MOLECULAR APPROACHES FOR CHEMOTHERAPEUTIC TARGETING

A thesis presented by

Krista Marie Wager

to

The Department of Chemistry & Chemical Biology

In partial fulfillment of the requirements for the degree of

Master of Science

in the field of

Chemistry

Northeastern University
Boston, Massachusetts
May, 2012
MOLECULAR APPROACHES FOR CHEMOTHERAPEUTIC TARGETING

by

Krista Marie Wager

ABSTRACT OF THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Chemistry in the Graduate School of Northeastern University

May, 2012
ABSTRACT

Given the pressures the pharmaceutical industry is under (budgetary, regulatory, environmental, etc.), it is vital that the methods to developing safe and innovative products are continually refined. Approaches to doing so may include the development of more efficient/streamlined synthetic routes to targeted compounds, novel techniques to improve the targeting or biological properties of pharmaceuticals, innovative repurposing at the end of their patent-life, or increased use of non-invasive imaging techniques to make an early go/no-go decision. This thesis highlights three such areas of research, each an avenue towards enhancing small-molecule pharmaceuticals or better (earlier) assessment of their potential as drug candidates. Chapter 1 describes the development and optimization of advantageous transition metal catalyzed cross coupling methodology for the synthesis of aryl thioethers. Enhancement of biological properties will be examined through PEGylation, antibody, and ammonium salt bioconjugation in Chapter 2. Chapter 3 will explore radio-iodination methodologies for the production of SPECT imaging agents, along with a discussion of SPECT imaging and important considerations in radiosynthesis. As a conclusion, Chapter 4 will discuss the current industry shift in focus away from small-molecule to biopharmaceutical development, including a comparison between these therapies as well as considerations pertinent to biopharmaceutical design, development, manufacturing, safety, regulation, and patent protection.
DEDICATION

This work is dedicated to Aelred A. Koepfer.
ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Graham Jones, for his guidance, encouragement, and genuine interest in helping me grow to my full potential, as well as all of the Jones lab members both past and present. You have all been a wonderful family that will truly be missed.

Additionally I would like to thank Merck Research Laboratories Boston for allowing me the opportunity to Co-op there and welcoming me with open arms. In particular, I would like to acknowledge Matthew Daniels, Jonathan Young, as well as the members of the MRL Boston Chemistry team for their continued guidance, support, and invaluable intellectual contributions to the work.
TABLE OF CONTENTS

ABSTRACT.................................................................................................................................................. 3

ACKNOWLEDGEMENTS.......................................................................................................................... 5

LIST OF SCHEMES .................................................................................................................................... 9

LIST OF TABLES ....................................................................................................................................... 11

PREFACE .................................................................................................................................................... 12

CHAPTER 1. Palladium-Catalyzed Cross-Coupling of Benzyl Thioacetates and Aryl Halides . 16
  1.1 Introduction ....................................................................................................................................... 16
  1.2 Method Development, Optimization, and Scope Exploration ......................................................... 17
  1.3 Summary .......................................................................................................................................... 26
  1.4 Acknowledgment .............................................................................................................................. 26
  1.5 Experimental .................................................................................................................................... 27
  1.6 Literature .......................................................................................................................................... 101

CHAPTER 2. Synthesis, Functionalization and Photo-bergman Chemistry of Enediyne
  Bioconjugates .......................................................................................................................................... 104
  2.1 Introduction ....................................................................................................................................... 104
  2.2 Previous Work in our Laboratory.................................................................................................... 105
    2.2.1 Template Design ....................................................................................................................... 105
    2.2.2 PEG-ylation and Antibody Conjugation of Mono-Alcohol Template................................. 106
  2.3 Synthesis of Bis-functionalized enediyne templates ....................................................................... 111
    2.3.1 Bis-Alcohol ............................................................................................................................... 111
  2.4 Synthesis of Water Soluble Ammonium Salts .............................................................................. 112
2.5 Succinic acid derivatization and PEG-ylation of the bis-alcohol templates .......... 115
2.6 Summary ............................................................................................................. 116
2.7 Experimental ...................................................................................................... 117
2.8 Literature .......................................................................................................... 146

CHAPTER 3. Radio-iodination Methods for the Production of SPECT Imaging Agents ...... 147
3.1 Introduction ......................................................................................................... 147
3.2 Background to SPECT Imaging ......................................................................... 147
3.3 General Considerations for Radio-iodination .................................................... 148
3.4 Nucleophilic Substitution Methods .................................................................... 150
  3.4.1 Isotopic Exchange ...................................................................................... 151
  3.4.2 Interhalogen Exchange .............................................................................. 152
  3.4.3 Radioiodo-dediazonization (using the Wallach triazene reaction) .............. 154
3.5 Electrophilic Substitution Methods .................................................................... 155
  3.5.1 Halo-demetallation (via Iododestannylation) .............................................. 155
  3.5.2 Halo-demetallation (via Radioiodethallation) ............................................ 161
3.6 Purification Considerations ................................................................................ 163
3.7 Summary ........................................................................................................... 165
3.8 Literature .......................................................................................................... 166

CHAPTER 4. Making Biosimilars a Reality: Comparison to Small-Molecules and Host Selection for Biopharmaceuticals and Biosimilar Development ............................................. 168
4.1 Introduction ......................................................................................................... 168
4.2 Biopharmaceutical Comparison to Small-molecule Pharmaceuticals .................... 177
4.3 Regulatory Considerations ................................................................................ 180
4.4 Immunogenicity & Post-Translational Modifications........................................... 184

4.4.1 Glycosylation.................................................................................................. 187

4.4.2 Glycoengineering............................................................................................ 190

4.4.3 Post-translational Modifications in Biosimilar Development......................... 192

4.5 Host Selection .................................................................................................... 192

4.5.1 Prokaryotic Expression Systems...................................................................... 195

  Bacteria................................................................................................................... 195

4.5.2 Eukaryotic Expression Systems....................................................................... 201

  Fungi....................................................................................................................... 201

  Plant......................................................................................................................... 212

  Insect...................................................................................................................... 221

  Mammalian............................................................................................................. 224

4.6 Biosimilar Outlook.............................................................................................. 231

4.7 Literature............................................................................................................ 233
LIST OF SCHEMES

Scheme 1.1. Proposed thioacetate coupling .......................... 17
Scheme 1.2. Proposed mechanism for Aryl/Aryl exchange ............... 21
Scheme 2.1. Bergman cyclization of (Z)-hex-3-ene-1,5-diyne .......... 104
Scheme 2.2. Initial synthesis of mono-alcohol enediyne template ......... 106
Scheme 2.3. PEGylation and antibody-enediyne conjugation ............... 109
Scheme 2.4. Photo-Bergman cyclization of enediyne core and bioconjugates ............... 110
Scheme 2.5. Preparation of gold modified nanoparticle-enediyne conjugate ............... 110
Scheme 2.6. Synthesis of enediyne diol .................................. 112
Scheme 2.7. Attempted conversion of enediyne benzyl alcohols to leaving groups ............... 114
Scheme 2.8. Synthesis of tri-alkyl ammonium salts .......................... 115
Scheme 2.9. Succinic acid derivatization and PEGylation of the bis-alcohol templates ............... 116

Scheme 3.1. Isotopic exchange radiolabeling of N-(2-diethylaminoethyl)-4-iodobenzamide Analogs .......................................................... 152
Scheme 3.2. Interhalogen exchange radio-iodination .......................... 153
Scheme 3.3. Radiiodo-dediazonization of leucinal derivatives ............... 154
Scheme 3.4. $^{123}$I radiolabeling of benzamides via chloramine-T induced halodestannylation .... 157
Scheme 3.5. $^{131}$I-IMTO radio-iodination using chloramine-T ............... 157
Scheme 3.6. $^{125}$I radio-iodination of benzodiazepines using iodobeads ............... 158
Scheme 3.7. Radiolabeling of quinazolinone derivatives with iodogen ............... 160
Scheme 3.8. Radiolabeling of aurones through halodestannylation using $\text{H}_2\text{O}_2$ and HCl ............... 161
Scheme 3.9. $^{125}$I labeling of styrylpyridine species using $\text{H}_2\text{O}_2$ and HCl ............... 161
Scheme 3.10. NCA labeling of benzamides with $^{123}$I via radioiododethallation ............... 162
Scheme 3.11. Radio-iodination of fluorous ureas following the FLS.
LIST OF TABLES

Table 1.1. Optimization of Thioacetate Coupling with PhBr. ...................................................... 19
Table 1.2. Pd-Catalyzed Reaction of 1a with Aryl Halides .......................................................... 23
Table 1.3. Pd-Catalyzed Reaction of Thioacetates 1 with 2c ...................................................... 24
Table 1.4. One-Port formation of Benzyl-Aryl Thioethers .......................................................... 25
PREFACE

Recent shifts in the pharmaceutical industry indicate the great deal of pressure it is experiencing. Evidence includes the many large acquisitions, such as Pfizer’s acquisition of Wyeth and Merck & Co.’s of Shering-Plough, and the growing trend of outsourcing as many pharmaceutical companies divest their R&D efforts which were once largely kept in-house. Pressure to increase revenue while decreasing the time and cost of drug development is intensified by time limitations on patents, increased scrutiny from the FDA on developers, and the need to produce innovative therapies that can be sold at affordable prices. Given the rising costs of development and, in many respects, antiquated approaches to doing so some are urging it is time for a new paradigm of drug development.

Strategies to address these pressures include the development of novel techniques to improve the targeting or biological properties pharmaceuticals, innovative repurposing at the end of their patent-life, more efficient/streamlined synthetic approaches to targeted compounds, or increased use of non-invasive imaging techniques to make an early go/no-go decision. This thesis highlights three such areas of research, each an avenue towards enhancing small-molecule pharmaceuticals or better (earlier) assessment of their potential as drug candidates.

Chapter 1 describes the development of transition metal catalyzed cross coupling conditions for the synthesis of aryl thioethers. Aryl thioethers attract significant interest in medicinal chemistry programs due to their prevalence in biologically active molecules. Although several approaches exist for the formation of alkyl-aryl thioethers, these frequently rely on cross coupling aryl
halides to alkyl thiols - compounds known to be both unpleasantly odorous as well as sensitive to oxidation and dimerization. A strategy to alleviate these disadvantages is the addition of a protecting group, one compatible with cross coupling chemistry, to the alkyl thiols, which can then undergo in situ deprotection and subsequent arylation. This approach was examined, resulting in the development of an improved method for the synthesis of aryl thioethers which is highly efficient and applicable to both alkyl thiols and aryl halides bearing diverse chemical functionalities\(^4\).

Beyond the development of more effective synthetic methods, another approach to increasing efficiency in small-molecule pharmaceutical development is enhancement of the biological properties of molecules known to be biologically active. This approach can involve more sophisticated molecular designs to enhance the desired biological effect prior to clinical testing, as well as the introduction of new and/or improved functionalities within an approved therapeutic molecule to reduce doses in the patient or for repurposing of drugs upon the expiration of their patents. In Chapter 2 examples of such strategies will be provided via the discussion of the rational design and synthesis of potential anti-tumor prodrugs for use as targeted photodynamic therapies\(^5,6\).

Although the number of candidate drugs entering clinical trials has increased, the number of candidates making it all the way to Phase III trials hasn’t changed much since 1997\(^3\). This is problematic and may be the result of insufficient capacity to valuate drug properties and clinical potential of candidates prior to entrance in clinical trials\(^3\). One strategy for reducing the number of reduce the amount of unsuitable candidates reaching clinical trials is via increased use of
imaging technologies, such as Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT), to provide *in vivo* evidence of a drug candidate’s applicability to a particular biological target or disease state early on in the drug development process\(^7\). These noninvasive techniques provide invaluable visualization of biological processes at tissue, cellular, and sub-cellular levels\(^8\). Chapter 3 will survey recent advances in radioiodination methodology, including synthetic chemistry and post-labeling purification methods, for the rapid and efficient synthesis of diagnostic imaging agents. The implementation of such imaging agents early into the drug discovery and development process is promising and may reduce the number of entrants into costly clinical trials.

In conclusion, this thesis will discuss what is perhaps the most dramatic shift the industry has seen in recent years. As biomedical and genetic researchers continue to decipher biological mechanisms, as well as the underlying causes of their breakdown which lead to disease, we are able to develop increasingly sophisticated biologic drugs, or biopharmaceuticals. Perhaps due to the stagnancy surrounding traditional small-molecule drugs at present or the fact that biopharmaceuticals have proven effective in treating many diseases once considered incurable\(^9\),\(^10\), a dramatic shift has been seen in the pharmaceutical industry away from small-molecule development and towards that of biopharmaceuticals\(^11\)-\(^13\). Regardless of what has led to this shift, biopharmaceuticals bring with them a unique set of complex issues relating to their design, development, manufacturing, safety, regulation, and patent protection. Given their relatively new introduction as pharmaceutical products compared to small-molecules, industry and regulatory members attempting to bring these products to market have a difficult road to navigate. Chapter 4 of this thesis will provide an overview of these issues in greater details, including a discussion
of the cell lines available for biopharmaceutical production.

Literature

CHAPTER 1. Palladium-Catalyzed Cross-Coupling of Benzyl Thioacetates and Aryl Halides

1.1 Introduction

The prevalence of alkyl aryl thioethers (and their derivatives) in natural products and pharmaceuticals has led to the development of many methods for C-S bond construction. Metal-catalyzed cross-coupling reactions have proven to be a robust technique, with Pd and Cu emerging as excellent catalysts in various applications. These methods, however, rely on the use of alkyl thiols, which can be odorous and sensitive to oxidation. To circumvent this issue, diaryl disulfides have been coupled with alkyl halides. There have also been reports of in situ cleavage of thioesters for use in SNAr reactions. Hartwig and co-workers have used Pd coupling to prepare unsymmetrical biaryl thioethers from aryl halides and TIPS thiol, while others have utilized copper and a xanthate salt. Potassium thioacetate has also been employed in a Cu-catalyzed coupling with an aryl halide followed by hydrolysis and reaction with an alkyl halide. Recognizing the utility of a protected thiol, we sought to optimize the coupling of S-benzyl thioesters with aryl halides, believing that an in situ deprotection would be compatible with known cross-coupling conditions (Scheme 1.1).
**Scheme 1.1.** Proposed thioacetate coupling.

![Scheme 1.1. Proposed thioacetate coupling.](image)

**1.2 Method Development, Optimization, and Scope Exploration**

The thioacetate protecting group, a common intermediate in the synthesis of thiols in a variety of applications\(^{53-56}\), was chosen for its ease of installation, simplicity of deprotection, and low molecular weight. Most examples of this transformation describe the proposed coupling as an undesired side reaction\(^{57}\) or are limited to a single example\(^{58, 59}\). After considering several known conditions\(^{38, 60, 61}\), we began our investigations with Pd\(_2\)dba\(_3\), Xantphos, and DIPEA in 1,4-dioxane\(^{27}\). Unfortunately, this was not effective in coupling thioacetate 1a with bromide 2a (Table 1.1, entry 1).

To enable this transformation, we examined various bases and solvents. Organic bases such as DIPEA and Me\(_2\)NH in various solvents were ineffective (entries 1-4). In light of this, both strong and weak inorganic bases were investigated using wet or dry 1,4-dioxane as a solvent. Aqueous conditions were favored, with the mild base potassium carbonate emerging as the preferred additive (entry 8). The addition of water likely improved the rate of hydrolysis of thioacetates. A solvent screen revealed aqueous THF as the superior solvent (entry 13), and it was determined that only 1 equiv of base was required to obtain good yields (entry 18). Attempts to convert 1a to 3a in the absence of
either base or palladium failed to give product. Further experiments (all with 2 equiv of 
K₂CO₃ in wet THF) revealed that a reaction temperature of 100 °C gave a slightly higher 
yield (85%), performing the reaction at 60 °C gave a reduced yield (20%), and using 
Pd(OAc)₂ or PdCl₂ also provided product (73% and 63%, respectively).
Table 1.1. Optimization of Thioacetate Coupling with PhBr$^a$.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Base</th>
<th>Solvent</th>
<th>Water$^b$</th>
<th>Yield$^c$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DIPEA</td>
<td>1,4-dioxane</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>DIPEA</td>
<td>MeOH</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>DIPEA</td>
<td>1,4-dioxane</td>
<td>4:1</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>DMA</td>
<td>1,4-dioxane</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>NaO$^t$Bu</td>
<td>1,4-dioxane</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>NaO$^t$Bu</td>
<td>1,4-dioxane</td>
<td>4:1</td>
<td>28</td>
</tr>
<tr>
<td>7</td>
<td>K$_2$CO$_3$</td>
<td>1,4-dioxane</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>8</td>
<td>K$_2$CO$_3$</td>
<td>1,4-dioxane</td>
<td>4:1</td>
<td>74</td>
</tr>
<tr>
<td>9</td>
<td>KOH</td>
<td>1,4-dioxane</td>
<td>0</td>
<td>66</td>
</tr>
<tr>
<td>10</td>
<td>KOH</td>
<td>1,4-dioxane</td>
<td>4:1</td>
<td>61</td>
</tr>
<tr>
<td>11</td>
<td>NaOMe</td>
<td>1,4-dioxane</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>NaOAc</td>
<td>1,4-dioxane</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>K$_2$CO$_3$</td>
<td>THF</td>
<td>4:1</td>
<td>78</td>
</tr>
<tr>
<td>14</td>
<td>K$_2$CO$_3$</td>
<td>Toluene</td>
<td>4:1</td>
<td>42</td>
</tr>
<tr>
<td>15</td>
<td>K$_2$CO$_3$</td>
<td>MeOH</td>
<td>4:1</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>K$_2$CO$_3$</td>
<td>DMF</td>
<td>4:1</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>K$_2$CO$_3$$^d$</td>
<td>THF</td>
<td>4:1</td>
<td>79</td>
</tr>
<tr>
<td>19</td>
<td>K$_2$CO$_3$</td>
<td>THF</td>
<td>4:1</td>
<td>60</td>
</tr>
<tr>
<td>20</td>
<td>K$_2$CO$_3$</td>
<td>THF</td>
<td>4:1</td>
<td>68</td>
</tr>
<tr>
<td>21</td>
<td>K$_2$CO$_3$</td>
<td>THF</td>
<td>4:1</td>
<td>27</td>
</tr>
</tbody>
</table>

$^a$ Reactions run on 0.5 mmol of 1a at 0.4 M for 14-18 h. Reactions were not optimized for time but were generally complete after 5 h. $^b$ Ratio of solvent to water. $^c$ Isolated yield of pure 3a. $^d$ Reaction run with 1 equiv of base.

With this optimization in hand, we elected to determine the scope of the reaction using the Pd$_2$dba/Xantphos system and 1 equiv of K$_2$CO$_3$ in 4:1 THF/water at 100 °C. To our surprise, initial experiments with 2b (see Table 1.2 for structures) provided a 7:1 mixture of the desired product 3b and compound 3a. Indeed, 3a was observed with a variety of
aryl bromide substrates even when using rigorously purified materials. At this point, we suspected that the phenyl group found in the 3a byproduct came from either degradation or transfer of a phenyl group from Xantphos. By closely monitoring the reaction of 1a and 2b, it was determined that 3a formed concomitantly with 3b. Meanwhile, subjecting compound 3b by itself to the standard reaction conditions gave no 3a, ruling out the degradation of 3b. Switching from Pd₂dba₃ to Pd(OAc)₂ did not change the product distribution, but using commercially available tert-butyl Xantphos gave solely compound 3b. This result suggested a process of aryl/aryl transfer between a phenyl group on Xantphos and Pd-bound 2b. Such a transfer has been observed previously, mostly with monodentate ligands such as triphenylphosphine⁶²-⁶⁴. Notably, Hartwig observed an aryl/aryl exchange process while studying palladium-catalyzed C-S coupling reactions²⁸,⁶⁵. A likely mechanism based on that proposed by Chenard et al. is shown in Scheme 1.2⁶².

Initial oxidative addition of 2b with Xantphos-bound Pd(0) would form 4, where the ligand X is either bromide or thiolate depending on whether the aryl/aryl exchange process occurs prior to or after the expected σ bond metathesis⁶⁶.

Reductive elimination provides species 5, which then undergoes oxidative addition either at a C-Ph bond (to regenerate 4) or at the C-Ar bond (providing 6), with 6 providing 3a.
To support this hypothesis, we reacted 1a with 2a-d5 under the previously optimized conditions (Table 1.1, entry 18); this provided a mixture of 3.5:1 3a-d5:3a. The inclusion of 3a again implicates Xantphos as the source of the phenyl group via an aryl/aryl exchange process. Additionally, 31P NMR analysis of Xantphos heated in the presence of Pd2dba3 and excess 2b confirmed the formation of phosphorus-containing species distinct from 467. After heating for several hours at 100 °C, we observed several doublets in the 31P NMR spectrum downfield of the both the free and Pd-bound Xantphos singlets. Though not conclusive, this is consistent with aryl/aryl exchange disrupting the symmetry of the phosphorus atoms in Xantphos.

A brief ligand screen showed that XPhos68 provided good yields of 3b (66%) under the previous reaction conditions with no formation of 3a. We proceeded to explore the substrate scope of our optimized one-pot deprotection/cross-coupling reaction (Table 1.2). These conditions were effective for a diverse array of substrates, including electron-poor (entries 1-6) and electron-rich (entries 7-8) aryl bromides. Sterically demanding aryl bromides (entries 10, 11, and 14) coupled efficiently, as did base-sensitive functional
groups including methyl esters and nitriles (entries 2, 3, and 14). No hydrolysis of methyl ester products to the corresponding carboxylic acids was observed. Heterocycles such as 6-bromoquinoline 2n coupled smoothly (entry 13), but 2-chloropyridine gave complex reaction mixtures. Finally, bromobenzene provided very good yields of 3a13.
Table 1.2. Pd-Catalyzed Reaction of 1a with Aryl Halides$^a$.

![Reaction Diagram]

<table>
<thead>
<tr>
<th>Entry</th>
<th>ArX</th>
<th>Product</th>
<th>Yield$^b$</th>
<th>Entry</th>
<th>ArX</th>
<th>Product</th>
<th>Yield$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Br-COMe</td>
<td>2b 3b</td>
<td>66</td>
<td>8</td>
<td>Br-NMe$_2$</td>
<td>2i 3i</td>
<td>68</td>
</tr>
<tr>
<td>2</td>
<td>Br-CO$_2$Me</td>
<td>2c 3c</td>
<td>89</td>
<td>9</td>
<td>Br-Cl</td>
<td>2j 3j</td>
<td>67</td>
</tr>
<tr>
<td>3</td>
<td>Br-CN</td>
<td>2d 3d</td>
<td>72</td>
<td>10</td>
<td>Br-Me</td>
<td>2k 3k</td>
<td>74</td>
</tr>
<tr>
<td>4</td>
<td>Br-F</td>
<td>2e 3e</td>
<td>88</td>
<td>11</td>
<td>Br-Me</td>
<td>2l 3l</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Br-CF$_3$</td>
<td>2f 3f</td>
<td>71</td>
<td>12</td>
<td></td>
<td>2m 3m</td>
<td>96</td>
</tr>
<tr>
<td>6</td>
<td>Br-NO$_2$</td>
<td>2g 3g</td>
<td>96</td>
<td>13</td>
<td></td>
<td>2n 3n</td>
<td>82</td>
</tr>
<tr>
<td>7</td>
<td>Br-OMe</td>
<td>2h 3h</td>
<td>53</td>
<td>14</td>
<td>Br-CO$_2$Me</td>
<td>2o 3o</td>
<td>63</td>
</tr>
</tbody>
</table>

$^a$Reactions run on 0.5 mmol 1a at 100 °C for 15-18 h. $^b$Isolated yield of pure compounds 3. $^c$Reaction run on 0.3 mmol of 1a.

Variation of the thioacetate component 1 was also explored (Table 1.3). Electron-rich (entry 1), electron-poor (entry 3), and sterically hindered (entry 5) thioacetates were tolerated. Thiobenzoates and thiopivalates were also effective (data not shown).
Table 1.3. Pd-Catalyzed Reaction of Thioacetates 1 with 2c.<sup>a</sup>

<table>
<thead>
<tr>
<th>Entry</th>
<th>RS&lt;sub&gt;S&lt;/sub&gt;Ac</th>
<th>Product</th>
<th>Yield (%)&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>phenyl&lt;sub&gt;C&lt;/sub&gt;H&lt;sub&gt;4&lt;/sub&gt;CO&lt;sub&gt;2&lt;/sub&gt;S&lt;sub&gt;S&lt;/sub&gt;R&lt;sub&gt;S&lt;/sub&gt;Ac</td>
<td>1b 3p</td>
<td>90&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>2</td>
<td>benzyl&lt;sub&gt;C&lt;/sub&gt;H&lt;sub&gt;2&lt;/sub&gt;CH&lt;sub&gt;2&lt;/sub&gt;</td>
<td>1c 3q</td>
<td>78</td>
</tr>
<tr>
<td>3</td>
<td>4-fluorobenzyl&lt;sub&gt;F&lt;/sub&gt;C&lt;sub&gt;2&lt;/sub&gt;S&lt;sub&gt;S&lt;/sub&gt;R&lt;sub&gt;S&lt;/sub&gt;Ac</td>
<td>1d 3r</td>
<td>55</td>
</tr>
<tr>
<td>4</td>
<td>4-chlorobenzyl&lt;sub&gt;Cl&lt;/sub&gt;S&lt;sub&gt;S&lt;/sub&gt;R&lt;sub&gt;S&lt;/sub&gt;Ac</td>
<td>1e 3s</td>
<td>62</td>
</tr>
<tr>
<td>5</td>
<td>4-bromobenzyl&lt;sub&gt;Br&lt;/sub&gt;S&lt;sub&gt;S&lt;/sub&gt;R&lt;sub&gt;S&lt;/sub&gt;Ac</td>
<td>1f 3t</td>
<td>73</td>
</tr>
</tbody>
</table>

<sup>a</sup>Reactions run on 0.5 mmol of 1a at 100 °C for 15-18 h and were not optimized for time. <sup>b</sup>Isolated yield of pure compounds 3. <sup>c</sup>Based on 90% pure product.

Thus far, we had demonstrated the ability to produce a variety of benzyl aryl thioethers through the palladium-catalyzed cross-coupling of aryl halides with thioacetates. We were interested to see if this procedure could be adapted to a more convenient one-pot protocol wherein an alkyl halide is reacted with potassium thioacetate, a readily available solid sulfur source, followed by immediate coupling in the same reaction vessel. Similar
one-pot, three-component couplings of an aryl halide, an alkyl halide, and a sulfur source are known\(^6\).

Gratifyingly, treatment of bromide \(7a\) with potassium thioacetate (THF, 2 h, 60 °C) followed by \textit{in situ} deprotection and cross-coupling \(1c\) with \(2o\) gave \(3u\) in 96% yield, highlighting the efficiency of both steps of this reaction. Table 1.4 shows several benzyl aryl thioethers prepared in good to excellent yields. Unfortunately, this reaction failed to give product with alkyl halides as starting material.

**Table 1.4. One-Port formation of Benzyl-Aryl Thioethers\(^a\).**

<table>
<thead>
<tr>
<th>Entry</th>
<th>(R)</th>
<th>(2)</th>
<th>Product</th>
<th>Yield (%)(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(H)</td>
<td>(7a)</td>
<td>(2o)</td>
<td>(3u) 96</td>
</tr>
<tr>
<td>2</td>
<td>(CF_3)</td>
<td>(7b)</td>
<td>(2c)</td>
<td>(3r) 4</td>
</tr>
<tr>
<td>3</td>
<td>(t)-Bu</td>
<td>(7c)</td>
<td>(2f)</td>
<td>(3v) 65</td>
</tr>
</tbody>
</table>

\(a\)Reactions run on 0.5 mmol \(7\) with 1.1 equiv of KSAc at 60 °C for 3 h, followed by addition of remaining reagents and heating at 100°C for 15-18 h. No Attempt was made to optimize for time. \(b\) Isolated yield of pure compounds 3.
1.3 Summary

In summary, a method for synthesizing benzyl aryl thioethers from easy to handle and readily available thioacetates has been developed. Deacylation and subsequent cross-coupling of thioacetates 1 can be achieved in the presence of mild aqueous base in up to 96% yield. In situ deprotection of thioacetates avoids the use of often odorous and unstable thiols in cross-coupling chemistry. As an even more streamlined procedure, we have demonstrated that a one-pot thioacetate formation/deacylation/ cross-coupling sequence provides compounds 3 in moderate to excellent yields.

1.4 Acknowledgment

This work was conducted at Merck Research Laboratories, Boston, as part of the Cooperative Education Program with Northeastern University, Boston, MA.

I would like to thank Matthew Daniels for his guidance and continued encouragement. I thank Merck colleagues Charles W. Ross III for HRMS analysis of all compounds, Bruce Adams for assistance with NMR experiments and analysis, and Xin Linghu for helpful discussions.
1.5 Experimental

**General experimental information:** All reagents and solvents were commercially obtained and used as-is. Reactions were run under anaerobic conditions and a nitrogen atmosphere unless otherwise noted. Water was degassed following the method of Buchwald and coworkers. Reactions were monitored by TLC (60Å 250 µm, fluorescent indicator, UV detection) and/or LC/MS using a C18 column (2.5 µm, 30 x 2.1 mm). For LC/MS, mobile phases consisted of water/0.1 % TFA (A) and MeCN/0.1 % TFA (B). A typical gradient consisted of 3 % B for 0.1 min, 3-98 % B over 2 min and 98 % B for 0.3 min (flow rate of 0.9-1.6 mL/min) followed by ion detection using a single quad MS with electrospray ionization running in positive ion mode. Compounds were purified by flash chromatography on silica gel. Yields reported are based on greater than 95 % pure product unless otherwise noted. $^1$H NMR and $^{13}$C NMR spectra were obtained using a 500 MHz NMR at effective frequencies of 499.811 and 125.678 MHz, respectively. NMR spectra are reported as follows: chemical shift δ (ppm), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet), coupling constant (Hz) and integration. High resolving power accurate mass measurements (AMM) were acquired by use of a Fourier transform ion cyclotron resonance mass spectrometer. Samples were dissolved in acetonitrile:water (50:50) with 0.1 % acetic acid (v/v) or methanol:dichloromethane (50:50) with 0.1 % acetic acid and ionized by use of electrospray ionization (ESI). External calibration was accomplished with oligomers of polypropylene glycol (average molecular weight 1000 Da). Samples not amenable to ESI were analyzed by use of a
AMM GC/MS. Samples were taken up in methanol:dichloromethane and filtered as needed. Internal calibration was accomplished with perfluorotributylamine (PFTBA).

