

DISCOVERY OF SMALL MOLECULE AND PEPTIDE-BASED LIGANDS FOR THE
METHYL-LYSINE BINDING PROTEINS 53BP1 AND PHF1/PHF19

Michael Thomas Perfetti

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Division of Chemical Biology and Medicinal Chemistry in the Doctoral Program of the UNC Eshelman School of Pharmacy.

Chapel Hill
2015

Approved by:
Stephen Frye
David Lawrence
Albert Bowers
Brian Strahl
Greg Gang Wang
Andrew Lee

© 2015
Michael Thomas Perfetti
ALL RIGHTS RESERVED

ABSTRACT

Michael Thomas Perfetti: Discovery of Small Molecule and Peptide-based Ligands for the Methyl-Lysine Binding Proteins 53BP1 and PHF1/PHF19
(Under the direction of Stephen V. Frye)

Improving the understanding of the role of chromatin regulators in the initiation, development, and suppression of cancer and other devastating diseases is critical, as they are integral players in the regulation of DNA integrity and gene expression. Developing chemical tools for histone binding proteins that possess cellular activity will allow for further elucidation of the specific function of this class of histone regulating proteins. This research specifically targeted two different classes of Tudor domain containing histone binding proteins that are directly involved in the DNA damage response and modulation of gene transcription activities. The first methyl-lysine binding protein targeted was 53BP1, which is a DNA damage response protein. 53BP1 uses a tandem tudor domain (TTD) to recognize histone H4 dimethylated on lysine 20 (H4K20me₂), a post-translational modification (PTM) induced by double-strand DNA breaks. Through a cross-screening approach, two different small molecule ligands were identified for the 53BP1 TTD. Medicinal chemistry and structure-based design techniques were used, in addition to the development of a fragment-based screen, toward the goal of optimizing the potency and selectivity of the identified small molecule ligands. The later small molecule ligand, **61** (UNC2170), was further optimized as a micromolar inhibitor of 53BP1, which demonstrated a 17-fold selectivity

for 53BP1 TTD as compared to other screened methyl-lysine (Kme) binding proteins. Structural studies revealed that the *N-tert*-butyl amine of **61** (UNC2170) anchors the compound in the Kme binding pocket of 53BP1 TTD, making it competitive with endogenous Kme substrates. X-ray crystallography demonstrated that **61** (UNC2170) bound at the interface between two tudor domains of a 53BP1 TTD dimer. Importantly, this compound functions as a 53BP1 antagonist in cellular lysates and shows cellular activity by suppressing class switch recombination, a process that requires a functional 53BP1 tudor domain. These results demonstrate that **61** (UNC2170) is a functionally active, fragment-like ligand for 53BP1 and the first selective small molecule ligand discovered for 53BP1 TTD.

Additional research efforts were focused toward development of peptide-based ligands for the histone binding proteins PHF1 and PHF19. These proteins play an active role in the modulation of PRC2 activity that in turn controls various gene expression and repression outcomes. Structure-based design techniques were used to design peptide-based inhibitors for PHF1 and PHF19. Research efforts were focused on conducting an initial Structure Activity Relationship (SAR) study to develop a short, low molecular weight peptide via truncation of the endogenous histone 3 peptide. Structure based design techniques were used to develop a shorter 6- or 7-mer peptide that possessed various amino acid substitutions that specifically bound within the Tudor domain of PHF1 and PHF19 with emphasis on determining a quaternary amine mimic for the tri-methyl lysine 36 residue. These efforts proved fruitful and a 7-mer peptide (**148**) was developed that was equipotent to the endogenous histone peptide.

Research efforts on both of these targeted proteins demonstrates that structure based design and medicinal chemistry techniques can successfully facilitate the development of small molecule and peptide based ligands for histone binding proteins.

To my family.

ACKNOWLEDGEMENTS

I would like to extend my sincerest gratitude to my family and friends for their continued support, encouragement, and continued love throughout my graduate training. This would have never been possible without their continued support. Most importantly, I would like to thank my Mom, Dad, and Brother (Patty, Tom and David), for always being there and supporting and encouraging me through this experience to achieve my goals. I could never have completed such a journey without you. Thank you!

I would like to thank my dissertation committee for their guidance and continued commitment to my scientific development. To my advisor, Dr. Stephen Frye: You have truly been an amazing mentor and advisor throughout my graduate training and I cannot thank you enough for all of the support, advice, and friendship that you have willingly provided me with over the years. Thank you so very much for helping me to achieve my goals. I wish to thank my committee chair, Dr. David Lawrence, who has always provided insightful advice and constructive humor in regards to on my scientific and professional development and for that, I am very grateful. I thank wish to thank the rest of my committee members, Dr. Brian Strahl, Dr. Greg Wang, and Dr. Albert Bowers, for all of their support and constructive suggestions and conversations during committee meetings, recruitment weekends, and our collaboration on the research projects described in this dissertation.

I wish to thank all of the current and past members of the Frye laboratory and the Center for Integrative Chemical Biology and Drug Discovery. Your support and friendship have made the challenges and obstacles encountered in graduate school easier to endure. I wish to sincerely thank Dr. Lindsey Ingerman, Dr. Brandi Baughman, Dr. Bradley Dickson, Dr. Scott Rothbart, Dr. Jacqueline L. Norris, Dr.

Samantha Pattenden, Mr. Chatura Jayakody, Dr. Nancy Cheng, Dr. Dmitri Kireev, Mr. William Janzen and other members of the CICBDD, Strahl, Bedford, McBride, Mer and SGC Toronto research labs for their contributions to this project. Lindsey, this would not have been possible without your excellent guidance and support, thank you! I extend my sincerest gratitude to Jake Stuckey, Kim Barnash, and Junghyun Lee for working together as a group of graduate students in the Frye Lab and helping each other grow. Barbara Dearry, Thank you so much for keeping me on schedule and motivated in the CICBDD. I wish to say thank you to my close friends Colin O'Banion, Matt Geden, Stephen Capuzzi, Dr. Deepak Jha, Dr. Aaron Nile, Dr. Cindy Gode, Dr. Lance Thurlow, Szymon Karwowski, Dr. Eric Montijo, and Ethan Suttles for their continuous love and support throughout the past few years. You've made this journey a blast and I could not thank you more for being there for me and for all the friendship, encouragement, advice, support during my graduate training.

The University of North Carolina at Chapel Hill has been and will always be my home and I could ask more of this amazing institution over the past five years of my graduate training. Thank you so very much!

PREFACE

This Ph.D. thesis was initiated and completed in the Division of Chemical Biology and Medicinal Chemistry in the Doctoral Program of the UNC Eshelman School of Pharmacy at UNC-Chapel Hill.

Work described in Chapter 2 including several figures and adapted text has been reproduced in part from “Identification of a Fragment-like Small Molecule Ligand for the Methyl-lysine Binding Protein, 53BP1”, (in press) with permission from ACS Chemical Biology. Unpublished work copyright 2015 American Chemical Society."

TABLE OF CONTENTS

LIST OF TABLES.....	xviii
LIST OF FIGURES.....	xix
LIST OF ABBREVIATIONS AND SYMBOLS.....	xxii
CHAPTER 1: DEVELOPMENT OF METHYL-LYSINE INHIBITORS.....	1
Background.....	1
Post-Translational Modifications (PTMs) and their role	
in chromatin regulation.....	3
Classes of methyl-lysine binding proteins.....	5
Interpretation of specific PTMs by methyl-lysine binding proteins.....	5
Tudor binding domain.....	6
Hypothesized outcomes of domain inhibition.....	8
Current state of ligand development for methyl-lysine binding proteins.....	9
Challenges toward development of inhibitory ligands for	
methyl-lysine binding proteins.....	9
Current examples of inhibitory ligands.....	10
Use of Hit discovery techniques to discover inhibitory ligands.....	13

Methyl-lysine binding protein biological screening methods.....	13
Structure-based drug design.....	13
Purpose of this work.....	14
CHAPTER 2: DESIGN, SYNTHESIS, AND CHARACTERIZATION OF SMALL MOLECULE LIGANDS FOR THE CHROMATIN METHYL-LYSINE BINDER, 53BP1.....	
Discovery and biological function of 53BP1.....	15
DNA Repair Pathways, BRCA Mutations, and Cancer.....	17
Use of medicinal chemistry techniques to discover inhibitory ligands for 53BP1.....	19
Purpose of this work.....	20
Benzimidazole-based small molecule ligand of 53BP1, UNC1554.....	22
Results and discussion.....	22
Structure based analysis of 53BP1.....	22
Discovery of UNC1554 hit.....	24
Comparison of UNC1554 to endogenous H4K20me2 chromatin peptide.....	26
Structure-activity relationship (SAR) studies.....	27
Synthetic strategy for SAR studies.....	28

Modification of amine head group.....	29
Modification of linker region.....	31
Modification of benzimidazole aromatic ring.....	33
Conclusions.....	37
Experimental section.....	38
General Procedure for Chemical Synthesis.....	38
Small molecule synthesis.....	40
AlphaScreen assay method.....	88
Isothermal titration calorimetry binding experiments.....	88
Future Work.....	89
Benzamide-based small molecule inhibitors of 53BP1.....	91
Results and discussion.....	91
Discovery of UNC2170 hit.....	91
Comparison to endogenous H4K20me2 chromatin peptide tail.....	92
Structure-activity relationship studies.....	93
Synthetic strategy for SAR studies.....	93
Modification of amine head group.....	93
Modification of linker region.....	98

Modification of aromatic ring.....	100
Disubstituted aromatic analogs of 61 (UNC2170).....	103
Non-phenyl ring based analogs.....	105
Tri-coordinate analogs.....	107
Ugi analogs.....	109
Structural binding analysis.....	110
X-ray co-crystallization experiments	110
Protein NMR binding experiments.....	113
<i>In vitro</i> and cellular assay characterization.....	115
Competitive in-solution peptide pull down assay.....	115
Competitive chromatin fractionation assay.....	116
Cell permeability and cell toxicity assay.....	117
Foci formation assay.....	118
Fluorescence recovery after photobleaching assay.....	119
Class switch recombination assay.....	123
Conclusions	126
Experimental section.....	127
General procedure for chemical synthesis.....	127

Small molecule synthesis.....	129
Analytical and structural characterization (LC-MS and NMR)	168
Protein expression and purification.....	168
AlphaScreen assay of synthesized ligands.....	168
Isothermal titration calorimetry binding experiments	171
Competitive in-solution peptide pull down assay.....	172
Fluorescence recovery after photobleaching assay.....	172
FRAP technique.....	172
Cell culture.....	173
Ionizing radiation dosing experiments.....	174
Protein NMR experiments	174
X-ray crystallization.....	174
Protein expression and purification.....	174
Crystallization.....	175
X-ray data collection and structural determination.....	176
Crystallography data and refinement statistics.....	176
Class-Switch Recombination assay.....	177
Chromatin fractionation assay.....	178

Foci formation assay.....	178
Cell permeability.....	179
Cell toxicity.....	179
Future Work.....	180
Fragment Based Screen for Improvement of 53BP1 Ligands.....	183
Rationale	184
Current fragment based screening methods.....	184
Fragment-based screening platform design.....	186
Description of Structural Genomics Consortium (SGC) fragment library.....	187
Results and discussion.....	187
Validation of assay method.....	188
Initial hits from screen of SGC fragment library.....	189
Outcome of assay upon addition of UNC1215	192
Conclusions	192
Experimental section	193
Preparation of fragment-only and fragment + UNC1215 IC50 plates for AlphaScreen assay.....	193
Isothermal titration calorimetry binding experiments	194

Future Directions.....	195
Chapter 3: DESIGN, SYNTHESIS, AND CHARACTERIZATION	
OF PEPTIDE-BASED INHIBITORS FOR THE METHYL-LYSINE BINDERS,	
PHF1 AND PHF19.....	197
Biology and Clinical significance of PHF1 and PHF19.....	197
Biological Function of PHF1 and PHF19.....	197
Clinical significance of PHF1 and PHF19.....	198
Purpose of this work.....	199
Development of peptide-based inhibitors	200
Results and discussion.....	200
Structure based analysis of PHF1 and PHF19.....	200
Design of peptide ligands for PHF1 and PHF19.....	202
Analysis of endogenous H3K36Kme3 chromatin peptide	
and protein binding pocket residues.....	202
Analysis of endogenous histone substrate binding affinity.....	204
Structure-activity relationship studies.....	205
Preparation of Non-natural H3K36me3 mimetics.....	205
Initial SAR round of peptide ligand	
amino acid substitutions.....	206

Second SAR round of peptide ligand	
amino acid substitutions.....	210
Conclusions	212
Experimental section	215
General Procedure for Chemical Synthesis.....	215
Non-natural trimethyl-lysine Fmoc protected monomer synthesis.....	216
Solid-Phase Peptide Synthesis of Designed Peptides.....	221
Analytical and structural characterization (LC-MS and NMR).....	231
AlphaScreen assay of synthesized peptide ligands.....	232
Isothermal titration calorimetry binding experiments.....	233
Future work.....	234
APPENDIX I – NMR SPECTRA FOR CHAPTERS II and III.....	238
FUNDING SOURCES.....	312
REFERENCES.....	313

LIST OF TABLES

Table 2.1. SAR Studies and Selectivity Data of the	
N-ethylamine Group of 2 (UNC1554)	30
Table 2.2 . SAR Studies and Selectivity Data for the Linker Group for 2 (UNC1554).....	32
Table 2.3. SAR Studies of the Nicotinic Acid Head Group of 2 (UNC1554).....	34
Table 2.4. Selectivity Data for 61 (UNC2170) and SAR Studies of the	
Terminal Amine Group	95
Table 2.5. Thermodynamic data from ITC experiments.....	97
Table 2.6. SAR Studies of Modifying the Linker of 61 (UNC2170).....	99
Table 2.7. SAR Studies of the Aromatic Head Group of 61 (UNC2170)	102
Table 2.8: SAR Studies of Disubstituted aromatic analogs of 61 (UNC2170)	103
Table 2.9. SAR Studies of the Non-Phenyl Ring Head Groups of 61 (UNC2170).....	105
Table 2.10: SAR Studies of 2-, 5-functionalized aromatic analogs of 61 (UNC2170).....	108
Table 2.11: SAR Studies of Ugi analogs of 61 (UNC2170)	109
Table 3.1: SAR of synthesized peptides for PHF1 and PHF19.....	213

LIST OF FIGURES

Figure 1.1: Structures of known small molecule inhibitors and chemical probes of histone Kme binding proteins.....	12
Figure 2.1: Structural analysis of 53BP1 TTD binding pocket, PDB 2LVM.....	23
Figure 2.2: Structure of UNC00000522.....	25
Figure 2.3: Unpublished NMR titration study of UNC1554 bound within the 53BP1 TTD.....	27
Figure 2.4: Focus areas of UNC1554 SAR studies.....	29
Figure 2.5: ITC binding data for 2 (UNC1554) against wt and D1521A inactive 53BP1 TTD.....	38
Figure 2.6: Region of synthetic modification on 61 (UNC2170)	93
Figure 2.7: <i>In silico</i> docking model of 61 (UNC2170) bound to 53BP1 TTD.....	94
Figure 2.8: ITC analysis of prepared analogs against wt and inactive D1521A 53BP1 TTD mutant.....	97
Figure 2.9: X-ray co-crystallization structure of 61 (UNC2170) bound to 53BP1 TTD dimer.....	111
Figure 2.10: Key interactions between 61 (UNC2170) and the 53BP1 Tandem Tudor domains.....	113
Figure 2.11: NMR structural analysis of 61 (UNC2170) binding to 53BP1 TTD.....	114
Figure 2.12: Competitive in-solution peptide pulldown assay.....	115
Figure 2.13: Chromatin fractionation assay.....	116
Figure 2.14: CellTiter-Glo luminescent cell viability assay	

for 61 (UNC2170) and 64 (UNC2892)	117
Figure 2.15: Foci formation assay.....	118
Figure 2.16: Initial dose range for FRAP experiments.....	120
Figure 2.17: Variation of concentration of 61 (UNC2170) at 10 Gy IR.....	121
Figure 2.18: Representative image of treated and untreated U2OS cell transfected with GFP-tagged isolated 53BP1 TTD response to photobleaching.....	122
Figure 2.19: FRAP experiments without IR.....	123
Figure 2.20: CSR effects upon treatment with 61 (UNC2170).....	125
Figure 2.21: FACS analysis of CSR assay cells 3.5 post treatment.....	126
Figure 2.22: Future proposed 53BP1 TTD inhibitors.....	182
Figure 2.23: Plate Map of Fragment Based Screen.....	189
Figure 2.24: Initial Fragment Hits.....	190
Figure 2.25: Representative IC ₅₀ curves and ITC curves of high potency fragment “Hits”.....	191
Figure 2.26: Representative image of fragment screen plate map and IC ₅₀ curves of hit compounds.....	194
Figure 3.1: Binding regions of PHF1 with endogenous H3K36me3 peptide substrate bound, PDB 4HCZ.....	201
Figure 3.2: Preparation of Non-natural H3K36me3 Mimetics.....	206
Figure 3.3: Focus areas of Peptide SAR Series I.....	207
Figure 3.4: ITC of Series I Peptide SAR ITC.....	208

Figure 3.5: Focus Areas of Peptide SAR Series II	211
Figure 3.6: Proposed Acetamide quaternary amine mimetic bound	
within Tudor domain	236
Figure 3.7: Synthesis of N'-Acetamide, N'-methyl N- α -Fmoc-Lysine.....	236
Figure 3.8: Double Mutant Analogs of Arg40 and Pro38 amino Acid Substitutions.....	237

LIST OF ABBREVIATIONS AND SYMBOLS

°	Degree
Å	Angstrom (10^{-10} meter)
53BP1	p53 Binding Protein 1
ATM	Ataxia Telangiectasia Mutated
CICBDD	Center for Integrative Chemical Biology and Drug Discovery
CSR	Class Switch Recombination
DMSO	Dimethyl sulfoxide
DCM	Dichloromethane
DIC	<i>N,N'</i> -diisopropylcarbodiimide
DMF	Dimethylformamide
EDTA	Ethylenediaminetetraacetic acid
EED	Embryonic ectoderm development
ESI	Electrospraying Ionization
EZH2	Enhancer of zeste homolog 2
FACS	Fluorescence Activated Cell Sorting
FBDD	Fragment-based drug design
FP	Fluorescence Polarization
FRAP	Fluorescence Recovery After Photobleaching

GFP	Green fluorescent protein
HPLC	High-performance liquid chromatography
HAT	Histone Acetyltransferase
HBTU	<i>N,N,N',N'</i> -Tetramethyl- <i>O</i> -(1 <i>H</i> -benzotriazol-1-yl) uronium hexafluorophosphate
HDAC	Histone Deacetylase
HOAt 1	-Hydroxy-7-azabenzotriazole
HTS	High-throughput screening
IC ₅₀	Half maximal inhibitory concentration
ITC	Isothermal titration calorimetry
kDa	Kilodalton
K _d	equilibrium dissociation constant
LC-MS	Liquid Chromatography – Mass Spectrometry
M	Molar
mg	Milligram
min	Minute
mmol	millimole
mM	Millimolar
MeOH	Methanol

MS	Mass spectrometry
MW	Molecular weight
NaCl	Sodium chloride
NMR	Nuclear Magnetic Resonance
PHD	Plant Homeodomain
PHF1	PHD Finger 1
PHF19	PHD Finger 19
PRC2	Polycomb repressive complex 2
rt	room temperature
RNA	Ribonucleic acid
SAR	Structure-Activity Relationship
SBDD	Structure-based drug design
SGC	Structural Genomics Consortium
SUZ12	Suppressor of zeste 12
TFA	Trifluoroacetic acid
TIPS	Tri-isopropylsilane
TLC	Thin Layer Chromatography
TTD	Tandem Tudor Domain
μg	microgram
UNC-CH	University of North Carolina at Chapel Hill

UV	Ultra violet light
nM	Nanomolar
PDB	Protein databank file
μ	10^{-6}
μ l	Microliter
μ M	Micro molar
wt	Wild-type

CHAPTER I

DEVELOPMENT OF INHIBITORS FOR METHYL-LYSINE BINDING PROTEINS

Background

The processes controlling the flow of genetic information are highly regulated and made up of complex networks of enzymes and proteins that control cellular gene expression activities. It is known that every cell in an organism contains the same genetic make-up, but the manner in which that genetic material is manipulated during the development, growth, and maintenance of that organism is controlled by the precise availability of DNA and the gene transcription activities conducted upon it. DNA is located in the nucleus of a cell and resides in a compacted form known as chromatin. Chromatin is a dynamic structure that undergoes architectural modulations to allow for un-winding of tightly pack DNA to allow for gene transcription events to occur on specific sequences of DNA.[1, 2] Chromatin is made up of highly organized nucleosomes that are in turn made up of DNA that is wrapped around histone proteins.[3] The dynamic nature of chromatin architecture influences access to DNA and gene transcription.[1] The process of how different sequences of DNA conformationally move between these two states involves the writing, reading, and erasing of post-translational marks (PTMs) including acetylations on histone peptides and methylations on CpG islands on DNA.[4] The process of methylating DNA at CpG dinucleotides within transcription start sites provides a signal that can be interpreted by DNA binding proteins and lead to gene silencing events.[5] The understanding

of the overall effects and functions of DNA methylation are still be elucidated.[5] Euchromatin is characterized by hyper-acetylation of the histone tails of histone 3 and 4 and low levels of cytosine methylation on CpG islands.[1] Heterochromatin opposes this with hyper-methylated CpG islands and low levels of histone acetylation.[1] This dynamic process of reversible chemical modifications on both DNA and associated histone proteins, where the DNA sequence is not directly modified, allows for specific gene transcription events to occur and this overall process is known as epigenetics.[6] Epigenetics is defined as a genetic process where heritable phenotypic traits are passed onto an organism's progeny, where those phenotypic traits are controlled by events that do not involve direct changes to the underlying DNA sequence of the organism. One mechanism that influences this dynamic process of passing on heritable phenotypic traits is the covalent modification, reading, and removal of PTMs on histone peptides by various enzymes and binding proteins. In 2000, a hypothesis known as the "histone code" was proposed and outlines a general way to view how histone PTMs function to modulate chromatin architecture.[7] This hypothesis has been expanded to gain an understanding how PTMs work to induce chromatin decompaction or condensation, or the recruitment of effector proteins such as Kme reader proteins to allow for defined gene transcription processes to occur.[8] Epigenetic regulation of an organism's genetic code is thought to occur through reading of histone PTMs that make-up the "language" of the "histone code." [7] Binding of specific histone PTMs, or binding of a specific combinations of histone PTMs have defined biological outcome due to the modulation of chromatin architecture and subsequent DNA transcription and silencing events. It has been shown that misregulation of this defined epigenetic process can occur in diseases such as cancer and can have devastating health outcomes.[9] It is of interest to develop chemical tools

that can help further understand this highly complex process and also potentially regulate different enzymes, or proteins within it that directly induce defined gene transcription outcomes.

Post-Translational Modifications (PTMs) and their role in chromatin regulation

There are several major biological processes that play a role in the regulation of gene expression such as the activity of transcription factors and the dynamic modulation of chromatin architecture.[10] A key mechanism involved chromatin modulation is the addition, recognition, and removal of PTMs on histone peptides and proteins by various enzymes and binding proteins. An example of one of the earliest PTMs to be identified was phosphorylation on residues on protein surfaces.[11, 12] Common forms of histone-based PTMs include phosphoryl, acetyl, ubiquitin, and methyl groups that are chemically installed onto specific residues, such as serine, threonine, and tyrosine, on histone peptides.[8] These types of chemical modifications are enzymatically installed, read, and removed in a dynamic process that allow for various types of biological processes to be initiated, enhanced, or silenced.[2, 13] Post-translational modifications on histones were identified in 1964 with the initial example being dynamic regulation of histone acetylation by the opposing actions of histone acetyltransferases (HATs) and histone deacetylases (HDACs).[14] It was observed that HATs enzymatically transferred an acetyl moiety from S-adenosyl methionine to the ϵ -NH₃⁺ group of histone lysine residues.[15, 16] The acetyl group could then be removed by HDAC enzymes.[15, 16] The process of linking histone PTMs to gene transcription processes was further elucidated by the discovery of p55/Gcn5 and HDAC1/Rpd3 which were identified as a histone associated HAT and HDAC.[17, 18] From that point, this field of research has grown immensely to identify a large number of other forms of histone PTMs and the enzymes and proteins that modulate those PTMs. There are currently

greater than a dozen types of PTMs that have been identified on various types of histone residues with new forms being discovered on a regular basis.[8] Lysine residues on histone peptides are able to be mono-, di-, or tri-methylated by methyltransferase enzymes that are responsible for the chemical installation of this type of PTM. The binding of these variants of Kme PTMs have been observed to modulate chromatin architecture and can lead to gene expression, or silencing events depending on which Kme PTM is present on specific histone tails. For example, the trimethylation of histone 3, lysine 9 (H3K9me3) and histone 3, lysine 27 (H3K27me3) has been shown to lead to decreased transcription of gene loci leading to a gene silencing event.[2] Specifically, H3K9 is methylated by the methyltransferase Suv39h and creates a binding site for the Kme reader protein HP1.[19, 20] The binding of HP1 leads to the recruitment of other effector proteins that overall contribute to the propagation of heterochromatic subdomains in chromatin.[19, 20] Conversely, the trimethylation of histone 3, lysine 4 (H3K4me3) is a hallmark of active transcription.[2, 21, 22] These examples show that the recognition of specific methylated lysine residues on histone peptides by Kme binding proteins is a critical event in chromatin regulation.[7, 23] The binding of these types of PTMs is often associated with the recruitment of various types of effector proteins leading to assembly of macromolecular complexes that support defined gene activation and repression outcomes.[24-26] The dynamic process of binding histone Kme PTMs is finely controlled in healthy cells and is essential to critical cellular processes. Interestingly, it has been shown in diseases such as cancer that this process becomes highly misregulated.[27-29]

Classes of methyl-lysine binding proteins

Currently there are greater than 200 identified histone Kme reader proteins and they are classified into three overarching classes of Kme binding proteins based upon the type of protein domain that they use to bind Kme PTMs on histone tails.[30] The three classes are plant homeodomain (PHD) zinc fingers, WD40 repeat domain proteins, and a super-family of binding proteins known as the Royal family of reader proteins. The Royal family consists of Tudor, chromodomain, Pro-Trp-Trp-Pro (PWWP) domain, and malignant brain tumor (MBT) domain proteins. The research described herein is focused on developing inhibitors for two members of the Tudor class of Royal family of reader proteins; 53BP1 and PHF1/PHF19.

Interpretation of specific PTMs by histone methyl-lysine binding proteins

Kme binding proteins act to continuously sample the residue sequence of chromatin, DNA, histone peptides, and protein surface residues to identify specific lysine PTMs that they are capable of binding. This sampling process works to identify a specific, or combination of specific, Kme PTMs on the same or multiple histone peptides through the use of a unique binding method. It is often the case that chromatin regulatory proteins possess multiple types of domains that are each individually capable of binding multiple histone PTMs on a specific nucleosome.[31] An example of a histone binding protein that reads multiple PTMs is BPTF.[32, 33] This protein contains a PHD Kme binding domain and also a bromodomain which selectively binds acetylated lysine residues.[32, 33] Work to understand how a histone reader protein can engage in the reading of multiple PTMs through various domains contained within the binding protein has been described as “multivalency”. The methyl-lysine reader protein UHRF1 has been described to bind two different PTMs via use of its PHD and Tandem Tudor

domains.[34, 35] Kme binding proteins have evolved a unique structural motif that is utilized to bind to Kme PTMs. The most common manner in which methylated histone lysine residues are bound is through the use of an aromatic cage binding pocket. The aromatic cage consists of 3-5 aromatic residues including phenylalanine, tyrosine, and tryptophan residues that form a cavity into which a methylated lysine residue can bind through cation- π and hydrophobic interactions. The physical structure of the Kme binding aromatic cage has been observed to sample through multiple conformation, but once a specific Kme PTM is recognized and bound, they primarily adopt either a static aromatic cage conformation, or the cage is formed through a substrate-induced conformational change. Examples of Kme binding proteins that possess preorganized aromatic Kme binding pockets include 53BP1 TTD and L3MBTL1. An example of a Kme binding protein that has a Kme binding pocket that forms through a substrate-induced conformational change is CBX7. Both types of Kme binding mechanisms act to recognize the positively charged ammonium cation on the methylated lysine residue on the histone peptide tail.[24, 25, 36] Specifically, this work focuses on a Kme domain known as the Tudor family of Kme binding proteins.[37]

Tudor binding domain

Currently there are approximately 37 different proteins that have been identified to contain Tudor domains; both tandem and single Tudor domains.[30, 37] This type of binding domain has been observed to selectively bind methylated histone lysine residues of varying methylation states. The Kme binding domains of these proteins have a similar conformational fold consisting of a β -barrel conformation that contains a series of aromatic residues that interact with the charged ammonium head group of the methylated lysine residue through cation- π

binding and hydrophobic interactions.[37] Additional hydrogen bond and hydrophobic/Van der Waals type interactions are also observed in structural studies to play a role in binding their endogenous substrate, but the key binding interaction is the binding of the methylated histone lysine residue.[24-26, 36] It has been observed that different Tudor domain containing Kme binding proteins have differing degrees of specificity for the extent of methylation present on the ϵ -NH₃ group on histone lysine residues. For example, the Kme binding protein 53BP1 selectively binds H4K20me₂ with decreased affinity for H4K20me₁, and no affinity for the non-methylated or trimethylated H4K20 residue.[36] The two Kme binding proteins that this dissertation is focused on are 53BP1 and PHF1/PHF19.[24-26, 36] 53BP1 and PHF1/PHF19 are made up of two different types of Tudor domains.[24-26, 36] 53BP1 contains a tandem Tudor domain, where only one of its two Tudor domains works to bind the histone 4, lysine 20 dimethyl (H4K20me₂) PTM in a well-defined aromatic cage in combination with hydrogen bonds between the protein surface and endogenous peptide backbone residues.[36] PHF1 and PHF19 differ in that it contains a shallow, single Tudor domain that acts to bind the histone 3, lysine 36 trimethyl (H3K36me₃) PTM in combination with hydrogen bonds between the protein surface and histone peptide backbone residues.[24-26] The different structural properties of these two types of Tudor domains may require different approaches toward development of inhibitory ligands. It has been shown that both of these proteins play a role in diseases such as cancer and development of inhibitory ligands could be used to explore whether their inhibition would lead to favorable biological outcomes.[38-40]

Hypothesized outcomes of domain inhibition

The biological process of binding Kme PTMs on histone tails has been observed to induce specific gene transcription activity including gene silencing as previously described.[2, 21] The misregulation of this highly controlled biological process is found in diseases such as cancer where the activity of Kme binding proteins may lead to suppression of tumor suppressor genes. If the PTM binding domain within a methyl-lysine binding protein could be inhibited through the use of a small molecule, or peptide-based ligand, then the protein would be unable to bind its endogenous substrate. Thus inhibition of this binding protein's activity may allow for a return to normal levels of gene activation/repression found in healthy cells and lead to a therapeutic effect. It has been shown that small molecule inhibition of lysine methyltransferases has potential favorable therapeutic outcomes.[41, 42] Currently there are several methyltransferase inhibitors in clinical trials for treatment of various types of cancer and inhibitors for the histone methyltransferases Dot1L and EZH2 have successfully progressed into Phase 1 and Phase 2 clinical trials.[43, 44] The highly misregulated activity of these two histone methyltransferases in various types of cancer has been shown to lead to uncontrolled cell proliferation.[45-48] The development of inhibitory ligands for this class of epigenetic proteins speaks favorably towards development of inhibitors for histone methyl lysine binding proteins, as the enzymatic activity of methyltransferases provides a defined substrate that allows for recruitment of a Kme reader. A similar type of histone binding protein that selectively binds acetylation marks on histone lysine residues are bromodomain acetyl-lysine binding proteins. Recent inhibitor development for the BET family of bromodomain proteins has resulted in the

development of JQ-1 chemical probe and also variants of that compound that have been commercialized and entered Phase 1/ 2 clinical trials.[49-51]

Current state of ligand development for methyl-lysine binding proteins

Challenges toward development of inhibitory ligands for methyl-lysine binding proteins

There are many challenges toward developing inhibitory ligands for this class of binding protein. An initial challenge for this process is to develop a robust screening assay that can be used in a high-throughput (HTS) assay format. This class of protein engages in non-enzymatic binding interactions. The standard screening methods for this type of interaction include fluorescent polarization (FP) assays, isothermal titration calorimetry (ITC), and surface plasmon resonance (SPR) experiments. These types of assays can prove costly in terms of time, instrument, and protein costs to successfully conduct a large scale screen of synthesized inhibitors. Histone Kme binding proteins engage their endogenous targets through low, to mid-micromolar protein-protein binding interactions.[24-26, 36] It has proved difficult to successfully develop a binding assay that is reliable and robust for this range of affinities. These two initial challenges were overcome through the development of a HTS format of the bead-based high avidity AlphaScreen assay (Perkin Elmer) for Kme reader proteins in the Center for Integrative Chemical Biology and Drug Discovery (CICBDD).[52] This assay was optimized to allow for a robust HTS quality screening assay capable of measuring low affinity interactions between protein and inhibitory compounds for future inhibitory ligand discovery efforts. The high avidity nature of the AlphaScreen assay leads to a very large signal increase once a single, low affinity binding interaction is detected. The signal detected from one binding interaction multiplies into a large signal increase, compared to a FP based assay where the detected signal is only occurring

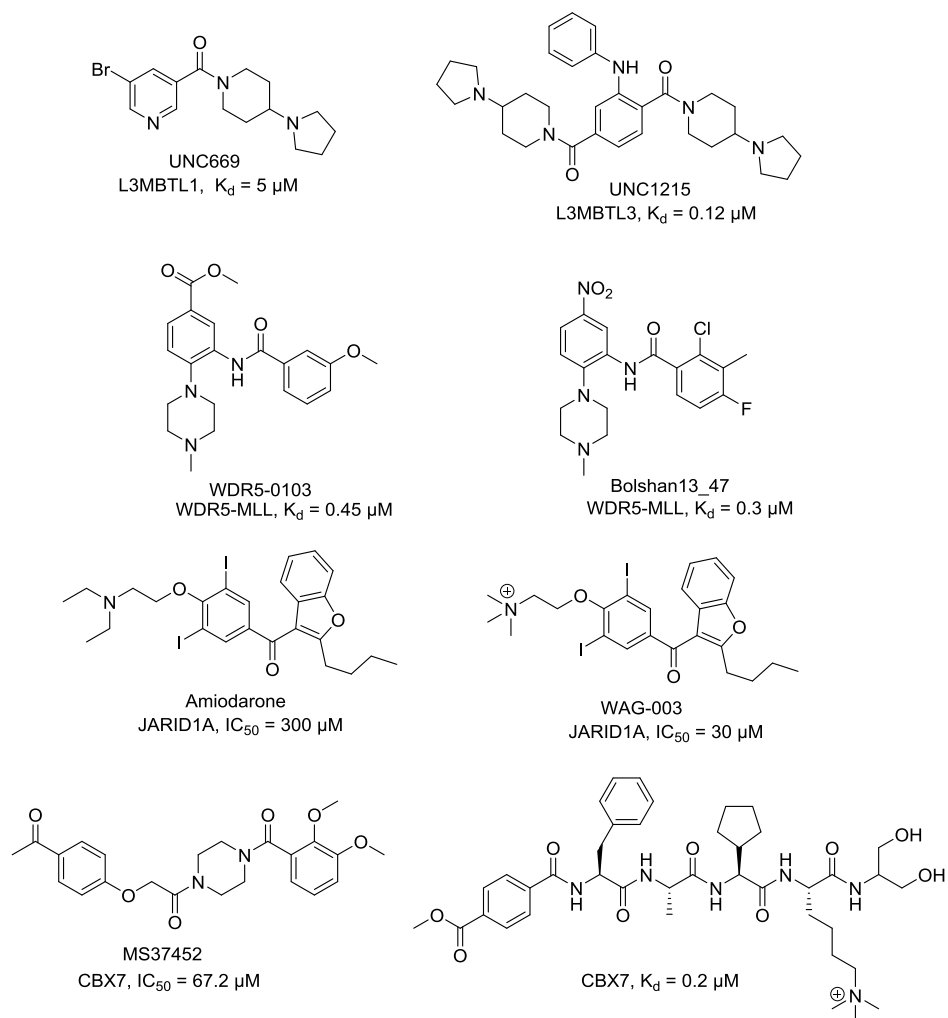
from a single binding event. Another difficulty in developing inhibitory ligands for this class of binding proteins is that there were no small molecule, or non-endogenous peptide-based ligands known when our efforts in the CIBDD began. Many x-ray crystal structures of Kme reader proteins bound to their endogenous substrate have been solved and have assisted greatly in structure based drug design methods that are the backbone for how inhibitory ligands are currently being developed for this class of binding proteins.

Current examples of Inhibitory ligands

Development of small molecule ligands for MBT domains were the first examples of Kme binding protein inhibitors. The initial development of the ligand, UNC669, for L3MBTL1 provided evidence that small molecules were able to mimic the methylated lysine residue of its endogenous peptide substrate, H4K20me1.[53] L3MBTL1 has a static aromatic cage in which H4K20me1 residue binds in a cavity insertion manner. The chemical structure of UNC669, Figure 1.1, has a piperidine-pyrrolidine head group that mimics the H4K20me2 and structural studies have shown in docks it the same fashion as the endogenous substrate. The pyrrolidine head group was key to this compound's ~5 fold greater affinity compared to the endogenous peptide. It was thought that the pyrrolidine head group allowed for greater affinity within the aromatic cage of L3BTL1 compared to endogenous substrate. The next development for inhibitors for the MBT family of proteins was the discovery of UNC1215.[54] This inhibitor was classified as a chemical probe due to its high affinity and selectivity and ability to be used in both *in vitro* and cellular based assays.[54-56] Recent work towards the development of inhibitors for the CBX family of chromodomain containing Kme reader proteins has produced both a small molecule inhibitor and a peptide-based inhibitor.[57, 58] The small molecule inhibitor,

MS37452, has a $IC_{50} = 67.2 \mu M$ for CBX7 in a FP assay.[58] The peptide-based inhibitor has $K_d = 0.20 \pm 0.04 \mu M$ for CBX7 and is observed to possess a 10-fold and 400-fold selectivity over CBX8 and CBX1, respectively.[57] In the WD40 class of proteins, there has been early work toward development of small molecule inhibitory ligands. Development of WDR-0101, WDR-103, and Bolshan13_47 all show low nanomolar potencies for the WDR5 protein of the WD40 family.[59, 60] Recent work for the PHD family of proteins is shown through Amiodarone and WAG-003 that are low micromolar inhibitors for JARID1A.[61] Figure 1.1 outlines the structures of the currently developed small molecule inhibitors and chemical probes for histone methyl lysine binding proteins. This work describes the discovery and development of a small molecule ligand for the 53BP1 TTD and initial work toward a peptide-based inhibitor for PHF1 and PHF19.

Figure 1.1: Structures of known small molecule inhibitors and chemical probes of histone Kme binding proteins.



Use of Hit discovery techniques to discover inhibitory ligands

Methyl-lysine binding protein biological screening methods

As previously mentioned, the Frye research group has utilized the bead-based high avidity AlphaScreen assay to measure the affinity and selectivity of prepared ligands against defined Kme binding proteins. This assay has been modulated to work with a common method of screening large numbers of proteins and enzymes, known as cross-screening, in order to obtain selectivity and potency measurements for all synthesized analogs. The current method of cross-screening over 12 Kme reader proteins, from all classes of Kme reader proteins, and has allowed for the development and optimization of UNC669 and UNC1215.[53, 54] It is hoped that future discoveries of inhibitory hits will be generated for other members of the Kme reader protein panel through the use of this method. Additionally, once a new Kme reader protein has a validated AlphaScreen assay in place, it is screened for binding activity against a focused chemical library of putative epigenetic inhibitors and their analogs. Screening results from this initial focused screen are used to identify initial “hit” molecules that can be further optimized for potency and selectivity.

Structure-based drug design

Once an initial “hit” molecule has been identified for a defined target, structure-based drug design techniques are utilized to optimize the potency and selectivity of the compound. To successfully use structure-based design techniques, an X-ray or NMR structure of the target protein bound to its endogenous substrate or a synthetic ligand is needed. Key binding interactions between the protein and peptide need to be identified, such as hydrogen bonds, and estimations on how the “hit” molecule replicates those binding interactions can be hypothesized.

A technique that is helpful in initial hypothesis development is the use of *in silico* docking techniques.[62, 63] A model of how the “hit” molecule binds to the protein can be developed and used to predict how the hit molecule docks to the target protein. From this initial modeling and structural analysis process, analogs of the hit molecule can be rationally designed and synthesized to develop a structure-activity relationship (SAR). The prepared analogs are then screened against the target protein to determine if structural changes of various analogs improves the potency compared to the “hit” molecule. It is important that each prepared analog only contains a single synthetic modification compared to the “hit” molecule. This is important so that the potency measurements for the analog can be compared to the initial “hit” molecule to determine if a single functional group change increases potency. Once a functional group change has been identified that improves potency and affinity, it is incorporated into the structure of the “hit” molecule and the process is continued over several cycles in an effort to optimize the hit.

Purpose of this work

The purpose of this work is to discover and optimize small molecule and peptide-based ligands for two different chromatin binding proteins, 53BP1 and PHF1/PHF19. Both proteins share a common type of Tudor Kme binding domain, but do not bind the same PTM on histones. Medicinal chemistry efforts will utilize structure-based design techniques to improve their selectivity and affinity towards their intended target. Once a ligand is optimized, its ability to engage and modulate the biological function of its target will be characterized through structural based techniques (X-ray co-crystallization and protein NMR) and through *in vitro* and *in vivo* assays.

Chapter II

Design, synthesis, and characterization of small molecule ligands for the chromatin

methyl-lysine binder, 53BP1

Discovery and biological function of 53BP1

The chromatin methyl-lysine binding protein p53-Binding Protein 1 (53BP1) was discovered, in 1994, as a p53-interacting protein in a yeast two-hybrid screen through its ability to bind and stabilize p53 at sites of DNA damage.[64] 53BP1 is a Kme binding protein that plays a central role in DNA damage repair (DDR) pathways and is recruited to sites of double-strand breaks (DSB). This protein was first detected in a pull-down assay where it was bound to p53 at a specific dimethylated lysine residue. The assay was conducted by induction of DNA double-strand break damage and showed initial evidence that 53BP1 played a role in a DNA Damage Repair (DDR) pathway.[64] It was found that 53BP1 was able to selectively bind three different dimethylated lysine residues on hyperphosphorylated p53, post DNA damage repair (DDR) activation. Those residues were p53 lysine 370 dimethyl (p53K370me₂), p53 lysine 371 dimethyl (p53K372me₂), and p53 lysine 382 dimethyl (p53K382me₂).[65] Additional histone peptide screening assays showed evidence that 53BP1 was able to bind histone 3, lysine 70 trimethyl (H3K70me₃), but was later shown that histone 4, lysine 20 dimethyl (H4K20me) was the primary PTM in histones that it selectively binds.[65, 66] This finding was subsequently

supported by structural studies with a binding $K_d = 19.7 \pm 0.7 \mu\text{M}$. [67] The process of how 53BP1 is activated in DDR is dependent on the activity of Ataxia Telangiectasia mutated (ATM) kinase that regulates the DNA damage signaling response.[68] The surface of the protein is hyperphosphorylated by Ataxia Telangiectasia mutated (ATM) kinase once DDR is activated (Ser6, Ser25, Ser29, Ser784).[69] Recruitment of 53BP1 can lead to cell cycle arrest and DNA damage repair, or lack thereof, if there are defects or deficiencies present in the enzymes or proteins involved in this signaling cascade [69-81]. 53BP1 undergoes homo-oligomerization at sites of DNA double strand breaks that is necessary for efficient foci formation to occur through binding of its residues 1231-1277.[68] There are various domains of 53BP1 interaction with DNA. There are two BRCT (breast cancer 1 (BRCA1) carboxy-terminal) domains (residues 1702-1972) that interact with the DNA binding domain of p53 (residues 94-292) and also bind DNA.[68, 82] Recognition of the H4K20me2 post-translational modification (PTM) is performed by the tandem Tudor domain (TTD) of 53BP1.[66] Only the first domain of tandem Tudor domain contains the Kme binding aromatic cage and binds the K4K20me2 post-translational modification; the second BRCT domain only interacts with 1 residue (H4H18) from H4K20me2 peptide.[66]

Recent evidence has shown that 53BP1 is also able to bind histone 2A, lysine 15 mono-ubiquitin (H2AK15ub) through use of a small binding region on the protein known as the Ubiquitin Binding Domain.[83] It is thought to undergo multivalent binding of H4K20me2 and H2AK15ub in the DDR response. Additionally, the TTD of 53BP1 has been shown to bind to p53K382me2 and thereby increase the concentration of p53 at DNA damage sites [65]. Unlike H2AX phosphorylation, global levels of H4K20me2 do not increase after DNA damage; instead,

53BP1 recognizes histone H2A ubiquitylated at lysine 15 by RNF168 in response to DNA double-strand breaks [83]. 53BP1 recruitment is also facilitated by RNF168 ubiquitylation-dependent removal of L3MBTL1 and JMJD2A from H4K20me2 after DNA damage [84, 85].

It has been observed that cells deficient in 53BP1 prove to be viable but display a phenotype resembling mild ATM kinase-signaling defects due to the cell's inability to properly respond to and repair DNA damage [86, 87] 53BP1 knockout mice are viable, but are growth retarded, immune deficient, radiation sensitive, and cancer prone [86]. In addition, these mice display a defect in antibody class switch recombination (CSR) which also requires 53BP1 activity [88]. By testing different fragments of the 53BP1 gene, it has been shown that the region that harbors its Kme binding tandem Tudor domain (amino acids 1235-1709) is recruited to ionizing radiation (IR)-induced foci as efficiently as full-length, endogenous 53BP1 [64].

DNA Repair Pathways, BRCA Mutations, and Cancer

The process of the repair of DNA DSBs can proceed through different repair pathways. Non-homologous end joining (NHEJ) recombination and homologous recombination (HR) are the two most common repair pathways. 53BP1 acts to assist in the identification of DNA damage break sites through recognition of H4K20me2 and then facilitates the repair of DNA through the use of either activation of NHEJ or HR.[89] NHEJ is a rapid, error-prone DNA repair pathway.[90] It acts to ligate two broken DNA ends without the requirement of sequence specificity.[90] DSBs are detected by Ku70-Ku80 dimer and bind to exposed break sites.[90] The Ku70-Ku80 dimer then recruits DNA-dependent protein kinase (DNA-PKcs) and in turn XRCC4 and DNA ligase IV that act to anneal broken stands together. NHEJ is a form of V(D)J recombination. HR is a different DNA repair pathway that operates in a higher fidelity method to

repair DNA breaks.[91] HR requires sequence specificity and reliance on the sister, or non-sister chromatic of the homologous chromosome as a template for repair of the break. DSBs are detected by ATM, H2AX is phosphorylated and MDC1 is recruited to γ H2AX at break sites.[92] MDC1 then recruits the MRN complex, in addition to exo1, CtIP, and BLM and 5'-end resection begins (Mre11/exo1/CtIP nuclease activity facilitated by BLM helixase).[73] As ssDNA are generated, they are bound by RAD51/BRCA2 and are eventually displaced by RPA. The generated 3'-ssDNA then begins strand invasion into sister chromatid (via RAD51 assistance) to generate a D-loop and a Holliday junction. Pol ζ then synthesizes new DNA off 3' end using sister chromatid template. Newly synthesized ssDNA is then ligated to original broken strand by DNA ligase and the Holliday junction is resolved providing repaired dsDNA. The decision to proceed through either pathway is controlled by the presence or absence of BRCA1. BRCA1 is a checkpoint and a DNA damage repair gene that is required for maintenance of genomic integrity. When BRCA1 is present, it acts to inhibit 53BP1 and leads to the induction of the HR repair pathway. When BRCA1 is not present or has undergone an inactivating mutation, 53BP1 is able to inhibit the activity of the MRN and CtIP complex and lead to the activation of NHEJ.[93] It has been shown in knock-out cell lines that when 53BP1 and BRCA1 are both absent, HR is the predominant pathway that is activated leading to high fidelity DNA repair.

BRCA1 is a checkpoint and DNA damage repair gene that is required for maintenance of genomic integrity, and the inheritance of mutated BRCA1 is a major risk factor for breast and ovarian cancer [94]. BRCA1 and BRCA2 defective cells (BRCA mutants) are unable to fully activate the HR pathway and this can lead to induction of cancer growth and metastasis.[93] Heavy reliance on NHEJ that can increase genomic instability caused by increased amounts of

mutations and improper gene rearrangements. Gene mutations predispose carriers to higher incidences of breast and ovarian cancer. BRCA1 knockout in mice are embryonic lethal [95], and conditional knockout in the mammary gland results in low frequency and long latency of mammary tumor formation [96]. It was recently reported that the BRCA1 knockout developmental phenotype is rescued when placed on a 53BP1-null background, and adult mice that are null for both the 53BP1 and BRCA1 genes age normally and display a very low incidence of tumor formation. [97] The genomic instability in the BRCA1 knockout can be overcome because the homologous recombination (HR) pathway is largely restored in cells lacking both BRCA1 and 53BP1. [93]

Use of medicinal chemistry techniques to discover inhibitory ligands for 53BP1

To date there have been no inhibitor ligands developed for 53BP1. The biological role that 53BP1 plays in cells possessing a BRCA1 mutation leading to the induction of NHEJ and error prone DNA repair is thought to play a deleterious role in the progression of diseases such as cancer.[98] Knockout models show that when both 53BP1 and BRCA1 are not present (analogous to an inactive BRCA1 mutation commonly found in cancer), the high fidelity HR DNA repair pathway is activated leading to increased genomic integrity.[86] It was hypothesized that a small molecule ligand that prevents Kme recognition by the 53BP1 TTD would antagonize its biological activity in cells possessing BRCA1 mutations and lead to rescue of their genomic stability by restoration of the HR pathway. Notwithstanding the clinical challenges of cancer prevention strategies, the discovery of a small molecule *in vivo* probe for 53BP1 would enable this hypothesis to be tested pre-clinically, and the devastating effects of BRCA1 mutations, including prophylactic surgical interventions [87], could perhaps be diminished. Therefore,

research was aimed at finding and synthesizing cell penetrant small molecules that would bind selectively to the tandem Tudor domain of 53BP1 using structure-based design and iterative medicinal chemistry. It is additionally hypothesized that 53BP1 TTD inhibitors would work in concert with other cancer inhibitors that are focused on BRCA1. Specifically, the use of inhibitors of Poly(ADP-ribose) polymerase (PARP).[99, 100] PARP rapidly co-localized at sites of DNA breaks and catalyzes formation of poly(ADP-ribe) polymers that induce relaxation of local chromatin structure.[99] PARP is necessary for efficient recruitment of MRE11-dependent end-resection, in concert with CtIP, Rad 50, and Nbs1 during S- and G2-phase. Inhibition of PARP will prevent SSB repair and lead to replication fork stalling/collapse in DSBs that are lethal to HR-deficient cells.[99, 100] BRCA1-deficient cells are vulnerable to poly(ADP-ribose) polymerase inhibitors (PARPi) due to inability to favor HR-mediated DDR as 53BP1 is not inhibited by CtIP due to lack of BRCA1.[99, 100] 53BP1 inhibitors will potentially re-capitulate the process of how CtIP is released from inhibition by 53BP1 and allow for restoration of HR repair. Herein, the initial results of those research efforts are reported.

Purpose of this work

The purpose of this work will be to identify highly selective and potent small molecule inhibitors for the methyl-lysine binding protein 53BP1. The molecules will focus on binding within the tandem Tudor domain of this protein and act to inhibit its ability to bind its endogenous substrates. Once-initial selective small molecules are identified, a candidate will be chosen and optimized through the use of iterative medicinal chemistry techniques and novel compound synthesis to develop a Structure-Activity Relationships (SAR) study. Further, biological assays will be used to determine whether candidate compounds are selective and

potent enough to disrupt the known biological function of 53BP1 in a series of *in vitro* and *in vivo* assays designed to quantitatively measure proteins biological functions.

Work on development for small molecule inhibitors of 53BP1 TTD will be outlined in three sections in Chapter II; UNC1554, UNC2170, and a small molecule fragment screen.

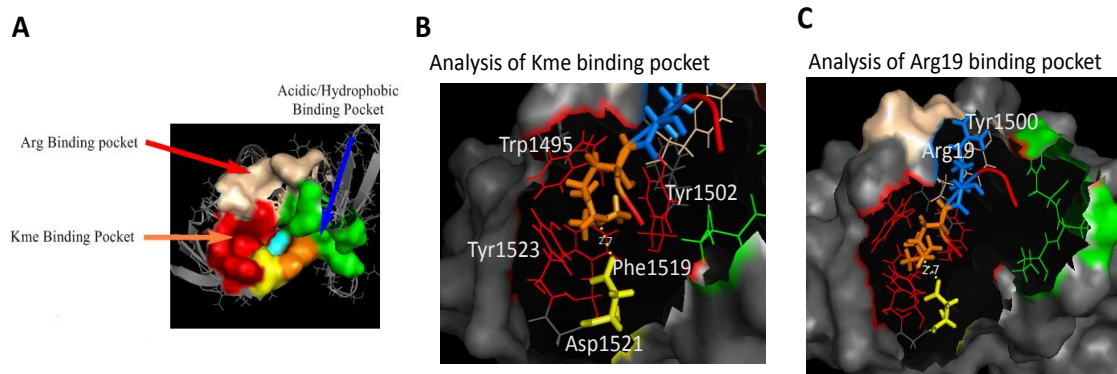
Benzimidazole-based small molecule ligand of 53BP1, UNC1554

Results and discussion

Structure based analysis of 53BP1

Several binding interactions were observed in the X-ray co-crystallization structure of 53BP1 TTD bound to endogenous H4K20me2 peptide. The endogenous peptide substrate binds in three main regions that include the Kme binding pocket (red), the Arg binding pocket (gray), and the Acidic/hydrophobic pocket (green), Figure 2.1 A. The acidic D1521A residue is highlighted in yellow, Figure 2.1 A. An additional serine (blue) and aspartic acid residue (orange) are highlighted as potential binding interactions that were not clearly defined in the X-ray structural studies. The key binding interactions were a hydrogen-bond between the Kme basic amine and an aspartic acid (Asp1521) and a cation- π interaction between Kme2 and phenylalanine, tyrosine, and tryptophan residues within the aromatic binding cage (Tyr1502, Tyr1523, Phe1519, Trp1495), as shown in the red regions in Figure 2.1 B. An additional binding interaction was a cation- π bond between arginine 19 on the H4 peptide tail and a tryptophan residue (Trp1500), white region in Figure 2.1 C.[66]

Figure 2.1: Structural analysis of 53BP1 TTD binding pocket, PDB 2LVM



Mutagenesis of the histone peptide had previously demonstrated that H18 also contributes to binding, but this interaction was not clearly defined in the X-ray crystal structure.[66] To develop a more complete understanding of the binding interactions of the H4 peptide (amino acids 14-27) bound to 53BP1, isotope enriched (^{13}C and ^{15}N) NMR spectroscopy [101] was applied to determine the structure of the central residues of the H4K20me2 peptide bound to 53BP1 (PDB 2LVM), Figure 2.1.[102] It was observed that this central region of the H4 peptide corresponding to residues 15 to 22 adopts a “U-turn” conformation. Notably, in addition to the binding interactions with H4K20me2 and H4R19 that were detected in the crystal structure, further analysis revealed a pocket containing both acidic and hydrophobic residues that accommodates H4R17 and H4V21, green regions in Figure 2.1. It has also been shown that acetylation of H4K16 diminishes 53BP1 binding by disrupting a salt bridge between H4K16 and Glu1551.[102] Therefore, a small molecule that could occupy the methyl-lysine binding cage of 53BP1 and interact favorably with some of the surrounding residues would be expected to block 53BP1 binding to H4K20me2.

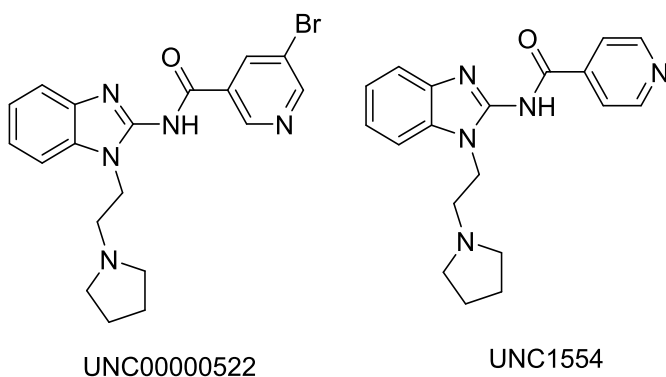
The key interactions that a small molecule inhibitor should mimic are the hydrogen-bond between the Kme basic amine and Asp1521 that works in combination with a cation- π interaction between Kme2. Additionally, it is proposed that the other key interaction to mimic would be the cation- π bond between Arg19 and Trp1500. The hydrophobic Val21 residue is bound within the acidic/hydrophobic binding pocket, but no direct binding interactions are noted in that region in the X-ray co-crystal structure. Initial small molecule inhibitor discovery efforts will focus on compounds that mimic the above mentioned binding interactions.

Discovery of UNC1554 hit

The process of discovering a small molecule inhibitor for the Tandem Tudor domain (TTD) of 53BP1 was initiated by analysis of the previously described structural characteristics of its Kme binding domain. It is known that the TTD binds the H4K20me2 residue on the histone 4 peptide tail (H4K20me2) which maintains a positive charge under physiological buffer conditions. This cation- π binding interaction is key to the overall binding process of the histone peptide and should be maintained in a small molecule inhibitor. Previous *in silico* docking studies that were conducted by Dr. Cen Gao and Dr. Martin Herold (both formerly of the Center for Integrative Chemical Biology and Drug Discovery, CICBDD) and Dr. Lindsey Ingerman James (Assistant Research Professor, CICBDD) show that flexible inhibitors that contained a pyrrolidine residue were highly likely to bind the 53BP1 TTD with the pyrrolidine docking within the Kme binding pocket. This pyrrolidine structural feature was common to a small molecule ligand that was discovered for L3MBTL1, UNC669. This molecule contains a pyrrolidine functional group that docks within the L3MBTL1 Kme binding pocket and mimics the H4K20me2 residue that it endogenously binds. 53BP1 TTD was initially screened against a

focused compound library at the CICBDD known as the EpiG library was screened to obtain potential “hit” molecules for 53BP1 TTD. The EpiG library of compounds possesses a large number of known and putative small molecule inhibitors and their analogs for epigenetic-related enzymes and proteins. From this focused screen that was conducted using the AlphaScreen assay platform, a hit compound known as UNC00000522 was observed to inhibit the binding activity of the 53BP1 TTD by 95% when screened at 10 μ M and had an $IC_{50} = 6.3 \pm 6.3$ μ M, Figure 2.2.

Figure 2.2: Structure of UNC00000522



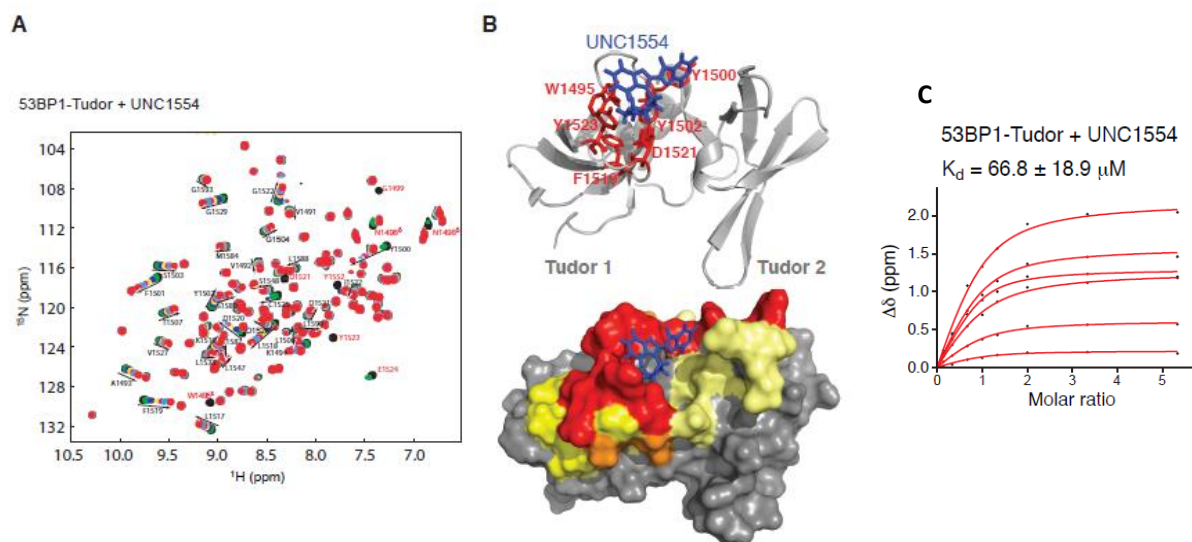
From this initial “hit”, structure-based design methods and also *in silico* techniques were used in tandem to begin the process of determining how UNC00000522 bound to 53BP1 TTD and also begin rationally designed new analogs for screening. *In silico* docking models to of how UNC00000522 binds to the 53BP1 TTD showed that the pyrrolidine functional group bound within the Kme binding pocket and the 5-bromo-nicotinic acid functional group bound either within the arginine binding pocket or in the acidic pocket. An initial set of nine UNC00000522 analogs were synthesized to determine if potency could be improved. It was determined that the AlphaScreen assay conditions for the initial screening of UNC00000522 did not provide IC_{50} 's with a high level of precision and improved conditions were used to screen the newly

synthesized analogs. From that set of synthesized analogs emerged UNC1554 with an $IC_{50} = 22.7 \pm 6.5 \mu M$, Figure 2.2. The benzimidazole scaffold that possessed an ethyl-pyrrolidine functional group was chosen as the starting point for SAR studies to discover a potent and selective small molecule inhibitor of the 53BP1 TTD.

Comparison of UNC1554 to endogenous H4K20me2 chromatin peptide

A structure-based design approach was initiated by analyzing available crystal structure data of 53BP1 bound to an H4K20me2 peptide (PDB 2IG0).[66] As previously described, the H4K20me2 residue binds within the Kme binding pocket with the Arg19 residue binding within a shallow pocket above the Kme2 pocket and Val21 binding in an acidic/hydrophobic pocket below the Kme binding site. It was hypothesized that the ethyl pyrrolidine functional group of UNC1554 would bind within the Kme binding site but it was uncertain how the remaining part of the small molecule would dock within the TTD. The compound was sent off to our collaborator, Prof. Georges Mer (Mayo Clinic) to undergo NMR studies in order to gain a better idea on how the molecule docked to the protein. Intermolecular NOEs experiments were conducted to determine direct interactions between UNC1554 and the 53BP1 TTD. These unpublished results showed that UNC1554 binds within the Kme aromatic cage of 53BP1, with the ethyl pyrrolidine group interacting with residues D1520, F1519, and Y1502 forming proposed cation- π interactions with the pyrrolidine ammonium group and D1521 in the Kme pocket. Additionally, the nicotinic acid functional group of UNC1554 was shown to interact with protein residue Tyr1500 where H4R19 binds, Figure 2.3. These interactions are noted in Figure 2.3 A and the binding pose is shown in Figure 2.3 A. These interactions are in agreement with HSQC titration experiments as shown in Figure 2.3B.

Figure 2.3: Unpublished NMR titration study of UNC1554 bound within the 53BP1 TTD



These unpublished results provided evidence that lead to a hypothesis that the nicotinic acid functional group of UNC1554 acted as a mimic of the H4R19 residue of the endogenous histone peptide substrate. These findings provided a basis to begin designing SAR experiments to optimize UNC1554. Additionally, NMR titration experiments show that UNC1554 has a $K_d = 66.8 \pm 18.9 \mu\text{M}$.

Structure-activity relationship (SAR) studies

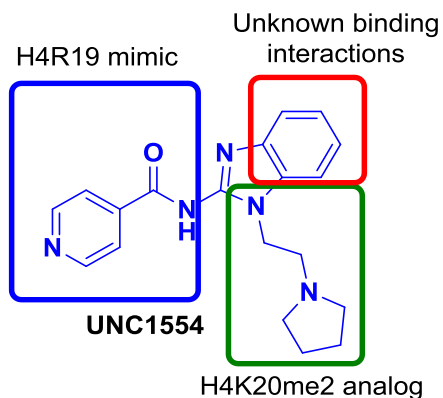
SAR studies were focused on optimizing the potency and selectivity of UNC1554 for the 53BP1 TTD. Medicinal chemistry efforts were primarily based on the X-ray structure of the 53BP1 TTD bound to endogenous H4K20me2 peptide substrate. Additionally, the NMR titration studies of UNC1554 bound to the 53BP1 TTD provided evidence of the small molecule's binding pose in a solution phase environment. This was very useful information as it allowed for development of a rational hypothesis of how synthetic changes of the structure of UNC1554 may

affect its ability to bind to various regions on the protein surface. These findings allowed for a more sound design of the small molecule SAR.

Synthetic strategy for SAR studies

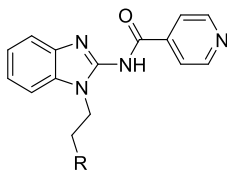
The NMR titration studies showed direct evidence that two interactions were involved in how UNC1554 binds to the 53BP1 TTD. The N-ethyl pyrrolidine functional group was shown to bind within the Kme binding pocket. This directly mimics the H4K20me endogenous substrate that also binds in the same manner. The second interaction observed was how the nicotinic acid functional group on UNC1554 bound in proximity to the H4R19 binding pocket on 53BP1 TTD. This functional group is hypothesized to mimic the peptide Arg19 binding interaction. These two interactions provided a basis for the key focus of medicinal chemistry efforts in SAR studies on UNC1554. Initial medicinal chemistry efforts will explore how variation of the functional group size on the amine head group of UNC1554 will affect binding potency and selectivity for the TTD of 53BP1. It was unknown what the steric tolerance of the Kme binding pocket of 53BP1 TTD is for the steric bulk of an amine head group of a small molecule. The next area of medicinal chemistry efforts explored how variations of the structural and electronic characteristics of the nicotinic acid functional group affect potency and selectivity. It is hypothesized that the nicotinic acid functional group mimics the binding activity of the H4R19 residue. The final region of UNC1554's structure that will be explored is how variation of the linker region between the N-ethyl pyrrolidine and nicotinic acid functional groups will affect compound potency and selectivity. These focus areas are shown in Figure 2.4.

Figure 2.4: Focus areas of UNC1554 SAR studies



Modification of amine head group

NMR structural data showed evidence that the N-ethyl pyrrolidine functional group of UNC1554 bound within the Kme binding pocket and provided in initial area of focus for SAR experiments. No previous studies have analyzed the steric tolerance that the Kme binding pocket has for amine functional groups with the exception of varying degrees of N-methylation in a histone peptide context. It was hypothesized that the Kme binding pocket would have a moderate degree of steric tolerance and be able to accept less sterically bulky amine functional groups. To determine the steric tolerance of the binding pocket, a small number of amine functional groups on the N-ethyl side chain of UNC1554 were prepared while maintaining the nicotinic acid functional group, Table 2.1. The smallest analog was the di-methyl N-ethyl amine, **1**, and the largest was a di-benzyl amine, **4**.

Table 2.1. SAR Studies and Selectivity Data of the N-ethylamine Group of **2** (UNC1554).^{a,b}

ID	R	53BP1	CBX7	JARID1A	L3MBTL1	L3MBTL3	MBTD1
1		>30	>30	>30	>30	4.5 ± 0.1	14.5 ± 0.7
2 (UNC1554)		22.7 ± 6.5	22.1 ± 8.4	>30	>30	10.6 ± 12.0	23.9 ± 9.2
3		>30	>30	>30	>30	>30	11.6 ± 1.1
4		>30	>30	>30	>30	>30	>30

^aIC₅₀ values are the average of at least 3 values ± the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 30 μM and compounds showing less than 50% inhibition at this concentration are labeled >30 μM.

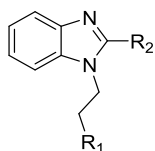
The IC₅₀ values and protein selectivity for these four analogs was measured by AlphaScreen assay against a panel of Kme binding proteins that consisted of tandem tudor, chromo, and MBT binding domains. It was observed that only analog **2** (UNC1554) had affinity for 53BP1 TTD with an IC₅₀ = 22.7 ± 6.5 μM while the other three analogs had IC₅₀ values >30 μM. The selectivity across a panel of Kme binding proteins showed that the N-ethylpyrrolidine head group provided a measure of selectivity for 53BP1 TTD, but in turn was also moderately potent for the other Kme reader proteins. It was also observed that as the N-ethyl amine head group increased in steric bulk, the selectivity and affinity of the analogs for the aromatic binding cage of 53BP1 TTD decreased. The di-benzyl analog, **4**, had no affinity for any of the screened Kme reader proteins. The small set of analogs showed that the Kme binding pocket was unable to bind sterically bulky amine functional groups, and was only able to successfully accommodate

a smaller group as shown with compound **2**. Further exploration of the steric tolerance of the Kme binding pocket was explored through SAR experiments where the nicotinic acid functional group was modified.

Modification of linker region

The next region of the UNC1554 structure to be modified through SAR experiments was the linker region between the benzimidazole core and nicotinic acid aromatic ring. In UNC1554, this linker was an amide bond. The NMR structure of UNC1554 bound to the 53BP1 TTD did not show any direct hydrogen bonds between the amide carbonyl group and the protein's surface residues. A clear hypothesis of whether the amide bond was necessary was not able to be generated and therefore a series of different linker groups were substituted in place of the amide bond to determine if modification affected potency and selectivity of the prepared analogs. A series of analogs was prepared that modified the amide bond into several different linkers, Table 2.2. Additional analogs were prepared where the amine functional group was modified from the N-ethyl pyrrolidine found on **2** (UNC1554). The types of linkers included a carbon-carbon bond between the benzimidazole ring and the aromatic ring, inclusion of a methylene spacer to that bond, and a N-benzyl functional group directly off the benzimidazole core.

Table 2.2 . SAR Studies and Selectivity Data for the Linker Group for **2 (UNC1554).^{a,b}**



ID	R ₁	R ₂	53BP1	CBX7	JARID1A	L3MBTL1	L3MBTL3	MBTD1
5			22.5 ± 7.1	>30	>30	>30	>30	>30
6			>30	>30	>30	>30	>30	>30
7			>30	>30	>30	>30	>30	>30
8			>30	>30	>30	>30	>30	>30
9			>30	5.3 ± 0.7	>30	>30	>30	>30
10			>30	12.3 ± 0.6	>30	>30	>30	>30
11			>30	>30	>30	26.3 ± 3.1	>30	>30
12			>30	>30	>30	>30	>30	>30
13			>30	>30	>30	20.0 ± 0.1	>30	>30
14			>30	>30	>30	>30	>30	>30
15			>30	>30	>30	18.9 ± 6.8	>30	>30
16			>30	>30	>30	>30	>30	>30
17			>30	>30	>30	>30	>30	>30
18			>30	>30	>30	13.0 ± 1.7	>30	>30
19			>30	>30	>30	20.1 ± 8.6	>30	>30

^aIC₅₀ values are the average of at least 3 values ± the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 30 μM and compounds showing less than 50% inhibition at this concentration are labeled >30 μM.

The IC₅₀ values and protein selectivity for these three analogs types were measured by AlphaScreen assay against a panel of Kme binding proteins. It was shown that only a single analog, **5**, provided any affinity for 53BP1 TTD with an IC₅₀ = 22.5 ± 7.1 μM. The majority of the linker analogs proved to be inactive regardless of the N-ethylamine head group or the type of modified linker that the analog possessed. Several of the analogs were shown to provide some affinity for CBX7 and L3MBTL1, analogs **9** and **10**, and analogs **11**, **13**, **15**, **18**, and **19**, respectively. The overall findings from this set of analogs were surprising as it showed that when the linker was modified from an amide bond, as found in **2**, the affinity and selectivity of the compound changed significantly. Regardless of whether the linker was rigidified as shown through analogs **5** - **16**, or made to be more flexible as in analogs **17** – **19**, the affinity for 53BP1 TTD did not increase with the exception of analog **5**. Additionally, the nicotinic acid head group was modified on a number of these analogs and only analog **5** was selective while retaining modest affinity for 53BP1 TTD. The aryl, morpholino functional group of **5** is proposed to mimic the aromatic and basic nitrogen properties found in the nicotinic acid of **2** (UNC1554). Overall, the modification of the linker region was found to be unsuccessful and the use of an amide bond was maintained in further SAR studies.

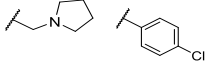
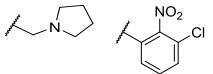
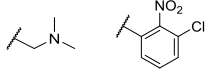
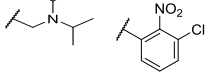
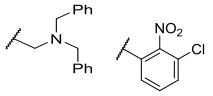
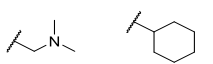
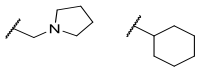
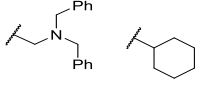
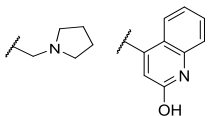
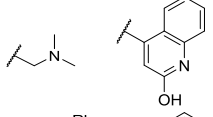
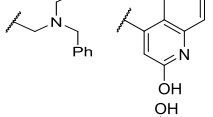
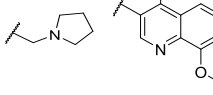
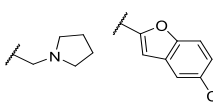
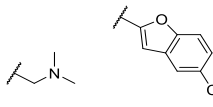
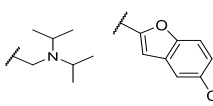
Modification of benzimidazole aromatic ring

Synthetic modifications of the N-ethyl amine functional group and the linker region did not provide an improvement in the potency and selectivity of **2** (UNC1554). The final region where SAR experiments were conducted was the aromatic nicotinic acid head group. The NMR studies of **2** (UNC1554) bound to 53BP1 TTD showed that the nicotinic acid functional group resides above the R19 binding pocket. It was hypothesized that substitution of different

functional groups at this position would increase the affinity of the compound for 53BP1 TTD by mimicking the R19 – Y1500 cation – π interactions of the endogenous peptide substrate. A large number of analogs were prepared in this series, Table 2.3, where the nicotinic acid was synthetically modified with functional groups that possessed varying degrees of steric bulk and different electronic characteristics. Modifications included substitution of the nicotinic acid by aromatic, non-aromatic, lactones, and sterically bulky acyl-acetamide analogs.

Table 2.3. SAR Studies of the Nicotinic Acid Head Group of **2** (UNC1554). ^{a,b}

ID	R ₁	R ₂	53BP1	CBX7	JARID1A	L3MBTL1	L3MBTL3	MBTD1
20			>30	>30	>30	>30	>30	>30
21			21.3 ± 9.1	>30	>30	>30	17.4 ± 11.5	>30
22			>30	>30	>30	>30	>30	>30
23			>30	>30	>30	>30	22.5 ± 13.0	>30
24			>30	>30	>30	>30	>30	>30
25			27 ± 5.2	>30	>30	>30	>30	13.7 ± 3.8
26			>30	>30	>30	>30	>30	>30
27			>30	>30	>30	>30	>30	>30
28			>30	>30	>30	>30	>30	>30
29			>30	>30	>30	>30	>30	>30

30		>30	>30	>30	>30	>30	>30
31		>30	24.7 ± 9.2	>30	>30	7.2 ± 0.9	>30
32		>30	20.6 ± 0.6	>30	>30	14.8 ± 10.2	>30
33		>30	>30	>30	>30	>30	>30
34		>30	>30	>30	>30	>30	>30
35		>30	>30	>30	>30	>30	>30
36		>30	>30	>30	>30	23.5 ± 9.2	>30
37		>30	>30	>30	>30	>30	>30
38		>30	>30	>30	>30	>30	>30
39		27.3 ± 4.3	24.7 ± 5.8	>30	>30	17.2 ± 12.0	26.2 ± 4.3
40		>30	>30	>30	>30	>30	>30
41		>30	>30	>30	>30	>30	>30
42		>30	>30	>30	>30	>30	>30
43		>30	>30	>30	>30	>30	>30
44		>30	>30	>30	>30	>30	>30

45		>30	>30	>30	>30	23.7 ± 5.7	>30
46		>30	>30	>30	>30	9.8 ± 6.6	>30
47		23.7 ± 7.4	>30	>30	>30	14.4 ± 14.2	>30
48		21.3 ± 7.7	22.5 ± 8.3	>30	>30	16.9 ± 14.3	22.5 ± 8.2
49		>30	>30	>30	>30	4.9 ± 0.1	25.5 ± 0.7
50		>30	>30	>30	>30	8.3 ± 1.4	>30
51		>30	>30	>30	>30	7.6 ± 0.8	>30
52		>30	>30	>30	>30	6.4 ± 0.7	>30

^aIC₅₀ values are the average of at least 3 values ± the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 30 μM and compounds showing less than 50% inhibition at this concentration are labeled >30 μM.

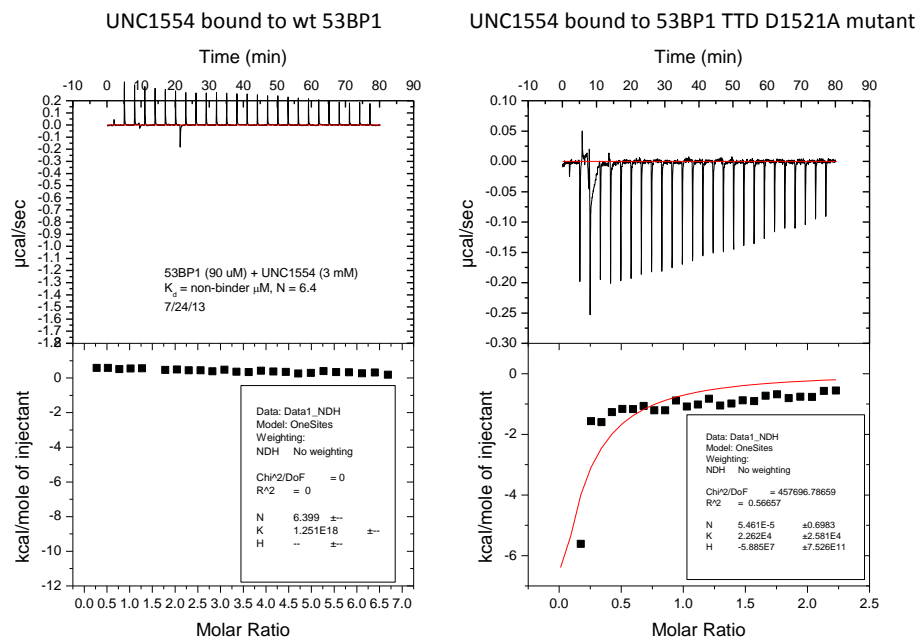
The IC₅₀ values and selectivity for these three analogs was measured by AlphaScreen assay against a panel of Kme binding proteins. It was observed that five different analogs had affinity for 53BP1 TTD, but unfortunately were not selective. Analogs **21**, **25**, **39**, **47**, and **48** were equipotent for 53BP1 TTD compared to **2**. The functional groups that were substituted for nicotinic acid were a phenyl, an isonicotinic acid, a hydroxyl quinolone, and two modified γ-lactones. It was proposed that the phenyl and isonicotinic acid groups would participate in a similar binding mode as the nicotinic acid in **2**. It was uncertain on how the hydroxyl quinolone and γ-lactones would mimic the binding action of the nicotinic acid in **2**. It was hypothesized that these three analogs may adopt a different binding pose than **2**. It was hypothesized that the N-ethylpyrrolidine or N-ethyltrimethylamine head group was thought to maintain its binding position within the aromatic cage, but the nicotinic acid mimics may rotate their binding position

to interact within the acidic/hydrophobic pocket of 53BP1 TTD. This type of binding mode would be reasonable for the γ -lactones that possessed a large hydrophobic heptane ring of the γ -lactone's amide group. This large and flexible hydrophobic group has the potential to bind in the acidic/hydrophobic pocket via hydrophobic and Van der Waals interactions. An additional structural characteristics of these analogs is that analogs **39** and **48** possessed an N-ethyldimethylamine head group and analogs **21**, **25**, and **49** possessed the N-ethylpyrrolidine head group. Overall this final SAR series showed only five analogs that possessed equipotent affinity for 53BP1 TTD.

Conclusions

A large number of analogs were prepared with different degrees of conformational rigidity, flexibility, and electronic characteristics in an effort to increase binding interactions based on structural data provided from NMR binding studies of **2** (UNC1554) bound to 53BP1 TTD. It was observed that over three SAR series, focused on different parts of the molecule, that the potency and affinity of **2** (UNC1554) was not improved. Synthetic modifications at three different position on the molecule; the N-ethyl amine head group, the linker region, and nicotinic acid region; were shown to be ineffective. During the final SAR series, the author of this work was introduced to the experimental methods of isothermal titration calorimetry (ITC) and upon conducting an ITC experiment using **2** (UNC1554) and 53BP1 TTD, it was found that no measureable binding was observed, Figure 2.5. An additional ITC experiment was conducted with the 53BP1 TTD D1521A inactive mutant and also found to provide no binding, Figure 2.5.

Figure 2.5: ITC binding data for **2** (UNC1554) against wt and D1521A inactive 53BP1 TTD



This binding data unfortunately showed that the key molecule of this set of SAR experiments provided no measureable binding data against 53BP1 TTD. The protein NMR binding data and NMR titration data counter this ITC observation. A potential explanation for why there was no measured binding is that the ΔH cancel each other out. The practice of validating the binding potency of any future “hit” molecules for 53BP1 TTD via the use of ITC was formally adopted. This binding data provided evidence of the intractable SAR efforts that were outlined above. The high avidity of the AlphaScreen bead-based assay has been known to be highly sensitive and can provide a more potent affinity measurement than other assay methods, however, it can also be prone to false positives.

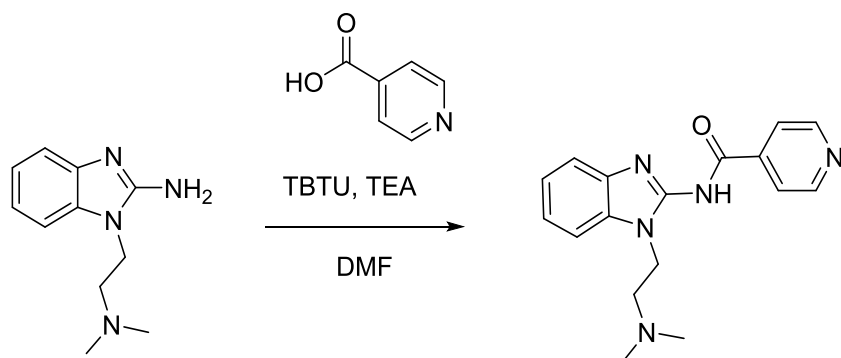
Experimental section

General Procedure for Chemical Synthesis

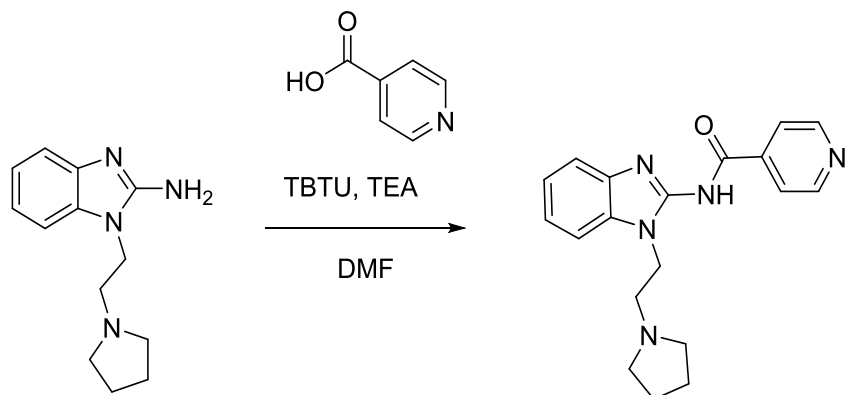
Analytical LCMS data for all compounds were acquired using an Agilent 6110 series system with the UV detector set to 220 and 254 nm. Samples were injected (<10 μ L) onto an Agilent Eclipse Plus 4.6 \times 50 mm, 1.8 μ m, C18 column at room temperature. A mobile phase of A (H₂O + 0.1% acetic acid) and B (MeOH + 0.1% acetic acid) was used with a linear gradient from 10% to 100% B in 5.0 min, followed by a flush at 100% B for another 2 minutes with a flow rate of 1.0 mL/min. Mass spectra data were acquired in positive ion mode using an Agilent 6110 single quadrupole mass spectrometer with an electrospray ionization source. Nuclear Magnetic Resonance (NMR) spectra were recorded on a Varian Mercury spectrometer at 400 MHz for proton (¹H NMR) and 100 MHz for carbon (¹³C NMR); chemical shifts are reported in ppm (δ). Analytical thin-layer chromatography (TLC) was performed with silica gel 60 F₂₅₄, 0.25 mm pre-coated TLC plates, generally using a 10% MeOH in DCM solvent system. TLC plates were visualized using UV₂₅₄, I₂ impregnated silica gel, potassium permanganate with charring, and phosphomolybdic acid with charring. Reverse phase chromatography was used to purify reaction mixtures to obtain final products using a Teledyne Isco CombiFlash Rf 200 chromatography unit equipped with the UV detector set to 220 nm and 254 nm. Samples were injected onto a RediSep Rf 30g C18 high performance Gold column at room temperature and collected at the previously mentioned wavelengths. Mobile phases of A (H₂O + 0.1% TFA) and B (MeOH) were used with a flow rate of 30 mL/min. A general gradient was used consisting of 0-2 minutes at 5% B, 5-15 minutes increasing from 5 to 100% B, and a 100% B flush for another 3 minutes. Small variations in this purification method were made as needed to achieve ideal

separation for each compound. All compounds that were evaluated in biochemical and biophysical assays had >95% purity as determined by ¹H NMR and LCMS.

Small molecule synthesis and analytical and structural characterization (LC-MS and NMR)

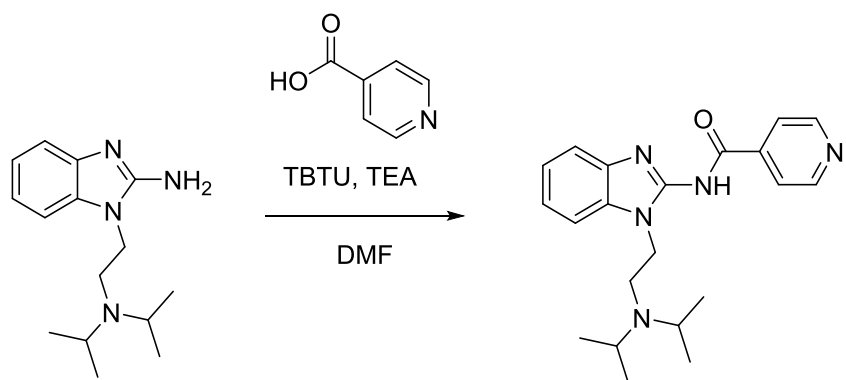


***N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)isonicotinamide (**1**):** A solution of 50 mg (0.41 mmol) isonicotinic acid, 169 mg (0.53 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 169 mL (1.22 mmol) triethylamine was stirred at rt for 15 min. Then 100 mg (0.41 mmol) **57** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 62 mg (60.3% yield) of pure product as a pale orange solid. ¹H NMR (400 MHz, cdcl₃) δ 8.70 (dd, *J* = 4.4, 1.6 Hz, 2H), 8.08 (dd, *J* = 4.4, 1.6 Hz, 2H), 7.35 – 7.16 (m, 4H), 4.35 (t, *J* = 7.1 Hz, 2H), 2.75 (t, *J* = 7.1 Hz, 2H), 2.34 (s, 6H). LC-MS (λ = 254 nm): 99%, *t*_R = 2.6 min. MS (ESI⁺): 310.2 [M+H]⁺.



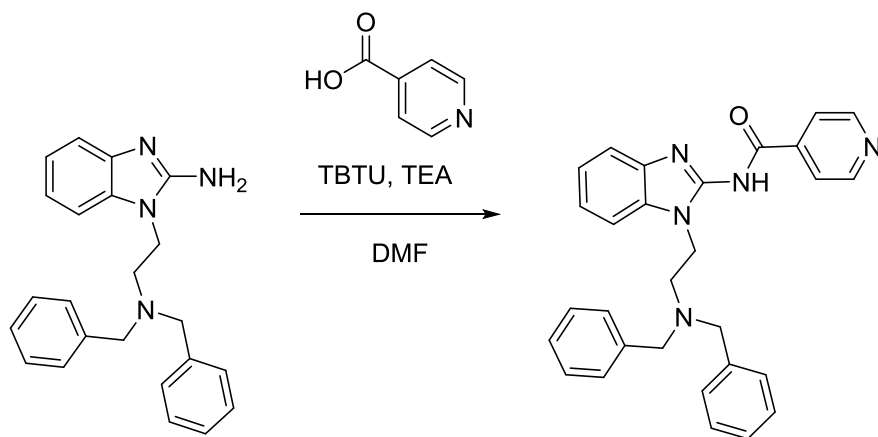
***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)isonicotinamide (2, UNC1554):**

UNC1554 was previously prepared by Dr. Lindsey J. Ingeman (CICBDD) and this sample was used as a standard throughout all screening assays. ^1H NMR (400 MHz, CDCl_3) δ 8.87 (dd, $J = 3.6, 1.6$ Hz, 2H), 8.57 (dd, $J = 3.1, 1.4$ Hz, 2H), 7.62 – 7.58 (m, 2H), 7.45 – 7.36 (m, 2H), 4.38 (t, $J = 6.3$ Hz, 2H), 3.81 (t, $J = 7.3$ Hz, 4H), 3.26 (s, 2H), 2.12 (s, 2H), 1.93 (s, 2H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 2.9$ min. MS (ESI $^+$): 337.2 $[\text{M}+\text{H}]^+$.



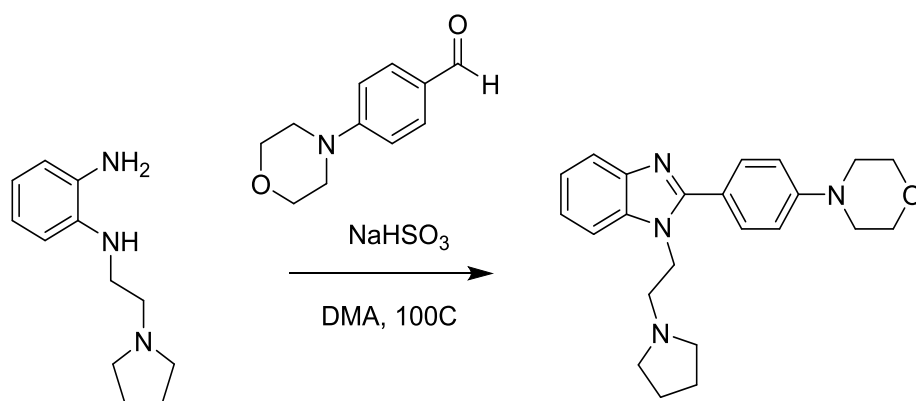
***N*-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)isonicotinamide (3):** A solution of 39 mg (0.32 mmol) isonicotinic acid, 132 mg (0.41 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 132 mL (0.95 mmol) triethylamine was

stirred at rt for 15 min. Then 100 mg (0.28 mmol) **58** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 68 mg (58.6% yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 8.70 (dd, *J* = 4.5, 1.6 Hz, 2H), 8.10 (dd, *J* = 4.4, 1.6 Hz, 2H), 7.34 – 7.18 (m, 4H), 4.24 – 4.15 (m, 2H), 3.05 (dt, *J* = 12.9, 6.5 Hz, 2H), 2.87 – 2.77 (m, 2H), 0.98 (d, *J* = 6.5 Hz, 12H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.2 min. MS (ESI⁺): 367.3 [M+H]⁺.

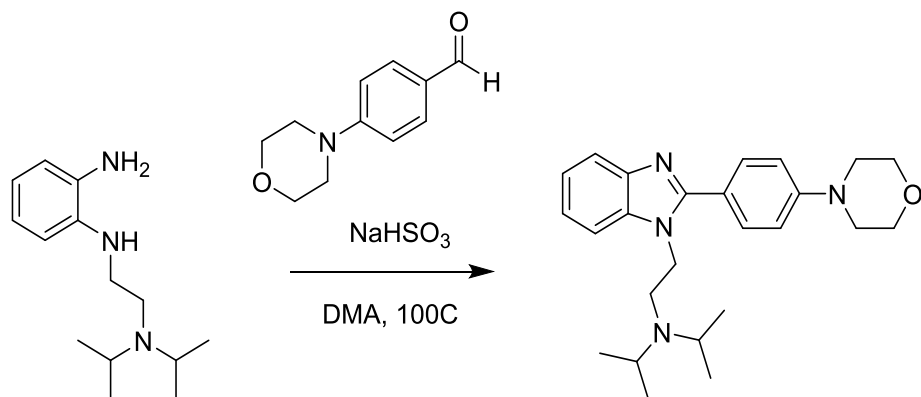


***N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)isonicotinamide (4):** A solution of 28 mg (0.23 mmol) isonicotinic acid, 95 mg (0.30 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 95 mL (0.68 mmol) triethylamine was stirred at rt for 15 min. Then 100 mg (0.28 mmol) **60** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 21 mg (20.1% yield) of pure product as a clear, yellow oil. ¹H NMR (400 MHz, cdcl₃) δ

8.66 (dd, $J = 4.4, 1.6$ Hz, 2H), 7.94 (dd, $J = 4.4, 1.6$ Hz, 2H), 7.32 (d, $J = 7.8$ Hz, 1H), 7.23 (td, $J = 7.8, 1.0$ Hz, 1H), 7.20 – 7.15 (m, 4H), 7.15 – 7.10 (m, 6H), 6.87 (d, $J = 8.0$ Hz, 1H), 4.31 (t, $J = 6.5$ Hz, 2H), 3.70 (s, 4H), 2.94 (t, $J = 6.5$ Hz, 2H), 1.74 (s, 1H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 6.1$ min. MS (ESI⁺): 462.3 [M+H]⁺.

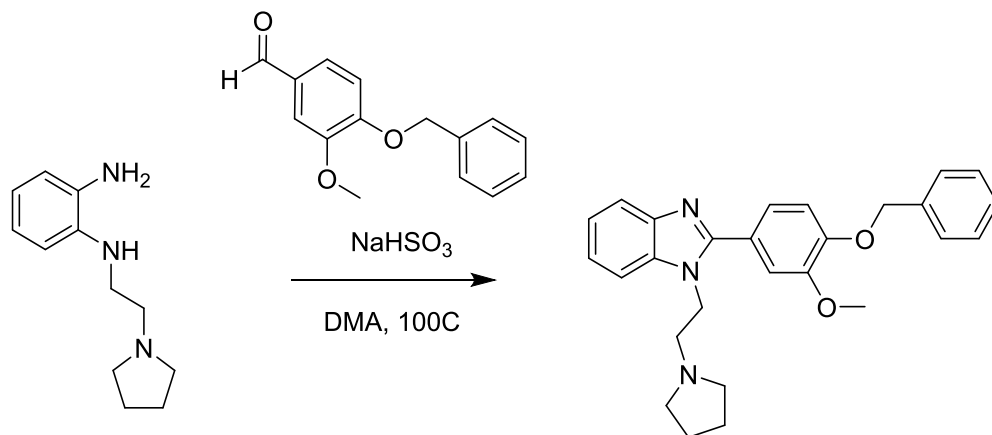


4-(4-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)phenyl)morpholine (5): To a solution of **53** (103 mg, 0.50 mmol) and 104 mg (1 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 96 mg (0.50 mmol) 4-morpholinobenzaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 78 mg (43.1% yield) of pure product as a pale orange solid. ¹H NMR (400 MHz, cd₃od) δ 7.69 – 7.62 (m, 3H), 7.60 – 7.54 (m, 1H), 7.30 (dq, $J = 14.7, 7.3, 1.3$ Hz, 2H), 7.13 (dd, $J = 9.3, 2.4$ Hz, 2H), 4.51 – 4.41 (m, 2H), 3.91 – 3.82 (m, 4H), 3.30 – 3.26 (m, 4H), 2.86 – 2.80 (m, 2H), 2.50 – 2.39 (m, 4H), 1.79 – 1.67 (m, 4H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 2.8$ min. MS (ESI⁺): 378.3 [M+2H]⁺.



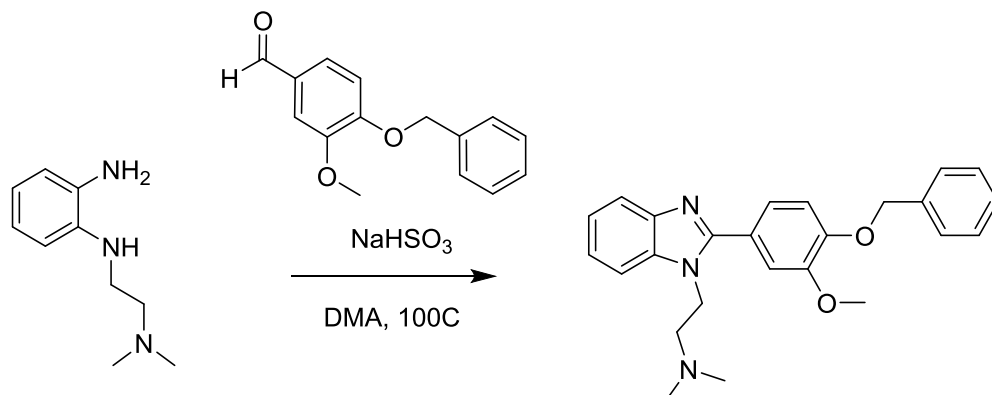
***N*-isopropyl-*N*-(2-(2-(4-morpholinophenyl)-1H-benzo[d]imidazol-1-yl)ethyl)propan-2-**

amine (6): To a solution of **55** (108 mg, 0.46 mmol) and 0.104 g (1 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 0.089 g (0.47 mmol) 4-morpholinobenzaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 75 mg (42.1% yield) of pure product as a pale orange solid. ¹H NMR (400 MHz, cd₃od) δ 7.71 – 7.62 (m, 3H), 7.51 (d, *J* = 7.3 Hz, 1H), 7.34 – 7.23 (m, 2H), 7.14 – 7.07 (m, 2H), 4.29 (t, *J* = 6.8 Hz, 2H), 3.88 – 3.81 (m, 4H), 3.28 – 3.22 (m, 4H), 2.84 (hept, *J* = 6.6 Hz, 2H), 2.65 (t, *J* = 6.8 Hz, 2H), 0.76 (d, *J* = 6.6 Hz, 12H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.3 min. MS (ESI⁺): 408.3 [M+2H]⁺.



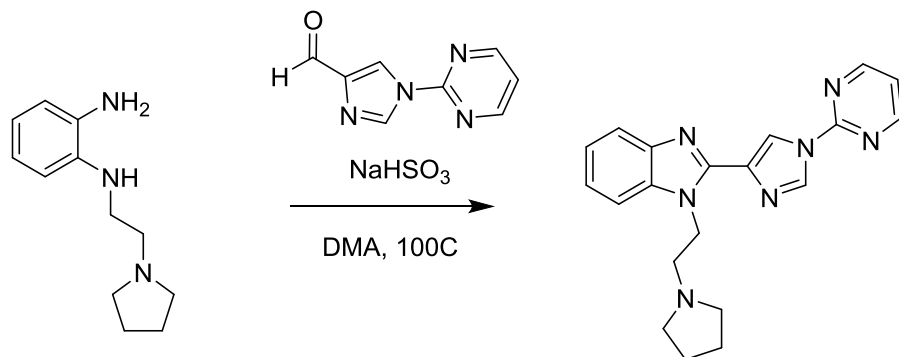
2-(4-(benzyloxy)-3-methoxyphenyl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (7):

To a solution of **53** (149 mg, 0.73 mmol) and 151 mg (1.45 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 176 mg (0.73 mmol) 4-(benzyloxy)-3-methoxybenzaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 90 mg (29.0% yield) of pure product as a clear, yellow oil. ¹H NMR (400 MHz, cdcl₃) δ 7.80 – 7.73 (m, 1H), 7.45 – 7.15 (m, 11H), 6.95 (d, *J* = 8.3 Hz, 1H), 5.18 (s, 2H), 4.40 – 4.30 (m, 2H), 3.95 – 3.89 (m, 3H), 2.88 – 2.80 (m, 2H), 2.43 (s, 4H), 1.76 – 1.64 (m, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.9 min. MS (ESI⁺): 429.3 [M+H]⁺.



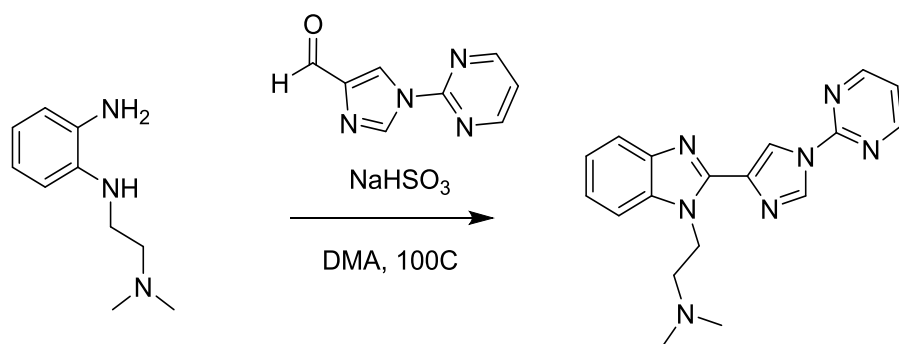
2-(2-(4-(benzyloxy)-3-methoxyphenyl)-1H-benzo[d]imidazol-1-yl)-N,N-dimethylethan-1-

amine (8): To a solution of **54** (70 mg, 0.39 mmol) and 81 mg (0.78 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 95 mg (0.39 mmol) 4-(benzyloxy)-3-methoxybenzaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 102 mg (65.3% yield) of pure product as a clear, orange oil. ¹H NMR (400 MHz, cdcl₃) δ 7.82 – 7.77 (m, 1H), 7.44 (d, *J* = 1.3 Hz, 2H), 7.42 – 7.33 (m, 4H), 7.33 – 7.11 (m, 5H), 6.98 (d, *J* = 8.5 Hz, 2H), 5.21 (s, 2H), 4.31 (t, *J* = 9.1 Hz, 2H), 3.95 (s, 3H), 2.69 (t, *J* = 9.5 Hz, 2H), 2.91 (s, 6H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.9 min. MS (ESI⁺): 403.3 [M+H]⁺.

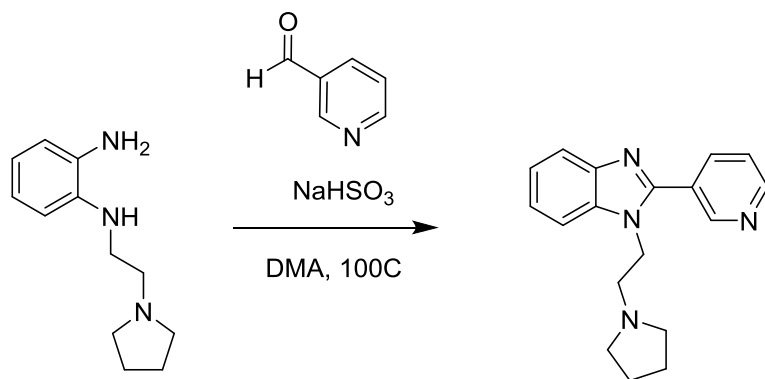


2-(1-(pyrimidin-2-yl)-1H-imidazol-4-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole

(9): To a solution of **53** (150 mg, 0.73 mmol) and 151 mg (1.45 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 126 mg (0.72 mmol) 1-(pyrimidin-2-yl)-1H-imidazole-4-carbaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 93 mg (35.6% yield) of pure product as a pale white solid. ¹H NMR (400 MHz, cdcl₃) δ 8.75 (dd, *J* = 8.3, 3.1 Hz, 3H), 8.70 (d, *J* = 1.4 Hz, 1H), 7.78 (ddd, *J* = 4.2, 2.7, 0.7 Hz, 1H), 7.49 – 7.44 (m, 1H), 7.31 – 7.26 (m, 2H), 7.25 (s, 0H), 5.04 – 4.91 (m, 2H), 2.97 (dd, *J* = 8.8, 7.1 Hz, 2H), 2.66 (s, 4H), 1.88 – 1.77 (m, 4H). LC-MS (λ = 254 nm): 99%, *t_R* = 3.4 min. MS (ESI⁺): 361.2 [M+H]⁺.

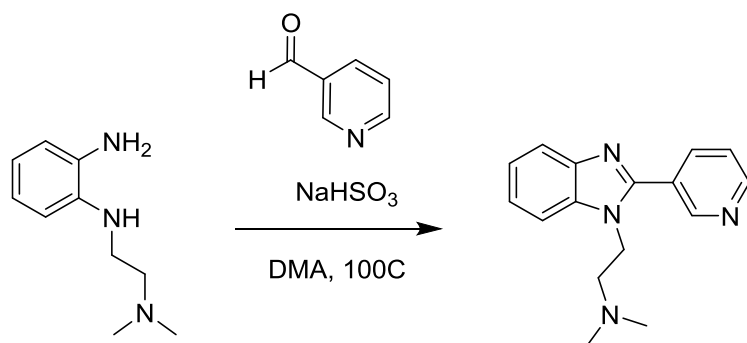


***N,N*-dimethyl-2-(2-(1-(pyrimidin-2-yl)-1H-imidazol-4-yl)-1H-benzo[d]imidazol-1-yl)ethan-1-amine (10):** To a solution of **55** (70 mg, 0.39 mmol) and 81 mg (0.78 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 68 mg (0.39 mmol) 1-(pyrimidin-2-yl)-1H-imidazole-4-carbaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 68 mg (52.3% yield) of pure product as a pale white solid. ¹H NMR (400 MHz, cdcl₃) δ 8.72 (d, *J* = 1.3 Hz, 1H), 8.70 – 8.63 (m, 3H), 7.74 (m, 1H), 7.41 (m, 1H), 7.28 – 7.22 (m, 2H), 7.18 (t, *J* = 4.8 Hz, 1H), 4.96 – 4.86 (m, 2H), 2.81 – 2.69 (m, 2H), 2.33 (s, 6H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.2 min. MS (ESI⁺): 335.2 [M+H]⁺.



2-(pyridin-3-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (11): To a solution of **53** (149 mg, 0.73 mmol) and 151 mg (1.45 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 0.07 mL (0.72 mmol) nicotinaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted

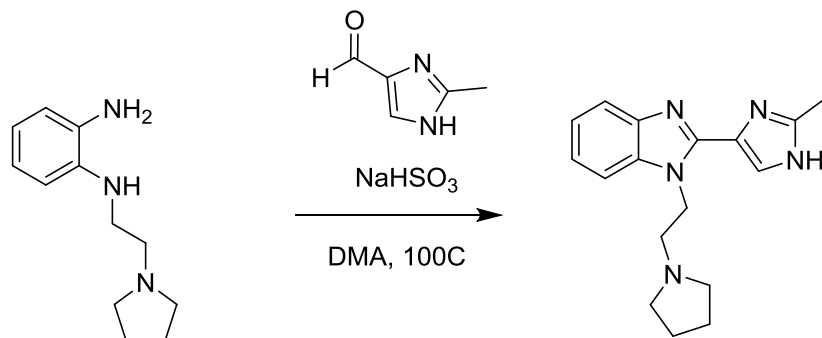
3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 145 mg (68.4% yield) of pure product as a clear, yellow oil. ¹H NMR (400 MHz, cdcl₃) δ 9.00 (dd, *J* = 2.3, 0.9 Hz, 1H), 8.67 (dd, *J* = 4.9, 1.7 Hz, 1H), 8.09 (m, 1H), 7.82 – 7.74 (m, 1H), 7.43 – 7.35 (m, 2H), 7.30 – 7.22 (m, 2H), 4.28 (t, *J* = 7.1 Hz, 2H), 2.88 – 2.77 (m, 2H), 2.42 – 2.29 (m, 4H), 1.64 (m, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 1.9 min. MS (ESI⁺): 293.2 [M+H]⁺.



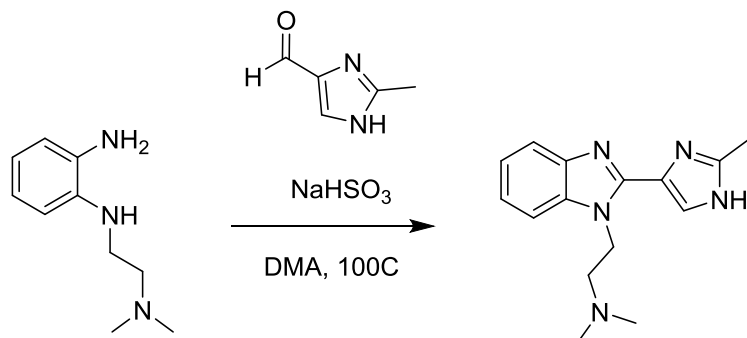
N,N-dimethyl-2-(2-(pyridin-3-yl)-1H-benzo[d]imidazol-1-yl)ethan-1-amine (12): To a solution of **54** (100 mg, 0.56 mmol) and 116 mg (1.1 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 0.05 mL (0.55 mmol) nicotinaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 135 mg (93.1% yield) of pure product as a clear oil. ¹H NMR (400 MHz, cdcl₃) δ 8.95 (dd, *J* = 2.2, 0.7 Hz, 1H), 8.64 (dd, *J* = 4.9, 1.7 Hz, 1H), 8.04 (m, 1H), 7.75 (m, 1H), 7.35 (m, 2H), 7.23 (m, 2H), 4.20 (t, *J* = 7.0 Hz, 2H), 2.58 (t, *J* =

7.0 Hz, 2H), 2.06 (s, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 1.9$ min. MS (ESI+): 268.2 $[M+H]^+$.

UNC2561A, MTP00140-27-1

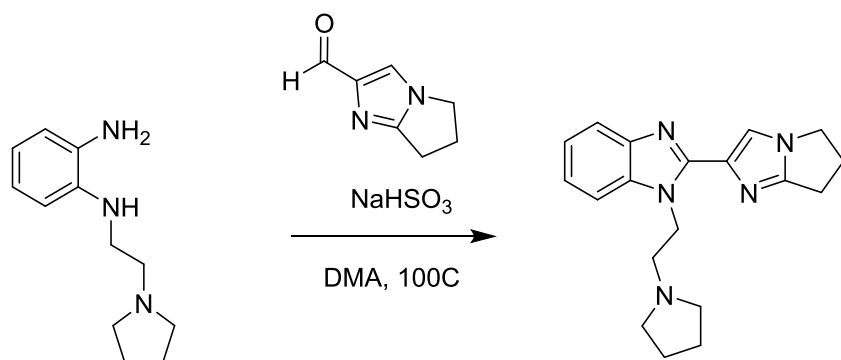


2-(2-methyl-1H-imidazol-4-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (13): To a solution of **53** (149 mg, 0.73 mmol) and 151 mg (1.45 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 79 mg (0.72 mmol) 2-methyl-1H-imidazole-4-carbaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 125 mg (58.4% yield) of pure product as a pale white solid. ¹H NMR (400 MHz, cdcl₃) δ 7.69 – 7.64 (m, 1H), 7.60 (s, 1H), 7.44 – 7.39 (m, 1H), 7.30 – 7.22 (m, 2H), 4.74 (s, 2H), 3.00 (t, $J = 7.3$ Hz, 2H), 2.70 – 2.60 (m, 4H), 2.40 (s, 3H), 1.87 – 1.72 (m, 4H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 2.2$ min. MS (ESI+): 297.3 $[M+H]^+$.



