It should be noted, however, that the importance of these molecules does not rest within the OPS in and of itself; rather, this polysaccharide exists as one component of an immunologically important surface molecule known as lipopolysaccharide, or LPS. Structurally, LPS can be divided into the following three regions: Lipid A, Core Oligosaccharide and OPS (Figure 2.2). As will be discussed, it is through the linking of these regions that several types of resistance/immunostimulatory activities are combined to give bacteria an effective means of invasion and defense against host systems. Thus, given that each region plays a role in the immunological relationship between a Gram-negative bacterium and its host, the structure and function of each will now be considered in detail.
Figure 2.2. Structure of a representative lipopolysaccharide.

2.2.1 Lipid A structure and function

The Lipid A region of LPS is composed of a β1,6-linked glucosamine disaccharide which is phosphorylated at the 1- and 4'-positions as well as acylated with fatty acids at the N-2/2' and 3/3'-positions (Figure 2.3). These structural modifications play critical roles in the survival and virulence of bacteria. For example, the fatty acid chains provide the necessary hydrophobic moieties for LPS to serve as the primary component of the outer lipid leaflet layer of the outer membrane (Figure 2.1). Additionally, studies on the use of Lipid A variants as vaccines and adjuvants found that the 1-phosphate group is critical to bacterial endotoxic activity. Recent crystallographic
Evidence suggests this importance arises from phosphate-dependent dimerization of an immunostimulatory complex as described below.\(^{27}\)

**Figure 2.3.** Structure of the LPS Lipid A region.

Initial interest in this region of LPS originated from a hypothesis that LPS in its entirety was responsible for the acute pathophysiological effects observed in patients suffering from severe bacterial infections.\(^{28}\) The pathway through which these effects are propagated has largely been discerned (Figure 2.4).\(^{13,27}\) Specifically, as LPS dissociates from bacterial membranes within a host, Serum LPS-binding protein (LBP) forms a complex with the free LPS. LBP acts as a shuttle, transferring this biomolecule to the CD14 protein on the surface of macrophages and endothelial animal cells.\(^{29,30}\) This CD14-LPS complex then interacts with toll-like receptor 4 (TLR4), ultimately transferring LPS to TLR4. This newly formed complex further interacts with the soluble accessory protein MD-2. Recently, the crystal structure of this TLR4-MD-2-LPS complex was solved, and it indicates that this entire complex exists a dimer. Such
Dimerization is promoted in part through interactions of the 1-phosphate group of Lipid A with TLR4 Lys and Arg residues. Given that endotoxic activity is lost upon removal of this phosphate moiety, as mentioned above, it is highly likely that dimerization is necessary to bring about the immunostimulatory effects associated with LPS.

**Figure 2.4.** Proposed pathway for the LPS initiated immune response.

Upon successful complex assembly, the production of IL1-β, TNF-α and other inflammation mediators, along with stimulatory molecules associated with the adaptive immune response, is triggered. While such a response is typically highly favorable in the setting of localized infections, the occurrence of severe sepsis causes the
aforementioned inflammation mediators and stimulatory molecules to be released systemically. This overproduction ultimately leads to blood vessel damage, Gram-negative septic shock, intravascular coagulation and possible multi-organ failure.\textsuperscript{13,33-36}

While the above pathway offers an elegant model for the induction of LPS-associated endotoxic effects, the supposition that the entire LPS molecule is necessary for activity has ultimately been shown to be incorrect. The following hypothesis offered by Westphal in 1954 has instead been shown to be more accurate: “…for induction of these effects the polysaccharide portion is dispensable and…the lipid A component represents the active center responsible for the endotoxic properties of lipopolysaccharide.”\textsuperscript{37} Unfortunately, Westphal’s hypothesis went untested for three decades until the Shiba group reported the first total synthesis of Lipid A in 1985.\textsuperscript{38} Westphal necessarily collaborated with Shiba and quickly demonstrated that the endotoxic activity propagated by natural Lipid A could be replicated with synthetic Lipid A (\textbf{Table 2.2}).\textsuperscript{39} Thus, Lipid A not only serves as the region through which LPS is appended to the outer membrane, but also as the inducer of endotoxic activity within host organisms.

<table>
<thead>
<tr>
<th>Preparation</th>
<th>Lethality with</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.1 (\mu g)</td>
</tr>
<tr>
<td>Synthetic Lipid A</td>
<td>10/10</td>
</tr>
<tr>
<td>Natural Lipid A</td>
<td>10/10</td>
</tr>
</tbody>
</table>

\textbf{Table 2.2.} Comparision of natural and synthetic Lipid A lethality in rats.

36
2.2.2 Core oligosaccharide structure and function

Appended to the Lipid A region of LPS is the Core oligosaccharide moiety, a region which is further subdivided into the outer (red) and inner (blue) core oligosaccharides. (Figure 2.5). The inner core oligosaccharide is highly conserved and typically consists of 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo) and L-glycero-D-mannoheptose residues. The outer core region, however, exhibits slightly increased structural diversity. For example, E. coli can possess one of five different core structures in which the inner region is constant while the outer region contains various arrangements of hexose and hexosamine residues. Such variability is not surprising, however, as the outer core oligosaccharides are more readily exposed to the selective pressures of the external environment.

While the Core region may initially appear to merely serve as an adaptor between the membrane-anchoring Lipid A and OPS moieties, recent work has demonstrated that this moiety is critical to bacterial survival. As shown in Figure 2.2, the aforementioned L-glycero-D-mannoheptose residues are phosphorylated. These phosphoryl groups provide a means by which adjacent LPS molecules can interact through a divalent cation bridge. Such interactions effectively create a layer around the bacterium which impedes the penetration of cationic peptides as well as hydrophobic compounds such as antibiotics. In the presence of EDTA, however, the integrity of the outer membrane is severely compromised.
2.2.3 O-polysaccharide structure and function

Of the three regions which comprise LPS, OPSs by far exhibit the greatest degree of structural diversity. Shown below is a representative repeating unit (RU) for a given OPS (Figure 2.6). While the example given here is a tetrasaccharide, RUs typically vary in size from two to five monosaccharides. Additionally, the order and identity of each of these monosaccharides can be varied. Variations can also occur through changes in the stereochemistry of the glycosidic linkage as well as which hydroxyl groups are united to form this linkage. Nonstoichiometric modifications (methylation, sulfation, etc.) are furthermore quite prevalent.\textsuperscript{13} As with the diversity seen in the outer core oligosaccharides, this multitude of possible polysaccharide structures arises from external selective pressures in the environment to which this region of LPS is exposed.
Given the presence of such adaptable structural units, the aforementioned studies of Westphal, which implicated only the Lipid A region of LPS in initiation of endotoxic activity, caused considerable confusion as to the biological significance of the O-polysaccharide region. Recent efforts using an *E. coli* mutant strain devoid of the OPS region, however, have shown that OPSs offer resistance to bile and other hydrophobic substances as well as bactericidal agents utilized by the innate immune system such as cationic peptides. OPSs have furthermore been shown to resist killing by the serum-mediated response and phagocytosis.\(^{42,43}\)

### 2.3 Polysaccharide biosynthetic pathways

With a detailed understanding of the biosynthesis of Core oligosaccharide-Lipid A having emerged largely through the efforts of Raetz, increasing efforts have similarly been put forth in the past several decades to better understand the biosynthesis of this immunologically relevant OPS region. Three major models have been proposed based on the genetic investigations of an increasing number of sequenced gene clusters dedicated
These models include the wzy-dependent pathway, ABC-transporter-dependent pathway and synthase-dependent pathway. Among them, the wzy-dependent pathway accounts for the synthesis of the vast majority of heteropolysaccharides. Specifically, the pathway is initiated by the addition of a sugar phosphate to Undecaprenyl-phosphate (Und-P) to form monosaccharide-PP-Und, a precursor for oligosaccharide RUs assembled by a series of dedicated glycosyltransferases on the cytoplasmic face of the inner membrane (Figure 2.7). The flippase Wzx then translocates these RU-PP-Und substrates to the periplasmic side of the inner membrane, where the putative polymerase Wzy catalyzes polymerization from the non-reducing end via a block-transfer mechanism. Finally, the Wzz protein is postulated to act as a chain length regulator such that certain polysaccharide chain lengths are generated for each bacterial strain.

Figure 2.7. wzy-dependent pathway of O-polysaccharide biosynthesis.
2.4 In vitro reconstitution of \textit{wzy}-dependent polysaccharide biosynthesis

Despite extensive genetic studies, this proposed \textit{wzy}-dependent model has yet to be biochemically verified \textit{in vitro}. This has hampered the mechanistic studies of polysaccharide polymerization as well as investigations concerning export and assembly of polysaccharides into the outer membrane. Specifically, the following difficulties have impeded efforts towards reconstituting the \textit{wzy}-dependent pathway \textit{in vitro}: 1) structurally-defined repeating unit substrates are not readily available, 2) over-expression and isolation of the purified Wzy membrane protein has been unsuccessful and 3) development of a straightforward and sensitive \textit{in vitro} assay to monitor polymerization has proven challenging.

Chemical approaches using homogenously synthesized substrates and purified enzymes offer a powerful complement to genetic studies, and can provide unambiguous biochemical evidence so as to help delineate the molecular details of the pathway. Such an approach to the reconstitution of polysaccharide biosynthesis \textit{in vitro} would minimally require two components, the putative polymerase Wzy and the naturally occurring RU-PP-Und substrate.

2.4.1 Expression and purification of Wzy

Initial attempts at expressing Wzy with a relatively simple expression system proved fruitless as only high molecular weight aggregates were discernible by Western-blot analysis (Figure 2.8). Given that Wzy is heavily imbedded in the inner membrane (10 predicted transmembrane domains), such aggregation can readily be attributed to
intermolecular interactions between these hydrophobic domains. Accordingly, researchers in the Wang laboratory developed a co-expression system which coupled expression of the \textit{wzy}-bearing plasmid with another plasmid containing the GroEL/ES chaperone system. This system affords an environment for hydrophobic proteins to correctly fold in the absence of problematic aggregative interactions. SDS-PAGE and Western blot analysis both indicated the successful expression and purification of Wzy using this system (\textbf{Figure 2.8}). Although a considerable amount of aggregate remained, the presence of monomeric and dimeric bands suggested that this expression system had likely produced functional Wzy. The identity of the proteins within these bands was furthermore confirmed to be Wzy through capillary-liquid chromatography-nanospray tandem mass spectrometry.
Figure 2.8. Expression of polymerase Wzy. a) wzy-containing plasmid and resulting Western blot. b) Co-expression system and resulting SDS-PAGE and Western blot.

2.4.2 Synthesis of RU-PP-Und

With the purified polymerase now in hand, attention was focused on the preparation of the RU-PP-Und substrate. Previous work in the Wang laboratory had involved detailed characterization of the four glycosyltransferases involved in elaboration
of initially formed \(N\)-Acetyl-D-galactosamine-PP-Und (GalNAc-PP-Und) to RU-PP-Und.\(^{51-54}\) Accordingly, the method of choice for preparing this molecule was postulated to be a chemoenzymatic route involving chemical synthesis of GalNAc-PP-Und followed by enzymatic elaboration to the full RU substrate.

Chemical synthesis of GalNAc-PP-Und was initiated by the preparation of GalNAc-1-phosphate\(^{55-60}\) and Und-P.\(^{61}\) Specifically, \(N\)-Acetyl-D-galactosamine (2.1) was first peracetylated to give 2.2. Selective anomeric deprotection with hydrazine acetate then afforded 2.3 (Scheme 2.1). Treatment of 2.3 with dibenyl \(N,N\)-diisopropylphosphoramidite in the presence of the acid/nucleophilic catalyst tetrazole effected phosphite formation. Oxidation to the corresponding phosphate 2.4 then proceeded via introduction of \(m\)CPBA. Final hydrogenolysis of 2.4 yielded the desired peracetylated GalNAc-1-phosphate as the diisopropylethylammonium salt (2.5).

**Scheme 2.1.** Synthesis of peracetylated GalNAc-1-phosphate (2.5).
Similar phosphoramidite chemistry was employed to furnish Und-P (2.8) (Scheme 2.2). However, the specific reagent had to necessarily be altered to the dicyanoethoxy $N,N$-diisopropylphosphoramidite equivalent so as to preserve the isoprene-based olefins in the deprotection step. Use of this phosphoramidite allowed deprotection to safely proceed via a base-mediated reaction.

Scheme 2.2. Synthesis of Undecaprenyl-phosphate (2.8).

Coupling of 2.5 and 2.8 made use of a classical pyrophosphate bond forming strategy originally developed by Danilov.62 Specifically, 2.5 was first activated for coupling through treatment with 1,1'-carbonyldiimidazole (CDI) (Scheme 2.3). Activation proceeds via initial displacement of the imidazolate from CDI upon attack at the carbonyl by a phosphate oxygen. Subsequent attack at the phosphorus atom in 2.10 by the previously released imidazolate affords the activated intermediate 2.12 with concomitant release of carbon dioxide and another equivalent of the imidazolate. Addition of the activated species as a solution in THF to 2.8 yielded the desired coupled
adduct (2.14) after a period of three days. Base-catalyzed deprotection of 2.14 to give 2.15 then completed the chemical portion of the synthesis (Scheme 2.4).

Scheme 2.3. Mechanism of formation of the coupled adduct peracetylated GalNAc-PP-Und (2.14). a) Activation of peracetylated GalNAc-1-phosphate (2.5). b) Coupling of activated peracetylated GalNAc-1-phosphate (2.12) and Und-Phos (2.8).
Enzymatic elaboration of 2.15 to RU-PP-Und commenced with WbnH-mediated formation of the disaccharide GalNAc-α1,3-GalNAc-PP-Und (2.16) (Scheme 2.5). WbnJ was then utilized to append galactose to the disaccharide so as to form Gal-β1,3-GalNAc-α1,3-GalNAc-PP-Und (2.17). Generation of Fuc-α1,2-Gal-β1,3-GalNAc-α1,3-GalNAc-PP-Und (2.18) with WbnK followed by use of WbnI yielded Gal-α1,3-(Fuc-α1,2)-Gal-β1,3-GalNAc-α1,3-GalNAc-PP-Und (2.19), the desired RU-PP-Und substrate. It should be noted that this enzymatic synthesis had to proceed in a sequential, multi-pot fashion. Attempts at producing RU-PP-Und in one pot failed, likely due to inhibition by the byproduct UDP.
2.4.3 Polymerization assays

While typical LPS molecules can readily be detected following gel electrophoresis through the use of silver staining techniques, 2.19 is devoid of the Core oligosaccharide-Lipid A moiety which renders such techniques effective. Accordingly, a radioactivity-based approach was pursued in which 2.19 was radiolabeled through use of UDP-[3H] Gal in the Wbnl catalyzed reaction. (Scheme 2.5).

The radio-labeled substrate 2.19 was incubated for 4 h at room temperature with purified Wzy. After quenching with cold methanol, the reaction mixture was spotted on chromatography paper. The control reaction, which contained Ni-NTA purified product from expression of an empty pBAD vector, showed migration of radioactivity to an R_f of 0.75 on the paper. However, in the reaction containing Wzy, a significant amount of
radioactivity appeared at a position with a considerably lower R_f value (Figure 2.9 (a)), indicating the formation of higher molecular weight compounds. At this point, SDS-PAGE and autoradiography were used to verify that the observed higher molecular weight compounds were in fact polysaccharides. The radioactive pattern observed using this technique consisted of a ladder-like pattern, thus indicating that polysaccharides with variable numbers of RUs had been produced (Figure 2.9 (b)). Up to 13 RUs can be resolved under the optimized experimental conditions. It is thus evident that Wzy is both necessary and sufficient to catalyze polymerization of RUs.

![Figure 2.9](image)

**Figure 2.9.** Analysis of Wzy activity using radiolabeled RU-PP-Und (2.19). a) Paper chromatography assay; b) SDS-PAGE assay, Left (+ Wzy); Right (-Wzy).

### 2.4.4 Regeneration of modal distributions

While a broad distribution of polysaccharide lengths was generated in the above *in vitro* system, cell surface polysaccharides generally have a strain-specific pattern of
chain lengths termed modality that has been shown to be closely associated with the virulence of pathogens. Genetic studies suggest that modality is conferred by the Wzz protein, possibly through concerted action with Wzy\textsuperscript{63-65}. To investigate this hypothesis \textit{in vitro}, purified Wzz proteins obtained from two different \textit{E. coli} O86 strains (B7 and H2) were introduced into the polymerization system. O-polysaccharides from these two strains share the same branched pentasaccharide repeating unit but differ in the anomeric configuration of the linkage\textsuperscript{53}. Furthermore, although Wzz proteins from these two strains share 90% amino acid identity, the LPS modality of the two strains is markedly different (Figure 2.10 (a)). B7 derived LPS exhibits a short modal distribution (2 to 5 units) (Lane 1), while H2 derived LPS exhibits an intermediate modal distribution (9 to 16 units). The same LPS modality (9-16 units) was observed for \textit{E. coli} O86:B7 \textit{Δwzz} complemented with plasmid pTR-102 (with \textit{wzz}_{H2} gene) (Lane 3).\textsuperscript{51} To demonstrate its effect on chain length, purified Wzz\textsubscript{B7} or Wzz\textsubscript{H2} was added to the previous polymerization reaction mixture containing the radiolabeled RU-PP-Und substrate \textit{2.19} and Wzy. The reaction was analyzed using SDS-PAGE/autoradiography. The radioactive signal on the gel indicated the presence of a very short modal distribution (3-6 units) in the case of Wzz\textsubscript{B7} (Figure 2.10 (b), Lane 1), whereas that for Wzz\textsubscript{H2} showed intermediate modality (10-14 units) (Figure 2.10 (b), Lane 3). These patterns are similar to those observed for LPS from \textit{E. coli} O86:B7 and O86:B7 \textit{Δwzz} complemented with \textit{wzz}_{H2}, respectively. The LPS of \textit{E. coli} O86:B7 \textit{Δwzz} was also provided (Figure 2.10 (a), Lane 2) and shows a similar modality as the \textit{in vitro} Wzy reaction without Wzz (Figure 2.10 (b), Lane 2).
Figure 2.10. Confirmation of the regulatory role of Wzz in polysaccharide biosynthesis.  
a) Silver stained LPS isolates from bacteria: Lane 1-LPS profile of *E. coli* O86:B7; Lane 2-LPS profile of *E. coli* O86:B7 Δwzz; Lane 3-LPS profile of *E. coli* O86:B7 Δwzz complemented with wzzH2 (plasmid pTR-102). b) Analysis of the polymerization reaction via SDS-PAGE and auto-radiography: Lane 1-Control reaction containing Ni-affinity elution from expression of empty pBAD vector; Lane 2-Reaction containing Ni-affinity elution of Wzy; Lane 3-Polymerization reaction containing Wzy and B7 strain Wzz; Lane 4-Polymerization reaction containing Wzy and H2 strain Wzz.

In contrast, incubation of Wzz with the repeating unit substrate in the absence of Wzy did not produce any detectable polysaccharides (Figure 2.11). Collectively, this result and those presented in Figure 2.10 demonstrate that our Wzy-Wzz system is sufficient to generate polysaccharides with specific chain lengths. This observation also opens up an exciting possibility of modulating polysaccharide chain lengths *in vitro*. 
2.5 Analysis of Wzy substrate specificity

While a system was in place for the *in vitro* biosynthesis of bacterial polysaccharides, several important questions lingered regarding the molecular details of the *wzy*-dependent pathway. For example, how is the activity of Wzy functionally regulated? Does Wzy interact with Wzz as suggested by previous *in vivo* studies? If so, is the interaction required for chain length regulation? First, however, given that Wzy was demonstrated to be both necessary and sufficient for polymerization, the polymerase itself was investigated in more detail through a substrate-specificity study.

2.5.1 Relevance of the pentasaccharide moiety

The naturally occurring substrate for Wzy, 2.19, can be considered to be composed of two distinct moieties, the RU and the Polyprenol. Of these two, the RU can
be much more easily modified through genetic modification of the assembly pathway. Accordingly, researchers in the Wang laboratory performed a preliminary evaluation of RU specificity in which WbnI-catalyzed addition of the terminal galactose residue was prevented through functional inactivation of the \( wbnI \) gene.\(^{66} \) Such gene inactivation effectively caused the bacteria to produce tetrasaccharide RUs rather than the natural pentasaccharide substrates. Extraction and analysis of the LPS, however, indicated that while these repeating units are formed, relatively little polymerization occurred (Figure 2.12, Lanes 3 and 4). This result suggests that Wzy possesses rather limited tolerance for variations in the RU moiety of the substrate. Furthermore, it should be noted that a possible alternative explanation involving poor translocation by Wzx is most likely invalid. Previous studies have indicated that only the first sugar residue is important for recognition and thus successful translocation.\(^{67} \)

**Figure 2.12.** Evaluation of RU specificity. Lane 1: *E. coli* O86 wild type; Lanes 2 and 3: \( \Delta wbnI \) mutant strains; Lane 4: \( \Delta wbnI \) mutant complemented with pUC-\( wbnI \) plasmid.
### 2.5.2 Relevance of the polyprenol moiety

To examine the Polyprenol substrate specificity of Wzy, a small library was proposed (Table 2.3). The first three polyprenols, Und, Heptaprenol (Hept) and Pentaprenol (Pent), were selected to evaluate the effect of length on polymerization. As can be seen, the number of isoprene units decreases from 11 to 7 to 5 along this series. However, it was also noted that as the length decreases in this series, the number of cis-olefins also decreases. Given that these olefins are adjacent to the hydroxy terminus and thus in close proximity to the bond breaking event during polymerization, it is reasonable to believe that they may serve as an essential recognition element. A derivative of Pent, cis-Pentaprenol (cis-Pent) was therefore proposed in which the number of cis-olefins is increased to four with the total number of isoprene units remaining constant. The hypothesis for this design was that if cis-olefins are an important recognition element, then the degree of polymerization should increase from Pent to cis-Pent-based substrates. In contrast, if double bond geometry is irrelevant, then no increase should be observed. The polyprenol Solanesol (Solan) was also proposed to support the results of the cis-Pent experiment. Finally, while bacteria utilize isoprene-based polyprenols in which each unit is unsaturated, eukaryotes instead utilize a polyprenol class generally referred to as Dolichols.\(^{68}\) Dolichols are almost structurally identical to bacterial polyprenols with the exception of the \(\alpha\)-isoprene unit which is saturated. Another derivative of Pent, Monosaturated-Pentaprenol (MS-Pent), was thus proposed to consider the importance of having an unsaturated \(\alpha\)-isoprene unit.
<table>
<thead>
<tr>
<th>Lipid</th>
<th>Geometry</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undecaprenol</td>
<td>7 cis, 3 trans</td>
<td><img src="2.19" alt="structure" /></td>
</tr>
<tr>
<td>Heptaprenol</td>
<td>4 cis, 2 trans</td>
<td><img src="2.20" alt="structure" /></td>
</tr>
<tr>
<td>Pentaprenol</td>
<td>2 cis, 2 trans</td>
<td><img src="2.21" alt="structure" /></td>
</tr>
<tr>
<td>cis-Pentaprenol</td>
<td>4 cis, 0 trans</td>
<td><img src="2.22" alt="structure" /></td>
</tr>
<tr>
<td>Solanesol</td>
<td>0 cis, 8 trans</td>
<td><img src="2.23" alt="structure" /></td>
</tr>
<tr>
<td>MS-Pentaprenol</td>
<td>1 cis, 2 trans</td>
<td><img src="2.24" alt="structure" /></td>
</tr>
</tbody>
</table>

**Table 2.3.** Proposed library of polyprenols for Wzy substrate specificity study.

While the RU domain can be easily modified through genetic manipulation, a similar strategy does not exist for modification of the Polyprenol moiety. cis-Polyprenols are furthermore not readily commercially available. Accordingly, attention was focused on identifying the best chemical route to synthesize these molecules.
2.5.2.1 Synthetic approaches to polyprenols

2.5.2.1.1 Solution phase synthetic methods

By the early 1980s, several reports had been published detailing the synthesis of all-trans-polyprenols.\textsuperscript{69-73} However, approaches for preparing polyprenols containing trisubstituted cis-olefins were quite limited. The first significant breakthrough in this area originated with two similar reports from Still and Sato in which various $\alpha$-alkoxyketones (\textsuperscript{2.25} and \textsuperscript{2.28}) were shown to yield cis-olefins (\textsuperscript{2.27} and \textsuperscript{2.30}) in high selectivity via Wittig olefination with unstabilized ylides (\textsuperscript{2.26} and \textsuperscript{2.29}) (Scheme 2.6, Top: Sato, Bottom: Still).\textsuperscript{74,75} Subsequent investigations demonstrated that the protected alcohol product could be converted to the corresponding bromide and coupled with additional cis-olefins which had been derivatized as $\alpha$-sulfonyl variants (Scheme 2.7).\textsuperscript{75}

![Scheme 2.6. Stereoselective formation of trisubstituted cis-olefins via Wittig olefination.](image-url)
With this impetus, Sato and coworkers went on to detail the preparation of a bifunctional all-cis-diene (2.39) which can be utilized as a common intermediate for both the bromide and sulfonyle coupling partners described above (Scheme 2.8). Preparation of this substrate requires two key steps, the first being regioselective epoxidation of the terminal olefin of benzylated Nerol (2.35). Secondly, the aforementioned Wittig olefination is employed to yield 2.39 with stereoselectivities $\geq 95:5$ (Z:E) depending on reaction conditions.

**Scheme 2.7.** Sato method for the assembly of trisubstituted cis-olefin polyenes.

**Scheme 2.8.** Preparation of the key bifunctional diene for the Sato method.
With this substrate in hand, acid mediated removal of the tetrahydropyranyl protecting group or reductive removal of the benzyl group affords an alcohol which can be converted to the corresponding chloride (2.41) or sulfonyl (2.42) derivative (Scheme 2.9). Base-mediated coupling of 2.41 and 2.42 yields the bifunctional substrate 2.43 which is similar to the key diene 2.39 but is now elongated to an all-cis-tetraene (Scheme 2.10). A pathway similar to that described above can then be repeated to generate longer all-cis substrates. Coupling of chloride 2.44 derived from 2.43 with a sulfone prepared from a polyprenol (i.e., Farnesol or Nerol) (2.45) yields the fully assembled polyprenol skeleton 2.46. Final reduction then affords the desired polyprenol, in this case Hept (2.47). Through this strategy, Sato and coworkers were able to prepare a variety of cis-polyprenols including (Z,Z)-farnesol,76 (Z,Z,Z)-tetraprenol,76 (Z,Z,E,E)-pentaprenol,77 betulaprenols-6, -7, -8 and -9,77 (Z,Z,Z,Z,Z,Z,Z,Z,E,E)-undecaprenol,78 and one polyprenol derived from a citronellol-based building block, (S)-(−)-dolichol-20.79

Scheme 2.9. Diversification of Sato’s key diene intermediate.
Scheme 2.10. Polyprenol formation via the Sato method. a) Generation of longer all-cis building blocks through sulfone-halide coupling; b) Attachment of the terminal polyprenol-derived tail also through sulfone-halide coupling.

While the Sato method has emerged as the protocol of choice for preparation of cis-olefin containing polyprenols, four additional approaches merit discussion. The first, initially reported by Moiseenkov, makes use of a C-5 homologation in which one isoprene unit is introduced in each coupling reaction. Specifically, hydroxysulfonamide 2.51, which is derived from isoprene (2.48), is converted to the
dianion and subsequently treated with 2.52 to afford δ-sulfonamide polyprenol 2.53 in a manner reminiscent of the Sato method (Scheme 2.11). Final dissolving metal reduction affords the polyprenol as describe above. Further modification of this substrate to the corresponding bromide followed by introduction of 2.51 as the dianion enables continued elongation of the polyprenol substrate.

![Scheme 2.11. Moiseenkov sulfonamide approach to polyprenol synthesis.](image)

While such iterative introduction of only one isoprene unit hinders the practicality of this method in comparison to the Sato approach, this method affords the opportunity to synthesize various unnatural polyprenols. For example, use of the corresponding (E)-isomer of 2.51 facilitates introduction of trans-olefins in place of the cis-counterparts. Polyprenols with alternating olefin stereochemistry can thus be readily prepared. As with the Sato method, this approach can also be extended to the preparation of dolichol-type lipids through use of the saturated variant of the initial hydroxysulfonamide.
The next approach, also introduced by Moiseenkov, utilizes a notably different way to unite the isoprene-based building blocks (Scheme 2.12). In this case, the terminal fragment is derived from Farnesyl phenyl sulfide 2.54. 2.54 is first alkylated and then desulfurized to afford 2.56 upon acidic workup. Conversion to the corresponding imine and subsequent treatment with base yields 2.57 to which is added 2.58. Such benzyloxyaldehydes are readily obtained through early termination of the pathway Sato used to generate 2.39 or via selective ozonolysis of 2.35. Generation of α,β-unsaturated aldehyde 2.59 upon acidic work-up is then followed by reduction to the methyl group using a multi-step reductive conversion outlined by Corey. Debenzylation and regeneration of the terminal imine afford a substrate suitable for further iteration to longer polyprenols. As before, dolichols can also be prepared through use of a saturated variant of the starting benzyloxyaldehyde.

Scheme 2.12. Moiseenkov imine approach to polyprenol synthesis.
More recently, Gibbs described a method which enabled the preparation of the various isomers of farnesol. The sequence is initiated through generation of β-ketoesters from Geranyl or Neryl bromide (2.61) and the dianion of ethyl acetoacetate (2.62) (Scheme 2.13, only Neryl bromide is shown). The subsequent key step involves stereoselective formation of a triflate derived from the β-ketoesters. Gibbs found that varying the solvent from THF to DMF yields predominantly the (Z)-triflate 2.64 or the (E)-triflate 2.67, respectively. Thus, all four farnesol precursors can be readily prepared from the Geranyl and Neryl bromide-derived β-ketoesters. Subsequent Pd-catalyzed coupling with tetramethylytin and DIBAL-mediated reduction affords all four farnesol stereoisomers. This strategy could similarly be used for iterative preparation of longer polyprenols. However, the addition of only a single isoprene unit during each iteration makes this approach less advantageous for such a goal. Much like the first method presented by Moiseenkov, this method is nonetheless ideal for preparation of unnatural polyprenols such as those with alternating olefin stereochemistry.
Scheme 2.13. Gibbs synthesis of farnesol stereoisomers.

