Bioactive Cyclic Peptides and Chemical Probes

Rebecca Courtemanche
Masters of Science
Department of Chemistry
University of Toronto
2014

Abstract

We report herein studies in aziridine aldehyde Ugi macrocyclization of peptides and their development as protein inhibitors, as well as the study of chemical probes for epigenetic proteins. Aziridine aldehyde peptide macrocyclizations are limited to N-terminal proline substrates. To expand this scope, the reactivity of pseudoprolines and secondary amine N-terminal building blocks in cyclizations will be reported. Cyclic peptides were explored as interrogators of EED/EZH2 (embryonic ectoderm protein/ enhancer of zeste homolog 2) by surveying the druggable surface of EED. Cyclic peptides were designed to inhibit the oncogenic tyrosine kinase Bcr-Abl (Break-point cluster region – Abelson). The computational design and synthesis of chemical probes for oncogenic proteins BRD1 (bromodomain-containing protein 1) and Setdb1 (SET domain bifurcated protein 1) will be reported. Results from biochemical assays including differential scanning fluorimetry (DSF), HSQC chemical shift mapping, isothermal titration calorimetry (ITC), and cocrystallization will be reported.
In loving memory of my grandparents Betty and Gordon Stubbs
Acknowledgments

I would like to thank Professor Andrei Yudin for his supervision and continued support throughout the course of my graduate studies. I am grateful for my second reader: Professor Jik Chin. Much of this research would not be possible with the support of my collaborators. Dr. Matthieu Schapira and Dr. Melissa Landon who performed computation modeling. Dr. Tony Pawson who we worked with on Bcr-Abl. Elena Dobrovetsky, who expressed and purified the proteins and taught me how to perform protein crystallizations and co-crystallizations. Dr. Aman Iqbal who performed the DSF experiments, and taught me crystallization techniques and ITC. Elena’s and Aman’s helpful discussions were very insightful and I was continually inspired by their optimism and passion. Aiping Dong who performed all of the protein X-ray work. Scott Houliston who performed all the protein NMR work and taught me the details of chemical shift mapping and 3D NMR. Finally, Dr. Peter Brown who oversaw the entire operation over at SGC. At the Department of Chemistry, I would like to thank the NMR staff including Dr. Timothy Burrow, Dr. Darcy Burns, and Dmitry Pichugin. Dr. Darcy Burns helped me to perform water suppression NMR experiments. Dr. Alan Lough for X-ray analysis. Anna Liza Villavelez and Kevin Ahn were of great assistance. Departmental funding and scholarships including NSERC CGS-M, as well as the various funding sources that allowed our lab to operate were greatly appreciated. Thank you to my fellow research colleagues, both past and present in the Yudin Group. In particular I would like to thank Dr. Conor Scully, who I worked closely with on the biologically inspired projects. His patience and guidance helped me from the beginning of my degree and his knowledge that he has passed down will continue to inspire my future work beyond graduation. Dr. Shinya Adachi and Victoria Corless who also helped out on this project through fragment synthesis and protein expression and purification. Thank you to my closest friends at U of T for their love and kindness: Jacque, Rita, Scott, Rhys, Frank, and Joy. Thank you to my family, friends outside the lab, and Bryan Boots – your love and support could not be overstated.
# Table of Contents

Acknowledgments........................................................................................................................................ iv

Table of Contents........................................................................................................................................ v

List of Abbreviations ................................................................................................................................ ix

List of Tables ................................................................................................................................................. xiii

List of Plates ................................................................................................................................................ xv

List of Figures ............................................................................................................................................... xvi

List of Appendices ...................................................................................................................................... xx

Chapter 1  \(N\)-terminus Proline Mimics for Peptide Cyclization ............................................................... 1

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

1.1.1 \(N\)-terminal amino acids in cyclization ................................................................................................. 1

1.1.2 Pseudoprolines and other secondary amine amino acids ................................................................. 2

1.1.3 Objectives ............................................................................................................................................ 5

1.2 Results .................................................................................................................................................... 5

1.2.1 Synthesis of pseudoproline substrates ............................................................................................... 5

1.2.2 Synthesis of pseudoproline piperazinones .......................................................................................... 8

1.2.3 Synthesis and cyclization of an \(N\)-terminal pseudoproline dipeptide .............................................. 9

1.2.4 \(N\)-benzyl amino acids ...................................................................................................................... 13

1.2.5 \(N\)-aryl terminal amino acids and cyclization reactions ................................................................. 15

1.2.6 Imidazolidinone peptides ................................................................................................................... 18

1.2.7 Using PGLGF as a benchmark – a final analysis of pseudoproline reactivity ..................................... 20

1.3 Conclusion and Future Work .................................................................................................................. 23

1.4 Experimental Details ............................................................................................................................ 23

1.4.1 General information ............................................................................................................................ 23
3.2 Methodology .......................................................................................................................... 69
  3.2.1 Glide docking .................................................................................................................. 69
    3.2.1.1 Preparation of protein ............................................................................................... 69
    3.2.1.2 Preparation of ligands ............................................................................................. 70
    3.2.1.3 Glide Docking ......................................................................................................... 70
  3.2.2 Evaluating binding efficiency ............................................................................................ 71
    3.2.2.1 Crystallization .......................................................................................................... 71
    3.2.2.2 Protein-ligand binding studies and structural characterization by NMR .............. 72
    3.2.2.3. Differential Scanning Fluorimetry, DSF and Stargazer ........................................ 73
    3.2.2.4 Isothermal Titration Calorimetry, ITC ................................................................. 73
  3.3 Results .................................................................................................................................. 74
    3.3.1 Pyrrolidine as a trimethyllysine mimic ....................................................................... 74
    3.3.2 Analysis of fragments by DSF and Stargazer ............................................................. 76
    3.3.3 Increasing fragment complexity using the Suzuki reaction: ...................................... 78
    3.3.4 Re-evaluating fragment design ..................................................................................... 81
    3.3.5 A new scaffold for BRD1 ............................................................................................. 87
    3.3.6 Analysis of compounds to BRD1 by $^1$H$^{15}$N HSQC NMR Chemical Shift Mapping ......................................................................................................................... 90
    3.3.7 Aggregation Studies ...................................................................................................... 98
    3.3.8 $^1$H$^{15}$N HSQC Binding Curve .................................................................................. 102
    3.3.9 3D NMR using C-13 N-15 BRD1 ............................................................................... 103
    3.3.10 Cocrystallization studies with BRD1 ......................................................................... 103
    3.3.11 Crystallization by soaking BRD1 crystals with compounds ..................................... 105
    3.3.11 Isothermal titration calorimetry, ITC ....................................................................... 107
    3.3.12 Reassigning the structure of 3.28 .............................................................................. 108
3.3.13 Setdb1 tetrahydroisoquinoline peptides ................................................................. 109

3.4 Conclusion and Future Work .................................................................................. 112

3.5 Experimental Details ............................................................................................ 113

3.5.1 General information ......................................................................................... 113

3.5.2 Characterization pyrrolidine analogues by microwave synthesis .................. 113

3.4.3 Synthesis of more pyrrolidine based fragments and other small molecules ................................................................................................. 114

3.6 References ............................................................................................................ 123

Appendix 1. Chapter 1 NMR Spectra, X-ray, and LCMS Data ........................................ 125

Appendix 2. Chapter 2 NMR Spectra and LCMS Data .................................................. 161

Appendix 3. Chapter 3 NMR Spectra ......................................................................... 165

Appendix 4. Red Wings Screen .................................................................................. 188
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Å</td>
<td>Angstrom</td>
</tr>
<tr>
<td>a.a.</td>
<td>amino acid</td>
</tr>
<tr>
<td>Abl</td>
<td>Abelson</td>
</tr>
<tr>
<td>Ac</td>
<td>acetyl</td>
</tr>
<tr>
<td>aq</td>
<td>aqueous</td>
</tr>
<tr>
<td>Ar</td>
<td>aryl</td>
</tr>
<tr>
<td>Bn</td>
<td>benzyl</td>
</tr>
<tr>
<td>Boc</td>
<td>di-tert-butyl dicarbonate</td>
</tr>
<tr>
<td>Bcr</td>
<td>Break point cluster</td>
</tr>
<tr>
<td>Bcr-Abl</td>
<td>Break point cluster Abelson fusion tyrosine kinase protein</td>
</tr>
<tr>
<td>BRD</td>
<td>bromodomains</td>
</tr>
<tr>
<td>BRD1</td>
<td>bromodomain containing protein one</td>
</tr>
<tr>
<td>tBu</td>
<td>tert-butyl</td>
</tr>
<tr>
<td>°C</td>
<td>degrees Celcius</td>
</tr>
<tr>
<td>cal</td>
<td>calories</td>
</tr>
<tr>
<td>Cbz</td>
<td>carboxybenzyl</td>
</tr>
<tr>
<td>CML</td>
<td>chronic myeloid leukemia</td>
</tr>
<tr>
<td>CS-Map</td>
<td>computational solvent mapping</td>
</tr>
<tr>
<td>cyc</td>
<td>cyclization</td>
</tr>
<tr>
<td>d</td>
<td>doublet (multiplicity in NMR)</td>
</tr>
<tr>
<td>dd</td>
<td>doublet of doublets (multiplicity in NMR)</td>
</tr>
<tr>
<td>dsnmax(r1)</td>
<td>maximum ratio of integrated signal to s/n</td>
</tr>
<tr>
<td>dt</td>
<td>doublet of triplets (multiplicity in NMR)</td>
</tr>
<tr>
<td>DCC</td>
<td>$N,N'$-dicyclohexylcarbodiimide</td>
</tr>
<tr>
<td>DCM</td>
<td>dichloromethane</td>
</tr>
<tr>
<td>DIAD</td>
<td>diisopropyl azocarboxylate</td>
</tr>
<tr>
<td>DIBAL-H</td>
<td>diisobutylaluminium hydride</td>
</tr>
<tr>
<td>DMA</td>
<td>dimethylacetamide</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-$(N,N$-dimethylamino)pyridine</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>DMF</td>
<td>N,N-dimethylformamide</td>
</tr>
<tr>
<td>DSF</td>
<td>differential scanning fluorimetry</td>
</tr>
<tr>
<td>dsnmax</td>
<td>calculate maximum signal-to-noise (VnmrJ Command)</td>
</tr>
<tr>
<td>EED</td>
<td>embryonic ectoderm development protein</td>
</tr>
<tr>
<td>Eqv.</td>
<td>equivalents</td>
</tr>
<tr>
<td>ESI</td>
<td>electron spray ionization (mass spectrometry)</td>
</tr>
<tr>
<td>Et</td>
<td>ethyl</td>
</tr>
<tr>
<td>EtOAc</td>
<td>ethyl acetate</td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>EZH2</td>
<td>enhancer of zeste homolog 2 (Drosophila)</td>
</tr>
<tr>
<td>FKBp</td>
<td>immunophilin and peptidyl-prolyl cis-trans isomerase protein</td>
</tr>
<tr>
<td>Fmoc</td>
<td>Fluorenlymethyloxycarbonyl</td>
</tr>
<tr>
<td>FRAP22</td>
<td>FKBP-rapamycin-associated protein 22</td>
</tr>
<tr>
<td>g</td>
<td>grams</td>
</tr>
<tr>
<td>Glide</td>
<td>grid-based docking with energetics</td>
</tr>
<tr>
<td>h</td>
<td>hours</td>
</tr>
<tr>
<td>H3</td>
<td>histone 3</td>
</tr>
<tr>
<td>H3K27</td>
<td>histone 3 lysine 27</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>histone 3 lysine 27 trimethylated</td>
</tr>
<tr>
<td>H3K36</td>
<td>histone 3 lysine 36</td>
</tr>
<tr>
<td>H3K36me3</td>
<td>histone 3 lysine 36 trimethylated</td>
</tr>
<tr>
<td>HFIP</td>
<td>1,1,1,3,3,3-hexafluoroisopropanol</td>
</tr>
<tr>
<td>HOBt</td>
<td>Hydroxylbenzotriazole</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>Hz</td>
<td>hertz</td>
</tr>
<tr>
<td>HRMS</td>
<td>high-resolution mass spectrometry</td>
</tr>
<tr>
<td>J</td>
<td>coupling constant in NMR</td>
</tr>
<tr>
<td>L</td>
<td>liters</td>
</tr>
<tr>
<td>lb</td>
<td>line broadening</td>
</tr>
<tr>
<td>m</td>
<td>mili</td>
</tr>
<tr>
<td>M</td>
<td>molar (moles/liter)</td>
</tr>
</tbody>
</table>
M  mega
M$^+$  parent molecular ion
Me  methyl
MeCN  acetonitrile
MeOH  methanol
mL  milliliter
mmol  millimole
MS  mass spectrometry
MTBE  methyl tert-butyl ether
MW  microwave
m/z  mass-to-charge ratio
NBS  N-bromosuccinamide
Nle  L-norleucine
NMR  Nuclear magnetic resonance
ns  number of scans
o  ortho
OAc  acetate
OtBu  $O$-tert-butoxy
P  para
PDB/pdb  protein data bank
Ph  phenyl
ppm  parts per million
PRC2  polycomb repressive complex two
Pyr  pyridine
q  quartet (NMR multiplicity)
RbAp48  retinoblastoma binding protein 4
Rf  retention factor
rg  receiver gain
R.O.  ring opening
rt  room temperature
s  singlet
s.d. simple dimer
SET Drosophila Su(var)3-9 and 'Enhancer of zeste' protein domain
SETdb1 SET domain containing protein bifurcated 1
sfrq spectrometer frequency
SGC Structural Genomics Consortium
SH2 scr homology 3-Scr homology 2
SH3 scr homology 3-Scr homology 3
s/n signal to noise ratio
SPPS solid phase peptide synthesis
SUZ12 Suppressor of zeste 12 homolog (Drosophila)
t triplet
tBuNC tert-butyl isocyanide
TIPS triisopropylsilane
TFA trifluoroacetic acid
TFE 2,2,2-trifluoroethanol
tt triplet of triplets
u micro
VLC vacuum liquid chromatography
WD40 beta-transducin protein
List of Tables

Table 1.1 Boc protection reactions of N-phenyl phenylalanine ........................................... 17

Table 1.2 Reaction conditions for synthesis an N-terminal imidazolidinone dipeptide .... 20

Table 1.3 Cyclization reaction screening for NAr-Phe-Gly-Leu-Gly-Phe ......................... 22

Table 1.4 Cyclization results of 5aa, 4aa, and 2aa linear peptide substrates .................. 23

Table 2.1 Synthesis of 7a.a. cyclic peptide PLMT(OtBu)GD(OtBu)T(OtBu) ................. 49

Table 2.2 Structure and sequence of 7a.a. linear peptides .............................................. 50

Table 2.3 Cyclization of 7a.a. sequences using general cyclization conditions .......... 51

Table 2.4 DMSO additive for cyclization of 2.3 ................................................................. 52

Table 2.5 Conditions used for cyclizing 2.3 with d-leucine aziridine aldehyde dimer .... 55

Table 3.1 Synthesized molecules using microwave reaction ........................................... 77

Table 3.2 Compounds analyzed by DSF and Stargazer .................................................. 79

Table 3.3 Isoindoline scaffolds and docking scores ......................................................... 83

Table 3.4 Isoindoline compounds and docked scores ..................................................... 84

Table 3.5 Compounds screened by ¹H¹⁵N HSQC NMR to assess binding .................... 92

Table 3.6 Acquisition and processing parameters ......................................................... 101

Table 3.7 dsnmax(r1) values for various concentration of DMAP and compounds of interest ........................................................................................................ 103

Table 3.8 Workable concentrations for compounds of interest .................................. 104

Table 3.9 NMR sample preparation for HSQC binding curve ................................... 105
Table 3.10 Results of plate 3.1, Red Wings screen of crystallization conditions .......... 106

Table 3.11 3.25 BRD1 co-crystals cryoprotectant optimizations .................................. 107

Table 3.12 Mounted BRD1 crystals soaked with compounds ........................................ 109
List of Plates

Plate 3.1 Cocrystallizations screen of 3.25, 3.27, and 3.28 ............................................. 106

Plate 3.2 Cocrystallizations of 3.25 and 3.27 with BRD1 using conditions H10 ................. 107

Plate 3.3 Hanging-drop crystallizations of BRD1 ................................................................ 108

Plate 3.4 Soaking of compounds, 8 μL at 4 mM with BRD1 apo crystal ............................. 109
## List of Figures

**Figure 1.1** Synthesis of cyclic peptides using an aziridine-aldehyde ........................................ 1

**Figure 1.2** Mechanism of macrocyclization .................................................................................. 1

**Figure 1.3** Pseudoprolines derived from serine, threonine, and cysteine ................................. 2

**Figure 1.4** Derivatives of serine from reactions with: formaldehyde (left), acetaldehyde (middle), and acetone (right) ......................................................................................... 3

**Figure 1.5** Kink in $2^\circ$ structure induced by pseudoproline in peptide 9b: Ac-Cys-Leu-Ala-Ala-Leu-Aib-Ala-Leu-Leu-Ala-Thr($\psi^{HH}$Pro)-Leu-Sib-Ala-Leu-Leu-Ala-NH$_2$ 3

**Figure 1.6** Fmoc-Gly-Ser($\psi^{HH}$Pro) ............................................................................................. 4

**Figure 1.7** Solid phase peptide synthesis, SPPS, synthetic utility, and variations utilizing pseudoprolines .................................................................................................................. 4

**Figure 1.8** Stereochemistry analysis of 1.5 by NMR; minor diastereomer $J$ coupling constants calculated by $^1$H NMR analysis ................................................................................................. 9

**Figure 1.9** Crystal structure and crystal lattice of 1.19 .................................................................. 16

**Figure 1.10** Comparison of predicted and observed nOe signals ............................................. 16

**Figure 1.11** Measured distance between two relevant protons = 3.84Å ....................................... 17

**Figure 2.1** PRC2 Complex ........................................................................................................... 42

**Figure 2.2** Left: Top view of EED top view. PDB Code 3JZN Right: EZH2 alpha helix (red) bound to EED. PDB Code 2QXV .................................................................................................. 42

**Figure 2.3** Inactive and active conformers of Abl ........................................................................ 44

**Figure 2.4** CS-Map of EED binding pocket ................................................................................ 45
Figure 2.5 9-membered cyclic peptide scaffolds used for computational modeling ........................................46

Figure 2.6 Gaussian distribution of docked cyclic peptides .................................................................46

Figure 2.7 Workflow to screen cyclic peptides by computational analysis ..........................................46

Figure 2.8 9-membered rings with best docking score ........................................................................47

Figure 2.9 Left: Crystal Structure Abl activation loop (highlighted in yellow) PDB Code 2FO0 Right: native loop of Beta-turn and cyclic peptide mimic .........................................................48

Figure 2.10 Left: SH2 kinase domain, red and yellow spheres highlight polar sites, aromatic regions are shown in blue. Middle: SH2 kinase domain, available binding surface illustrated in gray. Right: Isoleucine-164 sits on upper right ridge of pocket (residues in orange) ...........................................................................................................................57

Figure 2.11 List of 20 cyclic tetramer peptides selected from computational analysis ................57

Figure 2.12 All L-a.a. containing sequences .........................................................................................58

Figure 3.1 Nucleosome showing structure of histones and histone tails ............................................66

Figure 3.2 PTMs of histone tails ........................................................................................................67

Figure 3.3 Crystal structure of histone-3 peptide bound to BRD1 ......................................................69

Figure 3.4 BRD inhibitors IBET 762, IBET 151, and JQ1 .................................................................70

Figure 3.5 Representative glide grid centered on aromatic cage of BRD1. BRD1 pdb code3LYI ...............................................................................................................................................72

Figure 3.6 Chemdraw of lysine and the corresponding smile code ..................................................72

Figure 3.7 Left: hanging-drop and sitting drop. Right: Concentration of protein and precipitant to obtain nucleation (crystallization) state ..................................................................................74

Figure 3.8 Typical DSF curve ..............................................................................................................7
Figure 3.9 Left: Raw data from a typical ITC titration Right: ITC titration curve

Figure 3.10 Compounds used in dose response assay and BRD1 crystal soaks

Figure 3.11 Suzuki compounds docked in BRD1 with the best-fit

Figure 3.12 Benzyl isoindoline fragment (magenta) docked in BRD1

Figure 3.13 Retrosynthesis of pyrazole isoindoline

Figure 3.14 Retrosynthetic scheme towards target compound

Figure 3.15 Apo BRD1 (blue). 3.28:BRD1 (12:1) (red)

Figure 3.16 Apo BRD1 (blue). 3.25:BRD1 (12:1) (red)

Figure 3.17 Apo BRD1 (blue). 3.27:BRD1 (12:1) (red)

Figure 3.18a 3.28 bound to BRD1 (3:1) (red), 1:1 (green). Apo protein (blue)

Figure 3.18b Zoomed. 3.28 bound to BRD1 (3:1) (red), 1:1 (green) Apo BRD1 (blue)

Figure 3.19a 3.25 bound to BRD1 (3:1) (red), 1:1 (green). Apo protein (blue)

Figure 3.19b Zoomed: 3.25 bound to BRD1 (3:1) (red), 1:1 (green) Apo protein (blue)

Figure 3.20a 3.27 bound to BRD1 (3:1) (red), 1:1 (green). Apo protein (blue)

Figure 3.20b Zoomed: 3.27 bound to BRD1 (3:1) (red), 1:1 (green) Apo protein (blue)

Figure 3.21a. 3.24 and BRD1 (red) 12:1. BRD1 apo protein (blue)

Figure 3.21b. Zoomed region. 3.24 and BRD1 (red) 12:1. BRD1 apo protein (blue)

Figure 3.22 SGC compound 3.30 hit for BRD1

Figure 3.23a 3.30 and BRD1 (red) 10:1. BRD1 apo protein (blue)
Figure 3.23b Zoomed region. 3.30 and BRD1 (red) 10:1. BRD1 apo protein (blue). 100

Figure 3.24 DMAP stacked spectra at various concentrations. 102

Figure 3.25 Example data for compounds, stacked spectra of 3.27 at various concentrations. 102
List of Schemes

Scheme 1.1 Synthesis plan using an N-terminal pseudoproline ...........................................5

Scheme 1.2. Acid catalyzed synthesis of thiazolidine .................................................................5

Scheme 1.3 Base catalyzed synthesis of pseudoprolines ............................................................6

Scheme 1.4 Synthesis by crystallization of pseudoproline ..........................................................6

Scheme 1.5 Cyclization of 1.2b .....................................................................................................6

Scheme 1.6 Optimized synthesis of thiazolidine ...........................................................................7

Scheme 1.7 General synthesis of oxazolidine ..................................................................................7

Scheme 1.8 Synthetic plans for the synthesis of piperazinone, as well as di- and tri- cyclic peptides .....................................................................................................................................8

Scheme 1.9 Synthesis of 1.5 ..........................................................................................................8

Scheme 1.10 Proposed scheme using Boc and methyl ester protecting groups .........................9

Scheme 1.11 Synthesis of alanine and valine methyl ester hydrochloride ..................................10

Scheme 1.12 Synthesis of protected dipeptide .............................................................................10

Scheme 1.13 Synthesis of deprotected dipeptide ..........................................................................10

Scheme 1.14 Cyclization of Thz-Ala .............................................................................................11

Scheme 1.15 Cbz-based retrosynthesis .........................................................................................11

Scheme 1.16 Syntheses of Cbz-Thz, 1.11 .....................................................................................12

Scheme 1.17 Synthesis of dipeptide and subsequent deprotections ..........................................12

Scheme 1.18 Hydrogenation reaction of 1.11 ..............................................................................12

Scheme 1.19 Benchmark reaction for hydrogenation conditions ..............................................13
Scheme 1.20 Hydrogenation of 1.11 in the presence of Cbz-Val ........................................ 13

Scheme 1.21 Reductive amination of l-alanine to give 1.15 ........................................... 14

Scheme 1.22 Synthesis of Boc-protected benzyl alanine .................................................. 14

Scheme 1.23 Synthesis of Fmoc-N-benzyl-l-alanine ......................................................... 14

Scheme 1.24 Synthesis of N-aryl phenylalanine 1.18 ....................................................... 15

Scheme 1.25 Microwave synthesis of N-aryl phenylalanine 1.18 .................................. 15

Scheme 1.26 Synthesis of piperazinone 1.19 .................................................................. 15

Scheme 1.27 Boc protection of N-phenyl phenylalanine .................................................. 17

Scheme 1.28 Synthesis of an N-terminal N-aryl amino acid peptide .......................... 18

Scheme 1.29 Deprotection of an N-phenyl dipeptide linear substrate and subsequent cyclization .......................................................................................................................... 18

Scheme 1.30 Synthesis of an imidazolidinone N-terminus dipeptide substrate .......... 19

Scheme 1.31 Synthesis of a simple imidazolidinone ......................................................... 19

Scheme 1.32 Synthesis of an N-terminal imidazolidinone dipeptide substrate .......... 19

Scheme 1.33 Synthesis of methyl ester and reaction to give imidazolidinone ............ 20

Scheme 1.34 Synthesis of relevant amino acids ............................................................... 21

Scheme 1.35 Synthesis of cyclic XGLGF, X = Thz, N-Ar-Phe, or N-Bz-Ala ............... 22

Scheme 2.1 Synthesis of cyclic 7a.a. PLMT(OtBu)GD(OtBu)T(OtBu) using conditions (Table 2.1) ............................................................................................................................. 48

Scheme 2.2 Synthesis of cyclic peptide sequences (example for sequence 2.2) ........ 51

Scheme 2.3 Investigating effects of DMSO on cyclization ........................................... 52
Scheme 2.4 Synthesis of cyclic PL(Nle)TGDT, 2.6 .......................................................... 53

Scheme 2.5 Synthesis of linear PL(Nle)TGDT, TIPS = triisopropylsilane ..................... 54

Scheme 2.6 Linear peptide using d-leucine aziridine aldehyde dimer, tBuNC, and various conditions .................................................................................................................. 54

Scheme 2.7 Synthesis of lactone macrocycle 2.8 .......................................................... 56

Scheme 2.8 Solution phase benzylaion of 2.12 ................................................................ 59

Scheme 3.1 General synthesis of pyrrolidines using microwave synthesis ...................... 77

Scheme 3.2 Reductive amination syntheses ...................................................................... 77

Scheme 3.3 Synthesis of sulfonamides ........................................................................... 78

Scheme 3.4 Synthesis of pyrrolidine amide derivatives .................................................... 78

Scheme 3.5 Suzuki reaction using Pd(dba)₃ ................................................................... 81

Scheme 3.6 Synthesis of phenylurea ligand for Suzuki cross-coupling ............................ 81

Scheme 3.7 Suzuki cross-coupling using phenylurea ligand ............................................ 81

Scheme 3.8 Synthesis of stable Pd catalyst for Suzuki cross-couplings ............................ 82

Scheme 3.9 Cross-coupling synthesis of 3.16 .................................................................. 82

Scheme 3.10 Synthesis of 3.17 by microwave synthesis ................................................... 82

Scheme 3.11 Synthesis of 3.18 ....................................................................................... 83

Scheme 3.12 Synthesis of isoindoline compounds ............................................................ 85

Scheme 3.13 Synthesis of benzyl isoindoline using mild conditions .................................. 86

Scheme 3.14 Syntheses of 3.21 ....................................................................................... 87

Scheme 3.15 Cross-coupling reaction of 3.21 ................................................................. 87
List of Appendices

Appendix 1. Chapter 1 NMR Spectra, X-ray, and LCMS Data ........................................127

Appendix 2. Chapter 2 NMR Spectra and LCMS Data .........................................................174

Appendix 3. Chapter 3 NMR Spectra ...................................................................................180

Appendix 4. Red Wings .......................................................................................................210
Chapter 1 *N*-terminus Proline Mimics for Peptide Cyclization

1.1 Introduction

1.1.1 *N*-terminal amino acids in cyclization

There are several approaches to cyclizing a linear peptide including head-to-tail, or employing one or more side chain functionalities. Despite numerous creative synthetic advances including metal-ion assistance, sulfur-mediated cyclizations, ring-contractions utilizing lactones, azide-alkyne cycloadditions, ring-closing metathesis, and multicomponent reactions, there are still several synthetic challenges.\(^1\) Forcing a linear peptide precursor to adopt a cyclic geometry is an entropic battle. If successful, not only will the desired product form, but also oligomers and polymers resulting from undesired intermolecular processes.\(^1\) Although high dilution and pseudo-high dilution conditions are utilized to control side reactions, this is not the most practical approach nor does this method reliably give the desired product.

The synthetic approach developed by the Yudin research group addresses both the entropic and high-dilution limitations of macrocyclization. By employing an amphoteric aziridine aldehyde into a modified Ugi four-component condensation, the synthesis of small to large macrocyclic peptides in concentrated reaction media is enabled without oligomeric or polymeric side products.\(^2\)

![Figure 1.1 Synthesis of cyclic peptides using an aziridine-aldehyde](image1)

![Figure 1.2 Mechanism of macrocyclization](image2)
As shown in Figure 1.2 the $N$-terminus of the linear peptide will condense to form an iminium ion with the open dimer form aziridine aldehyde, which is then attacked by an isocyanide. The resulting nitrilium species is attacked by the $C$-terminus giving a mixed-anhydride macrocycle. A fast transannular attack by aziridine on the mixed anhydride finally gives an aziridine-containing macrocyclic peptide. The aziridine functionality can be further reacted with nucleophiles to give pure macrocyclic peptides following HPLC purification.

Currently, macrocyclization synthesis employs proline at the $N$-terminus as primary amines are slow to react with aziridine aldehyde dimer and give poor conversion to the cyclic peptide. This constrain on the $N$-terminus drastically limits what linear peptides can be selected for cyclization, as they must all contain an $N$-terminal proline. There are two questions that arise from these observations. Why does only proline work? And, can employing other ‘proline’-like or pseudoproline $N$-terminal amino acids expand the current toolkit for macrocyclizations. In the following chapter, the latter question will be investigated.

1.1.2 Pseudoprolines and other secondary amine amino acids

Pseudoprolines as evidenced from their name are simple proline mimics. Typically they are derived from serine, threonine, and cysteine.

![Figure 1.3 Pseudoprolines derived from serine, threonine, and cysteine](image)

Usually, these compounds are derived from a condensation reaction with formaldehyde, acetaldehyde, or acetone. The pseudoproline products from the condensation reaction are acid labile; those that are derived from formaldehyde are the most robust and require harsh conditions to convert the derivative back into its original amino acid. In this way the pseudoproline is ‘deprotected’.
A powerful usage of pseudoprolines is their ability to reduce intermolecular hydrogen bonding and therefore prevent peptide aggregation, self-association, and beta-sheet secondary structure formation. In fact, during the assembly of peptides these pseudoprolines will help solubilize fully protected peptides. Therefore, these simple molecules are ideal for longer linear sequences during synthesis and purification. The ability of pseudoprolines to induce turns have made them attractive building blocks for the synthesis of highly constrained cyclic tetrapeptides. Moreover, incorporating pseudoprolines into peptides can give rise to interesting secondary structures. Wohr et al., took advantage of pseudoprolines energetic preference for a cis-amide bond to induce a kink in an alpha helix.

