

# TARGETING THE MENIN-MLL-LEDGF INTERACTION WITH SMALL MOLECULE INHIBITORS

by

Jonathan Pollock

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Molecular and Cellular Pathology)  
in the University of Michigan  
2016

**Doctoral Committee:**

Assistant Professor Tomasz Cierpicki, Co-Chair  
Assistant Professor Jolanta Grembecka, Co-Chair  
Professor Anna K. Mapp  
Assistant Professor Andrew G. Muntean  
Associate Professor Zaneta Nikolovska-Coleska

© Jonathan Pollock

---

All rights reserved, 2016

## **Dedication**

For Melissa

## **Acknowledgements**

I would first like to thank my mentors Jolanta Grembecka and Tomasz Cierpicki for their guidance throughout my graduate school experience. The former and current members of the TCJG lab have provided a supportive community both experimentally and emotionally. I would like to especially thank the former lab members George Lund, Felicia Gray, David Rogawski and Dmitry Borkin for their friendship and shared experiences throughout this process. My thesis committee has helped me by providing instrumental feedback and support.

I would also like to thank my friends and family that motivated me and supported my efforts to embark on this path. My friends in Ann Arbor and in Chicago have provided a way to relieve stress, stay connected and I have had many enjoyable moments for which I am grateful. My parents Patricia and Ralph Pollock have always supported my decisions and provided a warm loving home environment which allowed me to get reach my goals, I am very lucky to have them. I would like to thank my brothers, Michael, Ben, and Kyle for their encouragement. I would also like to thank my chosen family, the Meyers, whom have always supported and loved me as one of their own. I am very lucky to have you all in my life.

Finally I would like to thank the most important person in my life, Melissa. You make me a better person every day. You have always built me up and seen my potential when I didn't know it was there. Your love and support through my graduate career is the only way I could have achieved this and I am so very thankful for you in my life. I love you and thank you for everything.

## Table of Contents

Dedication .....	ii
Acknowledgements .....	iii
List of Figures.....	vii
List of Tables .....	ix
List of Equations .....	x
List of Schematics .....	xi
List of Appendices.....	xii
List of Abbreviations .....	xiii
Abstract.....	xv
<b>Chapter 1: Introduction .....</b>	<b>1</b>
<b>A. Background.....</b>	<b>1</b>
<i>A.1 Wild-type MLL.....</i>	<i>1</i>
<i>A.2 MLL translocations in acute leukemia .....</i>	<i>2</i>
<i>A.3 MLL translocations .....</i>	<i>2</i>
<i>A.4 WT MLL in leukemogenesis.....</i>	<i>3</i>
<i>A.5 HOX genes in MLL rearranged leukemia .....</i>	<i>3</i>
<b>B. The menin-MLL-LEDGF ternary complex.....</b>	<b>5</b>
<i>B.1 Importance of the interaction between menin-MLL-LEDGF for leukemogenesis .....</i>	<i>5</i>
<i>B.2 Structural studies on menin.....</i>	<i>7</i>
<i>B.3 Structure characterization of the menin-MLL interaction .....</i>	<i>7</i>
<b>C. Targeting the menin-MLL interaction with small molecules .....</b>	<b>9</b>
<i>C.1 Identifying small molecule inhibitors of the menin-MLL interaction .....</i>	<i>9</i>
<i>C.2 Targeting the bivalent interaction.....</i>	<i>13</i>
<i>C.3 Potential applications of menin inhibitors .....</i>	<i>14</i>
<b>Chapter 2: Rational Design of Orthogonal Multipolar Interactions with Fluorine in Protein–Ligand Complexes.....</b>	<b>16</b>
<b>A. Abstract.....</b>	<b>16</b>
<b>B. Background.....</b>	<b>17</b>
<b>C. Results and Discussion.....</b>	<b>18</b>

C.1 Trifluoromethyl Groups in Menin–MLL Inhibitors Form Close Contacts with Protein Backbone.....	18
C.2 Development of FMAP Algorithm To Predict Multipolar C–F...C=O Interactions.....	19
C.3 Interactions of Trifluoromethyl–Thiadiazole Moiety with Menin.....	21
C.4. Interactions of Trifluoroethyl Group in Thienopyrimidine Core with Menin.....	24
<b>D. Conclusions.....</b>	<b>25</b>
<b>Chapter 3: Development of inhibitors targeting the menin-MLL interaction.....</b>	<b>28</b>
<b>A. Abstract.....</b>	<b>28</b>
<b>B. Background.....</b>	<b>29</b>
<b>C. Results and Discussion.....</b>	<b>30</b>
C.1 Modifications of MI-2-2 to explore the menin binding pocket.....	30
C.2 Investigation of indole nitrogen substituents on MI-136 scaffold.....	34
C.3. Exploring modifications at positions 3, 4, and 6 of the of indole ring in MI-136 scaffold.....	40
C.4 Generating analogs through combinations of optimal substituents on the indole ring in MI-136 scaffold.....	44
C.5 Pharmacokinetic analysis of menin-MLL inhibitors in a murine model.....	50
C.6 MI-503 treatment reduces MLL leukemia burden and improves survival in murine model.....	51
<b>D. Conclusion.....</b>	<b>52</b>
<b>Chapter 4: Design and characterization of the hydroxy- and aminomethylpiperidine class of menin-MLL inhibitors.....</b>	<b>54</b>
<b>A. Abstract.....</b>	<b>54</b>
<b>B. Background.....</b>	<b>55</b>
B.1 Identification of a chemically distinct class of menin-MLL inhibitors.....	55
<b>C. Results and Discussion.....</b>	<b>56</b>
C.1 HTS campaign and identification of the minimal pharmacophore.....	56
C.2 Structure based-design of MIV analogs.....	59
C.3 Hydroxymethylpiperidine Inhibitors Mimic the Most Critical Interactions of MLL with Menin.....	60
C.4 Design of orthogonal multipolar interaction of MIV analog with the menin backbone.....	63
C.5 Increasing the polarity of the MIV class of menin-MLL inhibitors.....	64
C.6 Identification of optimal substituents at the tail region.....	65
C.7 Exploring the hydrogen bond network in the head region.....	66
<b>D. Conclusions.....</b>	<b>67</b>
<b>Chapter 5: Development of covalent inhibitors of the menin-MLL interaction.....</b>	<b>71</b>
<b>A. Abstract.....</b>	<b>71</b>
<b>B. Background.....</b>	<b>72</b>
B.1 Advantages and disadvantages of covalent inhibitors.....	72
B.2 Design of covalent inhibitors of PPIs.....	72

<i>B.3 Biochemical characterization of covalent inhibitors .....</i>	73
<b>C. Results .....</b>	<b>75</b>
<i>C.1 Investigation of accessible cysteine residues near the small molecule binding pocket of menin.....</i>	75
<i>C.2 Structure based-design of covalent thienopyrimidine class of menin-MLL inhibitors .....</i>	77
<i>C.3 Characterizing the potency and rate of menin engagement of the covalent inhibitors. ....</i>	82
<i>C.4. Structure-based design and characterization of acrylamide covalent inhibitors .....</i>	85
<i>C.5 Investigating the reactivity of the covalent menin-MLL inhibitors with glutathione. ....</i>	88
<i>C.6 Activity of covalent inhibitors of the menin-MLL interaction in leukemic cells. ....</i>	89
<i>C.7 Detection of menin engagement in MLL leukemic cells and bone marrow in vivo. ..</i>	90
<b>D. Conclusion.....</b>	<b>94</b>
<b>Chapter 6: Conclusions and Future Directions.....</b>	<b>96</b>
<b>A. Conclusions .....</b>	<b>96</b>
<i>A.1 Fluorine-backbone interactions contribute to high affinity interaction with menin..</i>	97
<i>A.2 Optimization of small molecule inhibitors of the menin-MLL interaction to demonstrate in vivo efficacy.....</i>	98
<b>B. Future Directions.....</b>	<b>101</b>
<i>B.1 Future development of the menin-MLL inhibitors.....</i>	101
<i>B.2 Investigation of menin-MLL inhibitors in other malignancies.....</i>	101
<i>B.3 Structural investigation of the IBD interaction with MLL identifies new interaction motif and potential therapeutic target .....</i>	101
<b>Materials and Methods.....</b>	<b>103</b>
<b>Appendices.....</b>	<b>111</b>
<b>References .....</b>	<b>127</b>

## List of Figures

Figure 1.1 Crystal structure of the menin-MLL-LEDGF complex.....	6
Figure 1.2 Characterization of the menin-MLL interaction.....	10
Figure 1.3 Thienopyrimidine class of menin-MLL inhibitors.....	11
Figure 1.4 Reported inhibitors of the menin-MLL interaction.....	13
Figure 2.1 Inhibitors of the menin-MLL interaction containing CF <sub>3</sub> groups.....	19
Figure 2.2 Prediction of favorable C-F···C=O interactions using FMAP algorithm. ....	21
Figure 2.3 Effect of fluorine substitutions in thiadiazole moiety on activity of menin-MLL inhibitors. ....	23
Figure 2.4 Effect of fluorine substitutions in thienopyrimidine moiety on activity of menin-MLL inhibitors. ....	25
Figure 3.1 Crystal structure of MI-2-2 bound to menin.....	30
Figure 3.2 Crystal structure of MI-136 bound to menin.....	34
Figure 3.3 Structure-based design of indole nitrogen substituents. ....	36
Figure 3.4 Structure-based design of indole nitrogen substituents. ....	37
Figure 3.5 Structure-based design of substituents at position 4 of the indole. ....	40
Figure 3.6 Structure-based design of indole position 6 modifications in MI-136.....	42
Figure 3.7 Exploring combinations of substituent at positions 1 and 4 of indole. ....	46
Figure 3.8 Analysis of substituent combinations at positions 1 and 6 of indole. ....	47
Figure 3.9 Exploring combinations of substituent at positions 1, 4 and 6 of indole. ....	48
Figure 3.10 Treatment with MI-503 reduces tumor burden and demonstrates survival benefit..	51
Figure 4.1 Identification of the most potent HTS hit.....	57
Figure 4.2 Determining binding affinity of each MIV-3 enantiomer. ....	58
Figure 4.3 MIV-3 enantiomers bind to menin in the MLL binding site.....	60
Figure 4.4 MIV-3R binding mode to menin. ....	62
Figure 4.5 Crystal structure of MIV-4 bound to menin.....	64
Figure 4.6 Crystal structure of MIV-5 bound to menin.....	65



Figure 4.7 Crystal structure of MIV-5 and MIV-7 bound to menin. ....	66
Figure 4.8 MIV-3R and MIV-6R binding modes to menin are investigated.....	67
Figure 5.1 Crystal structure of MI-503 bound to menin.....	76
Figure 5.2 Design of BD615 with an acrylamide reactive moiety. ....	78
Figure 5.3 Design of irreversible inhibitor BD712 with a propynamide reactive moiety. ....	79
Figure 5.4 Design of irreversible inhibitor BD688 with a vinylsulfonamide reactive moiety. ....	81
Figure 5.5 Characterizing irreversible inhibitor rates of menin engagement. ....	84
Figure 5.6 Structure-based design of acrylamide covalent inhibitors.....	86
Figure 5.7 Structure-based design of aromatic substituted acrylamides with menin. ....	87
Figure 5.8 Structure-based optimization of polar interactions of aromatic substituted acrylamides with menin. ....	87
Figure 5.9 Menin engagement in cells.....	92
Figure 5.10 Determination of the rate of menin turnover with washout experiments.....	93
Figure 5.11 CETSA analysis to assess menin engagement in the bone marrow <i>in vivo</i> . ....	94
Figure 6.1 Development of menin-MLL inhibitors.....	100
Figure A.1 Structure determination and biophysical characterization of the MLL-IBD interaction. ....	113
Figure A.2 The IBD interface with MLL and IN overlaps.....	116
Figure B.1 Identification and characterization of fragment screening hits.....	121
Figure B.2 Solution structure of IBD with chemical shift perturbations mapped at 2mM 3CC7. ....	122

## List of Tables

Table 3.1 SAR and activities of analogues of the thienopyrimidine class of menin-MLL inhibitors. ....	32
Table 3.2 SAR and properties of analogues with substitutions at position 2 of indole. ....	33
Table 3.3 SAR and properties of analogues with substitutions at position 1 of indole. ....	38
Table 3.4 SAR and properties of analogues with substitutions at position 1 of indole. ....	39
Table 3.5 SAR and properties of analogues with substitutions at position 3, 4, or 6 of indole....	43
Table 3.6 SAR and properties of analogues with substitutions at positions 3, 4, or 6 of indole. .	45
Table 3.7 SAR and properties of analogues with substitutions at positions 3, 4, or 6 of indole. .	49
Table 3.8 Pharmacokinetic profile of menin-MLL inhibitors in blood plasma of murine models. ....	50
Table 4.1 SAR and activity of MIV analogues. ....	57
Table 4.2 SAR and activity of MIV analogues. ....	69
Table 5.1 SAR and activities of covalent analogues of the thienopyrimidine class of menin-MLL inhibitors. ....	77
Table 5.2 SAR and properties of analogues. ....	83
Table 5.3 SAR and properties of analogues. ....	86
Table 5.4 Reactivity of inhibitors with glutathione. ....	88
Table 5.5 SAR and properties of analogues. ....	90
Table B.1 SAR and 3PA values for 3CC7. ....	124

## List of Equations

Equation 5.1 Equilibrium equation .....	84
Equation 5.2 Pseudo-first order rate equation and half-life derivatization. ....	84
Equation B.1 Calculation of chemical shift perturbations .....	106
Equation B.2 Calculation of $K_D$ for ligand titration .....	106

## List of Schematics

Schematic 1.1 Menin-MLL-LEDGF ternary complex assembly. ....	4
Schematic 5.1 Design of the menin pulldown assay with MLL43 coated beads. ....	91

## List of Appendices

<b>Appendix A: The same site on IBD of LEDGF binds to both MLL and HIV-IN.....</b>	<b>111</b>
<b>A. Abstract.....</b>	<b>111</b>
<b>B. Background.....</b>	<b>112</b>
<b>C. Results and Discussion.....</b>	<b>113</b>
<i>C.1 Structural and biochemical analysis of the menin-MLL interaction with IBD. ....</i>	<i>114</i>
<i>C.2 HIV-IN binding to IBD overlaps with the MLL binding interface. ....</i>	<i>114</i>
<b>D. Conclusions .....</b>	<b>117</b>
<b>Appendix B: Fragment screening of IBD-LEDGF by NMR.....</b>	<b>118</b>
<b>A. Abstract.....</b>	<b>118</b>
<b>B. Background.....</b>	<b>119</b>
<b>C. Results and Discussion.....</b>	<b>120</b>
<i>C.1 Identification and characterization of small molecule fragments that bind to IBD. 120</i>	
<i>C.2 Development of the structure activity relationship (SAR) for 3CC7.....</i>	<i>122</i>
<b>D. Conclusion.....</b>	<b>125</b>
<b>Appendix C: Table of Crystallography Statistics .....</b>	<b>126</b>

## List of Abbreviations

<b>AlphaLISA</b>	Amplified Luminescence Proximity Homogenous Assay
<b>AML</b>	Acute myeloid leukemia
<b>ALL</b>	Acute lymphoblastic leukemia
<b>BMC</b>	Bone marrow cells
<b>BME</b>	$\beta$ -mercaptoethanol
<b>CETSA</b>	Cellular thermoshift assay
<b>D<sub>2</sub>O</b>	Deuterium oxide
<b>DMSO</b>	Dimethyl sulfoxide
<b>DTT</b>	Dithiothreitol
<b>FBDD</b>	Fragment based drug design
<b>FITC</b>	Fluorescein
<b>FP</b>	Fluorescence polarization
<b>FRET</b>	Förster resonance energy transfer
<b>GI<sub>50</sub></b>	Half-maximal growth inhibition
<b>GSH</b>	Glutathione
<b>H3K4</b>	Histone 3 lysine 4
<b>H3K4me3</b>	Histone 3 lysine 4 tri-methylation
<b>HM-2</b>	<i>Hoxa9</i> and <i>Meis1</i> transformed murine BMCs
<b>HTRF</b>	Homogeneous Time Resolved Fluorescence
<b>HTS</b>	High-throughput screen
<b>IBD</b>	Integrase binding domain
<b>IBM1</b>	IBD binding motif 1
<b>IBM2</b>	IBD binding motif 2
<b>IC<sub>50</sub></b>	Half-maximal inhibitory concentration
<b>IPTG</b>	Isopropyl $\beta$ -D-1-thiogalactopyranoside
<b>ITC</b>	Isothermal titration calorimetry
<b>LC-MS</b>	Liquid chromatography couple with mass spectrometry
<b>LEDGF</b>	Lens epithelial-derived growth factor
<b>MBM1</b>	Menin binding motif 1
<b>MBM2</b>	Menin binding motif 2
<b>MLL</b>	Mixed lineage leukemia 1
<b>Ni-NTA</b>	Nickle nitrilotriacetic acid
<b>NMR</b>	Nuclear Magnetic Resonance

<b>PDB</b>	Protein Data Bank
<b>PPI</b>	Protein-protein interaction
<b>SAR</b>	Structure Activity Relationship
<b>TCEP</b>	tris(2-carboxyethyl)phosphine
<b>TPR</b>	Tetratricopeptide repeat
<b>WT</b>	Wild type

## Abstract

Mixed Lineage Leukemia 1 (MLL) is a large multi-domain protein (430 kDa) encoded by the *MLL* gene that catalyzes the methylation of histone H3 lysine 4 (H3K4). Chromosomal translocations of the *Mixed Lineage Leukemia (MLL)* gene result in acute myeloid and lymphoblastic leukemias. MLL mediated leukemia is present in about 10% of adult acute leukemia and up to 70% of infant leukemias. Patients with MLL leukemia do not respond well to conventional treatment methods; therefore, have very poor prognosis with only about a 35% overall 5-year survival rate and high risk of relapse. This demonstrates the urgent need for novel targeted therapeutics to treat these leukemias.

Chromosomal translocation of the *MLL* gene with one of over 60 partner genes generates an oncogenic chimeric fusion protein. MLL fusion proteins lead to enhanced cell proliferation, up-regulation of *Hoxa9* and *Meis1* genes and block hematopoietic differentiation, ultimately leading to acute leukemia. The oncogenic function of MLL fusion proteins is reliant on the interaction with menin and with Lens Epithelium Derived Growth Factor (LEDGF). Menin functions as a scaffolding protein and interacts with wild type (WT) MLL or MLL fusion proteins which are localized to target genes through association with LEDGF. Formation of the menin-MLL-LEDGF ternary complex is critical for MLL associated gene regulation and leukemogenic transformation. Therefore, inhibition of the menin-MLL interaction should abrogate the development and progression of MLL leukemia.

Our laboratory was the first to report small molecule inhibitors of the menin-MLL interaction. We have identified two classes of reversible menin-MLL inhibitors and rationally designed a new class of covalent menin-MLL inhibitors. Using high resolution crystal structures



of menin in complex with inhibitors we applied structure-based design to further optimize inhibitors of the menin-MLL interaction. We performed extensive characterization and optimization of potency, solubility and pharmacokinetic properties of these compounds. These efforts led to the development of compounds suitable for *in vivo* analysis and demonstrated significant survival benefit in murine models of MLL leukemias. Overall, we have demonstrated that menin-MLL inhibitors may offer a novel therapeutic strategy for MLL leukemia patients and have developed a promising lead scaffold for clinical optimization.

# Chapter 1: Introduction

## A. Background

### A.1 Wild-type MLL

Mixed Lineage Leukemia 1 (MLL1, referred to as MLL and KMT2A) is a large multi-domain protein (430 kDa) encoded by the *MLL* gene that catalyzes the methylation of histone H3 lysine 4 (H3K4).<sup>2-4</sup> MLL is a human homolog of the trithorax (Trx) group proteins, which is required for body patterning in *drosophila*,<sup>5</sup> and is required for normal mouse development,<sup>6,7</sup> Importantly MLL expression is necessary for maintained expression of Homeotic (*Hox*) genes and plays a role in normal hematopoiesis.<sup>4</sup> MLL is post-translationally cleaved by the protease Taspase1 into two protein fragments: MLL<sup>C</sup> and MLL<sup>N</sup>.<sup>8</sup> These fragments re-associate non-covalently and migrate to the nucleus where they perform transcriptional regulatory functions.<sup>8</sup> Once in the nucleus, the wild type (WT) MLL is then recruited to target genes via the N-terminal domain of MLL by interactions with menin, lens epithelium derived growth factor LEDGF, and c-Myb, and through the CXXC domain, which leads to transcriptional activation.<sup>9, 10, 11, 12</sup> The H3K4me3 is catalyzed by the MLL C-terminal SET domain and association with essential cofactors Wdr5, Ash2L, and Rbbp5 is needed for catalytic activity resulting in gene activation.<sup>13, 14, 15</sup> H3K4me3 is an epigenetic conserved mark for transcriptionally active chromatin.<sup>16,17</sup> H3K4me3 mark is enriched near the transcription start sites of transcriptionally active genes.<sup>17,18</sup> *Hox* genes are considered important downstream targets of MLL and are tightly controlled through epigenetic modifications.<sup>2,19, 20</sup> *Hoxa9* and *Meis1* are essential for hematopoietic stem cells and progenitor cells self-renewal,<sup>21,22</sup> however *Hoxa9*<sup>-/-</sup> has only a minor hematopoietic defect.<sup>23</sup> *Hoxa9* and *Meis1* normally become down-regulated through the progression of

hematopoietic progenitor cells to a fully differentiated state.<sup>11,22</sup> Failure to repress *Hox* expression leads to the induction of self-renewal preventing differentiation and ultimately results in acute leukemia.<sup>24</sup>

## **A.2 MLL translocations in acute leukemia**

Chromosomal translocations that result in a chimeric fusion gene represent a common mechanism in several types of malignancies especially in acute leukemias.<sup>25</sup> Mixed Lineage Leukemia (MLL) is localized at chromosome band 11q23, and is a common target for chromosomal rearrangements. MLL-rearrangements found in patients with acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL) represent 10% of *de novo* cases<sup>26,27</sup> and 30% of therapy related acute leukemias.<sup>28</sup> MLL mediated leukemia is present in about 10% of adult acute leukemia<sup>29</sup> and about 70% of acute lymphoid leukemia (ALL) and 35-50% of acute myeloid leukemia (AML) in infants.<sup>30</sup> MLL leukemias have relatively few cooperating mutations potentially making them more amenable to therapeutics targeting the fusion protein.<sup>31</sup> Patients with MLL leukemia do not respond well to conventional treatment methods; therefore, have very poor prognosis with only about a 35% overall 5-year survival rate<sup>32</sup> and high risk of relapse.<sup>24</sup> This demonstrates the urgent need for novel therapeutic strategies to treat these leukemias.

## **A.3 MLL translocations**

Chromosomal translocation of the *MLL* gene results in an in-frame fusion with one of over 60 partner genes and the expression of a chimeric fusion protein.<sup>33, 27,34</sup> The most frequent MLL fusion partners are AF4, AF9, ENL, AF6, ELL and AF10.<sup>22,35</sup> The resultant fusion protein has the N-terminal domain of MLL conserved and the C-terminal domain of the partner protein. This fusion protein has aberrant regulatory activity and multiple downstream oncogenic effects. The chimeric fusion protein is not post-translationally modified, as is the case for WT MLL, because the Taspase1 protease-binding site is not conserved.<sup>36</sup> The gene that MLL is fused with can have varying effects on the oncogenic activity of the mutant MLL fusion protein. There are two main classes of fusion partners, nuclear and cytosolic. Both classes can have oncogenic downstream effects, but the nuclear pathway fusion proteins are most common and are the most thoroughly characterized. In the nuclear pathway, the chimeric MLL fusion protein disrupts the

activity of the normal epigenetic regulatory complex. This perturbation of WT MLL function results in the upregulation and overexpression of *Hox* genes that are essential for normal hematopoiesis.<sup>37, 38</sup> However, there are less common fusion partners that are cytosolic proteins including AF6, GAS7, EEN, septin proteins.<sup>39</sup> MLL-GAS7 is able to achieve leukemic transformation in the absence of *Hoxa7*, or *Hoxa9*,<sup>40</sup> however, this cytosolic fusion partner transformation mechanism is not well understood. Despite MLL fusions using different mechanisms for transformation, they all may be susceptible to menin-MLL inhibitors.<sup>41</sup> The critical dependence of MLL fusion proteins on the interaction with menin-LEDGF demonstrates the importance of the N-terminal region of MLL. The lack of the C-terminal portion of WT MLL in the MLL fusion proteins would appear to suggest that for the progression of leukemogenesis this region of MLL is not required.

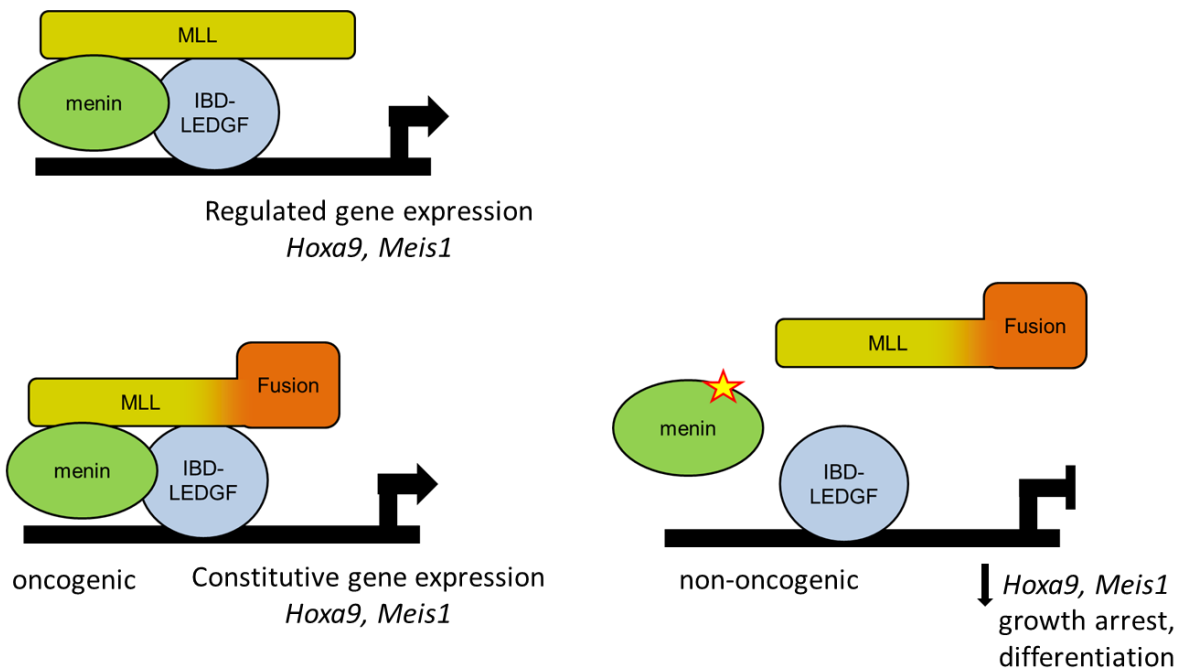
#### **A.4 WT MLL in leukemogenesis**

Cells with *MLL* rearrangements have one allele that remains intact and produces WT MLL.<sup>42,43</sup> Therefore, MLL fusion proteins lack the C-terminus of MLL that is necessary for gene activation but this region is still expressed in the remaining WT allele allowing for WT MLL function to be carried out. It has been demonstrated that the remaining WT allele of *MLL* remains intact in MLL-AF9 leukemias and plays a critical role.<sup>11, 42</sup> Knockdown of WT MLL in MLL-AF9 transformed cells leads to a decrease in *Hox* gene transcription and WT MLL mediated H3K4me3.<sup>42</sup> Together this indicates the WT MLL is required for the maintenance of the transformed state of MLL rearranged leukemias.<sup>42</sup> With the absence of a direct interaction between MLL-AF9 and WT MLL,<sup>11</sup> the WT MLL may play a role in laying the foundational epigenetic landscape that is required for MLL fusion proteins to mediate oncogenesis.

#### **A.5 HOX genes in MLL rearranged leukemia**

Disruption of WT MLL through a chromosomal translocation event can result in the upregulation of *Hox* genes, particularly *Hoxa7* and *Hoxa9*.<sup>38,44,45</sup> This leads to the blockage of hematopoietic differentiation and stimulates rapid cell proliferation resulting in acute leukemia.<sup>46</sup> The deregulation of *Hoxa7*, *Hoxa9* and *Meis1* represents a critical step in the progression of leukemogenesis by MLL fusion proteins.<sup>38,47-49</sup> The molecular mechanisms that HOXA9 uses to regulate subsequent gene expression and block cell differentiation are still largely unknown.

However, it is known that HOXA9 protein can dimerize with MEIS1, which increases the specificity of HOXA9's recognition of target genes.<sup>22</sup> During cell differentiation and the late stages of maturation, *Hoxa9* expression decreases, and overexpression results in the sustained ability for cells to self-renew.<sup>48</sup> However, in MLL translocated leukemias constitutive expression of *Hoxa9* and *Meis1* blocks hematopoietic cell differentiation.<sup>50</sup> The importance of *Hox* genes role in leukemogenic transformation, particularly, the dimerization partners HOXA9 and MEIS1 are the most well characterized targets.<sup>51-53</sup> In addition, the upregulation of *Hoxa9* and *Meis1* genes leads to superior proliferation and differentiation block in hematopoietic cells resulting in acute leukemia.<sup>37,38,50</sup> Transduction of *Hoxa9* and *Meis1* genes into bone marrow progenitor cells accelerates the progression of acute leukemia in mouse models and can mimic MLL-AF9 leukemic disease.<sup>38,54</sup> Importantly, *Hoxa9* and *Meis1* overexpression can rescue the bone marrow colony formation deficiencies caused by loss of menin, a critical interaction partner of MLL, or WT MLL.<sup>12,42</sup> Overall, this demonstrates the importance of *Hox* genes in mediating MLL rearranged leukemias.



**Schematic 1.1 Menin-MLL-LEDGF ternary complex assembly.**

The schematic highlighting the assembly of WT MLL and MLL fusion proteins with menin and LEDGF. The N-terminal region of MLL is maintained in the MLL-fusion proteins where the interaction sites for menin and LEDGF are located.

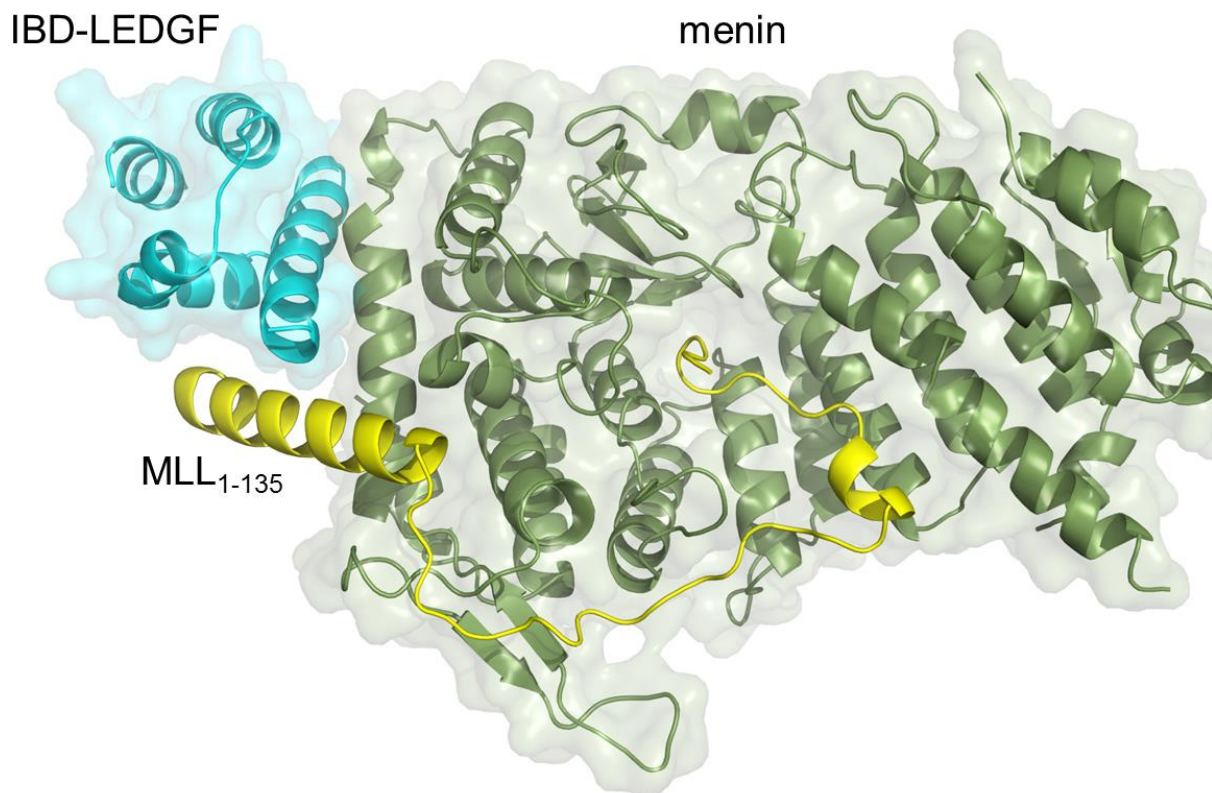
## **B. The menin-MLL-LEDGF ternary complex**

### **B.1 Importance of the interaction between menin-MLL-LEDGF for leukemogenesis**

The MLL fusion proteins oncogenic function is reliant on the interaction with menin<sup>12,55</sup> and with Lens Epithelium Derived Growth Factor (LEDGF).<sup>9</sup> Menin is encoded by the *Multiple Endocrine Neoplasia I (MEN1)* and is considered a tumor suppressor protein in cells of the endocrine lineage.<sup>56</sup> However, menin also is an essential oncogenic cofactor of MLL fusion proteins in the hematopoietic system for the development of acute leukemia.<sup>12,55,57 58</sup> Studies in our laboratory have revealed that menin binds to the N-terminal region of MLL with a very high binding affinity ( $K_d = 6.8$  nM).<sup>55,58,59</sup> Importantly, this N-terminal region of MLL is present in all of the MLL fusion proteins and this interaction is needed for recruiting MLL and MLL fusion proteins to the target genes.<sup>12,55,58,60</sup>

The menin-MLL interaction is necessary for the upregulation of *Hoxa9* and *Meis1* as it has been demonstrated that disruption of this interaction abolishes the aberrant expression of *Hox* genes and reverses the oncogenic characteristics of MLL-transformed leukemic blasts.<sup>55</sup> Therefore, targeting the menin-MLL interaction has been postulated as a possible therapeutic approach for the development of novel therapeutics for MLL rearranged leukemias (**Schematic 1.1**).<sup>22,55,61</sup>

Menin plays an important role as a scaffold protein in the menin-MLL-LEDGF ternary complex. Loss of menin or Lens Epithelial Derived Growth Factor (LEDGF) disrupts localization of MLL or MLL fusion proteins to target genes. Neither menin nor MLL<sup>N</sup> alone are capable of potent interactions with LEDGF.<sup>62</sup> The crystal structure of the ternary complex demonstrates the integrase binding domain (IBD) of LEDGF being involved in direct interactions with both menin and MLL (**Figure 1.1**).<sup>9,62</sup> LEDGF makes direct contacts with menin-MLL1 and disruption of this interaction abrogates acute leukemia in mouse models.<sup>9</sup> Thus, LEDGF is a critical cofactor for MLL associated gene regulation and leukemogenic transformation.<sup>63</sup> This data suggests the LEDGF interaction with menin-MLL may be a promising target for identifying novel therapeutics for MLL dependent leukemias.<sup>9</sup>



**Figure 1.1 Crystal structure of the menin-MLL-LEDGF complex.**

Menin (green) is primarily alpha-helical globular protein with a large central cavity that binds MLL (yellow). MLL makes additional contacts with IBD-LEDGF (cyan) to hold this ternary complex together. Figure is prepared based on the crystal structure of menin-MLL-LEDGF (PDB: 3U88).

## B.2 Structural studies on menin

NMR and bioinformatics studies demonstrate that menin is a globular protein with several regions of disordered fragments.<sup>64</sup> The crystal structure of menin from *Nematostella vectensis* was solved first by Cierpicki's laboratory and deletion of an internal loop fragment and the C-termini was required for this construct to crystalize.<sup>65</sup> Similarly, deletion of an internal loop was necessary to solve the human menin crystal structure.<sup>62</sup> However, the resolution of the human menin crystal structure was rather poor (above 2.5 Å)<sup>62</sup> and it only improved substantially when our laboratory performed more extensive deletion of loop regions to obtain crystals capable of diffracting to 1.3 Å resolution (1.3 Å).<sup>59,62</sup>

Structural analysis of menin revealed a primarily  $\alpha$ -helical protein with a tetratricopeptide repeat (TPR) structural motif.<sup>66,62,65</sup> There is high sequence homology in the structured core region between human menin and *Nematostella* menin exceeding 50%.<sup>65,67</sup> The differences in the sequence of these homologs are evident in the three loop regions and the C-terminus. Menin contains a large central cavity (5,000 Å) formed by the TPR motif. This cavity is relatively rigid and composed of hydrophobic side chains and lined with several acidic residues.<sup>66</sup> This cavity on menin represents the interface site for the MLL or JunD-menin protein-protein interaction.<sup>62,66,68</sup> Comparing the binding interactions of MLL or JunD to apo-menin demonstrate the structural rigidity of this central cavity with no conformational changes. This large, well-defined binding pocket of menin suggests this cavity is suitable to target with small molecule inhibitors.

## B.3 Structure characterization of the menin-MLL interaction

The N-terminal region of MLL is involved in the interaction with menin, specifically the fragment of MLL with the RWRFP sequence.<sup>55</sup> Further characterization of this interface revealed the interaction sequence of MLL required for the high affinity binding extends to the first 46 amino acid of the N-terminus.<sup>58</sup> The Isothermal Titration Calorimetry (ITC) characterization of the menin-MLL interaction performed by Grembecka et al.<sup>59</sup> determined that MLL (residues 1-46) binds to menin with low nanomolar potency ( $K_d = 6.8$  nM).<sup>59</sup> Further analysis of the sequence identified two binding motifs within this N-terminal fragment.<sup>59</sup> The menin binding motif 1 (MBM1) is defined as residues 4-15 of MLL that binds to menin with a



$K_d = 53$  nM and the second motif (MBM2) encompasses MLL residues 23-40 and binds to menin with much lower affinity  $K_d = 1.4$   $\mu$ M as assessed by ITC.<sup>59</sup> Interestingly, these two motifs bind with a negative cooperativity, that is, there is a competition between these two motifs for the binding interaction with menin.<sup>59</sup>

MBM1 binds to the large hydrophobic cavity of menin in an extended conformation forming both electrostatic and hydrophobic interactions.<sup>62,66</sup> Alanine scanning mutagenesis determined that residues F9, P10, and P13 were the most important residues for the high affinity binding of MBM1 (**Figure 1.2 B and C**).<sup>59</sup> F9 is the most important hydrophobic interaction in the MBM1 region with alanine replacement resulting in over 2200-fold decrease in the binding affinity, followed by P10 and P13 to alanine substitution leading to a 30- and 50-fold decrease in the binding affinity, respectively.

The MBM2 motif on MLL is composed of several arginine residues and forms an electrostatic interaction with the acidic residues that line the central cavity of menin.<sup>66</sup> Studies in Cierpicki's lab have identified that D252K and L289K point mutations resulted in a significant reduction of MBM2 binding to menin (**Figure 1.2 D**).<sup>66</sup> Based on this analysis MBM2 has been described as an "electrostatic anchor" to enhance the binding of the high affinity motif, MBM1, with menin.<sup>67</sup> Co-immunoprecipitation experiments have demonstrated a possible more distant menin interaction motif within MLL consisting of residues 640-1251 of MLL.<sup>55</sup> However, the N-terminus of MLL (containing both the MBM1 and MBM2) over expressed is adequate to block proliferation of bone marrow cells transformed with MLL-AF9.<sup>58</sup>

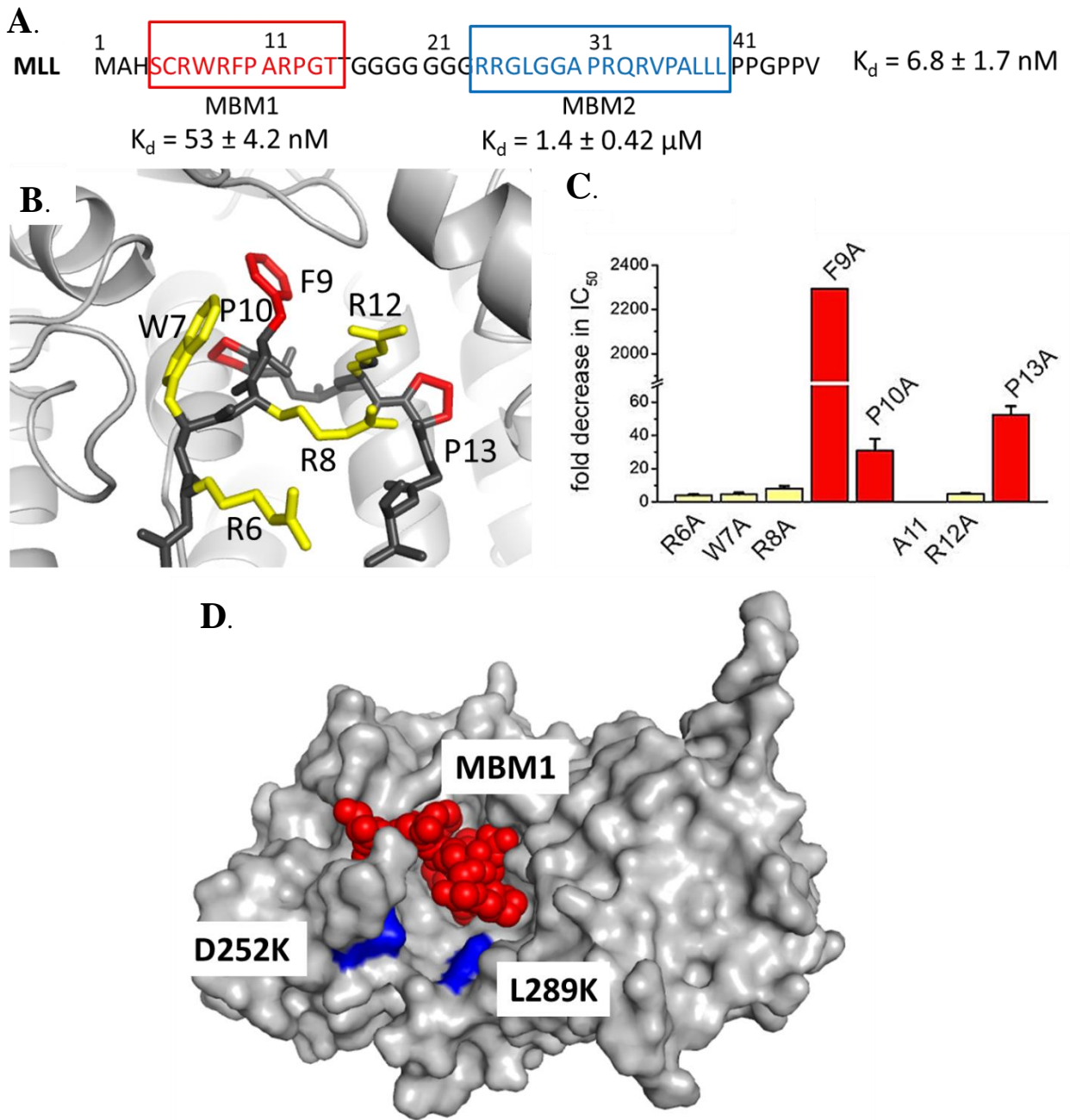
This complicated multivalent interaction of MLL to menin provides a potentially problematic system to inhibit with small molecules. Blocking the whole MLL interaction interface would be difficult to achieve with small molecules. The large hydrophobic cavity on menin provides a promising location for small molecules but the question arises whether disruption of the MBM1-menin interaction is sufficient to block the association of MLL with menin? Importantly, the deletion of just 5 residues of MBM1 was able to abrogate the progression of acute leukemia in mouse models of MLL leukemia in the context of MLL-ENL fusion protein,<sup>55</sup> demonstrating that interfering with the MBM1 interaction with menin is suitable for the effective disruption of the entire menin-MLL complex. Additionally, the alanine scanning mutagenesis identified the F9A mutant to severely impair the binding of the MLL fragment to

menin with a  $K_d = 1.5 \mu\text{M}$ .<sup>59</sup> This marked decrease in the binding affinity is reflected by over 150-fold weaker affinity when compared to the WT sequence.<sup>59</sup> Competition experiments with the MBM1 peptide show that disruption of the interaction of menin and MLL comprising both MBM1 and MBM2 regions is sufficient to completely dissociate the MLL-menin complex. Overall, disruption of the MBM1 high affinity site binding to menin is sufficient to disrupt the bivalent interaction motifs of MLL and completely eliminate the complex formation. Importantly, the large binding interface of the MBM1 region in the central region of menin may pose as a promising cavity to develop small molecule inhibitors.