Weimer has since modified the Gibbs strategy, offering an approach that does not require the use of tin reagents. Specifically, 2.61 is first converted to Nerylacetone (2.72) using a three step sequence of alkylation with ethyl acetoacetate, saponification, and decarboxylation (Scheme 2.14). A modified Wittig procedure based on the work of Schlosser and Corey is then employed to yield 2.75, the (Z,Z)-isomer (Scheme 2.15).
Preparation of the (E,Z)-isomer 2.78 is accomplished through a Horner-Wadsworth-Emmons reaction also involving 2.72. Use of Geranyl bromide similarly affords the (Z,E)- and (E,E)-isomers.

**Scheme 2.14.** Preparation of Nerylacetone (2.72) as a common starting material for the Weimer polyprenol synthetic strategy.
Scheme 2.15. Weimer synthesis of farnesol stereoisomers.

2.5.2.1.2 Solid phase synthetic methods

Given the iterative nature of Sato’s methodology, Cheng accordingly adapted it so as to provide a solid-phase approach to polyprenol and dolichol substrates. To accomplish this aim, terminal isoprene-based fragment 2.80 is first appended to a benzenesulfinate resin (2.79) (Scheme 2.16). The methodology from this point forward is essentially identical to that employed by Sato with the exception that the diene building
block 2.82 is protected as the TBDPS-ether rather than the benzyl ether. The polymer-bound polyprenol fragment is elongated through the same type of base-mediated halide/sulfone couplings as described above. Dolichols are similarly prepared through use of the corresponding saturated building block. After sufficient elongation, free polyprenols are then furnished via Pd-catalyzed reduction. It should also be noted that longer building blocks can be utilized to reduce the number of on-resin steps. However, the preparation of such building blocks in the solution phase is considerably more involved than the synthesis of their shorter counterparts.

**Scheme 2.16.** Solid phase adaptation of the Sato approach to polyprenol synthesis.
2.5.2.2 Synthesis of a polyprenol library

While several of the above methods afford the opportunity to construct exotic lipids, the polyprenol library proposed in Table 2.3 is primarily composed of obvious structural analogs of the natural polyprenol Und. Construction of this library therefore demanded a strategy in which cis-olefin fragments could be quickly assembled and then ligated to a terminal fragment. Given this need, the Sato method offered the best choice. Accordingly, synthesis of the previously reported bifunctional building block was initiated with benzylation of the commercially available polyprenol Nerol (2.34) to afford 2.35 (Scheme 2.17). Subsequent regioselective epoxidation of the terminal olefin followed by oxidative cleavage afforded the crude aldehyde. Direct reduction of this crude product furnished the alcohol 2.86 in 62% yield over three steps. Conversion of 2.86 to the corresponding iodide then facilitated generation of the Wittig salt 2.37 (88%, over steps). Final Wittig olefination using α-hydroxyacetone (2.38) gave the desired building block 2.39 in 82% yield.
Scheme 2.17. Synthesis of the bifunctional diene building block 2.39.

As described above, the benefit to using 2.39 is the ability to generate two dissimilar diene fragments from a common starting material. For example, the benzyl group was first reductively removed to afford alcohol 2.87 (Scheme 2.18). Mesylation of 2.87 followed by nucleophilic displacement of the mesylate yielded the alkyl chloride (2.88). Further nucleophilic displacement with the sodium salt of $p$-toluenesulfinic acid provided access to sulfone 2.42 in 66% yield over the two previous steps. Alternatively, the tetrahydropyranyl (THP) protecting group was removed under acidic conditions, giving 2.89 in 98% yield (Scheme 2.19). 2.89 was then used to form the corresponding alkyl chloride 2.41 as described above.
Having formed these differentially functionalized dienes, the backbones of the various desired polyprenols were assembled. In the case of Hept, these two dienes were first coupled via initial treatment of 2.42 with n-butyllithium to afford the carbanion (Scheme 2.20). Introduction of 2.41 led to formation of tetraene 2.43 through nucleophilic displacement. This all cis-tetraene now possessed the same bifunctional appearance as previously seen with the diene building block. One could thus consider similarly manipulating each end of the molecule to form two differentially functionalized tetraenes. These fragments could then be coupled to build longer all cis-olefin chains.
Scheme 2.20. Synthesis and elaboration of the cis-tetraene 2.43.

For the purposes of this study, however, tetraene 2.43 was of a suitable length such that the terminal trans-olefin containing tail could be appended to this molecule. First, the THP protecting group was removed, after which alcohol 2.90 was converted to the corresponding alkyl chloride (2.44) as previously discussed (Scheme 2.20). The carbanion of Farnesyl sulfone 2.45 (prepared in one step from commercially available Farnesyl bromide) was then generated, after which 2.44 was introduced to form the coupled adduct 2.46 in 90% yield (Scheme 2.21). Final reductive desulfonylation/debenzylation gave the desired polyprenol Hept (2.47). Unfortunately, the yield of this final transformation was considerably low. The primary reason for such a poor yield arises from the generation of regioisomers during the course of the reaction (~1.5%). While these isomers were effectively separated from 2.47, such removal required conversion of the alcohol to the acetate, purification over silica gel imbedded with silver nitrate, and final deprotection to afford pure Hept. Thus, the presence of the regioisomers
demanded the use of these additional transformations and surprisingly difficult form of
column chromatography, ultimately leading to considerable losses of material.

Scheme 2.21. Completion of the Heptaprenol (2.47) synthesis.

In contrast to Hept, preparation of Pentaprenol did not require use of tetraene
2.44. Rather, one of the diene derivatives (2.41) was instead utilized in the coupling
reaction with 2.45 (Scheme 2.22) to give 2.91. Reductive removal of the benzyl and
sulfonyl groups afforded Pent (2.92) without the presence of contaminating regioisomers.
Accordingly, the yield of this step was significantly improved to 74%.
Scheme 2.22. Synthesis of Pentaprenol (2.92).

Whereas 2.47 terminates in the triene derived from Farnesyl sulfone, 2.44 can instead be coupled with allyl sulfone 2.93 to form the all cis-variant of Pent (cis-Pent, 2.95). Specifically, 2.93 was first deprotonated, after which 2.44 was added to generate the coupled adduct (2.94) (Scheme 2.23). Final reduction yielded cis-Pent (2.95). As with Hept, the presence of multiple sulfonyl groups led to the formation of regioisomers (~1.5%) and thus necessitated the use of the purification strategy outlined above.

Scheme 2.23. Synthesis of cis-Pentaprenol (2.95).
While the above polyprenols were all prepared from the Nerol-derived building blocks, synthesis of the last desired polyprenol (MS-Pent, 2.105) demanded the use of a building block with a saturated α-isoprene unit. Such a building block was readily accessed through use of the commercially available polyprenol Citronellol (2.96). Specifically, 2.96 was substituted into the same sequence detailed above for the preparation of 2.39, ultimately leading to formation of the bifunctional building block 2.101 (Scheme 2.24).

![Scheme 2.24. Synthesis of the α-saturated diene building block 2.101.](image)

Removal of the THP protecting group from 2.101 to form 2.102 enabled generation of bromide 2.103, once again through initial mesylation followed by addition of a lithium halide salt (Scheme 2.25). Coupling of 2.103 and 2.45 yielded the backbone
of MS-Pent (2.104). Subsequent reductive debenzylolation/desulfonylation afforded the desired polypropenol 2.105 (Scheme 2.26). As with Pent, no significant amounts of regioisomers were observed, thus enhancing the yield for this transformation.

Scheme 2.25. Preparation of coupling partner 2.103 for the MS-Pentaprenol synthesis.


2.5.2.3 Interpretation of polymerization data

Having prepared the desired polypropenol library, these molecules were elaborated into the corresponding RU-PP-Polyprenol substrates. As can be seen (Table 2.4, entries 1-3),
the amount of substrate successfully polymerized decreased as the length of the lipid was shortened. However, given that the number of cis-double bonds, which are adjacent to the hydroxy terminus, also decreases as lipid length decreases, one must consider that the double bond geometry rather than lipid length is the primary source of decreased polymerization levels. Donor 2.22 was thus evaluated. The degree of polymerization observed, in comparison to that seen with 2.21 was notably greater (Table 2.4, entry 4). This indicates that double bond geometry is a more important factor than lipid length in determining polymerization efficiency. To further verify this hypothesis, donor 2.23 (Table 2.4, entry 5) was tested as it is only two isoprene units shorter than the natural lipid, Und, while containing only trans double bonds. No detectable polymerization was observed with this donor. Finally, the Dolichol-like substrate 2.24 (Table 2.4, entry 6) was evaluated to discern the importance of unsaturation in the \( \alpha \)-isoprene unit. The lack of activity observed using this donor verifies the expected preference for bacterial-based Und-containing substrates rather than eukaryotic-based Dolichol-containing substrates.
<table>
<thead>
<tr>
<th>Lipid</th>
<th>Geometry</th>
<th>Structure</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undecaprenol</td>
<td>7 cis, 3 trans</td>
<td><img src="image1" alt="Undecaprenol Structure" /></td>
<td>51 ± 5</td>
</tr>
<tr>
<td>Heptaprenol</td>
<td>4 cis, 2 trans</td>
<td><img src="image2" alt="Heptaprenol Structure" /></td>
<td>34 ± 4</td>
</tr>
<tr>
<td>Pentaprenol</td>
<td>2 cis, 2 trans</td>
<td><img src="image3" alt="Pentaprenol Structure" /></td>
<td>15 ± 2</td>
</tr>
<tr>
<td>cis-Pentaprenol</td>
<td>4 cis, 0 trans</td>
<td><img src="image4" alt="cis-Pentaprenol Structure" /></td>
<td>29 ± 4</td>
</tr>
<tr>
<td>Solanesol</td>
<td>0 cis, 8 trans</td>
<td><img src="image5" alt="Solanesol Structure" /></td>
<td>0</td>
</tr>
<tr>
<td>MS-Pentaprenol</td>
<td>1 cis, 2 trans</td>
<td><img src="image6" alt="MS-Pentaprenol Structure" /></td>
<td>3 ± 1</td>
</tr>
</tbody>
</table>

Table 2.4. Comparison of Wzy activity among RU-PP-Polyprenol analogs (“Percent” indicates the amount of RU substrate successfully polymerized).
2.6 Conclusions

Bacteria possess polysaccharides with enormous structural complexity. Efforts to understand their biosynthesis and export have been driven by their importance in host-bacterium interface biology and, in particular, their strong association with bacterial pathogenicity. While such efforts have revealed the \textit{wzy}-dependent pathway to arguably be the most widely distributed polysaccharide biosynthetic pathway in nature, convincing evidence that Wzy functions as a polymerase has not previously been presented. Rather, assignment of Wzy as the putative polymerase was based on the generation of a semi-rough LPS phenotype upon genetic depletion of the \textit{wzy} gene. Such assignment can be problematic, however, and warrants caution as deletion of the \textit{wzy} gene can potentially disrupt critical interactions and propagate changes in the overall biosynthetic pathway. Definitive assignment of polymerase activity has instead awaited a well-defined, reconstituted system that consists of a minimal yet sufficient number of components. The work described above demonstrates that purified Wzy is both necessary and sufficient to induce polymerization of chemically synthesized oligosaccharide substrates \textit{in vitro}.

While not necessary to effect polymerization, Wzz is also proposed to play an interesting, but as of yet mechanistically undefined, role in imparting a defined modal distribution to the polysaccharides. Generation of modal distributions consistent with those of \textit{E. coli} O86:H2 and O86:B7 upon introduction of the respective Wzz protein into the above \textit{in vitro} system further confirms the role of Wzz in this process. Determination of precisely how Wzz imparts modality, however, remains an even more intriguing topic. Currently, two possible mechanisms have been proposed. One possibility suggests that Wzz functions as a molecular timer, modulating Wzy activity between two states that
favor either the polymerization or ligation reaction. The alternative model, however, suggests that Wzz functions as a molecular ruler, regulating the oligomerization state of a complex between Wzy and the ligase WaaL to define the modality. Structural studies on periplasmic domain and full length Wzz seem to favor the molecular ruler model. Small-angle X-ray scattering data, crystallographic evidence and the most recent electron microscopy results have all suggested that Wzz proteins form oligomers. Nonetheless, no direct evidence illustrates how these oligomers modulate the chain length of O-polysaccharides. Several pieces of evidence do suggest the presence of multi-protein complexes in the wzy-dependent pathway. In addition, it has been shown in vivo that chain length modality is conferred prior to the ligation reaction, suggesting that regulation occurs on the polymerizing substrate. The in vitro reconstitution system described in this chapter further confirms such a hypothesis and suggests the existence of a meaningful interaction among Wzy, Wzz and polyisoprenoid lipid-linked polymers. Work is ongoing in the Wang lab to identify Wzy-Wzz interactions using a pull down assay, and to modulate chain length modality by varying the stoichiometry of Wzy-Wzz in the system. It is therefore expected that the newly developed in vitro system will provide further insights into the mechanism of action of Wzz.

Apart from Wzy and Wzz, also central to our studies was Und-P, the predominant lipid carrier in bacteria for the assembly of a diverse range of polysaccharides. Unlike typical glycosyltransferase-catalyzed reactions, where the diphospho-lipid moiety is part of the acceptor structure, in the Wzy-catalyzed polymerization, the diphospho-lipid moiety also serves as part of the donor substrate and is directly involved in bond breaking in the reaction. Therefore, it was speculated that varying the lipid structure might have a
more dramatic effect in this instance. This notion garnered support from lipid specificity studies of TGase and PglB, which both use similar diphospho-lipid sugar moieties as donors.\textsuperscript{57,98-99} The above specificity studies with Wzy revealed a preference similar to those observed in those studies, suggesting a defined mode of recognition for Wzy towards lipid structures in the sugar donor. Additionally, it is now apparent that much simpler lipid based donors can be utilized in future studies of this system, thus remedying problems associated with the use of the highly hydrophobic lipid, Und.

Since the wzy-dependent pathway had only been studied \textit{in vivo} prior to the studies described herein, many molecular details are yet to be revealed. Thus, this system has supplied the groundwork necessary for probing such details of Wzy polymerization and Wzz chain length regulation and represents one step forward towards understanding and re-defining the mechanisms of polysaccharide biosynthesis in general.
REFERENCES


3.1 Introduction

Proteins represent one of the major classes of biomacromolecules present in living organisms. Furthermore, these molecules are responsible for many of the diverse properties and processes which are encoded within the genome. Given that the size of the human genome is ~30,000 genes, one may be led to conclude that the proteome is of a similar size. However, a much higher degree of complexity is actually observed with the estimated number of different protein forms ranging from 300,000-3,000,000.\(^1\) Such heightened diversification results largely from the following three mechanisms: 1) Utilization of alternate mRNA promoter sequences, 2) Alternative splicing of mRNA and 3) Covalent modification of proteins.\(^1-3\)

Of these mechanisms, the covalent modification of proteins in particular allows for an incredible degree of diversity as evidenced by the plethora of modifications which have been identified (Table 3.1).\(^4\) Glycosylation (Entry 6), or the enzymatic attachment of a carbohydrate structure, represents arguably the most common and complex form of co-/posttranslational modification. For example, a variety of sophisticated pathways have already been described which lead to the generation of a multitude of glycoproteins.
differing in both the protein substrate as well as the composition of the transferred carbohydrate moiety. Common to all of these pathways, however, is the formation of a bond between the carbohydrate structure and an amino acid of the substrate.\(^5\)

<table>
<thead>
<tr>
<th>Modification</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acylation</td>
<td>Lys acetylation in the core histone proteins</td>
</tr>
<tr>
<td>Aminocarboxypropylation</td>
<td>Conversion of His to Diphthamide in eEF-2</td>
</tr>
<tr>
<td>ADP-Ribosylation</td>
<td>Alteration of Diphthamide in eEF-2</td>
</tr>
<tr>
<td>Carboxylation</td>
<td>Attachment of CO(_2) to Glu in blood coagulation proteins</td>
</tr>
<tr>
<td>Disulfide Bond Formation</td>
<td>Linking of Cys residues as proteins pass through the secretory pathway</td>
</tr>
<tr>
<td>Glycosylation</td>
<td>Trp mannosylation in plasma membrane proteins</td>
</tr>
<tr>
<td>Hydroxylation</td>
<td>Pro hydroxylation in collagen</td>
</tr>
<tr>
<td>Isomerization</td>
<td>Conversion of Asp to isoAsp</td>
</tr>
<tr>
<td>Methylation</td>
<td>Lys methylation in the core histone proteins</td>
</tr>
<tr>
<td>Nitration</td>
<td>Tyr ortho-nitration during the inflammatory response</td>
</tr>
<tr>
<td>Oxidation</td>
<td>Conversion of Met to the corresponding sulfoxide</td>
</tr>
<tr>
<td>Phosphopantetheinylation</td>
<td>Ser modification in fatty acid synthases</td>
</tr>
<tr>
<td>Polyglutamylation</td>
<td>Glu modification in tubulin</td>
</tr>
<tr>
<td>Polyglycination</td>
<td>Glu modification in tubulin</td>
</tr>
<tr>
<td>Prenylation</td>
<td>C(_{15}) farnesylation of the RAS family proteins</td>
</tr>
<tr>
<td>Splicing</td>
<td>Excision of inteins</td>
</tr>
<tr>
<td>Sulfation</td>
<td>Tyr sulfation during CCR5 receptor maturation</td>
</tr>
<tr>
<td>Transglutamination</td>
<td>Cross-linking of Gln residues</td>
</tr>
</tbody>
</table>

**Table 3.1.** Covalent modifications of proteins.\(^4\)
As shown below (Figure 3.1), there are numerous ways in which the carbohydrate-amino acid linkage can be established. Variability can exist, for example, in the type of bond formed (N-C, C-C, P-C or O-C), amino acid utilized (Asn, Arg, Trp, Ser, Thr, Tyr, Hyp or Hyl), and identity of the carbohydrate residue. While not specifically shown here, the configuration of the anomeric linkage also provides another point of variation.5

Figure 3.1. Diagrammatic representation of the five distinct types of sugar–peptide bonds that have currently been identified. (Spiro, R.G. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 12(4), 43R-56R (2002), by permission of Oxford University Press)
3.2 N-Glycosylation pathway overview

Despite the numerous potential linkages described above, by far the most common in eukaryotes is the Asn-carbohydrate linkage formed via the process known as N-glycosylation. In fact, estimates based on the Swiss-Prot database indicate that more than one half of all eukaryotic proteins are N-glycosylated. N-glycosylation is initiated at the cytoplasmic face of the endoplasmic reticulum (ER) membrane (Figure 3.2). Here, Alg7, an N-acetylglucosamine-1-phosphate transferase, catalyzes the transfer of GlcNAc-1-phosphate to the Dolichyl-phosphate lipid carrier which is embedded in the ER membrane. As discussed in Chapter 2, Dolichol type lipids consist of a series of isoprene units (typically 17-21) arranged in an acyclic, head-to-tail fashion in which the α-isoprene unit is saturated.

Further elaboration of the aforementioned GalNAc-PP-Dolichyl product by glycosyltransferases including Alg13/14, Alg1, Alg2 and Alg11, yields a Dolichyl-linked heptasaccharide (Man₅GlcNAc₂). Rft1p, an ATP-independent flippase, then translocates this lipid-linked glycan to the luminal face of the ER membrane. Additional sugar residues are added to the heptasaccharide, forming a tetradecasaccharide which is primed for transfer to a protein. It should be noted that the glycosyltransferases implicated in the addition of sugars within the ER lumen (Alg3, Alg9, Alg12, Alg6, Alg8 and Alg10) utilize Dolichyl-phosphate-linked sugars rather than the commonly encountered sugar-nucleotide donors such as UDP-GlcNAc or GDP-Man.

The next step in N-glycosylation involves the transfer of the synthesized Dolichyl-linked tetradecasaccharide to a protein substrate. Specifically, the oligosaccharide is transferred and appended to an Asn residue within a consensus N-
glycosylation amino acid sequence (Asn-X-Ser/Thr, where X ≠ Proline). This transfer is catalyzed by a membrane-associated multimeric enzyme complex known as an Oligosaccharytransferase, or OTase. In the case of S.cerevisiae, OTases are composed of at least eight subunits. Of these subunits, Stt3p is highly conserved among all eukaryotes and has been shown to be directly involved in catalysis.

Figure 3.2. The dolichol pathway of N-linked protein glycosylation in S. cerevisiae. (Eranthie Weerapana, et. al. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. Glycobiology 16(6), 91R-101R (2006), by permission of Oxford University Press)
While this glycosylation process is frequently denoted as a posttranslational modification, attachment of the tetradecasaccharide more typically occurs on the unfolded protein substrate as it enters the ER lumen. Eukaryotic glycosylation may thus more appropriately be referred to as a co-translational modification.\(^6\) In fact, the presence of this oligosaccharide helps regulate the correct folding of the polypeptide to which it is appended. Specifically, the glycan serves as necessary recognition element for entrance into the calnexin-calreticulin cycle.\(^{25-27}\) Further processing must first occur, however, in which the conserved tetradecasaccharide \((\text{Glc}_3\text{Man}_9\text{GlcNAc}_2)\) is trimmed by Glucosidase I and Glucosidase II to form a dodecasaccharide \((\text{GlcMan}_9\text{GlcNAc}_2)\) (Figure 3.3). Such processing of the oligosaccharide generates a substrate which is recognized by membrane-associated lectin calnexin (CNX) or its soluble counterpart calreticulin (CRT), both of which have chaperone functions. Each of these lectins is further complexed with the co-chaperone ERp57.\(^6\) Following correct folding, the glycoprotein is released from the CNX/CRT-ERp57 complex and trimmed once more by Glucosidase II. However, if the protein is misfolded, UDP-glucose:glycoprotein glucosyltransferase catalyzes addition of a glucose residue back onto the oligosaccharide. This re-glucosylation promotes complexation once again with CNX/CRT-ERp57, thus forcing incorrectly folded proteins to remain in the CNX-CRT cycle.\(^{25}\) It should be mentioned, however, that this cycle is not strictly necessary for correct protein folding and may slow down the process in some cases. The efficiency of generating a correctly folded protein is nevertheless much greater. Once folded into the native conformation, the glycoprotein is then transported to the Golgi for final processing and elaboration.\(^6,25\)
While the previous two figures only highlight the *S. cerevisiae* protein glycosylation pathway, the associated oligosaccharides involved are largely conserved among eukaryotes. This fact is somewhat surprising given the level of heterogeneity which has been demonstrated to occur in cell surface glycans. However, as mentioned above, processing does not terminate in the ER but rather is continued in the medial stacks of the Golgi. This additional processing involves an assortment of glycosidases and glycosyltransferases which generate the aforementioned structural heterogeneity typically seen with cell surface glycans. For example, mammalian systems readily incorporate *N*-acetylglucosamine, *N*-acetylgalactosamine, fucose, galactose, and most
notably N-acetyllneuraminic acid, while yeast typically generate highly branched mannose derivatives. Once these highly modified oligosaccharides are formed, the glycoproteins are at last primed for secretion or incorporation into the cell membrane.

3.3 Biological significance of eukaryotic N-glycosylation

The adornment of eukaryotic proteins with various oligosaccharides serves several biologically significant roles, a few of which will now be considered. First, as mentioned above, appended N-glycans have an indirect effect on protein folding by providing a recognition element for entry into the CNX-CRT cycle. The chaperone activities of CNX and CRT then act in concert with the co-chaperone ERp57 and UDP-glucose:glycoprotein glucosyltransferase to ensure that glycoproteins are properly folded prior to being transported to the Golgi.

Secondly, N-glycosylation exerts a direct effect on protein folding. Specifically, folding studies have demonstrated that N-glycans promote the proper folding of glycosylated proteins in comparison to the non-glycosylated counterparts. It should be noted, however, that although these glycans are important for the folding process, they do not possess a critical role for the maintenance of the properly folded protein.

Thirdly, N-glycans play a vital role in the proper functioning of lysosomes, hydrolase-containing organelles which are important for the degradation of various materials. Specifically, the lysosomal membrane proteins LAMP-1 and LAMP-2 are much more densely glycosylated than typical glycoproteins. Up to 20 N-glycans can be found on the luminal surface of these proteins, thus effectively creating a
“carbohydrate shell” around this portion of the protein. Given that lysosomes contain enzymes responsible for degradation, this “shell” serves to protect LAMP-1 and -2 from the action of these enzymes, specifically proteases. The protective nature of these pendant $N$-glycans enables LAMP-1 and -2 to not only help form a stable membrane, but more importantly maintain a controlled environment such that eukaryotes can selectively degrade various substances.

In addition to these functions, $N$-glycosylation also serves several immunologically significant roles in eukaryotes. One such example involves the recognition of antigens presented on the surface of dendritic cells. This recognition process is initiated when T-cells begin scanning the surface of dendritic cells to search for the specific peptide-MHC complex which they can recognize. To facilitate this sampling process, several types of intercellular interactions exist between the T-cells and dendritic cells. These interactions allow the two cells to remain in close enough proximity to find the desired complex. One of the more notable interactions involves a high-mannose $N$-linked oligosaccharide known as intracellular adhesion molecule 3, or ICAM-3 which is displayed on the surface of T-cells (Figure 3.4). As the immune cells scan the surface of dendritic cells in search of a MHC-presented antigen, ICAM-3 specifically interacts with a transmembrane protein known as dendritic cell-specific ICAM-3-grabbing non-integrin (DC-SIGN). This interaction is only temporary, however, unless the desired peptide-MHC complex is located. Interaction with this complex facilitates several ICAM-3/DC-SIGN binding events to not only promote adherence of the T-cell, but also to initiate signaling cascades which will trigger lymphocyte activation.

3.4 Development of an *in vitro* assay for eukaryotic oligosaccharide transfer

While considerable knowledge has been gained over the past two decades with regards to the formation of N-glycans, the presence of an OTase complex in *S. cerevisiae* has made detailed evaluation of the key step in glycoprotein formation quite challenging.
Attention has consequently been shifted to lower eukaryotes such as *Giardia* and kinetoplastids (*Trypanosoma, Leishmania*) as they have been proposed to contain single subunit OTases.\(^{37-40}\) Recently, this hypothesis was confirmed through evaluation of a series of *Leishmania major* STT3 paralogs in Δstt3 *S. cerevisiae*.\(^{41}\) Three of the four tested proteins were found to complement the *stt3* deletion. More importantly, these paralogs were demonstrated to exist as single subunits rather than being integrated into the *S. cerevisiae* OTase complex. Such a result indicates that the OTase of *L. major* is composed of only a single subunit, rendering it an ideal candidate for detailed studies of the glycosylation step. Accordingly, an *in vitro* system has been designed through collaborative efforts of the Wang and Aebi groups enabling protein glycosylation and the OTase itself to be investigated in more detail.

### 3.4.1 Design of the *in vitro* glycosylation system

In order to realize glycosylation *in vitro*, the system had to be minimally composed of the following two components: 1) Functional Stt3d and 2) Lipid linked oligosaccharide substrate. Preparation of the active enzyme will not be considered in detail here due to privacy limitations of the aforementioned collaboration. Nonetheless, the system will now be described in general terms.

*L. major* Stt3d was isolated from a Δstt3 *S. cerevisiae* strain. Specifically, the Δstt3 strain was prepared by first introducing the functional yeast *stt3* gene into a heterozygous *STT3* diploid strain via a complementing plasmid. Following sporulation, haploids containing both the deletion and plasmid were identified. A plasmid harboring
the *L. major* *stt3* gene was then introduced into the haploid strain. Selection for successful transformation was accomplished using Leu’ Ura’ minimal media, after which further selection was performed on FOA-media to isolate cells that had lost the plasmid containing the yeast *stt3* gene. This engineered yeast strain was then cultured, and microsomal extracts were obtained to serve as the source of *L. major* Stt3d.

With the desired OTase in hand, attention was focused on designing a substrate which could be recognized by Stt3d but was much less complex than the natural donor. It should be noted that such donors in trypanosomatid parasites such as *Leishmania* are somewhat different than the oligosaccharide discussed above for *S. cerevisiae* and most other eukaryotes. These donors are composed of only Man$_6$GlcNAc$_2$ and Man$_7$GlcNAc$_2$ and thus represent simpler versions of the previously described tetradecasaccharide. However, such octa- and nonasaccharides are still very difficult synthetic targets. This difficulty is further heightened when the full donor substrate is considered as it is also contains the highly hydrophobic and poorly available Dolichol moiety.

Given these issues of substrate complexity, a minimal oligosaccharide donor was proposed which contained only the conserved GlcNAc$_2$ disaccharide and the dolichol-type lipid MS-Pent which was described in Chapter 2. Accordingly, synthesis of the OTase substrate commenced with selective anomeric deprotection of commercially available Chitobiose Octaacetate. The crude product was then treated with dibenyl *N*,*N*-diisopropylphosphoramidite in the presence of tetrazole to effect phosphite formation. Oxidation to the corresponding phosphate 3.2 was subsequently realized via introduction of *m*CPBA. Final hydrogenolysis of 3.2 yielded the desired peracetylated Chitobiose-1-phosphate as the diisopropylethylammonium salt (3.3).
Scheme 3.1. Synthesis of the OTase substrate Chitobiose-PP-MS-Pent (3.4).

Coupling of peracetylated Chitobiose-1-phosphate and MS-Pent-phosphate (MS-Pent-P) made use of the strategy previously outlined in Chapter 2. Specifically, 3.3 was first activated for coupling through treatment with 1,1'-carbonyldiimidazole (CDI) (Scheme 3.1). Addition of this activated species as a solution in THF to MS-Pent-P yielded the coupled adduct after three days. Base-catalyzed deprotection then afforded Chitobiose-PP-MS-Pent (3.4) to complete the synthesis.
3.4.2 *In vitro* reconstitution of oligosaccharide transfer

Having successfully prepared both components of the proposed *in vitro* system, collaborative efforts with the Aebi group have since enabled a detailed characterization of Stt3d. The first goal, however, was to simply demonstrate that glycosylation could be achieved *in vitro* using the prepared minimal substrate 3.4. The microsomal extract containing Stt3d was thus combined with the sugar donor and a peptide acceptor (*Figure 3.5*). Additionally, a concentrated (4x) extract was also evaluated. As can be seen, glycosylation was only apparent when using the concentrated extract in the absence of EDTA. The necessary exclusion of EDTA indicates that the activity of Stt3d is likely cation dependent.

<table>
<thead>
<tr>
<th></th>
<th>+</th>
<th>+</th>
<th>+</th>
<th>-</th>
<th>+*</th>
<th>+*</th>
<th>+*</th>
<th>+*</th>
<th>-*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsomal extract</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+*</td>
<td>+*</td>
<td>+*</td>
<td>+*</td>
<td>-*</td>
</tr>
<tr>
<td>Peptide acceptor</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Synthetic sugar</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>EDTA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*concentrated solubilized microsomal (4X)*

*Figure 3.5.* Demonstration of *in vitro* Stt3d-mediated peptide glycosylation.