*S-(Quinolin-6-ylmethyl) ethanethioate (1a):* To an oven-dried round bottom flask with a stir bar was added 6-quinolinylmethanol (2.50 g, 15.7 mmol) in THF (16 mL). Once fully dissolved, thioacetic acid (1.34 mL, 18.9 mmol) was added and the reaction mixture was stirred for 15 minutes. A separate oven-dried round bottom flask was charged with di-tert-butyl azodicarboxylate (4.34 g, 18.9 mmol) and triphenylphosphine (4.94 g, 18.9 mmol) in THF (8 mL) and cooled to 0°C. The contents of the first flask were then added via syringe to the second. After stirring for 30 min at 0°C, the orange reaction mixture was warmed to room temperature and stirred overnight. Trifluoroacetic acid (4.84 mL, 62.8 mmol) was added and the reaction mixture was stirred at room temperature for 15 minutes, poured into saturated Na₂CO₃ solution (150 mL) at 0°C, stirred for 1.5 hours and extracted with EtOAc (3×100 mL). The combined organic layers were washed with brine (100 mL), dried over Na₂SO₄, filtered, concentrated in vacuo and purified by flash chromatography on silica gel (0-30 % EtOAc in hexanes) to yield 1a as a light brown oil (3.03 g, 89 %). ¹H NMR (CDCl₃) δ 8.85 (d, J = 3.7 Hz, 1H), 8.06 (d, J = 8.3 Hz, 1H), 8.04 (d, J = 8.7 Hz, 1H), 7.64 (s, 1H), 7.59 (d, J = 8.7 Hz, 1H), 7.42 – 7.30 (m, 1H), 4.25 (s, 2H), 2.52 (s, 3H). ¹³C NMR (CDCl₃) δ 195.1, 150.6, 147.7, 136.3, 136.1, 130.8, 130.1, 128.3, 127.5, 121.6, 33.6, 30.6. HRMS calc’d for C₁₂H₁₂NOS [M+H]⁺: 218.0634, found: 218.0628.
**General procedure for the synthesis of thioacetates 1b-c,e-f:** To an oven-dried microwave vial with a stir bar was added 4-dimethylaminopyridine (0.02 equiv, 7.7-9.8 mg, 0.063-0.081 mmol), pyridine (0.75 equiv, 0.19-0.24 mL, 2.4-3.0 mmol), mercaptan (0.50 g, 3.2-4.0 mmol), dichloromethane (2 mL) and acetic anhydride (1.2 equiv, 0.36-0.46 mL, 3.8-4.8 mmol). The reaction mixture was allowed to stir at room temperature overnight. The mixture was diluted with EtOAc (5 mL) and quenched with 1N HCl (5 mL). After separation, the aqueous layer was extracted three times with EtOAc (5 mL). The combined organic layers were then washed three times with H2O (5 mL) and brine (5 mL), dried over Na2SO4, and concentrated in vacuo. The crude product was purified by flash chromatography on silica gel to afford thioacetates 1b-c and 1e-f.

**S-(4-Methoxybenzyl) ethanethioate (1b):** Scale: 0.46 mL (3.24 mmol) 4-methoxybenzyl mercaptan. Time: 15 hours. Clear oil. Purification: 0-5 % EtOAc in hexanes. Yield: 0.49 g, 77 % yield. $^1$H NMR (CDCl$_3$) $\delta$ 7.16 (d, $J = 8.5$ Hz, 2H), 6.89 (d, $J = 8.7$ Hz, 2H), 4.08 (s, 2H), 3.78 (s, 3H), 2.34 (s, 3H). $^{13}$C NMR (DMSO-$d_6$) $\delta$ 195.5, 159.0, 130.5, 130.2, 114.5, 55.6, 32.7, 30.8. HRMS calc'd for C$_{10}$H$_{12}$O$_2$S [M]$^+$: 196.0558, found: 196.0546.

**S-Benzyl ethanethioate (1c):** Scale: 0.48 mL (4.0 mmol) benzyl mercaptan. Time: 15 hours. Clear oil. Purification: 0-5 % EtOAc in hexanes. Yield: 211 mg, 32 %. $^1$H NMR (CDCl$_3$) $\delta$ 7.53 – 7.04 (m, 5H), 4.15 (s, 2H), 2.36 (s, 3H). $^{13}$C NMR (CDCl$_3$) $\delta$ 195.3, 137.9, 77.7, 77.4, 77.2, 33.7, 30.6. HRMS calc'd for C$_9$H$_{10}$OS [M]$^+$: 166.0452, found: 166.0435.
**S-(4-Chlorobenzyl) ethanethioate (1e):** Scale: 0.42 mL (3.2 mmol) 4-chlorobenzyl mercaptan. Time: 15 hours. Clear oil. Purification: 0-3 % EtOAc in hexanes. Yield: 601 mg, 95 %. $^1$H NMR (CDCl$_3$) $\delta$ 7.23 (dd, $J = 7.9, 19.0$ Hz, 4H), 4.06 (s, 2H), 2.34 (s, 3H). $^{13}$C NMR (CDCl$_3$) $\delta$ 195.1, 136.5, 133.3, 130.4, 129.0, 33.0, 30.6. HRMS calc'd for C$_9$H$_9$ClOS [M]$^+$: 200.0063, found: 200.0058.

**S-(1-Phenylethyl) ethanethioate (1f):** Scale: 0.50 mL (3.6 mmol) 1-phenylethyl mercaptan. Time: 15 hours. Clear oil. Purification: 0-6 % EtOAc in hexanes. Yield: 518 mg, 79 %. $^1$H NMR (CDCl$_3$) $\delta$ 7.46 – 7.29 (m, 4H), 7.26 (t, $J = 7.0$ Hz, 1H), 4.78 (q, $J = 7.2$ Hz, 1H), 2.31 (s, 3H), 1.68 (d, $J = 7.2$ Hz, 3H). $^{13}$C NMR (CDCl$_3$) $\delta$ 195.2, 142.9, 77.7, 77.4, 77.2, 43.2, 30.7, 22.5. HRMS calc'd for C$_{10}$H$_{12}$OS [M]$^+$: 180.0609, found: 180.0595.

**S-[4-(Trifluoromethyl)benzyl] ethanethioate (1d):** To an oven-dried microwave vial with a stir bar was added 4-(trifluoromethyl)benzyl bromide (0.323 mL, 2.09 mmol), anhydrous acetone (15 mL), and potassium thioacetate (0.287 g, 2.51 mmol). This mixture was vigorously stirred at 60°C for 18 hours. The reaction mixture was concentrated in vacuo, diluted with H$_2$O (50 mL) and washed with DCM (50 mL). The organic layer was washed three times with H$_2$O (50 mL) and the combined aqueous layers were back extracted three times with DCM (50 mL). The organic layers were washed with brine (50 mL), dried over Na$_2$SO$_4$, and concentrated in vacuo. The crude product was purified via flash chromatography on silica gel (0-3 % EtOAc in hexanes) to
yield 1d as a pale yellow oil (305 mg, 62%). $^1$H NMR (DMSO-$_d6$) $\delta$ 7.63 (d, $J = 8.2$ Hz, 2H), 7.50 (d, $J = 8.1$ Hz, 2H), 4.17 (s, 2H), 2.33 (s, 3H). $^{13}$C NMR (DMSO-$_d6$) $\delta$ 195.2, 143.7, 130.1, 128.4 (q, C-CF$_3$, $J = 32.2$ Hz), 126.0 (q, C-C-CF$_3$, $J = 3.8$ Hz), 124.9 (q, CF$_3$, $J = 272.3$ Hz), 32.6, 30.8. HRMS calc’d for C$_{10}$H$_9$F$_3$OS $[M]^+$: 234.0326, found: 234.0339.

**General procedure for deacylation/cross-coupling (Tables 1.1-1.3, substrates 3a-3t)**: To an oven-dried microwave vial with a stir bar was added Pd$_2$dba$_3$ (0.023 g, 0.025 mmol), XPhos (0.024 g, 0.050 mmol), aryl halides 2 (0.55 mmol), and K$_2$CO$_3$ (0.069 g, 0.50 mmol). After evacuating and purging the vial with nitrogen three times, degassed thioacetate 1 (0.50 mmol), THF (1.0 mL), and water (0.25 mL) were added to the reaction mixture. Liquid aryl halides were added with the thioacetates. The reaction mixture was stirred vigorously at 100°C, monitored by TLC and/or LC/MS and upon completion was cooled to room temperature. The cooled reaction mixture was diluted with EtOAc (2 mL), filtered though Celite and rinsed with EtOAc (100 mL). The filtrate was concentrated *in vacuo* and purified by flash chromatography over silica gel to afford substrates 3a-3t.

**6-[(Phenylsulfanyl)methyl]quinoline (3a)**: Time: 15 hours. Pale yellow solid. Purification: 5 % EtOAc in DCM. Yield: 107 mg, 85 %. $^1$H NMR (DMSO-$_d6$) $\delta$ 8.83 (s, 1H), 8.26 (d, $J = 8.2$ Hz, 1H), 7.94 (d, $J = 8.6$ Hz, 1H), 7.86 (s, 1H), 7.76 (d, $J = 8.6$ Hz, 1H), 7.54 (q, $J = 8.3$ Hz, 1H), 7.34 (d, $J = 7.5$ Hz, 2H), 7.26 (t, $J = 7.6$ Hz, 2H), 7.18 (t, $J = 7.1$ Hz, 1H), 4.43 (s, 2H). $^{13}$C NMR (DMSO-$_d6$) $\delta$ 151.1, 147.6, 136.6, 136.4, 136.3,
131.4, 129.8, 129.7, 129.3, 128.3, 128.0, 127.0, 122.4, 37.3. HRMS calc'd for C\textsubscript{16}H\textsubscript{14}NS [M+H]\textsuperscript{+}: 252.0841, found: 252.0836.

1-\{(Quinolin-6-ylmethyl)sulfanyl\}phenyl\}ethanone (3b): Time: 15 hours. Pale yellow solid. Purification: 30 % EtOAc in hexanes. Yield: 97 mg, 66 %. \textsuperscript{1}H NMR (CDCl\textsubscript{3}) \(\delta\) 8.89 (dd, \(J = 4.1\) Hz, 1.4 Hz, 1H), 8.08 (t, \(J = 8.5\) Hz, 2H), 7.82 (d, \(J = 8.5\) Hz, 2H), 7.74 (m, 2H), 7.39 (dd, \(J = 8.1\) Hz, 4.2 Hz, 1H), 7.33 (d, \(J = 8.6\) Hz, 2H), 4.38 (s, 2H), 2.54 (s, 3H). \textsuperscript{13}C NMR (CDCl\textsubscript{3}) \(\delta\) 197.3, 150.8, 147.9, 143.7, 136.1, 135.0, 134.7, 130.6, 130.3, 129.0, 128.3, 127.54, 127.45, 121.7, 37.5, 26.7. HRMS calc'd for C\textsubscript{18}H\textsubscript{16}NOS [M+H]\textsuperscript{+}: 294.0947, found: 294.0941.

6-\{[(4-Methoxyphenyl)sulfanyl]methyl\}quinoline (3c): Time: 15 hours. Pale yellow solid. Purification: 30 % EtOAc in hexanes. Yield: 137 mg, 89 %. \textsuperscript{1}H NMR (DMSO-\textit{d}_6) \(\delta\) 8.85 (d, \(J = 2.6\) Hz, 1H), 8.29 (d, \(J = 7.7\) Hz, 1H), 8.03 – 7.91 (m, 2H), 7.88 – 7.75 (m, 3H), 7.55 – 7.41 (m, 3H), 4.56 (s, 2H), 3.79 (s, 3H). \textsuperscript{13}C NMR (DMSO-\textit{d}_6) \(\delta\) 166.5, 151.2, 147.7, 144.0, 136.5, 135.8, 131.3, 130.2, 129.9, 128.3, 128.2, 127.3, 127.0, 122.5, 52.7, 35.8. HRMS calc'd for C\textsubscript{18}H\textsubscript{16}NO\textsubscript{2}S [M+H]\textsuperscript{+}: 310.0896, found: 310.0891.

4-\{(Quinolin-6-ylmethyl)sulfanyl\}benzonitrile (3d): Time: 15 hours. Yellow solid. Purification: 60 % Et\textsubscript{2}O in hexanes. Yield: 101 mg, 72 %. \textsuperscript{1}H NMR (DMSO-\textit{d}_6) \(\delta\) 8.85 (dd, \(J = 1.6\) Hz, 4.2 Hz, 1H), 8.30 (d, \(J = 7.4\) Hz, 1H), 8.03 – 7.91 (m, 2H), 7.80 (dd, \(J = 2.0\) Hz, 8.7 Hz, 1H), 7.70 (d, \(J = 8.4\) Hz, 2H), 7.50 (dd, \(J = 4.4\) Hz, 8.4 Hz, 3H), 4.58 (s, 2H). \textsuperscript{13}C NMR (DMSO-\textit{d}_6) \(\delta\) 151.3, 147.7, 144.6, 136.5, 135.6, 133.1, 131.2, 130.0,
128.3, 128.2, 127.7, 122.5, 119.5, 108.1, 35.6. HRMS calc'd for C\textsubscript{17}H\textsubscript{13}N\textsubscript{2}S [M+H]\textsuperscript{+}: 277.0794, found: 277.0787.

**6-[(4-Fluorophenyl)sulfanyl]methylquinoline (3e):** Time: 15 hours. Yellow solid. Purification: 5 % EtOAc in DCM. Yield: 119 mg, 88 %. \(^1\)H NMR (CDCl\textsubscript{3}) \(\delta\) 8.88 (s, 1H), 8.05 (t, \(J = 8.4 \text{ Hz}, 2\text{H}\)), 7.65 (d, \(J = 8.7 \text{ Hz}, 1\text{H}\)), 7.50 (s, 1H), 7.46 – 7.33 (m, 1H), 7.32 – 7.20 (m, 2H), 6.91 (t, \(J = 8.2 \text{ Hz}, 2\text{H}\)), 4.18 (s, 2H). \(^{13}\)C NMR (DMSO-d\textsubscript{6}) \(\delta\) 162.5 (d, C-F, \(J = 247.5\)), 161.5, 150.6, 147.8, 136.2, 136.0, 134.21 (d, C-C-C-F, \(J = 8.4 \text{ Hz}\)), 130.9, 130.31, 130.28, 130.0, 128.2, 127.4, 121.6, 116.3 (d, C-C-F, \(J = 21.9 \text{ Hz}\)), 40.8. HRMS calc'd for C\textsubscript{16}H\textsubscript{13}FNS [M+H]\textsuperscript{+}: 270.0747, found: 270.0742.

**6-[(4-(Trifluoromethyl)phenyl)sulfanyl]methylquinoline (3f):** Time: 15 hours. Pale yellow solid. Purification: 5 % EtOAc in DCM. Yield: 114 mg, 71 %. \(^1\)H NMR (DMSO-d\textsubscript{6}) \(\delta\) 8.85 (s, 1H), 8.29 (d, \(J = 7.8 \text{ Hz}, 1\text{H}\)), 7.96 (d, \(J = 7.8 \text{ Hz}, 2\text{H}\)), 7.80 (d, \(J = 8.7 \text{ Hz}, 1\text{H}\)), 7.64 – 7.46 (m, 6H), 4.56 (s, 2H). \(^{13}\)C NMR (DMSO-d\textsubscript{6}) \(\delta\) 151.2, 147.7, 142.9, 136.4, 135.8, 131.3, 129.9, 128.3, 128.2, 128.0, 126.6, 126.3 (q, C-C-CF\textsubscript{3}, \(J = 3.9 \text{ Hz}\)), 124.9 (q, CF\textsubscript{3}, \(J = 270.8 \text{ Hz}\)), 122.5, 36.0. HRMS calc'd for C\textsubscript{17}H\textsubscript{13}F\textsubscript{3}NS [M+H]\textsuperscript{+}: 320.0715, found: 320.0711.

**6-[(4-Nitrophenyl)sulfanyl]methylquinoline (3g):** Time: 16 hours. Yellow-orange solid. Purification: 0-5 % EtOAc in DCM. Yield: 143 mg, 96 %. \(^1\)H NMR (DMSO-d\textsubscript{6}) \(\delta\) 8.85 (dd, \(J = 1.6 \text{ Hz}, 4.1 \text{ Hz}, 1\text{H}\)), 8.20 (d, \(J = 7.8 \text{ Hz}, 1\text{H}\)), 8.09 (d, \(J = 8.8 \text{ Hz}, 2\text{H}\)), 8.01 (d, \(J = 1.0 \text{ Hz}, 1\text{H}\)), 7.98 (d, \(J = 8.8 \text{ Hz}, 1\text{H}\)), 7.82 (dd, \(J = 1.9 \text{ Hz}, 8.8 \text{ Hz}, 1\text{H}\)), 7.57 (d, \(J = 8.8 \text{ Hz}, 1\text{H}\)), 7.50 (t, \(J = 8.8 \text{ Hz}, 1\text{H}\)), 7.33 (d, \(J = 8.4 \text{ Hz}, 2\text{H}\)), 6.91 (t, \(J = 8.2 \text{ Hz}, 2\text{H}\)), 4.18 (s, 2H). HRMS calc'd for C\textsubscript{17}H\textsubscript{13}N\textsubscript{2}S [M+H]\textsuperscript{+}: 277.0794, found: 277.0787.
= 9.0 Hz, 2H), 7.50 (m, 1H) 4.63 (s, 2H). $^{13}$C NMR (DMSO-d$_6$) δ 151.3, 147.7, 147.5, 145.3, 136.5, 135.3, 131.2, 130.0, 128.3, 128.3, 127.3, 124.6, 122.5, 35.7. HRMS calc'd for C$_{16}$H$_{13}$N$_2$O$_2$S $[M+H]^+$: 297.0692, found: 297.0686.

6-{{[(4-Methoxyphenyl)sulfanyl]methyl}quinoline (3h): Time: 15 hours. Purification: 30 % EtOAc in hexanes. Yield: 75 mg, 53 %. $^1$H NMR (DMSO-d$_6$) δ 8.83 (dd, $J = 4.1$ Hz, 1.7 Hz, 1H), 8.23 (d, $J = 7.4$ Hz, 1H), 7.92 (d, $J = 8.8$ Hz, 1H); 7.71 (s, 1H), 7.69 (dd, $J = 8.6$, 2.0 Hz, 1H), 7.47 (dd, $J = 8.4$ Hz, 4.2 Hz, 1H), 7.26 (d, $J = 8.8$ Hz, 2H); 6.82 (d, 8.8 Hz, 2H); 4.28 (s, 2H); 3.58 (s, 3H). $^{13}$C NMR (CDCl$_3$) δ 159.7, 150.4, 147.7, 136.9, 136.1, 134.8, 131.1, 129.8, 128.2, 127.3, 125.5, 121.5, 114.7, 55.5, 41.6. HRMS calc'd for C$_{17}$H$_{16}$NOS $[M+H]^+$: 282.0947, found: 282.0941.

$N,N$-Dimethyl-4-{{[(quinolin-6-ylmethyl)sulfanyl]aniline (3i): Time: 15 hours. Yellow solid. Purification: 20 % EtOAc in hexanes. Yield: 103 mg, 68 %. $^1$H NMR (DMSO-d$_6$) δ 8.83 (s, 1H), 8.21 (d, $J = 7.8$ Hz, 1H), 7.91 (d, $J = 9.3$ Hz, 1H), 7.66 (d, $J = 7.2$ Hz, 2H), 7.55 – 7.40 (m, 1H), 7.13 (d, $J = 8.8$ Hz, 2H), 6.57 (d, $J = 8.8$ Hz, 2H), 4.05 (s, 2H), 2.82 (s, 6H). $^{13}$C NMR (DMSO-d$_6$) δ 150.9, 150.5, 147.5, 137.5, 136.3, 134.3, 131.5, 129.5, 128.2, 127.8, 122.3, 119.7, 113.3, 100.5. HRMS calc'd for C$_{18}$H$_{19}$N$_2$S $[M+H]^+$: 295.1263, found: 295.1257.

6-{{[(4-Chlorophenyl)sulfanyl]methyl}quinoline (3j): Time: 15 hours. Yellow solid. Purification: 5 % EtOAc in DCM. Yield: 99 mg, 67 %. $^1$H NMR (DMSO-d$_6$) δ 8.84 (d, $J = 2.5$ Hz, 1H), 8.27 (d, $J = 8.4$ Hz, 1H), 7.90 (d, $J = 8.7$ Hz, 1H), 7.90 (s, 1H), 7.80 (s,
6-[(2-Methylphenyl)sulfanyl]methyl]quinoline (3k): Time: 15 hours. Pale yellow solid. Purification: 0-5 % EtOAc in DCM. Yield: 99 mg, 74 %. \(^1\)H NMR (DMSO-\(d_6\)) \(\delta\) 8.84 (d, \(J = 4.1\) Hz, 1H), 8.26 (d, \(J = 8.3\) Hz, 1H), 7.95 (d, \(J = 8.6\) Hz, 1H), 7.85 (s, 1H), 7.79 (d, 1H), 7.54 – 7.42 (m, 1H), 7.34 (d, \(J = 7.6\) Hz, 1H), 7.22 – 7.00 (m, 3H), 4.44 (s, 2H), 2.28 (s, 3H). \(^{13}\)C NMR (DMSO-\(d_6\)) \(\delta\) 151.1, 147.6, 137.2, 136.4, 136.4, 135.7, 131.4, 130.7, 129.8, 128.7, 128.3, 128.0, 127.2, 126.5, 122.4, 37.0, 20.5. HRMS calc'd for \(\text{C}_{17}\text{H}_{16}\text{NS}\) [M+H]^+: 266.0998, found: 266.0994.

6-[(2,6-Dimethylphenyl)sulfanyl]methyl]quinoline (3l): Time: 24 hours. Yellow-orange solid. Purification: 0-3 % EtOAc in DCM. Yield: 84 mg, 60 %. \(^1\)H NMR (CDCl\(_3\)) \(\delta\) 8.86 (d, \(J = 3.0\) Hz, 1H), 7.98 (t, \(J = 8.8\) Hz, 2H), 7.54 (dd, \(J = 8.6\) Hz, 1.8 Hz, 1H), 7.35 (dd, \(J = 8.3\) Hz, 4.2 Hz, 1H), 7.29 (s, 1H), 7.11 (t, \(J = 7.8\) Hz, 1H), 7.04 (d, \(J = 7.3\) Hz, 2H), 3.95 (s, 2H), 2.35 (s, 6H). \(^{13}\)C NMR (CDCl\(_3\)) \(\delta\) 150.4, 127.7, 143.8, 137.0, 135.9, 132.7, 131.2, 129.8, 128.8, 128.3, 128.2, 127.1, 121.5, 40.0, 22.1. HRMS calc'd for \(\text{C}_{18}\text{H}_{18}\text{NS}\) [M+H]^+: 280.1154, found: 280.1149.

6-[(Naphthalen-2-ylsulfanyl)methyl]quinoline (3m): Time: 16 hours. Yellow solid. Purification: 0-7 % EtOAc/DCM. Yield: 145 mg, 96 %. \(^1\)H NMR (DMSO-\(d_6\)) \(\delta\) 8.83 (dd, \(J = 1.7\) Hz, 4.1 Hz, 1H), 8.25 (dd, \(J = 1.2\) Hz, 8.3 Hz, 1H) 7.94 (m, 2H), 7.87 (d, \(J = 1.4\) Hz, 7.55 – 7.44 (m, 1H), 7.41 – 7.26 (m, 4H), 4.44 (s, 2H). \(^{13}\)C NMR (DMSO-\(d_6\)) \(\delta\) 151.1, 147.6, 136.4, 136.3, 135.4, 131.4, 131.3, 131.0, 129.8, 129.6, 128.3, 128.0, 122.4, 37.3. HRMS calc'd for \(\text{C}_{16}\text{H}_{13}\text{ClNS}\) [M+H]^+: 286.0452, found: 286.0447.
Hz, 1H), 8.80 (m, 4H), 7.44 (m, 4H), 4.55 (s, 2H). $^{13}$C (DMSO-$d_6$) $\delta$ 151.1, 147.6, 136.5, 136.4, 134.1, 134.0, 131.9, 131.4, 129.8, 129.0, 128.3, 128.1, 127.6, 127.4, 126.7, 126.4, 122.4, 37.1. HRMS calc'd for C$_{20}$H$_{16}$NS [M+H]$^+$: 302.0998, found: 302.0992.

6-[(Quinolin-6-ylmethyl)sulfanyl]quinoline (3n): Time: 15 hours. Pale yellow solid. Purification: 5 % MeOH in EtOAc. Yield: 129 mg, 82 %. $^1$H NMR (DMSO-$d_6$) $\delta$ 8.81 (dd, $J$ = 2.7, 16.9 Hz, 2H), 8.24 (dd, $J$ = 8.1, 18.8 Hz, 2H), 7.99 – 7.87 (m, 4H), 7.82 (d, $J$ = 8.7 Hz, 1H), 7.72 (d, $J$ = 6.9 Hz, 1H), 7.54 – 7.41 (m, 2H), 4.59 (s, 2H). $^{13}$C NMR (DMSO-$d_6$) $\delta$ 151.1, 150.8, 147.6, 146.8, 146.8, 136.4, 136.2, 135.7, 135.2, 131.4, 130.7, 130.0, 130.0, 129.0, 128.3, 128.2, 126.3, 122.7, 122.4, 37.0. HRMS calc'd for C$_{19}$H$_{15}$N$_2$S [M+H]$^+$: 303.095, found: 303.0942.

Methyl 2-[(quinolin-6-ylmethyl)sulfanyl]benzoate (3o): Time: 15 hours. Pale yellow solid. Purification: 30 % EtOAc in hexanes. Yield: 102 mg, 63 %. $^1$H NMR (DMSO-$d_6$) $\delta$ 8.86 (s, 1H), 8.31 (d, $J$ = 8.1 Hz, 1H), 8.05 – 7.91 (m, 2H), 7.91 – 7.78 (m, 2H), 7.63 – 7.44 (m, 3H), 7.22 (t, $J$ = 7.4 Hz, 1H), 4.44 (s, 2H), 3.79 (s, 3H). $^{13}$C NMR (DMSO-$d_6$) $\delta$ 166.8, 151.2, 147.7, 141.2, 136.4, 135.7, 133.4, 131.6, 131.4, 129.8, 128.4, 128.4, 127.8, 127.1, 125.1, 122.4, 52.8, 36.3. HRMS calc'd for C$_{18}$H$_{16}$NO$_2$S [M+H]$^+$: 310.0896, found: 310.0891.

6-[(Phenylsulfanyl)methyl]quinoline (3a): Prepared from iodobenzene. Time: 72 hours. Pale yellow solid. Purification: 5 % EtOAc in DCM. Yield: 78 mg, 61 %. $^1$H NMR (DMSO-$d_6$) $\delta$ 8.84 (d, $J$ = 2.7 Hz, 1H), 8.26 (d, $J$ = 8.1 Hz, 1H), 7.94 (d, $J$ = 8.7 Hz, 1H),
7.86 (s, 1H), 7.76 (d, J = 8.7 Hz, 1H), 7.54 – 7.42 (m, 1H), 7.34 (d, J = 7.8 Hz, 2H), 7.26 (t, J = 7.7 Hz, 2H), 7.15 (t, J = 7.3 Hz, 1H), 4.42 (s, 2H). 13C NMR (DMSO-d6) δ 151.1, 147.6, 136.6, 136.4, 136.3, 131.4, 129.8, 129.7, 129.3, 128.3, 128.0, 126.7, 122.4, 37.3.

HRMS calc’d for C_{16}H_{14}NS [M+H]^+: 252.0841, found: 252.0836.

6-[(Phenylsulfanyl)methyl]quinoline (3a): Prepared from chlorobenzene. Time: 120 hours. Pale yellow solid. Purification: 5 % EtOAc in DCM. Yield: 13 mg, 10 %. 1H NMR (CDCl₃) δ 8.88 (s, 1H), 8.07 (d, J = 8.5 Hz, 2H), 7.70 (d, J = 8.4 Hz, 1H), 7.66 (s, 1H), 7.44 – 7.35 (m, 1H), 7.35 – 7.14 (m, 5H), 4.27 (s, 2H). 13C NMR (CDCl₃) δ 150.2, 147.4, 136.4, 135.8, 131.1, 130.7, 129.7, 129.2, 128.3, 127.3, 127.0, 122.0, 39.5. C_{16}H_{14}NS [M+H]^+: 252.0841, found: 252.0836.

Methyl 4-[(4-methoxybenzyl)sulfanyl]benzoate (3p): Time: 15 hours. Pale yellow solid. Purification: 0-8 % EtOAc in hexanes. Yield: 130 mg, 81 %. 1H NMR (DMSO-d₆) δ 7.82 (d, J = 8.4 Hz, 2H), 7.37 (d, J = 8.4 Hz, 2H), 7.31 (d, J = 8.5 Hz, 2H), 6.85 (d, J = 8.6 Hz, 2H), 4.28 (s, 2H), 3.82 (d, J = 15.9 Hz, 3H), 3.70 (s, 3H). 13C NMR (DMSO-d₆) δ 166.5, 159.1, 144.7, 130.8, 130.2, 128.9, 127.1, 126.7, 114.6, 55.7, 52.7, 35.4.

HRMS calc’d for C_{16}H_{16}O₃S [M]^+: 288.0820, found: none.

Methyl 4-(benzylsulfanyl)benzoate (3q): Time: 15 hours. Pale yellow solid.

Purification: 0-3 % EtOAc in hexanes. Yield: 102 mg, 78 %. 1H NMR (CDCl₃) δ 7.91 (d, J = 8.4 Hz, 2H), 7.41 – 7.19 (m, 8H), 4.21 (s, 2H), 3.89 (s, 3H). 13C NMR (DMSO-d₆) δ
Methyl 4-[(4-(trifluoromethyl)benzyl)sulfanyl]benzoate (3r): Time: 15 hours. Yellow solid. Purification: 0-8 % EtOAc in hexanes. Yield: 90 mg, 55 %. $^1$H NMR (CDCl$_3$) $\delta$ 7.91 (d, $J = 8.1$ Hz, 2H), 7.55 (d, $J = 7.8$ Hz, 2H), 7.44 (d, $J = 7.8$ Hz, 2H), 7.28 (d, $J = 8.3$ Hz, 2H), 4.21 (s, 2H), 3.88 (s, 3H). $^{13}$C NMR (CDCl$_3$) $\delta$ 166.8, 142.9, 141.0, 130.3, 129.9 (q, C- CF$_3$, $J = 32.5$ Hz), 129.3, 127.8, 127.6, 125.8 (q, C-C-F$_3$, $J = 3.7$ Hz), 124.3 (q, CF$_3$, $J = 270.1$ Hz), 52.3, 37.2. HRMS calc'd for C$_{15}$H$_{13}$O$_2$S [M]$^+$: 326.0588, found: 326.0591.

Methyl 4-[(4-chlorobenzyl)sulfanyl]benzoate (3s): Time: 15 hours. Pale yellow solid. Purification: 0-5 % EtOAc in hexanes. Yield: 91 mg, 62 %. $^1$H NMR (DMSO-$d_6$) $\delta$ 7.82 (d, $J = 8.5$ Hz, 2H), 7.48 (dd, $J = 8.5$ Hz, 33.6 Hz, 6H), 4.35 (s, 2H), 3.80 (s, 3H). $^{13}$C NMR (DMSO-$d_6$) $\delta$ 166.5, 143.9, 136.7, 132.5, 131.4, 130.2, 129.1, 127.3, 127.0, 52.8, 35.0. HRMS calc'd for C$_{15}$H$_{13}$ClO$_2$S [M]$^+$: 292.0325, found: 292.0326.