Typically, pseudoprolines are used as substrates within a linear sequence and can be used as a reversible and orthogonal protecting group where the pseudoproline can be deprotected with strong acid to give the original serine, threonine, or cysteine. Pseudoprolines are typically used in the form of a building block: Fmoc-Xxx-Xaa(Ψ^{H, H} Pro), where Xxx is any amino acid, and Xaa is a pseudoproline and z is either an H or Me depending on the pseudoproline derivative (Figure 1.6).
In the synthesis of a linear peptide, this dipeptide substrate is usually incorporated in the middle of a sequence and the pseudoproline may be deprotected to give back serine, threonine, or cysteine either before or after the resin is cleaved. Furthermore, other protecting groups, PG, may be cleaved first depending or at the same time as the pseudoprolines is deprotected. These synthetic variations are all dependent on which pseudoproline is selected and its robustness to acidic cleavage cocktails.

Currently, macrocyclization synthesis with (S)-aziridine aldehyde dimer and tert-butyl isocyanide are limited to proline at the N-terminus, as primary amines are slow to react with dimer and give poor conversion to cyclic peptides. Pseudoprolines are secondary amino acids that are structurally related to proline and differ by an oxygen or sulphur. Utilizing pseudoprolines in macrocyclization may offer unique reactivity with (S)-aziridine aldehyde dimer and furthermore, the incorporation of a new oxygen or sulphur heteroatom may give rise to interesting structural properties of the resulting macrocycles. Moreover, heteroatoms of
pseudoprolines can offer new hydrogen bond contacts when these macrocycles are explored in biologically relevant systems

1.1.3 Objectives

Conventional dipeptide pseudoproline building blocks can’t be utilized, as the pseudoproline substrate is not on the \( N \)-terminus. Instead, they will be synthetically synthesized and coupled onto linear peptides prior to cyclization.

Scheme 1.1 Synthesis plan using an \( N \)-terminal pseudoproline

Pseudoprolines, \( N \)-benzyl amino acids, and \( N \)-aryl amino acids will be synthesized and studied in piperazinone formation, and in the cyclization of cyclic peptides as \( N \)-terminus substrates. Based on these results, pseudoprolines will be evaluated as possible \( N \)-terminal substrates for further use in peptide cyclization chemistry.

1.2 Results

1.2.1 Synthesis of pseudoproline substrates

Dimethyl- Ser, Thr, and Cys derived pseudo-prolines (\( \psi^{\text{Mc,Mc}}\text{Pro} \)) were first synthesized. An example procedure for synthesizing the Cys derived compound is shown below.\(^7\)

Scheme 1.2. Acid catalyzed synthesis of thiazolidine

Conducting this reaction with cysteine in hydrochloric acid and acetone gave a 2:1 mixture of product and starting material using \(^1\text{H} \) NMR analysis. Rather than optimizing this procedure,
formaldehyde-derived pseudoprolines were synthesized, as these are more robust to acidic conditions and therefore more applicable to solid phase linear peptide synthesis methodology.

Using a base-catalyzed procedure, the pseudoproline can first be isolated from the equilibrium reaction with a Boc group then formic acid is added to cleave the protecting group and isolate the final product. The Boc derivatives 1.1a-b were isolated cleanly, however upon addition of formic acid, a mixture of compounds was obtained as evidenced by $^1$H NMR. The formic acid was neutralized using sodium bicarbonate, yet by $^1$H NMR, the reaction was not clean. Repeating the reaction and carefully controlling the pH, gave the same result by NMR analysis.

![Scheme 1.3 Base catalyzed synthesis of pseudoprolines](image)

A different procedure from literature was attempted where the product could be selectively crystallized out of solution with addition of a weak base (Scheme 1.4).

![Scheme 1.4 Synthesis by crystallization of pseudoproline](image)

Scheme 1.5 Cyclization of 1.2b
Using standard reaction procedures, the oxazoline 1.2b was subjected to cyclization conditions using (S)-aziridine aldehyde dimer and tert-butyl isocyanide in TFE. The reaction was monitored by LCMS to completion. Purification by LCMS of the major product and analysis by NMR led to the conclusion that the major product was the TFE adduct of serine. Therefore, in the previous reaction, the oxazoline did not form and serine was recovered.

A review of literature and syntheses of pseudoprolines was done and it was found that only thiazolidine is commercially available. Employing a modified procedure, 1.4 was successfully synthesized in five minutes with high yield and purity.

Another method used to achieve both oxazolidine and methyloxazolidine uses anhydrous acidic conditions (Scheme 1.7). However, the preparation of 3N HCl in EtOAc is done by dripping sulphuric acid onto NaCl and the resulting formation of HCl gas bubbled into EtOAc.

Given the successful synthesis of thiazolidine, compound 1.4 was carried onwards to cyclization chemistry. Currently, there is one example of an N-terminal thiazolidine dipeptide in literature, which was achieved by coupling a methyl ester amino acid with Boc-thiazolidine followed by deprotection of Boc with trifluoroacetic acid. The initial goal with thiazolidine substrate, 1.4 was to first make a piperazinone, followed by a cyclic dipeptide, and finally a cyclic tripeptide using (S)-aziridine aldehyde dimer and tert-butyl isocyanide (Scheme 1.8).
Scheme 1.8 Synthetic plans for the synthesis of piperazinone, as well as di- and tri- cyclic peptides

1.2.2 Synthesis of pseudoproline piperazinones

Initially, a small-scale reaction to form a thiazolidine piperazinone was performed. After 15 min, the reaction showed product, yet the major peak by LCMS was m/z = 550.5 (unknown by-product). Performing this reaction on a larger amount of thiazolidine, only one major peak was observed by LCMS and TLC. However on inspection of the $^1$H and $^{13}$C NMR, it was observed that the reaction was not diastereoselective. 10% of the opposite diastereomer was present by NMR. Interestingly, synthesis of proline piperazinone was known to yield a single diastereomer so therefore the diastereoselectivity of thiazolidine could be a limiting factor in its incorporation in macrocyclization reactions.

Scheme 1.9 Synthesis of 1.5
Purification of 1.5 from column chromatography gave both diastereomers observable by NMR. A previous student in our lab, Dr. Ben Rotstein, confirmed that the SYN diastereomer of proline piperazinone gave a coupling constant of $J = 6.4$ Hz by $^1$H NMR and X-ray analysis. This correlates with the coupling constant of $J = 6.5$ Hz for the major product present in thiazolidine piperazinone. Therefore, it was concluded that the major diastereomer of 1.5 must also have a SYN relationship.

![Stereochemistry analysis of 1.5 by NMR; minor diastereomer $J$ coupling constants calculated by $^1$H NMR analysis](image)

**Figure 1.8** Stereochemistry analysis of 1.5 by NMR; minor diastereomer $J$ coupling constants calculated by $^1$H NMR analysis

### 1.2.3 Synthesis and cyclization of an $N$-terminal pseudoproline dipeptide

The synthesis of a dipeptide precursor for an $N$-terminal thiazolidine cyclic dipeptide was devised from a Boc-thiazolidine and alanine methyl ester protected amino acid synthesis (Scheme 1.10).

![Proposed retro-synthetic scheme using Boc and methyl ester protecting groups](image)

**Scheme 1.10** Proposed retro-synthetic scheme using Boc and methyl ester protecting groups
Alanine methyl ester 1.6 and valine methyl ester 1.7 as the hydrochloride salts were synthesized in high yield and purity.\textsuperscript{12}

\[ \text{Scheme 1.11 Synthesis of alanine and valine methyl ester hydrochloride} \]

Alanine methyl ester, 1.6 was reacted with Boc-thiazolidine 1.2a using, EDC, HOBt, and Et\textsubscript{3}N to give the protected dipeptide 1.8 in 43% yield after purification by column chromatography.

\[ \text{Scheme 1.12 Synthesis of protected dipeptide} \]

Saponification to cleave the ester group followed by deprotecting the $N$-terminal Boc group using trifluoroacetic acid gave 1.10 in 91 % yield (Scheme 1.13).\textsuperscript{13,14}

\[ \text{Scheme 1.13 Synthesis of deprotected dipeptide} \]

Two parallel cyclization reactions were setup with 1.10, one with and one without an equivalent of diisopropylethylamine (DIPEA) in order to neutralize trifluoroacetic acid. However although starting material was consumed in both reactions, no product could be identified.
Scheme 1.14 Cyclization of Thz-Ala

Rather than trying to neutralize the peptide in the reaction, the dipeptide was completely neutralized prior to cyclization using potassium carbonate. However the potassium carbonate most likely deprotonated the carboxylic acid and solubilized the resulting salt into the aqueous layer. Eliminating trifluoroacetic acid altogether, a new synthetic scheme was devised taking advantage of a carboxybenzyl, Cbz protecting group.

Scheme 1.15 Cbz-based retrosynthesis

Cbz protected thiazolidine was synthesized then later optimized to a 2 h reaction, which afforded Cbz-Thz 1.11 in near quantitative yield and high purity.\(^{15,16}\) Due to the steric bulk of the Cbz group, rotamers were present and VT-NMR confirmed this phenomenon (see Appendix II). Cbz-Thz was reacted with alanine methyl ester using DMAP, and later optimized using HOBT to give the dipetide.\(^{14,17}\) Finally, the dipeptide 1.12 was reacted with base to saponify the methyl ester and subjected to hydrogenation to cleave Cbz.\(^{17}\)
Scheme 1.16 Syntheses of Cbz-Thz, 1.11

Scheme 1.17 Synthesis of dipeptide and subsequent deprotections

Hydrogenation of Cbz proved unsuccessful and starting material was fully recovered. In order to properly study this reaction, the simpler Cbz-Thz substrate was hydrogenated. In 1.17 Again, this method was unsuccessful and pure starting material remained as evidenced by $^1$H and $^{13}$C NMR.

Scheme 1.18 Hydrogenation reaction of 1.11
In order to verify that Cbz-Thz was stable to hydrogenation or if the catalyst had decomposed, a hydrogenation was carried out on another Cbz substrate using the exact same conditions. Cbz-L-phenylalanol was available in-house and was pure by NMR. This benchmark reaction was successful and Cbz was completely cleaved from phenylalanol.

![Scheme 1.19 Benchmark reaction for hydrogenation conditions](image)

**Scheme 1.19** Benchmark reaction for hydrogenation conditions

Finally both Cbz-Val and Cbz-Thz were hydrogenated in the same vessel and analysis of the crude mixture by NMR showed that neither substrate underwent hydrogenation. This result was not expected as the catalyst was previously shown to hydrogenate Cbz-L-phenylalaninol. More work is needed to explore optimal conditions for the successful removal of the Cbz protecting group.

![Scheme 1.20 Hydrogenation of 1.11 in the presence of Cbz-Val](image)

**Scheme 1.20** Hydrogenation of 1.11 in the presence of Cbz-Val

### 1.2.4 *N*-benzyl amino acids

*N*-benzyl amino acids were synthesized as possible *N*-terminal amino acids in linear peptide substrates for cyclization. This substrate emerged in parallel with the project on the design and synthesis of cyclic peptide inhibitors for the SH2-kinase domain of Bcr-Abl (Chapter 2).

To being exploring this chemistry, a simple amino acid was selected and subjected to benzylation conditions. Using reductive amination conditions, *L*-alanine was reacted with benzaldehyde. However, over-reaction led to the dibenzylated derivative.\(^{18}\) Optimization of both the time and temperature gave the correct mono-benzylated derivative 1.15 in high yield.
Scheme 1.21 Reductive amination of L-alanine to give 1.15

\[
\begin{align*}
\text{O} & \quad \text{HN} \quad \text{OH} \\
\text{NH}_2 & \quad \text{C} \quad \text{NaBH}_4 \\
\text{O} & \quad 0^\circ \text{C}, \text{rt}, 16 \text{ h} \\
\text{O} & \quad \text{OH} \\
& \quad 1.15 (98%)
\end{align*}
\]

\[\text{N} - \text{benzyl alanine, } 1.15 \text{ was then Boc-protected and Fmoc-protected to produce two different substrates suitable for synthesis of peptide sequences (Scheme 1.22 and Scheme 1.23). Boc protection using Boc}_2\text{O and triethylamine in a mixture of dioxane and water gave 1.16 in 51\% yield. Synthesis of Fmoc-N-benzyl alanine gave the expected product though it contained } N\text{-hydroxysuccinimide, which was inseparable by chromatography. The Fmoc reaction was repeated using Fmoc-Cl as the reagent and 1.17 was obtained.}\]

\[
\begin{align*}
\text{HN} & \quad \text{OH} \\
\text{Boc}_2\text{O} & \quad \text{TEA, dioxane, H}_2\text{O} \\
\text{rt, o/n} & \quad 1.16 (51\%) \\
\text{OH} & \quad \text{N} \\
& \quad 1.15 \\
\text{OH} & \quad \text{Boc}_2\text{O} \\
\end{align*}
\]

Scheme 1.22 Synthesis of Boc-protected benzyl alanine

\[
\begin{align*}
\text{O} & \quad \text{HN} \quad \text{OH} \\
\text{C} & \quad \text{Na}_2\text{CO}_3, \text{Fmoc-Suc} \\
\text{O} & \quad \text{H}_2\text{O, dioxane} \\
\text{3h, rt, 0\^\circ C - rt} & \quad \text{Product not separable from } N\text{-hydroxysuccinimide} \\
\text{OH} & \quad \text{N} \\
& \quad 1.16 \\
\text{OH} & \quad \text{Boc}_2\text{O} \\
\text{H}_2\text{O, dioxane} & \quad \text{0\^\circ C - rt} \\
\text{1.17 (34\%)} & \quad \text{Na}_2\text{CO}_3, \text{Fmoc-Cl}
\end{align*}
\]

Scheme 1.23 Synthesis of Fmoc-N-benzyl-L-alanine
1.2.5  *N*-aryl terminal amino acids and cyclization reactions

*N*-aryl amino acids were synthesized using a Chan-Lam coupling reaction. Using anhydrous and anoxic conditions, as well as a long reaction time at high-temperature resulted in trace product identifiable by LCMS.\(^\text{19}\)

![Scheme 1.24 Synthesis of *N*-aryl phenylalanine 1.18](image)

This reaction was ideal to adapt to a microwave synthesis and to explore this, two different microwave reactions were set up: one with iodobenzene, 5 mol% copper (I) oxide and potassium carbonate in NMP; and another with bromobenzene, 5 mol% copper (I) iodide, and potassium carbonate in DMF. The former did not react while the latter proved successful (see Scheme 1.25).\(^\text{19}\)

![Scheme 1.25 Microwave synthesis of *N*-aryl phenylalanine 1.18](image)

*N*-aryl phenylalanine, 1.18, was subjected to piperazinone formation (Scheme 1.26) and following purification; crystals were successfully grown using a vapor diffusion method to confirm product stereochemistry.

![Scheme 1.26 Synthesis of piperazinone 1.19](image)
Figure 1.9 Crystal structure and crystal lattice of 1.19

NOESY NMR was performed on 1.19 to see if nOe interactions would corroborate the stereochemistry observed in the crystal structure. However, this was not the case as only one nOe out of the two predicted nOe interactions necessary to confirm stereochemistry (see Figure 1.10).

Figure 1.10 Comparison of predicted and observed nOe signals

NOESY data is often discussed in relation to stereochemistry however any two protons within four angstroms of each other, no matter their directionality, should theoretically have a nOe, this cross peak was not observed. Rendering the crystal structure into jmol, the distance between these protons was calculated as 3.84 angstroms (Figure 1.11). However, considering that the conformation of a molecule in a crystal structure does not necessarily correspond to the conformation of the molecule in solution the actual distance between these two protons in
solution could be greater. Furthermore, it is not only distance that affects nOe signals, but also a molecules viscosity or tumbling rate, as well as correlation time.

**Figure 1.11.** Measured distance between two relevant protons = 3.84Å

Since N-phenyl phenylalanine, 1.18, cyclized to give the piperazinone 1.19, an N-phenyl terminal dipeptide was synthesized and subjected to cyclization. To synthesize the dipeptide substrate, N-phenyl phenylalanine was Boc protected. Boc protection of 1.18 did not readily occur and screening of several reaction conditions all gave negative results (Table 1.1).

[![Scheme 1.27 Boc protection of N-phenyl phenylalanine](image)](image)

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Conditions</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1,4-dioxane, H₂O, NaOH</td>
<td>SM</td>
</tr>
<tr>
<td>2.</td>
<td>1,4-dioxane, H₂O, NaOH (microwave)</td>
<td>SM</td>
</tr>
<tr>
<td>3.</td>
<td>Neat, rt, o/n</td>
<td>SM</td>
</tr>
<tr>
<td>4.</td>
<td>Neat, DIPEA, rt. o/n (10 mg SM)</td>
<td>SM</td>
</tr>
<tr>
<td>5.</td>
<td>Neat, DIPEA, rt. o/n (30 mg SM)</td>
<td>SM</td>
</tr>
</tbody>
</table>

SM = Starting material. DIPEA = diisopropylethylamine

These results suggested that N-phenyl phenylalanine was very unreactive and that protection of the amine may not be required. To test this hypothesis, N-phenyl phenylalanine was reacted with previously synthesized valine methyl ester to give 1.20 in 74% yield (Scheme 1.28).
Scheme 1.28 Synthesis of an N-terminal N-aryl amino acid peptide

Dipeptide 1.20 was saponified to give the carboxylic acid in high yield and purity and this substrate was subjected to cyclization conditions (Scheme 1.29). Analysis by LCMS showed only trace amount of the hydrolyzed product (M+H$_2$O)$^+$ and many other overlapping peaks.

Scheme 1.29 Deprotection of an N-phenyl dipeptide linear substrate and subsequent cyclization

1.26 Imidazolidinone peptides

Imidazolidinones are secondary amine amino acids that, like pseudoprolines contain five-membered rings, and were seen as a potential substrate to incorporate in this project. In order to synthesize an imidazolidinone, the N-terminus of a free amine needs to react with the NH of the nearby amide. To begin, a dipeptide was synthesized as shown below using standard coupling procedures. Following a literature procedure, the dipeptide was converted to its sodium salt, isolated as a white powder and reacted with acetone under reflux. However this reaction proved unsuccessful.
Scheme 1.30 Synthesis of an imidazolidinone N-terminus dipeptide substrate

A different approach towards an imidazolidinone was then attempted using a simpler substrate (Scheme 1.31). Pivaldehyde was condensed with the primary amine, which then reacted with the NH of the amide. Rearrangement under acidic conditions gave an imidazolidinone and pure product in high yield was obtained without any recrystallization.

Scheme 1.31 Synthesis of a simple imidazolidinone

Adapting this reaction for a different substrate, Gly-Gly hydrochloride, was unsuccessful and only starting material was observed. Suspecting that the carboxylic acid might be consuming triethylamine preferentially over the amine, a series of reactions were set up to screen reaction conditions that led to product formation.

Scheme 1.32 Synthesis of an N-terminal imidazolidinone dipeptide substrate; see Table 1.2 for various reaction conditions
Table 1.2 Reaction conditions for synthesis an $N$-terminal imidazolidinone dipeptide

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Starting Material</th>
<th>Base, equiv.</th>
<th>Solvent</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GlyGly HCl</td>
<td>TEA, 1.5</td>
<td>MeCN</td>
<td>No Rxn.</td>
</tr>
<tr>
<td>2</td>
<td>GlyGly HCl</td>
<td>Pyr, 1.5</td>
<td>MeCN</td>
<td>No Rxn.</td>
</tr>
<tr>
<td>3</td>
<td>GlyGly HCl</td>
<td>TEA, 1.0</td>
<td>MeCN</td>
<td>No Rxn.</td>
</tr>
<tr>
<td>4</td>
<td>GlyGly HCl</td>
<td>Pyr, 1.0</td>
<td>MeCN</td>
<td>No Rxn.</td>
</tr>
<tr>
<td>5</td>
<td>GlyGly HCl</td>
<td>TEA, 1.0</td>
<td>DCM</td>
<td>No Rxn.</td>
</tr>
<tr>
<td>6</td>
<td>GlyGly HCl</td>
<td>Pyr, 1.5</td>
<td>DCM</td>
<td>No Rxn.</td>
</tr>
<tr>
<td>7</td>
<td>GlyGly HCl</td>
<td>Pyr, 1.0</td>
<td>DCM</td>
<td>No Rxn.</td>
</tr>
<tr>
<td>8</td>
<td>GlyGly</td>
<td>-</td>
<td>DCM</td>
<td>No Rxn.</td>
</tr>
<tr>
<td>9</td>
<td>GlyGly</td>
<td>-</td>
<td>MeCN</td>
<td>No Rxn.</td>
</tr>
</tbody>
</table>

Reactions 1 – 9 (Table 1.2) were set up in a parallel reaction block using four-angstrom molecular sieves to ensure that water was removed from the system. TLC analysis showed that all vials contained pure starting material and no product was observed. As an alternative to adjusting equivalencies of base, the carboxylic acid was esterified to make a methyl ester. The reactivity of the methyl ester derivative was then explored however, analysis showed only starting material. Imidazolidinone as $N$-terminus peptide substrates were not investigated further past this point.

![Scheme 1.33](image)

Scheme 1.33 Synthesis of methyl ester and reaction to give imidazolidinone

1.2.7 Using PGLGF as a benchmark – a final analysis of pseudoproline reactivity

The linear sequence PGLGF readily forms the cyclized product cleanly. We decided to use this sequence as a benchmark sequence as a final analysis of pseudoproline and secondary amine cyclization reactivity. Using the sequence XGLGF, where X corresponds to a pseudoproline or a secondary amine amino acid, a direct comparison of these substrates against proline could be achieved. To initiate this comparison, the relevant $N$-terminus amino acids (thiazolidine (Thz), $N$-benzylalanine (N-BzAla), and $N$-phenylphenylalanine (N-Ar-Phe)) were first synthesized.
Scheme 1.34 Synthesis of relevant amino acids

With the relevant amino acids in hand, the linear sequences were made using the automated solid phase peptide synthesizer and 2-chlorotrityl resin. GLGF-resin was synthesized by SPPS and then X was coupled on manually to make XGLGF-resin where X = Thz, N-Ar-Phe, or N-Bz-Ala. Cleavage of the resin was done using 25% HFIP in CH$_2$Cl$_2$ and purification of the resulting linear peptides was done by trituration with cold methyltertbutylether (MTBE) followed by repeated centrifugation and washing with cold MTBE. Cyclizations of these substrates were unsuccessful for both the Thz and N-Bz derivatives, while the N-Ar resulted in trace amounts of hydrolyzed product (Scheme 1.35).
Since a trace amount of product was detected for the N-Ar derivative, this reaction was screened with various conditions by modifying the equivalencies of (S)-aziridine aldehyde dimer, tBuNC, DIPEA, and HOAc (Table 1.3).

Table 1.3 Cyclization reaction screening for N-Ar-Phe-Gly-Leu-Gly-Phe

<table>
<thead>
<tr>
<th>Reaction</th>
<th>(S)-aziridine aldehyde dimer (eqv.)</th>
<th>tertbutyl isocyanide (eqv.)</th>
<th>Additive (eqv)</th>
<th>LCMS result (m/z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.6</td>
<td>1.2</td>
<td>-</td>
<td>SM</td>
</tr>
<tr>
<td>2.</td>
<td>1.2</td>
<td>2.4</td>
<td>-</td>
<td>SM, trace M+18</td>
</tr>
<tr>
<td>3.</td>
<td>0.6</td>
<td>1.2</td>
<td>DIPEA (1.0)</td>
<td>Decomposition</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td>2.4</td>
<td>DIPEA (2.0)</td>
<td>Decomposition</td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td>HOAc (1.0)</td>
<td>SM, decomposition</td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td>HOAc (2.0)</td>
<td>SM, decomposition</td>
</tr>
</tbody>
</table>

Eqv. = Equivalencies, DIPEA = diisopropylethylamine, HOAc = acetic acid

As indicated in Table 1.3, neither DIPEA nor HOAc were successful in yielding the cyclized product. Furthermore, additional amounts of (S)-aziridine aldehyde dimer or tBuNC increased the rate of hydrolyzed product formation.

Although the proline mimics failed to work for X-GLGF, linear substrates with varying lengths and containing different amino acids were synthesized and subjected to cyclization. Amino acids not requiring any protecting groups were selected and all linear peptides were synthesized by SPPS in which the key N-terminus amino acid was coupled on manually. The linear peptides were cleaved, then purified by trituration with cold MTBE followed by repeated centrifugation.
and washing with cold MTBE. All linear peptide substrates were subjected to cyclization and the results are summarized below (Table 1.4)

**Table 1.4 Cyclization results of 5aa, 4aa, and 2aa linear peptide substrates**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sequence</th>
<th>Mass of peptide (mg) / mmol, (1.0 eqv.)</th>
<th>Vol. (uL) of (S)-aziridine aldehyde dimer stock in TFE (0.6 eqv.)</th>
<th>Vol. (uL) of tBuNC stock in TFE (0.6 eqv.)</th>
<th>LCMS results at various time points</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Thz-GLGA</td>
<td>5 / 0.012</td>
<td>100</td>
<td>79</td>
<td>SM, SM, SM</td>
</tr>
<tr>
<td>B</td>
<td>Thz-LGF</td>
<td>5 / 0.011</td>
<td>95</td>
<td>75</td>
<td>SM, SM, SM</td>
</tr>
<tr>
<td>C</td>
<td>Thz-F</td>
<td>3 / 0.011</td>
<td>91</td>
<td>73</td>
<td>SM, SM, SM</td>
</tr>
<tr>
<td>D</td>
<td>NBzA-GLGA</td>
<td>5 / 0.010</td>
<td>89</td>
<td>71</td>
<td>SM, SM, SM</td>
</tr>
<tr>
<td>E</td>
<td>NBzA-LGF</td>
<td>5 / 0.011</td>
<td>91</td>
<td>72</td>
<td>SM+26, SM+26, SM+26</td>
</tr>
<tr>
<td>F</td>
<td>NBz-F</td>
<td>5 / 0.009</td>
<td>78</td>
<td>78</td>
<td>SM, SM, SM</td>
</tr>
<tr>
<td>G</td>
<td>NArF-GLGA</td>
<td>5 / 0.009</td>
<td>79</td>
<td>79</td>
<td>SM, SM, M+18, SM+18, SM+18</td>
</tr>
<tr>
<td>H</td>
<td>NArF-LGF</td>
<td>5 / 0.009</td>
<td>76</td>
<td>76</td>
<td>SM, SM, SM</td>
</tr>
<tr>
<td>I</td>
<td>NArF-F</td>
<td>3 / 0.008</td>
<td>66</td>
<td>66</td>
<td>SM, SM, SM</td>
</tr>
</tbody>
</table>

SM = starting material, M+18 = hydrolyzed product.

### 1.3 Conclusion and Future Work

While the oxazolidines derived from serine and threonine proved unsuccessful, thiazolidine was synthesized in high yield and purity and cyclized to a piperazinone with a 9:1 ($S$, $R$) diastereoselectivity. An $N$-benzyl terminus amino acid in a peptide sequence did not give a cyclized product however the $N$-phenyl terminus amino acid peptide sequences seemed more promising. An $N$-phenyl phenylalanine piperazinone was synthesized and its stereochemistry was fully assigned through NMR and its X-ray crystal structure. While imidazolidinone were attempted, initial reactions proved this substrate unviable. A final comparison of the $N$-terminal substrates again PGLGF as well as cyclization of various lengths of linear peptide substrates also proved ineffective.

### 1.4 Experimental Details

#### 1.4.1 General information

All reactions were carried out under an atmosphere of air unless otherwise stated. Reagents were used as obtained from commercial suppliers or purified according to standard procedures. Flash chromatography was performed using SiliaFlash P60 40 – 63 um silica. TLC was performed on EMD silica gel 60 F254 TLC glass plates and visualized with UV light, $I_2$ on silica, or
permanganate stain. All $^1$H and $^{13}$C NMR spectra were recorded using a Bruker 400 MHz or Varian 400, 600, 700, or 500 MHz cryoprobe spectrophotometer and internally referenced to a residual protiosolvent. 2D (COSY, HSQC) NMR spectroscopy was used where appropriate to assist in the assignment of $^1$H and $^{13}$C NMR spectra. LC-MS (ESI) was recorded on an Agilent 1200 Series quadrupole spectrometer.

1.4.2 Synthesis and characterization data

\[ \text{(S)-3-} \text{(tert-butoxycarbonyloxadizidine-4-carboxylic acid (1.1a)}^{21} \text{: NaOH (2M, 8.4mL) and 37\% formaldehyde (1.68 mmol) were added to serine (16.8 mmol) at 0\degree C. The reaction was stirred for 7 h at 0\degree C upon which NH}_2 \text{OHCl (1.68 mmol), acetone (9.8 mL), and NaOH (1.2 M, 4.2 mL). The reaction was stirred for an additional 15 at 0\degree C and then warmed to room temperature. (Boc)}_2 \text{O (18.15 mmol) was added and the reaction was allowed to stir at room temperature for 3 h. The reaction was diluted with H}_2 \text{O, washed with ether (3 x 10 mL), and the aqueous layer was acidified with citric acid (20\% solution) till the pH was 4. Finally this acidic aqueous layer was extracted with EtOAc (3 x 10 mL), dried with MgSO}_4 \text{ and isolated by rotary evaporation to give viscous oils that solidified to white solids after leaving overnight. Yield = 88\%} \text{.} \]

$^1$H NMR (400 MHz, DMSO) $\delta$ 12.80 (br s, 1H), 4.81 (d, 1H, $J = 3.6$ Hz), 4.78 – 4.74 (m, 1H), 4.25 – 4.13 (m, 2H), 4.00 (s, 1H), 1.38 (s, 9H).

\[ \text{(4S,5R)-3-} \text{(tert-butoxycarbonyl)-5-methyloxazolidine-4-carboxylic acid (1.1b)}^{21} \text{: NaOH (2M, 8.4mL) and 37\% formaldehyde (1.68 mmol) were added to threonine (16.8 mmol) at 0\degree C. The reaction was stirred for 7 h at 0\degree C upon which NH}_2 \text{OHCl (1.68 mmol), acetone (9.8 mL), and NaOH (1.2 M, 4.2 mL). The reaction was stirred for an additional 15 at 0\degree C and then warmed to room temperature. (Boc)}_2 \text{O (18.15 mmol) was added and the reaction was allowed to stir at room} \]
temperature for 3 h. The reaction was diluted with H₂O, washed with ether (3 x 10 mL), and the aqueous layer was acidified with citric acid (20% solution) till the pH was 4. Finally this acidic aqueous layer was extracted with EtOAc (3 x 10 mL), dried with MgSO₄ and isolated by rotary evaporation to give viscous oils that solidified to white solids after leaving overnight. Yield = 84%.