## C. Targeting the menin-MLL interaction with small molecules

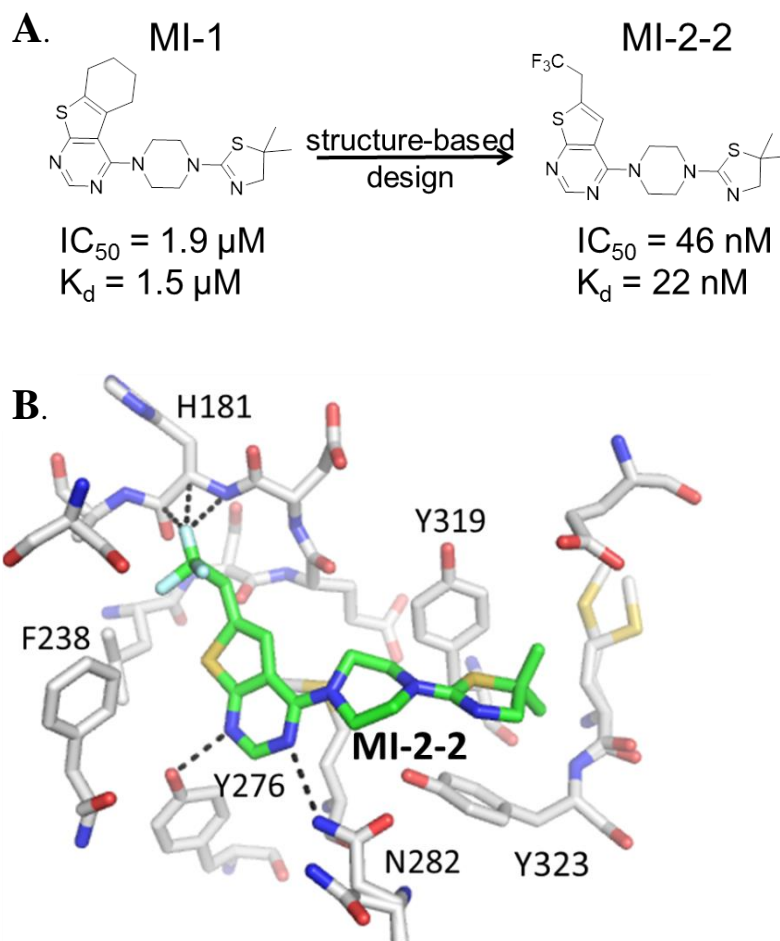
### C.1 Identifying small molecule inhibitors of the menin-MLL interaction

Our laboratory was the first to report small molecule inhibitors of the menin-MLL interaction. Before my involvement in this project, Grembecka and Cierpicki's laboratories have discovered and reported the thienopyrimidine class of menin-MLL inhibitors<sup>68</sup> This class of compounds was found through a High Throughput Screen (HTS) with a competitive fluorescence polarization (FP) assay using a fluorescein (FITC)-labeled peptide corresponding to the MBM1 sequence.<sup>68</sup> The HTS hit (**MI-1**) belonging to the thienopyrimidine scaffold was found as the most potent small molecule from the screen,  $IC_{50} = 1.9 \mu\text{M}$  (**Figure 1.3**).<sup>68</sup> Extensive medicinal chemistry efforts in Grembecka's laboratory combined with structure-based design resulted in **MI-2-2** (**Figure 1.3**) with about 40-fold improved inhibitory activity,  $IC_{50} = 46 \text{ nM}$ .<sup>66</sup>



**Figure 1.2 Characterization of the menin-MLL interaction.**

**A.** Figure adapted from (Cierpicki 2014)<sup>67</sup> The binding constants ( $K_d$ ) were determined by employing FP for the MBM1 peptide and MBM2 peptide while IC was used for the peptide incorporating both MBM1 and MBM2 motifs. All FP measurements were done for synthetic peptides labeled with the N-terminal fluorescein. The ITC experiments were carried out for unlabeled peptides. **B.** Figure adapted from (Cierpicki 2014)<sup>67</sup> The MBM1 binding interface with menin with the most important residues for the binding interaction as determined by alanine scanning mutagenesis followed by FP analysis is highlighted in red. **C.** Figure from (Cierpicki 2014)<sup>67</sup> Quantification of the fold decrease in potency of the mutant peptides determined by FP. **D.** Figure adapted from (Cierpicki 2014)<sup>67</sup> The surface representation of menin structure (gray) is displayed with MBM1 bound (red). Point mutants D252K and L289K which interfere with the MBM2 binding are highlighted in blue.



**Figure 1.3** Thienopyrimidine class of menin-MLL inhibitors.

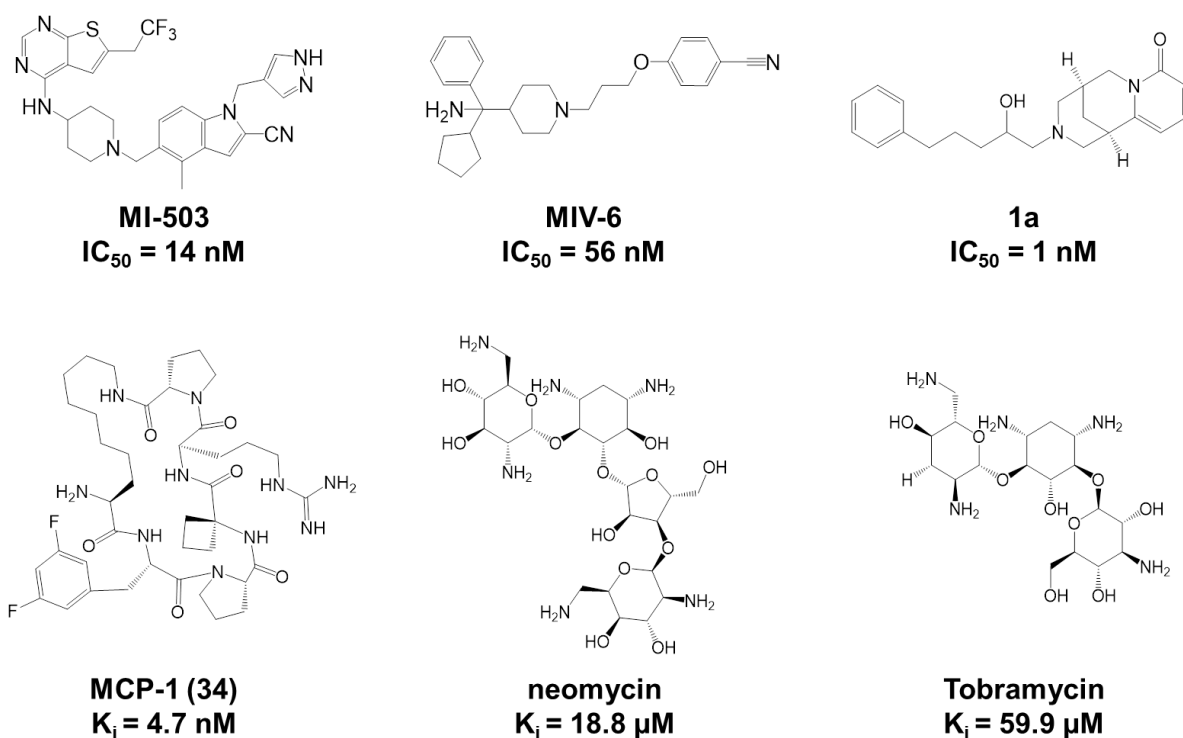
**A.** The development of the thienopyrimidine class of menin-MLL inhibitors from HTS hit compound, **MI-1** to **MI-2-2**. **B.** The co-crystal structure of **MI-2-2** bound to menin.

The crystal structure of **MI-2-2** in complex with menin solved in Dr. Cierpicki's laboratory revealed that **MI-2-2** binds to the MLL binding site on menin. Additionally, there are no major conformational changes of menin when comparing the apo structure with the **MI-2** bound menin structure; this demonstrates the rigidity of this binding pocket. **MI-2-2** forms several hydrogen bonds with the thienopyrimidine region to Tyr276 and Asn282 and hydrophobic contacts with Tyr319 and Tyr323. The trifluoroethyl substituent was in close distance to the protein backbone, resulting in an increased potency of this compound.<sup>66</sup> The fluorine atoms were making orthogonal multipolar contacts with His181 backbone atoms.<sup>1</sup> These types of contacts have been described previously in protein-ligand complexes and can contribute

significantly to the binding potency.<sup>69</sup> Overall, the thienopyrimidine class represents a promising scaffold for further development of menin-MLL inhibitors as a potential therapeutic strategy to treat patients with MLL translocations. Indeed, our additional lead optimization efforts discussed in detail in the following chapters have led to the identification of much more potent thienopyrimidine menin-MLL inhibitors, including **MI-503** ( $IC_{50} = 14\text{nM}$ , **Figure 1.4**) and several others with low nanomolar  $IC_{50}$  values (see Chapter 4).<sup>70,71</sup>

In addition to thienopyrimidine class, our laboratory has pursued the development of a different class of menin-MLL inhibitors, resulting from the HTS performed at the NIH MLPCN (Molecular Libraries Probe Production Centers Network).<sup>72</sup> The inhibitor identified from this screening, **MIV-1** ( $IC_{50} = 12\ \mu\text{M}$ ), which consist of a hydroxymethyl piperidine scaffold, was further optimized within this project, resulting in **MIV-6** ( $IC_{50} = 56\ \text{nM}$ )<sup>73</sup> (**Figure 5.1**), and will be discussed in details in Chapter 5.

During the course of this project other groups have also reported menin-MLL inhibitors. For example, studies in Wang's laboratory (University of Michigan) resulted in peptidomimetics that mimic the MBM1 region of MLL. The MBM1 fragment of MLL binds to menin in a U-shaped conformation with the sidechain of R8 oriented near the backbone carbonyl group of P13 providing a rationale for generating a macrocyclic analog.<sup>74</sup> The macrocyclic analogs significantly improved the potency, and further modification of the side chains produced **MCP-1** with a  $K_i = 4.7\ \text{nM}$  (**Figure 1.4**). The **MCP-1** analog is over ~600-fold more potent than the linear analog.<sup>74</sup> The cellular activity of these peptidomimetics, however, has not been reported suggesting that further optimization is likely necessary to yield cell-permeable drug-like molecules.



**Figure 1.4** Reported inhibitors of the menin-MLL interaction. IC<sub>50</sub> and K<sub>i</sub> values are from a Fluorescence Polarization Assay.

More recently few other reports were published on inhibitors targeting the menin-MLL interface. A natural product-like compound (**1a**) was identified through an *in silico* screening approach followed up with structure-based design and evaluated in biochemical assay for menin-MLL inhibition (**Figure 1.4**).<sup>75</sup> However, no validation of the direct binding of this compound to menin has been provided.<sup>75</sup> Another *in silico* screening approach screened FDA approved therapeutics that inhibits the menin-MLL interaction and identified neomycin and tobramycin (**Figure 1.4**).<sup>76</sup> Follow-up of these hits was then performed through biochemical characterization determining these antibiotics disrupt the menin-MLL binding event in the range of ~20-60 μM representing relatively weak menin-MLL inhibitors.<sup>76</sup> Further validation of these antibiotics for their direct binding to menin is required before pursuing additional optimization efforts.

## C.2 Targeting the bivalent interaction

The bivalent nature of MLL binding to menin embodies a difficult mechanism for inhibition through small molecules. The hydrophobic pocket on menin is relatively rigid and

highly supportive of binding small molecules based on the binding of the MBM1 peptide compared to menin without a binding partner. Evidence has demonstrated that loss of the high affinity MBM1 binding interaction with menin is sufficient for disruption of the entire menin-MLL complex.<sup>77</sup> This allows us to target the MBM1 interaction with small molecule inhibitors leading to abrogation of the leukemogenic mechanism of MLL oncogenic fusion proteins.<sup>59,68</sup> Grembecka et. al. have employed biophysical and biochemical methods to investigate the disruption of the bivalent binding of MLL with menin by small molecule inhibitors.<sup>59</sup> **MI-2-2** was explored in competition experiments with MBM1, MBM2 and MLL<sub>4-43</sub> which encompasses both binding fragments.<sup>68</sup> **MI-2-2** binds to menin with  $K_d = 22$  nM, and inhibits the MBM1-menin binding with an  $IC_{50} = 46$  nM.<sup>66</sup> However, no effect of **MI-2-2** on MBM2-menin interaction was detected.<sup>66</sup> This demonstrates the MBM2 binding interface is not inhibited by **MI-2-2**. Competition experiments with **MI-2-2** in the FP assay incorporating MLL<sub>4-43</sub> showed a 10-fold reduction in the **MI-2-2** potency as compared to the FP assay with the MBM1 peptide.<sup>66</sup> This has shown that even though the MBM2 binding to menin is weak ( $K_d = 1$   $\mu$ M), it does contribute to decreased inhibitory activity of **MI-2-2**. The MLL<sub>4-43</sub> bivalent binding to menin displays slower kinetics of binding in comparison with MBM1 alone and needs a longer incubation period to achieve a stable equilibrium.<sup>68</sup> This demonstrates that small molecule inhibitors targeting the MBM1 interface with menin are capable of disrupting the entire menin-MLL binding event but the presence of MBM2 decreases the potency of these compounds in blocking the menin-MLL interaction. This bivalent interaction produces a more complex system to inhibit the menin-MLL interaction. Inhibitors targeting the MBM1 site will have a reduced potency when targeting the bivalent MLL interaction with menin. As a result, inhibitors targeting solely the MBM1 site will likely require very potent in vitro binding affinity to demonstrate strong efficacy *in vivo*.

### C.3 Potential applications of menin inhibitors

Menin was postulated to be a promising target in acute leukemias as a result of the role of menin in the MLL fusion protein mediated leukemogenic transformation.<sup>55</sup> MLL fusion proteins require the interaction with menin for leukemic transformation and eliminating the menin from the MLL fusion protein inhibits the progression of leukemia. Therefore, small molecule inhibitors of the menin-MLL interaction might represent new potential therapeutic strategy for

MLL leukemia. Furthermore, more recent studies have shown that targeting the menin interface with WT MLL1 and MLL2 could hold possible therapeutic applications in solid cancers including liver,<sup>78</sup> brain,<sup>79</sup> colon,<sup>80</sup> breast<sup>81</sup> cancers and potentially provide a way to chemosensitize cancer cells to chemotherapeutic agents through the down regulation of MDR1.<sup>82</sup> Additionally, WT MLL has been found to be amplified in sub-types of AML and myelodysplasia and Hoxa9 have been found to be overexpressed in up to 50% of AMLs which may further expand the potential therapeutic applications of menin-MLL inhibitors.<sup>34,83-85</sup> Importantly, recent studies have provided evidence to suggest that menin may be a dispensable co-factor of MLL1 for normal hematopoiesis.<sup>86</sup> Therefore, it would be expected that no on-target toxicity would be evident, at least in the context of the normal hematopoietic system, through targeting the menin-MLL complex.<sup>86</sup> In the endocrine system, menin is a tumor suppressor protein requiring an additional investigation of the menin-MLL complex in the endocrine system.<sup>56,57</sup>

Menin is a viable target for small molecule inhibitor development. The central cavity of menin where MBM1 binds is a large, well-defined, relatively rigid pocket providing a promising site for binding small molecule inhibitor. In the subsequent chapters, we will discuss further development of two classes of menin-MLL inhibitors and derive a new class of inhibitors into potent small molecules targeting the menin-MLL interaction. Structure-based design utilizing the high resolution crystals structures of menin in complex with inhibitors greatly facilitated our inhibitor design as described below in detail.



## Chapter 2: Rational Design of Orthogonal Multipolar Interactions with Fluorine in Protein–Ligand Complexes

\*The text and figures in this chapter are adapted from the following manuscript: **Pollock, J.;** Borkin, D.; Lund, G.; Purohit, T.; Dyguda-Kazimierowicz, E.; Grembecka, J.; Cierpicki, T. Rational Design of Orthogonal Multipolar Interactions with Fluorine in Protein-Ligand Complexes. *J Med Chem* 2015, 58, 7465-74.

### A. Abstract

Multipolar interactions involving fluorine and the protein backbone have been frequently observed in protein–ligand complexes. Such fluorine–backbone interactions may substantially contribute to the high affinity of small molecule inhibitors. Here we found that introduction of trifluoromethyl groups into two different sites in the thienopyrimidine class of menin–MLL inhibitors considerably improved their inhibitory activity. In both cases, trifluoromethyl groups are engaged in short interactions with the backbone of menin. In order to understand the effect of fluorine, we synthesized a series of analogues by systematically changing the number of fluorine atoms, and we determined high-resolution crystal structures of the complexes with menin. We found that introduction of fluorine at favorable geometry for interactions with backbone carbonyls may improve the activity of menin–MLL inhibitors as much as 5- to 10-fold. In order to facilitate the design of multipolar fluorine–backbone interactions in protein–ligand complexes, we developed a computational algorithm named FMAP, which calculates fluorophilic sites in proximity to the protein backbone. We postulate that multipolar fluorine–backbone interactions may represent an attractive approach to improve inhibitors of protein–protein interactions.

## B. Background

Fluorine has been recognized as a valuable element in medicinal chemistry, and about 20–25% known drugs contain fluorine atoms.<sup>87-89</sup> Fluorine is the most electronegative element and has a strong effect on physicochemical and conformational properties of organic compounds.<sup>89</sup> As a consequence, introduction of fluorine atoms into ligands is a promising strategy in lead optimization to strengthen protein–ligand interactions. Furthermore, introduction of fluorine into ligand molecules affects physicochemical properties and modulates absorption, distribution, metabolism, and excretion in drug-like molecules.<sup>88,89</sup> Fluorine can enhance ligand affinity through interaction with both polar and hydrophobic groups in proteins.<sup>90</sup> Fluorine scanning has been proposed as an effective strategy for ligand optimization.<sup>69,91</sup> Systematic incorporation of fluorine at different positions in a series of thrombin inhibitors revealed that introduction of fluorine into the benzyl ring enhanced the binding affinity by 6-fold.<sup>91</sup> As a step toward the identification of fluorophilic hot-spots in proteins, it has been proposed to use <sup>19</sup>F NMR ligand-based screening of fluorinated fragments<sup>92</sup> and a combination of screening and computational analysis.<sup>93</sup> However, a rational approach for designing fluorinated ligands is missing.

We previously identified the thienopyrimidine class of compounds which directly bind to menin and inhibit the protein–protein interaction (PPI) between menin and MLL with nanomolar affinity.<sup>68</sup> Substitution of a propyl group on the thienopyrimidine scaffold with trifluoroethyl, which resulted in the **MI-2-2** compound, leads to a significant 10-fold increase in the binding affinity.<sup>66</sup> The crystal structure of **MI-2-2** bound to menin revealed that the CF<sub>3</sub> group is involved in close contacts with the protein backbone. This demonstrates that fluorine–backbone interactions offer excellent opportunities to enhance the activity of inhibitors targeting protein–protein interactions. However, introduction of fluorine atoms into ligand molecules might be synthetically challenging or may require multistep synthesis. Therefore, a method for rational design of favorable fluorine interactions in protein–ligand complexes would significantly facilitate inhibitor development in drug discovery projects. In order to understand the effect of fluorine substitutions, we synthesized series of **MI-2-2** analogues systematically changing the number of fluorine atoms in two different groups and determined high-resolution crystal structures of the inhibitors bound to menin. We found that when fluorine atoms in menin

inhibitors are involved in the orthogonal multipolar C–F $\cdots$ C=O interactions, it significantly enhances ligand binding affinity. On the basis of these findings, we developed a computational algorithm named FMAP to support structure-based design of favorable C–F $\cdots$ C=O interactions in protein–ligand complexes, and we demonstrated its applicability to known fluorine-containing small molecule inhibitors. This study should facilitate rational development of fluorinated ligands for drug discovery applications.

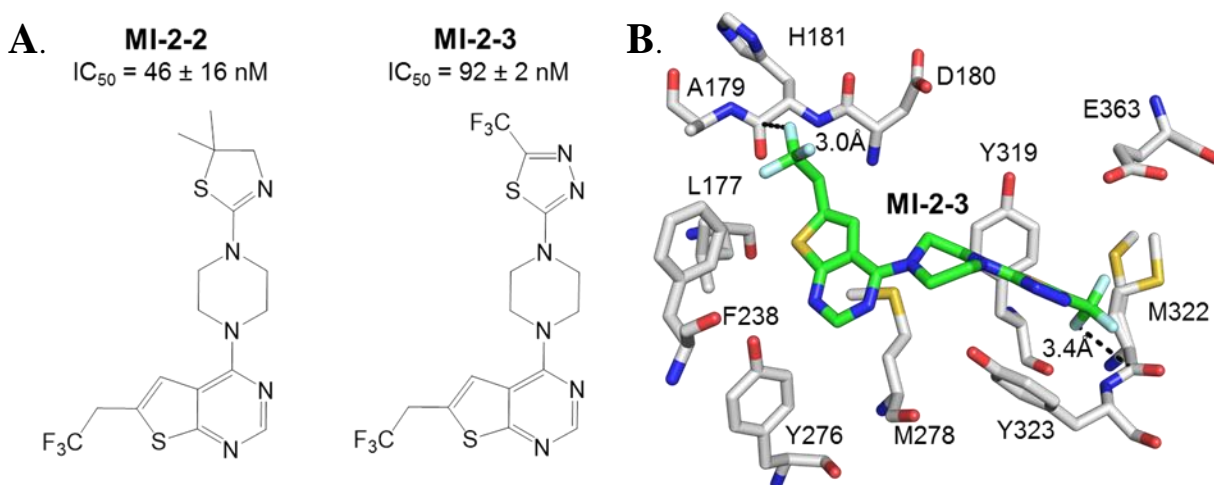
My contribution to this project was to crystalize the menin-inhibitor complexes, solve the crystal structures of these complexes and perform structural analysis of the thienopyrimidine class of inhibitors to characterize the fluorine-backbone interactions with menin. I then performed the FP assay to determine the potency of inhibitors targeting the menin-MLL interaction for each of the analogs generated. To further validate the FMAP algorithm, I performed FMAP analysis of several PDB structures to identify examples where a H to F substitution improved the potency of small molecules.

## C. Results and Discussion

### C.1 Trifluoromethyl Groups in Menin–MLL Inhibitors Form Close Contacts with Protein Backbone.

We previously performed extensive medicinal chemistry optimization of the thienopyrimidine class of menin–MLL inhibitors and found that substitution of propyl in the **MI-2** compound by trifluoroethyl group resulted in a substantial, 10-fold increase in the activity of **MI-2-2** (**Figure 2.1 A**).<sup>1,66</sup> Due to difficulties for further substitutions and potential metabolic liability of the thiazoline moiety, we modified this class of compounds by replacing thiazoline with an aromatic thiadiazole ring. Although the unsubstituted thiadiazole analogue is very weak,<sup>68</sup> we found that introduction of the trifluoromethyl group substantially improved the activity, resulting in **MI-2-3** with IC<sub>50</sub> = 92 nM (**Figure 2.1 A**). Both compounds, **MI-2-2** and **MI-2-3**, are potent inhibitors of the menin–MLL interaction with the IC<sub>50</sub> values below 100 nM (**Figure 2.1 A**). Our previous studies revealed that one fluorine atom from the trifluoroethyl group in **MI-2-2** forms close contacts with the backbone atoms on menin and is located within 3.0 Å distance to the backbone carbonyl of His181,<sup>66</sup> suggesting that this interaction might play an important role in increasing the inhibitory activity of **MI-2-2** over **MI-2**. To understand the molecular basis of high binding affinity of **MI-2-3**, we determined the crystal structure of its

complex with menin. The newly developed **MI-2-3** with an additional trifluoromethyl group binds to menin in a similar binding mode as **MI-2-2** (**Figure 2.1 B**). Interestingly, the new CF<sub>3</sub> group within the trifluoromethyl–thiadiazole moiety also forms close contacts with the menin backbone (**Figure 2.1 B**), and one of the fluorine atoms is located 3.4 Å from the carbonyl group of Met322. Therefore, the fluorine atoms in both CF<sub>3</sub> groups of **MI-2-3** are involved in orthogonal multipolar C–F···C=O interactions with the backbone atoms in two different regions on menin. To assess the contribution of the CF<sub>3</sub> group in **MI-2-3**, we synthesized **MI-326** by replacing trifluoromethyl with the methyl group and found that it led to ~8-fold decrease in the inhibitory activity (IC<sub>50</sub> = 779 nM for **MI-326**).<sup>1</sup> These two examples, **MI-2-2** and **MI-2-3**, emphasize that C–F···C=O contribute very favorably to the protein–ligand interactions.<sup>69,88,91</sup>



**Figure 2.1 Inhibitors of the menin–MLL interaction containing CF<sub>3</sub> groups.**

Figure adapted from (Pollock 2015)<sup>1</sup> **A.** Compound structures and IC<sub>50</sub> values were measured by fluorescence polarization assay. **B.** Crystal structure of **MI-2-3** bound to menin. Short C–F···C=O distances are shown using dashed lines.

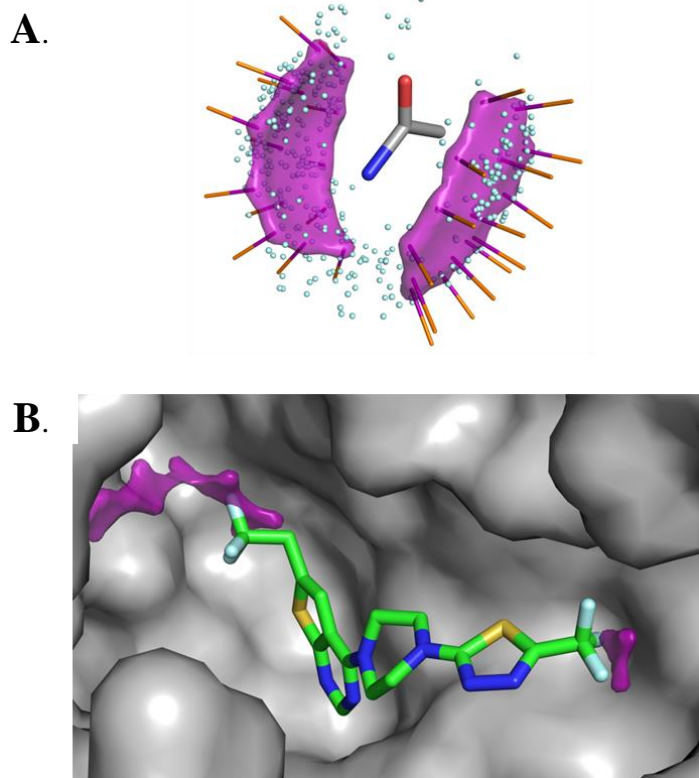
## C.2 Development of FMAP Algorithm To Predict Multipolar C–F···C=O Interactions.

Multipolar interactions involving fluorine atoms have been recognized for their pronounced effect on protein–ligand interactions, and well-placed fluorine may substantially enhance the activity of small molecule inhibitors.<sup>69,88,91,94</sup> Introduction of trifluoromethyl groups in menin inhibitors resulted in a significant improvement of inhibitory activity due to formation of short-distance multipolar interactions with the protein backbone. We therefore sought whether such interactions could be rationally designed. First, we analyzed the geometry of

fluorine–backbone interactions in known high resolution crystal structures of protein–ligand complexes.<sup>1</sup> Out of 2559 structures containing fluorinated ligands, we found 442 complexes with a fluorine atom within 3.5 Å of either the backbone carbonyl carbon or amide nitrogen. This search demonstrated that fluorine is frequently located within a short distance of the backbone carbonyl group with the C–F bond preferably oriented in the orthogonal arrangement relative to the plane of the peptide bond (**Figure 2.2 A**). This exemplifies a presence of multipolar C–F···C=O interactions as described in detail in the previous studies.<sup>88,94</sup>

On the basis of the analysis of protein–ligand complexes from the Protein Data Bank (PDB), we developed an algorithm (FMAP) for mapping sites for fluorine atoms on protein structures to form favorable C–F···C=O interactions with the protein backbone. The geometric criteria used in FMAP have been selected to encompass ~80% of fluorine sites found in the experimental structures in PDB. Fluorine sites are mapped onto a protein structure through a Pymol<sup>95</sup> extension and are represented as a surface spanning 2.8–3.2 Å range from the peptide bond (**Figure 2.2 A**). FMAP also eliminates unlikely fluorine positions through filters based on unfavorable geometry for multipolar interactions as well as steric clashes with protein atoms.<sup>1</sup>

We employed FMAP to analyze the inhibitor binding site on menin and found that there are two potential sites for accessing close contacts between fluorine and protein backbone. Importantly, both sites are occupied by the CF<sub>3</sub> groups in the complex of menin with **MI-2-3** (**Figure 2.2 B**), supporting the utility of FMAP for the prediction of fluorophilic sites in protein structures. The first site is relatively small and is occupied by the CF<sub>3</sub> group connected to the thiadiazole moiety, whereas the second site is much larger and is occupied by the trifluoroethyl group attached to the thienopyrimidine scaffold. Close inspection of the menin-**MI-2-3** crystal structure revealed that only a single fluorine in each CF<sub>3</sub> group has favorable geometry for C–F···C=O interactions with backbone. On the basis of this analysis, we concluded that most likely not all fluorines are needed for high-affinity interactions of menin with the **MI-2-3** and **MI-2-2** inhibitors.



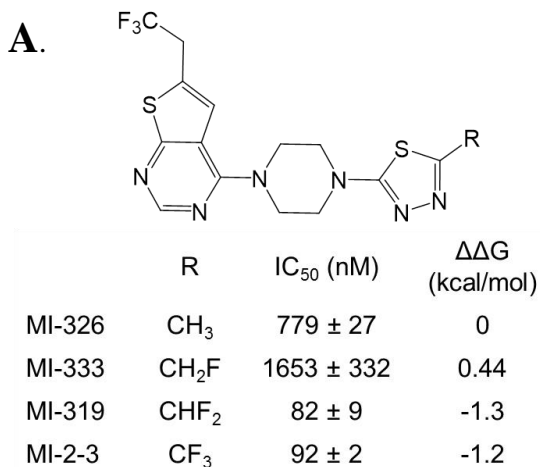
**Figure 2.2 Prediction of favorable C–F $\cdots$ C=O interactions using FMAP algorithm.**

Figure adapted from (Pollock 2015)<sup>1</sup> **A.** Combined analysis of protein–ligand structures from PDB, with FMAP predictions of the potential fluorine positions and their representative C–F bonds relative to backbone peptide bond. Positions of fluorine atoms derived from the protein–ligand complexes found in PDB are shown as cyan points. FMAP prediction is shown as purple surface with orange vectors shown for representative C–F bonds. (George Lund) **B.** FMAP prediction for the menin–MI-2-3 complex. Purple surface represents favorable positions for fluorine atoms to interact with the protein backbone.

**C.3 Interactions of Trifluoromethyl–Thiadiazole Moiety with Menin.**

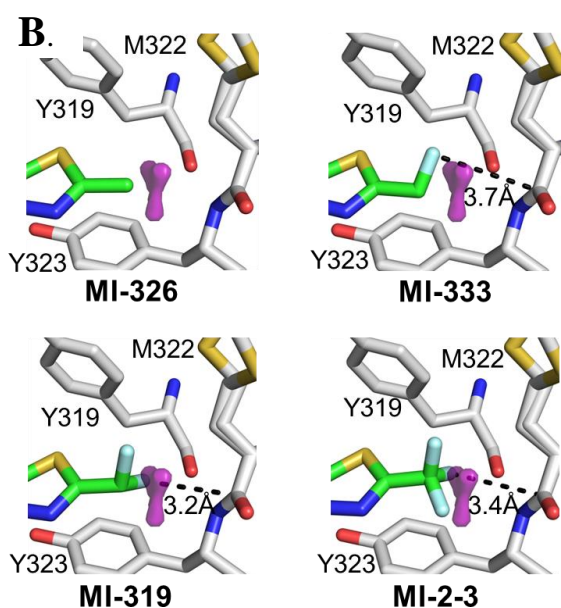
FMAP analysis suggested that only one fluorine atom in the CF<sub>3</sub> group within the trifluoromethyl-thiadiazole moiety of **MI-2-3** is capable of favorable interactions with the backbone carbonyl of Met332. In order to evaluate the contributions of fluorine atoms to the binding affinity of **MI-2-3**, we synthesized a series of analogues replacing CF<sub>3</sub> with CH<sub>3</sub>, CH<sub>2</sub>F, and CHF<sub>2</sub> groups (**Figure 2.3**). First, we assessed the effect of substituting CF<sub>3</sub> by CH<sub>3</sub> and found that the absence of the three fluorine atoms in **MI-326** results in >8 fold decrease in the inhibitory activity (**Figure 2.3 A**). The crystal structure revealed that **MI-326** binds to menin in a very similar manner as **MI-2-3** (**Figure 2.3 B**), and the difference in the binding affinity predominantly results from the loss of the fluorine atoms. We then synthesized and tested two

additional analogues with two (**MI-319**) and single (**MI-333**) fluorines. The inhibitory activity of **MI-319** is very similar to **MI-2-3** indicating no differences between  $\text{CF}_3$  and  $\text{CHF}_2$  groups (**Figure 2.3 A**). Surprisingly, **MI-333**, which harbors the  $\text{CH}_2\text{F}$  group, has about 20-fold weaker activity than **MI-2-3** and is even 2-fold less potent than **MI-326** with no fluorines (**Figure 2.3 A**). To explain this effect, we determined the crystal structures of **MI-333** and **MI-319** bound to menin. The  $\text{CHF}_2$  group in **MI-319** binds in a very similar manner as  $\text{CF}_3$  with one of the fluorine atoms in a short, 3.2 Å, distance to the backbone carbonyl of Met322 (**Figure 2.3 B**). On the contrary, the single fluorine in **MI-333** adopts a position that is tilted approximately  $38.5^\circ$  from the plane of the thiadiazole ring and points away from the protein backbone (3.7 Å distance to C=O of Met322) (**Figure 2.3 B**). As a consequence, the fluorine is too far to be involved in a favorable multipolar  $\text{C}-\text{F}\cdots\text{C}=\text{O}$  interactions, and no gain in the activity is observed for **MI-333** (**Figure 2.3 A**).



**Figure 2.3 Effect of fluorine substitutions in thiadiazole moiety on activity of menin-MLL inhibitors.**

Figure adapted from (Pollock 2015)<sup>1</sup> **A.** Compound structures and IC<sub>50</sub> values were measured by fluorescence polarization assay. ΔΔG values are calculated relative to **MI-326**. **B.** Crystal structures of inhibitors bound to menin showing the shortest distances between fluorine and menin backbone. FMAP prediction is shown as purple surface.



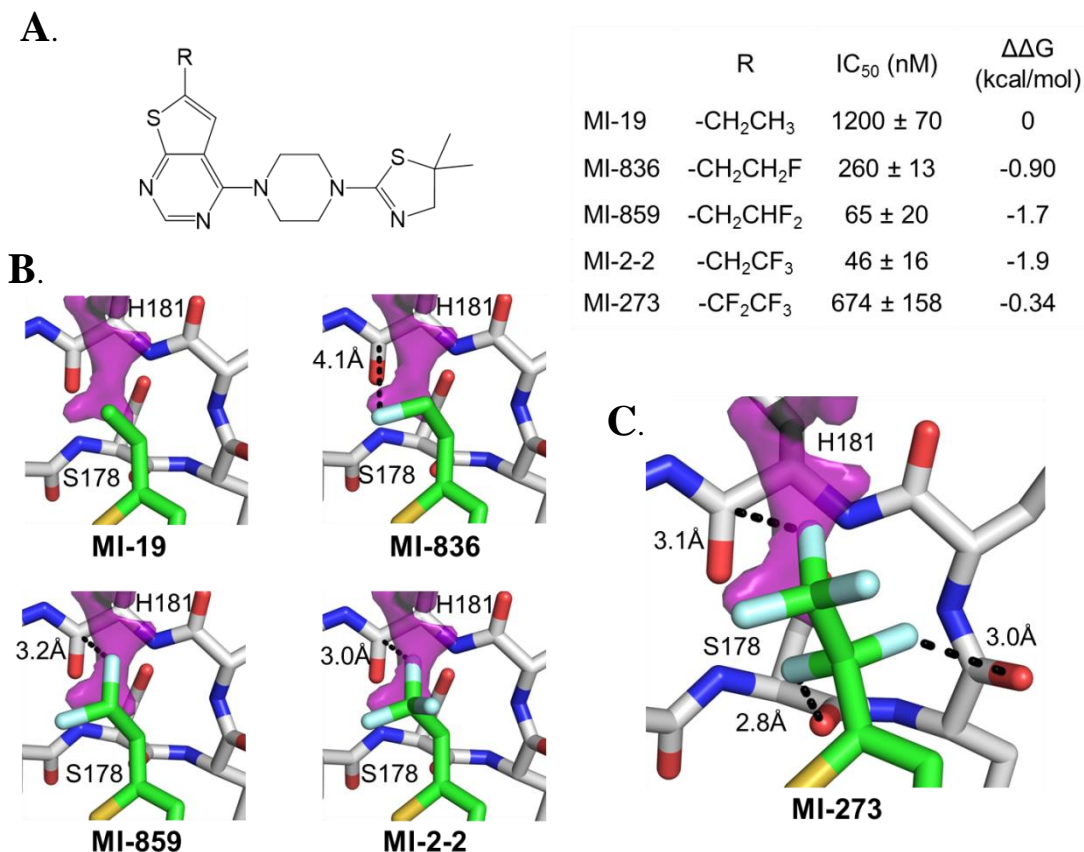
The orientation of the CH<sub>2</sub>F group relative to the thiadiazole ring was unexpected, emphasizing a strong conformational effect of the fluorine atom. As previously observed, substitution of H by F can profoundly change the conformational preferences of a small molecule because of the size and stereoelectronic effects.<sup>88</sup> Although we were able to predict the position of fluorine required for favorable interactions with the protein backbone using FMAP, we did not anticipate that CH<sub>2</sub>F can adopt an orientation where the fluorine points away from the backbone. Introduction of the second fluorine into the CHF<sub>2</sub> group was necessary to achieve an orientation of the C–F bond allowing for favorable C–F⋯C=O interactions and substantial improvement in activity. Analysis of the crystal structure of **MI-333** shows that S–C–C–F dihedral adapts 38.5° angle. Quantum mechanical energy calculations demonstrate that multiple fluorine atoms might be needed to stabilize the appropriate rotameric state.<sup>1</sup>



#### C.4. Interactions of Trifluoroethyl Group in Thienopyrimidine Core with Menin.

Comparison of the activities of **MI-2-2** and **MI-19** indicates that the trifluoroethyl group contributes significantly to the high activity of **MI-2-2**, and replacement of  $\text{CF}_3$  with  $\text{CH}_3$  results in over 20-fold loss in inhibitory activity (**Figure 2.4 A**). FMAP analysis for the menin binding site suggests that only single fluorine in  $\text{CF}_3$  group can form  $\text{C-F}\cdots\text{C=O}$  interactions with the backbone. To test the contributions of individual fluorine atoms in the  $\text{CF}_3$  group of **MI-2-2**, we synthesized two compounds with  $\text{CH}_2\text{F}$  (**MI-836**) or  $\text{CHF}_2$  (**MI-859**) groups. When compared to **MI-19**, addition of the first fluorine enhanced the activity nearly 5-fold, whereas addition of the second fluorine increased the activity further by about 4-fold, making it comparable to **MI-2-2** with  $\text{CF}_3$  group (**Figure 2.4 A**). To rationalize the effect of these modifications, we determined the crystal structures of **MI-836** and **MI-859** bound to menin (**Figure 2.4 B**). We found that the single fluorine in **MI-836** points toward a hydrophobic site formed by the side chains of Leu177, Phe238, Ala182, and Ser155, and therefore, the 5-fold gain in the activity likely results from favorable hydrophobic contacts. Introduction of  $\text{CHF}_2$  in **MI-859** allows for the second fluorine to be involved in the  $\text{C-F}\cdots\text{C=O}$  interactions with the backbone of His181, accounting for an additional 4-fold improvement in activity. Very similar  $\text{IC}_{50}$  values of **MI-859** (with  $\text{CHF}_2$ ) and **MI-2-2** (with  $\text{CF}_3$ ) indicates that the third fluorine is dispensable for binding. Furthermore, the  $\text{cLogP}$  value for **MI-859** is approximately 0.6 unit lower than for **MI-2-2** ( $\text{cLogP} = 3.89$  and  $3.32$  for **MI-2-2** and **MI-859**, respectively). Therefore, our approach based on the FMAP calculations may be used not only to predict fluorine substitutions in ligand molecules but also to design compounds with fewer number of fluorine atoms and reduced lipophilicity without compromising ligand binding affinity. Analysis of the **MI-2-2**-menin structure revealed that the methylene group in the  $\text{CH}_2\text{CF}_3$  moiety is positioned closely to the backbone carbonyl groups of Ser178 and Glu179 and may constitute a further site for fluorine substitutions. However, the FMAP analysis revealed that introduction of the  $\text{CF}_2$  group at this site would not be favorable due to poor geometry of the two fluorines with respect to the carbonyl groups of Ser178 and Glu179. To test this hypothesis, we synthesized **MI-273** with  $\text{CF}_2\text{CF}_3$  group and found that such a substitution results in a  $\sim 15$ -fold decrease in the activity when compared to **MI-2-2** (**Figure 2.4 A**). We determined the crystal structure of **MI-273** bound to menin and found that it binds in an identical manner as **MI-2-2** (**Figure 2.4 C**). The two additional fluorines in the  $\text{CF}_2$  group of **MI-273** are in close distances to the carbonyl oxygens of Ser178 and Glu179 leading to repulsive

interactions. This further emphasizes that favorable C–F···C=O interactions require optimal geometry, and the FMAP approach can filter-out the sites that are unfavorable for fluorine atoms in ligand molecules.



**Figure 2.4 Effect of fluorine substitutions in thienopyrimidine moiety on activity of menin–MLL inhibitors.**

Figure adapted from (Pollock 2015)<sup>1</sup> **A.** Compound structures and IC<sub>50</sub> values were measured by fluorescence polarization assay. ΔΔG values are calculated relative to **MI-19**. **B.** Crystal structures of inhibitors bound to menin showing the shortest distances between fluorine and the menin backbone. FMAP prediction is shown as purple surface. Model of **MI-19** has been made on the basis of the structure of **MI-2-2**-menin complex. **C.** Crystal structure of **MI-273** bound to menin showing close contacts of fluorines in CF<sub>2</sub>CF<sub>3</sub> group with menin. FMAP prediction is shown as purple surface.

## D. Conclusions

Fluorine scanning strategy has been previously proposed as an effective approach to improve the activity of small molecule inhibitors.<sup>69,88,91</sup> Such a strategy is synthetically demanding and requires synthesis of multiple analogues.<sup>69,91</sup> In an attempt to facilitate design of C–F···C=O interactions in protein–ligand complexes we developed the FMAP algorithm. FMAP

uses a crystal structure of a protein–ligand complex and calculates sites surrounding a ligand which could be favorably occupied by fluorine atoms. We demonstrated that FMAP could be used to rationalize improvement in activity upon introduction of fluorine in thienopyrimidine class of menin inhibitors as well as for several known inhibitors. FMAP may also represent a valuable tool for the design of new fluorine substitutions in protein ligands. FMAP relies solely on geometrical and structural criteria, and other effects, such as conformational or electronic changes resulting from fluorine substitution are not taken into account, which might represent a limitation of this approach. Nevertheless, we expect that FMAP can be very useful in the drug discovery projects to rationally design positions for fluorine atoms in ligand molecules. It may also support development of ligands with an optimal number of fluorine atoms to improve binding affinity while reducing ligand hydrophobicity and molecular weight.

Introduction of the CF<sub>3</sub> group in menin inhibitors as well as in several examples reviewed in this study results in a substantial gain in the affinity providing that optimal geometry of the C–F bond relative to the backbone carbonyl is achieved. Such C–F⋯C=O interactions provide unique opportunities to introduce favorable interactions between small molecule ligands and the polar protein backbone. Due to unique orthogonal geometry relative to the protein backbone, these interactions may be introduced into the binding sites where hydrogen bonds are not feasible. We found that substitution of CH<sub>3</sub> for CF<sub>3</sub> may increase ligand binding affinity as much as 10-fold. However, multipolar interactions involving the CF<sub>3</sub> group may not be solely responsible for the increase in binding affinity. The effect of desolvation of more hydrophobic CF<sub>3</sub> group is expected to lead to larger positive entropy of binding when compared with CH<sub>3</sub>.<sup>96</sup> The CF<sub>3</sub> is roughly twice the size of a methyl group<sup>1</sup> and due to a larger size and different shape; it may form more optimal van der Waals contacts within the binding site. Furthermore, the two additional fluorine atoms may participate in hydrophobic interactions with neighboring atoms. In the case of menin inhibitors, we found that not all fluorine atoms in the CF<sub>3</sub> group are needed for the high affinity interaction. However, introduction of CFH<sub>2</sub> or CF<sub>2</sub>H groups to achieve favorable C–F⋯C=O interactions may impact conformational equilibrium and favor a rotamer which cannot favorably interact with protein backbone or might cause high entropic cost of freezing out a desired rotamer. Despite that the H to F substitution represents a relatively minor modification; it may have a complex impact on ligand binding affinity. Our structural data

collected for the menin–inhibitor complexes offers a unique set of data which may facilitate better understanding of the C–F $\cdots$ C=O interactions.