Methyl 4-[(1-phenylethyl)sulfanyl]benzoate (3t): Time: 15 hours. Orange solid. Purification: 0-3 % EtOAc in hexanes. Yield: 101 mg, 73 %. $^1$H NMR (DMSO-$d_6$) $\delta$ 7.79 (d, $J = 8.5$ Hz, 2H), 7.49 – 7.35 (m, 4H), 7.29 (t, $J = 7.6$ Hz, 2H), 7.24 (t, $J = 7.3$ Hz, 1H), 4.81 (q, $J = 6.9$ Hz, 1H), 3.79 (s, 3H), 1.57 (d, $J = 7.0$ Hz, 3H). $^{13}$C NMR (DMSO-$d_6$) $\delta$ 166.5, 143.5, 143.1, 130.2, 129.2, 128.9, 128.0, 127.9, 127.3, 52.8, 45.1, 23.0. HRMS calc'd for C$_{16}$H$_{16}$O$_2$S [M]$^+$: 272.0871, found: 272.0877.
General procedure for the one-pot formation of alkyl-aryl thioethers (Table 1.4): To an oven-dried microwave vial with a stir bar was added benzyl bromide substrate 9 (0.80 mmol) and THF (1.0 mL). To this solution was added potassium thioacetate (101 mg, 0.880 mmol) and the mixture was vigorously stirred at 60°C. The reaction was monitored by TLC and/or LC/MS and upon completion was cooled to room temperature. XPhos (38.1 mg, 0.080 mmol) was added to the vial, which was evacuated and purged with nitrogen. To a separate oven-dried microwave vial with a stir bar was added Pd$_2$dba$_3$ (36.6 mg, 0.040 mmol) and K$_2$CO$_3$ (111 mg, 0.800 mmol). Degassed THF (1.00 mL), aryl halide substrate 2 (1.200 mmol), and H$_2$O (0.50 mL) were then added to the microwave vial. This mixture was stirred at room temperature for 15 minutes, and then added to the first microwave vial via syringe. This reaction was heated, vigorously stirred at 100°C, monitored by TLC and/or LC/MS, and upon completion was cooled to room temperature. The cooled reaction mixture was diluted with EtOAc (2 mL), filtered through a pad of Celite, and rinsed with EtOAc (100 mL). The filtrate was concentrated in vacuo and purified by flash chromatography over silica gel to afford thioethers 3.

Methyl 2-(benzylsulfanyl)benzoate (3u): Time: 3 hours followed by 24 hours. Orange solid. Purification: 0-3 % EtOAc in hexanes. Yield: 205 mg, 96 %. $^1$H NMR (DMSO-d$_6$) δ 7.85 (d, J = 7.5 Hz, 1H), 7.51 (d, J = 2.9 Hz, 2H), 7.41 (d, J = 7.3 Hz, 2H), 7.32 (t, J = 7.5 Hz, 2H), 7.28 – 7.18 (m, 2H), 4.22 (s, 2H), 3.79 (s, 3H). $^{13}$C NMR (CDCl$_3$) δ 167.1, 142.2, 136.3, 132.6, 131.4, 129.3, 128.8, 127.6, 126.1, 124.3, 105.0, 52.4, 37.5. HRMS calc'd for C$_{15}$H$_{14}$O$_2$S [M$^+$]: 258.0715, found: 258.0703.
Methyl 4-[(4-(trifluoromethyl)benzyl)sulfanyl]benzoate (3q): Time: 2 hours followed by 25 hours. White solid. Purification: 0-8 % EtOAc in hexanes. Yield: 115 mg, 44 %. $^1$H NMR (CDCl$_3$) δ 7.91 (d, $J = 8.3$ Hz, 2H), 7.55 (d, $J = 8.1$ Hz, 2H), 7.45 (d, $J = 7.8$ Hz, 2H), 7.28 (d, $J = 8.3$ Hz, 2H), 4.21 (s, 2H), 3.88 (s, 3H). $^{13}$C NMR (CDCl$_3$) δ 166.8, 142.9, 141.0, 130.3, 129.9 (q, C-CF$_3$, $J = 32.6$ Hz), 129.3, 127.8, 127.6, 125.8 (q, C-C-CF$_3$, $J = 3.7$ Hz), 124.3 (q, CF$_3$, $J = 271.6$ Hz), 52.3, 37.2. HRMS calc'd for C$_{16}$H$_{13}$F$_3$O$_2$S [M]$^+$: 259.0787, found: 259.0784.

(4-tert-Butylbenzyl)(4-(trifluoromethyl)phenyl)sulfane (3v): Time: 4 hours followed by 22 hours. White solid. Purification: 100 % hexanes. Yield: 168 mg, 65 %. $^1$H NMR (CDCl$_3$) δ 7.49 (d, $J = 8.3$ Hz, 2H), 7.41 – 7.32 (m, 4H), 7.29 (d, $J = 8.3$ Hz, 2H), 4.18 (s, 2H), 1.31 (s, 9H). $^{13}$C NMR (CDCl$_3$) δ 150.8, 142.8, 133.42, 127.8 (q, C-CF$_3$, $J = 32.6$ Hz), 127.7, 125.94, 125.86 (q, C-C-CF$_3$, $J = 3.8$ Hz), 124.5 (q, CF$_3$, $J = 271.5$ Hz), 37.4, 34.8, 31.6. HRMS calc'd for C$_{18}$H$_{19}$F$_3$S [M]$^+$: 324.1160, found 324.1154.
Sample name: T70
Notebook: 0310106; Page: 2170
Data acquired: 03-Nov-2010, 09:23:51
Data file name: Proton1.pdd
Data reprocessed: Mon, Feb 21, 12:14:34 MST 2011
Chemist: wakerk
YMRD-50D Farivar, probe: CEF
Solvent: None
Ambient temperature

**CHEST**
Proton, 400.13 MHz
Relax, delay 2.000 sec
Pulse 45.0 degrees
AEZ, TIME 3.000 sec
WIDE 2132.4 Hz
12 scans
Data Processing
Line broadening 0.2 Hz
**Notes**

Chemist: wakerk; Notebook: 0310106; Page: 2170
Plotted: 02/21 12:14:34
Restricted Confidential, Limited Access
Sample name: 0170
Notebook: 0310996, Page: 0270
Data acquired: 03-Nov-2010, 19:02:03
Data file name: Carbon01.fid
Chemist: weperk
VNAH-510 Carrar, Probe: CEN
Solvent: diacet
Ambient temperature

OBSERVE Carbon, 115.678 MHz
Relax. Delay 1.016 sec
Pulse 9.5° Half
وقع 1.111 Hz
Width 10457.9 Hz
256 scans
DECOUPLE: X1, 499.618 MHz
Power 90 cm
continuously on
X1, w(t), 499.618 MHz

DATA PROCESSING
Line broadening 0.5 Hz
F7 site 131073
Sample name: 9 Carbon
Notebook: 0310906, Pages: 0000
Data acquired: 04-Oct-2010, 22:16:38
Data file name: Carbon09.txt
Data reprocessed: Mon Feb 21 11:45:28 EST 2011
Chemist: wegerk
VEHICLES: 500 farvor, Probe: CNT
Solvent: deso
Ambient temperature

OBSSERVE Carbon: 33.678 MHz
Relax. Delay 1.001 sec
Pulse 45.0 degrees
Acq. time 1.000 sec
Width 10.467 Hz
500 scans
 Wendlandt 1.125 s 1.125 1.125 Hz
Power 45 dB
continuously on
21, worst 75 kHz modulated
DATA PROCESSING
Line broadening 0.5 Hz
FT size 131072
Sample name: 129
Notebook: 0101004, Page: 0109
Data acquired: 01-04-2012, 01:49:30
Data file name: techbook1.fds
Data reprocessed: Mon Feb 21 12:17:21 EST 2011
Chemical: wapex
VBRP=100 Farrar, Probe: CRF
Software: ndsl3
Ambient temperature

OBSERVE Carbon, 125.67 MHz
h.e. delay 1.600 sec
Pulse 45.0 degrees
acq. time 1.300 secw

SOURCE H, 491.614 MHz
Power 45.0
continuously on

DATA PROCESSING
Line broadening 0.0 Hz
TV size 133.972

Chemist: wapex; Notebook: 0101004; Pages: 0109
Requested: 02/21 12:17:02
Restricted Confidential, Limited Access
Sample name: 132
Notebook: 0310996, Page: 0132
Data acquired: 06-Oct-2010, 11:02:39
Data file name: Carbon1.file
Data reprocessed: Sun Nov 21 13:17:40 EST 2011
Chemist: wegehr
VNMRS-510 fararr, Probe: CHF
Solvent: ddeo
Ambient temperature

OBSERVE Carbon, 121.678 MHz
Relax. Delay 1.001 sec
Pulse 45.0 degrees
Acq. time 1.100 sec
Width 20487.8 Hz
5000 scans

DECOPES X1, 499.166 MHz
Power 45 dB
continuously on
X1_freq_spikes modulated
DATA PROCESSING
Line broadening 0.5 Hz
FT size 131072
Sample name: 166 Carbon
Notebook: 031096; Page: 014
Date acquired: 14-Oct-2011, 18:12:43
Data file name: Carbon166.fid
Date reprocessed: Mon Feb 21 12:18:29 EST 2011
Chemist: wepa
VNMRS-501 farwax
Probe: CHEF
Solvent: cdCl3
Ambient temperature

OBSERVE Carbon, 125.728 MHz
Relax. Delay 1.000 sec
Reson. 45.6 degrees
Avg. time 1.000 sec
Width 21.000 Hz
500 scans

DEPRESSOR 51, 499.614 MHz
Power 50 dB
continuously on
FI_wxcomp_7pMHz modulated
data processing
Line broadening 0.5 Hz
FT size 131072

Chemist: wepa
Notebook: 031096; Page: 014
VNMRS-501 farwax
Probe: CHEF
Date reprocessed: Mon Feb 21 12:18:29
Restricted Confidential, Limited Access
Sample name: 143 purity
Notebook: 010001, page: 003
Data acquired: 13-Oct-2010, 14:16:43
Date file name: PFTNOL.fid
Data reprocessed: Mon Feb 22 13:39:38 EST 2010
Chemist: Wagner
Detector: 500 MHz, probe: CF
Solvent: cdcl3
Ambient temperature

CHEMSELECT Proton, 400.021 MHz
Relax: Delay 1.000 sec
Pulse 60.0 degrees
Aver. time 5.000 sec
Aver. 8122.8 Hz
1st trace
Data Processing:
Line broadening 0.2 Hz
PT ratio 0.93

Chemist: Wagner; Notebook: 010001, page: 002
Plotted: 02/11 13:39:39
Restricted Confidential, Limited Access
Sample name: 72 Purity

Data file name: 2001-00051.dsf

Data reprocessed: Mon Feb 21 11:25:25 2000

Chemist: wegerk

Temperature: Ambient temperature

Proton, 400.011 MHz

Pulse: 10.000 sec

Spin lock: Continuous

Frequency: 4022.0 Hz

16 scans

Chemical shift: 4.2 Hz

PT size 65536
Sample name: 72 Purity 3
Notebook: 0310906, Page: 0072
Data acquired: 09-Aug-2012, 13:14:12
Date file name: Carbond03.616
Data reprocessed: Mon Feb 21 12:20:48 EST 2011
Chemist: wamph

**RESOVE** Carbon, 125.678 MHz
Relax. Delay 1.010 sec
Pulse 45.0 degrees
Acq. time 1.130 sec

Width 10487.8 Hz
156 scans

**DECOUPLED** H1, 499.616 MHz
Power 55 dB
continuously on
SI_wust_lystke modulation

Line broadening 1.5 Hz
FT size 131072

---

220 200 180 160 140 120 100 80 60 40 20 0 ppm

Chemist: wamph; Notebook: 0310906; Page: 0072
Plotted: 02/21/12 12:20:30

Restricted Confidential, Limited Access
Sample name: wepr-0110-3b-cdo13
Notebook: 0316439; Page: 5041
Data acquired: 2-Feb-2011, 20:19:04
Date reprocessed: Tue Feb 22 10:24:13 MST 2011
Chemist: danimet
VNMRS-500 Farrar, Probe: CHF
Solvent: cdcl3
Ambient temperature

DECREASE Carbon, 125.678 MHz
Relax. delay 1.400 sec
Pulse 45.0 degrees
Acq. time 1.300 sec
Width 10.487 Hz
2000 scans

DECREASE H1, 599.814 MHz
Power 45.00
continuously on
H1, wurel_7pbmbs modulated
DATA PROCESSES
Line broadening 0.5 Hz
FT size 151072
Sample name: 174
Molecule: unknown, page: 0174
Data acquired: 23-Nov-2010, 23:18:18
Data file name: Carbon13.016
Data reprocessed: Mon Feb 21 13:04:09 EST 2011
Chemist: vegar
VMODE: 500 fararr, Probes: CHF
Solvent: cdc113
Ambient temperature

genera Carbon, 125.678 MHz
Relax delay 1.000 sec
Pulse 90.0 degrees
Avg. time 1.000 sec
Width 30447.5 Hz
128 scans
DECoupling: 1H, 493.814 MHz
Power 45 dB
continuously on
1H wmrsc.7p5kHz modulated
Data Processing
line broadening 0.5 Hz
PT size 131072
Sample name: N4 Poly
Notebook: 0320308, Page: 0116
Data acquired: 29-Mar-2013, 13:40:42
Data file name: Proto055-ED
Data reprocessed: Mon Feb 25 13:24:44 EST 2013
Chemist: Weigarek
Instrument: Avance 400, Probe: TOP
Solvent: DMSO
Ambient temperature

**DESBTPE Proton, 400.13 MHz**
Decay: Delay 1.010 sec
Pulse 45.0 degrees
AQ: 4169.3 1.000 H
c/8.000 Hz
16 scans
DATA PROCEDURE
Line broadening 1.0 Hz
NC also 65536
Sample name: 116
Notebook: 03109046, Page: 0116
Data acquired: 27-Oct-2010, 19:02:50
Data file name: Carbon1of6d
Data reprocessed: Mon Feb 21 12:29:17 EST 2011
Chemist: wepark
VNMRS-500 fararr, Probe: CHEF
Solvent: DMSO
Ambient temperature

CARBON Carbon, 129.678 MHz
Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.333 sec
Width 20487.8 Hz
4000 scans
DECOUPLE sp, 499.916 MHz
Power 45 dB
continuously on
2L wurt_7ppm modulated
data processing
Line broadening 0.5 Hz
PT size 131072

Chemist: wepark; Notebook: 03109046; Page: 0116
Plotted: 02/21 12:29:17
Restricted Confidential, Limited Access
Sample name: 115
Notebook: 031590, Page: 0555
Data acquired: 21-Oct-2012, 18:02:47
Data file name: Cachon115_116
Data reprocessed: Mon Feb 21 12:30:01 EST 2011

Chemical: waver

VNMRS-500 farrrr, Probe: CHEF
Solvent: DMSO

Ambient temperature

Observation: 126.070 MHz
Relax. delay 1.500 sec
Pulse 45.0 degrees
Acq. time 1.500 sec
Width 30485.6 Hz
500 scans

DECOUPLING: R1, 499.0 Hz
Power 45 dB
Continuously on
R1 wave. Unshifted

Data processing
Line broadening 5.5 Hz
FT size 131072

220 200 180 160 140 120 100 80 60 40 20 0 ppm

Chemical: waveker
Notebook: 0311004, Page: 0652
Plotted: 03/31 10:36:01
Restricted Confidential, Limited Access
Sample name: 115 Purity
Notebook: 0319064, Page: 5113
Data acquired: 24-Sep-2011, 12:45:35
Data file name: Proton01.fid
Data reprocessed: Mon Feb 21 12:27:20 EST 2011
Chemist: sayark
VNMRS-500 Farrer 
Probe: CHEF
Solvent: dmeo
Ambient temperature

OBSERVE Proton, 499.813 MHz
Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 3.000 sec
Width 8012.0 Hz
16 scans

DATA PROCESSING
Line broadening 3.2 Hz
FT size 65536
Sample name: final
Molecule: 0312355, Page: 0102
Data acquired: 11-Nov-2010, 09:29:54
Data file name: proton1212.fid
Data reprocessed: 11-Feb-2013 09:12:21
Chemist: danriet
VNR H-500 farizar, Probe: CHP
Solvent: DMSO
Ambient temperature

OBSERVE Proton, 400.13 MHz
Slice. Delay 1.000 sec
Pulse 90.0 degrees
Amp. time 3.000 sec
Width 0.022.8 Hz
11 scans
Data processed
Line broadening 0.2 Hz
FT size 65536

Chemist: danriet; Notebook: 0312355; Page: 0112
Plotted: 02/21 13:29:07
Restricted Confidential, Limited Access
Sample name: 113 Purity
Notebook: 0310906, Page: 0112
Data acquired: 14-Oct-2010, 15:35:54
Data file name: protocola.fid
Data reprocessed: Mon Feb 21 12:30:33 BST 2011
Chemist: vapark
VENDOR: 300 farzar . Probe: CHF
Solvent: dmso
Ambient temperature
OBSERVE Proton, 499.63 MHz
Relax. Delay 1.000 sec
Pulse 49.0 degrees
Avg. time 5.500 sec
Width 8033.8 Hz
16 scans
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536

Chemist: vapark; Notebook: 0310906; Page: 0112
Plotted: 02/21 11:30:33
Restricted Confidential, Limited Access
Sample name: 140 Puriss
Notebook: 031000; Page: 0140
Data acquired: 12-Oct-2012, 09:38:06
Data file name: Protocol1.fid
Data reprocessed: Mon Feb 01 10:31:37 2013
Chemist: wagar

PROTONES-301 farrar, Probe: CHF
Solenet: dmsi
Ambient temperature

INSTRUMENT Proton, 400.13 MHz
Relax. delay 1.000 sec
Puls 45.0 degrees
Acq. Line 3.000 sec
Width 6115.8 Hz
16 scans
DATA PROCESSING
Line broadening 5.0 Hz
FT size 65536

Chemist: wagar; Notebook: 031000; Page: 0140
Plotted: 02/21 12:31:38
Restricted Confidential, Limited Access
Sample name: 140Carbon
Notebook: 031996, Page: 0140
Data acquired: 12-Oct-2010, 22:16:40
Data file name: Carbon1.5ks
Data reprocessed: Mon Feb 21 12:31:24 EST 2011
Chemist: wagsky
NMR-300 farvar, Probe: CIK
Solvent: DMSO
Ambient temperature

CHEMISTRY Carbon, 115.678 MHz
Relax. Delay 1.010 sec
Pulse 15.0 degrees
Acq. time 1.355 sec
Width 10487.8 Hz
256 scans
Resolution 64, 0.010 ppm
Power 7 dB
continuously on
ML, mult, 7p58khz modulated
data resolution
Line broadening 1.5 Hz
FT size 131072

Chemist: wagsky, Notebook: 031996, Page: 0140
Plotted: 02/21 12:31:25
Restricted Confidential, Limited Access
Sample name: 177 Purity
Notebook: 0310096, Page: 017
Data acquired: 29-Aug-2010, 13:05:03
Data file name: Proton02.fid
Data reprocessed: Mon Feb 21 13:32:20 2011
Chemist: vaperk
Magnetic field. Probes: CHEF
Solvent: ddeo
Ambient temperature

OBSERVE proton, 400.13 MHz
magnet. delay 1.964 sec
Pulse 45.0 degrees
Acq. time 3.000 sec
Mixt 102.0 Hz
16 scans
 Nicolet Dimension
Line broadening 0.2 Hz
FT size 65536

Chemist: vaperk; Notebook: 0310096; Page: 017
Plotted: 21/21 13:32:10
Restricted Confidential, Limited Access

71
Observe Carbon, 125.678 MHz
Relax. delay 1.000 sec
Pulse 45.0 degrees
Avg. time 8.988 sec
Width 40467.8 Hz
5000 scans
Decoupler XH, 499.816 MHz
Power 45 dB
continuously on
Hi.square, 3004A modulated
data sweep curves
Line broadening 0.5 Hz
PT size 113072
Sample name: 137 Purity
Notes: O31006, Page: 0137
Date acquired: 11-Oct-2008, 09:34:41
Data file name: Proton01.fid
Data reprocessed: Mon Feb 21 12:39:50 EST 2011
Chemist: vgsrk
VEHICLE: d2o, Probe: CHF
Solvent: d2o
Ambient temperature

OBSERVE Proton, 400.13 MHz
Relax. Delay 1.000 sec
Pulse 45.0 degrees
RADC. 6544.000 MHz
Width 801.2 Hz
12 scans
NMR PROCESSING
Line broadening 0.2 Hz
FT size 65536
Sample name: 137 Carbon
Notebook: 0310096; Pages: 0137
Date acquired: 12-Oct-2009, 19:02:38
Data file name: Carbon13.F16
Instrument: NMR-2000
Spectrometer: CFT
Solvent: DMO
Ambient temperature

CHEMICAL SHIFT, 131.678 MHz
Decay Delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.300 sec
WATER 100.6 Hz
5000 scans
Sweep rate 61 KHz
Power 45 dB
Continuously on
HI-west T200Hs modulated
Line broadening 0.5 Hz
FT size 131072

---

120 100 180 160 140 120 100 80 60 40 20 0 ppm

Chemist: wagon; Notebook: 0310096; Pages: 0137
Plotted: 02/21 12:34:02
Restricted Confidential, Limited Access
Sample name: nager-0110-31-cdd13
Notebook: unknown; Page: 0041
Data acquired: 31-Feb-2011, 23:38:57
Data file name: Protool3.r16
Data reprocessed: Thu Feb 24 10:26:26 WET 2011
Chemist: daniemart
VENDOR-005 factor; Probe: CWV
Solvent: cdd13
Ambient temperature

SPECTRUM Proton, 499.811 MHz
Spectrum delay 1.000 sec
Pulse 45.0 degrees
Acq. time 3.000 sec
Winch 0.033 Hz
16 scans
DATA PROCESSING
Line broadening 0.2 Hz
FO size 65536
Sample name: final
Notebook: 1993655, Pages: 0012
Date acquired: 11-Nov-2010, 09:43:25
Data file name: proton300D.M
Data resprocessed: Mon Feb 21 12:35:47 2011
Chemist: danemat
VENDOR: 500 (avex, Probe: CNW)
Solvent: dcmo
Ambient temperature

CHEMWEB Proton, 400.133 MHz
Scale: 1.000 ppm
Pulse 65.1 degrees
Hex. time 1.558 sec
VUNS 0.011 Hz
16 scans
DATA PROCESSING
Line broadening 0.1 Hz
FT size 65536
Sample name: final
Notebook: 0313355, Page: 0113
Data acquired: 11-Nov-2011, 19:02:07
Data file name: Carbono2.efd
data reprocessed: Mon Feb 21 13:15:54 EST 2011
Chemist: daniomat
VNMRS-500 server, Probe: CFN
Solvent: dmac
Ambient temperature

OBSERVE Carbon, 125.678 MHz
Relax. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.380 sec
Width 32417.0 Hz
3000 scans
DECOUPLING: Hz, 499.816 MHz
Power 45 dB
continuously on
Hz, ncut 50kHz modulated
DATA PROCESSING
Lines broadening 0.5 Hz
FT size 512072

220 200 180 160 140 120 100 80 60 40 20 0 ppm

Chemist: daniomat; Notebook: 0313355; Page: 0113
Plotted: 02/22 12:15:55
Restricted Confidential, Limited Access
Sample name: 156
Notebook: 031096; Page: 0136
Data acquired: 1-Oct-2010, 16:05:46
Data file name: proton02.fid
Data reprocessed: Mon Feb 21 12:41:27 BST 2011
Chemist: vegar
VNMRS-100 farrr , Probe: GHP
Solvent: dmso
Ambient temperature

OBSERVE Proton, 499.813 MHz
Relax. Delay 1.000 sec
Pulse 45.0 degrees
Avg. time 3.000 sec
Width 8012.8 Hz
32 scans
DATA PROCESSING
Line broadening 0.2 Hz
FT size 65536
Sample name: 151
Protocol: 0319006, Page: 038
Date acquired: 16-Oct-2011, 12:02:33
Date processed: Mon Feb 21 11:49:15 EST 2011
Chemist: egarsk
PREP3000 Spectra, Probe: CHEF
diluent: DMSO
Ambient temperature

OCVENUE Proton, 400.13 MHz
Pulse: 1.000 sec
Pulse width: 90.0 degrees
Acq. time: 3.001 sec
Sweep width: 8 MHz
32 scans
data processed
Line broadening: 0.2 Hz
PF size 65536
Sample name: 1171
Notebook: 0313894, Page: 0171
Date acquired: 03-Nov-2010, 21:50:18
Data file name: Carbon11.fid
Data reprocessed: Mon Feb 21 12:49:03 2011
Chemist: xwark
Machine: 500 farcr, Probe: CNF
Solute: cdcl3
Ambient temperature

CUMEN 1 HNMR 125.68 MHz
Balle. delay 1.000 sec
Pulse 45.0 degrees
Acq. time 1.100 sec
Average 10.000 sec
4000 scans
DECOUPLING Off, 149.944 MHz
Power 45 dB
computer on
8,1,3 meth, 7ethyls undinated
DATA PROCESSES
Line broadening 0.5 Hz
FF size 151372
Sample name: 143 Purity
Notebook: 0310304; Page: 0149
Data reprocessed: Mon Feb 21 12:16:58 EST 2011
Chemist: vexpert
$\text{H},\text{NMR}$ 500 Faxvar, Proton: $\text{CHF}$
Solvent: dcll
Ambient temperature

**Observations**
- Proton: 491.81 ppm
- Relax. delay: 1.001 sec
- Pulse: 45.0 degrees
- Acq. time: 3.000 sec
- Width: 8012.8 Hz
- 14 scans

**Data Processing**
- Line broadening: 0.2 Hz
- FT axes: 41614

Chemist: vexpert; Notebook: 0310304; Page: 0149
Plotted: 02/21 12:50:58
Restricted Confidential, Limited Access
Sample name: 148 Carbon
Notebook: 0310606, Page: 0348
Data acquired: 10-Oct-2010, 11:02:07
Data file name: carbon11.fld
Data reprocessed: Mon Feb 21 12:53:12 2011
Chemist: wepkr
VND50 1ppm inner, Probes: CRF

Acquire: 1sec
Ambient temperature

Operate Carbon, 120.676 MHz
Relax. delay 1.000 sec
Pulse 45.1 degrees
acq. time 1.300 sec
Width 3047.8 Hz
5000 scans
recorced 1, 493.910 ppm
Power 45 dB
continuously on
si. wets, pd1line modulated
DATA PROCESSING
Line broadening 0.5 Hz
pt wide 11.0772
Sample name: 349
Notebook: C510904, Page: 0149
Data acquired: 19-oct-0913, 13:48:39
Data file name: Proton22.fid
Data reprocessed: Mon Feb 21 12:43:50 EST 2011
Chemist: wagar
VENUS-500 farver, Probe: CHF
Solvent: DMSO
Ambient temperature

OBSERVE Proton, 499.813 MHz
Delay: delay 1.000 sec
Pulse 90.0 degrees
Acq. time 0.000 sec
NEXQ 0012.8 Hz
32 scans

DATA PROCESSING
Line broadening 2.2 Hz
RT size 0.1 Hz

Chemist: wagar; Notebook: C510904; Page: 0149
Plotted: 10/21 11:53:50
Restricted Confidential, Limited Access

91
Sample name: 181
Notebook: 03138945; Page: 0161
Data acquired: 12-Nov-2010, 18:03:15
Data file name: Carbon11.116
Data reprocessed: Mon Feb 21 12:36:50 EST 2011
Chemist: vaster
vnmr.900 Varian, Proton: CWF
Solvent: cdcl3
Ambient temperature

COSY Chart, 125.674 MHz
Relax. Delay 1.000 sec
Puls 45.0 degree
Acq. time 1.100 sec
FSBCH 35487.3 Hz
500 scans

DECOUPL, 67.011 MHz
Power 45 dB
continuously on
H1, 32KHz modulated

DATA PROCESSING
line broadening 0.5 Hz
FF size 131072
Sample name: 188
Notebook: 03122016. Pages: 018
Date acquired: 10-Dec-2016, 13:03:13
Date file name: Carbon116n.d
Date reprocessed: Mon Feb 21 12:55:24 NPT 2011
Chemist: wagem
Version: 9.2.0 factor, frun: ump
Solvent: dcm3
Ambient temperature

unsaturation 140.678 amu
Relax. Delay 1.000 sec
times 46.0 degrees
Acq. time 1.500 sec
Width 30687.8 Hz
S010 scans
DISORDER 81.649.814 MHz
Power 45 dB
continuously on
m1-meet, yrophs modulated
DATA PROCESSING
Line broadening 2.5 Hz
FT size 128072
<table>
<thead>
<tr>
<th>ppm</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Approximately 25.1 mg of 2a

Product masses from mass spectrometry.

![](image)

**Compound:** N,N-dimethyl-2,1-benzothiazolidine-3,4-dione

**Structure:**

![Chemical Structure](image)

**Synthesis:**

1. **Step 1:**
   - **Reagents:**
     - **N,N-dimethyl-2,1-benzothiazolidine-3,4-dione**
   - **Procedure:**
     - Dissolve the compound in a suitable solvent (e.g., DMF)
     - Add the reagent to the solution
     - Stir at 121°C for 5 h

2. **Step 2:**
   - **Reagents:**
     - **N,N-dimethyl-2,1-benzothiazolidine-3,4-dione**
   - **Procedure:**
     - Dissolve the compound in a suitable solvent (e.g., DMF)
     - Add the reagent to the solution
     - Stir at 121°C for 5 h

**Yield:**

- Approx. 90%}

**Notes:**

- The reaction was monitored by NMR spectroscopy.
- The crude product was purified by column chromatography (silica gel).
- The purity of the product was confirmed by mass spectrometry.
1.6 Literature


CHAPTER 2. Synthesis, Functionalization and Photo-Bergman Chemistry of Enediyne Bioconjugates

2.1 Introduction

Naturally occurring enediynes are a class of compounds that bear a characteristic (Z)-3-ene-1,5-diyne pharmacophore and are isolated from soil bacteria. This pharmacophore affords a convenient means by which the enediyne cytotoxic functionality can be initiated selectively, as well as makes them among the most potent of antineoplastic agents yet discovered. Upon irradiation with UV light the otherwise shelf-stable enediynes undergo changes to their overall geometry, converting to a cytotoxic cyclic diradical species via a process commonly referred to as the Bergman cycloaromatization. The diradical intermediate of this cyclization reaction abstracts hydrogen atoms from surrounding macromolecules, resulting in DNA damage and cell death.

Scheme 2.1. Bergman cyclization of (Z)-hex-3-ene-1,5-diyne.