***N,N*-dimethyl-2-(2-(2-methyl-1H-imidazol-4-yl)-1H-benzo[d]imidazol-1-yl)ethan-1-amine**

(14): To a solution of **54** (100 mg, 0.56 mmol) and 116 mg (1.1 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 0.06 mg (0.55 mmol) 2-methyl-1H-imidazole-4-carbaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 106 mg (70.7% yield) of pure product as a white solid. ¹H NMR (400 MHz, cdcl₃) δ 7.36 – 7.30 (m, 1H), 7.26 (s, 1H), 7.10 – 7.05 (m, 1H), 6.96 – 6.88 (m, 2H), 4.57 – 4.38 (m, 2H), 2.54 – 2.41 (m, 2H), 2.04 (s, 3H), 2.02 (s, 6H). LC-MS (λ = 254 nm): 99%, *t*_R = 1.3 min. MS (ESI⁺): 271.2 [M+H]⁺.



2-(6,7-dihydro-5H-pyrrolo[1,2-a]imidazol-2-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-

benzo[d]imidazole (15): To a solution of **53** (62 mg, 0.30 mmol) and 63 mg (0.61 mmol)

sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 41 mg (0.30

mmol) 6,7-dihydro-5H-pyrrolo[1,2-a]imidazole-2-carbaldehydewas slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt).

The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried

onto silica. The product was purified by normal phase Teledyne Isco automated column system

to afford 34 mg (35.0% yield) of pure product as a pale white solid. ¹H NMR (400 MHz, cdcl₃) δ

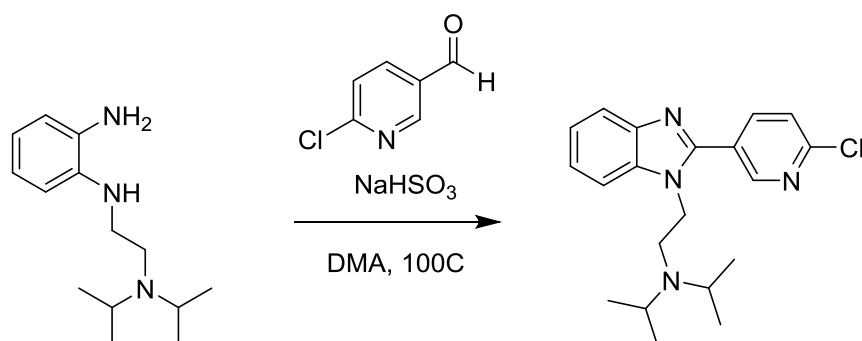
7.47 – 7.38 (m, 3H), 7.16 – 7.09 (m, 1H), 7.01 – 6.91 (m, 3H), 4.71 – 4.61 (m, 3H), 3.80 – 3.72

(m, 3H), 2.69 – 2.56 (m, 6H), 2.51 (s, 1H), 2.51 – 2.26 (m, 11H), 1.62 – 1.42 (m, 6H), 2.41 –

2.31 (m, 8H), 1.58 – 1.45 (m, 6H), 3.81 – 3.72 (m, 3H), 2.70 – 2.59 (m, 6H), 0.00 (s, 1H), 0.00

(ddd, *J* = 3.6, 3.2, 0.6 Hz, 1H), 4.73 – 4.59 (m, 3H), 7.16 – 7.10 (m, 1H), 6.98 – 6.92 (m, 3H).

LC-MS (λ = 254 nm): 99%, *t*_R = 2.6 min. MS (ESI⁺): 322.2 [M+H]⁺.



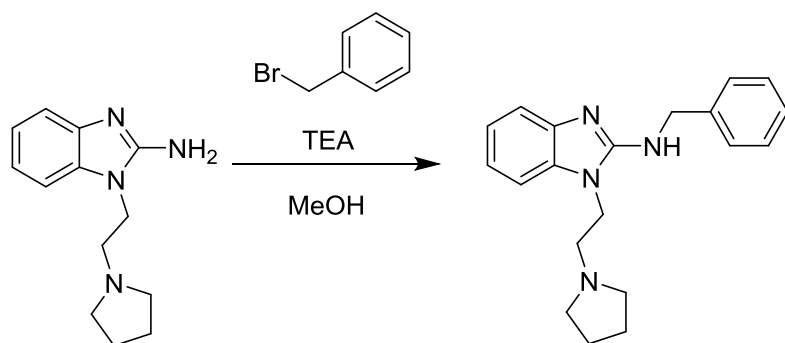
N-(2-(2-(6-chloropyridin-3-yl)-1H-benzo[d]imidazol-1-yl)ethyl)-N-isopropylpropan-2-

amine (16): To a solution of **55** (80 mg, 0.34 mmol) and 0.071 g (0.68 mmol) sodium bisulfite in

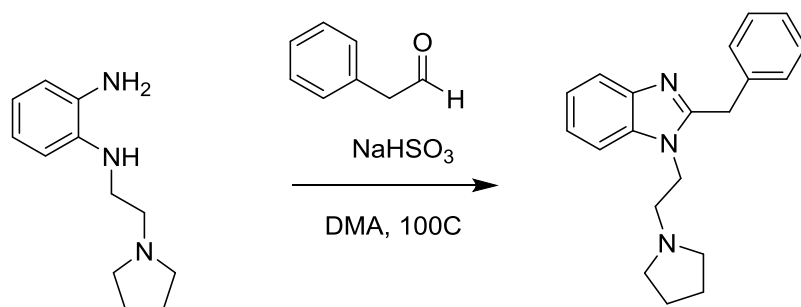
3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 0.048 g (0.34 mmol) 6-

chloronicotinaldehyde was slowly added over 2 min to the reaction. This mixture was then

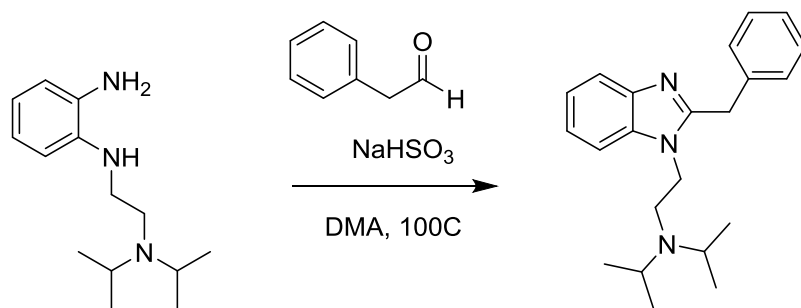
stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 67 mg (55.4% yield) of pure product as a clear yellow oil. ¹H NMR (400 MHz, cdcl₃) δ 8.91 (dd, *J* = 2.5, 0.7 Hz, 1H), 8.21 (dd, *J* = 8.3, 2.5 Hz, 1H), 7.82 – 7.78 (m, 1H), 7.47 (dd, *J* = 8.2, 0.7 Hz, 1H), 7.45 – 7.40 (m, 1H), 7.36 – 7.28 (m, 2H), 4.20 (t, *J* = 6.5 Hz, 2H), 2.92 – 2.80 (m, 3H), 2.75 (t, *J* = 6.5 Hz, 2H), 0.79 (d, *J* = 6.6 Hz, 12H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.4 min. MS (ESI⁺): 357.2 [M+H]⁺.



N-benzyl-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzodimidazol-2-amine (17): To a solution of **56** (0.1 g, 0.43 mmol) in 5 mL MeOH was added 0.08 mL (0.65 mmol) benzyl bromide and 0.24 mL (1.74 mmol) triethylamine and the solution was stirred overnight at rt. The crude reaction was dried to residue, dissolved in DCM and dried onto silica and purified by normal phase Teledyne Isco automated column system to afford 41 mg (29.5% yield) of pure product as a orange solid. ¹H NMR (400 MHz, cdcl₃) δ 8.16 (s, 1H), 7.49 (dd, *J* = 7.6, 0.9 Hz, 1H), 7.41 (dd, *J* = 7.8, 1.1 Hz, 2H), 7.36 – 7.21 (m, 3H), 7.08 (ddd, *J* = 7.8, 7.1, 1.6 Hz, 1H), 7.05 – 6.95 (m, 2H), 4.62 (d, *J* = 4.8 Hz, 2H), 3.98 – 3.91 (m, 2H), 2.87 – 2.78 (m, 2H), 2.52 – 2.41 (m, 2H), 1.41 – 1.28 (m, 2H). LC-MS (λ = 254 nm): 99%, *t*_R = 2.4 min. MS (ESI⁺): 321.2 [M+H]⁺.

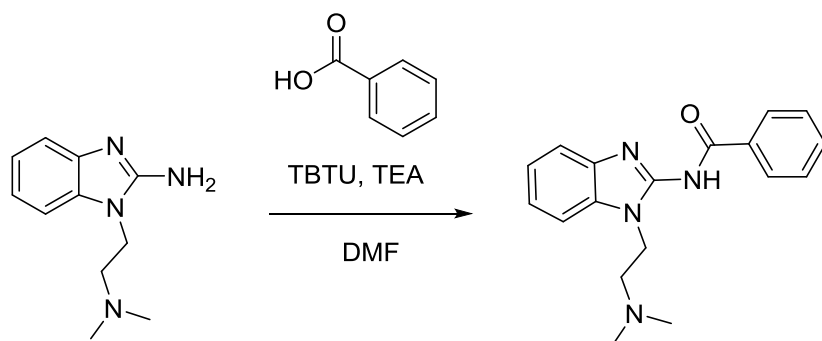


2-benzyl-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (18): To a solution of **53** (214 mg, 1.04 mmol) and 108 mg (1.04 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 0.12 mL (1.04 mmol) 2-phenylacetaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 110 mg (35.0% yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 7.80 – 7.74 (m, 1H), 7.34 – 7.20 (m, 9H), 4.35 (s, 2H), 4.17 – 4.11 (m, 2H), 2.57 – 2.50 (m, 2H), 2.50 – 2.42 (m, 4H), 1.81 – 1.71 (m, 4H), 1.66 (s, 2H). LC-MS (λ = 254 nm): 99%, t_R = 2.9 min. MS (ESI⁺): 306.3 [M+H]⁺.



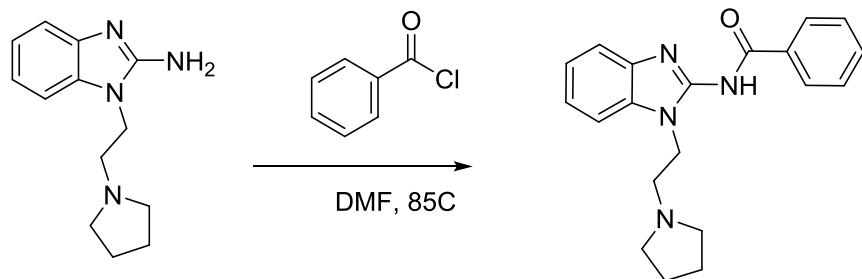
N-(2-(2-benzyl-1H-benzo[d]imidazol-1-yl)ethyl)-N-isopropylpropan-2-amine (19): To a solution of **55** (79 mg, 0.34 mmol) and 69 mg (0.66 mmol) sodium bisulfite in 3 mL dimethylacetamide was stirred at 100°C for 2 hr. Then 0.039 mL (0.33 mmol) 2-

phenylacetaldehyde was slowly added over 2 min to the reaction. This mixture was then stirred at 100°C for 2 hr and cooled to room temperature (rt). The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 58 mg (51.2% yield) of pure product as a clear yellow oil. ¹H NMR (400 MHz, cd₃od) δ 7.66 – 7.60 (m, 1H), 7.45 – 7.40 (m, 1H), 7.32 – 7.17 (m, 7H), 4.03 (t, *J* = 6.5 Hz, 2H), 3.00 – 2.87 (m, 2H), 2.61 (t, *J* = 6.5 Hz, 2H), 0.84 (d, *J* = 6.6 Hz, 12H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.6 min. MS (ESI⁺): 336.3 [M+H]⁺.

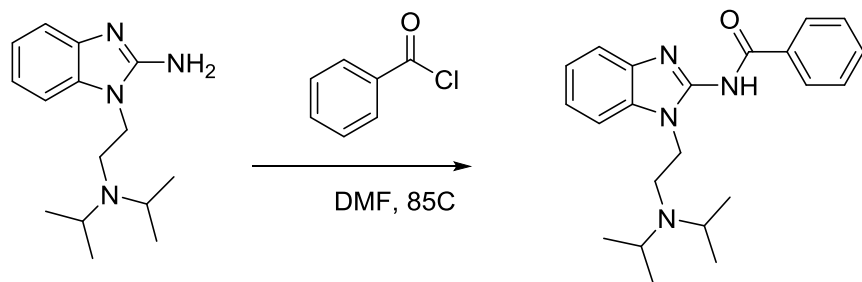


***N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (20):** A solution of **57** (28 mg, 0.23 mmol) benzoic acid, 95 mg (0.30 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 96 mL (0.68 mmol) triethylamine was stirred at rt for 15 min. Then 57 mg (0.28 mmol) **59** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 37 mg (43.0 % yield) of pure product as a pale white solid. ¹H NMR (400 MHz, cdcl₃)

δ 8.34 (d, J = 6.8 Hz, 2H), 7.44 (m, 3H), 7.33 – 7.16 (m, 4H), 4.41 – 4.33 (m, 2H), 2.82 – 2.73 (m, 2H), 2.37 (s, 6H). LC-MS (λ = 254 nm): 99%, t_R = 3.6 min. MS (ESI⁺): 310.2 [M+H]⁺.

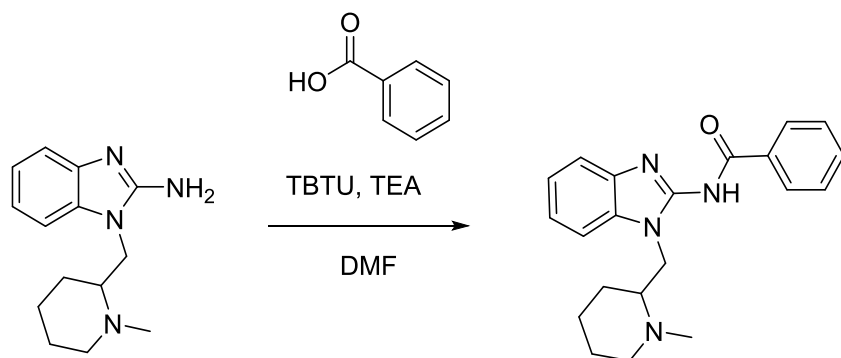


***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (21):** To a solution of **56** (119 mg, 0.52 mmol) in 6 mL DMF was added 0.08 mL (0.62 mmol) benzoyl chloride and the reaction was stirred at rt. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 14 mg (7.9% yield) of pure product as a white solid. ¹H NMR (400 MHz, cdcl₃) δ 8.41 – 8.31 (m, 2H), 7.54 – 7.41 (m, 3H), 7.36 – 7.18 (m, 5H), 4.47 – 4.37 (m, 2H), 3.00 – 2.91 (m, 2H), 2.77 – 2.64 (m, 4H), 1.82 (td, J = 6.4, 3.1 Hz, 4H), 1.72 (s, 1H). LC-MS (λ = 254 nm): 99%, t_R = 3.6 min. MS (ESI⁺): 335.3 [M+H]⁺.



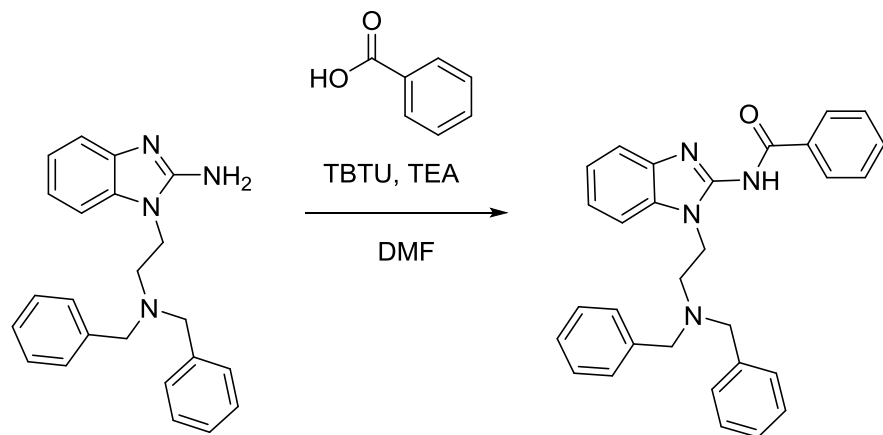
***N*-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (22):** To a solution of **58** (85 mg, 0.33 mmol) in 6 mL DMF was added 0.5 mL (0.39 mmol) benzoyl chloride and the reaction was stirred at rt. The reaction was quenched using sat. aqueous sodium bicarbonate

and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 21 mg (17.7% yield) of pure product as a clear oil. ¹H NMR (400 MHz, cdcl₃) δ 8.39 – 8.32 (m, 2H), 7.53 – 7.36 (m, 3H), 7.33 – 7.18 (m, 4H), 4.26 – 4.15 (m, 2H), 3.09 (dt, *J* = 13.0, 6.5 Hz, 2H), 2.91 – 2.83 (m, 2H), 1.04 (d, *J* = 6.5 Hz, 12H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.7 min. MS (ESI⁺): 365.3 [M+H]⁺.

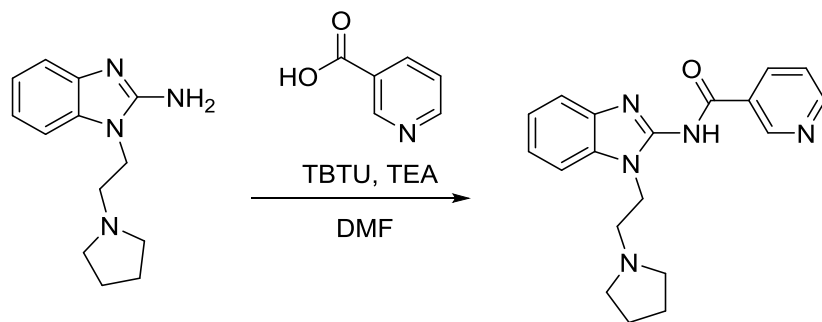


***N*-(1-((1-methylpiperidin-2-yl)methyl)-1H-benzo[d]imidazol-2-yl)benzamide (23):** A

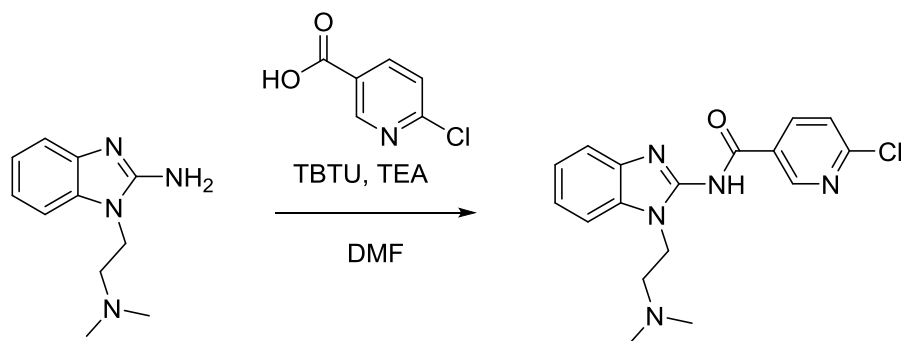
solution of **59** (42 mg, 0.34 mmol) benzoic acid, 144 mg (0.45 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 144 mL (1.03 mmol) triethylamine was stirred at rt for 15 min. Then 100 mg (0.41 mmol) **61** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 45 mg (37.8% yield) of pure product as a white solid. ¹H NMR (400 MHz, cdcl₃) δ 8.19 – 8.08 (m, 2H), 7.34 – 7.15 (m, 3H), 7.14 – 6.94 (m, 4H), 4.38 (dd, *J* = 13.7, 4.2 Hz, 1H), 3.91 (dd, *J* = 13.7, 9.7 Hz, 1H), 2.76 – 2.66 (m, 1H), 2.43 (tt, *J* = 9.8, 3.7 Hz, 1H), 2.37 (s, 3H), 2.04 – 1.94 (m, 1H), 1.52 – 1.36 (m, 3H), 1.31 – 1.12 (m, 2H), 1.03 – 0.86 (m, 1H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.8 min. MS (ESI⁺): 350.2 [M+H]⁺.



***N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (24):** A solution of **60** (75 mg, 0.61 mmol) benzoic acid, 256 mg (0.79 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 257 mL (1.8 mmol) triethylamine was stirred at rt for 15 min. Then 262 mg (0.73 mmol) **62** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 84 mg (29.7% yield) of pure product as a yellow solid. ¹H NMR (400 MHz, cd₃od) δ 8.16 – 8.09 (m, 2H), 7.36 – 7.23 (m, 3H), 7.15 – 7.07 (m, 5H), 7.07 – 6.99 (m, 7H), 6.94 (td, *J* = 7.9, 0.9 Hz, 1H), 6.68 (d, *J* = 7.9 Hz, 1H), 4.16 (t, *J* = 6.7 Hz, 2H), 3.58 (s, 4H), 2.81 (t, *J* = 6.7 Hz, 2H). LC-MS (λ = 254 nm): 99%, *t*_R = 6.5 min. MS (ESI⁺): 463.2 [M+H]⁺.

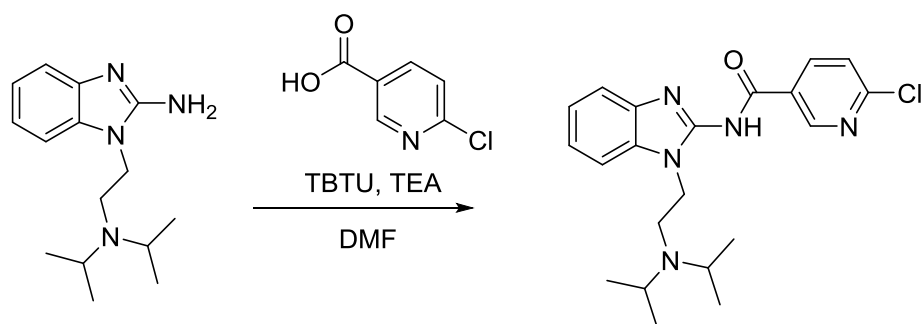


***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)nicotinamide (25):** A solution of **56** (50 mg, 0.41 mmol) nicotinic acid, 169 mg (0.53 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 169 mL (1.2 mmol) triethylamine was stirred at rt for 15 min. Then 112 mg (0.49 mmol) **59** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 70 mg (51.5 % yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 9.51 (dd, *J* = 2.1, 0.8 Hz, 1H), 8.65 (dd, *J* = 4.8, 1.8 Hz, 1H), 8.50 (dt, *J* = 7.9, 1.9 Hz, 1H), 7.36 – 7.13 (m, 5H), 4.42 – 4.29 (m, 2H), 2.89 (dd, *J* = 8.8, 6.1 Hz, 2H), 2.69 – 2.56 (m, 4H), 1.82 – 1.69 (m, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.0 min. MS (ESI⁺): 337.2 [M+H]⁺.

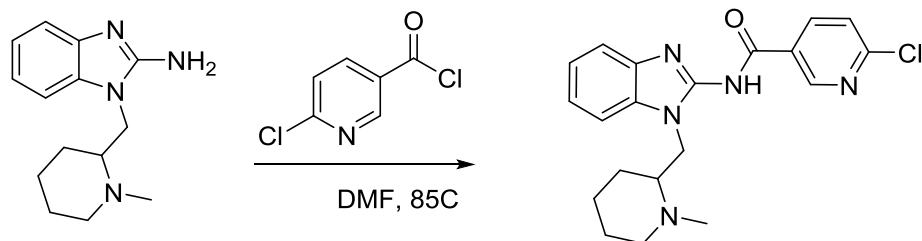


6-chloro-*N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)nicotinamide (26): A solution of **57** (100 mg, 0.63 mmol) 6-chloronicotinic acid, 265 mg (0.83 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 265 mL (1.9 mmol) triethylamine was stirred at rt for 15 min. Then 156 mg (0.76 mmol) **60** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase

Teledyne Isco automated column system to afford 132 mg (60.5 % yield) of pure product as a pale orange solid. ¹H NMR (400 MHz, cdcl₃) δ 9.29 (dd, *J* = 2.3, 0.7 Hz, 1H), 8.46 (dd, *J* = 8.3, 2.3 Hz, 1H), 7.37 (dd, *J* = 8.2, 0.6 Hz, 1H), 7.35 – 7.22 (m, 4H), 4.35 (t, *J* = 7.1 Hz, 2H), 2.75 (t, *J* = 7.1 Hz, 2H), 2.35 (s, 6H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.6 min. MS (ESI⁺): 346.2 [M+H]⁺.

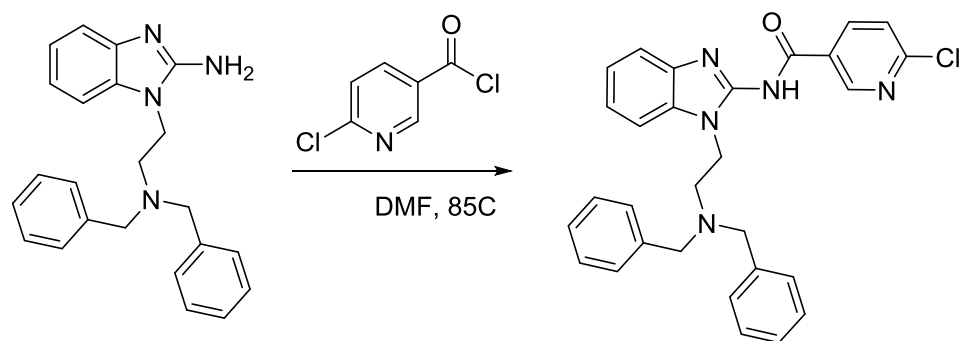


6-chloro-N-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)nicotinamide (27): A solution of **58** (102 mg, 0.64 mmol) 6-chloronicotinic acid, 265 mg (0.82 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 265 mL (1.9 mmol) triethylamine was stirred at rt for 15 min. Then 198 mg (0.76 mmol) **58** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 171 mg (67.3 % yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 9.28 (dd, *J* = 2.3, 0.7 Hz, 1H), 8.47 (dd, *J* = 8.3, 2.3 Hz, 1H), 7.33 (dd, *J* = 8.3, 0.6 Hz, 1H), 7.29 – 7.16 (m, 4H), 4.17 (dd, *J* = 8.1, 6.7 Hz, 2H), 3.12 – 2.96 (m, 2H), 2.80 (dd, *J* = 12.7, 5.4 Hz, 2H), 0.97 (d, *J* = 6.6 Hz, 12H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.8 min. MS (ESI⁺): 402.2 [M+H]⁺.



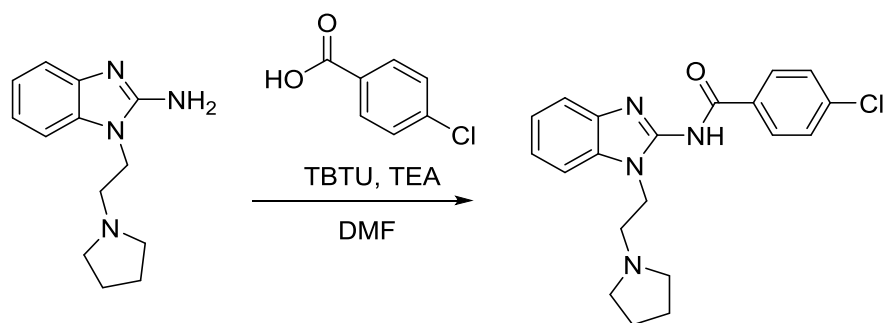
6-chloro-N-(1-((1-methylpiperidin-2-yl)methyl)-1H-benzo[d]imidazol-2-yl)nicotinamide (28):

To a solution of **59** (56 mg, 0.23 mmol) in 6 mL DMF was added 48 g (0.27 mmol) 6-chloronicotinoyl chloride and the reaction was stirred at rt. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 19 mg (21.5% yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 9.31 (d, *J* = 2.3 Hz, 1H), 8.48 (dd, *J* = 8.2, 2.3 Hz, 1H), 7.78 (d, *J* = 6.1 Hz, 1H), 7.39 (d, *J* = 8.2 Hz, 1H), 7.35 – 7.29 (m, 1H), 7.29 – 7.22 (m, 2H), 5.17 (s, 1H), 3.26 (s, 1H), 2.91 (dd, *J* = 13.8, 5.1 Hz, 1H), 2.78 (dd, *J* = 11.1, 6.3 Hz, 1H), 2.68 (d, *J* = 9.3 Hz, 1H), 2.43 (d, *J* = 14.0 Hz, 4H), 2.08 – 1.97 (m, 1H), 1.97 – 1.88 (m, 2H), 1.87 – 1.74 (m, 1H), 1.70 (dd, *J* = 15.5, 5.4 Hz, 2H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.1 min. MS (ESI⁺): 385.2 [M+H]⁺.



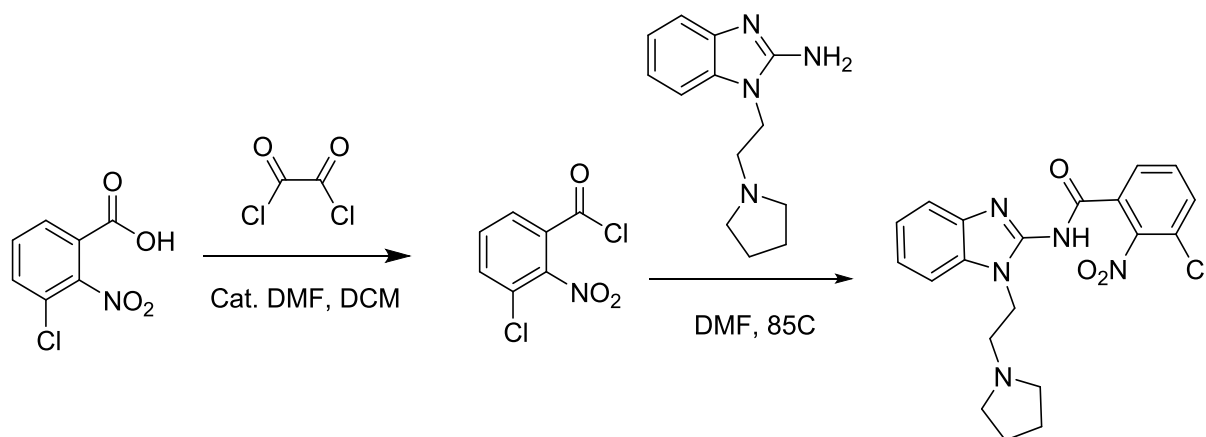
6-chloro-N-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)nicotinamide (29): To a solution of **60** (358 mg, 1 mmol) in 6 mL DMF was added 211 g (1.2 mmol) 6-chloronicotinoyl

chloride and the reaction was stirred at rt. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 150 mg (30.1% yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 9.12 (d, *J* = 2.2 Hz, 1H), 8.29 (dd, *J* = 8.3, 2.3 Hz, 1H), 7.31 (dd, *J* = 16.7, 8.1 Hz, 2H), 7.22 (d, *J* = 7.2 Hz, 1H), 7.19 – 7.06 (m, 11H), 6.87 (d, *J* = 8.0 Hz, 1H), 4.29 (t, *J* = 6.4 Hz, 2H), 3.68 (s, 4H), 2.93 (t, *J* = 6.4 Hz, 2H). LC-MS (λ = 254 nm): 99%, *t*_R = 6.4 min. MS (ESI⁺): 497.3 [M+H]⁺.



4-chloro-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzod[1,2-b:4,5-b']imidazol-2-yl)benzamide (30): A solution of 50 mg (0.32 mmol) 4-chlorobenzoic acid, 133 mg (0.41 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 133 mL (0.95 mmol) triethylamine was stirred at rt for 15 min. Then 88 mg (0.38 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 36 mg (30.5 % yield) of pure product as a white solid. ¹H NMR (400 MHz, cdcl₃) δ 8.27 (d, *J* = 8.4 Hz, 2H), 7.39 (d, *J* = 8.3 Hz, 2H), 7.32 – 7.18 (m,

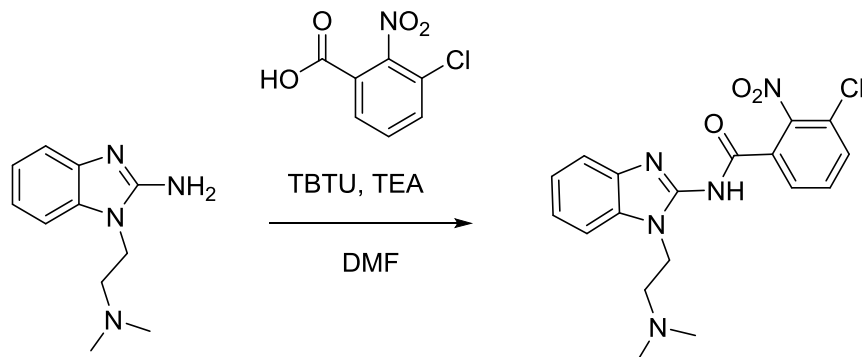
4H), 4.38 (t, $J = 7.4$ Hz, 2H), 2.92 (t, $J = 7.4$ Hz, 2H), 2.67 (m, 4H), 1.80 (m, 4H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.2$ min. MS (ESI⁺): 371.2 [M+H]⁺.



3-chloro-2-nitro-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (31):

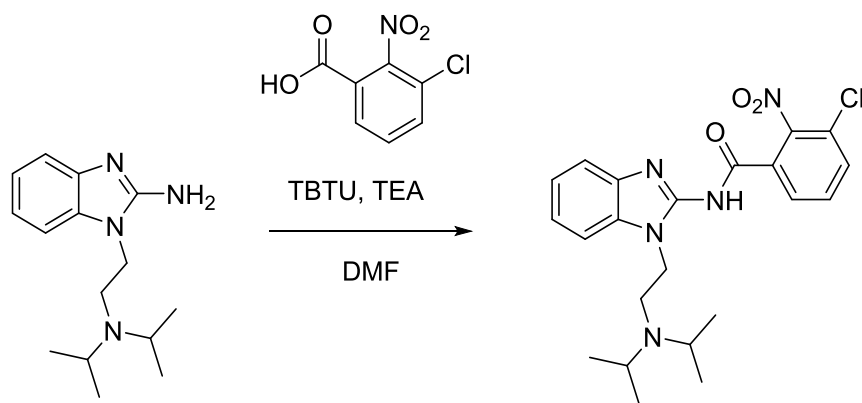
To a solution of 3-chloro-2-nitrobenzoic acid (250 mg, 1.24 mmol) and a catalytic amount of DMF (2 drops) in 6.5 mL DCM added 0.13 mL (1.49 mmol) oxalyl dichloride and stirred at rt for 2 hr. The reaction was then dried to residue as the crude acyl chloride intermediate and moved directly forward.

The crude intermediate was dissolved in 2 mL DMF and 0.15 g (0.65 mmol) **56** and stirred overnight at 85°C. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 59 mg (16.5% yield) of pure product as a white solid. ¹H NMR (400 MHz, cdCl₃) δ 8.25 (dd, $J = 7.9, 1.4$ Hz, 1H), 7.60 (dd, $J = 8.0, 1.4$ Hz, 1H), 7.48 (t, $J = 7.9$ Hz, 1H), 7.39 – 7.35 (m, 1H), 7.33 – 7.26 (m, 3H), 4.33 (t, $J = 8.7$, 2H), 2.96 (t, $J = 9.0$, 2H), 2.74 – 2.66 (m, 4H), 1.86 – 1.79 (m, 4H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.9$ min. MS (ESI⁺): 415.2 [M+H]⁺.



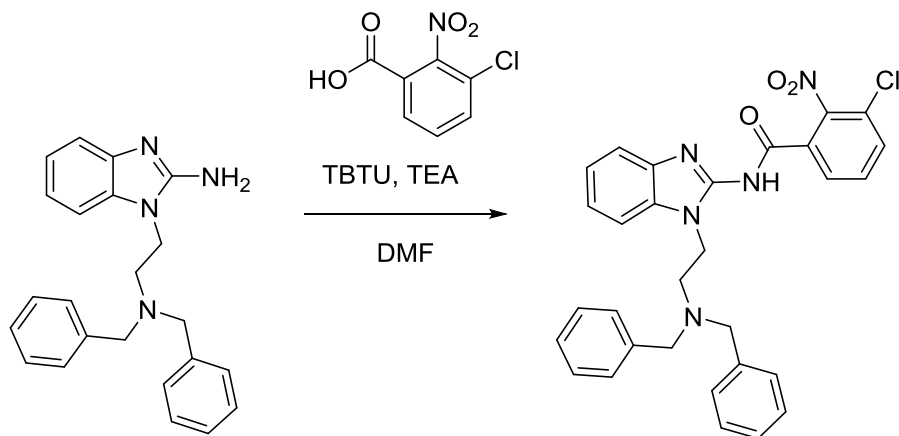
3-chloro-*N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-2-nitrobenzamide (32):

A solution of 100 mg (0.45 mmol) 3-chloro-2-nitrobenzoic acid, 189 mg (0.59 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 190 mL (1.35 mmol) triethylamine was stirred at rt for 15 min. Then 111 mg (0.54 mmol) **57** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 158 mg (89.2% yield) of pure product as a clear oil. ¹H NMR (400 MHz, cdcl₃) δ 8.17 (dd, *J* = 7.9, 1.4 Hz, 1H), 7.51 (dd, *J* = 8.0, 1.4 Hz, 1H), 7.39 (t, *J* = 7.9 Hz, 1H), 7.28 – 7.12 (m, 4H), 4.20 (t, *J* = 7.1 Hz, 2H), 2.66 (t, *J* = 7.1 Hz, 2H), 2.31 (s, 6H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.9 min. MS (ESI⁺): 390.2 [M+H]⁺.



3-chloro-*N*-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-2-nitrobenzamide

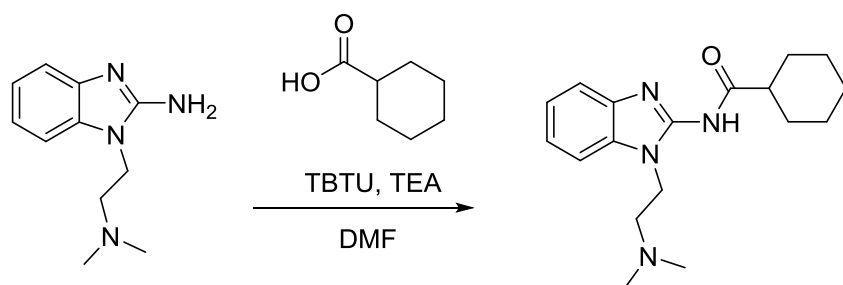
(33): A solution of **58** (75 mg, 0.34 mmol) 3-chloro-2-nitrobenzoic acid, 142 mg (0.44 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 143 mL (1.02 mmol) triethylamine was stirred at rt for 15 min. Then 106 mg (0.41 mmol) **58** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 40 mg (26.5% yield) of pure product as a pale brown solid. ¹H NMR (400 MHz, cdcl₃) δ 8.27 (dd, *J* = 7.9, 1.4 Hz, 1H), 7.56 (dd, *J* = 8.0, 1.4 Hz, 1H), 7.44 (t, *J* = 7.9 Hz, 1H), 7.34 – 7.20 (m, 3H), 4.11 (t, *J* = 6.6 Hz, 2H), 3.09 – 2.94 (m, 2H), 2.84 (t, *J* = 6.6 Hz, 2H), 0.91 (d, *J* = 6.6 Hz, 12H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.2 min. MS (ESI⁺): 446.2 [M+H]⁺.



3-chloro-*N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-2-nitrobenzamide (34):

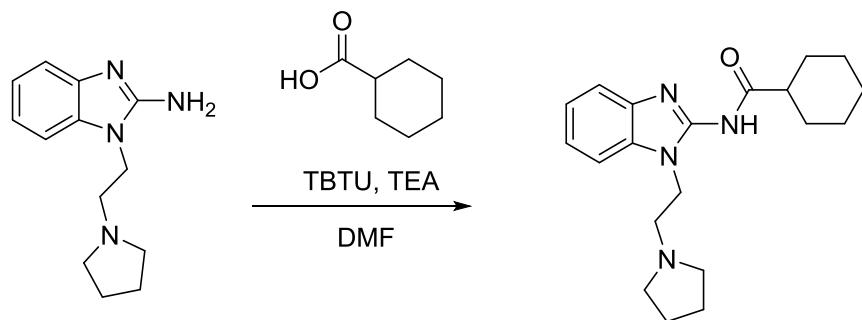
A solution of 75 mg (0.34 mmol) 3-chloro-2-nitrobenzoic acid, 142 mg (0.44 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 146 mL (1.0 mmol) triethylamine was stirred at rt for 15 min. Then 146 mg (0.41 mmol) **60** was added and

the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 196 mg (61.9% yield) of pure product as a yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 8.08 (dd, *J* = 7.9, 1.4 Hz, 1H), 7.53 (dd, *J* = 8.0, 1.4 Hz, 1H), 7.38 (t, *J* = 7.9 Hz, 1H), 7.29 – 7.08 (m, 13H), 6.82 (d, *J* = 8.0 Hz, 1H), 4.20 (t, *J* = 6.4 Hz, 2H), 3.68 (s, 4H), 2.90 (t, *J* = 6.4 Hz, 2H). LC-MS (λ = 254 nm): 99%, *t*_R = 6.6 min. MS (ESI+): 542.2 [M+H]⁺.

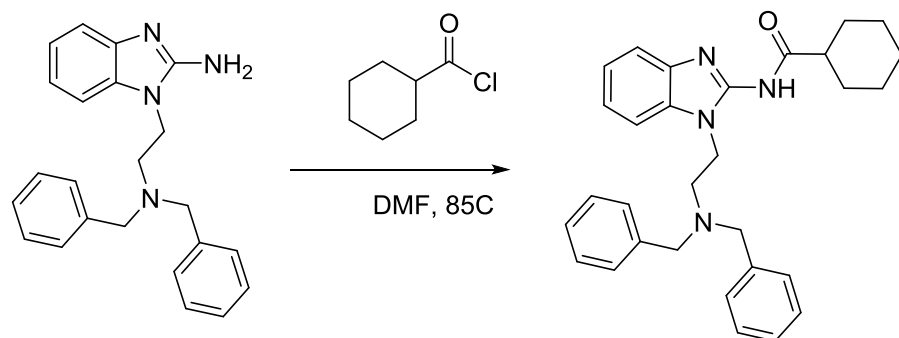


N-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)cyclohexanecarboxamide (35): A solution of 50 mg (0.39 mmol) cyclohexanecarboxylic acid, 163 mg (0.51 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 163 mL (1.2 mmol) triethylamine was stirred at rt for 15 min. Then 95 mg (0.46 mmol) **57** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 52 mg (42.1 % yield) of pure product as a red solid. ¹H NMR (400 MHz, cdcl₃) δ 7.20 (m, 38H), 4.22 (m, 18H), 3.04 (dd, *J* = 19.2, 13.1 Hz, 15H), 2.34 (s, 3H), 2.29 (m, 89H), 2.03 (s, 1H), 2.00 (m, 50H), 1.91 (dd, *J* = 22.4, 8.0 Hz, 36H), 1.70 (m,

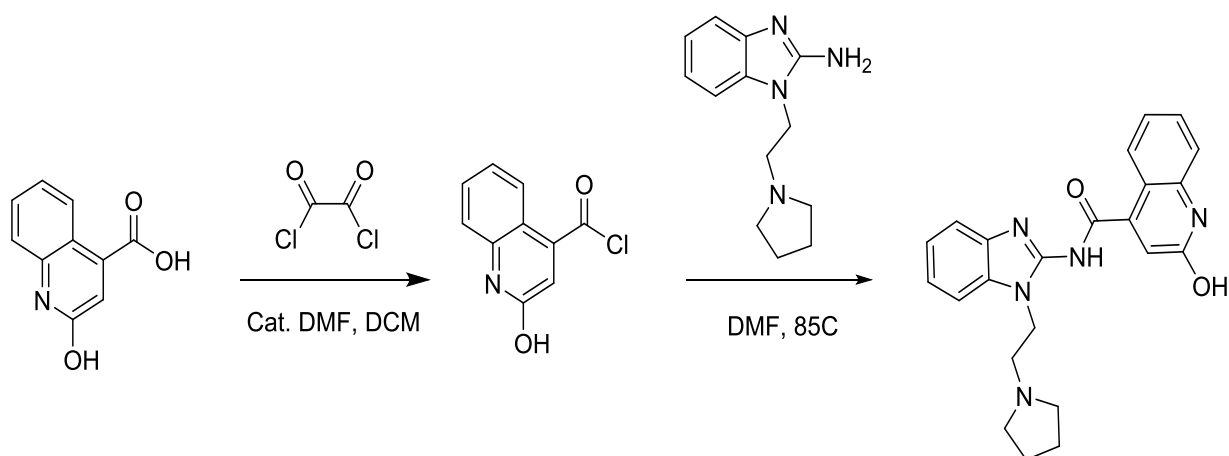
35H), 1.56 (s, 5H), 1.51 (m, 16H), 1.26 (m, 26H), 1.09 (s, 1H), 0.86 (m, 2H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.7$ min. MS (ESI⁺): 316.3 [M+H]⁺.



***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzodimidazol-2-yl)cyclohexanecarboxamide (36):** A solution of 50 mg (0.39 mmol) cyclohexanecarboxylic acid, 163 mg (0.51 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 154 mL (1.2 mmol) triethylamine was stirred at rt for 15 min. Then 107 mg (0.46 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 60 mg (44.7 % yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 7.20 (m, 38H), 4.22 (m, 18H), 3.04 (dd, $J = 19.2, 13.1$ Hz, 15H), 2.34 (s, 3H), 2.29 (m, 89H), 2.03 (s, 1H), 2.00 (m, 50H), 1.91 (dd, $J = 22.4, 8.0$ Hz, 36H), 1.70 (m, 35H), 1.56 (s, 5H), 1.51 (m, 16H), 1.26 (m, 26H), 1.09 (s, 1H), 0.86 (m, 2H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.9$ min. MS (ESI⁺): 342.2 [M+H]⁺.



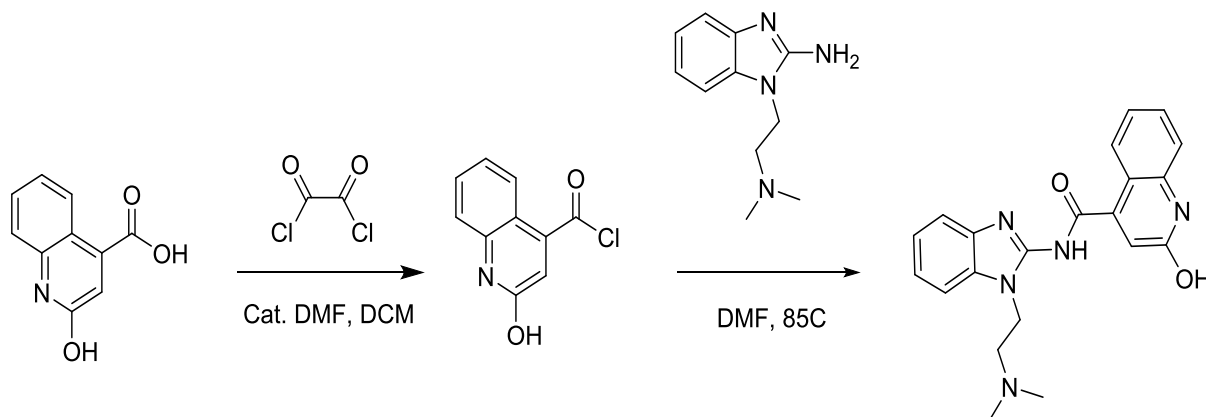
***N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)cyclohexanecarboxamide (37):** To a solution of **60** (106 mg, 0.30 mmol) in 6 mL DMF was added 0.05 mL (0.49 mmol) cyclohexanecarbonyl chloride and the reaction was stirred at rt. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 57 mg (41.0% yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 7.33 (d, *J* = 6.8 Hz, 1H), 7.26 (s, 10H), 7.24 – 7.18 (m, 2H), 7.14 – 7.07 (m, 1H), 6.86 (d, *J* = 8.0 Hz, 1H), 4.25 (t, *J* = 6.6 Hz, 2H), 3.74 (s, 4H), 2.92 (t, *J* = 6.6 Hz, 2H), 2.37 – 2.27 (m, 1H), 1.94 (d, *J* = 12.9 Hz, 2H), 1.86 – 1.76 (m, 2H), 1.71 (d, *J* = 9.1 Hz, 1H), 1.55 – 1.21 (m, 7H). LC-MS (λ = 254 nm): 99%, *t*_R = 6.3 min. MS (ESI⁺): 467.3 [M+H]⁺.



2-hydroxy-N-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)quinoline-4-

carboxamide (38): To a solution of 2-hydroxyquinoline-4-carboxylic acid (250 mg, 1.32 mmol) and a catalytic amount of DMF (2 drops) in 6.5 mL DCM added 0.85 mL (1.58 mmol) oxalyl dichloride and stirred at rt for 2 hr. The reaction was then dried to residue as the crude acyl chloride intermediate and moved directly forward.

The crude intermediate was dissolved in 2 mL DMF and 0.15 g (0.65 mmol) **56** and stirred overnight at 85°C. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 34 mg (13.0% yield) of pure product as a white solid. ¹H NMR (400 MHz, cdcl₃) δ 8.94 (dd, *J* = 8.5, 0.8 Hz, 1H), 8.08 – 8.03 (m, 1H), 8.02 (s, 1H), 7.72 (dd, *J* = 8.4, 1.4 Hz, 1H), 7.62 – 7.52 (m, 1H), 7.40 – 7.23 (m, 5H), 4.38 (t, *J* = 7.3 Hz, 2H), 2.99 – 2.90 (m, 2H), 2.72 – 2.58 (m, 4H), 1.85 – 1.73 (m, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.9 min. MS (ESI⁺): 422.2 [(M+H)+Na]⁺.

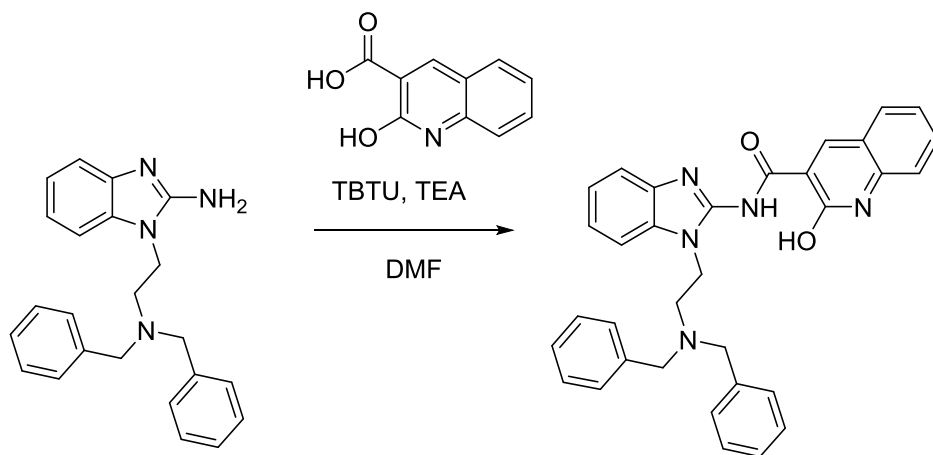


2-hydroxy-N-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)quinoline-4-

carboxamide (39): To a solution of 2-hydroxyquinoline-4-carboxylic acid (250 mg, 1.32 mmol)

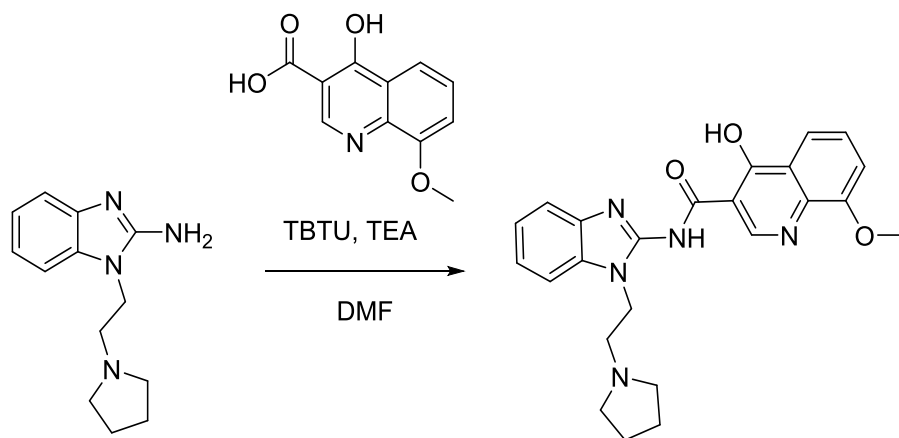
and a catalytic amount of DMF (2 drops) in 6.5 mL DCM added 0.85 mL (1.58 mmol) oxalyl dichloride and stirred at rt for 2 hr. The reaction was then dried to residue as the crude acyl chloride intermediate and moved directly forward.

The crude intermediate was dissolved in 2 mL DMF and 0.23 g (1.1 mmol) **57** and stirred overnight at 85°C. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 30 mg (7.2% yield) of pure product as a white solid. ¹H NMR (400 MHz, cdcl₃) δ 8.93 (dd, *J* = 8.6, 1.0 Hz, 1H), 8.08 – 7.98 (m, 2H), 7.71 (ddd, *J* = 8.4, 6.9, 1.4 Hz, 1H), 7.55 (ddd, *J* = 8.3, 6.9, 1.3 Hz, 1H), 7.34 – 7.24 (m, 3H), 7.23 – 7.17 (m, 1H), 4.32 (t, *J* = 7.0 Hz, 2H), 2.75 (t, *J* = 7.0 Hz, 2H), 2.33 (s, 6H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.4 min. MS (ESI⁺): 376.2 [M+H]⁺.



***N*-(1-(2-(diphenylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-2-hydroxyquinoline-3-carboxamide (40):** A solution of 50 mg (0.26 mmol) 2-hydroxyquinoline-3-carboxylic acid, 110 mg (0.34 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 0.11 mL (0.79 mmol) triethylamine was stirred at rt for 15 min. Then 113 mg (0.32 mmol)

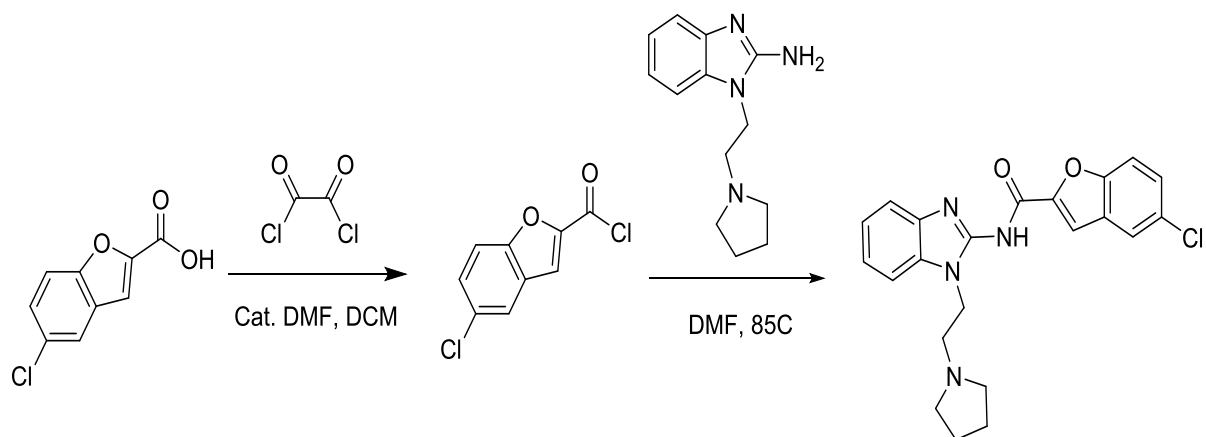
60 was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 82 mg (58.9% yield) of pure product as a yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 8.41 – 8.34 (m, 1H), 7.52 – 7.39 (m, 3H), 7.30 – 6.97 (m, 14H), 6.94 (d, *J* = 8.0 Hz, 1H), 4.25 (t, *J* = 6.2 Hz, 2H), 3.66 (s, 4H), 2.93 (t, *J* = 6.2 Hz, 2H). LC-MS (λ = 254 nm): 99%, *t*_R = 6.1 min. MS (ESI⁺): 529.3 [M+H]⁺.



4-hydroxy-8-methoxy-N-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)quinoline-

3-carboxamide (41): A solution of 75 mg (0.34 mmol) 4-hydroxy-8-methoxyquinoline-3-carboxylic acid, 143 mg (0.44 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 143 mL (1.03 mmol) triethylamine was stirred at rt for 15 min. Then 94 mg (0.41 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 89 mg (60.5% yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cdcl₃) δ 9.57 (s, 1H),

8.58 – 8.51 (m, 1H), 8.25 (dd, $J = 8.3, 1.2$ Hz, 1H), 7.42 (ddd, $J = 12.7, 5.2, 2.7$ Hz, 2H), 7.37 – 7.27 (m, 2H), 6.96 (dd, $J = 7.9, 1.0$ Hz, 1H), 4.45 (t, $J = 7.0$ Hz, 2H), 3.98 (s, 3H), 2.97 (t, $J = 7.0$ Hz, 2H), 2.74 – 2.60 (m, 4H), 1.81 – 1.68 (m, 4H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.4$ min. MS (ESI⁺): 415.2 [M+H]⁺.

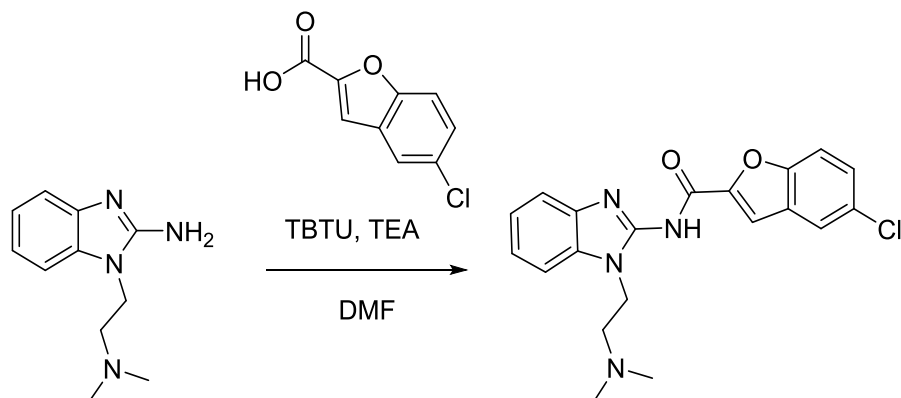


5-chloro-N-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)benzofuran-2-

carboxamide (42): To a solution of 5-chlorobenzofuran-2-carboxylic acid (250 mg, 1.27 mmol) and a catalytic amount of DMF (2 drops) in 6.5 mL DCM added 0.13 mL (1.53 mmol) oxalyl dichloride and stirred at rt for 2 hr. The reaction was then dried to residue as the crude acyl chloride intermediate and moved directly forward.

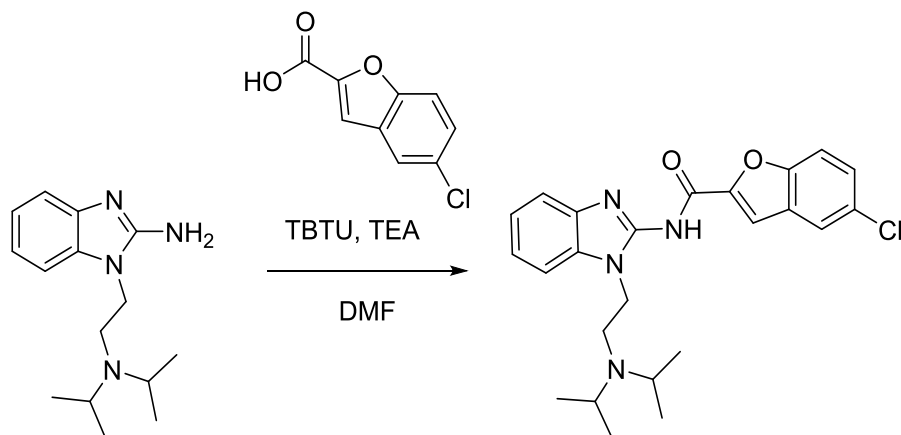
The crude intermediate was dissolved in 2 mL DMF and 0.15 g (0.65 mmol) **56** and stirred overnight at 85°C. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 40 mg (16.5% yield) of pure product as a white solid. ¹H NMR (400 MHz, cdcl₃) δ 7.66 – 7.64 (m, 1H), 7.53 (d, $J = 8.8$ Hz, 1H), 7.50 (d, $J = 0.9$ Hz, 1H),

7.41 – 7.27 (m, 5H), 4.46 – 4.37 (m, 2H), 3.01 – 2.90 (m, 2H), 2.74 – 2.65 (m, 4H), 1.87 – 1.77 (m, 4H), 1.57 (s, 5H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.4$ min. MS (ESI⁺): 410.2 [M+H]⁺.



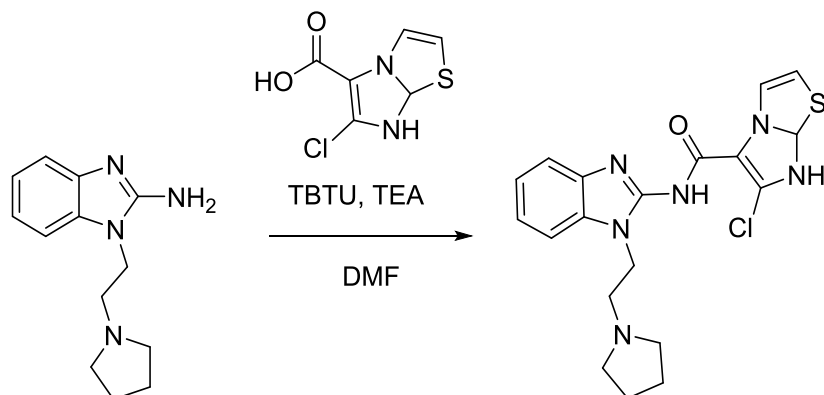
5-chloro-*N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzofuran-2-

carboxamide (43): A solution of 100 mg (0.49 mmol) 5-chlorobenzofuran-2-carboxylic acid, 204 mg (0.63 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 204 mL (1.47 mmol) triethylamine was stirred at rt for 15 min. Then 115 mg (0.58 mmol) **57** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 100 mg (53.4% yield) of pure product as a white solid. ¹H NMR (400 MHz, MeoD) δ 12.40 (s, 0H), 7.57 (dd, $J = 2.2, 0.4$ Hz, 1H), 7.47 – 7.42 (m, 1H), 7.30 – 7.23 (m, 3H), 7.15 (dd, $J = 8.8, 2.3$ Hz, 1H), 6.96 (pd, $J = 7.4, 1.4$ Hz, 2H), 4.08 (t, $J = 6.4$ Hz, 2H), 2.41 (t, $J = 6.0$ Hz, 2H), 1.94 (s, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.3$ min. MS (ESI⁺): 384.2 [M+H]⁺.



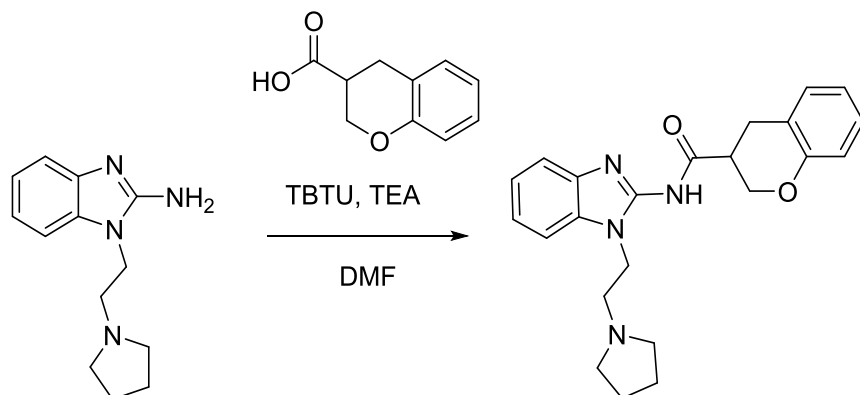
5-chloro-*N*-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzofuran-2-

carboxamide (44): A solution of 49 mg (0.25 mmol) 5-chlorobenzofuran-2-carboxylic acid, 104 mg (0.32 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 104 mL (0.76 mmol) triethylamine was stirred at rt for 15 min. Then 78 mg (0.3 mmol) **58** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 33 mg (30.2 % yield) of pure product as a white solid. ¹H NMR (400 MHz, cdcl₃) δ 7.64 – 7.61 (m, 1H), 7.51 – 7.46 (m, 1H), 7.44 (d, *J* = 0.9 Hz, 1H), 7.39 (dt, *J* = 7.3, 1.1 Hz, 1H), 7.32 (d, *J* = 2.2 Hz, 1H), 7.31 – 7.21 (m, 4H), 4.18 (dd, *J* = 8.3, 6.8 Hz, 2H), 3.15 – 3.02 (m, 2H), 2.88 – 2.79 (m, 2H), 1.03 (d, *J* = 6.5 Hz, 12H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.5 min. MS (ESI⁺): 441.2 [M+H]⁺.



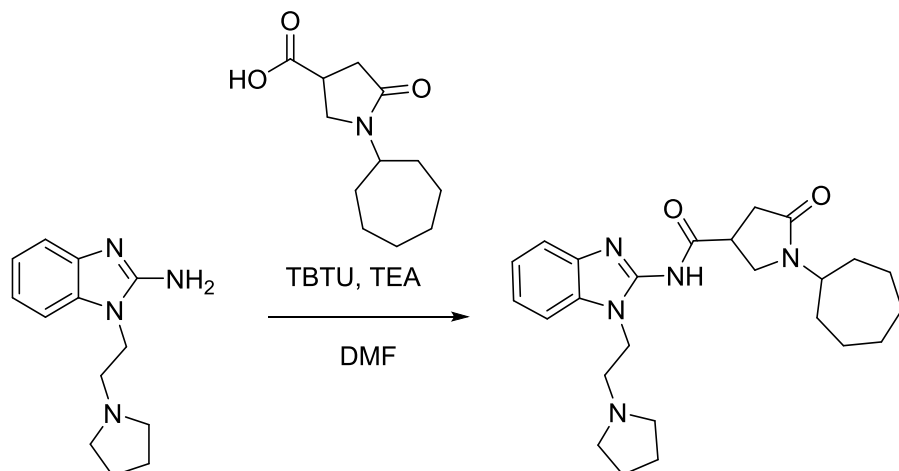
6-chloro-N-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)-7,7a-

dihydroimidazo[2,1-b]thiazole-5-carboxamide (45): A solution of 73 mg (0.36 mmol) 6-chloro-7,7a-dihydroimidazo[2,1-b]thiazole-5-carboxylic acid, 150 mg (0.47 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 151 mL (1.08 mmol) triethylamine was stirred at rt for 15 min. Then 100 mg (0.43 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 118 mg (78.7% yield) of pure product as a yellow solid. ¹H NMR (400 MHz, dmso) δ 12.73 (s, 1H), 8.55 (d, *J* = 4.4 Hz, 1H), 7.56 – 7.49 (m, 3H), 7.31 – 7.19 (m, 2H), 4.37 (t, *J* = 6.8 Hz, 2H), 2.86 (t, *J* = 6.4 Hz, 2H), 2.54 (s, 4H), 1.64 (t, *J* = 3.3 Hz, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.8 min. MS (ESI⁺): 417.3 [M+H]⁺.

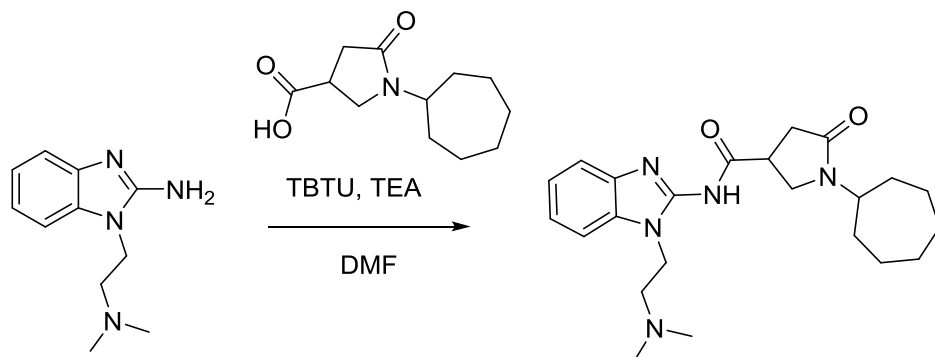


***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)chromane-3-carboxamide (46):** A

solution of 70 mg (0.39 mmol) chromane-3-carboxylic acid, 165 mg (0.51 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 164 mL (1.18 mmol) triethylamine was stirred at rt for 15 min. Then 108 mg (0.47 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 118 mg (77.1% yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, cd₃od) δ 7.50 – 7.42 (m, 2H), 7.33 – 7.22 (m, 2H), 7.12 – 7.01 (m, 2H), 6.83 (td, *J* = 7.4, 1.2 Hz, 1H), 6.75 (dd, *J* = 8.2, 1.0 Hz, 1H), 4.54 – 4.48 (m, 1H), 4.33 (t, *J* = 6.9 Hz, 2H), 4.20 (dd, *J* = 10.6, 9.0 Hz, 1H), 3.25 – 3.13 (m, 1H), 3.03 (dd, *J* = 10.4, 7.4 Hz, 1H), 2.91 (t, *J* = 7.0 Hz, 2H), 2.81 (s, 0H), 2.74 – 2.61 (m, 4H), 1.88 – 1.75 (m, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.1 min. MS (ESI⁺): 391.2 [M+H]⁺. UNC2373A

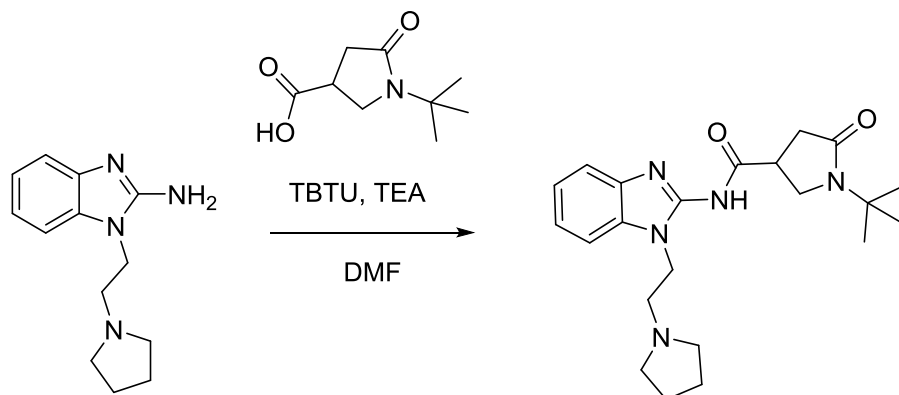


1-cycloheptyl-5-oxo-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzotriazol-2-yl)pyrrolidine-3-carboxamide (47): A solution of 70 mg (0.31 mmol) 1-cycloheptyl-5-oxopyrrolidine-3-carboxylic acid, 129 mg (0.40 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 129 mL (0.93 mmol) triethylamine was stirred at rt for 15 min. Then 86 mg (0.37 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 130 mg (92.8% yield) of pure product as a pale orange solid. ¹H NMR (400 MHz, dmso) δ 12.47 (s, 1H), 7.51 – 7.39 (m, 2H), 7.24 – 7.12 (m, 2H), 4.20 (dd, *J* = 6.5, 4.5 Hz, 2H), 3.97 – 3.85 (m, 1H), 3.64 – 3.50 (m, 2H), 3.14 (d, *J* = 5.0 Hz, 1H), 2.77 (d, *J* = 5.8 Hz, 2H), 2.52 (d, *J* = 10.5 Hz, 4H), 1.72 – 1.32 (m, 17H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.1 min. MS (ESI⁺): 439.3 [M+H]⁺.



1-cycloheptyl-N-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-5-oxopyrrolidine-

3-carboxamide (48): A solution of 75 mg (0.33 mmol) 1-cycloheptyl-5-oxopyrrolidine-3-carboxylic acid, 139 mg (0.43 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 139 mL (1.0 mmol) triethylamine was stirred at rt for 15 min. Then 81 mg (0.39 mmol) **57** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 66 mg (45.5% yield) of pure product as a red oil. ¹H NMR (400 MHz, cdcl₃) δ 7.31 – 7.17 (m, 4H), 4.21 (t, *J* = 6.9 Hz, 2H), 4.15 (d, *J* = 4.4 Hz, 1H), 3.70 (dd, *J* = 9.7, 6.3 Hz, 1H), 3.65 – 3.55 (m, 1H), 3.25 (d, *J* = 5.9 Hz, 1H), 3.01 (s, 1H), 2.86 (dd, *J* = 16.9, 7.2 Hz, 1H), 2.67 (td, *J* = 7.0, 3.3 Hz, 3H), 2.31 (s, 6H), 1.82 – 1.73 (m, 2H), 1.73 – 1.55 (m, 6H), 1.55 – 1.43 (m, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.1 min. MS (ESI⁺): 413.3 [M+H]⁺.



1-(tert-butyl)-5-oxo-N-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzodimidazol-2-yl)pyrrolidine-

3-carboxamide (49): A solution of 100 mg (0.54 mmol) 1-(tert-butyl)-5-oxopyrrolidine-3-

carboxylic acid, 225 mg (0.70 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium

tetrafluoroborate (TBTU), and 225 mL (1.6 mmol) triethylamine was stirred at rt for 15 min.

Then 149 mg (0.65 mmol) **56** was added and the mixture was stirred overnight. The reaction was

quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the

organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The

product was purified by normal phase Teledyne Isco automated column system to afford 78 mg

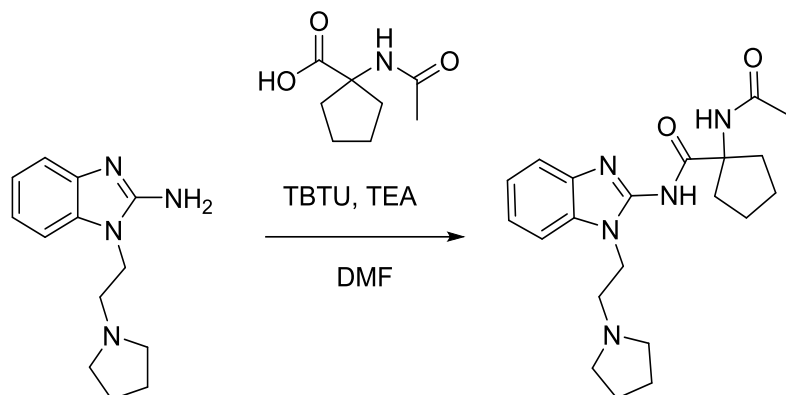
(30.3% yield) of pure product as a red oil. ¹H NMR (400 MHz, cdcl₃) δ 7.31 – 7.17 (m, 4H), 4.21

(t, *J* = 6.9 Hz, 2H), 4.15 (d, *J* = 4.4 Hz, 1H), 3.70 (dd, *J* = 9.7, 6.3 Hz, 1H), 3.65 – 3.55 (m, 1H),

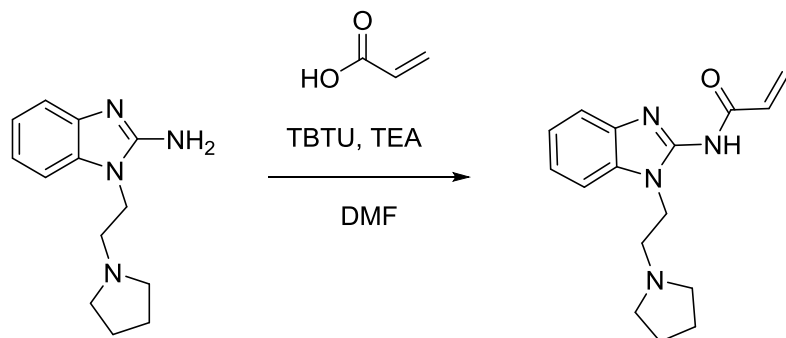
3.25 (d, *J* = 5.9 Hz, 1H), 3.01 (s, 1H), 2.86 (dd, *J* = 16.9, 7.2 Hz, 1H), 2.67 (td, *J* = 7.0, 3.3 Hz,

3H), 2.31 (s, 6H), 1.82 – 1.73 (m, 2H), 1.73 – 1.55 (m, 6H), 1.55 – 1.43 (m, 4H). LC-MS (λ =

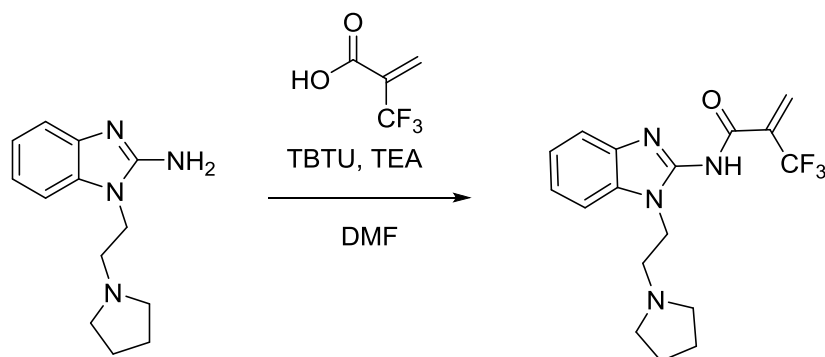
254 nm): 99%, *t*_R = 3.6 min. MS (ESI⁺): 399.3 [M+H]⁺.



1-acetamido-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)cyclopentane-1-carboxamide (50): A solution of 71 mg (0.41 mmol) 1-acetamidocyclopentane-1-carboxylic acid, 171 mg (0.53 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 171 mL (1.23 mmol) triethylamine was stirred at rt for 15 min. Then 113 mg (0.49 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 16 mg (10.2% yield) of pure product as a white solid. ¹H NMR (400 MHz, cd₃od) δ 7.45 (dd, *J* = 10.8, 8.1 Hz, 2H), 7.34 – 7.19 (m, 2H), 4.31 (t, *J* = 7.2 Hz, 2H), 3.01 – 2.85 (m, 3H), 2.74 – 2.63 (m, 4H), 2.43 – 2.31 (m, 2H), 2.07 – 2.00 (m, 2H), 1.99 (s, 3H), 1.90 – 1.73 (m, 9H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.3 min. MS (ESI⁺): 385.2 [M+H]⁺.



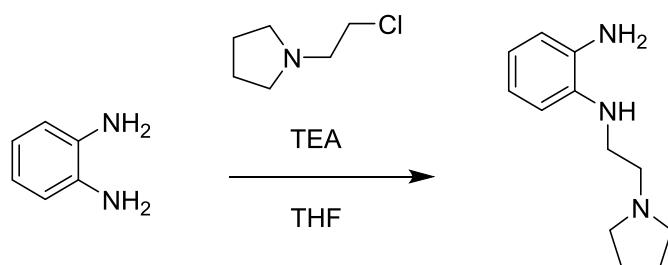
***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)acrylamide (51):** A solution of 75 mg (0.87 mmol) acrylic acid, 364 mg (1.13 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 364 mL (2.61 mmol) triethylamine was stirred at rt for 15 min. Then 241 mg (1.05 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 51 mg (20.5% yield) of pure product as a white solid. ¹H NMR (400 MHz, cd₃od) δ 7.42 (ddd, *J* = 16.1, 7.2, 1.5 Hz, 2H), 7.34 – 7.17 (m, 2H), 6.18 (dd, *J* = 2.2, 0.9 Hz, 1H), 5.50 – 5.42 (m, 1H), 4.89 (s, 2H), 4.34 (t, *J* = 7.0 Hz, 2H), 2.95 (t, *J* = 7.0 Hz, 2H), 2.79 – 2.64 (m, 4H), 1.88 – 1.75 (m, 4H). LC-MS (λ = 254 nm): 99%, *t_R* = 3.2 min. MS (ESI⁺): 300.2 [M+H]⁺.



***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)-2-(trifluoromethyl)acrylamide**

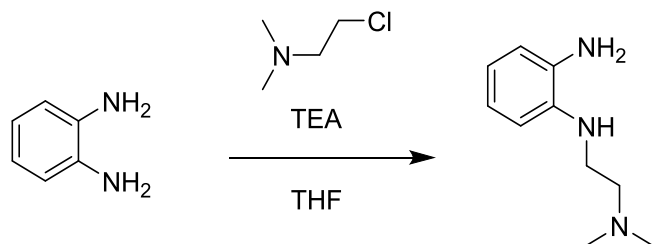
(52): A solution of 50 mg (0.36 mmol) 2-(trifluoromethyl)acrylic acid, 149 mg (0.46 mmol) *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU), and 150 mL (1.07 mmol) triethylamine was stirred at rt for 15 min. Then 98 mg (0.43 mmol) **56** was added and the mixture was stirred overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 56 mg (44.4% yield) of pure product as a yellow oil. ¹H NMR (400 MHz, cdcl₃) δ 7.29 – 7.08 (m, 4H), 4.33 – 4.23 (m, 2H), 4.20 (dd, *J* = 14.3, 7.4 Hz, 2H), 3.54 – 3.41 (m, 1H), 2.88 (td, *J* = 6.7, 3.5 Hz, 2H), 2.59 (d, *J* = 5.5 Hz, 4H), 1.76 – 1.64 (m, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 2.8 min. MS (ESI+): 354.2 [M+H]⁺.

Intermediates:

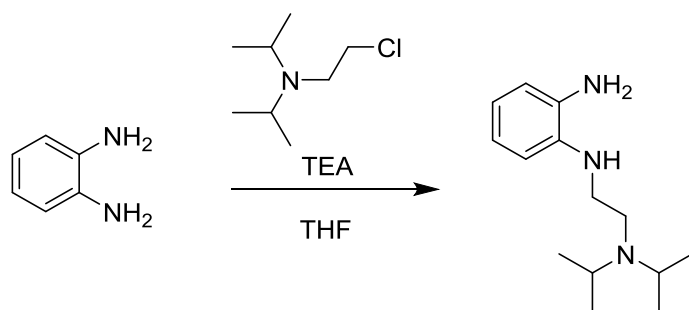


N¹-(2-(pyrrolidin-1-yl)ethyl)benzene-1,2-diamine (53): To a solution of benzene-1,2-diamine (2 g, 18.49 mmol) in 6 mL THF was added 0.37 g (2.78 mmol) 1-(2-chloroethyl)pyrrolidine and 1.2 mL (9.25 mmol) triethylamine and the solution was stirred overnight at rt. The crude reaction was dried to residue, dissolved in DCM and dried onto silica and purified by normal phase Teledyne Isco automated column system to afford 103 mg (21.7% yield) of pure product

as a dark yellow oil. This intermediate was found to not be stable after 1 day and quickly oxidize to a black residue; needs to be used the same day after purification.

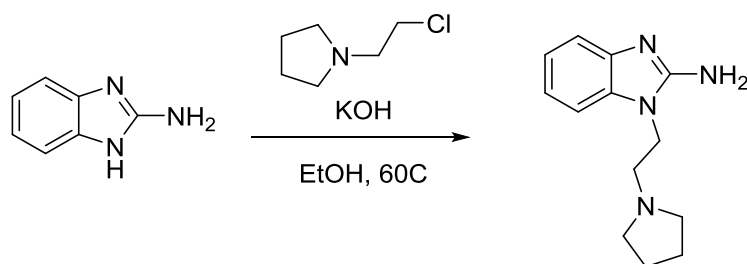


N¹-(2-(dimethylamino)ethyl)benzene-1,2-diamine (54): To a solution of benzene-1,2-diamine (0.5 g, 4.63 mmol) in 6 mL THF was added 0.59 g (5.54 mmol) 2-chloro-*N,N*-dimethylethan-1-amine and 2.6 mL (18.48 mmol) triethylamine and the solution was stirred overnight at rt. The crude reaction was dried to residue, dissolved in DCM and dried onto silica and purified by normal phase Teledyne Isco automated column system to afford 100 mg (12.1% yield) of pure product as a dark yellow oil. This intermediate was found to not be stable after 1 day and quickly oxidize to a black residue; needs to be used the same day after purification.

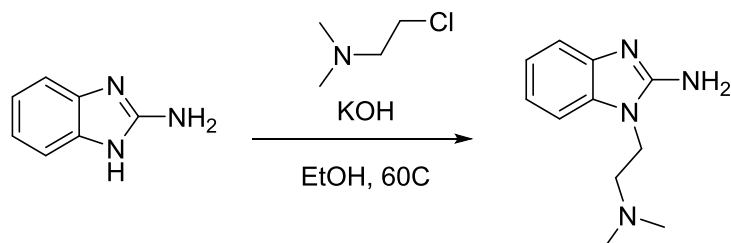


N¹-(2-(diisopropylamino)ethyl)benzene-1,2-diamine (55): To a solution of benzene-1,2-diamine (0.5 g, 4.63 mmol) in 6 mL THF was added 0.91 g (5.56 mmol) *N*-(2-chloroethyl)-*N*-isopropylpropan-2-amine and 2.6 mL (18.48 mmol) triethylamine and the solution was stirred overnight at rt. The crude reaction was dried to residue, dissolved in DCM and dried onto silica and purified by normal phase Teledyne Isco automated column system to afford 702 mg (64.3%

yield) of pure product as a dark yellow oil. This intermediate was found to not be stable after 1 day and quickly oxidize to a black residue; needs to be used the same day after purification.

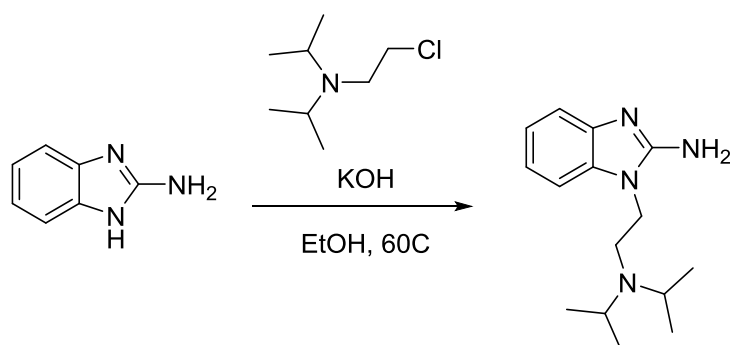


1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-amine (56): To a solution of 1H-benzo[d]imidazol-2-amine (2g, 14.1 mmol) and 1.2g (21.2 mmol) potassium hydroxide (KOH) in 6 mL EtOH was added 2.8 g (20.9 mmol) 1-(2-chloroethyl)pyrrolidine and the reaction was stirred at 60°C overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 1.5 g (46.1% yield) of pure product as a pale yellow solid. ¹H NMR (400 MHz, dmso) δ 7.14 – 7.06 (m, 2H), 6.95 – 6.85 (m, 4H), 6.39 (s, 2H), 4.06 (t, *J* = 7.9, 2H), 2.69 (t, *J* = 7.4, 2H), 2.54 – 2.47 (m, 4H), 1.68 – 1.62 (m, 4H). LC-MS (λ = 254 nm): 99%, *t*_R = 1.3 min. MS (ESI⁺): 231.2 [M+H]⁺. MTP00047-154-2

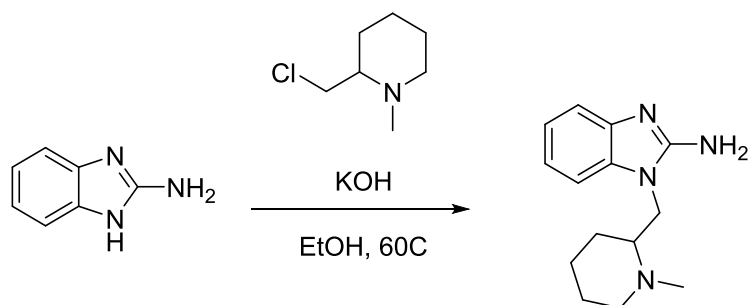


1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-amine (57): To a solution of 1H-benzo[d]imidazol-2-amine (2g, 14.1 mmol) and 1.2g (21.2 mmol) potassium hydroxide (KOH) in 6 mL EtOH was added 2.8 g (21.2 mmol) 2-chloro-N,N-dimethylethan-1-amine and the

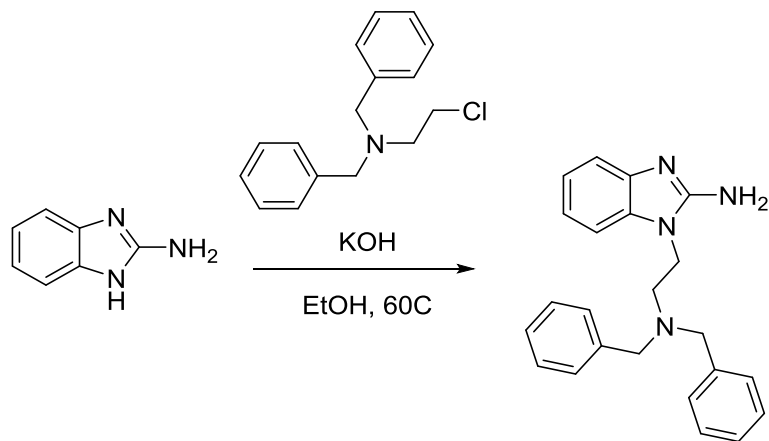
reaction was stirred at 60°C overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 1.15 g (39.9% yield) of pure product as a pale pink solid. ¹H NMR (400 MHz, cdcl₃) δ 7.45 – 7.38 (m, 1H), 7.10 (ddd, *J* = 7.8, 6.4, 2.2 Hz, 1H), 7.03 (ddd, *J* = 4.9, 2.8, 1.0 Hz, 2H), 6.00 (s, 2H), 4.04 – 3.95 (m, 2H), 2.72 – 2.66 (m, 2H), 2.32 (s, 6H). LC-MS (λ = 254 nm): 99%, *t*_R = 0.6 min. MS (ESI⁺): 205.2 [M+H]⁺.



1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-amine (58): To a solution of 1H-benzo[d]imidazol-2-amine (0.5g, 3.5 mmol) and 0.3 g (5.2 mmol) potassium hydroxide (KOH) in 6 mL EtOH was added 1.1 g (5.3 mmol) N-(2-chloroethyl)-N-isopropylpropan-2-amine and the reaction was stirred at 60°C overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 0.21 g (22.9% yield) of pure product as a pale pink solid. No proton NMR was obtained for this intermediate. LC-MS (λ = 254 nm): 99%, *t*_R = 2.9 min. MS (ESI⁺): 236.2 [M+H]⁺.



1-((1-methylpiperidin-2-yl)methyl)-1H-benzo[d]imidazol-2-amine (59): To a solution of 1H-benzo[d]imidazol-2-amine (350 mg, 2.47 mmol) and 209 mg (3.7 mmol) potassium hydroxide (KOH) in 6 mL EtOH was added 682 mg (3.7 mmol) 2-(chloromethyl)-1-methylpiperidine and the reaction was stirred at 60°C overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 208 mg (34.4% yield) of pure product as a pale yellow solid. *a clean proton NMR was not obtained for this intermediate.



1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-amine (60): To a solution of 1H-benzo[d]imidazol-2-amine (501 mg, 3.54 mmol) and 297 mg (5.29 mmol) potassium hydroxide (KOH) in 6 mL EtOH was added 1.34 g (5.15 mmol) N,N-dibenzyl-2-chloroethan-1-amine and

the reaction was stirred at 60°C overnight. The reaction was quenched using sat. aqueous sodium bicarbonate and extracted 3x with DCM (10 mL), the organic phase was combined, dried over Na₂SO₄, filtered, concentrated and dried onto silica. The product was purified by normal phase Teledyne Isco automated column system to afford 833 mg (66.1% yield) of pure product as a yellow solid. ¹H NMR (400 MHz, dmsO) δ 7.27 – 7.15 (m, 10H), 7.11 (d, *J* = 7.4, 1H), 6.92 – 6.85 (m, 2H), 6.76 – 6.69 (m, 2H), 6.29 (s, 2H), 4.1 (t, *J* = 7.6, 2H).

AlphaScreen assay method

The Alphascreen assay was performed as described elsewhere.[103, 104] In brief, compound plates (1 µl at 30 mM highest concentration) were diluted in 1X assay buffer (20 mM TRIS pH 8.0, 25 mM NaCl, 2 mM DTT and 0.05% Tween-20) over 2 steps using a Multimek robotic pipettor (Nanoscreen) and 1 µl was spotted into the wells of 384-well assay Proxiplates (PerkinElmer). To these plates 9 µl of protein- peptide mix in 1X assay buffer was added by Multidrop (Thermo) and incubated for 30 min at room temperature. At this point 2 uL of streptavidine- conjugate donor and nickel-chelate acceptor beads (45 µg/mL in 1X assay buffer) were added, the plates were allowed to incubate for an additional 30 min in the dark at room temperature. After incubation the plates were read on EnVision mulilabel reader equipped with HTS alpha screen laser (Perkin Elmer). The screens reported are performed up to 30 µM, and therefore it should be noted that those compounds referred to as inactive are indeed inactive only within the concentration range tested

Isothermal titration calorimetry binding experiments

All ITC measurements were recorded at 25°C with an AutoITC₂₀₀ microcalorimeter (MicroCal Inc.). All protein and compound stock samples were prepared in the target buffer (25

mM Tris-HCl, pH 8, 150 mM NaCl, and 2 mM β -mercaptoethanol), and then diluted in the same buffer to achieve the desired concentrations: 90 μ M protein and 1 mM compound depending on the expected dissociation constant. The concentration of protein stock solutions were established using the Edelhoch method, whereas 10 mM compound stock solutions were prepared gravimetrically based on molecular weight. A typical experiment included a single 0.2 μ L compound injection into a 200 μ L cell filled with protein, followed by 25 subsequent 1.5 μ L injections of compound. Injections were performed with a spacing of 180 seconds and a reference power of 8 μ cal/sec. The titration data was analyzed using Origin Software (MicroCal Inc.) by non-linear least squares, fitting the heats of binding as a function of the compound:protein ratio. The data were fit based on a one set of sites model. ITC experiments of remaining peptides are shown below.

Future work

Currently there are no current plans to continue synthetic efforts on the UNC1554 benzimidazole scaffold. The initial NMR protein binding studies (Figure 2.3 A & B) and NMR titration data (Figure 2.3 C) show evidence that UNC1554 interacts with the 53BP1 TTD and the N-ethyl pyrrolidine functional group mimics the endogenous H4K20me2 residue. Additional NMR titration data show that this is modest affinity for UNC1554 against 53BP1 TTD. Unfortunately, ITC experiments failed to show binding interactions. These contradictory pieces of data may potentially be explained by the thermodynamics of binding of UNC1554 in the ITC experiments where the measured ΔH and ΔS of binding may cancel each other out and provide no detectable binding measurement. Additionally, the lack of tractable SAR lead to the overall decision to no longer pursue future medicinal chemistry efforts on UNC1554. It should be noted

that all future medicinal chemistry efforts always included ITC measurements against any potent small molecule hits that were discovered to validate true binding interactions prior to initiation of chemistry efforts.

It is worth mentioning that one of the compounds that was synthesized during the SAR efforts became a hit for a different screening project within the Center for Integrative Chemical Biology and Drug Discovery (CICBDD). Compound **42** was shown in unpublished work to be capable of facilitating release of endosome-trapped siRNA oligomers. This molecule and others have been licensed by an UNC start-up firm for future therapeutic discovery efforts.

Benzamide -based small molecule inhibitors of 53BP1, UNC2170

Results and discussion

Discovery of UNC2170 hit

The Frye research group efforts utilize a comprehensive structure-based design and cross-screening approach in order to evaluate all synthesized ligands against a panel of Kme reader domains. The Kme reader panel consists of 10 reader proteins from four different families: Tudor domains (53BP1, UHRF1, PHF1, PHF19), chromodomains (CBX7), MBT domains (L3MBTL1, L3MBTL3, MBTD1), and PHD fingers (JARID1A, PHF23, UHRF1). These proteins were chosen largely based on available structural information, reader family representation, and biological relevance. Previously, the Frye research group reported an HTS optimized format for the AlphaScreen bead-based proximity assay (Perkin-Elmer) for Kme readers [105] and this screening tool was employed to initially assess 53BP1 binding. **61** (UNC2170), Table 1, emerged as a preliminary 53BP1 hit from these cross-screening efforts. Because of the modest affinity of **61** (UNC2170) ($29 \pm 7.4 \mu\text{M}$), and its fragment-like nature (MW = 313.24, ligand efficiency = 0.35, lipophilic ligand efficiency = 1.5)[106, 107], **61** (UNC2170) was profiled at concentrations up to 500 μM versus the Kme reader reader panel to better ascertain its selectivity, as this was of paramount importance for chemical probe development[56]. This approach revealed promising levels of selectivity within this small but diverse sampling of Kme reader proteins, as **61** (UNC2170) showed no affinity within the concentrations tested against any other members of the panel, including other tudor domain-containing proteins and other readers of methylated H4K20. Additionally, isothermal titration calorimetry (ITC) experiments confirmed ligand binding, revealing a K_d of $22 \pm 2.5 \mu\text{M}$ for

UNC2170 against the 53BP1 TTD, Figure 2.8. Overall, through a cross-screening approach **61** was identified (UNC2170) as a micromolar ligand of 53BP1, which demonstrates at least 17-fold selectivity for 53BP1 as compared to other methyl-lysine (Kme) binding proteins tested.

Comparison to endogenous H4K20me2 chromatin peptide tail

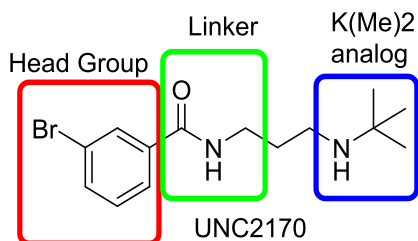
A structure-based design approach was initiated by analyzing available crystal structure data of 53BP1 bound to an H4K20me2 peptide (PDB 2IG0) [66]. The key interactions noted in this structure were a hydrogen-bond between the Kme basic amine and an aspartic acid (Asp1521), cation- π interactions between Kme2 and phenylalanine, tyrosine, and tryptophan residues within the aromatic binding cage (Tyr1502, Tyr1523, Phe1519, Trp1495), and a cation- π interaction between Arg19 on the H4 peptide tail and a tryptophan residue (Trp1500) [66]. Mutagenesis of the histone peptide had previously demonstrated that H18 also contributes to binding, but this interaction was not clearly defined in the X-ray crystal structure [66]. To develop a more complete understanding of the binding interactions of the H4 peptide (amino acids 14-27) bound to 53BP1, isotope enriched (^{13}C and ^{15}N) NMR spectroscopy [101] was applied to determine the structure of the central residues of the H4K20me2 peptide bound to 53BP1 (PDB 2LVM) [102]. It was observed that this central region of the H4 peptide corresponding to residues 15 to 22 adopts a “U-turn” conformation. Notably, in addition to the binding interactions with H4K20me2 and H4R19 that were detected in the crystal structure, further analysis revealed a pocket containing both acidic and hydrophobic residues that accommodates H4R17 and H4V21. It has also been shown that acetylation of H4K16 diminishes 53BP1 binding by disrupting a salt bridge between H4K16 and Glu1551 [102]. Therefore, a small molecule that could occupy the methyl-lysine binding cage of 53BP1 and interact

favorably with some of the surrounding residues would be expected to block 53BP1 binding to H4K20me2.

Structure-activity relationship studies

Structure activity studies began by analysis of the structure of **61** (UNC2170). There were three focus areas that were identified for synthetic modification, Figure 2.6. The Kme2 head group, the linker region and the aromatic head group.

Figure 2.6: Region of synthetic modification on **61** (UNC2170)



Initial synthetic efforts were conducted without knowledge of the defined binding pose of the ligand upon the 53BP1 TTD binding domain. It was hypothesized that the Kme2 head group would dock within the aromatic cage, but it was uncertain where the 3-bromo aromatic head group would bind. Use of structure based design efforts through study of the x-ray co-crystal structure of 53BP1 TTD bound to the endogenous H4K20me2 peptide (PDB 2LVM) was used as a primary source of inspiration for synthetic modifications. Synthetic efforts are outlined below.

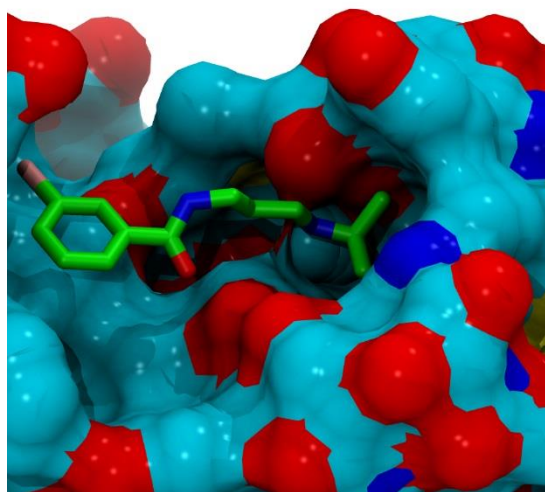
Synthetic strategy for SAR studies

Modification of amine head group

It was hypothesized that the basic amine head group of UNC2170 would bind within the Kme binding pocket of 53BP1, similar to the endogenous dimethyl-lysine substrate. *In silico*

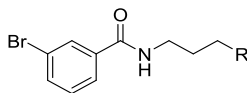
docking models conducted by Dr. Bradley Dickson (CICBDD) showed the N-*tert*-butyl amine head group bound within the aromatic head group in an analogous fashion as H4K20me2 and that the 3-bromo aromatic head group was docked within the acidic/hydrophobic binding pocket, Figure 2.7.

Figure 2.7: *In silico* docking model of **61** (UNC2170) bound to 53BP1 TTD



This docking pose differed from that proposed binding mode of **2** (UNC1554) in that the aromatic head group was docked in the acidic/hydrophobic binding pocket rather than the R19 binding pocket. This computer-based docking analysis provided an initial basis to begin a SAR study of how modifications in the amine head group affect the binding potency and selectivity of **61** (UNC2170) analogs within the aromatic cage of 53BP1.

It was hypothesized that the aromatic binding cage would possess a defined degree of steric tolerance for amine head groups as shown through work with **2** (UNC1554). Findings from that SAR study were used to design new analogues as shown in Table 2.4.

Table 2.4. Selectivity Data For **61** (UNC2170) and SAR Studies of the Terminal Amine Group.^{a,b}

ID	R	53BP1	CBX 7	JARID1 A	PHF1	PHF1 9	PHF2 3	UHRF 1	L3MBT L1	L3MBT L3	MBTD1
61^c (UNC2170)		29 ± 7.4	> 500	> 500	> 500	> 500	> 500	> 500	> 500	> 500	> 500
62		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
63		> 100	> 100	> 100	82 ± 13	61 ± 0.9	> 100	> 100	> 100	> 100	> 100
64 (UNC2892)		> 500	> 500	> 500	> 500	> 500	> 500	> 500	> 500	> 500	> 500
65		> 100	93 (1) >100 (3)	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
66		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
67		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
68		> 100	> 100	> 100	> 100	> 100	> 100	> 100	37 ± 12	> 100	78 ± 16

^aIC₅₀ values are the average of at least 3 values ± the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 100 μM and compounds showing less than 50% inhibition at this concentration are labeled >100 μM. ^cCompound **61** and **64** were tested at higher concentrations (500 μM) against the panel to further evaluate selectivity.

It was observed that variation of the size of the amine head group and also the substitution state of the amine (secondary vs. tertiary) had an effect on the affinity for 53BP1 TTD. Exploration of the steric tolerance of the basic amine began by replacement of the secondary *N-tert*-butyl amine group with dimethyl amine, **62**, to more closely mimic the endogenous H4K20me₂ substrate. Interestingly, this compound showed no binding activity by AlphaScreen and this finding was confirmed by ITC, Figure 2.8. This result led to the hypothesis that the 53BP1 Kme binding pocket may have a preference for secondary amines over tertiary

amines in the context of the **61** (UNC2170) framework, in contrast to its preference for dimethylated versus monomethylated lysine in histone peptides [66]. Preparation of the isopropyl amino compound, **63**, showed no activity against 53BP1 TTD but interestingly, it was weakly active against some of the other tudor domain containing proteins in the panel, Table 2.4. Incorporation of other tertiary amines as in compounds **65** - **66**, where the secondary amine of UNC2170 is methylated or ethylated, also showed no appreciable binding activity up to 500 μ M as measured by AlphaScreen, Table 2.4. ITC studies further confirmed that **64** has no measurable affinity for 53BP1 TTD, and therefore compound **64** (UNC2892) was selected as a negative control compound due to its inactivity and structural similarity to **61** (UNC2170). These analogs were prepared to determine if a basic amine was necessary for binding within the aromatic cage and it was shown that analog **6** had no affinity or selectivity for any of the assayed Kme reader proteins. A separate investigation was made to determine whether 53BP1 TTD had similar Kme mimetic preferences as the MBT containing protein, L3MBTL1, as both bind the H4K20me2 mark. Compounds incorporating pyrrolidine amine head groups have been shown to serve as effective ligands for such MBT domain containing proteins [53, 63, 108, 109]. To test this, a pyrrolidine amine group was introduced to give compound **68** and it was found that this modification did not improve affinity towards 53BP1 TTD, confirming that Kme readers of the same histone mark do not necessarily have similar small molecule ligand preferences. Overall, this SAR exploration of the basic amine of **61** (UNC2170) showed that a sterically bulky, lipophilic secondary amine is preferred for binding in the Kme pocket of 53BP1 TTD.

Prior to expanding SAR studies to other portions of the molecule, confirmation that the *N-tert*-butyl amine of **61** (UNC2170) was binding within the aromatic bind cage was necessary. It was hypothesized that a specific binding interaction between a basic head group and the

Asp1521 within the binding pocket was essential. To test this hypothesis, an active site mutant of 53BP1 (D1521A) was prepared that no longer binds H4K20me2 due to loss of a salt bridge and hydrogen bonding with Kme2 [66]. ITC demonstrated that binding of **61** (UNC2170) was abrogated by the 53BP1 D1521A mutant, Figure 2.8 and Table 2.5, which is consistent with the basic amine of **61** (UNC2170) occupying the Kme reader pocket.

Figure 2.8: ITC analysis of prepared analogs against wt and inactive D1521A 53BP1 TTD mutant

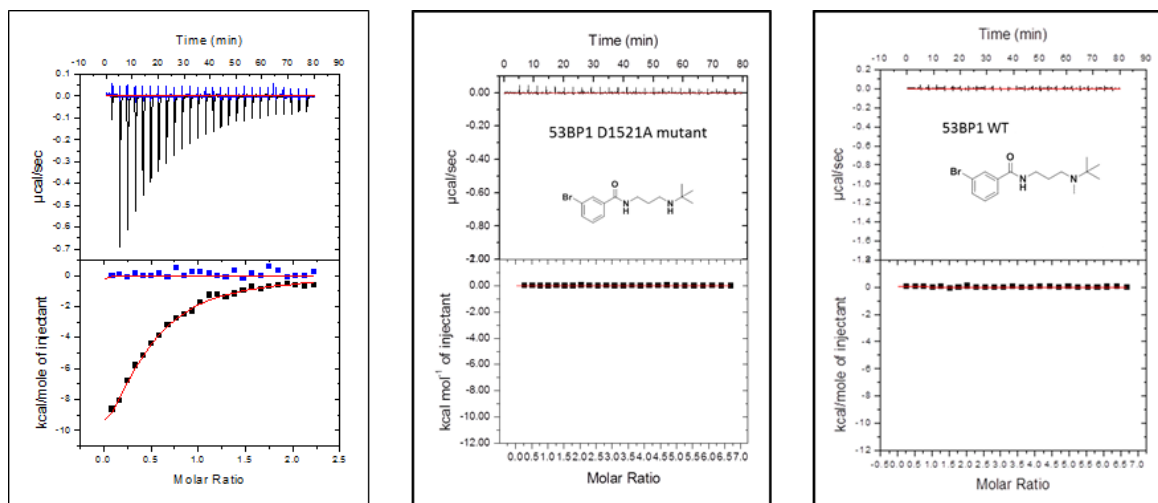


Table 2.5. Thermodynamic data from ITC experiments

Compound	K _d (μM)	ΔH kcal/mol	TΔS kcal/mol	N
61 (UNC2170)	24.6 ± 0.7	-13776.7 ± 496.5	-7.5 ± 0.5	0.6
64	No binding	-	-	-

The AlphaScreen and ITC experiments showed that there was a defined preference for secondary amine head groups over tertiary head groups. These findings are in contrast to previously conducted ITC experiments on endogenous histone 4 peptides where the K20 amine

head group was varied from 0-3 methyl groups. Those findings showed a preference for tertiary (Kme2) over secondary (*tert*-butyl) amine head groups. This was an interesting finding and overall, the SAR exploration of the basic amine of **61** (UNC2170) showed that a sterically bulky, lipophilic secondary amine is preferred for binding in the Kme pocket of 53BP1 TTD. Interestingly, the N-value for **61** (UNC2170) show that there is a less than 1 to 1 binding ratio of protein to ligand.