The C–F $\cdots$ C=O interactions have been typically reported for enzyme inhibitors.<sup>69,91 97,98</sup> With increasing interest and demand in development of PPI inhibitors, efficient approaches are needed to optimize protein–ligand interactions at solvent exposed interfaces. As we demonstrated for the menin–MLL inhibitors, fluorine interactions with the protein backbone may offer such opportunities, particularly at the interfaces involving  $\alpha$ -helical or  $\beta$ -sheet structures. In this study, we developed the FMAP approach to streamline the design of C–F $\cdots$ C=O interactions, which adds a new tool for structure-based design of new inhibitors targeting protein–protein interfaces as well as protein ligands in a more general context. The FMAP algorithm may facilitate prediction of fluorine substitutions in ligand molecules and support ligand optimization in drug discovery projects.

## Chapter 3: Development of inhibitors targeting the menin-MLL interaction

\*The text and figures in this chapter are adapted from the following manuscripts:

Borkin, D.; He, S.; Miao, H.; Kempinska, K.; **Pollock**, J.; Chase, J.; Purohit, T.; Malik, B.; Zhao, T.; Wang, J.; Wen, B.; Zong, H.; Jones, M.; Danet-Desnoyers, G.; Guzman, M. L.; Talpaz, M.; Bixby, D. L.; Sun, D.; Hess, J. L.; Muntean, A. G.; Maillard, I.; Cierpicki, T.; Grembecka, J. Pharmacologic inhibition of the Menin-MLL interaction blocks progression of MLL leukemia in vivo. *Cancer Cell* 2015, 27, 589-602.

Borkin, D.; **Pollock**, J.; Kempinska, K.; Purohit, T.; Li, X.; Wen, B.; Zhao, T.; Miao, H.; Shukla, S.; He, M.; Sun, D.; Cierpicki, T.; Grembecka, J. Property Focused Structure-Based Optimization of Small Molecule Inhibitors of the Protein-Protein Interaction between Menin and Mixed Lineage Leukemia (MLL). *J Med Chem* 2016, 59, 892-913.

### A. Abstract

MLL rearrangements generate oncogenic MLL fusion proteins and leads to the development of acute leukemia. The leukemic cells critically rely on the MLL fusion protein interaction with menin for sustained proliferative capacity. This provides the rationale for targeting the menin-MLL interaction with small molecule inhibitors to abrogate MLL-rearranged acute leukemias. We developed the thienopyrimidine class of small molecules to disrupt the menin-MLL interaction. **MI-2-2** served as a promising lead compound with demonstrative potency in leukemic cells but required improvements of drug-like properties for investigation of *in vivo* efficacy. Structure-based design was extensively utilized to design small molecule analogs of **MI-2-2** leading to **MI-136** as a promising pharmacophore. Systematic investigation of indole substituents was performed to improve potency, selectivity, solubility, and pharmacokinetic properties in mouse models. These efforts led to the identification of compounds suitable for *in vivo* analysis and demonstrated a significant survival benefit in murine

models of MLL leukemias. Overall, this study provides a promising scaffold for clinical optimization and potential therapeutic applications.

## B. Background

The menin-MLL fusion protein interaction is critical for the progression of MLL-rearranged acute leukemias.<sup>9,55</sup> Leukemias harboring chromosomal rearrangements of the MLL gene account for up to 10% of acute myeloid and acute lymphoblastic leukemias (AML and ALL)<sup>99</sup>. MLL rearrangements confers poor prognosis with a 5-year survival rate of about 35%<sup>32</sup> as a result of therapeutic intervention with limited efficacy.<sup>24,32,33</sup> This evidence demonstrates the urgent need for effective targeted therapies.

Menin is an essential oncogenic cofactor for MLL fusion proteins and plays the role as a scaffolding protein to bridge MLL and MLL fusion proteins to target genes.<sup>9,55,58</sup> The menin interaction motif of MLL is located at the N-terminus and is maintained in the chromosomal rearrangements of MLL.<sup>34,55,59,100</sup> Leukemic cell proliferation and blockage of hematopoietic cell differentiation is dependent on the MLL interaction with menin.<sup>55,58</sup>

The menin-MLL interaction has been well characterized, including our own work,<sup>62,64,65</sup> and we validated that this protein-protein interaction can be targeted by small molecules.<sup>66,68</sup> The large central binding pocket on menin provides a large hydrophobic cavity for the N-terminal region of MLL to bind. Our group has developed the first small molecules targeting the MLL binding site on menin and effectively blocking this protein-protein interaction.<sup>66,68</sup> However, no demonstration of *in vivo* efficacy has been reported for the menin-MLL inhibitors prior to this study described herein. This owes to the difficulty in developing small molecule inhibitors of PPIs with favorable drug-like properties<sup>101</sup> although examples of PPI inhibitors do exist.<sup>102,103</sup>

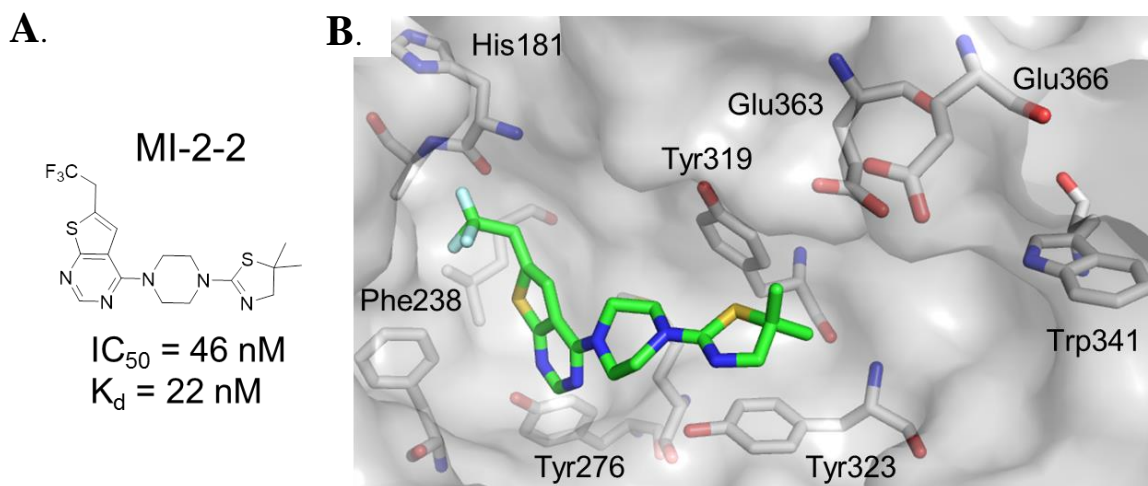
We developed the thienopyrimidine class of small molecule inhibitors with **MI-2-2**, the most potent analog described prior to the data presented in this thesis. This compound exhibited relatively potent activity in MLL rearranged leukemic cells<sup>66</sup> but required further optimization of drug-like properties. We sought to improve the drug-like properties of the **MI-2-2** compound of the thienopyrimidine class of menin-MLL inhibitors. Structural analysis of menin with bound inhibitors guided the rationale for modifications while characterization of these inhibitors was performed through multi-parameter analysis to improve potency, selectivity, solubility, half-life

and exposure for their characterization in mouse models of MLL-leukemias. My contribution to this project focused on structure-based design of small molecule analogs and characterization with biophysical and biochemical approaches.

My contribution to this project was to crystallize the menin-inhibitor complexes, solve the crystal structures of these complexes and perform structural characterization of the thienopyrimidine class of menin-MLL inhibitors for their binding to menin. I have also contributed to the rational design of new analogs targeting menin-MLL interaction to develop more potent and drug-like compounds. I have also performed biochemical characterization of these analogs in an FP assay and biophysical analysis of their direct binding to menin by ITC to determine the binding affinities of the thienopyrimidine class of menin-MLL inhibitors.

## C. Results and Discussion

### C.1 Modifications of MI-2-2 to explore the menin binding pocket



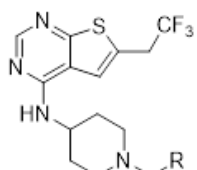
**Figure 3.1 Crystal structure of MI-2-2 bound to menin.**


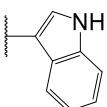
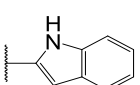
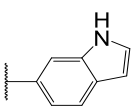
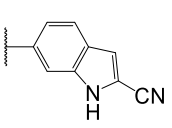
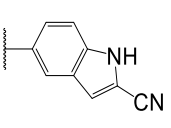
**A.** Structure of MI-2-2,  $IC_{50}$  values were measured by fluorescence polarization assay and binding affinity to menin determined by ITC. **B.** Crystal structure of menin with MI-2-2 bound. Menin is displayed as surface representation and key residues are highlighted.

MI-2-2 has poor drug-like properties and requires optimization of potency in MLL leukemia cells, half-life in microsomal stability assays, and polarity which are currently limiting its use for *in vivo* applications.<sup>70</sup> The crystal structure of MI-2-2 showed an unoccupied region of

menin near the 5,5-dimethyl-thiazolyl fragment, however further substitution of this moiety of **MI-2-2** is synthetically difficult (**Figure 3.1**). Therefore, we decided to leave the thienopyrimidine core unchanged and replaced the piperazine with an amino-piperidine linker substituted further with a benzyl group to generate compound **1**. The scaffold of compound **1** exhibited only a 2-fold reduction in potency but provided a more synthetically tractable molecule to possibly explore substitutions on the benzyl ring. Considering the size of the unoccupied space of **MI-2-2** near the 5,5-dimethyl-thiazolyl fragment, we explored larger substituents at this region of the binding site. Different isomers of the indole to replace the phenyl (compounds **2**, **3**, **4**) were explored; indole was chosen because of the synthetic feasibility of exploring additional substituents (**Table 3.1**). As I will describe in **Chapter 3**, we identified the importance of the nitrile in the MIV class of menin-MLL inhibitors with loss of this hydrogen bond with W341 resulting in a >20-fold decrease in potency. We therefore sought to mimic this interaction in the context of the thienopyrimidine scaffold. Additional substitution of the indole in compound **4** with a nitrile was performed to generate compound **5** with almost a 2-fold improvement in potency, **Table 3.1**. Changing the indole orientation to generate compound **6** resulted in an additional 5-fold improvement in potency (**Table 3.1**).





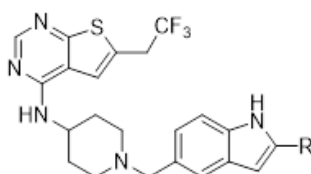
Compound	R	IC <sub>50</sub> (nM)
1		100 ± 7.8
2		237 ± 10
3		2350 ± 495
4		284 ± 44
5		153 ± 7.8
6		31 ± 3.5




**Table 3.1 SAR and activities of analogues of the thienopyrimidine class of menin-MLL inhibitors.** IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided.

into this pocket reflected by over a 500-fold decrease in potency compared to **MI-136**.

Therefore with the limited SAR at position 2 of the indole the nitrile was the most favorable group at this position.

To understand the molecular interactions involved in the **MI-136** binding to menin we performed structural studies of menin-inhibitor complexes. The thienopyrimidine core of **MI-136** binds to menin in the same manner as compared to **MI-2-2** (**Figure 3.2**). The nitrogen of the indole is involved in a hydrogen bond with the carboxyl group of Glu363 with a distance of 2.8 Å. Furthermore, the nitrile group forms a hydrogen bond with the sidechain of Trp341 with a distance of 3.1 Å. We then explored the replacement of the nitrile with other polar groups to maintain the hydrogen bond with Trp341 while lowering the clogP, **Table 3.2**. Replacement of the nitrile at position 2 of the indole with an amide (**8**) lead to 8-fold decrease in potency in the FP assay. The sulfone moiety further decreased the potency about 12-fold compared to **MI-136**. A more bulky 1*H*-tetrazole moiety does not fit well



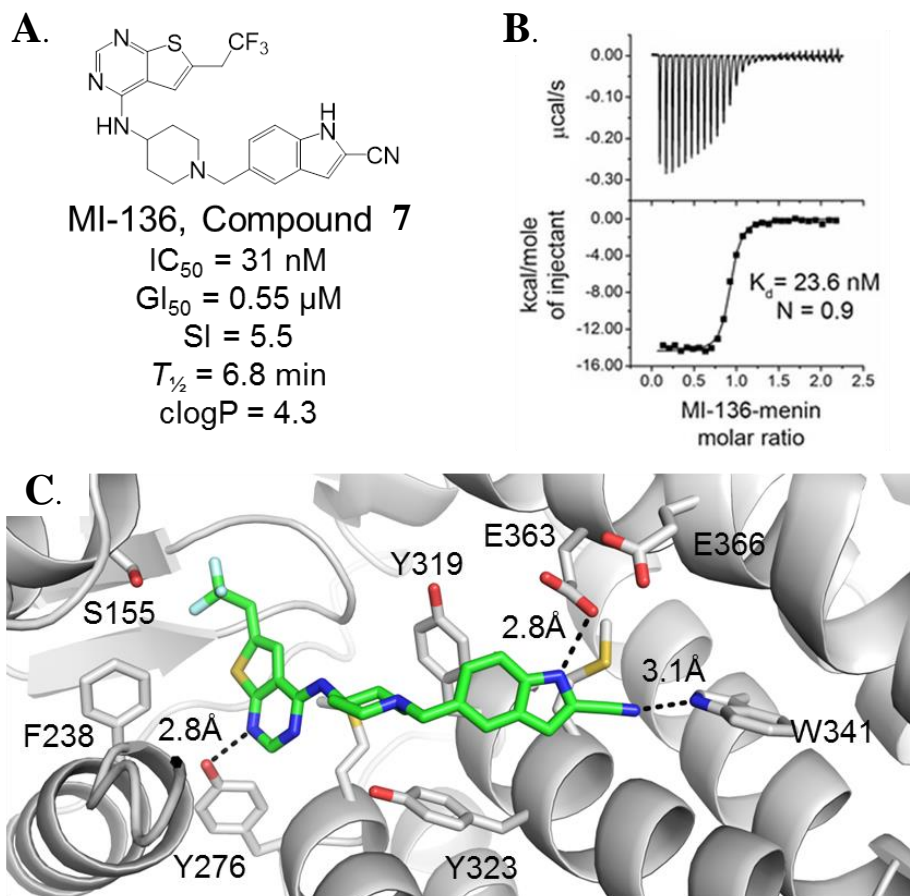
Compound	R	IC <sub>50</sub> (nM)	clogP
7 (MI-136)	-CN	31 ± 3.5	4.3
8		251 ± 78	3.4
9		359 ± 5.0	3.2
10 (MI-405)		15650 ± 3323	4.0

**Table 3.2 SAR and properties of analogues with substitutions at position 2 of indole.** IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided.

**MI-136** had an optimal indole isomer and nitrile substitution with an IC<sub>50</sub> of 31 nM in the FP assay (**Table 3.1, Table 3.2**). We explored the binding affinity of **MI-136** to menin by isothermal titration calorimetry (ITC) and obtained the K<sub>d</sub> = 24 nM (**Figure 3.2**). Cellular potency of **MI-136** was then assessed in MLL-AF9 transformed murine bone marrow cells (mBMCs) using the MTT cell viability assay with a growth inhibition at half maximal (GI<sub>50</sub>) of 0.55 μM (**Figure 3.2**). The half-life (T<sub>1/2</sub>) of **MI-136** in murine liver microsomes was also measured resulting in T<sub>1/2</sub> ~7 min (**Figure 3.2**). Further analysis of **MI-136** against a panel of GPCRs reveals not substantial off-target activation.<sup>71</sup> This data demonstrates **MI-136** is a

promising starting pharmacophore as a menin-MLL inhibitor.

**MI-136** has limited in vivo applications and requires additional development of potency and selectivity in MLL leukemia cells, metabolic stability, PK profile, and polarity of this class of menin inhibitors. We focused medicinal chemistry efforts on the optimization of the cyanoindole of the **MI-136** scaffold. The solubility of **MI-136** with a clogP = 4.3 needed optimization to decrease the lipophilicity which required the addition of more polar substituents. Furthermore, the potency of **MI-136** in MLL-rearranged leukemia cells needed improvement and better selectivity compared to the *Hoxa9/Meis1* transformed mBMCs (control cell line) as measured by the selectivity index (SI) (**Figure 3.2**). We also sought to improve the half-life of **MI-136** in microsomal stability assays. Based on the crystal structure, several positions of the cyanoindole (1, 3, 4, 6) in **MI-136** can be explored for further substitutions to form favorable contacts with menin. Therefore, a multi-parameter optimization was performed generating SAR around the **MI-136** scaffold to not only increase potency but also improve drug-like properties of this compound.



**Figure 3.2 Crystal structure of MI-136 bound to menin.**

**A.** Structure of compound 7, **MI-136** and  $IC_{50}$  values were measured by fluorescence polarization assay, average values from two to three independent measurements  $\pm$  SD are provided. Growth inhibition ( $GI_{50}$ ) measured in the MTT cell viability assay in MLL-AF9 transformed murine bone marrow cells after 7 days of treatment with compounds. SI, selectivity index calculated as a ratio of  $GI_{50}$  values measured in HM-2 cells and MLL-AF9 transformed cells. ( $T_{1/2}$ ) Half-life of compounds in mouse liver microsomes. **B.** Figure adapted from (Borkin 2015).<sup>70</sup> ITC analysis of **MI-136** shows menin binding isotherm.<sup>70</sup> **C.** Crystal structure of menin with **MI-136** bound. Key residues and interactions are highlighted.

## C.2 Investigation of indole nitrogen substituents on MI-136 scaffold

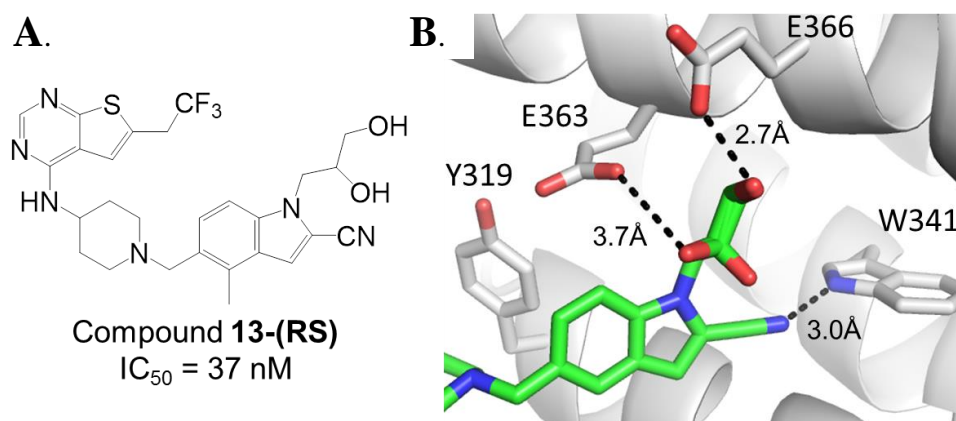
The nitrogen of the indole ring of **MI-136** is involved in a hydrogen bond the Glu363 sidechain and Glu366 sidechain is also in a close proximity (**Figure 3.2**). Based on analysis of several menin structures these glutamic acid side chains are flexible and this area should be susceptible to accommodating larger modifications at the indole nitrogen of **MI-136**. Therefore,

polar substitutions may adequately occupy this region of menin with hydrogen bond donor groups providing the potential for hydrogen bond formation with these acidic residues.

To form a hydrogen bond with the acidic residues near this site, we rationalized a hydroxyl substituent should form a hydrogen bond with these glutamic acids while lowering the clogP. We explored a hydroxymethyl moiety (**11**) at the indole nitrogen and found no change in potency (**Table 3.3**). Further extension of the hydroxyethyl (**12**) led to no improvement in potency compared to **MI-136**. Interestingly, compound **11** was about 4-fold more potent in growth inhibition of MLL-AF9 cells compared to compound **12** despite similar *in vitro* activity and clogP values (**Table 3.3**).

We then explored further extension of the linker to incorporate a 1,2-propanediol group to form compound **13**. Compound **13** had similar potency to compound **12** but was a racemic mixture. To identify if a single enantiomer was more potent, we synthesized each enantiomer. **13-R** and **13-S** had similar potencies and crystallographic analysis of each enantiomer in complex with menin revealed similar binding modes with the only difference being the positioning of the secondary hydroxyl (**Figure 3.3**). The **13-R** compound is the only enantiomer with the secondary hydroxyl group forming a hydrogen bond with Glu363 (**Figure 3.3**). The terminal hydroxyl of both enantiomers also forms a hydrogen bond with Glu366 (**Figure 3.3**).

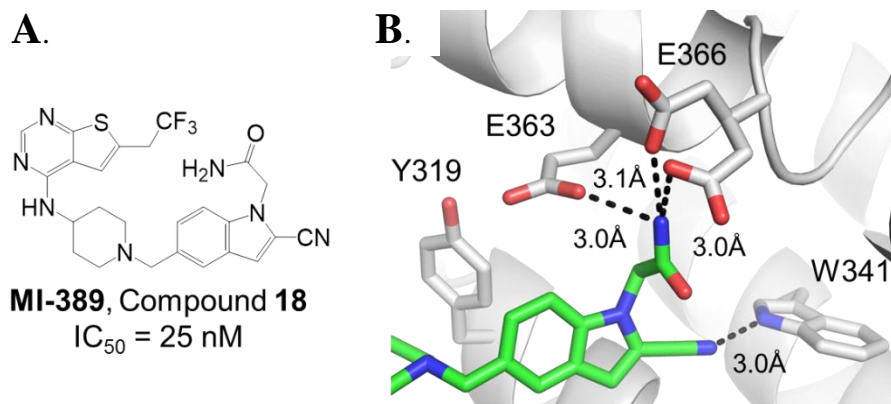
Modification of the secondary hydroxyl to a methoxy in the context of each enantiomer (**14-R**, **14-S**) led to no significant changes in binding affinity for either enantiomer (**Table 3.3**). We then explored substitution of the terminal hydroxyl to methoxy (**15**) which led to a 3-fold decrease in potency compared to compounds **14-R** and **14-S** (**Table 3.3**). Despite similar *in vitro* potencies for the **-R** and **-S** enantiomers of compounds **13** and **14** the cellular activity was about 2-fold more potent for the **-R** enantiomer in both cases (**Table 3.3**). Further modifications of the secondary hydroxyl with a bulkier branched substituent (**16**) led to almost a 2-fold increase in potency in the FP assay. Compound **16** also had 1.5 units higher clogP value compared to compound **13**. Importantly, compound **16** had a significantly shorter  $T_{1/2}$  in murine liver microsomes with <3 min compared to >60 min for compound **13** (**Table 3.3**). Therefore, hydrophobic additions to the secondary hydroxyl resulted in improved activity in MLL-AF9 cells but significantly reducing the half-life in microsomes limiting the application of these analogs *in vivo*.



**Figure 3.3 Structure-based design of indole nitrogen substituents.**

**A.** Structure of compound **13-(RS)** and  $IC_{50}$  values were measured by fluorescence polarization assay. **B.** Crystal structure of compound **13-(RS)** with both enantiomers bound to menin. Hydrogen bonds with Glu363, Glu366 and Trp341 are highlighted.

To explore additional hydrogen bond donors and electrostatic interactions with the glutamic acids Glu363 and Glu366 an aminoethyl moiety at the indole nitrogen was investigated (compound **17**). Compound **17** was 2.5-fold weaker than the hydroxyethyl analog (**12**) in the FP assay and 6-fold weaker in inhibiting proliferation of MLL-AF9 cells (**Table 3.3** and **3.4**). This positively charged amine may impact cell permeability explaining the decrease in cellular activity. To remove the positive charge of the amino group we explored an amide (**MI-389**, compound **18**) modification of the indole nitrogen. Compound **18** had slightly improved potency to compound **12** but improved cellular activity ( $GI_{50} = 0.78$ ) and exhibited a good half-life in mouse liver microsomes ( $T_{1/2} = 54 \text{ min}$ ) (**Table 3.4**). The crystal structure of compound **18** in complex with menin demonstrated the nitrogen of the amide is involved in hydrogen bonds with both Glu363 and Glu366 (**Figure 3.4**). **18** has also improved the polarity significantly with a  $clogP = 3$ . Addition of an  $\alpha$ -methylated amide (**19**) had a 10-fold reduction in potency (**Table 3.4**), likely due to steric clash with the protein.

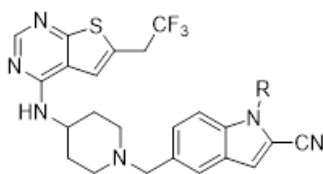


**Figure 3.4 Structure-based design of indole nitrogen substituents.**

**A.** Structure of compound **18** and  $IC_{50}$  values were measured by fluorescence polarization assay. **B.** Crystal structure of compound **18** bound to menin. Hydrogen bonds with Glu363, Glu366 and Trp341 are highlighted.

Next, several 5-member aromatic heterocycles with hydrogen bond donors were explored. A 4-methyl-1*H*-imidazole group in compound **20** had slight decrease in potency *in vitro* compared to compound **18** but almost a 4-fold improvement in cellular activity (**Table 3.4**). The half-life of **20** was only about 5 min in microsomes. A 3-methyl-1*H*-pyrazole in compound **21** had similar inhibitory activity to compound **20** *in vitro* and in cells with no improvement in half-life (**Table 3.4**). A 5-methyl-2*H*-tetrazole in compound **22** and the 3-methyl-1*H*-1,2,4-triazole in compound **23** had about a 5-fold decrease and 3-fold decrease in potency respectively compared to compounds **20** and **21**. The 4-methyl-1*H*-pyrazole in compound **24** had about a 2-fold improvement in the FP assay compared to compound **20**. The 4-methyl-1*H*-1,2,3-triazole in compound **25** had similar potency compared to compound **24**. Compounds **24** and **25** displayed the best microsomal stability half-life in this series of heterocycles with good cellular activity. Limitations in the microsomal stability half-life may possibly be overcome by exploring modifications at other regions of the cyanoindole.

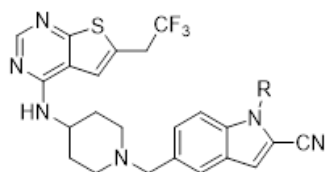
Modifications of the indole nitrogen demonstrate a valuable region for substitutions to increase potency, solubility and stability of these analogs of **MI-136**. The substitutions with the best potency in MLL-AF9 cells are 4-methylpyrazole (**24**) and 4-methyltriazole (**25**). The substitutions with the best solubility based on clogP include the hydroxyethyl (**12**) and amide (**18**).



Compound	R	IC <sub>50</sub> (nM)	GI <sub>50</sub> (uM)			T <sub>1/2</sub> (min)	clogP
			MLL-AF9	HM-2	SI		
7 (MI-136)	H	31 ± 3.5	0.55	3	5.5	6.8	4.3
11		31 ± 6	0.25	4	16	4.7	3.3
12		35 ± 2.8	1.1	6.2	5.6	18	3.5
13-(RS)		37 ± 2.8	3.4	>12	>3.5	>60	3.1
13-R		46 ± 2.8	1.7	>40	>24	48	3.1
13-S		42 ± 0.7	3.8	ND	ND	>60	3.1
14-R		40 ± 1.4	0.36	9.7	27	<3	3.9
14-S		37 ± 6.0	0.66	6.9	11	5.2	3.9
15		101 ± 13	2.1	ND	ND	<3	3.4
16		23 ± 1.4	0.18	3.5	19	<3	4.6

**Table 3.3 SAR and properties of analogues with substitutions at position 1 of indole.**

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided. Growth inhibition (GI<sub>50</sub>) measured in the MTT cell viability assay in MLL-AF9 or Hoxa9/Meis1 (HM-2) transformed murine bone marrow cells after 7 days of treatment with compounds. SI, selectivity index calculated as a ratio of GI<sub>50</sub> values measured in HM-2 cells and MLL-AF9 transformed cells. (T<sub>1/2</sub>) Half-life of compounds in mouse liver microsomes.



Compound	R	IC <sub>50</sub> (nM)	GI <sub>50</sub> (uM)		SI	T <sub>1/2</sub> (min)	clogP
			MLL-AF9	HM-2			
17		110 ± 16	6.3	>24	4	>60	3.6
18 (MI-389)		25 ± 4.9	0.78	25	32	54	3
19		244 ± 57	ND	ND	ND	ND	3.3
20		42 ± 9.2	0.2	15	75	4.8	3.7
21		31 ± 4.9	0.3	ND	ND	3	3.9
22		153 ± 7.1	ND	ND	ND	ND	3.2
23		94 ± 21	10	ND	ND	4	3.4
24		23 ± 4.2	0.26	6.9	27	8.3	3.9
25		27 ± 2.8	0.33	10	30	11	3.5

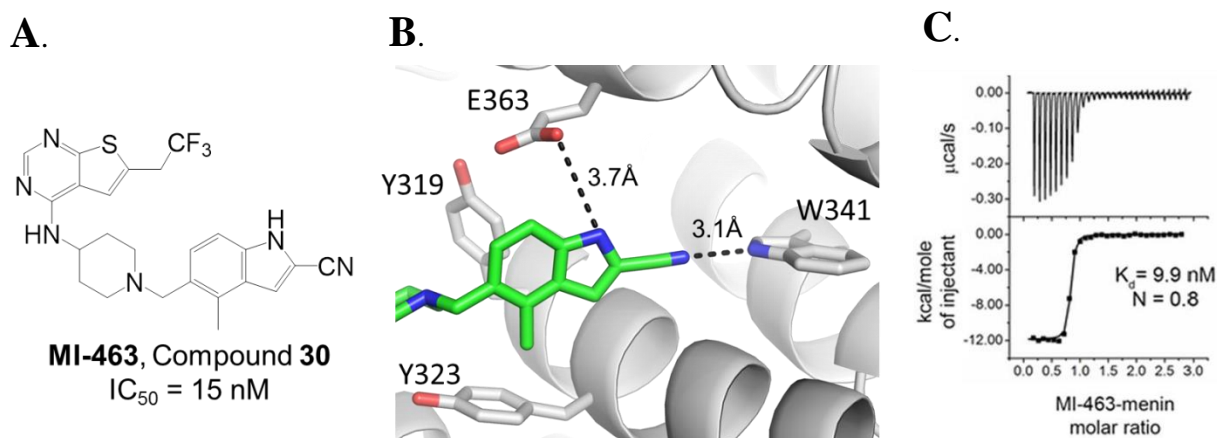
**Table 3.4 SAR and properties of analogues with substitutions at position 1 of indole.**

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided. Growth inhibition (GI<sub>50</sub>) measured in the MTT cell viability assay in MLL-AF9 or Hoxa9/Meis1 (HM-2) transformed murine bone marrow cells after 7 days of treatment with compounds. SI, selectivity index calculated as a ratio of GI<sub>50</sub> values measured in HM-2 cells and MLL-AF9 transformed cells. (T<sub>1/2</sub>) Half-life of compounds in mouse liver microsomes.



### C.3. Exploring modifications at positions 3, 4, and 6 of the of indole ring in MI-136 scaffold

The crystal structure of **MI-136** in complex with menin demonstrated that the carbon at position 3 of the indole is close to the protein backbone (**Figure 3.2**). We explored a few small substitutions at this position to possibly improve the potency of **MI-136**. Fluorine substitution (**26**) resulted in a 3-fold decrease in potency and poor cellular activity (**Table 3.5**). Amino substitution (**27**) and methyl (**28**) at this position were also not favorable and exhibited poor cellular activity (**Table 3.5**). Based on these data we concluded that position 3 on the indole is unable to accommodate further substitutions most likely due to lack of space for modifications.



**Figure 3.5 Structure-based design of substituents at position 4 of the indole.**

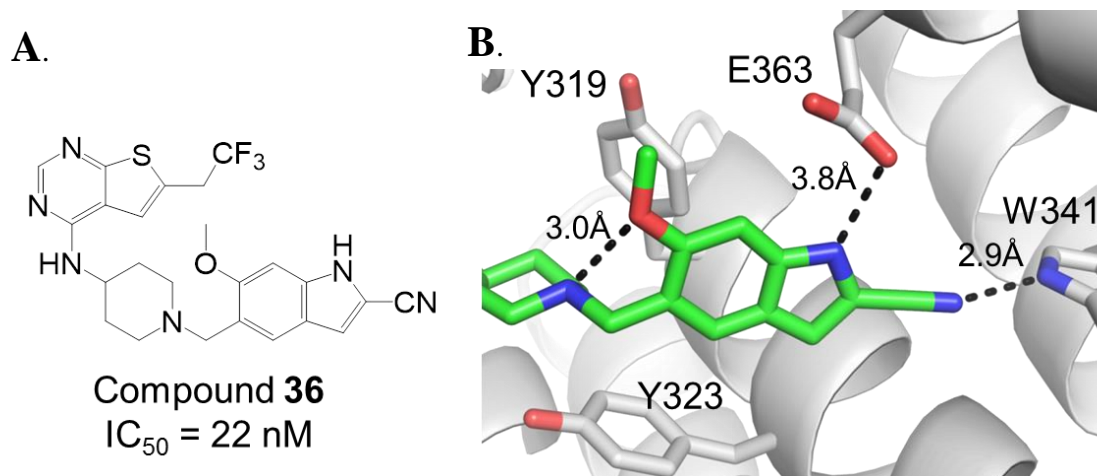
**A.** Structure of **MI-463** (compound **30**) and  $IC_{50}$  values were measured by fluorescence polarization assay is displayed. **B.** Crystal structure of compound **30** bound to menin. Hydrogen bonds with Glu363 and Trp341 are highlighted. **C.** Figure adapted from (Borkin 2015).<sup>70</sup> ITC analysis of compound **30** shows menin binding isotherm.<sup>70</sup>

In contrast, the crystal structure of menin-**MI-136** revealed that position 4 of the indole approaches a hydrophobic patch formed primarily by Tyr323. Analysis of the **MI-136** crystal structure suggests that small hydrophobic substitutions could fit in this region of the binding site. Chlorine at position 4 (compound **29**) had good in vitro activity ( $IC_{50} = 22$  nM) but cellular activity of this compound was relatively weak ( $GI_{50} = 1.1$   $\mu$ M in MLL-AF9 transformed cells). The microsomal stability half-life of compound **29** was improved 4-fold compared to **MI-136**. Interestingly, substitution with a methyl at position 3 of the indole (**MI-463**, compound **30**) increased inhibitory activity by about 2-fold in the FP assay compared to compound **29** and showed significant improvement in the potency in MLL-AF9 cells with the  $GI_{50} = 0.2$   $\mu$ M (**Table 3.5**). Compound **30** had about 2-fold decrease in half-life in microsomes compared to

compound **29**. Structural analysis of compound **30** demonstrates a similar binding mode to the one observed for **MI-136** with the methyl group forming hydrophobic contacts with Tyr323 (**Figure 3.5 B**). Further extension of the methyl to an ethyl (**31**) led to decrease in both *in vitro* inhibitory activity and cellular potency (**Table 3.5**). In attempt to maintain a lower clogP with substituents at this position a hydroxyl group was incorporated (**32**). Unfortunately, this led to a decrease in *in vitro* inhibitory activity versus **MI-463**. Replacement of the hydroxyl with methoxy (**36**) led to a slight improvement *in vitro* potency but a 2.5-fold improvement activity in MLL-AF9 cells compared to the ethyl (**31**).

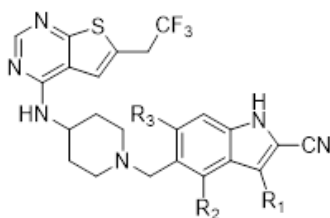
Position 6 of the indole is primarily solvent exposed and therefore small substitutions of the indole were explored to stabilize the conformation of the indole. Fluorine was not tolerated at this position as reflected by over 3-fold decrease in activity compared to unsubstituted **MI-136** (**Table 3.5**). Hydroxyl at this position (compound **39**) lowered the clogP = 3.7 and had excellent activity in MLL-AF9 cells with a GI<sub>50</sub> = 0.18 μM and improved SI over **MI-136**. Further substitution to a methoxy (**36**) or ethoxy (**37**) resulted in a slight decrease in growth inhibition in cells compared to **35** (**Table 3.5**).

In order to gain a thorough understanding of the SAR at this position we performed structural analysis of compound **36** in complex with menin (**Figure 3.6 B**). Compound **36** demonstrated an intramolecular hydrogen bond (3.0 Å) between the oxygen of the methoxy and the nitrogen of the piperidine ring. This interaction most likely leads to the stabilization of this indole conformation that is conducive to the binding conformation to menin (**Figure 3.6 B**). Overall, the hydroxyl and methoxy groups at position 6 were most optimal substituents with improvements in cellular activity and slight improvement in clogP. Improvements in the metabolic stability to achieve a longer half-life are required to investigate the efficacy of these molecules in the mouse models of MLL leukemia.



**Figure 3.6 Structure-based design of indole position 6 modifications in MI-136.**

**A.** Structure of compound **36**, compound **36** and  $IC_{50}$  values were measured by fluorescence polarization assay is displayed. **B.** Crystal structure of compound 40 bound to menin (right panel). Hydrogen bonds with Glu363 and Trp341 are highlighted.



Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	IC <sub>50</sub> (nM)	GI <sub>50</sub> (uM)			T <sub>1/2</sub> (min)	clogP
					MLL-AF9	HM-2	SI		
<b>7 (MI-136)</b>	H	H	H	31 ± 3.5	0.55	3	5.5	6.8	4.3
<b>26</b>	F	H	H	96 ± 0.7	1.7	ND	ND	5.7	4.5
<b>27</b>	NH <sub>2</sub>	H	H	124 ± 30	5.2	ND	ND	27	3.3
<b>28</b>	CH <sub>3</sub>	H	H	111 ± 16	2.6	ND	ND	4.8	4.9
<b>29</b>	H	Cl	H	22 ± 3.5	1.1	11	10	28	5
<b>30 (MI-463)</b>	H	CH <sub>3</sub>	H	15 ± 1.4	0.2	5.6	28	14	4.7
<b>31</b>	H	CH <sub>2</sub> CH <sub>3</sub>	H	21 ± 4.2	1	ND	ND	10	5.2
<b>32</b>	H	OH	H	138 ± 15	ND	ND	ND	4	3.7
<b>33</b>	H	OCH <sub>3</sub>	H	62 ± 17	0.4	5.5	14	10	4.3
<b>34</b>	H	H	F	109 ± 5.6	3	>12	>4	9	4.5
<b>35</b>	H	H	OH	20 ± 5	0.18	>10	>56	6.6	3.7
<b>36</b>	H	H	OCH <sub>3</sub>	22 ± 0.7	0.34	16	47	3.7	4.3
<b>37</b>	H	H	Oet	47 ± 4.2	0.45	5.3	12	11	4.8

**Table 3.5 SAR and properties of analogues with substitutions at position 3, 4, or 6 of indole.**

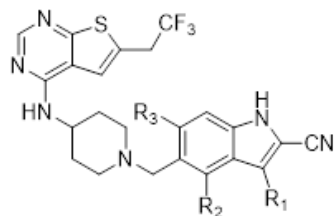
IC<sub>50</sub> values for inhibition of the menin-MLL interaction measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided. Growth inhibition (GI<sub>50</sub>) measured in the MTT cell viability assay in MLL-AF9 or Hoxa9/Meis1 (HM-2) transformed murine bone marrow cells after 7 days of treatment with compounds. SI, selectivity index calculated as a ratio of GI<sub>50</sub> values measured in HM-2 cells and MLL-AF9 transformed cells. (T<sub>1/2</sub>) Half-life of compounds in mouse liver microsomes.

#### C.4 Generating analogs through combinations of optimal substituents on the indole ring in MI-136 scaffold

Combinations of the optimal substituents at positions 1, 4, and 6 of the indole ring were explored to identify if the simultaneous integration of these moieties will improve the potency and drug-like properties of these inhibitors. We first explored the combinations at positions 4 and 6 of the indole. Compound **38** combined the methyl at position 4 and hydroxyl at position 6 which led to about a 2-fold decrease in potency (**Table 3.6**) with relatively similar activity in MLL-AF9 cells. Compound **38** had significantly improved microsomal metabolic stability compared to 4-methyl (**38**) or 6-hydroxyl (**35**) alone. Methoxy group at position 4 in the context of 4-methyl (compound **39**) had improved activity in MLL-AF9 cells but slight decreased  $T_{1/2}$  compared to **38**.

We went on to explore the combinations at position 4 and 6 in the context of hydroxyethyl at position 1 of the indole. Compound **12** with unsubstituted 4 and 6 positions had poor cellular activity despite a low  $\text{clogP} = 3.5$ . Combinations with methyl at position 4 and the absence of a substituent (**40**), hydroxy (**41**) or methoxy (**42**) at position 6 were explored to improve potency. The triple substituted compound (**42**) with the methoxy at position 6 was the most potent in the MTT cellular assay with slight improvement compared to compound **40** or **41**. The methyl at position 4 (**40**) was the only analog that improved microsomal stability compared to compound **12**.

The amide at position 1 (**18**) had sub-micromolar potency in MLL-AF9 cells and almost an hour half-life in murine microsomes. Combination with methyl at position 4 (**43**) or 4-methyl and 6-methoxy (**44**) resulted in improvement in the cellular activity with some improvement in half-life. This group of analogs had improved  $\text{clogP}$  values as well. There was not much of an improvement in cellular potency between compounds (**43**) and (**44**), therefore the methoxy at position 6 may not be required for these analogs.

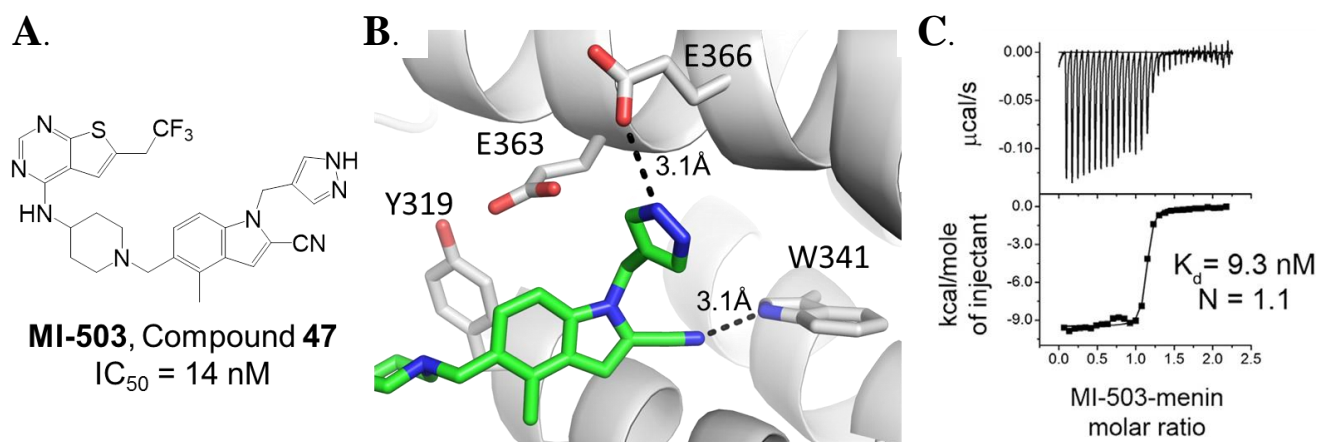


Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	IC <sub>50</sub> (nM)	GI <sub>50</sub> (uM)			T <sub>1/2</sub> (min)	clogP
					MLL-AF9	HM-2	SI		
7 (MI-136)	H	H	H	31 ± 3.5	0.55	3	5.5	6.8	4.3
30 (MI-463)	H	CH <sub>3</sub>	H	15 ± 1.4	0.2	5.6	28	14	4.7
38	H	CH <sub>3</sub>	OH	27 ± 1.4	0.22	12	55	56	4.1
39	H	CH <sub>3</sub>	OCH <sub>3</sub>	40 ± 0.9	0.1	11	110	39	4.7
12		H	H	35 ± 2.8	1.1	6.2	5.6	18	3.5
40		CH <sub>3</sub>	H	44 ± 8.4	0.44	5.6	13	23	3.9
41		CH <sub>3</sub>	OH	27 ± 0.5	0.4	10	25	11	3.7
42		CH <sub>3</sub>	OCH <sub>3</sub>	16 ± 2.0	0.36	15	69	7.8	4.1
18 (MI-389)		H	H	25 ± 4.9	0.78	25	32	54	3
43		CH <sub>3</sub>	H	27 ± 5.6	0.47	>25	>53	59	3.4
44		CH <sub>3</sub>	OCH <sub>3</sub>	37 ± 3.5	0.42	>25	>60	>60	3.5

**Table 3.6 SAR and properties of analogues with substitutions at positions 3, 4, or 6 of indole.**

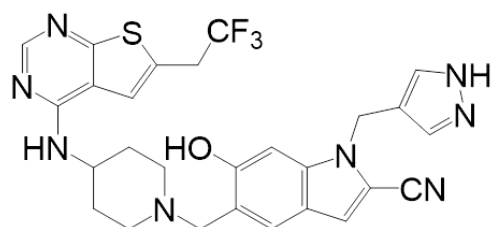
IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided. Growth inhibition (GI<sub>50</sub>) measured in the MTT cell viability assay in MLL-AF9 or Hoxa9/Meis1 (HM-2) transformed murine bone marrow cells after 7 days of treatment with compounds. SI, selectivity index calculated as a ratio of GI<sub>50</sub> values measured in HM-2 cells and MLL-AF9 transformed cells. (T<sub>1/2</sub>) Half-life of compounds in mouse liver microsomes.