Over twenty enediynes are already known, including dynemicin A, esperamicin, neocarzinostatin and calicheamicin. In the year 2000 an anti-CD33 antibody conjugate of
calicheamicin, Mylotarg®, became the first monoclonal (mAb) conjugated cytotoxin to receive approval by the FDA for its clinical benefit against acute myeloid leukemia (ALM)

Given their demonstrated potency as antineoplastic agents and the selective control light (UV) activated cyclization offers, compounds bearing the enediyne moiety bear considerable promise as candidates for targeted photodynamic therapies. Accordingly, a number of designed enediyne prodrugs have been developed, however its lipophilic nature limits the applications of these compounds to biological systems. We have therefore become interested in developing efficient synthetic routs to “building-block” templates bearing the enediyne moiety and investigating various bioconjugation chemistries to decorate these templates with groups capable of expanding their applications to biological systems and enhancing their targeting abilities. In pursuit of these goals we examined (i) the preparation versatile enediyne building blocks (ii) derivatization to form water-soluble analogues (iii) coupling to form PEG-ylated derivatives (iv) formation of monoclonal antibody (mAb) conjugated chimeras and (v) investigation of the potential to form nanoparticle coupled derivatives.

2.2 Previous Work in our Laboratory

2.2.1 Template Design

A synthetic route to the mono-alcohol enediyne template 11, capable of routine
derivatization for bioconjugation, had been previously investigated in our laboratory. This synthesis was accomplished as shown in scheme 2.2, and was scalable through multiple gram scale. The stability of 11 was demonstrated by its preserved fidelity over several weeks at room temperature.

**Scheme 2.2.** Initial synthesis of mono-alcohol enediyne template.

![Scheme 2.2](image)

2.2.2 PEGylation and Antibody Conjugation of Mono-Alcohol Template

PEGylation is a known strategy for improving the delivery of lipophilic drug candidates, and was a logical first step in examining the bioconjugation tolerance of our lipophilic enediyne templates. PEG conjugation is known to improve delivery via (i) increasing water solubility of lipophilic molecules to better accommodate drug delivery in biological systems, (ii) block liposomal recognition and removal of the molecule in both the circulatory and lymphatic systems to increase circulatory capacity, (iii) decrease
immunogenic responses to biological molecules by mimicking post-translational modifications via PEGylation, and (iv) enhance targeting to and uptake of the molecule in solid tumors via trapping the large PEG molecule in the often poorly formed/leaky vasculature common to rapidly grown tumors\textsuperscript{86-89}.

Mono-alcohol enediyne template \textbf{11} was coupled to a C-16 polyethylene glycol (PEG) chain under EDCI coupling as outlined in \textbf{Scheme 2.3}. These transformations were efficient and demonstrated a “remarkable” improvement in solubility upon PEGylation in both aqueous and buffered solution\textsuperscript{6}.

Further improvement on the biological and targeting applications of \textbf{11} included the installation of 2G4 and 2G5 mAbs to the PEG conjugated template (\textbf{Scheme 2.3}). These transformations were facile and ELISA confirmed the mAbs retained immunocompetence when conjugated\textsuperscript{75,90}.

The ability of the bioconjugated analogs of mono-substituted template \textbf{11} (\textbf{12, 14-15}) to undergo light activated Bergman cycloaromatization was confirmed (\textbf{Scheme 2.4}) via irradiation through a low-pressure Hg lamp (450 W) using isopropanol as hydrogen atom donor. Conversion to the cycloaromatized species was monitored with Ultraviolet–visible spectroscopy (UV/Vis) spectroscopy based on the characteristic increase in wavelength corresponding to aromatization. Nuclear magnetic resonance (NMR) and Mass spectrometry (MS) confirmed the identity of products, and Matrix-Assisted Laser Desorption/Ionization-MS (MALDI-MS) was used for the PEGylated derivatives and
bioconjugates.

A further demonstration of the versatility of these enediyne template building blocks included the development of a surface modified gold nanoparticle conjugate\textsuperscript{91,92}. In preparation for conjugation mono-alcohol template 11 was first converted to the thiol analogue 19. This was subsequently converted to the gold derivative (20) following the Brust two-phase method\textsuperscript{75}. The ability of the gold nanoparticle analogue to cyclize was confirmed as described above and followed by oxidative cleavage and Gas chromatography–mass spectrometry (GCMS) analysis of the enediyne moiety. Transmission electron microscopy (TEM) analysis was used to examine the morphology of the Au derivatives, which had an average core size of 5.0 nm and were shelf-stable up to a month following derivatization\textsuperscript{84}. 
Scheme 2.3. PEG-ylation and antibody-enediyne conjugation.

1. DCC, DCM (96%)
2. Amberlyst\(^{15}\)
   CH\(_3\)COCH\(_3\), H\(_2\)O (90%)
3. NaCNBH\(_3\), HEPES pH 7.4

2G4 mAb or 2G5 mAb

2G4 = 3.15
2G5 = 3.16
Scheme 2.4. Photo-Bergman cyclization of enediyne core and bioconjugates.

Scheme 2.5. Preparation of gold modified nanoparticle-enediyne conjugate.

The promising results from these early studies supported the decision to pursue further investigation of improving the enediyne pharmacophores as potential anti-tumor agents.

A bis-alcohol enediyne template for tandem bioconjugation and quaternary ammonium salt and succinic acid analogues were developed and evaluated.
2.3 Synthesis of Bis-functionalized enediyne templates

2.3.1 Bis-Alcohol

Synthesis of the bis-alcohol enediyne template 25 followed a similar approach to that of the mono-alcohol (Scheme 2.6). Coupling of 4-iodo-benzyl alcohol to 22, however, proved to be unsuccessful, even when the solvent and base were varied. This may be a result of Sonogashira incompatibility with protic functionalities. Alternatively, when 4-iodo-benzyl alcohol was protected with t-butyldimethylsilyl chloride (TBS-Cl) (23) prior to coupling with 22 the transformation did occur. Deprotection was then accomplished with tetra-n-butylammonium fluoride (TBAF), giving a 34% overall yield for the synthesis of 25. 
Scheme 2.6. Synthesis of enediyne diol.

2.4 Synthesis of Water Soluble Ammonium Salts

As investigational prodrugs for use in biological systems, which are composed primarily of water, water solubility of the bioconjugated enediynes is a necessity. Although both had an alcohol functionality, the mono- and bis-alcohol enediyne templates were still somewhat lipophilic. We believed the water solubility of these could be improved through conjugation with various amines to form ammonium salt analogues.
Initially we focused on converting the benzyl alcohol functionalities of 11 and 25 to easy-to-remove leaving groups that could double as counter ions to the ammonium salt. Attempts to convert the alcohols using sulfonyl chloride, boron tribromide, and tosyl chloride were unsuccessful, yielding complex reaction mixtures including unreacted starting material (Scheme 2.7). A variety of reaction conditions were probed with limited success and decomposition upon purification using flash column chromatography was noted. Full conversion to a 95% pure product was accomplished with methanesulfonyl chloride in dichloromethane (demonstrated via $^1$H NMR of the crude material) but decomposition upon purification by flash chromatography led to product loss. The crude material was therefore immediately carried on immediately in the synthesis. For this step, the mesylate precursor was added slowly to a pre-formed solution of its corresponding amine in acetonitrile (Scheme 2.8). A variety of amines were screened, with 1,4-diazabicyclo[2.2.2]octane (DABCO) emerging as the optimal amine in light of its resultant high yields from quick reaction times and low temperatures. The salts proved to be freely soluble in water, polar organic solvents, and buffering media. GCMS and UV/Vis confirmed 33 and 34 were able to undergo the light-activated cycloaromatization and it was noted that, perhaps due to steric strain, the bis-salts generally underwent slower conversion to their cyclized products than the mono-salts.$^6$
Scheme 2.7. Attempted conversion of enediyne benzyl alcohols to leaving groups.
Scheme 2.8. Synthesis of tri-alkyl ammonium salts.

2.5 Succinic acid derivatization and PEG-ylation of the bis-alcohol templates

We envisioned an application of the bis-alcohol templates in which differential bioconjugation via tandem PEG-ylation and/or various functionalities could be incorporated onto one enediyne core. Initial attempts to convert 25 to its PEG-ylated form via EDCI coupling, as was utilized for PEG-ylation of the mono-alcohol template, led to an inseparable mixture of mono- and bis-PEG-ylated products 36 and 37. Converting 25 first to a bis-succinate (38) and subsequent PEG-ylation via carbodiimide coupling efficiently produced the desired bis-PEG-ylated ester 39 (Scheme 2.9).
Scheme 2.9. Succinic acid derivatization and PEG-ylation of the bis-alcohol templates.

2.6 Summary

Readily available enediynes can be prepared and efficiently functionalized allowing conversion to a variety of stable derivatives. The inherent lipophilicity of the enediyne core has been addressed via PEG-ylation and derivatization to ammonium salts,
enhancing their applicability for use in biological systems. Bioconjugation of monoclonal antibodies to the PEG linker has also been accomplished, following with immunocompetence was demonstrated to be retained. The enediyne templates have been shown to undergo light activated Bergman-cycloaromatizations to yield arene products in both their template and derivatized forms. Finally, Au nanoparticle enediyne conjugates have been prepared using a two-phase method. The methodology described herein is flexible and efficient, with expected applications in bioorganic, medicinal, and materials photochemistry.

2.7 Experimental

**General Experimental Information:** NMR spectra were obtained either on a Varian Mercury 300 (300 MHz) or a Varian Inova 500 (500 MHz) spectrometer. Mass spectra were obtained on a Hewlett Packard 6890 Plus GC with 5973 Mass Selective Detector and Agilent HP Ultra 1, 25 m x 0.2 mm column with a 0.11 lm film thickness. The electron impact ionization source was run at 70 eV. MALDI was conducted on an Applied Biosystems 4800 MALDI-TOF/TOF. High-resolution TEM analyses were performed on JEOL, JEM 1010 and the images captured using an Ex41B 4MP bottom mount CCD camera system (Advanced Microscopy Techniques) and visualized using a 254 nm / 366 nm UV lamp, phosphomolybdic acid, Ninhydrin and/or iodine vapor. Preparative thin layer chromatography was carried with Silica Gel GF (Analtech, Inc.). Flash chromatography was performed using silica gel 60 (230-400 Mesh, Whatman Inc.). High performance liquid chromatography was performed using Waters 510 pump
and gradient controller equipped with Waters 486 tunable UV absorbance detector and
474 scanning fluorescence detectors. All reactions were carried out under anhydrous
conditions, under inert atmosphere (nitrogen or argon) with dry, freshly distilled solvents
and flame-dried glassware unless otherwise noted. Diethyl ether and THF were distilled
from sodium-benzophenone ketyl. Dichloromethane, chloroform, methanol, isopropanol,
ethanol, triethylamine, benzene, toluene and acetonitrile were distilled from calcium
hydride. DMF and DMSO were stirred with dry 4 Å molecular sieves for 12 h (unless
otherwise described), and distilled under reduced pressure. PEG reagents were purchased
from Laysan Bio, Inc.

**General Procedure for photo-cyclization of enediynes:** A solution of enediyne (5 mg,
0.005% in HPLC grade isopropanol) was added to a 14 mL quartz UV cuvette and
degassed (Ar, 30 min). In a photochemical box, a 0.01 M potassium chromate solution
filter was placed between the lamp and the solution, and irradiation initiated using a 450
W Hanovia low-pressure Hg lamp. The reactions were stirred for up to 3 h (N₂
atmosphere) and monitored by both UV/Vis and mass spectrometry to determine the rate
and extent of formation of cyclodaromatized product.

**1-iodo-2-(ethynylphenyl)benzene (8):** To a solution of 1,2-diiodobenzene (10 g, 30.03
mmol) in diethylether (30 mL) was added dichloro(bistriphenylphos-phine)palladium
(0.85 g, 1.21 mmol). The solution was degassed (N₂) for 45 min. In a second flask copper
iodide (0.49 g, 2.55 mmol) and n-butyl amine (1.78 g, 24.3 mmol) were added to a
degassed solution of phenyl acetylene (0.62 g, 6.0 mmol) in ether (20 mL). The
diiodobenzene solution was cannulated into the alkyne solution and the resulting mixture stirred for 12 h at 25 °C and then saturated ammonium chloride (20mL) added. Ethyl acetate (20 mL) was added to the biphasic mixture, and the organic layer washed with water (20 mL) and the combined aqueous layers were extracted with ethyl acetate (2 x 30 mL). The combined organic layers were washed with brine (30 mL), dried (MgSO₄) and concentrated in vacuo. The crude solid was subjected to flash chromatography (hexanes) to yield the title compound (1.08 g, 59%) as a white solid (mp 62–65 °C); Rₖ: 0.36 (hexanes); ¹H NMR (300 MHz, CDCl₃): 7.87 (dd, ¹H, J = 3.9, 6.3 Hz), 7.61 (m, 2H), 7.5 (dd, ¹H, J = 4.5, 7.2 Hz), 7.35 (m, 4H), 7.03 (dd, ¹H, J = 3.6, 6.0 Hz); ¹³C NMR (75 MHz, CDCl₃): 138.7, 132.7, 132.6, 129.5, 129.3, 128.7, 128.5, 127.8, 122.0, 101.2, 93.0, 91.6; HRMS (Cl), m/z C₁₄H₉I (M+H)+: calcd 304.9828, obsd 304.9828.

1-(trimethylsilylethynyl)-2-(ethynylphenyl)benzene (9): To a degassed solution of 1-iodo-2-(ethynylphenyl)benzene (1.08 g, 3.55 mmol) in triethylamine (10 mL) was added dichloro(bistriphenylphosphine)palladium (0.1 g, 0.14 mmol) and copper iodide (0.055 g, 0.29 mmol). The solution was degassed and trimethylsilylacetylene (0.46 g, 4.67mmol) added. The solution was again degassed and then stirred for 12 h at 25 °C. The solution was quenched with saturated ammonium chloride (10 mL). Ethyl acetate (20 mL) was added to the biphasic mixture. The organic layer was washed with water (20 mL) and the combined aqueous layers were extracted with ethyl acetate (2 x 30 mL). The combined organic layers were washed with brine (30 mL), dried (MgSO₄) and concentrated in vacuo. The residue was subjected to column chromatography (1:99, ethyl acetate/hexanes) to yield the title compound as a yellow oil (0.55 g, 57%); Rₖ: 0.25 (hexanes); ¹H NMR
(300 MHz, CDCl₃): 7.56 (m, 2H), 7.50 (m, 2H), 7.32 (m, 3H), 7.23 (m, 2H), 0.26 (9H); ¹³C NMR (75 MHz, CDCl₃): 132.5 131.9, 128.6, 128.5, 128, 0.0; HRMS (CI), m/z C19H18Si (M+H)+: calcd 275.1256, obsd 275.1256.

1-(Ethynylphenyl)-2-ethynylbezene (10): A solution of 1-(trimethylsilylethynyl)-2-(ethynylphenyl)benzene (0.73 g, 2.66 mmol) in tetrahydrofuran (20 mL) was cooled to 0 °C. Tetrabutylammonium fluoride (2.79 g, 10.7 mmol, 1.0 M in THF) was added drop-wise. The reaction mixture was stirred for 2.5 h at 0 °C. Water (15 mL) was added slowly to quench the reaction. The biphasic mixture was extracted with ethyl acetate (2 x 30 mL). The combined organic layers were washed with brine (30 mL), dried (MgSO₄) and the solution filtered through a short plug of silica gel and concentrated in vacuo to yield the title compound as a yellow oil (0.53 g, 100%). Rf: 0.27 (hexanes); ¹H NMR (300 MHz, CDCl₃): 7.59–7.52 (m, 4H), 7.36–7.25 (m, 5H), 3.36 (s, 1H); ¹³C NMR (75 MHz, CDCl₃): 132.8, 132.0, 128.6, 128.1, 127.9, 126.6, 124.9, 123.4, 93.78, 88.1, 88.3; HRMS (CI), m/z C16H10 (M+H)+: calcd 203.0863, obsd 203.0861.

4-(2-Phenylethynyl-phenylethynyl)-benzyl alcohol (11): To a degassed solution of 1-(ethynylphenyl)-2-ethynylbenzene (0.53 g, 2.62 mmol) in triethylamine (35 mL) was added 4-iodobenzyl alcohol (0.737 g, 3.15 mmol), dichloro(bis triphenylphosphine)palladium (0.074 g, 0.11 mmol) and copper iodide (0.04 g, 0.21 mmol). The solution was degassed again (45 min N₂ purge) and then stirred at room temperature for 4 h. Saturated ammonium chloride (30mL) was added followed by ethyl acetate (30 mL). The organic layer was washed with water (30 mL) and the combined
aqueous layers extracted with ethyl acetate (2 x 40 mL). The combined organic layers were washed with brine, dried (MgSO₄) and concentrated in vacuo. The residue was purified by flash chromatography (70:30, hexanes/ethyl acetate) to yield the title compound (0.65 g, 81%) as a tan solid. (mp 95–97 °C); ¹H NMR (300 MHz, CDCl₃): 7.59–7.25 (m, 13H), 4.71 (s, 2H), 1.75 (br s, 1H); ¹³C NMR (75 MHz, CDCl₃): 141.4, 132.1, 132.0, 131.9, 128.7, 128.6, 128.3, 127.1, 126.1, 123.5, 122.8, 93.8, 93.7, 88.6, 88.5, 65.2; HRMS (CI), m/z C23H16O (M+H)+: calcd 309.1281, obsd 309.1280.

1,2-bis((trimethylsilyl)ethynyl)benzene (21): Pd(PPh₃)₂Cl₂ (0.032 g, 3 mol%) and CuI (0.014 g, 5 mol%) were added to a degassed (Ar 30 min.) solution of 1,2-diiodobenzene (0.500 g, 198 µL, 1.516 mmol) in triethylamine (5 mL). The solution was degassed (Ar, 30 min.) then TMS-acetylene (535 µl, 2.5 eq) added and the mixture stirred for 16h at room temperature. Saturated aqueous NH₄Cl (5 mL) was added followed by ethyl acetate (10 mL). The layers were separated and the organic layer was washed with water (2x 10 mL). The aqueous layers were back extracted with ethyl acetate (2 x 10 mL) and the combined organic extracts washed with brine (10 mL), dried (MgSO₄) and concentrated in vacuo. The residue was purified by flash column chromatography to afford the title compound (0.272 g, 66.4%) as a yellow oil which was spectroscopically identical to that reported⁹³. ¹H NMR (75 MHz, CDCl₃): δ 0.27(s, 18H), 7.19-7.23(m, 2H), 7.45(dd, J=6.3Hz, 3.3Hz, 2H); ¹³C NMR (300MHz, CDCl₃): δ 0.00, 98.36, 103.23, 125.75, 127.99, 132.24.

1,2-diethynylbenzene (22): Potassium carbonate (0.542 g) was added to a solution of
1,2-bis((trimethylsilyl) ethynyl)benzene (0.212 g, 0.784 mmol) in methanol/THF (1:1, 80 mL). The resulting solution was stirred at 25 °C (Ar atmosphere) for 30 min then diluted with ether (300 mL). The mixture was washed with saturated ammonium chloride (4 x 150 mL), water (4 x 150 mL) then dried (MgSO₄) and condensed in vacuo to afford the title compound (0.098 g, 99%) as a clear oil. TLC (hexanes/ethyl acetate = 9:1): Rₕ: 0.42; ¹H NMR (300 MHz, CDCl₃): d 7.50 (dd, J = 6.3 Hz, 2H), 7.24–7.31 (m, 2H), 3.33 (s, 2H); ¹³C NMR (75 MHz, CDCl₃): d 132.87, 128.76, 125.30, 82.09, 81.49.

1,2-bis((4-(tert-butyldimethylsilyloxy)methyl)phenyl)ethynyl)benzene (24):
Pd(PPh)₂Cl₂. (0.0445 g, 8 mol %) and CuI (0.1208 g, 8 mol %) were added to a degassed (Ar, 30 min) solution of 1,2-diethynylbenzene (0.100 g, 0.793 mmol) and 4-((tertbutyldimethylsilyloxy)methyl)-phenyl iodide14 (0.607 g, 1.75 mmol) in triethylamine (25 mL). The resulting solution was stirred at 25 °C (Ar atmosphere) for 5 h then diluted with saturated ammonium chloride (5 mL) and ethyl acetate (300 mL). The organic layer was washed with brine (100 mL), dried (MgSO₄) and condensed in vacuo. The residue was purified by flash column chromatography (19:1 hexanes/ethyl acetate) to afford the title compound (0.240 g, 53%) as a yellow solid. Mp = 109–111 °C; TLC (hexanes/ethyl acetate = 19:1): Rₕ: 0.35; ¹H NMR (300 MHz, CDCl₃): d 7.48–7.60 (m, 4H), 7.21–7.36 (m, 8H), 4.76 (s, 4H), 0.95 (s, 18H), 0.11 (s, 12H); ¹³C NMR (75 MHz, CDCl₃): d 155.66, 147.16, 136.96, 136.83, 133.12, 131.21, 131.16, 127.04, 99.00, 93.30, 69.93, 31.20, 23.63, 0.0; HRMS (CI), m/z C₃₆H₄₆O₂Si₂ (M+H)+ calcd 567.3116, obsd 567.3115.
(4,4’-(1,2-phenylenebis(ethyne-2,1-diyl))bis(4,1-phenylene))dimethanol (25): Tetrabutylammonium fluoride (815 µl, 0.815 mmol, 1 M in THF) was added to a solution of 1,2-bis((4-((tert-butyl(dimethyl)silyloxy)methyl)phenyl)ethynyl)benzene in tetrahydrofuran at 0 °C (Ar atmosphere) and stirred for 30 min. Water (20 mL) was added. The solution was condensed in vacuo, diluted with ethyl acetate (300 mL) and the organic layer washed with water (200 mL) and brine (100 mL), dried (MgSO₄) and condensed in vacuo. The residue was purified by flash chromatography (2:1 ethyl acetate/hexanes) to afford the title compound (0.133 g, 98%) as a white solid. Mp = 146–147 °C; TLC (hexanes/ethyl acetate = 1:1): Rf: 0.18; ¹H NMR (500 MHz, CDCl₃): δ 7.54–7.59 (m, 6H), 7.34–7.38 (m, 4H), 7.30–7.33 (m, 2H), 4.73 (s, 4H), 1.4–1.9 (br s, 2H); ¹³C NMR (75MHz, CDCl₃): δ 141.32, 131.97, 131.91, 128.19, 127.00, 125.93, 122.66, 93.57, 88.52, 65.11; HRMS C24H18O2 (Cl), m/z (M)+: calcd 338.1307, obsd 338.1307.

1-(4-((2-(phenylethynyl)phenyl)ethynyl)benzyl)-4-aza-1-azoniabicyclo[2.2.2]octane methanesulfonate (salt) (32): Triethylamine (41 µl, 3 equiv) and methanesulfonyl chloride (15 µl, 2 equiv) was added to a solution of 4-(2-phenylethynyl-phenylethynyl)benzyl alcohol in dichloromethane (4 mL) at 0 °C (Ar atmosphere) then stirred at 25 °C for 1 h. The resulting solution was diluted with ethyl acetate (25 mL) and washed with citric acid (5%, 30 mL), sodium bicarbonate (5%, 30 mL) and brine (30 mL), dried (MgSO₄) and condensed in vacuo. This residue was added drop-wise to a solution of 1,4-diazabicyclo[2.2.2]octane (DABCO) (0.0086 g, 0.7666 mmol) in acetonitrile (2 mL) at 0
°C. The resulting solution was stirred at 0 °C for 30 min, warmed to 25 °C and condensed in vacuo. The residue was recrystallized from hexanes, methanol and ethyl acetate to afford the title compound 0.036 g, 94% as a yellow solid. Mp 60–61 °C; \(^1\)H NMR (500 MHz, CDCl\(_3\)): 7.51–7.57 (m, 8H), 7.34–7.39 (m, 2H), 7.24–7.33 (m, 3H), 4.89 (s, 2H), d 3.59 (t, J=7.5Hz, 6H), 3.13 (t, J = 7.5 Hz, 6H), 2.84 (s, 3H); \(^{13}\)C NMR (75 MHz, CD3OD): d 135.54, 134.20, 134.10, 134.02, 133.53, 130.69, 128.94, 128.01, 127.94, 127.14, 125.24, 95.39, 93.90, 92.08, 89.65, 69.60, 54.40, 46.98, 40.28; HRMS C29H27N2þ (EI), m/z (M)+: calcd 403.2169, obsd 403.2174.

N-benzyl-N,N-dimethyl-1-(4-((2-(phenylethynyl)phenyl)ethynyl)phenyl)methanaminium chloride (salt) (33)

Triethylamine (21.7 µl, 3 equiv) and methanesulfonyl chloride (12.1 µl, 3 equiv) was added to a solution of 4-(2-Phenylethynyl-phenylethynyl)-benzyl alcohol (0.016 g, 0.0519 mmol) in dichloromethane (4 mL) at 0 °C (Ar atmosphere) then stirred at 25 °C for 1 h. The resulting solution was diluted with ethyl acetate (30 mL) and washed with citric acid (5%, 30 mL), sodium bicarbonate (5%, 30 mL) and brine (30 mL), dried (MgSO\(_4\)) and condensed in vacuo. Acetonitrile (1 mL) was added and the solution was added drop-wise to a solution of N,N’-dimethylbenzylamine (7.8 µl, 1 equiv) in acetonitrile (1.5 mL) at 25 °C. The resulting solution was stirred at 60 °C for 4 h then condensed in vacuo. The residue was dissolved in ethyl acetate (10 mL) and the solution was washed with brine (10 x 10 mL). The brine washings were back extracted with dichloromethane (10 x 10 mL) then the combined organic layers were dried (MgSO\(_4\)) and condensed in vacuo to afford the title compound (0.020 g, 83.6%) as a white solid. Mp
112–115 °C; \(^1\)H NMR (500 MHz, CD3OD): 7.68 (d, \(J = 8.5\) Hz, 2H) 7.50–7.64 (m, 1\(^1\)H), 7.33–7.45 (m, 5H), 4.63 (s, 2H), 4.61 (s, 2H), d 2.95 (s, 6H); \(^{13}\)C NMR (125MHz, CD3OD): d 134.74, 134.50, 133.38, 133.16, 133.12, 132.72, 132.19, 130.58, 130.10, 130.00, 129.84, 129.67, 129.12, 128.87, 127.36, 127.33, 126.44, 124.54, 94.71, 93.21, 91.48, 88.94, 70.04, 69.43, 55.65; HRMS C32H28N+ (TOF MS ES+), m/z (M)+: calcd 426.2222, obsd 426.2210.

**N,N’-(4,4’-(1,2-phenylenebis(ethyne-2,1-diyl))bis(4,1-phenylene))bis(methylene)bis(N,N-dimethyl-1-phenylmethanaminium) methanesulfonate (salt) (34):** Triethylamine (16 µl, 6 equiv) and methanesulfonyl chloride (9 µl, 6 equiv) was added to a solution of (4,40 -(1,2-phenyl-enebis(ethyne-2,1-diyl))bis(4,1-phenylene))dimethanol (0.007 g, 0.02069 mmol) in dichloromethane (2 mL) at 0 °C (Ar atmosphere) then stirred at 25 °C for 1 h. The resulting solution was diluted with ethyl acetate (20mL) and washed with citric acid (5%, 30 mL), sodium bicarbonate (5%, 30 mL) and brine (30 mL), dried (MgSO\(_4\)) and condensed in vacuo. This residue was diluted with acetonitrile (1 mL) and added drop-wise to a solution of N,N’-dimethylbenzylamine (6.1 µl, 2 equiv) in acetonitrile (1.5 mL) at 0 °C. The resulting solution was stirred at 60 °C for 3 h then condensed in vacuo and diluted with 49:1 hexanes/methanol (10 mL). After vigorous stirring the methanol layer was removed, the hexanes layer was warmed to 25 °C and methanol (1 mL) added. This process was repeated three times and the methanol extracts were combined, dried (MgSO\(_4\)), and condensed in vacuo to afford the title compound (0.008 g, 52%) as a white solid. Mp 55–57 °C; \(^1\)H NMR (500 MHz, CD3OD): d 7.41–7.77 (m, 22H), 4.66 (s, 4H), 4.64 (s, 4H),
2.98 (s, 12H), 2.72 (s, 6H); $^{13}$C NMR (125 MHz, CD3OD): d 134.82, 134.50, 133.40, 132.20, 132.09, 130.58, 130.19, 129.27, 128.88, 127.22, 126.55, 93.35, 91.20, 70.02, 69.42, 43.09, 39.67; HRMS C42H42N2+p (TOF MS ES+), m/z (M)2+: calcd 287.1674, obsd 287.1667.

1,1’-(4,4’-(1,2-phenylenebis(ethyne-2,1-diyl))bis(4,1-phenylene))bis(methylene)bis(4-aza-1-azoniabicyclo[2.2.2]octane) methanesulfonate (salt) (35): Triethylamine (22.2 µl, 6 equiv) and methanesulfonyl chloride (12.42 µl, 6 equiv) was added to a solution of (4,40-(1,2-phenyl-enebis(ethyne-2,1-diyl))bis(4,1-phenylene))dimethanol (0.009 g, 0.0266 mmol) in dichloromethane (4 mL) at 0 °C (Ar atmosphere) then stirred at 25 °C for 1 h. The resulting solution was diluted with ethyl acetate (15mL) and washed with citric acid (5%, 20 mL), sodium bicarbonate (5%, 20 mL) and brine (20 mL), dried (MgSO₄) and condensed in vacuo. The residue was added drop-wise to a solution of 1,4-diazabicyclo[2.2.2]octane (DABCO) (0.006 g, 0.2 equiv) in acetonitrile (2 mL) at 0 °C. The resulting solution was stirred at 25 °C for 90 min then condensed in vacuo. The residue was recrystallized from hexanes, methanol and ethyl acetate to afford the title compound (0.0176 g, 92%) as a yellow solid. Mp 133–135 °C; $^1$H NMR (300 MHz, CDCl₃): 7.44–7.51 (m, 4H), 7.28–7.37 (m, 6H), 7.01–7.08 (m, 2H), 4.85 (s, 4H),3.58 (t, 12H, J=7.2Hz), d 3.14 (t, 12H, J=7.2Hz), 2.81 (s, 6H); $^{13}$C NMR (75 MHz, CDCl₃): 134.90, 134.85, 133.43, 130.20, 128.38, 127.12, 126.52, 93.28, 91.11, 68.86, 53.71, 46.29, 39.74; HRMS C36H40N2+p (TOF MS ES+), m/z (M+H)2+: calcd 264.1627, obsd 264.1614.
4,4’-(4,4’-(1,2-phenylenebis(ethyne-2,1-diyl))bis(4,1-phenylene))bis(methylene)bis(oxy)bis(4-oxobutanoic acid) (38): (4,4’-(1,2-Phenylenebis(ethyne-2,1-diyl))bis(4,1-phenylene))di-methanol (0.0165 g, 0.0488 mmol) was added to a solution of 4-dimethylaminopyridine (0.006 g, 1 equiv) and succinic anhydride (0.0059 g, 1.2 equiv) in dichloromethane (3 mL). The mixture was stirred at 25 °C (Ar atmosphere) for 12 h. The solution was diluted with dichloromethane (10 mL) and washed with aqueous sodium carbonate (5%, 25 mL). The dichloromethane layer was removed and the aqueous layer acidified by drop-wise addition of concentrated hydrochloric acid. This solution was back extracted with ethyl acetate (3 x 25 mL), washed with brine (25 mL), dried (MgSO₄), and condensed in vacuo to afford the title compound (0.0205 g, 99%) as a white solid. Mp 126–127 °C; TLC (dichloromethane/methanol/aqueous ammonium hydroxide = 80:19:1): Rf: 0.22; ^1H NMR (500 MHz, CDCl₃): d 7.53–7.60 (m, 6H), 7.29–7.35 (m, 6H), 5.17 (s, 4H), 2.64–2.75 (m, 8H); ^13C NMR (75MHz, CD3OD): d 176.16, 173.98, 138.32, 132.82, 132.66, 129.47, 129.22, 126.95, 124.16, 94.15, 89.36, 66.85, 30.05, 29.83; HRMS C32H26O8 (TOF MS ES+), m/z (M+Na)+: calcd 561.1536, obsd 561.1525.