\( ^1H \) NMR (400 MHz, DMSO) \( \delta \) 12.81 (br s, 1H), 4.99 (d, 1H, \( J = 3.6 \) Hz), 4.68 (s, 1H), 4.12 (q, \( J = 6.2 \) Hz), 3.80 – 3.75 (m, 1H), 1.40/1.37 (s, 9H), 1.34 (t, \( J = 6.2 \) Hz, 3H). \(^a^\) signals from two rotamers.

**\( R \)-thiazolidine-4-carboxylic acid (1.4):** Cysteine (1221 mg, 1 mmol) was added to a 2 mL eppendorff tube and 37% formaldehyde (400uL) was added. The reaction was sonicated and a white precipitate was observed after five minutes. The reaction was centrifuged and the supernatant removed. Successive rinses with EtOH (3 x 1 mL) and Et₂O afforded the thiazolidine as a white powder (118 mg, 89%). \(^1H\) NMR (400 MHz, D₂O) \( \delta \) 4.52 – 4.46 (m, 2 H), 4.40 – 4.37 (m, 1H), 3.48 – 3.34 (m, 2H). \(^{13}C\) NMR (400 MHz, D₂O) \( \delta \) 171.2, 100.0, 64.0, 48.6, 33.0.

**\( 5S, \ 5aS, \ 8aR \)-N-(\textit{tert}-butyl)-8-oxotetrahydro-1H,3H,5H-azirino[1,2-\textit{a}]thiazolo[3,4-\textit{d}]pyrazine-5-carboxamide (1.5):** Cysteine pseudo proline 1.4 (21.3 mg, 0.16 mmol) and (S)-aziridine aldehyde dimer (12.5 mg, 0.088 mmol) was dissolved in TFE (0.8 mL, 0.2M). Immediately, tBuNC (19.9uL, 0.18 mmol) was added to the mixture and the reaction as monitored by TLC (SiO₂; 100% EtOAc; \( R_f = 0.5 \); ninhydrin stain). After completion, 30 min, the reaction mixture was concentrated to dryness, taken up in EtOAc and purified by column chromatography (SiO₂, 100% EtOAc) to give an oil (26 mg, 64 %, 9:1 diastereoselectivity \textit{syn:anti}). \(^1H\) NMR (CDCl₃, 500 MHz) 6.24 (br s, 1H), 4.05 (dd, 7.2 Hz, 0.6 Hz, 1H), 3.57 (d, 6.5 Hz, 1H), 3.52 (d, 7.2 Hz, 1H), 3.36 – 3.30 (m, 3H), 3.14 (ddd, \( J = 6.5 \) Hz, 4.8 Hz, 4.1 Hz, 1H),
3.07 (dd, $J = 10.0, 6.2$ Hz, 1H), 2.48 (d, $J = 4.8$ Hz, 1H), 2.28 (dd, $J = 4.2, 0.6$ Hz, 1H), 1.35 (s, 9H). $^{13}$C NMR (CDCl$_3$) 181.3, 167.7, 64.5, 63.9, 60.0, 51.6, 37.3, 31.2, 28.9, 28.6.

$^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 8.67 (s, 2H), 4.29 (q, $J = 7.3$, 1H), 3.79, (d, $J = 7.2$ Hz, 3H).

$^{13}$C NMR (400 MHz, CDCl$_3$) $\delta$ 170.7, 53.3, 49.4, 16.1.

**Methyl-L-alaninate hydrochloride (1.6)$^{12}$**: L-alanine (1.00 g, 11.2 mmol) in MeOH (5.0 mL) was cooled to $0^\circ$C and SOCl$_2$ was slowly added. The cloudy white mixture was warmed to room temperature then heated to reflux for 3 h and the solution turned clear and colorless. The condenser was removed and the reaction was cooled to room temperature and allowed to stir at rt overnight. Reaction completion was monitored by TLC (SiO$_2$, ninhydrin stain), product rf = 0.59 (1-propanol: acetic acid: H$_2$O / 12:3:5). The reaction was concentrated to an oil under vacuum and triturating with hexanes (5 mL) gave a white powder which was filtered and washed with ether (3 x 5 mL) to give a white powder (1.451 g, 93%). $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 8.67 (s, 2H), 4.29 (q, $J = 7.3$, 1H), 3.79 (s, 3H), 3.71, (d, $J = 7.2$ Hz, 3H). $^{13}$C NMR (400 MHz, CDCl$_3$) $\delta$ 170.7, 53.3, 49.4, 16.1.

**Tert-butyl (R)-4-(((S)-1-methoxy-1-oxopropan-2-yl)carbamoyl)thiazolidine-3-carboxylate (1.8)$^{8,12,14}$**: Alanine methyl ester 1.6 (442 mg, 4.287 mmol), Boc-thiazolidine 1.2a (1 g, 4.287 mmol), EDC •HCl (986 mg, 5.144 mmol), HOBt•6HCl (872.4) were added to a round bottom
flask. Freshly distilled THF (20 mL) and Et$_3$N (1.2 mL) were added and the reaction was allowed to stir at room temperature for 24 h. The pale brown cloudy mixture was then concentrated to a brown residue by rotary evaporation, which was then diluted with EtOAc (40 mL) and washed with H$_2$O (40 mL), 5% aqueous citric acid (4 x 12 mL), 5% aqueous Na$_2$CO$_3$ (2 x 20 mL). The organic layer was dried with anhydrous Na$_2$SO$_4$, and concentrated by rotary evaporation to give a brown oil (585.6 mg, 43%). TLC (SiO$_2$, Hexanes: EtOAc, 50:50) Rf = 0.35 visualized with KMnO$_4$ stain.\(^1\)H NMR (400 MHz, CDCl$_3$): \(\delta\) 6.86 (br s, 1H), 4.66 – 4.59 (m, 2H), 4.54 – 4.49 (m, 1H), 4.31 (d, \(J = 9.5\), 1H), 3.69 (s, 3H), 3.31 (br s, 1H), 3.17 (br s), 1.98 (s, 1H), 1.43 (s, 9H), 1.34 (d, \(J = 7.2\) Hz). \(^13\)C NMR (400 MHz, CDCl$_3$): peaks in brackets correspond to rotamers \(\delta\) 172.9, 170.0, 154.0, 82.0, 62.9, 52.5, 49.6, 48.2, (34.7, 32.9), 28.2, 18.4.

\[
\text{((R)-3-(t} \text{ert-butoxycarbonyl)thiazolidine-4-carbonyl-L-alanine (1.9)}^{13}\text{): Boc-thiazolidine alanine methyl ester 1.8 (300 mg, 0.942 mmol) was dissolved in EtOH (6.4 mL) to give a pale yellow solution. NaOH (1 N, 1.23 mL) was added and a white precipitate formed. The reaction was left to stir for 24 h upon before concentrating to an oily residue by rotary evaporation. Distilled water (8 mL) was added to the residue followed by KHSO$_4$ (5%w/w in H$_2$O) to give a pH = 3. The acidic aqueous solution was extracted with EtOAc (3 x 4 mL), dried with anhydrous sodium sulfate, and concentrated to a white crystalline solid by rotary evaporation (189.6 mg, 66%). \(^1\)H NMR (CD$_3$OD, 400 MHz) \(\delta\) 8.32 (br s, 1H, COOH), 4.64 (d, \(J = 9.2\) Hz, 1H), 4.54 (br s, 1H), 4.48 – 4.40 (m, 2H), 3.40 (br s, 1H), 3.14 (dd, \(J = 11.9, 5.4\) Hz, 1H), 1.46 (s, 9H), 1.41 (d, \(J = 8.12\), 3H). \(^1\)H NMR (CD$_3$OD, 600 MHz) \(\delta\) 4.65 (d, \(J = 9.2\) Hz, 1H), 4.56 (m, 1H), 4.45 (d, 1H), 4.40 (q, \(J = 7.3\) Hz, 1H), 3.38 (dd, \(J = 12.1\) Hz, 7.5 Hz, 1H), 3.16 (d of d, \(J = 11.9\) Hz, 5.1 Hz, 1H), 1.46 (s, 9H), 1.42 (d, \(J = 7.3\) Hz, 3H). \(^13\)C NMR (CD$_3$OD, 600 MHz) \(\delta\)175.4, 172.9, 155.2, 82.7, 66.8, 64.0, 50.7, 28.6, 35.4, 27.1, 16.3.
(R)-4-(((S)-1-carboxyethyl)carbamoyl)thiazolidin-3-ium 2,2,2-trifluoroacetate (1.10): Boc-thiazolidine alanine 1.9 (100 mg, 0.33 mmol) was dissolved in trifluoroacetic acid (1 mL) and stirred at room temperature for 1 h. The reaction was monitored using mass spectroscopy. Upon completion, the mixture was concentrated to dryness under vacuum and gave a pale-yellow oil. Trituration of this oil with diethyl ether gave a white crystalline solid (95.4 mg, 91%). MS (ESI): m/z = 205.1 [M+H] \(^1\)H NMR (MeOD, 400 MHz) \(\delta\) 4.58 (t, 7.4 Hz, 1H), 4.45 (m, 2H), 3.60 (d of d, 12.0, 7.7 Hz, 1H), 3.29 (d of d, 12.0, 7.1 Hz, 1H), 1.44 (d, 7.4 Hz, 3H). \(^1^3\)C NMR (MeOD, 400 MHz) 175.2, 167.4, 63.4, 50.0, 49.8, 34.5, 17.3.

(R)-3-((benzyloxy)carbonyl)thiazolidine-4-carboxylic acid (1.11): Thiazolidine 1.4 (100 mg, 0.75 mmol) was dissolved in an aqueous solution of KOH (1.15 mL, 1.5 M) and benzyl chloroformate (118 uL, 0.83 mmol) was added along with another addition of KOH (1.0 mL). The reaction was stirred overnight at room temperature. The reaction was washed with diethyl ether (2 x 5 mL), then acidified with a dropwise addition of 6M HCl to achieve pH < 3 whereupon the acidic aqueous phase was extracted with ethyl acetate (3 x 5 mL), dried with Na\(_2\)SO\(_4\) and concentrated to dryness to yield a clear colourless oil (189.7 mg, 95 %). \(^1\)H NMR (DMSO, 400 MHz) \(\delta\) 12.8 (br s, 1H), 7.38 – 7.30 (m, 5H), 1.89 (s, 2H), 4.77 (dd, \(J = 7.2\) Hz, 3.2 Hz, 1H), 4.62 (d, \(J = 8.7\) Hz, 1H), 4.41 (d, \(J = 8.7\) Hz, 1H), 3.40 (dd, \(J = 11.5\) Hz, 7.1 Hz, 1H), 3.17 (dd, \(J = 11.5\) Hz, 2.7 Hz, 1H).

(R)-3-((benzyloxy)carbonyl)thiazolidine-4-carboxylic acid (1.11): Thiazolidine 1.4 (160 mg, 1.20 mmol) was dissolved in a cold aqueous solution of NaOH (1.00 mL, 2.0 M). Keeping the mixture between 0-5°C, benzyl chloroformate (118 uL, 0.83 mmol) was added dropwise with vigorous stirring over 10 min concurrently with another addition of NaOH (1.00 mL). The reaction was stirred for an additional 2 h whereupon the cloudy white suspension was washed with diethyl ether (2 x 5 mL). To the aqueous layer, concentrated HCl was added dropwise to
achieve pH < 3 and the resulting oily white precipitate was extracted with ethyl acetate (3 x 5 mL), dried with Na₂SO₄ and concentrated to dryness via rotary evaporation to yield a clear colourless oil (281.7 mg, 87.9 %). ¹H NMR (DMSO, 400 MHz) 12.8 (br s, 1H), 7.38 – 7.30 (m, 5H), 1.89 (s, 2H), 4.77 (dd, J = 7.2 Hz, 3.2 Hz, 1H), 4.62 (d, J = 8.7 Hz, 1H), 4.41 (d, J = 8.7 Hz, 1H), 3.40 (dd, J = 11.5 Hz, 7.1 Hz, 1H), 3.17 (dd, J = 11.5 Hz, 2.7 Hz, 1H). ¹³C NMR (CDCl₃, 125 MHz) δ 174.5, 154.5, 135.9, 128.7, 128.5, 128.2, 68.3, (62.0, 61.2), (49.4, 48.3), (34.5, 33.1).

Benzyl (R)-4-(((S)-1-methoxy-1-oxopropan-2-yl)carbamoyl)thiazolidine-3-carboxylate (1.12): Cbz-thiazolidine 1.11, (100 mg, 0.42 mmol), alanine methyl ester hydrochloride 1.6 (43 mg, 0.42 mmol), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (98 mg, 0.51 mmol), and hydroxybenzotriazole hexahydrochloride (86 mg, 0.51 mmol) were dissolved in CH₂Cl₂ (5 mL) under nitrogen. Triethylamine (116 uL, 0.83 mmol) was added to the reaction mixture. The mixture was stirred overnight and monitored by TLC (Rf = 0.3 in 50:50, EtOAc: hexanes, SiO₂). Upon completion, the reaction was filtered to remove the insoluble urea salts then purified by column chromatography (SiO₂, 50:50 EtOAc: hexanes) to afford a clear colourless oil (26.6 mg, 18 %)¹H NMR (CDCl₃, 400 MHz) δ 7.34 (m, 5H), 6.79 (br s, 1H), 5.19 (s, 2H), 4.76 – 4.71 (m, 2H), 4.55 – 4.52 (m, 1H), 4.45 – 4.43 (m, 1H), 3.72 (s, 3H), 3.40 (m, 1H), 3.21 (m, 1H), 1.36 (d, J = 7.1 Hz, 3H).

((R)-3-((benzylloxy)carbonyl)thiazolidine-4-carbonyl)-L-alanine (1.13): 1.12 (26 mg, 0.07 mmol) was dissolved in MeOH (1 mL) and the reaction was cooled to 0°C whereupon NaOH
(2N, 45 uL) was added and the reaction was allowed to warm to room temperature and left o/n to go to completion. MeOH was removed by rotary evaporation and the mixture was acidified with 1M HCl and extracted with EtOAc (3 x 1.5 mL), and concentrated to afford a clear colorless oil (18.7 mg, 79%). $^1$H NMR (500 MHz, MeOD, as rotamers) δ 7.35 – 7.22 (m, 5H), 4.97 (dd, $J$ = 6.6, 3.0) and 4.94 (dd, $J$ = 6.9, 2.9) 1H, 4.65 (d, $J$ = 7.7) and 4.63 (d, $J$ = 7.9), 1H, 4.43 (d, $J$ = 8.0) and 4.41 (d, $J$ = 7.6) 1H), 4.31 (q, $J$ = 7.3) and 4.31 (q, $J$ = 7.3) 1H, 1.42 (d, $J$ = 7.4) and 1.40 (d, $J$ = 7.3), 3 H. $^{13}$C NMR (500 MHz, MeOD, as rotamers) 177.1 and 177.0, 157.8, 142.7, 129.3, 128.2, 128.0, 63.2, 63.2 and 63.0, 50.7, 48.4, 34.2 and 33.9, 18.1 and 17.9.

(S)-2-amino-3-phenylpropan-1-ol (1.14)$^{17}$: Synthesized according to literature procedure.$^{10}$ $^1$H NMR (CDCl$_3$, 400 MHz) 7.35 – 7.18 (m, 5 H), 3.63 (dd, $J$ = 10.6, 3.9 Hz, 1H), 3.38 (dd, $J$ = 10.6, 7.1 Hz, 1H), 3.12 (tt, $J$ = 8.7, 4.5 Hz, 1H), 2.79 (dd, $J$ = 13.5, 5.2 Hz, 1H), 2.52 (dd, $J$ = 13.5, 8.6 Hz, 1H), 1.86 (br s, 2H). $^{13}$C NMR (100 MHz, CDCl$_3$) δ 138.82, 129.33, 128.70, 126.54, 66.50, 54.30, 41.09.

(R)-3-(((9H-fluoren-9-yl)methoxy)carbonyl)thiazolidine-4-carboxylic acid (1.27)$^{22}$: Thiazolidine (133.02mg, 1.0 mmol) was dissolved in Na$_2$CO$_3$(aq) (9%, 6 mL) and cooled to 0°C. A solution of Fmoc-OSuc (1.2mmol) in dioxane was added to the thiazolidine solution and the mixture was allowed to warm to room temperature and left to stir for 3 h. Upon completion, the cloudy white suspension was diluted with water and washed with ether (2 x 10mL) then EtOAc (2 x 10 mL). The aqueous layer was acidified to pH = 2 with a dropwise addition of concentrated HCl over ice, whereupon a white oil precipitated from solution, which was subsequently extracted with EtOAc (6 x 10 mL), dried over anhydrous Na$_2$SO$_4$ and concentrated to dryness via rotary evaporation to yield a white crystalline solid (226.3 mg, 63.8%). $^1$H NMR (400 MHz, DMSO-d$_6$) 13.08 (br s, 1H), 7.89 – 7.34 (m, 8H), 4.78 – 3.56 (m, 6 H), 3.56 – 3.15 (m, 2H) $^{13}$C
NMR (100 MHz, DMOS-\textsubscript{d\textdelta} doubling due to rotameric mixture) 171.4, 153.1, 143.6, 140.7, 127.7, 127.2, 125.2, 120.1, 67.3, 61.5, 47.6 and 46.5, 34.0 and 32.6.

**N-benzyl-L-alanine (1.15)**\textsuperscript{18}: Synthesized according to literature procedure. **Work Up:** The reaction was diluted with H\textsubscript{2}O (15 mL), washed with CH\textsubscript{2}Cl\textsubscript{2} (2 x 30 mL). To the aqueous layer, conc. HCl was added to give a white precipitate, which was collected by vacuum filtration. The residue was washed with H\textsubscript{2}O (2 x 5mL). Although the reference suggested to also wash with acetone, this actually slightly dissolved the benzylated compound as a white solid (3.6 g, 96%)\textsuperscript{1H}
NMR (300 MHz, MeOD) exists as rotamers 8.04 – 7.43 (m, 5 H), 4.34 – 4.05 (m, 2 H), 3.72 (q, $J = 7.2$ Hz), 3.64 (q, $J = 7.2$ Hz), 1H; 1.52 (d, $J = 7.2$ Hz) and 1.49 (d, $J = 7.2$ Hz), 1H. $^{13}$C (125 MHz, MeOD) exists as rotamers 173.9, 134.0, 132.9, 131.0, 130.7, 130.5, 130.2, 129.4, (58.5 and 51.0), 16.2.

$^{1}$H NMR at 50°C (400 MHz, DMSO-$d_6$) 12.44 (br s, 1H), 8.05 – 7.08 (m, 5H), 4.45 (d, $J = 14.8$ Hz, 2H), 4.04 (q, $J = 7.1$ Hz, 1H), 1.38 (s, 9H), 1.24 (d, $J = 7.4$ Hz, 3H). $^{13}$C NMR (101 MHz, dmso) δ 172.9, 170.0, 154.6, 132.5, 130.7, 129.0, 128.3, 127.9, 126.5, 79.2, 59.5, 55.0, 49.9, 27.8.

$N$-benzyl-$N$-(tert-butoxycarbonyl)-L-alanine (1.16)$^{23}$: $N$-benzyl alanine (1.00 g, 5.6 mmol) was suspended in 1:1 water: dioxane (15 mL) and TEA (2.4 mL was added). A clear colourless solution was obtained after stirring for 20 min at rt. Boc2O (1.30 g, 6.0 mmol) was added and the reaction was allowed to stir at rt for 24 h. The resulting solution contained a white precipitate and rotary evaporator first removed the dioxane before the aqueous was washed with diethyl ether (2 x 5 mL). The clear aqueous layer was acidified with conc. HCl to pH=3 and a white oily precipitate resulted which was then extracted into EtOAc (3 x 5mL), dried over anhydrous sodium sulfate, filtered, and concentrated to a clear colourless oil by rotary evaporation. Yield = 794.5 mg, 51%. $^1$H NMR at 50°C (400 MHz, DMSO-$d_6$) 12.44 (br s, 1H), 8.05 - 7.08 (m, 5H), 4.45 (d, $J = 14.8$ Hz, 2H), 4.04 (q, $J = 7.1$ Hz, 1H), 1.38 (s, 9H), 1.24 (d, $J = 7.4$ Hz, 3H). $^{13}$C NMR (101 MHz, dmso) δ 172.9, 170.0, 154.6, 132.5, 130.7, 129.0, 128.3, 127.9, 126.5, 79.2, 59.5, 55.0, 49.9, 27.8.

$N$-((9H-fluoren-9-yl)methoxy)carbonyl)-$N$-benzyl-L-alanine (1.17)$^{14}$: $N$-benzylalanine (200 mg, 1.12 mmol) was suspended in a 9% Na$_2$SO$_4$ (5 mL). This was cooled to 0°C and Fmoc-Cl (347 mg, 1.34 mmol) in dioxane (5 mL) was slowly added. Following addition, the reaction was brought to room temperature and allowed to stir overnight. Dioxane was removed by rotary evaporation and the reaction was diluted with H$_2$O. The aqueous phase was washed with ether (3 x 5 mL), and then acidified with conc. HCl producing a white oily precipitate that was extracted
into EtOAc. The organic layer was then dried with anhydrous Na$_2$SO$_4$, filtered, and concentrated to dryness using a rotary evaporator to give a white powder. Yield = 151 mg, 34%. VT NMR was performed to characterize the product. $^1$H NMR at 50°C (400 MHz, DMSO-$d_6$) 7.94 - 7.14 (m, 13H), 4.43 – 4.03 (m, 6H), 1.24 – 1.13 (m, 3H). $^{13}$C NMR (101 MHz, dmso) δ 186.5, 144.2, 141.2, 133.2, 129.6, 128.9, 128.6, 128.0, 127.4, 120.4, 105.0, 65.1, 62.8, 50.1, 45.0, 44.0, 29.3.

Phenyl-$L$-phenylalanine (1.18)$^{24}$: Phenylalanine (100 mg, 0.60 mmol), Cul (12 mg, 10 mol%), and K$_2$CO$_3$ (124 mg, 0.90 mmol) were added to a flame-dried 5.00 mL microwave vial charged with a stir bar. The flask was evacuated and back-filled with N$_2$ three times before anhydrous degassed DMF (2mL) was added via syringe followed by bromobenzene (63 uL, 0.60 mmol). The sealed vial was then microwaved for 40 min. The blue reaction mixture was diluted with H$_2$O and concentrated HCl was added to give a bright yellow suspension. The product was extracted with ethyl acetate (3 x 2 mL), dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated by rotary evaporation to give a brown-yellow oil, which was subsequently purified by column chromatography (SiO$_2$, 40 – 100% EtOAc in hexanes, R$_f$ (EtOAc/Hexanes, visualized with UV = 0.37) to give an off-white solid: 119 mg, 83%. $^1$H NMR (500 MHz, DMSO-$d_6$) 12.56 (br s, 1H), 7.36 - 6.49 (m, 10H), 5.89 (br s, 1H), 4.11 (dd, $J$ = 8.3, 5.7 Hz, 1H), 3.07 (dd, $J$ = 13.7, 5.7 Hz, 1H), 2.97 (dd, $J$ = 13.7, 8.3 Hz, 1H). $^{13}$C NMR (125 MHz, DMSO-d6) 174.7, 147.7, 137.9, 129.2, 128.9, 128.1, 126.4, 116.4, 112.5, 57.4, 37.8.

(3S,5S,6S)-3-benzyl-$N$-(tert-butyl)-2-oxo-4-phenyl-$1,4$-diazabicyclo[4.1.0]heptane-$5$-carboxamide (1.19): N-arylphenylalanine (50 mg, 0.21 mmol), (S)-aziridine aldehyde dimer (18 mg, 0.12 mmol), and a stir bar were added to a small vial. TFE (1 mL) and tBuNC (28 uL, 0.25 mmol) were added and a pale yellow suspension resulted. The reaction was monitored by LCMS
and after two hours, the reaction was concentrated to an oil and purified by column chromatography (SiO₂, hexanes:EtOAc 1:1). The fractions were analyzed for product by LCMS and the relevant fractions were combined, dried with anhydrous Na₂SO₄, filtered, and concentrated by rotary evaporation. Yield = 26.8 mg, 34%. The white solid was dissolved in a minimal acetone and slow diffusion of hexanes gave white needle like crystals. Analysis by X-ray crystallography proved the product identity as well as stereochemistry (S, S, S). ¹H NMR (500 MHz, DMSO-d₆) 7.38 – 6.83 (m, 10H, ArH), 4.74 (d, J = 7.4 Hz, 1H), 4.56 (dd, J = 8.6 Hz, 2.0 Hz, 1H), 3.59 (dd, J = 14.6 Hz, 8.5 Hz, 1H), 3.25 (ddd, J = 7.5 Hz, 4.3 Hz, 3.2 Hz, 1H), 2.71 (dd, J = 14.6 Hz, 2.0 Hz, 1H), 2.38 (d, J = 4.3, 1H), 2.27 (d, J = 3.2 Hz, 1H). ¹³C NMR (125 MHz, (CD₃)₂CO) 184.0, 169.5, 148.8, 141.5, 130.2, 129.0, 126.8, 120.4, 117.6, 61.7, 57.8, 51.9, 37.1, 36.8, 30.9, 28.8.

**Methyl phenyl-L-phenylalanyl-L-valinate (1.20):** N-aryl phenylalanine (30 mg, 0.12 mmol), valine methyl ester hydrochloride (21 mg, 0.12 mmol), EDC HCl (28 mg, 0.14 mmol), HOBt 6HCl (24 mg, 0.14 mmol), and Et₃N (33 uL, 0.24 mmol) were dissolved in dichloromethane (5 mL) and left to stir at rt. for 24 h. The reaction was concentrated to an oil then diluted with ethyl acetate (10 mL). The organic layer was washed with water (2 x 5 mL), 5% citric acid (2 x 5 mL), 5% Na₂CO₃ (2 x 5 mL). The organic layer was then dried over anhydrous sodium sulfate, filtered, and evaporated under reduced pressure to give a yellow crystalline solid. Yield = 31.5 mg, 74%. ¹H NMR (400 MHz, CDCl₃) 7.26 – 6.49 (m, 10H), 4.43 (dd, J = 8.9, 5.2 Hz, 1H), 3.94 (dd, J = 8.5, 4.7 Hz, 1H), 3.52 (s, 3H), 3.25 (dd, J = 14.2, 4.6 Hz, 1H), 2.98 (dd, J = 14.1, 8.5 Hz, 1H), 2.06 (pd, J = 6.8, 5.0 Hz, 1H), 0.80 (d, J = 6.9 Hz, 3H), 0.76 (d, J = 6.9 Hz, 3H). ¹³C NMR (100 MHz, CDCl₃) δ 173.0, 171.7, 146.7, 136.6, 129.3, 129.2, 129.0, 127.3, 119.6, 114.6, 77.2, 60.8, 57.3, 52.0, 38.8, 31.1, 19.1, 18.0.

**Phenyl-L-phenylalanyl-L-valine (1.21):** N-aryl-Phe-Val-OMe (31.5 mg, 0.089 mmol) was dissolved in MeOH and 2N NaOH was added and the reaction was left to stir o/n at room
temperature. By LCMS, the reaction was determined complete the next morning. The mixture was concentrated by rotary evaporation to remove MeOH and then diluted with water. Concentrated HCl was added dropwise to acidify the aqueous before extracting it with EtOAc (3 x 2 mL), drying over anhydrous sodium sulfate, and concentrating to a yellow solid. Yield = 25.8 mg, 85%. $^1$H NMR (300 MHz, CDCl$_3$) 7.27 – 6.61 (m, 10H), 4.34 – 4.25 (m, 1H) 4.17 (dd, $J = 8.4$, 5.4 Hz, 1H), 3.18 (dd, $J = 13.9$, 5.4 Hz, 1H), 2.99 (dd, $J = 13.9$, 8.3 Hz, 1H), 2.10 (pd, $J = 6.8$, 5.4 Hz, 1H), 0.89 (d, $J = 3.6$ Hz, 3H), 0.86 (d, $J = 3.6$ Hz, 2H). $^{13}$C NMR (75 MHz, CDCl$_3$) 174.1, 160.0, 148.6, 138.8 130.3, 130.0, 129.4, 127.7, 119.4, 115.1, 61.1, 58.8, 39.9, 32.0, 19.5, 18.2

**Methyl (tert-butoxycarbonyl)glycyl-L-valinate (1.22):** Boc-glycine (1.00 g, 5.7 mmol), valine methyl ester hydrochloride (955 mg, 5.7 mmol), EDC HCl (1.32 g, 6.9 mmol), HOBt 6HCl (1.2 g, 6.9 mmol), TEA (1.6 mL, 11.4 mmol), were dissolved in CH$_2$Cl$_2$ (30 mL) and left to stir for 24 at rt. The reaction was then concentrated to dryness by rotary evaporation and diluted with EtOAc (15 mL), washed with H$_2$O (2 x 10 mL), 5% citric acid (2 x 10 mL), and 5% Na$_2$CO$_3$ (2 x 10 mL). The organic layer was dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated to an oil. Purification by column chromatography (SiO$_2$, 1:1 hexanes/ethyl acetate, Rf = 0.41, visualized with ninhydrin) gave colourless oil. Yield = 1.25 g, 76%. $^1$H NMR (400 MHz, CDCl$_3$) 4.56 (d, 6.7 Hz, 1H), 5.11 (br s, 1H), 4.56 (dd, $J = 6.7$, 1.8 Hz, 1H), 3.86 (td, $J = 16.7$, 16.1, 5.8 Hz, 1H), 0.94 (d, $J = 6.8$ Hz, 3H), 0.90 (d, $J = 6.8$ Hz, 3H). $^{13}$C NMR (125 MHz, CDCl$_3$) $\delta$ 172.4, 169.7, 156.2, 57.1, 52.3, 44.8, 31.3, 28.4, 19.0, 17.8.

**Glycyl-L-valine (1.23):** Boc-Gly-Val-OMe (1.25g, 4.34 mmol) was stirred at room temperature overnight in TFA/TIPS/H$_2$O (95/2.5/2.5, 10 mL) and the reaction monitored by TLC (Rf of SM = 0.41 in 1:1 EtOAc/Hexanes). The reaction was concentrated to dryness by rotary evaporation and 2M NaOH and MeOH were added. The reaction was vigorously stirred and monitored by LCMS. Upon completion, the reaction was again concentrated by rotary evaporation to remove
MeOH. The basic aqueous solution was washed with EtOAc. The aqueous layer was concentrated to dryness and the residue was taken up in MeOH, dried with Na₂SO₄, filtered, and concentrated to an oil by rotary evaporation (760 mg, 93%). ¹H NMR (400 MHz, DMSO-d₆) 3.80 (dd, J = 17.8, 0.9 Hz, 1H), 3.62 (d, J = 17.7 Hz, 2H), 3.52 (d, J = 4.1 Hz, 1H), 2.10 (pd, J = 6.8, 4.0 Hz, 1H), 0.92 (d, J = 7.0 Hz, 3H), 0.85 (d, J = 6.8 Hz, 3H. ¹³C NMR (100 MHz, DMSO-d₆) 167.2, 166.1, 59.7, 44.0, 39.5, 32.3, 18.6, 17.1.