The series of 5-membered aromatic heterocycles resulted in analogs with optimal substituents at position 1 of the indole in terms of potency and growth inhibition in MLL-AF9 cells but required improvement in the microsomal stability half-life of this series. Therefore, investigation of combinations of substituents at positions 4 and 6 in the context of methylpyrazole (**24**) and methyltriazole (**25**) was explored. The methylpyrazole substituent present in **24** was systematically explored with optimal combinations of substituents at position 4 and 6. The 4-hydroxy substituent (**45**) improved cellular activity 4-fold with a  $GI_{50} = 83$  nM value and a  $SI > 100$  (**Table 3.7**). Methoxy group at position 4 (**46**) decreased the potency almost 3-fold but improved the microsomal stability half-life over 2-fold. Methyl at position 4 (compound **47**, **MI-503**) slightly decreased the microsomal stability half-life compared to (**46**) and slightly improved *in vitro* potency. Structure determination of compound **47** with menin shows the methylpyrazole involved in a hydrogen bond (3.1 Å) with Glu366 (**Figure 3.7**). The 4-methyl and 6-hydroxyl combination of substituents (**48**) did not lead to significant improvements in growth inhibition of MLL-AF9 cells. Compound **49** with the 4-methyl and 6-methoxy substituents improved potency in cells and improved the half-life to about 1h (**Table 3.7**).

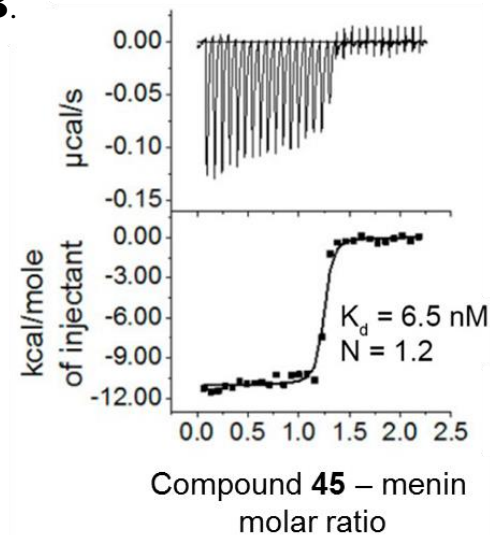


**Figure 3.7 Exploring combinations of substituent at positions 1 and 4 of indole.**

**A.** Structure of **MI-503**, compound **47** and  $IC_{50}$  values were measured by fluorescence polarization assay is displayed. **B.** Crystal structure of compound **47** bound to menin (middle panel). Hydrogen bonds with Glu366 and Trp341 are highlighted. **C.** Figure adapted from (Borkin 2015).<sup>70</sup> ITC analysis of compound **47** shows menin binding isotherm (right panel).<sup>70</sup>

**A.**

**Compound 45**  
 $IC_{50} = 21 \text{ nM}$

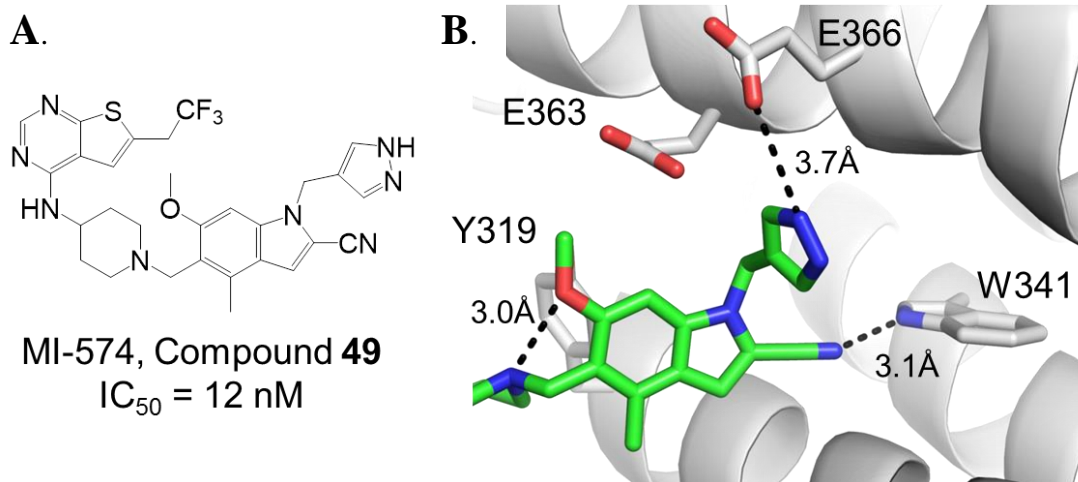
**B.**

**Figure 3.8 Analysis of substituent combinations at positions 1 and 6 of indole.**

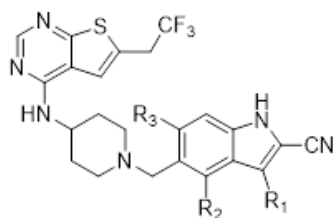
A. Structure of compound **45** and inhibitory activity for inhibition of the menin-MLL interaction as determined in FP assay is displayed. B. Figure adapted from (Borkin 2016).<sup>71</sup> ITC analysis of compound **45** shows menin binding isotherm.<sup>70</sup>

The 4-methyltriazole (**25**) was then explored in combination with the 4-methyl (**50**), the 4-methyl 6-hydroxyl (**51**), and the 4-methyl 6-methoxy (**52**). Compound **50** exhibited the most potent cellular activity and longest microsomal stability half-life, around 1h. Interestingly, comparing compound **50** and **47**, the impact of the 4-methyl on compound **50** is almost 3-fold higher compared to **47** despite relatively similar half-lives in murine microsomes of the unsubstituted 4 position of the indole (**Table 3.7**). Structure determination of menin in complex with compound **49** shows a similar binding mode to each singly substituted analog of the cyanoinidole ring (**Figure 3.9 B**).





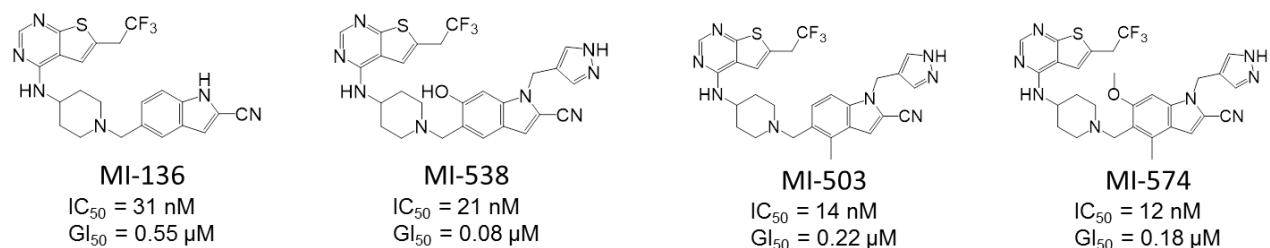
**Figure 3.9 Exploring combinations of substituent at positions 1, 4 and 6 of indole.**  
**A.** Structure of compound **49** and  $IC_{50}$  values were measured by fluorescence polarization assay is displayed. **B.** Crystal structure of compound **49** bound to menin. Intramolecular hydrogen bond and hydrogen bonds with Glu366 and Trp341 are highlighted.



Compound	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	IC <sub>50</sub> (nM)	GI <sub>50</sub> (uM)			T <sub>1/2</sub> (min)	clogP
					MLL-AF9	HM-2	SI		
24		H	H	23 ± 4.2	0.26	6.9	27	8.3	3.9
45 (MI-538)		H	OH	21 ± 1.4	0.083	9	108	14	3.7
46		H	OCH <sub>3</sub>	24 ± 3.0	0.29	6.7	23	35	4.1
47 (MI-503)		CH <sub>3</sub>	H	14 ± 2.1	0.22	>5	>23	21	4.4
48		CH <sub>3</sub>	OH	18 ± 1.4	0.27	13	48	25	4.2
49 (MI-574)		CH <sub>3</sub>	OCH <sub>3</sub>	12 ± 2.0	0.18	4.2	23	55	4.5
25		H	H	27 ± 2.8	0.33	10	30	11	3.5
50		CH <sub>3</sub>	H	23 ± 2.0	0.25	8	32	59	4
51		CH <sub>3</sub>	OH	17 ± 1.4	0.42	10	24	25	3.8
52		CH <sub>3</sub>	OCH <sub>3</sub>	16 ± 1.4	0.41	9	22	32	4.1

**Table 3.7 SAR and properties of analogues with substitutions at positions 3, 4, or 6 of indole.**

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided. Growth inhibition (GI<sub>50</sub>) measured in the MTT cell viability assay in MLL-AF9 or Hoxa9/Meis1 (HM-2) transformed murine bone marrow cells after 7 days of treatment with compounds. SI, selectivity index calculated as a ratio of GI<sub>50</sub> values measured in HM-2 cells and MLL-AF9 transformed cells. (T<sub>1/2</sub>) Half-life of compounds in mouse liver microsomes.



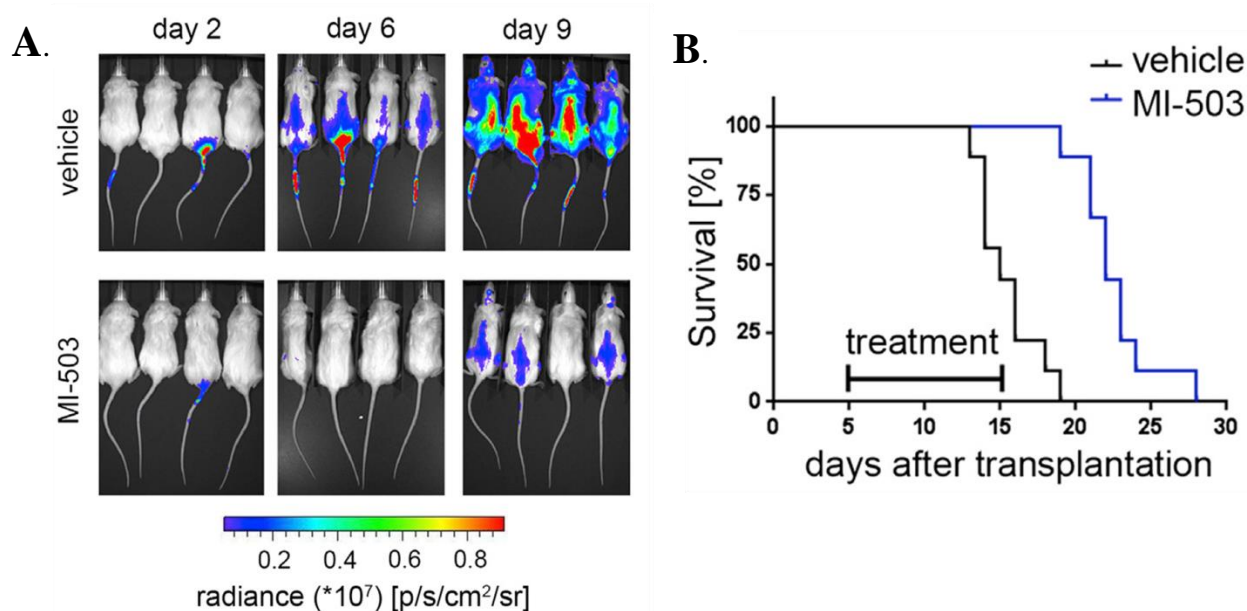
Compound	route of admin.	dose (mg/kg)	$T_{max}$ (h)	$C_{max}$ (ng/mL)	$AUC_{obs}$ (h·ng/mL)	$T_{1/2}$ (h)	$Cl_{obs}$ (mL min <sup>-1</sup> kg <sup>-1</sup> )	$V_{ss}$ (mL/kg)	$F$ (%)
MI-136	iv	15			6,576	3.1	38.1	6,516	
	po	30	1	1516	7,509	3.1			57
MI-503	iv	15			20,624	2.8	12	1,432	
	po	30	2	3975	30,192	2.9			73
MI-538	iv	15			17,897	1.6	14	2,347	
	po	30	2	4572	18,090	1.6			51
MI-574	iv	15			16,235	1.4	15	1,837	
	po	30	2	101	383	NA			1

**Table 3.8 Pharmacokinetic profile of menin-MLL inhibitors in blood plasma of murine models.**

Figure adapted from (Borkin 2015 and Borkin 2016)<sup>70,71</sup> (Top) Structures and  $aIC_{50}$  values were measured by fluorescence polarization assay. Growth inhibition ( $GI_{50}$  values) measured in the MTT cell viability assay in MLL-AF9 transformed murine bone marrow cells after 7 days of treatment with compounds. (Bottom)  $C_{max}$  (maximum compound concentration) in blood plasma,  $AUC_{obs}$  (area under the curve),  $Cl_{obs}$  (Clearance),  $T_{1/2}$  (half-life),  $V_{ss}$  (volume of distribution),  $F$  (oral bioavailability). **MI-574** was an 8 h experiment; the other compounds were analyzed with 24h experiments.

### C.5 Pharmacokinetic analysis of menin-MLL inhibitors in a murine model.

Compounds that exhibited good cellular activity, selectivity, and microsomal stability were selected for follow-up studies in mice to determine the PK profile. The PK profile for **MI-136** was determined to be used as a reference. Intravenous and oral routes of dosing were explored for all compounds. **MI-503** and **MI-538** demonstrated enhanced exposure (area under the curve, AUC),  $C_{max}$  (maximum compound concentration), and a low rate of clearance ( $Cl_{obs}$ ) (**Table 3.8**) compared to the parent compound, **MI-136**. There was almost a 2-fold difference in  $T_{1/2}$  between compounds **MI-503** and **MI-538**. Further investigation of this difference revealed **MI-538** undergoing phase II metabolism with a glucuronidation site at the hydroxyl substituent.<sup>71</sup> As a result, **MI-503** was chosen for further investigation *in vivo*.



**Figure 3.10 Treatment with MI-503 reduces tumor burden and demonstrates survival benefit.**

Figure adapted from (Borkin 2015).<sup>70</sup> **A.** Bioluminescent imaging of MV4:11 cells expressing luciferase transplanted into NSGS mice. The time points after the start of treatment with **MI-503** at 60mg/kg, twice daily i.p. or vehicle, n = 9 mice per group.<sup>70</sup> **B.** Kaplan-Meier survival curves of **MI-503** and vehicle treated C57BL/6 mice transplanted with  $1 \times 10^5$  syngeneic MLL-AF9 leukemic cells isolated from primary recipient mice (n=9 mice per group). Treatment was performed for 10 days as indicated on the graph. The p values were calculated using the log rank (Mantel-Cox) test and are <0.0001 for each compound.<sup>70</sup>

### C.6 MI-503 treatment reduces MLL leukemia burden and improves survival in murine model.

Treatment with **MI-503** was explored in a disseminated xenotransplantation murine model using MV4;11 cells expressing luciferase. NSGS mice were injected through the tail vein with the initiation of treatment 5 days later. The progression of leukemia was monitored via bioluminescence imaging. Twice daily treatment was used as a result of the PK profile of **MI-503** to maintain therapeutically relevant compound concentration in the blood plasma. Treatment was performed for ten consecutive days which demonstrated a decreased proliferation rate of MV4;11 cells compared to vehicle samples as evident by the reduction in bioluminescence intensity (**Figure 3.10 A**).

We then investigated the effect of **MI-503** on the survival benefit of a MLL-AF9 transformed BMCs murine model. Treatment with **MI-503** at 80mg/kg twice daily by oral

gavage began 5 days post transplantation and continued for 10 consecutive days. Treatment with **MI-503** demonstrated a median survival benefit of ~45% compared to vehicle treated mice. This evidence demonstrates the pronounced effect of **MI-503** on inhibiting MLL leukemia development *in vivo*.<sup>70</sup>

## D. Conclusion

MLL rearranged leukemias represent an area of urgent need for the development of novel therapeutics.<sup>32</sup> The menin-MLL interaction has been proposed to be a promising therapeutic target as a result of analysis by genetic manipulations.<sup>55,86</sup> We have demonstrated that inhibition of the menin-MLL interactions with pharmacological agents impairs the progression of MLL rearranged leukemia.

Overall we have performed optimization of the lead compound (**MI-136**) of the thienopyrimidine scaffold. A multi-parameter optimization was performed for this thienopyrimidine class of inhibitors to achieve an improved PK profile for *in vivo* utility. **MI-136** required improvement in potency, solubility, microsomal stability, and oral bioavailability for investigation *in vivo*. Unsubstituted position 3 of the indole ring was most optimal with all of the substituents explored at this position decreasing *in vitro* potency. At position 4 of the indole the methyl substituent was optimal and provided potent activity in MLL-AF9 cells. The hydroxyl at position 6 of the indole increased polarity and was potent in MLL-AF9 cells. The optimal analog at position 1 was the 4-methylpyrazole which improved solubility and potency in cells. Interestingly, when we combined the optimal substituents at each position we did not find a synergistic improvement in the half-life of murine liver microsomes. Ultimately the combinations of the methyl at position 4 and the 4-methylpyrazole at position 1 of the indole generating compound **MI-503** exhibited the best PK properties.

We were able to demonstrate that treatment with **MI-503** reduces the MLL leukemia tumor burden and contributes to a significant survival benefit. Importantly, despite the role of MLL in normal hematopoiesis,<sup>104</sup> the menin-MLL inhibitors do not effect this process.<sup>70</sup> This analysis supports the further development of the thienopyrimidine class of menin-MLL inhibitors for clinical development. Additional optimization is required to investigate the optimal formulation, PK analysis in other species and long-term toxicology. Additional analogs

surrounding substituents of the cyanoindole ring could be explored to circumvent any possible side-effects associated with **MI-503** or related analogs.

## **Chapter 4: Design and characterization of the hydroxy- and aminomethylpiperidine class of menin-MLL inhibitors**

\* The text and figures in this chapter are adapted from the following manuscripts:  
He, S.; Senter, T. J.; **Pollock, J.**; Han, C.; Upadhyay, S. K.; Purohit, T.; Gogliotti, R. D.; Lindsley, C. W.; Cierpicki, T.; Stauffer, S. R.; Grembecka, J. High-affinity small-molecule inhibitors of the menin-mixed lineage leukemia (MLL) interaction closely mimic a natural protein-protein interaction. *J Med Chem* 2014, 57, 1543-56.

### **A. Abstract**

The menin-MLL interaction is crucial for the development of MLL rearranged acute leukemias. Therefore, targeting this interaction with small molecule inhibitors represents a potential therapeutic method of modulating the disease progression. A high-throughput screening campaign led to the identification of a novel class of small molecules targeting the menin-MLL complex. The hydroxy- and aminomethylpiperidine class of menin-MLL inhibitors closely mimic the most critical residues of MLL for the binding interaction with menin. We used structure-based design to develop this lead compound into a potent inhibitor of the menin-MLL complex and demonstrate efficacy in leukemic cells. This novel scaffold provides a new class of inhibitors for the development of a potential therapeutic approach to targeting menin-MLL dependent leukemias.

## B. Background

### B.1 Identification of a chemically distinct class of menin-MLL inhibitors

Developing small molecule inhibitors of protein-protein interactions (PPIs) is a challenging task.<sup>105,106</sup> Small molecule inhibitors of PPIs typically require increased molecular weight to achieve sufficient potency.<sup>101</sup> As a result, they can provide challenges with cell-permeability, poor oral bioavailability, and optimized drug like properties.<sup>101</sup> Successful development of PPI inhibitors have been reported<sup>106-110</sup> providing evidence that some PPIs are susceptible to small molecule inhibition. Some examples have reached clinical trials,<sup>111,112</sup> supporting the effort for development of PPI inhibitors.

We and others have demonstrated that the menin-MLL interaction is an important target for MLL-rearranged leukemias.<sup>77,68</sup> However, there are several difficulties targeting this PPI. The large interface of the menin-MLL interaction poses a difficult system to inhibit with small molecule inhibitors. The most important hotspot residues have been characterized but simultaneous inhibition of the P10, P13 and F9 pockets may be difficult. Additionally, the bivalent interaction motifs provide another level of complexity and difficulty in the development of inhibitors of the interaction.<sup>66,67</sup>

We have described the development of the thienopyrimidine class of menin-MLL inhibitors in previous chapters but we sought to identify a novel pharmacophore to target this interaction. The thienopyrimidine classes of inhibitors do not bind to all of the hotspot regions important for the MLL high affinity interaction with menin.<sup>66</sup> The P10 pocket is unoccupied by the thienopyrimidine class of menin-MLL inhibitors with the F9 and P13 pockets occupied by these compounds. This limitation could provide a barrier to achieve potent menin-MLL inhibitors. Therefore, a chemically distinct scaffold from the thienopyrimidine class of inhibitors will provide an alternative method for developing inhibitors of this interaction. It is important in inhibitor discovery campaigns to have more than one lead compound because it is difficult to predict the potential for downstream toxicity or off-target effects that are associated with an entire class of compounds. In addition, recent other efforts to identify inhibitors of this interaction have been reported including peptidomimetics but the absence of cellular and *in vivo* data suggests possible issues around drug-like property optimization.<sup>67</sup> This demonstrates the



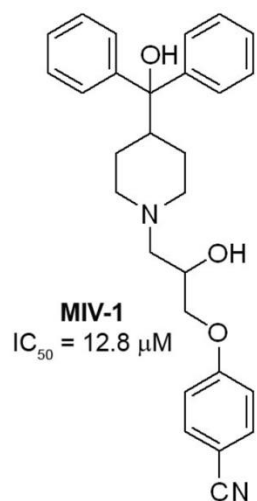
need for alternative pharmacophores to increase the potential of identifying a clinical candidate therapeutic.

My contribution to this project was to crystalize the menin-inhibitor complexes, solve the crystal structures of these complexes and perform structural characterization of the thienopyrimidine class of menin-MLL inhibitors for their binding to menin. I also contributed to the rational design of these inhibitors to develop more potent analogs with improved drug-like properties. I also performed biochemical characterization of these analogs in an FP assay to determine inhibitory activity of these compounds and biophysical analysis by ITC to assess the binding affinity of the methylpiperidine analogs for their binding to menin.

## **C. Results and Discussion**

### **C.1 HTS campaign and identification of the minimal pharmacophore.**

A HTS campaign was performed to identify novel small molecule inhibitors of the menin-MLL complex using a library of about 288,000 compounds at the NIH Molecular Libraries Probe Production Centers Network (MLPCN).<sup>113</sup> The primary screen utilized a fluorescence polarization (FP) assay with a fluorescein-labeled MBM1.<sup>73</sup> A follow-up secondary screen was performed with a homogenous time-resolved fluorescence (HTRF) assay.<sup>73</sup> Finally, NMR saturation transfer difference (STD) was used to validate the direct binding of these small molecules to the menin protein.<sup>73</sup>



**Figure 4.1 Identification of the most potent HTS hit.**

Figure adapted from (He 2014).<sup>73</sup>  
Structure of **MIV-1** and activity in FP assay.

The best inhibitor identified from this screen was **1**, **MIV-1**, (4-(2-hydroxy-3-(4-(hydroxydiphenylmethyl)piperidin-1-yl)-propoxy)benzoyl)benzotrile (**Figure 4.1**). This compound belongs to the hydroxymethylpiperidine class and exhibited an IC<sub>50</sub> value of 12.8 μM for inhibition of the menin-MLL interaction.

To identify the minimal pharmacophore for this class of compounds we investigated removal of the hydroxyl group in the linker of **MIV-1** to generate **MIV-2 (2)** (**Table 4.1**).

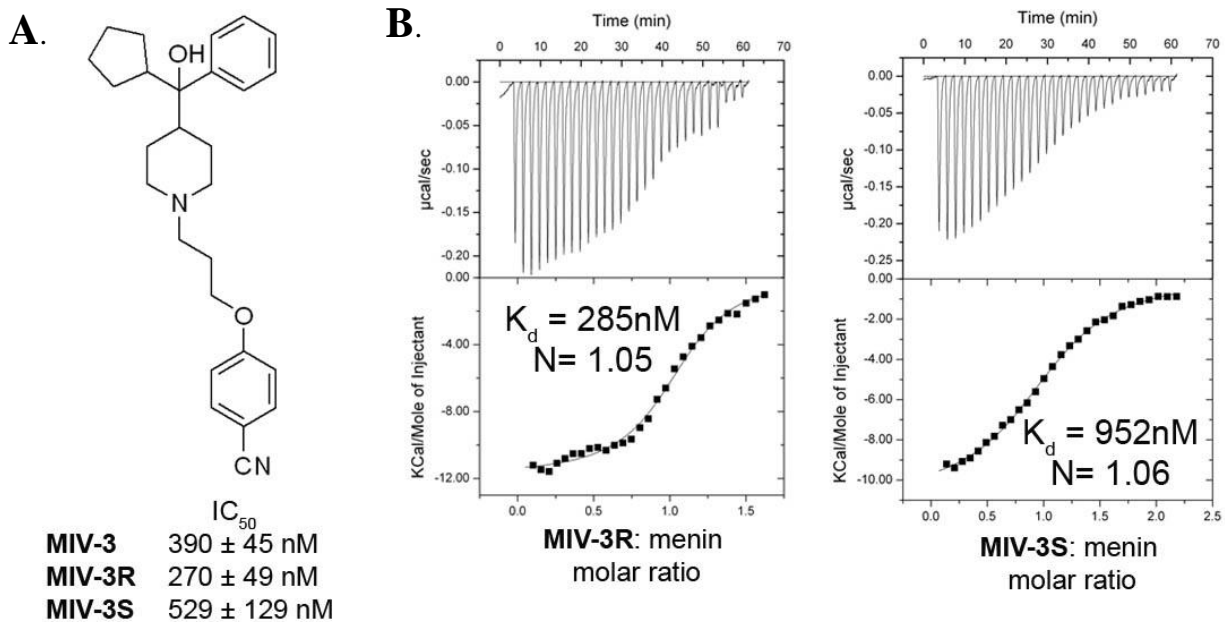
The potency for **MIV-2** was unaffected with an IC<sub>50</sub> = 10.8 μM. Further loss of the hydroxyl substituent from the quaternary carbon led to more than 20-fold decrease in potency (**3**). Additionally, removal of the nitrile group

resulted in an IC<sub>50</sub> = >250 μM in compound **4**. This analysis helped to identify the minimum scaffold that is needed for the potency of the initial HTS hit.

Compound	X	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	IC <sub>50</sub> (μM)
<b>1</b> ( <b>MIV-1</b> )	-OH			-CN	-OH	12.8 ± 1.6
<b>2</b> ( <b>MIV-2</b> )	-OH			-CN	-H	10.8 ± 0.2
<b>3</b>	-H			-CN	-H	205 ± 63
<b>4</b>	-OH			-H	-H	>250
<b>5</b> ( <b>MIV-nc</b> )	-OH		-H	-CN	-H	234 ± 22

**Table 4.1 SAR and activity of MIV analogues.**

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided.



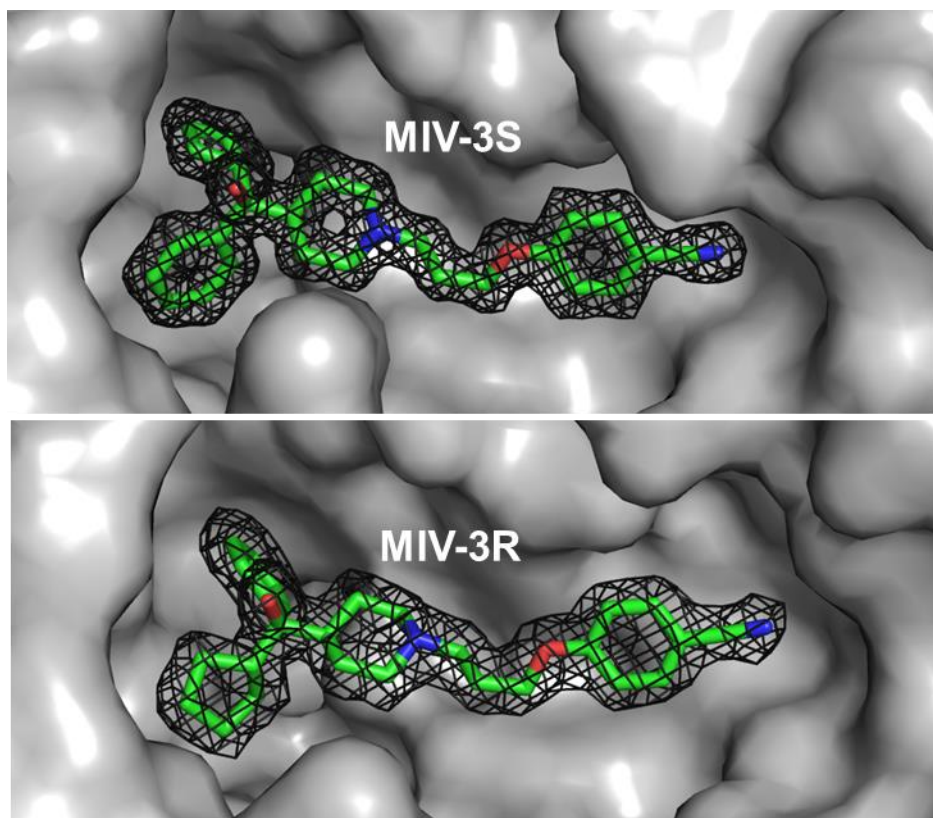
**Figure 4.2 Determining binding affinity of each MIV-3 enantiomer.**

Figure adapted from (He 2014).<sup>73</sup> **A.** Structure and IC<sub>50</sub> values were measured by fluorescence polarization assay for **MIV-3** and each enantiomer. **B.** ITC experiments demonstrating the direct binding of each enantiomer of **MIV-3** to menin.

## C.2 Structure based-design of MIV analogs

**MIV-2** was then used as an initial lead inhibitor for further medicinal chemistry development to increase inhibitory activity of these compounds for blocking the menin-MLL interaction. We first explored substitutions of a phenyl ring with various hydrophobic substituents listed in (**Table 4.2**). Replacement of the phenyl ring with hydrogen or a methyl led to a ~20-fold loss in potency; suggesting that larger bulky hydrophobic substituents at this site are most likely needed to maintain the inhibitory potency of these inhibitors. Therefore, we explored a few larger alkyl groups such as n-butyl (**7**) and iso-propyl (**9**) which resulted in  $IC_{50} = 15.3 \mu\text{M}$  and  $IC_{50} = 4.1 \mu\text{M}$  respectively. We performed systematical analysis of cycloalkyl substituents ranging from cyclopropyl to cyclohexyl. Cyclopentyl was the most potent saturated carbocycle we explored with an  $IC_{50} = 390 \text{ nM}$  for the racemic mixture (**Figure 4.2**). Further increase in the ring size resulted in decreased activity as demonstrated with the cyclohexyl analogue **12**, which resulted in about 4-fold decrease in *in vitro* potency. We concluded that cyclohexyl was the optimal substitution at this position based on the SAR performed at this site.

We further assessed the individual enantiomers of **MIV-3**, and **MIV-3R** has an  $IC_{50} = 270 \text{ nM}$ , while **MIV-3S** has an  $IC_{50} = 529 \text{ nM}$  (**Figure 4.2**). We then explored **MIV-3R** and **MIV-3S** by ITC with  $K_d = 285 \text{ nM}$  and  $K_d = 952 \text{ nM}$  respectively (**Figure 4.2**). **MIV-3R** demonstrates over 40-fold improvement compared to the initial HTS hit **MIV-1**.



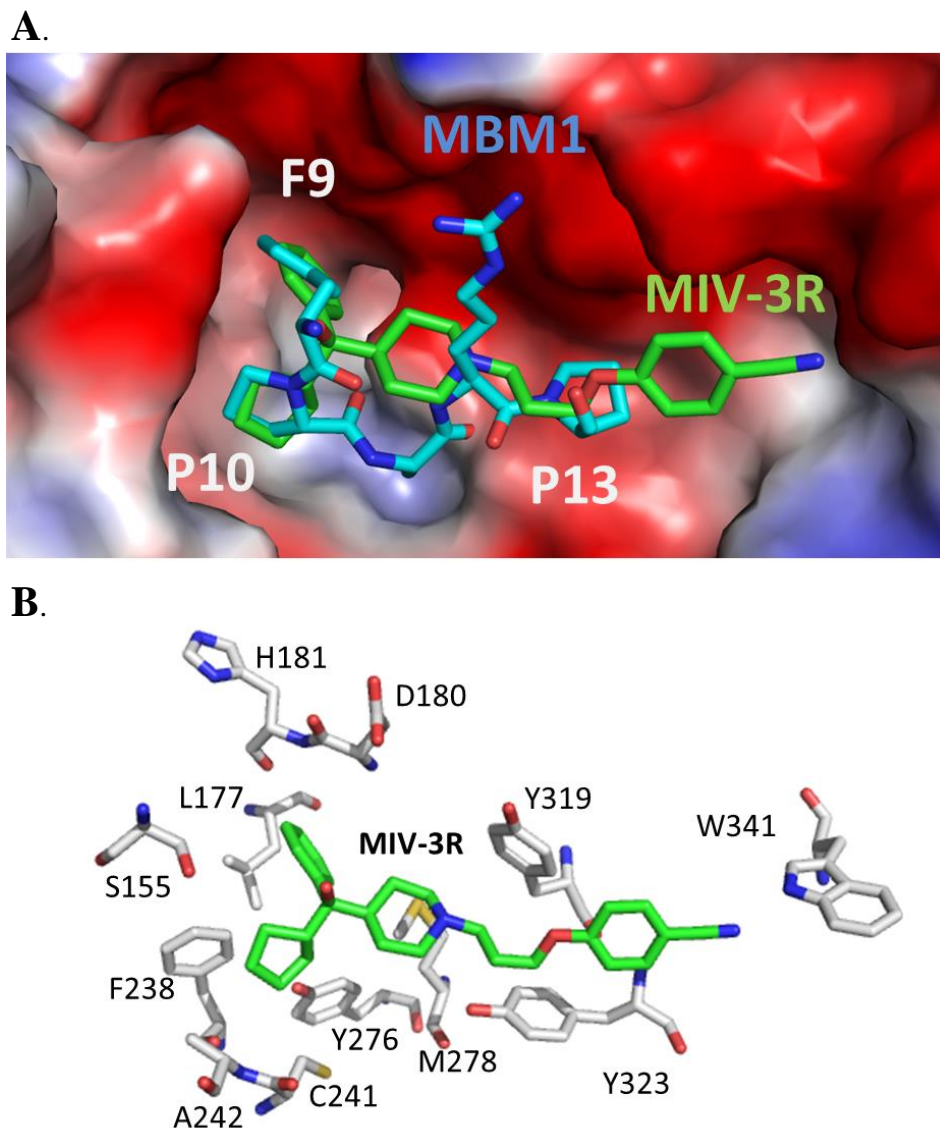
**Figure 4.3 MIV-3 enantiomers bind to menin in the MLL binding site.** Figure adapted from (He 2014).<sup>73</sup> Crystal structures of menin in complex with each enantiomer of **MIV-3**.  $2F_o - F_c$  electron density maps for ligands contoured at the 1 $\sigma$  level. Menin is shown as a surface representation.

### C.3 Hydroxymethylpiperidine Inhibitors Mimic the Most Critical Interactions of MLL with Menin.

To gain a complete understanding of the differences in the binding affinity of the **MIV-3** enantiomers we performed structural studies of each enantiomer bound to menin. Both enantiomers bind in a similar extended fashion (**Figure 4.3**). The majority of the contacts of the **MIV-3** with menin are hydrophobic in nature with only one direct hydrogen bond with the nitrile and the indole nitrogen of Trp341 of menin (**Figure 4.4**). There is a water-mediated hydrogen bond between the hydroxyl group of **MIV-3** and the Asp190 side chain on menin. Loss of either hydrogen bond resulted in about 20-fold decrease in inhibitory activity as shown by analogues **3** and **4** demonstrating the importance of this hydrogen bond network (**Table 4.1**).

The phenyl ring of **MIV-3R** fits into the hydrophobic cavity occupied by F9 of MBM1 formed by side chains of residues Leu177, Phe238, Ser155 and Met278 (**Figure 4.4 B**). The

cyclopentyl ring occupies the P10 region of MBM1 and is in close contacts with side chains of Ser155, Phe238, Ala242, Cys241, and Tyr276 (**Figure 4.4**). The piperidine ring and benzonitrile in both enantiomers have a similar binding mode with menin. The alkoxy region overlaps with the region occupied by P13 of MBM1 that is composed of Tyr319 and Tyr323 on menin and forms hydrophobic interactions. The benzonitrile reaches beyond the P13 region and forms a hydrogen bond with the side chain of Trp341 (**Figure 4.4**). Interestingly, this area of the binding site on menin is not engaged by MBM1 of MLL but forms important interactions with the compounds to maintain high potency.



**Figure 4.4 MIV-3R binding mode to menin.**

Figure adapted from (He 2014).<sup>73</sup> **A.** Superposition of the most important residues of MLL for binding to menin (cyan carbon atoms) and **MIV-3R** (green carbon atoms). Menin is displayed as a surface representation with electrostatics displayed as negative (red) and positive (blue). **B.** Molecular contacts involved in the **MIV-3R** binding interaction with menin are highlighted.

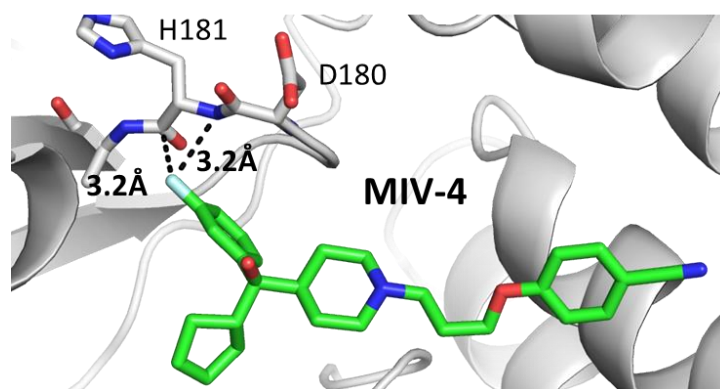
We have previously identified the three most critical residues for the MBM1 peptide to bind with high affinity to menin Phe9, Pro10, and Pro13.<sup>59</sup> Superposition of the MBM1 peptide with the **MIV-3R** inhibitor reveals a close overlay of these residues with the small molecule inhibitors of the hydroxymethylpiperidine class. The phenyl ring of **MIV-3R** overlaps with Phe9 in MBM1 and the cyclopentyl substituent adopts a comparable conformation as Pro10 in MBM1

(**Figure 4.4 A**). The alkoxy region in **MIV-3R** mimics the contacts of Pro13 with menin, while the benzonitrile extends outside of the P13 pocket and makes additional hydrogen bond interactions with W341 menin. The relatively strong potency of these inhibitors is derived from the close mimicking of the important interactions of MLL with menin specifically mimicking residues Phe9, Pro10, Pro13 and making additional contacts by extending outside of the P13 pocket.

#### **C.4 Design of orthogonal multipolar interaction of MIV analog with the menin backbone**

The **MIV-3** crystal structure allowed us to use structure based-design to help rationalize exploratory modification and generate a well-informed SAR. The class of thienopyrimidine analogs and the **MI-2-2** crystal structure demonstrated the importance of the trifluoro modification.<sup>66</sup> One of the fluorine atoms in **MI-2-2** approaches the protein backbone of His181 and is in close proximity to the amide peptide bond.<sup>1,66</sup> Characterization of this region of **MI-2-2** class of compounds exposed the importance of this interaction as replacement of all of the fluorine atoms to hydrogen atoms results in over a 25-fold decrease in *in vitro* potency.<sup>1,68</sup> To investigate if a fluorine-backbone interaction would be favorable with this scaffold we generated **MIV-4** by replacing the phenyl ring with 3-fluorobenzene modification (**Table 4.2**). This modification improved potency by 2-fold (**Table 4.2**) compared to **MIV-3**. The **MIV-4** menin structural analysis was performed to validate the binding mode (**Figure 4.5**). The structure confirmed the presence of the fluorine-backbone interaction with the carbonyl carbon of His181 within a distance of 3.2 Å. In addition we explored a (**15**) 3-chlorobenzene and (**14**) 3,5-fluorobenzene both of which resulted in decreased activity (**Figure 4.5**). Therefore, the 3-fluorobenzene substituent represents the optimal substituent identified through our SAR efforts at this position.



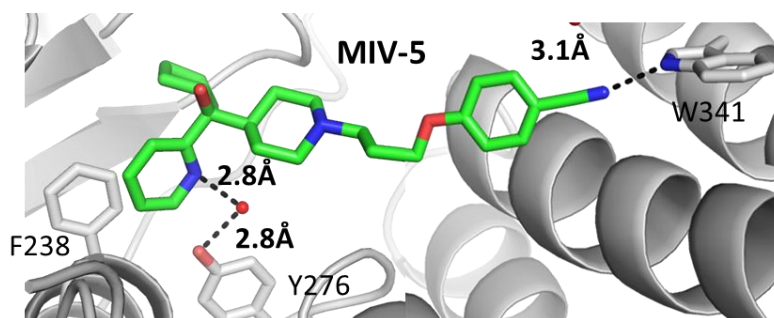


**Figure 4.5 Crystal structure of MIV-4 bound to menin.** Figure adapted from (He 2014).<sup>73</sup> Highlighting the fluorine of **MIV-4** orthogonal multipolar interaction with the menin backbone. Short distances to the protein backbone are shown as dashed lines.

### C.5 Increasing the polarity of the MIV class of menin-MLL inhibitors

The R<sub>2</sub> substituent that occupies the P10 pocket was also investigated to identify an optimal substituent. The cyclopentyl ring is very lipophilic with a clogP = 5.6 for **MIV-3**. We wanted to explore modifications at this position and **MIV-3S** only has 2-fold decrease in potency compared to **MIV-3R**. Therefore, to facilitate medicinal chemistry exploration **MIV-3S** was used as a starting molecule to design an analog. Both the **MIV-3S** co-crystal structure with menin and the menin apo structure showed a water molecule near the P10 pocket within hydrogen bond distance of Tyr276. Based on the **MIV-3S** crystal structure we rationalized that correct positioning of the nitrogen of a pyridine should allow interactions with this water molecule resulting in water mediated hydrogen bond with Tyr276. The introduction of this nitrogen atom should also help increase the hydrophilicity of the inhibitor (clogP = 3.7). We explored a pyridine (**MIV-5**, **16**) and thiazole (**17**) substituent at this R<sub>2</sub> position, both of which resulted in an increase in potency and improved polarity with a clogP of 3.7 and 3.5, respectively (**Table 4.2**). The **MIV-5** menin co-crystal structure confirmed the presence of this bridged hydrogen bond between pyridine the water molecule and Tyr276. The racemic mixture of **MIV-5** was used for crystallization and only the S enantiomer was seen as bound to menin (**Figure 4.6**). Additional separation of the enantiomers demonstrated that **MIV-5S** had about a 4-fold improved inhibitory activity (IC<sub>50</sub> = 195 nM) compared to the R enantiomer.

The optimal substituents identified for the R<sub>1</sub> and R<sub>2</sub> positions were then combined to generate a hybrid structure (**18**). This compound demonstrated about a 3-fold improvement in IC<sub>50</sub> value compared to **MIV-3R** (IC<sub>50</sub> = 90 nM of R enantiomer of **18**, **Table 4.2**). This compound is the most potent hydroxymethylpiperidine menin-MLL inhibitor.

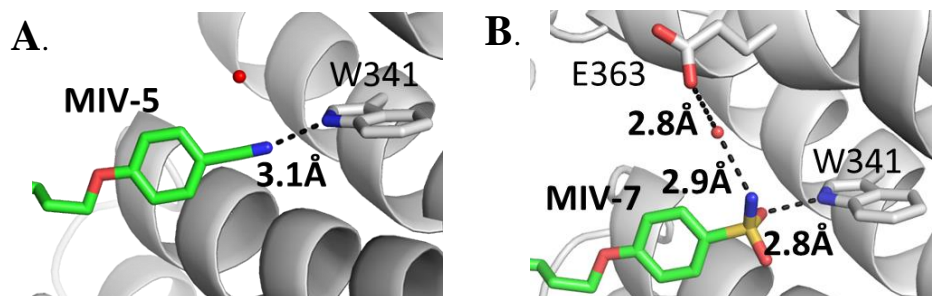


**Figure 4.6** Crystal structure of **MIV-5** bound to menin.

Figure adapted from (He 2014).<sup>73</sup> Highlighting the pyridine involved in the water mediated hydrogen bond with Tyr276. The nitrile is extending beyond the P13 pocket and involved in hydrogen bond interactions with Trp341.