40-(4,40-(1,2-Phenylenebis(ethyne-2,1-diyl))bis(4,1-phenylene))bis(methyleneoxy-polyethylene glycol alcohol) (39): 4,4’-(4,4’-(1,2-phenylenebis(ethyne-1,2-diyl))bis(4,1-phenylene))bis(methylene)bis(oxy)bis(4-oxobutanoic acid) (0.004 g, 0.00742 mmol) was added to a solution of 1-ethyl-3-(3’-dimethylaminopropyl)carbodiimide (EDCI) (0.0031 g, 2.2 eq) and 4-Dimethylaminopyridine (DMAP) (0.0020 g, 2.2 eq) in dichloromethane (1.5 mL) at 25°C
(Ar atmosphere). The mixture was heated at reflux for 4h then cooled to 25°C and 40
0Mn polyethylene glycol (5.8 µl, 2.2 eq) was added. The solution was stirred for 48h at
25°C, then diluted with dichloromethane (10 mL), washed with brine (2 x 10 mL), dried
(MgSO₄), and concentrated in vacuo. The crude residue was purified via preparative thin
layer chromatography (90:9:1 DCM:MeOH:aq) to yield the title compound (3.7mg, 39%)
as a yellow oil. TLC (dichloromethane:methanol:aq ammonium hydroxide = 80:19:1): Rf:
0.31; HRMS C66H90O24 (MALDI): m/z (M+2Na)2+ calcd 1285.43982, obsd
1287.44434.
1,2-bis((4-((tert-butyl(dimethyl)siloxy)methyl)phenyl)ethynyl)benzene
7/3/08
KW/JLL
STANDARD CARBON PARAMETERS

Pipe length: 21 cm
Solution: C6D6
Solvent: C6D6
Solvent temperature: 20°C
Pulse: 10.2 μs
Temperature: 296 K

Tubes: 9.7 degrees
Echo time: 0.40 ms
Delay time: 0.10 s

FID repetitions: 32
FID acquisition: 15,565.74 μs
Sample: 14.57 ppm
Phase: 10 kHz

On running acquisition:
SCF during delay

Acq. mode: PSD
Mix: 100 kHz

1H excitation 1.9 Hz
13C excitation 59 Hz

Total time: 31 ms, 31 ms, 2 ms

Compound 16

[Diagram of a molecule]
2.8 Literature


CHAPTER 3. Radio-iodination methods for the production of SPECT Imaging Agents

3.1 Introduction

Molecular imaging via Single Photon Emission Computed Tomography (SPECT) holds considerable promise for analysis of metabolism and bio-distribution of drugs in addition to imaging of diseased states. While the related Positron Emission Tomography (PET) imaging offers higher resolution images, SPECT is considered the more practical approach to nuclear imaging for routine diagnostic use. One of the attractive features of SPECT is its utilization of medium to long-lived radioisotopes, particularly radio-iodinated species. This mini-review will survey recent advances in radio-iodination, including synthetic chemistry and post-labeling purification methods.

3.2 Background to SPECT Imaging

Nuclear medicine plays a fundamental role in both the diagnosis and treatment of disease. This noninvasive technique provides invaluable in vivo visualization of biological processes at tissue, cellular and sub-cellular levels. Its benefits extend from presymptomatic detection of disease, through imaging with radionuclide tagged molecular probes, to radionuclide therapy, a practical alternative to traditional external beam radiation therapy. At the forefront of current imaging technology are Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT),
modalities with considerable promise for analysis of metabolism and biodistribution of drugs, in addition to imaging of diseased states. Since their introduction much debate has arisen as to which is most appropriate for general adoption. While PET offers higher resolution images and better selectivity\(^95\), its use has not yet become widespread outside of dedicated clinical centers\(^96\). In addition to the high capital cost of PET equipment, the tracers typically used have relatively short half-lives requiring specialist radiochemical facilities and access to nearby cyclotron sources. SPECT is considered the more practical approach to nuclear medical imaging for routine diagnostic use\(^97\). Infrastructure is less expensive than PET and the longer-lived tracer isotopes increase its convenience in the clinic as well as in the production of more elaborate radiopharmaceuticals from remote sites\(^95\). An additional advantage to SPECT is that numerous research institutions, medical centers, and hospitals regionally and locally possess necessary imaging and isotope handling equipment and thus the transition to establish a clinical SPECT imaging is a logical one\(^98\).

### 3.3 General Considerations for Radio-iodination

Radio-iodinated species are practical for SPECT labeling for a number of different reasons. Its chemistry is relatively well understood and has already been applied to label a wide range of molecules, either in \(^{123}\text{I}\) \((t_{1/2} = 13.2 \text{ h})\) or \(^{125}\text{I}\) \((t_{1/2} = 1440 \text{ h})\) form\(^99\). Covalently labeled iodinated compounds typically demonstrate high \textit{in vivo} stability - a necessity of any radiopharmaceutical. It is also possible to radio-iodinate with high specific activity, and the isotopes of iodine are relatively easy to use and commercially
available at nominal cost\textsuperscript{100}.

Numerous classical iodination methods can be readily applied to radio-iodination procedures. When being adapted, careful consideration to must be given to several key parameters that apply when dealing with any radioisotope\textsuperscript{96, 99, 101}.

1. Half-life and reaction time: \textit{Is it realistically possible to perform the labeling method without substantially compromising radiochemical yield through radioactive decay?}

2. Labeling site: \textit{Is there an appropriately activated position on the structure of the tracer molecule? Is it possible to use an organic precursor prior to labeling? Should the molecule be labeled in an aliphatic or an aromatic position?}

3. Radiochemical purity: \textit{Can the labeled product be easily and quickly separated from its precursor? Is there a potential for undesirable or toxic side products? Can the labeled product be easily separated from contaminants?}

4. Specific Activity: \textit{What specific activity is necessary for the radiopharmaceutical? Can such specific activity be achieved while still only injecting a minimal amount of tracer to ensure the biochemistry will not be altered in any way?}

5. Reaction scale: \textit{Can the tracer be realistically labeled in high radiochemical yield using the reaction concentrations typical of small-scale radiosynthesis?}
Careful attention must also be paid to the biological effect of labeling, as altering a molecule in any way can alter its biological behavior. Only radio-iodination techniques that necessitate minimal structural changes as far as possible from the pharmacophore of the molecule should be realistically considered\textsuperscript{8, 96}. It is also important to consider the stability of the carbon- iodine bond at the particular labeling site to ensure \textit{in vivo} stability of the radiopharmaceutical\textsuperscript{101, 102}. Research has shown that vinylic and aromatic iodo compounds enjoy the highest chemical, as well as frequently, \textit{in vivo} stability\textsuperscript{96}.

The high cost and time consuming process of discovery, development, validation, and implementation of new imaging probes into clinical use has led to considerable lag between the discovery of new targets and associated imaging agents\textsuperscript{8}. While numerous methods have been developed for the radio-iodination radio-iodination of various biological molecules, relatively few have reliably produced labeled compounds with both adequate chemical yield and specific radioactivity\textsuperscript{101}. As most current research utilizes these primary methods they are reviewed herein based on reaction class, and insight to future developments presented.

3.4 Nucleophilic Substitution Methods

Direct nucleophilic substitution of an efficient leaving group with radioactive iodine/iodide is a well-developed strategy. Nucleophilic substitution methods are most commonly used to label aliphatic sites\textsuperscript{96}, and while not as widely used as electrophilic substitution methods\textsuperscript{101}, their application has still produced a number of successfully
labeled compounds.

3.4.1 Isotopic Exchange

Isotopic Exchange methods benefit from the simplicity of substituting radioiodine for a non-radioactive iodine atom already present in the molecule. Such exchange can frequently be accomplished by heating the unlabelled compound in the presence of a radioactive iodide source and an appropriate solvent. Ideally the solvent selected should solubilize the substrate and inorganic iodide$^{103}$. The use of copper catalysts in these reactions has been shown to improve yield considerably$^{96}$.

Unfortunately only radiopharmaceuticals with modest specific activity can be prepared using this method due to competition between the radioactive and non-radioactive species$^{96,103}$. Chezal et al.$^{104}$ recently reporting labeling (hetero)aromatic analogues of $N$-(2-diethylaminoethyl)-4-iodobenzamide 40 with $^{125}$I in radiochemical yields as high as 92% (Scheme 3.1). The labeled analogs were purified via the use of an Extrelut column, and yielded products with a specific activity of 4.7-36.3 MBq/µmol.
3.4.2 Interhalogen Exchange

Interhalogen exchange reactions utilize displacement of a “cold” bromo (and to a lesser extent, chloro) precursor with radioactive iodide. These reactions yield labeled compounds that can be easily prepared under no-carrier-added conditions\textsuperscript{101}. The advantage to using this method (vs. isotopic exchange) is compounds with very high specific activity can be prepared. A disadvantage is the method lies in the potential for time-consuming separation of the radioiodinated product from its brominated precursor\textsuperscript{96,103}.

Vaidyanathan et al. recently reported labeling MIBG–octreotate conjugates \textbf{42} with \textsuperscript{131}I.
Radio-iodination was accomplished in up to 36% radiochemical yield by heating bromo derivatives of peptide conjugates in the presence of radioiodine and acetic acid (Scheme 3.2). While the radiochemical yield of product 43 was modest, they have suggested improvements may be attainable by optimization of reagent stoichiometry.\textsuperscript{105}

**Scheme 3.2.** Interhalogen exchange radio-iodination.\textsuperscript{105}
3.4.3 Radioiodo-dediazonization (using the Wallach triazene reaction)

This method represents an improvement on Sandmeyer-type reactions formerly used to label deactivated or non-activated aromatic rings with radioiodine in a particular position\textsuperscript{101}. Specifically, a diazotized amine is trapped by the formation of a triazene with a secondary amine\textsuperscript{103}. When required, the thermodynamically stable triazenyl species\textsuperscript{106} is then decomposed in acidic medium in the presence of radioiodine to yield the targeted labeled species. The method is an efficient alternative to other nucleophilic exchange techniques and offers easy purification of the radioiodinated product from its precursor\textsuperscript{106}. Vivier et al. reported success using this method in the synthesis of a series of $^{125}$I labeled leucine peptides linked to a benzamide structure 45, with radiochemical yields $\sim70\%$ and radiochemical purity of $>99\%$ (Scheme 3.3)\textsuperscript{107}.

Scheme 3.3. Radioiodo-dediazonization of leucinal derivatives\textsuperscript{107}.

![Scheme 3.3](image-url)
3.5 Electrophilic Substitution Methods

Electrophilic substitution methods are commonly employed in radiolabeling, typically via an oxidized iodine species (with a formal +1 charge) acting as an electrophile and attacking an electron rich species, resulting in the formation of a covalent carbon-iodine bond the loss of a leaving group. Ideally the leaving group should be able to comfortably accommodate a positive charge, making organometallic moieties especially useful in these reactions. The methods are most commonly employed in the synthesis of radioiodoarenes.96

3.5.1 Halo-demetallation (via Iododestannylation)

Iododestannylation reactions employ an organometallic precursor which undergoes subsequent electrophilic substitution. The attacking species is generated in situ by the oxidation of iodide with one of the many different oxidizing agents currently available.96 The method possesses many advantages over the others mentioned thus far. The polarized nature of the carbon-tin bond results in efficient reactions with high yields and also allows regioselective labeling. Furthermore, the highly labile organometallic precursor alleviates the need for harsh reaction conditions.96,101 Finally, the metallic precursor can typically be introduced into the molecule in the final stages of synthesis, reducing loss from radioactive decay and allowing for efficient “one-pot” radio-
iodination\textsuperscript{108}. A number of different oxidizing agents have been employed in iododestannylation reactions depending on the particular species being labeled. Their development over time has led to a few select reagents, which are thus highlighted.

**Chloramine-T**

One of the earliest oxidizing agents to be used in these reactions, chloramines-T was originally used by Hunter and Greenwood to radioiodinate antibodies in high specific activity using \textit{p}-\textit{toluene} sulfonochloramide\textsuperscript{103, 109}. In brief, chloramine-T releases hypochlorous acid in aqueous solution. This oxidizes iodine which generates the electrophilic iodide species in the form of an iodonium ion\textsuperscript{96}. While chloramine-T is an effective oxidizing agent, numerous byproducts can result from the harsh conditions involved. Among others these include products from chlorination, oxidation of thiol and thioester groups, cleavage of tryptophanyl peptide bonds, and denatured proteins\textsuperscript{103}. Despite these limitations, the process remains popular. Pham et al. recently used this method to label benzamides with \textsuperscript{123}I in high radiochemical yield (50-95\%) (Scheme 3.4). Purification of the radiolabeled species 47 was performed by semi-preparative C18 RP- HPLC, yielding the targeted \textsuperscript{123}I benzamides with a radiochemical purity of > 95\% and a specific activity of > 2 GBq/nmol\textsuperscript{110}. 


Scheme 3.4. $^{123}$I radiolabeling of benzamides via chloramine-T induced halodestannylation\textsuperscript{110}.

Using similar conditions, Zoelle et al. reported labeling of iodophenyl metomidate (IMTO), with $^{131}$I in both high radiochemical yield and specific activity (Scheme 3.5)\textsuperscript{111}. The product 49 is an inhibitor of steroid 11-hydroxylation in the adrenal cortex.

Scheme 3.5. $^{131}$I-IMTO radio-iodination using chloramine-T\textsuperscript{111}. 

\begin{align*}
\text{Scheme 3.4.} & \quad $^{123}$I radiolabeling of benzamides via chloramine-T induced halodestannylation\textsuperscript{110}. \\
\text{Using similar conditions, Zoelle et al. reported labeling of iodophenyl metomidate (IMTO), with $^{131}$I in both high radiochemical yield and specific activity (Scheme 3.5)\textsuperscript{111}. The product 49 is an inhibitor of steroid 11-hydroxylation in the adrenal cortex.} \\
\text{Scheme 3.5.} & \quad $^{131}$I-IMTO radio-iodination using chloramine-T\textsuperscript{111}. 
\end{align*}
Iodobeads

In order to alleviate the issues associated with harsh oxidizing conditions iodobeads, small spherical polystyrene beads coated with chloramine-T, were developed. The beads are dispersed in the reaction vial during oxidation, slowing release of the oxidant and keeping its overall concentration low. Iodobeads can easily be separated from the reaction mixture via simple filtration, thus the reaction can be terminated on demand, without the need of a reducing agent. Akgun et al. utilized the method in the radio-iodination a series of benzodiazepines (e.g. 50) with $^{125}$I, yielding the targeted tracer molecules in a specific activity of 2000 Ci/mmol. The procedure was effective in the enantioselective radio-iodination of both enantiomers of 51 (Scheme 3.6).

Scheme 3.6. $^{125}$I radio-iodination of benzodiazepines using iodobeads.

Iodogen

To reduce side reactions from harsh oxidizing conditions Fraker and Speck developed iodogen as a milder oxidizing agent in 1978. Iodogen (1, 3, 4, 6-tetrachloro-3α, 6α-
diphenylglycouril) is insoluble in water and used for biphasic reactions wherein it disperses as a thin layer on the walls of the iodinating vessel\(^96,103\). This allows effective oxidation but limits the exposure of the oxidizing agent. The labeling reaction can quickly and efficiently be terminated by simply removing the aqueous reaction mixture from the coated iodination vessel\(^96\). Iodogen has proven especially effective for labeling proteins and since its development has been extensively used for this purpose\(^101\). Iodogen is now available in a commercial kit (available from Pierce Protein Research Products)\(^113\). While milder oxidizing conditions are advantageous there are some drawbacks to using this particular method. Notably, since the labeling reaction relies on catalysis at the interface it is dependent on the surface area of contact between the oxidizing agent and the tracer being labeled. This can lead to lower yields and much slower reaction times in some instances\(^101\). Liu et al. used the iodogen method in radioiodination of an antigastric cancer mAb. They successfully developed and labeled a novel dual-modality molecular probe composed of \(^{125}\)I tagged mAb 3H11 and a Fe\(_3\)O\(_4\) nanoparticle, chemically bonded by a PEG linker\(^5,114\). The dual approach was designed to enhance upon results in previous MRI studies by combining MRI and SPECT imaging technologies. (MRI for its high temporal and spatial resolution, and SPECT for its high sensitivity). The introduction of SPECT not only enhanced imaging of the tumor site but also permitted studying the \textit{in vivo} behavior and biodistribution of the nanoparticle probes - something not easily accomplished via MRI studies alone\(^114\). Their labeling procedure quickly (5 min) produced the targeted radioiodinated mAb in 85% radiochemical yield, and >98% radiochemical purity following purification by a PD-10 Sephadex G-25 column\(^114\). Wang et al. also utilized iodogen for \(^{125}\)I labeling of a tumor
targeting quinazolinone prodrug 53 with an impressive radiochemical yield of >99% (Scheme 3.7)\textsuperscript{115}.

**Scheme 3.7.** Radiolabeling of quinazolinone derivatives with iodogen\textsuperscript{115}.

\[
\begin{align*}
\text{Scheme 3.7. Radiolabeling of quinazolinone derivatives with iodogen}^{115}.
\end{align*}
\]

Peracids

Peracids represent another class of oxidant frequently employed. They are generally formed *in situ* from the reaction of hydrogen peroxide with an organic acid, keeping the oxidant concentration low and labeling conditions mild, minimizing the risk of over-oxidation\textsuperscript{96}. While they are generally very effective their use frequently produces lower radiochemical yields than corresponding processes using *N*-halooxidants\textsuperscript{96}. Maya et al. reported use of hydrogen peroxide and hydrochloric acid in iododestannylation reactions. A series of aurones 54 were labeled with \textsuperscript{125}I in 25-57% radiochemical yield within two minutes, producing the targeted species 55 in 95% radiochemical purity after purification via HPLC on a C18 column (Scheme 3.8)\textsuperscript{97}.
Scheme 3.8. Radiolabeling of aurones through halodestannylation using H\textsubscript{2}O\textsubscript{2} and HCl\textsuperscript{97}.

![Scheme 3.8 diagram]

Qu et al. reported similar success labeling a series of styrylpyridines 57 with \textsuperscript{125}I in 50-70% radiochemical yield (Scheme 3.9)\textsuperscript{98}.

Scheme 3.9. \textsuperscript{125}I labeling of styrylpyidine species using H\textsubscript{2}O\textsubscript{2} and HCl\textsuperscript{98}.

3.5.2 Halo-demetallation (via Radioiodethallation)

Radio iodo-dethallation techniques are increasingly being investigated in the synthesis of
radioiodoarenes. Analogous to radioiododestannylation, an aryl thallium precursor, usually formed from reaction with thallium(III)triflufuoracetate (TTFA) in TFA, forms an intermediate thallium-radioiodo-complex in the presence of radioiodine$^{96}$ to generate the radioiodinated species. Pham et al. reported success with this method labeling a series of novel radioiodinated benzamides with no carrier added$^{123}$I producing regioselectively radiolabeled benzamides (59, 61) in 40-55% radiochemical yield (Scheme 3.10). The radioiodinated species were purified via solid-phase extraction (RP-SPE) followed by HPLC and possessed >95% radiochemical purity and specific activity of >2 GBq/nmol$^{110}$.

**Scheme 3.10.** NCA labeling of benzamides with $^{123}$I via radioiododethallation$^{110}$.  

![Scheme 3.10](image-url)
3.6 Purification Considerations

Numerous considerations need to be taken into account following synthesis of radiolabeled ligands for SPECT imaging. Most significantly, radioactive decay of product creates an inbuilt need for labeling strategies that do not require time-consuming purification. Additionally, any unlabeled starting material present during imaging studies could compete with the radioactive tracer molecule for site occupancy, diminishing image quality and potentially engaging in non-specific binding\textsuperscript{116,117}. Halogen exchange reactions, specifically those that require iodinated precursors, are particularly susceptible to this problem\textsuperscript{99,118}. Finally, potentially cytotoxic precursor residues such as those from iododestannylation reactions can also be problematic and, for obvious reasons, must be completely removed prior to any \textit{in vivo} imaging study can be performed\textsuperscript{119}. Spivey et al. developed a novel solution to the iodostannylation toxicity problem by binding the precursor molecule to an insoluble polymer prior to labeling. This was accomplished via a trialkylgermane linker, which can undergo iododegermylative radiolabeling with considerable flexibility\textsuperscript{119}. The method also provides for greatly improved (accelerated) workup procedures\textsuperscript{120}. Solid phase labeling strategies utilize a resin bound substrate that is only released into solution upon reaction with the selected radionuclide. The unreacted starting material remains bound to the resin support, and can be easily separated out by simple filtration\textsuperscript{116}. While this type of synthesis can be very effective at reducing the time spent on purification post labeling, it too presents some additional challenges. The preparation of solid phase precursors can often be problematic, and these substrates cannot be purified via traditional techniques\textsuperscript{117}. Additionally if loading of the precursor to
the support is not quantitative the risk of impurities forming during labeling exists - a problem that might negate the advantages of the method itself\textsuperscript{116}. An alternative to solid phase supports are soluble supports, such as those developed by Donovan et al. Their fluorous labeling strategy (FLS) follows a similar principle as most solid phase techniques. Upon labeling the radioligand is released from the support and can then be separated via a simple solid phase extraction techniques\textsuperscript{116}. FLS differs, however in that it utilizes a soluble perfluoroalkyl moiety in place of the traditional insoluble solid support resin. As the radioiodinated species is released from the support it becomes non-fluorous. Purification is quickly and easily accomplished by chemoselective filtration using a commercially available fluorous solid phase extraction cartridge. In this manner the desired radioligand is eluted through the cartridge while unreacted precursor and any fluorous by-products are retained\textsuperscript{121, 122}. A great advantage to the FLS over solid phase techniques is that non-time-sensitive fluorous precursors can be purified and characterized via traditional solution phase techniques \textit{prior} to labeling\textsuperscript{116}. This eliminates the requirement that loading be quantitative, and ensures labeled product will be free of impurities following a quick purification via fluorous solid phase extraction. The fluorous precursors can be prepared following traditional synthetic methods, and have shown similar reactivity profiles to iododestannylation type reactions\textsuperscript{122}. FLS offers the potential for libraries of radiopharmaceuticals to be prepared, expanding radioiodination possibilities to use in high throughput chemistry\textsuperscript{116}. Donovan and Valliant have utilized this strategy to radiolabel a library of iodo-benzenamides in both high yield and specific activity without the need for HPLC purification\textsuperscript{121}. They have also radiolabeled IUdR and MIBG radiopharmaceuticals with different isotopes of iodine (\textsuperscript{123}I
and $^{125\text{I}}$. Most recently, the method was expanded by diversifying the synthons available for radiolabeling to include a series of fluorous ureas (63, Scheme 3.11). Radiiodination of the ureas proceeded quickly, typically in 3 min., in radiochemical yield of 87-100%. The desired species were isolated in radiochemical purity no less than 98% without the need for HPLC purification.$^{121}$

**Scheme 3.11.** Radio-iodination of fluorous ureas following the FLS.$^{121}$

3.7 Summary

The preparation of radiochemical ligands for use in SPECT imaging is developing into a rapidly maturing field, with a number of options available for the production of diverse
structures. Radio-iodinated compounds are extremely versatile contrast agents for SPECT and the efficiency of labeling, both in terms of chemical and radiochemical yield has become extremely high based on emerging new methodologies. The advent of polymer bound reagents and fluorous technology look set to expand the possibilities of molecular design yet further, and this will no doubt provide access to new and exciting classes of SPECT image contrast agents.

3.8 Literature


CHAPTER 4. Making Biosimilars a Reality: Comparison to Small-Molecules and Host Selection for Biopharmaceuticals and Biosimilar Development

4.1 Introduction

Biopharmaceuticals are biological products with active ingredients derived from biological sources using techniques of biotechnology. (The FDA Public Health Service Act defines a biological product as a “virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, or analogous product… applicable to the prevention, treatment, or cure of a disease or condition of human beings." Biological products are also referred to interchangeably as "biologics," "biological drugs," and "protein products" throughout the literature. For clarity, they will be referred to as biopharmaceuticals throughout this review.) These products are typically versions of extracellular eukaryotic proteins, however they can also be polysaccharides, nucleic acids, or a complex mixture including any or all of the above. (For simplicity in this review, the word protein will be used to describe all biopharmaceuticals because proteins are the most common biopharmaceuticals and this is not expected to change in the near future.) Biopharmaceuticals include a vast span of biological products, such as hormones, enzymes, cytokines, vaccines, growth factors, coagulation factors, and, humanized or partially humanized monoclonal antibodies (mAbs). They are clinically administered to replace or supplement natural human proteins when patients possess mutant mis- or non-functional protein variants, or functional proteins in deficient concentrations. These are
"highly-effective, life-altering therapies" used in the treatment of over 200 diverse diseases such as anemia, diabetes, cancer, hepatitis, multiple sclerosis, AIDS, and chronic inflammatory disease to name a few. Over 150 biopharmaceuticals have received United States Food and Drug Administration (U.S. FDA) approval to date including Herceptin®(trastuzumab), Humira®(adalimumab), Cerezyme®(imiglucerase), Epogen®(erythropoietin), and Gardasil®([Human Papillomavirus Quadrivalent (Types 6, 11, 16, and 18) Vaccine, Recombinant].

Unlike small-molecule pharmaceuticals, which are typically synthesized from defined chemicals and reagents via organic synthesis, biopharmaceuticals are such large and structurally complex molecules that it would be impractical to synthesize them in this way. The natural (or, as will be described in greater detail engineered) biological processes of living organisms are therefore utilized to synthesize target biopharmaceuticals. In these systems the organism acts as a "biofactory" for biosynthesis of targeted proteins that can subsequently be extracted, purified, and administered to patients for therapeutic intervention.

Since the first biotechnology company, Genentech, was founded in 1976 biopharmaceuticals have transformed the pharmaceutical industry and "revolutionized modern medicine. They extend our ability to fight diseases once considered incurable and address unmet medical needs not treatable through conventional small-molecule therapy. This remarkable clinical utility of
biopharmaceuticals can be attributed to their high efficacy\textsuperscript{153}, potency\textsuperscript{144}, and degree of receptor specificity\textsuperscript{65,113}. Biopharmaceuticals possess the potential to cure rather than just treat symptoms of disease\textsuperscript{153} and create the unique opportunity to design biological products capable of binding to personalized targets\textsuperscript{10,12}, such as mAbs. Originally biopharmaceuticals were obtained as purified extracts from the host (i.e., the living organism producing the biopharmaceutical\textsuperscript{143}) organism's blood or tissue\textsuperscript{132, 134, 137, 155, 156}. Although this strategy was adequate for obtaining proteins it resulted in low yields\textsuperscript{157} that could not support the growing demand of the biopharmaceutical industry and thus patient access to these therapies was limited\textsuperscript{132, 157}.

Biotechnological advances\textsuperscript{158}, such as the advent of recombinant DNA (rDNA)\textsuperscript{10, 12, 113, 127, 133, 136, 138, 143, 157, 159-162} (rDNA is DNA from two or more sources that have been artificially recombined\textsuperscript{163} – an important advancement to the biotechnology industry because it allowed for the expression (and amplified production) of therapeutic proteins in organisms that would otherwise not naturally produce them\textsuperscript{132, 157, 161}), hybridoma\textsuperscript{113, 137, 151}, gene expression\textsuperscript{9,127}, and monoclonal antibody technologies\textsuperscript{9, 10, 127, 143} enabled scientists to specifically engineer a variety of hosts\textsuperscript{1} amendable to genetic manipulation\textsuperscript{11} for the production of human proteins\textsuperscript{9, 131, 164}. These technologies also afforded for much simpler and faster\textsuperscript{11} protein cultivation, resulting in economically favorable\textsuperscript{11} amplified production of therapeutic proteins on a commercially viable scale\textsuperscript{71, 159, 165}. The use of recombinant hosts to generate biopharmaceuticals also significantly improved the quality of the proteins obtained\textsuperscript{155}. 
The first recombinant biopharmaceutical to receive FDA approval for human use was Humulin® (Genentech, 1982), a human insulin clone extracted from bacteria (Escherichia coli) for the treatment of diabetes. In 1987 the first rDNA biopharmaceutical produced in mammalian cells, Activase® (Genentech, 1987), received FDA approval. Activase® is a tissue plasminogen activator (t-PA) used to treat heart attacks and pulmonary embolism and is biosynthesized in Chinese hamster ovary (CHO) cells. Biotechnological advances made since these early products received regulatory approval, including the development of more sophisticated characterization and analytical methodologies, improvements in bioprocess design, the establishment of new hosts (expression systems), and advances in genomics and proteomics have fueled rapid growth, evolution, and success for the biopharmaceutical industry. These advances led to a wave of highly successful first generation biopharmaceuticals, including many blockbusters, in the early 1990s.

Biopharmaceuticals now represent the fastest growing segment in the pharmaceutical industry with impressive worldwide growth in an ever widening span of therapeutic fields. In 2009 worldwide sales of biopharmaceuticals exceeded US$92 billion. These sales are expected to continually increase, at least twice as fast as those of small-molecule pharmaceuticals, to US$167 billion by 2015. The number of FDA approvals of biopharmaceuticals has also been rising and in 2002 the number of approved biopharmaceuticals exceeded that of small-molecule pharmaceuticals.
approvals for the first time. Currently there are over 400 biopharmaceuticals in late clinical development.

Despite this robust market growth, patient access to biopharmaceutical therapies has been limited by their high unit costs which average 22 times more than small-molecule pharmaceuticals. Herceptin® (trastuzumab) and Humira® (adalimumab), as examples, annual costs average US$37,000 and US$50,000, respectively. Exacerbating high annual costs is the fact that biopharmaceuticals are frequently used to treat chronic conditions requiring on-going or lifetime treatment. The cost of Cerezyme® (imiglucerase), for example, averages US$200,000 annually for life.