To confirm the glycosylation results presented above, the sample was subjected to analysis by mass spectrometry (*Figure 3.6*). Both the peptide acceptor (m/z = 1122.5) and glycosylated peptide (m/z = 1528.6) are clearly present in the reaction mixture, thus
confirming that this \textit{in vitro} system is functional. Furthermore, glycosylation was also apparent when the $\beta$–$N$-Acetylhexosaminidase treated substrate was utilized. This enzyme cleaves the terminal GlcNAc residue from the disaccharide, generating a MS-Pent-PP-linked monosaccharide. Taken together, these results indicate that Stt3d not only recognizes the conserved disaccharide core of \textit{L. major} \textit{N}-glycans, but also the simplest carbohydrate motif possible, a single GlcNAc residue.

\textbf{Figure 3.6.} Confirmation of Stt3d-mediated glycosylation. Lanes labeled with “+” were pretreated with $\beta$–$N$-Acetylhexosaminidase which cleaves the terminal GlcNAc residue.
With a functional system now in place, several features of this enzymatic reaction were evaluated. First, the effect of sugar donor concentration was investigated. An expected direct correlation was observed as the concentration was increased from 0 μM to 1 μM to 10 μM and finally to 100 μM (Figure 3.7). The most significant extent of glycosylation was observed with the concentrated microsomal extract in which 49% of the peptide was found to be glycosylated at a sugar donor concentration of 100 μM. At higher concentrations, however, both of the extracts showed a decrease in percent glycosylation, thus indicating inhibition of the OTase.

**Figure 3.7.** Effect of Chitobiose-PP-MS-Pent concentration on Stt3d-mediated peptide glycosylation.
A general evaluation of the kinetic parameters for this reaction was then performed (Figure 3.8). As can be seen, the obtained Lineweaver-Burk plots are relatively similar for both of the extracts. The OTase of the concentrated microsomal extract exhibits a maximum velocity (0.47 pmol/min) which is slightly greater than that observed with the OTase of the nonconcentrated extract (0.32 pmol/min). However, this maximum velocity is observed at a higher sugar donor concentration, thus leading to a higher Km value for the concentrated extract. A much broader time course for the reaction was also considered (Figure 3.9). The increase in percent glycosylation observed at each time point gradually decreased throughout the experiment. Such a decrease likely resulted from the associated decrease in sugar donor concentration as the experiment proceeded.

Figure 3.8. Kinetics of Stt3d-mediated peptide glycosylation at different Chitobiose-PP-MS-Pent concentrations.
Next, the effect of pH on Stt3d-mediated peptide glycosylation was considered. The amount of glycosylation observed was relatively consistent when the pH was held at or near neutral (Figure 3.10). Optimal glycosylation, under the conditions tested, occurred at a slightly acidic pH (pH = 6.5). As the reaction solution became more basic, however, the degree of glycosylation decreased.
A dependency on divalent cations was also observed for this enzyme (Figure 3.11). For example, when the reaction mixture contained EDTA, which effectively removes divalent cations from solution, no glycosylation was apparent. Use of Mn$^{2+}$, however, resulted in significant glycosylation. It also appears that Mn$^{2+}$ is the preferred divalent cation, as only minor levels of glycosylation were observed for the other cations with Ca$^{2+}$ and Mg$^{2+}$ exhibiting the highest levels among that group.
Two substances, which could theoretically result in OTase inhibition, (free sugars and UDP) were investigated next. In both cases, no significant effects were observed as the concentration of free sugar/UDP was increased up to 1 mM (Figures 3.12 and 3.13). Only when the concentration was increased to 10 mM, in the case of UDP, was OTase inhibition observed.
Figure 3.12. Effect of free monosaccharides on Stt3d-mediated peptide glycosylation.

Figure 3.13. Effect of free UDP on Stt3d-mediated peptide glycosylation.
Finally, the impact of temperature on glycosylation was investigated. No effects were observed as the temperature was increased from 20 to 24 °C (Figure 3.14). However, as the temperature was increased further, a gradual decline in glycosylation was noted. Optimal OTase activity thus occurs at or near room temperature.

![Figure 3.14. Effect of temperature on Stt3d-mediated peptide glycosylation.](image)

3.5 Conclusions

Considerable insight has been gained into the pathway for eukaryotic N-glycosylation over the course of the last two decades. Nevertheless, a detailed understanding of the key step in this pathway, protein glycosylation, has not been achieved. This limited knowledge has resulted primarily from the existence of OTase complexes in eukaryotes such as *S. cerevisiae*. The discovery of single subunit OTases in
lower eukaryotes such as *L. major*, however, has created an opportunity to study the glycosylation step in detail. As shown above, significant progress has now been made through the development of an *in vitro* system, which utilizes a synthetic sugar donor, Chitobiose-PP-MS-Pent (3,4), and concentrated microsomal extract containing the *L. major* OTase Stt3d. Use of this system allowed for a general characterization of the OTase-containing extract including dependency on sugar donor concentration, pH, divalent cations, free sugars, free UDP, and temperature. Some general kinetic parameters and the time course for the reaction were also obtained. All of this information will undoubtedly aid in the design of future experiments to study this enzyme. Revelation of the mechanistic details, however, awaits an OTase that has been purified to apparent homogeneity, a process which is currently in progress.
REFERENCES


4.1 Introduction

Well into the 20th century, glycosylation was presumed to be a eukaryotic-specific protein modification. This dogma was ultimately found to be incorrect, however, upon the discovery that an archeon, *Halobacterium salinarium*, possessed surface layer (S-layer) glycoproteins.1 Since this seminal discovery, glycoproteins have also been found throughout the bacterial domain.2 Such modified proteins are frequently found in surface appendages which include flagella and pili.3-4 Given the roles of these appendages in virulence, it is highly likely that glycoproteins are involved in bacterial pathogenicity to some extent.5 Such functional assignments remain speculative, however, as the functions of glycoproteins in general are poorly understood. Accordingly, efforts to better understand the process of bacterial protein glycosylation and its associated pathogenic implications have become quite prevalent.
4.2 Bacterial protein glycosylation

Once considered rare, glycosylation of bacterial proteins is now known to commonly occur in both Gram-negative and Gram-positive bacteria.5-17 Similar to eukaryotes, both N-glycosylation and O-glycosylation pathways have been identified and characterized.12,17-19 In general, the structure of these pathways is quite similar to that found in eukaryotes; however, as will be seen in the succeeding sections, several differences exist.

4.2.1 N-Glycosylation

While highly prevalent in eukaryotes, only one N-glycosylation pathway has been identified and well characterized in bacteria. This system occurs in *Campylobacter jejuni*, a human-gut pathogen which has been implicated as the primary causative agent of gastroenteritis.20-25 From a general perspective, this pathway is reminiscent of that described for eukaryotes in Chapter 3 in which a tetradecasaccharide is sequentially assembled on a Dolichyl-diphosphate lipid carrier and then co-translationally transferred to a target polypeptide. In the case of *C. jejuni*, however, the oligosaccharide is a conserved heptasaccharide with the structure GalNAc-α1,4-GalNAc-α1,4-(Glcβ1,3)-GalNAc-α1,4-GalNAc-α1,4-GalNAc-α1,3-Bac-β1, where Bac is bacillosamine (2,4-diacetamido-2,4,6-trideoxyglucopyranose). Rather than being synthesized on a Dolichyl-based carrier, the heptasaccharide is furthermore synthesized on an Und-diphosphate lipid carrier on the cytoplasmic face of the membrane (Figure 4.1).26 This glycoconjugate is subsequently flipped across the membrane to the periplasmic face, at which point the
heptasaccharide moiety is transferred to the Asn residue of a Asp/Glu-X-Asn-Z-Ser/Thr (X, Z ≠ Proline) consensus sequence. In contrast to what is observed with eukaryotes, such transfer occurs post-translationally.²⁷


Genes implicated in the synthesis and transfer of this oligosaccharide were found to originate in a single gene locus termed the “pgl gene cluster” for protein glycosylation locus.²¹ This cluster contains the genes *pglA* through *pglG*, all of which act to either 1) Synthesize the uncommon sugar bacillosamine, 2) Assemble the heptasaccharide, or 3) Transfer the oligosaccharide to the target protein (**Figure 4.2**).²⁷ The key step in this pathway, glycosylation of the target protein, is accomplished with the product of the *pglB*
gene, an OTase appropriately referred to as PglB. Although the natural substrate of this OTase is the heptasaccharide mentioned above, studies have revealed that PglB can also recognize various other carbohydrate moieties which contain a hexosamine at the reducing end. The specificity for the polyprenol moiety, however, is much more stringent and follows the same general trend as that outlined for Wzy (Chapter 2).


### 4.2.2 \(O\)-Glycosylation

While bacterial \(N\)-glycosylation has been extensively studied in \(C.\ jejunii\), it should be reiterated that the above pathway represents the only example which has been definitively identified and well characterized in bacteria. \(O\)-glycosylation, in contrast, is frequently found in both Gram-positive and Gram-negative bacteria. This type of glycosylation involves the transfer of a carbohydrate structure to a Ser or Thr residue, neither of which is present in a well-defined consensus sequence. Furthermore, such glycosylation is completed through one of three distinct pathways.
One such pathway is found in the aforementioned Gram-negative bacteria *C. jejuni*. Specifically, detailed NMR and MS studies have revealed that *C. jejuni* flagella are highly glycosylated (19 Ser/Thr residues), accounting for the 10% higher than expected mass of flagellar protein FlaA. Unlike the *N*-glycosylation pathways discussed above, mechanistic analysis suggests that this pathway occurs through direct transfer of carbohydrates from the corresponding sugar nucleotide donors. This particular example thus circumvents the need for initial lipid-linked oligosaccharide assembly.

A second pathway, described in *P. aeruginosa* 1244, similarly avoids the need for *O*-glycosylation specific oligosaccharide assembly. Rather than simply adding sugars directly from the sugar nucleotide donors, however, this pathway exploits the polysaccharide biosynthetic machinery. Specifically, the OTase PilO effectively robs the LPS biosynthetic pathway of assembled RU-PP-Und building blocks. These RUs are then transferred to the Ser148 residue of pilin.

The final pathway to be considered, which will be the focus of this chapter, makes use of a *pgl* gene cluster comparable to that described for bacterial *N*-glycosylation. Accordingly, the pathways for assembly and transfer of the oligosaccharide are quite similar (*Figure 4.3*). The oligosaccharide is assembled via sequential enzymatic glycosylation of an Und-P lipid carrier. In the case of *N. meningitidis* strain C311#3, this oligosaccharide is an unusual trisaccharide (Gal-β1,4-Gal-α1,3-DATDH) which incorporates the uncommon sugar 2,4-diacetamido-2,4,6-trideoxyhexose, or DATDH, at the reducing end. Once assembled on the cytoplasmic face of the membrane, the lipid-linked oligosaccharide is flipped to the periplasmic face. An OTase, PglL, then catalyzes transfer of the oligosaccharide to a pilin Ser residue.
Until recently, functional assignment of PglL as the OTase for this pathway was based purely on homology of the \textit{pglL} gene to that of another \(O\)-glycosylation OTase, PilO.\textsuperscript{12} This activity was confirmed first through analysis of pilin isolated from wildtype and \(\Delta pglL\) strains. Pilin taken from the knockout strain exhibited increased electrophoretic mobility, a property which is indicative of a loss of glycosylation.\textsuperscript{19} Secondly, PglL and pilin were coexpressed in \textit{E. coli} along with a \(\Delta pglB pgl\) gene cluster from \textit{C. jejuni}. Analysis of isolated pilin via Western blot and mass spectrometry verified that pilin had in fact been glycosylated.\textsuperscript{18} This result provided definitive evidence that PglL acts as an OTase.

\textbf{Figure 4.3.} Bacterial \(O\)-glycosylation in \textit{N. meningitidis} strain C311#3.

\textbf{4.3 Substrate specificity of PglL}

During the course of the aforementioned experiment in which \(O\)-glycosylation was reconstituted in \textit{E. coli}, it was demonstrated that \textit{E. coli O7} RU5 or full length \(O\)-
polysaccharides could be appended to pilin. These preliminary results suggested that this OTase may be able to recognize a broad range of substrates. To investigate this idea, a collaboration has been established between the Feldman and Wang groups so as to discern PglL specificity for both the carbohydrate and polyprenol moieties.

4.3.1 Relevance of the carbohydrate moiety

As described above, preliminary studies of the O-glycosylation OTase PglL indicated that strict substrate specificity may not exist. Several other poly-/oligosaccharide donors were accordingly incorporated into the E. coli glycosylation system (Figure 4.4). Western blot and mass spectrometry analysis of the resultant pilin confirmed that an incredible degree of flexibility exists in PglL-mediated carbohydrate recognition. For example, PglL is able to recognize substrates containing a hexosamine (GlcNAc, GalNAc, FucNAc, Bac) or hexose (Gal) at the reducing end. This result indicates that the 2-acetamindo group is not required for PglL-mediated recognition. Additionally, the second sugar in the chain can vary not only in identity but also the type of linkage it shares with the reducing end sugar, including which hydroxyl groups are attached and the associated stereochemistry. Branched and linear substrates are also suitable for recognition. Such branching involves different types of linkages (1,2 or 1,3) and stereochemistry (α or β). PglL therefore does not appear to discriminate among different potential poly-/oligosaccharide donors.
4.3.2 Relevance of the polyprenol moiety

Given the lack of specificity for the carbohydrate moiety, it was presumed that PglL must be specific for the polyprenol region of the substrate. Such investigation of polyprenol substrate specificity required the development of an *in vitro* glycosylation assay in which different lipid-linked oligosaccharides could be utilized. Given the Wang group’s success in preparing such substrates with variable lipid moieties, the Wang and Feldman groups were able to develop this type of *in vitro* system.
4.3.2.1 Preliminary in vitro results

The impetus for this collaboration did not actually begin with a goal of characterizing polyprenol specificity; rather, it was borne out of a request by the Feldman group for a lipid-linked sugar donor that would serve as a negative control for an in vitro reaction. Given the polyprenol specificity of PglB, a lipid of short length and (E)-olefin geometry appeared to be the ideal choice. Accordingly, Farnesol was selected as the polyprenol given that it is only three isoprene units in length and is composed of two (E)-olefins in addition to the terminal olefin.

RU-PP-Farnesol was synthesized from Farnesol using the strategy outlined in Chapter 2. The substrate was then screened against Und-linked C. jejuni heptasaccharide (Figure 4.2). As shown, both the heptasaccharide donor and RU donor were acceptable substrates for PglL (Figure 4.5). PglL thus appeared to have no specificity for either polyprenol length or olefin geometry. This result was incredibly surprising given the typical patterns of polyprenol specificity observed with enzymes such as PglB and Wzy.

To provide a more detailed look at this apparent lack of lipid specificity, a series of RU-PP-Lipid (4.1-4, 2.24) analogs were proposed to evaluate several structural features (Table 4.1). Specifically, the first three polyprenols, Geranylgeraniol, Farnesol and Geraniol were selected to evaluate the effect of length on OTase activity. The fourth polyprenol, cis-Farnesol (cis-Farn), was chosen to confirm whether or not olefin geometry was an essential recognition element. The hypothesis for this design was once again that if the cis-olefins are an important recognition element, then the degree of polymerization should increase from Farnesol to cis-Farnesol based substrates. In contrast, if double bond geometry is irrelevant, then no increase should be observed.
Finally, the relevance of an unsaturated α–isoprene unit was to be evaluated with the same MS-Pent derived donor described in Chapter 2.

Figure 4.5. Preliminary investigation of PglL polyrenol specificity (Adapted from Faridmoayer, A., Fentabil, M.A., Haurat, M.F., Yi, W., Woodward, R., Wang, P.G. & Feldman, M.F. Extreme Substrate Promiscuity of the Neisseria Oligosaccharyl Transferase Involved in Protein O-Glycosylation. J. Biol. Chem. 283, 34596-34604 (2008)).
<table>
<thead>
<tr>
<th>Lipid</th>
<th>Geometry</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geranylgeraniol</td>
<td>0 cis, 3 trans</td>
<td><img src="4.1" alt="Structure" /></td>
</tr>
<tr>
<td>Farnesol</td>
<td>0 cis, 2 trans</td>
<td><img src="4.2" alt="Structure" /></td>
</tr>
<tr>
<td>Geraniol</td>
<td>0 cis, 1 trans</td>
<td><img src="4.3" alt="Structure" /></td>
</tr>
<tr>
<td>cis-Farnesol</td>
<td>2 cis, 0 trans</td>
<td><img src="4.4" alt="Structure" /></td>
</tr>
<tr>
<td>MS-Pentaprenol</td>
<td>1 cis, 2 trans</td>
<td><img src="2.24" alt="Structure" /></td>
</tr>
</tbody>
</table>

Table 4.1. Proposed library of polyprenols for PglL substrate specificity study.

4.3.2.2 Synthesis of RU-PP-Lipid analog library

Of the proposed polyprenols, only two needed to be synthesized, cis-Farn and MS-Pent. The remaining polyprenols were commercially available and thus alleviated a considerable synthetic burden. Given that the preparation of MS-Pent was previously described in Chapter 2, only the preparation of cis-Farn will be considered here.

As with the polyprenols of Chapter 2, the Sato Method was utilized to prepare the desired polyprenol, cis-Farn. This synthesis made use of building block 2.41 that was
also used in the preparation of Pent (2.47), cis-Pent (2.92) and Hept (2.95). Specifically, 2.93 was first deprotonated with $n$-BuLi to form the corresponding carbanion. Introduction of 2.41 led to formation of coupled triene 4.7 through nucleophilic displacement of the chloride. Reductive removal of the benzyl and sulfonyl groups then afforded cis-Farn (4.8).

\[ \begin{align*}
2.93 & \quad \text{[n-BuLi, THF, HMPA]} \quad -78\, ^\circ\text{C, then 2.41 in THF]} \quad 92\% \\
4.7 & \quad \text{[Li, EtN}^+\text{]$_2$, -78\, ^\circ\text{C]} \quad 54\% \\
4.8 & \\
\end{align*} \]

**Scheme 4.1.** Synthesis of cis-Farnesol (4.8).

### 4.3.2.3 Polyprenol specificity of PglL

With these substrates in hand, the Feldman group recently completed a preliminary specificity screen (Figure 4.6). While these initial results are only qualitative, it is apparent that PglL recognizes all of the tested substrates. Glycosylation appears to occur to different degrees, however, with Geraniol-based donor 4.3 being a considerably poorer substrate than 4.4, 2.24, and especially 4.1 and 4.2. The Feldman group is currently in the process of developing a suitable quantitative assay to more accurately assess the relationship between polyprenol structure and OTase activity.
Figure 4.6. Polyrenol substrate specificity screen for PglL. Lane 1: Pilin, PglL; Lane 2: Pilin, *C. jejuni* heptasaccharide; Lane 3: Geranylgeraniol-based donor (4.1); Lane 4: Farnesol-based donor (4.2); Lane 5: Geraniol-based donor (4.3); Lane 6: cis-Farnesol-based donor (4.4); Lane 7: MS-Pentaprenol-based donor (2.24); Lane 8: *C. jejuni* heptasaccharide (Lanes 3-8 also contain Pilin and PglL).

4.4 Conclusions

Glycosylation of bacterial appendages including flagella and pili has been postulated to contribute to pathogenicity. Understanding the process through which these oligosaccharides become attached to the target proteins has accordingly become an area of significant interest. One of the enzymes involved in this key reaction, PglL, mediates the transfer of a trisaccharide to a Ser residue of the pilin protein in *N. meningitidis* strain C311#3. Detailed analysis of this reaction has now revealed that PglL possesses incredibly relaxed substrate specificity with respect to the carbohydrate moiety. Furthermore, an *in vitro* glycosylation assay was developed in the Feldman lab which enabled demonstration that RU-PP-Farnesyl could serve as a suitable substrate. A series of RU-PP-Lipid analogs were then designed and have since helped demonstrate that PglL is also highly flexible with respect to the polyrenol moiety. A quantitative assay is currently under development to more accurately describe the observed specificity.
REFERENCES


CHAPTER 5
EXPERIMENTAL PROCEDURES

5.1 HDACi chemical synthesis and compound characterization

All solvents were dried with a solvent-purification system from Innovative Technology, Inc. All reagents were obtained from commercial sources and used without further purification. Analytical TLC was carried out on silica gel 60 F254 aluminum-backed plates (E. Merck). The 230–400 mesh size of the same absorbent was utilized for all chromatographic purifications with the exception of the click HDACi which required use of 200-400 mesh silica gel 60 RP-18 (from EMD\textsuperscript{TM}). \textsuperscript{1}H and \textsuperscript{13}C NMR spectra were recorded at the indicated field strengths. The high-resolution mass spectra were collected at The Ohio State University Campus Chemical Instrumentation Center. The purities of the click HDACi were determined to be $\geq 97.7\%$ according to HPLC.

General procedure for preparation of click HDACi (1.15-1.21). Triisopropylsilane (3 mmol, 0.6 mL), PMB protected substrate (1.54-1.60, 1 mmol) and TFA (2 mL) were added in sequence to a reaction vessel containing CH\textsubscript{2}Cl\textsubscript{2} (40 mL). The mixture was stirred at room temperature and monitored by TLC. Upon completion, the reaction mixture was diluted with CH\textsubscript{3}CN (100 mL) and then neutralized with DOWEX\textsuperscript{®}
MARATHON® WBA anion exchange resin (Aldrich). The organic solution was filtered, after which the resin was washed with CH$_3$CN (50 mL x 2). The combined organic solution was evaporated to provide a solid, which was purified to afford the final product.

1-Benzyl-1H-[1,2,3]triazole-4-carboxylic acid hydroxamide (1.15). The crude product was purified via C-18 reverse phase column chromatography (H$_2$O : CH$_3$CN = 6 : 1) to provide a white solid (16 mg, 47% yield). $^1$H NMR (500 MHz, DMF-d$_7$): $\delta$ 11.40-9.20 (bd s, 1H), 8.64 (s, 1H), 7.58-7.28 (m, 5H), 5.76 (s, 2H); $^{13}$C NMR (125 MHz, DMF-d$_7$): $\delta$ 158.9, 143.1, 137.0, 129.9, 129.3, 129.1, 127.1, 54.4; HRMS (ESI) calcd for C$_{10}$H$_{10}$N$_4$O$_2$Na [M+Na]$^+$ 241.0701, found 241.0697.

1-Phenethyl-1H-[1,2,3]triazole-4-carboxylic acid hydroxamide (1.16). The crude product was purified via C-18 reverse phase column chromatography (H$_2$O : CH$_3$CN = 6 : 1) to provide a white solid (21 mg, 51% yield). $^1$H NMR (500 MHz, DMF-d$_7$): $\delta$ 11.21 (s, 1H), 9.30 (s, 1H), 8.48 (s, 1H), 7.40-7.20 (m, 5H), 4.78 (t, $J$ = 7.3 Hz, 2H), 3.29 (t, $J$ = 7.3 Hz, 2H); $^{13}$C NMR (125 MHz, DMF-d$_7$): $\delta$ 158.3, 141.9, 138.0, 129.0, 128.7, 126.9, 126.2, 51.3, 36.2; HRMS (ESI) calcd for C$_{11}$H$_{12}$N$_4$O$_2$Na [M+Na]$^+$ 255.0858, found 255.0862.
1-(3-Phenyl-propyl)-1H-[1,2,3]triazole-4-carboxylic acid hydroxamide (1.17). The crude product was purified via C-18 reverse phase column chromatography (H2O : CH3CN = 6 : 1) to provide a white solid (31 mg, 74\% yield). 1H NMR (500 MHz, DMF-d7): δ 11.27 (s, 1H), 9.38 (s, 1H), 8.61 (s, 1H), 7.36-7.20 (m, 5H), 4.54 (t, J = 7.1 Hz, 2H), 2.66 (t, J = 7.7 Hz, 2H), 2.26 (tt, J1 = J2 = 7.1 Hz, 2H); 13C NMR (125 MHz, DMF-d7): δ 159.1, 142.8, 142.0, 129.4, 127.0, 126.9, 55.7, 50.5, 33.2, 32.7; HRMS (ESI) calcd for C12H14N4O2Na [M+Na]+ 269.1014, found 269.1014.

1-(2-Phenoxy-ethyl)-1H-[1,2,3]triazole-4-carboxylic acid hydroxamide (1.18). The crude product was purified via C-18 reverse phase column chromatography (H2O : CH3CN = 6 : 1) to provide a white solid (15 mg, 56\% yield). 1H NMR (500 MHz, DMF-d7): δ 10.60-9.90 (bd s, 1H), 8.65 (s, 1H), 7.34-7.28 (m, 2H), 7.00-6.94 (m, 2H), 4.95 (t, J = 5.1 Hz, 2H), 4.53 (t, J = 5.1, 2H); 13C NMR (125 MHz, DMF-d7): δ 159.4, 159.1, 143.0, 130.6, 127.6, 122.2, 115.7, 67.3, 50.7; HRMS (ESI) calcd for C11H12N4O2Na [M+Na]+ 271.0807, found 271.0807.

1-(2-Oxo-2-phenyl-ethyl)-1H-[1,2,3]triazole-4-carboxylic acid hydroxyamide (1.19). The crude product was purified via C-18 reverse phase
column chromatography (H₂O : CH₃CN = 7 : 1) to provide a white solid (56 mg, 72% yield). ¹H NMR (500 MHz, DMF-d₇): δ 11.40-9.20 (bd s, 2H), 8.62 (s, 1H), 8.16 (d, J = 7.3 Hz, 2H), 7.78 (t, J = 7.4 Hz, 1H), 7.65 (dd, J₁ = J₂ = 7.8 Hz, 2H), 6.38 (s, 2H); ¹³C NMR (125 MHz, DMF-d₇): δ 192.9, 159.1, 143.0, 135.5, 135.3, 130.1, 129.3, 128.9, 57.3; HRMS (ESI) calcd for C₁₁H₁₀N₄O₃Na [M+Na]⁺ 269.0651, found 269.0629.

1-Phenylcarbamoylmethyl-1H-[1,2,3]triazole-4-carboxylic acid hydroxyamide (1.20). The crude product was purified via C-18 reverse phase column chromatography (H₂O : CH₃CN = 7 : 1) to provide a white solid (45 mg, 55% yield). ¹H NMR (500 MHz, DMF-d₇): δ 10.77 (s, 1H), 10.50-10.00 (bd s, 1H), 8.65 (s, 1H), 7.70 (d, J = 7.4 Hz, 2H), 7.35 (dd, J₁ = J₂ = 7.9 Hz, 2H), 7.11 (t, J = 7.4 Hz, 1H), 5.57 (s, 2H); ¹³C NMR (125 MHz, DMF-d₇): δ 165.3, 159.1, 142.8, 140.1, 129.9, 128.8, 124.8, 120.4, 53.6; HRMS (ESI) calcd for C₁₁H₁₁N₅O₃Na [M+Na]⁺ 284.0760, found 284.0743.

1-(3-Phenyl-allyl)-1H-[1,2,3]triazole-4-carboxylic acid hydroxyamide (1.21). The crude product was purified via C-18 reverse phase column chromatography (H₂O : CH₃CN = 5 : 1) to provide a white solid (21 mg, 50% yield). ¹H NMR (500 MHz, DMF-d₇): δ 10.60-9.80 (bd s, 2H), 8.61 (s, 1H), 7.53 (d, J = 7.6 Hz, 2H), 7.38 (dd, J₁ = J₂ = 7.5 Hz, 2H), 7.31 (t, J = 7.3 Hz, 1H), 6.79 (d, J = 15.9, 1H), 6.62
(td, $J_1 = 15.7$ Hz, $J_2 = 6.6$ Hz, 1H), 5.33 (d, $J = 6.5$, 2H); $^{13}$C NMR (125 MHz, DMF-d$_7$): δ 159.4, 159.1, 143.0, 130.6, 127.6, 122.2, 115.7, 67.3, 50.7; HRMS (ESI) calcd for C$_{12}$H$_{12}$N$_4$O$_2$Na [M+Na]$^+$ 267.0858, found 267.0855.

**N-(4-methoxybenzyloxy)-3-(trimethylsilyl)propiolamide (1.23).** Commercially available acid 1.41 (710 mg, 4.99 mmol) was added to a reaction vessel containing CH$_2$Cl$_2$ (10 mL). To this solution, oxalyl chloride (471 μL, 5.49 mmol) and a trace amount of DMF were added. The mixture was stirred at room temperature for 2 h. The solution was then placed in an ice bath, and most of the acidic gases (HCl & SO$_2$) were removed under vacuum to give a yellow solution. Freshly dried CH$_2$Cl$_2$ (10 mL) was then added (solution A). The amine 1.40 (950 mg, 5.01 mmol) and Hünig’s base (2.80 mL, 16.1 mmol) were added to a second reaction vessel containing CH$_2$Cl$_2$ (50 mL) (solution B). Solution B was cooled to 0 °C, after which solution A was added within 1 min. The mixture was then stirred for 30 min at 0 °C. The reaction mixture was poured into ice cold 0.1 N HCl (200 mL), before CH$_2$Cl$_2$ (200 mL) was also added. The organic layer was additionally washed with ice cold 0.1 N HCl (2 x 100 mL). The DCM solution was dried over Na$_2$SO$_4$, and the solvent was evaporated. The crude product was purified via flash column chromatography (Hexanes : EtOAc, 6 : 1 to 4 : 1) to provide 8 as pale color oil (0.85 g, 61%). $^1$H NMR (500 MHz, CDCl$_3$): δ 8.43 (bd s, 1H), 7.32 (d, $J = 8.5$ Hz, 2H), 6.89 (dd, $J = 8.5$ Hz, 2H), 4.85 (s, 2H), 3.80 (s, 3H), 0.20 (s, 9H); $^{13}$C NMR (125 MHz, CDCl$_3$): δ 160.1, 151.0, 131.0, 126.9, 114.0, 95.0, 94.4, 78.2, 55.3, 0.8; HRMS (ESI) calcd for C$_{14}$H$_{19}$NO$_3$SiNa [M+Na]$^+$ 300.1032, found 300.1033.
General procedure for the preparation of 1.54-1.60. The alkyne precursor 1.23 (589 mg, 2.12 mmol) was added to a reaction vessel containing THF (50 mL) and MeOH (50 mL). Cesium fluoride (322 mg, 2.12 mmol) was then added, after which the reaction was monitored by TLC analysis. The desilylation intermediate was slightly more polar than 1.23. Upon the disappearance of 8, the solution was split into ten equal fractions. These fractions were directly used in the ensuing reactions as work-up led to degradation of the alkyne. To each fraction, an azide-containing compound 1.24-1.30 (0.32 mmol), catalyst CuI(P(OEt)3 (10 mg, 0.028 mmol) and one drop of Hünig’s base were added sequentially. The reaction was allowed to proceed to completion by stirring overnight at room temperature. The solvent was removed, and the residue was partitioned between CH2Cl2 (50 mL) and 1% CuSO4 aqueous solution (50 mL). The aqueous phase was additionally extracted with CH2Cl2 (2 x 50 mL). The combined organic fractions were then dried over Na2SO4. After removal of the solvent, the crude product was purified via flash column chromatography.

