Development of Ligands for the TRIM33 Bromodomain



Larissa See

Thesis submitted for the degree

Doctor of Philosophy in Organic Chemistry

St John's College, University of Oxford

Michaelmas Term 2018

Under the supervision of Prof. Stuart Conway

Department of Chemistry

DECLARATION OF AUTHORSHIP

Name: Larissa See **Candidate number: 838531**

College: St John's College **Supervisor: Prof. Stuart Conway**

Title of thesis:
Development of Ligands for the TRIM33 Bromodomain

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*The University Student Handbook* Section 8.8; available at https://www.ox.ac.uk/students/academic/student-handbook). ☑

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at https://www.ox.ac.uk/students/academic/guidance/skills?wssl=1. ☑

The thesis I am submitting is entirely my own work except where otherwise indicated. ☑

It has not been submitted, either partially or in full, either for this Honour School or qualification or for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution. ☑

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs. ☑

I have clearly indicated the presence of all paraphrased material with appropriate references. ☑

I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisor. ☑

I have not copied from the work of any other candidate. ☑

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: http://www.admin.ox.ac.uk/statutes/352-051a.shtml). ☑

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted. ☑

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism. ☑

Candidate's signature: …….. Date: ……..

Table of Contents

Abstract ..... **i**

Acknowledgements ..... **iii**

Abbreviations ..... **v**

Chapter 1: Introduction **1**

1.1 Epigenetics and the Histone Code 1

1.1.1 Multiple Epigenetic Domains 3

1.2 Bromodomains 4

1.3 Plant Homeodomain (PHD) Finger 6

1.4 Tripartite Motif-Containing (TRIM) Proteins 8

1.4.1 TRIM33 8

1.5 Small Molecule Ligands for Bromodomains and PHD Fingers 10

1.5.1 Bromodomain Ligands 10

1.5.2 PHD Finger 12

1.5.3 Covalent Inhibitors of Proteins 13

1.6 Assays 15

1.6.1 AlphaScreen™ 16

1.6.2 NMR WaterLOGSY 18

1.6.3 Isothermal Titration Calorimetry (ITC) 19

1.7 Previous Work and Aims 21

Chapter 2: Biological Characterisation **23**

2.1 Introduction and Aims 23

2.2 The AlphaScreen™ Assay 23
2.2.1 Expression and Purification of TRIM33β 23
2.2.2 Optimisation of AlphaScreen™ Assay Conditions 24
2.3 AlphaScreen™ as a Technique for Probing Individual Reader Domains 27
2.4 Investigation of Protein-Peptide Interactions by ITC 29
2.5 Conclusions 35
Chapter 3: Hit Validation 37
3.1 Introduction and Aims 37
3.2 Follow-up of PPI-Net Collection Hit Compounds 37
3.3 Follow-up of Maybridge Library Hit Compounds 41
3.3.1 Selection of Compounds 41
3.3.2 Screening of Focused Selection by AlphaScreen™ 42
3.4 Further Validation Using Complementary Experiments 43
3.4.1 AlphaScreen™ TruHits Kit 43
3.4.1.1 Screening Hits for Assay Interference 44
3.4.2 NMR WaterLOGSY as an Orthogonal Technique 46
3.4.2.1 Compound Screening by NMR WaterLOGSY 47
3.4.2.2 NMR WaterLOGSY Experiments with Mutant Protein 51
3.5 Conclusions 54
Chapter 4: Development of Benzimidazolone Series 57
4.1 Introduction and Aims 57
4.2 Establishing Structure-Activity Relationships 58
4.2.1 Preliminary Docking Studies 58

4.2.2 Establishing Structure-Activity Relationships 1 60
4.2.3 Establishing Structure-Activity Relationships 2 63
4.2.3.1 Synthesis of 2-Carbon Amine Analogue 63
4.2.3.2 Evaluation of Compounds 64
4.3 Variation of Amine Chain 67
4.3.1 Structure-Activity Relationship Studies 1 67
4.3.1.1 Synthesis of Compounds 67
4.3.1.2 Evaluation of Compounds 69
4.3.2 Validation of Compound Binding 72
4.3.3 Computational Studies 76
4.3.4 Structure-Activity Relationship Studies 2 78
4.3.4.1 Synthesis of Compounds 78
4.3.4.2 Evaluation of Compounds 79
4.4 Exploration of Hydrophobic Region 86
4.4.1 Computational Studies 86
4.4.2 Investigation of Optimal Vector 1 86
4.4.2.1 Synthesis of Compounds 86
4.4.2.2 Evaluation of Compounds 90
4.4.3 Investigation of Optimal Vector 2 92
4.4.3.1 Synthesis of Compounds 92
4.4.3.2 Evaluation of Compounds 93
4.4.4 Structure-Activity Relationship Studies 95
4.4.4.1 Synthesis of Compounds 95

4.4.4.2 Evaluation of Compounds 98

4.4.4.3 Follow-Up of Cyclopentyl Compound 103

4.5 ITC Experiments 105

4.6 Selectivity Against BRD4 108

4.7 Conclusions 108

Chapter 5: Development of Covalent Ligands for Mutant TRIM33β-E984C 111

5.1 Introduction and Aims 111

5.2 Docking Studies 112

5.3 Synthesis of Alkylating Compounds 113

5.3.1 Acrylamide and Vinyl Sulfonamide Synthesis 113

5.3.2 Methyl Acrylate and Methyl Vinyl Sulfone Synthesis 114

5.4 Synthesis of Mutant Protein TRIM33β-E984C 116

5.4.1 DNA Cloning 116

5.4.2 Protein Expression and Purification 117

5.4.3 Verification of Protein Structure 117

5.4.3.1 Circular Dichroism 117

5.4.3.2 Testing of Protein Binding Ability to Control Compound **62** 118

5.5 Investigation of Surface-Exposed Cysteine Residues 119

5.6 Investigation of Compound Binding by AlphaScreen™ 122

5.7 Alkylation Experiments with TRIM33β-E984C 124

5.7.1 Confirmation of Compound Alkylation Site 129

5.7.1.1 Verification Using NMR WaterLOGSY Competition Experiments 129

5.7.1.2 Verification by Proteomics Analysis 131

5.8 Conclusions 134

Chapter 6: Development of 3-(2-Methylpyrimidin-4-yl)phenyl Series 137

6.1 Introduction and Aims 137

6.2 Initial Investigation of the 3-(2-Methylpyrimidin-4-yl)phenyl Series 137

6.2.1 Synthesis of 3-(2-Methylpyrimidin-4-yl)phenyl Series 137

6.2.2 Evaluation of Compounds Using AlphaScreen™ 139

6.2.3 Validation of Compound **171** 140

6.3 Follow-Up of Compound **171** 141

6.3.1 Synthesis of Compounds 141

6.3.2 Evaluation of Compounds by AlphaScreen™ 145

6.4 Conclusions 146

Chapter 7: Conclusions 147

Chapter 8: Experimental Section 151

8.1 General Chemistry Methods 151

8.2 Chemistry Experimental Procedures 154

8.3 General Biology Methods 260

8.3.1 Proteins and Plasmids 260

8.3.2 Protein Concentration Measurement 260

8.3.3 pH Measurement 260

8.4 Biology Experimental Procedures 261

8.4.1 DNA Cloning 261

8.4.1.1 DNA Amplification – Polymerase Chain Reaction (PCR) 261

8.4.1.2 DNA Purification – Agarose Gel Electrophoresis 261

8.4.1.3 DNA Assembly – HiFi DNA Assembly Kit 262
8.4.1.4 Plasmid DNA Isolation 262
8.4.2 Protein Production 263
8.4.2.1 Bacterial Transformation 263
8.4.2.2 Protein Expression 263
8.4.2.3 Protein Purification 264
8.4.2.4 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis 266
8.4.3 AlphaScreen™ 267
8.4.3.1 Peptides 267
8.4.3.2 Typical Procedure for AlphaScreen™ 268
8.4.3.3 Signal Response Experiment 268
8.4.3.4 Single-Concentration Screening Experiments 268
8.4.3.5 Dose Response Experiments 269
8.4.3.6 TruHits Experiments 269
8.4.4 NMR WaterLOGSY 270
8.4.5 ITC Experiments 271
8.4.5.1 Protein-Peptide ITC 271
8.4.5.2 Protein-Ligand ITC 272
8.4.6 Circular Dichroism Spectroscopy 272
8.4.7 Alkylation Experiments 273
8.4.7.1 Alkylation Experiments with *N*-Ethylmaleimide for Proteomics 273
8.4.7.2 Alkylation Experiments with Covalent Ligands 274
8.4.7.3 Alkylation Experiments with Covalent Ligands for Proteomics 274

8.5 Docking 274

References **277**

Appendices **289**

Abstract

Epigenetic information is encoded in a wide range of histone modifications, which gives rise to the "histone code" and brings an additional level of complexity to the control of gene expression. The proteins involved include "writers", "erasers", and "readers", which install, remove, and recognise histone modifications, respectively. Given the involvement of epigenetics in diseases, including cancer, epigenetic regulation presents opportunities for the development of novel therapeutics. Understanding the roles of proteins involved in epigenetic regulation is, therefore, important.

TRIM33 is a transcriptional regulator with E3 ubiquitin ligase activity that has been associated with leukaemia and DNA damage response pathways. It contains a tandem PHD finger-bromodomain cassette, which recognises methylated and acetylated lysine residues on histone proteins. The aim of this project was to develop ligands for the TRIM33 bromodomain, which can be used as chemical tools to further validate the role of TRIM33.

This work describes the identification of hits in an initial AlphaScreen™ assay, which were then validated and used as the starting point for further development. Compound **D664-0002** (Figure 1B) was identified as a promising compound and further studies focused on this. Iterations of synthesis and evaluation have enabled the exploration of structure-activity relationships (SAR) around the benzimidazolone core. A combination of biological assay techniques and computational studies have been used to guide investigation of the acetyl-lysine binding pocket.

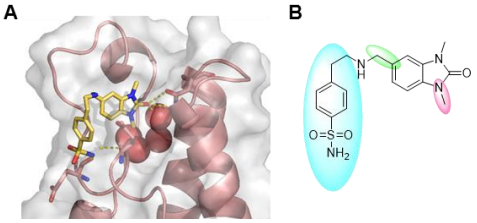


Figure 1. A) Hit compound **D664-0002** in TRIM33β bromodomain; B) Structure of **D664-0002**.

Interactions of the amine chain with glutamic acid resides E981 and E984 were identified as important for ligand design. Variation of the amine chain showed a relatively flat SAR with most ligands demonstrating low micromolar affinity, likely due to the solvent-exposed pocket. Computational studies also suggested the presence of structural water molecules in the ZA channel as a key consideration in understanding the precise binding mode and consequently, future design of a high-affinity ligand. The potential of gaining affinity *via* interaction with the region above the acetyl-lysine binding pocket was explored. This was achieved using an extension of the headgroup, which established the optimum vector combination between the extension and the amine chain. However, biological evaluation of compounds demonstrated that extensions larger than a methyl or ethyl group were not tolerated.

Covalent analogues of the TRIM33 bromodomain ligand were explored for their ability to alkylate a mutant variant TRIM33β-E984C, as a potential strategy to enable crystallography. Analogues containing an electrophile were designed and synthesised. Alkylation experiments with TRIM33β-E984C indicated selective alkylation of the mutated cysteine residue. This result remains to be confirmed through mass spectrometry studies. The exploitation of a cysteine residue to enable crystallography will increase our understanding of the TRIM33β bromodomain pocket. In particular, understanding the ligand binding mode in relation to the ZA channel water molecules will further guide the design of a high-affinity ligand.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Stuart Conway, for his support and guidance over the course of my DPhil. Thank you for the opportunity to do my DPhil project in the Conway group, for your patience and encouragement, and for the reminders that we are more stubborn than our compounds. I will always remember our group motto: *crustulum omnia solvit*.

Thank you, Angie, for laying the foundations of the TRIM project so that I could join this adventure, and for teaching me everything – from biology techniques and making detailed Excel sheets, to the quirks of using the multichannel pipette… Thank you for so generously sharing your knowledge with me.

A big thank you to the postdocs who have come into the group and brought a wealth of experience with them. Ewen, thank you for your willingness to answer my many questions, and for sharing your wisdom (and excitement over food) with me. Maze, I have greatly benefitted from your advice, especially with assays, and have enjoyed your understated humour which has only added to the fun in G12. James, the reaction encyclopaedia in your brain never ceases to amaze me (as do your "when I was in …" stories). Anthony, your enthusiasm is infectious; thank you for your help with all things biochemistry.

Working in G12 would not have been half as fun without Sarah and Mus. Sarah, I will miss our fumehood chats while columning. Thank you especially for teaching me how to use Microsoft Word properly, and for your help and patience with the quirks/frustrations of formatting. Mus, it's been great having you as my fellow G12 detective haha. I will miss our conversations about everything including our mundane TLCs and days of feeling sha'erred, and I can't wait to introduce you to the wonders of Singaporean hawker food.

To Coco and Catherine, it's been amazing starting together and sharing this journey with you both. Coco, thank you for all the tea/milo breaks we've had, sharing worries, stresses,

chemistry/biology ups and downs, and (gluten-free) cake. Catherine, meow. Thank you for always being there (also literally haha) to chat with, and for always being ready with some cheer and your loud infectious laugh.

To team TRIMendous: Amy, you are the best TRIM buddy I could have hoped for. There are no words to describe how thankful I am for your friendship (and everything else), and no words are needed *sends jumble of thoughts by telepathy*. Joe, thank you for always being willing to explain your computational work, and for teaching me how to make pretty PyMOL figures. Jess, it's been so great working on TRIM together!

To the many other postdocs, DPhils, Part IIs and rotations students who have been part of the group – especially Michael, Alex, Amelie, Pietro, Charles, Holly and Adam – thank you for all your help, life in the Conway group would not have been the same without you.

Many thanks to the NMR team, in particular Prof. Tim Claridge, for teaching me to run the waterLOGSY experiments and for patiently answering my questions. Thank you also the mass spectrometry team, especially Elisabete for kindly helping me with my proteomics samples.

I have been greatly blessed with the friendship of many from home and St Ebbe's *Thesis*, without whom this DPhil journey would not have been as enjoyable. Last but not least, my deepest gratitude goes to my parents and family for all their love and support.

Abbreviations

2D	Two-dimensional
4-DMAP	4-Dimethylaminopyridine
Å	Angstrom
a.u.	Arbitrary units
ÄKTAFPLC	ÄKTA fast protein liquid chromatography
ALC1	Amplified in liver cancer 1
AlphaScreen™	Amplified luminescent proximity homogeneous assay screen
APS	Ammonium persulfate
ATAD2	ATPase family AAA domain-containing protein 2
β-ME	β-Mercaptoethanol
BAZ2	Bromodomain adjacent to zinc finger domain protein 2
BET	Bromodomain and extra terminal domain
Boc	*tert*-Butoxycarbonyl
BPTF	Bromodomain PHD finger transcription factor
BRD2	Bromodomain-containing protein 2
BRD4	Bromodomain-containing protein 4
BSA	Bovine serum albumin
CDI	1,1'-Carbonyldiimidazole
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CI	Chemical ionisation
COSY	Homonuclear correlation spectroscopy
CV	Column volumes
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DIBAL	Diisobutylaluminium hydride
DMF	*N*,*N*-Dimethylformamide
DMSO	Dimethyl sulfoxide

DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E2	Ubiquitin-conjugating enzyme
EI	Electron impact
ERα	Oestrogen receptor alpha
ESI	Electrospray ionisation
FAC	Final assay concentration
FDA	Food and Drug Administration
GSK	GlaxoSmithKline
h	Hour(s)
H3	Histone 3
HAT	Histone acetyltransferase
HCV	Hepatitis C virus
HDAC	Histone deacetylase
HEPES	(4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid)
HIV-1	Human immunodeficiency virus type 1
HMBC	Heteronuclear multiple-bond correlation
HP1	Heterochromatin protein 1
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
HSQC	Heteronuclear single quantum correlation
IC_{50}	Half-maximal inhibitory concentration
IMAC	Immobilised metal affinity chromatography
ING2	Inhibitor of growth protein 2
IPA	Isopropanol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KAc	Acetylated lysine

kb	Kilobase
K_D	Dissociation constant
KDM5A	Lysine demethylase 5A
LC-MS	Liquid chromatography-mass spectrometry
LE	Ligand efficiency
LLE	Lipophilic ligand efficiency
LOCI™	Luminescent oxygen channelling assay
LRMS	Low-resolution mass spectrometry
m.p.	Melting point
MALDI-TOF	Matrix assisted laser desorption/ionization time-of-flight
min	Minute(s)
MW	Molecular weight
MWCO	Molecular weight cut-off
NEM	*N*-Ethylmaleimide
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
ns	Nanoseconds
NSCLC	Non-small cell lung carcinoma
o/n	Overnight
OD_{600}	Optical density at wavelength of 600 nm
p300/CBP	cAMP response element-binding (CREB) binding protein
PARP	Poly(ADP-ribose) polymerase
PBS	Phosphate-buffered saline
PCAF	p300/CBP-associated factor
PCR	Polymerase chain reaction
PDB	Protein data bank
PHD	Plant homeodomain
PPI	Protein-protein interaction

ppm	Parts per million
R_f	Retention factor
RING	Really interesting new gene
RNA	Ribonucleic acid
RP-HPLC	Reversed-phase high-performance liquid chromatography
rpm	Revolutions per minute
rt	Room temperature
SAR	Structure-activity relationships
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SFI	Solubility forecast index
SGC	Structural Genomics Consortium
SMAD	"Small" worm phenotype mothers against decapentaplegic
t	Time
TBAI	Tetrabutylammonium iodide
TBDMS	*tert*-Butyldimethylsilyl
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TGF-β	Transforming growth factor beta
THF	Tetrahydrofuran
TIF1	Transcription intermediary factor 1
TLC	Thin-layer chromatography
T_m	Melting temperature
TRIM	Tripartite motif-containing
uv	Ultraviolet
WaterLOGSY	Water-ligand observed *via* gradient spectroscopy
WT	Wild-type

Chapter 1

Introduction

1.1 Epigenetics and the Histone Code

A growing understanding of epigenetics has led to a deeper appreciation of the complexity of cell biology, and the therapeutic opportunities that arise from this. The idea of epigenetics was first proposed by Waddington in 1942, referring to the mechanisms relating genotype and phenotype.[1] More recently, epigenetics has been defined as "a stably heritable phenotype resulting from changes in a chromosome without alterations in the deoxyribonucleic acid (DNA) sequence".[2] This is mediated through covalent DNA modification, histone marks or non-coding ribonucleic acids (RNAs) in RNA interference, which results in gene silencing.[3]

The first-discovered class of epigenetic modification is DNA methylation, where position C5 of cytosine in CpG dinucleotides are methylated (**1**, Figure 1.1).[4] The maintenance of DNA methylation state is important and has implications in gene expression, development, and disease. More recently, additional DNA modifications have been identified and studied in light of their potential functional significance (Figure 1.1).[5] In contrast, the other main class of epigenetic modification, histone modifications, does not involve direct modification of DNA.

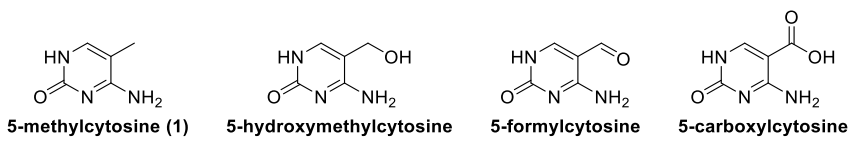


Figure 1.1. 5-Methylcytosine and other derivatives.[5]

In eukaryotic cells, DNA is wound around histone octamers (comprising two of each histone, H2A, H2B, H3 and H4), forming nucleosomes.[6] These nucleosomes, forming "beads on a string", are further coiled and packed into chromatin. These structures exist in two forms – euchromatin and heterochromatin – reflecting the lower and higher densities of packing,

which are associated with an "on" and "off" transcription state, respectively.[7] Histone proteins, initially considered to play only a structural role in the packing of DNA into chromatin, have since been shown to play an important role in the regulation of gene transcription.[8] Histone proteins undergo a wide range of more than 60 post-translational modifications on the amino acid residues, including methylation, acetylation, phosphorylation, ubiquitylation, and crotonylation.[9,10] However, the functional relevance of the various modifications remains to be established.[10,11] Early understanding of these modifications attributed downstream cellular effects to the difference in charge. For example, acetylation neutralises the positive charge on lysine residues, lowering their affinity for negatively-charged DNA, resulting in more loosely-packed chromatin that permits access of transcription machinery.[12] However, this model has changed following the demonstration that specific post-translational modifications regulate binding of chromatin factors involved in transcription and chromatin remodelling.[13] The ability of these histone modifications to modulate biological activity has given rise to the idea of the "histone code" and brings an additional level of complexity to understanding gene expression. Formally proposed by Strahl and Allis in 2000, the "histone code" refers to a language encoded on histone tails by distinct patterns of modifications with associated biological specificity, leading to a downstream phenotype.[14]

Enzymes that add and remove these post-translational modifications are known as "writers" and "erasers", respectively.[15] "Writers" install the chemical group on the histone proteins, which form histone tails protruding out of the nucleosome complex (Figure 1.2). An additional class of proteins are the "reader" proteins, which bind to the chemical group installed and relate epigenetic information to the cell by modulating protein-protein interactions with cellular machinery.[16] Given the role of these modifications in the control of gene expression, they have implications in a wide range of diseases including cancer,[17] asthma,[18] parasitic diseases[19] and Alzheimer's disease.[20] Consequently, understanding

epigenetic regulation is of great importance and presents opportunities for the development of novel therapeutics.

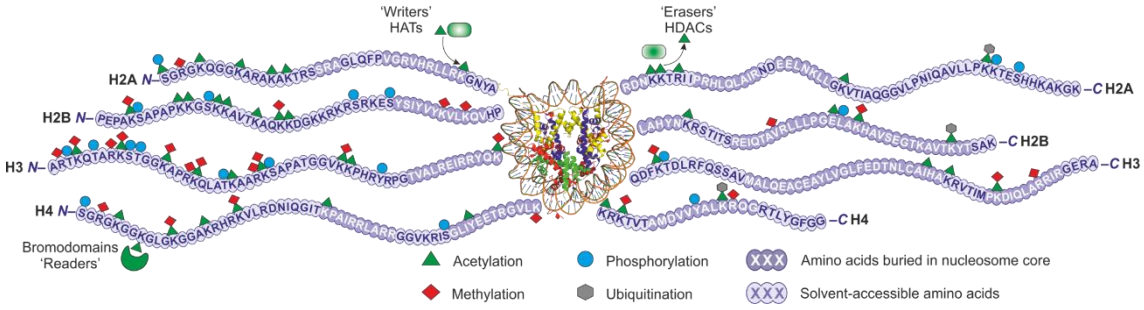


Figure 1.2. Schematic of a nucleosome with post-translational modifications on histone tails, indicated by amino acid sequences. DNA is wound around histone octamers with histone tails protruding out. Acetyl groups (green triangles) are a common chemical modification on lysine residues and are installed by histone acetyltransferases (HATs), removed by histone deacetylases (HDACs) and read by bromodomains. Reprinted with permission from Jennings *et al.*, Future Medicinal Chemistry.[16] Copyright 2014 Future Science Group.

1.1.1 Multiple Epigenetic Domains

Epigenetic proteins often contain more than one domain capable of recognising, installing or removing post-translational modifications, permitting chromatin interaction involving multiple specific histone modifications. Consequently, combinatorial readout is observed, with several options possible.[21] For instance, the binding of a protein to a specific modification can be dependent on another modification in close proximity. This was observed in heterochromatin protein 1 (HP1), which binds to histone H3 trimethylated lysine 9 ($H3K9Me_3$) but not during mitosis when H3S10 is phosphorylated.[22] Many proteins have dual-"reader" domains that recognise multiple modifications simultaneously through multivalent binding.[21] This simultaneous recognition leads to context-dependent binding where there is cooperativity of the binding event, possibly on the same histone tail, or across different histone tails on the same nucleosome as seen with bromodomain plant homeodomain (PHD) finger transcription factor (BPTF) (Figure 1.3).[23]

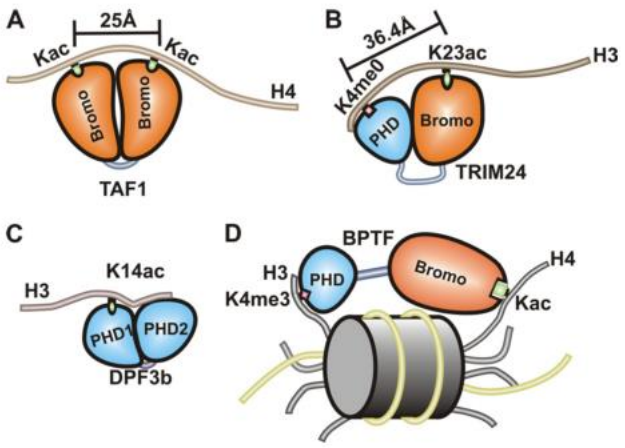


Figure 1.3. Models of combinatorial readout of two histone marks by paired reader domains. A) The TAF1 double bromodomain reads two acetylated lysines (KAc) on the H4 histone tail; B) The TRIM24 PHD-bromodomain cassette reads H3K4Me$_0$ and H3K23Ac modifications; C) The DPF3b double PHD reads N-terminal H3 and the H3K14Ac mark simultaneously; D) The BPTF PHD-bromodomain cassette reads H3K4Me$_3$ and H4KAc paired histone marks on a single nucleosome. Reprinted with permission from Wang *et al.*, Journal of Biological Chemistry.[23] Copyright 2011 The American Society for Biochemistry and Molecular Biology.

Further contributing to the complex dynamic of histone modifications are proteins with "reader"-"writer" or "reader"-"eraser" modules on the same protein.[23] For example, the binding of PHD finger 1 of lysine demethylase 5A (KDM5A) to H3K4Me$_0$, stimulates the removal of methyl marks from H3K4Me$_3$ by KDM5A, resulting in a positive feedback mechanism.[24] Therefore, development of ligands for specific domains would be a valuable tool in enabling detailed study that could further clarify the relationships involved in dual readership.

1.2 Bromodomains

Bromodomains are "readers" of acetyl-lysine (KAc) comprising approximately 110 amino acids.[25] There are 62 known bromodomains in the human proteome, found within 46 different proteins. Sequence alignment reveals a phylogenetic tree with eight families (Figure 1.4).[26]

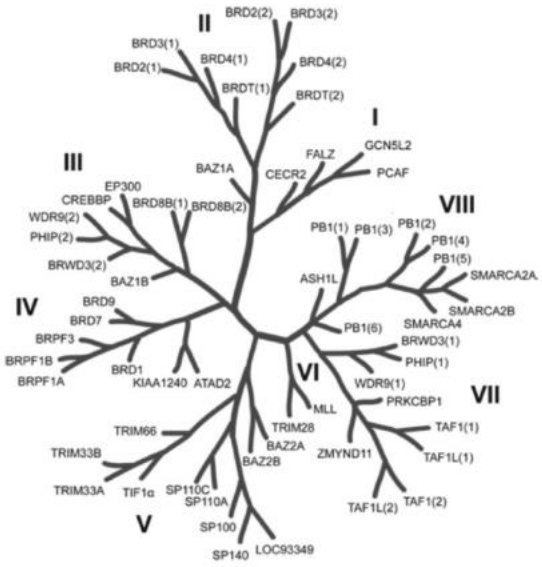


Figure 1.4. Phylogenetic tree of the human bromodomain family.[27] Reprinted with permission from Hewings *et al.*, Journal of Medicinal Chemistry.[16] Copyright 2012 American Chemical Society.

Structurally, bromodomains are characterised by a left-handed bundle of four α-helices (αA, αB, αC and αZ) connected by two loop regions (ZA and BC loops) (Figure 1.5A).[16] The acetyl-lysine binding site of most bromodomains contains a conserved asparagine, and four or five water molecules forming the base of the pocket.[16] The conserved asparagine, which can be a threonine or tyrosine in some bromodomains, forms a key hydrogen bond with the acetyl-lysine residue of the histone peptide (Figure 1.5B). While the secondary and tertiary structure of bromodomains is highly conserved, they have low sequence homology, resulting in diversity in the bromodomain binding partners.[26]

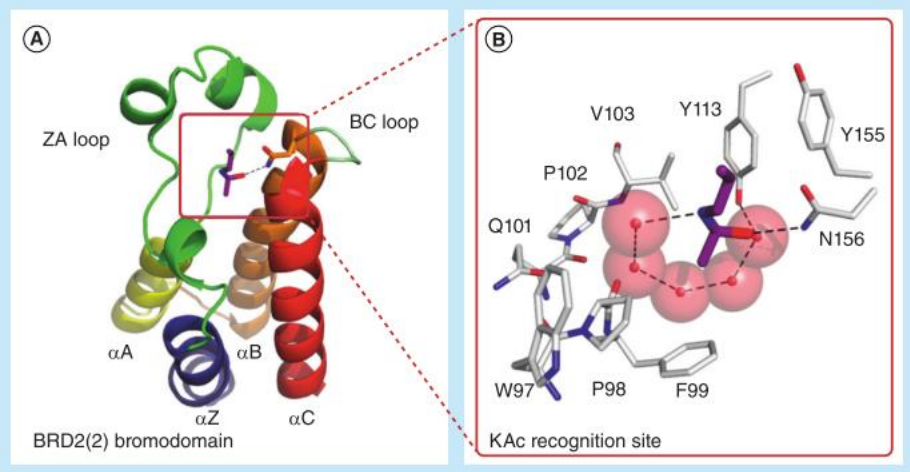


Figure 1.5. Bromodomain structure. A) X-Ray crystal structure of the second bromodomain of bromodomain-containing protein 2 (BRD2) in complex with acetylated lysine (carbon = purple, PDB ID: 2DVQ).[16] The bromodomain structural motif is formed of four helices αZ (blue), αA (yellow), αB (orange) and αC (red), and two loop regions known as the ZA loop (green) and the BC loop (light green); B) The acetylated lysine residue binds in a well-defined hydrophobic pocket, forming interactions with a highly conserved asparagine residue and several tightly bound water molecules (red spheres), which form the base of the pocket. Reprinted with permission from Jennings *et al.*, Future Medicinal Chemistry.[16] Copyright 2014 Future Science Group.

Bromodomain-containing proteins are involved in a wide variety of functions relating to the formation of chromatin complexes, including chromatin modification and transcription co-activation. Bromodomain-containing protein 4 (BRD4), for example, is a co-activator in the transcriptional activation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) and the activation of RNA polymerase II,[28] while the bromodomain-containing P300/CBP-associated factor (PCAF) also has histone acetyltransferase activity.[29]

1.3 Plant Homeodomain (PHD) Finger

Plant homeodomain (PHD) fingers are "readers" of methylated and unmethylated lysine residues or, to a lesser extent, methylated arginine residues, on histone tails.[30] PHD fingers, comprising 50 to 80 amino acids, contain a zinc-binding motif where two antiparallel β-sheets are stabilised by two zinc ions bound to the $Cys_4HisCys_3$ motif (Figure 1.6A and Figure 1.6B).[31] Despite a high sequence variability in PHD fingers, this secondary structure is maintained while the diversity in primary sequence enables PHD fingers to distinguish between methylation states, resulting in specificity in ligand binding.[31]

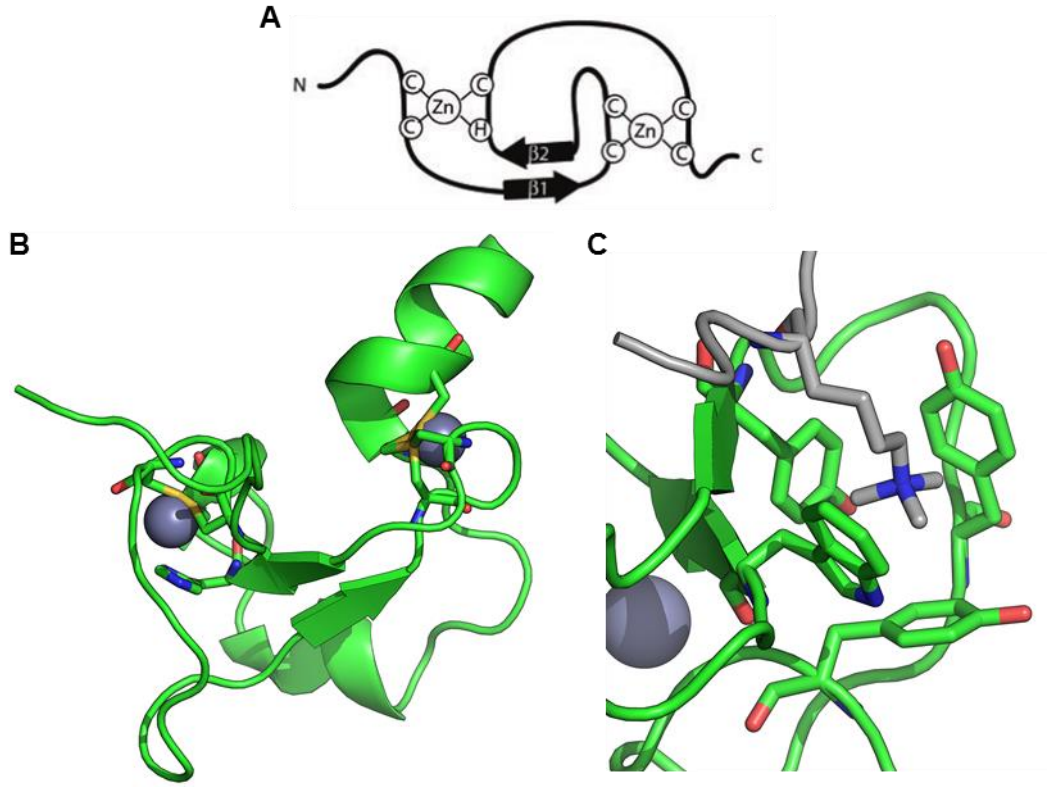


Figure 1.6. Structure of PHD finger. A) The PHD finger is characterised by two antiparallel β-sheets stabilised by two zinc ions bound to a $Cys_4HisCys_3$ motif. Reprinted with permission from Musselman *et al.*, Molecular Interventions.[30] Copyright 2009 American Society for Pharmacology and Experimental Therapeutics; B) X-Ray crystal structure of the BPTF PHD finger (PDB ID: 2FUU) showing the characteristic structures; C) X-Ray crystal structure of the BPTF PHD finger (PDB ID: 2FUU) with three tyrosine residues and one tryptophan residue forming an aromatic cage that recognises trimethylated lysine (histone peptide in grey).

PHD fingers recognise lysine residues with varying levels of methylation, often through interaction with an aromatic cage of three or four residues consisting of phenylalanine, tyrosine and tryptophan residues (Figure 1.6C).[32] In contrast, binding to unmethylated $H3K4Me_0$ is often achieved through interaction with acidic or hydrophobic residues.[31] Like bromodomains, PHD fingers are often associated with proteins involved in chromatin modification and remodelling. For example, inhibitor of growth protein 2 (ING2) is a tumour suppressor with a PHD finger that binds $H3K4Me_3$, securing the mSin3a/HDAC1 histone deacetylase complex, resulting in histone deacetylation and subsequent gene repression.[30]

1.4 Tripartite Motif-Containing (TRIM) Proteins

The tripartite motif-containing (TRIM) family of proteins is a diverse group of approximately 80 proteins classified into 11 subfamilies based on differences in their C-terminal domain.[33] TRIM proteins are characterised by the N-terminal RBCC motif, consisting of a really interesting new gene (RING) domain, one or two B-box domains followed by a coiled-coil domain.[34] The RING domain catalyses ubiquitination through recruiting ubiquitin-conjugating enzyme (E2), enabling TRIM proteins to act as E3 ubiquitin ligases.[33] TRIM proteins are involved in many different functions in the cell and have been implicated in cancer, cell signalling, immunity, and autophagy.[35]

Subgroup C-VI of the TRIM family of proteins, defined by the tandem PHD-bromodomain cassette at their C-terminal, comprises TRIM24, TRIM28 and TRIM33 (Figure 1.7).[33] These proteins are also known as transcription intermediary factor 1 (TIF1)-α, TIF1-β and TIF1-γ, respectively. All three proteins are cofactors that interact with transcription factors.[36] TRIM24, for example, interacts with retinoic acid receptor alpha (RARα) and is a co-activator of oestrogen receptor alpha (ERα), which is associated with cell proliferation in breast cancer.[37] More recently, TRIM24 has also been found to act as a transcriptional activator of genes involved in cell proliferation in prostate cancer.[38]

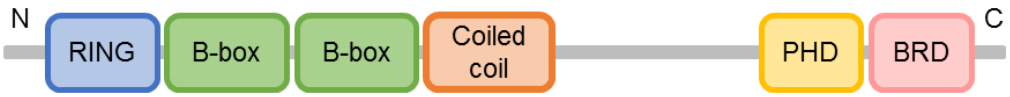


Figure 1.7. Schematic of TRIM protein subgroup VI architecture.

1.4.1 TRIM33

TRIM33 and TRIM24, together with other PHD finger-containing proteins, belong to group V of the bromodomain family of proteins. They are structurally similar and interact with methylated and acetylated lysine residues *via* their PHD finger and bromodomain, respectively.[37,39] TRIM33 is a repressor of the transforming growth factor beta (TGF-β) family and binds histone 3 where lysine 4 is unmethylated and lysine 18 and 23 are

acetylated.[39] This binding to chromatin is important as studies suggest that this has implications in the mothers against decapentaplegic (SMAD) pathway[40] and regulates its E3 ubiquitin ligase activity.[39] TRIM33 also acts as a tumour repressor through its involvement in mitosis[41] and haematopoiesis.[42] Additionally, TRIM33 interacts with amplified in liver cancer 1 (ALC1) protein and regulates its function in the poly(ADP-ribose) polymerase (PARP)-dependent DNA damage response pathway.[43]

TRIM33 exists in two isoforms known as TRIM33α and TRIM33β. The bromodomains of both isoforms are defined by the characteristic left-handed bundle of four α-helices, with the only difference being a 17 amino acid insertion in the BC loop of TRIM33α (Figure 1.8A). The PHD finger of TRIM33, like TRIM24, lacks an aromatic cage, and instead binds H3K4Me_0 through aspartate and glutamate residues (Figure 1.8B).[40] Interaction with H3K9Me_3 is achieved through a cation-π interaction with a conserved tryptophan residue (W889) (Figure 1.8B).[40] Binding to acetylated lysine H3K18Ac occurs in the typical manner *via* a hydrogen bonding interaction with the conserved asparagine (N1039). Given the limited understanding of TRIM33 function, particularly with respect to the individual isoforms, and its suggested implications in cancer, a TRIM33 ligand would serve as a useful tool in the validation of TRIM33 function in cells.

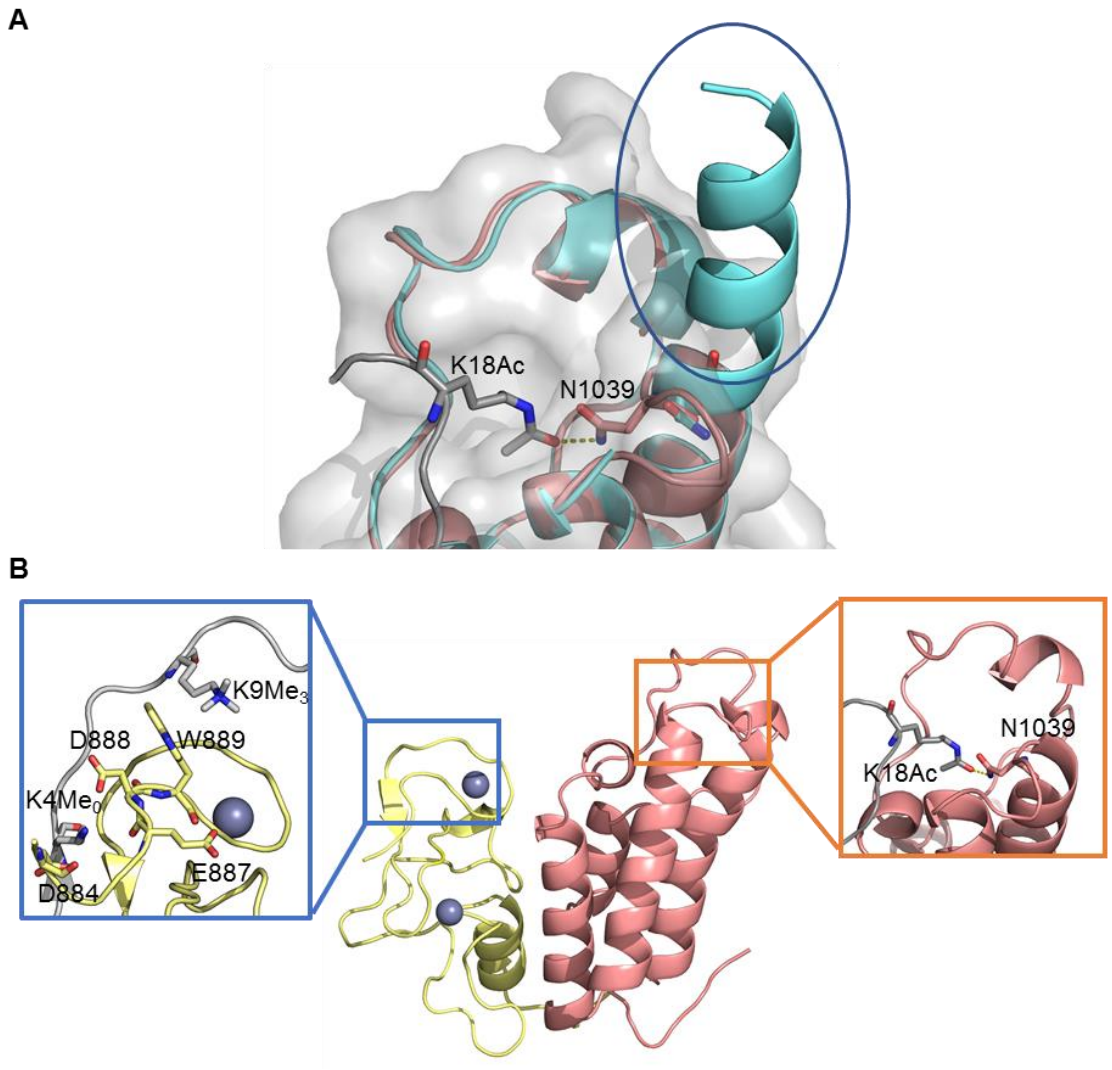


Figure 1.8. Structure of TRIM33. A) Overlay of X-ray crystal structure of TRIM33α (PDB ID: 3U5O, blue) and TRIM33β (unpublished, Professor Stefan Knapp, pink), with 17 amino acid insertion indicated by the blue circle; B) X-Ray crystal structure of TRIM33β showing the bromodomain (pink) binding acetyl-lysine through conserved asparagine N1039 and the PHD finger (yellow) binding trimethylated lysine 9 and unmethylated lysine 4 (peptide overlay from crystal structure of TRIM33α, PDB ID: 3U5O) through conserved tryptophan W889 and acidic aspartate and glutamate residues D884, E887 and D888.

1.5 Small Molecule Ligands for Bromodomains and PHD Fingers

1.5.1 Bromodomain Ligands

The first bromodomain ligand **2** (Figure 1.9) was reported in 2007,[44] following the discovery in 2002 that human immunodeficiency virus type 1 (HIV-1) trans-activator protein Tat transcription activity is dependent on acetylation at lysine 50, and subsequent association with bromodomain-containing proteins P300/CBP (CREB binding protein) and PCAF.[45] In 2010, **(+)-JQ1** (**3**)[46] and **I-BET762** (**4**)[47] were reported as ligands for the

bromodomain and extra terminal domain (BET) family bromodomains by the Structural Genomics Consortium (SGC) with the Dana-Farber Cancer Institute, and GlaxoSmithKline (GSK), respectively. Shortly after this, our group reported the discovery of 3,5-dimethylisoxazole **OXFBD02** (**5**) as acetyl-lysine mimetics and bromodomain ligands.[48,49]

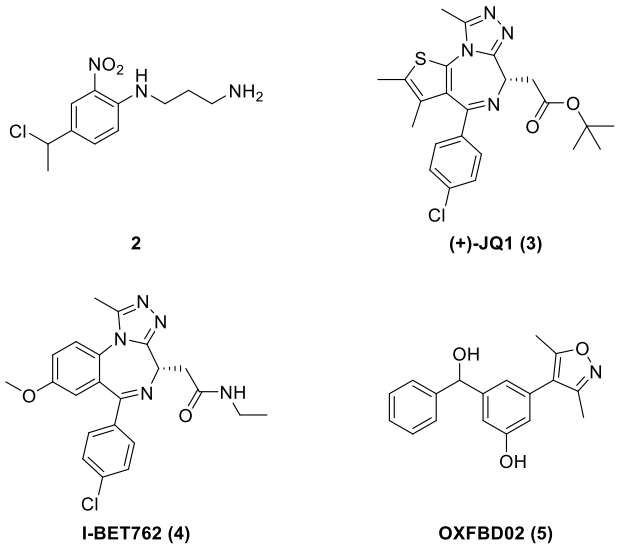


Figure 1.9. Examples of bromodomain ligands.[44,46,47,48]

The success of **(+)-JQ1** in treating NUT midline carcinoma in patient-derived mouse xenograft models[46] has since sparked a large interest in the development of bromodomain ligands for oncology. The use of known bromodomain ligands as tool compounds has also revealed the relevance of bromodomains in a wide range of biological processes including angiogenesis[50] and immunity.[51] The ligands have also enabled target validation[52] and deeper understanding of cell biology.[53] This has further fuelled the discovery of more selective ligands of the BET family proteins and ligands of non-BET proteins, such as CBP,[54] ATPase family AAA domain-containing protein 2 (ATAD2),[55] and BRPF.[56] With many bromodomain ligands currently in clinical trials,[57,58] it is of ongoing interest to develop bromodomain ligands both for therapeutic indications and to aid understanding of their role in disease.

In contrast to certain classes of protein-protein interactions that involve large and relatively flat interfaces, bromodomains are considered more druggable due to the defined binding pocket and key interaction with acetylated lysine. This has been affirmed by the variety of bromodomain ligands that have been developed over the years. Despite this, in a bromodomain druggability study by Vidler *et al.*[59] based on binding site grouping, it was noted that TRIM24 and TRIM33 were difficult to develop ligands for as compared to some other bromodomains, due to their relatively open binding site. While there are still no published ligands of the TRIM33 bromodomain, there are two known ligands of the related TRIM24 bromodomain, **IACS-9571**[60] (**6**) and **7**[61] developed by the MD Anderson Cancer Centre and the SGC, respectively (Figure 1.10).

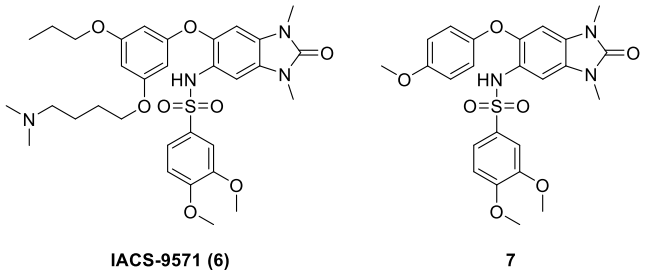


Figure 1.10. TRIM24 ligands.[60,61]

1.5.2 PHD Finger

Methyl-lysine reader proteins are an important class of epigenetic proteins, playing a key role in chromatin regulation.[62] Currently, there are ligands for methyl-lysine reader domains including Tudor domains and chromodomains, but inhibition of these proteins are not yet well understood, particularly because they bind to their substrate with low affinities (low- to mid-micromolar) and often involve an induced-fit mechanism.[63] Despite the large number of PHD fingers that exist on a more than 100 proteins,[64] there are very few reported ligands of the PHD finger.

Amiodarone (**8**, Figure 1.11) analogues were first identified in 2012 as ligands of the JARID1A (also known as KDM5A) PHD finger[65] which binds H3K4Me$_3$. Since then, amiodarone analogues have been used in further studies to better understand the

mechanism of its catalytic activity.[66] Benzothiazoles (**9**, Figure 1.11) also bind in a narrow cleft on the Pygo PHD finger.[67] More recently, a combination of *in silico* docking and nuclear magnetic resonance (NMR)-based fragment screening resulted in the development of four fragments (**10**–**13**, Figure 1.11) that bind to two different pockets on the bromodomain adjacent to zinc finger domain protein 2 (BAZ2).[68] These fragments could serve as a starting point for the development of higher-affinity ligands, improving understanding of PHD finger inhibition.

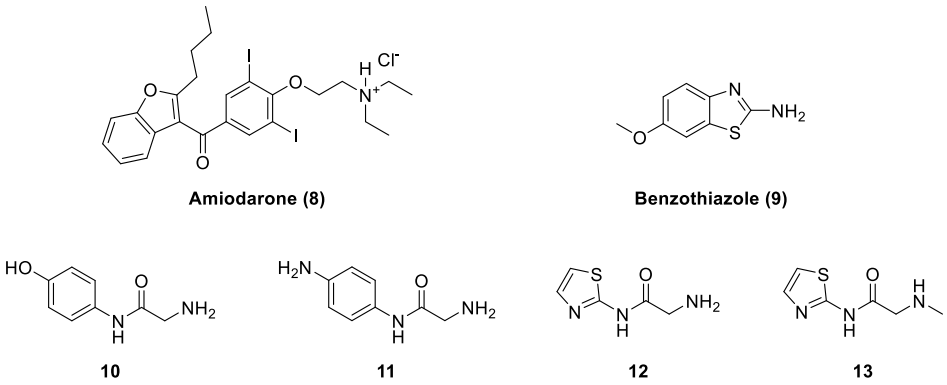


Figure 1.11. Examples of PHD finger ligands.[65,67,68]

1.5.3 Covalent Inhibitors of Proteins

Covalent inhibitors are compounds containing functional groups that react with the protein target through an amino acid residue. This amino acid side chain may be involved in the catalytic mechanism of the enzyme or may be a non-catalytic side chain situated in close proximity to the binding pocket.[69] First discovered from natural products, many of which were antibiotics, covalent inhibitors in the context of drug discovery were approached with caution due to the potential for off-target effects and toxic liability.[69] Telaprevir (**14**, Figure 1.12) is an example of a covalent drug targeting a catalytic serine residue of the hepatitis C virus (HCV) protease that was approved by the Food and Drug Administration (FDA) in 2011.[70] Additionally, the rise of target-based discovery has seen an increase in rationally-designed covalent inhibitors.[71] This has led to the discovery of drugs including

afatinib (**15**, Figure 1.12), a tyrosine kinase inhibitor which was recently approved for the treatment of non-small cell lung carcinoma (NSCLC).[72]

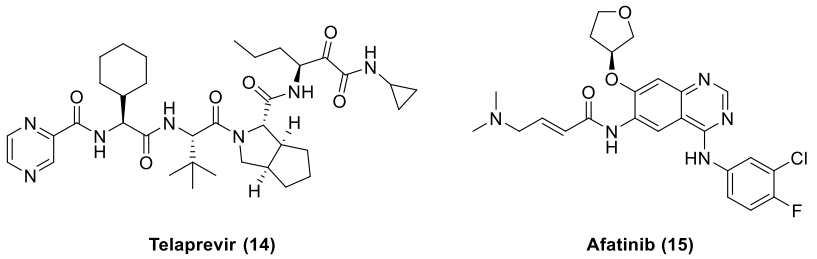


Figure 1.12. Examples of FDA approved covalent drugs.[70,72]

One of the most common amino acid residues targeted in the design of covalent inhibitors is cysteine.[73] This is unsurprising because cysteine has the most nucleophilic amino acid side chain. A wide range of electrophiles including a variety of Michael acceptors have been developed as cysteine-targeting moieties (Figure 1.13).[69,73]

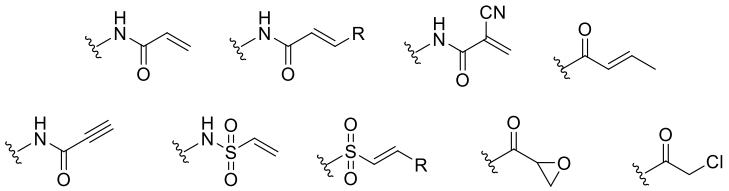


Figure 1.13. Examples of cysteine-targeting electrophiles.[69,73]

To take advantage of these reactive cysteine residues, tethering approaches toward ligand discovery have been developed. This approach involves the use of a native or mutated cysteine residue to identify ligands from a library of disulfide-forming compounds.[74] Compounds that interact non-covalently with the binding site would react with the cysteine residue situated nearby, allowing their binding mode to be studied. Erlanson *et al*. demonstrated that this strategy enabled the discovery of a series of ligands for *Escherichia coli* thymidylate synthase.[74]

There are three examples of covalent ligands for bromodomains reported in the literature. In 2015, Daguer *et al*. identified **8a1b** (**16**) and **19a1b** (**17**) (Figure 1.14) as covalent ligands of the PCAF bromodomain.[75] Cell lysate enrichment studies suggested that many other

bromodomains, including TRIM24 and TRIM33, contain reactive cysteine residues that could be exploited to develop tools for further understanding the role of bromodomains. More recently, the design of a covalent inhibitor **ZEN-3862** (**18**, Figure 1.14) targeting a methionine residue in BRD4 *via* a reactive epoxide has also been reported.[76] Moreover, tropolone derivatives (such as **19**, Figure 1.14) have been shown to covalently label BRD4 upon photoactivation, suggesting that these photosensitive ligands could also be exploited to enable direct target validation.[77]

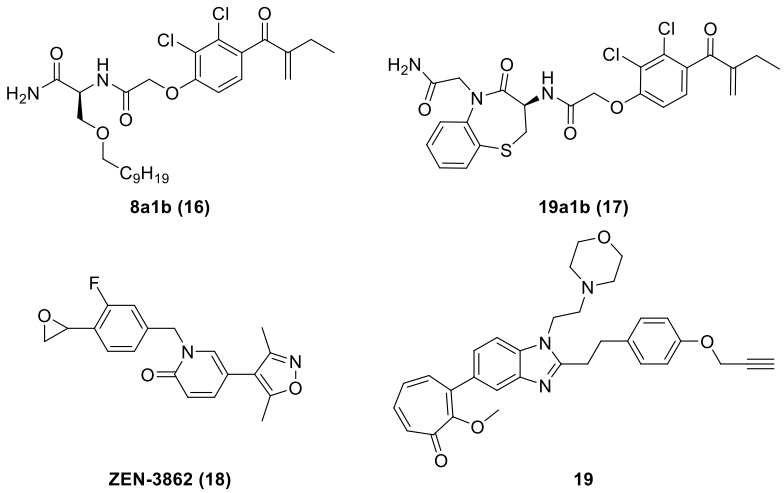


Figure 1.14. Covalent bromodomain ligands.[75,76]

1.6 Assays

In the identification and development of ligands, it is crucial to have reliable assays that are appropriate for the target protein and that enable assessment and validation of compound binding. In instances when the target protein is an enzyme, enzyme inhibition assays are often used. However, this is not possible in probing a protein-protein interaction. In studying protein-protein interactions, the most commonly used primary assay often relies on ligand displacement as this also enables high throughput screening.[78] Assays that directly probe binding, such as NMR techniques, are also employed as orthogonal assays to further confirm binding.

1.6.1 AlphaScreen™

AlphaScreen™ (Amplified Luminescent Proximity Homogeneous Assay Screen) is a proximity-based displacement assay that relies on the displacement of a binding partner resulting in a corresponding change in fluorescence readout.[79] This technology has been adapted for use in a range of cell-based and *in vitro* formats studying a variety of protein types including kinases and receptors.[80] The development and utility of this assay has been central to the identification and development of the first BRD4 bromodomain ligands.[81] Following the use of AlphaScreen™ in the development of BRD4 ligand **(+)-JQ1** (**3**),[46] this has become a widely-used assay playing a key role in the discovery and development of other bromodomain ligands. This is also the set-up on which the TRIM24 and TRIM33 AlphaScreen™ assay has been based.

AlphaScreen™ was developed from the luminescent oxygen channelling assay (LOCI™) method, where one particle contains a photosensitiser that produces singlet oxygen on exposure to light, and another particle reacts with the singlet oxygen to release a chemiluminescent signal.[82] The intensity of emission is thus dependent on the acceptor particle being sufficiently close to the donor particle to react with the singlet oxygen. In the AlphaScreen™, the donor bead containing the photosensitiser phthalocyanine produces singlet oxygen when irradiated with light at 680 nm (Figure 1.15). When the acceptor bead is within 200 nm, the singlet oxygen is able to diffuse to it and reacts with thioxene in the bead to produce a chemiluminescent signal at 370 nm.[79,80] Transfer of energy to anthracene and rubrene results in fluorescence emission at 520–620 nm, which can then be detected and measured.

In this AlphaScreen™ assay, where the target protein is an epigenetic reader protein, the binding partner for displacement is a peptide sequence from the histone tail, which is a known binding partner. The donor beads are coated with streptavidin, which binds to the peptide binding partner that has a terminal biotin tag. The acceptor beads are

nickel-chelated, enabling the target protein containing a histidine tag to bind to it. When the target protein is bound to the peptide binding partner, the beads are in close proximity and hence, on irradiation, release a chemiluminescent signal (Figure 1.15). However, in the presence of a ligand that inhibits this protein-peptide interaction, the distance between the donor and acceptor beads increases, resulting in a reduction in emission. This technique allows the affinity of the ligands to be measured and quantified through dose-response experiments with increasing compound concentration.

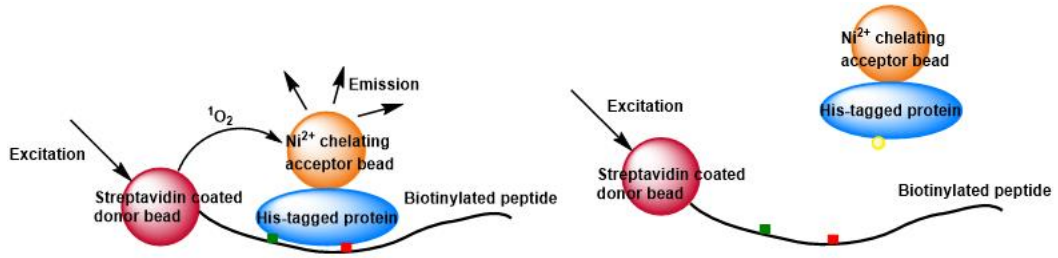


Figure 1.15. Schematic diagram of AlphaScreen™ assay. Donor beads (red) contain a photosensitiser, which produces singlet oxygen on irradiation. The interaction of singlet oxygen with acceptor beads (orange) results in a chemiluminescent signal. The biotinylated peptide binds to streptavidin coated donor beads (red) while His-tagged protein (blue) bind to nickel chelating acceptor beads (orange). The His_6-tagged protein also binds to specific modifications on the peptide (green and red squares). However, in the presence of a ligand (yellow), this interaction is disrupted resulting in no emission.

In this study, TRIM24, TRIM33α and TRIM33β are the three target proteins while the peptide binding partners are $H3_{1-27}K9Me_3$, $H3_{1-27}K18Ac$ and $H3_{1-27}K9Me_3K18Ac$. These peptides, respectively, bind to the PHD finger, bromodomain, and both domains. While compounds are tested against all three proteins each with all three peptides, in developing a ligand for the bromodomain, the peptide of particular interest is $H3_{1-27}K18Ac$. However, as AlphaScreen™ is a displacement assay, the IC_{50} values measured are dependent on the affinity of the peptide binding partner. Consequently, IC_{50} values can be compared between ligands in the same protein-peptide combination, but not across protein-peptide combinations. This is an inherent limitation of measuring IC_{50} values as they can vary depending on the assay set-up.

1.6.2 NMR WaterLOGSY

Water-ligand observed *via* gradient spectroscopy (waterLOGSY) is an NMR technique based on the transfer of magnetisation of bulk water molecules to the ligand.[83] In this ligand-observed experiment, ligands that bind to the target protein can be detected directly without the need for protein or ligand labelling, or a fluorescent probe. When the compound does not bind to the protein, the nuclear Overhauser effect (NOE) observed is small and positive, resulting in negative peaks in the waterLOGSY spectrum (Figure 1.16).[84] Conversely, when the compound binds to the protein, the NOE observed is stronger and negative due to slow tumbling, resulting in positive peaks in the waterLOGSY spectrum. This inversion, when compared to the control spectrum in the absence of protein, allows compound binding to be assessed.

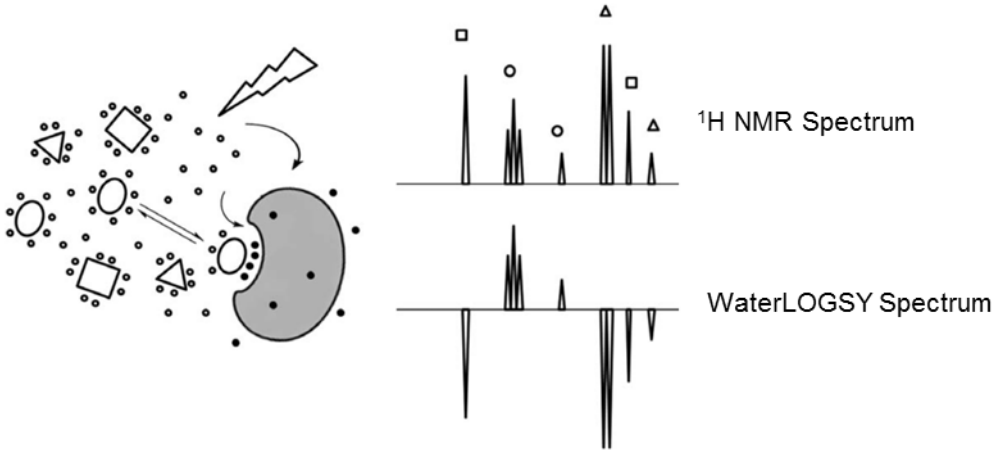


Figure 1.16. Schematic diagram of waterLOGSY with the compound NMR spectrum (top) and resulting waterLOGSY spectrum (bottom). Excitation of bulk water (small circles) results in transfer of magnetisation directly to the ligand or via the protein, resulting in negative peaks in the waterLOGSY spectrum (bottom) for binding compounds (circle) or positive peaks for non-binding compounds (triangle and square). Reprinted with permission from del Carmen Fernandez-Alonso *et al*., Current Protein and Peptide Science.[85] Copyright 2012 Bentham Science Publishers Ltd.

In the absence of protein, the waterLOGSY spectrum is negative due to the small and positive NOEs. However, there are two instances where positive peaks may be detected.[83] In the first instance, chemical exchange of labile protons provide an alternative magnetisation transfer mechanism, leading to positive signals in the waterLOGSY spectrum corresponding to the peaks of the labile protons. In the second case, when all compound

peaks are positive, this suggests that the formation of soluble aggregates has occurred due to the limited solubility of the compound. As a result, the compound concentration should be decreased until no aggregation is observed. This often requires an increased number of scans. Unfortunately, if the signals are too weak, or if further dilution is not sufficient to prevent the formation of aggregates, binding of the compound to the target protein cannot be assessed by waterLOGSY. As with NMR techniques that are ligand observed, waterLOGSY is not able to detect tightly binding ligands with low dissociation constant (K_D) (less than 1 μM) due to the slow dissociation rates resulting in difficulty detecting the NOEs on dissociation.[83]

1.6.3 Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC) is an orthogonal technique that has been widely used to directly measure the binding affinity between a protein and a peptide or ligand. The ITC instrument consists of two identical cells, one sample cell and one reference cell. These are contained in an adiabatic jacket with a thermocouple to detect small differences in temperature between the two cells.[86] Each cell is also connected to a heater that is activated and supplies heat to the cell when required to maintain the predefined temperature. Additionally, the sample cell is attached to a syringe that injects a small amount titrant (the ligand, for example) into the cell (containing the protein) at predetermined intervals (Figure 1.17A).

At the start of the experiment, a constant power is supplied to the sample cell. When a binding event occurs, heat is either absorbed or released depending on whether it is endothermic or exothermic. In the case of an exothermic reaction, the power supplied is momentarily reduced to allow the sample cell temperature to match that of the reference cell (Figure 1.17B). Following each injection, as the ligand concentration increases, the protein binding sites are gradually saturated. Consequently, the heat evolved with each injection decreases until the binding sites are fully saturated. After this point, any residual

peaks are due to the heat of dilution of the ligand. Analysis of the integration of the area under each peak against the molar ratio (Figure 1.17C) enables the estimation of the K_D and thermodynamic parameters such as enthalpy and entropy changes.[87]

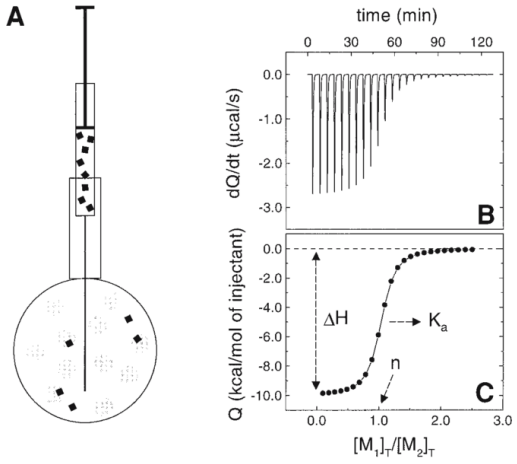


Figure 1.17. Schematic of an ITC experiment. A) Sample cell and syringe injecting the titrant (black squares) into the cell; B) Graph of the power supplied to the sample cell at regular time intervals following injection of the titrant; C) Graph of the integration under the peaks against the molar ratio following each injection enabling the determination of thermodynamic parameters. Reprinted with permission from Velaquez-Campoy *et al.*, Protein-Protein Interactions.[87] Copyright 2004 Springer.

As ITC is a very sensitive technique, it is important to ensure that the buffer of both the titrant and the sample in the cell is the same. This is because buffer mismatch, particularly in the pH or salt concentration, will affect the graphs obtained and affect the reliability of the data.[88] Therefore, it is essential to buffer exchange the protein of interest from storage buffer to the desired ITC buffer, and to supplement this with dimethyl sulfoxide (DMSO) to match the DMSO concentration resulting from the dilution of compounds stored in DMSO.

Another key consideration in the set-up of an ITC experiment is the *c*-value ($c = K_D \times [M]$, where [M] is the concentration of binding sites of the macromolecule in the cell).[88] The *c*-value characterises the binding sigmoidal isotherm and a value between 1 and 1000 is recommended.[86] When the *c*-value is too high, the transition is too sharp as saturation occurs too quickly, resulting in inaccurate determination of the binding constant. In contrast, when the *c*-value is too low, the transition is too broad, resulting in difficulty identifying the equivalence point.[86] Consequently, the K_D of weakly-binding ligands are

difficult to determine by ITC as a much higher concentrations of protein are required to obtain a binding isotherm from which the binding affinity can be reliably calculated.

1.7 Previous Work and Aims

In previous work by Dr Angelina Sekirnik,[89] the AlphaScreen™ was developed for TRIM24 and TRIM33α with three peptides – H3K9Me$_3$, H3K18Ac, H3K9Me$_3$K18Ac. This was then used to screen the Protein-Protein Interaction (PPI-net) collection of compounds as well as a focused library of 200 compounds from the Maybridge fragment library pre-selected by docking. This resulted in the identification of seven hits (Figure 1.18A) and four hits (Figure 1.18B), respectively.

Consequently, the aim of this project is to validate these hits and use them as a starting point for the development ligands for the TRIM33 bromodomain. Optimised ligands would enable a deeper understanding of the TRIM33 bromodomain and its cellular functions. This will be achieved through:

- Optimisation of the AlphaScreen™ assay for TRIM33β and characterise the protein-peptide interactions that form the basis of this assay.
- Validation of hits followed by screening of a focused library of compounds based on the hits.
- Development of a ligand in iterative cycles of design, synthesis and evaluation.
- Development of a covalent ligand that would enable detailed study of ligand binding mode with respect to the binding pocket.

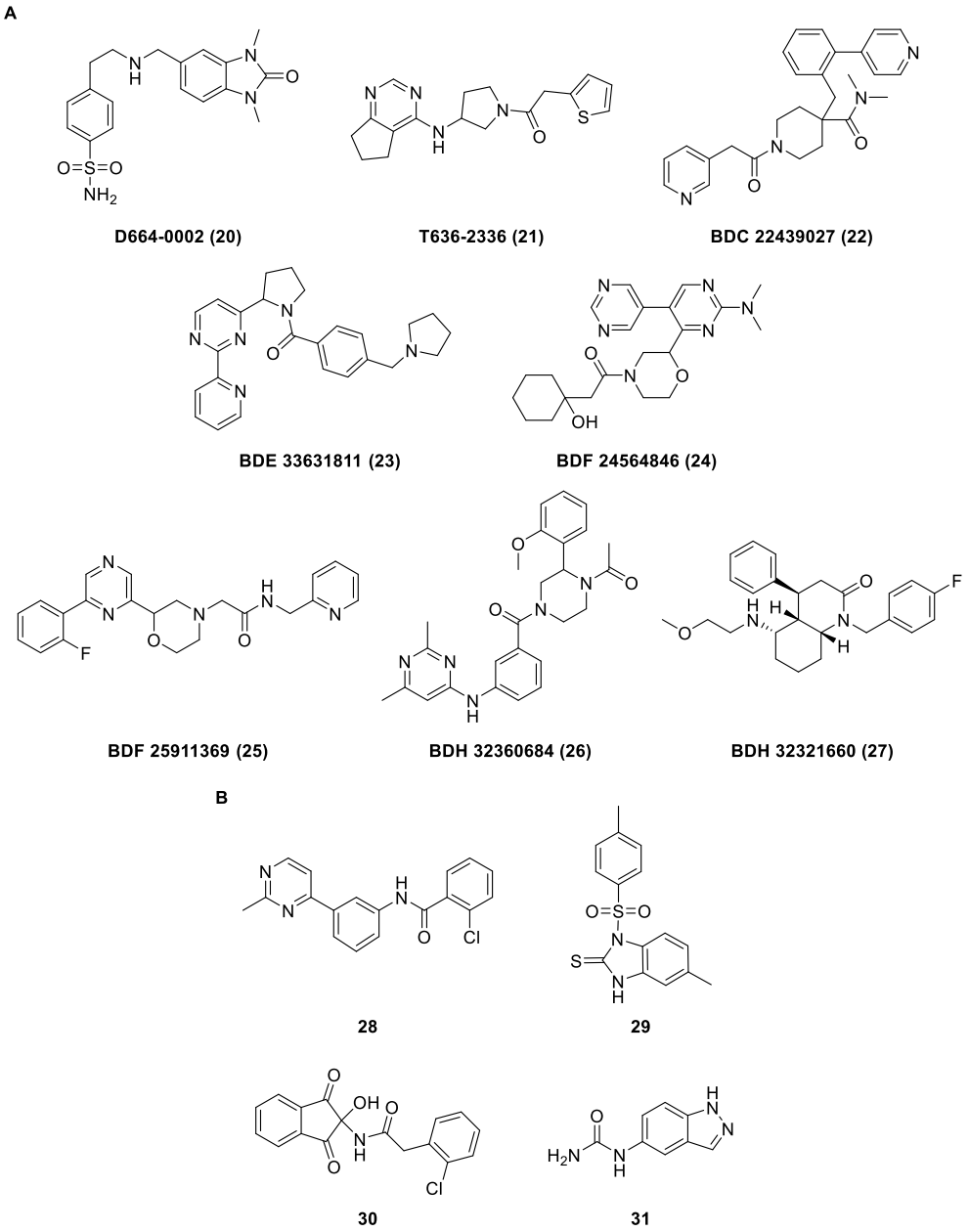


Figure 1.18. Hit compounds identified in AlphaScreen™. A) Compounds from PPI-net library; B) Compounds from Maybridge fragment library.

Chapter 2

Biological Characterisation

2.1 Introduction and Aims

As discussed in Section 1.7.1, the AlphaScreen™ assay set up by Dr Angelina Sekirnik[89] was intended to be used as a biological assay to screen for compounds that interfere with the protein-peptide interaction, and has previously been used for TRIM24 and TRIM33α. Following the development of a suitable construct for TRIM33β (Conway and Sekirnik, unpublished data), the aim of this work was to optimise the AlphaScreen™ assay conditions for TRIM33β to be used in compound screening. With a known TRIM24 bromodomain ligand **7** (provided by the SGC),[61] this study aims to assess the peptides used in the AlphaScreen™ assay as a means of probing compound binding to the individual PHD finger domain or bromodomain. As the AlphaScreen™ assay is a peptide displacement assay, the measured IC_{50} value is a function of the protein-peptide affinity. Consequently, this work aims to characterise the protein-peptide interactions of interest using isothermal titration calorimetry (ITC).

2.2 The AlphaScreen™ Assay

Based on known interactions between the proteins of interest (TRIM24 and TRIM33α) and their histone peptide binding partners (H3K9Me_3 – binds to the PHD finger, H3K9K18Ac – binds to the bromodomain, and H3K9Me_3K18Ac – binds to both), an AlphaScreen™ assay was developed by Dr Angelina Sekirnik.[89] To screen compounds against TRIM33β, the AlphaScreen™ would have to be set up for TRIM33β as well.

2.2.1 Expression and Purification of TRIM33β

The TRIM33β plasmid was transformed into *Escherichia coli* BL21 (DE3) cells and a single colony was used to inoculate a starter culture. This was then used for a large-scale recombinant protein expression with induction by isopropyl β-D-1-thiogalactopyranoside

(IPTG). The media was also supplemented with zinc chloride, which is required for the proper folding of the PHD finger. Protein purification using immobilised metal affinity chromatography (IMAC) followed by size exclusion chromatography (SEC) on an ÄKTA fast protein liquid chromatography (ÄKTAFPLC) system yielded TRIM33β. The purity of the protein was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and protein mass spectrometry (Figure 2.1).

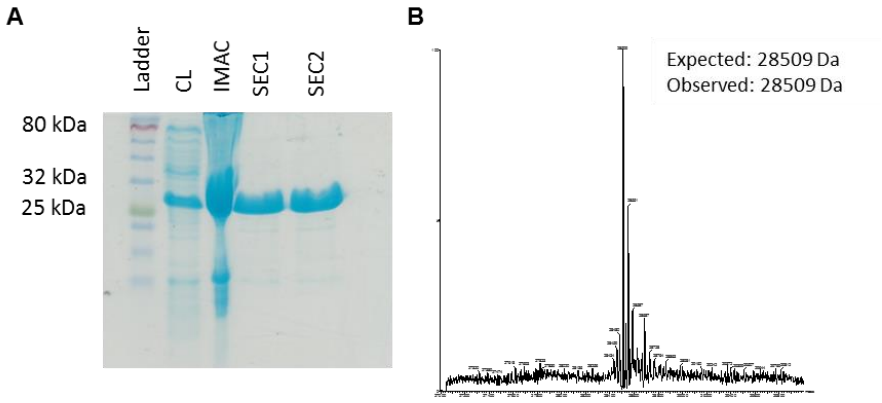


Figure 2.1. Expression of mutant protein TRIM33β-E984C. A) SDS-PAGE gel showing purity of TRIM33β and its size relative to the protein marker (ladder), where CL = cell lysate; B) Protein mass spectrum of TRIM33β.

2.2.2 Optimisation of AlphaScreen™ Assay Conditions

During assay development, it was found that the bead concentration required was less than the recommended concentration of 20 µg/mL. Instead, a 1:400 dilution (final concentration 5.0 µg/mL) or 1:300 dilution (final concentration 6.7 µg/mL) was sufficient for TRIM24 and TRIM33α respectively. The lower bead concentration tolerated for TRIM24 is due to the higher affinity of TRIM24 for the peptides. Following the successful expression of a TRIM33β protein construct, it was decided that for ease of compound screening, the bead concentration would be standardised at 6.7 µg/mL for all three proteins. As such, a signal response experiment was repeated for all three proteins with each peptide – a total of nine protein-peptide combinations. In this experiment, protein and peptide concentrations were varied so as to determine the appropriate final assay concentration (FAC) to be used in future AlphaScreen™ assays. For each protein-peptide combination, a graph of emission

against peptide concentration was plotted for three different protein concentrations (Figure 2.2, Figure B1 in the Appendices).

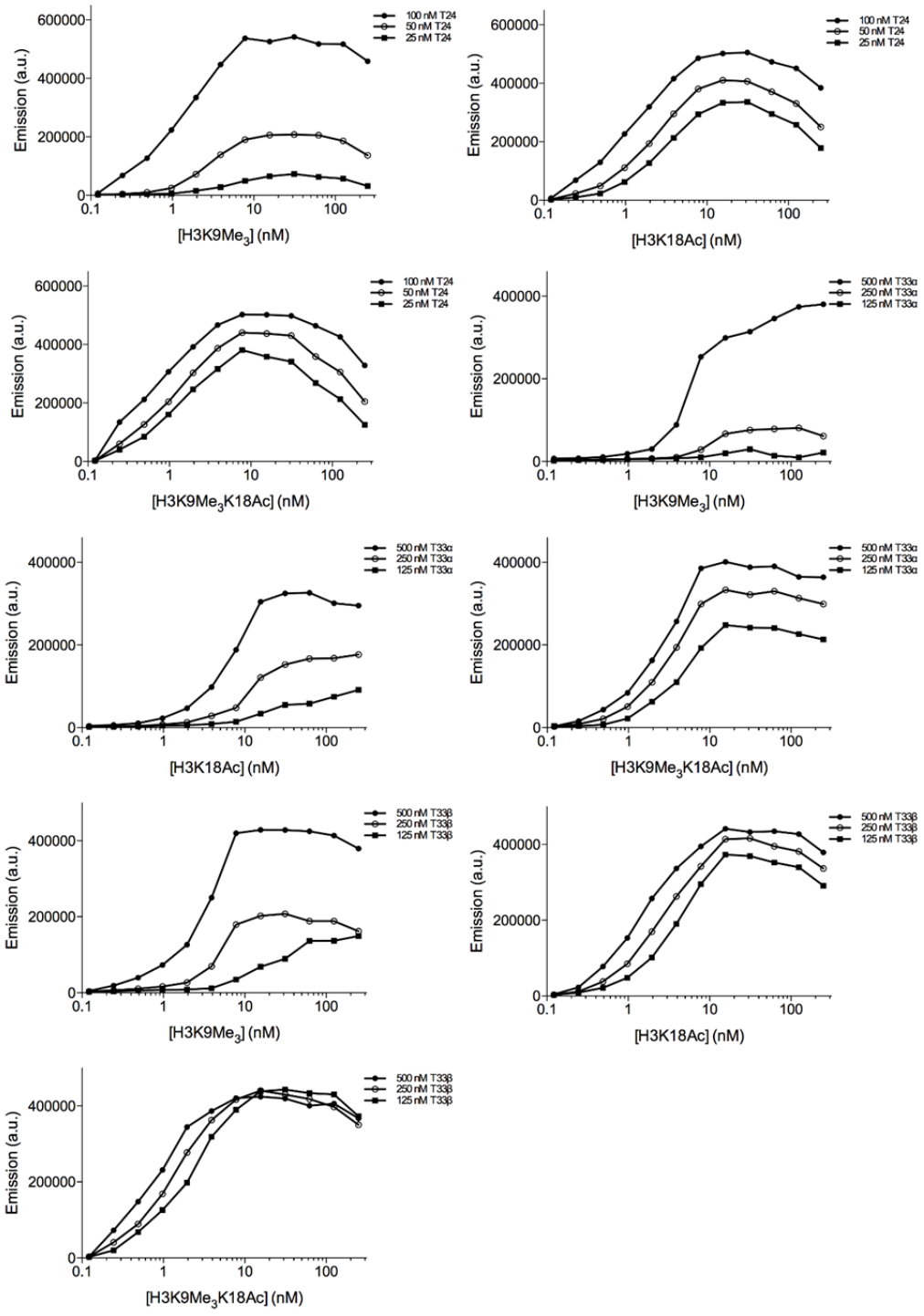


Figure 2.2. AlphaScreen™ signal response curves. Graphs of emission were plotted as a function of peptide concentration for each combination. Each protein (T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β) was tested at three different concentrations in combination with all three peptides ($H3_{1-27}K9Me_3$, $H3_{1-27}K18Ac$ and $H3_{1-27}K9Me_3K18Ac$) in a 1:2 serial dilution.

As the protein or peptide concentration increases, the emission observed also increases due to a larger amount of protein or peptide available to bind to the beads. However, the hook effect, where the emission intensity decreases at high peptide concentration, is clearly observed in the signal response curves. This is because at high peptide concentration, the beads are saturated and the amount of unbound peptide increases. The unbound peptides compete with the bead-bound peptides for binding to the protein but do not result in any emission, resulting in an observed decrease in emission intensity.

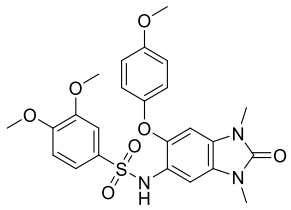
The protein and peptide FACs selected for each combination is summarised in Table 2.1 below. In selecting the assay conditions, it is important to consider the assay window, which results from a sufficiently high signal-to-noise ratio that would allow ligands to be accurately identified. Hence, to maximise the sensitivity of the assay, the concentrations were selected to achieve emission intensity of approximately 300000 to 500000 a.u. An emission intensity that is too high will limit the ability of the assay to detect weakly bound ligands, and result in a higher number of false negatives. Conversely, an emission intensity that is too low will increase the number of false positives due to difficulties in identifying decreases in emission intensity caused by actual binding. Furthermore, a phenomenon known as the edge effect, where some of the solution in wells along the edge of the assay plate evaporates, was observed. This results in inconsistencies in volume and concentration, and leads to errors in the emission recorded and hence, IC_{50} values calculated. To avoid this, the rows and columns closest to the edge were filled with buffer, forming a wall that limits the impact of the edge effect on the results obtained.

Table 2.1. Summary of protein and peptide FAC for AlphaScreen™ assay.

Protein	Peptide	Peptide FAC (nM)	Protein FAC (nM)
TRIM 24	$H3K9Me_3$	4	100
	H3K18Ac	8	25
	$H3K9Me_3K18Ac$	4	25
TRIM33α	$H3K9Me_3$	15	500
	H3K18Ac	15	500
	$H3K9Me_3K18Ac$	8	250
TRIM33β	$H3K9Me_3$	8	500
	H3K18Ac	15	125
	$H3K9Me_3K18Ac$	8	125

2.3 AlphaScreen™ as a Technique for Probing Individual Reader Domains

TRIM24 and TRIM33 contain both a PHD finger domain and a bromodomain, giving rise to the possibility of a combinatorial readout of histone H3 modifications. This could lead to a context-dependent cellular response, which is particularly interesting, and having ligands of the individual domains would greatly enable the study of the biological effects arising from the individual or dual modifications. The flexibility of setting up AlphaScreen™ assays for specific protein-peptide combinations presents it as a useful tool to screen compounds for affinity to the PHD finger, bromodomain or a region that could affect histone binding to both domains. This can be done simply by changing the peptide binding partner used in the assay. In order to investigate this potential, an AlphaScreen™ assay was carried out with all three peptides to obtain dose response curves for a known TRIM24 bromodomain ligand[61] provided by the SGC (Figure 2.3).



7

Figure 2.3. TRIM24 ligand 7.[61]

Due to the low solubility of compound **7** in aqueous buffer, the maximum FAC used was 100 μM. The dose response curves obtained are shown in Figure 2.4. The compound inhibits the TRIM24-H3K18Ac interaction (IC_{50} = 1.29 ± 0.19 μM) and TRIM24-H3K9Me$_3$K18Ac interaction (IC_{50} = 1.75 ± 0.34 μM) but not the TRIM24-H3K9Me$_3$ interaction. Given that H3K18Ac binds to the bromodomain region, and H3K9Me$_3$ binds to the PHD finger domain, these data show that ligand **7**, a bromodomain ligand, is unable to displace the peptide that binds mainly to the PHD finger domain. This result also suggests that choosing the appropriate peptide to screen against could provide information about the compound binding site. The dually-modified peptide H3K9Me$_3$K18Ac is presumed to bind both the PHD finger and bromodomain, and has a stronger affinity for the protein by ITC. Consequently, the peptide is less easily displaced.

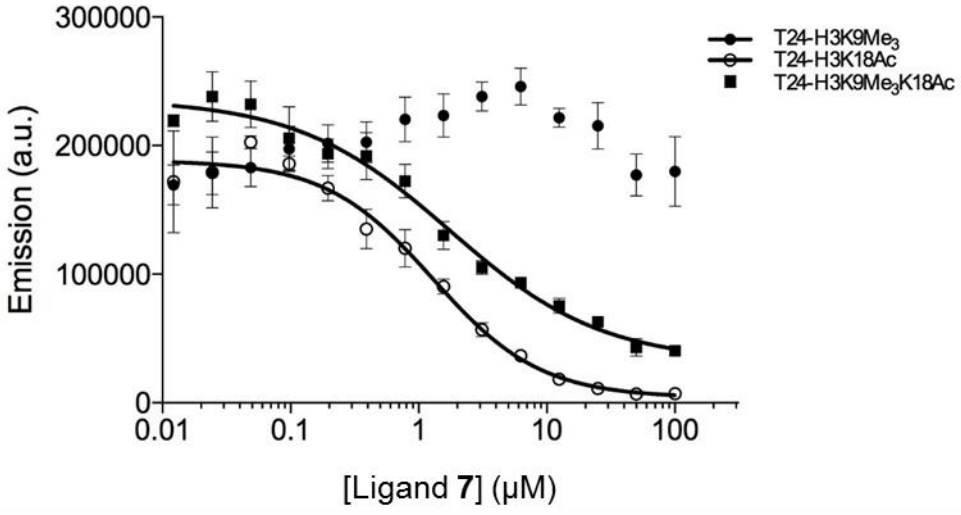


Figure 2.4. Dose response curves of ligand 7 with TRIM24.

Compound **7** was also tested against the TRIM33α and TRIM33β PHD-bromodomain cassettes (Figure 2.5) but showed no binding to either. This is not surprising given that the probe was developed for TRIM24. This also confirms that **7** cannot be used as a positive control for TRIM33.

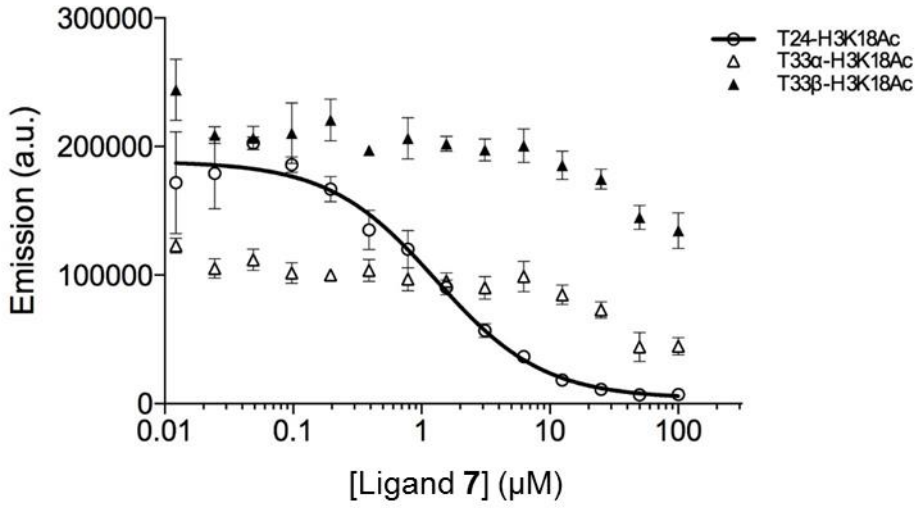


Figure 2.5. Dose response curves of ligand 7 with TRIM24, TRIM33α and TRIM33β.

2.4 Investigation of Protein-Peptide Interactions by ITC

As AlphaScreen™ is a peptide-displacement assay, the IC_{50} values obtained are dependent on the ability of the compounds tested to displace the peptide binding partner. Consequently, the results obtained are a function of the affinity of each peptide for each individual protein. To quantify the strength of the protein-peptide interaction using an orthogonal technique, ITC experiments were carried out to measure the binding affinity of TRIM24, TRIM33α and TRIM33β, each with the peptide binding partners (H3K9Me_3, H3K18Ac and H3K9Me_3K18Ac). All proteins and peptides used were dialysed overnight into HEPES buffer (20 mM HEPES, 150 mM NaCl) before use to ensure a good buffer match for the titration. HEPES buffer has a relatively high enthalpy of ionisation compared to inorganic buffers, which are usually preferred for ITC. However, the proteins are stable in

this buffer and heat of dilution experiments indicated that the HEPES buffer is suitable for this protein-peptide system.

Based on the preliminary results by Dr Angelina Sekirnik for TRIM24 and TRIM33α, each with H3K9Me$_3$ and H3K18Ac,[89] it was expected that the protein-peptide interactions with H3K9Me$_3$ would be much weaker. For ITC experiments, it is important to use estimated K_D and enthalpy values to guide design of a good ITC experiment. This is because the *c*-value of the isotherm, which relates to the shape of the sigmoidal curve, is dependent on the K_D and the concentration in the cell (i.e. protein concentration). As such, K_D and enthalpy values guide the ideal protein and peptide concentrations to use for each combinations (within the limits of the amount of material available). In some cases where the interaction is very weak, a much higher concentration is needed to achieve a *c*-value within the desired range. However, at high concentrations, complications arise from aggregation and dissociation effects, and should also be considered when deciding what concentrations to use. Consequently, the concentrations used for this set of experiments is 160 μM peptide with 16 μM protein for H3K18Ac and H3K9Me$_3$K18Ac, and 500 μM peptide with 50 μM protein for H3K9Me$_3$.

The titration curves obtained from the ITC experiments are shown below (Figure 2.6–Figure 2.8) and the results summarised in Table 2.2. Sigmoidal titration curves with clear equivalence points were obtained for TRIM24 due to the peptides binding most strongly to TRIM24. While the curves obtained for TRIM33β maintained the distinct sigmoidal shape, it was clear that TRIM33β does not bind as strongly to the same peptides. Sigmoidal curves were not achieved for TRIM33α and by using the preliminary K_D and enthalpy values to guide the design of possible follow-up experiments, it was suggested that the weak binding of TRIM33α, especially for the peptide H3K18Ac, greatly limits the possibility of obtaining a good binding curve and hence, more accurate data. This result is not surprising as TRIM33α has a 17 amino acid insertion as compared to TRIM24 and TRIM33β, which causes

the conserved asparagine of the bromodomain to point outward rather than into the bromodomain binding pocket. As such, it is unlikely to interact with the acetylated lysine residue with as high an affinity as TRIM24 and TRIM33β. The order of affinities is as expected based on observations in the AlphaScreen™ assay, with all proteins binding to the dually-modified peptide most strongly.

It was observed that the N-values obtained for all the experiments were below the expected value of 1. As ITC experiments are particularly sensitive to the cell and syringe concentrations, and calculations are based on the concentrations used, errors in the accuracy of protein and/or peptide concentration easily accumulate and result in errors in the N-values obtained. For the TRIM33α-H3K9Me$_3$ combination, however, the titration curve suggests that the binding is too weak to obtain a reliable fitting and hence, the N-value is much lower than that of the other combinations.

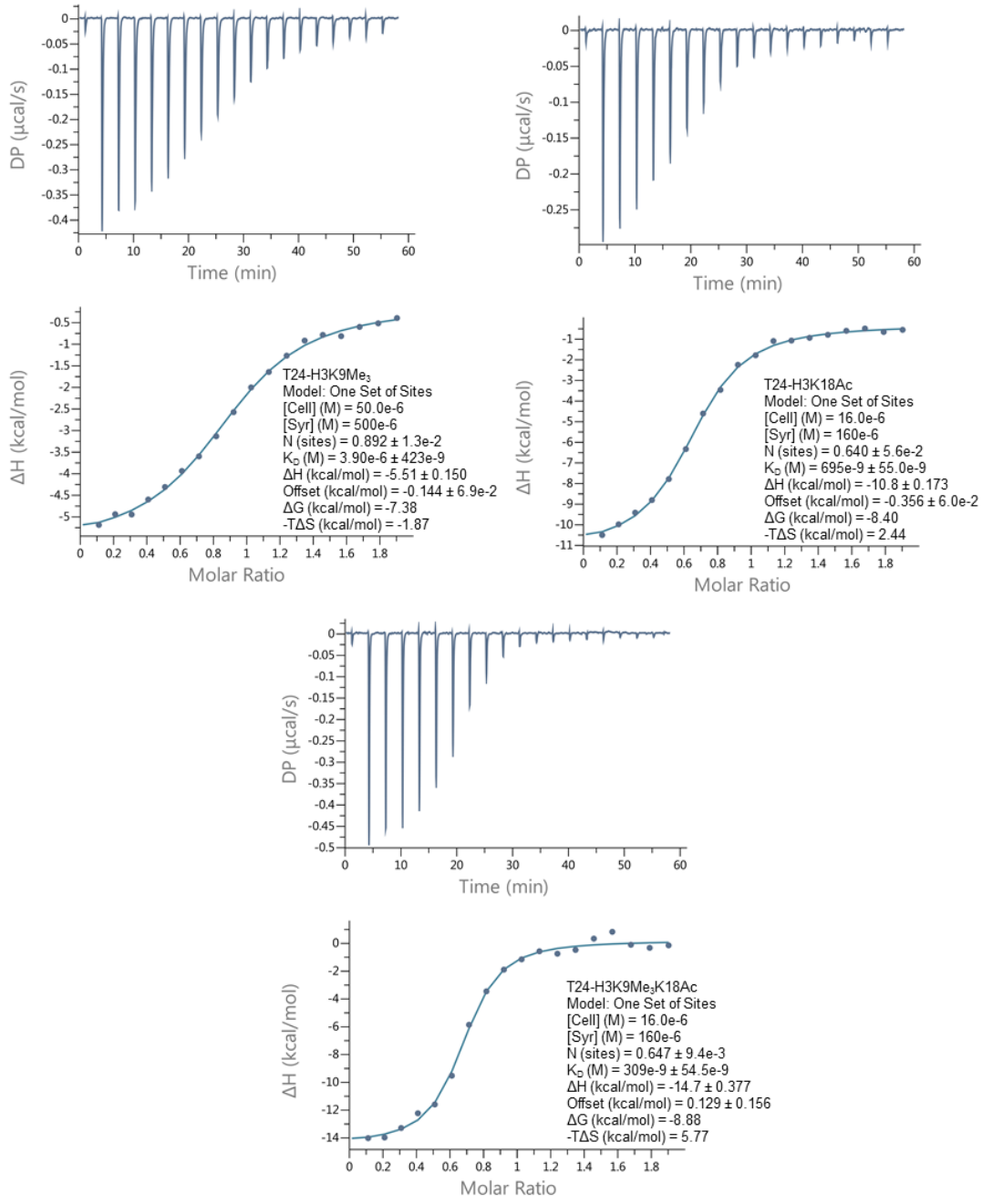


Figure 2.6. ITC titration curves of TRIM24 with the PHD finger-binding peptide (H3K9Me$_3$), the bromodomain binding peptide (H3K18Ac) and the dually-modified peptide (H3K9Me$_3$K18Ac).

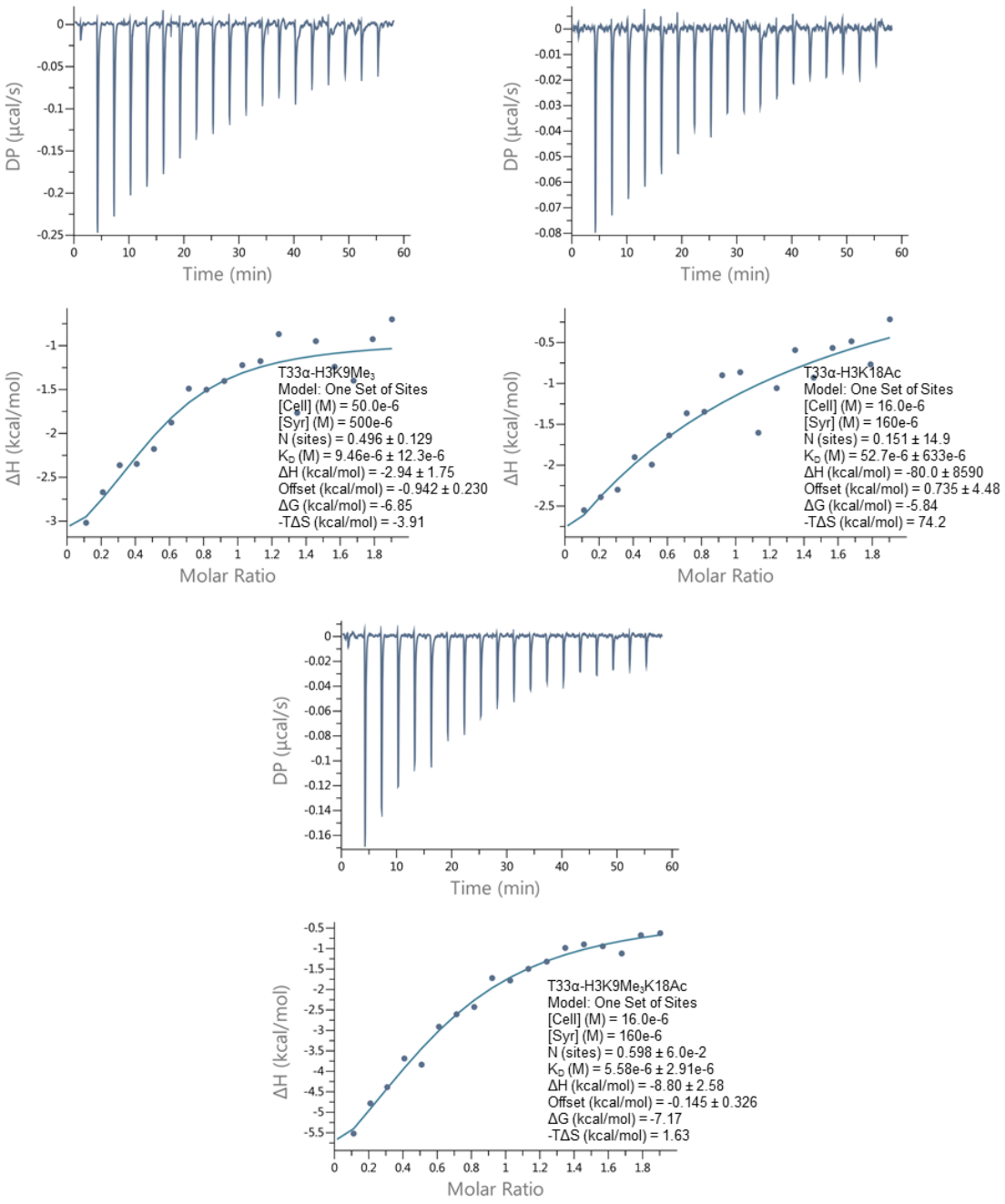


Figure 2.7. ITC titration curves of TRIM33α with the PHD finger-binding peptide (H3K9Me$_3$), the bromodomain binding peptide (H3K18Ac) and the dually-modified peptide (H3K9Me$_3$K18Ac).

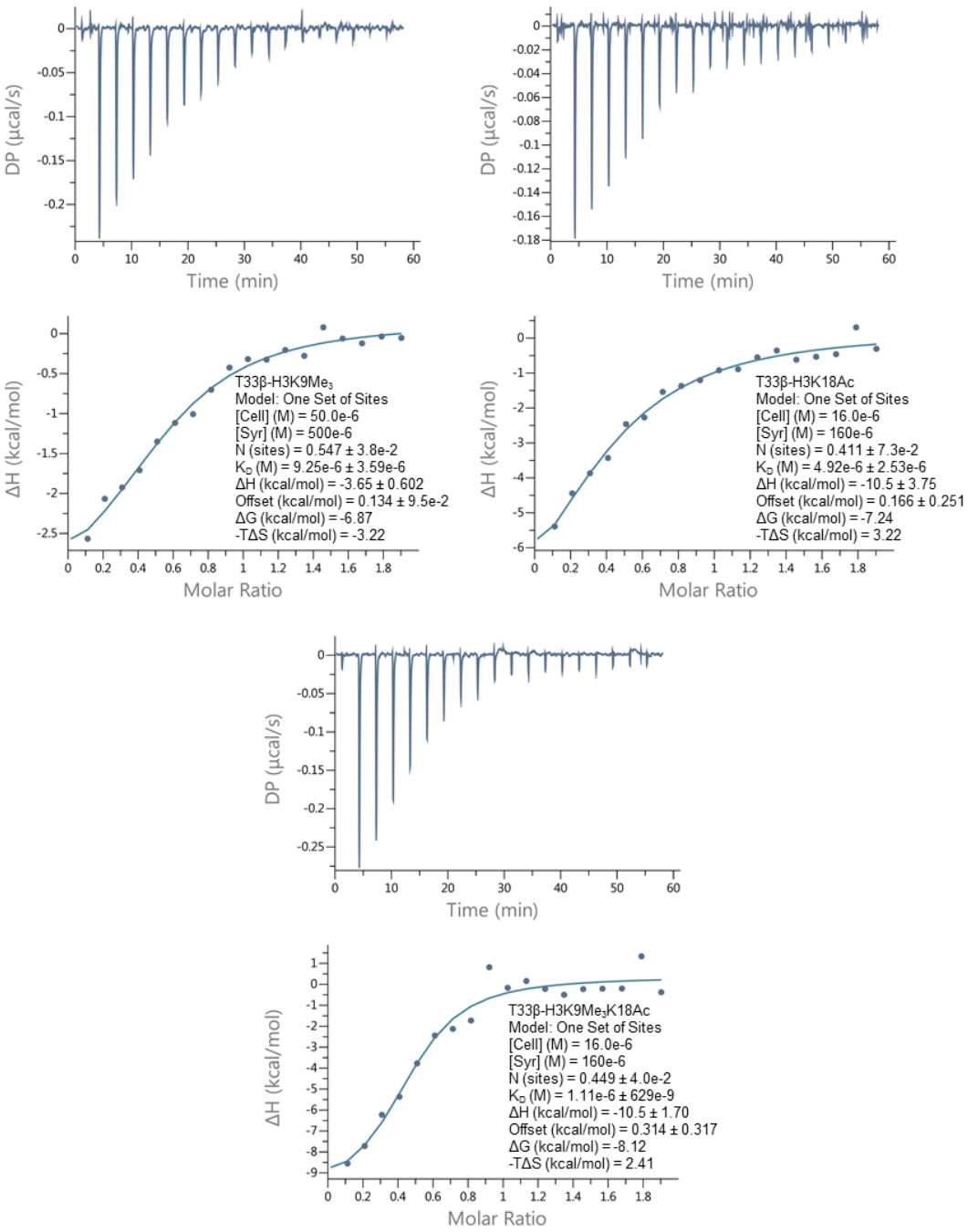


Figure 2.8. ITC titration curves of TRIM33β with the PHD finger-binding peptide (H3K9Me$_3$), the bromodomain binding peptide (H3K18Ac) and the dually-modified peptide (H3K9Me$_3$K18Ac).

Table 2.2. Summary of ITC results.

		$H3K9Me_3$	H3K18Ac	$H3K9Me_3K18Ac$
TRIM24	N	0.892	0.640	0.647
	K_D	3.90 μM	695 nM	309 nM
	ΔH (kcal/mol)	–5.51	–10.8	–14.7
	ΔG (kcal/mol)	–7.38	–8.40	–8.88
	-TΔS (kcal/mol)	–1.87	2.44	5.77
TRIM33α	N	0.496	0.158	0.598
	K_D	9.46 μM	54.7 μM	5.58 μM
	ΔH (kcal/mol)	–2.94	–80	–8.80
	ΔG (kcal/mol)	–6.85	–5.82	–7.17
	-TΔS (kcal/mol)	–3.91	74.2	1.63
TRIM33β	N	0.547	0.411	0.449
	K_D	9.25 μM	4.92 μM	1.11 μM
	ΔH (kcal/mol)	–3.65	–10.5	–10.5
	ΔG (kcal/mol)	–6.87	–7.24	–8.12
	-TΔS (kcal/mol)	–3.22	3.22	2.41

2.5 Conclusions

The AlphaScreen™ assay was optimised for TRIM24, TRIM33α and TRIM33β to determine the protein and peptide FAC that will be used in subsequent assays for screening compounds and assessing binding affinity. Using the TRIM24 ligand **7**, it was also demonstrated that peptides with differing modifications, trimethylated lysine and acetylated lysine, which bind specifically to the PHD finger domain and bromodomain respectively, can provide insight into the ligand binding region.

While the higher TRIM33 concentrations required for the AlphaScreen™ suggested a lower protein-peptide affinity compared to TRIM24, ITC experiments provided deeper understanding of the relative strength of binding. All three peptides bind most strongly to TRIM24. Notably, while the measured K_D of H3K9Me$_3$ to both isoforms of TRIM33 are similar, H3K18Ac binds more strongly to TRIM33β than TRIM33α. This difference between the peptide affinity for the proteins has implications in the AlphaScreen™ because the IC_{50} values obtained are dependent on the ability of the compound to displace the peptide binding partner. When evaluating a compound, the peptide H3K18Ac would be significantly easier to displace from TRIM33α than TRIM33β, resulting in lower compound concentrations needed to displace peptide H3K18Ac from TRIM33α, and therefore leading to a lower IC_{50} value. As such, when interpreting AlphaScreen™ data, it is important to bear in mind that the apparent strength of binding to TRIM33α as compared to TRIM33β is a result of the low affinity of H3K18Ac for TRIM33α.

Chapter 3

Hit Validation

3.1 Introduction and Aims

In an AlphaScreen™ performed by Dr Angelina Sekirnik,[89] eight compounds from the PPI-net collection and four from the Maybridge fragment library were identified as potential hits of TRIM24 and TRIM33α. The work demonstrated here aimed to validate the binding of these compounds to the TRIM24, TRIM33α and TRIM33β bromodomains. Initial screening saw 200 compounds from the Maybridge fragment library screened following docking studies. Consequently, this study aims to screen a focused selection of compounds from the Maybridge fragment library, selecting compounds based on their similarity to the initial hits, with the aim of exploring focused structure-activity relationships (SAR).

3.2 Follow-up of PPI-Net Collection Hit Compounds

Of the 8 PPI-net hits (Figure 3.1A), compound **27** was no longer commercially available, but the remaining seven were purchased for re-testing. A dose response AlphaScreen™ assay was carried out with all nine protein-peptide combinations using a maximum compound final assay concentration (FAC) of 500 μM. For compounds **21**, **22** and **24**, however, the maximum FAC used was 300 μM due to the limited solubility of these compounds. The solubility forecast index (SFI = $cLogD_{pH7.4}$ + no. of aromatic rings)[90,91] was calculated and is indicated on Figure 3.1A for each compound. Notably, compounds **21**and **24** are not far from the suggested guide of SFI < 5[90] and compound **22** has a high SFI of 6.42, indicating that it is possible for these compounds to be less soluble. While it is not appropriate to take SFI as an absolute measure of solubility, and the cut-off can vary between compound series, it is a useful guide for avoiding compounds that might not be reliably tested because of their solubility limitations. A representative set of dose response curves is shown in Figure 3.1B.

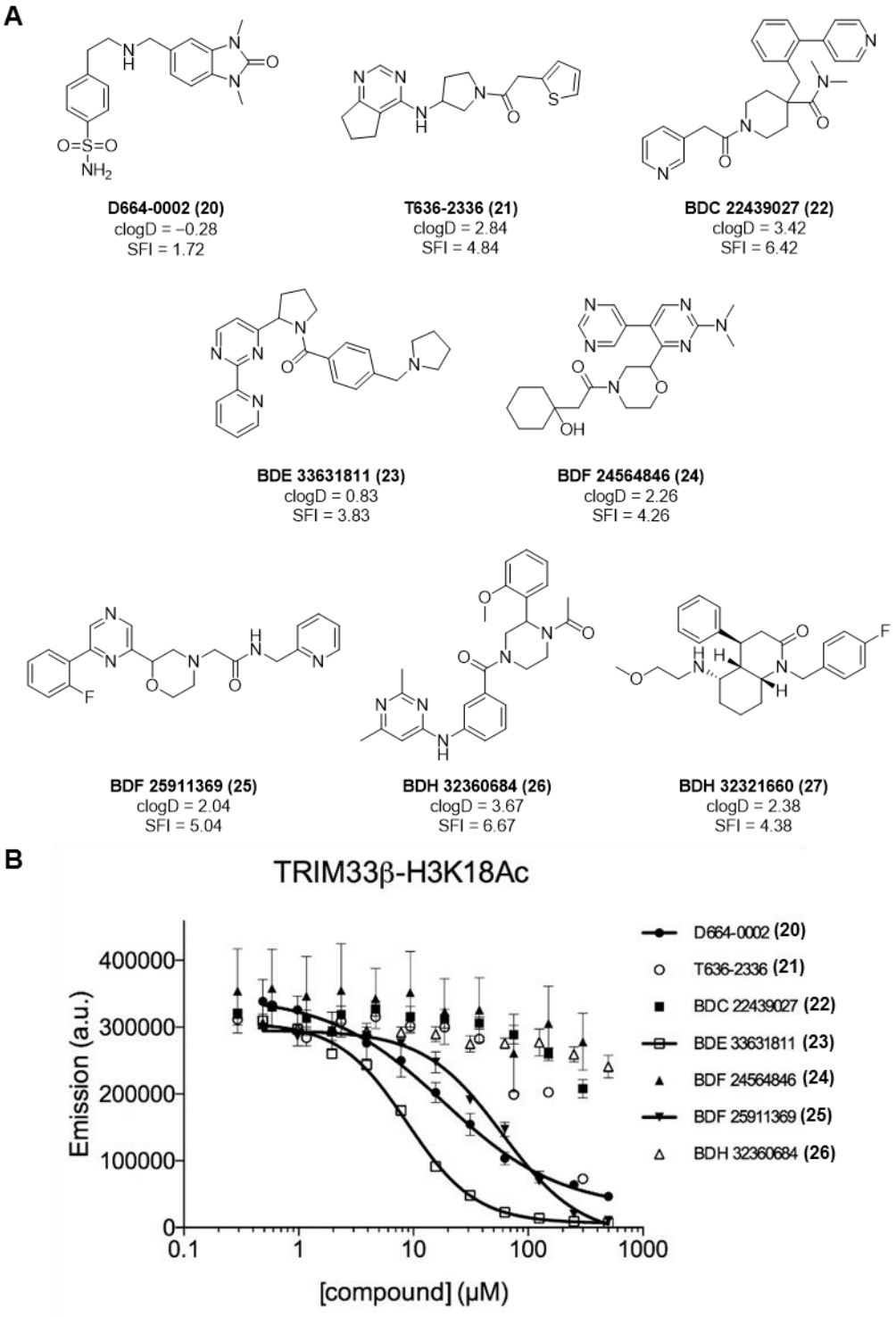
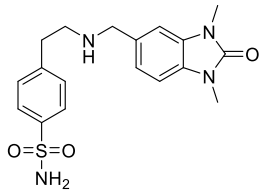
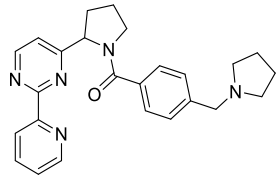
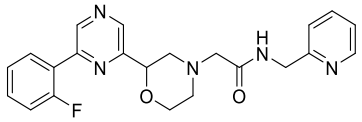


Figure 3.1. Follow-up of PPI-net collection compounds in dose-response AlphaScreen™ to determine IC_{50} values. A) PPI-net hit compounds. B) A representative example of the dose response curves obtained against TRIM33β with peptide H3K18Ac.

IC_{50} values were calculated and are summarised in Table 3.1. Compounds **21**, **22**, **24** and **26** showed little or no binding to all three proteins (TRIM24, TRIM33α and TRIM33β), and no IC_{50} curves were obtained for them. Compounds **20**, **23** and **25** were most effective in

disrupting the protein-peptide interactions. While IC_{50} values of compounds **23** and **25** could be obtained for all nine protein-peptide combinations, IC_{50} values of compound **20** could not be obtained for TRIM24, suggesting possible selectivity for TRIM33. Additionally, in comparison to when H3K18Ac was the peptide binding partner, compound **20** (IC_{50} (TRIM33α) = 2.79 μM, IC_{50} (TRIM33β) = 18.4 μM) did not significantly disrupt the interaction between TRIM33 and H3K9Me$_3$ (IC_{50} (TRIM33α) >250 μM, IC_{50} (TRIM33β) >250 μM). This suggests that compound **20** is binding to bromodomain region, which binds acetylated lysine H3K18Ac and not the PHD finger domain, which binds the trimethylated lysine H3K9Me$_3$. Notably, the benzimidazolone headgroup in compound **20** is also present in other known TRIM24 bromodomain ligands,[61,60] raising the likelihood of compound **20** binding in the acetyl-lysine binding pocket of the bromodomain. In contrast, when the binding peptide was H3K18Ac, compound **23** had similar IC_{50} values (IC_{50} (TRIM33α) = 7.69 μM, IC_{50} (TRIM33β) = 8.84 μM) while compound **25** demonstrated much weaker binding (IC_{50} (TRIM33α) = 51.8 μM, IC_{50} (TRIM33β) = 59.4 μM). Moreover, both compounds **23** and **25** bound to TRIM24 with IC_{50} values of 6.68 μM and 41.7 μM respectively.

Table 3.1. AlphaScreen™ IC_{50} values of PPI-net hit compounds. Colour shown as a heatmap; from low IC_{50} values (red) to high IC_{50} values (green); H3(PHD) – $H3K9Me_3$, H3(BRD) – H3K18Ac, H3(Dual) – $H3K9Me_3K18Ac$.

Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **20**	>250	>250	>250	>250	2.79 ± 0.83	14.1 ± 1.62	>250	18.4 ± 2.06	>250
 **23**	5.90 ± 0.28	6.68 ± 0.41	6.08 ± 0.43	10.3 ± 0.59	7.69 ± 0.62	5.41 ± 0.64	12.8 ± 0.78	8.84 ± 0.60	22.2 ± 1.56
 **25**	41.8 ± 5.11	41.7 ± 7.82	62.2 ± 4.15	43.2 ± 5.67	51.8 ± 3.87	56.5 ± 10.3	78.6 ± 9.27	59.4 ± 5.73	111 ± 13.5

3.3 Follow-up of Maybridge Library Hit Compounds

3.3.1 Selection of Compounds

Of the 200 compounds tested from the Maybridge fragment library, four compounds were found to bind to TRIM24 or TRIM33α (Figure 3.2).[89] To investigate SAR around these four hits as a starting point for further ligand development, a small focused library of 50 compounds was selected from the Maybridge library (Figure B2 in the Appendices). To aid selection, the software OpenBabel[92] was used to rank the Maybridge fragment library of 30000 compounds based on their chemical similarity (using the Tanimoto coefficient)[93] with respect to each of the four compounds. The Tanimoto coefficient is an index that assesses and quantifies two-dimensional (2D) chemical similarity between compounds, assigning them a value from 0 to 1, with 1 being exactly the same as the query compound. To avoid solubility issues affecting assay reliability, SFI[90,91] values were calculated for the shortlisted compounds and those with SFI > 6.5 were discarded.

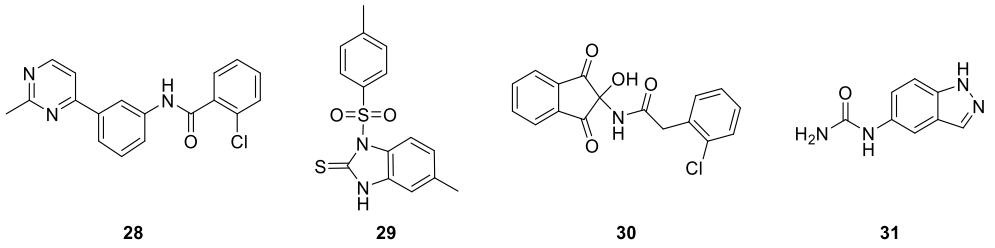


Figure 3.2. Hit compounds from the Maybridge fragment library identified previously in an AlphaScreen™.[89]

Compounds containing a ninhydrin structure, such as compound **30**, were ignored as attempts to re-synthesise it were met with difficulties in purification due to the ease of hydrolysis. Therefore, this compound is likely to have hydrolysed under the aqueous assay conditions. Notably, Jayatunga *et al.* reported that similar compounds were shown to inhibit the hypoxia-inducible factor 1-alpha/histone acetyltransferase p300 (HIF1-α/p300) interaction, but were later found to undergo hydrolysis in aqueous conditions.[94] Further studies confirmed that ninhydrin was the active species, interfering with the zinc ions in the protein.[94] Given that TRIM24, TRIM33α and TRIM33β contain two zinc ions associated with

the PHD finger, the positive result observed for compound **30** in the initial screen is likely to be due to a similar effect.

3.3.2 Screening of Focused Selection by AlphaScreen™

The 50 compounds selected were purchased and tested in triplicate against all 9 protein-peptide combinations at 30 μM and 150 μM FAC. Based on previous observations, most compounds would be soluble in the aqueous buffer at 150 μM. The percentage reduction in fluorescence emission in the assay was calculated for each compound with reference to the DMSO control for the respective protein-peptide combination (Table B1 in the Appendices). Compounds that showed a dose-dependent decrease in emission and resulted in ≥50% reduction in emission intensity (i.e. emission less than 50% of the DMSO control) were taken on for further validation. Compounds **29**, **32**, **33**, **34**, **35**, **36**, **37** and **38** were selected for further validation (Figure 3.3). Compounds **39** and **40** were discarded due to the presence of the ninhydrin motif and the consequent likelihood of them interacting with TRIM24 and TRIM33 *via* binding to the zinc ions in the PHD finger domain.

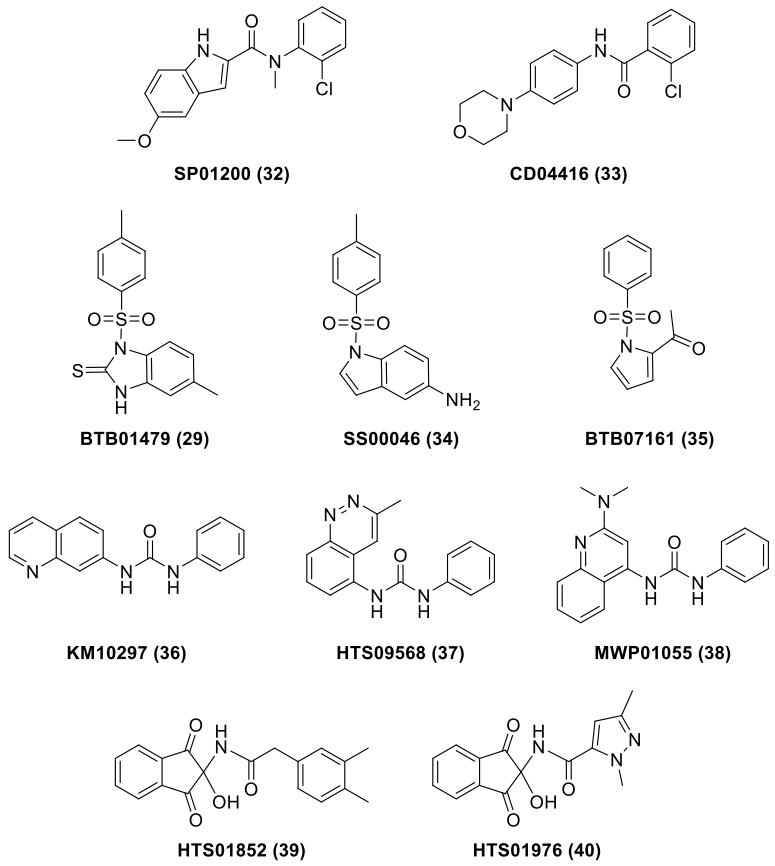


Figure 3.3. Hit compounds from expanded Maybridge selection.

3.4 Further Validation Using Complementary Experiments

3.4.1 AlphaScreen™ TruHits Kit

As AlphaScreen™ is a fluorescence-based assay, compounds interfering with this would result in an apparent reduction in emission and a false positive. Examples of such compounds include singlet oxygen quenchers, fluorescence quenchers or light scatterers.[81] The AlphaScreen™ TruHits kit was used to identify compounds that resulted in a reduced emission intensity due to assay interference rather than actual protein-ligand binding. Similar to the AlphaScreen™ assay, a streptavidin donor bead is used. However, instead of a nickel-chelating acceptor bead, a biotinylated acceptor bead is provided. The strong streptavidin-biotin interaction ensures that the beads will be brought in close enough proximity for emission to be observed (Figure 3.4). Consequently, any reduction in fluorescence results from assay interference.

By pre-incubating both beads before addition of the compound of interest, compounds resulting in a reduction of emission by more than two standard deviations with respect to a DMSO control can be identified as false positives. Compounds can also first be pre-incubated with the streptavidin donor beads before addition of the biotinylated acceptor beads. This enables screening for compounds that are potentially biotin-mimetic, which in the AlphaScreen™ assay would prevent biotinylated peptides from binding to the streptavidin donor beads.

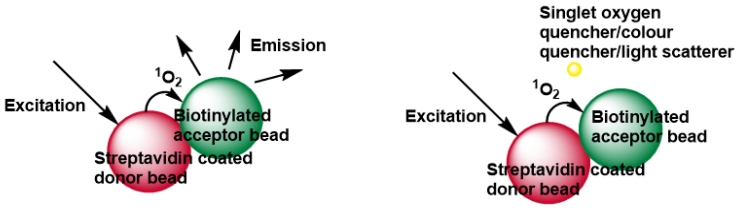


Figure 3.4. Schematic diagram of AlphaScreen™ TruHits kit. A streptavidin coated and biotinylated donor and acceptor bead, respectively, ensure that the beads interact with each other, resulting in emission. In the presence of a compound that interferes with the fluorescence process, no emission is observed.

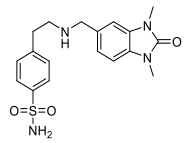
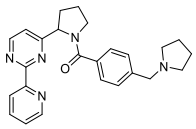
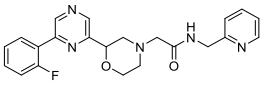
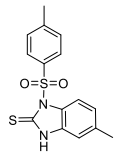
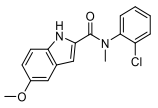
3.4.1.1 *Screening Hits for Assay Interference*

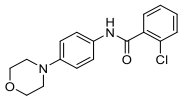
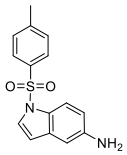
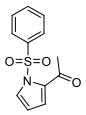
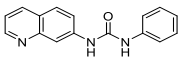
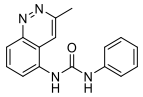
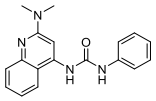
The three hits from the PPI-net collection and the eight hits from the Maybridge library were tested in a TruHits experiment. Compounds were evaluated using both incubation strategies at five different concentrations in 1:2 serial dilutions with 150 µM as the maximum FAC. For the second incubation method, a biotin control was also included with a maximum FAC of 5 µM. The difference between the emission with the compounds and that of the DMSO control was then calculated and normalised against the standard deviation of the DMSO control (Table 3.2).

From this experiment, it is clear that compounds **29**, **32**, **33**, **34**, and **38** (Figure 3.3) interfered with the fluorescence output of the assay. The dose-dependent trend observed further confirms that these compounds are false positives and were not binding to TRIM24 or TRIM33. Particularly for compounds **29**, **34** and **38**, the results of the AlphaScreen™

experiment in Section 3.3.2 where all three compounds showed a large reduction in emission across all nine protein-peptide combinations correlates well with the large difference (up to approximately 15 times the standard deviation of the DMSO control) observed here. While there is a significant difference recorded for the second pre-incubation method, this is likely to be due to compound interference in the chemiluminescent signal rather than with the biotin-streptavidin interaction between the beads (Table B2 in the Appendices).

Table 3.2. Results from AlphaScreen™ TruHits experiment. Differences of greater than two standard deviations are indicated in pink.

Compound	Number of standard deviations from DMSO average, normalised against standard deviation of DMSO control				
	Concentration				
	9.38 µM	18.8 µM	37.5 µM	75.0 µM	150 µM
 **20**	0.19	0.68	0.58	0.67	0.33
 **23**	0.73	1.03	1.33	1.55	1.65
 **25**	0.56	1.28	1.30	1.31	0.86
 **29**	2.99	3.50	6.50	10.66	16.43
 **32**	1.10	2.30	3.42	4.74	5.58

 **33**	1.15	0.12	0.44	0.82	2.04
 **34**	0.74	2.76	5.73	9.18	13.09
 **35**	1.30	0.10	0.55	0.35	0.77
 **36**	0.31	0.14	0.55	0.96	1.54
 **37**	0.75	0.11	0.31	0.81	0.86
 **38**	2.44	4.39	7.07	10.59	14.61

Based on the TruHits experiment, compounds **29**, **32**, **33**, **34**, and **38** were taken as false positives resulting from assay interference and compounds **20**, **23**, **25**, **35**, **36**, and **37** progressed for further validation.

3.4.2 NMR WaterLOGSY as an Orthogonal Technique

Water-ligand observed *via* gradient spectroscopy (waterLOGSY) is a nuclear magnetic resonance (NMR) technique used to identify ligands that bind to the target protein and has the advantage of being a direct observation method requiring neither ligand- or protein-labelling,[95] nor the use of reporter molecules. It is a ligand-observed method based on the saturation of bulk water magnetisation, followed by transfer of its magnetisation *via*

the intermolecular nuclear Overhauser effect (NOE) to detect ligand-protein interactions (Figure 3.5).[96,97] Ligands that bind to the protein experience strong negative NOEs whereas ligands that do not bind experience weaker and positive transfer.[98] Consequently, in the presence of protein, binding ligands exhibit a positive (or less negative) spectrum whereas non-binding ligands exhibit a negative spectrum identical to that recorded in the absence of protein.[99] As such, it is important to record a spectrum of the ligand without protein as a control for valid comparison. This is particularly important because compounds that have limited solubility and form soluble aggregates exhibit a positive spectrum, causing false positives.[100]

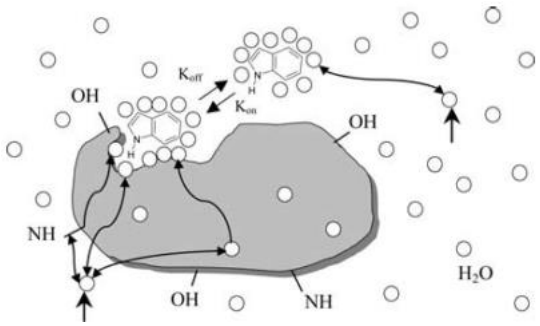


Figure 3.5. Schematic diagram of waterLOGSY. Excitation of bulk water (circles) results in transfer of magnetisation directly to the ligand or *via* the protein. Reprinted with permission from Dalvit *et al.*, Journal of Biomolecular NMR.[83] Copyright 2001 Springer.

3.4.2.1 *Compound Screening by NMR WaterLOGSY*

Following the AlphaScreen™ TruHits experiment, compounds **20**, **23**, **25**, **35**, **36** and **37** were progressed for further validation using waterLOGSY as an orthogonal technique. In the waterLOGSY experiment, a water-suppressed proton NMR is first recorded followed by a waterLOGSY NMR with the compound (500 μM) in the absence of protein. Ligand peaks are negative in the absence of protein, showing that there is no compound aggregation. In some cases, a positive spectrum was observed in the absence of protein (Figure 3.6). This indicates the formation of soluble aggregates that are not visible in the form of precipitation.[100] The larger size that results from compound aggregation mimics the NOEs experienced by ligands bound to protein. It is important to obtain a negative spectrum in

the absence of protein so that a positive spectrum in the presence of protein can be attributed to compound binding and not compound aggregation. Consequently, these were tested at lower concentrations in order to establish a concentration with no compound aggregation.

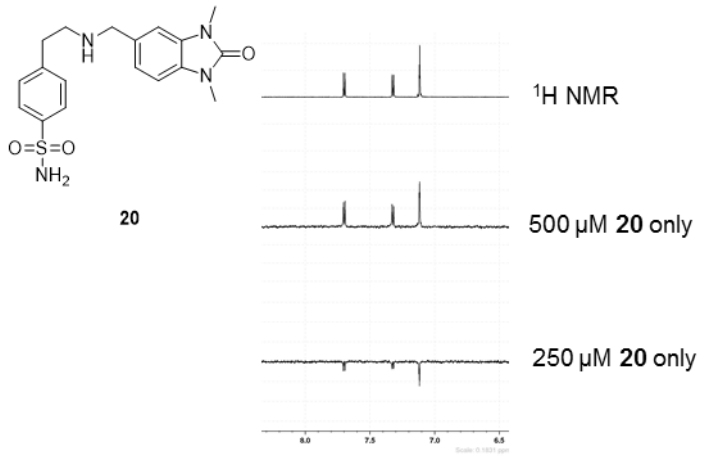


Figure 3.6. NMR waterLOGSY spectra of compound 20 in the absence of protein. Sample prepared in PBS buffer. Aggregation is observed at 500 µM, as indicated by a positive spectrum. This effect is removed by dilution to 250 µM.

WaterLOGSY experiments were then carried out individually with 20 µM of each protein, which had been buffer exchanged from HEPES storage buffer into phosphate-buffered saline (PBS buffer). An inorganic buffer was preferred to avoid the appearance of peaks resulting from HEPES buffer that would further complicate the waterLOGSY spectra. An example of the spectra obtained for compound **25** is shown in Figure 3.7. Ligand peaks (highlighted in blue) remained negative in the presence of TRIM24 but became inverted in the presence of TRIM33α and TRIM33β, indicating that compound **25** binds to TRIM33α and TRIM33β but not TRIM24. Labile protons (highlighted in green) experience transfer of magnetisation through chemical exchange[83] and hence result in a positive signal even in the absence of protein. Differences in the extent to which the compound peaks change on binding can be attributed to the binding mode of the compound, which unfortunately cannot be easily probed by NMR waterLOGSY. Whether part of the compound is buried in a protein pocket or exposed to the bulk solvent on the protein surface affects the way magnetisation

from surrounding water is transferred to the protons on the compound. This contributes to the slight difference in signal intensity of the ligand peaks that is sometimes observed.

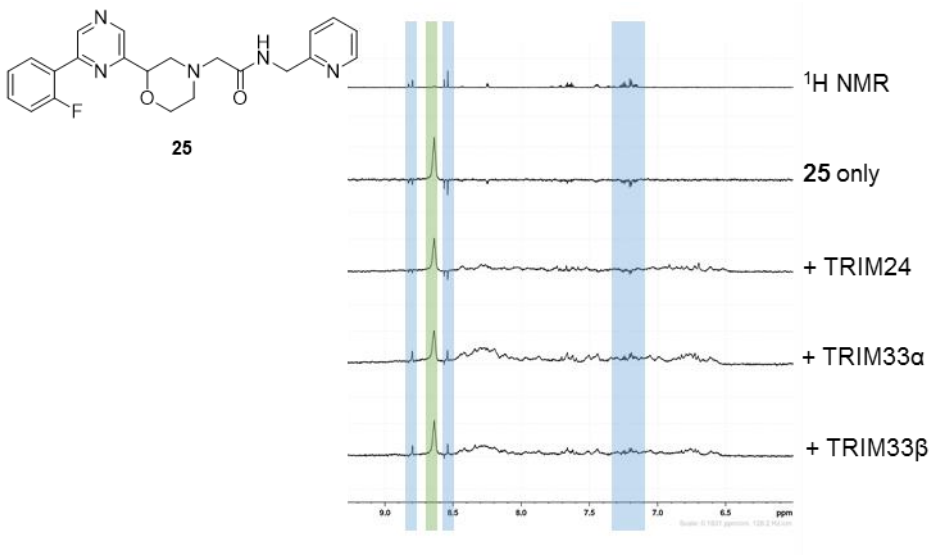
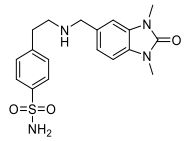
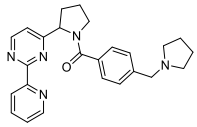
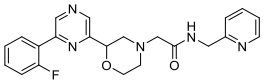
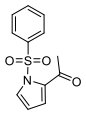
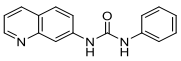
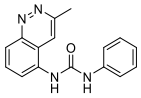


Figure 3.7. NMR waterLOGSY spectra of compound 25. Ligand peaks are highlighted in blue, labile proton peaks are highlighted in green. Sample prepared in PBS buffer, with 500 μM compound and 20 μM protein.

The results are summarised in Table 3.3 below, with the compound concentration used indicated. Where the compound concentration is lowered, the binding is less clear due to the weaker signal obtained, particularly in the absence of protein. For compound **35**, the low concentration resulted in difficulty distinguishing compound peaks from the protein peaks and hence, the results are inconclusive by NMR waterLOGSY.

Table 3.3. Results of NMR waterLOGSY experiments. Compounds that bind by waterLOGSY are indicated in pale blue; ✓ – binding, × – not binding, ~ – inconclusive.

Compound	Concentration (μM)	TRIM24	TRIM33α	TRIM33β
 **20**	250	✓	✓	✓
 **23**	500	×	×	×
 **25**	500	×	✓	✓
 **35**	150	~	~	~
 **36**	150	✓	✓	✓
 **37**	300	✓	✓	✓

All compounds except compounds **23** and **35** were confirmed to bind to both TRIM33α and TRIM33β. On closer inspection of the AlphaScreen™ TruHits experiment data (Table 3.2), the difference in emission as compared to the DMSO control follows an increasing trend with increasing compound concentration (up to 150 μM, the highest concentration tested). While this did not reach the predefined cut-off of two standard deviations from the DMSO control, it is likely that this would have been exceeded at a higher compound concentration

of 500 µM. Therefore, compound **23** could also be interfering with the assay chemiluminescence and thus is a false positive from the AlphaScreen™ assay.

3.4.2.2 *NMR WaterLOGSY Experiments with Mutant Protein*

The conserved asparagine (N980 in TRIM24, and N1039 in TRIM33α and TRIM33β) of the bromodomain and the conserved tryptophan (W828 in TRIM24, and W889 in TRIM33α and TRIM33β) of the PHD finger domain play a key role in binding to histone H3K18Ac and histone H3K9Me$_3$ respectively (Figure 3.8). When the asparagine residue has been mutated to a phenylalanine, or the tryptophan residue has been mutated to alanine, the protein has limited binding to the histone peptide binding partners.[89] As such, the NF and WA mutant proteins can be used to probe the compound binding site. While an NA (asparagine to alanine) mutant protein had also been expressed, the NF mutant protein was found to be more effective in preventing binding to the bromodomain due to the phenyl group taking up more space in the binding pocket.[89]

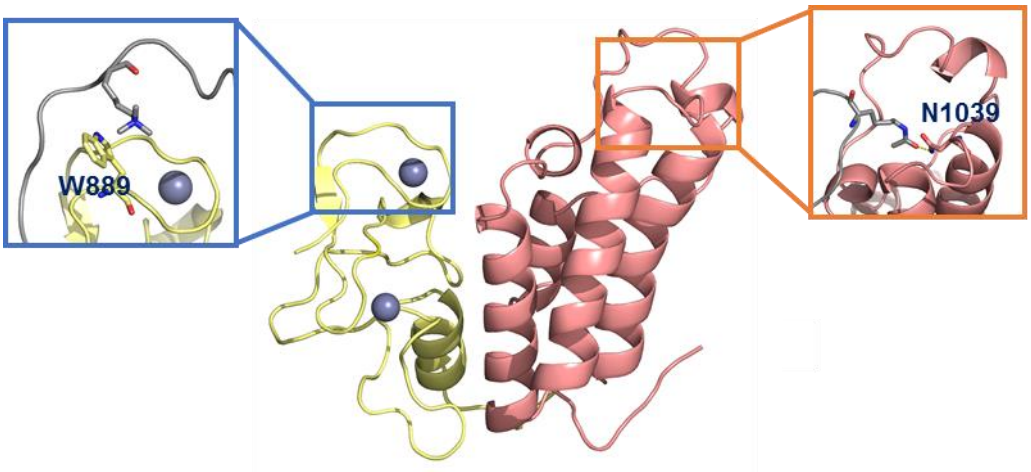


Figure 3.8. X-Ray crystal structure of the TRIM33β PHD-bromodomain cassette. The bromodomain (pink) binds to acetyl-lysine through conserved asparagine N1039 and the PHD finger (yellow) binds trimethylated lysine (peptide overlay from crystal structure of TRIM33α, PDB ID: 3U5O) through conserved tryptophan W889.

The TRIM24, TRIM33α and TRIM33β WA mutant proteins were provided by Dr Angelina Sekirnik, as well as the TRIM24 and TRIM33α NF mutant proteins.[89] The TRIM33β-NF mutant protein was produced using large scale recombinant protein expression in

Escherichia coli BL21 (DE3) cells using TRIM33β-NF plasmid. Protein purification using IMAC followed by SEC on an ÄKTAFPLC system yielded TRIM33β-NF. It was observed that this protein was more prone to aggregation than wild-type TRIM33β and had to be stored at relatively lower concentrations. The purity of the protein was confirmed by SDS-PAGE and protein mass spectrometry (Figure 3.9).

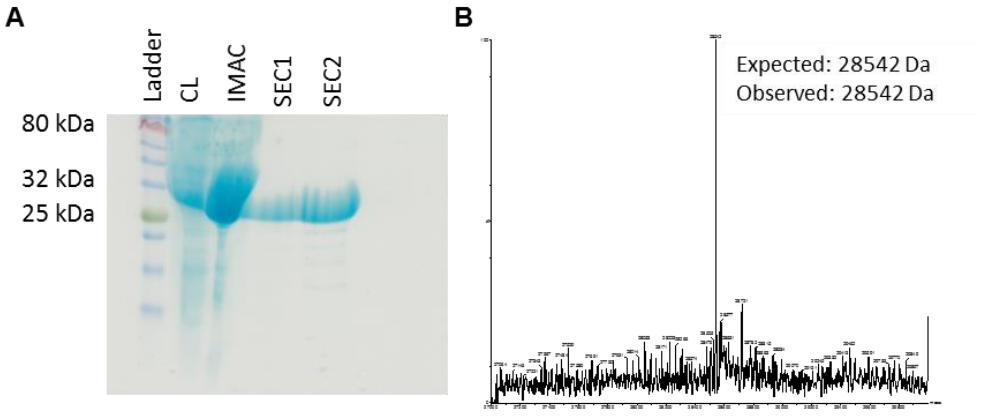


Figure 3.9. Expression of mutant protein TRIM33β-N1039F. A) SDS-PAGE gel showing purity of TRIM33β-N1039F and its size relative to the protein marker (ladder), where CL = cell lysate; B) Protein mass spectrum of TRIM33β-N1039F.

WaterLOGSY experiments were carried out with compounds **20**, **25**, **36** and **37** using the wild-type protein as well as NF and WA mutants of TRIM24, TRIM33α and TRIM33β. Compound binding between the mutant proteins and the wild-type protein was compared. Compounds **36** and **37** showed no difference and were able to bind to the proteins with either mutation.

Although the IC_{50} value of compound **20** in the AlphaScreen™ was beyond the limit of detection for TRIM24 as a full inhibition curve was not obtained, NMR waterLOGSY clearly indicates that it is binding to all 3 proteins (Figure 3.10A). This is because waterLOGSY has a large range of detection (100 nM to 1 mM)[101] and can be used to detect weakly binding ligands. Compound **20** does not bind to the TRIM24-NF mutant protein (Figure 3.10B), further confirming that this compound binds in the bromodomain pocket. Additionally, although binding was not completely negated in the TRIM33β-NF mutant protein, the peaks

observed are significantly weaker, suggesting weaker binding (Figure 3.10C). TRIM33α has an additional 17 amino acids in the BC loop of the bromodomain which pushes the conserved asparagine in a different direction, pointing out of the bromodomain binding pocket. This suggests that the phenylalanine residue could also be oriented away from the acetyl-lysine binding pocket, hence still permitting compound **20** to bind such that it is unaffected by the replacement of the asparagine residue with phenylalanine.

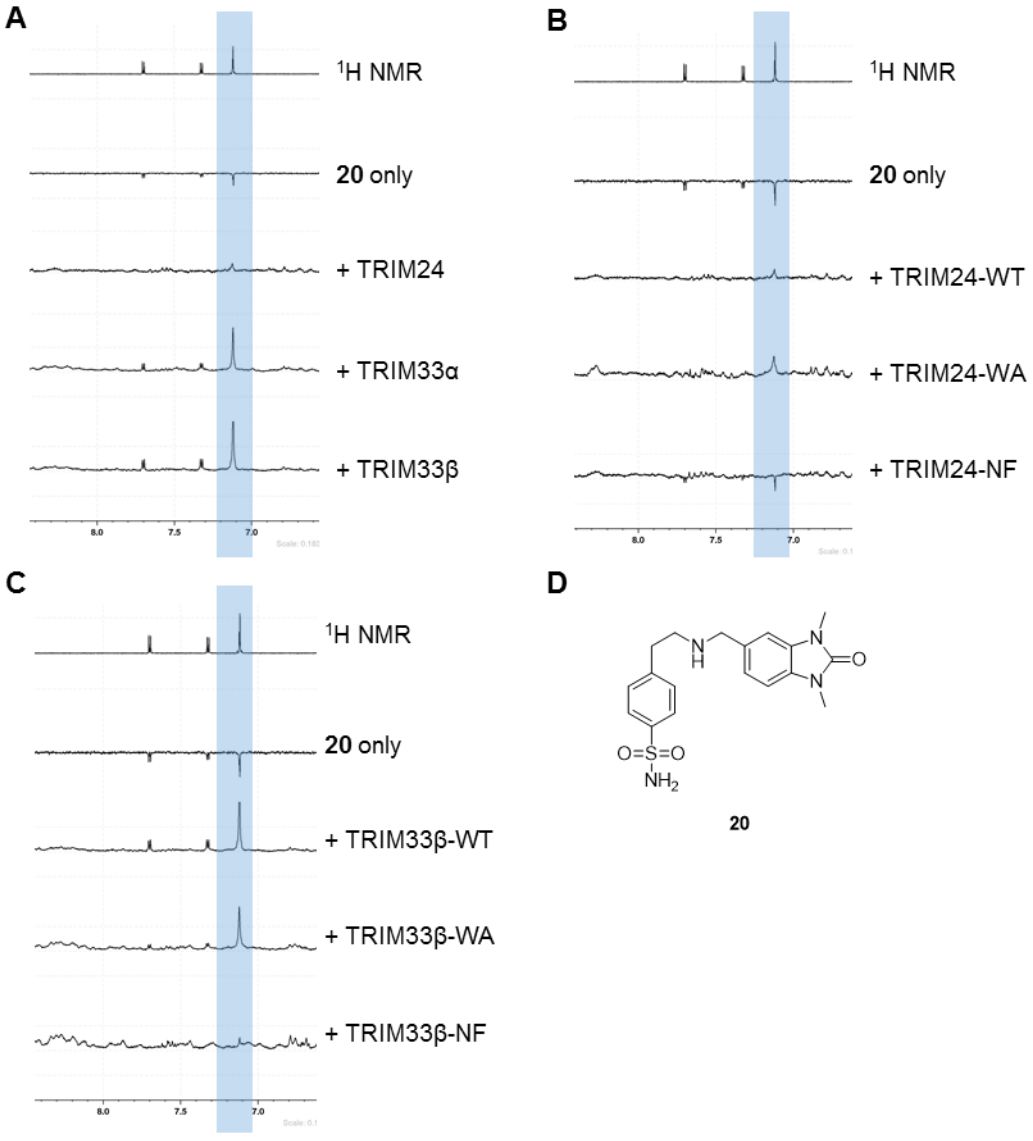
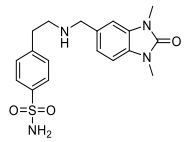
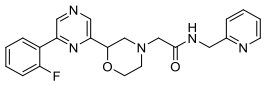
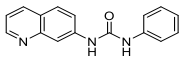
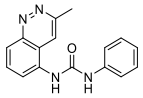


Figure 3.10. NMR waterLOGSY spectra of compound 20. Ligand peaks are highlighted in blue. Sample prepared in PBS buffer, with 500 μM compound and 20 μM protein. A) Spectra with wild-type (WT) TRIM24, TRIM33α and TRIM33β, showing binding to all three proteins. B) Spectra with TRIM24 wild-type and mutants, showing no binding to TRIM24-NF. C) Spectra with TRIM33β wild-type and mutants, showing no binding to TRIM33β-NF.

For all other compounds, it is unclear from the waterLOGSY experiments where they are binding (Table 3.4). All spectra can be found in Appendices E.

Table 3.4. Summary of NMR waterLOGSY results with mutant proteins. Compounds that bind by waterLOGSY are indicated in pale blue; ✓ – binding, × – not binding, ~ – inconclusive. WT – wild-type, NF – N980F mutant for TRIM24 and N1039F mutant for TRIM33, WA – W828A mutant for TRIM24 and W889A mutant for TRIM33.

Compound	TRIM24			TRIM33α			TRIM33β		
	WT	NF	WA	WT	NF	WA	WT	NF	WA
 **20**	✓	×	✓	✓	✓	✓	✓	Less	✓
 **25**		×		✓	✓	✓	✓	×	×
 **36**	✓	✓	✓	✓	✓	✓	✓	✓	✓
 **37**	✓	✓	✓	✓	✓	✓	✓	✓	✓

3.5 Conclusions

Follow-up AlphaScreen™ experiments indicated that compounds **20**, **23** and **25** had sufficiently high affinity to be quantified and brought forward for further validation. At the same time, the Tanimoto similarity index was used to select compounds from the Maybridge fragment library, forming a small focused library to probe SAR around the Maybridge fragment hits identified in the initial screening. These compounds were tested at two concentrations in the AlphaScreen™ and were narrowed down to compounds **29**, **32**, **34**, **33**, **35**, **36**, **37**, and **38**.

All 11 of these compounds were tested using the AlphaScreen™ TruHits kit to identify compounds that are false positives as they cause a reduction in emission due to assay interference rather than actual binding to the protein of interest. Compounds **20**, **23**, **25**, **35**, **36** and **37** were progressed to further testing. NMR waterLOGSY was then used as an orthogonal assay for confirming binding of these compounds. WaterLOGSY experiments confirmed that compounds **20**, **36** and **37** bound to all three proteins while compound **25** bound to both isoforms of TRIM33 but not TRIM24. Furthermore, waterLOGSY experiments with mutant protein indicated that compound **20** bound to the bromodomain region of TRIM24 and TRIM33β. Therefore, further development was focused on compound **20** with a view to optimise this for the TRIM33 bromodomain.

Chapter 4

Development of Benzimidazolone Series

4.1 Introduction and Aims

Following the identification of hit compound **20** (Figure 4.1A) as a TRIM33 bromodomain ligand, this chapter focuses on work to investigate structure-activity relationships (SAR) around compound **20** with the aim of improving affinity to TRIM33. Different amine chains (blue), vectors (green) and substituents extending off the benzimidazolone headgroup (pink) were explored. This study involves iterative cycles of ligand design, synthesis of analogues and testing of compounds using AlphaScreen™ and NMR waterLOGSY assays (Figure 4.1B). Throughout the workflow, computational techniques including docking and molecular dynamics simulations were employed to guide ligand design as well as to aid understanding of assay results. Building on our understanding of the acetyl-lysine binding pocket in other bromodomains, this study aims to deepen understanding of the acetyl-lysine binding pocket in TRIM33, further enabling ligand development.

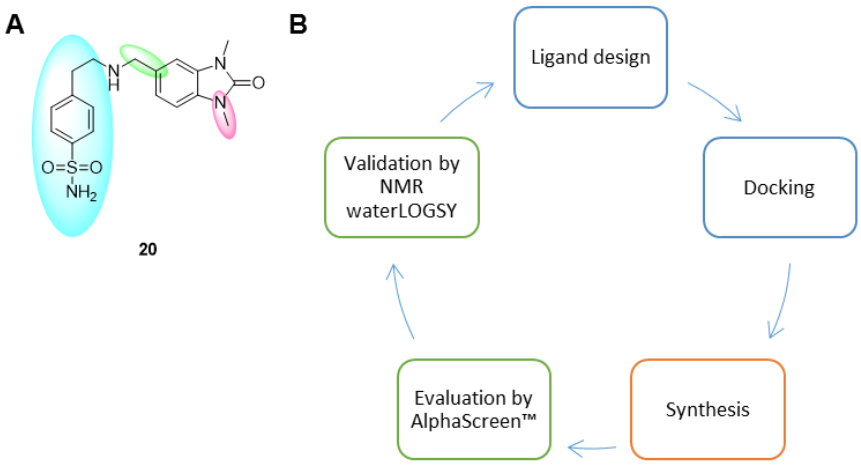


Figure 4.1. Overview of development of benzimidazolone series. A) Hit compound **20** on which further analogues are based; B) Workflow summarising development of the benzimidazolone series.

4.2 Establishing Structure-Activity Relationships

4.2.1 Preliminary Docking Studies

To better understand the binding mode of hit compound **20**, it was first docked into the TRIM33β acetyl-lysine binding pocket using CLC Drug Discovery Workbench and a TRIM33β bromodomain and PHD finger X-ray crystal structure from Prof. Stefan Knapp (unpublished). It is known that there are five key conserved water molecules that line the bottom of the acetyl-lysine binding pocket of the bromodomain and form a water network.[25] These water molecules have also been observed in TRIM24, which is a closely related protein. Two TRIM24 bromodomain ligands based on the benzimidazolone headgroup (Figure 4.2A) are bound to the acetyl-lysine binding pocket with the carbonyl oxygen making a hydrogen bond with the conserved asparagine N980 (Figure 4.2B and Figure 4.2C, PDB ID: 4YC9[60], 4ZQL[61]). Consequently, these water molecules were taken into consideration when setting up the binding site for docking. Additionally, it was observed that there are three water molecules present in the ZA channel in the TRIM33β crystal structure (labelled 5–7, Figure 4.2D).

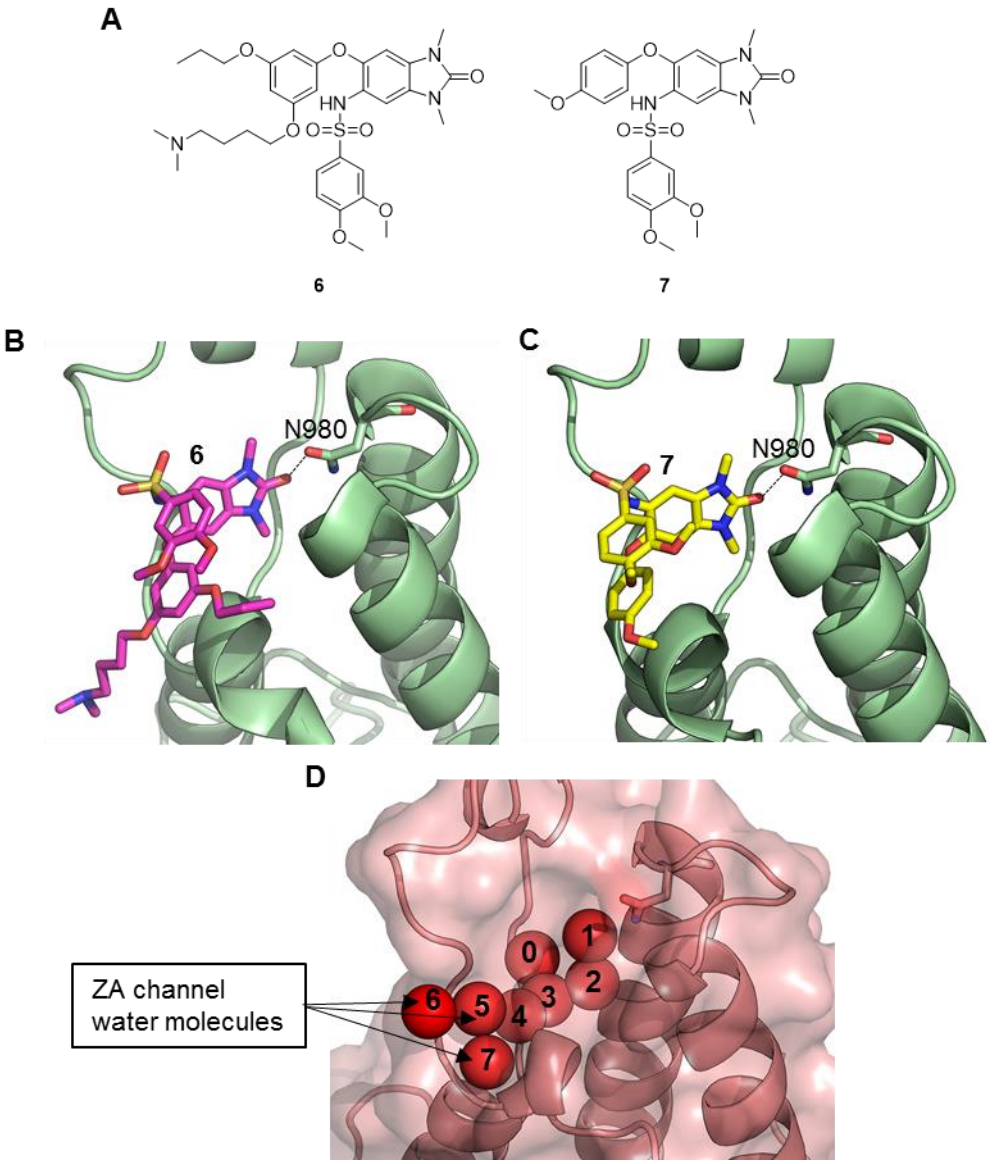


Figure 4.2. TRIM24 and TRIM33β binding site. A) Structure of TRIM24 ligands **6** and **7**; B) X-Ray crystal structure of **6** (carbon = magenta) bound to TRIM24 (PDB ID: 4YC9, green);[60] C) X-Ray crystal structure of **7** (carbon = yellow) in TRIM24 (PDB ID: 4QZL, green);[61] D) X-Ray crystal structure of TRIM33β (unpublished, Prof. Stefan Knapp, pink) with binding site water molecules indicated by the red spheres.