These astounding costs have led to an unmet clinical demand for biopharmaceuticals worldwide ((Insert footnote here: The unmet demand for biopharmaceuticals has shown to be problematic in the past, as evident by the 2001 shortage of Enbrel®, a recombinant protein (tumor necrosis factor-alpha, TNF-alpha) commercialized by Immunex (Immunex was later acquired by Amgen in 2001) in 1998 for the treatment of rheumatoid arthritis. Nearly 1.3 million patients in the United States are affected by rheumatoid arthritis. Enbrel® was so successful that by March of 2002 a waiting list of 13,000 patients (existed), causing Immunex to ration Enbrel® to pharmacies until the patient demand could be met. Experience, such as that evidenced by the Enbrel® shortage, demonstrate the need for improved/adequate biomanufacturing capacity that can support the growing demand for recombinant protein drugs.)) and can be
attributed to the significant financial burden\textsuperscript{12, 179} and high barriers of entry characteristic of biopharmaceutical development\textsuperscript{123, 126, 171}. The establishment of host cell lines for biopharmaceutical production can be a tedious, time consuming, and costly task\textsuperscript{10, 159, 169, 172}. Even after an expression system has been developed and optimized, biopharmaceutical production costs are high\textsuperscript{9, 159, 172}. Production protocols can easily include hundreds of critical steps in fermentation alone\textsuperscript{135}, can be costly, and typically require laborious maintenance\textsuperscript{135}. In 2006, for example, the cost to produce a monoclonal antibody was estimated to be US$300-3000 per gram\textsuperscript{180}. Expression systems are also extremely sensitive to changes in their environment, and even minimal “tweaking” of production processes has been shown to cause dramatic changes in the structure and quality of the proteins produced. Since biopharmaceuticals are difficult to fully characterize with the techniques currently available the quality of final products are typically concomitantly established via strict regulation of their manufacturing protocols as well as with the analytical tools available\textsuperscript{135}. Following regulatory approval, biopharmaceuticals are held under strict regulatory supervision and stringent quality control mechanisms to assure these protocols do not change, or experience genetic drift, over time\textsuperscript{123, 126, 128, 143, 153, 165, 173, 181, 182}. These complexities characteristic of biopharmaceutical development quickly add up. It costs an estimated US$1.2 - 1.3 billion to bring an innovator biopharmaceutical to market\textsuperscript{175}. Given the extensive time and financial investments involved in bringing a biopharmaceutical to market most biopharmaceutical companies will patent, and never fully publically disclose\textsuperscript{128}, not only their final product but also the process by which it is produced\textsuperscript{9, 71, 154}.
Hope of alleviating the high costs of these treatment lies in the upcoming expiration of many biopharmaceutical patents. Since patents covering newly approved drugs generally last 20 years from the time of filling a submission many of the commercially successful first generation biopharmaceuticals developed during the 1990s have recently gone off patent. Patents on more complex biopharmaceuticals worth an estimated $US64 billion will also be expiring within the coming five years. This "patent cliff" has opened the biopharmaceutical industry to the unprecedented opportunity for generic competition.

The anticipated generic competition will come in the form of biosimilars, biopharmaceutical products that have been determined, via comprehensive comparability studies, to be “highly similar” to or “interchangeable” with an already FDA-approved biological product. (So-called “generic biologics” have several names in the literature. They were previously referred to as “follow-on” products in the United States but are now considered “biosimilars.” They are also spoken of as "biosimilars" in Europe, "subsequent entry biologics" in Canada, "biocomparables" in Mexico, and "follow-on pharmaceuticals" in Japan. For consistency they will be referred to as biosimilars throughout this review. A biosimilar is defined in The Patient Protection and Affordable Care Act as “a biological product approved under an abbreviated application for a license of a biological product that relies in part on data or information in an application for another biological product licensed under section 351 of the Public Health Service Act.” Emphasis should be placed on the fact that unlike generic small-molecule pharmaceuticals biosimilar products are similar, and not
identical, to innovator products\textsuperscript{9, 10, 127, 171}. An identical biopharmaceutical product would be considered a "biogeneric," not to be confused with a biosimilar.

Competition brought about by biosimilars is expected to drive down biopharmaceutical prices\textsuperscript{188} and will likely increase patient access to them\textsuperscript{31, 123, 143}. The US government also stands to benefit from reduced biopharmaceutical costs. For example, the Medicare Payment Advisory Commission recently reported that 43\% of the entire Medicare Part B drug budget is expended on payments for only six biopharmaceuticals\textsuperscript{12, 142, 175, 189}.

“Generic” versions of these six could lead to much needed savings. While the estimated savings are debated\textsuperscript{71, 123, 143, 149, 181}, they are generally expected to fall below those generated from generic small-molecule pharmaceuticals\textsuperscript{139}. The reduction in savings is credited to the complexity involved in bringing a biopharmaceutical to market, which will limit the number of biosimilar entrants in competition\textsuperscript{13, 149}. Regardless, generic pharmaceuticals saved an estimated US$931 billion in the past ten years\textsuperscript{190} alone. Even a fraction of these savings would be socially and politically beneficial. A governmental push towards biosimilars is evident in the\textsuperscript{175} recent focus of managing the costs of healthcare\textsuperscript{10, 12, 123, 124, 127, 131, 136, 143, 159, 172, 186, 191, 192}. The Biologics Price Competition and Innovation (BPCI) Act of 2009, for example, was aimed at improving patient access to biopharmaceuticals\textsuperscript{139} by reducing the cost of these therapies\textsuperscript{124, 126} as well as the time it takes to get them approved and to market\textsuperscript{175}.

Considering the high clinical utility of biopharmaceuticals, the evolving establishment of refined techniques for their development, and the desire for biosimilar products from
patients and the government alike, it is not surprising that a notable shift in the pharmaceutical industry is underway. Once dominated by the development of small-molecule pharmaceuticals\textsuperscript{11, 129}, a number of companies are now shifting their focus to the research and development (R&D) of biopharmaceuticals instead\textsuperscript{11-13}. Evidence of pharmaceutical companies transitioning to become bigger players in the biopharmaceutical arena\textsuperscript{13} can be found in many of the mergers and acquisitions that have recently taken place. Pfizer's acquisition of Wyeth (October 2009), as an example, was their effort to "shift toward becoming a broader, more competitive biopharma company\textsuperscript{13}.” A number of large pharmaceutical firms have also invested substantial resources\textsuperscript{159} specifically into the development of biosimilars. This is their effort to strategically place themselves in the growing global biosimilar market\textsuperscript{65, 129, 154}, which is estimated to grow to $US19.4 billion by the year 2014\textsuperscript{159, 193}. “The world’s largest generics manufacturer\textsuperscript{159},” Teva, for example, has partnered with the Lonza Group to begin making and selling biosimilars\textsuperscript{129, 159}. Sandoz, the generic pharmaceuticals division of Novartis\textsuperscript{194}, has increased its biomanufacturing capacity in anticipation of biosimilar efforts\textsuperscript{129, 159}. Furthermore, in 2008 Merck & Co. dedicated an entire division (Merck BioVentures) to biosimilar development\textsuperscript{129, 149, 159}. Other firms making moves to increase their biosimilar pipelines include Roche, Novo Nordisk, Amgen, Biogen Idec\textsuperscript{159} and Samsung\textsuperscript{126, 143, 195}. 
4.2 Biopharmaceutical comparison to small-molecule pharmaceuticals

Given the many stakeholders who stand to be affected by biosimilar development and regulation these are currently topics of active debate\(^3\). Their production in living hosts\(^{123, 127, 134, 137}\), characteristic complex structures, intensive development, and potential to invoke immunogenic reactions make biopharmaceuticals particularly complex relative to small-molecule pharmaceuticals.

When compared to small-molecule pharmaceuticals, biopharmaceuticals are typically much larger and possess great structural variation and complexity\(^9, 65, 123, 124, 127, 130, 135, 139, 159, 160, 172\). Small-molecule pharmaceuticals have defined molecular weights in the range of 50-1000 Da\(^{172}\), compared to biopharmaceuticals which have molecular weights ranging from 5-200 kDa\(^{127, 172}\). The active ingredient in small-molecules is generally a single molecular entity\(^{137, 175}\) that can be identified and completely characterized\(^{65, 123, 135, 160}\). Biopharmaceuticals, on the other hand, have heterogeneous\(^{123, 172}\) active substances made of a collection of protein isoforms\(^{137}\) that are usually not possible to fully characterize with the analytical tools currently available\(^{127, 135, 137, 139, 159, 175, 185}\).

The multifaceted biopharmaceutical manufacturing processes add to their complexity and further distinguish them from small-molecule pharmaceuticals. Biopharmaceutical manufacturing can be several orders of magnitude more complicated\(^{137}\) and result in costs of development easily 100-fold those of small-molecule pharmaceuticals\(^{127}\). A great deal of care must be taken to insure biopharmaceutical manufacturing processes are well
understood and reproducible because they are extremely sensitive to the environment in which they are produced\textsuperscript{123, 127, 128}. The industry-accepted mantra "the process defines the product\textsuperscript{71, 134, 154}" cannot be overstated when speaking of biopharmaceuticals, as even minor changes can alter the final product attributes\textsuperscript{12}. The affects environmental changes will have on the final protein are very difficult to predict and may possess clinical consequences\textsuperscript{10, 137, 152} such as altering a biopharmaceutical’s biological activity\textsuperscript{12, 135, 175}, efficacy, and even clinical safety\textsuperscript{9, 12, 128, 137, 143, 156}.

Perhaps the most critical distinction between biopharmaceuticals and small-molecule pharmaceuticals is the former’s increased risk for inducing immunogenic reactions\textsuperscript{9, 127, 137, 139}. Briefly, immunogenicity is the ability of a substance to induce an immune response\textsuperscript{196}. Since biopharmaceuticals are biologically active molecules\textsuperscript{159} each has the potential to affect nearly 100 physiological processes in patients\textsuperscript{123}, a dramatic increase over the "mere handful of reactions\textsuperscript{123}" of a small-molecule pharmaceutical\textsuperscript{123}. Biopharmaceuticals are therefore more prone than small-molecule pharmaceuticals to induce an immune response, either acute or chronic, when administered to patients\textsuperscript{137, 159}. Examples of immune reactions include hypersensitivity, anaphylaxis, pseudo-allergic/anaphylactoid, serum sickness, infusion reactions, decreased therapeutic efficacy\textsuperscript{197}, the generation of antibodies against the therapeutic, and cross reactions between antibodies to the therapeutic and endogenous proteins\textsuperscript{137, 196}. Complicating this issue is biopharmaceuticals that are highly similar in structure can invoke remarkably different immunogenic reactions\textsuperscript{10, 127, 137}.
The likelihood and severity of an immunogenic reaction is influenced by the biopharmaceutical product, the patient receiving treatment, or a mixture of both of these. Product related causes include impurities or contaminants in the biopharmaceutical (such as host-cell proteins, leachables, and extractables), variation in protein structure from manufacturing changes (such as in amino acid sequence or PTMs), the formulation of the biopharmaceutical, or improper storage and handling following its manufacture. Patient-related factors include the patient's age, genetics (human leukocyte antigen (HLA) expression), immune status at the time of treatment, concomitant medication, disease state, previous exposure to biological agents, and the biopharmaceutical route of administration. (Intravenous administration is generally less immunogenic than intramuscular or subcutaneous administration of biopharmaceuticals.) Biopharmaceuticals are therefore most commonly administered in this way. The frequency of treatment and the length of time treatment has been administered can also influence the immunogenicity of a biopharmaceutical. The longer and more frequent treatment is administered the more likely they are to induce an immune response. None of these factors, however, can be isolated as definite "predictors of immunogenicity in humans," and predicting immunogenicity remains a difficult task that will likely be a consideration of the FDA as they shape the biosimilar approval pathway.
4.3 Regulatory Considerations

The FDA regulates small-molecule pharmaceuticals under the Food, Drug, and Cosmetics Act (FDCA), and generics are approved following the guidelines outlined by the Hatch-Waxman Act of 1984\textsuperscript{123, 139, 159, 175}. This act, an amendment to the FDCA intended to spur generic competition\textsuperscript{123}, outlines an abbreviated approval pathway in which generic sponsors must only demonstrate “bioequivalence\textsuperscript{123, 175},” with innovator products\textsuperscript{9, 65, 124, 143} to receive approval\textsuperscript{183, 200}. This effectively streamlines the approval process by eliminating the need for the full clinical trials necessary to bring a first generation pharmaceutical to market\textsuperscript{12, 127, 128, 137, 139, 143, 153, 159, 172, 175}. The FDA reasons that bioequivalent products will share the same properties in terms of safety and efficacy and therefore can be considered fully interchangeable with the innovator product\textsuperscript{9, 123, 137, 159, 175}. Bioequivalence can be (relatively) easily established for small-molecule pharmaceuticals because their active substances are simple, can be replicated exactly, are unequivocally characterizable, and are not nearly as sensitive to environmental changes as biopharmaceuticals\textsuperscript{9, 124 135}.

The Hatch-Waxman Act has been a great success and led to dramatic cost savings\textsuperscript{139}. Generic pharmaceuticals have saved an estimated US$931 billion over the past ten years\textsuperscript{190} alone. Despite this, nearly thirty years after it was passed, no similar pathway exists for biopharmaceuticals\textsuperscript{123}. Most biopharmaceuticals are regulated under the Public Health Service Act (PHSA)\textsuperscript{12, 123} and not the FDCA, thus the Hatch-Waxman Act did not automatically extend to them\textsuperscript{201}. 
The Biologics Price Competition and Innovation Act of 2009 (BPCIA), part of the Patient Protection and Affordable Care Act (PPACA), tasked the FDA with developing an abbreviated pathway for biosimilars. The Congressional Budget Office has estimated that such a pathway could save the federal government US$9-12 billion over a ten year span. Due to the complexity particular to biopharmaceuticals, as outlined above, developing and implementing this pathway will be quite the arduous task. The mode of action of biopharmaceuticals is not always well understood so its safety profile is dependent on its manufacturing process and composition. Biosimilars manufactured under different processes therefore cannot automatically be assumed to have the same safety and efficacy profile as innovator products. Furthermore, given the sensitivity of biopharmaceuticals to their environment and the fact that these “environments” are typically proprietary information it is unlikely that a biosimilar developer will be able to replicate an innovator biopharmaceutical exactly. Environmental variation in the production process may include, but is certainly not limited to, alterations in primary structure, expression system, fermentation, purification, DNA vectors, cell culture, and formulations (such as particles for inhalation, adjuvants, buffers, packaging and delivery systems, or the use of stabilizers such as albumin). Inevitably, biosimilars will differ from innovator products to some extent. This, along with their complicated structures, inevitable heterogeneity, and the current inability to fully characterize biopharmaceuticals, make it impossible to demonstrate bioequivalence as can be done for generic small-molecule pharmaceuticals.
Since the Hatch Waxman Act is not an applicable model for the abbreviated biosimilar pathway, some are looking towards the way in which the FDA has handled manufacturing changes to already approved biopharmaceuticals as an alternative model. During the successful wave of biopharmaceuticals in the 1990s many companies had to scale-up their production in order to keep up with the demand\textsuperscript{12} for these products. To assess whether the scaled-up manufacturing protocols induced changes in the products the FDA set forth “comparability” guidelines in their 1996 “Demonstration of Comparability of Human Biological Products, including Therapeutic Biotechnology-derived Products\textsuperscript{12}.” Under these guidelines a head-to-head product comparability exercise must be undertaken to demonstrate that the biopharmaceutical produced under the new manufacturing process can be considered "comparable\textsuperscript{12}”, in terms of safety and efficacy\textsuperscript{128, 202}, to the original product\textsuperscript{71, 124, 139, 203}. In order to do so innovators must only conduct limited clinical testing\textsuperscript{12, 203} to prove the scaled-up products can confidently be considered comparable or identical\textsuperscript{128} to the original. Some members of the generics industry have pointed out the similarities between these guidelines and the Hatch-Waxman Act. They argue that biopharmaceutical manufacturers are essentially making biosimilars of their own products when the manufacturing process is changed\textsuperscript{203}. Perhaps the “Demonstration of Comparability of Human Biological Products, including Therapeutic Biotechnology-derived Products” guidelines will provide a better model for how an abbreviated biosimilar pathway could take shape?
An illustrative example depicting the use of these guidelines and the impact of manufacturing changes on product attributes can be seen in the way the FDA handled the approval of a scaled-up version of the biopharmaceutical Myozyme (alglucosidase-α)\textsuperscript{129}. Even though it was produced following an otherwise identical process, the FDA required the scaled-up version be approved as a separate biopharmaceutical with a different name (Lumizyme) because the scaling procedure caused slight chemical variations between the original and scaled-up biopharmaceuticals\textsuperscript{129, 143, 196, 204}.

Until the BPCIA all biologics had to receive approval under the guidelines outlined in the Biologics License Application (BLA) of the PHSA, even if they were technically biosimilar products\textsuperscript{123, 175}. The BLA pathway requires full-scale testing and clinical trials comparable those necessary for approval of new small-molecule pharmaceuticals\textsuperscript{123, 175}. Due to uncertainty in the industry as to how the FDA will move forward with an abbreviated approval pathway many developers have opted to apply for biosimilar approval under the existing BLA pathway\textsuperscript{126, 175, 202, 203}. Their decision to do so highlights the importance that the FDA develop and release a biosimilar abbreviated pathway as soon as possible. The continued postponement of such a pathway will undoubtedly stifle the cost savings possible from biosimilars\textsuperscript{175}. Furthermore, it will leave little incentive for biotechnology firms to innovate biosimilar products if they will be subject to the same costly approval requirements as innovator products\textsuperscript{123}. The FDA must therefore strike an intricate balance of inciting scientific innovation while protecting patient safety and managing the rising costs of healthcare\textsuperscript{175, 205}. 
In order to develop the abbreviated pathway the FDA must critically examine (i) how “bio-similar” is similar enough, (ii) what type(s) of differences are clinically significant and what type(s) aren’t, (iii) what preclinical and clinical trials will be needed/sufficient to demonstrate safety, efficacy, similarity, and/or interchangeability, (iv) to what extent will data on safety and efficacy from clinical trials on innovator products be applied to biosimilars, and (v) what control mechanisms must be in place to ensure consistency/reliability within the manufacturing of the biosimilar product. The issue of clinical trials is of particular gravity since it seems, for now at least, clinical trials may be inherent to assuring safety of biopharmaceuticals. This is especially true considering their ability to unpredictably provoke immunogenic reactions.

4.4 Immunogenicity & Post-Translational Modifications

Central to the risk of immunogenicity are post-translational modifications (PTMs), covalent modifications to a protein that occur during the final stage of protein synthesis. (Covalent modification of eukaryotic proteins can occur during (co-translational) or after (post-translational) the protein is synthesized. For clarity, these will both be referred to as post-translational modifications throughout this review.) PTMs aid the nascent protein, a polypeptide chain, in folding into is biologically active secondary, tertiary, and quaternary structures.
More than 300 types of PTMs have been identified\textsuperscript{130, 150} including glycosylation, acetylation, acylation, phosphorylation, sulfonation, proteolytic processing, carboxylation, hydroxylation, amidation, and hypermannosylation to name a few\textsuperscript{130, 138, 152, 153, 172, 196, 207, 209, 210}. These manifest from both time- and signal- dependent factors and vary depending on developmental state, location, stress, disease state, and cellular environment\textsuperscript{11, 150}. The extent of PTM a protein undergoes is also subject to great diversity. A protein can undergo relatively little PTM with only simple chemical alterations\textsuperscript{150} or large and highly elaborate modifications that end up being bigger than the unmodified protein itself\textsuperscript{150}.

PTMs play a role in nearly all protein activities\textsuperscript{150} including regulation of gene expression\textsuperscript{208}, tissue targeting, ligand recognition, stability, solubility, immunogenicity, and regulation of protein serum half-life\textsuperscript{11, 113, 130, 138, 150, 156, 165, 174, 206, 211}. Two proteins built from identical amino acid sequences can therefore exhibit vastly different final structures and functions\textsuperscript{123}. Although all of this diversity enables "an efficient and cost-effective mechanism for the exponential diversification of the genome\textsuperscript{150}" by expanding the proteome\textsuperscript{169}, PTMs introduce a great deal of complexity in biopharmaceutical development\textsuperscript{150}. Most biopharmaceuticals exhibit a combination of at least two or more PTMs\textsuperscript{130, 147} that are essential to its therapeutic function\textsuperscript{150, 202, 206}. If these are not correctly assembled the biopharmaceutical may lose efficacy\textsuperscript{127, 176, 202, 209, 212} or cause immunogenicity. Even "miniscule" variations in the PTM profile have been shown to cause severe immunogenic ramifications\textsuperscript{137, 143, 156, 159, 202}. 
Most biopharmaceuticals are known to induce an immune response of some sort\textsuperscript{124, 127}. However these range in severity and many are not considered clinically significant\textsuperscript{124, 127, 135, 137, 176, 196, 207}. Regardless, a biopharmaceutical’s “antigenic potential\textsuperscript{9,6}” is impossible to fully predict given the methods available currently\textsuperscript{11, 127, 137, 172, 179, 196}. Once developed and clinically approved a biopharmaceutical’s PTM profile must therefore remain consistent.

Since proteins that differ only in their PTM profile may possess different final structures, functions, and, therefore, clinical utility, they are subject to separate intellectual property rights\textsuperscript{140, 153}. PTMs are thus a primary consideration for biosimilar developers\textsuperscript{130, 164, 211}. A variety of different factors can affect the type, location, and extent of PTMs a protein expresses\textsuperscript{169}. Both up- and downstream conditions of the manufacturing process can influence this\textsuperscript{169}. In brief, upstream conditions such as temperature, oxygen levels, pH, culture medium, fermentation\textsuperscript{196} and feeling strategies\textsuperscript{206, 207, 213} have been shown to affect PTM profile. Downstream conditions\textsuperscript{169} including purification\textsuperscript{207}, final product formulation\textsuperscript{206}, and storage and handling\textsuperscript{196} may as well. The host organism, however, bears the greatest influence on PTMs and dictates what kind of PTMs, if any, the product will express\textsuperscript{196}. Selecting a host capable of generating an appropriate PTM profile is therefore by far the most important consideration in biosimilar development\textsuperscript{11, 157, 169, 173, 206}. 
4.4.1 Glycosylation

Of the PTMs considered during biopharmaceutical development, perhaps the most pressing is glycosylation. An estimated 70% of biopharmaceuticals currently in drug development are glycosylated proteins\textsuperscript{140, 168}. Glycosylation is therefore a critical concern for biopharmaceutical and biosimilar developers\textsuperscript{140, 169, 209} and is thus the PTM of primary focus in this review.

Protein glycosylation is a co- and post-translational modification in which oligosaccharide\textsuperscript{140, 169}, also called "glycan\textsuperscript{11, 214}," groups are added to the protein polypeptide backbone\textsuperscript{11}. It is the most widespread\textsuperscript{169, 196} and complicated\textsuperscript{140} PTM observed in both natural proteins and biopharmaceuticals\textsuperscript{11, 130, 140, 207}, as well as the most common PTM of eukaryotic cells\textsuperscript{150}. An estimated 50-70\%\textsuperscript{150, 169, 196, 207} of human proteins are glycosylated in an incredible variety of subtly different forms\textsuperscript{140, 150, 214}. Defects in glycosylation pathways\textsuperscript{169} are known to be associated with many, over 30\textsuperscript{169}, diseases such as cancer\textsuperscript{11} and rheumatoid arthritis\textsuperscript{140, 169}, and are used as disease markers for such\textsuperscript{169}. Although there are actually five types of glycosylation (N-, O-, P-, C-, and G-linked glycosylation, listed in relative order of importance), the two most commonly mentioned in the literature are N- and O-glycosylation\textsuperscript{157}. The addition of glycan groups to the protein core is a common feature of all of the above mentioned types of glycosylation but they differ in their respective glycan attachment sites\textsuperscript{157}. N-glycosylation involves the covalent addition of glycans to the nitrogen of Asparagine and Arginine side chains\textsuperscript{150, 157, 215}. It is much more common and better understood than the
other types of glycosylation, and thus will be the main focus of glycosylation in this review\textsuperscript{215,216}. \textit{O}-linked glycosylation involves the covalent attachment of glycans to the hydroxyl oxygen of Serine and Threonine side chains\textsuperscript{150,216}. Its biological role is not currently well understood\textsuperscript{216}.

Glycosylation is not directly coded for in the genome, i.e. "template driven." Unlike other cellular macromolecules which are synthesized in a template driven fashion (e.g. proteins, a string of amino acids coded for by nucleic acids in the DNA), the decision of which and where glycan groups will be added to a glycoprotein is ultimately decided by a complex array of enzymes including glycosyltransferases and glycosidases present in the cell\textsuperscript{176}. Enzymatic coupling of glycan groups occurs in either the endoplasmic reticulum (ER) or the Golgi apparatus\textsuperscript{11}. The coupled glycans are further modified by complicated multi-enzyme pathways\textsuperscript{11,217}, sometimes involving two to three dozen different enzymes\textsuperscript{130,169}. The same protein is capable of expressing diverse variation of core and outer-arm\textsuperscript{196} glycan structures (microheterogeneity)\textsuperscript{140,176,218}, to any given extent, at any given glycosylation site\textsuperscript{144} (macroheterogeneity)\textsuperscript{140,168,176}. Glycoproteins are therefore typically biosynthesized as a heterogeneous mixture of glycoforms\textsuperscript{169,176,219}, isoforms differing in their glycosylation profiles\textsuperscript{140,144,169}. These glycoforms can vary in activity\textsuperscript{219} and each possess its own pharmacokinetic, pharmacodynamics, and efficacy profile\textsuperscript{169,216,219}. Glycosylation is extremely dependent on the environmental conditions of the cell during protein synthesis\textsuperscript{11}. Examples of environmental conditions affecting glycoform expression are the culture media, mode of culture, specific protein being synthesized\textsuperscript{209}, host organism, and physiological conditions such as disease state\textsuperscript{140,214}.
Although glycosylation occurring in ER has been shown to be highly conserved in both lower and higher eukaryotes\textsuperscript{140, 168, 169, 216, 220}, the Golgi modifications are species\textsuperscript{140, 169}, tissue, cell type\textsuperscript{220}, and, sometimes even gender specific\textsuperscript{150, 162, 176, 212, 214}. The repertoire of glycan modifications exhibited by bacteria, yeast, fungi, plants, and mammalian cells are very different\textsuperscript{169, 214}. If non-human glycans are incorporated into biopharmaceuticals undesired, and potentially deadly, immunogenic reactions can result\textsuperscript{156, 169, 176, 214}. Biopharmaceuticals expressing foreign glycans are more prone to elicit an immunogenic reaction because the foreign epitopes are more likely to be recognized as "non-self" by the immune system\textsuperscript{196}. Even biopharmaceuticals bearing completely humanized glycans can elicit an immune response in patients who do not naturally produce that protein in sufficient concentrations\textsuperscript{196}. The cumulative complexity of incredible variation in glycosylation patterns and the unpredictable immunogenic responses these can provoke create a particularly complex obstacle for developers to navigate\textsuperscript{165, 166, 196}.

Although glycosylation has complicated biopharmaceutical development, it also offers the opportunity to create biopharmaceuticals with enhanced therapeutic utility\textsuperscript{132, 150, 156, 196, 219}. Developers look to selectively exclude unwanted PTMs while enhancing desired PTMs via expression system and glycoengineering\textsuperscript{130, 140, 150, 165, 184, 196}.
4.4.2 Glycoengineering

Glycosylation engineering, or glycoengineering, is the strategic alterations of where and what glycans are expressed on a protein\textsuperscript{140, 217, 220}. The two primary goals of glycoengineering in biopharmaceutical development are (i) engineer a host organism able to produce proteins decorated with “human” glycans and (ii) ensure this glycosylation is as reproducible and homogeneous as possible\textsuperscript{140, 150, 169, 196}. Although these goals seem straightforward, alteration of glycosylation can be quite complex\textsuperscript{166} and requires a comprehensive understanding of the glycosylation reactions of that particular organism (their “glycome\textsuperscript{219}”). Our ability to effectively glycoengineer hosts is therefore limited by our understanding of glycosylation in the host itself\textsuperscript{166}.

Early efforts at glycoengineering were generally more “extracellular” in approach. Modification of the culture conditions such as pH, temperature, and the concentrations of oxygen, ammonia, and glucose\textsuperscript{136, 144, 169, 207, 218} have been attempted for this purpose. Other "extracellular" strategies include the addition of "post-synthesis\textsuperscript{206}" engineering steps, such as \textit{in vitro}\textsuperscript{176, 218} glycan alteration via downstream enzyme treatment\textsuperscript{130, 207, 218}. This frequently involves the covalent attachment of one or more poly(ethylene glycol) (PEG) molecules to the protein backbone\textsuperscript{144, 206, 207, 219}, “PEGylation”\textsuperscript{113, 138, 144, 196}.

Glycoengineering has also been accomplished via direct chemical conjugation of synthetic oligosaccharides to the protein backbone itself\textsuperscript{130}, as well as the incorporation of additional glycosylation consensus sites into the protein to promote desired hyperglycosylation\textsuperscript{130, 144, 156}.
More recent approaches to glycoengineering have shifted focus to “intracellular” metabolic engineering of the host itself\textsuperscript{133, 161, 166, 173, 221}. With this approach the host is capable of biosynthesizing the optimized glycoprofile directly into the protein product\textsuperscript{130}, in theory eliminating the need for additional downstream glycoengineering steps. Intracellular glycoengineering is generally accomplished via elimination of undesired host glycosylation pathways and concomitant introduction of desired glycosylation reactions\textsuperscript{166, 173, 216, 219}. Due to regulatory considerations, developers will generally attempt extracellular approaches to glycoengineering before moving to the intracellular strategies\textsuperscript{71, 169}.

Examples of enhanced characteristics that can be acquired through glycoengineering include increased solubility, stability\textsuperscript{144}, potency, half-life\textsuperscript{207, 10, 11, 130}, and tissue distribution\textsuperscript{166, 184}, as well as reduced immunogenicity\textsuperscript{144, 207}, protease degradation\textsuperscript{144, 207}, and aggregation\textsuperscript{196}. These enhanced properties have been associated with reducing the necessary biopharmaceutical frequency and dose of administration\textsuperscript{144, 156, 206}. Functional glycomics is an area of research that aims at understanding the functions of glycans and their influence on human diseases\textsuperscript{72, 176} and has aided in the development of biopharmaceuticals with such enhanced properties.
4.4.3 Post-translational Modifications in Biosimilar Development

Since proteins that differ only in their PTM profile may possess different final structures, functions, and, therefore, clinical utility, they are subject to separate intellectual property rights\textsuperscript{140, 153}. PTMs are thus a primary consideration for biosimilar developers\textsuperscript{130, 164, 211}.

A variety of different factors can affect the type, location, and extent of PTMs a protein expresses\textsuperscript{169}. Both up- and downstream conditions of the manufacturing process can influence this\textsuperscript{169}. In brief, upstream conditions such as temperature, oxygen levels, pH, culture medium, fermentation\textsuperscript{196} and feeling strategies\textsuperscript{206, 207, 213} have been shown to affect PTM profile. Downstream conditions\textsuperscript{169} including purification\textsuperscript{207}, final product formulation\textsuperscript{206}, and storage and handling\textsuperscript{196} may as well.

The host organism, however, bears the greatest influence on PTMs and dictates what kind of PTMs, if any, the product will express\textsuperscript{196}. Selecting a host capable of generating an appropriate PTM profile is therefore by far the most important consideration in biosimilar development\textsuperscript{11, 140, 157, 164, 169, 173, 196, 206, 207, 222}.