2-(tert-butyl)imidazolidin-4-one (1.24)²⁰: A 100 mL round bottom was fitted with a soxhlet apparatus and was flame-dried under vacuum and allowed to cool under nitrogen. Glycine amide hydrochloride (500 mg, 4.52 mmol) suspended in CH₂Cl₂, pivaldehyde (746 uL, 6.78 mmol), and freshly distilled TEA (946 uL, 6.78 mmol) were added. Four-angstrom molecular sieves were added to the thimble for continual removal of water and the reaction was set to reflux overnight. The reaction was then cooled to 0°C, and any TEA HCl salts were removed by filtration. The filtrate was diluted with CH₂Cl₂ and TFA (522 uL, 6.78 mmol) was added and the reaction was allowed to stir for 24 h resulting in a white precipitate. 2N NaOH (40 mL) was added to give a clear solution. The aqueous layer was then extracted with CH₂Cl₂ (3 x 10 mL) This was dried over anhydrous sodium sulfate, filtered, and concentrated by rotary evaporation to give a white crystalline solid. Yield = 597 mg, 93%. ¹H NMR (400 MHz, Chloroform-d) 8.3 (s, 1H), 4.29 (s, 1H), 3.54 - 3.30 (m, 2H), 2.18 (s, 1H), 0.88 (s, 9H). ¹³C NMR (100 MHz, cdcl₃) δ 177.8, 80.1, 49.7, 35.6, 24.4.

Methyl glycyglycinate hydrochloride (1.25)²⁵: Gly-Gly (1.00g, 7.57 mmol) was suspended in MeOH, cooled to 0°C with an ice-bath, and SOCl₂ (823 uL, 11.35 mmol) was added dropwise to give a clear solution. The mixture was allowed to warm to room temperature and stir for another 3 h. Concentration via rotary evaporation gave a white solid, which was washed with ether. Upon drying under vacuum, a white crystalline powder resulted. Yield = 1.31 g, 94%. ¹H NMR (300
MHz, DMSO-d$_6$) 9.02 (t, $J$ = 5.9 Hz, 1H), 8.31 (s, 3H), 3.94 (d, $J$ = 5.8 Hz, 2H), 3.64 (s, 3H), 3.59 (s, 2H). $^{13}$C NMR (75 MHz, dmso-d$_6$) δ 169.9, 166.5, 51.9, 40.5, 39.9.

GLGF-O-trityl-resin

1. X, HATU, DIPEA, NMP

2. HFIP/DCM (1/4)

$X = $ Fmoc-Thz
$X = $ N-Bz-Ala
$X = $ N-Ar-Phe

**Example procedure for Fmoc-Thz:** The linear peptide on resin (GLGF-resin, 50 mg, 1.0 equiv.) was added to a shaker tube along with NMP (5 mL), HATU (30 mg, 0.078 mmol, 1.1 equiv.), and Fmoc-Thz (28 mg, 0.078 mmol, 1.1 equiv.). The shaker tube was shook for 1 h at room temperature. The reaction was monitored by LCMS by micro-cleaving a small aliquot with 25% HFIP in CH$_2$Cl$_2$. Upon completion as determined by LCMS, the peptide was washed with NMP (3 x 5 mL). 20% piperidine in NMP (5 mL) were added and the cleavage of Fmoc from Thz was again monitored by LCMS using micro-cleavage of a small aliquot using 25% HFIP in CH$_2$Cl$_2$. Upon completion of the Fmoc-deprotection reaction, the peptide was again rinsed well with NMP. 25% HFIP in CH$_2$Cl$_2$ (2 mL) was added and the shaker tube was allowed to shake for an additional 15 min before the reaction was filtered. The residue was washed with CH$_2$Cl$_2$ and the filtrate was concentrated to dryness and cold MTBE was added to the oily residue to triturate a solid linear peptide. The mixture was centrifuged, the supernatant removed, and the pellet triturated again with MTBE. This last step was repeated a total of three times and the peptide dried under high-vacuum overnight to remove excess MTBE. The final product was identified by LCMS (m/z = 508.2 = M+H)$^+$.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sequence</th>
<th>m/z</th>
<th>Amount (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Thz-GLGA</td>
<td>432.2 = (M+H)$^+$</td>
<td>19.3</td>
</tr>
<tr>
<td>B</td>
<td>Thz-LGF</td>
<td>451.2 = (M+H)$^+$</td>
<td>28.2</td>
</tr>
<tr>
<td>C</td>
<td>Thz-F</td>
<td>281.8 = (M+H)$^+$</td>
<td>3.8</td>
</tr>
<tr>
<td>D</td>
<td>NBzA-GLGA</td>
<td>479.3 = (M+H)$^+$</td>
<td>20.4</td>
</tr>
<tr>
<td>E</td>
<td>NBzA-LGF</td>
<td>497.2 = (M+H)$^+$</td>
<td>24.5</td>
</tr>
<tr>
<td>F</td>
<td>NBz-F</td>
<td>327.2 = (M+H)$^+$</td>
<td>4.1</td>
</tr>
<tr>
<td>G</td>
<td>NArF-GLGA</td>
<td>540.3 = (M+H)$^+$</td>
<td>19.3</td>
</tr>
<tr>
<td>H</td>
<td>NArF-LGF</td>
<td>559.2 = (M+H)$^+$</td>
<td>28.1</td>
</tr>
<tr>
<td>I</td>
<td>NArF-F</td>
<td>389.2 = (M+H)$^+$</td>
<td>5.8</td>
</tr>
</tbody>
</table>
1.5 References


Chapter 2. Designing and Synthesis of Cyclic Peptides as Therapeutic Agents: Targeting EZH2/EED and Bcr-Abl

2.1 Introduction

2.1.1 Cyclic peptides as novel therapeutic targets

The area of cyclic peptides is rapidly expanding and has fueled a surge of interest in academic research and catalyzed the growth of new pharmaceutical-based biotech companies.¹ Excitement for macrocyclic peptides comes from their potential to target diseases currently inaccessible by small molecules and larger biological therapies.² Intersecting the boundary of biological and small molecule therapies, macrocyclic peptides are small enough to achieve cell permeability and yet large enough to extend over a druggable region of a protein for enhanced affinity and selectivity. Furthermore, macrocyclic peptides form native peptide interactions with protein surfaces offers enhanced binding associations. Macrocyclic peptides overcome large entropic barriers to bind proteins, as they are pre-organized in a cyclic structure. There are several examples of macrocyclic peptides as therapeutic agents. Rapamycin is a macrocyclic protein that binds to the protein FRAP22.³ Macrocycles laulimalide and peloruside A recognize tubulin.⁴ Cyclosporine, an 11-amino acid containing macrocycle was proved to successfully target cyclophilin A and is now prescribed as an immunosuppressant by Novartis AG.¹ While macrocycles have vast therapeutic potential, they are limited from the lack of high-throughput synthetic methodology. Using our lab’s macrocyclic methodology, previously explained in chapter one, several macrocycles can be synthesized and then analyzed for their ability to interact with relevant proteins. Two proteins will be the focus of this research section: the polycomb protein embryonic ectoderm development, EED, as well as the oncogenic fusion tyrosine kinase, break point cluster Abelson, Bcr-Abl.

2.1.2 Targeting the protein-protein interaction of EZH2 and EED

EED is a noncatalytic subunit of the polycomb repressive complex, PRC2. PRC2 modifies chromatin by inducing changes on histones. One PRC2’s most well studied mechanism of action is the complex’s ability to trimethylate histone-3 on lysine-27 (H3K27me3).⁵ H3K27me3 is
associated with the transcriptional silencing of genes that code for tumor suppressor proteins. The role of PRC2 in cancer epigenetics has been well researched\textsuperscript{6,7,8}. Methylation activity is controlled by four subunits that make up PRC2: RbAp48, SUZ12, EZH2, and EED. EZH2 contains the SET domain that ‘writes’ methylations on lysine.\textsuperscript{5}

![Figure 2.1 PRC2 Complex\textsuperscript{9}]

The noncatalytic subunit EED is a beta-transducin or WD40 (Trp-Asp) repeat protein that has a domain that folds into a beta propeller, which acts as a surface to which multiple protein complexes can be reversibly bind and assemble.\textsuperscript{10} One face of the WD propeller of EED binds the N-terminal region of EZH2 while the central pocket of the WD repeat propeller recognizes trimethyllysine through its aromatic cage.

![Figure 2.2 Left: Top view of EED top view. PDB Code 3JZN\textsuperscript{11} Right: EZH2 alpha helix (red) bound to EED. PDB Code 2QXY\textsuperscript{12}]

Furthermore, EED is specific for trimethyllysine that are within the sequence ARKS, unique to lysine residues K9 and K27, both of which are associated with gene silencing. When EED binds to trimethyllysine, the activity of EZH2 increases leading to amplified histone methylation. Therefore, EED acts as a factor in conserving H3K27me3 as well as propagating methylation in the chromatin region surrounding this mark. Additional observations whereby loss of the EED subunit in PRC2 led to severe inhibition of EZH2 activity suggests that the role of EED may be more substantial.

2.1.3 Bcr-Abl and the SH2 kinase domain

The phosphorylation of serine, threonine, and tyrosine residues is a well-studied mechanism that can turn the function of an enzyme on or off, thereby altering a wide range of cellular processes. Disrupting these highly interconnected networks results in downstream signal modifications and global alternations. Abl is a protein tyrosine kinase and is associated with cell division and cell homeostasis. Within Abl there exists an SH3-SH2-TK (Scr homology 3-Src homology 2-tyrosine kinase) domain. The SH3 domain controls Abl’s tyrosine kinase activity.

Abl exists in two conformations: a closed inactive conformer as well as an open active conformer (Figure 2.3). The inactive closed conformation is stabilized by several intramolecular interactions. The SH3 domain interacts with the proline rich motif in the SH2-kinase linker, the SH2 domain interacts with the C-lobe of the kinase domain, and the myristoylation region binds to a deep hydrophobic pocket in the C-lobe. The N-terminus of Abl, denoted as the Cap region, also interacts with the SH3-SH2 kinase connector. When one of these interactions is disrupted it becomes energetically favourable for Abl to disrupt all closed form intramolecular interactions and open up into the active form. This process is referred to as catalytic activation.
Chronic myeloid leukemia, CML, has been found to be associated with the deregulation of cytoplasmic Abl kinases. In CML, the chromosome that contains the gene that codes for Abl is translocated next to the BCR gene to produce the Philadelphia chromosome. This Philadelphia chromosome codes for the Bcr-Abl oncoprotein, which is constantly phosphorylated on Tyr245 and Tyr412 leading to stabilization of the open active form of Abl driving aberrant phosphorylation activity. Current prescribed inhibitors for CML inhibit the ATP binding site on Abl. Imatinib is prescribed as a first-line treatment for CML, however, many patients become resistant to this treatment due mutations such as the phosphorylation of Abl by tyrosine kinase SFK, which has been found to reduce the drugs potency by 50-fold. Current pharmaceutical research efforts target mutant phosphorylated Abl (dasatinib).

2.1.4 Research Objectives

Macro cyclic peptides will be designed and computationally modeled to best-fit EED in order to disrupt the proteins binding to EZH2. Computational analysis will be performed by Dr. Matthieu Schapira and Dr. Melissa Landon at the Structural Genomics Consortium (SGC). The potential of cyclic peptides to interrogate both Abl’s activation loop and the SH2 kinase domain will be studied by computational modeling in-house and at SGC. Synthesis of the cyclic peptides will be conducted using the methodology developed by the Yudin group. Finally, the Pawson research group will perform biochemical analysis of our compounds with Abl.
2.2 Results

2.2.1 Computational Modeling of Cyclic Peptides with EZH2/EED

Hotspot analysis was done using computational solvent mapping (CS-Map).\textsuperscript{16} This algorithm docks small organic functional groups in various positions around a protein surface to locate their most energetically favorable binding region. This method has been proven to reproduce the results of multiple solvent crystal structures that involve soaking a protein in a series of organic solvents then identifying patterns in the regions where the solvents tend to crystallize with the protein. Furthermore, this method has been used on proteins with known binding regions. CS-Map predicted those same areas as druggable regions.\textsuperscript{16}

![Figure 2.4 CS-Map of EED binding pocket](image)

The surface of the protein EED was computationally scanned for binding sites with 16 small molecules for energetically favourable binding regions.\textsuperscript{17,18} These small molecules include: ethanol, isopropanol, isobutanol, acetone, acetaldehyde, dimethyl ether, cyclohexane, ethane, acetonitrile, urea, methylamine, phenol, benzaldehyde, benzene, acetamido, and $N,N$-dimethylformamide.\textsuperscript{17} Small fragments were found to preferentially bind a pocked located at the EZH2/EED interface (Figure 2.4). Intriguingly, the collection of small fragments adopt a cyclic distribution at the binding local reinforcing our original idea of using a cyclic peptide. This favourable binding region was then screened with various cyclic peptides.
Our original plan was to dock 9-membered macrocyclic peptides in the identified binding region of EED. Scaffolds of 9-membered cyclic peptides incorporating the components of our unique cyclization methodology were created (see Chapter 1).

**Figure 2.5** 9-membered cyclic peptide scaffolds used for computational modeling.

From these five templates with 20 different possible R1 amino acid residues, 100 cyclic peptides were generated. All conformers of each cyclic peptide were generated which gave a total of 5521 structures to computationally model in the binding pocket using GOLD (Genetic Optimization for Ligand Docking) program. The mean docking score of these compounds was -9.5 where a lower number indicated a stronger affinity. In order to filter all the conformers, those that satisfied a score that were more than or equal to one standard of deviation less than the mean were retained. Further analysis based on structural analysis and mode of binding resulted in a succinct list of macrocycles for synthesis (Figures 2.6 – 2.8).

**Figure 2.6** Gaussian distribution of docked cyclic peptides and **Figure 2.7** Workflow to screen cyclic peptides by computational analysis
Figure 2.8 9-membered rings with best docking score

An in-depth analysis of the mode of binding between EED and EZH2 was done. A recent article used fragment analysis of EZH2 to understand which points of contact were the more potent and some important observations where made. Although EZH2 makes a fair amount of contact along EED, only residues 39-68 of EZH2 are important for activity. In fact, even if only these residues bind and the rest of EZH2 is not making any contact, methyltransferase activity can exist. The druggable region on EED that makes contact with EZH2, that we are attempting to interrupt with cyclic peptides, is where the C-terminus of EZH2 binds. Based on the EZH2 fragment analysis, intersecting this area of contact has a high possibility of allowing EZH2 to bind on its N-terminus and induce methyltransferase activity.

Based on these findings as well as the ambiguity of the computational work, this project was deprioritized. Further computational modeling should be done on a potential druggable area of EED where residues 39 – 68 of EZH2 make contact. Whether or not this area of EED, if deemed druggable, would be suitable for a cyclic peptide needs considering.
2.2.2 Bcr-Abl activation loop inhibitors (7a.a. cyclic peptides)

Together with Dr. Conor Scully, we found that 7a.a. cyclic peptides best mimic the native beta-turn of the Abl activated loop conformation (Figure 2.9). Mimicking the activation loop in its active conformation will stabilize the inactive form of Abl.

**Figure 2.9** *Left:* Crystal Structure Abl activation loop (highlighted in yellow) PDB Code 2FO0\(^{19}\)  
*Right:* native loop of Beta-turn and cyclic peptide mimic

To test this hypothesis a 7a.a. cyclic peptide was first synthesized. Using SPPS, PLMT(OtBu)GD(OtBu)T(OtBu), **2.1**, was synthesized and purified. Various cyclization conditions were screened and these results are summarized below.

**Scheme 2.1** synthesis of cyclic 7a.a. PLMT(OtBu)GD(OtBu)T(OtBu) using conditions (Table 2.1)
Table 2.1 Synthesis of 7a.a. cyclic peptide PLMT(OtBu)GD(OtBu)T(OtBu)

<table>
<thead>
<tr>
<th>Reaction</th>
<th>(S)-aziridine aldehyde dimer (eqv.)</th>
<th>tBuNC (eqv.)</th>
<th>Solvent</th>
<th>Additive</th>
<th>Results&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.55</td>
<td>1.0</td>
<td>1:1 DCM:HFIP</td>
<td>-</td>
<td>SM</td>
</tr>
<tr>
<td>2</td>
<td>0.55</td>
<td>1.0</td>
<td>HFIP</td>
<td>-</td>
<td>SM</td>
</tr>
<tr>
<td>3</td>
<td>0.55</td>
<td>1.0</td>
<td>HFIP</td>
<td>DIPEA (2.0 eqv.)</td>
<td>SM, trace M+18</td>
</tr>
<tr>
<td>4</td>
<td>0.55</td>
<td>1.0</td>
<td>HFIP</td>
<td>Yb(OTf)&lt;sub&gt;3&lt;/sub&gt; 20mol%</td>
<td>No product</td>
</tr>
<tr>
<td>5</td>
<td>1.0</td>
<td>2.0</td>
<td>TFE</td>
<td></td>
<td>SM</td>
</tr>
<tr>
<td>6</td>
<td>1.0</td>
<td>2.0</td>
<td>TFE</td>
<td>DIPEA (1.0 eqv.)</td>
<td>SM</td>
</tr>
<tr>
<td>7</td>
<td>1.0&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.0</td>
<td>TFE</td>
<td>-</td>
<td>SM</td>
</tr>
<tr>
<td>8 (PGLGF)</td>
<td>1.0</td>
<td>2.0</td>
<td>TFE</td>
<td>-</td>
<td>M, M+18, trace SM</td>
</tr>
</tbody>
</table>

<sup>a</sup>Reactions monitored by LCMS. DIPEA= diisopropylethylamine, HFIP=1,1,1,3,3,3-hexafluoroisopropanol. <sup>b</sup>d-valine aziridine aldehyde dimer, (2R,4R,5S,6R)-6-isopropyl-2-((2S,3R)-3-isopropylaziridin-2-yl)-3-oxa-1-azabicyclo[3.1.0]hexan-4-ol.

All reactions gave negative results (Table 2.1). A reason for the lack of reactivity could be attributed to the C-terminal threonine with a tert-butyl-protecting group that hindered cyclization. PGLGF is a well-behaved amino acid sequence that consistently cyclizes well and adopting the PG N-terminus and C-terminal Phe of this sequence may encourage cyclization of the 7a.a. peptides. Moreover, considering potential difficulties with methionine during azide ring opening of the aziridine (inert atmosphere required) the original amino acid sequence of the 7a.a. peptide was modified whereby norleucine (Nle) was used instead of Met. Four different linear peptides were designed considering these observations and hypotheses (Table 2.2).
Table 2.2 Structure and sequence of 7a.a. linear peptides

<table>
<thead>
<tr>
<th>Code</th>
<th>Structure and Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2</td>
<td>Pro-Leu-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Thr(OtBu)</td>
</tr>
<tr>
<td>2.3</td>
<td>Pro-Gly-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Thr(OtBu)</td>
</tr>
<tr>
<td>2.4</td>
<td>Pro-Leu-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Phe</td>
</tr>
<tr>
<td>2.5</td>
<td>Pro-Gly-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Phe</td>
</tr>
</tbody>
</table>

Following SPPS, LCMS showed the linear peptides 2.2, 2.3, and 2.4 as the major product, and 2.3 showed both the product as well as a deletion sequence impurity. Purification of all linear sequences by VLC proved challenging, as the peptides were not fully soluble in eluents of MeCN/H₂O or 100% MeCN. The peptides precipitated out of the eluent onto the column and as a result, not only was product lost, the sequences became contaminated. Semi-preparative RP-LCMS was performed to purify the first three sequences and a minimal amount of DMSO was added to solubilize the peptides before injecting onto the column. With the purified linear peptides in hand, cyclization reactions were set up according to the general cyclization procedure (Table 2.2). Observation of the reactions clearly showed that the linear sequences had poor solubility in TFE and analysis by LCMS showed starting material (Table 2.2).
Scheme 2.2 Synthesis of cyclic peptide sequences (example for sequence 2.2)

Table 2.3 Cyclization of 7a.a. sequences using general cyclization conditions

<table>
<thead>
<tr>
<th>Code</th>
<th>Sequence</th>
<th>Result&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2</td>
<td>Pro-Leu-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Thr(OtBu)</td>
<td>SM</td>
</tr>
<tr>
<td></td>
<td>PL(Nle)T(OtBu)GD(OtBu)T(OtBu)</td>
<td>SM, trace M&lt;sup&gt;+&lt;/sup&gt;, trace (M+18)&lt;sup&gt;+&lt;/sup&gt;</td>
</tr>
<tr>
<td>2.3</td>
<td>PG(Nle)T(OtBu)GD(OtBu)T(OtBu)</td>
<td>No product or starting material</td>
</tr>
<tr>
<td>2.5</td>
<td>PG(Nle)T(OtBu)GD(OtBu)F</td>
<td></td>
</tr>
</tbody>
</table>

<sup>c</sup>Cyclization conditions: 0.55 eqv. (S)-aziridine aldehyde dimer, 1.1 eqv. tBuNC, in TFE at 0.2M with respect to the linear peptide. TFE = 2,2,2-trifluoroethanol, SM = starting material, M<sup>+</sup> = parent molecular ion, (M+18)<sup>+</sup> = hydrolyzed product ion. <sup>c</sup>Results based on LCMS analysis.
As indicated in Table 2.2, mainly starting material was observed after cyclization. After 3.5 h, trace cyclized product was observed by LCMS from linear peptide \textit{2.3} as well as the hydrolyzed product, (M+18)$^+$, where H$_2$O opens up the aziridine ring of the macrocycle. Another 0.55 equivalencies of dimer along with 1.1 equivalencies of tBuNC were added to the reaction to favor formation of product. More optimization was performed to increase conversion to the cyclized product.

The cyclization reactions were optimized using two different approaches. Firstly, by addition of DMSO to solubilize the peptide, and secondly, by using a different aziridine aldehyde dimer. A cyclization reaction screen of \textit{2.3} was set up to investigate the addition of trace amounts of DMSO to help solubilize the linear peptide in TFE (Scheme 2.3). To do this, the linear peptides were first dissolved in a minimal amount of DMSO, diluted with a 1:1 solution of acetonitrile: water and lyophilized to give a white powder. The results are summarized in Table 2.4 and indicate that while trace amounts of DMSO increase the rate of linear peptide consumption, the product to hydrolyzed cyclized product ratio did not change substantially.

\begin{table}[h]
\centering
\begin{tabular}{l l l}
\hline
Additive & LCMS analysis [SM: M$^+$: (M+18)$^+$] \\
\hline
- & 1:2.5:0.5 \\
Trace DMSO & 1:10:5 \\
\hline
\end{tabular}
\caption{DMSO additive for cyclization of \textit{2.3}}
\end{table}

SM = starting material, M$^+$ = parent molecular ion, (M+18)$^+$ = hydrolyzed product ion.

Subjecting the crude cyclized product from \textit{2.3} to thiobenzoic acid-mediated aziridine ring-opening conditions gave the aziridine ring-opened product, however RaNi desulphurization proved unsuccessful. Despite an overall success of the small-scale reaction, the reaction was
scaled up to produce a practical amount for further optimization and purification. Starting from 100 mg of the linear peptide, the cyclization reaction proved successful and the ring-opening the aziridine macrocycle using thiobenzoic acid followed by cleavage of the C-S bond by RaNi, followed by purification using prep LCMS gave the cyclized product (Scheme 2.4).

Scheme 2.4 Synthesis of cyclic PL(Nle)TGDT, 2.6

As a control for biochemical assays, the fully deprotected linear peptide PG(Nle)TGDT 2.7 was also synthesized and purified. This way the biological effects of a macrocyclic versus linear peptide could by studied by the Pawson group.

Scheme 2.5 Synthesis of linear PL(Nle)TGDT, TIPS = triisopropylsilane
Previous linear 7a.a. sequences (2.2, 2.4, and 2.5) were resynthesized by SPPS, and cyclized in parallel using \( d \)-leucine aziridine aldehyde dimer to achieve the lactone-containing cyclic peptide product. All reactions were monitored by LCMS and the major peak was the linear peptide and a side product equal to the linear mass plus 109 amu was also observed. Based on discussions, it was suggested that the side product is due to a reaction with proline and a backbone amide NH. Since there was no obvious explanation as to why the linear peptides behaved this way, the integrity of \( d \)-leucine aziridine aldehyde dimer was evaluated by conducting a known proline piperazinone reaction. Although this proved successful, and the integrity of \( d \)-leucine aziridine aldehyde dimer was validated, it did not explain why the linear peptides did not cyclize. Various solvents were screened on a similar linear peptide by adjusting the solvent system as well as the amount of DIPEA (Table 2.4).

![Scheme 2.6](image)

**Table 2.5** Conditions used for cyclizing 2.3 with \( d \)-leucine aziridine aldehyde dimer

<table>
<thead>
<tr>
<th>Vial #</th>
<th>Solvent (10 uL)</th>
<th>Additive</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TFE/HFIP</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>2</td>
<td>HFIP</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>3</td>
<td>10% H(_2)O in HFIP</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>4</td>
<td>1/1 TFE/DCE(^a)</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>5</td>
<td>1/1 TFE/MeNO(_2)</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>6</td>
<td>1/1 TFE/DMF</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>7</td>
<td>1/1 TFE/DMSO</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>8</td>
<td>1/1 IPA/MeCN</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>9</td>
<td>MeCN</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>10</td>
<td>10% H(_2)O in MeCN</td>
<td>1 uL DIPEA</td>
<td>SM</td>
</tr>
<tr>
<td>11</td>
<td>10% DIPEA in TFE</td>
<td>-</td>
<td>(M+26)(^+)</td>
</tr>
<tr>
<td>12</td>
<td>30% DIPEA in TFE</td>
<td>-</td>
<td>No product</td>
</tr>
<tr>
<td>13</td>
<td>Toluene</td>
<td>-</td>
<td>No product</td>
</tr>
</tbody>
</table>
The cyclized product was not observed initially, but by analysis of a small aliquot using TFA to remove protecting groups (micro-cleavage), the deprotected cyclized product could be ionized and detected by LCMS. Using conditions from vial 17 (Table 2.4) as well as analysis using micro-cleavage, cyclizing of linear peptide 2.3 was performed and 2.8 was purified by RP-HPLC and isolated by lyophilization in 29% overall yield (Scheme 2.7).

Scheme 2.7 Synthesis of lactone macrocycle 2.8

During the pursuit of 7aa cyclic peptides, there were also ongoing mechanistic and structural characterization studies on peptide cyclization reactions. Serge Zaretsky and Dr. Jen Hickey are
investigating various intermediates and side products that arise during cyclization. Current cyclic peptide work is focused on understanding the mechanism, kinetics, and structures of unwanted side-products using $^{13}$C and $^{15}$N labeled reagents before continuing the pursuit of libraries of biologically active peptide macrocycles.

### 2.2.3 SH2 kinase domain peptides

The SH2 kinase domain of Abl has been found integral for Bcr-Abl activity. This region of Abl provides a suitable molecular basis for the design of inhibitors. The SH2 domain contains a mixture of aromatic and polar residues as well as a hydrophobic cleft. The available binding surface was calculated as 800 – 1000 Å$^2$. Within the SH2 kinase domain, an isoleucine (I164), has been found to be imperative to any kinase activity. This key residue exists on the ridge of the pocket in a hydrophobic region and points outwards from the pocket.

**Figure 2.10** *Left:* SH2 kinase domain, red and yellow spheres highlight polar sites. Aromatic regions are shown in blue. *Middle:* SH2 kinase domain, available binding surface illustrated in gray. *Right:* Isoleucine-164 sits on upper right ridge of pocket (*residues in orange*).

Cyclic peptides were designed computationally to achieve a steric and electronic fit in this SH2 domain. A tetracyclic peptide was the optimal size with an isoleucine present as at least one of the amino acid residues. All possible combinations of tetracyclic peptides using at least one of D-isoleucine or L-isoleucine were enumerated. Hundreds of possible tetramer cyclic peptides resulted and each one, along with its various conformers, was docked into the SH2 kinase domain using GLIDE Software. Each docked posed was analyzed manually and ranked based on the docking conformation, hydrogen bond contacts, and favourable electrostatic interactions.
Furthermore, the feasibility of synthesizing the amino acid sequence was also considered where sequences containing more than one charged amino acid were disfavored. Another consideration was tetracyclic peptides that contained all L amino acids, as these building blocks are less expensive and readily available.

1. Lile-Lleu-Lser-Dser
2. Dleu-Lile-Lasp-Lorn
3. Lile-Dasn-Lasp-Dleu
4. Lile-Darg-Lglu-Dpro
5. Dthr-Dhis-Lile-Dleu
6. Dlys-gly-Dleu-Lile
7. Lile-Dser-Larg-Lglu
8. Dgln-Dthr-Lile-Dleu
9. Dleu-Darg-Lile-Lhis
10. Lasp-Lile-Lglu-Lser
12. Dlys-Lile-Dasn-Ltrp
13. Larg-Dser-gly-Lile
14. Lile-Dasn-gly-Lgln
15. Lile-Lthr-Larg-Lser
16. Lile-Dglu-gly-Dgln
17. Lasn-Lile-Llys-Lpro
18. Lpro-Lile-Lglu-Larg
19. Lile-Larg-Lala-Ltrp
20. Lile-Dser-Larg-Lglu

**Figure 2.11** List of 20 cyclic tetramer peptides selected from computational analysis

Since the majority of these sequences do not contain an N-terminal proline, we were wary of attempting any cyclization reactions. Our strategy was to first make the linear peptides and then benzylate the N-terminus to make a secondary amine. With an N-terminal secondary amine, cyclization should theoretically be feasible. To begin, five amino acid sequences were focused on; all of which contained only L-amino acids. These were prepared by SPPS and the sequences were kept on resin to perform solid phase benzylations.
1. Larg – Lorn – Lasp – Lile (2.7)
2. Lile – Lthr – Larg – Lser (2.8)
3. Lasn – Lile – Llys – Lpro (2.9)
4. Lpro – Lile – Lglu – Larg (2.10)
5. Lile – Larg – Lala – Ltrp (2.11)

**Figure 2.12** All L-a.a. containing sequences

The linear peptides on solid phase were slow to react and further optimization needs to be done. When the D-amino acids became available, all remaining sequences where prepared by SPPS (see Figure 2.11). These sequences were cleaved from the resin and purified using a C-18 solid phase extraction (SPE) columns and a gradient elution 0 - 100% MeCN in H₂O. Fractions were analyzed for product using LCMS. A benzylation reaction was attempted with one sequence in solution phase using benzaldehyde and sodium borohydride (Scheme 2.8).