As the displaceability of these water molecules were not known, docking was carried out with none, all, or combinations of the ZA channel water molecules present in the binding site. Docking was conducted with the binding site set-up with a radius of 13 Å (Figure 4.3A). Docking results were then visually inspected. Based on comparison to the binding mode of TRIM24 ligands **6** and **7**, results were ignored when the benzimidazolone headgroup was not in the acetyl-lysine binding pocket or not interacting with the conserved asparagine

through a hydrogen bond. Hit compound **20** was seen to dock in the acetyl-lysine binding pocket regardless of the number of water molecules in the ZA channel (Figure 4.3B). This observation was confirmed in molecular dynamics simulations conducted by Joe Bluck (Figure 4.3C). It was predicted that the benzylic amine (Figure 4.3B) is likely to be hydrogen bonded to glutamic acid 981 (E981), and that there is potential for an additional hydrogen-bonding or ionic interaction with glutamic acid 984 (E984).

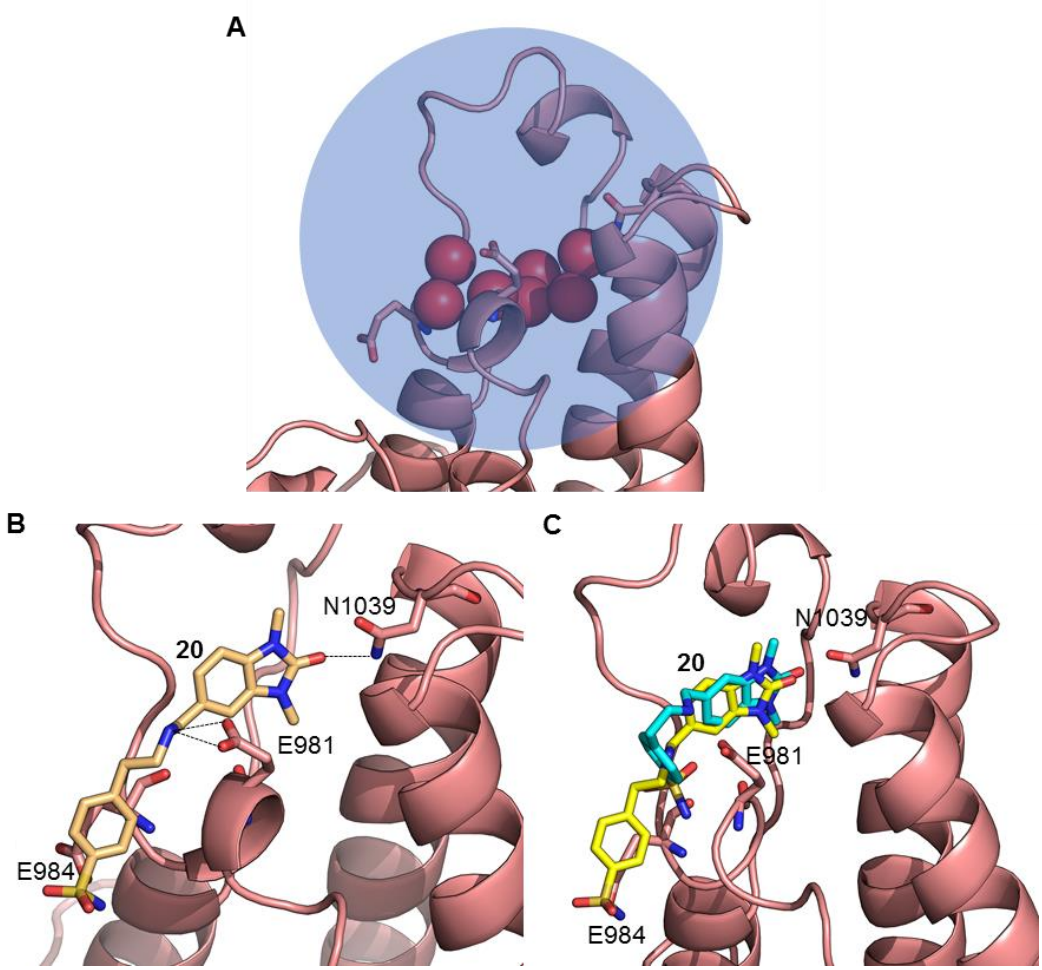


Figure 4.3. Docking of 20 into TRIM33β binding site. A) Binding site set-up as indicated by the blue sphere; B) Compound **20** (carbon = orange) docked into the X-ray crystal structure of TRIM33β (pink) using CLC Drug Discovery software; C) Compound **20** in two poses (carbon = cyan; carbon = yellow) in TRIM33β obtained from molecular dynamics simulations (50 ns simulations in triplicate, using GROMACS 2016.3).

4.2.2 Establishing Structure-Activity Relationships 1

Six compounds were synthesised by Amy Scorah to establish initial understanding of the binding mode by confirming the conclusions drawn from preliminary docking studies

(Figure 4.4). Firstly, the importance of the benzylic amine was investigated by replacing the amine with an amide bond (compound **41**). Amide analogue **42** with an unmethylated benzimidazolone headgroup was also synthesised for evaluation. Secondly, to test the importance of the benzylic amine, benzimidazolone analogues were synthesised with a methyl ester, carboxylic acid and alcohol (compounds **43**, **44** and **45**, respectively). Finally, aniline analogue **46** was synthesised to investigate the effect of amine position and basicity (pK_a (benzylic amine of **20**) = 8.73; pK_a (aniline of **46**) = 5.86).

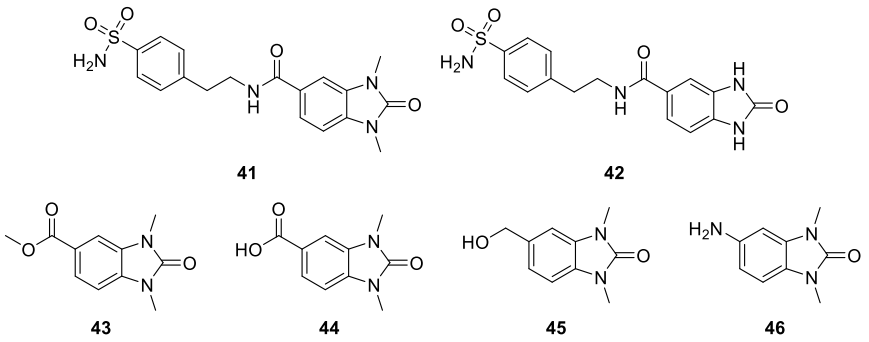
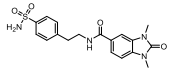
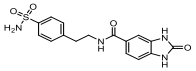
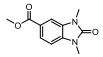
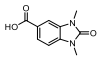


Figure 4.4. Compounds synthesised for initial SAR studies.

These six compounds were tested at two different concentrations (30 μM and 150 μM) in an AlphaScreen™ assay and the emission was calculated as a percentage of the DMSO control. A reduction in emission indicates binding to the protein. The results are summarised in Table 4.1. It was observed that replacing the amine with an amide (**41**) and removing the methyl groups of the benzimidazolone headgroup (**42**) resulted in less binding compared to the original hit or no binding at all. As expected based on computational analysis, carboxylic acid **44** did not bind as the presence of glutamic acid (E981) close to the binding pocket is likely to repel it. Alcohol **45** as well as ester **43** also showed less binding than amine **46**, indicating that a basic group is important in forming this interaction with E981. Therefore, in subsequent studies, this basic group was maintained.

Table 4.1. Results from AlphaScreen™. Colour scale from low emission indicating binding (red) to high emission indication no binding (green); T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β, H3(PHD) – H3K9Me$_3$, H3(BRD) – H3K18Ac, H3(Dual) – H3K9Me$_3$K18Ac.

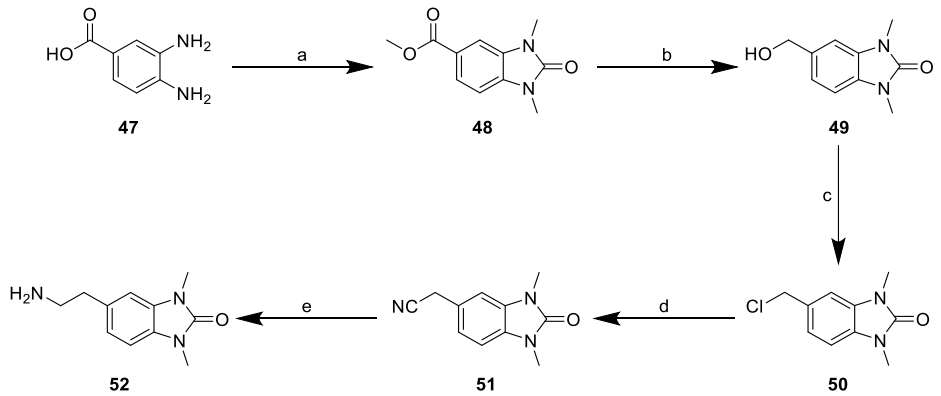
Concentration (μM) / Compound	Emission (as a percentage of DMSO control)																	
	T24-H3(PHD)		T24-H3(BRD)		T24-H3(Dual)		T33α-H3(PHD)		T33α-H3(BRD)		T33α-H3(Dual)		T33β-H3(PHD)		T33β-H3(BRD)		T33β-H3(Dual)	
	30	150	30	150	30	150	30	150	30	150	30	150	30	150	30	150	30	150
 **41**	117	105	97	97	107	98	240	248	71	49	97	79	104	104	143	116	97	88
 **42**	110	98	91	85	113	103	148	133	123	105	121	113	107	96	156	145	100	96
 **43**	105	96	85	71	104	93	132	101	86	41	103	69	101	100	145	117	96	89
 **44**	112	108	120	107	115	110	158	176	121	93	122	106	108	105	152	134	102	96
 **45**	111	108	124	118	117	111	124	104	87	39	102	66	106	103	143	120	99	89
 **46**	95	83	113	95	107	87	97	77	73	23	90	44	100	87	129	94	87	72

4.2.3 Establishing Structure-Activity Relationships 2

While it was established that a basic group was essential, the impact of distance of the amine from the headgroup, or whether other basic groups would be tolerated, was not known. Investigations were next focused on exploring this.

4.2.3.1 *Synthesis of 2-Carbon Amine Analogue*

The benzimidazolone headgroup was first synthesised by reacting 3,4-diaminobenzoic acid (**47**) with 1,1'-carbonyldiimidazole (CDI) in excess (Scheme 4.1). Side reaction of CDI with the carboxylic acid of **47** resulted in formation of the imidazole amide, which, on addition of methanol, reacted to form the methyl ester with elimination of imidazole. Methylation with iodomethane and potassium carbonate afforded the benzimidazolone headgroup methyl ester **48** in 77% yield. Reduction to the alcohol **49** using diisobutylaluminium hydride (DIBAL), followed by chlorination with methanesulfonyl chloride, gave compound **50**. This was carried forward to make nitrile **51**, without further purification, in a reaction with sodium cyanide.[102] Nitrile reduction was first attempted using lithium aluminium hydride but no product was obtained. It is possible that the benzimidazolone core is not stable to lithium aluminium hydride. Instead, a hydrogenation reaction using palladium on carbon under acidic conditions[103] was carried out to afford the desired amine **52**.



Scheme 4.1. Synthesis of 2-carbon amine analogue. *Reagents and conditions:* a) (i) CDI, CH_2Cl_2, rt, 2 h, (ii) MeOH, rt, 1 h, (iii) MeI, K_2CO_3, DMF, 0 °C to rt, o/n, 77% over three steps; b) DIBAL-H (1 M in hexanes), THF, –78 °C to rt, o/n, 85%; c) Methanesulfonyl chloride, Et_3N, CH_2Cl_2, 0 °C, 0.5 h; d) NaCN, DMF, 70 °C, 1.5 h, 40% over two steps; e) H_2, Pd/C (10 mol%), HCl, EtOH, rt, o/n, 49%.

4.2.3.2 *Evaluation of Compounds*

Four other compounds were synthesised by Amy Scorah (Figure 4.5) to investigate the effect of proximity of the basic group from the headgroup (**46**, **52** and **53**). Analogues where the basic amine was replaced with pyrazole and imidazole (**54** and **55**, respectively) were also synthesised. These five compounds were tested in a dose-response AlphaScreen™ in 1:2 serial dilutions from 500 μM FAC. In all dose-response AlphaScreen™ assays, the original hit compound **20** was included on each plate as positive control for the assay.

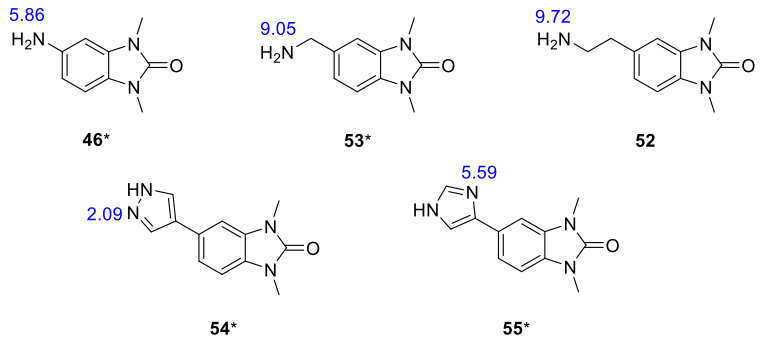


Figure 4.5. Compounds tested in a dose-response AlphaScreen™ assay. pK_a values (in blue) were calculated using ACD/I-Lab 2.0. *Compounds synthesised by Amy Scorah.

Where full sigmoidal curves were obtained, IC_{50} values were calculated using GraphPad Prism 7 software by fitting data points to a sigmoidal four-parameter equation:

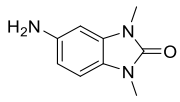
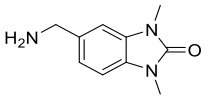
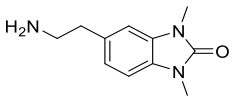
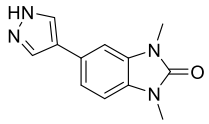
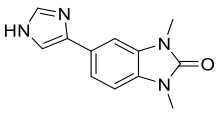
$$Y = Minimum + \frac{(Maximum - Minimum)}{1+10^{(\log IC_{50} - X)(Hill\ Slope)}}$$, where X is log(compound concentration).

In some instances, the binding affinity of the compound was not high enough for full IC_{50} curves to be obtained. This is indicated by > 250 μM affinity. Compounds were always tested against all three proteins (TRIM24, TRIM33α and TRIM33β) in combination with all three peptides (H3K9Me_3 – binds to PHD finger, H3K18Ac – binds to bromodomain, H3K9Me_3K18Ac – binds to both domains). However, given that the compounds bind to the bromodomain, the main peptide of interest is H3K18Ac. Therefore, for ease of comparison, this will be the focus in analysing the results, and data for the other combinations can be found in Appendices B. Following dose-response AlphaScreen™ experiments, compounds

showing binding were validated using NMR waterLOGSY experiments as an orthogonal technique. This enabled binding to be directly detected through ligand observation, confirming binding without the potential for assay interference. For compounds that formed soluble aggregates, compounds concentration was reduced until a negative spectrum was obtained in the absence of protein. The results obtained are presented in Table 4.2.

AlphaScreen™ results indicate that pyrazole **54** and imidazole **55** are both tolerated and are basic groups suitable for maintaining interaction with E981. With their small size (MW = 228), despite their and low affinity for TRIM33α (6.51 μM and 3.35 μM, respectively), these compounds are extremely ligand efficient and suitable candidates for further ligand optimisation. Ligand efficiency (LE = $\frac{1.4(-logIC_{50})}{Number\ of\ non-hydrogen\ atoms}$) is a metric for assessing binding affinity while accounting for the size of the molecule.[104] As ligand efficiency often decreases during ligand optimisation due to the addition of substituents, a ligand efficiency of > 0.3 is desirable and the ligand efficiency of 0.4 here makes this an ideal starting point for ligand optimisation.

Table 4.2. AlphaScreen™ IC_{50} values and NMR waterLOGSY binding data. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green). Compounds that bind by waterLOGSY are indicated in pale blue; ✓ – binding, × – not binding, – – compound not tested in waterLOGSY experiments. Ligand efficiency is indicated in parentheses.

Compound	NMR waterLOGSY			AlphaScreen™ IC_{50} (μM)		
	T24	T33α	T33β	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **46**	✓	✓	✓	22.5 ± 1.45 (0.50)	41.8 ± 1.13 (0.47)	168 ± 1.96 (0.41)
 **53**	×	✓	✓	> 250	7.83 ± 1.10 (0.51)	30.6 ± 1.10 (0.45)
 **52**	–	–	–	> 250	65.1 ± 12.4 (0.39)	> 250
 **54**	✓	✓	✓	92.1 ± 1.23 (0.33)	6.51 ± 1.07 (0.43)	23.3 ± 1.27 (0.38)
 **55**	✓	✓	✓	52.4 ± 1.21 (0.35)	3.35 ± 1.11 (0.41)	26.5 ± 1.10 (0.38)

Comparing the proximity of the basic nitrogen from the benzimidazolone headgroup, having a benzylic amine (**53**, IC_{50} (TRIM33α) = 7.83 μM, IC_{50} (TRIM33β) = 30.6 μM) is more effective than having an aniline with the nitrogen attached directly to the headgroup (**46**, IC_{50} (TRIM33α) = 41.8 μM, IC_{50} (TRIM33β) = 168 μM). Similarly, these compounds are very ligand efficient and present compound **53** as a good fragment for ligand optimisation. When there is an additional carbon atom in between (compound **52**), however, binding of the compound is greatly reduced with an IC_{50} value of 65.1 μM for TRIM33α and no binding to

TRIM24 or TRIM33β. This result suggests that one carbon is the optimum chain length for interaction with E981 when considering the primary amine. Notably, while compounds **46**, **54**, and **55** bind to TRIM24, compound **53** does not bind to TRIM24. This observation is supported by waterLOGSY experiments and can be rationalised by comparing the crystal structures of TRIM24 and TRIM33β. Overlaying TRIM24 and TRIM33β (Figure 4.6A) revealed that TRIM24 has an alanine residue (A923) where E981 is on TRIM33β. In **46**, **54**, and **55** the basic group is oriented differently from that in **53** (Figure 4.6B). It is likely that the basic amine of **53** is able to form a hydrogen bond interaction with E981 in TRIM33α and TRIM33β, but not A923 on TRIM24, giving rise to the selectivity observed. Hence, this amine plays an important role in the affinity as well as selectivity of this compound series and will be maintained throughout ligand optimisation.

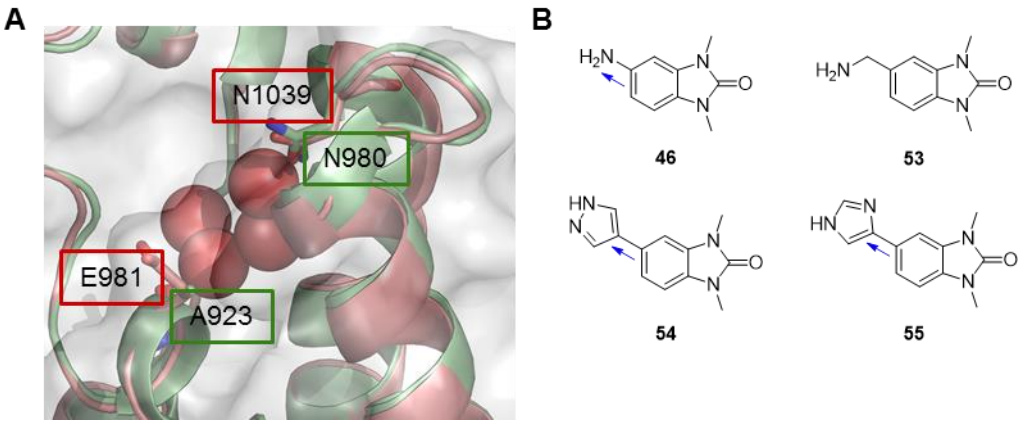


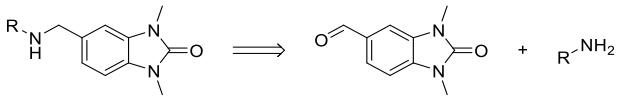
Figure 4.6. A) Overlay of TRIM24 (PDB ID: 3O33, green) with TRIM33β (pink) showing difference in a key glutamic acid E981 in TRIM33β that is A923 in TRIM24; B) Comparison of orientation of basic group in compounds tested. Blue arrow indicates direction of basic group.

4.3 Variation of Amine Chain

4.3.1 Structure-Activity Relationship Studies 1

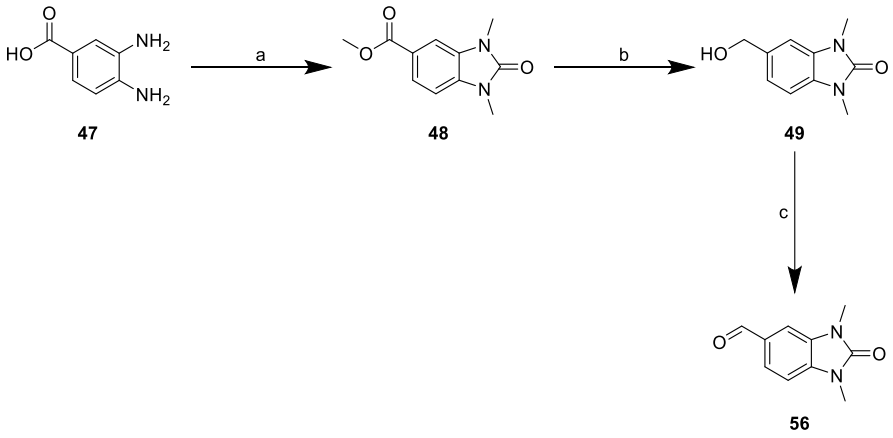
4.3.1.1 *Synthesis of Compounds*

In the first phase of ligand optimisation, a series of benzimidazolone analogues were synthesised with different amine chains extending off the amine. These can be easily obtained through a common aldehyde headgroup intermediate, enabling a variety of compounds to be synthesised *via* reductive amination with different amines (Scheme 4.2).



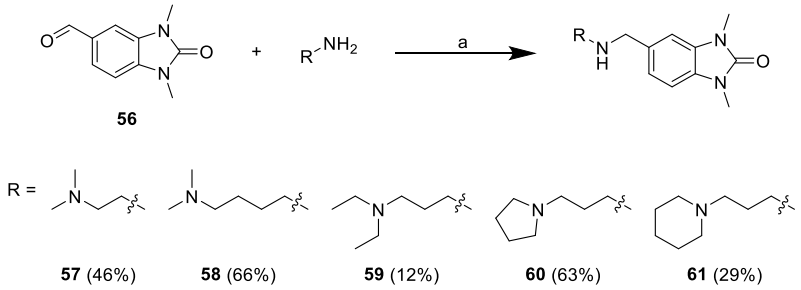
Scheme 4.2. Proposed disconnection to allow efficient synthesis of analogues.

The ester **48** was obtained by reaction of 3,4-diaminobenzoic acid **47** with CDI, followed by esterification with methanol and methylation with methyl iodide (Scheme 4.3). This sequence did not require isolation and purification and worked well on a 6 g scale (40 mmol), resulting in yields of 67–98%. The ester **48** was reduced to alcohol **49** using DIBAL on a scale of up to 2.4 g (11 mmol). It was observed that reduction of the ester headgroup **48** using DIBAL resulted in unpredictable and inconsistent yields, where if product could be isolated, yields of 33–85% were obtained, but in some instances, no product or a negligible amount of product was isolated. This was found to be independent of the work-up used (Rochelle salt or Fieser). It was subsequently established that lithium borohydride reduction was a more effective choice and reliably resulted in yields of 79–96% on up to 3 g scale. The benzylic alcohol **49** was oxidised to the key aldehyde intermediate **56** using manganese(IV) dioxide. This reaction was carried out on a scale of 5 g, resulting in yields of 66–95%.



Scheme 4.3. Synthesis of aldehyde benzimidazolone headgroup. *Reagents and conditions:* a) (i) CDI, CH_2Cl_2, rt, 2 h, (ii) MeOH, rt, 1 h, (iii) MeI, K_2CO_3, DMF, 0 °C to rt, o/n, 67–98% over three steps, n = 3; b) DIBAL-H (1 M in hexanes), THF, –78 °C to rt, o/n, 33–85%, n = 3, or $LiBH_4$ (2 M in THF), THF, MeOH, reflux, o/n, 79–96%, n = 3; c) Manganese(IV) oxide, CH_2Cl_2, rt, 2 h, 66–95%, n = 4.

The aldehyde **56** was then used in a series of reductive aminations with five different amines to afford compounds **57**–**61** (Scheme 4.4). Both the amine chain length and the alkyl groups on the terminal amine were varied.



Scheme 4.4. Reductive amination with various amines. *Reagents and conditions:* a) (i) 1,2-Dichloroethane, acetic acid, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 12–66%.

4.3.1.2 *Evaluation of Compounds*

Besides the five compounds discussed, another four compounds were synthesised by Amy Scorah and together, these nine compounds (Figure 4.7) were evaluated in a dose-response AlphaScreen™ and where appropriate, in an NMR waterLOGSY experiment (Table 4.3).

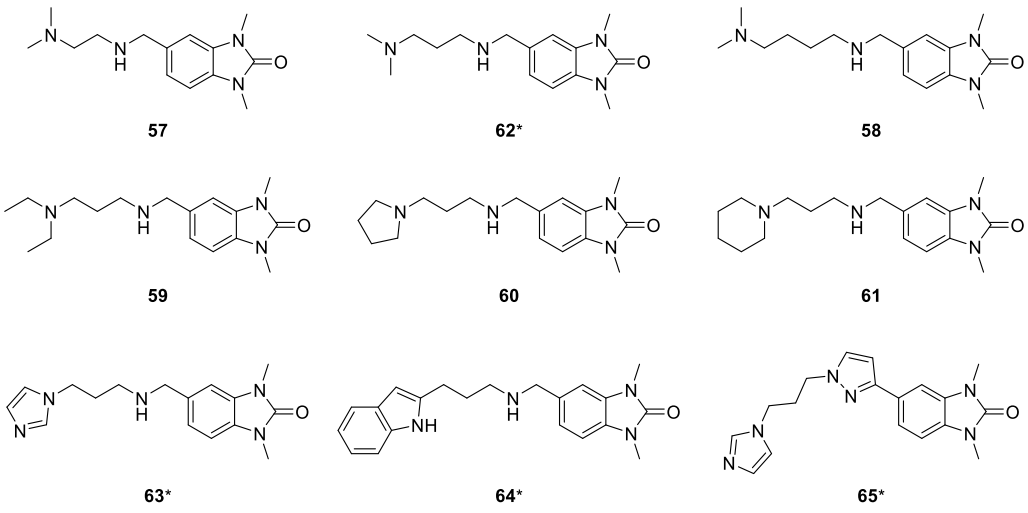
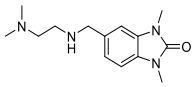
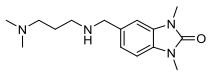
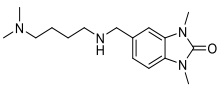
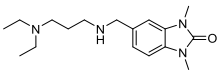
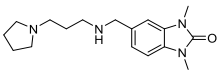
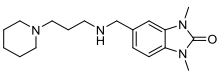
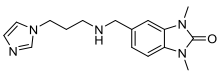
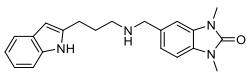
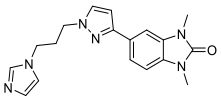


Figure 4.7. Compounds tested in dose-response AlphaScreen™ to test the effect of different amine chains. *Compounds synthesised by Amy Scorah.

Table 4.3. AlphaScreen™ IC_{50} values and NMR waterLOGSY data. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} values (green). Compounds that bind by waterLOGSY are indicated in pale blue; ✓ – binding, × – not binding, ~ – inconclusive, – – compound not tested in waterLOGSY experiments. Ligand efficiency is indicated in parentheses. *These IC_{50} values are the mean and standard deviation of multiple repeats.

Compound	NMR waterLOGSY			AlphaScreen™ IC_{50} (μM)		
	T24	T33α	T33β	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **57**	~	✓	✓	> 250	5.89 ± 1.09 (0.39)	15.2 ± 1.07 (0.35)
 **62**	~	✓	✓	> 250	3.61* ± 1.15 (0.38)	13.7* ± 9.34 (0.34)
 **58**	~	✓	✓	> 250	7.96 ± 1.11 (0.34)	2.73 ± 1.14 (0.37)
 **59**	~	✓	✓	> 250	4.74 ± 1.14 (0.34)	2.02 ± 1.06 (0.36)
 **60**	~	✓	✓	24.3 ± 1.36 (0.29)	4.11 ± 1.31 (0.34)	9.30 ± 1.47 (0.32)
 **61**	~	✓	✓	> 250	2.71 ± 1.08 (0.34)	9.24 ± 1.06 (0.31)
 **63**	×	✓	✓	> 250	7.97 ± 1.06 (0.32)	11.9 ± 1.12 (0.31)
 **64**	~	✓	✓	> 250	10.5 ± 1.09 (0.28)	65.5 ±1.18 (0.23)
 **65**	~	✓	✓	34.5 ± 1.35 (0.25)	2.46 ± 1.07 (0.31)	15.3 ± 1.10 (0.27)

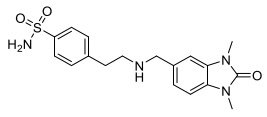
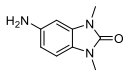
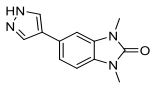
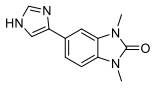
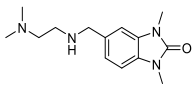
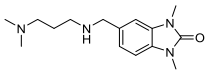
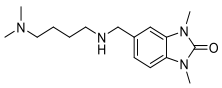
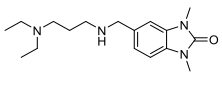
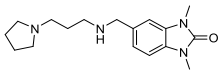
By comparing compounds **57**, **58**, and **62**, it is clear that compound **59** has the most improved affinity with an IC_{50} value of 2.02 μM against TRIM33β, compared to compound **53** (IC_{50} = 30.6 μM) which has no amine chain. Comparing the IC_{50} values obtained for TRIM33α suggests that there is little impact of chain length. Although compounds **57**, **58**, and **62**, have IC_{50} values of 5.89 μM, 7.96 μM and 3.61 μM for TRIM33α, a limitation of AlphaScreen™ is that being a fluorescence-based assay, there is often some variation between repeats and differences of 2 μM is not significant. The terminal diethyl in compound **59** was well-tolerated while cyclopentyl (**60**) and cyclohexyl (**61**) groups were less preferred for TRIM33β. All three were well-tolerated for TRIM33α with no significant difference in IC_{50} values. Pleasingly, for all these compounds, ligand efficiency was maintained at > 0.3. Additionally, an aromatic terminal basic group, the imidazole in **63**, was tolerated though binding was slightly weaker (IC_{50} (TRIM33α) = 7.97 μM, IC_{50} (TRIM33β) = 11.9 μM). In contrast, tryptamine **64** showed a significant drop in affinity, resulting in an IC_{50} value of 65.5 μM.

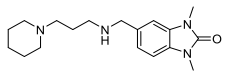
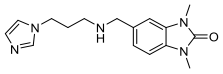
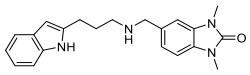
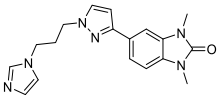
Compound **65**, on the other hand, had comparable affinity as compared to the 2- and 3-carbon chain compounds (**57** and **62**, respectively). Compound **65** also showed binding to TRIM24 with an IC_{50} value of 34.5 μM. In the previous AlphaScreen™ (Table 4.2), it was observed that compounds **54** and **55**, with the pyrazole and imidazole similarly directly attached to the headgroup, as well as compound **46**, with the amine directly attached, also had some binding to TRIM24. The observed binding to TRIM24 could be accounted for by π-stacking as observed in the known TRIM24 ligands. It is also possible that this is due to the difference in pK_a values of the conjugate acid of imidazole (pK_a = 7.0), pyrazole (pK_a = 2.5), aniline (pK_a = 4.6) and benzylamine (pK_a = 9.3). As the pK_a of the glutamate side chain is 4.3, it is partially deprotonated under the assay conditions of pH 7.5. Only benzylamine, with a pK_a greater than 7.5, would be mainly protonated and hence better able to form a hydrogen bond or ionic interaction with E981.

4.3.2 Validation of Compound Binding

As further ligand optimisation would be based on these compounds, it was important to confirm that compounds identified as binders were not false positives due to interference with the AlphaScreen™ assay. Compounds **20**, **46**, **54**, **55**, **57**, **58**, **59**, **60**, **61**, **62**, **63**, **64**, and **65** were evaluated using the AlphaScreen™ TruHits kit. After incubation with the beads, the difference between the emission with the compounds and that of the DMSO control was then calculated and normalised against the standard deviation of the DMSO control (Table 4.4).

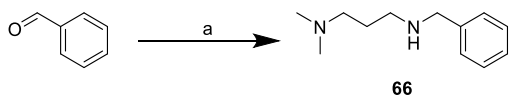
Table 4.4. Results from AlphaScreen™ TruHits experiment. Differences of greater than 2 standard deviations are indicated in pink.

Compound \ Concentration	Difference from DMSO average, normalised against standard deviation of DMSO control				
	Concentration				
	31.3 µM	62.5 µM	125 µM	250 µM	500 µM
 **20**	1.801	0.782	0.409	0.607	2.020
 **46**	0.620	1.323	3.706	5.706	11.569
 **54**	1.876	1.219	0.376	0.170	1.718
 **55**	1.473	0.117	2.910	5.707	11.423
 **57**	0.650	1.215	1.185	1.680	0.972
 **62**	0.352	0.114	1.008	0.945	2.441
 **58**	0.274	1.067	0.460	1.859	3.612
 **59**	0.449	1.382	0.527	0.569	2.096
 **60**	0.749	0.476	1.119	0.605	2.667

 **61**	0.131	0.943	1.361	1.495	2.759
 **63**	1.227	0.109	0.827	1.586	4.179
 **64**	0.602	2.255	5.534	8.835	14.964
 **65**	0.715	2.345	3.933	5.707	9.882

For compounds **20**, **58**, **59**, **60**, **61**, and **62**, differences of greater than two standard deviations were recorded but only at high compound concentration of 500 μM and the increase in difference was not found to be concentration dependent. Compounds **46**, **55**, **63**, **64** and **65**, however, showed significant interference with the assay. Compound **46** contains an aniline, compound **64** contains a tryptamine while compounds **55**, **63** and **65** feature an imidazole moiety. All three groups are known to undergo photooxygenation by reaction with singlet oxygen.[105,106,107,108] These compounds are singlet oxygen quenchers and interfere with the AlphaScreen™ assay, which depends on diffusion of singlet oxygen to produce the luminescent signal observed. NMR waterLOGSY experiments confirm binding of these compounds, therefore they are not false positives. However, the low IC_{50} values obtained for these compounds are likely to be due to assay interference and it cannot be taken as a measure of the actual binding affinity. Based on the results of this assay, then benzimidazolone headgroup with an amine chain as in **62** does not interfere with the AlphaScreen™ and can be used as the valuable starting point for ligand optimisation.

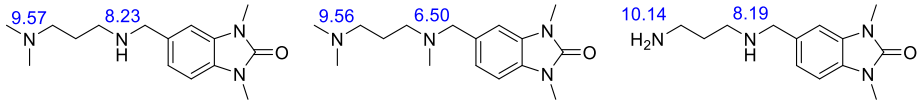
In addition, to confirm that the benzimidazolone headgroup is essential in the binding of **62** to TRIM33, compound **66**, without the carbonyl and two methyl groups, was synthesised *via* reductive amination with benzaldehyde (Scheme 4.5).



66

Scheme 4.5. Synthesis of negative control compound 66. *Reagents and conditions:* a) (i) 3-(Dimethylamino)-1-propylamine, Et_3N, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 19%.

Two other compounds were synthesised by Amy Scorah – compounds **67**, with the benzylic amine methylated, and **68** with the terminal amine unmethylated (Figure 4.8). It was hypothesised that methylating the benzylic amine would reduce the hydrogen bonding ability of the ligand due to no proton being available, particularly as pK_a calculations suggested that the methylated benzylic amine would only be partially protonated at pH 7.5.



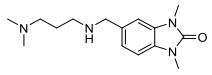
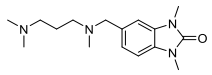
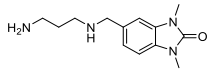
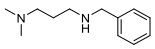
62 **67** **68**

Figure 4.8. Compounds tested with different methylation states. pK_a values (in blue) calculated using ACD/I-Lab 2.0.

A dose-response AlphaScreen™ assay was carried out and results are shown in Table 4.5. As expected, compound **67** has the weakest binding of all three compounds, with IC_{50} values of 5.69 μM for TRIM33α and 23.5 μM for TRIM33β. Methylation of the benzylic amine does not seem to negate binding, suggesting that the methylated amine is still able to interact with E981 through an ionic interaction rather than a hydrogen-bonding interaction, though possibly to a lesser extent. Compound **68** has slightly stronger binding to TRIM33β, with an IC_{50} value of 9.33 μM as compared to 13.7 μM for **62**. The unmethylated terminal amine of compound **68** is more basic than the dimethylated terminal amine of compound **62** and has additional protons available for hydrogen bonding. This could allow compound **68** to interact more strongly with E984, resulting in a higher affinity. The negative control

compound **66** did not show binding to any of the proteins, confirming that the benzimidazolone headgroup plays a key role in the binding of the ligand to TRIM33.

Table 4.5. AlphaScreen™ IC_{50} values and NMR waterLOGSY data. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} values (green). Compounds that bind by waterLOGSY are indicated in pale blue; ✓ – binding, **×** – not binding, ~ – inconclusive, – – compound not tested in waterLOGSY experiments. Ligand efficiency is indicated in parentheses. *These IC_{50} values are the mean and standard deviation of multiple repeats.

Compound	NMR waterLOGSY			AlphaScreen™ IC_{50} (μM)		
	T24	T33α	T33β	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **62**	~	✓	✓	> 250	3.61* ± 1.15 (0.38)	13.7* ± 9.34 (0.34)
 **67**	~	✓	✓	> 250	5.69 ± 0.33 (0.35)	23.5 ± 2.80 (0.31)
 **68**	–	–	–	> 250	3.73 ± 0.83 (0.42)	9.33 ± 1.50 (0.39)
 **66**	–	–	–	> 250	> 250	> 250

4.3.3 Computational Studies

As mentioned previously, the key interactions between the ligand and the protein are 1) the hydrogen bond between the carbonyl oxygen of the benzimidazolone headgroup with the conserved asparagine (N1039), and 2) the ionic interactions between proton of the benzylic amine with E981. At the same time, there is potential for an additional ionic interaction to be formed between the terminal amine and E984 (Figure 4.9A) or with E981 in a "pincer" conformation (Figure 4.9B). These interactions were confirmed in molecular dynamics simulations as well as in distance plots calculated from the simulations by Joe Bluck (Figure 4.9C). In the molecular dynamics simulation experiment of **62** in the acetyl-lysine

binding site of the TRIM33β bromodomain, the distances measured were between 1) the carbonyl oxygen of the headgroup and N1039, 2) the benzylic amine nitrogen and E981, 3) the terminal amine nitrogen and E981, and 4) the terminal amine nitrogen and E984. The experiment was carried out in triplicate and the distances are plotted below (Figure 4.9C), with purple indicating a smaller distance and hence, stronger interaction, and yellow indicating a larger distance and no interaction.

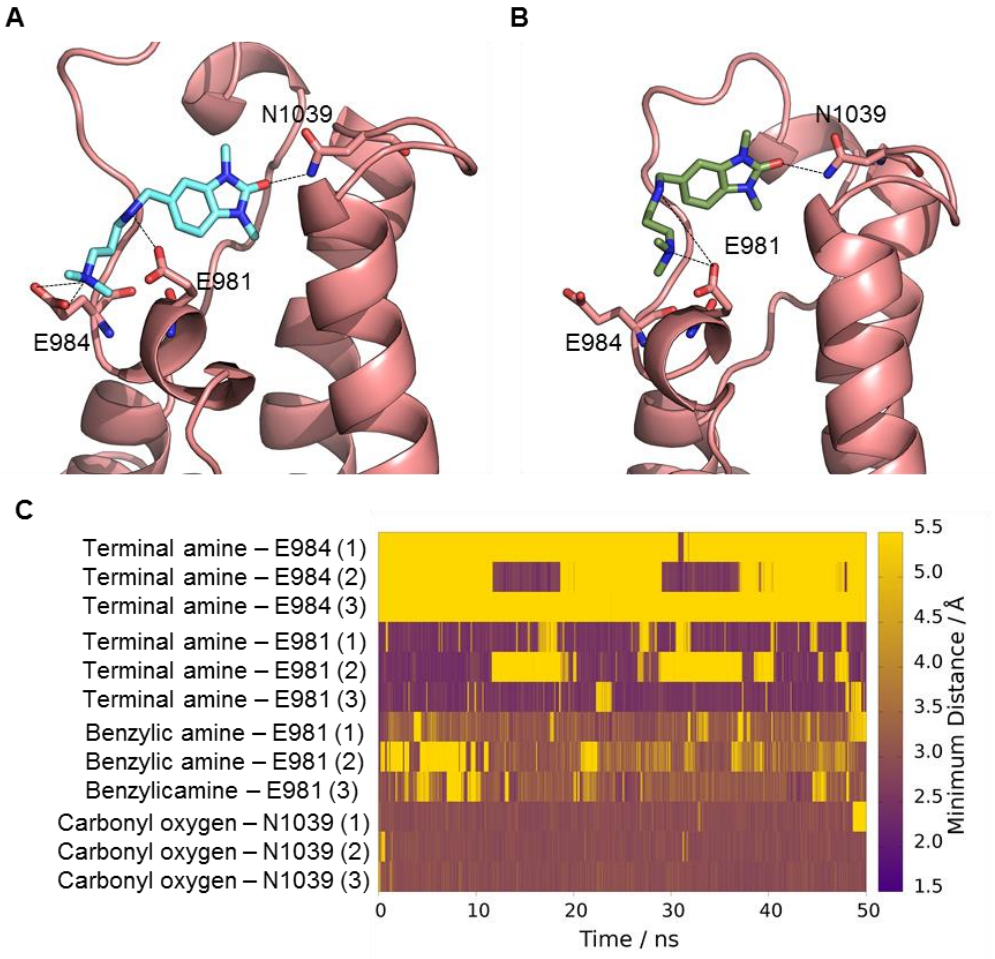


Figure 4.9. Molecular dynamics simulation of 62 in TRIM33β. The flexible amine chain is shown in two possible configurations – (A) terminal amine interacting with E984 and (B) terminal amine interacting with E981; C) Distance plot from molecular dynamics simulations showing distance between the compound (terminal amine, benzylic amine and carbonyl oxygen) and key amino acid residues E984, E981 and N1039.

From the distance plots, it is evident that the hydrogen bonds between the carbonyl oxygen of the headgroup and N1039, and that between the benzylic amine nitrogen and E981 are maintained throughout the simulation. This suggests that these interactions are key in

holding the ligand in place in the acetyl-lysine binding pocket. Interestingly, the terminal amine spends a large portion of the time bent around in a "pincer" formation and interacts more with E981 than E984. In the second replicate it is clear that the terminal amine flicks between E984 and E981 rather than maintaining an interaction with E984. As the binding pocket is close to the TRIM33β protein surface, the amine chain is not in a defined pocket but rather, is exposed to the solvent. This allows the amine chain to move more freely rather than forming a stable interaction that would help improve affinity. Consequently, it was hypothesised that analogues containing a more rigid amine would be the way forward. Alternatively, branched amines with substituents that would favour a conformation where the terminal amine would be orientated toward E984 could also help improve binding.

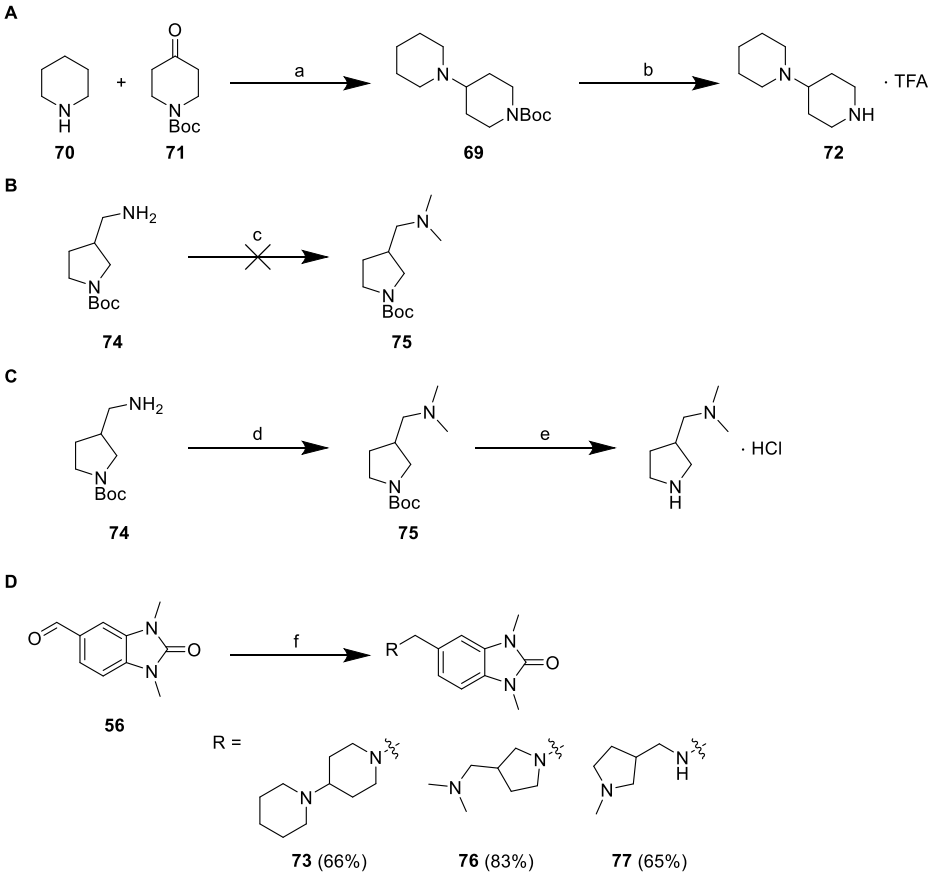
4.3.4 Structure-Activity Relationship Studies 2

4.3.4.1 *Synthesis of Compounds*

To explore the potential of more rigid amines to restrict amine conformation and orientate the terminal amine to make a more stable ionic interaction with E981 and E984, a series of amines incorporating cyclopentyl or cyclohexyl rings were synthesised. It was thought that a *cis*-alkene in the amine chain would be ideal to test this hypothesis. However, given the amine chain length, this would be an enamine, which is likely to be unstable. Amine **69** was synthesised in 81% yield *via* a reductive amination between piperidine (**70**) and the *tert*-butoxycarbonyl (Boc)-protected piperidone (**71**) (Scheme 4.6A). Compound **69** was then deprotected using trifluoroacetic acid (TFA) and excess TFA was removed using an azeotrope of toluene. The deprotected amine **72** was used immediately, as a TFA salt, in the reductive amination (Scheme 4.6D). As the amine was a TFA salt, the reductive amination was carried out under basic conditions following the procedure reported by Hart *et al*.[109] to yield compound **73**.

An initial attempt to dimethylate amine **74** using the Eschweiler-Clarke reaction[110] was unsuccessful (Scheme 4.6B). This reaction was selected to avoid over-alkylating the amine

to form the quaternary amine. The dimethylated amine **75** was subsequently obtained *via* reductive amination of amine **74** using formaldehyde and sodium cyanoborohydride,[111] to avoid overalkylation, in 84% yield (Scheme 4.6C). Amine **75** was deprotected with 4 M HCl in dioxane, and used immediately in a reductive amination with the headgroup aldehyde **56** (under basic conditions) to yield the desired final product **76** (Scheme 4.6D). Compound **77** was similarly synthesised by reacting the headgroup aldehyde **56** with the corresponding amine.



Scheme 4.6. Synthesis of analogues with different amine chains. *Reagents and conditions:* a) (i) 3 Å molecular sieves, Et_3N, 1,2-dichloroethane, 0 °C to rt, 3 h, (ii) $NaBH(OAc)_3$, rt, o/n, 81%; b) TFA, triisopropylsilane, CH_2Cl_2, rt, 0.5 h; c) Formic acid, formaldehyde (37% wt in H_2O), 80 °C, 24 h; d) (i) Formaldehyde (37% wt in H_2O), MeOH, rt, 2 h, (ii) $NaBH_3CN$, rt, o/n, 84%; e) 4 M HCl in dioxane, rt, 1 h; f) (i) amine, 3 Å molecular sieves, Et_3N, 1,2-dichloroethane, 0 °C to rt, 3 h, (ii) $NaBH(OAc)_3$, rt, o/n.

4.3.4.2 *Evaluation of Compounds*

In addition to compounds **73**, **76**, and **77**, compounds **78–85** (Figure 4.10) were synthesised by Amy Scorah and tested in a dose-response AlphaScreen™ assay followed by

NMR waterLOGSY validation of selected compounds (Table 4.6). The fused rings in compounds **81**–**83** restricted flexibility of the amine and placed the terminal methylated amine in specific orientations with respect to the headgroup. For compound **84**, the additional dimethyl on the amine chain could favour a "kinked" chain, causing the terminal amine to be folded towards E981. Compounds **85**–**89** allow the effect of phenyl group position and stereochemistry to be investigated.

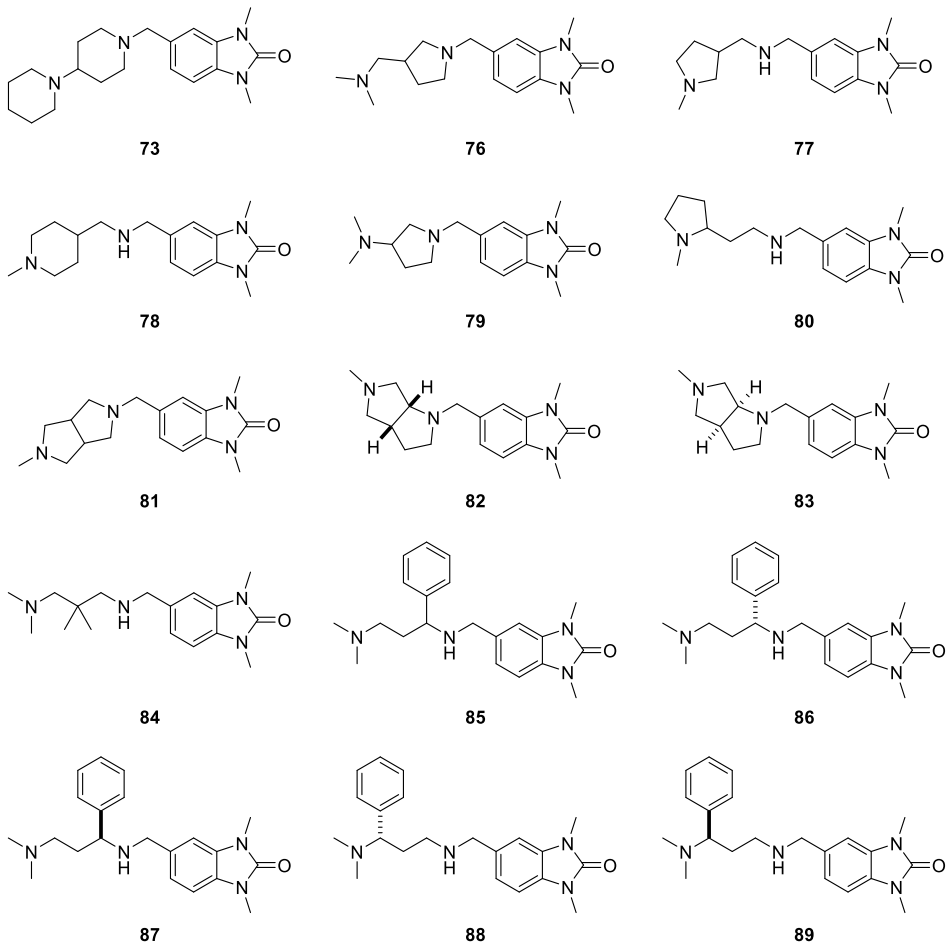
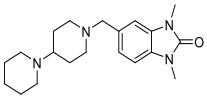
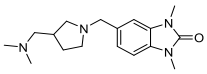
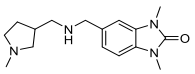
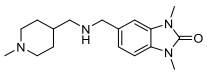
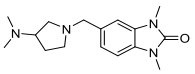
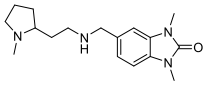
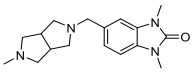
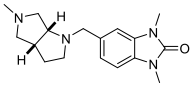
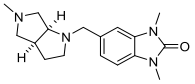
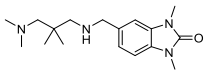
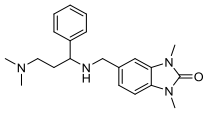
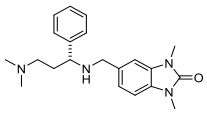
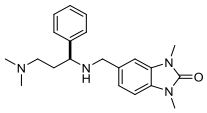
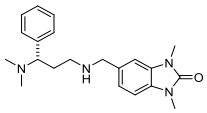
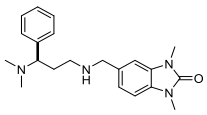


Figure 4.10. Compounds evaluated in a dose-response AlphaScreen™ assay.

Table 4.6. AlphaScreen™ IC50 values and NMR waterLOGSY data. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} values (green). Compounds that bind by waterLOGSY are indicated in pale blue; ✓ – binding, ✗ – not binding, ~ – inconclusive, – – compound not tested in waterLOGSY experiments. Ligand efficiency is indicated in parentheses.

Compound	NMR waterLOGSY			AlphaScreen™ IC_{50} (μM)		
	T24	T33α	T33β	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **73**	~	✓	✓	> 250	7.96 ± 1.11 (0.29)	40.7 ± 1.09 (0.25)
 **76**	~	✓	✓	> 250	5.81 ± 0.21 (0.33)	15.2 ± 0.79 (0.31)
 **77**	~	✓	✓	> 250	1.60 ± 0.14 (0.39)	8.53 ± 0.56 (0.34)
 **78**	~	✓	✓	> 250	5.49 ± 1.12 (0.33)	12.0 ± 1.12 (0.31)
 **79**	~	✓	✓	> 250	6.61 ± 0.54 (0.35)	12.8 ± 2.09 (0.33)
 **80**	~	✓	✓	> 250	1.52 ± 0.09 (0.37)	1.62 ± 0.14 (0.37)
 **81**	–	–	–	> 250	25.6 ± 6.82 (0.29)	100.4 ± 20.7 (0.25)
 **82**	~	✓	✓	44.25 ± 3.43 (0.28)	3.97 ± 0.86 (0.34)	12.25 ± 3.13 (0.31)
 **83**	~	✓	✓	> 250	5.79 ± 0.84 (0.33)	9.57 ± 1.74 (0.32)

 **84**	~	✓	✓	> 250	19.8 ± 3.56 (0.30)	39.2 ± 8.41 (0.28)
 **85**	✓	✓	✓	> 250	1.56 ± 0.10 (0.31)	4.83 ± 0.51 (0.29)
 **86**	~	✓	✓	> 250	2.45 ± 0.24 (0.30)	7.75 ± 0.60 (0.28)
 **87**	~	✓	✓	> 250	1.33 ± 0.06 (0.32)	4.52 ± 0.60 (0.29)
 **88**	~	✓	✓	> 250	5.94 ± 0.62 (0.28)	9.63 ± 2.00 (0.27)
 **89**	–	–	–	> 250	5.56 ± 0.56 (0.28)	12.2 ± 1.55 (0.26)

Compared to compound **62** (IC_{50} (TRIM33α) = 3.61 μM, IC_{50} (TRIM33β) = 13.7 μM), compound **73** has a lower binding affinity (IC_{50} (TRIM33α) = 7.96 μM, IC_{50} (TRIM33β) = 40.7 μM), especially for TRIM33β, while compounds **76**, **78** and **79** have similar binding affinity (IC_{50} (TRIM33α) = 5–7 μM, IC_{50} (TRIM33β) = 12–15 μM). This could be due to compound **73** having two piperidine rings and being too rigid or not sterically tolerated. Based on compounds **76**, **78** and **79**, there seems to be little difference in affinity whether the benzylic amine is secondary or tertiary. Comparing compounds **77** and **78** indicates that there is a small positive impact on having the 5-membered ring (**77**) as compared to the 6-membered ring (**78**). Given the surface-exposed area of the binding site, it is more likely that this is an effect of the relative orientation of the methylated amine

rather than that of ring size. Compound **80** has one carbon more than compound **77** between the terminal amine and the benzylic amine. While their IC_{50} values are similar for TRIM33α (1.52 μM and 1.60 μM, respectively), compound **80** has a lower IC_{50} value for TRIM33β (1.62 μM) than that of compound **77** (8.53 μM), suggesting that the additional carbon placed the amine in a position to bind better to TRIM33β. Calculation of the lipophilic ligand efficiency (LLE, where LLE = pIC_{50} – logP) demonstrates that compounds **77** and **80** have LLE values of 3.91 and 4.24, respectively. This indicates that the increase in affinity is not simply due to an increase in hydrophobic interactions.

The fused ring-containing compound **81** had a significantly lower binding affinity (IC_{50} (TRIM33α) = 25.6 μM, IC_{50} (TRIM33β) = 100.4 μM), indicating that this was not tolerated. On the other hand, compounds **82** and **83** had the fused rings in a different connectivity with respect to the headgroup. Both had similar binding to TRIM33α and TRIM33β, with the (*R*)-enantiomer **82** having a slightly lower IC_{50} value than the (*S*)-enantiomer **83**.

Interestingly, the dimethyl analogue **84** was not well tolerated and had IC_{50} values of 19.8 μM and 39.2 μM with TRIM33α and TRIM33β, respectively. It is possible that having a substituent at the middle carbon on the amine chain is not preferred. Conversely, compound **85** had improved binding with IC_{50} values of 1.56 μM and 4.83 μM with TRIM33α and TRIM33β, respectively. Analysis of the TRIM33β crystal structure highlighted the possibility that the phenyl group might be interacting with a proline-valine-proline (PVP) shelf (Figure 4.11) and that this additional hydrophobic interaction contributes to an increase in affinity. Furthermore, the additional phenyl group encourages the terminal amine to form an interaction with E981. As **85** was synthesised as a racemate, the (*R*)- and (*S*)-enantiomers (**86** and **87**, respectively) were then synthesised to investigate whether there was a preference for one enantiomer over the other. The IC_{50} values obtained suggested that there was a slight preference for the (*S*)-enantiomer **87** but that the

difference is minimal. Shifting the phenyl group to the end of the carbon chain (compounds **89** and **88**, (*R*)- and (*S*)-enantiomer, respectively) resulted in higher IC_{50} values and was not desirable.

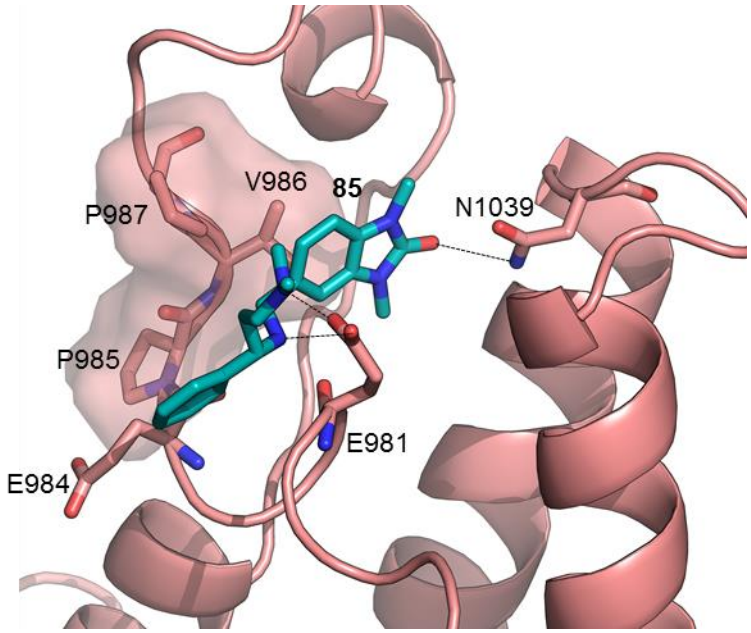


Figure 4.11. X-Ray crystal structure of TRIM33β (pink) with 85 (carbon = blue) docked into it showing potential interaction of the phenyl group with P985. This also encourages the terminal amine to form a hydrogen bond with E981 (black dashed lines).

While some improvement has been made in optimising the amine chain, the changes have indicated a relatively flat SAR and no drastic increase in affinity. On inspection of the crystal structure, there are several factors that might contribute to this. Although the benzimidazolone headgroup is in the acetyl-lysine binding pocket, the amine chain extends towards a much more solvent exposed region of the protein. This permits the flexible amine chain and E984 greater degree of freedom, resulting in less likely formation of a stable specific interaction. Furthermore, computational analysis of water molecules at binding site by Joe Bluck revealed that there are three stable water molecules in the ZA channel (Figure 4.12). The binding energy of each water molecule was calculated and was found to be 1.71, 0.82 and –3.33 kcal mol^{-1}. The two water molecules with binding energy 1.71 and 0.82 kcal mol^{-1} are likely to be displaceable. However, displacing these water molecules may involve a desolvation penalty. Conversely, a binding energy of –3.33 kcal mol^{-1} indicates that

the water is very stable in the protein structure and is unlikely to be displaceable. This is situated just beside E981, where the amine chain would be and hence, could provide an explanation for the relatively flat SAR observed. This water molecule edges the amine chain away from the protein and further out into the solvent. Therefore, any changes made to the amine chain has minimal effect on the compound binding to the protein.

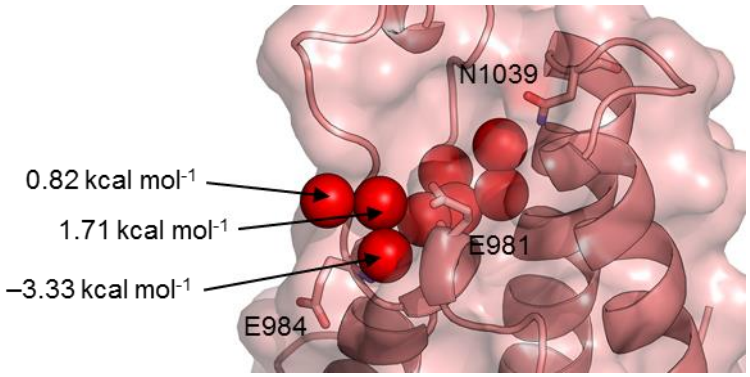


Figure 4.12. Water molecules (indicated by red spheres) in the acetyl-lysine binding pocket and ZA channel of the TRIM33β bromodomain (pink). The binding energies of the water molecules in the ZA channel are indicated and are a measure of its displaceability.

Taken together, the water molecules in the ZA channel are likely to significantly impact the binding of the ligand and are an important consideration in the design of further analogues for ligand optimisation. Structural waters at the base of the acetyl-lysine binding pocket have played a role in bromodomain ligand design over the years, and recent studies have demonstrated the value of understanding the ZA channel waters and how they might also affect ligand design.[112,113] While the ZA channel waters could hinder ligand binding to the protein, it also presents an opportunity as displacing the displaceable water molecules could bring a positive entropic contribution. Additionally, the amine chain could be designed to include a hydrogen bond acceptor in an appropriate configuration that would be able to make a hydrogen bond to the stable structural water. Therefore, understanding the implications of the ZA channel waters is key to improving ligand affinity.

4.4 Exploration of Hydrophobic Region

4.4.1 Computational Studies

Despite the modifications made and analogues synthesised so far, the SAR observed is relatively flat. As such, larger changes need to be made in order to probe potential interactions that would increase binding affinity. Computational studies by Joe Bluck were carried out to assess the favoured interactions that can be made with the protein. The region with potential for hydrophobic interactions is indicated in Figure 4.13A by a red mesh. The hydrophobic wall above the acetyl-lysine binding pocket (circled) presents an opportunity to improve affinity by installing an extension to one of the nitrogen atoms of the benzimidazolone headgroup (Figure 4.13B, vector indicated with a blue arrow). Efforts were then focused on synthesising and testing analogues with a variety of groups attached to the headgroup.

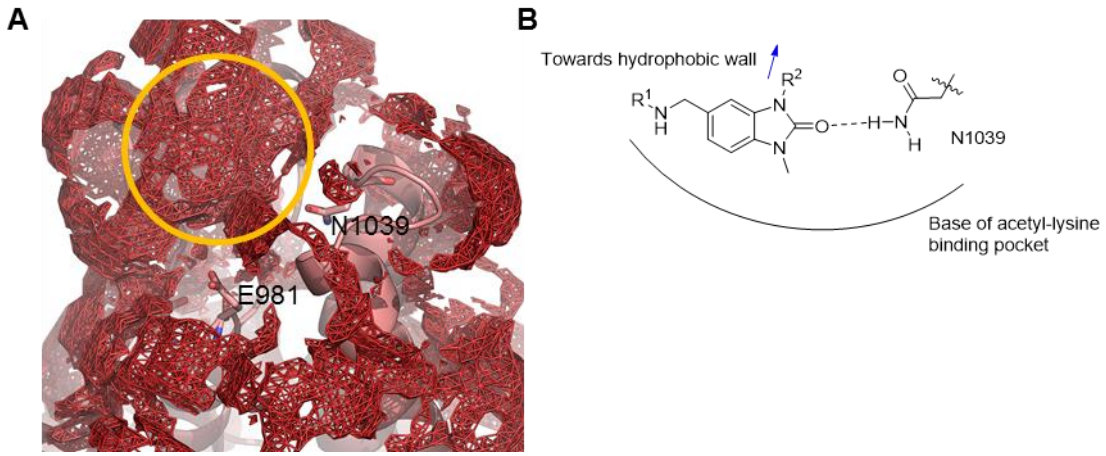


Figure 4.13. Potential for the exploration of hydrophobic region. A) X-Ray crystal structure of TRIM33β (pink) overlaid with mesh indicating regions for potential hydrophobic interactions (red). Area of interest is marked by a yellow circle. B) Schematic representation of proposed compounds in the binding pocket with the vector for expansion indicated by a blue arrow.

4.4.2 Investigation of Optimal Vector 1

4.4.2.1 *Synthesis of Compounds*

Up until this point, the orientation of the ligand in the binding pocket (with the amine either up or down, Figure 4.14A) had not been investigated as the dimethyl benzimidazolone headgroup is symmetrical. Molecular dynamics simulations by Joe Bluck also suggest that

both orientations of ligand binding can be accommodated by TRIM33β. However, with the addition of a hydrophobic group on one nitrogen, the headgroup is no longer symmetrical (Figure 4.14B). As we now wish to make unsymmetrical acetyl-lysine mimics, it is important to experimentally verify the optimal vector (blue arrows) for the amine chain. For clarity and consistency in discussion, the vector numbering will be as indicated in Figure 4.14B.

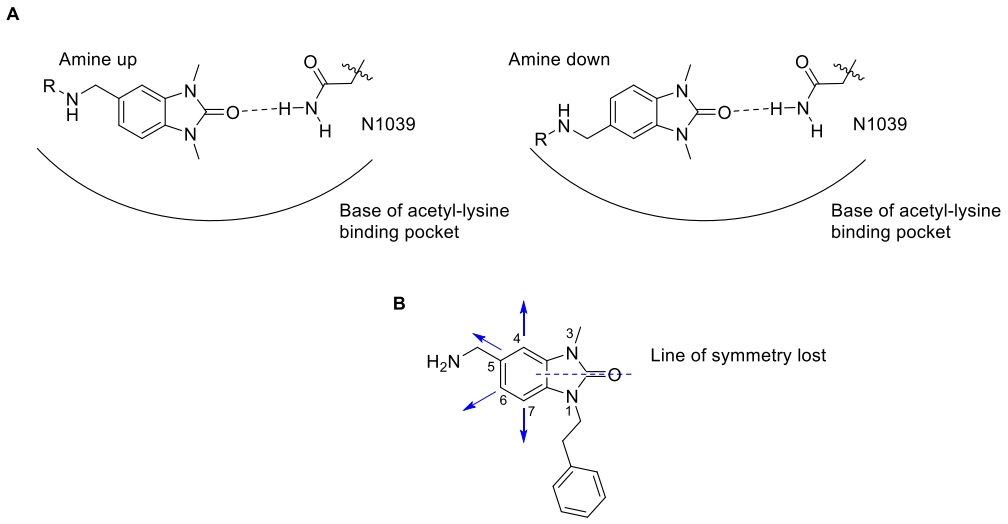
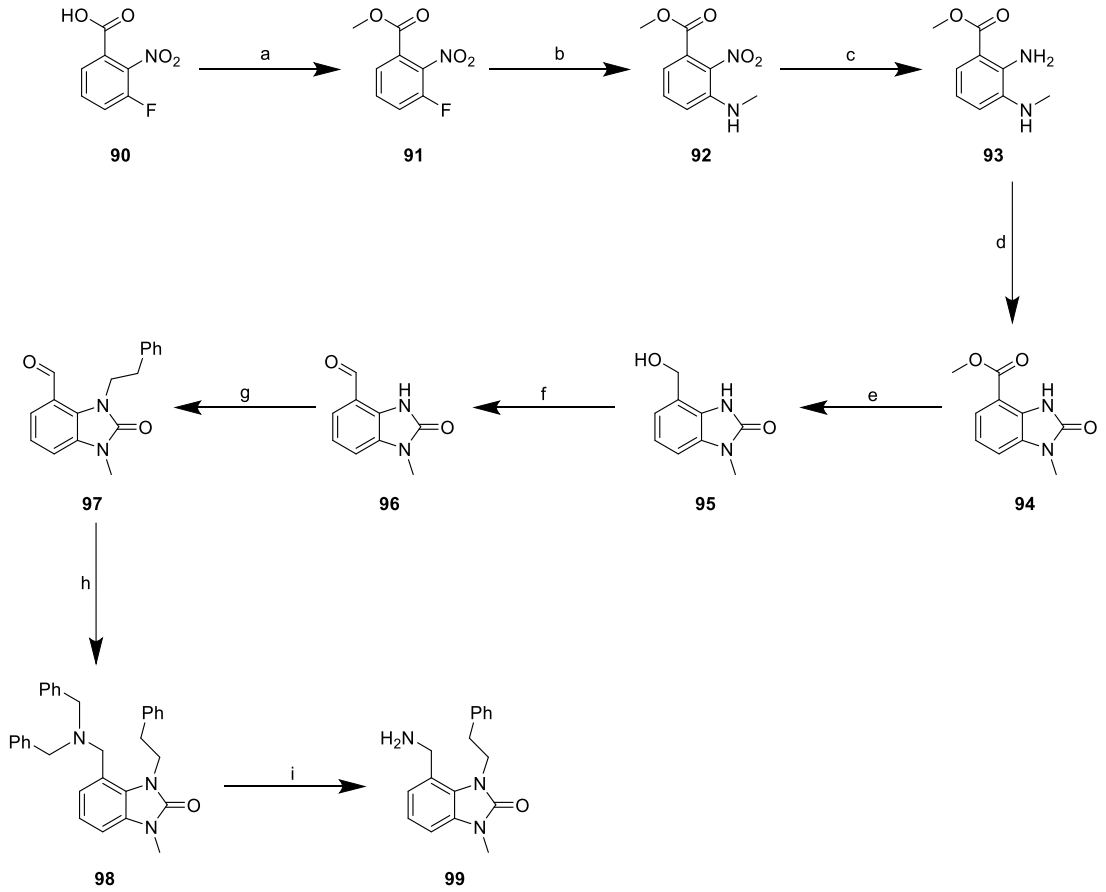


Figure 4.14. Different orientations of amine chain. A) Schematic representation showing two possible orientations of the ligand in the binding pocket; B) Schematic representation of the four possible vectors for the amine chain.

To compare the binding of the four vectors, each compound was synthesised as the free amine with a phenyl group attached to the headgroup by a 2-carbon chain. The amine in the 7-position was synthesised by first converting 3-fluoro-2-nitrobenzoic acid **90** to the methyl ester **91** (Scheme 4.7). The presence of the nitro group enabled synthesis of the monomethylated headgroup *via* nucleophilic aromatic substitution of the aromatic fluorine using methylamine hydrochloride, followed by hydrogenation of the nitro group to the primary amine **93**. Reaction of compound **93** with CDI afforded the benzimidazolone headgroup **94** in 82% yield. The key monomethylated headgroup aldehyde intermediate **96** was obtained by DIBAL reduction of methyl ester **94** to benzylic alcohol **95** then re-oxidation to the aldehyde using manganese(IV) oxide. The monomethylated headgroup aldehyde **96** was alkylated by deprotonation with potassium carbonate followed by

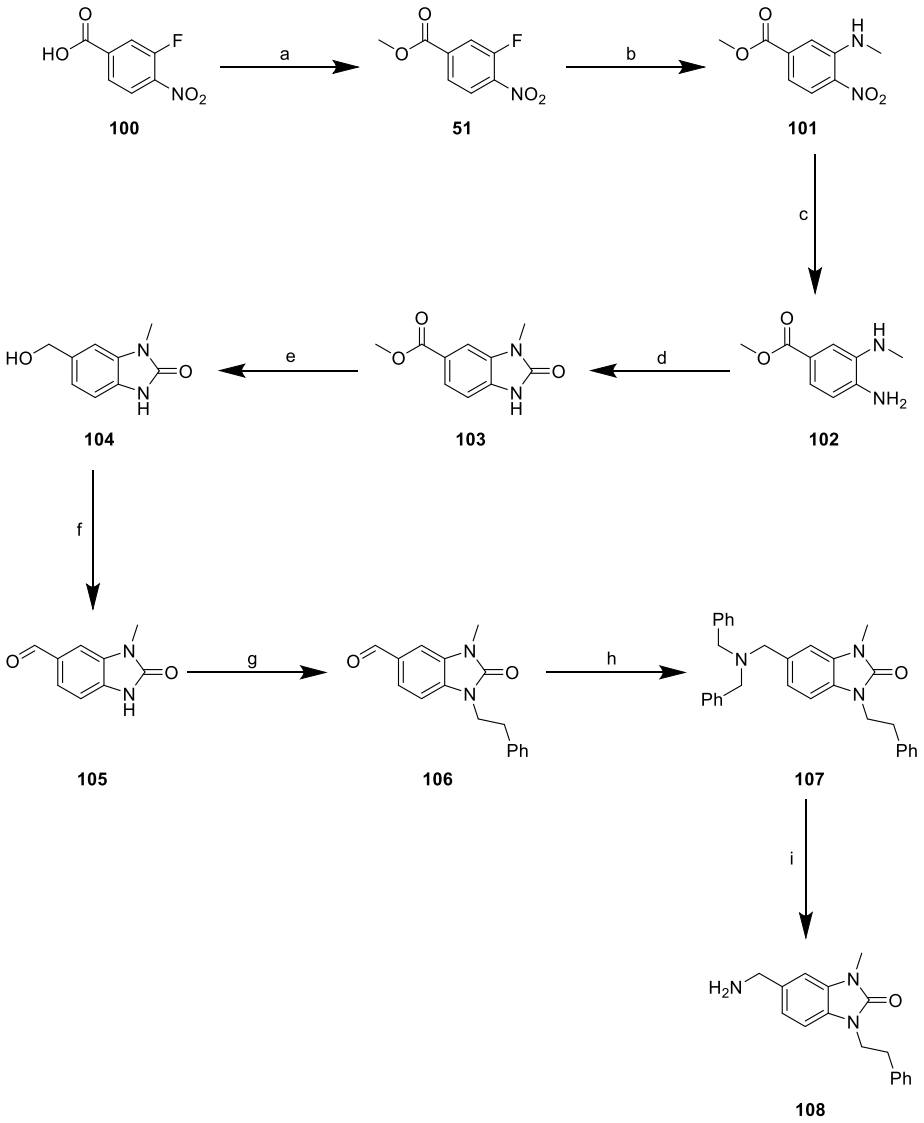
addition of (2-bromoethyl)benzene. Free amine **99** was obtained *via* a reductive amination with dibenzylamine to afford **98** followed by hydrogenation to remove both benzyl groups. Palladium black with HCl was found to be higher yielding than palladium on carbon or palladium hydroxide on carbon (Pearlman's catalyst), allowing the reaction to go to completion.



Scheme 4.7. Synthesis of compound with amine in the 7-position. *Reagents and conditions:* a) MeI, K_2CO_3, DMF, rt, 1 h, 94%; b) Methylamine hydrochloride, *N*,*N*-diisopropylethylamine, DMF, rt, o/n, 73%; c) H_2, Pd/C, EtOAc/EtOH 2:1, rt, 4 h, 73%; d) CDI, CH_2Cl_2, rt, 1.5 h, 82%; e) DIBAL-H (1 M in hexanes), THF, –78 °C to rt, o/n, 72%; f) MnO_2, CH_2Cl_2, rt, 2 h, 96%; g) (i) K_2CO_3, DMF, rt, 0.5 h, (ii) (2-Bromoethyl)benzene, 80 °C, o/n, 82%; h) (i) Dibenzylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 66%; i) Pd black, H_2, HCl, EtOH, rt, o/n, 84%.

The analogue with the amine in the 5-position was synthesised in the same manner from 3-fluoro-4-nitrobenzoic acid (**100**) (Scheme 4.8). The synthesis worked well on a 7 g (38 mmol) scale up to the DIBAL reduction, which was carried out on a smaller scale (~3 g, 16 mmol). As with compound **99**, the free amine **108** was synthesised by a reductive

amination of the aldehyde **106** with dibenzylamine, followed by benzyl deprotection of compound **107** using palladium black in the presence of HCl.



Scheme 4.8. Synthesis of compound with amine in the 5-position. *Reagents and conditions:* a) MeI, K_2CO_3, DMF, rt, 1 h, 89–97%, n = 2; b) Methylamine hydrochloride, *N*,*N*-diisopropylethylamine, DMF, rt, o/n, 94–99%, n = 2; c) H_2, Pd/C, EtOAc/EtOH 2:1, rt, 4 h; d) CDI, CH_2Cl_2, rt, 5 h, 70–87% over two steps, n = 2; e) DIBAL-H (1 M in hexanes), THF, –78 °C to rt, 6 h, 55–91%, n = 3; f) MnO_2, CH_2Cl_2, rt, o/n, 76–91%, n = 2; g) (i) K_2CO_3, DMF, rt, 0.5 h, (ii) (2-Bromoethyl)benzene, 80 °C, o/n, 85%; h) (i) Dibenzylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 67%; i) Pd black, H_2, HCl, EtOH, rt, o/n, 85%.

4.4.2.2 *Evaluation of Compounds*

Analogues with the free amine in positions 4 and 6 (**109** and **110**, respectively), as well as compound **111**, in which the benzimidazolone headgroup was unmethylated, were synthesised by Amy Scorah (Figure 4.15). These compounds were tested in a dose-response AlphaScreen™ (Table 4.7).

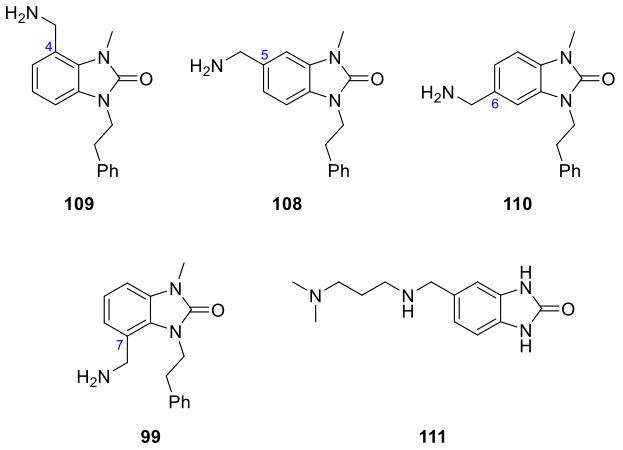
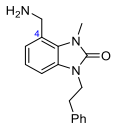
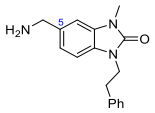
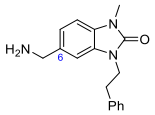
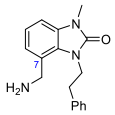
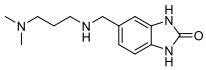


Figure 4.15. Compounds tested in dose-response AlphaScreen™. The vector position number is indicated in blue.

Table 4.7. AlphaScreen™ IC_{50} values. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} values (green). Ligand efficiency is indicated in parentheses.

Compound	AlphaScreen™ IC_{50} (μM)		
	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **109**	> 250	> 250	> 250
 **108**	> 250	12.1 ± 1.12 (0.33)	> 250
 **110**	> 250	23.3 ± 2.34 (0.31)	> 250
 **99**	> 250	> 250	> 250
 **111**	> 250	> 250	> 250

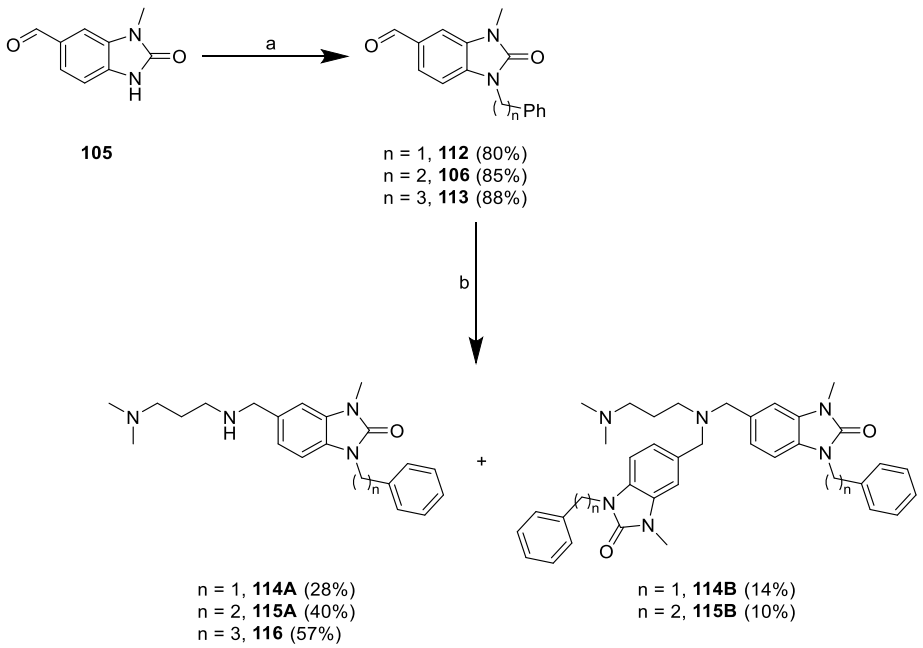
Having the amine in position 4 and 7 (**109** and **99**) resulted in no binding, which is not surprising based on the hypothesised binding mode. Compounds **108** and **110** both did not bind to TRIM33β but bound to TRIM33α with IC_{50} values of 23.3 μM and 12.1 μM, respectively. This indicates that position 5 is favoured over position 6. However, the dramatically weaker binding could be due to the absence of an amine chain or the hydrophobic group on the headgroup being sub-optimal. The large difference in affinity observed suggest that when developing this series of compounds with an unsymmetrical

benzimidazolone headgroup, it is important that the orientation of the amine chain with respect to the headgroup extension is optimised.

4.4.3 Investigation of Optimal Vector 2

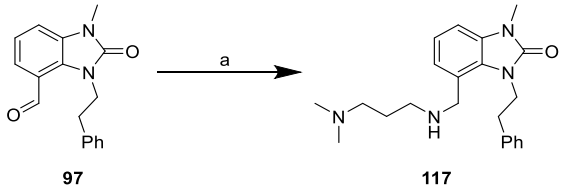
4.4.3.1 *Synthesis of Compounds*

In the next set of compounds, the number of carbon atoms in the headgroup extension was varied (n = 1–3) and an amine chain was installed. The amine chain from **62** was selected for consistency and for its robustness in the reductive amination. The monomethylated benzimidazolone headgroup aldehyde **105** could be easily alkylated with the relevant bromide before submitting to a reductive amination with 3-(dimethylamino)-1-propylamine (Scheme 4.9). It was found that some side product was formed where the amine product (**114A** and **115A**) had reacted with the aldehyde starting material **112** and **106** to form side products **114B** and **115B**, respectively, possibly due to the accidental use of the aldehyde in excess.



Scheme 4.9. Synthesis of analogues with varying carbon chain lengths on headgroup extension. *Reagents and conditions:* a) (i) K_2CO_3, DMF, rt, 0.5 h, (ii) Benzyl bromide or (2-bromoethyl)benzene or 1-bromo-3-phenylpropane, 80 °C, o/n; h) (i) 3-(Dimethylamino)-1-propylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n.

Additionally, to confirm that the lack of binding observed in compound **99** was due to the position of the amine rather than the lack of an amine chain, compound **117** was synthesised as well (Scheme 4.10).

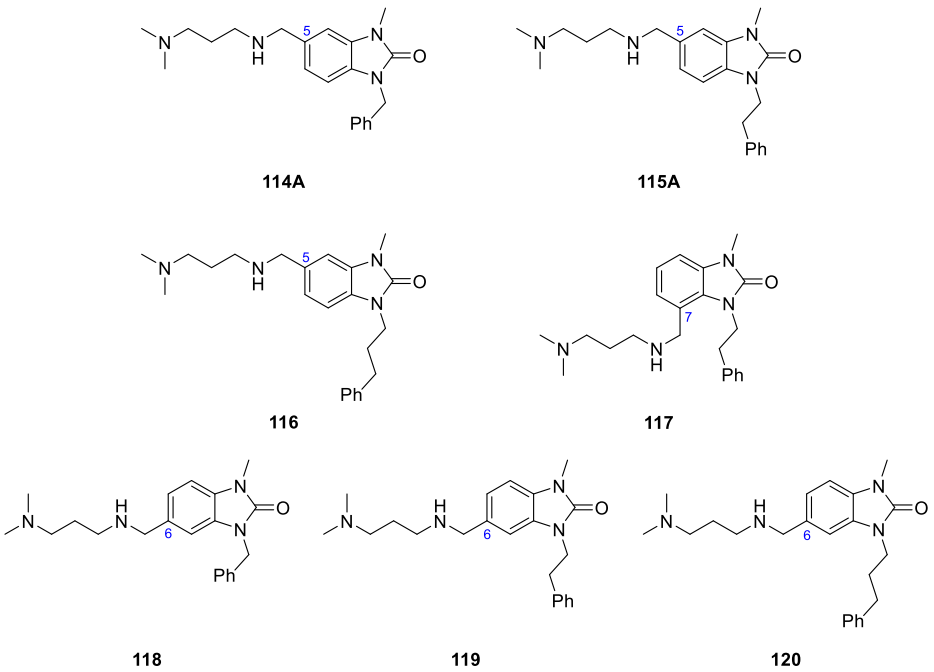


97 **117**

Scheme 4.10. Synthesis of analogue with vector at position 7. *Reagents and conditions:* a) (i) 3-(Dimethylamino)-1-propylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 86%.

4.4.3.2 *Evaluation of Compounds*

In addition to compounds **114A**–**117**, analogues with the amine chain in position 6 (compounds **118**–**120**) were synthesised by Amy Scorah (Figure 4.16). These compounds were tested in a dose-response AlphaScreen™ assay (Table 4.8).



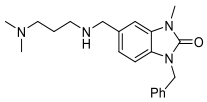
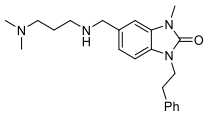
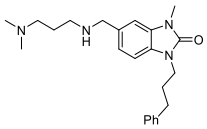
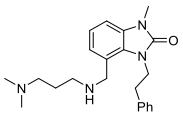
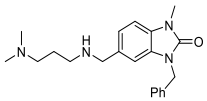
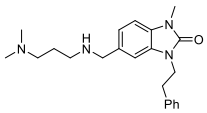
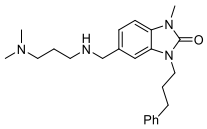
114A **115A**

116 **117**

118 **119** **120**

Figure 4.16. Compounds tested in dose-response AlphaScreen™.

Table 4.8. AlphaScreen™ IC_{50} values and NMR waterLOGSY data. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} values (green). Compounds that bind by waterLOGSY are indicated in pale blue; ✓ – binding, × – not binding, ~ – inconclusive, – – compound not tested in waterLOGSY experiments. Ligand efficiency is indicated in parentheses.

Compound	NMR waterLOGSY			AlphaScreen™ IC_{50} (μM)		
	T24	T33α	T33β	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **114A**	~	✓	✓	> 250	3.64 ± 1.11 (0.29)	86.0 ± 2.12 (0.22)
 **115A**	~	✓	✓	> 250	4.34 ± 1.08 (0.28)	59.7 ± 1.12 (0.22)
 **116**	~	✓	✓	> 250	2.47 ± 1.08 (0.28)	100.0 ± 1.10 (0.20)
 **117**	–	–	–	> 250	> 250	> 250
 **118**	–	–	–	> 250	14.8 ± 1.12 (0.26)	> 250
 **119**	–	–	–	> 250	5.37 ± 1.10 (0.27)	> 250
 **120**	–	–	–	> 250	8.76 ± 1.15 (0.25)	> 250

Based on the AlphaScreen™ results, compounds **118–120** (amine chain in 6-position) had no binding to TRIM33β but retained some binding to TRIM33α. Compounds **114A–116** (amine chain in 5-position) had similar IC_{50} values (2–4 μM) to compound **62** for TRIM33α (IC_{50} = 3.61 μM) but much higher IC_{50} values (60–100 μM) for TRIM33β (as compared to compound **62** (IC_{50} = 13.7 μM)). This observation might be accounted for by the difference in the loop region of the two isoforms. The 17-amino acid insertion in TRIM33α results in the conserved asparagine (N1039) pointing away from the acetyl-lysine binding pocket which could lead to a larger space available to accommodate the pendant hydrophobic group. Compounds **114A–116** exhibit stronger binding than compounds **118–120**, indicating that the amine chain in position 5 is preferred over position 6. While this confirms the hypothesis that there is a clear preference for one vector over the other, the addition of the hydrophobic group has not enhanced binding affinity. It is possible that the carbon chain allows flexibility, giving the pendant hydrophobic group freedom of movement in the solvent exposed loop region above the acetyl-lysine binding pocket. Consequently, different groups with less flexibility will be explored.

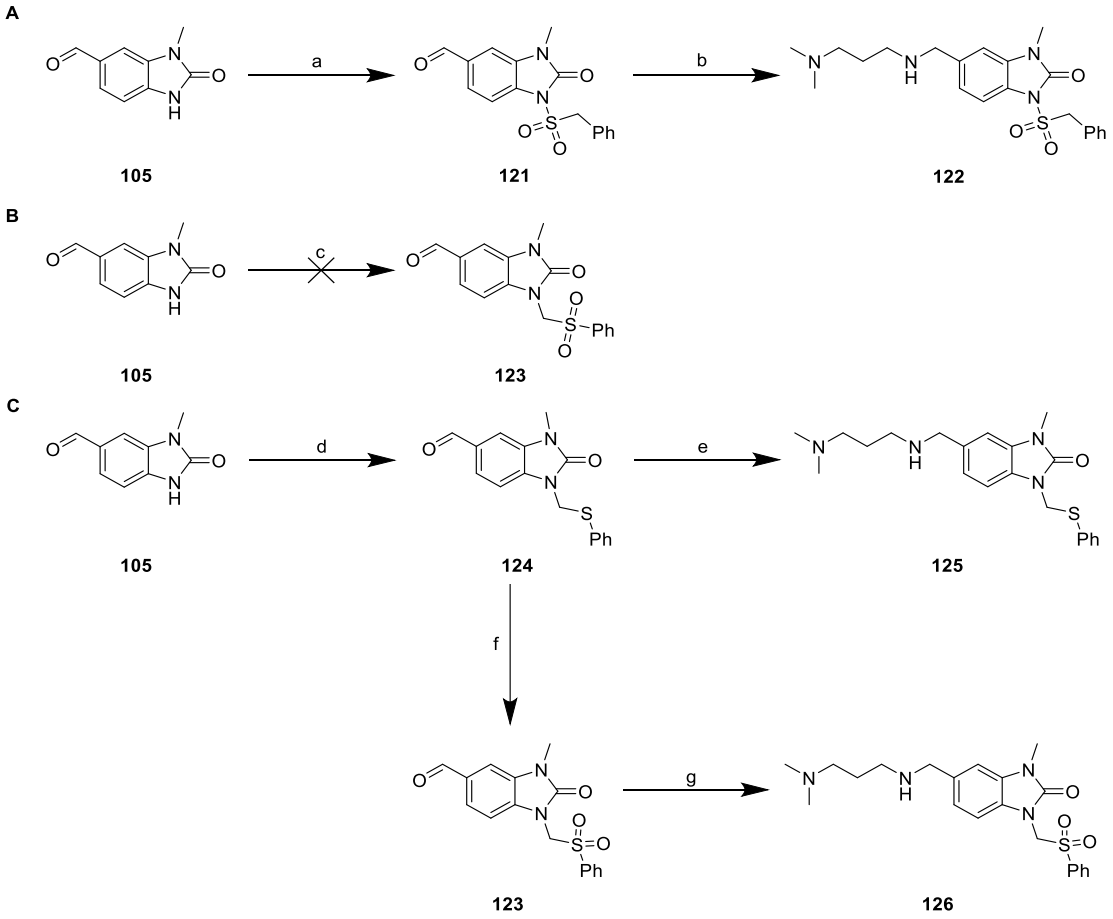
Furthermore, comparison of the carbon chain lengths in compounds **114A** (n = 1), **115A** (n = 2) and **116** (n = 3) revealed that compound **115A** had the lowest IC_{50} value of 59.7 μM as compared to 86.0 μM and 100 μM in compounds **114A** and **116**, respectively. This suggests that having the hydrophobic phenyl three bonds away from the headgroup is ideal and this was taken into consideration in the design of subsequent analogues.

4.4.4 Structure-Activity Relationship Studies

4.4.4.1 *Synthesis of Compounds*

To introduce some rigidity to the hydrophobic extension, sulfone analogues of **115A** were synthesised. Sulfonamide analogue **122** was synthesised by reaction of monomethylated headgroup aldehyde **105** with phenylmethanesulfonyl chloride using 4-dimethylaminopyridine (4-DMAP) as a catalyst followed by a reductive amination of the

resulting aldehyde[114] (Scheme 4.11A). Attempts to alkylate monomethylated headgroup **105** with chloromethyl phenyl sulfone using the standard alkylation conditions were unsuccessful, even on addition of tetrabutylammonium iodide (TBAI), heating up to 85 °C[115] (Scheme 4.11B). Compound **126** was eventually synthesised by alkylation of monomethylated headgroup **105** chloromethyl phenyl sulfide and subsequent oxidation to sulfone **123** using 3-chloroperbenzoic acid[116] (Scheme 4.11C).



Scheme 4.11. Synthesis of sulfone analogues. *Reagents and conditions:* a) Phenylmethanesulfonyl chloride, Et_3N, 4-dimethylaminopyridine, THF, rt, o/n, 39%; b) (i) 3-(Dimethylamino)-1-propylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 96%; c) (i) K_2CO_3 or Cs_2CO_3, DMF, rt, 0.5 h, (ii) Chloromethyl phenyl sulfone, with or without TBAI, rt up to 85 °C, o/n; d) (i) K_2CO_3, DMF, rt, 0.5 h, (ii) Chloromethyl phenyl sulfide, rt °C, o/n, 75%; e) (i) 3-(Dimethylamino)-1-propylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 71%; f) 3-Chloroperbenzoic acid, CH_2Cl_2, 0 °C to rt, 2 h, 94%; g) (i) 3-(Dimethylamino)-1-propylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 63%.

Computational studies by Joe Bluck proposed that a phenyl group attached directly to the benzimidazolone would favour hydrophobic interactions with the region of interest above

the acetyl-lysine binding pocket. However, the close proximity of the two aromatic protons (Figure 4.17) suggested that the phenyl group would be at an angle relative to the benzimidazolone headgroup. Consequently, a smaller cyclopentyl group was attached by reacting aldehyde **105** with bromocyclopentane in the presence of TBAI in 42% yield, followed by reductive amination to furnish compound **129** in 96% yield (Scheme 4.12).

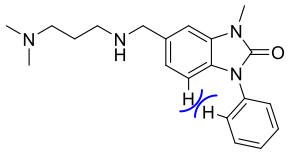
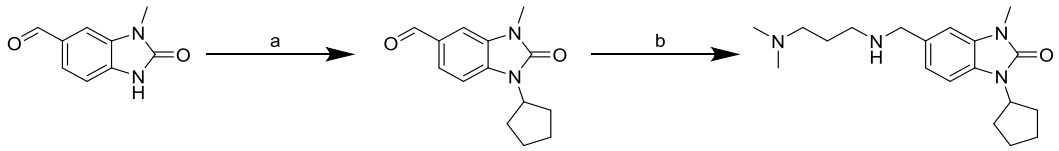


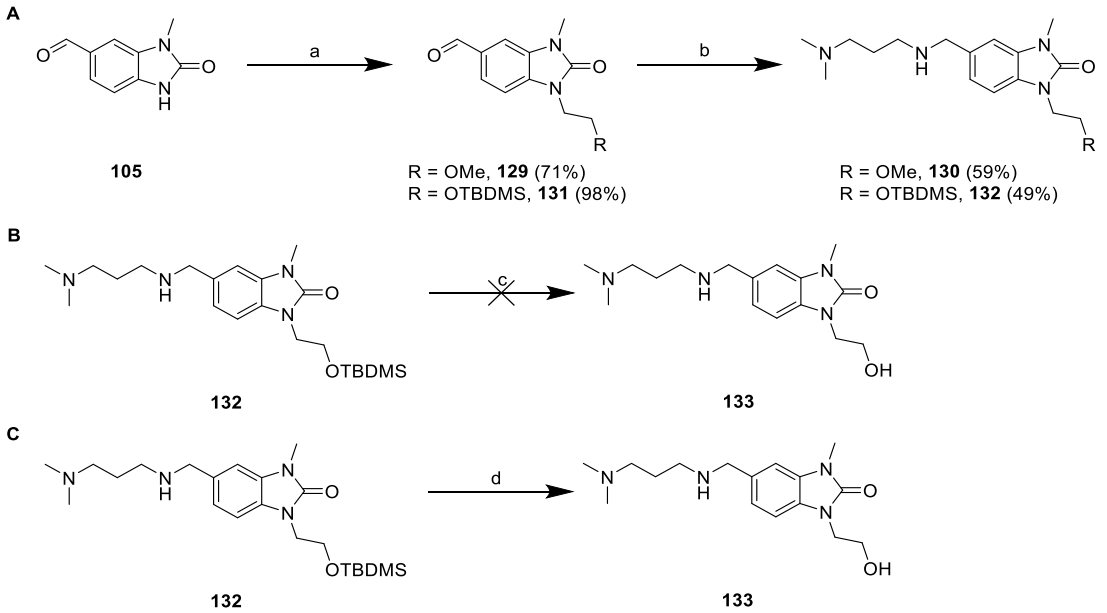
Figure 4.17. Compound proposed by computation modelling studies.



Scheme 4.12. Synthesis of cyclopentyl analogue. *Reagents and conditions:* a) (i) K_2CO_3, DMF, rt, 0.5 h, (ii) Bromocyclopentane, rt °C, o/n, 42%; b) (i) 3-(Dimethylamino)-1-propylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n, 96%.

In addition, computational studies by Joe Bluck suggested that a 2-carbon alkyl chain with a methoxy or hydroxyl group had the potential to make favourable interactions with the peptide backbone of the hydrophobic wall above the acetyl-lysine binding pocket, or an additional hydrogen bond with the conserved asparagine (N1039). Alkylation of monomethylated headgroup **105** with 2-bromoethyl methyl ether furnished aldehyde **129**, which enabled access to compound **130** in 59% yield (Scheme 4.12A). To synthesise the hydroxyl analogue, monomethylated headgroup **105** was reacted with the protected alcohol using 2-bromoethoxy-*tert*-butyl dimethylsilane to give compound **131**. Subsequent reductive amination yielded compound **132** which could then be deprotected. Test deprotection of the *tert*-butyldimethylsilyl (TBDMS) protecting group using tetrabutylammonium fluoride (TBAF) was unsuccessful and it seemed that the headgroup

was not stable to TBAF (Scheme 4.12B). Deprotection using 1 M HCl in methanol proved successful and afforded **133** in 87% yield (Scheme 4.12C).



Scheme 4.13. Synthesis of potential hydrogen bond forming analogues. *Reagents and conditions:* a) (i) K_2CO_3, DMF, rt, 0.5 h, (ii) 2-Bromoethyl methyl ether or 2-bromoethoxy-*tert*-butyl dimethylsilane, rt °C, o/n; b) (i) 3-(Dimethylamino)-1-propylamine, acetic acid, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n; c) Tetrabutylammonium fluoride (1 M in THF), THF, rt, 1 h; d) HCl, MeOH, rt, 1 h, 87%.

4.4.4.2 *Evaluation of Compounds*

Compounds **122**, **125**, **126**, **128**, **130**, and **133** were tested together with compounds synthesised by Amy Scorah and Jessica Reynolds (Figure 4.18) in a dose-response AlphaScreen™ assay. Binding of relevant compounds were then confirmed in NMR waterLOGSY experiments.

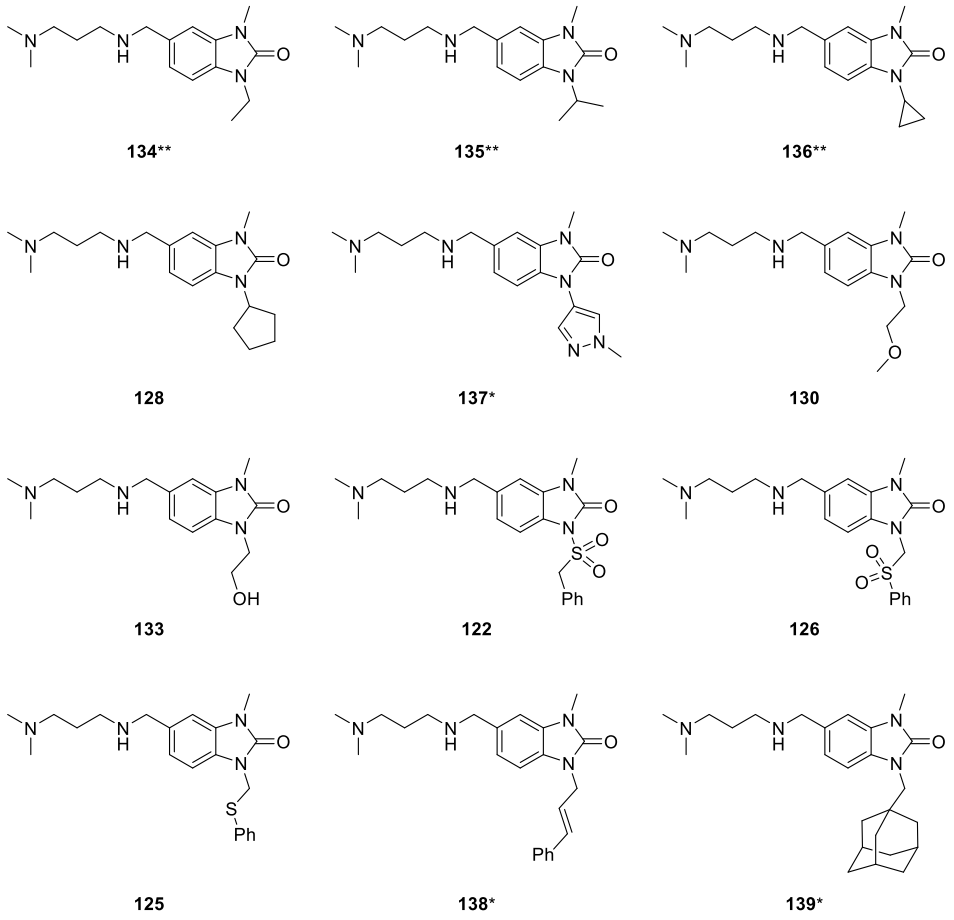
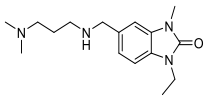
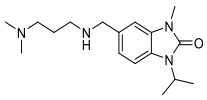
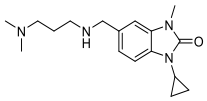
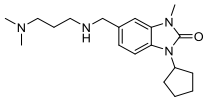
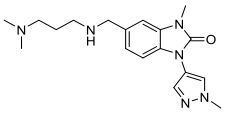
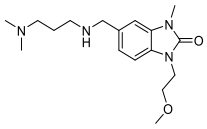
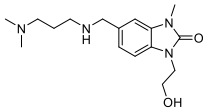
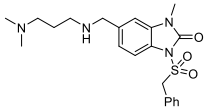
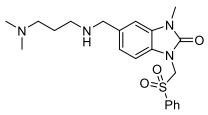
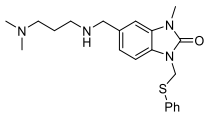
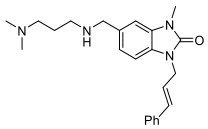
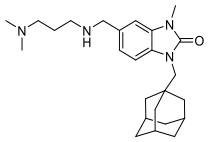


Figure 4.18. Compounds tested in dose-response AlphaScreen™. *Compounds synthesised by Amy Scorah. **Compounds synthesised by Jessica Reynolds.

Table 4.9. AlphaScreen™ IC_{50} values and NMR waterLOGSY data. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} values (green). Compounds that bind by waterLOGSY are indicated in pale blue; ✓ – binding, × – not binding, ~ – inconclusive, – – compound not tested in waterLOGSY experiments. Ligand efficiency is indicated in parentheses.

Compound	NMR waterLOGSY			AlphaScreen™ IC_{50} (μM)		
	T24	T33α	T33β	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **134**	×	✓	✓	> 250	2.07 ± 0.20 (0.38)	14.7 ± 1.47 (0.32)
 **135**	×	✓	✓	> 250	3.22 ± 0.25 (0.35)	19.6 ± 1.64 (0.30)
 **136**	–	–	–	124 ± 32.2 (0.25)	9.01 ± 1.66 (0.32)	96.0 ± 24.9 (0.26)
 **128**	✓	✓	✓	> 250	4.32 ± 0.35 (0.30)	60.2 ± 7.42 (0.20)
 **137**	–	–	–	> 250	> 250	> 250
 **130**	–	–	–	> 250	16.4 ± 1.10 (0.29)	> 250
 **133**	✓	✓	✓	> 250	8.74 ± 0.82 (0.32)	20.6 ± 1.94 (0.30)
 **122**	–	–	–	> 250	> 250	> 250

 **126**	–	–	–	> 250	> 250	> 250
 **125**	–	–	–	> 250	1.42 ± 0.13 (0.30)	> 250
 **138**	–	–	–	> 250	2.01 ± 0.24 (0.28)	53.3 ± 7.89 (0.21)
 **139**	–	–	–	> 250	> 250	> 250

Ethyl and isopropyl analogues (**134** and **135**, respectively) had comparable IC_{50} values to compound **62** (IC_{50} (TRIM33α) = 3.61 μM, IC_{50} (TRIM33β) = 13.7 μM). Cyclopropyl **136**, however, had weaker binding to TRIM33α (IC_{50} = 9.01 μM) and dramatically weaker binding to TRIM33β (IC_{50} = 96.0 μM). Compound **136**, with an IC_{50} value of 124 μM, also showed some binding to TRIM24. This is in contrast with compounds **134** and **135** which clearly showed no binding to TRIM24 by both AlphaScreen™ and waterLOGSY. Compounds **130** and **133** both showed weaker binding to TRIM33α (IC_{50} = 16.4 μM and 8.74 μM, respectively). Compound **130** showed no binding to TRIM33β while compound **133** had an IC_{50} value of 20.6 μM. Taken together, the addition of this extension has not improved binding to either isoform and in **130**, negates binding. This result suggests that, contrary to the initial hypothesis, it is unlikely that the methoxy or hydroxyl group has been able to form any hydrogen bonds that might improve binding affinity. Given the solvent exposure of the area above the acetyl-lysine binding pocket, it is possible that any potential favourable

interactions is not able to compensate for the desolvation penalty that would be involved in a binding event.

More rigid analogues of **115A** (**122** and **126**) showed no binding as compared to **115A**, indicating that the larger sulfone moiety is not tolerated and interferes with the interaction of the hydrophobic extension. The addition of an adamantyl group (**139**) results in no binding and could indicate that larger bulk, including that observed in the sulfones, is not tolerated. Thioether analogue **125**, however, binds to only TRIM33α and not TRIM33β, possibly due to the difference in the loop region. Interestingly, compound **138**, with an *E*-alkene, had similar binding affinity to TRIM33α (IC_{50} = 2.01 μM) as the alkane analogue **116** (IC_{50} = 2.47 μM), and significantly stronger binding to TRIM33β (IC_{50} = 53.3 μM) than **116** (IC_{50} = 100 μM). Despite this, binding is still significantly weaker than having only the methyl groups on the headgroup.

Cyclopentyl **128** (IC_{50} (TRIM33α) = 4.32 μM and IC_{50} (TRIM33β) = 60.2 μM) had stronger binding to TRIM33β than cyclopropyl (IC_{50} = 96.0 μM). Molecular dynamics studies by Joe Bluck suggested that this could be accounted for by its potential to form an isoleucine zipper-like interaction with a nearby isoleucine (I990) (blue circle in Figure 4.19). This was found to be a stable interaction in simulations and would help enhance ligand affinity. However, the expected improvement in binding as compared to dimethyl analogue **62** was not observed. Further molecular dynamics studies by Joe Bluck suggested that the cyclopentyl group could be pulling the headgroup in a slightly different angle, placing the benzylic amine in a less than optimal distance from E981. Hence, the two interactions are not optimised.

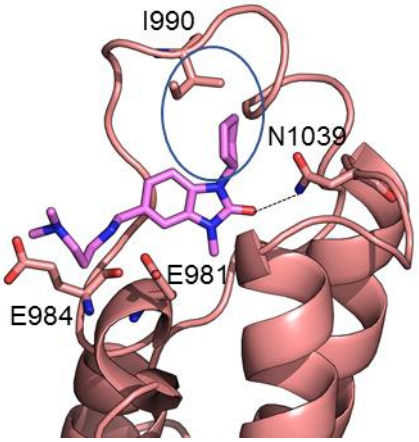
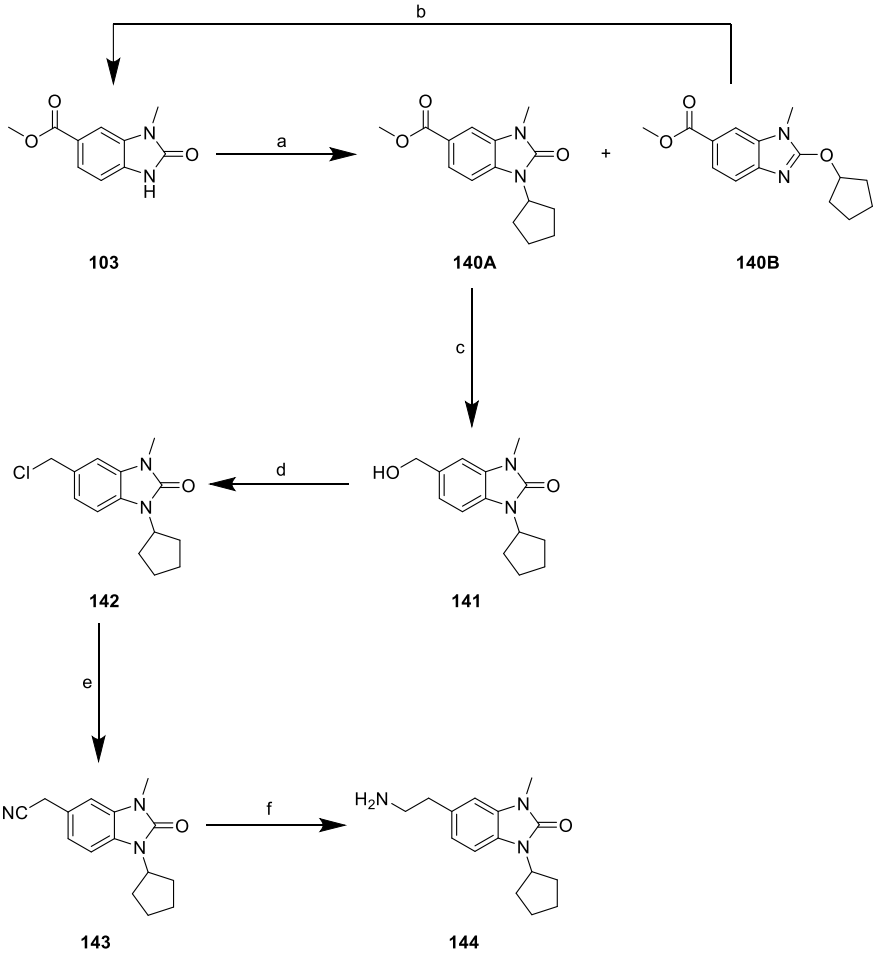


Figure 4.19. Compound 128 (carbon = bright pink) in X-ray crystal structure of TRIM33β (pink) with potential isoleucine zipper interaction circled in blue.

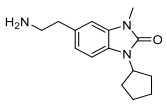
4.4.4.3 *Follow-Up of Cyclopentyl Compound*

To investigate this, an analogue was synthesised containing the cyclopentyl as well as a γ-amine (Scheme 4.14). Alkylation of the monomethylated headgroup ester **103** with bromocyclopentane afforded cyclopentyl **140A** in 78% yield and an *O*-alkylated side product **140B** in 14% yield, which was hydrolysed to compound **103** by heating under reflux with hydrobromic acid. The ester **140A** was reduced to the alcohol using DIBAL then chlorinated using thionyl chloride before submitting to a nucleophilic substitution reaction with sodium cyanide to yield nitrile **143**. Nitrile **143** was then hydrogenated with palladium on carbon in the presence of HCl to furnish γ-amine **144** in 75% yield. Compound **144** was tested in a dose-response AlphaScreen™ assay and was found to bind to TRIM33α with an IC_{50} value of 41.4 μM but not to TRIM33β (Table 4.10). This is lower than that of dimethyl analogue **52**, which was previously found to be 65.1 μM (Table 4.2). Dimethyl analogue **52** similarly had no binding to TRIM33β. This suggests that the cyclopentyl group has a positive effect on binding affinity but the distance of the amine chain would have to be optimised in order to maximise binding. Unfortunately, due to time constraints, the cyclopentyl analogue with a γ-amine chain was not synthesised.



Scheme 4.14. Synthesis of γ-amine cyclopentyl analogue 144. *Reagents and conditions:* a) (i) K_2CO_3, DMF, rt, 0.5 h, (ii) Bromocyclopentane, rt, o/n, 71%; b) HBr (48% wt in H_2O), EtOH, reflux, 4 h, 78% (side product 14%); c) DIBAL-H (1 M in hexanes), THF, –78 °C to 0 °C, 2 h, 93%; d) Thionyl chloride, CH_2Cl_2, rt, 1 h; e) NaCN, TBAI, DMF, 70 °C, o/n, 86% over two steps; f) H_2, Pd/C (10 mol%), HCl, EtOH, rt, o/n, 75%.

Table 4.10. AlphaScreen™ IC_{50} values. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} values (green). Ligand efficiency is indicated in parentheses.

Compound	AlphaScreen™ IC_{50} (μM)		
	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **144**	> 250	41.4 ± 0.32 (0.32)	> 250

4.5 ITC Experiments

Throughout this project, AlphaScreen™ was used as a reliable experiment for comparing IC_{50} values, enabling compounds to be ranked easily. However, as AlphaScreen™ is a peptide displacement assay, the IC_{50} values obtained are a function of the inherent binding affinity of the protein for its peptide binding partner. This impacts interpretation of binding affinity in two ways: 1) the absolute binding affinity is not known and an easily displaceable peptide would give a false understanding of the compound affinities, and 2) compound affinities cannot be compared between protein-peptide combinations and hence, selectivity between TRIM33α and TRIM33β cannot be quantified.

To investigate compound binding affinity independent of the peptide binding partner, five compounds with the lowest IC_{50} values by AlphaScreen™ (**59**, **58**, **77**, **80**, and **87**) and control compound **62** were tested in ITC experiments with TRIM33α and TRIM33β. Compounds were initially tested with 160 μM protein in the syringe and 16 μM compound in the cell. However, due to the weak binding affinity of these compounds, the heat change on each injection of protein was too small for titration curves to be obtained. This suggests that the binding affinity is close to the limit of detection by ITC. Consequently, the protein and ligand concentrations were increased to 363 μM for TRIM33α and 368 μM for TRIM33β, and 40 μM, respectively. Pleasingly, this enabled ITC curves to be obtained and an example of that for compound **80** is shown in Figure 4.20 (other ITC curves can be found in Appendices C) and the results are summarised in Table 4.11 below.

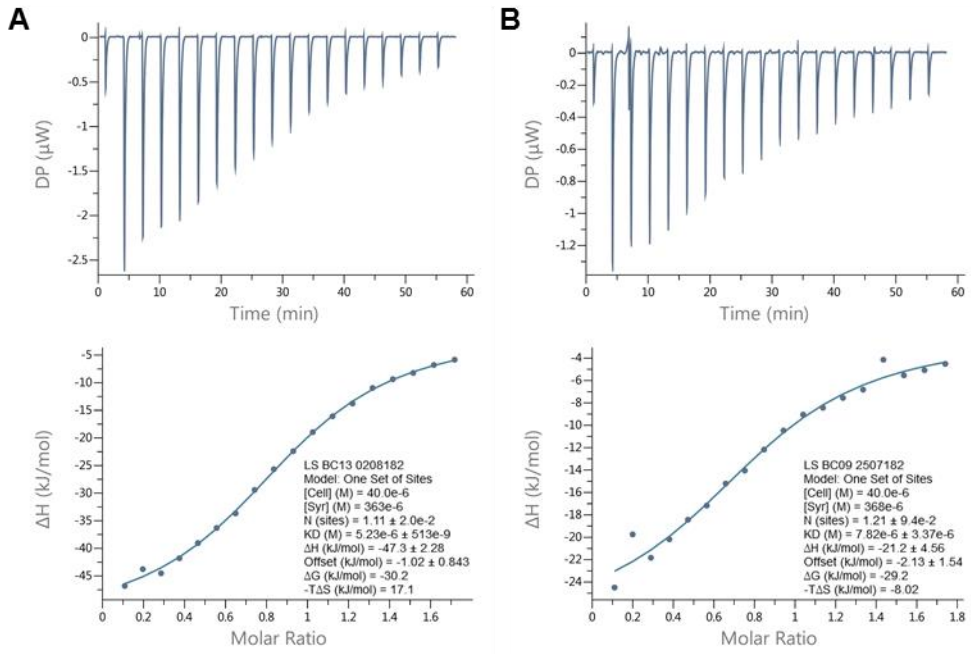
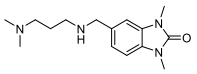
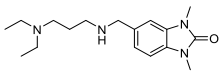
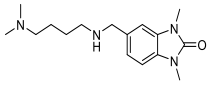
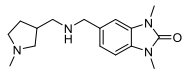
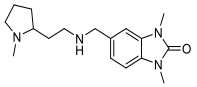
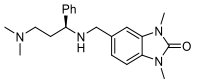


Figure 4.20. Example of ITC curves obtained for compound 80 with TRIM33α (A) and TRIM33β (B).

Comparison of the ITC K_D values with AlphaScreen™ IC_{50} values facilitates understanding of how the AlphaScreen™ data relates to absolute binding affinity and indicates that in general, the AlphaScreen™ data tends to overstate the affinity of the compound. For example, control compound **62** has AlphaScreen™ IC_{50} values of 3.61 µM and 13.7 µM but K_D of 9.20 µM and 16.6 µM for TRIM33α and TRIM33β, respectively. While the ITC K_D values do not map the exact same trend as the AlphaScreen™ IC_{50} values, the compound with the highest affinity identified using both techniques is compound **87** with IC_{50} values of 1.33 µM and 4.52 µM and K_D of 2.71 µM and 7.33 µM for TRIM33α and TRIM33β, respectively.

Table 4.11. ITC results for TRIM33α and TRIM33β. AlphaScreen™ IC_{50} values have been included for comparison. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} values (green). *These IC_{50} values are the mean and standard deviation of multiple repeats.

Compound	ITC K_D (μM)		AlphaScreen™ IC_{50} (μM)		
	TRIM33α	TRIM33β	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **62**	9.20 ± 1.22	16.6 ± 4.54	> 250	3.61* ± 1.15	13.7* ± 9.34
 **59**	5.46 ± 0.46	11.3 ± 2.55	> 250	4.74 ± 1.14	2.02 ± 1.06
 **58**	9.92 ± 0.82	18.3 ± 2.71	> 250	7.96 ± 0.90	2.73 ± 0.37
 **77**	5.26 ± 0.27	7.81 ± 1.10	> 250	1.60 ± 0.14	8.53 ± 0.56
 **80**	5.23 ± 0.51	7.82 ± 3.37	> 250	1.52 ± 0.01	1.62 ± 0.14
 **87**	2.71 ± 0.17	7.33 ± 1.18	> 250	1.33 ± 0.06	4.52 ± 0.60

The difference observed between K_D values and IC_{50} values is because K_D refers to the dissociation constant associated with a ligand binding to a protein and is a direct measurement of the binding affinity of the ligand for its target protein. In contrast, IC_{50} values refer to the concentration at which the biological response (in this case in the AlphaScreen™, binding of the protein for the peptide binding partner) is reduced by half. Therefore, IC_{50} values often vary between assays and are dependent on the particular assay set-up and conditions used.

4.6 Selectivity Against BRD4

BRD4, a member of the BET family, is a protein with a very strong phenotype. To test if the TRIM33 ligands are selective against BRD4, compounds **20** and **62** were evaluated in a dose-response AlphaScreen™ BRD4 and its peptide binding partner $H4_{1\text{-}20}(KAc)_4$ up to 500 μM compound FAC. Both compounds showed little inhibition of the protein-peptide interaction and full IC_{50} curves could not be obtained (Figure 4.21), indicating that these compounds do not bind to BRD4 and hence, are selective.

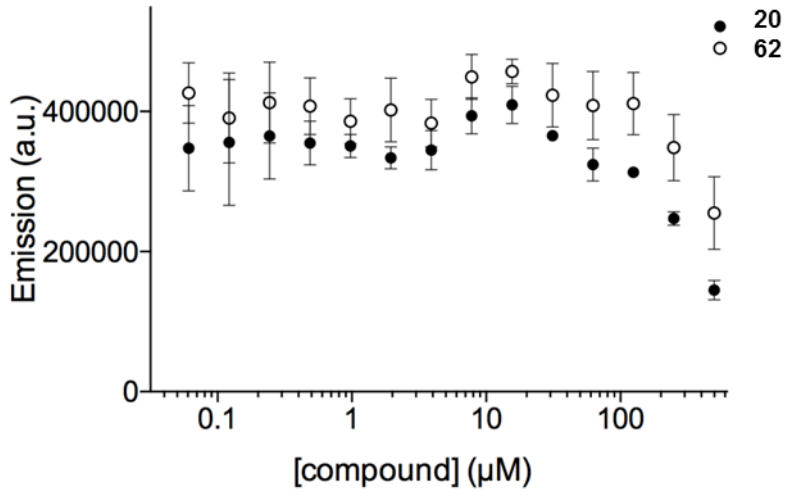


Figure 4.21. AlphaScreen™ data from dose-response experiment of compounds 20 and 62 with BRD4.

4.7 Conclusions

This chapter describes work to investigate the SAR in relation to hit compound **20** and to improve its binding affinity to TRIM33. Initial SAR studies established the benzimidazolone moiety as a ligand efficient headgroup that forms a hydrogen bond with N1039 through the carbonyl oxygen. The benzylic amine was also identified as making a key hydrogen bond interaction with E981. Both interactions were maintained in the design of analogues for testing. Although the addition of an amine chain was found to be beneficial while maintaining ligand efficiency at ~0.3, a flat SAR was observed with most compounds having IC_{50} values between 2–8 μM and 4–15 μM for TRIM33α and TRIM33β, respectively. Computational studies revealed that this is likely to be due to the presence of three water

molecules in the ZA channel, one of which is unlikely to be displaceable. Consequently, the amine chain is likely to be extending into the solvent rather than forming specific interactions that would increase affinity, resulting in little improvement in binding affinity. The implications of the water molecules should then be taken into consideration in the optimisation of the amine chain.

Additionally, the hydrophobic area above the acetyl-lysine binding pocket was explored following the synthesis of a series of compounds with varying groups extending off the nitrogen. The resulting non-symmetric headgroup led to the experimental confirmation of the optimal vector combination between the headgroup extension and the amine chain. Unfortunately, despite the successful identification of the most favoured amine chain vector, AlphaScreen™ experiments indicated that only a methyl or ethyl group was tolerated, and larger groups resulted in weaker binding. ITC experiments were carried out as an independent assay for measuring binding affinity without relying on additional components in the assay. It was shown that the AlphaScreen™ IC_{50} values tend to imply stronger binding than the K_D values obtained by ITC, particularly for TRIM33α because of its weaker affinity to the peptide binding partner.

Taken together, compound **87** (Figure 4.22) was found to have an IC_{50} values of 1.33 μM and 4.52 μM and K_D of 2.71 μM and 7.33 μM for TRIM33α and TRIM33β, respectively.

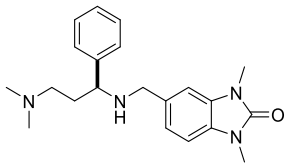


Figure 4.22. Compound 87.

In the process of optimising the benzimidazolone series of compounds, a deeper understanding of the acetyl-lysine binding pocket in TRIM33 has been achieved. Further efforts to improve ligand affinity should capitalise on this, particularly in light of the structural water molecules present in the bromodomain.

Chapter 5

Development of Covalent Ligands for Mutant TRIM33β-E984C

5.1 Introduction and Aims

Structure-activity relationship (SAR) studies in Chapter 4 suggested that an increased understanding of the TRIM33β acetyl-lysine binding site was necessary for further development of an improved ligand for the bromodomain. A crystal structure with a ligand bound to the TRIM33β bromodomain would provide a deeper understanding of its binding mode and enable structure-guided design. However, attempts to obtain an X-ray crystal structure, in collaboration with Professor Stefan Knapp, were unsuccessful.

The aim of the work described in this chapter was to develop covalent ligand analogues of compound **62** with an electrophile on the amine chain, that would react with a cysteine nucleophile. It was hypothesised that glutamic acid E984 could be replaced with cysteine, which could then react with the covalent ligand. E984 is well-placed as it is in close proximity to the terminal amine of **62** but sufficiently distant from the benzimidazolone core and benzylic amine, enabling the binding mode, particularly with respect to the ZA channel water molecules, to be studied. Moreover, SAR studies indicate that E984 is situated in a solvent-exposed region, and likely has minimal impact on the binding of ligand **62** to TRIM33β. By designing the ligand with an appropriate distance between the electrophile and benzimidazolone headgroup, the bromodomain ligand in the acetyl-lysine binding pocket could react with an E984C mutant (Figure 5.1). In this study, analogues of varying electrophiles and chain lengths were synthesised, and alkylation experiments with mutant TRIM33β-E984C were carried out to validate this approach.

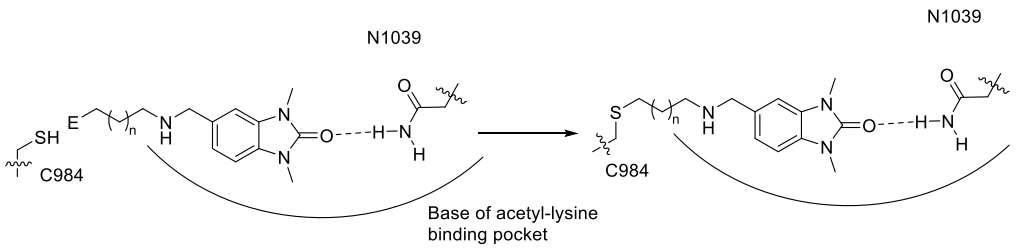


Figure 5.1. Schematic representation of reaction between electrophile (E) on the covalent ligand with mutated residue cysteine C984.

5.2 Docking Studies

A series of covalent ligands retaining the benzimidazolone headgroup and benzylic amine were designed with an electrophilic group replacing the terminal amine. Based on known covalent inhibitors[117,118] and selected for their ease of synthesis, these included acrylamides (**145**), vinyl sulfonamides (**146**), methyl acrylates (**147**) and methyl vinylsulfones (**148**) (Figure 5.2A). Using the CLC Drug Discovery Workbench, these compounds were docked into TRIM33β with an E984C mutation introduced using the PyMOL mutagenesis wizard. An example of the result obtained with compound **146** is shown in Figure 5.2B. Due to the flexibility of the amine chain and amino acid residues (E981 and C984), it is difficult to predict the optimum chain length that would place the electrophile in the correct position for reaction with C984.

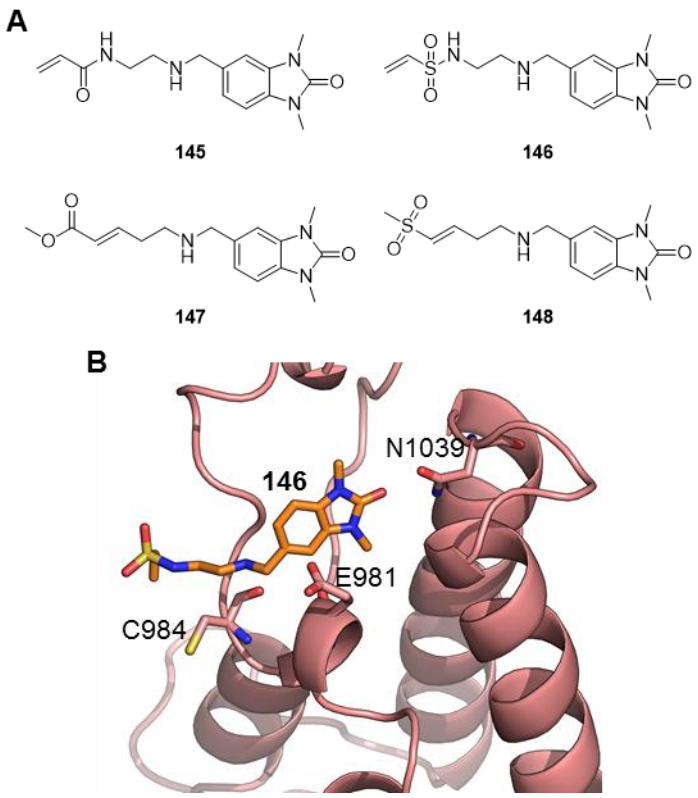
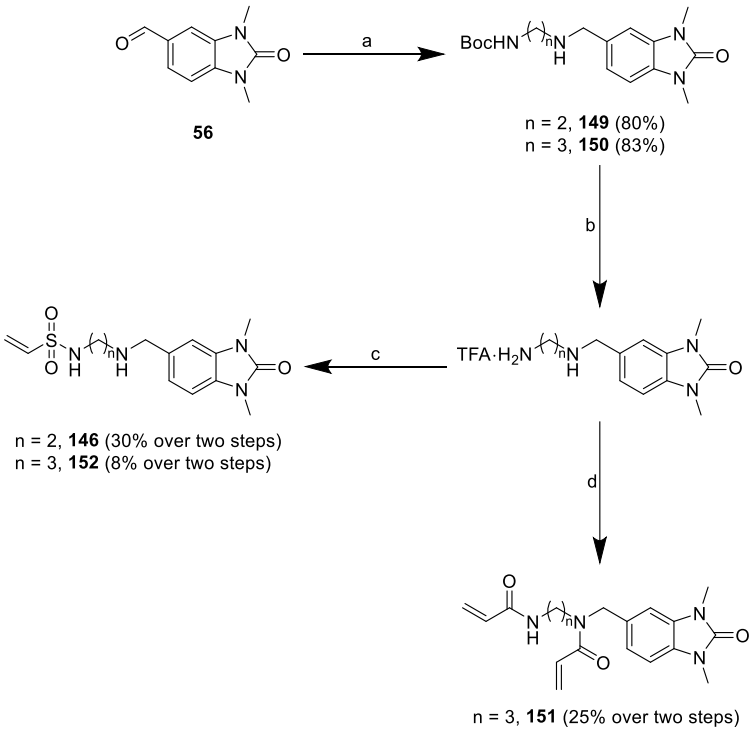


Figure 5.2. Docking to aid design of covalent ligands. A) Selection of compounds used in docking; B) X-Ray crystal structure of TRIM33β (pink) with E984C mutation introduced using PyMOL mutagenesis wizard, with sulfonamide **146** (carbon = orange) docked using CLC Drug Discovery Workbench.

5.3 Synthesis of Alkylating Compounds

5.3.1 Acrylamide and Vinyl Sulfonamide Synthesis

To synthesise acrylamide **151** and vinyl sulfonamides **146** and **152**, the key Boc-protected intermediates, **149** and **150**, were obtained *via* a reductive amination of aldehyde **56** with the corresponding Boc-protected amine (Scheme 5.1). Amines **149** and **150** were deprotected using TFA, and reacted immediately with 2-chloroethanesulfonyl chloride or acryloyl chloride.[118] Vinyl sulfonamides **146** and **152** (2-carbon and 3-carbon chain respectively) were obtained in a low yield due to reaction of the secondary amine, forming the di-sulfonamide product. Acylation with acryloyl chloride exclusively formed the diacylated product **151**, even on reduction of equivalents of acryloyl chloride and base.

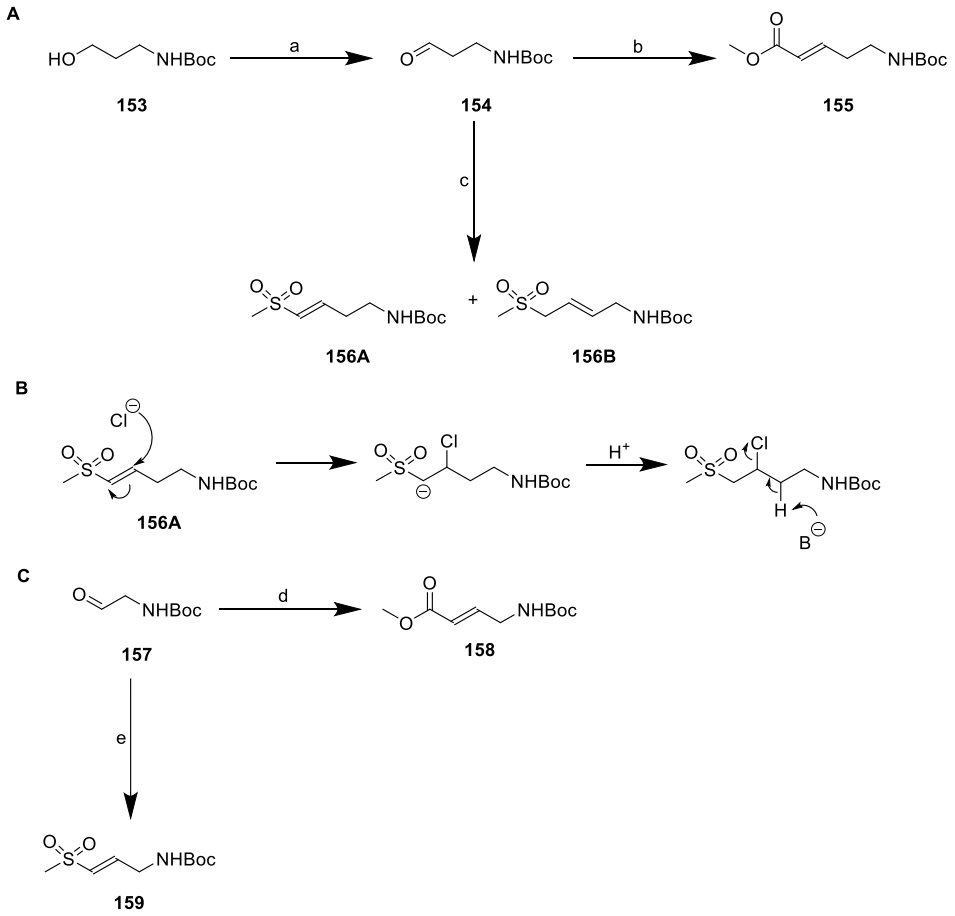


Scheme 5.1. Synthesis of vinyl sulfonamide and acrylamide. *Reagents and conditions:* a) (i) *N*-Boc-ethylenediamine or *N*-Boc-1,3-propanediamine, Et_3N, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n; b) TFA, Triisopropylsilane, CH_2Cl_2, rt, 0.5 h; c) 2-Chloroethanesulfonyl chloride, Et_3N, CH_2Cl_2, 0 °C, 3 h; d) Acryloyl chloride, *N*,*N*-diisopropylethylamine, CH_2Cl_2, 0 °C, 2 h.

5.3.2 Methyl Acrylate and Methyl Vinyl Sulfone Synthesis

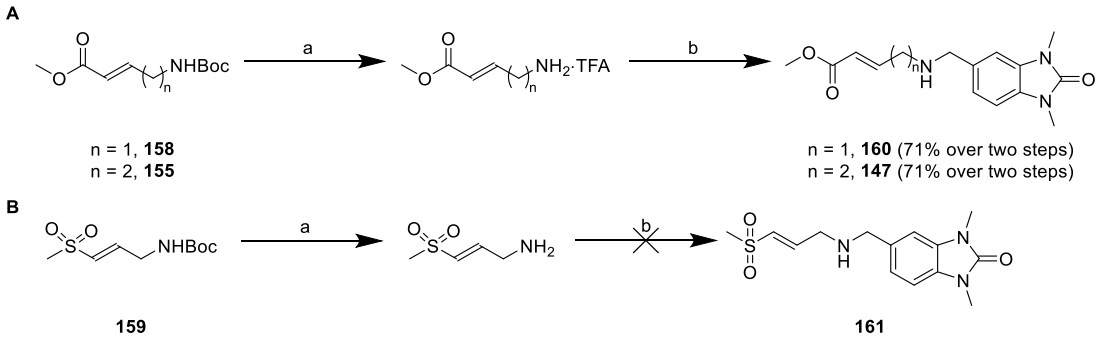
Oxidation of 3-(Boc-amino)-1-propanol (**153**)[119] to aldehyde **154** provided acrylate **155** *via* a Horner-Wadsworth-Emmons reaction with trimethylphosphonoacetate in 41% yield (Scheme 5.2A). Reaction of the aldehyde **154** with diethyl (methylsulfonylmethyl)phosphonate to form vinyl sulfone **156A**, however, also resulted in the formation of isomer **156B** as an inseparable isomer in a 3:1 ratio. Analysis of the NMR coupling constants confirmed that **156B** was, in fact, the *E*-isomer and not the *Z*-isomer. Given that lithium chloride is present in the reaction, it is proposed that the chloride ion acts as a nucleophile and reacts with the vinyl sulfone (Scheme 5.2B) resulting in isomerisation. To avoid this side reaction, the Horner-Wadsworth-Emmons reaction could be carried out with a non-nucleophilic counterion in the lithium salt, $LiBF_4$ for instance. Unfortunately, due to time constraints and limited starting material, this was not pursued. The shorter chain

analogues **158** and **159** were similarly synthesised directly from *N*-Boc-2-aminoacetaldehyde (**157**) (Scheme 5.2C).



Scheme 5.2. Synthesis of Boc-protected amines. *Reagents and conditions:* a) (i) Dess-Martin periodinane, CH_2Cl_2, 0 °C to rt, 3 h, 50%; b) Trimethylphosphonoacetate, lithium chloride, DBU, MeCN, 0 °C to rt, 1 h, 41%; c) Diethyl (methylsulfonylmethyl)phosphonate, lithium chloride, DBU, MeCN, 0 °C to rt, 2 h; d) Trimethylphosphonoacetate, lithium chloride, DBU, MeCN, 0 °C to rt, 2 h, 50%; e) Diethyl (methylsulfonylmethyl)phosphonate, sodium hydride (60% in mineral oil), THF, 0 °C to rt, 4 h, 26%.

The Boc-protected amines **155** and **158** were deprotected using TFA, then carried forward in reductive amination reactions with the benzimidazolone aldehyde **56** (Scheme 5.3) to afford **160** and **147**, respectively. For the Boc deprotection of amines **155**, **158** and **159** (Scheme 5.3A), TFA was used instead of 4 M HCl in dioxane because it was found that the HCl salt was insoluble, resulting in no reaction in the reductive amination. Reductive amination of vinyl sulfone **159**, on the other hand, yielded no product (Scheme 5.3B).



Scheme 5.3. Synthesis of methyl acrylate and vinyl sulfone. *Reagents and conditions:* a) TFA, triisopropylsilane, CH_2Cl_2, rt, 0.5 h; b) (i) Aldehyde **56**, Et_3N, 1,2-dichloroethane, 0 °C to rt, 5 h, (ii) $NaBH(OAc)_3$, rt, o/n.

5.4 Synthesis of Mutant Protein TRIM33β-E984C

5.4.1 DNA Cloning

To synthesise mutant protein TRIM33β-E984C, a DNA gene fragment between amino acids 888 and 1100 was purchased with the desired mutation at residue 984, and was amplified by PCR (Figure 5.3). The DNA fragment was excised and extracted using a GeneJET Gel Extraction Kit. The DNA was then assembled into a pNIC28 plasmid using a HiFi DNA assembly kit and transformed into *Escherichia coli* NEB5α competent cells. The plasmid was isolated from a small-scale culture of a single colony using a GeneJET Plasmid MiniPrep Kit then submitted for sequencing, which confirmed that the gene fragment had been correctly inserted into the plasmid.

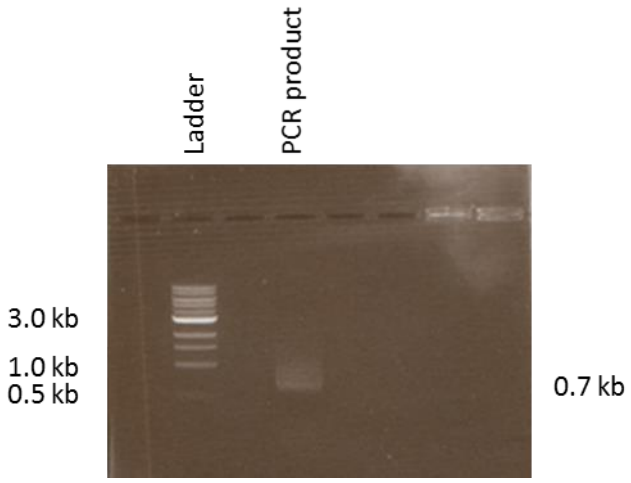


Figure 5.3. DNA gel showing desired TRIM33β-E984C DNA gene fragment.

5.4.2 Protein Expression and Purification

The TRIM33β-E984C plasmid was transformed into *Escherichia coli* BL21 (DE3) cells and a single colony was used to inoculate a starter culture, which was then used for large-scale recombinant protein expression. On induction by IPTG, the media was supplemented with zinc chloride which is required for the proper formation of the PHD finger. Protein purification using IMAC followed by SEC on an ÄKTAFPLC system yielded TRIM33β-E984C. The purity of mutant protein TRIM33β-E984C was confirmed by SDS-PAGE and protein mass spectrometry (Figure 5.4).

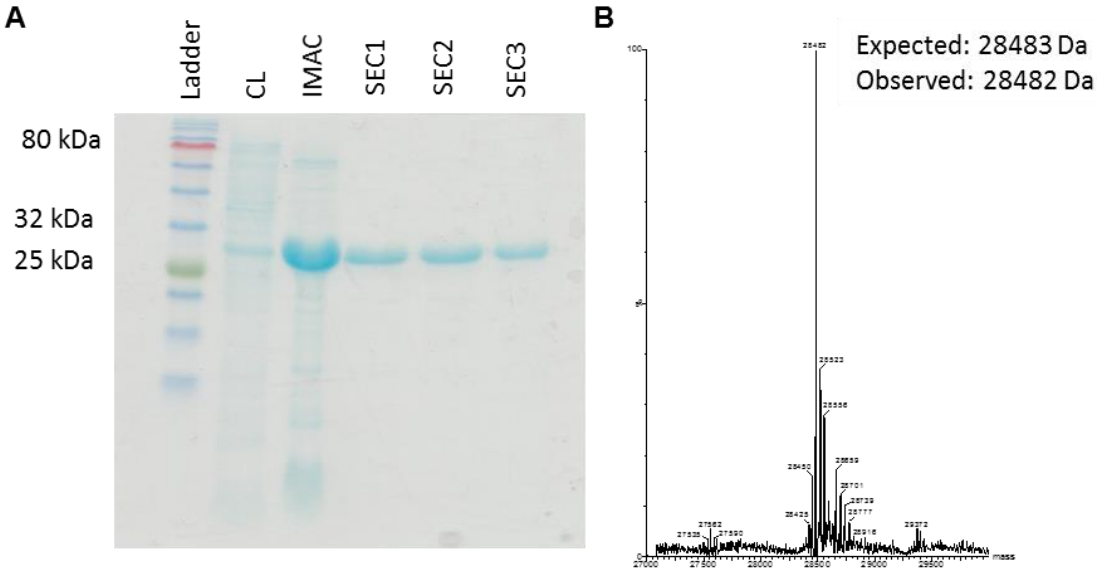


Figure 5.4. Synthesis of mutant protein TRIM33β-E984C. A) SDS-PAGE gel showing purity of TRIM33β-E984C and its size relative to the protein marker (ladder), where CL = cell lysate; B) Protein mass spectrum of TRIM33β-E984C.

5.4.3 Verification of Protein Structure

5.4.3.1 *Circular Dichroism*

To confirm that the mutant protein was folded in a similar manner to the wild-type protein, a circular dichroism experiment was carried out. In this experiment, the difference in absorption of left- and right-handed circularly polarised light is recorded from 185 to 240 nm. The spectra obtained for the wild-type protein (WT, blue) and mutant protein (EC, pink) can be seen in (Figure 5.5). The spectrum obtained for the mutant protein is very similar to that of the wild-type protein, and displays the characteristic double minimum of

α-helices. This is expected of the bromodomain, which is a left-handed bundle of four α-helices. This suggested that the mutant protein was correctly folded and could be used in further experiments.

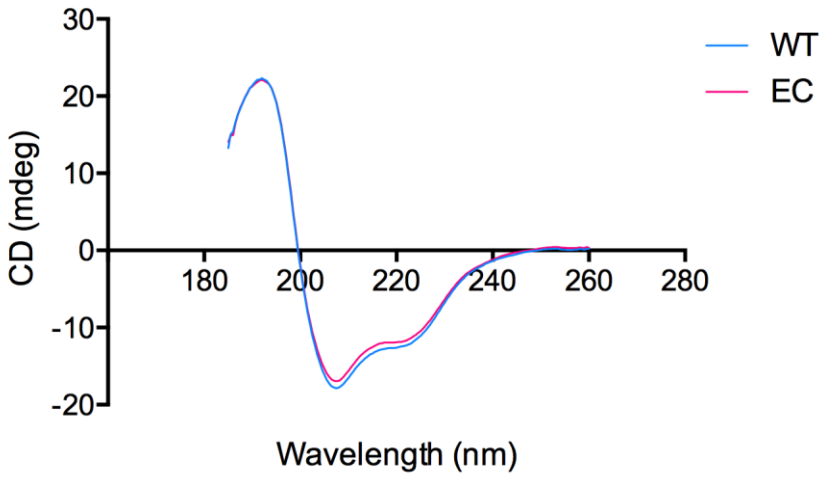


Figure 5.5. Circular dichroism spectra of wild-type TRIM33β (WT, blue) and mutant protein TRIM33β-E984C (EC, pink).

5.4.3.2 *Testing of Protein Binding Ability to Control Compound 62*

Before testing TRIM33β-E984C with the alkylating compounds, it was important to ensure that the mutant protein bound to control compound **62** in a similar manner to wild-type TRIM33β. A signal response AlphaScreen™ was first carried out with TRIM33β-E984C and histone peptide H3K18Ac (Figure 5.6A). The protein concentration selected was 250 nM and the peptide concentration selected was 8 nM. These assay conditions were then used to evaluate compound **62** in 1:2 serial dilutions from 500 μM. Compound **62** bound to TRIM33β-E984C with an IC_{50} of 21.4 μM (Figure 5.6B, data for other peptide combinations can be found in Appendices B), which is similar to the IC_{50} value for TRIM33β of 13.7 μM. The mutation of glutamic acid E984 to cysteine could result in the slight reduction in binding affinity due to the lack of hydrogen bond-forming ability of the cysteine residue with the terminal amine of compound **62**.

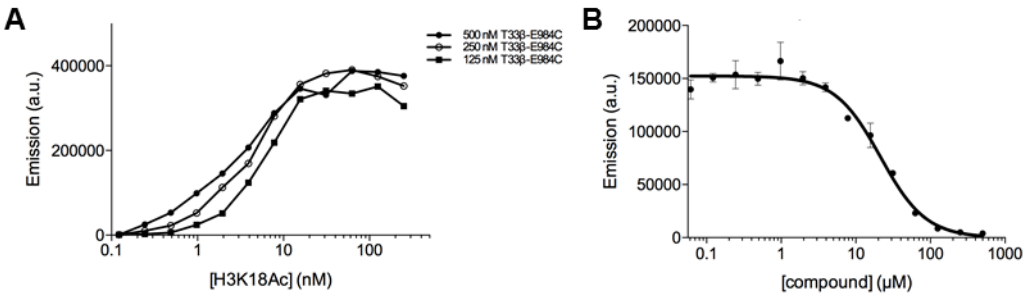


Figure 5.6. AlphaScreen™ with TRIM33β-E984C. A) Signal response experiment with histone peptide H3K18Ac to determine protein-peptide concentration for testing compound **62**; B) Dose-response curve obtained with compound **62**.

5.5 Investigation of Surface-Exposed Cysteine Residues

As the TRIM33β PHD-bromodomain construct has twelve cysteine residues, it was important to identify surface-exposed cysteine residues other than E984C that might react also with the alkylating compounds. The presence of these other cysteine residues would result in difficulties distinguishing whether the alkylating compounds had reacted with the target cysteine mutation. To test for other surface-exposed cysteine residues, 50 μM TRIM33β was incubated at 37 °C for 1 hour with 250 μM *N*-ethylmaleimide (NEM), a common reagent for labelling cysteine residues in proteins.[120] It was observed that within 1 hour, most of the wild-type protein had been doubly-alkylated, with little unalkylated protein and no mono-alkylated protein detected (Figure 5.7). Given that the mass of the TRIM33β construct is 28509, the expected mass in the case of monoalkylation is 28634 and that of dialkylation is 28759. The mass spectrum at t = 1 h has a clear peak of 28758, confirming that the wild-type TRIM33β has two surface exposed cysteine residues that are readily alkylated.