4.5 Host Selection

The selection and optimization of a suitable host organism must be considered separately for each individual biopharmaceutical due to their characteristic complexities\textsuperscript{164, 207}. No
one system should be considered superior to any another, as each host organism’s
applicability to the biosynthesis of a biopharmaceutical is entirely dependent on the traits
of that biopharmaceutical itself\textsuperscript{157, 164, 207}.

An ideal host organism would be able to rapidly grow on inexpensive media, have well
understood genetics and undergo facile genetic manipulation, produce high yields of
correctly (human) PTM proteins that are easy to extract and purify, and be accepted as
safe by regulatory agencies\textsuperscript{157}. There is currently no one host that possesses all of these
characteristics so biopharmaceutical developers must strike an intricate balance between
product, production, and regulatory considerations. This balancing act is largely a process
of trial and error\textsuperscript{223}, in which many hosts may be tried before the optimal selection can be
identified\textsuperscript{158}.

The desired biopharmaceutical product in development holds a large influence over
which host organism will be appropriate for its biosynthesis\textsuperscript{219}. Factors such as (i)
intended therapeutic effect\textsuperscript{136, 169}, (ii) desired size\textsuperscript{136}, structure , and function\textsuperscript{132, 133, 136, 169},
(iii) PTMs\textsuperscript{136, 158}, (iv) destination and route of administration\textsuperscript{136, 158}, (v) quality \textsuperscript{132, 133, 161, 164}, and (vi) quantity\textsuperscript{158, 161, 164} of the product are all important considerations. These must,
however, be balanced by production considerations such as (i) familiarity ("genetic
know-how and tools\textsuperscript{164"}), (ii) safety, (iii) cost, (iv) convenience\textsuperscript{161}, (v) yield\textsuperscript{133}, (vii)
production speed\textsuperscript{132}, (vii) and production space needed\textsuperscript{161, 164}. Regulatory considerations
include the acceptance of a particular host organism by approval agencies, which may
accelerate or delay approval of the biopharmaceutical\textsuperscript{224}. 

193
Several prokaryotic and eukaryotic host organisms have already been developed into high producing expression systems. Broadly, these include cell cultures of viruses, bacteria, yeast, molds, plants, insects, and mammals. Higher eukaryotic organisms such as transgenic plants and animals can also be utilized as hosts.

Although each host must be selected on a case-by-case basis, a few trends in host selection have been noted. Microbial systems, Escherichia coli (bacterial) and Saccharomyces Cerevisiae (yeast) in particular, are most commonly employed for the production of biopharmaceuticals treating infectious, endocrine, nutritional, and metabolic diseases. Insect cells, filamentous fungi, and plant systems have primarily been used for the production of vaccines, although they are not limited to this and have been used to produce glycoproteins as well. Larger proteins (> 100 kD) tend to be best suited to biosynthesis in eukaryotic systems whereas smaller proteins (< 30 kD) tend to be expressed in prokaryotic systems well. Bacterial, mammalian, and yeast cell lines have historically experienced the most abundant usage. It is estimated that approximately 90% of biological products are currently produced in one of these organisms.

Due to their influence over the protein products obtained from biosynthesis, the remainder of this review will be dedicated to examining some of the most common cellular hosts in the current literature. Advantages and disadvantages to each will be
discussed as well as historical successes, currently approved therapeutic applications, and engineering attempts to improve PTM.

4.5.1 Prokaryotic Expression Systems

**Bacteria**

Bacterial hosts are attractive options for biopharmaceutical developers. They are inexpensive, rapidly grow to high densities on a variety of substrates, have well characterized genetics, and have many commercially available cloning vectors and mutant host strains\(^{225}\).

One critical disadvantage to bacterial hosts is they are not able to carry out many PTMs\(^{155,161,168,225}\). Bacteria are therefore unable to biosynthesize complex proteins, such as monoclonal antibodies and coagulation blood factors, which require PTMs for their bioactivity and stability \emph{in vivo}\(^{155}\). Bacteria are also prone to improper protein folding and misassembly of protein subunits\(^{155}\). Their use is thus generally restricted to the biosynthesis of small recombinant proteins that are not glycosylated\(^{164}\).

**E. coli**

The gram-negative enterobacteria \emph{Escherichia coli}\(^{138}\) (\emph{E. coli}) were the first host organisms used for the expression of recombinant proteins\(^{158}\) in 1977\(^{210}\). They have since enjoyed the most extensive use of all the prokaryotes for rDNA protein expression\(^{132,157}\).
and are generally considered the "first choice microorganism" for this purpose. This can be attributed to the many advantageous characteristics for biosynthesis they possess, the fact that large-scale E. coli expression systems have already been established, and that many diverse E. coli expression vectors are commercially available. E. coli are most efficient for the production of soluble proteins smaller than 60 kDa and are commonly used for the expression of simple low molecular weight polypeptides with few or no PTMs.

E. coli systems hold many advantages over other hosts currently in use. Considering much of our understanding of microbial genetics was gleaned from studies on E. coli, their genetics and molecular biology are incredibly well characterized and much better understood than any other microorganism. E. coli are further advantageous because they offer high product yields and are cost effective. E. coli can be rapidly grown with expression in a mere matter of days, have rapid biomass accumulation, and express high cell density in continuous fermentations. They are relatively inexpensive, allow for cheap and easy culture, and simple process scale up. Furthermore, E. coli offer genetic flexibility and are amendable to glycoengineering.

Since they are prokaryotes, E. coli lack the cellular machinery needed for the complex PTMs of eukaryotic proteins, including glycosylation, phosphorylation, amidation, hydroxylation, proteolytic processing, and disulfide bond formation. Recombinant proteins expressed in E. coli therefore lack PTMs that
may be essential to a biopharmaceuticals therapeutic activity\textsuperscript{138, 210, 226}. Of particular consequence are the PTMs N-glycosylation\textsuperscript{226} and disulfide bond formation. Aglycosylated versions of glycoproteins\textsuperscript{169} and proteins missing disulfide bonds are prone to misfolding\textsuperscript{169, 210, 219}. This can lead to the formation of insoluble protein aggregates\textsuperscript{138, 158} referred to as “inclusion bodies” (IBs\textsuperscript{138, 158, 161, 169, 226}). Inclusion bodies contain not only the recombinant protein of interest but also biological contaminants such as chaperones, DNA, RNA, and lipids\textsuperscript{138, 210}.

Biopharmaceuticals produced as inclusion bodies can be insoluble, inactive, unstable, rapidly cleared from the blood\textsuperscript{132, 136, 168, 169}, and may require refolding\textsuperscript{132, 161}. Successful strategies for refolding of recombinant proteins from inclusion bodies have been developed. These, however, can be difficult, costly, may cause denaturing\textsuperscript{155, 157}, and are not always effective\textsuperscript{138}. Expression system engineering has been applied to \textit{E. coli} to alleviate this issue\textsuperscript{157, 158, 225, 227}. Briefly, optimization of the growth temperature\textsuperscript{132, 158}, media composition, promoter system\textsuperscript{132, 164}, plasmid copy number\textsuperscript{132, 138}, the rate of protein synthesis\textsuperscript{132}, and the co-expression of chaperones and/or foldases\textsuperscript{158, 164, 223} have been successful. Several commercially available engineered strains, such as the \textsuperscript{\textit{OrigamiTM}} Competent Cells (EMD-Novagen\textsuperscript{®})\textsuperscript{158, 228} and \textit{E. coli} SHuffle\textsuperscript{®} Competent Cells (New England BioLabs Inc.)\textsuperscript{229}, have also been developed. Finally, glycoengineering of \textit{E. coli} has been successful and includes the development of \textit{E. coli} strains possessing a functional N-glycosylation pathway\textsuperscript{161, 168}.
Codon bias presents another potential hurdle to biopharmaceutical production in *E. coli*. Different organisms possess biases in the frequencies specific codons appear in their genes\(^{138, 158}\). This is closely related to the abundance of specific tRNAs available in their cells\(^{138, 158, 225}\). For a recombinant protein to be produced in high quantities it must have access to the necessary tRNAs to facilitate translation of the desired protein\(^{138, 158, 225}\). A tRNA shortage could cause protein translation to stall\(^{225}\), premature termination of translation\(^{138, 158, 225}\), frameshifts in translation\(^{158, 225}\), or amino acid misincorporation\(^{138, 158, 225}\).

This can be overcome through expression system engineering of *E. coli* strains to include increased pools of the rare tRNAs\(^{158}\), codon optimization\(^{158, 223}\), or selection of a bacterial host that has more similar codon usage to humans\(^{223, 225}\). A commercially available host strain designed to alleviate problems associated with codon bias is Rosetta™ from EMD-Novagen®\(^{158, 210}\).

A final noteworthy disadvantage to *E. coli* as biopharmaceutical hosts is they produce lipopolysaccharides (LPS), also referred to as "endotoxins"\(^{132, 210, 225}\), which are pyrogenic to humans and other mammals. Recombinant proteins used as biopharmaceuticals therefore require extensive and thorough purification to remove any endotoxins prior to patient administration\(^{157, 225}\).

Despite these drawbacks *E. coli* have enjoyed extensive use in the production of biopharmaceuticals. An estimated 50% of approved biopharmaceuticals are
biosynthesized in *E. coli*\(^{130}\). Examples include Apidra® (insulin glulisine [rDNA origin] injection), a rapid-acting insulin analog for the treatment of diabetes mellitus\(^{230}\), and Accretropin\(^{TM}\) (recombinant human growth hormone (r-hGH); somatropin), used for the treatment of growth failure in patients with Turner syndrome\(^{230}\). Other therapeutics produced in *E. coli* include recombinant thrombolytics and anticoagulants (e.g. tissue plasminogen activator), hormones (e.g. insulin, human growth hormone, parathyroid hormone, calcitonin, glucagons), growth factors (e.g. granulocyte/granulocyte–macrophage colony-stimulating factor), and interferons (e.g. interferon-α and interferon-β)\(^{132, 136, 138, 140, 164, 210, 230}\).

*B. subtilis*

*Bacillus subtilis* is a gram-positive bacterium that is an attractive alternative to *E. coli*\(^{132, 164, 225-227, 231}\). It is a well characterized\(^{210, 225}\) soil bacterium\(^{210}\) and has been granted "generally regarded as safe" (GRAS) status\(^{132, 161, 210}\) by the FDA\(^{228, 231}\).

*B. subtilis* has many advantageous characteristics as a host for biopharmaceutical biosynthesis and is gaining popularity as a choice for such\(^{157}\). They can be grown in simple and cheap media\(^{210}\), have well studied\(^{210, 225}\) genetics, and several established expression systems are already commercially available\(^{132, 210}\). *B. subtilis* undergo facile transformation with bacteriophages and plasmids\(^{132, 157, 210}\). Compared to gram negative bacteria, such as *E. coli*\(^{132}\), their outer membranes do not contain lipopolysaccharides\(^{132, 157, 210}\) meaning their proteins are not produced with pyrogenic endotoxins\(^{132, 157}\). *B.*
*B. subtilis* are naturally able to grow to high cell densities\textsuperscript{132, 210} and secrete large quantities of protein directly into the extracellular medium\textsuperscript{157, 161, 210, 225, 227, 231}. Extracellular protein secretion, as opposed secretion into the cytoplasm, is advantageous because it eliminates the need\textsuperscript{132} for costly downstream purification and processing\textsuperscript{157} such as cell rupture, denaturation, and refolding\textsuperscript{132, 164, 227}.

*B. subtilis* possess their share of inherent disadvantages as well. They are known produce proteases that can degrade the recombinant proteins of interest\textsuperscript{132, 161, 210, 227}. Problems with plasmid instability\textsuperscript{157} and the lack available expression vectors and sustainable protocols for "high cell density cultivation (HCDC)"\textsuperscript{161} have also been reported for *B. subtilis* systems\textsuperscript{161, 210}. Compared to *E. coli* cells, transformation of *B. subtilis* is generally more difficult and requires Polyethylene glycoyo-mediated protoplasting\textsuperscript{225}. Finally since they are prokaryotic organisms *B. subtilis* don't possess the machinery for glycosylation\textsuperscript{157}, making them inefficient for production of glycosylated biopharmaceuticals. Given these issues systems of *B. subtilis* are more often utilized for the production of recombinant enzymes such as proteases, amylases, esterases, and lipases\textsuperscript{132, 227}.

Other *bacillus* mentioned in the current literature include the *B. megaterium*\textsuperscript{132, 157, 164, 225}, *B. brevis*\textsuperscript{132, 157, 225}, and *B. licheniformis*\textsuperscript{132}. Other bacterial hosts include *R. eutropha*\textsuperscript{132, 133, 157, 164, 210} (formally *A. eutrophus*\textsuperscript{164}), *Sta. carnosus*\textsuperscript{164} used for the production of organophosphohydrolase\textsuperscript{157}, and *Pseudomonas putida*\textsuperscript{132, 157}, used for the production of antibody fragments\textsuperscript{157}. The lack of genomic and metabolic information and expression
vectors for these bacteria may be to blame for their limited use thus far in biopharmaceutical production\textsuperscript{157}.

4.5.2 Eukaryotic Expression Systems

\textit{Fungi}

Fungi have many appealing characteristics for biosynthesis and are already a host of choice for some applications\textsuperscript{157, 162, 169}. Since they are eukaryotic organisms fungi can naturally perform the complex PTMs\textsuperscript{157, 158, 161, 162, 164}, such as glycosylation, phosphorylation, and disulfide bond formation\textsuperscript{132, 138, 140, 169, 219}, that may be important to biopharmaceutical structure and function\textsuperscript{162}. As with bacteria they are generally cheap and easy to culture\textsuperscript{157, 162, 169}, have short fermentation times on the order of days\textsuperscript{216}, are amendable to genetic manipulation, and possess a variety of established molecular tools and know-how for doing so\textsuperscript{169}. Furthermore, fungi can be grown in serum-free chemically defined media\textsuperscript{216} which eliminates the risk for viral or prion contamination of the cell lines and the potential this contamination will spread to the final product\textsuperscript{162, 169, 216, 219}. Fungi are resistant to shearing due to their rigid cell wall\textsuperscript{162, 169} and are therefore capable of growing to higher cell densities than mammalian and insect cells, making them apt for use in large-scale productions\textsuperscript{162}.

The primary drawback to fungal systems is they express glycosylation patterns that differ from those of the higher eukaryotes\textsuperscript{132, 136, 138, 155, 210, 216, 219, 226}. Yeast and other fungi\textsuperscript{162} are prone to hypermannosylation of \textit{N}-glycans\textsuperscript{113, 130, 140, 147, 157, 162, 169, 174, 209, 213, 216, 219}. 
Glycoproteins with high mannose content\textsuperscript{138} have affinity for human endothelial cell and macrophage mannose receptors\textsuperscript{162,219}, leading to rapid blood clearance\textsuperscript{113,130,136,169,219} and the potential for immunogenicity\textsuperscript{113,136,169,174,180,209,213,219}.

Advances in glycoengineering of fungi\textsuperscript{162,230} towards the expression of humanized glycoproteins\textsuperscript{113,136,162,174,209,210} have alleviated this issue. Glycoengineering approaches generally focus on eliminating the genes responsible for hypermannosylation and introducing genes of the human glycosylation pathway\textsuperscript{130,162,216,219,232}. These have resulted in the development of fungal hosts, such as \textit{Pichia pastoris}, capable of humanized glycosylation.

\textbf{Yeast}

Yeast are non-pathogenic\textsuperscript{157} unicellular eukaryotes\textsuperscript{132,155,233} of the\textsuperscript{210} kingdom fungi. Although convenient and advantageous hosts, yeast are typically not employed for the production of biopharmaceuticals until the simpler prokaryotic systems have proven to be ineffective\textsuperscript{132,138,219}. Yeast have still enjoyed extensive historical use in the food and agriculture industries\textsuperscript{157,219} and several non-glycosylated recombinant proteins expressed in yeast are already on the market\textsuperscript{174,209}. The most common yeast hosts for this purpose are \textit{Saccharomyces cerevisiae (S. Cerevisiae)} and \textit{Pichia pastoris (P. pastoris)}\textsuperscript{132,162}.

Yeast share many advantages with bacteria as hosts\textsuperscript{158,226}. They have well characterized genetics\textsuperscript{132,136}, undergo easy genetic manipulation\textsuperscript{136,157,158,161}, and high producing
mutant strains are already commercially available\textsuperscript{138,219}. Many molecular tools, such as “genome sequences, the ability to perform direct gene knockouts, gene replacements and site-specific chromosomal integration, and the availability of recyclable genetic markers allowing for the repeated introduction of new genetic constructs\textsuperscript{216}” are available for yeast. They also grow rapidly\textsuperscript{136} to high cell densities\textsuperscript{133,140,164} on relatively simple\textsuperscript{210} and inexpensive media\textsuperscript{132,138,161,174}, and are amendable for use in large-scale commercial fermentations\textsuperscript{168,174,210}.

Beyond their ability to carry out eukaryotic PTMs, yeast posses other advantages over the prokaryotic hosts. Yeast have been shown to express mammalian proteins that were not expressible or were expressed in insoluble forms from \textit{E. coli}\textsuperscript{161}. They are more efficient at secreting proteins directly into the extracellular space\textsuperscript{132,136,158,164,210} and thus may produce proteins with improved solubility and folding\textsuperscript{158,210}. Yeast are also not as limited as \textit{E. coli} in the size of rDNA proteins they can express, and are capable of biosynthesizing recombinant proteins larger than 50 kDa\textsuperscript{132}. Finally, they do not produce proteins contaminated by endotoxins\textsuperscript{136,210}.

Compared to insect and mammalian hosts, yeast undergo more facile genetic modifications, are easier and less expensive to culture\textsuperscript{210}, can grown to higher cell densities\textsuperscript{132,210}, and are considered generally more robust\textsuperscript{209}. Although they possess many attractive advantages yeast have traditionally not been used for the expression of glycoprotein biopharmaceuticals due to the high mannose content of their glycans\textsuperscript{130,169,216}. 
**S. Cerevisiae**

*S. Cerevisiae*, also known as baker's or brewer's yeast\(^{113, 169}\), has enjoyed three decades of use in the expression of recombinant proteins\(^{210, 234}\). Owing to this extensive use, *S. Cerevisiae* have well established industrial fermentation, purification, and formulation protocols\(^{113, 132, 157, 161}\). Furthermore they have been awarded GRAS status by the US FDA\(^{113, 140}\) and are not pathogenic or pyrogenic\(^{140, 164}\) to humans. Another advantage to *S. Cerevisiae* is they have, arguably, the best characterized genetics of all the eukaryotic organisms\(^{161, 164, 233}\). Compared to *E. coli*, however, *S. Cerevisiae* are still not as well characterized or as genetically understood\(^{138}\). Recombinant proteins expressed in *S. Cerevisiae* are also known to occasionally misfold\(^{138}\), and proteins larger than 30 kDa are not secreted extracellularly\(^{164}\).

Despite these drawbacks *S. Cerevisiae* has historically been the most commonly used yeast for the expression of biopharmaceuticals\(^{157, 161, 164, 169}\). In fact, the majority of FDA and European Medicines Agency (EMA) approved recombinant therapeutic proteins expressed in yeast are done so in *S. Cerevisiae*\(^{113, 133, 169}\). Examples include insulin, human growth hormone, glucagon, urate oxidase, hirudin, and granulocyte macrophage colony stimulating factor (GM-CSF)\(^{132, 133, 138, 140, 230}\). Examples of biopharmaceuticals include Merck's blockbuster recombinant human papillomavirus (HPV) vaccine Gardasil([Human Papillomavirus Quadrivalent (Types 6, 11, 16, and 18) Vaccine, Recombinant]\(^{113, 130, 139, 143, 144, 230}\), and Novo Nordisk’s NovoLog® Mix, an insulin aspart
mix used for the treatment of diabetes mellitus\textsuperscript{130, 140, 230}. \textit{S. Cerevisiae} are also used in the production of recombinant Hepatitis B vaccines\textsuperscript{210, 230}.

\textit{P. pastoris}

\textit{P. pastoris} is another established fungal host for the expression of recombinant proteins\textsuperscript{136, 140, 164, 169, 210}. It is a methylotrophic yeast\textsuperscript{131-133, 136, 157, 174} that is gaining popularity\textsuperscript{226} and is already used in the expression of over 400\textsuperscript{157} diverse recombinant proteins\textsuperscript{158, 210}.

Although both yeast are prone to hypermannosylation, it is much less extensive in \textit{P. pastoris} than in \textit{S. Cerevisiae}\textsuperscript{132, 157, 169}. \textit{P. pastoris} produces much shorter versions of the problematic high-mannose glycan chains. (\textit{P. pastoris} will usually only add about 8-20 residues\textsuperscript{157, 158} whereas \textit{S. Cerevisiae} is known to add closer to 50-150 residues\textsuperscript{132, 157, 169}.) Furthermore, unlike \textit{S. Cerevisiae}, \textit{P. pastoris} does not add \(\alpha\)-1,3-linked mannosyl terminal linkages to their glycan chains, which are highly immunogenic\textsuperscript{132} and contribute to rapid clearance of glycoproteins from blood\textsuperscript{157, 162}. Together these characteristics make \textit{P. pastoris} a much more appealing candidate for glycoengineering\textsuperscript{169}. It is the only yeast that has been glycoengineered to produce recombinant glycoproteins with homogeneous "human-type"\textsuperscript{174} \(N\)-glycans\textsuperscript{113, 132, 133, 157, 158, 161, 162, 168, 169, 174} to date. Humanization of the \textit{P. pastoris} glycosylation pathway was accomplished by the biotechnology firm GlycoFi who knocked out four genes responsible for glycosylation of the yeast type and introduced fourteen others to mimic the human glycosylation pathway\textsuperscript{201, 209, 216, 230}. Only
one protein glycoform\textsuperscript{180, 216, 219, 235} is generally obtained from this system, an improvement over glycoproteins obtained from mammalian cell culture\textsuperscript{216, 232} which tend to bear more heterogeneous glycosylation patterns.

Another advantage to \textit{P. pastoris} is its secretion efficiency\textsuperscript{132, 158, 164}, which can produce recombinant proteins in yields 10 to 100 times higher than \textit{S. Cerevisiae}\textsuperscript{157, 158}. Systems of \textit{P. pastoris} also do not require extensive process development\textsuperscript{158, 164}, are cheap and easy to maintain\textsuperscript{132}, and control of protein expression can be easily regulated via media manipulation\textsuperscript{132}. Even extractions and purifications are simplified in \textit{P. pastoris}, as they do not secrete many endogenous proteins\textsuperscript{158} and can also be grown in methanol solutions\textsuperscript{226} that are inhospitable to potential microorganism contaminants\textsuperscript{132}.

\textit{P. pastoris} systems, too, come with their inherent drawbacks. Unlike \textit{S. Cerevisiae}, \textit{P. pastoris} has not been awarded GRAS status by the FDA\textsuperscript{164, 236}. They have also been reported to occasionally misfold or incorrectly PTM eukaryotic proteins\textsuperscript{226}, such as with the addition of \textit{O}-glycans not found in endogenous human glycoproteins\textsuperscript{158}.

Despite these drawbacks \textit{P. pastoris} is one of the most extensively used rDNA expression hosts\textsuperscript{131, 236}, albeit with modest commercial success thus far. It has been used for the production of trypsin\textsuperscript{164}, tumor necrosis factor, serum albumin, and tetanus toxin fragment C\textsuperscript{131, 133}. In 2009\textsuperscript{157} the FDA approved the first biopharmaceutical biosynthesized in \textit{P. pastoris}, Kalbitor\textsuperscript{®}(ecallantide)\textsuperscript{157, 230, 235}. Kalbitor\textsuperscript{®} is a plasma kallikrein inhibitor used for the treatment of acute hereditary angioedema attacks\textsuperscript{157, 237}.
Other Yeast Hosts

Another promising yeast for the production of biopharmaceuticals is Hansenula polymorpha (H. polymorpha). H. polymorpha is a methylotrophic yeast closely related to P. pastoris that has many advantageous characteristics. As with P. pastoris, hypermannosylation is much less extensive in H. polymorpha than S. Cerevisiae. H. polymorpha is capable of efficiently producing proteins with molecular masses up to 150 kDa, and has been used for the biosynthesis of a hepatitis B vaccine (Rhein-Biotech). Other yeast include Schizosaccharomyces pombe, Kluveromyces lactis, Yarrowia lipolytica, Pichia methanolica, Pichia stipitis, Zygosaccharomyces rouxii, Zygosaccharomyces bailii, Candida boidinii, Schwanniomyces (Debaryomyces) occidentalis, and Arxula adeninivorans.

Filamentous Fungi

Filamentous fungi, known as molds, are also eukaryotic microorganisms. They too have enjoyed extensive historical use in the production of commercial metabolites, enzymes, and recombinant proteins. Although frequently mentioned in the literature as a host option for the production of biopharmaceuticals, filamentous fungi are currently "rarely" utilized for this purpose.
Their limited use in biopharmaceutical production to date should not be misinterpreted as an indication of their aptitude for doing so. Filamentous fungi are extremely attractive candidates for biosynthesis hosts and a number of laboratories are working towards the development of such systems at present. The most notable advantage to filamentous fungi is their ability to naturally "hyperproduce and secrete" extracellular proteins at levels superior to any of the other hosts. They are able to grow to high densities on relatively low cost and simple media, providing for "moderate to high production rates" of recombinant proteins with fermentation taking only a few days. Since they are higher organisms than the yeast, filamentous fungi possess more complex machinery for post-translational modification. As a consequence, filamentous fungi are able to glycosylate proteins in a way which more closely resembles endogenous human proteins.

Filamentous fungi are also seen as an enormous source of metabolic diversity that will contribute "novel biosynthetic pathways" applicable to the synthesis of a myriad of protein products. Indeed, only a small number of the estimated 1.5 million species of filamentous fungi have even been isolated, a demonstration of the currently untapped potential in filamentous fungi hosts.

Their extensive commercial use in the food and enzyme industries also makes filamentous fungi an attractive option for biopharmaceutical production. Many companies already have platforms in place for low-cost, large-scale, and high yielding
fermentations\textsuperscript{219, 224, 228}. Robust protocols for genetic transformations and scale-up procedures are also available, making their transition from industrial to pharmaceutical use in the near future all the more feasible\textsuperscript{219}.

Although filamentous fungi have enjoyed several decades of commercial use they are not physiologically or genetically characterized to the level of many of the simpler microorganisms\textsuperscript{164, 210}. Genomic sequence information\textsuperscript{228} and genetic tools\textsuperscript{224} are limited for most filamentous species which may impair the development, optimization, and efficiency of filamentous fungi hosts\textsuperscript{224}. Furthermore, although the extent of hypermannosylation observed from filamentous fungi is less pronounced than in yeast\textsuperscript{136, 157, 213, 228}, hypermannosylation is still known to occur\textsuperscript{157, 224, 228} and produces high-mannose $N$-glycans bearing approximately 20 mannose units\textsuperscript{132, 169, 219, 228}. A number of glycoengineering attempts at optimizing and/or humanizing the glycosylation pathways in filamentous fungi are underway\textsuperscript{228}. These pathways are highly complex and include a great deal of cellular processes not yet fully understood in filamentous fungi\textsuperscript{164, 224} and glycoengineering attempts thus far have proven to be quite difficult\textsuperscript{228}.

Proteases are another major drawback to filamentous fungi, which are known to produce and secrete proteases capable of degrading the recombinant protein products of interest\textsuperscript{132, 228}. Several engineering strategies to prevent proteolysis of recombinant proteins are being explored\textsuperscript{132}, including the selection and construction of protease mutant strains\textsuperscript{224} and fine-tuning of fermentation conditions\textsuperscript{228}. These, however, have not been met with much success to date\textsuperscript{132, 228}. Proteases may be to blame for the observation
that, although they are known for their ability for hyperproduction and secretion\textsuperscript{228}, these are not always observed from filamentous fungi hosts\textsuperscript{224}. In particular, limits in secretion have been seen when non-fungal recombinant proteins\textsuperscript{132, 164, 210, 224, 228} are expressed in filamentous fungi.

Another possible reason for limited recombinant protein secretion from filamentous fungi is the formation of “bottlenecks.” Any proteins with incorrect structures (improper protein assembly, PTMs, or folding) are kept in the fungal ER for degradation. This can cause bottlenecks in the filamentous fungi secretion pathway to form, both limiting its secretion capacity as well as stressing the cell\textsuperscript{224, 228}. Finally, transformation in Filamentous fungi can be difficult and may yield low transformation frequencies\textsuperscript{228} as well as random genetic integration of the genes\textsuperscript{164}. Some molecular strategies to improve recombinant protein expression in filamentous fungi have been met with success\textsuperscript{132}. These include the use of strong promoters\textsuperscript{132}, recombinant gene fusion\textsuperscript{210, 228} to a fungal protein naturally produced and secreted in high amounts\textsuperscript{132}, and the overproduction of foldases and chaperones\textsuperscript{228}.

A final general and significant disadvantage to filamentous fungi hosts is many of them are pathogenic to animals, humans, and other organisms\textsuperscript{228}. This is an obvious issue if filamentous fungi are to be seriously considered as expression systems for the production of biopharmaceuticals. Due to the aforementioned drawbacks to filamentous fungi they are alternatively being considered as novel sources of useful biosynthetic pathways. These pathways could be transformed into other (microbial or non-microbial) hosts that
may be better suited to carry out the actual biosynthesis\textsuperscript{228}. Until then, other fungal systems such as the yeast \textit{P. pastoris} are generally accepted to be better options as hosts for biopharmaceuticals production\textsuperscript{228}.

\textit{Aspergillus}

Members of the filamentous fungi genus \textit{Aspergillus} are receiving a great deal of attention as potential hosts for the production of biopharmaceuticals\textsuperscript{157, 224}. They have been utilized for over 1,500 years\textsuperscript{224} in food production and are currently used to produce both homologous and heterologous enzymes\textsuperscript{228} such as amylases, lipases, proteases, cellulases, and phytases\textsuperscript{224}. Of the 180 species of \textit{Aspergillus} that have been officially recognized\textsuperscript{228} some of the most promising as biopharmaceutical hosts include \textit{A. niger}\textsuperscript{132, 140, 169, 232}, \textit{A. awamori}, \textit{A. oryzae}\textsuperscript{133, 140, 228, 232}, \textit{A. nidulans}, and \textit{A. terreus}\textsuperscript{228, 22, 54, 122, 12}.

\textit{Aspergillus} produce correctly folded recombinant eukaryotic proteins with PTMs similar to those of endogenous human proteins\textsuperscript{132, 228}. Furthermore, the characteristic hypermannosylation of other fungi is usually not observed in \textit{Aspergillus}\textsuperscript{224}. Other advantages to \textit{Aspergilli} as hosts include their FDA granted GRAS status\textsuperscript{157} and, compared to other filamentous fungi, their genetics are better studied\textsuperscript{228}. ((Footnote\textsuperscript{?}: \textit{Aspergilli} species which currently have GRAS status include \textit{A. nidulans}, \textit{A. niger}, \textit{A. sydowii}, and \textit{A. awamori}\textsuperscript{164}.) A major drawback to \textit{Aspergilli} is the potential for proteases. \textit{Asperigilli} are known to secrete an abundant array of extracellular proteases\textsuperscript{228}.}
A. nidulans, for example, has an estimated 80 protease encoding genes\textsuperscript{132}. Furthermore, transformation of Aspergilli can be inefficient compared to yeast and E. coli due to the presence of the strong fungal cell wall\textsuperscript{224}. The use of enzymes to degrade the cell wall prior to transformation have been shown to help alleviate this issue but may introduce heterogeneity in transformation and offers less control over the transformation altogether\textsuperscript{224,228}.