1-Benzyl-1H-[1,2,3]triazole-4-carboxylic acid (4-methoxy-benzylloxy)-amide (1.54). A white solid was obtained (60 mg, 83% yield) via flash column chromatography (CH3OH in CH2Cl2, 0.5% to 0.8%). 1H NMR (500 MHz, DMF-d7): δ 11.76 (bd s, 1H), 8.72 (s, 1H), 7.60-7.25 (m, 7H), 6.96 (d, J = 7.3 Hz, 2H), 5.77 (s, 2H), 4.95 (s, 2H), 3.82 (s, 3H); 13C NMR (125 MHz, DMF-d7): δ 161.0, 159.1, 142.8, 137.1, 131.8, 130.0, 129.5, 129.3, 127.7, 114.7, 78.5, 56.0, 54.5; HRMS (ESI) calcd for C18H18N4O3Na [M+Na]+ 361.1277, found 361.1263.
1-Phenethyl-1H-[1,2,3]triazole-4-carboxylic acid (4-methoxy-benzyloxy)-amide (1.55). A white solid was obtained (63 mg, 95% yield) via flash column chromatography (CH$_3$OH in CH$_2$Cl$_2$, 0.5% to 0.75 %). $^1$H NMR (500 MHz, DMF-d$_7$): δ 11.69 (bd s, 1H), 8.56 (s, 1H), 7.42 (d, $J = 8.6$ Hz, 2H), 7.32-7.20 (m, 5H), 6.96 (d, $J = 8.7$ Hz, 2H), 4.93 (s, 2H), 4.77 (t, $J = 7.3$ Hz, 2H), 3.81 (s, 3H), 3.28 (t, $J = 7.4$ Hz, 2H); $^{13}$C NMR (125 MHz, DMF-d$_7$): δ 161.0, 142.4, 138.8, 131.8, 129.9, 129.6, 129.3, 127.8, 127.7, 114.8, 78.5, 56.0, 52.2, 37.0; HRMS (ESI) calcd for C$_{19}$H$_{20}$N$_4$O$_3$Na [M+Na]$^+$ 375.1433, found 375.1425.

1-(3-Phenyl-propyl)-1H-[1,2,3]triazole-4-carboxylic acid (4-methoxy-benzyloxy)-amide (1.56). A white solid was obtained (53 mg, 67% yield) via flash column chromatography (CH$_3$OH in CH$_2$Cl$_2$, 0.5% to 0.6%). $^1$H NMR (500 MHz, DMF-d$_7$): δ 11.74 (bd s, 1H), 8.69 (s, 1H), 7.45 (d, $J = 8.4$ Hz, 2H), 7.36-7.20 (m, 5H), 6.98 (d, $J = 8.5$ Hz, 2H), 4.97 (s, 2H), 4.54 (t, $J = 7.0$ Hz, 2H), 3.83 (s, 3H), 2.67 (t, $J = 7.7$ Hz, 2H), 2.27 (tt, $J_1 = J_2 = 7.4$ Hz, 2H); $^{13}$C NMR (125 MHz, DMF-d$_7$): δ 161.0, 159.2, 142.6, 142.1, 131.8, 129.5, 129.3, 127.6, 127.2, 114.8, 78.5, 56.0, 50.7, 33.3, 32.8 (One less peak due to overlap); HRMS (ESI) calcd for C$_{20}$H$_{22}$N$_4$O$_3$Na [M+Na]$^+$ 389.1590, found 389.1585.
1-(2-Phenoxy-ethyl)-1H-[1,2,3]triazole-4-carboxylic acid (4-methoxy-benzyloxy)-amide (1.57). A white solid was obtained (44 mg, 56 % yield) via flash column chromatography (CH$_3$OH in CH$_2$Cl$_2$, 0.5% to 2%). $^1$H NMR (500 MHz, DMF-d$_7$): $\delta$ 11.81-11.72 (bd s, 1H), 8.73 (s, 1H), 7.44 (d, $J = 8.6$ Hz, 2H), 7.31 (t, $J = 8.0$ Hz, 2H), 7.01-6.95 (m, 5H), 4.98-4.94 (m, 4H), 4.54 (t, $J = 5.2$ Hz, 2H), 3.82 (s, 3H); $^{13}$C NMR (125 MHz, DMF-d$_7$): $\delta$ 161.0, 159.4, 159.1, 142.6, 131.8, 130.7, 129.3, 128.2, 122.3, 115.7, 114.7, 78.5, 63.7, 56.0, 50.8; HRMS (ESI) calcd for C$_{19}$H$_{20}$N$_4$O$_4$Na $[M+Na]^+$ 391.1382, found 391.1379.

1-(2-Oxo-2-phenyl-ethyl-1H-[1,2,3]triazole-4-carboxylic acid (4-methoxy-benzyloxy)-amide (1.58). A pale yellow solid was obtained (74 mg, 76% yield) via flash column chromatography (CH$_3$OH in CH$_2$Cl$_2$, 0.25% to 1%). $^1$H NMR (500 MHz, DMF-d$_7$): $\delta$ 11.83 (s, 1H), 8.69 (s, 1H), 8.17 (d, $J = 7.6$ Hz, 2H), 7.78 (t, $J = 7.3$ Hz, 1H), 7.65 (dd, $J_1 = J_2 = 7.5$ Hz, 2H), 7.46 (d, $J = 8.2$ Hz, 2H), 6.99 (d, $J = 8.2$ Hz, 2H), 6.39 (s, 2H), 4.99 (s, 2H), 3.84 (s, 3H); $^{13}$C NMR (125 MHz, DMF-d$_7$): $\delta$ 192.9, 161.0, 159.2, 142.6, 135.6, 135.4, 131.9, 130.1, 129.5, 129.3, 114.8, 78.5, 57.4, 56.0 (One less peak due to overlap); HRMS (ESI) calcd for C$_{19}$H$_{18}$N$_4$O$_4$Na $[M+Na]^+$ 389.1266, found 389.1218.
1-Phenylcarbamoylmethyl-1H-[1,2,3]triazole-4-carboxylic acid (4-methoxy-benzyloxy)-amide (1.59). The crude product was dissolved in a minimal amount of 5% CH₃OH in CH₂Cl₂. The pure compound was then precipitated via addition of hexane (2:1 = Hexane : Solution). Following centrifugation (5000 rpm, 5 min), the supernatant was removed, and the solid was sonicated in the presence of a 1% EDTA aqueous solution. The solution was then filtered to afford the pure compound (82 mg, 81% yield). ¹H NMR (500 MHz, DMF-d₇): δ 11.90-11.70 (bd s, 1H), 10.65 (s, 1H), 8.72 (s, 1H), 7.69 (d, J = 8.0 Hz, 2H), 7.45 (d, J = 8.5 Hz, 2H), 7.36 (dd, J₁ = J₂ = 7.8 Hz, 2H), 7.12 (t, J = 7.3 Hz, 1H), 6.98 (d, J = 8.5 Hz, 2H), 5.57 (s, 2H), 4.98 (s, 2H), 3.83 (s, 3H); ¹³C NMR (125 MHz, DMF-d₇): δ 165.2, 161.0, 159.2, 142.4, 140.0, 131.8, 129.4, 129.3, 124.9, 120.4, 114.8, 78.5, 56.0, 53.7 (One less peak due to overlap); HRMS (ESI) calcd for C₁₉H₁₉N₅O₄Na [M+Na]⁺ 404.1335, found 404.1319.

1-(3-Phenyl-allyl)-1H-[1,2,3]triazole-4-carboxylic acid (4-methoxy-benzyloxy)-amide (1.60). A white solid was obtained (62 mg, 80% yield) via flash column chromatography (CH₃OH in CH₂Cl₂, 0.2% to 0.5%). ¹H NMR (500 MHz, DMF-d₇): δ 11.76 (s, 1H), 8.69 (s, 1H), 7.54 (d, J = 7.4 Hz, 2H), 7.44 (d, J = 8.6 Hz, 2H), 7.38 (dd, J₁ = J₂ = 7.5 Hz, 2H), 7.31 (t, J = 7.3 Hz, 1H), 6.97 (d, J = 8.5 Hz, 2H), 6.80 (d, J = 15.9 Hz, 1H), 6.63 (td, J₁ = 15.8 Hz, J₂ = 6.6 Hz, 1H), 5.34 (d, J = 6.5, 2H), 4.96 (s, 2H), 3.82 (s, 3H); ¹³C NMR (125 MHz, DMF-d₇): δ 161.0, 159.1, 142.7,
137.3, 135.6, 131.8, 129.8, 129.4, 129.3, 127.8, 127.5, 124.4, 114.7, 78.5, 56.0, 53.0;
HRMS (ESI) calcd for C_{20}H_{20}N_{4}O_{3}Na [M+Na]^+ 387.1433, found 387.1430.

5.2 Biological studies of HDACi

For the HDAC1 and HDAC8 inhibition studies, the assay kits were obtained from Biomol International LP. All assays were repeated at least three times. Briefly, the assay was performed in two stages. In the first stage, HDAC8 solution (15 μL, 1 U total), HDACi solution (10 μL) and HDAC8 substrate solution (25 μL) were added to the 96-well microplate. The reaction proceeded at 30 °C for 45 min. The second stage was initiated by the addition of 50 μL of Developer which stopped HDAC activity and produced the fluorescent signal (λ<sub>ex</sub> = 350 nm; λ<sub>em</sub> = 450 nm). The cell line assays for the selected six target compounds, 5b (NSC746454), 5e (NSC746455), 5f (NSC746456), 5g (NSC746457), 6e (NSC746458) and 6g (NSC746459), were performed by the NIH. Cell lines tested included the following: CCRF-CEM, HL-60(TB), K-562, MOLT-4, RPMI-8226 and SR (Leukemia); A549/ATCC, EKVX, HOP-62, HOP-92, NCI-H226, NCI-H23, NCI-H322M and NCI-H460 (Non-Small Cell Lung Cancer); COLO 205, HCT-116, HCT-15, HT29, KM12 and SW-620 (Colon Cancer); SF-628, SF-295, SF-539, SNB-19, SNB-75 and U251 (CNS Cancer); LOX IMVI, MALME-3M, M14, SK-MEL-28, SK-MEL-5, UACC-257 and UACC-62 (Melanoma); IGROV1, OVCAR-3, OVCAR-4, OVCAR-5, OVCAR-8 and SK-OV-3 (Ovarian Cancer); 786-0, A498, ACHN, CAKI-1, RXF 393, SN12C, TK-10 and UO-31 (Renal Cancer); PC-3 and DU-145 (Prostate Cancer); MCF7, NCI/ADR-RES, MDA-MB-231/ATCC, HS578T, MDA-MB-435, BT-

5.3 Sugar-PP-Polyprenol chemical synthesis and compound characterization

All solvents were dried with a solvent-purification system from Innovative Technology, Inc. All reagents were obtained from commercial sources and used without further purification. The 200-400 mesh silica gel 60 RP-18 (from EMD™) was utilized for purification. 1H, 13C and 31P NMR spectra were recorded at the indicated field strengths. Mass spectral data was collected in our laboratory using a Shimadzu LCMS-2010A Liquid Chromatograph Mass Spectrometer as well as at The Ohio State University mass spectrometry facility on a Bruker micrOTOF Instrument provided by a grant from the Ohio BioProducts Innovation Center.

Acetic acid 3,4,6-triacetoxy-5-acetylamino-tetrahydro-pyran-2-ylmethyl ester (2.2). A reaction vessel was charged with 2.1 (0.900 g, 4.07 mmol) and was flushed with nitrogen. CH3CN (25 mL) and triethylamine (6.65 mL, 47.6 mmol) were then introduced in sequence, after which the reaction was cooled to 0 °C. Acetic anhydride (6.65 mL, 70.4 mmol) was added dropwise, and the reaction was stirred at 0 °C until the solution became clear (~ 2 h). At this point, the reaction was warmed to room temperature and stirred for an additional 2 h. The reaction was cooled once more to 0 °C,
and CH$_3$OH (9.1 mL) was added. The reaction was stirred for 30 min, after which the solvent was removed. The residue was then dissolved in CH$_2$Cl$_2$ (50 mL) and washed with 1 M HCl (aq.) (25 mL), 5% NaHCO$_3$ (aq.) (2 x 25 mL), and brine (25 mL). The organic layer was dried over Na$_2$SO$_4$. The crude product was subjected to flash column chromatography (Hexanes : Acetone, 1:1 to 1:2) to afford 2.2 as a white foam (1.450 g, 92 % yield). $^1$H NMR (500 MHz, CD$_3$Cl): δ 6.18 (d, $J$ = 3.6 Hz, 1H), 5.48 (d, $J$ = 9.1 Hz, 1H), 5.39 (d, $J$ = 2.7 Hz, 1H), 5.18 (dd, $J$ = 11.5 Hz, $J$ = 3.2 Hz, 1H), 4.72-4.66 (m, 1H), 4.21 (t, $J$ = 6.7 Hz, 1H), 4.10-4.00 (m, 1H), 2.14 (s, 6H), 2.00 (s, 3H), 1.99 (s, 3H), 1.91 (s, 3H); $^{13}$C NMR (125 MHz, CD$_3$Cl): δ 171.3, 170.5, 170.4, 170.2, 169.0, 91.5, 68.7, 68.0, 65.9, 61.5, 47.2, 23.3, 21.1, 20.9, 20.8, 20.8.

Acetic acid 3-acetoxy-2-acetoxymethyl-5-acetylamino-6-hydroxy-tetrahydro-pyran-4-yl ester (2.3). A reaction vessel was charged with 2.2 (1.450 g, 3.724 mmol) and DMF (20 mL). Hydrazine acetate (0.4116 g, 4.469 mmol) was then added in one portion, and the reaction was stirred at room temperature for 40 min. The solvent was evaporated, and the crude product was subjected to flash column chromatography (Hexanes : Acetone, 2:1 to 1:1) to afford 2.3 as a white foam (1.031 g, 80 % yield). $^1$H NMR (500 MHz, CD$_3$Cl): δ 5.94 (d, $J$ = 9.5 Hz, 1H), 5.35 (d, $J$ = 2.4 Hz, 1H), 5.28 (t, $J$ = 2.9 Hz, 1H), 5.22 (dd, $J$ = 11.4 Hz, $J$ = 3.3 Hz, 1H), 4.53-4.47 (m, 1H), 4.40 (t, $J$ = 6.6 Hz, 1H), 4.29 (bs, 1H), 4.10-4.01 (m, 2H), 2.13 (s, 3H), 2.02 (s, 3H), 1.97 (s, 3H), 1.95 (s, 3H); $^{13}$C NMR (125 MHz, CD$_3$Cl): δ 171.3, 171.0, 170.8, 170.6, 92.3,
Acetic acid 3,4-diacetoxy-5-acetylamino-6-(bis-benzylxophosphoryloxy)-tetrahydro-pyran-2-ylmethyl ester (2.4). A reaction vessel was charged with 2.3 (2.275 g, 6.550 mmol) and tetrazole (4.558 g, 65.49 mmol), and was flushed with nitrogen. CH₂Cl₂ (75 mL) was added, after which the reaction was cooled to -40 °C. Dibenzyl N,N-diisopropylphosphoramidite (10.76 mL, 32.75 mmol) was then added dropwise over a period of 5 min. The reaction was stirred for 3 h as it warmed to room temperature. The reaction was subsequently cooled to -78 °C, and mCPBA (11.303, 65.499 mmol) was added in two portions. The reaction was allowed to warm to room temperature overnight. This solution was then washed with Na₂SO₃ (30 mL), H₂O (30 mL) and brine (30 mL). The organic layer was dried over Na₂SO₄. The crude product was subjected to flash column chromatography (CH₃OH in CH₂Cl₂, 0% to 2%) to afford 2.4 as a white foam (7.975 g, 70 % yield). ¹H NMR (500 MHz, CD₃Cl): δ 8.59 (bs, 1H), 7.39-7.31 (m, 10H), 5.70 (dd, J = 5.8 Hz, J = 3.3 Hz, 1H), 5.48 (d, J = 9.4 Hz, 1H), 5.36 (d, J = 2.0 Hz, 1H), 5.12-5.00 (m, 4H), 4.63-4.56 (m, 1H), 4.24 (t, J = 6.6 Hz, 1H), 4.05 (dd, J = 11.3 Hz, J = 6.6 Hz, 1H), 3.92 (dd, J = 11.3 Hz, J = 6.5 Hz, 1H), 2.12 (s, 3H), 1.97 (s, 3H), 1.91 (s, 3H), 1.70 (s, 3H); ¹³C NMR (125 MHz, CD₃Cl): δ 170.9, 170.5, 170.4, 170.3, 129.1, 129.0, 129.0, 128.3, 128.3, 97.2, 70.2, 70.1, 70.0, 68.9, 67.5, 67.0, 61.5, 47.7, 23.1, 20.9, 20.9, 20.8; ³¹P NMR (162 MHz, CD₃Cl): δ -3.5.
GalNAc-1-PO$_4^{2-}$ diisopropylethlammonium salt

(2.5). A reaction vessel was charged with **2.4** (90 mg, 0.15 mmol), 20% Pd(OH)$_2$ on carbon (15 mg), and CH$_3$OH (5 mL). The vessel was placed under a H$_2$ atmosphere (1 atm), and the reaction was stirred for 30 min. Hünig’s base (0.1 mL) was then added to the reaction mixture followed by CH$_3$OH (5 mL). The reaction was stirred for 30 min before the solution was filtered through a pad of Celite. The crude product was concentrated and utilized directly in the coupling step.

**General procedure for preparation of Polyprenol-phosphates.** A reaction vessel was charged with a polyprenol (0.152 mmol) was flushed with argon. Dicyanoethoxy N,N-diisopropylamine (50 mg, 0.18 mmol) was then added, and the reaction was cooled to 0 °C. Tetrazole (42 mg, 0.59 mmol) was added in one portion. The reaction was stirred at 0 °C for 5 min, after which it was warmed to room temperature and stirred for an additional 1 h. The reaction was cooled once again to 0 °C, and pyridine (0.040 mL, 0.53 mmol) and NaIO$_4$ (65 mg, 0.30 mmol) were added sequentially. After stirring for 15 min at 0 °C, the reaction was warmed to room temperature and stirred for an additional 1 h. The reaction was then diluted with EtOAc (30 mL) and washed with H$_2$O (30 mL), 5% Na$_2$SO$_3$ (aq.) (30 mL) and brine (30 mL). The organic layer was dried over Na$_2$SO$_4$ and concentrated. The residue was dissolved in a 1% NaOCH$_3$ in CH$_3$OH solution (10 mL), and the reaction was allowed to proceed for 3 days. The solvent was removed, and the
crude product was purified via ion exchange on DEAE cellulose. In the case of \textbf{2.105}, however, a different procedure was employed. Specifically, reaction vessel A was charged with MS-Pent (24 mg, 0.067 mmol) and hexanes (3.5 mL). POCl$_3$ (0.030 mL, 0.36 mmol), triethylamine (0.050 mL, 0.36 mmol), and hexanes (1.75 mL) were added to reaction vessel B. The reaction in vessel B was stirred for 5 min, after which the contents of reaction vessel A was cannulated into vessel B over a period of 15 min. The reaction was stirred for 15 min. Acetone : triethylamine : H$_2$O (88:10:2) (~ 60 mL) was added, and the reaction was stirred overnight. The solvent was removed, after which the residue was washed several times with \textit{n}-butanol (1.5 mL) and benzene (1.5 mL). A final portion of benzene (1.5 mL) was added, and the solution was filtered. The filtrate was concentrated and purified via ion exchange on DEAE cellulose.

\textbf{General procedure of preparation of GalNAc-PP-Polyrenol (Precursors to 2.19-2.24).} 1,1'-Carbonyldiimidazole (CDI) (23 mg, 0.14 mmol) was added to a reaction vessel containing \textbf{2.5} (23 mg, 0.034 mmol). The vessel was flushed with argon, after which anhydrous THF was added (3 mL). The reaction was stirred at room temperature for 2 h. Addition of dry CH$_3$OH (42 μL) followed by stirring for an additional 1 h at room temperature served to remove excess CDI. The reaction mixture was concentrated and then anhydrous THF (3 mL) was once again added. To a separate reaction vessel, lipid phosphate (0.026 mmol) was added, and the vessel was then flushed with argon. The activated GalNAc-1- PO$_4^{2-}$ was transferred into this vessel via syringe. The reaction was stirred at room temperature for 3 days, after which the reaction mixture was concentrated. Dissolution of the residue in water, followed by filtration through a short
bed of C-18 reverse phase silica gel (H₂O then 1:1 H₂O/Isopropanol) afforded the crude product.

The crude product was treated with a 0.1% NaOCH₃ in CH₃OH solution (3 mL). The reaction was stirred at room temperature for 40 min. The reaction mixture was concentrated, after which the residue was dissolved in water. Purification via C-18 reverse phase column chromatography (Solvent A: Isopropanol (Lipids: Und, Solan) or CH₃CN (Lipids: Hept, Pent, cis-Pent, MS-Pent); Solvent B = 1.7% NH₄HCO₃; A/B = 0%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% and 50% (10 mL each)) provided GalNAc-PP-Lipid as a white solid following lyophilization.

GalNAc-PP-Lipid

(Precursor to 2.19). GalNAc-PP-Und was prepared as described above. However, the two purification steps were reversed as the final product proved too insoluble to obtain suitable spectral data. The crude product was thus purified via C-18 reverse phase column chromatography to afford Peracetylated GalNAc-PP-Und as a white solid (20 mg, 59%).

^1H NMR (500 MHz, CD₃OD): δ 5.60 (dd, J = 3.2 Hz, J = 7.3 Hz, 1H), 5.42 (t, J = 6.1 Hz, 2H), 5.21 (dd, J = 3.2 Hz, J = 11.3 Hz, 1H), 5.14-5.05 (m, 10H), 4.58 (t, J = 6.9 Hz, 1H), 4.54-4.45 (m, 3H), 4.21 (dd, J = 8.1 Hz, J = 10.9 Hz, 1H), 4.04 (dd, J = 5.8 Hz, J = 10.8 Hz, 1H), 2.10 (s, 3H), 2.08-2.00 (m, 34H), 1.98 (s, 3H), 1.97 (s, 3H), 1.96-1.92 (m, 6H), 1.89 (s, 3H), 1.71 (s, 3H), 1.66-1.62 (m, 21H), 1.58 (s, 3H), 1.57 (s, 9H); ^13C NMR (125 MHz, CD₃OD): δ 174.4, 172.3, 172.2, 171.9, 140.6, 136.5, 136.4, 136.3, 136.3, 136.1, 147
GalNAc-PP-Heptaprenyl (Precursor to 2.20). GalNAc-PP-Hept was prepared as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-Hept as a white solid (22 mg, 36% over two steps). $^1$H NMR (500 MHz, D$_2$O): $\delta$ 5.61 (dd, $J = 2.7$ Hz, $J = 5.9$ Hz, 1H), 5.50 (t, $J = 5.5$ Hz, 1H), 5.24-5.08 (m, 6H), 4.51 (t, $J = 5.7$ Hz, 2H), 4.29 (d, $J = 10.8$ Hz, 1H), 4.23 (t, $J = 5.5$ Hz, 1H), 4.07 (d, $J = 1.4$ Hz, 1H), 4.00 (dd, $J = 2.2$ Hz, $J = 10.7$ Hz, 1H), 3.88-3.75 (m, 2H), 2.20-1.96 (m, 27H), 1.79 (s, 3H), 1.72 (s, 3H), 1.70 (s, 3H), 1.68 (s, 3H), 1.62 (s, 3H), 1.60 (s, 9H);
$^{13}$C NMR (125 MHz, D$_2$O): $\delta$ 174.5, 172.4, 172.2, 171.9, 140.8, 136.3, 135.9, 135.9, 135.8, 132.1, 125.7, 125.7, 125.6, 125.4, 122.7, 122.6, 96.5, 70.6, 68.7, 64.0, 64.0, 62.3, 41.0, 41.0, 40.9, 40.8, 34.5, 33.2, 30.9, 30.8, 30.7, 30.6, 30.5, 28.0, 27.9, 27.8, 27.7, 27.7, 27.6, 26.0, 24.0, 23.8, 23.0, 20.8, 20.7, 20.7, 17.9, 16.8, 16.2, 15.7, 14.5;
174.7, 165.3, 134.8, 134.6, 134.3, 130.4, 125.2, 125.0, 124.8, 124.5, 124.3, 124.3, 121.2, 94.6, 72.1, 68.7, 67.7, 62.6, 61.3, 49.9, 39.7, 39.7, 32.0, 31.9, 31.9, 29.9, 29.8, 29.7, 29.5, 26.7, 26.6, 26.2, 25.4, 23.3, 23.2, 23.0, 23.0, 22.3, 17.3, 15.7; $^{31}$P NMR (162 MHz, D$_2$O): 
$\delta$ -9.7, -11.4; LRMS (m/z): [M-H]- calcd. for C$_{43}$H$_{72}$NO$_{12}$P$_2$, 856.5; found, 856.7.

**GalNAc-PP-Pentaprenyl (Precursor to 2.21).** GalNAc-PP-Pent was prepared as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-Pent as a white solid (22 mg, 49% over two steps). $^1$H NMR (500 MHz, CD$_3$OD): $\delta$ 5.67 (dd, $J = 3.1$ Hz, $J = 6.8$ Hz, 1H), 5.56 (t, $J = 6.6$ Hz, 1H), 5.30-5.16 (m, 4H), 4.58 (t, $J = 6.5$ Hz, 2H), 4.36 (d, $J = 10.9$ Hz, 1H), 4.30 (dd, $J = 6.0$ Hz, 1H), 4.14 (d, $J = 3.0$ Hz, 1H), 4.07 (dd, $J = 2.9$ Hz, $J = 10.9$ Hz, 1H), 3.94-3.84 (m, 2H), 3.20-2.23 (m, 2H), 2.22-2.11 (m, 13H), 2.10-2.02 (m, 4H), 1.85 (s, 3H), 1.80 (s, 3H), 1.75 (s, 3H), 1.71 (s, 3H), 1.68 (s, 6H); $^{13}$C NMR (125 MHz, D$_2$O): $\delta$ 174.7, 160.3, 141.2, 135.8, 134.9, 134.4, 130.6, 124.6, 124.5, 124.3, 121.2, 94.7, 72.2, 68.7, 67.7, 62.7, 61.3, 49.9, 48.9, 39.7, 39.7, 32.0, 31.9, 26.7, 26.6, 26.1, 25.4, 23.1, 22.9, 22.3, 17.4, 15.8, 15.7; $^{31}$P NMR (162 MHz, D$_2$O): $\delta$ -10.4, -13.1; LRMS (m/z): [M-H]- calcd. for C$_{33}$H$_{56}$NO$_{12}$P$_2$, 720.3; found, 720.3.
GalNAc-PP-cis-Pentaprenyl (Precursor to 2.22). GalNAc-PP-cis-Pent was prepared as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-cis-Pent as a white solid (15 mg, 20% over two steps). $^1$H NMR (500 MHz, D$_2$O): $\delta$ 5.62 (dd, $J = 3.0$ Hz, $J = 6.5$ Hz, 1H), 5.51 (t, $J = 6.0$ Hz, 1H), 5.24-5.12 (m, 4H), 4.53 (t, $J = 6.2$ Hz, 2H), 4.30 (d, $J = 11.1$ Hz, 1H), 4.25 (t, $J = 5.9$ Hz, 1H), 4.09 (d, $J = 2.4$ Hz, 1H), 4.01 (dd, $J = 2.6$ Hz, $J = 10.9$ Hz, 1H), 3.88-3.78 (m, 2H), 2.22-2.00 (m, 19H), 1.79 (s, 3H), 1.73 (s, 3H), 1.72-1.68 (m, 9H), 1.63 (s, 3H); $^{13}$C NMR (125 MHz, D$_2$O): $\delta$ 174.7, 141.3, 135.4, 135.4, 134.8, 134.7, 134.7, 125.2, 125.1, 124.9, 124.4, 124.4, 94.8, 72.2, 68.6, 67.7, 62.8, 61.3, 49.8, 32.1, 32.1, 31.8, 26.6, 26.4, 26.2, 25.5, 25.4, 23.2, 23.2, 23.0, 22.3, 17.3, 15.7; $^{31}$P NMR (162 MHz, D$_2$O): $\delta$ -10.1, -12.3; LRMS ($m/z$): [M-H]- calcd. for C$_{33}$H$_{56}$NO$_{12}$P$_2$, 720.3; found, 720.4.