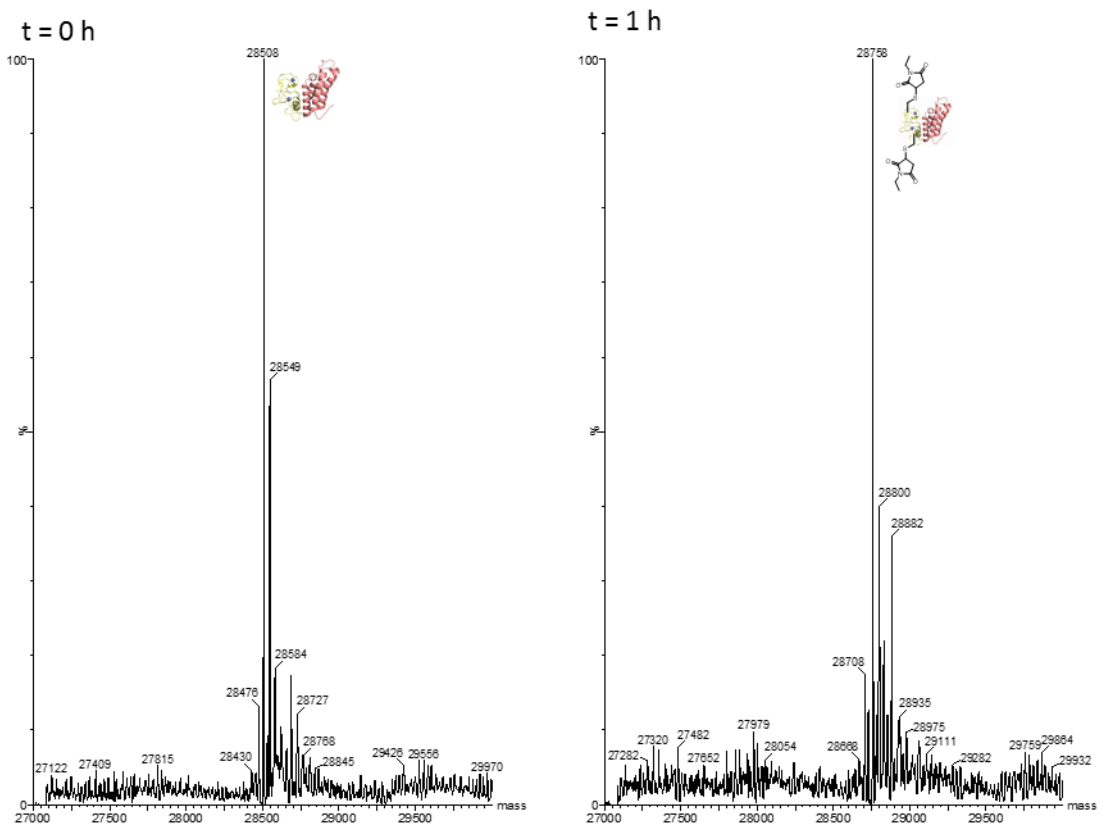


Figure 5.7. Protein mass spectrum of TRIM33β incubated with NEM at t = 0 h and t = 1 h. Molecular mass of TRIM33β = 28509; NEM = 125; TRIM33β + (1 × NEM) = 28634; TRIM33β + (2 × NEM) = 28759.

To determine which cysteine residues were alkylated, the experiment was repeated and after 1 hour, the protein was buffer exchanged into a Tris digestion buffer suitable for the protease chymotrypsin. This process also removed excess NEM to prevent further alkylation on reduction of the protein with dithiothreitol (DTT) and digestion with chymotrypsin. The resulting peptide mixture was prepared for mass spectrometry using Pierce C18 spin columns. The sample was submitted to nano LC-MS/MS followed by proteomics analysis. Proteomics analysis showed that four cysteine residues were alkylated (Figure 5.8, detailed results can be found in Appendices D). This is in contrast with the intact protein mass spectrometry, which suggested that only two cysteine residues had been alkylated by NEM. It is possible that the buffer exchange did not completely remove NEM and a small amount of residual NEM was still present. This was then able to react with

cysteine residues that had been reduced by DTT and exposed on digestion of TRIM33β by chymotrypsin.

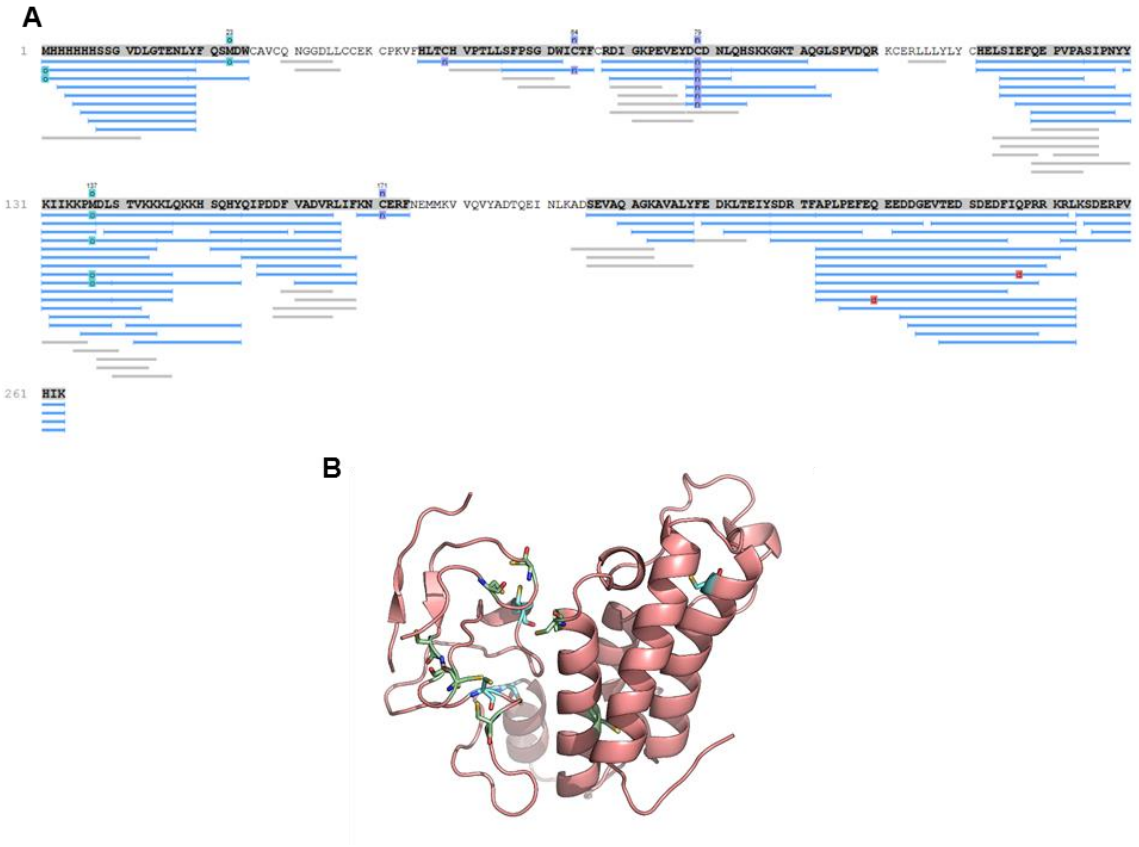


Figure 5.8. Alkylation of surface exposed cysteine residues with *N*-ethylmaleimide. A) Protein sequence indicating peptide sequences with modifications detected (blue), where n = *N*-ethylmaleimide, o = oxidation, d = deamidation; B) X-Ray crystal structure of TRIM33β with cysteine residues highlighted to indicate those alkylated by NEM (blue) and those not alkylated (green).

The experiment was repeated and TRIM33β was incubated with NEM. The protein was buffer exchanged to a low concentration HEPES buffer and submitted to intact protein mass spectrometry using a matrix assisted laser desorption/ionisation (MALDI) technique. This fragments the protein upon ionisation and limits the exposure of reduced cysteine residues, enabling identification of labelled cysteine residues in the native state. As MALDI detects peptides that have been fragmented from the N- and C-termini, there is a limitation in the coverage achieved as the residues that can be identified is dependent on the length of peptide fragmented. In this experiment, coverage was low and only peptides up to 15 amino acids long from the N- and C-termini were detected. This result suggests that MALDI

analysis is not suitable for this system and the protein should be digested as before. Due to time constraints, the procedure could not be optimised. Key to the improvement of this procedure would be ensuring complete removal of the alkylating agent. This could be achieved using dialysis with regular replacement of dialysis buffer instead of buffer exchange columns.

5.6 Investigation of Compound Binding by AlphaScreen™

To measure the binding affinity of the alkylating compounds (Figure 5.9) to TRIM33β, a dose-response AlphaScreen™ experiment was carried out (Table 5.1). This would provide an estimation of binding affinity to TRIM33β relative to the control compound **62**.

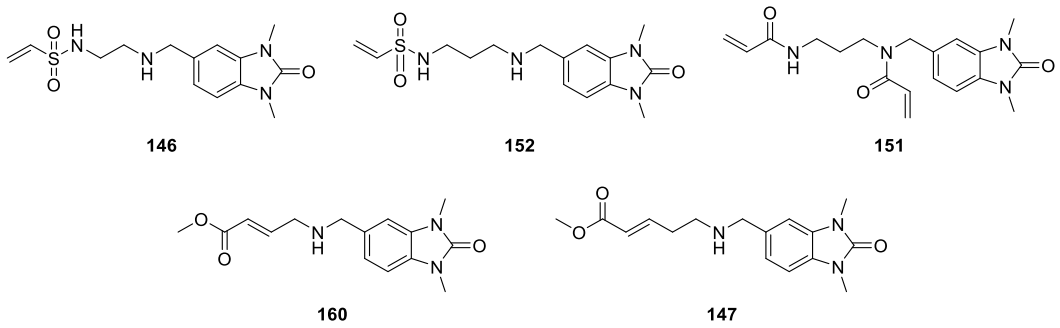
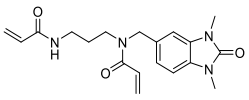
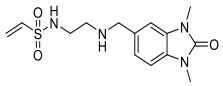
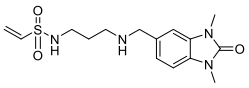
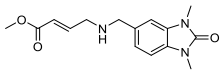
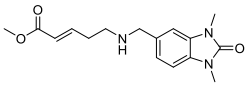


Figure 5.9. Compounds tested in dose-response AlphaScreen™.

Table 5.1. AlphaScreen™ IC_{50} values and NMR waterLOGSY data. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β; H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green). Ligand efficiency is indicated in parentheses.

Compound	AlphaScreen™ IC_{50} (μM)		
	T24-H3(BRD)	T33α-H3(BRD)	T33β-H3(BRD)
 **151**	> 250	> 250	> 250
 **146**	> 250	> 250	> 250
 **152**	> 250	> 250	> 250
 **160**	> 250	18.7 ± 3.16 (0.32)	85.2 ± 20.1 (0.27)
 **147**	> 250	3.81 ± 0.68 (0.34)	12.6 ± 1.88 (0.31)

AlphaScreen™ revealed that compounds **146**, **151** and **152** had significantly weaker binding to TRIM33 than compounds **160** and **147**. Full IC_{50} curves could not be obtained for compounds **146** and **152**, while **151** showed almost no binding to both TRIM33α and TRIM33β. The lack of TRIM33β binding observed for compound **151** is unsurprising given that the benzylic amine is now a tertiary amine. In contrast, compound **147** had IC_{50} values of 3.81 μM for TRIM33α and 12.6 μM for TRIM33β, which are comparable to that of control compound **62** (IC_{50} (TRIM33α) = 3.61 μM, IC_{50} (TRIM33β) = 13.7 μM). Compound **160** still showed binding to TRIM33α and TRIM33β albeit with a lower affinity (IC_{50} (TRIM33α) = 18.7 μM, IC_{50} (TRIM33β) = 85.2 μM). These results indicated that some weaker

binding was observed. Consequently, it is likely that the alkylation of TRIM33β-E984C observed, or lack thereof, would be due to the alkylating ability of the vinyl sulfonamide or methyl acrylate, or the position of the electrophile relative to the cysteine residue rather than due to lack of binding of the ligand.

5.7 Alkylation Experiments with TRIM33β-E984C

In this experiment, 50 μM TRIM33β or TRIM33β-E984C was incubated at 37 °C with 250 μM alkylating compound (**146**, **151**, **152**, **160** or **147**).[120] A small sample of each was removed at pre-defined timepoints (t = 0, 0.5, 1, 3, and 24 h), diluted with buffer and submitted to intact protein mass spectrometry. Any alkylation would be indicated by a peak corresponding to the mass of protein plus compound. This is illustrated in Figure 5.10, where methyl acrylate **147** was incubated with TRIM33β-E984C and mass spectra were obtained at t = 0.5, 1, 3, and 24 h. It is evident that by t = 0.5 h, some monoalkylation was detected, and by t = 3 h, no more unalkylated protein was detected. At t = 24 h, the main peak observed was that corresponding to the dialkylated protein.

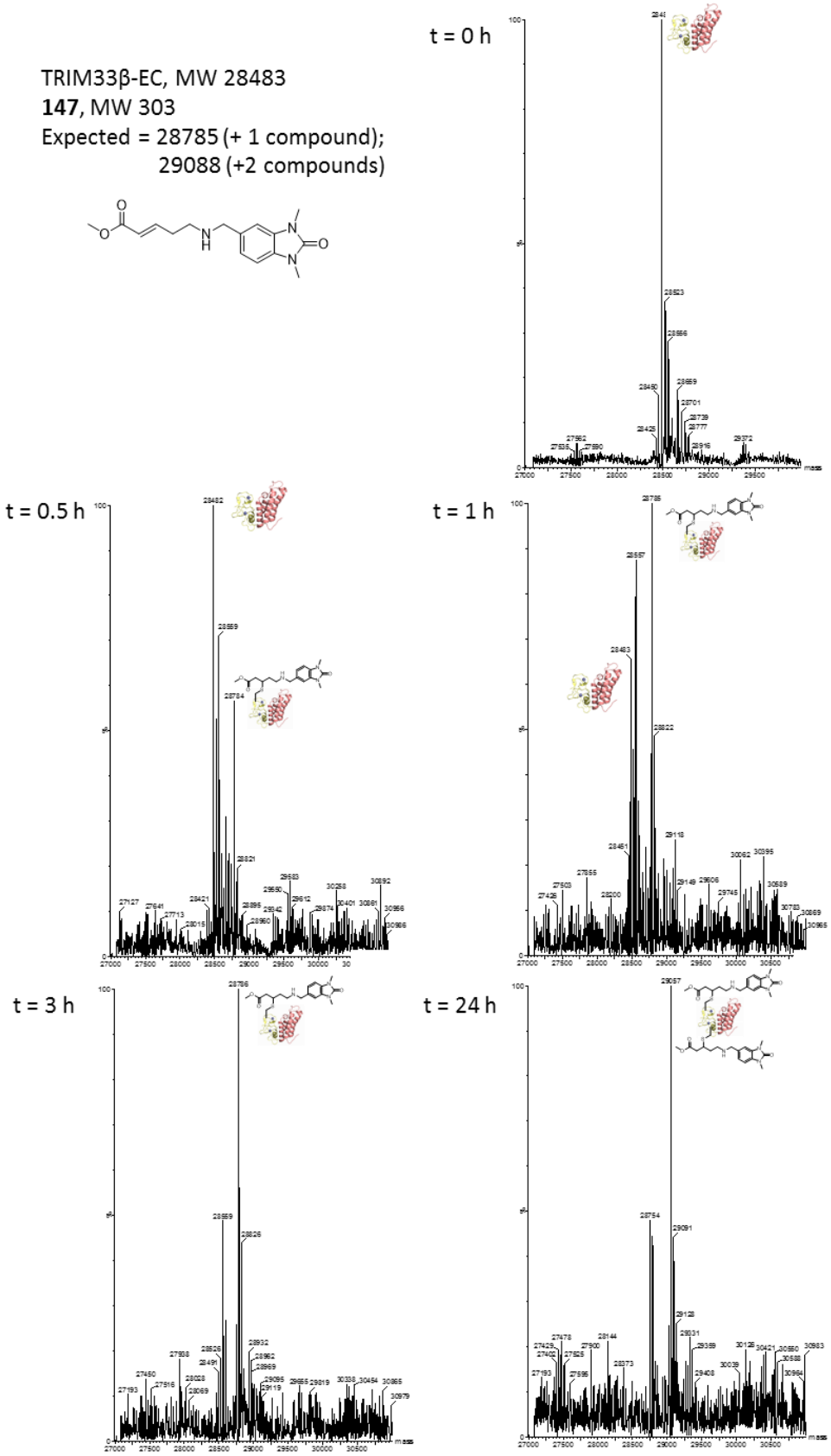


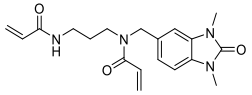
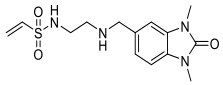
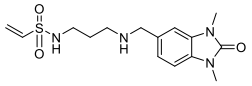
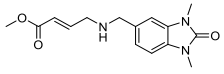
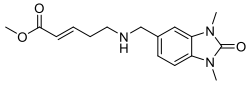
Figure 5.10. Example of mass spectra at predefined time points of alkylation experiment of TRIM33β-E984C using compound 147.

The results for all compounds are summarised in Table 5.2. As expected from the lack of binding observed for compound **151**, no reaction was observed for the wild-type TRIM33β, and only a small amount of alkylation was observed for the mutant TRIM33β-E984C after 1 hour, with the amount of monoalkylation increasing over 24 hours. Vinyl sulfonamides **146** and **152** both reacted once with TRIM33β-E984C up to 3 hours, with no reaction with wild-type TRIM33β, suggesting that the vinyl sulfonamide had selectively reacted with the cysteine mutation. After incubation for 24 hours, a peak corresponding to dialkylated protein was detected for both compounds. The observation of the dialkylated protein peak can be accounted for by comparison to that of wild-type TRIM33β, where some mono- and dialkylation was observed after 24 hours, indicating the reaction of other surface exposed cysteine residues. Notably, compound **152** (with the longer carbon chain) was more reactive than compound **146**, with little or no unalkylated protein detected at t = 0.5 h. Furthermore, it was noted that for every compound reacted with the protein, there was an additional mass of 97. This was similarly observed in attempts to obtain mass spectrometry characterisation data of the pure compound. While NMR confirmed the identity of the compound, high-resolution mass spectrometry was unable to detect the mass of compounds **146** and **152**. Instead, the mass recorded was also that of compound+97. Consequently, this additional mass of 97 is attributed to a mass spectrometry artefact of the vinyl sulfonamide compounds.

In the mass spectrum for incubation with compounds **160** and **147**, at t = 24 h, the peak corresponded to the mass of protein+(2 compound)–30. This is likely to be due to hydrolysis of the methyl ester. As the protein is dialkylated, the loss of two methyl groups (one on each compound), corresponds to a mass of 30. Methyl acylates **160** and **147** were noticeably more reactive than vinyl sulfonamides **146** and **152**, with full dialkylation of TRIM33β-E984C after 24 hours, and quicker monoalkylation of TRIM33β, with monoalkylated TRIM33β detected at t = 0.5 h. This result is in contrast to the relative reaction rates reported by Kathman *et al*.[118] Given the difference in binding affinity as

measured by AlphaScreen™, the difference in rate of alkylation observed could be due to the higher affinity of methyl acrylates **160** and **147** as compared to vinyl sulfonamides **146** and **152**. Similar to the vinyl sulfonamides, when comparing alkylation of wild-type TRIM33β to mutant TRIM33β-E984C at the same time points, it is suggested that methyl acrylates **160** and **147** selectively react with the mutated cysteine residue C984.

Table 5.2. Summary of results from alkylation experiments with TRIM33β and TRIM33β-E984C. Peaks detected in mass spectra are indicated; * – mass of unalkylated protein, # – mass of protein + 1 compound, × – mass of protein + 2 compound.

Compound	TRIM33β				TRIM33β-E984C			
	t = 0.5 h	t = 1 h	t = 3 h	t = 24 h	t = 0.5 h	t = 1 h	t = 3 h	t = 24 h
 **151**	No reaction *	No reaction *	No reaction *	No reaction *	No reaction *	Some mono-alkylation *#	More mono-alkylation *#	Mostly mono-alkylation *#
 **146**	No reaction *	No reaction *	No reaction *	Some mono- and dialkylation *#×	Some mono-alkylation *#	Some mono-alkylation *#	Mostly mono-alkylation *#	Mono- and dialkylation #×
 **152**	No reaction *	No reaction *	No reaction *	Some mono- and dialkylation *#×	Mono-alkylation #	Mono-alkylation #	Mono-alkylation #	Mono- and dialkylation #×
 **160**	No reaction *	Some mono-alkylation *#	More mono-alkylation *#	Mono-alkylation #	Some mono-alkylation *#	Mostly mono-alkylation *#	Mono- and dialkylation #×	Dialkylation ×
 **147**	No reaction *	Some mono-alkylation *#	More mono-alkylation *#	Mostly mono-alkylation *#	Some mono-alkylation *#	Mostly mono-alkylation *#	Mono-alkylation #	Dialkylation ×

5.7.1 Confirmation of Compound Alkylation Site

5.7.1.1 *Verification Using NMR WaterLOGSY Competition Experiments*

If the compound had reacted with the mutated cysteine residue (C984) and is situated in the acetyl-lysine binding site, it should prevent binding of control compound **62**. To verify this, waterLOGSY experiments were conducted to see whether alkylation of the protein would result in the peaks reverting from positive (when binding) to negative (when not binding). AlphaScreen™ was not chosen as it is a peptide displacement assay and binding of the peptide would be affected by the presence of the alkylating compound. As a result, AlphaScreen™ would not be a suitable technique for the competition experiments.

In this experiment, mutant protein TRIM33β-E984C was first incubated with the alkylating compounds for at 37 °C for 2 hours, then buffer exchanged into PBS buffer to remove excess compound. Based on the alkylation experiments, the proteins should have been mostly monoalkylated by this time. A longer incubation time was avoided as when dialkylation has taken place, it would be difficult to confirm that any absence of binding is due to alkylation at the desired site. WaterLOGSY spectra were first obtained for control compound **62** with mutant protein TRIM33β-E984C (Figure 5.11) to confirm binding. The compound peaks observed were positive and hence, are in agreement with the AlphaScreen™ data which showed that compound **62** binds to TRIM33β-E984C. 500 μM of compound **62** was then added to TRIM33β-E984C that had been pre-alkylated, each with compounds **146**, **152**, **160** and **147**. An example of **146** is shown in Figure 5.11. Positive peaks were observed for compound **62**, indicating that binding of compound **62** to the protein was still present. However, it is possible that full alkylation had not taken place and a small percentage of TRIM33β-E984C was left unalkylated, allowing some compound **62** to bind. This is because waterLOGSY is based on transfer of magnetisation from bulk water to the ligand, which is detected after the ligand dissociates from the protein.[121] Consequently, a small fraction of protein that allows the ligand to bind is sufficient to result in a positive signal in the waterLOGSY spectrum.

In an alternative set up, a lower concentration (125 μM) of control compound **62** was incubated with TRIM33β-E984C and the NMR signal was found to be strong enough for detection. 500 μM of alkylating compound was then added. It was thought that the high concentration of alkylating compound might allow it to outcompete control compound **62**, resulting in less or no binding to mutant protein TRIM33β-E984C. As expected, negative peaks were observed for the alkylating compounds (**146** in Figure 5.11, highlighted in green). When compound **146** binds to TRIM33β-E984C, it reacts with the protein and therefore, no positive peaks will be detected. Given the large excess of compound **146**, only unbound molecules of **146** are detected, resulting in the negative peaks. Unfortunately, as indicated by AlphaScreen™, the alkylating compounds have a much lower binding affinity for TRIM33β, limiting its ability to displace control compound **62**. As a result, it was observed that compound **62** was still binding, indicating that waterLOGSY is not a suitable technique to probe the alkylation site of TRIM33β-E984C.

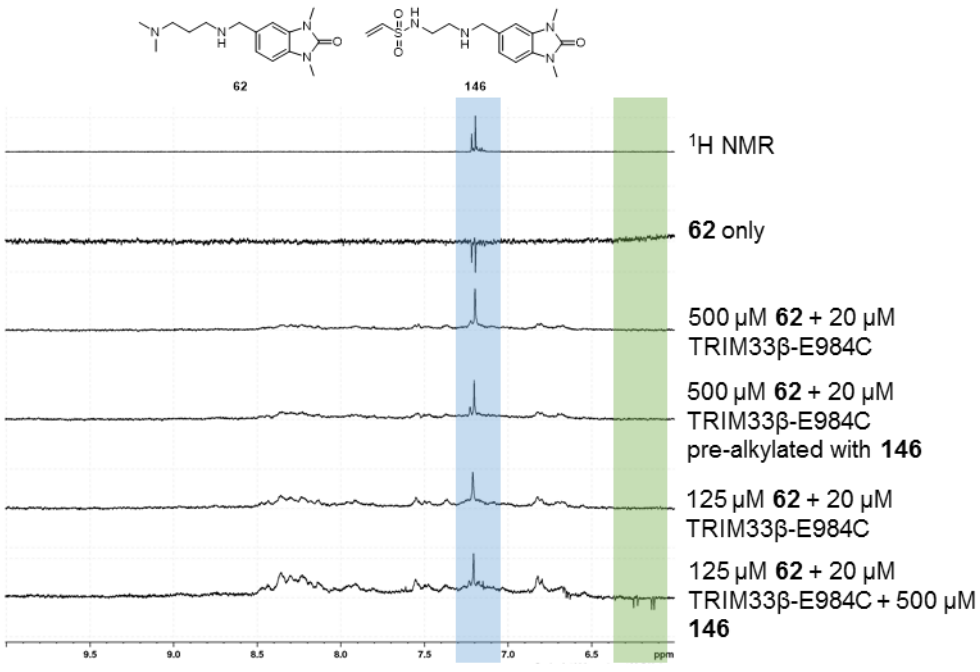


Figure 5.11. WaterLOGSY spectra of competition experiments using control compound 62, mutant protein TRIM33β-E984C and alkylating compound 146. Sample prepared in PBS buffer, with compound (concentration as stated) and 20 μM protein. Peaks corresponding to 62 and 146 have been highlighted in blue and green, respectively.

5.7.1.2 *Verification by Proteomics Analysis*

Intact mass spectrometry with wild-type TRIM33β and mutant TRIM33β-E984C suggested that selective alkylation of cysteine C984 was occurring. To confirm this, the alkylation experiment with proteomics analysis was repeated with compounds **160** and **147**. Vinyl sulfonamides **146** and **152** were not used due to the presence of a mass spectrometry artefact that would complicate proteomics analysis. This time, the compounds were incubated for 2 hours as the time course experiment suggested that at t = 1 h, about half the protein was still unalkylated, but at t = 3 h, some dialkylation is observed. After 2 hours, the protein was buffer exchanged into a Tris buffer suitable for chymotrypsin digestion to remove excess compound. The alkylated protein was then reduced with DTT and digested with chymotrypsin. The resulting peptide mixture was prepared for mass spectrometry using Pierce C18 spin columns and the resulting sample was then submitted to nano LC-MS/MS followed by proteomics analysis.

The expected masses of **160** (+289) and **147** (+303) were not observed. This is not surprising as in characterising these compounds, it was noted that the benzimidazolone headgroup often fragments, and its mass (175) is detected in LRMS. As a result, the data was reprocessed using the masses for (mass of compound)–175. Additionally, intact mass spectrometry had suggested that the methyl acrylate was easily hydrolysed over the course of the experiment. Consequently, another possible mass of modification expected was (mass of compound)–175–CH_2. The peptide sequences and the modifications detected are shown in Figure 5.12 and Figure 5.13. Both compounds **160** and **147** indicated alkylation of several native cysteine residues. However, this is in contrast to the intact protein mass spectroscopy data, which suggested only monoalkylation at t = 2 h. It is possible that removal of the covalent ligand during buffer exchange was not sufficiently thorough and an alternative method such as dialysis should be employed.

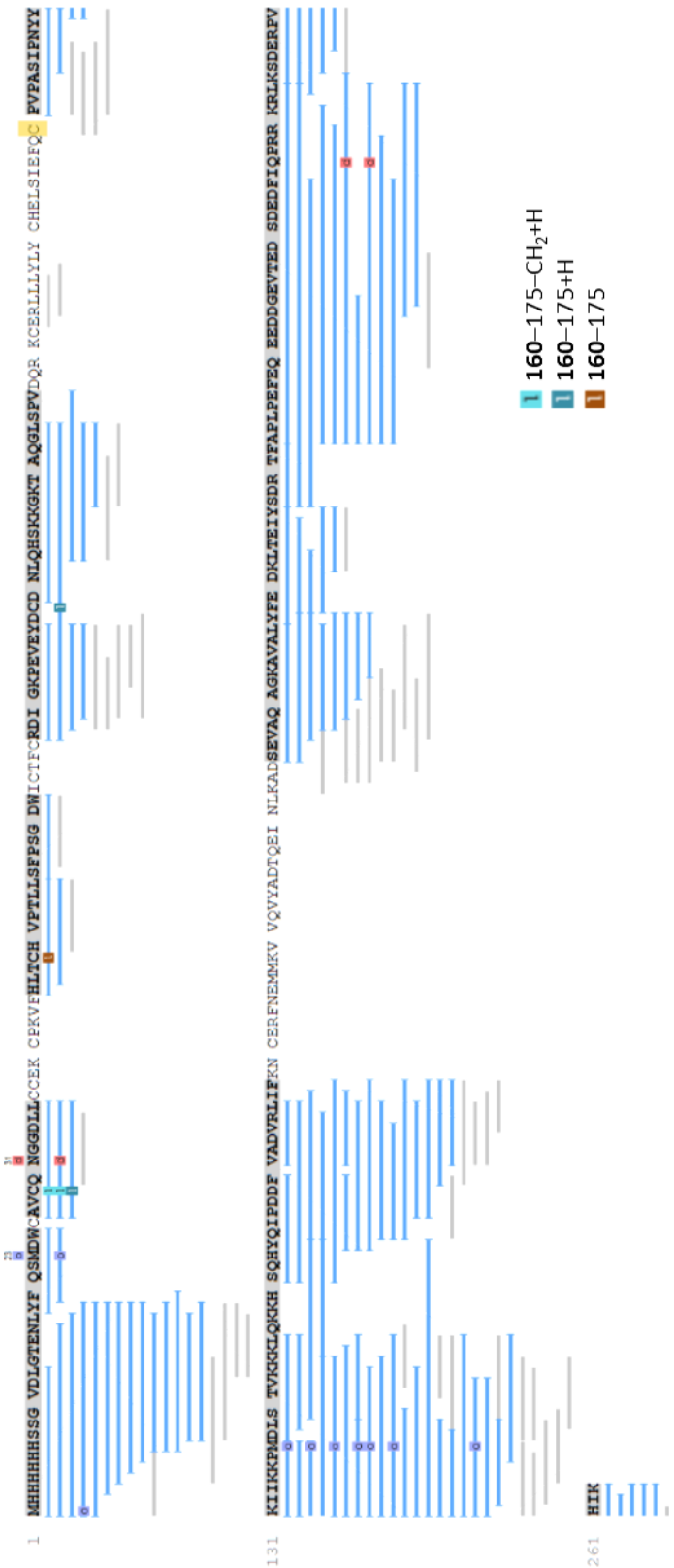


Figure 5.12. Alkylation of TRIM33β-E984C with covalent ligand 160. Protein sequence indicating peptide sequences with modifications detected, where n = ligand, o = oxidation, d = deamidation. E984C mutation is highlighted in yellow.

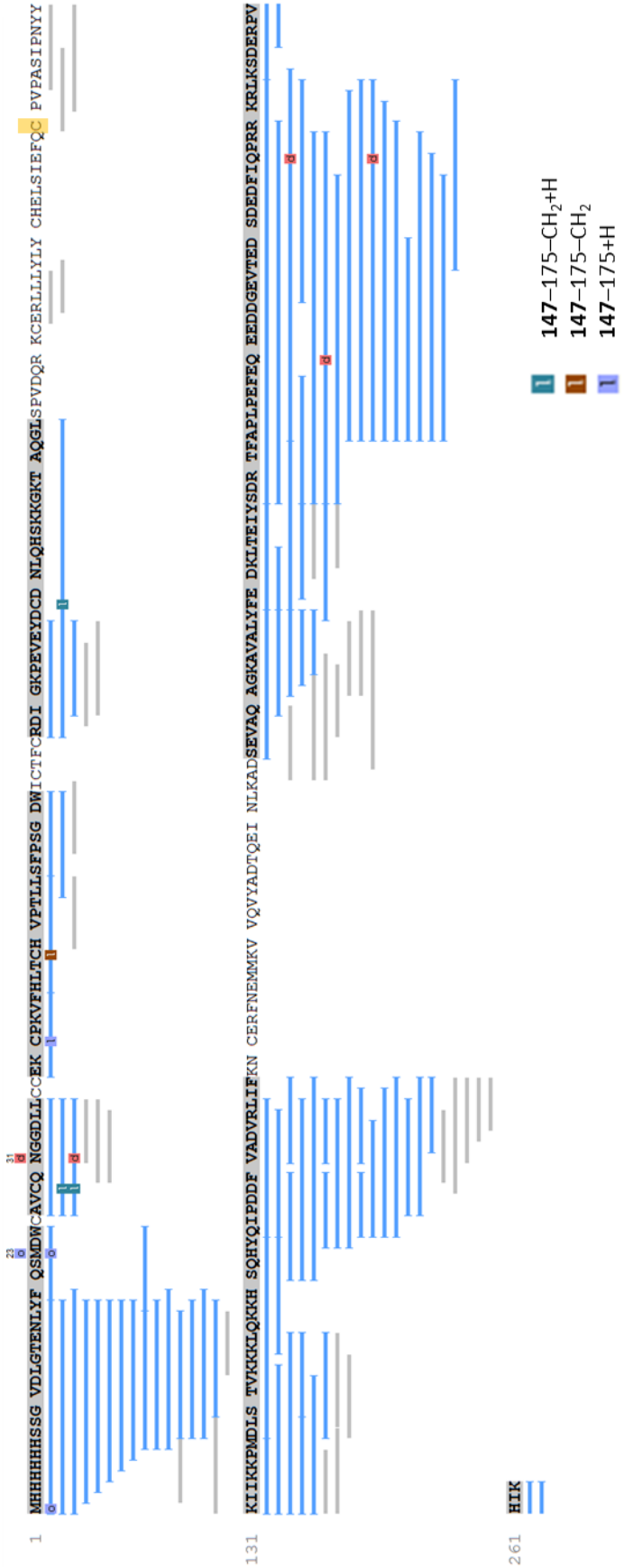


Figure 5.13. Alkylation of TRIM33β-E984C with covalent ligand 147. Protein sequence indicating peptide sequences with modifications detected, where n = ligand, o = oxidation, d = deamidation. E984C mutation is highlighted in yellow.

Analysis of the region containing the E984C mutation (highlighted in yellow) showed that the peptide corresponding to this region of the protein sequence was not detected. This is because the mass of the peptide is too large to be detected. This was not expected as chymotrypsin cleaves peptides after the residues phenylalanine, tyrosine, tryptophan and leucine, suggesting that that region should have been cleaved into two. Earlier proteomics experiments using wild-type TRIM33β with *N*-ethylmaleimide (Section 5.5) confirmed this and the peptide containing E984 was detected. This could be due to the E984C mutation, or the alkylated E984C mutation, affecting the way it fits into the chymotrypsin binding site. Inefficient cleavage by chymotrypsin could also be improved by optimising the protein:protease ratio, or the incubation conditions. Alternatively, the use of a different protease could be explored.

Samples of TRIM33β-E984C alkylated with compounds **160** and **147** were submitted for MALDI analysis to investigate the potential of detecting modifications using a top-down approach. However, similar to the TRIM33β experiment with *N*-ethylmaleimide, the peptide fragments detected were close to the N- and C-termini and hence, did not provide any information of the modification states of the cysteine residues of interest.

5.8 Conclusions

This chapter set out to investigate the potential of enabling crystallographic studies through the use of a covalent analogue of known binder **62**. The covalent ligand can react with mutant TRIM33β-E984C, where a glutamic acid residue in the solvent exposed region of the binding site is mutated to cysteine. Vinyl sulfonamide analogues **146** and **152**, as well as methyl acrylate analogues **160** and **147** were synthesised. Mutant TRIM33β-E984C was produced using bacterial protein expression, and its conformation and binding ability to **62** were verified through circular dichroism and AlphaScreen™. AlphaScreen™ indicated that the alkylating compounds **146**, **152** and **160** bound more weakly to TRIM33β than **62**, while **147** had comparable binding affinity. Despite this, all four compounds alkylated

TRIM33β-E984C but not TRIM33β when incubated for a short period of less than 2 hours. This suggests that alkylation had taken place selectively on the cysteine residue introduced as an E984C mutation. However, as alkylation experiments suggested that wild-type TRIM33β had two surface-exposed cysteine residues that were readily alkylated by *N*-ethylmaleimide, it was desirable to further confirm the selective alkylation.

Consequently, attempts were made to investigate the site of alkylation and binding using NMR waterLOGSY. Unfortunately, due to the difficulty of ensuring that all proteins were only monoalkylated, it was deemed unsuitable as a technique for determining the site of alkylation. Proteomics was attempted as a means of confirming the site of alkylation, however further optimisation of the sample preparation procedure is needed to ensure complete removal of excess ligand and efficient cleavage of the alkylated protein.

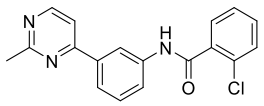
This approach is promising as a means of obtaining detailed information about the binding mode of the compound in spite of the relatively weak binding of compound **62** (K_D = 7 μM) and unsuccessful crystallography trials. Knowledge of other exposed cysteine residues that are prone to alkylation could direct the development of a construct where selective alkylation of E984C is achieved. Obtaining selectively-alkylated TRIM33β-E984C would allow crystallographic studies to be carried out, enabling structure-guided design of a more potent ligand.

Chapter 6

Development of 3-(2-Methylpyrimidin-4-yl)phenyl Series

6.1 Introduction and Aims

Following an initial AlphaScreen™ by Dr Angelina Measures seeking to identify TRIM33 bromodomain or PHD finger ligands,[89] compound **28** (Figure 6.1) was identified as a hit and was validated using NMR waterLOGSY experiments. Attempts to quantify the affinity using AlphaScreen™, however, suggested that this was a very weakly binding compound (IC_{50} (TRIM33α) = 113 μM). It was not clear where on the bromodomain or PHD finger this compound might be binding. However, a ligand that binds in a location other than the bromodomain would be useful as a tool for understanding contextual histone readout and the biological outcomes associated with the PHD finger and bromodomain both individually and together. In light of this, the development of a ligand with greater binding would enable further investigation and consequently, this chapter is a follow-up study of this compound. This study aims to develop a ligand of improved binding and investigate its structure-activity relationships (SAR).



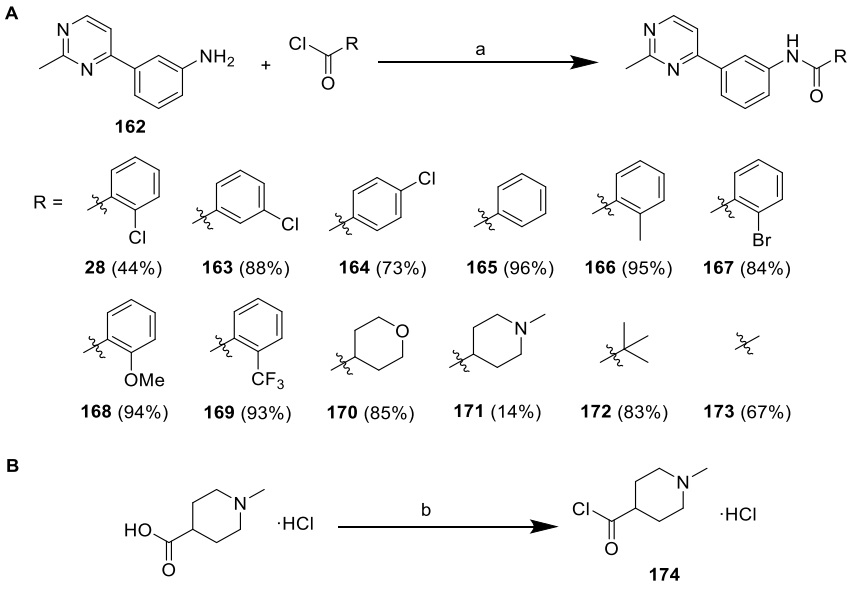
28

Figure 6.1. Hit compound 28.

6.2 Initial Investigation of the 3-(2-Methylpyrimidin-4-yl)phenyl Series

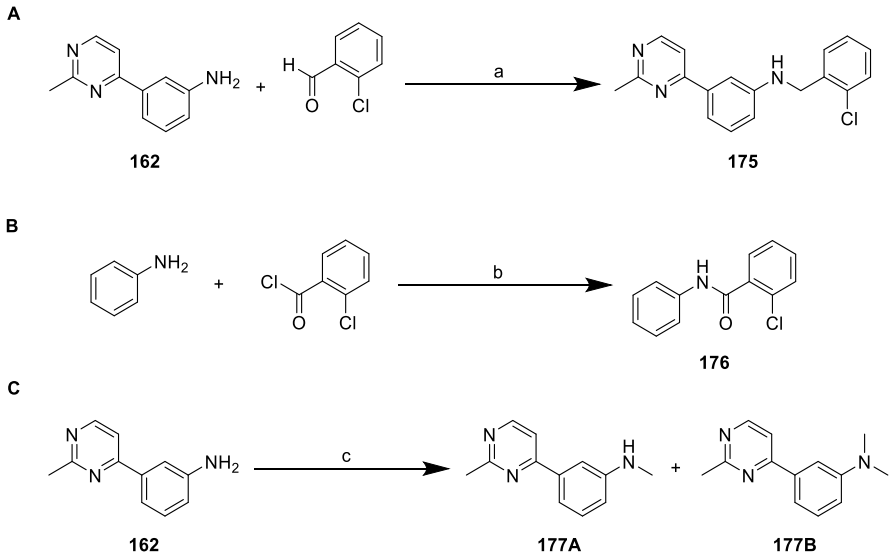
6.2.1 Synthesis of 3-(2-Methylpyrimidin-4-yl)phenyl Series

Hit compound **28** was first synthesised by coupling amine **162** with 2-chlorobenzoyl chloride in a 44% yield. A series of related compounds (**28**, **163**–**173**) were subsequently synthesised in a similar manner using a variety of acid chlorides (Scheme 6.1A). Compound **171** was synthesised by first converting the carboxylic acid to the acid chloride salt **174**, which was used without purification (Scheme 6.1B).



Scheme 6.1. Synthesis of 3-(2-methylpyrimidin-4-yl)phenyl series using a variety of acid chlorides. *Reagents and conditions:* a) Et_3N, *N*,*N*-dimethylacetamide, rt, o/n; b) $SOCl_2$, DMF, toluene, 75 °C, 4 h.

To investigate if the amide carbonyl was necessary for protein binding, amine analogue **175** was synthesised *via* reductive amination (Scheme 6.2A). Analogues of **28** were also synthesised to determine the relative importance of the aromatic rings. Compound **176** was synthesised using aniline and following the general procedure as in Scheme 6.1A with a 92% yield (Scheme 6.2B). The amine **162** was methylated to give compounds **177A** and **177B** (Scheme 6.2C).



Scheme 6.2. Synthesis of amine analogue of compound 28. *Reagents and conditions:* a) (i) Toluene, reflux, 4 h, (ii) $NaBH_4$, MeOH, rt, 0.5 h, 74% over two steps; b) Et_3N, *N*,*N*-dimethylacetamide, rt, o/n, 92%; c) K_2CO_3, MeI, DMF, rt, o/n, 56% (**177A**), 22% (**177B**).

6.2.2 Evaluation of Compounds Using AlphaScreen™

The 3-(2-methylpyrimidin-4-yl)phenyl analogues (**28**, **163**–**173**, **175**–**177B**) were evaluated in an AlphaScreen™ assay in triplicate at two different compound concentrations (30 μM and 150 μM). For each compound, the percentage reduction in emission was calculated reference to the DMSO control for the respective protein-peptide combination (Table B14 in the Appendices). With the exception of compound **171**, none of the compounds showed more than 50% reduction as compared to the DMSO control for any protein-peptide combination, even at 150 μM. This includes the hit compound **28**, which only had a 20% decrease in emission for TRIM33α with histone peptide H3K18Ac. While waterLOGSY experiments by Dr Angelina Measures had confirmed that **28** binds to both TRIM24 (IC_{50} out of range of detection) and TRIM33α (IC_{50} = 113 μM),[89] NMR waterLOGSY has a large range of detection and can detect binders of up to K_D = 1 mM, which is beyond the limit of detection of the AlphaScreen™ assay.

The 2-bromo analogue **167** and 2-methoxy analogue **168** showed a small reduction of 23% and 33%, respectively, for TRIM33α in combination with H3K18Ac. In general, exploration of the phenyl substituent did not bring about any significant change in binding affinity,

suggesting that it is not strongly involved in binding to the protein. For compound **171**, on the other hand, a 46% and 20% reduction in emission was observed for TRIM33α with histone peptide H3K18Ac and H3K9Me$_3$K18Ac, respectively. This result could be due to the three-dimensional shape that is in contrast to the flat aromatic compounds **28**, and **163**–**169**. Consequently, compound **171** was progressed for further validation.

6.2.3 Validation of Compound 171

A waterLOGSY experiment was carried out with compound **171** at 500 μM in the presence of 20 μM of protein (Figure 6.2). A negative spectrum in the absence of protein confirmed that there was no formation of soluble aggregates. In the presence of TRIM24, a negative spectrum was observed indicating that **171** does not bind to TRIM24. In the presence of TRIM33α and TRIM33β, the ligand peaks were inverted, showing that **171** binds to both isoforms of TRIM33. Interestingly, as compared to compound **28** which also bound to TRIM24 in initial waterLOGSY experiments, **171** does not bind to TRIM24 suggesting that replacement of the 2-chlorophenyl with a methylpiperidine removes the observed binding to TRIM24. It is difficult, however, to make further conclusions from this due to the binding site of the compound being unknown.

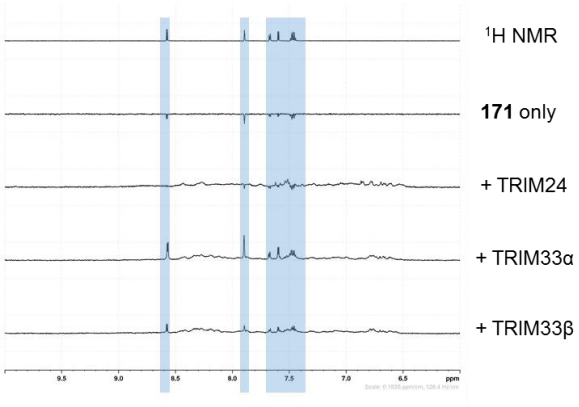


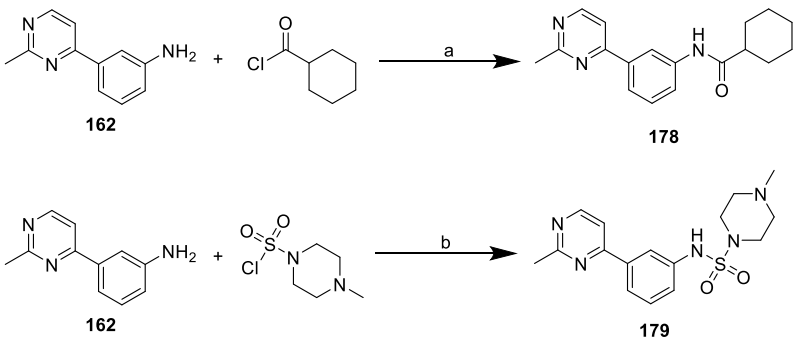
Figure 6.2. **WaterLOGSY spectra of 171.** Ligand peaks are highlighted in blue. Sample prepared in PBS buffer, with 500 μM compound and 20 μM protein. A negative spectrum (in the presence of TRIM24) indicates no binding while a positive spectrum (as with TRIM33α and TRIM33β) indicate binding.

Compound **171** was tested in the AlphaScreen™ TruHits experiment at five different concentrations in 1:2 serial dilutions, with 150 μM as the maximum FAC. This experiment enables the identification of false positives by assessing the compound's ability to interfere with the assay fluorescence. The difference between the emission with compound **171** and that of the DMSO control was then normalised against the standard deviation of the DMSO control and indicated that compound **171** did not interfere with the AlphaScreen™ assay. This confirmed that compound **171** is not a false positive and was used as the basis for design of analogues for further SAR studies.

6.3 Follow-Up of Compound 171

6.3.1 Synthesis of Compounds

As a variation of the methylpiperidine, a cyclohexyl analogue **178** and a sulfonamide analogue **179** were synthesised. Compound **178** was synthesised by reacting amine **162** with cyclohexane carbonyl chloride, in a similar manner to that discussed above, to investigate the role of the methyl piperidine. Compound **179** was synthesised by reacting amine **162** with the sulfonyl chloride following the procedure reported by Guo *et al*. (Scheme 6.1).[122] This compound enables comparison of the sulfamide in place of the amide, which also gives the methylpiperazine a different vector.



Scheme 6.3. Synthesis of 171 analogues. *Reagents and conditions:* a) Et_3N, *N*,*N*-dimethylacetamide, rt, o/n, 91%; b) 3,5-Lutidine, CH_2Cl_2, rt, o/n, 92%.

To investigate the role of other vectors (Figure 6.3), pyrimidine-5-boronic acid (**180**) was reacted with bromoaniline using a Suzuki coupling,[123] affording compounds **181**, **182** and **183** (Scheme 6.4A).

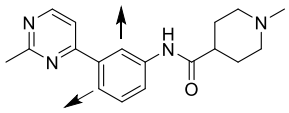
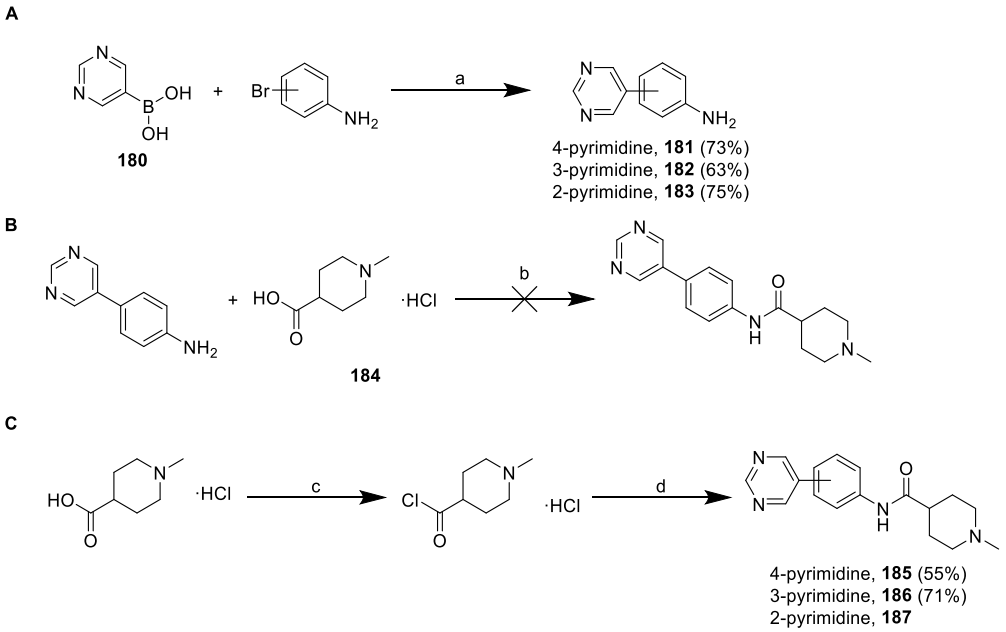


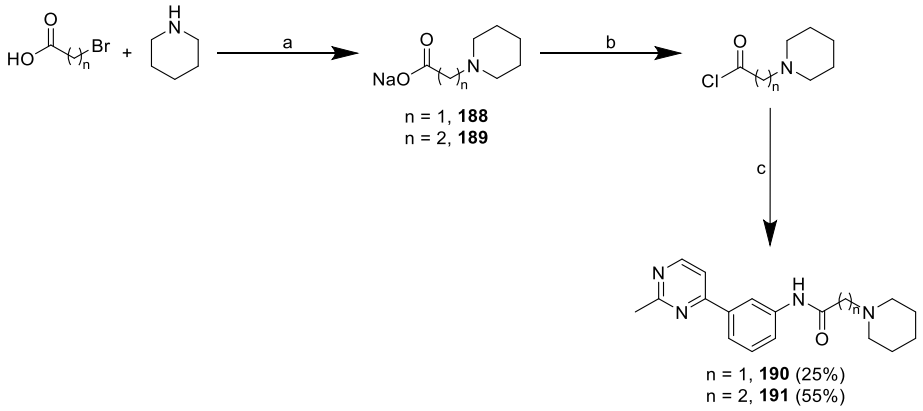
Figure 6.3. Vectors of interest to be investigated as analogues of 171.

Although pyrimidine-4-boronic acid would make for a more similar comparison to **171**, pyrimidine-5-boronic acid was chosen due to its commercial availability. Compound **181** was then reacted with carboxylic acid **184** using CDI as the coupling agent (Scheme 6.4B). This reaction was unsuccessful, possibly due to the aniline nitrogen not being nucleophilic enough. Consequently, carboxylic acid **184** was converted to the corresponding acid chloride using thionyl chloride before reacting with amines **181**, **182** and **183** to afford amides **185**, **186** and **187**, respectively (Scheme 6.4C). Unfortunately, despite multiple attempts, compound **187** could not be sufficiently purified to 95% purity and hence, was not tested in the AlphaScreen™.



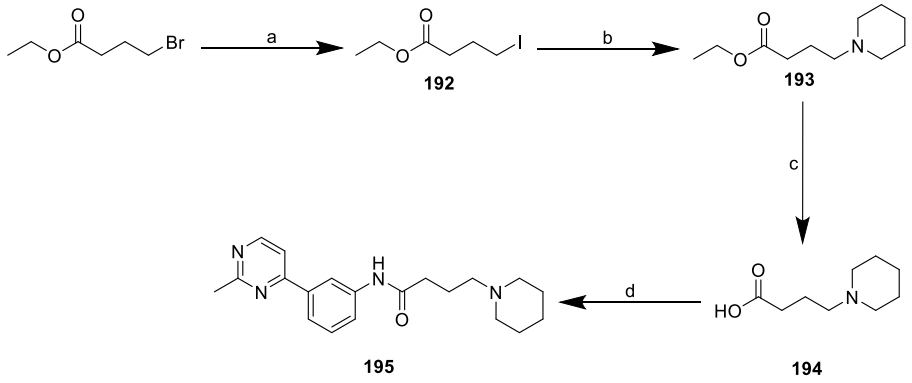
Scheme 6.4. Synthesis of 171 analogues with varying pyrimidine vectors. *Reagents and conditions:* a) $Pd_2(dba)_3$ (2 mol%), PCy_3 (4.8 mol%), K_3PO_4, dioxane/H_2O 2:1, 100 °C, o/n; b) CDI, CH_2Cl_2, rt, o/n; c) $SOCl_2$, toluene, 70 °C, 1 h; d) Et_3N, CH_2Cl_2, rt, o/n.

To investigate the effect of chain length between the aromatic core and the methyl piperidine on TRIM33 binding, a series of analogues of varying chain lengths were synthesised. For ease of synthesis, a piperidine group replaced the *N*-methyl piperidyl group at the end. A bromocarboxylic acid was first reacted with piperidine[124] to yield compounds **188** and **189** as the 1- and 2-carbon analogues, respectively, which were simply lyophilised and reacted with thionyl chloride in the next step. The resulting acid chloride was then reacted with amine **162**, without further purification, to furnish the final compounds **190** and **191**, respectively (Scheme 6.5).



Scheme 6.5. Synthesis of 1- and 2- carbon chain analogues of 171. *Reagents and conditions:* a) NaOH (aq, 3.3 M), H_2O, rt, o/n; b) $SOCl_2$, 40 °C, 1 h; c) Amine **162**, Et_3N, *N*,*N*-dimethylacetamide, rt, o/n.

The 3-carbon chain analogue was not as easily obtained and initial attempts using 4-bromobutyric acid with piperidine did not give the desired product. This could be due to the bromide being less reactive than the 1- and 2-carbon analogues as it is further away from the carboxylic acid. Consequently, the bromide of ethyl-4-bromobutyrate was first converted to iodide **192**, *via* a Finkelstein reaction,[125] before reacting with piperidine[126] to yield compound **193** (Scheme 6.6). The ethyl ester **193** was then hydrolysed to carboxylic acid **194**, which was then used in the same way as for the other analogues, affording **195** in a 46% yield.



Scheme 6.6. Synthesis of 3-carbon chain analogue of 171. *Reagent and conditions:* a) KI, acetone, reflux, 6 h, 90%; b) Piperidine, MeCN, reflux, 4 h, 80%; c) NaOH (aq, 6 M), MeOH, reflux, 2 h; d)(i) $SOCl_2$, toluene, 70 °C, 2 h, (ii) Amine **162**, Et_3N, *N*,*N*-dimethylacetamide, rt, o/n, 46% over two steps.

6.3.2 Evaluation of Compounds by AlphaScreen™

Following synthesis, compounds **178**, **179**, **185**, **186**, **190**, **191**, and **195** were evaluated in an AlphaScreen™ assay against TRIM24, TRIM33α and TRIM33β in triplicate at 30 μM and 150 μM. For each compound, the percentage reduction in emission was calculated with reference to the DMSO control for the respective protein-peptide combination (Table B14 in the Appendices). Compounds **178** and **179** displayed no reduction in emission with any of the protein-peptide combinations. Comparing compounds **178** with **171** suggested that the methylpiperidine contribute to protein binding, while compound **179** indicated that the sulfonamide replacement was not tolerated and the amide was preferred. Compounds **185** and **186** showed no binding to any of the proteins. By comparing compound **186**, which is the most similar to **171**, it can be seen that the replacement of the pyrimid-4-yl with pyrimid-5-yl caused a significant change and prevented binding. This result suggests that the position of the nitrogen atoms of the pyrimidine ring plays a role in the ligand-protein interactions. While the analogues with varying chain lengths (**190**, **191**, **195**) were tolerated, the 2-carbon chain (compound **191**) was the optimum length and showed greater reduction in emission than compound **171** across the TRIM33α protein-peptide combinations.

Compound **191** was subsequently tested in an NMR waterLOGSY experiment with all three proteins. Although all the ligand peaks were positive in the presence of protein (Figure 6.4), only some peaks were inverted on addition of protein (highlighted in blue). It was observed that two peaks (highlighted in green) were not distinctly negative in the absence of protein. This could be due to formation of some soluble aggregates and the very weak spectrum recorded. Consequently, it cannot be reliably concluded whether compound **191** binds to TRIM24, TRIM33α and TRIM33β.

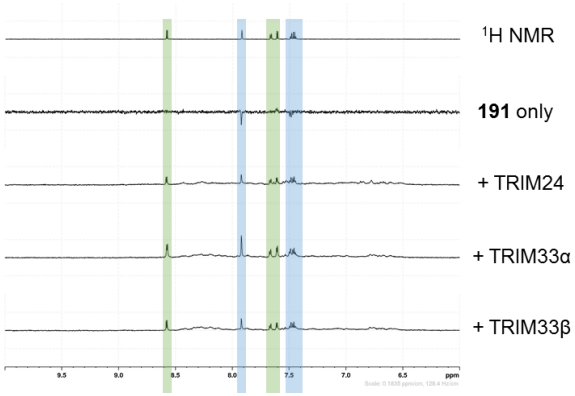


Figure 6.4. WaterLOGSY spectra of 191. Ligand peaks are highlighted in blue. Sample prepared in PBS buffer, with 500 μM compound and 20 μM protein. Peaks that inverted on addition of protein are highlighted in blue while peaks that were not clear in the compound only spectrum are highlighted in green.

6.4 Conclusions

Analogues of compound **28** were synthesised by reacting amine **162** with a variety of acid chlorides in order to better understand the SAR around the compound. Aromatic acid chlorides with different substituents were tested, as well as several non-aromatic ones, to investigate whether they would be tolerated. AlphaScreen™ data suggested that compound **28** itself does not bind strongly to any of the proteins. In spite of this, compound **171** was identified as a binder to TRIM33α and TRIM33β, possibly because the methylpiperidine enables the ligand to fit better into the binding site than the flat aromatic analogues **28**, and **163**–**169**. Analogues of **171** were then synthesised to explore changing the vector of the pyrimidyl ring and the distance of the methylpiperidine from the aromatic core. Given the limited understanding of the binding site, optimising this would have provided a deeper understanding of the shape of the binding site. However, none of the compounds showed significant improvement. Due to the lack of information with respect to the compound binding site, it would be difficult to further optimise this compound and hence, this series was not pursued.

Chapter 7

Conclusions

Based on hits previously identified in an AlphaScreen™ assay, this study aimed to validate these hits for use as a starting point in the development of ligands for the TRIM33. The dual-reader domain found on TRIM33 – the PHD-bromodomain cassette – is of particular interest in light of progress in the understanding of context-dependent histone readout. A ligand for either, or both, domains would therefore be a valuable chemical tool in further deepening our understanding of TRIM33 and its role in epigenetic regulation and disease.

Chapter 2 sought to optimise the AlphaScreen™ assay for TRIM33β and establish an understanding of the assay in relation to the protein-peptide binding affinities for each combination. ITC experiments demonstrated that the bromodomain-binding peptide, H3K18Ac, has significantly weaker binding for TRIM33α and TRIM33β as compared to TRIM24, and that this should be taken into account when using IC_{50} values to study SAR. Moreover, using a known TRIM24 bromodomain ligand, it was shown that by altering the competitor peptide binding partner, some insight into the ligand binding location can be revealed.

In Chapter 3, the binding affinity of hits from the PPI-net collection was quantified. Additionally, a focused library from the Maybridge fragment library was selected and evaluated in an AlphaScreen™ assay. Hit compounds were then progressed for validation *via* an orthogonal NMR waterLOGSY assay, and an AlphaScreen™ TruHits assay was used to test for assay interference. Compound **20** was found to bind to the TRIM33 bromodomain, with selectivity over TRIM24, and was selected for further development.

Chapter 4 explored the SAR in relation to hit compound **20**, and initial studies identified the benzimidazolone headgroup as a ligand efficient starting point with a key hydrogen bonding interaction between the benzylic amine and glutamic acid E981. This interaction was

preserved in the design and synthesis of analogues. Optimisation of the amine chain demonstrated relatively flat SAR, largely due to the relatively open and solvent-exposed pocket. Furthermore, computational studies suggested the importance of structural water molecules in the ZA channel, and understanding their displaceability and position relative to the ligand binding mode is key to the design of a high affinity ligand. Synthesis of headgroup analogues containing an extension on the nitrogen atom enabled investigation of the hydrophobic region above the acetyl-lysine pocket. While this established the optimum vector combination between the extension and amine chain, extensions larger than a methyl or ethyl were not tolerated. This study presented compound **87** as the most promising ligand.

Following unsuccessful attempts to obtain an X-ray crystal structure of TRIM33β containing a ligand, Chapter 5 explored the possibility of exploiting a covalent ligand strategy as a means to achieve this. This would provide new insight into the binding mode of the ligand, particularly with reference to the ZA channel water molecules, and would further guide the development of a TRIM33 bromodomain ligand. Covalent analogues of compound **62**, containing an electrophile suitable for reaction with cysteine, were synthesised and evaluated in alkylation experiments with a TRIM33β-E984C mutant. Intact mass spectrometry studies comparing alkylation with wild-type and mutant TRIM33β indicated selective alkylation in the first few hours. However, attempts to confirm this in proteomics experiments were unsuccessful and require further optimisation for this system.

In Chapter 6, a 3-(2-methylpyrimidin-4-yl)phenyl fragment **28** identified from the Maybridge fragment library was studied. In addition to the low affinity of the ligand, its binding site was not known, resulting in difficulties optimising the ligand. While a ligand binding in a location other than the bromodomain would provide insight into context-dependent histone readout, SAR studies resulted in minimal improvement and consequently, this series was not pursued.

Taken together, SAR studies of hit compound **20** have resulted in a deeper understanding of the TRIM33 bromodomain binding pocket. Detailed knowledge of the binding mode, particularly with consideration of the ZA channel waters, would greatly aid the development of a high affinity TRIM33 bromodomain ligand. It is envisioned that such a ligand could enable study of the cellular functions of TRIM33 and the role that the bromodomain plays in it, further expanding our understanding of the complex mechanisms involved in epigenetics.

Chapter 8

Experimental Section

8.1 General Chemistry Methods

All dry reactions were carried out using standard techniques under an argon atmosphere with flame-dried glassware. Anhydrous solvents were obtained from an MBRAUN Solvent Purification System 5 and stored under an argon atmosphere over 3 Å molecular sieves. Concentration *in vacuo* was performed at 40 °C using a Buchi™ rotary evaporator. Brine refers to a saturated solution of sodium chloride. Rochelle's salt refers to potassium sodium tartrate. Commercial reagents were used as supplied without further purification. Petroleum ether refers to the fraction boiling at 40–60 °C. Celite® refers to Celite® 545 filter aid, treated with sodium carbonate, flux-calcined (Sigma Aldrich).

Analytical thin-layer chromatography (TLC) analysis was carried out on Merck silica gel 60 F_{254} pre-coated aluminium-backed plates and visualised by UV light (λ_{max} = 254 nm) and stained with an aqueous basic potassium permanganate solution, an ethanolic solution of ninhydrin or iodine.

Reaction progress was monitored at appropriate times by TLC analysis.

Flash column chromatography was carried out on Geduran® silica gel (60 Å, 240–400 mesh) under a positive pressure of nitrogen. When used, Biotage column chromatography was run on a Biotage SP1 automated column chromatography system using KP-Sil® SNAP flash silica cartridges.

^{1}H NMR spectra were recorded on a Bruker AVIII HD 400 (400 MHz) spectrometer or a Bruker AVII 500 (500 MHz) spectrometer in $CDCl_3$ (unless otherwise stated). Chemical shifts δ are quoted in parts per million (ppm) relative to tetramethylsilane (δ = 0.00 ppm), to the nearest 0.01 ppm. ^{1}H NMR spectra is reported as follows: chemical shift (number of protons, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet,

br = broad), coupling constant *J* (to the nearest 0.1 Hz), assignment). Residual protic solvent was used as the internal reference: $CDCl_3$ (δ_H = 7.26 ppm), CD_3OD (δ_H = 3.31 ppm), $(CD_3)_2CO$ (δ_H = 2.05 ppm), $(CD_3)_2SO$ (δ_H = 2.50 ppm).[127] Spectra were assigned using COSY, HSQC and HMBC experiments as necessary. Where coupling constants are identical, mean values are quoted.

^{13}C NMR spectra were recorded on a Bruker AVIII HD 400 (101 MHz) spectrometer or a Bruker AVII 500 (126 MHz) spectrometer in $CDCl_3$. Chemical shifts δ are quoted in parts per million (ppm) relative to tetramethylsilane (δ = 0.00 ppm), to the nearest 0.1 ppm, unless peaks are indistinguishable, in which case they are quoted to the nearest 0.01 ppm. The multiplicity of each signal is singlet unless otherwise stated. Residual protic solvent was used as the internal reference: $CDCl_3$ (δ_C = 77.2 ppm), CD_3OD (δ_C = 49.0 ppm), $(CD_3)_2CO$ (δ_C = 29.8 ppm), $(CD_3)_2SO$ (δ_C = 39.5 ppm). NMR data were processed using Mestrenova 10.0 or Bruker Topspin v3.5 software. Spectra were assigned using HSQC and HMBC experiments, as necessary.

^{19}F NMR spectra were recorded on Bruker AVII 500 (470 MHz) spectrometer in $CDCl_3$. Chemical shifts δ are quoted in parts per million (ppm), to the nearest 0.1 ppm. The multiplicity of each signal is singlet unless otherwise stated.

Electrospray ionisation (ESI) mass spectra were obtained using an Agilent 6120 spectrometer for low resolution or Bruker MicroTOF spectrometer for high resolution from solutions of methanol. *m/z* values are reported in Daltons and followed by their percentage abundance in parentheses. For compounds which did not ionise with ESI, spectra were acquired on a Waters GCT classic spectrometer using electron impact (EI) or chemical ionisation (CI).

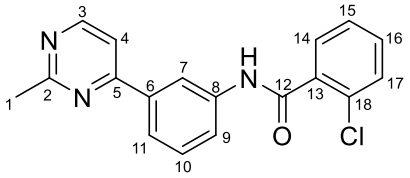
Infrared spectra were recorded from neat samples on a Bruker Tensor 27 spectrometer with a diamond ATR Spectrometer. Absorption maxima ($\bar{\nu}_{max}$) are reported in wavenumbers (cm^{-1}) where s = strong, m = medium, w = weak, br = broad.

Melting points were obtained using a Gallenkamp MF-370 capillary tube melting point apparatus and are uncorrected. The solvent from which each compound was crystallised is noted in parentheses.

Analytical high-performance liquid chromatography (HPLC) was performed on a PerkinElmer Flexar system with a Binary LC Pump and UV/VIS LC Detector set at 254 nm. For determination of purity, a Dionex Acclaim® 120 C18 reverse phase column [5 μm, 12 Å, 150 mm × 4.6 mm] was used with a gradient method (95:5 H_2O:MeCN to 5:95 H_2O:MeCN; 10 min; 5 min hold; 1.5 mL min^{-1}). Samples injected were prepared by dissolving in MeCN and filtered. Compounds run without 0.1% TFA unless otherwise stated.

8.2 Chemistry Experimental Procedures

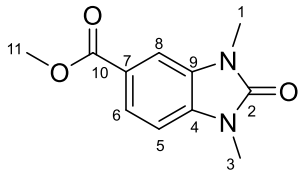
2-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (29)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (185.3 mg, 1.0 mmol, 1.0 eq.) in *N*,*N*-dimethylacetamide (2.5 mL) was added 2-chlorobenzoyl chloride (152 µL, 1.2 mmol, 1.2 eq.) and triethylamine (168 µL, 1.2 mmol, 1.2 eq.). The reaction solution was stirred at room temperature for 19 hours. After this time, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added and the mixture was stirred at room temperature. After 1 hour, water (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic layer was then washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (80% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (142.4 mg, 0.44 mmol, 44% yield): R_f 0.20 (petroleum ether/EtOAc 1:3); m.p. 157–158 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3246 (N-H, w), 3059 (w), 1687 (C=O, m), 1677 (m), 1609 (m), 1572 (s), 1556 (s), 1488 (s), 1435 (s), 1332 (s), 1311 (s), 1264 (w), 1251 (w), 1172 (w), 1085 (w), 1034 (w), 924 (w), 887 (m), 806 (m), 772 (s), 748 (s), 734 (s), 693 (s), 629 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.67 (1H, d, *J* 5.4, C^4-*H*), 8.36 (1H, dd, *J* 1.7, 1.7, C^7-*H*), 8.14 (1H, br s, N-*H*), 7.89–7.84 (2H, m, C^9-*H*, C^{11}-*H*), 7.76 (1H, dd, *J* 7.5, 1.7, C^{17}-*H*), 7.54 (1H, d, *J* 5.4, C^3-*H*), 7.51 (1H, dd, *J* 8.0, 8.0, C^{10}-*H*), 7.48–7.36 (3H, m, C^{14}-*H*, C^{15}-*H*, C^{16}-*H*), 2.79 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.6 (C^2), 164.8 (C^{12}), 163.5 (C^5), 157.7 (C^4), 138.5 (C^8), 138.1 (C^6), 135.1 (C^{18}), 132.0, 130.6, 127.5 (C^{14}, C^{15}, C^{16}), 130.8 (C^{13}), 130.5 (C^{17}), 129.9 (C^{10}), 123.6 (C^9), 122.5 (C^{11}), 118.8 (C^7), 114.3 (C^3), 26.4 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 324.0895, $C_{18}H_{14}{}^{35}ClN_3O$ requires [M+H]$^+$ 324.0898); LRMS *m/z*

(ESI+) 324 ([M^{35}Cl+H]$^+$, 100%), 326 ([M^{37}Cl+H]$^+$, 34%), 669 ([2M^{35}Cl+Na]$^+$, 19%); RP-HPLC (TFA), retention time = 10.5 min, 99.5%.

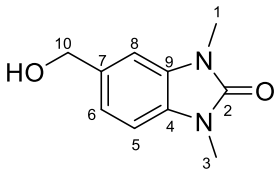
Methyl 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (48)



To a suspension of 3,4-diaminocarboxylic acid (6.09 g, 40.0 mmol, 1.0 eq.) in dry CH_2Cl_2 (50 mL) under argon atmosphere, 1,1'-carbonyldiimidazole (14.27 g, 88.0 mmol, 2.2 eq.) was added. The reaction mixture was stirred at room temperature for 2 hours. After this time, a brown precipitate formed and MeOH (36.0 mL) was added. A pinkish precipitate was formed, and the suspension was stirred for 1 hour. After this time, the reaction mixture was cooled in an ice/water bath, filtered, and the precipitate was transferred to a flask and rinsed with MeOH. The solvent was removed *in vacuo* and the residue was suspended in *N*,*N*-dimethylformamide (40 mL) and cooled to 0 °C. Potassium carbonate (22.13 g, 160.0 mmol, 4.0 eq.) was added followed by iodomethane (10.0 mL, 22.71 g, 160.0 mmol, 4.0 eq.) dropwise while cooling on ice. The reaction mixture was warmed to room temperature and stirred overnight. After 14 hours, a beige precipitate formed. The reaction mixture was diluted with H_2O (200 mL) and EtOAc (200 mL). The aqueous layer was extracted with EtOAc (6 × 30 mL) then extracted with CH_2Cl_2 (6 × 30 mL). The organic components were then washed with 0.5 M aqueous LiCl solution (5 × 30 mL), brine (30 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* to afford the title compound without further purification as a fluffy cream solid (7.08 g, 32.2 mmol, 98% yield): R_f 0.58 (CH_2Cl_2/EtOH 9:1); m.p. 139–141 °C (from EtOAc); $\bar{\nu}_{max}$/cm^{-1} 2949 (w), 1704 (C=O, s), 1625 (w), 1511 (w), 1444 (w), 1302 (m), 1104 (w), 981 (w), 762 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.87 (1H, dd, *J* 8.2, 1.6, C^6-*H*), 7.66 (1H, d, *J* 1.6, C^8-*H*), 6.98 (1H, d, *J* 8.2, C^5-*H*), 3.92 (3H, s, C^{11}-*H*$_3$), 3.45 (3H, s, C^3-*H*$_3$), 3.44 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$)

δ 167.3 (C^{10}), 154.9 (C^{2}), 133.9 (C^{9}), 130.0 (C^{4}), 124.0 (C^{6}), 123.4 (C^{7}), 108.7 (C^{8}), 106.8 (C^{5}), 52.3 (C^{11}), 27.51, 27.46 (C^{1} or C^{3}); HRMS *m/z* (ESI^{+}) (Found $[M+H]^{+}$ 221.0921, $C_{11}H_{12}N_2O_3$ requires $[M+H]^{+}$ 221.0921); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 7.2 min, 99.0%. This compound has been synthesised but not characterised.[102]

5-(Hydroxymethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (49)



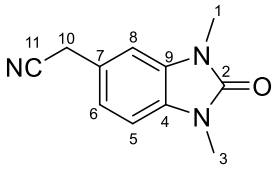
To a solution of methyl 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (**48**) (991.2 mg, 4.5 mmol, 1.0 eq.) suspended in dry toluene (40 mL) under argon atmosphere cooled to –78 °C was added DIBAL-H (1 M in hexanes, 45 mL, 45 mmol, 10.0 eq.) dropwise over 30 minutes. The reaction mixture was stirred at –78 °C for 4 hours. After this time, the reaction was quenched with an aqueous saturated solution of Rochelle's salt (added dropwise initially, in an ice/water bath). The reaction mixture was stirred until the yellow clumps dissolved and two distinct layers were formed. The aqueous layer was then extracted with CH_2Cl_2 (15 × 25 mL). The organic components were then washed with H_2O (20 mL), brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* to afford the title compound as a pale cream solid (842.6 mg, 4.39 mmol, 85% yield): R_f 0.19 (petroleum ether/EtOAc 1:3); m.p. 125–126 °C (from EtOAc/EtOH); $\bar{\nu}_{max}/cm^{-1}$ 3373 (w), 2942 (w), 2858 (w), 1678 (C=O, s), 1508 (m), 1464 (m), 1243 (w), 1191 (m), 1029 (w), 984 (w), 811 (w), 745 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.07 (1H, dd, *J* 7.9, 1.5, C^{6}-*H*), 7.01 (1H, d, *J* 1.5, C^{8}-*H*), 6.90 (1H, d, *J* 7.9, C^{5}-*H*), 4.72 (2H, d, *J* 4.3, C^{10}-H_2), 3.39 (6H, s, C^{1}-H_3, C^{3}-H_3), 2.08 (1H, t, *J* 5.3, O-*H*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.9 (C^{2}), 134.6 (C^{7}), 130.3 (C^{4}), 129.7 (C^{9}), 120.4 (C^{6}), 107.2 (C^{5}), 106.6 (C^{8}), 65.8 (C^{10}), 27.4, 27.3 (C^{1}, C^{3}); HRMS *m/z* (ESI^{+}) (Found $[M+H]^{+}$ 221.09217, $C_{10}H_{12}N_2O_2$ requires $[M+H]^{+}$ 221.09207); LRMS *m/z* (ESI^{-}) 175 ([M–

OH]⁻, 100%); RP-HPLC (TFA), retention time = 5.3 min, 99.9%. This compound has been synthesised but not characterised.[102]

Alternative reduction procedure using lithium borohydride:

To a solution of methyl 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (**48**) (1.32 g, 6.0 mmol, 1.0 eq.) suspended in dry tetrahydrofuran (30 mL) under argon atmosphere was added dropwise lithium borohydride (2 M in tetrahydrofuran, 6.0 mL, 12.0 mmol, 10.0 eq.) followed by methanol (0.75 mL). The reaction mixture was stirred under reflux overnight. After 16 hours, the reaction mixture was cooled to room temperature and quenched with a saturated aqueous solution of $NaHCO_3$ (10mL). The aqueous layer was extracted with CH_2Cl_2 (5 × 30 mL) then extracted with 10% IPA in $CHCl_3$ (5 × 30 mL). The organic components were dried over Na_2SO_4 and filtered. The solvent was removed *in vacuo* to afford the title compound as a pale cream solid (1.10 g, 5.73 mmol, 96% yield). ^{1}H NMR data and mass spectrum is in agreement with the data obtained in the above synthesis.

(1,3-Dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)acetonitrile (50)

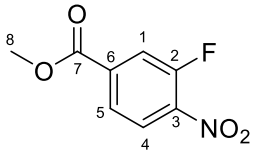


To a solution of 5-(hydroxymethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (**49**) (153.7 mg, 0.8 mmol, 1.0 eq.) in dry CH_2Cl_2 (8.0 mL) under argon atmosphere at 0 °C was added triethylamine (0.42 mL, 303.6 mg, 2.4 mmol, 3.0 eq.) and methanesulfonyl chloride (74 µL, 110.0 mg, 0.96 mmol, 1.2 eq.). After stirring at 0 °C for 1 hour, a saturated aqueous solution of $NaHCO_3$ (10 mL) was added. The aqueous layer was extracted with CH_2Cl_2 (3 × 20 mL), dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo* to afford

5-(chloromethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one, which was used without further purification.

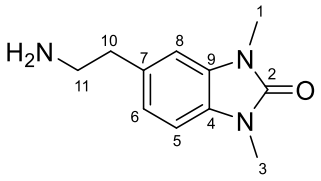
To a solution of 5-(chloromethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one in dry *N*,*N*-dimethylformamide (5.0 mL) under argon atmosphere was added sodium cyanide (42.3 mg, 0.84 mmol, 1.05 eq.). The reaction mixture was stirred at 70 °C for 1 hour. After this time, a saturated aqueous solution of $NaHCO_3$ (10 mL) was added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were then dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (65–80% EtOAc in petroleum ether gradient) to afford the title compound as a colourless solid (64.8 mg, 0.32 mmol, 40% yield over 2 steps): R_f 0.42 (petroleum ether/EtOAc 1:4); m.p. 159–160 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3056 (w), 2917 (w), 2849 (w), 2244 (C≡N, w), 1693 (C=O, s), 1511 (m), 1443 (m), 1393 (m), 1268 (w), 1241 (w), 1141 (w), 1083 (w), 961 (w), 879 (w), 800 (m), 770 (m), 743 (w), 728 (w), 660 (w), 617 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.05 (1H, ddt, *J* 7.9, 1.6, 0.7, C^6-*H*), 6.971 (1H, dd, *J* 1.6, 0.7, C^8-*H*), 6.968 (1H, d, *J* 7.9, C^5-*H*), 3.83 (2H, dd, *J* 0.7, 0.7, C^{10}-*H*$_2$), 3.46 (3H, s, C^3-*H*$_3$), 3.45 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.8 (C^2), 130.7 (C^4), 129.9 (C^9), 123.0 (C^7), 121.0 (C^6), 118.3 (C^{11}), 107.7 (C^5), 107.0 (C^8), 27.4 (C^1, C^3), 23.8 (C^{10}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 202.09759, $C_{11}H_{11}N_3O$ requires $[M+H]^+$ 202.09749); LRMS *m/z* (ESI$^+$) 175 ($[M–N]^+$, 100%); RP-HPLC (TFA), retention time = 6.6 min, 99.2%. This compound has been synthesised but not characterised.[102]

Methyl 3-fluoro-4-nitrobenzoate (51)[128,129]



To a suspension of 3-fluoro-4-nitrobenzoic acid (7.03 g, 38.0 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (45.0 mL) under argon atmosphere was added potassium carbonate (26.3 g, 190.0 mmol, 5.0 eq.). Iodomethane (4.70 mL, 10.8 g, 76.0 mmol, 2.0 eq.) was added dropwise. After stirring at room temperature for 2 hours, H_2O (70 mL) and EtOAc (70 mL) was added. The aqueous layer was extracted with Et_2O (10 × 30 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (5 × 20 mL). The organic components were then dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo* to afford the title compound as a yellow crystalline solid (7.37 g, 37.0 mmol, 97% yield): R_f 0.88 (petroleum ether/EtOAc 1:4); m.p. 40–41 °C (from EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.10 (1H, dd, *J* 7.9, 7.9, C^5-*H*), 7.98–7.90 (2H, m, C^1-*H*, C^4-*H*), 3.98 (3H, s, C^8-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 164.1 (d$_F$, ^4J_F 2.0, C^7), 155.1 (d$_F$, ^1J_F 265.9, C^2), 140.0 (d$_F$, ^2J_F 8.0, C^3), 136.6 (d$_F$, ^3J_F 7.6, C^6), 126.2 (d$_F$, ^4J_F 2.8, C^5), 125.5 (d$_F$, ^3J_F 4.4, C^4), 119.7 (d$_F$, ^2J_F 22.5, C^1), 53.1 (C^8); ^{19}F NMR (470 MHz, $CDCl_3$) δ –116.5; HRMS *m/z* (APCI) (Found $[M]^-$ 199.02820, $C_8H_6FNO_4$ requires $[M]^-$ 199.02863); Compound could not be detected by LRMS. ^{1}H NMR spectrum is in good agreement with literature.[128]

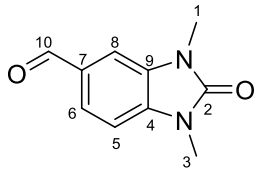
5-(2-Aminoethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (52)



A suspension of (1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)acetonitrile (**50**) (7.0 mg, 0.035 mmol, 1.0 eq.) in dry EtOH (2.0 mL) was degassed under nitrogen for 5 minutes. Palladium on carbon (10% wt, 3.7 mg, 0.0035 mmol, 0.1 eq.) was added, followed by 10% HCl in EtOH (*v/v*) (0.1 mL). The reaction flask was purged with 3 balloons of hydrogen before inserting another balloon of hydrogen. The reaction mixture was stirred at room temperature for 16 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH (100 mL). The solvent was removed in vacuo and the crude

material was purified by flash column chromatography on silica gel (flushed with 100% petroleum ether with 2% triethylamine, followed by 100% EtOAc with 2% triethylamine, then 100% EtOH with 2% triethylamine) to afford the title compound as a colourless oil (3.5 mg, 0.017mmol, 49% yield): R_f 0.25 (EtOH with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3414 (N-H, w), 3368 (N-H, w), 2922 (w), 2852 (w), 1693 (C=O, s), 1509 (m), 1464 (m), 1397 (w), 1318 (w), 1269 (w), 1141 (w), 1084 (w), 1029 (w), 808 (w), 743 (w), 651 (w) ; ^{1}H NMR (500 MHz, MeOD) δ 7.07 (1H, d, *J* 7.9, C^5-*H*), 7.03 (1H, d, *J* 1.6, C^8-*H*), 7.02 (1H, dd, *J* 7.9, 1.6, C^6-*H*), 3.41 (3H, s, C^1-*H$_3$* or C^3-*H$_3$*), 3.40 (3H, s, C^1-*H$_3$* or C^3-*H$_3$*), 2.96–2.91 (2H, m, C^{11}-*H$_2$*), 2.84 (2H, t, *J* 7.4, C^{10}-*H$_2$*); ^{13}C NMR (126 MHz, MeOD) δ 156.4 (C^2), 134.3 (C^7), 131.4, 129.8 (C^4, C^9), 123.3 (C^6), 109.3 (C^8), 108.9 (C^5), 44.2 (C^{11}), 39.4 (C^{10}), 27.44, 27.41 (C^1, C^3); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 206.12888, $C_{11}H_{15}N_3O$ requires [M+H]$^+$ 206.12879); LRMS *m/z* (ESI$^+$) 189 ([M–NH_2]$^+$, 100%); RP-HPLC (TFA), retention time = 4.9 min, 100.0%.

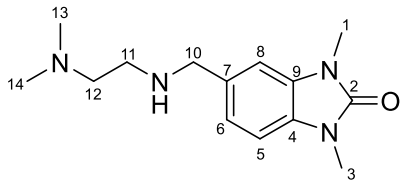
1,3-Dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (56)



To a solution of 5-(hydroxymethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (**49**) (4.58 g, 23.8 mmol, 1.0 eq.) dissolved in CH_2Cl_2 (250 mL) was added manganese(IV) oxide (20.7 g, 238.0 mmol, 10.0 eq.). The reaction mixture was stirred at room temperature for 4 hours. After this time, manganese(IV) oxide (3.11 g, 35.7 mmol, 1.5 eq.) was added and the reaction mixture was stirred at room temperature overnight. After 17 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* to afford the title compound as a peach crystalline solid (4.30 g, 22.6 mmol, 95% yield): R_f 0.33 (petroleum ether/EtOAc 1:3); m.p. 144–145 °C (from EtOH); $\bar{\nu}_{max}$/cm^{-1} 3056 (w), 2801 (w), 2727 (w), 1703 (C=O, s), 1680 (C=O, s), 1618 (m), 1511 (m), 1487 (m), 1400 (w), 1354 (w), 1271 (w), 1193 (m), 809 (w), 726 (w), 670 (w); ^{1}H NMR

(500 MHz, $CDCl_3$) δ 9.95 (1H, s, C^{10}-*H*), 7.65 (1H, dd, *J* 8.0, 1.5, C^6-*H*), 7.54 (1H, d, *J* 1.5, C^8-*H*), 7.08 (1H, d, *J* 8.0, C^5-*H*), 3.48 (6H, s, C^1-*H*$_3$, C^3-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.4 (C^{10}), 154.9 (C^2), 135.5 (C^9), 130.9 (C^4), 130.8 (C^7), 127.0 (C^6), 107.1 (C^5), 106.5 (C^8), 27.63, 27.55 (C^1, C^3); HRMS *m/z* (CI) (Found $[M+H]^+$ 191.0817, $C_{10}H_{10}N_2O_2$ requires $[M+H]^+$ 191.0817); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 6.3 min, 99.9%.

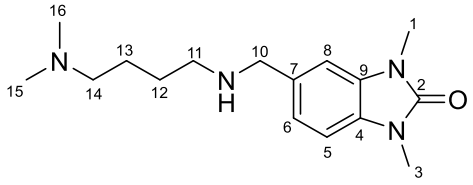
5-({[2-(Dimethylamino)ethyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (57)



To a solution of 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (56.5 mg, 0.3 mmol, 1.5 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and *N*,*N*-dimethylethylenediamine (22 μL, 17.6 mg, 0.2 mmol, 1.0 eq.) and acetic acid (23 μL, 24.0 mg, 0.4 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 3 hours. After this time, sodium triacetoxyborohydride (212.2 mg, 1.0 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature overnight. After 15 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) followed by flash column chromatography on C18-reversed phase silica gel (11% MeCN in H_2O with 0.1% TFA isocratic) to afford the title compound as a colourless oil (24.0 mg, 0.09 mmol, 46% yield): R_f 0.15 (EtOH with 2% triethylamine); $\bar{\nu}_{max}/cm^{-1}$ 3393 (w), 2927 (w), 2851 (w), 1685 (C=O, s), 1509 (m), 1456 (m), 1398 (w), 1269 (w), 1139 (w), 1027 (w), 727 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.03 (1H, dd, *J* 7.9, 1.5, C^6-*H*), 6.99 (1H, d, *J* 1.5,

C^{8}-*H*), 6.88 (1H, d, *J* 7.9, C^{5}-*H*), 3.84 (2H, s, C^{10}-*H*$_2$), 3.403 (3H, s, C^{1}-*H*$_3$), 3.396 (3H, s, C^{3}-*H*$_3$), 2.69 (2H, t, *J* 6.1, C^{11}-*H*$_2$), 2.43 (2H, t, *J* 6.1, C^{12}-*H*$_2$), 2.19 (6H, s, C^{13}-*H*$_2$, C^{14}-*H*$_2$), 1.95 (1H, br s, N-*H*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.0 (C^{2}), 134.0 (C^{7}), 130.3 (C^{9}), 129.2 (C^{4}), 121.3 (C^{6}), 107.5 (C^{8}), 107.1 (C^{5}), 59.2 (C^{12}), 54.5 (C^{10}), 46.8 (C^{11}), 45.7 (C^{13}, C^{14}), 27.3 (C^{1}, C^{3}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 263.18648, $C_{14}H_{22}N_4O$ requires $[M+H]^+$ 263.18664); LRMS *m/z* (ESI$^+$) 175 ($[M–C_4H_{11}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 4.6 min, 99.9%.

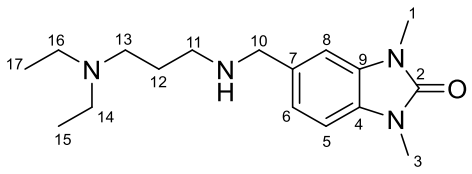
5-({[4-(Dimethylamino)butyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (58)



To a solution of 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (37.6 mg, 0.3 mmol, 1.5 eq.) in dry DCE (5.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 4-(dimethylamino)butylamine (46.5 mg, 0.4 mmol, 2.0 eq.) and triethylamine (56 µL, 40.5 mg, 0.4 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 5 hours. After this time, sodium triacetoxyborohydride (211.7 mg, 1.0 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature overnight. After 15 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified using an Isolute® SCX-2 amine capture-release column (equilibrated with MeOH (10 mL), loaded in MeOH (2 mL), washed with MeOH (10 mL) and eluted with 2 M NH_3 in MeOH (10 mL)) to afford the title compound as a colourless oil (38.2 mg, 0.13 mmol, 66% yield): R_f 0.08 (EtOH with 2% triethylamine);

$\bar{\nu}_{max}$/cm^{-1} 2939 (w), 2823 (w), (1689 (C=O, s), 1508 (m), 1463 (m), 1397 (w), 1267 (w), 1139 (w), 1102 (w), 1038 (w), 806 (w), 743 (w), 727 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.03 (1H, dd, *J* 7.9, 1.5, C^{6}-*H*), 6.98 (1H, d, *J* 1.5, C^{8}-*H*), 6.89 (1H, d, *J* 7.9, C^{5}-*H*), 3.82 (2H, s, C^{10}-*H*$_2$), 3.41 (3H, s, C^{1}-*H*$_3$), 3.40 (3H, s, C^{3}-*H*$_3$), 2.65 (2H, t, *J* 6.8, C^{11}-*H*$_2$), 2.25 (2H, t, *J* 7.1, C^{11}-*H*$_2$), 2.20 (6H, s, C^{15}-*H*$_3$, C^{16}-*H*$_3$), 1.57–1.46 (4H, m, C^{12}-*H*$_2$, C^{13}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.0 (C^{2}), 134.1 (C^{7}), 130.3 (C^{9}), 129.2 (C^{4}), 121.3 (C^{6}), 107.4 (C^{8}), 107.1 (C^{5}), 59.9 (C^{14}), 54.4 (C^{10}), 49.5 (C^{11}), 45.7 (C^{15}, C^{16}), 28.1, 25.8 (C^{12}, C^{13}), 27.3 (C^{1}, C^{3}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 291.21793, $C_{16}H_{26}N_4O$ requires $[M+H]^+$ 291.21794); LRMS *m/z* (ESI$^+$) 175 ($[M-C_6H_{15}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 4.7 min, 98.5%.

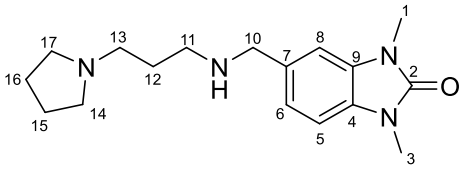
5-({[3-(Diethylamino)propyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (59)



To a solution of 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (56.3 mg, 0.3 mmol, 1.5 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(diethylamino)propylamine (32 µL, 26.0 mg, 0.2 mmol, 1.0 eq.) and acetic acid (23 µL, 24.0 mg, 0.4 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 5 hours. After this time, sodium triacetoxyborohydride (212.8 mg, 1.0 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature overnight. After 15 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by

flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) to afford the title compound as a colourless oil (11.2 mg, 0.04 mmol, 12% yield): R_f 0.19 (EtOH with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3401 (w), 2934 (w), 2818 (w), 1691 (C=O, s), 1508 (m), 1462 (m), 1396 (w), 1269 (w), 1196 (w), 1139 (w), 1084 (w), 807 (w), 744 (w), 727 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.03 (1H, dd, *J* 8.0, 1.5, C^6-*H*), 6.98 (1H, d, *J* 1.5, C^8-*H*), 6.89 (1H, d, *J* 7.9, C^5-*H*), 3.81 (2H, s, C^{10}-*H*$_2$), 3.41 (3H, s, C^3-*H*$_3$), 3.40 (3H, s, C^1-*H*$_3$), 2.67 (2H, t, *J* 6.9, C^{11}-*H*$_2$), 2.50 (4H, q, *J* 7.2, C^{14}-*H*$_2$, C^{16}-*H*$_2$), 2.49–2.45 (2H, m, C^{13}-*H*$_2$), 1.85 (1H, br s, N-*H*), 1.68 (2H, tt, *J* 7.0, 7.0, C^{12}-*H*$_2$), 1.00 (6H, t, *J* 7.2, C^{15}-*H*$_2$, C^{17}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.0 (C^2), 134.1 (C^7), 130.3 (C^4), 129.2 (C^7), 121.3 (C^6), 107.5 (C^8), 107.1 (C^5), 54.5 (C^{10}), 51.4 (C^{13}), 48.5 (C^{11}), 47.0 (C^{14}, C^{16}), 27.4, 27.34, 27.32 (C^1, C^3, C^{12}), 11.8 (C^{15}, C^{17}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 305.23344, $C_{17}H_{28}N_4O$ requires $[M+H]^+$ 305.23359); LRMS *m/z* (ESI$^+$) 175 ($[M-C_6H_{15}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 4.8 min, 96.6%.

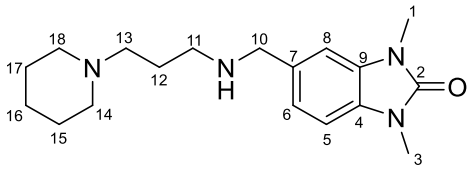
1,3-Dimethyl-5-({[3-(pyrrolidin-1-yl)propyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (60)



To a solution of 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (38.3 mg, 0.2 mmol, 1.0 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 1-(3-aminopropyl)pyrrolidine (57 µL, 51.3 mg, 0.4 mmol, 2.0 eq.) and acetic acid (23 µL, 24.0 mg, 0.4 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 6 hours. After this time, sodium triacetoxyborohydride (211.4 mg, 1.0 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature overnight. After 18 hours, the reaction mixture was filtered through Celite® and washed with EtOH (50 mL) and H_2O (50 mL). The solvent was removed *in vacuo* and

the crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) followed by flash column chromatography on C18-reversed phased silica gel (15% MeCN in H_2O with 0.1% TFA isocratic) to afford the title compound as a colourless oil (38.3 mg, 0.13 mmol, 63% yield): R_f 0.16 (EtOH with 2% triethylamine); $\bar{\nu}_{max}/cm^{-1}$ 3420 (w), 2981 (w), 2802 (w), 1676 (C=O, s), 1510 (w), 1458 (w), 1202 (m), 1133 (m), 838 (w), 801 (w), 723 (w); ^{1}H NMR (500 MHz, DMSO-D6) δ 7.10 (1H, d, *J* 1.4, C^8-*H*), 7.04 (1H, d, *J* 8.0, C^5-*H*), 7.01 (1H, dd, *J* 8.0, 1.4, C^6-*H*), 3.70 (2H, s, C^{10}-*H*$_2$), 3.31 (3H, s, C^1-*H*$_3$ or C^3-*H*$_3$), 3.30 (3H, s, C^1-*H*$_3$ or C^3-*H*$_3$), 2.53–2.50 (2H, m, C^{11}-*H*$_2$), 2.41–2.35 (6H, m, C^{13}-*H*$_2$, C^{14}-*H*$_2$, C^{17}-*H*$_2$), 1.67–1.61 (4H, m, C^{15}-*H*$_2$, C^{16}-*H*$_2$), 1.58 (2H, tt, *J* 8.0, 1.4, C^{12}-*H*$_2$); ^{13}C NMR (126 MHz, DMSO-D6) δ 154.0 (C^2), 134.1 (C^7), 129.6, 128.4 (C^4, C^9), 120.5 (C^6), 107.2 (C^8), 107.0 (C^5), 54.1 (C^{13}), 53.7 (C^{14}, C^{17}), 53.3 (C^{10}), 47.1 (C^{11}), 28.8 (C^{12}), 26.92, 26.86 (C^1, C^3), 23.1 (C^{15}, C^{16}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 303.21784, $C_{17}H_{26}N_4O$ requires $[M+H]^+$ 303.21794); LRMS *m/z* (ESI$^+$) 175 ($[M–C_7H_{15}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 4.7 min, 97.8%.

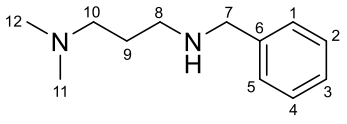
1,3-Dimethyl-5-({[3-(piperidin-1-yl)propyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (61)



To a solution of 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (57.0 mg, 0.3 mmol, 1.5 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and *N*-(3-aminopropyl)piperidine (29 µL, 25.7 mg, 0.2 mmol, 1.0 eq.) and acetic acid (23 µL, 24.0 mg, 0.4 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 3 hours. After this time, sodium triacetoxyborohydride (212.5 mg, 1.0 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room

temperature overnight. After 15 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) followed by flash column chromatography on C18-reversed phase silica gel (20% MeCN in H_2O with 0.1% TFA isocratic) to afford the title compound as a colourless oil (18.3 mg, 0.06 mmol, 29% yield): R_f 0.22 (EtOH with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3397 (w), 2935 (w), 2811 (w), 1691 (C=O, s), 1508 (m), 1456 (m), 1397 (w), 1269 (w), 1126 (w), 954 (w), 809 (w), 744 (w), 728 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.03 (1H, dd, *J* 7.9, 1.5, C^6-*H*), 6.98 (1H, d, *J* 1.5, C^8-*H*), 6.89 (1H, d, *J* 7.9, C^5-*H*), 3.82 (2H, s, C^{10}-*H$_2$*), 3.41 (3H, s, C^1-*H$_3$* or C^3-*H$_3$*), 3.40 (3H, s, C^1-*H* or C^3-*H*), 2.68 (2H, t, *J* 7.0, C^{11}-*H$_2$*), 2.39–2.32 (6H, m, C^{13}-*H$_2$*, C^{14}-*H$_2$*, C^{18}-*H$_2$*), 2.00 (1H, br s, N-*H*), 1.72 (2H, tt, *J* 7.0, 7.0, C^{12}-*H$_2$*), 1.55 (4H, tt, *J* 5.6, 5.6, C^{15}-*H$_2$*, C^{17}-*H$_2$*), 1.45–1.38 (2H, m, C^{16}-*H$_2$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.0 (C^2), 134.0 (C^9), 130.3 (C^4), 129.2 (C^7), 121.2 (C^6), 107.4 (C^8), 107.1 (C^5), 58.0 (C^{13}), 54.8 (C^{14}, C^{18}), 54.4 (C^{10}), 48.5 (C^{11}), 27.33, 27.32 (C^1, C^3), 27.1 (C^{12}), 26.1 (C^{15}, C^{17}), 24.6 (C^{16}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 317.23359, $C_{18}H_{28}N_4O$ requires $[M+H]^+$ 317.23359); LRMS *m/z* (ESI$^+$) 175 ($[M–C_8H_{17}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 4.8 min, 98.6%

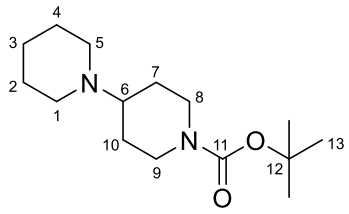
N^3-Benzyl-N^1,N^1-dimethylpropane-1,3-diamine (66)



To a solution of benzaldehyde (71 µL, 74.3 mg, 0.7 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (176 µL, 143.0 mg, 1.4 mmol, 2 eq.) and triethylamine (195 µL, 141.7 mg, 1.4 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 5 hours. After this time, sodium triacetoxyborohydride (741.3 mg, 3.5 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 16 hours. After this time, the reaction mixture was

filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (20 mL) and $CHCl_3$ (20 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (4 × 20 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified using an Isolute® SCX-2 amine capture-release column (equilibrated with MeOH (10 mL), loaded in MeOH (2 mL), washed with MeOH (10 mL) and eluted with 2 M NH_3 in MeOH (10 mL)) to afford the title compound as a colourless oil (25.6 mg, 0.13 mmol, 19% yield): R_f 0.08 (petroleum ether/EtOAc 1:3); $\bar{\nu}_{max}$/cm^{-1} 2942 (w), 2816 (w), 2779 (w), 1455 (m), 1377 (w), 1262 (w), 1156 (w), 1116 (w), 1027 (w), 839 (w), 736 (s), 698 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.34–7.30 (4H, m, C^1-*H*, C^2-*H*, C^4-*H*, C^5-*H*), 7.26–7.22 (1H, m, C^3-*H*), 3.78 (2H, s, C^7-*H*$_2$), 2.67 (2H, t, *J* 7.2, C^8-*H*$_2$), 2.31 (2H, t, *J* 7.3, C^{10}-*H*$_2$), 2.21 (6H, s, C^{11}-*H*$_3$, C^{12}-*H*$_3$), 1.68 (2H, tt, *J* 7.3, 7.2, C^9-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 140.6 (C^6), 128.5, 128.2 (C^1, C^2, C^4, C^5), 127.0 (C^3), 58.2 (C^{10}), 54.2 (C^7), 48.0 (C^8), 45.7 (C^{11}, C^{12}), 28.3 (C^9); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 193.17003, $C_{12}H_{20}N_2$ requires [M+H]$^+$ 193.16993); LRMS *m/z* (ESI$^+$) 193 ([M+H]$^+$, 100%); RP-HPLC (TFA), retention time = 4.4 min, 97.6%.

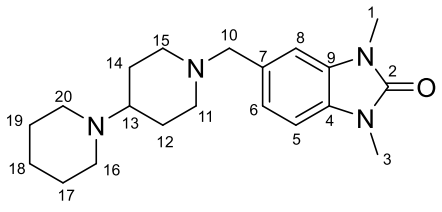
tert-Butyl [1,4'-bipiperidine]-1'-carboxylate (69)



To a solution of 1-Boc-4-piperidone (996.0 mg, 5.0 mmol, 1.0 eq.) in dry 1,2-dichloroethane (30.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and piperidine (1.0 mL, 10.0 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 2 hours. After this time, sodium triacetoxyborohydride (5.30 g, 25.0 mmol, 5.0 eq.) was added portionwise. The reaction

mixture was then stirred at room temperature overnight. After 16 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was suspended in CH_2Cl_2 (20 mL) and quenched with a saturated aqueous solution of $NaHCO_3$ (20 mL). The aqueous layer was extracted with CH_2Cl_2 (4 × 20 mL). The organic components were then washed with H_2O (20 mL), brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (petroleum ether/EtOAc 1:6 to EtOAc to EtOAc/EtOH 1:1), followed by flash column chromatography on silica gel (1–100% EtOH in CH_2Cl_2 gradient) to afford the title compound as a pale yellow oil (1.09 g, 4.1 mmol, 81% yield): R_f 0.33 (CH_2Cl_2/EtOH 9:1); $\bar{\nu}_{max}$/cm^{-1} 2932 (w), 2853 (w), 1694 (C=O, s), 1421 (m), 1365 (m), 1244 (w), 1158 (s), 1105 (m), 1021 (w), 872 (w), 768 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.24–4.00 (2H, m, C^8-*H$_a$*, C^9-*H$_a$*), 2.65 (2H, t, *J* 12.4, 12.4, C^8-*H$_b$*, C^9-*H$_b$*), 2.48 (4H, t, *J* 5.5, C^1-*H$_2$*, C^5-*H$_2$*), 2.36 (1H, tt, *J* 11.6, 3.6, C^6-*H*), 1.76 (2H, ddd, *J* 13.1, 2.8, 2.8, C^7-*H$_a$*, C^{10}-*H$_a$*), 1.56 (4H, tt, *J* 5.6, 5.5, C^2-*H$_2$*, C^4-*H$_2$*), 1.42 (9H, s, C^{13}-*H$_3$*, C^{14}-*H$_3$*, C^{15}-*H$_3$*), 1.45–1.34 (4H, m, C^3-*H$_2$*, C^7-*H$_b$*, C^{10}-*H$_b$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.8 (C^{11}), 79.4 (C^{12}), 62.9 (C^6), 50.3 (C^1, C^5), 43.6 (br, C^8, C^9), 28.5 (C^{13}), 28.0 (br, C^7, C^{10}), 26.4 (C^2, C^4), 24.8 (C^3); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 269.22228, $C_{15}H_{28}N_2O_2$ requires [M+H]$^+$ 269.22235); Compound could not be detected by LRMS; No HPLC data obtained as the compound is not UV active.

5-[([1,4'-Bipiperidin]-1'-yl)methyl]-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (73)



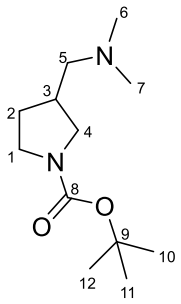
To a solution of *tert*-butyl [1,4'-bipiperidine]-1'-carboxylate (**69**) (108.0 mg, 0.4 mmol, 1.0 eq.) in CH_2Cl_2 (8.0 mL) was added triisopropylsilane (8.2 µL, 0.04 mmol, 0.1 eq.) and TFA

(2.0 mL). The reaction mixture was stirred at room temperature for 30 minutes. After this time, the solvent was removed *in vacuo* and residual TFA was azeotroped with toluene to afford 1,4'-bipiperidine as a TFA salt, which was used in the next step without further purification.

To a solution of 1,4'-bipiperidine TFA salt (0.4 mmol, 2.0 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves and triethylamine (140 µL, 0.5 mmol, 2.5 eq.). The solution was cooled to 0 °C and 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (38.5 mg, 0.2 mmol, 1.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 2 hours. After this time, sodium triacetoxyborohydride (212.0 mg, 1.0 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature overnight. After 18 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was suspended in CH_2Cl_2 (10 mL) and quenched with a saturated aqueous solution of $NaHCO_3$ (10 mL). The aqueous layer was extracted with CH_2Cl_2 (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) to afford the title compound as a pale-yellow oil (45.0 g, 0.13 mmol, 66% yield): R_f 0.10 (EtOH with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 2933 (w), 2853 (w), 2802 (w), 1695 (C=O, s), 1507 (m), 1445 (m), 1396 (w), 1342 (s), 1267 (m), 1137 (w), 1100 (w), 776 (w), 727 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.00 (1H, dd, *J* 7.9, 1.5, C^6-*H*), 6.97 (1H, d, *J* 1.5, C^8-*H*), 6.88 (1H, d, *J* 7.9, C^5-*H*), 3.52 (2H, s, C^{10}-H_2), 3.41 (3H, s, C^3-H_3), 3.40 (3H, s, C^1-H_3), 2.97 (2H, ddd, *J* 11.7, 2.4, C^{11}-H_a, C^{15}-H_a), 2.65 (4H, br s, C^{17}-H_2, C^{19}-H_2), 2.47 (1H, t, *J* 10.9, C^{13}-*H*), 1.98 (2H, dt, *J* 11.7, 2.4, C^{11}-H_b, C^{15}-H_b), 1.87 (2H, d, *J* 11.9, C^{12}-H_a, C^{14}-H_a), 1.76–1.62 (6H, m, C^{12}-H_a, C^{14}-H_a, C^{16}-H_2, C^{20}-H_2), 1.52–1.43 (2H, m, C^{18}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.9 (C^2), 131.7 (C^7), 130.3 (C^4), 129.4 (C^9), 122.3 (C^6), 108.2 (C^8), 106.9 (C^5), 63.3 (C^{13}), 63.2 (C^{10}), 53.2 (C^{11}, C^{15}), 50.2 (C^{12}, C^{14}), 27.3 (C^1, C^3), 25.5 (C^{16}, C^{20}), 24.3 (C^{18}); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$

343.24894, $C_{15}H_{28}N_2O_2$ requires $[M+H]^+$ 343.24924); LRMS *m/z* (ESI+) 175 ($[M-C_{10}H_{19}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 4.8 min, 95.5%.

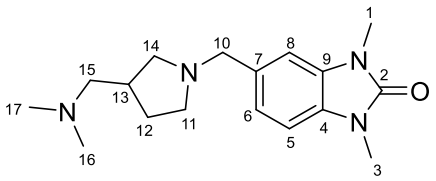
tert-Butyl 3-[(dimethylamino)methyl]pyrrolidine-1-carboxylate (75)



To a solution of *tert*-butyl 3-(aminomethyl)pyrrolidine-1-carboxylate (201.0 mg, 1.0 mmol, 1.0 eq.) in MeOH (9.0 mL) was added formaldehyde (37% wt in H_2O, 0.75 mL, 300.3 mg, 10.0 mmol, 10.0 eq.) and the reaction mixture was stirred at room temperature for 2 hours. After this time, sodium cyanoborohydride (265.0 mg, 4.2 mmol, 4.2 eq.) was added. The reaction mixture was then stirred at room temperature for 22 hours. After this time, the reaction was quenched with a saturated aqueous solution of $NaHCO_3$ (5 mL). The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (60–100% EtOAc in petroleum ether gradient to 0–75% EtOH in EtOAc gradient) to afford the title compound as a colourless amorphous solid (191.0 mg, 0.84 mmol, 84% yield): R_f 0.11 (EtOAc/EtOH 3:1); $\bar{\nu}_{max}/cm^{-1}$ 2974 (w), 2943 (w), 2870 (w), 2819 (w), 2768 (w), 1695 (C=O, s), 1478 (w), 1405 (m), 1366 (w), 1256 (w), 1167 (m), 1122 (m), 1041 (w), 882 (w), 773 (w); ^{1}H NMR (500 MHz, MeOD) δ 3.59–3.51 (2H, m, major and minor rotamer, C^4-H_a), 3.46–3.40 (2H, m, major and minor rotamer, C^1-H_a), 3.30–3.22 (2H, m, major and minor rotamer, C^1-H_b), 3.00–2.91 (2H, m, major and minor rotamer, C^4-H_b), 2.48–2.36 (6H, m, major and minor rotamer, C^3-H, C^5-H_2), 2.29 (6H, s, minor rotamer, C^6-H_3, C^7-H_3), 2.28 (6H, s, major rotamer, C^6-H_3, C^7-H_3), 2.08–1.99 (2H, m, major and minor rotamer, C^2-H_a), 1.66–1.52 (2H, m, major and minor rotamer, C^2-H_b), 1.46 (18H, s, major and minor rotamer, C^{10}-H_3, C^{11}-H_3, C^{12}-H_3); ^{13}C NMR (126 MHz, MeOD) δ 156.40 (major rotamer, C^8),

156.36 (minor rotamer, C^8), 80.8 (major rotamer, C^9), 80.7 (minor rotamer, C^9), 63.7 (major rotamer, C^5), 63.6 (minor rotamer, C^5), 51.7 (minor rotamer, C^4), 51.2 (major rotamer, C^4), 46.6 (minor rotamer, C^1), 46.2 (major rotamer, C^1), 45.8 (major and minor rotamer, C^6, C^7), 38.1 (major rotamer, C^3), 37.4 (minor rotamer, C^3), 31.3 (minor rotamer, C^2), 30.5 (major rotamer, C^2), 28.8 (major and minor rotamer, C^{10}, C^{11}, C^{12}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 229.19106, $C_{12}H_{24}N_2O_2$ requires $[M+H]^+$ 229.19105); LRMS *m/z* (ESI^+) 229 ($[M+H]^+$, 100%); No HPLC data obtained as compound is not UV active.

5-({3-[(Dimethylamino)methyl]pyrrolidin-1-yl}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (76)

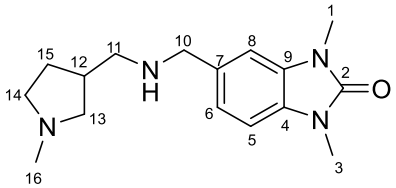


To *tert*-butyl 3-[(dimethylamino)methyl]pyrrolidine-1-carboxylate (**75**) (101.2 mg, 0.44 mmol, 1.0 eq.) in a flask under argon atmosphere was added hydrogen chloride (4 M in dioxane, 3.0 mL). The reaction mixture was stirred at room temperature for 1 hour. After this time, the solvent was removed *in vacuo* to afford *N*,*N*-dimethyl-1-(pyrrolidin-3-yl)methanamine as a HCl salt, which was used in the next step without further purification.

To a solution of *N*,*N*-dimethyl-1-(pyrrolidin-3-yl)methanamine HCl salt (0.44 mmol, 2.0 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves and triethylamine (123 μL, 89.0 mg, 0.88 mmol, 4.0 eq.). The solution was cooled to 0 °C and 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (42.3 mg, 0.22 mmol, 1.0 eq.) was added. The reaction mixture was stirred at room temperature for 2 hours. After this time, sodium triacetoxyborohydride (233.6 mg, 1.10 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature overnight. After 18 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The

solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (3 × 20 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) followed by flash column chromatography on silica gel (35% EtOH in $CHCl_3$ with 2% triethylamine isocratic) to afford the title compound as a colourless oil (55.3 mg, 0.18 mmol, 83% yield): R_f 0.17 (EtOH with 2% triethylamine); $\bar{\nu}_{max}/cm^{-1}$ 2938 (w), 2815 (w), 2774 (w), 1708 (C=O, s), 1508 (m), 1460 (m), 1395 (w), 1336 (w), 1266 (w), 1137 (w), 1081 (w), 1041 (w), 913 (w), 847 (w), 805 (w), 774 (w), 744 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.01 (1H, dd, *J* 7.9, 1.5, C^6-*H*), 6.98 (1H, d, *J* 1.5, C^8-*H*), 6.85 (1H, d, *J* 7.9, C^5-*H*), 3.62 (2H, q, *J* 12.6, C^{10}-H_2), 3.39 (3H, s, C^1-H_3), 3.38 (3H, s, C^3-H_3), 2.81 (1H, dd, *J* 9.3, 7.5, C^{14}-H_a), 2.70–2.63 (1H, m, C^{11}-H_a), 2.48–2.40 (1H, m, C^{11}-H_b), 2.40–2.31 (1H, m, C^{13}-*H*), 2.28–2.11 (3H, m, C^{14}-H_b, C^{15}-H_2), 2.18 (6H, s, C^{16}-H_3, C^{17}-H_3), 2.03–1.93 (1H, m, C^{12}-H_a), 1.49–1.40 (1H, m, C^{12}-H_b); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.9 (C^2), 132.4 (C^7), 130.2 (C^9), 129.2 (C^4), 122.0 (C^6), 108.1 (C^8), 106.9 (C^5), 65.4 (C^{15}), 61.0 (C^{10}), 59.3 (C^{14}), 53.9 (C^{11}), 45.8 (C^{16}, C^{17}), 35.5 (C^{13}), 29.2 (C^{12}), 27.31, 27.26 (C^1, C^3); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 303.2188, $C_{17}H_{26}N_4O$ requires $[M+H]^+$ 303.2185); LRMS *m/z* (ESI$^+$) 303 ($[M+H]^+$, 100%); RP-HPLC (TFA), retention time = 4.6 min, 99.5%.

1,3-Dimethyl-5-({[(1-methylpyrrolidin-3-yl)methyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (77)

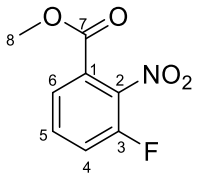


To a solution of 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (37.4 mg, 0.20 mmol, 1.0 eq.) in dry 1,2-dichloroethane (3.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and C-(1-methyl-

pyrrolidine-3-yl)-methylamine (46.0 mg, 0.40 mmol, 2.0 eq.) and triethylamine (56 μL, 40.5 mg, 0.40 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 7 hours. After this time, sodium triacetoxyborohydride (212.0 mg, 1.0 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 17 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. ^{1}H NMR indicated that the product obtained was the imine. Consequently, the crude material was dissolved in dry methanol (4.0 mL) under argon atmosphere and cooled to 0 °C. Sodium borohydride (37.8 mg, 1.0 mmol, 5.0 eq.) was added and the reaction mixture was stirred at room temperature for 2 hours. After this time, the reaction was quenched with a saturated aqueous solution of $NaHCO_3$ (10 mL) and the aqueous components extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified using an Isolute® SCX-2 amine capture-release column (equilibrated with MeOH (10 mL), loaded in MeOH (2 mL), washed with MeOH (10 mL) and eluted with 2 M NH_3 in MeOH (10 mL)) to afford the title compound as a colourless oil (36.7 mg, 0.13 mmol, 64% yield): R_f 0.06 (EtOH with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 2941 (w), 2789 (w), 1688 (C=O, s), 1508 (m), 1451 (m), 1397 (w), 1269 (w), 1242 (w), 1141 (w), 808 (w), 743 (w), 727 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.02 (1H, dd, *J* 7.9, 1.5, C^6-*H*), 6.97 (1H, d, *J* 1.5, C^8-*H*), 6.89 (1H, d, *J* 7.9, C^5-*H*), 3.82 (2H, s, C^{10}-*H*$_2$), 3.41 (3H, s, C^1-*H*$_3$), 3.40 (3H, s, C^3-*H*$_3$), 2.68 (1H, dd, *J* 9.2, 7.8, C^{13}-*H*$_a$), 2.61 (2H, d, *J* 7.3, C^{11}-*H*$_2$), 2.55–2.41 (2H, m, C^{14}-*H*$_2$), 2.41–2.33 (1H, m, C^{12}-*H*), 2.32 (3H, s, C^{16}-*H*$_3$), 2.22 (1H, dd, *J* 9.2, 6.0, C^{13}-*H*$_b$), 2.05–1.95 (1H, m, C^{15}-*H*$_a$), 1.45 (1H, ddt, *J* 12.3, 8.0, 6.1, C^{15}-*H*$_b$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.0 (C^2), 134.1 (C^7), 130.4 (C^9), 129.2 (C^4), 121.3 (C^6), 107.4 (C^8), 107.1 (C^5), 61.1 (C^{13}), 56.2 (C^{14}), 54.7 (C^{11}), 54.4 (C^{10}), 42.4 (C^{16}), 38.5 (C^{12}), 29.9 (C^{15}), 27.4 (C^1,

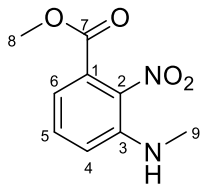
C[3]); HRMS *m/z* (ESI[+]) (Found [M+H][+] 289.20231, $C_{16}H_{24}N_4O$ requires [M+H][+] 289.20229); LRMS *m/z* (ESI[+]) 175 ([M–$C_6H_{13}N_2$][+], 100%); RP-HPLC (TFA), retention time = 4.6 min, 99.4%.

Methyl 3-fluoro-2-nitrobenzoate (91)



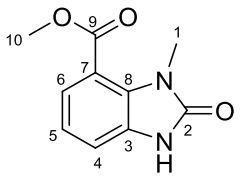
To a solution of 3-fluoro-2-nitrobenzoic acid (0.74 g, 4.0 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (6.0 mL) under argon atmosphere was added potassium carbonate (2.77 g, 20.0 mmol, 5.0 eq.) and the reaction mixture was stirred at 0 °C for 10 minutes. After this time, iodomethane (0.50 mL, 1.14 g, 8.0 mmol, 2.0 eq.) was added dropwise at 0 °C and allowed to warm to room temperature. After stirring at room temperature for 2 hours, H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (4 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (4 × 20 mL), dried over Na_2SO_4, filtered. The solvent was removed *in vacuo* to afford the title compound as a yellow crystalline solid (0.75 g, 3.76 mmol, 94% yield): R_f 0.78 (petroleum ether/EtOAc 1:3); m.p. 47–49 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3082 (w), 2964 (w), 2910 (w), 1727 (C=O, s), 1591 (w), 1542 (N=O, s), 1460 (w), 1371 (N=O, m), 1294 (s), 1216 (m), 1008 (s), 902 (m), 854 (w), 818 (w), 799 (w), 761 (s), 743 (w), 729 (w), 694 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.81 (1H, ddd, *J* 7.9, 1.3, $^5J_{HF}$ 1.2, C[6]-*H*), 7.58 (1H, ddd, *J* 8.6, 7.9, $^4J_{HF}$ 5.0, C[5]-*H*), 7.46 (1H, ddd, *J* 8.6, 1.3, $^3J_{HF}$ 8.7, C[4]-*H*), 3.91 (3H, s, C[8]-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.9 (d, $^4J_{CF}$ 3.1, C[7]), 153.7 (d, $^1J_{CF}$ 258.3, C[3]), 139.4 (d, $^2J_{CF}$ 15.0, C[2]), 131.8 (d, $^3J_{CF}$ 7.8, C[5]), 126.6 (d, $^4J_{CF}$ 3.5, C[6]), 125.3 (C[1]), 121.5 (d, $^2J_{CF}$ 19.4, C[4]), 53.5 (C[8]); ^{19}F NMR (377 MHz, $CDCl_3$) δ –123.2; HRMS *m/z* (EI) (Found [M][+] 199.0281, $C_8H_6FNO_4$ requires [M][+] 199.0275); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 9.1 min, 98.1%.

Methyl 3-(methylamino)-2-nitrobenzoate (92)



To a solution of methyl 3-fluoro-2-nitrobenzoate (**91**) (745.0 mg, 3.74 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (4.0 mL) under argon atmosphere was added methylamine hydrochloride (505.3 mg, 7.48 mmol, 2.0 eq.) and *N*,*N*-diisopropylethylamine (2.0 mL, 1.45 g, 11.22 mmol, 3.0 eq.). The reaction solution was stirred at room temperature for 20 hours. After this time, TLC showed that starting material had not been consumed. Additional methylamine hydrochloride (253.0 mg, 3.74 mmol, 1.0 eq.) and *N*,*N*-diisopropylethylamine (0.67 mL, 725.0 mg, 3.74 mmol, 1.0 eq.) was added and the reaction solution was stirred at room temperature for 20 hours. After this time, a saturated aqueous solution of $NaHCO_3$ (10 mL) and H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (4 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (4 × 20 mL), followed by brine (20 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (30–100% CH_2Cl_2 in petroleum ether gradient) to afford the title compound as a bright orange oil (575.3 mg, 2.74 mmol, 73% yield): R_f 0.25 (petroleum ether/CH_2Cl_2 1:1); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.53 (1H, br s, N-*H*), 7.43 (1H, ddd, *J* 8.7, 7.2, 0.7, C^5-*H*), 6.92 (1H, dd, *J* 8.7, 1.3, C^6-*H*), 6.77 (1H, dd, *J* 7.2, 1.3, C^4-*H*), 3.89 (3H, s, C^8-H_3), 3.01 (3H, s, C^9-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.0 (C^7), 145.5 (C^3), 134.7 (C^5), 132.3 (C^1), 130.5 (C^2), 116.1 (C^6), 115.5 (C^4), 53.2 (C^8), 30.2 (C^9); HRMS *m/z* (CI) (Found $[M]^+$ 210.0643, $C_9H_{10}N_2O_4$ requires $[M]^+$ 210.0641); Compound could not be detected by LRMS. The data is in good agreement with literature.[130]

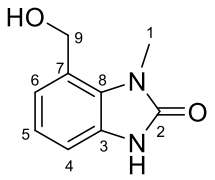
Methyl 1-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-4-carboxylate (94)



A solution of methyl 3-(methylamino)-2-nitrobenzoate (**92**) (574.8 mg, 2.74 mmol, 1.0 eq.) in EtOAc (9.0 mL) and EtOH (4.5 mL) was degassed under nitrogen for 5 minutes. Palladium on carbon (10% wt, 145.8 mg, 0.14 mmol, 0.05 eq.) was added and the reaction flask was purged with 3 balloons of hydrogen before inserting another balloon of hydrogen. The reaction mixture was stirred at room temperature for 4 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH (100 mL). The solvent was removed *in vacuo* to afford methyl 2-amino-3-(methylamino)benzoate (**93**), which was used without further purification.

To a solution of methyl 2-amino-3-(methylamino)benzoate (**93**) (362.1 mg, 2.01 mmol, 1.0 eq.) in CH_2Cl_2 (4.0 mL) was added 1,1′-carbonyldiimidazole (717.0 mg, 4.42 mmol, 2.2 eq.). The reaction mixture was stirred at room temperature for 3 hours. After this time, a precipitate formed and the reaction mixture was filtered and washed with H_2O (50 mL) to afford the title compound as a fluffy pink solid (338.9 mg, 1.64 mmol, 82% yield): R_f 0.53 (petroleum ether/EtOAc 1:3); m.p. 214–217 °C (from MeOH); $\bar{\nu}_{max}$/cm^{-1} 3151 (w), 3030 (w), 2954 (w), 1696 (C=O, s), 1623 (w), 1481 (w), 1431 (m), 1388 (m), 1308 (m), 1199 (w), 1122 (s), 1053 (w), 883 (w), 744 (s), 724 (s); ^{1}H NMR (500 MHz, MeOD) δ 7.67 (1H, dd, *J* 8.0, 1.1, C^6-*H*), 7.32 (1H, dd, *J* 7.9, 1.1, C^4-*H*), 7.17 (1H, dd, *J* 8.0, 7.9, C^5-*H*), 3.96 (3H, s, C^{10}-*H*$_3$), 3.42 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, MeOD) δ 167.4 (C^9), 156.8 (C^2), 133.3 (C^8), 130.4 (C^3), 123.6 (C^6), 121.2 (C^5), 113.3 (C^4), 112.8 (C^7), 52.6 (C^{10}), 27.1 (C^1); HRMS *m/z* (ESI$^-$) (Found [M+H]$^+$ 207.07642, $C_{10}H_{10}N_2O_3$ requires [M+H]$^+$ 207.07642); LRMS *m/z* (ESI$^+$) 215 ([M–Me+2H]$^+$, 100%); RP-HPLC (TFA), retention time = 6.9 min, 100.0%.

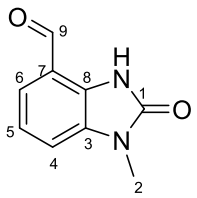
4-(Hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (95)



To a solution of methyl 1-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-4-carboxylate (**94**) (331.5 mg, 1.61 mmol, 1.0 eq.) suspended in dry toluene (13.0 mL) under argon atmosphere cooled to –78 °C was added DIBAL-H (1 M in hexanes, 11.3 mL, 11.3 mmol, 7.0 eq.) dropwise over 10 minutes. The reaction solution was stirred at –78 °C then warmed to room temperature. After stirring for 15 hours, the reaction solution was diluted with Et_2O (20 mL) and cooled to 0 °C. The reaction solution was quenched with H_2O (0.46 mL) dropwise (Fieser work-up), followed by aqueous NaOH solution (15% *w/w*, 0.46 mL) dropwise. H_2O (1.2 mL) was added and the reaction mixture was stirred at room temperature for 15 minutes. Due to insoluble clumps, saturated Rochelle's salt solution (100 mL) was added and the reaction mixture was stirred at room temperature until two distinct phases formed. The mixture was extracted with EtOAc (5 × 20 mL) and the organic components were washed with brine (20 mL), dried over Na_2SO_4 and filtered. The solvent was removed *in vacuo* and the crude material was purified with flash column chromatography on silica gel (80–100% EtOAc in petroleum ether with 2% triethylamine gradient to 0–10% EtOH in EtOAc with 2% triethylamine gradient) to afford the title compound as a fluffy colourless solid (207.0 mg, 1.16 mmol, 72% yield): R_f 0.12 (petroleum ether/EtOAc 1:4); m.p. 185–186 °C (from EtOH); $\bar{\nu}_{max}/cm^{-1}$ 3393 (br, O-H), 3006 (w), 2927 (w), 1678 (C=O, s), 1479 (m), 1438 (m), 1311 (w), 1200 (w), 1146 (w), 1026 (m), 994 (m), 896 (w), 789 (w), 730 (s), 664 (m), 634 (m); ^{1}H NMR (500 MHz, MeOD) δ 7.12–7.06 (2H, m, C^4-*H*, C^5-*H*), 7.03 (1H, dd, *J* 6.6, 2.4, C^6-*H*), 4.73 (2H, s, C^9-H_2), 3.39 (3H, s, C^1-H_3); ^{13}C NMR (126 MHz, MeOD) δ 157.1 (C^2), 132.4 (C^3), 127.2, 124.3 (C^7, C^8), 122.5. 121.8 (C^4, C^5), 108.2 (C^6), 62.5 (C^9), 27.1 (C^1); HRMS

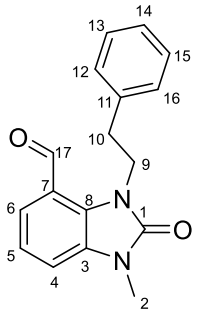
m/z (APCI) (Found $[M+H]^+$ 179.08131, $C_9H_{10}N_2O_2$ requires $[M+H]^+$ 179.08150); LRMS *m/z* (ESI^-) 177 ($[M-H]^-$, 100%); RP-HPLC (TFA), retention time = 5.4 min, 100.0%.

1-Methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-4-carbaldehyde (96)



To a solution of 4-(hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (**95**) (168.4 mg, 0.95 mmol, 1.0 eq.) suspended in CH_2Cl_2 (25 mL) was added manganese(IV) oxide (826.5 mg, 9.50 mmol, 10.0 eq.). The reaction mixture was stirred at room temperature for 3 hours. After this time, additional manganese(IV) oxide (165.2 mg, 1.90 mmol, 2.0 eq.) was added. After stirring at room temperature for 2 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* to afford the title compound as a pale yellow crystalline solid (163.9 mg, 0.93 mmol, 98% yield): R_f 0.58 (EtOAc); m.p. 180–181 °C (from EtOH); $\bar{\nu}_{max}/cm^{-1}$ 3000 (w), 2824 (w), 2765 (w), 1682 (C=O, s), 1613 (m), 1476 (w), 1414 (w), 1389 (m), 1310 (w), 1234 (m), 1177 (w), 1141 (w), 991 (w), 892 (m), 781 (s), 732 (m), 624 (w); ^{1}H NMR (500 MHz, MeOD) δ 10.05 (1H, s, C^9-*H*), 7.59 (1H, dd, *J* 7.9, 1.1, C^4-*H* or C^6-*H*), 7.40 (1H, dd, *J* 7.9, 1.1, C^4-*H* or C^6-*H*), 7.29 (1H, dd, *J* 7.9, 7.9, C^5-*H*), 3.44 (3H, s, C^2-H_3); ^{13}C NMR (126 MHz, MeOD) δ 193.2 (C^9), 157.3 (C^1), 133.3 (C^3), 128.7 (C^7), 126.3, 114.4 (C^4, C^6), 122.4 (C^5), 120.4 (C^8), 27.1 (C^2); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 177.06605, $C_9H_8N_2O_2$ requires $[M-H]^-$ 177.06585); LRMS *m/z* (ESI^+) 215 ($[M+K]^+$, 100%); RP-HPLC (TFA), retention time = 6.0 min, 100.0%.

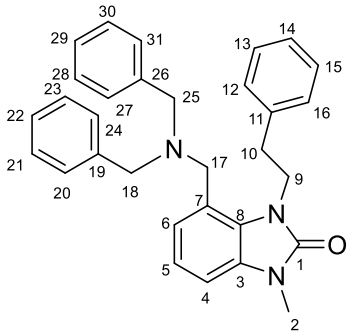
1-Methyl-2-oxo-3-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-4-carbaldehyde (97)



To a solution of 4-(hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (**96**) (164.0 mg, 0.93 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (4.0 mL) under argon atmosphere was added potassium carbonate (385.7mg, 2.79 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, (2-bromoethyl)benzene (0.26 mL, 344.2 mg, 1.86 mmol, 2.0 eq.) was added. The reaction mixture was stirred at 80 °C for 5 hours. A saturated aqueous solution of $NaHCO_3$ (10 mL) and H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (25–75% EtOAc in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (213.3 mg, 0.76 mmol, 82% yield): R_f 0.20 (petroleum ether/EtOAc 2:1); m.p. 121–123 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3062 (w), 2922 (w), 2852 (w), 1711 (C=O, s), 1617 (w), 1479 (m), 1444 (w), 1401 (w), 1145 (w), 982 (w), 902 (w), 749 (w), 702 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.12 (1H, s, C^{17}-*H*), 7.58 (1H, dd, *J* 7.7, 1.6, C^4-*H* or C^6-*H*), 7.30–7.26 (4H, m, C^{12}-*H*, C^{13}-*H*, C^{15}-*H*, C^{16}-*H*), 7.25–7.19 (2H, m, C^5-*H*, C^{14}-*H*), 7.20 (1H, dd, *J* 7.7, 1.5, C^4-*H* or C^6-*H*), 4.62–4.56 (2H, m, C^9-H_2), 3.46 (3H, s, C^2-H_3), 2.99–2.92 (2H, m, C^{10}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 190.4 (C^{17}), 154.7 (C^1), 138.2 (C^{11}), 131.6 (C^3), 129.3 (C^{12}, C^{16}), 128.6 (C^8, C^{13}, C^{15}), 127.2, 112.7 (C^4, C^6), 126.7 (C^{14}),

120.8 (C^5), 120.5 (C^7), 46.5 (C^9), 35.8 (C^{10}), 27.6 (C^1); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 281.12789, $C_{17}H_{16}N_2O_2$ requires $[M+H]^+$ 281.12845); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 9.6 min, 100.0%.

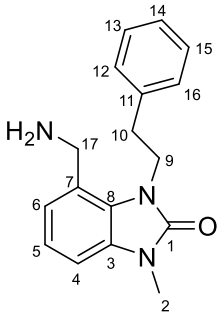
4-[(Dibenzylamino)methyl]-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (98)



To a solution of 1-methyl-2-oxo-3-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-4-carbaldehyde (**97**) (84.2 mg, 0.30 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and dibenzylamine (87 µL, 88.8 mg, 0.45 mmol, 1.5 eq.) and acetic acid (0.03 mL, 36.0 mg, 0.60 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 5 hours. After this time, sodium triacetoxyborohydride (319.8 mg, 1.50 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 18 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (10–50% EtOAc in petroleum ether gradient) followed by flash column chromatography on silica gel (66% EtOAc in cyclohexane isocratic) to afford the title compound as a colourless oil (90.8 mg, 0.20 mmol, 66% yield):

R_f 0.17 (petroleum ether/EtOAc 9:1); $\bar{\nu}_{max}$/cm^{-1} 3028 (w), 2924 (w), 2852 (w), 1705 (C=O, s), 1603 (w), 1475 (m), 1399 (w), 1241 (w), 1153 (w), 1070 (w), 1029 (w), 965 (w), 775 (w), 749 (m), 700 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.30–7.15 (13H, m, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*, C^{20}-*H*, C^{21}-*H*, C^{22}-*H*, C^{23}-*H*, C^{24}-*H*, C^{27}-*H*, C^{28}-*H*, C^{29}-*H*, C^{30}-*H*, C^{31}-*H*), 7.09 (1H, dd, *J* 7.8, 1.3, C^4-*H* or C^6-*H*), 7.02–6.95 (3H, m, C^5-*H*, C^{12}-*H*, C^{16}-H), 6.83 (1H, dd, *J* 7.8, 1.3, C^4-*H* or C^6-*H*), 4.45–4.39 (2H, m, C^9-*H*$_2$), 3.72 (2H, s, C^{17}-*H*$_2$), 3.49 (4H, s, C^{18}-*H*$_2$, C^{25}-*H*$_2$), 3.38 (3H, s, C^2-*H*$_3$), 2.83–2.78 (2H, m, C^{10}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.9 (C^2), 138.6 (C^{19}, C^{26}), 138.4 (C^{11}), 130.7 (C^3), 129.2, 128.6, 128.3, 127.2, 126.6 (C^{13}–C^{15}, C^{20}–C^{24}, C^{27}–C^{31}), 129.1 (C^{12}, C^{16}), 127.8 (C^8), 124.6, 106.9 (C^4, C^6), 120.8 (C^5), 57.8 (C^{18}, C^{25}), 56.0 (C^{17}), 44.3 (C^9), 36.3 (C^{10}), 27.3 (C^2); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 462.25336, $C_{31}H_{31}N_3O$ requires [M+H]$^+$ 462.25399); LRMS *m/z* (ESI$^+$) 265 ([M–$C_{14}H_{14}N_2$]$^+$, 100%); RP-HPLC (TFA), retention time = 9.3 min, 100.0%.

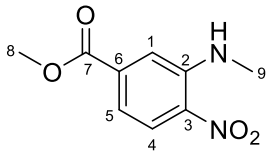
4-(Aminomethyl)-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (99)



EtOH (2.0 mL) in a flask was degassed under nitrogen for 5 minutes. Palladium black (36.6 mg, 0.34 mmol, 2.0 eq. (1.0 eq. per benzyl group)) was added and pre-loaded with hydrogen by purging the reaction flask with 2 balloons of hydrogen before inserting another balloon of hydrogen. A solution of 4-[(dibenzylamino)methyl]-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (**98**) (79.2 mg, 0.17 mmol, 1.0 eq.) in EtOH (7.0 mL) was degassed under nitrogen and added to the palladium suspension. The reaction mixture was stirred at room temperature for 48 hours. TLC showed that the reaction was not complete. The reaction flask was purged with 2 balloons of nitrogen before adding

palladium black (18.0 mg, 0.17 mmol, 1.0 eq.). The reaction flask was then purged with 2 balloons of hydrogen before inserting another balloon of hydrogen then stirred at room temperature for 20 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH (100 mL). The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (0–20% EtOH in EtOAc gradient) to afford the title compound as a colourless crystalline solid (40.7 mg, 0.15 mmol, 84% yield): R_f 0.28 (EtOAc with 2% triethylamine); m.p. 53–54 °C (from EtOH); $\bar{\nu}_{max}$/cm^{-1} 3373 (N-H, w), 3295 (N-H, w), 3059 (w), 2924 (w), 2853 (w), 1697 (C=O, s), 1604 (w), 1474 (m), 1400 (w), 1362 (w), 1156 (w), 1031 (w), 749 (m), 702 (m); ^{1}H NMR (500 MHz, MeOD) δ 7.29–7.23 (2H, m, C^{13}-*H*, C^{15}-*H*), 7.23–7.17 (3H, m, C^{12}-*H*, C^{14}-*H*, C^{16}-*H*), 7.16–7.06 (3H, m, C^4-*H*, C^5-*H*, C^6-*H*), 4.33–4.27 (2H, m, C^9-*H*$_2$), 4.04 (2H, s, C^{17}-*H*$_2$), 3.41 (3H, s, C^2-*H*$_3$), 3.02–2.97 (2H, m, C^{10}-*H*$_2$); ^{13}C NMR (126 MHz, MeOD) δ 156.3 (C^1), 139.3 (C^{11}), 131.8 (C^3), 130.0 (C^{12}, C^{16}), 129.6 (C^{13}, C^{15}), 127.8 (C^{14}), 127.4 (C^8), 125.2 (C^7), 124.1 (C^6), 122.9, 108.3 (C^4, C^5), 45.6 (C^9), 42.5 (C^{17}), 37.0 (C^{10}), 27.5 (C^2); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 282.16007, $C_{17}H_{19}N_3O$ requires [M+H]$^+$ 282.16009); LRMS *m/z* (ESI$^+$) 265 ([M–NH_2]$^+$, 100%); RP-HPLC (TFA), retention time = 6.4 min, 96.6%.

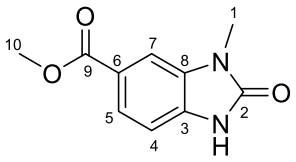
Methyl 3-(methylamino)-4-nitrobenzoate (101)[131]



To a solution of methyl 3-fluoro-4-nitrobenzoate (**51**) (4.72 g, 23.7 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (25.0 mL) under argon atmosphere was added methylamine hydrochloride (3.21 g, 47.4 mmol, 2.0 eq.) and *N*,*N*-diisopropylethylamine (13.0 mL, 9.20 g, 71.1 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 16 hours. After this time, H_2O (20 mL) and EtOAc (30 mL) was added. The aqueous layer was extracted with EtOAc (5 × 30 mL) and the organic components were washed with 0.5 M

aqueous LiCl solution (5 × 20 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo* to afford the title compound as an orange crystalline solid (4.91 g, 23.4 mmol, 99% yield): R_f 0.48 (petroleum ether/EtOAc 4:1); m.p. 125–126 °C (from EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.22 (1H, d, *J* 8.9, C^4-*H*), 8.00 (1H, br s, N-*H*), 7.55 (1H, d, *J* 1.8, C^1-*H*), 7.24 (1H, dd, *J* 8.9, 1.8, C^5-*H*), 3.95 (3H, s, C^8-*H*$_3$), 3.09 (3H, d, *J* 5.1, C^9-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.4 (C^{10}), 154.9 (C^2), 135.5 (C^9), 130.9 (C^4), 130.8 (C^7), 127.0 (C^6), 107.1 (C^5), 106.5 (C^8), 27.63, 27.55 (C^1, C^3); HRMS *m/z* (APCI) (Found $[M+H]^+$ 211.07139, $C_9H_{10}N_2O_4$ requires $[M+H]^+$ 211.07133); Compound could not be detected by LRMS. ^{1}H NMR spectrum is in good agreement with literature. [131]

Methyl 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (103)

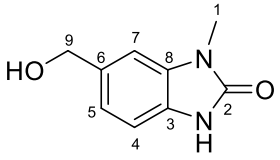


A suspension of methyl 3-(methylamino)-4-nitrobenzoate (**101**) (7.33 g, 34.9 mmol, 1.0 eq.) in EtOAc (90.0 mL) and EtOH (45.0 mL) was degassed under nitrogen for 5 minutes. Palladium on carbon (10% wt, 1.22 g, 1.15 mmol, 0.03 eq.) was added and the reaction flask was purged with 3 balloons of hydrogen before inserting another balloon of hydrogen. The reaction mixture was stirred at room temperature overnight. After 16 hours, the reaction mixture was filtered through Celite® and washed with EtOH (100 mL). The solvent was removed *in vacuo* to afford methyl 4-amino-3-(methylamino)benzoate (**102**), which was used without further purification.

To a solution of methyl 4-amino-3-(methylamino)benzoate (**102**) (6.26 g, 34.7 mmol, 1.0 eq.) in CH_2Cl_2 (45 mL) was added 1,1′-carbonyldiimidazole (12.4 g, 76.3 mmol, 2.2 eq.). The reaction mixture was stirred at room temperature for 5 hours. After this time, a precipitate was formed and the reaction mixture was filtered and washed with H_2O (50 mL) to afford the title compound as an off-white solid (6.20 g, 30.1 mmol, 87% yield over two

steps): R_f 0.56 (petroleum ether/EtOAc 1:4); m.p. 237–238 °C (from MeOH); $\bar{\nu}_{max}$/cm^{-1} 2932 (w), 2853 (w), 1694 (C=O, s), 1421 (m), 1365 (m), 1244 (w), 1158 (s), 1105 (m), 1021 (w), 872 (w), 768 (w); ^{1}H NMR (500 MHz, MeOD) δ 7.82 (1H, dd, *J* 8.3, 1.6, C^5-*H*), 7.74 (1H, dd, *J* 1.6, 0.6, C^7-*H*), 7.12 (1H, dd, *J* 8.3, 0.6, C^4-*H*), 3.91 (3H, s, C^{10}-*H$_3$*), 3.43 (3H, s, C^1-*H$_3$*) ; ^{13}C NMR (126 MHz, MeOD) δ 168.7 (C^9), 157.2 (C^2), 133.8 (C^3), 132.3 (C^8), 125.2 (C^5), 124.5 (C^6), 110.0 (C^7), 109.8 (C^4), 52.5 (C^{10}), 27.1 (C^1); HRMS *m/z* (ESI$^-$) (Found [M–H]$^-$ 205.06173, $C_{10}H_{10}N_2O_3$ requires [M–H]$^-$ 205.06187); LRMS *m/z* (ESI$^-$) 205 ([M–H]$^-$, 100%); RP-HPLC (TFA), retention time = 6.5 min, 96.3%.

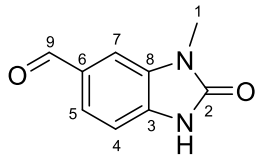
6-(Hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (104)



To a solution of methyl 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (**103**) (3.25 g, 15.7 mmol, 1.0 eq.) suspended in dry toluene (130 mL) under argon atmosphere cooled to –78 °C was added DIBAL-H (1 M in hexanes, 110 mL, 110 mmol, 7.0 eq.) dropwise over 45 minutes. The reaction solution was stirred at –78 °C then warmed to 0 °C. After 6 hours, the reaction solution was diluted with Et_2O (100 mL) and cooled to 0 °C. The reaction solution was quenched at 0 °C with H_2O (4.4 mL) dropwise (Fieser work-up), followed by aqueous NaOH solution (15% *w/w*, 4.4 mL) dropwise. H_2O (11 mL) was added and the reaction mixture was stirred at room temperature for 15 minutes. The mixture was dried over Na_2SO_4 and filtered. The solvent was removed *in vacuo* to afford the title compound as a colourless crystalline solid (2.54 g, 14.3 mmol, 91% yield): R_f 0.19 (petroleum ether/EtOAc 1:4); m.p. 196–197 °C (from EtOH); $\bar{\nu}_{max}$/cm^{-1} 3234 (br w, O-H), 2937 (w), 2876 (w), 1683 (C=O, s), 1504 (w), 1475 (w), 1394 (w), 1308 (w), 1192 (w), 1012 (w), 806 (w), 687 (w); ^{1}H NMR (500 MHz, MeOD) δ 7.12 (1H, d, *J* 1.5, C^7-*H*), 7.07 (1H, dd, *J* 8.0, 1.5, C^5-*H*), 7.03 (1H, d, *J* 8.0, C^4-*H*), 4.63 (2H, s, C^9-*H$_2$*), 3.39 (3H, s, C^1-*H$_3$*); ^{13}C NMR

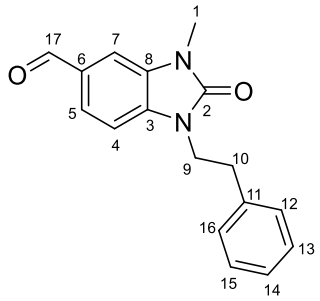
(126 MHz, MeOD) δ 157.2 (C^2), 136.4 (C^6), 132.4 (C^8), 128.8 (C^3), 122.0 (C^5), 110.0 (C^4), 108.1 (C^7), 65.5 (C^9), 27.0 (C^1); HRMS *m/z* (ESI^-) (Found $[M–H]^-$ 177.06678, $C_9H_{10}N_2O_2$ requires $[M–H]^-$ 177.06695); LRMS *m/z* (ESI^-) 177 ($[M–H]^-$, 100%); RP-HPLC (TFA), retention time = 4.8 min, 98.7%.

3-Methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (105)



To a solution of 6-(hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (**104**) (1.03 g, 5.8 mmol, 1.0 eq.) suspended in CH_2Cl_2 (100 mL) was added manganese(IV) oxide (5.04 g, 58.0 mmol, 10.0 eq.). The reaction mixture was stirred at room temperature overnight. After 16 hours, additional manganese(IV) oxide (1.51 g, 17.4 mmol, 3.0 eq.) was added. After stirring at room temperature for 4 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* to afford the title compound as a colourless solid (0.93 g, 5.3 mmol, 91% yield): R_f 0.46 (petroleum ether/EtOAc 1:4); m.p. 190 °C (decomposed) (from EtOH); $\bar{\nu}_{max}/cm^{-1}$ 2883 (w), 2850 (w), 2748 (w), 1682 (C=O, s), 1621 (m), 1504 (w), 1479 (m), 1391 (w), 1323 (w), 1174 (m), 1012 (w), 819 (w), 746 (w), 691 (w), 656 (w); ^1H NMR (500 MHz, MeOD) δ 9.91 (1H, s, C^9-*H*), 7.69 (1H, dd, *J* 8.0, 1.5, C^5-*H*), 7.65 (1H, d, *J* 1.5, C^7-*H*), 7.22 (1H, d, *J* 8.0, C^4-*H*), 3.45 (3H, s, C^1-H_3); ^{13}C NMR (126 MHz, MeOD) δ 193.3 (C^9), 157.2 (C^2), 135.3 (C^3), 133.1 (C^8), 132.3 (C^6), 127.3 (C^5), 110.2 (C^4), 108.8 (C^7), 27.2 (C^1); HRMS *m/z* (ESI^-) (Found $[M–H]^-$ 175.05113, $C_9H_8N_2O_2$ requires $[M–H]^-$ 175.05130); LRMS *m/z* (ESI^-) 175 ($[M–H]^-$, 100%); RP-HPLC (TFA), retention time = 5.6 min, 100.0%.

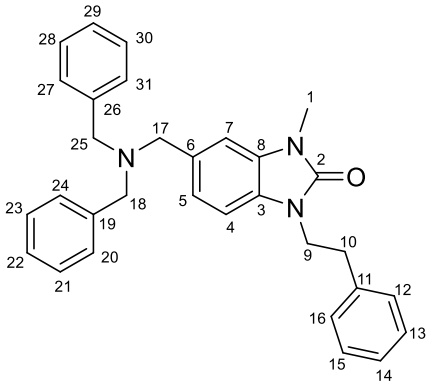
3-Methyl-2-oxo-1-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (106)



To a solution of 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**105**) (180.2 mg, 1.02 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (3.0 mL) under argon atmosphere was added potassium carbonate (422.7 mg, 3.06 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, (2-bromoethyl)benzene (0.28 mL, 377.5 mg, 2.04 mmol, 2.0 eq.) was added. The reaction mixture was stirred at 80 °C for 5 hours. A saturated aqueous solution of $NaHCO_3$ (10 mL) and H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were washed with brine (20 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (25–75% EtOAc in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (242.2 mg, 0.87 mmol, 85% yield): R_f 0.19 (petroleum ether/EtOAc 3:1); m.p. 128–129 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3062 (w), 2938 (w), 2816 (w), 1712 (C=O, s), 1682 (C=O, s), 1619 (m), 1506 (m), 1480 (m), 1397 (w), 1348 (w), 1185 (m), 802 (w), 748 (w), 701 (w), 668 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.91 (1H, s, C^{17}-*H*), 7.53 (1H, dd, *J* 8.0, 1.4, C^5-*H*), 7.51 (1H, d, *J* 1.4, C^7-*H*), 7.29–7.17 (5H, m, C^{11}-*H*, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*), 6.84 (1H, d, *J* 8.0, C^4-*H*), 4.14 (2H, dd, *J* 7.5, 7.4, C^9-H_2), 3.47 (3H, s, C^1-H_3), 3.05 (2H, dd, *J* 7.5, 7.4, C^{10}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.4 (C^{17}), 154.5 (C^2), 138.0 (C^{11}), 134.8 (C^3), 130.7, 130.6 (C^6, C^8), 128.94 (C^{12}, C^{16}), 128.85, 127.0

(C^{13}, C^{14}, C^{15}), 126.7 (C^{5}), 107.3 (C^{4}), 106.6 (C^{7}), 43.3 (C^{9}), 34.9 (C^{10}), 27.5 (C^{1}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 281.12843, $C_{17}H_{16}N_2O_2$ requires $[M+H]^+$ 281.12845); LRMS *m/z* (ESI$^+$) 251 ($[M–CO]^+$, 100%), 281 ($[M+H]^+$, 29%); RP-HPLC (TFA), retention time = 9.0 min, 99.0%.