Modification of linker region

Once an optimal amine head group was determined, modifications of the linker region UNC2170 that affected binding activity were explored. There were two components of the linker region that were explored that included modification of the amide bond of UNC2170 into other type of linker analogs and also modification of the aliphatic propyl chain to determine the optimal distance and flexibility that was necessary between the basic amine of UNC2170 and the aromatic functional group to maintain/increase binding affinity. It was uncertain whether the amide bond present in **61** was involved in any necessary binding interactions with the surface of 53BP1 TTD since an x-ray crystal structure or NMR binding pose was not available. It was hypothesized that modification of the amide bond to other types of linkers would allow for increased binding interactions. Additionally it was uncertain whether the flexible propyl linker of **61** was necessary to maintain binding. To explore this region of compound **61**, several conformationally rigid and also shortened and extended analogs were prepared. The findings are discussed below in Table 2.6.

Table 2.6. SAR Studies of Modifying the Linker of **61** (UNC2170).^{a,b}

ID	R	53BP1	CBX7	JARI D1A	PHF1	PHF19	PHF23	UHR F1	L3MBT L1	L3MBT L3	MBT D1
69		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
70		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	> 100	> 100
71		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
72		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	> 100	> 100
73		> 100 ⁷	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
74		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
75		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
76		> 100	> 100	> 100	NA	NA	FP	> 100	> 100	FP	> 100
77		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
78		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
79		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
80		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
81		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100

^aIC₅₀ values are the average of at least 3 values ± the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 100 μM and compounds showing less than 50% inhibition at this concentration are labeled >100 μM. FP: FP designates assay plates that failed assay controls with Z' < 0.5. NA: NA designates runs where the screened protein was not available for screening due; PHF1 and PHF19 were not validated for screening until after several analogs were synthesized.

Synthetic modification of the amide bond component of the linker region for SAR efforts began with the substitution of a sulfonamide linker for the amide bond to give compounds **69** and **70**. These analogs showed a complete loss of binding affinity for all assayed Kme reader proteins

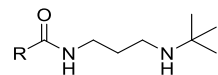
including 53BP1 TTD. Conversion of the amide into a benzylamine, **71**, showed similarly abrogated binding affinities. It was hypothesized that the urea and thiourea linkers would provide additional hydrogen bonding interactions with the protein surface. A series of five urea and thiourea analogs were prepared, **72** – **76**, and unfortunately showed no binding affinity for any of the screened Kme reader proteins. When the linker region between the amide bond and the basic nitrogen was rigidified using two different cyclic variants, analogs **77** and **78**, all affinity was again lost. Additionally, lengthening the aliphatic chain between the amide bond and the amine as in analogs **79** and **80** resulted in inactivity, and shortening the linker, **81**, had a similar effect. These findings show that modification of the amide bond to these linkers results in complete lack of affinity and selectivity for all Kme reader proteins assayed. Additionally, the lengthening or shortening of the initial propyl linker of **61** (UNC2170) also showed complete loss of affinity and selectivity for all assayed Kme reader proteins. Within this series, the amide bond linker and propyl alkyl chain in **61** (UNC2170) was found to be optimal. The results showed that a certain degree of flexibility and the inherent electronic characteristics of the linker region are necessary to maintain binding affinity for 53BP1 TTD.

Modification of aromatic ring





The next region explored in SAR studies was the aromatic ring of **61** (UNC2170). Initial synthetic efforts were made by determining the contribution of the bromine to binding. It was uncertain whether the large bromine group was involved in any Van der Waals interactions or whether its large steric bulk was necessary for binding. Removal of the bromine from the aromatic ring as shown in analog **82** was observed to abolish all binding affinity, Table 2.7. It was uncertain whether the position of the bromine functional group on the aromatic ring provides

specific contributions to binding. Movement of the bromine to the 2- and 4-positions on the aromatic ring via preparation of analogs **83** and **84** showed no measurable binding to 53BP1 TTD. The inactivity of both the bromine regioisomers suggested that the bromine of **61** (UNC2170) at the 3-position of the aromatic ring is making a specific interaction with 53BP1-TTD and that the large size of bromine make unfavorable interactions at the 2- and 4-positions. A halogen scan was then conducted via preparation of analogs **85**, **86**, and **87**. These analogs demonstrated that chlorine (**85**) and fluorine (**86**) provided no improvement in affinity, but replacement of the bromine with iodine was well tolerated (**87**, $IC_{50} = 13 \pm 4.8 \mu M$). Consistent with this, it was found that binding affinity was maintained when an isopropyl group (**89**) or a trifluoromethyl substituent (**91**) occupied the 3-position, suggesting that larger, more lipophilic substituents were preferred. The binding of **91** to 53BP1 was further confirmed by ITC (Experimental, ITC section), which resulted in a K_d that correlated well with the AlphaScreen results ($K_d = 10 \pm 1.0 \mu M$). While the 3-trifluoromethyl analog was roughly equipotent to **61** (UNC2170), activity was also specific to this regioisomer, as compound **92** with CF_3 para to the amide is inactive. Surprisingly, a single methyl group at this position did not provide as high affinity, **88**, nor an ethyl group at the 4-position, **90**, which further underscores the importance of functional group placement at the 3-position. Replacement of the Br of **61** (UNC2170) with a various other functional groups including nitro, amine, naphthylene, and branched aromatic groups at the 3-position (**93** - **97**) also resulted in a loss in potency. Two different pyridine derivatives, **98** and **99** did not provide any affinity.

Table 2.7. SAR Studies of the Aromatic Head Group of **61** (UNC2170). ^{a,b}



ID	R	53BP1	CBX 7	JARID1 A	PHF1	PHF19	PHF23	UHRF 1	L3MBTL 1	L3MBTL 3	MBTD 1
82		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
83		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
84		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
85		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
86		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
87		13 ± 4.8	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
88		91.6 ± 18.8	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
89		14 ± 7.6	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
90		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
91		22 ± 14	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
92		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
93		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
94		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
95		>33 (1) >100 (3)	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
96		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
97		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100

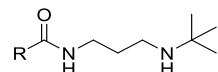
98		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
99		> 100	> 100	> 100	> 100	> 100	> 100	> 100	>33 (1) >100 (4)	> 100	> 100
100		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
101		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100

^aIC₅₀ values are the average of at least 3 values \pm the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 100 μ M and compounds showing less than 50% inhibition at this concentration are labeled >100 μ M. FP: FP designates assay plates that failed assay controls with Z' < 0.5. NA: NA designates runs where the screened protein was not available for screening due; PHF1 and PHF19 were not validated for screening until after several analogs were synthesized.

It was hypothesized that the halogenated 4-substituted pyridine derivatives, compounds **100** and **101**, would increase the positive charge of the σ -hole of the bromine in **61** (UNC2170) in turn increasing the likelihood of a halogen-hydrogen bond with the protein and improving binding [110]. These pyridine analogs resulted in a decrease in potency relative to their phenyl counterparts, suggesting that a potential halogen bond with 53BP1 is unlikely and instead favorable hydrophobic interactions are contributing to binding.

Disubstituted aromatic analogs of **61** (UNC2170)

Synthetic efforts to this point had focused on preparing analogs that had one or two functional groups present in the aromatic ring. In this way, electronic and steric effects associated with various functional groups were tested at various positions on the ring to determine their effect on overall potency. It was hypothesized that additional functional groups would be necessary to pick up further binding interactions to increase potency. Several analogs were prepared that maintained a bromine group at the 3-position of the aromatic ring and additional functional groups were added at the 4-, 5-, and 6-positions of the ring, Table 2.8. The goal of these analogs was to gain new hydrogen-bond or Van der Waals interactions with residues in the acidic/hydrophobic binding pocket.

Table 2.8: SAR Studies of Disubstituted aromatic analogs of **61** (UNC2170). ^{a,b}

ID	R	53BP1	CBX 7	JARID1 A	PHF1	PHF19	PHF23	UHRF 1	L3MBTL 1	L3MBTL 3	MBTD 1
102		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
103		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> ³³ (1) >100 (5)	> 100	> 100
104		22 (1) >100 (1)	> 100	> 100	> 100	NA	> 100	> 100	> 100	FP	> 100
105		77.3 ± 24.8	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	94.0 ± 13.4
106		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
107		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
108		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
109		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
110		59 (1)	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100

^aIC₅₀ values are the average of at least 3 values ± the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 100 μM and compounds showing less than 50% inhibition at this concentration are labeled >100 μM. FP: FP designates assay plates that failed assay controls with Z' < 0.5. NA: NA designates runs where the screened protein was not available for screening due; PHF1 and PHF19 were not validated for screening until after several analogs were synthesized.

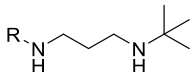
The series of disubstituted aromatic analogs explored the placement of amine and hydroxyl group at the 2- and 5-positions of the ring while maintaining the 3-bromo functional group. It was hypothesized that these functional group would allow for increased hydrogen bonding interactions. It was observed that only **105** showed any affinity for 53BP1 TTD. A small series of other bi-functionalized analogs, **102** – **110**, were also prepared to test the above

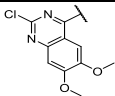
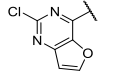
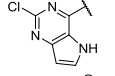
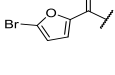
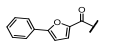
hypothesis, but unfortunately provide no increase in affinity. This SAR series provided evidence that different, functionalized, non-phenyl ring based analogs would potentially have improved binding characteristics compared to the previously prepared and screened analogs.




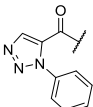








Non-phenyl ring based analogs

The findings from the initial exploration of the aromatic ring showed that the 3-position on an aromatic ring is important to maintaining binding interactions. The aromatic ring of **61** (UNC2170) was proposed to dock within the acidic/hydrophobic pocket of 53BP1 TTD. It was uncertain whether the substituted phenyl ring could be modified into a non-phenyl ring functional group. It was hypothesized that non-aromatic rings, or functional group of different steric bulk and possessing different functional groups may provide increased binding interactions within the acidic/hydrophobic pocket. To test this hypothesis, a number of additional analogs were prepared, Table 2.9.

Table 2.9. SAR Studies of the Non-Phenyl Ring Head Groups of **61** (UNC2170). ^{a,b}



ID	R	53BP1	CBX 7	JARID1 A	PHF1	PHF19	PHF23	UHRF 1	L3MBTL 1	L3MBTL 3	MBTD 1
111		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
112		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
113		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
114		32 (1) >100 (3)	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
115		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100

116		> 100	> 100	> 100	NA	NA	90.7 ± 16.2	> 100	> 100	FP	> 100
117		40 (1) >100 (3)	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
118		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
119		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
120		38 (1) >100 (3)	> 100	> 100	> 100	> 100	> 100	> 100	>33 (1) >100 (4)	> 100	> 100
121		> 100	> 100	> 100	> 100	> 100	> 100	> 100	>33 (1), 62 (1), >100 (3)	> 100	> 100
122		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
123		> 100	> 100	> 100	NA	NA	> 100	> 100	> 100	FP	> 100
124		52 (1) >100 (5)	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
125		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
126		57 (1), 70 (1), >100 (2)	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
127		> 100	> 100	> 100	NA	NA	FP	> 100	> 100	FP	> 100

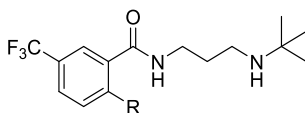
^aIC₅₀ values are the average of at least 3 values ± the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 100 μM and compounds showing less than 50% inhibition at this concentration are labeled >100 μM. FP: FP designates assay plates that failed assay controls with Z' < 0.5. NA: NA designates runs where the screened protein was not available for screening due; PHF1 and PHF19 were not validated for screening until after several analogs were synthesized.

This series of non-phenyl ring analogs explored a large diversity of structures that would re-place the 3-bromo benzoic acid ring in **61** (UNC2170). Analogs **111** – **113** tested bicyclic nitrogen containing rings that showed no affinity for any of the screened Kme reader proteins. Two functionalized furan analogs, **114** and **115**, were prepared and it was observed that neither were active, versus the Kme reader panel. Three analogs that possessed 1 or 2 methylene spacers between the aromatic ring and the amide bond were prepared, **116** – **118**. These unfortunately

also showed little activity against the Kme reader panel. Analogs **119** – **127**, were prepared that contained functionalized triazole, thiazole, benzofuran, thiofuran, 1,2,3,4-tetrahydronaphthalene, chromane, and γ -lactone groups and also showed no activity. This series was surprising as the large diversity of functional groups screened provided no evidence of increased binding interactions against 53BP1 TTD. Findings from this series of analogs suggested that 3-position functionalized aromatic rings were necessary to maintain potency.

Tri-coordinate analogs

SAR efforts focusing on mono-substituted aromatic rings showed some promise in their ability for additional interactions on the protein surface outside of the Kme binding pocket. As a consequence of this observation, a series of di-substituted aromatic analogs that focused on adding a functional group to the aromatic ring at the 2- position while maintain the 5-CF₃ functional group were prepared and tested, Table 2.10. These analogs were created by an aromatic ring displacement of a fluorine group that was para to an electron withdrawing group such as a CF₃ functional group. It was hypothesized that these disubstituted analogs would be able to pick up binding interactions in both the acidic/hydrophobic pocket, via CF₃ binding interactions, maintain the N-*tert*-butyl amine interactions in the Kme binding pocket, and pick up an additional binding interaction of the additional functional group within the arginine binding pocket. It was proposed that these analogs would act in a tri-coordinate manner to bind the three binding pockets simultaneously.

Table 2.10: SAR Studies of 2-, 5-functionalized aromatic analogs of **61** (UNC2170). ^{a,b}

ID	R	53BP1	CBX 7	JARID1 A	PHF1	PHF19	PHF23	UHRF 1	L3MBTL 1	L3MBTL 3	MBTD 1
128		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
129		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
130		79.0 ± 24.6	> 100	> 100	> 100	> 100	> 100	> 100	34.5 ± 9.5	> 100	> 100
131		> 100	> 100	> 100	> 100	> 100	> 100	> 100	29.0 ± 7.8	> 100	> 100

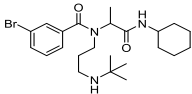
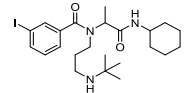
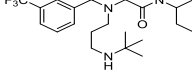
^aIC₅₀ values are the average of at least 3 values ± the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 100 μM and compounds showing less than 50% inhibition at this concentration are labeled >100 μM.

It was observed that only one of the analogs in this series, **130**, provided any affinity for 53BP1 TTD. Interestingly, two of the analogs, **130** and **131**, were observed to provide modest affinity for L3MBTL1. It is uncertain why this set of compounds did not increase affinity for 53BP1 TTD, but it is proposed that the position of the functional groups that are located on the 2-position of the aromatic ring did conform to a proper geometry to allow for binding into the Arg19 binding pocket. Findings from this series were used in development of a different set of tri-coordinate analogs that were synthesized using a different type of synthetic reaction. It was hoped that placement of the third functional group, proposed to bind within the Arg19 binding pocket, would allow for a different conformation geometry to allow for better binding properties to all three binding pockets.

Ugi analogs

Synthetic efforts to this point had focused on preparing analogs that had one or two functional groups present upon an aromatic, or a non-aromatic ring. It was hypothesized that functional groups at different locations other than the aromatic ring of **61** (UNC2170) would be necessary to pick up additional binding interactions for increase potency. Three analogs were prepared that maintained a bromine group at the 3-position of the aromatic ring and a functional group off of the nitrogen on the amide bond with the goal of gaining new hydrogen-bond or Van der Waals interactions. The core structure of these analogs would possess the back-bone structure of **61** (UNC2170), but would include a third functional group located on the amide bond. To prepare these analogs, an Ugi multicomponent reaction was used to synthesize amide-nitrogen functionalized analogs. The goal of this series was to gain new binding interactions with the arginine binding pocket from addition of the third functional group while maintaining the previously discovered 3-bromo position and *N*-propyl, *N*-*tert*-butyl amine functional groups of **61** (UNC2170). These variants differed from the UNC2170 core structure in that the aliphatic propyl linker between the amide bond and amine head group was not modified to α -methyl *N*-functionalized amide, Table 2.11.

Table 2.11: SAR Studies of Ugi analogs of **61** (UNC2170). ^{a,b}

ID	R	53BP1	CBX 7	JARI D1A	PHF1	PHF19	PHF 23	UHR F1	L3MBT L1	L3MBT L3	MBT D1
132		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
133		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100
134		> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100	> 100

^aIC₅₀ values are the average of at least 3 values \pm the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 100 μ M and compounds showing less than 50% inhibition at this concentration are labeled >100 μ M.

It was unfortunately found that all of these analogs were inactive versus the screened Kme reader proteins. It is proposed that the altered structure of these molecules that possessed an α -methyl N-functionalized amide were not able to obtain an optimal conformation to all three binding pockets simultaneously.

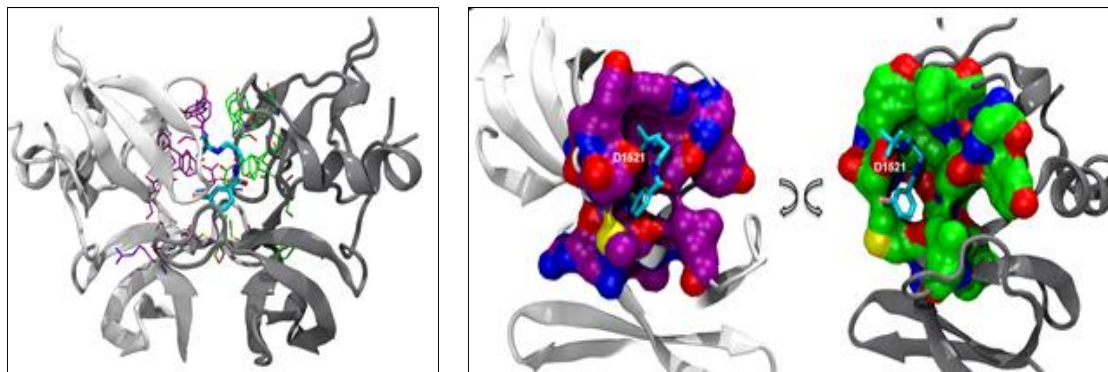
At this point in the medicinal chemistry efforts, it was shown that a large number of synthetic modifications on all of the focus regions of **61** (UNC2170) did not provide an increase in the potency of the molecule. It was therefore decided to stop medicinal chemistry efforts and further characterize the **61** (UNC2170) to define how it bound to 53BP1 TTD and to also determine whether it possessed the ability to inhibit the biological activity and function of 53BP1 TTD in both *in vitro* and *in vivo* assays.

Structural Binding Analysis

X-Ray co-crystallization experiments

To better understand the structural basis for UNC2170 binding, efforts to co-crystallize the 53BP1 tandem Tudor domain with **61** (UNC2170) were conducted. A crystal was successfully obtained by our collaborators at SGC Toronto (Dr. Cheryl Arrowsmith, Dr. Aiping Dong, and Dr. Pavel Made), and its structure was re-solved by X-ray crystallography at 1.5 Å resolution (PDB 4RG2). Interestingly, it was found that under these conditions **61** (UNC2170) was bound to a 53BP1 TTD dimer, making distinct interactions with both Tudor domains, Figure 2.9.

Figure 2.9: X-ray co-crystallization structure of **61** (UNC2170) bound to 53BP1 TTD dimer



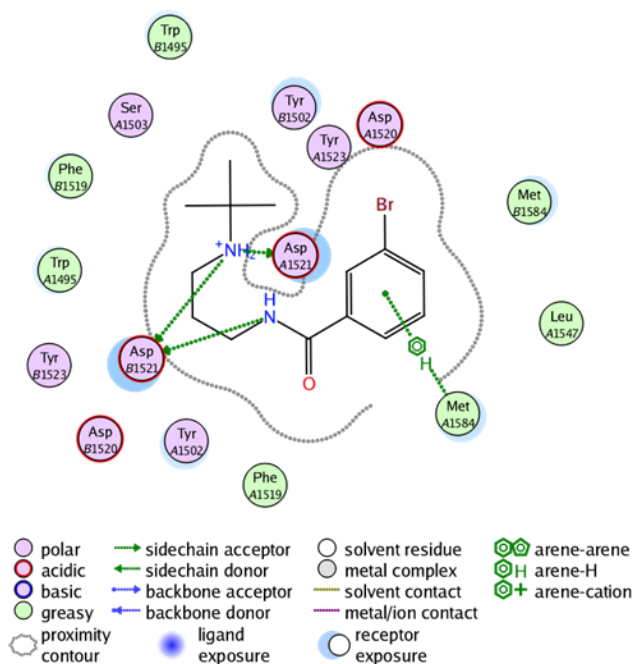
The images in Figure 2.9 show a 2:1 dimer of 53BP1 TTD to **61** (UNC2170). This correlates well with the molar ratio observed by ITC (Figure 2.8), which consistently produced an N value less than 1, suggesting a 2:1 ratio of protein to ligand. The structure revealed that the *N-tert*-butyl amine of **61** (UNC2170) was buried in the Kme binding pocket of one tudor domain. It is proposed that this binding interaction was driven largely by van der Waals interactions, hydrogen bonds, and electrostatic interactions. Due to the fact that the aromatic residues of both Kme pockets were greater than 5 Å from the amine, it was anticipated that cation- π interactions are a minor contribution to binding relative to the hydrogen bonds formed between the ligand and protein. The basic amine of **61** (UNC2170) was about equidistant from D1521 in each protein binding pocket (2.8 Å and 2.6 Å), and was well positioned to interact with and make a critical hydrogen bond to D1521 in both proteins. The importance of this interaction was consistent with ITC data that showed that the Kme pocket mutant, D1521A, does not bind to **61** (UNC2170). Additionally, the amide nitrogen is also within hydrogen bonding distance of the D1521 residue of one tudor domain, which is likely to further facilitate binding. The distance between the amine and the amide of **61** (UNC2170), therefore, appears to be quite important in enabling a number of key interactions, and this sheds light on why changing the length of the

aliphatic linker (analogs **79** – **81**) impairs binding. The aliphatic linker and aromatic ring of UNC2170 is also nicely sandwiched between the dimer interfaces, packing against the surface of each protein unit.

The co-crystal structure greatly assisted in understanding why there was a very constrained SAR in the studies described above, as the interaction of **61** (UNC2170) with the 53BP1 dimer reveals that the ligand is encircled by both proteins and that there is limited space for modification. This is in contrast to our preliminary structural predictions based on **61** (UNC2170) bound to a single 53BP1 Tudor domain. Interestingly, this is the second example of how Kme binding proteins have been observed to dimerize around a small molecule ligand; the other example being the small molecule chemical probe UNC1215 that binds via L3MBTL3 protein dimerization around the molecule. It is uncertain at this time whether this is a property of this type of binding protein or an artifact of X-ray and NMR coordination studies.

Additionally, the binding interactions of **61** (UNC2170) and the 53BP1 TTD dimer were modeled by Dr. Bradley Dickson and showed a large number of the predicted binding interactions, Figure 2.10.

Figure 2.10: Key interactions between **61** (UNC2170) and the 53BP1 Tandem Tudor domains.



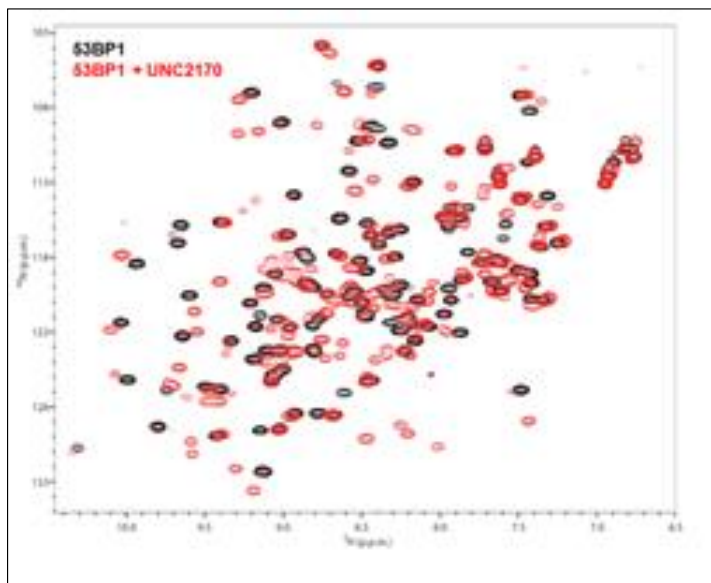
These interactions were determined by the modeling program MOE based on the **61** (UNC2170) co-crystal structure (PDB 4RG2). The residues from one Tudor domain are labeled as *A*, and the residues from the second Tudor domain are labeled as *B*. The *N-tert*-butyl group is buried in the Kme pocket of Tudor *B*.

Protein NMR binding experiment

In addition to the X-ray structure, a ^1H - ^{15}N HSQC NMR correlation spectrum of the 53BP1 TTD was also obtained in the absence and presence of a 10-fold excess of **61** (UNC2170). This assay was conducted by our collaborator, Prof. Georges Mer (Mayo Clinic). Extensive changes in chemical shifts were observed, Figure 2.11, further confirming specific binding interactions of the ligand to 53BP1 TTD. Exchange between the bound and unbound

state was observed to be slow on the NMR chemical shift time scale which is surprising given the low measured affinity of **61** (UNC2170).

Figure 2.11: NMR structural analysis of **61** (UNC2170) binding to 53BP1 TTD



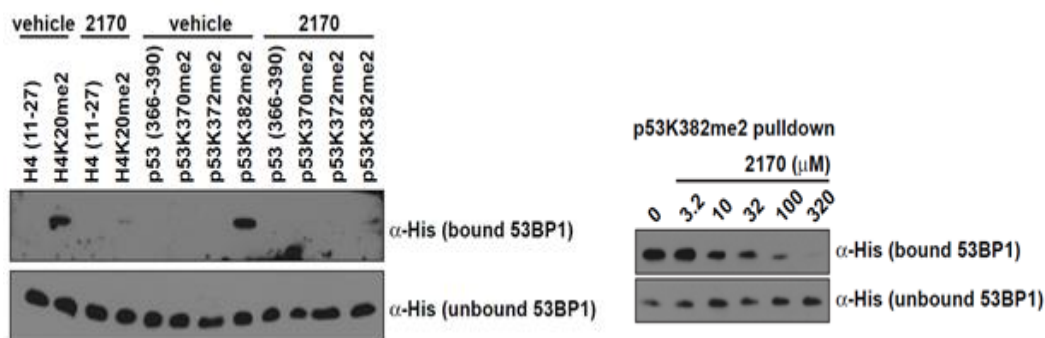
This NMR coordination study showed that there were two populations of NMR signals present for the 53BP1 TTD residues and also signify the potential presence of a protein dimer. Unfortunately, the exact residue coordination was not properly identified, nor was titration of **61** (UNC2170) conducted, so a NMR measurement of a K_d value was not obtained. Both of these structural studies support the existence of a dimeric inhibitor/53BP1 interaction and provide direct evidence that the small molecule ligand directly engages the 53BP1 TTD through a series of well-defined binding interactions.

In Vitro and Cellular Assay Characterization

Competitive In-solution Peptide Pulldown Assays

To further evaluate whether **61** (UNC2170) was able to compete with endogenous Kme2 peptide substrates for binding, H4K20me2 and p53K382me2 peptide pull down assays were conducted. This assay was conducted by our collaborator, Dr. Scott Rothbart (Strahl lab, UNC-CH). **61** (UNC2170) was observed to successfully displace the tandem Tudor domain of a His-53BP1 fusion protein from both H4K20me2 and p53K382me2 peptides, Figure 2.12. Additionally, **61** (UNC2170) displaces His-53BP1 TTD from immobilized p53K382me2 in a dose dependent fashion, resulting in an apparent IC_{50} of about 30 μ M which is consistent with the results from other *in vitro* assays described above.

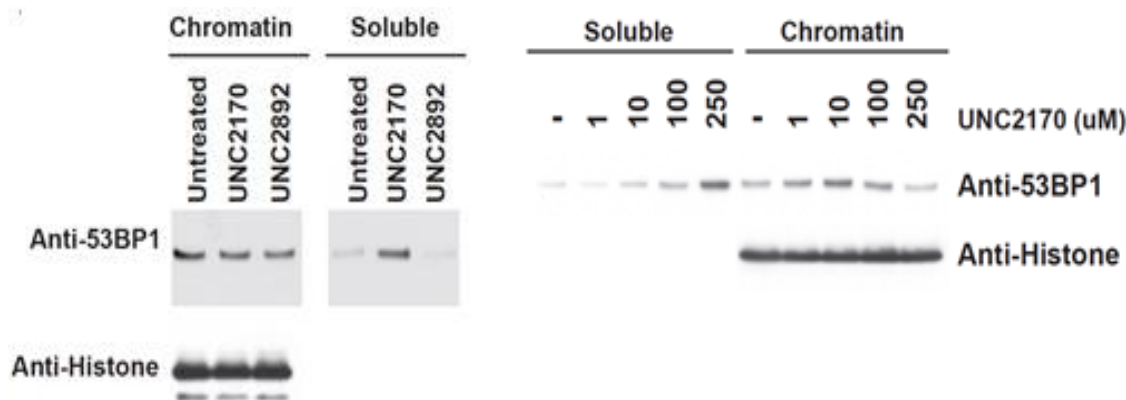
Figure 2.12: Competitive in-solution peptide pulldown assay



Competitive Chromatin Fractionation Assay

To further study whether **61** (UNC2170) was able to competitively bind 53BP1 TTD, chromatin release assays were conducted in HEK293 cell lysates to determine if UNC2170 could influence the amount of endogenous 53BP1 bound to chromatin. This assay was conducted by our collaborator, Prof. Kevin McBride (MD Anderson). Purification of chromatin from HEK293 cells was followed by compound treatment for two hours at room temperature and subsequent separation of the solubilized proteins from those remaining on chromatin. Western blot analysis revealed that treatment with **61** (UNC2170) (500 μ M) resulted in a significant increase in soluble 53BP1 as compared to untreated lysates or lysates treated with **64** (UNC2892), the negative control compound, Figure 2.13.

Figure 2.13: Chromatin fractionation assay

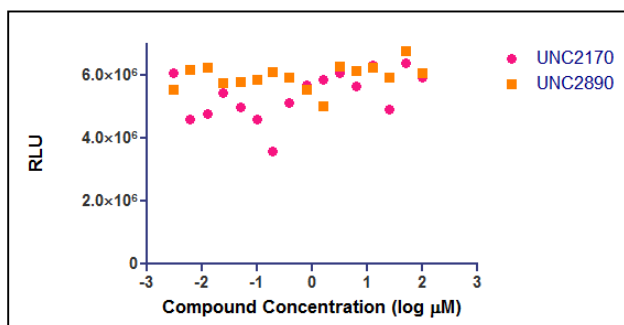


This effect was found to be concentration dependent, as an increase in soluble 53BP1 was observed with increasing concentrations of **61** (UNC2170), and partial release of 53BP1 was detected at concentrations as low as 10 μ M. This result is consistent with **61** (UNC2170) antagonizing the interaction between the Tudor domain of full-length 53BP1 and chromatin.

Cell permeability and cell toxicity assays

After conducting several *in vitro* assays, it was necessary to evaluate the potential of **61** (UNC2170) to engage 53BP1 TTD in a cellular context. The low molecular weight (313.24 g/mol) and clogP (~3.27) of the small molecule ligand suggested *a priori* that **61** (UNC2170) would be cell permeant. To confirm this and assess the suitability of the compound for cellular studies, the cell permeability of **61** (UNC2170) was conducted by an external lab (Absorption Systems, Exton, PA). It was observed that **61** (UNC2170) was indeed highly cell permeant with no significant measureable cellular efflux (efflux ratio = 1.2) as determined by a bi-directional Caco-2 cell permeability assay {Smart, 2011 #160}. Once cell permeability was determined, the toxicity of the compound was measured using a CellTiter-Glo luminescent cell viability assay. This assay was conducted with the assistance of Dr. Nancy Cheng (CICBDD). It was observed that **61** (UNC2170) had no measurable toxicity within the desired concentrations for cellular experiments, with cell toxicity only beginning to take effect at >10 mM, Figure 2.14.

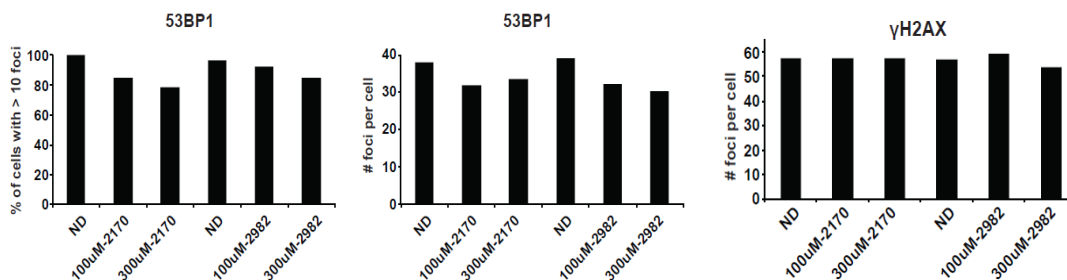
Figure 2.14: CellTiter-Glo luminescent cell viability assay for **61** (UNC2170) and **64** (UNC2892)



Foci Formation Assay

It is known that DNA double-strand (dsDNA) breaks occur after treatment with ionizing radiation (IR), and the functional tandem Tudor domain of 53BP1 is required for foci formation at those break sites [64]. Foci are denoted as oligomerized 53BP1 at break sites in response to identification of DNA double strand breaks. It was hypothesized that treatment with cell permeant and non-toxic **61** (UNC2170) would impair the ability of 53BP1 TTD to form foci at IR-induced DNA double strand break sites. To test this hypothesis, our collaborator Prof. Mark Bedford (MD Anderson) pretreated U2OS cells that were transfected with GFP-tagged full length 53BP1 with **61** (UNC2170) (1 hour, 100 – 300 μ M). Once the treatment period was over, the cells were irradiated with 5 Gy irradiation, followed by analysis of the extent of foci formation in cell culture and also the number of foci formed in each cell. Unfortunately, it was observed that treatment with **61** (UNC2170) did not significantly decrease the extent of endogenous 53BP1 foci formed following 5 Gy irradiation, and similar results were observed with the negative control analog **64** (UNC2982), Figure 2.15. An additional aspect of this experiment was to measure the formation of γ H2AX foci at break sites and they were quantified in parallel after compound treatment and no changes were observed.

Figure 2.15: Foci formation assay

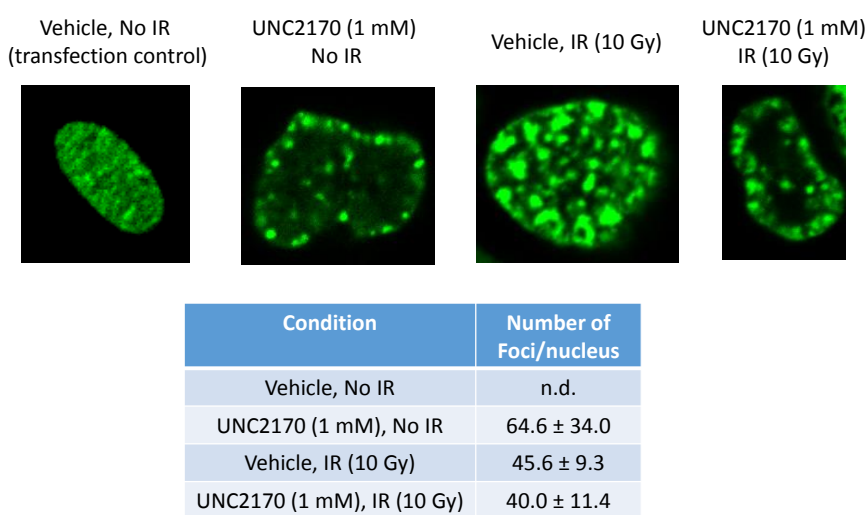


Fluorescence recovery after photobleaching assay

This assay was used as a follow-up to the foci formation assay that was proposed to determine whether the biological activity of the Tudor domain of 53BP1 was affected upon inhibition by treatment with **61** (UNC2170). The FRAP assay is based on measuring the kinetics of how a system responds to photobleaching of a defined region of the cell that contains a GFP-tagged protein. The region of interest that contains the GFP-tagged protein, i.e. the GFP-tagged isolated 53BP1 TTD, is photobleached using a high power laser on a confocal microscope leaving a “photobleached region” where there is no GFP activity. The confocal microscope is then used to monitor how quickly unbleached GFP-tagged 53BP1 TTD can respond to this event via the movement of non-photobleached GFP-tagged protein into that photobleached region. It was hypothesized that mobility of the GFP-53BP1 isolated Tudor domain would be increased upon treatment with **61** (UNC2170) as the Tudor domain would be unable to bind its endogenous target; H4K20me2 PTMs at the sites of dsDNA breaks; and be more mobile within the nucleus and better able to diffuse and recover the fluorescence at the photobleached site of damage. These experiments were conducted by transfecting U2OS cells with a GFP-tagged isolated 53BP1 TTD and then treating with 1 mM of **61** (UNC2170) with 15 min incubation, then irradiating the cells over a range of different IR doses (0, 2, 5, and 10 Gy). The concentration of **61** (UNC2170) was chosen as it was well above the *in vitro* IC₅₀ and it was uncertain how much was necessary to induce a cellular response and also that concentration was well within the a non-toxic concentration range for U2OS cells. After IR treatment, the cells were incubated for 1 hr to allow for foci formation. These DNA damaged cells were then exposed to defined regions of photobleaching (405 nm wavelength at 100% power, 10 flashes at 1 sec intervals) in their

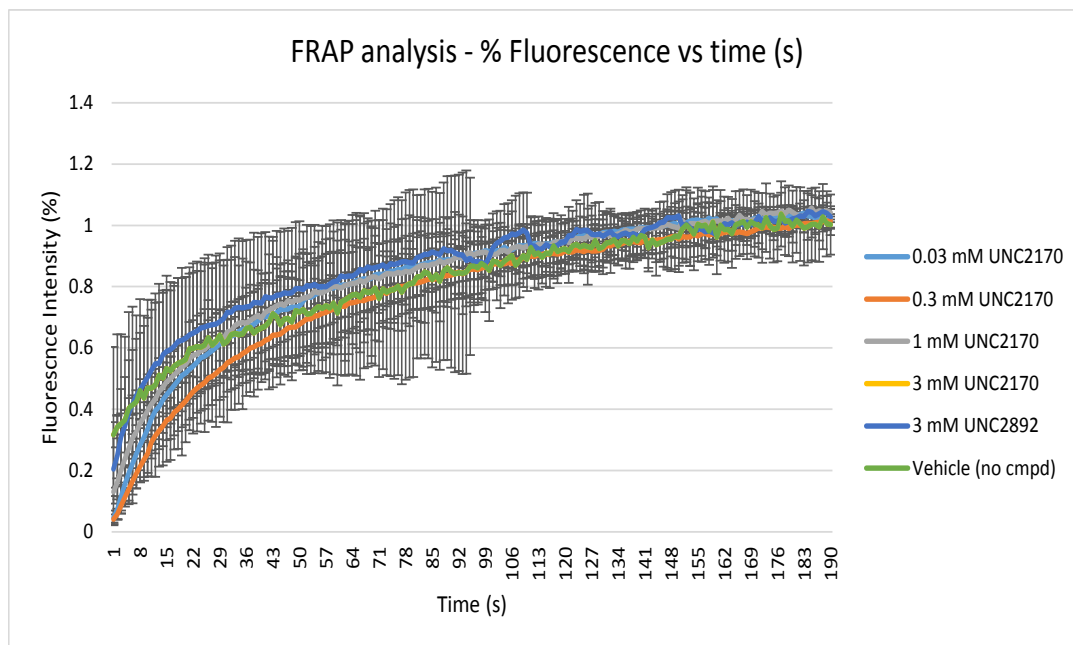
nucleus (3 – 5 photobleached foci /nucleus, with 3-5 cells tested per irradiation condition) using a Zeiss LSM 700 confocal microscope. Initial IR dosing experiments without inhibitor treatment showed that 2 Gy did not lead to measurable foci formation and that 5 and 10 Gy were optimal. Additional treatment both with and without IR (10 Gy) and with and without 1 mM treatment with **61** (UNC2170) showed the following regarding foci formation, Figure 2.16.

Figure 2.16: Initial dose range for FRAP experiments



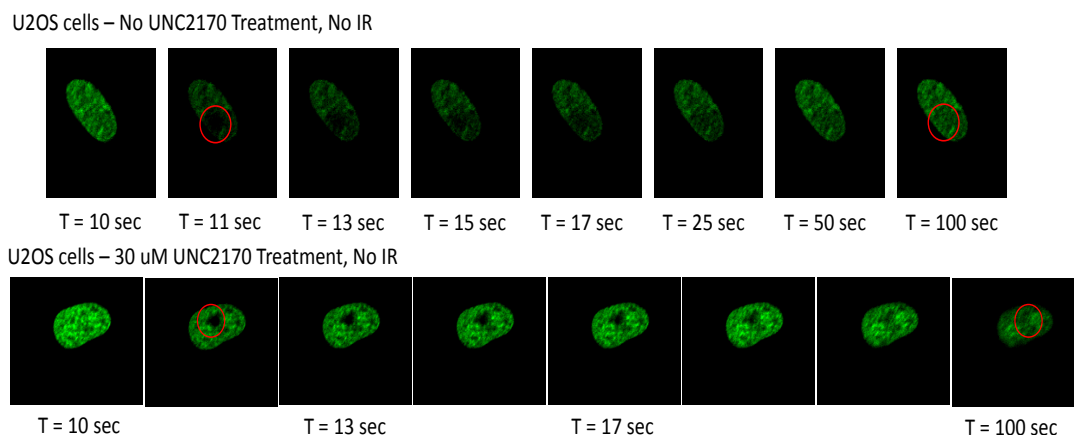
From these initial findings, a series of inhibitor experiments were conducted that varied the concentration (30 μ M, 300 μ M, 1 mM, 3 mM). Additionally, the treatment of the negative control analog **61** (UNC2892) at 3 mM showed highly inconsistent results in the presence of IR treatment, Figure 2.17. It was noted that the IR treatment often led to drastic cell morphology changes compared to non-IR treated cells.

Figure 2.17: Variation of concentration of **61** (UNC2170) at 10 Gy IR



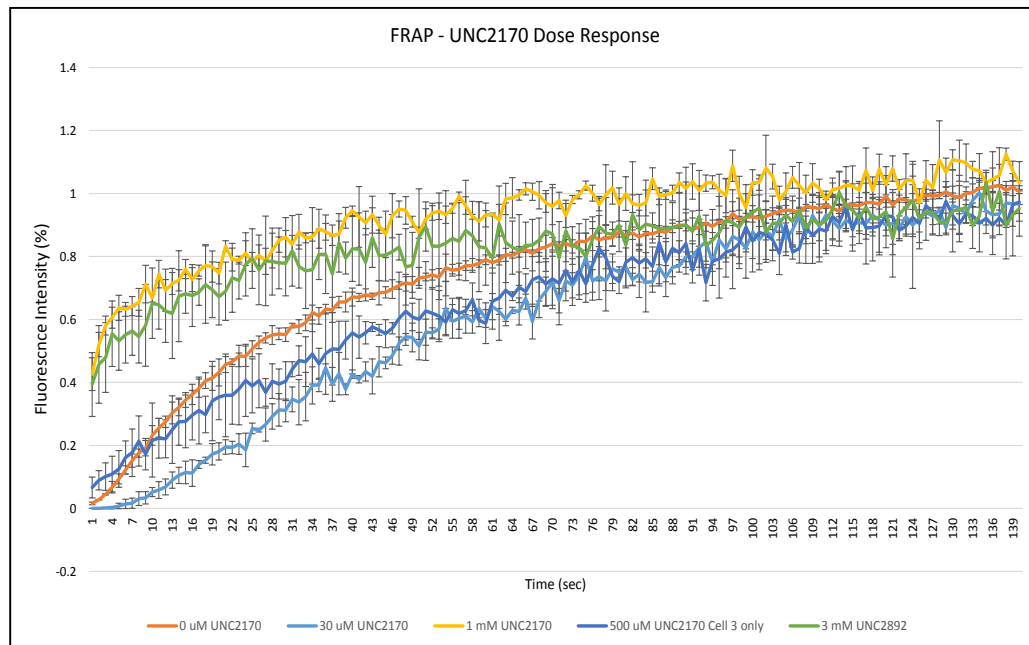
A different approach was conducted that did not include the use of IR treatments. This approach was based on a previous literature example of how a GFP-tagged inactive Y1487L mutant 53BP1 TTD was exposed to FRAP study using the same conditions and showed that the mutant has a higher degree of mobility than wild type GFP-tagged 53BP1 TTD.[77] It was hypothesized that treatment of **61** (UNC2170) would mimic the inactive form Y1487L 53BP1 used in that study as the small molecule ligand would bind to the TTD and lead to its inability to bind chromatin. This hypothesis was tested by transfecting U2OS cells with a GFP-tagged isolated 53BP1 TTD and then incubation these cells overnight with various concentrations of **61** (UNC2170). Post incubation, the cells were washed with fresh media and then FRAP experiments were conducted on both non-inhibitor treated cells and treated cells. A representative image of the recovery process is shown in figure 2.18.

Figure 2.18: Representative image of treated and untreated U2OS cell transfected with GFP-tagged isolated 53BP1 TTD response to photobleaching



Using these new conditions in the absence of IR treatment, experiments were conducted that varied the concentration of **61** (UNC2170) over a 24 hr incubation period, followed by FRAP. The concentrations tested were 30 μ M, 500 μ M, and 1 mM **61** (UNC2170), along with a 3 mM **64** (UNC2892) as a negative control. It was observed that cell morphology changed and there was increased toxicity as the concentration of **61** (UNC2170) increased. Additionally, with high concentrations of UNC2170, photobleached regions do not fully recover (~85-90%), or recover at all (by visual inspection). The negative control compound unfortunately did not replicate the no-treatment control (vehicle). Interestingly, it was observed that low doses of **61** (UNC2170 (30 and 500 μ M) appear to have slow recovery after photobleaching, but the high dose (1 mM) acted like the Y1487L mutant, Figure 2.19.

Figure 2.19: FRAP experiments without IR



Unfortunately, the two different forms of FRAP experiments that were conducted failed to show any consistent affect that **61** (UNC2170) has upon GFP-tagged isolated 53BP1 TTD after photobleaching. It is hypothesized that the modest potency of the inhibitor is not sufficient to induce the desired effect of preventing 53BP1 TTD from binding its defined target. A more rapid recovery after photobleaching was anticipated, due to more rapid diffusion of non-chromatin bound 53BP1 TTD, since it would itself be bound by **61** (UNC2170) and free to move to photobleached regions.

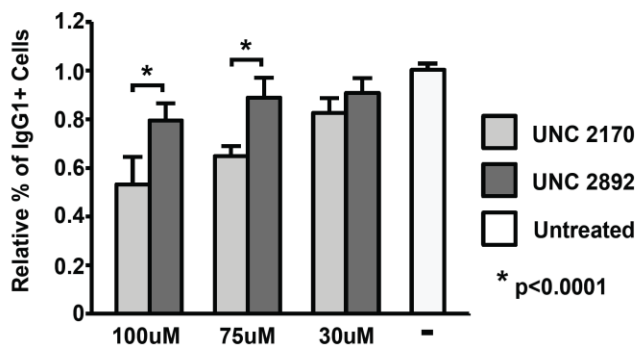
A different approach that was not utilized to determine whether NHEJ or HR was being either up- or down-regulated would be the use of a novel genome engineering reporter system.[111] This system known as 'traffic light' uses a construct containing two fluorescent protein sequences that are opposing expressed upon nucleosome activity in response to DDR activation, in conjunction with flow-cytometry, to determine which specific DDR pathway is

active.[111] Use of this quantitative DDR analysis system would provide a different method to analyze whether UNC2170 has any specific effect upon modulation of the selection of either HR or NHEJ.

Class Switch Recombination Assay

Efficient B cell antibody class switch recombination (CSR) requires 53BP1 to have an active tudor domain (48). It has been observed that 53BP1 knockout mice are severely impaired in their ability to undergo CSR [112, 113] and this led to the hypothesis that treatment of naïve B-cells with **61** (UNC2170) would induce this same defect. This hypothesis was tested through the use of naïve IgM⁺ B cells that were stimulated to undergo isotype switching from IgM to IgG1 via use of LPS and Il-4 in combination with treatment of either UNC2170, UNC2892, or vehicle. Relative switching to IgG1 was then determined by FACS analysis 3.5 days after compound treatment. B cells treated with UNC2170 (75 μ M) underwent CSR 64% as effectively as untreated cells, while treatment with UNC2892 had a much less significant inhibitory effect on CSR (89%), Figure 2.20.

Figure 2.20: CSR effects upon treatment with **61** (UNC2170)

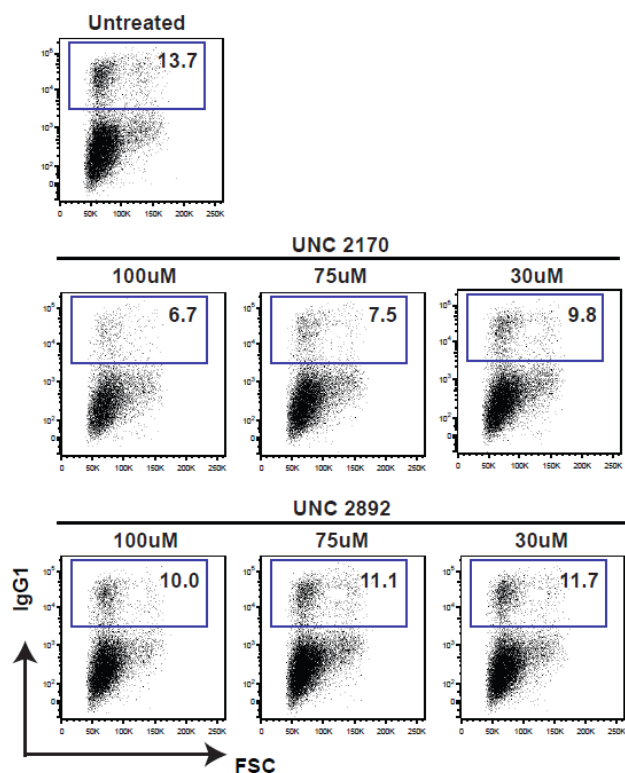


Although slightly higher and lower doses were tested (100 μ M and 30 μ M, respectively), treatment at 75 μ M appears to be optimal in that the effects observed are specific to the 53BP1 active compound. Attempts to probe the effect on CSR upon treatment with 300 μ M of UNC2170 were unsuccessful, as this resulted in a reduction in cell viability. However, this effect could be mechanism related as UNC2892 does not have the same anti-proliferative effects, and B cells undergoing CSR endure a high rate of DNA damage while proliferating rapidly, and therefore, represent a particularly sensitive cell subset. It has been reported previously that B cells stimulated from mice bearing a single 53BP1 D1518R amino acid substitution (this mutant is equivalent to D1521R in humans) switch at about 10% of wild type levels [112]. While the effect of UNC2170 is not as pronounced, the treatment with UNC2170 clearly phenocopies the reduction in CSR seen in 53BP1 mutant B cells.

Analysis of CSR to IgG1 switching in wildtype splenocytes cultured with LPS and IL-4 and incubated with indicated compounds for 3.5 days was conducted. Treatment with UNC2170 decreased the number of cells that underwent isotope switching to IgG1, whereas UNC2892 did not have the same effect. CSR was measured by IgG1⁺ cell surface expression and plotted versus

forward size scatter (FSC). The percentage of IgG1⁺ cells is indicated on the top right of each graph, Figure 2.21.

Figure 2.21: FACS analysis of CSR assay cells 3.5 post treatment



Conclusions

It was demonstrated that the small molecule 53BP1 TTD ligand **61** (UNC2170) possesses modest potency and a 17-fold selectivity for 53BP1 TTD as compared to nine other Kme reader proteins. The binding of **61** (UNC2170) was also dependent upon a functional tandem tudor domain as demonstrated by site directed mutagenesis and ITC studies. A co-crystal structure revealed that the small molecule ligand binds within the same Kme binding pocket as

endogenous Kme peptides. However, distinct from Kme peptide ligands, **61** (UNC2170) engages the binding pockets of two 53BP1 tudor domains. This simple, relatively non-toxic ligand also exhibits modest activity as a 53BP1 TTD antagonist in cellular lysates and has functional consequences in CSR assays consistent with its weak *in vitro* activity. Unfortunately, no biological consequences of **61** (UNC2170) were observed in the cellular foci formation assay, nor in the kinetics of the proteins in two different FRAP studies. These findings suggest that **61** (UNC2170) may not be potent enough to exhibit biological outcomes in these two cellular assay platforms. A closely related negative control compound, **64** (UNC2892), lacks both *in vitro* and cellular effects, suggesting that the results observed with **61** (UNC2170) are indeed due to the inhibition of 53BP1 TTD. Thus, **61** (UNC2170) is a novel small molecule ligand of 53BP1 TTD, a Kme reader protein that has gained much attention recently due to its integral role in DNA damage repair and its link to BRCA1. Our SAR studies revealed only modest improvements in potency, and further exploitation of the structural data should facilitate the development of a high-quality chemical probe for 53BP1. This work also aids in demonstrating the potential for Kme readers to be modulated via small molecule intervention, despite the limited research in this field to date.

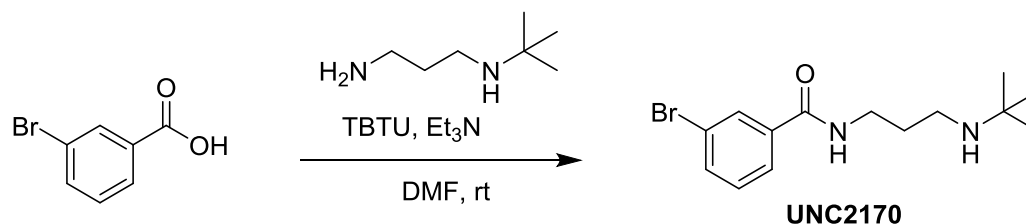
Experimental section

General Procedure for Chemical Synthesis

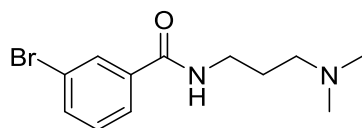
Analytical LCMS data for all compounds were acquired using an Agilent 6110 series system with the UV detector set to 220 and 254 nm. Samples were injected (<10 μ L) onto an Agilent Eclipse Plus 4.6 \times 50 mm, 1.8 μ m, C18 column at room temperature. A mobile phase of

A (H₂O + 0.1% acetic acid) and B (MeOH + 0.1% acetic acid) was used with a linear gradient from 10% to 100% B in 5.0 min, followed by a flush at 100% B for another 2 minutes with a flow rate of 1.0 mL/min. Mass spectra data were acquired in positive ion mode using an Agilent 6110 single quadrupole mass spectrometer with an electrospray ionization source. Nuclear Magnetic Resonance (NMR) spectra were recorded on a Varian Mercury spectrometer at 400 MHz for proton (¹H NMR) and 100 MHz for carbon (¹³C NMR); chemical shifts are reported in ppm (δ). Analytical thin-layer chromatography (TLC) was performed with silica gel 60 F₂₅₄, 0.25 mm pre-coated TLC plates, generally using a 10% MeOH in DCM solvent system. TLC plates were visualized using UV₂₅₄, I₂ impregnated silica gel, *potassium permanganate* with charring, and phosphomolybdic acid with charring. Reverse phase chromatography was used to purify reaction mixtures to obtain final products using a Teledyne Isco CombiFlash Rf 200 chromatography unit equipped with the UV detector set to 220 nm and 254 nm. Samples were injected onto a RediSep Rf 30g C18 high performance Gold column at room temperature and collected at the previously mentioned wavelengths. Mobile phases of A (H₂O + 0.1% TFA) and B (MeOH) were used with a flow rate of 30 mL/min. A general gradient was used consisting of 0-2 minutes at 5% B, 5-15 minutes increasing from 5 to 100% B, and a 100% B flush for another 3 minutes. Small variations in this purification method were made as needed to achieve ideal separation for each compound. All compounds that were evaluated in biochemical and biophysical assays had >95% purity as determined by ¹HNMR and LCMS.

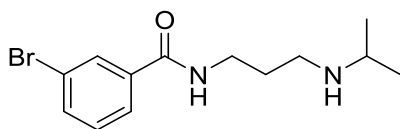
Small molecule synthesis



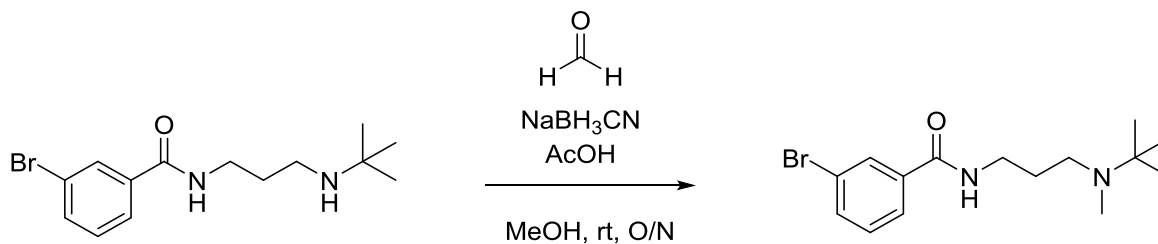
3-Bromo-*N*-(3-(*tert*-butylamino)propyl)benzamide (61, UNC2170): To a solution of 400 mg (1.98 mmol) of 3-bromobenzoic acid in 2 mL of DMF was added 831 mg (2.6 mmol) of TBTU and the mixture was stirred for 5 min at rt. Upon complete dissolution, 387 mg (2.4 mmol) of *N*-(*tert*-butyl)propane-1,3-diamine and 0.6 mL of Et₃N (6 mmol) was added and the solution was stirred at rt overnight. The reaction was quenched by addition of 10 mL of sat. aq NaHCO₃ and diluted with 15 mL of CH₂Cl₂. The aqueous phase was extracted with CH₂Cl₂ (4 × 15 mL). The combined organic extracts were washed with brine, dried over Na₂SO₄ and filtered. The solvent was evaporated and the crude product was dissolved in 8 mL of CH₂Cl₂ followed by the addition of silica gel. The solvent was again removed by rotary evaporation and the crude material was purified by reverse phase column chromatography using an automated Teledyne Isco chromatography unit to afford 117 mg (69%) of the TFA salt as a yellow, waxy solid. ¹H NMR (400 MHz, CD₃OD) δ 8.02 (t, *J* = 1.8 Hz, 1H), 7.82 (m, 1H), 7.73 (m, 1H), 7.42 (t, *J* = 7.9 Hz, 1H), 3.51 (t, *J* = 6.5 Hz, 2H), 3.08 – 3.00 (m, 2H), 2.03 – 1.92 (m, 2H), 1.40 (s, 9H). ¹³C NMR (101 MHz, CD₃OD) δ 169.4, 137.1, 135.9, 131.6, 131.5, 127.1, 123.6, 58.0, 39.9, 37.7, 28.3, 25.9. LC-MS (λ = 254 nm): 100%, *t*_R = 3.7 min. HRMS calculated for C₁₄H₂₀N₂OBr - H: 311.0759; found: 311.0762 [M-H]⁻.



3-Bromo-*N*-(3-(dimethylamino)propyl)benzamide (62): Compound **62** was prepared from 3-bromobenzoic acid (0.15 g, 0.75 mmol) and *N,N*-dimethylpropane-1,3-diamine (0.12 g, 1.1 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (112 mg, 38%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.02 (s, 1H), 7.82 (d, $J = 7.0$ Hz, 1H), 7.69 (d, $J = 7.2$ Hz, 1H), 7.39 (t, $J = 7.4$ Hz, 1H), 3.48 (t, $J = 6.6$ Hz, 2H), 3.19 (t, $J = 8.1$ Hz, 2H), 2.90 (s, 6H), 2.13 – 1.96 (m, 2H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.1$ min. MS (ESI $^+$): 285.1 + 287.1 $[\text{M}+\text{H}]^+$.

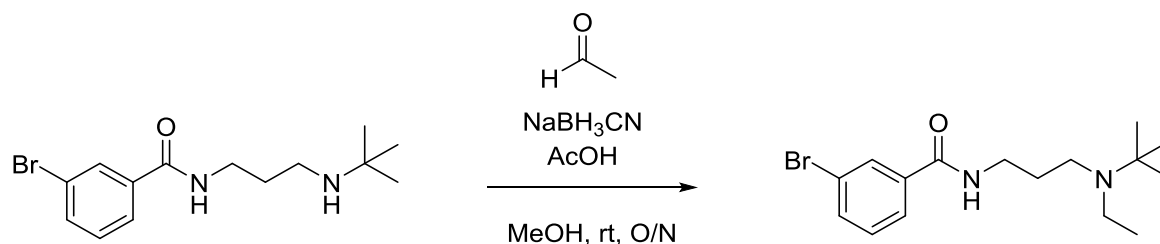


3-Bromo-*N*-(3-(isopropylamino)propyl)benzamide (63): Compound **63** was prepared from 3-bromobenzoic acid (0.15 g, 0.75 mmol) and *N*-isopropylpropane-1,3-diamine (0.10 g, 0.90 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (113 mg, 37%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.01 (t, $J = 1.8$ Hz, 1H), 7.84 – 7.79 (m, 1H), 7.70 (ddd, $J = 8.0, 1.9, 0.8$ Hz, 1H), 7.39 (t, $J = 7.9$ Hz, 1H), 3.50 (t, $J = 6.5$ Hz, 2H), 3.36 (m, 1H), 3.06 (t, $J = 7.4$ Hz, 2H), 2.04 – 1.93 (m, 2H), 1.34 (d, $J = 6.6$ Hz, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.5$ min. MS (ESI $^+$): 299.1 + 301.1 $[\text{M}+\text{H}]^+$.



3-Bromo-*N*-(3-(*tert*-butyl(methyl)amino)propyl)benzamide (64, UNC2892): To a solution of **61** (84 mg, 0.27 mmol) in 4 mL of MeOH was added 33 μL (0.44 mmol) formaldehyde solution (37% wt in H_2O), 101 mg (1.6 mmol) sodium cyanoborohydride, and 77 μL (1.3 mmol) acetic

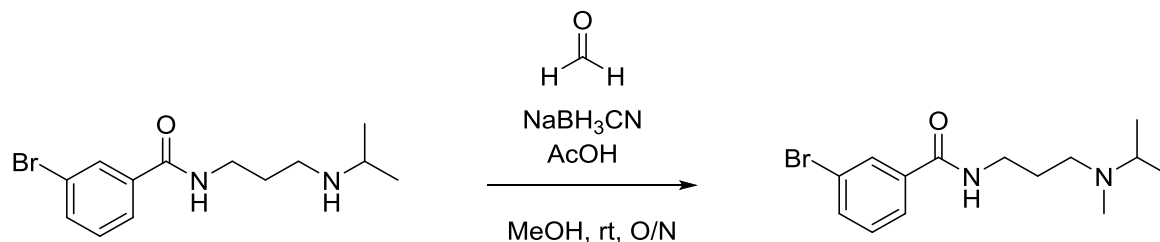
acid, and the mixture was stirred overnight at rt. The reaction was monitored by LC-MS until completion and then dried to residue. The crude material was dissolved in 1 mL of DCM on silica and purified by reverse phase column chromatography using the Teledyne Isco automated column system to afford 106 mg (90% yield) as the TFA salt as a cream white solid. ^1H NMR (400 MHz, CD_3OD) δ 8.02 (t, $J = 1.8$ Hz, 1H), 7.82 (m, 1H), 7.71 (m, 1H), 7.40 (t, $J = 7.9$ Hz, 1H), 3.58 – 3.40 (m, 3H), 2.95 (ddd, $J = 15.2, 13.4, 6.9$ Hz, 1H), 2.83 (s, 3H), 2.22 – 2.07 (m, 1H), 2.07 – 1.92 (m, 1H), 1.43 (s, 9H). ^{13}C NMR (101 MHz, CD_3OD) δ 169.0, 137.2, 135.8, 131.5, 131.4, 127.1, 123.6, 65.3, 50.0, 38.0, 35.1, 27.0, 24.6. LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.4$ min. HRMS calculated for $\text{C}_{15}\text{H}_{22}\text{N}_2\text{OBr} - \text{H}$: 325.0916; Found: 325.0921 $[\text{M}-\text{H}]^-$.



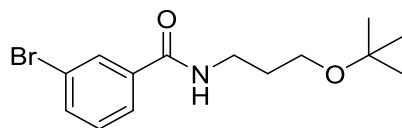
3-bromo-N-(3-(tert-butyl(ethyl)amino)propyl)benzamide (65): To a solution of **61** (130 mg, 0.42 mmol) in 4 mL of MeOH was added 102 μL (1.8 mmol) acetaldehyde 156 mg (2.5 mmol) sodium cyanoborohydride, and 120 μL (2.1 mmol) acetic acid, and the mixture was stirred overnight at rt. The reaction was monitored by LC-MS until completion and then dried to residue. The crude material was dissolved in 1 mL of DCM on silica and purified by reverse phase column chromatography using the Teledyne Isco automated column system to afford 15.3 mg (10.8 % yield) as the TFA salt as a clear yellow oil. ^1H NMR (400 MHz, cd_3od) δ 8.01 (t, $J = 1.8$ Hz, 1H), 7.84 – 7.77 (m, 1H), 7.72 (m, 1H), 7.41 (t, $J = 7.9$ Hz, 1H), 3.47 (m, 5H), 3.16 – 2.98 (m, 3H), 2.22 – 1.95 (m, 3H), 1.47 – 1.34 (m, 13H), 0.00 (t, $J = 7.9, 7.9$ Hz, 1H),

0.00 (s, 9H), 7.83 – 7.79 (m, 1H), 2.09 – 1.95 (m, 2H), 3.15 – 3.02 (m, 2H), 2.21 – 2.09 (m, 1H).

LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.8$ min. MS (ESI+): 342.3 $[M+H]^+$.

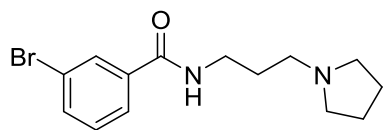


3-bromo-N-(3-(isopropyl(methyl)amino)propyl)benzamide (66): To a solution of **63** (36 mg, 0.12 mmol) in 4 mL of MeOH was added 15 μ L (0.67 mmol) formaldehyde solution (37% wt in H₂O), 45 mg (0.72 mmol) sodium cyanoborohydride, and 35 μ L (0.60 mmol) acetic acid, and the mixture was stirred overnight at rt. The reaction was monitored by LC-MS until completion and then dried to residue. The crude material was dissolved in 1 mL of DCM on silica and purified by reverse phase column chromatography using the Teledyne Isco automated column system to afford 35 mg (94.5% yield) as the TFA salt as a clear oil. ¹H NMR (400 MHz, cd₃od) δ 8.01 (t, $J = 1.8$ Hz, 1H), 7.84 – 7.79 (m, 1H), 7.70 (m, 1H), 7.39 (t, $J = 7.9$ Hz, 1H), 3.50 (t, $J = 6.5$ Hz, 2H), 3.40 – 3.32 (m, 1H), 3.06 (t, $J = 7.4$ Hz, 2H), 2.04 – 1.93 (m, 2H), 1.34 (d, $J = 6.6$ Hz, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.8$ min. MS (ESI+): 316.1 $[M+Na]^+$.

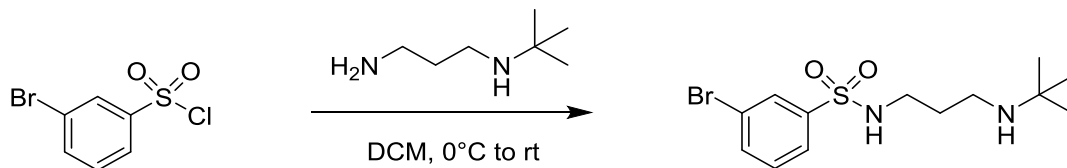


3-Bromo-N-(3-(tert-butoxy)propyl)benzamide (67): Compound **67** was prepared from 3-bromobenzoic acid (0.10 g, 0.50 mmol) and 3-(tert-butoxy)propan-1-amine HCl (0.10 g, 0.60 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (146 mg, 69%) after purification as the TFA salt. ¹H NMR (400 MHz, CD₃OD) δ 7.96 (t, $J = 1.8$ Hz, 1H), 7.77 (m, 1H), 7.67 (m, 1H), 7.37 (t, $J = 7.9$ Hz, 1H), 3.49 (t, $J = 6.4$ Hz, 2H), 3.45 (t, J

= 7.3 Hz, 2H), 1.87 – 1.78 (m, 2H), 1.19 (s, 9H). LC-MS (λ = 254 nm): 99%, t_R = 5.8 min. MS (ESI+): 336.1 + 338.1 [M+Na]⁺.

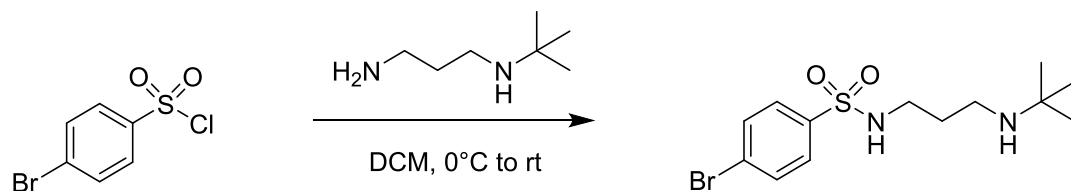


3-Bromo-N-(3-(pyrrolidin-1-yl)propyl)benzamide (68): Compound **68** was prepared from 3-bromobenzoic acid (0.5 g, 2.5 mmol) and 3-(pyrrolidin-1-yl)propan-1-amine (0.38 g, 3.0 mmol) by the same procedure as compound **61**. The product was obtained as a yellow oil (657 mg, 85%) after purification as the TFA salt. ¹H NMR (400 MHz, CD₃OD) δ 7.97 (t, J = 1.8 Hz, 1H), 7.81 – 7.75 (m, 1H), 7.69 (ddd, J = 8.0, 2.0, 1.0 Hz, 1H), 7.39 (t, J = 7.9 Hz, 1H), 3.43 (t, J = 6.9 Hz, 2H), 2.68 – 2.58 (m, 6H), 1.84 (m, 6H). LC-MS (λ = 254 nm): 99%, t_R = 2.7 min. MS (ESI+): 311.1 + 313.1 [M+H]⁺.

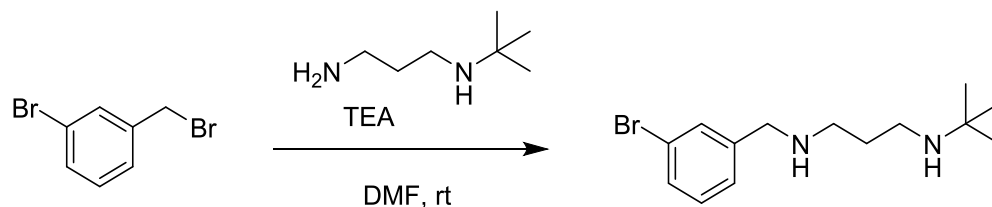


3-Bromo-N-(3-(*tert*-butylamino)propyl)benzenesulfonamide (69): To a solution of 3-bromobenzenesulfonyl chloride (0.12 g, 0.47 mmol) in dichloromethane at 0°C was slowly added *N*-(*tert*-butyl)propane-1,3-diamine (0.063 g, 0.48 mmol). The reaction was stirred at 0°C for 30 min and then allowed to warm to room temperature. The reaction was monitored by TLC until it had gone to completion, and then diluted with MeOH, concentrated down, re-suspended in methanol and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a white solid (147 mg, 68%) after purification as the TFA salt. ¹H NMR (400 MHz,

CD₃OD) δ 8.00 (s, 1H), 7.81 (m, 2H), 7.51 (t, J = 7.9 Hz, 1H), 3.02 (dd, J = 10.8, 5.1 Hz, 4H), 1.93 – 1.79 (m, 2H), 1.35 (s, 9H). LC-MS (λ = 254 nm): 99%, t_R = 3.6 min. MS (ESI⁺): 349.1 + 351.10 [M+H]⁺.

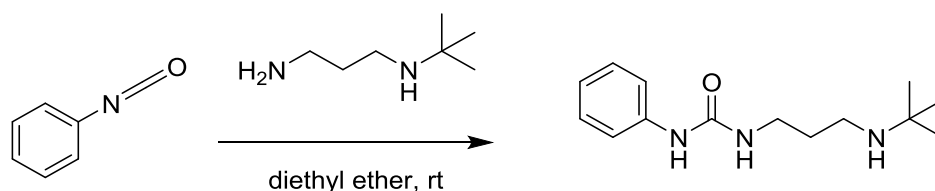


4-bromo-N-(3-(tert-butylamino)propyl)benzenesulfonamide (70): To a solution of 3-bromobenzenesulfonyl chloride (0.12 g, 0.47 mmol) in dichloromethane at 0°C was slowly added *N*-(*tert*-butyl)propane-1,3-diamine (0.063 g, 0.48 mmol). The reaction was stirred at 0°C for 30 min and then allowed to warm to room temperature. The reaction was monitored by TLC until it had gone to completion, and then diluted with MeOH, concentrated down, re-suspended in methanol and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a white solid (320.7 mg, 91.4%) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.85 – 7.66 (m, 4H), 3.00 (t, J = 6.7 Hz, 4H), 1.92 – 1.79 (m, 2H), 1.34 (s, 9H). LC-MS (λ = 254 nm): 99%, t_R = 3.7 min. MS (ESI⁺): 349.1 + 351.10 [M+H]⁺.

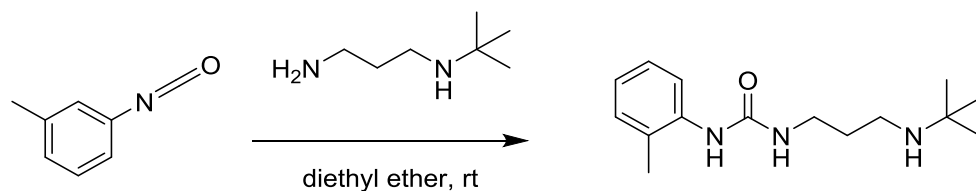


N¹-(3-Bromobenzyl)-N³-(*tert*-butyl)propane-1,3-diamine (71): To a solution of 1-bromo-3-(bromomethyl)benzene (0.23 g, 0.92 mmol) and triethylamine (0.233 g, 2.3 mmol) in 2 mL dimethylformamide at room temperature was slowly added *N*-(*tert*-butyl)propane-1,3-diamine (0.1 g, 0.77 mmol) and the reaction was stirred overnight. Once the reaction had gone to

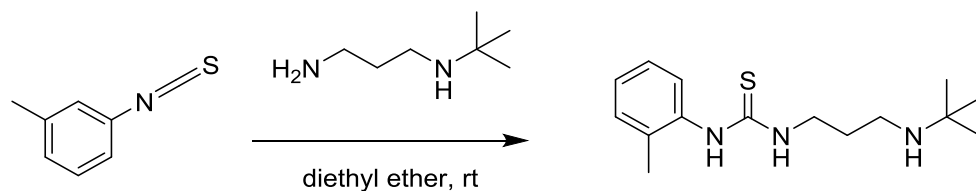
completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear yellow oil (37 mg, 11%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.73 (s, 1H), 7.62 (d, $J = 8.0$ Hz, 1H), 7.49 (d, $J = 7.7$ Hz, 1H), 7.38 (t, $J = 7.8$ Hz, 1H), 4.23 (s, 2H), 3.20 (t, $J = 7.6$ Hz, 2H), 3.10 (dd, $J = 9.2, 8.3$ Hz, 2H), 2.12 (dt, $J = 15.5, 7.9$ Hz, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 1.8$ min. MS (ESI $^+$): 299.2 + 301.2 $[\text{M}+\text{H}]^+$.



1-(3-(tert-butylamino)propyl)-3-phenylurea (72): To a solution of *N*-(*tert*-butyl)propane-1,3-diamine (0.131 g, 1.0 mmol) in 4 mL diethyl ether was added isocyanatobenzene (0.1 g, 0.84 mmol) drop-wise and the reaction was stirred for 4 hr. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear oil (289 mg, 91.3%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.39 (m, 2H), 7.24 (m, 2H), 6.97 (m, 1H), 3.31 (m, 2H), 2.99 (t, $J = 7.2$ Hz, 2H), 1.88 (m, 2H), 1.34 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.5$ min. MS (ESI $^+$): 251.2 $[\text{M}+\text{H}]^+$.

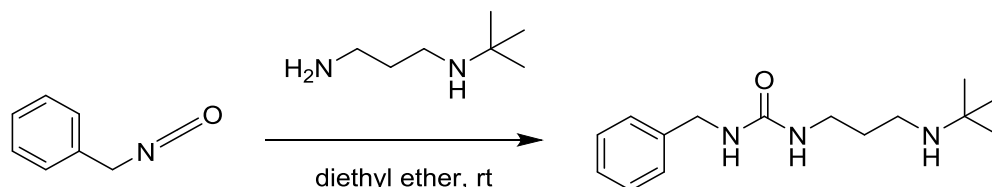


1-(3-(tert-butylamino)propyl)-3-(o-tolyl)urea (73): To a solution of *N*-(*tert*-butyl)propane-1,3-diamine (0.117 g, 0.89 mmol) in 4 mL diethyl ether was added 1-isocyanato-3-methylbenzene (0.1 g, 0.75 mmol) drop-wise and the reaction was stirred for 4 hr. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear oil (276 mg, 48.5%) after purification as the TFA salt. ^1H NMR (400 MHz, CDCl_3) δ 7.42 (dd, $J = 7.9, 0.9$ Hz, 1H), 7.11 (m, 2H), 6.98 (td, $J = 7.4, 1.2$ Hz, 1H), 3.27 (m, 2H), 2.96 (t, $J = 7.3$ Hz, 2H), 2.20 (s, 3H), 1.83 (m, 2H), 1.30 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.4$ min. MS (ESI $^+$): 265.3 $[\text{M}+\text{H}]^+$.

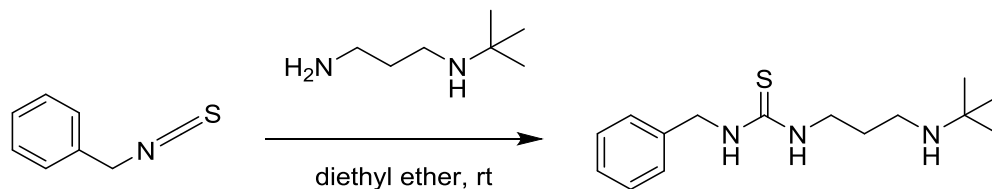


1-(3-(tert-butylamino)propyl)-3-(o-tolyl)thiourea (74): To a solution of *N*-(*tert*-butyl)propane-1,3-diamine (0.104 g, 0.79 mmol) in 4 mL diethyl ether was added 1-isothiocyanto-3-methylbenzene (0.1 g, 0.67 mmol) drop-wise and the reaction was stirred for 4 hr. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear oil (254 mg, 96.5%) after

purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.26 (t, $J = 7.7$ Hz, 1H), 7.11 (d, $J = 8.7$ Hz, 2H), 7.05 (d, $J = 7.6$ Hz, 1H), 3.75 (t, $J = 6.4$ Hz, 2H), 3.02 (t, $J = 7.3$ Hz, 2H), 2.34 (s, 3H), 1.98 (m, 2H), 1.38 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.5$ min. MS (ESI $^{+}$): 281.2 $[\text{M}+\text{H}]^{+}$.

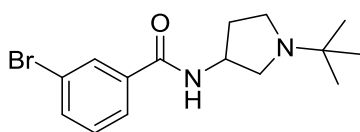


1-benzyl-3-(3-(*tert*-butylamino)propyl)urea (75): To a solution of *N*-(*tert*-butyl)propane-1,3-diamine (0.117 g, 0.89 mmol) in 4 mL diethyl ether was added (isocyanatomethyl)benzene (0.1 g, 0.75 mmol) drop-wise and the reaction was stirred for 4 hr. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear oil (259 mg, 82.4%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.26 (m, 5H), 4.31 (s, 2H), 3.27 (m, 2H), 2.93 (t, $J = 7.0$ Hz, 2H), 1.83 (m, 2H), 1.28 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.4$ min. MS (ESI $^{+}$): 265.2 $[\text{M}+\text{H}]^{+}$.

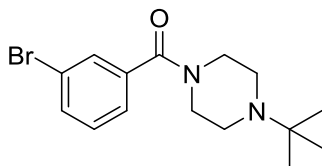


1-benzyl-3-(3-(*tert*-butylamino)propyl)thiourea (76): To a solution of *N*-(*tert*-butyl)propane-1,3-diamine (0.105 g, 0.81 mmol) in 4 mL diethyl ether was added (isothiocyatomethyl)benzene (0.1 g, 0.67 mmol) drop-wise and the reaction was stirred for 4

hr. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear oil (235 mg, 88.7%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.27 (m, 5H), 4.70 (s, 2H), 3.72 (t, $J = 5.5$ Hz, 2H), 2.96 (s, 2H), 1.94 (m, 2H), 1.33 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.8$ min. MS (ESI $^+$): 281.2 $[\text{M}+\text{H}]^+$.

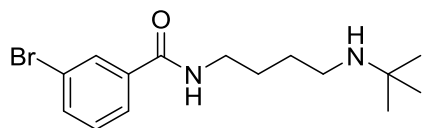


(3-Bromophenyl)(3-(*tert*-butyl)imidazolidin-1-yl)methanone (77): Compound **77** was prepared from 3-bromobenzoic acid (0.10 g, 0.50 mmol) and 1-(*tert*-butyl)pyrrolidin-3-amine (0.09 g, 0.63 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (158 mg, 73%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.03 (d, $J = 1.6$ Hz, 1H), 7.86 – 7.80 (m, 1H), 7.70 (d, $J = 7.9$ Hz, 1H), 7.39 (m, 1H), 4.73 – 4.58 (m, 1H), 3.82 (m, 1H), 3.73 – 3.61 (m, 1H), 3.54 (m, 1H), 3.46 – 3.25 (m, 1H), 2.58 – 2.36 (m, 1H), 2.22 (m, 1H), 1.42 (s, 9H). LC-MS ($\lambda = 254$ nm): 97%, $t_R = 3.7$ min. MS (ESI $^+$): 325.1 + 327.1 $[\text{M}+\text{H}]^+$.

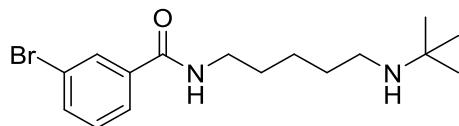


(3-Bromophenyl)(4-(*tert*-butyl)piperazin-1-yl)methanone (78): Compound **78** was prepared from 3-bromobenzoic acid (0.10 g, 0.50 mmol) and 1-(*tert*-butyl)piperazine (0.09 g, 0.63 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (201 mg,

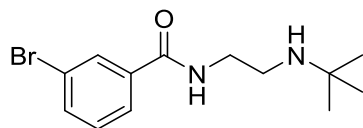
92%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.70 (m, 1H), 7.67 (m, 1H), 7.47 (dt, $J = 7.7, 1.3$ Hz, 1H), 7.40 (m, 1H), 4.71 (m, 1H), 4.23 – 3.25 (m, 5H), 3.17 (m, 2H), 1.43 (s, 9H). LC-MS ($\lambda = 254$ nm): 95%, $t_{\text{R}} = 3.4$ min. MS (ESI $^{+}$): 325.1 + 327.1 $[\text{M}+\text{H}]^{+}$.



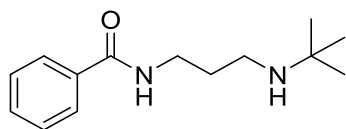
3-Bromo-N-(4-(*tert*-butylamino)butyl)benzamide (79): Compound **79** was prepared from 3-bromobenzoic acid (40 mg, 0.19 mmol) and **136** (34 mg, 0.091 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (23 mg, 57%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.99 (t, $J = 1.7$ Hz, 1H), 7.80 (m, 1H), 7.71 (m, 1H), 7.40 (t, $J = 7.9$ Hz, 1H), 3.45 (t, $J = 6.4$ Hz, 2H), 3.07 – 2.99 (m, 2H), 1.79 – 1.67 (m, 4H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 95%, $t_{\text{R}} = 4.2$ min. MS (ESI $^{+}$): 327.1 + 329.1 $[\text{M}+\text{H}]^{+}$.



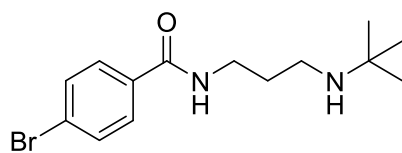
3-Bromo-N-(5-(*tert*-butylamino)pentyl)benzamide (80): Compound **80** was prepared from 3-bromobenzoic acid (0.075 g, 0.37 mmol) and **139** (0.071 g, 0.18 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (47 mg, 57%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.99 (t, $J = 1.8$ Hz, 1H), 7.82 – 7.75 (m, 1H), 7.70 (m, 1H), 7.40 (t, $J = 7.9$ Hz, 1H), 3.41 (t, $J = 7.0$ Hz, 2H), 3.01 – 2.94 (m, 2H), 1.71 (m, 4H), 1.54 – 1.43 (m, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 4.3$ min. MS (ESI $^{+}$): 341.3 + 343.1 $[\text{M}+\text{H}]^{+}$.



3-Bromo-*N*-(2-(*tert*-butylamino)ethyl)benzamide (81): Compound **81** was prepared from 3-bromobenzoic acid (0.10 g, 0.50 mmol) and *N*-(*tert*-butyl)ethane-1,2-diamine (0.07 g, 0.59 mmol) by the same procedure as compound **61**. The product was obtained as a white solid (119 mg, 58%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.05 (m, 1H), 7.85 (m, 1H), 7.74 – 7.68 (m, 1H), 7.40 (t, $J = 7.9$ Hz, 1H), 3.70 (t, $J = 6.2$ Hz, 2H), 3.24 (t, $J = 6.2$ Hz, 2H), 1.39 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.8$ min. MS (ESI $^+$): 299.1 + 301.1 $[\text{M}+\text{H}]^+$.

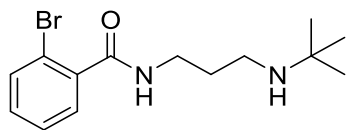


***N*-(3-(*Tert*-butylamino)propyl)benzamide (82):** Compound **82** was prepared from benzoic acid (0.10 g, 0.82 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.13 g, 0.98 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (192 mg, 67%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.85 (m, 2H), 7.59 – 7.54 (m, 1H), 7.51 – 7.45 (m, 2H), 3.53 (t, $J = 6.5$ Hz, 2H), 3.04 (t, $J = 7.3$ Hz, 2H), 2.05 – 1.92 (m, 2H), 1.39 (s, 9H). LC-MS ($\lambda = 254$ nm): 96%, $t_R = 3.4$ min. MS (ESI $^+$): 235.2 $[\text{M}+\text{H}]^+$.

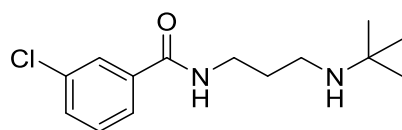


4-Bromo-*N*-(3-(*tert*-butylamino)propyl)benzamide (83): Compound **83** was prepared from 4-bromobenzoic acid (0.10 g, 0.50 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.10 g, 0.74 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (120 mg, 56%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.76 (d, $J = 8.6$ Hz, 2H), 7.63 (d, $J = 8.6$ Hz, 2H), 3.50 (t, $J = 6.5$ Hz, 2H), 3.03 (t, $J = 7.4$ Hz, 2H), 2.04 – 1.93

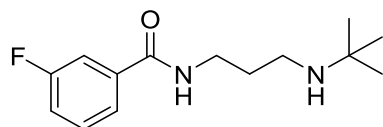
(m, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.7$ min. MS (ESI⁺): 313.1 + 315.1 [M+H]⁺.



2-Bromo-N-(3-(*tert*-butylamino)propyl)benzamide (84): Compound **84** was prepared from 2-bromobenzoic acid (0.10 g, 0.50 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.08 g, 0.60 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (191 mg, 55%) after purification as the TFA salt. ¹H NMR (400 MHz, CD₃OD) δ 7.64 (m, 1H), 7.45 – 7.40 (m, 2H), 7.36 (m, 1H), 3.50 (t, $J = 6.2$ Hz, 2H), 3.17 – 3.07 (m, 2H), 2.06 – 1.95 (m, 2H), 1.38 (s, 9H). LC-MS ($\lambda = 254$ nm): 95%, $t_R = 3.3$ min. MS (ESI⁺): 313.1 + 315.1 [M+H]⁺.

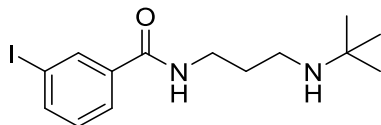


N-(3-(*Tert*-butylamino)propyl)-3-chlorobenzamide (85): Compound **85** was prepared from 3-chlorobenzoic acid (0.10 g, 0.64 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.10 g, 0.76 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (115 mg, 47%) after purification as the TFA salt. ¹H NMR (400 MHz, CD₃OD) δ 7.86 (m, 1H), 7.77 (m, 1H), 7.55 – 7.51 (m, 1H), 7.44 (m, 1H), 3.51 (t, $J = 6.5$ Hz, 2H), 3.03 (t, $J = 7.4$ Hz, 2H), 2.05 – 1.94 (m, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.4$ min. MS (ESI⁺): 269.2 [M+H]⁺.

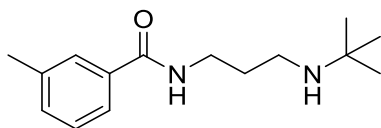


N-(3-(*Tert*-butylamino)propyl)-3-fluorobenzamide (86): Compound **86** was prepared from 3-fluorobenzoic acid (0.10 g, 0.71 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.11 g, 0.85

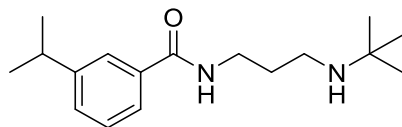
mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (68 mg, 26%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.68 (d, $J = 7.8$ Hz, 1H), 7.61 – 7.56 (m, 1H), 7.49 (td, $J = 8.0, 5.7$ Hz, 1H), 7.29 (td, $J = 8.4, 2.5$ Hz, 1H), 3.52 (t, $J = 6.5$ Hz, 2H), 3.04 (t, $J = 7.4$ Hz, 2H), 2.06 – 1.94 (m, 2H), 1.38 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.1$ min. MS (ESI $^{+}$): 253.2 $[\text{M}+\text{H}]^{+}$.



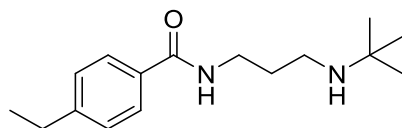
***N*-(3-(*Tert*-butylamino)propyl)-3-iodobenzamide (87):** Compound **87** was prepared from 3-iodobenzoic acid (0.10 g, 0.40 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.06 g, 0.99 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (90 mg, 47%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.20 (t, $J = 1.6$ Hz, 1H), 7.91 – 7.81 (m, 2H), 7.23 (t, $J = 7.8$ Hz, 1H), 3.50 (t, $J = 6.5$ Hz, 2H), 3.03 (t, $J = 7.4$ Hz, 2H), 2.04 – 1.93 (m, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.8$ min. MS (ESI $^{+}$): 361.1 $[\text{M}+\text{H}]^{+}$.



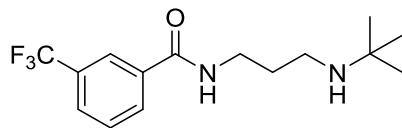
***N*-(3-(*tert*-butylamino)propyl)-3-methylbenzamide (86):** Compound **86** was prepared from 3-methylbenzoic acid (0.10 g, 0.73 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.114 g, 0.88 mmol) by the same procedure as compound **61**. The product was obtained as a pale yellow solid (216 mg, 57.7%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.69 – 7.66 (m, 1H), 7.64 (m, 1H), 7.37 – 7.30 (m, 2H), 3.51 (t, $J = 6.5$ Hz, 2H), 3.03 (t, $J = 7.4$ Hz, 2H), 2.38 (s, 3H), 2.00 (dd, $J = 7.3, 6.3$ Hz, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.7$ min. MS (ESI $^{+}$): 250.2 $[\text{M}+\text{H}]^{+}$.



***N*-(3-(*Tert*-butylamino)propyl)-3-isopropylbenzamide (89):** Compound **89** was prepared from 3-isopropylbenzoic acid (95 mg, 0.58 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.1 g, 0.73 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (171 mg, 76%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.75 (t, J = 1.8 Hz, 1H), 7.67 (dt, J = 7.4, 1.6 Hz, 1H), 7.42 (m, 1H), 7.37 (t, J = 7.6 Hz, 1H), 3.53 (t, J = 6.5 Hz, 2H), 3.04 (t, J = 7.3 Hz, 2H), 2.95 (m, 1H), 2.07 – 1.95 (m, 2H), 1.38 (s, 9H), 1.26 (d, J = 6.9 Hz, 6H). LC-MS (λ = 254 nm): 99%, t_{R} = 4.2 min. MS (ESI $^{+}$): 277.2 $[\text{M}+\text{H}]^{+}$.

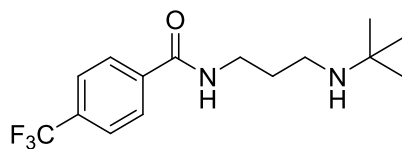


***N*-(3-(*tert*-butylamino)propyl)-4-ethylbenzamide (90):** Compound **90** was prepared from 4-ethylbenzoic acid (0.104 g, 0.69 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.108 g, 0.83 mmol) by the same procedure as compound **61**. The product was obtained as a yellow solid (230 mg, 97.1%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.78 (m, 2H), 7.29 (d, J = 8.4 Hz, 2H), 3.52 (t, J = 6.4 Hz, 2H), 3.02 (t, J = 7.3 Hz, 2H), 2.68 (q, J = 7.6 Hz, 2H), 2.00 (m, 2H), 1.37 (s, 9H), 1.23 (t, J = 7.6 Hz, 3H). LC-MS (λ = 254 nm): 99%, t_{R} = 3.9 min. MS (ESI $^{+}$): 264.2 $[\text{M}+\text{H}]^{+}$.

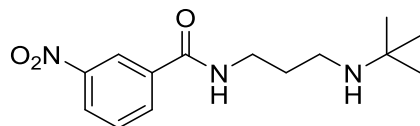


***N*-(3-(*Tert*-butylamino)propyl)-3-(trifluoromethyl)benzamide (91):** Compound **91** was prepared from 3-(trifluoromethyl)benzoic acid (0.10 g, 0.48 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.08 g, 0.63 mmol) by the same procedure as compound **61**. The product was

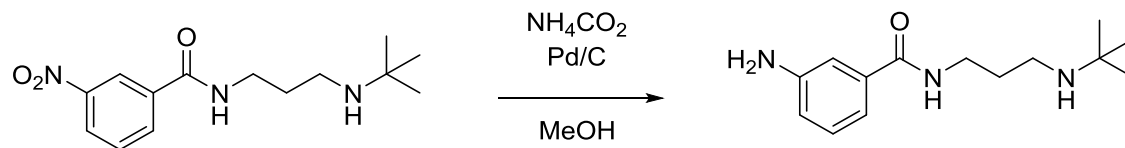
obtained as a clear oil (193 mg, 97%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.17 (s, 1H), 8.12 (d, $J = 7.9$ Hz, 1H), 7.85 (d, $J = 7.5$ Hz, 1H), 7.68 (t, $J = 7.8$ Hz, 1H), 3.55 (t, $J = 6.5$ Hz, 2H), 3.06 (t, $J = 7.5$ Hz, 2H), 2.02 (m, 2H), 1.40 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 4.1$ min. MS (ESI $^{+}$): 303.2 $[\text{M}+\text{H}]^{+}$.



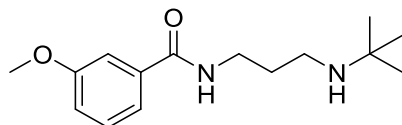
***N*-(3-(*Tert*-butylamino)propyl)-4-(trifluoromethyl)benzamide (92):** Compound **92** was prepared from 4-(trifluoromethyl)benzoic acid (0.10 g, 0.48 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.09 g, 0.72 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (128 mg, 64%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.03 (d, $J = 8.1$ Hz, 2H), 7.78 (d, $J = 8.2$ Hz, 2H), 3.54 (t, $J = 6.6$ Hz, 2H), 3.06 (t, $J = 7.5$ Hz, 2H), 2.02 (m, 2H), 1.39 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 4.1$ min. MS (ESI $^{+}$): 303.2 $[\text{M}+\text{H}]^{+}$.



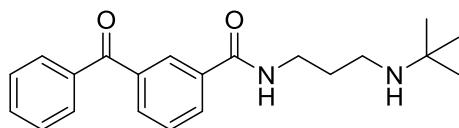
***N*-(3-(*Tert*-butylamino)propyl)-3-nitrobenzamide (93):** Compound **93** was prepared from 3-nitrobenzoic acid (0.10 g, 0.6 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.09 g, 0.73 mmol) by the same procedure as compound **61**. The product was obtained as a yellow solid (43 mg, 18%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.71 (t, $J = 2.0$ Hz, 1H), 8.41 (m, 1H), 8.25 (m, 1H), 7.75 (t, $J = 8.0$ Hz, 1H), 3.56 (t, $J = 6.6$ Hz, 2H), 3.11 – 3.04 (m, 2H), 2.02 (m, 2H), 1.39 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.5$ min. MS (ESI $^{+}$): 280.2 $[\text{M}+\text{H}]^{+}$.



3-amino-N-(3-(tert-butylamino)propyl)benzamide (94): To a solution of **93** (129 mg, -46 mmol) in 10 mL MeOH was added 0.146 g (2.3 mmol) ammonium formate and ~0.02 g of Pd/carbon. The reaction was stirred overnight at rt, then filtered and purified at compound **61**. ^1H NMR (400 MHz, cd_3od) δ 7.88 (m, 2H), 7.61 (m, 1H), 7.54 (m, 1H), 3.54 (t, $J = 6.5$ Hz, 2H), 3.05 (t, $J = 7.5$ Hz, 2H), 2.01 (m, 2H), 1.38 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 2.7$ min. MS (ESI $^{+}$): 251.2 $[\text{M}+\text{H}]^{+}$.

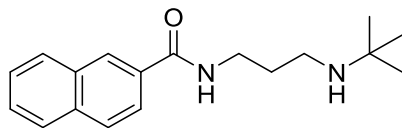


N-(3-(tert-butylamino)propyl)-3-methoxybenzamide (95): Compound **95** was prepared from 3-methoxybenzoic acid (0.10 g, 0.65 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.103 g, 0.79 mmol) by the same procedure as compound **61**. The product was obtained as a pale yellow oil (184 mg, 61.3%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.41 (m, 2H), 7.37 (dd, $J = 11.7, 4.0$ Hz, 1H), 7.09 (m, 1H), 3.82 (s, 3H), 3.51 (t, $J = 6.5$ Hz, 2H), 3.03 (t, $J = 7.4$ Hz, 2H), 2.05 – 1.95 (m, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.6$ min. MS (ESI $^{+}$): 266.2 $[\text{M}+\text{H}]^{+}$.

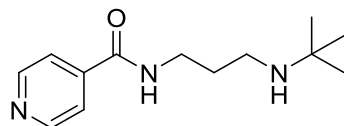


3-benzoyl-N-(3-(tert-butylamino)propyl)benzamide (96): Compound **96** was prepared from 3-benzoylbenzoic acid (0.10 g, 0.43 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.069 g, 0.52 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (72 mg,

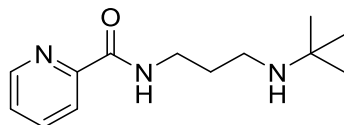
34%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.17 (s, 1H), 8.12 (d, J = 7.9 Hz, 1H), 7.85 (d, J = 7.5 Hz, 1H), 7.68 (t, J = 7.8 Hz, 1H), 3.55 (t, J = 6.5 Hz, 2H), 3.06 (t, J = 7.5 Hz, 2H), 2.02 (m, 2H), 1.40 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 4.2 min. MS (ESI⁺): 340.2 $[\text{M}+\text{H}]^+$.



***N*-(3-(tert-butylamino)propyl)-2-naphthamide (97):** Compound **97** was prepared from 2-naphthoic acid (0.10 g, 0.58 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.091 g, 0.69 mmol) by the same procedure as compound **61**. The product was obtained as a yellow solid (232 mg, 98.7%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.40 (d, J = 0.8 Hz, 1H), 7.89 (m, 4H), 7.54 (m, 2H), 3.57 (t, J = 6.5 Hz, 2H), 3.04 (t, J = 7.3 Hz, 2H), 2.03 (m, 2H), 1.36 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 4.2 min. MS (ESI⁺): 286.2 $[\text{M}+\text{H}]^+$.

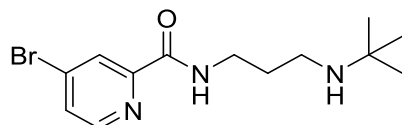


***N*-(3-(tert-butylamino)propyl)picolinamide (98):** Compound **98** was prepared from nicotinic acid (0.10 g, 0.81 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.16 g, 0.97 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (68 mg, 35.6%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.93 (d, J = 6.0 Hz, 2H), 8.24 (d, J = 6.3 Hz, 2H), 3.56 (t, J = 6.6 Hz, 2H), 3.12 – 3.04 (m, 2H), 2.08 – 1.96 (m, 2H), 1.37 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 2.2 min. MS (ESI⁺): 237.2 $[\text{M}+\text{H}]^+$.

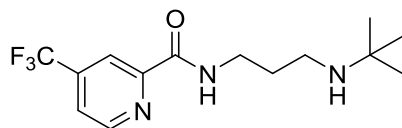


***N*-(3-(tert-butylamino)propyl)picolinamide (99):** Compound **99** was prepared from picolinic acid (0.10 g, 0.81 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.127 g, 0.97 mmol) by the

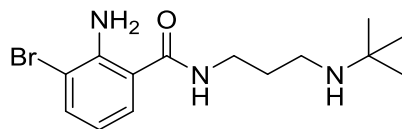
same procedure as compound **61**. The product was obtained as a clear yellow oil (232 mg, 65.7%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.64 (m, 1H), 8.11 (d, J = 7.8 Hz, 1H), 7.98 (m, 1H), 7.57 (m, 1H), 3.56 (t, J = 6.5 Hz, 2H), 3.05 (t, J = 7.5 Hz, 2H), 2.02 (td, J = 13.2, 6.9 Hz, 2H), 1.37 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 2.9 min. MS (ESI $^{+}$): 237.2 $[\text{M}+\text{H}]^{+}$.



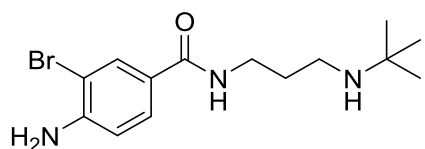
4-Bromo-*N*-(3-(*tert*-butylamino)propyl)picolinamide (100): Compound **100** was prepared from 4-bromopicolinic acid (0.13 g, 0.65 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.1 g, 0.78 mmol) by the same procedure as compound **61**. The product was obtained as a yellow solid (43 mg, 14%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.92 (d, J = 5.4 Hz, 1H), 8.34 (s, 1H), 7.89 (d, J = 4.9 Hz, 1H), 3.58 (t, J = 6.5 Hz, 2H), 3.04 (t, J = 7.5 Hz, 2H), 2.02 (m, 2H), 1.38 (s, 9H). LC-MS (λ = 254 nm): 98%, t_{R} = 3.6 min. MS (ESI $^{+}$): 314.1 + 316.1 $[\text{M}+\text{H}]^{+}$.



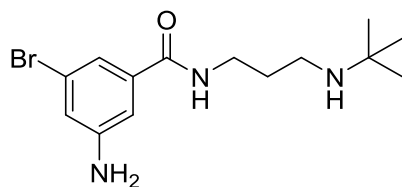
***N*-(3-(*Tert*-butylamino)propyl)-4-(trifluoromethyl)picolinamide (101):** Compound **101** was prepared from 4-(trifluoromethyl)picolinic acid (0.10 g, 0.52 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.08 g, 0.63 mmol) by the same procedure as compound **61**. The product was obtained as a yellow solid (40 mg, 18%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.91 (d, J = 5.0 Hz, 1H), 8.33 (s, 1H), 7.89 (d, J = 5.0, 1H), 3.58 (t, J = 6.5 Hz, 2H), 3.11 – 3.02 (m, 2H), 2.07 – 1.96 (m, 2H), 1.38 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 3.9 min. MS (ESI $^{+}$): 304.2 $[\text{M}+\text{H}]^{+}$.



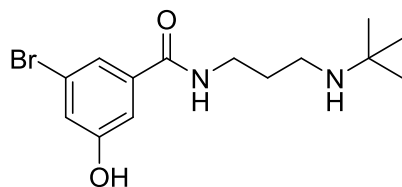
2-amino-3-bromo-*N*-(3-(tert-butylamino)propyl)benzamide (102): Compound **102** was prepared from 2-amino-3-bromobenzoic acid (0.15 g, 0.69 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.108 g, 0.83 mmol) by the same procedure as compound **61**. The product was obtained as a clear red oil (53 mg, 17.4%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 0.00 (t, $J = 7.4$, 7.4 Hz, 2H), 0.00 (d, $J = 6.2$ Hz, 9H), 0.00 (t, $J = 6.5$, 6.5 Hz, 2H), 2.02 – 1.92 (m, 2H), 0.00 (dd, $J = 7.9$, 1.4 Hz, 1H), 0.00 (dd, $J = 7.8$, 1.4 Hz, 1H), 0.00 (t, $J = 7.9$, 7.9 Hz, 1H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.8$ min. MS (ESI $^+$): 229.2 $[\text{M}+\text{H}]^+$.



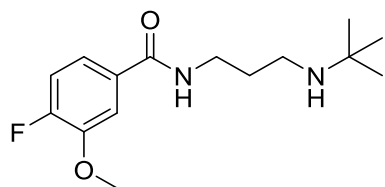
4-amino-3-bromo-*N*-(3-(tert-butylamino)propyl)benzamide (103): Compound **103** was prepared from 4-amino-3-bromobenzoic acid (0.1 g, 0.46 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.072 g, 0.55 mmol) by the same procedure as compound **61**. The product was obtained as a clear orange oil (37 mg, 18.1 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.93 (d, $J = 2.1$ Hz, 1H), 7.61 (dd, $J = 8.5$, 2.1 Hz, 1H), 6.82 (d, $J = 8.5$ Hz, 1H), 3.47 (t, $J = 6.4$ Hz, 2H), 3.00 (t, $J = 7.2$ Hz, 2H), 2.01 – 1.90 (m, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.3$ min. MS (ESI $^+$): 329.2 $[\text{M}+\text{H}]^+$.



3-amino-5-bromo-*N*-(3-(*tert*-butylamino)propyl)benzamide (104): Compound **104** was prepared from 3-amino-5-bromobenzoic acid (0.4 g, 1.8 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.352 g, 2.7 mmol) by the same procedure as compound **61**. The product was obtained as a tan oil (260 mg, 42.7 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.13 (t, J = 1.5 Hz, 1H), 7.01 (m, 1H), 6.96 (t, J = 1.9 Hz, 1H), 3.41 (t, J = 6.7 Hz, 2H), 2.61 (t, J = 7.2 Hz, 2H), 1.76 (m, 2H), 1.12 (s, 9H). LC-MS (λ = 254 nm): 99%, t_R = 3.3 min. MS (ESI $^+$): 331.0 $[\text{M}+\text{H}]^+$.

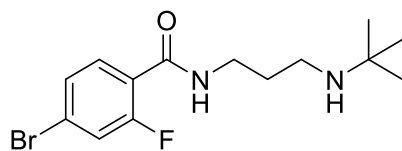


3-bromo-*N*-(3-(*tert*-butylamino)propyl)-5-hydroxybenzamide (105): Compound **105** was prepared from 3-bromo-5-hydroxybenzoic acid (0.15 g, 0.69 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.108 g, 0.83 mmol) by the same procedure as compound **61**. The product was obtained as a white solid (86.2 mg, 42.2 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.44 – 7.42 (m, 1H), 7.24 – 7.21 (m, 1H), 7.14 – 7.11 (m, 1H), 3.48 (t, J = 7.1 Hz, 2H), 3.02 (t, J = 8.3 Hz, 2H), 1.97 (m, 2H), 1.37 (s, 9H). LC-MS (λ = 254 nm): 99%, t_R = 3.6 min. MS (ESI $^+$): 332.2 $[\text{M}+\text{H}]^+$.

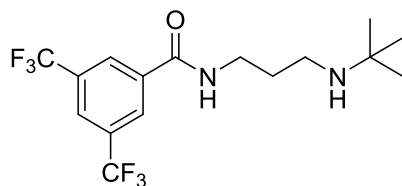


***N*-(3-(*tert*-butylamino)propyl)-4-fluoro-3-methoxybenzamide (106):** Compound **106** was prepared from 4-fluoro-3-methoxybenzoic acid (0.1 g, 0.58 mmol) and *N*-(*tert*-butyl)propane-

1,3-diamine (0.092 g, 0.71 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (87 mg, 37.3 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.95 (dd, $J = 8.7, 6.9$ Hz, 1H), 6.95 (dd, $J = 11.0, 2.4$ Hz, 1H), 6.80 (m, 1H), 3.97 (s, 3H), 3.54 (m, 2H), 3.03 (t, $J = 7.2$ Hz, 2H), 1.99 (m, 2H), 1.38 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.6$ min. MS (ESI $^+$): 284.2 $[\text{M}+\text{H}]^+$.

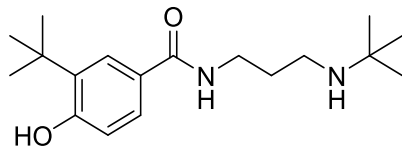


4-bromo-*N*-(3-(tert-butylamino)propyl)-2-fluorobenzamide (107): Compound **107** was prepared from 4-bromo-2-fluorobenzoic acid (0.1 g, 0.46 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.071 g, 0.54 mmol) by the same procedure as compound **61**. The product was obtained as a pale yellow solid (183 mg, 89.2 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.64 (dd, $J = 11.1, 5.3$ Hz, 1H), 7.47 (m, 2H), 3.52 (t, $J = 6.5$ Hz, 2H), 3.06 (m, 2H), 2.00 (tt, $J = 13.3, 6.8$ Hz, 2H), 1.38 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.8$ min. MS (ESI $^+$): 334.1 $[\text{M}+\text{H}]^+$.