GalNAc-PP-Solanesyl (Precursor to 2.23). GalNAc-PP-Solan was prepared as described above. However, the two purification steps were reversed as the final product proved too insoluble to obtain suitable spectral data. The crude product was thus purified via C-18 reverse phase column chromatography to afford Peracetylated GalNAc-PP-Solan as a white solid (35 mg, 50%). $^1$H NMR (500
MHz, CD$_3$OD): $\delta$ 5.61 (dd, $J = 3.3$ Hz, $J = 7.3$ Hz, 1H), 5.43 (t, $J = 6.6$ Hz, 2H), 5.22 (dd, $J = 3.1$ Hz, $J = 11.3$ Hz, 1H), 5.14-5.04 (m, 8H), 4.59 (t, $J = 7.0$ Hz, 1H), 4.53 (t, $J = 6.3$ Hz, 3H), 4.21 (dd, $J = 8.2$ Hz, $J = 10.8$ Hz, 1H), 4.04 (dd, $J = 5.8$ Hz, $J = 10.8$ Hz, 1H), 2.11 (s, 3H), 2.10-2.02 (m, 18H), 1.99 (s, 3H), 1.99-1.94 (m, 17H), 1.90 (s, 3H), 1.69 (s, 3H), 1.65 (s, 3H), 1.59 (s, 3H), 1.58 (s, 21H); $^{13}$C NMR (125 MHz, CD$_3$OD): $\delta$ 174.5, 172.4, 172.2, 171.9, 140.8, 136.3, 135.9, 135.9, 135.9, 135.8, 132.1, 125.7, 125.7, 125.6, 125.4, 122.7, 122.6, 96.5, 70.6, 68.7, 64.0, 64.0, 62.3, 41.0, 41.0, 41.0, 40.9, 40.8, 34.5, 33.2, 30.9, 30.8, 30.7, 30.6, 30.5, 28.0, 27.9, 27.8, 27.7, 27.7, 27.7, 26.0, 24.0, 23.8, 23.0, 20.8, 20.7, 20.7, 17.9, 16.8, 16.2, 15.7, 14.5; $^{31}$P NMR (162 MHz, CD$_3$OD): $\delta$ -9.6, -12.5; LRMS (m/z): [M-H]- calcd. for C$_{59}$H$_{94}$O$_{15}$P$_2$, 1118.6; found, 1118.5. Final deprotection with 0.1% NaOCH$_3$ in CH$_2$OH and filtration though a short bed of C-18 reverse phase silica gel afforded GalNAc-PP-Solan (31 mg, 99%). LRMS (ESI) calcd for C$_{53}$H$_{88}$O$_{12}$P$_2$ [M-H]$^-$ 992.6, found 992.6.

**GalNAc-PP-MS-Pentaprenyl**

(Precursor to 2.24). GalNAc-PP-MS-Pent was prepared as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-MS-Pent as a white solid (27 mg, 26% over two steps). $^1$H NMR (500 MHz, D$_2$O): $\delta$ 5.55 (dd, $J = 3.3$ Hz, $J = 7.1$ Hz, 1H), 5.22-5.05 (m, 4H), 4.26 (dt, $J = 2.9$ Hz, $J = 11.0$ Hz, 1H), 4.19 (dd, $J = 5.2$ Hz, $J = 7.1$ Hz, 1H), 4.06-3.93 (m, 4H), 3.84-3.72 (m, 2H), 2.11-2.02 (m, 11H), 2.01-1.92 (m, 6H), 1.69 (s, 3H), 151
1.66 (s, 3H) 1.61 (s, 3H), 1.58 (s, 6H), 1.54-1.24 (m, 4H), 1.22-1.10 (m, 1H), 0.92 (d, \( J = 6.5 \) Hz, 3H); \(^{13}\)C NMR (125 MHz, D\(_2\)O): \( \delta \) 174.5, 135.0, 134.8, 134.5, 130.8, 125.7, 124.4, 124.2, 94.7, 72.1, 68.6, 67.7, 64.9, 61.3, 49.9, 49.8, 39.6, 37.3, 37.1, 31.8, 29.0, 26.6, 26.6, 25.4, 24.9, 23.2, 22.3, 18.8, 18.8, 17.4, 15.7, 15.7; \(^{31}\)P NMR (162 MHz, D\(_2\)O): \( \delta \) -9.9, -11.8; LRMS (m/z): [M-H]- calcd. for C\(_{33}\)H\(_{58}\)NO\(_{12}\)P\(_2\), 722.3; found, 722.3.

\[
\text{BnO} \quad \text{(3,7-Dimethyl-octa-2,6-dienyloxymethyl)-benzene (2.35).}
\]
A reaction vessel was charged with sodium hydride (4.6677 g, 194.49 mmol) and flushed with argon. THF (200 mL) was added to the flask, after which the vessel was cooled to 0 °C. \textbf{2.34} (30.0 g, 194 mmol) was then added dropwise to the solution. The reaction was allowed to proceed for 30 min at 0 °C, at which point benzyl bromide (23.13 mL, 194.5 mmol) and NaI (2.9152 g, 19.449 mmol) were introduced into the reaction. The reaction was heated to reflux and stirred for 5 h. After quenching the reaction with 1.0 M HCl (aq.), the desired product was extracted with a Hexane : EtOAc mixture (10:1). The organic layer was dried over Na\(_2\)SO\(_4\). The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5.0%) to afford \textbf{2.35} as a colorless oil (43.8 g, 92%). \(^1\)H NMR (500 MHz, CD\(_3\)Cl): \( \delta \) 7.38-7.25 (m, 5H), 5.41 (td, \( J = 6.9 \) Hz, \( J = 1.8 \) Hz, 1H), 5.10-5.04 (m, 1H), 4.49 (s, 2H), 4.00 (d, \( J = 6.9 \) Hz, 2H), 2.07-2.02 (m, 4H), 1.75 (s, 3H), 1.66 (s, 3H), 1.57 (s, 3H); \(^{13}\)C NMR (125 MHz, CD\(_3\)Cl): \( \delta \) 140.8, 138.8, 132.1, 128.5, 128.0, 127.7, 124.1, 122.1, 72.3, 66.7, 32.5, 26.9, 25.9, 23.7, 17.8.
6-Benzyloxy-4-methyl-hex-4-en-1-ol (2.86). A reaction vessel was charged with 2.35 (43.80 g, 179.2 mmol). Dioxane (250 mL) and H₂O (30 mL) were added to the vessel, after which the reaction was cooled to 0 °C. N-Bromosuccinimide (31.89 g, 179.2 mmol) was then added in three portions, and the reaction was stirred for 5 h. A solution of KOH (10.05 g, 179.2 mmol) in CH₃OH (60 mL) was added dropwise to the vessel. After stirring for 3 h, the reaction mixture was diluted with water (150 mL) and extracted with Et₂O (200 mL). The organic layer was dried over Na₂SO₄. The crude product was concentrated and utilized directly in the next step. Specifically, a reaction vessel was charged with 2.36, after which THF (250 mL) and water (50 mL) were added. Sodium periodate (19.16 g, 89.60 mmol) and periodic acid (49.01 g, 215.0 mmol) were then added in sequence. The reaction was stirred for 30 min at room temperature and subsequently diluted with saturated Na₂SO₄ (aq.) (100 mL). The crude product was extracted with EtOAc (200 mL). The organic layer was dried over Na₂SO₄ before being concentrated. The crude product was utilized directly in the next step. Specifically, a reaction vessel was charged with the aldehyde (39.12 g, 179.2 mmol). CH₃OH (250 mL) was then added, and the reaction was cooled to 0 °C. NaBH₄ (5.425 g, 143.4 mmol) was added in three portions. The reaction was stirred for 30 min, after which the solvent was removed. The residue was dissolved in EtOAc (200 mL) and washed with water (100 mL). The organic layer was dried over Na₂SO₄. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 15% to 25%) to afford 2.86 as a colorless oil (2.1734 g, 32% over three steps). ¹H NMR (500 MHz, CD₂Cl): δ 7.33 (d, J = 4.4 Hz, 4H), 7.30-7.25 (m, 1H), 5.49 (d, J = 7.2 Hz, 4H),
4.50 (s, 2H), 3.97 (d, \( J = 7.2 \) Hz, 2H), 3.55 (t, \( J = 6.5 \) Hz, 2H), 2.29 (bs, 1H), 2.17 (t, \( J = 7.3 \) Hz, 2H), 1.73 (s, 3H), 1.64 (q, \( J = 6.7 \) Hz, 2H); \( ^{13} \)C NMR (125 MHz, CD\(_3\)Cl): \( \delta \) 141.8, 138.3, 128.6, 128.2, 127.9, 122.0, 72.6, 66.0, 61.5, 30.3, 28.0, 23.3.

(6-Benzylxy-4-methyl-hex-4-enyl)-triphenylphosphonium iodide (2.37). A reaction vessel was charged with 2.86 (13.00 g, 59.01 mmol) and benzene (300 mL). Triphenylphosphine (20.12 g, 76.71 mmol), imidazole (7.231 g, 106.2 mmol) and iodine (17.97 g, 70.81 mmol) were then added sequentially. The reaction was stirred for 20 min, after which the solvent was removed. The residue was diluted with ether (200 mL) and washed with H\(_2\)O (50 mL), 20% Citric Acid (50 mL), sat. NaHCO\(_3\) (aq.) (50 mL) and brine (50 mL). The organic layer was concentrated, suspended in hexanes and then filtered through a short bed of silica gel. The solvent was removed, and the crude product was utilized directly in the next step. Specifically, a reaction vessel was charged with the crude product (19.49 g, 59.01 mmol) and triphenylphosphine (21.69 g, 82.68 mmol). Benzene (100 mL) was added, and the solution was brought to reflux and stirred for 20 h. The solvent was then removed, yielding a white powder. The powder was triturated with Et\(_2\)O to remove residual triphenylphosphine and afforded 2.37 as a white powder (30.7 g, 88% yield over two steps).
1-(Tetrahydro-pyran-2-yloxy)-propan-2-one (2.38). A reaction vessel was charged with dihydropyran (24.7 mL, 137 mmol), hydroxyacetone (9.35 mL, 270 mmol) and pyridinium p-toluenesulfonate (0.339 g, 1.35 mmol). The reaction was allowed to proceed at room temperature for 2.5 h, after which it was diluted with hexanes and subjected to flash column chromatography (EtOAc in Hexanes, 0 to 6%) to afford 2.38 as a colorless oil (17.2 g, 80% yield). \[^1\]H NMR (500 MHz, CD\textsubscript{3}Cl): \(\delta\) 4.58 (t, \(J = 3.2\) Hz, 1H), 4.18 (d, \(J = 17.3\) Hz, 1H), 4.05 (d, \(J = 17.3\) Hz, 1H), 3.80-3.74 (m, 1H), 3.48-3.42 (m, 1H), 2.11 (s, 3H), 1.90-1.44 (m, 6H); \[^13\]C NMR (125 MHz, CD\textsubscript{3}Cl): \(\delta\) 206.8, 98.9, 72.4, 62.5, 30.4, 26.6, 25.4, 19.3.

2-(8-Benzylloxy-2,6-dimethyl-octa-2,6-dienyloxy)-tetrahydro-pyran (2.39). A reaction vessel was charged with 2.37 (14.89 g, 25.13 mmol) and flushed with nitrogen. THF was added (100 mL), and the vessel was cooled to -78 °C. \(n\)-BuLi (13.35 mL, 21.36 mmol) was then added dropwise, after which the reaction was stirred for 30 min. 2.38 (4.929 g, 31.16 mmol) was added dropwise, and the reaction mixture was allowed to warm to room temperature overnight. The reaction was quenched with saturated NH\textsubscript{4}Cl (aq.) (50 mL), after which the product was extracted with EtOAc (10 mL). The organic layer was dried over Na\textsubscript{2}SO\textsubscript{4}. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 0% to 5%) to afford 2.39 as a colorless oil (7.120 g, 82%). \[^1\]H NMR (500 MHz, CD\textsubscript{3}Cl): \(\delta\) 7.33-7.25 (m, 5H), 5.40 (t, \(J = 6.8\) Hz, 1H), 5.30 (t, \(J = 7.1\) Hz, 1H), 4.55 (t, \(J = 3.6\) Hz, 1H), 4.47 (s, 2H), 4.12-
4.01 (m, 2H), 3.98 (d, \( J = 6.9 \) Hz, 2H), 3.88-3.82 (m, 1H), 3.51-3.45 (m, 1H), 2.16-2.03 (m, 4H), 1.86-1.77 (m, 1H), 1.74 (s, 3H), 1.72-1.46 (m, 5H); \(^{13}\)C NMR (125 MHz, CD\(_3\)Cl): \( \delta \) 140.3, 138.8, 132.6, 128.9, 128.6, 128.0, 127.7, 122.4, 97.7, 72.3, 66.6, 65.5, 62.3, 32.6, 30.9, 26.5, 25.7, 23.7, 22.0, 19.7.

![Structure](image)

(8-Chloro-3,7-dimethyl-octa-2,6-dienyloxymethyl)-benzene (2.41). A reaction vessel was charged with 2.89 (1.3384 g, 5.1404 mmol) and flushed with nitrogen. CH\(_2\)Cl\(_2\) (20 mL) and THF (10 mL) were added, and the vessel was cooled to -40 °C. Triethylamine (0.93 mL, 6.7 mmol) was then added, and the reaction was stirred for 5 min. Dropwise addition of mesyl chloride (0.48 mL, 6.2 mmol) was followed by stirring of the reaction for 30 min. A solution of LiCl (2.1790 g, 51.400 mmol) in THF (35 mL) was transferred to the vessel, and the vessel was allowed to warm to 0 °C over 3 h. The solvent was removed, and the remaining residue was diluted with EtOAc (100 mL). The organic layer was washed with saturated NaHCO\(_3\) (50 mL) and dried over NaSO\(_4\). The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 1% to 5%) to afford 2.41 as a colorless oil (1.2917 g, 90%). \(^1\)H NMR (400 MHz, CD\(_3\)Cl): \( \delta \) 7.43-7.25 (m, 5H), 5.44 (d, \( J = 6.8 \) Hz, 1H), 5.33 (d, \( J = 6.6 \) Hz, 1H), 4.49 (s, 2H), 4.00 (s, 2H), 3.98 (d, \( J = 7.0 \) Hz, 2H), 2.19-2.06 (m, 4H), 1.79 (s, 3H), 1.75 (s, 3H); \(^{13}\)C NMR (100 MHz, CD\(_3\)Cl): \( \delta \) 140.0, 138.7, 132.1, 130.6, 128.6, 128.0, 127.7, 122.7, 72.4, 66.5, 43.7, 32.2, 26.7, 23.6, 21.8.
2-[2,6-Dimethyl-8-(toluene-4-sulfonyl)-octa-2,6-dienyloxy]-tetrahydro-pyran (2.42). A reaction vessel was charged with 2.87 (1.698 g, 6.676 mmol) and flushed with nitrogen. CH₂Cl₂ (20 mL) and THF (45 mL) were added, and the vessel was cooled to -40 °C. Triethylamine (1.21 mL, 8.68 mmol) was then added, and the reaction was stirred for 5 min. Dropwise addition of mesyl chloride (0.62 mL, 8.0 mmol) was followed by stirring of the reaction for 30 min. A solution of LiCl (3.096 g, 73.03 mmol) in THF (10 mL) was transferred to the vessel, and the vessel was allowed to warm to 0 °C over 3 h. The solvent was removed, and the remaining residue was diluted with 10% EtOAc in Hexanes (100 mL). The organic layer was washed with H₂O (50 mL) and saturated NaHCO₃ (50 mL), and then dried over NaSO₄. The crude product was used directly in the next step. Specifically, a reaction vessel was charged with crude 2.88 and DMF (40 mL). Sodium p-toluenesulfinate (4.496 g, 8.011 mmol) was then added in one portion, and the reaction was stirred for 3 h at room temperature. The solution was concentrated and the residue dissolved in CH₂Cl₂ (100 mL). This solution was then washed with water (50 mL), after which the organic layer was dried over NaSO₄. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 10% to 15%) to afford 2.42 as colorless oil (1.730 g, 66% over two steps). ¹H NMR (400 MHz, CDCl₃): δ 7.72 (d, J = 8.2 Hz, 2H), 7.30 (d, J = 8.2 Hz, 2H), 5.23-5.14 (m, 2H), 4.53 (t, J = 3.4 Hz, 1H), 4.03 (d, J = 11.4 Hz, 1H), 3.98 (d, J = 11.5 Hz, 1H), 3.88-3.80 (m, 1H), 3.76 (d, J = 8.1 Hz, 1H), 3.53-3.44 (m, 1H), 2.42 (s, 3H), 2.00-1.90 (m, 2H), 1.84-1.74 (m, 3H), 1.70 (s, 3H), 1.70 (s, 3H), 1.60-1.46
(m, 5H); $^{13}$C NMR (100 MHz, CDCl$_3$) δ 145.6, 144.6, 136.2, 133.0, 129.8, 128.7, 128.3, 111.5, 97.5, 65.3, 62.3, 56.1, 32.3, 30.9, 25.9, 25.7, 23.7, 21.9, 21.8, 19.7

\[
\text{BnO} \quad \text{Cl} \quad \text{SO} \quad \text{Cl}
\]

1-(16-Benzylkoxy-1-chloro-2,6,10,14-tetramethyl-hexadeca-2,6,10,14-tetraene-8-sulfonyl)-4-methyl-benzene (2.44). A reaction vessel was charged with 2.90 (0.206 g, 0.374 mmol) and flushed with nitrogen. CH$_2$Cl$_2$ (6 mL) and THF (3 mL) were added, and the vessel was cooled to -40 °C. Triethylamine (0.070 mL, 0.49 mmol) was then added, and the reaction was stirred for 5 min. Dropwise addition of mesyl chloride (0.040 mL, 0.45 mmol) was followed by stirring of the reaction for 30 min. A solution of LiCl (0.200 g, 4.72 mmol) in THF (10 mL) was transferred to the vessel, and the vessel was allowed to warm to 0 °C over 3 h. The solvent was removed, and the remaining residue was diluted with 10% EtOAc in Hexanes (20 mL). The organic layer was washed with H$_2$O (10 mL) and saturated NaHCO$_3$ (10 mL), and then dried over NaSO$_4$. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5% to 11%) to afford 2.41 as a colorless oil (0.180 g, 85%). $^1$H NMR (400 MHz, CDCl$_3$): δ 7.70 (d, $J = 8.3$ Hz, 2H), 7.34-7.25 (m, 7H), 5.39 (td, $J = 6.8$ Hz, $J = 1.0$ Hz, 1H), 5.21-5.10 (m, 2H), 4.97 (d, $J = 10.5$ Hz, 1H), 4.47 (s, 3H), 3.99-3.91 (m, 4H), 3.82 (td, $J = 10.9$ Hz, $J = 2.8$ Hz, 1H), 2.64 (dd, $J = 13.2$ Hz, $J = 2.4$ Hz, 1H), 2.41 (s, 3H), 2.03-1.94 (m, 4H), 1.93-1.79 (m, 2H), 1.76 (s, 3H), 1.69 (s, 3H), 1.66 (s, 3H), 1.65-1.58 (m, 2H), 1.57 (s, 3H); $^{13}$C NMR (100 MHz, CDCl$_3$)
1-Methyl-4-(3,7,11-trimethyl-dodeca-2,6,10-triene-1-sulfonyl)-benzene (2.45). A reaction vessel was charged with Farnesyl bromide (1.9235 g, 6.742 mmol) and DMF (20 mL). Sodium p-toluenesulfinate (1.4417 g, 8.0908 mmol) was then added in one portion, and the reaction was stirred for 3 h at room temperature. The solution was concentrated and the residue dissolved in CH₂Cl₂ (100 mL). This solution was then washed with water (50 mL), after which the organic layer was dried over Na₂SO₄. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5% to 12%) to afford 2.45 as a yellow oil (2.2094 g, 91% yield). ¹H NMR (400 MHz, CD₃Cl): δ 7.71 (d, J = 8.3 Hz, 2H), 7.29 (d, J = 8.2 Hz, 2H), 5.15 (t, J = 7.9 Hz, 1H), 5.09-4.99 (m, 2H), 3.76 (d, J = 7.9 Hz, 2H), 2.41 (s, 3H), 2.08-1.90 (m, 8H), 1.65 (s, 3H), 1.57 (s, 3H), 1.55 (s, 3H), 1.31 (s, 3H); ¹³C NMR (100 MHz, CD₃Cl): δ 146.4, 144.6, 136.1, 136.0, 131.6, 129.7, 128.7, 124.4, 123.6, 110.7, 56.3, 39.9, 26.9, 26.4, 25.9, 21.8, 17.9, 16.4, 16.2.

3,7,11,15,19,23,27-Heptamethyl-octacosa-2,6,10,14,18,22,26-heptaen-1-ol (2.47). Ethylamine (25 mL) was condensed in a reaction vessel under nitrogen at -78 °C. Lithium metal (114 mg, 16.4
mmol) was then added, after which the mixture was stirred for 1 h to afford a dark blue solution. Crude 2.46 (977 mg, 1.09 mmol) in anhydrous THF (5 mL) was added dropwise over 1 min. The reaction was immediately quenched with NH₄Cl and then brought to room temperature. Following evaporation of ethylamine, the residue was dissolved in EtOAc (35 mL) and washed with water. The organic layer was subsequently dried over Na₂SO₄, and concentrated. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 2% to 5%) to afford Heptaprenol 2.47 as a colorless oil (193 mg, 35%). ¹H NMR (500 MHz, CD₃Cl): δ 5.42 (td, J = 7.1 Hz, J = 1.2 Hz, 1H), 5.17-5.04 (m, 6H), 4.07 (d, J = 7.1 Hz, 2H), 2.09-2.00 (m, 20H), 1.99-1.93 (m, 4H), 1.73 (s, 3H), 1.67 (s, 12H), 1.59 (s, 3H), 1.58 (s, 6H); ¹³C NMR (125 MHz, CD₃Cl): δ 140.1, 136.3, 135.7, 135.6, 135.4, 135.2, 131.5, 125.2, 125.1, 124.8, 124.7, 124.6, 124.5, 124.4, 59.3, 40.0, 40.0, 32.5 (2), 32.4, 32.2, 27.0, 26.9, 26.9, 26.7, 26.6, 26.6, 25.9, 23.7 (2), 23.7, 23.6, 17.9, 16.2

3,7-Dimethyl-8-(tetrahydro-pyran-2-yloxy)-octa-2,6-dien-1-ol (2.87). NH₃ (g) (100 mL) was condensed in a reaction vessel at -78 °C. Na (2.002 g, 87.09 mmol) was added and the reaction was stirred for 1 h. A THF (2 mL) solution of 2.39 (3.000 g, 8.709 mmol) was then added dropwise. The reaction was stirred for 5 h before it was quenched with CH₃OH (50 mL). NH₃ (l) was allowed to boil off over night, after which the remaining solvent was removed. The remaining residue was dissolved in CH₂Cl₂ (100 mL). The organic layer was washed with water (50 mL) and dried over Na₂SO₄. The crude product was subjected to flash column chromatography
(EtOAc in Hexanes, 10% to 50%) to afford **2.87** as a colorless oil (1.698 g, 77%). $^1$H NMR (400 MHz, CDCl$_3$): $\delta$ 5.42 (t, $J = 7.1$ Hz, 1H), 5.34 (t, $J = 7.0$ Hz, 1H), 4.57 (t, $J = 3.4$ Hz, 1H), 4.12-3.98 (m, 4H), 3.90-3.78 (m, 1H), 3.54-3.44 (m, 1H), 2.20-2.00 (m, 4H), 1.88-1.78 (m, 1H), 1.75 (s, 3H), 1.72 (s, 3H), 1.70-1.46 (m, 6H); $^{13}$C NMR (100 MHz, CDCl$_3$) $\delta$ 139.5, 132.8, 129.1, 125.0, 97.5, 65.4, 62.2, 59.1, 32.3, 30.8, 26.6, 25.7, 23.8, 22.0, 19.5.

8-Benzylonoxy-2,6-dimethyl-octa-2,6-dien-1-ol (2.89).

A reaction vessel was charged with **2.39** (1.800 g, 5.225 mmol). CH$_3$OH (50 mL) was added followed by $p$-toluenesulfonic acid monohydrate (0.350 g, 1.84 mmol). The reaction was stirred at room temperature overnight. The solvent was removed, and the remaining residue was dissolved in CH$_2$Cl$_2$ (75 mL). The organic layer was washed with saturated NaHCO$_3$ (30 mL) and brine (30 mL), after which it was dried over Na$_2$SO$_4$. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5% to 30%) to afford **2.89** as a colorless oil (1.3379 g, 98% yield). $^1$H NMR (500 MHz, CD$_3$Cl): $\delta$ 7.36-7.30 (m, 4H), 7.29-7.25 (m, 1H), 5.43 (t, $J = 6.9$ Hz, 1H), 5.23 (t, $J = 7.1$ Hz, 1H), 4.48 (s, 2H), 4.02 (s, 3H), 3.95 (d, $J = 6.9$ Hz, 2H), 2.14-2.04 (m, 4H), 1.75 (s, 3H), 1.74 (s, 3H); $^{13}$C NMR (125 MHz, CD$_3$Cl): $\delta$ 140.5, 138.6, 135.4, 128.6, 128.0, 127.8, 127.6, 122.4, 72.4, 66.5, 61.6, 32.5, 26.3, 23.8, 21.5.
16-Benzylxoy-2,6,10,14-tetramethyl-8-(toluene-4-sulfonyl)-hexadeca-2,6,10,14-tetraen-1-ol (2.90). A reaction vessel was charged with **2.42** (0.299 g, 0.762 mmol) and flushed with argon. Freshly distilled THF (6 mL) and HMPA (2 mL) were then added, after which the vessel was cooled to -78 °C. *n*-BuLi (0.48 mL, 0.76 mmol) was added dropwise and the reaction was stirred for 10 min. **2.41** (0.177 g, 0.635 mmol) in THF (6 mL) was then added dropwise over 1 h via an automated injector. After warming overnight to room temperature, the reaction was quenched via the addition of 1% AcOH (aq.) (50 mL). The product was subsequently extracted with diethyl ether (100 mL), and the organic layer was dried over NaSO₄. The crude product (**2.43**) was utilized directly in the next step. Specifically, a reaction vessel was charged with crude **2.43** (0.250 g, 0.393 mmol). CH₃OH (15 mL) was added followed by *p*-toluenesulfonic acid monohydrate (0.095 g, 0.55 mmol). The reaction was stirred at room temperature overnight. The solvent was removed, and the remaining residue was dissolved in CH₂Cl₂ (30 mL). The organic layer was washed with saturated NaHCO₃ (15 mL) and brine (15 mL), after which it was dried over Na₂SO₄. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 7% to 15%) to afford **2.90** as a colorless oil (0.206 g, 95% yield). ¹H NMR (500 MHz, CD₃Cl): δ 7.68 (d, *J* = 8.2 Hz, 2H), 7.31 (d, *J* = 4.4 Hz, 4H), 7.30-7.25 (m, 3H), 5.39 (t, *J* = 6.8 Hz, 1H), 5.13 (t, *J* = 6.0 Hz, 1H), 5.09 (t, *J* = 7.1 Hz, 1H), 4.92 (d, *J* = 10.4 Hz, 1H), 4.47 (s, 2H), 4.04 (d, *J* = 11.8 Hz, 1H), 3.98 (d, *J* = 11.9 Hz, 1H), 3.95 (d, *J* = 6.8 Hz, 2H), 3.82 (td, *J*
= 10.8 Hz, \( J = 2.8 \) Hz, 1H), 2.63 (dd, \( J = 13.2 \) Hz, \( J = 2.4 \) Hz, 1H), 2.41 (s, 3H), 2.04-1.93 (m, 4H), 1.92-1.83 (m, 2H), 1.82-1.74 (m, 2H), 1.73 (s, 3H), 1.69 (s, 3H), 1.65 (s, 3H), 1.64-1.57 (m, 2H), 1.56 (s, 3H); \(^{13}\)C NMR (125 MHz, CD\(_3\)Cl): \( \delta \) 145.1, 144.7, 140.2, 138.7, 135.4, 135.1, 130.7, 129.6, 129.4, 128.5, 128.3, 127.7, 127.3, 122.4, 117.9, 72.3, 66.5, 63.6, 61.2, 32.6, 32.3, 30.8, 26.7, 25.9, 23.9, 23.6, 23.6, 21.8, 21.5.

1-(1-Benzyllox-3,7,11,15,19-pentamethyl-icosa-2,6,10,14,18-pentaene-9-sulfonyl)-4-methyl-benzene (2.91). A reaction vessel was charged with 2.45 (465 mg, 1.29 mmol) and flushed with argon. Freshly distilled THF (6 mL) and HMPA (2 mL) were then added, after which the vessel was cooled to -78 °C. \( n \)-BuLi (0.806 mL, 1.29 mmol) was added dropwise and the reaction was stirred for 10 min. 2.41 (300 mg, 1.08 mmol) in THF (6 mL) was then added dropwise over 1 h via an automated injector. After warming overnight to room temperature, the reaction was quenched via the addition of 1% AcOH (aq.) (50 mL). The product was subsequently extracted with diethyl ether (100 mL), and the organic layer was dried over NaSO\(_4\). The organic layer was concentrated, and the crude product was purified via flash column chromatography (EtOAc in hexanes, 3% to 5%) to afford 2.91 as a colorless oil (400 mg, 64%). \(^1\)H NMR (400 MHz, CD\(_3\)Cl): \( \delta \) 7.69 (d, \( J = 8.2 \) Hz, 2H), 7.34-7.30 (m, 4H), 7.29-7.23 (m, 3H), 5.40 (td, \( J = 6.8 \) Hz, \( J = 1.1 \) Hz, 1H), 5.13 (t, \( J = 5.6 \) Hz, 1H), 5.04 (dtt, \( J = 18.3 \) Hz, \( J = 6.8 \) Hz, \( J = 1.2 \) Hz, 2H), 4.92 (d, \( J = 10.4 \) Hz, 1H), 163
4.47 (s, 2H), 3.97 (d, J = 6.6 Hz, 2H), 3.82 (td, J = 10.8 Hz, J = 3.1 Hz, 1H), 2.74 (dd, J = 13.3 Hz, J = 2.6 Hz, 1H), 2.46 (dd, J = 13.2 Hz, J = 11.4 Hz, 1H), 2.41 (s, 3H), 2.10-1.99 (m, 6H), 1.71 (s, 3H), 1.66 (s, 3H), 1.57 (s, 6H), 1.56 (s, 3H), 1.17 (s, 3H); $^{13}$C NMR (100 MHz, CD$_3$Cl): $\delta$ 145.3, 144.5, 140.1, 138.8, 135.8, 135.2, 131.6, 130.9, 129.5, 129.4, 128.5, 128.1, 128.0, 127.7, 124.4, 123.6, 122.4, 117.3, 72.3, 66.6, 63.8, 40.0, 39.9, 32.4, 26.9, 26.7, 26.4, 25.9, 23.7, 23.7, 23.0, 21.8, 17.9, 16.6, 16.1.