Although Aspergilli are being explored as potential biopharmaceutical hosts they have not produced one to date\textsuperscript{157}. A number of recombinant human proteins have, however, successfully been expressed using systems of Asperigilli. These include human granulocyte macrophage colony stimulating factor (GM-CSF), interferon-\(\alpha\)-2, and humanized IgG1 antibodies in A. niger, interleukin-6 and lactoferrin in A. awamori, and epidermal growth factor (EGF), parathyroid hormone, and tissue plasminogen activator (tPA) in A. nidulan\textsuperscript{224}.

Other filamentous fungi mentioned in the literature include Trichoderma reesei\textsuperscript{157,219,228,232}, Acremonium chrysogenum\textsuperscript{133}, and Chrysosporium lucknowense\textsuperscript{133}.

\textbf{Plant}

Plant extracts were utilized as the main source of human pharmaceuticals for centuries\textsuperscript{155} until the 1800s when drug developers began using industrial chemical synthesis\textsuperscript{147}. Plants regained attention for the production of pharmaceuticals with the introduction of
recombinant DNA technology in the 1970s, and are now heavily investigated as expression systems for this purpose\textsuperscript{135, 147, 150, 171, 174}. The use of plants for recombinant protein production is also referred to as plant molecular farming (PMF)\textsuperscript{171, 238}. There are two general approaches to this process: via transgenic whole-plants or in cultures of plant cells.

Since plants are higher eukaryotic organisms\textsuperscript{157} their protein products share both “architectural and functional similarities\textsuperscript{210}” with human proteins above and beyond those of the prokaryotes and lower eukaryotes\textsuperscript{132, 147, 150, 217}. Transgenic plants are therefore capable of carrying out a range of complicated eukaryotic PTMs\textsuperscript{138, 150}, such as glycosylation\textsuperscript{132}, disulfide bond formation\textsuperscript{147}, oligomerization, and proteolytic cleavage\textsuperscript{147}, as well as producing complex, properly folded\textsuperscript{171} proteins\textsuperscript{132, 136, 147, 155, 209}.

Although plants can carry out complex eukaryotic post-translational modifications, their glycosylation patterns are known to consistently\textsuperscript{207} vary from mammalian proteins in both the type and degree of glycosylation observed\textsuperscript{136, 138, 140, 150, 155, 156, 210}. The plant $N$-glycosylation pathway has been relatively well characterized\textsuperscript{150} and it is proposed that these variations may result from processing steps that occur in the Golgi apparatus and beyond. These process steps are carried out by glycosidase and glycosyltransferase enzymes that differ significantly between plant and mammalian cells\textsuperscript{140}. More specifically, recombinant mammalian proteins expressed in plants tend to exhibit $N$-glycosylation at the same sites\textsuperscript{147} however the $N$-glycans are structurally different from their native human forms\textsuperscript{147}. Particularly problematic are the additions of plant-specific
fucose (α(1,3) fucose) and xylose (β(1,2) xylose) residues. These are immunogenic to many laboratory mammals, as well as elicit immunogenic responses from glycan specific antibodies in humans. Beyond the risk of inducing an immune response "the activated immune system" may cause the recombinant proteins to be more rapidly cleared from the blood and reducing therapeutic effectiveness. This possibility is exacerbated by the absence of sialic (neuraminic) acid in plant-made glycoproteins, a residue common to human glycoproteins and known to prevent protein blood clearance. Attempts to introduce a pathway capable of sialylation are underway but have, thus far, been met with limited success. Finally, human N-glycans contain galactose, a key residue important to the human immune system. Plant cells don't possess the enzymes necessary to make galactose. Galactose must therefore be added to plant N-glycans in order to humanize them, adding additional steps, and costs, in the production process.

Plants have been proven tolerant to various glycoengineering attempts, however strategies/techniques for doing so are much less developed than those for the microbe expression systems. Glycoengineering of the plant specific glycosyltransferases present in the golgi (α1,3 fucosyltransferase and β1,2 xylosyltransferase) to "knock-out" or "silence" the expression of plant-specific N-glycans have been successfully accomplished in mutant lines of Arabidopsis via insertional mutation and Physcomitrella patens via targeted gene inactivation. RNA-interference has also been used to for this purpose. Furthermore, several human
glycotransferases have been successfully expressed in plants in order to induce the
generation of recombinant glycoproteins bearing human-type N-glycans\textsuperscript{140, 150}. Another
glycoengineering strategy is intracellular targeting of recombinant proteins to specific
compartments of the cell. Although initially only used to increase yields this strategy
seems to alter the protein's glycosylation pattern expressed and may also be
advantageous\textsuperscript{140, 150} (e.g. if proteins do not enter the golgi apparatus they will not be
subject to the plant-specific protein modifications that occur there)\textsuperscript{140}.

\textbf{Transgenic Plants – Whole Plants}

Transgenic plants are plants that have undergone genetic modification via the
introduction of foreign DNA (a foreign gene or "transgene\textsuperscript{238}") using rDNA
technology\textsuperscript{238}. Transgenesis of plants is usually accomplished in one of two ways:
Agrobacterium mediated gene transformation via the insertion of the desired foreign gene
into Agrobacterium which then transfers it to the plants, or particle bombardment
(biolistics)\textsuperscript{140} in which the plant cells are literally bombarded with the foreign genes to be
transformed\textsuperscript{152, 155, 239}.

An advantage of transgenic plant expression systems is the opportunity for
directed/targeted accumulation of recombinant proteins to specific organs and subcellular
compartments, such as the cytosol, chloroplasts, vacuoles, and ER\textsuperscript{140, 147, 210}. This can
increase yields of the recombinant protein as well as simplify the protein purification
protocol\textsuperscript{140, 147}. It also creates the possibility for proteins to be stored in more
advantageous locations. Such locations may include (i) those with low proteolytic activity so as to reduce product losses from proteolysis, (ii) plant seeds or leaves which could enable more facile protein cultivation or administration\(^\text{140, 155}\). Expression systems of transgenic plants are therefore extremely cost-effective\(^\text{132, 136, 147, 155, 156, 174, 238, 240}\) in this sense, combining the benefits of the low cost and high yielding microbial systems with the capability for complex PTMs of the higher eukaryotes.

The advantage of transgenic plants as expression systems for the production of biopharmaceuticals over mammalian or human systems is plants are able to perform the complex post-translational modifications but still require only simple media, involving only minerals, sunlight, and water\(^\text{132, 138, 140, 209}\).

Transgenic plants are also known to be high producers\(^\text{132, 138, 140, 147, 156}\), with practically unlimited\(^\text{147, 155}\) production capacities\(^\text{140}\). The agriculture techniques used for growing\(^\text{155}\) transgenic plants offer the distinct\(^\text{210}\) advantage of flexible\(^\text{138}\), facile\(^\text{210}\), and inexpensive commercial\(^\text{238}\) scale up\(^\text{132, 155, 210}\) since it can be accomplished by simply growing more plants\(^\text{210}\) (increasing plant density in the field)\(^\text{140, 147}\). This is much easier to accomplish than scale-up in the other “culture-, fermentor- or reactor-based\(^\text{210}\)” expression systems, which can be quite difficult\(^\text{140, 210}\). Other advantages to transgenic plants is they undergo relatively facile transgenesis\(^\text{155}\), require only minimal extraction and purification\(^\text{132, 210}\), and the recombinant protein products can be easily and stably stored in plant seeds\(^\text{132, 210}\).

Last, but certainly not least, transgenic plant expression systems are safe\(^\text{140, 147}\) and
present an extremely low risk of contamination with human or zoonotic, pathogenic viruses and prions because no plant viruses have been identified as pathogenic to humans.

The huge commercial potential to these systems is one which "cannot be ignored," especially for the production of therapeutic glycoproteins and other proteins requiring more complicated human-like PTMs. A great deal of research in the development of transgenic plant expression systems is currently underway. "Seed crops," such as safflower, maize, and rice and "leaf biomass plants," such as tobacco, Arabidopsis thaliana and Medicago sativa (alfalfa) are the most commonly used transgenic plants in the production of therapeutic proteins. Other promising plants gaining attention for this purpose are duckweed, algae, and moss.

Disadvantages to transgenic plant expression systems include, thus far, slow commercial progress in their development. Low yields, transformation efficiency, and expression efficiency have also been observed for some recombinant proteins. Proteases, which are found in the compartments of many plants, may be partially to blame for these drawbacks. They have the potential to reduce the stability of and potentially degrade recombinant proteins, as well as complicate extraction and purification procedures. Strategies to increase the yield of transgenic plant expression systems have been developed. These include transgene optimization, codon optimization, selective suppression of plant RNA silencing, and targeting of recombinant proteins to particular organs and subcellular compartments of the plant cell.
Low recombinant protein productivity in transgenic plants necessitates that large amounts of plants be grown outside in open fields in order for quantities of biopharmaceuticals sufficient to meet the clinical demand be produced\textsuperscript{174}. Field containment, however, is an issue for transgenic crops\textsuperscript{174} as it is difficult to ensure the genetally engineered crop will not spread\textsuperscript{136, 138, 155, 171} via its pollen, seeds, and/or fruits\textsuperscript{171}. Greenhouses are one alternative to open field growth and may be more consistent with environmental and regulatory guidelines for the manufacturing of biopharmaceuticals\textsuperscript{174}. However, these become increasingly expensive as the production need rises\textsuperscript{155, 174} and may offset the cost-effectiveness of transgenic plant systems\textsuperscript{155}.

Human growth hormone, produced in tobacco, was the first therapeutic recombinant protein to be expressed in transgenic plants\textsuperscript{132, 138, 140}. Shortly thereafter they were utilized for the expression of more complex recombinant proteins, including full antibodies\textsuperscript{132, 140}. Since then a number of antibody fragments and complete antibodies have been produced in a variety of transgenic plants for both therapeutic and diagnostic applications\textsuperscript{132, 138, 140, 147, 155, 238, 240}. It is noteworthy to mention here that transgenic plants are exceptionally promising candidates for the production of edible recombinant vaccines\textsuperscript{140, 174, 210, 240} and topically administered biopharmaceuticals\textsuperscript{140}.

\textbf{Plant Cell Cultures}

An alternative to transgenic plants is cultured plant cells\textsuperscript{132, 147, 155}. This alternative offers
a more contained\textsuperscript{140} option for biopharmaceutical production in plants compared to the field or greenhouse grown plants\textsuperscript{140}, and allows for greater control in the biopharmaceutical production process. The additional control offers several advantages for plant cell culture systems. The biopharmaceuticals can be grown under sterile conditions\textsuperscript{147, 171} and secretion of the protein products directly into culture medium, as opposed to harvesting them from field crops, may substantially reduce downstream production costs\textsuperscript{147, 171}. Furthermore, the uniformity of plant cells in culture (as opposed to whole plants) reduces some of the glycan heterogeneity characteristic of field grown plants (e.g. glycosylation patterns have been shown to vary depending on the age of the plant\textsuperscript{147}), and thus may increase batch-to batch consistency\textsuperscript{136, 171}. Finally, plant cells have been shown to be capable of growth in high cell densities in culture\textsuperscript{140}.

Although the need for fermenters may increase production costs in these systems, this additional cost may be offset by alleviating the costs associated with containing field grown crops\textsuperscript{136, 138, 155, 171} or building and maintaining large greenhouses\textsuperscript{147, 155, 174}. This option may therefore sit better with regulatory agencies, or silence the environmental, societal, and political concerns surrounding field-grown transgenic crops\textsuperscript{140, 147}.

Proprietary systems have been developed for suspension cultures of plants \textit{P. patens} (moss, Greenovation Freiburg, Germany) and \textit{L. minora} (duckweed, Biolex Therapeutics Pittsboro, NC)\textsuperscript{140}. A novel system developed by Rincon Pharmaceuticals is capable of producing complex proteins right in the chloroplasts of \textit{Chlamydomonas reinhardtii} (green algae). Other cultured plants include \textit{Arabidopsis}, rice, soybeam, alfalfa, tomato,
and many tobacco culture systems (BY-2, NT-1, SR1, and Xanthi)\textsuperscript{140}. Issues, such as a lack of post-translational modifications occurring, have been reported for these systems. Plant cultures capable of producing biopharmaceuticals with engineered glycosylation do exist however\textsuperscript{230}. Other drawbacks to plant cell culture for biopharmaceutical production include low yields due to proteolytic activity and the limited availability of well-characterized cell lines\textsuperscript{171}.

Despite these, suspension cultures of plant cells have been successfully used for the production of variety biopharmaceuticals\textsuperscript{140, 147}. Protalix’s recombinant glucocerebrosidase (taliglucerase alfa), a treatment for Gaucher disease produced in carrot cells, is one such example\textsuperscript{140, 147, 171, 241}. An NDA was filed for taliglucerase alfa and granted an action date of May 1, 2012 based on positive Phase III clinical trial data\textsuperscript{140, 242}. Protalix, along with their partner Pfizer, have already made taliglucerase alfa available under Expanded Access Programs for adults suffering from Gaucher disease\textsuperscript{242}, making it the first plant-cell produced biopharmaceutical approved for human use\textsuperscript{140, 241}.

Plant cell produced recombinant proteins in preclinical development include $\alpha$-galactosidase for treating Fabry disease, follicle stimulating hormone for treating infertility, and acetylcholinesterase\textsuperscript{140}. The expression of full-size human antibodies (tobacco and rice cell cultures), granulocyte-macrophage colony stimulating factor (hGM-CSF) (tobacco and tomato cells), $\alpha$1-antitrypsin lysozyme (rice cell culture), tissue transglutaminase (tobacco), and dust mite allergens (tobacco)\textsuperscript{140, 147} have also been accomplished via plant cell culture expression systems.
Insect cells have been used as hosts for the production of over 200\textsuperscript{132, 133, 210} recombinant proteins of viral, bacterial, fungal, plant, and animal origin\textsuperscript{132, 133, 158, 169}. There are over 500 established insect cell lines, the majority of which were derived from\textsuperscript{220} the orders Lepidoptera and Diptera\textsuperscript{220}. The most common insect cell lines are the fall armyworm\textsuperscript{132, 133} (*Spodoptera frugiperda*), the cabbage looper (*Trichoplusia ni*), and the fruit fly (*Drosophila melanogaster*)\textsuperscript{158, 169, 174, 220}.

Heterologous gene expression in insect cells is typically accomplished via infection with viruses called baculoviruses\textsuperscript{132, 133, 138, 215, 220, 221, 243}. Baculoviruses are large double-stranded DNA viruses\textsuperscript{158} that are highly specific to insects and thus not infectious to plants, humans, or other animals\textsuperscript{132, 133, 210, 243}. Various baculovirus based expression systems and appropriate transfer vectors are already commercially available, including the Bac-to-Bac® (Invitrogen)\textsuperscript{158, 243}, flashBAC\textsuperscript{TM} (Oxford Expression Technologies)\textsuperscript{158, 220}, and BacPak\textsuperscript{TM} (Clontech) systems\textsuperscript{210}. An alternative for heterologous protein expression in insect cells is through stable transfection. Examples of commercially available transfection-based insect cell systems include InsectSelect\textsuperscript{TM} (Invitrogen),DES® (Invitrogen), and Mimic® (Invitrogen)\textsuperscript{168, 169, 243}.

Insect cells are an attractive alternative to the aforementioned hosts and possess many advantageous characteristics for the biosynthesis of biopharmaceuticals. Since they are eukaryotic organisms, insects possess the PTM machinery needed to express complex
PTMs more similar to human proteins than the microbial systems can\textsuperscript{132, 133, 136, 169, 210, 217, 220, 221}. These include the majority of mammalian PTMs\textsuperscript{158}, such as phosphorylation, N-glycosylation, O-glycosylation\textsuperscript{158}, carboxymethylation, amidation\textsuperscript{243}, correct signal peptide cleavage, proper proteolytic processing, acylation\textsuperscript{132, 243}, as well as proper protein folding and disulfide bond formation\textsuperscript{132, 158, 243}. Furthermore, the sites of these PTMs tend to more closely correlate with those of endogenously produced human proteins\textsuperscript{169, 215, 243}. Insect cells are also not limited in the size of proteins they can produce\textsuperscript{132, 243}, and can simultaneously express multiple genes\textsuperscript{132} to make multimeric proteins\textsuperscript{243}. 

Insect hosts are moderately quick producers. The time from cultivation to product harvest is generally on the order of weeks\textsuperscript{158, 232}. Insects have already successfully been employed for industrial scale production of recombinant proteins\textsuperscript{210, 220} and produce in high-density suspension cultures\textsuperscript{132}. These hosts are able to express protein at a higher level than any of the other eukaryotic systems\textsuperscript{132, 136, 158, 174, 210, 243}, and at moderate costs relative to mammalian hosts\textsuperscript{132, 133, 220}. Finally, insect cells are considered more stress resistant\textsuperscript{164}, tolerant of temperature fluctuations\textsuperscript{169}, and "easier to handle"\textsuperscript{164,165} than many of the other higher eukaryotes.

Insects do not come without their own inherent drawbacks, many of which can be attributed to their transformation procedure. The baculovirus is pathogenic to insect cells and eventually causes them to die\textsuperscript{138, 155, 210}. This means protein production in these systems is only transient\textsuperscript{169} and each “batch” of rDNA protein produced requires new insect cells and re-infection with the baculovirus\textsuperscript{138, 210}. Baculovirus infection can also be
an issue during scale-up and may cause the “passage effect,” in which protein production begins to decrease with each viral passage. Although the alternative stable transfection procedures offer a solution to these problems they tend to necessitate time consuming set-up.

Furthermore, although the glycosylation patterns of insect cells more closely resemble those of humans they are not identical. For example, N-glycosylation in mammalian cells yields glycoproteins that have terminal sialic acid galactose residues on their N-glycans. These are added via glycosyltransferase enzymes not present in insect cells, causing insect proteins to express “incomplete” N-glycosylation with simpler and shorter N-glycans compared to the mammalian glycoproteins. Insect cells also produce N-glycans that contain terminal mannose residues instead of the mammalian sialic acid galactose residues. These differences can cause adverse immunogenic affects and alter the in vivo bioactivity of the protein. In an attempt to alleviate these issues and fully utilize the benefits of insects, characterization of their N-glycosylation pathway and glycoengineering attempts to optimize it are currently underway.

Finally although they are quick producers compared to the higher eukaryotes, insect hosts exhibit relatively slow cell growth (18-24 hours) compared to the simpler microorganisms. This can lead to insufficient protein expression. It has been shown, however, that the addition of insect secretion signals may help alleviate this issue.
Despite the disadvantages associated with insect hosts they have found use in the biopharmaceutical industry, albeit somewhat limited thus far\textsuperscript{157}. Insect hosts have been used for the biosynthesis of vaccines, insecticides, and products for gene therapy\textsuperscript{174, 210}. Cervarix\textsuperscript{TM} [Human Papillomavirus Bivalent (Types 16 and 18) Vaccine, Recombinant] (GlaxoSmithKline)\textsuperscript{244}, a vaccine for the prevention of cervical cancer, is also produced in insect cells\textsuperscript{138, 180, 244}.

**Mammalian**

Mammalian cells are extremely attractive hosts for the production of biopharmaceuticals, especially those requiring complex PTMs\textsuperscript{132} and are occasionally the only viable option for such\textsuperscript{164}. They are higher eukaryotes\textsuperscript{157} and the host option most closely related to humans. Mammalian hosts are therefore able to perform PTMs identical, or the most similar, to human PTMs\textsuperscript{113, 132, 164, 168, 210} including glycosylation\textsuperscript{113, 132, 133, 140, 144, 155, 169, 210, 216, 219, 245, 246}, carboxylation\textsuperscript{155}, disulfide bond formation\textsuperscript{158}, and protein subunit assembly\textsuperscript{155, 247}. Mammalian rDNA proteins are the most likely to be produced in their correctly folded\textsuperscript{113, 132, 155, 210} forms. This may reduce the need for further protein processing and minimizes the risk for immunogenicity from foreign glycans\textsuperscript{164} and other epitopes. Currently, glycosylated biopharmaceuticals are most commonly biosynthesized in mammalian hosts\textsuperscript{132, 165, 169, 170, 216}. 
Many "platform technologies" for the stable transfection, gene amplification, and selection of mammalian cell lines are available. Despite the existence of these technologies, the genetics, biology, and physiology of many mammalian hosts are still not well understood. Moreover, although they have been used in biopharmaceutical production for a number of years much of the know-how pertaining to mammalian host systems is proprietary knowledge. This means, in many cases, new mammalian systems must be developed and optimized for each new biopharmaceutical product.

The development of mammalian hosts, including cell transfection, clonal selection, protein cultivation, and protein harvest, can be cumbersome, costly and time consuming compared to the systems of the microbes. It also requires that large expensive facilities be available and equipped with complicated technology. Even once a system has been developed and optimized mammalian cell growth is characteristically slow (24 hours). They require more time from cultivation to product harvesting and offer less control over the final protein product than the simpler systems. Mammalian hosts also tend to secrete rather low volumetric yields of rDNA protein products, an estimated 10-100 times lower than the microbes. Limits in their cell growth and productivity can be due to the presence of harmful metabolic by-products that can accumulate in cell cultures.

Ammonia, for example, is one such by-product and can affect glycosylation, inhibit cell growth, or even be toxic to mammalian cells. Efficient scale-up of these systems is therefore a necessity but can be complicated and costly. Perfusion cell culture has been shown to address these issues and help extend cell growth.
Beyond the costs associated with development, optimization, and scale-up of mammalian systems, they also require precise and labor intensive maintenance leading to high operating costs. Care must be taken to ensure these precarious cells are kept under stringent environmental conditions and are not contaminated with pathogenic viruses and prions during production. Serum-free, animal component-free, protein-free, and even chemically-defined media have been developed to relieve this issue and are now routinely employed in mammalian expression systems.

Another drawback to mammalian hosts is, although they express mammalian PTMs, their PTMs do not always match those of endogenous human proteins exactly. Slight differences can lead to immunogenic side effects or may necessitate downstream processing. Finally, mammalian cells are known to secrete heterogeneous distributions of protein glycoforms. Variations in pH, growth rate, temperature, concentrations of oxygen, ammonia, glucose, glutamine, and nucleotide sugars have been reported to increase the heterogeneity of glycoproteins.

Many strategies for glycoengineering of mammalian expression systems have been explored. Briefly, examples include fine tuning of culture conditions, the manipulation of glycosylation enzymes present in the cell, the use of enzymes in vitro to alter the glycosylation pattern after the protein has been purified, or downstream purification to remove the any unwanted glycoforms. Although effective, these
strategies can lead to extensive product loss. An estimated 80% of erythropoietin biosynthesized in CHO cells, for example, must be discarded because of improper glycosylation.

Despite the disadvantages characteristic to mammalian hosts, they have enjoyed use in the production of biopharmaceuticals treating various conditions including multiple sclerosis, anemia, and cancers. Biopharmaceutical examples include recombinant interferon-β (Avonex® (interferon beta-1a), Biogen Idec), erythropoietin (Procrit® (Epoetin Alfa), Ortho Biotech), and monoclonal antibodies (Herceptin® (trastuzumab), Genentech).

**Chinese Hamster Ovary**

CHO cells are immortalized ovarian cells from the Chinese hamster *Cricetulus griseus*, a rodent species native to the Chinese and Mongolian deserts. They have a history rich with "life-saving biomedical research… [and] are now credited with saving thousands of lives from illnesses like cancer every year."

CHO cells were introduced for laboratory use typing pneumococci in 1919 and into the world of protein manufacturing in 1957. Mutant CHO strains deficient in dihydrofolate reductase (DHFR) available in the early 1980s parlayed their entry into the production of recombinant proteins and enabled them to be one of the first mammalian cells to undergo stable gene transfection. Other than refinements, such as the
development of more productive and specific sub-cell lines, not a great deal of the technology used for CHO rDNA protein production has changed since the first DHFR-deficient mutants were transfected. They are still "the major host for the generation of recombinant cell lines" and have enjoyed decades of successful use in the biosynthesis of biopharmaceuticals.

The standard method for developing a CHO cell line for rDNA protein production begins with transfection of the cells by an exogenous gene of interest as well as, most commonly, the DHFR reductase gene. Clones expressing the exogenous gene are then selected for in media lacking glycine, hypoxanthine, and thymidine. The selected clones are amplified to increase their specific productivity of the exogenous gene and the most promising clones, in terms of production capability and growth rates, are isolated, expanded, and optimized.

There are many advantages to CHO cells as hosts for the biosynthesis of biopharmaceuticals. In particular, their glycoproteins express a pattern of glycans that is predominantly of the human type. CHO cells are generally considered the most dependable and, often times, preferred hosts for the production of biopharmaceuticals. Not surprisingly, they are the most commonly used mammalian host for the production of glycoproteins, including commercialized monoclonal antibodies and other recombinant proteins requiring complex PTMs. Their long history of regulatory acceptance and approval, considerable data available from safety testing, and
proven track record as hosts for the biosynthesis of biopharmaceuticals make them attractive from a regulatory perspective. Since they have “withstood the test of time\textsuperscript{170},” in this sense biopharmaceuticals produced in mammalian cells may have an easier time receiving FDA approval\textsuperscript{170}. Finally CHO cells may not be as susceptible to contamination by certain human pathogens, such as HIV, polio, herpes, influenza, and measles\textsuperscript{170}.

Examples from the impressive list of CHO produced biopharmaceuticals include Myozyme\textsuperscript{®}(alglucosidase alfa) (Genzyme), a glycogen-specific lysosomal enzyme used for the treatment of Pompe disease\textsuperscript{238}, blockbuster monoclonal antibody Rituxan\textsuperscript{®}(rituximab)\textsuperscript{249} (Genentech, Biogen Idec) for non-hodgkin's lymphoma\textsuperscript{140, 170, 249}, and Enbrel\textsuperscript{®}(Etanercept), a TNF-\(\alpha\) receptor fusion protein used in the treatment of rheumatoid arthritis\textsuperscript{170, 250}.

A disadvantage to CHO hosts is they necessitate time consuming development of cell lines\textsuperscript{170}. Selection and screening of clones is reported to take two months\textsuperscript{158}, and the entire process of recombinant CHO cell line generation can take more than six\textsuperscript{170, 245}. The recent development of high-throughput screening and selection methods are hoped to help speed up CHO cell line development\textsuperscript{158, 173, 176}. Also, compared to the well characterized microbes and better characterized mammals (mice and rats\textsuperscript{170}), a drawback particular to CHO systems is the CHO genome sequence is not publically available\textsuperscript{168, 170, 245}. 
Another drawback to CHO hosts is their potential to execute heterogeneous and non-human glycosylation\(^{168, 169, 213, 245}\). Some of the glycotransferase enzymes characteristic of human cells are known to be absent from CHO cells\(^{150}\). This can lead to large protein product loss during purification as any incorrect glycoforms must be removed from the product\(^{150}\) or may necessitate downstream glycan modification\(^{219}\). Successes in glycoengineering have alleviated some issues associated with glycosylation in CHO cells as well as enhanced the therapeutic properties of CHO produced biopharmaceuticals. The development of the Potelligent® (BioWa) cell line, for example, was one such breakthrough in CHO production\(^{130, 209, 251}\). The presence of certain fucosylated N-glycans on monoclonal IgG antibodies is known to dramatically reduce their antibody dependent cellular cytotoxicity (ADCC)\(^{169, 251}\), the primary means by which monoclonal antibodies initiate cancer cell death\(^{130}\). In the Potelligent® cell line, the gene (FUT8) that encodes the fucosyltransferase enzyme responsible for attaching the problematic fucosylated N-glycans to the antibody\(^{130}\) is completely "knocked-out"\(^{174}\) via "double siRNA knock-down"\(^{169, 166, 209}\). Completely fucose-free antibodies with as much as 100-fold\(^{130}\) improved ADCC activity\(^{209}\) are produced from these cell lines. This was an important breakthrough as it could lead to decreased mAb dose requirements\(^{173}\), a potential for cost reduction of treatment as well as reducing the unmet clinical need for these biopharmaceutical therapies\(^{169}\).
Other Mammalian Hosts

Ten percent of the biopharmaceuticals produced in 2010 were biosynthesized in mammalian hosts other than CHO cells. Examples of these include baby hamster kidney cells (BHK), mouse myeloma cells (NS0), and the human retina derived (PER-C6) and embryonic kidney (HEK) cells. BHK cell lines are noteworthy for their human-like N-glycosylation patterns and have been used for the biosynthesis of many vaccines and biopharmaceuticals including coagulation factors. The HEK-293 cell line is also noteworthy for its high transfection ability.

4.6 Biosimilar Outlook

The enormous complexity surrounding biosimilars makes their development a daunting investment, estimated to take 8-10 years and cost $US75-500 million. Regardless, there is an undeniable need for access to biosimilars worldwide.

Perhaps the most substantial barrier to biosimilars at present is not their development but rather their regulation. Biosimilar sponsors must navigate through uncharted legal, approval, and regulatory territory in bringing these products to market. Even once the FDA releases guidelines outlining an abbreviated pathway, many fear it may take another 5-10 years for the “kinks” in this pathway to be worked out. Until this time many developers may continue to apply for approval under the familiar BLA pathway.
on the basis of “superiority” rather than similarity, or instead in Europe using the EMA approval guidelines that have been in place since 2005.

Much needed advancements in biotechnology may also help speed up biosimilar development and approval. The establishment, validation, and standardization of enhanced analytical technology may reduce the depth of, or need for, clinical trials on biosimilars. Advances in glycomics would lead to improved understanding of glycoforms and their structure-function relationships. This may lead to advances in predictive immunology for more accurate prediction of a biopharmaceuticals antigenic potential prior to clinical trials. Finally, continued approaches to creative host engineering will undoubtedly lead to the establishment of improved hosts. These improvements may increase yields, purity, and simplicity, enhance PTM profiles, and decrease costs. Such biotechnological advancements may also open the industry to the opportunity for “biobetters,” second-generation biopharmaceuticals with improved characteristics over the reference product.

Despite the pressing need for biosimilars, patient safety must remain as the primary concern of their development. Until these products are better understood and their approval pathway has had time to settle rigorous pharmacovigilance and post-marketing surveillance programs must be in place. Judicious education of the physicians and clinicians who will be responsible for prescribing biosimilars will also be imperative to patient safety. This is especially true
considering the matter of “interchangeability” between biopharmaceutical products which, in many respects, will differ significantly from interchangeability as understood for generic pharmaceuticals\textsuperscript{127, 137}.

\textbf{4.7 Literature}


Roger, S. D., Biosimilars: how similar or dissimilar are they? *Nephrology (Carlton)* **2006**, *11*(4), 341-646.


186. Biosimilars. 


252. Cai, X.; Gouty, D.; Baughman, S.; Ramakrishnan, M.; Cullen, C., Recommendations and Requirements for the Design of Bioanalytical Testing used in