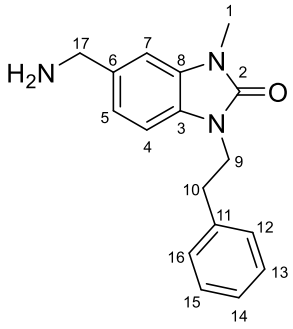
5-[(Dibenzylamino)methyl]-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (107)



To a solution of 3-methyl-2-oxo-1-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**106**) (78.5 mg, 0.28 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and dibenzylamine (81 µL, 82.8 mg, 0.42 mmol, 1.5 eq.) and acetic acid (0.03 mL, 33.6 mg, 0.56 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 6 hours. After this time, sodium triacetoxyborohydride (297.0 mg, 1.40 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 18 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (20–60% Et_2O in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (87.1 mg, 0.19 mmol, 67% yield): R_f 0.39 (petroleum ether/Et_2O 1:1); m.p. 62–64 °C (from petroleum ether/Et_2O); $\bar{\nu}_{max}$/cm^{-1}

3061 (w), 2924 (w), 2793 (w), 1704 (C=O, s), 1605 (w), 1504 (m), 1454 (w), 1398 (w), 1334 (w), 1166 (w), 1120 (w), 1073 (w), 972 (w), 807 (w), 745 (m), 698 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.42–7.38 (4H, m, C^{21}-*H*, C^{23}-*H*, C^{28}-*H*, C^{30}-*H*), 7.35–7.30 (4H, m, C^{20}-*H*, C^{24}-*H*, C^{27}-*H*, C^{31}-*H*), 7.29–7.18 (7H, m, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*, C^{16}-*H*, C^{22}-*H*, C^{29}-*H*), 7.06 (1H, dd, *J* 8.0, 1.4, C^{5}-*H*), 7.00 (1H, d, *J* 1.4, C^{7}-*H*), 6.78 (1H, d, *J* 8.0, C^{4}-*H*), 4.10–4.05 (2H, m, C^{9}-H_2), 3.59 (2H, s, C^{17}-H_2), 3.57 (4H, s, C^{18}-H_2, C^{25}-H_2), 3.43 (3H, s, C^{1}-H_3), 3.04–2.99 (2H, m, C^{10}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.5 (C^{2}), 139.7 (C^{19}, C^{26}), 138.5 (C^{11}), 133.0 (C^{6}), 130.2 (C^{8}), 129.0, 126.8 (C^{13}, C^{14}, C^{15}), 128.9 (C^{20}, C^{24}, C^{27}, C^{31}), 128.8 (C^{12}, C^{16}), 128.40, 128.38 (C^{21}, C^{23}, C^{28}, C^{30}, C^{3}), 127.1 (C^{22}, C^{29}), 121.8 (C^{5}), 108.0 (C^{7}), 107.2 (C^{4}), 58.2 (C^{17}), 58.0 (C^{18}, C^{25}), 43.0 (C^{9}), 35.0 (C^{10}), 27.3 (C^{1}); HRMS *m/z* (Found $[M+H]^+$ 462.25345, $C_{31}H_{31}N_3O$ requires $[M+H]^+$ 462.25399); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 9.1 min, 99.8%.

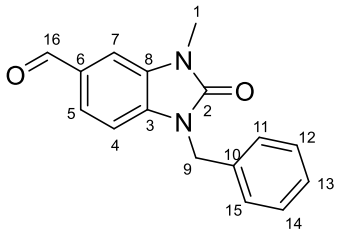
5-(Aminomethyl)-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (108)



EtOH (2.0 mL) in a flask was degassed under nitrogen for 5 minutes. Palladium black (11.0 mg, 0.10 mmol, 2.0 eq. (1.0 eq. per benzyl group)) was added and pre-loaded with hydrogen by purging the reaction flask with 2 balloons of hydrogen before inserting another balloon of hydrogen. A solution of 5-[(dibenzylamino)methyl]-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (**107**) (24.0 mg, 0.05 mmol, 1.0 eq.) in EtOH (7.0 mL) was degassed under nitrogen and added to the palladium suspension. The reaction mixture was stirred at room temperature for 44 hours. After this time, the reaction mixture was

filtered through Celite® and washed with EtOH (70 mL). The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) followed by a second flash column chromatography on silica gel (0–5% EtOH in $CHCl_3$ with 2% triethylamine) to afford the title compound as a colourless oil (11.9 mg, 0.40 mmol, 85% yield): R_f 0.07 (EtOH with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3368 (N-H, w), 3296 (N-H, w), 3027 (w), 2926 (w), 2854 (w), 1699 (C=O, s), 1505 (w), 1477 (m), 1400 (w), 1364 (w), 1169 (w), 806 (w), 747 (w), 701 (w); ^{1}H NMR (500 MHz, MeOD) δ 7.23–7.13 (6H, m, C^7-*H*, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*, C^{16}-*H*), 7.04 (1H, dd, *J* 8.0, 1.6, C^5-*H*), 6.96 (1H, d, *J* 8.0, C^4-*H*), 4.12 (2H, t, *J* 7.2, C^9-*H*$_2$), 3.86 (2H, s, C^{17}-*H*$_2$), 3.40 (3H, s, C^1-*H*$_3$), 3.01 (2H, t, *J* 7.2, C^{10}-*H*$_2$); ^{13}C NMR (126 MHz, MeOD) δ 156.1 (C^2), 139.6 (C^{11}), 136.0 (C^6), 131.3 (C^8), 130.0 (C^{12}, C^{16}), 129.6, 129.5 (C^3, C^{13}, C^{15}), 127.6 (C^{14}), 122.3 (C^5), 109.3 (C^4), 108.5 (C^7), 45.53 (C^{17}), 43.8 (C^9), 35.5 (C^{10}), 27.4 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 282.16048, $C_{17}H_{19}N_3O$ requires [M+H]$^+$ 282.16009); LRMS *m/z* (ESI$^+$) 265 ([M–NH_2]$^+$, 100%); RP-HPLC (TFA), retention time = 9.1 min, 96.7%.

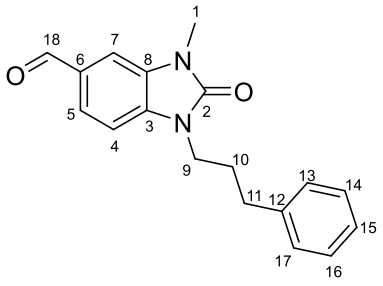
1-Benzyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (112)



To a solution of 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**105**) (180.0 mg, 1.02 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (3.0 mL) under argon atmosphere was added potassium carbonate (422.9 mg, 3.06 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, benzyl bromide (0.25 mL, 249.0 mg, 2.04 mmol, 2.0 eq.) was added. The reaction mixture was stirred at 80 °C for 5 hours. A saturated aqueous solution of $NaHCO_3$ (10 mL) and H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic

components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were washed with brine (20 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (20–75% EtOAc in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (218.1 mg, 0.82 mmol, 80% yield): R_f 0.16 (petroleum ether/EtOAc 3:1); m.p. 130–133 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3063 (w), 2938 (w), 2815 (w), 1712 (C=O, s), 1683 (C=O, s), 1618 (m), 1505 (m), 1479 (m), 1345 (w), 1303 (w), 1185 (m), 809 (w), 745 (w), 701 (w), 668 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.91 (1H, s, C^{16}-*H*), 7.55–7.52 (2H, m, C^5-*H*, C^7-*H*), 7.35–7.26 (5H, m, C^{11}-*H*, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*), 6.98 (1H, d, *J* 8.4, C^4-*H*), 5.11 (2H, s, C^9-H_2), 3.51 (3H, s, C^1-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.3 (C^{16}), 154.8 (C^2), 135.7 (C^{10}), 134.6 (C^3), 130.9 (C^6, C^8), 129.0, 128.2, (C^{12}, C^{13}, C^{14}), 127.6 (C^{11}, C^{15}), 127.0 (C^5), 108.0 (C^4), 106.6 (C^7), 45.3 (C^9), 27.6 (C^1); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 267.11287, $C_{16}H_{14}N_2O_2$ requires $[M+H]^+$ 267.11280); LRMS *m/z* (ESI^+) 251 ($[M-O]^-$, 65%), 267 ($[M+H]^+$, 100%); RP-HPLC (TFA), retention time = 8.7 min, 99.8%.

3-Methyl-2-oxo-1-(3-phenylpropyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (113)

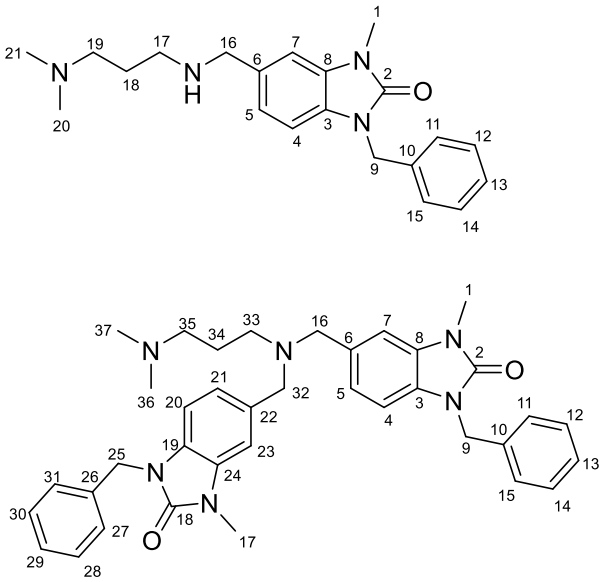


To a solution of 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**105**) (180.2 mg, 1.02 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (3.0 mL) under argon atmosphere was added potassium carbonate (422.7 mg, 3.06 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, 1-bromo-3-phenylpropane (0.31 mL, 406.1 mg, 2.04 mmol, 2.0 eq.) was added. The reaction mixture

was stirred at 80 °C for 5 hours. A saturated aqueous solution of $NaHCO_3$ (10 mL) and H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (20–75% EtOAc in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (265.7 mg, 0.90 mmol, 88% yield): R_f 0.15 (petroleum ether/EtOAc 3:1); m.p. 81–84 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3061 (w), 2939 (w), 2815 (w), 1714 (C=O, s), 1683 (C=O, s), 1619 (m), 1506 (m), 1480 (m), 1398 (w), 1348 (w), 1187 (m), 810 (w), 747 (w), 701 (w), 668 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.94 (1H, s, C^{18}-*H*), 7.60 (1H, dd, *J* 8.0, 1.4, C^5-*H*), 7.52 (1H, d, *J* 1.4, C^7-*H*), 7.29–7.25 (2H, m, C^{14}-*H*, C^{16}-*H*), 7.21–7.16 (3H, m, C^{13}-*H*, C^{15}-*H*, C^{17}-*H*), 6.98 (1H, d, *J* 8.0, C^4-*H*), 3.96 (2H, dd, *J* 7.4, 7.3, C^9-*H$_2$*), 3.46 (3H, s, C^1-*H$_3$*), 2.72 (2H, dd, *J* 7.7, 7.6, C^{11}-*H$_2$*), 2.12 (2H, tt, *J* 7.6, 7.4, C^{10}-*H$_2$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.4 (C^{18}), 154.7 (C^2), 140.8 (C^{12}), 134.8 (C^3), 130.8 (C^8), 130.7 (C^6), 128.6 (C^{14}, C^{16}), 128.4 (C^{13}, C^{17}), 126.8 (C^5), 126.3 (C^{15}), 107.3 (C^4), 106.6 (C^7), 41.2 (C^9), 33.1 (C^{11}), 29.7 (C^{10}), 27.5 (C^1); HRMS *m/z* (APCI) (Found $[M+H]^+$ 295.14400, $C_{18}H_{18}N_2O_2$ requires $[M+H]^+$ 295.14410); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 9.6 min, 99.1%.

1-Benzyl-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (114A) and

5,5'-[{[3-(Dimethylamino)propyl]azanediyl}bis(methylene)]bis(1-benzyl-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one) (114B)

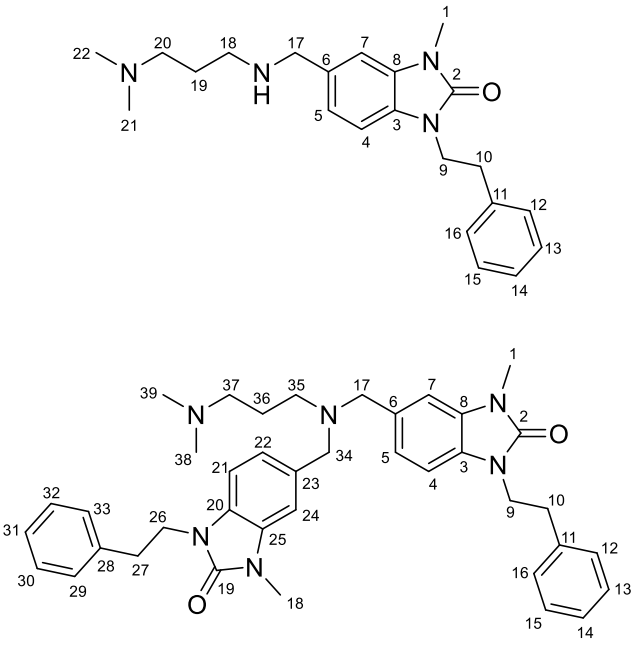


To a solution of 1-benzyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**112**) (93.3 mg, 0.35 mmol, 1.5 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (29 µL, 23.5 mg, 0.23 mmol, 1.0 eq.) and acetic acid (26 µL, 27.6 mg, 0.46 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 6 hours. After this time, sodium triacetoxyborohydride (243.8 mg, 1.15 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature overnight. After 16 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (0–100% EtOH in EtOAc with 2%

triethylamine) to afford **114A** as a colourless oil (22.7 mg, 0.06 mmol, 28% yield) and the side product **114B** as a colourless oil (19.9 mg, 0.03 mmol, 14% yield). **114A**: R_f 0.06 (EtOAc/EtOH 4:1 with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3481 (w), 2940 (w), 2819 (w), 1698 (C=O, s), 1505 (m), 1441 (m), 1399 (m), 1253 (w), 1178 (w), 1110 (w), 1027 (w), 808 (w), 744 (w), 700 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32–7.22 (5H, m, C^{11}-*H*, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*), 7.00 (1H, d, *J* 1.5, C^7-*H*), 6.93 (1H, dd, *J* 8.0, 1.5, C^5-*H*), 6.79 (1H, d, *J* 8.0, C^4-*H*), 5.06 (2H, s, C^9-*H*$_2$), 3.79 (2H, s, C^{16}-*H*$_2$), 3.45 (3H, s, C^1-*H*$_3$), 2.66 (2H, t, *J* 7.2, C^{17}-*H*$_2$), 2.50 (2H, t, *J* 7.3, C^{19}-*H*$_2$), 2.20 (6H, s, C^{20}-*H*$_3$, C^{21}-*H*$_3$), 1.80 (1H, br s, N-*H*), 1.67 (2H, tt, *J* 7.2, 7.3, C^{18}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.9 (C^2), 136.5 (C^{10}), 134.2 (C^6), 130.5 (C^8), 128.8, 127.6 (C^{12}, C^{13}, C^{14}), 128.3 (C^3), 127.8 (C^{11}, C^{15}), 121.3 (C^5), 108.0 (C^4), 107.5 (C^7), 58.2 (C^{19}), 54.4 (C^{16}), 48.1 (C^{17}), 45.7 (C^{20}, C^{21}), 45.1 (C^9), 28.2 (C^{18}), 27.4 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 353.23316, $C_{21}H_{28}N_4O$ requires [M+H]$^+$ 353.23359); LRMS *m/z* (ESI$^+$) 251 ([M–$C_5H_{13}N_2$]$^+$, 100%); RP-HPLC (TFA), retention time = 5.9 min, 98.0%; **114B**: R_f 0.64 (EtOAc/EtOH 4:1 with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3505 (w), 2929 (w), 2814 (w), 1703 (C=O, s), 1504 (m), 1440 (w), 1398 (w), 1256 (w), 1176 (w), 1099 (w), 1029 (w), 809 (w), 744 (w), 700 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36–7.25 (10H, m, C^{11}-*H*, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*, C^{27}-*H*, C^{28}-*H*, C^{29}-*H*, C^{30}-*H*, C^{31}-*H*), 7.00 (2H, d, *J* 1.4, C^7-*H*, C^{23}-*H*), 6.99 (2H, dd, *J* 7.9, 1.4, C^5-*H*, C^{21}-*H*), 6.80 (2H, d, *J* 7.9, C^4-*H*, C^{20}-*H*), 5.08 (4H, s, C^9-*H*$_2$, C^{25}-*H*$_2$), 3.59 (4H, s, C^{16}-*H*$_2$, C^{32}-*H*$_2$), 3.47 (6H, s, C^1-*H*$_3$, C^{17}-*H*$_3$), 2.46 (2H, t, *J* 7.3, C^{33}-*H*$_2$), 2.24 (2H, t, *J* 7.7, C^{35}-*H*$_2$), 2.18 (6H, s, C^{36}-*H*$_3$, C^{37}-*H*$_3$), 1.70 (2H, m, C^{34}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.7 (C^2, C^{18}), 136.4 (C^{10}, C^{26}), 133.2 (C^6, C^{22}), 130.2 (C^8, C^{24}), 128.7 (C^{11}, C^{15}, C^{27}, C^{31}), 128.2 (C^3, C^{19}), 128.3 (C^3), 127.5, 121.7 (C^{12}, C^{13}, C^{14}, C^{28}, C^{29}, C^{30}), 121.7 (C^4, C^{20}), 107.8 (C^5, C^{21}), 107.7 (C^7, C^{23}), 58.3 (C^{16}, C^{32}), 57.9 (C^{35}), 51.5 (C^{33}), 45.5 (C^{36}, C^{37}), 45.0 (C^9, C^{25}), 27.3 (C^1, C^{17}), 25.2 (C^{34}); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 603.34404, $C_{37}H_{42}N_6O_2$ requires [M+H]$^+$ 603.34420); LRMS *m/z* (ESI$^+$) 251 ([M–$C_{21}H_{27}N_4O$]$^+$, 100%); RP-HPLC (TFA), retention time = 7.2 min, 98.9%.

5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (115A) and

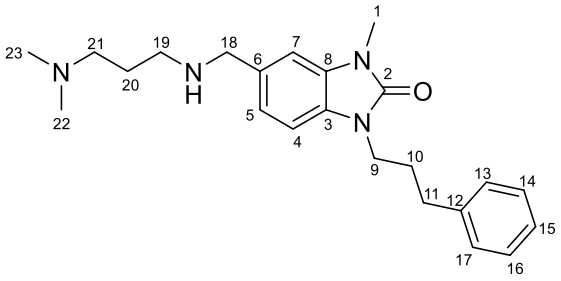
5,5'-[{[3-(Dimethylamino)propyl]azanediyl}bis(methylene)]bis[3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one] (115B)



To a solution of 3-methyl-2-oxo-1-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**106**) (98.0 mg, 0.35 mmol, 1.5 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (29 µL, 23.5 mg, 0.23 mmol, 1.0 eq.) and acetic acid (26 µL, 27.6 mg, 0.46 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 6 hours. After this time, sodium triacetoxyborohydride (244.1 mg, 1.15 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for overnight. After 15 hours, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was

purified by flash column chromatography on silica gel (0–100% EtOH in EtOAc with 2% triethylamine) to afford **115A** as a colourless oil (33.4 mg, 0.09 mmol, 40% yield) and the side product **115B** as a colourless oil (14.8 mg, 0.02 mmol, 10% yield). **115A**: R_f 0.06 (EtOAc/EtOH 4:1 with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3393 (w), 2939 (w), 2818 (w), 1698 (C=O, s), 1505 (m), 1456 (m), 1400 (m), 1283 (w), 1169 (w), 1102 (w), 1034 (w), 807 (w), 747 (w), 701 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.31–7.19 (5H, m, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*, C^{16}-*H*), 6.99 (1H, d, *J* 1.5, C^7-*H*), 6.97 (1H, dd, *J* 7.8, 1.5, C^5-*H*), 6.79 (1H, d, *J* 7.8, C^4-*H*), 4.10–4.15 (2H, m, C^9-*H*$_2$), 3.82 (2H, s, C^{17}-*H*$_2$), 3.41 (3H, s, C^1-*H*$_3$), 3.04–2.99 (2H, m, C^{10}-*H*$_2$), 2.69 (2H, t, *J* 7.1, C^{18}-*H*$_2$), 2.34 (2H, t, *J* 7.2, C^{20}-*H*$_2$), 2.22 (6H, s, C^{21}-*H*$_3$, C^{22}-*H*$_3$), 1.70 (2H, tt, *J* 7.1, 7.2, C^{19}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.5 (C^2), 138.4 (C^{11}), 133.6 (C^6), 130.4 (C^8), 129.0 (C^{12}, C^{16}), 128.7, 126.8 (C^{13}, C^{14}, C^{15}), 128.4 (C^3), 121.2 (C^5), 107.6 (C^7), 107.3 (C^4), 58.2 (C^{20}), 54.3 (C^{17}), 48.0 (C^{18}), 45.6 (C^{21}, C^{22}), 43.0 (C^9), 35.0 (C^{10}), 27.9 (C^{19}), 27.3 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 367.24877, $C_{22}H_{30}N_4O$ requires [M+H]$^+$ 367.24924); LRMS *m/z* (ESI$^+$) 265 ([M–$C_5H_{13}N_2$]$^+$, 100%); RP-HPLC (TFA), retention time = 6.1 min, 98.5%; **115B**: R_f 0.63 (EtOAc/EtOH 4:1 with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3504 (w), 2935 (w), 2814 (w), 1702 (C=O, s), 1504 (m), 1455 (w), 1398 (w), 1260 (w), 1166 (w), 1033 (w), 858 (w), 808 (w), 746 (w), 700 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.30–7.19 (10H, m, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*, C^{16}-*H*, C^{29}-*H*, C^{30}-*H*, C^{31}-*H*, C^{32}-*H*, C^{33}-*H*), 7.02 (2H, dd, *J* 8.0, 1.5, C^5-*H*, C^{22}-*H*), 6.99 (2H, d, *J* 1.5, C^7-*H*, C^{24}-*H*), 6.79 (2H, d, *J* 8.0, C^4-*H*, C^{21}-*H*), 4.11–4.05 (4H, m, C^9-*H*$_2$, C^{26}-*H*$_2$), 3.60 (4H, s, C^{17}-*H*$_2$, C^{34}-*H*$_2$), 3.42 (6H, s, C^1-*H*$_3$, C^{18}-*H*$_3$), 3.05–3.00 (4H, m, C^{10}-*H*$_2$, C^{27}-*H*$_2$), 2.47 (2H, t, *J* 7.2, C^{35}-*H*$_2$), 2.25 (2H, t, *J* 7.6, C^{37}-*H*$_2$), 2.19 (6H, s, C^{38}-*H*$_3$, C^{39}-*H*$_3$), 1.74–1.68 (2H, m, C^{36}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.5 (C^2, C^{19}), 138.4 (C^{11}, C^{28}), 133.1 (C^6, C^{23}), 130.2 (C^8, C^{25}), 129.0 (C^{12}, C^{16}, C^{29}, C^{33}), 128.8, 127.8 (C^{13}, C^{14}, C^{15}, C^{30}, C^{31}, C^{32}), 128.4 (C^3, C^{20}), 121.8 (C^5, C^{22}), 108.0 (C^7, C^{24}), 107.2 (C^4, C^{21}), 58.5 (C^{17}, C^{34}), 58.1 (C^{37}), 51.5 (C^{35}), 45.7 (C^{38}, C^{39}), 42.9 (C^9, C^{26}), 27.3 (C^1, C^{18}), 25.3 (C^{36}); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 631.37442, $C_{39}H_{46}N_6O_2$ requires [M+H]$^+$ 631.37550); LRMS *m/z* (ESI$^+$) 265 ([M–$C_{22}H_{29}N_4O$]$^+$, 100%); RP-HPLC (TFA), retention time = 6.1 min, 98.5%.

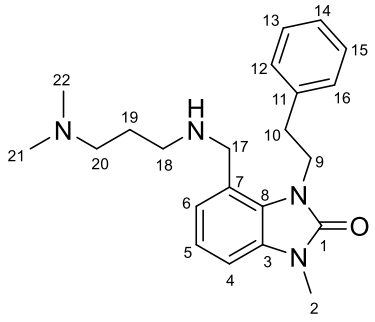
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(3-phenylpropyl)-1,3-dihydro-2*H*-benzimidazol-2-one (116)



To a solution of 3-methyl-2-oxo-1-(3-phenylpropyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**113**) (95.0 mg, 0.32 mmol, 1.0 eq.) in dry 1,2-dichloroethane (5.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (61 µL, 49.5 mg, 0.48 mmol, 1.5 eq.) and acetic acid (37 µL, 38.4 mg, 0.64 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 5 hours. After this time, sodium triacetoxyborohydride (339.0 mg, 1.60 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 16 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (0–100% EtOH in EtOAc with 2% triethylamine) to afford the title compound as a colourless oil (69.1 mg, 0.18 mmol, 57% yield): R_f 0.06 (EtOAc/EtOH 4:1 with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3481 (w), 2938 (w), 2817 (w), 1697 (C=O, s), 1505 (m), 1454 (m), 1399 (w), 1258 (w), 1165 (w), 1102 (w), 1031 (w), 805 (w), 745 (w), 700 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.28–7.24, 7.20–7.15 (5H, m, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*, C^{16}-*H*, C^{17}-*H*), 7.00 (1H, dd, *J* 7.8, 1.6, C^5-*H*), 6.98 (1H, d, *J* 1.6, C^7-*H*), 6.82 (1H, d, *J* 7.8, C^4-*H*), 3.90 (2H, t, *J* 7.2, C^9-*H*$_2$), 3.82 (2H, s, C^{18}-*H*$_2$), 3.41 (3H, s, C^1-*H*$_3$), 2.693 (2H, t, *J* 7.8, C^{11}-*H*$_2$), 2.686 (2H, t, *J* 7.1, C^{19}-*H*$_2$), 2.35–2.30 (2H, m, C^{21}-*H*$_2$), 2.21 (6H, s, C^{22}-*H*$_3$,

C^{23}-H_3), 2.12–2.04 (2H, m, C^{10}-H_2), 1.69 (2H, tt, *J* 7.2, 7.2, C^{20}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.7 (C^2), 141.2 (C^{12}), 133.8 (C^6), 130.4 (C^8), 128.50, 128.45, 128.4, 126.1 (C^3, C^{13}, C^{14}, C^{15}, C^{16}, C^{17}), 121.2 (C^5), 107.5 (C^7), 107.3 (C^4), 58.2 (C^{21}), 58.2 (C^{18}), 48.0 (C^{19}), 45.7 (C^{22}, C^{23}), 40.9 (C^9), 33.1 (C^{11}), 30.0 (C^{10}), 28.1 (C^{20}), 27.3 (C^1); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 381.2658, $C_{23}H_{32}N_4O$ requires $[M+H]^+$ 381.2654); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 6.4 min, 97.9%.

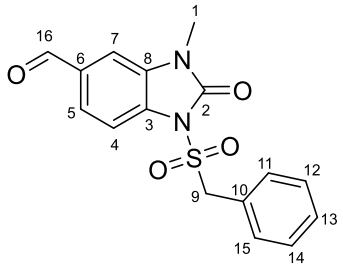
4-({[3-(Dimethylamino)propyl]amino}methyl)-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (117)



To a solution of 1-methyl-2-oxo-3-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-4-carbaldehyde (**97**) (70.2 mg, 0.25 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (48 µL, 38.3 mg, 0.38 mmol, 1.5 eq.) and acetic acid (29 µL, 30 mg, 0.50 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 5 hours. After this time, sodium triacetoxyborohydride (264.9 mg, 1.25 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 16 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (0–25% EtOH in EtOAc with 2%

triethylamine gradient) to afford the title compound as a colourless oil (79.0 mg, 0.22 mmol, 86% yield): R_f 0.23 (EtOAc with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 2940 (w), 2816 (w), 2766 (w), 1700 (C=O, s), 1475 (m), 1401 (w), 1157 (w), 1113 (w), 1030 (w), 750 (w), 702 (w); ^{1}H NMR (500 MHz, CD_2Cl_2) δ 7.33–7.20 (5H, m, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*, C^{15}-*H*, C^{16}-*H*), 7.02 (1H, dd, *J* 7.7, 7.6, C^5-*H*), 6.96 (1H, dd, *J* 7.7, 1.5, C^4-*H*), 6.92 (1H, dd, *J* 7.6, 1.5, C^6-*H*), 4.42–4.37 (2H, m, C^9-*H*$_2$), 3.88 (2H, s, C^{17}-*H*$_2$), 3.38 (3H, s, C^2-*H*$_3$), 3.04–2.98 (2H, m, C^{10}-*H*$_2$), 2.66 (2H, t, *J* 6.9, C^{18}-*H*$_2$), 2.25 (2H, t, *J* 7.1, C^{20}-*H*$_2$), 2.12 (6H, s, C^{21}-*H*$_3$, C^{22}-*H*$_3$), 1.79 (1H, br s, N-*H*), 1.59 (2H, tt, *J* 7.1, 6.9, C^{19}-*H*$_2$); ^{13}C NMR (126 MHz, CD_2Cl_2) δ 155.0 (C^1), 139.2 (C^{11}), 131.2 (C^3), 129.2 (C^{12}, C^{16}), 128.9 (C^{13}, C^{15}), 127.9 (C^8), 126.8 (C^{14}), 124.1 (C^4), 122.3 (C^7), 120.9 (C^5), 107.1 (C^6), 58.3 (C^{20}), 51.9 (C^{17}), 48.6 (C^{18}), 45.5 (C^{21}, C^{22}), 44.4 (C^9), 37.0 (C^{10}), 28.2 (C^{19}), 27.4 (C^2); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 367.24941, $C_{22}H_{30}N_4O$ requires $[M+H]^+$ 367.24924); LRMS *m/z* (ESI$^+$) 265 ($[M–C_5H_{13}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 6.0 min, 98.7%.

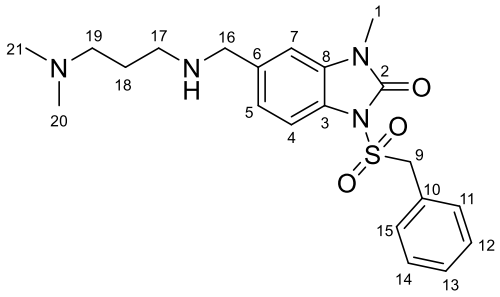
3-Methyl-2-oxo-1-(phenylmethanesulfonyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (121)



To a solution of 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**105**) (61.3 mg, 0.35 mmol, 1.0 eq.) in dry tetrahydrofuran (3.0 mL) under argon atmosphere was added triethylamine (98 μL, 70.8 mg, 0.70 mmol, 2.0 eq.), 4-(dimethylamino)pyridine (2.4 mg, 0.018 mmol, 0.05 eq.) and phenylmethanesulfonyl chloride (74.2 mg, 0.39 mmol, 1.1 eq.). The reaction mixture was stirred at room temperature for 16 hours. After this time, H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic components were washed with brine (20 mL), dried over Na_2SO_4, filtered, and

the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (30–60% EtOAc in petroleum ether gradient) to afford the title compound as a yellow crystalline solid (44.7 mg, 0.14 mmol, 39% yield): R_f 0.19 (petroleum ether/ EtOAc 2:1); m.p. 155–156 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3063 (w), 2925 (w), 2850 (w), 1738 (C=O, s), 1698 (C=O, s), 1618 (w), 1599 (w), 1473 (m), 1380 (S=O, m), 1303 (w), 1274 (w), 1170 (S=O, m), 1115 (w), 967 (w), 913 (w), 776 (w), 743 (w), 699 (w), 672 (w), 624 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.91 (1H, s, C^{16}-*H*), 7.55 (1H, d, *J* 8.2, C^4-*H*), 7.46 (1H, dd, *J* 8.2, 1.5, C^5-*H*), 7.45 (1H, d, *J* 1.5, C^7-*H*), 7.32–7.29 (2H, m, C^{11}-*H*, C^{15}-*H*), 7.25–7.20 (3H, m, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*), 4.89 (2H, s, C^9-*H*$_2$), 3.47 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.0 (C^{16}), 151.6 (C^2), 132.8 (C^6), 131.6 (C^3), 130.7 (C^{11}, C^{15}), 130.6 (C^8), 129.7 (C^{12}, C^{14}), 129.1 (C^{13}), 127.5 (C^5), 126.5 (C^{10}), 112.7 (C^4), 106.6 (C^7), 59.8 (C^9), 27.7 (C^1); HRMS *m/z* (APCI) (Found $[M+H]^+$ 331.07456, $C_{16}H_{14}N_2O_4S$ requires $[M+H]^+$ 331.07470); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 9.1 min, 98.8%.

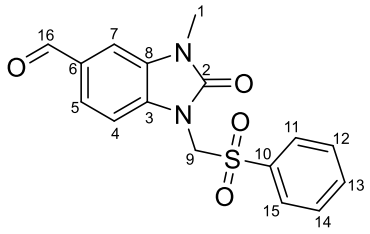
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(phenylmethanesulfonyl)-1,3-dihydro-2*H*-benzimidazol-2-one (122)



To a solution of 3-methyl-2-oxo-1-(phenylmethanesulfonyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**121**) (33.0 mg, 0.10 mmol, 1.0 eq.) in dry 1,2-dichloroethane (3.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (19 µL, 15.3 mg, 0.15 mmol, 1.5 eq.) and acetic acid (0.01 mL, 12.0 mg, 0.20 mmol, 2.0 eq.) was added. The

reaction mixture was warmed to room temperature and stirred for 6 hours. After this time, sodium triacetoxyborohydride (106.7 mg, 0.50 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 19 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) followed by flash column chromatography on silica gel (35–40% EtOH in $CHCl_3$ gradient with 2% triethylamine) to afford the title compound as a clear colourless oil (20.0 mg, 0.048 mmol, 48% yield): R_f 0.07 (EtOH with 2% triethylamine); $\bar{\nu}_{max}/cm^{-1}$ 2920 (m), 2850 (w), 1732 (C=O, s), 1624 (w), 1471 (m), 1367 (S=O, m), 1281 (w), 1169 (s=O, m), 1139 (w), 894 (w), 785 (w), 700 (w), 628 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36 (1H, d, *J* 8.3, C^4-*H*), 7.34–7.30 (2H, m, C^{11}-*H*, C^{15}-*H*), 7.26–7.21 (3H, m, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*), 6.94 (1H, d, *J* 1.6, C^7-*H*), 6.89 (1H, dd, *J* 8.3, 1.6, C^5-*H*), 4.84 (2H, s, C^9-H_2), 3.76 (2H, s, C^{16}-H_2), 3.40 (3H, s, C^1-H_3), 2.64 (2H, t, *J* 7.1, C^{17}-H_2), 2.32 (2H, t, *J* 7.2, C^5-H_2), 2.22 (6H, s, C^{20}-H_3, C^{21}-H_3), 1.00 (1H, s, N-*H*), 1.68 (2H, tt, *J* 7.2, 7.1, C^{18}-H_2) ^{13}C NMR (126 MHz, $CDCl_3$) δ 151.7 (C^2), 137.0 (C^6), 130.8 (C^{11}, C^{15}), 130.0 (C^8), 129.4 (C^{12}, C^{14}), 128.9 (C^{13}), 126.8 (C^{10}), 125.6 (C^3), 122.5 (C^5), 112.6 (C^4), 107.7 (C^7), 59.5 (C^9), 58.2 (C^{19}), 53.9 (C^{16}), 48.0 (C^{17}), 45.6 (C^{20}, C^{21}), 28.0 (C^{18}), 27.5 (C^1); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 417.19460, $C_{21}H_{28}N_4O_3S$ requires $[M+H]^+$ 417.19549); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 6.3 min, 99.2%.

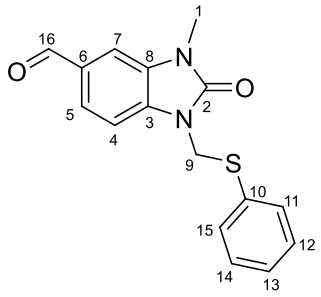
1-[(Benzenesulfonyl)methyl]-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (123)[116]



To a solution of 3-methyl-2-oxo-1-[(phenylsulfanyl)methyl]-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**124**) (74.5 mg, 0.25 mmol, 1.0 eq.) in dry CH_2Cl_2 (2.5 mL) under argon atmosphere cooled to 0 °C was added 3-chloroperbenzoic acid (124.8 mg, 0.50 mmol, 2.0 eq.) dissolved in dry CH_2Cl_2 (2.5 mL) dropwise. The reaction mixture was warmed to room temperature over 30 minutes then stirred at room temperature for 1.5 hours. After this time, the reaction was quenched with an aqueous solution of sodium thiosulfate (20% *w/v*, 10 mL). The aqueous layer was extracted with CH_2Cl_2 (3 × 20 mL) and the organic components were washed with a saturated aqueous solution of $NaHCO_3$ (10 mL), followed by brine (10 mL), dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (60–100% EtOAc in petroleum ether gradient to 0–10% EtOH in EtOAc gradient) to afford the title compound as a colourless solid (77.6 mg, 0.24 mmol, 94% yield): R_f 0.23 (petroleum ether/EtOAc 1:2); m.p. 195 °C (decomposed) (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3071 (w), 2996 (w), 2929 (w), 1726 (C=O, s), 1685 (C=O, s), 1617 (w), 1506 (w), 1477 (w), 1438 (w), 1416 (w), 1352 (w), 1315 (w), 1285 (m), 1191 (w), 1145 (s), 1084 (w), 875 (w), 823 (w), 754 (w), 738 (m), 690 (w), 669 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.98 (1H, s, C^{16}-H), 7.81–7.76 (2H, m, C^{11}-*H*, C^{15}-*H*), 7.72–7.66 (2H, m, C^5-*H*, C^{13}-*H*), 7.56–7.50 (3H, m, C^7-*H*, C^{12}-*H*, C^{14}-*H*), 7.39 (1H, d, *J* 8.1, C^4-*H*), 5.24 (2H, s, C^9-*H*$_2$), 3.32 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.2 (C^{16}), 153.2 (C^2), 136.5 (C^{10}), 135.0 (C^{13}), 133.0 (C^3), 132.1 (C^6), 130.9 (C^8), 129.6 (C^{12}, C^{14}), 129.1 (C^{11}, C^{15}), 127.5 (C^5), 109.4 (C^4), 106.8 (C^7), 62.1 (C^9), 27.7

(C[1]); HRMS *m/z* (ESI+) (Found $[M+Na]^+$ 353.05675, $C_{16}H_{14}N_2O_2S$ requires $[M+Na]^+$ 353.05665); LRMS *m/z* (ESI+) 331 ($[M+H]^+$, 9%), 353 ($[M+Na]^+$, 22%), 683 ($[2M+Na]^+$, 100%); RP-HPLC (TFA), retention time = 7.9 min, 99.7%.

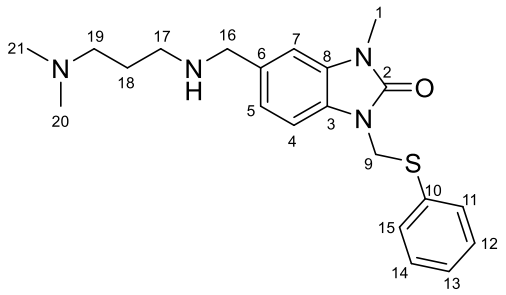
3-Methyl-2-oxo-1-[(phenylsulfanyl)methyl]-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (124)



To a solution of 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**105**) (158.7 mg, 0.90 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (6.0 mL) under argon atmosphere was added potassium carbonate (374.0 mg, 2.70 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, chloromethyl phenyl sulfide (241 μL, 285.6 mg, 1.80 mmol, 2.0 eq.) was added. The reaction mixture was stirred at room temperature for 18 hours. A saturated aqueous solution of $NaHCO_3$ (10 mL) and EtOAc (10 mL) were added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (50–100% Et_2O in petroleum ether gradient to 0–20% EtOAc in Et_2O gradient) to afford the title compound as a colourless solid (200.5 mg, 0.67 mmol, 75% yield): R_f 0.22 (petroleum ether/Et_2O 1:3); m.p. 102–104 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3060 (w), 2939 (w), 2812 (w), 1720 (C=O, s), 1685 (C=O, s), 1618 (w), 1504 (w), 1477 (m), 1437 (w), 1394 (w), 1347 (w), 1283 (w), 1186 (m), 1024 (w), 867 (w), 809 (w), 782 (w), 744 (w), 692 (w), 667 (w); ^1H NMR (500 MHz,

$CDCl_3$) δ 9.93 (1H, s, C^{16}-H), 7.54 (1H, dd, *J* 8.0, 1.5, C^5-*H*), 7.50 (1H, d, *J* 1.5, C^7-*H*), 7.42–7.39 (2H, m, C^{11}-*H*, C^{15}-*H*), 7.31–7.23 (3H, m, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*), 6.96 (1H, d, *J* 8.0, C^4-*H*), 5.29 (2H, s, C^9-H_2), 3.40 (3H, s, C^1-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.4 (C^{16}), 153.9 (C^2), 133.5 (C^3), 133.1 (C^{11}, C^{15}), 132.5 (C^{10}), 131.2 (C^6), 130.9 (C^8), 129.4 (C^{13}), 128.6 (C^{12}, C^{14}), 126.8 (C^5), 108.9 (C^4), 106.6 (C^7), 46.5 (C^9), 27.6 (C^1); HRMS *m/z* (ESI^+) (Found $[M+Na]^+$ 321.06687, $C_{16}H_{14}N_2O_2S$ requires $[M+Na]^+$ 321.06682); LRMS *m/z* (ESI^+) 321 ($[M+Na]^+$, 5%), 619 ($[2M+Na]^+$, 100%); RP-HPLC (TFA), retention time = 9.3 min, 98.9%.

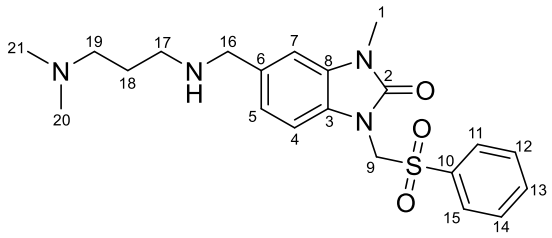
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-[(phenylsulfanyl)methyl]-1,3-dihydro-2*H*-benzimidazol-2-one (125)



To a solution of 3-methyl-2-oxo-1-[(phenylsulfanyl)methyl]-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**124**) (35.3 mg, 0.12 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (23 µL, 18.4 mg, 0.18 mmol, 1.5 eq.) and acetic acid (0.01 mL, 14.4 mg, 0.24 mmol, 2.0 eq.) were added. The reaction mixture was warmed to room temperature and stirred for 7 hours. After this time, sodium triacetoxyborohydride (128.2 mg, 0.60 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 17 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (3 × 20 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was

removed *in vacuo*. The crude material was purified using an Isolute® SCX-2 amine capture-release column (equilibrated with MeOH (10 mL), loaded in MeOH (2 mL), washed with MeOH (10 mL) and eluted with 2 M NH_3 in MeOH (10 mL)) to afford the title compound as a colourless oil (32.8 mg, 0.085 mmol, 71% yield): R_f 0.04 (petroleum ether/EtOAc 1:4); $\bar{\nu}_{max}$/cm^{-1} 3056 (w), 2938 (w), 2814 (w), 2765 (w), 1707 (C=O, s), 1628 (w), 1504 (m), 1475 (m), 1433 (m), 1394 (m), 1278 (w), 1165 (w), 1099 (w), 1027 (w), 806 (w), 742 (m), 682 (w), 657 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.44–7.39 (2H, m, C^{11}-*H*, C^{15}-*H*), 7.26–7.23 (3H, m, C^{12}-*H*, C^{13}-*H*, C^{14}-*H*), 6.98–6.94 (2H, m, C^5-*H*, C^7-*H*), 6.84 (1H, d, *J* 8.5, C^4-*H*), 5.27 (2H, s, C^9-*H*$_2$), 3.81 (2H, s, C^{16}-*H*$_2$), 3.34 (3H, s, C^1-*H*$_3$), 2.68 (2H, t, *J* 7.2, C^{17}-*H*$_2$), 2.32 (2H, t, *J* 7.3, C^{19}-*H*$_2$), 2.22 (6H, s, C^{20}-*H*$_3$, C^{21}-*H*$_3$), 1.69 (2H, tt, *J* 7.3, 7.2, C^{18}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.0 (C^2), 134.6 (C^6), 133.2 (C^{10}), 132.5 (C^{11}, C^{15}), 130.5 (C^8), 129.2 (C^{13}), 128.0 (C^{12}, C^{14}), 127.2 (C^3), 121.3, 107.6 (C^5, C^7), 108.9 (C^4), 58.2 (C^{19}), 54.4 (C^{16}), 48.1 (C^{17}), 46.2 (C^9), 45.7 (C^{20}, C^{21}), 28.2 (C^{18}), 27.4 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 385.20514, $C_{21}H_{28}N_4OS$ requires [M+H]$^+$ 385.20566); LRMS *m/z* (ESI$^+$) 385 ([M+H]$^+$, 100%); RP-HPLC (TFA), retention time = 6.3 min, 97.9%.

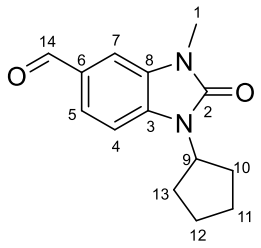
1-[(Benzenesulfonyl)methyl]-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (126)



To a solution of 1-[(benzenesulfonyl)methyl]-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**123**) (36.0 mg, 0.11 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (21 µL, 16.8 mg, 0.17 mmol, 1.5 eq.) and acetic acid (0.01 mL, 13.2 mg, 0.22 mmol, 2.0 eq.) was added. The

reaction mixture was warmed to room temperature and stirred for 3 hours. After this time, sodium triacetoxyborohydride (117.3 mg, 0.55 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 17.5 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (3 × 20 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified using an Isolute® SCX-2 amine capture-release column (equilibrated with MeOH (10 mL), loaded in MeOH (2 mL), washed with MeOH (10 mL) and eluted with 2 M NH_3 in MeOH (10 mL)) to afford the title compound as a colourless solid (28.8 mg, 0.069 mmol, 63% yield): R_f 0.08 (EtOH with 2% triethylamine); m.p. 75–78 °C (from MeOH); $\bar{\nu}_{max}$/cm^{-1} 22938 (w), 2858 (w), 2816 (w), 2767 (w), 1718 (C=O, s), 1627 (w), 1505 (w), 1474 (w), 1433 (w), 1398 (m), 1325 (w), 1146 (s), 1084 (w), 1040 (w), 934 (w), 867 (m), 809 (w), 738 (w), 688 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79–7.75 (2H, m, C^{11}-*H*, C^{15}-*H*), 7.25 (1H, dddd, *J* 7.5, 7.5, 1.3, 1.3, C^{13}-*H*), 7.52–7.47 (2H, m, C^{12}-*H*, C^{14}-*H*), 7.17 (1H, d, *J* 8.0, C^4-*H*), 7.05 (1H, dd, *J* 8.0, 1.5, C^5-*H*), 6.99 (1H, d, *J* 1.5, C^7-*H*), 5.19 (2H, s, C^9-*H*$_2$), 3.83 (2H, s, C^{16}-*H*$_2$), 3.25 (3H, s, C^1-*H*$_3$), 2.70 (2H, t, *J* 7.1, C^{17}-*H*$_2$), 2.33 (2H, t, *J* 7.2, C^{19}-*H*$_2$), 2.22 (6H, s, C^{20}-*H*$_3$, C^{21}-*H*$_3$), 1.70 (2H, tt, *J* 7.2, 7.1, C^{18}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 153.4 (C^2), 136.8 (C^{10}), 135.7 (C^6), 134.7 (C^{13}), 130.3 (C^8), 129.4 (C^{12}, C^{14}), 129.1 (C^{11}, C^{15}), 126.9 (C^3), 122.0 (C^5), 109.4 (C^4), 107.8 (C^7), 62.4 (C^9), 58.2 (C^{19}), 54.3 (C^{16}), 48.2 (C^{17}), 45.7 (C^{20}, C^{21}), 28.2 (C^{18}), 27.5 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 417.1973, $C_{21}H_{28}N_4O_3S$ requires [M+H]$^+$ 417.1960); LRMS *m/z* (ESI$^+$) 417 ([M+H]$^+$, 100%); RP-HPLC (TFA), retention time = 5.7 min, 98.9%.

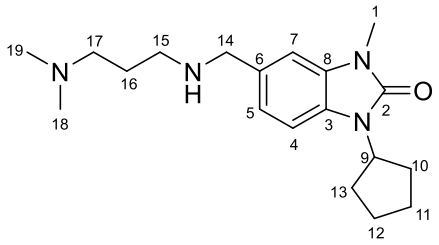
1-Cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (127)



To a solution of 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**105**) (88.2 mg, 0.50 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (3.0 mL) under argon atmosphere was added potassium carbonate (208.0 mg, 1.5 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, bromocyclopentane (108 µL, 149.0 mg, 1.0 mmol, 2.0 eq.) and tetrabutylammonium iodide (370.1 mg, 1.0 mmol, 2.0 eq.) were added. The reaction mixture was stirred at room temperature for 17 hours. After this time, a saturated aqueous solution of $NaHCO_3$ (10 mL) and H_2O (10 mL) was added. The aqueous layer was extracted with EtOAc (3 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (35–65% Et_2O in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (51.5 mg, 0.21 mmol, 42% yield): R_f 0.32 (petroleum ether/ Et_2O 1:2); m.p. 87–89 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 2955 (w), 2872 (w), 1709 (C=O, s), 1682 (C=O, s), 1619 (m), 1598 (m), 1502 (m), 1479 (m), 1429 (w), 1391 (w), 1343 (w), 1185 (m), 869 (w), 808 (w), 750 (w), 691 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.94 (1H, s, C^{14}-*H*), 7.60 (1H, dd, *J* 8.1, 1.5, C^5-*H*), 7.53 (1H, d, *J* 1.5, C^7-*H*), 7.16 (1H, d, *J* 8.1, C^4-*H*), 4.88 (1H, tt, *J* 8.8, 8.8, C^9-*H*), 3.45 (3H, s, C^1-*H*$_3$), 2.15–1.93 (6H, m, C^{10}-*H*$_2$, C^{11}-*H*$_a$, C^{12}-*H*$_a$, C^{13}-*H*$_2$), 1.80–1.70 (2H, m, C^{11}-*H*$_b$, C^{12}-*H*$_b$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.4 (C^{14}), 154.5 (C^2), 133.8 (C^3), 131.0 (C^8), 130.3 (C^6), 126.4 (C^5), 108.4 (C^4), 106.6 (C^7), 54.0 (C^9), 29.1 (C^{10}, C^{13}), 27.4 (C^1), 25.2 (C^{11}, C^{12}); HRMS *m/z* (APCI) (Found [M+H]$^+$ 245.12798, $C_{14}H_{16}N_2O_2$

requires $[M+H]^+$ 245.12845); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 8.8 min, 97.0%.

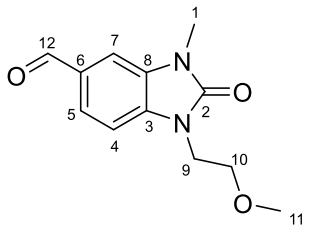
1-Cyclopentyl-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (128)



To a solution of 1-cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**127**) (45.0 mg, 0.18 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (35 µL, 28.6 mg, 0.28 mmol, 1.5 eq.) and acetic acid (0.02 mL, 22.2 mg, 0.37 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 4 hours. After this time, sodium triacetoxyborohydride (195.3 mg, 0.92 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 20 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) to afford the title compound as a colourless oil (57.3 mg, 0.17 mmol, 96% yield): R_f 0.10 (EtOH with 2% triethylamine); $\bar{\nu}_{max}/cm^{-1}$ 2947 (w), 2870 (w), 2816 (w), 2779 (w), 1700 (C=O, s), 1503 (w), 1477 (w), 1432 (w), 1395 (w), 1166 (w), 1100 (w), 1044 (w), 805 (w), 748 (w); ^{1}H NMR (500 MHz, MeOD) δ 7.18–7.15 (2H, m, C^4-*H*, C^7-*H*), 7.10 (1H, dd, *J* 8.2, 1.6, C^5-*H*), 4.86–4.78 (1H, m, C^9-*H*), 3.83 (2H, s, C^{14}-H_2), 3.41 (3H, s, C^1-H_3), 2.63 (2H, t, *J* 7.4,

C^{15}-H_2), 2.40–2.35 (2H, m, C^{17}-H_2), 2.25 (6H, s, C^{18}-H_3, C^{19}-H_3), 2.17–2.08 (2H, m, C^{10}-H_a, C^{13}-H_a), 2.04–1.95 (4H, m, C^{10}-H_b, C^{11}-H_a, C^{12}-H_a, C^{13}-H_b), 1.80–1.69 (4H, m, C^{11}-H_b, C^{12}-H_b, C^{16}-H_2); ^{13}C NMR (126 MHz, MeOD) δ 155.8 (C^2), 133.6 (C^6), 131.5 (C^8), 128.5 (C^3), 123.0 (C^5), 110.0, 109.4 (C^4, C^7), 58.7 (C^{17}), 55.2 (C^9), 54.4 (C^{14}), 48.0 (C^{15}), 45.4 (C^{18}, C^{19}), 29.8 (C^{10}, C^{13}), 27.7 (C^{16}), 27.3 (C^1), 26.0 (C^{11}, C^{12}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 331.24824, $C_{19}H_{30}N_4O$ requires $[M+H]^+$ 331.24924); LRMS *m/z* (ESI$^+$) 161 ($[M–C_5H_{13}N_2–\text{cyclopentyl}]^+$, 100%); RP-HPLC (TFA), retention time = 5.7 min, 96.7%.

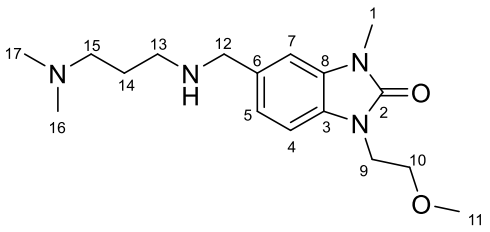
1-(2-Methoxyethyl)-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (129)



To a solution of 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**105**) (88.6 mg, 0.50 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (3.0 mL) under argon atmosphere was added potassium carbonate (207.5 mg, 1.50 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, 2-bromoethyl methyl ether (94 µL, 139.0 mg, 1.00 mmol, 2.0 eq.) was added. The reaction mixture was stirred at room temperature for 22 hours. After this time, a saturated aqueous solution of $NaHCO_3$ (5 mL) and H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (50–75% EtOAc in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (83.1 mg, 0.36 mmol, 71% yield): R_f 0.21 (petroleum ether/EtOAc 2:1); m.p. 103–104 °C (from petroleum

ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 2938 (w), 2892 (w), 2818 (w), 1713 (C=O, s), 1681 (C=O, s), 1619 (m), 1507 (m), 1480 (m), 1397 (w), 1351 (w), 1189 (m), 1120 (w),812 (w), 775 (w), 749 (w), 668 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.95 (1H, s, C^{12}-*H*), 7.63 (1H, dd, *J* 8.0, 1.5, C^{5}-*H*), 7.53 (1H, d, *J* 1.5, C^{7}-*H*), 7.23 (1H, d, *J* 8.0, C^{4}-*H*), 4.10 (2H, t, *J* 5.3, C^{9}-*H*$_2$), 3.69 (2H, t, *J* 5.3, C^{10}-*H*$_2$), 3.48 (3H, s, C^{1}-*H*$_3$), 3.32 (3H, s, C^{11}-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.5 (C^{12}), 154.8 (C^{2}), 135.6 (C^{3}), 130.8 (C^{6}, C^{8}), 126.9 (C^{5}), 108.4 (C^{4}), 106.5 (C^{7}), 70.9 (C^{10}), 59.2 (C^{11}), 41.9 (C^{9}), 27.5 (C^{1}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 235.10726, $C_{12}H_{14}N_2O_3$ requires $[M+H]^+$ 235.10772); Compound could not be detected by LRMS; RP-HPLC (TFA), retention time = 6.7 min, 99.0%.

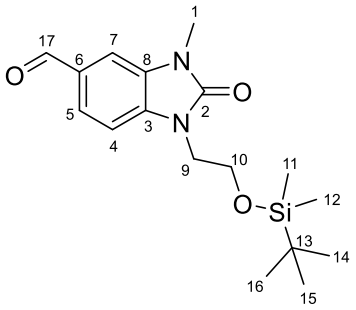
5-({[3-(Dimethylamino)propyl]amino}methyl)-1-(2-methoxyethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (130)



To a solution of 1-(2-methoxyethyl)-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**129**) (47.1 mg, 0.2 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (38 μL, 30.7 mg, 0.30 mmol, 1.5 eq.) and acetic acid (0.02 mL, 24.0 mg, 0.4 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 5 hours. After this time, sodium triacetoxyborohydride (212.5 mg, 1.00 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 18 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried

over Na_2SO_4, filtered, and the solvent removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (EtOH with 2% triethylamine isocratic) to afford the title compound as a pale yellow oil (37.8 mg, 0.12 mmol, 59% yield): R_f 0.06 (EtOH with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 2937 (w), 2860 (w), 2818 (w), 2779 (w), 1699 (C=O, s), 1506 (w), 1461 (w), 1400 (w), 1260 (w), 1181 (w), 1121 (m), 1041 (m), 971 (m), 808 (w), 747 (w), 725 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.03–7.00 (2H, m, C^4-*H*, C^5-*H* or C^7-*H*), 6.99–6.97 (1H, m, C^5-*H* or C^7-*H*), 4.04 (2H, t, *J* 5.6, C^9-*H*$_2$), 3.82 (2H, s, C^{12}-*H*$_2$), 3.67 (2H, t, *J* 5.6, C^{10}-*H*$_2$), 3.41 (3H, s, C^1-*H*$_3$), 3.32 (3H, s, C^{11}-*H*$_3$), 2.69 (2H, t, *J* 7.1, C^{13}-*H*$_2$), 2.32 (2H, t, *J* 7.2, C^{15}-*H*$_2$), 2.22 (6H, s, C^{16}-*H*$_3$, C^{17}-*H*$_3$), 1.69 (2H, tt, *J* 7.2, 7.1, C^{14}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.8 (C^2), 133.8 (C^6), 130.4 (C^8), 129.0 (C^3), 121.3, 107.4 (C^5, C^7), 108.1 (C^4), 70.8 (C^{10}), 59.1 (C^{11}), 58.3 (C^{15}), 54.4 (C^{12}), 48.1 (C^{13}), 45.7 (C^{16}, C^{17}), 41.4 (C^9), 28.1 (C^{14}), 27.3 (C^1); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 321.22849, $C_{17}H_{28}N_4O_2$ requires $[M+H]^+$ 321.22850); LRMS *m/z* (ESI$^+$) 219 ($[M–C_5H_{13}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 4.7 min, 96.8%.

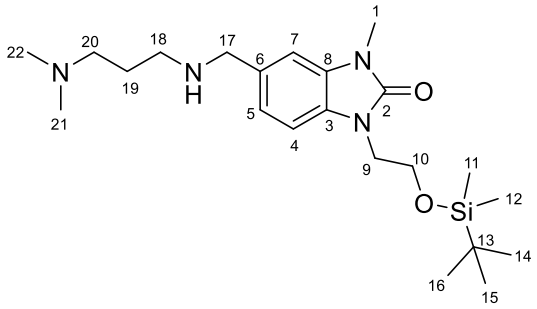
1-{2-{[*tert*-Butyl(dimethyl)silyl]oxy}ethyl)-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (131)



To a solution of 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**105**) (106.0 mg, 0.60 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (3.0 mL) under argon atmosphere was added potassium carbonate (249.2 mg, 1.80 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, 2-bromoethoxy-*tert*-butyl dimethylsilane (258 μL, 287.1 mg, 1.20 mmol, 2.0 eq.) was added. The reaction mixture was stirred at room temperature for 20 hours. After this time, a saturated aqueous

solution of $NaHCO_3$ (5 mL) and H_2O (10 mL) was added. The aqueous layer was then extracted with EtOAc (3 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (30–50% EtOAc in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (197.0 mg, 0.59 mmol, 98% yield): R_f 0.47 (petroleum ether/EtOAc 1:1); m.p. 81–83 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 2930 (w), 2856 (w), 2818 (w), 1715 (C=O, s), 1685 (C=O, s), 1618 (m), 1506 (m), 1480 (m), 1396 (w), 1349 (w), 1254 (w), 1189 (m), 948 (w), 900 (w), 834 (m), 809 (m), 777 (m), 749 (w), 724 (w), 699 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.94 (1H, s, C^{17}-*H*), 7.61 (1H, dd, *J* 8.1, 1.3, C^5-*H*), 7.53 (1H, d, *J* 1.3, C^7-*H*), 7.24 (1H, d, *J* 8.1, C^4-*H*), 4.05 (2H, t, *J* 5.2, C^9-H_2), 3.92 (2H, t, *J* 5.2, C^{10}-H_2), 3.48 (3H, s, C^1-H_3), 0.77 (9H, s, C^{14}-H_3, C^{15}-H_3, C^{16}-H_3), –0.11 (6H, s, C^{11}-H_3, C^{12}-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 191.5 (C^{17}), 154.7 (C^2), 135.9 (C^3), 130.64, 130.63 (C^6, C^8), 126.7 (C^5), 108.7 (C^4), 106.5 (C^7), 61.7 (C^{10}), 44.3 (C^9), 27.5 (C^1), 25.9 (C^{14}, C^{15}, C^{16}), 18.3 (C^{13}), –5.5 (C^{11}, C^{12}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 335.17869, $C_{17}H_{26}N_2O_3Si$ requires $[M+H]^+$ 335.17855); LRMS *m/z* (ESI^+) 203 ($[M–C_6H_{15}OSi]^+$, 100%); RP-HPLC (TFA), retention time = 11.1 min, 95.1%.

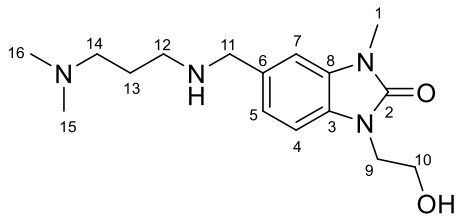
1-(2-{[*tert*-Butyl(dimethyl)silyl]oxy}ethyl)-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (132)



To a solution of 1-(2-{[*tert*-butyl(dimethyl)silyl]oxy}ethyl)-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**131**) (100.1 mg, 0.3 mmol, 1.0 eq.) in dry 1,2-dichloroethane (4.0 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 3-(dimethylamino)-1-propylamine (57 µL, 46.0 mg, 0.45 mmol, 1.5 eq.) and acetic acid (0.03 mL, 36.0 mg, 0.60 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 4 hours. After this time, sodium triacetoxyborohydride (318.1 mg, 1.50 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 20 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solvent was removed *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (10 mL) and $CHCl_3$ (10 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (5 × 15 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (10–60% EtOH in EtOAc with 2% triethylamine gradient) flash column chromatography on silica gel (50–70% EtOH in $CHCl_3$ with 2% triethylamine gradient) to afford the title compound as a pale yellow oil (61.9 mg, 0.15 mmol, 49% yield): R_f 0.08 (10% EtOH in EtOAc with 2% triethylamine); $\bar{\nu}_{max}/cm^{-1}$ 2929 (m), 2856 (w), 2816 (w), 2779 (w), 1707 (C=O, s), 1505 (w), 1464 (w), 1399 (w), 1255 (w), 1191 (w), 1108 (m), 837 (m), 778

(m), 748 (w), 725 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.04 (1H, d, *J* 8.4, C^4-*H*), 7.02–6.98 (2H, m, C^5-*H*, C^7-*H*), 4.01–3.97 (2H, m, C^9-H_2), 3.91–3.86 (2H, m, C^{10}-H_2), 3.85 (2H, s, C^{17}-H_2), 3.41 (3H, s, C^1-H_3), 2.72 (2H, t, *J* 7.0, C^{18}-H_2), 2.35 (2H, t, *J* 7.1, C^{20}-H_2), 2.23 (6H, s, C^{21}-H_3, C^{22}-H_3), 1.72 (2H, tt, *J* 7.1, 7.0, C^{19}-H_2), 0.79 (9H, s, C^{14}-H_3, C^{15}-H_3, C^{16}-H_3), –0.10 (6H, s, C^{11}-H_3, C^{12}-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.7 (C^2), 130.3 (C^6, C^8), 129.5 (C^3), 121.3, 107.4 (C^5, C^7), 108.5 (C^4), 61.6 (C^{10}), 58.4 (C^{20}), 54.2 (C^{17}), 48.0 (C^{18}), 45.6 (C^{21}, C^{22}), 43.9 (C^9), 27.5 (C^{19}), 27.3 (C^1), 26.0 (C^{11}, C^{12}), 18.3 (C^{13}), –5.5 (C^{14}, C^{15}, C^{16}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 421.29875, $C_{22}H_{40}N_4O_2Si$ requires $[M+H]^+$ 421.29933); LRMS *m/z* (ESI$^+$) 319 ($[M–C_5H_{13}N_2]^+$, 100%); RP-HPLC (TFA), retention time = 7.1 min, 97.5%.

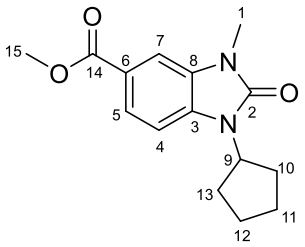
5-({[3-(Dimethylamino)propyl]amino}methyl)-1-(2-hydroxyethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (133)



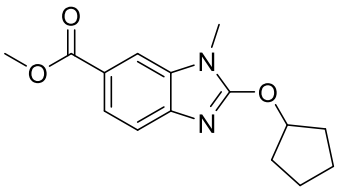
To 1-(2-{[*tert*-butyl(dimethyl)silyl]oxy}ethyl)-5-({[3-(dimethylamino)propyl]amino} methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (**132**) (19.5 mg, 0.046 mmol, 1.0 eq.) in a flask was added hydrogen chloride (1 M in MeOH) (4.0 mL). The reaction mixture was stirred at room temperature for 1 hour. After this time, a saturated aqueous solution of $NaHCO_3$ was added until pH 7 was reached and the solvent was removed *in vacuo*. The residue was stirred in EtOH to extract the product and the solvent was removed *in vacuo*. The crude material was purified using an Isolute® SCX-2 amine capture-release column (equilibrated with MeOH (10 mL), loaded in MeOH (2 mL), washed with MeOH (10 mL) and eluted with 2 M NH_3 in MeOH (10 mL)) to afford the title compound as a colourless oil (12.3 mg, 0.040 mmol, 87% yield): R_f 0.11 ($CHCl_3$/EtOH 1:1 with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3372 (O-H, br), 2938 (m), 2859 (w), 2817 (w), 2781 (w), 1698

(C=O, s), 1505 (w), 1461 (w), 1399 (w), 1183 (w), 1109 (m), 1070 (w), 805 (w), 748 (w), 668 (w); ^{1}H NMR (500 MHz, MeOD) δ 7.19–7.16 (2H, m, C^{4}-*H*, C^{7}-*H*), 7.11 (1H, dd, *J* 8.0, 1.6, C^{5}-*H*), 4.01 (2H, t, *J* 5.6, C^{9}-*H*$_2$), 3.84–3.79 (4H, m, C^{10}-*H*$_2$, C^{11}-*H*$_2$), 3.43 (3H, s, C^{1}-*H*$_3$), 2.61 (2H, t, *J* 7.5, C^{12}-*H*$_2$), 2.38–2.32 (2H, m, C^{24}-*H*$_2$), 2.24 (6H, s, C^{15}-*H*$_3$, C^{16}-*H*$_3$), 1.72 (2H, tt, *J* 7.5, 7.5, C^{13}-*H*$_2$); ^{13}C NMR (126 MHz, MeOD) δ 156.4 (C^{2}), 134.1 (C^{6}), 131.4 (C^{8}), 130.3 (C^{3}), 123.2 (C^{5}),109.4, 109.2 (C^{4}, C^{7}), 60.8 (C^{9}), 58.7 (C^{14}), 54.5 (C^{11}), 48.0 (C^{12}), 45.4 (C^{15}, C^{16}), 44.8 (C^{10}), 27.8 (C^{13}), 27.4 (C^{1}); HRMS *m/z* (ESI^{+}) (Found [M+H]$^{+}$ 307.2133, $C_{16}H_{26}N_4O_2$ requires [M+H]$^{+}$ 307.2134); LRMS *m/z* (ESI^{+}) 307 ([M+H]$^{+}$, 100%); RP-HPLC (TFA), retention time = 4.68 min, 97.8%.

Methyl 1-cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (140A)



Side product:



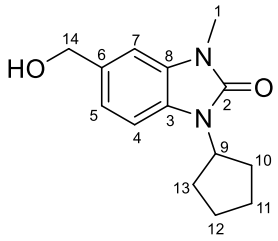
To a solution of methyl 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (**103**) (2.06 g, 10.0 mmol, 1.0 eq.) in dry *N*,*N*-dimethylformamide (40 mL) under argon atmosphere was added potassium carbonate (4.15 g, 30.0 mmol, 3.0 eq.). The reaction mixture was stirred at room temperature for 30 minutes. After this time, bromocyclopentane (3.20 mL, 4.47 g, 30.0 mmol, 3.0 eq.) and tetrabutylammonium iodide (3.69 g, 10 mmol, 1.0 eq.) were added. The reaction mixture was stirred at room temperature for 18 hours. After this time, TLC indicated that the reaction was incomplete. Bromocyclopentane (1.10 mL, 1.49 g, 10.0 mmol, 1.0 eq.) was added and the reaction mixture was stirred at 40 °C for 48 hours. H_2O (50 mL) and EtOAc (50 mL) were added. The aqueous layer was then extracted with EtOAc (10 × 40 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (6 × 20 mL). The organic components were

then washed with brine (20 mL), dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo*. The crude material was purified with flash column chromatography on silica gel (20–70% Et_2O in petroleum ether gradient) to afford the title compound as a colourless crystalline solid (1.96 g, 7.14 mmol, 71% yield): R_f 0.28 (petroleum ether/Et_2O 1:1); m.p. 115–118 °C (from petroleum ether/Et_2O); $\bar{\nu}_{max}/cm^{-1}$ 2952 (w), 2874 (w), 1704 (C=O, s), 1625 (w), 1503 (w), 1479 (w), 1432 (w), 1393 (w), 1297 (w), 1298 (m), 1114 (w), 764 (w); ^1H NMR (500 MHz, $CDCl_3$) δ 7.81 (1H, dd, *J* 8.3, 1.6, C^5-*H*), 7.64 (1H, d, *J* 1.6, C^7-*H*), 7.05 (1H, d, *J* 8.3, C^4-*H*), 4.85 (1H, tt, *J* 8.9, 8.9, C^9-*H*), 3.91 (3H, s, C^{15}-H_3), 3.43 (3H, s, C^1-H_3), 2.13–1.90 (6H, m, C^{10}-H_2, C^{11}-H_a, C^{12}-H_a, C^{13}-H_2), 1.77–1.67 (2H, m, C^{11}-H_b, C^{12}-H_b); ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.2 (C^{14}), 154.5 (C^2), 132.1 (C^3), 130.1 (C^8), 123.5 (C^5), 122.8 (C^6), 108.7 (C^7), 108.0 (C^4), 53.8 (C^9), 52.2 (C^{15}), 29.0 (C^{10}, C^{13}), 27.3 (C^1), 25.1 (C^{11}, C^{12}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 275.1398, $C_{15}H_{19}N_2O_3$ requires $[M+H]^+$ 175.1396); LRMS *m/z* (ESI^+) 275 ($[M+H]^+$, 30%), 571 ($[2M+Na]^+$, 100%); RP-HPLC (TFA), retention time = 9.6 min, 100.0%.

The *O*-alkylated product was also formed in 15% yield and hydrolysed according to the following procedure:

To a solution of methyl 2-(cyclopentyloxy)-1-methyl-1*H*-benzimidazole-6-carboxylate (**140B**) (404.4 mg, 1.48 mmol, 1.0 eq.) in EtOH (15 mL) heated to reflux was added hydrobromic acid (48% wt in H_2O, 167 µL, 248.7 mg, 1.48 mmol, 1.0 eq.). The reaction mixture was under reflux for 4 hours. After this time, the reaction mixture was cooled to room temperature. Colourless crystals formed and were filtered under vacuum and washed with cold EtOH to recover the title compound as a colourless crystalline solid (237.0 mg, 1.15 mmol, 78% yield).

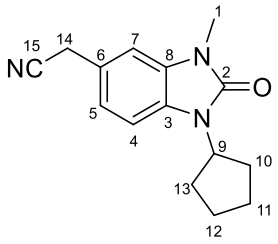
1-Cyclopentyl-5-(hydroxymethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (141)



To a solution of methyl 1-cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (**140A**) (1.95 g, 7.11 mmol, 1.0 eq.) dissolved in dry toluene (58.0 mL) under argon atmosphere cooled to −78 °C was added DIBAL-H (1 M in hexanes, 50.0 mL, 49.8 mmol, 7.0 eq.) dropwise over 30 minutes. The reaction mixture was stirred at −78 °C for 4 hours then slowly warmed to 0 °C. After stirring for 2 hours, the reaction was quenched with an aqueous saturated solution of Rochelle's salt (added dropwise initially, at 0 °C). The reaction mixture was stirred overnight until the clumps dissolved and two distinct layers were formed. The toluene/hexane layer was separated from the aqueous layer, which was then extracted with CH_2Cl_2 (10 × 20 mL). The organic components were each washed with H_2O (20 mL), brine (20 mL), combined, dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by filtering through a short silica gel plug (100% petroleum ether, followed by 50% EtOAc in petroleum ether, then 100% EtOAc) to afford the title compound as a fluffy colourless crystalline solid (1.62 g, 6.58 mmol, 93% yield): R_f 0.19 (petroleum ether/EtOAc 1:2); m.p. 145–146 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3403 (br, O-H), 2954 (w), 2871 (w), 1679 (C=O, s), 1502 (w), 1475 (w), 1434 (w), 1395 (m), 1189 (w), 1165 (w), 1026 (w), 811 (w), 748 (w), 669 (w), 611 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.06–7.01 (3H, m, C^4-*H*, C^5-*H*, C^7-*H*), 4.87 (1H, tt, *J* 8.9, 8.9, C^9-*H*), 4.72 (2H, d, *J* 5.9, C^{14}-H_2), 3.38 (3H, s, C^1-H_3), 2.14–1.90 (6H, m, C^{10}-H_2, C^{11}-H_a, C^{12}-H_a, C^{13}-H_2), 1.83 (1H, t, *J* 5.9, O-*H*), 1.78–1.68 (2H, m, C^{11}-H_b, C^{12}-H_b); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.5 (C^2), 134.1 (C^6), 130.7 (C^8), 127.8 (C^3), 120.1, 106.7 (C^5, C^7), 108.7 (C^4), 65.8 (C^{14}), 53.7 (C^9),

28.9 (C^{10}, C^{11}), 27.2 (C^{1}), 25.2 (C^{11}, C^{12}); HRMS *m/z* (ESI^+) (Found $[M+Na]^+$ 269.12609, $C_{14}H_{18}N_2O_2$ requires $[M+Na]^+$ 269.12605); LRMS *m/z* (ESI^+) 247 ($[M+H]^+$, 17%), 515 ($[2M+Na]^+$, 100%); RP-HPLC (TFA), retention time = 7.3 min, 100.0%.

(1-Cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)acetonitrile (143)

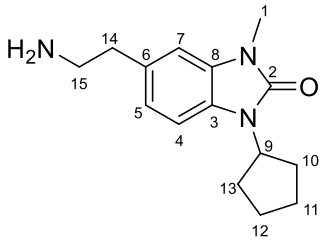


To a solution of 1-cyclopentyl-5-(hydroxymethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (**141**) (491.8 mg, 2.0 mmol, 1.0 eq.) in dry CH_2Cl_2 (10.0 mL) under argon atmosphere was added thionyl chloride (292 µL, 475.9 mg, 4.0 mmol, 2.0 eq.). After stirring at room temperature for 1 hour, the solvent was removed *in vacuo* to afford 5-(chloromethyl)-1-cyclopentyl-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (**142**), which was used without further purification.

To a solution of 5-(chloromethyl)-1-cyclopentyl-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (**142**) in dry *N*,*N*-dimethylformamide (10.0 mL) under argon atmosphere was added sodium cyanide (198.0 mg, 4.0 mmol, 2.0 eq.). The reaction mixture was stirred at 70 °C for 2.5 hours. After this time, TLC suggested that no reaction had taken place. Tetrabutylammonium iodide (296.0 mg, 0.8 mmol, 0.4 eq.) was added and the reaction mixture was stirred at 70 °C for 19 hours. After this time, H_2O (20 mL) was added. The aqueous layer was then extracted with EtOAc (4 × 20 mL) and the organic components were washed with 0.5 M aqueous LiCl solution (4 × 20 mL). The organic components were then washed with brine (20 mL), dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (35–

45% EtOAc in petroleum ether gradient) to afford the title compound as a white solid (525.7 mg, 2.06 mmol, 86% yield over 2 steps): R_f 0.42 (petroleum ether/EtOAc 1:2); m.p. 101–102 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 2954 (w), 2873 (w), 1694 (C=O, s), 1630 (w), 1504 (m), 1432 (w), 1394 (m), 1333 (w), 1254 (w), 1168 (w), 802 (w), 747 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.03 (1H, d, *J* 8.1, C^4-*H*), 6.98 (1H, dd, *J* 8.1, 1.7, C^5-*H*), 6.95 (1H, dd, *J* 1.7, 0.7, C^7-*H*), 4.86 (1H, tt, *J* 8.9, 8.9, C^9-*H*$_2$), 3.80 (2H, s, C^{14}-*H*$_2$), 3.41 (3H, s, C^1-*H*$_3$), 2.12–1.90 (6H, m, C^{10}-*H*$_2$, C^{11}-*H*$_a$, C^{12}-*H*$_a$, C^{13}-*H*$_2$), 1.79–1.68 (2H, m, C^{11}-*H*$_b$, C^{12}-*H*$_b$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.4 (C^2), 131.1 (C^8), 128.1 (C^3), 122.6 (C^6), 120.6 (C^5), 118.3 (C^{15}), 109.1 (C^4), 107.1 (C^7), 53.7 (C^9), 29.0 (C^{10}, C^{13}), 27.3 (C^1), 25.2 (C^{11}, C^{12}), 23.8 (C^{14}); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 202.09759, $C_{11}H_{11}N_3O$ requires [M+H]$^+$ 202.09749); LRMS *m/z* (ESI$^+$) 175 ([M–N]$^+$, 100%); RP-HPLC (TFA), retention time = 8.7 min, 97.8%.

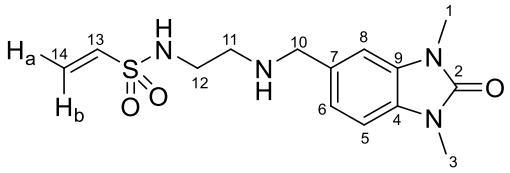
5-(2-Aminoethyl)-1-cyclopentyl-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (144)



(1-cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)acetonitrile (**143**) (50.5 mg, 0.20 mmol, 1.0 eq.) in EtOH (2.0 mL) was degassed under nitrogen for 5 minutes. Palladium on carbon (10% wt, 21.3 mg, 0.02 mmol, 0.1 eq.) was added, followed by 10% HCl in EtOH (*v/v*) (0.25 mL). The reaction flask was purged with 3 balloons of hydrogen before inserting another balloon of hydrogen. The reaction mixture was stirred at room temperature for 16 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH (100 mL). The solvent was removed in vacuo and the crude material was purified using an Isolute® SCX-2 amine capture-release column (equilibrated with MeOH (10 mL), loaded in MeOH (2 mL), washed with MeOH (10 mL) and eluted with 2 M

NH_3 in MeOH (10 mL)) to afford the title compound as a colourless oil (38.7 mg, 0.15 mmol, 75% yield): R_f 0.19 (EtOAc/EtOH 1:1); $\bar{\nu}_{max}$/cm^{-1} 3361 (N-H, w), 2950 (w), 2871 (w), 1687 (C=O, s), 1603 (w), 1503 (m), 1477 (w), 1433 (w), 1395 (m), 1329 (w), 1167 (w), 1048 (w), 804 (w), 747 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.98 (1H, d, *J* 8.0, C^4-*H*), 6.89 (1H, dd, *J* 8.0, 1.6, C^5-*H*), 6.82 (1H, d, *J* 1.6, C^7-*H*), 4.85 (1H, tt, *J* 8.9, 8.9, C^9-*H*), 3.39 (3H, s, C^1-*H$_3$*), 2.98 (2H, t, *J* 6.8, C^{15}-*H$_2$*), 2.78 (1H, t, *J* 6.8, C^{14}-*H$_2$*), 2.13–1.88 (6H, m, C^{10}-*H$_2$*, C^{11}-*H$_a$*, C^{12}-*H$_a$*, C^{13}-*H$_2$*), 1.77–1.66 (2H, m, C^{11}-*H$_b$*, C^{12}-*H$_b$*), 1.18 (2H, s, N-*H$_2$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.4 (C^2), 132.8 (C^6), 130.7 (C^6), 126.6 (C^3), 121.3 (C^5), 108.8 (C^4), 108.0 (C^7), 55.6 (C^9), 44.1 (C^{15}), 40.2 (C^{14}), 28.9 (C^{10}, C^{13}), 27.2 (C^1), 25.2 (C^{11}, C^{12}); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 260.17580, $C_{15}H_{21}N_3O$ requires [M+H]$^+$ 260.17574); LRMS *m/z* (ESI$^+$) 260 ([M+H]$^+$, 100%); RP-HPLC (TFA), retention time = 6.2 min, 98.2%.

_N_-(2-{[(1,3-Dimethyl-2-oxo-2,3-dihydro-1_H_-benzimidazol-5-yl)methyl]amino}ethyl)ethenesulfonamide (146)

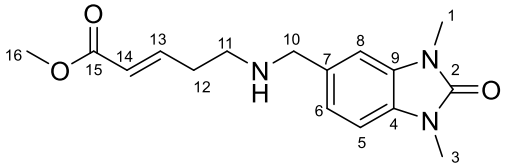


To a solution of *tert*-butyl (2-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}ethyl)carbamate (**149**) (100.9 mg, 0.3 mmol, 1.0 eq.) in CH_2Cl_2 (6.0 mL) was added triisopropylsilane (6 μL, 4.8 mg, 0.03 mmol, 0.1 eq.) and TFA (1.5 mL). The reaction mixture was stirred at room temperature for 30 minutes. After this time, the solvent was removed *in vacuo* and residual TFA was azeotroped with toluene to afford 5-{[(2-aminoethyl)amino]methyl}-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one as a TFA salt, which was used in the next step without further purification.

To a solution of 5-{[(2-aminoethyl)amino]methyl}-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one TFA salt (0.3 mmol, 1.0 eq.) in dry CH_2Cl_2 (4.0 mL) under argon

atmosphere cooled to 0 °C was added triethylamine (0.3 mL, 2.1 mmol, 7.0 eq.). The reaction mixture was stirred at 0 °C and a solution of 2-chloroethanesulfonyl chloride (63 µL, 97.8 mg, 0.6 mmol, 2.0 eq.) in dry CH_2Cl_2 (4.0 mL) was added dropwise slowly. The reaction mixture was stirred at 0 °C for 4 hours. After this time, the reaction was quenched with a saturated aqueous solution of $NaHCO_3$ (10 mL). The aqueous layer was extracted with 10% IPA in $CHCl_3$ (3 × 30 mL). The organic components were then dried over Na_2SO_4, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65–100% EtOAc in petroleum ether with 2% triethylamine gradient to 0–10% EtOH in EtOAc with 2% triethylamine gradient) to afford the title compound as a colourless oil (19.6 g, 0.60 mmol, 30% yield): R_f 0.32 (EtOAc with 2% triethylamine); $\bar{\nu}_{max}$/cm^{-1} 3084 (w), 2941 (w), 1694 (C=O, s), 1558 (w), 1510 (w), 1450 (w), 1398 (w), 1336 (S=O, w), 1188 (m), 1149 (S=O, m), 761 (w), 728 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.05 (1H, dd, *J* 7.9, 1.6, C^6-*H*), 6.99 (1H, d, *J* 1.6, C^8-*H*), 6.93 (1H, d, *J* 7.9, C^5-*H*), 6.45 (1H, dd, *J* 16.5, 9.8, C^{13}-*H*), 6.31 (1H, d, *J* 16.5, C^{14}-H_b), 6.05 (1H, d, *J* 9.8, C^{14}-H_a), 4.38 (2H, s, C^{10}-H_2), 3.42 (6H, s, C^1-H_3, C^3-H_3), 3.35–3.31 (2H, m, C^{11}-H_2), 3.29–3.23 (2H, m, C^{12}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.0 (C^2), 133.7 (C^{13}), 130.9 (C^9), 130.5 (C^4), 128.7 (C^7), 128.5 (C^{14}), 121.8 (C^6), 107.61 (C^8), 107.56 (C^5), 53.7 (C^{10}), 46.5 (C^{11}), 39.1 (C^{12}), 27.42, 27.40 (C^1, C^3); Compound could not be detected by HRMS ESI, EI, or CI; LRMS *m/z* (ESI$^+$) 175 ([M–$C_4H_9N_2O_2S$]$^+$, 100%); RP-HPLC (TFA), retention time = 7.8 min, 97.6%.

Methyl (2*E*)-5-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}pent-2-enoate (147)



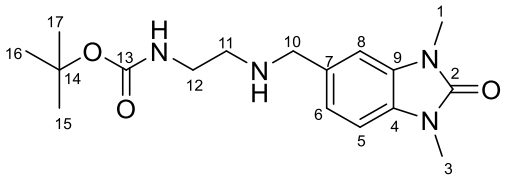
To a solution of methyl (2*E*)-5-[(*tert*-butoxycarbonyl)amino]pent-2-enoate (**155**) (125.0 mg, 0.55 mmol, 1.0 eq.) in CH_2Cl_2 (10.0 mL) was added triisopropylsilane (11 µL,

8.7 mg, 0.06 mmol, 0.1 eq.) and TFA (2.5 mL). The reaction mixture was stirred at room temperature for 1 hour. After this time, the solvent was removed *in vacuo* and residual TFA was azeotroped with toluene to afford methyl (2*E*)-5-aminopent-2-enoate as a TFA salt, which was used in the next step without further purification.

To a solution of (2*E*)-5-aminopent-2-enoate TFA salt (0.10 mmol, 1.5 eq.) and triethylamine (43 μL, 31.4 mg, 0.31 mmol, 1 eq.) in dry 1,2-dichloroethane (4 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (60.3 mg, 0.31 mmol, 1.0 eq.) and triethylamine (87 μL, 62.7 mg, 0.62 mmol, 2.0 eq.) were added. The reaction mixture was warmed to room temperature and stirred for 3 hours. After this time, sodium triacetoxyborohydride (328.3 g, 1.55 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 16 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solution was concentrated *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (15 mL) and $CHCl_3$ (15 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (4 × 30 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (5–30% EtOH in EtOAc gradient) followed by a second flash column chromatography on silica gel (2–7% EtOH in $CHCl_3$ gradient) to afford the title compound as a colourless oil (14.3 mg, 0.05 mmol, 71% yield): R_f 0.17 (EtOAc/EtOH 4:1); $\bar{\nu}_{max}$/cm^{-1} 3159 (w), 2942 (w), 1687 (C=O, s), 1508 (m), 1447 (w), 1397 (w), 1333 (w), 1263 (m), 1155 (m), 1137 (m), 1091 (w), 1021 (w), 969 (w), 808 (w), 696 (m), 620 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.02 (1H, dd, *J* 7.8, 1.5, C^6-*H*), 6.97 (1H, d, *J* 1.5, C^8-*H*), 6.96 (1H, dt, *J* 15.6, 7.0, C^{13}-*H*), 6.90 (1H, d, *J* 8.0, C^5-*H*), 5.89 (1H, dt, *J* 15.6, 1.6, C^{14}-*H*), 3.83 (2H, s, C^{10}-*H*$_2$), 3.72 (3H, s, C^{15}-*H*$_3$), 3.42 (3H, s, C^1-*H*$_3$), 3.41 (3H, s, C^3-*H*$_3$), 2.79 (2H, t, *J* 6.9, C^{11}-*H*$_2$), 2.43 (2H, t, *J* 7.0, 6.9, 1.6, C^{12}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.0 (C^{15}), 155.0 (C^2), 147.2 (C^{13}), 133.6 (C^7), 130.4 (C^9), 129.3 (C^4), 122.5 (C^{14}), 121.3 (C^6), 107.4 (C^8), 107.1 (C^5), 54.2 (C^{10}), 51.6 (C^{16}), 47.6

(C^{11}), 33.0 (C^{12}), 27.36, 27.34 (C^{1}, C^{3}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 304.16575, $C_{16}H_{21}N_3O_3$ requires $[M+H]^+$ 304,16557); LRMS *m/z* (ESI$^+$) 175 ($[M–C_6H_{10}NO_2]^+$, 100%); RP-HPLC (TFA), retention time = 5.3 min, 98.7%.

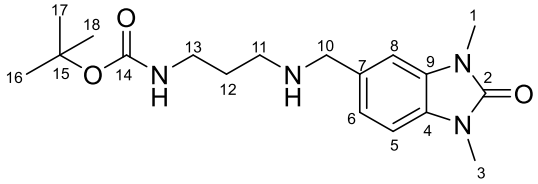
tert-Butyl (2-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}ethyl)carbamate (149)



To a solution of 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (437.5 mg, 2.3 mmol, 1.0 eq.) in dry 1,2-dichloroethane (25 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and *N*-Boc-ethylenediamine (728 µL, 737.0 mg, 4.6 mmol, 2.0 eq.) and triethylamine (641 µL, 465.5 mg, 4.6 mmol, 2.0 eq.) were added. The reaction mixture was warmed to room temperature and stirred for 6 hours. After this time, sodium triacetoxyborohydride (2.44 g, 11.5 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 17 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH and H_2O. The solution was concentrated *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (15 mL) and $CHCl_3$ (15 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (4 × 30 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (80–100% EtOAc in petroleum ether with 2% triethylamine gradient to 0–20% EtOH in EtOAc with 2% triethylamine gradient), followed by an Isolute® SCX-2 amine capture-release column (equilibrated with MeOH (40 mL), loaded in MeOH (6 mL), washed with MeOH (40 mL) and eluted with 2 M NH_3 in MeOH (40 mL)) to afford the title compound as a colourless oil (613.5 mg, 1.84 mmol, 80% yield): R_f 0.10 (EtOAc/EtOH 4:1); $\bar{\nu}_{max}$/cm^{-1} 2975 (w), 2932 (w), 1688 (C=O, s), 1509 (m), 1454 (w),

1396 (w), 1365 (w), 1271 (w), 1250 (w), 1171 (m), 865 (w), 805 (w), 744 (w), 727 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.03 (1H, dd, *J* 7.9, 1.5, C^{6}-*H*), 6.97 (1H, d, *J* 1.5, C^{8}-*H*), 6.90 (1H, d, *J* 7.9, C^{5}-*H*), 4.89 (1H, br s, N-*H*), 3.83 (2H, s, C^{10}-H_2), 3.42 (3H, s, C^{1}-H_3), 3.41 (3H, s, C^{3}-H_3), 3.25 (2H, dt, *J* 5.9, 5.9, C^{12}-H_2), 2.76 (2H, t, *J* 5.9, C^{11}-H_2), 1.44 (9H, s, C^{15}-H_3, C^{16}-H_3, C^{17}-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.2 (C^{13}), 155.0 (C^{2}), 133.8 (C^{7}), 130.4 (C^{9}), 129.3 (C^{4}), 121.3 (C^{6}), 107.4 (C^{8}), 107.2 (C^{5}), 79.4 (C^{14}), 53.9 (C^{10}), 48.7 (C^{11}), 40.4 (C^{12}), 28.6 (C^{15}, C^{16}, C^{17}), 27.4, 27.3 (C^{1}, C^{3}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 335.20716, $C_{17}H_{26}N_4O_3$ requires $[M+H]^+$ 335.20777); LRMS *m/z* (ESI$^+$) 175 ($[M–C_7H_{15}N_2O_2]^+$, 100%); RP-HPLC (TFA), retention time = 6.2 min, 89.1%.

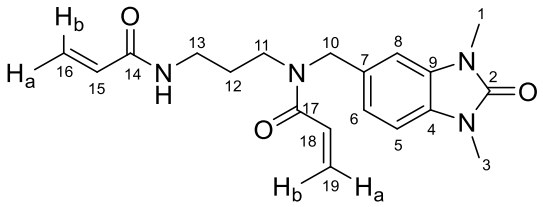
tert-Butyl (3-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}propyl)carbamate (150)



To a solution of 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (437.7 mg, 2.3 mmol, 1.0 eq.) in dry 1,2-dichloroethane (25 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and *N*-Boc-1,3-propanediamine (801.5 mg, 4.6 mmol, 2.0 eq.) and triethylamine (641 µL, 465.5 mg, 4.6 mmol, 2.0 eq.) was added. The reaction mixture was warmed to room temperature and stirred for 3 hours. After this time, sodium triacetoxyborohydride (2.44 g, 11.5 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 17 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH and H_2O. The solution was concentrated *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (15 mL) and $CHCl_3$ (15 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (4 × 30 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was

purified by flash column chromatography on silica gel (80–100% EtOAc in petroleum ether with 2% triethylamine gradient to 0–20% EtOH in EtOAc with 2% triethylamine gradient) to afford the title compound as a colourless oil (661.3 mg, 1.90 mmol, 83% yield): R_f 0.13 (petroleum ether/EtOAc 1:4); $\bar{\nu}_{max}$/cm^{-1} 2975 (w), 2933 (w), 1690 (C=O, s), 1509 (m), 1453 (w), 1396 (w), 1365 (w), 1272 (w), 1251 (w), 1171 (m), 863 (w), 808 (w), 745 (w), 728 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.04 (1H, dd, *J* 7.9, 1.5, C^6-*H*), 6.98 (1H, d, *J* 1.5, C^8-*H*), 6.89 (1H, d, *J* 7.9, C^5-*H*), 5.08 (1H, br s, N-*H*), 3.81 (2H, s, C^{10}-*H*$_2$), 3.41 (3H, s, C^1-*H*$_3$), 3.40 (3H, s, C^3-*H*$_3$), 3.21 (2H, dt, *J* 6.5, 6.3, C^{13}-*H*$_2$), 2.70 (2H, t, *J* 6.6, C^{11}-*H*$_2$), 1.68 (2H, tt, *J* 6.6, 6.5, C^{12}-*H*$_2$), 1.42 (9H, s, C^{15}-*H*$_3$, C^{16}-*H*$_3$, C^{17}-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.3 (C^{14}), 155.0 (C^2), 133.6 (C^7), 130.3 (C^9), 129.3 (C^4), 121.4 (C^6), 107.5 (C^8), 107.1 (C^5), 79.2 (C^{15}), 54.4 (C^{10}), 47.2 (C^{11}), 39.2 (C^{13}), 30.0 (C^{12}), 28.6 (C^{16}, C^{17}, C^{18}), 27.34, 27.32 (C^1, C^3); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 349.22339, $C_{18}H_{28}N_4O_3$ requires [M+H]$^+$ 349.22342); LRMS *m/z* (ESI$^+$) 175 ([M–$C_8H_{17}N_2O_2$]$^+$, 100%); RP-HPLC (TFA), retention time = 6.3 min, 96.2%.

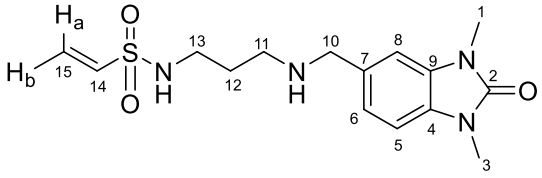
_N_-[(1,3-Dimethyl-2-oxo-2,3-dihydro-1_H_-benzimidazol-5-yl)methyl]-_N_-{3-[(prop-2-enoyl)amino]propyl}prop-2-enamide (151)



To a solution of *tert*-butyl (3-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}propyl)carbamate (**150**) (198.8 mg, 0.57 mmol, 1.0 eq.) in CH_2Cl_2 (12.0 mL) was added triisopropylsilane (12 μL, 9.0 mg, 0.06 mmol, 0.1 eq.) and TFA (3.0 mL). The reaction mixture was stirred at room temperature for 1 hour. After this time, the solvent was removed *in vacuo* and residual TFA was azeotroped with toluene to afford 5-{[(3-aminopropyl)amino]methyl}-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one as a TFA salt, which was used in the next step without further purification.

To a solution of 5-{[(3-aminopropyl)amino]methyl}-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one TFA salt (0.57 mmol, 1.0 eq.) in dry CH_2Cl_2 (8.0 mL) under argon atmosphere cooled to 0 °C was added *N*,*N*-diisopropylethylamine (99 µL, 73.7 mg, 0.57 mmol, 1.0 eq.). The reaction mixture was stirred at 0 °C and a solution of acryloyl chloride (46 µL, 51.6 mg, 0.57 mmol, 1.0 eq.) and 4-dimethylaminopyridine (1 drop, cat.) in dry CH_2Cl_2 (8.0 mL) was added dropwise slowly. The reaction mixture was stirred at 0 °C for 2 hours. After this time, the reaction was quenched with a saturated aqueous solution of $NaHCO_3$ (10 mL). The aqueous layer was extracted with 10% IPA in $CHCl_3$ (3 × 30 mL). The organic components were then dried over Na_2SO_4, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (0–50% EtOH in EtOAc gradient) followed by a second flash column chromatography on silica gel (50–60% acetone in petroleum ether gradient) to afford the title compound as a colourless oil (51.5 g, 0.14 mmol, 25% yield): R_f 0.37 (EtOAc/EtOH 4:1); $\bar{\nu}_{max}$/cm^{-1} 3071 (w), 2935 (w), 1694 (C=O, s), 1661 (C=O, s), 1607 (m), 1552 (w), 1509 (m), 1451 (m), 1397 (w), 1243 (w), 982 (w), 797 (w), 727 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.07 (1H, t, *J* 6.0, N-*H*), 6.94 (1H, d, *J* 8.1, C^5-*H*), 6.90 (1H, dd, *J* 8.1, 1.6, C^6-*H*), 6.74 (1H, d, *J* 1.6, C^8-*H*), 6.61 (1H, dd, *J* 16.7, 10.3, C^{18}-*H*), 6.44 (1H, dd, *J* 16.7, 2.0, C^{19}-H_b), 6.29 (1H, dd, *J* 17.1, 1.7, C^{16}-H_b), 6.17 (1H, dd, *J* 17.1, 10.1, C^{15}-*H*), 5.75 (1H, dd, *J* 10.3, 2.0, C^{19}-H_a), 5.63 (1H, dd, *J* 10.1, 1.7, C^{16}-H_a), 4.64 (2H, s, C^{10}-H_2), 3.54 (2H, t, *J* 6.2, C^{11}-H_2), 3.42 (3H, s, C^3-H_3), 3.40 (3H, s, C^1-H_3), 3.31 (2H, dt, *J* 6.2, 6.2, C^{13}-H_2), 1.74 (2H, tt, *J* 6.2, 6.2, C^{12}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.0 (C^{17}), 165.8 (C^{14}), 154.9 (C^2), 131.6 (C^{15}), 130.9 (C^9), 130.0 (C^4), 129.6 (C^7, C^{19}), 127.5 (C^{18}), 125.9 (C^{16}), 119.5 (C^6), 107.7 (C^5), 105.4 (C^8), 51.0 (C^{10}), 42.9 (C^{11}), 35.9 (C^{13}), 27.43, 27.41 (C^1, C^3), 27.1 (C^{12}); HRMS *m/z* (ESI$^+$) (Found [M+Na]$^+$ 379.17387, $C_{19}H_{24}N_4O_3$ requires [M+Na]$^+$ 379.17406); LRMS *m/z* (ESI$^+$) 379 ([M+Na]$^+$, 100%); RP-HPLC (TFA), retention time = 6.1 min, 96.5%.

_N_-(3-{[(1,3-Dimethyl-2-oxo-2,3-dihydro-1_H_-benzimidazol-5-yl)methyl]amino}propyl)ethenesulfonamide (152)

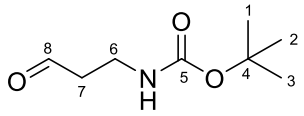


To a solution of *tert*-butyl (3-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}propyl)carbamate (**150**) (383.0 mg, 1.1 mmol, 1.0 eq.) in CH_2Cl_2 (16.0 mL) was added triisopropylsilane (23 μL, 17.4 mg, 0.11 mmol, 0.1 eq.) and TFA (4.0 mL). The reaction mixture was stirred at room temperature for 1 hour. After this time, the solvent was removed *in vacuo* and residual TFA was azeotroped with toluene to afford 5-{[(3-aminopropyl)amino]methyl}-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one as a TFA salt, which was used in the next step without further purification.

To a solution of 5-{[(3-aminopropyl)amino]methyl}-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one TFA salt (1.1 mmol, 1.0 eq.) in dry CH_2Cl_2 (16.0 mL) under argon atmosphere cooled to 0 °C was added triethylamine (0.77 mL, 556.5 mg 5.5 mmol, 5.0 eq.). The reaction mixture was stirred at 0 °C and a solution of 2-chloroethanesulfonyl chloride (115 μL, 179.3 mg, 1.1 mmol, 1.0 eq.) in dry CH_2Cl_2 (16.0 mL) was added dropwise slowly. The reaction mixture was stirred at 0 °C for 2 hours. After this time, the reaction was quenched with a saturated aqueous solution of $NaHCO_3$ (10 mL). The aqueous layer was extracted with 10% IPA in $CHCl_3$ (3 × 30 mL). The organic components were then dried over Na_2SO_4 and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (70–100% EtOAc in petroleum ether gradient) followed by a second flash column chromatography on silica gel (5–20% acetone in Et_2O gradient) to afford the title compound as a colourless oil (30.2 g, 0.09 mmol, 8% yield): R_f 0.38 (EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3093 (w), 2943 (w), 1693 (C=O, s), 1558 (w), 1509 (w), 1449 (w), 1397 (w), 1336 (S=O, w), 1186 (m), 1148 (S=O, s), 783 (w), 726 (w); ^{1}H NMR

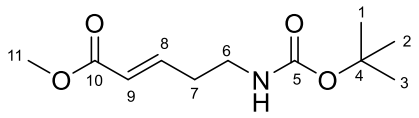
(500 MHz, $CDCl_3$) δ 7.21 (1H, t, *J* 6.4, SO_2N-*H*), 7.05 (1H, dd, *J* 7.9, 1.6, C^6-*H*), 6.98 (1H, d, *J* 1.6, C^8-*H*), 6.92 (1H, d, *J* 7.9, C^5-*H*), 6.41 (1H, dd, *J* 16.5, 9.9, C^{14}-*H*), 6.27 (1H, d, *J* 16.5, C^{15}-H_b), 6.00 (1H, d, *J* 9.9, C^{15}-H_a), 4.35 (2H, s, C^{10}-H_2), 3.42 (6H, s, C^1-H_3, C^3-H_3), 3.31 (2H, dt, *J* 6.3, 6.1, C^{11}-H_2), 3.18 (2H, t, *J* 6.3, C^{13}-H_2), 1.54 (2H, tt, *J* 6.1, 5.9, C^{12}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.8 (C^2), 134.0 (C^{14}), 130.7 (C^9), 130.4 (C^4), 129.0 (C^7), 128.0 (C^{15}), 121.9 (C^6), 107.8 (C^8), 107.4 (C^5), 53.1 (C^{10}), 45.2 (C^{13}), 36.0 (C^{11}), 27.43, 27.41 (C^1, C^3), 27.3 (C^{12}); Compound could not be detected by HRMS ESI, EI, or CI; LRMS *m/z* (ESI$^+$) 175 ([M–$C_5H_{11}N_2O_2S$]$^+$, 100%); RP-HPLC (TFA), retention time = 7.9 min, 99.2%.

***tert*-Butyl (3-oxopropyl)carbamate (154)**[119]



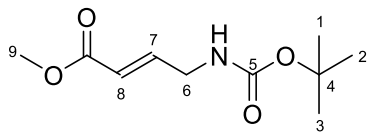
To a solution of 3-(Boc-amino)-1-propanol (1.05 g, 6.0 mmol, 1.0 eq.) dissolved in CH_2Cl_2 (35 mL) cooled to 0 °C was added Dess-Martin periodinane (3.31 g, 7.8 mmol, 1.3 eq.). The reaction mixture was stirred at room temperature for 3 hours. After this time, the reaction was quenched with aqueous $Na_2S_2O_3$ solution (10% *w/v*, 50 mL). The aqueous component was extracted with CH_2Cl_2 (5 × 30 mL). The organic components were washed with a saturated aqueous solution of $NaHCO_3$ (20 mL), followed by brine (10 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel using the Biotage system (20–40% EtOAc in petroleum ether gradient) to afford the title compound as a colourless oil (0.52 g, 3.02 mmol, 50% yield): R_f 0.45 (petroleum ether/EtOAc 1:1); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.81 (1H, t, *J* 0.9, C^8-*H*), 4.87 (1H, br s, N-*H*), 3.92 (2H, dt, *J* 6.0, 6.0, C^7-H_2), 2.70 (2H, t, *J* 6.0, C^6-H_2), 1.43 (9H, s, C^1-H_3, C^2-H_3, C^3-H_3); HRMS *m/z* (ESI$^-$) (Found [M–H]$^-$ 172.09746, $C_8H_{15}NO_3$ requires [[M–H]$^-$ 172.09792); Could not be detected by LRMS (ESI); No HPLC data obtained as compound is not UV active. ^{1}H NMR data is in agreement with the literature.[132]

Methyl (2*E*)-5-[(*tert*-butoxycarbonyl)amino]pent-2-enoate (155)



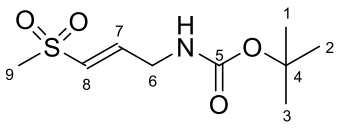
To a solution of lithium chloride (76.8 mg, 1.8 mmol, 1.2 eq.) suspended in dry acetonitrile (10 mL) under argon atmosphere was added trimethylphosphonoacetate (243 µL, 273.2 mg, 1.8 mmol, 1.2 eq.) and 1,8-diazabicyclo[5.4.0]undec-7-ene (224 µL, 228.4 mg, 1.5 mmol, 1.0 eq.). The reaction mixture was cooled to 0 °C and *tert*-butyl (3-oxopropyl)carbamate (**154**) (260.1 mg, 1.5 mmol, 1.0 eq.) was added. The solution was stirred at room temperature for 1 hour. After this time, the reaction mixture was concentrated *in vacuo* to half the original volume then diluted with EtOAc (15 mL) and H_2O (15 mL). The aqueous component was extracted with EtOAc (2 × 20 mL). The organic components were washed with brine (10 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel using the Biotage system (0–20% EtOAc in petroleum ether gradient) to afford the title compound as a colourless oil (141.1 mg, 0.63 mmol, 41% yield): R_f 0.33 (petroleum ether/EtOAc 9:1); $\bar{\nu}_{max}$/cm^{-1} 2978 (w), 2935 (w), 1695 (C=O, s), 1659 (C=O, s), 1521 (m), 1437 (w), 1392 (w), 1366 (w), 1272 (m), 1251 (m), 1169 (s), 1042 (w), 976 (w), 855 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.90 (1H, dt, *J* 15.7, 7.0, C^8-*H*), 5.88 (1H, dt, *J* 15.7, 1.6, C^9-*H*), 4.59 (1H, br s, N-*H*), 3.72 (3H, s, C^{11}-*H*$_3$), 3.26 (2H, dt, *J* 6.7, 6.5, C^6-*H*$_2$), 2.40 (2H, ddt, *J* 7.0, 6.7, 1.6, C^7-*H*$_2$), 1.43 (9H, s, C^1-*H*$_3$, C^2-*H*$_3$, C^3-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.8 (C^{10}), 155.9 (C^5), 145.8 (C^8), 123.1 (C^9), 79.6 (C^4), 51.7 (C^{11}), 39.2 (C^6), 32.9 (C^7), 28.5 (C^1, C^2, C^3); HRMS *m/z* (ESI$^+$) (Found [M+Na]$^+$ 252.12032, $C_{11}H_{19}NO_4$ requires [M+Na]$^+$ 252.12063); LRMS *m/z* (ESI$^+$) 252 ([M+Na]$^+$, 100%); No HPLC data obtained as compound is not UV active.

Methyl (2*E*)-4-[(*tert*-butoxycarbonyl)amino]but-2-enoate (158)[133]



To a solution of lithium chloride (76.8 mg, 1.8 mmol, 1.2 eq.) suspended in dry acetonitrile (18 mL) under argon atmosphere was added trimethylphosphonoacetate (243 μL, 273.2 mg, 1.8 mmol, 1.2 eq.) and 1,8-diazabicyclo[5.4.0]undec-7-ene (224 μL, 228.4 mg, 1.5 mmol, 1.0 eq.). To the reaction mixture was added *N*-Boc-2-aminoacetaldehyde (238.9 mg, 1.5 mmol, 1.0 eq.). The solution was stirred at room temperature for 1.5 hours. After this time, the reaction mixture was concentrated *in vacuo* to half the original volume then diluted with EtOAc (15 mL) and H_2O (15 mL). The aqueous component was extracted with EtOAc (2 × 20 mL). The organic components were washed with brine (10 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel using the Biotage system (20–100% EtOAc in petroleum ether gradient) to afford the title compound as a colourless oil (160.4 mg, 0.75 mmol, 50% yield): R_f 0.20 (petroleum ether/EtOAc 4:1); ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.91 (1H, dt, *J* 15.7, 5.3, C^7-*H*), 5.94 (1H, dt, *J* 15.7, 1.9, C^8-*H*), 4.68 (1H, br s, N-*H*), 3.92 (2H, dd, *J* 5.7, 5.3, C^6-*H*), 3.74 (3H, s, C^9-H_3), 1.45 (9H, s, C^1-H_3, C^2-H_3, C^3-H_3); LRMS *m/z* (ESI$^+$) 238 ([M+Na]$^+$, 100%); No HPLC data obtained as compound is not UV active. ^{1}H NMR data is in agreement with the literature.[118]

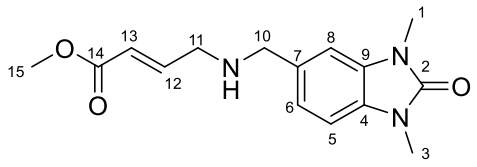
***tert*-Butyl [(2*E*)-3-(methanesulfonyl)prop-2-en-1-yl]carbamate (159)**[118]



To a solution of sodium hydride (60% dispersion in mineral oil, 60.5 mg, 1.5 mmol, 1.0 eq.) suspended in dry tetrahydrofuran (1.0 mL) under argon atmosphere cooled to 0 °C was

added dropwise diethyl (methylsulfonylmethyl)phosphonate (345.8 mg, 1.5 mmol, 1.0 eq.) in dry tetrahydrofuran (2.0 mL). The solution was stirred at room temperature for 20 minutes. A solution of *N*-Boc-2-aminoacetaldehyde (238.1 mg, 1.5 mmol, 1.0 eq.) dissolved in dry tetrahydrofuran (1.5 mL) was then added dropwise and the mixture was warmed to room temperature and stirred for 4 hours. After this time, the reaction mixture was concentrated *in vacuo*. The residue was dissolved in EtOAc (15 mL) and H_2O (15 mL) and the aqueous component was extracted with EtOAc (2 × 20 mL). The organic components were washed brine (10 mL), dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel using the Biotage system (20–65% EtOAc in petroleum ether gradient) to afford the title compound as a colourless oil (92.8 mg, 0.39 mmol, 26% yield): R_f 0.18 (petroleum ether/EtOAc 2:3); $\bar{\nu}_{max}$/cm^{-1} 2979 (w), 2929 (w), 1701 (C=O, s), 1519 (m), 1368 (w), 1307 (S=O, m), 1282 (m), 1252 (w), 1167 (m), 1133 (S=O, s), 970 (C=C, w), 749 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.90 (1H, dt, *J* 15.2, 4.9, C^7-*H*), 6.49 (1H, dt, *J* 15.2, 2.0, C^8-*H*), 4.90 (1H, br s, N-*H*), 3.97 (2H, dd, *J* 5.4, 4.9, C^6-*H*$_2$), 2.93 (3H, s, C^9-*H*$_3$), 1.44 (9H, s, C^1-*H*$_3$, C^2-*H*$_3$, C^3-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.6 (C^5), 145.2 (C^7), 129.8 (C^8), 80.4 (C^4), 42.9 (C^9), 40.8 (C^6), 28.4 (C^1, C^2, C^3); LRMS *m/z* (ESI$^+$) 234 ([M–H]$^-$, 100%); No HPLC data obtained as compound is not UV active. ^{1}H NMR data and mass spectrum is in agreement with the literature.[134]

Methyl (2*E*)-4-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}but-2-enoate (160)



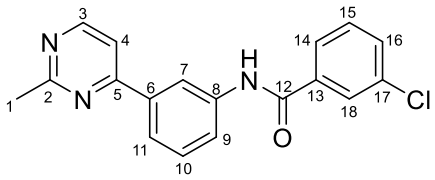
To a solution of methyl (2*E*)-4-[(*tert*-butoxycarbonyl)amino]but-2-enoate (**158**) (21.3 mg, 0.10 mmol, 1.0 eq.) in CH_2Cl_2 (2.0 mL) was added triisopropylsilane (2 µL, 1.6 mg, 0.01 mmol, 0.1 eq.) and TFA (0.5 mL). The reaction mixture was stirred at room

temperature for 1 hour. After this time, the solvent was removed *in vacuo* and residual TFA was azeotroped with toluene to afford methyl (2*E*)-4-aminobut-2-enoate as a TFA salt, which was used in the next step without further purification.

To a solution of methyl (2*E*)-4-aminobut-2-enoate TFA salt (0.10 mmol, 1.5 eq.) and triethylamine (7 µL, 5.0 mg, 0.05 mmol, 0.7 eq.) in dry 1,2-dichloroethane (4 mL) under argon atmosphere was added 3 Å molecular sieves. The solution was cooled to 0 °C and 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (**56**) (13.1 mg, 0.07 mmol, 1.0 eq.) and triethylamine (13 µL, 9.2 mg, 0.09 mmol, 1.3 eq.) were added. The reaction mixture was warmed to room temperature and stirred for 3 hours. After this time, sodium triacetoxyborohydride (74.2 g, 0.35 mmol, 5.0 eq.) was added. The reaction mixture was then stirred at room temperature for 16 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOH. The solution was concentrated *in vacuo* and the residue was dissolved in a saturated aqueous solution of $NaHCO_3$ (15 mL) and $CHCl_3$ (15 mL). The aqueous layer was then extracted with 10% IPA in $CHCl_3$ (4 × 30 mL). The organic components were then dried over Na_2SO_4, filtered, and the solvent was removed *in vacuo*. The crude material was purified by flash column chromatography on silica gel (5–20% EtOH in EtOAc gradient) followed by a second flash column chromatography on silica gel (2–5% EtOH in $CHCl_3$ gradient) to afford the title compound as a colourless oil (14.3 mg, 0.05 mmol, 71% yield): R_f 0.23 (EtOAc/EtOH 4:1); $\bar{\nu}_{max}$/cm^{-1} 3159 (w), 2937 (w), 1684 (C=O, s), 1508 (m), 1447 (m), 1396 (w), 1332 (w), 1263 (m), 1154 (m), 1137 (m), 1090 (w), 1022 (w), 968 (w), 929 (w), 801 (w), 706 (m), 620 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.034 (1H, dd, *J* 8.0, 1.4, C^6-*H*), 7.032 (1H, dt, *J* 15.8, 5.5, C^{12}-*H*), 6.99 (1H, d, *J* 1.4, C^8-*H*), 6.90 (1H, d, *J* 8.0, C^5-*H*), 6.03 (1H, dt, *J* 15.8, 1.8, C^{13}-*H*), 3.85 (2H, s, C^{10}-*H$_2$*), 3.75 (3H, s, C^{15}-*H$_3$*), 3.45 (2H, dd, *J* 5.5, 1.8, C^{11}-*H$_2$*), 3.42 (3H, s, C^1-*H$_3$*), 3.41 (3H, s, C^3-*H$_3$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.0 (C^{14}), 155.0 (C^2), 147.1 (C^{12}), 133.3 (C^7), 130.5 (C^9), 129.4 (C^4), 121.4 (C^{13}), 121.3 (C^6), 107.4 (C^8), 107.2 (C^5), 53.7 (C^{10}), 51.7 (C^{15}), 49.7 (C^{11}), 27.4 (C^1, C^3); HRMS *m/z* (ESI$^+$) (Found

$[M+H]^+$ 290.14999, $C_{15}H_{19}N_3O_3$ requires $[M+H]^+$ 290.14992); LRMS *m/z* (ESI^+) 175 ($[M-C_5H_8NO_2]^+$, 100%); RP-HPLC (TFA), retention time = 5.2 min, 93.7%.

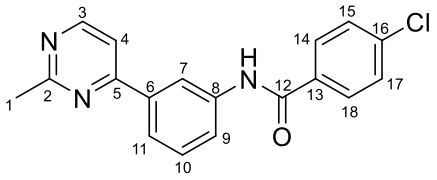
3-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (163)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (56.0 mg, 0.30 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (0.75 mL), was added 3-chlorobenzoyl chloride (46 µL, 0.36 mmol, 1.2 eq.) and triethylamine (50 µL, 0.36 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 16 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (86.2 mg, 0.27 mmol, 88% yield): R_f 0.37 (petroleum ether/EtOAc 1:3); m.p. 154–156 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3294 (N-H, w), 1652 (C=O, m), 1567 (s), 1529 (s), 1489 (w), 1432 (m), 1358 (w), 1303 (m), 1260 (w), 1079 (w), 886 (w), 836 (w), 795 (m), 714 (m), 684 (s), 626 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.66 (1H, d, *J* 5.4, C^4-*H*), 8.33 (1H, dd, *J* 2.1, 1.6, C^7-*H*), 8.11 (1H, br s, N-*H*), 7.88 (1H, ddd, *J* 8.0, 2.1, 1.2, C^9-*H*), 7.87 (1H, dd, *J* 2.1, 1.7, C^{18}-*H*), 7.82 (1H, ddd, *J* 7.9, 1.6, 1.2, C^{11}-*H*), 7.75 (1H, ddd, *J* 7.8, 1.7, 1.2, C^{14}-*H*), 7.52 (1H, ddd, *J* 8.0, 2.1, 1.2, C^{16}-*H*), 7.52 (1H, d, *J* 5.4, C^3-*H*), 7.49 (1H, dd, *J* 8.0, 7.9, C^{10}-*H*), 7.42 (1H, dd, *J* 8.0, 7.8, C^{15}-*H*), 2.78 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.6 (C^2), 164.7 (C^{12}), 163.4 (C^5), 157.8 (C^4), 138.5 (C^8), 138.0 (C^6), 136.6 (C^{13}), 135.2 (C^{17}), 132.2 (C^{16}), 130.3 (C^{15}), 129.9 (C^{10}), 127.6 (C^{14}), 123.6 (C^{11}), 122.8

(C^9), 119.0 (C^7), 114.2 (C^3), 26.4 (C^1); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 324.0897, $C_{18}H_{14}{}^{35}ClN_3O$ requires $[M+H]^+$ 324.0898); LRMS *m/z* (ESI^+) 324 ($[M^{35}Cl+H]^+$, 10%), 326 ($[M^{37}Cl+H]^+$, 36%); RP-HPLC (TFA), retention time = 11.2 min, 99.9%.

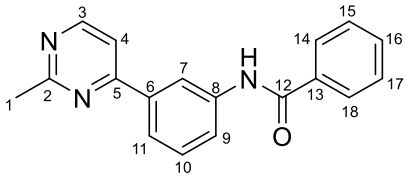
4-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (164)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (74.3 mg, 0.40 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (1.0 mL), was added 4-chlorobenzoyl chloride (62 µL, 0.48 mmol, 1.2 eq.) and triethylamine (67 µL, 0.48 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 16 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (94.3 mg, 0.29 mmol, 73% yield): R_f 0.42 (petroleum ether/EtOAc 1:3); m.p. 156–158 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3278 (N-H, w), 1650 (C=O, s), 1593 (w), 1575 (s), 1551 (m), 1486 (w), 1427 (m), 1358 (w), 1298 (m), 1261 (w), 1093 (m), 1013 (w), 916 (w), 879 (w), 832 (m), 791 (s), 756 (m), 703 (m), 643 (s), 625 (m); ^1H NMR (500 MHz, $CDCl_3$) δ 8.67 (1H, d, *J* 5.5, C^4-*H*), 8.33 (1H, dd, *J* 2.1, 1.9, C^7-*H*), 8.06 (1H, br s, N-*H*), 7.88 (1H, ddd, *J* 8.0, 2.1, 1.0, C^9-*H*), 7.85–7.81 (3H, m, C^{11}-*H,* C^{15}-*H*, C^{17}-*H*), 7.52 (1H, d, *J* 5.5, C^3-*H*), 7.49 (1H, dd, *J* 8.0, 7.9, C^{10}-*H*), 7.47–7.44 (2H, m, C^{14}-*H*, C^{18}-*H*), 2.79 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.6 (C^2), 165.0 (C^{12}), 163.5 (C^5), 157.8 (C^4), 138.6 (C^8), 138.5 (C^{16}), 138.0 (C^6), 133.2 (C^{13}), 130.0 (C^{10}), 129.3 (C^{15}, C^{17}), 128.6 (C^{14},

C^{18}), 123.5 (C^{11}), 122.7 (C^{9}), 119.0 (C^{7}), 114.2 (C^{3}), 26.4 (C^{1}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 324.0895, $C_{18}H_{14}{}^{35}ClN_3O$ requires $[M+H]^+$ 324.0898); LRMS *m/z* (ESI^+) 324 ($[M^{35}Cl+H]^+$, 100%), 326 ($[M^{37}Cl+H]^+$, 36%); RP-HPLC (TFA), retention time = 11.2 min, 100.0%.

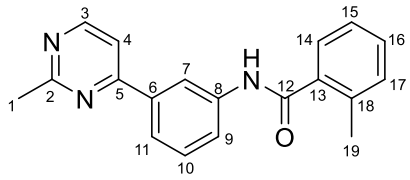
_N_-[3-(2-Methylpyrimidin-4-yl)phenyl]benzamide (165)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (93.1 mg, 0.50 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (1.3 mL), was added benzoyl chloride (70 µL, 0.60 mmol, 1.2 eq.) and triethylamine (84 µL, 0.60 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 17 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (138.1 mg, 0.48 mmol, 96% yield): R_f 0.34 (petroleum ether/EtOAc 1:3); m.p. 178–179 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3244 (N-H, w), 3061 (w), 1673 (C=O, m), 1610 (w), 1573 (m), 1550 (m), 1487 (m), 1438 (m), 1331 (m), 1305 (m), 1260 (w), 1171 (w), 993 (w), 856 (m), 837 (w), 797 (m), 774 (m), 710 (s), 692 (s), 646 (w), 629 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.67 (1H, d, *J* 5.3, C^4-*H*), 8.35 (1H, dd, *J* 1.9, 1.9, C^7-*H*), 8.09 (1H, br s, N-*H*), 7.92–7.88 (3H, m, C^{14}-*H*, C^{18}-*H*, C^9-*H*), 7.82 (1H, ddd, *J* 7.8, 1.6, 1.0, C^{11}-*H*), 7.57 (1H, dddd, *J* 7.4, 7.4, 1.3, 1.3, C^{16}-*H*), 7.53 (1H, d, *J* 5.3, C^3-*H*), 7.52–7.47 (3H, m, C^{10}-*H*, C^{15}-*H*, C^{17}-*H*), 2.79 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.5 (C^2), 166.0

(C^{12}), 163.6 (C^{5}), 157.7 (C^{4}), 138.9 (C^{8}), 138.0 (C^{6}), 134.8 (C^{13}), 132.2 (C^{16}), 129.9 (C^{10}), 129.0 (C^{15}, C^{17}), 127.2 (C^{14}, C^{18}), 123.3 (C^{11}), 122.6 (C^{9}), 118.9 (C^{7}), 114.2 (C^{3}), 26.4 (C^{1}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 290.1286, $C_{18}H_{15}N_3O$ requires $[M+H]^+$ 290.1288); LRMS *m/z* (ESI^+) 290 ($[M+H]^+$, 100%), 601 ($[2M+Na]^+$, 16%); RP-HPLC (TFA), retention time = 10.2 min, 99.8%.

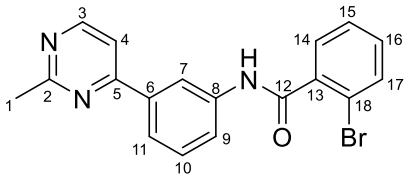
2-Methyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (166)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (55.9 mg, 0.30 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (0.75 mL), was added *o*-toluoyl chloride (47 µL, 0.36 mmol, 1.2 eq.) and triethylamine (50 µL, 0.36 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 16 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (86.9 mg, 0.29 mmol, 95% yield): R_f 0.37 (petroleum ether/EtOAc 1:3); m.p. 136–138 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3231 (N-H, w), 3072 (w), 1655 (C=O, s), 1573 (s), 1552 (s), 1449 (m), 1382 (m), 1353 (m), 1326 (w), 1305 (m), 1264 (w), 877 (w), 834 (w), 793 (s), 767 (s), 734 (m), 711 (m), 688 (m), 670 (m), 630 (m); ^1H NMR (500 MHz, $CDCl_3$) δ 8.66 (1H, d, *J* 5.3, C^4-*H*), 8.33 (1H, s, C^7-*H*), 7.87 (1H, d, *J* 8.0, C^{11}-*H*), 7.82 (1H, ddd, *J* 7.7, 1.3, 1.3, C^9-*H*), 7.76 (1H, br s, N-*H*), 7.53 (1H, d, *J* 5.3, C^3-*H*), 7.50 (1H, dd, *J* 8.1, 8.1, C^{10}-*H*), 7.52–7.47 (1H, m, C^{14}-*H*), 7.37 (1H, ddd, *J* 7.5, 7.5, 1.4, C^{15}-*H*),

7.29–7.23 (2H, m, C^{16}-*H*, C^{17}-*H*), 2.78 (3H, s, C^{1}-H_3), 2.51 (3H, s, C^{19}-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.5 (C^2), 168.4 (C^{12}), 163.6 (C^5), 157.7 (C^4), 139.0 (C^8), 138.0 (C^6), 136.7 (C^{18}), 136.3 (C^{13}), 131.5 (C^{17}), 130.6 (C^{14}), 129.9 (C^{10}), 126.7 (C^{15}), 126.1 (C^{16}), 123.3 (C^9), 122.3 (C^{11}), 118.5 (C^7), 114.3 (C^3), 26.4 (C^1), 20.0 (C^{19}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 304.1441, $C_{19}H_{17}N_3O$ requires $[M+H]^+$ 304.1444); LRMS *m/z* (ESI$^+$) 304 ($[M+H]^+$, 100%); RP-HPLC (TFA), retention time = 10.4 min, 98.8%.

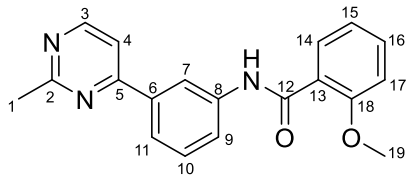
2-Bromo-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (167)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (55.5 mg, 0.30 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (0.75 mL), was added 2-bromobenzoyl chloride (47 µL, 0.36 mmol, 1.2 eq.) and triethylamine (50 µL, 0.36 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 15 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (92.5 mg, 0.25 mmol, 84% yield): R_f 0.36 (petroleum ether/EtOAc 1:3); m.p. 134–137 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3341 (N-H, w), 3055 (w), 1672 (C=O, m), 1608 (w), 1554 (s), 1489 (s), 1435 (m), 1332 (m), 1309 (m), 1251 (w), 1130 (w), 1013 (w), 858 (w), 835 (w), 805 (w), 770 (m), 743 (s), 689 (m), 628 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.67 (1H, d, *J* 5.4, C^4-*H*), 8.36 (1H, dd, *J* 2.0, 2.0, C^7-*H*), 7.92 (1H, br s, N-*H*), 7.89–7.84 (2H, m, C^9-*H*, C^{11}-*H*), 7.67–7.63 (2H, m, C^{14}-*H*, C^{17}-*H*), 7.54 (1H, d,

J 5.7, C^3-*H*), 7.52 (1H, dd, *J* 8.0, 8.0, C^{10}-*H*), 7.42 (1H, ddd, *J* 7.7, 7.6, 1.2, C^{15}-*H* or C^{16}-*H*), 7.34 (1H, ddd, *J* 7.7, 7.7, 1.8, C^{15}-*H* or C^{16}-*H*), 2.79 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.6 (C^2), 165.8 (C^{12}), 163.5 (C^5), 157.8 (C^4), 138.5 (C^8), 138.1 (C^6), 137.7 (C^{13}), 133.7 (C^{17}), 132.0, 128.0 (C^{15}, C^{16}), 130.0, 129.9 (C^{10}, C^{14}), 123.6 (C^{11}), 122.4 (C^9), 119.4 (C^{18}), 118.7 (C^7), 114.3 (C^3), 26.4 (C^1); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 368.0398, $C_{18}H_{14}{}^{79}BrN_3O$ requires $[M+H]^+$ 368.0393, found $[M+H]^+$ 370.0376, $C_{18}H_{14}{}^{81}BrN_3O$ requires $[M+H]^+$ 370.0375); LRMS *m/z* (ESI$^+$) 368 ($[M^{79}Br+H]^+$, 100%), 370 ($[M^{81}Br+H]^+$, 96%); RP-HPLC (TFA), retention time = 10.4 min, 99.9%.

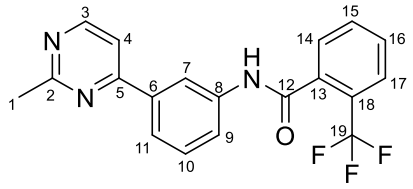
2-Methoxy-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (168)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (55.3 mg, 0.30 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (0.75 mL), was added 2-methoxybenzoyl chloride (54 μL, 0.36 mmol, 1.2 eq.) and triethylamine (50 μL, 0.36 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 16 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in CH_2Cl_2 isocratic) to afford the title compound as a colourless solid (85.7 mg, 0.27 mmol, 90% yield): R_f 0.25 (petroleum ether/EtOAc 1:3); m.p. 78–81 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3342 (N-H, w), 3054 (w), 2926 (w), 1689 (C=O, s), 1600 (w), 1551 (s), 1491 (m), 1437 (m), 1305 (m), 1236 (m), 1187 (w), 1162 (w), 1130 (w), 1047 (w), 1013 (m), 896 (w), 835 (w), 749 (s), 686 (s),

627 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.94 (1H, br s, N-*H*), 8.67 (1H, d, *J* 5.3, C^4-*H*), 8.40 (1H, dd, *J* 1.9, 1.9, C^7-*H*), 8.31 (1H, dd, *J* 7.8, 1.9, C^{14}-*H*), 7.86–7.81 (2H, m, C^9-*H*, C^{11}-*H*), 7.56 (1H, d, *J* 5.2, C^3-*H*), 7.52 (1H, ddd, *J* 8.4, 7.3, 1.9, C^{15}-*H*), 7.49 (1H, dd, *J* 7.9, 7.9, C^{10}-*H*), 7.15 (1H, ddd, *J* 7.9, 7.4, 1.0, C^{16}-*H*), 7.06 (1H, dd, *J* 8.4, 1.0, C^{17}-*H*), 4.09 (3H, s, C^{19}-H_3), 2.81 (3H, s, C^1-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.5 (C^2), 163.9 (C^5), 163.6 (C^{12}), 157.7 (C^4), 157.4 (C^{19}), 139.3 (C^8), 137.9 (C^6), 133.6 (C^{15}), 132.7 (C^{14}), 129.7 (C^{10}), 123.1 (C^{11}), 123.0 (C^9), 121.9 (C^{16}), 121.7 (C^{13}), 119.2 (C^7), 114.4 (C^3), 111.7 (C^{17}), 56.5 (C^{19}), 26.5 (C^1); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 320.1393, $C_{19}H_{17}N_3O_2$ requires $[M+H]^+$ 320.1394); LRMS *m/z* (ESI$^+$) 320 ($[M+H]^+$, 100%), 661 ($[2M+Na]^+$, 25%); RP-HPLC (TFA), retention time = 10.8 min, 99.5%.

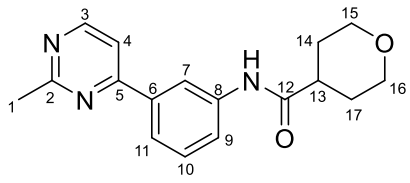
_N_-[3-(2-Methylpyrimidin-4-yl)phenyl]-2-(trifluoromethyl)benzamide (169)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (55.5 mg, 0.30 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (0.75 mL), was added 2-(trifluoromethyl)benzoyl chloride (53 µL, 0.36 mmol, 1.2 eq.) and triethylamine (50 µL, 0.36 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 16 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (99.8 mg, 0.28 mmol, 93% yield): R_f 0.34 (petroleum ether/EtOAc 1:3); m.p. 173–174 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3290 (N-H w), 1650 (C=O, m), 1572 (m), 1527 (m), 1429 (m), 1360 (w), 1311 (s), 1265 (w), 1160 (s), 1124

(s), 1053 (m), 1034 (m), 882 (w), 833 (w), 800 (m), 765 (s), 709 (m), 627 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.66 (1H, d, *J* 5.3, C^{4}-*H*), 8.30 (1H, dd, *J* 1.9, 1.9, C^{7}-*H*), 7.87–7.83 (2H, m, C^{9}-*H,* C^{11}-*H*), 7.76 (1H, d, *J* 7.2, C^{17}-*H*), 7.75 (1H, br s, N-*H*), 7.68–7.58 (3H, m, C^{14}-*H*, C^{15}-*H*, C^{16}-*H*), 7.53 (1H, d, *J* 5.4, C^{3}-*H*), 7.51 (1H, dd, *J* 7.9, 7.9, C^{10}-*H*), 2.78 (3H, s, C^{1}-*H$_3$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.5 (C^{2}), 166.0 (C^{12}), 163.5 (C^{5}), 157.7 (C^{4}), 138.4 (C^{8}), 138.1 (C^{6}), 135.7 (C^{13}), 132.4, 130.5, 128.7 (C^{14}, C^{15}, C^{16}), 130.0 (C^{10}), 127.5 (q$_F$, J_F 32.0, C^{18}), 126.8 (q$_F$, J_F 4.9, C^{17}), 123.8 (C^{11}), 123.7 (q$_F$, J_F 273.7, C^{19}), 122.6 (C^{9}), 118.8 (C^{7}), 114.3 (C^{3}), 26.4 (C^{1}); ^{19}F NMR (470 MHz, $CDCl_3$) δ –58.9 (s, C*F$_3$*); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 358.1164, $C_{19}H_{14}F_3N_3O$ requires [M+H]$^+$ 358.1162); LRMS *m/z* (ESI$^+$) 358 ([M+H]$^+$, 100%); RP-HPLC (TFA), retention time = 10.7 min, 100.0%.

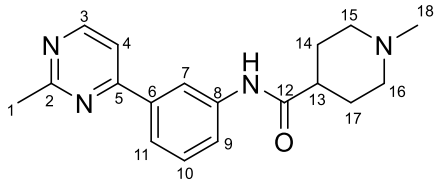
_N_-[3-(2-Methylpyrimidin-4-yl)phenyl]tetrahydro-2_H_-pyran-4-carboxamide (170)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (55.6 mg, 0.30 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (0.75 mL), was added tertrahydro-2*H*-pyran-4-carbonyl chloride (54.5 mg, 0.36 mmol, 1.2 eq.) and triethylamine (50 µL, 0.36 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 16 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (75% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (75.3 mg, 0.25 mmol, 85% yield): R_f 0.10 (petroleum ether/EtOAc 1:3); m.p. 160–161 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1}

3264 (N-H, w), 2959 (w), 2841 (w), 1655 (C=O, s), 1572 (m), 1539 (s), 1431 (m), 1393 (w), 1356 (w), 1300 (w), 1235 (m), 1192 (w), 1121 (m), 1085 (m), 1014 (w), 972 (m), 917 (w), 839 (w), 793 (m), 761 (s), 684 (s), 626 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.66 (1H, d, *J* 5.3, C^{4}-*H*), 8.26 (1H, dd, *J* 2.0, 2.0, C^{7}-*H*), 7.78 (1H, ddd, *J* 7.8, 1.3, 1.3, C^{11}-*H*), 7.75 (1H, dd, *J* 7.9, 1.5, C^{9}-*H*), 7.51 (d, *J* 5.3, C^{3}-*H*), 7.45 (1H, dd, *J* 7.9, 7.8, C^{10}-*H*), 7.42 (1H, br s, N-*H*), 4.07 (2H, ddd, *J* 11.6, 4.4, 2.2, C^{15}-H_2 or C^{16}-H_2), 3.46 (2H, ddd, *J* 11.6, 2.4, 2.4, C^{15}-H_2 or C^{16}-H_2), 2.79 (3H, s, C^{1}-H_3), 2.52 (1H, dddd, *J* 11.4, 11.3, 4.1, 4.1, C^{13}-*H*), 1.94 (2H, dddd, *J* 13.3, 11.5, 11.4, 4.4, C^{14}-H_a, C^{17}-H_a), 1.86 (2H, dddd, *J* 13.0, 4.4, 4.4, 2.2, C^{14}-H_b, C^{17}-H_b); ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.8 (C^{12}), 168.5 (C^{2}), 163.5 (C^{5}), 157.7 (C^{4}), 138.7 (C^{6}), 138.0 (C^{8}), 129.9 (C^{10}), 123.1 (C^{11}), 122.2 (C^{9}), 118.5 (C^{7}), 114.2 (C^{3}), 67.3 (C^{15}, C^{16}), 43.5 (C^{13}), 29.4 (C^{14}, C^{17}), 26.4 (C^{1}); HRMS *m/z* (ESI^{+}) (Found $[M+H]^{+}$ 298.1551, $C_{17}H_{19}N_3O_2$ requires $[M+H]^{+}$ 298.1550); LRMS *m/z* (ESI^{+}) 298 ($[M+H]^{+}$, 100%); RP-HPLC (TFA), retention time = 8.5 min, 99.9%.

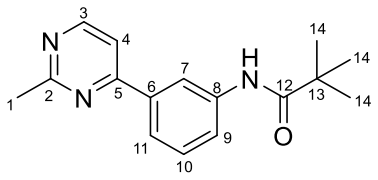
1-Methyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]piperidine-4-carboxamide (171)



To a solution of 1-methylpiperidine-4-carboxylic acid hydrochloride (90 mg, 0.50 mmol, 1.7 eq.) dissolved in dry toluene (2.0 mL) under an argon atmosphere, was added thionyl chloride (44 uL, 71 mg, 0.60 mmol, 2.0 eq.) and *N*,*N*-dimethylformamide (cat., 2 drops). The reaction solution was stirred at 75 °C for 6 hours. After this time, the solvent was removed *in vacuo* and *N*,*N*-dimethylacetamide (0.75 mL) was added. 4-(3-aminophenyl)-2-methylpyrimidine (56 mg, 0.30 mmol, 1.0 eq.) and triethylamine (121 µL, 88 mg, 0.87 mmol, 2.9 eq.) were then added and the reaction solution was stirred at room temperature overnight. After 15 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the organic components were extracted with EtOAc

(3 × 15 mL). The organic components were then washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (H_2O/IPA/EtOAc 1:1:1 isocratic). The product, which was found to be protonated, was then dissolved in EtOAc (10 mL) and washed with a saturated aqueous solution of $NaHCO_3$ (10 mL), followed by brine (10 mL), dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo* to afford the title compound as a pale brown solid (14 mg, 0.044 mmol, 15% yield): R_f 0.17 (H_2O/IPA/EtOAc 1:1:1); m.p. 143–145 °C (from H_2O/IPA/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3290 (N-H, w), 2966 (w), 2939 (w), 2853 (w), 2788 (w), 1664 (C=O, m), 1573 (s), 1546 (s), 1489 (m), 1433 (m), 1378 (m), 1303 (m), 1277 (w), 1231 (w), 1182 (w), 1134 (w), 1068 (w), 965 (w), 910 (w), 839 (w), 798 (w), 770 (m), 730 (s), 693 (m), 631 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.64 (1H, d, *J* 5.2, C^4-*H*), 8.25 (1H, dd, *J* 1.9, 1.9, C^7-*H*), 7.72 (1H, ddd, *J* 7.9, 1.9, 1.0, C^{11}-*H*), 7.73 (1H, dd, *J* 7.9, 1.9, C^9-*H*), 7.50 (1H, br s, N*H*), 7.50 (1H, d, *J* 5.2, C^3-*H*), 7.43 (1H, dd, *J* 7.9, 7.9, C^{10}-*H*), 2.94 (2H, ddd, *J* 11.4, 3.4, 3.4, C^{15}-*H$_a$*, C^{16}-*H$_a$*), 2.78 (3H, s, C^1-*H$_3$*), 2.29 (3H, s, C^{18}-*H$_3$*), 2.24 (1H, dddd, *J* 10.8, 10.8, 4.5, 4.5, C^{13}-*H*), 2.02–1.85 (6H, m, C^{14}-*H$_2$*, C^{15}-*H$_b$*, C^{16}-*H$_b$*, C^{17}-*H$_2$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 173.5 (C^{12}), 168.5 (C^2), 163.6 (C^5), 157.7 (C^4), 138.8 (C^8), 137.9 (C^6), 129.8 (C^{10}), 123.0 (C^{11}), 122.2 (C^9), 118.4 (C^7), 114.2 (C^3), 55.2 (C^{15}, C^{16}), 46.5 (C^{18}), 44.0 (C^{13}), 29.1 (C^{14}, C^{17}), 26.4 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 311.1863, $C_{18}H_{22}N_4O$ requires [M+H]$^+$ 311.1866); LRMS *m/z* (ESI$^+$) 156 ([M+2H]$^{2+}$, 20%), 311 ([M+H]$^+$, 100%), 333 ([M+Na]$^+$, 24%); RP-HPLC (TFA), retention time = 8.5 min, 99.9%.

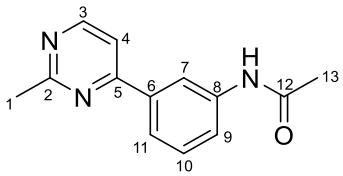
2,2-Dimethyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]propanamide (172)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (74.3 mg, 0.40 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (1.0 mL), was added trimethylacetyl chloride (59 μL,

0.48 mmol, 1.2 eq.) and triethylamine (67 µL, 0.48 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 16 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (50% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (89.1 mg, 0.33 mmol, 83% yield): R_f 0.36 (petroleum ether/EtOAc 1:3); m.p. 161–162 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3308 (N-H, w), 2975 (w), 1666 (C=O, s), 1574 (s), 1547 (s), 1488 (s), 1440 (m), 1325 (m), 1302 (s), 1219 (m), 1161 (m), 992 (w), 905 (m), 866 (m), 808 (m), 775 (s), 699 (s), 632 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.65 (1H, d, *J* 5.4, C^4-*H*), 8.24 (1H, dd, *J* 2.0, 2.0, C^7-*H*), 7.79–7.76 (2H, m, C^9-*H,* C^{11}-*H*), 7.52 (1H, d, *J* 5.4, C^3-*H*), 7.50 (1H, br s, N-*H*), 7.45 (1H, dd, *J* 7.9, 7.9, C^{10}-*H*), 2.79 (3H, s, C^1-*H$_3$*), 1.34 (9H, s, (C^{14}-*H$_3$*)$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 176.9 (C^{12}), 168.5 (C^2), 163.7 (C^5), 157.7 (C^4), 139.0 (C^8), 137.8 (C^6), 129.8 (C^{10}), 122.9 (C^9), 122.4 (C^{11}), 118.5 (C^7), 114.3 (C^3), 39.7 (C^{13}), 27.8 (C^{14} × 3), 26.4 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 270.1602, $C_{16}H_{19}N_3O$ requires [M+H]$^+$ 270.1601); LRMS *m/z* (ESI$^+$) 270 ([M+H]$^+$, 100%); RP-HPLC (TFA), retention time = 9.9 min, 100.0%.

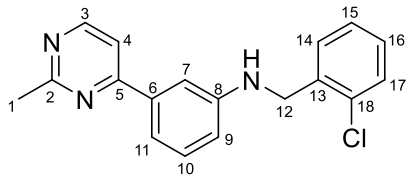
_N_-[3-(2-Methylpyrimidin-4-yl)phenyl]acetamide (173)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (74.4 mg, 0.40 mmol, 1.0 equiv.) dissolved in *N*,*N*-dimethylacetamide (1.0 mL), was added acetyl chloride (34 µL, 0.48 mmol, 1.2 eq.) and triethylamine (67 µL, 0.48 mmol, 1.2 eq.). The reaction solution was stirred at

room temperature overnight. After 16 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (60.9 mg, 0.27 mmol, 67 % yield): R_f 0.20 (petroleum ether/EtOAc 1:3); m.p. 163–165 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3257 (N-H, w), 3099 (w), 1666 (C=O, m), 1611 (m), 1567 (s), 1493 (s), 1437 (m), 1391 (w), 1369 (m), 1312 (m), 1265 (w), 1025 (w), 991 (w), 895 (w), 851 (m), 798 (w), 767 (s), 686 (s), 627 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.65 (1H, d, *J* 5.3, C^4-*H*), 8.24 (1H, dd, *J* 1.9, 1.9, C^7-*H*), 7.77 (1H, ddd, *J* 7.8, 1.3, 1.3, C^{11}-*H*), 7.73 (1H, dd, *J* 8.1, 1.3, C^9-*H*), 7.56 (1H, br s, N-*H*), 7.49 (1H, d, *J* 5.3, C^3-*H*), 7.44 (1H, dd, *J* 8.1, 7.8, C^{10}-*H*), 2.78 (3H, s, C^1-*H*$_3$), 2.21 (3H, s, C^{13}-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.6 (C^{12}), 168.5 (C^2), 163.6 (C^5), 157.7 (C^4), 138.8 (C^8), 137.9 (C^6), 129.8 (C^{10}), 123.1 (C^{11}), 122.3 (C^9), 118.5 (C^7), 114.2 (C^3), 26.4 (C^1), 24.7 (C^{13}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 228.1130, $C_{13}H_{13}N_3O$ requires $[M+H]^+$ 228.1131); LRMS *m/z* (ESI^+) 228 ($[M+H]^+$, 100%), 266 ($[M+K]^+$, 30%); RP-HPLC (TFA), retention time = 8.0 min, 99.9%.

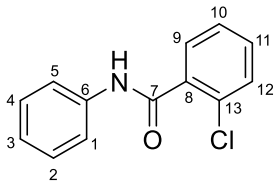
_N_-(2-Chlorobenzyl)-3-(2-methylpyrimidin-4-yl)aniline (175)



Using a soxhlet apparatus with 3 Å molecular sieves under argon atmosphere, 4-(3-aminophenyl)-2-methylpyrimidine (94 mg, 0.50 mmol, 1.0 eq.) and 2-chlorobenzaldehyde (56 µL, 70 mg, 0.50 mmol, 1.0 eq.) were heated under reflux in dry toluene (7.0 mL) for 4 hours. After this time, the reaction solution was cooled to room

temperature. Sodium borohydride (57 mg, 1.5 mmol, 3.0 eq.) and methanol (7.0 mL) were added. After stirring for 30 minutes, the solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (0–20% EtOAc in petroleum ether gradient) to afford the title compound as a pale yellow solid (115 mg, 0.37 mmol, 74% yield): R_f 0.16 (petroleum ether/EtOAc 3:1); m.p. 133–134 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3276 (N-H, w), 3052 (w), 2918 (w), 1573 (s), 1553 (m), 1495 (w), 1437 (m), 1337 (m), 1253 (w), 1129 (w), 1047 (w), 996 (w), 837 (m), 748 (s), 691 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.63 (1H, d, *J* 5.4, C^4-*H*), 7.45–7.38 (4H, m, C^3-*H*, C^7-*H*, C^{14}-*H* or C^{15}-*H* or C^{16}-*H* or C^{17}-*H*,), 7.34 (1H, ddd, *J* 7.7, 1.7, 1.0, C^{11}-*H*), 7.28 (1H, dd, *J* 7.8, 7.8, C^{10}-*H*), 7.24–7.20 (2H, m, C^{14}-*H* or C^{15}-*H* or C^{16}-*H* or C^{17}-*H*,), 6.72 (1H, ddd, *J* 8.0, 2.5, 1.0, C^9-*H*), 4.52 (2H, d, *J* 5.5, C^{12}-*H*), 4.36 (1H, t, *J* 5.5, N-*H*), 2.79 (3H, s, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.4 (C^2), 164.5 (C^5), 157.5 (C^4), 148.4 (C^8), 138.1 (C^6), 136.5, 133.5 (C^{13}, C^{18}), 129.8, 129.3, 128.7, 127.1 (C^{14}, C^{15}, C^{16}, C^{17}), 123.0 (C^{10}), 116.7 (C^{11}), 115.1 (C^9), 114.3 (C^3), 111.8 (C^7), 46.0 (C^{12}), 26.5 (C^1); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 310.1106, $C_{18}H_{16}{}^{35}ClN_3$ requires $[M+H]^+$ 310.1106); LRMS *m/z* (ESI$^+$) 310 ($[M^{35}Cl+H]^+$, 100%), 312 ($[M^{37}Cl+H]^+$, 36%); RP-HPLC (TFA), retention time = 11.7 min, 100.0%.

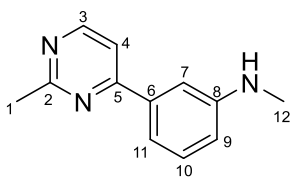
2-Chloro-*N*-phenylbenzamide (176)

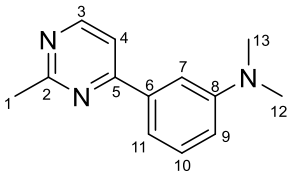


To a solution of aniline (27 μL, 0.30 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (0.75 mL), was added 2-chlorobenzoyl chloride (46 μL, 0.36 mmol, 1.2 eq.) and triethylamine (50 μL, 0.36 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 17 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc

(3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (20% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (63.5 mg, 0.28 mmol, 92% yield): R_f 0.21 (petroleum ether/EtOAc 4:1); m.p. 117–120 °C (from petroleum ether/EtOAc) [lit.[135] 115–117 °C]; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.89 (1H, br s, N-*H*), 7.76 (1H, dd, *J* 7.4, 1.9, C^9-*H* or C^{12}-*H*), 7.67–7.63 (2H, m, C^1-*H*, C^5-*H*), 7.47–7.36 (5H, m, C^2-*H*, C^4-*H*, C^{10}-*H*, C^{11}-*H*, C^9-*H* or C^{12}-*H*), 7.18 (tt, *J* 7.5, 7.5, 1.3, 1.3, C^3-*H*); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 232.0523, $C_{13}H_{10}{}^{35}ClNO$ requires $[M+H]^+$ 232.0524); LRMS *m/z* (ESI$^+$) 232 ($[M^{35}Cl+H]^+$, 83%), 234 ($[M^{37}Cl+H]^+$, 38%), 254 ($[M^{35}Cl+Na]^+$, 66%), 256 ($[M^{35}Cl+Na]^+$, 21%), 485 ($[2M^{35}Cl+Na]^+$, 100%), 487 ($[M^{35}Cl+M^{37}Cl+Na]^+$, 70%), 489 ($[2M^{37}Cl+Na]^+$, 14%); RP-HPLC retention time = 11.9 min, 99.7%. These data are in agreement with the literature.[135]

_N_-Methyl-3-(2-methylpyrimidin-4-yl)aniline (177A) and _N,N_-dimethyl-3-(2-methylpyrimidin-4-yl)aniline (177B)

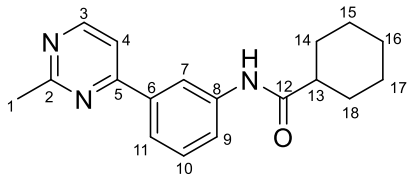




To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (93 mg, 0.5 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylformamide (1.0 mL), iodomethane (31 µL, 71 mg, 0.5 mmol, 1.0 eq.) was added and the reaction solution was stirred at room temperature overnight. After 16 hours, additional iodomethane (31 µL, 71 mg, 0.5 mmol, 1.0 eq.) was added and the reaction solution was stirred for 2 hours. After this time, H_2O (10 mL) was added and the organic components were extracted with EtOAc (3 × 20 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified

by flash column chromatography on silica gel (0–25% EtOAc in petroleum ether gradient) to afford the title compounds **177A** as a yellow solid (33 mg, 0.28 mmol, 56% yield) and **177B** as a yellow oil (25 mg, 0.12 mmol, 23% yield): **177A** R_f 0.11 (petroleum ether/EtOAc 3:1); m.p. 103–106 °C (from petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3297 (N-H, w), 2928 (w), 1575 (m), 1549 (s), 1465 (w), 1429 (m), 1392 (m), 1359 (w), 1309 (m), 1276 (m), 1173 (w), 995 (w), 838 (m), 790 (w), 761 (s), 689 (s), 653 (m), 628 (s); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.64 (1H, d, *J* 5.4, C^4-*H*), 7.47 (1H, d, *J* 5.4, C^3-*H*), 7.36–7.28 (3H, m, C^7-*H,* C^{10}-*H,* C^9-*H* or C^{11}-*H*), 6.74 (1H, ddd, *J* 7.5, 2.5, 1.5, C^9-*H* or C^{11}-*H*), 3.89 (1H, br s, N-*H*), 2.92 (3H, d, *J* 3.7, C^{12}-*H$_3$*), 2.79 (3H, s, C^1-*H$_3$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.4 (C^2), 164.8 (C^5), 157.4 (C^4), 150.0 (C^8), 138.1 (C^6), 129.9 (C^{10}), 116.3, 114.9 (C^9, C^{11}), 114.3 (C^3), 110.9 (C^7), 30.9 (C^{12}), 26.5 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 200.1182, $C_{12}H_{13}N_3$ requires [M+H]$^+$ 200.1182); LRMS *m/z* (ESI$^+$) 200 ([M+H]$^+$, 100%); RP-HPLC (TFA), retention time = 7.3 min, 99.6%; **177B** R_f 0.15 (petroleum ether/EtOAc 3:1); $\bar{\nu}_{max}$/cm^{-1} 2981 (w), 2924 (w), 2851 (w), 2804 (w), 1604 (m), 1570 (s), 1548 (s), 1499 (m), 1434 (m), 1396 (w), 1353 (m), 1231 (w), 1180 (w), 1001 (w), 836 (w), 788 (w), 766 (w), 690 (w), 629 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.64 (1H, d, *J* 5.3, C^4-*H*), 7.49 (1H, d, *J* 5.3, C^3-*H*), 7.45–7.47 (1H, m, C^7-*H* or C^{10}-*H*), 7.32–7.37 (2H, m, C^{11}-*H*, C^7-*H* or C^{10}-*H*), 6.87 (1H, ddd, *J* 7.3, 2.4, 2.4, C^9-*H*), 3.04 (6H, s, C^{12}-*H$_3$*, C^{13}-*H$_3$*), 2.80 (3H, s, C^1-*H$_3$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.4 (C^2), 165.1 (C^5), 157.4 (C^4), 151.1 (C^8), 137.9 (C^6), 129.7 (C^{10}), 115.6 (C^{11}), 115.1 (C^9), 114.4 (C^3), 111.1 (C^7), 40.8 (C^{12}, C^{13}), 26.5 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 214.1339, $C_{13}H_{15}N_3$ requires [M+H]$^+$ 214.1339); LRMS *m/z* (ESI$^+$) 214 ([M+H]$^+$, 100%); RP-HPLC (TFA), retention time = 7.4 min, 97.6%.