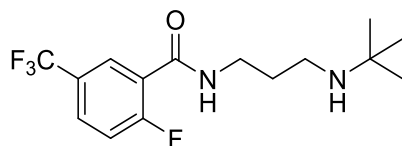
***N*-(3-(tert-butylamino)propyl)-3,5-bis(trifluoromethyl)benzamide (108):** Compound **108** was prepared from 3,5-bis(trifluoromethyl)benzoic acid (0.1 g, 0.39 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.074 g, 0.57 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (76 mg, 40.6 %) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 8.46 (s, 2H), 8.19 (s, 1H), 3.56 (t, $J = 6.1$ Hz, 1H), 3.08 (t, $J = 8.9$ Hz,

2H), 2.01 (m, 2H), 1.39 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.4$ min. MS (ESI⁺): 372.0 [M+H]⁺.



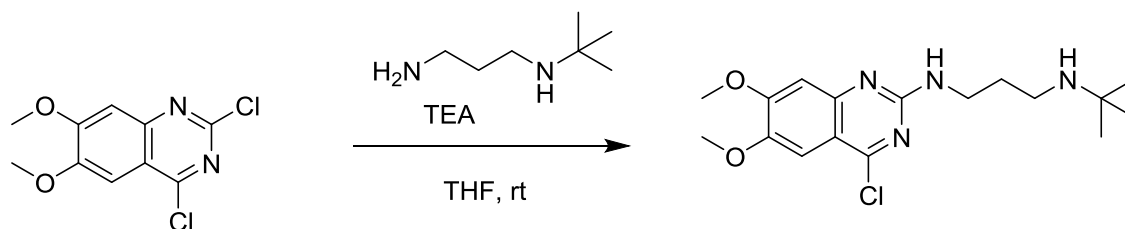
3-(tert-butyl)-N-(3-(tert-butylamino)propyl)-4-hydroxybenzamide (109): Compound **109**

was prepared from 3-(tert-butyl)-4-hydroxybenzoic acid (0.1 g, 0.51 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.080 g, 0.61 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (209 mg, 76.3 %) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.60 (d, $J = 2.3$ Hz, 1H), 7.39 (dd, $J = 8.4, 2.3$ Hz, 1H), 6.63 (d, $J = 8.4$ Hz, 1H), 3.33 (t, $J = 6.4$ Hz, 2H), 2.84 (t, $J = 7.1$ Hz, 2H), 1.82 (m, 2H), 1.22 (d, $J = 10.9$ Hz, 18H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.3$ min. MS (ESI⁺): 308.2 [M+H]⁺.

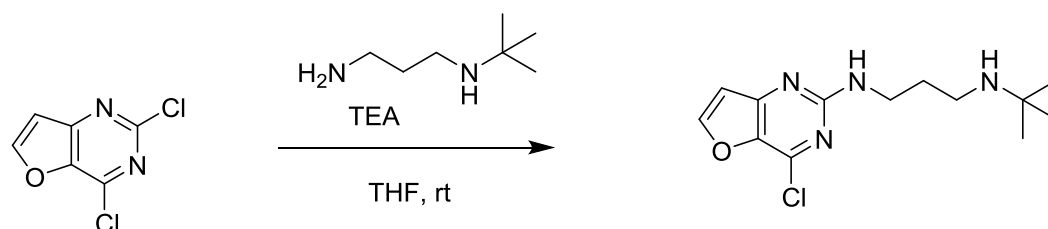


***N*-(3-(tert-butylamino)propyl)-2-fluoro-5-(trifluoromethyl)benzamide (110):** Compound **110**

was prepared from 2-fluoro-5-(trifluoromethyl)benzoic acid (0.3 g, 1.4 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.23 g, 1.7 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (276mg, 63.5 %) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 8.03 (dd, $J = 6.4, 2.4$ Hz, 1H), 7.86 (m, 1H), 7.43 (m, 1H), 3.55 (t, $J = 6.5$ Hz, 2H), 3.08 (m, 2H), 2.02 (tt, $J = 13.2, 6.7$ Hz, 2H), 1.40 (d, $J = 11.7$ Hz, 9H).. LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.9$ min. MS (ESI⁺): 322.2 [M+H]⁺.

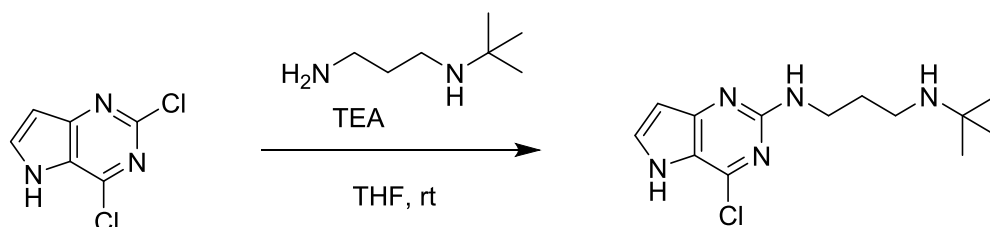


***N*¹-(*tert*-butyl)-*N*³-(4-chloro-6,7-dimethoxyquinazolin-2-yl)propane-1,3-diamine (111):** To a solution of 2,4-dichloro-6,7-dimethoxyquinazoline (0.15 g, 0.58 mmol) and triethylamine (0.064 g, 0.63 mmol) in 2 mL dimethylformamide at room temperature was slowly added *N*-(*tert*-butyl)propane-1,3-diamine (0.075 g, 0.58 mmol) and the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear yellow oil (133 mg, 49.2%) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.48 (s, 1H), 7.04 (s, 1H), 3.97 (d, *J* = 2.0 Hz, 7H), 3.75 (t, *J* = 6.8 Hz, 2H), 3.19 – 3.08 (m, 2H), 2.18 – 2.03 (m, 2H), 1.39 (s, 9H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.8 min. MS (ESI⁺): 355.1 [M+H]⁺.



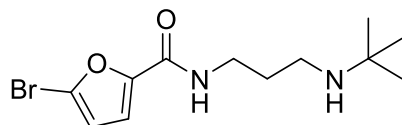
***N*¹-(*tert*-butyl)-*N*³-(4-chlorofuro[3,2-d]pyrimidin-2-yl)propane-1,3-diamine (112):** To a solution of 2,4-dichlorofuro[3,2-d]pyrimidine (0.15 g, 0.79 mmol) and triethylamine (0.241 g, 2.3 mmol) in 2 mL dimethylformamide at room temperature was slowly added *N*-(*tert*-butyl)propane-1,3-diamine (0.103 g, 0.79 mmol) and the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH,

concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a yellow solid (237 mg, 49.2%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.0 (s, 1H), 6.77 (d, $J = 1.7$ Hz, 1H), 3.17 – 3.05 (m, 2H), 3.69 (t, $J = 5.8, 5.8$ Hz, 2H), 2.17 – 2.02 (m, 2H), 1.36 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.1$ min. MS (ESI+): 283,2 $[\text{M}+\text{H}]^+$.

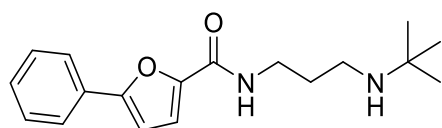


N^1 -(tert-butyl)- N^3 -(4-chloro-5H-pyrrolo[3,2-d]pyrimidin-2-yl)propane-1,3-diamine (113):

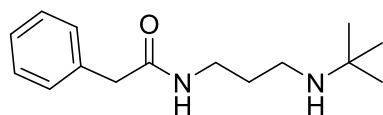
To a solution of 2,4-dichloro-5H-pyrrolo[3,2-d]pyrimidine (0.15 g, 0.79 mmol) and triethylamine (0.242 g, 2.3 mmol) in 2 mL dimethylformamide at room temperature was slowly added N -(tert-butyl)propane-1,3-diamine (0.104 g, 0.79 mmol) and the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a white solid (80 mg, 25.3%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.63 (d, $J = 3.0$ Hz, 1H), 6.45 (d, $J = 3.0$ Hz, 1H), 3.80 (t, $J = 6.7, 6.7$ Hz, 2H), 3.15 (dd, $J = 15.2, 7.2$ Hz, 2H), 2.09 – 2.15 (m, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.1$ min. MS (ESI+): 284.2 $[\text{M}+\text{H}]^+$.



5-bromo-*N*-(3-(tert-butylamino)propyl)furan-2-carboxamide (114): Compound **114** was prepared from 5-bromofuran-2-carboxylic acid (0.1 g, 0.52 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.082 g, 0.63 mmol) by the same procedure as compound **61**. The product was obtained as a pale yellow solid (99 mg, 46.4 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.12 (d, $J = 3.6$ Hz, 1H), 6.60 (d, $J = 3.6$ Hz, 1H), 3.46 (t, $J = 6.5$ Hz, 2H), 3.02 (m, 2H), 1.96 (m, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.5$ min. MS (ESI $^{+}$): 306.1 $[\text{M}+\text{H}]^{+}$.

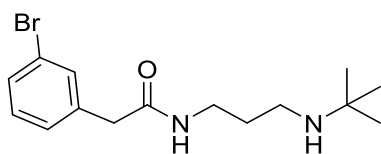


***N*-(3-(tert-butylamino)propyl)-5-phenylfuran-2-carboxamide (115):** Compound **115** was prepared from 5-phenylfuran-2-carboxylic acid (0.1 g, 0.53 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.083 g, 0.63 mmol) by the same procedure as compound **61**. The product was obtained as a clear yellow oil (180 mg, 81.8 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.84 (m, 2H), 7.41 (m, 2H), 7.33 (m, 1H), 7.21 (d, $J = 3.6$ Hz, 1H), 6.90 (d, $J = 3.6$ Hz, 1H), 3.52 (t, $J = 6.5$ Hz, 2H), 3.04 (t, $J = 7.4$ Hz, 2H), 2.01 (m, 2H), 1.36 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 4.3$ min. MS (ESI $^{+}$): 302.2 $[\text{M}+\text{H}]^{+}$.

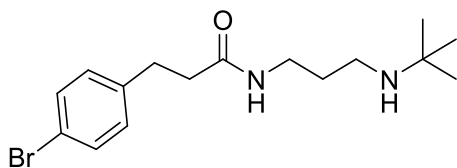


***N*-(3-(tert-butylamino)propyl)-2-phenylacetamide (116):** Compound **116** was prepared from 2-phenylacetic acid (0.1 g, 0.65 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.084 g, 0.63

mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (80 mg, 35.5 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.26 (m, 5H), 3.52 (d, J = 2.7 Hz, 2H), 3.30 (m, 2H), 2.84 (t, J = 7.4 Hz, 2H), 1.84 (td, J = 13.1, 7.1 Hz, 2H), 1.25 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 3.4 min. MS (ESI $^{+}$): 250.2 $[\text{M}+\text{H}]^{+}$.

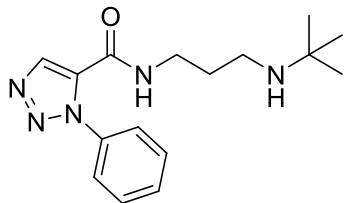


2-(3-bromophenyl)-N-(3-(tert-butylamino)propyl)acetamide (117): Compound **117** was prepared from 2-(3-bromophenyl)acetic acid (0.1 g, 0.46 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.073 g, 0.56 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (85 mg, 41.5 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.51 (t, J = 1.7 Hz, 1H), 7.41 (m, 1H), 7.29 (dt, J = 7.7, 1.3 Hz, 1H), 7.23 (t, J = 7.7 Hz, 1H), 3.52 (s, 2H), 3.31 (m, 3H), 2.87 (t, J = 7.4 Hz, 2H), 1.86 (td, J = 13.2, 7.0 Hz, 2H), 1.28 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 3.9 min. MS (ESI $^{+}$): 330.2 $[\text{M}+\text{H}]^{+}$.

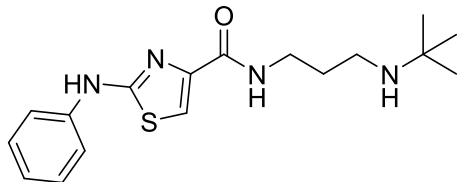


3-(4-bromophenyl)-N-(3-(tert-butylamino)propyl)propanamide (118): Compound **118** was prepared from 3-(4-bromophenyl)propanoic acid (0.1 g, 0.43 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.068 g, 0.52 mmol) by the same procedure as compound **61**. The product was obtained as a yellow solid (201 mg, 98.5 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.40 (m, 2H), 7.14 (m, 2H), 3.25 (t, J = 6.5 Hz, 2H), 2.87 (m, 4H), 2.53 (t, J = 7.5

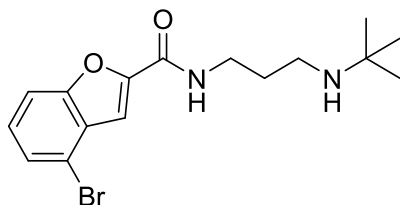
Hz, 2H), 1.82 (m, 2H), 1.32 (s, 9H).LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.3$ min. MS (ESI+): 344.2 [M+H]⁺.



***N*-(3-(tert-butylamino)propyl)-1-phenyl-1H-1,2,3-triazole-5-carboxamide (119):** Compound **119** was prepared from 1-phenyl-1H-1,2,3-triazole-5-carboxylic acid (0.1 g, 0.52 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.083 g, 0.63 mmol) by the same procedure as compound **61**. The product was obtained as a white solid (37 mg, 16.8 %) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 8.20 (s, 1H), 7.54 (m, 5H), 3.43 (t, $J = 6.7$ Hz, 2H), 2.98 (m, 2H), 1.93 (m, 2H), 1.30 (s, 9H).LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.4$ min. MS (ESI+): 303.2 [M+H]⁺.

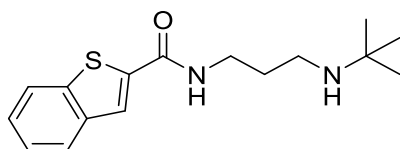


***N*-(3-(tert-butylamino)propyl)-2-(phenylamino)thiazole-4-carboxamide (120):** Compound **120** was prepared from 2-(phenylamino)thiazole-4-carboxylic acid (0.1 g, 0.45 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.071 g, 0.54 mmol) by the same procedure as compound **61**. The product was obtained as a white solid (73 mg, 36.0 %) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.62 (m, 2H), 7.50 (s, 1H), 7.33 (m, 2H), 7.01 (m, 1H), 3.53 (t, $J = 6.4$ Hz, 2H), 3.02 (t, $J = 7.3$ Hz, 2H), 1.99 (m, 2H), 1.37 (s, 9H).LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.1$ min. MS (ESI+): 334.2 [M+H]⁺.



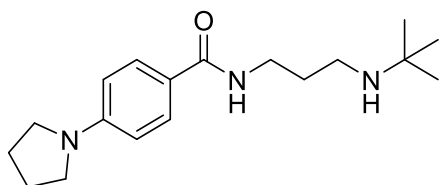
4-bromo-*N*-(3-(tert-butylamino)propyl)benzofuran-2-carboxamide (121): Compound **121**

was prepared from 4-bromobenzofuran-2-carboxylic acid (0.1 g, 0.41 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.065 g, 0.50 mmol) by the same procedure as compound **61**. The product was obtained as a yellow solid (122 mg, 63.2 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.58 (dt, $J = 8.4, 0.8$ Hz, 1H), 7.48 (m, 2H), 7.36 (m, 1H), 3.55 (t, $J = 6.6$ Hz, 2H), 3.07 (m, 2H), 2.02 (dd, $J = 8.0, 6.9$ Hz, 2H), 1.38 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.4$ min. MS (ESI $^+$): 356.1 $[\text{M}+\text{H}]^+$.

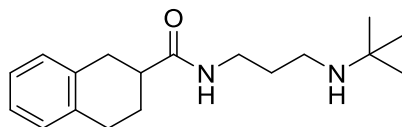


***N*-(3-(tert-butylamino)propyl)benzo[b]thiophene-2-carboxamide (122):** Compound **122** was

prepared from benzo[b]thiophene-2-carboxylic acid (0.1 g, 0.56 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.087 g, 0.66 mmol) by the same procedure as compound **61**. The product was obtained as a white solid (168 mg, 74.3 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.97 (s, 1H), 7.88 (m, 2H), 7.42 (m, 2H), 3.53 (t, $J = 6.5$ Hz, 2H), 3.05 (t, $J = 7.4$ Hz, 2H), 2.01 (dt, $J = 13.5, 6.9$ Hz, 2H), 1.37 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.9$ min. MS (ESI $^+$): 292.2 $[\text{M}+\text{H}]^+$.

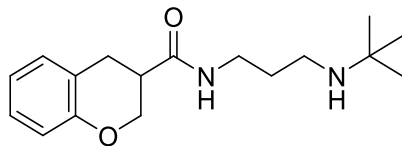


***N*-(3-(*tert*-butylamino)propyl)benzo[*b*]thiophene-2-carboxamide (123):** Compound **123** was prepared from 4-(pyrrolidin-1-yl)benzoic acid (0.1 g, 0.52 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.082 g, 0.62 mmol) by the same procedure as compound **61**. The product was obtained as a brown solid (173 mg, 79.3 %) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.72 (d, *J* = 8.9 Hz, 2H), 6.54 (d, *J* = 9.0 Hz, 2H), 3.48 (m, 2H), 3.29 (m, 4H), 2.98 (t, *J* = 7.0 Hz, 2H), 1.98 (m, 6H), 1.36 (s, 8H).LC-MS (λ = 254 nm): 99%, *t*_R = 4.2 min. MS (ESI⁺): 305.3 [M+H]⁺.



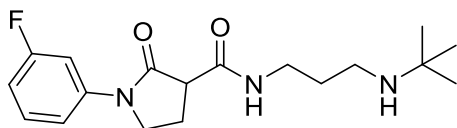
***N*-(3-(*tert*-butylamino)propyl)-1,2,3,4-tetrahydronaphthalene-2-carboxamide (124):**

Compound **124** was prepared from 1,2,3,4-tetrahydronaphthalene-2-carboxylic acid (0.1 g, 0.56 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.088 g, 0.67 mmol) by the same procedure as compound **61**. The product was obtained as a yellow solid (205 mg, 89.9 %) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.05 (s, 4H), 3.35 (t, *J* = 6.5 Hz, 2H), 2.97 (dd, *J* = 12.5, 5.3 Hz, 3H), 2.90 (m, 1H), 2.84 (m, 2H), 2.63 (m, 1H), 2.05 (m, 1H), 1.90 (m, 2H), 1.83 (m, 1H), 1.36 (s, 9H).LC-MS (λ = 254 nm): 99%, *t*_R = 3.9 min. MS (ESI⁺): 290.3 [M+H]⁺.



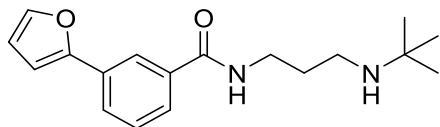
***N*-(3-(*tert*-butylamino)propyl)chromane-3-carboxamide (125):** Compound **125** was prepared from chromane-3-carboxylic acid (0.1 g, 0.56 mmol) and *N*-(*tert*-butyl)propane-1,3-diamine (0.088 g, 0.67 mmol) by the same procedure as compound **61**. The product was obtained as a yellow oil (217 mg, 95.6 %) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.01

- 7.08 (m, 2H), 6.83 (td, $J = 7.4, 1.2$ Hz, 1H), 6.74 (dd, $J = 8.1, 1.0$ Hz, 1H), 4.31 (ddd, $J = 10.8, 3.2, 1.8$ Hz, 1H), 4.09 (dd, $J = 10.7, 8.3$ Hz, 1H), 3.33 (m, 2H), 3.06 (dd, $J = 17.4, 10.5$ Hz, 1H), 2.92 (m, 4H), 1.88 (m, 2H), 1.33 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.7$ min. MS (ESI+): 292.2 $[M+H]^+$.



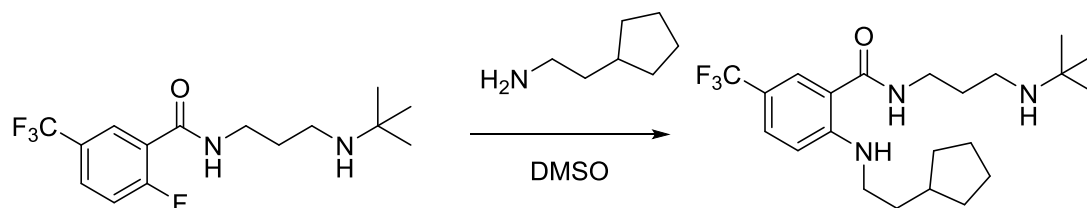
***N*-(3-(tert-butylamino)propyl)-1-(3-fluorophenyl)-2-oxopyrrolidine-3-carboxamide (126):**

Compound **126** was prepared from 1-(3-fluorophenyl)-2-oxopyrrolidine-3-carboxylic acid (0.1 g, 0.44 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.070 g, 0.53 mmol) by the same procedure as compound **61**. The product was obtained as a clear oil (60 mg, 20.1 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.38 (m, 2H), 7.23 (m, 2H), 3.87 (dt, $J = 9.5, 5.1$ Hz, 2H), 3.59 (m, 2H), 3.25 (dt, $J = 14.2, 5.8$ Hz, 1H), 3.03 (m, 2H), 2.60 (dq, $J = 12.8, 8.4$ Hz, 1H), 2.42 (m, 1H), 1.91 (m, 2H), 1.34 (s, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.6$ min. MS (ESI+): 337.2 $[M+H]^+$.



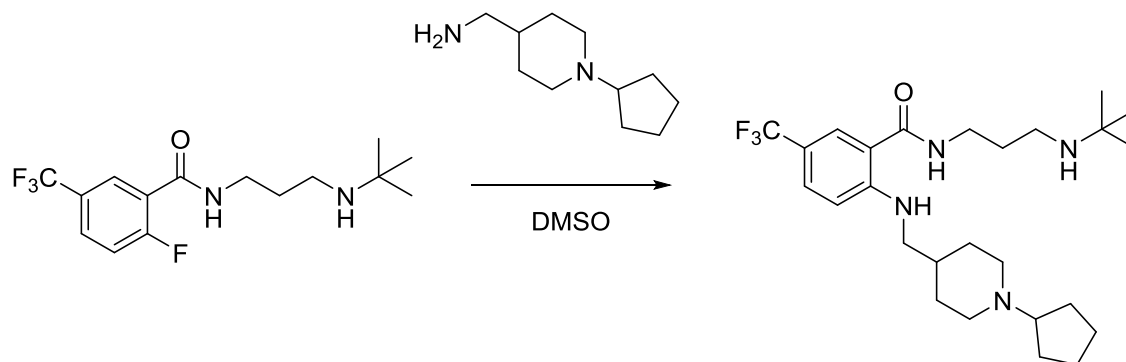
***N*-(3-(tert-butylamino)propyl)-3-(furan-2-yl)benzamide (127):** Compound **127** was prepared from 3-(furan-2-yl)benzoic acid (0.11 g, 0.59 mmol) and *N*-(tert-butyl)propane-1,3-diamine (0.092 g, 0.713 mmol) by the same procedure as compound **61**. The product was obtained as a yellow oil (177 mg, 80.4 %) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.18 (t, $J = 1.6$ Hz, 1H), 7.83 (ddd, $J = 7.8, 1.6, 1.1$ Hz, 1H), 7.74 (ddd, $J = 7.8, 1.7, 1.1$ Hz, 1H), 7.57 (dd, $J = 1.8, 0.6$ Hz, 1H), 7.48 (m, 1H), 6.84 (dd, $J = 3.4, 0.6$ Hz, 1H), 6.52 (dd, $J = 3.4, 1.8$ Hz,

1H), 3.53 (t, $J = 6.5$ Hz, 2H), 3.03 (t, $J = 7.4$ Hz, 2H), 2.01 (m, 2H), 1.36 (s, 9H).LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.2$ min. MS (ESI⁺): 302.2 [M+H]⁺.



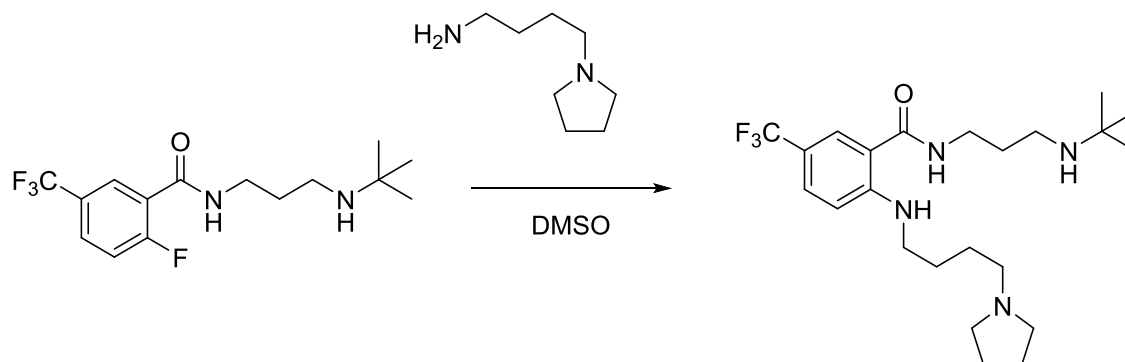
***N*-(3-(tert-butylamino)propyl)-2-((2-cyclopentylethyl)amino)-5-(trifluoromethyl)benzamide**

(128): To a solution of **110** (0.051 g, 0.16 mmol) in 4 mL DMSO at room temperature was slowly added 2-cyclopentylethan-1-amine (0.019 g, 0.17 mmol) and the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear yellow oil (45 mg, 27.2%) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.82 (d, $J = 1.3$ Hz, 1H), 7.53 (dd, $J = 8.9, 2.1$ Hz, 1H), 6.82 (d, $J = 8.9$ Hz, 1H), 3.47 (t, $J = 6.5$ Hz, 2H), 3.22 (t, $J = 7.1$ Hz, 2H), 3.04 (t, $J = 7.4$ Hz, 2H), 1.97 (dt, $J = 14.9, 7.4$ Hz, 3H), 1.85 (m, 2H), 1.62 (m, 6H), 1.38 (s, 9H), 1.19 (ddd, $J = 8.0, 6.0, 3.3$ Hz, 2H).LC-MS ($\lambda = 254$ nm): 99%, $t_R = 5.0$ min. MS (ESI⁺): 415.2 [M+H]⁺.



***N*-(3-(tert-butylamino)propyl)-2-(((1-cyclopentylpiperidin-4-yl)methyl)amino)-5-**

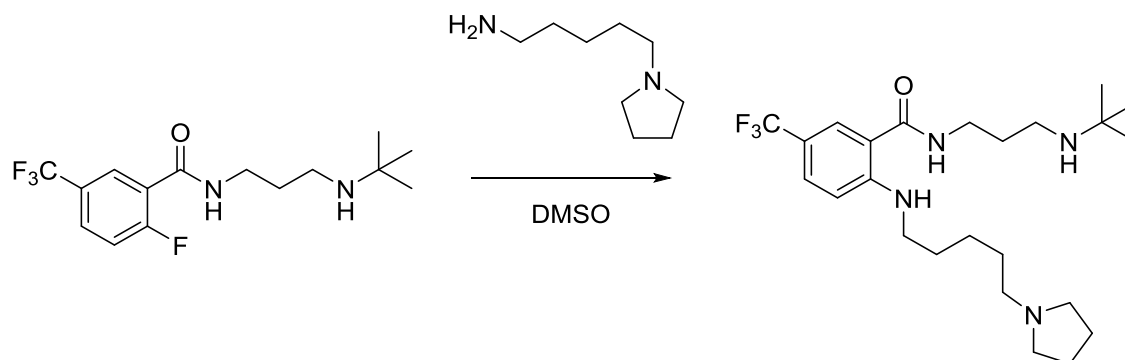
(trifluoromethyl)benzamide (129): To a solution of **110** (0.041 g, 0.13 mmol) in 4 mL DMSO at room temperature was slowly added (1-cyclopentylpiperidin-4-yl)methanamine (0.026 g, 0.14 mmol) and the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a dark yellow oil (65 mg, 85.5%) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.83 (d, *J* = 1.3 Hz, 1H), 7.54 (dd, *J* = 8.9, 2.0 Hz, 1H), 6.89 (d, *J* = 8.9 Hz, 1H), 3.65 (d, *J* = 12.5 Hz, 2H), 3.48 (m, 3H), 3.34 (m, 1H), 3.23 (t, *J* = 7.2 Hz, 2H), 3.00 (ddd, *J* = 15.0, 14.2, 4.8 Hz, 4H), 2.14 (m, 4H), 1.99 (m, 3H), 1.82 (dd, *J* = 9.8, 5.5 Hz, 2H), 1.65 (m, 6H), 1.40 (m, 9H). LC-MS (λ = 254 nm): 99%, *t*_R = 3.6 min. MS (ESI⁺): 484.3 [M+H]⁺.



***N*-(3-(tert-butylamino)propyl)-2-((4-(pyrrolidin-1-yl)butyl)amino)-5-**

(trifluoromethyl)benzamide (130): To a solution of **110** (0.071 g, 0.22 mmol) in 4 mL DMSO at room temperature was slowly added 4-(pyrrolidin-1-yl)butan-1-amine (0.035 g, 0.24 mmol) and the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-

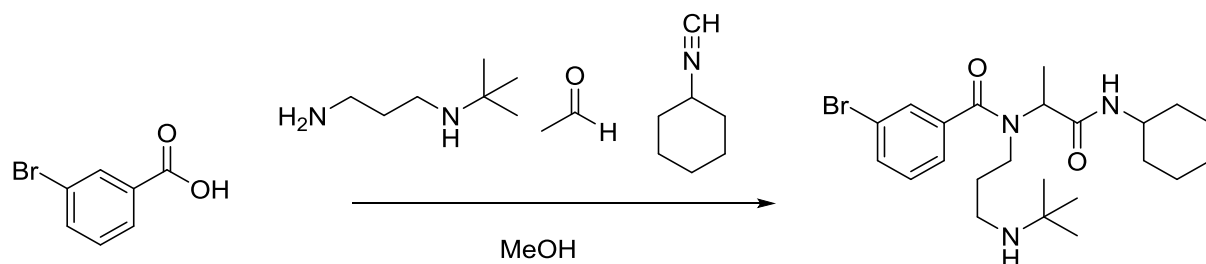
suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a red oil (63 mg, 51.2%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.81 (d, $J = 1.3$ Hz, 1H), 7.54 (dd, $J = 8.9, 1.9$ Hz, 1H), 6.86 (d, $J = 8.9$ Hz, 1H), 3.66 (m, 2H), 3.60 (m, 4H), 3.47 (t, $J = 6.6$ Hz, 2H), 3.31 (ddd, $J = 8.7, 5.2, 3.6$ Hz, 3H), 3.24 (dd, $J = 9.2, 6.8$ Hz, 2H), 3.21 (m, 9H), 3.07 (dt, $J = 20.8, 7.8$ Hz, 4H), 2.13 (dt, $J = 11.3, 6.5$ Hz, 2H), 2.01 (m, 4H), 1.95 (m, 10H), 1.86 (ddd, $J = 10.8, 9.4, 5.9$ Hz, 2H), 1.75 (dt, $J = 14.4, 7.1$ Hz, 2H), 1.37 (d, $J = 8.5$ Hz, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 3.7$ min. MS (ESI $^{+}$): 443.3 $[\text{M}+\text{H}]^{+}$.



***N*-(3-(tert-butylamino)propyl)-2-((5-(pyrrolidin-1-yl)pentyl)amino)-5-**

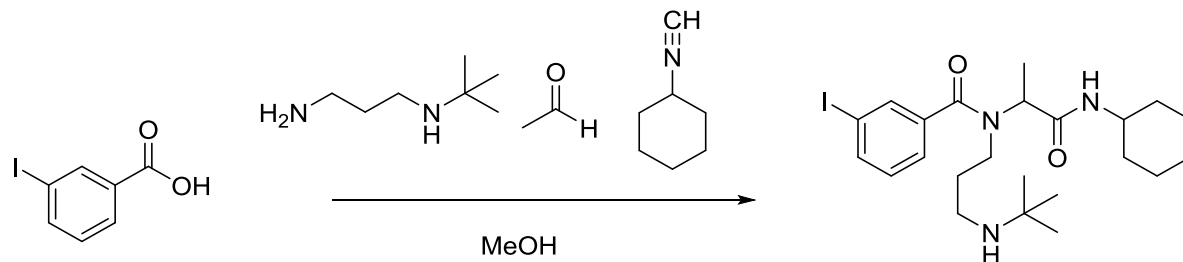
(trifluoromethyl)benzamide (131): To a solution of **110** (0.054 g, 0.17 mmol) in 4 mL DMSO at room temperature was slowly added 5-(pyrrolidin-1-yl)pentan-1-amine (0.032 g, 0.19 mmol) and the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a red oil (92 mg, 95.8%) after purification as the TFA salt. ^1H NMR

(400 MHz, cd_3od) δ 7.81 (d, $J = 1.2$ Hz, 1H), 7.53 (dd, $J = 8.9, 1.9$ Hz, 1H), 6.84 (d, $J = 8.9$ Hz, 1H), 3.65 (d, $J = 5.2$ Hz, 2H), 3.47 (t, $J = 6.5$ Hz, 2H), 3.26 (t, $J = 6.9$ Hz, 2H), 3.19 (m, 2H), 3.06 (dd, $J = 16.7, 9.2$ Hz, 4H), 2.14 (m, 2H), 2.00 (m, 4H), 1.76 (m, 4H), 1.54 (dd, $J = 15.2, 7.9$ Hz, 2H), 1.37 (d, $J = 9.2$ Hz, 9H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.9$ min. MS (ESI⁺): 457.3 [M+H]⁺.



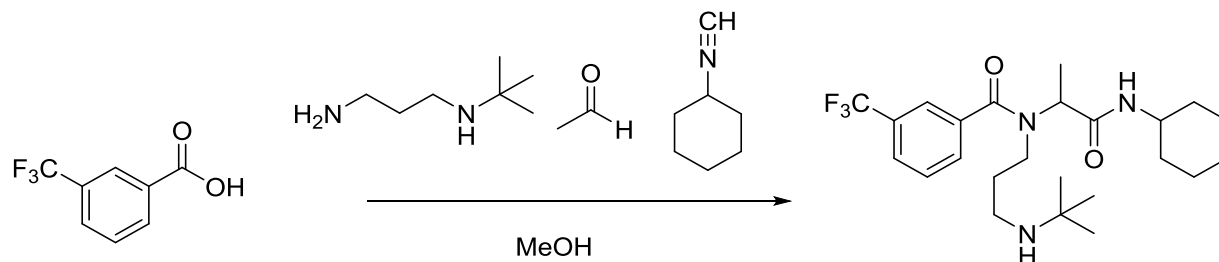
3-bromo-N-(3-(tert-butylamino)propyl)-N-(1-(cyclohexylamino)-1-oxopropan-2-

yl)benzamide (132): To a solution of 3-bromobenzoic acid (0.1 g, 0.49 mmol) in 4 mL DMSO at room temperature was slowly added N¹-(tert-butyl)propane-1,3-diamine (0.07 g, 0.52 mmol), acetaldehyde (0.03 mL, 0.49 mmol), and lastly cyclohexyl isocyanide (0.06 mL, 0.49 mmol) the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear yellow oil (161 mg, 69.3%) after purification as the TFA salt. ¹H NMR (400 MHz, cd_3od) δ 7.65 (m, 1H), 7.39 (m, 3H), 4.30 (s, 1H), 3.64 (d, $J = 37.6$ Hz, 3H), 3.12 (s, 2H), 2.11 (d, $J = 21.4$ Hz, 2H), 1.43 (m, 24H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.5$ min. MS (ESI⁺): 470.2 [M+H]⁺.



***N*-(3-(tert-butylamino)propyl)-*N*-(1-(cyclohexylamino)-1-oxopropan-2-yl)-3-**

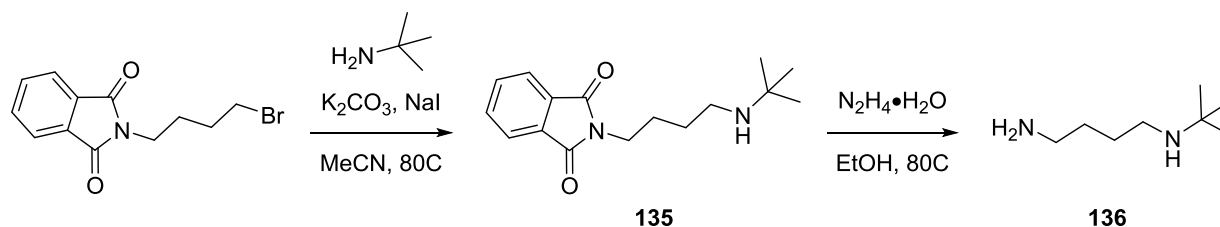
iodobenzamidebenzamide (133): To a solution of 3-iodobenzoic acid (0.1 g, 0.49 mmol) in 4 mL DMSO at room temperature was slowly added *N*¹-(tert-butyl)propane-1,3-diamine (0.07 g, 0.52 mmol), acetaldehyde (0.03 mL, 0.49 mmol), and lastly cyclohexyl isocyanide (0.06 mL, 0.49 mmol) the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear yellow oil (130 mg, 54.3%) after purification as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.85 (ddd, *J* = 7.9, 1.6, 1.2 Hz, 1H), 7.66 (s, 1H), 7.35 (d, *J* = 7.3 Hz, 1H), 7.24 (t, *J* = 7.8 Hz, 1H), 4.30 (s, 1H), 3.63 (d, *J* = 32.9 Hz, 3H), 3.11 (s, 2H), 2.07 (m, 2H), 1.75 (ddd, *J* = 53.4, 44.7, 11.5 Hz, 5H), 1.37 (m, 14H), 1.19 (s, 3H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.7 min. MS (ESI⁺): 515.2 [M+H]⁺.



***N*-(3-(tert-butylamino)propyl)-*N*-(1-(cyclohexylamino)-1-oxopropan-2-yl)-3-**

(trifluoromethyl)benzamide (134): To a solution of 3-(trifluoromethyl)benzoic acid (0.1 g, 0.49

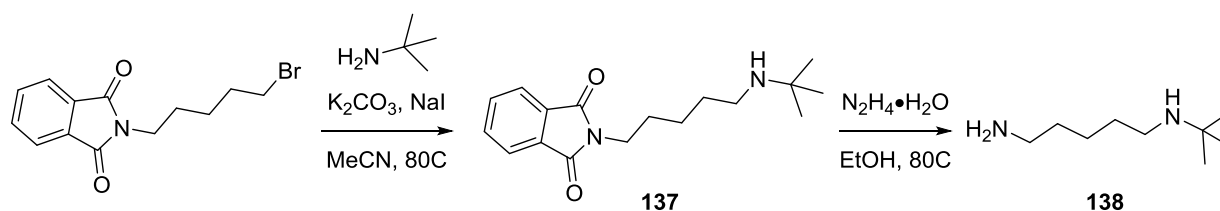
mmol) in 4 mL DMSO at room temperature was slowly added N^1 -(tert-butyl)propane-1,3-diamine (0.07 g, 0.52 mmol), acetaldehyde (0.03 mL, 0.49 mmol), and lastly cyclohexyl isocyanide (0.06 mL, 0.49 mmol) the reaction was stirred overnight. Once the reaction had gone to completion via TLC analysis, the reaction mixture was diluted with MeOH, concentrated down to a residue, re-suspended in methanol, and dried onto silica. The reaction mixture was purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a clear yellow oil (130 mg, 54.3%) after purification as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.81 (d, $J = 7.7$ Hz, 1H), 7.67 (dd, $J = 17.2, 9.5$ Hz, 3H), 4.29 (s, 1H), 3.64 (d, $J = 49.7$ Hz, 3H), 3.13 (s, 2H), 2.09 (m, 2H), 1.74 (d, $J = 13.0$ Hz, 4H), 1.63 (d, $J = 12.7$ Hz, 1H), 1.50 (d, $J = 6.0$ Hz, 3H), 1.33 (m, 12H), 1.20 (dd, $J = 26.2, 9.5$ Hz, 3H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.7$ min. MS (ESI $^+$): 458.3 $[\text{M}+\text{H}]^+$.



2-(4-(Tert-butylamino)butyl)isoindoline-1,3-dione (135): To a solution of 2-(4-bromobutyl)isoindoline-1,3-dione (0.40 g, 1.43 mmol) in acetonitrile (10 ml) was added *N*-tert-butylamine (0.14 ml, 1.34 mmol), potassium carbonate (0.39 g, 2.84 mmol), and sodium iodide (0.43 g, 2.83 mmol) and the reaction was stirred overnight at 80°C. After 24 hrs the reaction was cooled to room temperature, dried down to a crude residue, dissolved in water (10 ml), and then extracted with dichloromethane 3 times. The organic layers were combined, dried over Na_2SO_4 , and filtered. The crude material was dried onto silica and purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was

obtained as a white solid (117 mg, 22%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.85 – 7.78 (m, 4H), 3.74 (t, J = 6.7 Hz, 2H), 3.06 - 3.02 (m, 2H), 1.84 - 1.78 (m, 2H), 1.73 - 1.65 (s, 2H), 1.36 (s, 9H). LC-MS (λ = 254 nm): 99%, t_R = 3.6 min. MS (ESI $^+$): 275.2 $[\text{M}+\text{H}]^+$.

***N*-(*Tert*-butyl)butane-1,4-diamine (136):** To a solution of **135** (0.12 g, 0.31 mmol) in 10 mL of ethanol was added hydrazine monohydrate (0.03 ml, 0.61 mmol) and stirred for 2 hrs at 80°C. The reaction was cooled, dried to a residue, dissolved in chloroform, and filtered. The resulting solution was dried onto silica and purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column and an ELSD detector. The product was obtained as a white solid (34 mg, 29%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 3.03 – 2.97 (m, 4H), 1.81 – 1.71 (m, 4H), 1.37 (s, 9H). LC-MS (λ = 254 nm): 99%, t_R = 0.6 min. MS (ESI $^+$): 145.3 $[\text{M}+\text{H}]^+$.



2-(5-(*Tert*-butylamino)pentyl)isoindoline-1,3-dione (137): To a solution of 2-(5-bromopentyl)isoindoline-1,3-dione (0.41 g, 1.37 mmol) in acetonitrile (10 ml) was added *N*-*tert*-butylamine (0.14 ml, 1.34 mmol), potassium carbonate (0.39 g, 2.70 mmol), and sodium iodide (0.41 g, 2.74 mmol) and the reaction was stirred overnight at 80°C. After 24 hrs the reaction was cooled to room temperature, dried down to a crude residue, dissolved in water (10 ml), and then extracted with dichloromethane 3 times. The organic layers were combined, dried over Na_2SO_4 ,

and filtered. The crude material was dried onto silica and purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column. The product was obtained as a white solid (156 mg, 29%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 7.85 – 7.78 (m, 4H), 3.70 (t, J = 8.1 Hz, 2H), 2.98 – 2.94 (m, 2H), 1.78 - 1.68 (m, 4H), 1.52 - 1.42 (m, 2H), 1.36 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 4.0 min. MS (ESI+): 289.2 $[\text{M}+\text{H}]^+$.

***N*-(*Tert*-butyl)pentane-1,5-diamine (138):** To a solution of **137** (0.16 g, 0.40 mmol) in 10 mL of ethanol was added hydrazine monohydrate (0.07 mL, 1.43 mmol) and stirred for 2 hrs at 80°C. The reaction was cooled, dried to a residue, dissolved in chloroform, and filtered. The resulting solution was dried onto silica and purified by reverse phase column chromatography using a Teledyne Isco reverse phase chromatography column and an ELSD detector. The product was obtained as a white solid (71 mg, 47%) after purification as the TFA salt. ^1H NMR (400 MHz, CD_3OD) δ 2.99 – 2.92 (m, 4H), 1.75 – 1.67 (m, 4H), 1.53 – 1.46 (m, 2H), 1.37 (s, 9H). LC-MS (λ = 254 nm): 99%, t_{R} = 0.6 min. MS (ESI+): 159.2 $[\text{M}+\text{H}]^+$.

Analytical and structural characterization (LC-MS and NMR)

Analytical LCMS data for all compounds were acquired using an Agilent 6110 series system with the UV detector set to 220 and 254 nm. Samples were injected (<10 μL) onto an Agilent Eclipse Plus 4.6 \times 50 mm, 1.8 μm , C18 column at room temperature. A mobile phase of A (H_2O + 0.1% acetic acid) and B (MeOH + 0.1% acetic acid) was used with a linear gradient from 10% to 100% B in 5.0 min, followed by a flush at 100% B for another 2 minutes with a

flow rate of 1.0 mL/min. Mass spectra data were acquired in positive ion mode using an Agilent 6110 single quadrupole mass spectrometer with an electrospray ionization source. Nuclear Magnetic Resonance (NMR) spectra were recorded on a Varian Mercury spectrometer at 400 MHz for proton (^1H NMR) and 100 MHz for carbon (^{13}C NMR); chemical shifts are reported in ppm (δ). All synthesized peptides were dissolved in deuterated MeOH.

Protein Expression and Purification

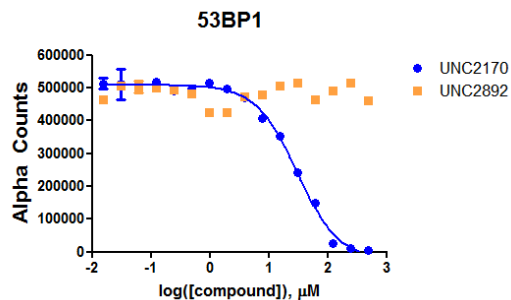
53BP1, L3MBTL1, L3MBTL3, CBX7, MBTD1, and UHRF1 (tandem tudor-PHD construct) were expressed and purified as previously described.[53, 105] The proteins constructs for the aforementioned proteins were provided by the Structural Genomics Consortium. PHF1, PHF19, PHF23, and JARID1A protein constructs were provided by Greg Wang (UNC) and prepared in the same manner. The construct for the 53BP1 D1521A mutant (residues 1484 - 1603) was obtained from George Mer (Mayo Clinic) and was expressed and purified in the same manner as 53BP1.[66, 105]

AlphaScreen assay of synthesized ligands

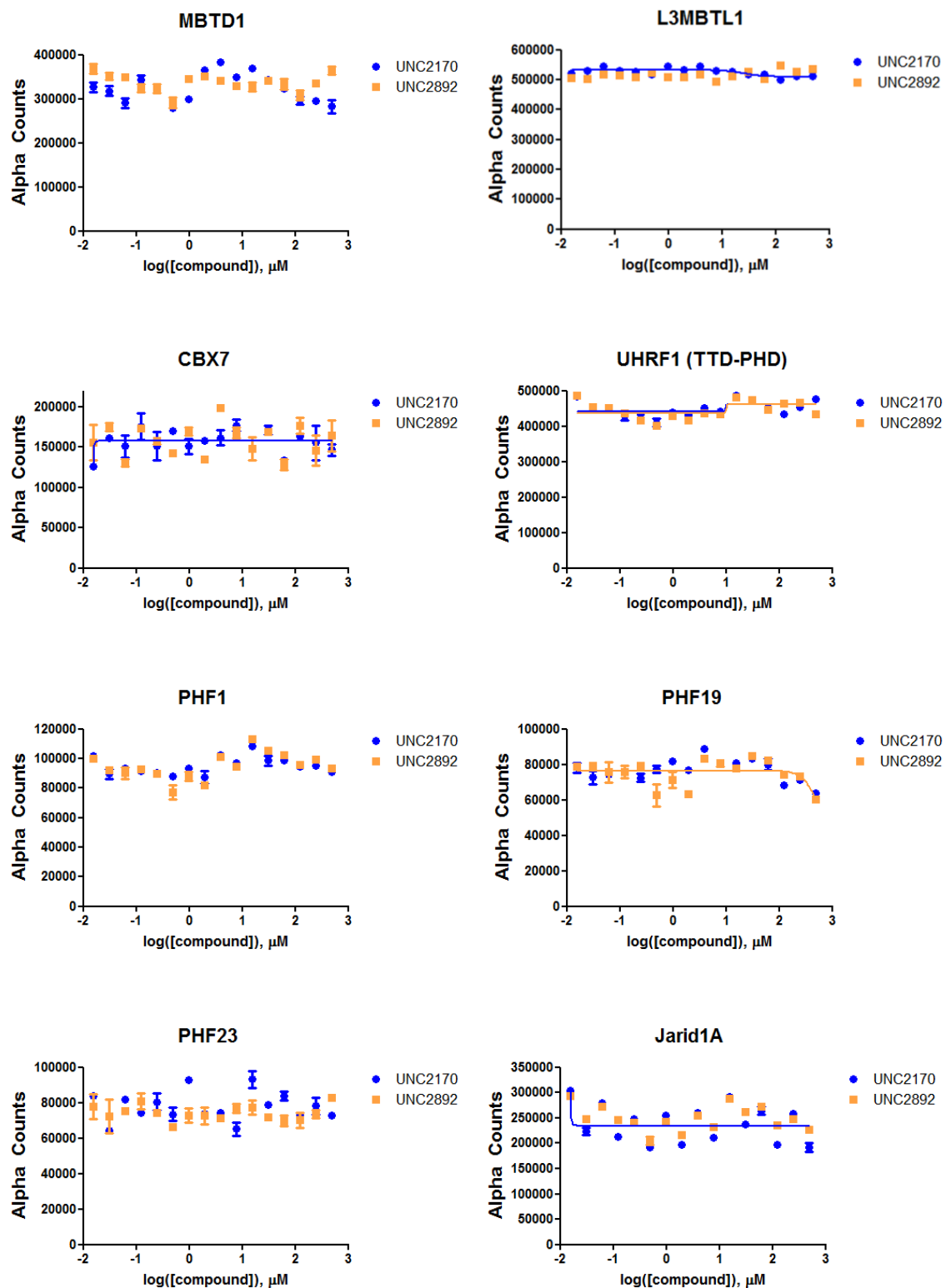
The AlphaScreen assay was generally performed as previously described.[105] The general method for screening synthesized peptide ligands is as follows. Peptides were weighed out as solids and then dissolved in de-ionized water to a concentration of 10 mM. Compound plates were then prepared using an automated liquid handling system. Compound plates (1 μL at 10 mM highest concentration; 3-fold, 10-point dilutions in DMSO) were then diluted in 1X assay buffer (20 mM TRIS pH 8.0, 25 mM NaCl, 2 mM DTT and 0.05% Tween-20) to 1 mM using a Multimek robotic pipettor (Nanoscreen) and 1 μL was spotted into the wells of 384-well low-volume Proxiplates (PerkinElmer). To these plates 9 μL of protein-peptide mix in 1X assay

buffer was added by Multidrop (Thermo) to bring the final compound concentration to 100 μM and incubated for 30 min at room temperature. Next, 2 μL of a 1:1 mixture of streptavidin-conjugate donor and nickel-chelate acceptor beads (45 $\mu\text{g/mL}$ in 1X assay buffer) were added and the plates were allowed to incubate for an additional 30 min in the dark at room temperature. After incubation, the plates were read on an EnVision multi-label reader equipped with an HTS AlphaScreen laser (Perkin Elmer). The IC_{50} values reported are the average of at least 2 values \pm the standard deviation. When IC_{50} values for a single compound were not all active ($< 100 \mu\text{M}$) or inactive ($> 100 \mu\text{M}$), the IC_{50} values were calculated using 4-parameter curve fitting (GraphPad Prism 5) from replicate runs using averaged response values for each compound concentration.

Averaged AlphaScreening binding curves for UNC2170 (**61**) and UNC2892 (**64**) against 53BP1. The average IC_{50} for UNC2170 is $28.48 \pm 7.44 \mu\text{M}$.



UNC2170 (**61**) and UNC2892 (**64**) selectivity data against the Kme panel. Both compounds were inactive against every panel member up to 500 μM .

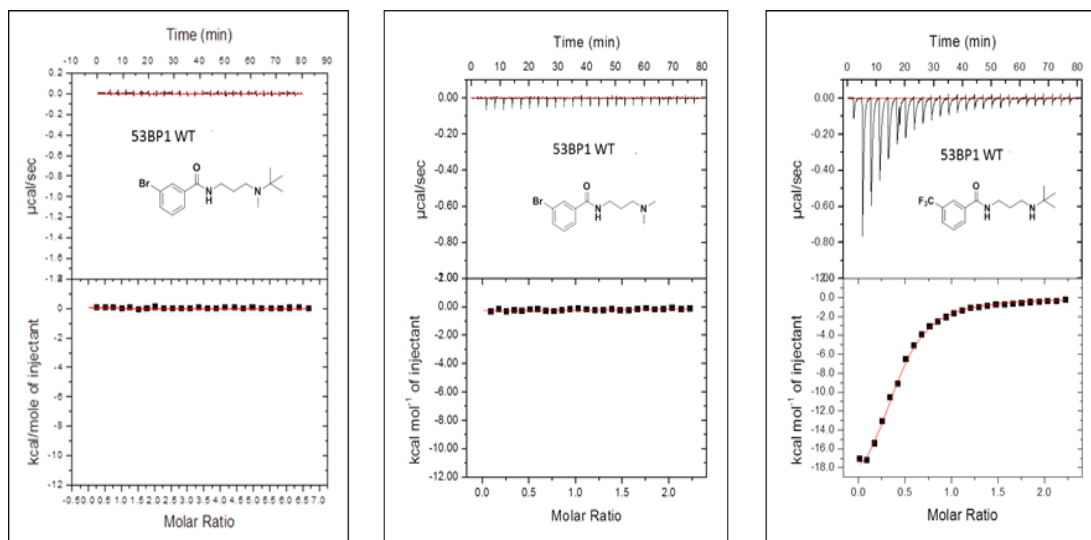


AlphaScreen proteins and their corresponding peptide substrates.

Protein	Peptide	Peptide sequence
53BP1	H4K20Me2	Biotin-AHX-KGGAKRHRK(Me2)VLRDNIQ-OH
L3MBTL1	H4K20Me1	Biotin-AHA-KGGAKRHRK(Me1)VLRDNIQ-OH
L3MBTL3	H4K20Me2	Biotin-AHX-KGGAKRHRK(Me2)VLRDNIQ-OH
MBTD1	H4K20Me1	Biotin-AHA-KGGAKRHRK(Me1)VLRDNIQ-OH
CBX7	H3K9Me3	ARTKQTARK(Me3)STGGKAPRKQL-K(Biotin)-NH2
UHRF1	H3K9Me3	ARTKQTARK(Me3)STGGKAPRKQL-K(Biotin)-NH2
PHF23	H3K4Me3	ARTK(Me3)QTARKSTGGKAPRKQYT-K(Biotin)-NH2
JARID1A	H3K4Me3	ARTK(Me3)QTARKSTGGKAPRKQYT-K(Biotin)-NH2
PHF1	H3K36me3	KSAPSTGGVK(Me3)KPHRYRPGTV-K(biotin)-NH2
PHF19	H3K36me3	KSAPSTGGVK(Me3)KPHRYRPGTV-K(biotin)-NH2

Isothermal titration calorimetry binding experiments

All ITC measurements were recorded at 25 °C with an AutoITC₂₀₀ microcalorimeter (MicroCal Inc.). All protein and compound stock samples were prepared in the target buffer (25 mM Tris-HCl, pH 8, 150 mM NaCl, and 2 mM β -mercaptoethanol), and then diluted in the same buffer to achieve the desired concentrations: 90 – 200 μ M protein and 1 – 2 mM compound depending on the expected dissociation constant. The concentration of protein stock solutions were established using the Edelhoch method, whereas 10 mM compound stock solutions were prepared gravimetrically based on molecular weight. A typical experiment included a single 0.2 μ L compound injection into a 200 μ L cell filled with protein, followed by 25 subsequent 1.5 μ L injections of compound. Injections were performed with a spacing of 180 seconds and a reference power of 8 μ cal/sec. The titration data was analyzed using Origin Software (MicroCal Inc.) by non-linear least squares, fitting the heats of binding as a function of the compound:protein ratio. The data were fit based on a one set of sites model.



Competitive In-solution Peptide Pulldown Assays

A 5 μL slurry of streptavidin magnetic beads (Pierce) was equilibrated in binding buffer containing 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 0.1% NP-40 before being saturated with 500 pmoles of biotinylated peptide for 1 h at room temperature with rotation. Unbound peptide was washed with binding buffer. To pre-complex protein with inhibitor, an 80 pmol solution of His-53BP1 TTD (residues 1485-1611) in binding buffer containing 0.5% BSA was incubated with rotation at 4°C with the indicated concentration of UNC2170 or vehicle (H_2O). The pre-complex was combined with peptide-saturated resin and incubated for 3 h at 4 °C with rotation. Unbound protein was collected and bound protein was washed with binding buffer before being eluted from beads by boiling in 1x SDS loading buffer. Proteins were detected with α -His antibody (Bethyl).

Fluorescence recovery after photobleaching assay

FRAP techniques

FRAP experiments were conducted based on literature procedures[77, 83, 114, 115], using a Zeiss CLSM 700 confocal laser scanning microscope equipped with a 20x oil objective and Biopetechs open dish live cell imaging system (maintained at 37°C with low CO₂ atmosphere). FRAP experiments were conducted 24hr after transfection with GFP-tagged plasmid. Time-lapse image acquisition was performed on cell nuclei containing GFP-53BP1 tudor domain. Initially 53BP1 foci were scanned for 10 cycles at a 1 sec time interval and then photobleached for 20 iterations using a 405 Å laser set at 100% power and then scanned for 200 additional iterations after photobleaching. All confocal and FRAP experimental data and images were acquired and processed using Zess ZEN2011 software package with the recovery time obtained by fitting a single exponential equation. Each data point for recovery measurements was derived from the mean intensities of 3-4 different foci measurements from 3-5 different cells per experimental condition.

Cell culture

Human U2OS were seeded onto 25-mm round cover-slips at $\sim 9 \times 10^5$ cell/mL and cultured overnight at 39.5°C, 5% CO₂ in McCoy's medium with 10% FBS. The cells were then transiently transfected with a GFP-tagged 53BP1 tudor domain containing residues 1146 – 1709 which also encompasses the UDR, Accession #: NP_001135452.1 provided by the Bedford lab (MD Anderson); approximately 1 µg using Lipofectamine® 2000 (Life Technologies) and incubated overnight with or without treatment of **61** (UNC2170) at varying concentrations. After 24 hrs post transfection and compound treatment, the media was changed and then the cells were subjected to FRAP experiments.

Ionizing radiation dosing experiments

All IR treatments were conducted using a Rad Source Technology RS2000 X-ray unit located in the UNC-CH Lineberger Comprehensive Cancer Center. U2OS cell cultures that were dosed with either 2, 5, or 10 Gy of IR for instrument calibrated periods of time under standard X-ray safety conditions. Once IR-treatment was completed inside of the RS2000, IR-damaged cell cultures were then incubated 1 hr prior to FRAP experiments using confocal microscopy.

Protein NMR experiments

All NMR spectra were recorded at 298 K using a Bruker Avance III 700 MHz spectrometer equipped with a cryoprobe. The 53BP1-Tudor samples were in 25 mM sodium phosphate buffer at pH 7.5, 90% H₂O/10% D₂O, 0.3 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), 1.5 mM NaN₃. ¹H-¹⁵N HSQC-based NMR titrations were conducted by gradual addition of 10 mM UNC2170 to ¹⁵N-labeled 53BP1-tudor prepared at a starting concentration of 0.3 mM. The NMR spectra were processed and analyzed using NMRPipe and NMRView.[116, 117]

X-ray co-crystallization

Protein Expression and Purification

The expression construct for 53BP1 tudor domain (residues 1438 – 1603) in pET28-MHL vector was transformed into BL21(DE3)-V2R-pRARE2 cells. The cells were incubated in Terrific Broth medium (TB) in the presence of 50 µg/mL kanamycin and 34 µg/mL chloramphenicol at 37 °C. When the OD₆₀₀ reached 1.5, the overexpression of 53BP1 was induced by addition of isopropyl-1-thio-D-galactopyranoside (IPTG), final concentration 0.5

mM, and incubated overnight at 16 °C. Next day, the cells were harvested by centrifugation at $12,227 \times g$ (10 min, 4°C) and the cell pellets were flash frozen in liquid N₂ and stored at -80°C.

The cell pellet (18.8 g) was thawed and resuspended in 190 mL of lysis buffer (20 mM HEPES, pH 7.5, 500 mM NaCl, 5 mM imidazole, 0.5 mM TCEP, and 5% glycerol). The cell suspension was supplemented with 0.5% (w/v) CHAPS, 5 µl of benzonase (EMD Millipore, cat. no. 70746), protease inhibitor cocktail (Roche) and the cells were sonicated on ice for 5 min total (10 s pulses with 5s interruptions). The lysate was clarified by centrifugation at $20,000 \times g$, 4°C, 60 min and the resulting supernatant was filtered through 0.45 µm filter and applied onto 5 mL HisTrap HP column (GE). The column was washed with 10 CV of wash buffer (20 mM HEPES, pH 7.5, 500 mM NaCl, 40 mM imidazole, 0.5 mM TCEP, and 5% glycerol) and the protein was eluted using elution buffer (20 mM HEPES, pH 7.5, 500 mM NaCl, 250 mM imidazole, 0.5 mM TCEP, and 5% glycerol). TEV protease was added to the eluted protein (at ratio of 1 mg of TEV protease per 50 mg protein), and incubated overnight at 4°C during dialysis against 20 mM HEPES, pH 7.5, 500 mM NaCl, 0.5 mM TCEP, and 5% glycerol. The uncleaved protein (and His-tagged TEV protease) were removed by passing through 1 mL HisTrap HP column (GE). The cleaved protein was collected in the flow-through fraction. As the final purification step, the cleaved protein was applied on 16/600 Superdex 200 (GE) column equilibrated with 20 mM HEPES, pH 7.5, 150 mM NaCl, 0.5 mM TCEP. Final purification yield was 25 mg of protein per 1 L of culture and the purity of the protein was over 95%. The MW (14,103.9 Da) of the purified construct was confirmed by LC/MSD TOF (Agilent).

Crystallization

Purified 53BP1 (42 mg/mL) in 20 mM HEPES, pH 7.5, 150 mM NaCl, 0.5 mM TCEP was pre-incubated with 10 mM UNC2170 (dissolved in water) and the best crystals were

obtained by vapor diffusion technique at 20°C in sitting drops by mixing 1 μ L of protein solution with 1 μ L of reservoir solution containing 19% PEG3350, 150 mM DL-malic acid pH 7.2. For cryoprotection, the crystals were soaked in the reservoir solution supplemented with 15% ethylene glycol (v/v) for 60 s before flash freezing in liquid N₂.

X-ray Data Collection and Structure Determination

X-ray diffraction data for 53BP1 + UNC2170 was collected at 100K at beam line 19ID of Advanced Photon Source (APS), Argonne National Laboratory. Data sets were processed using the HKL-3000 suite[118]. The structures of 53BP1 + UNC2170 was solved by molecular replacement using PHASER[119] with PDB entry 2G3R as the search template. REFMAC[120] and MOLPROBITY[93] were used for structure refinement. Geometry constraints for the compound refinement were prepared with GRADE{Smart, 2011 #160} developed at Global Phasing Ltd. Graphics program COOT[121] was used for model building and visualization.

Crystallography data and refinement statistics

	53BP1+UNC2170
PDB Code	4RG2
DATA COLLECTION	
Space group	C222 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	73.5, 100.5, 83.3
α , β , γ (°)	90.0,90.0,90.0
Resolution (Å) (highest resolution shell)	50.00-1.50(1.53-1.50)
Measured reflections	377996
Unique reflections	48740
<i>R</i> _{merge}	4.9(67.0)
<i>I</i> / Σ <i>i</i>	45.0
Completeness(%)	98.1(82.9)
Redundancy	7.8(5.8)
REFINEMENT	
Resolution (Å)	50.0-1.50
No. reflections (test set)	47270(1433)
<i>R</i> _{work} / <i>R</i> _{free} (%)	22.3/19.8

No. atoms	
Protein	2064
Ligand	36
Water	199
B-factors (Å ²)	
Protein	29.0
Ligand	21.1
Water	35.9
RMSD	
Bond lengths (Å)	0.009
Bond angles (°)	1.350
Ramachandran plot % residues	
Favored	98.2
Additional allowed	1.8
Generously allowed	0
Disallowed	0

Class-Switch Recombination assay

Mice and Cells: B cell isolation and culture have been previously described [122]. In brief, primary naïve B-lymphocytes from WT C57/BL6 mouse spleens were purified by negative selection with anti-CD43 beads (Miltenyi Biotec) and cultured in RPMI 1640, 1 mM sodium pyruvate, 10% fetal bovine serum (Atlanta Biologicals,), 50 uM 2-mercaptoethanol, 25 µg/ml Lipopolysaccharide(LPS) and 5 ng/ml IL-4 (Sigma-Aldrich). Inhibitors were added at 12 hours culture and cells were analyzed 72 hours later.

Class switch recombination assay: Cultured splenocytes were stained with anti-mouse IgG1 antibodies (BD). Dead cells were excluded on the basis of forward-side scatter and propidium iodide staining. Cells were analyzed on a LSRFortessa (BD) and data was analyzed with FloJo software. Data is the summary of triplicate culture analysis of 100uM (n=4), 75uM (n=3) and 30uM (n=3) independent experiments. Statistical significance was determined by a two-tailed Student's *t* test assuming unequal variance, p values indicated.

Chromatin fractionation assay

Chromatin was isolated from HEK293 cells. Cells were lysed in EBC-1 buffer (Tris pH 7.5, 100mM NaCl, 0.5% NP-40, 1mM EDTA), nuclear pellets were collected, washed in EBC-1 and resuspended in EBC-2 (Tris pH7.5, 300mM NaCl, 5mM CaCl₂) for 30 minutes. The insoluble chromatin was pelleted and resuspended in EBC-2 for use in the release assay. This chromatin was incubated with the indicated concentration of compounds for 2 hrs at RT. The supernatant was collected and analyzed as the soluble fraction while the remaining pellet was resuspended in loading buffer, sonicated, and analyzed as the chromatin bound fraction. Anti-53BP1 (Sigma B4436) and anti-histone 3 (loading control) were used in western blots.

Foci formation assay

U2OS cells were seeded on a cover-slip 24 h prior to the experiment. Cells were pre-treated with 100 μ M of the compound for 1h and then exposed to 5 Gy of ionizing radiation (IR) using a RS-2000 Biological Irradiator (Rad Source), or not treated. Cells were rinsed with cold PBS at the indicated times after IR and fixed with 3% formaldehyde (Sigma), followed by permeabilization with 0.5% Triton X-100 (Sigma) and blocking with 10% FCS (Hyclone). The foci staining was performed in 10% FCS in PBS (anti- γ H2AX from Millipore #05-636 and anti-53BP1 from Bethyl Labs #A300-272A at a 1:1000 dilution) for 1h at room temperature followed by 3 washes with PBS. Cells were incubated with secondary antibodies conjugated to AlexaFluor 488 (Invitrogen) for green and to Alexa Fluor 596 (Invitrogen) for red foci staining. Secondary antibodies were diluted 1:500 in 10% FCS in PBS and the coverslips were incubated for 1 h at room temperature. Coverslips were washed 3 times with PBS and air-dried. After drying, the coverslips were mounted on glass slides using ProLong Gold with DAPI (Invitrogen) and

pictures were taken using a Zeiss LSM510 META confocal microscope (63x oil objective). Foci quantification was performed using ImageJ software and at least 150 cells per time point were analyzed per experiment for the calculation.

Cell permeability

Bi-directional Caco-2 Cell Permeability Assay: This cellular permeability assay was conducted by Absorption Systems (Exton, PA). Assay parameters are available through their website at <http://www.absorption.com>. The assay results for UNC2170 (**1**) for the bi-directional assay are below, where A is the top aqueous cellular layer and B is lower cellular layer.

Test Article	Direction	Recovery (%)	P _{app} (10 ⁻⁶ cm/s)			Efflux Ratio	Permeability Classification	Significant Efflux
			R1	R2	AVG			
UNC2170	A-to-B	93	19.3	17.5	18.4	1.2	High	No
	B-toA	92	20.6	22.0	21.3			

Interpretation and Advancement Potential:

Permeability classification: (P_{app} A-B) < 1.0 x 10⁻⁶ cm/s: **Low**

(P_{app} B-A) ≥ 1.0 x 10⁻⁶ cm/s: **High**

Significant Efflux: Efflux ration ≥ 2.0 and (P_{app} B-A) ≥ 1.0 x 1.0 x 10⁻⁶ cm/s

Overall:

Compound	Permeability Classification	Significant Efflux
UNC2170	High	No

Cell toxicity

The effect of 61 (UNC2170) and negative control analog 64 (UNC2892) on cell viability was determined using a CellTiter-Glo ATP detection system (Promega #7573). Ten-point, 1:3 dilution curves of compounds starting at 100 μM final concentration were diluted to 5X final concentration in PBS (vehicle control) and then 5 μL were added to 384-well white, clear bottom

tissue culture plates (Corning #3707) with a Multimek automated liquid handling device (Nanoscreen, Charleston, SC). Twenty μ l of low passage, subconfluent HEK293T (ATCC CRL-11268) grown in Dulbecco's Modified Eagle's Medium without phenol red (Gibco #31053) and supplemented with 10% Fetal Bovine Serum (GIBCO #26140) were immediately added at a density of 2,500 cells per well using a Multidrop 384 (Thermo-Fisher). Cell plates were incubated for 48 hours at 37°C and 5% CO₂, and then lysed with 25 μ l of CellTiter-Glo™ reagent. Luminescence was measured on an Envision platereader (Perkin Elmer) after 15 minutes at room temperature.

Future Work

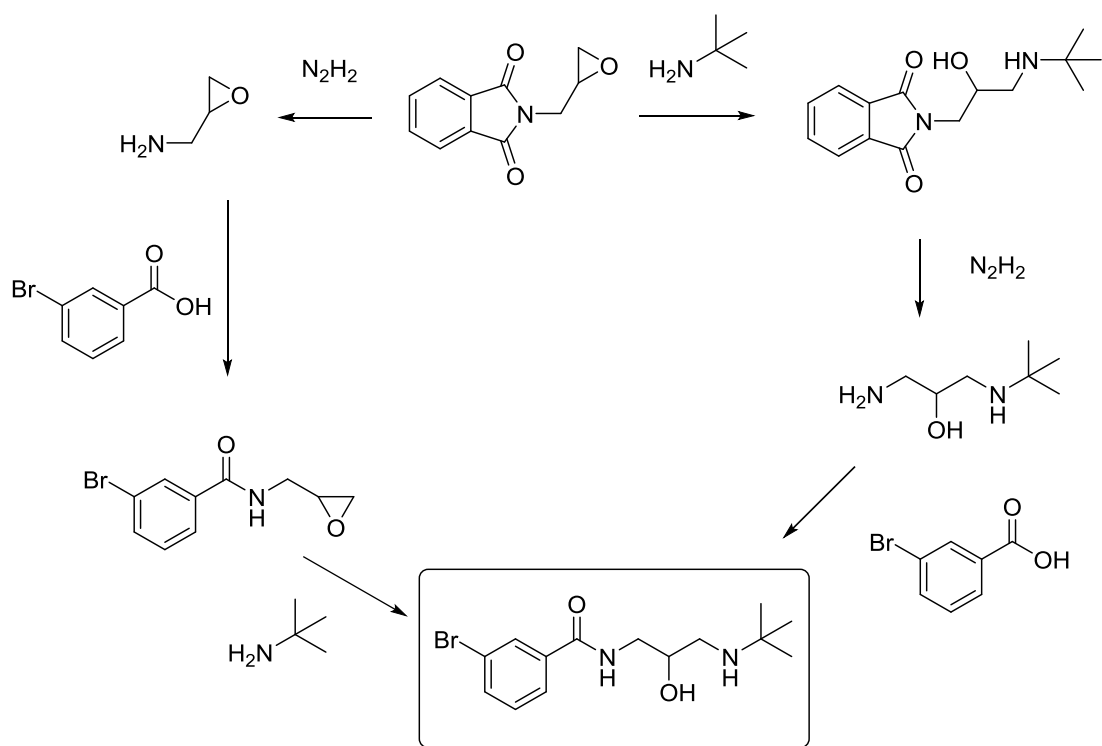
The findings described above show that development of a small molecule inhibitor for the 53BP1 TTD has proven challenging. Future efforts towards inhibitor optimization will need to take into account the protein dimerization effect that was observed in both X-ray co-crystallization and suggested by protein NMR studies. It would be optimal to develop an inhibitor that bound in a 2:1 protein to ligand ratio. The two described approaches below may allow for such an inhibitor to be obtained as their unique structural properties would potentially have the ability to prevent protein dimerization.

A small molecule/peptide like inhibitor as described in Figure 2.22 would incorporate **61** (UNC2170) with the addition of an amino acid residue. It has been shown that the N-tert butyl group of incorporate **61** (UNC2170) mimics the H4K20me2 residue of the endogenous peptide and it is proposed that the large 3-bromo group mimics the H4V21 residue. An analog that could synthetically incorporate a leucine residue off of the aromatic ring may have the potential to further mimic the H4L22 residue on the endogenous peptide substrate. Such an inhibitor would

have the potential to prevent protein dimerization by fully engaging the acidic/hydrophobic pocket and possess enough steric bulk to prevent a second 53BP1 TTD protein from dimerizing.

A second approach would focus on the aliphatic propyl chain of **61** (UNC2170). It was observed that the propyl chain is sandwiched between the two proteins in the dimer. If an aliphatic or hydrogen bond donor group could be introduced into that chain at the 2-position, then dimerization could be prevented. Use of an oxirane containing phthalamide starting material would allow for two different routes to be approached to obtain the final aliphatic functionalized inhibitor.

H4 peptide sequence:
...RHRKme2VLD...



Fragment Based Screen for Improvement of 53BP1 Ligands

Rationale

Early structure-based drug design efforts towards development of a small molecule inhibitor of 53BP1 TTD unfortunately did not provide a highly potent molecule (Chapter 2, sections B and C). Efforts were successful in increasing the potency of the known fragment-like small molecule ligand (UNC2170), yet a more potent molecule was desired. Alternative methods toward discovering a more potent molecule, or discovering an additional molecular fragment that could be coupled to UNC2170 were desired. These efforts were conducted in parallel with ongoing UNC2170 SAR efforts and it was decided to use a known small molecule ligand of 53BP1 TTD as a basis for this new small molecule fragment screen. Initially, when this fragment-based screen was developed, an x-ray co-crystal of UNC2170 and the 53BP1 TTD was not solved and the true binding pose of the small molecule ligand was unknown. Because of the modest affinity of UNC2170 ($29 \pm 7.4 \mu\text{M}$) and its fragment-like nature (MW = 313.24, ligand efficiency = 0.35, lipophilic ligand efficiency = 1.5)[106, 107], it was hypothesized that it could be used in a fragment screen to identify other small molecule fragments that worked in a synergistic manner to further inhibit 53BP1 TTD. Initial assay validation methods showed low precision and reproducibility of IC_{50} values for UNC2170. Therefore it was decided that a previously discovered small molecule chemical probe, UNC1215, be used that possessed potency against 53BP1 TTD, but that had consistent AlphaScreen IC_{50} values.

The existing fragment screening methods were modified by incorporating UNC1215 into the fragment screening process. The overall goal of this method was to initially screen a fragment collection, provided by SGC Toronto, for their single fragment potency against 53BP1

TTD, followed by a screen where each individual fragment was combined with UNC1215 to discover fragments that might bind in a synergistic manner to increase their combined potency against 53BP1. That work will be described below.

Current Fragment-based Screening Methods

The concept of discovering and then synthetically combining small molecule fragments with weak affinity toward a defined enzyme/protein target to create a high affinity inhibitor was first proposed in the early 1980s. This idea was formulated by Prof. William Jencks and generally stated that the total binding energy of an inhibitory ligand (generated by synthetically combining small molecule fragments) is composed of the combined binding energies of its component fragments and their interactions with the target. This concept was further improved in both academia and in industry and the process of fragment-based drug design (FBDD) was coined in the early 1990s. FBDD was eventually shown to be a proven inhibitor discovery technique in 1996 when scientists at the Abbott Laboratories successfully created a highly selective potent inhibitor for BCL-2/BCL-XL. Their efforts incorporated drug discovery screening methods including high-throughput screening (HTS) technologies, X-ray co-crystallization, and primarily NMR screening of fragments bound to the target protein. The combined use of these screening and spectroscopic techniques quickly identified small molecule fragments and determined their defined binding pose within BCL-2. Synthetic chemistry methods were then used to modify and eventually combine these fragments to create the highly selective and potent inhibitor ABT-737. The process of FBDD has been widely accepted in the therapeutic development arena and has allowed for the creation of highly selective and potent inhibitors for various enzyme and protein targets including kinases, metalloproteinases, and

polymerases. Each of these screening discoveries is based on the process of discovering small molecule fragments that possess low mM to high μM potency for a defined target, determining their binding pose, and then synthetically combining or ‘growing’ the fragments to create a highly selective and potent inhibitor.

There are several fragment based screening methods that are commonly used in FBDD efforts to determine binding potency of fragments and their binding locations on an enzyme/protein target. The most commonly used technique is target-based NMR spectroscopy; also known as SAR by NMR. This technique screens the protein target that is isotopically labeled (C^{13} and N^{15} isotopes) in combination with “pooled” fragments to identify sets of fragments that possess affinity for the target. The “pool” of fragments is then deconvoluted to determine sets of those fragments that provide affinity and then a binding pose of the fragments when bound to target. This technique provides highly detailed structural data of the binding interactions between fragments and target and is very useful for later medicinal chemistry efforts. Another common technique is X-ray co-crystallization.[123-126] This technique also provides highly detailed structural information on fragment-enzyme/protein interactions. Two techniques that provide measures of binding kinetics and thermodynamic information include surface plasmon resonance (SPR) and isothermal titration calorimetry (ITC). These two techniques unfortunately do not provide structural information and require either immobilization of target or fragment and high protein concentrations. Another technique that is known as high concentration screening uses high concentration of fragments and screens for inhibition of the biochemical activity of a defined enzyme/protein target. This technique requires a validated biochemical assay for the target and high concentrations of fragment. Other methods include mass-

spectrometry to show evidence of fragment binding to target and virtual screening of fragment collections by *in silico* docking programs. Combinations of these techniques are used in the overall process of FBDD as they each have their own strengths and weaknesses and provide different types of structural and binding data.

Fragment-based Screening Platform Design

Exploration of existing fragment-based drug discovery methods that have been used to discover an initial fragment hit, or a combination of fragments that could be combined into a single molecule provided inspiration for this modified fragment-based screening assay. The hypothesis for undertaking this modified fragment-based screening assay is as follows. The assay will use the AlphaScreen bead-based avidity assay platform where inhibition of protein-histone peptide interaction by fragments results in a decrease in the chemiluminescence read-out to provide a measure of fragment's potency. Since an NMR structure of UNC2170 bound to 53BP1 TTD was not solved when we initiated this approach, traditional NMR fragment screening techniques could not be used. However, there was an existing AlphaScreen assay for 53BP1 TTD in place that used low concentrations with an adequate level of precision and robustness that would allow for easy incorporation into a fragment based screen. The overall screen was conducted both in the presence and absence of UNC1215. The reason behind screening in the absence of UNC1215 is that it will allow for a comparison of IC_{50} values as a means to determine if a fragment is bound in a competitive or non-competitive manner to UNC1215. If there is non-competitive binding, then there might be synergy. If not, there might be no improvement in potency. The initial screen where UNC1215 is absent will also help to determine new types of fragments that may provide varying degrees of potency. In the assay, where

UNC1215 is present, each fragment will be screened at 150 μM in the presence of UNC1215 (at its EC_{20} concentration of 10 μM) in order to determine if there is an additive effect between the two compounds to provide an increase in inhibition of 53BP1 TTD. Once a fragment or combination of fragments is discovered that shows increased potency compared to UNC1215 alone, then determination of their binding pose will be conducted using X-ray co-crystallization. If a defined binding pose of the fragment(s) in combination with UNC1215 can be resolved, then synthetic chemistry efforts will be used to combine the fragment(s) to UNC1215 to create a higher potency inhibitor.

Description of Structural Genomics Consortium (SGC) fragment library

A small molecule fragment library composed of 2500 fragments with molecular weights below 300 g/mol at 15 mM concentration in DMSO was provided by the SGC Toronto group. This library has fragments of varying molecule weights and was composed of acids, bases, hydrophobic, aromatic, and hydrophilic fragments that are all commercially available. Each fragment in the library possesses functional groups that allow for chemical modification to be achieved through straightforward synthetic methods such as amide bond coupling, metal-catalyzed aromatic ring substitutions, and $\text{S}_{\text{N}}2$ reactions.

Results and discussion

Overall, 1846 compounds from the fragment collection were screened in duplicate in two different assays. The initial assay screened the fragments alone against the 53BP1 TTD and the second assay screened a combination of a single fragment and UNC1215 to determine if there was any synergy between the two compounds. The remaining 654 fragments were unable to be

screened due to solubility issues with the biological assay due to the low DMSO tolerance of 53BP1.

Validation of assay method

The fragment screen was conducted using a modified form of the existing 53BP1 TTD AlphaScreen assay. This fragment screening assay was validated for high through-put screen (HTS) quality through the use of two different assay controls. DMSO served as the negative control that provided a high signal and a positive control used biotin. The solution based biotin binds to the streptavidin linked donor bead in the assay and prevents binding of the biotin-conjugated target peptide sequence to the Kme binding protein. Additionally, a standard curve for UNC1215 was generated on each plate to verify that an accurate IC_{50} measurement for UNC1215 was obtained in order to confirm that UNC1215's IC_{50} values were consistent across all screen fragment plates. The fragment based screen was setup on plates as outlined in Figure 2.23.

Figure 2.23: Plate Map of Fragment Based Screen

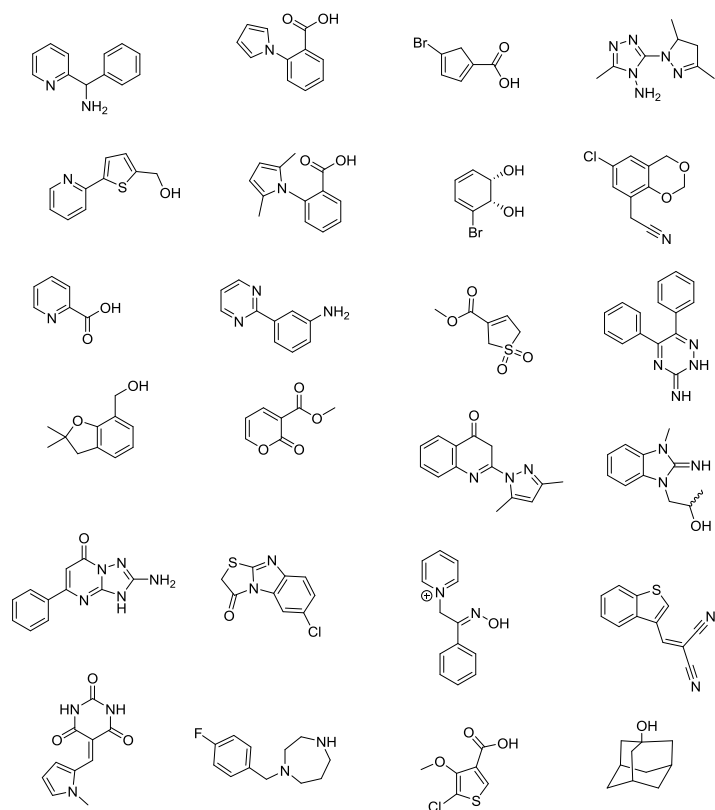
Fragments arranged from top to bottom (2-fold serial dilution from 100 uM)	Biotin
	DMSO + UNC1215
	UNC1250 IC50 Dose Response
	DMSO

This plate setup was used for the fragment-only screen with the absence of UNC1215 in columns 1-20, 22, and 24. It was also used in a counterscreen assay where no protein was added to any of the columns so as to verify that the fragments did not interfere with the components of the assay. The counterscreens were run in duplicate and provided Z' values of 0.86 and 0.73. These Z' values are well within the measure of a robust HTS assay where the generally accepted minimum Z' value is ~ 0.5 . The Z' value for the UNC1215 + fragment assays were 0.74 and 0.81 and the counterscreen assay showed no fragments that interfered with the assay components.

Initial hits from screen of SGC fragment library

The initial fragment-only screen on the SGC fragment library provided 30 compounds that showed 100% inhibition of 53BP1 activity at 150 μM . The chemical structures of these initial hits were visually screened to determine if there were any chemically promiscuous functional groups present or known non-selective binders present, Figure 2.24.

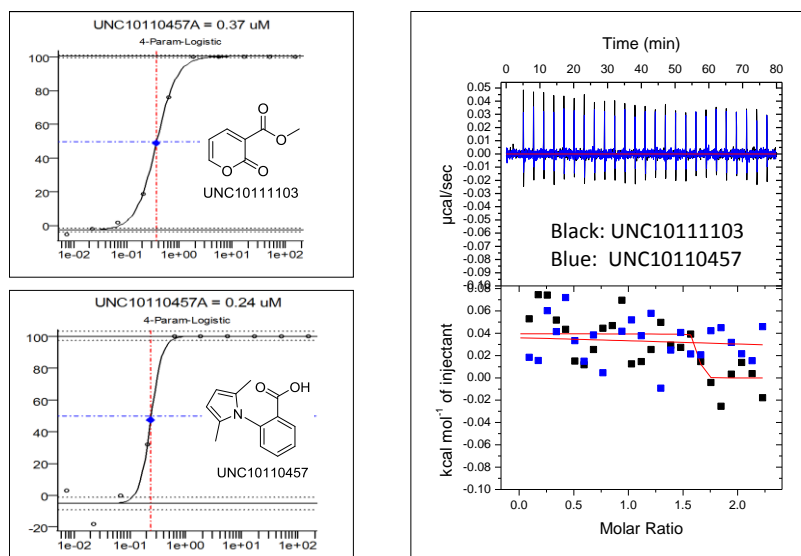
Figure 2.24: Initial Fragment Hits



Of the initial 30 hits, 24 fragments (Figure 2.24) were selected to conduct dose-response analysis for their inhibition of 53BP1 +/- UNC1215. The other 6 fragments were excluded as they possessed functional groups that were potentially problematic in biological assays; these functional groups included quinones, Michael acceptors, or highly electrophilic centers that might be susceptible to nucleophilic attack by either protein or other assay components which could lead to a false-positive assay signal.

The 24 selected fragments were assayed in the absence of UNC1215 in a dose-response assay using the AlphaScreen to gain a true measure of their IC_{50} values. Two of the fragments were observed to possess very high potency against 53BP1 TTD, Figure 2.25.

Figure 2.25: Representative IC₅₀ curves and ITC curves of high potency fragment “Hits”



A modified pyran fragment, **UNC10110457**, had an IC₅₀ = 0.78 ± 0.55 μM and a pyrrrole-containing fragment, **UNC10111103**, had an IC₅₀ = 0.30 ± 0.07 μM. ITC measurements were taken for these two fragments against the 53BP1 TTD and unfortunately showed no binding activity for either fragment, Figure 2.25. It is uncertain why the ITC measurements disagreed with the AlphaScreen assay values. It was noted that **UNC10111103** was not active in the counterscreen and therefore did not interfere with the assay. Unfortunately, **UNC10110457** was active in the counterscreen with a IC₅₀ = 29 μM. It is predicted that the high bead avidity properties of the AlphaScreen assay is more sensitive to binding interactions than other types of screening assays. Therefore, a compound that shows very high potency in AlphaScreen may not show the same degree of potency as in other assays such as ITC. The other 22 fragments showed modest potency values, while UNC1215 and UNC2170 provided potency measurements in agreement with their known IC₅₀ values against 53BP1 TTD.

Outcome of assay upon addition of UNC1215

The next step in this assay was to determine if there was any synergistic activity between any of the identified fragment “hits” and a known inhibitor of 53BP1 TTD. Once the IC₅₀ values for each of the initial 24 fragments was measured, an additional assay was run in combination with UNC1215. The concentration of UNC1215 was held at 10 µM and each fragment was assayed at 150 µM. The fragment screening concentration was 15-fold higher than that of UNC1215 to allow for identification of any synergistic active hits. Once this combination assay was run, it was observed that there was unfortunately no combined synergistic activity between the any of the identified fragments and UNC1215.

Conclusions

Overall, this fragment screen provided two interesting fragments that showed high potency levels against the 53BP1 TTD in the AlphaScreen assay, but unfortunately did not confirm that potency in ITC. Unfortunately, the goal of identifying a fragment that worked in synergy with a previously identified 53BP1 TTD inhibitor proved unsuccessful. This screening assay was validated and was shown to possess high efficiency given the Z' values obtained during assay optimization experiments. It is believed that the assay has promise in aiding in the process of discovering new fragments that could be combined with ligands that are in the process of compound optimization.

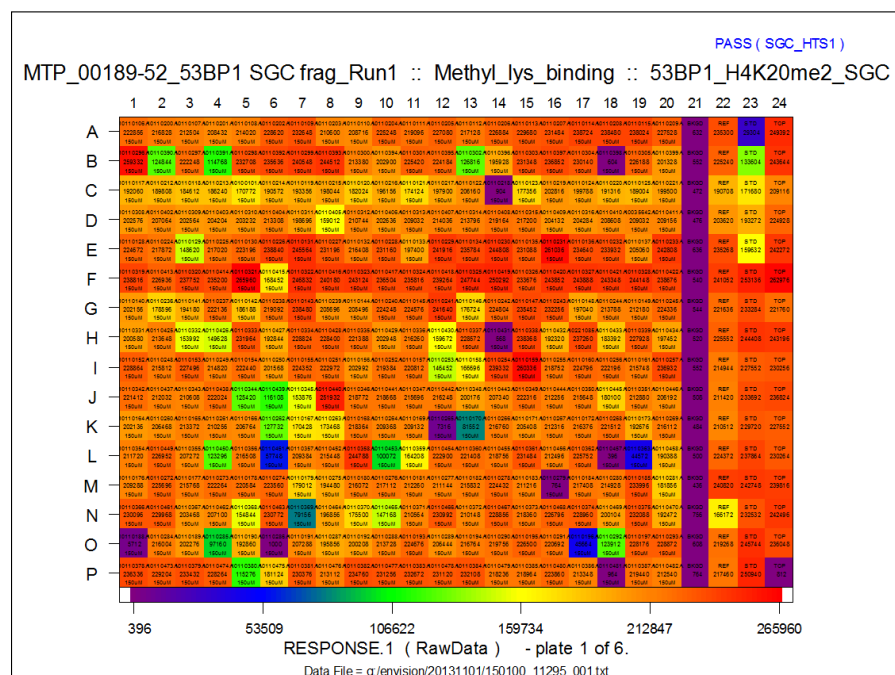
Experimental section

Preparation of fragment-only and fragment + UNC1215 IC50 plates for AlphaScreen assay

Preparation of fragment plates for the screening assay began by preparing dilution plates from the parent fragment plates received from SGC Toronto. Dilution plates were prepared using a Nano-drop instrument (ThermoScientific) to stamp out 1 μ L of fragments from the parent fragment plates (at 15 mM in DMSO) into columns 1 – 20 in XC50 plates, followed by addition of 9 μ L of buffer (150 mM NaCl, 25 mM TRIS, 2 mM DTT) to the same columns to provide a fragment concentration of 1.5 mM. Into column 21 was added 10 μ L of biotin control (150 mM NaCl, 25 mM TRIS, 2 mM DTT, 1 mM DMSO), into column 24 was added 10 μ L of DMSO control (150 mM NaCl, 25 mM TRIS, 2 mM DTT, 1 mM DMSO, 0.1 mM biotin). In column 22 was added 10 μ L of DMSO and 10 μ L of UNC1215 at 10 μ M, and in column 23 was added 10 μ L of a freshly prepared UNC1215 standard curve (2-fold dilution from 10 mM in distilled water). The plates were placed onto a centrifuge and spun for 2 min at 2000 rpm. Then 1 μ L from these dilution plates was transferred to a Proxi plate (performed in duplicate). In screening runs in the absence of UNC1215, to columns 1 - 24 were added 9 μ L of a prepared 53BP1 TTD + H4 peptide solution (0.13 μ M 53BP1 TTD, 0.2 μ M peptide P565 in buffer). For the screens that included UNC1215, a final concentration of 10 μ M was pre-incubated with 53BP1 TTD + H4 peptide solution for 30 min prior and added to the proxy plates in the same manner as described above. The prepared plate was incubated at room temperature for 30 min and then 2 μ L of a prepared AlphaScreen bead mixture (final concentration 7.5 μ g/ μ L donor bead + 7.5 μ g/ μ L acceptor bead in buffer) was added to all columns. The plate was then incubated in the dark for 20 min. After incubation, the plates for both experiments were read on an EnVision

multi-label reader equipped with an HTS AlphaScreen laser (Perkin Elmer). The IC₅₀ values reported are the average of at least 2 values \pm the standard deviation. When IC₅₀ values for a single compound were not all active (< 100 μ M) or inactive (> 100 μ M), the IC₅₀ values were calculated using 4-paramter curve fitting (GraphPad Prism 5) from replicate runs using averaged response values for each compound concentration. A representative heat map of a fragment assay plate that passed all controls, Figure 2.26.

Figure 2.26: Representative image of fragment screen plate map and IC₅₀ curves of hit compounds



Isothermal titration calorimetry (ITC) binding experiments

All ITC measurements were recorded at 25°C with an AutoITC₂₀₀ microcalorimeter (MicroCal Inc.). All protein and compound stock samples were prepared in the target buffer (25 mM Tris-HCl, pH 8, 150 mM NaCl, and 2 mM β -mercaptoethanol), and then diluted in the same

buffer to achieve the desired concentrations: 90 μ M protein and 1 mM compound depending on the expected dissociation constant. The concentration of protein stock solutions were established using the Edelhoch method, whereas 10 mM compound stock solutions were prepared gravimetrically based on molecular weight. A typical experiment included a single 0.2 μ L compound injection into a 200 μ L cell filled with protein, followed by 25 subsequent 1.5 μ L injections of compound. Injections were performed with a spacing of 180 seconds and a reference power of 8 μ cal/sec. The titration data was analyzed using Origin Software (MicroCal Inc.) by non-linear least squares, fitting the heats of binding as a function of the compound:protein ratio. The data were fit based on a one set of sites model.

Future Directions

This fragment-based screening approach shows promise, in that; it was able to identify unique fragments that were able to show inhibitory activity against the 53BP1 TTD, But were inactive by ITC. Unfortunately, it was unable to show any synergistic activity in combination with a known inhibitor of the 53BP1 TTD. There are several possible explanations for these findings. One explanation for the lack of synergistic activity was that the 53BP1 TTD is shown to form a 2:1 protein to inhibitor dimer around at least one known small molecule ligand, UNC2170. This dimerization property of the protein target was unknown when this assay was designed and conducted. Unfortunately, there is no x-ray co-crystal structure of UNC1215 bound to 53BP1 TTD. An *in silico* model could be used to dock UNC1215 onto 53BP1 TTD and predict whether it binds in a similar manner to UNC2170. If UNC1215 binds to 53BP1 TTD in a manner similar to UNC2170, then protein dimerization could be occurring with UNC1215. If dimerization does occur in this fragment screen, then this would present a problem for the

AlphaScreen assay platform. This assay is currently setup to predict a 1:1 binding ratio of protein to inhibitor and protein dimerization would lead to inaccurate inhibitory measurements. If docking models show that UNC1215 binds in the same manner as UNC2170, then ITC measurements of UNC1215 against 53BP1 TTD would provide additional evidence of protein stoichiometry. Another potential explanation for the lack of synergistic activity is that the fragments bind competitively with UNC1215 for 53BP1. If the fragments bound in a non-competitive manner with UNC1215, then increased inhibitory measurements might have been observed. Under optimal assay conditions, it would have been favorable to use the smaller inhibitory ligand UNC2170 as its smaller size would potentially have allowed for binding of fragments to other locations within the 53BP1 TTD binding site; within the acidic/hydrophobic pocket or Arg19 binding pocket. Future efforts with this assay should try to modulate it to use UNC2170 and re-screen with the same fragment collection to obtain fragments that bind in different locations on the protein. If fragments are identified with this modified assay, then X-ray co-crystallization experiments with UNC2170 and the fragment should be conducted to determine the binding pose and location of each. Once a binding pose for each is determined, then synthetic chemistry methods could then be utilized to insert a linker region between the two inhibitory components to develop an improved inhibitor small molecule. An additional method that could be incorporated is the use of NMR fragment screening techniques. Now that a NMR structure of UNC2170 bound to 53BP1 TTD has been developed, a NMR screen of UNC2170 combined with the identified fragments and others from the SGC fragment library could be conducted.

Chapter III:

Design, Synthesis, and Characterization of Peptide-based Inhibitors for the Methyl-lysine Binders, PHF1 and PHF19

Biology and Clinical Significance of PHF1 and PHF19

Biological Function of PHF1 and PHF19

The active expression and repression of genes is a highly regulated process that is in part modulated through the writing, reading, and erasing of post-translational modifications (PTMs) on chromatin. Different PTMs are associated with active transcription such as H3K4me3 and H3K36me3, and also with repressed chromatin, such as H3K27me3.[2, 21, 127] These PTMs are written, read, or erased by larger protein complexes such as Polycomb repressive complex 2 (PRC2). PRC2 has a tetrameric core of proteins consisting of the catalytic Enhancer of zeste homolog 2 (EZH2), Suppressor of zeste 12 (SUZ12), Embryonic ectoderm development (EED), and the histone binding proteins RBAP46 and NURF55 (NURF55/RbAp48).[25, 26, 128-132] This core set of proteins works to actively write and maintain the H3K27me3 repressive mark on chromatin leading to gene silencing. The PRC2 complex is assisted by a small number of associating proteins including JARID2, AEBP2, PHF1, and PHF19.[24, 26, 128] These associating proteins are not always present in PRC2, but assist in modulating its ability to

identify and target active chromatin, induce movement of this complex into a gene locus that is being actively transcribed, and then act to tether the complex at that location until a repressive PTM is written and established. PHF1 and PHF19 play central roles in these functions through the use of a common Tudor domain that is able to read and bind the H3K36me3 PTM.[24-26, 129] This binding event recruits the PRC2 complex into the active locus and allows for the methyltransferase activity of SUZ1 and EZH2 to trimethylate H3K27.[25] Further enzymatic activity and instillation of this repressive mark leads to the removal of the H3K36me3 mark and an overall gene-silencing event. Both PHF1 and PHF19 act to identify H3K36me3, assist in insertion of the PRC2 complex, participate in anchoring its associated proteins within gene loci undergoing active transcription and allows for the PRC2 complex to install a repressive PTM within that region at H2K27me3, resulting in a gene silencing event.[24-26] Interestingly, even though PHF1 and PHF19 share a very high degree of structural similarity, they are not found together in the same PRC2 complex and differ in that PHF1 is associated with PRC2 complexes that have JARID2. PHF19 is only found in PRC2 complexes that lack JARID2.[24] Additionally, PHF1 has been shown to be active in NHEJ through the binding the broken ends of DNA double strand breaks.[26] However, its roles in the repair of double strand DNA breaks and in the maintenance of genomic stability is not thorough understood.

Clinical significance of PHF1 and PHF19

PHF1 and PHF19 play an active role in the modulation of PRC2 activity that in turn controls various gene expression and repression outcomes. Active PRC2 can induce gene-silencing events that have the potential to increase the expression of deleterious genes directly related to various diseases. This is caused by PRC2's repression of other genes, such as

repression of tumor-suppressing gene that lead to up-regulation of deleterious genes. It has been shown that in various types of human tumors there is an increased level of both PHF1 and PHF19 present.[38, 39] These increased levels may be leading to increased levels of PRC2 activity, and in turn increased levels of H3K27me3 PTMs on essential tumor suppressing gene loci. Increased levels of repression of these tumor-suppressing genes can lead to increased tumor growth and potentially downstream metastasis events. The core biological function of both PHF1 and PHF19; *i.e.* the binding of their Tudor domain to H3K36me3 on transcriptionally active chromatin; represents a potential target for the development of novel cancer therapeutics. It has been shown that knockdown of either PHF1 or PHF19 in various cell lines leads to global decrease in the amount of the H3K27me3.[24] Inhibitors of the Tudor domain of either of these proteins would provide a chemical tool to test whether their inhibitory characteristics could lead to a favorable therapeutic outcome in the treatment of tumors.

Purpose of this work

The purpose of this work is to develop a low molecular weight, peptide ligand for the methyl-lysine binding proteins PHF1 and PHF19 based upon the co-crystal structure of their endogenous chromatin histone peptide binding substrate. This work focuses on conducting an initial Structure Activity Relationship (SAR) study to develop a short, low molecular weight peptide via truncating the endogenous histone 3 peptide into a shorter 6- or 7-mer peptide that binds specifically within the Tudor domain of PHF1 and PHF19 with emphasis on determining a quaternary amine mimic for the tri-methyl lysine 36 residue.

Development of peptide-based inhibitors

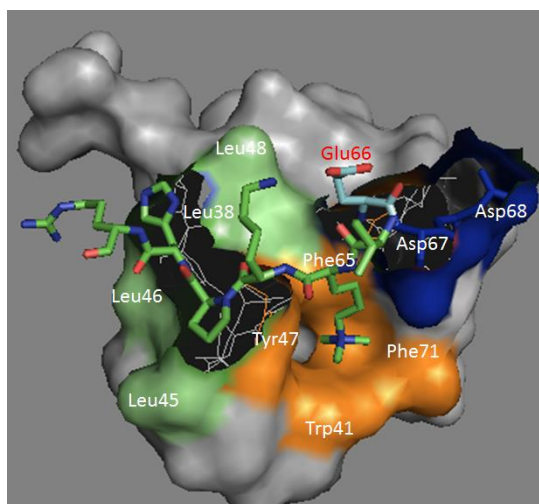
Results and discussion

Structure based analysis of the Tudor domain of PHF1 and PHF19

The methyl-lysine binding protein PHF1 is a ~50 kDa protein that contains two plant homeodomain (PHD) fingers and a single Tudor domain. A thorough structural analysis of the Tudor domain bound to an H3K36Kme2 peptide conducted by Musselman *et al* provided deeper understanding into the key structural elements of the Tudor domain.[26] The domain is made up of three unique regions that work in tandem to selectively bind its endogenous peptide substrate. The regions are shown in Figure 3.1 and are an aromatic methyl-lysine binding cage (orange), a hydrophobic patch (green), and an acidic groove (blue). The Tudor domain adopts a five-stranded β -barrel conformation providing a defined binding pocket for the trimethylated H3K36 residue within the aromatic cage. The aromatic cage is located between the acidic and hydrophobic regions of this domain and is highly selective for the trimethylated H3K36 residue. The Tudor domain binds the histone peptide in an extended conformation with C-terminal portion of the peptide interacting with the hydrophobic patch, the trimethylated K36 residue bound within the aromatic cage, and the N-terminal sequence interacting with the acidic groove. The aromatic cage consists of the Tyr47, Trp41, Phe65, and Phe71 residues that work together in binding the K36me3 residue through cation- π and hydrophobic interactions. Structural studies contrasting the bound and unbound Tudor domain show that the aromatic cage is not preformed, but adopts its conformation upon binding peptide substrate. The hydrophobic patch consists of a group of four leucine residues (Leu38, Leu45, Leu46, and Leu48) and forms a hydrophobic “wall” that allows for hydrophobic interactions between peptide substrate and protein. The crystal structure

shows an intermolecular hydrogen bond between the backbone NH bond of His39 and the carbonyl group of Leu46. The acidic groove consists of Glu66, Asp67, and Asp68. The peptide residues Thr32, Gly33, and Gly34 bind within this groove and an intermolecular hydrogen bond was observed between the NH bond of Val35 and one of the oxygen of the carboxylate of Glu66. This Glu66 residue also form a salt bridge via its other carboxylate oxygen group with the ammonium group of Lys37 of the histone peptide. The Arg34 peptide residue is shown to point away from the hydrophobic patch and is solvent exposed. These initial binding interactions provide a basis for the rationale behind amino acid substitutions in the peptide SAR.

Figure 3.1: Binding regions of PHF1 with endogenous H3K36me3 peptide substrate bound, PDB 4HCZ



The PHF19 structure is very similar to PHF1 and binds the H3 peptide in the same fashion with all of the same binding interactions. Peptide-based inhibitors for both targets will be developed using a single core peptide structure. Given that both proteins bind the same endogenous peptide sequence, it is hypothesized that a single core peptide sequence that is specific for both proteins is achievable through structure-based design techniques.

Design of peptide ligands for PHF1 and PHF19

Structure-based techniques were used to design peptide-based inhibitors for PHF1 and PHF19. Therefore, the high degree of structural similarity between these two proteins was used to an advantage. As described previously, a single core peptide structure can selectively bind the Tudor domains of both proteins. Secondly, medicinal chemistry efforts were conducted through analysis of the core section of the H3 histone peptide that PHF1 and PHF19 are known to bind. Key peptide residue interactions with the protein's surface were maintained with specific amino acid substitutions that were hypothesized to increase the affinity of the peptide to the Tudor domain. An additional goal of the peptide SAR was to discover an appropriate quaternary amine mimic to replace the K36me3 residue which maintains a permanent cationic charge. This cationic residue is a barrier to cell membrane permeability, therefore efforts were made to replace it with a tertiary amine mimic that was highly selective for binding within the aromatic cage of the Tudor domain. A final goal was to truncate the large histone peptide into a shorter 6- or 7-mer peptide sequence with high affinity and selectivity for the Tudor domain of PHF1 and PHF19. Medicinal chemistry efforts toward achieving these goals are described below.

Analysis of endogenous H3K36Kme3 chromatin peptide and protein binding pocket residues

Initial medicinal chemistry efforts focused on the analysis of the endogenous histone peptide substrate of PHF1 and PHF19. Both targets are known to bind the H3K36me3 and this shared histone peptide substrate provided an initial core residue sequence for inhibitor development. Previously conducted experiments determined the binding specificity, specific peptide-protein binding interactions, and ligand affinity for the protein target using several H3

peptide sequences of varying amino acid sequence lengths and peptide regions. These experiments were conducted on the purified Tudor domain of PHF1 that were produced from different plasmid constructs of varying residue lengths. Peptide pull down assays used a Tudor domain containing residues 14-87, the X-ray co-crystallization structures, data from isothermal titration calorimetry (ITC) experiments, data from nuclear magnetic resonance (NMR) titration experiments contained from residues 28-87, and IC₅₀ measurements from the AlphaScreen assay used on a Tudor domain with residues 26-87. These varying domain sizes may have caused slight variations in the affinity measurements, but overall, the core aromatic cage, and the acidic and hydrophobic patches of the domain were present. Experiments conducted on PHF19 used Tudor domain residues 37-95 for the same set of experiments. Additionally, the H3 histone peptide sequence lengths and sequence regions used in these experiments also varied. The NMR titration and X-ray co-crystallization experiments for PHF1 used peptide residues 31-40. However, the pull down assays used peptide residues 21-44, while ITC measurements used peptide residues 27-46. Overall analysis of the peptide sequences used in the above studies demonstrates how the core peptide region shown to bind within the Tudor domain consisted of peptide residues 35-41. The N-terminal Val35 residue was shown to bind within the acidic groove of the Tudor domain, Lys36me3 within the aromatic cage, and Lys37, Pro38, His39 bound within the hydrophobic pocket. Arg40 does not directly interact with the protein's surface but is solvent exposed in the x-ray co-crystal. Try41 was not clearly evident in the x-ray, but was included to provide a UV-active residue, allowing for increased ease in peptide purification. Additional key peptide-protein residue interactions that were noted were a salt bridge between peptide Lys37 and protein Glu66 and H-bonds that exist between the peptide backbone amide N-H of His39 and Val35 with protein carbonyl oxygen.

An alanine scan of the H3 peptide sequence was conducted to determine which residues play an essential role in binding to the Tudor domain of PHF1. (Kycia *et al*) A Celluspot peptide array was used where individual peptide residues between 30-44 were separately mutated to an alanine in individual peptides via a traditional alanine scan, all while maintaining the remaining endogenous residue sequence.[129] These prepared peptide sequences were then used in a PHF1 pull down assay to determine how Tudor domain binding affinity varied for each peptide depending on which amino acid was mutated. This study showed four residues; outside of K36me3; were essential in the peptide sequence allowing for PHF1 binding. These residues were Pro38, Thr32, Val35, and His39, with Pro38 providing the greatest role when influencing binding. Lastly, when Arg40 was mutated to alanine, there was no effect on binding.[129] This result agrees with the x-ray co-crystal structure that shows that Arg40 is solvent exposed. All but one of these residues, Tyr41, is included in the core peptide inhibitor sequence used to begin peptide SAR studies.

Analysis of endogenous histone substrate binding affinity

The identification of the endogenous histone peptide substrate for both PHF1 and PHF19 was previously conducted in two different studies by using histone peptide microarrays.[25, 26] Both studies showed the target of both proteins was H3K36me3. The PHF1 study showed that the Tudor domain was essential for binding that mark through the use of an inactive Tudor domain mutant, Y47A.[25] Additionally, it was shown that both PHF1 and PHF19 showed a distinct preference for H3K36me3 over the di- and monomethyl-lysine peptides and no binding for non-methylated lysine.[25] It was observed that the Tudor domain of PHF1 bound a 12-mer peptide sequence containing the H3K36me3 mark with a $K_d = 36 \pm 4 \mu\text{M}$, H3K36me2 with a K_d

= 167 ± 34 μ M, and had no affinity for the non-methylated peptide.[26] These values were confirmed with a 19-mer peptide containing the same marks. The Tudor domain of PHF19 was shown in ITC experiments to bind a 19-mer H3K36me3 peptide with a $K_d = 6.2 \pm 0.6$ μ M and provided no binding for either H3K36me1 and non-methylated H3K36.[25] These qualitative peptide binding studies were validated for both proteins by NMR titration measurements.