## C.6 Identification of optimal substituents at the tail region

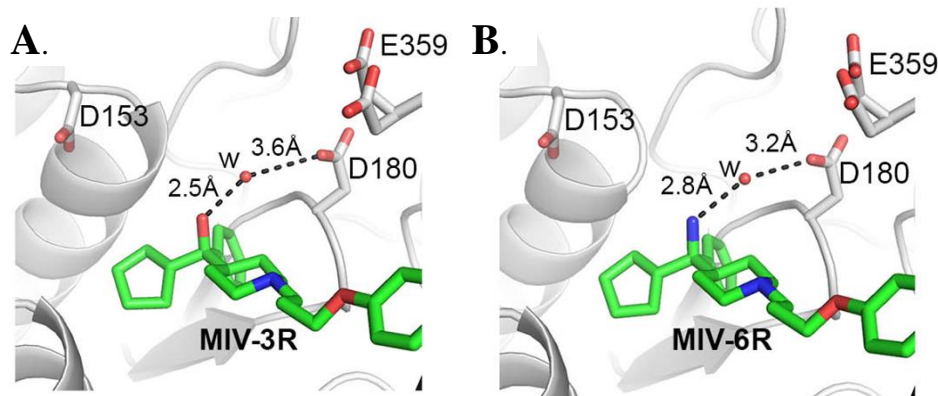
To further increase the polarity of these compounds we explored modifications of the tail region. The nitrile of the MIV class of compounds is involved in a hydrogen bond with Trp341 (**Figure 4.6**). In order to maintain this hydrogen bond while increasing the polarity of this tail group, replacement of the nitrile with sulfonamide (**MIV-7**, **20**) was explored. **MIV-7** increased the potency (IC<sub>50</sub> = 114nM) about 3-fold compared to **MIV-5** (**Table 4.2**). The **MIV-7** crystal structure confirms that the hydrogen bond is maintained with Trp341. Additionally, the sulfonamide introduces another hydrogen bond with the side chain of Glu363 (**Figure 4.6**) most likely leading to the improvement in potency. This SAR demonstrates the sulfonamide as a feasible substitution for the nitrile resulting in an increase in the hydrophilicity of these compounds (clogP = 3.7 for **MIV-5** and clogP = 2.5 for **MIV-7**).



**Figure 4.7 Crystal structure of MIV-5 and MIV-7 bound to menin.**  
 Figure adapted from (He 2014).<sup>73</sup> **A.** Highlighting the nitrile of **MIV-5** involved in hydrogen bond interactions with Trp341. **B.** **MIV-7** sulfonamide interaction with Trp341 and an additional water-mediated hydrogen bond with E363.

### C.7 Exploring the hydrogen bond network in the head region

The hydroxyl group in the crystal structure of **MIV-3** is solvent exposed and involved in water mediated hydrogen bond with Asp180 (**Figure 4.8**). Introduction of a positively charged amino group at this position should maintain this hydrogen bond while introducing longer range electrostatic interactions. The amino moiety was introduced to develop compound **MIV-6** and demonstrated 6-fold improvement in *in vitro* potency ( $IC_{50} = 67$  nM) compared to **MIV-3**. The enantiomers were isolated with the  $IC_{50} = 56$  nM of the R enantiomer being the more potent in this class of menin-MLL inhibitors. The co-crystal structure of **MIV-6R**-menin demonstrated the binding mode of this class of compounds is not affected by this modification (**Figure 4.8**). **MIV-6R** is the most potent analog reported with a  $K_d = 85$  nM (**Table 4.2**). The importance of the long range electrostatic interactions is evident in this analog and may more broadly provide an important method for chemical tools to achieve efficient binding for ligands inhibiting protein-protein interactions.



**Figure 4.8 MIV-3R and MIV-6R binding modes to menin are investigated.** Figure adapted from (He 2014).<sup>73</sup> The crystal structure of menin in complex with **A. MIV-3R** or **B. MIV-6R**. The polar interactions with the head group are shown. E359 is present in two conformations. Both compounds are involved in a water mediated hydrogen bond.

## D. Conclusions

Menin is an essential interaction partner for MLL fusion proteins providing an important target for small molecule inhibitor development.<sup>12,55,58</sup> This PPI is difficult to inhibit with small molecule inhibitors as a result of the large menin-MLL binding interface and the limitation that a small molecule inhibitors can only occupy a small region of this site.<sup>59</sup> Characterization of this interaction identified that the residues F9, P10, and P13 are the most critical for MBM1 binding to menin.<sup>59</sup> Therefore, small molecule inhibitors do not need to disrupt this entire menin-MLL interaction but specific targeting of these hot spot residues can provide disruption of the entire complex.

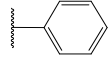

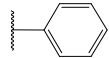

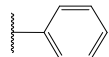

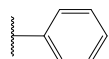

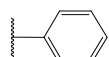

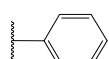

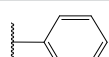

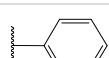

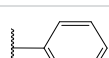

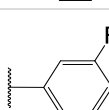
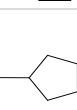
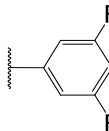
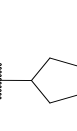
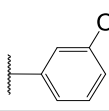
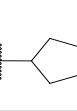
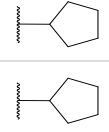
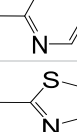
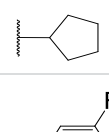
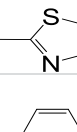
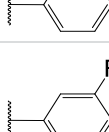
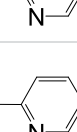
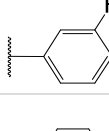
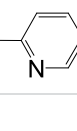
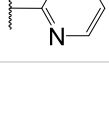
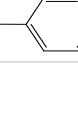
Extensive structure-based design was employed to optimize the **MIV-1** compound identified in the HTS screen into efficacious inhibitors of the menin-MLL complex. We identified the minimal pharmacophore from the **MIV-1** hit and then optimized the hydrophobic interactions of the phenyl groups by exploring several carbocycles. Cyclopentyl replacements of the phenyl ring generated MIV-3 and lead to a significant 30-fold improvement in the inhibitory activity of the menin-MLL interaction. Structural analysis of the **MIV-3** enantiomers in complex with menin helped rationalize modifications for improvements in potency. These small molecule inhibitors mimic the three most important interactions of MBM1 with menin (F9, P10, P13). This class of compounds are capable of extending beyond the P13 pocket with the

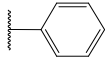

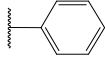

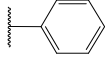

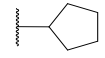

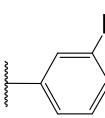
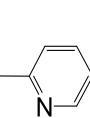
nitrile forming an additional hydrogen bond with the side chain of W341 and is very important for the binding affinity, reflected by an  $IC_{50} > 250 \mu M$  in the absence of this nitrile.

Development of the thienopyrimidine SAR, as described in **Chapter 2** and **Chapter 3**, helped to identify the importance of orthogonal multipolar interactions of the inhibitor-menin fluorine-backbone interactions.<sup>1,68</sup> We were able to apply this knowledge to the improvement of the methylpiperidine class of inhibitors by introducing a fluorine atom into the phenyl ring (**MIV-4**) and form orthogonal multipolar interactions with the backbone of His181. We improved solubility of these inhibitors with pyridine replacement of the phenyl (**MIV-5**) and sulfonamide replacement of the nitrile (**MIV-7**). The most substantial modification was seen with the replacement of the hydroxyl group for the amino which resulted in long-ranged electrostatic interactions with Asp180. **MIV-6R** demonstrates about a 230-fold improvement in inhibitory activity compared to **MIV-1**.

These compounds represent a promising scaffold for further optimization. All of the MLL hotspot residues are occupied by this class of inhibitors including the P10 pocket, which remains unoccupied in the thienopyrimidine scaffold of menin-MLL inhibitors.<sup>68</sup> However, the methylpiperidine class of inhibitors requires further optimization of potency in MLL leukemia cells and PK properties to explore applications in murine models of leukemia. This study provides a good basis for the further development of these inhibitors into valuable chemical tools or potential therapeutics.

**Table 4.2 SAR and activity of MIV analogues**

Compound	X	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	IC <sub>50</sub>
6	-OH			-CN	295 ± 50 μM
7	-OH			-CN	15.3 ± 3.8 μM
8	-H			-CN	11.2 ± 3.9 μM
9	-OH			-CN	4.1 ± 0.6 μM
10	-OH			-CN	4.0 ± 1.4 μM
11 (MIV-3)	-OH			-CN	390 ± 45 nM
11a (MIV-3R)	-OH			-CN	270 ± 49 nM
11b (MIV-3S)	-OH			-CN	529 ± 129 nM
12	-OH			-CN	1.7 ± 0.6 μM
13 (MIV-4)	-OH			-CN	1.7 ± 0.6 μM
14	-OH			-CN	436 ± 9 nM
15	-OH			-CN	226 ± 35 nM
16 (MIV-5)	-OH			-CN	302 ± 49 nM
17	-OH			-CN	242 ± 27 nM
18 (RS)	-OH			-CN	225 ± 28 nM
18R	-OH			-CN	90 ± 6 nM
18S	-OH			-CN	1.4 ± 0.2 μM

<b>19</b> <b>(MIV-6)</b>	-NH <sub>2</sub>			-CN	67 ± 10 nM
<b>19a</b> <b>(MIV-6S)</b>	-NH <sub>3</sub>			-CN	81 ± 5 nM
<b>19b</b> <b>(MIV-6R)</b>	-NH <sub>4</sub>			-CN	56 ± 3 nM
<b>20</b> <b>(MIV-7)</b>	-OH			-SO <sub>2</sub> NH <sub>2</sub>	114 ± 2 nM
<b>21</b>	-OH			-SO <sub>2</sub> NH <sub>2</sub>	148 ± 19 nM

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided.

## **Chapter 5: Development of covalent inhibitors of the menin-MLL interaction**

### **A. Abstract**

Rational design of covalent inhibitors has gained attention as recent examples of targeted covalent therapeutics have demonstrated clinical efficacy in oncology. Covalent inhibitors exhibit unique properties including increased residence time, require different PK profile compared to reversible inhibitors, and allow for the assessment of target engagement. There are no published examples of covalent inhibitors targeting protein-protein interactions. We have rationally designed covalent inhibitors targeting the menin-MLL interaction using the thienopyrimidine class of menin-MLL inhibitors as a starting pharmacophore. We employed structure-based design to develop covalent inhibitors of the menin-MLL interaction focusing on the indole nitrogen substituents to engage the sidechain of Cys329. Several reactive chemotypes were investigated and characterized to assess the rate of menin engagement and nonspecific reactivity rate with a biologically relevant thiol containing biomolecule, glutathione. In cellular studies, covalent inhibitors exhibited increased potency and improved selectivity index compared to the noncovalent analogs. Through the irreversible binding mechanism of covalent inhibitors, we were able to assess menin engagement *in vivo*. Overall, we have developed rationally designed covalent inhibitors of the protein-protein interaction between menin-MLL. This work provides a detailed approach for the characterization and assessment of covalent inhibitors targeting protein-protein interactions.



## **B. Background**

### **B.1 Advantages and disadvantages of covalent inhibitors**

Traditionally, small molecule inhibitor campaigns were not intentionally explored for covalent modifiers.<sup>114-116</sup> This is due to major concerns with covalent inhibitors surrounding the potential for off-target toxicity as a result of nonspecific reactivity with biomolecules or non-target proteins.<sup>117</sup> Additionally, activation of the immune system can occur through haptenization of proteins as a result of nonspecific macromolecular reactions.<sup>118</sup> These safety concerns have led to a disinclination towards the development of covalent small molecule therapeutics.<sup>119</sup> However recently covalent inhibitors targeting kinases have been successful in clinical trials<sup>120</sup> and FDA approved targeted therapies have shown efficacy with limited toxicity shedding new light to this approach.<sup>121</sup>

If there are so many concerns with the safety and toxicity of covalent inhibitors why would anyone rationally design a covalent drug? Covalent inhibitors have unique features including long residence time, do not require a long half-life, and allow for the assessment of target engagement.<sup>119,122,123 124, 125</sup> While safety concerns with covalent inhibitors needs to be investigated,<sup>126</sup> these inhibitors offer a unique approach to inhibition of the target protein and may provide a therapeutic benefit.

### **B.2 Design of covalent inhibitors of PPIs**

Cysteine residues are frequently targeted for covalent modification because they are the most intrinsically nucleophilic residue.<sup>127</sup> It is difficult to target residues other than cysteine because less nucleophilic residues require stronger reactive groups for covalent modification and may increase the potential for off-target reactivity. Activated cysteine residues in proteases are more reactive therefore the small molecule reactive groups can be more weakly electrophilic. Cysteine residues that are not activated, such as the cysteines on menin, require a different strategy to develop selective covalent molecules.

Cysteine reactive moieties that may be viable candidates for applications in humans are being assessed.<sup>128,129</sup> Michael acceptors are typically used as cysteine reactive moieties where the  $\beta$ -carbon is an electrophile. The electrophilicity or reactivity of these groups can be tuned through steric and inductive effects by modifications of the small molecule substituents.<sup>128</sup>

Cysteine reactive moieties such as acrylamides, vinylsulfonamides, propynamides and halomethylketones can have low rates of non-specific reactivity with free thiols, or intrinsic reactivity. Clinically approved rationally designed covalent molecules that target non-active site cysteine residues in kinases contain acrylamide reactive groups (Afatinib, Ibrutinib).<sup>130,131</sup> Whether other reactive moieties that target cysteine may be viable candidates for drug discovery applications and have good safety profile is currently being explored.<sup>128</sup>

### **B.3 Biochemical characterization of covalent inhibitors**

Covalent inhibitors bind irreversibly to the protein target; therefore, are able to achieve a long residence time.<sup>122</sup> Once the target is inhibited by the covalent inhibitor the efficacy is maintained until physiologically relevant levels of the target protein is produced through *de novo* synthesis. Therefore, the efficacy of covalent inhibitors is dependent on the rate of re-synthesis of the target protein as long as the drug dosage allows for engagement of the target protein. The rate of protein re-synthesis is an important parameter to quantify because the duration of the biological effect is only a result of the protein-inhibitor complex.<sup>122,132</sup> This unique mechanism of action is particularly attractive for proteins that have a long half-life or need to be inhibited to near entirety to demonstrate efficacy. This irreversible binding mechanism of these inhibitors can also be a major concern if other proteins or biomolecules are covalently modified.

Covalent inhibitors require a different pharmacokinetic profile as compared to reversible inhibitors.<sup>119</sup> Irreversible covalent inhibitors do not reach a dynamic equilibrium; once they react there is no dissociation. Therefore, a long half-life of covalent inhibitors *in vivo* may not be required.<sup>119</sup> Covalent inhibitors with a long half-life could contribute to increased toxicity as a result of increasing the amount to reactivity to other biomolecules. Covalent inhibitors should be designed to rapidly covalently modify the target protein and then become quickly metabolized and excreted to reduce the potential for non-specific reactivity.<sup>119 133</sup>

Reactive moieties that are used for the covalent modification of the inhibitor to the target protein should be chosen carefully. These covalent chemotypes should have low intrinsic rates of reactivity to reduce the potential for off-target reactivity.<sup>128</sup> The intrinsic rate of reactivity of covalent inhibitors is assessed by the non-specific reactivity with biologically relevant nucleophiles such as glutathione.<sup>128,129,134</sup> Glutathione (GSH) is present ubiquitously at high

concentrations in the cell. Studies have demonstrated that a high rate of covalent reactivity with GSH correlates with off-target binding burden in hepatocytes and hepatotoxicity.<sup>126,135,136</sup> Additionally, off-target reactivity will diminish the concentration of functional covalent inhibitors to covalently engage the target protein. Modifying the electrophilicity of the Michaelis acceptor, or electrophilic tuning, can help decrease the intrinsic rate of reactivity of these inhibitors minimizing non-specific reactivity.

Low rates of reactivity can help reduce off-target covalent modifications but the molecules must still react with the target protein. Covalent small molecule inhibitors with low intrinsic rates of reactivity rely on the reversible interactions to drive the specificity of these inhibitors with the target protein. This allows for selective covalent modification only when the reactive moiety is positioned in close proximity to the targeted residue. One of the key properties of covalent inhibitors is the rate of reactivity with the protein target (target engagement). Importantly, the rate of target engagement for covalent inhibitors is an important metric to perform direct comparison between covalent analogs.

The irreversible binding of these inhibitors allows for the assessment of target engagement *in situ*.<sup>123 124, 125</sup> Determining target engagement enables the correlation of inhibited target protein to other inhibitor induced phenotypic effects e.g. growth inhibition. Importantly identifying off-target proteins that are covalently modified can yield insight into the potential toxicity. These off-target interactions may potentially be ameliorated through modifications in the noncovalent binding interactions but also tuning the electrophilicity of the Michaelis acceptor.

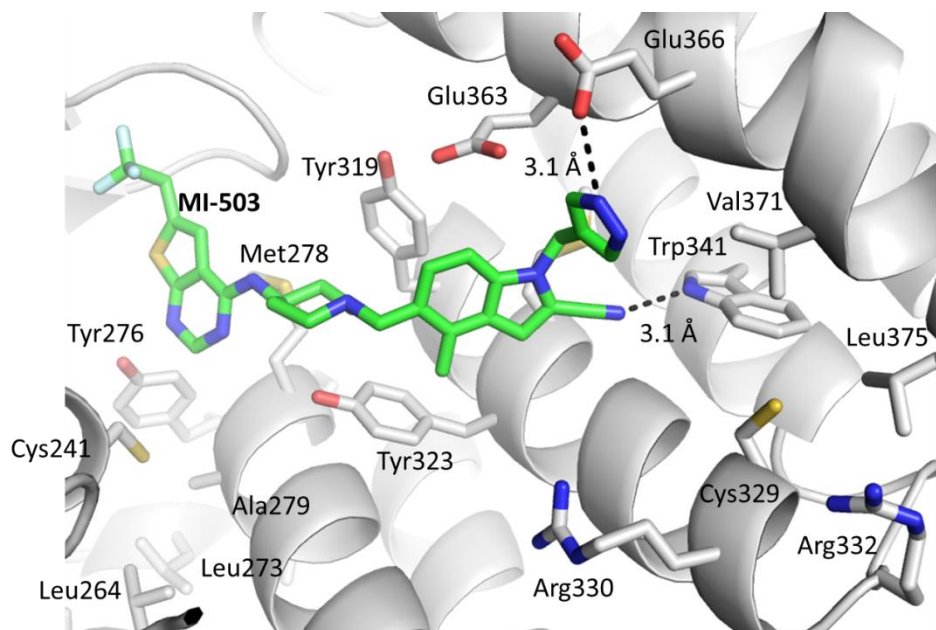
Structure-based design of small molecules can greatly facilitate design rationale for chemical modifications to form favorable interactions that improve potency.<sup>137,138</sup> This structural analysis can be used to identify a cysteine that is accessible to target with a small molecule covalent modifier.<sup>134,137,139</sup> Starting with a non-covalent inhibitor and targeting a cysteine residue near the small molecule binding pocket is an effective strategy to develop covalent inhibitors.<sup>137</sup> This strategy allows for a more heavy reliance on the non-covalent interactions to drive the specificity of the covalent linkage.<sup>119</sup> Therefore the electrophile must be in correct orientation and close proximity to the target nucleophile for the formation of the covalent linkage. Structure based-design can greatly facilitate the rational design of noncovalent inhibitors into covalent inhibitors through targeting a residue adjacent to the inhibitor binding pocket.

My contribution to this project was the rational design of covalent menin-MLL inhibitors performed by converting the reversible thienopyrimidine class of menin-MLL inhibitors into irreversible compounds by exploring several different reactive moieties. I also performed crystallization of the small molecules in complex with menin, solved the crystal structures of the complexes and performed structural analysis of the complexes. I characterized the covalent inhibitors in a biochemical FP assay and using biophysical methods to compare the potency of covalent thienopyrimidine analogs, rate of menin engagement, and general reactivity with glutathione. Finally, I determined the rate of menin engagement with covalent thienopyrimidine menin-MLL inhibitors in MLL leukemia cell lines and in mouse bone marrow samples isolated from mice treated with these compounds.

## **C. Results**

### **C.1 Investigation of accessible cysteine residues near the small molecule binding pocket of menin**

We applied structure-based drug design to develop potent inhibitors of the menin-MLL interaction. Through the design process of the reversible inhibitors, we identified regions of the menin binding pocket that could be explored to attain more potent inhibitors. Targeting Cys329 was selected for covalent modification using the thienopyrimidine class of reversible inhibitors as a starting pharmacophore. The thienopyrimidine class was chosen based on the potential for modifications of the indole nitrogen to design a linker and reach Cys329 (**Figure 5.1**).

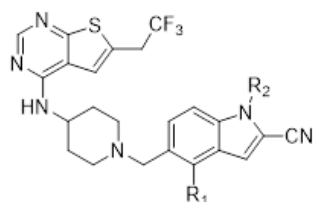


**Figure 5.1 Crystal structure of MI-503 bound to menin.**

Highlighting the Cys329 local landscape. Modifications of the indole nitrogen can be made to extend to adjacent areas of the inhibitor binding pocket as evident from **MI-503** binding mode

## C.2 Structure based-design of covalent thienopyrimidine class of menin-MLL inhibitors

**MI-503** of the thienopyrimidine class was used as a starting compound to explore longer substitutions of the indole nitrogen to reach Cys329. Cys329 is surrounded on one side by a hydrophobic patch formed by Val371 and Leu375 and on the adjacent side two arginine residues Arg330 and Arg332. Arg330 is involved in a salt bridge with Glu288. An overview analysis of menin co-crystal structures revealed the Arg330 side chain exhibits a stable orientation as displayed in **Figure 5.1**. This suggested that might be able to exploit the surrounding landscape of Cys329 to make favorable noncovalent interactions that result in the correct geometric positioning of a covalent moiety to react with the side chain of Cys329.



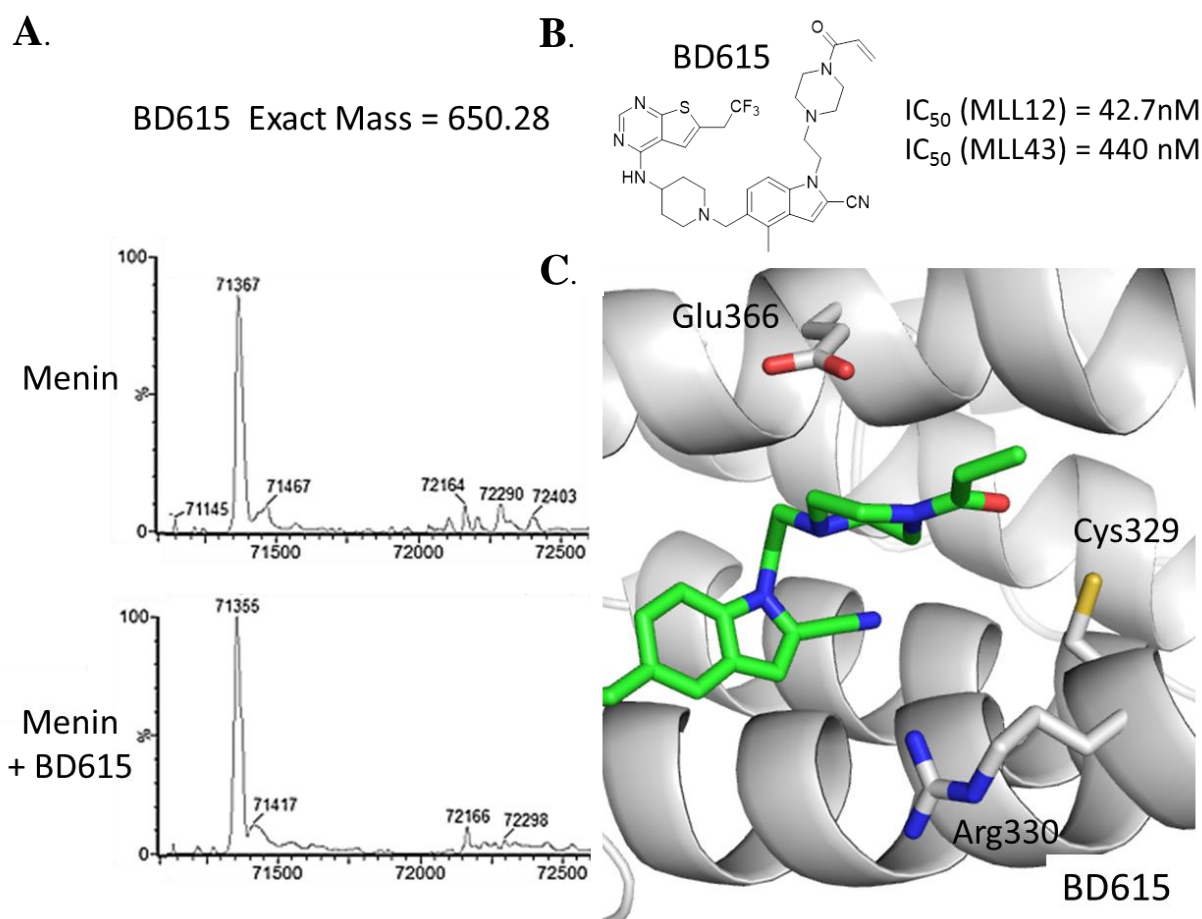
Compound	R <sub>1</sub>	R <sub>2</sub>	MLL12
			IC <sub>50</sub> (nM)
BD613	H		23.0 ± 8.5
BD615	H		42.7 ± 20.2
BD712	Me		42 ± 31.1
BD688	Me		21.4 ± 6.9

**Table 5.1 SAR and activities of covalent analogues of the thienopyrimidine class of menin-MLL inhibitors.**

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided.

We removed the methylpyrazole of **MI-503** and replaced it with a N-ethylpiperazine (**BD613**) with less than a 2-fold decrease in potency as determined by the MLL12 FP assay (**Table 5.1**). This N-ethylpiperazine linker was utilized to explore the different covalent moieties. We first explored acrylamide reactive moieties which were chosen as a result of recent FDA approval of irreversible therapeutics harboring acrylamide moieties that have demonstrated safety and clinical utility.<sup>130,131</sup> The first acrylamide menin-MLL inhibitor we explored was **BD615** which

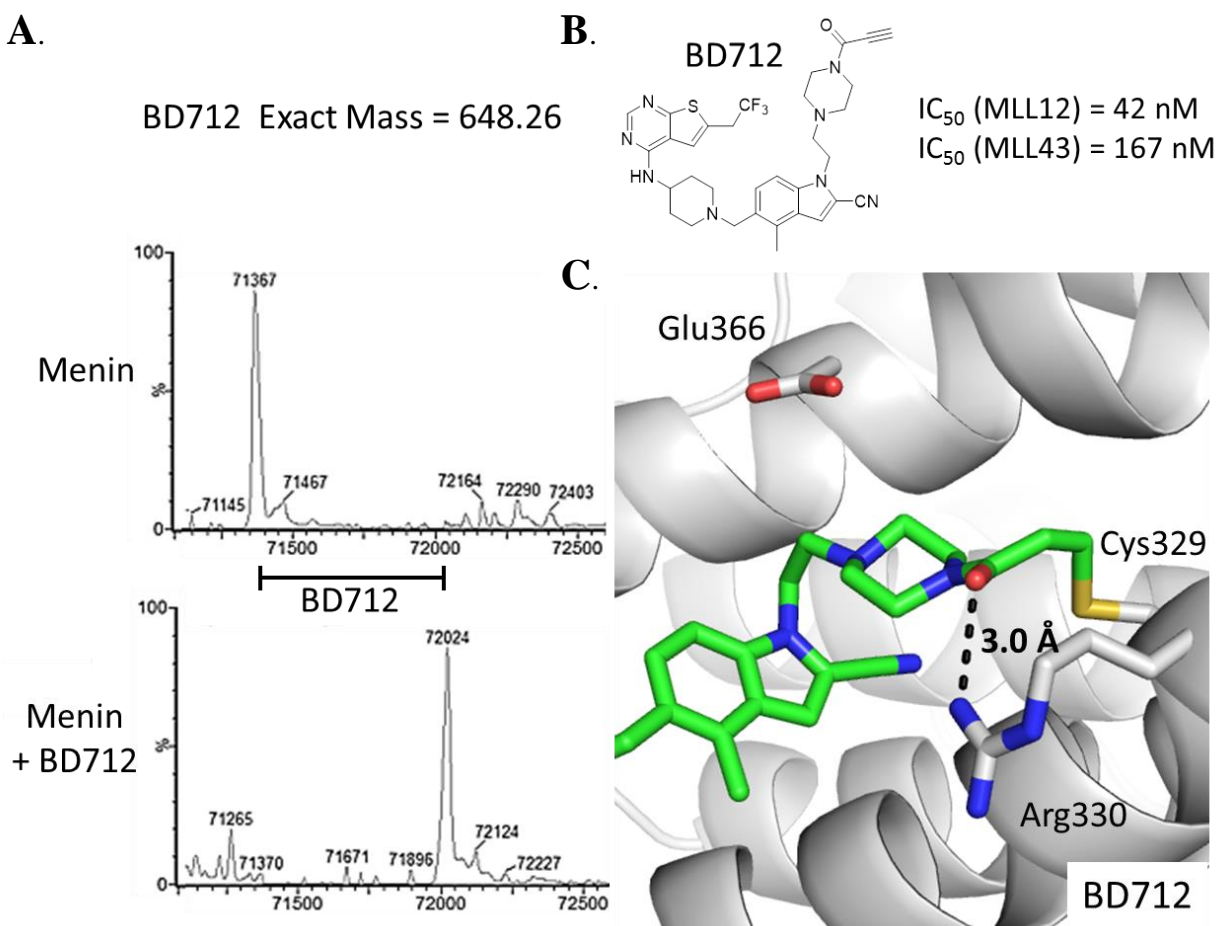
used the N-ethylpiperazine linker with an acrylamide reactive group. To verify a covalent adduct was formed between the inhibitor and menin, we used whole protein molecular weight (MW) determination by mass spectrometry (MS) (see **Methods section: LC-MS determination of covalent adducts to menin**). We found that menin alone has a MW of 71.3 kDa and if **BD615** was covalently bound we would expect an addition mass of 650 Da but the MW corresponding to the covalent adduct with menin was not identified (**Figure 5.2 A**). We performed structural analysis of **BD615** in complex with menin and found it did not bind covalently to menin with the acrylamide pointing towards the solvent (**Figure 5.2 B**).



**Figure 5.2 Design of BD615 with an acrylamide reactive moiety.**

**A.** MS analysis of menin with in the presence and absence of **BD615**. **B.** Structure and activity of **BD615** in FP assay. **C.** Crystal structure of **BD615** in complex with menin.

We then explored a propynamide reactive moiety (**BD712**) using the same linker as in **BD615**. By MS, **BD712** was shown to form a covalent adduct to menin with an overnight incubation of a 2-fold excess **BD712** and all of the menin was fully reacted (**Figure 5.3 B**). The additional MW corresponds to one molecule of **BD712** attached to menin validating this reaction is selective to only cysteine. We determined the crystal structure of **BD712** in complex with menin to analyze the molecular interactions (**Figure 5.3 A**). We identified **BD712** forms a covalent bond with Cys329 with the propynamide oriented with the oxygen in close distance (3.0 Å) to Arg330 side chain. The Cys329 side chain in the **BD712** crystal structure is in a different orientation when compared to menin alone and menin in complex with inhibitors that do not interact with this Cys329 (**Figure 5.3**). Therefore, the cysteine side chain must adopt a different conformation to allow for the covalent reaction with **BD712**.

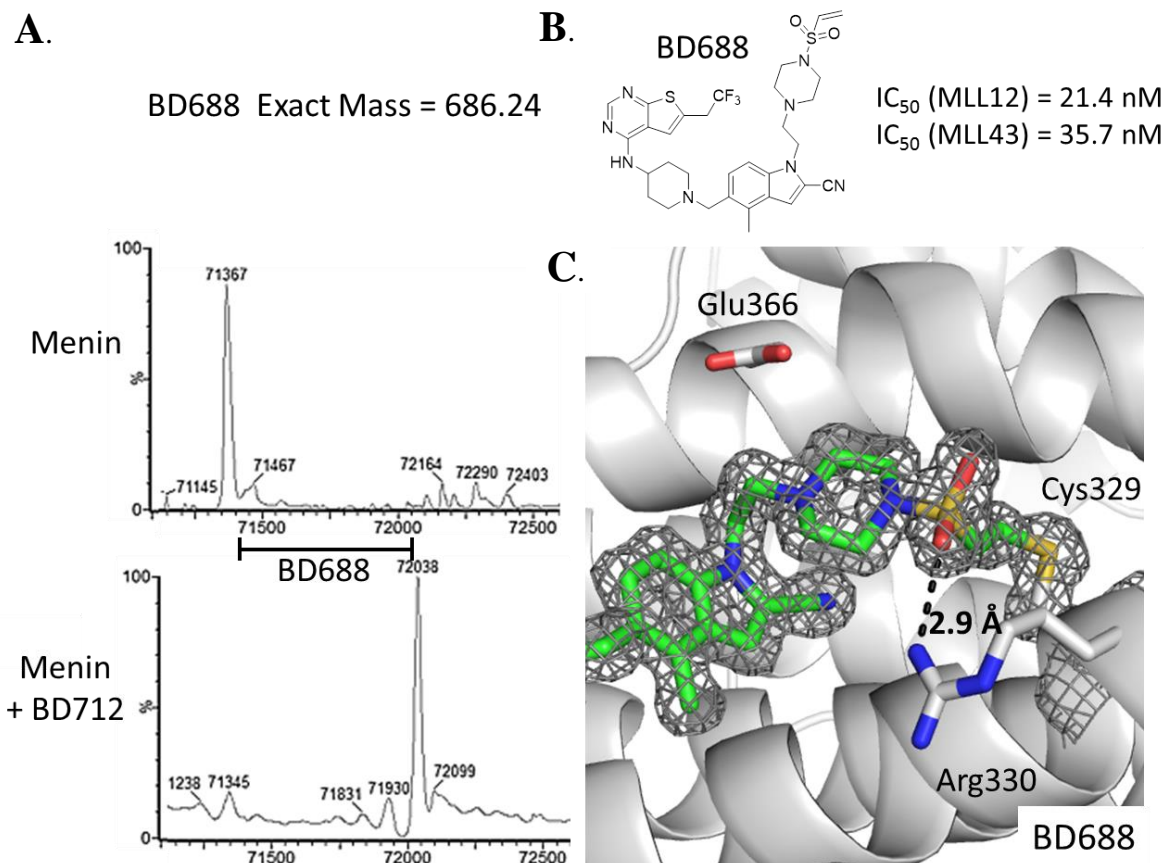


**Figure 5.3 Design of irreversible inhibitor BD712 with a propynamide reactive moiety.**  
**A.** MS analysis of menin with in the presence and absence of **BD712**. **B.** Structure and activity of **BD712** in FP assay. **C.** Crystal structure of **BD712** in complex with menin.



We then explored the vinylsulfonamide reactive moiety (**BD688**). **BD688** covalently modified menin with an overnight incubation at 2-fold molar excess of compound and identified all of the menin was covalently bound to **BD688** (**Figure 5.4 A**). We then determined the crystal structure of **BD688** bound to menin to validate the covalent linkage of Cys329 with **BD688** (**Figure 5.4 B**). There is a continuous electron density from **BD688** extending through Cys329 sidechain as shown in **Figure 5.4 B** indicating a covalent bond has been formed. The sulfonamide moiety is positioned with an oxygen atom of the sulfonamide involved in a direct hydrogen bond with Arg330 at 2.9 Å. The linker was long enough to allow for the vinylsulfonamide reactive moiety to covalently modify Cys329.

We have demonstrated the linker used in **BD615** is sufficient for the vinylsulfonamide (**BD688**) and propynamide (**BD712**) to form covalent bonds with Cys329. The linker was likely too short or there was not enough flexibility to allow for the acrylamide to correctly orient itself to form the covalent bond with Cys329 of menin. The acrylamide reactive moieties require further optimization to develop into covalent modifiers.



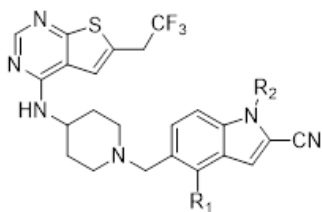
**Figure 5.4 Design of irreversible inhibitor BD688 with a vinylsulfonamide reactive moiety.**

**A.** MS analysis of menin in the presence and absence of **BD688**. **B.** Structure and activity of **BD688** in FP assay. **C.** Crystal structure of **BD688** in complex with menin.

### C.3 Characterizing the potency and rate of menin engagement of the covalent inhibitors.

We explored biochemical analysis of the activity of the covalent menin-MLL inhibitors with the MLL12 FP assay. However, the  $IC_{50}$  values in the MLL12 FP assay do not reflect the covalent nature of these analogs as a result of only an hour incubation of these inhibitors with menin. Therefore, we developed the MLL43 FP assay to more accurately characterize the potency of these inhibitors (**Table 5.2**). In the MLL43 assay both MLL binding motifs (MBM1 and MBM2) are present, which leads to a stronger binding interaction of MLL43 peptide with menin compared to the MLL12 (MBM1) peptide. This leads to higher  $IC_{50}$  values, and with longer incubation times more accurately reflects the relative potency of these analogs. We found that **BD615** is 3-fold and 12-fold weaker than the covalent inhibitors **BD712** and **BD688**, respectively (**Table 5.2**). However, we are unable to determine whether the modifications are increasing the noncovalent interactions or increasing the rate of covalent adducts formation.

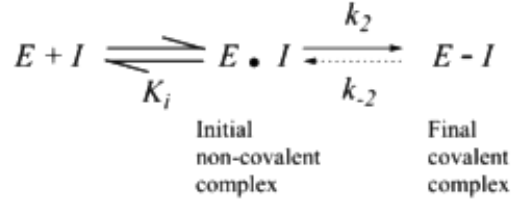
An important parameter to compare these analogs is the reactivity rates with the target protein or rate of menin engagement. We have used MS previously to determine whether the developed inhibitors bind covalently. To determine pseudo-first order reaction rates we changed the conditions to allow for saturating concentrations of compound (10-fold molar excess) over menin concentration. This will allow us to determine pseudo-first order reaction kinetics and characterize the rate of covalent adduct formation to menin (**Equation 5.1 and 5.2**). We used MS to explore the rate of reactivity and characterize the SAR for the covalent inhibitors with saturating concentrations of inhibitor relative to menin (**Figure 5.5**). We found **BD688** to engage menin 4-fold faster when compared to **BD712** (**Table 5.2**). We can utilize this method to quantify the rates of menin engagement to compare analogs.



Compound	R <sub>1</sub>	R <sub>2</sub>	MLL12	MLL43	Menin Engagement	
			IC <sub>50</sub> (nM)	IC <sub>50</sub> (nM)	k <sub>obs</sub> (s <sup>-1</sup> )	t <sub>1/2</sub> (min)
BD615	H		42.7 ± 20.2	440 ± 103	x	x
BD712	Me		42 ± 31.1	167 ± 50.3	1.36E-03	8.5
BD688	Me		21.4 ± 6.9	35.7 ± 10.7	5.26E-03	2.2

**Table 5.2 SAR and properties of analogues.**

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided. Pseudo-first order reaction kinetics (k<sub>obs</sub>) was used to calculate the rate of menin engagement and half-life (t<sub>1/2</sub>).



**Equation 5.1 Equilibrium equation.**<sup>119</sup>

Inhibition of menin occurs through a two-step mechanism of action. First the non-covalent interactions followed by the electrophile positioned in close proximity to a cysteine 329 residue for the covalent reaction to occur. The rate of covalent bond formation ( $k_2$ ) can be determined using pseudo-first order reaction kinetics.

$$\ln [A] = -kt + \ln [A]_0$$

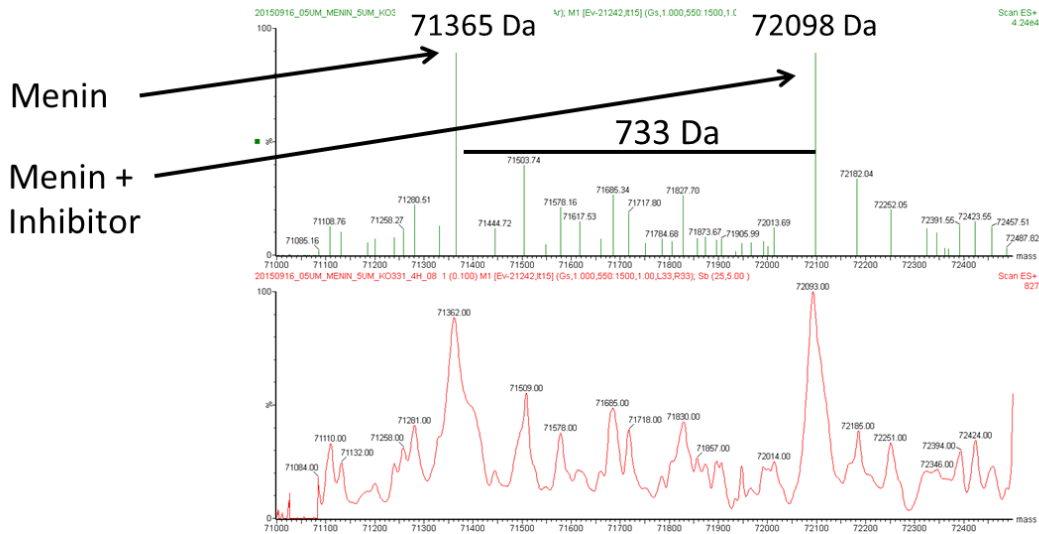
$$k = \frac{\ln [A]_0 - \ln [A]}{t}$$

$$k = \frac{\ln [100] - \ln [100 - (\% \text{ reacted})]}{t}$$

$$t_{1/2} = \frac{\ln(2)}{k}$$

**Equation 5.2 Pseudo-first order rate equation and half-life derivatization.**

Using a 10-fold molar excess of compound over menin protein we can determine the rate of covalent linkage ( $k_2$ ) by mass spectrometry to quantify the rate of covalent adduct formation (% reacted).



**Figure 5.5 Characterizing irreversible inhibitor rates of menin engagement.**

Ratio of menin to menin with a covalent adduct is quantified by whole protein molecular weight determination by mass spectrometry. Top is a centered spectra of the bottom spectra plotting relative intensity vs molecular weight. The centered spectrum is used to compare the intensity levels as a percent relative to the total intensity from menin and menin adduct intensities.

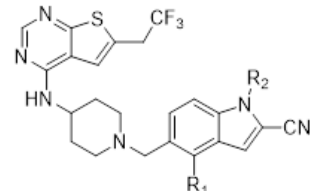
#### C.4. Structure-based design and characterization of acrylamide covalent inhibitors

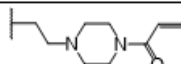
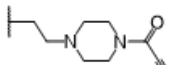
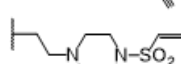
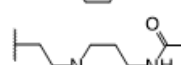
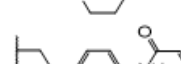
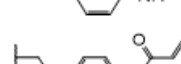
The vinylsulfonamide (**BD688**) and the propynamide (**BD712**) were found to bind menin covalently and had the same linker, however this was not an optimal linker for the acrylamide (**BD615**) to form a covalent adduct with menin. To investigate if the linker length was too short we extended the linker through an exocyclic acrylamide (**BD964**) to increase the flexibility near the reactive moiety. **BD964** was validated to bind covalently to menin by MS and structural characterization was performed to explore the molecular contacts (**Figure 5.6**). The oxygen of the amide does not point towards Arg330 as in **BD712**; instead the nitrogen of the amide is involved in a water-mediated hydrogen bond with Arg330. The cysteine sidechain is in a similar orientation compared to menin without inhibitors therefore the sidechain of Cys329 does not need to adapt a different rotameric state as in **BD712** to form a covalent linkage.

We next investigated the impact of a phenyl replacement (**BD1044**) of the piperidine in **BD964**. **BD1044** binds covalently to menin by MS with a rate of reactivity 5-fold faster than **BD964** (**Table 5.3**). To explore the orientation of the amide and the linker we performed structural studies which demonstrated the amide oxygen of **BD1044** points away from Arg330 similar to **BD964** (**Figure 5.7**).

We then pursued a pyridine replacement of the phenyl to form **BD1119**. **BD1119** exhibited an increased rate of menin engagement with over a 20-fold increase compared to phenyl ring (**BD1044**) and a half-life of less than a minute (**Table 5.3**). We performed structural analysis to explore the molecular contacts of this modification. The nitrogen of the pyridine was oriented towards Arg330 and involved in a water mediated hydrogen bond with Arg330 (**Figure 5.8**). This same water molecule is also in hydrogen bond distance to the nitrogen of the amide allowing for both these nitrogen atoms to coordinate the water mediated hydrogen bond with Arg330 of menin (**Figure 5.8**).