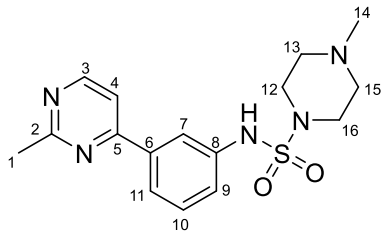
_N_-[3-(2-Methylpyrimidin-4-yl)phenyl]cyclohexanecarboxamide (178)



To a solution of 4-(3-aminophenyl)-2-methylpyrimidine (55.4 mg, 0.30 mmol, 1.0 eq.) dissolved in *N*,*N*-dimethylacetamide (0.75 mL), was added cyclohexane carbonyl chloride (48 µL, 0.36 mmol, 1.2 eq.) and triethylamine (50 µL, 0.36 mmol, 1.2 eq.). The reaction solution was stirred at room temperature overnight. After 17 hours, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added, and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$, and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (65% EtOAc in petroleum ether isocratic) to afford the title compound as a colourless solid (79.9 mg, 0.27 mmol, 91% yield): R_f 0.35 (petroleum ether/EtOAc 1:3); m.p. 149–151 °C (from petroleum ether /EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3290 (N-H, w), 2929 (w), 2852 (w), 1656 (C=O, m), 1599 (w), 1573 (s), 1551 (s), 1433 (m), 1390 (m), 1357 (w), 1280 (w), 1229 (w), 1193 (w), 958 (m), 794 (s), 769 (m), 746 (m), 701 (s), 632 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.65 (1H, d, *J* 5.4, C^4-*H*), 8.27 (1H, dd, *J* 1.6, 1.6, C^7-*H*), 7.77 (1H, ddd, *J* 8.0, 1.6, 1.3, C^{11}-*H*), 7.74 (1H, dd, *J* 8.0, 1.6, C^9-*H*), 7.51 (d, *J* 5.4, C^3-*H*), 7.44 (1H, dd, *J* 8.0, 8.0, C^{10}-*H*), 7.38 (1H, br s, N*H*), 2.79 (3H, s, C^1-*H$_3$*), 2.26 (2H, tt, *J* 11.7, 3.4, C^{13}-*H*), 2.01–1.95 (2H, m, C^{14}-*H$_2$* or C^{18}-*H$_2$*), 1.88–1.82 (2H, m, C^{15}-*H$_2$* or C^{16}-*H$_2$* or C^{17}-*H$_2$*), 1.74–1.68 (1H, m, C^{15}-*H$_2$* or C^{16}-*H$_2$* or C^{17}-*H$_2$*), 1.56 (2H, tdd, *J* 12.1, 12.1, 3.4, C^{14}-*H$_2$* or C^{18}-*H$_2$*), 1.37–1.21 (3H, m, C^{15}-*H$_2$* or C^{16}-*H$_2$* or C^{17}-*H$_2$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 174.7 (C^{12}), 168.5 (C^2), 163.7 (C^6), 157.7 (C^4), 139.0 (C^8), 137.9 (C^5), 129.8 (C^{10}), 122.9 (C^{11}), 122.1 (C^9), 118.4 (C^7), 114.2 (C^3), 46.8 (C^{13}), 29.8 (C^{14}, C^{18}), 26.4 (C^1), 25.8 (C^{15}, C^{16}, C^{17}); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$

296.1757, $C_{18}H_{21}N_3O$ requires $[M+H]^+$ 296.1685); LRMS *m/z* (ESI$^-$) 294.1 ($[M-H]^-$, 100%); RP-HPLC, retention time = 9.3 min, 99.9%.

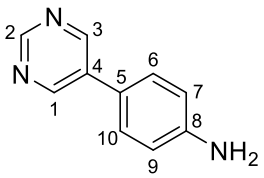
4-Methyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]piperazine-1-sulfonamide (179)



To a solution of 4-methyl-1-piperazinesulfonyl chloride hydrochloride (70.6 mg, 0.3 mmol, 1.0 eq.) dissolved in CH_2Cl_2 (1.0 mL) was added 4-(3-aminophenyl)-2-methylpyrimidine (112.0 mg, 0.6 mmol, 2.0 eq.) in 3,5-lutidine (1.0 mL). The reaction mixture was stirred at room temperature overnight. After 16 hours, the reaction mixture was diluted with EtOAc (20 mL) and a saturated aqueous solution of $NaHCO_3$ (10 mL) was added. H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (0–20% EtOH in EtOAc gradient) to afford the title compound as a colourless solid (95.4 mg, 0.27 mmol, 92% yield): R_f 0.04 (EtOAc/EtOH 9:1); m.p. 99–101 °C (from EtOAc/EtOH); $\bar{\nu}_{max}$/cm^{-1} 3024 (w), 2954 (w), 2923 (w), 1572 (m), 1437 (w), 1349 (w), 1285 (w), 1150 (m), 1066 (w), 936 (m), 842 (w), 786 (w), 748 (s), 727 (s), 691 (m), 667 (w), 651 (w), 628 (m), 615 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.69 (1H, d, *J* 5.4, C^4-*H*), 7.93 (1H, dd, *J* 2.2, 1.9, C^7-*H*), 7.79 (1H, ddd, *J* 7.9, 1.9, 1.0, C^9-*H* or C^{11}-*H*), 7.49 (1H, d, *J* 5.4, C^3-*H*), 7.44 (1H, dd, *J* 8.0, 7.9, C^{10}-*H*), 7.35 (1H, ddd, *J* 8.0, 2.2, 1.0, C^9-*H*), 6.75 (1H, br s, N-*H*), 3.33 (4H, t, *J* 5.0, 5.0, C^{12}-*H$_2$* and C^{16}-*H$_2$* or C^{13}-*H$_2$* and C^{15}-*H$_2$*), 2.80 (3H, s, C^1-*H$_3$*), 2.41 (4H, t, *J* 5.0, 5.0, C^{12}-*H$_2$* and C^{16}-*H$_2$* or C^{13}-*H$_2$* and C^{15}-*H$_2$*), 2.25 (3H, s, C^{14}-*H$_3$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 168.6 (C^2), 163.1 (C^6), 157.9 (C^4), 138.4 (C^5), 138.1 (C^8), 130.1 (C^{10}), 123.4, 122.5 (C^9, C^{11}), 119.0 (C^7), 114.1 (C^3), 54.3, 46.4 (C^{12}, C^{13}, C^{15}, C^{16}), 46.0 (C^{14}), 26.5 (C^1); HRMS *m/z*

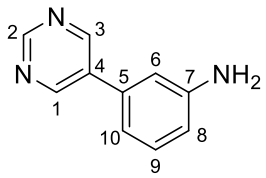
(ESI$^+$) (Found [M+H]$^+$ 348.1489, $C_{16}H_{21}N_5O_2S$ requires [M+H]$^+$ 348.1489); LRMS *m/z* (ESI$^-$) 346.0 ([M–H]$^-$, 100%); RP-HPLC (TFA), retention time = 5.6 min, 99.9%.

4-(Pyrimidin-5-yl)aniline (181)



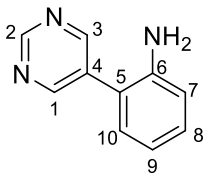
Pyrimidine-5-boronic acid (147.0 mg, 1.1 mmol, 1.1 eq.), 4-bromoaniline (172.4 mg, 1.0 mmol, 1.0 eq.), tris(dibenzylideneacetone)dipalladium (0) (9.8 mg, 0.01 mmol, 0.01 eq.) and tricyclohexylphosphine (6.7 mg, 0.024 mmol, 0.024 eq.) were placed in a flask and flushed with argon. Dry 1,4-dioxane (2.7 mL) was then added, followed by K_3PO_4 (361.9 mg, 1.7 mmol, 1.7 eq.) dissolved in H_2O (1.3 mL). The reaction mixture was stirred at 95 °C for 17 hours. After this time, the reaction mixture was filtered through Celite® and washed with EtOAc. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (60–75% EtOAc in petroleum ether gradient) to afford the title compound as a yellow-orange solid (125.2 mg, 0.73 mmol, 73% yield): R_f 0.14 (petroleum ether/EtOAc 3:2); m.p. 202–204 °C (petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3386 (N-H, w), 3300 (N-H, w), 3172 (w), 3022 (w), 2922 (w), 1643 (w), 1612 (m), 1557 (m), 1519 (m), 1412 (m), 1288 (m), 1186 (m), 999 (s), 723 (s), 631 (s); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.11 (1H, s, C^2-*H*), 8.88 (2H, s, C^1-*H*, C^3-*H*), 7.41–7.38 (2H, m, C^6-*H*, C^{10}-*H*), 6.82–6.78 (2H, m, C^7-*H*, C^9-*H*), 3.89 (2H, s, N-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.6 (C^2), 154.2 (C^1, C^3), 147.6 (C^5), 134.4 (C^4), 128.1 (C^6, C^{10}), 124.1 (C^7, C^9), 115.8 (C^8); HRMS *m/z* (EI) (Found [M]$^+$ 171.0792, $C_{10}H_9N_3$ requires [M]$^+$ 171.0796); Could not be detected by LRMS (ESI); RP-HPLC (TFA), retention time = 5.8 min, 96.7%.

3-(Pyrimidin-5-yl)aniline (182)



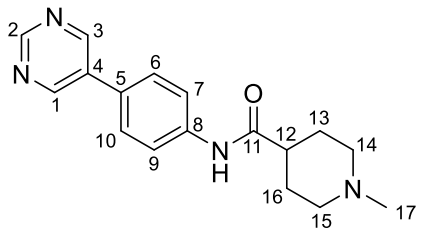
Pyrimidine-5-boronic acid (147.8 mg, 1.1 mmol, 1.1 eq.), tris(dibenzylideneacetone)dipalladium (0) (18.6 mg, 0.02 mmol, 0.02 eq.) and tricyclohexylphosphine (13.8 mg, 0.048 mmol, 0.048 eq.) were placed in a flask and flushed with argon. 3-Bromoaniline (109 µL, 172.0 mg, 1.0 mmol, 1.0 eq.) in 1,4-dioxane (2.7 mL) was then added, followed by K_3PO_4 (363.0 mg, 1.7 mmol, 1.7 eq.) dissolved in H_2O (1.3 mL). The reaction mixture was then stirred at 95 °C overnight. After 16 hours, the reaction mixture was filtered through Celite® and washed with EtOAc. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel using the Biotage system (50–75% EtOAc in petroleum ether gradient) to afford the title compound as a yellow-orange solid (107.8 mg, 0.63 mmol, 63% yield): R_f 0.20 (petroleum ether/EtOAc 2:1); m.p. 166–168 °C (petroleum ether/EtOAc); $\bar{\nu}_{max}$/cm^{-1} 3387 (N-H, w), 3312 (N-H, w), 3197 (w), 3047 (w), 3026 (w), 1638 (w), 1604 (w), 1585 (m), 1559 (w), 1412 (s), 1237 (m), 1189 (w), 992 (w), 788 (s), 726 (s), 697 (s), 637 (s); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.18 (1H, s, C^2-*H*), 8.92 (2H, s, C^1-*H*, C^3-*H*), 7.30 (1H, dd, *J* 7.9, 7.8, C^9-*H*), 6.95 (1H, ddd, *J* 7.8, 1.9, 0.9, C^{10}-*H*), 6.86 (1H, dd, *J* 2.2, 1.9, C^6-*H*), 6.78 (1H, ddd, *J* 7.9, 2.2, 0.9, C^8-*H*), 3.85 (2H, br s, N-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 157.6 (C^2), 155.0 (C^1, C^3), 147.5 (C^7), 135.5 (C^5), 134.7 (C^4), 130.6 (C^9), 117.3 (C^{10}), 115.7 (C^8), 113.3 (C^6); HRMS *m/z* (EI) (Found [M]$^+$ 171.0792, $C_{10}H_9N_3$ requires [M]$^+$ 171.0796); Could not be detected by LRMS (ESI); RP-HPLC (TFA), retention time = 6.0 min, 98.5%.

2-(Pyrimidin-5-yl)aniline (183)



Pyrimidine-5-boronic acid (200.6 mg, 1.5 mmol, 1.1 eq.), tris(dibenzylideneacetone)dipalladium (0) (25.3 mg, 0.02 mmol, 0.02 eq.) and tricyclohexylphosphine (13.8 mg, 0.048 mmol, 0.048 eq.) were placed in a flask and flushed with argon. 3-bromoaniline (109 µL, 172.0 mg, 1.0 mmol, 1.0 eq.) in 1,4-dioxane (2.7 mL) was then added, followed by K_3PO_4 (363.0 mg, 1.7 mmol, 1.7 eq.) dissolved in H_2O (1.3 mL). The reaction mixture was stirred at 95 °C overnight. After 18 hours, the reaction mixture was filtered through Celite® and washed with EtOAc. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel using the Biotage system (50–75% EtOAc in petroleum ether gradient) to afford the title compound as a cream solid (107.8 mg, 0.63 mmol, 63% yield): R_f 0.20 (petroleum ether/EtOAc 2:1); m.p. 94–95 °C (petroleum ether/EtOAc); $\bar{\nu}_{max}/cm^{-1}$ 3387 (N-H, w), 3342 (N-H, w), 3230 (w), 3026 (w), 1652 (w), 1606 (w), 1554 (w), 1493 (m), 1403 (m), 1295 (w), 1181 (w), 999 (w), 842 (w), 745 (s), 725 (s), 634 (s); ^{1}H NMR (500 MHz, acetone-D6) δ 9.08 (1H, s, C^2-*H*), 8.83 (2H, s, C^1-*H*, C^3-*H*), 7.17 (1H, ddd, *J* 8.1, 7.5, 1.4, C^9-*H*), 7.09 (1H, dd, *J* 7.5, 1.6, C^{10}-*H*), 6.88 (1H, dd, *J* 8.1, 1.1, C^8-*H*), 6.76 (1H, ddd, *J* 7.5, 7,4, 1.4, C^7-*H*), 4.71 (2H, br s, N-*H*$_2$); ^{13}C NMR (126 MHz, acetone-D6) δ 157.8 (C^2), 157.5 (C^1, C^3), 146.6 (C^5), 134.8 (C^4), 131.3 (C^6), 130.6 (C^9), 120.5 (C^{10}), 118.6 (C^7), 116.8 (C^8); HRMS *m/z* (EI) (Found $[M]^+$ 171.0792, $C_{10}H_9N_3$ requires $[M]^+$ 171.0790); Could not detect by LRMS (ES); RP-HPLC (TFA), retention time = 6.4 min, 98.8%.

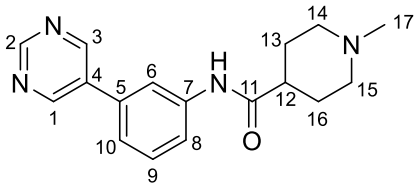
1-Methyl-*N*-[4-(pyrimidin-5-yl)phenyl]piperidine-4-carboxamide (185)



To a solution of 1-methylpiperidine-4-carboxylic acid hydrochloride (538.6 mg, 3.0 mmol, 3.0 eq.) in dry toluene (3.0 mL) under argon atmosphere, thionyl chloride (1.1 mL, 15.0 mmol, 15.0 eq.) was added. The reaction solution was stirred at 70 °C for 1.5 hour. After this time, the solvent was removed and CH_2Cl_2 (4.0 mL) was added. The solution was cooled to 0 °C and 4-(pyrimidin-5-yl)aniline (**181**) (170.9 mg, 1.0 mmol, 1.0 eq.) in CH_2Cl_2 (2.0 mL) was added dropwise at 0 °C. Triethylamine (697 µL, 5.0 mmol, 5.0 eq.) was then added and the reaction solution was allowed to warm to room temperature and stirred for 20 hours. After this time, a saturated aqueous solution of $NaHCO_3$ (10 mL) was added and the aqueous layer was extracted with $CHCl_3$ (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (H_2O/IPA/EtOAc/Et_3N 1:2:9:0.5 to 1:2:7:0.5 gradient) followed by flash column chromatography on silica gel (EtOAc/EtOH/Et_3N 7.5:2:0.5 to 6:3.5:0.5 gradient) to afford the title compound as a colourless solid (161.6 mg, 0.55 mmol, 55% yield): R_f 0.31 (EtOAc/EtOH/Et_3N 4.5:5:0.5); m.p. 198 °C (decomposed) (from EtOAC/EtOH); $\bar{\nu}_{max}$/cm^{-1} 3230 (N-H, w), 2941 (w), 2790 (w), 1673 (C=O, m), 1597 (w), 1498 (m), 1414 (m), 1236 (m), 1186 (w), 908 (m), 836 (m), 724 (s), 644 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.17 (1H, s, C^2-*H*), 8.92 (2H, s, C^1-*H*, C^3-*H*), 7.72–7.67 (2H, m, C^7-*H*, C^9-*H*), 7.56–7.52 (2H, m, C^6-*H*, C^{10}-*H*), 7.40 (1H, br s, N-*H*), 2.95 (2H, dt, *J* 11.8, 3.4, C^{14}-H_a, C^{15}-H_a), 2.30 (3H, s, C^{17}-H_3), 2.30–2.22 (1H, m, C^{12}-*H*), 2.05–1.83 (6H, m, C^{14}-H_b, C^{15}-H_b, C^{13}-H_2, C^{16}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 173.5 (C^{11}), 157.4 (C^2), 154.7 (C^1, C^3), 139.0 (C^8), 133.8 (C^4), 130.0 (C^5), 127.7 (C^6, C^{10}),

120.6 (C^7, C^9), 55.2 (C^{14}, C^{15}), 46.5 (C^{17}), 44.0 (C^{12}), 29.1 (C^{13}, C^{16}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 297.17102, $C_{17}H_{20}N_4O$ requires $[M+H]^+$ 297.17099); LRMS *m/z* (ESI^-) 295 ($[M-H]^-$, 100%); RP-HPLC (TFA), retention time = 5.2 min, 99.9%.

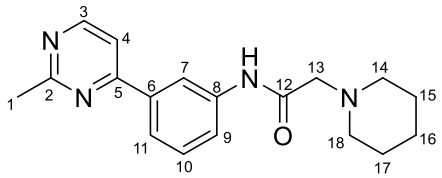
1-Methyl-*N*-[3-(pyrimidin-5-yl)phenyl]piperidine-4-carboxamide (186)



To a solution of 1-methylpiperidine-4-carboxylic acid hydrochloride (538.5 mg, 3.0 mmol, 8.3 eq.) dissolved in dry toluene (3.0 mL) under argon atmosphere, thionyl chloride (1.1 mL, 15.0 mmol, 41.7 eq.) was added. The reaction mixture was stirred at 70 °C for 1.5 hour. After this time, the solvent was removed and CH_2Cl_2 (4.0 mL) was added. The solution was cooled in to 0 °C and 3-(pyrimidin-5-yl)aniline (**182**) (62.0 mg, 0.36 mmol, 1.0 eq.) in CH_2Cl_2 (2.0 mL) was added dropwise at 0 °C. Triethylamine (697 µL, 5.0 mmol, 13.9 eq.) was then added and the reaction mixture was allowed to warm to room temperature overnight. After 16 hours, a saturated aqueous solution of $NaHCO_3$ (10 mL) was added and the aqueous layer was extracted with $CHCl_3$ (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (H_2O/IPA/EtOAc/Et_3N 1:2:9:0.5 isocratic) followed by flash column chromatography on silica gel (CH_2Cl_2/EtOH/Et_3N 9.5:0:0.5 to 9.1:0.4:0.5 gradient) to afford the title compound as a colourless solid (76.0 mg, 0.26 mmol, 71% yield): R_f 0.24 (CH_2Cl_2/EtOH/Et_3N 9.5:0.5:0.5); m.p. 128–130 °C (from CH_2Cl_2/EtOH); $\bar{\nu}_{max}$/cm^{-1} 3286 (N-H, w), 3040 (w), 2938 (w), 2791 (w), 1669 (C=O, m), 1596 (m), 1553 (s), 1449 (m), 1415 (s), 1232 (w), 1134 (w), 789 (w), 726 (m), 697 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.20 (1H, s, C^2-*H*), 8.94 (2H, s, C^1-*H*, C^3-*H*), 7.95 (1H, dd, *J* 2.1, 1.7, C^6-*H*), 7.51 (1H, ddd, *J* 8.0, 2.1, 1.3, C^{10}-*H*), 7.49 (1H,br s, N-*H*), 7.46 (1H, dd, *J* 8.0, 7.7, C^9-*H*), 7.31 (1H, ddd, *J* 7.7, 1.7, 1.3, C^8-*H*), 2.97 (2H, dt, *J* 11.6, 3.4,

C^{14}-H_a, C^{15}-H_a), 2.32 (3H, s, C^{17}-H_3), 2.31–2.50 (1H, m, C^{12}-*H*), 2.08–1.88 (6H, m, C^{14}-H_b, C^{15}-H_b, C^{13}-H_2, C^{16}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 173.5 (C^{11}), 157.8 (C^2), 155.1 (C^1, C^3), 139.2 (C^7), 135.4 (C^5), 134.1 (C^4), 130.2 (C^9), 122.9 (C^8), 120.2 (C^{10}), 118.4 (C^6), 55.1 (C^{14}, C^{15}), 46.4 (C^{17}), 43.8 (C^{12}), 29.0 (C^{13}, C^{16}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 297.17108, $C_{17}H_{20}N_4O$ requires $[M+H]^+$ 297.17099); LRMS *m/z* (ESI^-) 295 ($[M-H]^-$, 100%); RP-HPLC (TFA), retention time = 5.5 min, 98.7%.

_N_-[3-(2-Methylpyrimidin-4-yl)phenyl]-2-(piperidin-1-yl)acetamide (190)



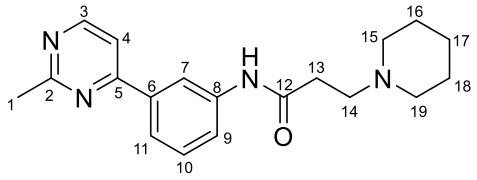
To a solution of bromoacetic acid (139.4 mg, 1.0 mmol, 1.0 eq.) in H_2O (2.0 mL) at 0 °C was added 3.3 M aqueous NaOH solution (1.0 mL) followed by piperidine (99 μL, 85.2 mg, 1.0 mmol, 1.0 eq.). The reaction solution was stirred at room temperature for 16 hours and then lyophilised to afford (piperidin-1-yl)acetic acid which was used without further purification.

Thionyl chloride (2.0 mL) was added dropwise to the crude (piperidin-1-yl)acetic acid and stirred at 40 °C for 1 hour. After this time, the solvent was removed, and the acid chloride was used without further purification.

To the crude (piperidin-1-yl)acetyl chloride was added 4-(3-aminophenyl)-2-methylpyrimidine (92.4 mg, 0.5 mmol, 1.0 eq.) and triethylamine (84 μL, 60.7 mg, 0.6 mmol, 1.2 eq.) dissolved in *N*,*N*-dimethylacetamide (1.0 mL). The reaction solution was stirred at room temperature for 17 hours. After this time, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were then washed with 0.5 M aqueous LiCl

solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (75–90% EtOAc in petroleum ether gradient) to afford the title compound as a yellow oil (38.0 mg, 0.12 mmol, 25% yield): R_f 0.08 (petroleum ether/EtOAc 1:3); $\bar{\nu}_{max}$/cm^{-1} 3284 (N-H, w), 2935 (w), 2855 (w), 2812 (w), 1685 (C=O, m), 1573 (s), 1550 (s), 1525 (s), 1490 (w), 1436 (m), 1304 (w), 797 (w), 772 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.40 (1H, br s, N-*H*), 8.66 (1H, d, *J* 5.2, C^4-*H*), 8.22 (1H, dd, *J* 2.0, 2.0, C^7-*H*), 7.82–7.76 (2H, m, C^9-*H*, C^{11}-*H*), 7.52 (1H, d, *J* 5.2, C^3-*H*), 7.46 (1H, dd, *J* 7.9, 7.9, C^{10}-*H*), 3.10 (2H, s, C^{13}-*H*$_2$), 2.80 (3H, s, C^1-*H*$_3$), 2.61–2.50 (4H, dd, *J* 5.3, 5.3, C^{14}-*H*$_2$, C^{18}-*H*$_2$), 1.66 (4H, dddd, *J* 5.6, 5.6, 5.6, 5.6, C^{15}-*H*$_2$, C^{17}-*H*$_2$), 1.53–1.45 (2H, m, C^{16}-*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.3 (C^{12}), 168.5 (C^2), 163.8 (C^6), 157.7 (C^4), 138.6 (C^8), 137.9 (C^5), 129.8 (C^{10}), 122.9, 121.9 (C^9, C^{11}), 118.1 (C^7), 114.3 (C^3), 63.0 (C^{13}), 55.1 (C^{14}, C^{18}), 26.42 (C^1), 26.35 (C^{15}, C^{17}), 23.7 (C^{16}); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$ 311.1863, $C_{18}H_{22}N_4O$ requires [M+H]$^+$ 311.1866); LRMS *m/z* (ESI$^+$) 311 ([M+H]$^+$, 8%), 333 ([M+Na]$^+$, 100%); RP-HPLC (TFA) retention time = 5.4 min, 96.0%.

_N_-[3-(2-Methylpyrimidin-4-yl)phenyl]-3-(piperidin-1-yl)propanamide (191)

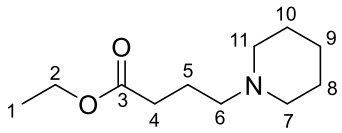


To a solution of 3-bromopropionic acid (154.0 mg, 1.0 mmol, 1.0 eq.) in H_2O (2.0 mL) at 0 °C was added 3.3 M aqueous NaOH solution (1.0 mL) followed by piperidine (99 μL, 85.2 mg, 1.0 mmol, 1.0 eq.). The reaction solution was stirred at room temperature for 17 hours and then lyophilised to afford 3-(piperidin-1-yl)propanoic acid which was used without further purification.

Thionyl chloride (2.0 mL) was added dropwise to the crude 3-(piperidin-1-yl)propanoic acid and stirred at 40 °C for 1 hour. After this time, the solvent was removed, and the acid chloride was used without further purification.

To the crude 3-(piperidin-1-yl)propanoyl chloride was added 4-(3-aminophenyl)-2-methylpyrimidine (92.5 mg, 0.5 mmol, 1.0 eq.) and triethylamine (84 µL, 60.7 mg, 0.6 mmol, 1.2 eq.) dissolved in *N*,*N*-dimethylacetamide (1.5 mL). The reaction solution was stirred at room temperature for 16 hours. After this time, EtOAc (1.5 mL) and a saturated aqueous solution of $NaHCO_3$ (9 mL) were added and the mixture was stirred at room temperature for 1 hour. After this time, H_2O (10 mL) was added and the aqueous layer was extracted with EtOAc (3 × 15 mL). The organic components were then washed with 0.5 M aqueous LiCl solution (3 × 20 mL), followed by brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (75–100% EtOAc in petroleum ether gradient to 0–20% EtOH in EtOAc gradient) to afford the title compound as a brown oil (98.3 mg, 0.28 mmol, 55% yield): R_f 0.08 (petroleum ether/EtOAc 1:3); $\bar{\nu}_{max}$/cm^{-1} 3296 (N-H, w), 2934 (w), 2852 (w), 2808 (w), 1675 (C=O, m), 1572 (s), 1550 (s), 1492 (w), 1437 (m), 1304 (w), 798 (w), 772 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.59 (1H, br s, N-*H*), 8.65 (1H, d, *J* 5.5, C^4-*H*), 8.34 (1H, dd, *J* 2.0, 1.9, C^7-*H*), 7.76 (1H, ddd, *J* 7.9, 2.0, 1.2, C^{11}-*H*), 7.68 (1H, ddd, *J* 8.1, 2.0, 1.9, C^9-*H*), 7.50 (1H, d, *J* 5.5, C^3-*H*), 7.42 (1H, dd, *J* 8.1, 7.9, C^{10}-*H*), 2.77 (3H, s, C^1-*H$_3$*), 2.70 (2H, dd, *J* 6.7, 5.1, C^{14}-*H$_2$*), 2.65–2.50 (4H, m, C^{15}-*H$_2$*, C^{19}-*H$_2$*), 2.54 (2H, dd, *J* 6.7, 5.1, C^{13}-*H$_2$*), 1.74 (4H, dddd, *J* 5.7, 5.7, 5.7, 5.7, C^{16}-*H$_2$*, C^{18}-*H$_2$*), 1.57 (2H, m, C^{17}-*H$_2$*); ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.1 (C^{12}), 168.4 (C^2), 163.7 (C^5), 157.6 (C^4), 139.8 (C^8), 137.7 (C^6), 129.7 (C^{10}), 122.2 (C^{11}), 121.8 (C^9), 118.1 (C^7), 114.1 (C^3), 54.4 (C^{14}), 53.7 (C^{15}, C^{19}), 32.6 (C^{13}), 26.4 (C^1), 26.3 (C^{16}, C^{18}), 24.3 (C^{17}); HRMS *m/z* (ESI$^+$) (Found $[M+H]^+$ 325.2021, $C_{19}H_{24}N_4O$ requires $[M+H]^+$ 325.2023); LRMS *m/z* (ESI$^+$) 347 ($[M+Na]^+$, 100%); RP-HPLC (TFA), retention time = 5.5 min, 96.6%.

Ethyl 4-(piperidin-1-yl)butanoate (193)

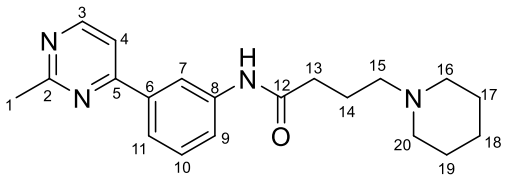


To a solution of ethyl-4-bromobutyrate (1.45 mL, 1.95 g, 10.0 mmol, 1.0 eq.) dissolved in acetone (25 mL), potassium iodide (4.99 g, 30.0 mmol, 3.0 eq.) was added. The reaction mixture was stirred under reflux for 6 hours. After this time, the solvent was removed *in vacuo* and the residue was dissolved in Et_2O (40 mL). The organic component was washed with aqueous NaS_2O_3 solution (10% *w/v*, 3 × 15 mL), followed by H_2O (2 × 15 mL), then brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* to afford ethyl 4-iodobutanoate (**192**) as a yellow oil without further purification (2.19 g, 9.04 mmol, 89% yield).

To a solution of ethyl 4-iodobutanoate (**192**) (2.19 g, 9.04 mmol, 1.0 eq.) in MeCN (15.0 mL) was added piperidine (1.80 mL, 1.54 g, 18.1 mmol, 2.2 eq.). The reaction solution was stirred under reflux for 4 hours. After this time, the solvent was removed *in vacuo* and the residue was dissolved in EtOAc (20 mL). The organic component was washed with aqueous NaS_2O_3 solution (10% *w/v*, 2 × 15 mL), followed by H_2O (2 × 15 mL), then brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* to afford ethyl 4-(piperidin-1-yl)butanoate as a yellow oil (1.43 g, 7.18 mmol, 79% yield): R_f 0.09 (petroleum ether/EtOAc 4:1); $\bar{\nu}_{max}$/cm^{-1} 2934 (m), 2762 (w), 1733 (C=O, s), 1444 (w), 1372 (m), 1156 (s), 1114 (m), 1037 (m), 953 (w), 860 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.10 (2H, q, *J* 7.2, C^2-*H*$_2$), 2.38–2.30 (4H, m, C^7-*H*$_2$, C^{11}-*H*$_2$), 2.29 (2H, t, *J* 7.5, C^4-*H*$_2$ or C^6-*H*$_2$), 2.28 (2H, t, *J* 7.5, C^4-*H*$_2$ or C^6-*H*$_2$), 1.82–1.75 (2H, m, C^5-*H*$_2$), 1.54 (4H, tt, *J* 5.6, 5.6, C^8-*H*$_2$, C^{10}-*H*$_2$), 1.43–1.35 (2H, m, C^9-*H*$_2$), 1.23 (3H, t, *J* 7.2, C^1-*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 173.7 (C^3), 60.3 (C^2), 58.6, 32.6 (C^4, C^6), 54.7 (C^7, C^{11}), 26.1 (C^8, C^{10}), 24.6 (C^9), 22.4 (C^5), 14.4 (C^1); HRMS *m/z* (ESI$^+$) (Found [M+H]$^+$

200.16463, $C_{11}H_{21}NO_2$ requires $[M+H]^+$ 200.16451); LRMS *m/z* (ESI$^-$) 244 ($[M+FA–H]^-$, 100%); No HPLC data obtained as compound is not UV active.

_N_-[3-(2-Methylpyrimidin-4-yl)phenyl]-4-(piperidin-1-yl)butanamide (195)



To a solution of ethyl 4-(piperidin-1-yl)butanoate (LS B13) (597.7 mg, 3.0 mmol, 3.0 eq.) in MeOH (6.0 mL) was added aqueous 6 M NaOH solution (0.8 mL). The reaction mixture was stirred under reflux for 2 hours. After this time, the solvent was removed in vacuo and the remaining solution was lyophilised to afford 4-(piperidin-1-yl)butanoic acid which was used without further purification.

To a solution of the crude 4-(piperidin-1-yl)butanoic acid in dry toluene (3.0 mL) under argon atmosphere, thionyl chloride (1.1 mL, 15.0 mmol, 15.0 eq.) was added. The reaction solution was stirred at 70 °C for 2.5 hours. After this time, the solvent was removed and CH_2Cl_2 (4.0 mL) was added. The solution was cooled to 0 °C and 4-(3-aminophenyl)-2-methylpyrimidine (185.9 mg, 1.0 mmol, 1.0 eq.) in CH_2Cl_2 (2.0 mL) was added dropwise while stirring at 0 °C. Triethylamine (418 µL, 3.0 mmol, 3.0 eq.) was then added and the reaction solution was allowed to warm to room temperature overnight. After 18 hours, a saturated aqueous solution of $NaHCO_3$ (10 mL) was added and the aqueous layer was extracted with $CHCl_3$ (3 × 20 mL). The organic components were then washed with brine (20 mL), dried over $MgSO_4$ and filtered. The solvent was removed *in vacuo* and the crude material was purified by flash column chromatography on silica gel (EtOAc/EtOH/Et_3N 8.5:1:0.5 to 7.5:2:0.5 gradient) to afford the title compound as a yellow-brown oil (155.1 mg, 0.46 mmol, 46% yield): R_f 0.22 (EtOAc/EtOH/Et_3N 8:1.5:0.5); $\bar{\nu}_{max}$/cm^{-1} 3274 (N-H, w), 2932 (w), 2853 (w), 2804 (w), 1665 (C=O, m), 1571 (s), 1547 (s), 1490 (m), 1434 (m), 1304 (w),

1121 (w), 798 (w), 772 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.66 (1H, br s, N-*H*), 8.65 (1H, d, *J* 5.4, C^4-*H*), 8.24 (1H, dd, *J* 2.1, 1.6, C^7-*H*), 7.81 (1H, ddd, *J* 8.0, 2.1, 1.1, C^9-*H*), 7.75 (1H, ddd, *J* 8.0, 1.6, 1.1, C^{11}-*H*), 7.50 (1H, d, *J* 5.4, C^3-*H*), 7.43 (1H, dd, *J* 8.0, 8.0, C^{10}-*H*), 2.77 (3H, s, C^1-H_3), 2.53–2.43 (8H, m, C^{13}-H_2, C^{15}-H_2, C^{16}-H_2, C^{20}-H_2), 1.91 (2H, tt, *J* 6.4, 6.4, C^{14}-H_2), 1.69 (4H, tt, *J* 5.9, 5.9, C^{17}-H_2, C^{19}-H_2), 1.52 (2H, tt, *J* 5.9, 5.9, C^{18}-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.0 (C^{12}), 168.4 (C^2), 163.7 (C^5), 157.6 (C^4), 139.5 (C^8), 137.7 (C^6), 129.7 (C^{10}), 122.4 (C^{11}), 122.3 (C^9), 118.5 (C^7), 114.1 (C^3), 57.7, 36.4 (C^{13}, C^{15}), 54.5 (C^{16}, C^{20}), 26.4 (C^1), 26.0 (C^{17}, C^{19}), 24.5 (C^{18}), 22.2 (C^{14}); HRMS *m/z* (ESI^+) (Found $[M+H]^+$ 339.21771, $C_{20}H_{26}N_4O$ requires $[M+H]^+$ 339.21794); LRMS *m/z* (ESI^-) 337 ($[M-H]^-$, 100%); RP-HPLC (TFA), retention time = 5.6 min, 99.9%.

8.3 General Biology Methods

8.3.1 Proteins and Plasmids

His_6-TRIM24 protein was provided by Dr Angelina Measures. The pNIC28-Bsa4 TRIM24, pNIC28-Bsa4 TRIM33α and pNIC28-Bsa4 TRIM33β plasmid constructs were provided by Dr Angelina Sekirnik.

8.3.2 Protein Concentration Measurement

Protein concentration was measured using Thermo Scientific Nanodrop Lite (TRIM24, $\varepsilon = 25160\ M^{-1}\ cm^{-1}$; TRIM33α, $\varepsilon = 26650\ M^{-1}\ cm^{-1}$; TRIM33β, $\varepsilon = 25160\ M^{-1}\ cm^{-1}$).

8.3.3 pH Measurement

The pH of buffers were determined using a VWR pH 100 pH meter. Calibration was carried out between pH 7.0–10.0 before use, with buffer solutions of phosphate (pH 7.0) and borate (pH 10.0) from Fisher Scientific. Electrodes were stored in a phthalate buffer solution (pH 4.0).

8.4 Biology Experimental Procedures

8.4.1 DNA Cloning

The DNA fragment TRIM33β-E984C was purchased from ThermoFisher Life Technologies through GeneArt® gene synthesis and diluted to 20 ng/μL.

8.4.1.1 *DNA Amplification – Polymerase Chain Reaction (PCR)*

The DNA fragment TRIM33β-E984C was amplified using Q5® High-Fidelity DNA Polymerase (New England Biolabs) following the recommended procedure.

Component	50 μL Reaction
5× Q5 reaction buffer	10.0 μL
10 mM dNTPs	1.00 μL
10 μM forward primer	2.50 μL
10 μM reverse primer	2.50 μL
Template DNA (20 ng/μL)	2.00 μL
Q5 High-Fidelity DNA Polymerase	0.50 μL
Nuclease-free water (autoclaved)	31.5 μL

The mixture was incubated in a Labnet Multigene II Personal Thermal Cycler, with an initial denaturation step at 98 °C for 3 minutes followed by 30 cycles of denaturing at 98 °C for 30 seconds, followed by annealing at 63 °C ((T_m of primer, calculated using NEB T_m calculator) + 3 °C) for 10 seconds, extension at 72 °C for 21 seconds (30 seconds/kb, TRIM33β-E984C DNA fragment 0.718 kb), then final extension at 72 °C for 2 minutes.

8.4.1.2 *DNA Purification – Agarose Gel Electrophoresis*

The resulting PCR mixture was analysed and purified by gel electrophoresis of the DNA mixture using 1% agarose in TAE buffer (50 mL) diluted from 50× TAE buffer with 5 μL SYBR Safe™ DNA gel stain (Invitrogen) added. The sample was loaded after mixing with 6× gel loading dye and run in TAE buffer on a Bio-Rad system at 100 V for 25 minutes. The

gel was visualised using a UV transilluminator (Bio-Rad Universal Hood III Transilluminator) and the size of the DNA fragment was referenced to a 1 kb GeneRuler DNA Ladder (ThermoFisher). The desired DNA fragment was cut out and the DNA was extracted using a GeneJET Gel Extraction Kit (ThermoFisher) according to the manufacturer instructions. DNA concentration was measured using Thermo Scientific Nanodrop Lite with the DNA programme.

8.4.1.3 *DNA Assembly – HiFi DNA Assembly Kit*

Cloning was performed using a one-step isothermal *in vitro* recombination reaction consisting of the DNA fragment (0.08 pmole), linearised pNIC28-Bsa4 vector (0.04 pmole provided by Dr Matthias Schiedel), NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) (10 µL) and deionised water to make up a total volume of 20 µL. The mixture was incubated in a Labnet Multigene II Personal Thermal Cycler at 50 °C for 1 hour. 5 µL of the resulting mixture was directly used to transform *Escherichia coli* NEB5α competent cells provided in the HiFi DNA Assembly Kit.

Competent cells were thawed on ice before addition of plasmid DNA was added. The mixture was flicked 10 times then incubated on ice for 20 minutes. The cells were then placed in 42 °C (heat shock treatment) for 30 seconds then returned to ice for 5 minutes. 500 µL SOC outgrowth media was added before incubating the cells at 37 °C for 1 hour. The bacteria (100 µL and 400 µL) were then streaked on solid 2× TY medium with 50 µg/mL kanamycin under aseptic conditions using sterile equipment. The plate cultures were inverted and grown overnight at 37 °C, then stored at 4 °C.

8.4.1.4 *Plasmid DNA Isolation*

A small-scale culture (100 mL of 2× TY media containing 50 µg/mL kanamycin in 250 mL baffled flask), was inoculated with a single colony from the agar plate and incubated at 37 °C overnight. The plasmid was isolated using a GeneJET Plasmid MiniPrep Kit (ThermoFisher) according to the manufacturer instructions. Plasmid sequencing was then performed by

Source BioScience Sequencing facility in the Department of Biochemistry (University of Oxford) to confirm the DNA sequence. The T7 promoter sequence primer was used. DNA sequence was analysed on CLC Main Workbench (v7.0).

8.4.2 Protein Production

8.4.2.1 *Bacterial Transformation*

Escherichia coli BL21 (DE3) competent cells were thawed on ice before addition of plasmid DNA was added. The mixture was flicked 10 times then incubated on ice for 20 minutes. The cells were then placed in 42 °C (heat shock treatment) for 30 seconds then returned to ice for 5 minutes. 500 µL SOC outgrowth media was added before incubating the cells at 37 °C for 1 hour. The bacteria (10 µL and 40 µL) were then streaked on solid 2× TY medium with 50 µg/mL kanamycin under aseptic conditions using sterile equipment. The plate cultures were inverted and grown overnight at 37 °C, then stored at 4 °C.

8.4.2.2 *Protein Expression*

Media	Reagent	Amount required (g)
2× TY (1 L)	Tryptone	16.0
	Yeast extract	10.0
	NaCl	5.00
	Agar bacteriological (Oxoid Ltd)	15.0

All growth media were autoclaved (*Thermo Live Sciences* MAT90 LEI) at 121 °C before use. Small-scale growths for starter cultures (100 mL of 2× TY media containing kanamycin in 500 mL baffled flask), were inoculated with a single colony from an agar plate and incubated at 37 °C overnight.

Large-scale recombinant protein expression was performed using 2000 mL PYREX® narrow-mouth baffled Tunair™ flasks containing 1000 mL of 2× TY medium containing 50 µg/mL kanamycin. Flasks were inoculated with 10 mL (1% by volume) of the bacterial

starter culture and grown at 37 °C (180 rpm.) until they reached an OD_{600} of 1.0. Optical density readings were taken in 1.0 mL cuvettes with 2× dilution in 2× TY against a reference sample of the 2× TY growth media. The absorbance was measured at 600 nm using a Novaspec® II spectrophotometer. After reaching an optical density of OD_{600} = 1.0, cultures were equilibrated to 16 °C and induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) (0.1 mM) and $ZnCl_2$ (0.2 mM), before incubation overnight. After 21 hours post-IPTG induction, cultures were centrifuged at 11325 ×g, 4 °C for 5 minutes (Rotor: Beckman F10-6 × 500y, Avanti™ J-25 centrifuge), and the resulting bacterial pellet frozen in a resealable plastic bag at –80 °C.

8.4.2.3 *Protein Purification*

Reagents	IMAC buffers (pH 7.6)			SEC buffer (pH 7.6)
	Binding	**Elution**	**Strip**	**Gel filtration**
HEPES	50.0 mM	50.0 mM	50.0 mM	50.0 mM
NaCl	500 mM	500 mM	500 mM	500 mM
Glycerol	10% *v/v*	10% *v/v*	–	–
Imidazole	–	500 mM	–	–
EDTA	–	–	50 mM	–

All buffers were filtered under vacuum through a 0.2 μm filter paper (Sartorius UK) before use and stored at 4 °C.

Cell pellets were resuspended in Extraction Buffer (1 Sigma*FAST*™ protease inhibitor cocktail tablet EDTA free (Sigma Aldrich) dissolved in 100 mL Binding Buffer), 100 mL per 20 g cell mass. 250 U Benzonase® nuclease (Novagen®) was added, and the cells were lysed by sonication (50% amplitude), applying thirty cycles of 5 second bursts interrupted by 5 second pauses (Fisherbrand™ Q500 Sonicator). Polyethyleneimine (Sigma Aldrich) (0.15% *v/v*) was added from a 5% (*v/v*) stock and the mixture was incubated on ice for

15 minutes to precipitate DNA. The lysate was clarified by centrifugation and the supernatant decanted, filtered through a 0.45 μm syringe filter (Merck Millipore), then supplemented with 10 mM β-mercaptoethanol (β-ME) (Bio-Rad Laboratories) and 20 mM imidazole.

The protein was purified at 4 °C by immobilised metal affinity chromatography (IMAC) followed by size-exclusion chromatography (SEC) operated through a computerised ÄKTA fast protein liquid chromatography (ÄKTAFPLC) system (GE Healthcare). IMAC was performed with a HisTrap™ HP 5 mL column (GE Healthcare) rinsed with 50 mL (10 CV) Milli-Q water, charged with 10 mL (2 CV) of 100 mM $NiSO_4$ and equilibrated with 50 mL (10 CV) of Binding Buffer (supplemented with 10 mM β-ME) prior to loading the cell lysate at 1 mL/min. After loading the sample onto the column, the column was washed with a step gradient of 50 mL (10 CV) 9% Elution Buffer in Binding Buffer (i.e. 45 mM imidazole) and 50 mL (10 CV) 12% Elution Buffer in Binding Buffer (i.e. 60 mM imidazole). The protein of interest was then batch-eluted with 20 mL (4 CV) 60% Elution Buffer in Binding Buffer (i.e. 300 mM imidazole) followed by 10 mL (2 CV) 100% Elution Buffer (i.e. 500 mM imidazole), and was collected in 2 mL fractions. Following this, 50 mL (10 CV) Milli-Q water was used to wash the column, then 50 mL (10 CV) of Strip Buffer was used to strip the bound Ni^{2+} ions off the affinity column. The column was stored in this condition at 4 °C until the next use.

Further purification was achieved by SEC. The desired fractions were spin concentrated using Corning® Spin-X® UF Centrifugal Concentrator, 10k MWCO in a Beckman Coulter Allegra X-12R benchtop centrifuge at 3270 ×g, 4 °C until it reached 4 or 6 mL. 2 mL of the concentrated protein solution was loaded *via* a 2 mL injection loop onto a gel filtration column XK 16/70 (GE Healthcare) packed with 120 mL Superdex 75 resin (GE Healthcare) that had been rinsed with 150 mL degassed Milli-Q water and equilibrated with 150 mL degassed gel filtration buffer. Protein was eluted with 150 mL degassed gel filtration buffer with a flow rate of 1 mL/min and collected in 5 mL fractions. Fraction collection was started

at the 20th mL and stopped at the 120th mL. Fractions containing the highest levels of pure protein based on the SEC trace were spin concentrated using Corning® Spin-X® UF Centrifugal Concentrator, 10k MWCO in a Beckman Coulter Allegra X-12R benchtop centrifuge at 3270 ×g, 4 °C then aliquoted, frozen and stored at –80 °C.

Protein concentration was estimated by measuring the absorbance at 280 nm using Thermo Scientific Nanodrop Lite and calculated using the Beer-Lambert law (TRIM33α, $\varepsilon = 256650\ M^{-1}\ cm^{-1}$; TRIM33β, $\varepsilon = 25160\ M^{-1}\ cm^{-1}$; TRIM33β-N1039F, $\varepsilon = 25160\ M^{-1}\ cm^{-1}$). Protein mass spectrometry was carried out on a Waters LCT Premier mass spectrometer diluted to 0.02 mg/mL and 0.1 mg/mL with methanol or gel filtration buffer. Protein molecular weight and molar extinction coefficient were calculated using ProtParam on the ExPASy Bioinformatics Resource Portal.

8.4.2.4 *Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis*

Gel composition (volume for 2 gels):

Reagents	Resolving Gel (16%)	Stacking Gel (4%)
Milli-Q water	2.10 mL	2.90 mL
1.5 M Tris-HCl, pH 8.8	2.50 mL	–
0.5 M Tris-HCl, pH 6.8	–	1.30 mL
30% Acrylamide/bis	5.30 mL	0.70 mL
10% Sodium dodecyl sulfate (SDS)	100 µL	50.0 µL
TEMED	8.00 µL	5.00 µL
10% APS	100 µL	50.0 µL

Protein purity was evaluated using SDS-PAGE 16% with Coomassie Blue staining. Gels were prepared using 70 mm × 100 mm glass plates, with 0.75 mm spacers. TEMED and 10% ammonium persulfate (APS) were added just before casting gels. The resolving gel was cast with the addition of a layer of methanol (to ensure a level surface). Once the resolving layer

had polymerized, the methanol was removed and the boundary rinsed with Milli-Q water before casting the stacking gel. For protein denaturation, samples were prepared by mixing with Laemmli sample loading buffer (4×, with addition of 10% β-ME), and incubating at 100 °C for 10 min. Gels were loaded with colour protein standard, broad range (NEB) molecular weight ladder then run on a mini-PROTEAN Tetra Electrophoresis System (Bio-Rad) at a constant potential of 100 V. Following electrophoresis, gels were stained for 15 minutes with InstantBlueTM Coomasie® stain (Expedeon) and afterwards rinsed with Milli-Q water then destained multiple times with destaining solution (40% methanol and 10% acetic acid in Milli-Q water).

8.4.3 AlphaScreen™

8.4.3.1 *Peptides*

Peptides were supplied as >95% purity (Cambridge Research Biochemicals or GenScript) and dissolved to 1 mM and 0.1 mM stocks in Milli-Q water before being stored at –80 °C in 5 µL aliquots.

Peptide	Amino acid sequence
$H3_{1-27}K9Me_3$	H_2N-ARTKQTARK(Me_3)STGGKAPRKQLATKAARKY-CO_2H
$H3_{1-27}K18Ac$	H_2N-ARTKQTARKSTGGKAPRK(Ac)QLATKAARKY-CO_2H
$H3_{1-27}K9Me_3K18Ac$	H_2N-ARTKQTARK(Me_3)STGGKAPRK(Ac)QLATKAARKY-CO_2H

8.4.3.2 *Typical Procedure for AlphaScreen™*

HEPES buffer (pH 7.6) composition (50 mL):

Reagent	Composition	Mass (mg)
HEPES	25 mM	298
NaCl	100 mM	292
CHAPS	0.05 % (*w/v*)	25
Bovine serum albumin (BSA)	0.1 % (*w/v*)	50

The assay buffer was stored as a 5× solution (without BSA) and stored at 4 °C until use. BSA was freshly added and the buffer filter sterilised through a 0.22 μm filter before use. For incubation steps, the plate was sealed, shaken for 10 seconds at 600 rpm on a plate oscillator and kept at room temperature in the dark. Compounds, protein, and peptides were dispensed into a ProxiPlate-384 Plus (Perkin Elmer) using a Thermo-Fisher electronic multi-channel pipette. The plate was read using an Envision 2104 Multilabel Plate Reader (Perkin Elmer) using the built-in AlphaScreen™ 384 ProxiPlate function: excitation 680 nm, 0.18 sec; emission 570 nm, 0.37 sec. Ni^{2+} chelate acceptor and streptavidin donor beads were prepared as a mixture in a 1:300 dilution (0.007 mg/mL final assay concentration (FAC)).

8.4.3.3 *Signal Response Experiment*

Proteins (7 μL) were dispensed into wells as 1:2 serial dilutions from 100 nM FAC for TRIM24 and 500 nM FAC for TRIM33α and TRIM33β. Peptides (5 μL) were dispensed as 1:2 serial dilutions from 250 nM FAC. Assay beads (8 μL) were added to attain a final assay volume of 20 μL.

8.4.3.4 *Single-Concentration Screening Experiments*

Compounds (5 μL) were dispensed into wells from 50 mM DMSO stock solutions and screened in triplicate at 30 μM and 150 μM FAC. Protein-peptide mixes (7 μL) for each

combination were prepared following the specified FAC and then dispensed into wells. Assay beads (8 µL) were added to attain a final assay volume of 20 µL. Data were processed by calculating the percentage decrease in emission intensity with respect to the average of the DMSO control wells on each respective plate, which were set as 100%.

Protein	Peptide	Peptide FAC (nM)	Protein FAC (nM)
TRIM 24	$H3K9Me_3$	4	100
	H3K18Ac	8	25.0
	$K3K9Me_3K18Ac$	4	25.0
TRIM33α	$H3K9Me_3$	15	500
	H3K18Ac	15	500
	$K3K9Me_3K18Ac$	8	250
TRIM33β	$H3K9Me_3$	8	500
	H3K18Ac	15	125
	$K3K9Me_3K18Ac$	8	125

8.4.3.5 *Dose Response Experiments*

Compounds (5 µL) were dispensed into wells as serial dilutions from the specified concentrations from 50 mM DMSO stock solutions. Protein-peptide mixes (7 µL) for each combination were prepared following the specified FAC and then dispensed into wells. Assay beads (8 µL) were added to attain a final assay volume of 20 µL. Dose response curves were obtained for compounds in triplicate, with serial 1:2 dilutions. Data were processed by fitting a four-parameter equation to calculate IC_{50} values using Prism software.

8.4.3.6 *TruHits Experiments*

Compounds (5 µL) were dispensed into wells as serial 1:2 dilutions from the specified concentrations from 50 mM DMSO stock solutions. Assay beads (15 µL) were pre-incubated for 30 minutes and added to the first plate to attain a final assay volume of 20 µL, and then

incubated for a further 10 minutes. To the second plate, streptavidin donor beads (7.5 µL) were added to the compounds and incubated for 30 minutes. Biotinylated acceptor beads (7.5 µL) were then added to the second plate to attain a final assay volume of 20 µL, and then incubated for a further 30 minutes. Data were processed by calculating the percentage decrease in emission intensity with respect to the average of the DMSO control wells on each respective plate, which were set as 100%.

8.4.4 NMR WaterLOGSY

10 mM Na/K phosphate buffer (PBS) buffer (pH 7.6) composition:

Reagent	Amount required per litre (g)
KH_2PO_4	0.26
Na_2HPO_4	1.15
NaCl	8.71

The buffer was filtered under vacuum through a 0.2 µm filter paper (Sartorius UK) before use and stored at 4 °C.

Proteins were buffer exchanged from HEPES storage buffer into PBS buffer (0.9% NaCl, pH 7.6) using GE Healthcare PD SpinTrap G-25 columns. For each compound, a water-suppressed ^1H spectrum was recorded as a reference, followed by a waterLOGSY spectrum of the compound in the absence of the protein, and then another waterLOGSY spectrum in the presence of the specified protein. Samples were prepared to a total sample volume of 160 µL in PBS buffer (pH 7.6), with 500 µM compound (unless otherwise specified) from 50 mM DMSO-D_6 stock solutions, 10% D_2O for signal locking, and 20 µM protein (unless otherwise specified) for samples with protein. After loading into 3 mm Hilgenberg NMR tubes for Bruker Match™ system, samples were briefly centrifuged in a

manual benchtop rotor. Spectra were recorded on a Bruker AVIII 700 MHz spectrometer with inverse TCI cryoprobe using a standard waterLOGSY pulse sequence.[96]

8.4.5 ITC Experiments

ITC HEPES buffer (pH 7.4) composition:

Reagent	Composition	Amount required per litre (g)
HEPES	50 mM	11.92
NaCl	150 mM	8.77

The buffer was filtered under vacuum through a 0.2 μm filter paper (Sartorius UK) before use and stored at 4 °C.

8.4.5.1 *Protein-Peptide ITC*

All ITC experiments were performed on a MicroCal PEAQ-ITC (Malvern) and analysed with the MicroCal software package using a single binding site model. Proteins and peptides were dialysed at 4 °C overnight in a Slide-A-Lyzer™ Mini Dialysis Device (2K MWCO, Thermo Fisher Scientific) into the ITC HEPES buffer (pH 7.6). Concentrations after dialysis were calculated by measuring the absorbance at 280 nM using Nanodrop Lite (Thermo Fisher Scientific). Samples were then diluted to the concentrations required for titration using the dialysis buffer. The cell was stirred at 750 rpm at 25 °C, with reference power set to 4 μcal/sec. after an initial delay of 60 seconds, the first injection (0.2 μL, discarded in data analysis) was performed and followed by 18 × 2 μL injections with a spacing of 180 seconds. Heats of dilution were measured but not used in data analysis. This is because sufficient saturation was achieved by the titration and the heats of dilution at the end of the experiment were used to offset the titration. Protein solutions in the calorimetric cell (250 μL) were titrated with the peptide solutions in the syringe (60 μL) at the following

concentrations: 160 μM protein and 16 μM H3K18Ac; 160 μM protein and 16 μM H3K9Me$_3$K18Ac; 500 μM protein and 50 μM H3K9Me$_3$.

8.4.5.2 *Protein-Ligand ITC*

All ITC experiments were performed on a MicroCal PEAQ-ITC Automated (Malvern) and analysed with the MicroCal software package using a single binding site model. Proteins and peptides were dialysed at 4 °C overnight in a Slide-A-Lyzer™ Mini Dialysis Device (2K MWCO, Thermo Fisher Scientific) into the ITC HEPES buffer supplemented with 0.2% DMSO (pH 7.6). Concentrations after dialysis were calculated by measuring the absorbance at 280 nM using Nanodrop Lite (Thermo Scientific). Samples were then diluted to the concentrations required for titration using the dialysis buffer. The cell was stirred at 750 rpm at 25 °C, with reference power set to 4 μcal/sec. after an initial delay of 60 seconds, the first injection (0.2 μL, discarded in data analysis) was performed and followed by 18 × 2 μL injections with a spacing of 180 seconds. Heats of dilution were measured but not used in data analysis. This is because sufficient saturation was achieved by the titration and the heats of dilution at the end of the experiment were used to offset the titration. Protein solutions in the calorimetric cell (250 μL) were titrated with the peptide solutions in the syringe (60 μL) at the following concentrations: 160 μM protein and 20 μM compound (diluted from 10 mM DMSO stocks).

8.4.6 Circular Dichroism Spectroscopy

Circular dichroism measurements were acquired using an Applied Photophysics Chirascan spectrometer with a Peltier temperature-controlled cell holder. All experiments were performed in a 1 mm path length cuvette using 0.2 mg/mL protein in 10 mM sodium phosphate buffer (pH 7.5). Data was recorded from 185–240 nm, at 0.5 nm intervals, and each data point was averaged for 3 seconds.

8.4.7 Alkylation Experiments

Buffer	Reagent	Composition	Amount required (mg)
Alkylation buffer (pH 7.4, 50 mL)	HEPES	25 mM	297.9
	NaCl	100 mM	292.2
Digestion buffer (pH 8.0, 50 mL)	Tris	100 mM	605.7
	$CaCl_2$	10 mM	55.5

The buffers were filtered under vacuum through a 0.22 μm syringe filter (Merck Millipore) before use and stored at 4 °C.

8.4.7.1 *Alkylation Experiments with N-Ethylmaleimide for Proteomics*

Alkylation experiments were performed in Eppendorf tubes in alkylation buffer and incubated in a 37 °C water bath. Experiments were carried out with 50 μM (1.4 mg/mL) TRIM33β and 250 μM *N*-ethylmaleimide (from 25 mM DMSO stocks), and blown over with nitrogen. After 1 hour, a small amount of reaction sample was removed, diluted to 0.1 mg/mL and submitted to mass spectrometry analysis on a Waters LCT Premier mass spectrometer. The remaining protein was buffer exchanged from alkylation buffer into digestion buffer using GE Healthcare PD SpinTrap G-25 columns to remove excess *N*-ethylmaleimide. The protein concentration was ~1.0 mg/mL after buffer exchange. 5 μL of sample was submitted to MALDI-TOF analysis.

To prepare for protein digest, alkylated TRIM33β was reduced with 10 mM dithiothreitol (DTT, from 500 mM aqueous stock solution) and incubated at 60 °C for 45 minutes. Chymotrypsin (Thermo Fisher, reconstituted using 1 M HCl to 1 mg/mL) was added in protein:protease 1:20 ratio and incubated at 37 °C overnight. The resulting mixture was cleaned up using Pierce C18 Spin Columns (Thermo Fisher), eluting with 70% MeCN + 0.1% formic acid, and submitted to Nano LC-MS/MS analysis.

8.4.7.2 *Alkylation Experiments with Covalent Ligands*

Alkylation experiments were performed in Eppendorf tubes and incubated in a 37 °C water bath. Experiments were carried out with 10 μM protein (TRIM33β or TRIM33β-E984C) and 50 μM compound (from 50 mM DMSO stocks), and blown over with nitrogen. At defined time points, a small amount of reaction sample was removed, diluted to 0.1 mg/mL and submitted to mass spectrometry analysis on a Waters LCT Premier mass spectrometer.

8.4.7.3 *Alkylation Experiments with Covalent Ligands for Proteomics*

Alkylation experiments were performed in Eppendorf tubes and incubated in a 37 °C water bath. Experiments were carried out with 10 μM protein (TRIM33β or TRIM33β-E984C) and 50 μM compound (from 50 mM DMSO stocks), and blown over with nitrogen. After 2 hours, a small amount of reaction sample was removed, diluted to 0.1 mg/mL and submitted to mass spectrometry analysis on a Waters LCT Premier mass spectrometer. The remaining protein was buffer exchanged from alkylation buffer into digestion buffer using GE Healthcare PD SpinTrap G-25 columns to remove excess ligand. The protein concentration was ~1.0 mg/mL after buffer exchange. 5 μL of sample was submitted to MALDI-TOF analysis.

To prepare for protein digest, alkylated TRIM33β was reduced with 10 mM dithiothreitol (DTT, from 500 mM aqueous stock solution) and incubated at 60 °C for 45 minutes. Chymotrypsin (Thermo Fisher, reconstituted using 1 M HCl to 1 mg/mL) was added in protein:protease 1:20 ratio and incubated at 37 °C overnight. The resulting mixture was cleaned up using Pierce C18 Spin Columns (Thermo Fisher), eluting with 70% MeCN + 0.1% formic acid, and submitted to Nano LC-MS/MS analysis.

8.5 Docking

Docking was performed using CLC Drug Discovery Workbench (v2.5). The crystal structure of TRIM33β (unpublished) was provided by Prof. Stefan Knapp and protonated at physiological pH within CLC. All binding site crystallographic waters were retained. The

region for docking was specified as a 13 Å sphere centred on the end of side chain E981. Docking was interpreted using the CLC score and manually inspected. The binding poses were filtered based on knowledge of headgroup binding from TRIM24.[61,60] For docking with TRIM33β-E984C, the x-ray crystal structure was prepared from that of TRIM33β provided by Prof. Stefan Knapp, with introduction of the E984C mutation using the PyMOL mutagenesis tool.

References

1 C. H. Waddington, *Int. J. Epidemiol.*, 2012, **41**, 10–13.

2 S. L. Berger, T. Kouzarides, R. Shiekhattar and A. Shilatifard, *Genes Dev.*, 2009, **23**, 781–783.

3 A. D. Goldberg, C. D. Allis and E. Bernstein, *Cell*, 2007, **128**, 635–638.

4 K. D. Robertson, *Nat. Rev. Genet.*, 2005, **6**, 597–610.

5 A. Breiling and F. Lyko, *Epigenetics and Chromatin*, 2015, **8**, 1–9.

6 B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology of the Cell, 4th edition*, Garland Science, 4th edn., 2002.

7 T. Jenuwein, *Science (80-.).*, 2001, **293**, 1074–1080.

8 A. Lennartsson and K. Ekwall, *Biochim. Biophys. Acta*, 2009, **1790**, 863–868.

9 M. Tan, H. Luo, S. Lee, F. Jin, J. S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, Z. Lu, Z. Ye, Q. Zhu, J. Wysocka, Y. Ye, S. Khochbin, B. Ren and Y. Zhao, *Cell*, 2011, **146**, 1016–1028.

10 M. Z. Atassi, *Protein Reviews*, Springer Singapore, Singapore, 2017, vol. 18.

11 V. W. Zhou, A. Goren and B. E. Bernstein, *Nat. Rev. Genet.*, 2011, **12**, 7–18.

12 M. Grunstein, *Nature*, 1997, **389**, 349–352.

13 A. J. Bannister and T. Kouzarides, *Cell Res.*, 2011, **21**, 381–395.

14 B. D. Strahl and C. D. Allis, *Nature*, 2000, **403**, 41–45.

15 P. Filippakopoulos and S. Knapp, *Nat. Rev. Drug Discov.*, 2014, **13**, 337–356.

16 L. E. Jennings, A. R. Measures, B. G. Wilson and S. J. Conway, *Future Med. Chem.*, 2014, **6**, 179–204.

17 P. A. Jones and S. B. Baylin, *Nat. Rev. Genet.*, 2002, **3**, 415–428.

18 P. Bégin and K. C. Nadeau, *Allergy, Asthma Clin. Immunol.*, 2014, **10**, 27.

19 G. S. Hailu, D. Robaa, M. Forgione, W. Sippl, D. Rotili and A. Mai, *J. Med. Chem.*, 2017, **60**, 4780–4804.

20 I. Ali, R. J. Conrad, E. Verdin and M. Ott, *Chem. Rev.*, 2018, **118**, 1216–1252.

21 Z. Su and J. M. Denu, *ACS Chem. Biol.*, 2016, **11**, 564–574.

22 W. Fischle, S. T. Boo, H. L. Dormann, B. M. Ueberheide, B. A. Garcia, J. Shabanowitz, D. F. Hunt, H. Funabiki and C. D. Allis, *Nature*, 2005, **438**, 1116–1122.

23 Z. Wang and D. J. Patel, *J. Biol. Chem.*, 2011, **286**, 18363–18368.

24 I. O. Torres, K. M. Kuchenbecker, C. I. Nnadi, R. J. Fletterick, M. J. S. Kelly and D. G. Fujimori, *Nat. Commun.*, 2015, **6**, 1–10.

25 M. Brand, A. M. Measures, B. G. Wilson, W. A. Cortopassi, R. Alexander, M. Höss, D. S. Hewings, T. P. C. Rooney, R. S. Paton and S. J. Conway, *ACS Chem. Biol.*, 2015, **10**, 22–39.

26 P. Filippakopoulos, S. Picaud, M. Mangos, T. Keates, J. P. Lambert, D. Barsyte-Lovejoy, I. Felletar, R. Volkmer, S. Müller, T. Pawson, A. C. Gingras, C. H. Arrowsmith and S. Knapp, *Cell*, 2012, **149**, 214–231.

27 D. S. Hewings, T. P. C. Rooney, L. E. Jennings, D. A. Hay, C. J. Schofield, P. E. Brennan, S. Knapp and S. J. Conway, *J. Med. Chem.*, 2012, **55**, 9393–9413.

28 S. Muller, P. Filippakopoulos and S. Knapp, *Expert Rev. Mol. Med.*, 2011, **13**, e29.

29 R. Sanchez and M.-M. Zhou, *Curr. Opin. Drug Discov. Devel.*, 2009, **12**, 659–665.

30 C. A. Musselman and T. G. Kutateladze, *Mol. Interv.*, 2009, **9**, 314–323.

31 R. Sanchez and M.-M. Zhou, *Trends Biochem. Sci.*, 2011, **36**, 364–372.

32 J. E. Beaver and M. L. Waters, *ACS Chem. Biol.*, 2016, **11**, 643–653.

33 M. Watanabe and S. Hatakeyama, *J. Biochem.*, 2017, **161**, 135–144.

34 K. Ozato, D.-M. Shin, T.-H. Chang and H. C. Morse, *Nat. Rev. Immunol.*, 2008, **8**, 849–860.

35 S. Hatakeyama, *Trends Biochem. Sci.*, 2017, **42**, 297–311.

36 B. Herquel, K. Ouararhni, K. Khetchoumian, M. Ignat, M. Teletin, M. Mark, G. Bechade, A. Van Dorsselaer, S. Sanglier-Cianferani, A. Hamiche, F. Cammas, I. Davidson and R. Losson, *Proc. Natl. Acad. Sci.*, 2011, **108**, 8212–8217.

37 W.-W. Tsai, Z. Wang, T. T. Yiu, K. C. Akdemir, W. Xia, S. Winter, C. Tsai, X. Shi, D. Schwarzer, W. Plunkett, B. Aronow, O. Gozani, W. Fischle, M. Hung, D. J. Patel and M. C. Barton, *Nature*, 2010, **468**, 927–932.

38 A. C. Groner, L. Cato, J. de Tribolet-Hardy, T. Bernasocchi, H. Janouskova, D. Melchers, R. Houtman, A. C. B. Cato, P. Tschopp, L. Gu, A. Corsinotti, Q. Zhong, C. Fankhauser, C. Fritz, C. Poyet, U. Wagner, T. Guo, R. Aebersold, L. A. Garraway, P. J. Wild, J.-P. Theurillat and M. Brown, *Cancer Cell*, 2016, 1–13.

39 E. Agricola, R. A. Randall, T. Gaarenstroom, S. Dupont and C. S. Hill, *Mol. Cell*, 2011, **43**, 85–96.

40 Q. Xi, Z. Wang, A. Zaromytidou, X. H.-F. Zhang, L.-F. Chow-Tsang, J. X. Liu, H. Kim, A. Barlas, K. Manova-Todorova, V. Kaartinen, L. Studer, W. Mark, D. J. Patel and J. Massagué, *Cell*, 2011, **147**, 1511–1524.

41 R. M. Pommier, J. Gout, D. F. Vincent, L. B. Alcaraz, N. Chuvin, V. Arfi, S. Martel, B. Kaniewski, G. Devailly, G. Fourel, P. Bernard, C. Moyret-Lalle, S. Ansieau, A. Puisieux, U. Valcourt, S. Sentis and L. Bartholin, *Cancer Res.*, 2015, **75**, 4335–4350.

42 R. Aucagne, N. Droin, J. Paggetti, B. Lagrange, A. Largeot, A. Hammann, A. Bataille, L. Martin, K. P. Yan, P. Fenaux, R. Losson, E. Solary, J. N. Bastie and L. Delva, *J. Clin. Invest.*, 2011, **121**, 2361–2370.

43 A. Kulkarni, J. Oza, M. Yao, H. Sohail, V. Ginjala, A. Tomas-Loba, Z. Horejsi, A. R. Tan, S. J. Boulton and S. Ganesan, *J. Biol. Chem.*, 2013, **288**, 32357–32369.

44 C. Pan, M. Mezei, S. Mujtaba, M. Muller, L. Zeng, J. Li, Z. Wang and M. M. Zhou, *J. Med. Chem.*, 2007, **50**, 2285–2288.

45 S. Mujtaba, Y. He, L. Zeng, A. Farooq, J. E. Carlson, M. Ott, E. Verdin and M. M. Zhou, *Mol. Cell*, 2002, **9**, 575–586.

46 P. Filippakopoulos, J. Qi, S. Picaud, Y. Shen, W. B. Smith, O. Fedorov, E. M. Morse, T. Keates, T. T. Hickman, I. Felletar, M. Philpott, S. Munro, M. R. McKeown, Y. Wang, A. L. Christie, N. West, M. J. Cameron, B. Schwartz, T. D. Heightman, N. La Thangue, C. A. French, O. Wiest, A. L. Kung, S. Knapp and J. E. Bradner, *Nature*, 2010, **468**, 1067–1073.

47 E. Nicodeme, K. L. Jeffrey, U. Schaefer, S. Beinke, S. Dewell, C.-W. Chung, R. Chandwani, I. Marazzi, P. Wilson, H. Coste, J. White, J. Kirilovsky, C. M. Rice, J. M. Lora, R. K. Prinjha, K. Lee and A. Tarakhovsky, *Nature*, 2010, **468**, 1119–1123.

48 D. S. Hewings, M. Wang, M. Philpott, O. Fedorov, S. Uttarkar, P. Filippakopoulos, S. Picaud, C. Vuppusetty, B. Marsden, S. Knapp, S. J. Conway and T. D. Heightman, *J. Med. Chem.*, 2011, **54**, 6761–6770.

49 D. S. Hewings, O. Fedorov, P. Filippakopoulos, S. Martin, S. Picaud, A. Tumber, C. Wells, M. M. Olcina, K. Freeman, A. Gill, A. J. Ritchie, D. W. Sheppard, A. J. Russell, E. M. Hammond, S. Knapp, P. E. Brennan and S. J. Conway, *J. Med. Chem.*, 2013, **56**, 3217–3227.

50 T. Fujisawa and P. Filippakopoulos, *Nat. Rev. Mol. Cell Biol.*, 2017, **18**, 246–262.

51 H. S. Bandukwala, J. Gagnon, S. Togher, J. A. Greenbaum, E. D. Lamperti, N. J. Parr, A. M. H. Molesworth, N. Smithers, K. Lee, J. Witherington, D. F. Tough, R. K. Prinjha, B. Peters and A. Rao, *Proc. Natl. Acad. Sci.*, 2012, **109**, 14532–14537.

52 C. Chung, H. Coste, J. H. White, O. Mirguet, J. Wilde, R. L. Gosmini, C. Delves, S. M. Magny, R. Woodward, S. A. Hughes, E. V. Boursier, H. Flynn, A. M. Bouillot, P. Bamborough, J.-M. G. Brusq, F. J. Gellibert, E. J. Jones, A. M. Riou, P. Homes, S. L. Martin, I. J. Uings, J. Toum, C. A. Clément, A.-B. Boullay, R. L. Grimley, F. M. Blandel, R. K. Prinjha, K. Lee, J. Kirilovsky and E. Nicodeme, *J. Med. Chem.*, 2011, **54**, 3827–3838.

53 L. N. Gechijian, D. L. Buckley, M. A. Lawlor, J. M. Reyes, J. Paulk, C. J. Ott, G. E. Winter, M. A. Erb, T. G. Scott, M. Xu, H.-S. Seo, S. Dhe-Paganon, N. P. Kwiatkowski, J. A. Perry, J. Qi, N. S. Gray and J. E. Bradner, *Nat. Chem. Biol.*, 2018, **14**, 405–412.

54 T. P. C. Rooney, P. Filippakopoulos, O. Fedorov, S. Picaud, W. A. Cortopassi, D. a. Hay, S. Martin, A. Tumber, C. M. Rogers, M. Philpott, M. Wang, A. L. Thompson, T. D. Heightman, D. C. Pryde, A. Cook, R. S. Paton, S. Müller, S. Knapp, P. E. Brennan and S. J. Conway, *Angew. Chemie Int. Ed.*, 2014, **53**, 6126–6130.

55 D. C. Miller, M. P. Martin, S. Adhikari, A. Brennan, J. A. Endicott, B. T. Golding, I. R. Hardcastle, A. Heptinstall, S. Hobson, C. Jennings, L. Molyneux, Y. Ng, S. R. Wedge, M. E. M. Noble and C. Cano, *Org. Biomol. Chem.*, 2018, **16**, 1843–1850.

56 E. H. Demont, P. Bamborough, C. W. Chung, P. D. Craggs, D. Fallon, L. J. Gordon, P. Grandi, C. I. Hobbs, J. Hussain, E. J. Jones, A. Le Gall, A. M. Michon, D. J. Mitchell, R. K. Prinjha, A. D. Roberts, R. J. Sheppard and R. J. Watson, *ACS Med. Chem. Lett.*, 2014, **5**, 1190–1195.

57 D. B. Doroshow, J. P. Eder and P. M. LoRusso, *Ann. Oncol.*, 2017, **28**, 1776–1787.

58 M. Pérez-Salvia and M. Esteller, *Epigenetics*, 2017, **12**, 323–339.

59 L. R. Vidler, N. Brown, S. Knapp and S. Hoelder, *J. Med. Chem.*, 2012, **55**, 7346–7359.

60 W. S. Palmer, G. Poncet-Montange, G. Liu, A. Petrocchi, N. Reyna, G. Subramanian, J. Theroff, A. Yau, M. Kost-Alimova, J. P. Bardenhagen, E. Leo, H. E. Shepard, T. N. Tieu, X. Shi, Y. Zhan, S. Zhao, M. C. Barton, G. Draetta, C. Toniatti, P. Jones, M. Geck Do and J. N. Andersen, *J. Med. Chem.*, 2016, **59**, 1440–1454.

61 J. Bennett, O. Fedorov, C. Tallant, O. Monteiro, J. Meier, V. Gamble, P. Savitsky, G. A. Nunez-Alonso, B. Haendler, C. Rogers, P. E. Brennan, S. Müller and S. Knapp, *J. Med. Chem.*, 2016, **59**, 1642–1647.

62 L. I. James and S. V. Frye, *Curr. Opin. Chem. Biol.*, 2016, **33**, 135–141.

63 N. Milosevich and F. Hof, *Biochemistry*, 2016, **55**, 1570–1583.

64 S. S. Oliver, C. A. Musselman, R. Srinivasan, J. P. Svaren, T. G. Kutateladze and J. M. Denu, *Biochemistry*, 2012, **51**, 6534–6544.

65 E. K. Wagner, N. Nath, R. Flemming, J. B. Feltenberger and J. M. Denu, *Biochemistry*, 2012, **51**, 8293–8306.

66 B. Bhushan, A. Erdmann, Y. Zhang, R. Belle, C. Johannson, U. Oppermann, R. J. Hopkinson, C. J. Schofield and A. Kawamura, *Bioorganic Med. Chem.*, 2018, **26**, 2984–2991.

67 T. C. R. Miller, T. J. Rutherford, K. Birchall, J. Chugh, M. Fiedler and M. Bienz, *ACS Chem. Biol.*, 2014, **9**, 2864–2874.

68 A. Amato, X. Lucas, A. Bortoluzzi, D. Wright and A. Ciulli, *ACS Chem. Biol.*, 2018, **13**, 915–921.

69 R. Lonsdale and R. A. Ward, *Chem. Soc. Rev.*, 2018, **47**, 3816–3830.

70 A. D. Kwong, R. S. Kauffman, P. Hurter and P. Mueller, *Nat. Biotechnol.*, 2011, **29**, 993–1003.

71 J. Singh, R. C. Petter, T. A. Baillie and A. Whitty, *Nat. Rev. Drug Discov.*, 2011, **10**, 307–317.

72 F. Solca, G. Dahl, A. Zoephel, G. Bader, M. Sanderson, C. Klein, O. Kraemer, F. Himmelsbach, E. Haaksma and G. R. Adolf, *J. Pharmacol. Exp. Ther.*, 2012, **343**, 342–350.

73 Z. Zhao and P. E. Bourne, *Drug Discov. Today*, 2018, **23**, 727–735.

74 D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon and J. A. Wells, *Proc. Natl. Acad. Sci.*, 2000, **97**, 9367–9372.

75 J. P. Daguer, C. Zambaldo, D. Abegg, S. Barluenga, C. Tallant, S. M??ller, A. Adibekian and N. Winssinger, *Angew. Chemie - Int. Ed.*, 2015, **54**, 6057–6061.

76 O. A. Kharenko, R. G. Patel, S. D. Brown, C. Calosing, A. White, D. Lakshminarasimhan, R. K. Suto, B. C. Duffy, D. B. Kitchen, K. G. McLure, H. C. Hansen, E. H. van der Horst and P. R. Young, *J. Med. Chem.*, 2018, **61**, 8202–8211.

77 E. C. Hett, E. L. Piatnitski Chekler, A. Basak, P. D. Bonin, R. A. Denny, A. C. Flick, K. F. Geoghegan, S. Liu, M. T. Pletcher, P. Sahasrabudhe, S. C. Salter, I. A. Stock, A. P. Taylor and L. H. Jones, *Medchemcomm*, 2015, **6**, 1018–1023.

78 G. Zinzalla and D. E. Thurston, *Future Med. Chem.*, 2009, **1**, 65–93.

79 J. Peppard, F. Glickman, Y. He, S. I. Hu, J. Doughty and R. Goldberg, *J. Biomol. Screen.*, 2003, **8**, 149–156.

80 R. M. Eglen, T. Reisine, P. Roby, N. Rouleau, C. Illy, R. Bossé and M. Bielefeld, *Curr. Chem. Genomics*, 2008, **1**, 2–10.

81 J. M. Roberts and J. E. Bradner, *Curr. Protoc. Chem. Biol.*, 2015, **7**, 263–278.

82 E. F. Ullman, H. Kirakossian, A. C. Switchenko, J. Ishkanian, M. Ericson, C. A. Wartchow, M. Pirio, J. Pease, B. R. Irvin, S. Singh, R. Singh, R. Patel, A. Dafforn, D. Davalian, C. Skold, N. Kurn and D. B. Wagner, *Clin. Chem.*, 1996, **42**, 1518–1526.

83 C. Dalvit, G. Fogliatto, A. Stewart, M. Veronesi and B. Stockman, *J. Biomol. NMR*, 2001, **21**, 349–359.

84 C. Ludwig and U. L. Guenther, *Front Biosci (Landmark Ed)*, 2009, **14**, 4565–4574.

85 M. del Carmen Fernández-Alonso, D. Díaz, M. Á. Berbis, F. Marcelo, J. Cañada and J. Jiménez-Barbero, *Curr. Protein Pept. Sci.*, 2012, **13**, 816–830.

86 M. M. Pierce, C. S. Raman and B. T. Nall, *Methods*, 1999, **19**, 213–221.

87 A. Velazquez-Campoy, S. A. Leavitt and E. Freire, in *Protein-Protein Interactions*, Humana Press, New Jersey, vol. 261, pp. 35–54.

88 M. R. Duff, J. Grubbs, E. E. Howell, M. R. Duff, Jr., J. Grubbs and E. E. Howell, *J. Vis. Exp.*, 2011, 5–8.

89 A. Sekirnik, DPhil Thesis, University of Oxford, 2016.

90 A. P. Hill and R. J. Young, *Drug Discov. Today*, 2010, **15**, 648–655.

91 R. J. Young, D. V. S. Green, C. N. Luscombe and A. P. Hill, *Drug Discov. Today*, 2011, **16**, 822–830.

92 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminform.*, 2011, **3**, 1–14.

93 X. Chen and C. H. Reynolds, *J. Chem. Inf. Comput. Sci.*, 2002, **42**, 1407–1414.

94 M. K. P. Jayatunga, S. Thompson, T. C. McKee, M. C. Chan, K. M. Reece, A. P. Hardy, R. Sekirnik, P. T. Seden, K. M. Cook, J. B. McMahon, W. D. Figg, C. J. Schofield and A. D. Hamilton, *Eur. J. Med. Chem.*, 2015, **94**, 506–516.

95 A. Ciulli, G. Williams, A. G. Smith, T. L. Blundell and C. Abell, *J. Med. Chem.*, 2006, **49**, 4992–5000.

96 C. Dalvit, P. Pevarello, M. Tatò, M. Veronesi, A. Vulpetti and M. Sundström, *J. Biomol. NMR*, 2000, **18**, 65–68.

97 L. Fielding, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2007, **51**, 219–242.

98 L. Fielding, S. Rutherford and D. Fletcher, *Magn. Reson. Chem.*, 2005, **43**, 463–470.

99 E. H. Mashalidis, P. Śledź, S. Lang and C. Abell, *Nat. Protoc.*, 2013, **8**, 2309–2324.

100 R. Campos-Olivas, *Curr Top Med Chem*, 2011, **11**, 43–67.

101 C. A. Lepre, *Practical aspects of NMR-based fragment screening*, Elsevier Inc., 1st edn., 2011, vol. 493.

102 M. A. Dombroski, M. A. Letavic, K. F. McClure, J. T. Barberia, T. J. Carty, S. R. Cortina, C. Csiki, A. J. Dipesa, N. C. Elliott, C. A. Gabel, C. K. Jordan, J. M. Labasi, W. H. Martin, K. M. Peese, I. A. Stock, L. Svensson, F. J. Sweeney and C. H. Yu, *Bioorganic Med. Chem. Lett.*, 2004, **14**, 919–923.

103 J. D. White, K. M. Yager and T. Yakura, *J. Am. Chem. Soc.*, 1994, **116**, 1831–1838.

104 S. Schultes, C. De Graaf, E. E. J. Haaksma, I. J. P. De Esch, R. Leurs and O. Krämer, *Drug Discov. Today Technol.*, 2010, **7**, 157–162.

105 H. H. Wasserman, K. Stiller and M. B. Floyd, *Tetrahedron Lett.*, 1968, **9**, 3277–3280.

106 I. Saito, T. Matsuura, M. Nakagawa and T. Hino, *Acc. Chem. Res.*, 1977, **10**, 346–352.

107 E. Haggi, N. Blasich, L. Gutiérrez, G. Vázquez, S. Criado, S. Miskoski, G. Ferrari, M. Paulina Montaña and N. A. García, *J. Photochem. Photobiol. B Biol.*, 2012, **113**, 22–28.

108 J. Al-Nu'Airat, M. Altarawneh, X. Gao, P. R. Westmoreland and B. Z. Dlugogorski, *J. Phys. Chem. A*, 2017, **121**, 3199–3206.

109 P. T. Hart, L. H. J. Kleijn, G. De Bruin, S. F. Oppedijk, J. Kemmink and N. I. Martin, *Org. Biomol. Chem.*, 2014, **12**, 913–918.

110 S. H. Pine and B. L. Sanchez, *J. Org. Chem.*, 1971, **36**, 829–832.

111 WO2004/043903A1, 2004, Propanamine Derivatives as Serotonin and Norepinephrine Reuptake Inhibitors.

112 X. Zhang, K. Chen, Y. D. Wu and O. Wiest, *PLoS One*, 2017, **12**, 1–16.

113 N. Bharatham, P. J. Slavish, B. M. Young and A. A. Shelat, *J. Mol. Graph. Model.*, 2018, **81**, 197–210.

114 WO2008/023736A1, 2008, Substituted Benzimidazolone Derivatives, Medicaments Comprising Them and Their Use.

115 Y. M. Zhang, X. Fan, S. M. Yang, R. H. Scannevin, S. L. Burke, K. J. Rhodes and P. F. Jackson, *Bioorganic Med. Chem. Lett.*, 2008, **18**, 405–408.

116 S. S. Bag, R. Kundu, A. Basak and Z. Slanina, *Org. Lett.*, 2009, **11**, 5722–5725.

117 J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells and K. M. Shokat, *Nature*, 2013, **503**, 548–551.

118 S. G. Kathman, Z. Xu and A. V. Statsyuk, *J. Med. Chem.*, 2014, **57**, 4969–4974.

119 WO2016/096686A1, 2016, Antibacterial Compounds Having Broad Spectrum of Activity.

120 M. W. Crankshaw and G. A. Grant, in *Current Protocols in Protein Science*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 1996, p. 15.1.1-15.1.18.

121 A. L. Skinner and J. S. Laurence, *J. Pharm. Sci.*, 2008, **97**, 4670–4695.

122 C. Guo, L. Dong, S. Kephart and X. Hou, *Tetrahedron Lett.*, 2010, **51**, 2909–2913.

123 N. Kudo, M. Perseghini and G. C. Fu, *Angew. Chemie - Int. Ed.*, 2006, **45**, 1282–1284.

124 N. Otto and T. Opatz, *Beilstein J. Org. Chem.*, 2012, **8**, 1105–1111.

125 T. Nakamura, H. Matsuyama, M. Takahashi and N. Kamigata, *Phosphorus. Sulfur. Silicon Relat. Elem.*, 1992, **66**, 59–65.

126 WO2007/95588 A1, 2007, PI-3 Kinase Inhibitors and Methods of Their Use.

127 G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw and K. I. Goldberg, *Organometallics*, 2010, **29**, 2176–2179.

128 M. Saitoh, J. Kunitomo, E. Kimura, Y. Hayase, H. Kobayashi, N. Uchiyama, T. Kawamoto, T. Tanaka, C. D. Mol, D. R. Dougan, G. S. Textor, G. P. Snell and F. Itoh, *Bioorg. Med. Chem.*, 2009, **17**, 2017–2029.

129 E. Ayral, P. Gloanec, G. Bergé, G. de Nanteuil, P. Mennecier, A. Rupin, T. J. Verbeuren, P. Fulcrand, J. Martinez and J. F. Hernandez, *Bioorganic Med. Chem. Lett.*, 2009, **19**, 1386–1391.

130 Y. Li, C. M. M. Soe, J. J. Wilson, S. L. Tuang, U. P. Apfel and S. J. Lippard, *Eur. J. Inorg. Chem.*, 2013, 2011–2019.

131 M. P. Hay, B. M. Sykes, W. A. Denny and C. J. O'Connor, *J. Chem. Soc. Perkin Trans. 1*, 1999, 2759–2770.

132 M. Kitajima, Y. Murakami, N. Takahashi, Y. Wu, N. Kogure, R. P. Zhang and H. Takayama, *Org. Lett.*, 2014, **16**, 5000–5003.

133 M. A. Blanchette, W. Choy, J. T. Davis, A. P. Essenfeld, S. Masamune, W. R. Roush and T. Sakai, *Tetrahedron Lett.*, 1984, **25**, 2183–2186.

134 S. Liu and R. P. Hanzlik, *J. Med. Chem.*, 1992, **35**, 1067–1075.

135 E. Racine, F. Monnier, J.-P. Vors and M. Taillefer, *Chem. Commun. (Camb).*, 2013, **49**, 7412–7414.

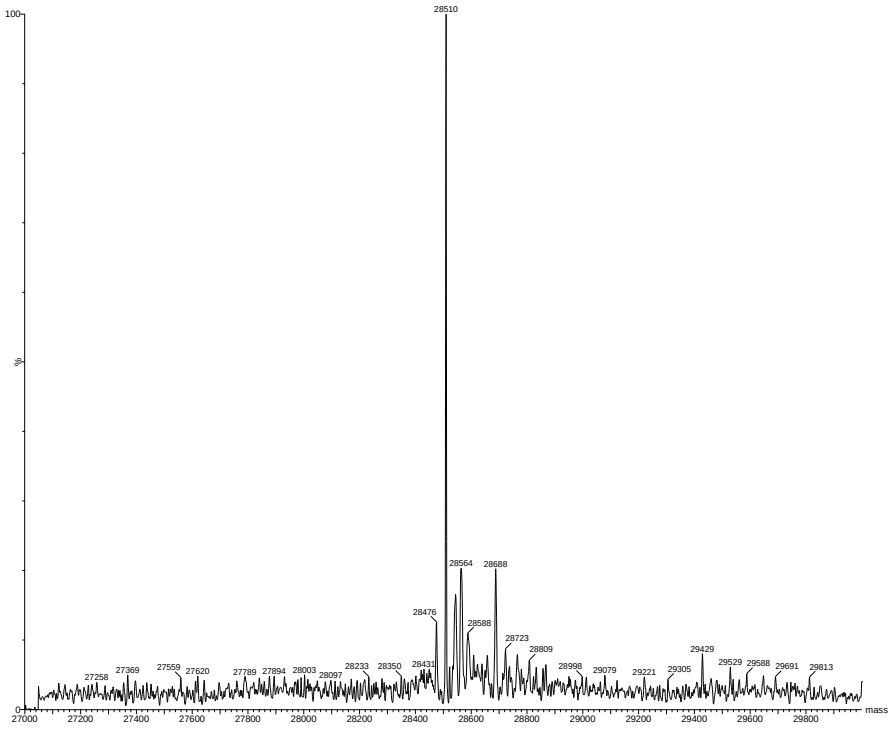
Appendices

A. Protein Characterisation

A1 TRIM33β

Protein	His_6-TRIM33β
Plasmid vector	pNIC28-Bsa4 (Kanamycin)
DNA sequence	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAAC CTGTACTTCCAATCCATGGACTGGTGTGCTGTCTGCCAAAACGGAGGAGAT CTCTTGTGCTGCGAAAAATGTCCAAAGGTCTTTCATCTAACTTGTCATGTT CCAACACTACTTAGCTTTCCAAGTGGGGACTGGATATGCACATTTTGTAGA GATATTGGAAAGCCAGAAGTTGAATATGATTGTGATAATTTGCAACATAG TAAGAAGGGGAAAACTGCGCAGGGGTTAAGCCCCGTGGACCAAAGGAAAT GTGAACGTCTTCTGCTTTACCTCTATTGCCATGAATTAAGTATTGAATTCC AGGAGCCTGTTCCTGCTTCGATACCAAACTACTATAAAATTATAAAGAAA CCAATGGATTTATCCACCGTGAAAAAGAAGCTTCAGAAAAAACATTCCCA ACACTACCAAATCCCGGATGACTTTGTGGCCGATGTCCGTTTGATCTTCAA GAACTGTGAAAGGTTTAATGAGGCTGATTCAGAAGTAGCTCAGGCAGGGA AAGCAGTTGCATTGTACTTTGAAGATAAACTCACAGAGATCTACTCAGAC AGGACCTTCGCACCTTTGCCAGAGTTTGAGCAGGAAGAGGATGATGGTGA GGTAACTGAGGACTCTGATGAAGACTTTATACAGCCCCGCAGAAAACGCCT AAAGTCAGATGAGAGACCAGTACATATAAAGTAAAAGCTTGCGGCCGCAC TCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAA A
Amino acid sequence	MHHHHHHSSGVDLGTENLYFQSMDWCAVCQNGGDLLCCEKCPKVFHLTC HVPTLLSFPSGDWICTFCRDIGKPEVEYDCDNLQHSKKGKTAQGLSPVDQR KCERLLLYLYCHELSIEFQEPVPASIPNYYKIIKKPMDLSTVKKKLQKKHSQH YQIPDDFVADVRLIFKNCERFNEADSEVAQAGKAVALYFEDKLTEIYSDRTF APLPEFEQEEDDGEVTEDSDEDFIQPRRKRLKSDERPVHIK*
MW (Da)	28509.2
ε (M-1 cm-1)	25160

Mass spectrum



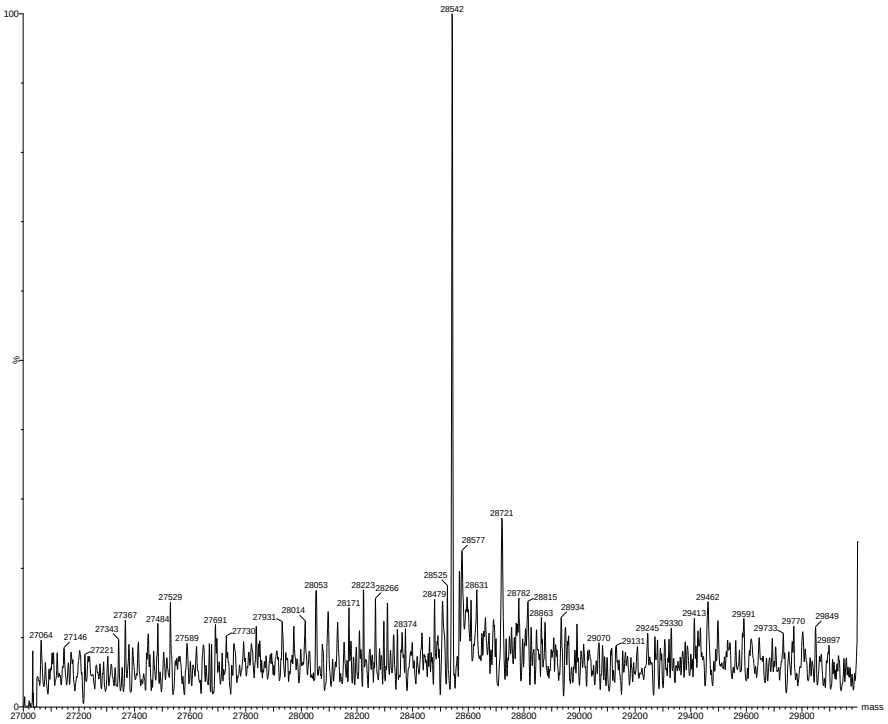
A2 TRIM33β-N1039F

Protein	His_6-TRIM33β-N1039F
Plasmid vector	pNIC28-Bsa4 (Kanamycin)
DNA sequence	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAAC CTGTACTTCCAATCCATGGACTGGTGTGCTGTCTGCCAAAACGGAGGAGAT CTCTTGTGCTGCGAAAAATGTCCAAAGGTCTTTCATCTAACTTGTCATGTT CCAACACTACTTAGCTTTCCAAGTGGGGACTGGATATGCACATTTTGTAGA GATATTGGAAAGCCAGAAGTTGAATATGATTGTGATAATTTGCAACATAG TAAGAAGGGGAAAACTGCGCAGGGGTTAAGCCCCGTGGACCAAAGGAAAT GTGAACGTCTTCTGCTTTACCTCTATTGCCATGAATTAAGTATTGAATTCC AGGAGCCTGTTCCTGCTTCGATACCAAACTACTATAAAATTATAAAGAAA CCAATGGATTTATCCACCGTGAAAAAGAAGCTTCAGAAAAAACATTCCCA ACACTACCAAATCCCGGATGACTTTGTGGCCGATGTCCGTTTGATCTTCAA GAACTGTGAAAGGTTTTTTGAGGCTGATTCAGAAGTAGCTCAGGCAGGGA AAGCAGTTGCATTGTACTTTGAAGATAAACTCACAGAGATCTACTCAGAC AGGACCTTCGCACCTTTGCCAGAGTTTGAGCAGGAAGAGGATGATGGTGA GGTAACTGAGGACTCTGATGAAGACTTTATACAGCCCCGCAGAAAACGCCT AAAGTCAGATGAGAGACCAGTACATATAAAGTAA
Amino acid sequence	MHHHHHHSSGVDLGTENLYFQSMDWCAVCQNGGDLLCCEKCPKVFHLTC HVPTLLSFPSGDWICTFCRDIGKPEVEYDCDNLQHSKKGKTAQGLSPVDQR KCERLLLYLYCHELSIEFQEPVPASIPNYYKIIKKPMDLSTVKKKLQKKHSQH

YQIPDDFVADVRLIFKNCERFFEADSEVAQAGKAVALYFEDKLTEIYSDRTF
APLPEFEQEEDDGEVTEDSDEDFIQPRRKRLKSDERPVHIK*

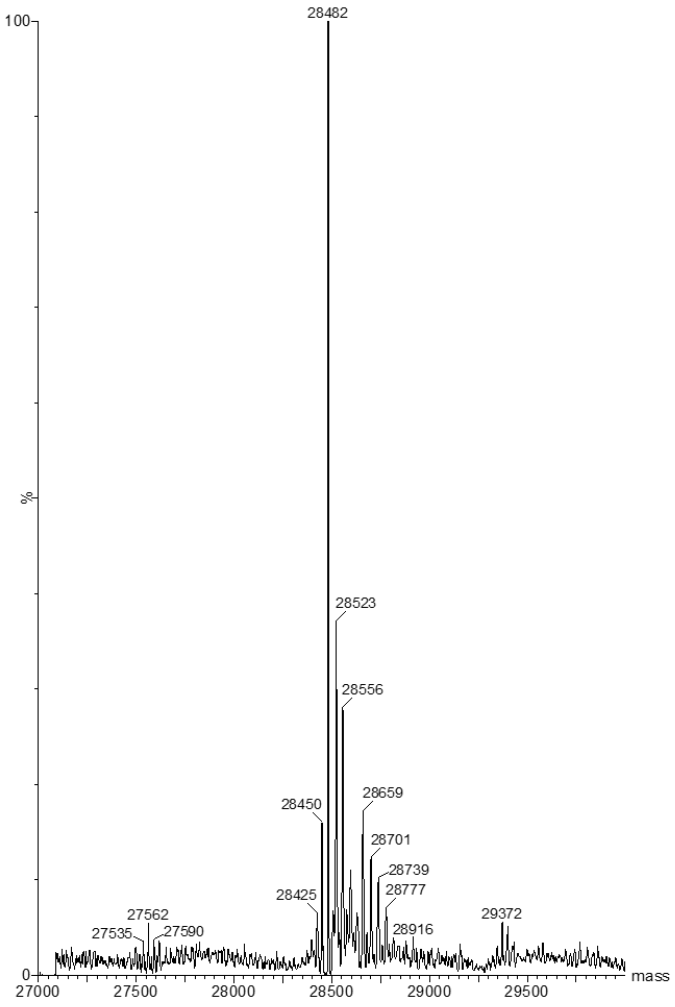
MW (Da)	28542.3
ε (M-1 cm-1)	25160

Mass spectrum



A3 TRIM33β-E984C

Protein	His_6-TRIM33β-E984C
Plasmid vector	pNIC28-Bsa4 (Kanamycin)
DNA sequence	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAAC CTGTACTTCCAATCCATGGACTGGTGTGCTGTCTGCCAAAACGGAGGAGAT CTCTTGTGCTGCGAAAAATGTCCAAAGGTCTTTCATCTAACTTGTCATGTT CCAACACTACTTAGCTTTCCAAGTGGGGACTGGATATGCACATTTTGTAGA GATATTGGAAAGCCAGAAGTTGAATATGATTGTGATAATTTGCAACATAG TAAGAAGGGGAAAACTGCGCAGGGGTTAAGCCCCGTGGACCAAAGGAAAT GTGAACGTCTTCTGCTTTACCTCTATTGCCATGAATTAAGTATTGAATTCC AGTGCCCTGTTCCTGCTTCGATACCAAACTACTATAAAATTATAAAGAAAC CAATGGATTTATCCACCGTGAAAAAGAAGCTTCAGAAAAAACATTCCCAA CACTACCAAATCCCGGATGACTTTGTGGCCGATGTCCGTTTGATCTTCAAG AACTGTGAAAGGTTTTTTGAGGCTGATTCAGAAGTAGCTCAGGCAGGGAA AGCAGTTGCATTGTACTTTGAAGATAAACTCACAGAGATCTACTCAGACA GGACCTTCGCACCTTTGCCAGAGTTTGAGCAGGAAGAGGATGATGGTGAG GTAACTGAGGACTCTGATGAAGACTTTATACAGCCCCGCAGAAAACGCCTA AAGTCAGATGAGAGACCAGTACATATAAAGTAA

Amino acid sequence	MHHHHHHSSGVDLGTENLYFQSMDWCAVCQNGGDLLCCEKCPKVFHLTC HVPTLLSFPSGDWICTFCRDIGKPEVEYDCDNLQHSKKGKTAQGLSPVDQR KCERLLLYLYCHELSIEFQCPVPASIPNYYKIIKKPMDLSTVKKKLQKKHSQH YQIPDDFVADVRLIFKNCERFNEADSEVAQAGKAVALYFEDKLTEIYSDRTF APLPEFEQEEDDGEVTEDSDEDFIQPRRKRLKSDERPVHIK*
MW (Da)	28483.3
ε (M-1 cm-1)	25160
Mass spectrum	

B. AlphaScreen™ Data

	[Peptide] nM		0	0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250	500	1000	0	0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250	500	1000		
		FAC	0	0.24414	0.48828	0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250	0	0.24414	0.48828	0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250		
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		
[Protein]		A																										
		B																										
		C																										
		D																										
T24 100 nM	$H3K9Me_3$	E	7416	67727	126905	223350	334276	447372	537079	525843	541877	517419	516891	458421	7426	7796	10873	18536	30247	88609	253054	299022	314036	345858	374321	380560	T33a 500 nM	$H3K9Me_3$
	H3K18Ac	F	7496	68685	129548	226459	319153	415968	485391	502118	505103	473180	450875	384188	4221	6693	10756	22912	46856	98477	188406	304569	324861	326505	309958	295537		H3K18Ac
	$H3K9Me_3K18Ac$	G	3668	134823	212284	306739	392014	466490	502447	502047	498018	464007	426026	329424	3396	15980	43366	84200	162888	257195	385580	401198	388209	390441	364763	364022		$H3K9Me_3K18Ac$
T24 50 nM	$H3K9Me_3$	H	3046	4761	10081	24459	72268	138988	190591	206078	207625	205640	185785	136546	2311	3942	4863	6088	7078	10243	28566	68833	76004	78709	80948	61429	T33a 250 nM	$H3K9Me_3$
	H3K18Ac	I	2772	22913	48940	111573	193888	295402	360744	410273	406329	370715	330520	250251	2635	2723	4321	7571	13094	28457	47936	121257	152831	156867	158193	177059		H3K18Ac
	$H3K9Me_3K18Ac$	J	2694	60555	126455	204317	303186	387184	440389	437237	430530	359066	305974	204761	2420	8426	21457	50799	109634	194098	298233	333231	321679	330300	313316	259167		$H3K9Me_3K18Ac$
T24 25 nM	$H3K9Me_3$	K	2311	3196	3961	5470	15380	27799	49728	65040	72626	63355	57394	31459	2154	3123	3777	5153	5962	7716	10211	19544	29517	13963	9721	21567	T33a 125 nM	$H3K9Me_3$
	H3K18Ac	L	2568	10040	22972	62821	127140	213366	293985	333137	335944	295065	257903	178302	2120	2291	3038	4740	5880	9018	14319	34021	55139	57965	74784	91285		H3K18Ac
	$H3K9Me_3K18Ac$	M	2574	40131	84746	160840	246732	316681	381120	358195	341702	268515	213059	125562	3575	3973	6947	22124	62855	110044	192653	248122	241909	241100	226385	213510		$H3K9Me_3K18Ac$
		N																										
		O																										
		P																										

	[Peptide] nM		0	0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250	500	1000
		FAC	0	0.24414	0.48828	0.97656	1.95313	3.90625	7.8125	15.625	31.25	62.5	125	250
			1	2	3	4	5	6	7	8	9	10	11	12
[Protein]		A												
		B												
		C												
		D												
T33β 500 nM	$H3K9Me_3$	E	4960	19047	39768	73222	126295	250092	419845	428348	427943	424751	413595	379243
	H3K18Ac	F	4188	22649	77739	153178	257137	336375	394863	441650	432854	434985	427296	378792
	$H3K9Me_3K18Ac$	G	3385	72569	148024	231329	344404	386596	420751	424534	419552	400575	405984	367331
T33β 250 nM	$H3K9Me_3$	H	3353	6448	10452	16486	26930	69752	179360	202204	207865	188533	188606	161697
	H3K18Ac	I	2957	11561	38033	84851	169917	262691	342496	414055	415924	395052	381397	336401
	$H3K9Me_3K18Ac$	J	3116	40927	89360	168369	277360	362602	416685	440376	429622	417918	397352	350384
T33β 125 nM	$H3K9Me_3$	K	2586	3816	5984	7186	8590	12029	34597	68434	89351	136273	136669	149209
	H3K18Ac	L	2808	9107	21340	48097	101470	190007	295100	373188	369307	352307	339672	291151
	$H3K9Me_3K18Ac$	M	2542	20361	68245	125858	198027	318634	389813	439517	442935	433604	430212	372796
		N												
		O												
		P												

Figure B1. AlphaScreen™ signal response data. T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β.

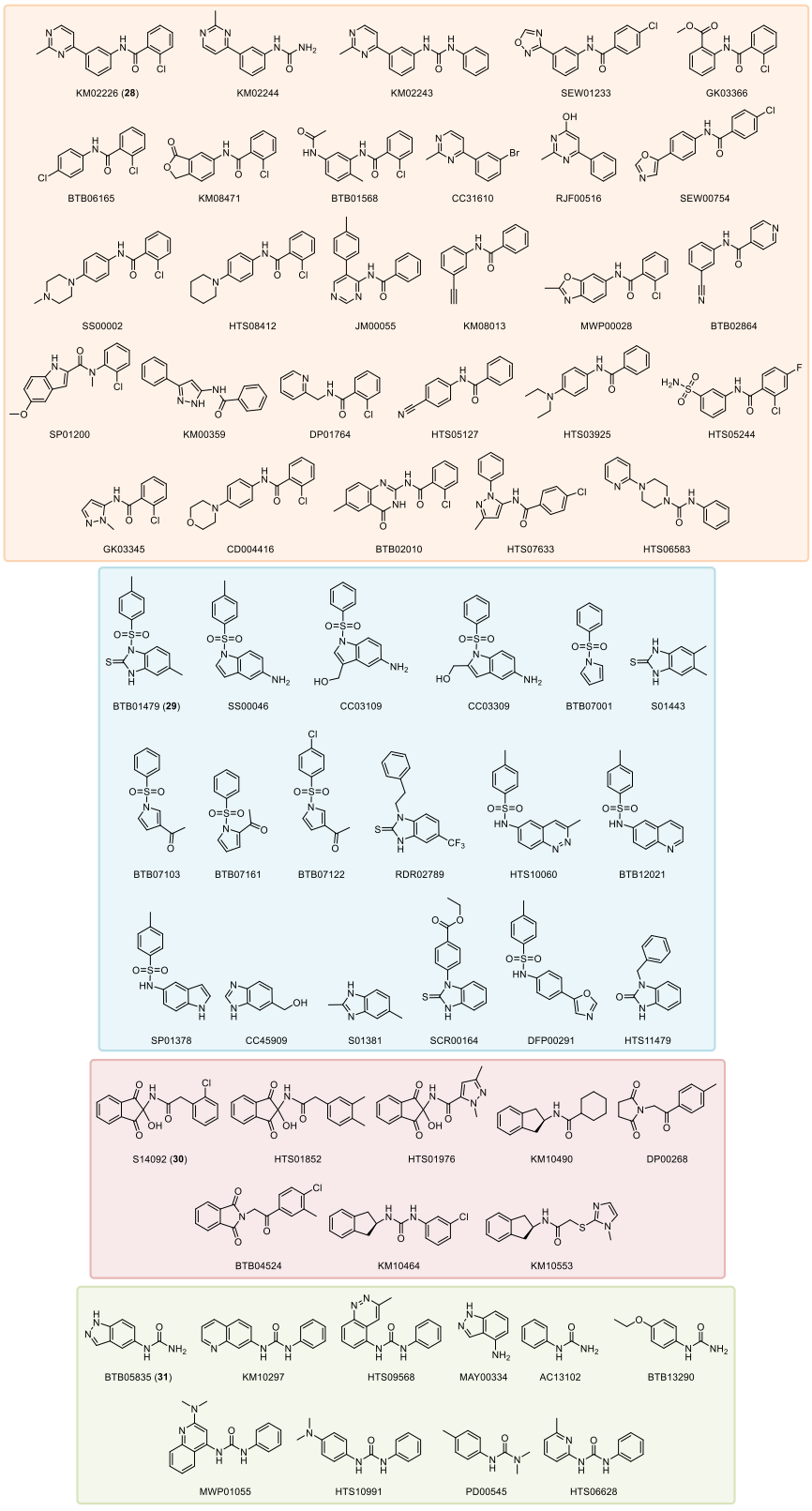


Figure B2. Compounds screened as part of the focused selection from the Maybridge fragment library, grouped by their similarity to hit compounds 28–31.

Table B1. Results from AlphaScreen™ (Chapter 3). Colour scale from low emission (red) to high emission (green); T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β, H3(PHD) – H3K9Me$_3$, H3(BRD) – H3K18Ac, H3(Dual) – H3K9Me$_3$K18Ac. Compounds taken forward for further validation are indicated by bold boxes.

Concentration (μM) / Compound	Emission (as a percentage of DMSO control)																	
	T24-H3(PHD)		T24-H3(BRD)		T24-H3(Dual)		T33α-H3(PHD)		T33α-H3(BRD)		T33α-H3(Dual)		T33β-H3(PHD)		T33β-H3(BRD)		T33β-H3(Dual)	
	30	150	30	150	30	150	30	150	30	150	30	150	30	150	30	150	30	150
KM02226 (**28**)	88	97	86	86	82	85	111	114	96	78	110	100	106	100	94	92	97	87
KM02244	70	78	74	78	71	72	84	93	70	60	84	75	93	90	92	86	101	97
KM02243	78	64	77	74	78	82	101	87	77	71	87	90	96	97	96	93	105	105
SEW01233	83	77	74	77	96	93	89	87	89	87	68	70	89	85	86	97	104	101
GK03366	101	101	88	104	97	115	107	124	95	123	74	97	107	124	105	133	122	142
BTB06165	83	71	90	97	106	103	99	101	100	104	94	104	98	105	102	106	107	114
KM08471	89	89	79	78	80	78	92	97	84	90	79	78	96	87	90	93	103	99
BTB01568	68	80	78	80	80	92	72	71	81	79	79	84	93	92	81	89	95	104
CC31610	93	104	77	91	86	102	78	111	75	88	63	69	83	106	90	111	93	112
SEW00754	111	74	115	73	131	95	105	69	106	80	89	62	111	89	111	87	133	100
SS00002	76	71	77	67	81	62	77	60	79	74	67	64	86	78	88	80	93	80

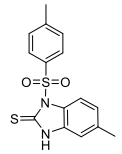
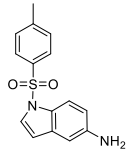
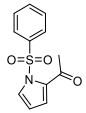
HTS08412	69	63	65	59	81	77	76	69	72	67	60	54	76	73	78	72	86	77
RJF00516	86	82	81	82	86	82	74	84	92	79	60	66	89	90	91	91	96	90
JM00055	81	78	65	61	90	83	78	70	80	93	66	67	83	84	93	95	97	93
KM08013	93	76	91	91	92	86	98	98	89	97	85	85	102	105	99	103	113	116
MWP00028	94	86	87	85	92	94	85	82	85	82	70	60	93	93	100	96	96	102
BTB02864	79	80	74	80	85	97	86	88	73	84	82	79	92	90	91	95	103	103
SP01200 (**32**)	66	61	63	55	77	66	72	59	69	50	54	44	77	61	84	71	83	69
KM00359	96	80	111	83	132	95	102	87	104	88	109	96	122	97	125	97	84	106
DP01764	79	68	79	75	80	78	77	74	81	71	67	67	88	88	89	88	92	86
HTS05127	77	79	77	81	81	73	85	90	62	57	62	62	84	83	90	87	98	82
HTS03925	70	70	79	69	92	84	79	76	85	80	76	80	82	82	87	85	88	77
HTS05244	124	91	127	82	138	105	131	104	107	82	102	79	121	101	119	96	130	110
GK03345	76	82	73	70	89	84	87	90	81	83	71	75	93	97	83	93	102	102
CD04416 (**33**)	84	64	75	65	79	74	87	74	72	53	71	48	87	81	90	87	95	80
BTB02010	96	95	86	87	91	93	98	93	82	91	73	80	94	95	92	93	93	104
HTS07633	82	95	70	71	89	89	82	75	68	89	67	66	84	88	95	80	93	96

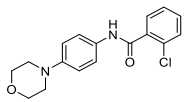
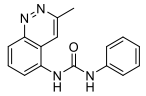
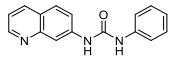
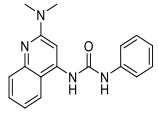
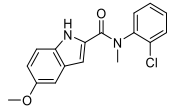
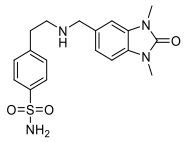
HTS06583	83	84	82	81	88	89	96	96	78	78	73	70	91	89	92	91	105	103
BTB01479 (**29**)	129	5	123	4	135	3	136	21	118	10	160	7	102	69	134	23	143	20
SS00046 (**34**)	54	29	52	32	61	35	58	37	58	22	57	28	71	40	70	39	73	37
CC03109	92	73	88	76	92	85	92	84	75	76	82	77	92	91	92	86	101	94
CC03309	79	88	89	87	90	93	96	71	89	90	86	96	92	97	92	95	102	104
BTB07001	85	89	89	96	92	99	94	94	92	95	90	97	89	95	97	101	100	109
S01443	69	54	61	55	82	70	77	72	76	54	68	63	83	77	95	80	89	79
BTB07103	81	88	85	87	89	91	86	90	92	88	83	98	88	95	94	100	99	95
BTB07161 (**35**)	137	86	122	76	118	87	115	27	103	15	104	61	125	83	121	89	109	99
BTB07122	132	102	98	66	122	92	98	96	108	79	91	79	125	107	123	97	108	92
RDR02789	90	82	67	72	94	92	78	68	82	81	60	57	88	84	92	94	97	96
HTS10060	77	75	76	80	93	88	92	74	92	76	74	62	92	84	91	85	96	93
BTB12021	89	85	82	77	82	80	74	70	81	70	82	74	98	92	90	88	100	93
SP01378	79	68	73	67	76	70	67	54	83	67	62	57	79	77	83	83	92	82
CC45909	79	76	87	81	86	86	88	40	95	88	90	94	88	89	91	94	77	88
S01381	87	76	67	66	89	93	84	83	85	74	63	64	90	87	93	92	97	95

SCR00164	72	78	63	63	85	83	69	51	67	52	61	60	86	83	81	88	80	89
DFP00291	73	75	120	80	136	96	111	87	107	82	101	78	117	91	112	90	131	102
HTS11479	76	70	70	69	77	81	83	82	55	52	55	54	85	84	93	81	94	92
S14092 (**30**)	105	92	95	93	90	83	92	71	87	60	108	84	97	93	96	95	88	86
HTS01852 (**39**)	85	70	80	74	87	71	59	51	70	31	73	54	89	76	92	85	98	89
HTS01976 (**40**)	79	60	79	75	90	83	54	49	72	32	77	57	83	76	92	82	96	88
KM10490	74	75	73	75	86	92	81	66	63	65	59	58	86	89	83	94	85	82
DP00268	78	68	77	77	90	95	86	84	90	66	76	67	94	96	89	93	39	104
BTB04524	79	71	87	85	95	101	92	88	87	91	85	88	92	90	92	94	101	101
KM10464	81	83	67	77	88	93	86	91	86	87	80	85	93	96	95	98	99	101
KM10553	107	88	109	77	137	96	101	84	89	74	94	71	120	93	128	98	131	101
BTB05835 (**31**)	95	81	79	74	92	78	105	94	84	58	103	86	107	93	95	79	86	81
KM10297 (**36**)	83	76	69	65	79	72	108	113	64	25	71	60	100	84	92	84	95	86
HTS09568 (**37**)	80	70	78	69	92	81	84	78	82	46	69	53	91	83	89	78	97	85
MAY00334	84	81	66	70	88	82	74	79	73	69	65	61	92	88	63	91	103	95
AC13102	114	87	113	83	123	84	98	68	104	87	117	84	121	101	110	91	132	105

BTB13290	91	89	76	83	86	90	98	92	83	71	81	83	96	95	91	93	102	103
MWP01055 (**38**)	100	48	78	37	86	41	188	113	101	50	64	29	73	40	70	35	78	36
HTS10991	65	58	71	63	81	74	72	69	71	66	57	59	74	72	74	74	71	78
PD00545	119	86	116	82	137	98	103	83	104	89	99	73	118	95	120	99	126	100
HTS06628	79	80	74	71	90	91	100	92	87	87	76	73	82	88	93	93	105	105

Table B2. Results from AlphaScreen™ TruHits experiment. Differences of greater than two standard deviations are indicated in pink.

Compound \ Concentration (μM)	Number of standard deviations from DMSO average, normalised against standard deviation of DMSO control									
	Pre-incubate beads					Pre-incubate donor beads with compound				
	9.38	18.8	37.5	75.0	150	9.38	18.8	37.5	75.0	150
 **29**	2.99	3.50	6.50	10.66	16.43	1.07	1.70	2.97	4.07	4.67
 **34**	0.74	2.76	5.73	9.18	13.09	0.85	0.09	0.47	1.75	3.02
 **35**	1.30	0.10	0.55	0.35	0.77	1.32	1.28	1.48	1.42	0.45

 **33**	1.15	0.12	0.44	0.82	2.04	1.10	1.62	1.31	1.32	0.61
 **37**	0.75	0.11	0.31	0.81	0.86	0.48	0.30	1.10	0.15	1.62
 **36**	0.31	0.14	0.55	0.96	1.54	0.79	0.29	0.72	0.78	0.26
 **38**	2.44	4.39	7.07	10.59	14.61	0.20	0.28	0.56	1.78	3.30
 **32**	1.10	2.30	3.42	4.74	5.58	0.95	0.59	0.29	0.05	0.29
 **20**	0.19	0.68	0.58	0.67	0.33	0.43	0.27	0.41	1.22	0.75

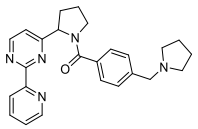
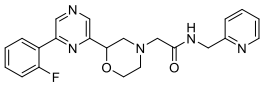
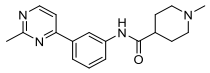
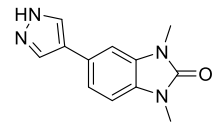
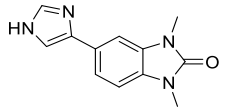
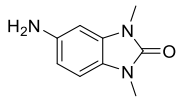
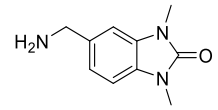
 **23**	0.73	1.03	1.33	1.55	1.65	0.41	0.14	0.26	0.58	0.11
 **25**	0.56	1.28	1.30	1.31	0.86	0.94	0.31	0.66	0.37	0.13
 **172**	0.70	0.11	0.22	0.26	0.75	0.11	0.21	0.03	0.72	1.41
Biotin control	-	-	-	-	-	5.18	5.36	5.55	5.61	5.74

Table B3. AlphaScreen™ IC_{50} values (corresponding to Table 4.2). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **54**	>250	92.1 ± 1.23 (0.33)	>250	>250	5.34 ± 1.10 (0.43)	6.501 ± 1.07 (0.43)	>250	23.3 ± 1.27 (0.38)	>250
 **55**	>250	52.4 ± 1.21 (0.35)	>250	>250	3.35 ± 1.11 (0.45)	10.4 ± 1.12 (0.41)	>250	26.5 ± 1.10 (0.38)	177 ± 1.34 (0.31)
 **46**	>250	22.5 ± 1.45 (0.50)	>250	>250	41.8 ± 1.13 (0.47)	>250	>250	168 ± 1.96 (0.41)	>250
 **53**	>250	>250	>250	9.63 ± 1.18 (0.50)	7.83 ± 1.10 (0.51)	16.3 ± 1.12 (0.48)	19.1 ± 1.18 (0.47)	30.6 ± 1.09 (0.45)	86.4 ± 1.32 (0.41)

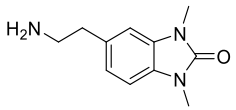
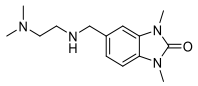
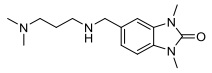
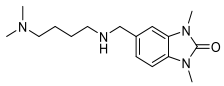
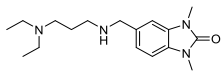
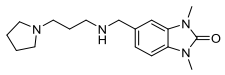
 **52**	>250	>250	>250	>250	65.1 ± 12.4 (0.39)	>250	>250	>250	>250

Table B4. AlphaScreen™ IC_{50} values (corresponding to Table 4.3). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **57**	8.72 ± 1.19 (0.37)	71.9 ± 1.30 (0.31)	45.1 ± 1.08 (0.32)	17.1 ± 1.17 (0.35)	5.89 ± 1.09 (0.39)	9.68 ± 1.11 (0.37)	6.79 ± 1.11 (0.38)	15.2 ± 1.07 (0.35)	26.10 ± 1.05 (0.34)
 **62***	>250	>250	>250	4.47 ± 1.78 (0.37)	3.61 ± 1.15 (0.38)	6.80 ± 2.98 (0.36)	7.48 ± 3.18 (0.36)	13.7 ± 9.34 (0.34)	32.7 ± 14.8 (0.31)
 **58**	>250	>250	>250	4.212 ± 1.11 (0.36)	7.96 ± 1.11 (0.34)	2.39 ± 1.12 (0.37)	2.67 ± 1.22 (0.37)	2.73 ± 1.14 (0.37)	3.86 ± 1.07 (0.36)
 **59**	>250	>250	>250	2.38 ± 1.12 (0.36)	4.743 ± 1.14 (0.34)	1.65 ± 1.08 (0.37)	1.58 ± 1.28 (0.37)	2.024 ± 1.06 (0.36)	6.24 ± 1.04 (0.33)
 **60**	4.54 ± 1.37 (0.34)	24.23 ± 1.36 (0.29)	156.0 ± 1.61 (0.31)	5.06 ± 1.54 (0.34)	4.11 ± 1.31 (0.34)	6.65 ± 1.48 (0.33)	5.32 ± 1.47 (0.34)	9.30 ± 1.47 (0.32)	12.9 ± 1.48 (0.31)

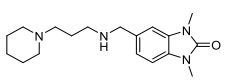
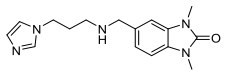
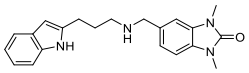
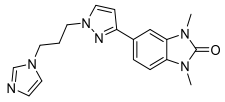
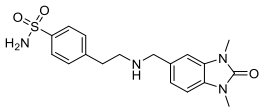
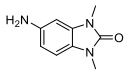
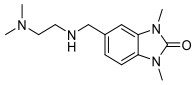
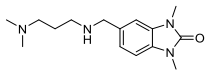
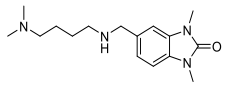
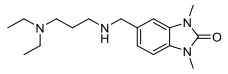
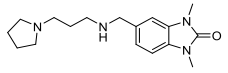
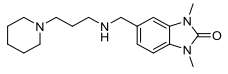
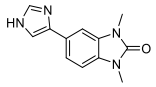
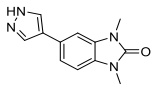
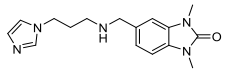
 **61**	22.7 ± 1.16 (0.28)	130 ± 10.3 (0.24)	102 ±1.12 (0.24)	4.60 ± 1.14 (0.32)	2.71 ± 1.08 (0.34)	4.89 ± 1.12 (0.32)	3.81 ± 1.08 (0.33)	9.24 ±1.06 (0.31)	24.6 ± 1.12 (0.28)
 **63**	>250	>250	>250	75.4 ± 1.66 (0.26)	7.97 ± 1.06 (0.32)	3.37 ± 1.20 (0.35)	54.3 ± 1.28 (0.27)	11.9 ± 1.12 (0.31)	72.8 ± 1.19 (0.26)
 **64**	>250	>250	>250	>250	10.5 ± 1.09 (0.28)	33.3 ± 1.32 (0.25)	>250	65.5 ± 1.18 (0.23)	>250
 **65**	>250	34.5 ± 1.35 (0.25)	>250	9.49 ± 1.37 (0.28)	2.46 ± 1.07 (0.31)	5.56 ± 1.11 (0.29)	21.1 ± 1.40 (0.26)	15.3 ± 1.10 (0.27)	43.6 ± 1.22 (0.24)

Table B5. Results from AlphaScreen™ TruHits experiment. Differences of greater than two standard deviations are indicated in pink.

Compound \ Concentration (μM)	Number of standard deviations from DMSO average, normalised against standard deviation of DMSO control									
	Pre-incubate beads					Pre-incubate donor beads with compound				
	31.25	62.5	125	250	500	31.25	62.5	125	250	500
 **20**	1.801	0.782	0.409	0.607	2.020	0.536	0.371	0.198	0.433	0.706
 **46**	0.620	1.323	3.706	5.706	11.569	0.749	0.095	0.246	0.221	1.668
 **57**	0.650	1.215	1.185	1.680	0.972	0.334	0.067	0.186	0.196	0.692
 **62**	0.352	0.114	1.008	0.945	2.441	0.010	0.160	0.065	0.256	0.806

 **58**	0.274	1.067	0.460	1.859	3.612	0.309	0.507	1.634	0.425	0.337
 **59**	0.449	1.382	0.527	0.569	2.096	0.364	0.991	1.039	0.077	0.669
 **60**	0.749	0.476	1.119	0.605	2.667	0.833	1.340	0.590	0.118	0.435
 **61**	0.876	0.002	0.970	1.040	2.180	0.417	1.334	0.413	0.137	0.533
 **55**	1.473	0.117	2.910	5.707	11.423	1.250	0.211	0.884	1.606	2.688
 **54**	1.876	1.219	0.376	0.170	1.718	0.457	0.698	1.202	1.571	2.018
 **63**	1.227	0.109	0.827	1.586	4.179	1.027	0.454	0.129	0.222	0.600

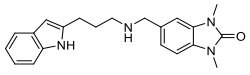
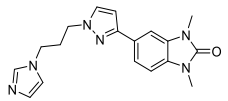
 **64**	0.602	2.255	5.534	8.835	14.964	0.294	0.036	0.431	1.056	2.035
 **65**	0.715	2.345	3.933	5.707	9.882	0.245	0.306	0.574	1.295	2.179

Table B6. AlphaScreen™ IC_{50} values (corresponding to Table 4.5). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

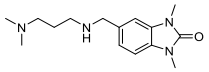
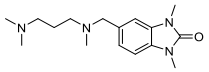
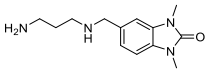
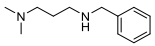
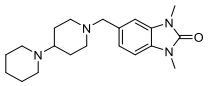
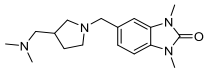
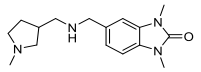
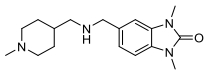
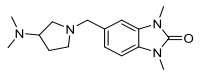
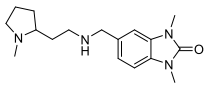
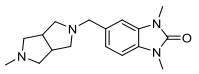
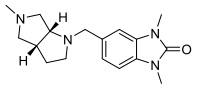
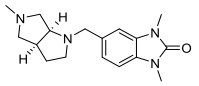
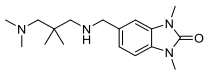
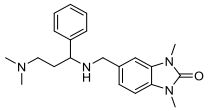
Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **62***	>250	>250	>250	4.47 ± 1.78 (0.37)	3.61 ± 1.15 (0.38)	6.80 ± 2.98 (0.36)	7.48 ± 3.18 (0.36)	13.7 ± 9.34 (0.34)	32.7 ± 14.8 (0.31)
 **67**	>250	>250	>250	9.42 ± 1.41 (0.34)	5.69 ± 0.33 (0.35)	12.8 ± 1.32 (0.33)	15.5 ± 1.79 (0.32)	23.5 ± 2.80 (0.31)	>250
 **68**	>250	>250	>250	>250	3.73 ± 0.38 (0.42)	>250	7.71 ± 0.79 (0.40)	13.8 ± 1.00 (0.38)	37.5 ± 8.01 (0.34)
 **66**	>250	>250	>250	>250	>250	>250	>250	>250	>250

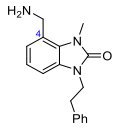
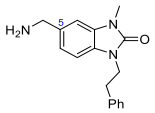
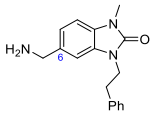
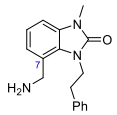
Table B7. AlphaScreen™ IC_{50} values (corresponding to Table 4.6). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **73**	>250	>250	>250	20.9 ± 1.12 (0.26)	7.96 ± 1.11 (0.29)	31.1 ± 1.28 (0.25)	30.9 ± 1.17 (0.25)	40.7 ± 1.09 (0.25)	>250
 **76**	>250	>250	>250	11.9 ± 1.78 (0.31)	5.81 ± 0.21 (0.33)	11.6 ± 0.97 (0.31)	18.3 ± 4.38 (0.30)	15.2 ± 0.79 (0.31)	72.8 ± 30.1 (0.26)
 **77**	>250	>250	>250	2.84 ± 0.47 (0.37)	1.60 ± 0.14 (0.39)	4.58 ± 0.23 (0.36)	7.50 ± 0.54 (0.34)	8.53 ± 0.56 (0.34)	29.7 ± 3.37 (0.30)
 **78**	>250	>250	>250	5.82 ± 1.17 (0.33)	5.49 ± 1.12 (0.33)	17.1 ± 1.07 (0.30)	12.2 ± 1.15 (0.31)	12.0 ± 1.12 (0.31)	51.4 ± 1.27 (0.27)
 **79**	>250	>250	>250	>250	6.61 ± 0.54 (0.35)	15.2 ± 3.61 (0.32)	12.6 ± 1.04 (0.33)	12.8 ± 2.09 (0.33)	38.5 ± 17.3 (0.29)

 **80**	>250	>250	>250	4.97 ± 0.74 (0.34)	1.52 ± 0.10 (0.37)	2.29 ± 0.26 (0.36)	2.83 ± 0.20 (0.35)	1.62 ± 0.14 (0.37)	7.50 ± 0.54 (0.33)
 **81**	>250	>250	>250	43.0 ± 7.00 (0.28)	25.6 ± 6.82 (0.29)	70.3 ± 12.2 (0.26)	>250	100 ± 20.7 (0.25)	>250
 **82**	45.7 ± 5.04 (0.28)	44.3 ± 3.43 (0.28)	55.6 ± 7.54 (0.27)	4.51 ± 0.62 (0.34)	3.97 ± 0.86 (0.34)	16.2 ± 1.41 (0.30)	13.8 ± 3.74 (0.31)	12.3 ± 3.13 (0.31)	>250
 **83**	>250	>250	>250	7.91 ± 0.81 (0.32)	5.79 ± 0.84 (0.33)	10.3 ± 2.21 (0.32)	12.8 ± 3.04 (0.31)	9.57 ± 1.74 (0.32)	>250
 **84**	>250	>250	>250	>250	19.8 ± 3.56 (0.30)	43.3 ± 11.7 (0.28)	51.6 ± 5.24 (0.27)	39.2 ± 8.41 (0.28)	>250
 **85**	>250	>250	>250	6.56 ± 1.63 (0.28)	1.56 ± 0.10 (0.31)	3.61 ± 0.29 (0.29)	5.36 ± 0.50 (0.28)	4.83 ± 0.51 (0.29)	12.1 ± 1.39 (0.26)

86	>250	>250	>250	4.44 ± 0.65 (0.28)	2.45 ± 0.24 (0.30)	5.89 ± 0.65 (0.28)	9.21 ± 2.60 (0.27)	7.75 ± 0.60 (0.28)	>250	
87	>250	>250	>250	1.97 ± 0.10 (0.31)	1.33 ± 0.06 (0.32)	2.70 ± 0.24 (0.30)	4.32 ± 1.25 (0.29)	4.52 ± 0.60 (0.29)	29.0 ± 2.69 (0.24)	
88	>250	>250	>250	7.13 ± 0.68 (0.28)	4.31 ± 0.24 (0.29)	12.3 ± 1.85 (0.26)	21.6 ± 6.45 (0.25)	16.9 ± 2.25 (0.26)	>250	
89	>250	>250	>250	>250	5.56 ± 0.56 (0.28)	>250	15.8 ± 1.46 (0.26)	12.2 ± 1.55 (0.26)	105 ± 25.8 (0.21)	

Table B8. AlphaScreen™ IC_{50} values (corresponding to Table 4.7). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **109**	>250	>250	>250	>250	>250	>250	>250	>250	>250
 **108**	>250	>250	>250	>250	12.1 ± 1.12 (0.33)	19.7 ± 1.86 (0.31)	>250	>250	>250
 **110**	>250	>250	>250	>250	23.3 ± 2.34 (0.31)	39.2 ± 5.97 (0.29)	>250	>250	>250
 **99**	>250	>250	>250	>250	>250	>250	>250	>250	>250

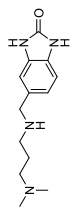
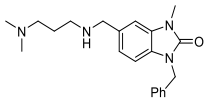
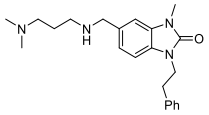
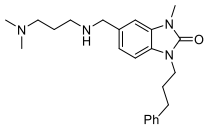
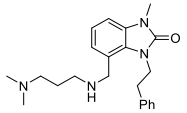
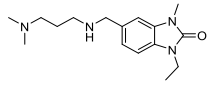
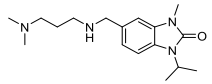
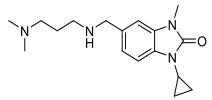
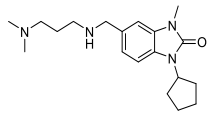
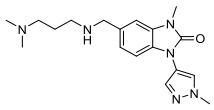
111		>250	>250	>250	>250	>250	>250	>250	>250

Table B9. AlphaScreen™ IC_{50} values (corresponding to Table 4.8). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **114A**	>250	>250	>250	2.39 ± 1.20 (0.30)	3.64 ± 1.11 (0.29)	5.77 ± 1.10 (0.28)	35.7 ± 1.12 (0.24)	86.0 ± 1.15 (0.22)	>250
 **115A**	>250	>250	>250	2.81 ± 1.19 (0.29)	4.34 ± 1.08 (0.28)	6.07 ± 1.10 (0.27)	19.2 ± 1.30 (0.24)	59.7 ± 1.12 (0.22)	150 ± 1.26 (0.20)
 **116**	>250	>250	>250	0.710 ± 2.56 (0.31)	2.47 ± 1.08 (0.28)	3.89 ± 1.05 (0.27)	27.8 ± 1.14 (0.23)	100 ± 1.10 (0.20)	>250
 **117**	>250	>250	>250	>250	>250	>250	>250	>250	>250

118	>250	>250	>250	>250	3.35 ± 1.11 (0.29)	9.76 ± 1.09 (0.27)	>250	>250
119	>250	>250	>250	>250	2.46 ± 1.07 (0.29)	4.25 ± 1.12 (0.28)	>250	>250
120	>250	>250	>250	>250	1.55 ± 1.07 (0.29)	8.36 ± 1.13 (0.25)	>250	>250

Table B10. AlphaScreen™ IC_{50} values (corresponding to Table 4.9). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **134**	>250	>250	>250	>250	2.07 ± 0.20 (0.38)	5.99 ± 0.41 (0.35)	10.5 ± 1.31 (0.33)	14.7 ± 1.47 (0.32)	>250
 **135**	>250	>250	>250	>250	3.22 ± 0.25 (0.35)	7.47 ± 0.59 (0.33)	15.6 ± 2.71 (0.31)	19.6 ± 1.64 (0.30)	>250
 **136**	>250	123 ± 32.2 (0.25)	155 ± 36.7 (0.24)	9.51 ± 1.04 (0.33)	9.01 ± 1.66 (0.32)	22.1 ± 3.59 (0.30)	81.3 ± 16.6 (0.26)	96.0 ± 24.9 (0.26)	>250
 **128**	>250	>250	>250	7.83 ± 2.16 (0.30)	4.32 ± 0.35 (0.31)	15.6 ± 1.93 (0.28)	58.9 ± 20.3 (0.25)	60.2 ± 7.42 (0.20)	>250
	>250	>250	>250	>250	>250	>250	>250	>250	>250

| | | |

138	>250	>250	>250	>250	2.01 ± 0.24 (0.28)	5.94 ± 0.61 (0.26)	>250	53.3 ± 7.89 (0.21)	>250	>250	>250	>250
139	>250	>250	>250	>250	>250	23.8 ± 4.58 (0.22)	>250	>250	>250	>250	>250	>250

Table B11. AlphaScreen™ IC_{50} values (corresponding to Table 4.10). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

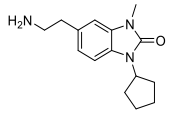
Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **144**	>250	>250	>250	>250	41.4 ± 0.32 (0.32)	118 ± 71.0 (0.29)	>250	182 ± 50.9 (0.28)	>250

Table B12. AlphaScreen™ IC_{50} values of control compound 62 with TRIM33β-E984C. H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

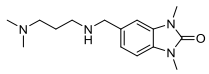
Compound	AlphaScreen™ IC_{50} (μM)		
	TRIM33β-E984C		
	H3(PHD)	H3(BRD)	H3(Dual)
 **62**	>250	21.4 ± 2.67	65.4 ± 22.3

Table B13. AlphaScreen™ IC_{50} values (corresponding to Table 5.1). H3(BRD) – H3K18Ac. AlphaScreen™ colour scale from low (red) to high IC_{50} (green).

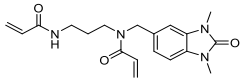
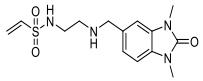
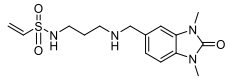
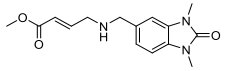
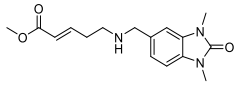
Compound	AlphaScreen™ IC_{50} (μM)								
	TRIM24			TRIM33α			TRIM33β		
	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)	H3(PHD)	H3(BRD)	H3(Dual)
 **151**	>250	>250	>250	>250	>250	>250	>250	>250	>250
 **146**	>250	>250	>250	>250	>250	>250	>250	>250	>250
 **152**	>250	>250	>250	>250	>250	>250	>250	>250	>250
 **160**	>250	>250	>250	>250	18.7 ± 3.16 (0.32)	>250	>250	85.2 ± 20.1 (0.27)	>250
 **147**	>250	>250	>250	>250	3.81 ± 0.68 (0.34)	26.1 ± 3.50 (0.29)	>250	12.6 ± 1.88 (0.31)	>250

Table B14. Results from AlphaScreen™ (Chapter 6). Colour scale from low emission (red) to high emission (green); T24 – TRIM24, T33α – TRIM33α, T33β – TRIM33β, H3(PHD) – H3K9Me$_3$, H3(BRD) – H3K18Ac, H3(Dual) – H3K9Me$_3$K18Ac.

Compound \ Concentration (μM)	Emission (as a percentage of DMSO control)																	
	T24-H3(PHD)		T24-H3(BRD)		T24-H3(Dual)		T33α-H3(PHD)		T33α-H3(BRD)		T33α-H3(Dual)		T33β-H3(PHD)		T33β-H3(BRD)		T33β-H3(Dual)	
	30	150	30	150	30	150	30	150	30	150	30	150	30	150	30	150	30	150
28	106	106	101	105	94	95	121	127	103	79	107	104	101	101	103	102	95	97
163	110	108	100	102	92	91	131	130	124	86	115	104	100	100	100	99	98	100
164	114	110	101	97	93	96	129	128	107	79	105	94	99	104	99	105	98	100
165	108	105	102	100	97	96	127	133	109	89	104	95	93	98	102	96	95	92
166	107	106	91	91	93	92	107	147	113	85	103	104	104	101	98	105	98	96
167	114	114	95	92	98	99	121	131	107	76	101	79	103	102	98	100	96	93
168	110	117	91	91	94	95	123	132	93	67	99	87	98	98	103	98	99	98
169	106	106	99	97	90	92	133	113	116	102	113	105	92	97	96	94	99	101
170	105	107	89	96	114	97	109	104	68	85	69	97	100	98	108	100	90	104
171	98	104	95	96	94	101	96	78	102	46	105	80	95	93	95	95	100	98
172	115	111	107	94	113	103	119	118	115	94	120	111	121	86	123	117	101	97

173	103	103	91	90	96	93	117	125	118	96	100	98	107	98	109	104	91	92
175	104	111	102	70	103	105	110	130	109	133	100	106	108	102	111	102	109	105
176	117	123	108	114	102	131	124	108	132	62	98	89	98	115	125	122	123	91
177A	101	93	91	89	94	89	116	122	109	92	99	97	90	90	94	87	93	101
177B	101	91	91	85	94	85	114	92	106	74	96	77	88	80	85	80	94	86
178	99	106	108	124	116	130	110	157	53	97	61	94	106	108	133	152	104	108
179	109	129	126	139	126	138	134	126	123	116	107	108	110	123	149	156	107	115
185	105	106	93	82	129	122	105	92	126	102	112	96	109	106	139	134	100	91
186	111	126	84	103	128	131	128	119	130	131	101	115	110	117	141	145	86	87
190	105	110	128	132	122	126	113	70	110	70	98	84	106	105	139	143	101	98
191	96	79	122	102	113	93	91	41	97	39	94	61	99	85	129	106	87	72
195	101	101	109	106	119	116	112	78	118	81	95	84	105	102	122	117	86	81

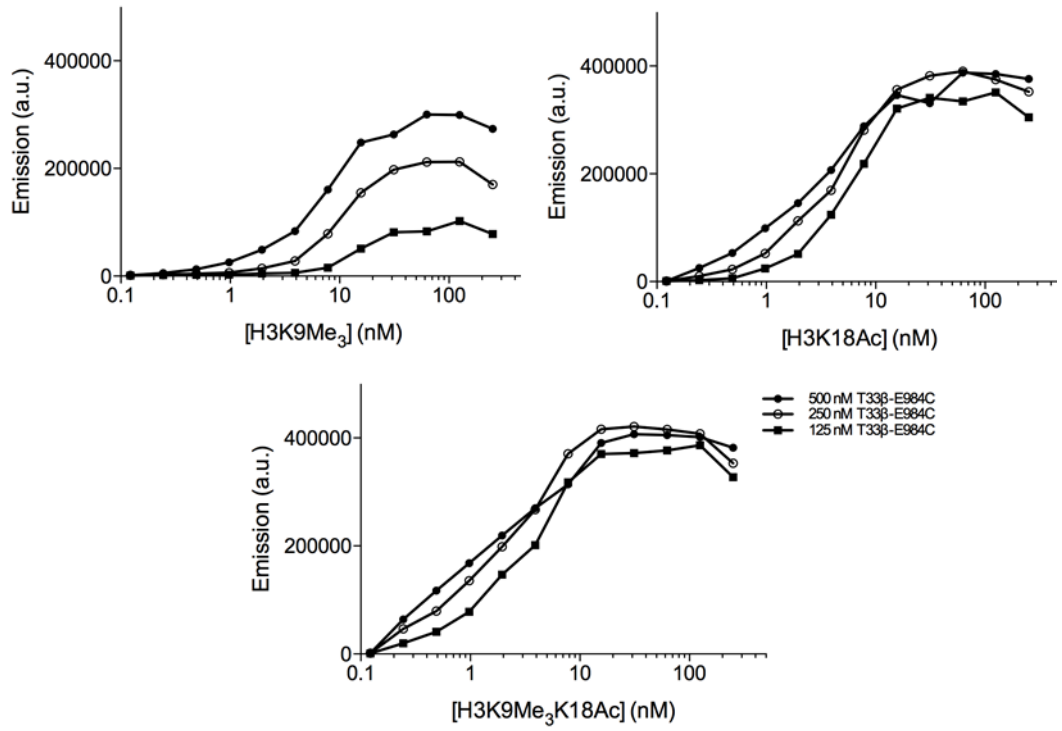


Figure B3. AlphaScreen™ signal response experiments for TRIM33β-E984C with each of the three peptide binding partners.

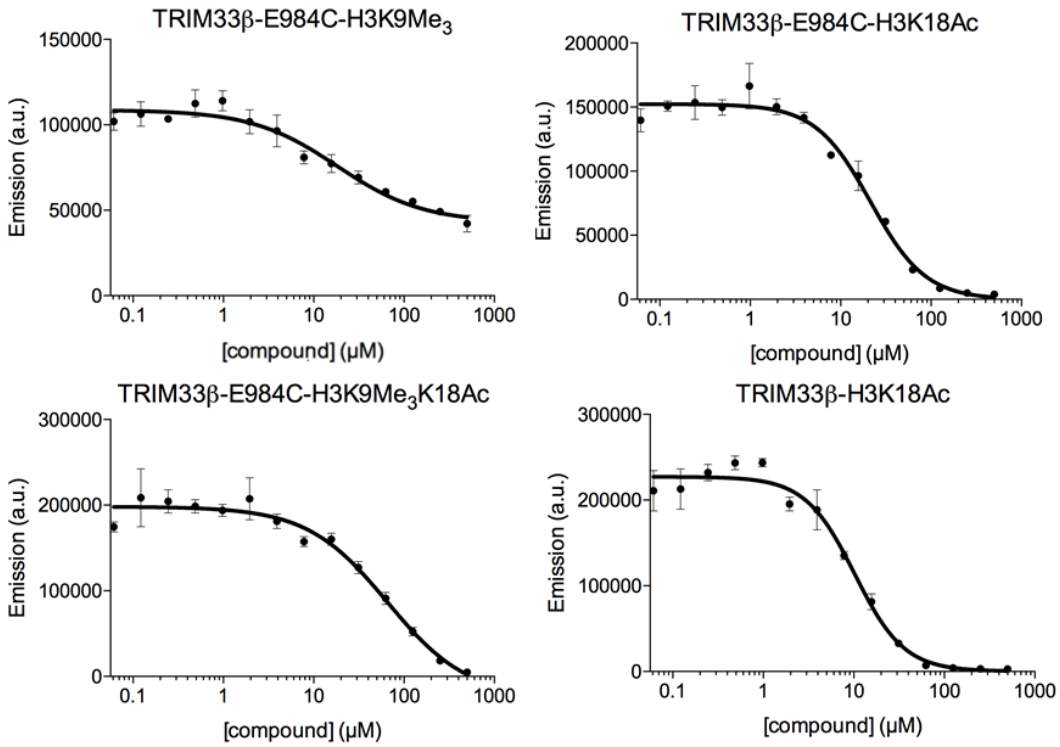
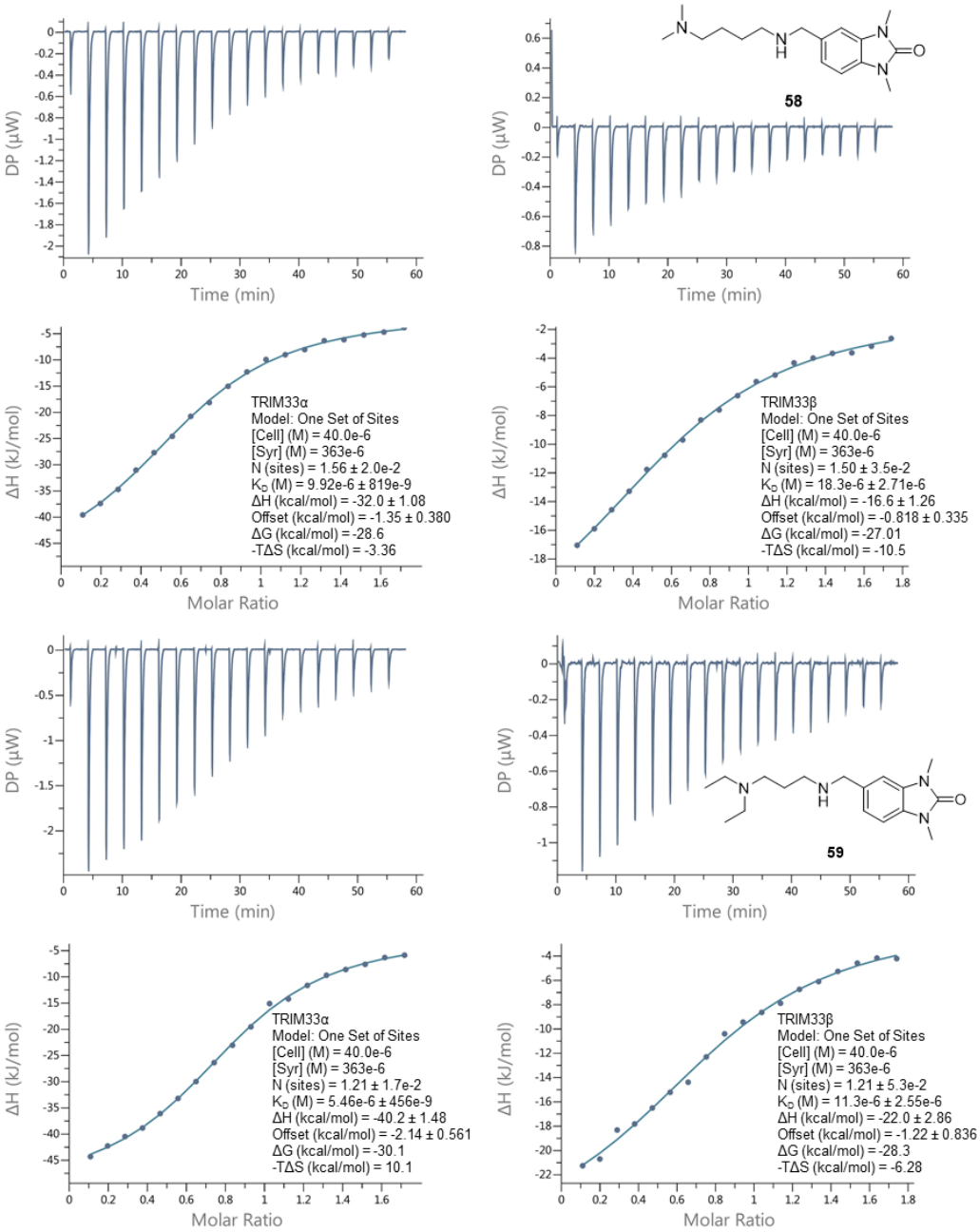
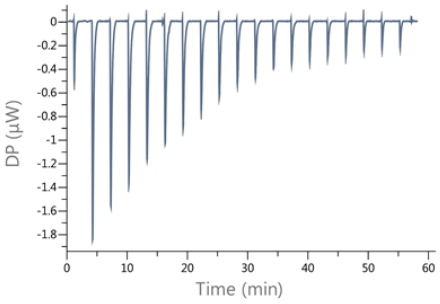
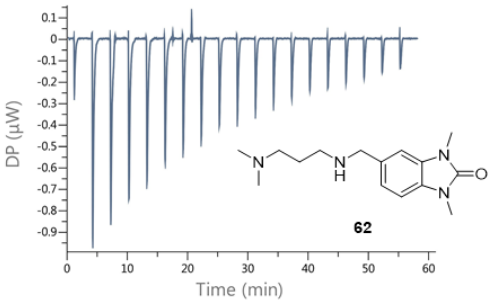


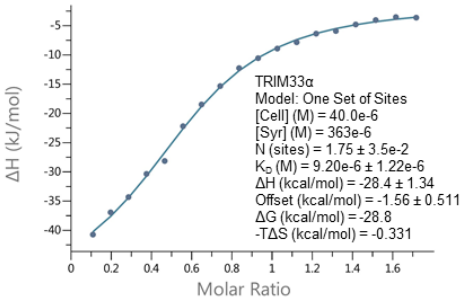
Figure B4. AlphaScreen™ dose response experiments for TRIM33β-E984C with control compound 62 each of the three peptide binding partners. Dose response curve with TRIM33β included for comparison.