Structure-Activity Relationship (SAR) studies

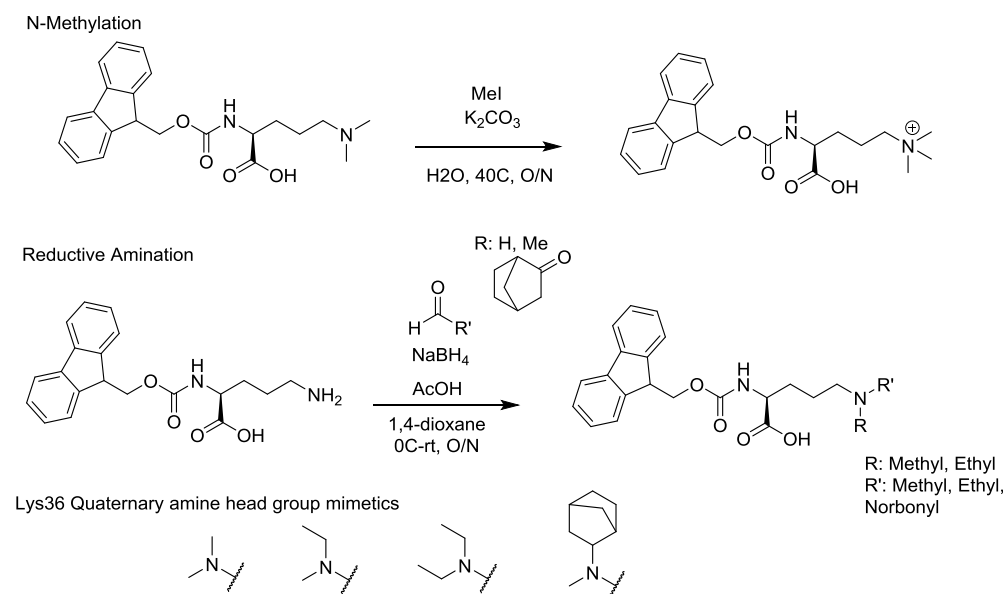
The structural properties and endogenous peptide binding characteristics for both proteins, PHF1 and PHF19, were primarily based on X-ray structure of PHF1 bound to endogenous H3K36me3 peptide substrate.[26] Additionally, an NMR titration studies showing key the interactions in a solution phase environment. Additionally, work based on an alanine scan of the H3K36 histone peptide using a Celluspot peptide array assay provided additional evidence on the various residues that PHF1 would tolerate substitutions for within the H3 histone tail. These findings allowed for a more sound design of the peptide SAR.

Preparation of Non-natural H3K36me3 Mimetics

Analysis of both the X-ray co-crystallization of the H3K36me3 peptide with the Tudor domain of PHF1 (PDB 4HCZ) and the apo crystal structure (PDB 2E5P) of the PHF1 Tudor domain showed that the aromatic cage Kme binding pocket is not static, but instead undergoes a substrate-induced conformational change upon binding the peptide ligand.[26] Within the aromatic pocket, there are cation- π and hydrophobic interactions, as previously outlined. Because this pocket undergoes a conformational change when binding a peptide substrate, it was hypothesized that the aromatic Kme binding pocket would be an initial area to target and to determine if the positively charged K36me3 residue could be substituted for an alternative non-

cationic quaternary amine functional group. Ultimately, this functional group should mimic the trimethyl-lysine residue while maintaining peptide binding affinity. Experiments were planned to determine the steric tolerance of the Kme binding pocket by synthesizing a small series of four different K36me3 analogs, Figure 3.2. These analogs were prepared from N- α -Fmoc-Lysine via standard medicinal chemistry techniques such as N-methylations and reductive aminations as shown in Figure 3.2.

Figure 3.2: Preparation of Non-natural H3K36me3 Mimetics

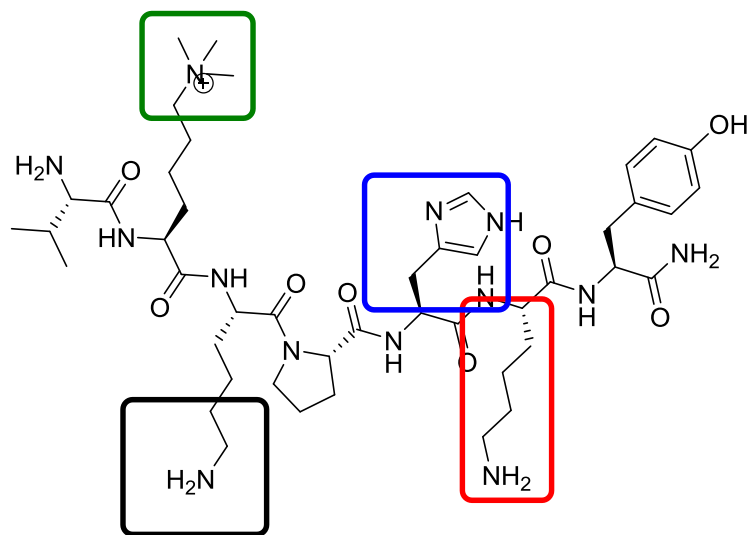


Each N- α -Fmoc protected analog was prepared, purified, and characterized prior to incorporation into the various peptides of the two peptide SAR series that were synthesized.

Initial SAR round of peptide ligand amino acid substitutions

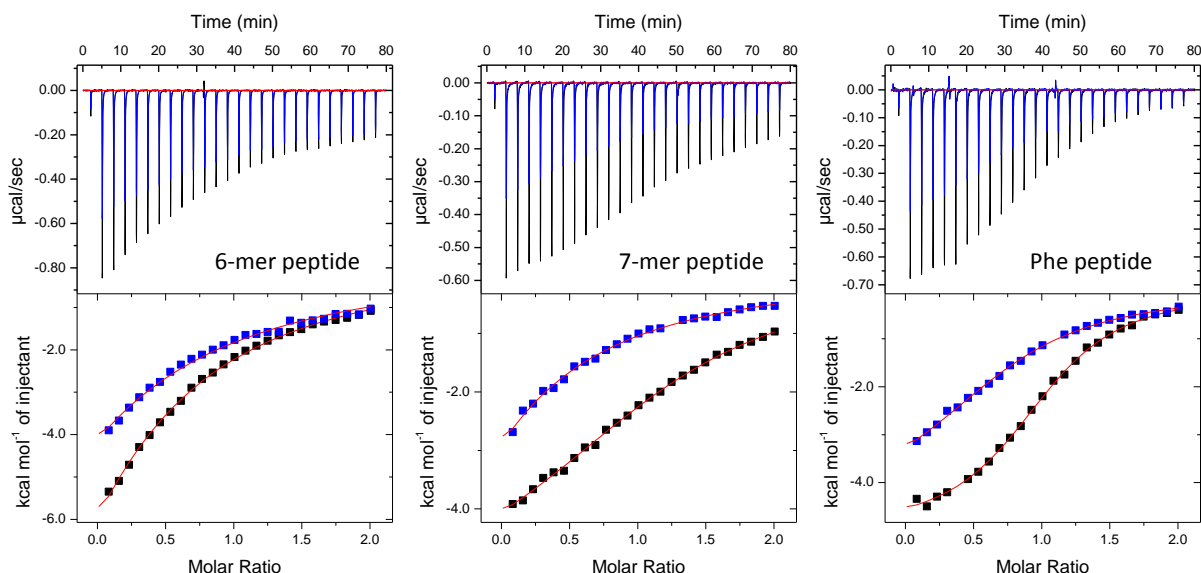
The first round of peptide SAR studies was designed to identify key focus areas on the core peptide substrate as shown in Figure 3.3. Focus areas of this series included the K36 quaternary amine mimics (green), K37 substitutions (black), His39 substitutions (blue), and Arg40(Lys40) substitutions (red).

Figure 3.3: Focus areas of Peptide SAR Series I



Two peptides were prepared containing a 6-mer (**145**) and 7-mer peptide (**146**) sequences of the endogenous H3 peptide possessing the residues Lys36-Tyr41 and Val35-Tyr41, respectively. The lysine residue was substituted for the Arg40 residue to assist in SPPS chemistry. It has been shown that arginine residues can undergo an intramolecular rearrangement during the coupling phase to the resin. The sequence of **145** consisted of Tyr-Lys-His-Pro-Lys-Lys(Kme3) and the sequence of **146** was identical with the addition of a Val residue following K36me3. These initial two peptides were screened for IC_{50} values in AlphaScreen and validated by ITC measurements. Peptide **145** provided an IC_{50} of $>100\ \mu\text{M}$ for both PHF1 and PHF19 with K_d measurements of $124.5\ \mu\text{M}$ (black) and 251.1 (blue), Figure 3.4

Figure 3.4: ITC of Series I Peptide SAR ITC



Peptide **146** provided IC_{50} measurements of $56.0 \pm 4.2 \mu M$ for PHF1 and $37.5 \pm 2.1 \mu M$ for PHF19. K_d measurements by ITC were $49.6 \mu M$ for PHF1(black) and $151.1 \mu M$ for PHF19 (blue). From these initial two peptides, it was decided to proceed forward with the remaining initial set of peptides as 7-mer peptides. The valine residue appears to increase affinity of the 7-mer peptide which is likely due to the N-H amide bond of the valine amide forming a hydrogen bond with a protein surface residue as was observed in the X-ray co-crystal structure.

The next set of peptides were synthesized to determine how amino acid substitutions at various locations on the 7mer sequence would affect binding affinity. Peptide **147** was prepared with a substitution of ornithine for Lys37. It was observed in the crystal structure that the peptide residue Lys37 had constrained conformation in its aliphatic chain that cause it to be bent outward from the protein surface. It was proposed that shortening this aliphatic chain would decrease this conformational strain in the aliphatic section of the residue and lead to increased affinity while maintaining the salt bridge with Glu66 on PHF1. This substitution was observed to be equipotent

to **146** with an IC_{50} of $34 \pm 9.9 \mu M$ for PHF1 and $47 \pm 22.6 \mu M$ for PHF19 in the AlphaScreen assay. The next amino acid substitution was at His38 and was aimed at determining if His38 substitutions were tolerated. This residue was believed to be essential for binding based on the Cellspot assay results. A phenylalanine was substituted for His38 in peptide **148** to maintain an aromatic residue at this location. It was shown to have an IC_{50} of $14.0 \pm 0.0 \mu M$ for PHF1 and $7.0 \pm 1.4 \mu M$ for PHF19 in AlphaScreen. Measured K_d values provide $10.9 \mu M$ for PHF1 (black) and $33.5 \mu M$ for PHF19 (blue), Figure 3.4. Comparison of this analog to **146** shows a ~5 fold improvement in K_d for both PHF1 and PHF19. It is proposed that increasing the hydrophobic character of the amino acid at this location increased the overall hydrophobic interactions between the peptide and tetra-leucine side wall of the protein where this peptide residue interacts. Substitutions were then begun at the Arg40 (Lys40) position to determine if a positively charged residue was necessary at this position and to determine if a hydrophobic residue would increase peptide affinity. Arg40 also binds within the tetra-leucine side wall, but the positively charged arginine head group was shown in the crystal structure to be solvent exposed pointing away from the protein's surface. It was hypothesized that by removing the positive charge, while increasing the hydrophobic nature of the residue by substituting it for a leucine that it would result in increased peptide affinity. Peptide **149** was shown to have an IC_{50} of $>100 \mu M$ for PHF1 and $32 \pm 2.1 \mu M$ for PHF19 in AlphaScreen and K_d values of $52.3 \mu M$ and $336.9 \mu M$ respectively. This modification did not prove advantageous for increased potency for either protein. Another experiment was conducted to determine whether a double mutant, His39Phe and Lys40Phe, could affect peptide potency. The double mutant peptide, **150**, was shown to be equipotent to peptide **148** and had K_d values of $14.6 \mu M$ and $21.5 \mu M$, respectively.

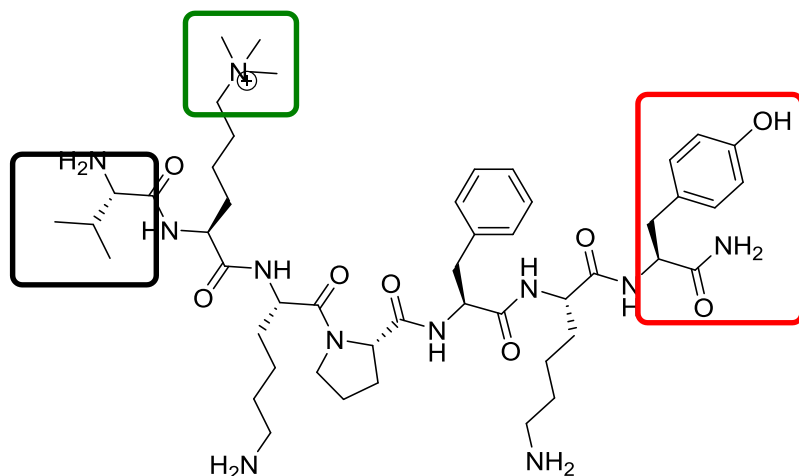
The phenylalanine substitution appears to have the strongest effect on increasing affinity for both PHF1 and PHF19.

The final set of experiments on peptide modifications was to determine the steric tolerance within the aromatic binding pocket by modulating the steric bulk of the Lys36 head group. Two different head groups were prepared, a *N',N'*-diethyl amine (**152**) and a more bulky *N'*-methyl, *N'*-norbornyl amine (**151**), see Figure 3.2 for head group structure. These two bulkier amine head groups were successfully applied in a peptide SAR project for a different Kme binding protein CBX7 (Jake Stuckey, unpublished results). Peptide **152** was shown to have an IC_{50} of $90 \pm 14 \mu M$ for PHF1, and $94.5 \pm 7.8 \mu M$ for PHF19 in AlphaScreen. Peptide **151** was shown to have an IC_{50} of $>100 \mu M$ for both PHF1 and PHF19 in AlphaScreen. These two peptides provide evidence that there is a defined degree of steric tolerance within the aromatic cage and larger R-groups on the lysine amine are not well tolerated. The findings of this initial set of peptides were used to design a second series of analogs.

Second SAR round of peptide ligand amino acid substitutions

A second series of peptides was designed based upon the the key finding that the phenylalanine substitution for Arg40 (Lys40) provided an increase in peptide affinity as shown in **150**. This second series continued to explore the steric tolerance of the Kme aromatic cage, to determine whether substitutions at Val35 would allow for increased interactions and to determine if Tyr41 was essential for peptide binding, Figure 3.5. The focus areas of the series II peptide SAR experiments included continued exploration of quaternary amine mimetics for K36 (green), substitution of Val35 (black), and truncation of Tyr41 (red).

Figure 3.5: Focus Areas of Peptide SAR Series II



It was hypothesized that the Kme aromatic cage would not tolerate large, sterically bulky functional groups on the amine head group of Lys36 based on the findings that the initial two amine heads groups showed poor binding affinity. We therefore decided to prepare two additional Lys36 analogs that possessed less bulky functional groups. The initial N'-ethyl, N'-methyl lysine analog, **153**, provided K_d values of 74.4 μ M PHF1 and was not determined for PHF19 due to synthetic difficulty to preparation large quantities of peptide. The N',N'-dimethyl lysine analog, **154**, provided K_d values of 112 μ M and 1.33 mM for PHF1 and PHF19, respectively. It is observed from these two analogs that there does appear to be a size preference for the Lys36 functional groups, however those most appropriate to increase affinity have not yet been discovered.

The next residue explored in the series II peptide SAR was Val35. This residue is known to make an amide N-H hydrogen bond with the surface of the protein. It was hypothesized that additional hydrogen bonding could be introduced at this residue site with the protein residue Glu66 by modifying the Val35 residue for either a threonine or serine residue. The Thr35 analog,

155, provided K_d values of 22.8 μM and 40.9 μM for PHF1 and PHF19 respectively. The Ser35 analog, **156**, provided K_d values of 16.1 μM and 40.9 μM for PHF1 and PHF19 respectively. These two modifications did not appear to directly increase peptide affinity for the target, but were equipotent compared to the phenylalanine analog in series I.

The final region that was explored was the Tyr41 residue. As previously mentioned, this residue was included to simplify peptide purification by providing a UV active residue. An experiment was planned to determine whether this residue was essential for peptide binding, therefore a truncated analog that deleted the Try41 residue, peptide **157** was prepared. This peptide provided K_d values of 24.1 μM and 241 μM for PHF1 and PHF19 respectively. This final analog for the series II peptide SAR experiments shows that the Try41 residue does not play a major role peptide affinity for the target as it is equipotent to the phenylalanine mutant analog.

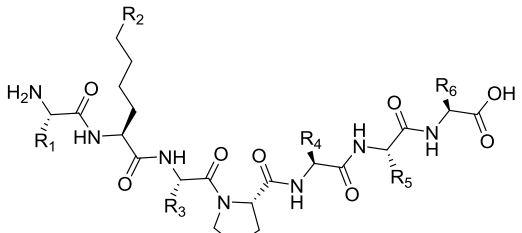
Conclusions











Work toward development of a highly potent and selective peptide-based inhibitor for PHF1 and PHF19 has thus far shown initial promise. The series I peptide SAR shows that it is possible to decrease the length of the histone peptide fragment from a large 12- or 19-mer sequence into a 7-mer peptide that is equipotent for the Tudor domain of PHF1 and PHF19 compared to the larger 12-mer and 19-mer peptide sequences. This represents a very large increase in ligand efficiency for the initial peptide inhibitor.




Efforts toward discovery of a quaternary amine mimetic for the Lys36Kme3 residues proved difficult. Four different Lys36 analogs that possessed amine functional groups of varying steric bulk were prepared within two peptide SAR series, Table 3.1. These analogs showed that the Tudor domain appears to have a low degree of steric tolerance for the size and steric bulk of

the lysine's amine functional groups within the Kme aromatic pocket of the Tudor domains of PHF1 and PHF19. The two larger analogs, **151** and **152**, showed very little affinity for the domain. It is proposed that this low affinity is due to the small volume of the Kme aromatic cage. Analysis of the co-crystal structure of the H3K36me3 peptide bound to PHF1's Tudor domain shows that the pocket is $\sim 4\text{\AA}$ in depth

Table 3.1: SAR of synthesized peptides for PHF1 and PHF19^{a, b, c}



ID	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆	PHF1 IC ₅₀ (μM)	PHF19 IC ₅₀ (μM)	PHF1 <i>K_d</i> (μM)	PHF19 <i>K_d</i> (μM)
145	-		Lys	His	Lys	Tyr	>100	>100	124.5	251.1
146	Val		Lys	His	Lys	Tyr	56.0 \pm 4.2	37.5 \pm 2.1	49.6	151.1
147	Val		Orn	His	Lys	Tyr	34.0 \pm 9.9	47.0 \pm 22.6	ND	ND
148	Val		Lys	Phe	Lys	Tyr	14.0 \pm 0.0	7.0 \pm 1.4	10.9	33.5
149	Val		Lys	His	Leu	Tyr	>100	32.0 \pm 2.1	52.3	336.9
150	Val		Lys	Phe	Leu	Tyr	ND	ND	14.6	21.5
151	Val		Lys	His	Lys	Tyr	32.0 \pm 2.1	32.0 \pm 2.1	ND	ND
152	Val		Lys	His	Lys	Tyr	90.0 \pm 14.0	94.5 \pm 7.8	ND	ND
153	Val		Lys	Phe	Lys	Tyr	ND	ND	74.4	ND
154	Val		Lys	Phe	Lys	Tyr	ND	ND	112	1330.0

155	Thr		Lys	Phe	Lys	Tyr	ND	ND	22.8	40.9
156	Ser		Lys	Phe	Lys	Tyr	ND	ND	16.1	40.9
157	Val		Lys	Phe	Lys	-	ND	ND	24.1	241.0

^aIC₅₀ values are the average of at least 2 values \pm the standard deviation as determined by AlphaScreen. ^bThe maximum concentration in the assay was 100 μ M and compounds showing less than 50% inhibition at this concentration are labeled >100 μ M. ^cK_d values are determined by ITC. ND: no determined.

Two additional Kme3 analogs of slightly decreased steric bulk were prepared to determine if the substrate-induced protein conformational change would allow for binding of the small functional groups. The IC₅₀ and ITC binding data showed otherwise. The two smaller H3K36 analogs, **153** and **154**, also showed poor affinity for the Tudor domain. The di-methyl analog, **154**, was expected to have lower affinity as compared to previous endogenous peptide studies that showed the Tudor domain selects for Kme3 over Kme2 lysine residues. The N'-ethyl,N'-methyl analog, **153**, showed a small increase in potency, but overall had a weaker affinity compared to the Kme3 analog, **146**. Future synthetic efforts toward discovery of an H3K36me3 mimetic will need to take these findings into account.

Exploration of the ability to mutate specific residues from the endogenous peptide sequence in the 7-mer core peptide (residues Val35-Tyr41) showed how analysis of the protein surface/peptide binding interactions allows for an increase in peptide affinity. The specific mutation of Arg40 for phenylalanine provide a large increase in peptide potency for the Tudor domains of PHF1 and PHF19. This substitution was chosen as the area where the Arg bound on the protein consisted of a large hydrophobic patch of four leucine residues. It was hypothesized that increasing the hydrophobic character of the peptide residue in this area would potentially increase peptide affinity. This substitution proved successful in increasing peptide affinity and

was maintained into series II peptide SAR. Unfortunately, no additional peptide amino acid substitutions provided any additional affinity for the target. It is curious to note that the truncation of Tyr41 provided an equipotent peptide compared to as the phenylalanine mutant. This is a promising finding as it provides a 6-mer peptide sequence. Future work on additional series of peptide SAR should focus on developing new Lys36Kme3 analogs to determine whether additional amino acid substitutions for hydrophobic residues that bind within the protein's hydrophobic patch can provide an increase in peptide affinity; specifically Pro38.

Experimental section

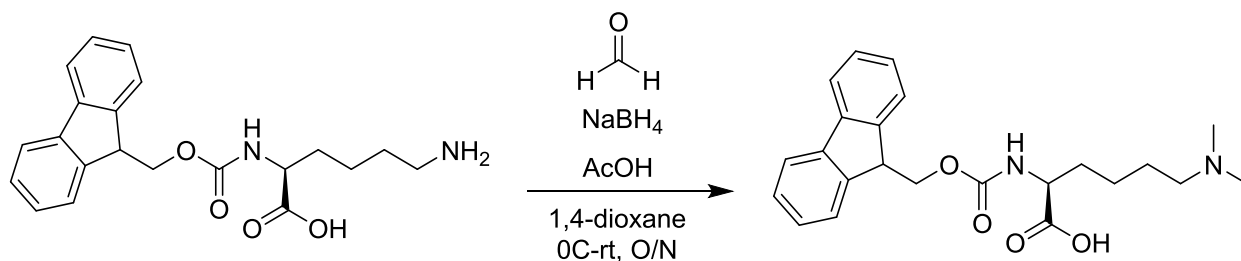
General Procedure for Chemical Synthesis

Analytical LCMS data for all compounds were acquired using an Agilent 6110 series system with the UV detector set to 220 and 254 nm. Samples were injected (<10 μ L) onto an Agilent Eclipse Plus 4.6 \times 50 mm, 1.8 μ m, C18 column at room temperature. A mobile phase of A (H₂O + 0.1% acetic acid) and B (MeOH + 0.1% acetic acid) was used with a linear gradient from 10% to 100% B in 5.0 min, followed by a flush at 100% B for another 2 minutes with a flow rate of 1.0 mL/min. Mass spectra data were acquired in positive ion mode using an Agilent 6110 single quadrupole mass spectrometer with an electrospray ionization source. Nuclear Magnetic Resonance (NMR) spectra were recorded on a Varian Mercury spectrometer at 400 MHz for proton (¹H NMR) and 100 MHz for carbon (¹³C NMR); chemical shifts are reported in ppm (δ). Analytical thin-layer chromatography (TLC) was performed with silica gel 60 F₂₅₄, 0.25 mm pre-coated TLC plates, generally using a 10% MeOH in DCM solvent system. TLC plates were visualized using UV₂₅₄, I₂ impregnated silica gel, *potassium permanganate* with charring,

and phosphomolybdic acid with charring. Reverse phase chromatography was used to purify reaction mixtures to obtain final products using a Teledyne Isco CombiFlash Rf 200 chromatography unit equipped with the UV detector set to 220 nm and 254 nm. Samples were injected onto a RediSep Rf 30g C18 high performance Gold column at room temperature and collected at the previously mentioned wavelengths. Mobile phases of A ($\text{H}_2\text{O} + 0.1\% \text{ TFA}$) and B (MeOH) were used with a flow rate of 30 mL/min. A general gradient was used consisting of 0-2 minutes at 5% B, 5-15 minutes increasing from 5 to 100% B, and a 100% B flush for another 3 minutes. Small variations in this purification method were made as needed to achieve ideal separation for each compound. All compounds that were evaluated in biochemical and biophysical assays had >95% purity as determined by ^1H NMR and LCMS.

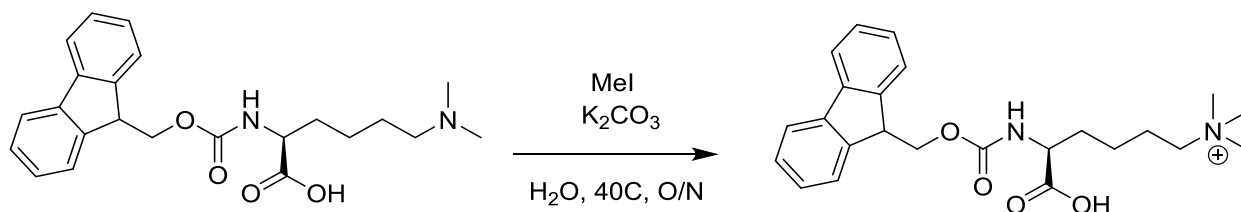
Non-natural trimethyl-lysine Fmoc protected monomer synthesis

Several non-natural trimethyl-lysine N- α -Fmoc protected amino acid monomers were synthesized. These monomers were prepared primarily from Fmoc-protected lysine HCl. Synthetic methods focused on reductive aminations under acidic conditions and N-methylations to prepare di- and tri-substituted analogs. A general method is described for each below.



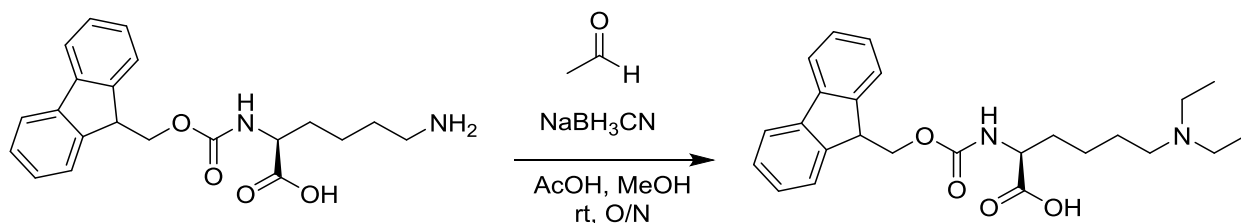
N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶,N⁶-dimethyl-L-lysine TFA (139): A solution of Fmoc-lysine HCl (0.30 g, 0.81 mmol), 0.30 ml (4.1 mmol) formaldehyde solution (37% wt in

H₂O), and 0.23 mL acetic acid (4.1 mmol) was stirred at 0°C for 30 min in 3 mL of 1,4-dioxane . Then 277 mg (7.3 mmol) sodium borohydride was slowly added in portions at 0°C until half was added to the reaction. Then another 0.30 ml (4.1 mmol) formaldehyde solution (37% wt in H₂O) was added followed by the remaining sodium borohydride being slowly added in portions at 0°C. The mixture was slowly warmed to room temperature and stirred overnight. The reaction was monitored by LC-MS until completion and then dried to residue. The crude material was dissolved in 0.250 mL of ACN and purified by reverse phase HPLC using the Teledyne Isco automated column system to afford 0.29 mg (90% yield) as the TFA salt as a white foam. ¹H NMR (400 MHz, CD₃OD) δ 7.74 (d, *J* = 7.9 Hz, 2H), 7.63 (t, *J* = 7.3 Hz, 2H), 7.35 (t, *J* = 8.1 Hz, 2H), 7.28 (t, *J* = 9.5 Hz, 2H), 4.41 – 4.27 (m, 2H), 4.16 (t, *J* = 7.8 Hz, 2H), 3.02 (t, *J* = 7.3 Hz, 2H), 2.9 (s, 6H), 1.95 – 1.84 (m, 1H), 1.79 – 1.56 (m, 3H), 1.51 – 1.35 (m, 2H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.6 min. MS (ESI⁺): 397.2 [M+H]⁺.

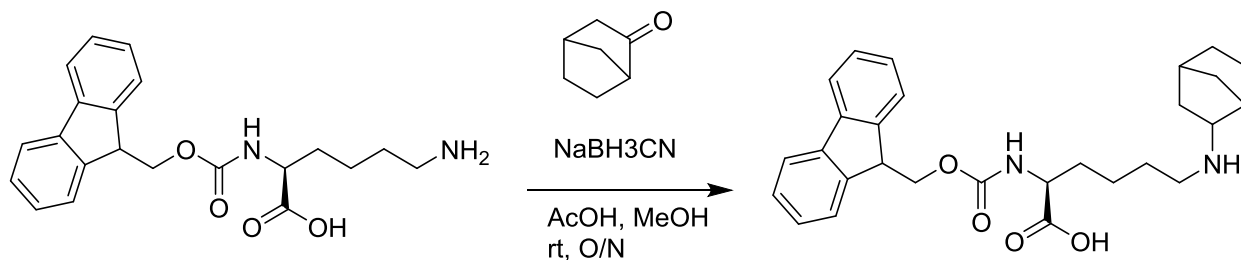


(S)-5-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-5-carboxy-N,N,N-trimethylpentan-1-aminium TFA (140): To a solution of 139 (0.2 g, 0.82 mmol) in 4 mL of H₂O was added 0.32 mL (5.1 mmol) MeI, and (0.21 g, 1.5 mmol) K₂CO₃, and the mixture was stirred at 40°C overnight in a heat bath. The reaction was monitored by LC-MS until completion and then dried to residue. The crude material was dissolved in 0.250 mL of ACN and purified by reverse phase HPLC using the Teledyne Isco automated column system to afford 160 mg (77 % yield) as the TFA salt as a white solid. ¹H NMR (400 MHz, CD₃OD) δ 7.77 (d, *J* = 6.9 Hz, 2H), 7.64 (t, *J* =

8.8 Hz, 2H), 7.39 (d, $J = 8.2$ Hz, 2H), 7.30 (t, $J = 7.4$ Hz, 1H), 4.33 (d, $J = 5.2$ Hz, 2H), 4.8 (t, $J = 8.6$ Hz, 1H), 4.02 (d, $J = 7.8$ Hz, 2H), 3.02 (m, 2H), 3.06 (s, 9H), 1.94 – 1.22 (m, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.6$ min. MS (ESI⁺): 411.3 [M+H]⁺. LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.6$ min. MS (ESI⁺): 411.3 [M+H]⁺.

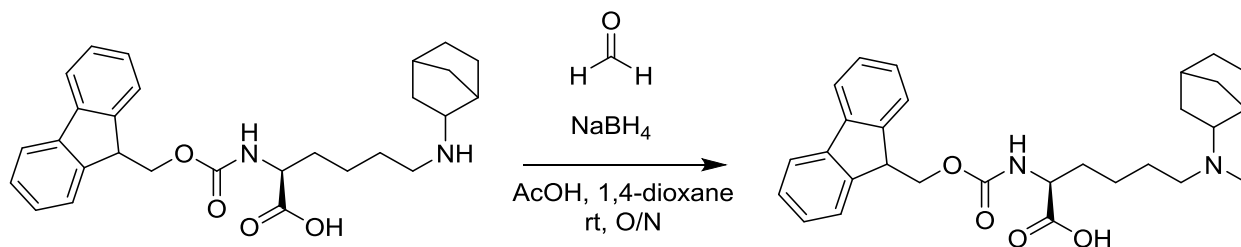


N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶,N⁶-diethyl-L-lysine TFA(141): A solution of Fmoc-lysine HCl (0.31 g, 0.77 mmol), 0.12 ml (2.2 mmol) acetaldehyde, and 0.14 mL (3.7 mmol) acetic acid were stirred at room temperature for 2 hr in 2 mL of MeOH. Then 139 mg (2.2 mmol) sodium cyanoborohydride was added in portions and the mixture was stirred for overnight at room temperature (rt). The reaction was monitored by LC-MS until completion and then dried to residue. The crude material was dissolved in 0.250 mL of ACN and purified by reverse phase HPLC using the Teledyne Isco automated column system to afford 307 mg (98 yield) as the TFA salt as a clear oil. ¹H NMR (400 MHz, CD₃OD) δ 7.81 (d, $J = 6.8$ Hz, 2H), 7.67 (t, $J = 9.7$ Hz, 2H), 7.4 (t, $J = 7.7$ Hz, 2H), 7.32 (t, $J = 7.7$ Hz, 2H), 4.42 (dd, $J = 10.1, 6.9$ Hz, 1H), 4.32 (dd, $J = 7.0, 6.9$ Hz, 1H), 4.23 (t, $J = 7.4$ Hz, 1H), 4.18 (dd, $J = 9.3, 4.8$ Hz, 1H), 3.2 (q, $J = 14.9, 7.1$ Hz, 4H), 3.14 – 3.02 (m, 2H), 3.14 – 3.02 (m, 1H), 2.0 – 1.87 (m, 4H), 1.81 – 1.63 (m, 2H), 1.29 (t, $J = 8.2$, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 4.7$ min. MS (ESI⁺): 425.3 [M+H]⁺.



N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶-(bicyclo[2.2.1]heptan-2-yl)-L-lysine TFA

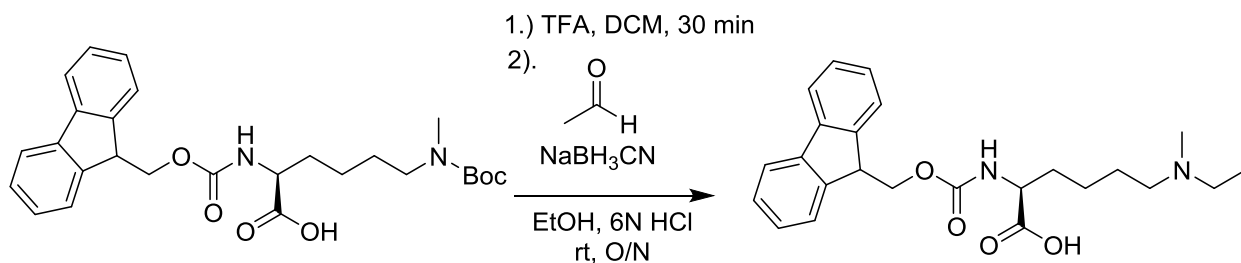
(142): A solution of Fmoc-Lysine HCl (0.40 g, 0.99 mmol), 0.87 g (2.2 mmol) 2-Norbornanone, and 0.32 mL (4.9 mmol) acetic acid were stirred at room temperature for 2 hr in 2 mL of MeOH. Then 186 mg (3.0 mmol) sodium cyanoborohydride was added in portions and the mixture was stirred for overnight at rt. The reaction was monitored by LC-MS until completion and then dried to residue. The crude material was dissolved in 0.250 mL of ACN and purified by reverse phase HPLC using the Teledyne Isco automated column system to afford 406 mg (89% yield) as the TFA salt as a clear oil. ¹H NMR (400 MHz, CD₃OD) δ 7.81 (d, *J* = 7.0 Hz, 2H), 7.67 (t, *J* = 9.5 Hz, 2H), 7.4 (t, *J* = 7.5 Hz, 2H), 7.31 (t, *J* = 7.9 Hz, 2H), 4.42 (dd, *J* = 9.9, 6.5 Hz, 1H), 4.34 (dd, *J* = 9.4, 7.1 Hz, 1H), 4.24 (t, *J* = 7.6 Hz, 1H), 4.18 (dd, *J* = 9.9, 4.7 Hz, 1H), 3.49 – 3.41 (m, 1H), 3.0 – 2.88 (m, 2H), 2.54 (t, *J* = 4.8, 1H), 2.33 (t, *J* = 5.1, 1H), 2.12–2.02 (m, 1H), 1.97 – 1.86 (m, 1H), 1.85 – 1.56 (m, 6H), 1.55 – 1.44 (m, 4H), 1.43 – 1.35 (m, 1H), 1.1 – 1.02 (m, 1H). LC-MS (λ = 254 nm): 99%, *t_R* = 4.9 min. MS (ESI⁺): 463.2 [M+H]⁺.



N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶-(bicyclo[2.2.1]heptan-2-yl)-N⁶-methyl-L-

lysine TFA (143): A solution of **142** (0.154 g, 0.33 mmol), 0.13 mL (1.7 mmol)

formaldehyde solution (37% wt in H₂O), and 0.09 mL acetic acid (1.7 mmol) was stirred at room temperature for 30 min in 4 mL of 1,4-dioxane. Then 113 mg (2.9 mmol) sodium borohydride was slowly added in portions. The mixture was stirred at room temperature overnight. The reaction was monitored by LC-MS until completion and then dried to residue. The crude material was dissolved in 0.250 mL of ACN and purified by reverse phase HPLC using the Teledyne Isco automated column system to afford 91 mg (57% yield) as the TFA salt as a clear oil. ¹H NMR (400 MHz, CD₃OD) δ 7.79 (d, *J* = 7.3 Hz, 2H), 7.66 (t, *J* = 9.2 Hz, 2H), 7.39 (t, *J* = 7.9 Hz, 2H), 7.30 (t, *J* = 7.6 Hz, 2H), 4.42 – 4.39 (m, 2H), 4.25 – 4.15 (m, 2H), 3.45 – 3.37 (m, 1H), 3.24 – 3.06 (m, 1H), 3.04 – 2.88 (m, 1H), 2.8 (s, 3H), 2.58 (t, *J* = 6.1, 1H), 2.31 (t, *J* = 7.6, 1H), 2.1- 2.0 (m, 1H), 1.9 – 1.0 (m, 13H), 1.21 – 1.1 (m, 1H). LC-MS (λ = 254 nm): 99%, *t*_R = 4.89 min. MS (ESI⁺): 477.3 [M+H]⁺.



N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶-ethyl-N⁶-methyl-L-lysine HCl (144). To a

solution of N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶-(tert-butoxycarbonyl)-N⁶-methyl-L-lysine (0.53 g, 0.58 mmol) in 10 mL of TFA and was stirred for 1 hr at room temperature. The solvent was then removed by rotovaporation to provide a viscous yellow oil. This crude material was moved directly forward by dissolving it in 20 mL EtOH added 0.31 mL (5.5 mmol)

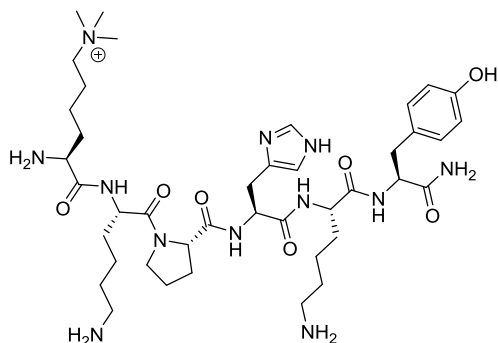
acetaldehyde solution and the mixture was stirred for 30 min at room temperature. Then 0.206 g (3.3 mmol) sodium cyanoborohydride and the mixture was stirred overnight at rt. The reaction was monitored by LC-MS until completion and then dried to residue. Once complete, the reaction was acidified by addition of 10 mL of 6N HCl to provide a white solid. The solid was filtered out to afford 263 g (59% yield) of the HCl salt as a white solid. ^1H NMR (400 MHz, dmsO) δ 7.90 (d, J = 7.6 Hz, 2H), 7.72 (dd, J = 7.3, 3.8 Hz, 1H), 7.66 (d, J = 8.2 Hz, 1H), 7.42 (t, J = 7.2 Hz, 2H), 7.33 (t, J = 7.4 Hz, 2H), 4.33 – 4.19 (m, 3H), 3.98 – 3.90 (m, 1H), 3.15 – 2.88 (m, 4H), 2.66 (d, J = 4.9 Hz, 3H), 1.79 – 1.57 (m, 4H), 1.41 – 1.29 (m, 2H), 1.20 (t, J = 7.2 Hz, 3H). LC-MS (λ = 254 nm): 98%, t_R = 4.9 min. MS (ESI $^+$): 411.2 $[\text{M}+\text{H}]^+$.

Solid-Phase Peptide Synthesis (SPPS) of Designed Peptides

Structure-based drug design methods were used to design two series of peptides as inhibitors of PHF1 and PHF19. All synthesized peptides were prepared through standard SPPS methods using N- α -Fmoc protected L-amino acids (AnaSpec Inc., Chem-Impex Int'l), HBTU and HOAt coupling reagents (AnaSpec Inc., CreoSalus Life Science Co.) in DMF solvent, and DIPEA (Sigma) as base. Rink amide MBHA resin (AnaSpec Inc) with 0.6 mmol/g loading was used as the solid phase support and coupling was conducted as described through the following general procedure. A small 25 mL disposable pipette was fitted with a silica wafer to allow for a convenient SPPS synthesis container. To this synthesis vessel was added 0.15 g of Rink amide MBHA resin. The resin was then swollen in DCM (4 mL, 2 x 5 min), followed by DMF (4 mL, 2 x 5 min) with gentle shaking during each step by placing the pipette on a shaking plate. DIPEA (5 mL, 2 x 5 min) was then used to deprotect the swollen resin, followed by washing cleaved resin by DMF (4 mL, 2 x 5 min), MeOH (4 mL, 2 x 5 min), DMF (4 mL, 2 x 5 min). During

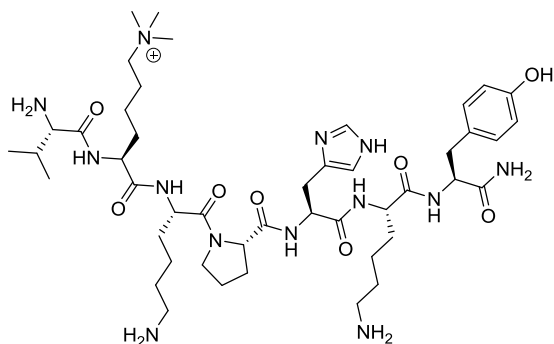
this washing step, the initial N- α -Fmoc protected L-amino acid was pre-activated by dissolved in DMF (3 mL) with HBTU (2 equiv), HOAt (2 equiv), and DIPEA (4 equiv) in a scintillation vial that was placed on a shaking plate. After the deprotection and washing of the resin was complete, the activated initial N- α -Fmoc protected L-amino acid was introduced onto the deprotected resin and coupled for 1 hr with gentle shaking. This coupling was repeated one additional time, followed by washing with DMF (4 mL, 2 x 1 min), MeOH (4 mL, 2 x 1 min), DMF (4 mL, 2 x 1 min). The completion of the coupling was confirmed using a Chloronil test to determine if there were any uncoupled amine groups remaining on the resin. If so, an additional coupling was conducted to fully saturate the resin with the initial amino acid. Once the initial coupling was completed and the resin was fully saturated, the N- α -Fmoc group on that amino acid was then deprotected was conducted over two steps using a 20% solution of piperidine in DMF (4 mL, 1 x 10 min, followed by 1 x 20 min) followed by three-part washing step as previously described to provide an active amine group upon which to couple the next amino acid. The next amino acid was then coupled following the same pre-activation and coupling procedure. This process was repeated until all desired amino acids were coupled in succession. Non-natural amino acids were incorporated using 3 equiv of amino acid with the above described standard coupling reagents and sequence. The coupling agent *N,N'*-diisopropylcarbodiimide (DIC) was used to only couple residues following proline; here 3 equiv DIC was used, 2 equiv HOAt, and 4 equiv DIPEA in 4 mL DMF for two iterations of 1 hr couplings. After all amino acids have been successfully coupled, the N- α -Fmoc group on the last amino acid was deprotected using the standard 20% piperidine in DMF solution, followed by the three-part wash step, and then the peptide was cleaved from the resin using a 95:2.5:2.5 TFA/TIPS/H₂O solution (5 mL) over a three hour period. The cleaved solution was concentrated to an oily residue and then dissolved in 0.5 mL

ACN and then run on semi-prep HPLC. Eluted peaks were analyzed by LC-MS to identify desired product peak(s). The desired peaks were combined, concentrated to residue, and then lyophilized overnight to provide a white peptide residue.



145: Peptide **145** was prepared and purified using

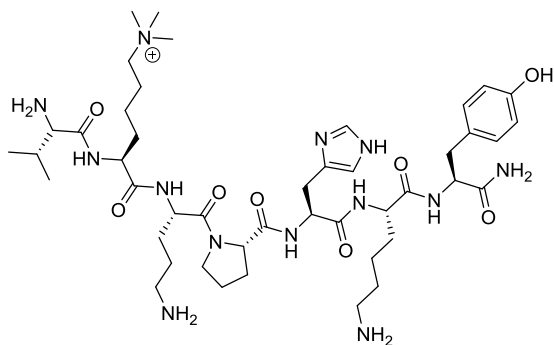
the general peptide synthesis procedure described above. Synthesis was started from 300 mg of Rink amide MBHA resin, 25 mg (17% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.71 (d, $J = 1.4$ Hz, 1H), 7.35 (d, $J = 1.3$, 1H), 7.07 (d, $J = 9.1$ Hz, 2H), 6.68 (d, $J = 8.7$ Hz, 2H), 4.71 – 4.47 (m, 3H), 4.47 – 4.30 (m, , 1H), 4.30 – 4.12 (m, 1H), 4.06 – 3.92 (m, 1H), 3.92 – 3.77 (m, 1H), 3.77 – 3.52 (m, 1H), 3.23 – 3.19 (m, 1H), 3.12 - 3.07 (m, 11H), 3.03 (dd, $J = 13.9, 5.7$ Hz, 2H), 2.95 – 2.82 (m, 6H), 2.31 – 1.14 (m, 28H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 0.73$ min. MS (ESI $^+$): $\text{C}_{41}\text{H}_{69}\text{N}_{12}\text{O}_7^+$ Calculated m/z 841.5, found 841.8.



146: Peptide **146** was prepared and purified

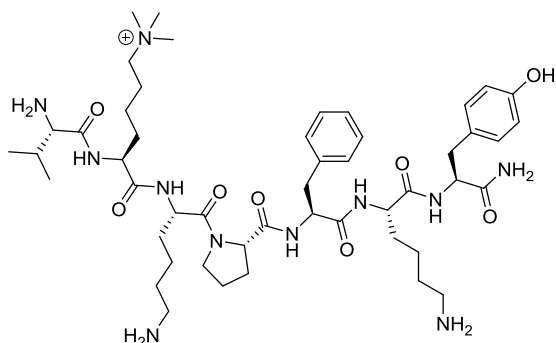
using the same procedure as for peptide **145**. Synthesis was started from 300 mg of Rink amide

MBHA resin, 44 mg (26% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.74 (d, $J = 1.4$ Hz, 1H), 7.35 (d, $J = 1.2$ Hz, 1H), 7.10 (d, $J = 8.1$ Hz, 2H), 6.70 (d, $J = 8.4$ Hz, 2H), 4.71 – 4.60 (m, 2H), 4.56 (dd, $J = 8.8, 5.6$ Hz, 1H), 4.49 – 4.36 (m, 2H), 4.25 (t, $J = 7.1$ Hz, 1H), 3.88 – 3.61 (m, 1H), 3.26 (dd, $J = 15.2, 5.2$ Hz, 1H), 3.18 – 3.07 (m, 14H), 3.05 (dd, $J = 5.5$ Hz, 1H), 2.98 – 2.87 (m, 7H), 2.27 – 2.14 (m, 1H), 2.10 – 1.27 (m, 24H), 1.05 (dd, $J = 8.4, 7.0$ Hz, 7H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 0.74$ min. MS (ESI $^{+}$): $\text{C}_{46}\text{H}_{78}\text{N}_{13}\text{O}_8^{+}$ Calculated m/z 940.6, found 941.6 (M+H).



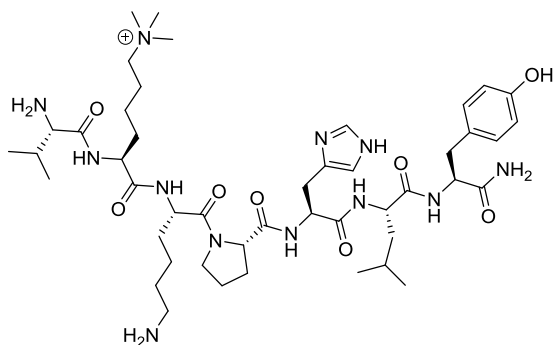
147: Peptide **147** was prepared and purified

using the same procedure as for peptide **145**. Synthesis was started from 100 mg of Rink amide MBHA resin, 3.45 mg (6.0% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.73 (d, $J = 12.6$ Hz, 1H), 7.37 (d, $J = 9.5$ Hz, 1H), 7.09 (d, $J = 8.5$ Hz, 2H), 6.70 (d, $J = 8.5$ Hz, 2H), 4.75 – 4.63 (m, 3H), 4.61 – 4.52 (m, 1H), 4.47 – 4.37 (m, 2H), 4.31 – 4.19 (m, 1H), 3.97 – 3.86 (m, 1H), 3.86 – 3.78 (m, 1H), 3.75 (t, $J = 7.2$ Hz, 1H), 3.72 – 3.62 (m, 1H), 3.28 – 3.19 (m, 1H), 3.19 – 3.02 (m, 7H), 3.02 – 2.82 (m, 6H), 2.30 – 2.12 (m, 2H), 2.12 – 1.17 (m, 21H), 1.11 – 0.98 (m, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_{\text{R}} = 0.68$ min. MS (ESI $^{+}$): $\text{C}_{45}\text{H}_{76}\text{N}_{13}\text{O}_8^{+}$ Calculated m/z 926.6, found 926.6.



148: Peptide **148** was prepared and purified

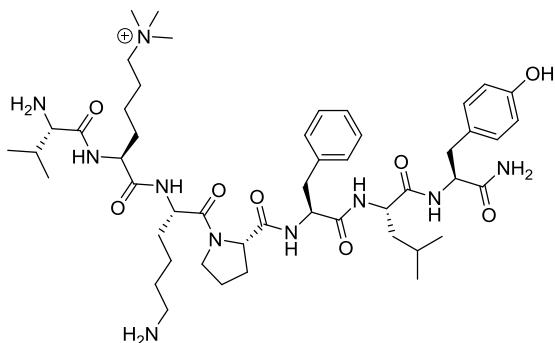
using the same procedure as for peptide **145**. Synthesis was started from 100 mg of Rink amide MBHA resin, 3.45 mg (6.1% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.30 – 7.16 (m, 5H), 7.08 (d, $J = 7.9$ Hz, 2H), 6.72 (d, $J = 8.5$ Hz, 2H), 4.60 – 4.48 (m, 8H), 4.44 – 4.37 (m, 4H), 4.23 (dd, $J = 12.3, 6.8$ Hz, 2H), 3.86 – 3.76 (m, 2H), 3.73 (d, $J = 5.8$ Hz, 2H), 3.65 – 3.56 (m, 2H), 3.15 – 3.12 (m, 1H), 3.10 (s, 9H), 3.09 – 3.00 (m, 2H), 2.99 – 2.87 (m, 6H), 2.86 – 2.81 (m, 1H), 2.24 – 2.11 (m, 3H), 1.99 – 1.39 (m, 24H), 1.35 – 1.23 (m, 3H), 1.05 (dd, $J = 8.9, 6.9$ Hz, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 0.64$ min. MS (ESI $^+$): $\text{C}_{49}\text{H}_{80}\text{N}_{11}\text{O}_8^+$. Calculated m/z 950.6, found 951.5 (M+H).



149: Peptide **149** was prepared and purified

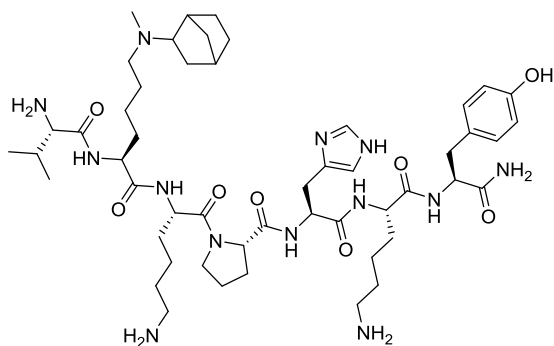
using the same procedure as for peptide **145**. Synthesis was started from 100 mg of Rink amide MBHA resin, 6.2 mg (11% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.71 (d, $J = 1.4$ Hz, 1H), 7.97 (s, 1H), 7.08 (d, $J = 8.5$ Hz, 2H), 6.70 (d, $J = 8.6$ Hz, 2H), 4.68 - 4.58

(m, 2H), 4.57 – 4.27 (m, 2H), 3.91 - 3.78 (s, 1H), 3.77 - 3.60 (m, 2H), 3.27 – 3.01 (m, 11H), 2.99 – 2.87 (m, 3H), 2.3 - 2.11 (m, 2H), 2.11 – 1.93 (m, 3H), 1.94 - 1.36 (m, 16H), 1.06 (dd, $J = 8.7$, 7.0 Hz, 6H), 0.91 (dd, $J = 18.3$, 6.4 Hz, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 0.65$ min. MS (ESI+): $C_{46}H_{77}N_{12}O_8^+$ Calculated m/z 925.6, found 926.5.



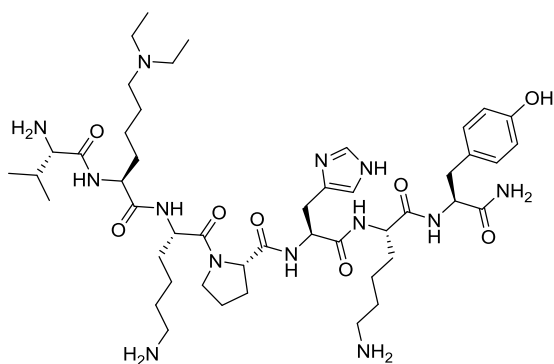
150: Peptide **150** was prepared and purified

using the same procedure as for peptide **145**. Synthesis was started from 150 mg of Rink amide MBHA resin, 14 mg (17% yield) white solid as the TFA salt. 1H NMR (400 MHz, cd_3od) δ 7.32 – 7.17 (m, 5H), 7.05 (d, $J = 8.6$ Hz, 2H), 6.70 (d, $J = 8.6$ Hz, 2H), 4.62 – 4.46 (m, 3H), 4.46 – 4.39 (m, 1H), 4.36 (dd, $J = 8.4$, 4.8 Hz, 1H), 4.32 – 4.24 (m, 1H), 3.85 – 3.76 (m, 1H), 3.73 (d, $J = 5.8$ Hz, 1H), 3.66 – 3.58 (m, 1H), 3.18 – 3.08 (m, 10H), 3.08 – 2.91 (m, 4H), 2.91 – 2.80 (m, 1H), 2.25 – 2.08 (m, 2H), 2.01 – 1.62 (m, 11H), 1.58 – 1.37 (m, 7H), 1.03 (ddd, $J = 12.4$, 9.3, 7.0 Hz, 6H), 0.89 (dd, $J = 18.0$, 6.2 Hz, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.29$ min. MS (ESI+): $C_{49}H_{79}N_{10}O_8^+$ Calculated m/z 935.6, found 936.5 (M+H).



151: Peptide **151** was prepared and purified

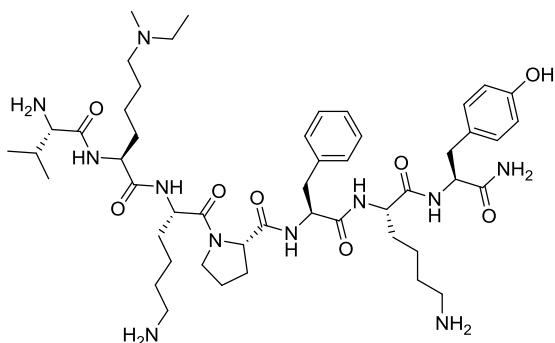
using the same procedure as for peptide **145**. Synthesis was started from 100 mg of Rink amide MBHA resin, 1.8 mg (3.0% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 8.73 (d, $J = 12.6$ Hz, 1H), 7.37 (d, $J = 9.5$ Hz, 1H), 7.09 (d, $J = 8.5$ Hz, 2H), 6.70 (d, $J = 8.5$ Hz, 2H), 4.75 – 4.63 (m, 3H), 4.61 – 4.52 (m, 1H), 4.47 – 4.37 (m, 2H), 4.31 – 4.19 (m, 1H), 3.97 – 3.86 (m, 1H), 3.86 – 3.78 (m, 1H), 3.75 (t, $J = 7.2$ Hz, 1H), 3.72 – 3.62 (m, 1H), 3.28 – 3.19 (m, 1H), 3.19 – 3.02 (m, 7H), 3.02 – 2.82 (m, 6H), 2.30 – 2.12 (m, 2H), 2.12 – 1.17 (m, 21H), 1.11 – 0.98 (m, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 0.65$ min. MS (ESI $^+$): $\text{C}_{51}\text{H}_{83}\text{N}_{13}\text{O}_8^+$ Calculated m/z 1005.7 (M), 502.8 (M+2), 335.2 (M+3), 251.4 (M+4), found M (ND), 503.9 ((M+2)+H), 336.3 ((M+3)+H), 252.5 ((M+4)+H).



152: Peptide **152** was prepared and purified

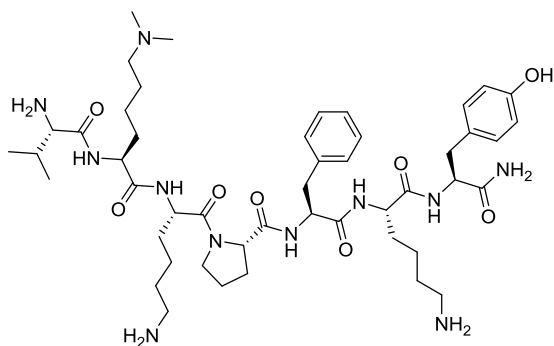
using the same procedure as for peptide **145**. Synthesis was started from 100 mg of Rink amide MBHA resin, 3.5 mg (6.1% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ

8.78 - 8.71 (m, 1H), 7.35 (s, 1H), 7.10 (d, $J = 8.5$ Hz, 2H), 6.70 (d, $J = 8.5$ Hz, 2H), 4.72 – 4.50 (m, 4H), 4.46 – 4.35 (m, 2H), 4.25 (t, $J = 7.0$ Hz, 1H), 3.90 – 3.76 (m, 1H), 3.76 - 3.61 (m, 2H), 3.27 – 3.17 (m, 5H), 3.17 – 3.02 (m, 4H), 2.98 – 2.87 (m, 5H), 2.28 - 2.14 (m, 2H), 2.10 – 1.80 (m, 6H), 1.80 - 1.58 (m, 10H), 1.58 – 1.40 (m, 5H), 1.37 (dd, $J = 6.6, 3.5$ Hz, 4H), 1.34 - 1.26 (m, 6H), 1.18 – 0.94 (m, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 0.64$ min. MS (ESI⁺): $C_{47}H_{79}N_{13}O_8^+$ Calculated m/z 953.6(M), 476.8 (M+2), 317.9 (M+3), 238.4 (M+4), found M (ND), 477.9 ((M+2)+H), 319.0 ((M+3)+H), 239.5 ((M+4)+H).



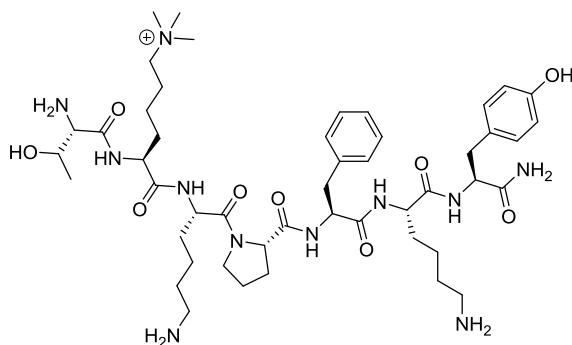
153: Peptide **153** was prepared and purified

using the same procedure as for peptide **145**. Synthesis was started from 50 mg of Rink amide MBHA resin, 2.5 mg (8.8% yield) white solid as the TFA salt. 1H NMR (400 MHz, cd_3od) δ 7.29 - 7.16 (m 5H), 7.08 (dd, $J = 8.5, 1.9$ Hz, 2H), 6.71 (d, $J = 8.0$ Hz, 2H), 4.62 – 4.45 (m, 5H), 4.45 - 4.33 (m, 2H), 4.27 - 4.19 (m, 1H), 3.88 - 3.76 (m, 1H), 3.70 (dd, $J = 12.6, 6.0$ Hz, 1H), 3.67 3. 57 (m, 2H), 3.22 – 2.78 (m, 15H), 2.24 - 2.09 (m, 2H), 2.01 – 1.58 (m, 21H), 1.55 - 1.38 (m, 5H), 1.36 - 1.25 (m, 5H), 1.09 – 0.97 (m, 6H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.25$ min. MS (ESI⁺): $C_{49}H_{79}N_{11}O_8^+$ Calculated m/z 949.6 (M), 474.8 (M+2), 316.5 (M+3), 237.4 (M+4), found M (ND), 475.9 ((M+2)+H), 317.7 ((M+3)+H), 238.6 ((M+4)+H).



154: Peptide **154** was prepared and purified

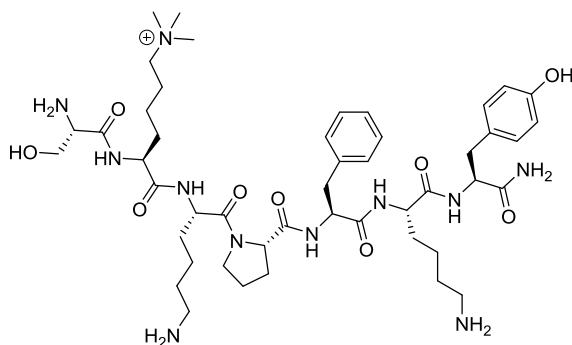
using the same procedure as for peptide **145**. Synthesis was started from 50 mg of Rink amide MBHA resin, 5.2 mg (19% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.31 - 7.16 (m, 5H), 7.08 (dd, J = 8.4, 2.1 Hz, 2H), 6.71 (d, J = 8.3 Hz, 2H), 4.613 - 4.47 (m, 4H), 4.45 - 4.33 (m, 2H), 4.26 - 4.18 (m, 2H), 3.86 - 3.75 (m, 1H), 3.71 (dd, J = 13.4, 5.9 Hz, 2H), 3.67 - 3.57 (m, 1H), 3.17 - 3.0 (m, 5H), 2.94 - 2.84 (m, 13H), 2.24 - 2.08 (m, 1H), 2.0 - 1.57 (m, 20H), 1.56 - 1.23 (m, 7H), 1.10 - 0.97 (m, 6H). LC-MS (λ = 254 nm): 99%, t_{R} = 3.11 min. MS (ESI $^{+}$): $\text{C}_{48}\text{H}_{77}\text{N}_{11}\text{O}_8^{+}$ Calculated m/z 935.6 (M), 467.8 (M+2), 311.9 (M+3), 233.9 (M+4), found M (ND), 468.9 ((M+2)+H), 313.0 ((M+3)+H), 235.0 ((M+4)+H).



155: Peptide **155** was prepared and purified

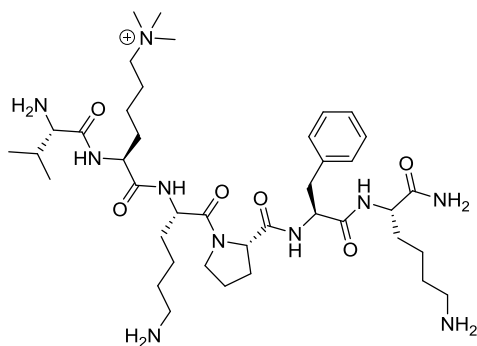
using the same procedure as for peptide **145**. Synthesis was started from 25 mg of Rink amide MBHA resin, 4.2 mg (29% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.31 - 7.17 (m, 5H), 7.08 (d, J = 8.8 Hz, 2H), 6.71 (d, J = 8.8 Hz, 2H), 4.61 - 4.46 (m, 3H), 4.46 -

4.36 (m, 2H), 4.22 (t, $J = 7.0$ Hz, 1H), 4.07 (dd, $J = 12.6, 6.3$ Hz, 1H), 3.87 -3.73 (m, 2H), 3.67 – 3.57 (m, 1H), 3.17 – 3.06 (m, 9H), 3.06 – 2.80 (m, 7H), 2.22 -2.08 (m, 1H), 2.02 – 1.56 (m, 15H), 1.56 – 1.39 (m, 4H), 1.38 – 1.22 (m, 5H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 2.9$ min. MS (ESI⁺): C₄₈H₇₈N₁₁O₉⁺ Calculated m/z 952.6 (M), 476.3 (M+2), 317.5 (M+3), 238.2 (M+4), found M (ND), 476.9 ((M+2)+H), 318.4 ((M+3)+H), 239.5 ((M+4)+H).



156: Peptide **156** was prepared and purified

using the same procedure as for peptide **145**. Synthesis was started from 25 mg of Rink amide MBHA resin, 3.6 mg (26% yield) white solid as the TFA salt. ¹H NMR (400 MHz, cd₃od) δ 7.24 (t, $J = 7.1$ Hz, 5H), 7.08 (t, $J = 8.4$, 2H), 6.71 (t, $J = 8.5$, 2H), 4.61 – 4.35 (m, 6H), 4.23 (t, $J = 7.0$ Hz, 1H), 4.10 – 3.99 (m, 1H), 3.99 – 3.86 (m, 3H), 3.85 -3.69 (m, 2H), 3.67 - 3.54 (m, 1H), 3.16 – 3.06 (m, 10H), 3.06 – 2.77 (m, 8H), 2.22 - 2.07(m, 2H), 1.99 – 1.56 (m, 18H), 1.56 -1.39 (m, 5H), 1.39 - 1.22 (m, 3H). LC-MS ($\lambda = 254$ nm): 99%, $t_R = 3.1$ min. MS (ESI⁺): C₄₇H₇₆N₁₁O₉⁺ Calculated m/z 938.6, found 938.5 (M+H).



157: Peptide **157** was prepared and purified using

the same procedure as for peptide **145**. Synthesis was started from 150 mg of Rink amide MBHA resin, 5.6 mg (7.9% yield) white solid as the TFA salt. ^1H NMR (400 MHz, cd_3od) δ 7.35 – 7.20 (m, 5H), 4.63 – 4.48 (m, 3H), 4.47 - 4.36 (m, 2H), 4.30 -4.24 (m, 1H), 3.86 - 3.76 (m, 1H), 3.73 (d, J = 6.4 Hz, 1H), 3.69 – 3.57 (m, 1H), 3.15 – 3.10 (m, 5H), 2.93 (m, 4H), 2.27 – 2.08 (m, 2H), 2.07 – 1.89 (m, 4H), 1.89 - 1.76 (m, 6H), 1.76 – 1.59 (m, 8H), 1.59 – 1.32 (m, 7H), 1.12 – 0.97 (m, 6H). LC-MS (λ = 254 nm): 99%, t_R = 2.6 min. MS (ESI $^+$): $\text{C}_{40}\text{H}_{71}\text{N}_{10}\text{O}_6^+$ Calculated m/z 787.6, found 787.5 (M).

Analytical and structural characterization (LC-MS and NMR)

Analytical LCMS data for all compounds were acquired using an Agilent 6110 series system with the UV detector set to 220 and 254 nm. Samples were injected (<10 μL) onto an Agilent Eclipse Plus 4.6×50 mm, 1.8 μm , C18 column at room temperature. A mobile phase of A (H_2O + 0.1% acetic acid) and B (MeOH + 0.1% acetic acid) was used with a linear gradient from 10% to 100% B in 5.0 min, followed by a flush at 100% B for another 2 minutes with a flow rate of 1.0 mL/min. Mass spectra data were acquired in positive ion mode using an Agilent 6110 single quadrupole mass spectrometer with an electrospray ionization source. Nuclear Magnetic Resonance (NMR) spectra were recorded on a Varian Mercury spectrometer at

400 MHz for proton (^1H NMR) and 100 MHz for carbon (^{13}C NMR); chemical shifts are reported in ppm (δ). All synthesized peptides were dissolved in deuterated MeOH.

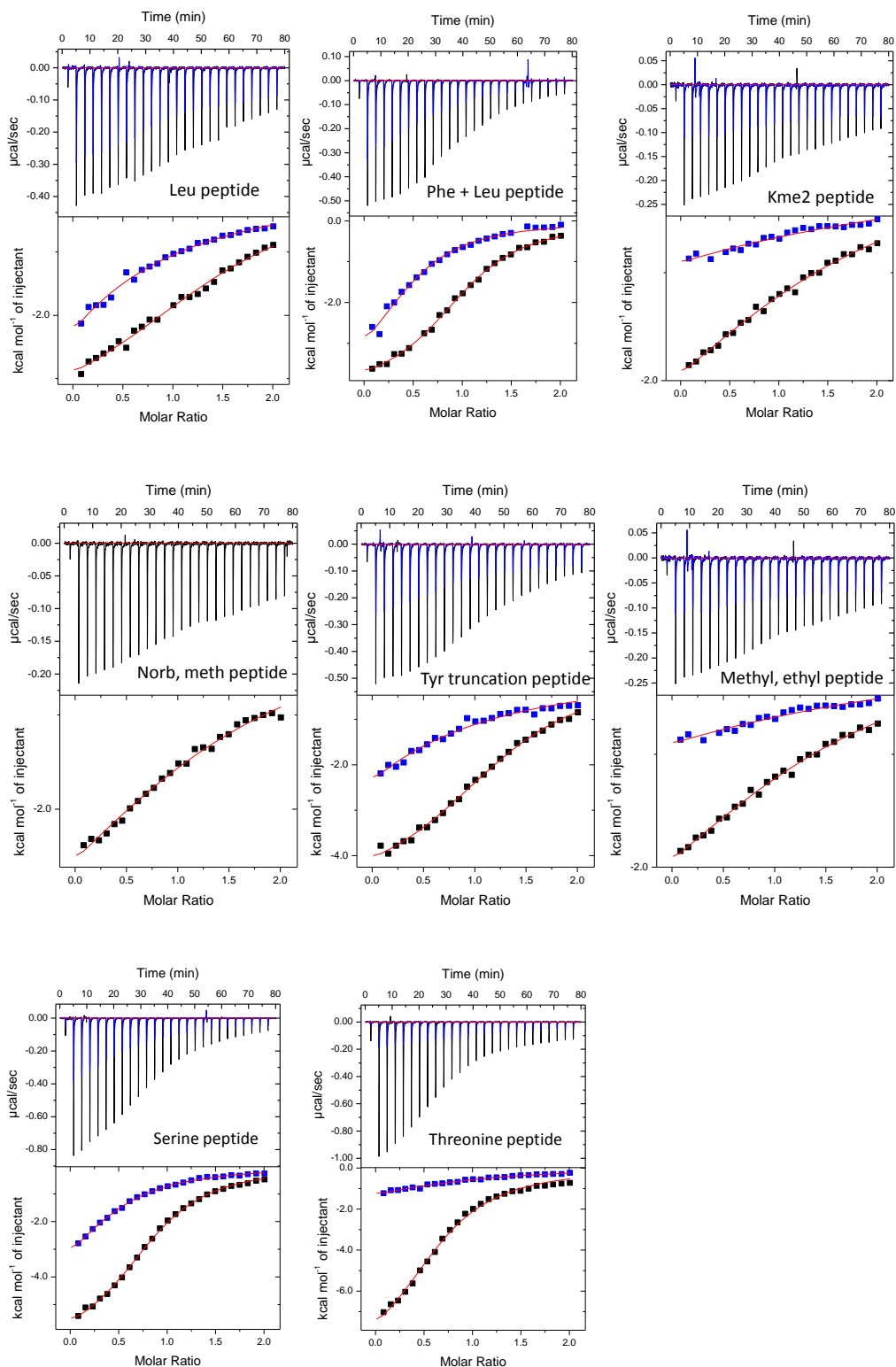
AlphaScreen assay of synthesized peptide ligands

The AlphaScreen assay was generally performed as previously described.[105] The general method for screening synthesized peptide ligands is as follows. Peptides were weighed out as solids and then dissolved in de-ionized water to a concentration of 10 mM. Compound plates were then prepared using an automated liquid handling system (Tecan, Freedom Evo). Compound plates (1 μL at 10 mM highest concentration; 3-fold, 10-point dilutions in DMSO) were then diluted in 1X assay buffer (20 mM TRIS pH 8.0, 25 mM NaCl, 2 mM DTT and 0.05% Tween-20) to 1 mM using a Multimek robotic pipettor (Nanoscreen) and 1 μL was spotted into the wells of 384-well low-volume Proxiplates (PerkinElmer). To these plates 9 μL of protein-peptide mix in 1X assay buffer was added by Multidrop (Thermo) to bring the final compound concentration to 100 μM and incubated for 30 min at room temperature. Next, 2 μL of a 1:1 mixture of streptavidin-conjugate donor and nickel-chelate acceptor beads (45 $\mu\text{g}/\text{mL}$ in 1X assay buffer) were added and the plates were allowed to incubate for an additional 30 min in the dark at room temperature. After incubation, the plates were read on an EnVision multi-label reader equipped with an HTS AlphaScreen laser (Perkin Elmer). The IC_{50} values reported are the average of at least 2 values \pm the standard deviation. When IC_{50} values for a single compound were not all active ($< 100 \mu\text{M}$) or inactive ($> 100 \mu\text{M}$), the IC_{50} values were calculated using 4-paramter curve fitting (GraphPad Prism 5) from replicate runs using averaged response values for each compound concentration.

Isothermal titration calorimetry binding experiments

All ITC measurements were recorded at 25 °C with an AutoITC₂₀₀ microcalorimeter (MicroCal Inc.). All protein and compound stock samples were prepared in the target buffer (25 mM Tris-HCl, pH 8, 150 mM NaCl, and 2 mM β -mercaptoethanol), and then diluted in the same buffer to achieve the desired concentrations: 90 μ M protein and 1 mM compound depending on the expected dissociation constant. The concentration of protein stock solutions were established using the Edelhoch method, whereas 10 mM compound stock solutions were prepared gravimetrically based on molecular weight. A typical experiment included a single 0.2 μ L compound injection into a 200 μ L cell filled with protein, followed by 25 subsequent 1.5 μ L injections of compound. Injections were performed with a spacing of 180 seconds and a reference power of 8 μ cal/sec. The titration data was analyzed using Origin Software (MicroCal Inc.) by non-linear least squares, fitting the heats of binding as a function of the compound:protein ratio. The data were fit based on a one set of sites model.

ITC experiments of remaining peptides are shown below. In the figure, the black curve represents PHF1 binding and the blue line represents PHF19 binding.



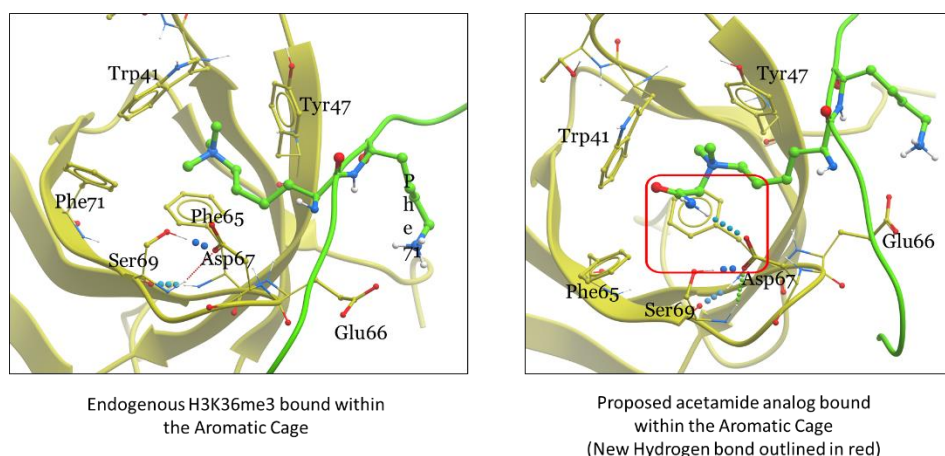
Future work

Current efforts toward developing a highly potent and selective peptide ligand for PHF1 and PHF19 have provided a micromolar ligand for PHF1 and PHF19. This ligand (**148**) is a 7-mer peptide that has a single amino acid substitution of a phenylalanine for His39 residue position on the peptide. Efforts toward discovering a mimic of the trimethyl-lysine residue has so far proven unsuccessful and that will be a key focus of future research efforts. It has been shown that the methyl-lysine binding pocket has a limited tolerance for the steric bulk of the trimethyl-lysine mimetic. Both the *N',N'*-di-ethyl and *N'*-methyl, *N'*-norbornyl derivatives appear to be too large to successfully bind within the pocket. Additionally, the *N'*-methyl, *N'*-ethyl and *N',N'*-dimethyl analogs are respectively 17 and 7 times less potent than the phenylalanine mutant peptide that contains Kme3. Future efforts should consider synthesizing trimethyl-lysine analogs that utilize a substituted ornithine residue as the decreased residue length may assist in docking within the aromatic binding pocket. An additional functional group to try would be an azetidine head group.

Another quaternary amine mimetic for H3K36me3 was suggested by Prof. Dmitri Kireev (Director of Computational Biophysics and Molecular Design, Center for Integrative Chemical Biology and Drug Discovery) and takes a different approach than efforts conducted so far. Prof. Kireev docked the *N',N'*-methyl, ethyl analog and noted that only the methyl functional group was bound within the aromatic cage. Further comparison of this analog to the endogenously bound trimethyl-lysine peptide shows a similar binding pattern with only a single methyl group bound within the pocket. Additionally, he noted that outside of the aromatic pocket was an acidic residue (Asp67) that may be of use to gain additional hydrogen bonding. With these observations, Prof. Kireev proposed a lysine analog that possess a single N-methyl group and an N-acetamide

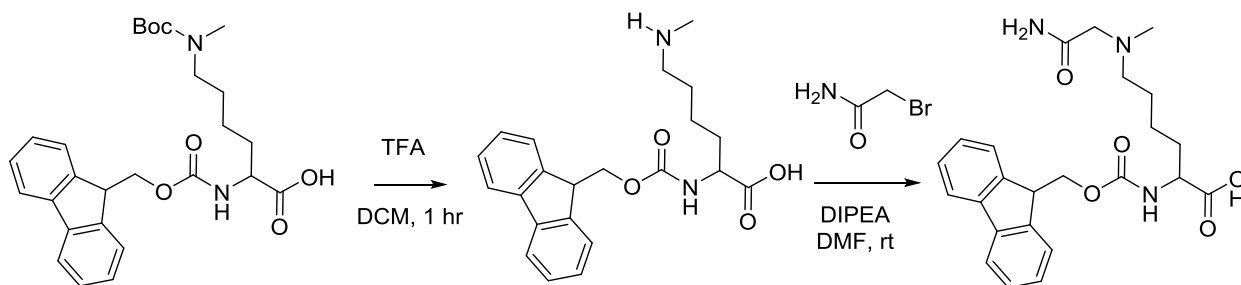
functional group. The acetamide functional group would potentially be able to gain 1 – 2 new hydrogen bonds with the acidic residues outside of the aromatic pocket while the methyl group would be bound within the aromatic pocket, Figure 3.6.

Figure 3.6: Proposed Acetamide quaternary amine mimetic bound within Tudor domain



Initial efforts to synthesize this unique analog have so far proven unsuccessful, Figure 3.7.

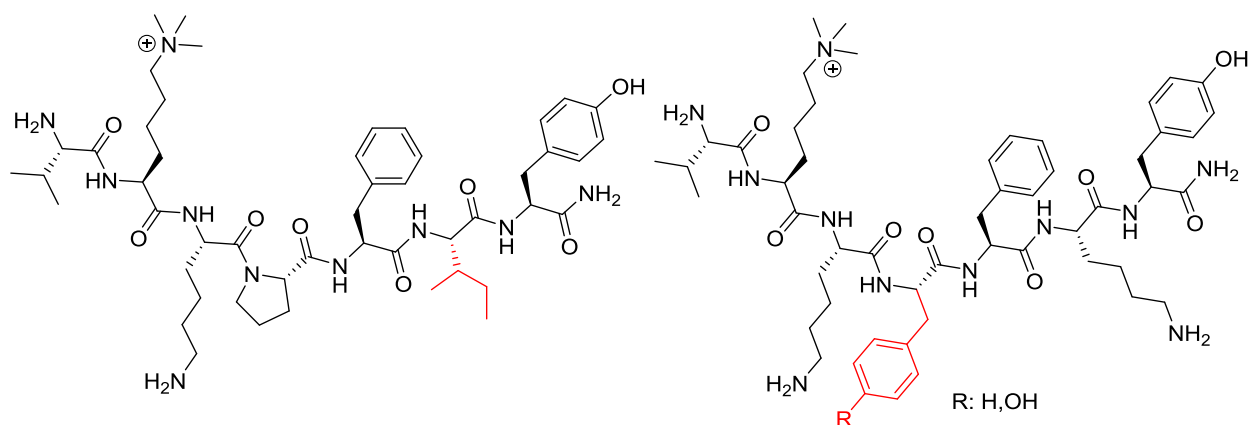
Figure 3.7: Synthesis of N'-Acetamide, N'-methyl N- α -Fmoc-Lysine



An additional area of the peptide to continue SAR exploration would be to synthesize additional Arg40 analogs that are in combination with the phenylalanine mutant. The initial SAR series showed that that a positively charged residue at this position was not necessary for binding activity. Additionally, when the lysine was mutated to a leucine it was observed that this peptide was an equipotent peptide compared to the endogenous 7-mer. The double mutant (Phe and Leu)

was observed to be slightly less potent than the single phenylalanine mutant. Since it is hypothesized that the phenylalanine and leucine residues are increasing the hydrophobic interactions with the 4 leucine residues within the protein's hydrophobic pocket, a double mutant that contained the observed phenylalanine for histidine and phenylalanine for lysine substitutions may increase binding affinity. Another residue to try at this double mutant would be isoleucine. A final area where amino acid substitutions may prove beneficial would be at the Pro38 position. This residue was shown to be important for Tudor domain binding in the Celluspot assay and substitution of a phenylalanine or tyrosine at this position may pick up additional hydrophobic interactions with the tetra-leucine wall in the protein's hydrophobic patch, Figure 3.8.

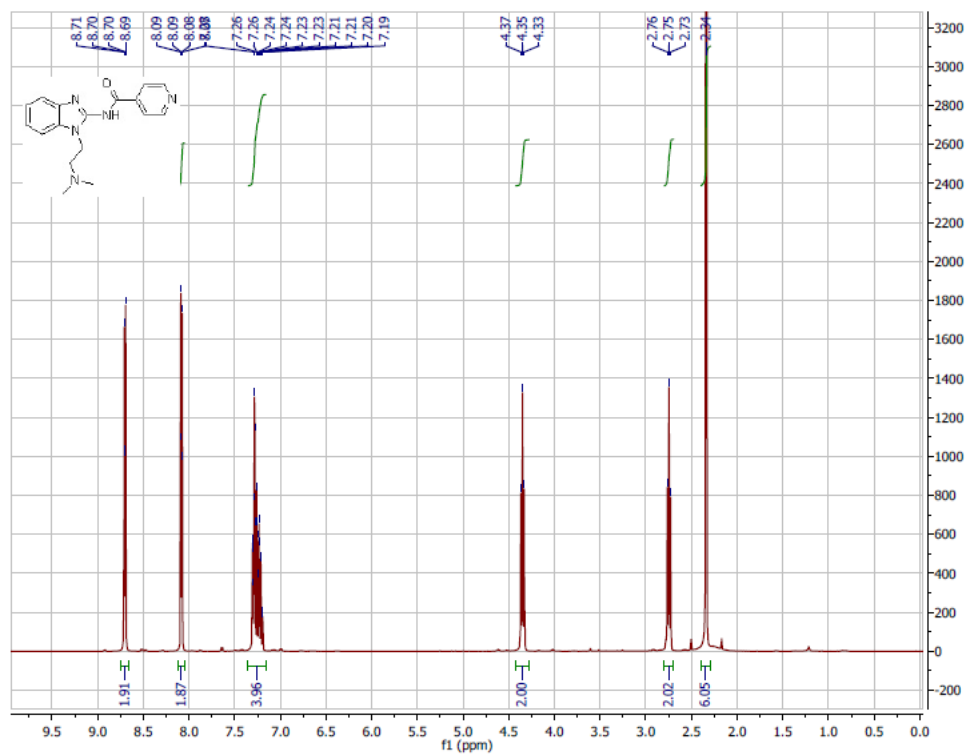
Figure 3.8: Double Mutant Analogs of Arg40 and Pro38 amino Acid Substitutions



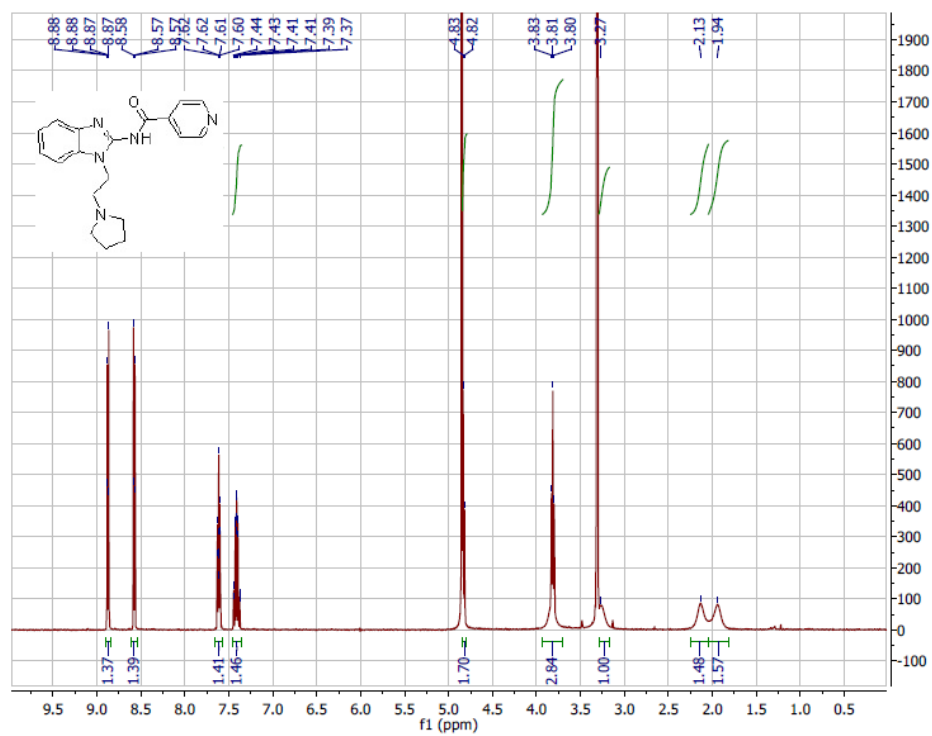
Appendix I

NMR Spectra for Chapters II and III

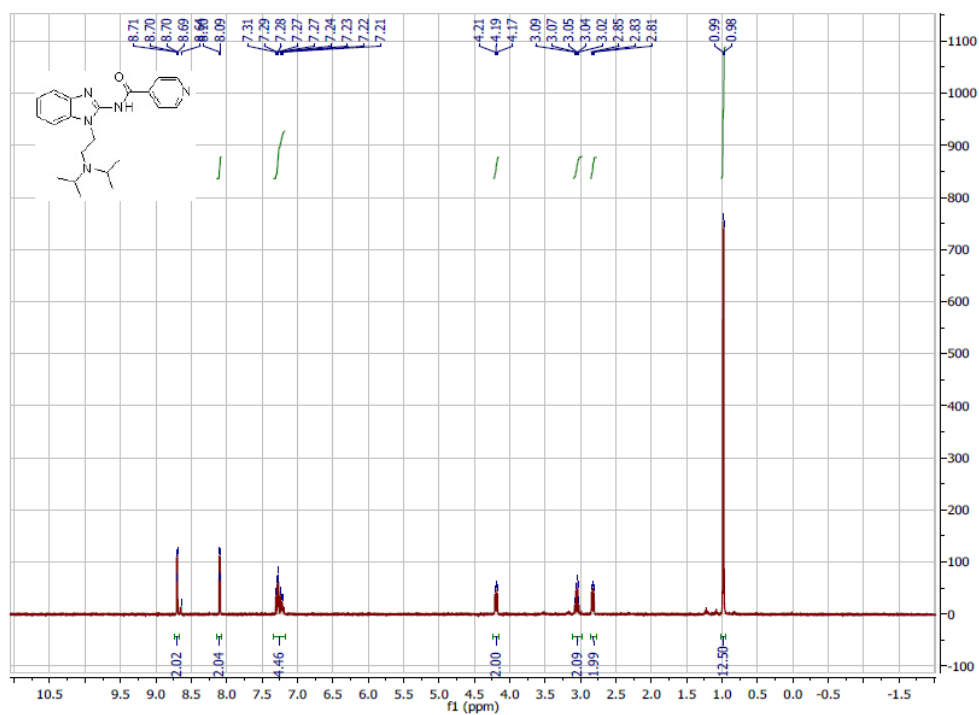
N-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)isonicotinamide (1)



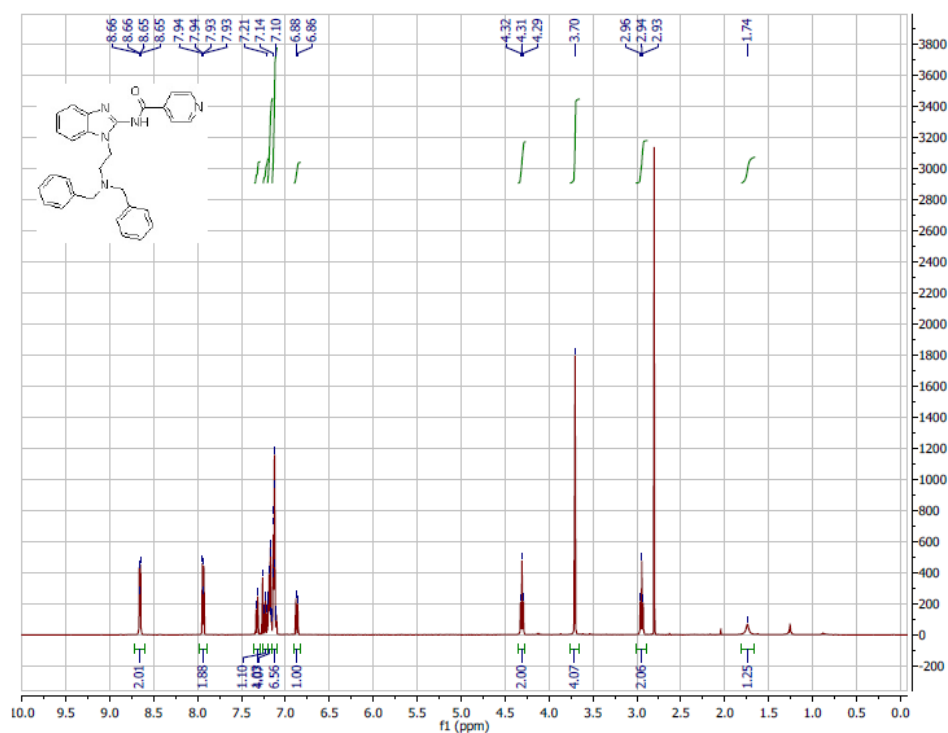
***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)isonicotinamide (2, UNC1554)**



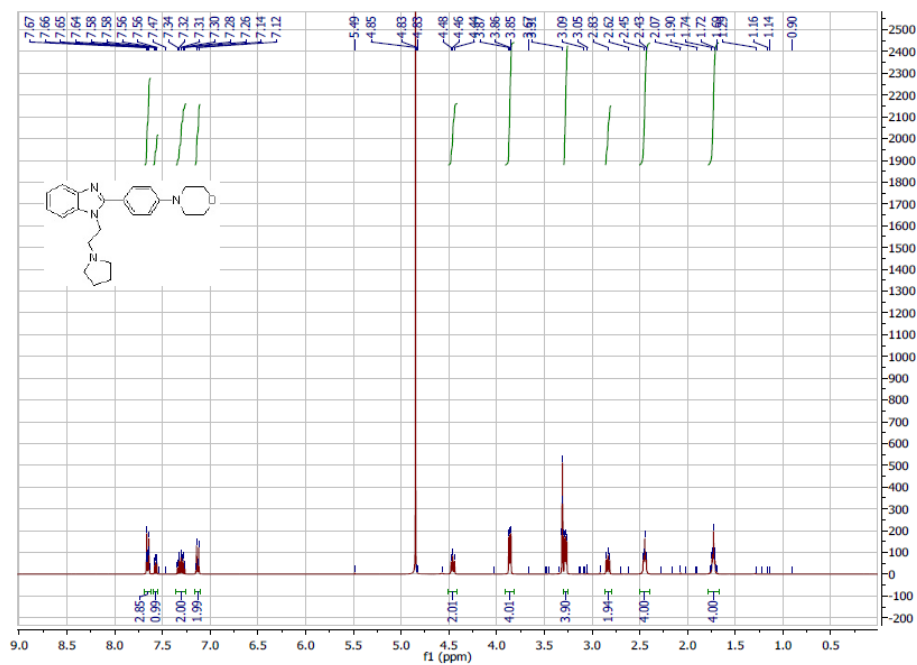
***N*-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)isonicotinamide (3)**



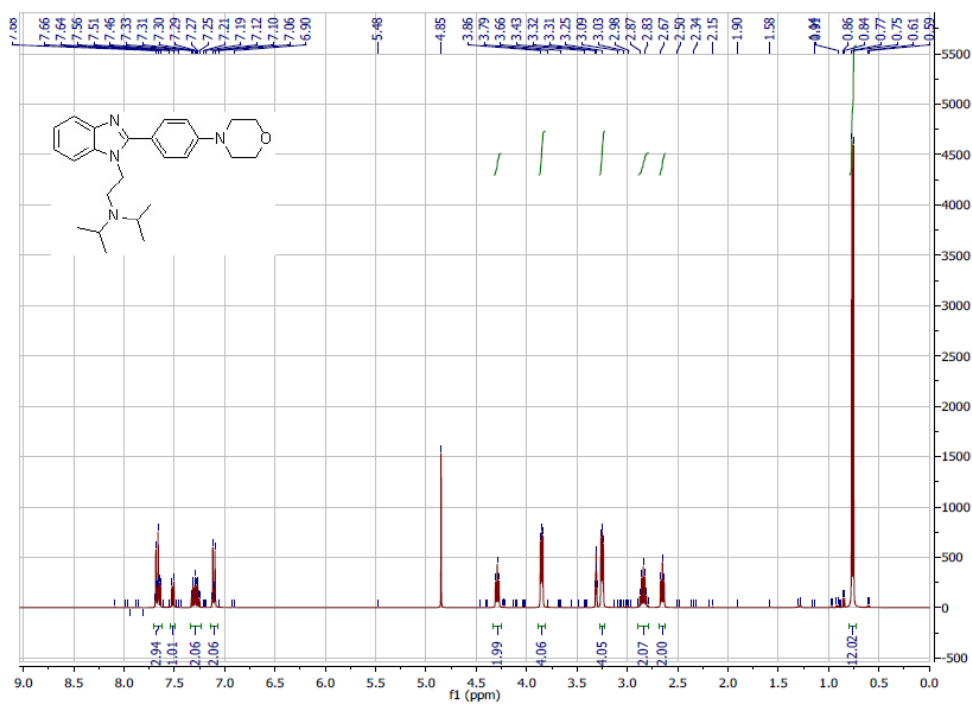
***N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)isonicotinamide (4)**



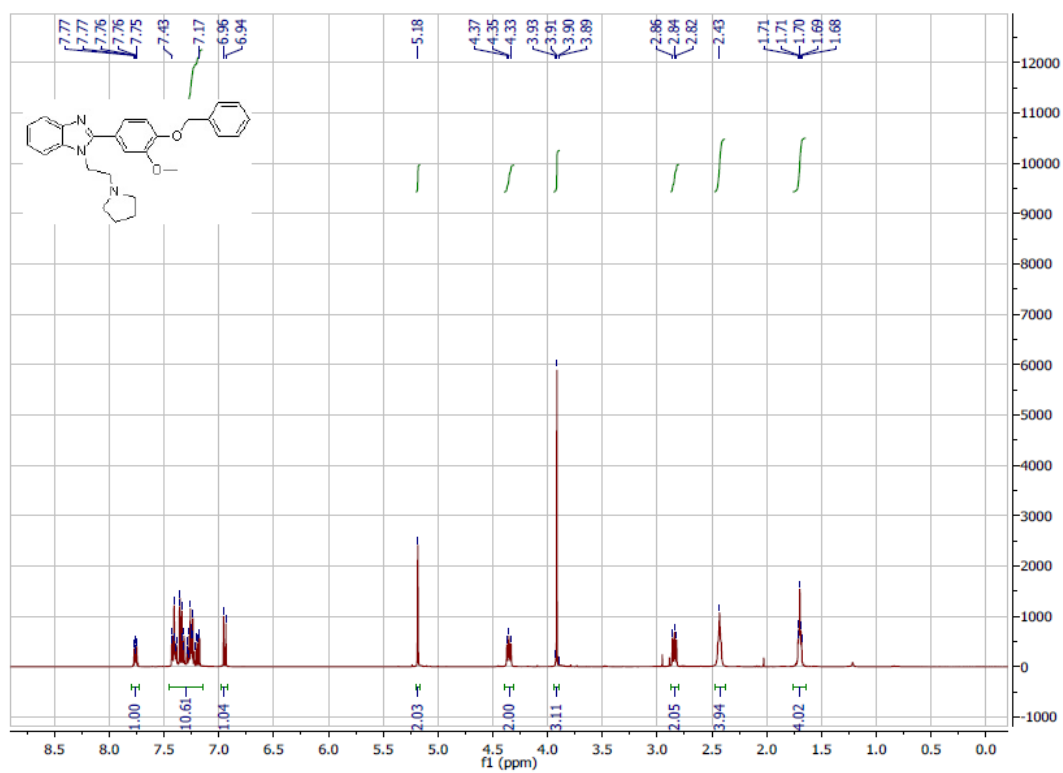
4-(4-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)phenyl)morpholine (5)



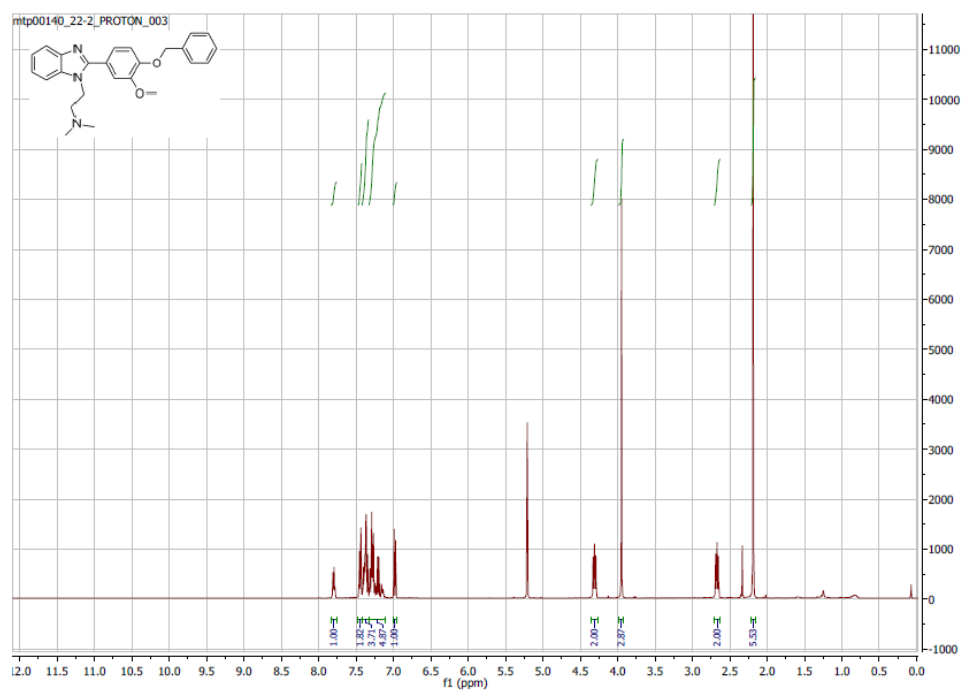
***N*-isopropyl-*N*-(2-(2-(4-morpholinophenyl)-1*H*-benzo[d]imidazol-1-yl)ethyl)propan-2-amine (6)**



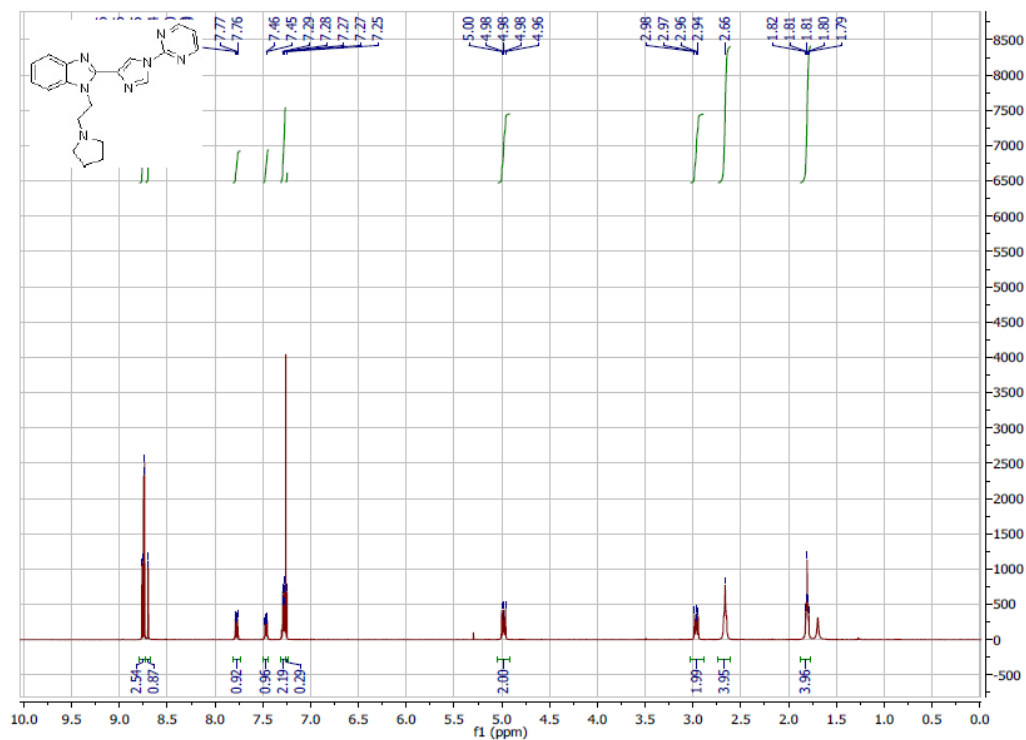
2-(4-(benzyloxy)-3-methoxyphenyl)-1-(2-(pyrrolidin-1-yl)ethyl)-1*H*-benzo[d]imidazole (7)



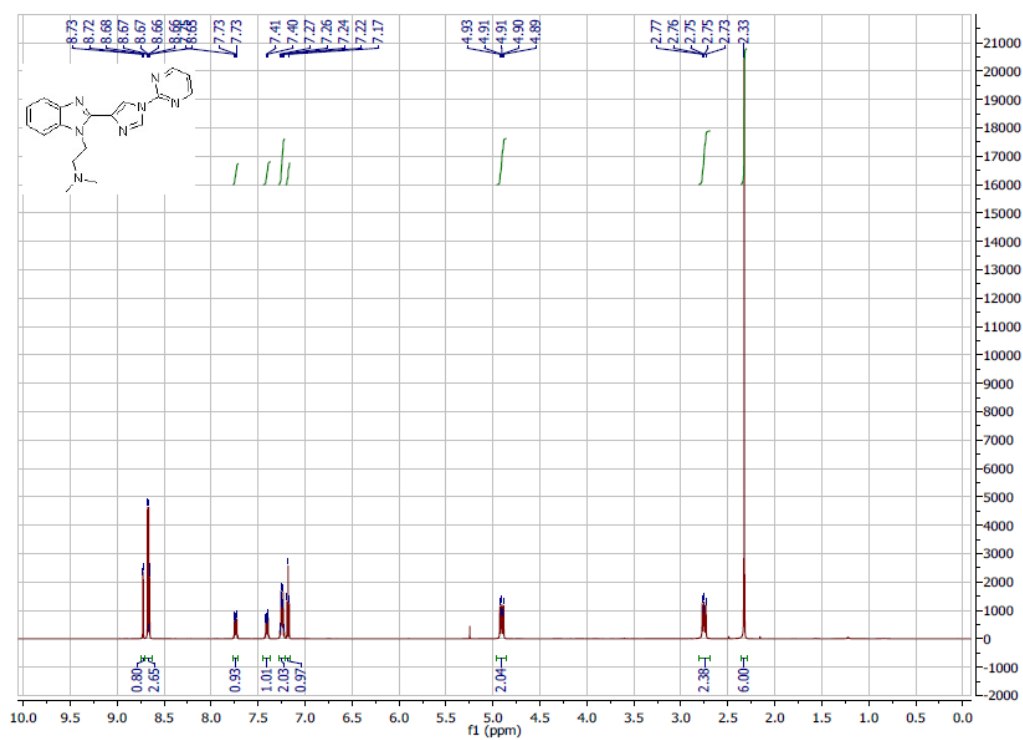
2-(2-(4-(benzyloxy)-3-methoxyphenyl)-1H-benzo[d]imidazol-1-yl)-N,N-dimethylethan-1-amine (8)



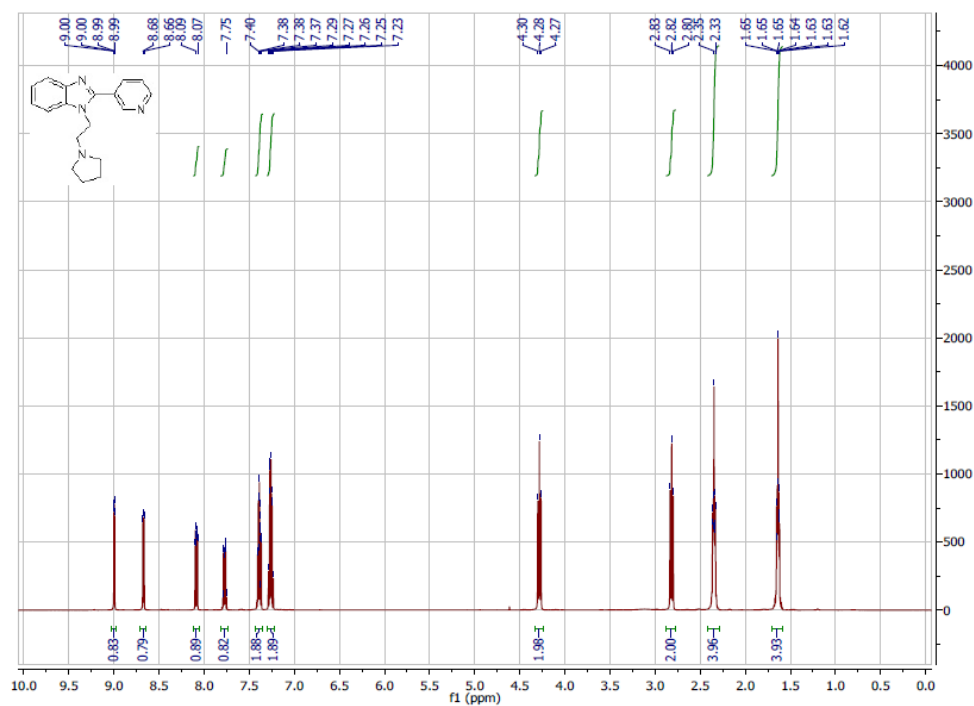
2-(1-(pyrimidin-2-yl)-1H-imidazol-4-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (9)



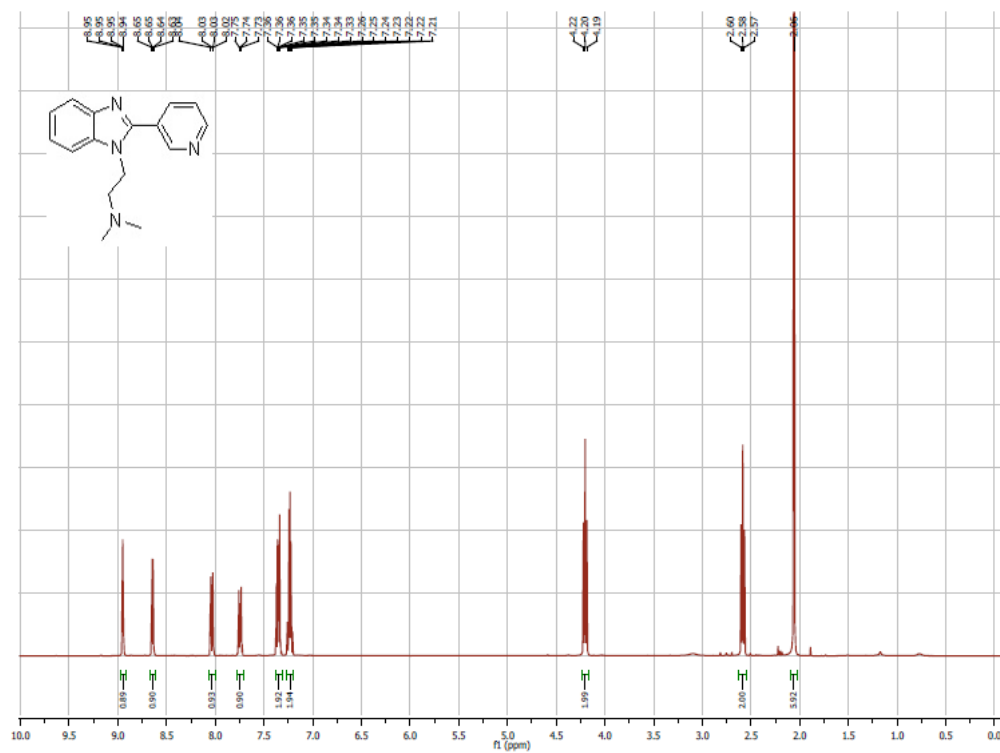
***N,N*-dimethyl-2-(2-(1-(pyrimidin-2-yl)-1H-imidazol-4-yl)-1H-benzo[d]imidazol-1-yl)ethan-1-amine (10)**