![Scheme 2.8 Solution phase benzylation of 2.12](image)

Unfortunately, this trial benzylation reaction proved unsuccessful. Before trying out several conditions and modifying the reaction to try and obtain the benzylated peptide, it was suggested that we should first focus on cyclizing peptides that already contained N-terminal benzylated amino acids in their sequence. This N-terminal amino acid work was a focus in chapter one. Furthermore, as a consequence of the mechanistic insights gained earlier, more understanding of peptide cyclization was required before exploring large library syntheses.
2.3 Conclusions and Future Work

Using computation analysis, a druggable region on EED was successfully identified. Cyclic peptides based on our methodology were designed and enumerated and computationally modeled with the identified binding pocket. The key region for biochemical activity on EED that binds to EZH2 was later identified as a different region on EED. More computational work is needed to identify a different druggable region on EED in the key area, which makes contact with EZH2. The activation loop of Abl was identified as a potential target suitable for inhibition by peptide macrocycles. Computational design and analysis showed that a 7a.a. cyclic peptide using our methodology had potential. A 7a.a. macrocycle was synthesized using (S)-aziridine aldehyde dimer, after overcoming solubility challenges. Using d-leucine aziridine aldehyde dimer gave a lactone containing macrocyclic product, however it is unlikely that this is the true species. We now suspect that d-leucine aziridine aldehyde dimer reacts with linear peptides to give various intermediates that yield a product equal in mass to the lactone containing macrocycle but differs in structure. More work is being done to isolate and characterize the various intermediates of macrocyclization and their reactivity. Once the synthetic foundation for macrocyclization has been laid, it will be exciting to continue our pursuit of bioactive cyclic peptides as therapeutic targets.

2.4 Experimental Details

2.4.1 General information

All reactions were carried out under an atmosphere of air unless otherwise indicated. Reagents were used as obtained from commercial suppliers or purified according to standard procedures. Aziridine-2-carboxaldehyde, (S)-aziridine aldehyde dimer was synthesized according to literature procedures. All $^1$H and $^{13}$C NMR spectra were recorded using a Bruker 400 MHz or Varian 400, 600, 700, or 500 MHz cryoprobe spectrophotometer and internally referenced to a residual protiosolvent. 2D (COSY, HSQC) NMR spectroscopy was used where appropriate to assist in the assignment of $^1$H and $^{13}$C NMR spectra.

**General Solid Phase Peptide Synthesis (SPPS):** Linear peptides were prepared on an Apex 396 Multiple Peptide Synthesizer (AAPPTec, LLC) using the RV40 reaction block (suitable for 0.15
mmol of peptide resin in each well). Peptides were assembled on 2-chlorotrityl chloride resin with manual loading of the first amino acid. Linear peptide precursors were then synthesized by conventional Fmoc solid-phase based peptide synthesis using coupling steps with HBTU in NMP. Peptides were cleaved from the resin by three successive washes, ½ hour apart, with 1:3 HFIP/DCM. The cleavage cocktail was then evaporated and peptides precipitated from chilled methyl tert-butyl ether, MTBE.

**Analytical HPLC/MS Chromatography:** Low-resolution mass spectra (ESI) were collected on an Agilent Technologies 1200 series HPLC paired to a 6130 Mass Spectrometer. Compounds were resolved on an Agilent Poroshell 120 EC-C18, 2.7 µm, 4.6 x 50 mm² column at room temperature with a flow of 1 mL/min. The gradient consisted of eluents A (0.1% formic acid in double distilled water) and B (0.1% formic acid in HPLC-grade acetonitrile). The gradient method started at 5% of B for the first 0.99 minutes, followed by a linear gradient from 5% to 95% B in 10.0 minutes. The column was then washed with 95 % B for 1.0 minutes and equilibrated at 5% B for 1.5 minutes.

**RP-HPLC Semi-Preparative:** Compounds were resolved on an Agilent Zorbax SB-C18, 5.0 µm, 9.4 x 250 mm² column at room temperature with a flow of 5 mL/min using a gradient consisting of eluents A (0.1% Formic acid in double distilled water) and B (0.1% Formic acid in HPLC-grade acetonitrile). The column was then washed with 95 % B for 5.0 minutes and equilibrated at 5 % B for 2 minutes.

**2.4.2 Synthesis and characterization data**

![Chemical structure](image)

**PLMT(OtBu)GD(OtBu)T(OtBu) (2.1):** Synthesized according to general SPPS procedure. $^1$H NMR (400 MHz, DMSO) $\delta$ 8.60 (br s, 1H), 8.43 (br s, 1H), 8.30 (d, $J$ = 8.2 Hz, 1H), 7.91 (t, $J$ =
5.2 Hz, 1H), 7.54 (d, \( J = 7.8 \) Hz, 1H), 7.49 (d, \( J = 8.8 \) Hz, 1H), 4.72 (q, \( J = 7.6 \) Hz, 1H), 4.43 – 4.36 (m, 2H), 4.24 (dd, \( J = 8.0 \) Hz, 3.4 Hz, 1H), 4.10 (dd, \( J = 6.2, 2.9 \) Hz, 1H), 3.96 – 3.93 (m, 2H), 3.77 (d, \( J = 5.2 \) Hz, 2H), 3.07 (br s, 2H), 2.71 – 2.65 (m, 1H), 2.46 – 2.40 (m, 2H), 2.16 (m, 1H), 1.95 (m, 1H), 1.81 – 1.77 (m, 3H), 1.58 (m, 1H), 1.51 – 1.47 (m, 2H), 1.38 (s, 6H), 1.13 (s, 6H), 1.10 (s, 6H), 1.04 (d, \( J = 6.2 \) Hz, 2H), 0.98 (d, \( J = 5.8 \) Hz, 2H), 0.88 (dd, \( J = 9.8 \) Hz, 6.5 Hz, 4H). Mass Spec (ESI-MS) m/z = 902.6 = (M+H)⁺.

Pro-Leu-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Thr(OtBu) (2.2): Synthesis by SPPS according to general procedure. Mass Spec (ESI-MS) m/z = 884.6 = (M+H)⁺.

Pro-Gly-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Thr(OtBu) (2.3): Synthesis by SPPS according to general procedure. Mass Spec (ESI-MS) m/z = 828.6 = (M+H)⁺.

Pro-Leu-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Phe (2.4): Synthesis by SPPS according to general procedure. Mass Spec (ESI-MS) m/z = 874.6 = (M+H)⁺.
Pro-Gly-Nle-Thr(OtBu)-Gly-Asp(OtBu)-Phe (2.5): Synthesis by SPPS according to general procedure. Mass Spec (ESI-MS) m/z = 818.5 = (M+H)^+.

Cyclic Pro-Leu-Nle-Thr-Gly-Asp-Thr (2.6): Linear peptide PL(Nle)T(OtBu)D(OtBu)T(OtBu) 2.3 (100 mg, 0.121 mmol) and (S)-aziridine aldehyde dimer (9.4 mg, 0.066 mmol) were added to a 1 dram vial containing a stir bar. To this, TFE (1.0 mL) was added followed by tBuNC (15 uL, 0.133 mmol). The reaction was allowed to stir at room temperature and was monitored by LCMS. After 4 h, additional (S)-aziridine aldehyde dimer (3.0 mg, 0.02 mmol) and tBuNC (7.5 uL, 0.065 mmol) were added. After a total of 5 h, the reaction was concentrated to dryness, diluted with CH$_2$Cl$_2$ (1.0 mL) and thiobenzoic acid (113 uL, 0.968 mmol) was added. The reaction was left to stir overnight at room temperature. Using a stream of nitrogen, the solvent was removed and 50:50 RaNi aq. slurry: 80% aq. EtOH (1 mL) was added. Following completion of hydrogenation, the reaction was filtered through Celite to remove RaNi, and the residue was washed successively with MeOH (3 x 2 mL) and EtOAc (3 x 2 mL). The filtrate was concentrated to dryness under a stream of nitrogen and a cleavage cocktail (TFA/TIPS/H$_2$O (95:2.5:2.5), 1 mL) was added. After 2 h, the reaction was diluted 10 fold (H$_2$O: MeCN (1:1) and the deprotected peptide was isolated by lyophilization. Purification by semi-preparative RP-HPLC using two injections of 2000 uL (200 uL DMSO: 800 uL H$_2$O: 1000 uL MeCN) a gradient of 5 – 95% MeCN in H$_2$O over 45 minutes followed by lyophilization the pure product as a white solid (0.5 mg, 0.5%). Mass Spec (ESI-MS) m/z = 784.5 = (M+H)^+. 
**Pro-Leu-Nle-Thr-Gly-Asp-Thr (2.7):** The linear peptide (10 mg, 0.012 mmol, 1.0 eqv) was dissolved in TFA/TIPS/H$_2$O (100 ul (95/2.5/2.5) and allowed to stir at room temperature for 2 h. H$_2$O/MeCN (1:1) was added to quench the reaction and the resulting mixture was freeze dried overnight to isolate the product, m/z = 716.5 = (M+H)$^+$. 

**Cyclic lactone containing PG(Nle)TGDT macrocycle 2.8:** The linear peptide (2.3, 42.5 mg, 0.048 mmol, 1.0 eqv.), and d-leucine aziridine aldehyde dimer (7.3 mg, 0.029 mmol, 0.6 eqv.) were added to a small vial along with a stir bar. TFE (240 uL) was added followed by tBuNC (6.5 uL, 0.058 mmol, 1.2 eqv.). The reaction was monitored by LCMS using micro cleavage. To form the lactone species, 3% formic acid in 1/1 MeCN/ H$_2$O was added and the reaction was stirred at room temperature until completion by LCMS (m/z = 1021.6) (M+H)$^+$. The amino acid protecting groups were then cleaved using 1.5 mL of TFA/H$_2$O/TIPS 95/2.5/2.5 stirring overnight at room temperature. Finally the reaction was concentrated to an oil and purified by semi-prep LCMS using a gradient of 20 – 100% MeCN in H$_2$O over 45 min and isolating the product as a white power, 11.7 mg, 28.6% yield.
2.5 References

18. FTMAP can be found at [http://ftmap.bu.edu/](http://ftmap.bu.edu/).
Chapter 3. Design, Synthesis, and Analysis of Small Molecules as Probes for Epigenetic Proteins

3.1 Introduction

3.1.1 Epigenetics and histone modifications

DNA is organized into chromosomes, which contain various genes that code for multiple proteins. Within eukaryotic cells, DNA is wound around large proteins known as histones. This can be envisioned as thread wound around a spool and this entire complex is referred to as a nucleosome. The spools or histones themselves are composed of various subunits that make up the core structure: H2A, H2B, H3 and H4. Both H3 and H4 have long chains of amino acids that extend outside the nucleosome. These chains are known as histone tails and have an integral role in gene regulation.

![Nucleosome showing structure of histones and histone tails.](image)

Gene regulation can occur by either the direct modification of DNA or by changes in gene expression. Epigenetics is defined as changes in gene expression independent of DNA modification. Two examples are DNA methylations as well as histone modifications, more commonly referred to post-translational modifications, PTMs. The covalent modification of a histone directly influences the accessibility of transcribing DNA by a promoter thereby either preventing transcription of a gene or allowing a gene to be overly transcribed. This silencing and/or activating of genes from PTMs can have massive downstream consequences.
The N-terminus of histone tails are alkaline and as a result susceptible to frequent and multiple PTMs\(^1\). The C-terminal regions can also be susceptible to PTM if they are flexible, while the core region of a histone is the least prone to modifications.\(^1\)

Histone modifications include (ac acetylation, \(me1\)-3 mono-, di-, and trimethylation, \(P\) phosphorylation, ribo ADP ribosylation, \(ub\) ubiquitination/ubiquitylation, bio biotinylation, iso proline isomerization, prop propionylation, cit citrulination, and but butylation).\(^1\) These modifications can occur in any number and frequency on a histone thereby regulating changes in DNA transcription. Various combinations of histone modifications are collectively known as the ‘histone code’ and there have been several efforts to decode what these combinations mean and how they impact gene transcriptional activation or silencing.

![Figure 3.2 PTMs of histone tails.](image)

Within the histone code, there are certain proteins that make modifications to various amino acids (writers), those that remove modifications (erasers), as well proteins that recognize modifications (readers). Histone modification by acetylation and methylation on lysine residues is particularly important and has become a focus of research aimed at understanding epigenetic regulation. Acetylation of lysine is almost always correlated with transcriptional activation of genes. The current known major sites of acetylation on H3 include: K9, K14, K18, K23, and
Methylation of lysine on H3 has been correlated with both transcriptional activation and repression. H3K4me3 leads to activation, whereas H3K9me3 results in gene silencing.

Methylation of H3K4 (H3K4me3) is regulated by methyltransferases Set1/COMPASS-family proteins and demethylases KDM5A-D. This methylation state of H3K4 directly influences gene transcription and deregulation of this fine balance has been implicated in leukemia and other cancers. It is important to point out however, that whether aberrant epigenetic mechanisms are a cause or an effect of cancer remains a mystery.

3.1.2 SETdb1

There are several proteins that are writers of methylations. Proteins that contain a lysine specific SET domain are a common catalyst of histone methylations. These SET-domain containing proteins bind a methyltransferase, S-Adenosyl methionine (SAM), as well as a lysine residue in the SET pocket thereby catalyzing the methylation reaction.

The gene that codes for SETdb1 is located in the chromosome 1q21 interval, the same region that is overly active in cancerous zebra fish models and likely contains a human melanoma susceptibility gene. SETdb1 has been shown to be specific for catalyzing the methylation of H3K9 (H3K9me3). H3K9me3 is associated with the repression of euchromatic genes (consistently actively transcribed chromatin) and retroelements. SETdb1 has also been found to participate in DNA methylation, and is over expressed in cancerous broncoepithelial cells.

3.1.3 Bromodomains

Bromodomains are readers of histone tails that are most often found in conjunction with other epigenetic reader domains such as PHD (plant homeodomain), PWWP (proline-tyrosine-tyrosine-proline), and Tudor domains. 44 of the 56 bromodomains that are encoded by the human genome function to bind acetyllysine.

3.1.4 BRD1

Acute myeloid leukemia and mixed-lineage leukemia are associated with chromosomal translocations or rearrangements caused by fusion proteins. MOZ (MYST3, monocytic leukemia histone acetyltransferase) is a fusion protein believed to induce leukemia by its ability
to aberrantly acetylate histones.\textsuperscript{8} This in turn, alters HOX gene expression (homeobox protein that controls activation/repression of genes) during hematopoiesis (formation of blood cells).\textsuperscript{8} Bromodomain-containing protein 1, BRD1 acts a scaffold that recruits MOZ fusion protein and other co-factors to various regions of active chromatin that become aberrantly acetylated.\textsuperscript{8}

Histone acetylation and methylation are thought to act in conjunction with each other to affect gene transcription activation or silencing.\textsuperscript{8} While BRD1 is associated with acetylation of histones, it has recently been implicated in reading methylated lysine. BRD1, contains both a PWWP domain as well as a PHD domain. BRD1 has been proven to bind H3K36me3 through its PWWP domain.\textsuperscript{8} In the research conducted by Vezzolli et al., an H3K36me3 peptide interaction with BRD1 was studied by \textsuperscript{1}H-\textsuperscript{15}N HSQC NMR.\textsuperscript{8} A $K_d$ value of 2.7 ± 0.2 mM was calculated by plotting the change in chemical shift upon increasing the ratio of peptide: BRD1.\textsuperscript{8} To further understand the interaction of H3K36me3 and BRD1, crystal structures of the apo protein and complexed with the peptide was obtained. The trimethyllysine fits into an aromatic cage formed from three key residues of BRD1: Tyr1095, Tyr1099, and Phe1147. This aromatic cage motif can be found in other PWWP Kme3 binding motifs. Complexation of the trimethyllysine is driven by energetically favourable cation-pi electrostatic interactions and complimentary fit.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{image.png}
\caption{Crystal structure of histone-3 peptide bound to BRD1.\textsuperscript{8}}
\end{figure}

\subsection{3.1.5 Molecular probes}

Molecular probes are molecules that can perturb or ‘probe’ complicated biological systems. By setting up specific biological assays and determining the effect of a molecular probe on the analysis, it is possible to learn more about how a native interaction controls, or is influenced by various biochemical pathways.
Previous work in the area of molecular probes has led to three potent bromodomains inhibitors: IBET-762, IBET-151, and JQ1. These specifically target BRD4 and BRD4-NUT respectively.\(^7\) IBET-762 blocks BRD4 from binding acetyllysine and JQ1 blocks the oncogenetic BRD4-NUT fusion protein from binding to acetyllysine.\(^6\) Not only were these two inhibitors found to be potent in assays; they resulted in a paradigm shift in how epigenetic proteins are researched. Typically, only writers and erasers have been targeted but BRD4 is a reader protein. IBET-762, IBET151, and JQ1 have paved the way to view epigenetic reader proteins as valid targets.

![IBET 762, IBET 151, and JQ1](image)

**Figure 3.4** BRD inhibitors IBET 762, IBET 151, and JQ1

SETdb1 and BRD1 are well studied in how they bind trimethyllysine yet, knowledge is lacking on how epigenetic proteins are involved in the larger landscape of oncogenic signaling pathways. Better understanding of oncogenic pathways through probe design and analysis can better guide the design and function of future therapeutic agents.

### 3.1.6 Research Objectives

Using pdb files obtained from crystal structures of SETdb1 and BRD1, each protein will be computationally modeled using GLIDE software.\(^9\) Using Glide docking, various ligands will be evaluated in how they bind in the aromatic cage of each protein. From these results, a list of molecules to synthesize and evaluate will be determined. Once the compounds are obtained they will be sent to Dr. Peter Brown’s group at the Structural Genomics Consortium in Toronto (SGC) for evaluation. Based on positive preliminary experimental results obtained from \(^1\text{H}-^{15}\text{N}\) HSQC NMR, cocrystallization experiments will be attempted in order to gain insight into the mode of binding, as well as isothermal titration calorimetry to understand the thermodynamics of binding. A full assignment of the protein backbone will be attempted by 3D NMR techniques using doubly labeled (\(^{13}\text{C},^{15}\text{N}\)) protein to understand the solution state mode of binding.
3.2 Methodology

3.2.1 Glide docking

Using crystal structure pdb (protein data bank) files it is possible to look at the crystal structure of a protein with compatible software. Furthermore, by studying the known binding mode of a native ligand or another entity in the crystal structure, it is possible to rationally design molecules to electronically and sterically fit a particular binding pocket.

Molecules will be assessed as a potential inhibitor of a protein by fitting or ‘docking’ them into the protein and quantifying their potency. Glide software (grid-based docking with energetics), a powerful computational program, will be used to dock molecules or ligands into each protein and score them based on their electronic and structural fit. By performing an exhaustive search of a ligand’s conformation, orientation, and position in space in a particular region a Glide Score will be assigned.

3.2.1.1 Preparation of protein

The crystal structure of the proteins was first imported from the RSC PDB bank. Using the Protein Preparation Wizard the proteins were prepped for docking by the addition of missing protons, fixed bond orders, optimized protonation states, and hydrogen-bonding networks. It also removed insignificant water molecules that were beyond a certain distance from the protein. Finally an energy minimization was performed.

Rather than working with the entire structure, Glide docking utilized a glide grid, which represents the properties of a region on the protein (electrostatics potential at each grid point, van der Walls contacts, etc…). It was possible to generate the grid using the receptor grid generation function in glide. The center of binding pocket was selected on the protein so that the glide receptor grid would be centered at this point. A large box representative of the glide grid was formed, and its size was such that it did not overly constrain the placement of the ligands during the docking process.
3.2.1.2 Preparation of ligands

Ligands were drawn as cdx files and then converted into smiles and saved as a smi file. This procedure was fairly straightforward and fairly intuitive. As an example, lysine was converted into a smile where the notation `@@` depicts S-stereochemistry around the alpha carbon and `C(O)=O` is the acid group. Hydrogen atoms were left out of smile codes unless they were important for stereochemistry.

\[
\text{NH}_2\text{N}\text{O}\text{OH} \quad \text{N[\text{C}@\text{H}](\text{CCC}[\text{N}+](\text{C})(\text{C})\text{C(O)=O}}
\]

**Figure 3.6** Chemdraw of lysine and the corresponding smile code

Once the smile file (.smi) was imported into a project table in Glide. The ligands were in 2D and were rendered into 3D and their conformation energy minimized. This was all conducted using the ligprep wizard, where the number of lowest energy conformations were set and these conformers were later docked in the glide grid.

3.2.1.3 Glide Docking

Finally it was possible to dock the ligands in the receptor grid using the Glide Ligand Docking function.\textsuperscript{10} Glide used several filters to achieve a final calculated score for each ligand.\textsuperscript{10} The
first step was to find the lowest energy location for the ligand within the glide grid. The receptor was represented by the grid sets of fields. The position and orientation, “pose” of the ligand was refined as more sets of fields were taking into account. The second step involved producing a set of ligand conformations, which were pre-screened or filtered based on those that gave the best fit in the region of space where the poses were scored. Starting from these select poses, the ligand was minimized in the field of the receptor using a standard molecular mechanics energy function. The lowest poses were then subjected to further energy minimization using Monte-Carlo that adjusted torsional minima. Finally the ligands and their poses were scored and these numbers gave a prediction for their binding affinity. Glide scoring took into account all the ligand-receptor lipophilic interactions, hydrogen bonding interactions, Columb and van der Walls interaction energies, and solvent effects for charged groups.

3.2.2 Evaluating binding efficiency

Various experimental methods will be used to evaluate how well ligands bind to the proteins including crystallization, differential scanning fluorimetry (DSF), $^1$H$^{15}$N-HSQC NMR, and isothermal titration calorimetry (ITC). The experimental data obtained from these methods will be used to evaluate structures designed using computational methodology.

3.2.2.1 Crystallization

Structural insight into the mode of binding can be attained from a crystal structure of the bound compound and protein of interest. A wide variety of techniques and methods of crystallization optimization exist. One commonly employed method is cocrystallization where the ligand and protein in separate solutions are added together and allowed to crystallize. Additives, temperature, time, and concentrations of protein and ligand all affect crystallization. Temperature can play an important role in crystal growth and heating a protein-ligand complex may allow the protein to fold properly and bind the ligand before crystallizing.

In vapour diffusion cocrystallization experiments, equal volumes of precipitant and protein were present in the drop and as the water diffuse out, the concentration of both the precipitant and protein steadily increased. Eventually, the conditions were such that the protein crystallizes. Two different techniques were employed in vapour diffusion experiments: sitting-drop and hanging-drop.
Figure 3.7 Left: hanging-drop and sitting drop. Right: Concentration of protein and precipitant to obtain nucleation (crystallization) state.\textsuperscript{12}

Soaking ligands into existing crystals and was used for both proteins. Once the conditions were optimized to grow diffraction-quality protein crystals, these crystals were transferred to solutions containing a ligand of interest. If the ligand bound the protein, it should have theoretically diffused into the crystal lattice. It was necessary to consider the integrity of the protein crystals, as they were not always stable in the solvent used to dissolve a particular ligand. This affected the allowable soaking time. Occasionally, additives were added to increase a protein’s stability during soaking or during cryoprotectant exchange.\textsuperscript{12}

3.2.2.2 Protein-ligand binding studies and structural characterization by NMR

It was possible to chemically map all the amino acids in a protein apart from proline using N-15 enriched protein and subjecting it $^1$H-$^{15}$N HSQC NMR. Each N-H can give a unique chemical shift or resonance as determined from its electronic environment in the protein. When a protein became bound by a ligand, the electronic environment of the N-H atoms is different and therefore resonates at a different frequency. Depending on the proportions of unbound and bound protein, the magnitude of chemical shift movement was different; where higher concentrations of bound protein resulted in the largest changes in chemical shift. In this way it was possible to qualitatively observe binding and screen several compounds as potential inhibitors.

Another powerful tool using NMR was the ability to fully characterize an entire protein backbone using complicated 3D techniques. This was not a trivial experiment, as the protein must first be expressed as a doubly labeled protein ($^{13}$C and $^{15}$N). Once the backbone has been characterized, it is possible to know which $^1$H-$^{15}$N cross-peaks in the HSQC experiments belong
to which amino acids’ NH backbone. From this, it is now reasonable to know which subset of peaks are binding to the ligand and therefore which part of the protein is bound to the ligand.

3.2.2.3. Differential Scanning Fluorimetry, DSF and Stargazer

A fast high-throughput screening technique to assess protein-ligand binding can be done by differential scanning fluorimetry, DSF. The principle of this method relies on the thermal melt temperature of a protein and the stabilizing effect of small compounds. If a small molecule or fragments binds to a protein, this event is energetically stabilizing. Therefore, as the bound protein is heated, the temperature at which it melts will be higher than the temperature required to melt the original protein. As the difference in temperature increases so does the strength of binding between the small molecule and the protein.

Fluorescence is used to measure the thermal melt temperature whereby a dye, SYPRO Orange, is added to the solution containing the protein. As the protein unfolds, the dye binds to the exposed hydrophobic sites. When the dye is bound to the protein, it fluoresces. The fluorescence is then plotted versus temperature and the thermal melt temperature is extrapolated from the curve.

![Figure 3.8 Typical DSF curve.](image)

Stargazer, a complimentary technique to DSF, was also used. This method relies on classical light principles to measure aggregated protein and is advantageous over DSF as any intrinsic fluorescence of a ligand will not affect the results.

3.2.2.4 Isothermal Titration Calorimetry, ITC

ITC is the best measure of the mechanism of how a compound binds to a protein. It provides a
direct measurement of the heat generated or absorbed upon a binding event. The quantities gained from an ITC experiment include the binding affinity ($K_d$), enthalpy ($\Delta H$), entropy ($\Delta S$), and the number of binding sites ($n$). As the ligand or compound is titrated into the protein, the temperature will increase if the binding event is exothermic. A reference cell is kept at the same temperature of the reaction cell so as the temperature of the reaction cell increases, more power is needed to adjust the reference cell to the correct temperature. The amount of power per mol of titrant (ligand) is plotted as a function of the molar ratio of ligand to protein, and from the resulting titration curve, various thermodynamic parameters are determined. The free energy of binding can be calculated from $\Delta G = -RT \ln K$ and the entropy of binding can be found from $\Delta G = \Delta H - T \Delta S$.

**Figure 3.9** *Left:* Raw data from a typical ITC titration *Right:* ITC titration curve.\textsuperscript{14}

3.3 Results

3.3.1 Pyrrolidine as a trimethyllysine mimic

Fragments were designed using pyrrolidine as a trimethyllysine mimic. The first fragments were of low molecular weight and synthesized from compounds readily available in our lab. All fragments were modeled in GLIDE.

Starting with aniline and 1,4-dibromobutane, a pyrrolidine could be synthesized using a microwave synthesis and basic aqueous conditions. This general reaction was useful in two ways, first as it was applicable to a small library synthesis and secondly it invoked green
chemistry methodology. The general scheme is shown below as well as table of synthesized compounds.

\[
R'\text{NH}_2 + X(CH_2)_nX \xrightarrow{\text{Aq. K}_2\text{CO}_3, \text{MW}} R-N(CH_2)_n\text{Aq.}.
\]

**Scheme 3.1** General synthesis of pyrrolidines using microwave synthesis\(^\text{15}\)

<table>
<thead>
<tr>
<th>Table 3.1 Synthesized molecules using microwave reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compound</strong></td>
</tr>
<tr>
<td>3.1</td>
</tr>
<tr>
<td>3.2</td>
</tr>
<tr>
<td>3.3</td>
</tr>
<tr>
<td>3.4</td>
</tr>
</tbody>
</table>

Other pyrrolidine fragments were synthesized via a reductive amination where the products could be isolated with minimal purification. However, in both of these reactions, only starting material was observed.

**Scheme 3.2** Reductive amination syntheses\(^\text{16}\)

Sulfonamides were synthesized in quantitative yield using the following green chemistry reactions. A piperidine derivative was also successful synthesized to investigate a possible steric difference between a five and six membered ring.
Amide derivatives were also synthesized. The bis-amide derivative was also attempted however this reaction was unsuccessful as the starting material was found to be impure.

Scheme 3.4 Synthesis of pyrrolidine amide derivatives

3.3.2 Analysis of fragments by DSF and Stargazer

To measure how strongly our fragments bound to BRD1, we performed differential scanning fluorimetry (DSF). This thermal shift assay measures the thermal stability of a protein where at a particular temperature, the protein will become denatured. If a compound binds to the protein it will stabilize the protein and increase the temperature at which the protein denatures. The thermal shift is a measure of the difference in temperature (°C) at which the protein denatures with and without the compound. A large positive value means that the compound bound and
stabilized the protein whereas a negative value means the compound actually destabilized the protein. DSF relies on a fluorescent dye, SYPRO Orange, and because some of our compounds are fluorescent themselves we were concerned about the accuracy of our measurements. To corroborate or disprove our values obtained from DSF, we also preformed thermal shift experiments using Stargazer. Stargazer is complimentary to DSF but does not rely on fluorescence. The results are shown below as well as a comparison of the experimental data obtained from DSF, Stargazer, with the computational docking scores (Table 3.2).

Table 3.2 Compounds analyzed by DSF and Stargazer

<table>
<thead>
<tr>
<th>Reference</th>
<th>Structure</th>
<th>ΔT (DSF) (°C)</th>
<th>ΔT (Stargazer) (°C)</th>
<th>Glide Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td><img src="image" alt="Structure 3.1" /></td>
<td>2.56</td>
<td>1.3</td>
<td>-4.846</td>
</tr>
<tr>
<td>3.2</td>
<td><img src="image" alt="Structure 3.2" /></td>
<td>2.35</td>
<td>2.15</td>
<td>-</td>
</tr>
<tr>
<td>3.3</td>
<td><img src="image" alt="Structure 3.3" /></td>
<td>2.47</td>
<td>2.3</td>
<td>-4.766</td>
</tr>
<tr>
<td>3.5</td>
<td><img src="image" alt="Structure 3.5" /></td>
<td>2.19</td>
<td>2.4</td>
<td>-4.65</td>
</tr>
<tr>
<td>3.7</td>
<td><img src="image" alt="Structure 3.7" /></td>
<td>1.30</td>
<td>1.26</td>
<td>-4.605</td>
</tr>
<tr>
<td>3.8</td>
<td><img src="image" alt="Structure 3.8" /></td>
<td>-0.06&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.12&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-5.282</td>
</tr>
<tr>
<td>3.6</td>
<td><img src="image" alt="Structure 3.6" /></td>
<td>2.23</td>
<td>1.4</td>
<td>-4.585</td>
</tr>
<tr>
<td>3.9</td>
<td><img src="image" alt="Structure 3.9" /></td>
<td>2.30</td>
<td>2.2</td>
<td>-4.084</td>
</tr>
</tbody>
</table>

<sup>a</sup>Negative values are meaningless and do not indicate destabilization of the protein

Four compounds were selected for a dose response assay where we would be able to obtain K<sub>d</sub> values. However we did not obtain any dissociation constants from this. BRD1 crystal soakings were set up using the following compounds; however these were also unsuccessful.
In theory, the more surface contact a molecule or fragment can make with the protein, the higher the interaction energy should be. We sought to increase the probability of our fragments binding to BRD1 by synthesizing more complicated fragments, while still maintaining the pyrrolidine trimethyllysine mimic core moiety.