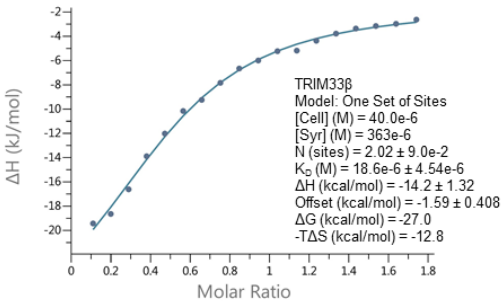
C. ITC Curves

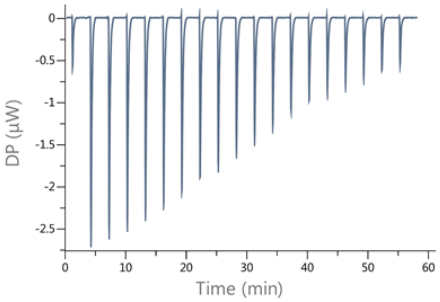

DP (µW)
Time (min)
58
ΔH (kJ/mol)
Molar Ratio
TRIM33α
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 1.56 ± 2.0e-2
K_D (M) = 9.92e-6 ± 819e-9
ΔH (kcal/mol) = -32.0 ± 1.08
Offset (kcal/mol) = -1.35 ± 0.380
ΔG (kcal/mol) = -28.6
-TΔS (kcal/mol) = -3.36
TRIM33β
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 1.50 ± 3.5e-2
K_D (M) = 18.3e-6 ± 2.71e-6
ΔH (kcal/mol) = -16.6 ± 1.26
Offset (kcal/mol) = -0.818 ± 0.335
ΔG (kcal/mol) = -27.01
-TΔS (kcal/mol) = -10.5
59
TRIM33α
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 1.21 ± 1.7e-2
K_D (M) = 5.46e-6 ± 456e-9
ΔH (kcal/mol) = -40.2 ± 1.48
Offset (kcal/mol) = -2.14 ± 0.561
ΔG (kcal/mol) = -30.1
-TΔS (kcal/mol) = 10.1
TRIM33β
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 1.21 ± 5.3e-2
K_D (M) = 11.3e-6 ± 2.55e-6
ΔH (kcal/mol) = -22.0 ± 2.86
Offset (kcal/mol) = -1.22 ± 0.836
ΔG (kcal/mol) = -28.3
-TΔS (kcal/mol) = -6.28

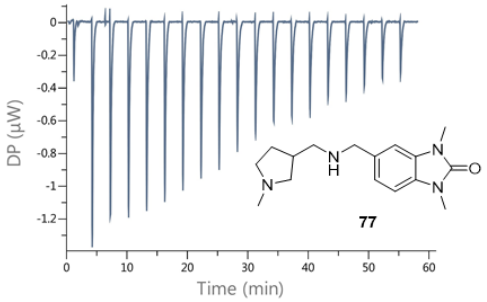

DP (µW)
Time (min)

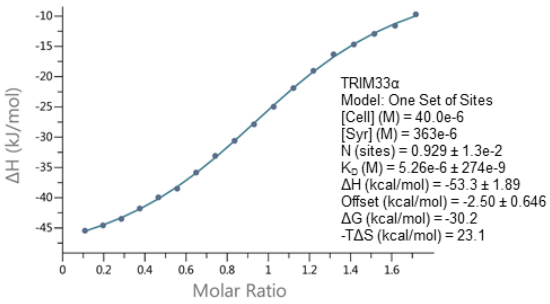

DP (µW)
Time (min)
62

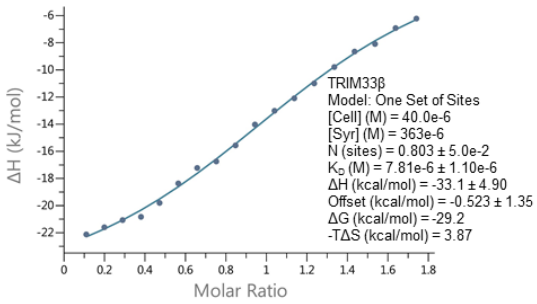

ΔH (kJ/mol)
Molar Ratio
TRIM33α
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 1.75 ± 3.5e-2
K_D (M) = 9.20e-6 ± 1.22e-6
ΔH (kcal/mol) = -28.4 ± 1.34
Offset (kcal/mol) = -1.56 ± 0.511
ΔG (kcal/mol) = -28.8
-TΔS (kcal/mol) = -0.331

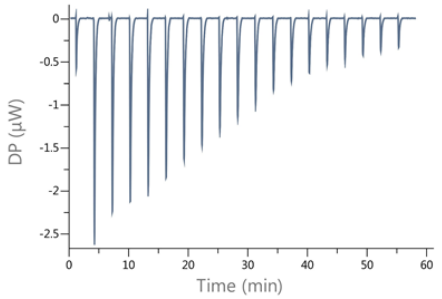

ΔH (kJ/mol)
Molar Ratio
TRIM33β
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 2.02 ± 9.0e-2
K_D (M) = 18.6e-6 ± 4.54e-6
ΔH (kcal/mol) = -14.2 ± 1.32
Offset (kcal/mol) = -1.59 ± 0.408
ΔG (kcal/mol) = -27.0
-TΔS (kcal/mol) = -12.8

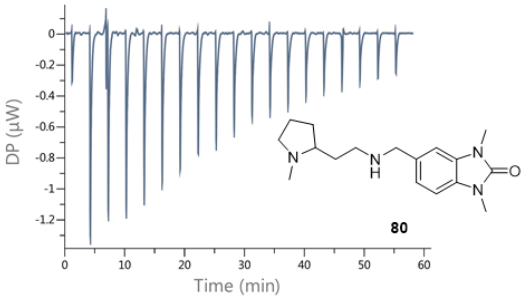

DP (µW)
Time (min)

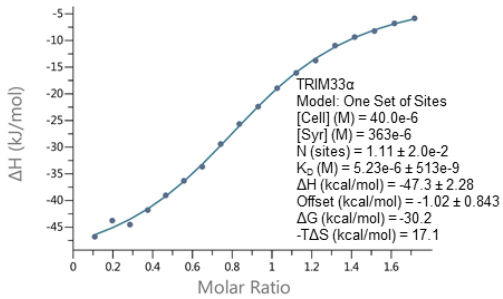

DP (µW)
Time (min)
77

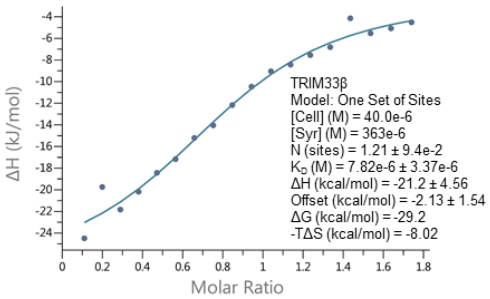

ΔH (kJ/mol)
Molar Ratio
TRIM33α
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 0.929 ± 1.3e-2
K_D (M) = 5.26e-6 ± 274e-9
ΔH (kcal/mol) = -53.3 ± 1.89
Offset (kcal/mol) = -2.50 ± 0.646
ΔG (kcal/mol) = -30.2
-TΔS (kcal/mol) = 23.1

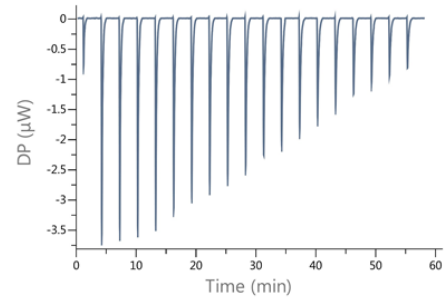

ΔH (kJ/mol)
Molar Ratio
TRIM33β
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 0.803 ± 5.0e-2
K_D (M) = 7.81e-6 ± 1.10e-6
ΔH (kcal/mol) = -33.1 ± 4.90
Offset (kcal/mol) = -0.523 ± 1.35
ΔG (kcal/mol) = -29.2
-TΔS (kcal/mol) = 3.87

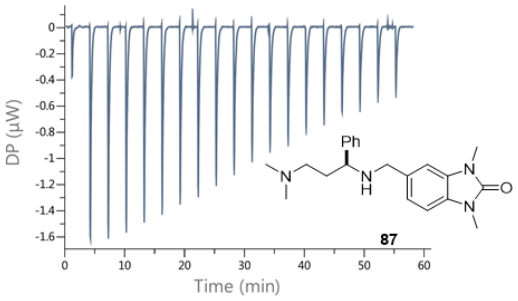

DP (µW)
Time (min)

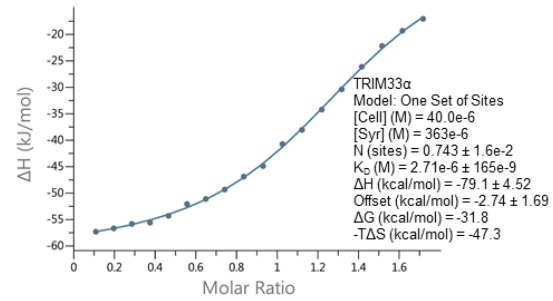

DP (µW)
Time (min)
80

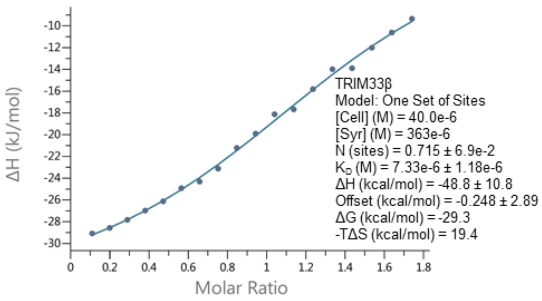

ΔH (kJ/mol)
Molar Ratio
TRIM33α
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 1.11 ± 2.0e-2
K_D (M) = 5.23e-6 ± 513e-9
ΔH (kcal/mol) = -47.3 ± 2.28
Offset (kcal/mol) = -1.02 ± 0.843
ΔG (kcal/mol) = -30.2
-TΔS (kcal/mol) = 17.1


ΔH (kJ/mol)
Molar Ratio
TRIM33β
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 1.21 ± 9.4e-2
K_D (M) = 7.82e-6 ± 3.37e-6
ΔH (kcal/mol) = -21.2 ± 4.56
Offset (kcal/mol) = -2.13 ± 1.54
ΔG (kcal/mol) = -29.2
-TΔS (kcal/mol) = -8.02


DP (µW)
Time (min)

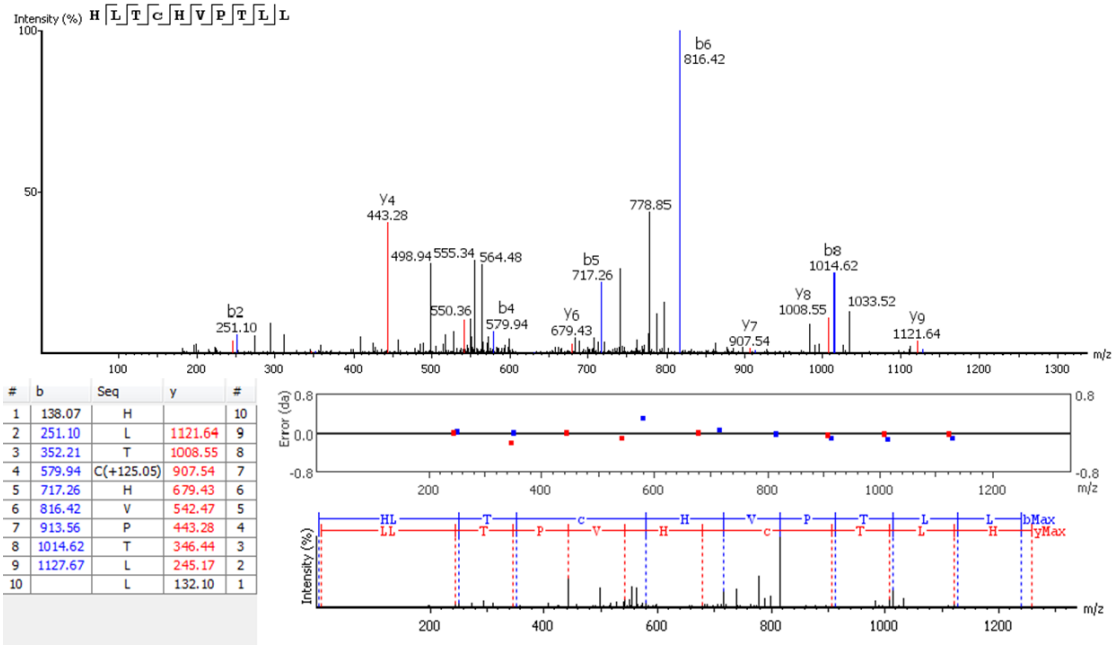

DP (µW)
Time (min)
Ph
87


ΔH (kJ/mol)
Molar Ratio
TRIM33α
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 0.743 ± 1.6e-2
K_D (M) = 2.71e-6 ± 165e-9
ΔH (kcal/mol) = -79.1 ± 4.52
Offset (kcal/mol) = -2.74 ± 1.69
ΔG (kcal/mol) = -31.8
-TΔS (kcal/mol) = -47.3

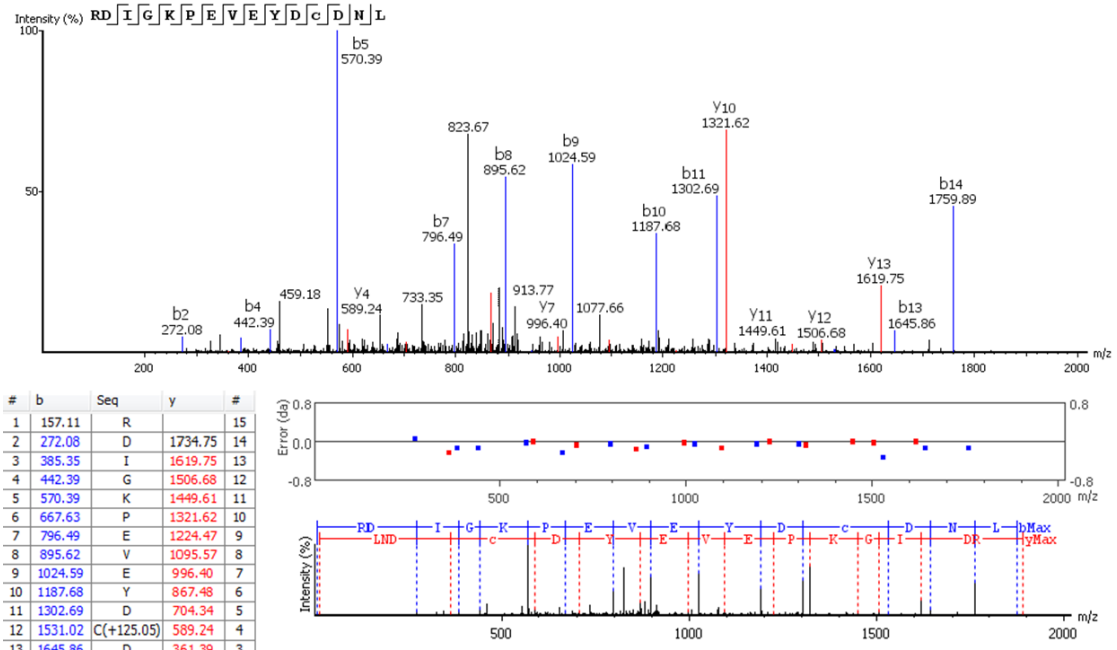

ΔH (kJ/mol)
Molar Ratio
TRIM33β
Model: One Set of Sites
[Cell] (M) = 40.0e-6
[Syr] (M) = 363e-6
N (sites) = 0.715 ± 6.9e-2
K_D (M) = 7.33e-6 ± 1.18e-6
ΔH (kcal/mol) = -48.8 ± 10.8
Offset (kcal/mol) = -0.248 ± 2.89
ΔG (kcal/mol) = -29.3
-TΔS (kcal/mol) = 19.4

D. Alkylation Experiments

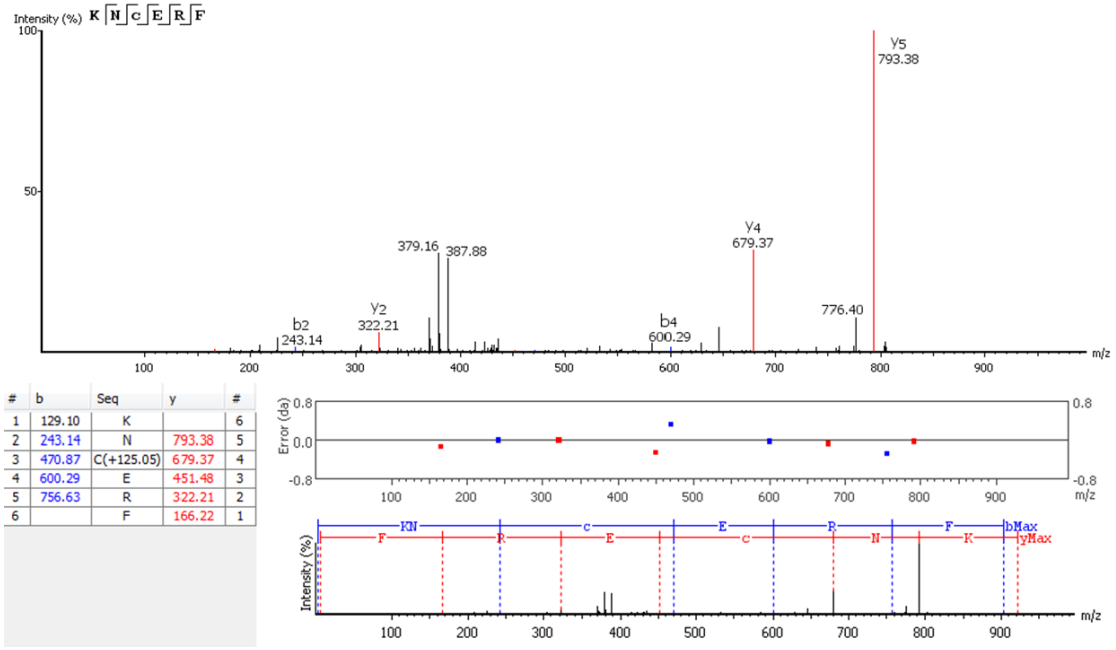
D1 Proteomics Data from Alkylation with *N*-Ethylmaleimide



#	b	Seq	y	#
1	138.07	H		10
2	251.10	L	1121.64	9
3	352.21	T	1008.55	8
4	579.94	C(+125.05)	907.54	7
5	717.26	H	679.43	6
6	816.42	V	542.47	5
7	913.56	P	443.28	4
8	1014.62	T	346.44	3
9	1127.67	L	245.17	2
10		L	132.10	1



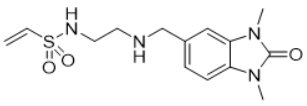
#	b	Seq	y	#
1	157.11	R		15
2	272.08	D	1734.75	14
3	385.35	I	1619.75	13
4	442.39	G	1506.68	12
5	570.39	K	1449.61	11
6	667.63	P	1321.62	10
7	796.49	E	1224.47	9
8	895.62	V	1095.57	8
9	1024.59	E	996.40	7
10	1187.68	Y	867.48	6
11	1302.69	D	704.34	5
12	1531.02	C(+125.05)	589.24	4
13	1645.86	D	361.39	3
14	1759.89	N	246.14	2
15		L	132.10	1



#	b	Seq	y	#
1	129.10	K		6
2	243.14	N	793.38	5
3	470.87	C(+125.05)	679.37	4
4	600.29	E	451.48	3
5	756.63	R	322.21	2
6		F	166.22	1

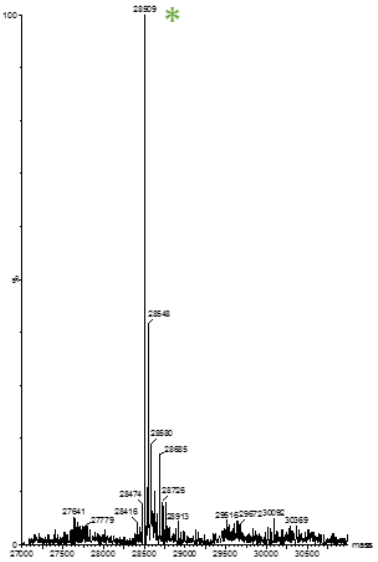
D2 Intact Mass Spectrometry Data from Alkylation with Covalent Ligands

TRIM33β-WT, MW 28509
146, MW 324
Expected = 28833 (+ 1 compound);
29157 (+2 compounds)

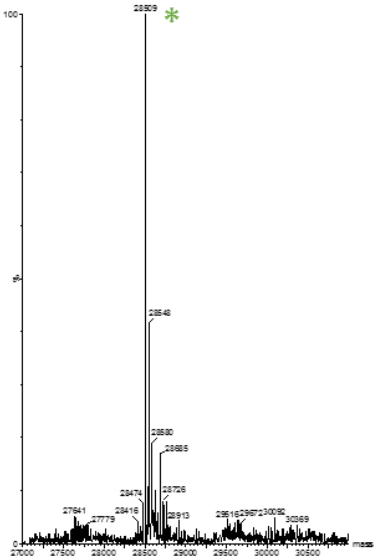


* Protein
\# Protein + 1 compound
× Protein + 2 compound

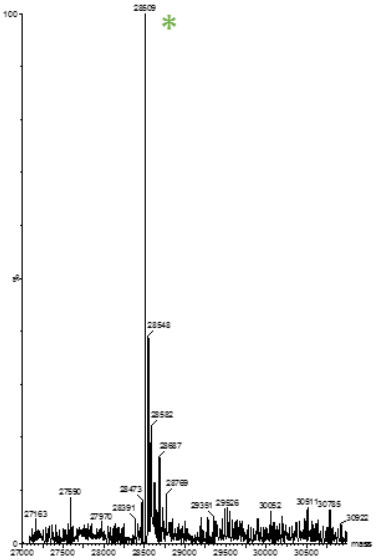
t = 0 h



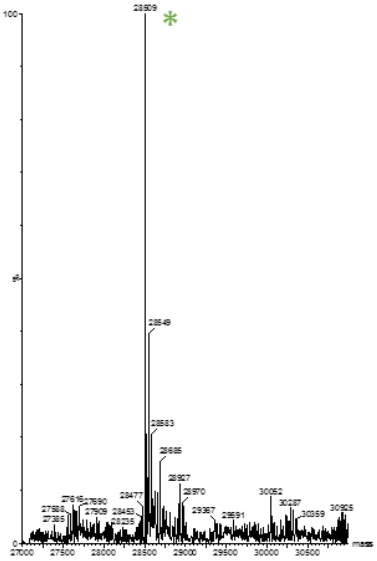
t = 0.5 h



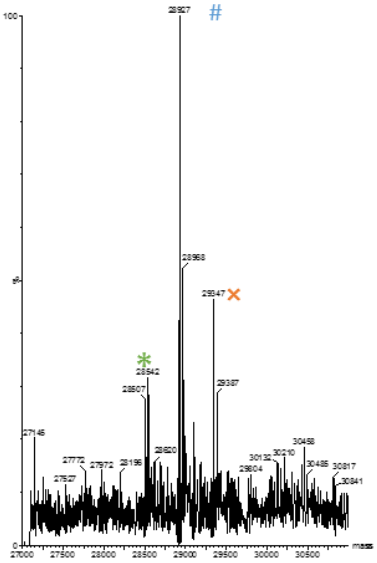
t = 1 h



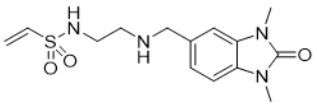
t = 3 h



t = 24 h

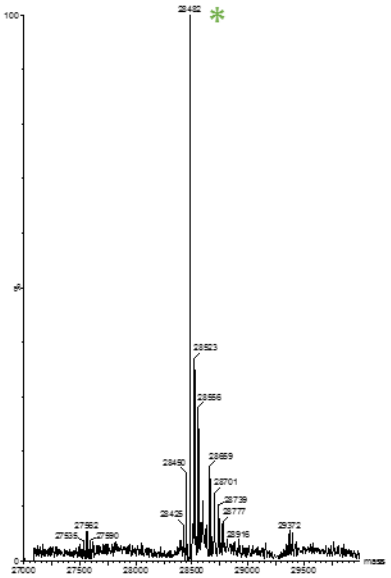


TRIM33β-EC, MW 28483
146, MW 324
Expected = 28806 (+ 1 compound);
29130 (+2 compounds)

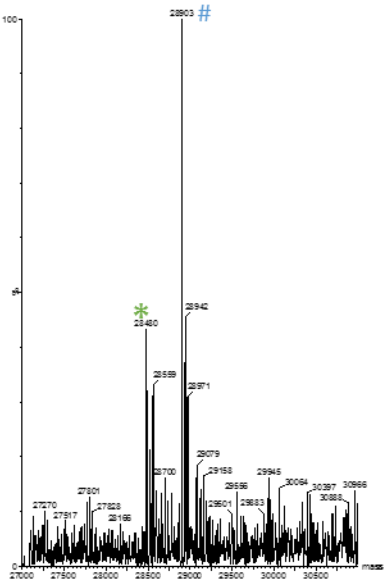


* Protein
\# Protein + 1 compound
× Protein + 2 compound

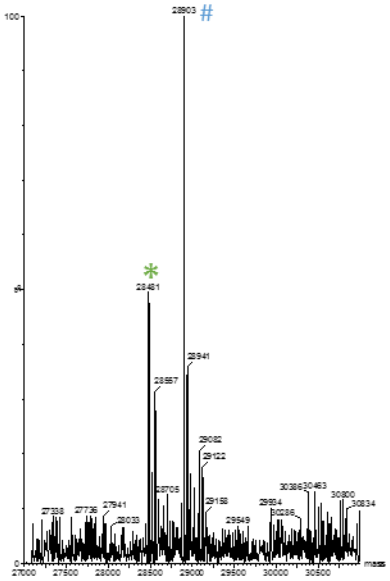
t = 0 h



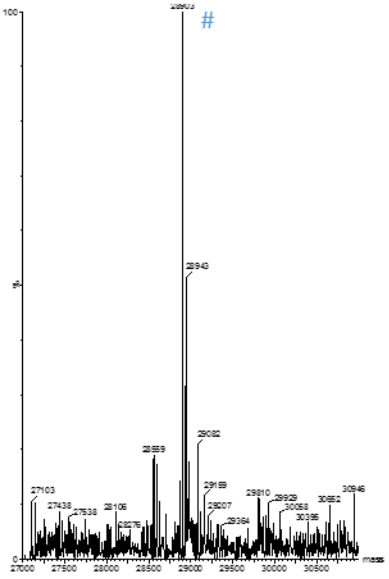
t = 0.5 h



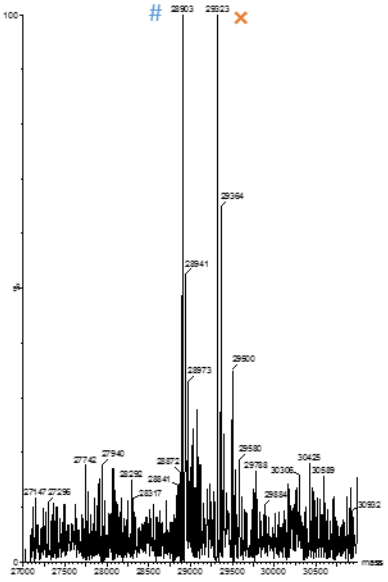
t = 1 h



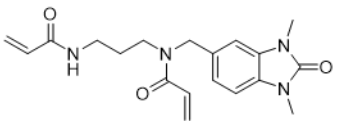
t = 3 h



t = 24 h

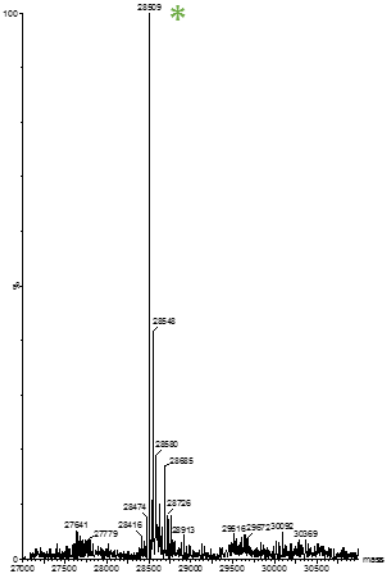


TRIM33β-WT, MW 28509
151, MW 356
Expected = 28864 (+ 1 compound);
29220 (+2 compounds)

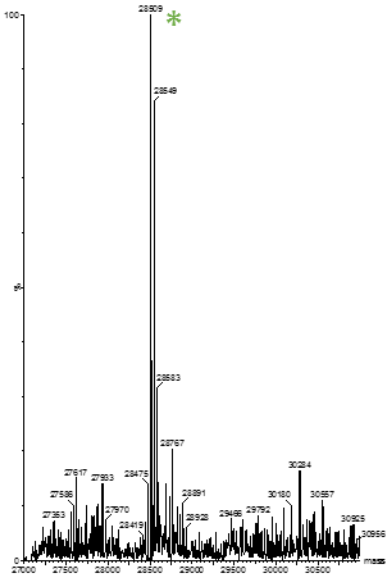


* Protein
\# Protein + 1 compound
× Protein + 2 compound

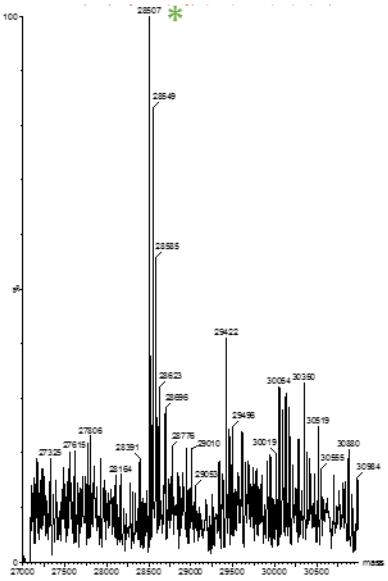
t = 0 h



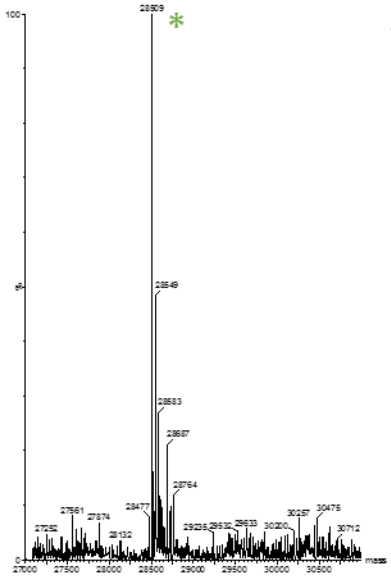
t = 0.5 h



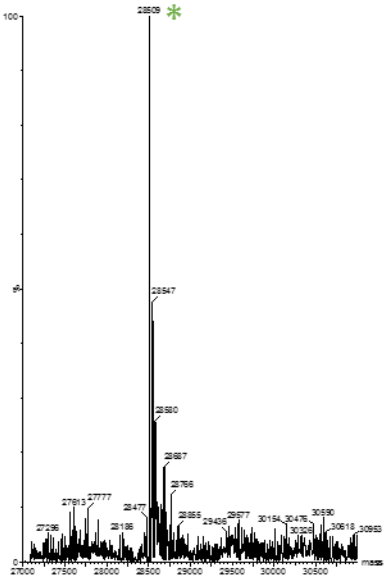
t = 1 h



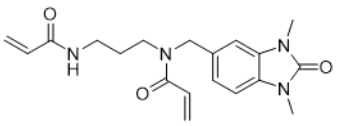
t = 3 h



t = 24 h

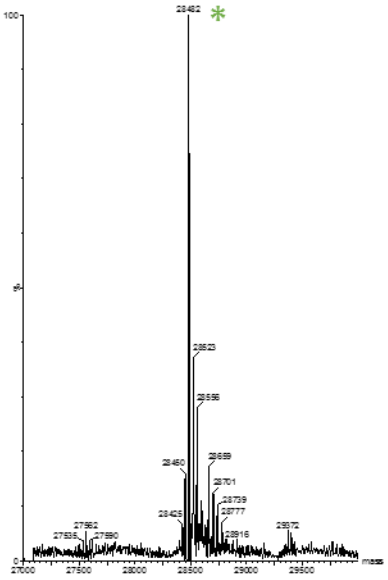


TRIM33β-EC, MW 28483
151, MW 356
Expected = 28838 (+ 1 compound);
29194 (+2 compounds)

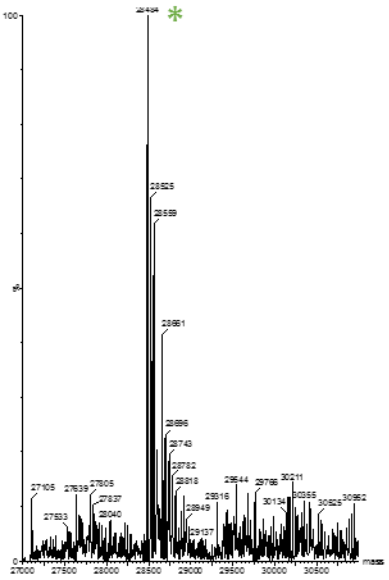


* Protein
\# Protein + 1 compound
× Protein + 2 compound

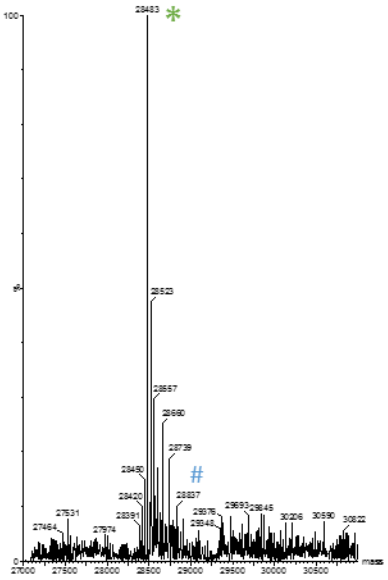
t = 0 h



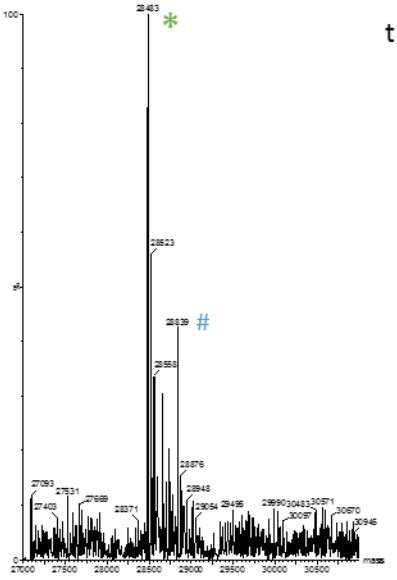
t = 0.5 h



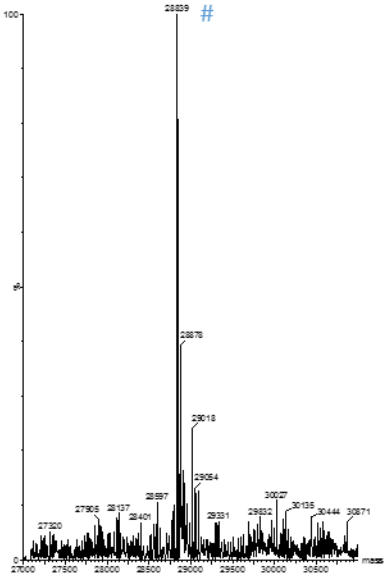
t = 1 h



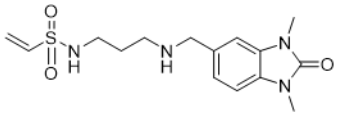
t = 3 h



t = 24 h

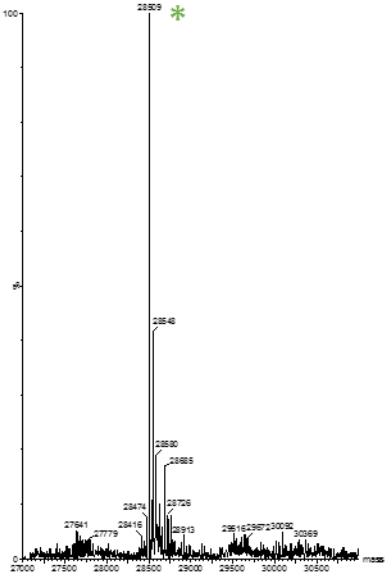


TRIM33β-WT, MW 28509
152, MW 338
Expected = 28846 (+ 1 compound);
29184 (+2 compounds)

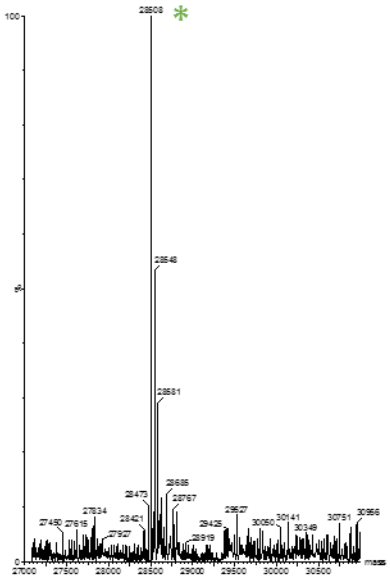


* Protein
\# Protein + 1 compound
× Protein + 2 compound

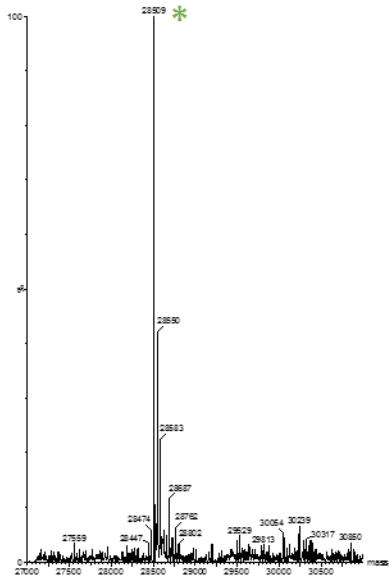
t = 0 h



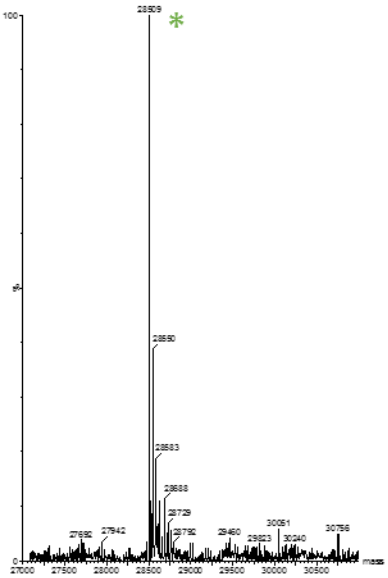
t = 0.5 h



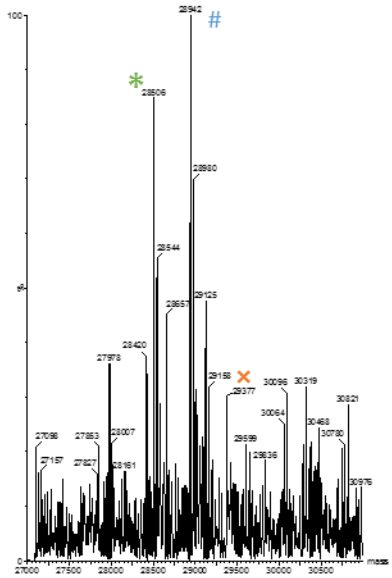
t = 1 h



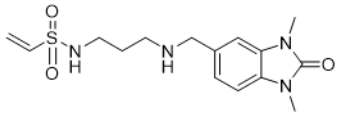
t = 3 h



t = 24 h

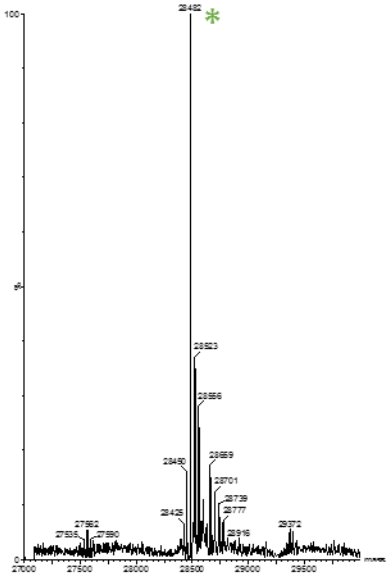


TRIM33β-EC, MW 28483
152, MW 338
Expected = 28820 (+ 1 compound);
29158 (+2 compounds)

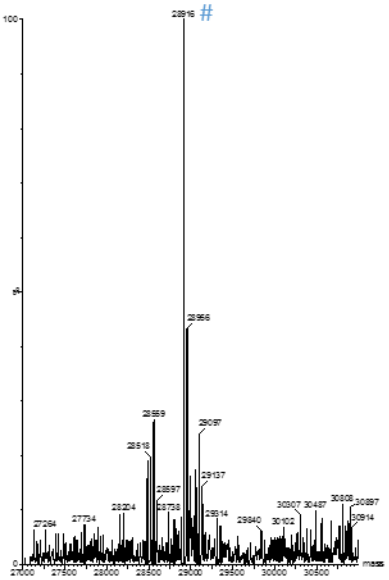


* Protein
\# Protein + 1 compound
× Protein + 2 compound

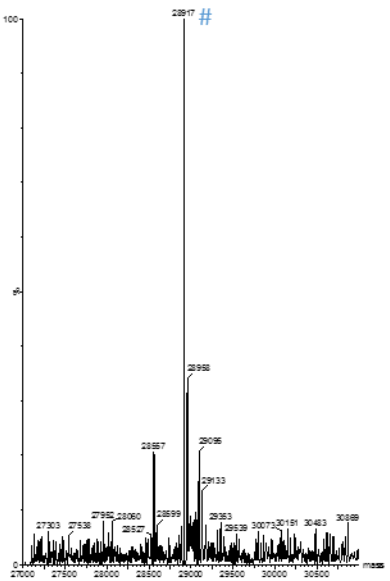
t = 0 h



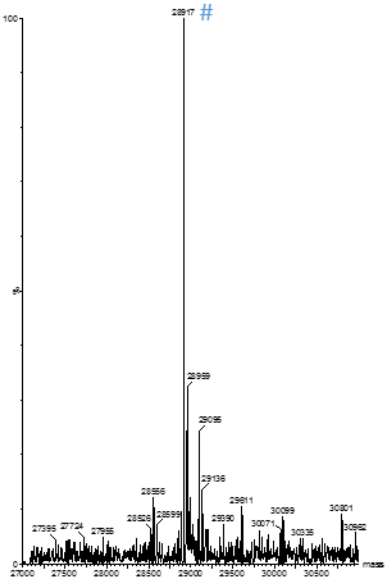
t = 0.5 h



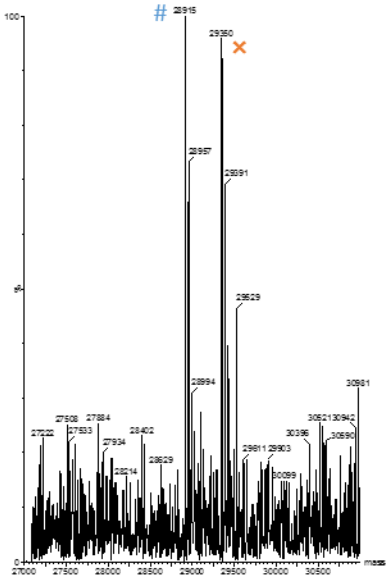
t = 1 h



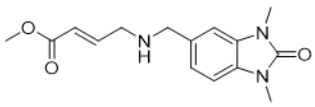
t = 3 h



t = 24 h

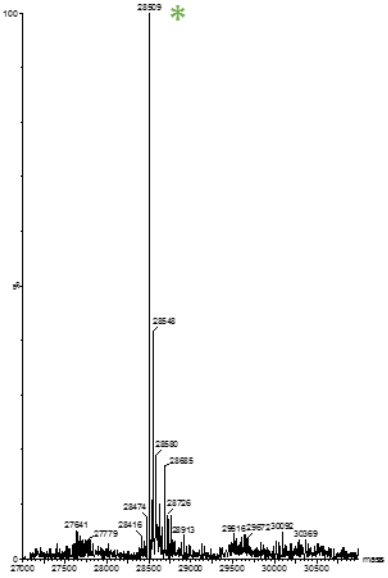


TRIM33β-WT, MW 28509
160, MW 289
Expected = 28797 (+ 1 compound);
29086 (+2 compounds)

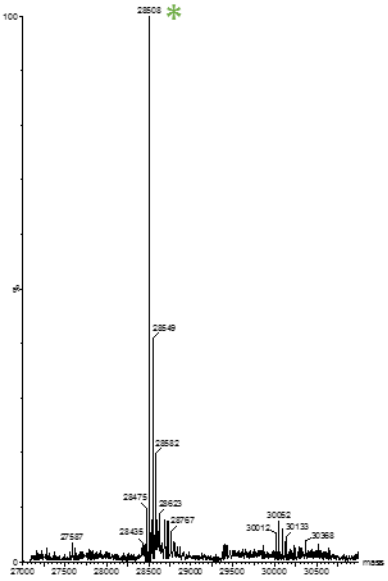


* Protein
\# Protein + 1 compound
× Protein + 2 compound

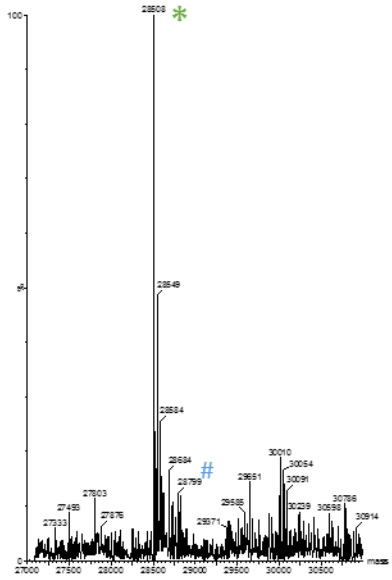
t = 0 h



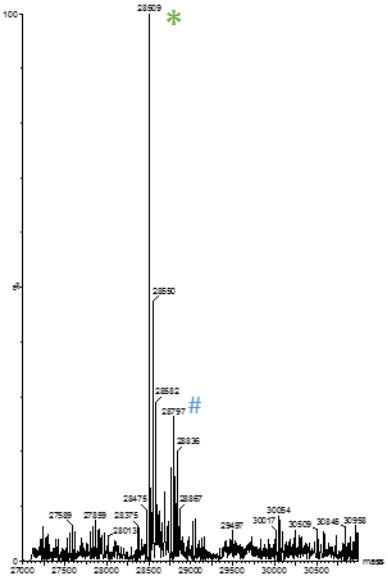
t = 0.5 h



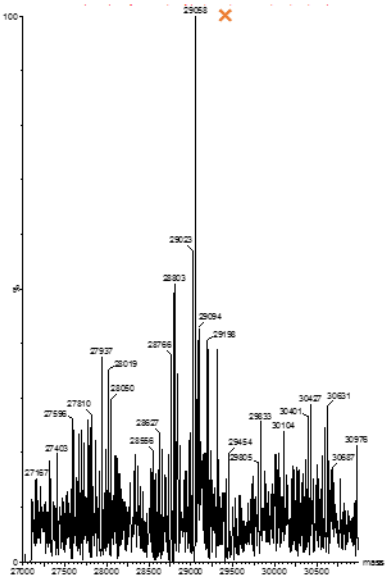
t = 1 h



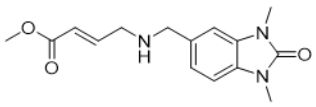
t = 3 h



t = 24 h

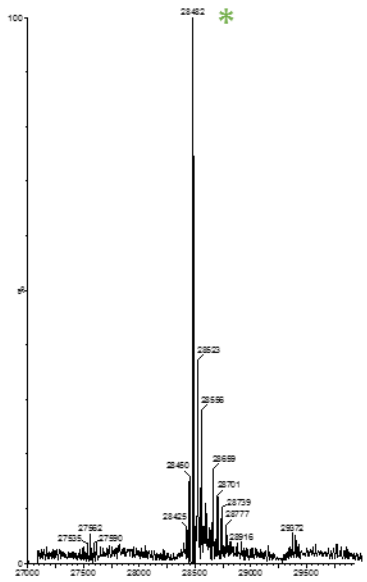


TRIM33β-EC, MW 28483
160, MW 289
Expected = 28771 (+ 1 compound);
29060 (+2 compounds)

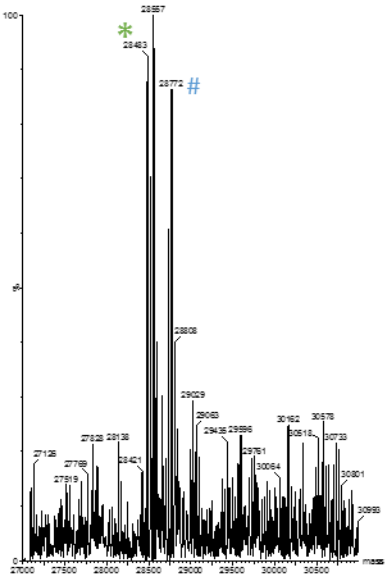


* Protein
\# Protein + 1 compound
× Protein + 2 compound

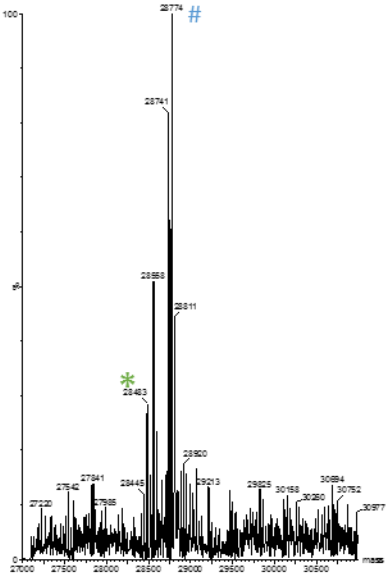
t = 0 h



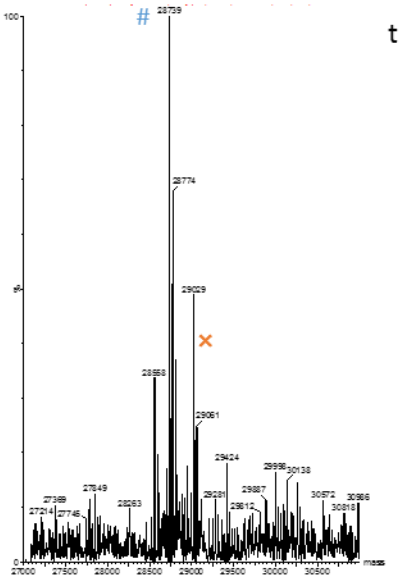
t = 0.5 h



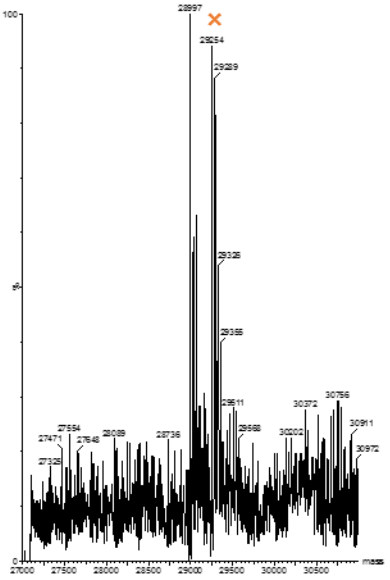
t = 1 h



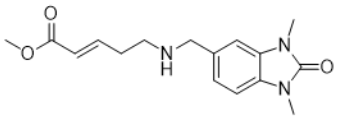
t = 3 h



t = 24 h

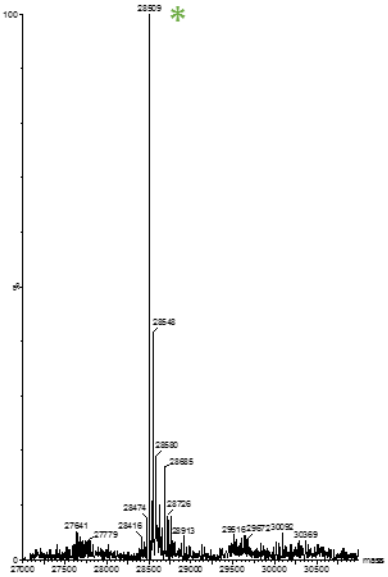


TRIM33β-WT, MW 28509
147, MW 303
Expected = 28811 (+ 1 compound);
29114 (+2 compounds)

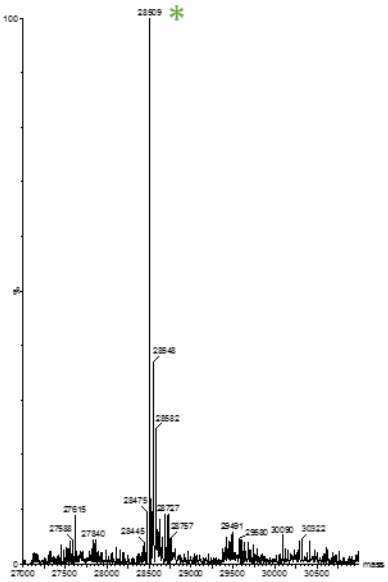


* Protein
\# Protein + 1 compound
× Protein + 2 compound

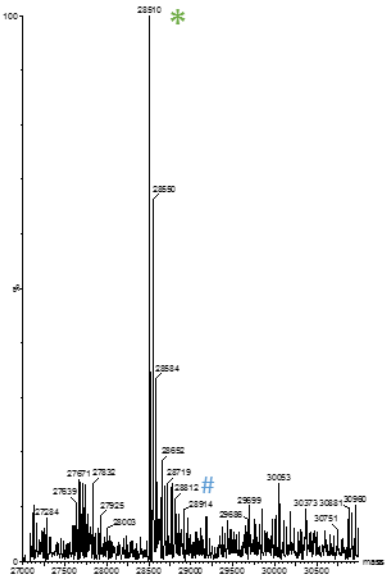
t = 0 h



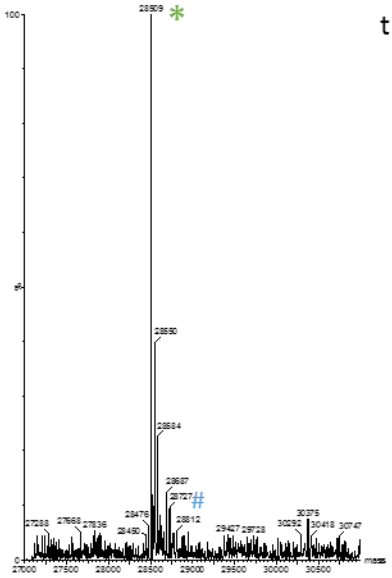
t = 0.5 h



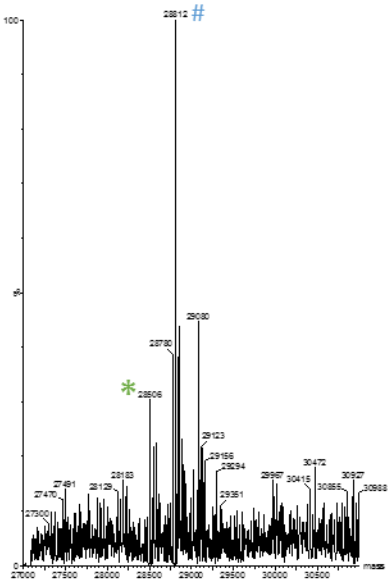
t = 1 h



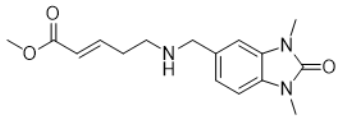
t = 3 h



t = 24 h

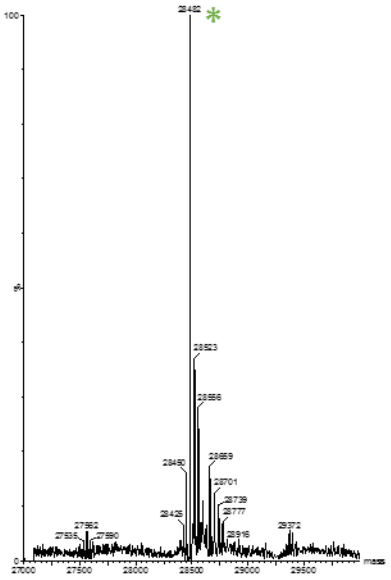


TRIM33β-EC, MW 28483
147, MW 303
Expected = 28786 (+ 1 compound);
29089 (+2 compounds)

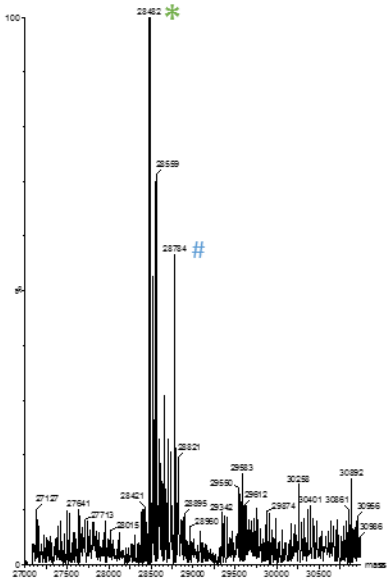


* Protein
\# Protein + 1 compound
× Protein + 2 compound

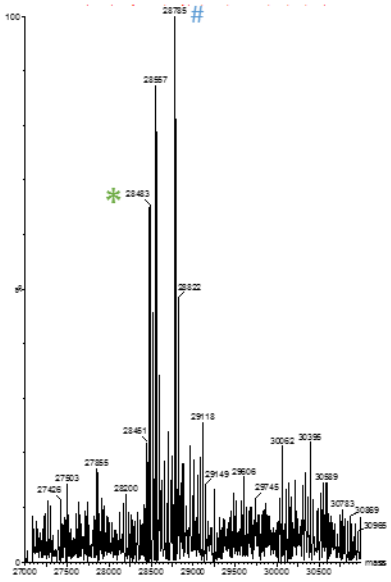
t = 0 h



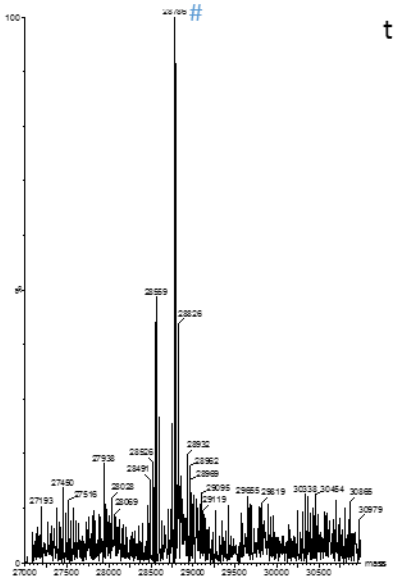
t = 0.5 h



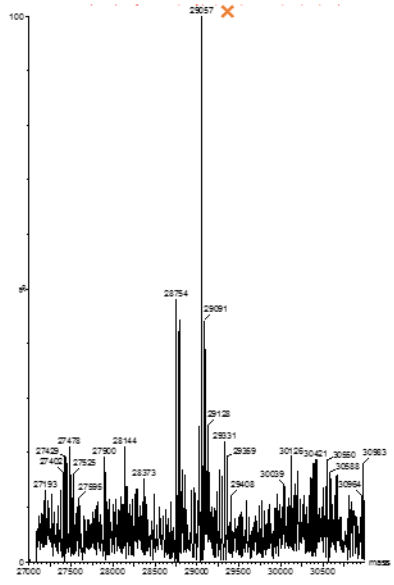
t = 1 h



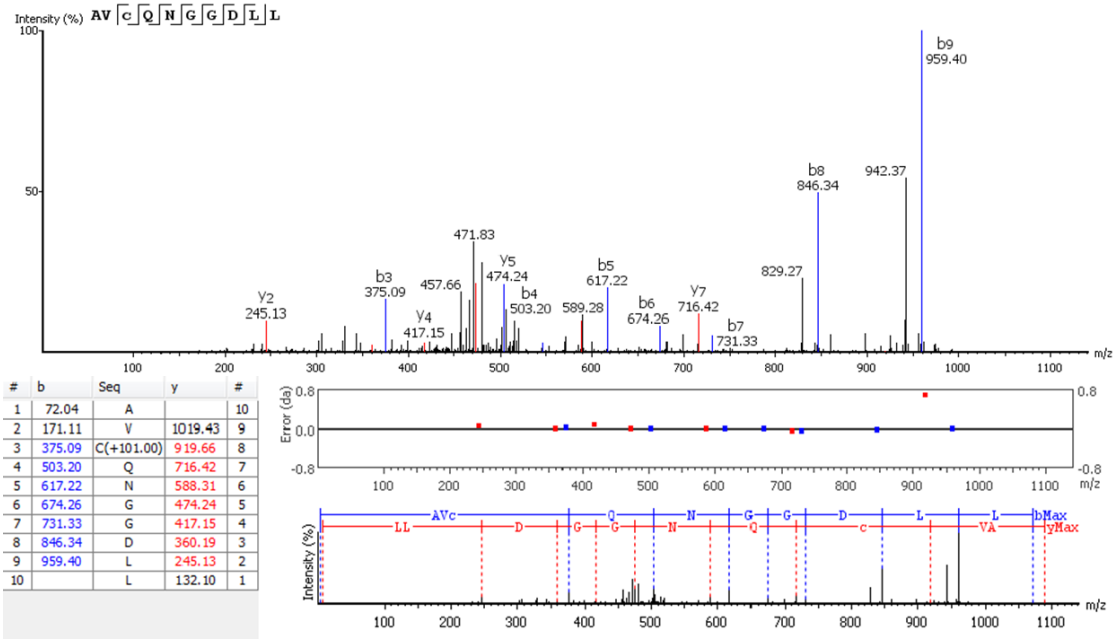
t = 3 h



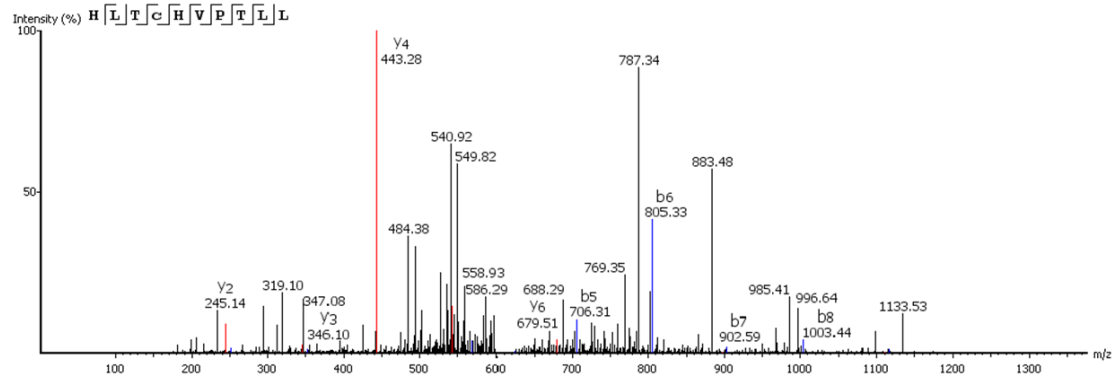
t = 24 h



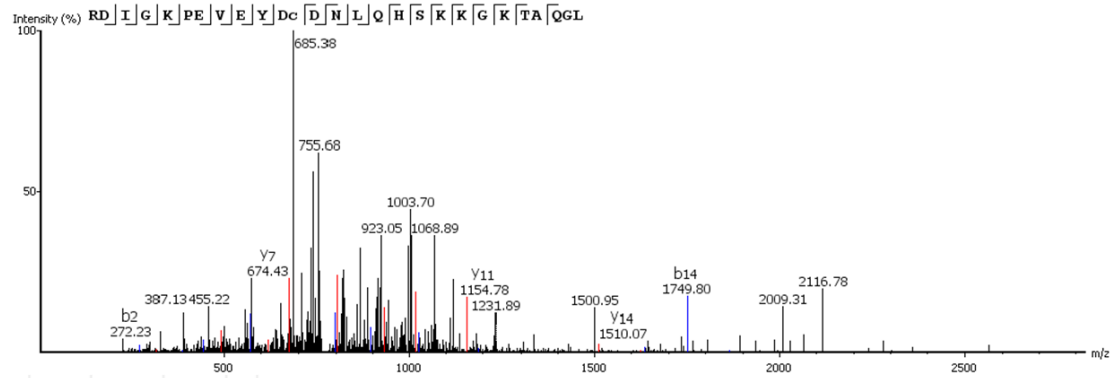
D3 Proteomics Data from Alkylation with Covalent Ligand 160



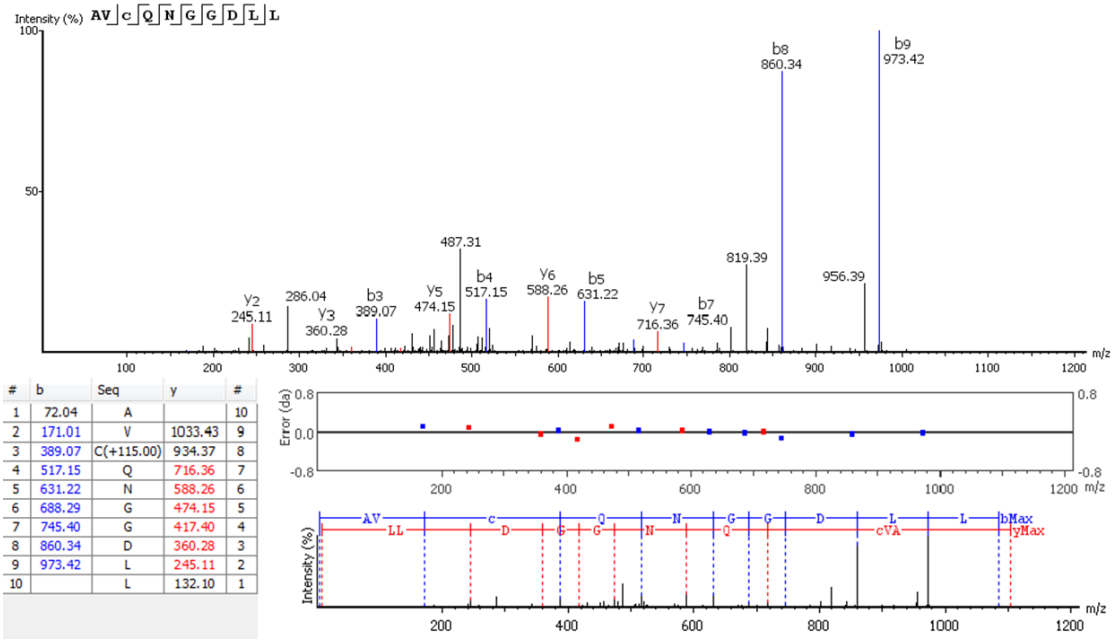
#	b	Seq	y	#
1	72.04	A		10
2	171.11	V	1019.43	9
3	375.09	C(+101.00)	919.66	8
4	503.20	Q	716.42	7
5	617.22	N	588.31	6
6	674.26	G	474.24	5
7	731.33	G	417.15	4
8	846.34	D	360.19	3
9	959.40	L	245.13	2
10		L	132.10	1



#	b	Seq	y	#
1	138.07	H		10
2	251.27	L	1110.76	9
3	352.52	T	997.47	8
4	568.71	C(+114.00)	896.92	7
5	706.31	H	679.51	6
6	805.33	V	542.34	5
7	902.59	P	443.28	4
8	1003.44	T	346.10	3
9	1116.59	L	245.14	2
10		L	132.10	1

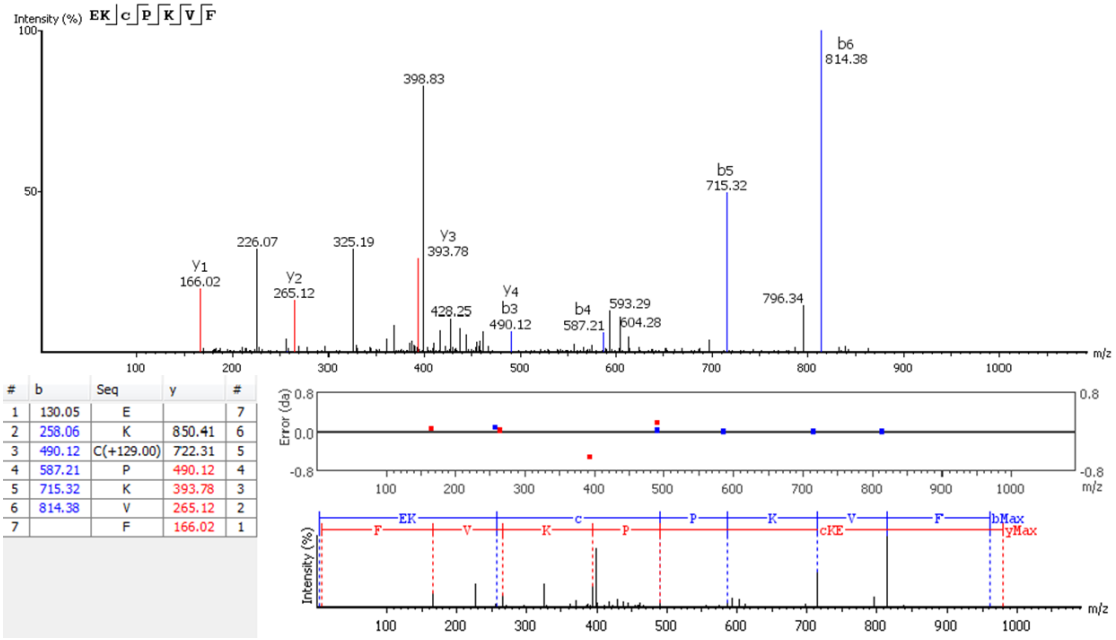


#	b	Seq	y	#
1	157.11	R		27
2	272.23	D	2988.41	26
3	385.33	I	2873.38	25
4	442.52	G	2760.30	24
5	570.28	K	2703.28	23
6	667.39	P	2575.18	22
7	796.42	E	2478.13	21
8	894.95	V	2349.09	20
9	1024.93	E	2250.02	19
10	1187.44	Y	2120.98	18
11	1302.63	D	1957.91	17
12	1520.64	C(+115.00)	1842.89	16
13	1635.79	D	1624.79	15
14	1749.80	N	1510.07	14
15	1862.64	L	1395.81	13
16	1990.85	Q	1282.72	12
17	2127.91	H	1154.78	11
18	2214.95	S	1017.78	10
19	2343.04	K	930.66	9
20	2471.14	K	802.57	8
21	2528.16	G	674.43	7

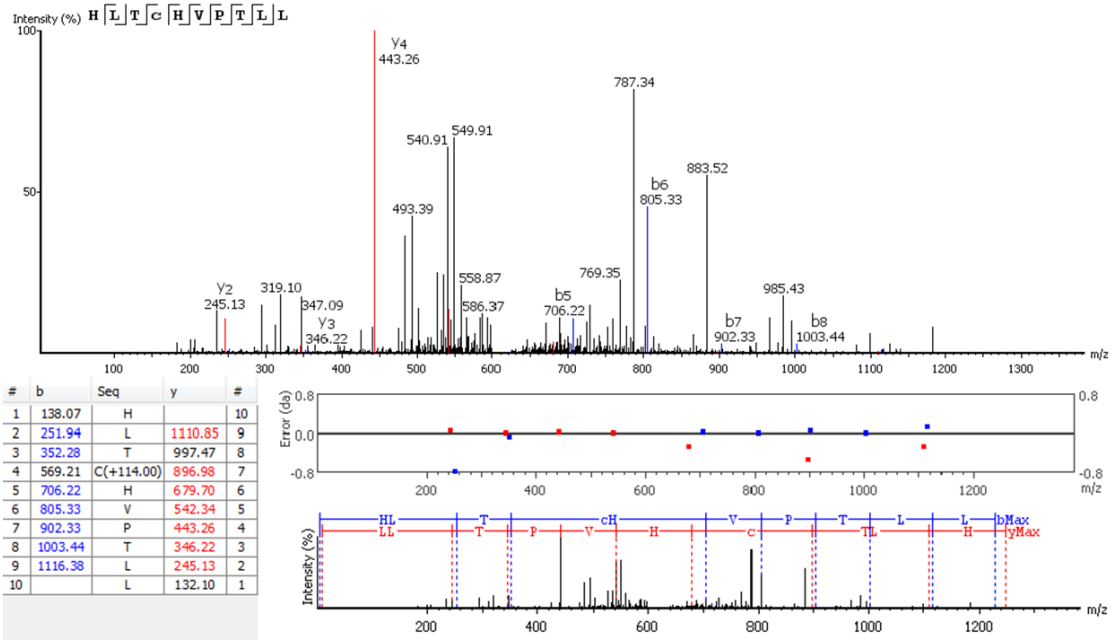


#	b	Seq	y	#
1	72.04	A		10
2	171.01	V	1033.43	9
3	389.07	C(+115.00)	934.37	8
4	517.15	Q	716.36	7
5	631.22	N	588.26	6
6	688.29	G	474.15	5
7	745.40	G	417.40	4
8	860.34	D	360.28	3
9	973.42	L	245.11	2
10		L	132.10	1

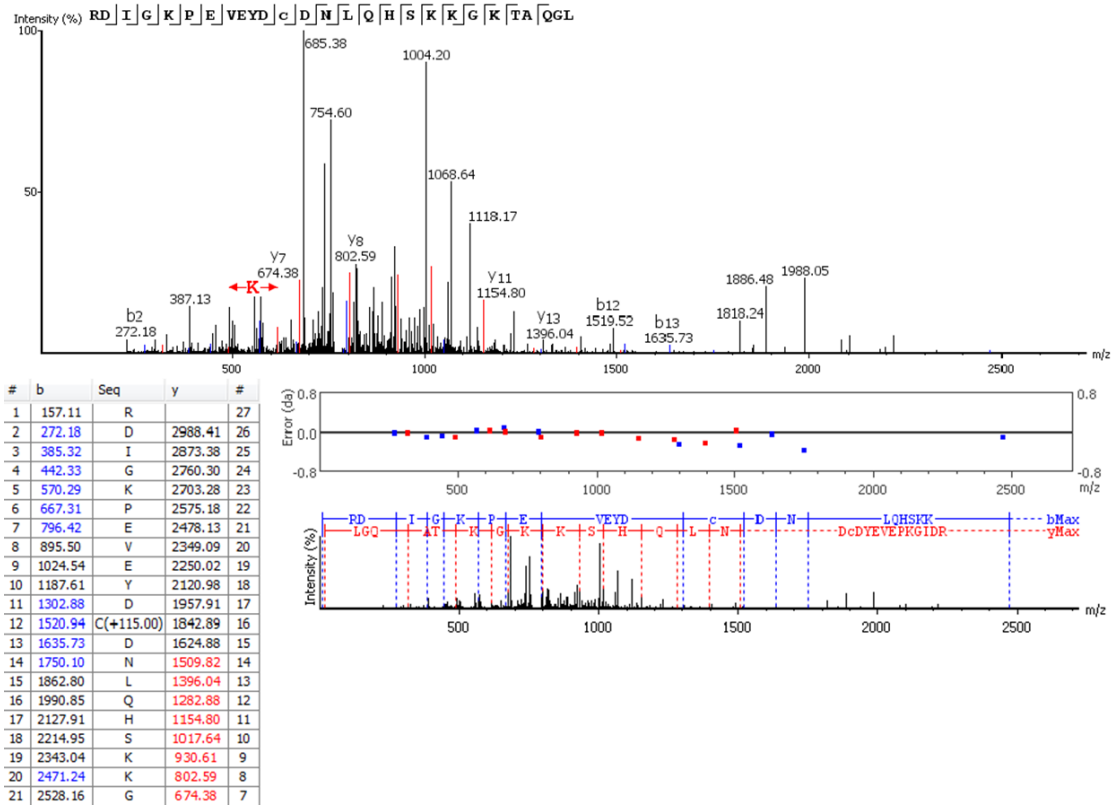
D4 Proteomics Data from Alkylation with Covalent Ligand 147



#	b	Seq	y	#
1	130.05	E		7
2	258.06	K	850.41	6
3	490.12	C(+129.00)	722.31	5
4	587.21	P	490.12	4
5	715.32	K	393.78	3
6	814.38	V	265.12	2
7		F	166.02	1



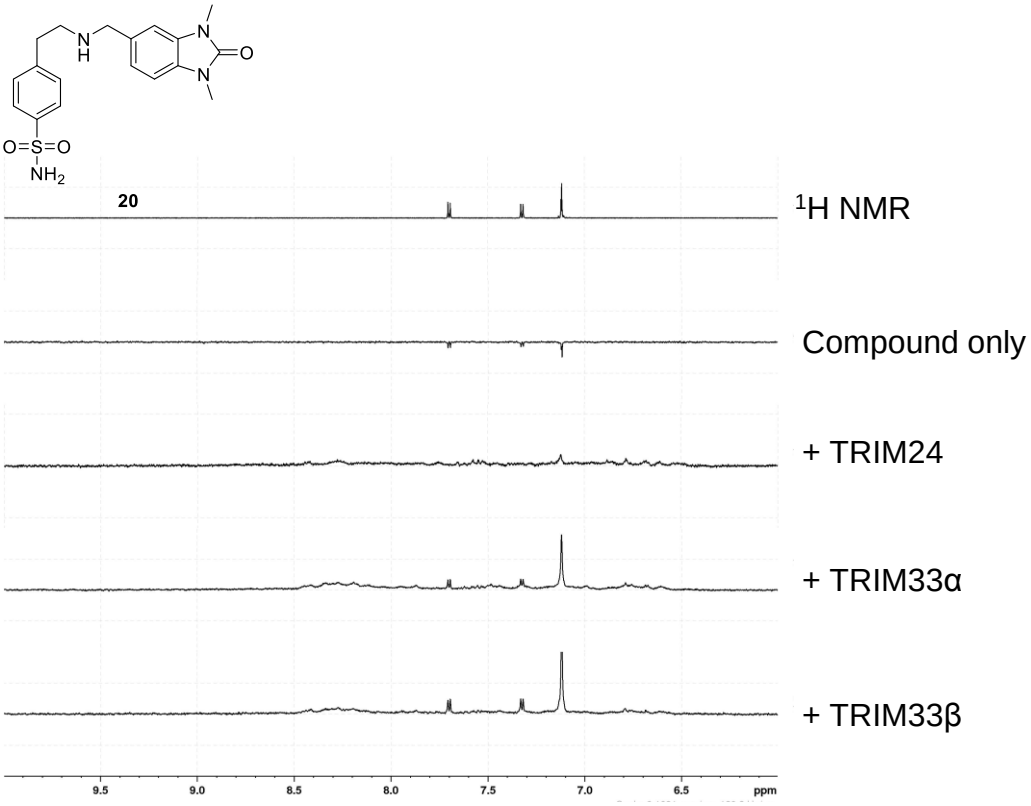
#	b	Seq	y	#
1	138.07	H		10
2	251.94	L	1110.85	9
3	352.28	T	997.47	8
4	569.21	C(+114.00)	896.98	7
5	706.22	H	679.70	6
6	805.33	V	542.34	5
7	902.33	P	443.26	4
8	1003.44	T	346.22	3
9	1116.38	L	245.13	2
10		L	132.10	1

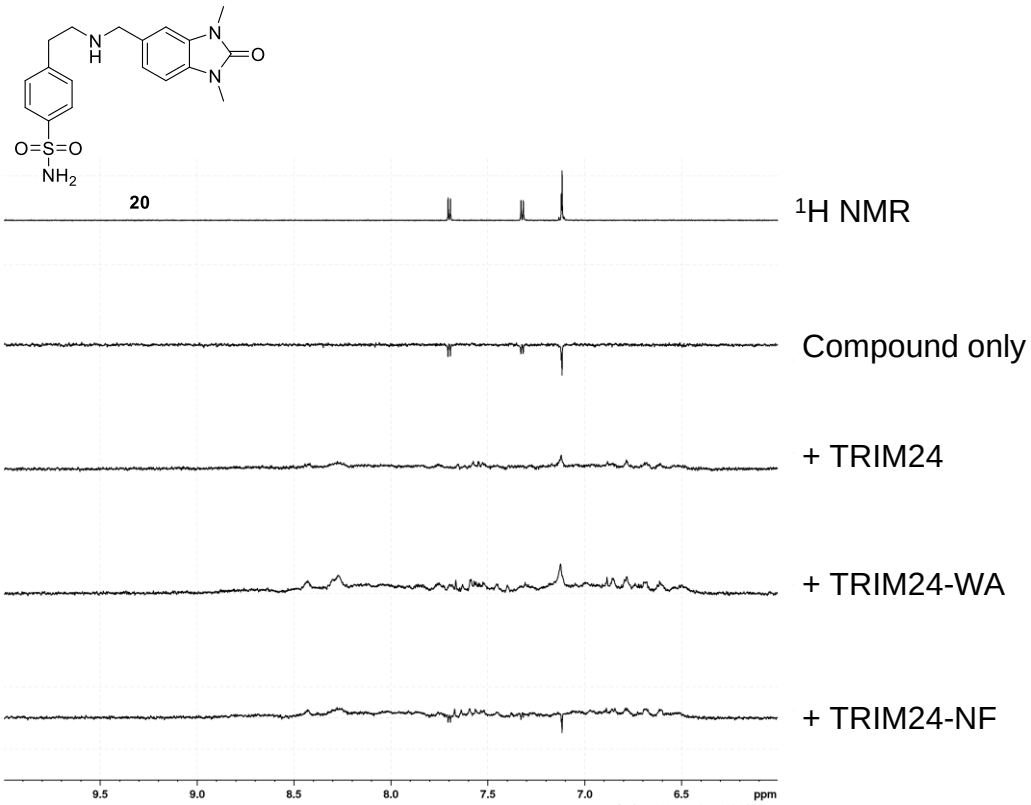


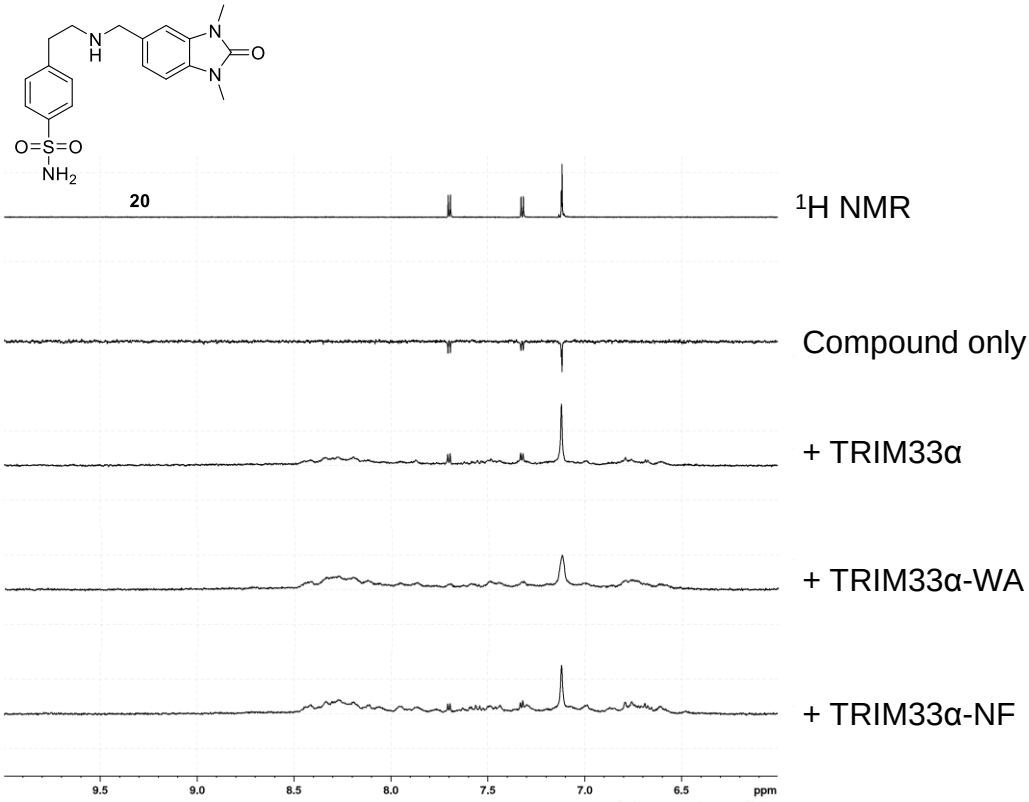
#	b	Seq	y	#
1	157.11	R		27
2	272.18	D	2988.41	26
3	385.32	I	2873.38	25
4	442.33	G	2760.30	24
5	570.29	K	2703.28	23
6	667.31	P	2575.18	22
7	796.42	E	2478.13	21
8	895.50	V	2349.09	20
9	1024.54	E	2250.02	19
10	1187.61	Y	2120.98	18
11	1302.88	D	1957.91	17
12	1520.94	C(+115.00)	1842.89	16
13	1635.73	D	1624.88	15
14	1750.10	N	1509.82	14
15	1862.80	L	1396.04	13
16	1990.85	Q	1282.88	12
17	2127.91	H	1154.80	11
18	2214.95	S	1017.64	10
19	2343.04	K	930.61	9
20	2471.24	K	802.59	8
21	2528.16	G	674.38	7

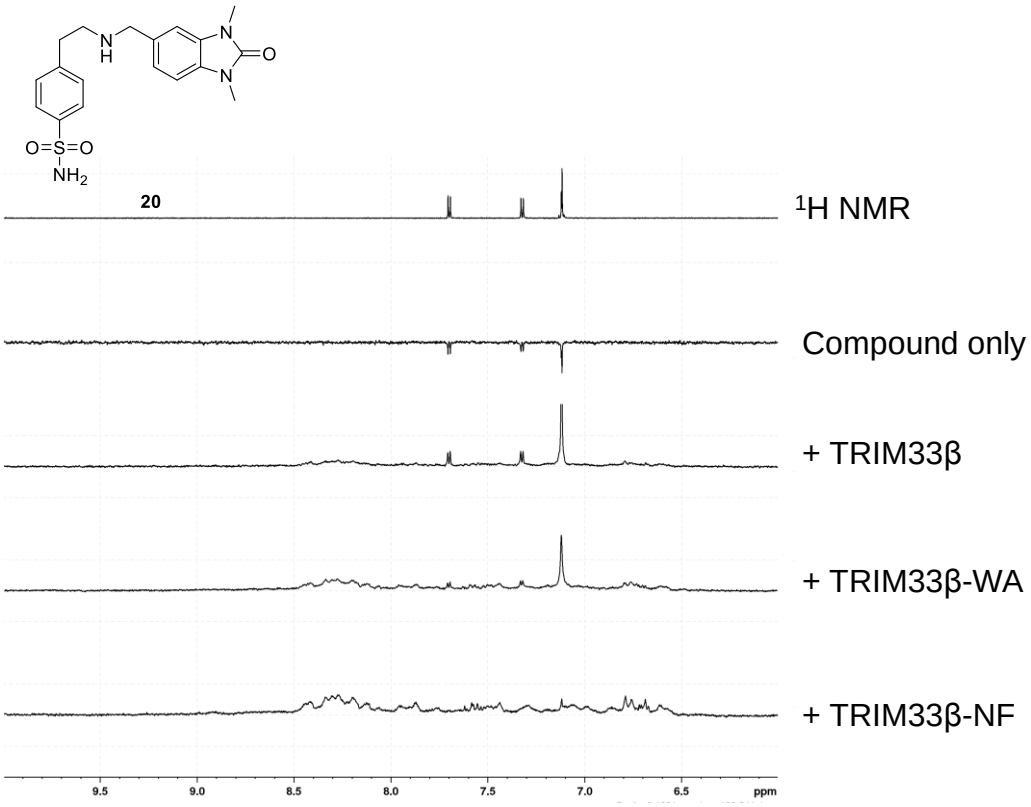
E. NMR WaterLOGSY Spectra

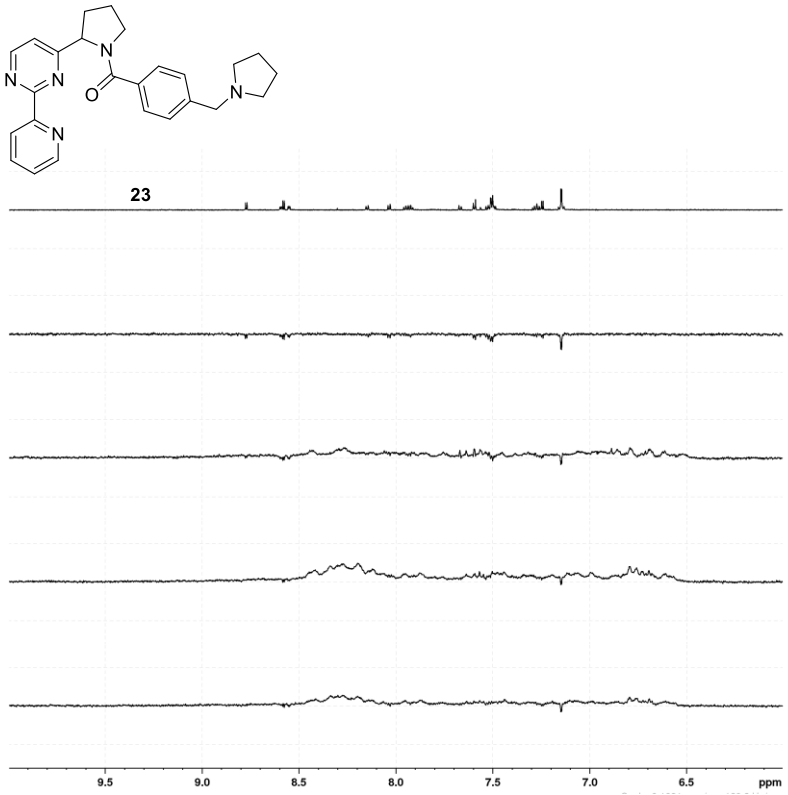
Water-suppressed ^{1}H NMR and waterLOGSY spectra with and without protein are reported. Spectra are presented in compound numerical order.

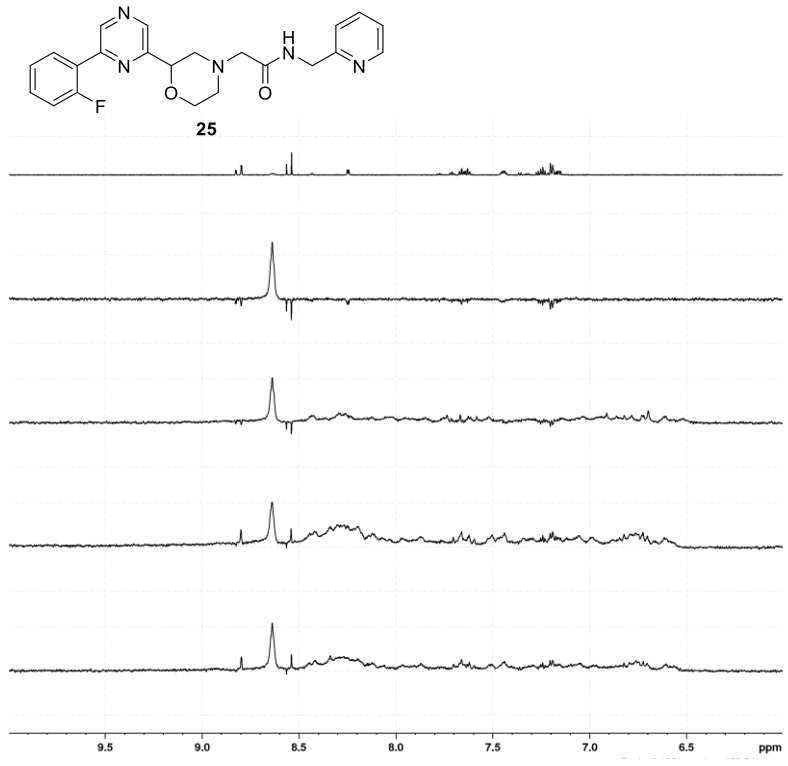

20
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

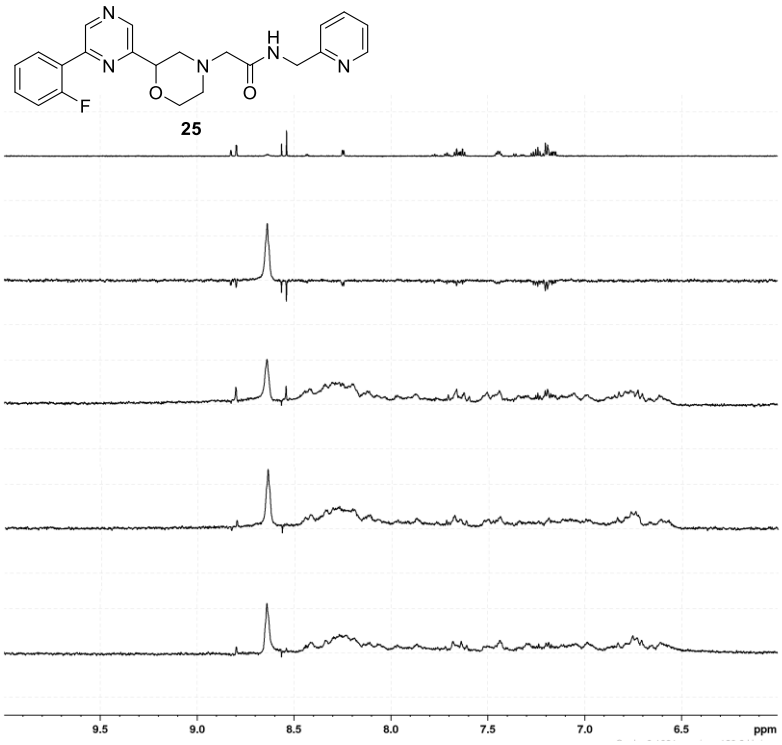

20
^{1}H NMR
Compound only
+ TRIM24
+ TRIM24-WA
+ TRIM24-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

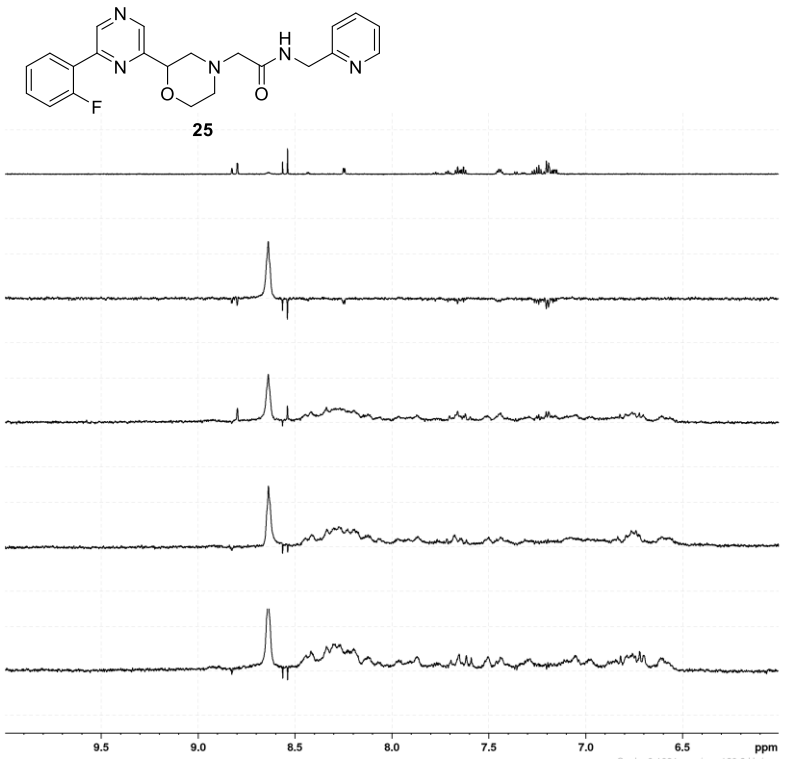

20
^{1}H NMR
Compound only
+ TRIM33α
+ TRIM33α-WA
+ TRIM33α-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

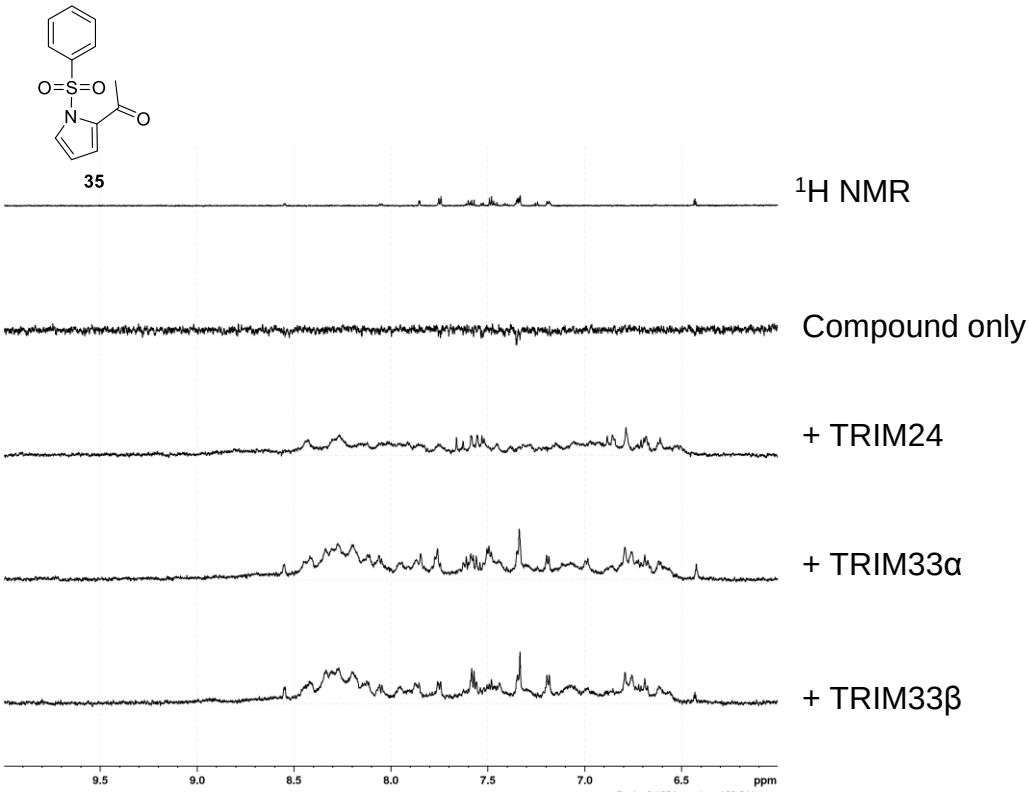

20
^{1}H NMR
Compound only
+ TRIM33β
+ TRIM33β-WA
+ TRIM33β-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

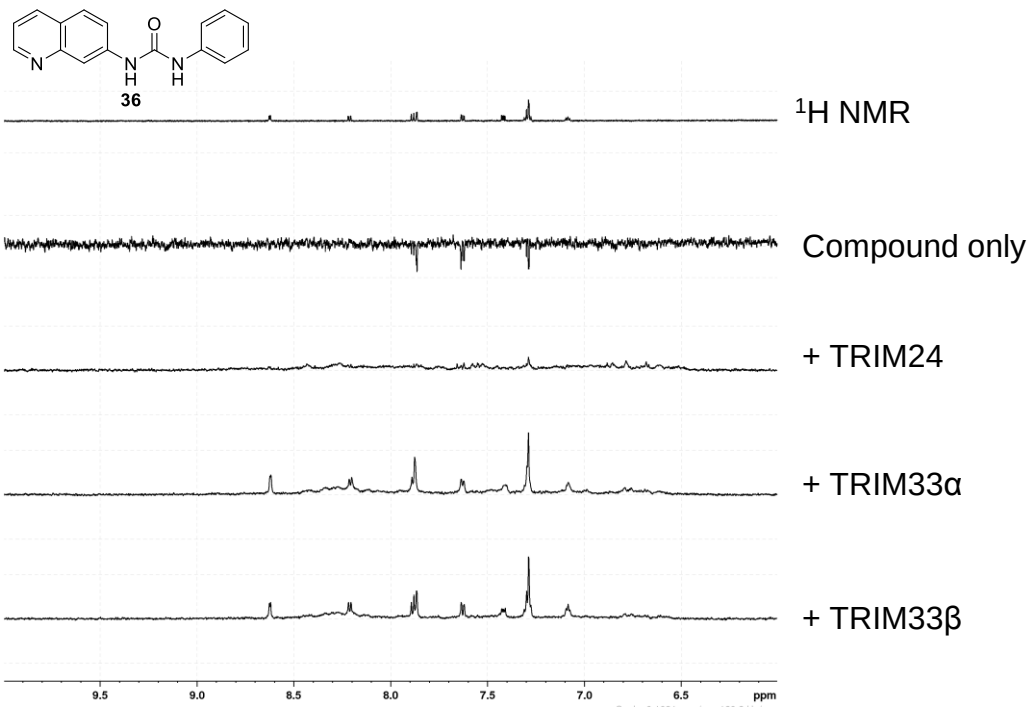

23
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

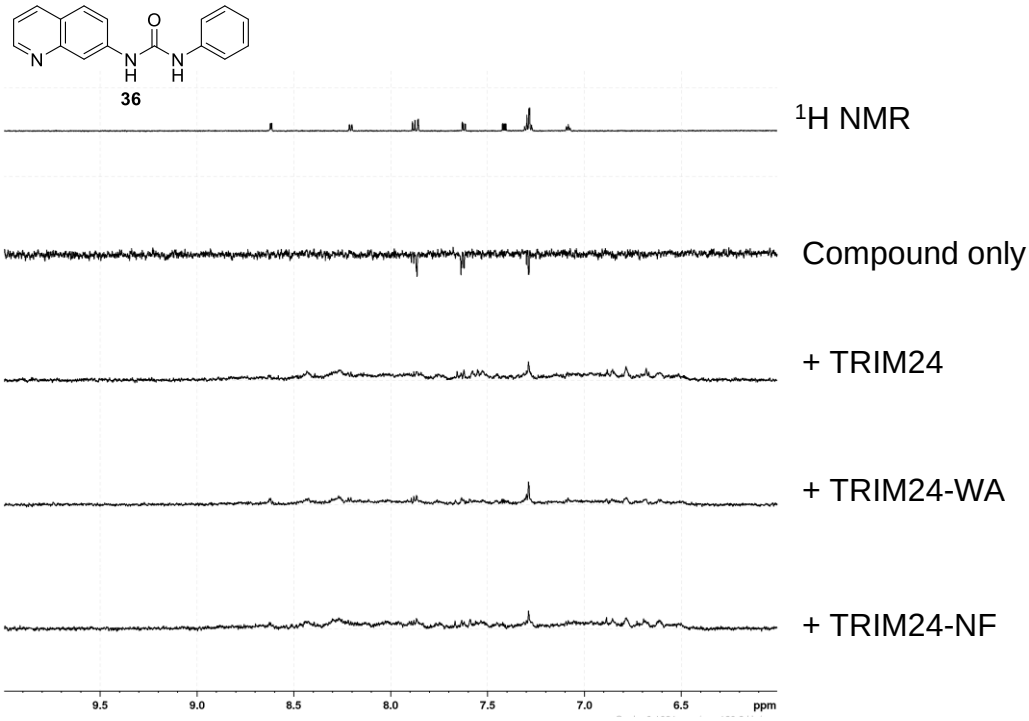

25
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

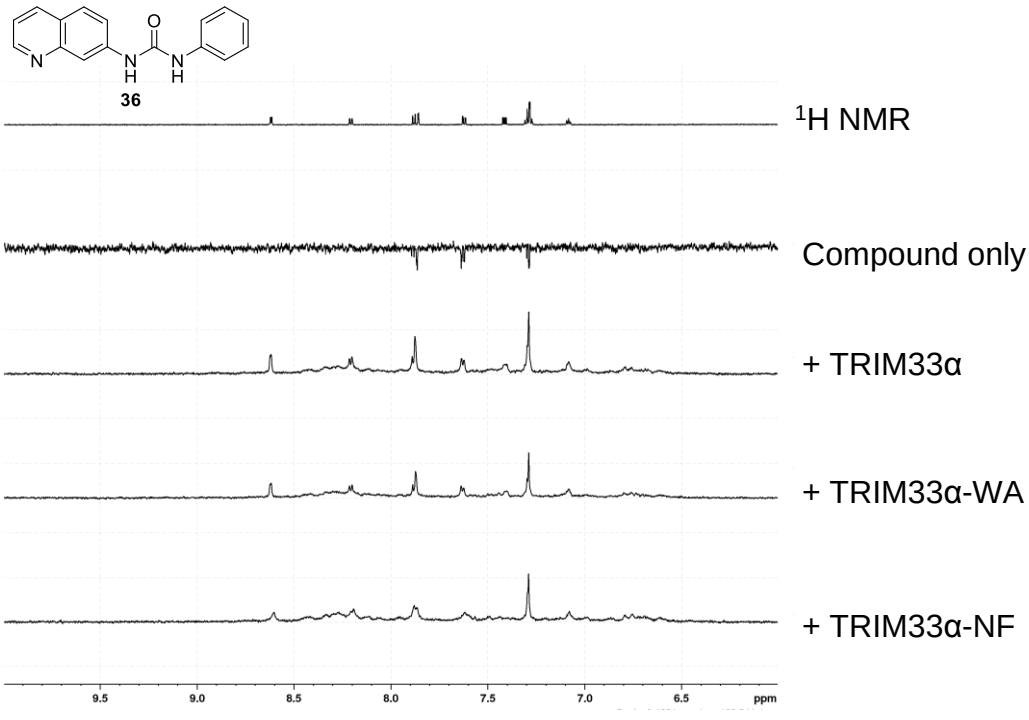

25
^{1}H NMR
Compound only
+ TRIM33α
+ TRIM33α-WA
+ TRIM33α-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

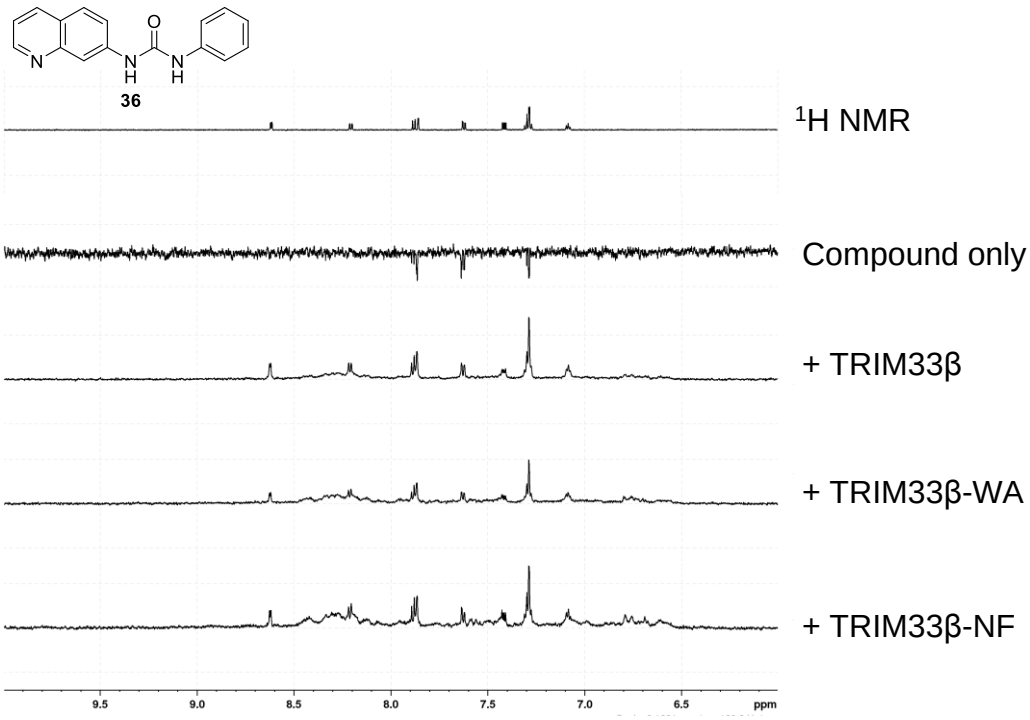

25
^{1}H NMR
Compound only
+ TRIM33β
+ TRIM33β-WA
+ TRIM33β-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

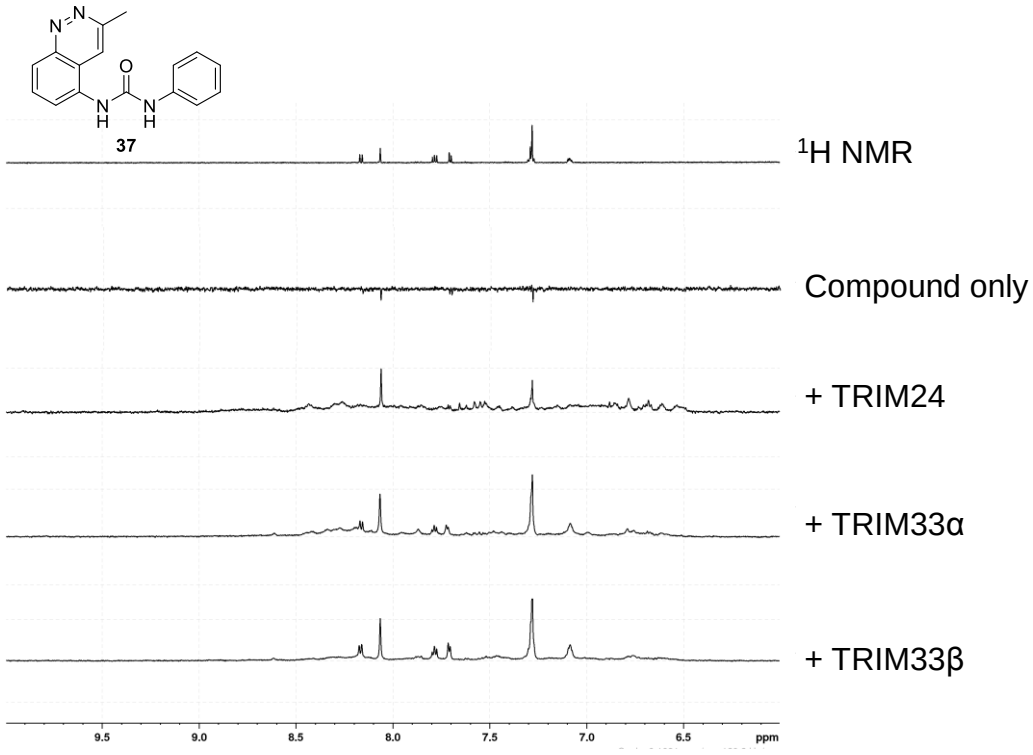

35
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

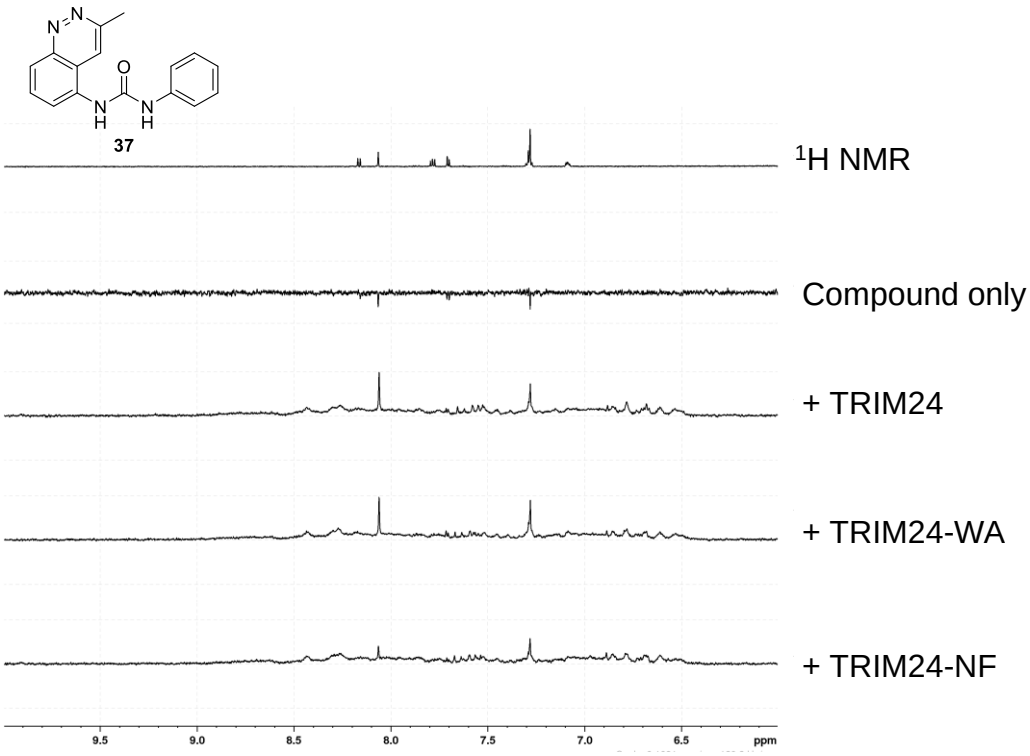

36
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

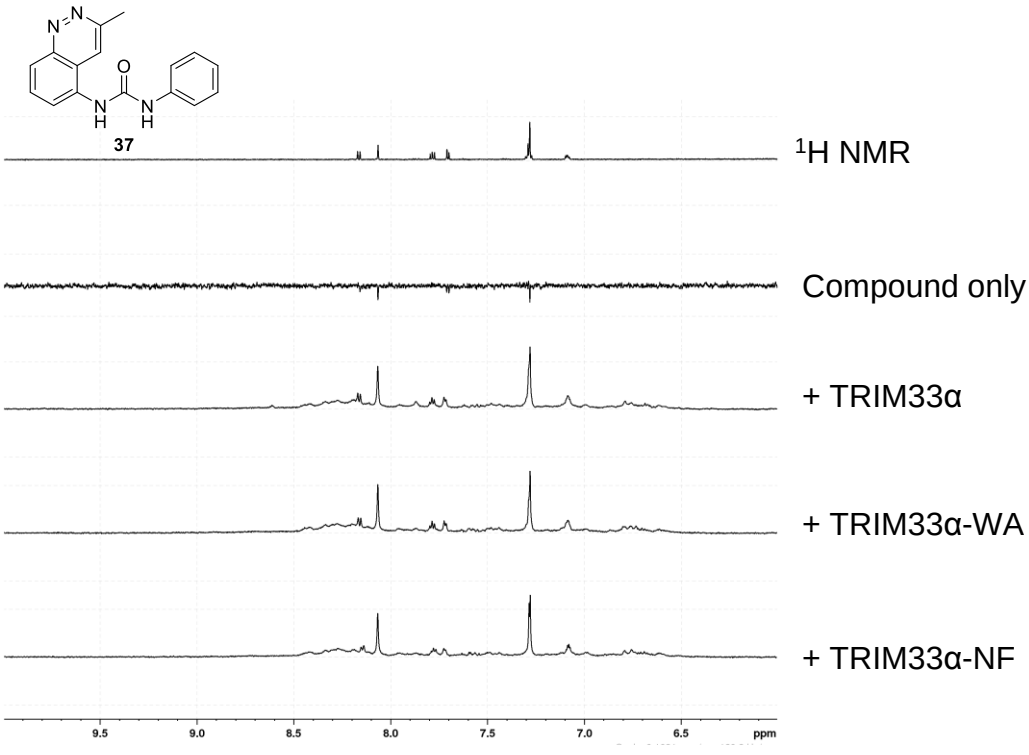

36
^{1}H NMR
Compound only
+ TRIM24
+ TRIM24-WA
+ TRIM24-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

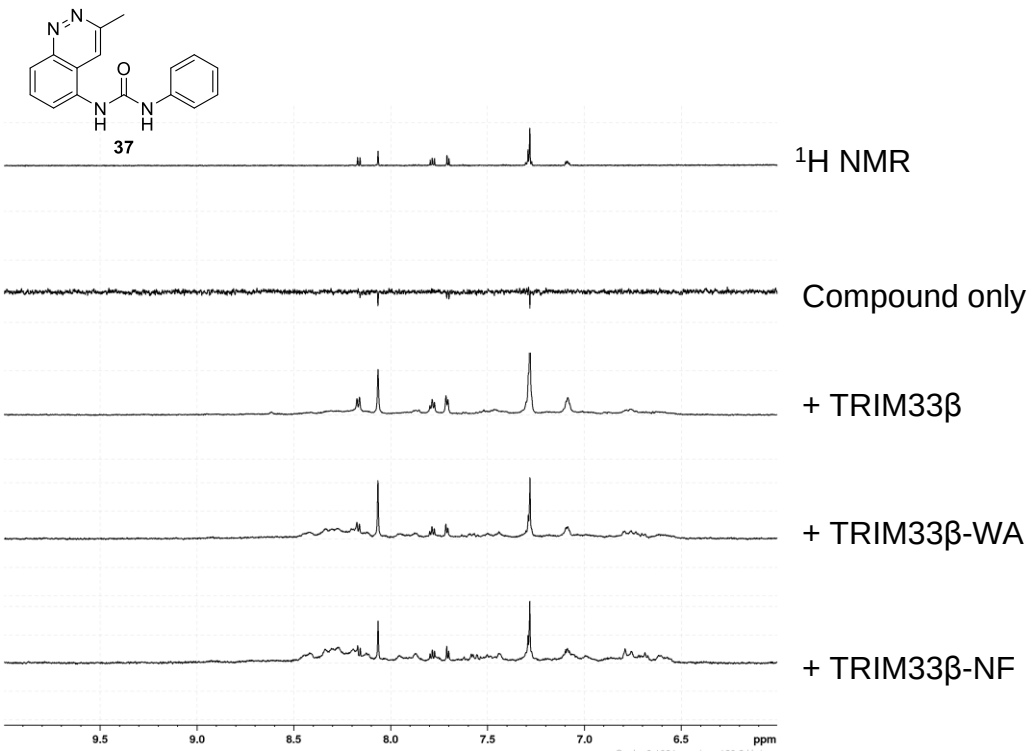

36
^{1}H NMR
Compound only
+ TRIM33α
+ TRIM33α-WA
+ TRIM33α-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

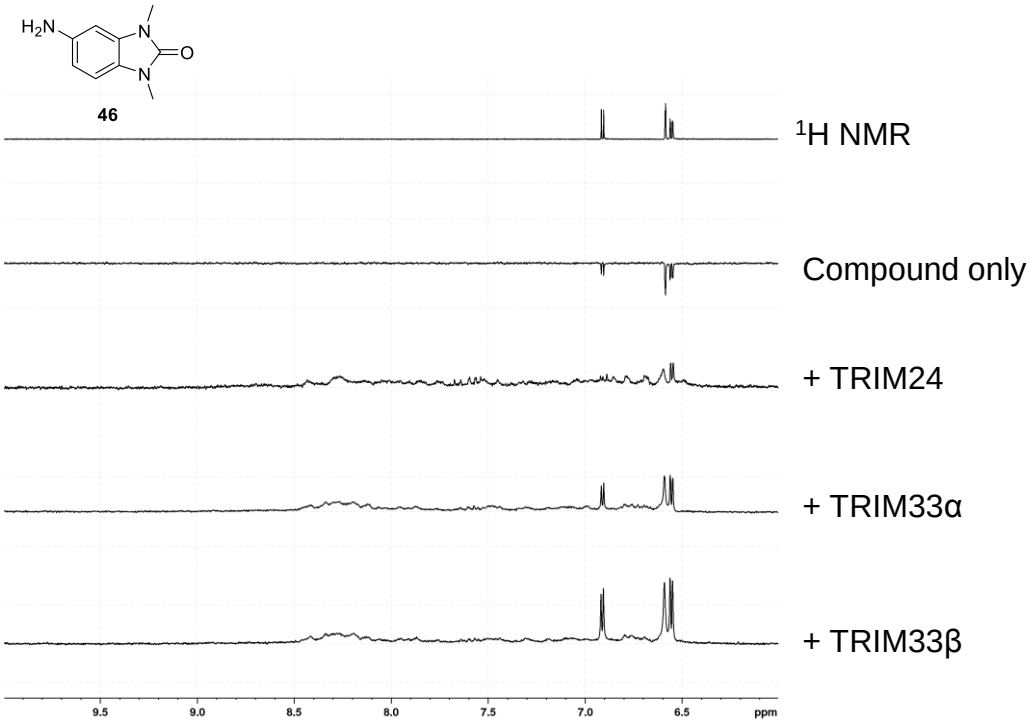

36
^{1}H NMR
Compound only
+ TRIM33β
+ TRIM33β-WA
+ TRIM33β-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

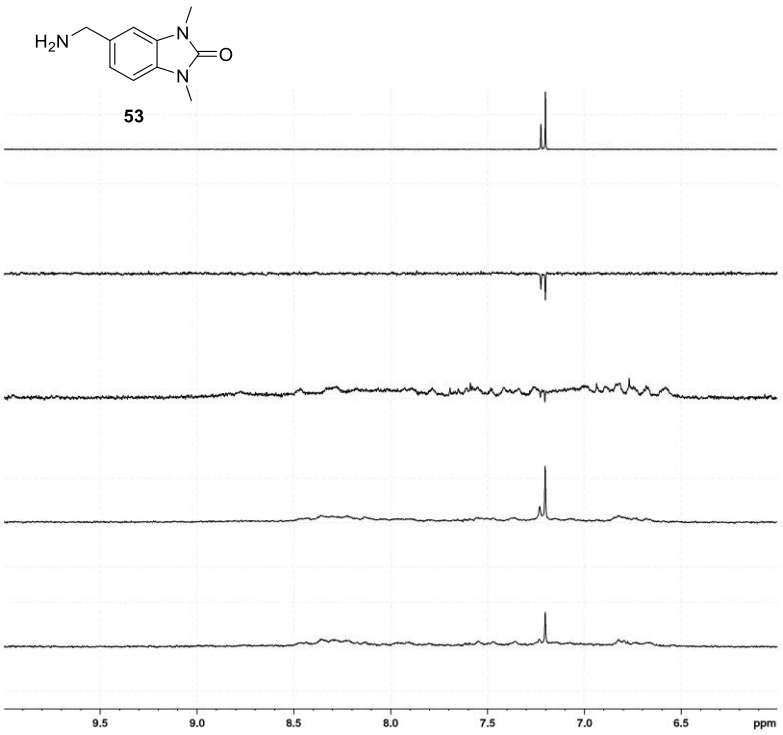

37
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

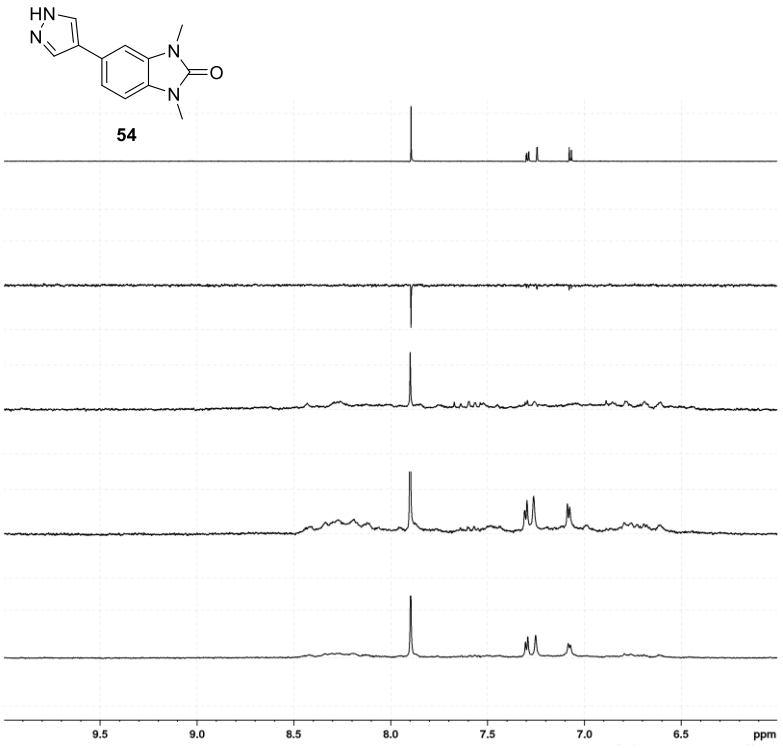

37
^{1}H NMR
Compound only
+ TRIM24
+ TRIM24-WA
+ TRIM24-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

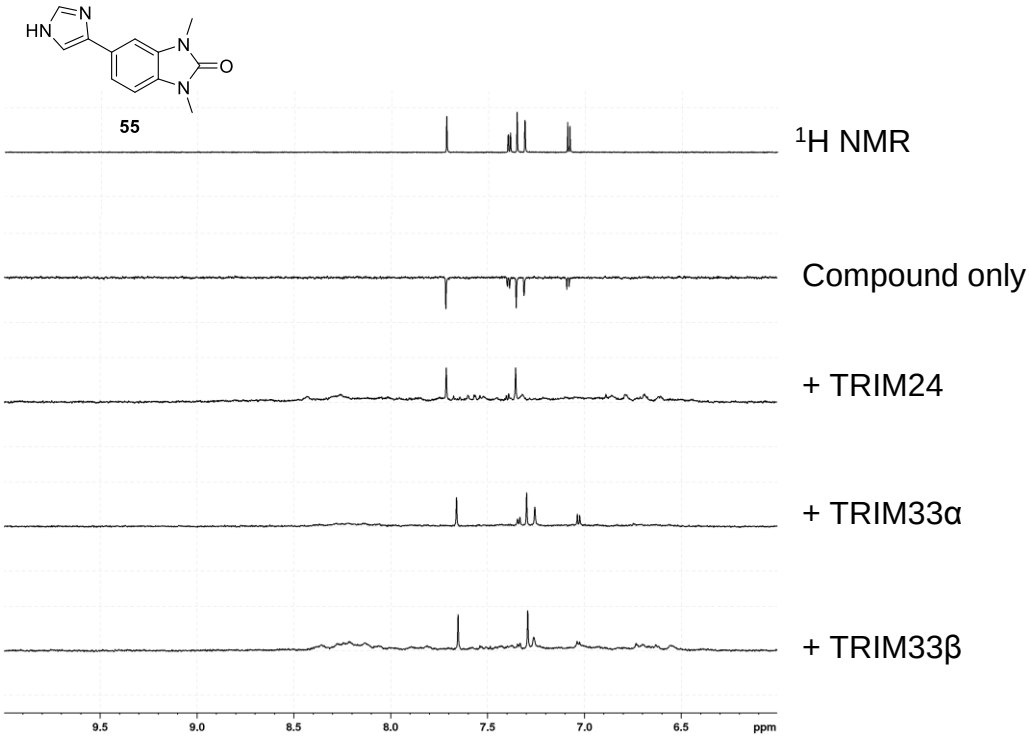

37
^{1}H NMR
Compound only
+ TRIM33α
+ TRIM33α-WA
+ TRIM33α-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

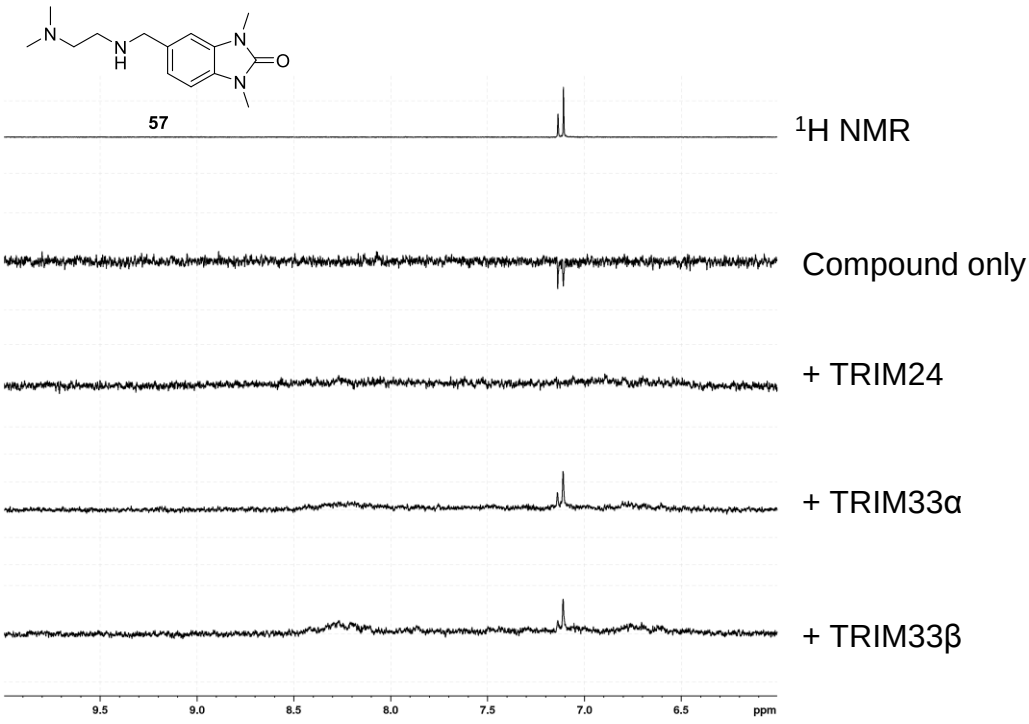

37
^{1}H NMR
Compound only
+ TRIM33β
+ TRIM33β-WA
+ TRIM33β-NF
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