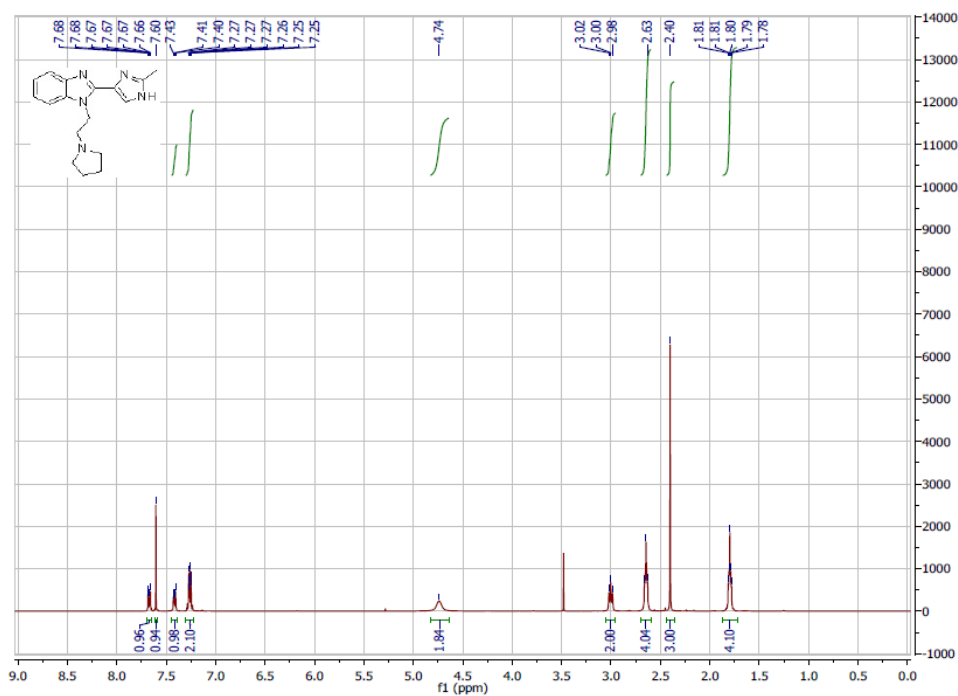
2-(pyridin-3-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (11)



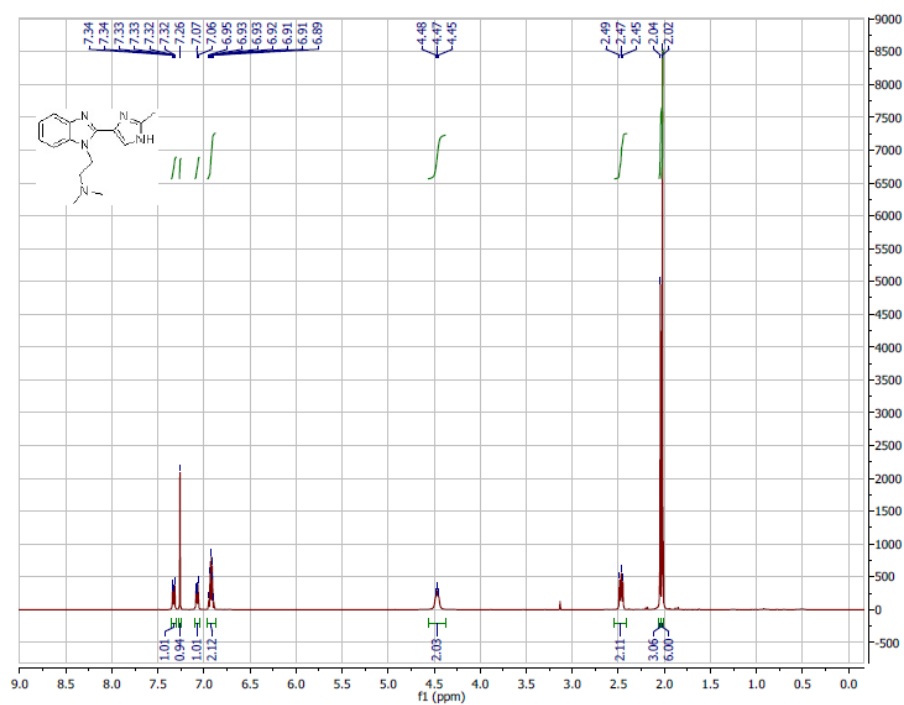
***N,N*-dimethyl-2-(2-(pyridin-3-yl)-1H-benzo[d]imidazol-1-yl)ethan-1-amine (12)**



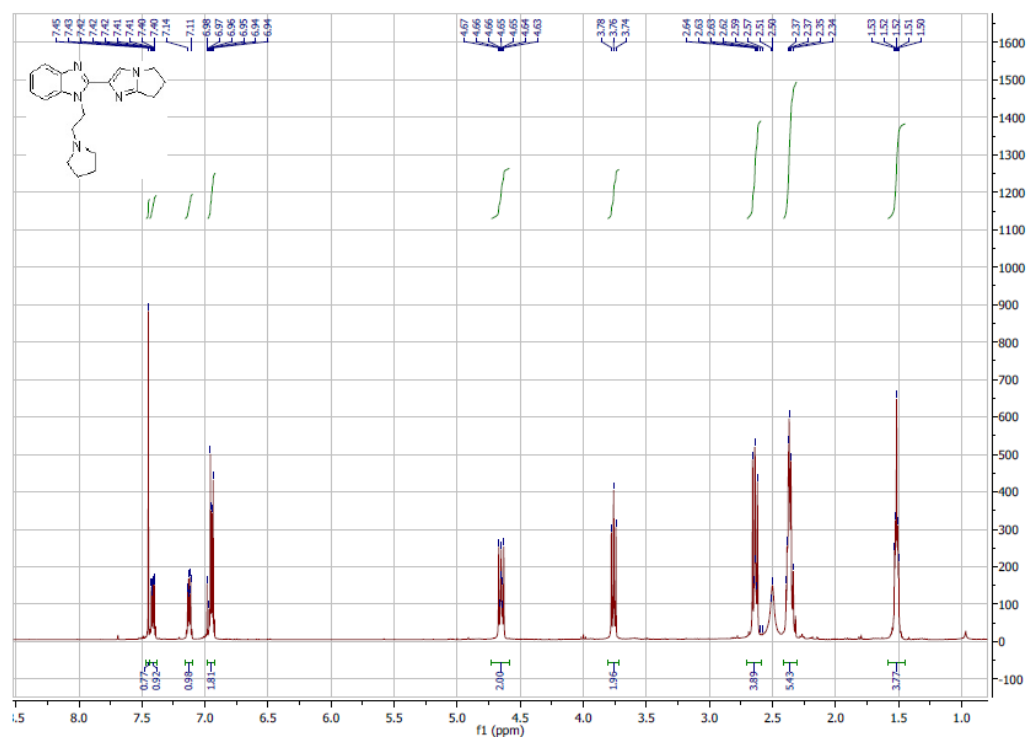
2-(2-methyl-1H-imidazol-4-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (13)



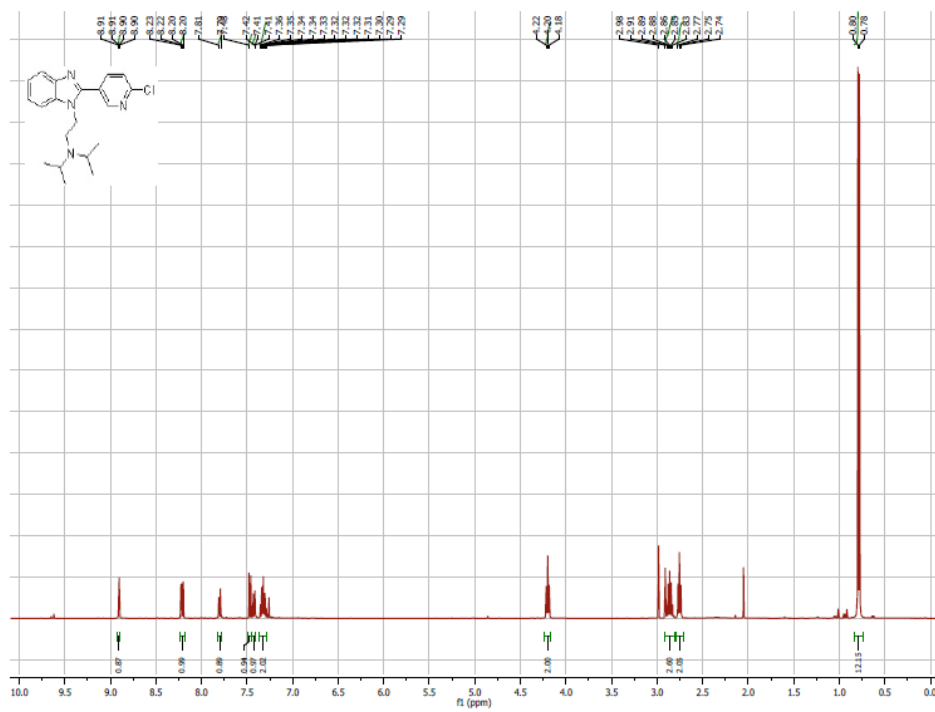
N,N-dimethyl-2-(2-(2-methyl-1H-imidazol-4-yl)-1H-benzo[d]imidazol-1-yl)ethan-1-amine (14)



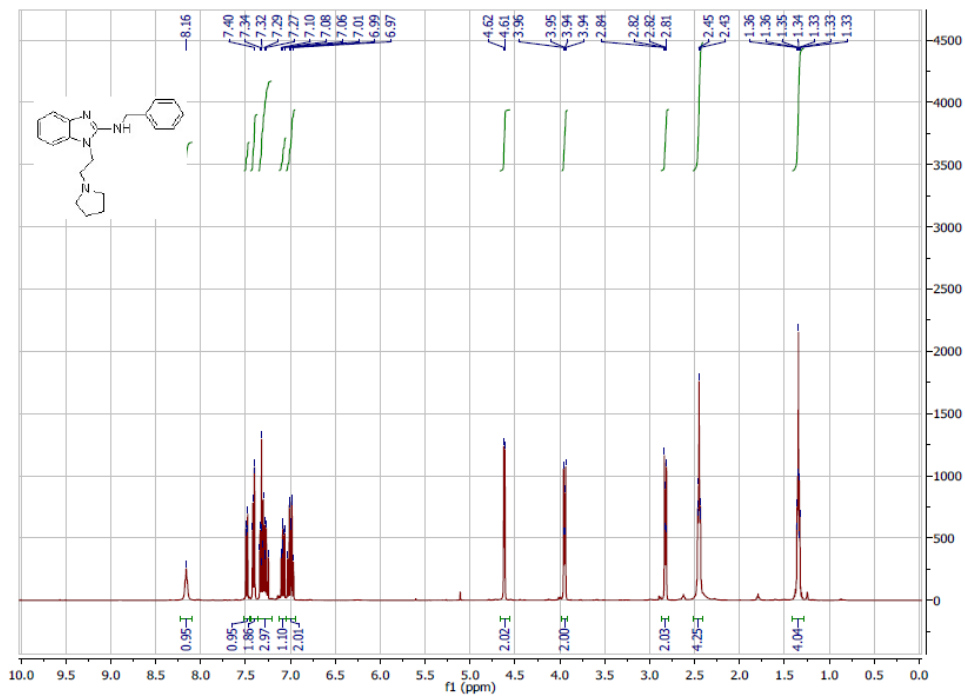
2-(6,7-dihydro-5H-pyrrolo[1,2-a]imidazol-2-yl)-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (15)



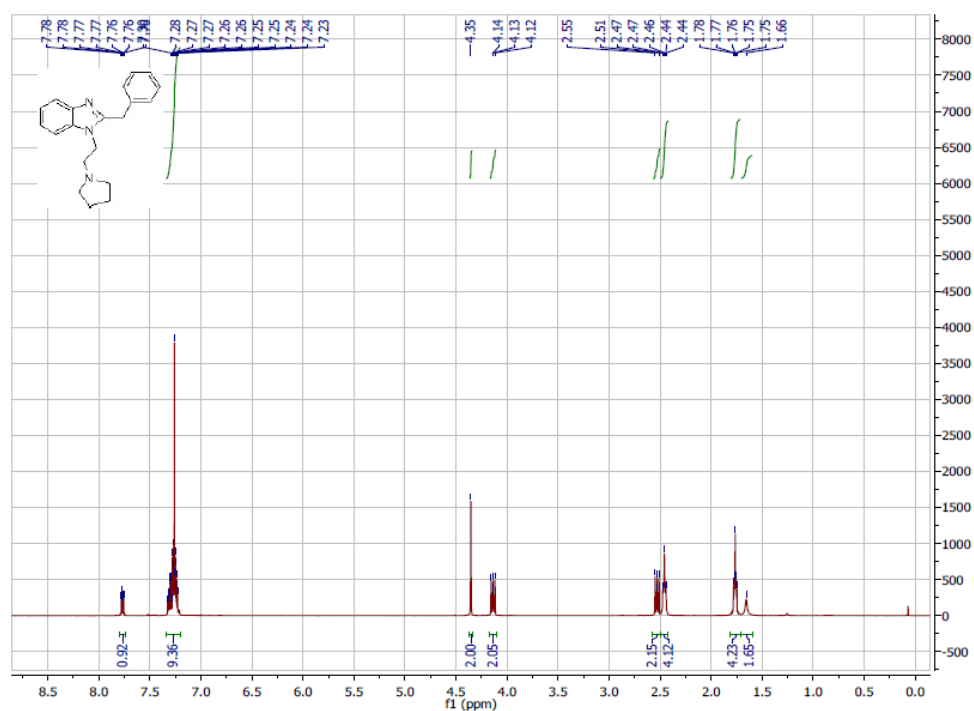
***N*-(2-(2-(6-chloropyridin-3-yl)-1H-benzo[d]imidazol-1-yl)ethyl)-*N*-isopropylpropan-2-amine (16)**



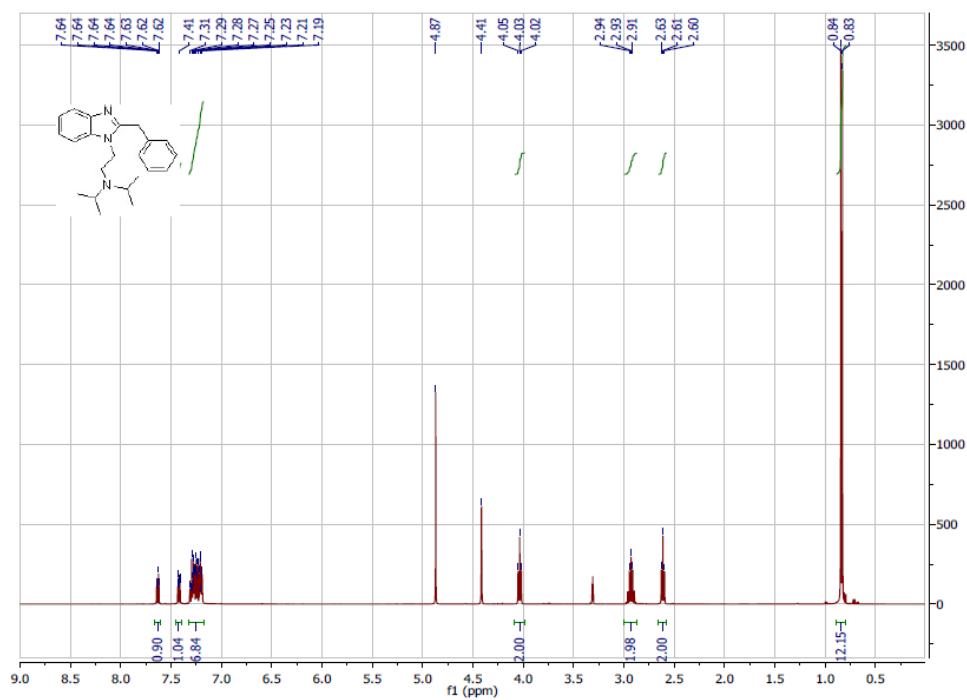
***N*-benzyl-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-amine (17)**



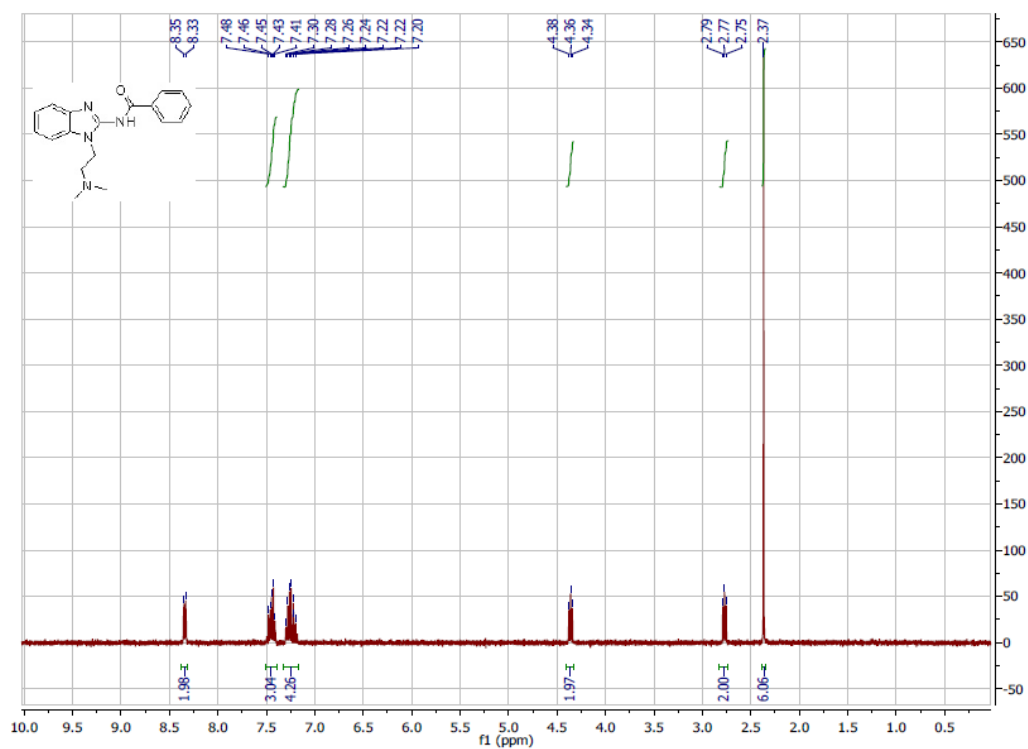
2-benzyl-1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazole (18)



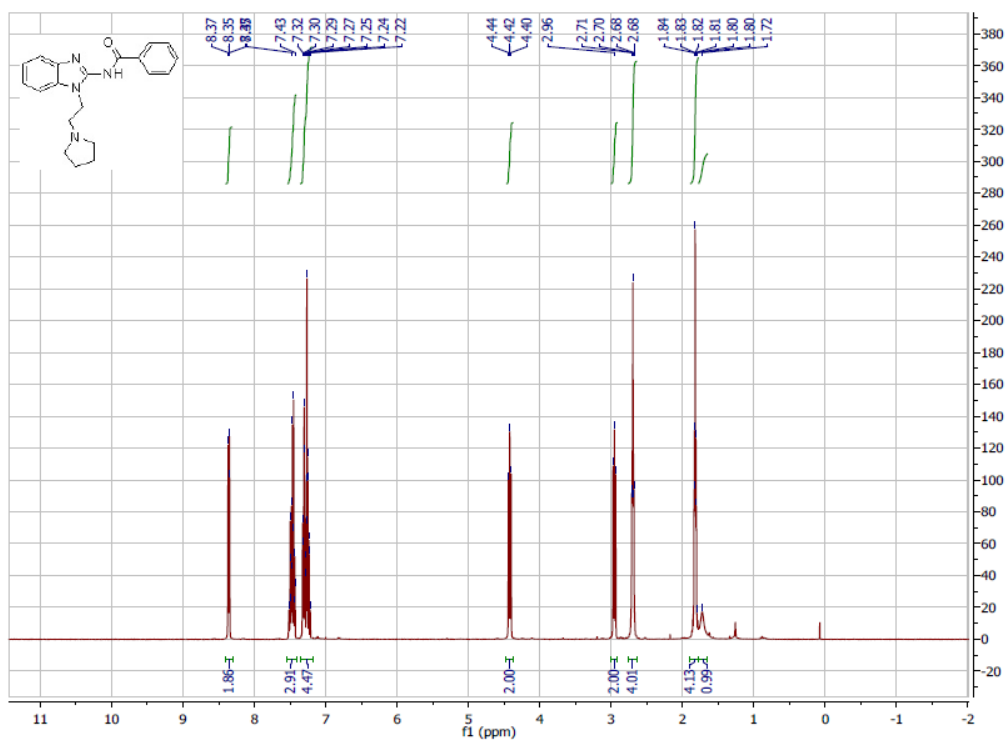
***N*-(2-(2-benzyl-1H-benzo[d]imidazol-1-yl)ethyl)-*N*-isopropylpropan-2-amine (19)**



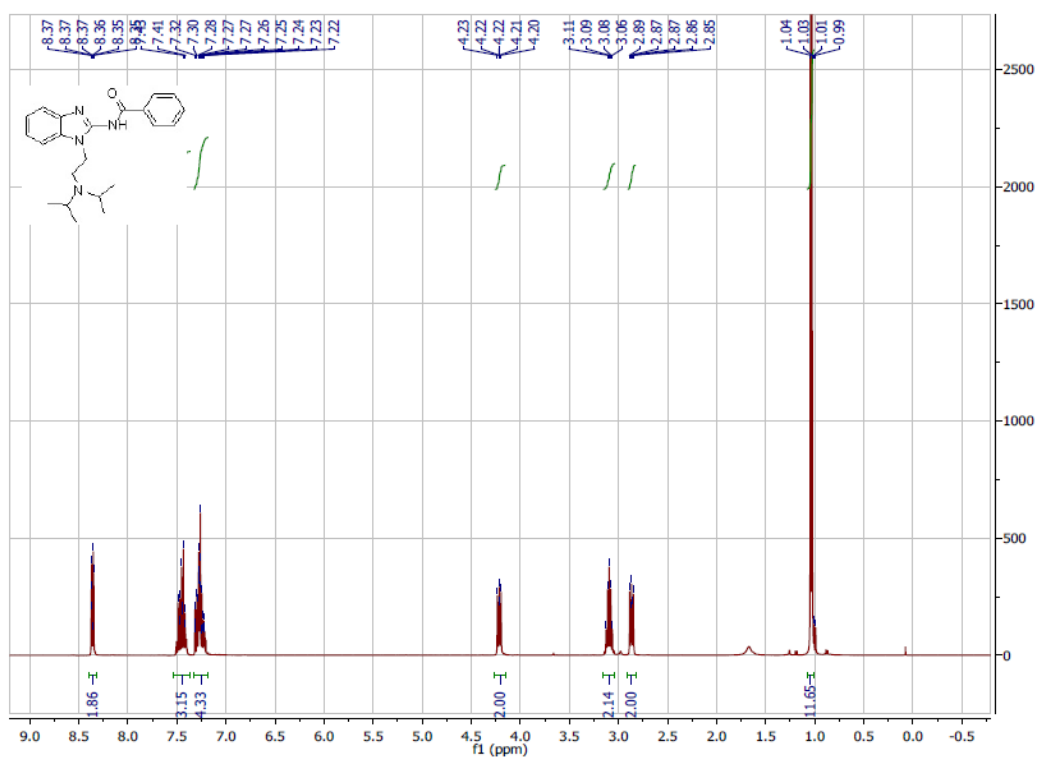
***N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (20)**



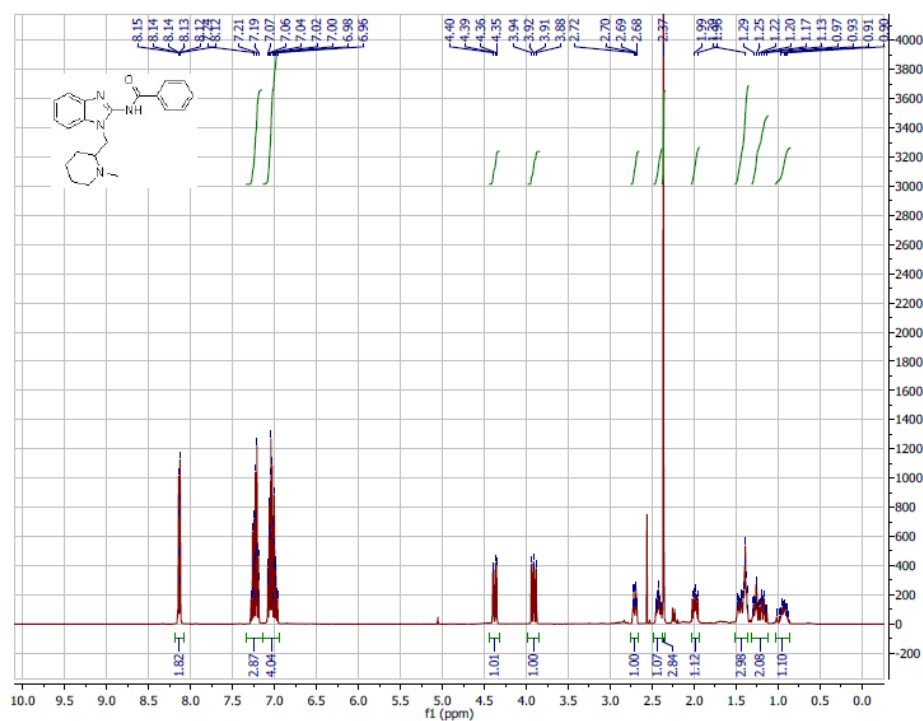
***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (21)**



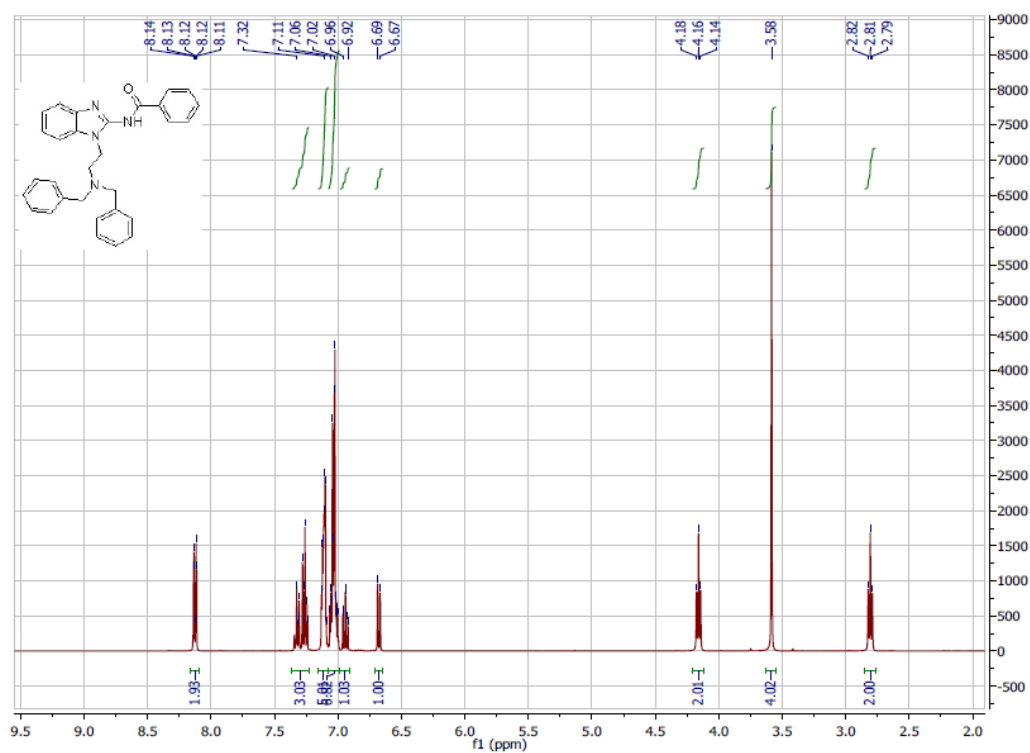
***N*-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (22)**



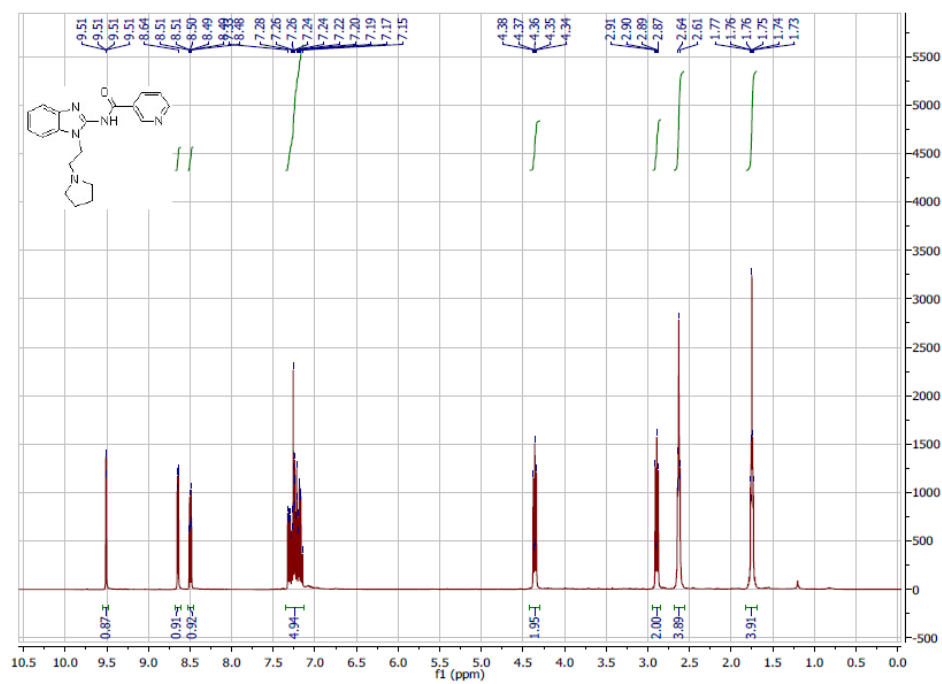
***N*-(1-((1-methylpiperidin-2-yl)methyl)-1H-benzo[d]imidazol-2-yl)benzamide (23)**



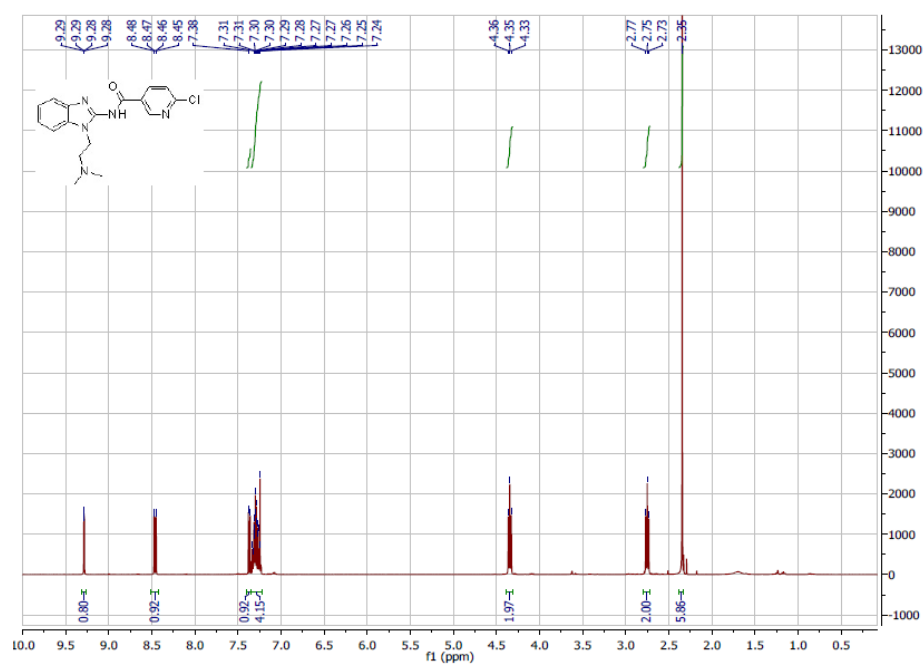
***N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (24)**



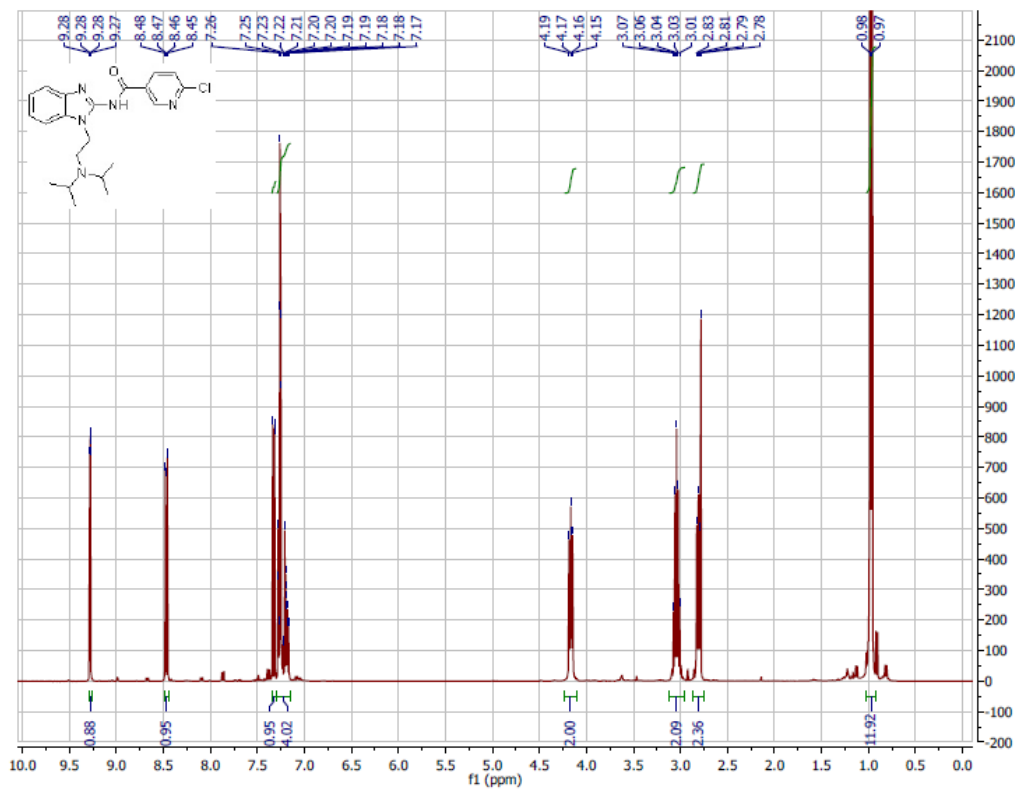
***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)nicotinamide (25)**



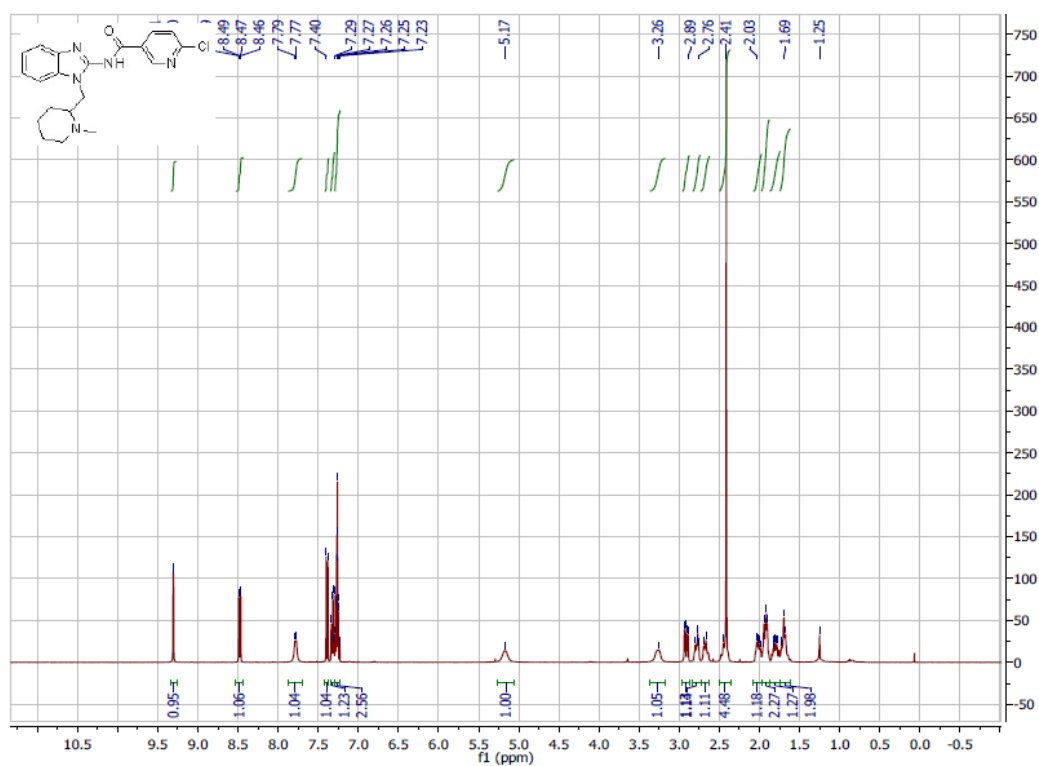
6-chloro-*N*-(1-(2-(dimethylamino)ethyl)-1*H*-benzo[d]imidazol-2-yl)nicotinamide (26)



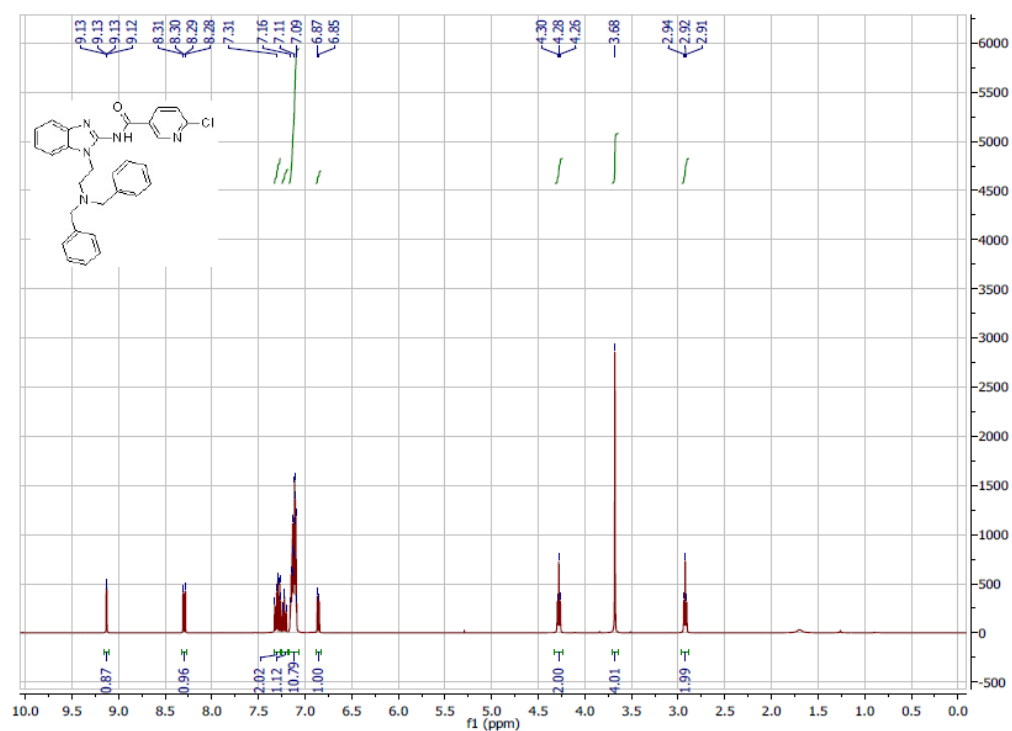
6-chloro-*N*-(1-(2-(diisopropylamino)ethyl)-1*H*-benzo[d]imidazol-2-yl)nicotinamide (27)



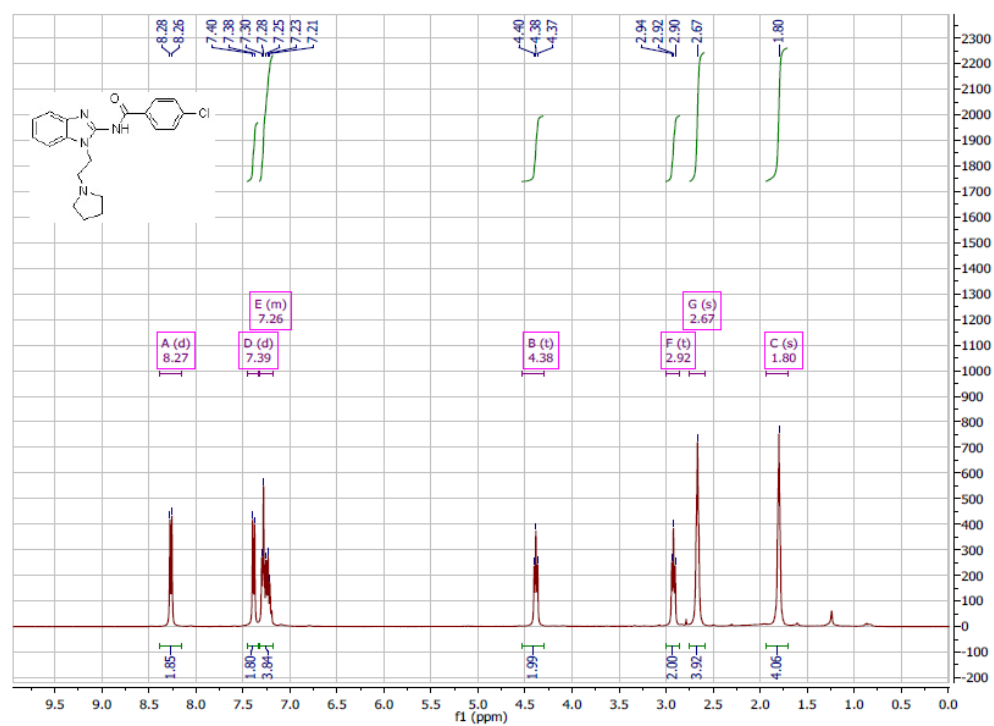
6-chloro-*N*-(1-((1-methylpiperidin-2-yl)methyl)-1H-benzo[d]imidazol-2-yl)nicotinamide (28)



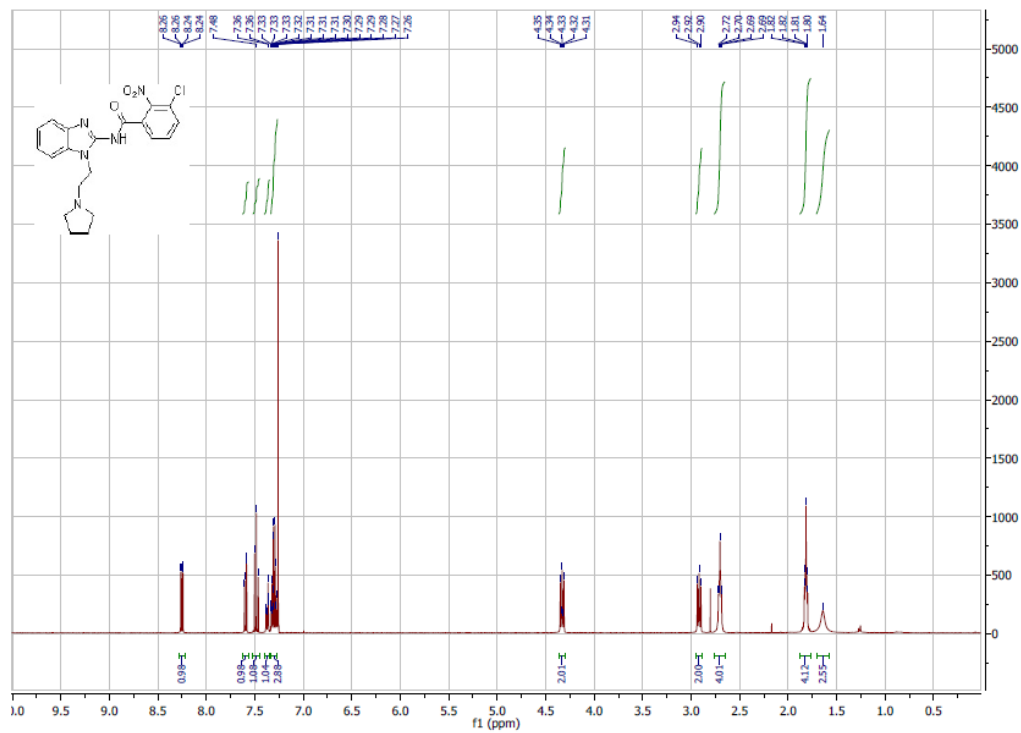
6-chloro-*N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)nicotinamide (29)



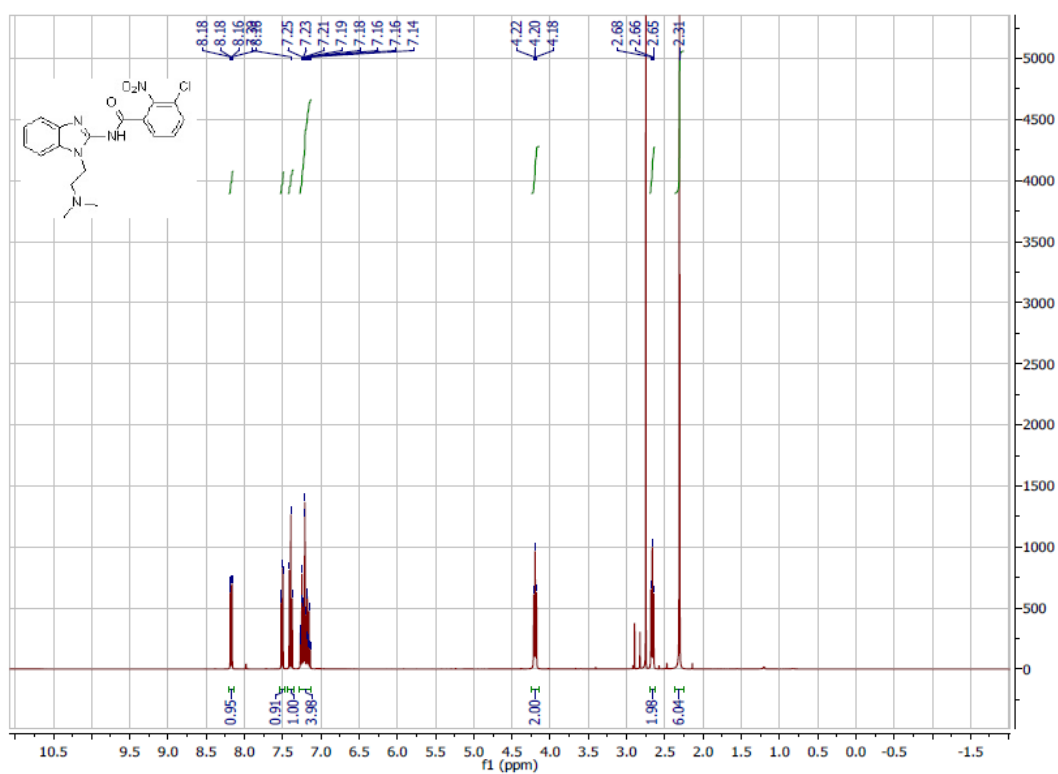
4-chloro-N-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (30)



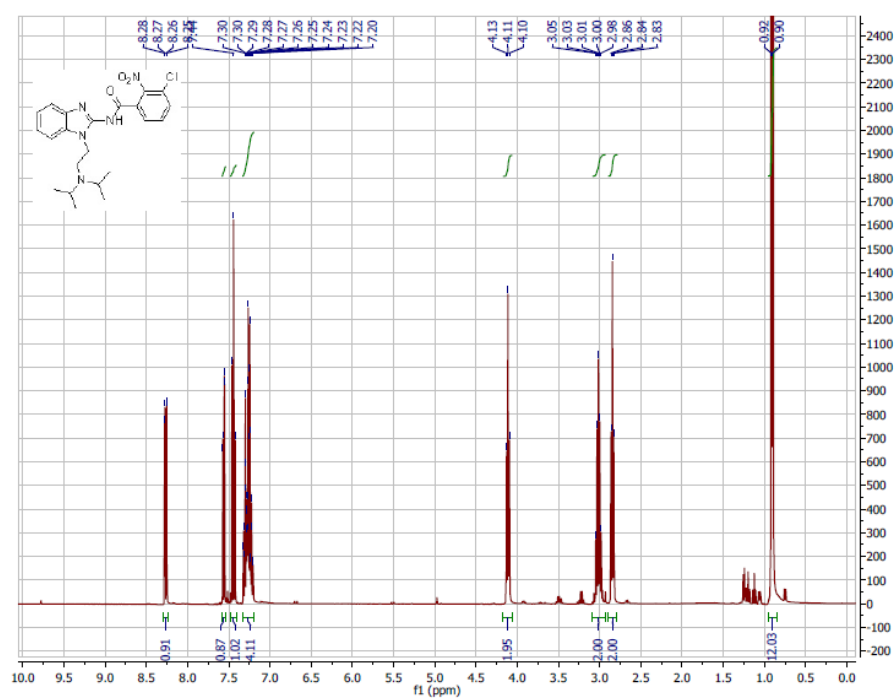
3-chloro-2-nitro-N-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)benzamide (31)



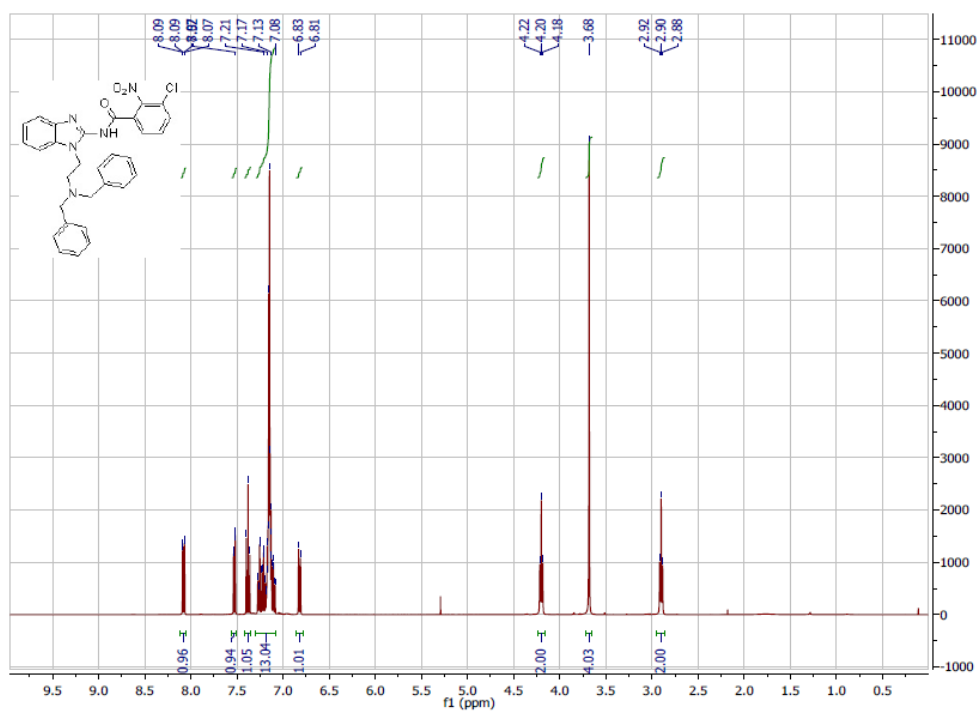
3-chloro-*N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-2-nitrobenzamide (32)



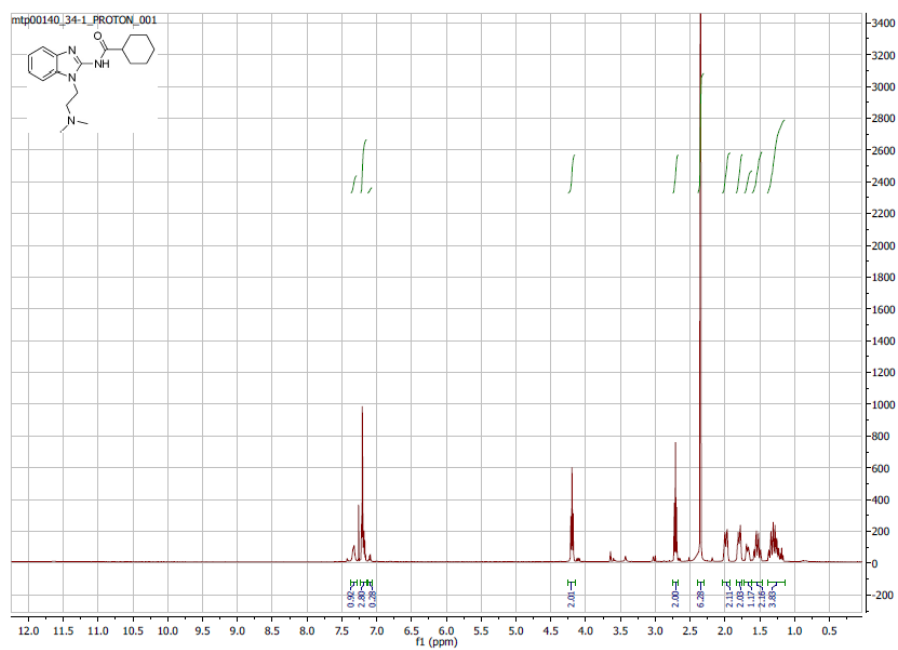
3-chloro-*N*-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-2-nitrobenzamide (33)



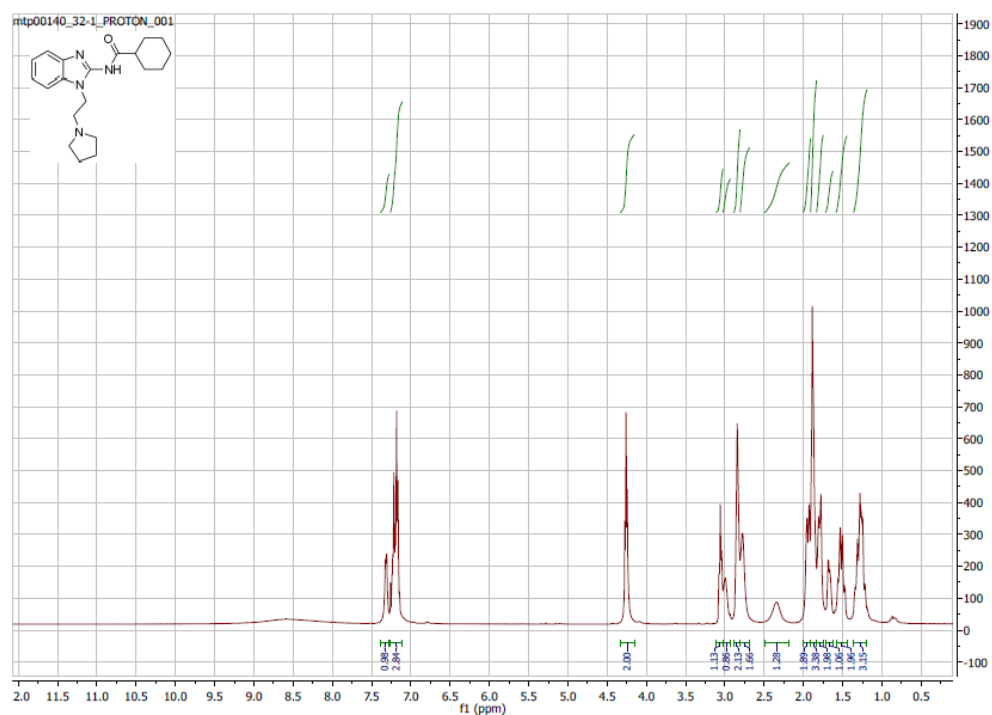
3-chloro-*N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-2-nitrobenzamide (34)



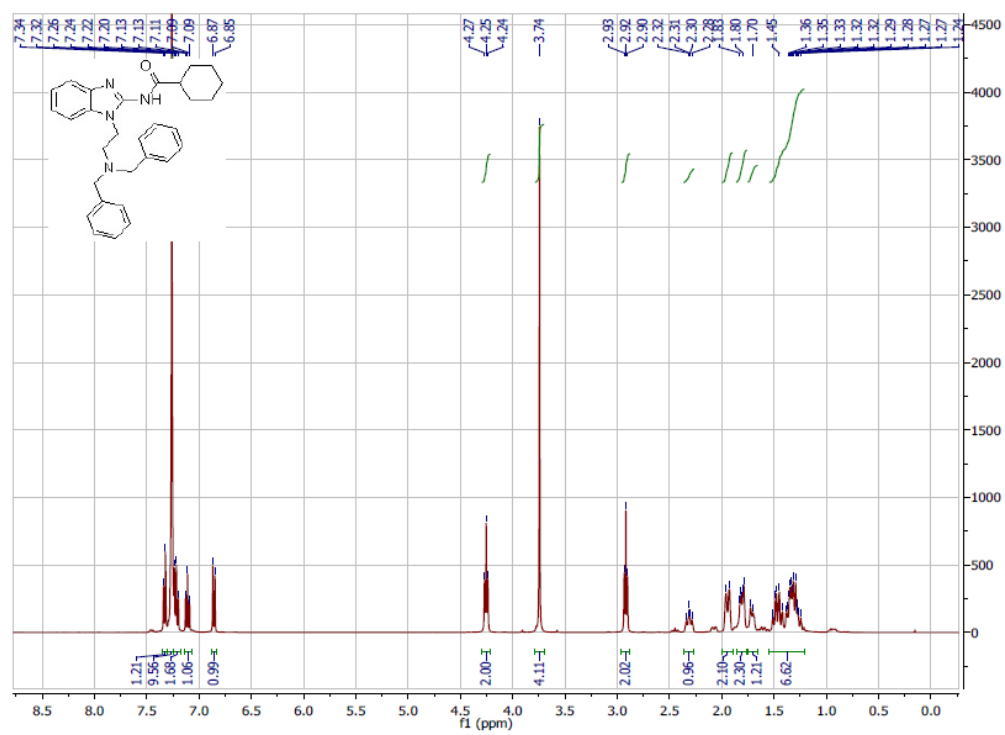
***N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)cyclohexanecarboxamide (35)**



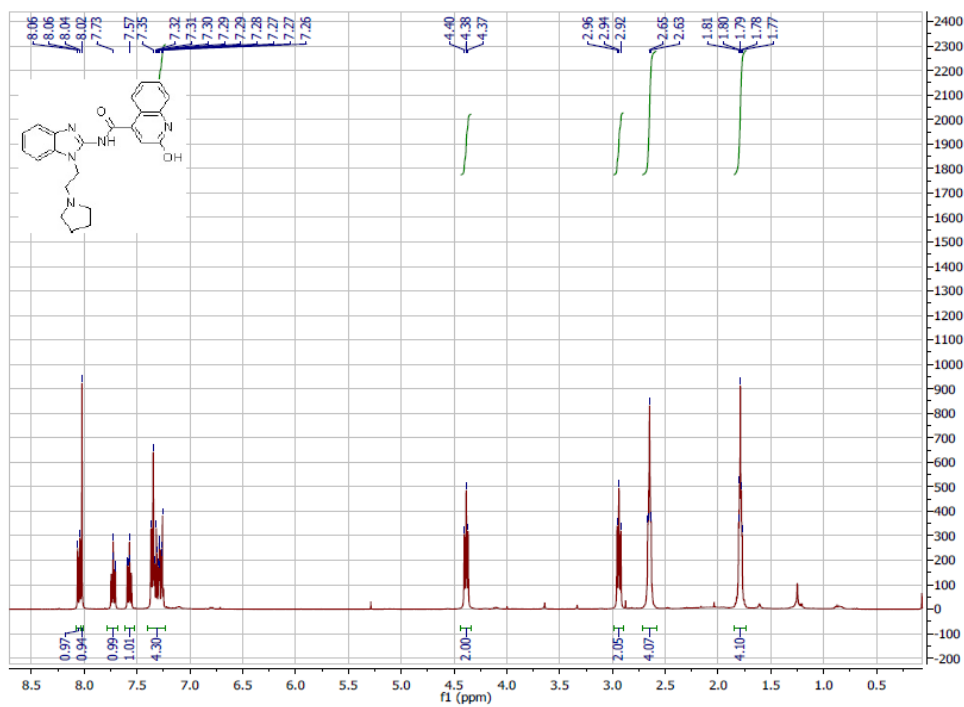
***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)cyclohexanecarboxamide (36)**



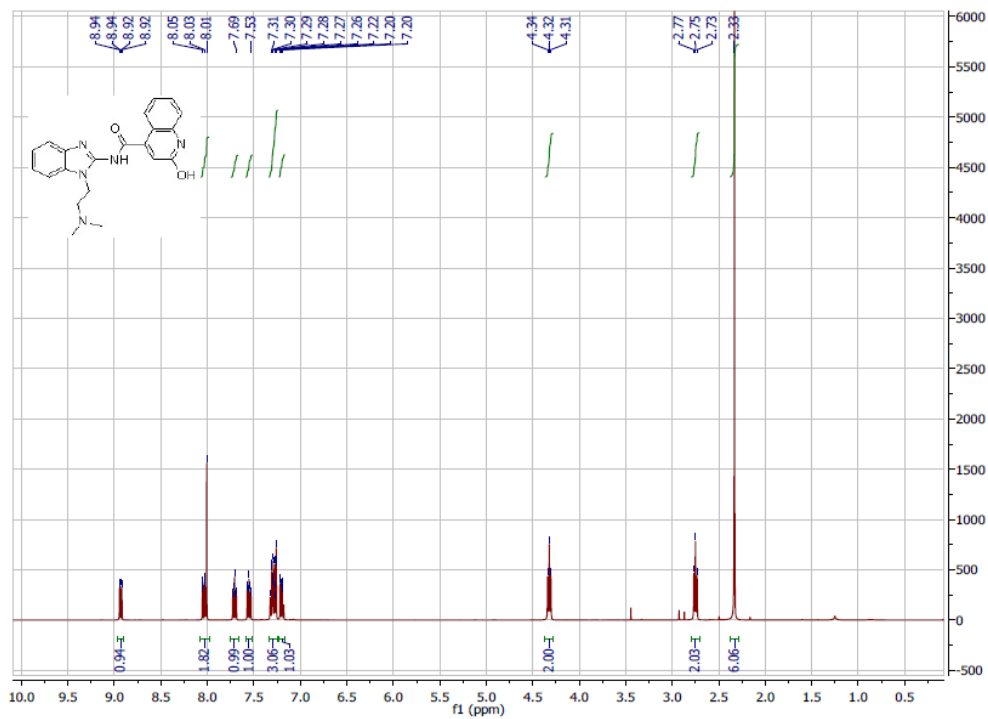
***N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)cyclohexanecarboxamide (37)**



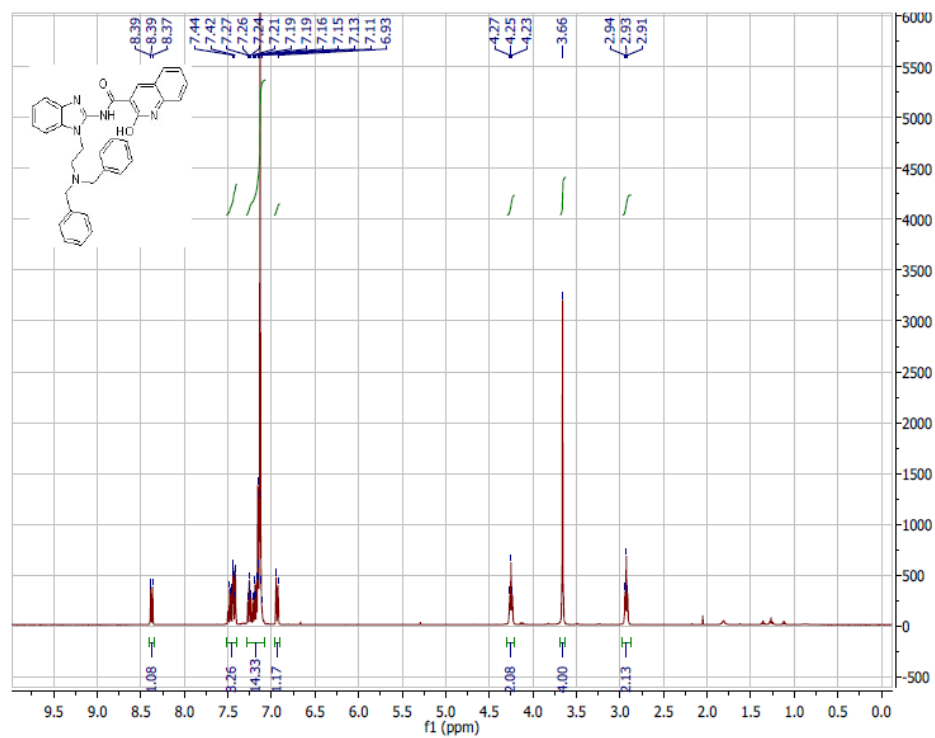
2-hydroxy-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)quinoline-4-carboxamide (38)



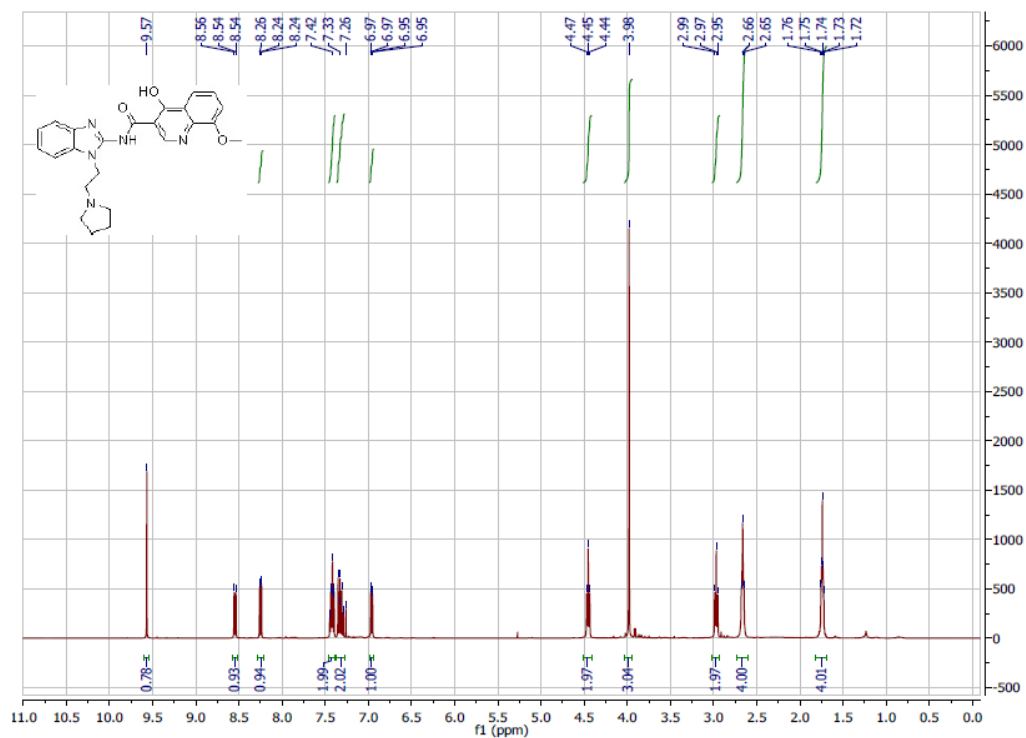
2-hydroxy-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)quinoline-4-carboxamide (39)



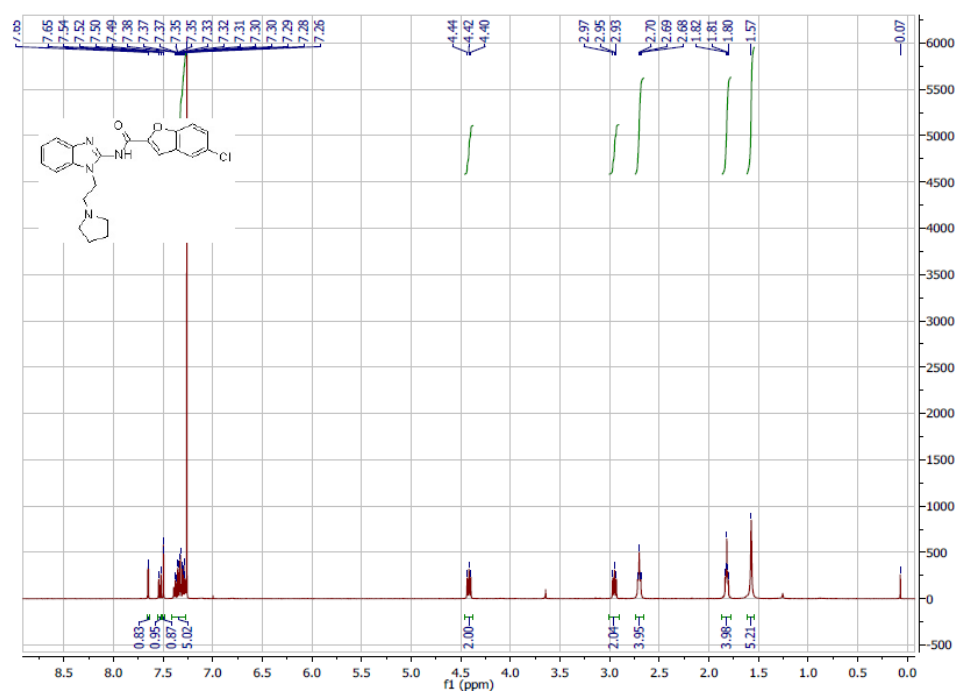
***N*-(1-(2-(dibenzylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-2-hydroxyquinoline-3-carboxamide (40)**



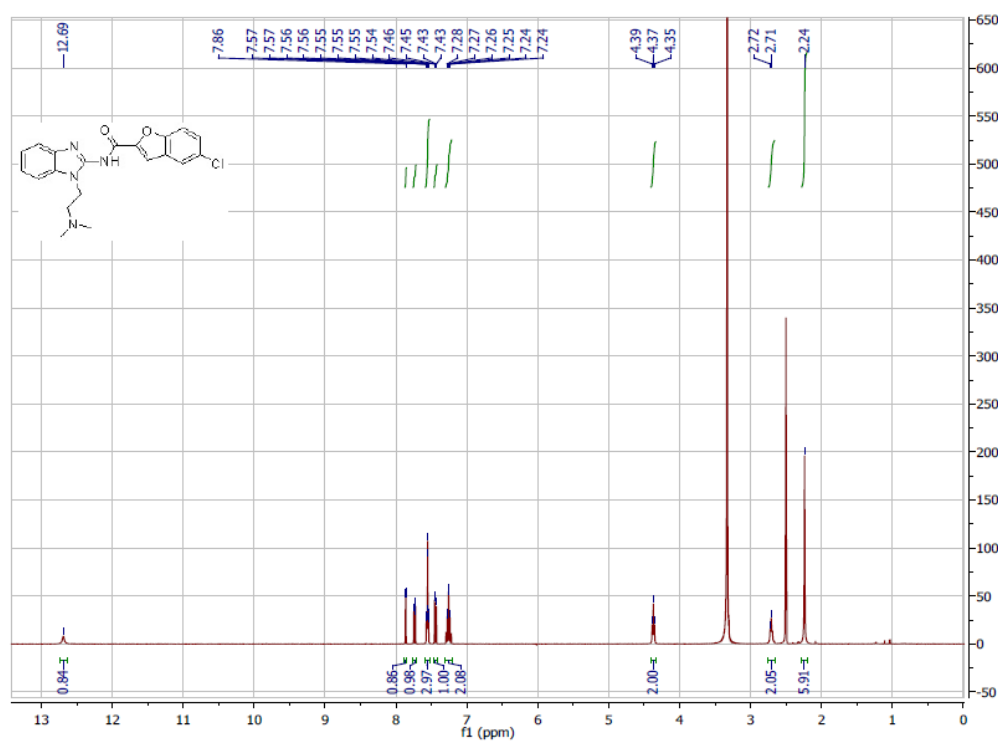
4-hydroxy-8-methoxy-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)quinoline-3-carboxamide (41)



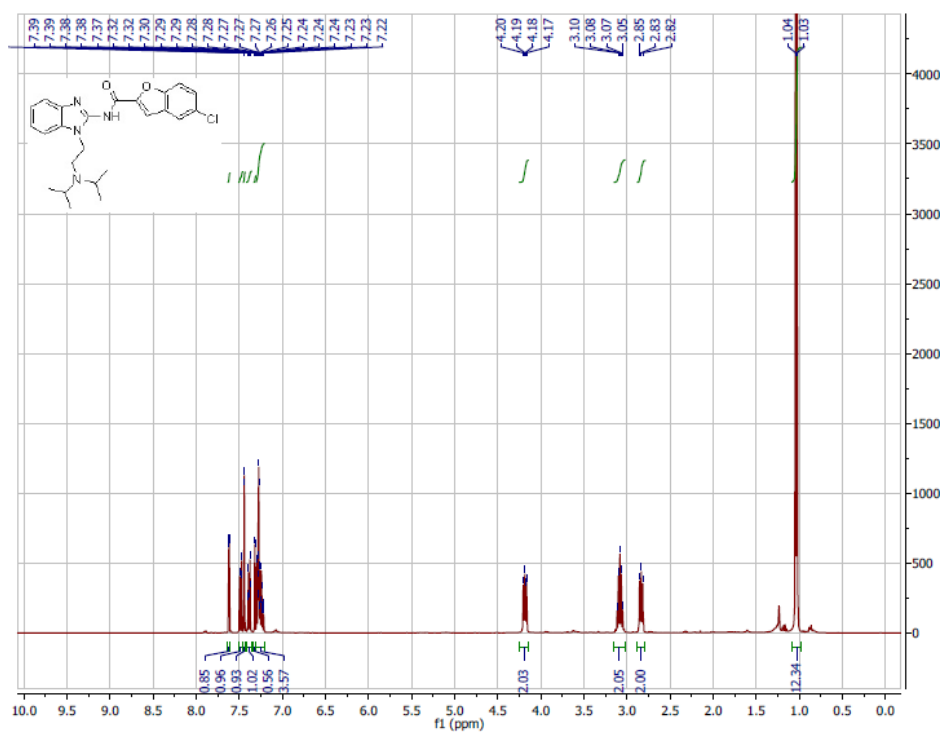
5-chloro-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)benzofuran-2-carboxamide (42)



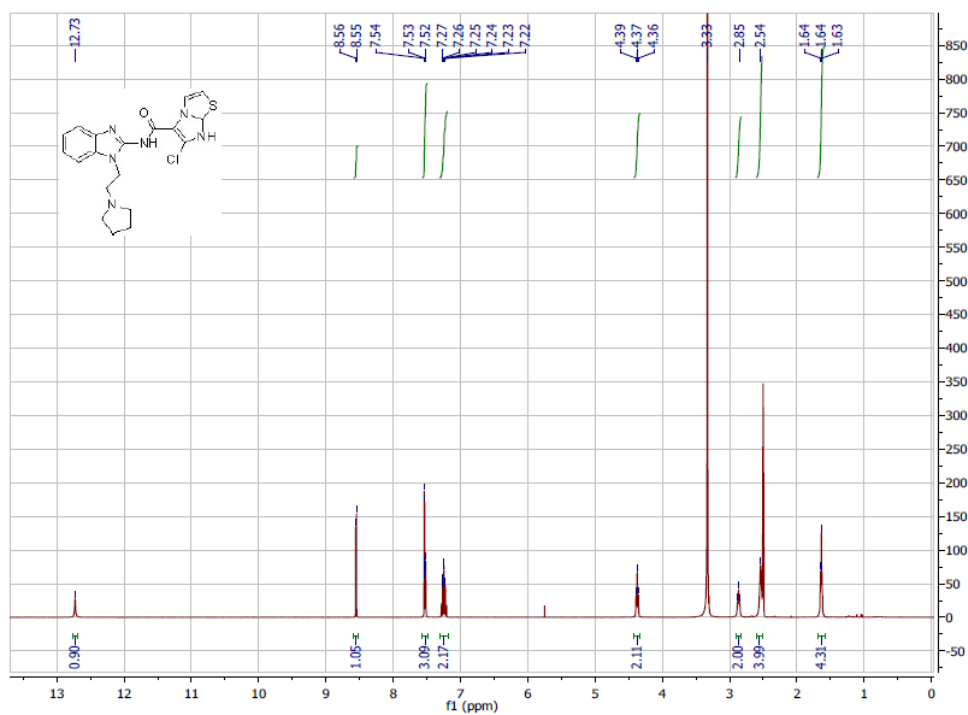
5-chloro-*N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzofuran-2-carboxamide (43)



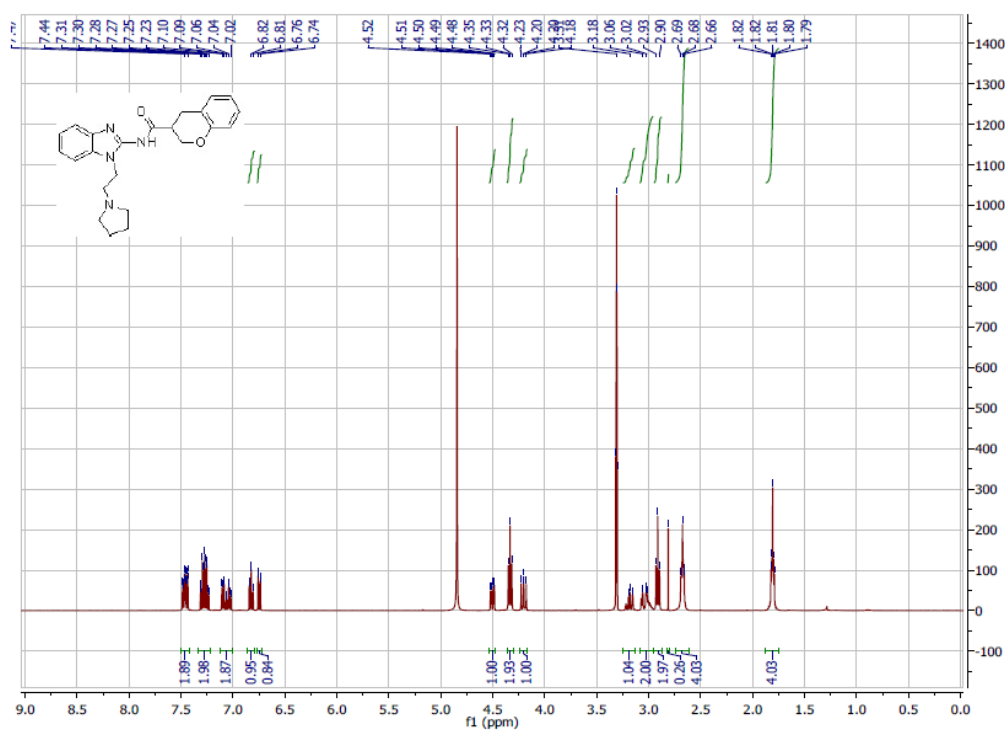
5-chloro-*N*-(1-(2-(diisopropylamino)ethyl)-1H-benzo[d]imidazol-2-yl)benzofuran-2-carboxamide (44)



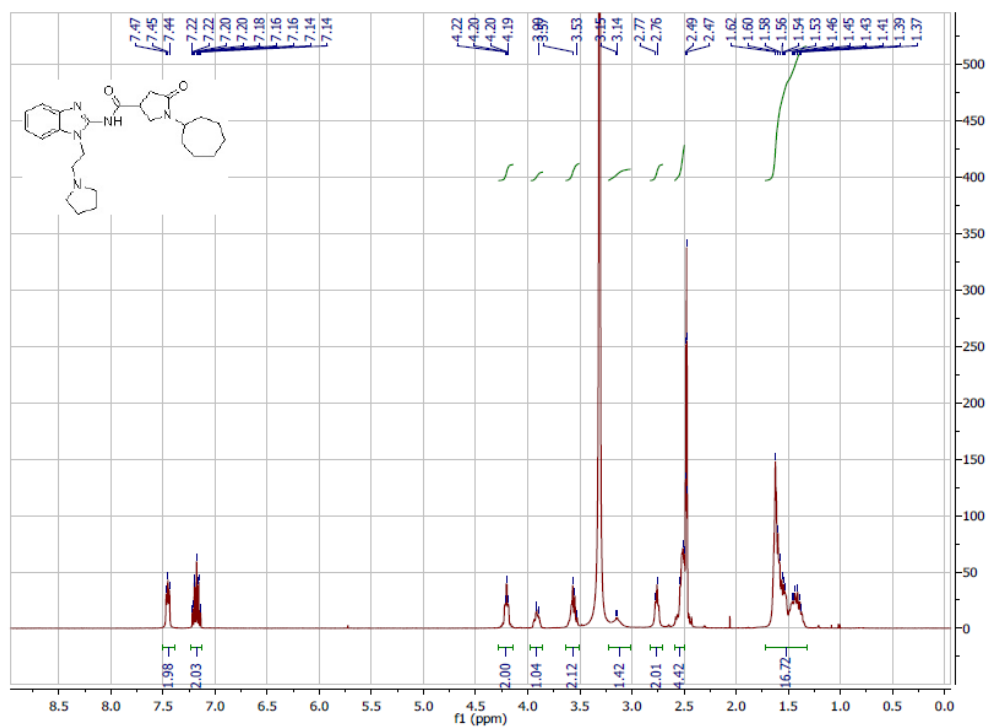
6-chloro-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)-7,7a-dihydroimidazo[2,1-b]thiazole-5-carboxamide (45)



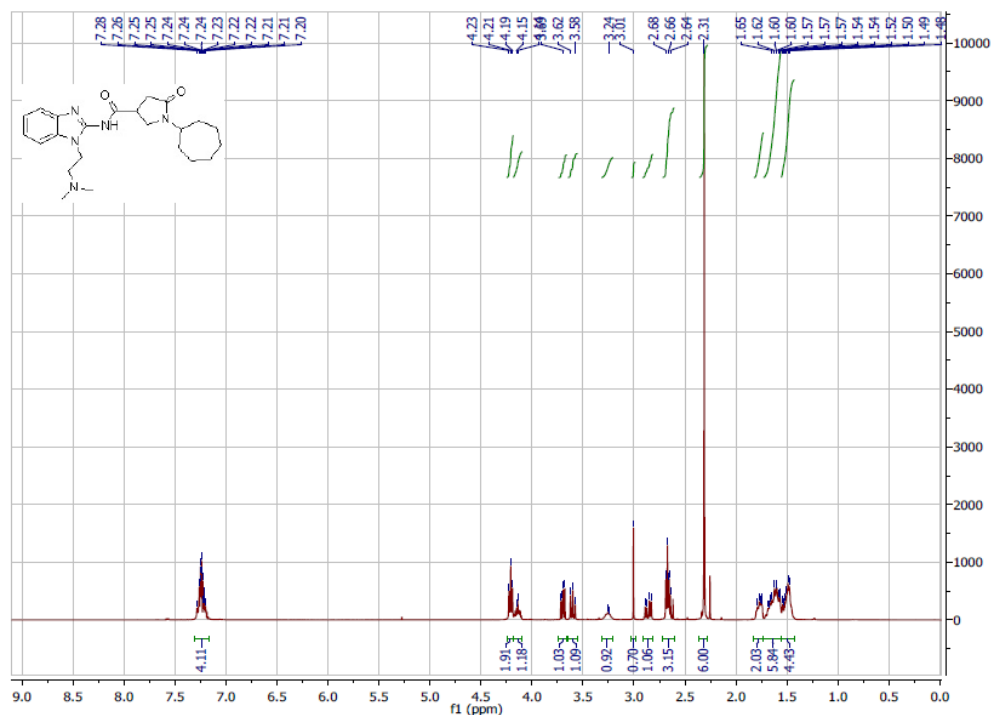
***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)chromane-3-carboxamide (46)**



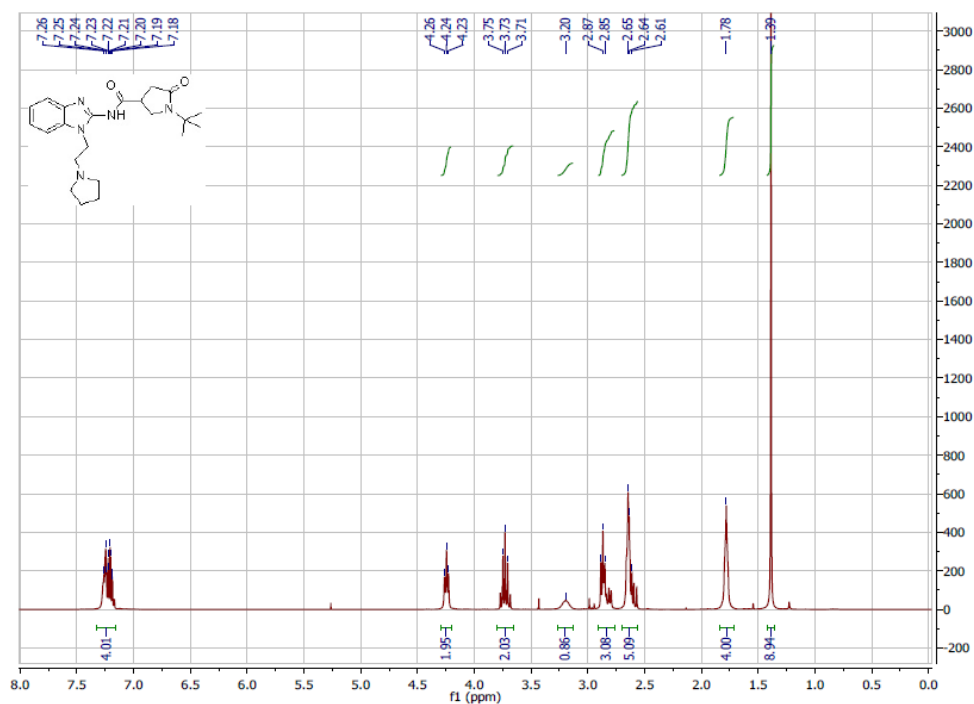
1-cycloheptyl-5-oxo-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)pyrrolidine-3-carboxamide (47)



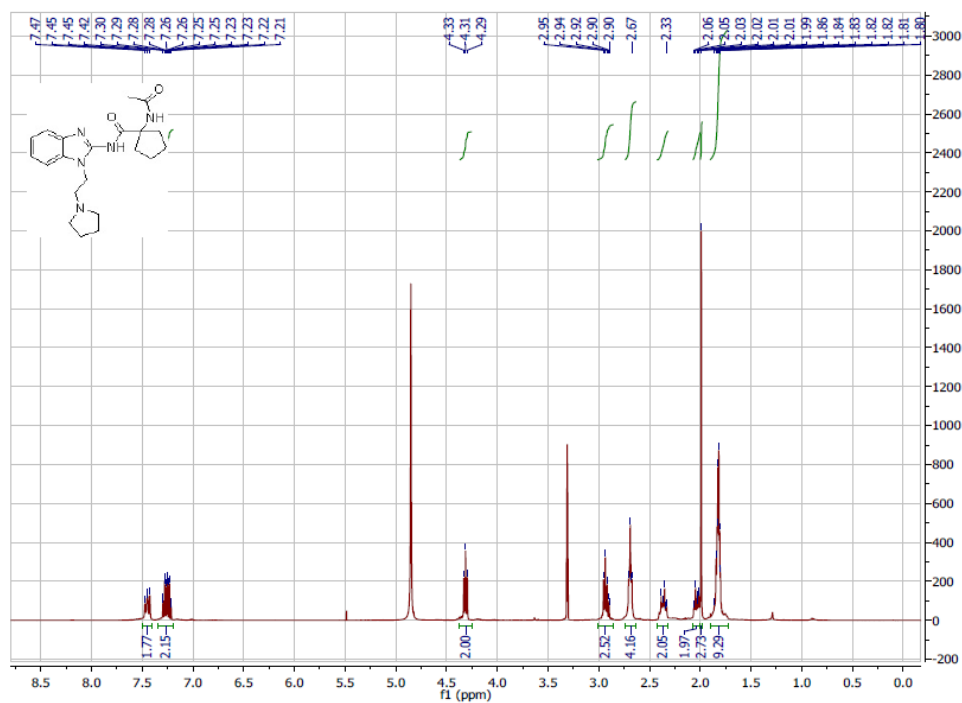
1-cycloheptyl-*N*-(1-(2-(dimethylamino)ethyl)-1H-benzo[d]imidazol-2-yl)-5-oxopyrrolidine-3-carboxamide (48)



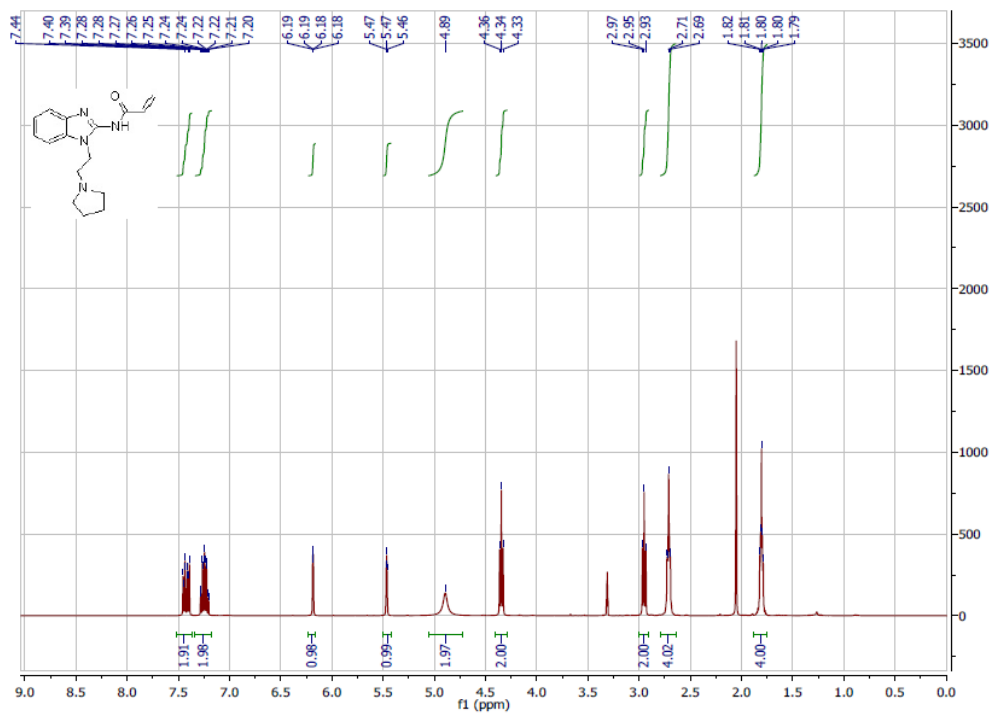
1-(tert-butyl)-5-oxo-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)pyrrolidine-3-carboxamide (49)



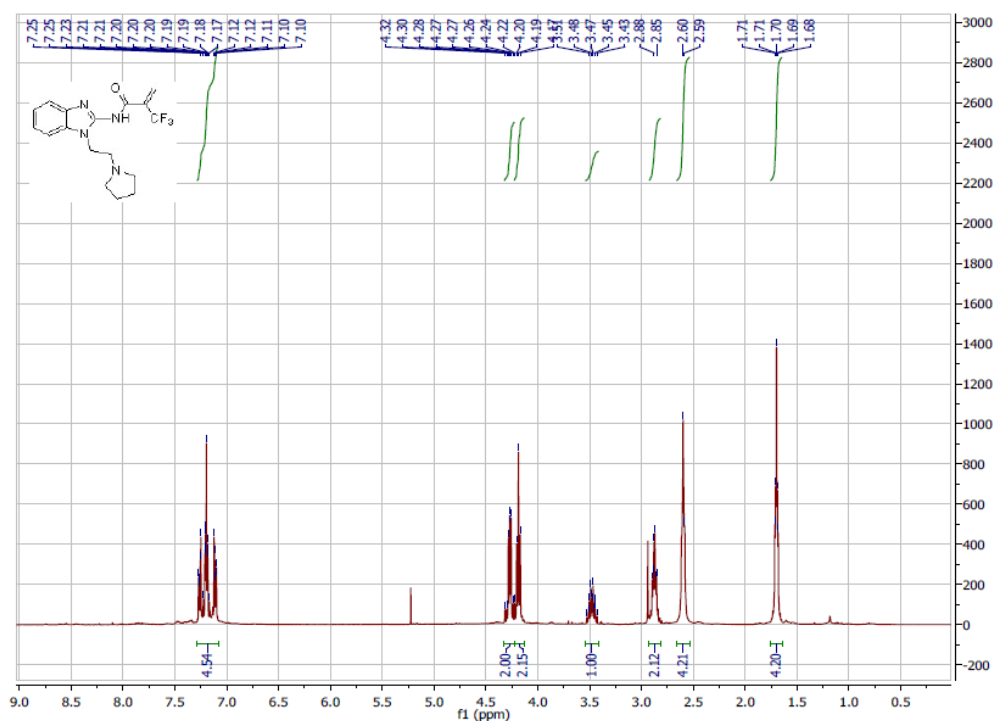
1-acetamido-*N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)cyclopentane-1-carboxamide (50)



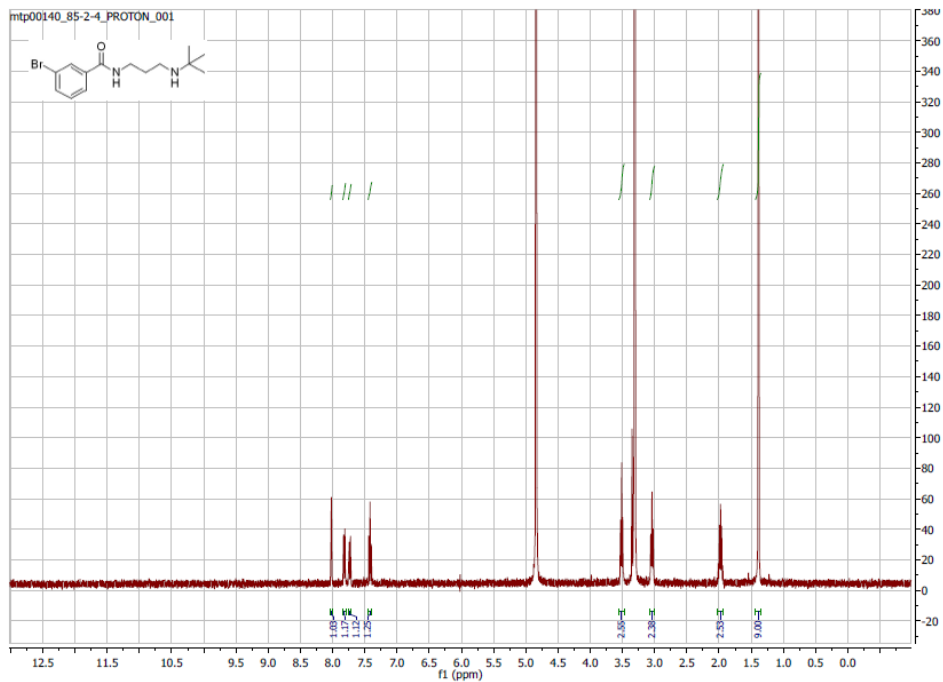
***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)acrylamide (51)**

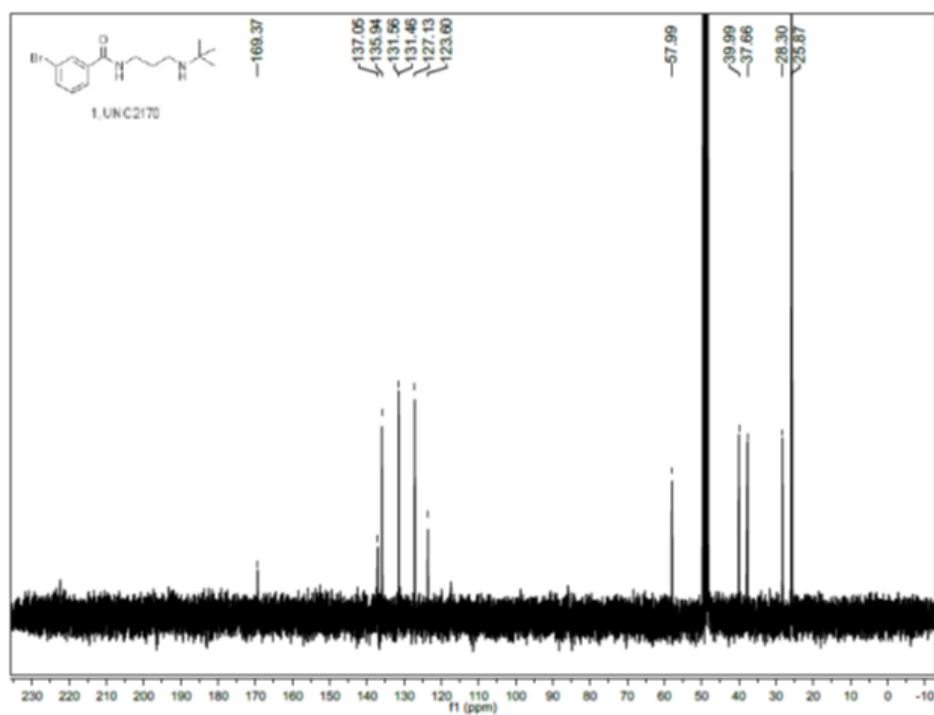


***N*-(1-(2-(pyrrolidin-1-yl)ethyl)-1H-benzo[d]imidazol-2-yl)-2-(trifluoromethyl)acrylamide (52)**

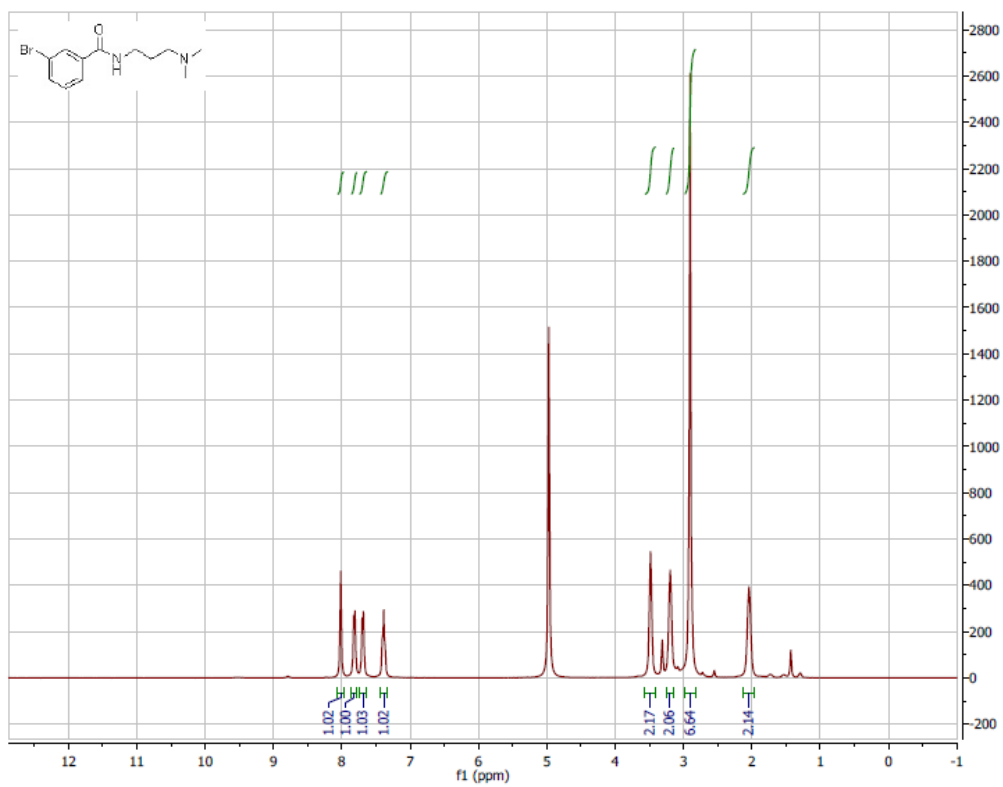


3-Bromo-*N*-(3-(*tert*-butylamino)propyl)benzamide (61, UNC2170)

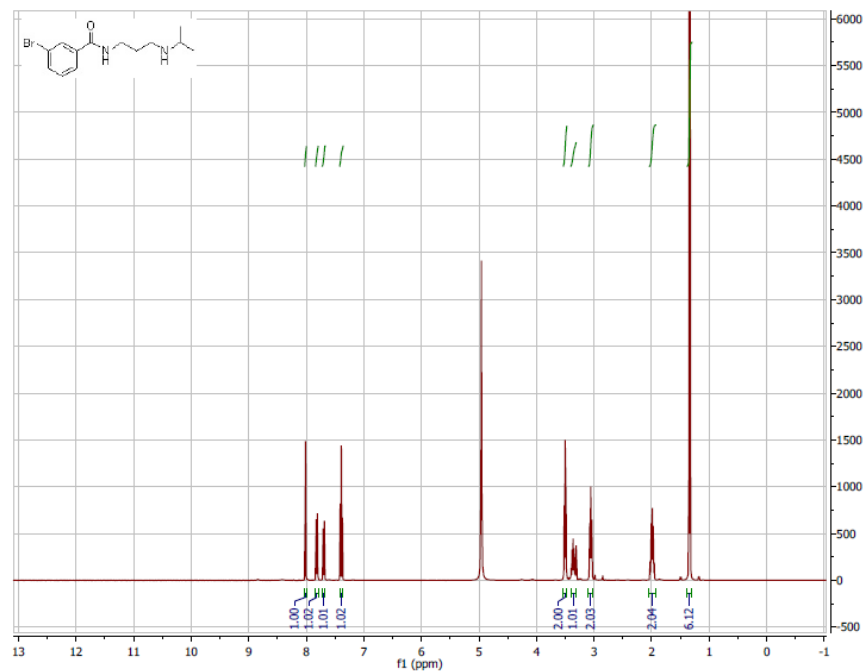




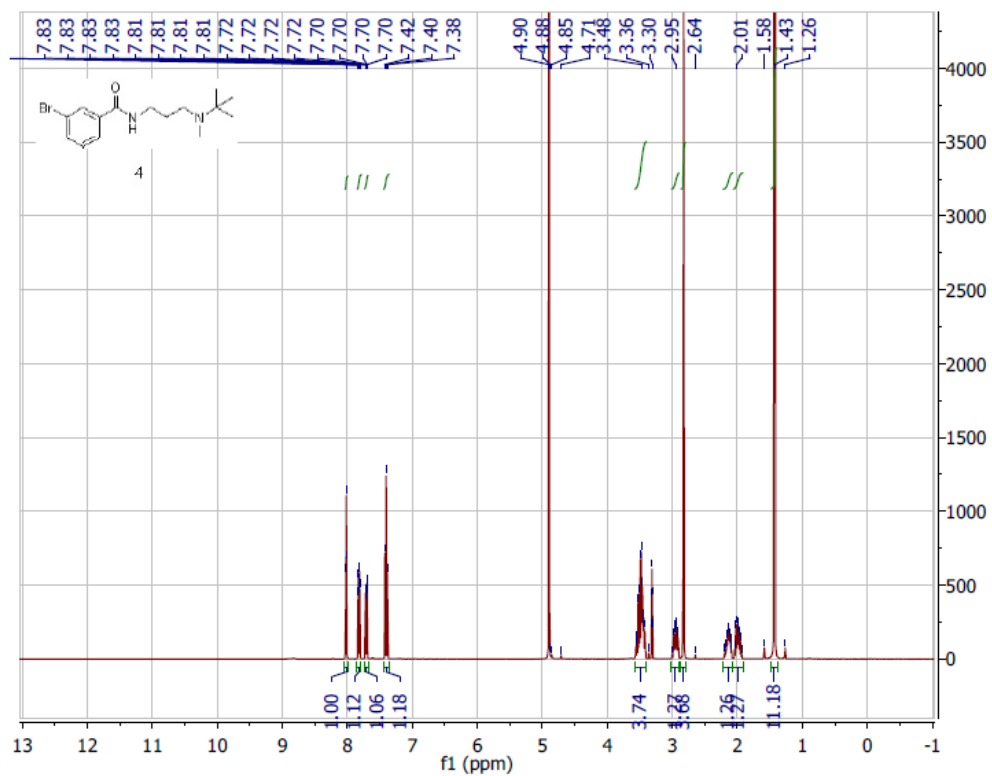
3-Bromo-N-(3-(dimethylamino)propyl)benzamide (62)

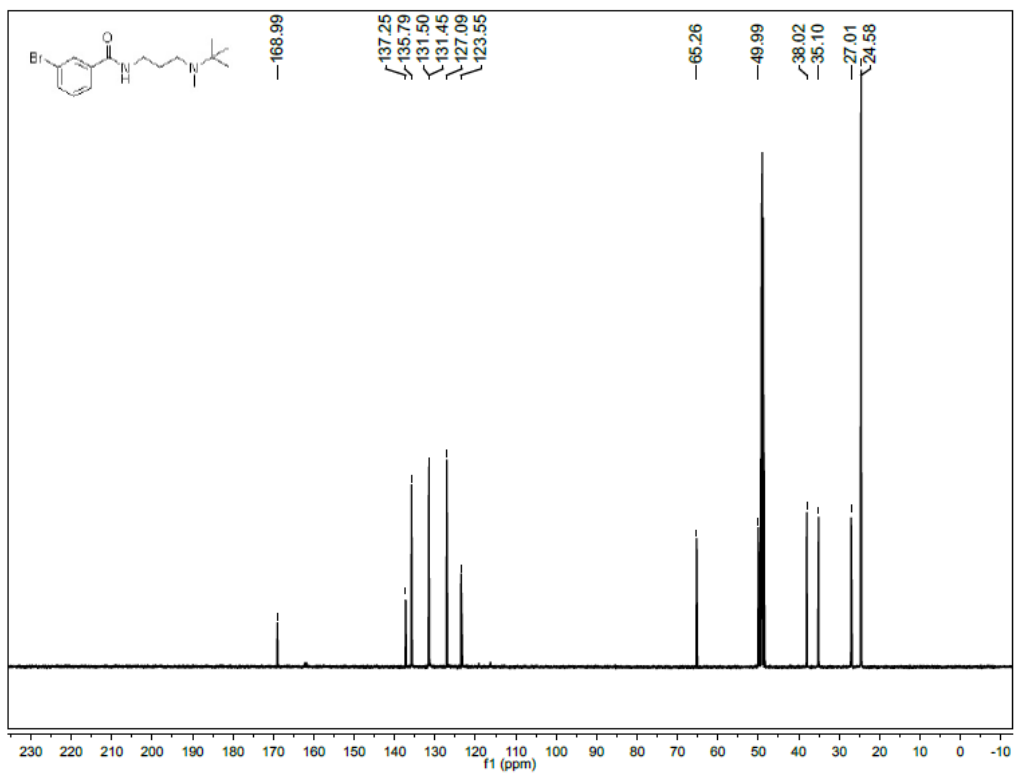


3-Bromo-N-(3-(isopropylamino)propyl)benzamide (63)

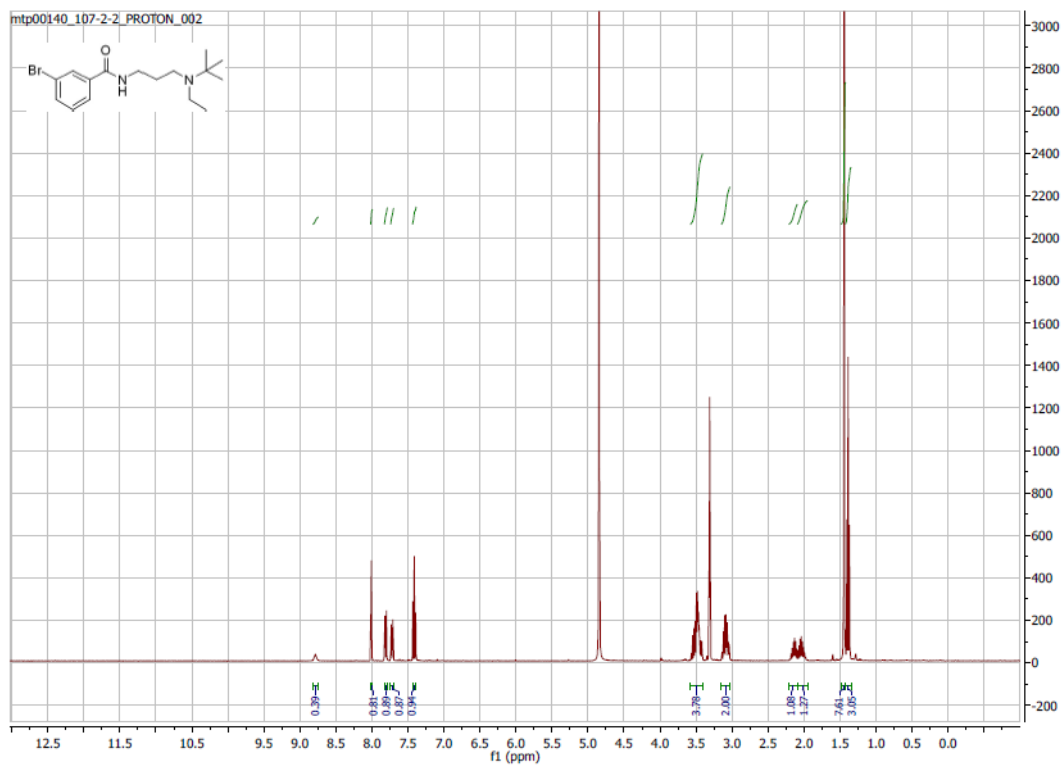


3-Bromo-N-(3-(*tert*-butyl(methyl)amino)propyl)benzamide (64, UNC2892)