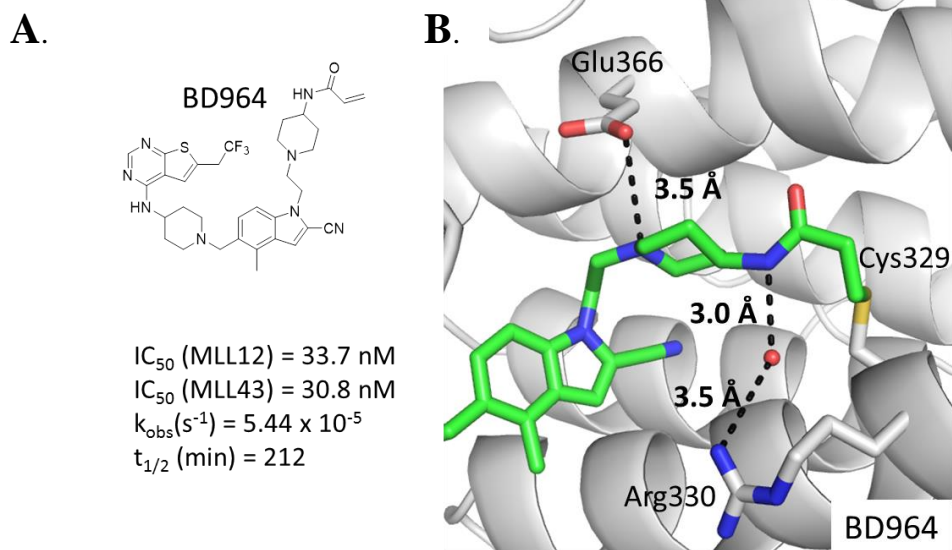
Acrylamide reactive moieties were explored in the context of the thienopyrimidine scaffold to develop covalent inhibitors of the menin-MLL interaction. These acrylamides moieties have very different rates of menin engagement with half-lives ranging from less than a minute to several hours. We have generated covalent inhibitors with improved rates of menin engagement (**BD1119**) but also wanted to explore the intrinsic reactivity of these compounds.



Compound	R <sub>1</sub>	R <sub>2</sub>	MLL12	MLL43	Menin Engagement	
			IC <sub>50</sub> (nM)	IC <sub>50</sub> (nM)	k <sub>obs</sub> (s <sup>-1</sup> )	t <sub>1/2</sub> (min)
BD615	H		42.7 ± 20.2	440 ± 103	x	x
BD712	Me		42 ± 31.1	167 ± 50.3	1.36E-03	8.5
BD688	Me		21.4 ± 6.9	35.7 ± 10.7	5.26E-03	2.2
BD964	Me		24.7 ± 9.5	30.8 ± 13.2	5.44E-05	212
BD1044	Me		17.8 ± 1.3	7.0 ± 2.3	1.02E-03	11.3
BD1119	Me		22.7 ± 4.5	34.4 ± 14.3	1.59E-02	0.7

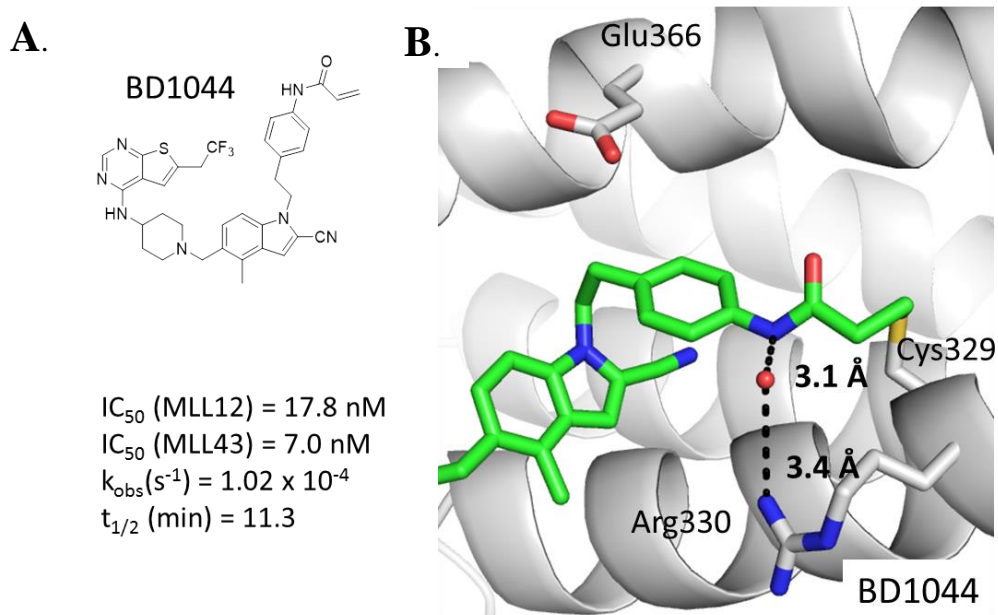
**Table 5.3 SAR and properties of analogues.**

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided. Pseudo-first order reaction kinetics (k<sub>obs</sub>) was used to calculate the rate of menin engagement and half-life (t<sub>1/2</sub>).



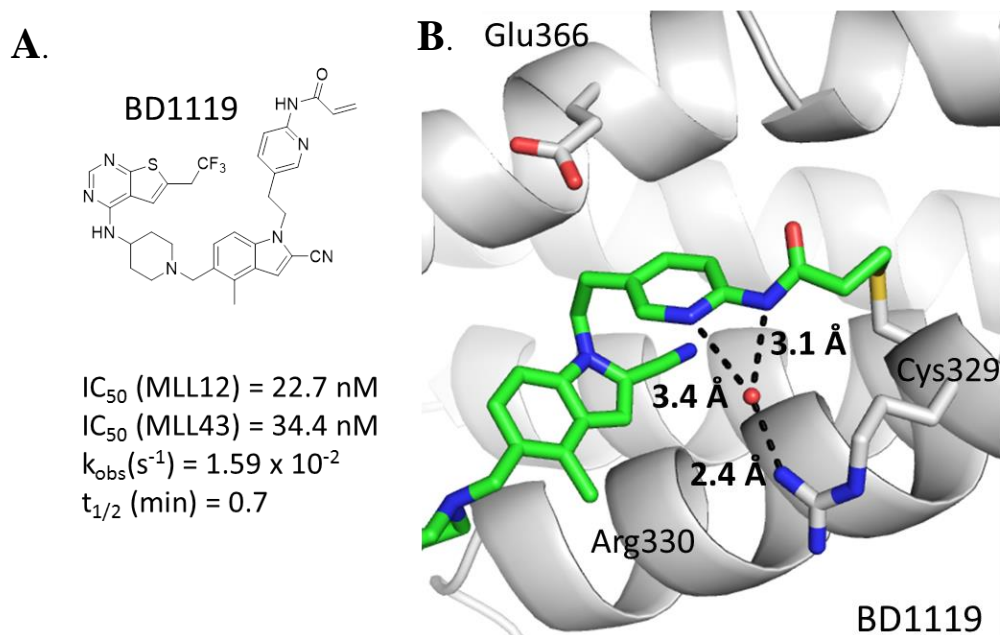
**Figure 5.6 Structure-based design of acrylamide covalent inhibitors.**

**A.** Structure of **BD964** with IC<sub>50</sub> values measured by fluorescence polarization assay, pseudo-first order reaction kinetics (k<sub>obs</sub>) was used to calculate the rate of menin engagement and half-life (t<sub>1/2</sub>). **B.** Extension of the reactive moiety through an exocyclic acrylamide forming **BD964**. The orientation of the chair is opposite of **BD688** with the nitrogen of the pyridine involved in a hydrogen bond with Glu366.



**Figure 5.7 Structure-based design of aromatic substituted acrylamides with menin.**

**A.** Structure of **BD1044** with  $IC_{50}$  values measured by fluorescence polarization assay, pseudo-first order reaction kinetics ( $k_{obs}$ ) was used to calculate the rate of menin engagement and half-life ( $t_{1/2}$ ). **B.** Crystal structure of **BD1044** bound to menin. The phenyl replacement of piperidine in **BD964** orients the amide oxygen away from Arg330.



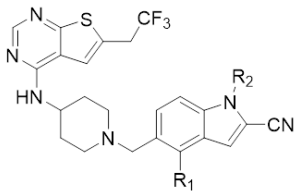
**Figure 5.8 Structure-based optimization of polar interactions of aromatic substituted acrylamides with menin.**

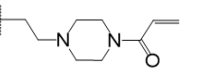
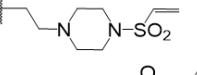
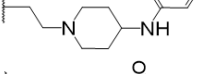
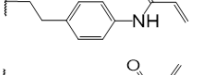
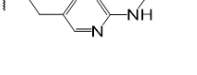
**A.** Structure of **BD1119** with  $IC_{50}$  values measured by fluorescence polarization assay, pseudo-first order reaction kinetics ( $k_{obs}$ ) was used to calculate the rate of menin engagement and half-life ( $t_{1/2}$ ). **B.** Crystal structure of **BD1119** bound to menin. Substitution of the phenyl with a pyridine allows both the nitrogen of the pyridine and the nitrogen of the amide forms a bridged hydrogen bond with a water molecule to the side chain of Arg330 on menin.

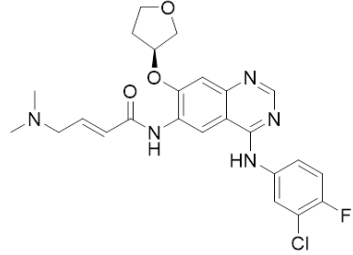


## C.5 Investigating the reactivity of the covalent menin-MLL inhibitors with glutathione.

The formation of covalent adducts between the menin inhibitors and small molecule nucleophiles can be used to quantify the intrinsic rate of reactivity to gain insight into the potential toxicity of these inhibitors. Covalent inhibitors should not exhibit significant rates of reactivity towards thiol containing molecules such as glutathione (GSH). High GSH reactivity has been correlated with hepatocellular toxicity,<sup>126,135,136</sup> and if the inhibitor is reacting quickly with GSH this will reduce the amount of functional inhibitor to engage the target protein. GSH is the primary intracellular nucleophile with important antioxidant functions and is found at high concentrations in the cell (up to 5mM). Therefore, we analyzed the reactivity of the menin-MLL inhibitors with GSH to assess intrinsic reactivity.



Compound	R <sub>1</sub>	R <sub>2</sub>	GSH Reactivity	
			k <sub>obs</sub> (s <sup>-1</sup> )	t <sub>1/2</sub> (min)
BD615	H		6.17E-04	18.7
BD688	Me		6.08E-04	19.0
BD964	Me		4.83E-05	239.0
BD1044	Me		1.50E-05	770.2
BD1119	Me		4.92E-04	23.5



Compound	k <sub>obs</sub> (s <sup>-1</sup> )	t <sub>1/2</sub> (min)
Afatinib	9.88E-04	11.7

**Table 5.4 Reactivity of inhibitors with glutathione.**

Rates of compound reactivity with GSH quantified by area under the curve for the LC trace at A280nm. Pseudo-first order reaction kinetic rates ( $k_{obs}$ ) were used to calculate the rate of loss of the parent compound with half-life ( $t_{1/2}$ ) values.

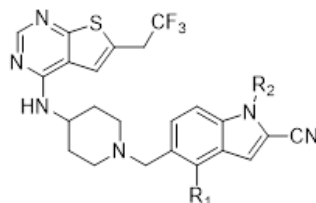
Reactive menin inhibitors were incubated at 25 $\mu$ M with 5mM GSH at various time points and the rate of parent compound loss as a result of adduct formation with GSH was analyzed by LC-MS to determine the pseudo-first order reaction rates. We used an FDA approved covalent inhibitor Afatinib as a benchmark to compare the intrinsic reactivity of the menin-MLL inhibitors. Afatinib is an acrylamide-based covalent inhibitor targeting human epidermal growth

factor receptor 2 (Her2) and epidermal growth factor receptor (EGFR) kinases. We sought to characterize the covalent inhibitor analogs of menin and compare to Afatinib (**Table 5.4**).

**BD615** was not able to covalently engage menin, but we identified that this compound is very reactive with GSH ( $t_{1/2} = 19$  min). **BD964** and **BD1044** had the slowest rates of GSH reactivity. **BD964** exhibited about a 3-fold faster rate of reactivity with GSH despite almost a 19-fold decreased rate in menin engagement compared to **BD1044**. While **BD688**, **BD615** and **BD1119** all displayed similar rates of GSH reactivity this was about 2-fold slower compared to Afatinib. This evidence demonstrates that the menin-MLL inhibitors we have developed are less reactive towards GSH when compared to an FDA approved and clinically efficacious therapeutic and therefore should be suitable to explore *in vivo*.

### **C.6 Activity of covalent inhibitors of the menin-MLL interaction in leukemic cells.**

We then explored the potency of the covalent menin-MLL inhibitors in MLL-AF9 leukemia cells. *Hoxa9/Meis1* (HM-2) transformed cells were used as a negative control. An MTT viability assay was used to determine the growth inhibition at 50% maximum ( $GI_{50}$ ) of MLL-AF9 cells and HM-2 cells with the selectivity index (SI) calculated as a ratio of  $GI_{50}$  values obtained in HM-2 cells to MLL-AF9 cells. We found that even though **BD615** does not covalently engage menin, it is only 3-fold weaker in  $GI_{50}$  compared to the covalent acrylamide **BD964** (**Table 5.5**). **BD964** had a 2-fold reduced  $GI_{50}$  compared to the vinylsulfonamide **BD688** and aromatic acrylamide **BD1044**. **BD1044** had the best selectivity index with a 43-fold more potent activity in MLL-AF9 cells compared to the HM-2 control cell line. The most potent analog was **BD1119** with a  $GI_{50} = 20$ nM in MLL-AF9 transformed murine BMCs (**Table 5.5**). **BD964** has a slow rate of menin engagement with a half-life of about 4h but the  $GI_{50}$  is similar to **BD688** and **BD1044** which have 100-fold and 20-fold faster rates of menin engagement respectively (**Table 5.5**). **BD1119** demonstrated a 16-fold faster rate of menin engagement and 30-fold faster rate of GSH reactivity compared to **BD1044** but only a 2-fold improvement in potency in MLL-AF9 cells (**Table 5.5**). Overall, the covalent inhibitors demonstrated more potent  $GI_{50}$  values in MLL-AF9 cells as compared to **BD615** an analog that does not engage menin. **BD1119** had fast rates of menin engagement with a half-life <1 minute and was the most potent analog in MLL-AF9 cells.



Compound	R <sub>1</sub>	R <sub>2</sub>	MLL43	Menin Engagement		GSH Reactivity		GI <sub>50</sub>		SI
			IC <sub>50</sub> (nM)	k <sub>obs</sub> (s <sup>-1</sup> )	t <sub>1/2</sub> (min)	k <sub>obs</sub> (s <sup>-1</sup> )	t <sub>1/2</sub> (min)	MLL-AF9 BMC	HM-2 BMC	
BD615	H		440 ± 103	x	x	6.17E-04	18.7	250nM	ND	ND
BD712	Me		167 ± 50.3	1.36E-03	8.5	ND	ND	150nM	ND	ND
BD688	Me		35.7 ± 10.7	5.26E-03	2.2	6.08E-04	19.0	50nM	1.5μM	30
BD964	Me		30.8 ± 13.2	5.44E-05	212	4.83E-05	239.0	90nM	2.5μM	28
BD1044	Me		7.0 ± 2.3	1.02E-03	11.3	1.50E-05	770.2	40nM	1.7μM	43
BD1119	Me		34.4 ± 14.3	1.59E-02	0.7	4.92E-04	23.5	20nM	0.5μM	25

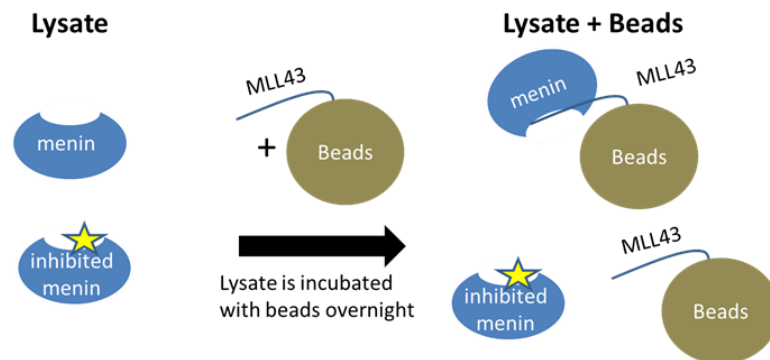
**Table 5.5 SAR and properties of analogues.**

IC<sub>50</sub> values were measured by fluorescence polarization assay, average values from two to three independent measurements ± SD are provided. Pseudo-first order reaction kinetics (k<sub>obs</sub>) was used to calculate the rate of menin engagement and half-life (t<sub>1/2</sub>). Growth inhibition (GI<sub>50</sub>) measured in the MTT cell viability assay in MLL-AF9 or Hoxa9/Meis1 (HM-2) transformed murine bone marrow cells after 7 days of treatment with compounds. Activity in HM-2 cells was measured for compounds with best activity in MLL-AF9 cells. SI, selectivity index calculated as a ratio of GI<sub>50</sub> values measured in HM-2 cells (control cell line) and MLL-AF9 transformed cells.

### C.7 Detection of menin engagement in MLL leukemic cells and bone marrow *in vivo*.

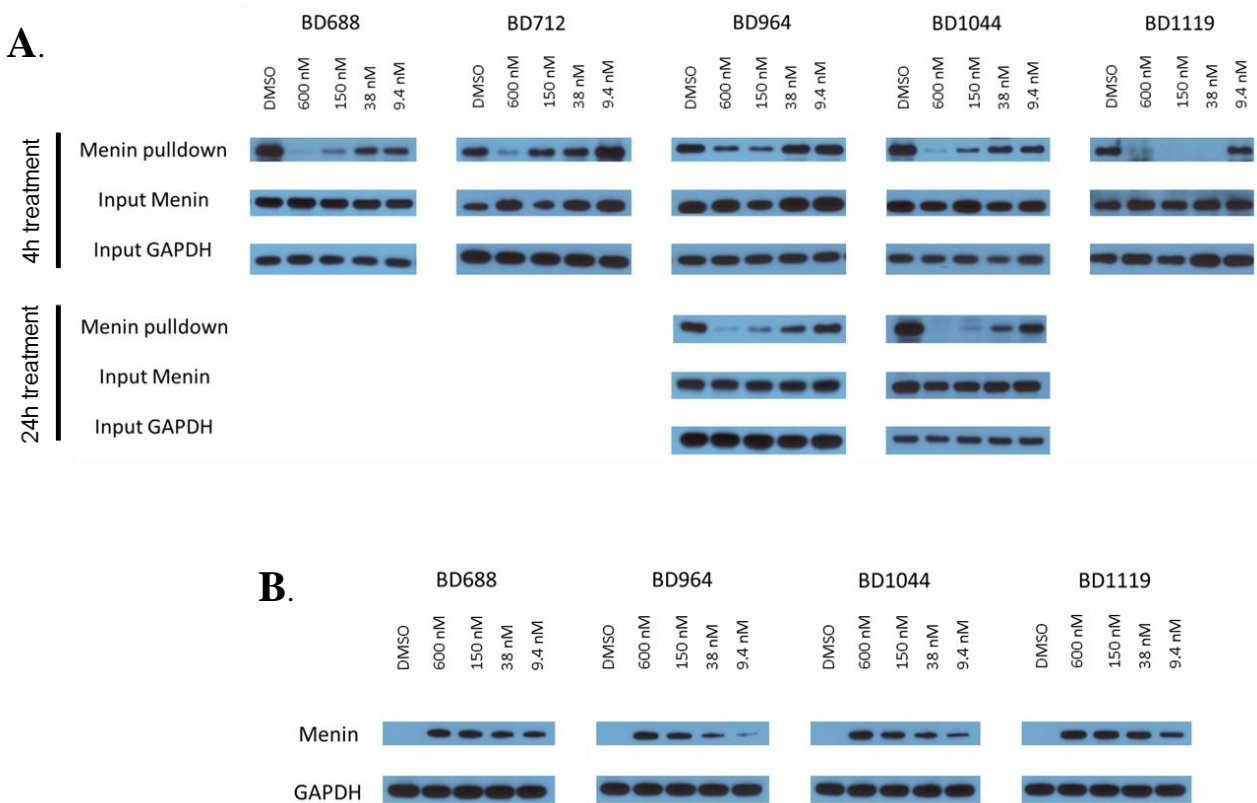
To explore the level of menin engagement following treatment with covalent inhibitors we developed a pulldown assay using biotinylated-MLL43 peptide immobilized to magnetic beads through a biotin-streptavidin complex (**Schematic 5.1**). The immobilized MLL43 on the beads will bind to uninhibited menin which we can then analyze by western blot. With this pulldown assay, we can quantify the menin that is not inhibited upon treatment with the inhibitors. We used KOPN8 cells to explore the covalent inhibitors menin engagement. CETSA<sup>140</sup> analysis was also performed to investigate the levels of menin engagement, which utilizes the heating of cells to unfold menin that is not stabilized through inhibitor binding. Cell samples were heated for 3min at 51°C. At the temperature 51°C, we found selective unfolding of menin that is not bound to an inhibitor while menin bound to an inhibitor is stabilized and

remains in solution. This approach allows us to quantify the level of menin that is bound to the inhibitor.



**Schematic 5.1 Design of the menin pulldown assay with MLL43 coated beads.**

Menin is pulldown from cell lysate with biotinylated-MLL43 and coupled to streptavidin coated beads. Uninhibited menin will be quantified by western blot analysis.

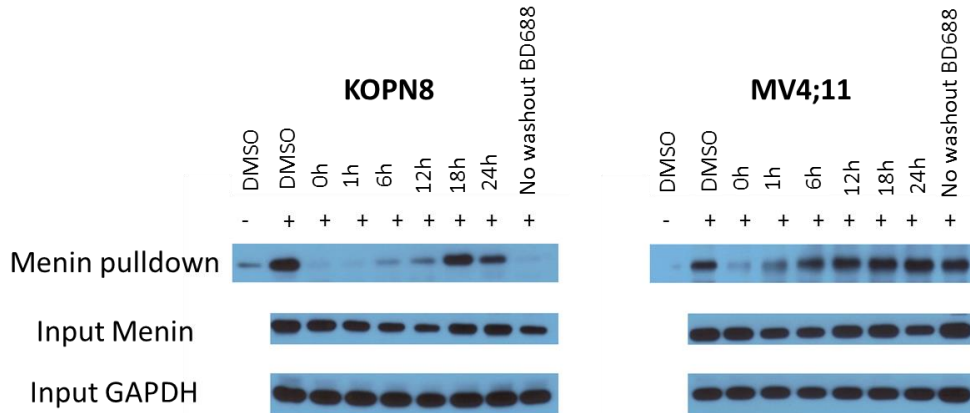


**Figure 5.9 Menin engagement in cells.**

**A.** KOPN8 cells were used to explore a 4h and 24h treatment of selected covalent inhibitors. Menin was pulled down with MLL43 coated beads and western blot analysis was performed to assess the level of uninhibited menin. **B.** CETSA analysis of menin engagement was performed in KOPN8 cells with a 4h treatment. Samples were heated at 51°C for 3min.

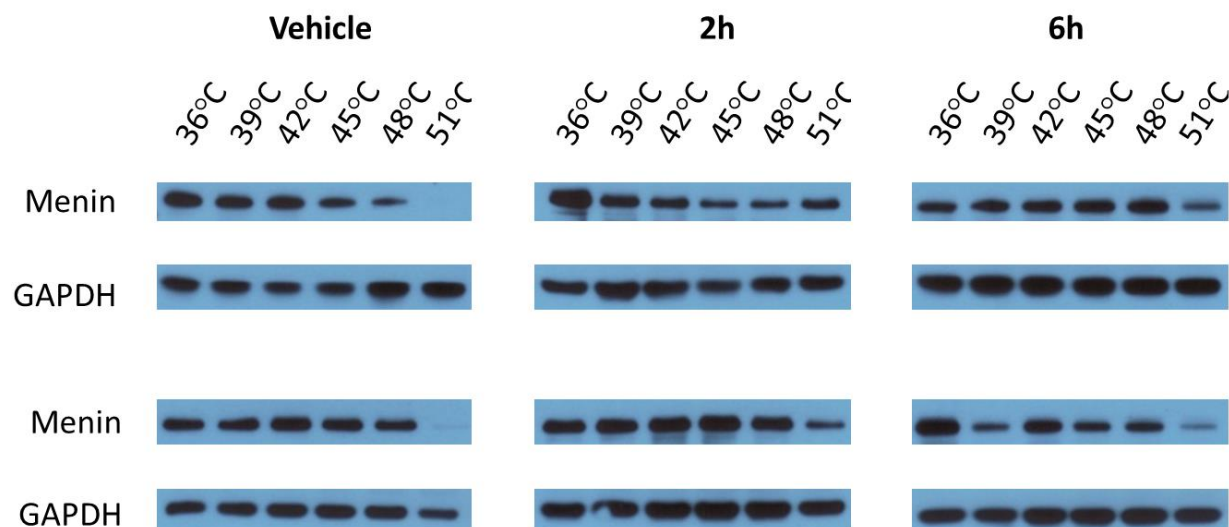
To determine the levels of menin engagement we explored the covalent inhibitors in KOPN8 cells analyzed by pull-down and CETSA (**Figure 5.9**). Similar results from both the pull-down and CETSA analysis of menin engagement. **BD1119** has the highest level of menin engagement with an  $IC_{50}$  of about 10nM. **BD1044** and **BD688** show similar levels of engagement with an  $IC_{50} = 150$ nM. **BD964** has an  $IC_{50} = 600$ nM with a 4h incubation and this improves about 4-fold when compared to the 24h incubation. When the incubation time is increased from 4h to 24h for **BD964**, the levels of menin engagement has a similar  $IC_{50}$  to **BD1044** while there is a little improvement between the  $IC_{50}$  of **BD1044** at 4h and 24h time points. This is because **BD1044** is able to react covalently to menin within the 4h incubation based on the half-life of menin engagement ( $t_{1/2} = 11$  min) whereas **BD964** ( $t_{1/2} = 4$ h) requires a longer incubation of 24h. Overall the  $IC_{50}$  values for menin engagement in MLL leukemia cells from the CETSA and pull-down analysis correlate well with the  $GI_{50}$  values in MLL-AF9 transformed murine BMCs.

We then explored menin turnover in KOPN8 and MV4;11 cells with 1  $\mu$ M **BD688** for 4 h and assessed the rate of newly synthesized menin using the pull down assay (**Figure 5.10**). In KOPN8 and MV4;11 cells the rate of menin turnover is between 6h and 12h. This analysis will help to determine the dosing regimen for *in vivo* treatment in mice to demonstrate a therapeutic response. *In vivo* the covalent inhibitors bind irreversibly to menin and remain bound until the protein is degraded; therefore their activity is dependent on the rate of *de novo* menin synthesis.



**Figure 5.10 Determination of the rate of menin turnover with washout experiments.**

KOPN8 and MV4;11 human leukemia cell lines with MLL rearrangements were treated with 1 $\mu$ M **BD688** for 4h to allow for full menin engagement. The remaining unreacted **BD688** was washed out and the rate of menin turnover was assessed with the MLL43 pulldown assay.



**Figure 5.11** CETSA analysis to assess menin engagement in the bone marrow *in vivo*.

C57BL/6 mice were treated with a single dose of **BD1119** (50 mg/kg) or vehicle via i.p. Bone marrow samples were harvested and levels of menin engagement was assessed by pull-down with the MLL43 coated beads and western blot analysis. n = 2 mice per treatment group and time point.

We went on to investigate the menin engagement in the bone marrow *in vivo* by CETSA analysis. We explored the target engagement of menin *in vivo* with **BD1119** based on the fast rate of menin engagement of this compound ( $t_{1/2} = 0.7$  min). C57BL/6 mice were dosed with 50mg/kg **BD1119** via i.p. injection and levels of menin engagement were explored at 2h and 6h time points compared to vehicle (**Figure 5.11**). In the vehicle treated samples no menin is present at 51°C while at 2h time point there is a strong band for menin demonstrating **BD1119** is bound to menin. At 6h time points the levels of menin that remain bound to the inhibitor starts to decrease as a result of menin turnover. At early time points we can measure significant levels of menin engagement but at later time points menin becomes resynthesized and the levels of menin bound to the inhibitor are decreasing. This data demonstrates that menin is covalently engaged in the bone marrow *in vivo* following treatment with **BD1119**.

## D. Conclusion

The work presented in this chapter describes the development of rationally designed covalent small molecule inhibitors of the menin-MLL interaction. We used structure-based design to convert the non-covalent thienopyrimidine class of inhibitors into covalent inhibitors targeting Cys329 of menin. We explored different reactive moieties and found the same linker

used for the vinylsulfonamide and propynamide reactive groups was not suitable for the correct positioning of the acrylamide to react with Cys329. The sidechain of Arg330 is an important hydrogen bonding partner with either direct or water mediated hydrogen bonds formed in all of the reactive groups explored. This hydrogen bonding interaction may help provide appropriate geometry for the reactive group to form the covalent bond with the sidechain of Cys329.

The reactive moieties we explored in the context of the thienopyrimidine scaffold were developed into potent inhibitors of the menin-MLL interaction and exhibit varying rates of menin engagement *in vitro*. The acrylamide of **BD1119** was found to covalently engage menin with the fastest rate, which correlated well with the potency in MLL-AF9 cells. All of the covalent menin inhibitors explored exhibited slower rates of GSH reactivity compared to Afatinib, an FDA approved covalent therapeutic. **BD1119** was the most reactive menin inhibitor with GSH but the rate was 2-fold lower than Afatinib, suggesting **BD1119** may be tolerated with respect to hepatocellular toxicity.<sup>126,135,136</sup> Inhibition of menin-MLL interaction is correlated with the degree of growth inhibition demonstrating the importance of the menin-MLL interaction in sustaining leukemic cell proliferation. The half-life of menin is relatively short with complete turnover of menin occurring within 6-12 hours. Importantly, in mouse models we have demonstrated that treatment with **BD1119** allows for rapid menin engagement in the bone marrow *in vivo*. The thienopyrimidine compounds that contain reactive moieties represent the first covalent inhibitors of the menin-MLL interaction and lay the groundwork for further development into a therapeutic strategy for leukemic patients with MLL-translocations.



## Chapter 6: Conclusions and Future Directions

### A. Conclusions

Translocations of MLL resulting in the onset of acute leukemias are associated with poor prognosis demonstrating an urgent need for novel therapies.<sup>32</sup> The MLL interaction with menin-LEDGF has been demonstrated to be a vital component in MLL fusion leukemia.<sup>9,55</sup> Genetic manipulation of the menin-MLL interaction which leads to the abrogation of the menin-MLL-LEDGF complex demonstrated the importance of the ternary complex formation for sustained leukemic cell proliferation.<sup>55,86</sup> Therefore, the reliance of MLL fusion proteins on the interaction with menin-LEDGF can be exploited to develop a novel therapeutic approach.<sup>9,55</sup> We have performed structural investigation of the molecular interactions involved in forming this ternary complex and has provided insight into the specific molecular interactions that need to be targeted by small molecule inhibitors.<sup>62</sup> Our own studies have demonstrated that pharmacologic agents can target the menin-MLL interaction and inhibit leukemic cell proliferation.<sup>59,66</sup> Studies in our laboratory have identified small molecule inhibitors of the menin-MLL interaction through HTS campaign, which after optimization have demonstrated *in vitro* inhibition of the menin-MLL interaction and modest cellular activity.<sup>59,66,73</sup> These inhibitors required further development to improve potency and PK properties for analysis *in vivo* in mice models of MLL leukemia.

The results I have presented within this thesis highlight the thorough characterization of the menin-MLL-LEDGF interaction to give insight into the development of small molecule inhibitors. We developed the first small molecule inhibitors of the menin-MLL interaction,

providing a proof-of-concept that menin is able to bind small molecules that could disrupt the menin-MLL interaction.<sup>68,73</sup> Additionally, we have further optimized the small molecule inhibitors of the menin-MLL interaction to serve as lead compounds for clinical development.<sup>1,66,68,70,71,73,141</sup>

### **A.1 Fluorine-backbone interactions contribute to high affinity interaction with menin**

The H to F substitution is a relatively minor modification, but as we demonstrated it can have a significant impact on binding affinity. We characterized the importance of the trifluoro moiety of the thienopyrimidine class of menin-MLL inhibitors which contributes significantly to the binding affinity of this class of molecules. We performed systematic exploration of the number of fluorine atoms at two distinct sites followed by structural and biochemical analysis. We found that not all fluorine atoms present in the CF<sub>3</sub> group are required for the high affinity interaction. Our analysis of the fluorine-backbone interaction of the menin inhibitors and investigation of PDB structures that contain these orthogonal multipolar interactions allowed us to develop the FMAP algorithm. We designed the FMAP algorithm to predict sites of a protein backbone that should interact favorably with fluorine atoms of a ligand through an orthogonal multipolar interaction. We have validated FMAP can predict favorable fluorine interaction sites of menin inhibitors with the protein backbone and of several other published protein-inhibitor complexes. Not all of the fluorine atoms of the trifluoro moiety are necessary for the high affinity interaction with menin. Therefore, FMAP can facilitate the rationale design of ligands with the optimal number of fluorine atoms to improve potency. The development of FMAP algorithm should aid in the rational design of fluorine-backbone interactions for protein ligands to improve small molecule binding affinity.<sup>1</sup>

The development of small molecule inhibitors targeting PPIs is difficult due to relatively large and flat interfaces frequently observed at PPI interfaces.<sup>110,142</sup> Importantly, PPI interfaces often engage the secondary structure elements such as  $\alpha$ -helices and  $\beta$ -sheets.<sup>143,144</sup> Our analysis using the FMAP algorithm highlights the potential to form orthogonal multipolar interactions between small molecule ligands and protein backbone in  $\alpha$ -helices and  $\beta$ -sheets and demonstrates that  $\beta$ -sheets have a much larger accessible area to form these interactions. This validates the importance of exploring the possibility of introducing the orthogonal multipolar interactions in the development of PPI inhibitors. C–F $\cdots$ C=O interactions with optimal geometry

can provide a substantial gain in the binding affinity of small molecules to proteins. For example a simple CH<sub>3</sub> to CF<sub>3</sub> substitution in the context of the menin inhibitors demonstrated over 10-fold increase in potency. Considering small molecule inhibitors that target PPIs typically have a high molecule weight to competitively bind the target protein, this approach can be applied to increase the binding affinity of the small molecule to the target protein without substantial increase in the molecular weight of such compounds. Additionally, orthogonal multipolar interactions require that the C-F bond has perpendicular geometry relative to the protein backbone; therefore, this may provide an opportunity to form favorable ligand interactions with the polar protein backbone when the introduction of hydrogen bonds would not be possible due to geometric limitations. With the increasing need to develop inhibitors of PPIs, application of the FMAP algorithm should help to efficiently design fluorine orthogonal multipolar interactions of small molecule inhibitors with the protein backbone.

## **A.2 Optimization of small molecule inhibitors of the menin-MLL interaction to demonstrate *in vivo* efficacy**

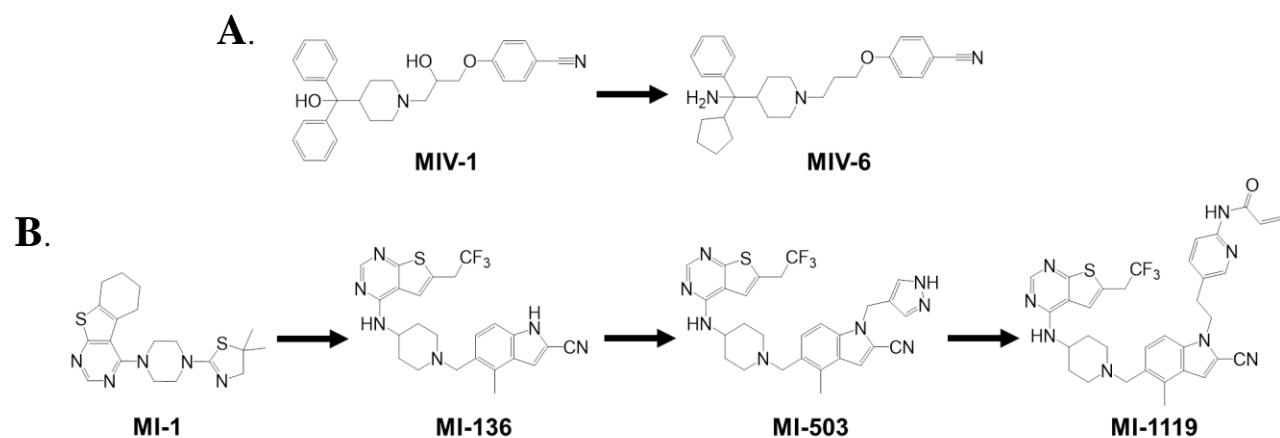
We performed optimization of several classes of menin-MLL inhibitors into potent inhibitors using structure-based drug design. Two distinct chemotypes were identified through the HTS, the hydroxymethylpiperidine class and thienopyrimidine class. The hydroxymethylpiperidine class of inhibitors (MIV series) was significantly improved potency upon, with the aminomethylpiperidine inhibitor (**MIV-6**) representing the most potent inhibitor of this class (**Figure 6.1**).<sup>73</sup> A multi-parameter optimization of the thienopyrimidine class of menin-MLL inhibitors was performed to improve potency, cellular activity, and PK properties.<sup>70,71</sup> These efforts led to the development of a series of compounds demonstrating *in vivo* efficacy in murine models of leukemia with **MI-503** demonstrating a survival benefit up to 45% (**Figure 6.1**).<sup>70,71</sup> Importantly, normal hematopoiesis in murine models tolerates the menin inhibitors demonstrating a therapeutic opportunity to treat MLL leukemias.<sup>70</sup>

We then used the thienopyrimidine class of inhibitors as an initial scaffold to develop covalent inhibitors of the menin-MLL interaction and to target Cys329 of menin. These covalent compounds were characterized to determine the potency, rates of reactivity to menin, intrinsic reactivity towards GSH, and menin engagement *in vivo*. We have demonstrated that inhibition of the menin-MLL interaction with covalent inhibitors can be achieved and contributes to

improved growth inhibition in MLL leukemia cells. The covalent inhibitors developed have different rates of menin engagement with the fastest compound **BD1119**  $t_{1/2} = <1\text{min}$ . All of the covalent menin inhibitors exhibited relatively low rates of intrinsic reactivity against GSH compared to Afatinib (**BD1119** has a 2-fold slower reactivity rate towards GSH compared to Afatinib). In mice, we have demonstrated that treatment with **BD1119** engages menin in the bone marrow cells *in vivo*. The half-life of menin in KOPN8 and MV4;11 leukemia cells was determined to be between a 6-12 h half-life which represents a relatively quick turnover. Therefore, covalent menin inhibitors as a potential therapeutic approach would be limited by the rate of menin turnover and would require dosing regimens 2-4 times per day to maintain inhibition of the menin-MLL interaction.

There are no published examples of rationally designed covalent inhibitors targeting PPIs. Therefore, the rational design of the covalent menin-MLL inhibitors and the methodology we used to characterize these compounds can serve as a set of guidelines to aid in the development and characterization of covalent small molecule inhibitors targeting other PPIs. We found the rate of menin turnover in cells to be an important factor in applying covalent menin inhibitors in biological studies. Therefore, before starting a covalent small molecule inhibitor campaign, it is important to determine the rate of the target protein *de novo* synthesis. Covalent inhibitors are only effective when the protein is covalently bound to the small molecule, therefore the rate of protein synthesis to physiologically relevant levels may be the limiting factor in the application of covalent inhibitors. Ideally this rate of *de novo* synthesis of the target protein should be slow ( $t_{1/2} > 12\text{h}$ ) to consider covalent inhibitors as a potential therapeutic strategy to block a specific target protein. To reduce the potential off target effects, the selection of covalent moieties should begin with weak electrophiles to generate covalent small molecule analogs. Weak electrophiles can be more easily explored when starting with a potent reversible inhibitor because the noncovalent interactions of the small molecule drive the potency and selectivity to position the electrophile in close proximity to the cysteine residue to allow for the covalent bond to form. This may reduce the potential for general reactivity of these covalent moieties towards other proteins and other thiol containing biomolecules. Next, it is important to characterize the potency, reactivity, selectivity and cellular activity of covalent small molecule inhibitors, and various biochemical, biophysical and cell biology methods can be applied for this purpose. Here, we utilized a biochemical FP assay to quantify inhibition of the targeted

interaction and the rate of covalent bond formation to the target protein by LC-MS to determine the rate of irreversible binding. The rate of target engagement should be relatively fast ( $t_{1/2} < 10\text{min}$ ) to allow for rapid and complete inhibition of the target protein. To characterize the inhibitors general reactivity towards proteins that contain surface cysteines and biomolecules containing free thiols the reactivity with GSH can be used as a surrogate. This reactivity ideally should be slow and we used Afatinib as a benchmark for comparison ( $t_{1/2} > 11\text{min}$  for Afatinib). Finally, we determined the rate of target engagement in both cellular studies and *in vivo* to validate whether inhibitors can covalently engage menin *in situ*. We developed the pull-down assay to quantify the uninhibited target protein and used CETSA to quantify the inhibited target protein. We found the  $\text{IC}_{50}$  values for both the pull-down and CETSA agreed well with each other. Overall, the methodology we developed and applied to characterize the covalent menin inhibitors will help set the foundation for the rational design and thorough characterization of irreversible inhibitors targeting PPIs. Our work also demonstrates that targeting PPI interfaces with covalent inhibitors is feasible and can represent a new direction to overcome some of the challenges associated with developing PPIs inhibitors.



**Figure 6.1 Development of menin-MLL inhibitors.**

**A.** The hydroxymethylpiperidine class of menin inhibitors with **MIV-6** the most potent analog developed in this series. **B.** **MI-1** the HTS hit developed to **MI-136** to serve as a lead scaffold. **MI-503** exhibited improved PK properties for analysis *in vivo*. **MI-1119** was a covalent inhibitor derived from the thienopyrimidine core scaffold.

## **B. Future Directions**

### **B.1 Future development of the menin-MLL inhibitors**

The menin-MLL inhibitors are a successful example of targeting protein-protein interactions with the reversible and irreversible inhibitors of the thienopyrimidine class and reversible methylpiperidine class. We have demonstrated the oncogenic effects of MLL-rearrangements can be reversed through disruption of the menin-MLL interaction. Overall, we have developed a series of menin-MLL inhibitors that exhibit a good PK profile to demonstrate *in vivo* efficacy supporting the effort for further exploration in clinical applications.<sup>71</sup> Additional analysis needs to be performed with the formulation of these inhibitors, long-term toxicology, and investigation in additional species such as rats and dogs. Menin-MLL inhibitors represent a promising therapeutic approach for the *MLL*-rearranged leukemias.

### **B.2 Investigation of menin-MLL inhibitors in other malignancies**

The menin-MLL interaction has a well characterized role in the progression of MLL-rearranged leukemias but other cancers that do not have MLL translocations may be dependent on the wildtype-MLL interaction with menin.<sup>11,42</sup> WT MLL is amplified in types of AML and myelodysplasia, and therefore menin-MLL inhibitors may provide a therapeutic benefit for these malignancies, particularly that have high *Hoxa9* expression.<sup>145</sup> In addition, menin interacts with MLL1 and MLL2, which have important roles in solid cancers including liver,<sup>78</sup> brain,<sup>79</sup> colon,<sup>80</sup> breast.<sup>81</sup> Considering the role of MLL1 in the regulation of MDR1, inhibitors of menin-MLL interaction may be important to chemo-sensitize cancer cells.<sup>82</sup> This evidence provides rationale for the exploration of the menin-MLL inhibitors in the context of various genetic backgrounds other than MLL translocations. Overall, the menin-MLL inhibitors may provide broader therapeutic advantages for the treatment of other malignancies.

### **B.3 Structural investigation of the IBD interaction with MLL identifies new interaction motif and potential therapeutic target**

The results I have shown in **Appendix A** demonstrate thorough characterization of the menin-MLL-LEDGF interaction with the identification of a novel interaction motif of MLL with IBD.<sup>141</sup> We identified a novel interaction motif of MLL binding to IBD, IBM2. Interestingly, the binding region of IBM2 on IBD overlaps with the HIV-IN binding site on IBD. In AlphaLISA

based competition experiments we determined that the MLL and IN interaction with IBD compete with each other and are mutually exclusive. Targeting the IBM2 site with inhibitors may provide a viable approach to develop therapeutics for MLL leukemias as well as HIV pathogenesis.<sup>141</sup>

Our structural characterization of the MLL-IBD interaction led to the identification of the IBM2 binding site on MLL that binds to IBD domain of LEDGF. The importance of the IBM1 and IBM2 binding motifs on MLL demonstrates that inhibition of either of these sites results in the disruption of the MLL interaction with IBD. Additionally, the IBM2 motif of MLL has an overlapping binding interface on IBD with HIV-IN suggesting this IBM2 binding site could be a promising therapeutic target for both MLL and HIV pathogenesis. This led us to explore small molecules that bind to the IBD of LEDGF through performing a fragment screen, of which details can be found in **Appendix B**. Targeting the interactions of IBD may provide an alternative therapeutic approach to disrupt the menin-MLL ternary complex rather or in addition to targeting the menin-MLL interaction.