### 3.3.3 Increasing fragment complexity using the Suzuki reaction:

A simple synthetic method to add to the complexity of the pyrrolidine-based compounds is by Suzuki reactions of bromo-phenylpyrrolidine derivatives with readily available aryl boronic acids. This will allow us to extend the compounds out of the aromatic cage pocket and engage in more surface interactions. Searching through literature for Suzuki couplings using bromo substituted dimethylated aniline; there were only a few reaction examples. Searching through our database of available arylboronic acids, all possible compounds were mapped out and docked them into BRD1 in GLIDE. The compounds in the shown below had the best fit in the aromatic cage and the best docking score (Figure 3.11).

![Figure 3.11 Suzuki compounds docked in BRD1 with the best-fit](image-url)

A simple cross coupling reaction was attempted using a Pd(db)3 catalyst and a phosphine ligand however, only starting material was observed.
Scheme 3.5 Suzuki reaction using Pd(dba)$_3$\textsuperscript{20}

A phenylurea compound was proven to be a simple and affective ligand.\textsuperscript{22} Phenylurea was first synthesized in near quantitative yield and used in a subsequent coupling reaction. However, analysis of the coupling reaction by TLC showed no product.

\[ K^+\text{N}=\text{C}=\text{O} + \text{H}_2\text{N}-\text{C}=\text{O} \rightarrow \text{2M HCl} \quad \text{(2-20)°C, 4.5 h}} \]

Scheme 3.6 Synthesis of phenylurea ligand for Suzuki cross-coupling\textsuperscript{21}

From observation of the reaction mixture, a black precipitate had formed indicating that the palladium catalyst had decomposed to form insoluble palladium(0). There were two issues with the reaction. Firstly, the palladium catalyst should have been mixed with the ligand for at least 30 minutes to allow the ligand to coordinate with the palladium to give the active catalyst. Secondly, the reaction should have been set-up more carefully with regards to the effect of air on the palladium. With regards to these observations, the reaction was repeated, yet the same outcome occurred.

Going back to the literature, another reaction seemed reasonable for our system that utilized the catalyst Pd(\(N,N\)-dimethyl-\(\beta\)-alaninate)$_2$. The first step is to make the palladium catalyst, which is a two-step procedure. Dimethylated betaalanine must be synthesized followed by coordination to Pd using basic conditions and Na$_2$PdCl$_4$ as the palladium source. The synthesis of dimethylated betaalanine was done using formic acid and formaldehyde. To make the catalyst, dimethylated
beta-alanine was deprotonated with base and added to Na$_2$PdCl$_4$. Adjusting the pH caused precipitation of the product, which was isolated by extraction to give a white air and water stable catalytic species.

**Scheme 3.8** Synthesis of stable Pd catalyst for Suzuki cross-couplings$^{23,24}$

With Pd($N,N$-dimethyl-β-alaninate)$_2$ in hand, the following coupling reaction successfully achieved and the product was isolated as a white crystalline solid in 78% yield.

**Scheme 3.9** Cross-coupling synthesis of 3.16$^{24}$

In parallel to performing Suzuki reactions, smaller molecules were also synthesized. 3.17 was synthesized in 56% isolated yield following purification and the para-bromo ortho-nitro aniline derivative failed to react. This could be attributed to the strongly withdrawing ortho-nitro group deactivated the nucleophilicity of the nitrogen amine.

**Scheme 3.10** Synthesis of 3.17 by microwave synthesis
Benzylpyrrolidine 3.18 was successfully synthesized using sodium triacetoxyborohydride, STAB according to the following reaction scheme.

![Chemical reaction scheme]

Scheme 3.11 Synthesis of 3.18

3.3.4 Re-evaluating fragment design

Our next course of action was to build our molecules such that they go deeper into the aromatic pocket of the protein host. In order to achieve this, an isoindoline scaffold was designed in which the aromatic moiety reaches further back into the pocket to engage in additional pi-stacking interactions. The phenyl isoindoline structure gave a much higher docking score as compared to the benzyl derivative since the methylene group could act as a ‘hinge’ allowing the aromatic ring to rotate and achieve the ideal pi-stacking position within the pocket. Table 3.3 depicts scaffolds that illustrate this trend.

Table 3.3 Isoindoline scaffolds and docking scores

<table>
<thead>
<tr>
<th>Compound</th>
<th>GLIDE SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Chemical structure]</td>
<td>-4.766</td>
</tr>
<tr>
<td>![Chemical structure]</td>
<td>-5.123</td>
</tr>
<tr>
<td>![Chemical structure]</td>
<td>-6.075</td>
</tr>
</tbody>
</table>
Using the benzyl isoindoline as the core architecture, the following compounds were designed and docked with BRD1.

**Table 3.4** Isoindoline compounds and docked scores

<table>
<thead>
<tr>
<th>Structure</th>
<th>Docked Score</th>
<th>Structure</th>
<th>Docked Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Structure" /></td>
<td>-6.526</td>
<td><img src="image2.png" alt="Structure" /></td>
<td>-6.241</td>
</tr>
<tr>
<td><img src="image3.png" alt="Structure" /></td>
<td>-6.818</td>
<td><img src="image4.png" alt="Structure" /></td>
<td>-6.197</td>
</tr>
<tr>
<td><img src="image5.png" alt="Structure" /></td>
<td>-6.348</td>
<td><img src="image6.png" alt="Structure" /></td>
<td>-6.150</td>
</tr>
<tr>
<td><img src="image7.png" alt="Structure" /></td>
<td>-6.279</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The pyrazole-containing derivative was found to have the best docking score. In the docked structure, a hydrogen bond forms between the acidic pyrazole and a backbone C=O of the protein. A retrosynthetic analysis for this pyrazole molecule is illustrated below (Figure 3.13).
Figure 3.13 Retrosynthesis of pyrazole isoindoline

The basic isoindoline scaffold was synthesized using a microwave reaction and 3.19 was isolated in 33% yield. The synthesis of the free isoindoline compound was also attempted, however no product could be detected. The isoindoline compound was found to be unstable and most likely oxidized to the fully aromatic compound.26

Scheme 3.12 Synthesis of isoindoline compounds

Rather than using microwave conditions, milder conditions were attempted to achieve the benzyl isoindoline compound.27
Dioxane (2.5 mL) was added to a round bottom flask equipped with a stir bar. To this, and α,α’-dibromo-o-xylene (132 mg, 0.5 mmol) was added followed by benzylamine (55 uL, 0.5 mmol) and NaOH (48 mg, 1.3 mmol). Upon addition of benzylamine, a precipitate formed after two minutes and persisted after the addition of NaOH. A small aliquot of the reaction was taken and the precipitate isolated. NMR analysis of this precipitate with DMSO showed a mixture of starting materials and dioxane. A mass spec of the crude reaction showed product (m/z = 210.1 = M+H)+. However, upon isolating the product by evaporation the solvent, the residue formed a dark-brown oil that did not correspond to the product. In order to increase the stability of the product, the hydrochloride salt of the compound was pursued. The reaction was set up according to the microwave protocol. Following completion of the reaction, the mixture was allowed to cool and by mass spec gave evidence of product formation (m/z = 196, 419.3 = M+H, 2M+H)+. The crude NMR revealed product although in very low purity. With addition of HCl to achieve precipitation of the hydrochloride salt, the mixture turned black. Analysis by NMR revealed decomposition.

Rather than trying to continue to make the benzyl isoindoline compound and coupling it successively to make the target compound, a different synthetic strategy was attempted using a phthalamide derivative, which would later be reduced in the last step to form the target compound (Figure 3.14). Using this plan, para-bromobenzyl amine, 3.20 was synthesized in 63% isolated yield using reductive amination of para-bromobenzaldehyde. The amine was then reacted with phthalic anhydride to give 3.21 in 90% yield. The same product was also achieved through an alkylation reaction in 63% yield following purification.
3.21 was then reacted with pyrazole-5-boronic acid using Pd(N,N-dimethyl-β-alaninate)$_2$ catalyst. However, only starting material could be detected after several hours of heating.

**Scheme 3.14** Syntheses of 3.21$^{28,29}$

The reaction was then screened with various bases including Na$_2$CO$_3$, K$_2$CO$_3$, and KF however in all instances no product was observed. The stability of pyrazole-5-boronic acid was questioned as it contains a boron atom directly attached to a heterocycle. These are known to be very unstable and to circumvent this problem heterocycles of this type are commonly converted into their MIDA derivatives (MIDA = $N$-methyliminodiacetic acid). Pyrazole-5-boronic acid was synthesized into its MIDA derivative in near quantitative yield and then used in a coupling reaction (Scheme 3.22 and Scheme 3.23)

**Scheme 3.16** Synthesis of MIDA pyrazole
Scheme 3.17 Cross-coupling using MIDA pyrazole

The palladium catalyst decomposed during the reaction. Preparatory TLC was used to purify the crude reaction mixture and starting material was isolated as the main component and no product was observed. The reaction was then repeated using a dual metal catalysis methodology (Scheme 3.24). No product was observed by mass spectroscopy or by TLC. There was still some blue material on the inside of the microwave vial, which suggested that the copper did not properly react with the boron.

Scheme 3.18 Dual metal catalysis cross-coupling

After unsuccessful Suzuki couplings, a MIDA-pyrazole with the pyrazole NH-Boc-protected was considered, as this substrate may work better. Boc anhydride and triethylamine were used but no product formation was detected after 48 h at room temperature in DMF.

More work still needs to be done for the Suzuki based analogues. The Boc-protected MIDA pyrazole should be synthesized starting from pyrazole, and then Boc protection of the NH should be done before converting the species to the MIDA boronate.

Isoindolines are inherently unstable due to their tendency to oxidize and no matter which path of the retrosynthesis was developed; the final target compound’s stability would remain a concern.
3.3.5 A new scaffold for BRD1

Going back to the drawing board, a phthalamide core architecture was designed and computationally modeled using GLIDE. A new target compound utilizing this new scaffold had an even better GLIDE score (<7), than previous designed compounds. A viable synthetic pathway towards the target molecule was thought out and is illustrated below.

![Chemical structure](image)

**Figure 3.14** Retrosynthetic scheme towards target compound

Two simple substrates containing the core phthalamide architecture were synthesized as small fragments to assay with BRD1. The fluoro derivative was made to compare electronics of the compounds.

![Chemical structures](image)

**Scheme 3.19** Synthesis of compounds containing core phthalamide architecture for biochemical assays

Towards the target compound, potassium phthalamide was reacted with 1,3-bis(bromomethyl)benzene in acetone at reflux to give 3.25 in 18% yield. This was then
converted to the nitrile derivative by $S_N2$ substitution with sodium cyanide (30% yield) and subjected to basic hydrolysis to achieve the carboxylic acid derivative.

Scheme 3.20 Synthesis of bromobenzyl and phenylacetonitrile phthalamide derivatives 3.25 and 3.26

Scheme 3.21 Basic hydrolysis leading to a dual acid product

The undesired hydrolyzed product was determined by $^1$H and $^{13}$C NMR. By $^1$H NMR the protons of the phthalamide were de-symmetrized. By $^{13}$C, the de-symmetrization of the aromatic carbons was present. Furthermore, there were three different carbonyl peaks suggesting that the base had attacked the amide functionality. Furthermore, by mass spec in negative mode, m/z =312.1 = (M-H)$^-$, corresponding to the mass of the hydrolyzed species. To avoid hydrolyzing the phthalamide, acidic-hydrolysis conditions can be used. A relevant literature protocol was found although the main difference was that this substrate contains an aryl CN and not benzyl.

Scheme 3.22 Literature reference for acid hydrolysis

Subjecting the benzyl nitrile derivative to these conditions gave the desired hydrolyzed product.
Scheme 3.23 Acid hydrolysis of phenyl acetonitrile phthalamide to give 3.27

When the acid catalyzed procedure was repeated on a larger scale, the fully hydrolyzed species resulted. In order to investigate this reaction without consuming large amounts of the phthalamide benzylnitrile starting material, the reaction was optimized using a model system. Both benzyl cyanide as well as N-benzylcyanide were added together with various concentrations of sulphuric acid and were reacted at either room temperature or at 50°C. However there was an inherent flaw in this experimental setup as the N-benzylphthalamide was not soluble. A second reaction was set up where the same two reagents were used but they were allowed to react with 70% H₂SO₄ at 120°C. By TLC, benzyl cyanide was consumed after 20 min.

Scheme 3.24 Optimization of hydrolysis using model substrates

From these observations, these conditions were used to hydrolyze the original starting material. Acid hydrolysis was complete after 45 minutes and the product was isolated as a white precipitate in quantitative yield.

Scheme 3.25 Optimized acid hydrolysis procedure to give 3.27

With the hydrolyzed product in hand, this was converted to the acid chloride and reacted with aniline in situ to give an amide derivative. However, although the acid chloride formed, no amide
product was observed. The reaction was repeated using benzylamine and the addition of a coupling reagent, DMAP. Again, no amide product was observed.

Scheme 3.26 Synthesis towards target compound

Although the target molecule still needs to be synthesized, the precursors were all tested in biochemical assays with BRD1 and two of the compounds both showed binding to BRD1.

3.3.6 Analysis of compounds to BRD1 by $^1$H$^{15}$N HSQC NMR Chemical Shift Mapping

The following table is a list of compounds screened for binding to BRD1 by HSQC NMR. All binding studies were performed at a 12:1 ratio of compound to BRD1.

Table 3.5 Compounds screened by $^1$H$^{15}$N HSQC NMR to assess binding

<table>
<thead>
<tr>
<th>Reference</th>
<th>Structure</th>
<th>Glide Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.28$^a$</td>
<td><img src="image" alt="3.28" /></td>
<td>-5.115</td>
</tr>
<tr>
<td>3.29$^a$</td>
<td><img src="image" alt="3.29" /></td>
<td>-6.417</td>
</tr>
<tr>
<td>3.19</td>
<td><img src="image" alt="3.19" /></td>
<td>-5.123</td>
</tr>
<tr>
<td>3.25</td>
<td><img src="image" alt="3.25" /></td>
<td>-5.948</td>
</tr>
</tbody>
</table>
a$^{3.28}$ and 3.29 synthesized and characterized by Dr. Shinya Adachi

From HSQC NMR, three compounds showed binding at a 12:1 ratio: 3.28, 3.25, and 3.27 (Figures 3.15 – 3.17). Based on these preliminary results the concentration of the compounds were lowered and HSQC measurements were performed at both a 1:1 and 1:3 ratios (Figures 3.18 – 3.20).

**Figure 3.15** Apo BRD1 (blue). 3.28:BRD1 (12:1) (red)
Figure 3.16 Apo BRD1 (blue). 3.25:BRD1 (12:1) (red)

Figure 3.17 Apo BRD1 (blue). 3.27:BRD1 (12:1) (red)
**Figure 3.18a** 3.28 bound to BRD1 (3:1) (red), 1:1 (green). Apo protein (blue).

**Figure 3.18b** Zoomed. 3.28 bound to BRD1 (3:1) (red), 1:1 (green). Apo BRD1 (blue).
Figure 3.19a 3.25 bound to BRD1 (3:1) (red), 1:1 (green). Apo protein (blue).

Figure 3.19b Zoomed: 3.25 bound to BRD1 (3:1) (red), 1:1 (green). Apo protein (blue).
Figure 3.20a 3.27 bound to BRD1 (3:1) (red), 1:1 (green). Apo protein (blue).

Figure 3.20b Zoomed: 3.27 bound to BRD1 (3:1) (red), 1:1 (green). Apo protein (blue).
Figure 3.21 is a representative HSQC of a compound that did not show any binding as evidence by the lack peak movement.

Figure 3.21a. 3.24 and BRD1 (red) 12:1. BRD1 apo protein (blue)

Figure 3.21b. Zoomed region. 3.24 and BRD1 (red) 12:1. BRD1 apo protein (blue)
As a positive control, the original hit 3.30 discovered by soaking crystallization at SGC was analyzed by HSQC NMR. Peak movement was observed as expected (Figure 3.27a and Figure 3.27b).

**Figure 3.22** SGC compound 3.30 hit for BRD1

![Figure 3.22 SGC compound 3.30 hit for BRD1](image.png)

**Figure 3.23a** 3.30 and BRD1 (*red*) 10:1. BRD1 apo protein (*blue*)

![Figure 3.23a 3.30 and BRD1 (red) 10:1. BRD1 apo protein (blue)](image.png)
3.3.7 Aggregation Studies

Previous $^1$H$^{15}$N HSQC NMR of a 12:1 ratio of 3.25: BRD1 showed a similar amount of peak movement to the 3:1 ratio of 3.25: BRD1. This suggested that there is a problem knowing the effective concentration of the compound in solution. Furthermore, when the NMR samples were centrifuged some precipitate was observed suggesting that not all the compound was dissolved in solution.

To investigate these observations, aggregation studies were performed on all three compounds. DMAP, a water-soluble organic compound, was used as a control to validate the methodology. The compounds were studied over the same concentration ranges used to perform the HSQC NMR experiments (0.5 – 2.0) mM and the exact same conditions were used except that BRD1 was not present.

Rather than compare signal intensity versus concentration, the ratio of an integrated signal versus the signal-to-noise ratio was plotted against concentration. The ratio of signal-to-noise was used
as an internal standard to corrected for any deviations in sample acquisition and processing. Each acquisition was done using the same parameters so that the signal-to-noise ratio would be identical in every experiment independent of a compounds concentration.

Table 3.6 Acquisition and processing parameters

<table>
<thead>
<tr>
<th>Instrument and processing parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrophotometer</td>
<td>600 MHz</td>
</tr>
<tr>
<td>Receiver gain (rg)</td>
<td>48</td>
</tr>
<tr>
<td>Number of scans</td>
<td>128</td>
</tr>
<tr>
<td>Line broadening</td>
<td>1/at (at = acquisition time)</td>
</tr>
<tr>
<td>r1</td>
<td>sfrq*0.2</td>
</tr>
</tbody>
</table>

r1 = spectral region; sfrq = signal frequency

Since all the samples were prepared in aqueous solutions identical to those prepared for HSQC measurements any spectrum would be dominated by a large a water peak. To account for this, a NMR technique known as water suppression was used. In this method, the $^1\text{H}_2\text{O}$ signal is presaturated with a weak radio frequency field. This process equalizes the spin states so that no signal is detected during acquisition.

![Figure 3.24 DMAP stacked spectra at various concentrations](image-url)
Figure 3.25 Example data for compounds, stacked spectra of 3.27 at various concentrations

In order to calculate the ratio of an integrated signal to the signal-to-noise ratio, the following processing steps were done. A selected region was chosen that contained the signal of interest (integrated) as well as a region of baseline. Dsnmax(r1), is defined as the ratio of the integral to the maximum signal-to-noise ratio over a defined region and this command was used to calculate various values for each spectrum.

Table 3.7 dsnmax(r1) values for various concentration of DMAP and compounds of interest

<table>
<thead>
<tr>
<th>Concentration (mM)</th>
<th>DMAP</th>
<th>3.28</th>
<th>3.25</th>
<th>3.27</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>28.618</td>
<td>4.483</td>
<td>19.396</td>
<td>5.681</td>
</tr>
<tr>
<td>1.0</td>
<td>34.265</td>
<td>8.966</td>
<td>20.718</td>
<td>11.485</td>
</tr>
<tr>
<td>1.5</td>
<td>68.244</td>
<td>11.659</td>
<td>20.216</td>
<td>19.976</td>
</tr>
<tr>
<td>2.0</td>
<td>79.756</td>
<td>21.706(^a)</td>
<td>24.049</td>
<td>22.181</td>
</tr>
</tbody>
</table>

\(^a\) Outlier, point excluded from plot
**Figure 3.26** Change in dsnmax(r1) with increasing concentrations of DMAP

**Figure 3.27** Change in dsnmax(r1) with increasing concentrations of 3.28

**Figure 3.28** Change in dsnmax(r1) with increasing concentrations of 3.25
Figure 3.29 Change in dsnmax(r1) with increasing concentrations of 3.27

Table 3.8 Workable concentrations for compounds of interest

<table>
<thead>
<tr>
<th>Compound</th>
<th>Workable concentration range (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.28</td>
<td>0.5 – 1.0</td>
</tr>
<tr>
<td>3.25</td>
<td>~ 0.5&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>3.27</td>
<td>0.5 – 1.5</td>
</tr>
</tbody>
</table>

<sup>a</sup>Sample not very soluble at low concentrations

3.3.8 ¹H¹⁵N HSQC Binding Curve

Six HSQC NMR samples were prepared using 150µM of N-15 BRD1 and the following conditions. By plotting the changes in chemical shift as the concentration of 3.28 increases, a binding constant can be derived.

Table 3.9 NMR sample preparation for HSQC binding curve

<table>
<thead>
<tr>
<th>Sample</th>
<th>Concentration of 3.28 (µM)</th>
<th>Ratio of 3.28: BRD1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0:1</td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>2:1</td>
</tr>
<tr>
<td>3</td>
<td>600</td>
<td>4:1</td>
</tr>
<tr>
<td>4</td>
<td>900</td>
<td>6:1</td>
</tr>
<tr>
<td>5</td>
<td>1200</td>
<td>8:1</td>
</tr>
<tr>
<td>6</td>
<td>1500</td>
<td>10:1</td>
</tr>
</tbody>
</table>

We are currently analyzing these results and experiment may be repeated for accuracy and precision.
3.3.9 3D NMR using C-13 N-15 BRD1

BRD1 was expressed and purified using C-13 glucose and N-15 ammonium chloride to give doubly labeled BRD1. We are currently in the process of assigning the backbone using the following 3D NMR experiments: HNCA, HNCO, CBCACONH, HBHACOHN and $^{15}$N-NOESY-HSQC.

3.3.10 Cocrystallization studies with BRD1

Based on the promising HSQC binding results the exact mode of binding was investigated through crystallization efforts. Cocrystallization experiments were first set up using the sitting drop technique. Each drop contained 1µL of BRD1 at 10mg/mL (700 µM final concentration), 1µL of the reservoir solution (E08 or F08), and 0.2 µL of the compound (3500 µM final concentration). This gave a ratio of compound: BRD1 of 5:1. The reservoir solutions were chosen as they had previous been proven to give crystallization of the apo protein. Two different techniques were used for each compound. The first technique allowed for the protein and compound to be incubated prior to mixing with the reservoir solution and the other technique was adding each component successively without any incubation time. All samples were set up in duplicate and the apo protein was used a control in four of the wells.

<table>
<thead>
<tr>
<th></th>
<th>1. 3.25</th>
<th>2. 3.25</th>
<th>3. 3.27</th>
<th>4. 3.27</th>
<th>5. 3.28</th>
<th>6. 3.28</th>
<th>7. BRD1 only</th>
<th>8. BRD1 only</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>RWE8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 µL BRD1 + 1 µL reservoir</td>
<td>1 µL BRD1 + 1 µL reservoir</td>
</tr>
<tr>
<td>B</td>
<td>RWF8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 µL BRD1 + 1 µL reservoir</td>
<td>1 µL BRD1 + 1 µL reservoir</td>
</tr>
<tr>
<td>C</td>
<td>RWE8</td>
<td>1 µL BRD1 + 1 µL reservoir + 0.2 µL compound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>RWF8</td>
<td>1 µL BRD1 + 1 µL reservoir + 0.2 µL compound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rows A/B – protein incubated with compound prior to mixing; Rows C/D – 1 µL BRD1 + 1 µL reservoir + 0.2 µL compound. BRD1 Buffer: 20 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer at pH 7.5 containing 200 mM NaCl and 1mM TCEP (tris(2-carboxylethyl)phosphine). Reservoir solution E08: 30.0 PEG2kM, 0.2 mM KBr Reservoir solution F08: 30.0 PEF 2kM, 0.1 mM K Thiocyanate

Plate 3.1 Cocrystallizations screen of 3.25, 3.27, and 3.28

No crystals formed in the wells containing the compounds of interest however the apo protein crystallized in wells A7-A8 and B7-B8.
A second screen of cocrystallization was done using the Red Wings Screen (96-Well screen). The same ratio (5:1) of compound: BRD1 was used and the protein was allowed to incubate with the compound overnight. Any compound that did not dissolve in the protein solution was removed by centrifugation. The sitting drop technique was used and 0.8 µL of both the solution containing the compound and BRD1 and the reservoir solution were added together. After four days crystals were observed (see Table 3.10).

**Table 3.10** Results of plate 3.1, Red Wings screen of crystallization conditions; TMANO (Trimethylamine-N-oxide dihydrate)

<table>
<thead>
<tr>
<th>Well</th>
<th>Compound</th>
<th>Condition</th>
<th>Quality of crystals</th>
</tr>
</thead>
<tbody>
<tr>
<td>G8</td>
<td>3.25</td>
<td>20.0 PEG 2kM, 0.2 TMANO, 0.1 Tris pH 8.5</td>
<td>Medium crystal</td>
</tr>
<tr>
<td>H10 (upper drop)</td>
<td>3.25</td>
<td>28.0 PEG 2kM, 0.1 Bis-Tris, pH 6.5</td>
<td>Large, well defined.</td>
</tr>
<tr>
<td>H10 (lower drop)</td>
<td>3.27</td>
<td>28.0 PEG 2kM, 0.1 Bis-Tris, pH 6.5</td>
<td>Several small crystals</td>
</tr>
</tbody>
</table>

The crystals in well H10 (upper drop) corresponding to 3.25 and BRD1 were frozen and mounted for diffraction. Because the crystals have a tendency to break when transferring into the cryobuffer as well as in the cryobuffer, it is possible to mount several crystals from one or two original crystals. This was the case for our experiment and we were able to screen various cryoprotectants to optimize X-ray resolution. The crystals are also covered in paratone oil, which helps remove water from the protein crystals.

**Table 3.11** 3.25 BRD1 co-crystals cryoprotectant optimizations

<table>
<thead>
<tr>
<th>Position on pack PY6</th>
<th>Cryoprotectant (% in H10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>12% glycerol</td>
</tr>
<tr>
<td>4</td>
<td>12% glycerol</td>
</tr>
<tr>
<td>5</td>
<td>15% ethylene glycol</td>
</tr>
<tr>
<td>6</td>
<td>15% ethylene glycol</td>
</tr>
<tr>
<td>7</td>
<td>15% ethylene glycol</td>
</tr>
<tr>
<td>8</td>
<td>15% ethylene glycol</td>
</tr>
</tbody>
</table>

All crystals also protected in paratone and frozen in liquid nitrogen.

Diffraction of the crystals proved that 3.25 did not co-crystallize with BRD1. However, 15% ethylene glycol was found to best the best cryoprotectant giving a resolution of 2.1Å.
Co-crystallizations of 3.27 with BRD1 were optimized to give better crystals by varying the concentration of the compound with BRD1. More co-crystals for 3.25 with BRD1 were also set up. Furthermore, small crystals for 3.25 in well G8 of plate 3.1 and these conditions will be optimized for crystal growth.

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 3.25</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>1:1</td>
<td>-</td>
</tr>
<tr>
<td>B 3.27</td>
<td>1:1.1</td>
<td>1:1.2</td>
<td>1:1.4</td>
<td>1:1.6</td>
<td>1:1.4</td>
</tr>
</tbody>
</table>

**Plate 3.2** Cocry stallizations of 3.25 and 3.27 with BRD1 using conditions H10. Ratio of BRD1: Compound indicated. BRD1 Buffer: 20mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer at pH 7.5 containing 200mM NaCl and 1mM TCEP (tris(2-carboxylethyl)phosphine).

### 3.3.11 Crystallization by soaking BRD1 crystals with compounds

To prepare BRD1 crystals for soaking the hanging drop technique was used. BRD1 protein (10mg/mL) was combined with Red Wing reservoir solutions E08 and F08 compatible for BRD1 crystal growth.

<table>
<thead>
<tr>
<th></th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A E08</td>
<td>1.5 +1.5 1.0+1.0</td>
<td>1.5 +1.5 1.0+1.0</td>
<td>1.0+1.0</td>
<td>1.0+1.0</td>
<td>1.0+1.0</td>
<td>1.0+1.0</td>
</tr>
<tr>
<td>B F08</td>
<td>1.5 +1.5 1.0+1.0</td>
<td>1.5 +1.5 1.0+1.0</td>
<td>1.0+1.0</td>
<td>1.0+1.0</td>
<td>1.0+1.0</td>
<td>1.0+1.0</td>
</tr>
</tbody>
</table>

**Plate 3.3** Hanging-drop crystallizations of BRD1. Volumes (µL) of BRD1 and reservoir solution indicated for each drop. Two drops were prepared for wells A1, A2, B1, and B2. BRD1 concentration at 10mg/mL. BRD1 Buffer: 20mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer at pH 7.5 containing 200mM NaCl and 1mM TCEP (tris(2-carboxylethyl)phosphine).

All wells gave crystals however only some gave well defined crystals suitable for soaking. Another hanging-drop crystallization plate for BRD1 was set up, which varied the relative concentrations of BRD1 with reservoir solution. The condition: 1 µL BRD1 (10 mg/mL) with 1.5 µL E08 gave very large well defined crystals. By increasing the volume of reservoir solution, the concentration of BRD1 was more dilute thereby optimizing crystallization growth.
To prepare the soaking experiment, the compound was first prepared by combing 48 μL of the reservoir solution (E08) with 2 μL of the compound (100 mM in DMSO). Since the reservoir solution does not contain any buffer, each solution pH was measure with a pH micrometer and dilute HCl or dilute NaOH was added dropwise to obtain a pH in the range 6.5 – 8.5. In each well, 8 μL of the prepared solution (4 mM compound) was added and a crystal was carefully transferred into each well using a fiber loop from the hanging drop crystallization plate (Plate 3.4) with the aide of a microscope.

<table>
<thead>
<tr>
<th>1.</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3.28</td>
</tr>
<tr>
<td>B</td>
<td>3.25</td>
</tr>
<tr>
<td>C</td>
<td>3.27</td>
</tr>
</tbody>
</table>

Plate 3.4 Soaking of compounds (8 μL at 4 mM with BRD1 apo crystal)

The soaking experiments were allowed to proceed over two days. The crystals were then transferred to the optimized cryoprotectant (15% ethylene glycol) and mounted. X-ray diffraction of all 12 crystals showed only the apo protein without any compound bound.
The apo crystal structure of BRD1 contains a compound bound in the Ar-cage. Although we have achieved a resolution of 2.1 Å in the crystal structure it is not possible to know what the density seen by X-ray corresponds to. From a design standpoint, knowing the structure of this molecule would be very useful, as a lot of insight would be gained from its mode of binding. However, from an inhibitor standpoint, if we were to design a fragment that out-competed the currently bound fragment then the structure of the original one is not as important. If the first approach is taken, it may be possible to figure out what the original fragment is by extracting it out of the protein by denaturing and removing the protein. The remaining solution could be analyzed by mass spectroscopy.