3,7,11,15,19-Pentamethyl-icosa-2,6,10,14,18-pentaen-1-ol (2.92). Ethylamine (10 mL) was condensed in a reaction vessel under nitrogen at -78 °C. Lithium metal (57 mg, 8.2 mmol) was then added, after which the mixture was stirred for 1 h to afford a dark blue solution. 2.91 (330 mg, 0.547 mmol) in anhydrous THF (1 mL) was added dropwise over 1 min. The reaction was immediately quenched with NH$_4$Cl and then brought to room temperature. Following evaporation of ethylamine, the residue was dissolved in EtOAc (20 mL) and washed with water. The organic layer was subsequently dried over Na$_2$SO$_4$, and concentrated. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5% to 10%) to afford Pentaprenol 2.92 as a colorless oil (145 mg, 74%). $^1$H NMR (500 MHz, CD$_3$Cl): $\delta$ 5.42 (t, J = 7.1 Hz, 1H), 5.16-5.04 (m, 4H), 4.07 (d, J = 7.1 Hz, 2H), 2.12-2.01 (m, 12H), 2.00-1.92 (m, 4H), 1.72 (s, 3H), 1.67 (s, 3H), 1.66 (s, 3H), 1.59 (s, 3H), 1.58 (s, 6H); $^{13}$C NMR (125 MHz, CD$_3$Cl): $\delta$ 140.0, 136.4, 135.5, 135.2, 131.5, 124.7, 124.7, 124.6, 124.4, 124.2, 59.2, 39.9 (2), 32.4, 32.2, 27.0, 26.8, 26.8, 26.5, 25.9, 23.6, 23.6, 17.9, 16.2.
1-Methyl-4-(3-methyl-but-2-ene-1-sulfonyl)-benzene (2.93). A reaction vessel was charged with 3,3-Dimethylallyl bromide (2.0000 g, 13.491 mmol) and DMF (20 mL). Sodium p-toluenesulfinate (2.8694 g, 16.103 mmol) was then added in one portion, and the reaction was stirred for 3 h at room temperature. The solution was concentrated and the residue dissolved in CH₂Cl₂ (100 mL). This solution was then washed with water (50 mL), after which the organic layer was dried over Na₂SO₄. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 10% to 15%) to afford 2.93 as a crystalline white solid (2.7280 g, 91%). ¹H NMR (500 MHz, CD₃Cl): δ 7.71 (d, J = 8.2 Hz, 2H), 7.30 (d, J = 8.0 Hz, 2H), 5.15 (t, J = 7.9 Hz, 1H), 3.74 (d, J = 7.9 Hz, 2H), 2.42 (s, 3H), 1.69 (s, 3H), 1.32 (s, 3H); ¹³C NMR (125 MHz, CD₃Cl): δ 144.6, 142.9, 136.2, 129.8, 128.7, 110.8, 56.5, 26.0, 21.8, 18.0.

3,7,11,15,19-Pentamethyl-icosa-2,6,10,14,18-pentaen-1-ol (2.95). Ethylamine (25 mL) was condensed in a reaction vessel under nitrogen at -78 °C. Lithium metal (63 mg, 9.1 mmol) was then added, after which the mixture was stirred for 1 h to afford a dark blue solution. Crude 2.94 (459 mg, 0.606 mmol) in anhydrous THF (4 mL) was added dropwise over 1 min. The reaction was immediately quenched with NH₄Cl and then brought to room temperature. Following evaporation of ethylamine, the residue was dissolved in EtOAc (25 mL) and washed with water. The organic layer was subsequently dried over NaSO₄, and concentrated. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5% to
10%) to afford cis-Pentaprenol 2.95 as a colorless oil (98 mg, 45%). $^1$H NMR (500 MHz, CD$_3$Cl): $\delta$ 5.43 (t, $J = 7.1$ Hz, 1H), 5.17-5.02 (m, 4H), 4.07 (d, $J = 6.4$ Hz, 2H), 2.13-1.92 (m, 16H), 1.73 (s, 3H), 1.67 (s, 12H), 1.59 (s, 3H); $^{13}$C NMR (125 MHz, CD$_3$Cl): $\delta$ 140.1, 136.3, 135.6, 135.6, 131.8, 125.2, 125.1, 124.8, 124.7, 124.5, 59.3, 32.5, 32.4, 32.2, 29.9, 26.9, 26.6, 26.6, 26.5, 23.7, 23.7, 23.6, 23.6, 23.6, 17.9; HRMS (ESI) calcd for C$_{25}$H$_{42}$ONa [M+Na]$^+$ 381.3128, found 381.3124.

(3,7-Dimethyl-oct-6-enyloxymethyl)-benzene (2.97). A reaction vessel was charged with sodium hydride (1.5358 g, 63.992 mmol) and flushed with argon. THF (50 mL) was added to the flask, after which the vessel was cooled to 0 °C. 2.96 (10.050 g, 63.992 mmol) was then added dropwise to the solution. The reaction was allowed to proceed for 30 min at 0 °C, at which point benzyl bromide (10.944 g, 63.992 mmol) and sodium iodide (0.9592 g, 6.399 mmol) were introduced into the reaction. The reaction was heated to reflux and stirred for 5 h. After quenching the reaction with 1.0 M HCl (aq.), the desired product was extracted with a Hexane : EtOAc mixture (10:1). The organic layer was dried over Na$_2$SO$_4$. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 2.5%) to afford 2.97 as a colorless oil (12.845 g, 81%). $^1$H NMR (500 MHz, CD$_3$Cl): $\delta$ 7.33 (d, $J = 4.4$ Hz, 4H), 7.28-7.23 (m, 1H), 5.09 (t, $J = 7.1$ Hz, 1H), 4.49 (s, 2H), 3.54-3.44 (m, 2H), 2.04-1.89 (m, 2H), 1.72-1.52 (m, 2H), 1.67 (s, 3H), 1.59 (s, 3H), 1.47-1.39 (m, 1H), 1.37-1.29 (m, 1H), 1.20-1.10 (m, 1H), 0.80 (d, $J = 6.6$ Hz, 3H); $^{13}$C NMR (125 MHz, CD$_3$Cl): $\delta$ 139.0, 131.3, 128.5, 127.8, 127.7, 125.1, 73.1, 69.0, 37.4, 37.0, 29.8, 25.9, 25.7, 19.8, 17.8..
6-Benzylxy-4-methyl-hexan-1-ol (2.99). A reaction vessel was charged with 2.97 (7.5270 g, 30.549 mmol). Dioxane (75 mL) and H₂O (10 mL) were added to the vessel, after which the reaction was cooled to 0 °C. N-Bromosuccinimide (5.4373 g, 30.549 mmol) was then added in three portions, and the reaction was stirred for 5 h. A solution of KOH (1.7166 g, 30.549 mmol) in MeOH (10 mL) was added dropwise to the vessel. After stirring for 3 h, the reaction mixture was diluted with water (100 mL) and extracted with Et₂O (100 mL). The organic layer was dried over Na₂SO₄. The crude product was concentrated and utilized directly in the next step. Specifically, a reaction vessel was charged with the epoxide, after which THF (75 mL) and water (15 mL) were added. Sodium periodate (3.2670 g, 15.274 mmol) and periodic acid (8.3561 g, 36.659 mmol) were then added in sequence. The reaction was stirred for 30 min at room temperature and subsequently diluted with saturated Na₂SO₄ (aq.) (60 mL). The crude product was extracted with EtOAc (100 mL). The organic layer was dried over Na₂SO₄ before being concentrated. The crude product was utilized directly in the next step. Specifically, a reaction vessel was charged with the aldehyde. CH₃OH (75 mL) was then added, and the reaction was cooled to 0 °C. NaBH₄ (0.9246 g, 24.44 mmol) was added in three portions. The reaction was stirred for 30 min, after which the solvent was removed. The residue was dissolved in EtOAc (100 mL) and washed with water (50 mL). The organic layer was dried over Na₂SO₄. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 15% to 20%) to afford 2.99 as a colorless oil (2.1734 g, 32% over three steps). ¹H NMR (500 MHz, CD₃Cl): δ 7.32 (d, J = 4.6 Hz, 1H), 7.29-7.25 (m, 1H), 4.48 (s, 2H), 3.55 (t, J = 6.7 Hz, 2H), 1.27 (s, 3H), 0.96 (t, J = 7.2 Hz, 3H).
2H), 3.53-3.45 (m, 2H), 2.06 (bs, 1H), 1.70-1.32 (m, 6H), 1.20-1.12 (m, 1H), 0.88 (d, J = 6.6 Hz, 3H); $^{13}$C NMR (125 MHz, CD$_3$Cl): δ 138.7, 128.5, 127.8, 127.6, 73.0, 68.7, 63.1, 36.8, 33.1, 30.2, 29.8, 19.7.

![Chemical Structure of (6-Benzylloxoy-4-methyl-hexyl)-triphenyl-phosphonium Iodide](image)

**iodide (2.100).** A reaction vessel was charged with **2.99** (1.500 g, 6.747 mmol) and benzene (30 mL). Triphenylphosphine (2.301 g, 8.771 mmol), imidazole (0.826 g, 12.1 mmol) and iodine (2.055 g, 8.096 mmol) were then added sequentially. The reaction was stirred for 20 min, after which the solvent was removed. The residue was diluted with ether (20 mL) and washed with H$_2$O (10 mL), 20% Citric Acid (10 mL), sat. NaHCO$_3$ (aq.) (10 mL) and brine (10 mL). The organic layer was concentrated, suspended in hexanes and then filtered through a short bed of silica gel. The solvent was removed, and the crude product was utilized directly in the next step. Specifically, a reaction vessel was charged with the crude product (2.228 g, 6.747 mmol) and triphenylphosphine (2.478 g, 9.446 mmol). Benzene (10 mL) was added, and the solution was brought to reflux and stirred for 20 h. The solvent was then removed, yielding a white powder. The powder was triturated with Et$_2$O to remove residual triphenylphosphine and afforded **2.100** as a white powder (1.795 g, 44% yield over two steps).

![Chemical Structure of 2-(8-Benzylloxoy-2,6-dimethyl-oct-2-enyloxy)-tetrahydropyran](image)

**2-(8-Benzylloxoy-2,6-dimethyl-oct-2-enyloxy)-tetrahydropyran (2.101).** A reaction vessel was charged with **2.100** (0.8500 g, 1.430
mmol) and flushed with nitrogen. THF was added (10 mL), and the vessel was cooled to -78 °C. n-BuLi (0.76 mL, 1.2 mmol) was then added dropwise, after which the reaction was stirred for 30 min. 2.38 (0.2715 g, 1.716 mmol) was added dropwise, and the reaction mixture was allowed to warm to room temperature overnight. The reaction was quenched with saturated NH₄Cl (aq.) (10 mL), after which the product was extracted with EtOAc (30 mL). The organic layer was dried over Na₂SO₄. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 1% to 10%) to afford 2.101 as a colorless oil (0.3000 g, 61%). ¹H NMR (500 MHz, CD₃Cl): δ 7.36-7.30 (m, 4H), 7.28-7.25 (m, 1H), 5.33 (t, J = 7.2 Hz, 1H), 4.56 (t, J = 3.6 Hz, 1H), 4.48 (s, 2H), 4.10 (d, J = 11.3 Hz, 1H), 4.05 (dd, J = 11.4 Hz, J = 2.1 Hz, 1H), 3.90-3.82 (m, 1H), 3.52-3.44 (m, 3H), 2.13-1.96 (m, 2H), 1.88-1.77 (m, 1H), 1.74 (3H), 1.72-1.48 (m, 7H), 1.44-1.29 (m, 2H), 1.20-1.11 (m, 1H), 0.86 (d, J = 6.5 Hz, 3H); ¹³C NMR (125 MHz, CD₃Cl): δ 138.9, 131.8, 130.0, 128.5, 127.8, 127.7, 97.8, 73.1, 68.9, 65.6, 62.4, 37.6, 36.9, 30.9, 29.8, 25.7, 25.4, 22.0, 19.8, 19.7.

8-Benzylxy-2,6-dimethyl-oct-2-en-1-ol (2.102). A reaction vessel was charged with 2.101 (0.3000 g, 0.8658 mmol). CH₃OH (10 mL) was added followed by p-toluenesulfonic acid monohydrate (0.1980 g, 1.039 mmol). The reaction was stirred at room temperature overnight. The solvent was removed, and the remaining residue was dissolved in CH₂Cl₂ (30 mL). The organic layer was washed with saturated NaHCO₃ (15 mL) and brine (15 mL), after which it was dried over Na₂SO₄. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5% to
12%) to afford 2.102 as a colorless oil (0.2181 g, 96%). $^1$H NMR (500 MHz, CD$_3$Cl): $\delta$ 7.35-7.29 (m, 4H), 7.28-7.24 (m, 1H), 5.24 (t, $J = 7.3$ Hz, 1H), 4.48 (s, 2H), 4.14 (d, $J = 11.8$ Hz, 1H), 4.03 (d, $J = 11.8$ Hz, 1H), 3.52-3.42 (m, 2H), 2.16-2.06 (m, 1H), 2.03-1.94 (m, 1H), 1.77 (s, 3H), 1.72-1.54 (m, 3H), 1.38-1.27 (m, 2H), 1.24-1.14 (m, 1H), 0.85 (d, $J = 6.6$ Hz, 3H); $^{13}$C NMR (125 MHz, CD$_3$Cl): $\delta$ 138.7, 134.5, 128.7, 128.6, 127.9, 127.7, 73.1, 68.5, 61.7, 37.4, 36.5, 29.2, 25.1, 21.5, 19.8.

(8-Bromo-3,7-dimethyl-oct-6-enyloxymethyl)-benzene

(2.103). A reaction vessel was charged with 2.102 (0.218 g, 0.831 mmol) and flushed with nitrogen. CH$_2$Cl$_2$ (15 mL) was added, and the vessel was cooled to -40 °C. Triethylamine was then added, and the reaction was stirred for 5 min. Dropwise addition of mesyl chloride (0.080 mL, 1.0 mmol) was followed by stirring of the reaction for 30 min. A solution of LiBr (0.3568 g, 4.156 mmol) in THF (2.8 mL) was transferred to the vessel, and the vessel was allowed to warm to 0 °C over 3 h. The solvent was removed, and the remaining residue was diluted with EtOAc (30 mL). The organic layer was washed with saturated NaHCO$_3$ (15 mL) and dried over Na$_2$SO$_4$. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 0% to 2.5%) to afford 2.103 as a pale yellow oil (0.2680 g, 99%). $^1$H NMR (500 MHz, CD$_3$Cl): $\delta$ 7.37-7.30 (m, 4H), 7.29-7.25 (m, 4H), 5.36 (td, $J = 1.4$ Hz, $J = 7.4$ Hz, 1H), 4.48 (s, 2H), 3.96 (s, 2H), 3.55-3.45 (m, 2H), 2.14-1.98 (m, 2H), 1.81 (s, 3H), 1.73-1.55 (m, 2H), 1.47-1.35 (m, 2H), 1.26-1.17 (m, 1H), 0.891 (d, $J = 6.6$ Hz, 3H); $^{13}$C NMR (125 MHz, CD$_3$Cl): $\delta$ 138.8, 132.1, 131.5, 128.5, 127.8, 127.7, 73.1, 68.8, 36.8, 36.8, 32.6, 29.7, 25.7, 22.1, 19.7.
1-(1-Benzylhydroxy-3,7,11,15,19-pentamethyl-icosa-6,10,14,18-tetraene-9-sulfonyl)-4-methyl-benzene (2.104). A reaction vessel was charged with 2.45 (357 mg, 0.989 mmol) and flushed with argon. Freshly distilled THF (9 mL) and HMPA (3 mL) were then added, after which the vessel was cooled to -78 °C. n-BuLi (0.62 mL, 0.99 mmol) was added dropwise, and the reaction was stirred for 10 min. 2.103 (268 mg, 0.824 mmol) in THF (9 mL) was then added dropwise over 1 h via an automated injector. After warming overnight to room temperature, the reaction was quenched via the addition of 1% AcOH (aq.) (30 mL). The product was subsequently extracted with diethyl ether (75 mL), and the organic layer was dried over NaSO₄. The organic layer was concentrated, and the crude product was purified via flash column chromatography (EtOAc in hexanes, 4% to 6%) to afford 2.104 as a colorless oil. ¹H NMR (500 MHz, CD₃Cl): δ 7.70 (d, J = 8.2 Hz, 2H), 7.32 (d, J = 4.4 Hz, 4H), 7.29-7.25 (m, 3H), 5.15 (t, J = 7.1 Hz, 1H), 5.06 (tt, J = 7.5 Hz, J = 1.3 Hz, 1H), 5.02 (t, J = 6.0 Hz, 1H), 4.92 (d, J = 10.4 Hz, 1H), 4.48 (s, 2H), 3.84 (td, J = 10.8 Hz, J = 3.0 Hz, 1H), 3.52-3.43 (m, 2H), 2.72 (dt, J = 14.2 Hz, J = 4.2 Hz, 1H), 2.54-2.44 (m, 1H), 2.40 (s, 3H), 2.08-1.80 (m, 11H), 1.72-1.49 (m, 1H), 1.66 (s, 3H), 1.58 (s, 6H), 1.56 (s, 3H), 1.44-1.35 (m, 4H), 1.33-1.23 (m, 1H), 1.19 (s, 3H), 1.14-1.04 (m, 1H), 0.85 (d, J = 6.6 Hz, 3H); ¹³C NMR (125 MHz, CD₃Cl): δ 145.3, 144.5, 138.9, 135.8, 135.3, 131.6, 130.1, 129.6, 129.2, 128.6, 127.8, 127.7, 124.5, 123.7, 117.3, 73.1, 68.9,
Ethylamine (30 mL) was condensed in a reaction vessel under nitrogen at -78 °C. Lithium metal (89 mg, 13 mmol) was then added, after which the mixture was stirred for 1 h to afford a dark blue solution. 2.104 (517 mg, 0.855 mmol) in anhydrous THF (2 mL) was added dropwise over 1 min. The reaction was immediately quenched with NH₄Cl and then brought to room temperature. Following evaporation of ethylamine, the residue was dissolved in EtOAc (25 mL) and washed with water. The organic layer was subsequently dried over NaSO₄, and concentrated. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5% to 7%) to afford MS-Pentaprenol 2.105 as a colorless oil (185 mg over two steps, 62%). 1H NMR (500 MHz, CD₃Cl): δ 5.15-5.05 (m, 4H), 3.72-3.60 (m, 2H), 2.10-2.01 (m, 8H), 2.00-1.90 (m, 6H), 1.66 (s, 3H), 1.66 (s, 3H), 1.62-1.52 (m, 2H), 1.59 (s, 3H), 1.58 (s, 6H), 1.41-1.27 (m, 2H), 1.24-1.12 (m, 1H), 0.89 (d, J = 6.6 Hz, 3H); 13C NMR (125 MHz, CD₃Cl): δ 135.4, 135.3, 135.2, 131.5, 125.6, 124.6, 124.5, 124.4, 61.4, 40.2, 39.9, 37.7, 32.2, 29.5, 27.0, 26.9, 26.8, 25.9, 25.5, 23.6, 19.7, 17.9, 16.2, 16.2.
5.4 RU-PP-Polypropenol enzymatic synthesis and product characterization

The enzymatic synthesis procedures for 2.19-2.24 have been previously described for the example of RU-PP-Und (2.19) (Wen Yi, doctoral dissertation); however, the product characterizations are given below. MALDI-Mass spectra for the di- (a), tri- (b), tetra- (c) and pentasaccharide (d) products in route to 2.19. The Und-PP moiety was hydrolyzed, and the products were labeled with 2-AB.

a)
b) Calculated: M+K 745.7  
Measured:    M+K 745.7

c) Calculated: M+Na 875.8  
Measured:    M+Na 875.8
Calculated: M+Na  1038.0
Measured:    M+Na  1037.7
ESI-MS for RU-PP-Lipid analogs (2.19-2.24).

RU-PP-Undecaprenyl (2.19):
LRMS (m/z): [M-2H]2- calcd. for C_{89}H_{146}N_{2}O_{31}P_{2}, 900.5; found, 900.4.

RU-PP-Solanesyl (2.23):
LRMS (m/z): [M-2H]2- calcd. for C_{79}H_{130}N_{2}O_{31}P_{2}, 832.4; found, 832.4.

RU-PP-Heptaprenyl (2.20):
LRMS (m/z): [M-2H]2- calcd. for C_{69}H_{114}N_{2}O_{31}P_{2}, 764.3; found, 764.4.
RU-PP-Pentaprenyl (2.21):
LRMS ($m/z$): [M-H]- calcd. for $C_{59}H_{99}N_2O_{31}P_2$, 1393.6; found, 1393.5.

RU-PP-cis-Pentaprenyl (2.22):
LRMS ($m/z$): [M-H]- calcd. for $C_{59}H_{99}N_2O_{31}P_2$, 1393.6; found, 1393.5.
5.5 Polymerization assays

Detailed protocols for the polymerization assays have been previously described for the example of RU-PP-Und (2.19) (Wen Yi, doctoral dissertation).

5.6 Chemical synthesis and characterization of the Stt3d substrate and associated precursors

Acetic acid 2-acetoxyethyl-5-acetylamino-6-(bis-benzyloxy-phosphoryloxy)-3-(4,5-diacetoxy-6-acetoxyethyl-3-acetylamino-tetrahydro-pyran-2-yloxy)-tetrahydro-pyran-4-yl ester (3.2). A reaction vessel was
charged 3.1 (254 mg, 0.375 mmol) and was flushed with nitrogen. DMF (2 mL) was then added to the vessel followed by dimethylamine (5 mL, 10 mmol). The reaction was allowed to proceed overnight at room temperature. The reaction mixture was concentrated, and the resulting residue was utilized directly in the subsequent step. Specifically, a reaction vessel was charged with the crude product and tetrazole (263 mg, 3.75 mmol), after which it was flushed with argon. CH2Cl2 (5 mL) was added, and the reaction was cooled to -40 °C. Dibenzyldiisopropylphosphoramidite (0.62 mL, 5.0 mmol) was added dropwise to the vessel, and the reaction was allowed to warm to room temperature over 3 h. The reaction was then cooled to -78 °C, and mCPBA was added (647 mg, 3.75 mmol). The reaction was allowed to warm to room temperature overnight. The reaction mixture diluted with CH2Cl2 (10 mL) was then washed successively with saturated Na2SO3 (25 mL), water (25 mL) and brine (25 mL). The organic layer was dried over Na2SO4 and concentrated. The crude product was purified via flash column chromatography (CH3OH in CH2Cl2, 0% to 2%) to afford 3.2 as a white foam (215 mg, 64%). 1H NMR (500 MHz, CD3Cl): δ 7.38-7.28 (m, 10H), 6.05 (d, J = 8.9 Hz, 1H), 5.71 (d, J = 9.3 Hz, 1H), 5.60 (dd, J = 5.9 Hz, J = 3.3 Hz, 1H), 5.19 (dd, J = 10.4 Hz, J = 9.4 Hz, 1H), 5.12 (dd, J = 10.9 Hz, J = 9.2 Hz, 1H), 5.08-4.98 (m, 5H), 4.55 (d, J = 8.3 Hz, 1H), 4.34 (dd, J = 12.4 Hz, J = 4.2 Hz, 1H), 4.27-4.19 (m, 2H), 4.07 (dd, J = 12.2 Hz, J = 1.9 Hz, 1H), 3.97 (dd, J = 12.4 Hz, J = 2.2 Hz, 1H), 3.93 (ddd, J = 10.1 Hz, J = 3.5 Hz, J = 2.0 Hz, 1H), 3.82-3.75 (m, 1H), 3.70 (t, J = 9.6 Hz, 1H), 3.59 (ddd, J = 9.9 Hz, J = 4.1 Hz, J = 2.4 Hz, 1H), 2.03 (s, 3H), 2.02 (s, 3H), 1.99 (s, 3H), 1.97 (s, 3H), 1.96 (s, 3H), 1.89 (s, 3H), 1.67 (s, 3H); 13C NMR (125 MHz, CD3Cl): δ 171.2, 171.1, 170.9, 170.7, 170.5, 170.5, 169.5, 135.5 (d, J = 6.4 Hz), 135.4 (d, J = 7.0 Hz), 129.1, 129.1, 129.0,
128.9, 128.3, 128.2, 101.3, 96.2 (d, $J = 6.3$ Hz), 75.7, 72.6, 72.1, 70.8, 70.3, 70.1 (d, $J = 5.5$ Hz), 70.0 (d, $J = 5.4$ Hz), 68.3, 61.9, 61.6, 55.1, 52.1 (d, $J = 8.2$ Hz), 23.3, 22.9, 21.0, 20.8, 20.8 (2), 20.7; $^{31}$P NMR (162 MHz, CD$_3$Cl): $\delta$ -0.9.

**Peracetylated Chitobiose-PP-MS-Pentaprenyl (Precursor to 3.4).** A reaction vessel was charged with 3.2 (85 mg, 0.095 mmol), 20% Pd(OH)$_2$ on carbon (15 mg), and CH$_3$OH (5 mL). The vessel was placed under a H$_2$ atmosphere (1 atm), and the reaction was stirred for 30 min. Hünig’s base (0.1 mL) was then added to the reaction mixture followed by CH$_3$OH (5 mL). The reaction was stirred for 30 min before the solution was filtered through a pad of Celite. The crude product was concentrated and utilized directly in the next step. Specifically, 1,1'-carbonyldiimidazole (CDI) (23 mg, 0.14 mmol) was added to a reaction vessel containing 3.3 (92 mg, 0.095 mmol). The vessel was flushed with argon, after which anhydrous THF was added (3 L). The reaction was stirred at room temperature for 2 h. Addition of dry CH$_3$OH (40 µL) followed by stirring for an additional 1 h at room temperature served to remove excess CDI. The reaction mixture was concentrated, and then anhydrous THF (3 mL) was once again added. To a separate reaction vessel, MS-Pentaprenyl phosphate$^8$ (41 mg, 0.086 mmol) was added, and the vessel was then flushed with argon. The activated Peracetylated Chitobiose-1-PO$_4$$^{2-}$ was transferred into this vessel via syringe. The reaction was stirred at room temperature for 3 days, after which
the reaction mixture was concentrated. The residue was then dissolved in water. Purification via C-18 reverse phase column chromatography (Solvent A: CH$_3$CN; Solvent B = 0.85% NH$_4$HCO$_3$; A/B = 0%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% and 50% (10 mL each)) afforded Peracetylated Chitobiose-PP-MS-Pentaprenyl as a white solid following lyophilization. $^1$H NMR (500 MHz, CD$_3$OH): $\delta$ 5.53 (dd, $J = 7.2$ Hz, $J = 3.3$ Hz, 1H), 5.40 (t, $J = 9.8$ Hz, 1H), 5.26 (dd, $J = 10.5$ Hz, $J = 9.2$ Hz, 1H), 5.18-5.07 (m, 4H), 4.95 (t, $J = 9.7$ Hz, 1H), 4.57 (dd, $J = 9.9$ Hz, $J = 1.5$ Hz, 1H), 4.43 (dd, $J = 12.5$ Hz, $J = 4.1$ Hz, 1H), 4.26-4.20 (m, 2H), 4.08 (dd, $J = 12.3$ Hz, $J = 3.6$ Hz, 1H), 4.05-3.98 (m, 3H), 3.92 (t, $J = 9.6$ Hz, 1H), 3.80 (ddd, $J = 10.0$ Hz, $J = 4.0$ Hz, $J = 2.5$ Hz, 1H), 3.60-3.54 (m, 1H), 2.12-1.95 (m, 14H), 2.11 (s, 3H), 2.06 (s, 3H), 2.04 (s, 3H) 1.98 (s, 3H), 1.97 (s, 3H), 1.90 (s, 3H), 1.68 (s, 3H), 1.67 (s, 3H), 1.62 (s, 3H), 1.60 (s, 6H), 1.52-1.12 (m, 5H), 0.92 (d, $J = 6.6$ Hz, 3H); $^{13}$C NMR (125 MHz, CD$_3$OH): $\delta$ 174.3, 173.8, 172.8, 172.4, 172.3, 172.0, 171.5, 136.2, 136.1, 132.2, 126.8, 125.7 (2), 125.7 (2), 101.8, 95.7 (d, $J = 6.2$ Hz), 77.1, 73.7, 73.1, 73.0, 71.1, 70.2, 66.0 (d, $J = 6.0$ Hz), 63.5, 63.3, 56.9, 53.5 (d, $J = 8.4$ Hz), 41.1, 39.0, 39.0, 38.9, 33.1, 30.7, 28.0, 27.8, 27.8, 26.6, 26.1, 24.0, 23.2, 23.0, 21.3, 21.1, 20.9, 20.8, 20.8, 20.0, 18.0, 16.4, 16.3; $^{31}$P NMR (162 MHz, CD$_3$OH): $\delta$ -8.9 (d, $J = 19.0$ Hz), -11.9 (d, $J = 20.0$ Hz); LRMS (m/z): [M-2H]$^2$- calcd. for C$_{51}$H$_{80}$N$_2$O$_{22}$P$_2$, 567.2; found, 567.3.
Chitobiose-PP-MS-
Pentaprenyl (3.4). A reaction vessel was charged with Peracetylated Chitobiose-PP-MS-Pentaprenyl (49 mg, 0.042 mmol). A 0.1% NaOCH\textsubscript{3} in CH\textsubscript{3}OH solution (3 mL) was then added to the vessel. The reaction was stirred at room temperature for 40 min. The reaction mixture was concentrated, after which the residue was dissolved in water. Purification via C-18 reverse phase column chromatography (Solvent A: CH\textsubscript{3}CN; Solvent B = 1.7% NH\textsubscript{4}HCO\textsubscript{3}; A/B = 0%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45% and 50% (10 mL each)) provided 3.4 as a white solid following lyophilization (32 mg, 39% over two steps). \textsuperscript{1}H NMR (500 MHz, D\textsubscript{2}O): \(\delta\) 5.56 (dd, \(J = 6.9\) Hz, \(J = 2.3\) Hz, 1H), 5.26-5.08 (m, 4H), 4.72 (d, \(J = 8.3\) Hz, 1H), 4.12-3.94 (m, 6H), 3.89 (t, \(J = 11.0\) Hz, 1H), 3.85-3.74 (m, 4H), 3.70 (t, \(J = 10.0\) Hz, 1H), 3.64-3.56 (m, 1H), 3.53 (t, \(J = 9.2\) Hz, 1H), 2.13 (s, 3H), 2.12 (s, 3H), 2.12-2.06 (m, 8H), 2.05-1.95 (m, 6H), 1.73 (s, 3H), 1.70 (s, 3H), 1.65 (s, 3H), 1.62 (s, 6H), 1.56-1.40 (m, 2H), 1.38-1.14 (m, 3H), 0.96 (d, \(J = 6.2\) Hz, 3H); \textsuperscript{13}C NMR (125 MHz, D\textsubscript{2}O): \(\delta\) 174.2, 174.2, 134.8, 134.8, 134.5, 130.7, 125.7, 124.5, 124.3, 124.3, 101.4, 94.0 (d, \(J = 6.0\) Hz), 79.2, 76.0, 73.5, 71.6, 69.8, 69.5, 65.0 (d, \(J = 5.3\) Hz), 60.6, 59.8, 55.7, 53.3 (d, \(J = 8.4\) Hz), 39.6, 37.5, 37.4, 37.2, 31.9, 29.0, 26.7, 26.6, 26.5, 25.4, 25.0, 23.3, 22.3, 22.2, 18.8, 17.4, 15.8, 15.7; \textsuperscript{31}P NMR (162 MHz, D\textsubscript{2}O): \(\delta\) -10.0 (d, \(J = 16.1\) Hz), -11.9 (d, \(J = 17.6\) Hz); LRMS (m/z): [M-2H]^2- calcd. for C\textsubscript{41}H\textsubscript{70}N\textsubscript{2}O\textsubscript{17}P\textsubscript{2}, 462.2; found, 462.2.
5.7 Chemical synthesis and characterization of PglL substrates and associated precursors

GalNAc-PP-Geranylgeranyl (Precursor to 4.1). GalNAc-PP-Geranylgeranyl was prepared as described above. However, the two purification steps were reversed as the final product proved too unstable to obtain suitable spectral data. The crude product was thus purified via C-18 reverse phase column chromatography to afford Peracetylated GalNAc-PP-Geranylgeranyl as a white solid (12 mg, 17% yield). $^1$H NMR (500 MHz, D$_2$O): $\delta$ 5.63 (dd, $J = 7.3$ Hz, $J = 3.2$ Hz, 1H), 5.47-5.42 (m, 2H), 5.24 (dd, $J = 11.3$ Hz, $J = 3.2$ Hz, 1H), 5.17-5.07 (m, 3H), 4.63-4.58 (m, 1H), 4.57-4.52 (m, 3H), 4.23 (dd, $J = 10.9$ Hz, $J = 8.1$ Hz, 1H), 4.06 (dd, $J = 10.8$ Hz, $J = 5.8$ Hz, 1H), 2.20-1.94 (m, 12H), 2.13 (s, 3H), 2.01 (s, 3H), 2.00 (s, 3H), 1.92 (s, 3H), 1.71 (s, 3H), 1.67 (s, 3H), 1.61 (s, 3H), 1.60 (s, 6H); $^{13}$C NMR (125 MHz, D$_2$O): $\delta$ 174.4, 172.3, 172.2, 172.0, 140.9, 136.3, 136.0, 132.2, 125.6 (2), 125.4, 122.6 (d, $J = 9.2$ Hz), 96.5 (d, $J = 6.1$ Hz), 70.5, 68.7, 68.7, 64.0 (d, $J = 5.4$ Hz), 62.3, 40.9 (d, $J = 1.6$ Hz), 40.8, 27.9, 27.8, 27.7, 26.0, 23.0, 20.8, 20.7, 20.7, 18.3, 17.9, 16.7, 16.2 (m); $^{31}$P NMR (162 MHz, D$_2$O): $\delta$ -10.1 (d, $J = 21.8$ Hz), -13.0 (d, $J = 21.6$ Hz); LRMS (ESI) calcd for C$_{34}$H$_{54}$NO$_{15}$P$_2$ [M-H]$^-$ 778.3, found 778.3. Final deprotection with 0.1% NaOCH$_3$ in CH$_3$OH and filtration though a
short bed of C-18 reverse phase silica gel afforded GalNAc-PP-Geranylgeranyl (3 mg, 30%). LRMS (ESI) calcd for \( \text{C}_{28}\text{H}_{48}\text{NO}_{12}\text{P}_{2} \) [M-H]- 652.3, found 652.4.