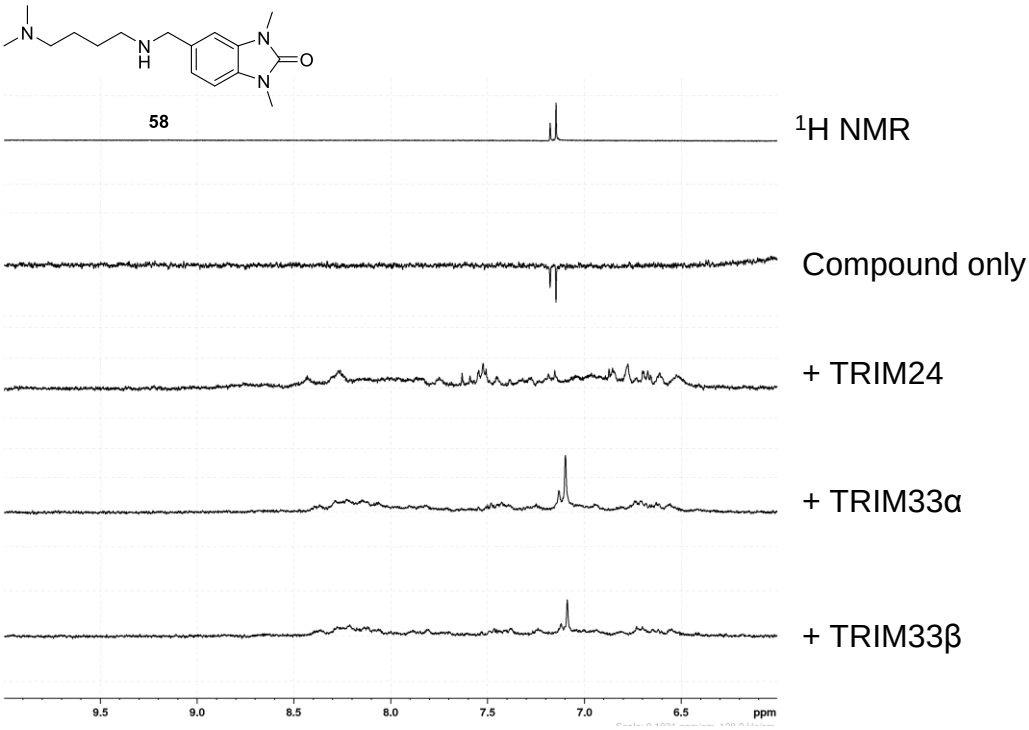

46
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

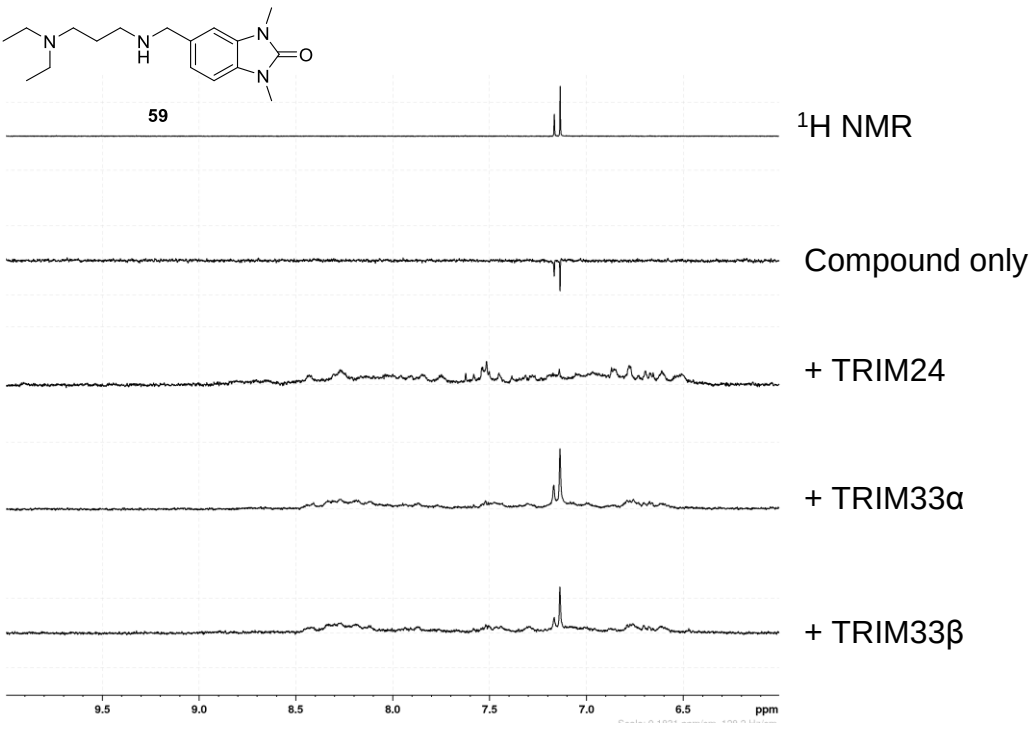

H2N
N
O
N
53
1H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

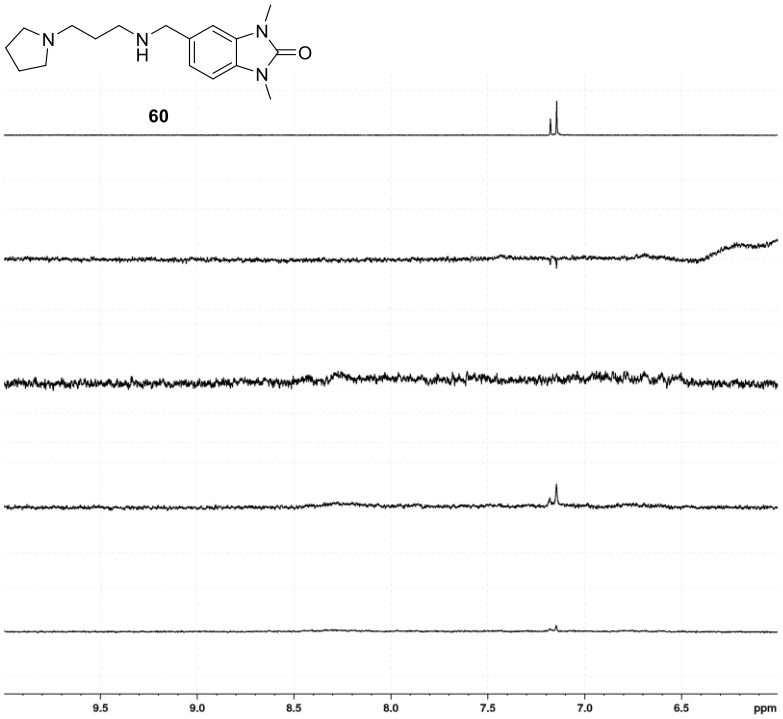

HN
N
N
O
N
54
1H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

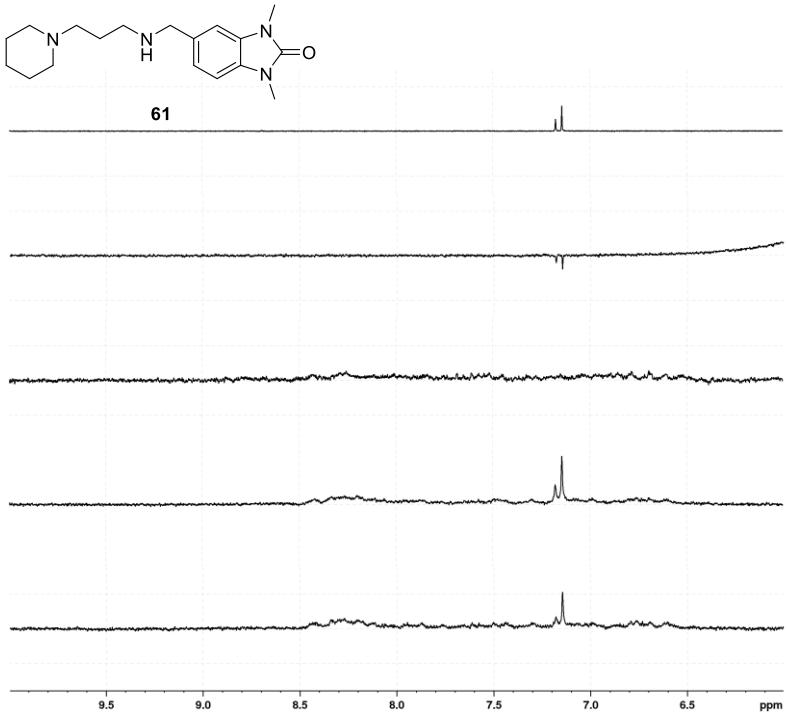

HN
N
N
O
N
55
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

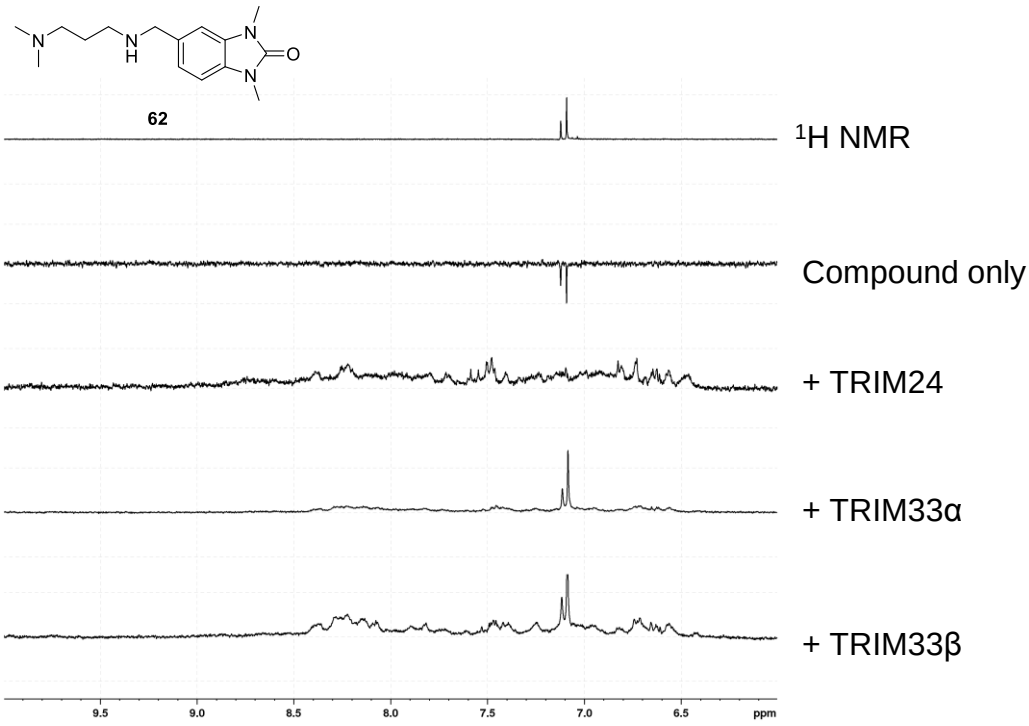

N
N
H
N
N
O
57
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

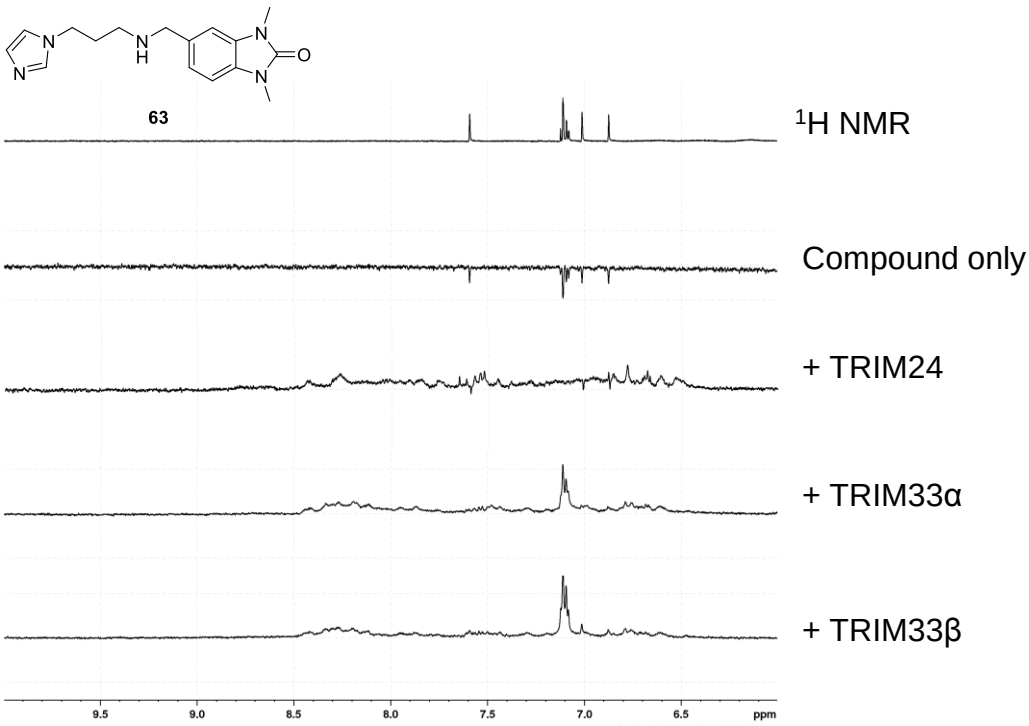

58
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

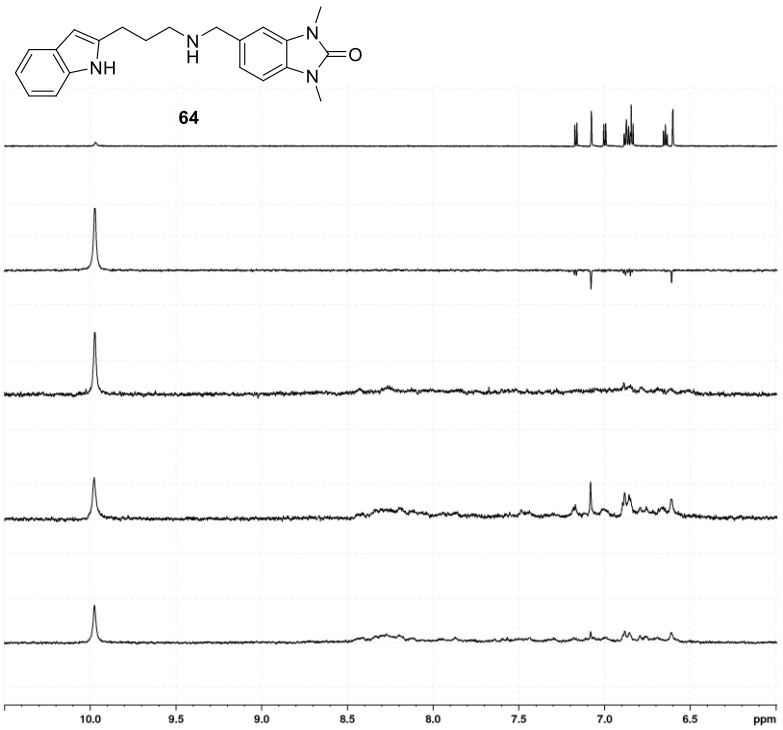

59
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

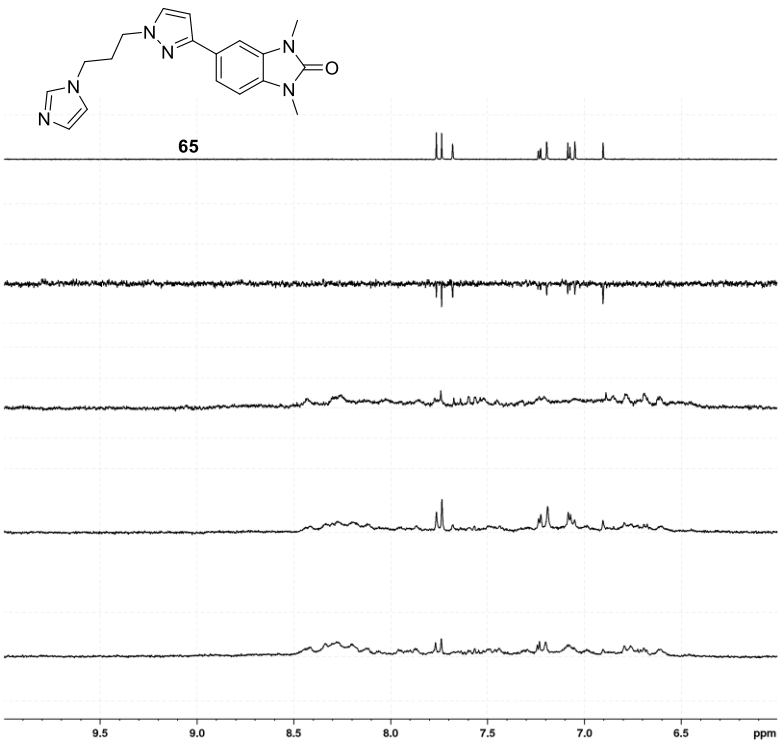

60
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

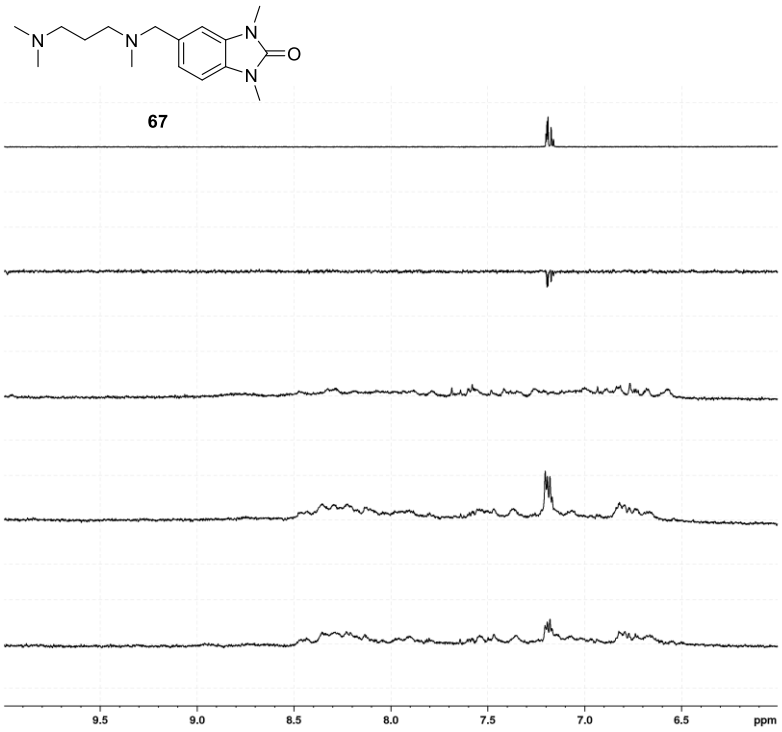

61
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

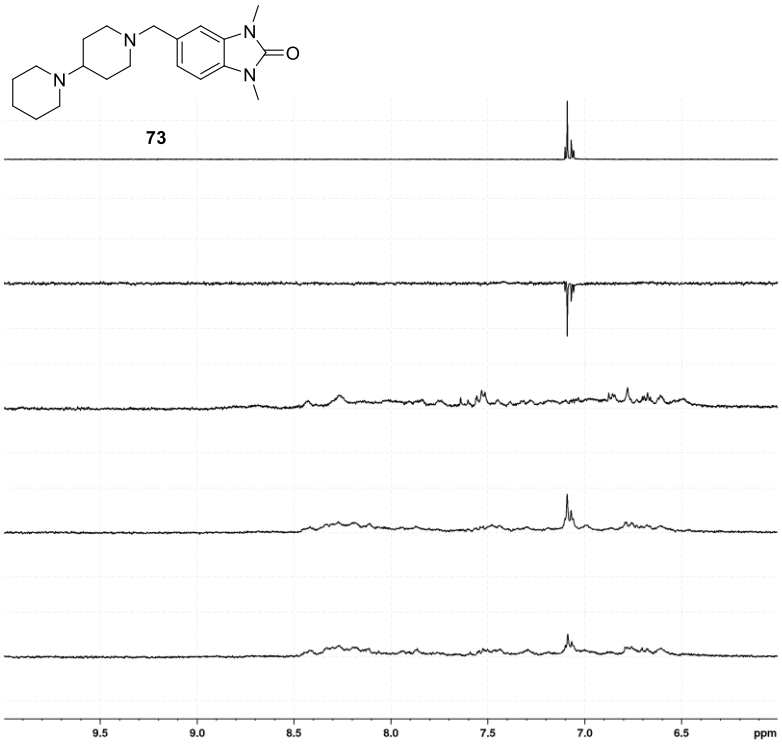

62
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

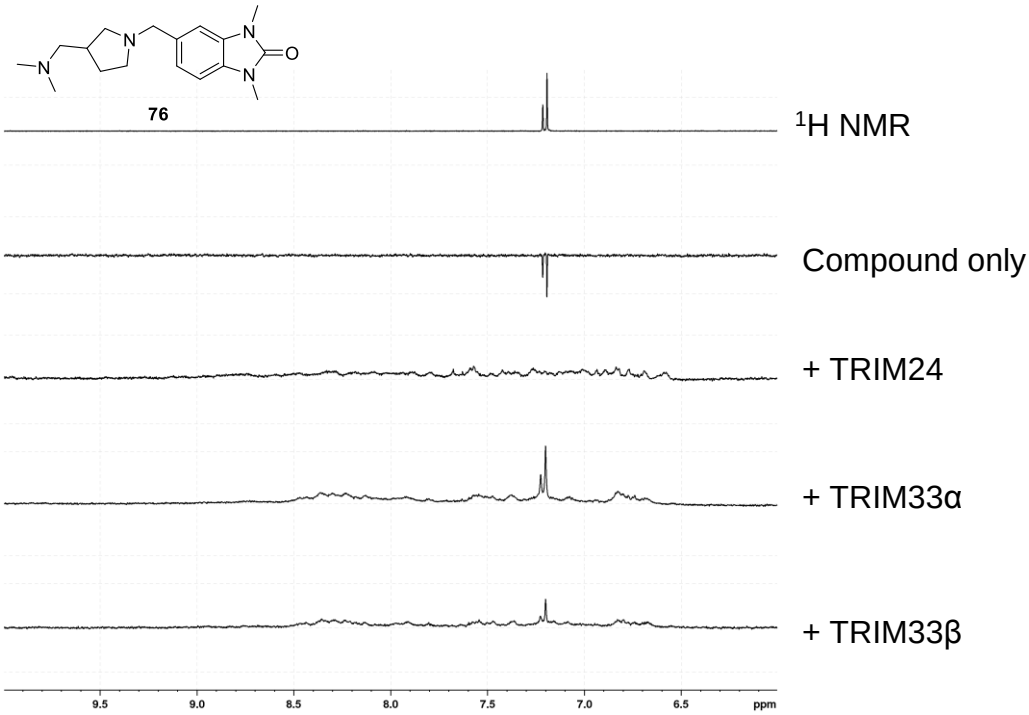

63
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

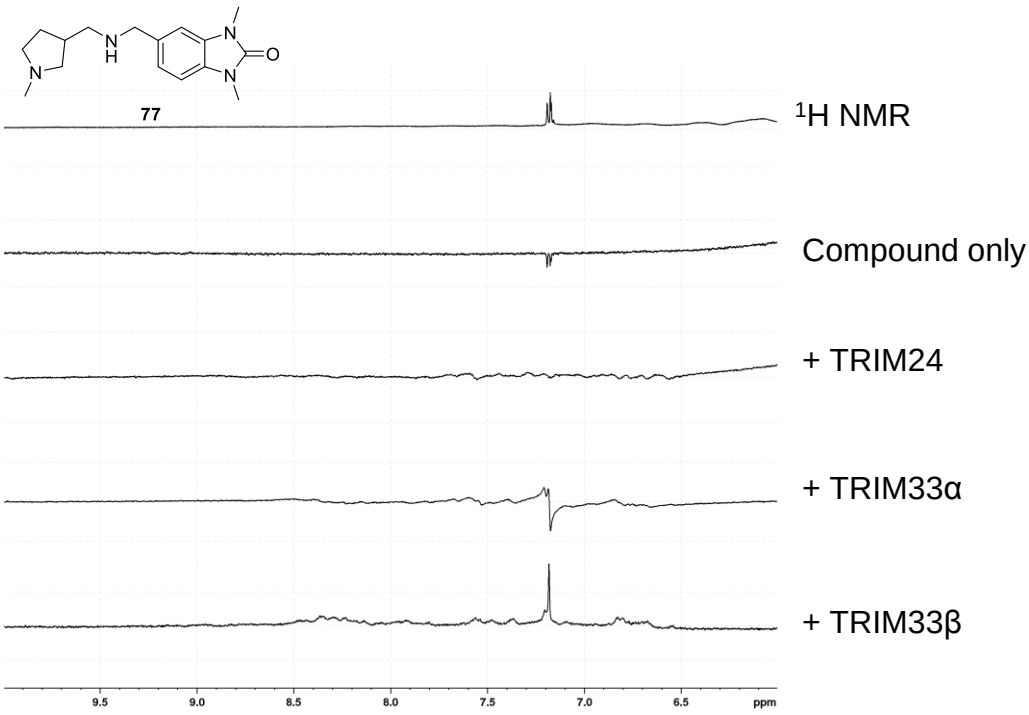

64
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

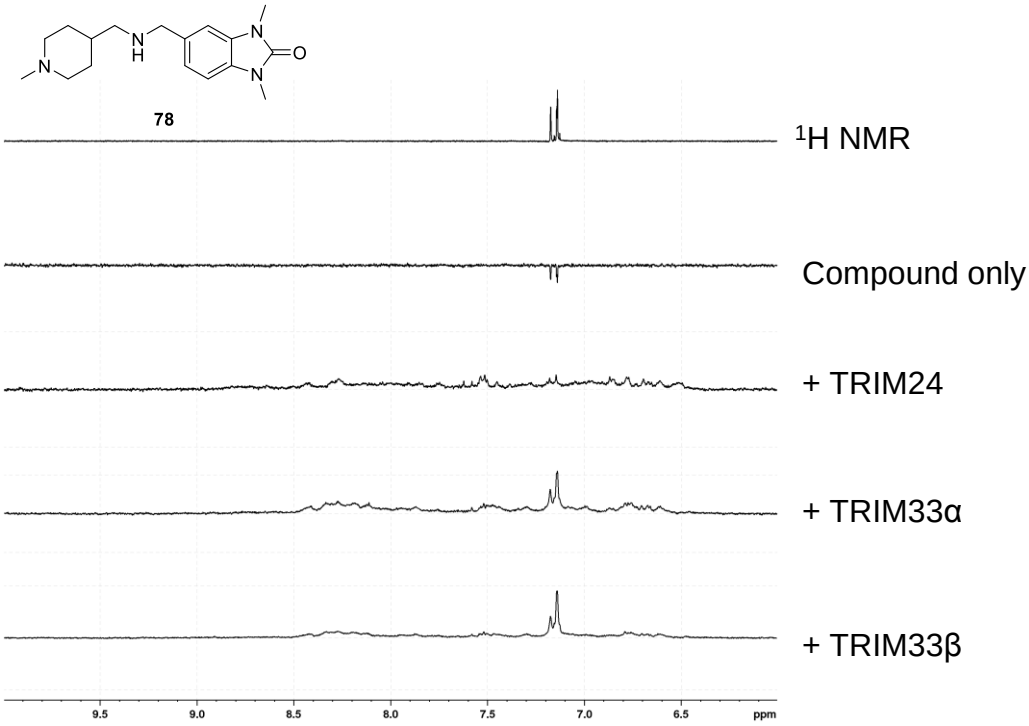

65
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

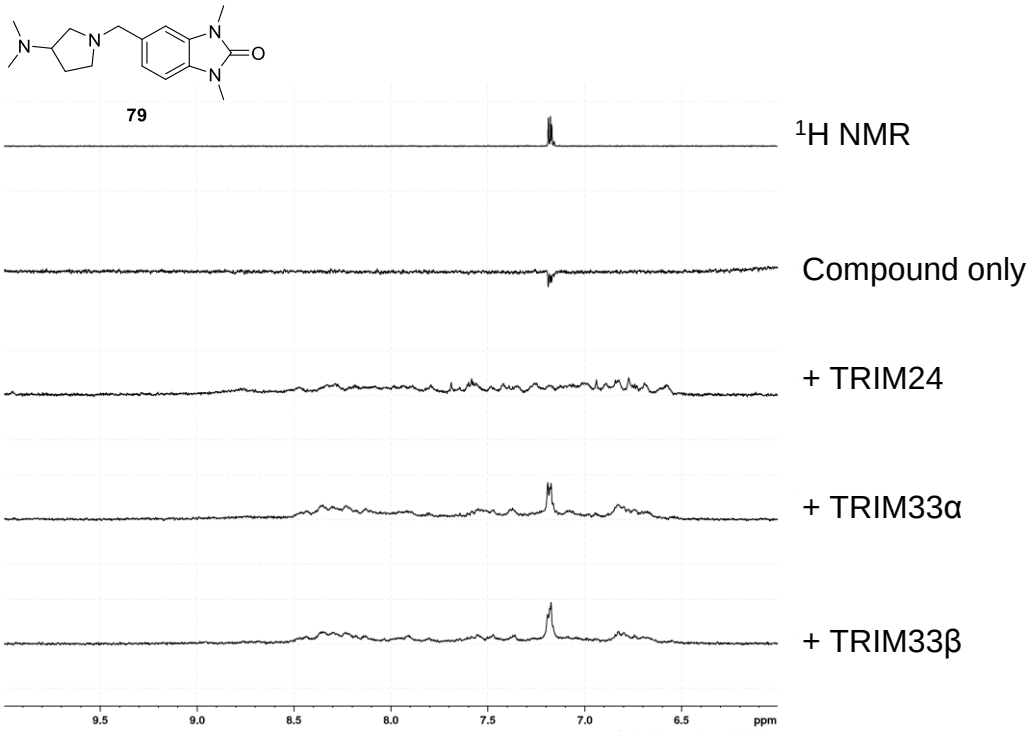

67
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

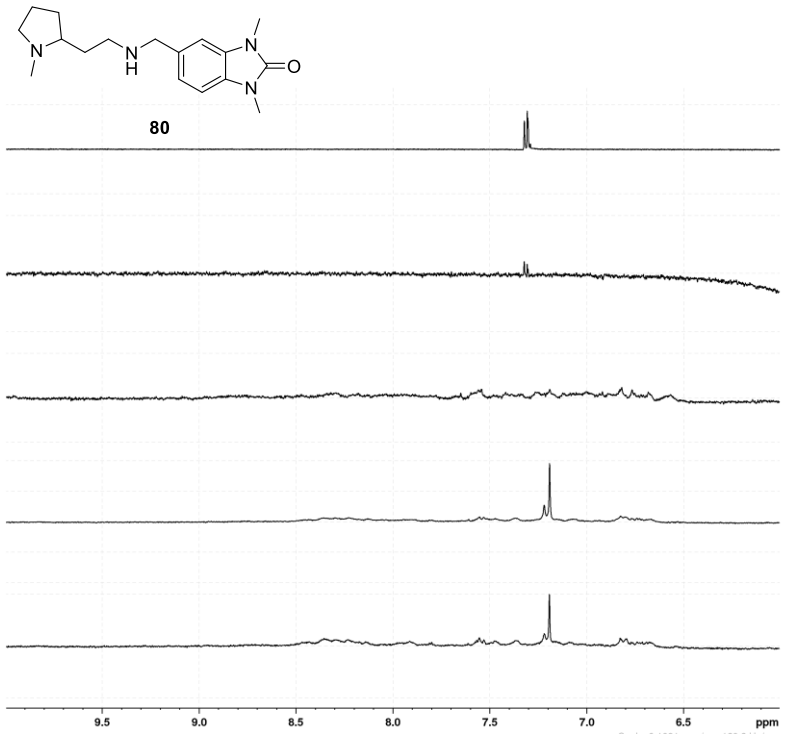

73
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

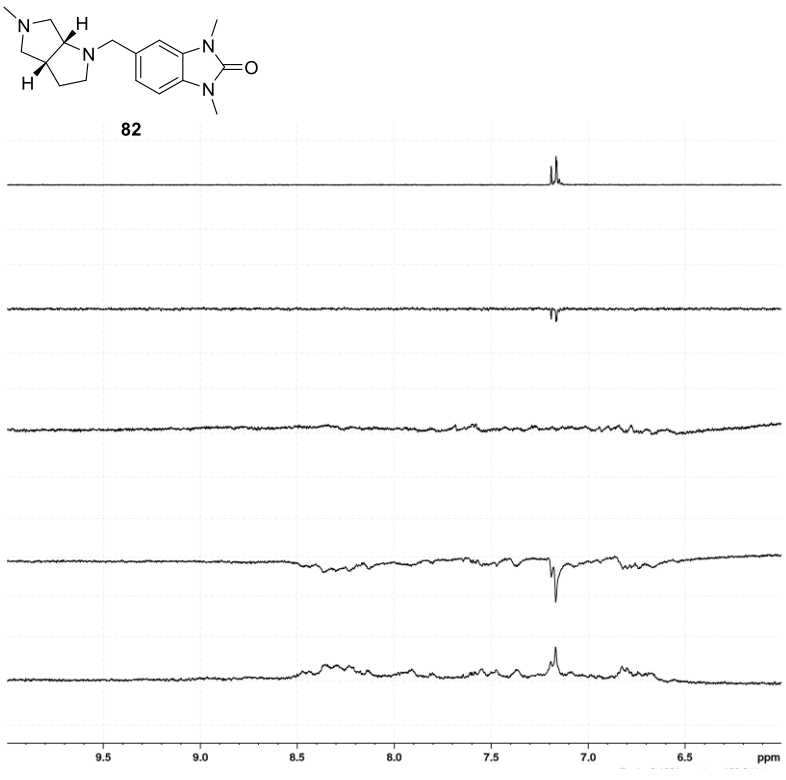

76
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

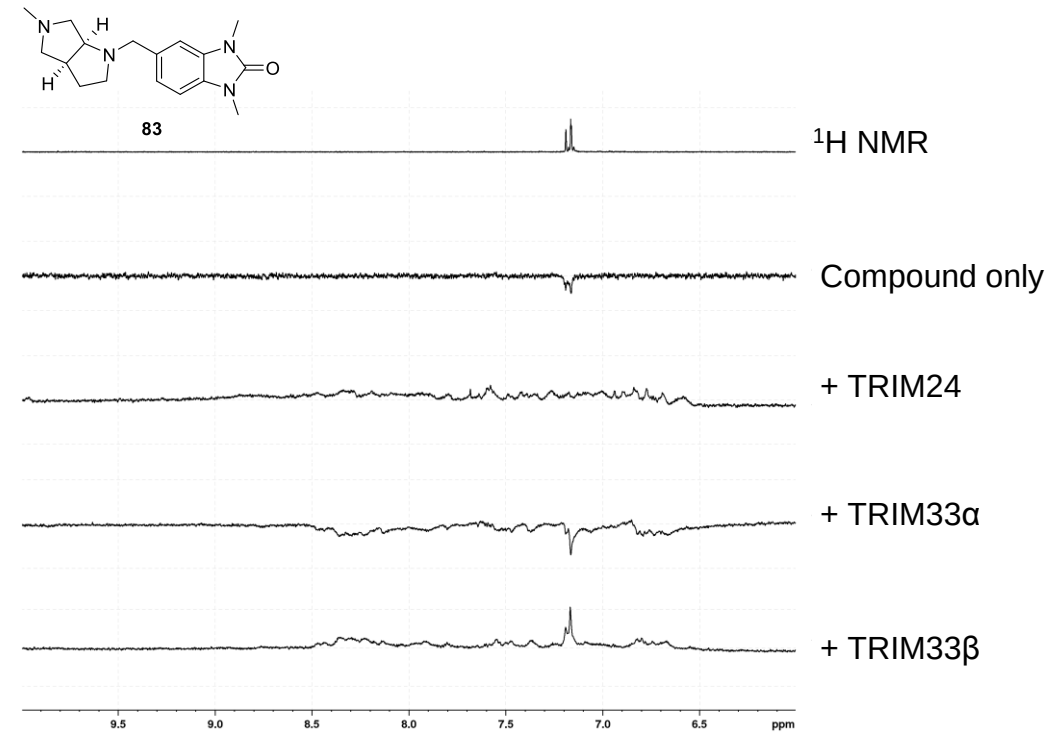

77
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

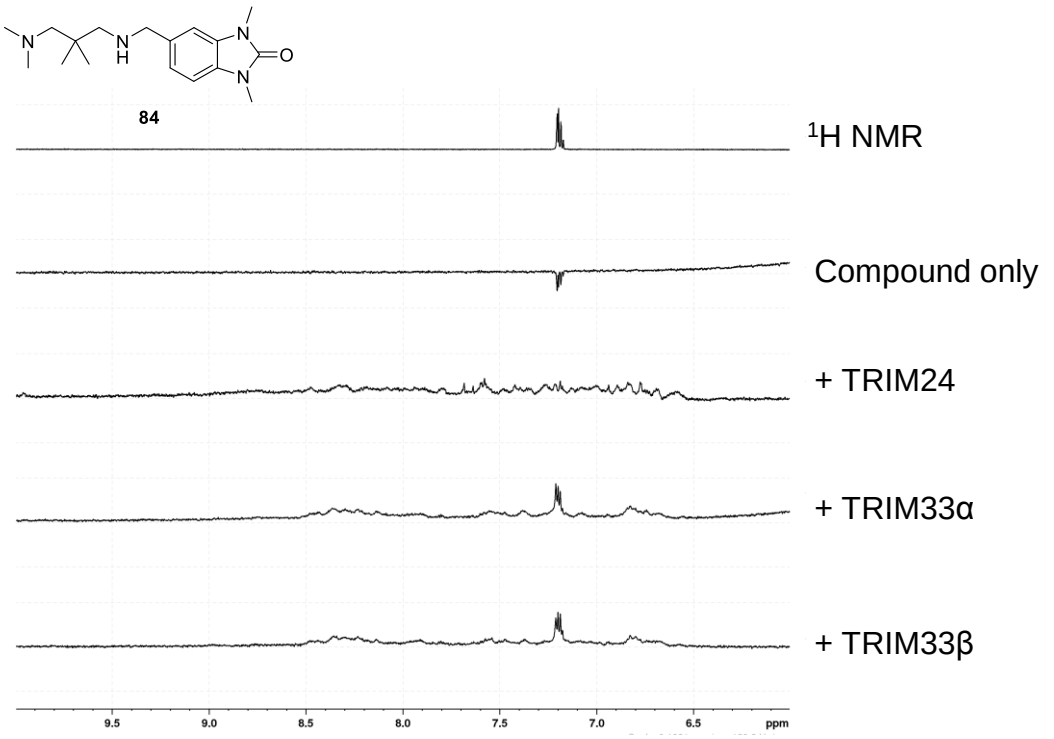

78
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

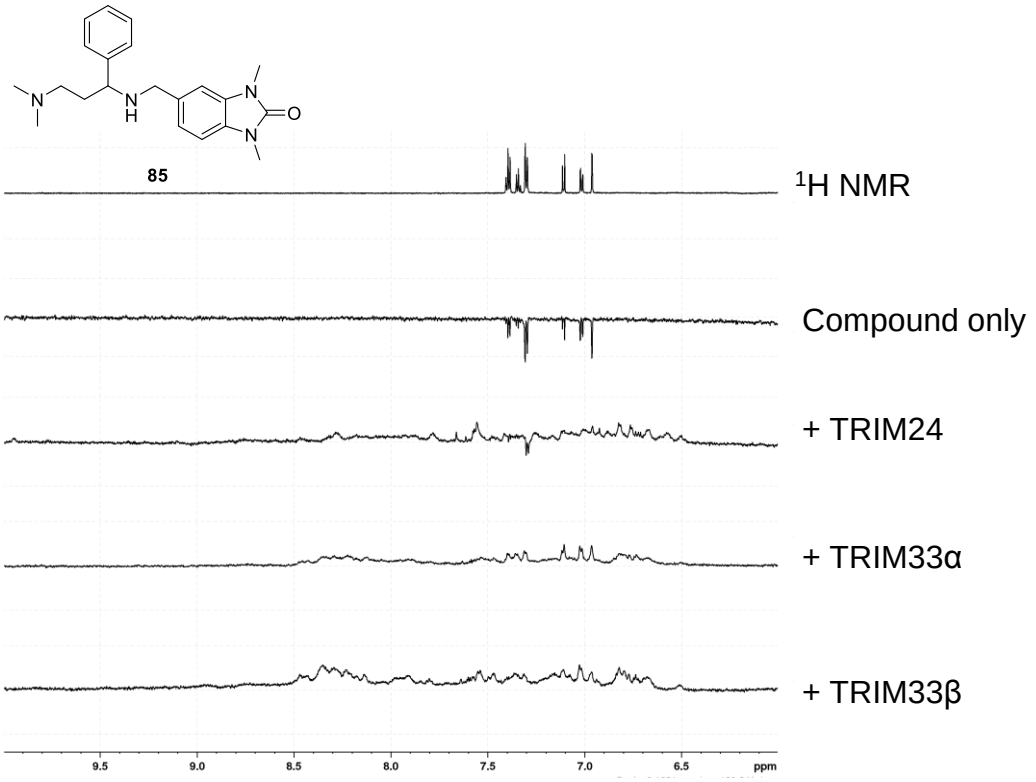

79
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

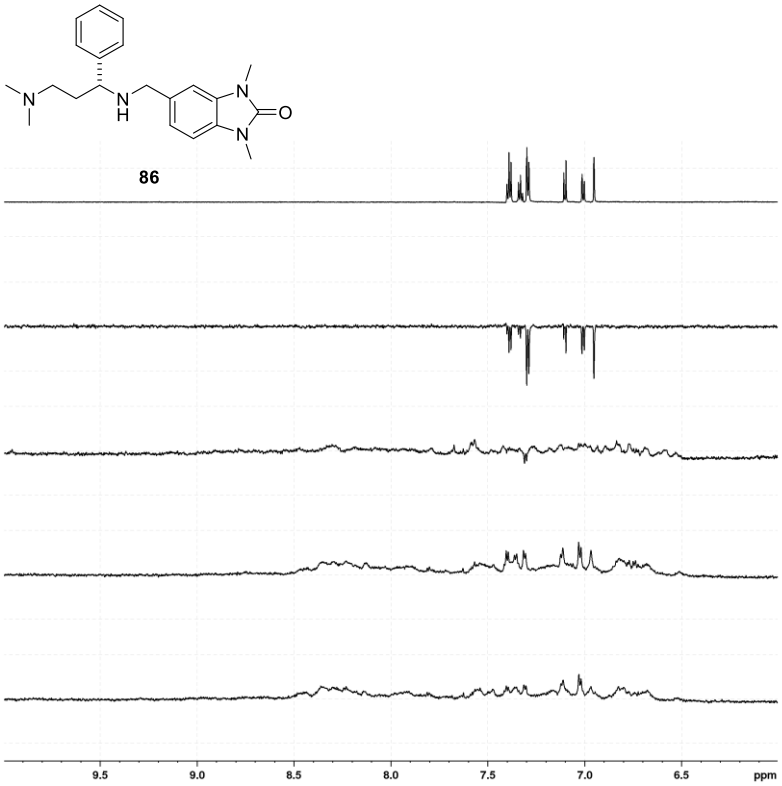

80
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

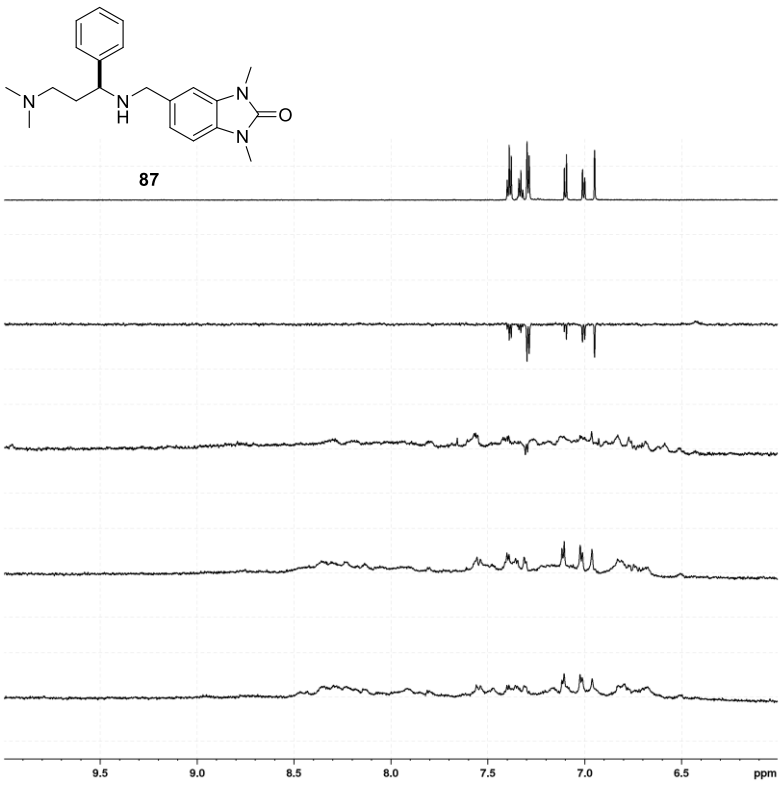

82
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

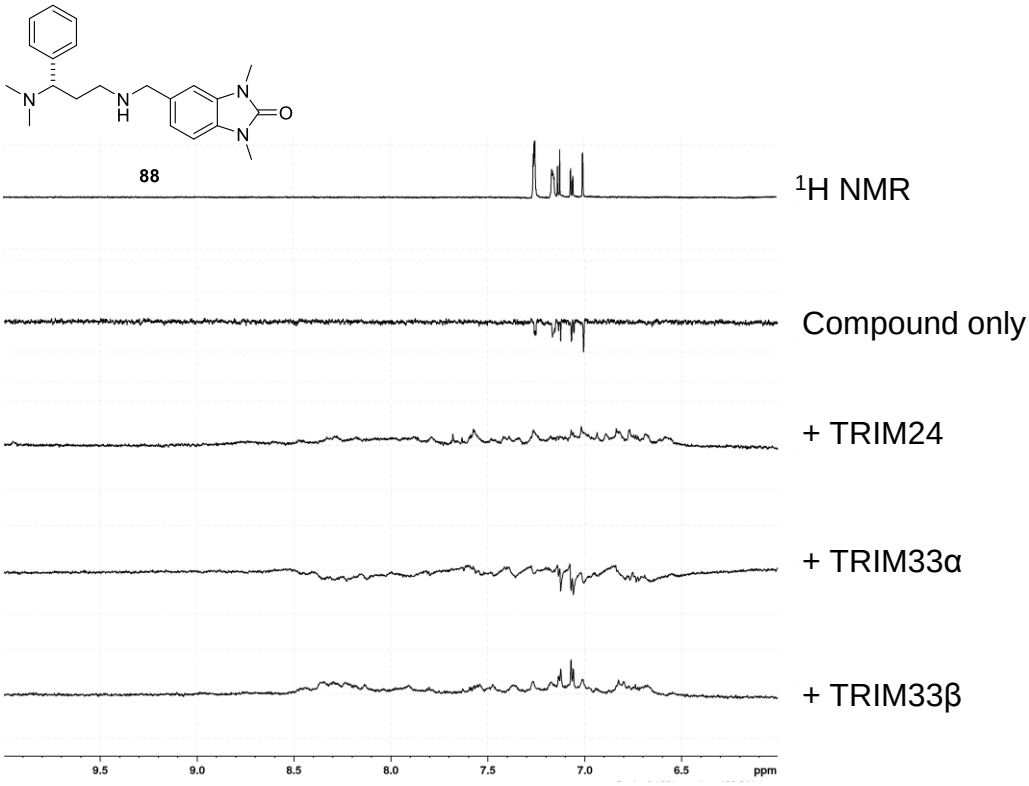

83
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

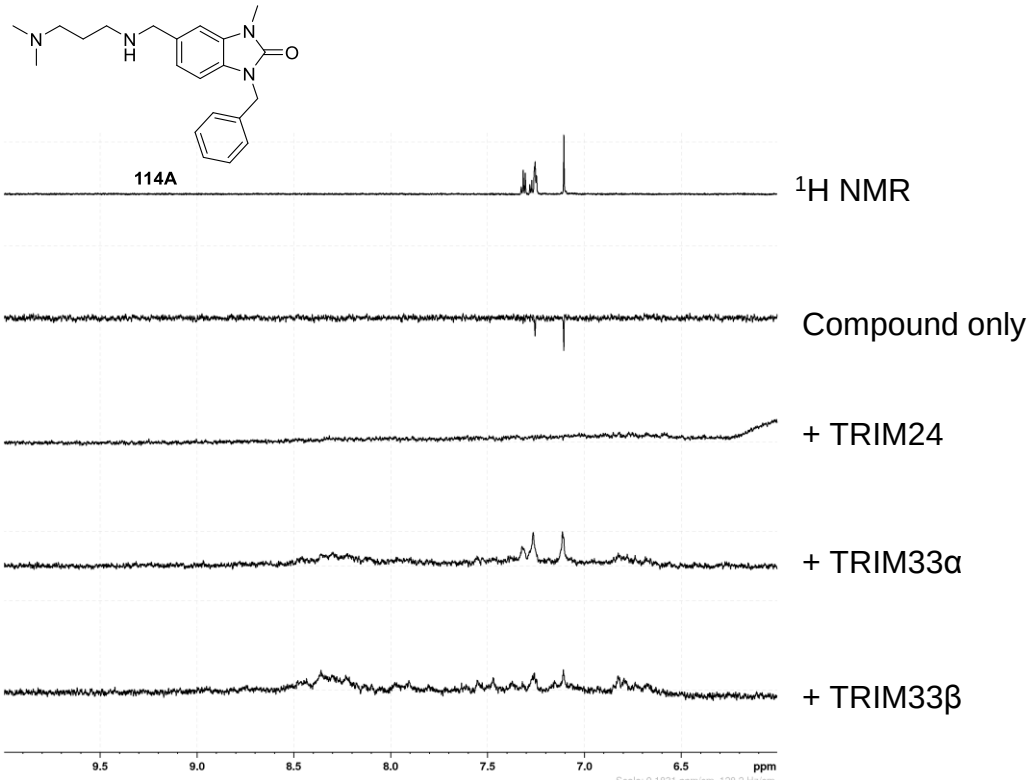

84
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

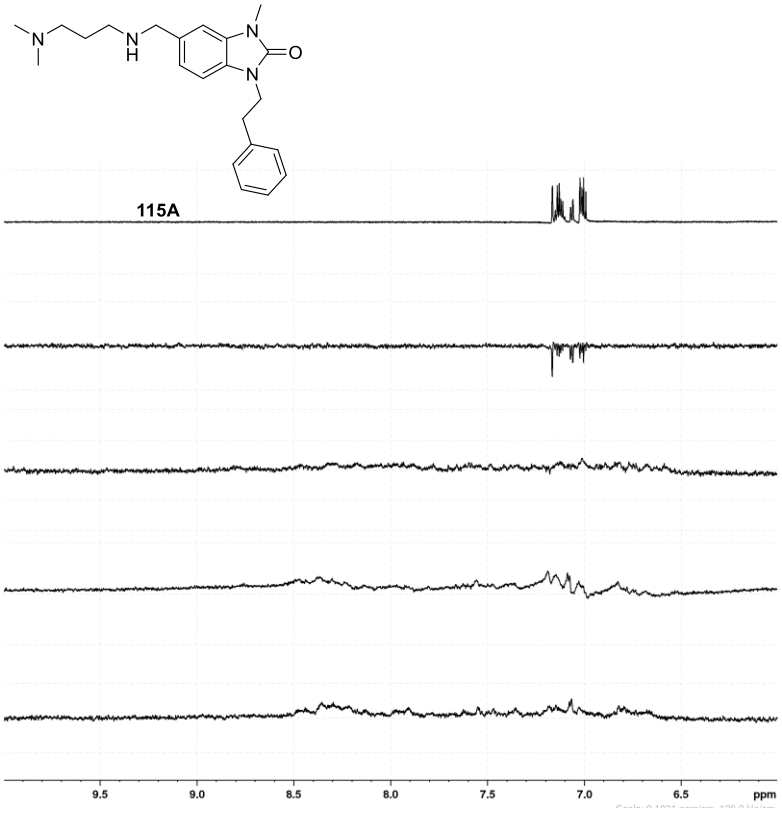

85
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

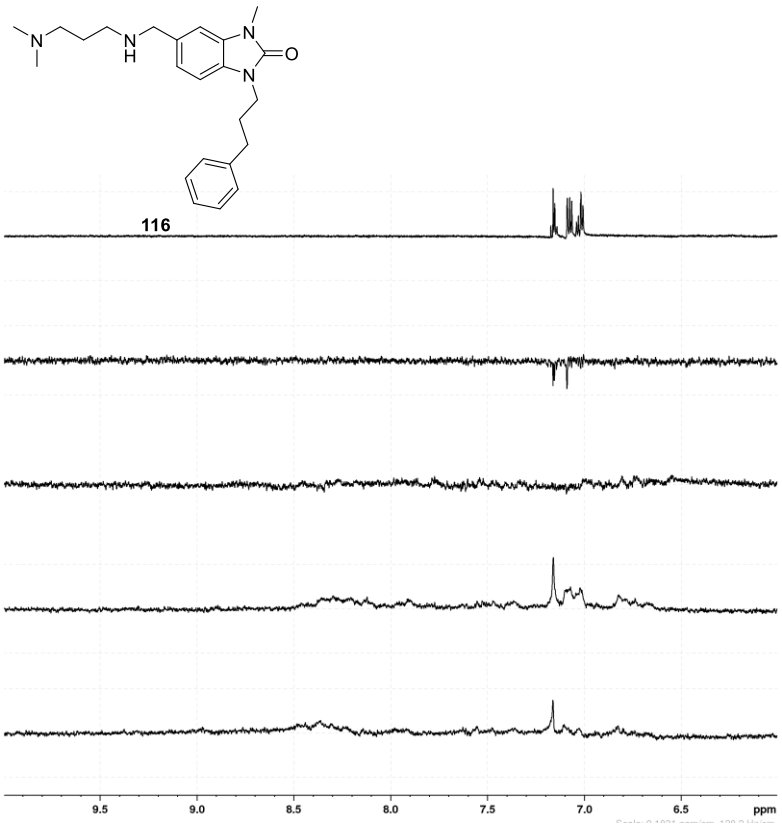

86
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

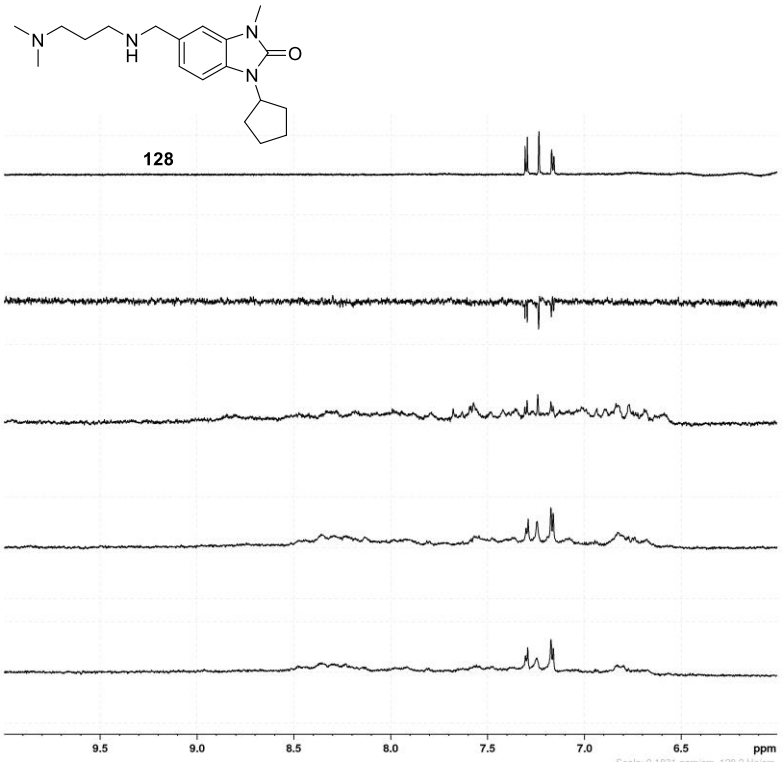

87
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

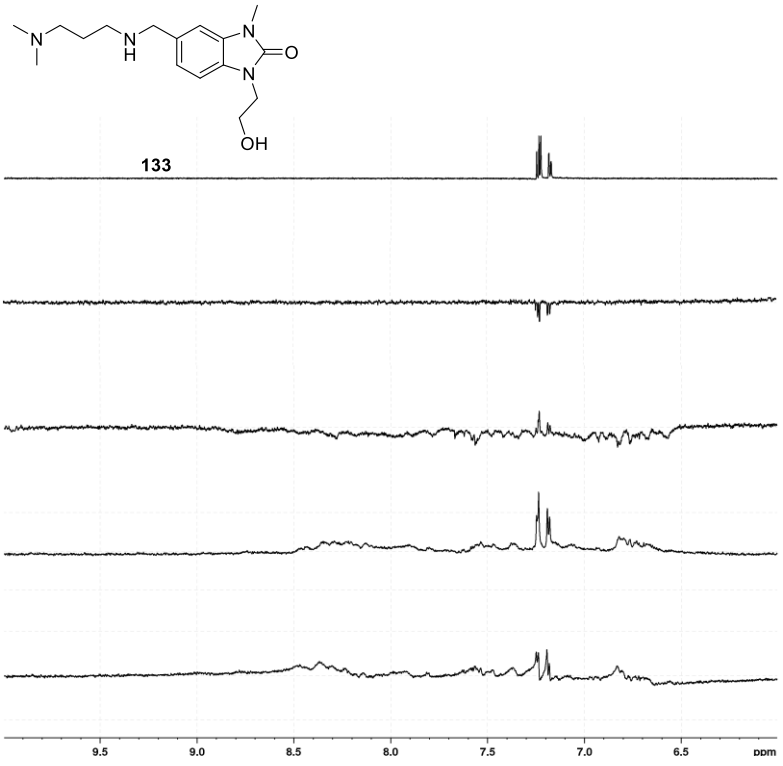

88
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

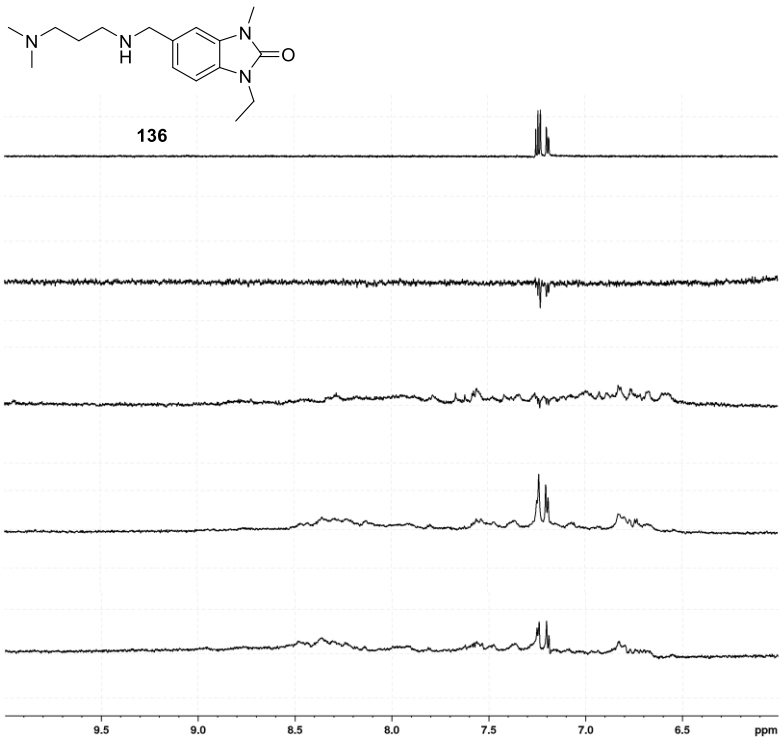

114A
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

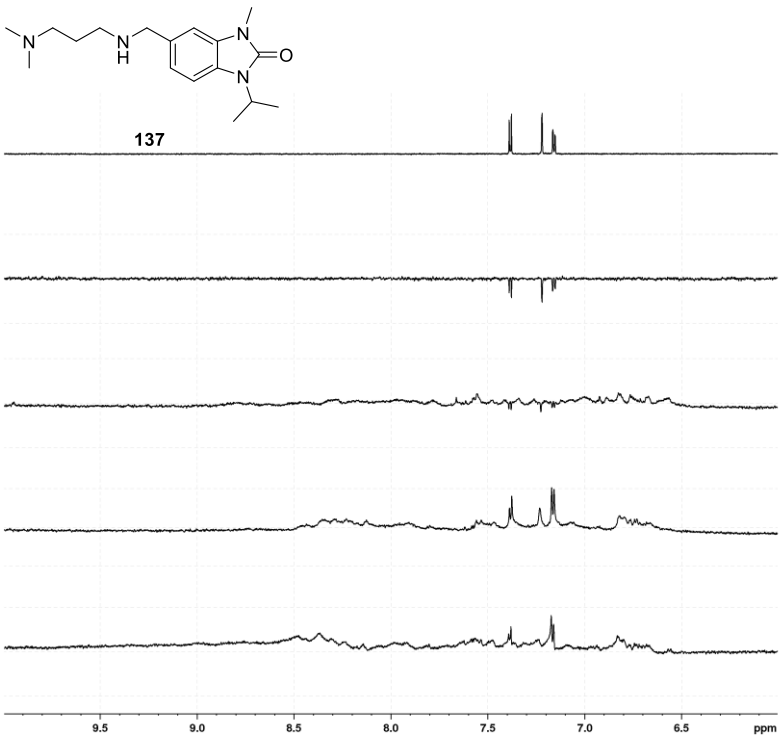

115A
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm


116
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm


128
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm


133
OH
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

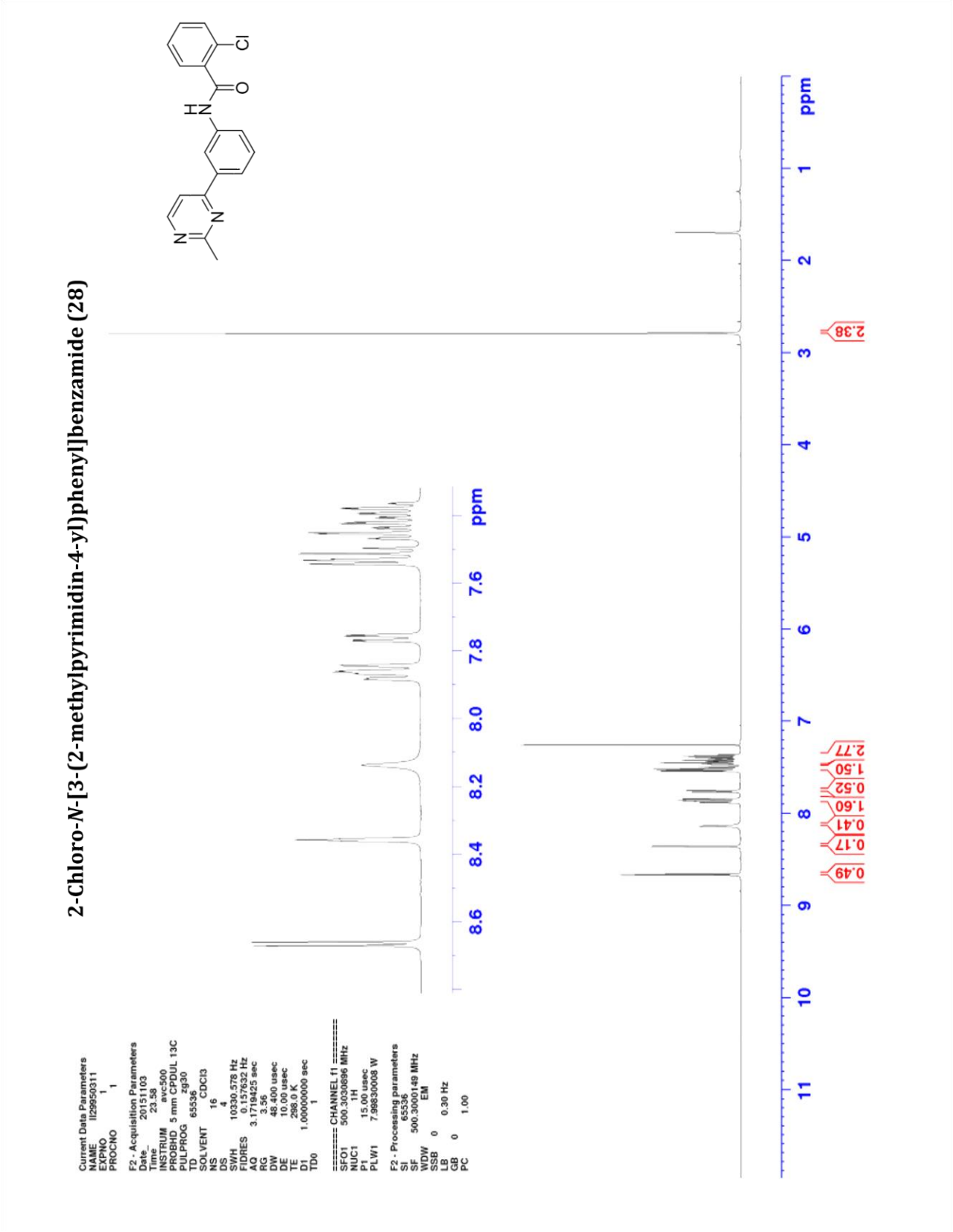

136
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm


137
^{1}H NMR
Compound only
+ TRIM24
+ TRIM33α
+ TRIM33β
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

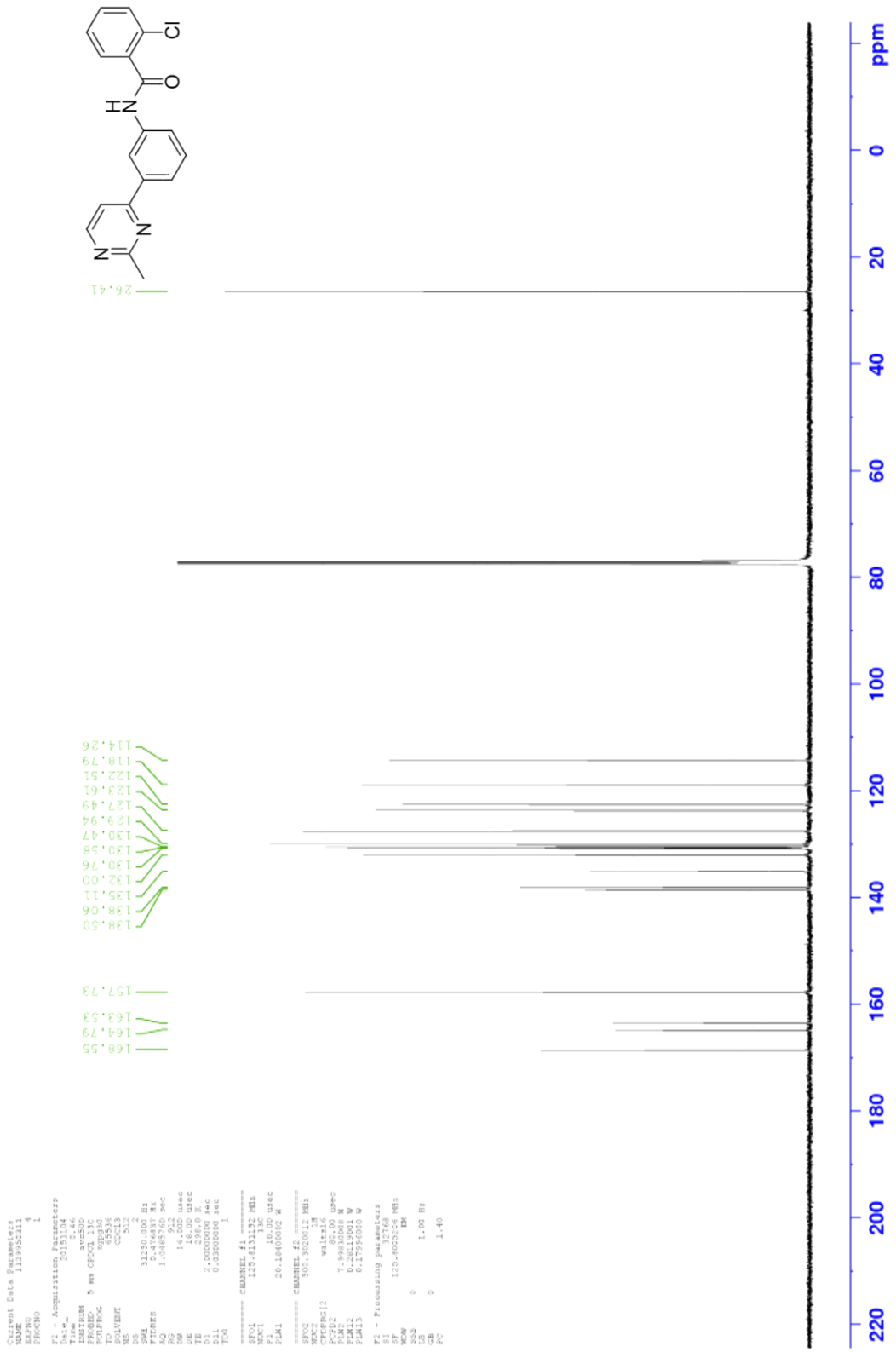
F. NMR Spectra of Novel Compounds

^{1}H, ^{13}C and ^{19}F NMR (where appropriate) are reported for all novel compounds. Spectra are presented in compound numerical order.

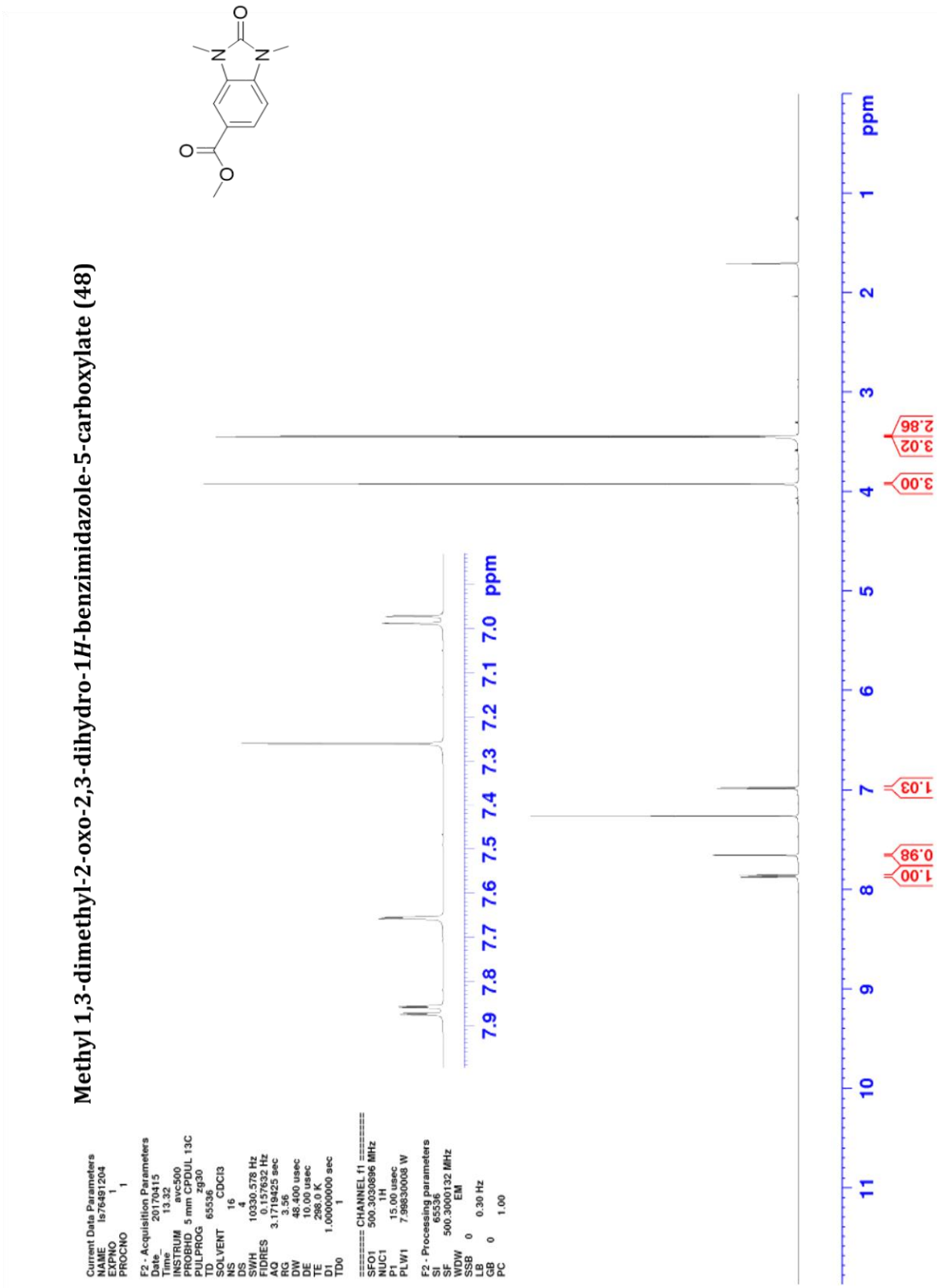
2-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (28)



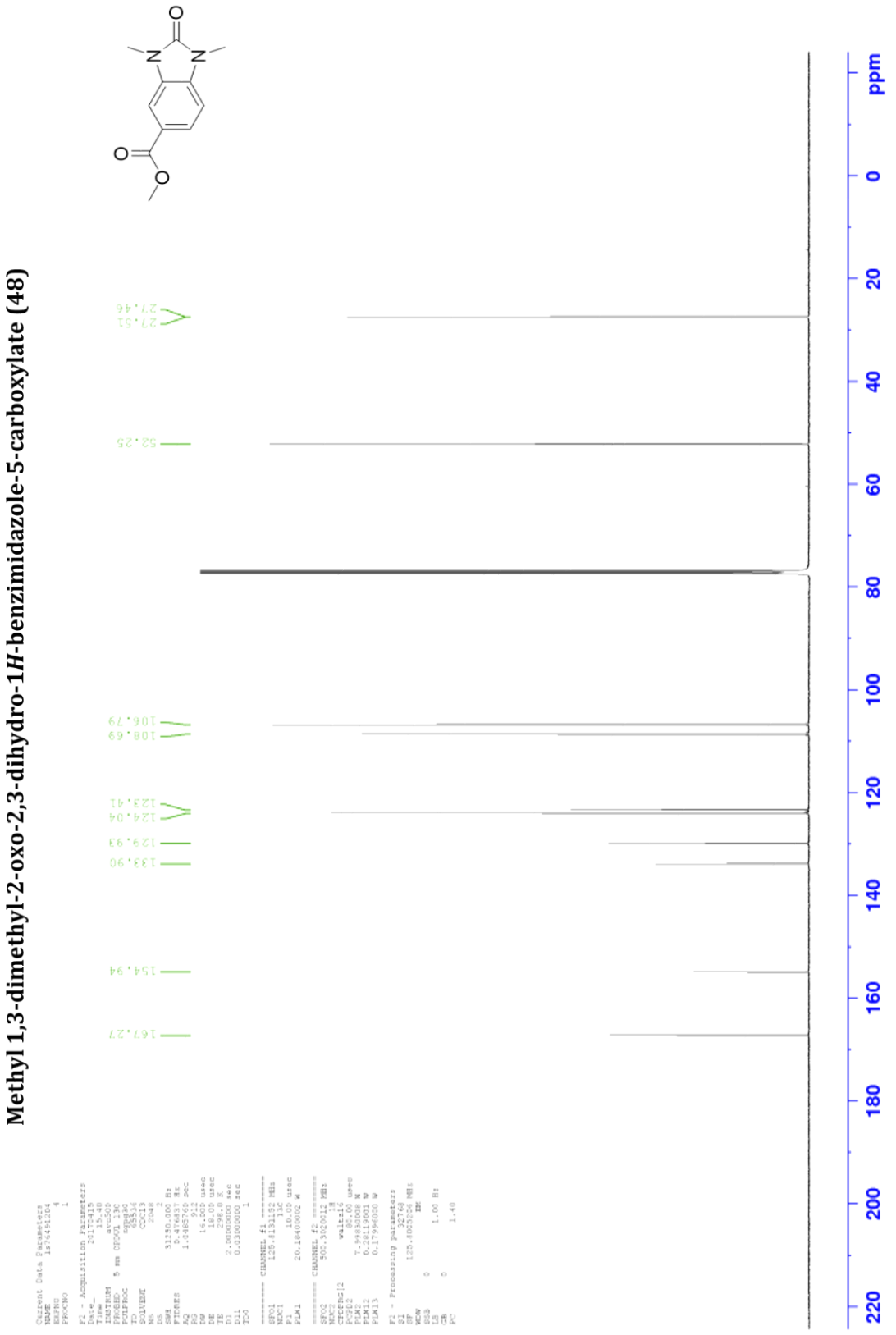
2-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (28)



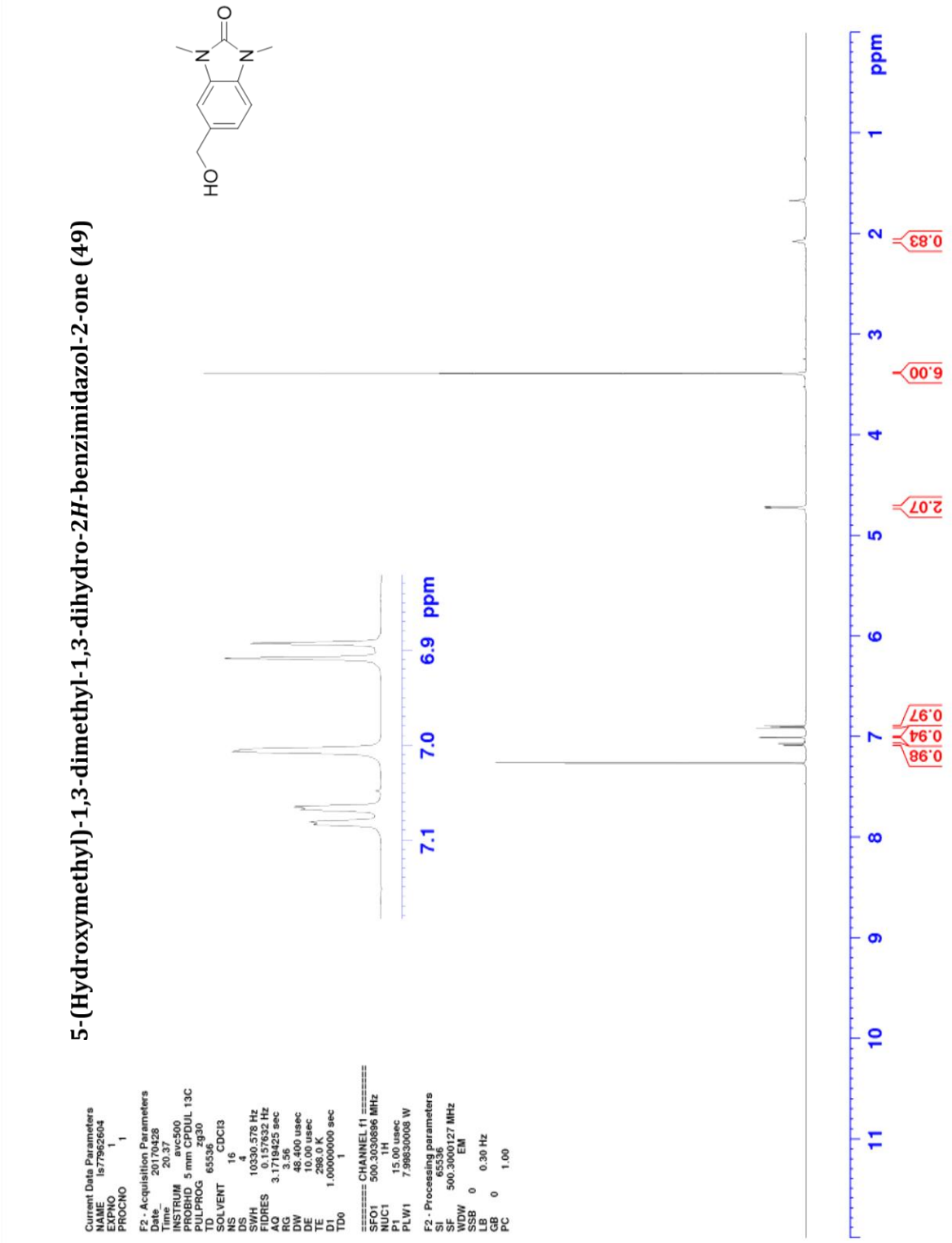
Methyl 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (48)



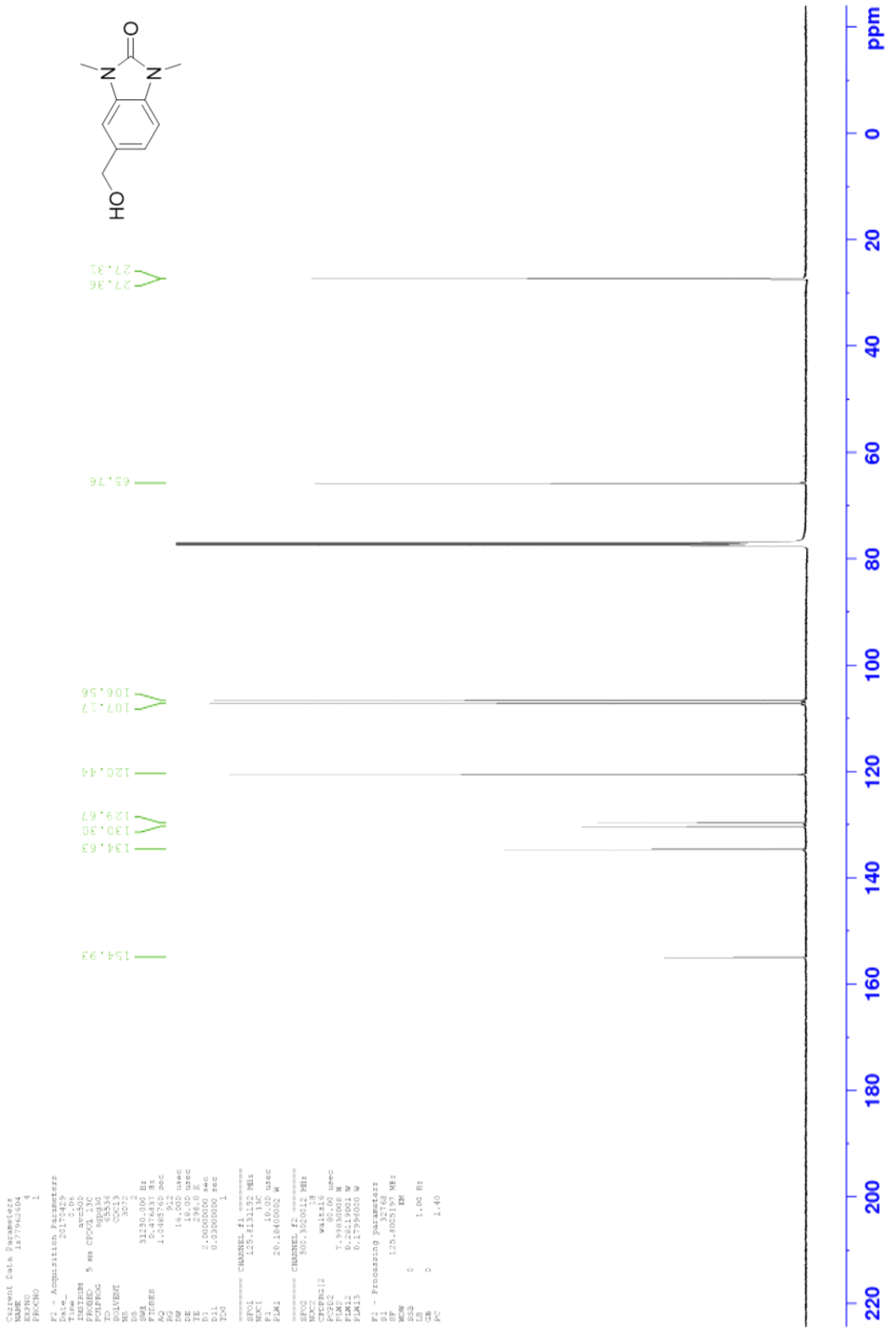
Methyl 1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (48)



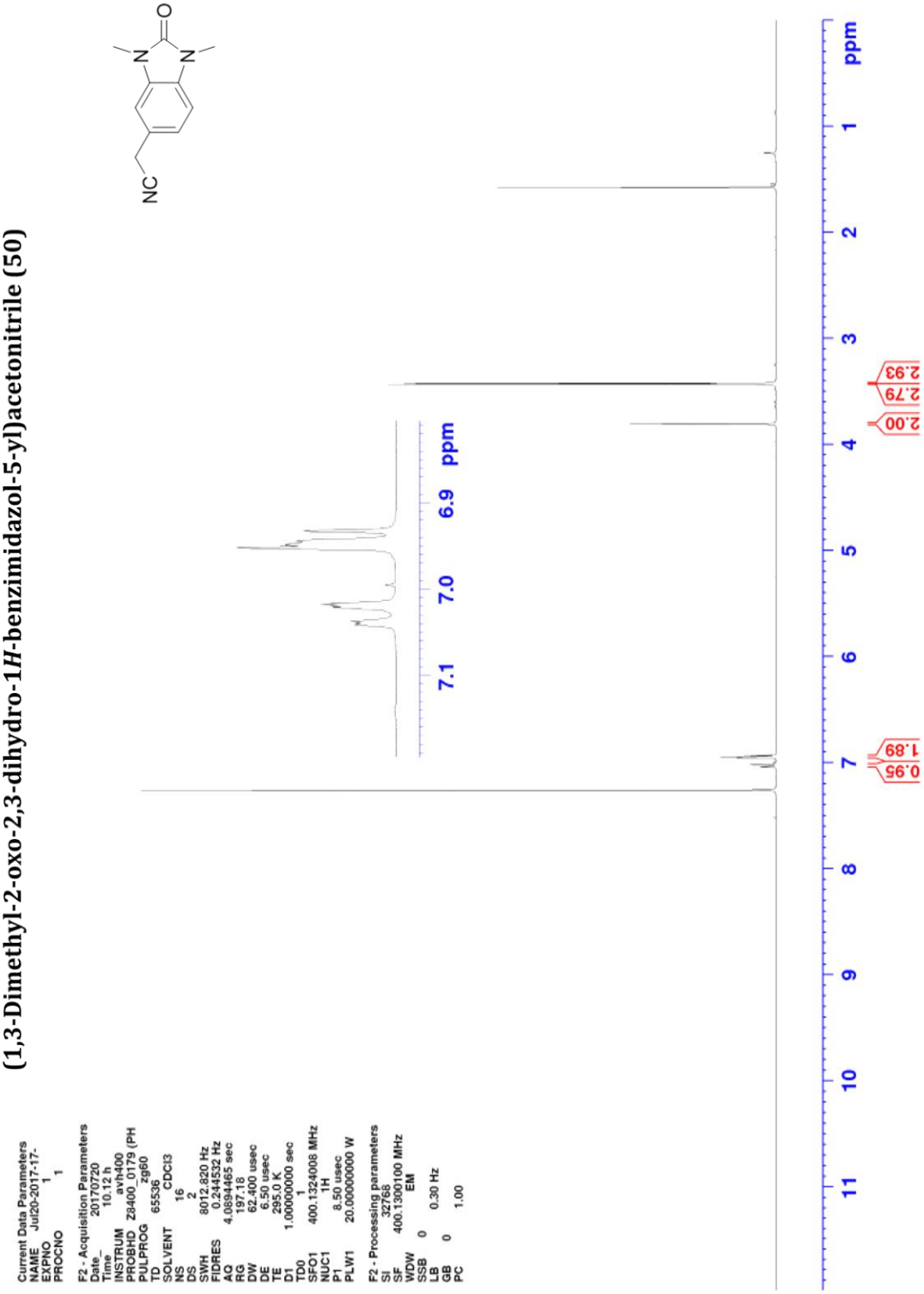
5-(Hydroxymethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (49)



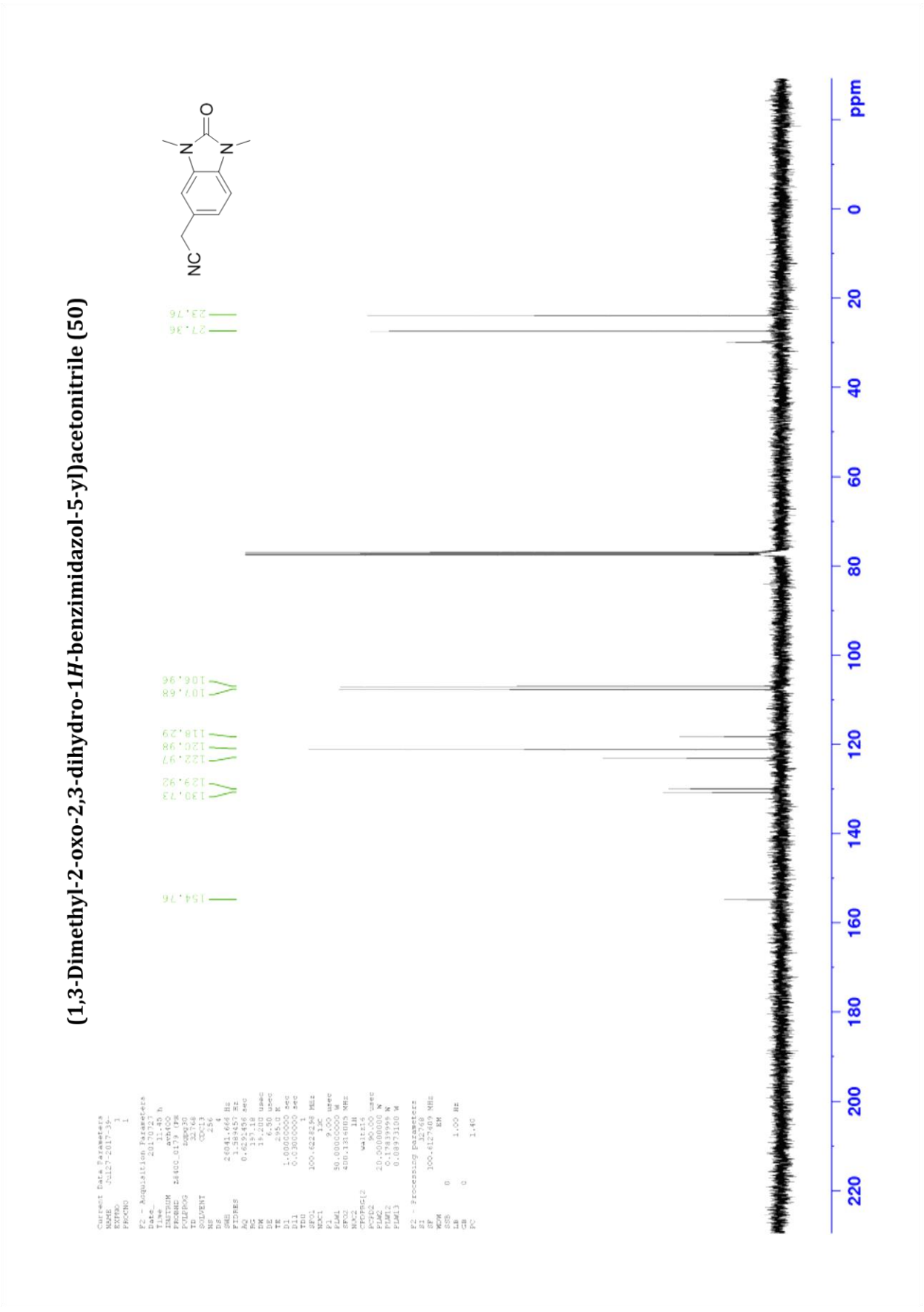
5-(Hydroxymethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (49)



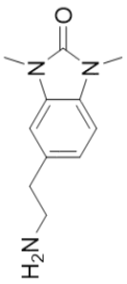
(1,3-Dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)acetonitrile (50)



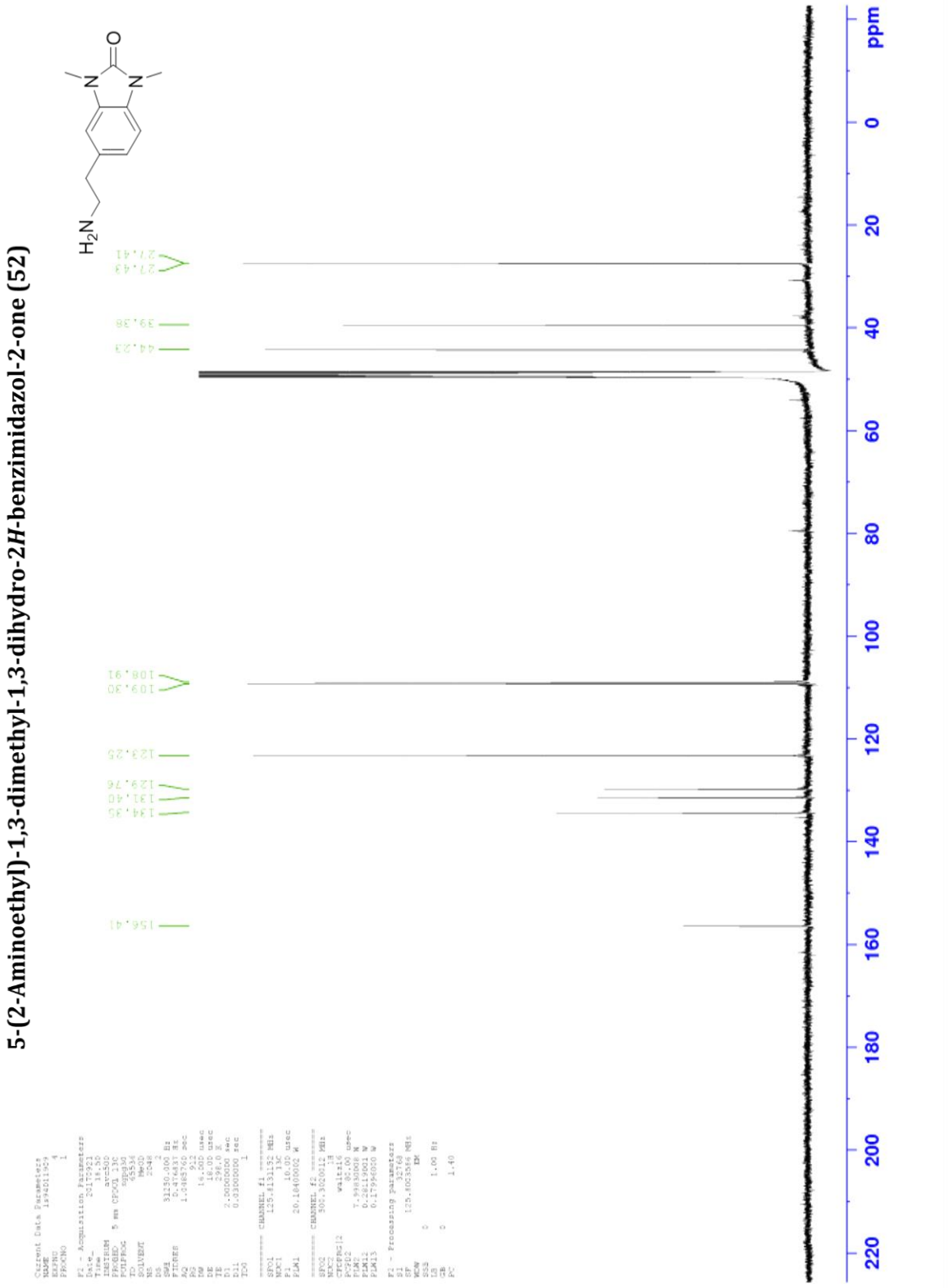
(1,3-Dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)acetonitrile (50)

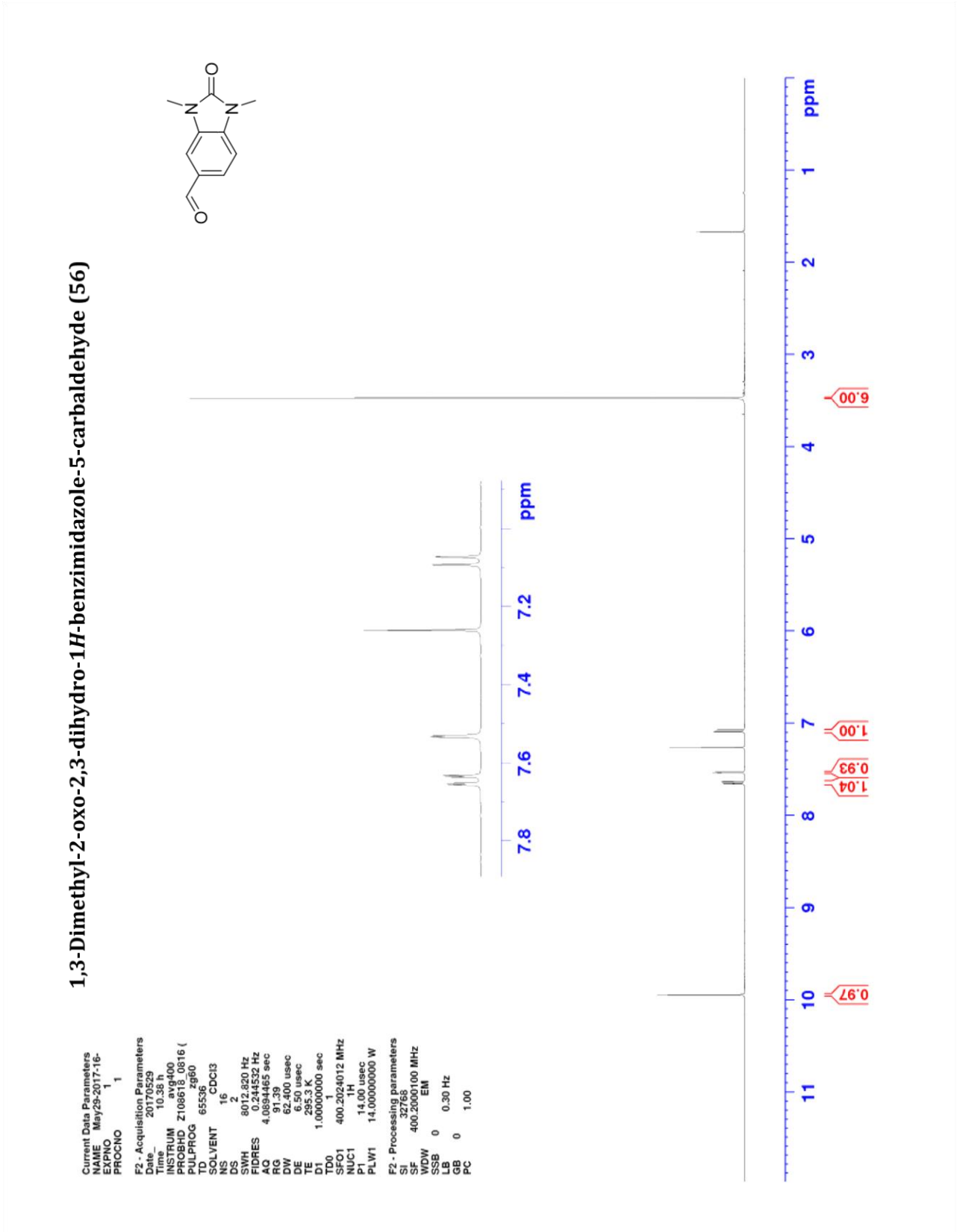


5-(2-Aminoethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (52)

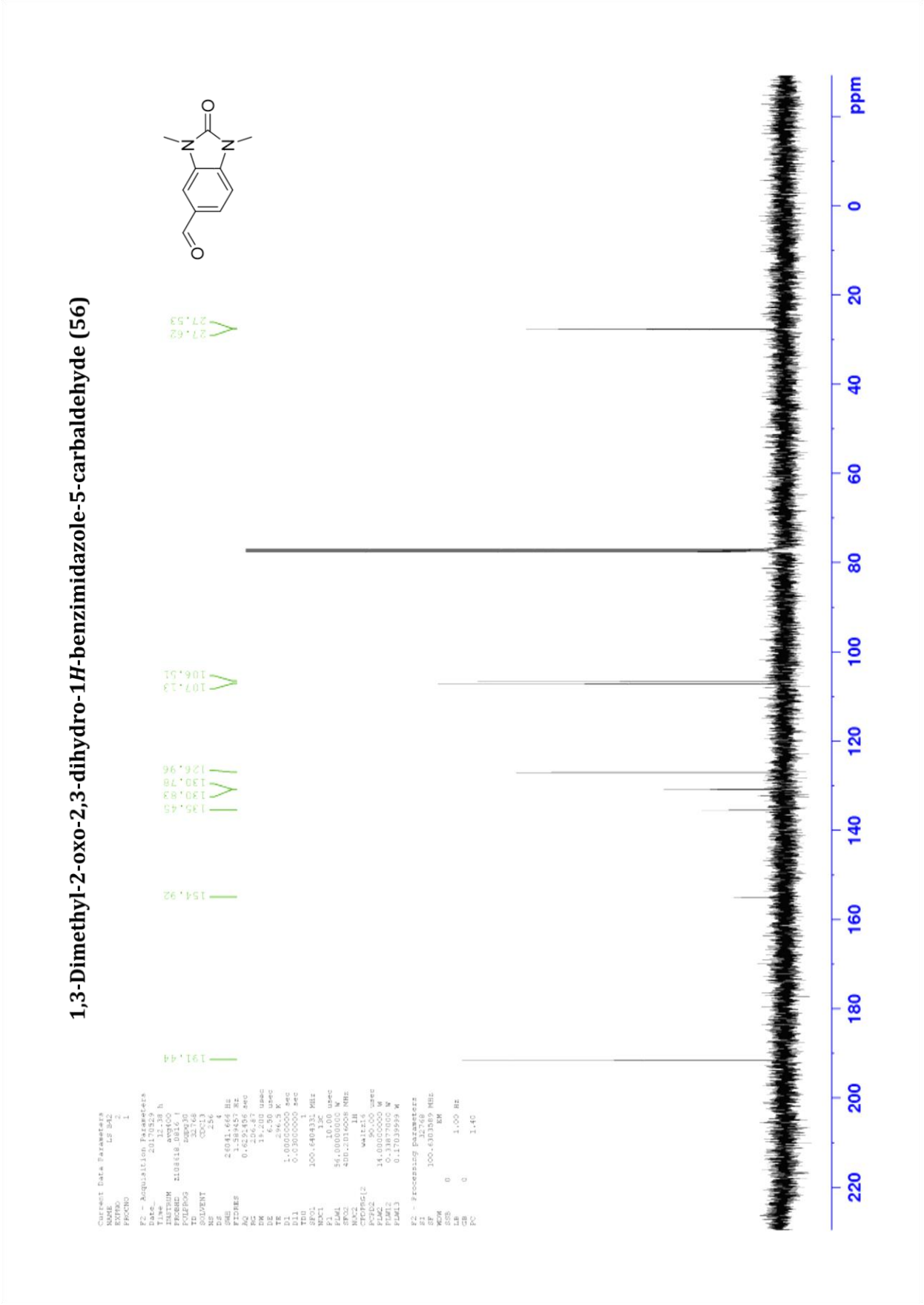


5-(2-Aminoethyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (52)

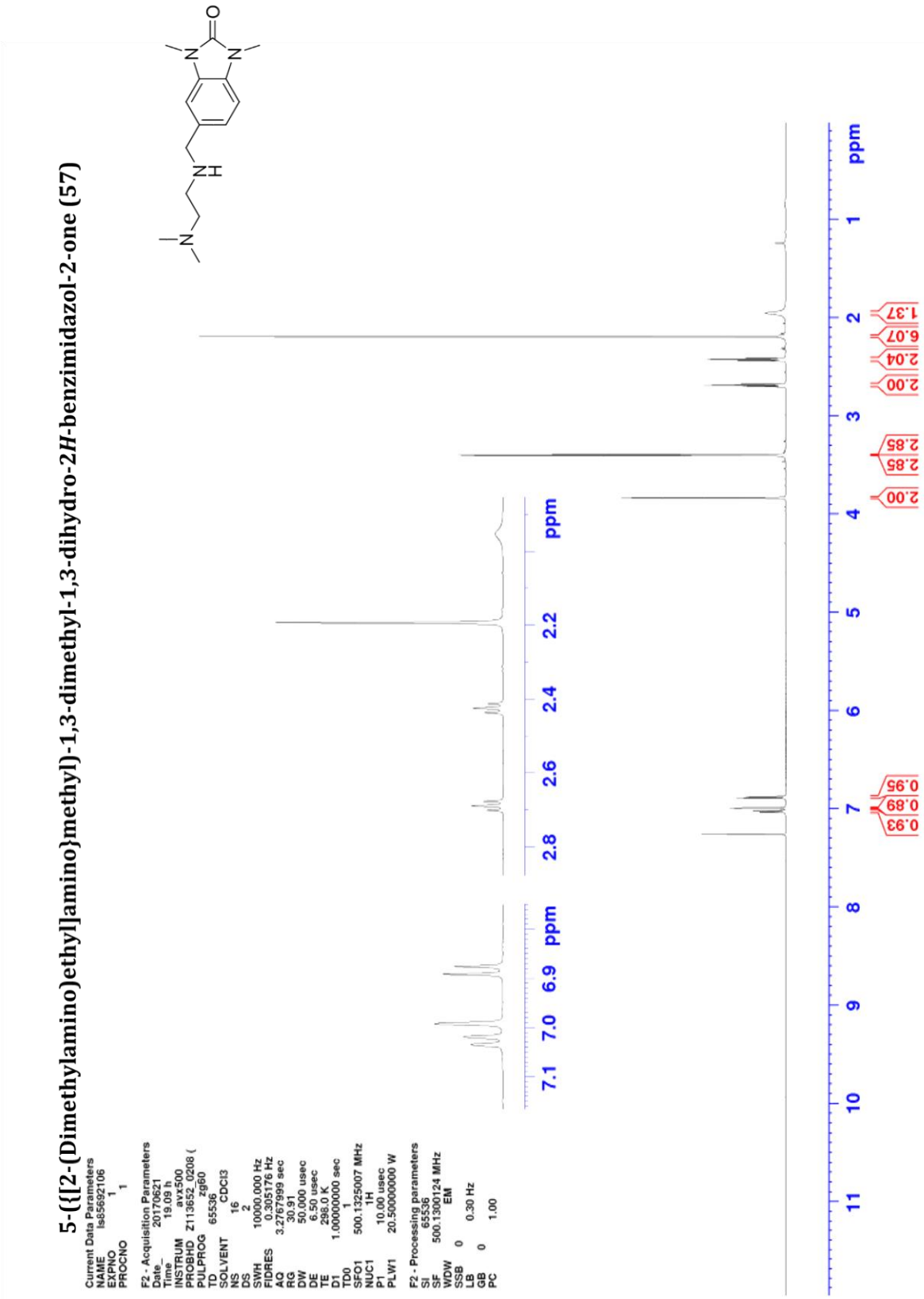



1,3-Dimethyl-2-oxo-2,3-dihydro-1H-benzimidazole-5-carbaldehyde (56)
Current Data Parameters
NAME May29-2017-16-
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20170529
Time 10.38 h
INSTRUM avg400
PROBHD Z108618_0816 (
PULPROG zg60
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8012.820 Hz
FIDRES 0.244532 Hz
AQ 4.0894465 sec
RG 91.39
DW 62.400 usec
DE 6.50 usec
TE 295.3 K
D1 1.00000000 sec
TD0 1
SFO1 400.20240012 MHz
NUC1 1H
P1 14.00 usec
PLW1 14.00000000 W
F2 - Processing parameters
SI 32768
SF 400.2000100 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.8
7.6
7.4
7.2
ppm
0.97
1.04
0.93
1.00
6.00
11
10
9
8
7
6
5
4
3
2
1
ppm

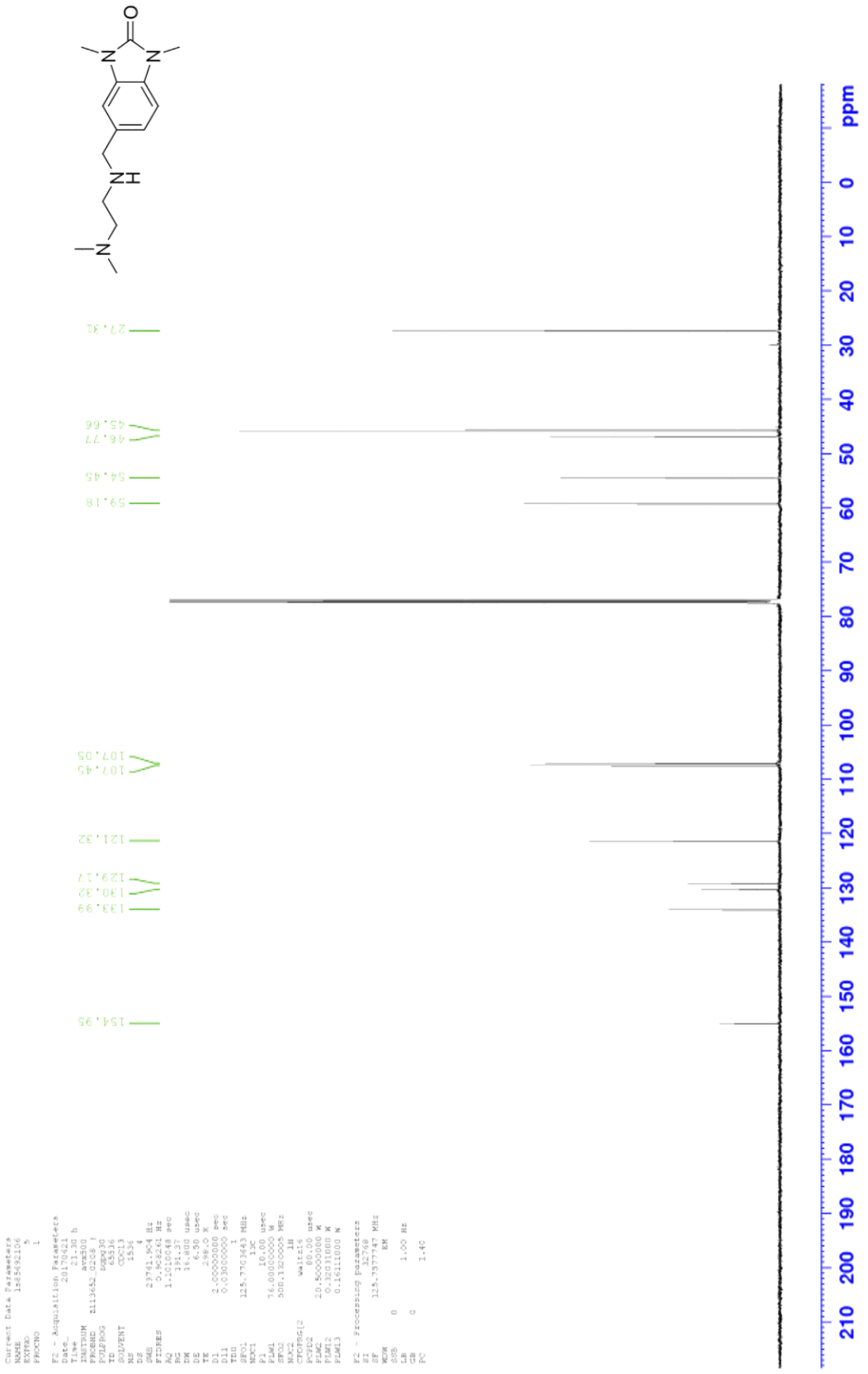
1,3-Dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (56)



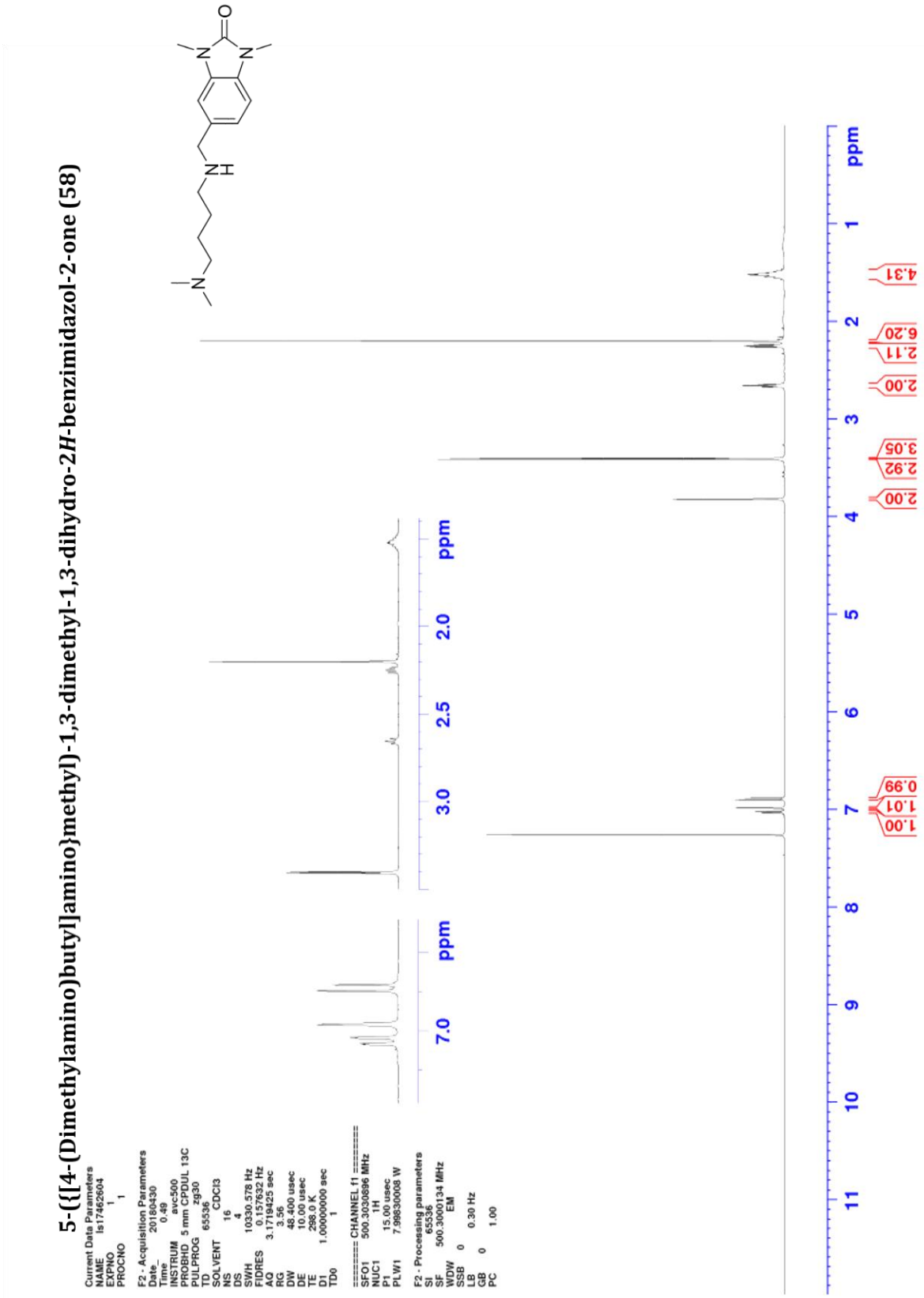
5-({[2-(Dimethylamino)ethyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (57)



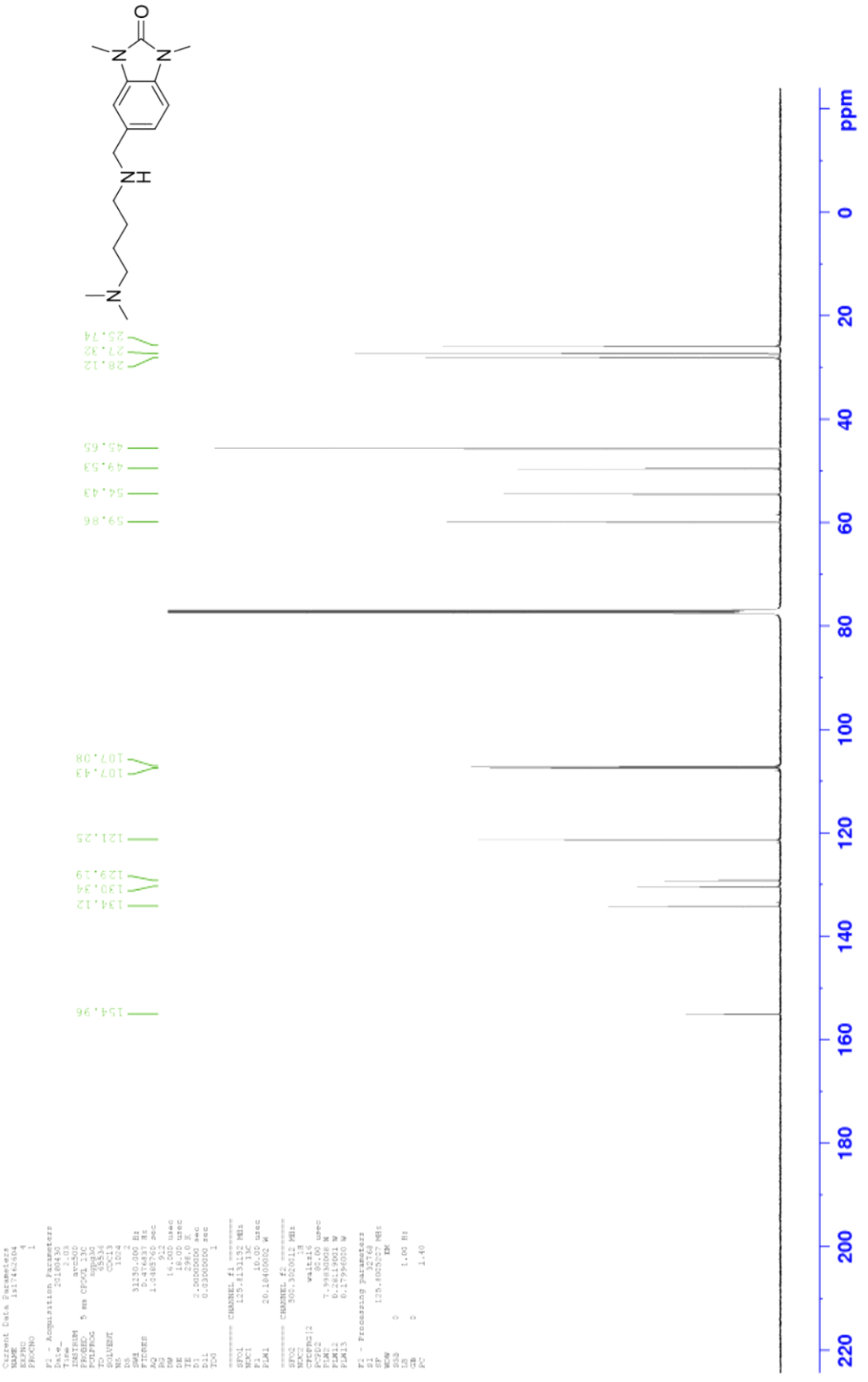
5-({[2-(Dimethylamino)ethyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (57)



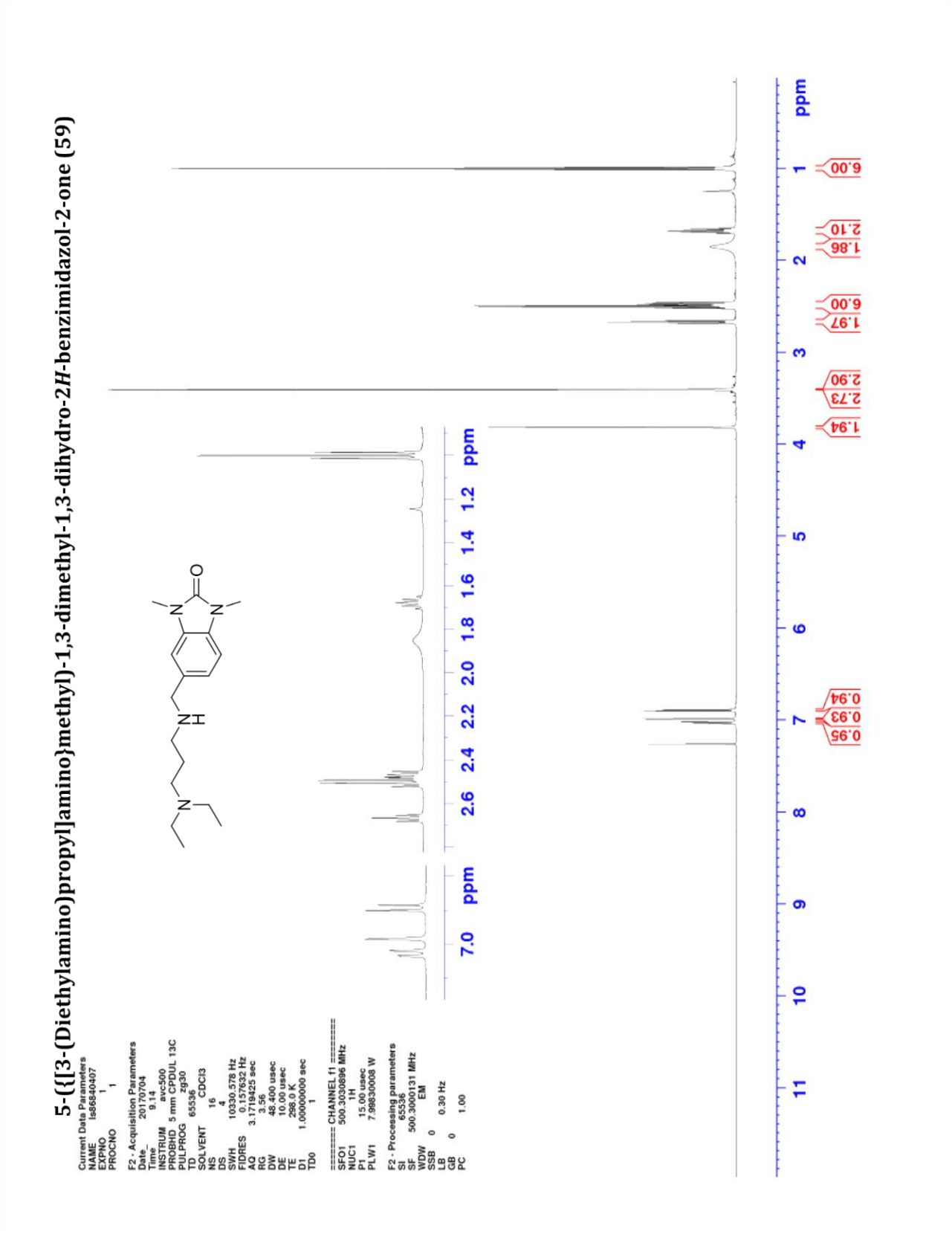
5-({[4-(Dimethylamino)butyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (58)



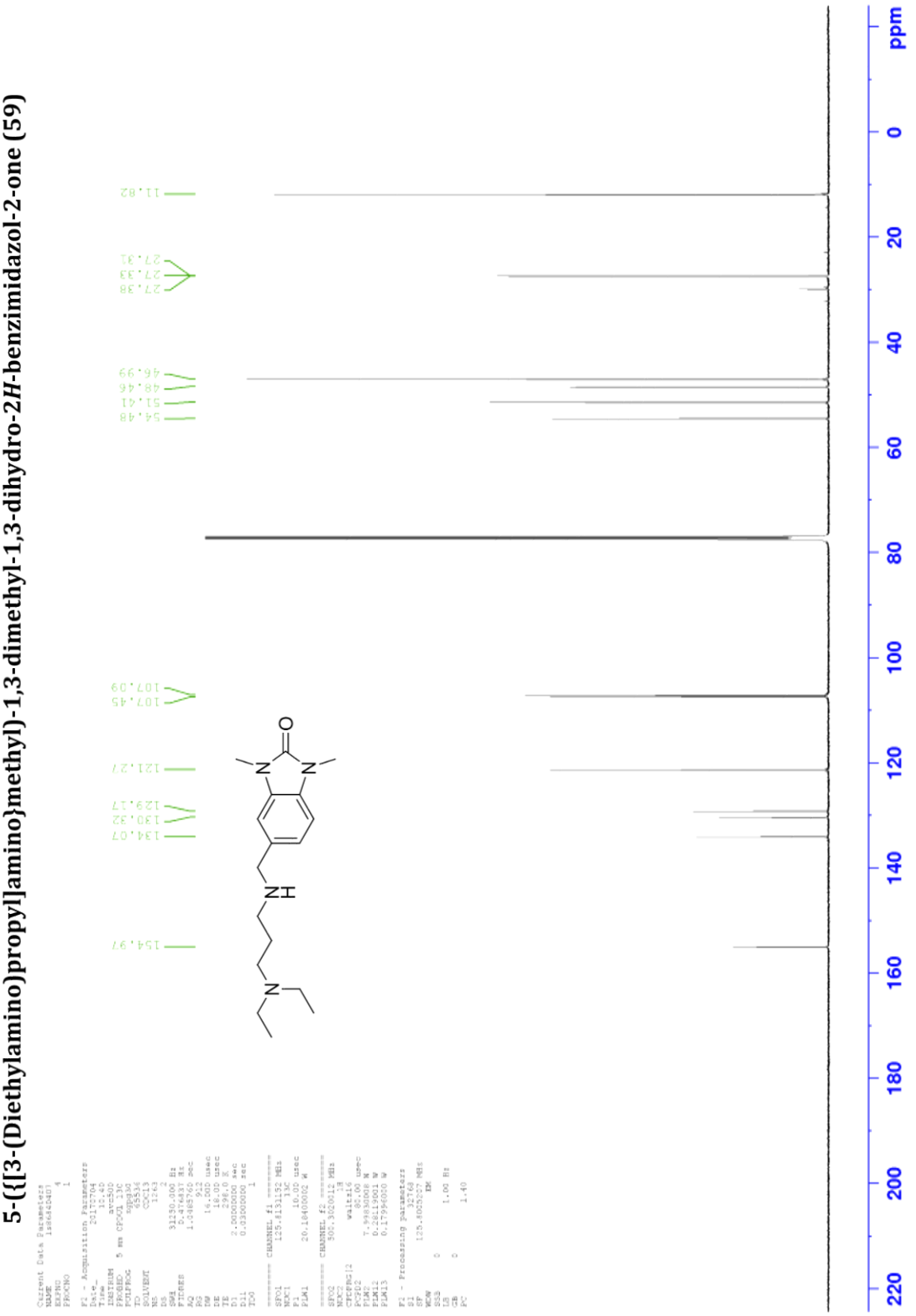
5-({[4-(Dimethylamino)butyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (58)



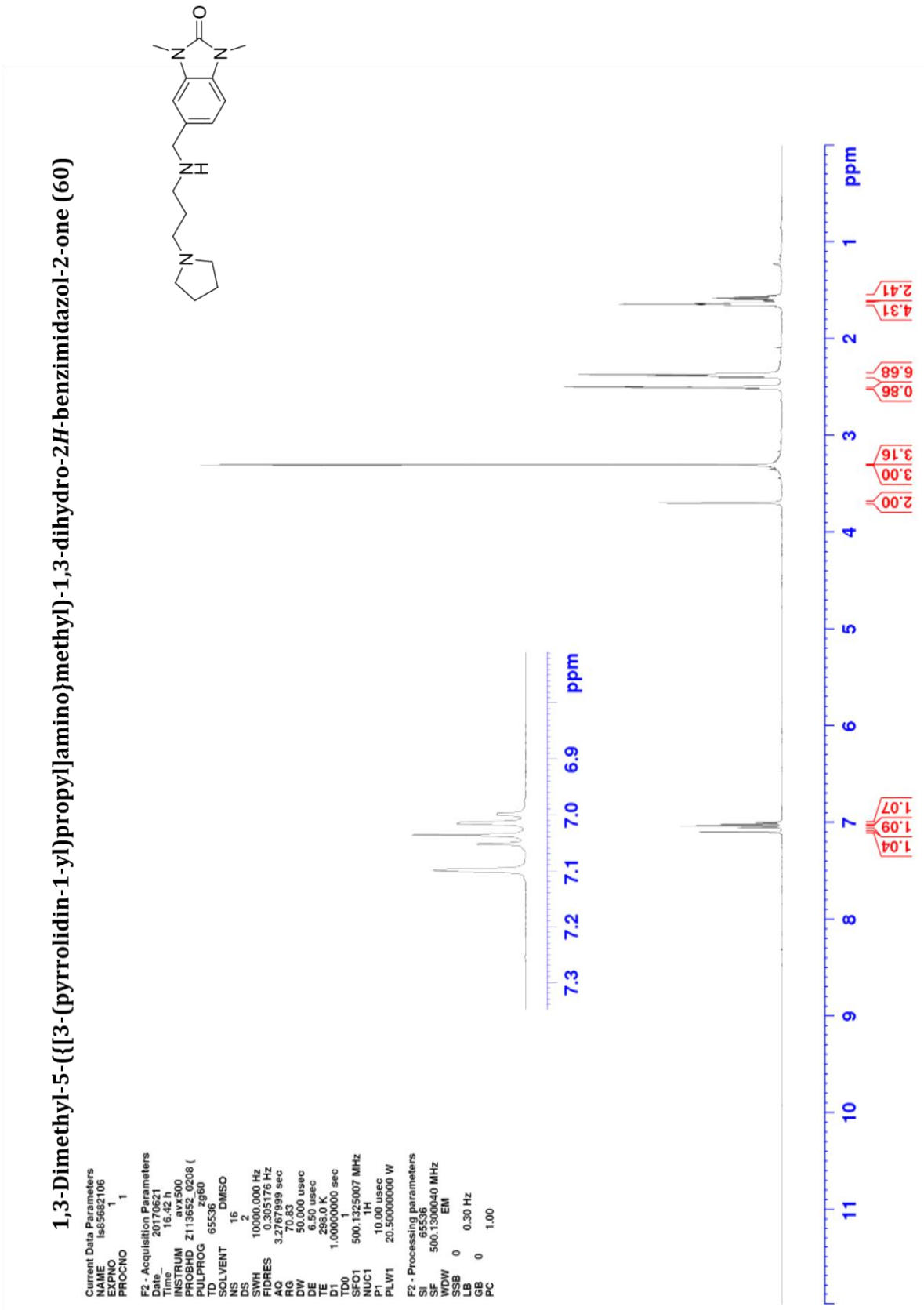
5-({[3-(Diethylamino)propyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (59)



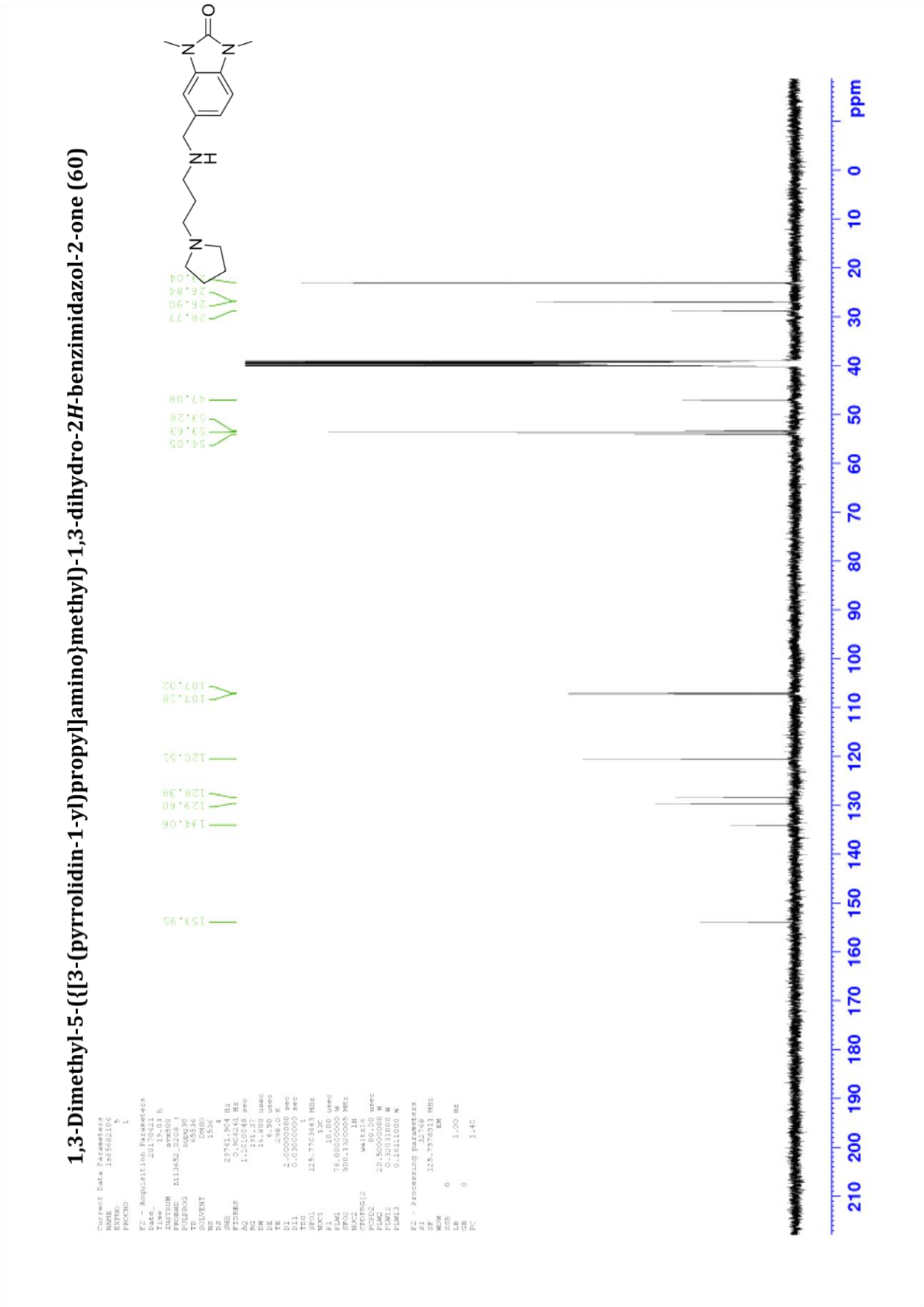
5-({[3-(Diethylamino)propyl]amino}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (59)



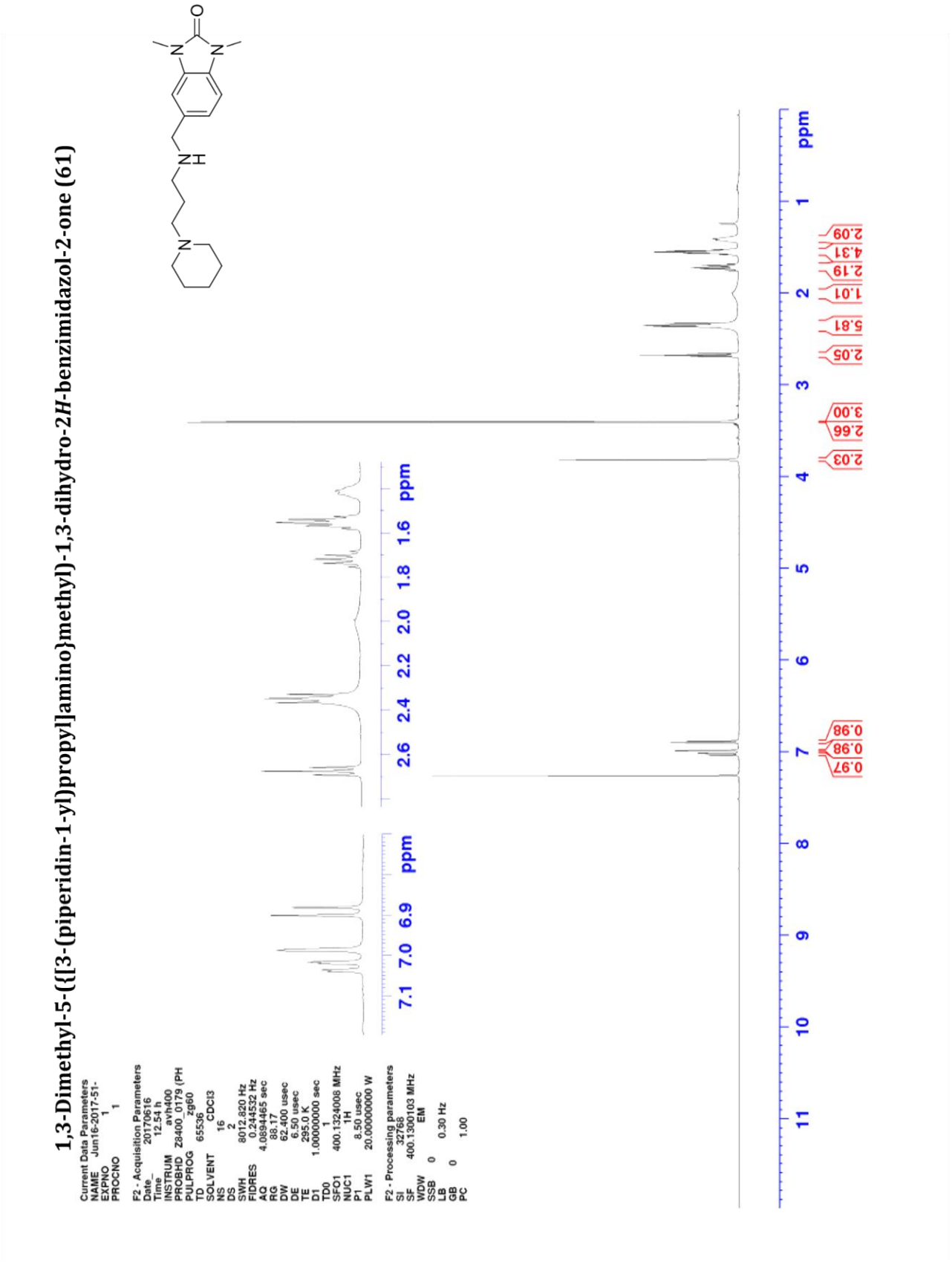
1,3-Dimethyl-5-({[3-(pyrrolidin-1-yl)propyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (60)



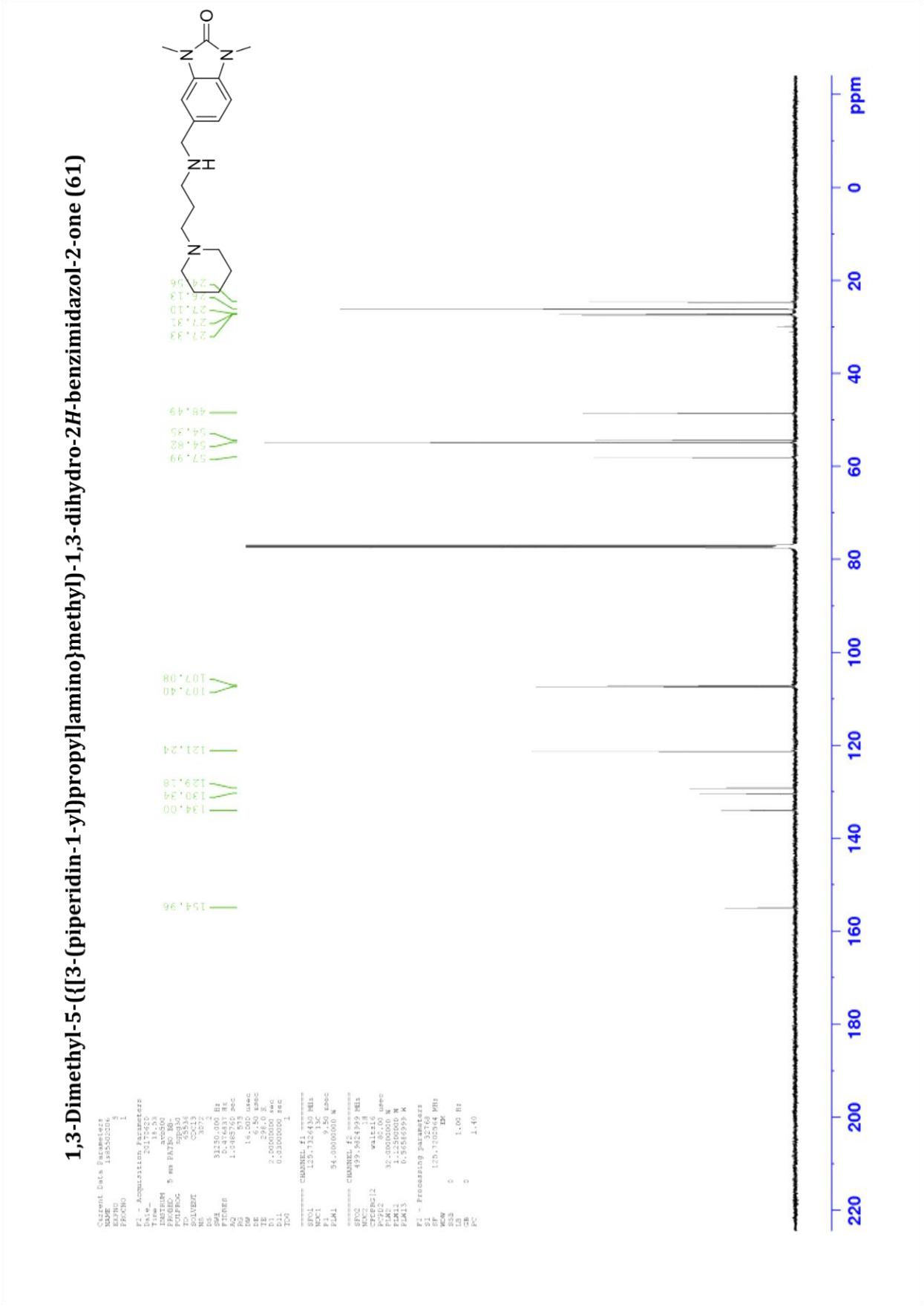
1,3-Dimethyl-5-({[3-(pyrrolidin-1-yl)propyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (60)



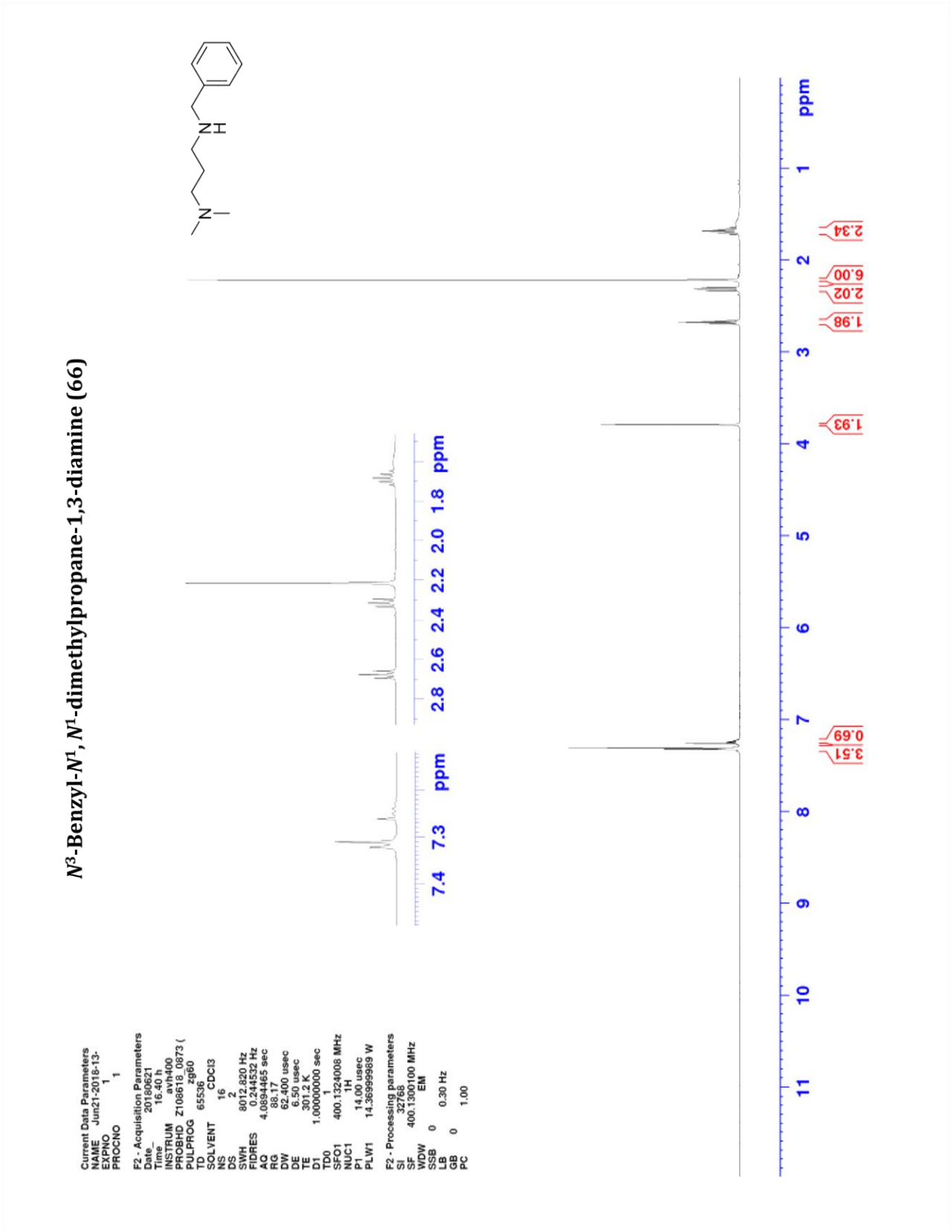
1,3-Dimethyl-5-({[3-(piperidin-1-yl)propyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (61)



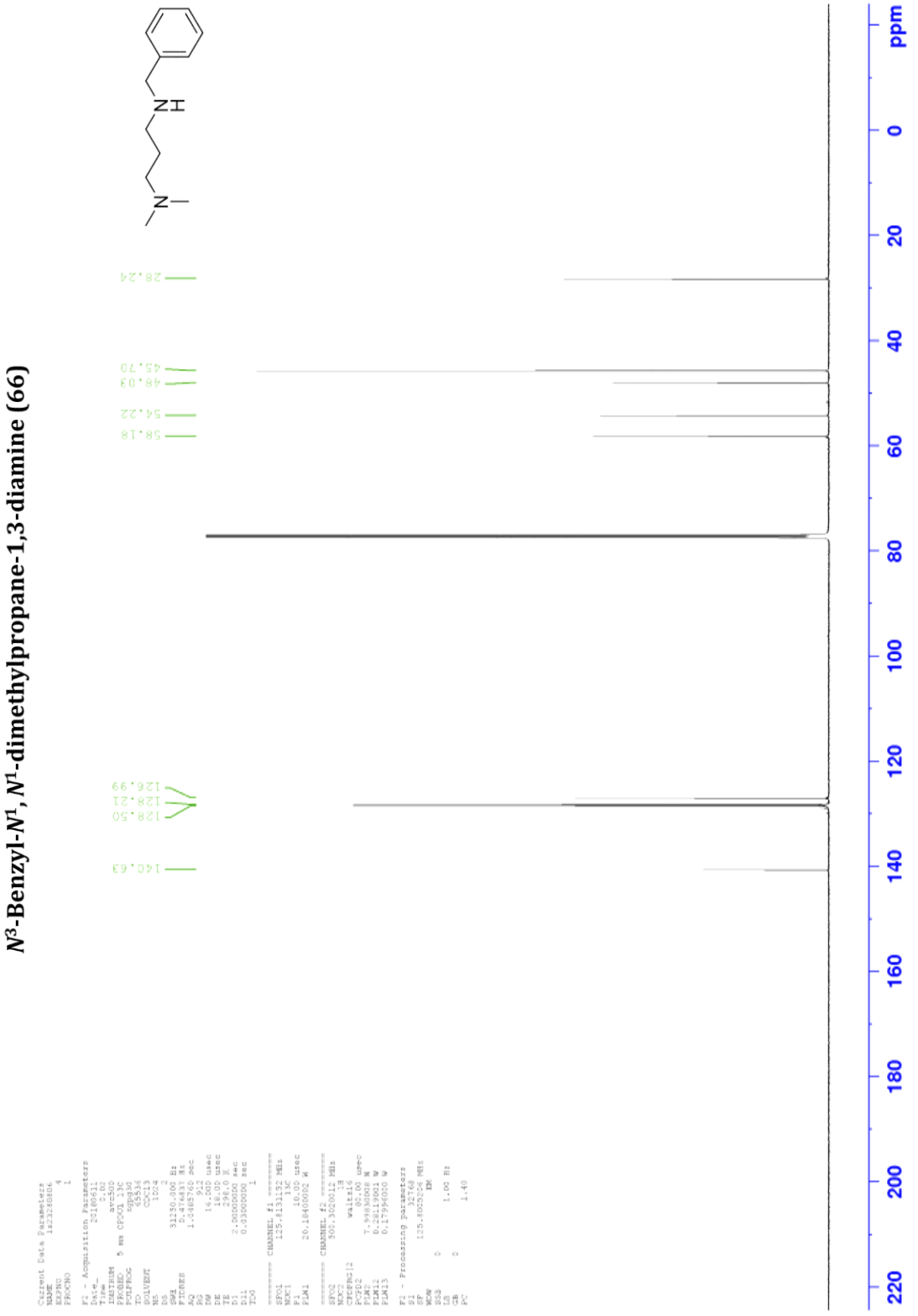
1,3-Dimethyl-5-({[3-(piperidin-1-yl)propyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (61)



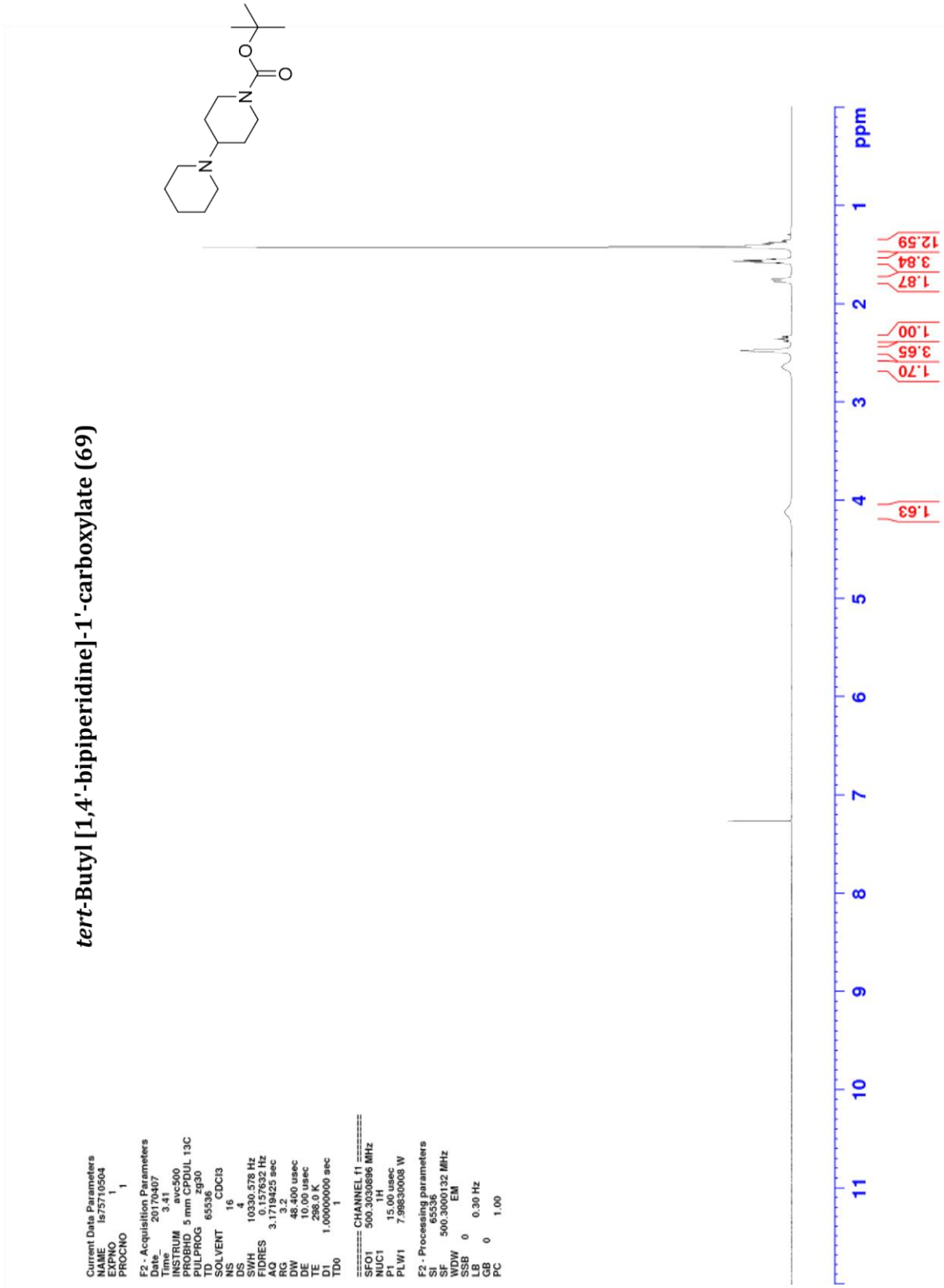
N^3-Benzyl-N^1, N^1-dimethylpropane-1,3-diamine (66)



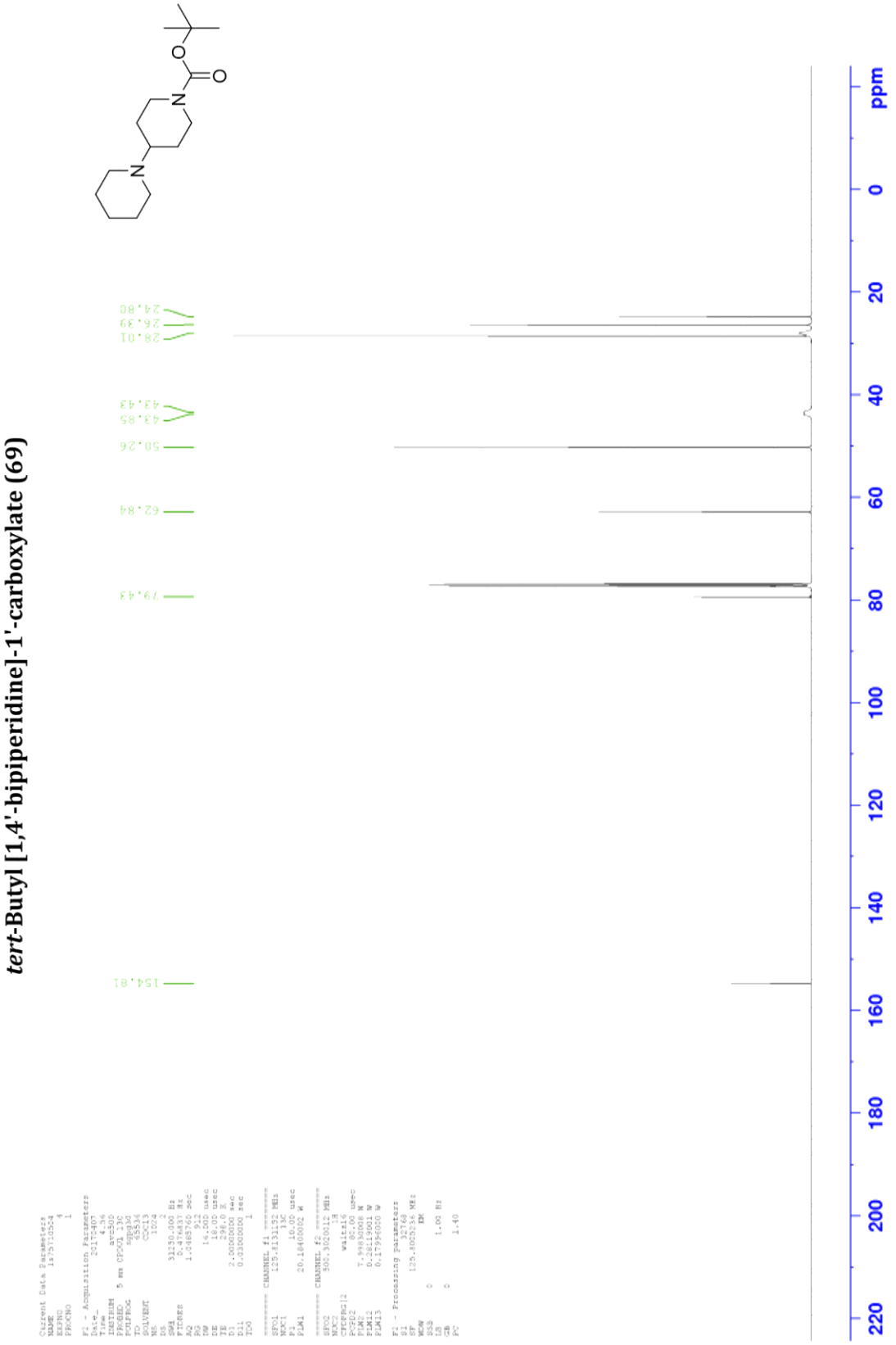
N^3-Benzyl-N^1, N^1-dimethylpropane-1,3-diamine (66)



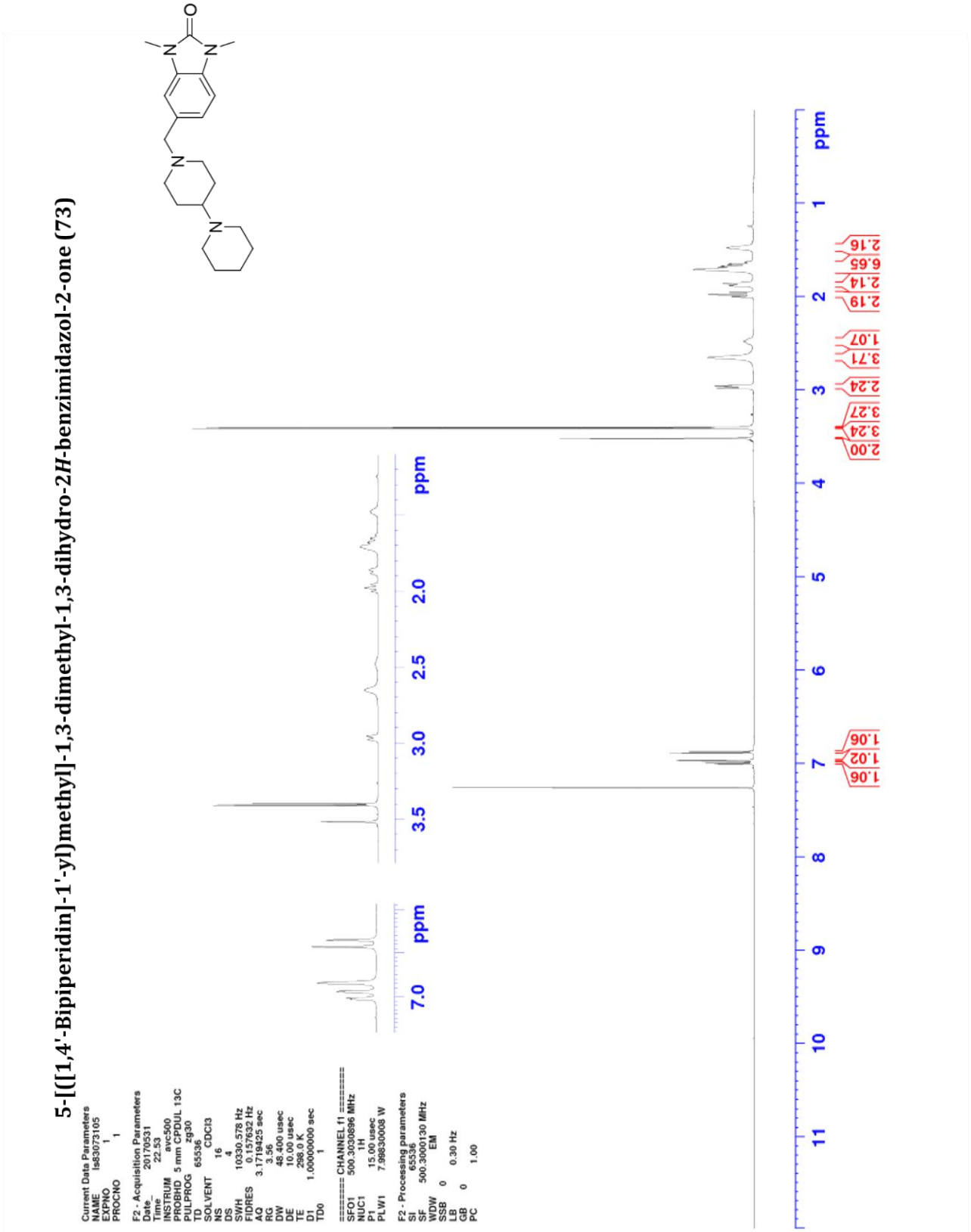
tert-Butyl [1,4'-bipiperidine]-1'-carboxylate (69)



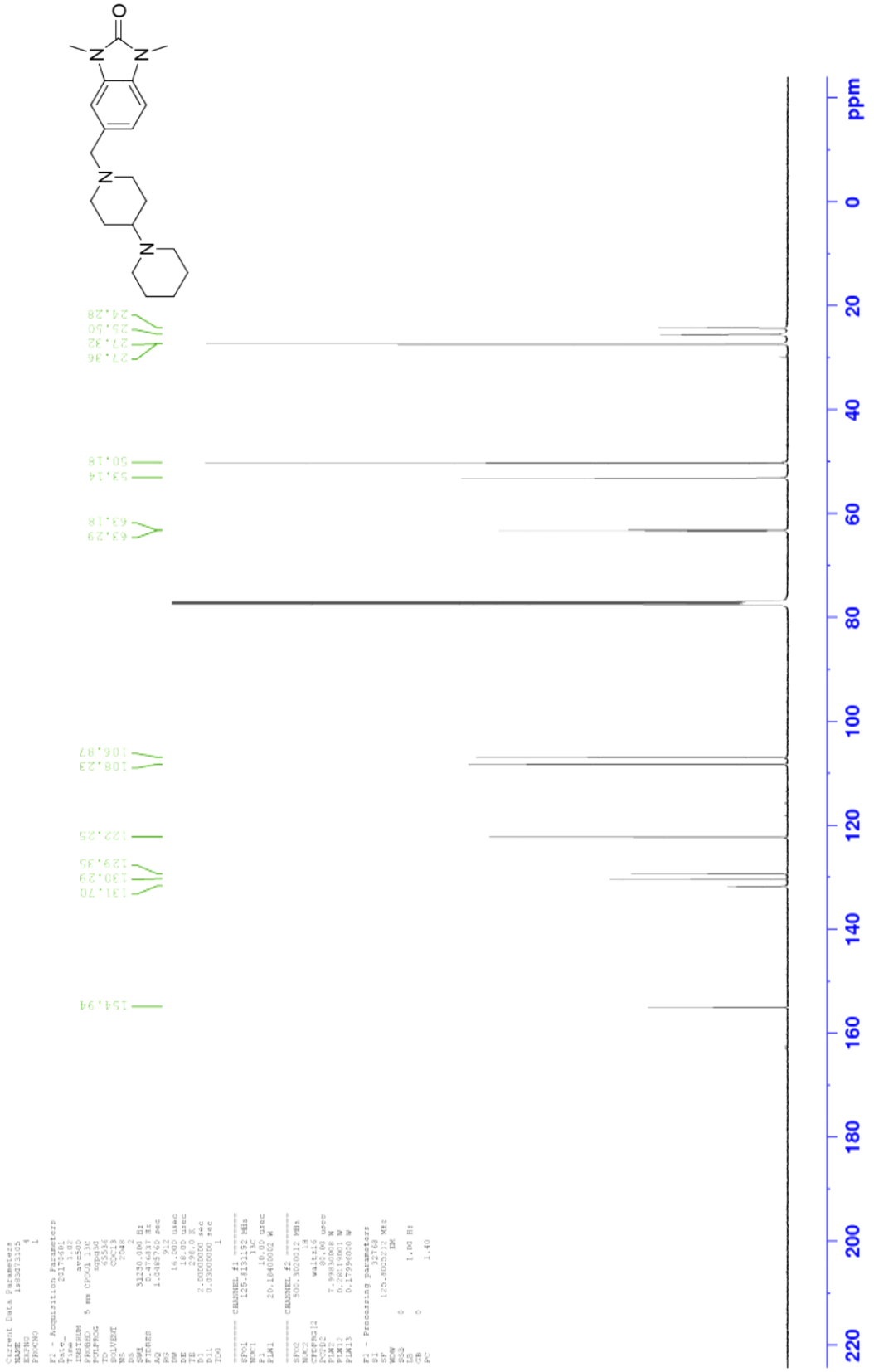
tert-Butyl [1,4'-bipiperidine]-1'-carboxylate (69)



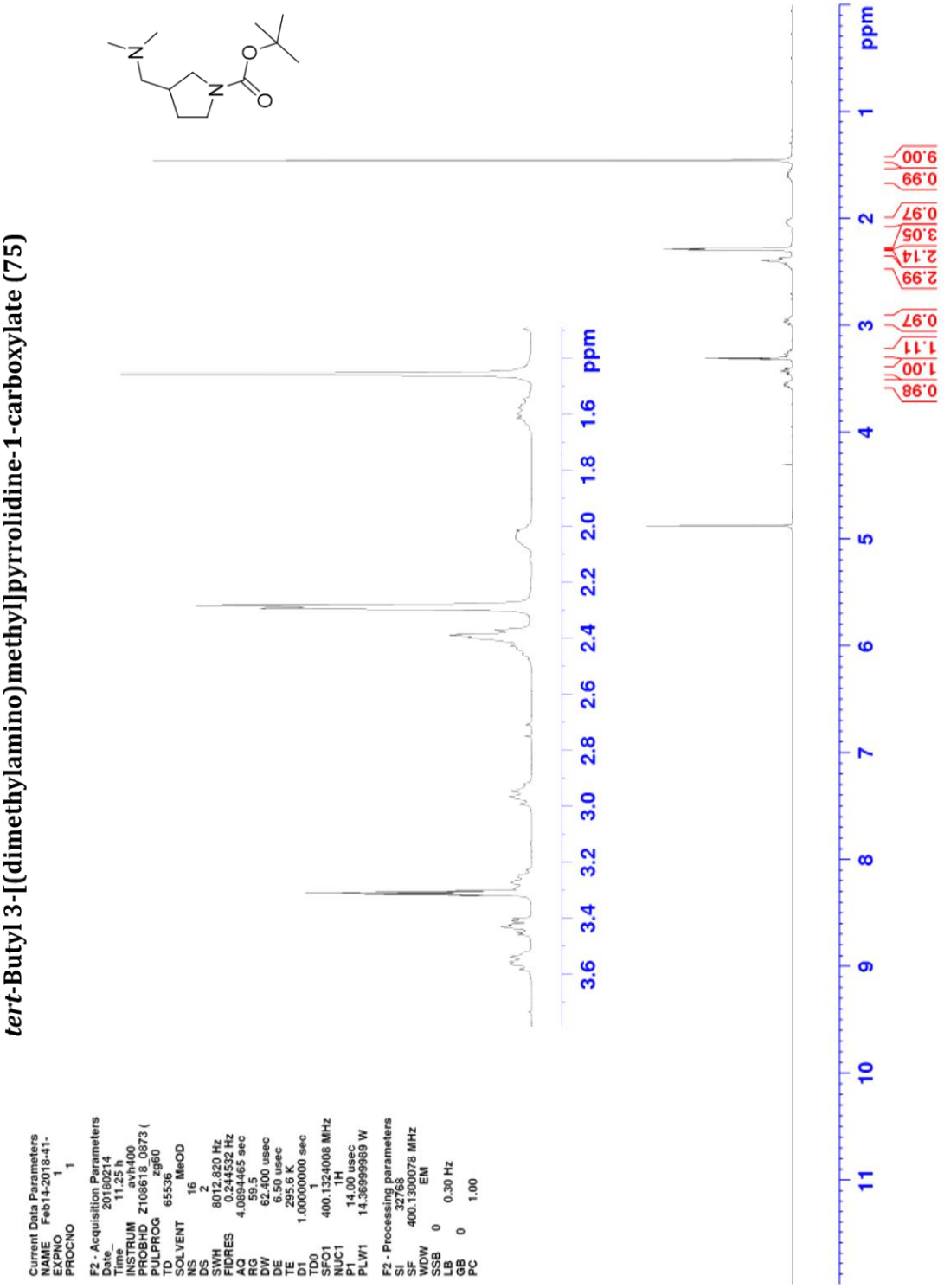
5-[([1,4'-Bipiperidin]-1'-yl)methyl]-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (73)



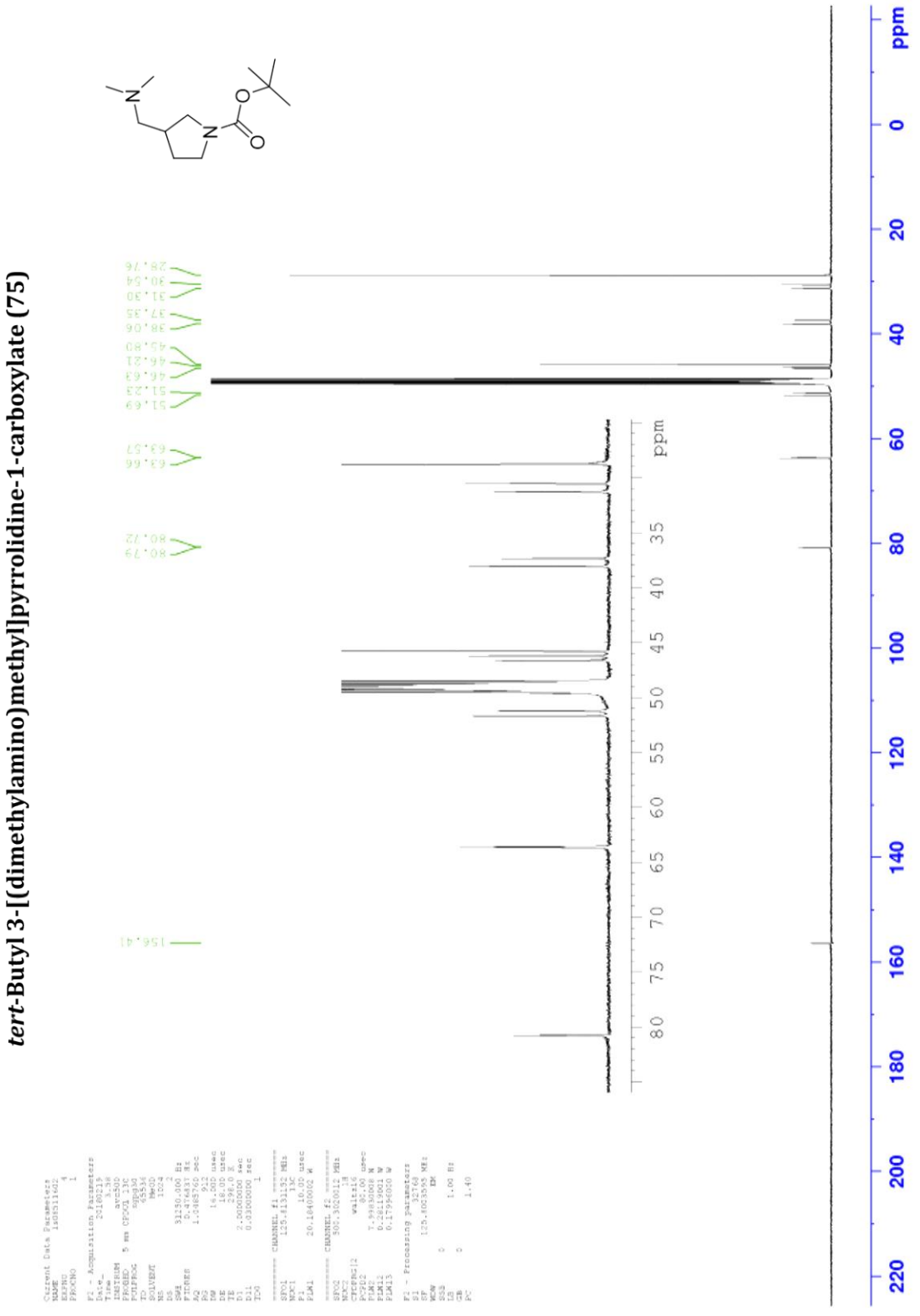
5-[([1,4'-Bipiperidin]-1'-yl)methyl]-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (73)