3.3.11 Isothermal titration calorimetry, ITC

Although no cocrystal structures were resolved with any of the compounds, the 1H15N HSQC NMR data still proved that the compounds were binding to BRD1. To investigate the strength and thermodynamics of binding, ITC runs were conducted using a MicroCal VP-ITC and titrating our compounds into BRD1.

The first run was done using a positive control, 3.30. A $K_d$ value of 400 $\mu$M was previously calculated for this compound. Based on the larger magnitude of chemical shift changes for 3.28 compared to 3.30 from $^1$H$^{15}$N HSQC, we expected the $K_d$ value for 3.28 to be smaller. Further more 3.30 had a $\Delta T$ of 1.5$^o$C whereas 3.28 had a delta $\Delta T$ of 2.9$^o$C by DSF.
BRD1 containing his-tag was dialyzed in a phosphate buffer pH 7.5, however the protein precipitated from solution overnight. Agitating this solution seemed to re-dissolved BRD1 and we attempted to use it for our measurements (1mM final concentration). **3.30** was prepared in a solution of the same buffer at 2.5mM concentration. Since this solution contained 2.5% DMSO, DMSO was added to the protein so that it would be identical to the solution used for the compound. Titrating **3.30** into BRD1 resulted in no binding. Unsure whether or not the instrument was dirty or if the protein was unsound, the instrument was re-cleaned and a second run using **3.28** was set up. Unfortunately, this compound was poorly soluble in the buffer conditions and no binding was observed with this compound as well.

The ITC conditions were optimized for both the protein and the compounds. Since the protein is known to be stable in a Tris-HCl buffer, this was selected and the cut BRD1 protein was used (no his-tag). The compounds are fairly hydrophobic so the concentration of DMSO was kept at 2.5% to ensure solubility.

Dialyzing BRD1 in freshly prepared buffer (20 mM Tris-HCl pH 7.5 containing 200 mM NaCl) was successful and the protein remained fully dissolved. The protein concentration was calculated to be 0.91mM using the NanoDrop Spectrophotometer. During preparation of BRD1 for ITC, DMSO was added to give a 2.5% DMSO solution. The compounds were prepared using 10 μL of the 100 mM concentrated sample and diluting to a volume of 500 μL with the buffer solution. The first ITC run was conducted on **3.28** and no binding was detected. A second ITC run was done using **3.27** with the same parameters and again, no binding was observed.

### 3.3.12 Reassigning the structure of **3.28**

Recently, the structure of **3.28** synthesized by Dr. Shinya Adachi has been re-assigned as the following:

![Reassigned structure of 3.28 by Dr. Shinya Adachi](image)
Although the structure was not as expected, it is very interesting nonetheless as it binds BRD1. This would not have been predicted based on the sterics about the amine. Modeling this compound computationally shows that it does fit inside the aromatic cage and the methylene and amine that connect the two aromatic moieties together act as a rotational point to achieve the best conformer. The glide score for the reassigned structure was calculated to be -4.469 compared to the original monomer unit of -5.115. More work is being done to synthesize the original monomer structure and study its mode of binding by NMR and crystallography.

![Docked compound of revised structure](image)

**Figure 3.32** Docked compound of revised structure 3.28 with BRD1

### 3.3.13 Setdb1 tetrahydroisoquinoline peptides

From soaking experiments, we found that the tetrahydroisoquinoline core gave a positive hit. This fragment was also found to bind to Setdb1 by $^1$H$^{15}$N HSQC NMR experiments.

![Tetrahydroisoquinoline](image)

**Figure 3.33** Setdb1 tetrahydroisoquinoline, TIC, fragment hit 3.29
Figure 3.34 TIC 3.29 (orange) bound to Setdb1 (green) cocrystal structure. Hydrogen bond contact (pink)

Figure 3.35a. Overlay of 3.29 bound to setdb1 (3:1) (red) and apo protein (blue)
Figure 3.35b. Zoomed regions: Overlay of 3.29 fragment bound to Setdb1 (3:1) (red) and apo protein (blue)

A peptide was also found to bind Setdb1 by $^1$H$^{15}$N HSQC NMR and X-ray crystallography (Figure 3.36 and Figure 3.37).

Figure 3.36 Overlay of peptide bound to Setdb1 (3:1) (red) and apo protein (blue).
Taking these two observations together, we sought to synthesis various peptides containing the tetrahydroisoquinoline moiety. These hybrid-peptides would be advantageous as the TIC component would make favourable contacts in the aromatic cage of the protein and the peptide would extend outside of the aromatic cage and make complimentary secondary interactions. Currently these peptides are being designed and synthesized.

3.4 Conclusion and Future Work

Using Glide software, pyrrolidine fragments were successfully designed and docked in the aromatic cage of BRD1. Pyrrolidine analogues were synthesized using a microwave reaction and screened by DSF and Stargazer. Four compounds showed binding however, a dose response assay and cocrystallization by soaking both gave negative results. An isoindoline scaffold was thought and although a benzyl isoindoline proved advantageous by computational docking, the synthesis of these compounds was obstructed by their inherent instability. Finally, a phthalamide based compound was designed, which gave better docking scores compared to any of our previous molecules. Although the target molecule still needs to be synthesized, each intermediate along with select compounds synthesized by Dr. Shinya Adachi were assessed for binding with BRD1 using HSQC chemical shift mapping. Three compounds were identified to bind BRD1 and further analysis by crystalloography and ITC is ongoing.
3.5 Experimental Details

3.5.1 General information

All reactions were carried out under an atmosphere of air unless otherwise stated. Reagents were used as obtained from commercial suppliers or purified according to standard procedures. Flash chromatography was performed using SiliaFlash P60 40 – 63 um silica. TLC was performed on EMD silica gel 60 F254 TLC glass plates and visualized with UV light, I$_2$ on silica, or permanganate stain. All $^1$H and $^{13}$C NMR spectra were recorded using a Bruker 400 MHz or Varian 400, 600, 700, or 500 MHz cryoprobe spectrophotometer and internally referenced to a residual protiosolvent. 2D (COSY, HSQC) NMR spectroscopy was used where appropriate to assist in the assignment of $^1$H and $^{13}$C NMR spectra. LC-MS (ESI) was recorded on an Agilent 1200 Series quadrupole spectrometer.

3.5.2 Characterization pyrrolidine analogues by microwave synthesis

*General procedure for microwave synthesis*\textsuperscript{15}

The aniline derivative (1.0 mmol) and 1,4-dibromobutane (131 uL, 1.1 mmol) were added to a solution of K$_2$CO$_3$ in H$_2$O (2 mL) in a 2-5 mL MW vial containing a stir bar. The vial was sealed and MW at 120$^\circ$C for 30 minutes. Upon completion, the vial was opened and EtOAc was added to the mixture and vigorously shaken. The organic layer was removed and the aqueous re-extracted with EtOAc (2 x 1 mL) and the combined organic layers were dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated by rotary evaporation and purified by column chromatography (20%EtOAc in hexanes or 100% CH$_2$Cl$_2$). The purified compound was then characterized by $^1$H and $^{13}$C NMR in CDCl$_3$.

\chem{\begin{picture}(30,30)
  \put(15,15){\line(1,0){5}}
  \put(15,15){\line(0,1){5}}
  \put(20,20){\line(0,1){5}}
  \put(15,15){\line(1,1){5}}
  \put(15,15){\line(-1,1){5}}
  \put(18,17.5){N}
\end{picture}}

**1-phenylpyrrolidine (3.1)\textsuperscript{15}**: Yield = 33 mg, 23%. $^1$H (400 MHz, CDCl$_3$) $\delta$ 7.16 – 7.13 (m, 2H), 6.59 – 6.56 (m, 1H), 6.50 – 6.48 (m, 2H), 3.21 – 3.18 (m, 4H), 1.93 – 1.90 (m, 4H). $^{13}$C (100 MHz, CDCl$_3$) $\delta$ 148.1, 129.2, 115.5, 111.8, 47.7, 25.6
1-(4-bromophenyl)pyrrolidine (3.2)\textsuperscript{15}: Yield = 895 mg, 40%. \textsuperscript{1}H (400 MHz, CDCl\textsubscript{3}) \(\delta\) 7.28 (d, \(J = 9.0\) Hz, 2H), 6.43 (d, \(J = 8.9\) Hz, 2H), 3.26 – 3.22 (m, 4H), 2.02 – 1.99 (m, 4H). \textsuperscript{13}C (100 MHz, CDCl\textsubscript{3}) \(\delta\) 146.9, 131.8, 113.4, 47.9, 25.6.

1-(4-methoxy-2-nitrophenyl)pyrrolidine (3.3)\textsuperscript{15}: Isolated 4 mg. \textsuperscript{1}H (400 MHz, CDCl\textsubscript{3}) \(\delta\) 7.30 (d, \(J = 3.0\) Hz, 1H), 7.05 (dd, \(J = 9.3, 3.1\) Hz, 1H), 6.91 (d, \(J = 9.3\) Hz, 1H), 3.78 (s, 3H), 3.20 – 3.14 (m, 4H), 2.00 – 1.94 (m, 4H). \textsuperscript{13}C (100 MHz, CDCl\textsubscript{3}) \(\delta\) 150.2, 138.5, 136.5, 122.5, 117.5, 109.4, 56.1, 50.8, 25.9.

1-(4-fluorophenethyl)pyrrolidine (3.4)\textsuperscript{15}: Yield = 86 mg, 45%. \textsuperscript{1}H (400 MHz, CDCl\textsubscript{3}) \(\delta\) 7.16 (dd, \(J = 8.5, 5.5\) Hz, 2H), 6.96 (t, \(J = 8.7\) Hz, 2H), 2.88 – 2.64 (m, 4H), 2.58 (dd, \(J = 4.9, 2.4\) Hz, 4H), 1.88 – 1.76 (m, 4H). \textsuperscript{13}C (100 MHz, CDCl\textsubscript{3}) \(\delta\) 162.7, 130.1, 115.4, 115.1, 58.5, 54.4, 35.0, 23.6.

3.4.3 Synthesis of more pyrrolidine based fragments and other small molecules

1-tosylpyrrolidine (3.5)\textsuperscript{17}: Pyrrolidine (45 uL, 0.5 mmol) and sodium p-toluenesulfonate monohydrate (267 mg, 1.5 mmol) was dissolved in H\textsubscript{2}O in a flask. The flask was cooled using an ice-water bath and NaHClO\textsubscript{4} was added dropwise to give a white precipitate. The reaction was allowed to warm to room temperature and was left to stir overnight. The white solid was isolated by filtration and then dried in the dessicator. Yield (107 mg, 95%). \textsuperscript{1}H NMR (400 MHz, CDCl\textsubscript{3}) \(\delta\) 7.71 (d, \(J = 8.2\) Hz, 2H), 7.31 (d, \(J = 7.9\) Hz, 2H), 3.29 – 3.16 (m, 4H), 2.43 (s, 3H), 1.78 – 1.68 (m, 4H). \textsuperscript{13}C NMR (100 MHz, CDCl\textsubscript{3}) \(\delta\) 143.4, 134.1, 129.7, 127.7, 48.0, 25.3, 21.6.
Phenyl(pyrrolidin-1-yl)methanone (3.7): Pyrrolidine (45 uL, 0.5 mmol) and Et$_3$N (83.7 uL, 0.6 mmol) were dissolved in CH$_2$Cl$_2$ (1 mL) and the reaction was cooled to 0°C. Benzylchloride (43 uL, 0.5 mmol) was added dropwise and following this addition the mixture was allowed to warm to room temperature and stirred for an additional 2 h. Following completion, as determined by TLC (100% CH$_2$Cl$_2$), the reaction was hydrolyzed with H$_2$O and extracted with EtOAc (3 x 2 mL). The combined organic layers were washed with diluted HCl (3 x 2 mL), 9% NaHCO$_3$ (3 x 2 mL), dried over anhydrous Na$_2$SO$_4$, filtered, and concentrated by rotary evaporation to give a clear pale yellow liquid. Yield = 66.6 mg, 76%. $^1$H NMR (400 MHz, CDCl$_3$) δ 7.60 – 7.33 (m, 5H), 3.53 (m, 4H), 1.92 (m, 4H). $^{13}$C NMR (100 MHz, CDCl$_3$) δ 169.9, 137.3, 129.9, 128.4, 127.2, 49.9, 21.2.

(E)-3-phenyl-1-(pyrrolidin-1-yl)prop-2-en-1-one (3.8): Pyrrolidine (45 uL, 0.5 mmol) was dissolved in CH$_2$Cl$_2$ and cooled to 0°C. Pyridine (40 uL, 0.5 mmol) was added followed by cinnamoylchloride (100 mg, 0.6 mmol). The reaction was stirred for 15 min at 0°C then warmed to room temperature and allowed to stir for an additional 2 h. Following completion as observed by TLC, the reaction was washed with H$_2$O, HCl, NaOH, and brine in succession, dried over anhydrous Na$_2$SO$_4$, filtered and concentrated to a pale yellow liquid by rotary evaporation. Yield = 73.5 mg, 73%. $^1$H NMR (400 MHz, CDCl$_3$) δ 7.93 – 7.31 (m, 7H), 6.78 – 6.67 (m, 1H), 3.53 (m, 4H), 1.92 (m, 4H). $^{13}$C NMR (100 MHz, cdcl$_3$) δ 164.8, 129.6, 128.8, 118.9, 46.7, 26.2.

1-tosylpiperidine (3.6): Piperidine (49 uL, 0.5 mmol) and sodium p-toluenesulfonate monohydrate (267 mg, 1.5 mmol) was dissolved in H$_2$O in a flask. The flask was cooled using an ice-water bath and NaHClO$_4$ was added dropwise to give a white precipitate. The reaction was allowed to warm to room temperature and was left to stir overnight. The white solid was isolated.
by filtration and then dried in the dessicator. Yield (110 mg, 92%). $^1$H NMR (400 MHz, CDCl$_3$) δ 7.64 (d, $J = 8.2$ Hz, 2H), 7.31 (d, $J = 8.2$ Hz, 2H), 3.02 – 2.91 (m, 4H), 2.43 (s, 3H), 1.63 (p, $J = 5.8$ Hz, 4H), 1.41 (t, $J = 5.8$ Hz, 2H). $^{13}$C NMR (100 MHz, cdcl$_3$) δ 143.4, 133.5, 129.7, 127.9, 77.2, 47.1, 25.3, 23.7, 21.7.

1-(pyrrolidin-1-yl)pentan-1-one (3.9): Pyrrolidine (45 uL, 0.5 mmol) was dissolved in CH$_2$Cl$_2$ and cooled to 0°C. Triethylamine (84 uL, 0.6 mmol) was added followed by valeryl chloride (59 uL, 0.5 mmol). The reaction was stirred for 15 min at 0°C then warmed to room temperature and allowed to stir for an additional 2 h. Following completion as observed by TLC, the reaction was washed with H$_2$O, HCl, NaOH, and brine in succession, dried over anhydrous Na$_2$SO$_4$, filtered and concentrated to a pale yellow liquid by rotary evaporation. Yield = 67 mg, 86%. $^1$H NMR (400 MHz, CDCl$_3$) δ 3.44 – 3.30 (m, 4H), 2.25 – 2.14 (m, 2H), 1.95 – 1.84 (m, 2H), 1.84 – 1.72 (m, 2H), 1.57 (dddd, $J = 8.9$, 7.6, 6.9, 5.8 Hz, 2H), 1.38 – 1.25 (m, 2H), 0.87 (t, $J = 7.3$ Hz, 3H). $^{13}$C NMR (100 MHz, CDCl$_3$) δ 171.9, 77.2, 46.6, 45.6, 34.6, 27.1, 26.2, 24.4, 22.6, 13.9.

Pd(N,N-dimethyl-B-alaninate)$_2$ (3.15)$^{23}$: A mixture of N,N-dimethylbetaalanine hydrochloride (9.2 mg, 0.06 mmol) and Na$_2$PdCl$_4$ (8.8 mg, 0.03 mmol) were stirred in H$_2$O (0.5 mL) at room temperature in a 1-dram vial for 10 minutes. 10% NaOH$_{(aq)}$ was added dropwise the mixture to achieve a pH = 8. The reaction was concentrated to dryness via rotary evaporation and the yellow oil was taken up in CHCl$_3$ and the insoluble particulates were removed by filtration. The filtrate was concentrated to dryness via rotary evaporation to the product. Yield = 7.6 mg, 75%. $^1$H NMR (400 MHz, CDCl$_3$) 2.59 – 2.56 (m, 16 H), 2.50 – 2.42 (m, 4H). $^{13}$C NMR (101 MHz, CDCl$_3$) 176.1, 60.1, 49.0, 34.6.
1-Phenylurea (3.13): Aniline (980 uL, 10.7 mmol) was dissolved in 2M HCl (6 mL). The reaction was cooled to 2°C with an ice-bath and KNCO (1.04 g, 12.84 mmol) was added in portions over 15 min with stirring. The mixture was warmed to room temperature and allowed to stir for an additional four hours. Upon completion, the white precipitate was isolated by vacuum filtration and washed with ice-cold distilled water. The compound was dried overnight in a desiccator under vacuum. Yield = 1.43 g, 98%. $^1$H NMR (400 MHz, DMSO-$d_6$) δ 8.48 (s, 1H), 7.37 (dd, $J$ = 8.6, 1.3 Hz, 1H), 7.19 (dd, $J$ = 8.6, 7.3 Hz, 2H), 6.91 – 6.82 (m, 1H), 5.80 (s, 2H). $^{13}$C NMR (101 MHz, DMSO) δ 155.9, 140.5, 128.5, 121.0, 117.7.

1-benzylpyrrolidine (3.18): Pyrrolidine (98 uL, 1.1 mmol), benzaldehyde (102 uL, 1.0 mmol) were added to THF (10 mL). To this mixture, sodium triacetoxyborohydride (297 mg, 1.4 mmol) was added portion-wise to give a cloudy suspension. The reaction was allowed to proceed at room temperature and was monitored by TLC. Mass spec of the crude mixture showed product (m/z = 162.2 = M+H)$^+$. The reaction was then concentrated by rotary evaporation, basified with NaOH, and the product extracted with EtOAc, dried with anhydrous sodium sulfate, and concentrated by rotary evaporation. The crude product was isolated as a yellow oil: 140 mg, 87%. $^1$H NMR (400 MHz, Chloroform-$d$) δ 7.98 – 7.29 (m, 6H), 3.74 (s, 2H), 2.70 – 2.57 (m, 4H), 1.88 – 1.79 (m, 4H). $^{13}$C NMR (101 MHz, CDCl$_3$) δ 129.5, 128.5, 127.6, 60.0, 53.6, 23.5.

2-phenylisoindoline (3.19): Aniline (91 uL, 1.0 mmol) and $\alpha,\alpha'$-dibromo-o-xylene (290 mg, 1.1 mmol) were added to a solution of K$_2$CO$_3$ (152 mg, 1.1 mmol) in H$_2$O (2 mL) in a 2 mL microwave vial equipped with a stir bar. The vial was sealed and the reaction was microwaved for 30 min at 120°C. Following completion the mixture was allowed to cool and a quick analysis by both TLC and mass spec gave evidence to product formation (m/z = 196 = M+H)$^+$. The mixture was extracted with EtOAc (3 x 1.5 mL), dried with anhydrous Na$_2$SO$_4$, filtered, and isolated by rotary evaporation. The crude residue was purified by column chromatography (SiO$_2$, 10% EtOAc in hexanes) to give a pink solid: 64 mg, 33%. $^1$H NMR (400 MHz, Chloroform-$d$) δ
7.28 (d, $J = 8.8$ Hz, 1H), 6.74 – 6.57 (m, 2H), 4.59 (s, 3H). $^{13}$C NMR (101 MHz, CDCl$_3$) $\delta$ 147.3, 138.0, 129.5, 127.3, 122.8, 116.4, 111.8, 54.0.

1-[(1,1'-biphenyl]-4-yl)pyrrolidine (3.16)$^{25}$: $N$-(4-bromophenyl)pyrrolidine (20 mg, 0.088 mmol), phenylboronic acid (21.6 mg, 0.176 mmol), Pd($N,N$-dimethyl-B-alaninate)$_2$, (0.01 mol%), and K$_3$PO$_4$ (56 mg, 12 mmol) were added to a small vial containing a stir bar. To this, a solution of EtOH: H$_2$O (1:1, 292 uL) was added and the reaction was heated to 50°C for 1 h. After 20 min, the reaction afforded a black precipitate identified as Pd(0). By mass spec of the crude reaction mixture, the product was identified (m/z = 224.1= M+H)$^+$. The reaction was cooled to room temperature and then extracted with diethyl ether. The column conditions were re-optimized and 10% diethyl ether in hexanes was chosen as the best eluent. The product was isolated using this method and gave a white crystalline solid: 15.4 mg, 78%. $^1$H NMR (400 MHz, Chloroform-$d$) $\delta$ 7.65 – 7.21 (m, 7H), 6.74 – 6.62 (m, 2H), 3.43 – 3.34 (m, 4H), 2.13 – 1.97 (m, 4H). $^{13}$C NMR (101 MHz, CDCl$_3$) $\delta$ 147.5, 141.6, 131.8, 128.7, 127.9, 126.3, 125.9, 112.1, 47.8, 25.6.

(4-bromophenyl)methanamine (3.20)$^{29}$: NH$_4$OH·HCl (83 mg, 1.2 mmol) was added to EtOH (600 uL) and H$_2$O was added dropwise with vigorous stirring to ensure dissolution. 4-bromobenzaldehyde (185 mg, 1.0 mmol) was added and the reaction was allowed to stir at room temperature for 30 min. Concentrated HCl (300 uL, 4.0 mmol) was carefully added dropwise followed by zinc dust (166 mg, 2.5 mmol). After a further 15 minutes of reaction, NH$_3$·H$_2$O (30%, 286 uL) was added to give a clear colourless solution. This was extracted with CH$_2$Cl$_2$ (3 x 4mL), dried with anhydrous Na$_2$SO$_4$, filtered, and concentrated by rotary evaporation to give a pure white semisolid with no further purification necessary (117 mg, 63%). $^1$H NMR (400 MHz, Chloroform-$d$) $\delta$ 7.44 (d, $J = 8.3$ Hz, 2H), 7.19 (d, $J = 8.3$ Hz, 2H), 3.82 (s, 2H). $^{13}$C NMR (101 MHz, CDCl$_3$) $\delta$ 142.2, 131.7, 129.0, 120.6, 45.9.
2-(4-bromobenzyl)isoindoline-1,3-dione (3.21): Phthalic anhydride (56.9 mg, 0.38 mmol) was suspended in glacial acetic acid (4 mL) and 4-bromobenzylamine (100 mg, 0.54 mmol) was slowly added. The reaction flask was fitted with a condenser was heated to reflux for 4.5 h. Upon completion, the hot solution was poured over ice and a white crystalline solid immediately formed. This was isolated by filtration, washed with cold H₂O and dried overnight in a dessicator equipped with a vacuum. Yield = 108 mg, 90%. No further purification was needed.¹H NMR (400 MHz, Chloroform-d) δ 7.85 (dd, J = 5.4, 3.1 Hz, 1H), 7.72 (dd, J = 5.5, 3.0 Hz, 1H), 7.44 (d, J = 8.4 Hz, 1H), 7.31 (d, J = 8.4 Hz, 1H), 4.79 (s, 1H).

¹³C NMR (101 MHz, CDCl₃) δ 168.1, 135.5, 134.2, 132.2, 132.0, 130.6, 123.6, 122.1, 41.1.

1H-pyrazole-5-boronic acid MIDA ester (3.22): The boronic acid (100 mg, 1.79 mmol) and MIDA (145 mg, 1.97 mmol) were dissolved in DMF and heated at 80°C for 20 h. Upon completion of the reaction, the solvent was removed en vacuo to give a white crystalline powder in quantitative yield. ¹H NMR (400 MHz, DMSO-d₆) δ 7.48 (d, J = 1.6 Hz, 1H), 6.30 (d, J = 1.6 Hz, 1H), 4.34 (dd, J = 17.3, 1.3 Hz, 2H), 4.12 – 4.01 (m, 2H), 2.63 – 2.55 (m, 4H).

2-(4-bromobenzyl)isoindoline-1,3-dione (3.21): Potassium phthalalide (100 mg, 0.54 mmol) and p-bromobenzylbromide (148 mg, 0.59 mmol) were added to DMF (2 mL) and were allowed to stir at room temperature for 24 h. The reaction was then concentrated to dryness en vacuo and CHCl₃ (10 mL) was added to the residue. This was then washed with H₂O (3 x 5 mL), dried over anhydrous sodium sulfate, filtered, and concentrated en vacuo. This residue was then purified by column chromatography (SiO2, 100% CHCl₃) to give a white solid, 106.3 mg, 63%. ¹H NMR (300 MHz, Chloroform-d) δ 7.84 (dd, J = 5.5, 3.0 Hz, 1H), 7.70 (dd, J = 5.4, 3.1 Hz, 1H), 7.46 –
7.39 (m, 1H), 7.35 – 7.28 (m, 1H), 4.78 (s, 1H). $^{13}$C NMR (75 MHz, CDCl$_3$) δ 168.0, 135.4, 134.2, 132.1, 131.9, 130.5, 123.5, 122.0, 41.1.

2-(3-(bromomethyl)benzyl)isoindoline-1,3-dione (3.25)$^{32}$: Potassium phthalamide (500 mg, 2.7 mmol) was added portion-wise to 1,3-bis bromobenzene in acetone while stirring. After complete addition, the reaction was heated at reflux for 18 h. The reaction was then concentrated en vacuo. The residue, CH$_2$Cl$_2$ (20 mL) was added and this was washed with H$_2$O (3 x 25 mL), dried over anhydrous sodium sulfate, filtered, and concentrated by rotary evaporation. The crude material was subsequently purified by column chromatography (SiO$_2$, 9:1 hexanes: ethyl acetate; Rf = 0.57) to give a white crystalline solid, 161 mg, 18 %. $^1$H NMR (399 MHz, Chloroform-d) δ 7.79 (dd, $J = 55.3$, 8.5 Hz, 2H), 7.51 – 7.29 (m, 2H), 4.84 (s, 1H), 4.46 (s, 1H). $^{13}$C NMR (100 MHz, cdcl$_3$) δ 168.1, 138.4, 137.1, 135.2, 134.2, 132.2, 129.4, 128.9, 128.8, 123.6, 41.5, 33.2.

2-benzylisoindoline-1,3-dione (3.23)$^{35}$: Phthalic anhydride (50 mg, 0.34 mmol) was suspended in glacial acetic acid (4 mL). Benzylamine (51 uL, 0.47 mmol) was slowly added and the mixture was heated to reflux for 4.5 h. Upon completion, the hot mixture was poured over ice and the resultant white precipitate was isolated by filtration, washed with ice-cold water and dried overnight. Yield = 79.4 mg, 94 %. $^1$H NMR (400 MHz, Chloroform-d) δ 7.85 (dd, $J = 5.5$, 3.1 Hz, 1H), 7.70 (dd, $J = 5.5$, 3.0 Hz, 1H), 7.47 – 7.40 (m, 1H), 7.36 – 7.27 (m, 1H), 4.85 (s, 1H). $^{13}$C NMR (101 MHz, CDCl$_3$) δ 168.2, 136.5, 134.1, 132.3, 128.8, 128.0, 123.5, 41.8.

2-(3-((1,3-dioxoisindolin-2-yl)methyl)phenyl)acetonitrile (3.26)$^{32}$: 3-bromobenzylphthalamide (50 mg, 0.152 mmol), and NaCN (37 mg, 0.76 mmol) were added to a
round bottom flask containing DMSO (1.5 mL). The reaction was allowed to proceed at ambient temperature for 3 days. Upon completion the DMSO was removed en vacuo and CH$_2$Cl$_2$ (10 mL) was added to the residue. This was washed with H$_2$O (3 x 5 mL), followed by brine (1 x 5 mL), dried over anhydrous sodium sulfate, filtered, and isolated by rotary evaporation. The crude material was subsequently purified by column chromatography (SiO$_2$, 100% CHCl$_3$, R$_f$ = 0.27) to give a white solid (15.1 mg, 30%). $^1$H NMR (300 MHz, Chloroform-$d$) $\delta$ 7.85 (dd, $J$ = 5.4, 3.2 Hz, 1H), 7.72 (dd, $J$ = 5.5, 3.0 Hz, 1H), 7.45 – 7.28 (m, 2H), 4.84 (s, 1H), 3.72 (s, 1H). $^{13}$C NMR (75 MHz, CDCl$_3$) $\delta$ 168.1, 137.5, 134.3, 132.1, 130.5, 129.7, 128.5, 128.4, 127.5, 123.6, 117.8, 41.4, 23.7.

2-benzyl-4,5,6,7-tetrafluoroisoindoline-1,3-dione (3.24)$^{36}$: 4-fluorophthalic anhydride (37 mg, 0.17 mmol), was suspended in glacial acetic acid (2 mL). Benzylamine (25 uL, 0.24 mmol) was slowly added and the mixture was heated to reflux for 4.5 h. Upon completion, the hot mixture was poured over ice and the resultant white precipitate was isolated by filtration, washed with ice-cold water and dried overnight. Yield = 50 mg, 95 %. $^1$H NMR (400 MHz, Chloroform-$d$) $\delta$ 7.47 – 7.29 (m, 5H), 4.81 (s, 2H). $^{13}$C NMR (101 MHz, CDCl$_3$) $\delta$ 162.3, 135.3, 129.1, 129.0, 128.5, 42.5. $^{19}$F NMR (377 MHz, CDCl$_3$) $\delta$ -135.4, -142.2.

2-(3-((1,3-dioxoisindolin-2-yl)methyl)phenyl)acetic acid (3.27)$^{33}$: Benzyl nitrile phthalimide (25 mg, 0.09 mmol) and 70% aqueous H$_2$SO$_4$ (1 mL) were heated for 2 h at 120°C. At the end of this time, the reaction was cooled and NH$_4$OH was added dropwise to achieve pH = 2. A white precipitate formed, which was collected by filtration, washed with cold distilled H$_2$O and left in the dessicator to dry overnight. Yield = 7. 1 mg, 27 %. m/z = 294.0 = (M-H)$^-$. $^1$H NMR (500 MHz, DMSO-$d_6$) $\delta$ 7.94 – 7.80 (m, 4H), 7.32 – 7.13 (m, 4H), 4.75 (s, 2H), 3.54 (s, 2H). $^{13}$C
NMR (126 MHz, DMSO) δ 172.6, 167.7, 136.7, 135.4, 134.6, 131.6, 128.5, 128.5, 128.2, 125.7, 123.3, 40.8, 40.5.
3.6 References

3. Department of Biochemistry and Molecular Biology, Research, Chromatin structure:
   alterations of epigenetic mechanisms in cancer and aging.
   biochem.slu.edu/faculty/gonzalo/researchprograms.shtml
9. GLIDE Software – Maestro.
   bcmp.med.harvard.edu/sites/bcmp.med.harvard.edu/files/facilities/ITC200%20training_pdf.pdf
34. Water Suppression – nmrfam. 
   http://www.nmrfam.wisc.edu/Documents/bchm800/notes/Water_suppression.pdf
Appendix 1. Chapter 1 NMR Spectra, X-ray, and LCMS Data
9:1
RC-01-125.3.8d
RC-01-125. columned
cChem_Carbon_Day CDCl3 /opt/data Please amend this.