GalNAc-PP-Farnesyl (Precursor to 4.2). GalNAc-PP-Farnesyl was prepared as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-Farnesyl as a white solid (15 mg, 14% over two steps). \(^1\)H NMR (500 MHz, D\(_2\)O): \( \delta \) 5.59 (dd, \( J = 7.1 \) Hz, \( J = 3.4 \) Hz, 1H), 5.51 (t, \( J = 6.7 \) Hz, 1H), 5.30-5.22 (m, 2H), 4.56 (t, \( J = 6.8 \) Hz, 2H), 4.32 (dt, \( J = 10.9 \) Hz, \( J = 3.1 \) Hz, 1H), 4.26 (t, \( J = 6.2 \) Hz, 1H), 4.10 (d, \( J = 2.8 \) Hz, 1H), 4.02 (dd, \( J = 10.9 \) Hz, \( J = 3.1 \) Hz, 1H), 3.88-3.78 (m, 2H), 2.26-2.15 (m, 6H), 2.13 (s, 3H), 2.11-2.07 (m, 2H), 1.79 (s, 3H), 1.75 (s, 3H), 1.69 (s, 6H); \(^{13}\)C NMR (125 MHz, D\(_2\)O): \( \delta \) 175.0, 143.1, 136.7, 133.5, 124.5, 124.2, 119.5 (d, \( J = 8.4 \) Hz), 94.8 (d, \( J = 6.2 \) Hz), 72.1, 68.5, 67.8, 63.1 (d, \( J = 5.9 \) Hz), 61.1, 49.8 (d, \( J = 8.3 \) Hz), 38.8, 25.8, 25.6, 24.9, 22.2, 17.0, 15.7, 15.3; \(^{31}\)P NMR (162 MHz, D\(_2\)O): \( \delta \) -10.2 (d, \( J = 21.4 \) Hz), -12.5 (d, \( J = 21.7 \) Hz); LRMS (ESI) calcd for \( \text{C}_{23}\text{H}_{40}\text{NO}_{12}\text{P}_{2} \) [M-H]- 584.2, found 584.2.
GalNAc-PP-Geranyl (Precursor to 4.3). GalNAc-PP-Geranyl was prepared as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-Geranyl as a white solid (10 mg, 11% over two steps). $^1$H NMR (500 MHz, D$_2$O): $\delta$ 5.57 (dd, $J = 3.4$ Hz, $J = 7.0$ Hz, 1H), 5.48 (td, $J = 1.1$ Hz, $J = 7.1$ Hz, 1H), 5.28-5.22 (m, 1H), 4.53 (t, $J = 6.9$ Hz, 2H), 4.29 (dt, $J = 3.0$ Hz, $J = 10.9$ Hz, 1H), 4.23 (t, $J = 6.2$ Hz, 1H), 4.08 (d, $J = 2.6$ Hz, 1H), 4.01 (dd, $J = 3.2$ Hz, $J = 10.9$ Hz, 1H), 3.85-3.78 (m, 2H), 2.23-2.17 (m, 2H), 2.16-2.12 (m, 2H), 2.11 (s, 3H), 1.76 (s, 3H), 1.73 (s, 3H), 1.67 (s, 3H); $^{13}$C NMR (125 MHz, D$_2$O): $\delta$ 160.3, 143.0, 133.7, 124.1, 121.6, 94.7, 72.1, 68.4, 67.8, 62.5, 61.1, 49.8, 38.8, 25.6, 24.8, 22.2, 17.0, 15.6; $^{31}$P NMR (162 MHz, D$_2$O): $\delta$ -10.2, -12.5; LRMS (ESI) calcd for C$_{18}$H$_{32}$NO$_{12}$P$_2$ [M-H]$^-$ 516.1, found 516.1.

GalNAc-PP-cis-Farnesyl (Precursor to 4.4). GalNAc-PP-cis-Farnesyl was prepared as described above. Following the deprotection step, the crude product was purified via C-18 reverse phase column chromatography to afford GalNAc-PP-cis-Farnesyl as a white solid (11 mg, 6% over two steps). $^1$H NMR (500 MHz, D$_2$O): $\delta$ 5.58 (dd, $J = 3.4$ Hz, $J = 7.0$ Hz, 1H), 5.51 (t, $J = 6.1$ Hz, 1H),...
Hz, 1H), 5.32-5.22 (m, 2H), 4.52 (t, J = 7.0 Hz, 2H), 4.31 (dt, J = 3.0 Hz, J = 10.9 Hz, 1H), 4.25 (t, J = 6.2 Hz, 1H), 4.09 (d, J = 3.0 Hz, 1H), 4.01 (dd, J = 3.2 Hz, J = 10.9 Hz, 1H), 3.88-3.75 (m, 2H), 2.24-2.13 (m, 8H), 2.12 (s, 3H), 1.83 (s, 3H), 1.75 (s, 6H), 1.69 (s, 3H); 13C NMR (125 MHz, D$_2$O): δ 175.0, 142.8, 137.2, 133.7, 124.8, 124.4, 120.5, 94.7, 72.1, 68.5, 67.9, 62.9, 61.1, 49.8, 32.1, 31.6, 31.2, 25.9, 24.9, 22.7, 22.5, 22.2, 17.0; $^{31}$P NMR (162 MHz, D$_2$O): δ -10.2, -12.5; LRMS (ESI) calcd for C$_{23}$H$_{40}$NO$_{12}$P$_2$ [M-H]$^-$ 584.2, found 584.2.

1-(12-Benzyl oxy-2,6,10-trimethyl-dodeca-2,6,10-triene-4-sulfonyl)-4-methyl-benzene (4.5). A reaction vessel was charged with 2.93 (305 mg, 1.36 mmol) and flushed with argon. Freshly distilled THF (6 mL) and HMPA (2 mL) were then added, after which the vessel was cooled to -78 °C. n-BuLi (0.850 mL, 1.36 mmol) was added dropwise and the reaction was stirred for 10 min. The bromide equivalent of 2.44 (366 mg, 1.13 mmol) in THF (6 mL) was then added dropwise over 1 h via an automated injector. After warming overnight to room temperature, the reaction was quenched via the addition of 1% AcOH (aq.) (25 mL). The product was subsequently extracted with diethyl ether (50 mL), and the organic layer was dried over Na$_2$SO$_4$. The organic layer was concentrated, and the crude product was purified via flash column chromatography (EtOAc in hexanes, 5% to 7%) to afford 4.5 as a colorless oil (487 mg, 92%). $^1$H NMR (500 MHz, D$_2$O): δ 7.68 (d, J = 8.2 Hz, 2H), 7.33-7.22 (m, 7H), 5.40 (td, 186
\[ J = 6.8 \text{ Hz, } J = 1.0 \text{ Hz, } 1H \], 5.13 (t, \( J = 6.7 \text{ Hz, } 1H \)), 4.90 (dt, \( J = 10.4 \text{ Hz, } J = 1.1 \text{ Hz, } 1H \)), 4.47 (s, 2H), 3.97 (d, \( J = 6.6 \text{ Hz, } 2H \)), 3.79 (td, \( J = 10.8 \text{ Hz, } J = 3.0 \text{ Hz, } 1H \)), 2.68 (dd, \( J = 13.3 \text{ Hz, } J = 2.6 \text{ Hz, } 1H \)), 2.48 (dd, \( J = 13.2 \text{ Hz, } J = 11.4 \text{ Hz, } 1H \)), 2.41 (s, 3H), 2.06-1.97 (m, 4H), 1.71 (s, 3H), 1.62 (s, 3H), 1.56 (s, 3H), 1.16 (s, 3H); \(^{13}\text{C} \text{ NMR (125 MHz, } \text{D}_2\text{O}): } \delta 144.5, 141.8, 140.2, 138.8, 135.2, 130.8, 129.5, 129.4, 128.5, 128.2, 128.0, 127.7, 122.5, 117.5, 72.3, 66.6, 63.9, 32.4, 30.0, 26.7, 26.0, 23.6, 23.6, 21.8, 18.1.

\[
\text{HO} \quad \begin{array}{c} \text{cis-Farnesol (4.6)} \end{array} \quad \text{Ethylamine (20 mL) was condensed into a reaction vessel under nitrogen at -78 °C. Lithium metal was then added, after which the mixture was stirred for 1 h to afford a dark blue solution. 4.5 (462 mg, 0.990 mmol) in anhydrous THF (1 mL) was added dropwise over 1 min. The reaction was immediately quenched with } \text{NH}_4\text{Cl and then brought to room temperature. Following evaporation of ethylamine, the residue was dissolved in EtOAc (25 mL) and washed with water. The organic layer was subsequently dried over } \text{Na}_2\text{SO}_4, \text{ and concentrated. The crude product was subjected to flash column chromatography (EtOAc in Hexanes, 5\% to 7\%) to afford cis-Farnesol 4.6 as a colorless oil (119 mg from 462 mg, 54\%). } \]

\(^1\text{H} \text{ NMR (500 MHz, } \text{D}_2\text{O}): } \delta 5.42 (t d, \( J = 7.1 \text{ Hz, } J = 1.2 \text{ Hz, } 1H \)), 5.12-5.06 (m, 2H), 4.07 (d, \( J = 7.1 \text{ Hz, } 2H \)), 2.11-2.05 (m, 2H), 2.04-1.96 (m, 2H), 1.73 (s, 3H), 1.67 (s, 6H), 1.59 (s, 3H); \(^{13}\text{C} \text{ NMR (125 MHz, } \text{D}_2\text{O): } \delta 140.1, 136.4, 131.9, 124.7, 124.7, 124.4, 59.2, 32.5, 32.1, 26.8, 26.5, 25.9, 23.7, 23.5, 17.8.\]


APPENDIX

HPLC DATA, NCI SCREENING RESULTS, AND NMR SPECTRA
**HPLC Analysis of Target compounds (1.15-1.21).**

<table>
<thead>
<tr>
<th>Cmpd.</th>
<th>Column Condition (^{(a)})</th>
<th>Retention time</th>
<th>Purity result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.15</td>
<td>12 % solvent B constant 12 % to 90 % solvent B gradient</td>
<td>7.10 min</td>
<td>&gt; 99.0 %</td>
</tr>
<tr>
<td>1.16</td>
<td>12 % solvent B constant 12 % to 90 % solvent B gradient</td>
<td>12.22 min</td>
<td>98.0 % (^{(b)})</td>
</tr>
<tr>
<td>1.17</td>
<td>20 % solvent B constant 20 % to 90 % solvent B gradient</td>
<td>8.99 min</td>
<td>&gt; 99.0 %</td>
</tr>
<tr>
<td>1.18</td>
<td>12 % solvent B constant 12 % to 90 % solvent B gradient</td>
<td>12.71 min</td>
<td>&gt; 99.0 %</td>
</tr>
<tr>
<td>1.19</td>
<td>12 % solvent B constant 12 % to 90 % solvent B gradient</td>
<td>6.57 min</td>
<td>&gt; 99.0 %</td>
</tr>
<tr>
<td>1.20</td>
<td>10 % solvent B constant 10 % to 90 % solvent B gradient</td>
<td>6.83 min</td>
<td>&gt; 99.0 %</td>
</tr>
<tr>
<td>1.21</td>
<td>20 % solvent B constant 20 % to 90 % solvent B gradient</td>
<td>8.79 min</td>
<td>97.7 % (^{(c)})</td>
</tr>
</tbody>
</table>

\(^{(a)}\) solvent A, 0.4 % TFA in water; solvent B, 0.4 % TFA in acetonitrile; total flow rate: 0.5 mL/min; column: Kromasil®C18, 3.5um, 30 X 4.6 from Restek Corporation; detector: UV absorption at 256 nm; \(^{(b)}\) impurities (~1.6 % and ~0.4 %) at 10.04 min and 13.42 min respectively; \(^{(c)}\) impurities (~1.1 % and ~1.2 %) at 7.20 min and 11.74 min respectively;
UV Absorption @ 256 nm (uV)

Time (min)

Compound 6g

6g
<table>
<thead>
<tr>
<th>Tumor Type</th>
<th>SAHA 1.9</th>
<th>SAHA 1.12</th>
<th>SAHA 1.13</th>
<th>SAHA 1.14</th>
<th>SAHA 1.19</th>
<th>SAHA 1.21</th>
<th>NSI (c)</th>
<th>NSI (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung</td>
<td>0 %</td>
<td>45 %</td>
<td>53 %</td>
<td>68 %</td>
<td>1 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon</td>
<td>-10 %</td>
<td>51 %</td>
<td>78 %</td>
<td>83 %</td>
<td>-25 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>-3 %</td>
<td>40 %</td>
<td>78 %</td>
<td>71 %</td>
<td>8 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ovarian</td>
<td>-4 %</td>
<td>48 %</td>
<td>65 %</td>
<td>74 %</td>
<td>1 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leukemia</td>
<td>0 %</td>
<td>7 %</td>
<td>104 %</td>
<td>84 %</td>
<td>-6 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renal</td>
<td>-23 %</td>
<td>46 %</td>
<td>65 %</td>
<td>90 %</td>
<td>-20 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melanoma</td>
<td>-18 %</td>
<td>33 %</td>
<td>60 %</td>
<td>69 %</td>
<td>-43 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>6 %</td>
<td>36 %</td>
<td>66 %</td>
<td>84 %</td>
<td>-9 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNS</td>
<td>5 %</td>
<td>58 %</td>
<td>67 %</td>
<td>89 %</td>
<td>-18 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>-7 %</td>
<td>42 %</td>
<td>70 %</td>
<td>78 %</td>
<td>-12 %</td>
<td>101 %</td>
<td>103 %</td>
<td></td>
</tr>
</tbody>
</table>

(a) The data listed were the average percentage by which the 10 µM HDACi treatment inhibited the growth of cancer cell lines. Without the HDACi treatment, after a 48 hour incubation, the cell number corresponds to 100%. The cell number at the initial moment corresponds to 0%.

(b) A negative percentage indicates that the HDACi at 10 µM decreases the cell number in comparison to the number seen at the initial moment.

(c) NSI means no significant inhibition of cell growth (> 95 %).
The following graphs present the inhibition values for 10 μM HDACi treatments of the NIH cancer cell lines. The average inhibition values presented above was derived from these graphs.
# Developmental Therapeutics Program

## One Dose Mean Graph

**NSC:** 746455 / 1  
**Conc:** 1.00E-5 Molar  
**Test Date:** Oct 22, 2007  
**Experiment ID:** 07100538  
**Report Date:** Nov 15, 2007

<table>
<thead>
<tr>
<th>Panel/Cell Line</th>
<th>Growth Percent</th>
<th>Mean Growth Percent - Growth Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Small Cell Lung Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A549/ATCC</td>
<td>54.08</td>
<td></td>
</tr>
<tr>
<td>PVCX</td>
<td>54.61</td>
<td></td>
</tr>
<tr>
<td>HOP-62</td>
<td>30.05</td>
<td></td>
</tr>
<tr>
<td>HOP-82</td>
<td>44.77</td>
<td></td>
</tr>
<tr>
<td>NCI-H226</td>
<td>88.38</td>
<td></td>
</tr>
<tr>
<td>NCI-H23</td>
<td>69.16</td>
<td></td>
</tr>
<tr>
<td>NCI-H322M</td>
<td>41.32</td>
<td></td>
</tr>
<tr>
<td>NCI-H460</td>
<td>62.51</td>
<td></td>
</tr>
<tr>
<td>NCI-H622</td>
<td>31.11</td>
<td></td>
</tr>
<tr>
<td>Colon Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLO 205</td>
<td>104.61</td>
<td></td>
</tr>
<tr>
<td>HCC-2998</td>
<td>71.12</td>
<td></td>
</tr>
<tr>
<td>HCT-116</td>
<td>30.07</td>
<td></td>
</tr>
<tr>
<td>HCT-115</td>
<td>116.77</td>
<td></td>
</tr>
<tr>
<td>HT29</td>
<td>60.05</td>
<td></td>
</tr>
<tr>
<td>KM12</td>
<td>77.97</td>
<td></td>
</tr>
<tr>
<td>SW-620</td>
<td>61.85</td>
<td></td>
</tr>
<tr>
<td>Breast Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BT-20</td>
<td>72.13</td>
<td></td>
</tr>
<tr>
<td>HS 578T</td>
<td>211.32</td>
<td></td>
</tr>
<tr>
<td>MCF7</td>
<td>77.80</td>
<td></td>
</tr>
<tr>
<td>MDA-MB-231/ATCC</td>
<td>56.67</td>
<td></td>
</tr>
<tr>
<td>MDA-MB-435</td>
<td>23.30</td>
<td></td>
</tr>
<tr>
<td>MDA-MB-468</td>
<td>64.58</td>
<td></td>
</tr>
<tr>
<td>NCI-AO2-RES</td>
<td>14.85</td>
<td></td>
</tr>
<tr>
<td>T-47D</td>
<td>64.70</td>
<td></td>
</tr>
<tr>
<td>Ovarian Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOROV1</td>
<td>64.19</td>
<td></td>
</tr>
<tr>
<td>OVCAR-3</td>
<td>57.74</td>
<td></td>
</tr>
<tr>
<td>OVCAR-4</td>
<td>71.78</td>
<td></td>
</tr>
<tr>
<td>OVCAR-5</td>
<td>64.96</td>
<td></td>
</tr>
<tr>
<td>OVCAR-8</td>
<td>45.14</td>
<td></td>
</tr>
<tr>
<td>SK-OV-3</td>
<td>81.89</td>
<td></td>
</tr>
<tr>
<td>Leukemia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCRF-CEM</td>
<td>48.59</td>
<td></td>
</tr>
<tr>
<td>HL-60(TB)</td>
<td>258.41</td>
<td></td>
</tr>
<tr>
<td>K-562</td>
<td>72.30</td>
<td></td>
</tr>
<tr>
<td>MOLT-4</td>
<td>63.02</td>
<td></td>
</tr>
<tr>
<td>RPMI-8226</td>
<td>55.64</td>
<td></td>
</tr>
<tr>
<td>Renal Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>786-O</td>
<td>39.93</td>
<td></td>
</tr>
<tr>
<td>A498</td>
<td>101.83</td>
<td></td>
</tr>
<tr>
<td>ACHN</td>
<td>82.47</td>
<td></td>
</tr>
<tr>
<td>CAKI-1</td>
<td>63.93</td>
<td></td>
</tr>
<tr>
<td>SN12C</td>
<td>86.20</td>
<td></td>
</tr>
<tr>
<td>TK-10</td>
<td>46.74</td>
<td></td>
</tr>
<tr>
<td>UO-11</td>
<td>45.30</td>
<td></td>
</tr>
<tr>
<td>Melanoma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOX IMVI</td>
<td>68.28</td>
<td></td>
</tr>
<tr>
<td>M14</td>
<td>48.52</td>
<td></td>
</tr>
<tr>
<td>NALME-3M</td>
<td>6.42</td>
<td></td>
</tr>
<tr>
<td>SK-NEL-2</td>
<td>61.86</td>
<td></td>
</tr>
<tr>
<td>SK-NEL-28</td>
<td>82.17</td>
<td></td>
</tr>
<tr>
<td>SK-NEL-3</td>
<td>92.91</td>
<td></td>
</tr>
<tr>
<td>UACC-257</td>
<td>43.03</td>
<td></td>
</tr>
<tr>
<td>UACC-62</td>
<td>51.22</td>
<td></td>
</tr>
<tr>
<td>Prostate Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DU-145</td>
<td>75.19</td>
<td></td>
</tr>
<tr>
<td>PC-3</td>
<td>57.76</td>
<td></td>
</tr>
<tr>
<td>CNS Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-208</td>
<td>71.67</td>
<td></td>
</tr>
<tr>
<td>SF-295</td>
<td>15.26</td>
<td></td>
</tr>
<tr>
<td>SF-63</td>
<td>86.36</td>
<td></td>
</tr>
<tr>
<td>SNB-19</td>
<td>101.20</td>
<td></td>
</tr>
<tr>
<td>SNB-75</td>
<td>53.92</td>
<td></td>
</tr>
<tr>
<td>U251</td>
<td>76.00</td>
<td></td>
</tr>
</tbody>
</table>

| Mean | 66.57 |
| Delta | 63.16 |
| Range | 20.46 |
### Developmental Therapeutics Program

#### One Dose Mean Graph

<table>
<thead>
<tr>
<th>Panel/Cell Line</th>
<th>Growth Percent</th>
<th>Mean Growth Percent - Growth Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Small Cell Lung Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A549/ATCC</td>
<td>-8.34</td>
<td></td>
</tr>
<tr>
<td>EKVX</td>
<td>25.24</td>
<td></td>
</tr>
<tr>
<td>HOPR-62</td>
<td>-31.70</td>
<td></td>
</tr>
<tr>
<td>HOPR-92</td>
<td>-61.96</td>
<td></td>
</tr>
<tr>
<td>NC1-H226</td>
<td>23.89</td>
<td></td>
</tr>
<tr>
<td>NC1-H226M</td>
<td>21.42</td>
<td></td>
</tr>
<tr>
<td>NC1-H322M</td>
<td>9.59</td>
<td></td>
</tr>
<tr>
<td>NC1-H460</td>
<td>7.36</td>
<td></td>
</tr>
<tr>
<td>NC1-H622</td>
<td>8.32</td>
<td></td>
</tr>
<tr>
<td>Colon Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLO 205</td>
<td>-75.70</td>
<td></td>
</tr>
<tr>
<td>HCC-2998</td>
<td>-1.33</td>
<td></td>
</tr>
<tr>
<td>HCT-116</td>
<td>-69.05</td>
<td></td>
</tr>
<tr>
<td>HCT-15</td>
<td>6.70</td>
<td></td>
</tr>
<tr>
<td>HT29</td>
<td>12.02</td>
<td></td>
</tr>
<tr>
<td>KM12</td>
<td>8.61</td>
<td></td>
</tr>
<tr>
<td>SW-620</td>
<td>-61.50</td>
<td></td>
</tr>
<tr>
<td>Breast Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BT-549</td>
<td>43.31</td>
<td></td>
</tr>
<tr>
<td>HS 578T</td>
<td>25.98</td>
<td></td>
</tr>
<tr>
<td>MCF7</td>
<td>13.57</td>
<td></td>
</tr>
<tr>
<td>MDA-MB-231/ATCC</td>
<td>20.86</td>
<td></td>
</tr>
<tr>
<td>MDA-MB-435</td>
<td>-21.31</td>
<td></td>
</tr>
<tr>
<td>MDA-MB-468</td>
<td>45.76</td>
<td></td>
</tr>
<tr>
<td>NC1A-D9R-RES</td>
<td>-14.15</td>
<td></td>
</tr>
<tr>
<td>T-47D</td>
<td>-11.87</td>
<td></td>
</tr>
<tr>
<td>Ovarian Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOROV1</td>
<td>-13.75</td>
<td></td>
</tr>
<tr>
<td>OVCAR-3</td>
<td>-16.17</td>
<td></td>
</tr>
<tr>
<td>OVCAR-4</td>
<td>55.74</td>
<td></td>
</tr>
<tr>
<td>OVCAR-5</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td>OVCAR-8</td>
<td>1.24</td>
<td></td>
</tr>
<tr>
<td>SK-OV-3</td>
<td>-26.04</td>
<td></td>
</tr>
<tr>
<td>Leukemia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CORF-CEM</td>
<td>9.16</td>
<td></td>
</tr>
<tr>
<td>HL-60(TB)</td>
<td>-5.98</td>
<td></td>
</tr>
<tr>
<td>K562</td>
<td>-16.07</td>
<td></td>
</tr>
<tr>
<td>MOLT-4</td>
<td>-7.12</td>
<td></td>
</tr>
<tr>
<td>RPMI-8226</td>
<td>-13.94</td>
<td></td>
</tr>
<tr>
<td>Renal Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>786-0</td>
<td>-36.11</td>
<td></td>
</tr>
<tr>
<td>A498</td>
<td>-74.42</td>
<td></td>
</tr>
<tr>
<td>ACHN</td>
<td>-81.23</td>
<td></td>
</tr>
<tr>
<td>CAKI-1</td>
<td>7.42</td>
<td></td>
</tr>
<tr>
<td>SN12C</td>
<td>24.88</td>
<td></td>
</tr>
<tr>
<td>TK-10</td>
<td>8.37</td>
<td></td>
</tr>
<tr>
<td>UO-11</td>
<td>6.73</td>
<td></td>
</tr>
<tr>
<td>Melanoma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOX IMVI</td>
<td>-72.56</td>
<td></td>
</tr>
<tr>
<td>M14</td>
<td>1.63</td>
<td></td>
</tr>
<tr>
<td>MALME-3M</td>
<td>-28.71</td>
<td></td>
</tr>
<tr>
<td>SK-MEL-2</td>
<td>-51.08</td>
<td></td>
</tr>
<tr>
<td>SK-MEL-28</td>
<td>-14.96</td>
<td></td>
</tr>
<tr>
<td>SK-MEL-3</td>
<td>-78.99</td>
<td></td>
</tr>
<tr>
<td>UACC-257</td>
<td>-32.09</td>
<td></td>
</tr>
<tr>
<td>UACC-62</td>
<td>-71.24</td>
<td></td>
</tr>
<tr>
<td>Prostate Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DU-145</td>
<td>-26.21</td>
<td></td>
</tr>
<tr>
<td>PC-3</td>
<td>11.04</td>
<td></td>
</tr>
<tr>
<td>CNS Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-208</td>
<td>-0.08</td>
<td></td>
</tr>
<tr>
<td>SF-295</td>
<td>-55.50</td>
<td></td>
</tr>
<tr>
<td>SF-538</td>
<td>5.16</td>
<td></td>
</tr>
<tr>
<td>SNB-19</td>
<td>-12.39</td>
<td></td>
</tr>
<tr>
<td>SNB-75</td>
<td>46.71</td>
<td></td>
</tr>
<tr>
<td>U251</td>
<td>-89.91</td>
<td></td>
</tr>
</tbody>
</table>

**Mean**: -12.02  
**Delta**: 77.89  
**Range**: 146.05
The table below lists the growth percent for various cell lines across different cancer types. The mean growth percent is calculated for each cell line and compared to the growth percent.