## Materials and Methods

**Plasmid construction.** A cDNA encoding the full-length human menin was cloned into the pET32a vector (Promega) containing an N-terminal thioredoxin-His6 tag.<sup>59</sup> cDNA encoding HIV-IN and MLL110-160 constructs were synthesized by GeneArt Strings (Invitrogen).<sup>141</sup> LEDGF-IBD construct was ordered from Genscript and cloned into pET32a plasmid for bacterial expression.<sup>141</sup>

**Protein expression and purification.** Full-length human menin was expressed in a pET32a vector (Promega) containing N-terminal thioredoxin His6-tag in Rosetta (DE3) cells. Menin was purified using nickel-agarose (GE Healthcare) followed by ion exchange with Q-Sepharose (GE Healthcare). Protein was subsequently cleaved with thrombin followed by nickel agarose purification to separate the thioredoxin tag from menin. Purified protein was dialyzed to 50 mM Tris, 50 mM NaCl, 1 mM TCEP pH 7.5 buffer. Details of menin purification have been published previously.<sup>59</sup>

MLL constructs were expressed in Rosetta<sup>TM</sup> (DE3) cells (EMD) as an insoluble protein. The inclusion bodies were solubilized in 6M guanidine HCl and refolded on-column bound to Ni-NTA resin (Qiagen). The thioredoxin-His6x tag was cleaved by TEV protease, and proteins were purified by ion exchange chromatography with SP Sepharose resin (GE Healthcare). MLL was concentrated and dialyzed into 50 mM Tris, pH 7.5, 50 mM NaCl, and 1 mM TCEP. The generation of deletion constructs was performed by mutagenesis according to the QuikChange Site-Directed Mutagenesis kit protocol to introduce different internal deletions and a stop codon. Details have been described previously.<sup>141</sup>



IBD was expressed in *Escherichia coli* as an insoluble protein. The inclusion bodies were solubilized in 6M guanidine HCl and refolded on-column bound to Ni-NTA resin (Qiagen). The thioredoxin-His6x tag was cleaved by TEV protease, and proteins were purified by ion exchange chromatography with SP Sepharose resin (GE Healthcare). IBD was concentrated and dialyzed into 50 mM Tris, pH 7.5, 50 mM NaCl, and 1 mM TCEP. IBD labeled with stable isotope (<sup>15</sup>N) was expressed in M9 minimal media in *E. coli*. The same purification procedure above was used. HIV-IN was purified with HisTrap HP (GE Healthcare) and then by ion exchange chromatography on a Q Sepharose column (GE Healthcare). Details of IBD and HIV-IN purification have been described previously.<sup>141</sup>

**AlphaLISA assays.**<sup>141</sup> Inhibition of the menin-MLL-IBD interaction was determined with an AlphaLISA assay with 62.5 nM His-Trx-MLL160, 62.5 nM menin, and 62.5 nM Flag-IBD. 1h incubation at room temperature in 50 mM MOPS (pH 7.25), 50 mM NaCl, 1 mM TCEP, 0.02% (w/v) BSA, and 0.025% (v/v) Tween-20. Ni-chelate-coated donor beads and anti-Flag acceptor beads (10 mg/mL final concentration) were added and incubated for 1 hour. The luminescence signal was recorded using excitation at 680 nm and emissions at 615 nm using the PHERAstar microplate reader (BMG). Competition experiments with either integrase domain binding motif (IBM) 2 peptide or HIV-IN were performed with twofold serial dilutions. Data were analyzed using the Origin 7.0.<sup>167</sup> To assess the inhibition of HIV-IN interaction with IBD, we employed a very similar protocol as discussed previously using 62 nM His-HIV-IN and 1 mM Flag-IBD. Ni-chelate-coated donor beads and anti-Flag acceptor beads were used at 2.5 mg/mL concentration. Competition experiments were performed with IBM2 using threefold serial dilutions. Details have been described previously.<sup>141</sup>

**Biochemical Characterization of Menin-MLL Inhibitors.** Inhibition of the menin-MLL interaction by small molecules was assessed by fluorescence polarization (FP) assay. Fluorescein-labeled MLL (MBM1) peptide (MLL12) at 10 nM, menin at 100 nM, and varying concentrations of compounds were used for IC<sub>50</sub> determination in the FP buffer (50 mM Tris, pH 7.5, 50mM NaCl, 1 mM DTT). Compounds (5% final DMSO concentration) were added to the menin-MLL peptide complex and incubated for 1 h. Fluorescein-labeled MLL (MBM1 and MBM2) peptide (MLL43) at 4 nM, menin at 4 nM, and varying concentrations of compounds were used for IC<sub>50</sub> determination in the FP buffer (50 mM Tris, pH 7.5, 50mM NaCl, 1 mM

DTT). Compounds (5% final DMSO concentration) were added to the menin–MLL peptide complex and incubated for 3h. Changes in fluorescence polarization were measured using the PHERAstar microplate reader (BMG). A detailed protocol has been described previously.<sup>59,68</sup>

**Isothermal Titration Calorimetry (ITC).** Menin was extensively dialyzed at 4 °C against ITC buffer (50 mM phosphate, pH 7.5, 50 mM NaCl, 1 mM  $\beta$ -mercaptoethanol) and degassed prior to measurement. Compounds were dissolved in DMSO and diluted with the ITC buffer to final concentrations (50–100  $\mu$ M, 5% DMSO). Protein solution was adjusted to contain 5% DMSO final concentration. The titrations were performed using a VP-ITC titration calorimetric system (MicroCal) at 25 °C. The calorimetric cell, containing menin (concentrations in the range 5–10  $\mu$ M), was titrated with the compounds (50–100  $\mu$ M) injected in 10  $\mu$ L aliquots. Data was analyzed using Origin 7.0 (OriginLab) to obtain  $K_d$  and stoichiometry.<sup>70,71,73</sup>

Menin-MLL1-160 complex was purified by size exclusion chromatography and titrated with IBD using a VP–Isothermal Titration Calorimeter. Menin-MLL1-160 complex was used at 20 $\mu$ M concentration, 1000-fold above  $K_d$  for menin-MLL1-160 interaction.<sup>141</sup> A detailed protocol has been described previously.<sup>59,141</sup>

**Crystallization of the Menin Complexes with Small Molecule Inhibitors.** Co-crystallization of menin with small molecule inhibitors was performed with 2.5 mg/mL menin15 incubated with 3-fold molar excess small-molecule inhibitors. Crystals were obtained using a sitting-drop technique at 10 °C in 0.2 M ammonium acetate, 0.1 M HEPES, pH 7.5, and 25% w/v PEG 3350. Prior to data collection, crystals were transferred to cryosolution containing 20% PEG550 MME and flash-frozen in liquid nitrogen as described previously.<sup>1,66,70,71,73</sup>

**Crystallographic Data Collection and Structure Determination.** X-ray diffraction data for co-crystals of menin with small molecule inhibitors were collected at 21-ID-D, 21-ID-F, and 21-ID-G beamlines at the Life Sciences Collaborative Access Team at the Advanced Photon Source. Data was processed with HKL-2000.<sup>168</sup> Structures of the complexes were determined by molecular replacement using MOLREP with the apo structure of human menin (PDB code: 4GPQ) as a search model. The model was refined using REFMAC,<sup>169</sup> COOT,<sup>170</sup> and the CCP4 package.<sup>171</sup> In the final stages, refinement was performed with addition of the TLS groups defined by the TLSMD server.<sup>172</sup> Validation of the structures was performed using

MOLPROBITY<sup>173</sup> and ADIT.<sup>174</sup> Details of data processing and refinement are summarized in (Pollock 2015),<sup>1</sup> (He 2014),<sup>73</sup> (Borkin 2015),<sup>70</sup> (Borkin 2016),<sup>71</sup> coordinates and structure factors have been deposited in the Protein Data Bank.

**NMR Spectroscopy.** NMR experiments were done on the Bruker Advance III 600-MHz spectrometer with a 5 mm TCI cryogenic probe. All <sup>1</sup>H-<sup>15</sup>N heteronuclear single quantum coherence (HSQC) experiments were performed at 30°C. Spectra was analyzed with NMRPipe<sup>175</sup> and Sparky.<sup>176</sup>

**NMR Fragment Screen.** Samples were made with 80 μM concentration <sup>15</sup>N IBD in 50 mM Tris, pH 7.5, 50 mM NaCl, 1 mM TCEP, 5% DMSO with the final NMR sample containing 7% D<sub>2</sub>O. Compound mixtures contained 20 compounds with a final concentration of 250uM. Samples that contained hits were deconvoluted to identify the single compound that binds to IBD.

**Chemical shift perturbations.** The magnitude of the fast-exchanging chemical shifts perturbations was calculated as the change of the <sup>1</sup>H and <sup>15</sup>N chemical shifts in Hz.<sup>165</sup>

$$\Delta\delta_{NH} = \sqrt{((\delta N^{15} free - \delta N^{15} bound)^2 + (\delta H^1 free - \delta H^1 bound)^2)}$$

**Equation B.1** Calculation of chemical shift perturbations

**Determination of K<sub>D</sub> by analysis of chemical shift perturbations of small molecule titrations.** Chemical shift perturbations for small molecule titrations were fit in Origin<sup>167</sup> employing a binding isotherm described by the **Equation B.1**. [IBD] = 80 μM was used in all titrations; [Ligand] represents varied ligand concentration. B<sub>max</sub> was determined by the fitting of ligand titrations with a saturating binding isotherm.

$$\Delta\delta_{NH} = \frac{\left\{ \left[ 1 + \left( \frac{1}{K_D} + ([Ligand] + [IBD]) \right) \right] - \sqrt{\left[ 1 + \left( \frac{1}{K_D} + ([Ligand] + [IBD]) \right) \right]^2 - \left( \frac{4 \left( \frac{1}{K_D} \right) [IBD]}{B_{max}} \right) * \left( B_{max} \frac{1}{K_D} * [Ligand] \right)} \right\}}{\left\{ \frac{2 \left( \frac{1}{K_D} \right) [IBD]}{B_{max}} \right\}}$$

**Equation B.2** Calculation of K<sub>D</sub> for ligand titration.<sup>165</sup>

**NMR 3PA Method of quantifying SAR.** The 3PA method for comparing the relative binding affinity of small molecules was calculated using the average value for the chemical shift

perturbations in Hz at  $3^{15}\text{N}$ - $^1\text{H}$  HSQC peaks (T398, A390, Q391) with 500  $\mu\text{M}$  or 1 mM compound in 5% DMSO and compared to DMSO control. Chemical shift perturbations caused by compound binding were calculated as in **Equation B.1**.

#### **LC-MS determination of covalent adducts to menin.**

**Sample Preparation:** Samples contained 5 $\mu\text{M}$  compound, 500nM menin in PBS at pH7.4 with 5% DMSO. Reactions were incubated at room temperature for variable time and quenched with 2.5 $\mu\text{L}$  of 4% Formic acid (0.2% FA final).

**Method:** A Thermo Finnigan Surveyor Autosampler, PDA Plus UV detector and MS Pump along with an LTQ linear ion trap mass spectrometer were used to collect sample data under XCalibur software control. A 5 $\mu\text{L}$  sample in “no waste” mode was injected onto a Phenomenex Jupiter 5u 300A C5 (guard column) 2 x 4.00 mm at 45 °C. Mobile phase composition: Buffer A (95:5 water:acetonitrile, 0.1% FA) and Buffer B (acetonitrile, 0.1% FA). Gradient elution was used with an initial mobile phase of 85:15 (Buffer A:B) and a flow rate of 250 $\mu\text{L}/\text{min}$ . Upon injection, 85:15 A:B was held for 1.3min, Buffer B was increased to 90% over 3.2min, held for 1min, and then returned to initial conditions in 0.1min and held for 2.4min. The total run time was 8min. A post-column divert valve employed to direct void volume salts to waste was used for the first 2min of the sample method. Blank injection of Buffer A was used in between each of the sample injections. A needle wash of 1:1 acetonitrile:water with 0.1% FA was used.

The electrospray ionization (ESI) source used a 300 °C capillary temperature, 40 units sheath gas flow, 20 units aux gas flow, 3 units sweep gas flow, 3.5 kV spray voltage, 120 V tube lens.

**Data Collection:** Data collection was performed in the positive ion full scan mode 550-1500 Da, 10 microscans, 200 ms max ion time.

**Data Processing:** Protein mass spectra were acquired as XCalibur data files. The best scans were added together using XCalibur Qual Browser. The spectra were displayed using “View/Spectrum List with a Display option to display all peaks. The Edit/Copy cell menu was used to copy the mass spectrum into the PC clipboard. The spectrum in the PC clipboard was pasted into Excel. The first two columns (m/z and Intensity) were kept and the third column (Relative) was deleted.

The remaining two columns were then saved as a tab delimited file (m/z and intensity) as filename.txt from Excel. The Masslynx Databridge program was then used to convert the filename.txt tab delimited file to Masslynx format. In some cases an external calibration using a (similarly converted) myoglobin spectrum was applied in Masslynx to correct the m/z values of the menin protein m/z data. MaxEnt1 software from the MassLynx software suite was used for deconvolution of the mass spectrum to yield the average MW of the protein(s). The percentage of covalent adduct formation was determined from the deconvoluted result and used to calculate the rate of the covalent reaction.

### **LC-MS determination of covalent adducts to GSH.**

**Sample Preparation:** Samples were prepared with 500uL of 5mM Glutathione, 25uM compound in PBS at pH7.4 with 5% DMSO and incubated at 37 °C.

**Method:** A Thermo Finnigan Surveyor Autosampler, PDA Plus UV detector and MS Pump along with an LTQ linear ion trap mass spectrometer were used to collect sample data under XCalibur software control. A 20µL sample in “full inject” mode was injected onto a Phenomenex Jupiter 5u 300A C18 2 x 4.00 mm at 45 °C. Mobile phase composition: Buffer A (95:5 water:acetonitrile, 0.1% FA) and Buffer B (acetonitrile, 0.1% FA). For menin-MLL inhibitors, the gradient elution was used with an initial mobile phase of 95:5 (Buffer A:B) and a flow rate of 250µL/min. Upon injection, Buffer B was increased to 90% over 12min, held for 0.5min, and then returned to initial conditions in 0.1min and held for 2.4min. The total run time is 18.5min. For analysis of Afatinib the gradient elution method was adjusted to a an initial mobile phase of 92:8 (Buffer A:B) and a flow rate of 250µL/min. Upon injection, Buffer B was increased to 18% over 14min, and then returned to initial conditions in 0.1min and held for 4.5min. The total run time was 28min. A post-column divert valve employed to direct void volume salts to waste was used for the first 2min of the sample method. A needle wash of 1:1 acetonitrile:water with 0.1% FA was used.

The electrospray ionization (ESI) source used a 300 °C capillary temperature, 40 units sheath gas flow, 20 units aux gas flow, 3 units sweep gas flow, 3.5 kV spray voltage, 120 V tube lens.

**Data Collection:** Data collection was performed in the positive ion full scan mode 550-1500 Da, 10 microscans, 200 ms max ion time.

**Sample preparation of murine bone marrow cells.** C57BL/6 mice were treated with 50 mg/kg of BD1119 via intraperitoneal (ip) administration. Compounds were dissolved in the vehicle containing 25% (v/v) DMSO, 25% (v/v) PEG-400, and 50% (v/v) PBS. Vehicle mice were sacrificed at 2h after treatment and BD1119 treated mice were sacrificed at 2h and 6h time points after administration of the compound. Bone marrow cells (BMCs) were isolated from the tibia and femur. BMCs were washed with PBS and resuspended in PBS for CETSA analysis.

**Sample preparation of human leukemia cell lines.** Human leukemia cell lines: MV4;11 and KOPN8 were cultured in RPMI-1640 medium with 10% FBS and 1% penicillin/streptomycin.  $4 \times 10^6$  cells were seeded in T-25 cell culture flasks (BD Biosciences, San Jose, CA, USA) in appropriate volume of culture medium, cells were treated with compounds (0.25% final DMSO concentration) or DMSO for 4 hours or 24 hours at 37°C, 5% CO<sub>2</sub>. Following the incubation the cells were harvested and washed with PBS in order to remove excess drug. Cells were pelleted followed by the addition of lysis buffer (50mM Hepes pH 7.5, 250 mM NaCl, 1mM EDTA, 2.5mM EGTA, 1mM NEM, 1mM NaF, 100uM Na<sub>3</sub>Vo<sub>4</sub>, 10mM Beta Glycerophosphate, 10% glycerol, 1% NP-40) at 30uL volume and lysed using 3 cycles of freeze-thawing with liquid nitrogen. The lysates were centrifuged at 20000 x g for 10 minutes at 4°C and supernatants were transferred to new tubes.

**Cellular Thermal Shift Assay.** Cellular thermal shift assay (CETSA) was performed as described previously.<sup>140</sup> Briefly equal amounts of cell suspensions were aliquoted into a 0.2 mL PCR tubes. These 0.2mL PCR tubes were heated individually at different temperatures for 3 minutes (Eppendorf, Hamburg) followed by cooling for 3 minutes at room temperature. Cells were pelleted followed by the addition of lysis buffer (30uL) and lysed using 3 cycles of freeze-thawing with liquid nitrogen. The lysates were centrifuged at 20,000 x g for 10 minutes at 4°C and supernatants were transferred to new tubes and analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) followed by western blot analysis.

**MLL43 Bead preparation.** Biotinylated MLL43 peptide was ordered from GenScript with the sequence: GSGMAHSARWRFPARPGTTGGGGGGGRRGLGGAPRQRVPALLPPG with C-Terminal Amidation and N-Terminal Biotin.

Biotinylated MLL43 beads were prepared with 100uL of 10mg/mL Streptavidin Magnetic beads (Pierce) were washed according to manufacturer's protocol. Beads were resuspended in 200uL of 0.5mM biotinylated-MLL43 in PBS and incubated for 18 hours at 4°C while rotating. Beads were washed extensively with PBS and stored in PBS containing 0.05% sodium azide.

**MLL43 pulldown of menin assay.** Whole cell lysates were incubated with MLL43 coated beads at 4°C for 18h while rotating. Beads were isolated and washed extensively in lysis buffer containing an additional 0.5M NaCl. The beads were then applied to SDS-PAGE electrophoresis and Western blotting. Antibodies used include rabbit anti-menin (A300-115A, Bethyl Laboratories, Inc.), mouse anti-GAPDH (AM4300, Ambion), horse anti-mouse HRP (7076, Cell Signaling), goat anti-rabbit HRP (7074, Cell Signaling).

## **Appendix A: The same site on IBD of LEDGF binds to both MLL and HIV-IN.**

\* The text and figures in this chapter are adapted from the following manuscript: Murai, M. J.; **Pollock, J.**; He, S.; Miao, H.; Purohit, T.; Yokom, A.; Hess, J. L.; Muntean, A. G.; Grembecka, J.; Cierpicki, T. The same site on the integrase-binding domain of lens epithelium-derived growth factor is a therapeutic target for MLL leukemia and HIV. *Blood* 2014, 124, 3730-7.

### **A. Abstract**

Lens epithelium-derived growth factor (LEDGF) plays a vital role in the pathogenesis of both MLL leukemias and HIV type-1 viral replication. In MLL rearranged leukemias, the integrase-binding domain (IBD) of LEDGF is involved in the interaction with MLL. To fully characterize this interaction we performed structural analysis of the IBD domain, which revealed a novel interaction motif of MLL, termed IBD binding motif 2 (IBM2). Validation of this interaction was performed through biochemical and biophysical analysis. The IBM2 site of MLL was identified to overlap with the HIV-IN binding interface on IBD. We have demonstrated the interaction with IBD to MLL or HIV-IN is mutually exclusive. Importantly, a peptide corresponding to the IBM2 sequence is capable of disrupting the IBD interaction with MLL or HIV-IN. This evidence suggests targeting the IBM2 binding site on LEDGF represents a novel approach for the development of therapeutics for MLL-dependent and HIV pathogenesis.



## B. Background

Lens epithelium-derived growth factor (LEDGF) is a chromatin associated protein<sup>9</sup> and acts as a transcriptional co-activator regulating stress related genes.<sup>146,147</sup> LEDGF has been implicated in cancer,<sup>148</sup> autoimmunity,<sup>149</sup> and HIV infection.<sup>150</sup> LEDGF translocations have been reported with nucleoporin NUP98 gene in leukemia patients,<sup>151,152</sup> and are upregulated in subset of acute myeloid leukemias (AML) resistant to chemotherapy.<sup>153</sup>

Evidence for the role of LEDGF in leukemogenic transformation by MLL fusion proteins has been demonstrated.<sup>9</sup> LEDGF co-precipitates with MLL wild type and MLL fusion proteins along with menin<sup>9</sup> and co-localizes with MLL-menin complex at *Hoxa9* and *Meis1* gene loci.<sup>9</sup> *Hoxa9* and *Meis1* genes are important downstream targets of the menin-MLL-LEDGF complex and are required for normal hematopoiesis and have been linked to MLL fusion pathologies.<sup>2,19</sup> The oncogenic function of MLL fusion proteins is critically dependent on the association with menin and LEDGF.<sup>9</sup> The Integrase Binding Domain (IBD) of LEDGF is the domain of LEDGF involved in the interaction with MLL variants and menin.<sup>9</sup> Disruption of the menin-MLL fusion protein interaction with LEDGF abrogates the progression of acute leukemia *in vivo*.<sup>9</sup> This evidence demonstrates the potential of targeting the LEDGF interaction with MLL as a therapeutic approach for MLL-rearranged leukemias.

Additionally, the IBD of LEDGF interaction with HIV-1 integrase (IN) and the importance of LEDGF in HIV type 1 (HIV-1) pathogenesis has been well characterized.<sup>154,155</sup> The IBD domain of LEDGF interacts with the catalytic domain of HIV-IN.<sup>154,156 157</sup> Small molecules and peptidomimetics have been developed to bind the HIV-IN and disrupt the interaction with IBD of LEDGF inhibiting HIV-IN activity and viral replication.<sup>158,159</sup> This evidence suggests the IBD of LEDGF plays an important role in both MLL-rearranged leukemias and HIV pathogenesis.

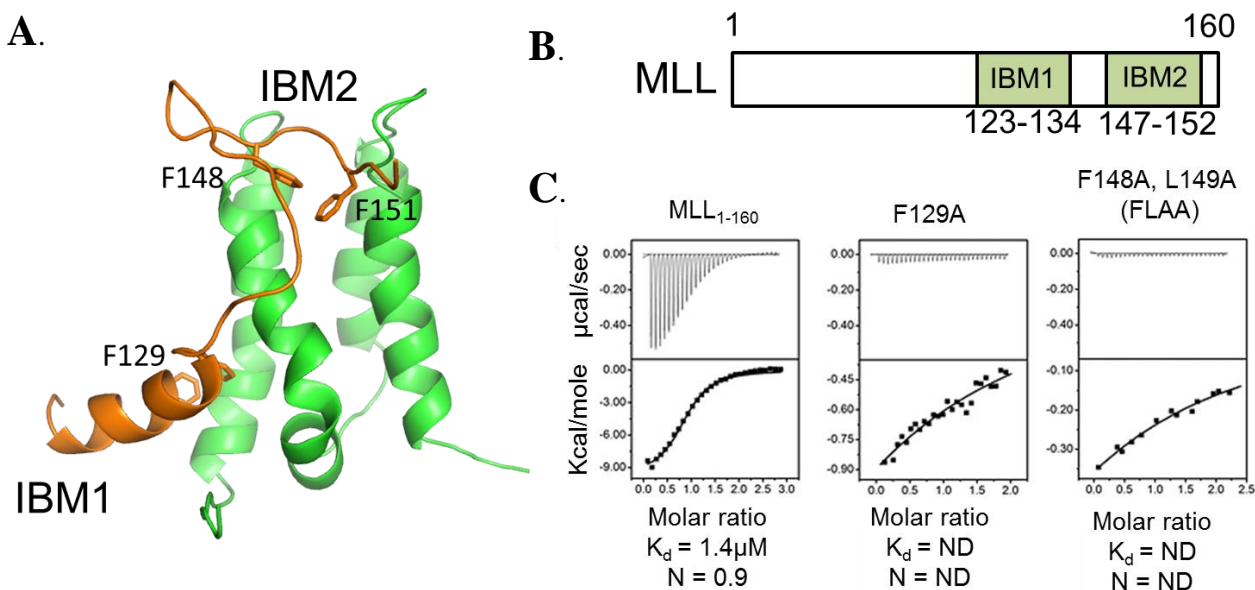
The molecular interactions of IBD of LEDGF with MLL are not thoroughly characterized. Structural analysis (PDB: 3U88) of the menin-MLL-IBD ternary complex clearly highlights the interactions of MLL residues 113-134 with IBD.<sup>62</sup> The Phe129 and Phe133 are clearly forming hydrophobic interactions with a groove of IBD. This data correlates well with previous studies describing the importance of Phe129 for binding to IBD.<sup>9</sup> ITC analysis of the

menin-MLL complex, with the MLL construct containing residues 6-153, binds to IBD with a  $K_d = 470$  nM. Although the MLL construct used for crystallization consisted of MLL residues 6-153, in the crystal structure electron density for residues 135-153 were not clearly defined and left unmodeled. It has also been demonstrated that deletion of residues 123-153 of MLL abolished the interaction of MLL with IBD suggesting the presence of additional interactions which were absent in the crystal structure.<sup>9</sup>

To investigate this interaction we determined the solution structure of MLL in complex with IBD; revealing a novel interaction motif of MLL<sup>141</sup> Importantly, this new motif of MLL overlaps with the binding site of HIV-IN on IBD of LEDGF.<sup>141</sup> This suggests that targeting this site of IBD may have therapeutic applications in both MLL rearranged leukemias and HIV pathogenesis.

My contribution to this project was to characterize the binding affinity of MLL mutants in complex with menin to the IBD-LEDGF by ITC. I have also developed the biochemical AlphaLISA assays to probe the interactions of the IBM2 site of MLL with IBD or HIV-IN.

### C. Results and Discussion



**Figure A.1 Structure determination and biophysical characterization of the MLL-IBD interaction.** A. The solution of IBD in complex with MLL (PDB: 2MTN)<sup>141</sup>. B. Figure adapted from (Murai 2014)<sup>141</sup> Construct of MLL residues 1-160 highlighting IBM1 and IBM2. C. Figure adapted from (Murai 2014)<sup>141</sup> Binding

isotherms of IBD titrated to menin in complex with different constructs of MLL.

### **C.1 Structural and biochemical analysis of the menin-MLL interaction with IBD.**

The construct used for the solution structure was designed to include residues 110-160 of MLL fused to residues 337-442 of LEDGF which corresponds to IBD. The solution structure of MLL-IBD was then solved by Marcelo Murai and Tomasz Cierpicki (**Figure 7.1 A**).<sup>141</sup> This structure highlighted two separate fragments of MLL forming interactions with IBD, IBD binding motif 1 (IBM1) (residues 123-134) and IBM2 (residues 147-152) (**Figure 7.1 A and B**). The IBM1 site was determined in the crystal structure of the ternary complex but the IBM2 site was absent.<sup>62</sup> IBM1 and IBM2 are separated by a 12 amino acid linker. IBM2 residues F148 and F151 form interactions with a hydrophobic region of IBD.

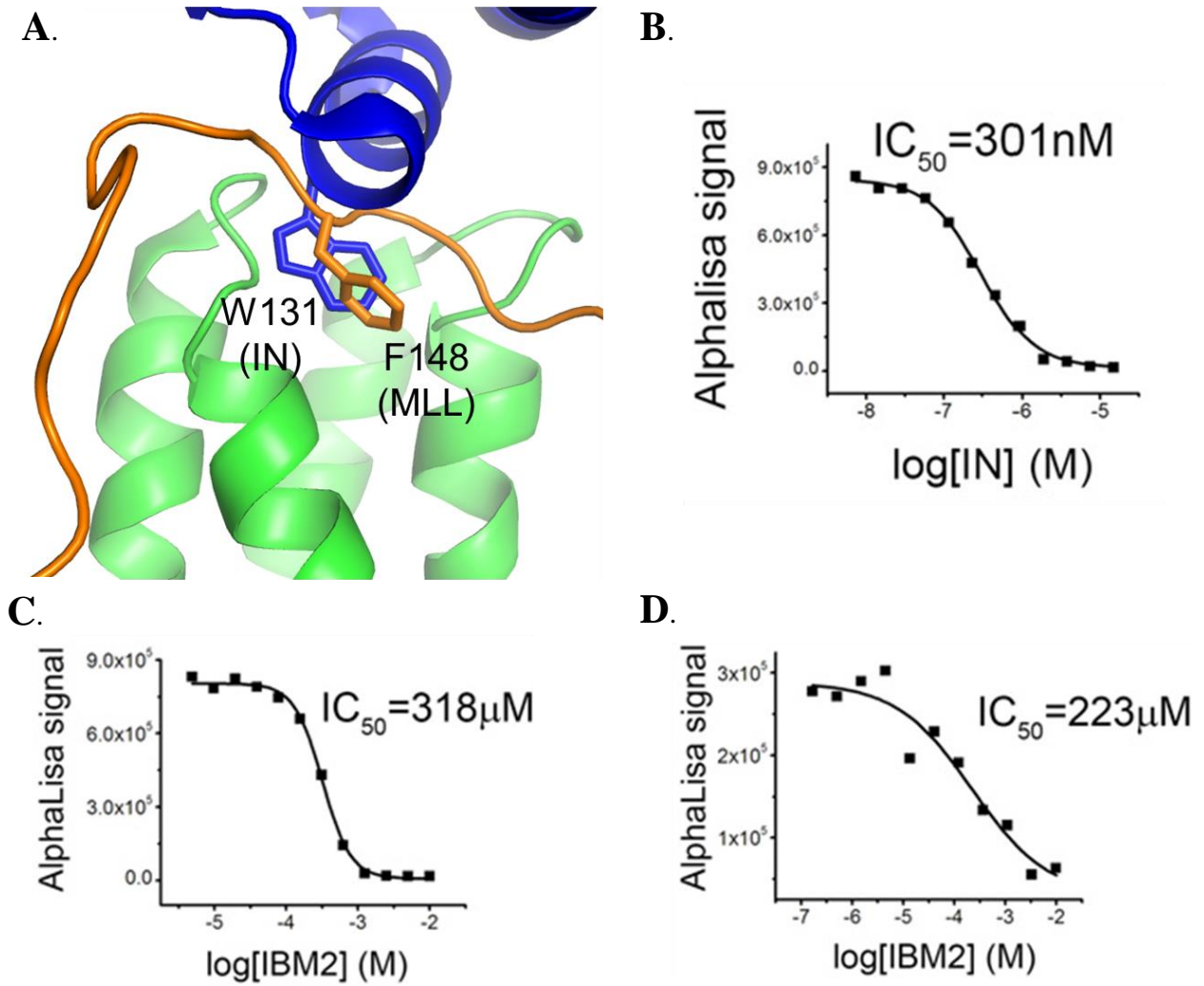
I validated these interaction motifs of MLL for binding to IBD using isothermal titration calorimetry (ITC). First, I characterized the interaction of IBD with menin-MLL1-160 which had a binding affinity of 1.4  $\mu\text{M}$ , approximately 10-fold higher compared to the interaction between IBD and MLL in the absence of menin (14.7  $\mu\text{M}$ ) (**Figure A.1 C**).<sup>141</sup> Mutagenesis was then performed to validate the IBM1 and IBM2 interaction motifs (**Figure A.1 C**). The IBM1 interaction with menin-MLL was abolished with the F129A mutation and F148A and L149A double mutant eliminated the IBM2 interaction with IBD. Each of these mutants in IBM1 and IBM2 independently was sufficient to completely disrupt the IBD interaction with menin-MLL. This analysis demonstrates both the IBM1 and IBM2 sites are necessary for the high-affinity interaction of menin-MLL with IBD.<sup>141</sup>

### **C.2 HIV-IN binding to IBD overlaps with the MLL binding interface.**

In HIV pathogenesis, HIV-IN binds to the IBD of LEDGF and is necessary for the integration of the HIV genome.<sup>9,154</sup> In MLL transformations, the IBD of LEDGF is involved in ternary complex formation with menin-MLL and is critical for the progression of MLL rearranged leukemia.<sup>9,154</sup> Both MLL and HIV-IN are involved with protein-protein interactions with IBD of LEDGF. The crystal structure of IN-IBD complex<sup>160</sup> (PDB: 2B4J) was superimposed with the MLL-IBD solution structure<sup>141</sup> (PDB: 2MTN) demonstrating the IBM2 site of MLL overlaps with the IN binding site on IBD (**Figure A.2 A**). W131 of IN binds to the same hydrophobic region on IBD as F148 of MLL. To investigate whether IN will inhibit the

MLL interaction with IBD I developed an AlphaLISA assay to probe the IBD-MLL interaction with IN. The AlphaLISA assay utilized a Ni-chelate donor and anti-Flag acceptor beads with the proteins 62.5 nM His-MLL160, 62.5 nM menin, and 62.5 nM Flag-IBD. The competition experiment with IN was done with 2-fold serial dilutions. It has been reported that IN binds to IBD with an affinity of about 10nM<sup>161</sup> and we found in our AlphaLISA assay IN competes with menin-MLL binding to IBD with an  $IC_{50} = 301$  nM. Therefore, IN can compete with MLL for the binding to IBD.

To further characterize the overlap of the binding interface of MLL and IN to IBD, I explored an IBM2 peptide to disrupt these interactions with IBD. Again an AlphaLISA assay was used to explore the potential for the IBM2 peptide to compete with the IBD interactions using the assay described above for the ternary complex and developed another AlphaLISA assay to probe the IBD-IN interaction. To assess the inhibition of HIV-IN interaction with IBD, the AlphaLISA assay was optimized with 62 nM His-HIV-IN and 1 mM Flag-IBD. I identified that the IBM2 peptide was able to compete with IBD binding to menin-MLL ( $IC_{50} = 318$   $\mu$ M) and IBD binding to HIV-IN ( $IC_{50} = 223$   $\mu$ M). The binding affinity of IBM2 peptide was estimated to be about 10-fold weaker than the binding of MLL1-160 to IBD.<sup>141</sup> Although the interaction of the isolated IBM2 peptide is not very potent, this evidence demonstrates the potential to develop an inhibitor to disrupt both the IN and MLL interaction with IBD. Therefore, this provides proof of concept that the IBM2 binding site on IBD can be targeted with potentially a small molecule or a peptidomimetic to disrupt the IBD interaction with both MLL and IN.



**Figure A.2 The IBD interface with MLL and IN overlaps.**

**A.** Figure adapted from (Murai 2014)<sup>141</sup> Structures of IBD-MLL overlain with IBD-IN. W131 of IN overlaps with F148 of IBM2 of MLL. HIV-IN (blue), IBD (green), MLL (orange). **B.** Figure from (Murai 2014)<sup>141</sup> AlphaLISA assay of the menin-MLL-IBD ternary complex demonstrated disruption of the complex when IN is titrated in. **C.** Figure adapted from (Murai 2014)<sup>141</sup> menin-MLL-IBD AlphaLISA assay with IBM2 competition. **D.** Figure adapted from (Murai 2014)<sup>141</sup> IBD-IN AlphaLISA assay with IBM2 competition.

## D. Conclusions

The solution structure of the MLL-IBD complex allowed us to identify a novel interaction site of MLL with IBD, termed IBM2. The IBM2 site is required for the menin-MLL complex to bind to IBD.<sup>141</sup> Additionally, the IBM2 motif overlaps with the binding site of IN to IBD demonstrating the MLL and IN interaction with IBD is mutually exclusive.<sup>141</sup> Importantly, the IBM2 peptide is able to inhibit the IBD interactions with both MLL and IN demonstrating the potential for an inhibitor targeting this site of IBD. This evidence validates the IBM2 binding site of IBD as a potential therapeutic target in both MLL leukemias and HIV pathogenesis.

Antiviral drugs have been developed to target the IBD-IN interaction but all bind to HIV-IN.<sup>162</sup> Viral proteins can more quickly develop mutations and become resistant to current therapies therefore inhibitors that bind to IBD of LEDGF may serve as a promising therapeutic.<sup>163</sup> Targeting the menin-MLL-LEDGF complex has been demonstrated to be an attractive therapeutic target for MLL rearranged leukemias.<sup>9,164</sup> Inhibitors targeting the menin-MLL interaction have been developed<sup>66,68</sup> however concerns surrounding menin's role as a tumor suppressor protein in endocrine tissues<sup>57</sup> may have undesired consequences if inhibited. Overall, this analysis demonstrates an opportunity of an alternative approach to target the menin-MLL-LEDGF ternary complex for potential therapeutic applications.<sup>141</sup>

## **Appendix B: Fragment screening of IBD-LEDGF by NMR**

### **A. Abstract**

The menin-MLL-LEDGF complex assembly is critical for the progression of *MLL* rearranged acute leukemias. Targeting the menin-MLL interaction to disrupt this complex has been achieved but the binding of small molecules to menin raise concerns surrounding the role of menin as a tumor suppressor protein in the endocrine system. Therefore, targeting the IBD-MLL interaction with small molecule inhibitor may serve as a promising alternative approach to disrupt this ternary complex. To identify if small molecules can bind to IBD we employed a fragment-based screening campaign. We identified two chemically distinct small molecules and validated binding to IBD. The binding site on IBD for both small molecules overlaps well with the IBM1 motif of MLL. With modest affinity of these small molecules, further optimization is required to demonstrate inhibition of the menin-MLL interaction with IBD. This study provides proof of concept that small molecule binding to IBD at a relevant site for MLL binding can be achieved.

## B. Background

LEDGF has been demonstrated to play a critical role in *MLL*-rearranged pathogenesis through the association with menin-*MLL*.<sup>9</sup> The integrase binding domain (IBD) of LEDGF has been determined to interact with menin-*MLL*.<sup>9,62</sup> Importantly disruption of the menin-*MLL* fusion protein interaction with LEDGF abolishes the progression of acute leukemia.<sup>9</sup> This evidence highlights the potential for inhibitors targeting the IBD interaction with menin-*MLL* to have therapeutic applications in *MLL* leukemias.

Our lab has demonstrated that disruption of menin with small molecule inhibitors inhibits *MLL* leukemia cell proliferation.<sup>66,68</sup> However, we have demonstrated the importance of the IBD-*MLL* interaction, of which; inhibitors have yet to be identified. The LEDGF interaction with the menin-*MLL* complex involves multiple points of contact.<sup>62,141</sup> We have performed extensive characterization of the IBD interaction with menin-*MLL* and have identified two IBD binding motifs 1 (IBM1) and IBM2 of *MLL*.<sup>62,141</sup> These interactions were validated by mutagenesis demonstrating the importance of both of these interactions.<sup>141</sup> Furthermore, this provides rationale for targeting either of these sites with small molecule inhibitors.<sup>62,141</sup>

In this study we performed a fragment-based screening approach to identify inhibitors of the IBD interaction with menin-*MLL*. Two small molecules were found to bind to IBD and overlap with the IBM1 binding site of *MLL*. This work is highly innovative because to date there is no report of small drug like molecules that bind to LEDGF. Disruption of IBD interactions with menin-*MLL* fusion proteins with small molecules is a novel strategy to develop targeted therapeutics for *MLL*-related leukemias.

My contributions to this project were to perform the fragment-based screening of the IBD by NMR using <sup>1</sup>H-<sup>15</sup>N HSQC experiments. I then selected commercially available analogs of the small molecule hits identified from the fragment screen and characterized the potency of these analogs by <sup>1</sup>H-<sup>15</sup>N HSQC experiments. I performed K<sub>d</sub> analysis of 3CC7 binding to IBD using <sup>1</sup>H-<sup>15</sup>N HSQC chemical shift perturbations and mapped these perturbations onto the IBD solution structure to identify the binding region of 3CC7 to IBD.

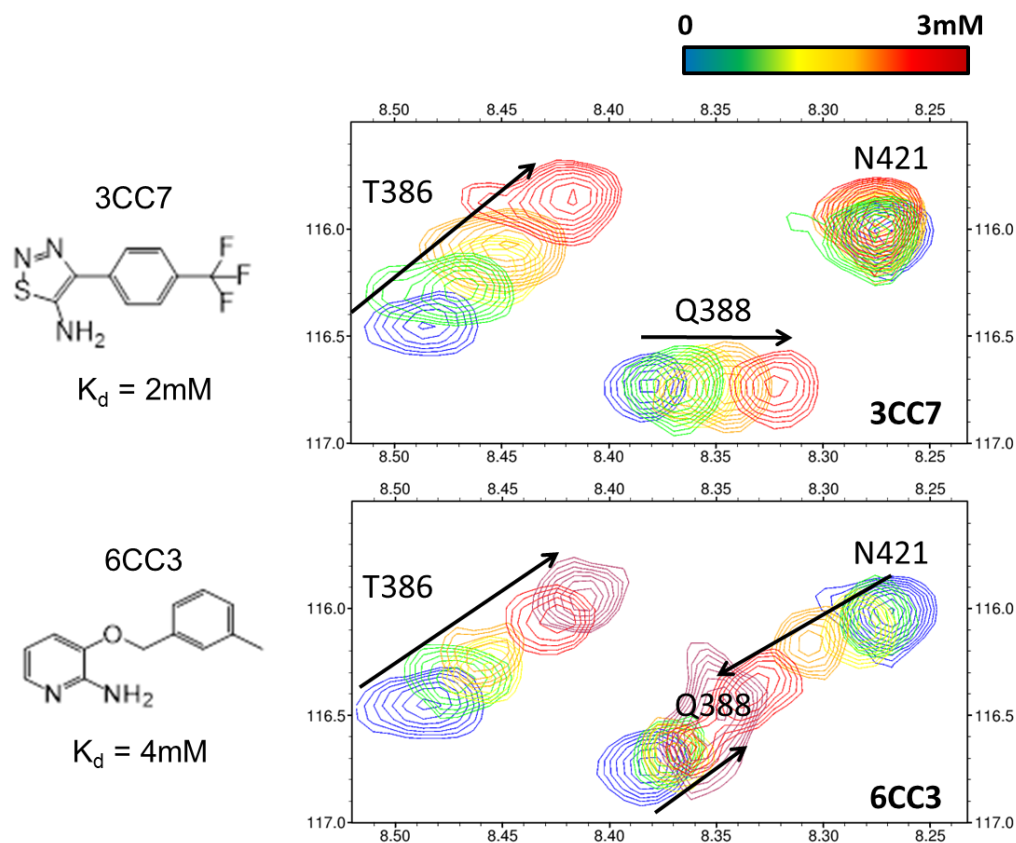


## C. Results and Discussion

### C.1 Identification and characterization of small molecule fragments that bind to IBD.

To explore the potential for small molecules to bind to IBD of LEDGF we performed a fragment-based screen. A library of ~500 chemically diverse fragment-like compounds was selected to screen by NMR. A  $^1\text{H}$  and  $^{15}\text{N}$  HSQC experiment was utilized to observe the  $^1\text{H}$ - $^{15}\text{N}$  correlations for uniformly labeled  $^{15}\text{N}$  IBD. The screening was performed in a multiplex format with 20 compounds per NMR sample. Hits were defined when chemical shift perturbations with greater than 0.03ppm shifts in the  $^1\text{H}$  dimension or greater than 0.03ppm shifts in the  $^{15}\text{N}$  dimension compared to DMSO control. The two compounds we identified from this screen were **3CC7** and **6CC3** which are shown in **Figure B.1**.

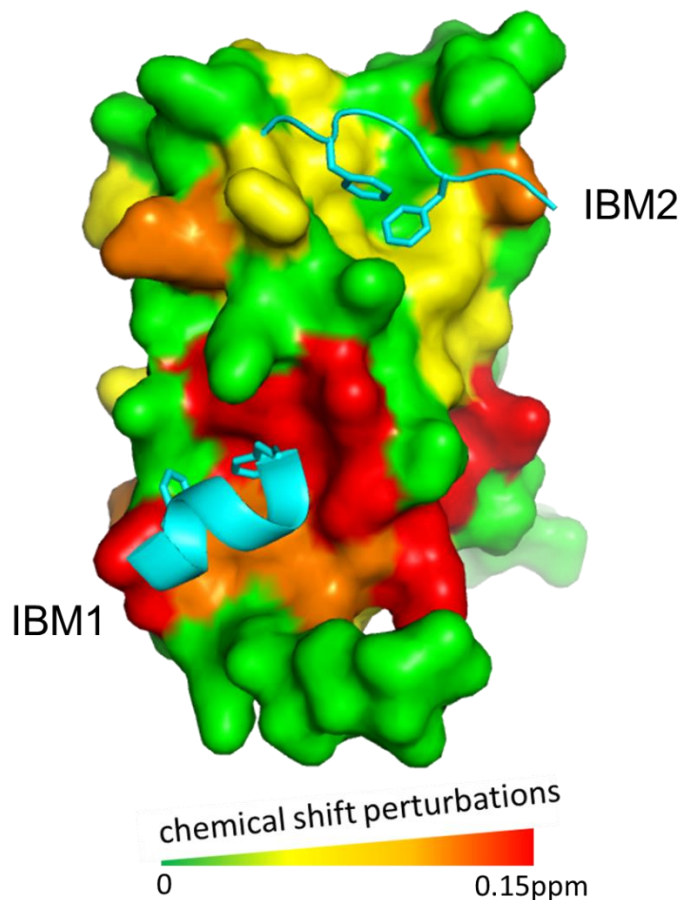
We performed titration experiments with **3CC7** from 250uM-2mM and **6CC3** up to 3mM to determine the  $K_d$  by NMR utilizing the chemical shift perturbations and **Equation B.2** (See Methods section). Over 20 chemical shifts were used to calculate the average binding affinity. The most potent small molecule identified was **3CC7** with a binding affinity of about 2mM, and **6CC3** was 2-fold weaker. Small molecule fragments typically exhibit very weak binding and require optimization but first we sought to identify if these small molecules bind to a site that might overlap with the MLL binding interface.