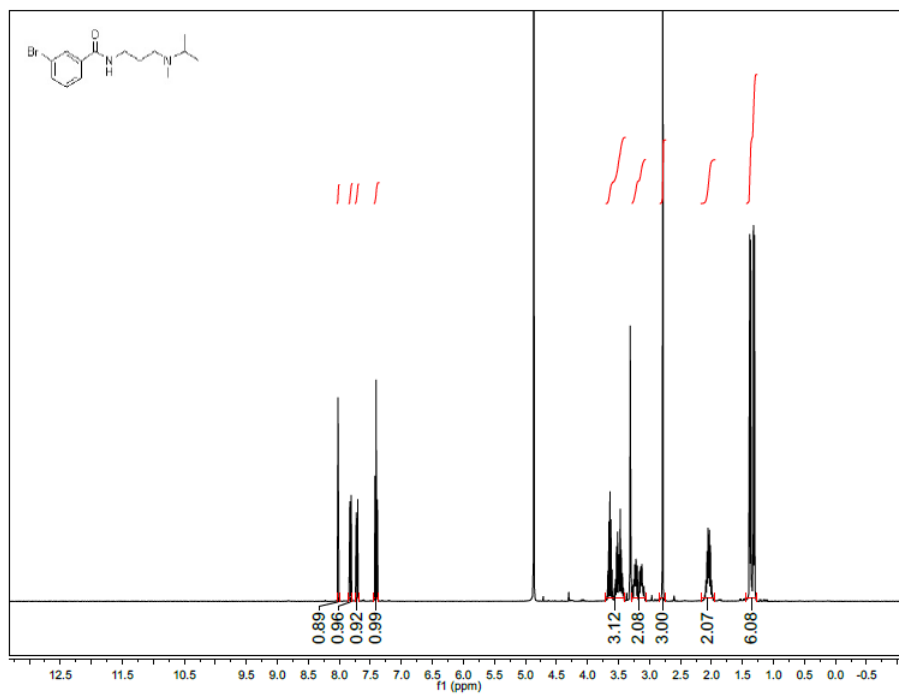




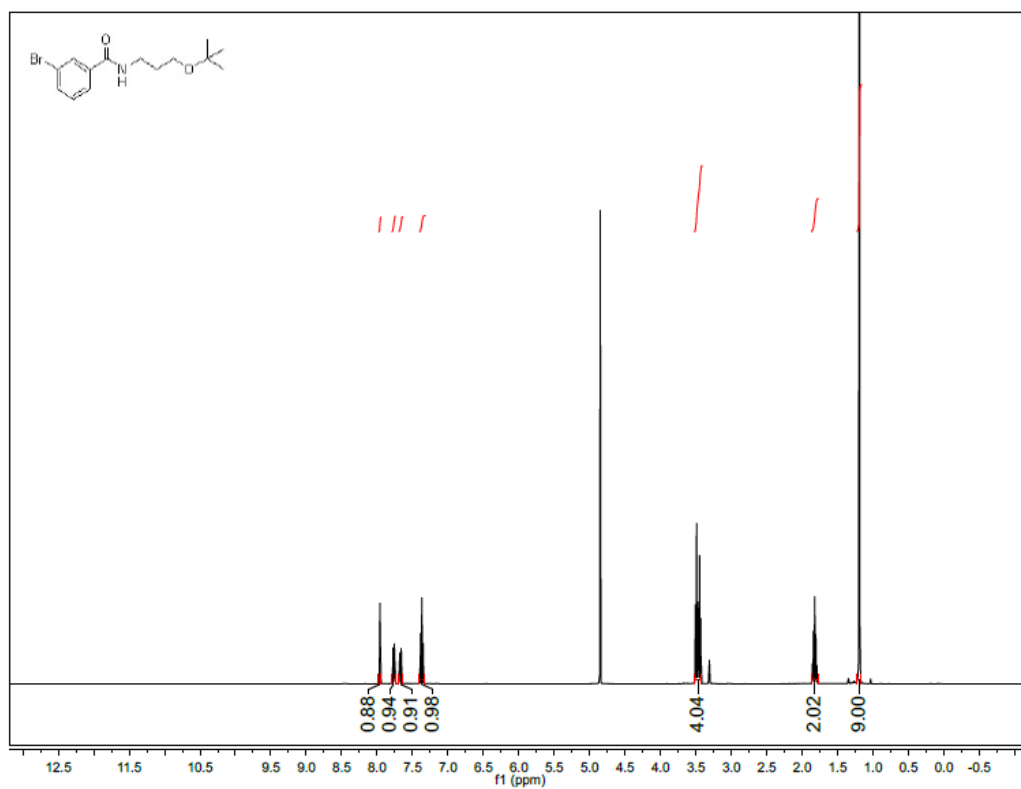
3-bromo-N-(3-(tert-butyl(ethyl)amino)propyl)benzamide (65)



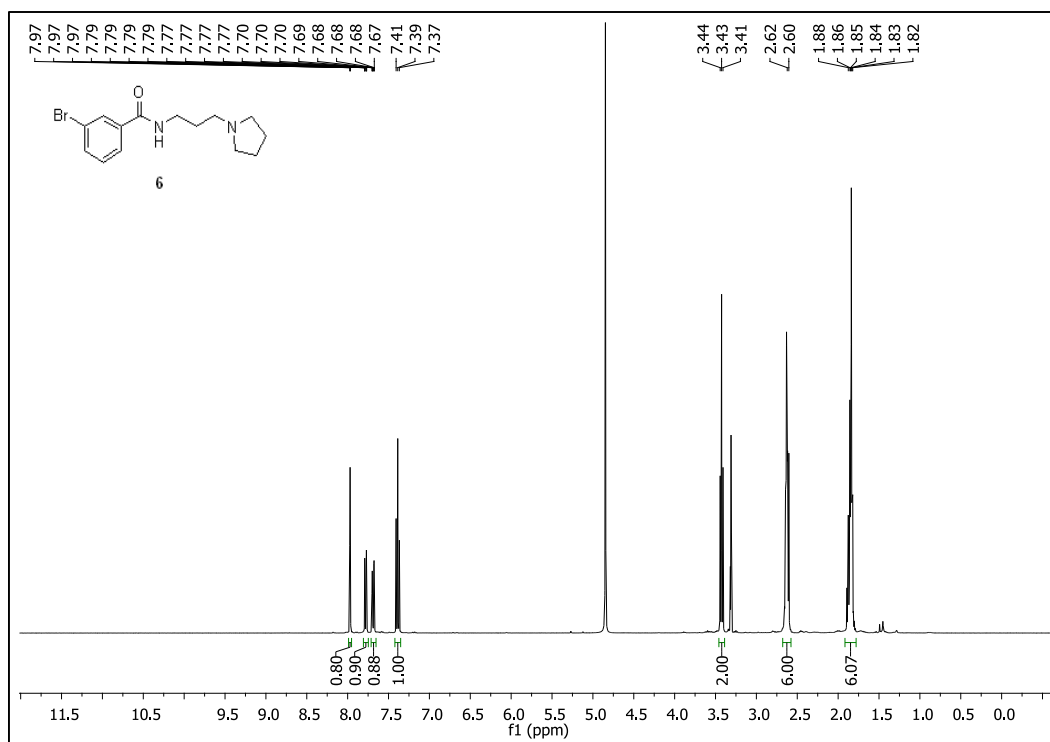
3-bromo-*N*-(3-(isopropyl(methyl)amino)propyl)benzamide (66)



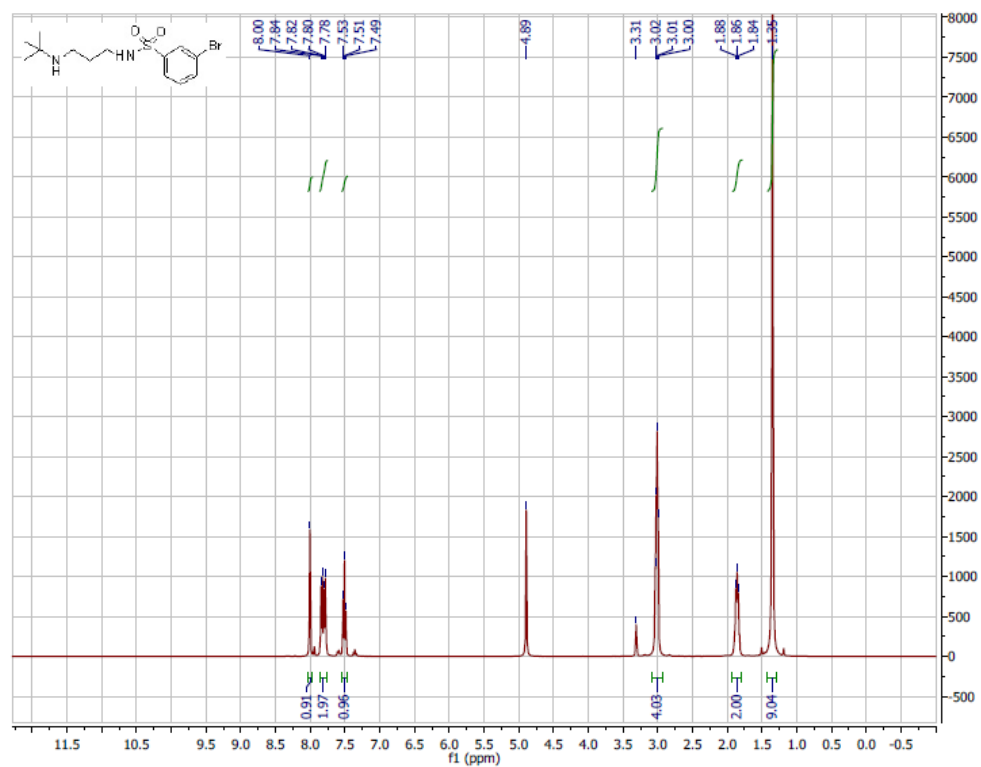
3-Bromo-*N*-(3-(*tert*-butoxy)propyl)benzamide (67)



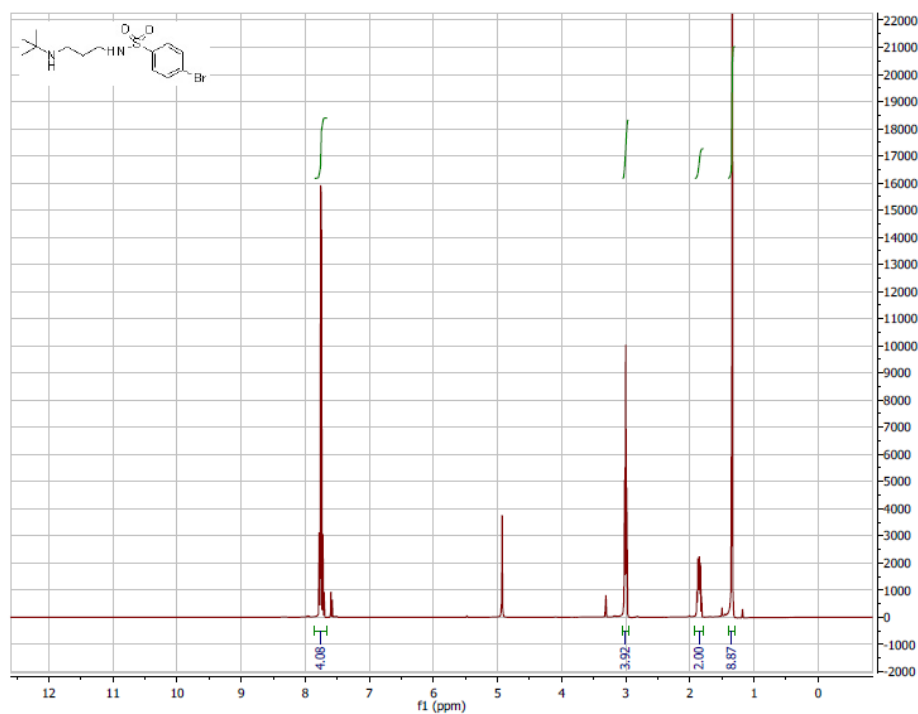
3-Bromo-N-(3-(pyrrolidin-1-yl)propyl)benzamide (68)



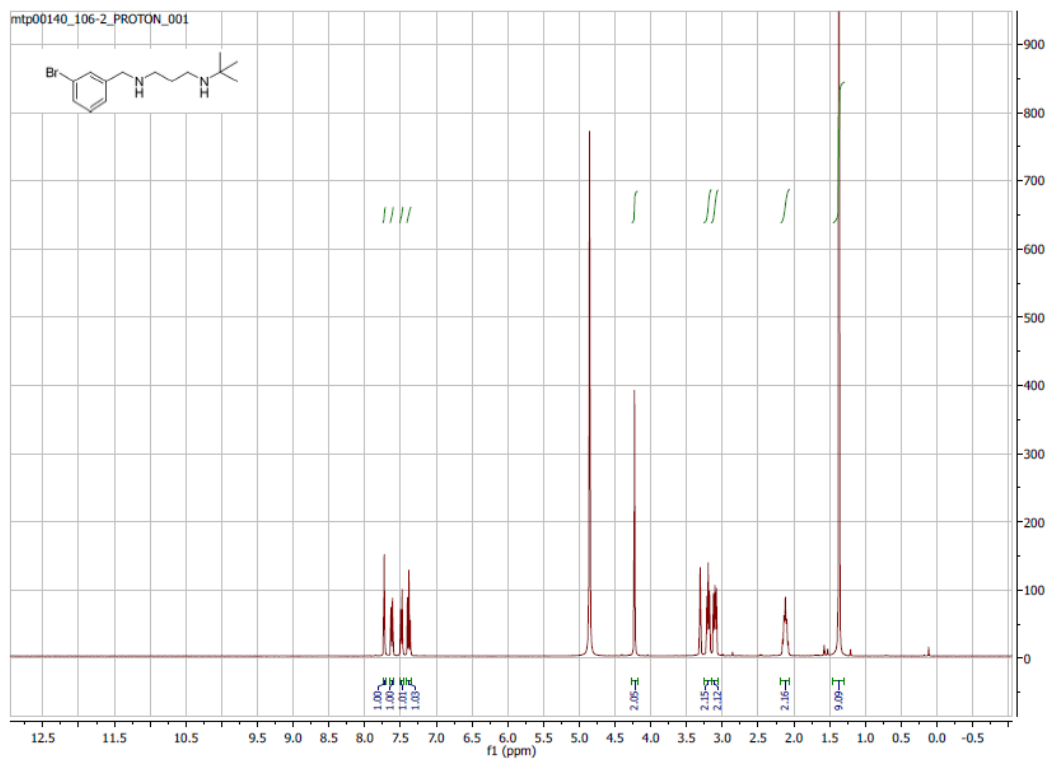
3-Bromo-N-(3-(*tert*-butylamino)propyl)benzenesulfonamide (69)



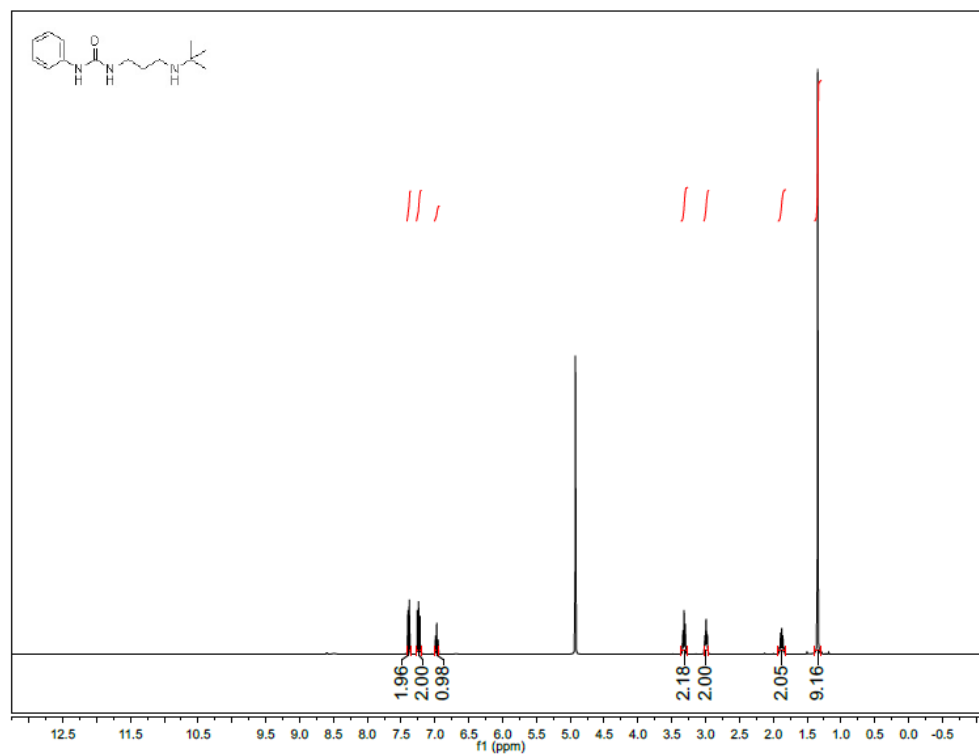
4-bromo-*N*-(3-(*tert*-butylamino)propyl)benzenesulfonamide (70)



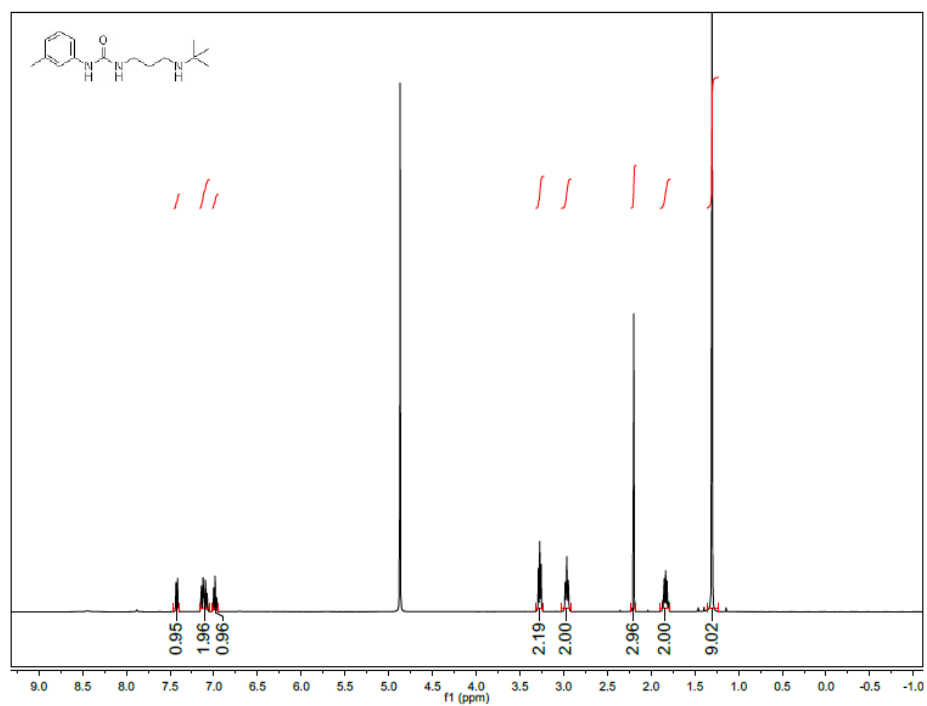
***N*¹-(3-Bromobenzyl)-*N*³-(*tert*-butyl)propane-1,3-diamine (71)**



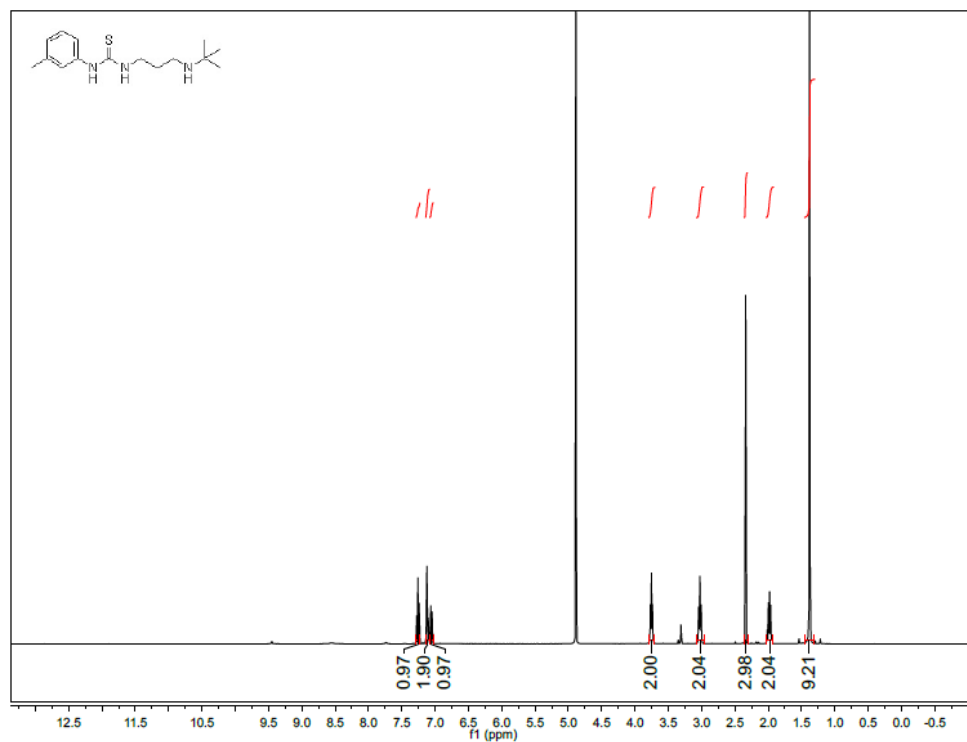
1-(3-(tert-butylamino)propyl)-3-phenylurea (72)



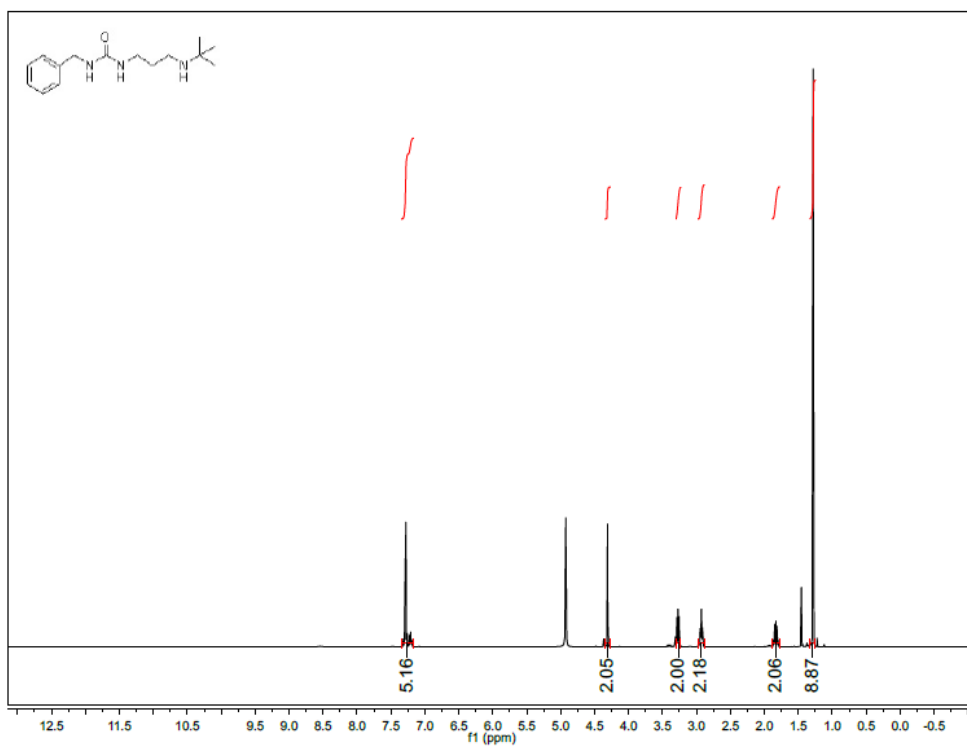
1-(3-(tert-butylamino)propyl)-3-(o-tolyl)urea (73)



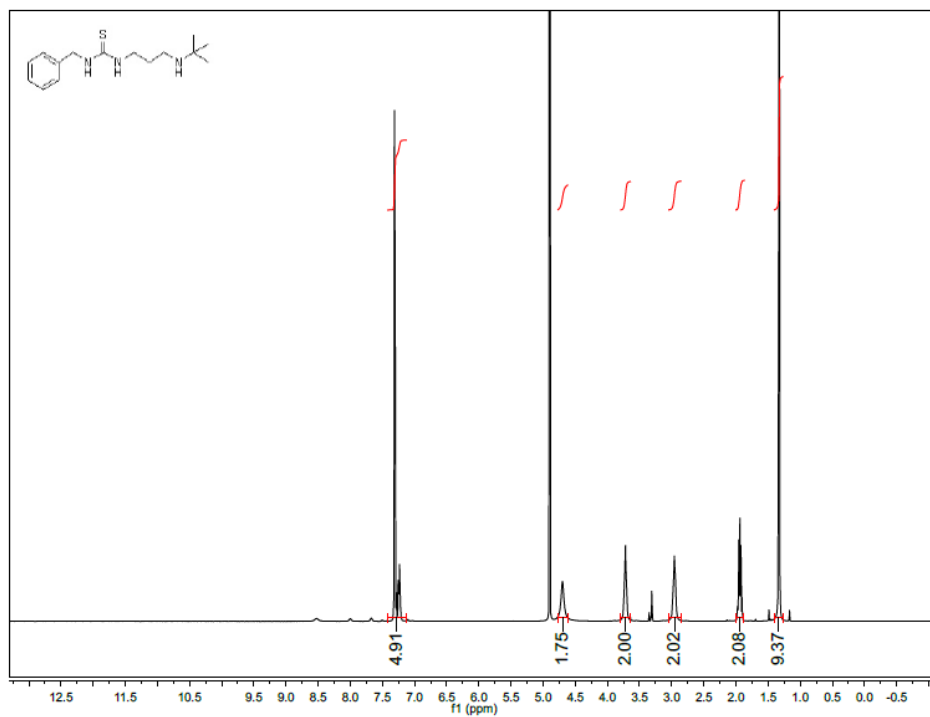
1-(3-(tert-butylamino)propyl)-3-(o-tolyl)thiourea (74)



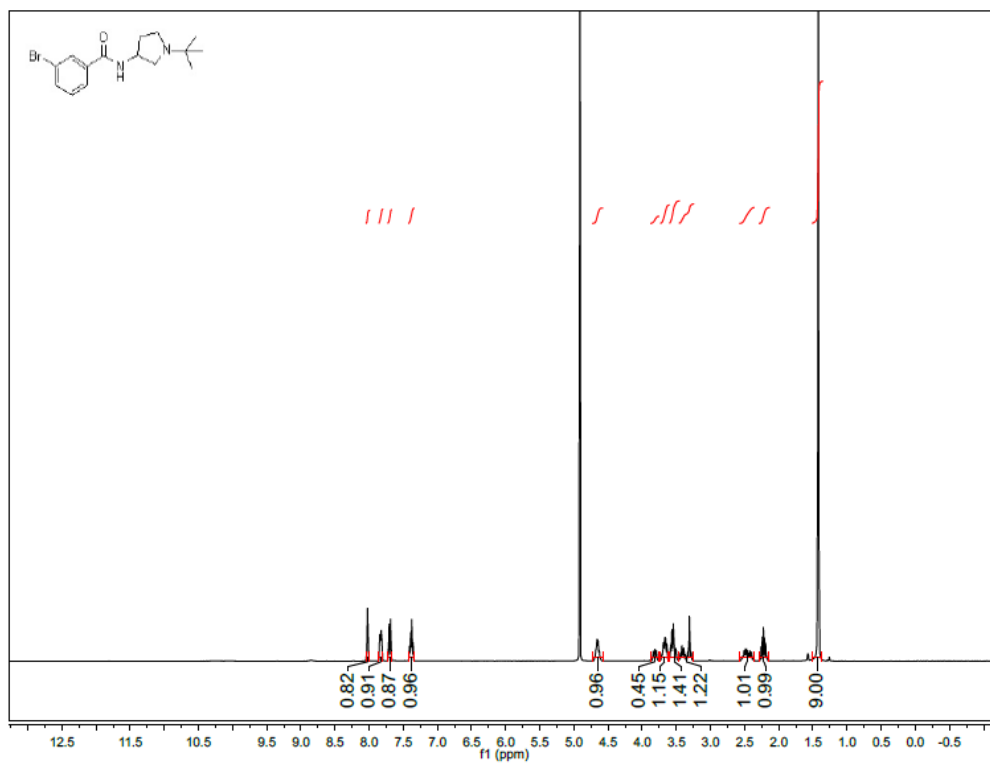
1-benzyl-3-(3-(tert-butylamino)propyl)urea (75)



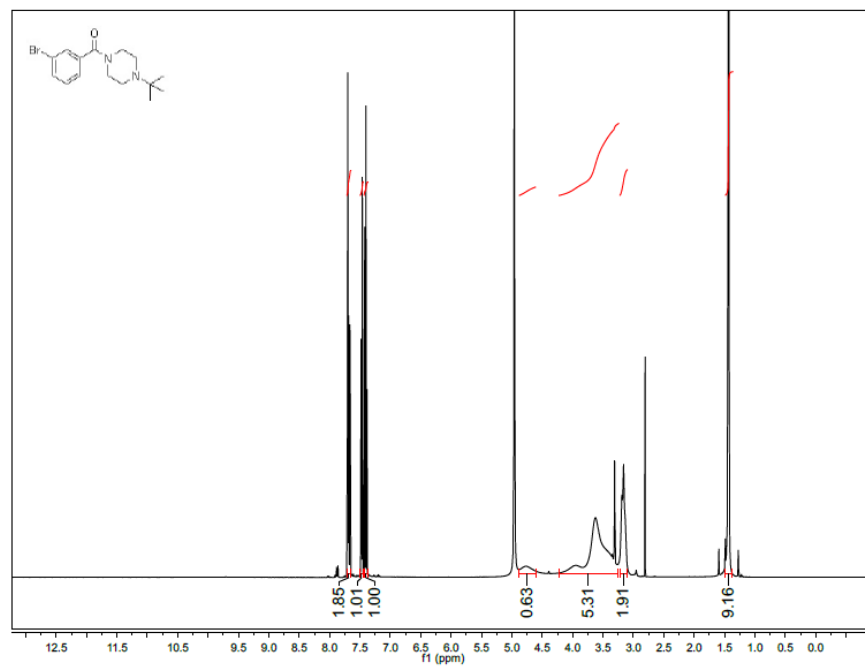
1-benzyl-3-(3-(tert-butylamino)propyl)thiourea (76)



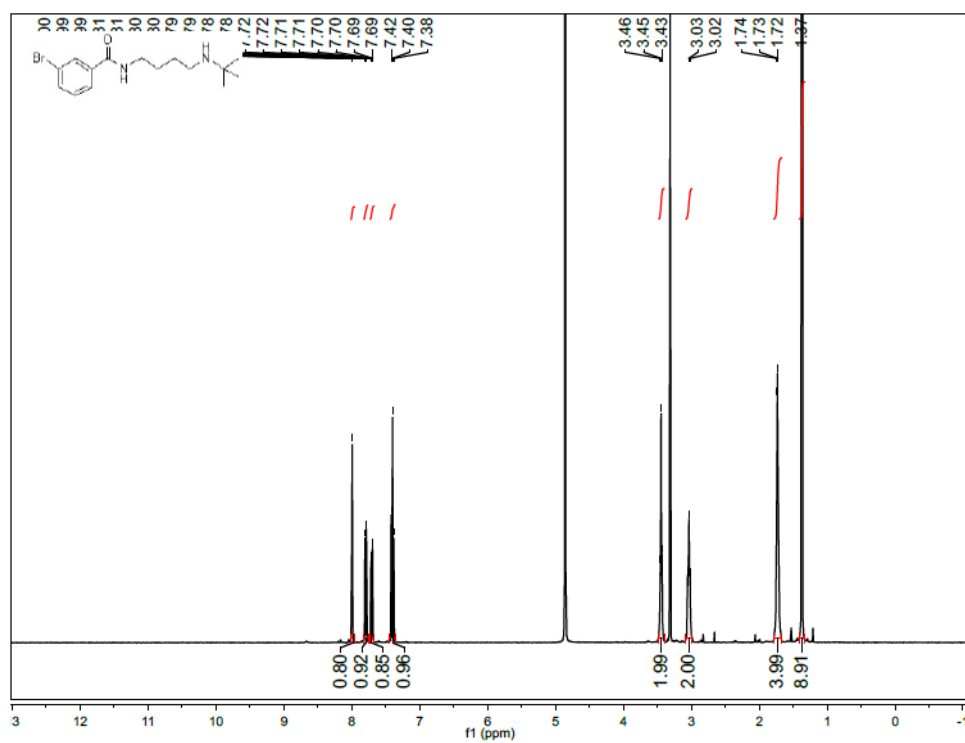
(3-Bromophenyl)(3-(tert-butyl)imidazolidin-1-yl)methanone (77)



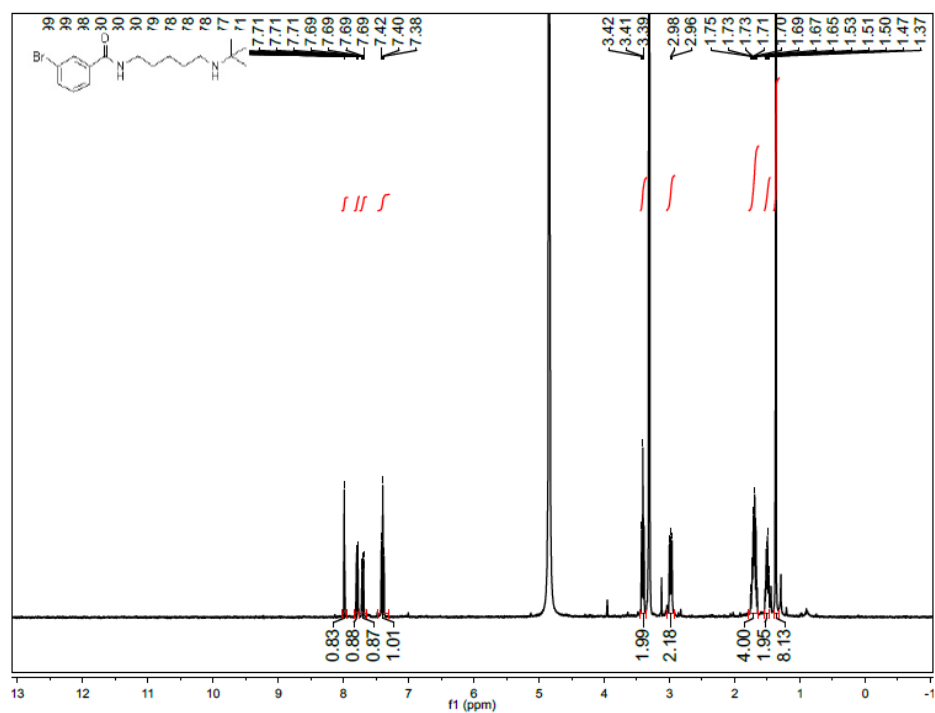
(3-Bromophenyl)(4-(*tert*-butyl)piperazin-1-yl)methanone (78)



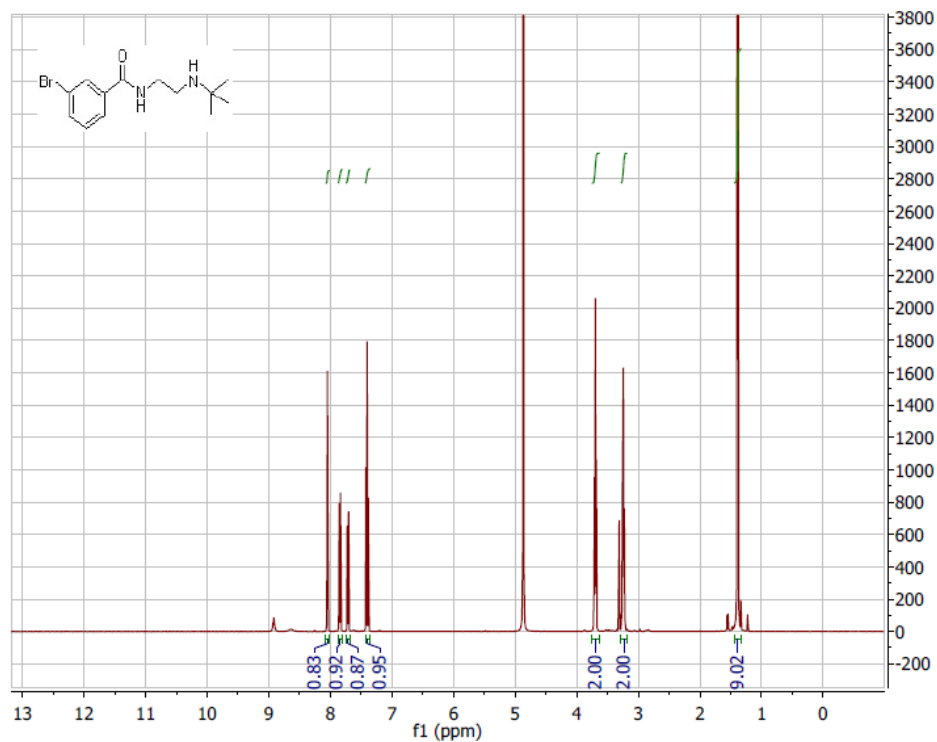
3-Bromo-N-(4-(*tert*-butylamino)butyl)benzamide (79)



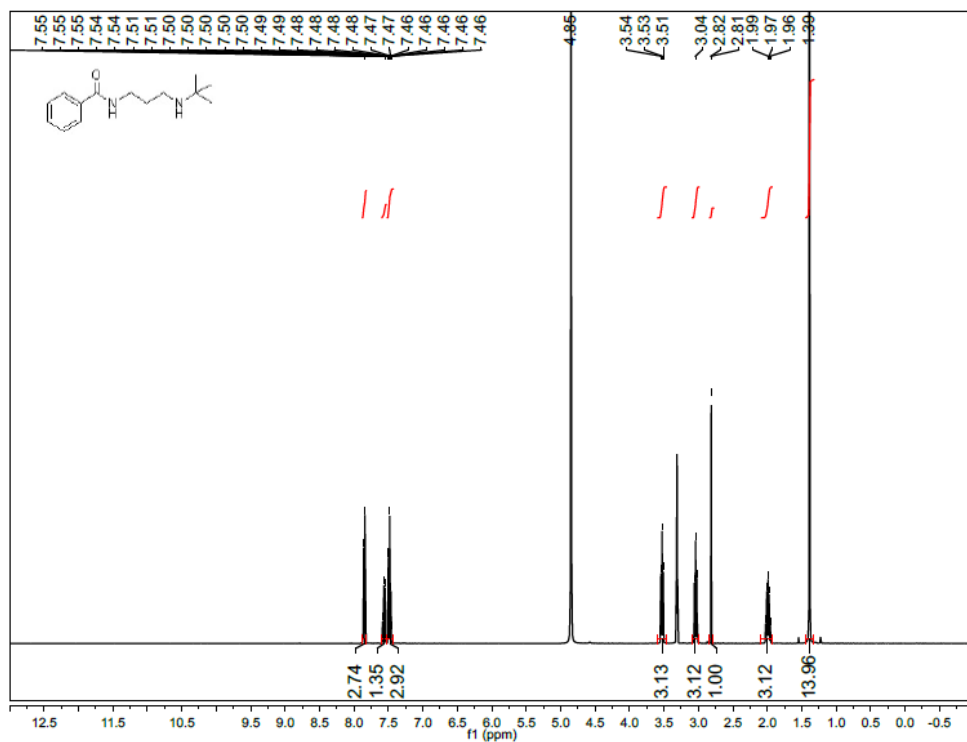
3-Bromo-N-(5-(*tert*-butylamino)pentyl)benzamide (80)



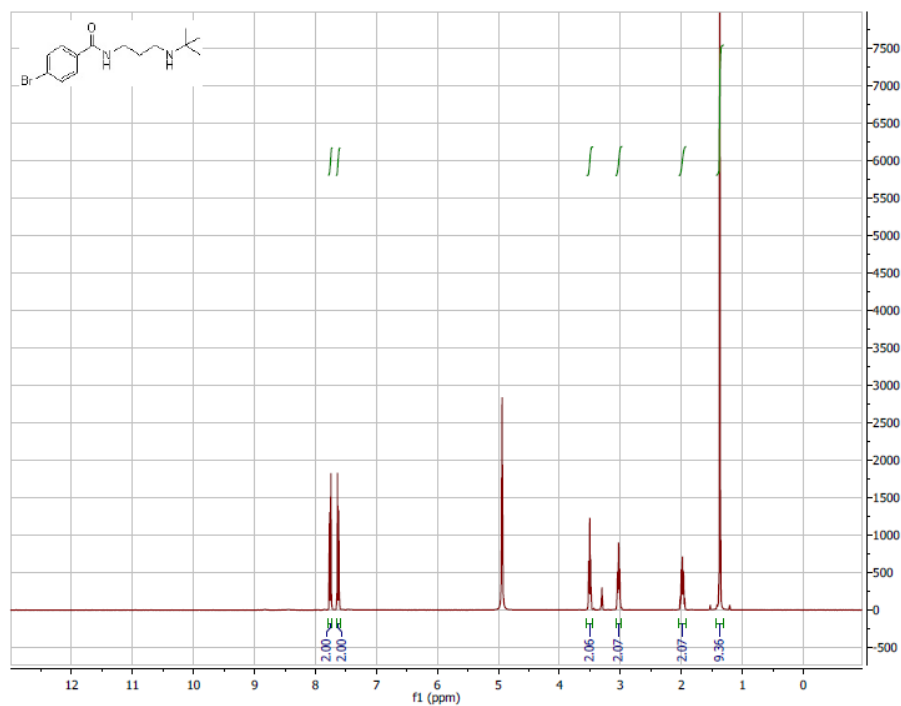
3-Bromo-N-(2-(*tert*-butylamino)ethyl)benzamide (81)



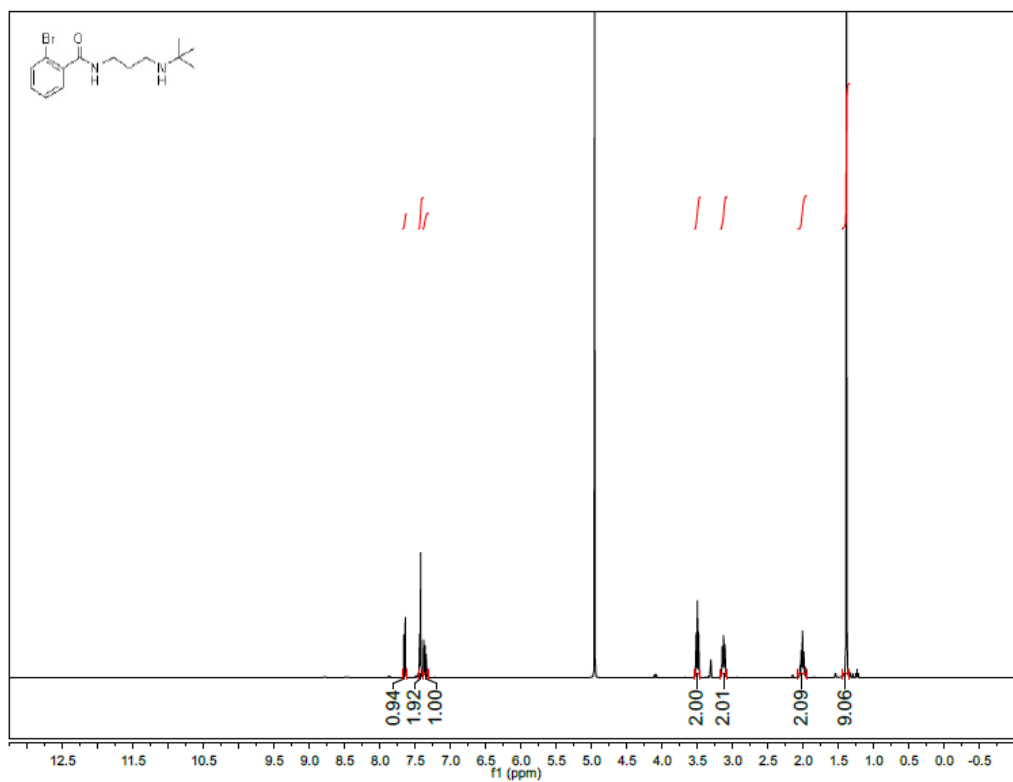
***N*-(3-(*Tert*-butylamino)propyl)benzamide (82)**



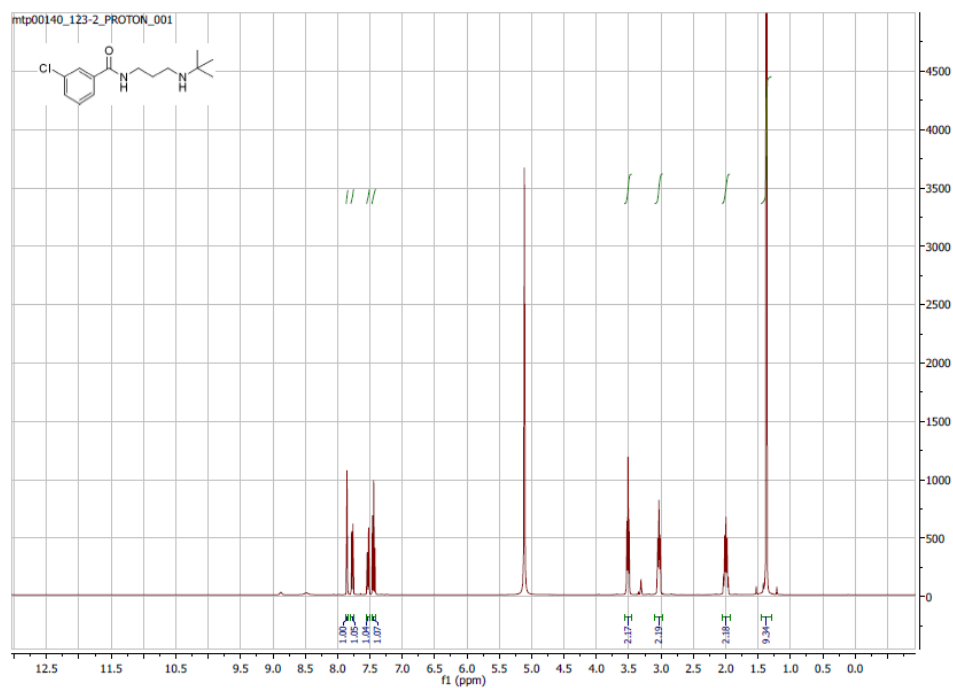
4-Bromo-*N*-(3-(*tert*-butylamino)propyl)benzamide (83)



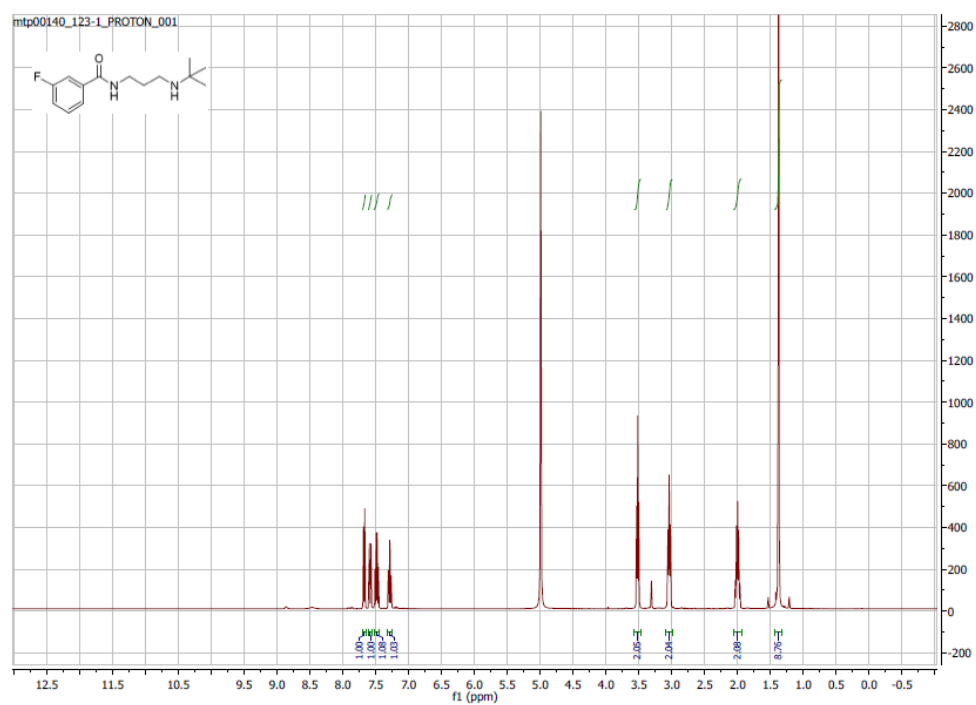
2-Bromo-N-(3-(*tert*-butylamino)propyl)benzamide (84)



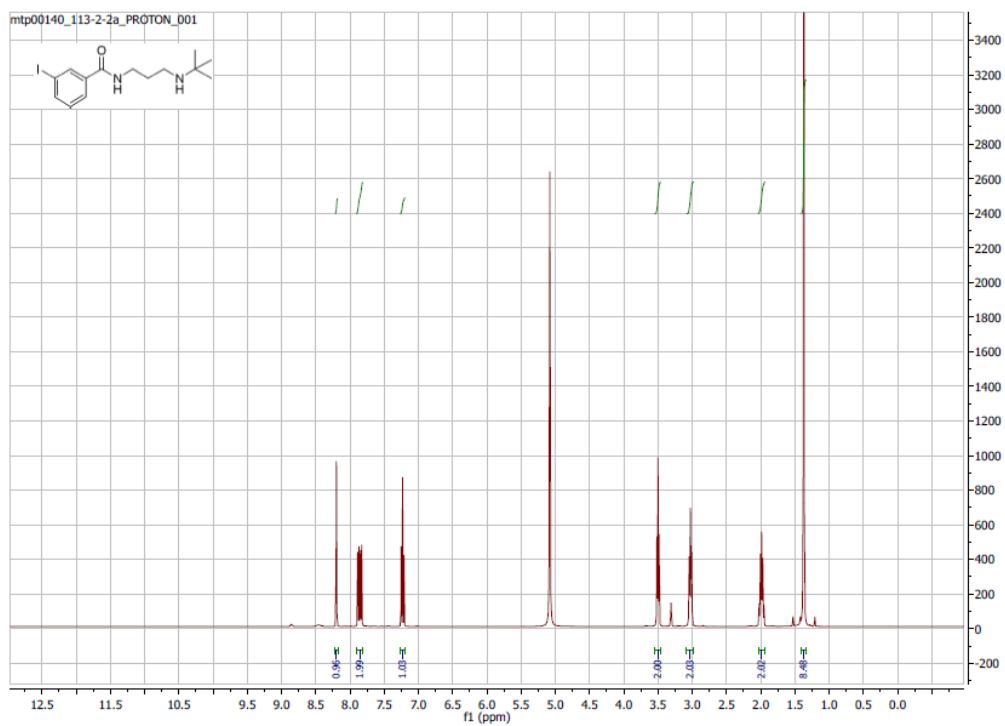
N-(3-(*Tert*-butylamino)propyl)-3-chlorobenzamide (85)



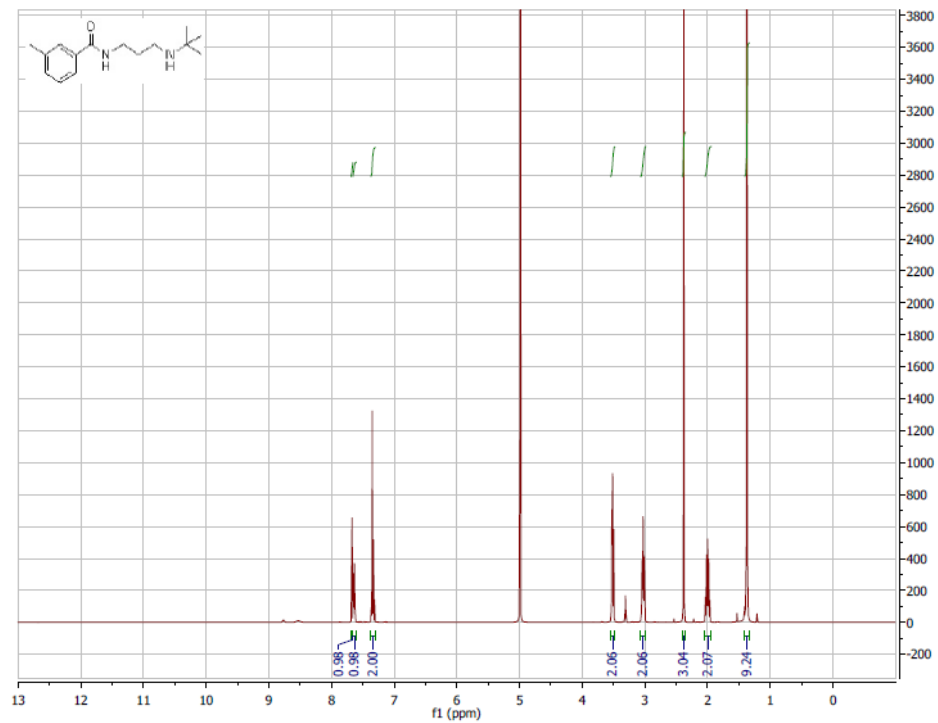
***N*-(3-(*Tert*-butylamino)propyl)-3-fluorobenzamide (86)**



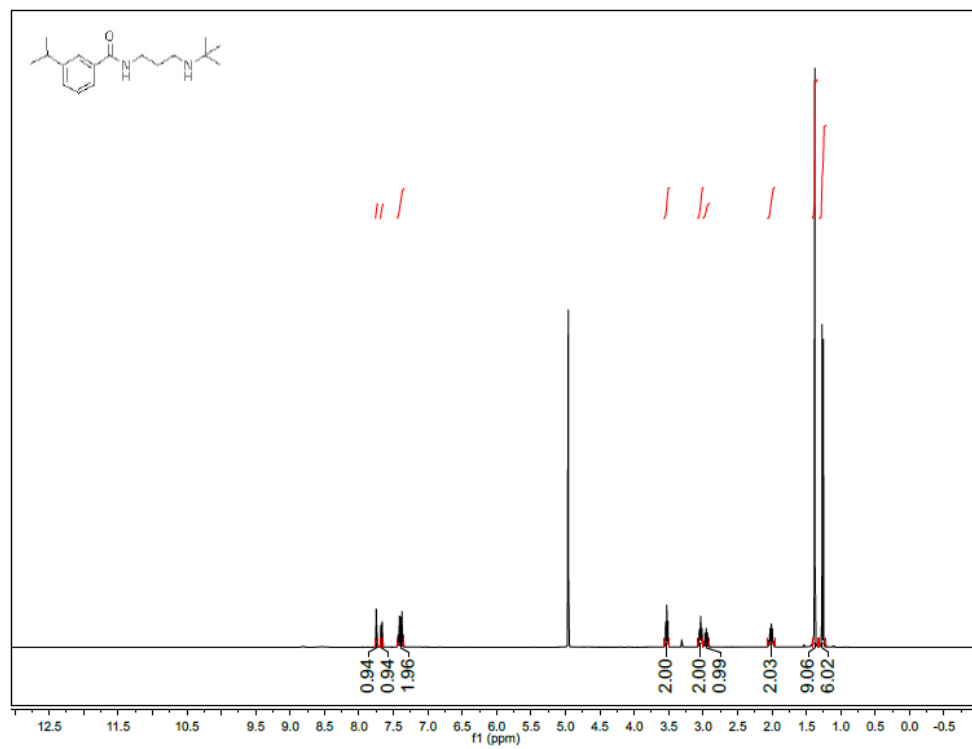
***N*-(3-(*Tert*-butylamino)propyl)-3-iodobenzamide (87)**



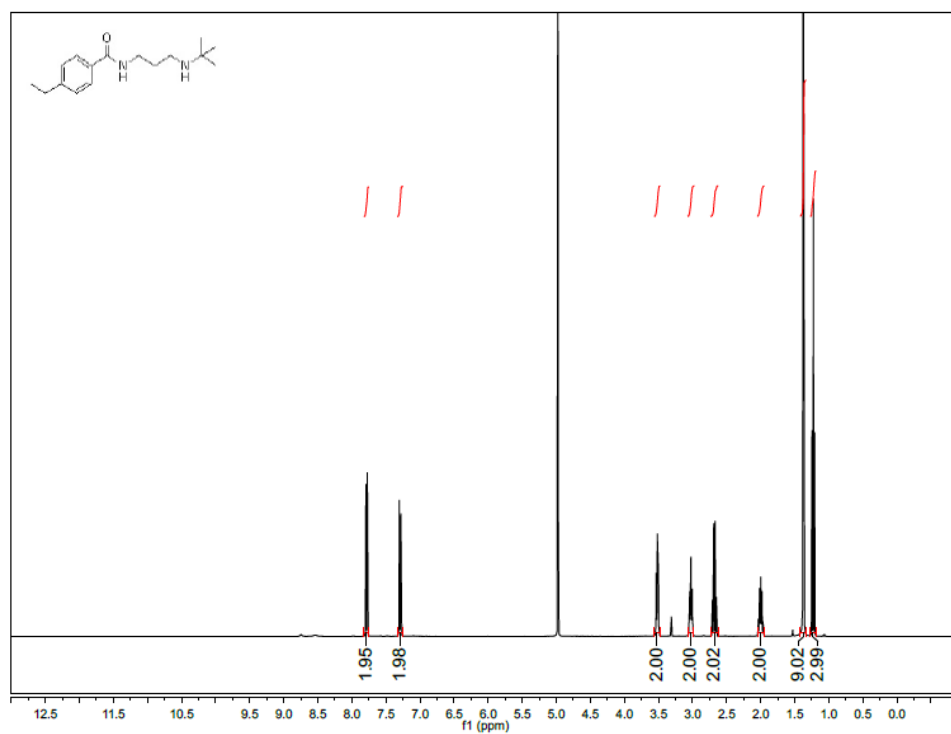
***N*-(3-(*tert*-butylamino)propyl)-3-methylbenzamide (88)**



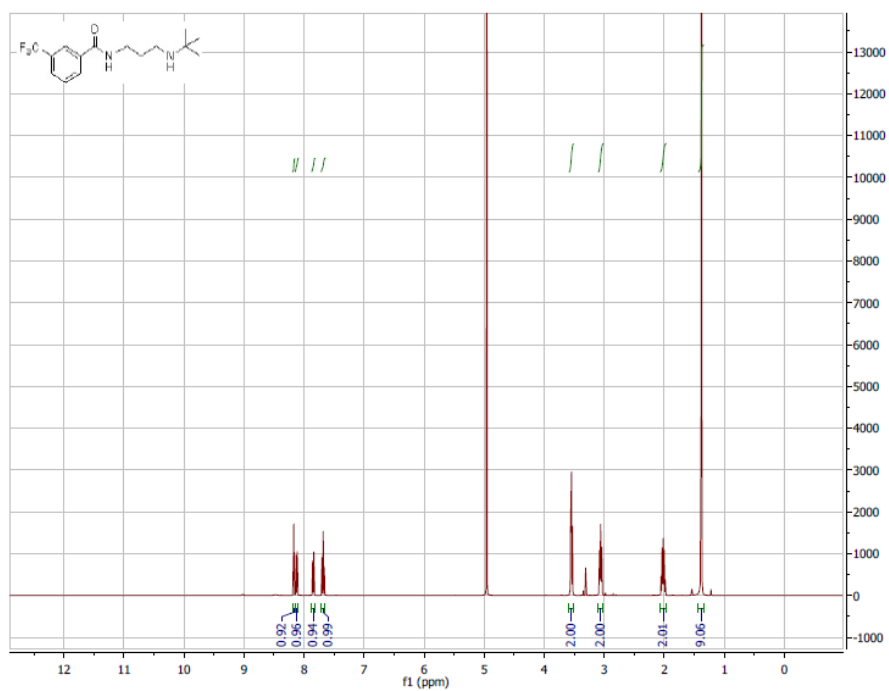
***N*-(3-(*Tert*-butylamino)propyl)-3-isopropylbenzamide (89)**



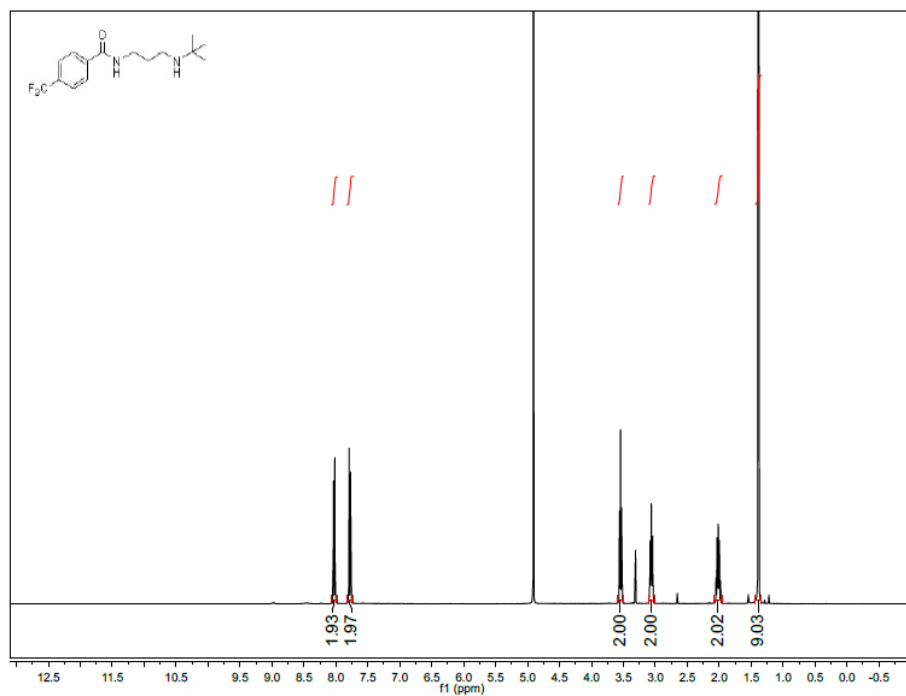
***N*-(3-(*tert*-butylamino)propyl)-4-ethylbenzamide (90)**



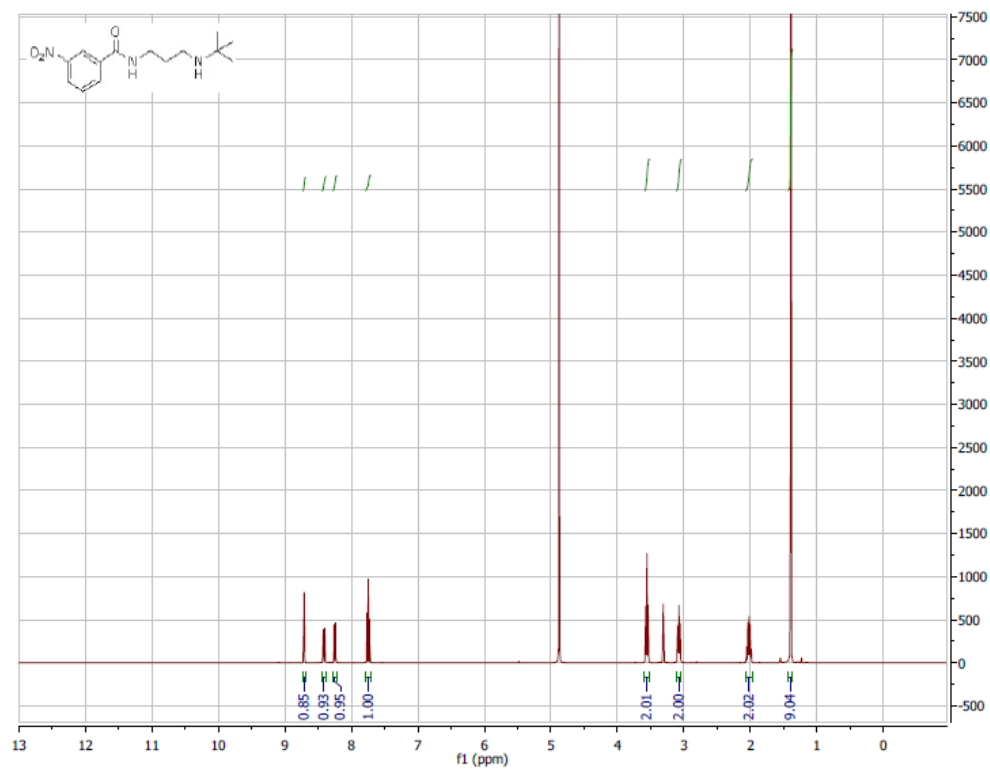
***N*-(3-(*Tert*-butylamino)propyl)-3-(trifluoromethyl)benzamide (91)**



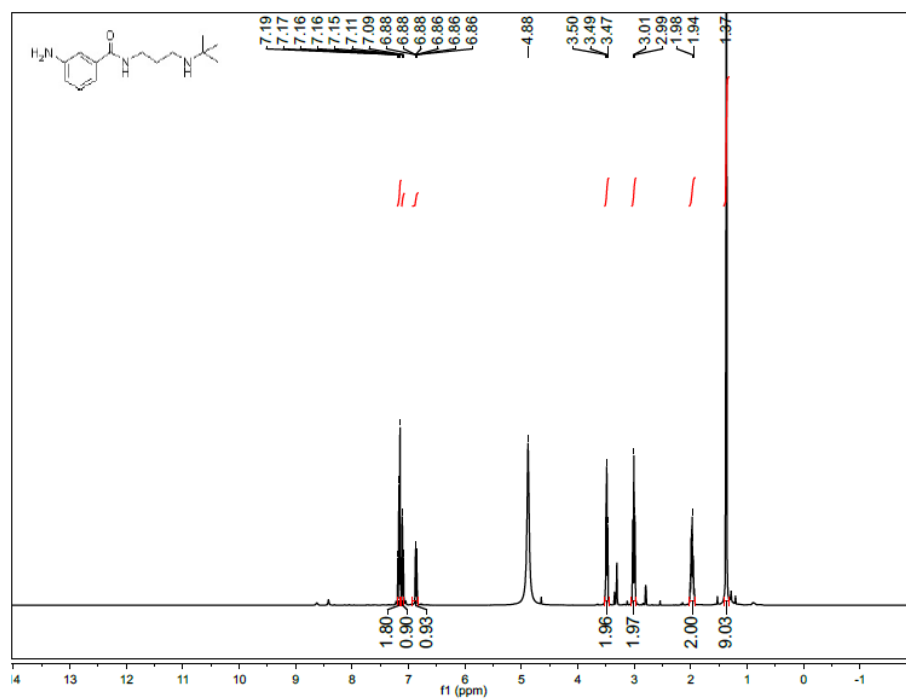
***N*-(3-(*Tert*-butylamino)propyl)-4-(trifluoromethyl)benzamide (92)**



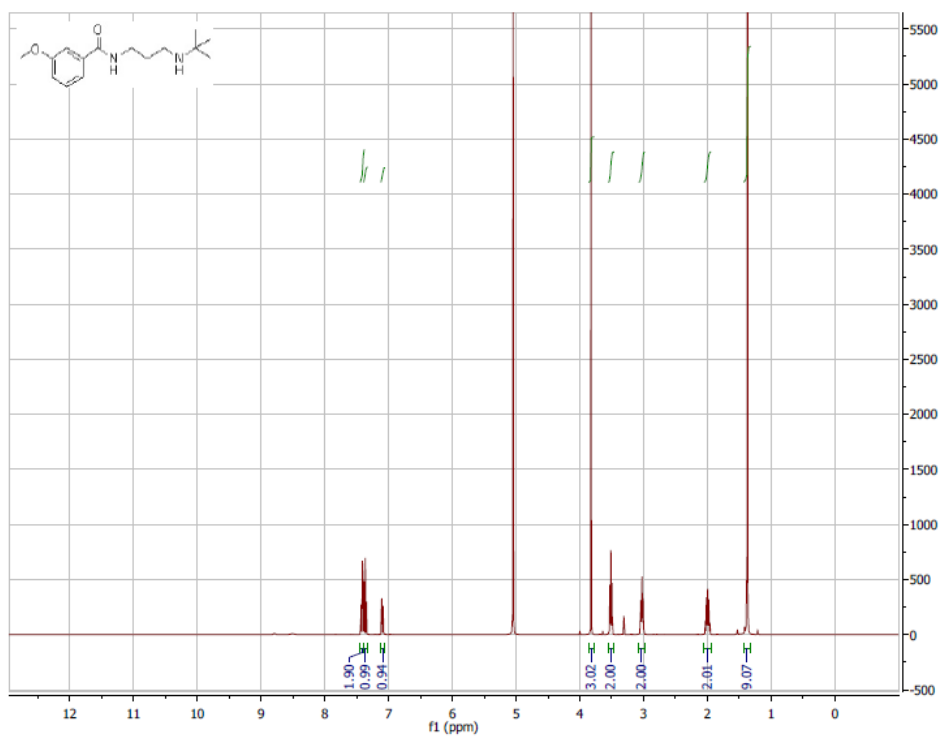
***N*-(3-(*Tert*-butylamino)propyl)-3-nitrobenzamide (93)**



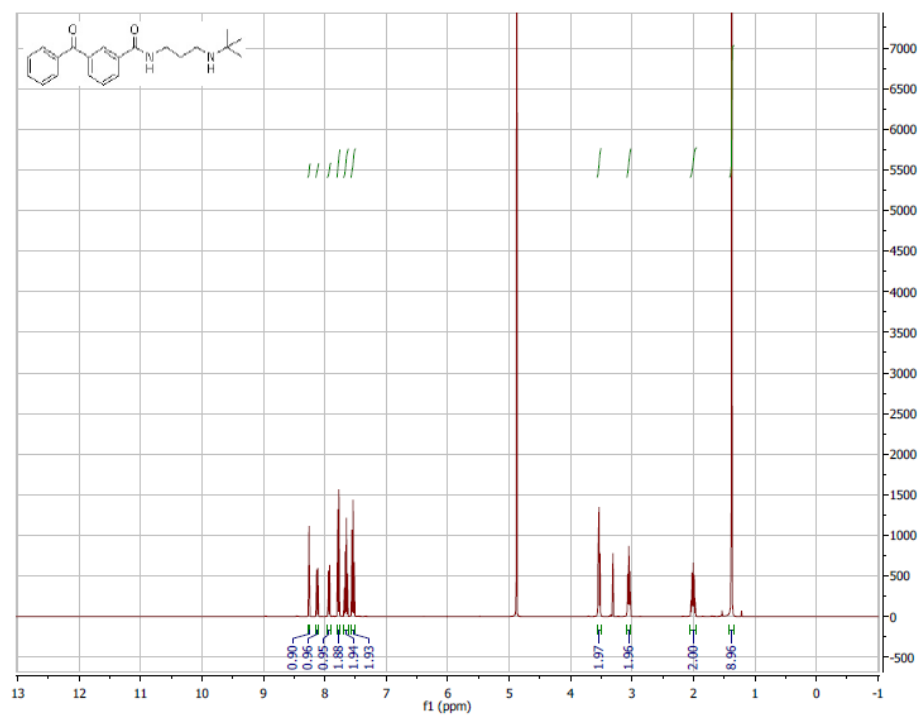
3-amino-N-(3-(tert-butylamino)propyl)benzamide (94)



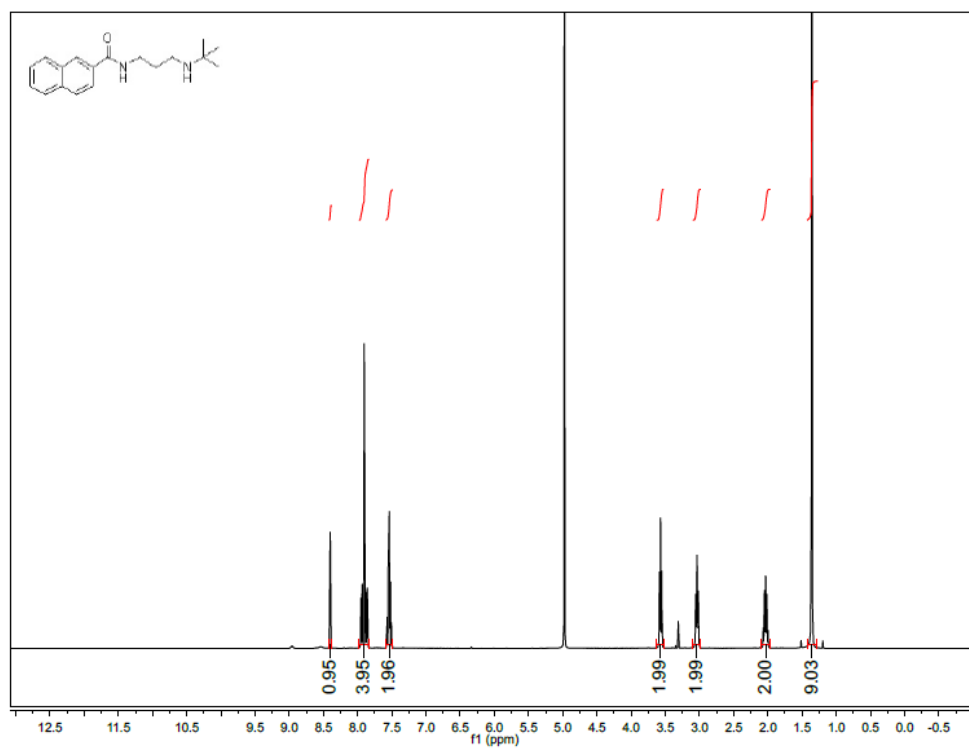
N-(3-(tert-butylamino)propyl)-3-methoxybenzamide (95)



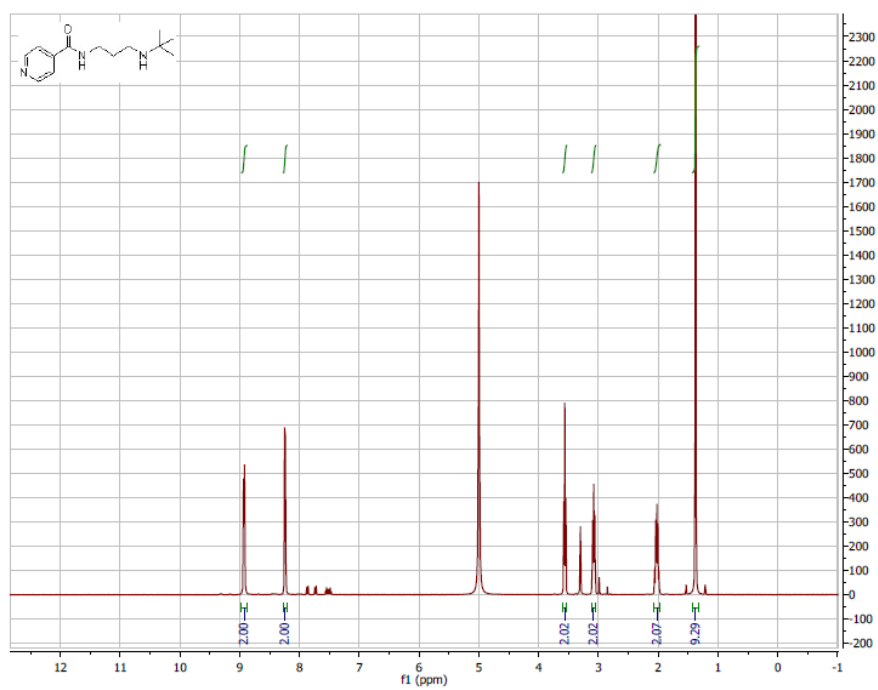
3-benzoyl-*N*-(3-(*tert*-butylamino)propyl)benzamide (96)



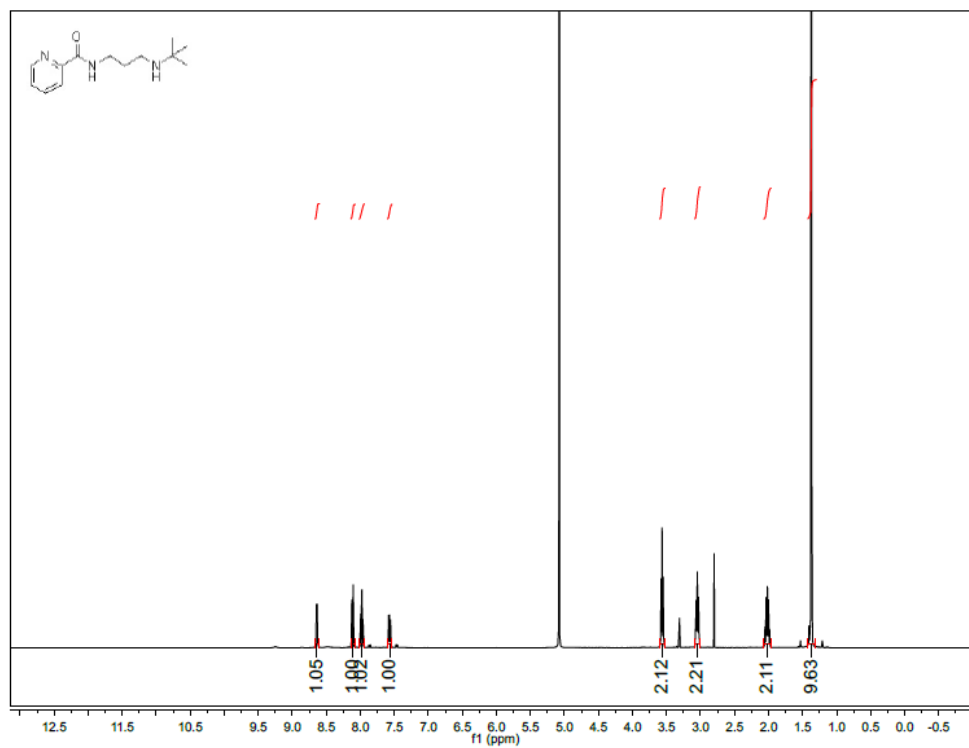
***N*-(3-(*tert*-butylamino)propyl)-2-naphthamide (97)**



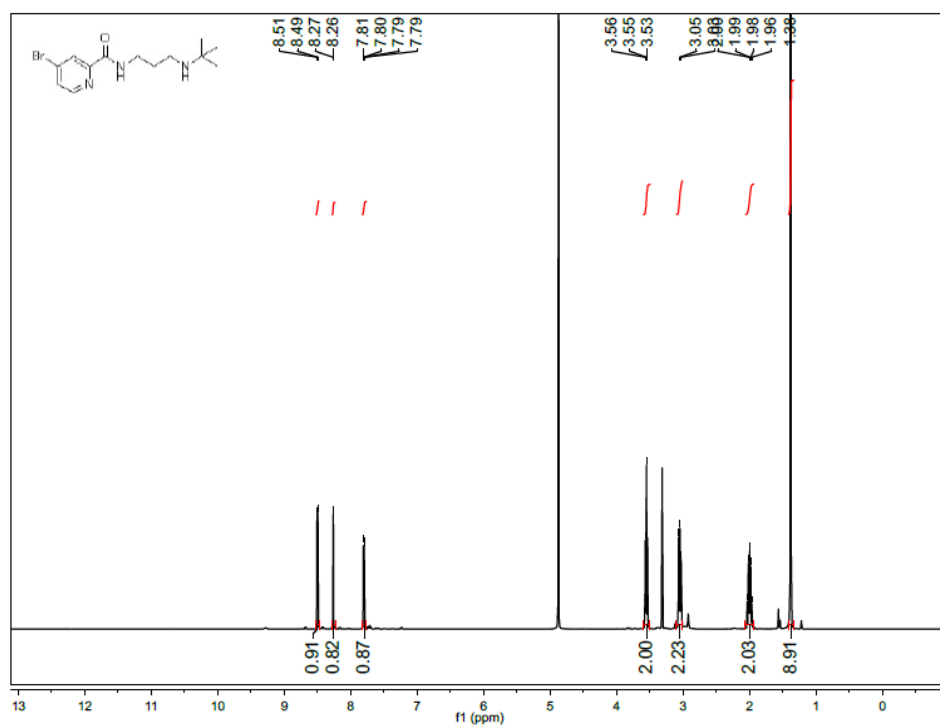
***N*-(3-(tert-butylamino)propyl)picolinamide (98)**



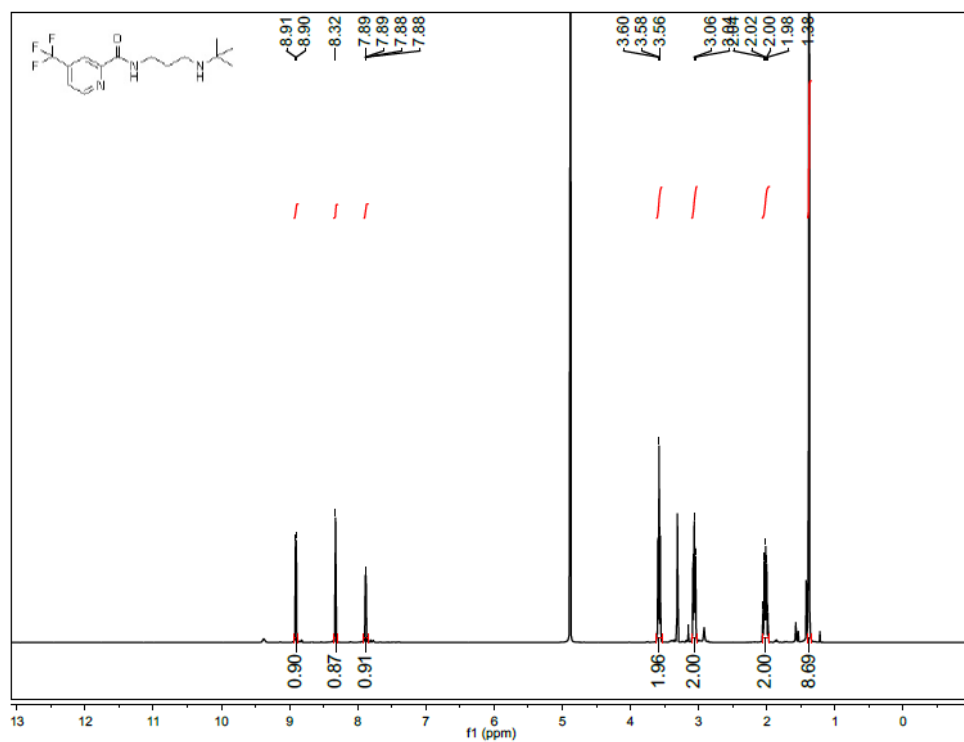
***N*-(3-(tert-butylamino)propyl)picolinamide (99)**



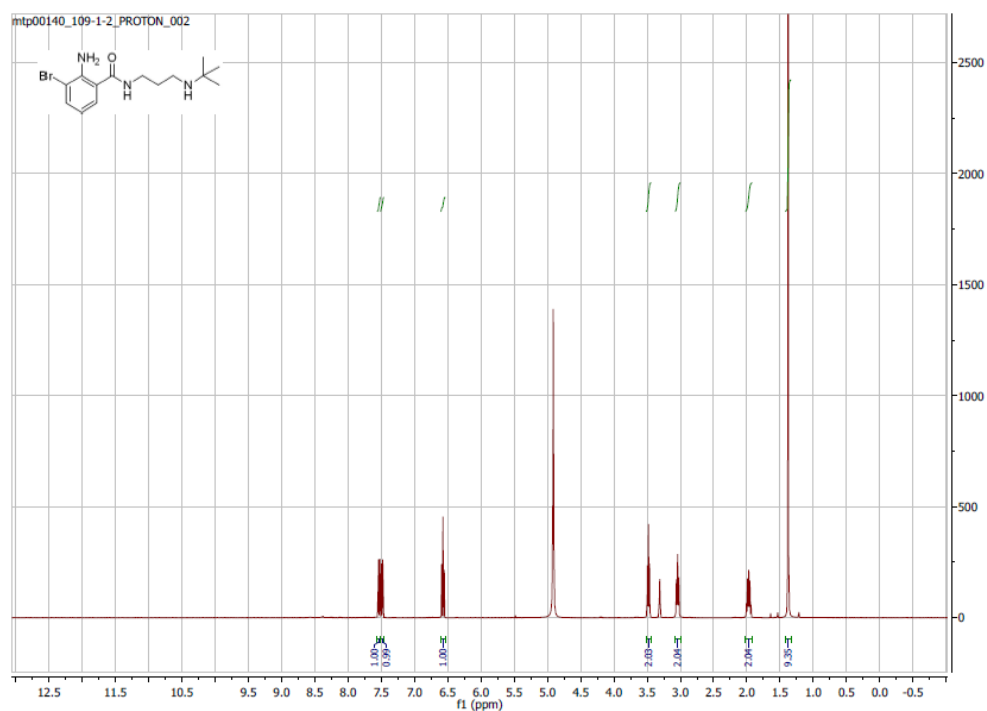
4-Bromo-*N*-(3-(*tert*-butylamino)propyl)picolinamide (100)



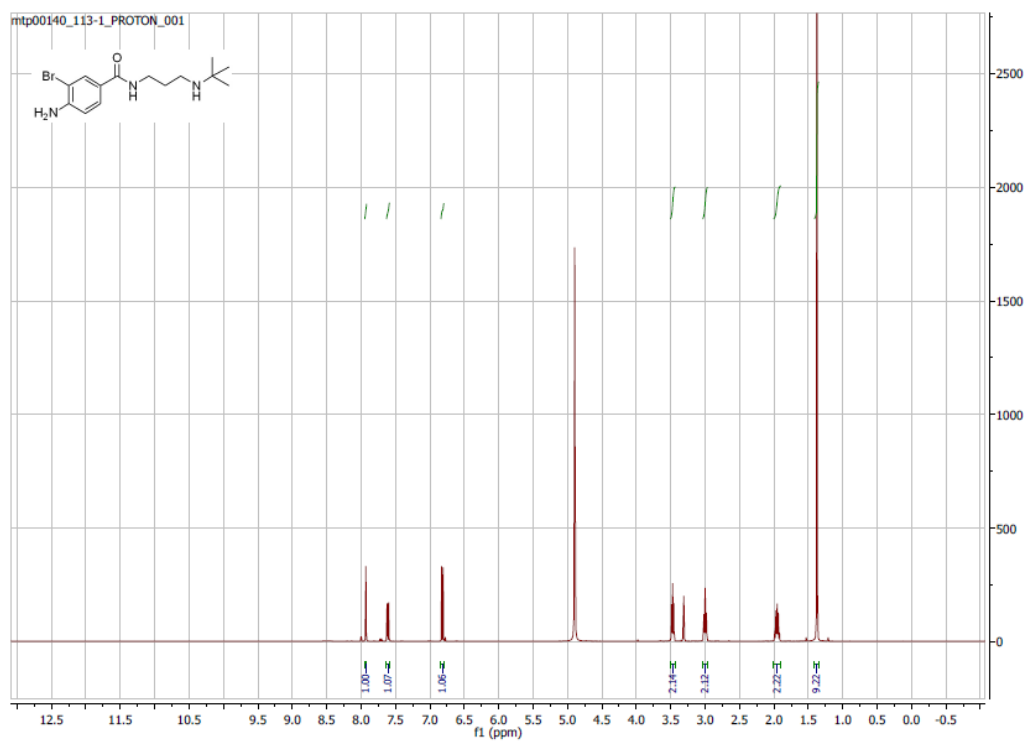
***N*-(3-(*Tert*-butylamino)propyl)-4-(trifluoromethyl)picolinamide (101)**



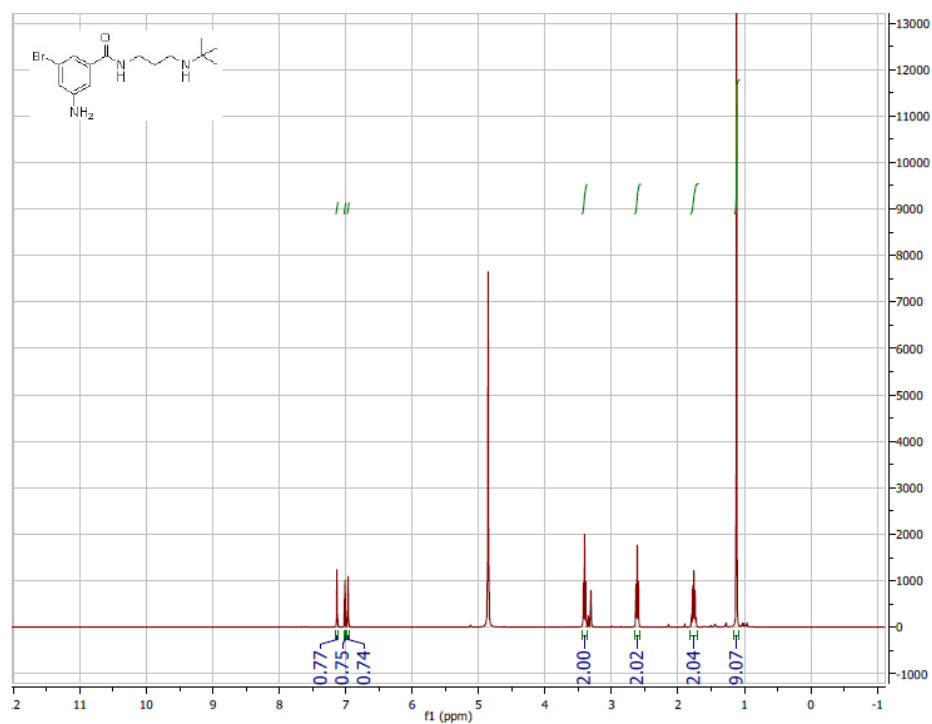
2-amino-3-bromo-N-(3-(tert-butylamino)propyl)benzamide (102)



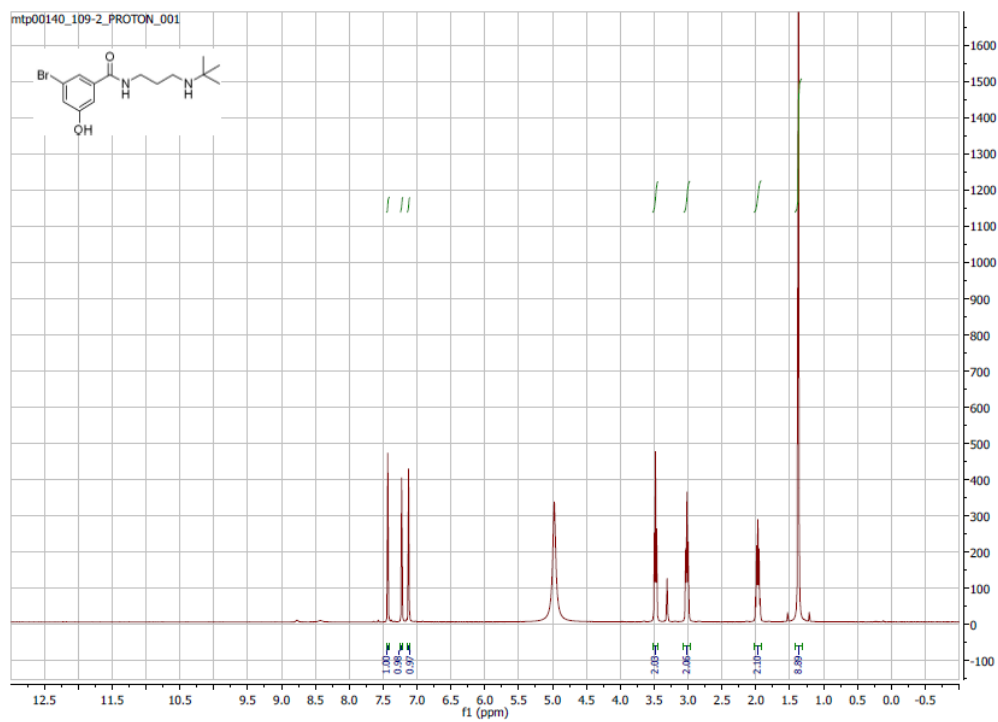
4-amino-3-bromo-N-(3-(tert-butylamino)propyl)benzamide (103)



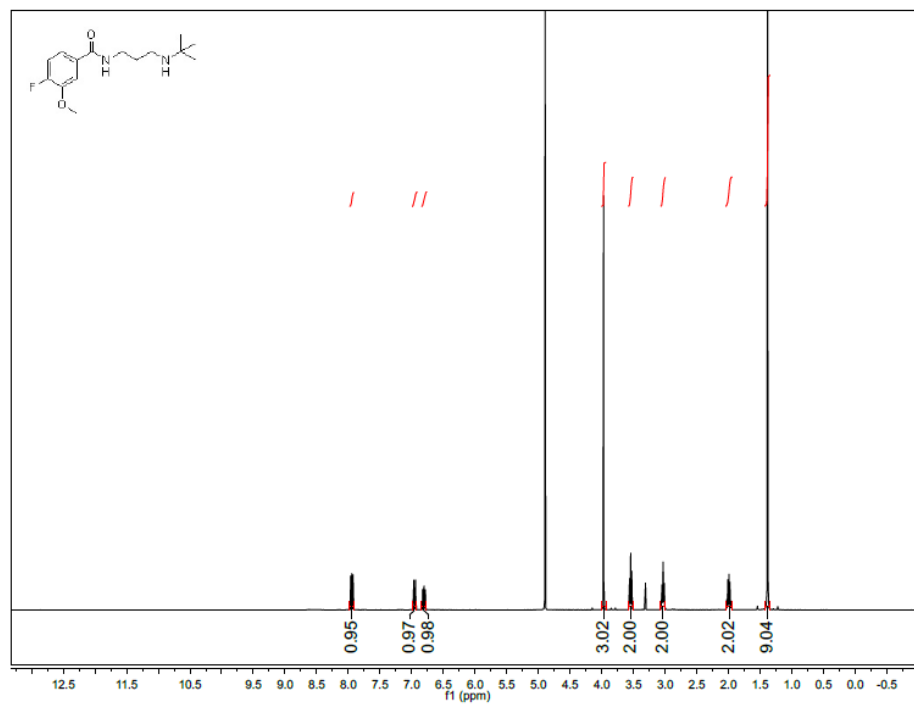
3-amino-5-bromo-N-(3-(tert-butylamino)propyl)benzamide (104)



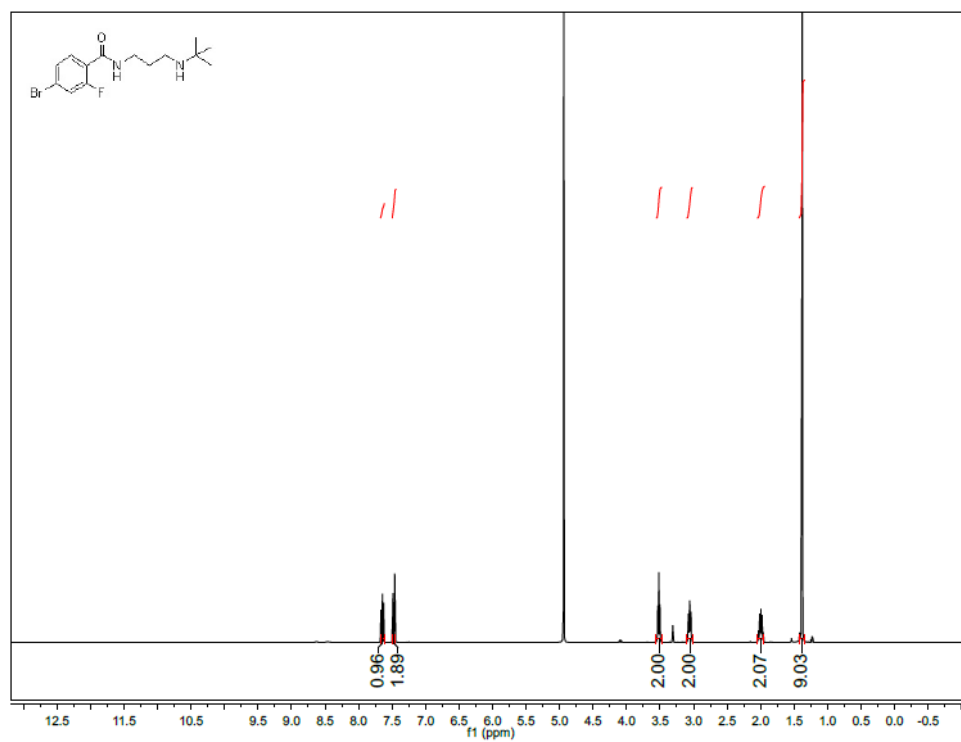
3-bromo-N-(3-(tert-butylamino)propyl)-5-hydroxybenzamide (105)



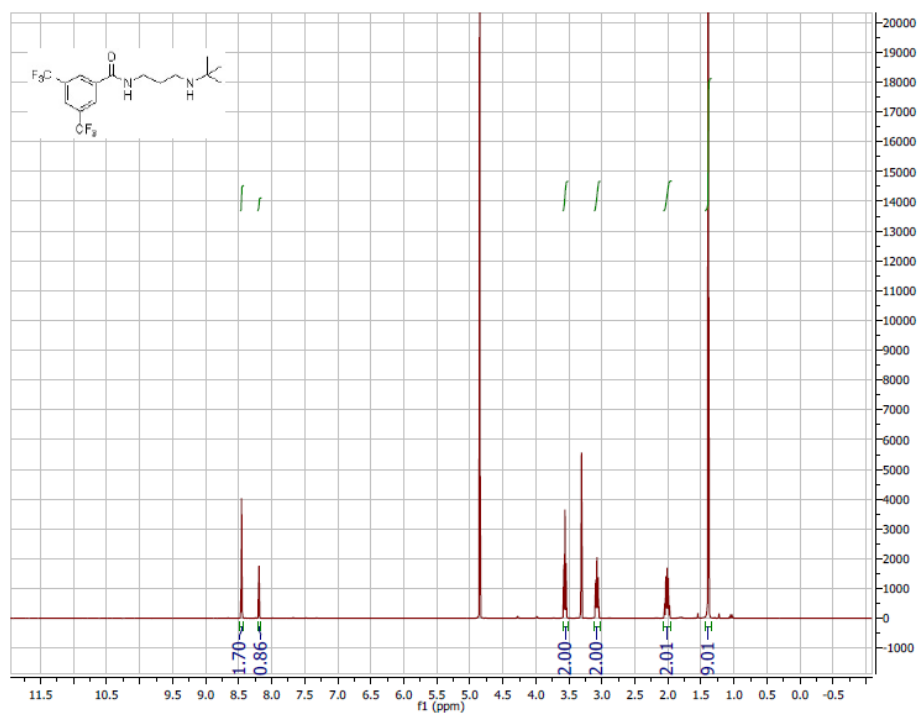
***N*-(3-(tert-butylamino)propyl)-4-fluoro-3-methoxybenzamide (106)**



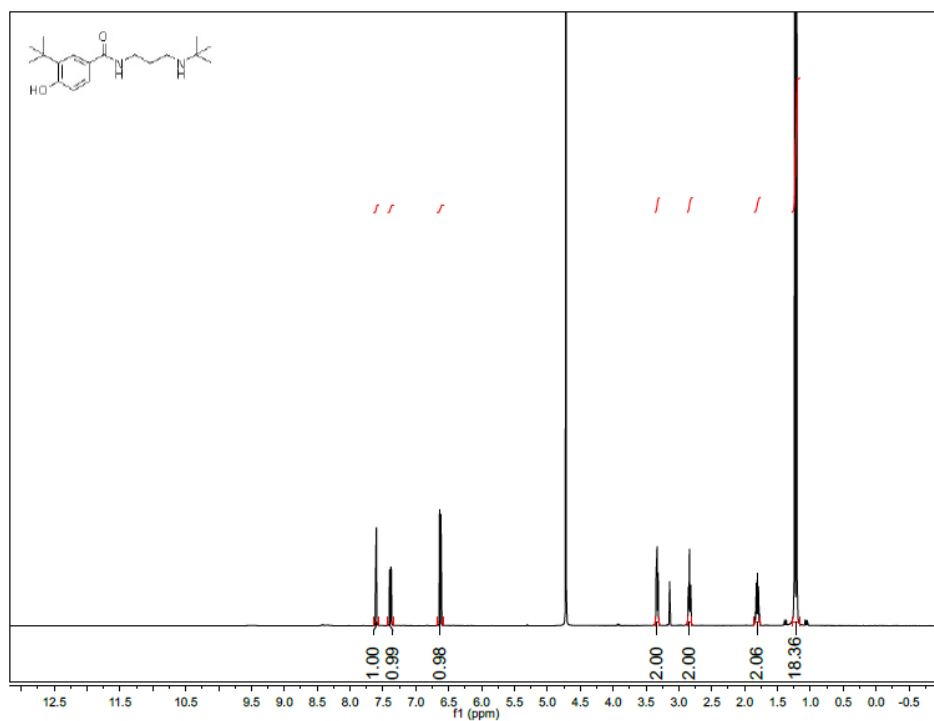
4-bromo-*N*-(3-(tert-butylamino)propyl)-2-fluorobenzamide (107)



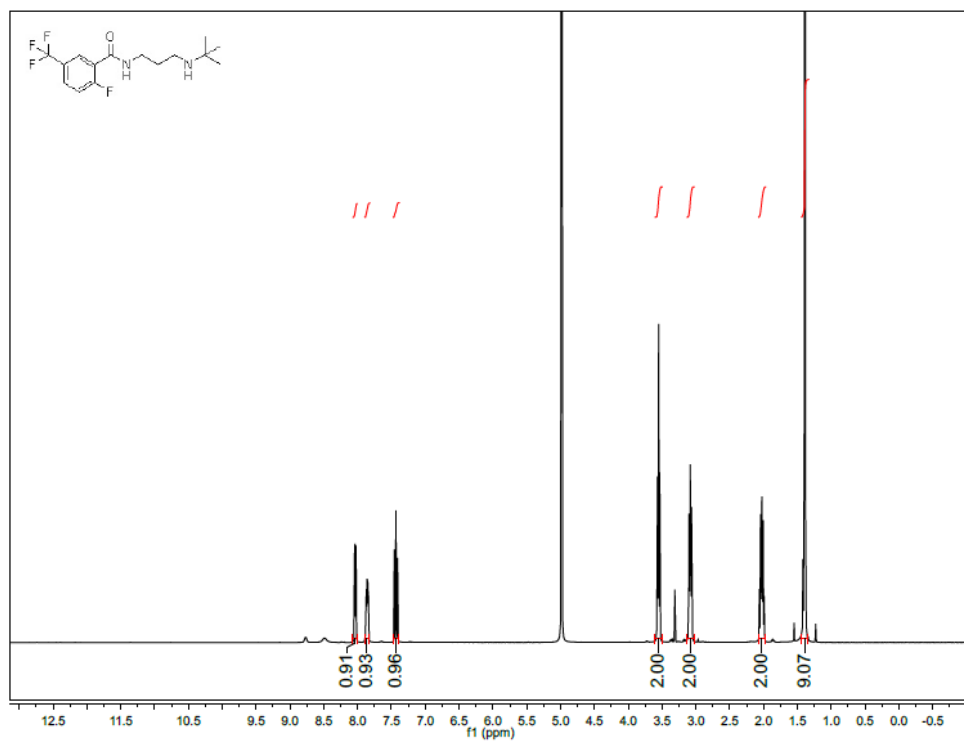
***N*-(3-(tert-butylamino)propyl)-3,5-bis(trifluoromethyl)benzamide (108)**



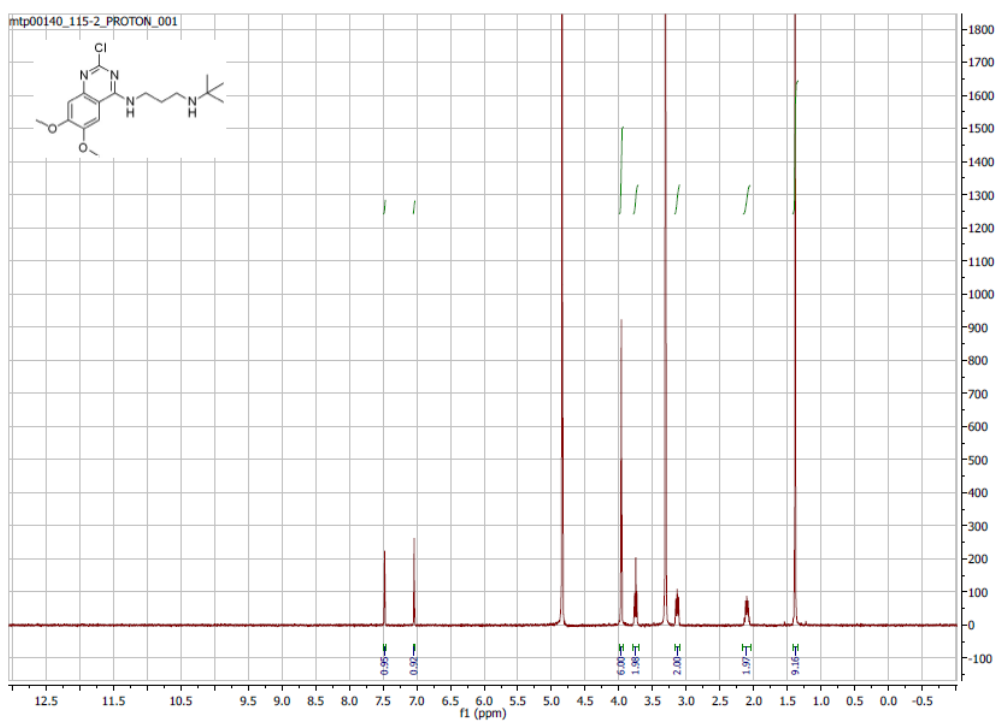
3-(tert-butyl)-*N*-(3-(tert-butylamino)propyl)-4-hydroxybenzamide (109)



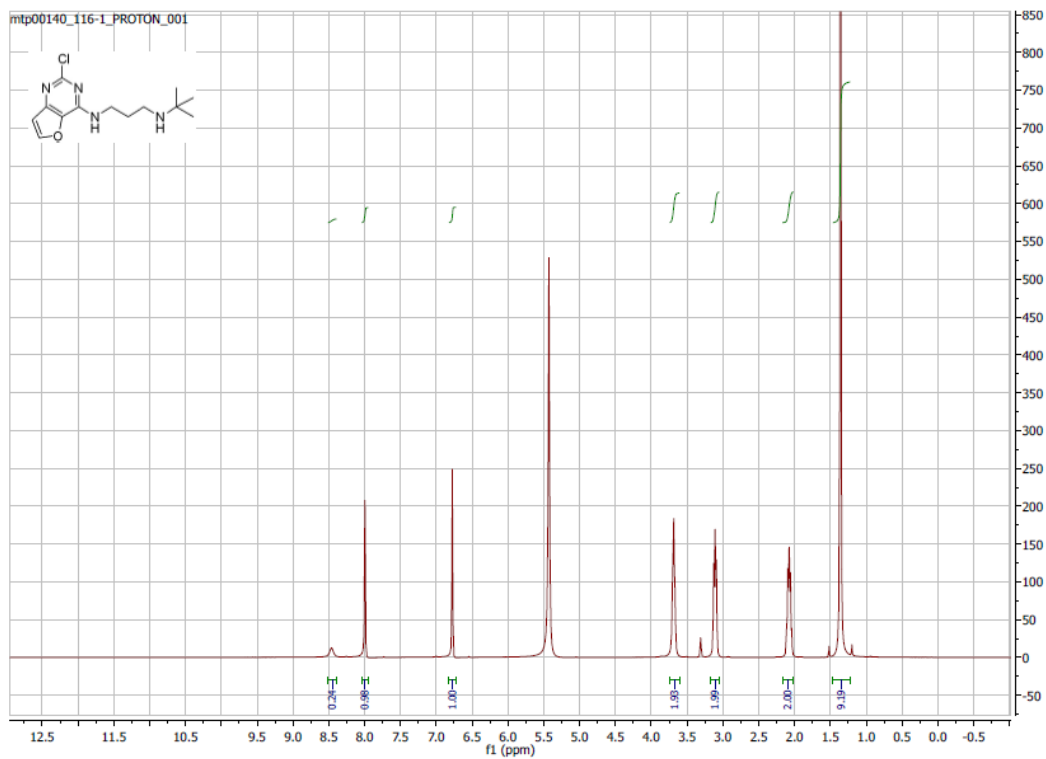
***N*-(3-(tert-butylamino)propyl)-2-fluoro-5-(trifluoromethyl)benzamide (110)**



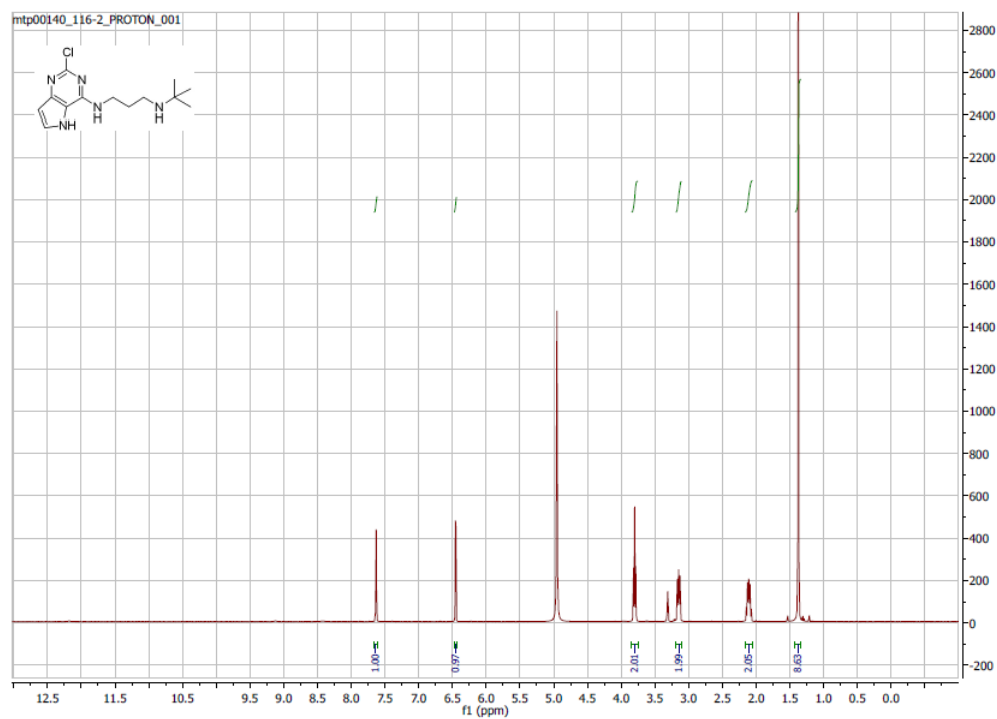
***N*¹-(tert-butyl)-*N*³-(4-chloro-6,7-dimethoxyquinazolin-2-yl)propane-1,3-diamine (111)**



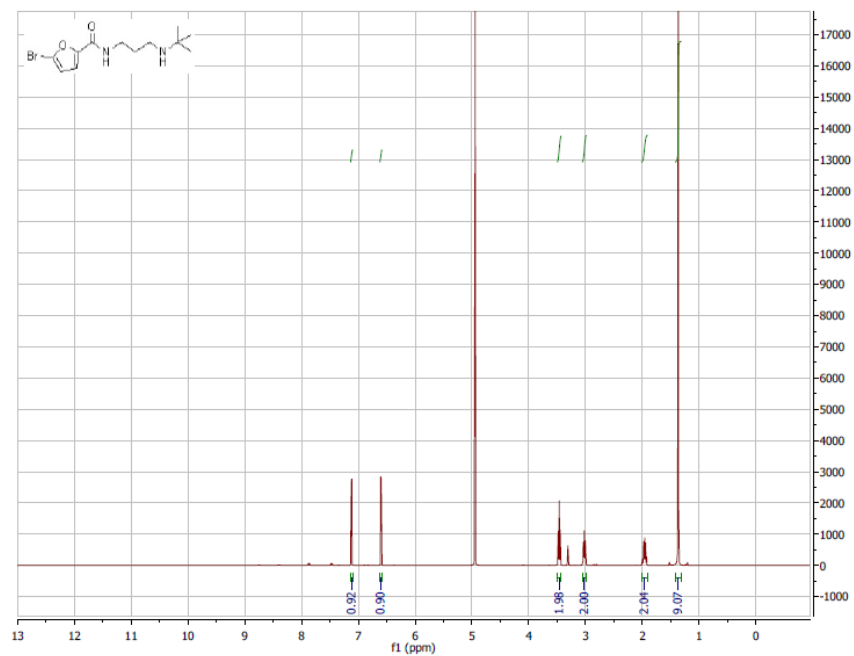
***N*¹-(tert-butyl)-*N*³-(4-chlorofuro[3,2-d]pyrimidin-2-yl)propane-1,3-diamine (112)**