### Growth Percent

<table>
<thead>
<tr>
<th>Panel/Cell Line</th>
<th>Growth Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Small Cell Lung Cancer</td>
<td></td>
</tr>
<tr>
<td>A549/ATCC</td>
<td>102.74</td>
</tr>
<tr>
<td>CCAK</td>
<td>108.39</td>
</tr>
<tr>
<td>HOP-62</td>
<td>118.03</td>
</tr>
<tr>
<td>HOP-92</td>
<td>69.38</td>
</tr>
<tr>
<td>NC1-H226</td>
<td>87.62</td>
</tr>
<tr>
<td>NC1-H23</td>
<td>109.64</td>
</tr>
<tr>
<td>NC1-H322M</td>
<td>95.93</td>
</tr>
<tr>
<td>NC1-H480</td>
<td>120.24</td>
</tr>
<tr>
<td>NC1-H522</td>
<td>99.22</td>
</tr>
<tr>
<td>Colon Cancer</td>
<td></td>
</tr>
<tr>
<td>COLO 205</td>
<td>122.84</td>
</tr>
<tr>
<td>HCC-2398</td>
<td>106.84</td>
</tr>
<tr>
<td>HCT-116</td>
<td>97.29</td>
</tr>
<tr>
<td>HCT-156</td>
<td>107.60</td>
</tr>
<tr>
<td>HT29</td>
<td>115.39</td>
</tr>
<tr>
<td>KM12</td>
<td>108.48</td>
</tr>
<tr>
<td>SW-483</td>
<td>94.82</td>
</tr>
<tr>
<td>Breast Cancer</td>
<td></td>
</tr>
<tr>
<td>BT-549</td>
<td>108.25</td>
</tr>
<tr>
<td>HS 578T</td>
<td>90.01</td>
</tr>
<tr>
<td>MCF7</td>
<td>112.57</td>
</tr>
<tr>
<td>MDA-MB-231/ATCC</td>
<td>100.67</td>
</tr>
<tr>
<td>MDA-MB-435</td>
<td>105.56</td>
</tr>
<tr>
<td>MDA-MB-468</td>
<td>110.33</td>
</tr>
<tr>
<td>NC1/ADR-RES</td>
<td>91.27</td>
</tr>
<tr>
<td>T-47D</td>
<td>102.42</td>
</tr>
<tr>
<td>Ovarian Cancer</td>
<td></td>
</tr>
<tr>
<td>IPOV1</td>
<td>100.50</td>
</tr>
<tr>
<td>OVCAR-3</td>
<td>145.83</td>
</tr>
<tr>
<td>OVCAR-4</td>
<td>111.53</td>
</tr>
<tr>
<td>OVCAR-5</td>
<td>84.12</td>
</tr>
<tr>
<td>OVCAR-8</td>
<td>96.16</td>
</tr>
<tr>
<td>SK-OV-3</td>
<td>109.33</td>
</tr>
<tr>
<td>Leukemia</td>
<td></td>
</tr>
<tr>
<td>CCRF-CEM</td>
<td>92.14</td>
</tr>
<tr>
<td>HL-60(TB)</td>
<td>136.00</td>
</tr>
<tr>
<td>K-562</td>
<td>85.09</td>
</tr>
<tr>
<td>MOLT-4</td>
<td>110.29</td>
</tr>
<tr>
<td>RPMI-8226</td>
<td>79.85</td>
</tr>
<tr>
<td>Renal Cancer</td>
<td></td>
</tr>
<tr>
<td>786-0</td>
<td>104.85</td>
</tr>
<tr>
<td>A498</td>
<td>67.63</td>
</tr>
<tr>
<td>ACHN</td>
<td>99.25</td>
</tr>
<tr>
<td>CAK-1</td>
<td>105.00</td>
</tr>
<tr>
<td>SN12C</td>
<td>97.73</td>
</tr>
<tr>
<td>TK10</td>
<td>118.02</td>
</tr>
<tr>
<td>UO-31</td>
<td>84.68</td>
</tr>
<tr>
<td>Melanoma</td>
<td></td>
</tr>
<tr>
<td>LOX IMVI</td>
<td>92.42</td>
</tr>
<tr>
<td>M14</td>
<td>94.08</td>
</tr>
<tr>
<td>MAL-ME-3M</td>
<td>106.78</td>
</tr>
<tr>
<td>SK-MEL-2</td>
<td>-26.79</td>
</tr>
<tr>
<td>SK-MEL-28</td>
<td>98.05</td>
</tr>
<tr>
<td>SK-MEL-3</td>
<td>111.24</td>
</tr>
<tr>
<td>UACC-257</td>
<td>112.16</td>
</tr>
<tr>
<td>UACC-62</td>
<td>93.26</td>
</tr>
<tr>
<td>Prostate Cancer</td>
<td></td>
</tr>
<tr>
<td>DU-145</td>
<td>134.64</td>
</tr>
<tr>
<td>PC-3</td>
<td>82.12</td>
</tr>
<tr>
<td>CNS Cancer</td>
<td></td>
</tr>
<tr>
<td>SF-268</td>
<td>103.96</td>
</tr>
<tr>
<td>SF-266</td>
<td>126.38</td>
</tr>
<tr>
<td>SF-539</td>
<td>118.27</td>
</tr>
<tr>
<td>SNB-19</td>
<td>95.67</td>
</tr>
<tr>
<td>SNB-75</td>
<td>116.94</td>
</tr>
<tr>
<td>U251</td>
<td>96.87</td>
</tr>
</tbody>
</table>

**Mean:** 101.21  **Delta:** 128.00  **Range:** 172.62
SpinWorks 2.3: C-17c in DMF-d7 at 500 MHz

File: C:\Users\BW\Desktop\NMR Data\Click Chemistry HDACi\C-17c\1\fid
Expt: zg30
Transmitter freq.: 500.023088 MHz
time domain size: 65536 points
processed size: 32768 complex points

Freq. of 0 ppm: 500.020004 MHz
number of scans: 16
LB: 0.300 G B: 0.0000
SpinWorks 2.3:  C-17c in DMF-d7 at 500 M PPM

file: C:\Users\BW\Desktop\N M R  DAT A\Click Chemistry H DACi\C-17c\2\fid

transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 908
freq. of 0 ppm: 125.730012 MHz
processed size: 32768 complex points
LB:    1.000    GB: 0.0000
SpinWorks 2.3: C-17d in DMF-d7 at 500 MHz

File: C:\Users\BW\Desktop\NMR DATA\Click Chemistry\HDACi\C-17d (2)\fid
Expt: <zg30>
Transmitter freq.: 500.023088 MHz
time domain size: 65536 points
width: 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
number of scans: 16

Freq. of 0 ppm: 500.020008 MHz
processed size: 32768 complex points
LB: 0.000 GB: 0.0000

1.18

- N- NHOH
  - O
  - O

Peaks of 0 ppm: 500.02000 MHz
refined to allow: 3058 complex points
LB: 0.000 GB: 0.0000
number of fits: 10
SpinWorks 2.3: C-17d in DMF-d7 at 500 MHz

File: C:\Users\BW\Desktop\NM R DATA\Click Chemistry H DAC\C-17d (2)\2\fid
Expt: <zgpg30>
Transmitter freq.: 125.742702 MHz
Time domain size: 65536 points
Width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
Number of scans: 800
Freq. of 0 ppm: 125.730001 MHz
Processed size: 32768 complex points
LB: 1.000 G B: 0.000

Species: 2.3 - C-NH-Me-C-NH-Me
SpinWorks 2.3: C-16e in DMF-d7 at 500 M

1.19

In: C:\Users\BW\Desktop\SpinData\NMRT\SpinWorks\Click Chemistry\HCCOC-C-FmAtNl_exp\exp30
Transmitted freq. 500.020090 MHz
Time domain 65536 points
Width 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
Number of scans 16
Freq. of 0 ppm 500.02009 MHz
Plot: 128000 complex points
LB 0.0000 GB 0.0000

Hz of 40 ppm 500.0000 MHz
Plot: 128000 complex points
LB 0.0000 GB 0.0000

Number of scans 16
SpinWorks 2.3: C-17e in DMF-d7 at 500 M

[Chemical structure image]

1.19
SpinWorks 2.3: C-17f in DMF-d7 at 500 M

10.7700
10.3176
8.6527
8.0308
7.7130
7.6977
7.3692
7.3542
7.3375
7.1266
7.1119
7.0971
5.5664

file: C:\Users\BW\Desktop\Spec Data\NMR DATA\Click Chemistry HDACi\C-17f\1\fid
expt: zg30
transmitter freq.: 500.023088 MHz
time domain size: 65536 points
width: 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
number of scans: 16
SpinWorks 2.3: C-17g in DMF-d7 at 500 MHz

file: C:\Users\BW\Desktop\NM R DATA\Click Chemistry HDACi\C-17g\2\fid  
transmitter freq.: 125.742702 MHz

time domain size: 65536 points

number of scans: 6940

freq. of 0 ppm: 125.730000 MHz

processed size: 32768 complex points

LB: 1.000    GB: 0.0000
SpinWorks 2.3: C-16a in DMF-d7 at 500M

File: C:\Users\BW\Desktop\NMR DATA\Click Chemistry\HDAOC\C\FullHPLC\expt<zg30>
Transmitter freq.: 500.023088 MHz
time domain size: 65536 points
width: 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
number of scans: 16

Freq. of 0 ppm: 500.020010 MHz
processed size: 32768 complex points
LB: 0.000 G B: 0.0000

1.54
SpinWorks 2.3: o-16c in DMF-d7 at 500 MHz

N=N

NHOPMB

1.56

PPM

0.803

1.230

1.414

1.56

1.81

2.026

2.041

2.025

2.021

1.981

3.085

2.026

2.021

1.9689

5.069

2.021

2.041

1.981

7.4374

7.3368

7.3072

7.2744

7.2380

7.2237

6.9679

4.5303

4.5443

4.5584

4.9662

2.6512

2.6673

2.6819

2.2838

2.2688

2.2539

2.2838

2.6819

2.6673

2.6512

2.2838

2.2688

2.2539

file: C:\Users\BW\Desktop\NMR DATA\Click Chemistry HDACi\C-16c\1\fid
expt: <zg30>
transmitter freq.: 500.023088 MHz
time domain size: 65536 points
width: 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
number of scans: 16
d.2744

7.2380

7.2237

6.9679

4.5303

4.5443

4.5584

4.9662

2.6512

2.6673

2.6819

2.2838

2.2688

2.2539

2.2838

2.6819

2.6673

2.6512

2.2838

2.2688

2.2539

number of points: 65536 complex points

Freq. of ppm: 100.023051 MHz
processing: 1024 complex points
LB: 0.500 - 1.500
GB: 0.500 - 0.500
number of scans: 16

239
SpinWorks 2.3: C-16c in DMF-d7 at 500 MHz

file: C:\Users\BW\Desktop\NMR DATA\Click Chemistry HDACi\C-16c\2\fid
expt: zgpg30
transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 1301

freq. of 0 ppm: 125.730019 MHz
processed size: 32768 complex points
LB: 1.000 G B: 0.0000

159.2459
142.6034
142.1246
131.8446
129.5429
129.3279
127.6091
127.1624
114.7529
78.4923
56.0079
50.6825
33.2872
32.7509

1.56

C-16c

N
N

NHOPMB

PPM

161.0013
159.2459
142.6034
142.1246
131.8446
129.5429
129.3279
127.6091
127.1624
114.7529
78.4923
56.0079
50.6825
33.2872
32.7509

1.56
SpinWorks 2.3: C-16d in DMF-d7 at 500M PPM

file: C:\Users\BW\Desktop\NM R DATA\Click Chemistry HDACi\C-16d\2\fid
transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 4783

Freq. of 0 ppm: 125.730019 MHz
processed size: 32768 complex points
LB: 1.000 G B: 0.0000
SpinWorks 2.3: C-16e in DMF-d7 at 500 MHz

file: C:\Users\BW\Desktop\NMR DATA\Click Chemistry HDACi\C-16e\1\fid   expt: <zg30>
transmitter freq.: 500.023088 MHz
time domain size: 65536 points
width: 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
number of scans: 16
freq. of 0 ppm: 500.020090 MHz
processed size: 32768 complex points
LB: 0.300  G B: 0.0000

1.58

No: C:\Users\BW\Desktop\NMR DATA\Click Chemistry HDACi\C-16e\1\fid   expt: <zg30>
transmitter freq.: 500.023088 MHz
time domain size: 65536 points
width: 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
number of scans: 16
freq. of 0 ppm: 500.020090 MHz
processed size: 32768 complex points
LB: 0.300  G B: 0.0000
SpinWorks 2.3: C-16g in DMF-d7 at 500 MHz

File: C:\Users\BW\Desktop\NM R DATA\Click Chemistry HDACi\C-16g\2\fid
Experiment: zgpgp30
Transmitter freq.: 125.7427 MHz
time domain size: 65536 points
width: 300.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 10240

Freq. of 0 ppm: 125.729999 MHz
processed size: 32768 complex points
LB: 1.000    GB: 0.0000

1.60

NHOPMB

PPM 170.0 160.0 150.0 140.0 130.0 120.0 110.0 100.0 90.0 80.0 70.0 60.0 50.0 40.0 30.0 160.9957 159.1376 142.6995 137.2509 135.6250 131.8500 129.8000 129.3348 129.3087 127.8307 127.4914 124.3720 114.7421 78.4726 55.9968 53.0068
SpinWorks 2.3:

file: J:\AcGalNAc Protected Phosphate\2nd expt: zgpg30
transmitter freq.: 161.967474 MHz
time domain size: 65536 points
width: 64935.06 Hz = 400.914228 ppm = 0.990830 Hz/pt
number of scans: 57

freq. of 0 ppm: 161.975770 MHz
processed size: 32768 complex points
LB:    1.000    GB: 0.0000
SpinWorks 2.3:

File: J:\GalN AcPPH eptaPhos\1\fid
Experiment: zgpg30
Transmitter freq.: 161.967474 MHz
time domain size: 65536 points
width: 64935.06 Hz = 0.990830 ppm
number of scans: 6118

freq. of 0 ppm: 161.975498 MHz
processed size: 32768 complex points
LB: 1.000 GB: 0.0000

Precursor to 2.20

[Chemical structure image]
SpinWorks 2.3:

Precursor to 2.21
SpinWorks 2.3:

Precursor to 2.22
SpinWorks 2.3:

file: J:\GalN AcPPFarnPhos\1\fid
expt: zgpg30
transmitter freq.: 161.967474 MHz
time domain size: 65536 points
width: 64935.06 Hz = 400.914228 ppm = 0.990830 Hz/pt
number of scans: 246

freq. of 0 ppm: 161.975498 MHz
processed size: 32768 complex points
LB: 1.000 GB: 0.0000

HO
OH
O
AcNH

O
O

NH4 NH4

Precursor to 2.22
SpinWorks 2.3:

AcO
O
AcO
O
AcNH

Precursor to 2.23

PPM

file: J:\AcGalNAc\PPSolanesol\2\expt <zgpg30>
transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 12970

freq. of 0 ppm: 125.729938 MHz
processed size: 32768 complex points
LB: 0.000 GB: 0.0000
SpinWorks 2.3:

file: J:\AcGalNAcPPSolan400\3\fid   expt: <zgpg30>
transmitter freq.: 161.967474 MHz
time domain size: 65536 points
width: 64935.06 Hz = 400.914228 ppm = 0.990830 Hz/pt
number of scans: 211

freq. of 0 ppm: 161.975460 MHz
processed size: 32768 complex points
LB: 1.000    GB: 0.0000

Precursor to 2.23
Precursor to 2.24
Precursor to 2.24
SpinWorks 2.3:

NMR spectrum of BnO showing a peak at 2.35 ppm.

File: C:\Users\BW\Desktop\Spec Data\NMR Data\Project\BnO Experiment\BnO Expt 1999\transmitter freq.: 125.7427 MHz

Processed size: 32768 complex points

number of scans: 185
SpinWorks 2.3: 041201 in MeOH-d4 at rt on 500 M

BnO

transmitter freq.: 500.023088 MHz

number of scans: 16

freq. of 0 ppm: 500.020022 MHz

processed size: 32768 complex points
SpinWorks 2.3: 041201 in MeOH-d$_4$ at rt on 500 M

**BnO-**

\[ \text{OHP} \]

2.39
SpinWorks 2.3: add after purification the second spot in CDCl3 at rt on 400 MHz

file: C:\Users\BW\Desktop\Ram u NMR\prs-3-1\1\fid
expt: zg30
transmitter freq.: 400.132471 MHz
time domain size: 32768 points
width: 8278.15 Hz = 20.688513 ppm = 0.252629 Hz/pt
number of scans: 16

freq. of 0 ppm: 400.130018 MHz
processed size: 16384 complex points
LB: 0.300 GB: 0.0000
<table>
<thead>
<tr>
<th>PPM</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>150.0</td>
<td>140.070</td>
</tr>
<tr>
<td>140.0</td>
<td>138.6970</td>
</tr>
<tr>
<td>130.0</td>
<td>132.0848</td>
</tr>
<tr>
<td>120.0</td>
<td>130.5510</td>
</tr>
<tr>
<td>110.0</td>
<td>128.5591</td>
</tr>
<tr>
<td>100.0</td>
<td>128.0158</td>
</tr>
<tr>
<td>90.0</td>
<td>127.7449</td>
</tr>
<tr>
<td>80.0</td>
<td>122.6555</td>
</tr>
<tr>
<td>70.0</td>
<td>72.4278</td>
</tr>
<tr>
<td>60.0</td>
<td>66.5013</td>
</tr>
<tr>
<td>50.0</td>
<td>43.6969</td>
</tr>
<tr>
<td>40.0</td>
<td>32.1821</td>
</tr>
<tr>
<td>30.0</td>
<td>26.6900</td>
</tr>
<tr>
<td>20.0</td>
<td>23.6135</td>
</tr>
<tr>
<td>10.0</td>
<td>21.7753</td>
</tr>
</tbody>
</table>

**SpinWorks 2.3:**

**file:** C:\Users\BW\Desktop\Ramu NMR\prs-3-1\2\fid  
**expt:** zgpg30  
**transmitter freq.:** 100.622830 MHz  
**time domain size:** 65536 points  
**width:** 23980.82 Hz = 238.323802 ppm = 0.365918 Hz/pt  
**number of scans:** 137  
**freq. of 0 ppm:** 100.612749 MHz  
**processed size:** 32768 complex points  
**LB:** 1.000  
**GB:** 0.0000
SpinWorks 2.3: after purification the second spot in CDCl3 at rt on 400 MHz.

**NMR spectrum details:**
- Frequency of 0 ppm: 400.130018 MHz
- Processed size: 16384 complex points
- Number of scans: 16

**Chemical Structure:**
```
O
```

**Signal Peaks:**
- Between 8.0 and 7.0 ppm
- Between 4.3 and 3.0 ppm
- Between 2.0 and 1.5 ppm
SpinWorks 2.3:

File: C:\Users\BW\Desktop\Ramu NMR\prs-3-6\2\fid
Experiment: <zgpg30>
Transmitter freq.: 100.622830 MHz
time domain size: 65536 points
Width: 23980.82 Hz = 0.365918 Hz/pt
Number of scans: 204

Frequency of 0 ppm: 100.612747 MHz
Processed size: 32768 complex points
LB: 0.000 GB: 0.0000

2.42

---

The image contains a chemical structure labeled "2.42". The spectrum is labeled with various frequency values, and the experimental conditions are noted at the bottom of the page.
SpinWorks 2.3: add after purification the second spot in CDCl3 at rt on 400 MHz.

**NMR Spectrogram:**
- **Transmitter freq.:** 400.132471 MHz
- **Time domain size:** 32768 points
- **Width:** 8278.15 Hz = 20.688513 ppm = 0.252629 Hz/pt
- **Number of scans:** 16

**Chemical Structure:**

![Chemical Structure Image]

**Selected Peaks (ppm):**
- 7.7208, 7.7001, 7.2980, 7.2776
- 5.1722, 5.1527, 5.1328
- 3.7661, 3.7463
- 2.4102, 1.9824, 1.6467, 1.5687, 1.5548, 1.3106
SpinWorks 2.3: azid after purification the second spot in CDCl3 at rt on 400 MHz.

BnO

2.44

O=S=O

Cl
SpinWorks 2.3: 041201 in MeOH-d4 at rt on 500 M

transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 145

freq. of 0 ppm: 125.730106 MHz
processed size: 32768 complex points
LB: 1.000    GB: 0.0000

BnO\rightleftharpoons\text{OH}

2.86
SpinWorks 2.3: 17 in CD3OD at 400 MHz

<table>
<thead>
<tr>
<th>PPM</th>
<th>139.4810</th>
<th>132.7881</th>
<th>129.1031</th>
<th>125.0129</th>
<th>97.5026</th>
<th>65.3857</th>
<th>62.1818</th>
<th>59.1338</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>139.4810</td>
<td>132.7881</td>
<td>129.1031</td>
<td>125.0129</td>
<td>97.5026</td>
<td>65.3857</td>
<td>62.1818</td>
<td>59.1338</td>
</tr>
</tbody>
</table>

file: C:\Users\BW\Desktop\Ramu NMR\prs-3-3\2\fid  
transmitter freq.: 100.622830 MHz  
time domain size: 65536 points 
width: 23980.82 Hz = 238.323802 ppm = 0.365918 Hz/pt  
number of scans: 108  
freq. of 0 ppm: 100.612747 MHz  
processed size: 32768 complex points 
LB: 0.000  
B: 0.0000

HO-\(\text{C}=-\text{C}\)-\(\text{C}=-\text{C}\)-O\(\text{THP}\)

2.87
SpinWorks 2.3: 041201 in MeOH-d4 at rt on 500 M

BnO

2.89
SpinWorks 2.3:

file: C:\Users\BW\Desktop\Spec Data\NMR2\Ram\NMR2pm-3-Me26.txt

transmitter freq.: 125.742702 MHz

time domain size: 65536 points

width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt

number of scans: 4442

freq. of 0 ppm: 125.730104 MHz

processed size: 32768 complex points

LB: 1.000    GB: 0.0000
SpinWorks 2.3: add after purification the second spot in CDCl₃ at rt on 400 MHz

file: C:\Users\BW\Desktop\NDM\pr-3-8\1\fid  
transmitter freq.: 400.132471 MHz

time domain size 51200 points

freq. of 0 ppm: 400.130018 MHz
processed size: 16384 complex points

number of scans: 16

2.92
SpinWorks 2.3:

file: C:\Users\BW\Desktop\Ramu NMR\prs-3-8\2\fid  
expt: zgpg30
transmitter freq.: 100.622830 MHz

time domain size: 65536 points
width: 23980.82 Hz = 238.323802 ppm = 0.365918 Hz/pt

number of scans: 160

freq. of 0 ppm: 100.612748 MHz
processed size: 32768 complex points

LB: 1.000  GB: 0.0000

2.92
SpinWorks 2.3:

138.9661
131.3352
128.5457
127.8152
127.6627
125.0665
73.1259
68.9714
37.4336
36.9572
29.8159
25.9226
25.7029
19.7852
17.8438

file: C:\Users\BW\Desktop\Spec Data\NM R DATA\Wzy Project\Dolichol Derivative\BnCitronellol\2\fid   expt: <zgpg30>
transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 546

freq. of 0 ppm: 125.730102 MHz
processed size: 32768 complex points
LB: 1.000 G B: 0.0000
<table>
<thead>
<tr>
<th>PPM</th>
<th>7.3144</th>
<th>7.2600</th>
<th>5.2520</th>
<th>5.2373</th>
<th>5.2227</th>
<th>4.4806</th>
<th>4.1556</th>
<th>4.1320</th>
<th>4.0370</th>
<th>4.0134</th>
<th>3.4717</th>
<th>2.1024</th>
<th>1.9907</th>
<th>1.7661</th>
<th>1.6732</th>
<th>1.5843</th>
<th>1.3073</th>
<th>1.1864</th>
<th>0.8588</th>
<th>0.8456</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpinWorks 2.3</td>
<td>PPM</td>
<td>7.2</td>
<td>6.8</td>
<td>6.4</td>
<td>6.0</td>
<td>5.6</td>
<td>5.2</td>
<td>4.8</td>
<td>4.4</td>
<td>4.0</td>
<td>3.6</td>
<td>3.2</td>
<td>2.8</td>
<td>2.4</td>
<td>2.0</td>
<td>1.6</td>
<td>1.2</td>
<td>0.8</td>
<td>0.4</td>
<td>1.015</td>
</tr>
</tbody>
</table>
SpinWorks 2.3:

![NMR Spectrum](image)

BnO—\(\text{CH=CH}\)Br

2.103
SpinWorks 2.3:


file: C:\Users\BW\Desktop\Spec Data\NM R DATA\Wzy Project\Dolichol Derivative\Dolichol Coupling Bromide\2\fid
expt: <zgpg30>
transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 300.30.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 176

freq. of 0 ppm: 125.730106 MHz
processed size: 32768 complex points
LB: 1.000 G B: 0.0000

BnO-CH2-CH=CH-Br

2.103
SpinWorks 2.3:

<table>
<thead>
<tr>
<th>PPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.7112</td>
</tr>
<tr>
<td>7.6948</td>
</tr>
<tr>
<td>7.3241</td>
</tr>
<tr>
<td>7.3152</td>
</tr>
<tr>
<td>7.2624</td>
</tr>
<tr>
<td>5.1648</td>
</tr>
<tr>
<td>5.1506</td>
</tr>
<tr>
<td>5.1364</td>
</tr>
<tr>
<td>5.0786</td>
</tr>
<tr>
<td>5.0760</td>
</tr>
<tr>
<td>5.0732</td>
</tr>
<tr>
<td>5.0650</td>
</tr>
<tr>
<td>5.0623</td>
</tr>
<tr>
<td>5.0597</td>
</tr>
<tr>
<td>5.0541</td>
</tr>
<tr>
<td>5.0510</td>
</tr>
<tr>
<td>5.0484</td>
</tr>
<tr>
<td>5.0458</td>
</tr>
<tr>
<td>5.0202</td>
</tr>
<tr>
<td>5.0083</td>
</tr>
<tr>
<td>4.9324</td>
</tr>
<tr>
<td>4.9117</td>
</tr>
<tr>
<td>4.4805</td>
</tr>
<tr>
<td>3.8622</td>
</tr>
<tr>
<td>3.8559</td>
</tr>
<tr>
<td>3.8402</td>
</tr>
<tr>
<td>3.8345</td>
</tr>
<tr>
<td>3.8189</td>
</tr>
<tr>
<td>3.8126</td>
</tr>
<tr>
<td>3.4717</td>
</tr>
<tr>
<td>2.7451</td>
</tr>
<tr>
<td>2.7377</td>
</tr>
<tr>
<td>2.7282</td>
</tr>
<tr>
<td>2.7187</td>
</tr>
<tr>
<td>2.7092</td>
</tr>
<tr>
<td>2.7021</td>
</tr>
<tr>
<td>2.4935</td>
</tr>
<tr>
<td>2.4030</td>
</tr>
<tr>
<td>1.9303</td>
</tr>
<tr>
<td>1.6592</td>
</tr>
<tr>
<td>1.5762</td>
</tr>
<tr>
<td>1.5597</td>
</tr>
<tr>
<td>1.3938</td>
</tr>
<tr>
<td>1.2795</td>
</tr>
<tr>
<td>1.1918</td>
</tr>
<tr>
<td>1.1033</td>
</tr>
<tr>
<td>0.8554</td>
</tr>
<tr>
<td>0.8365</td>
</tr>
</tbody>
</table>

**File:** C:\Users\BW\Desktop\Spec Data\NMR Data\Wzy Project\Dolichol Derivative\Dolichol Coupled\fid expt: zg30

transmitter freq.: 500.023088 MHz
time domain size: 65536 points
width: 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
number of scans: 128

freq. of 0 ppm: 500.020022 MHz
processed size: 32768 complex points
LB: 0.000 G B: 0.0000
SpinWorks 2.3:

PPM

170.0
160.0
150.0
140.0
130.0
120.0
110.0
100.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0

21524.196
21507.097
21483.011
21462.568
21439.880
21436.796
21313.039
17041.599
17035.239
17030.439
17023.462
16234.190
16232.556
16221.351
16212.253
16127.152
16116.387
12736.948
12103.932
12097.630
9519.910
9126.067
9060.787
8903.688
8838.447
8817.468
8811.987
8805.509
8800.095
8586.131
7783.849
7744.526
6924.025
6553.940
6545.764
2929.571
2879.944
2644.325
2618.962
2612.201
2607.774

file: J:\Chitobiose Protected Phosphate 3\2\fid
expt: <zgpg30>

transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt

number of scans: 393

freq. of 0 ppm: 125.730107 MHz
processed size: 32768 complex points
LB:    1.000    GB: 0.0000
SpinWorks 2.3:

file: C:\Users\BW\Desktop\Spec Data\NMR Data\Project\Mocatetol\GalNAc\ prefs\2\expt <zgpg30>

transmitter freq.: 125.742702 MHz

time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt

number of scans: 13177

freq. of 0 ppm: 125.729939 MHz

processed size: 32768 complex points

LB: 0.000 G B: 0.0000

Precursor to 4.1
SpinWorks 2.3: 041201 in MeOH-d4 at rt on 500 MHz

Precursor to 4.2
SpinWorks 2.3:

transmitter freq.: 125.742702 MHz

file: C:\Users\BW\Desktop\SpecData\NMRData\NMRData\Expt=gpg30

processed size: 32768 complex points

freq. of 0 ppm: 125.730129 MHz

processed expt. G950-spectrograms

number of scans: 642
Precursor to 4.4
Precursor to 4.4

SpinWorks 2.3:

HO       P-O       P-O
\   \     \      |
AcNH  O     NH\      |
  OH     O     NH_4   |

file: C:\Users\BW\Desktop\Spec Data\NMR DATA\Wzy Project\cis-Farn Synthesis\GalNAcPPcisFarnPhos\1\fid   expt: <zgpg30>
transmitter freq.: 161.967474 MHz
time domain size: 65536 points
width: 64935.06 Hz = 400.914228 ppm = 0.990830 Hz/pt
number of scans: 101

freq. of 0 ppm: 161.975498 MHz
processed size: 32768 complex points
LB:    0.000    GB: 0.0000
SpinWorks 2.3:

- Transmitter freq.: 125.742702 MHz
- Time domain size: 65536 points
- Width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
- Number of scans: 894
- Freq. of 0 ppm: 125.730104 MHz
- Processed size: 32768 complex points
- LB: 1.000000 GB: 0.000000

**Bn**

![NMR spectrum of a sample](image)

**4.5**
SpinWorks 2.3:

HO\[-CH=CH-CH=CH-CH=CH-\]

4.6

file: C:\Users\BW\Desktop\Spec Data\NM R Data\Project.cs\FarnSynth1a\transmitter freq.: 500.023088 MHz
time domain size: 65536 points
width: 10330.58 Hz = 20.660203 ppm = 0.157632 Hz/pt
number of scans: 64
freq. of 0 ppm: 500.020022 MHz
processed size: 32768 complex points
SpinWorks 2.3:

file: C:\Users\BW\Desktop\Spec Data\NM R DATA\Wzy Project\FarnSynt\cisFarn\2\fid
transmitter freq.: 125.742702 MHz
time domain size: 65536 points
width: 30030.03 Hz = 238.821256 ppm = 0.458222 Hz/pt
number of scans: 354

freq. of 0 ppm: 125.730102 MHz
processed size: 32768 complex points
LB: 1.000
GB: 0.0000

4.6