**Figure B.1 Identification and characterization of fragment screening hits.**

Structures of the hits identified in the fragment screen. Titrations of the small molecules analyzed by  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments with a selection of the spectra shown on the right. Blue = DMSO, green = 125 $\mu\text{M}$ , yellow = 250 $\mu\text{M}$ , orange = 500 $\mu\text{M}$ , red = 1mM, dark red = 2mM. Binding affinity was determined with using Equation 8.2.<sup>165</sup>

The IBD domain consists of a small 4-helix bundle therefore small molecules could potentially bind at the interface of any of these helices. Based on the similar residues resulting in chemical shift perturbations for both **3CC7** and **6CC3** we concluded both hits bound to the same region on IBD. To identify the region of IBD these small molecules were binding to, we mapped the difference in chemical shift perturbations of **3CC7** at 2mM compared to DMSO. We found the largest chemical shifts perturbations localized at the IBM1 binding site (**Figure B.2**). Interestingly, the small molecules identified in this screen bind to an important site on IBD for the MLL interaction.



**Figure B.2 Solution structure of IBD with chemical shift perturbations mapped at 2mM 3CC7.**

The solution structure highlights the residues with chemical shift perturbations upon the addition of 3CC7 relative to DMSO control. IBM regions 1 and 2 of MLL are highlighted in cyan.

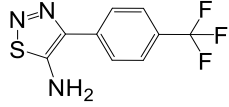
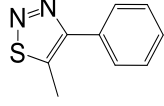
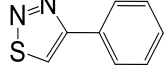
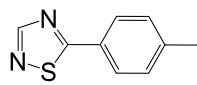
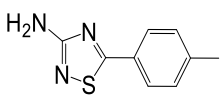
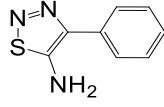
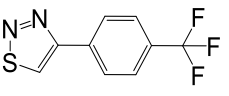
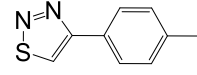
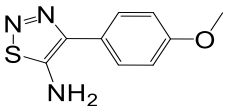
## C.2 Development of the structure activity relationship (SAR) for 3CC7

When comparing the structural similarities between compounds **6CC3** with **3CC7**, both of these molecules are composed of two aromatic rings with an amine substituted heterocycle and a phenyl ring substituted with a small hydrophobic moiety. We sought to further explore analogs of the more potent hit compound **3CC7**. Commercially available analogs of 3CC7 was limited to only a few molecules and no analogs of **6CC3** were explored. To compare analogs we used an NMR-based method 3 Peak Affinity (3PA).<sup>166</sup> This 3PA approach uses the average value in Hz of 3 HSQC peaks (T398, A390, Q391) in the presence of compound (500  $\mu$ M or 1mM). The peaks for these residues (T398, A390, Q391) had some of the largest chemical shift

perturbations upon the addition of compounds with IBD. **3CC7** has a 3PA (500  $\mu$ M) = 34.7Hz and 3PA (1mM) = 56.7Hz.

We first explored similar compounds that were present in our fragment screen to understand the critical features of **3CC7** binding to IBD. **3BB6**, **6HH5**, **2CC9**, and **3EE11** had similar structures with limited to no binding to IBD. **3EE11** was the most potent analog that was in our screen with a 3PA (1mM) = 8.46Hz. **3EE11** had a different arrangement of the thiadiazole substituted with an amine and a methyl in place of the trifluoromethyl. Despite these differences **3EE11** still exhibited some binding to IBD but was almost 7-fold weaker binding than **3CC7**.

We then pursued some small molecule analogs of **3CC7** that were commercially available. **3CC7\_1** with the absence of the trifluoromethyl had no binding to IBD demonstrating the importance of this hydrophobic region of **3CC7**. Loss of the amine substitution of the thiadiazole (**3CC7\_2**) also exhibited 6-fold decrease in binding. Interestingly, replacement of the trifluoromethyl with methyl (**3CC7\_3**) demonstrated no significant change in binding. Therefore, we sought to explore the trifluoromethyl with methoxy (**3CC7\_4**) and found it was about 2-fold weaker than **3CC7**. This limited SAR analysis demonstrates the importance of the amine in the thiadiazole moiety and the requirement of the hydrophobic interactions at the para-position of the phenyl ring. Based on the comparison with **6CC3** there is potentially more room near the phenyl ring region of **3CC7** and linkers between these two rings should be explored.

Compound	Structure	3PA [500uM] (Hz)	3PA [1mM] (Hz)
3CC7		34.7	56.7
3BB6			4.76
6HH5			5.4
2CC9			8.46
3EE11			1.43
3CC7_1		2	
3CC7_2		5.9	
3CC7_3		5.46	
3CC7_4		14.59	

**Table B.1 SAR and 3PA values for 3CC7.**

Binding of these inhibitors to IBD is quantified by the 3PA method by averaging 3 chemical shift perturbations corresponding to residues T398, A390, Q391.

## D. Conclusion

We performed fragment-based screening of IBD with a library of chemically diverse small molecules. This effort resulted in the identification of two chemically distinct small molecules. The binding site of **3CC7** on IBD was mapped through the assignment of  $^1\text{H}$ - $^{15}\text{N}$  chemical shifts, quantifying the magnitude of chemical shift perturbations using **Equation B.1**, and mapping the magnitude of the chemical shift perturbations onto the solution structure of IBD. The binding site was localized to the IBM1 binding site of MLL onto IBD. These small molecules were too weak to test in the AlphaLISA assay and require further optimization of potency before inhibition of the menin-MLL interaction with IBD in biochemical assays can be determined. Limited **3CC7** analog SAR was explored via commercially available analogs. Therefore further investigation will require medicinal chemistry efforts to explore analogs.

This study demonstrates that small molecules can bind to the IBM1 binding site of MLL onto IBD. Based on mutagenesis analysis of the IBM1 site with F148A resulting in the abrogation of the menin-MLL association with IBD we expect inhibition of this IBM1 interaction should mimic these results. Additionally in previous studies<sup>141</sup> it was determined the peptide of IBM2 was sufficient to disrupt the menin-MLL interaction with IBD. This suggests that small compounds or peptides can bind to both of these sites on IBD that interface with the MLL binding motifs. Additionally, this provides an alternative approach to disrupt the menin-MLL-LEDGF ternary complex compared to targeting the menin-MLL interaction.<sup>66,68</sup>

## Appendix C: Table of Crystallography Statistics

Compound	<b>BD615</b>	<b>BD688</b>	<b>BD712</b>	<b>BD964</b>	<b>BD1044</b>	<b>BD1119</b>
Space group	P212121	P212121	P212121	P212121	P212121	P212121
<b>Cell Dimensions</b>						
a, b, c (Å)	48.1, 79.7, 125.0	49.3, 80.1, 124.6	49.2, 79.9, 124.2	48.4, 79.8, 124.9	47.8, 80.4, 124.3	48.0, 80.4, 123.9
Solvent (%)	44.0	45.5	45.1	44.4	44.0	44.0
Resolution (Å)	1.95 (1.98-1.95)	1.50 (1.53-1.50)	1.62 (1.65-1.62)	2.04 (2.10-2.04)	2.21 (2.25-2.21)	2.15 (2.19-2.15)
Unique reflections	34975 (1690)	78547 (3670)	63035 (3097)	46273 (2570)	26109 (1267)	26733 (1295)
Total Reflections	236421	534936	457523	256955	184564	194211
Rsym	0.142 (0.686)	0.081 (0.582)	0.114 (0.909)	0.133 (0.779)	0.147 (0.930)	0.093 (0.811)
I/Sigma*I	20.59 (2.23)	24.24 (2.32)	20.31 (2.15)	15.9 (1.67)	20.33 (1.82)	22.77 (2.06)
Completeness (%)	97.3 (96.2)	98.4 (93.5)	99.9 (99.5)	88.1 (99.0)	99.9 (100)	98.8 (96.2)
Redundancy	6.8 (6.8)	6.8 (6.0)	7.3 (7.1)	5.6 (6.7)	7.1 (6.4)	7.3 (6.6)
<b>Refinement</b>						
Rwork/Rfree	17.9/21.3	15.4/17.9	17.7/20.3	21.5/27.9	19.9/24.3	21.42/26.93
<b>No. atoms</b>						
Protein	3640	3694	3736	3612	3586	3505
Water	250	601	579	114	76	104
Mean B-factor (Å <sup>2</sup> )	31.86	15.69	15.46	47.32	46.56	41.07
<b>R.m.s. Dev.</b>						
Bond lengths (Å)	0.016	0.016	0.016	0.016	0.018	0.018
Bond angles(°)	1.68	1.77	1.67	1.59	1.96	1.80
<b>Ramachandran plot</b>						
Most favored regions (%)	98.06	98.06	97.65	98.06	96.52	95.19
Additional allowed regions (%)	1.94	1.94	2.35	1.94	3.48	4.81

## References

- 1 Pollock, J. *et al.* Rational Design of Orthogonal Multipolar Interactions with Fluorine in Protein-Ligand Complexes. *J Med Chem* **58**, 7465-7474, doi:10.1021/acs.jmedchem.5b00975 (2015).
- 2 Guenther, M. G. *et al.* Global and Hox-specific roles for the MLL1 methyltransferase. *Proc Natl Acad Sci U S A* **102**, 8603-8608, doi:10.1073/pnas.0503072102 (2005).
- 3 Dou, Y. *et al.* Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**, 873-885, doi:10.1016/j.cell.2005.04.031 (2005).
- 4 Milne, T. A. *et al.* MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell* **10**, 1107-1117 (2002).
- 5 Steffen, P. A. & Ringrose, L. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nat Rev Mol Cell Biol* **15**, 340-356, doi:10.1038/nrm3789 (2014).
- 6 Hess, J. L., Yu, B. D., Li, B., Hanson, R. & Korsmeyer, S. J. Defects in yolk sac hematopoiesis in Mll-null embryos. *Blood* **90**, 1799-1806 (1997).
- 7 Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. & Korsmeyer, S. J. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505-508, doi:10.1038/378505a0 (1995).
- 8 Hsieh, J. J., Ernst, P., Erdjument-Bromage, H., Tempst, P. & Korsmeyer, S. J. Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. *Mol Cell Biol* **23**, 186-194 (2003).
- 9 Yokoyama, A. & Cleary, M. L. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell* **14**, 36-46, doi:10.1016/j.ccr.2008.05.003 (2008).



- 10 Jin, S. *et al.* c-Myb binds MLL through menin in human leukemia cells and is an important driver of MLL-associated leukemogenesis. *J Clin Invest* **120**, 593-606, doi:10.1172/jci38030 (2010).
- 11 Milne, T. A. *et al.* Multiple interactions recruit MLL1 and MLL1 fusion proteins to the HOXA9 locus in leukemogenesis. *Mol Cell* **38**, 853-863, doi:10.1016/j.molcel.2010.05.011 (2010).
- 12 Chen, Y. X. *et al.* The tumor suppressor menin regulates hematopoiesis and myeloid transformation by influencing Hox gene expression. *Proc Natl Acad Sci U S A* **103**, 1018-1023, doi:10.1073/pnas.0510347103 (2006).
- 13 Dou, Y. *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* **13**, 713-719, doi:10.1038/nsmb1128 (2006).
- 14 Nakamura, T. *et al.* ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell* **10**, 1119-1128 (2002).
- 15 Terranova, R., Agherbi, H., Boned, A., Meresse, S. & Djabali, M. Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. *Proc Natl Acad Sci U S A* **103**, 6629-6634, doi:10.1073/pnas.0507425103 (2006).
- 16 Sims, R. J., 3rd, Nishioka, K. & Reinberg, D. Histone lysine methylation: a signature for chromatin function. *Trends Genet* **19**, 629-639, doi:10.1016/j.tig.2003.09.007 (2003).
- 17 Santos-Rosa, H. *et al.* Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407-411, doi:10.1038/nature01080 (2002).
- 18 Bernstein, B. E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169-181, doi:10.1016/j.cell.2005.01.001 (2005).
- 19 Pineault, N., Helgason, C. D., Lawrence, H. J. & Humphries, R. K. Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. *Exp Hematol* **30**, 49-57 (2002).
- 20 Maillard, I. *et al.* Menin regulates the function of hematopoietic stem cells and lymphoid progenitors. *Blood* **113**, 1661-1669, doi:10.1182/blood-2009-01-135012 (2009).
- 21 Wang, P. *et al.* Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol Cell Biol* **29**, 6074-6085, doi:10.1128/mcb.00924-09 (2009).
- 22 Dou, Y. & Hess, J. L. Mechanisms of transcriptional regulation by MLL and its disruption in acute leukemia. *Int J Hematol* **87**, 10-18, doi:10.1007/s12185-007-0009-8 (2008).
- 23 Slany, R. K. The molecular mechanics of mixed lineage leukemia. *Oncogene*, doi:10.1038/onc.2016.30 (2016).

- 24 Slany, R. K. When epigenetics kills: MLL fusion proteins in leukemia. *Hematol Oncol* **23**, 1-9, doi:10.1002/hon.739 (2005).
- 25 So, C. W., Lin, M., Ayton, P. M., Chen, E. H. & Cleary, M. L. Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. *Cancer Cell* **4**, 99-110 (2003).
- 26 Pui, C. H. *et al.* Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. *Lancet* **359**, 1909-1915, doi:10.1016/s0140-6736(02)08782-2 (2002).
- 27 Liu, H., Cheng, E. H. & Hsieh, J. J. MLL fusions: pathways to leukemia. *Cancer Biol Ther* **8**, 1204-1211 (2009).
- 28 Rowley, J. D. & Olney, H. J. in *Genes Chromosomes Cancer* Vol. 33 331-345 (2002).
- 29 Cox, M. C. *et al.* Chromosomal aberration of the 11q23 locus in acute leukemia and frequency of MLL gene translocation: results in 378 adult patients. *Am J Clin Pathol* **122**, 298-306, doi:10.1309/rx27-r8gj-qm33-0c22 (2004).
- 30 Sorensen, P. H. *et al.* Molecular rearrangements of the MLL gene are present in most cases of infant acute myeloid leukemia and are strongly correlated with monocytic or myelomonocytic phenotypes. *J Clin Invest* **93**, 429-437, doi:10.1172/jci116978 (1994).
- 31 Lavalley, V. P. *et al.* The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. *Nat Genet* **47**, 1030-1037, doi:10.1038/ng.3371 (2015).
- 32 Dimartino, J. F. & Cleary, M. L. Mll rearrangements in haematological malignancies: lessons from clinical and biological studies. *Br J Haematol* **106**, 614-626 (1999).
- 33 Popovic, R. & Zeleznik-Le, N. J. MLL: how complex does it get? *J Cell Biochem* **95**, 234-242, doi:10.1002/jcb.20430 (2005).
- 34 Hess, J. L. MLL: a histone methyltransferase disrupted in leukemia. *Trends Mol Med* **10**, 500-507, doi:10.1016/j.molmed.2004.08.005 (2004).
- 35 Meyer, C. *et al.* The MLL recombinome of acute leukemias. *Leukemia* **20**, 777-784, doi:10.1038/sj.leu.2404150 (2006).
- 36 Ono, R., Nosaka, T. & Hayashi, Y. Roles of a trithorax group gene, MLL, in hematopoiesis. *Int J Hematol* **81**, 288-293, doi:10.1532/ijh97.04196 (2005).
- 37 Ayton, P. M. & Cleary, M. L. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes Dev* **17**, 2298-2307, doi:10.1101/gad.1111603 (2003).
- 38 Zeisig, B. B. *et al.* Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Mol Cell Biol* **24**, 617-628 (2004).

- 39 Krivtsov, A. V. & Armstrong, S. A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat Rev Cancer* **7**, 823-833, doi:10.1038/nrc2253 (2007).
- 40 So, C. W., Karsunky, H., Wong, P., Weissman, I. L. & Cleary, M. L. Leukemic transformation of hematopoietic progenitors by MLL-GAS7 in the absence of Hoxa7 or Hoxa9. *Blood* **103**, 3192-3199, doi:10.1182/blood-2003-10-3722 (2004).
- 41 He, S. *et al.* Menin-MLL inhibitors block oncogenic transformation by MLL-fusion proteins in a fusion partner-independent manner. *Leukemia* **30**, 508-513, doi:10.1038/leu.2015.144 (2016).
- 42 Thiel, A. T. *et al.* MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. *Cancer Cell* **17**, 148-159, doi:10.1016/j.ccr.2009.12.034 (2010).
- 43 Wang, Q. F. *et al.* MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood* **117**, 6895-6905, doi:10.1182/blood-2010-12-324699 (2011).
- 44 Lavau, C., Szilvassy, S. J., Slany, R. & Cleary, M. L. Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. *Embo j* **16**, 4226-4237 (1997).
- 45 Wei, J. *et al.* Microenvironment determines lineage fate in a human model of MLL-AF9 leukemia. *Cancer Cell* **13**, 483-495, doi:10.1016/j.ccr.2008.04.020 (2008).
- 46 Eguchi, M., Eguchi-Ishimae, M. & Greaves, M. Molecular pathogenesis of MLL-associated leukemias. *Int J Hematol* **82**, 9-20, doi:10.1532/ijh97.05042 (2005).
- 47 Milne, T. A. *et al.* MLL associates specifically with a subset of transcriptionally active target genes. *Proc Natl Acad Sci U S A* **102**, 14765-14770, doi:10.1073/pnas.0503630102 (2005).
- 48 Slany, R. K. The molecular biology of mixed lineage leukemia. *Haematologica* **94**, 984-993, doi:10.3324/haematol.2008.002436 (2009).
- 49 Wong, P., Iwasaki, M., Somervaille, T. C., So, C. W. & Cleary, M. L. Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. *Genes Dev* **21**, 2762-2774, doi:10.1101/gad.1602107 (2007).
- 50 Armstrong, S. A. *et al.* MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* **30**, 41-47, doi:10.1038/ng765 (2002).
- 51 Owens, B. M. & Hawley, R. G. HOX and non-HOX homeobox genes in leukemic hematopoiesis. *Stem Cells* **20**, 364-379, doi:10.1634/stemcells.20-5-364 (2002).
- 52 Argiropoulos, B. & Humphries, R. K. Hox genes in hematopoiesis and leukemogenesis. *Oncogene* **26**, 6766-6776, doi:10.1038/sj.onc.1210760 (2007).

- 53 Rice, K. L. & Licht, J. D. HOX deregulation in acute myeloid leukemia. *J Clin Invest* **117**, 865-868, doi:10.1172/jci31861 (2007).
- 54 Kroon, E. *et al.* Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *Embo j* **17**, 3714-3725, doi:10.1093/emboj/17.13.3714 (1998).
- 55 Yokoyama, A. *et al.* The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell* **123**, 207-218, doi:10.1016/j.cell.2005.09.025 (2005).
- 56 Thakker, R. V. Multiple endocrine neoplasia type 1 (MEN1) and type 4 (MEN4). *Mol Cell Endocrinol* **386**, 2-15, doi:10.1016/j.mce.2013.08.002 (2014).
- 57 Marx, S. J. Molecular genetics of multiple endocrine neoplasia types 1 and 2. *Nat Rev Cancer* **5**, 367-375, doi:10.1038/nrc1610 (2005).
- 58 Caslini, C. *et al.* Interaction of MLL amino terminal sequences with menin is required for transformation. *Cancer Res* **67**, 7275-7283, doi:10.1158/0008-5472.can-06-2369 (2007).
- 59 Grembecka, J., Belcher, A. M., Hartley, T. & Cierpicki, T. Molecular basis of the mixed lineage leukemia-menin interaction: implications for targeting mixed lineage leukemias. *J Biol Chem* **285**, 40690-40698, doi:10.1074/jbc.M110.172783 (2010).
- 60 Hughes, C. M. *et al.* Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. *Mol Cell* **13**, 587-597 (2004).
- 61 Liedtke, M. & Cleary, M. L. Therapeutic targeting of MLL. *Blood* **113**, 6061-6068, doi:10.1182/blood-2008-12-197061 (2009).
- 62 Huang, J. *et al.* The same pocket in menin binds both MLL and JUND but has opposite effects on transcription. *Nature* **482**, 542-546, doi:10.1038/nature10806 (2012).
- 63 Llano, M., Morrison, J. & Poeschla, E. M. Virological and cellular roles of the transcriptional coactivator LEDGF/p75. *Curr Top Microbiol Immunol* **339**, 125-146, doi:10.1007/978-3-642-02175-6\_7 (2009).
- 64 Gray, F. L., Murai, M. J., Grembecka, J. & Cierpicki, T. Detection of disordered regions in globular proteins using (1)(3)C-detected NMR. *Protein Sci* **21**, 1954-1960, doi:10.1002/pro.2174 (2012).
- 65 Murai, M. J., Chruszcz, M., Reddy, G., Grembecka, J. & Cierpicki, T. Crystal structure of menin reveals binding site for mixed lineage leukemia (MLL) protein. *J Biol Chem* **286**, 31742-31748, doi:10.1074/jbc.M111.258186 (2011).
- 66 Shi, A. *et al.* Structural insights into inhibition of the bivalent menin-MLL interaction by small molecules in leukemia. *Blood* **120**, 4461-4469, doi:10.1182/blood-2012-05-429274 (2012).

- 67 Cierpicki, T. & Grembecka, J. Challenges and opportunities in targeting the menin-MLL interaction. *Future Med Chem* **6**, 447-462, doi:10.4155/fmc.13.214 (2014).
- 68 Grembecka, J. *et al.* Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia. *Nat. Chem. Biol.* **8**, 277-284, doi:10.1038/nchembio.773 nchembio.773 [pii] (2012).
- 69 Olsen, J. A. *et al.* Fluorine interactions at the thrombin active site: protein backbone fragments H-C(alpha)-C=O comprise a favorable C-F environment and interactions of C-F with electrophiles. *Chembiochem* **5**, 666-675, doi:10.1002/cbic.200300907 (2004).
- 70 Borkin, D. *et al.* Pharmacologic inhibition of the Menin-MLL interaction blocks progression of MLL leukemia in vivo. *Cancer Cell* **27**, 589-602, doi:10.1016/j.ccell.2015.02.016 (2015).
- 71 Borkin, D. *et al.* Property Focused Structure-Based Optimization of Small Molecule Inhibitors of the Protein-Protein Interaction between Menin and Mixed Lineage Leukemia (MLL). *J Med Chem* **59**, 892-913, doi:10.1021/acs.jmedchem.5b01305 (2016).
- 72 Manka, J. *et al.* in *Probe Reports from the NIH Molecular Libraries Program* (National Center for Biotechnology Information (US), 2010).
- 73 He, S. *et al.* High-affinity small-molecule inhibitors of the menin-mixed lineage leukemia (MLL) interaction closely mimic a natural protein-protein interaction. *J Med Chem* **57**, 1543-1556, doi:10.1021/jm401868d (2014).
- 74 Zhou, H. *et al.* Structure-based design of high-affinity macrocyclic peptidomimetics to block the menin-mixed lineage leukemia 1 (MLL1) protein-protein interaction. *J Med Chem* **56**, 1113-1123, doi:10.1021/jm3015298 (2013).
- 75 Zhong, H. J. *et al.* Structure-based screening and optimization of cytosine derivatives as inhibitors of the menin-MLL interaction. *Chem Commun (Camb)*, doi:10.1039/c6cc01079b (2016).
- 76 Li, L. *et al.* Discovery of two aminoglycoside antibiotics as inhibitors targeting the menin-mixed lineage leukaemia interface. *Bioorg Med Chem Lett* **24**, 2090-2093, doi:10.1016/j.bmcl.2014.03.055 (2014).
- 77 Yokoyama, A. *et al.* Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Mol Cell Biol* **24**, 5639-5649, doi:10.1128/mcb.24.13.5639-5649.2004 (2004).
- 78 Takeda, S. *et al.* HGF-MET signals via the MLL-ETS2 complex in hepatocellular carcinoma. *J Clin Invest* **123**, 3154-3165, doi:10.1172/jci65566 (2013).
- 79 Gallo, M. *et al.* A tumorigenic MLL-homeobox network in human glioblastoma stem cells. *Cancer Res* **73**, 417-427, doi:10.1158/0008-5472.can-12-1881 (2013).

- 80 Ansari, K. I., Kasiri, S., Mishra, B. P. & Mandal, S. S. Mixed lineage leukaemia-4 regulates cell-cycle progression and cell viability and its depletion suppresses growth of xenografted tumour in vivo. *Br J Cancer* **107**, 315-324, doi:10.1038/bjc.2012.263 (2012).
- 81 Imachi, H. *et al.* Menin, a product of the MEN1 gene, binds to estrogen receptor to enhance its activity in breast cancer cells: possibility of a novel predictive factor for tamoxifen resistance. *Breast Cancer Res Treat* **122**, 395-407, doi:10.1007/s10549-009-0581-0 (2010).
- 82 Huo, H., Magro, P. G., Pietsch, E. C., Patel, B. B. & Scotto, K. W. Histone methyltransferase MLL1 regulates MDR1 transcription and chemoresistance. *Cancer Res* **70**, 8726-8735, doi:10.1158/0008-5472.can-10-0755 (2010).
- 83 Andreeff, M. *et al.* HOX expression patterns identify a common signature for favorable AML. *Leukemia* **22**, 2041-2047, doi:10.1038/leu.2008.198 (2008).
- 84 Golub, T. R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531-537 (1999).
- 85 Collins, C. *et al.* C/EBPalpha is an essential collaborator in Hoxa9/Meis1-mediated leukemogenesis. *Proc Natl Acad Sci U S A* **111**, 9899-9904, doi:10.1073/pnas.1402238111 (2014).
- 86 Li, B. *et al.* Hit-to-lead optimization and kinase selectivity of imidazo[1,2-a]quinoxalin-4-amine derived JNK1 inhibitors. *Bioorg Med Chem Lett* **23**, 5217-5222, doi:10.1016/j.bmcl.2013.06.087 (2013).
- 87 Bohm, H. J. *et al.* Fluorine in medicinal chemistry. *Chembiochem* **5**, 637-643, doi:10.1002/cbic.200301023 (2004).
- 88 Muller, K., Faeh, C. & Diederich, F. Fluorine in pharmaceuticals: looking beyond intuition. *Science* **317**, 1881-1886, doi:10.1126/science.1131943 (2007).
- 89 Purser, S., Moore, P. R., Swallow, S. & Gouverneur, V. Fluorine in medicinal chemistry. *Chem Soc Rev* **37**, 320-330, doi:10.1039/b610213c (2008).
- 90 Bissantz, C., Kuhn, B. & Stahl, M. A medicinal chemist's guide to molecular interactions. *J Med Chem* **53**, 5061-5084, doi:10.1021/jm100112j (2010).
- 91 Olsen, J. A. *et al.* A fluorine scan of thrombin inhibitors to map the fluorophilicity/fluorophobicity of an enzyme active site: Evidence for C-F center dot center dot center dot C=O interactions. *Angewandte Chemie-International Edition* **42**, 2507-2511, doi:DOI 10.1002/anie.200351268 (2003).
- 92 Vulpetti, A., Hommel, U., Landrum, G., Lewis, R. & Dalvit, C. Design and NMR-based screening of LEF, a library of chemical fragments with different local environment of fluorine. *J Am Chem Soc* **131**, 12949-12959, doi:10.1021/ja905207t (2009).
- 93 Vulpetti, A., Schiering, N. & Dalvit, C. Combined use of computational chemistry, NMR screening, and X-ray crystallography for identification and characterization of

- fluorophilic protein environments. *Proteins* **78**, 3281-3291, doi:10.1002/prot.22836 (2010).
- 94 Paulini, R., Muller, K. & Diederich, F. Orthogonal multipolar interactions in structural chemistry and biology. *Angew Chem Int Ed Engl* **44**, 1788-1805, doi:10.1002/anie.200462213 (2005).
- 95 The PyMOL Molecular Graphics System v. Version 1.2r3pre, Schrödinger, LLC.
- 96 Biffinger, J. C., Kim, H. W. & DiMugno, S. G. The polar hydrophobicity of fluorinated compounds. *Chembiochem* **5**, 622-627, doi:10.1002/cbic.200300910 (2004).
- 97 Murray, J. *et al.* Tailoring small molecules for an allosteric site on procaspase-6. *ChemMedChem* **9**, 73-77, 72, doi:10.1002/cmde.201300424 (2014).
- 98 Eidam, O. *et al.* Fragment-guided design of subnanomolar beta-lactamase inhibitors active in vivo. *Proc Natl Acad Sci U S A* **109**, 17448-17453, doi:10.1073/pnas.1208337109 (2012).
- 99 Marschalek, R. Mechanisms of leukemogenesis by MLL fusion proteins. *Br J Haematol* **152**, 141-154, doi:10.1111/j.1365-2141.2010.08459.x (2011).
- 100 Krivtsov, A. V. *et al.* Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442**, 818-822, doi:10.1038/nature04980 (2006).
- 101 Azzarito, V., Long, K., Murphy, N. S. & Wilson, A. J. Inhibition of alpha-helix-mediated protein-protein interactions using designed molecules. *Nat Chem* **5**, 161-173, doi:10.1038/nchem.1568 (2013).
- 102 Nero, T. L., Morton, C. J., Holien, J. K., Wielens, J. & Parker, M. W. Oncogenic protein interfaces: small molecules, big challenges. *Nat Rev Cancer* **14**, 248-262, doi:10.1038/nrc3690 (2014).
- 103 Bai, L. & Wang, S. Targeting apoptosis pathways for new cancer therapeutics. *Annu Rev Med* **65**, 139-155, doi:10.1146/annurev-med-010713-141310 (2014).
- 104 Li, B. E., Gan, T., Meyerson, M., Rabbitts, T. H. & Ernst, P. Distinct pathways regulated by menin and by MLL1 in hematopoietic stem cells and developing B cells. *Blood* **122**, 2039-2046, doi:10.1182/blood-2013-03-486647 (2013).
- 105 Fry, D. C. Protein-protein interactions as targets for small molecule drug discovery. *Biopolymers* **84**, 535-552, doi:10.1002/bip.20608 (2006).
- 106 Buchwald, P. Small-molecule protein-protein interaction inhibitors: therapeutic potential in light of molecular size, chemical space, and ligand binding efficiency considerations. *IUBMB Life* **62**, 724-731, doi:10.1002/iub.383 (2010).
- 107 Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450**, 1001-1009, doi:10.1038/nature06526 (2007).

- 108 Morelli, X., Bourgeas, R. & Roche, P. Chemical and structural lessons from recent successes in protein-protein interaction inhibition (2P2I). *Curr Opin Chem Biol* **15**, 475-481, doi:10.1016/j.cbpa.2011.05.024 (2011).
- 109 Arkin, M. R. & Whitty, A. The road less traveled: modulating signal transduction enzymes by inhibiting their protein-protein interactions. *Curr Opin Chem Biol* **13**, 284-290, doi:10.1016/j.cbpa.2009.05.125 (2009).
- 110 Smith, M. C. & Gestwicki, J. E. Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Rev Mol Med* **14**, e16, doi:10.1017/erm.2012.10 (2012).
- 111 Tse, C. *et al.* ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res* **68**, 3421-3428, doi:10.1158/0008-5472.can-07-5836 (2008).
- 112 Ray-Coquard, I. *et al.* Effect of the MDM2 antagonist RG7112 on the P53 pathway in patients with MDM2-amplified, well-differentiated or dedifferentiated liposarcoma: an exploratory proof-of-mechanism study. *Lancet Oncol* **13**, 1133-1140, doi:10.1016/s1470-2045(12)70474-6 (2012).
- 113 Senter, T. *et al.* in *Probe Reports from the NIH Molecular Libraries Program* (National Center for Biotechnology Information (US), 2010).
- 114 Potashman, M. H. & Duggan, M. E. Covalent modifiers: an orthogonal approach to drug design. *J Med Chem* **52**, 1231-1246, doi:10.1021/jm8008597 (2009).
- 115 Robertson, J. G. Mechanistic basis of enzyme-targeted drugs. *Biochemistry* **44**, 5561-5571, doi:10.1021/bi050247e (2005).
- 116 Lipinski, C. & Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **432**, 855-861, doi:10.1038/nature03193 (2004).
- 117 Erve, J. C. Chemical toxicology: reactive intermediates and their role in pharmacology and toxicology. *Expert Opin Drug Metab Toxicol* **2**, 923-946, doi:10.1517/17425255.2.6.923 (2006).
- 118 Uetrecht, J. Immune-mediated adverse drug reactions. *Chem Res Toxicol* **22**, 24-34, doi:10.1021/tx800389u (2009).
- 119 Singh, J., Petter, R. C., Baillie, T. A. & Whitty, A. The resurgence of covalent drugs. *Nat Rev Drug Discov* **10**, 307-317, doi:10.1038/nrd3410 (2011).
- 120 Liu, Q. *et al.* Developing irreversible inhibitors of the protein kinase cysteinome. *Chem Biol* **20**, 146-159, doi:10.1016/j.chembiol.2012.12.006 (2013).
- 121 Wong, K. K. *et al.* A phase I study with neratinib (HKI-272), an irreversible pan ErbB receptor tyrosine kinase inhibitor, in patients with solid tumors. *Clin Cancer Res* **15**, 2552-2558, doi:10.1158/1078-0432.ccr-08-1978 (2009).



- 122 Tummino, P. J. & Copeland, R. A. Residence time of receptor-ligand complexes and its effect on biological function. *Biochemistry* **47**, 5481-5492, doi:10.1021/bi8002023 (2008).
- 123 Herman, S. E. *et al.* Bruton tyrosine kinase represents a promising therapeutic target for treatment of chronic lymphocytic leukemia and is effectively targeted by PCI-32765. *Blood* **117**, 6287-6296, doi:10.1182/blood-2011-01-328484 (2011).
- 124 Troxler, T. *et al.* Discovery of novel indolinone-based, potent, selective and brain penetrant inhibitors of LRRK2. *Bioorg Med Chem Lett* **23**, 4085-4090, doi:10.1016/j.bmcl.2013.05.054 (2013).
- 125 Evans, E. K. *et al.* Inhibition of Btk with CC-292 provides early pharmacodynamic assessment of activity in mice and humans. *J Pharmacol Exp Ther* **346**, 219-228, doi:10.1124/jpet.113.203489 (2013).
- 126 Nakayama, S. *et al.* A zone classification system for risk assessment of idiosyncratic drug toxicity using daily dose and covalent binding. *Drug Metab Dispos* **37**, 1970-1977, doi:10.1124/dmd.109.027797 (2009).
- 127 Shin, N. Y., Liu, Q., Stamer, S. L. & Liebler, D. C. Protein targets of reactive electrophiles in human liver microsomes. *Chem Res Toxicol* **20**, 859-867, doi:10.1021/tx700031r (2007).
- 128 Flanagan, M. E. *et al.* Chemical and computational methods for the characterization of covalent reactive groups for the prospective design of irreversible inhibitors. *J Med Chem* **57**, 10072-10079, doi:10.1021/jm501412a (2014).
- 129 Lanning, B. R. *et al.* A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat Chem Biol* **10**, 760-767, doi:10.1038/nchembio.1582 (2014).
- 130 Wirth, S. M. Afatinib in Non-Small Cell Lung Cancer. *J Adv Pract Oncol* **6**, 448-455 (2015).
- 131 Jain, N. & O'Brien, S. Targeted therapies for CLL: Practical issues with the changing treatment paradigm. *Blood Rev* **30**, 233-244, doi:10.1016/j.blre.2015.12.002 (2016).
- 132 Copeland, R. A., Pompliano, D. L. & Meek, T. D. Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* **5**, 730-739, doi:10.1038/nrd2082 (2006).
- 133 Gonzales, A. J. *et al.* Antitumor activity and pharmacokinetic properties of PF-00299804, a second-generation irreversible pan-erbB receptor tyrosine kinase inhibitor. *Mol Cancer Ther* **7**, 1880-1889, doi:10.1158/1535-7163.mct-07-2232 (2008).
- 134 Wissner, A. *et al.* Synthesis and structure-activity relationships of 6,7-disubstituted 4-anilinoquinoline-3-carbonitriles. The design of an orally active, irreversible inhibitor of the tyrosine kinase activity of the epidermal growth factor receptor (EGFR) and the human epidermal growth factor receptor-2 (HER-2). *J Med Chem* **46**, 49-63, doi:10.1021/jm020241c (2003).

- 135 Dahal, U. P., Obach, R. S. & Gilbert, A. M. Benchmarking in vitro covalent binding burden as a tool to assess potential toxicity caused by nonspecific covalent binding of covalent drugs. *Chem Res Toxicol* **26**, 1739-1745, doi:10.1021/tx400301q (2013).
- 136 Bauman, J. N. *et al.* Can in vitro metabolism-dependent covalent binding data distinguish hepatotoxic from nonhepatotoxic drugs? An analysis using human hepatocytes and liver S-9 fraction. *Chem Res Toxicol* **22**, 332-340, doi:10.1021/tx800407w (2009).
- 137 Zhang, T. *et al.* Discovery of potent and selective covalent inhibitors of JNK. *Chem Biol* **19**, 140-154, doi:10.1016/j.chembiol.2011.11.010 (2012).
- 138 Patricelli, M. P. *et al.* Selective Inhibition of Oncogenic KRAS Output with Small Molecules Targeting the Inactive State. *Cancer Discov* **6**, 316-329, doi:10.1158/2159-8290.cd-15-1105 (2016).
- 139 Zhou, W. *et al.* A structure-guided approach to creating covalent FGFR inhibitors. *Chem Biol* **17**, 285-295, doi:10.1016/j.chembiol.2010.02.007 (2010).
- 140 Martinez Molina, D. *et al.* Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84-87, doi:10.1126/science.1233606 (2013).
- 141 Murai, M. J. *et al.* The same site on the integrase-binding domain of lens epithelium-derived growth factor is a therapeutic target for MLL leukemia and HIV. *Blood* **124**, 3730-3737, doi:10.1182/blood-2014-01-550079 (2014).
- 142 Cierpicki, T. & Grembecka, J. Targeting protein-protein interactions in hematologic malignancies: still a challenge or a great opportunity for future therapies? *Immunol Rev* **263**, 279-301, doi:10.1111/imr.12244 (2015).
- 143 Jochim, A. L. & Arora, P. S. Assessment of helical interfaces in protein-protein interactions. *Mol Biosyst* **5**, 924-926, doi:10.1039/b903202a (2009).
- 144 Remaut, H. & Waksman, G. Protein-protein interaction through beta-strand addition. *Trends Biochem Sci* **31**, 436-444, doi:10.1016/j.tibs.2006.06.007 (2006).
- 145 Poppe, B. *et al.* Expression analyses identify MLL as a prominent target of 11q23 amplification and support an etiologic role for MLL gain of function in myeloid malignancies. *Blood* **103**, 229-235, doi:10.1182/blood-2003-06-2163 (2004).
- 146 Ge, H., Si, Y. & Roeder, R. G. Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *Embo j* **17**, 6723-6729, doi:10.1093/emboj/17.22.6723 (1998).
- 147 Singh, D. P., Fatma, N., Kimura, A., Chylack, L. T., Jr. & Shinohara, T. LEDGF binds to heat shock and stress-related element to activate the expression of stress-related genes. *Biochem Biophys Res Commun* **283**, 943-955, doi:10.1006/bbrc.2001.4887 (2001).
- 148 Daugaard, M. *et al.* Lens epithelium-derived growth factor is an Hsp70-2 regulated guardian of lysosomal stability in human cancer. *Cancer Res* **67**, 2559-2567, doi:10.1158/0008-5472.can-06-4121 (2007).

- 149 Ganapathy, V. & Casiano, C. A. Autoimmunity to the nuclear autoantigen DFS70 (LEDGF): what exactly are the autoantibodies trying to tell us? *Arthritis Rheum* **50**, 684-688, doi:10.1002/art.20095 (2004).
- 150 Ciuffi, A. & Bushman, F. D. Retroviral DNA integration: HIV and the role of LEDGF/p75. *Trends Genet* **22**, 388-395, doi:10.1016/j.tig.2006.05.006 (2006).
- 151 Hussey, D. J., Moore, S., Nicola, M. & Dobrovic, A. Fusion of the NUP98 gene with the LEDGF/p52 gene defines a recurrent acute myeloid leukemia translocation. *BMC Genet* **2**, 20 (2001).
- 152 Morerio, C. *et al.* t(9;11)(p22;p15) with NUP98-LEDGF fusion gene in pediatric acute myeloid leukemia. *Leuk Res* **29**, 467-470, doi:10.1016/j.leukres.2004.09.002 (2005).
- 153 Huang, T. S. *et al.* LEDGF/p75 has increased expression in blasts from chemotherapy-resistant human acute myelogenous leukemia patients and protects leukemia cells from apoptosis in vitro. *Mol Cancer* **6**, 31, doi:10.1186/1476-4598-6-31 (2007).
- 154 Llano, M. *et al.* An essential role for LEDGF/p75 in HIV integration. *Science* **314**, 461-464, doi:10.1126/science.1132319 (2006).
- 155 Maertens, G. *et al.* LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J Biol Chem* **278**, 33528-33539, doi:10.1074/jbc.M303594200 (2003).
- 156 Cherepanov, P., Devroe, E., Silver, P. A. & Engelman, A. Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. *J Biol Chem* **279**, 48883-48892, doi:10.1074/jbc.M406307200 (2004).
- 157 De Rijck, J. *et al.* Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. *J Virol* **80**, 11498-11509, doi:10.1128/jvi.00801-06 (2006).
- 158 Christ, F. *et al.* Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat Chem Biol* **6**, 442-448, doi:10.1038/nchembio.370 (2010).
- 159 Christ, F. *et al.* Small-molecule inhibitors of the LEDGF/p75 binding site of integrase block HIV replication and modulate integrase multimerization. *Antimicrob Agents Chemother* **56**, 4365-4374, doi:10.1128/aac.00717-12 (2012).
- 160 Cherepanov, P., Ambrosio, A. L., Rahman, S., Ellenberger, T. & Engelman, A. Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. *Proc Natl Acad Sci U S A* **102**, 17308-17313, doi:10.1073/pnas.0506924102 (2005).
- 161 Vandegraaff, N., Devroe, E., Turlure, F., Silver, P. A. & Engelman, A. Biochemical and genetic analyses of integrase-interacting proteins lens epithelium-derived growth factor (LEDGF)/p75 and hepatoma-derived growth factor related protein 2 (HRP2) in

- preintegration complex function and HIV-1 replication. *Virology* **346**, 415-426, doi:10.1016/j.virol.2005.11.022 (2006).
- 162 Christ, F. & Debysier, Z. The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. *Virology* **435**, 102-109, doi:10.1016/j.virol.2012.09.033 (2013).
- 163 Hombrouck, A. *et al.* Virus evolution reveals an exclusive role for LEDGF/p75 in chromosomal tethering of HIV. *PLoS Pathog* **3**, e47, doi:10.1371/journal.ppat.0030047 (2007).
- 164 Mereau, H. *et al.* Impairing MLL-fusion gene-mediated transformation by dissecting critical interactions with the lens epithelium-derived growth factor (LEDGF/p75). *Leukemia* **27**, 1245-1253, doi:10.1038/leu.2013.10 (2013).
- 165 Tugarinov, V. & Kay, L. E. Quantitative NMR studies of high molecular weight proteins: application to domain orientation and ligand binding in the 723 residue enzyme malate synthase G. *J Mol Biol* **327**, 1121-1133 (2003).
- 166 Rogawski, D. *The function of the ASH1L histone methyltransferase in cancer: A chemical biology approach* PhD thesis, University of Michigan, (2016).
- 167 ( Origin (OriginLab, Northampton, MA)).
- 168 Otwinowski, Z. & Minor, W. in *Macromolecular Crystallography, Pt A Vol. 276 Methods in Enzymology* 307-326 (1997).
- 169 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255, doi:10.1107/s0907444996012255 (1997).
- 170 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, doi:10.1107/s0907444904019158 (2004).
- 171 Collaborative Computational Project, N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-763, doi:10.1107/s0907444994003112 (1994).
- 172 Painter, J. & Merritt, E. A. TLSMD web server for the generation of multi-group TLS models. *J. Appl. Cryst.* **39**, 109-111 (2006).
- 173 Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-383, doi:10.1093/nar/gkm216 (2007).
- 174 Yang, H. *et al.* Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* **60**, 1833-1839, doi:10.1107/s0907444904019419 (2004).
- 175 Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293 (1995).

- 176 Lee, W., Westler, W. M., Bahrami, A., Eghbalnia, H. R. & Markley, J. L. PINE-SPARKY: graphical interface for evaluating automated probabilistic peak assignments in protein NMR spectroscopy. *Bioinformatics* **25**, 2085-2087, doi:10.1093/bioinformatics/btp345 (2009).