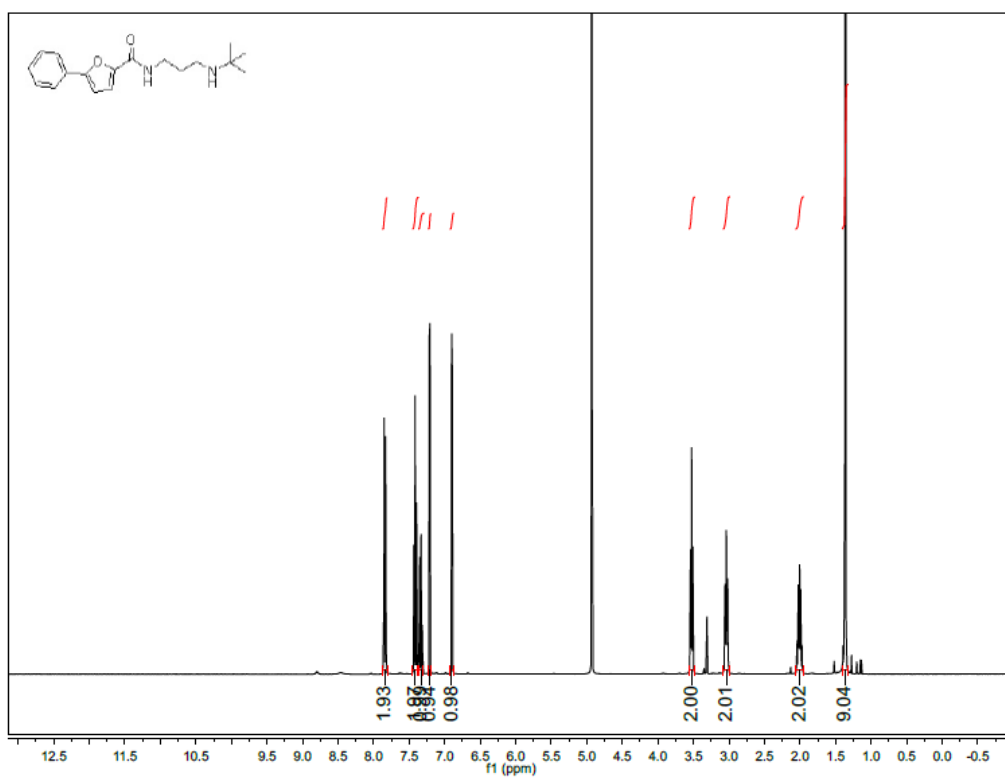
***N*¹-(tert-butyl)-*N*³-(4-chloro-5H-pyrrolo[3,2-d]pyrimidin-2-yl)propane-1,3-diamine (113)**



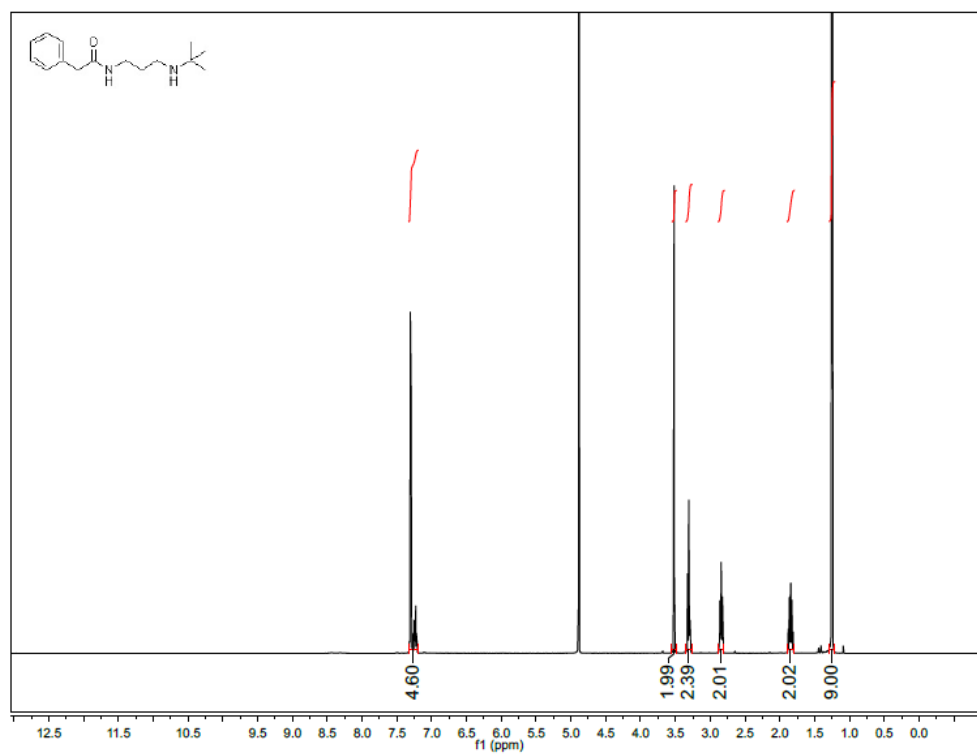
5-bromo-*N*-(3-(*tert*-butylamino)propyl)furan-2-carboxamide (114)



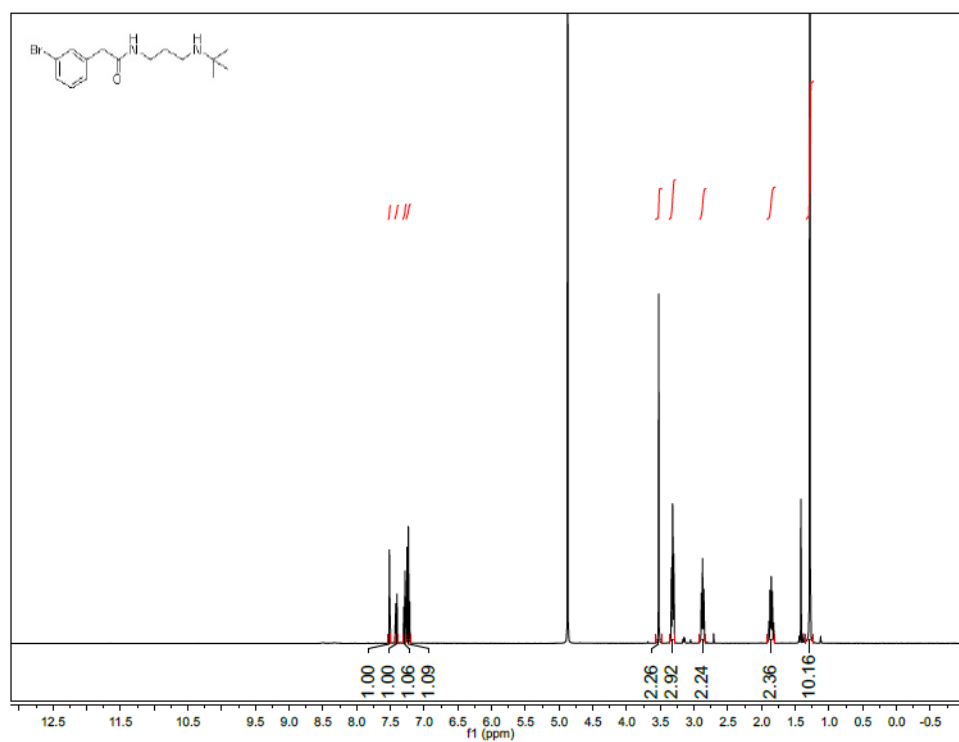
***N*-(3-(*tert*-butylamino)propyl)-5-phenylfuran-2-carboxamide (115)**



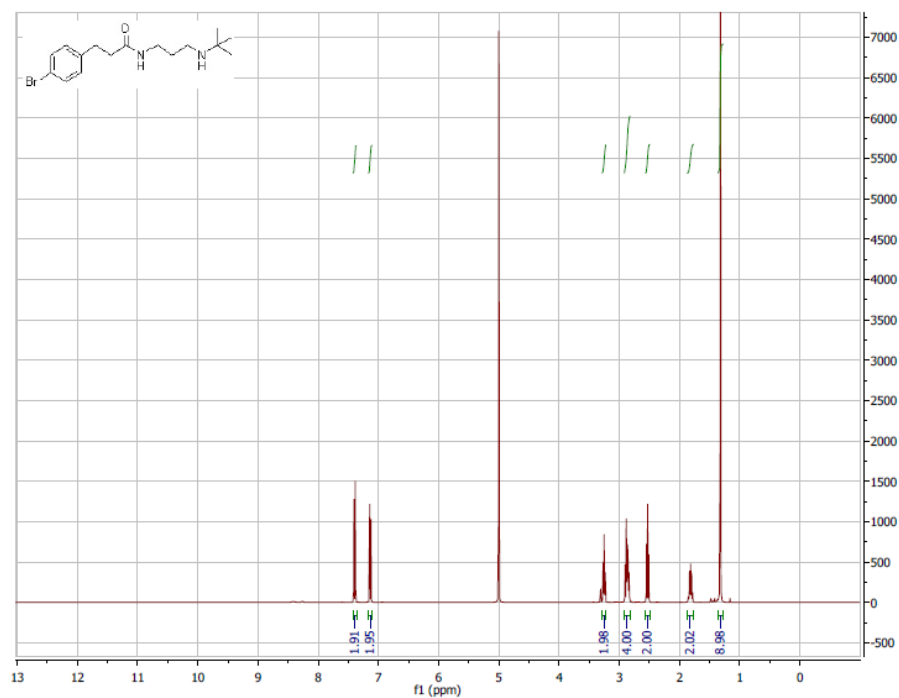
***N*-(3-(tert-butylamino)propyl)-2-phenylacetamide (116)**



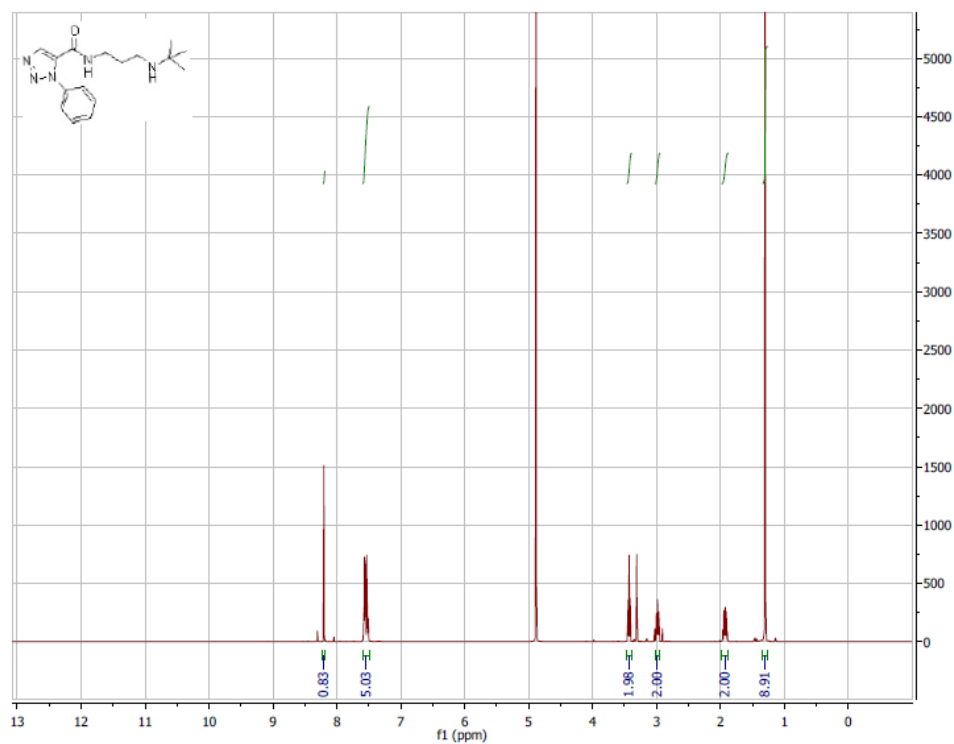
2-(3-bromophenyl)-*N*-(3-(tert-butylamino)propyl)acetamide (117)



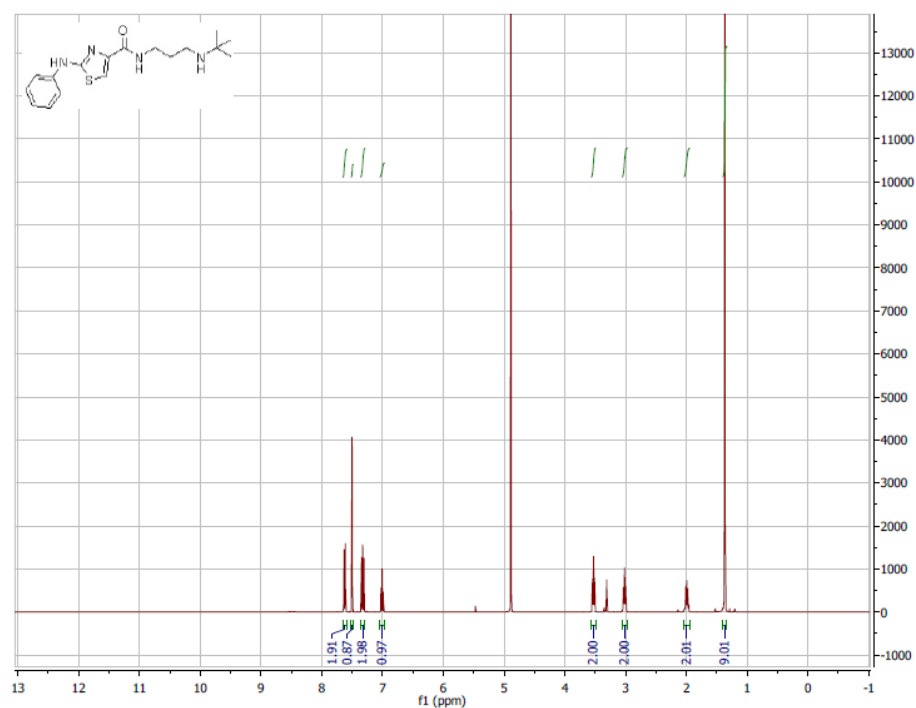
3-(4-bromophenyl)-N-(3-(tert-butylamino)propyl)propanamide (118)



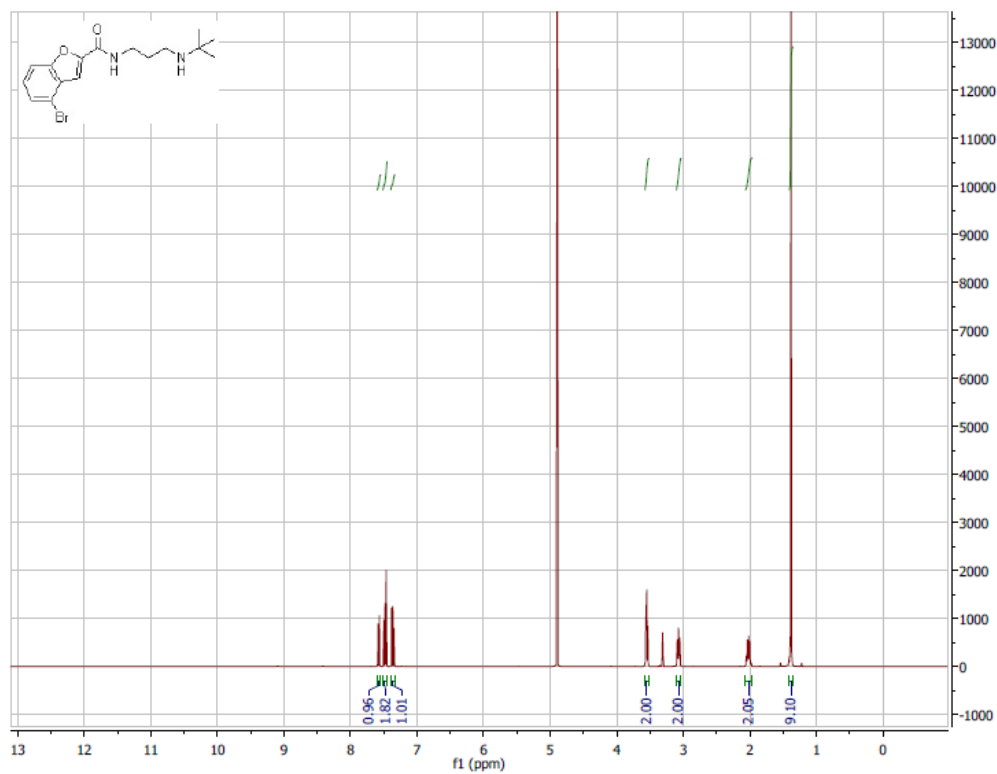
N-(3-(tert-butylamino)propyl)-1-phenyl-1H-1,2,3-triazole-5-carboxamide (119)



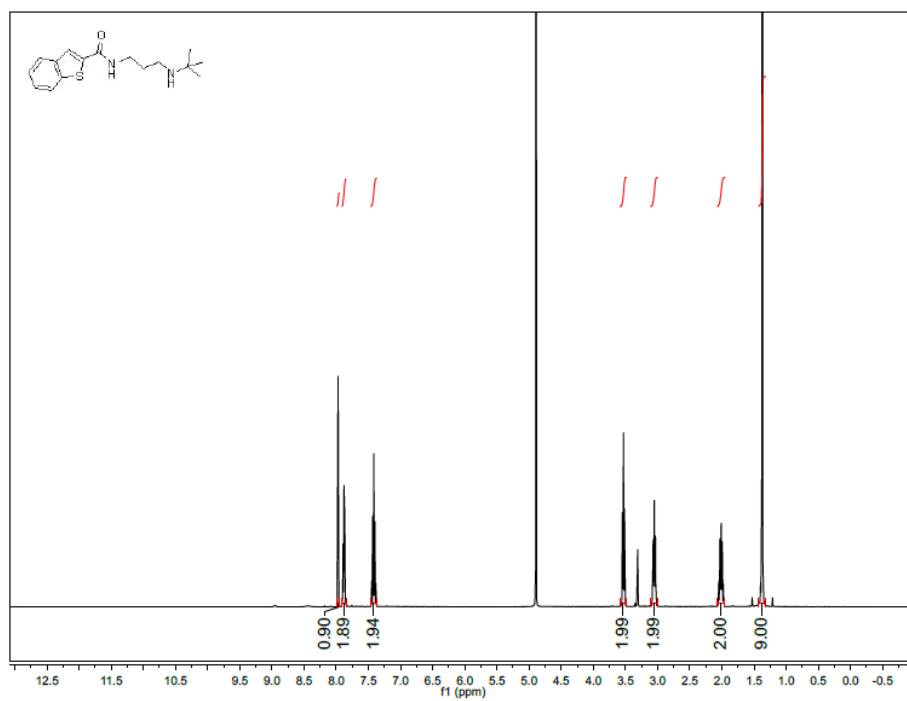
***N*-(3-(tert-butylamino)propyl)-2-(phenylamino)thiazole-4-carboxamide (120)**



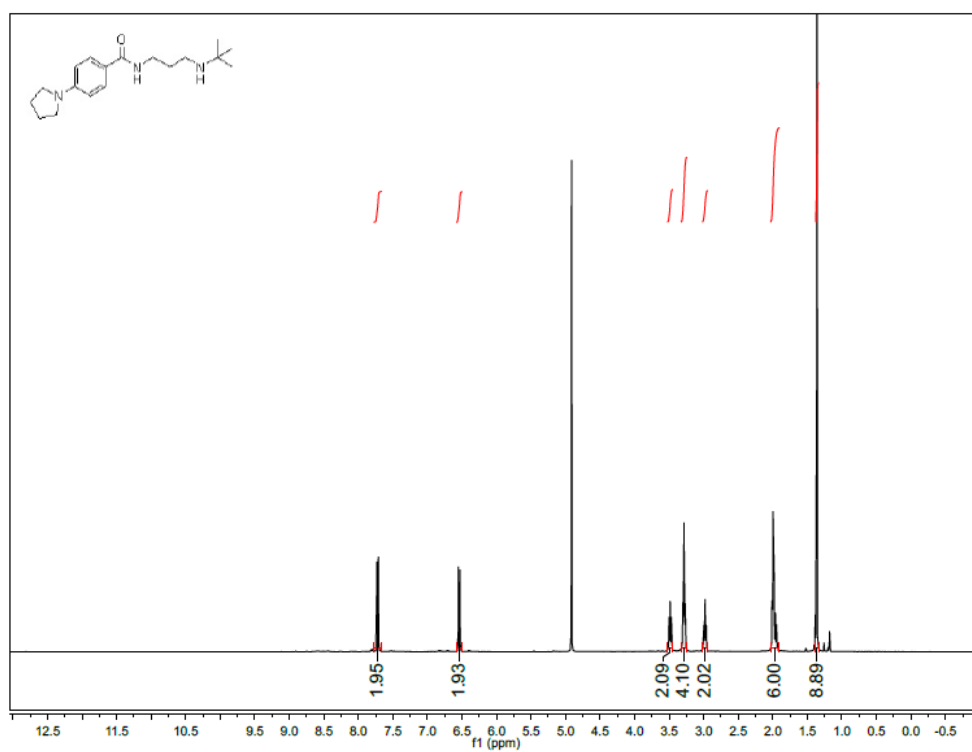
4-bromo-*N*-(3-(tert-butylamino)propyl)benzofuran-2-carboxamide (121)



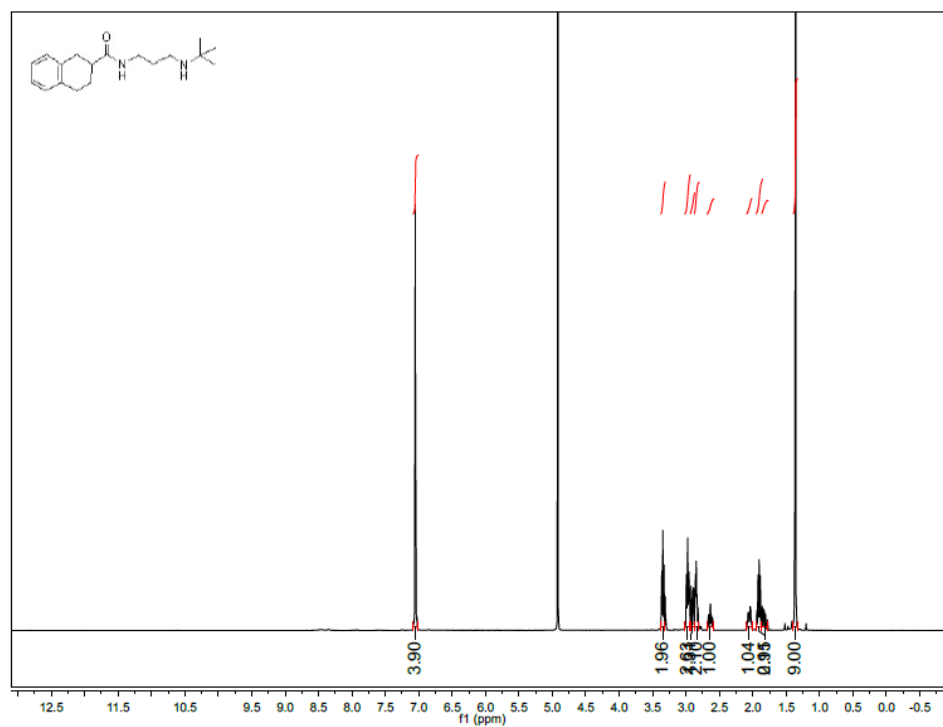
***N*-(3-(tert-butylamino)propyl)benzo[*b*]thiophene-2-carboxamide (122)**



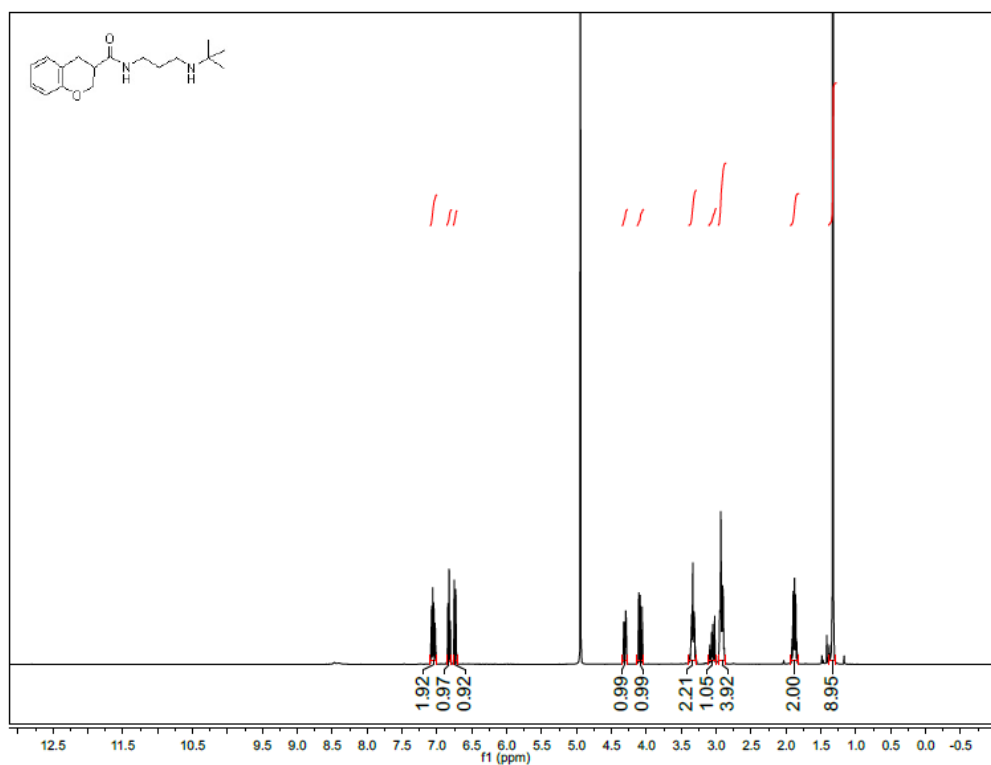
***N*-(3-(tert-butylamino)propyl)benzo[*b*]thiophene-2-carboxamide (123)**



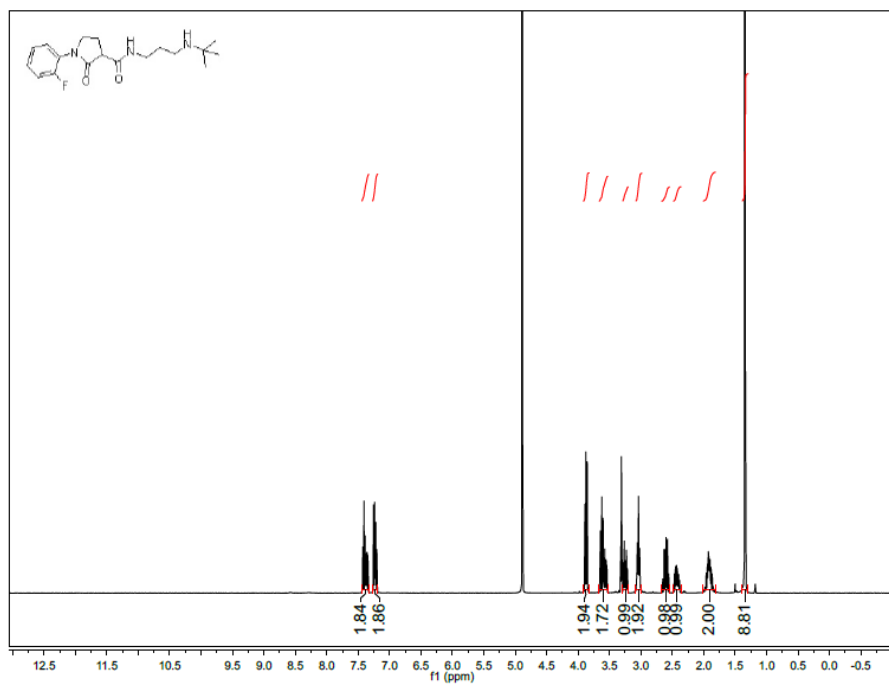
***N*-(3-(tert-butylamino)propyl)-1,2,3,4-tetrahydronaphthalene-2-carboxamide (124)**



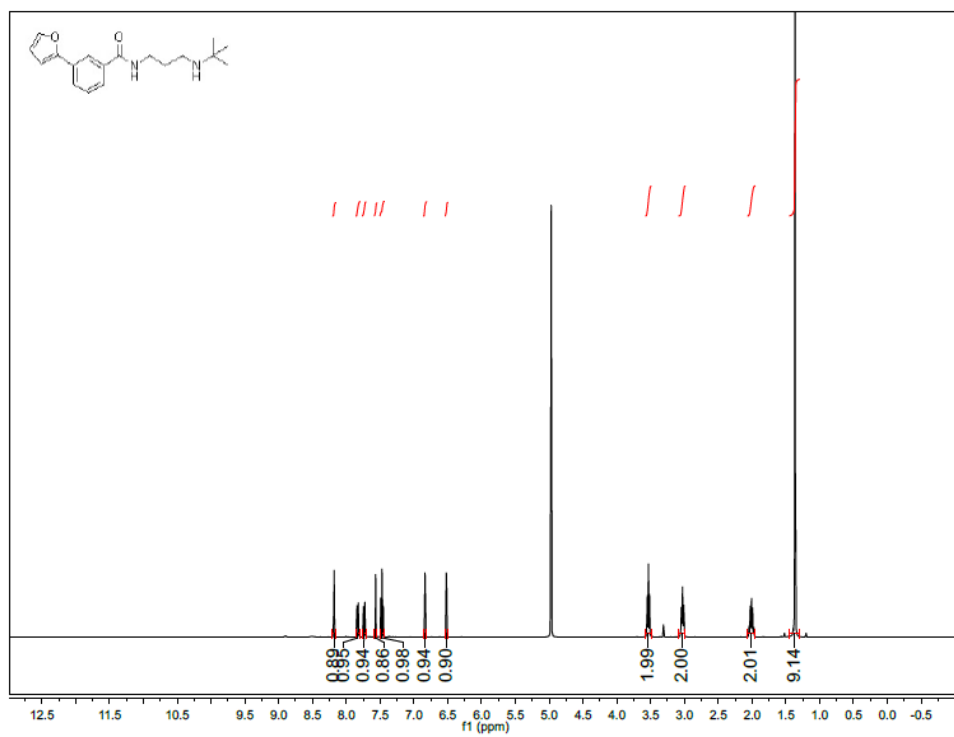
***N*-(3-(tert-butylamino)propyl)chromane-3-carboxamide (125)**



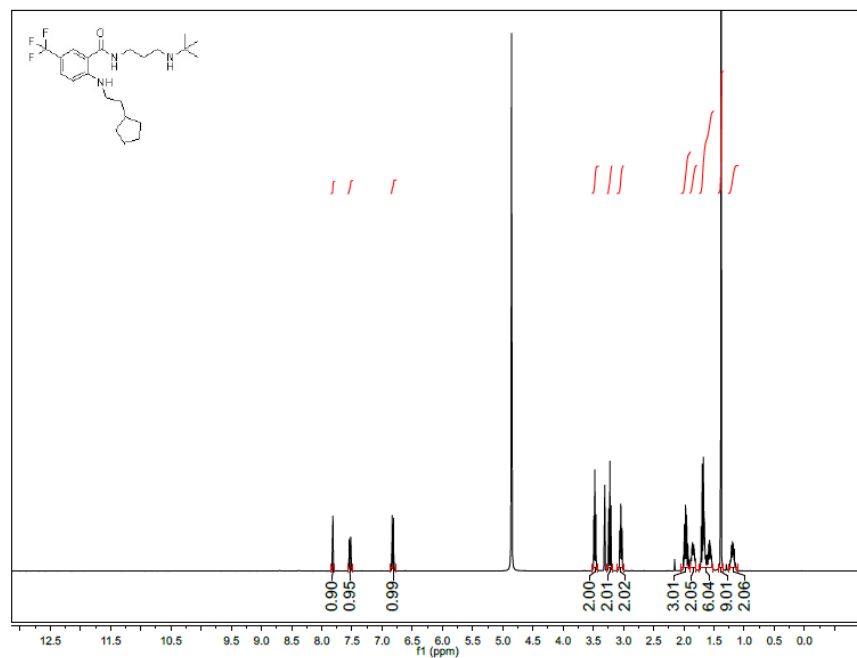
***N*-(3-(tert-butylamino)propyl)-1-(3-fluorophenyl)-2-oxopyrrolidine-3-carboxamide (126)**



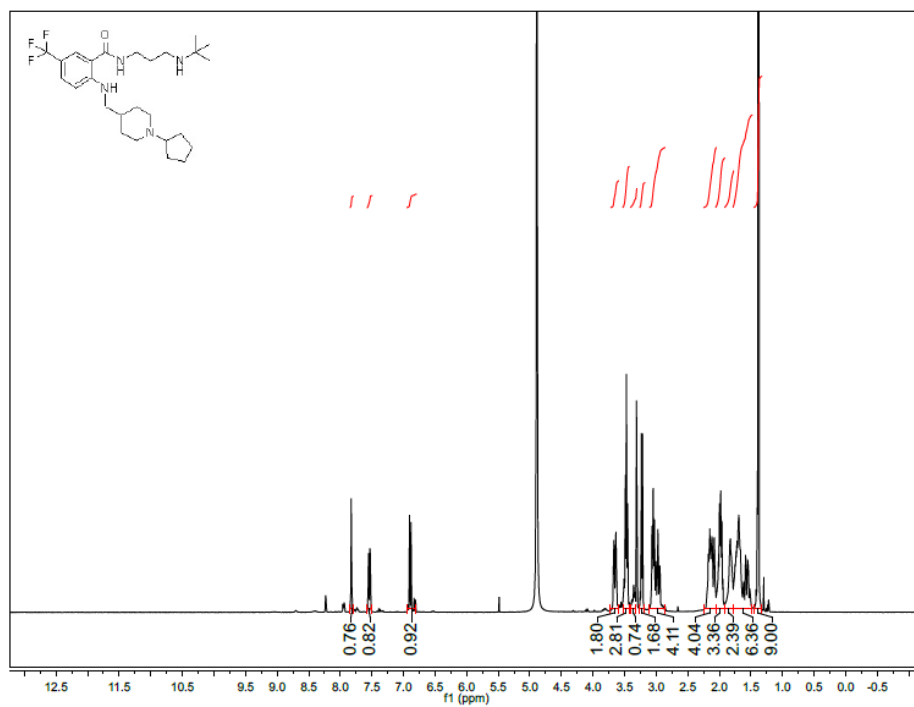
***N*-(3-(tert-butylamino)propyl)-3-(furan-2-yl)benzamide (127)**



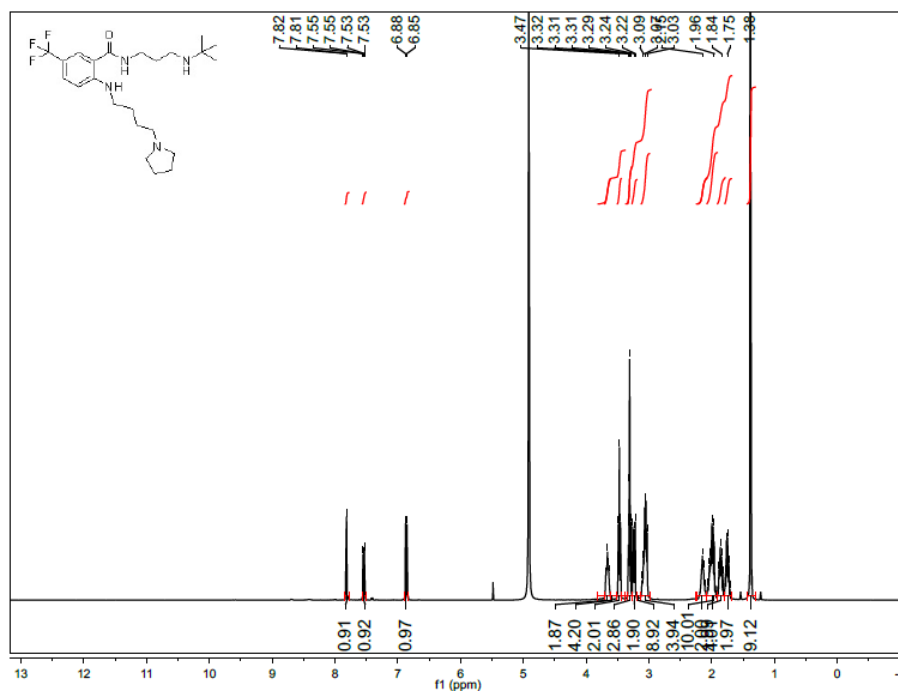
***N*-(3-(tert-butylamino)propyl)-2-((2-cyclopentylethyl)amino)-5-(trifluoromethyl)benzamide (128)**



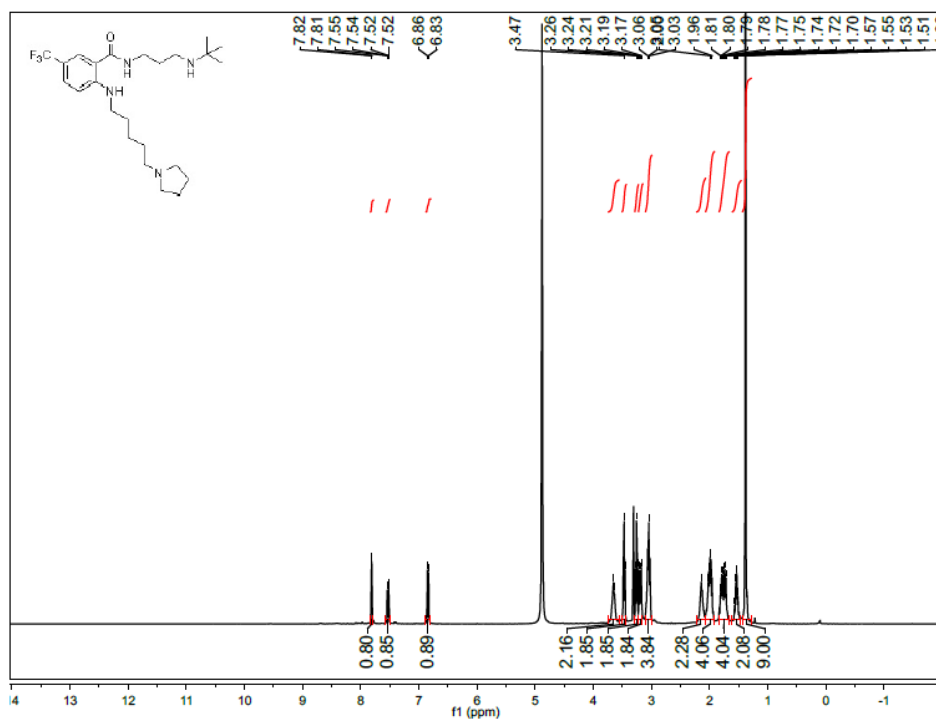
***N*-(3-(tert-butylamino)propyl)-2-(((1-cyclopentylpiperidin-4-yl)methyl)amino)-5-(trifluoromethyl)benzamide (129)**



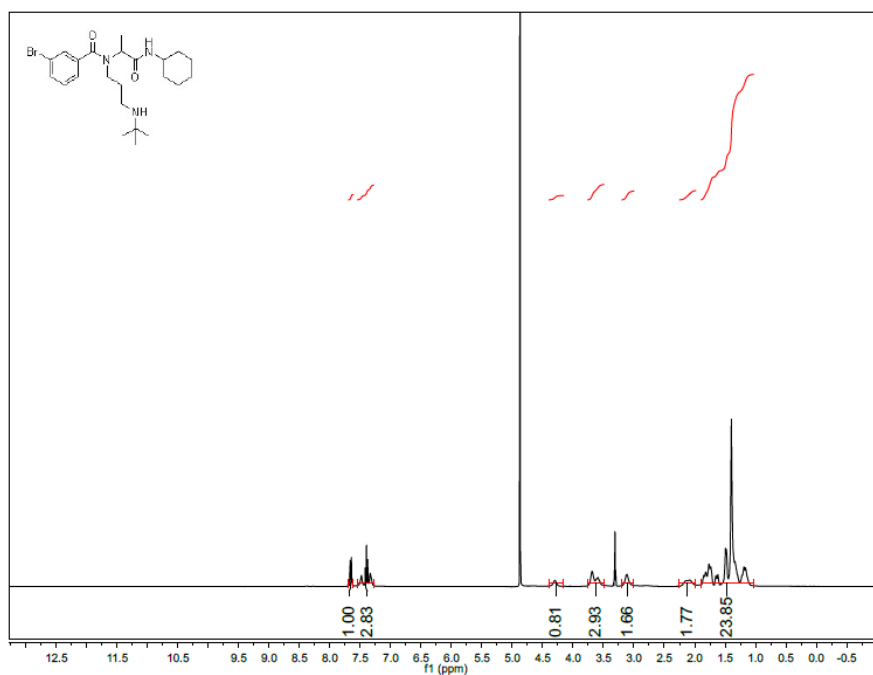
***N*-(3-(tert-butylamino)propyl)-2-((4-(pyrrolidin-1-yl)butyl)amino)-5-(trifluoromethyl)benzamide (130)**



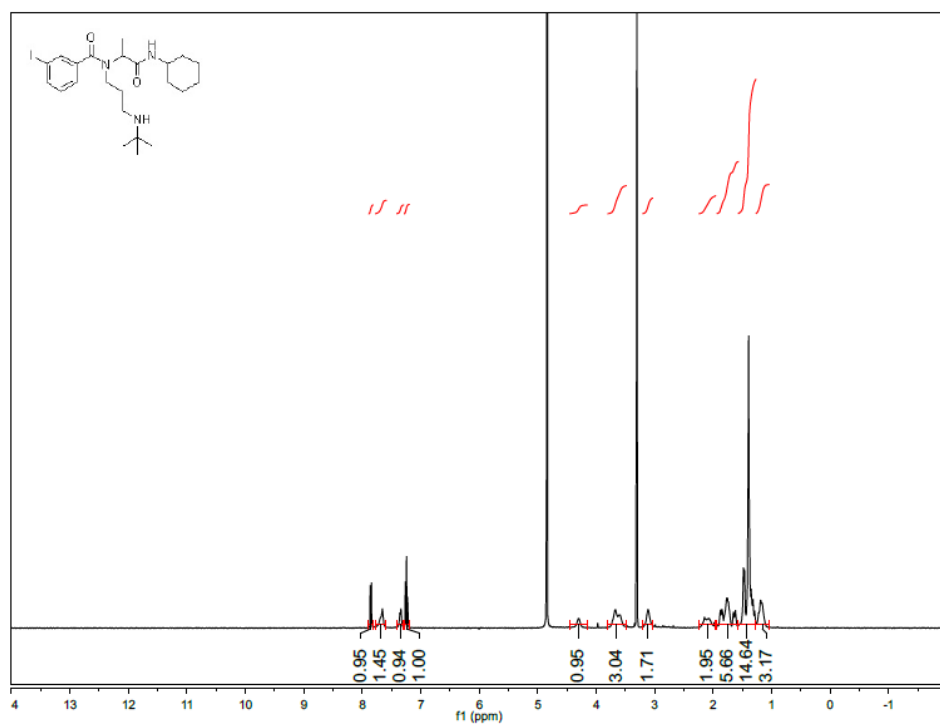
***N*-(3-(tert-butylamino)propyl)-2-((5-(pyrrolidin-1-yl)pentyl)amino)-5-(trifluoromethyl)benzamide (131)**



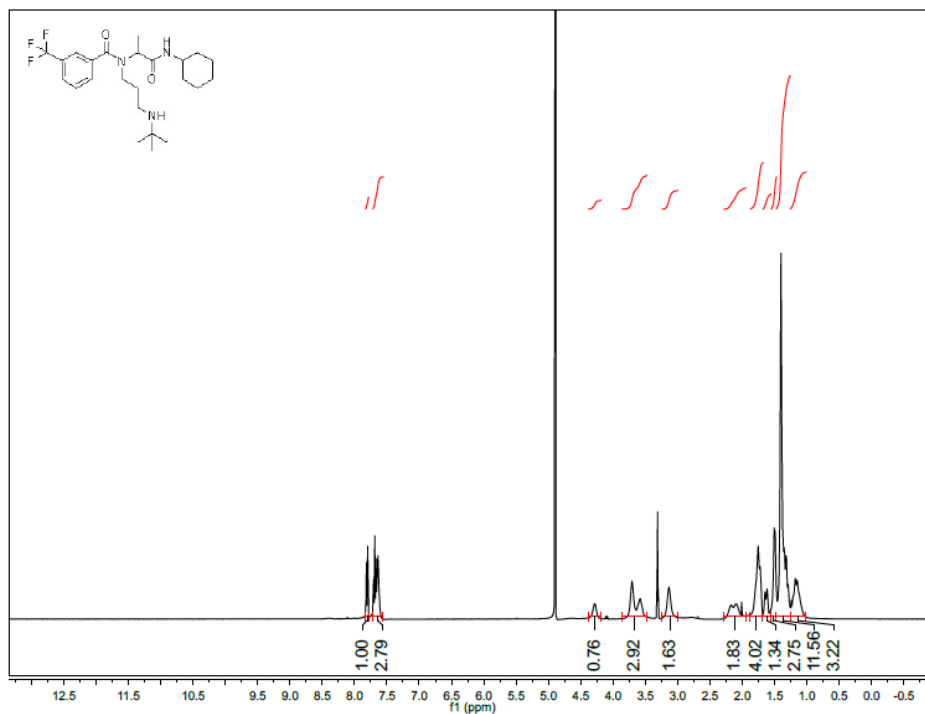
3-bromo-*N*-(3-(tert-butylamino)propyl)-*N*-(1-(cyclohexylamino)-1-oxopropan-2-yl)benzamide (132)



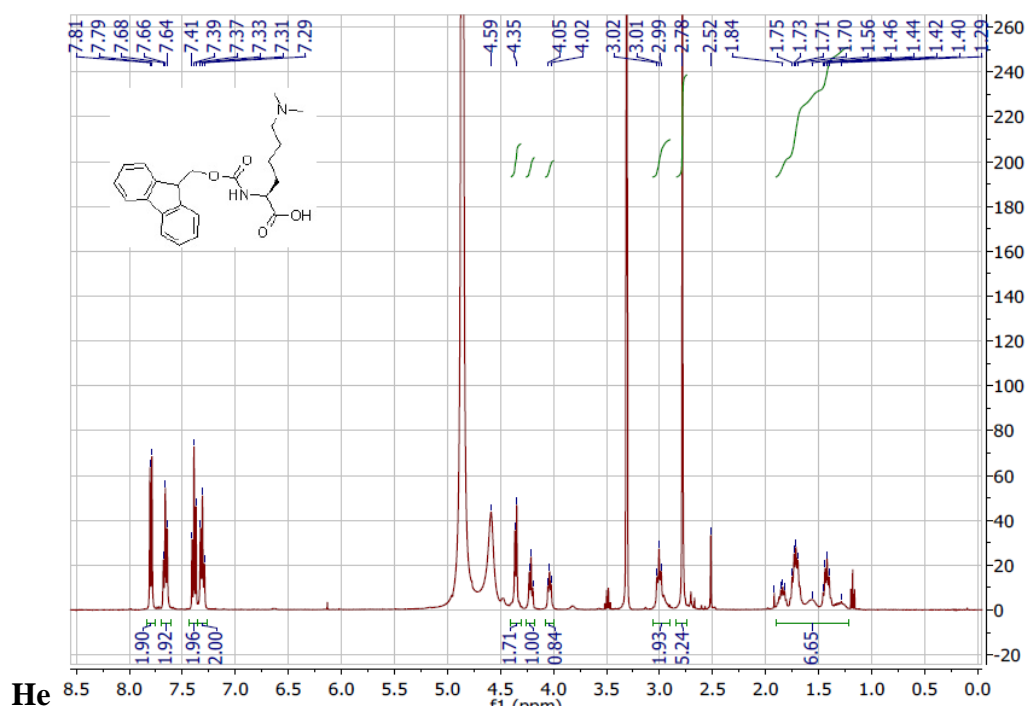
***N*-(3-(tert-butylamino)propyl)-*N*-(1-(cyclohexylamino)-1-oxopropan-2-yl)-3-iodobenzamidebenzamide (133)**



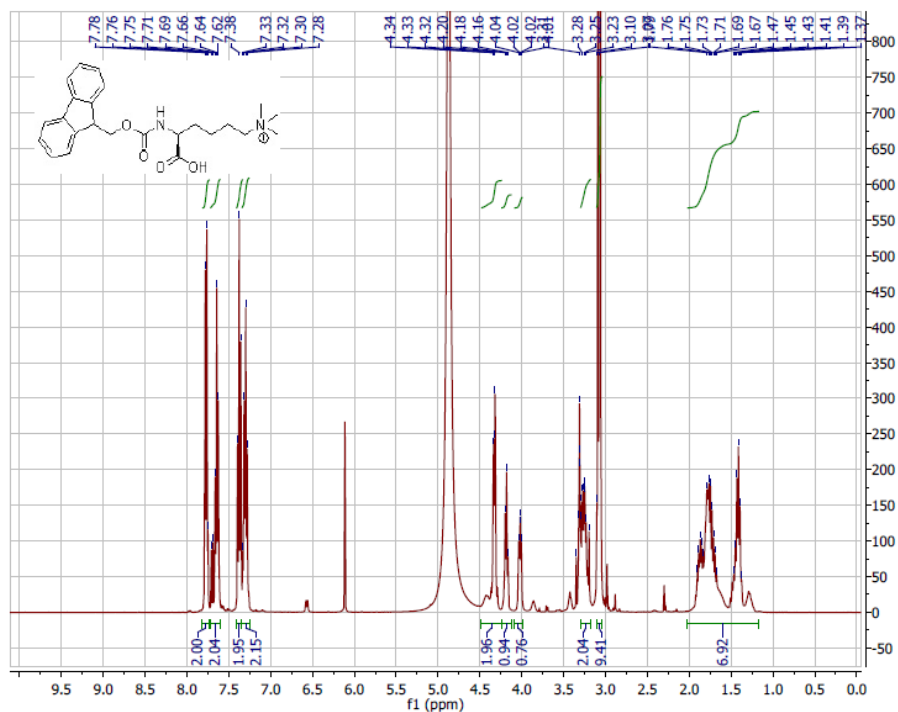
***N*-(3-(tert-butylamino)propyl)-*N*-(1-(cyclohexylamino)-1-oxopropan-2-yl)-3-(trifluoromethyl)benzamide (134)**



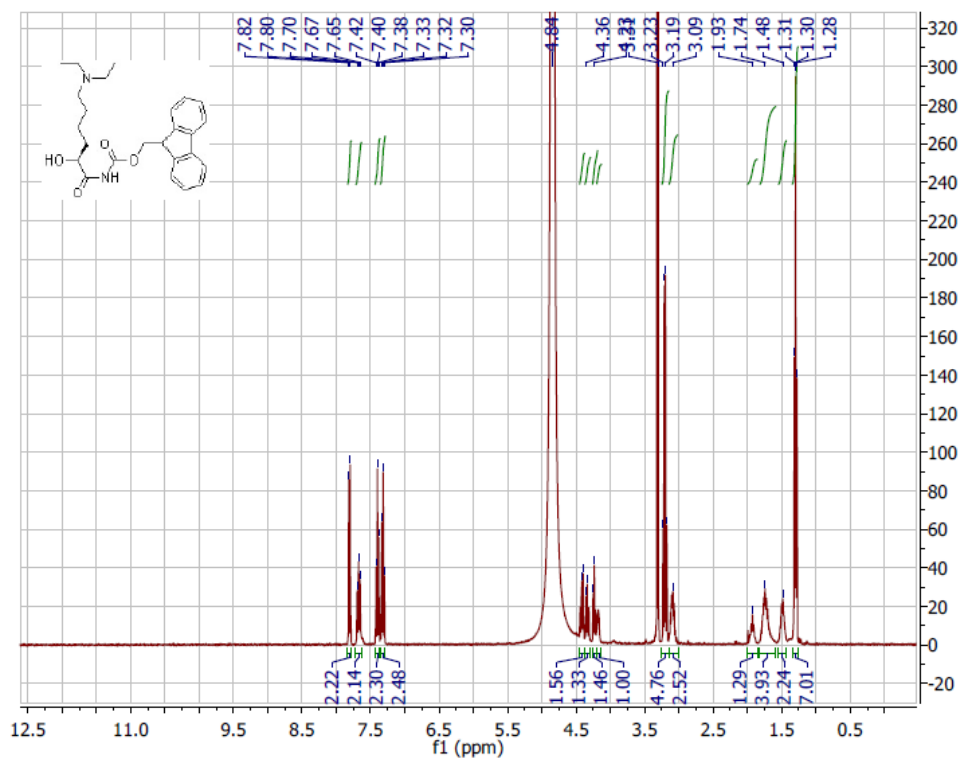
***N*²-(((9H-fluoren-9-yl)methoxy)carbonyl)-*N*⁶,*N*⁶-dimethyl-L-lysine TFA (139)**



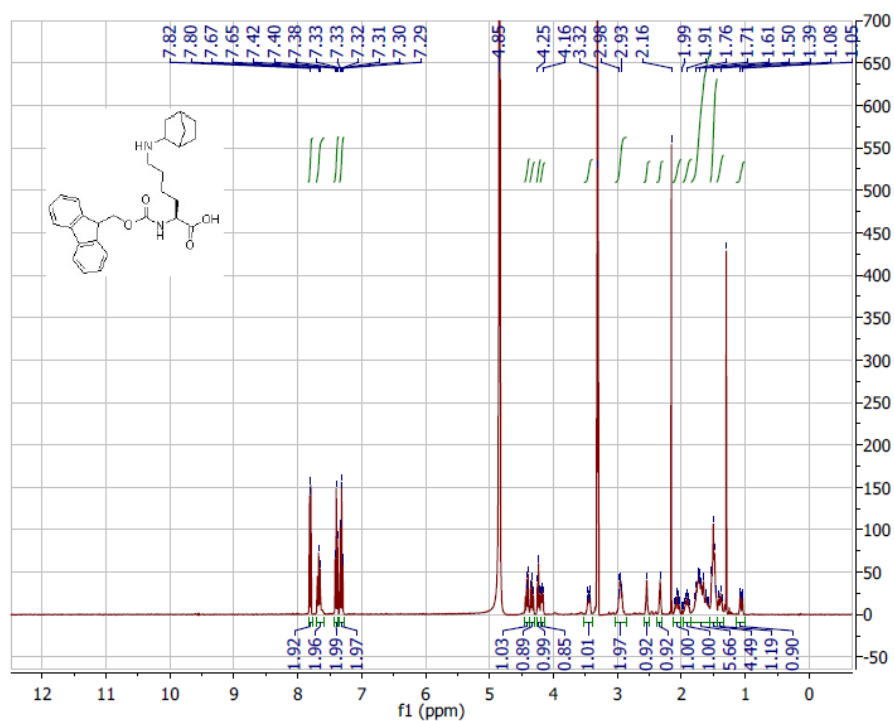
(S)-5-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-5-carboxy-N,N,N-trimethylpentan-1-aminium TFA (140)



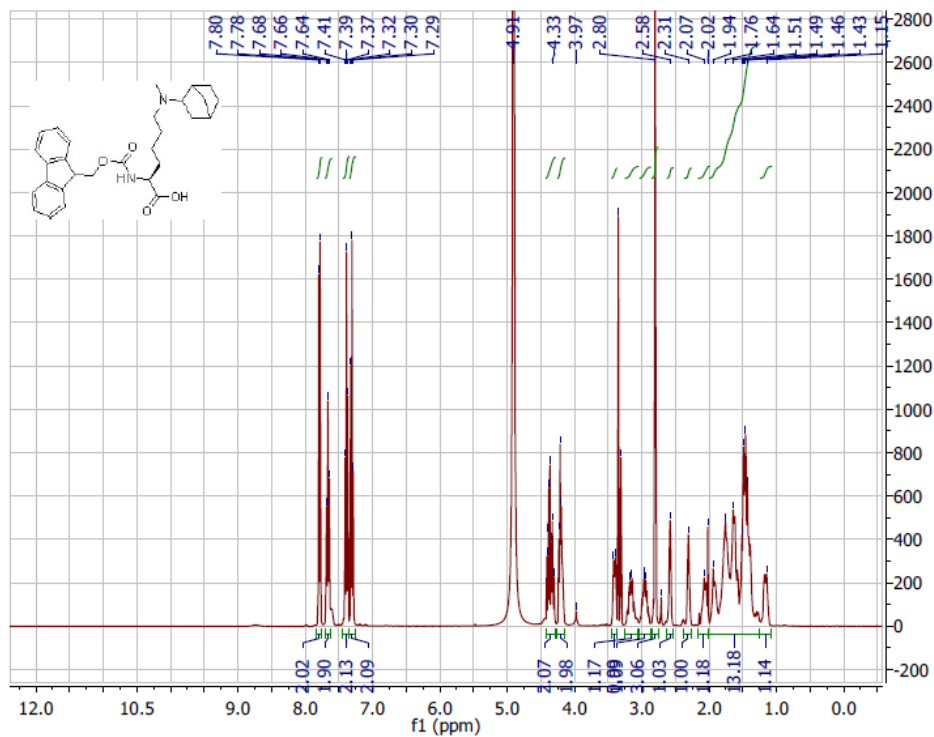
N²-((((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶,N⁶-diethyl-L-lysine TFA(141)



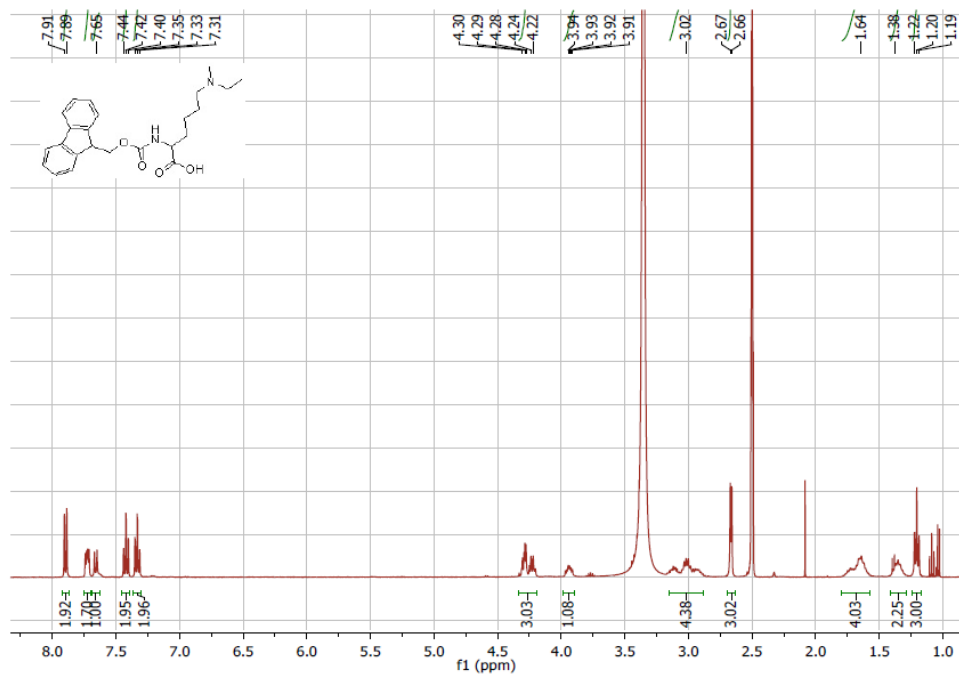
N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶-(bicyclo[2.2.1]heptan-2-yl)-L-lysine TFA(142):



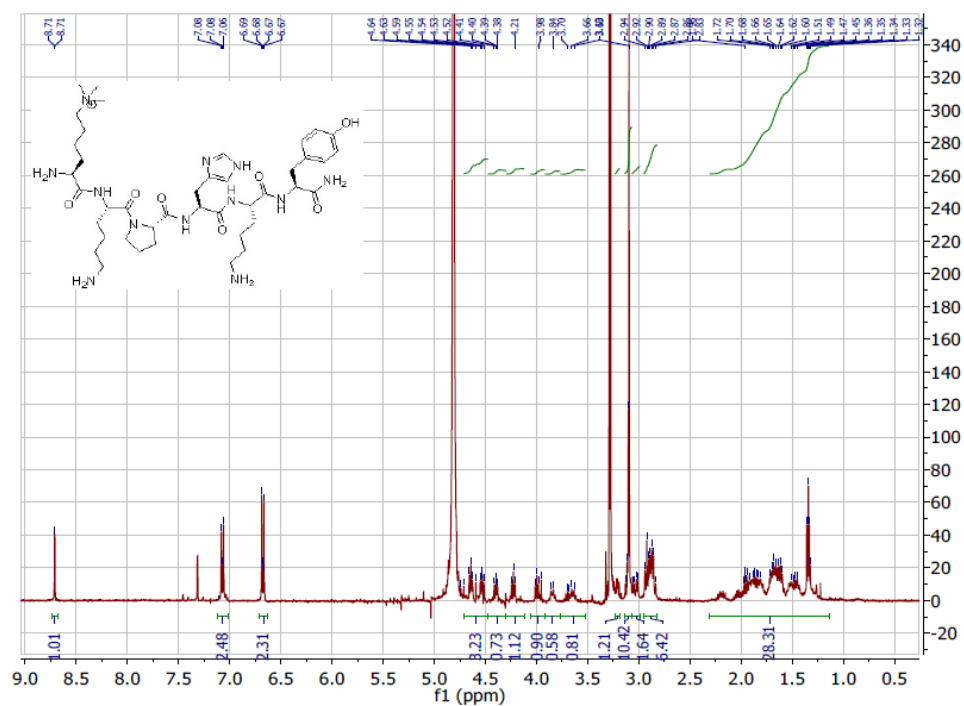
N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶-(bicyclo[2.2.1]heptan-2-yl)-N⁶-methyl-L-lysine TFA (143)



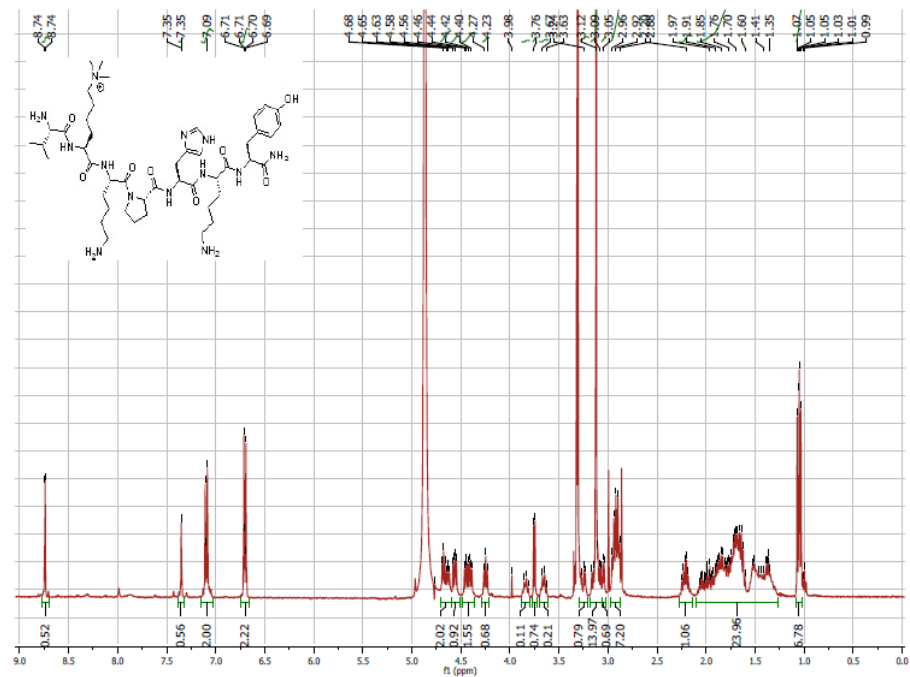
N²-(((9H-fluoren-9-yl)methoxy)carbonyl)-N⁶-ethyl-N⁶-methyl-L-lysine HCl (144)



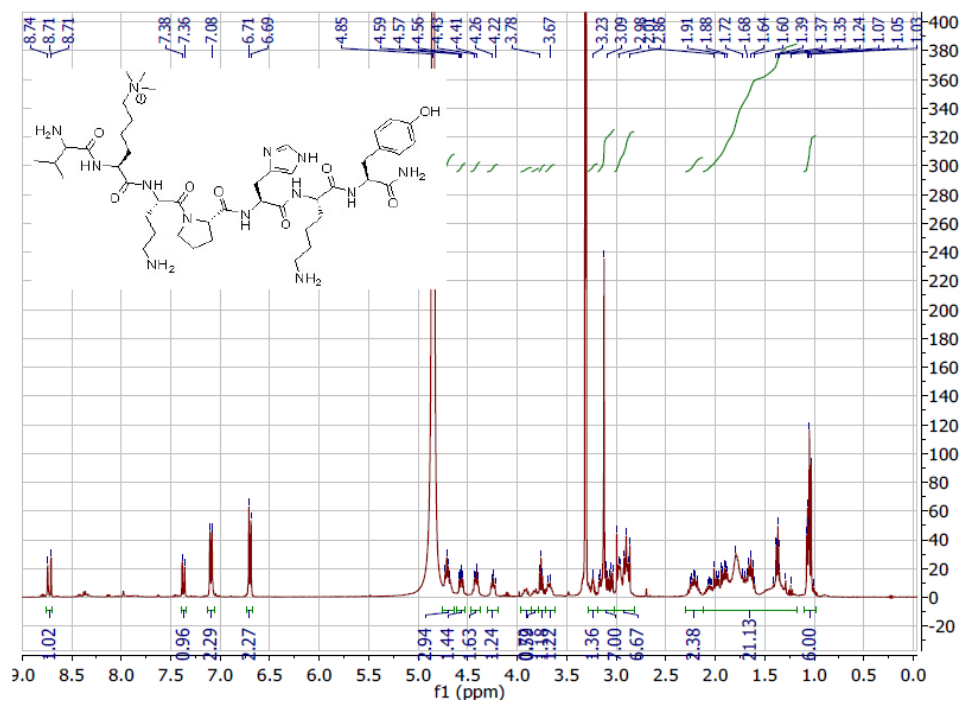
6-mer peptide (145)



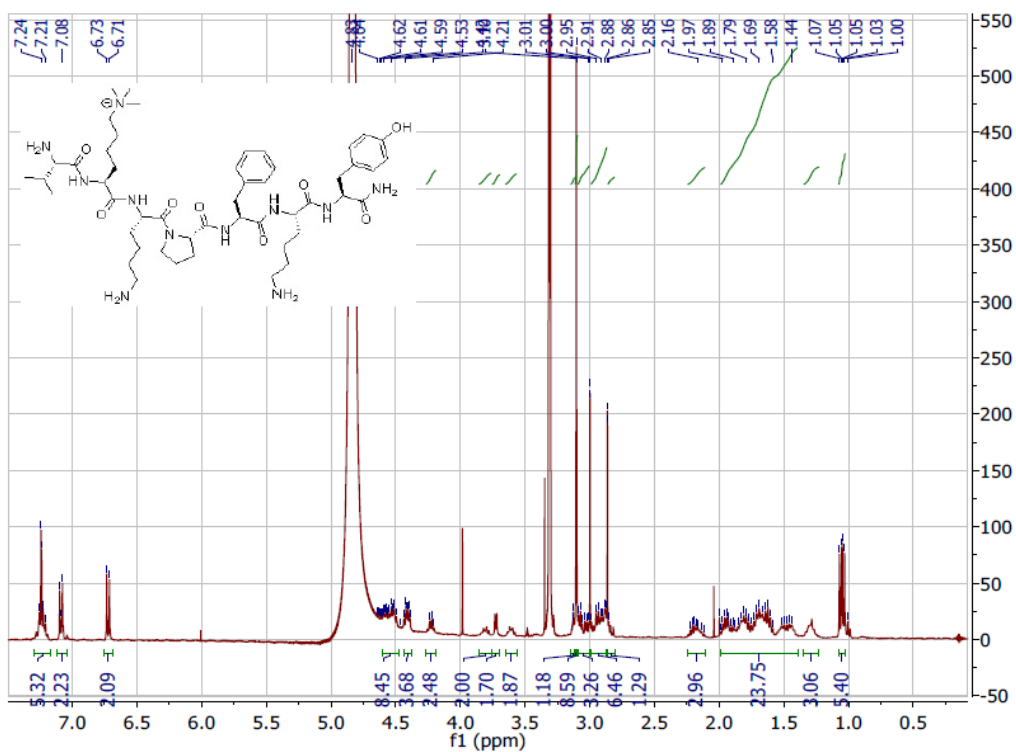
7-mer peptide (147)



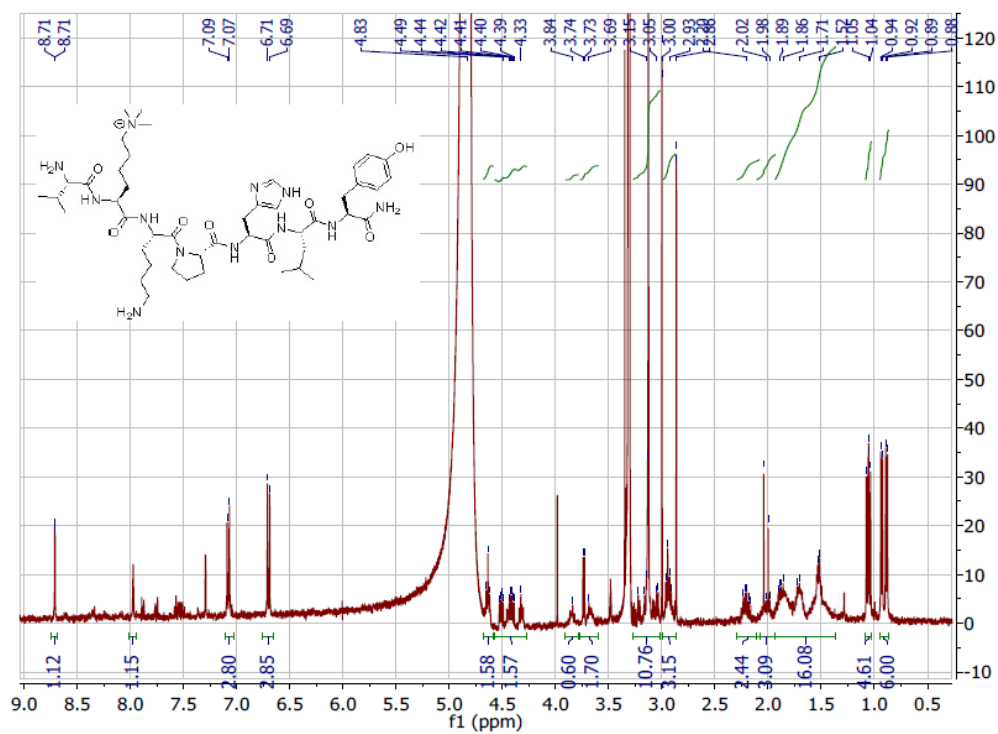
Orn peptide (148)



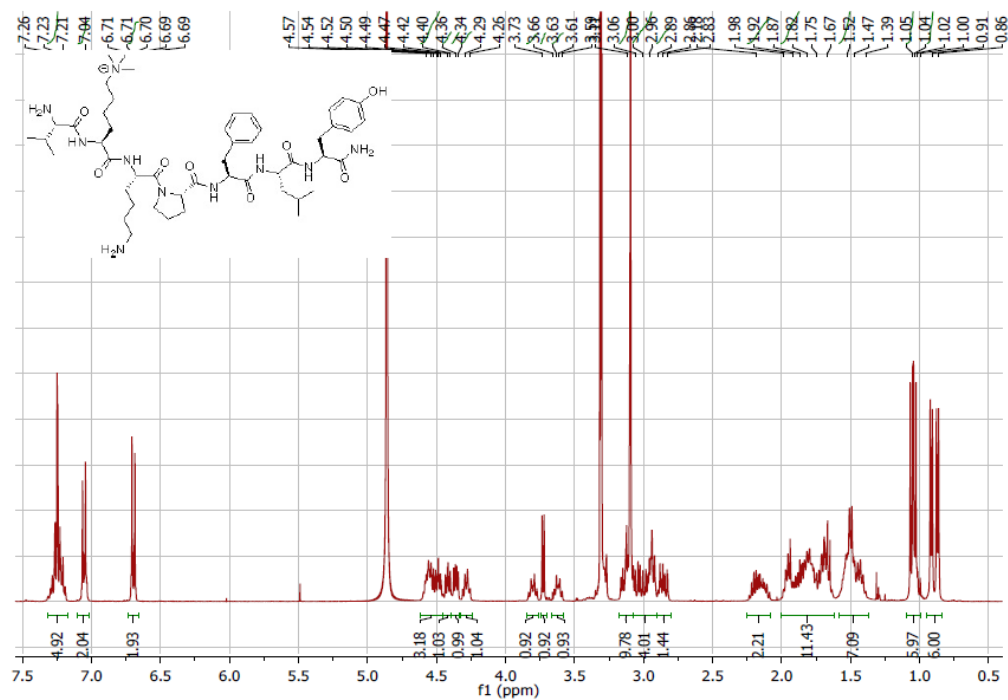
Phe analog (149)



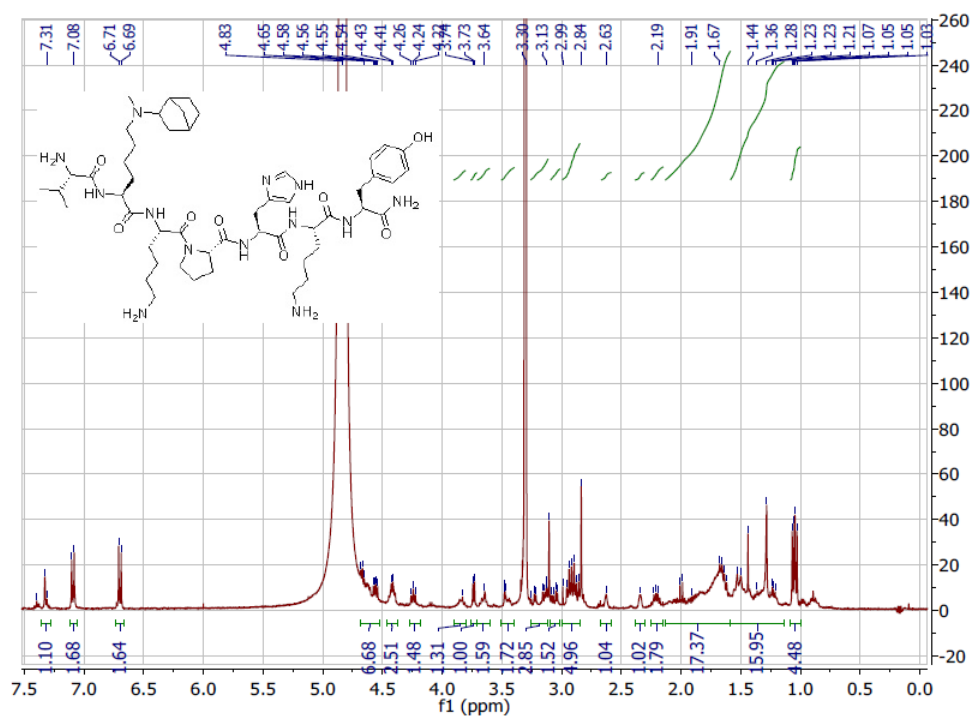
Leu Analog (150)



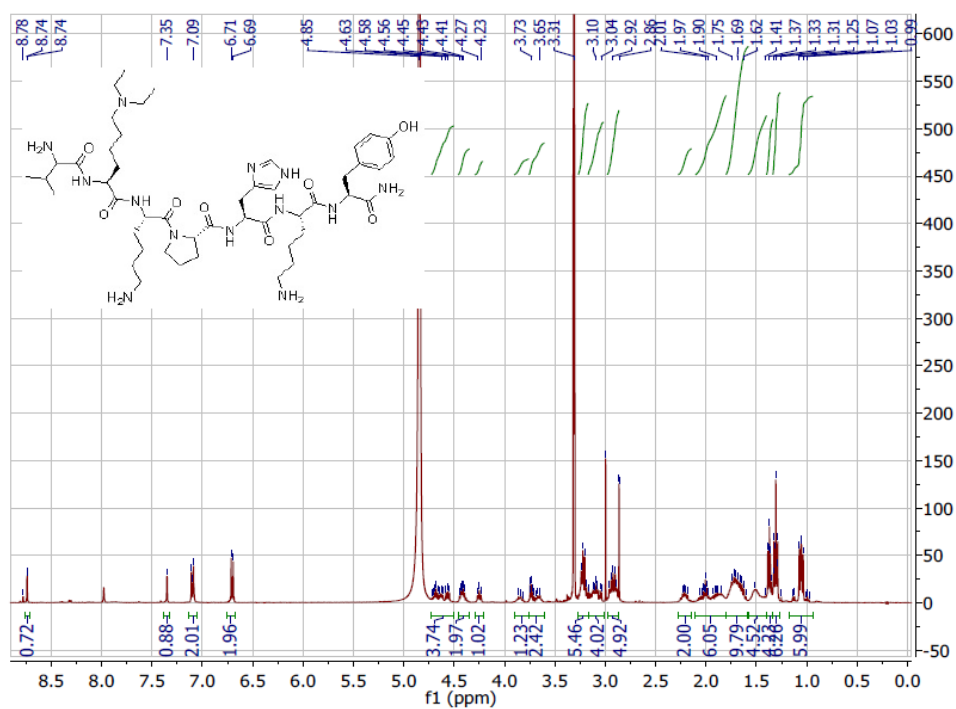
Phe + Leu analog (151)



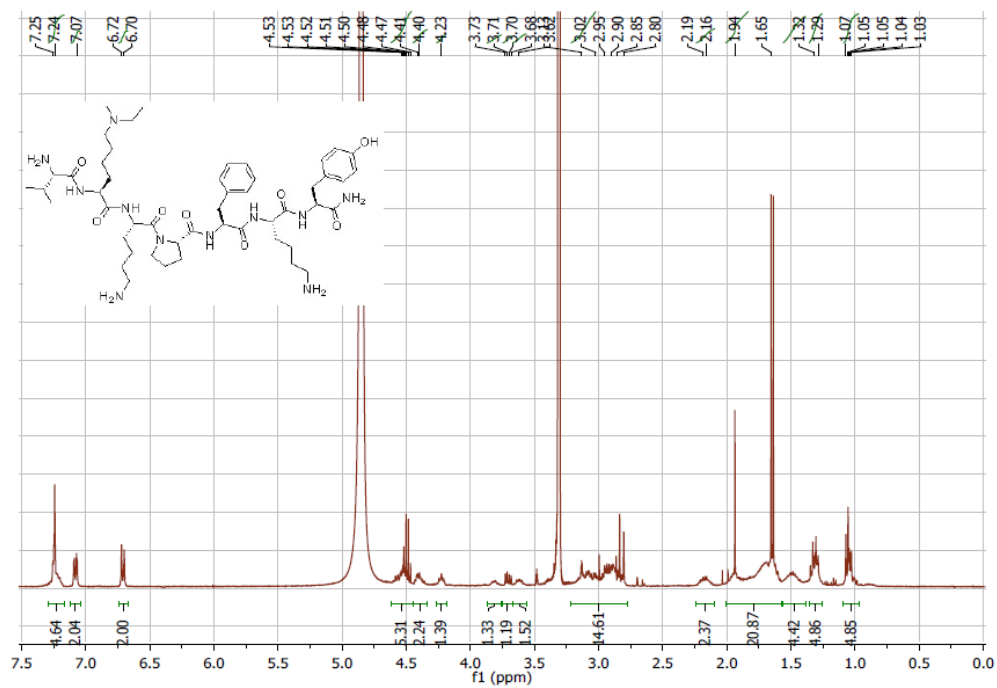
Norb, Meth analog (152)



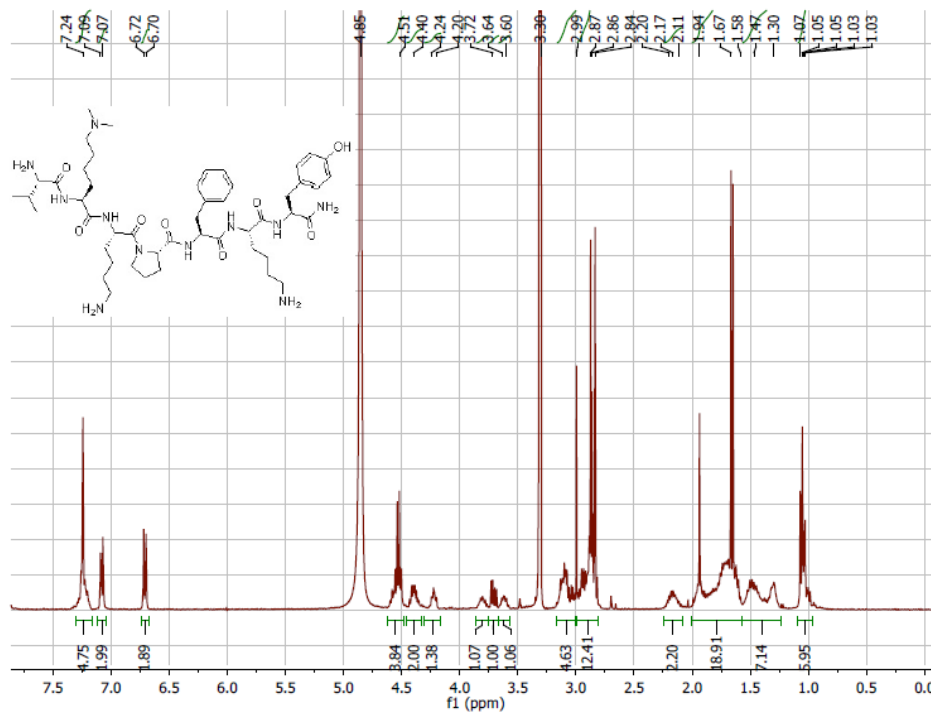
DiEthyl analog (153)



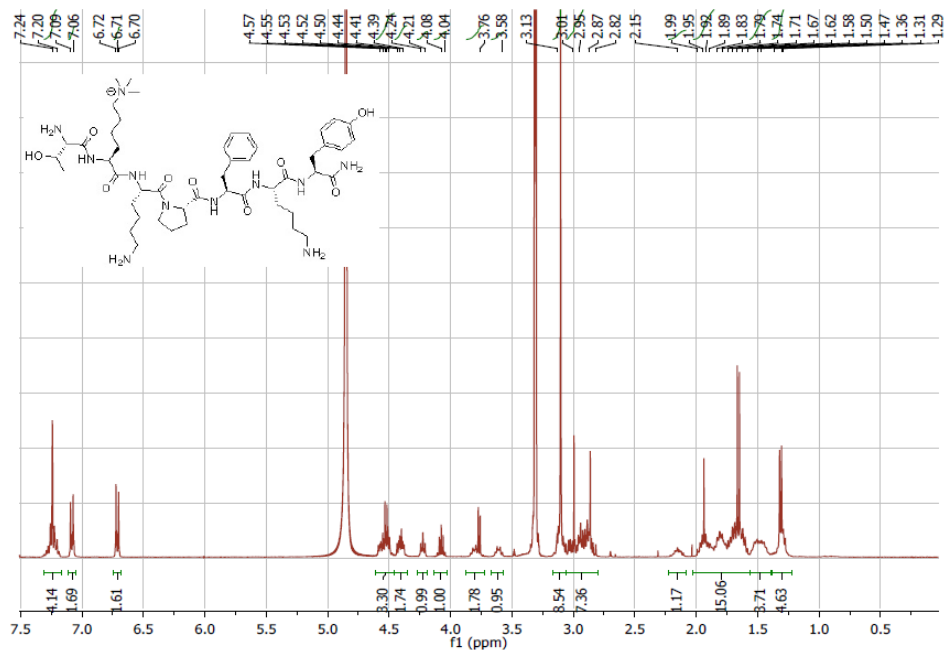
Methyl, Ethyl analog (154)



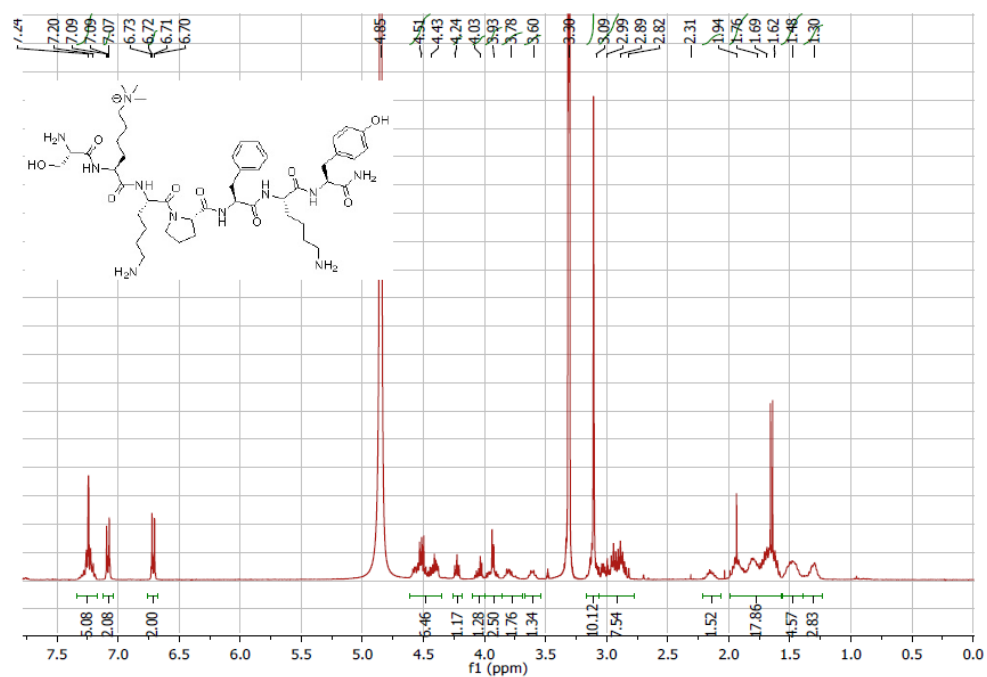
Dimethyl analog (155)



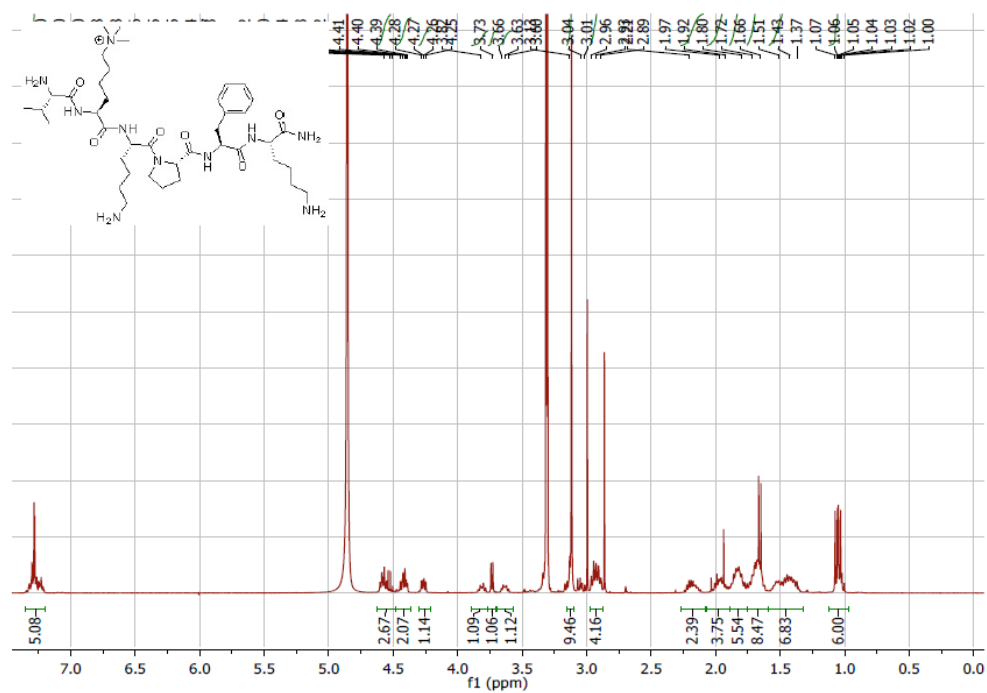
Thr analog (156)



Ser analog (157)



Tyr truncation (158)



FUNDING SOURCES

The research described here was supported by the National Institute of General Medical Sciences, US National Institutes of Health (NIH, grant R01GM100919), the Carolina Partnership and the University Cancer Research Fund, University of North Carolina at Chapel Hill, and the Welch Foundation (G-1847). The SGC is a registered charity (no. 1097737) that receives funds from AbbVie, Boehringer Ingelheim, the Canada Foundation for Innovation (CFI), the Canadian Institutes of Health Research (CIHR), Genome Canada, Ontario Genomics Institute Grant OGI-055, GlaxoSmithKline, Janssen, Lilly Canada, the Novartis Research Foundation, the Ontario Ministry of Economic Development and Innovation, Pfizer, Takeda, and Wellcome Trust Grant 092809/Z/10/Z. G.M. acknowledges support from NIH grant R01 CA132878 and Mayo Clinic NCI SPORE programs P50 CA116201 and P50 CA108961. G.C. is supported by a Mayo Clinic Cancer Center Fraternal Order of Eagles Fellowship. B.D.S. acknowledges support from the W.M. Keck Foundation. S.B.R. acknowledges support from NIH grant K99CA181343.

REFERENCES

1. Quina, A.S., M. Buschbeck, and L. Di Croce, *Chromatin structure and epigenetics*. Biochem Pharmacol, 2006. **72**(11): p. 1563-9.
2. Bannister, A.J. and T. Kouzarides, *Regulation of chromatin by histone modifications*. Cell Res, 2011. **21**(3): p. 381-95.
3. Li, G. and D. Reinberg, *Chromatin higher-order structures and gene regulation*. Curr Opin Genet Dev, 2011. **21**(2): p. 175-86.
4. Deaton, A.M. and A. Bird, *CpG islands and the regulation of transcription*. Genes Dev, 2011. **25**(10): p. 1010-22.
5. Jones, P.A., *Functions of DNA methylation: islands, start sites, gene bodies and beyond*. Nat Rev Genet, 2012. **13**(7): p. 484-92.
6. Margueron, R. and D. Reinberg, *Chromatin structure and the inheritance of epigenetic information*. Nat Rev Genet, 2010. **11**(4): p. 285-96.
7. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications*. Nature, 2000. **403**(6765): p. 41-5.
8. Rothbart, S.B. and B.D. Strahl, *Interpreting the language of histone and DNA modifications*. Biochim Biophys Acta, 2014. **1839**(8): p. 627-43.
9. Virani, S., et al., *Cancer epigenetics: a brief review*. ILAR J, 2012. **53**(3-4): p. 359-69.
10. Ptashne, M., *Epigenetics: core misconception*. Proc Natl Acad Sci U S A, 2013. **110**(18): p. 7101-3.
11. Burnett, G. and E.P. Kennedy, *The enzymatic phosphorylation of proteins*. J Biol Chem, 1954. **211**(2): p. 969-80.
12. Cori, G.T.G., A.A., J. Biol. Chem., 1943. **151**: p. 31-38.
13. Cheng, X., *Structure and function of DNA methyltransferases*. Annu Rev Biophys Biomol Struct, 1995. **24**: p. 293-318.
14. Allfrey, V.G., R. Faulkner, and A.E. Mirsky, *Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis*. Proc Natl Acad Sci U S A, 1964. **51**: p. 786-94.
15. Kuo, M.H. and C.D. Allis, *Roles of histone acetyltransferases and deacetylases in gene regulation*. Bioessays, 1998. **20**(8): p. 615-26.

16. Yang, X.J. and E. Seto, *HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention*. *Oncogene*, 2007. **26**(37): p. 5310-8.
17. Brownell, J.E., et al., *Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation*. *Cell*, 1996. **84**(6): p. 843-51.
18. Taunton, J., C.A. Hassig, and S.L. Schreiber, *A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p*. *Science*, 1996. **272**(5260): p. 408-11.
19. Lachner, M., et al., *Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins*. *Nature*, 2001. **410**(6824): p. 116-20.
20. Peters, A.H., et al., *Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability*. *Cell*, 2001. **107**(3): p. 323-37.
21. Kouzarides, T., *Chromatin modifications and their function*. *Cell*, 2007. **128**(4): p. 693-705.
22. Arrowsmith, C.H., et al., *Epigenetic protein families: a new frontier for drug discovery*. *Nat Rev Drug Discov*, 2012. **11**(5): p. 384-400.
23. Jenuwein, T. and C.D. Allis, *Translating the histone code*. *Science*, 2001. **293**(5532): p. 1074-80.
24. Ballare, C., et al., *Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity*. *Nat Struct Mol Biol*, 2012. **19**(12): p. 1257-65.
25. Cai, L., et al., *An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting*. *Mol Cell*, 2013. **49**(3): p. 571-82.
26. Musselman, C.A., et al., *Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1*. *Nat Struct Mol Biol*, 2012. **19**(12): p. 1266-72.
27. Jackson, S.P. and J. Bartek, *The DNA-damage response in human biology and disease*. *Nature*, 2009. **461**(7267): p. 1071-8.
28. Kelly, T.K., D.D. De Carvalho, and P.A. Jones, *Epigenetic modifications as therapeutic targets*. *Nat Biotechnol*, 2010. **28**(10): p. 1069-1078.
29. Portela, A. and M. Esteller, *Epigenetic modifications and human disease*. *Nat Biotechnol*, 2010. **28**(10): p. 1057-68.
30. Wagner, T., et al., *Mind the methyl: methyllysine binding proteins in epigenetic regulation*. *ChemMedChem*, 2014. **9**(3): p. 466-83.
31. Ruthenburg, A.J., et al., *Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions*. *Cell*, 2011. **145**(5): p. 692-706.

32. Li, H., et al., *Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF*. Nature, 2006. **442**(7098): p. 91-5.
33. Wysocka, J., et al., *A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling*. Nature, 2006. **442**(7098): p. 86-90.
34. Rothbart, S.B., et al., *Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation*. Genes & Development, 2013. **27**(11): p. 1288-1298.
35. Nady, N., et al., *Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein*. J Biol Chem, 2011. **286**(27): p. 24300-11.
36. Botuyan, M.V., et al., *Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair*. Cell, 2006. **127**(7): p. 1361-73.
37. Lu, R. and G.G. Wang, *Tudor: a versatile family of histone methylation 'readers'*. Trends Biochem Sci, 2013. **38**(11): p. 546-55.
38. Micci, F., et al., *Consistent rearrangement of chromosomal band 6p21 with generation of fusion genes JAZF1/PHF1 and EPC1/PHF1 in endometrial stromal sarcoma*. Cancer Res, 2006. **66**(1): p. 107-12.
39. Wang, S., G.P. Robertson, and J. Zhu, *A novel human homologue of Drosophila polycomblike gene is up-regulated in multiple cancers*. Gene, 2004. **343**(1): p. 69-78.
40. Lowndes, N.F., *The interplay between BRCA1 and 53BP1 influences death, aging, senescence and cancer*. DNA Repair (Amst), 2010. **9**(10): p. 1112-6.
41. Brueckner, B., D. Kuck, and F. Lyko, *DNA methyltransferase inhibitors for cancer therapy*. Cancer J, 2007. **13**(1): p. 17-22.
42. Singh, V., P. Sharma, and N. Capalash, *DNA methyltransferase-1 inhibitors as epigenetic therapy for cancer*. Curr Cancer Drug Targets, 2013. **13**(4): p. 379-99.
43. Epizyme, I., *Epizyme DOT1L Inhibitor EPZ-5676 Shows Clinical and Biological Activity in Adult Patients with Acute Leukemias in Phase 1 Trial*. 2015.
44. Epizyme, I., *Epizyme Presents Pre-Clinical Data and Early Clinical Observations from Ongoing Phase 1 Trial of EZH2 Inhibitor EPZ-6438 (E7438) at ASH Meeting on Lymphoma Biology*. 2015.
45. Bernt, K.M., et al., *MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L*. Cancer Cell, 2011. **20**(1): p. 66-78.
46. Daigle, S.R., et al., *Potent inhibition of DOT1L as treatment of MLL-fusion leukemia*. Blood, 2013. **122**(6): p. 1017-25.

47. Sneeringer, C.J., et al., *Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas*. Proc Natl Acad Sci U S A, 2010. **107**(49): p. 20980-5.
48. Wigle, T.J., et al., *The Y641C mutation of EZH2 alters substrate specificity for histone H3 lysine 27 methylation states*. FEBS Lett, 2011. **585**(19): p. 3011-4.
49. *Merck acquires Swiss oncology drug development company OncoEthix*. 2014.
50. Delmore, J.E., et al., *BET bromodomain inhibition as a therapeutic strategy to target c-Myc*. Cell, 2011. **146**(6): p. 904-17.
51. Filippakopoulos, P., et al., *Selective inhibition of BET bromodomains*. Nature, 2010. **468**(7327): p. 1067-73.
52. Wigle, T.J., et al., *Screening for Inhibitors of Low-Affinity Epigenetic Peptide-Protein Interactions: An AlphaScreen (TM)-Based Assay for Antagonists of Methyl-Lysine Binding Proteins*. Journal of Biomolecular Screening, 2010. **15**(1): p. 62-71.
53. Herold, J.M., et al., *Small-molecule ligands of methyl-lysine binding proteins*. J Med Chem, 2011. **54**(7): p. 2504-11.
54. James, L.I., et al., *Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain*. Nature Chemical Biology, 2013. **9**(3): p. 184-191.
55. James, L.I., et al., *Small-Molecule Ligands of Methyl-Lysine Binding Proteins: Optimization of Selectivity for L3MBTL3*. Journal of Medicinal Chemistry, 2013. **56**(18): p. 7358-7371.
56. Frye, S.V., *The art of the chemical probe*. Nat Chem Biol, 2010. **6**(3): p. 159-161.
57. Simhadri, C., et al., *Chromodomain Antagonists That Target the Polycomb-Group Methyllysine Reader Protein Chromobox Homo log 7 (CBX7)*. Journal of Medicinal Chemistry, 2014. **57**(7): p. 2874-2883.
58. Ren, C., et al., *Small-Molecule Modulators of Methyl-Lysine Binding for the CBX7 Chromodomain*. Chem Biol, 2015. **22**(2): p. 161-8.
59. Bolshan, Y., et al., *Synthesis, Optimization, and Evaluation of Novel Small Molecules as Antagonists of WDR5-MLL Interaction*. ACS Med Chem Lett, 2013. **4**(3): p. 353-7.
60. Senisterra, G., et al., *Small-molecule inhibition of MLL activity by disruption of its interaction with WDR5*. Biochem J, 2013. **449**(1): p. 151-9.
61. Wagner, E.K., et al., *Identification and Characterization of Small Molecule Inhibitors of a Plant Homeodomain Finger*. Biochemistry, 2012. **51**(41): p. 8293-8306.

62. Da, C. and D. Kireev, *Structural protein-ligand interaction fingerprints (SPLIF) for structure-based virtual screening: method and benchmark study*. J Chem Inf Model, 2014. **54**(9): p. 2555-61.
63. Kireev, D., et al., *Identification of non-peptide malignant brain tumor (MBT) repeat antagonists by virtual screening of commercially available compounds*. J Med Chem, 2010. **53**(21): p. 7625-31.
64. Iwabuchi, K., et al., *Potential role for 53BP1 in DNA end-joining repair through direct interaction with DNA*. J Biol Chem, 2003. **278**(38): p. 36487-95.
65. Kachirskaja, I., et al., *Role for 53BP1 Tudor domain recognition of p53 dimethylated at lysine 382 in DNA damage signaling*. J Biol Chem, 2008. **283**(50): p. 34660-6.
66. Botuyan, M.V., et al., *Structural Basis for the Methylation State-Specific Recognition of Histone H4-K20 by 53BP1 and Crb2 in DNA Repair*. Cell, 2006. **127**(7): p. 1361-1373.
67. Kim, J., et al., *Tudor, MBT and chromo domains gauge the degree of lysine methylation*. EMBO Rep, 2006. **7**(4): p. 397-403.
68. Zgheib, O., et al., *An oligomerized 53BP1 tudor domain suffices for recognition of DNA double-strand breaks*. Mol Cell Biol, 2009. **29**(4): p. 1050-8.
69. Mochan, T.A., et al., *53BP1 and NFB1/MDC1-Nbs1 function in parallel interacting pathways activating ataxia-telangiectasia mutated (ATM) in response to DNA damage*. Cancer Res, 2003. **63**(24): p. 8586-91.
70. Shibata, A., et al., *Role of ATM and the damage response mediator proteins 53BP1 and MDC1 in the maintenance of G(2)/M checkpoint arrest*. Mol Cell Biol, 2010. **30**(13): p. 3371-83.
71. Stewart, G.S., et al., *MDC1 is a mediator of the mammalian DNA damage checkpoint*. Nature, 2003. **421**(6926): p. 961-6.
72. Di Virgilio, M., et al., *Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching*. Science, 2013. **339**(6120): p. 711-5.
73. Escribano-Diaz, C., et al., *A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice*. Mol Cell, 2013. **49**(5): p. 872-83.
74. Sartori, A.A., et al., *Human CtIP promotes DNA end resection*. Nature, 2007. **450**(7169): p. 509-14.
75. Kousholt, A.N., et al., *CtIP-dependent DNA resection is required for DNA damage checkpoint maintenance but not initiation*. J Cell Biol, 2012. **197**(7): p. 869-76.
76. Zimmermann, M., et al., *53BP1 regulates DSB repair using Rif1 to control 5' end resection*. Science, 2013. **339**(6120): p. 700-4.

77. Bekker-Jensen, S., et al., *Dynamic assembly and sustained retention of 53BP1 at the sites of DNA damage are controlled by Mdc1/NFBD1*. J Cell Biol, 2005. **170**(2): p. 201-11.
78. Huen, M.S., et al., *RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly*. Cell, 2007. **131**(5): p. 901-14.
79. Doil, C., et al., *RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins*. Cell, 2009. **136**(3): p. 435-46.
80. Ramachandran, S., et al., *The RNF8/RNF168 ubiquitin ligase cascade facilitates class switch recombination*. Proc Natl Acad Sci U S A, 2010. **107**(2): p. 809-14.
81. Chowdhury, D., et al., *gamma-H2AX dephosphorylation by protein phosphatase 2A facilitates DNA double-strand break repair*. Mol Cell, 2005. **20**(5): p. 801-9.
82. Joo, W.S., et al., *Structure of the 53BP1 BRCT region bound to p53 and its comparison to the Brca1 BRCT structure*. Genes Dev, 2002. **16**(5): p. 583-93.
83. Fradet-Turcotte, A., et al., *53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark*. Nature, 2013. **499**(7456): p. 50-4.
84. Acs, K., et al., *The AAA-ATPase VCP/p97 promotes 53BP1 recruitment by removing L3MBTL1 from DNA double-strand breaks*. Nat Struct Mol Biol, 2011. **18**(12): p. 1345-50.
85. Mallette, F.A., et al., *RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites*. EMBO J, 2012.
86. Ward, I.M., et al., *p53 Binding protein 53BP1 is required for DNA damage responses and tumor suppression in mice*. Mol Cell Biol, 2003. **23**(7): p. 2556-63.
87. Narod, S.A., *BRCA mutations in the management of breast cancer: the state of the art*. Nat Rev Clin Oncol, 2010. **7**(12): p. 702-7.
88. Manis, J.P., et al., *53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination*. Nat Immunol, 2004. **5**(5): p. 481-7.
89. Shiloh, Y., *ATM and related protein kinases: safeguarding genome integrity*. Nat Rev Cancer, 2003. **3**(3): p. 155-68.
90. Lieber, M.R., et al., *Mechanism and regulation of human non-homologous DNA end-joining*. Nat Rev Mol Cell Biol, 2003. **4**(9): p. 712-20.
91. Li, X. and W.D. Heyer, *Homologous recombination in DNA repair and DNA damage tolerance*. Cell Res, 2008. **18**(1): p. 99-113.

92. Giunta, S., R. Belotserkovskaya, and S.P. Jackson, *DNA damage signaling in response to double-strand breaks during mitosis*. J Cell Biol, 2010. **190**(2): p. 197-207.
93. Bunting, S.F., et al., *53BP1 Inhibits Homologous Recombination in Brca1-Deficient Cells by Blocking Resection of DNA Breaks*. Cell, 2010. **141**(2): p. 243-254.
94. Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection*. Nat Rev Cancer, 2012. **12**(1): p. 68-78.
95. Hakem, R., et al., *The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse*. Cell, 1996. **85**(7): p. 1009-23.
96. Xu, X., et al., *Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation*. Nat Genet, 1999. **22**(1): p. 37-43.
97. Cao, L., et al., *A selective requirement for 53BP1 in the biological response to genomic instability induced by Brca1 deficiency*. Mol Cell, 2009. **35**(4): p. 534-41.
98. Bouwman, P., et al., *53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers*. Nat Struct Mol Biol, 2010. **17**(6): p. 688-95.
99. Aly, A. and S. Ganesan, *BRCA1, PARP, and 53BP1: conditional synthetic lethality and synthetic viability*. J Mol Cell Biol, 2011. **3**(1): p. 66-74.
100. Tutt, A.e.a., *Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial*. Lancet, 2010. **376**: p. 235-244.
101. Cui, G., M.V. Botuyan, and G. Mer, *Preparation of recombinant peptides with site- and degree-specific lysine (13)C-methylation*. Biochemistry, 2009. **48**(18): p. 3798-800.
102. Tang, J., et al., *Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination*. Nat Struct Mol Biol, 2013. **20**(3): p. 317-25.
103. Francke, A., et al., *Generation of mature murine monocytes from heterogeneous bone marrow and description of their properties*. J Histochem Cytochem, 2011. **59**(9): p. 813-25.
104. Janzen, W.P., et al., *Epigenetics: Tools and Technologies*. Drug Discov Today Technol, 2010. **7**(1): p. e59-e65.
105. Wigle, T.J., et al., *Screening for inhibitors of low-affinity epigenetic peptide-protein interactions: an AlphaScreen-based assay for antagonists of methyl-lysine binding proteins*. J Biomol Screen, 2010. **15**(1): p. 62-71.
106. Shultz, M.D., *Setting expectations in molecular optimizations: Strengths and limitations of commonly used composite parameters*. Bioorg Med Chem Lett, 2013. **23**(21): p. 5980-91.

107. Reynolds, C.H., B.A. Tounge, and S.D. Bembenek, *Ligand binding efficiency: trends, physical basis, and implications*. J Med Chem, 2008. **51**(8): p. 2432-8.
108. James, L.I., et al., *Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain*. Nat Chem Biol, 2013. **9**(3): p. 184-91.
109. Herold, J.M., et al., *Structure-activity relationships of methyl-lysine reader antagonists*. Medchemcomm, 2012. **3**(1): p. 45-51.
110. Wilcken, R., et al., *Using halogen bonds to address the protein backbone: a systematic evaluation*. J Comput Aided Mol Des, 2012. **26**(8): p. 935-45.
111. Certo, M.T., et al., *Tracking genome engineering outcome at individual DNA breakpoints*. Nat Methods, 2011. **8**(8): p. 671-6.
112. Bothmer, A., et al., *Regulation of DNA end joining, resection, and immunoglobulin class switch recombination by 53BP1*. Mol Cell, 2011. **42**(3): p. 319-29.
113. Ward, I.M., et al., *53BP1 is required for class switch recombination*. J Cell Biol, 2004. **165**(4): p. 459-64.
114. Choi, S., et al., *Quantitative proteomics reveal ATM kinase-dependent exchange in DNA damage response complexes*. J Proteome Res, 2012. **11**(10): p. 4983-91.
115. Pryde, F., et al., *53BP1 exchanges slowly at the sites of DNA damage and appears to require RNA for its association with chromatin*. J Cell Sci, 2005. **118**(Pt 9): p. 2043-55.
116. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J Biomol NMR, 1995. **6**(3): p. 277-93.
117. Johnson, B.A. and R.A. Blevins, *NMR View: A computer program for the visualization and analysis of NMR data*. J Biomol NMR, 1994. **4**(5): p. 603-14.
118. Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode*. Macromolecular Crystallography, Pt A, 1997. **276**: p. 307-326.
119. McCoy, A.J., et al., *Phaser crystallographic software*. J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
120. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method*. Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.
121. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.

122. Gazumyan, A., et al., *Amino-terminal phosphorylation of activation-induced cytidine deaminase suppresses c-myc/IgH translocation*. Mol Cell Biol, 2011. **31**(3): p. 442-9.
123. Blundell, T.L., H. Jhoti, and C. Abell, *High-throughput crystallography for lead discovery in drug design*. Nat Rev Drug Discov, 2002. **1**(1): p. 45-54.
124. Davies, T.G. and I.J. Tickle, *Fragment screening using X-ray crystallography*. Top Curr Chem, 2012. **317**: p. 33-59.
125. Jhoti, H., *High-throughput X-ray techniques and drug discovery*. Ernst Schering Res Found Workshop, 2003(42): p. 43-58.
126. Winter, A., et al., *Biophysical and computational fragment-based approaches to targeting protein-protein interactions: applications in structure-guided drug discovery*. Q Rev Biophys, 2012. **45**(4): p. 383-426.
127. Mikkelsen, T.S., et al., *Genome-wide maps of chromatin state in pluripotent and lineage-committed cells*. Nature, 2007. **448**(7153): p. 553-60.
128. Margueron, R. and D. Reinberg, *The Polycomb complex PRC2 and its mark in life*. Nature, 2011. **469**(7330): p. 343-9.
129. Kycia, I., et al., *The Tudor domain of the PHD finger protein 1 is a dual reader of lysine trimethylation at lysine 36 of histone H3 and lysine 27 of histone variant H3t*. J Mol Biol, 2014. **426**(8): p. 1651-60.
130. Sarma, K., et al., *Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo*. Mol Cell Biol, 2008. **28**(8): p. 2718-31.
131. Cao, R. and Y. Zhang, *SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex*. Mol Cell, 2004. **15**(1): p. 57-67.
132. Margueron, R., et al., *Role of the polycomb protein EED in the propagation of repressive histone marks*. Nature, 2009. **461**(7265): p. 762-7.