[Chemical Structure Image]

20130706_mercury_400 RC-01-127 PROTON 01

[Chemical Structure Image]
(3S,5S,6S)-3-benzyl-N-(tert-butyl)-2-oxo-4-phenyl-1,4-diazabicyclo[4.1.0]heptane-5-carboxamide (1.19)

Table 1. Crystal data and structure refinement for d13120.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification code</td>
<td>d13120</td>
</tr>
<tr>
<td>Empirical formula</td>
<td>C23 H27 N3 O2</td>
</tr>
<tr>
<td>Formula weight</td>
<td>377.47</td>
</tr>
<tr>
<td>Temperature</td>
<td>147(2) K</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.54178 Å</td>
</tr>
<tr>
<td>Crystal system</td>
<td>Orthorhombic</td>
</tr>
<tr>
<td>Space group</td>
<td>P 2 1 2 1 2 1</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>a = 9.090(2) Å</td>
</tr>
</tbody>
</table>

\(a = 90^\circ\).
\[ b = 10.774(3) \, \text{Å} \quad \text{and} \quad b = 90^\circ. \]
\[ c = 21.624(5) \, \text{Å} \quad \text{and} \quad g = 90^\circ. \]

Volume: \( 2117.8(10) \, \text{Å}^3 \)

\[ Z = 4 \]

Density (calculated): 1.184 Mg/m\(^3\)

Absorption coefficient: 0.607 mm\(^{-1}\)

\[ F(000) = 808 \]

Crystal size: 0.260 \times 0.100 \times 0.040 \, \text{mm}^3

Theta range for data collection: 4.089 to 66.285°.

Index ranges: \(-10 \leq h \leq 9, \quad -12 \leq k \leq 12, \quad -15 \leq l \leq 25\)

Reflections collected: 15303

Independent reflections: 3626 [R(int) = 0.0280]

Completeness to theta = 67.679°: 96.2 %

Absorption correction: Semi-empirical from equivalents

Max. and min. transmission: 0.7528 and 0.6804

Refinement method: Full-matrix least-squares on F\(^2\)

Data / restraints / parameters: 3626 / 0 / 257

Goodness-of-fit on F\(^2\): 1.073

Final R indices [I>2\sigma(I)]: R1 = 0.0274, wR2 = 0.0692

R indices (all data): R1 = 0.0290, wR2 = 0.0705

Absolute structure parameter: -0.09(8)

Extinction coefficient: n/a

Largest diff. peak and hole: 0.103 and -0.165 e.Å\(^{-3}\)

**Table 2.** Atomic coordinates (x \times 10^4) and equivalent isotropic displacement parameters (Å\(^2\)x 10\(^3\)) for d13120. U(eq) is defined as one third of the trace of the orthogonalized \( U_{ij} \) tensor.

<table>
<thead>
<tr>
<th></th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>U(eq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O(1)</td>
<td>5646(2)</td>
<td>6942(1)</td>
<td>3541(1)</td>
<td>26(1)</td>
</tr>
<tr>
<td>O(2)</td>
<td>4388(2)</td>
<td>7667(1)</td>
<td>2158(1)</td>
<td>27(1)</td>
</tr>
<tr>
<td>N(1)</td>
<td>6484(2)</td>
<td>8864(1)</td>
<td>3280(1)</td>
<td>22(1)</td>
</tr>
<tr>
<td>N(2)</td>
<td>3591(2)</td>
<td>9557(1)</td>
<td>3024(1)</td>
<td>21(1)</td>
</tr>
<tr>
<td>N(3)</td>
<td>4608(2)</td>
<td>9212(1)</td>
<td>1441(1)</td>
<td>21(1)</td>
</tr>
<tr>
<td>C(1)</td>
<td>3857(2)</td>
<td>8586(2)</td>
<td>3490(1)</td>
<td>20(1)</td>
</tr>
<tr>
<td>C(2)</td>
<td>5389(2)</td>
<td>8026(2)</td>
<td>3418(1)</td>
<td>21(1)</td>
</tr>
<tr>
<td>C(3)</td>
<td>7327(2)</td>
<td>8833(2)</td>
<td>2700(1)</td>
<td>26(1)</td>
</tr>
<tr>
<td>C(4)</td>
<td>6183(2)</td>
<td>9809(2)</td>
<td>2810(1)</td>
<td>21(1)</td>
</tr>
<tr>
<td>C(5)</td>
<td>4658(2)</td>
<td>9757(1)</td>
<td>2528(1)</td>
<td>18(1)</td>
</tr>
<tr>
<td>C(6)</td>
<td>4531(2)</td>
<td>8771(2)</td>
<td>2020(1)</td>
<td>19(1)</td>
</tr>
<tr>
<td>C(7)</td>
<td>4489(2)</td>
<td>8463(2)</td>
<td>866(1)</td>
<td>24(1)</td>
</tr>
<tr>
<td>C(8)</td>
<td>5757(2)</td>
<td>7539(2)</td>
<td>835(1)</td>
<td>35(1)</td>
</tr>
<tr>
<td>C(9)</td>
<td>4572(2)</td>
<td>9389(2)</td>
<td>336(1)</td>
<td>39(1)</td>
</tr>
<tr>
<td>C(10)</td>
<td>3017(2)</td>
<td>7788(2)</td>
<td>851(1)</td>
<td>33(1)</td>
</tr>
<tr>
<td></td>
<td>Bond lengths [Å] and angles [°] for d13120.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O(1)-C(2)</td>
<td>1.220(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O(2)-C(6)</td>
<td>1.233(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(1)-C(2)</td>
<td>1.377(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(1)-C(4)</td>
<td>1.464(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(1)-C(3)</td>
<td>1.470(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(2)-C(11)</td>
<td>1.395(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(2)-C(5)</td>
<td>1.461(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(2)-C(1)</td>
<td>1.472(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(3)-C(6)</td>
<td>1.339(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(3)-C(7)</td>
<td>1.486(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(3)-H(3N)</td>
<td>0.91(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(1)-C(2)</td>
<td>1.526(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(1)-C(17)</td>
<td>1.534(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(1)-H(1A)</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(3)-C(4)</td>
<td>1.498(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(3)-H(3A)</td>
<td>0.9900</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(3)-H(3B)</td>
<td>0.9900</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(4)-C(5)</td>
<td>1.516(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(4)-H(4A)</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(5)-C(6)</td>
<td>1.533(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(5)-H(5A)</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(7)-C(9)</td>
<td>1.523(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(7)-C(10)</td>
<td>1.523(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(7)-C(8)</td>
<td>1.525(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(8)-H(8A)</td>
<td>0.9800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(8)-H(8B)</td>
<td>0.9800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(8)-H(8C)</td>
<td>0.9800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(9)-H(9A)</td>
<td>0.9800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(9)-H(9B)</td>
<td>0.9800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(9)-H(9C)</td>
<td>0.9800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(10)-H(10A)</td>
<td>0.9800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(10)-H(10B)</td>
<td>0.9800</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
C(10)-H(10C) 0.9800
C(11)-C(12) 1.401(2)
C(11)-C(16) 1.402(2)
C(12)-C(13) 1.386(3)
C(12)-H(12A) 0.9500
C(13)-C(14) 1.379(3)
C(13)-H(13A) 0.9500
C(14)-C(15) 1.383(3)
C(14)-H(14A) 0.9500
C(15)-C(16) 1.386(3)
C(15)-H(15A) 0.9500
C(16)-H(16A) 0.9500
C(17)-C(18) 1.515(2)
C(17)-H(17A) 0.9900
C(17)-H(17B) 0.9900
C(18)-C(19) 1.390(3)
C(18)-C(23) 1.393(3)
C(19)-C(20) 1.389(3)
C(19)-H(19A) 0.9500
C(20)-C(21) 1.378(3)
C(20)-H(20A) 0.9500
C(21)-C(22) 1.382(3)
C(21)-H(21A) 0.9500
C(22)-C(23) 1.393(3)
C(22)-H(22A) 0.9500
C(23)-H(23A) 0.9500
C(2)-N(1)-C(4) 118.18(14)
C(2)-N(1)-C(3) 123.19(14)
C(4)-N(1)-C(3) 61.42(11)
C(11)-N(2)-C(5) 119.40(13)
C(11)-N(2)-C(1) 120.33(14)
C(5)-N(2)-C(1) 119.82(14)
C(6)-N(3)-C(7) 125.80(14)
C(6)-N(3)-H(3N) 115.7(13)
C(7)-N(3)-H(3N) 118.2(13)
N(2)-C(1)-C(2) 111.27(13)
N(2)-C(1)-C(17) 112.62(14)
C(2)-C(1)-C(17) 110.63(13)
N(2)-C(1)-H(1A) 107.3
C(2)-C(1)-H(1A) 107.3
C(17)-C(1)-H(1A) 107.3
O(1)-C(2)-N(1) 122.49(17)
O(1)-C(2)-C(1) 122.13(16)
N(1)-C(2)-C(1) 114.93(14)
N(1)-C(3)-C(4) 59.09(11)
N(1)-C(3)-H(3A) 117.9
C(4)-C(3)-H(3A) 117.9
N(1)-C(3)-H(3B)    117.9
C(4)-C(3)-H(3B)    117.9
H(3A)-C(3)-H(3B)   115.0
N(1)-C(4)-C(3)     59.49(11)
N(1)-C(4)-C(5)     115.12(14)
C(3)-C(4)-C(5)     123.02(15)
N(1)-C(4)-H(4A)    115.7
C(3)-C(4)-H(4A)    115.7
C(5)-C(4)-H(4A)    115.7
N(2)-C(5)-C(4)     108.51(13)
N(2)-C(5)-C(6)     111.94(13)
C(4)-C(5)-C(6)     112.54(14)
N(2)-C(5)-H(5A)    107.9
C(4)-C(5)-H(5A)    107.9
C(6)-C(5)-H(5A)    107.9
O(2)-C(6)-N(3)     124.99(15)
O(2)-C(6)-C(5)     120.16(14)
N(3)-C(6)-C(5)     114.84(14)
N(3)-C(7)-C(9)     105.74(14)
N(3)-C(7)-C(10)    109.97(15)
C(9)-C(7)-C(10)    109.88(17)
N(3)-C(7)-C(8)     109.69(15)
C(9)-C(7)-C(8)     110.92(16)
C(10)-C(7)-C(8)    110.53(15)
C(7)-C(8)-H(8A)    109.5
C(7)-C(8)-H(8B)    109.5
H(8A)-C(8)-H(8B)   109.5
C(7)-C(8)-H(8C)    109.5
H(8A)-C(8)-H(8C)   109.5
H(8B)-C(8)-H(8C)   109.5
C(7)-C(9)-H(9A)    109.5
C(7)-C(9)-H(9B)    109.5
H(9A)-C(9)-H(9B)   109.5
C(7)-C(9)-H(9C)    109.5
H(9A)-C(9)-H(9C)   109.5
H(9B)-C(9)-H(9C)   109.5
C(7)-C(10)-H(10A)  109.5
C(7)-C(10)-H(10B)  109.5
H(10A)-C(10)-H(10B) 109.5
C(7)-C(10)-H(10C)  109.5
H(10A)-C(10)-H(10C) 109.5
H(10B)-C(10)-H(10C) 109.5
N(2)-C(11)-C(12)   120.86(15)
N(2)-C(11)-C(16)   121.64(15)
C(12)-C(11)-C(16)  117.50(16)
C(13)-C(12)-C(11)  120.73(16)
C(13)-C(12)-H(12A) 119.6
<table>
<thead>
<tr>
<th>Bond</th>
<th>Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(11)-C(12)-H(12A)</td>
<td>119.6</td>
</tr>
<tr>
<td>C(14)-C(13)-C(12)</td>
<td>121.34(17)</td>
</tr>
<tr>
<td>C(14)-C(13)-H(13A)</td>
<td>119.3</td>
</tr>
<tr>
<td>C(12)-C(13)-H(13A)</td>
<td>119.3</td>
</tr>
<tr>
<td>C(13)-C(14)-C(15)</td>
<td>118.38(17)</td>
</tr>
<tr>
<td>C(13)-C(14)-H(14A)</td>
<td>120.8</td>
</tr>
<tr>
<td>C(15)-C(14)-H(14A)</td>
<td>120.8</td>
</tr>
<tr>
<td>C(14)-C(15)-C(16)</td>
<td>121.24(18)</td>
</tr>
<tr>
<td>C(14)-C(15)-H(15A)</td>
<td>119.4</td>
</tr>
<tr>
<td>C(16)-C(15)-H(15A)</td>
<td>119.4</td>
</tr>
<tr>
<td>C(15)-C(16)-C(11)</td>
<td>120.69(17)</td>
</tr>
<tr>
<td>C(15)-C(16)-H(16A)</td>
<td>119.7</td>
</tr>
<tr>
<td>C(11)-C(16)-H(16A)</td>
<td>119.7</td>
</tr>
<tr>
<td>C(18)-C(17)-C(1)</td>
<td>117.91(15)</td>
</tr>
<tr>
<td>C(18)-C(17)-H(17A)</td>
<td>107.8</td>
</tr>
<tr>
<td>C(1)-C(17)-H(17A)</td>
<td>107.8</td>
</tr>
<tr>
<td>C(18)-C(17)-H(17B)</td>
<td>107.8</td>
</tr>
<tr>
<td>C(1)-C(17)-H(17B)</td>
<td>107.8</td>
</tr>
<tr>
<td>H(17A)-C(17)-H(17B)</td>
<td>107.2</td>
</tr>
<tr>
<td>C(19)-C(18)-C(23)</td>
<td>117.60(17)</td>
</tr>
<tr>
<td>C(19)-C(18)-C(17)</td>
<td>124.22(17)</td>
</tr>
<tr>
<td>C(23)-C(18)-C(17)</td>
<td>118.09(17)</td>
</tr>
<tr>
<td>C(20)-C(19)-C(18)</td>
<td>121.3(2)</td>
</tr>
<tr>
<td>C(20)-C(19)-H(19A)</td>
<td>119.3</td>
</tr>
<tr>
<td>C(18)-C(19)-H(19A)</td>
<td>119.3</td>
</tr>
<tr>
<td>C(21)-C(20)-C(19)</td>
<td>120.5(2)</td>
</tr>
<tr>
<td>C(21)-C(20)-H(20A)</td>
<td>119.7</td>
</tr>
<tr>
<td>C(19)-C(20)-H(20A)</td>
<td>119.7</td>
</tr>
<tr>
<td>C(20)-C(21)-C(22)</td>
<td>119.12(18)</td>
</tr>
<tr>
<td>C(20)-C(21)-H(21A)</td>
<td>120.4</td>
</tr>
<tr>
<td>C(22)-C(21)-H(21A)</td>
<td>120.4</td>
</tr>
<tr>
<td>C(21)-C(22)-C(23)</td>
<td>120.4(2)</td>
</tr>
<tr>
<td>C(21)-C(22)-H(22A)</td>
<td>119.8</td>
</tr>
<tr>
<td>C(23)-C(22)-H(22A)</td>
<td>119.8</td>
</tr>
<tr>
<td>C(22)-C(23)-C(18)</td>
<td>121.0(2)</td>
</tr>
<tr>
<td>C(22)-C(23)-H(23A)</td>
<td>119.5</td>
</tr>
<tr>
<td>C(18)-C(23)-H(23A)</td>
<td>119.5</td>
</tr>
</tbody>
</table>

Symmetry transformations used to generate equivalent atoms:
**Table 4.** Anisotropic displacement parameters $\left(Å^2 \times 10^3\right)$ for d13120. The anisotropic displacement factor exponent takes the form: $-2p^2 [h^2 a^*2U_{11} + ... + 2hk a^* b^* U_{12}]$

<table>
<thead>
<tr>
<th></th>
<th>$U_{11}$</th>
<th>$U_{22}$</th>
<th>$U_{33}$</th>
<th>$U_{23}$</th>
<th>$U_{13}$</th>
<th>$U_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>O(1)</td>
<td>40(1)</td>
<td>16(1)</td>
<td>22(1)</td>
<td>1(1)</td>
<td>-1(1)</td>
<td>2(1)</td>
</tr>
<tr>
<td>O(2)</td>
<td>44(1)</td>
<td>15(1)</td>
<td>22(1)</td>
<td>1(1)</td>
<td>2(1)</td>
<td>-4(1)</td>
</tr>
<tr>
<td>N(1)</td>
<td>24(1)</td>
<td>18(1)</td>
<td>23(1)</td>
<td>2(1)</td>
<td>-1(1)</td>
<td>2(1)</td>
</tr>
<tr>
<td>N(2)</td>
<td>25(1)</td>
<td>20(1)</td>
<td>17(1)</td>
<td>4(1)</td>
<td>2(1)</td>
<td>3(1)</td>
</tr>
<tr>
<td>N(3)</td>
<td>31(1)</td>
<td>15(1)</td>
<td>17(1)</td>
<td>0(1)</td>
<td>2(1)</td>
<td>-2(1)</td>
</tr>
<tr>
<td>C(1)</td>
<td>28(1)</td>
<td>18(1)</td>
<td>13(1)</td>
<td>1(1)</td>
<td>0(1)</td>
<td>-1(1)</td>
</tr>
<tr>
<td>C(2)</td>
<td>32(1)</td>
<td>16(1)</td>
<td>12(1)</td>
<td>-2(1)</td>
<td>-4(1)</td>
<td>-2(1)</td>
</tr>
<tr>
<td>C(3)</td>
<td>27(1)</td>
<td>21(1)</td>
<td>28(1)</td>
<td>2(1)</td>
<td>3(1)</td>
<td>3(1)</td>
</tr>
<tr>
<td>C(4)</td>
<td>26(1)</td>
<td>14(1)</td>
<td>21(1)</td>
<td>2(1)</td>
<td>1(1)</td>
<td>0(1)</td>
</tr>
<tr>
<td>C(5)</td>
<td>24(1)</td>
<td>14(1)</td>
<td>17(1)</td>
<td>1(1)</td>
<td>2(1)</td>
<td>0(1)</td>
</tr>
<tr>
<td>C(6)</td>
<td>22(1)</td>
<td>16(1)</td>
<td>18(1)</td>
<td>0(1)</td>
<td>0(1)</td>
<td>-1(1)</td>
</tr>
<tr>
<td>C(7)</td>
<td>33(1)</td>
<td>23(1)</td>
<td>17(1)</td>
<td>-3(1)</td>
<td>2(1)</td>
<td>-5(1)</td>
</tr>
<tr>
<td>C(8)</td>
<td>40(1)</td>
<td>33(1)</td>
<td>33(1)</td>
<td>-10(1)</td>
<td>9(1)</td>
<td>1(1)</td>
</tr>
<tr>
<td>C(9)</td>
<td>62(1)</td>
<td>38(1)</td>
<td>17(1)</td>
<td>2(1)</td>
<td>2(1)</td>
<td>-8(1)</td>
</tr>
<tr>
<td>C(10)</td>
<td>36(1)</td>
<td>38(1)</td>
<td>27(1)</td>
<td>-4(1)</td>
<td>-3(1)</td>
<td>-8(1)</td>
</tr>
<tr>
<td>C(11)</td>
<td>20(1)</td>
<td>18(1)</td>
<td>22(1)</td>
<td>-2(1)</td>
<td>-1(1)</td>
<td>-2(1)</td>
</tr>
<tr>
<td>C(12)</td>
<td>24(1)</td>
<td>23(1)</td>
<td>22(1)</td>
<td>1(1)</td>
<td>0(1)</td>
<td>-3(1)</td>
</tr>
<tr>
<td>C(13)</td>
<td>23(1)</td>
<td>26(1)</td>
<td>31(1)</td>
<td>5(1)</td>
<td>-6(1)</td>
<td>-3(1)</td>
</tr>
<tr>
<td>C(14)</td>
<td>20(1)</td>
<td>36(1)</td>
<td>40(1)</td>
<td>-3(1)</td>
<td>-2(1)</td>
<td>4(1)</td>
</tr>
<tr>
<td>C(15)</td>
<td>25(1)</td>
<td>53(1)</td>
<td>29(1)</td>
<td>-6(1)</td>
<td>4(1)</td>
<td>7(1)</td>
</tr>
<tr>
<td>C(16)</td>
<td>26(1)</td>
<td>41(1)</td>
<td>21(1)</td>
<td>0(1)</td>
<td>-2(1)</td>
<td>3(1)</td>
</tr>
<tr>
<td>C(17)</td>
<td>31(1)</td>
<td>25(1)</td>
<td>23(1)</td>
<td>3(1)</td>
<td>-2(1)</td>
<td>-6(1)</td>
</tr>
<tr>
<td>C(18)</td>
<td>26(1)</td>
<td>23(1)</td>
<td>28(1)</td>
<td>5(1)</td>
<td>4(1)</td>
<td>1(1)</td>
</tr>
<tr>
<td>C(19)</td>
<td>54(1)</td>
<td>29(1)</td>
<td>28(1)</td>
<td>0(1)</td>
<td>12(1)</td>
<td>-6(1)</td>
</tr>
<tr>
<td>C(20)</td>
<td>65(2)</td>
<td>38(1)</td>
<td>26(1)</td>
<td>1(1)</td>
<td>16(1)</td>
<td>1(1)</td>
</tr>
<tr>
<td>C(21)</td>
<td>48(1)</td>
<td>33(1)</td>
<td>38(1)</td>
<td>13(1)</td>
<td>21(1)</td>
<td>5(1)</td>
</tr>
<tr>
<td>C(22)</td>
<td>43(1)</td>
<td>31(1)</td>
<td>50(1)</td>
<td>15(1)</td>
<td>6(1)</td>
<td>-7(1)</td>
</tr>
<tr>
<td>C(23)</td>
<td>38(1)</td>
<td>29(1)</td>
<td>38(1)</td>
<td>7(1)</td>
<td>-2(1)</td>
<td>-7(1)</td>
</tr>
</tbody>
</table>

**Table 5.** Hydrogen coordinates ($x \times 10^4$) and isotropic displacement parameters ($Å^2 \times 10^3$) for d13120.

<table>
<thead>
<tr>
<th></th>
<th>$x$</th>
<th>$y$</th>
<th>$z$</th>
<th>$U$(eq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H(1A)</td>
<td>3816</td>
<td>8992</td>
<td>3905</td>
<td>24</td>
</tr>
<tr>
<td>H(3A)</td>
<td>7099</td>
<td>8167</td>
<td>2399</td>
<td>31</td>
</tr>
<tr>
<td>H(3B)</td>
<td>8373</td>
<td>9087</td>
<td>2715</td>
<td>31</td>
</tr>
<tr>
<td>H(4A)</td>
<td>6566</td>
<td>10659</td>
<td>2903</td>
<td>25</td>
</tr>
<tr>
<td>H(5A)</td>
<td>4448</td>
<td>10584</td>
<td>2337</td>
<td>22</td>
</tr>
<tr>
<td>---</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>6694</td>
<td>7988</td>
<td>844</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>2948</td>
<td>7227</td>
<td>1206</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>1742</td>
<td>7950</td>
<td>3347</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>4650(20)</td>
<td>10050(20)</td>
<td>1406(9)</td>
<td>30(5)</td>
</tr>
</tbody>
</table>
Table 6. Torsion angles [°] for d13120.

<table>
<thead>
<tr>
<th>Bond</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(11)-N(2)-C(1)-C(2)</td>
<td>179.72(14)</td>
</tr>
<tr>
<td>C(5)-N(2)-C(1)-C(2)</td>
<td>7.5(2)</td>
</tr>
<tr>
<td>C(11)-N(2)-C(1)-C(17)</td>
<td>54.9(2)</td>
</tr>
<tr>
<td>C(5)-N(2)-C(1)-C(17)</td>
<td>-117.38(17)</td>
</tr>
<tr>
<td>C(4)-N(1)-C(2)-O(1)</td>
<td>143.58(15)</td>
</tr>
<tr>
<td>C(3)-N(1)-C(2)-O(1)</td>
<td>70.9(2)</td>
</tr>
<tr>
<td>C(4)-N(1)-C(2)-C(1)</td>
<td>-43.97(19)</td>
</tr>
<tr>
<td>C(3)-N(1)-C(2)-C(1)</td>
<td>-116.64(16)</td>
</tr>
<tr>
<td>N(2)-C(4)-C(2)-O(1)</td>
<td>-147.67(15)</td>
</tr>
<tr>
<td>C(17)-C(1)-C(2)-O(1)</td>
<td>-21.7(2)</td>
</tr>
<tr>
<td>N(2)-C(1)-C(2)-N(1)</td>
<td>39.85(18)</td>
</tr>
<tr>
<td>C(17)-C(1)-C(2)-N(1)</td>
<td>165.83(13)</td>
</tr>
<tr>
<td>C(2)-N(1)-C(3)-C(4)</td>
<td>106.64(18)</td>
</tr>
<tr>
<td>C(2)-N(1)-C(3)-C(3)</td>
<td>-114.54(16)</td>
</tr>
<tr>
<td>C(2)-N(1)-C(4)-C(5)</td>
<td>0.4(2)</td>
</tr>
<tr>
<td>C(3)-N(1)-C(4)-C(5)</td>
<td>114.98(16)</td>
</tr>
<tr>
<td>N(1)-C(3)-C(4)-C(5)</td>
<td>-101.81(17)</td>
</tr>
<tr>
<td>C(11)-N(2)-C(5)-C(4)</td>
<td>140.37(15)</td>
</tr>
<tr>
<td>C(1)-N(2)-C(5)-C(4)</td>
<td>-47.32(19)</td>
</tr>
<tr>
<td>C(11)-N(2)-C(5)-C(6)</td>
<td>-94.83(17)</td>
</tr>
<tr>
<td>C(1)-N(2)-C(5)-C(6)</td>
<td>77.47(18)</td>
</tr>
<tr>
<td>N(1)-C(4)-C(5)-N(2)</td>
<td>43.59(18)</td>
</tr>
<tr>
<td>C(3)-C(4)-C(5)-N(2)</td>
<td>112.44(16)</td>
</tr>
<tr>
<td>N(1)-C(4)-C(5)-C(6)</td>
<td>-80.85(17)</td>
</tr>
<tr>
<td>C(3)-C(4)-C(5)-C(6)</td>
<td>-12.2(2)</td>
</tr>
<tr>
<td>C(7)-N(3)-C(6)-O(2)</td>
<td>2.0(3)</td>
</tr>
<tr>
<td>C(7)-N(3)-C(6)-C(5)</td>
<td>-178.85(16)</td>
</tr>
<tr>
<td>N(2)-C(5)-C(6)-O(2)</td>
<td>-43.9(2)</td>
</tr>
<tr>
<td>C(4)-C(5)-C(6)-O(2)</td>
<td>78.6(2)</td>
</tr>
<tr>
<td>N(2)-C(5)-C(6)-N(3)</td>
<td>136.88(15)</td>
</tr>
<tr>
<td>C(4)-C(5)-C(6)-N(3)</td>
<td>-100.59(17)</td>
</tr>
<tr>
<td>C(6)-N(3)-C(7)-C(9)</td>
<td>177.42(19)</td>
</tr>
<tr>
<td>C(6)-N(3)-C(7)-C(10)</td>
<td>58.8(2)</td>
</tr>
<tr>
<td>C(6)-N(3)-C(7)-C(8)</td>
<td>-62.9(2)</td>
</tr>
<tr>
<td>C(5)-N(2)-C(11)-C(12)</td>
<td>18.7(2)</td>
</tr>
<tr>
<td>C(1)-N(2)-C(11)-C(12)</td>
<td>-153.59(16)</td>
</tr>
<tr>
<td>C(5)-N(2)-C(11)-C(16)</td>
<td>-160.47(16)</td>
</tr>
<tr>
<td>C(1)-N(2)-C(11)-C(16)</td>
<td>27.3(2)</td>
</tr>
<tr>
<td>N(2)-C(11)-C(12)-C(13)</td>
<td>-175.18(15)</td>
</tr>
<tr>
<td>C(16)-C(11)-C(12)-C(13)</td>
<td>4.0(3)</td>
</tr>
<tr>
<td>C(11)-C(12)-C(13)-C(14)</td>
<td>-2.0(3)</td>
</tr>
<tr>
<td>C(12)-C(13)-C(14)-C(15)</td>
<td>-1.1(3)</td>
</tr>
<tr>
<td>C(13)-C(14)-C(15)-C(16)</td>
<td>2.2(3)</td>
</tr>
<tr>
<td>C(14)-C(15)-C(16)-C(11)</td>
<td>-0.1(3)</td>
</tr>
<tr>
<td>N(2)-C(11)-C(16)-C(15)</td>
<td>176.20(17)</td>
</tr>
</tbody>
</table>
C(12)-C(11)-C(16)-C(15) -3.0(3)
N(2)-C(1)-C(17)-C(18) -153.02(16)
C(2)-C(1)-C(17)-C(18) 81.76(19)
C(1)-C(17)-C(18)-C(19) 20.9(3)
C(1)-C(17)-C(18)-C(23) -162.58(17)
C(23)-C(18)-C(19)-C(20) -0.8(3)
C(17)-C(18)-C(19)-C(20) 175.8(2)
C(18)-C(19)-C(20)-C(21) -1.0(4)
C(19)-C(20)-C(21)-C(22) 1.6(4)
C(20)-C(21)-C(22)-C(23) -0.4(3)
C(21)-C(22)-C(23)-C(18) -1.4(3)
C(19)-C(18)-C(23)-C(22) 1.9(3)
C(17)-C(18)-C(23)-C(22) -174.8(2)

Symmetry transformations used to generate equivalent atoms:

Table 7. Hydrogen bonds for d13120 [Å and °].

<table>
<thead>
<tr>
<th>D-H...A</th>
<th>d(D-H)</th>
<th>d(H...A)</th>
<th>d(D...A)</th>
<th>&lt;(DHA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N(3)-H(3N)...O(1)#1</td>
<td>0.91(2)</td>
<td>2.06(2)</td>
<td>2.951(2)</td>
<td>167.4(19)</td>
</tr>
</tbody>
</table>

Symmetry transformations used to generate equivalent atoms:
#1 -x+1,y+1/2,-z+1/2
Appendix 2. Chapter 2 NMR Spectra and LCMS Data
Appendix 3. Chapter 3 NMR Spectra
20130923_mercury_400_RC-02-13-PROTON_01

20130923_mercury_400_RC-02-13-CARBON_01
Appendix 4. Red Wings Screen

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>