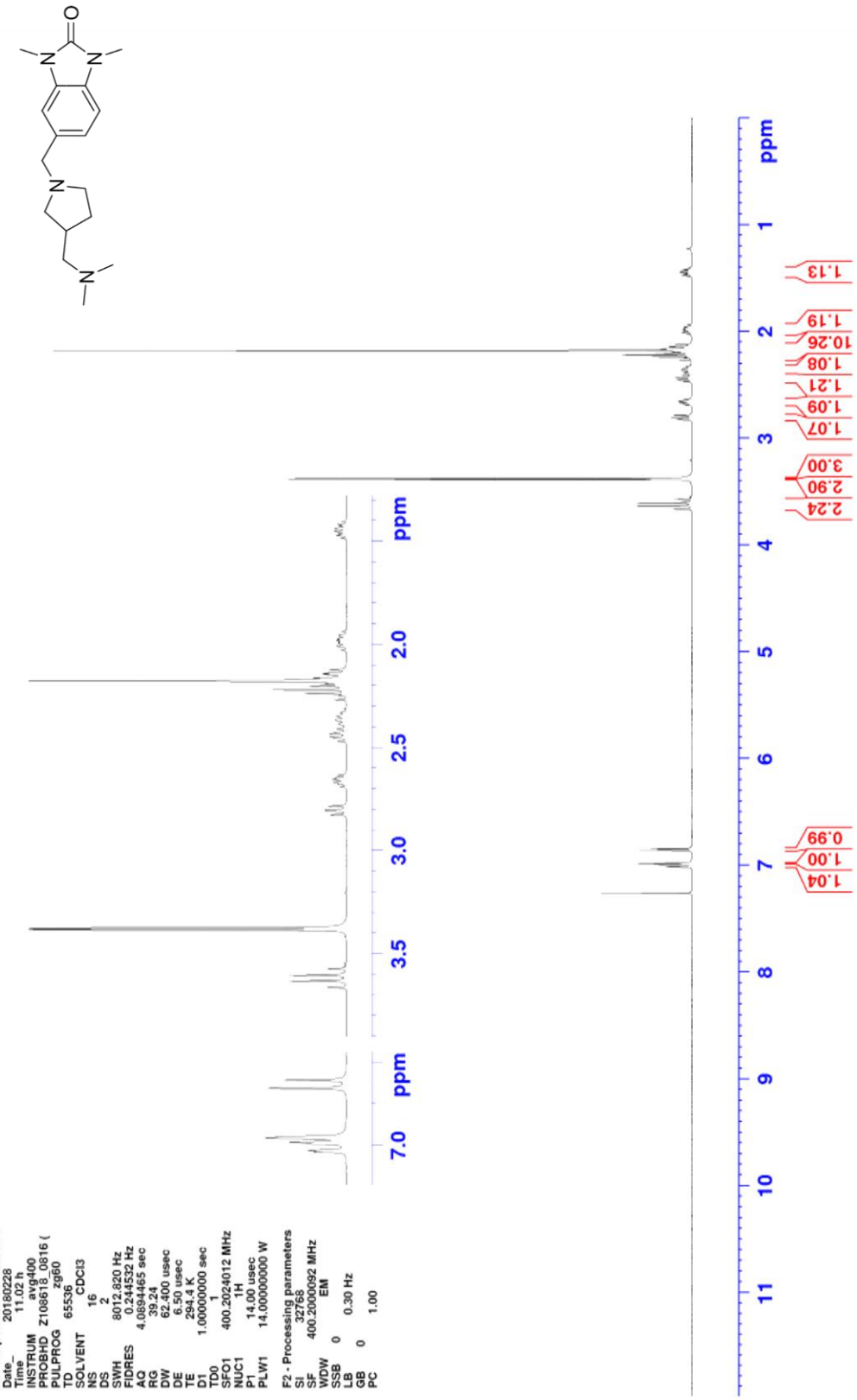
***tert*-Butyl 3-[(dimethylamino)methyl]pyrrolidine-1-carboxylate (75)**



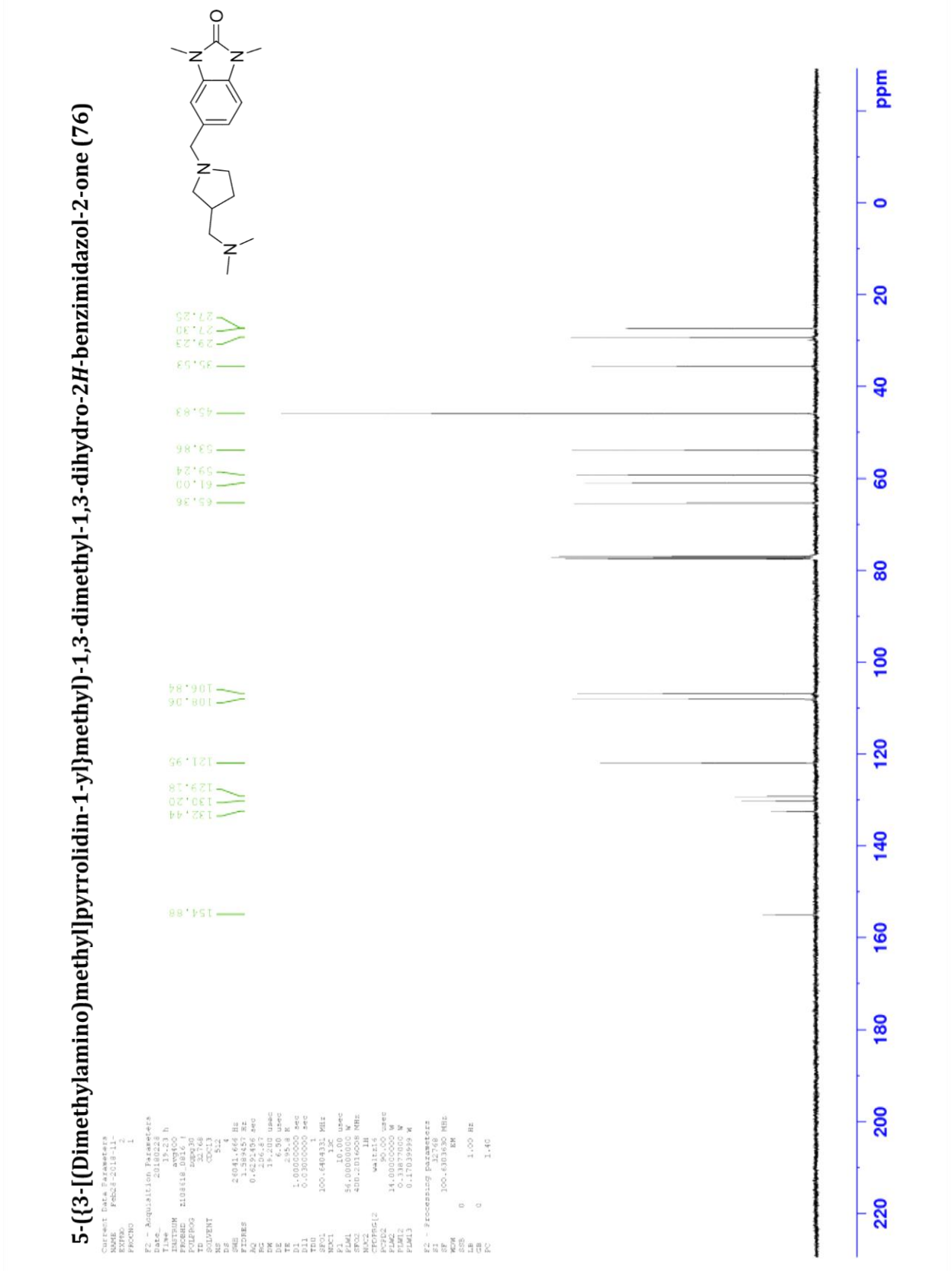
***tert*-Butyl 3-[(dimethylamino)methyl]pyrrolidine-1-carboxylate (75)**



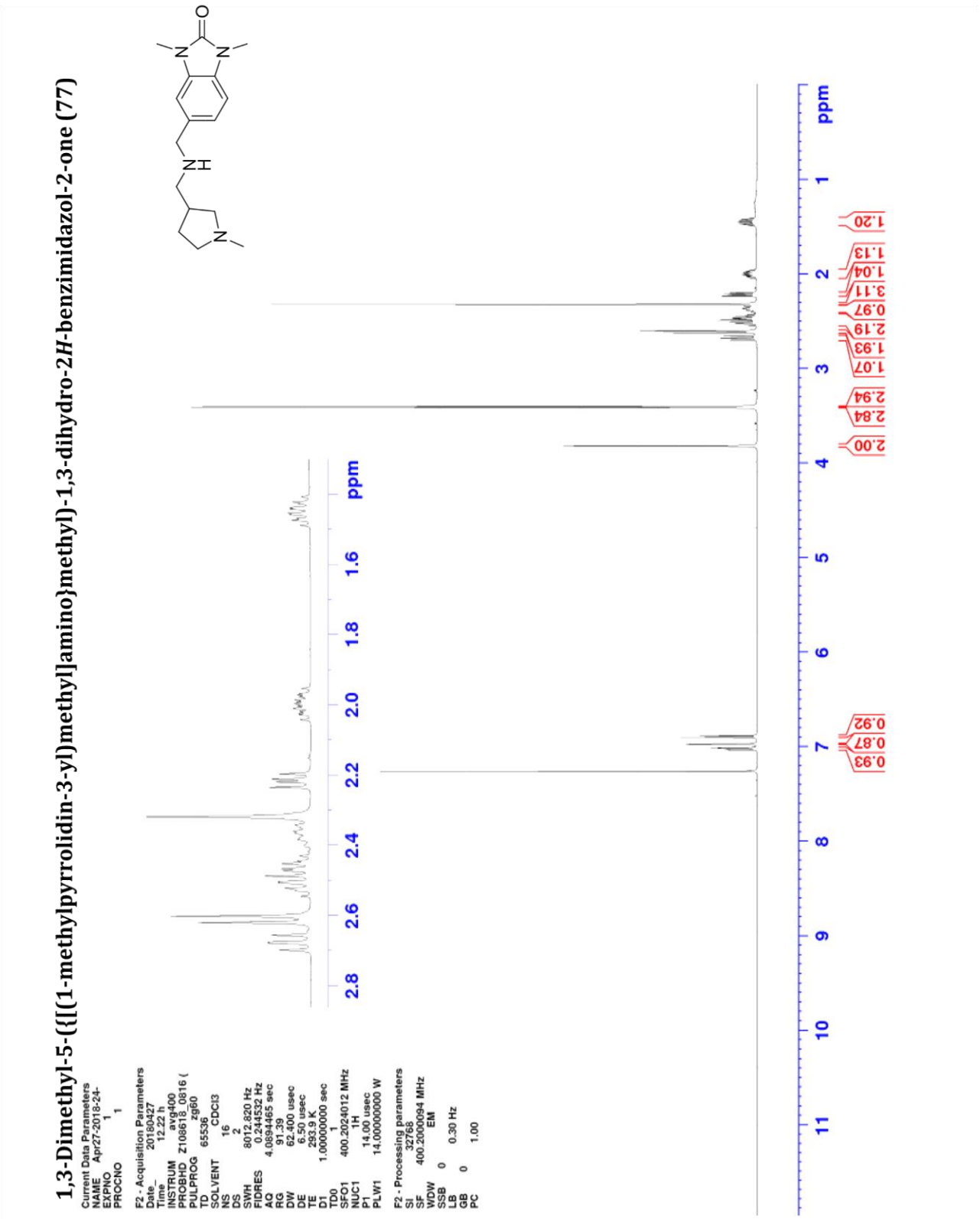
5-({3-[(Dimethylamino)methyl]pyrrolidin-1-yl}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (76)



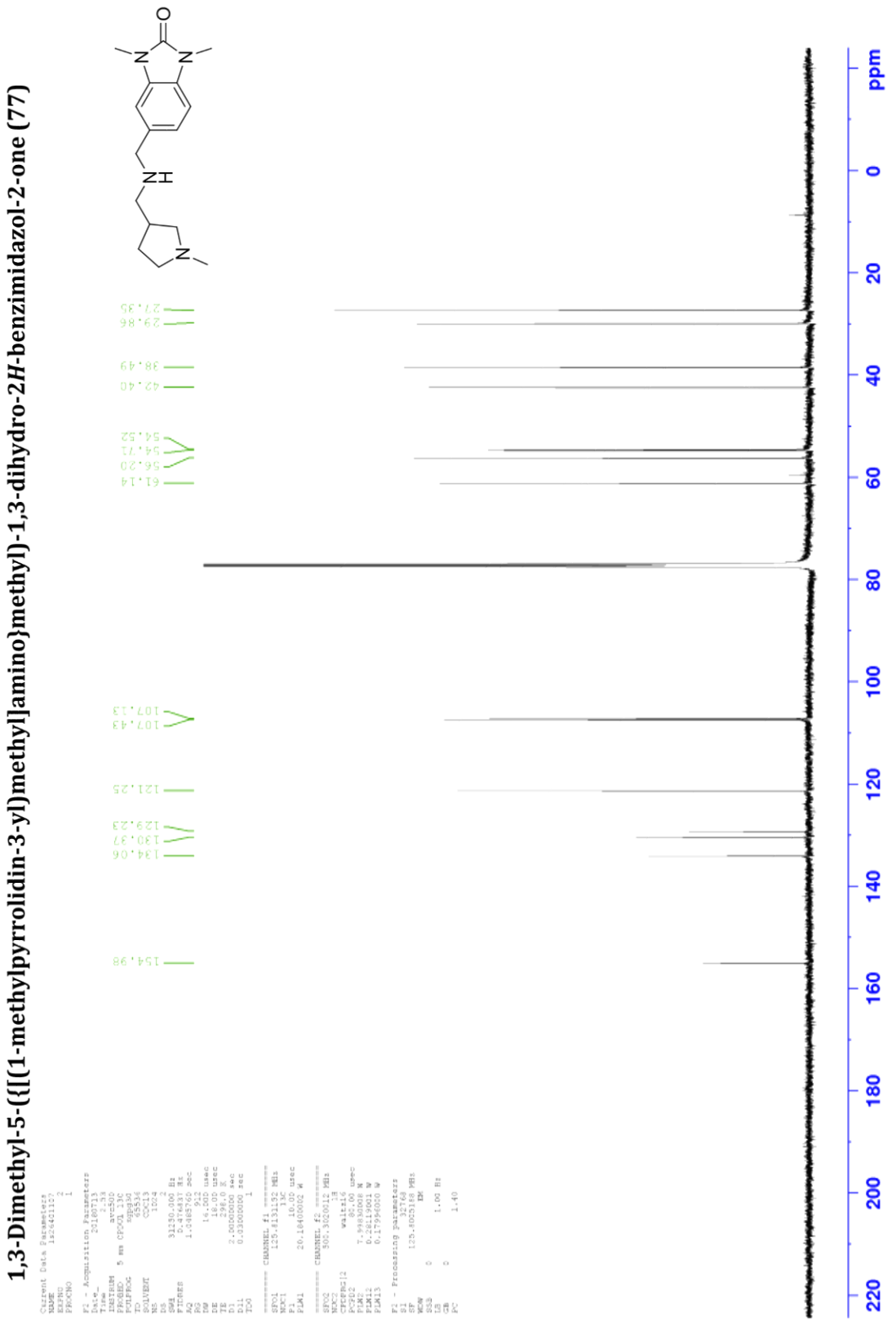
5-({3-[(Dimethylamino)methyl]pyrrolidin-1-yl}methyl)-1,3-dimethyl-1,3-dihydro-2*H*-benzimidazol-2-one (76)



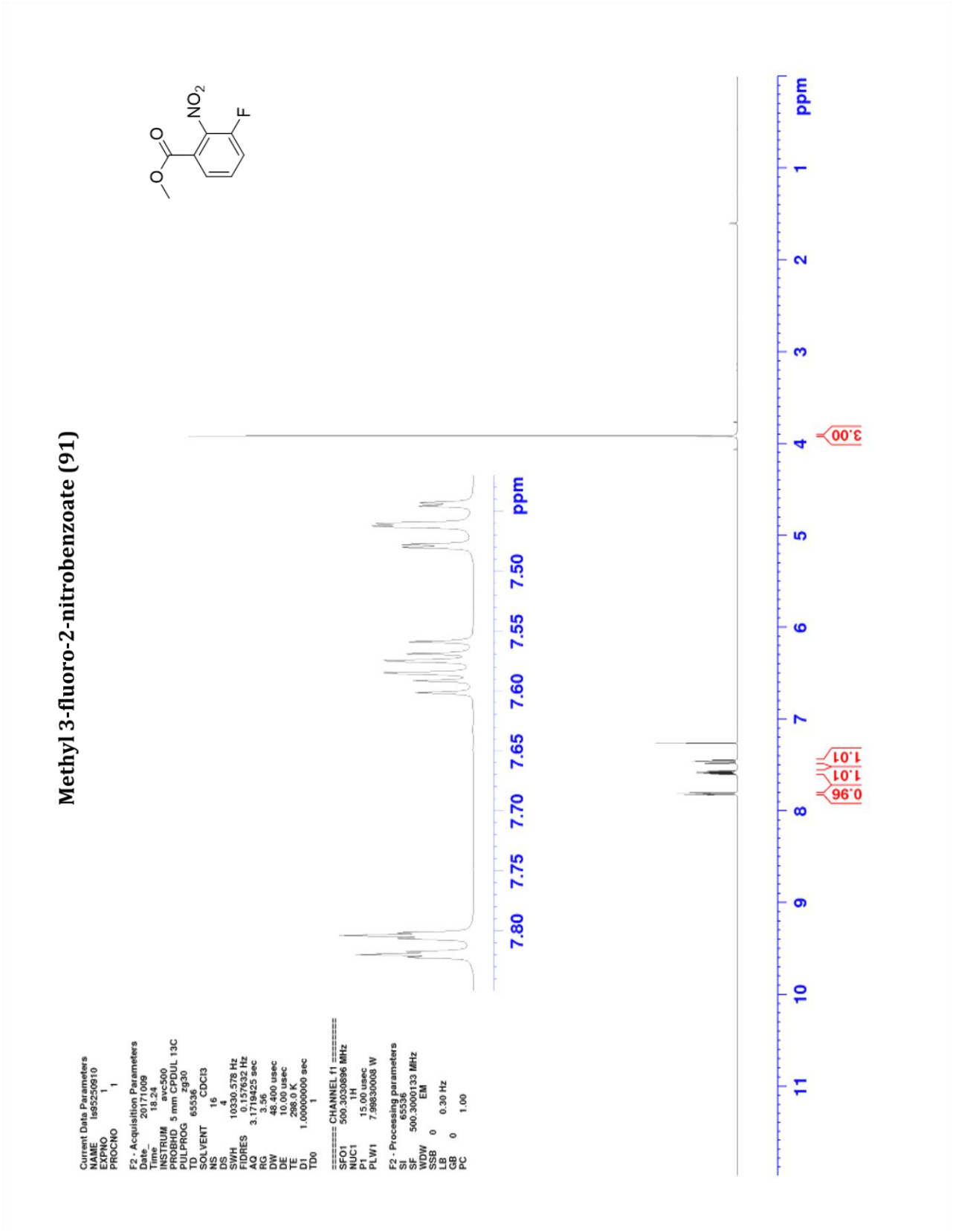
1,3-Dimethyl-5-({[(1-methylpyrrolidin-3-yl)methyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (77)



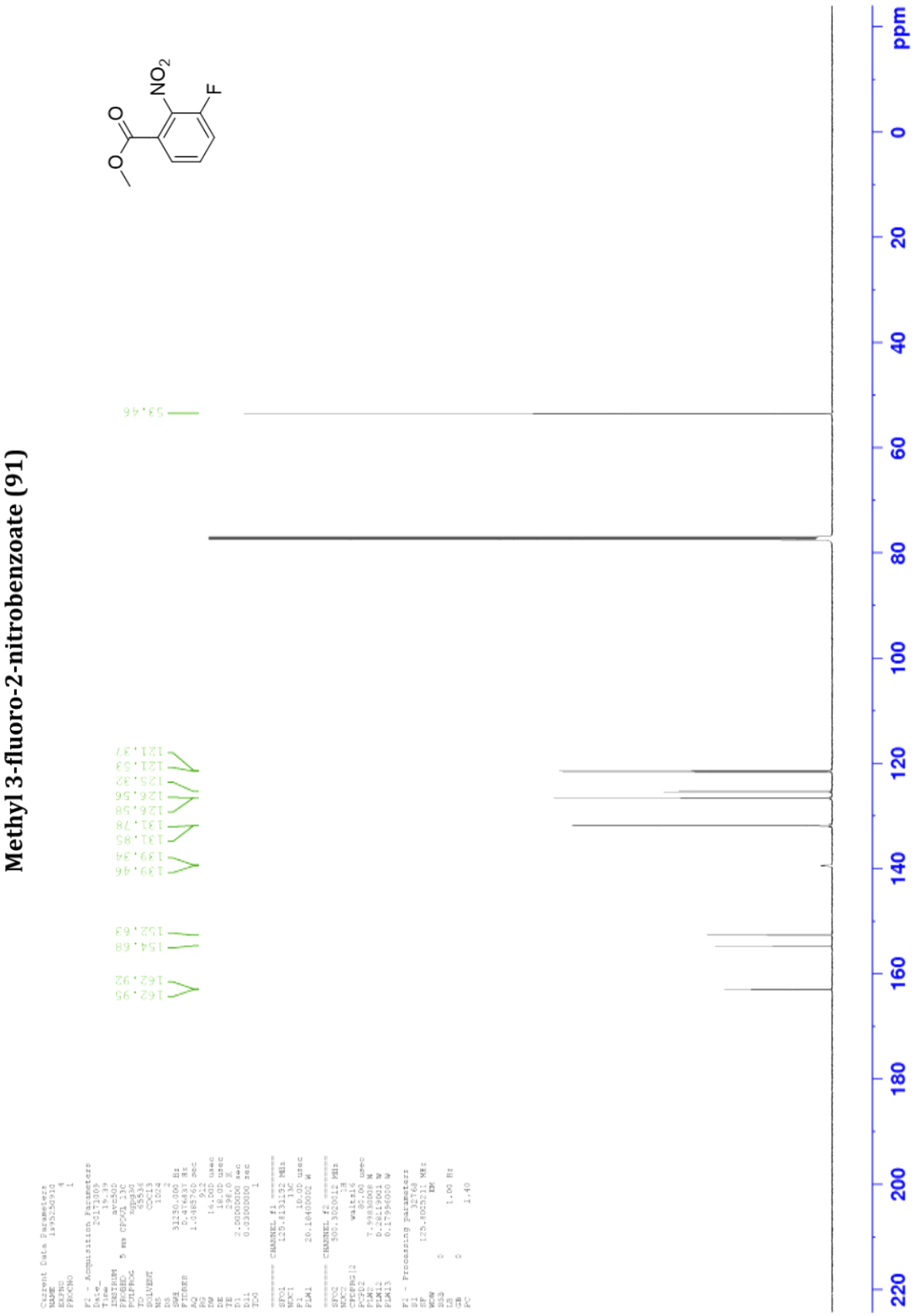
1,3-Dimethyl-5-({[(1-methylpyrrolidin-3-yl)methyl]amino}methyl)-1,3-dihydro-2*H*-benzimidazol-2-one (77)



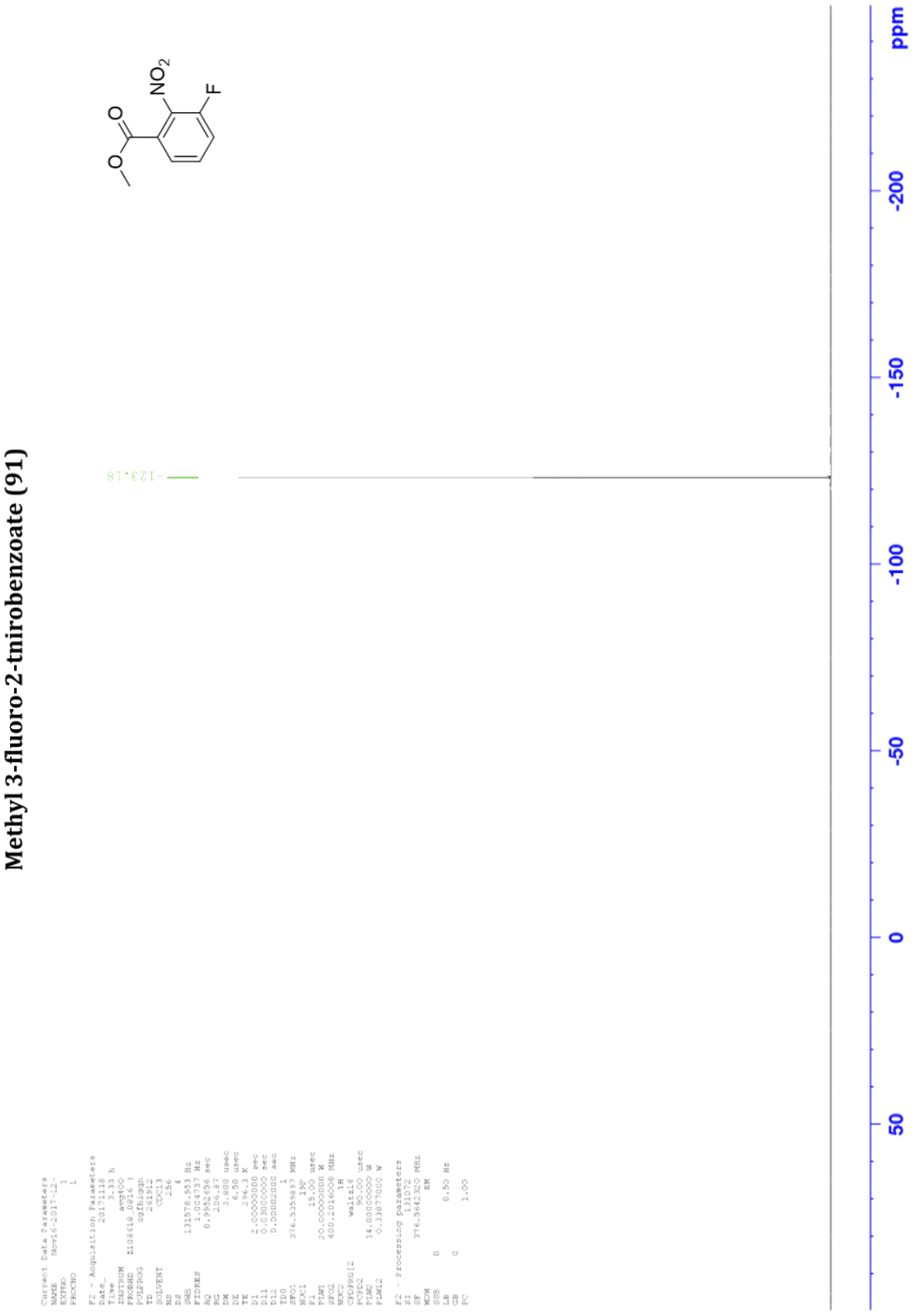
Methyl 3-fluoro-2-nitrobenzoate (91)

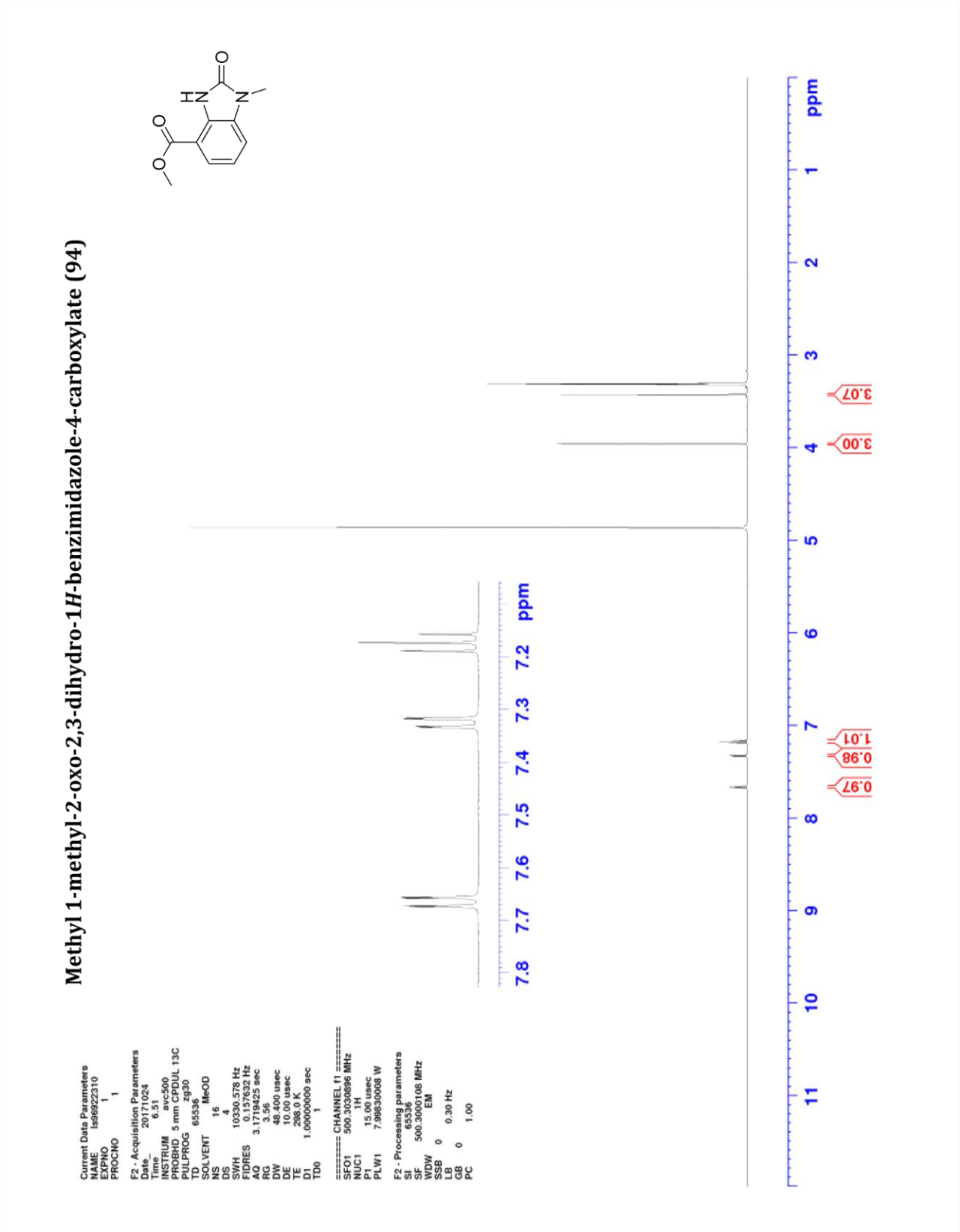


Methyl 3-fluoro-2-nitrobenzoate (91)

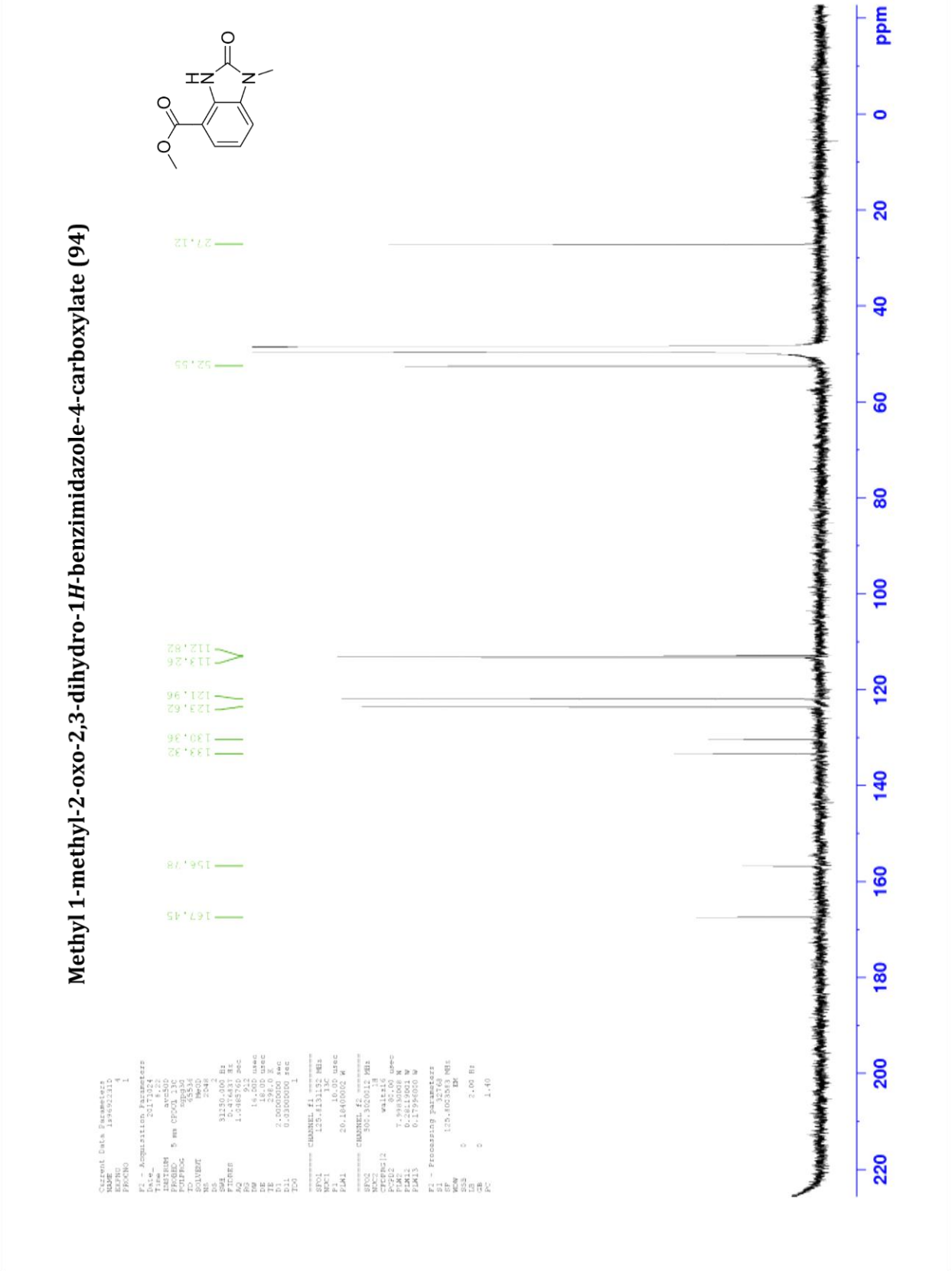


Methyl 3-fluoro-2-tnirobenzoate (91)

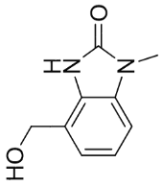



Methyl 1-methyl-2-oxo-2,3-dihydro-1H-benzimidazole-4-carboxylate (94)
Current Data Parameters
NAME ls96922310
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20171024
Time 6.51
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT MeOD
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.1719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000108 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.8 7.7 7.6 7.5 7.4 7.3 7.2 ppm
11 10 9 8 7 6 5 4 3 2 1 ppm
0.97
1.01
0.98
3.00
3.07

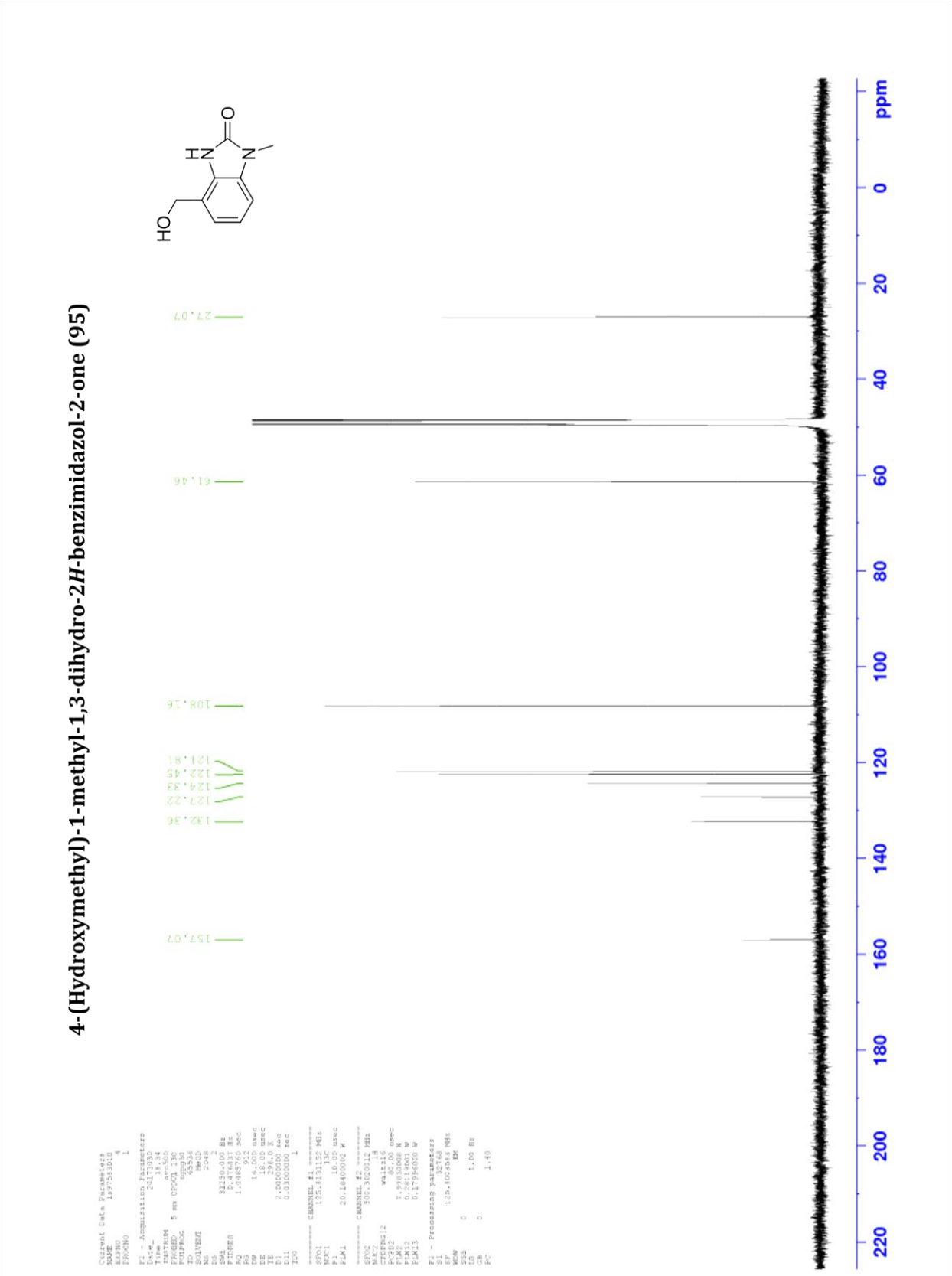
Methyl 1-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-4-carboxylate (94)

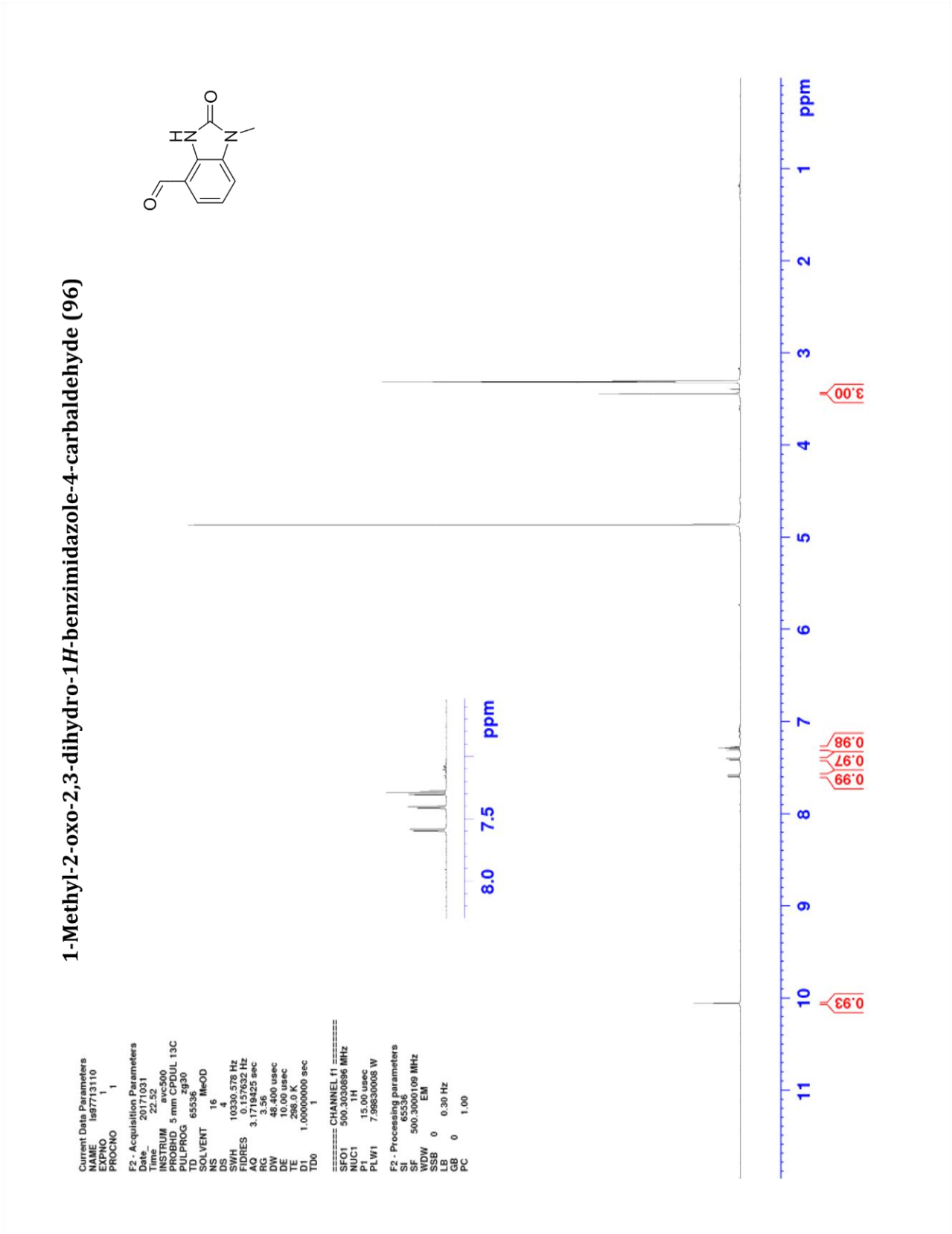


4-(Hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (95)

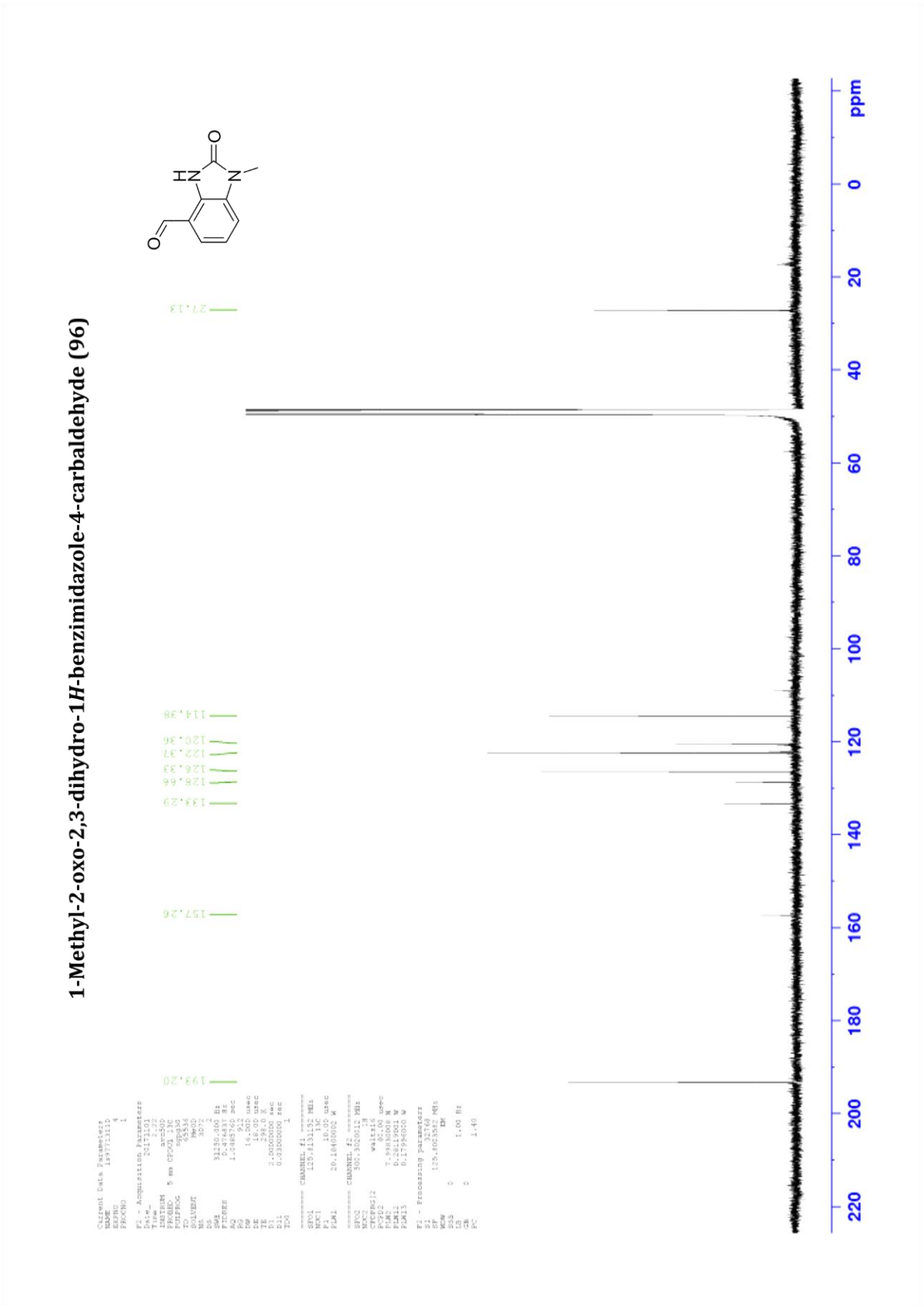


4-(Hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (95)

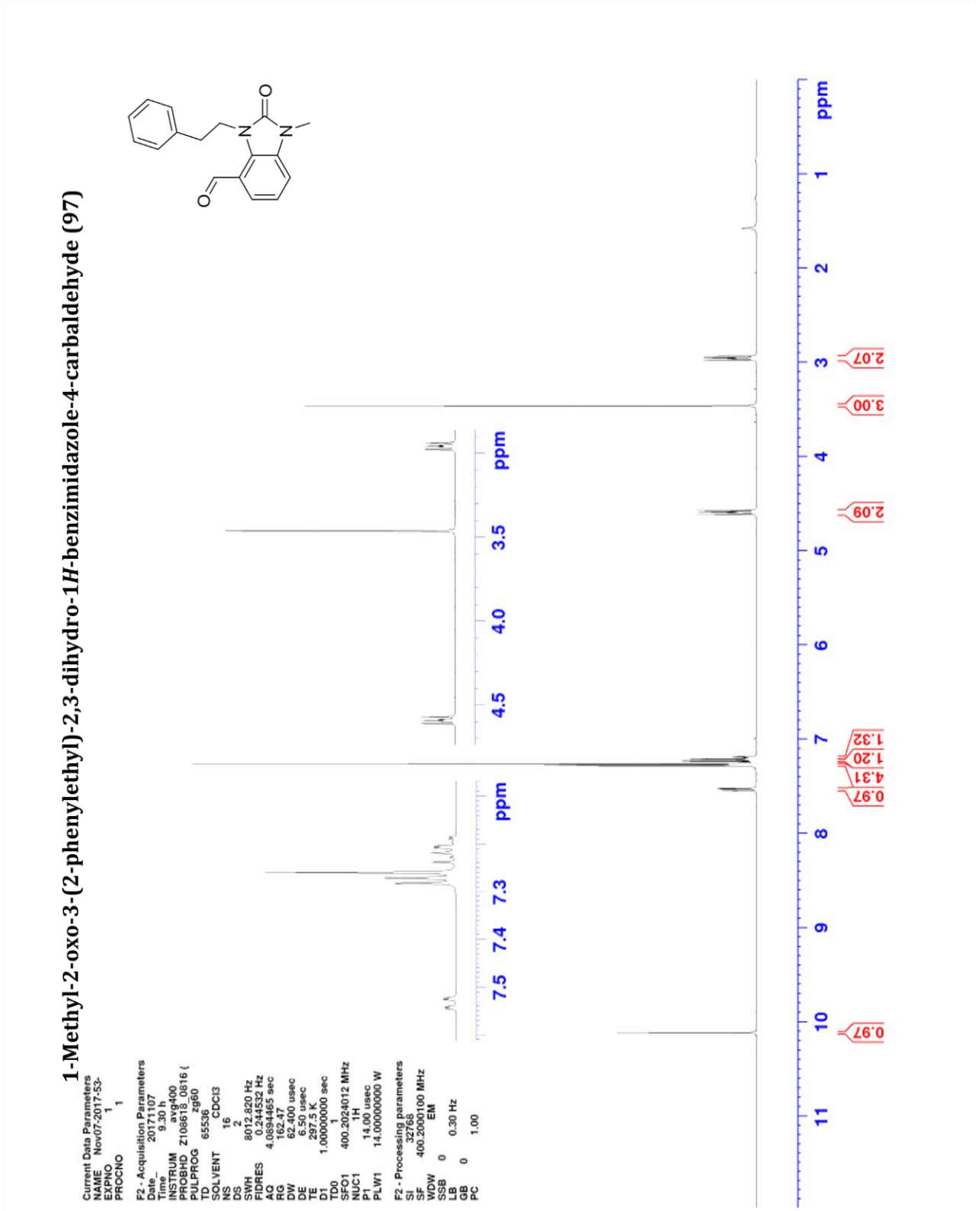



1-Methyl-2-oxo-2,3-dihydro-1H-benzimidazole-4-carbaldehyde (96)
Current Data Parameters
NAME ls9771310
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20171031
Time 22.52
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT MeOD
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17189425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000109 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
3.00
0.99
0.97
0.98
0.93
ppm
8.0
7.5

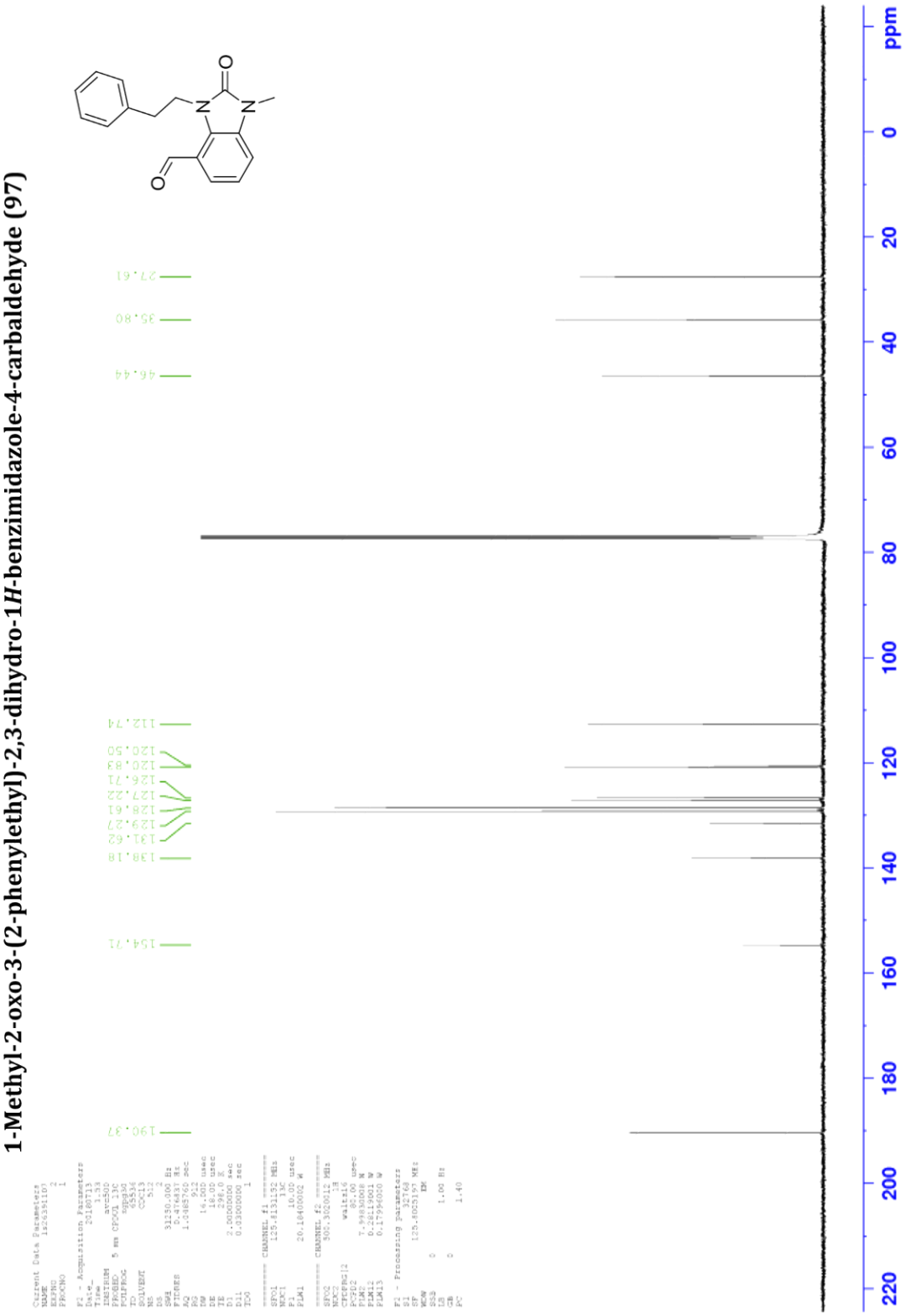
1-Methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-4-carbaldehyde (96)



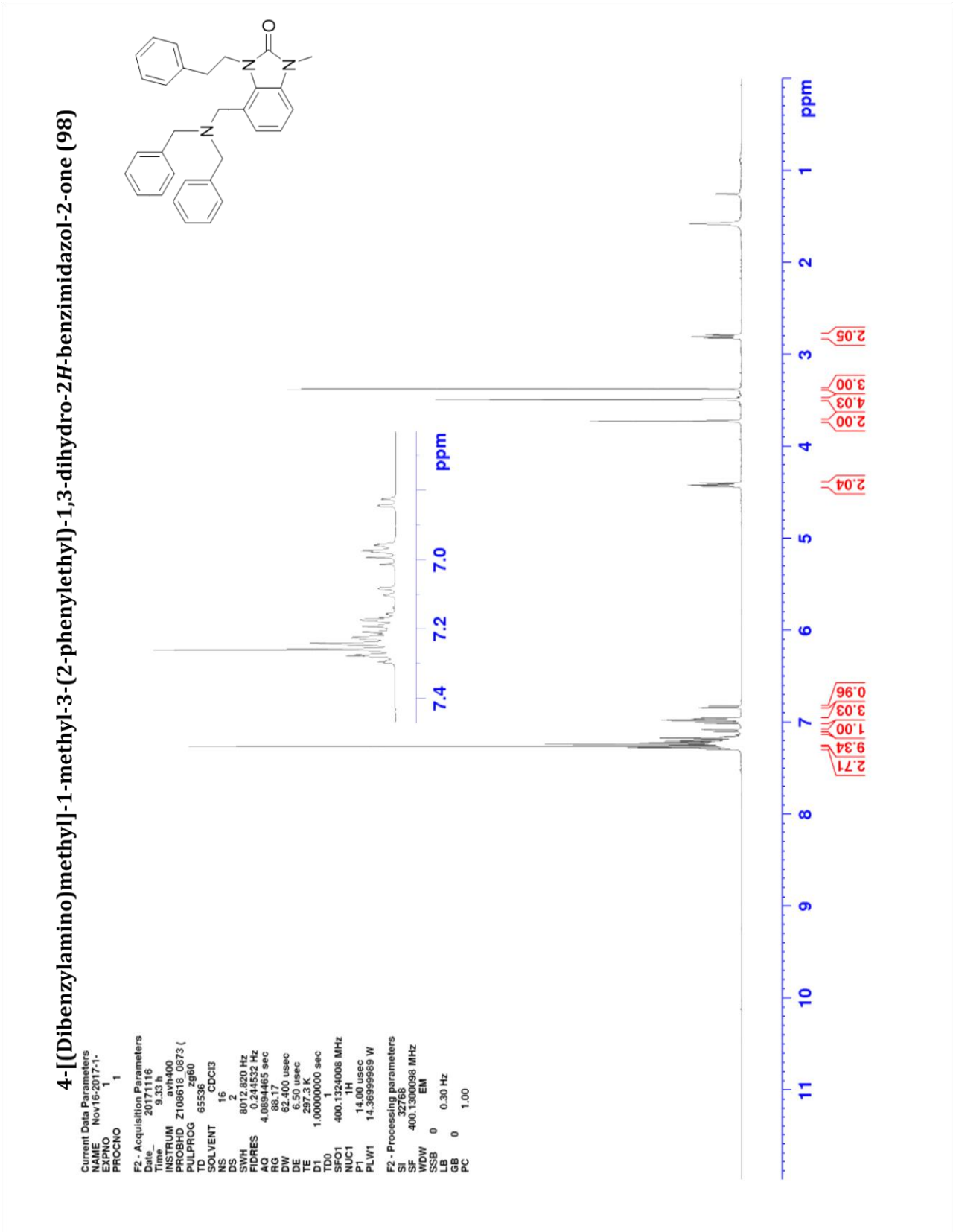
1-Methyl-2-oxo-3-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-4-carbaldehyde (97)



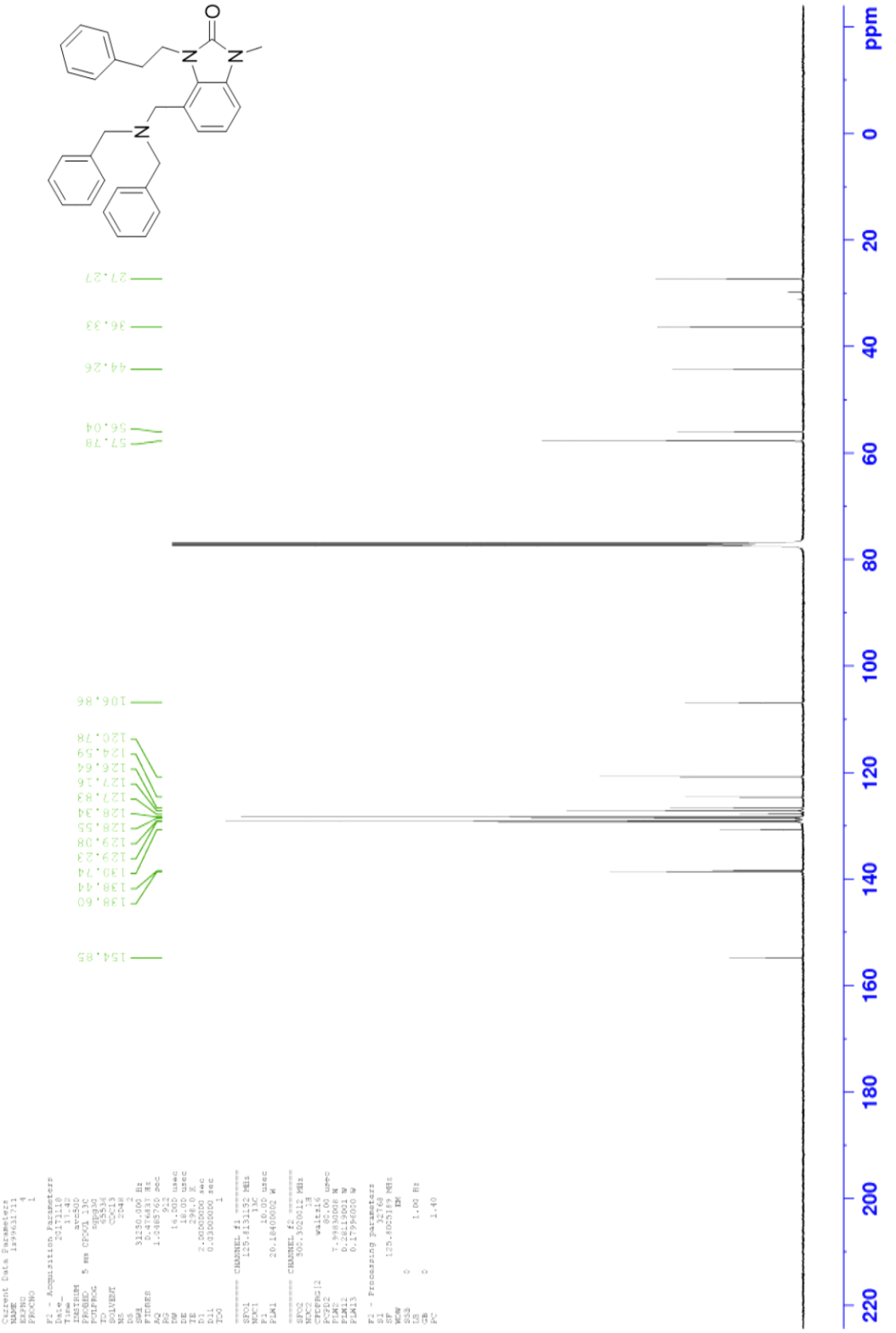
1-Methyl-2-oxo-3-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-4-carbaldehyde (97)



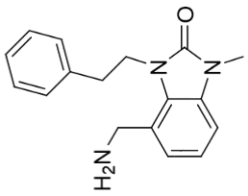
4-[(Dibenzylamino)methyl]-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (98)



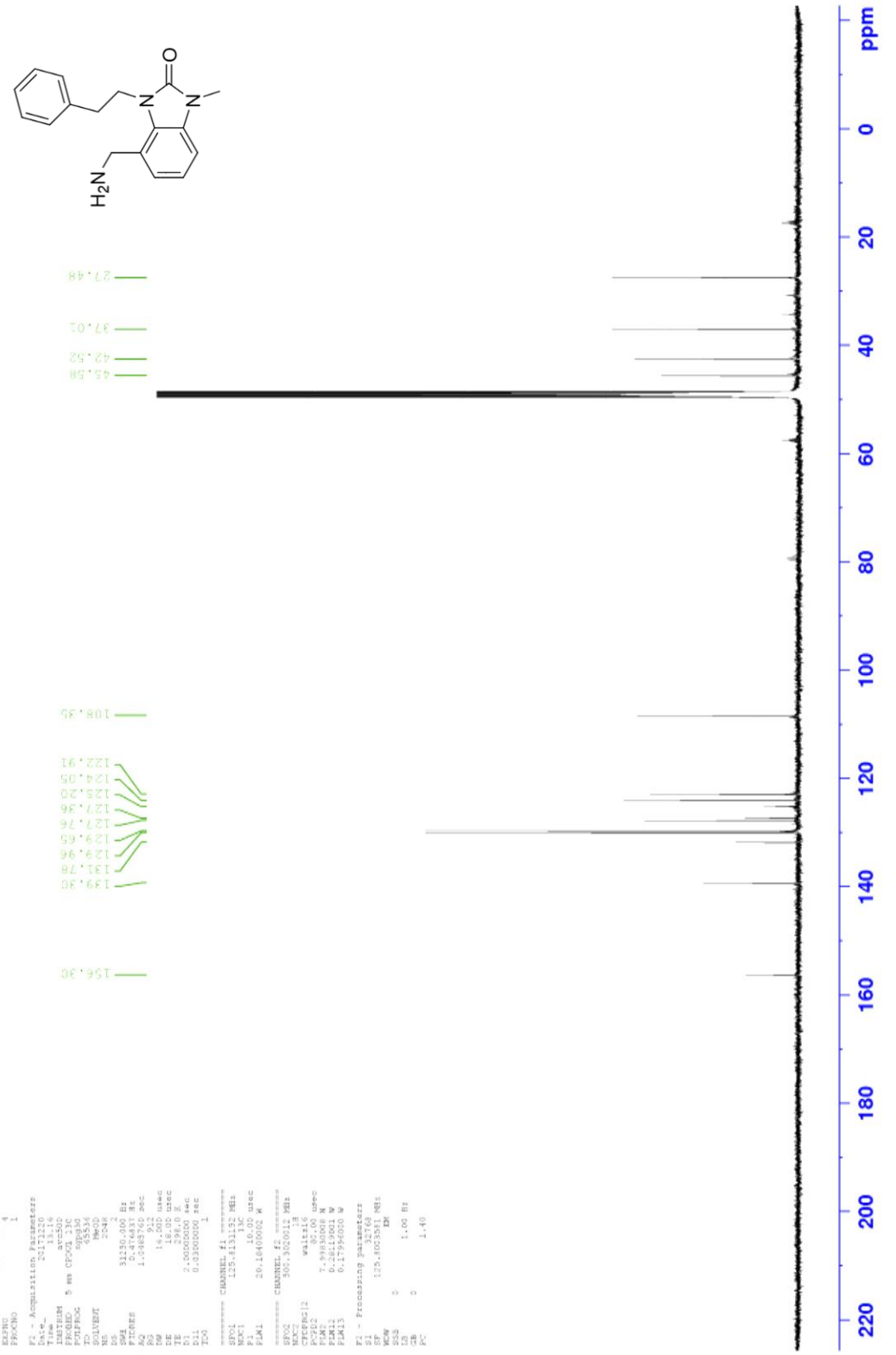
4-[(Dibenzylamino)methyl]-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (98)

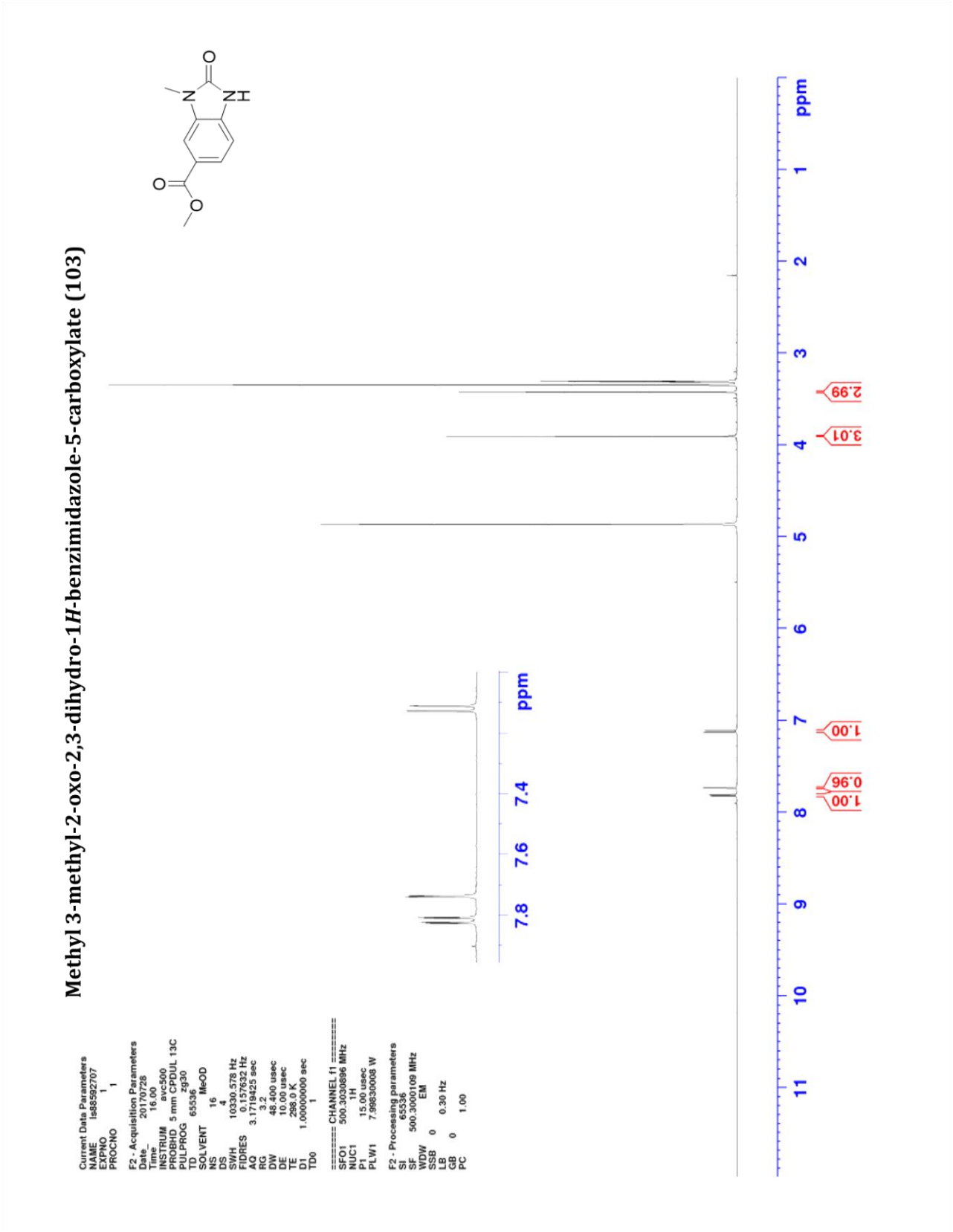


4-(Aminomethyl)-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (99)

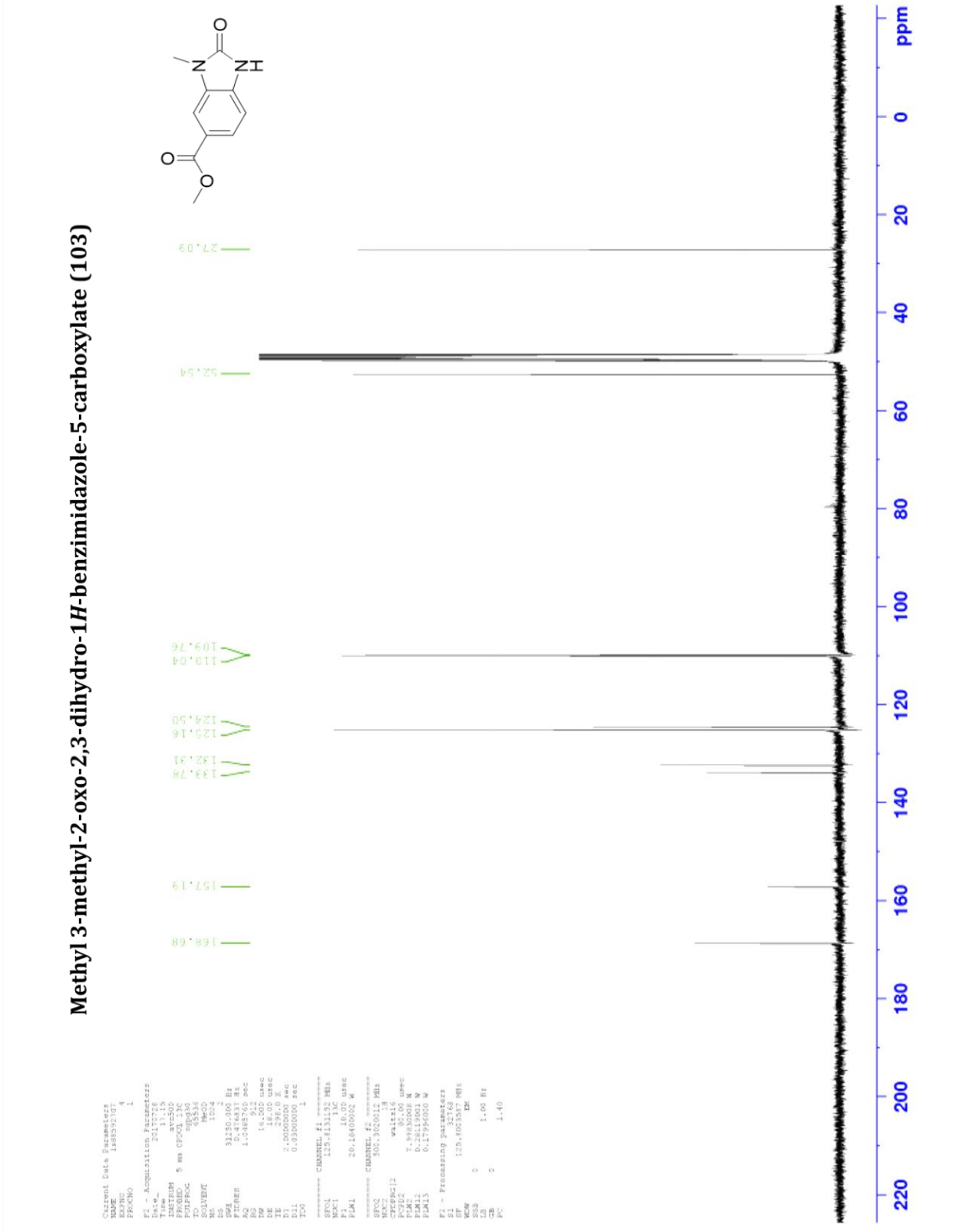


4-(Aminomethyl)-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (99)

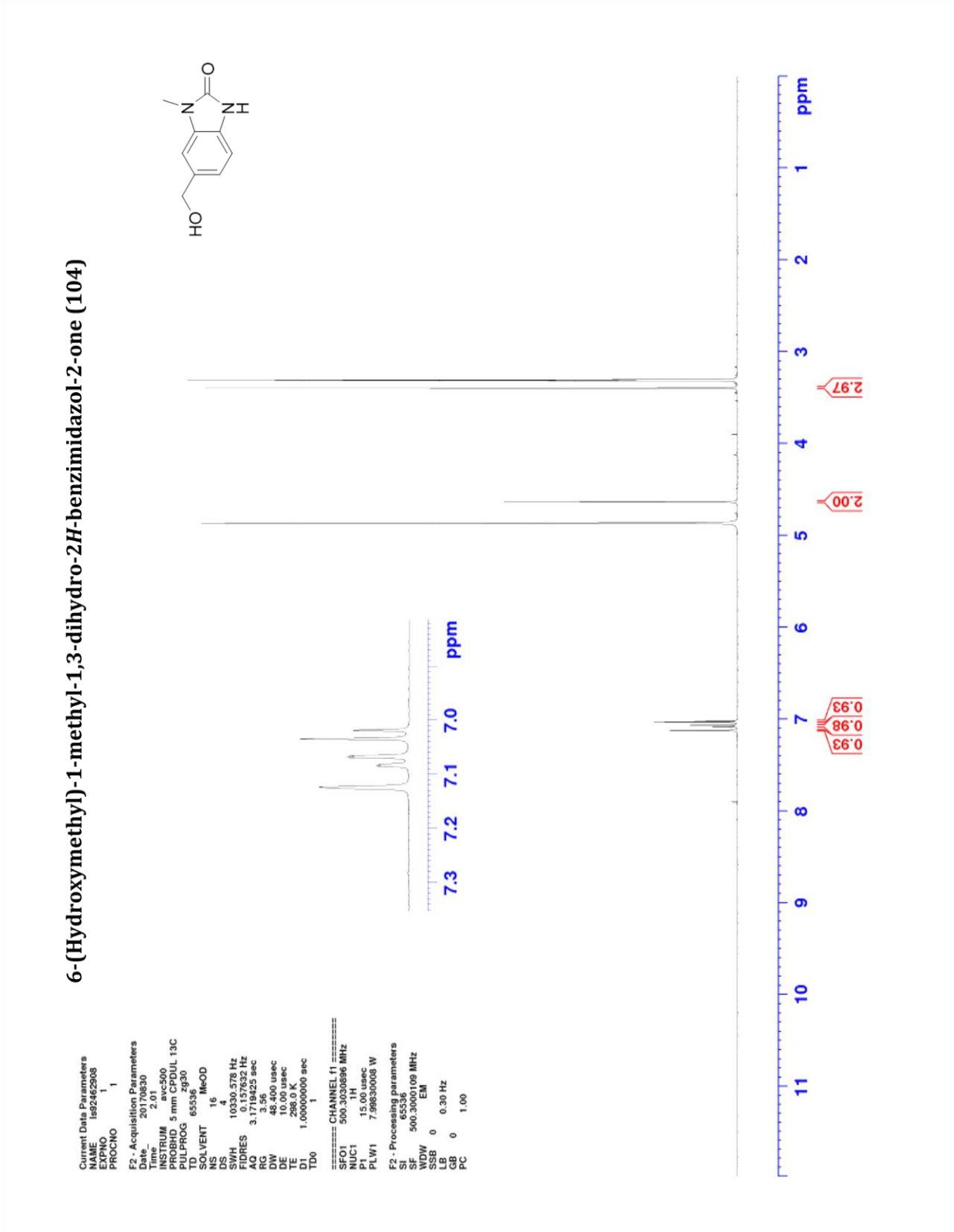



Methyl 3-methyl-2-oxo-2,3-dihydro-1H-benzimidazole-5-carboxylate (103)
Current Data Parameters
NAME ls88592707
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20170728
Time 16.00
INSTRUM av-500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT MeOD
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.1719425 sec
RG 3.2
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000109 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.8
7.6
7.4
ppm
11
10
9
8
7
6
5
4
3
2
1
ppm
1.00
0.96
1.00
2.99
3.01

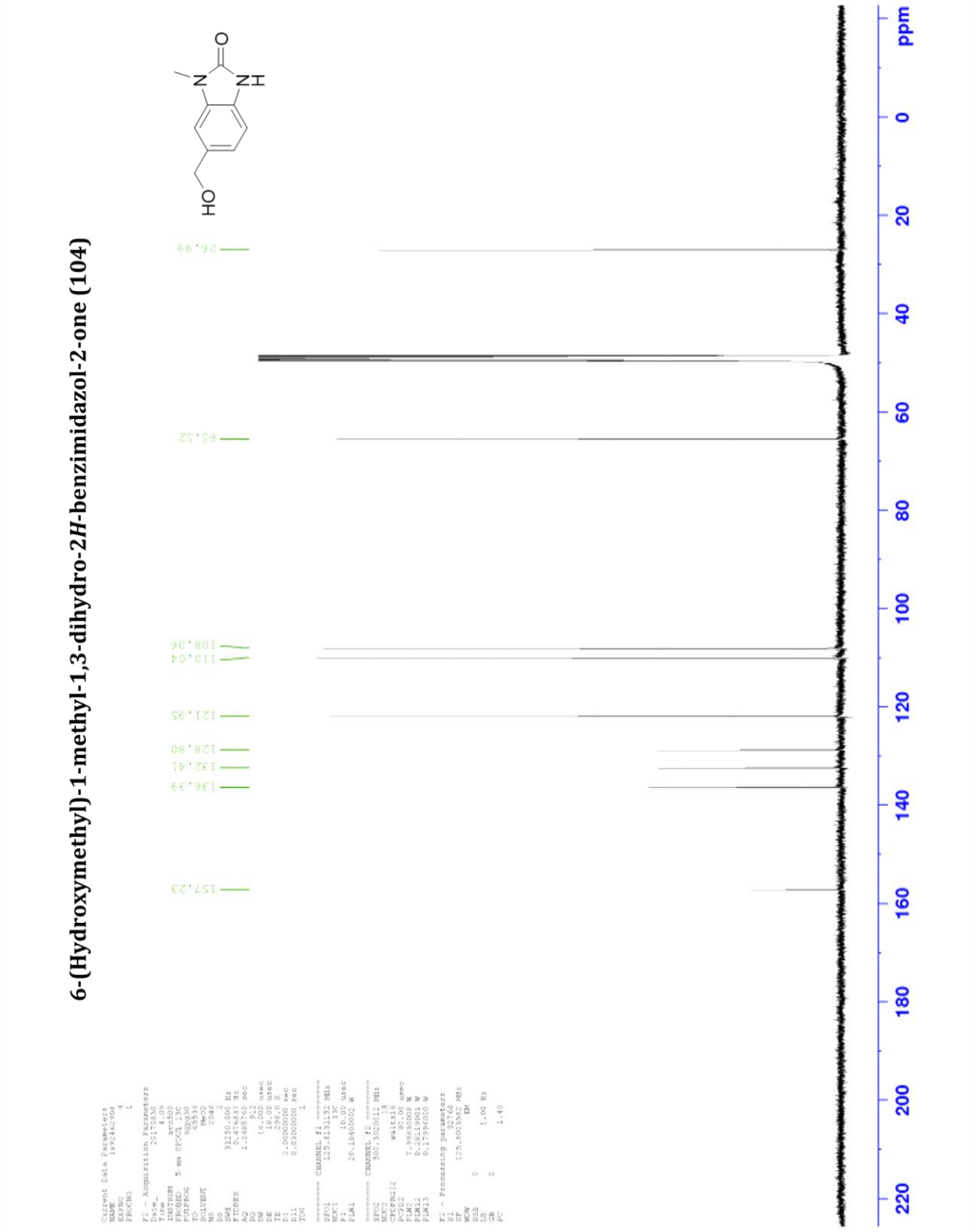
Methyl 3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (103)

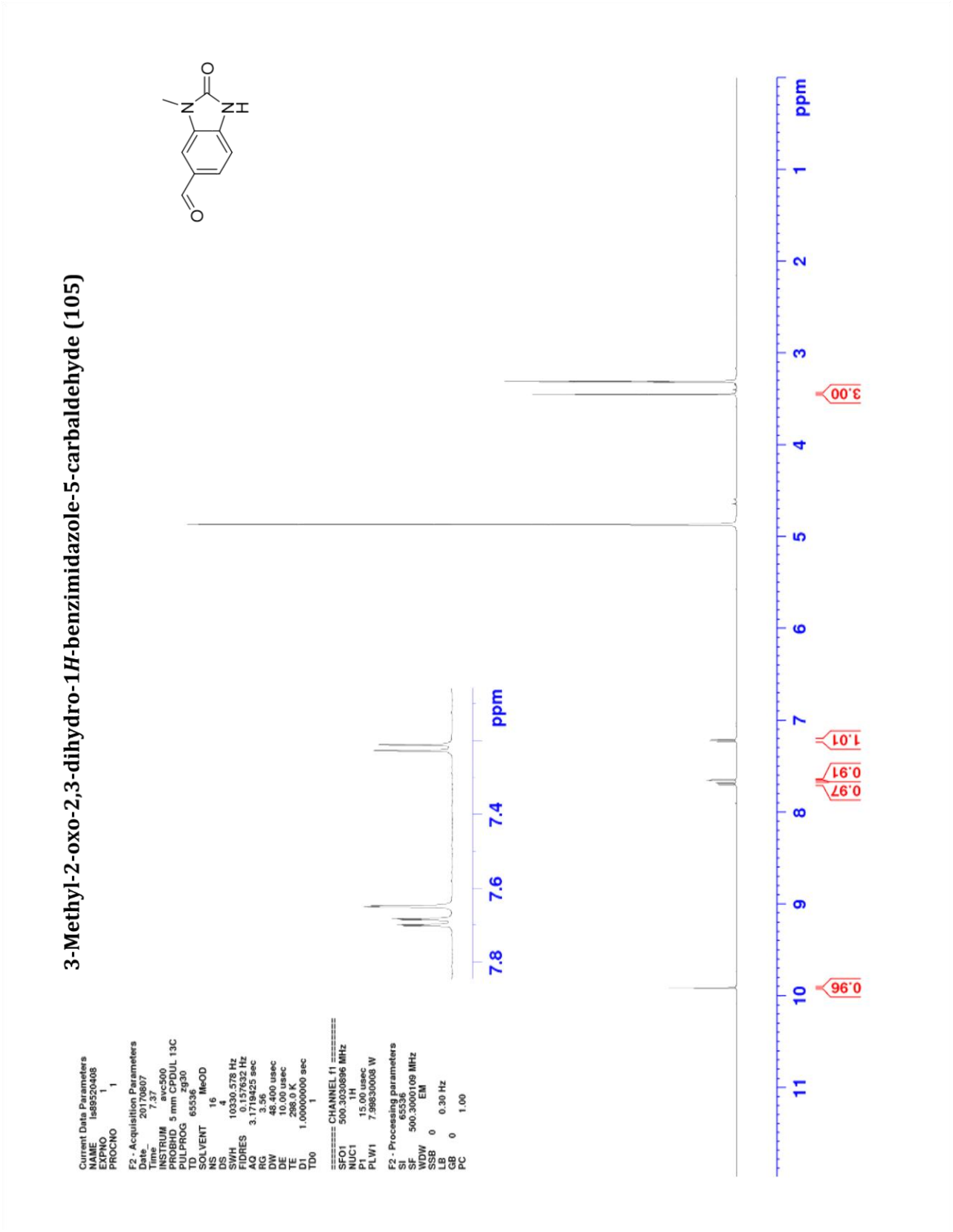


6-(Hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (104)

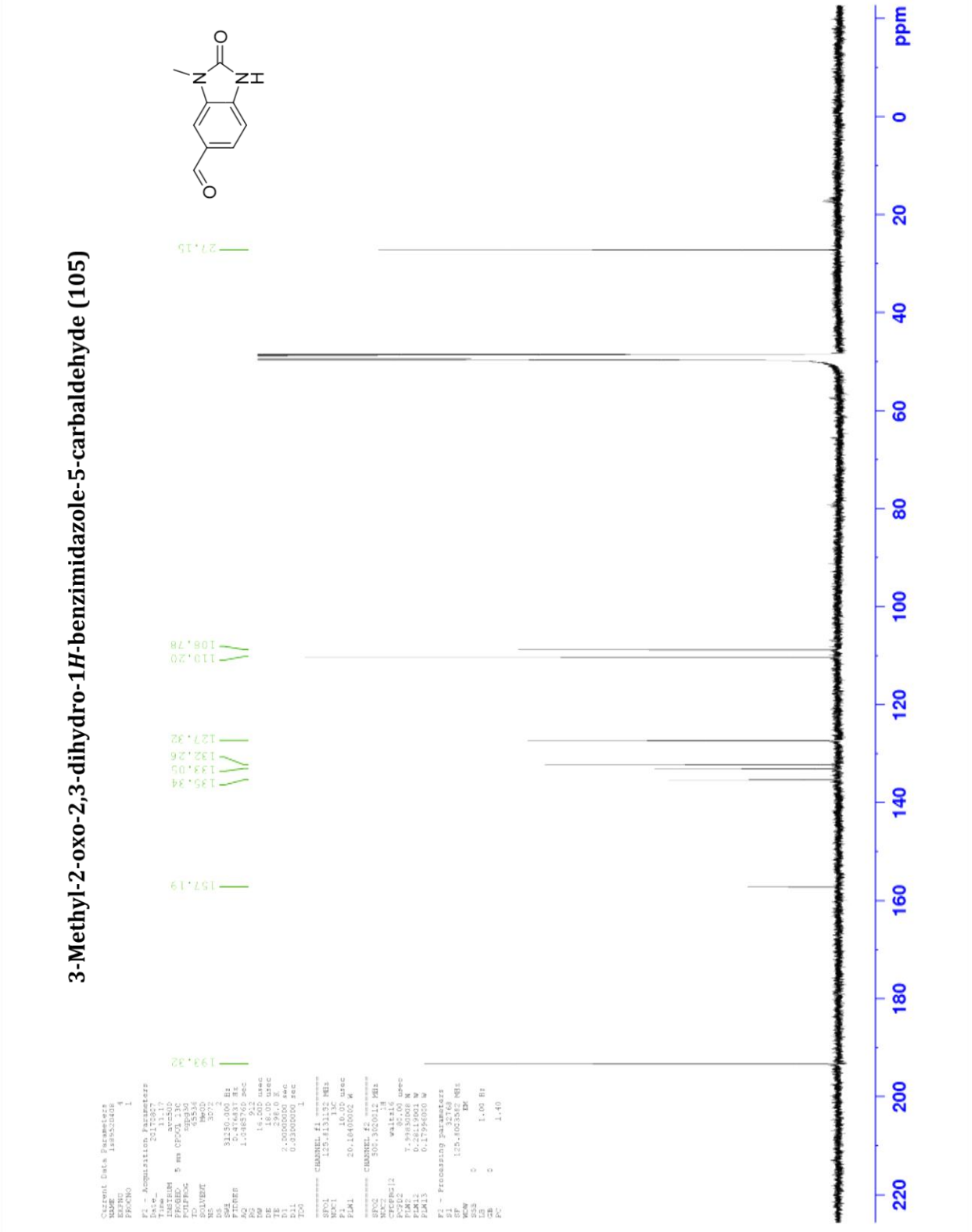


6-(Hydroxymethyl)-1-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (104)

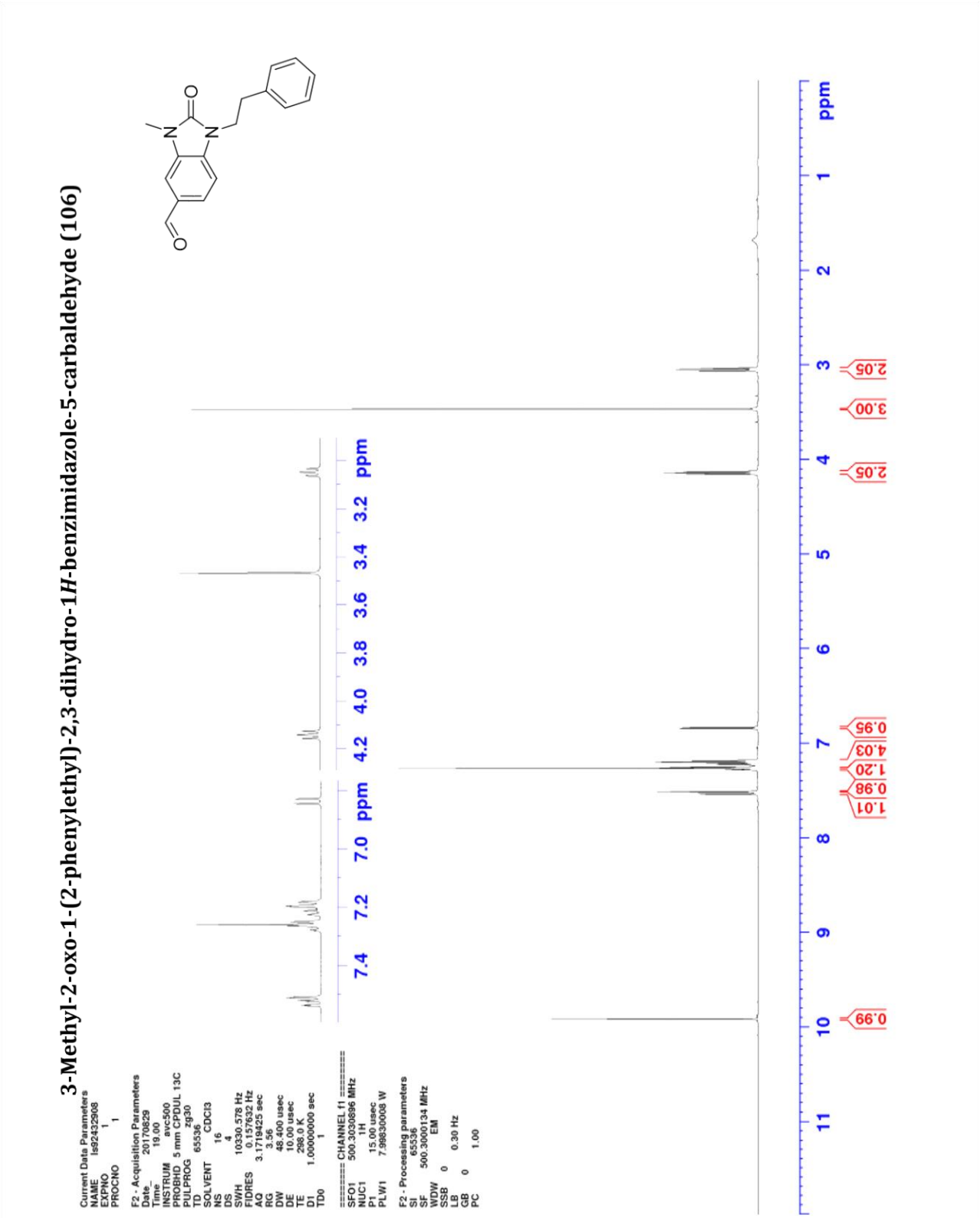



3-Methyl-2-oxo-2,3-dihydro-1H-benzimidazole-5-carbaldehyde (105)
Current Data Parameters
NAME ls89520408
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20170607
Time 7.37
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT MeOD
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000109 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
O
N
NH
O
7.8
7.6
7.4
ppm
11
10
9
8
7
6
5
4
3
2
1
ppm
0.96
0.97
0.91
1.01
3.00

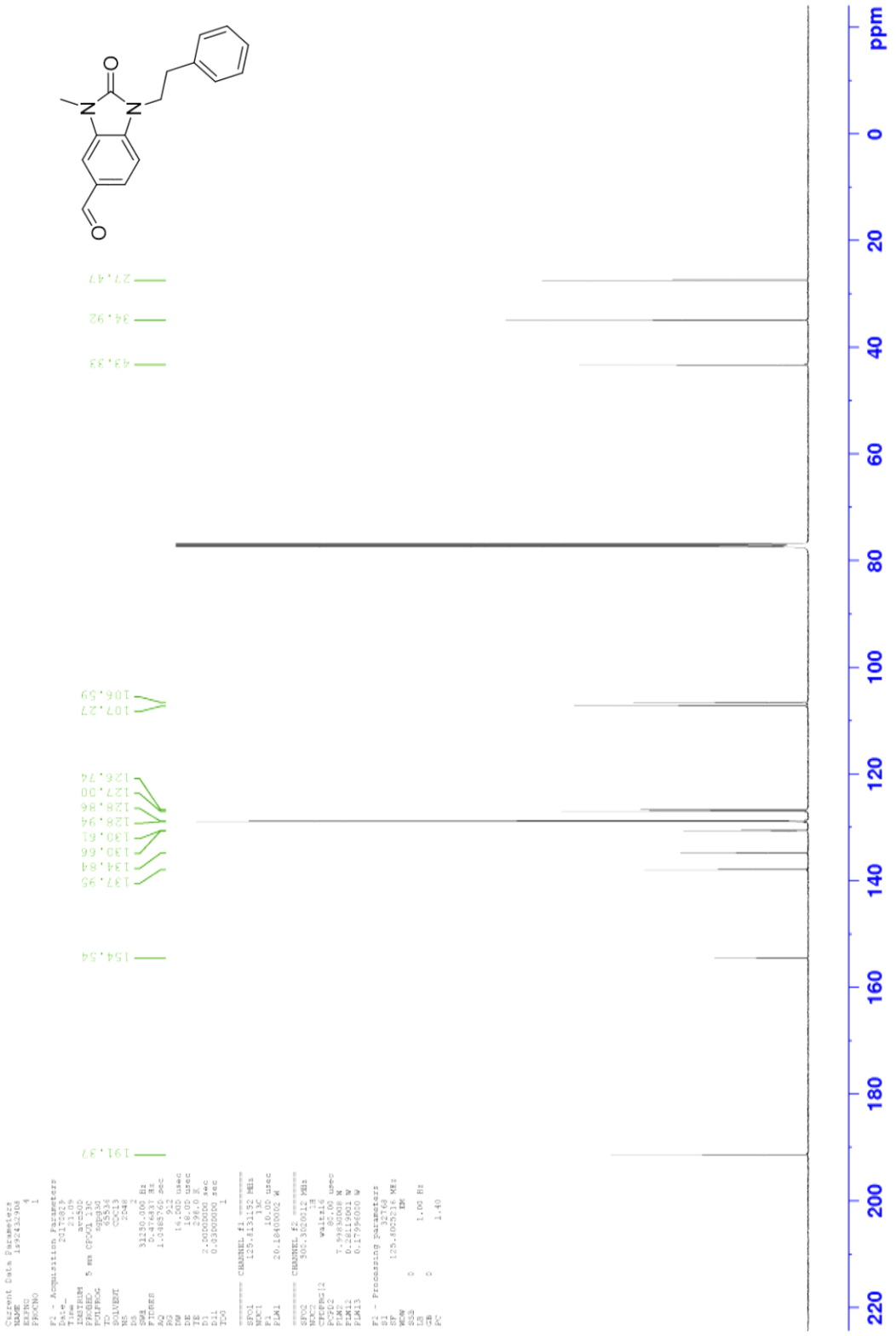
3-Methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (105)



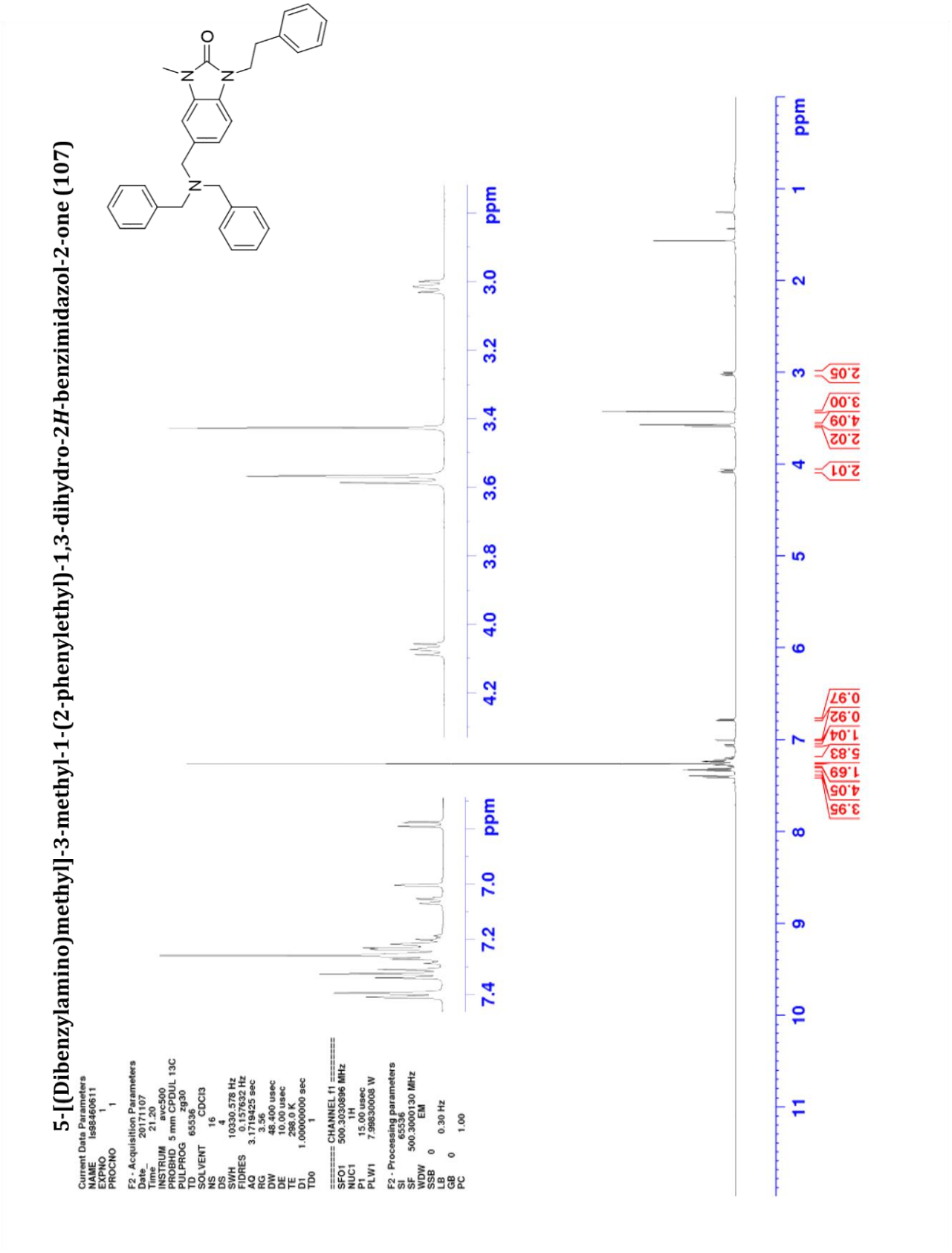
3-Methyl-2-oxo-1-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (106)



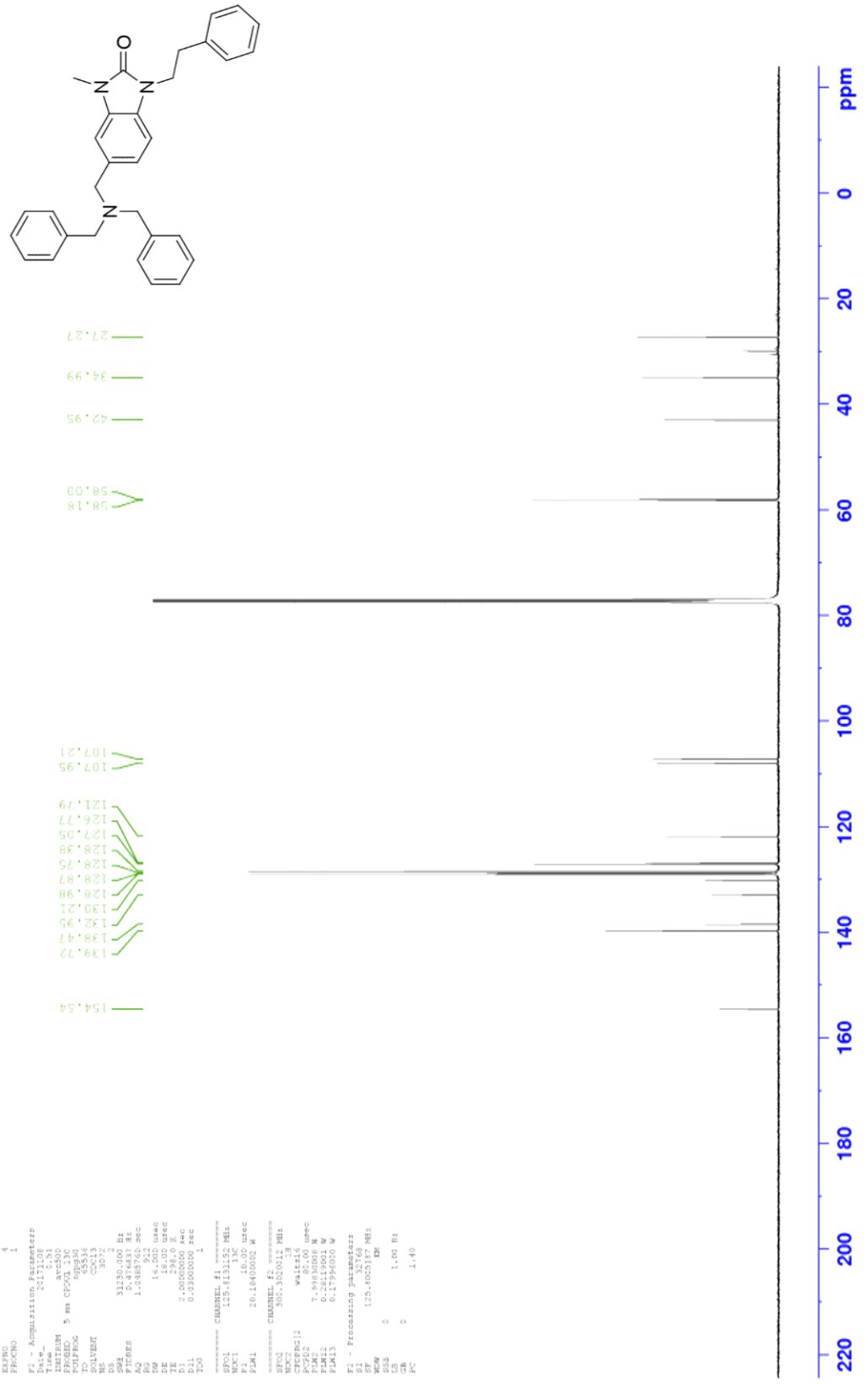
3-Methyl-2-oxo-1-(2-phenylethyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (106)



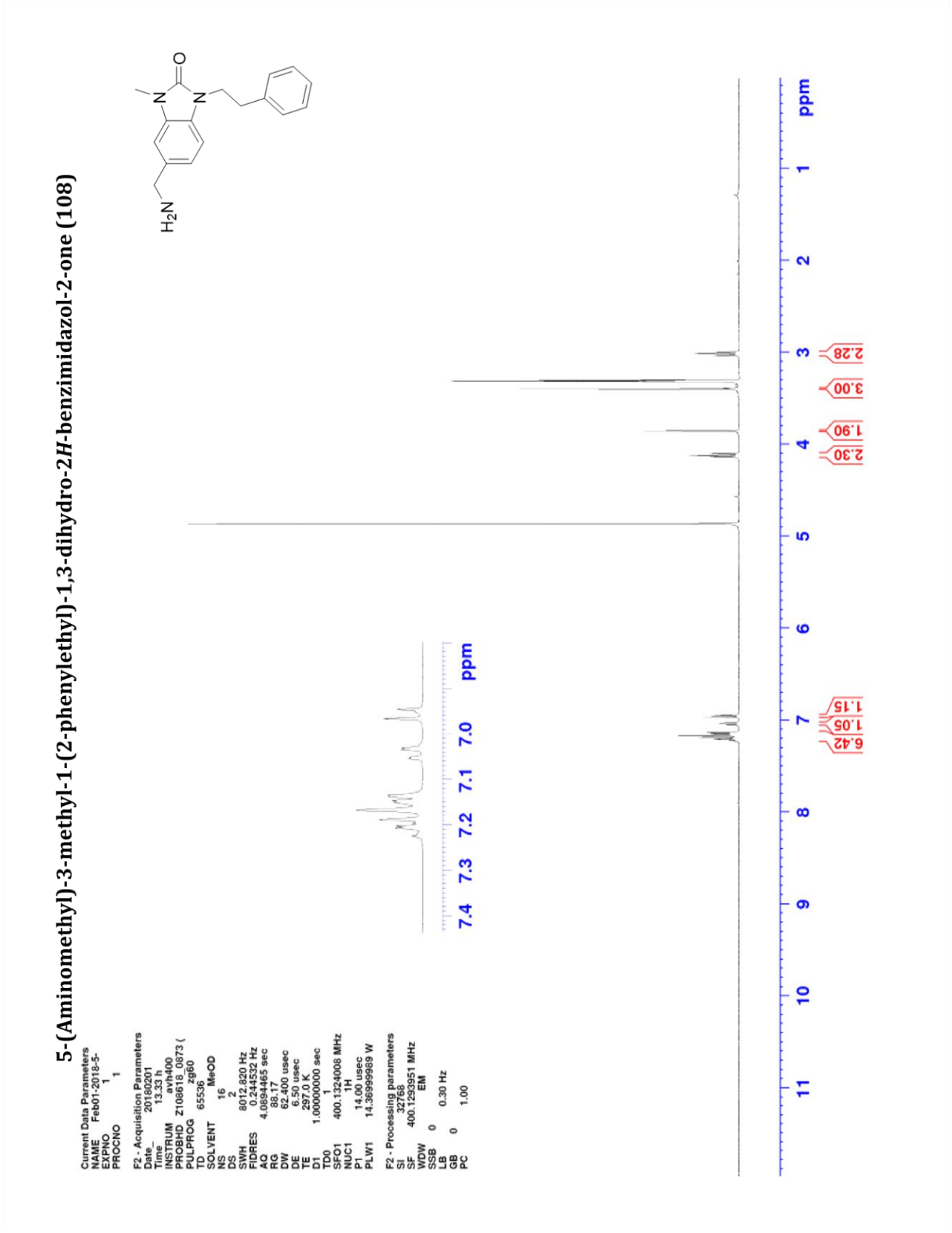
5-[(Dibenzylamino)methyl]-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (107)



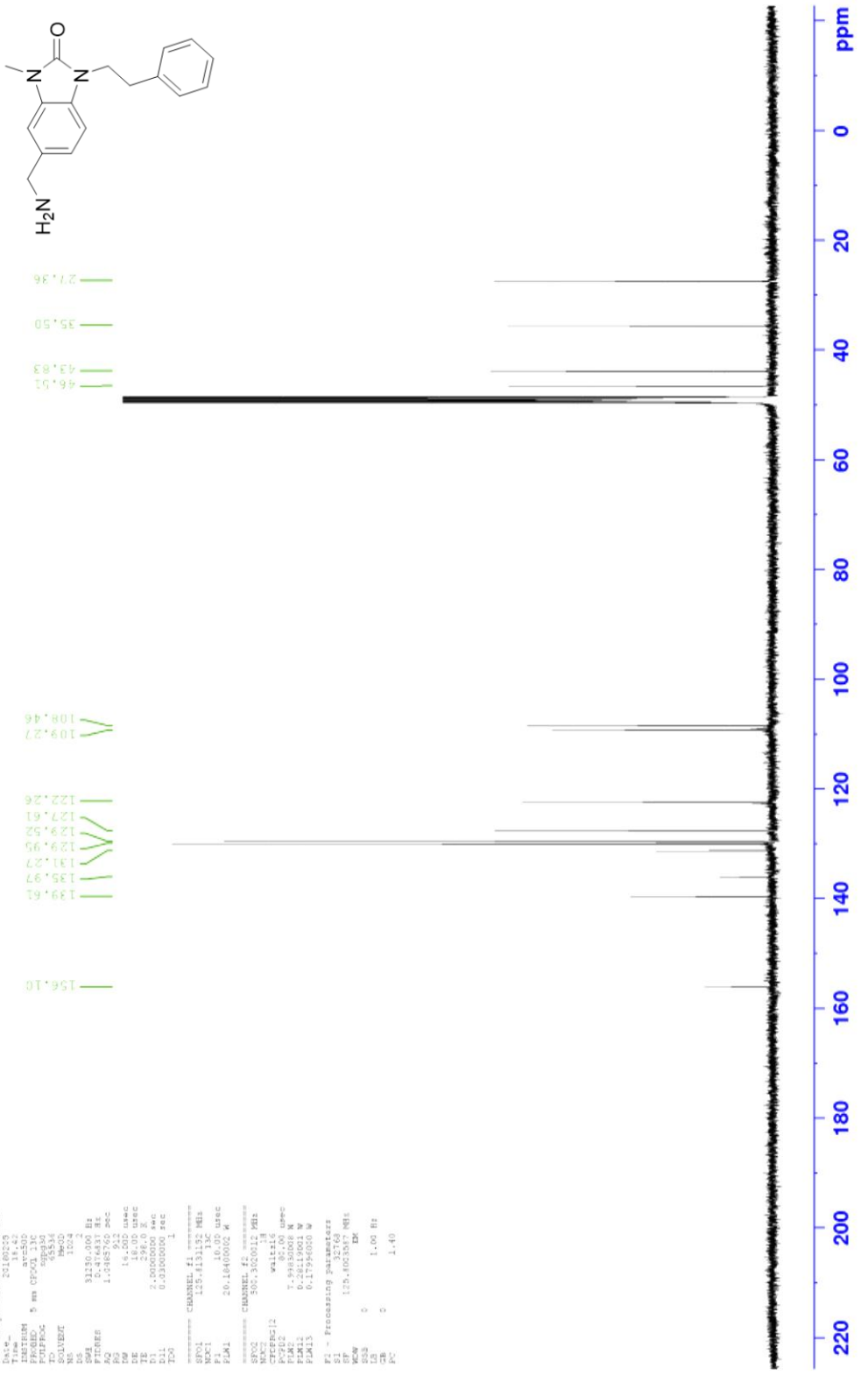
5-[(Dibenzylamino)methyl]-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (107)



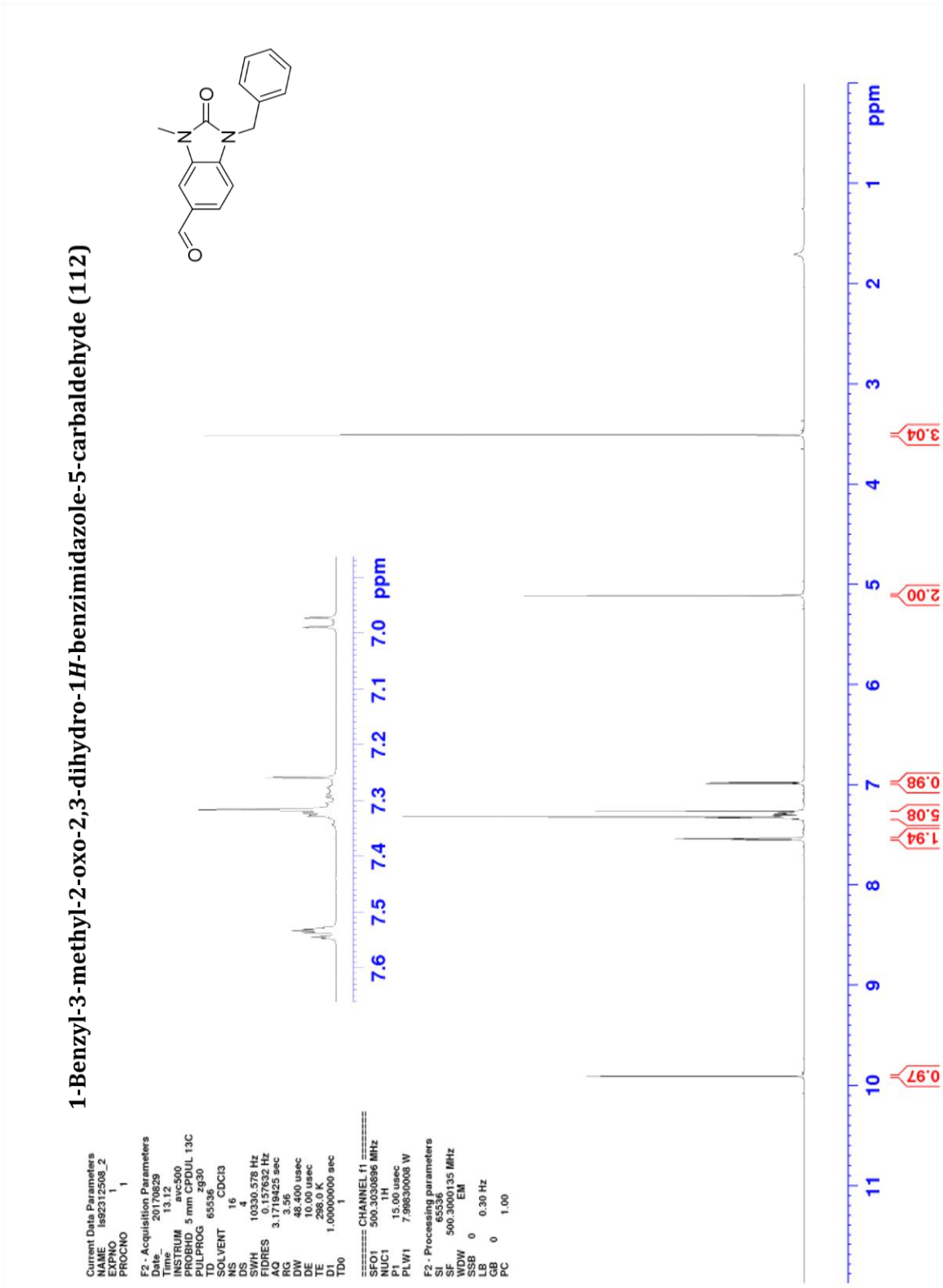
5-(Aminomethyl)-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (108)



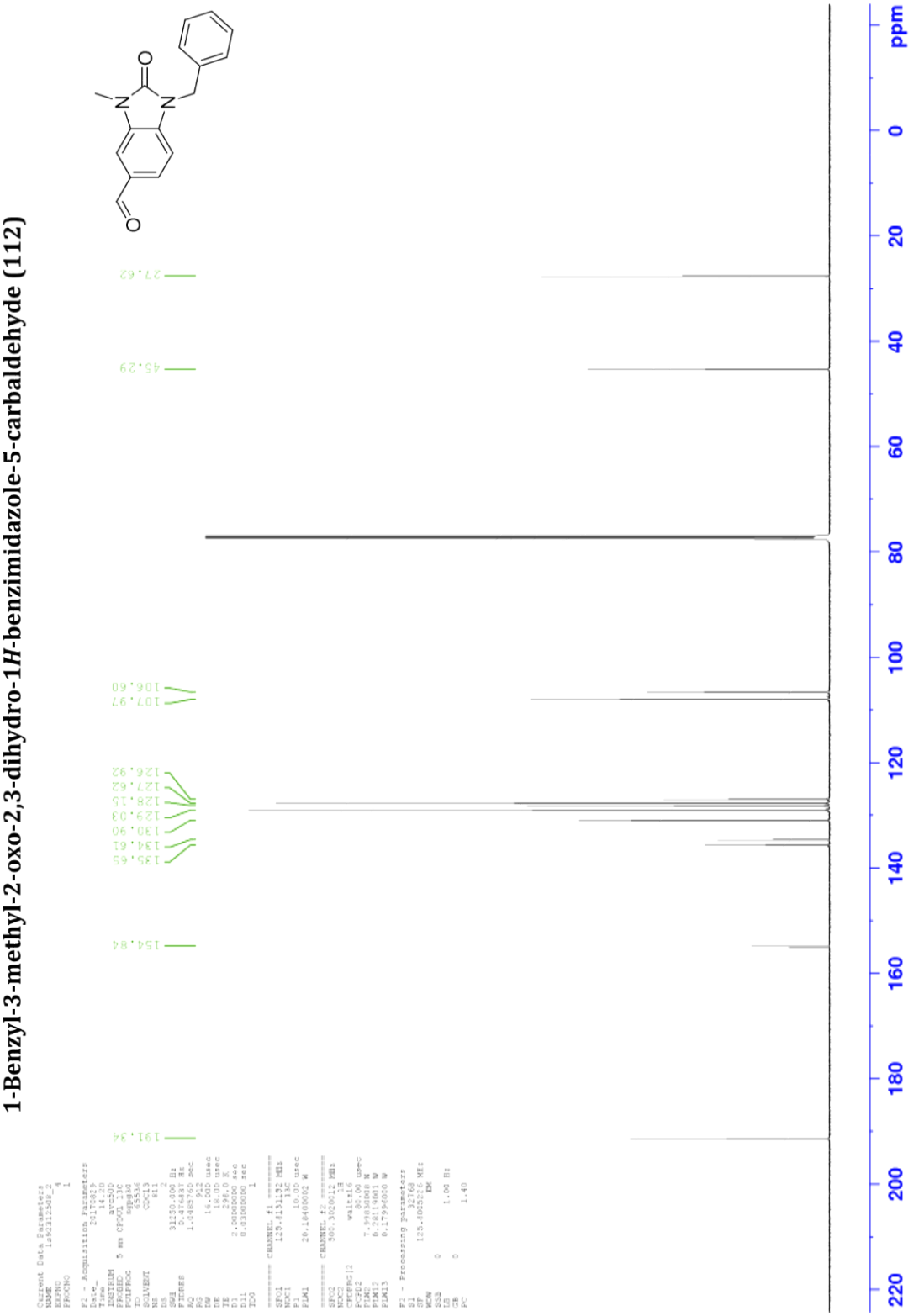
5-(Aminomethyl)-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (108)



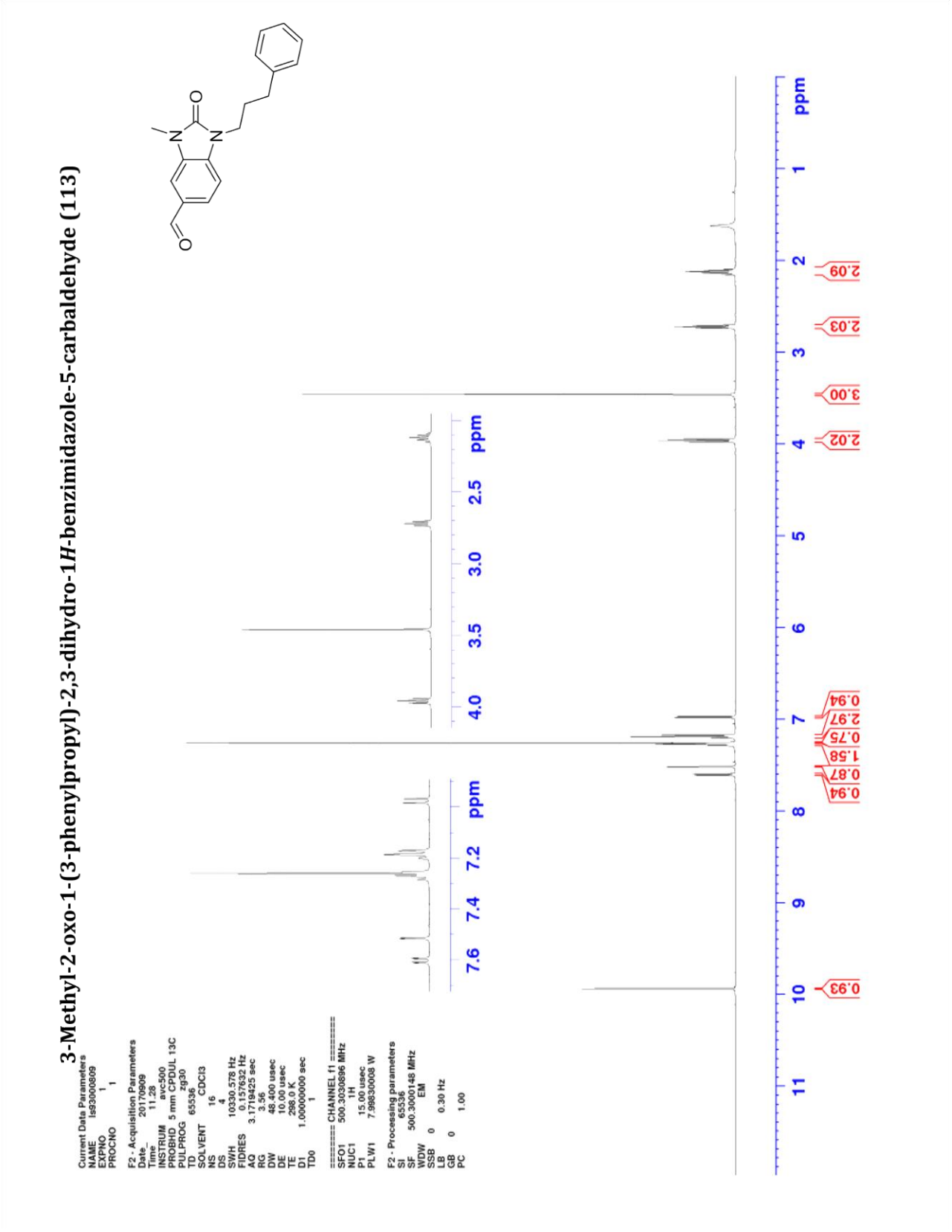
1-Benzyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (112)



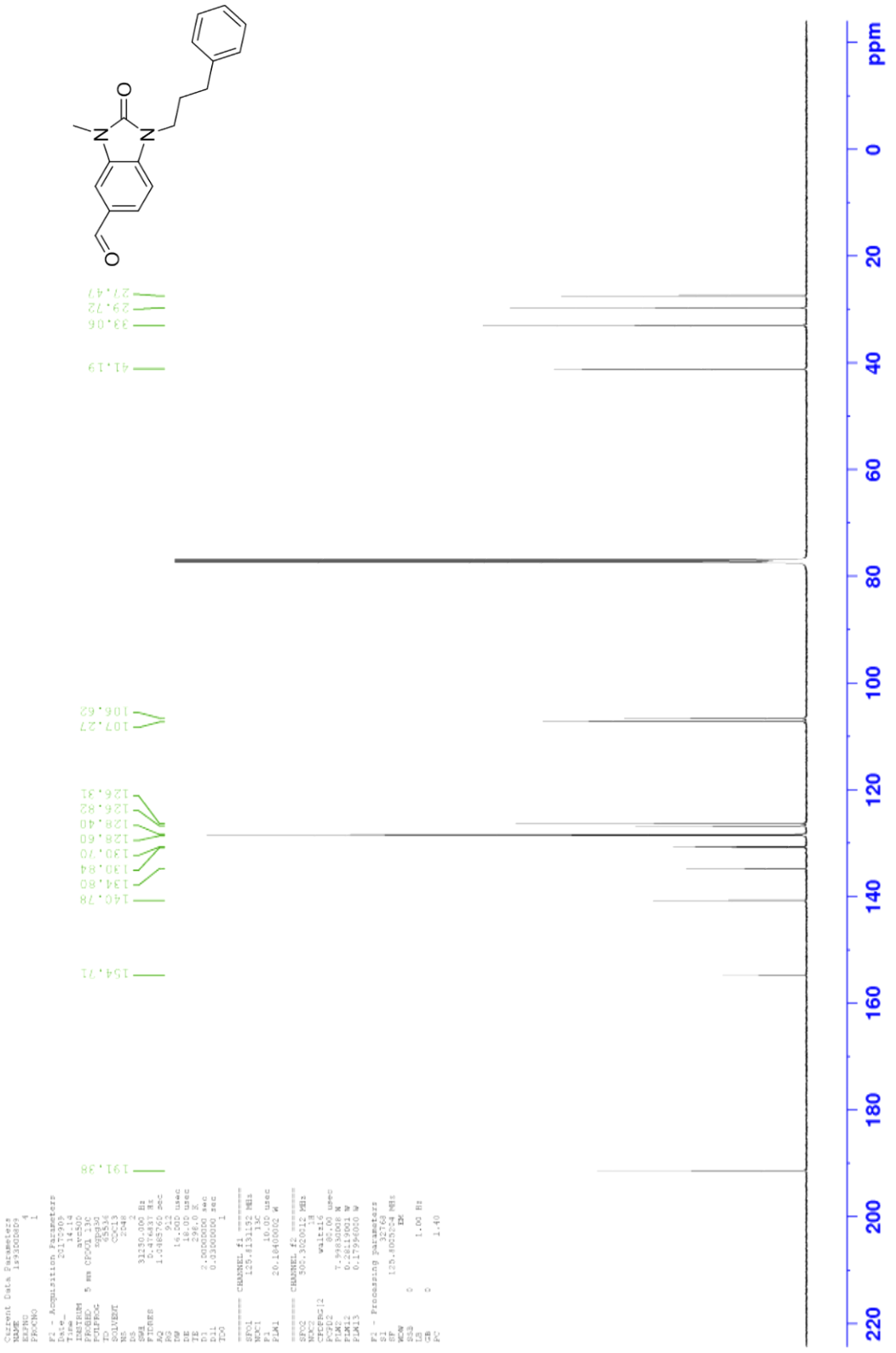
1-Benzyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (112)

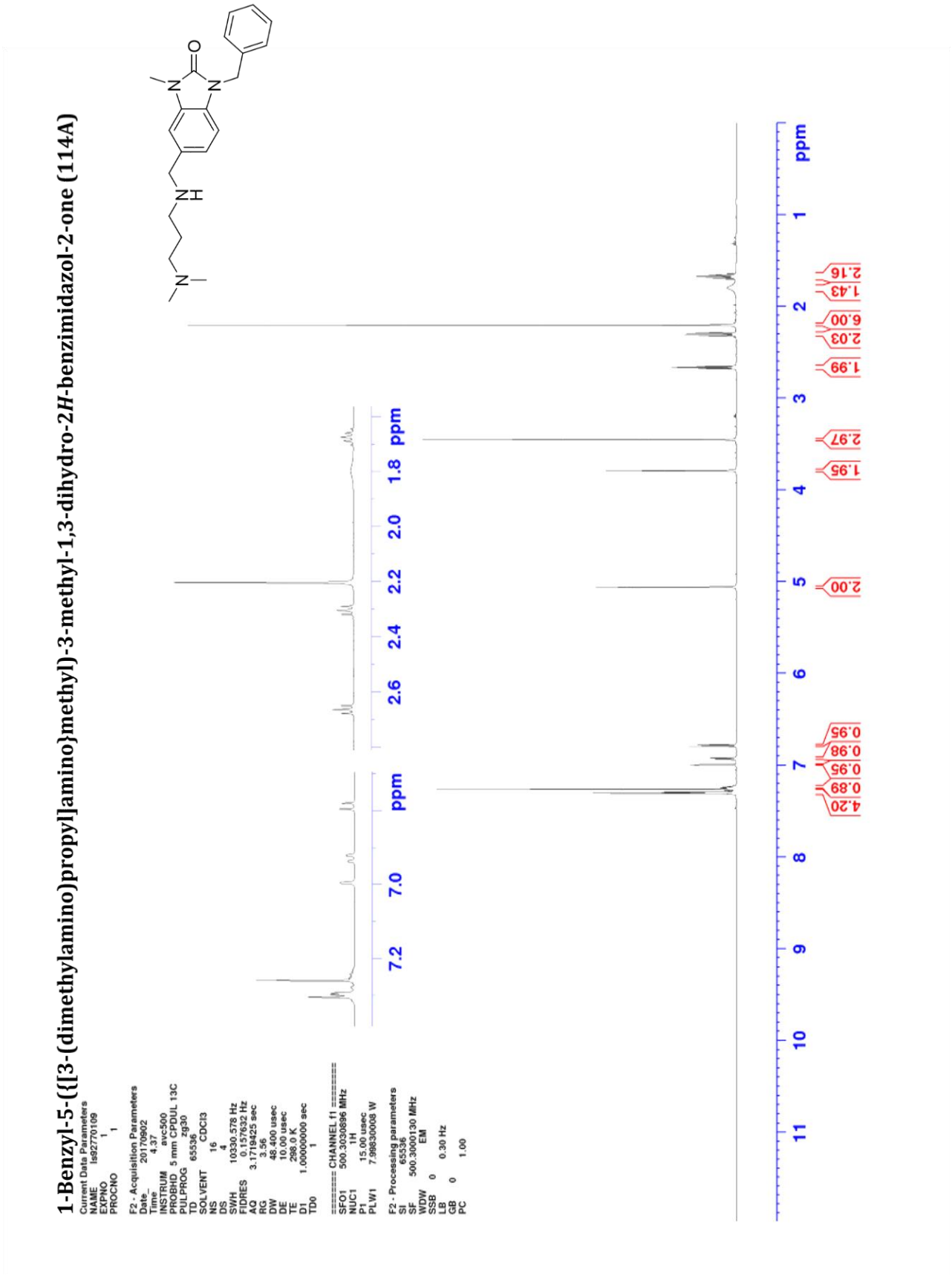


3-Methyl-2-oxo-1-(3-phenylpropyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (113)

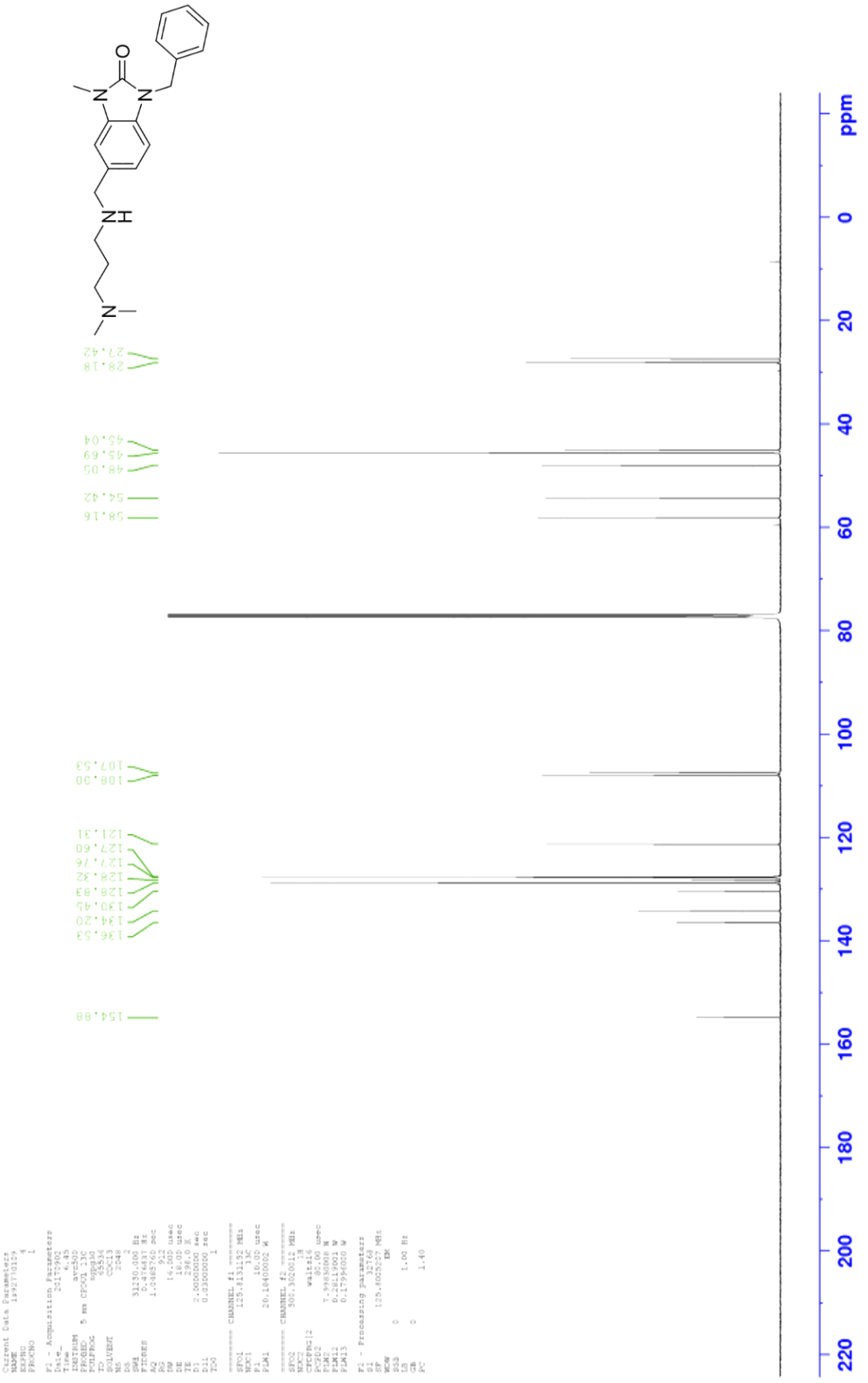


3-Methyl-2-oxo-1-(3-phenylpropyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (113)

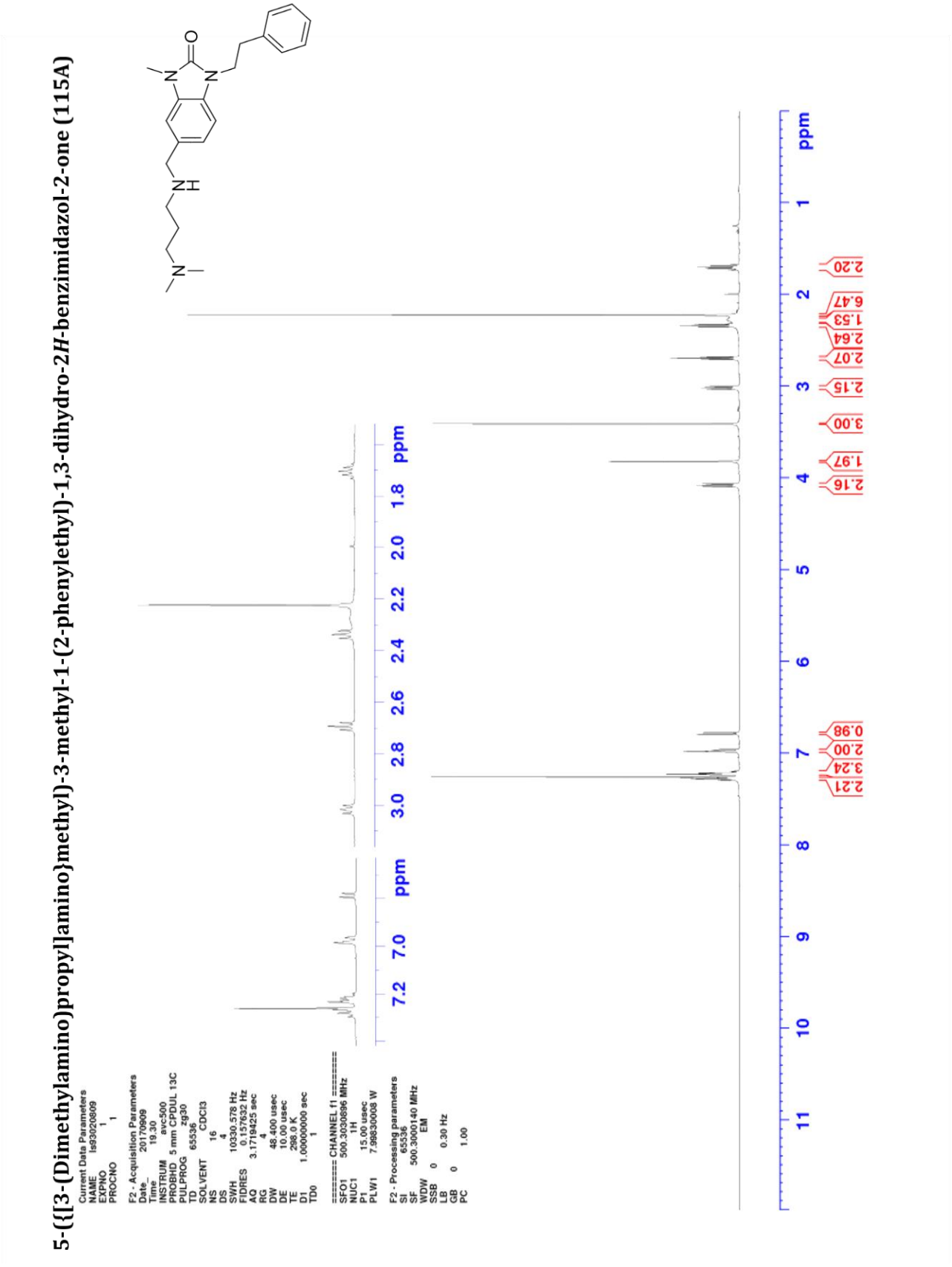



1-Benzyl-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2H-benzimidazol-2-one (114A)
Current Data Parameters
NAME ls92770109
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20170602
Time 4.37
INSTRUM av500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.1719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000130 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.2
7.0
ppm
2.6
2.4
2.2
2.0
1.8
ppm
11
10
9
8
7
6
5
4
3
2
1
ppm
4.20
0.89
0.95
0.98
0.95
2.00
1.95
2.97
1.99
2.03
6.00
1.43
2.16

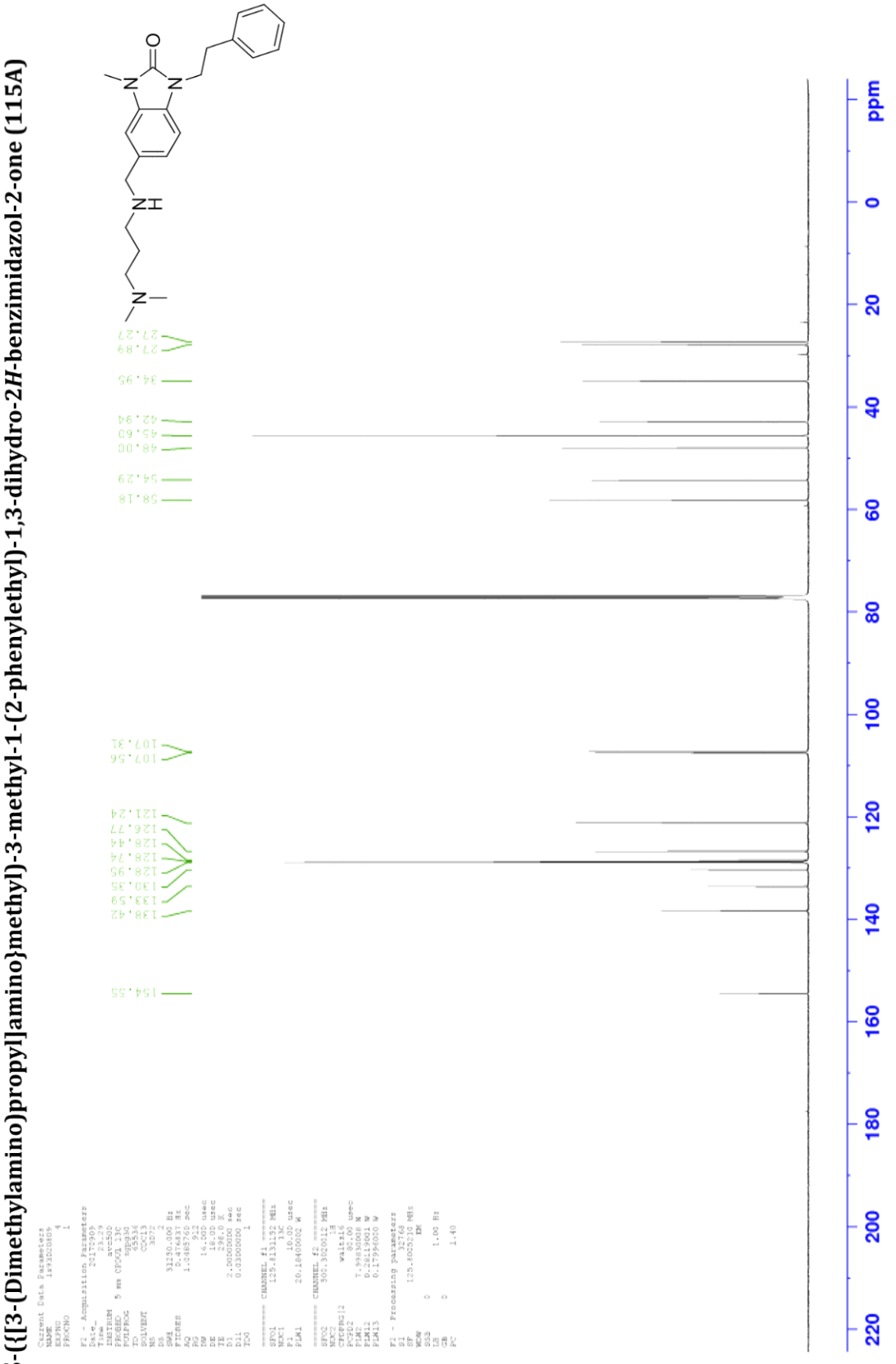
1-Benzyl-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (114A)



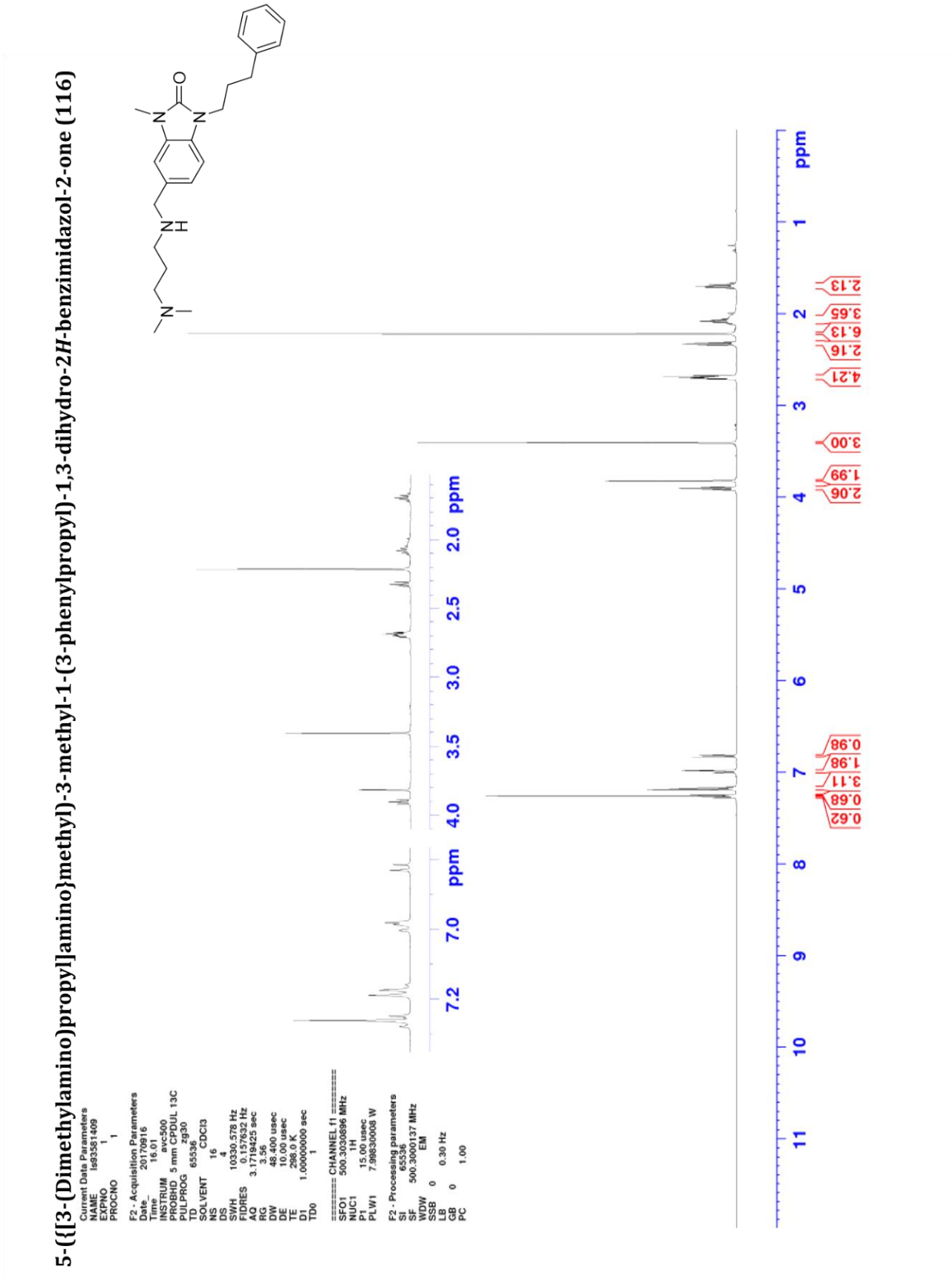
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (115A)



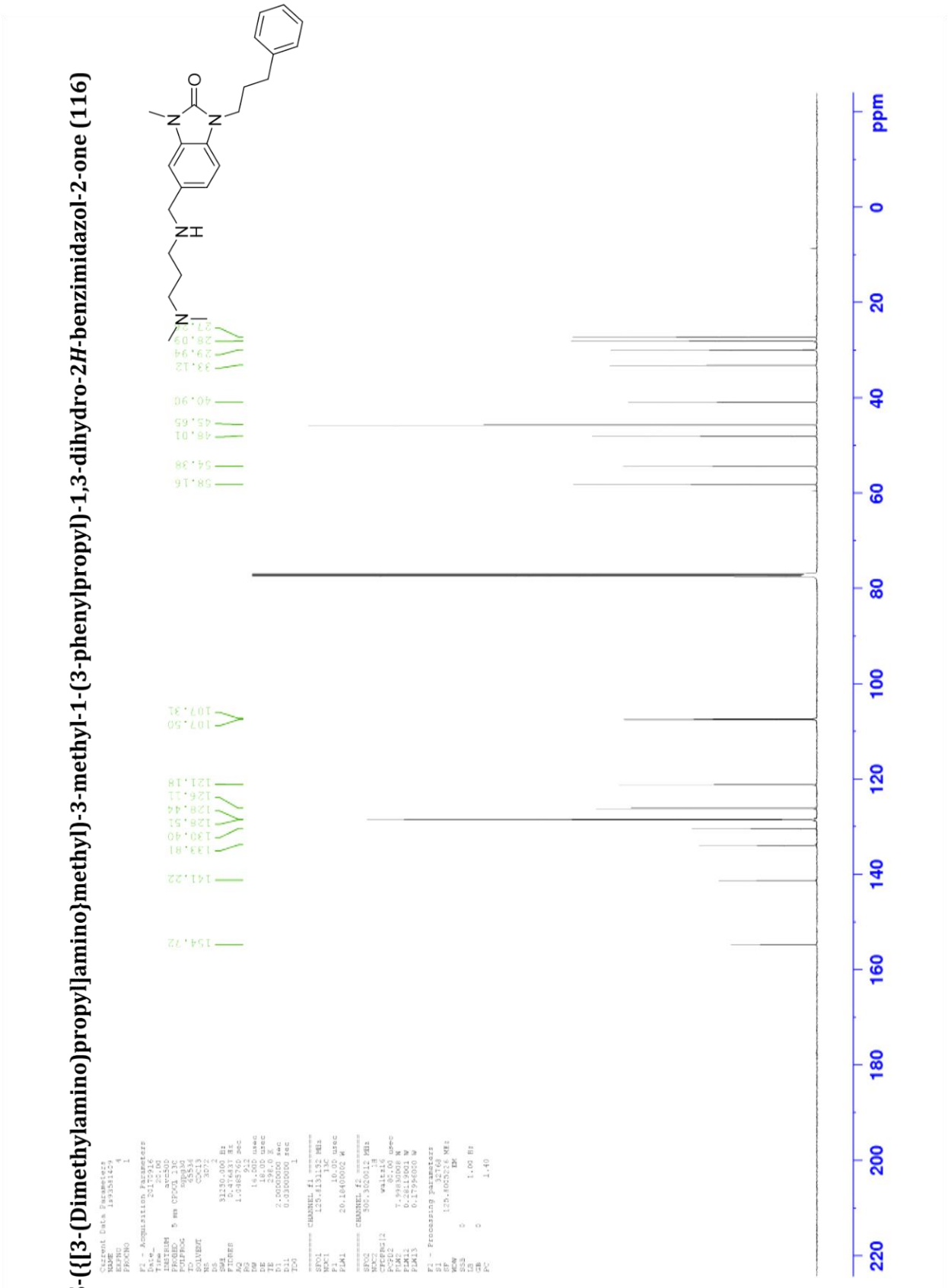
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (115A)



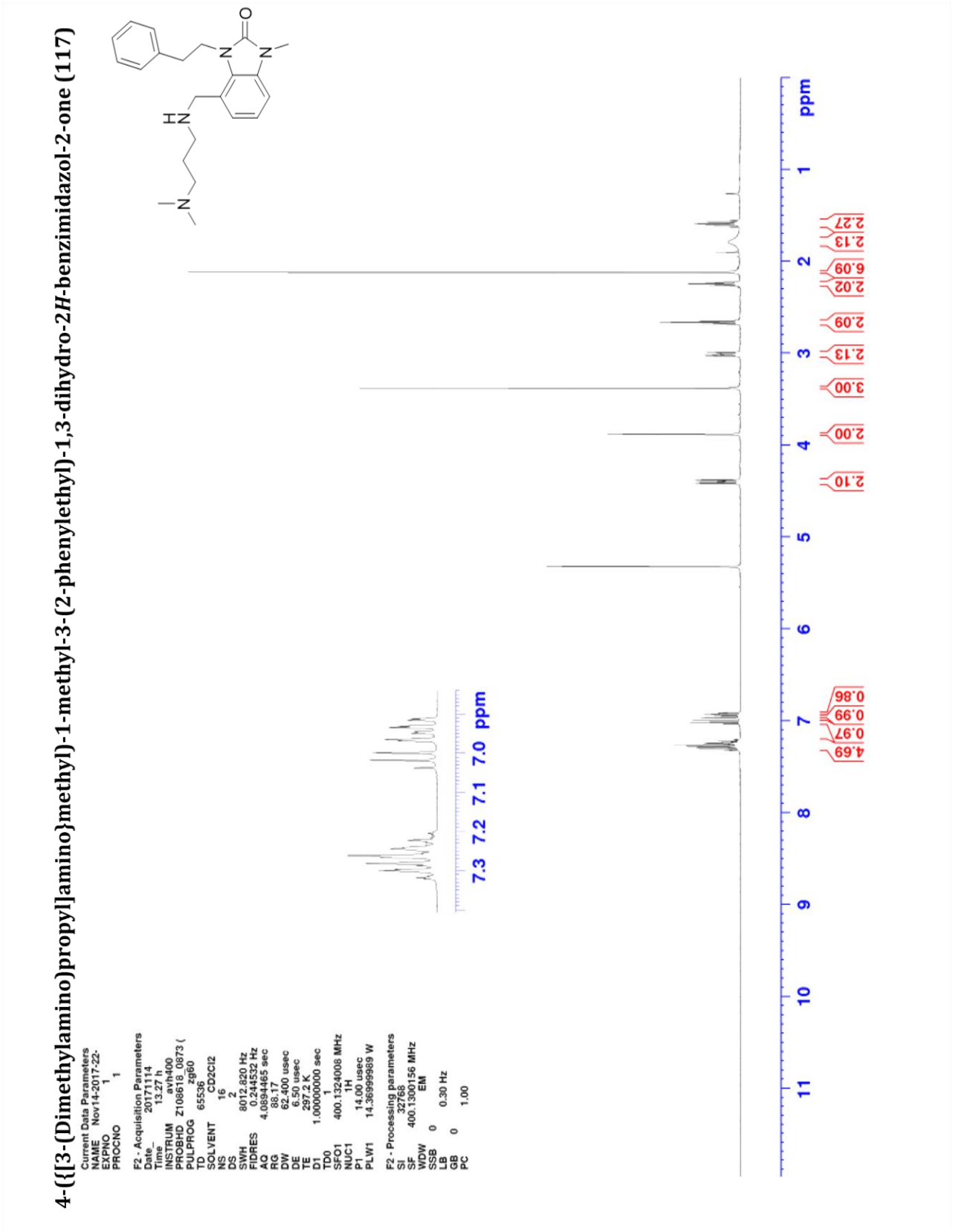
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(3-phenylpropyl)-1,3-dihydro-2*H*-benzimidazol-2-one (116)



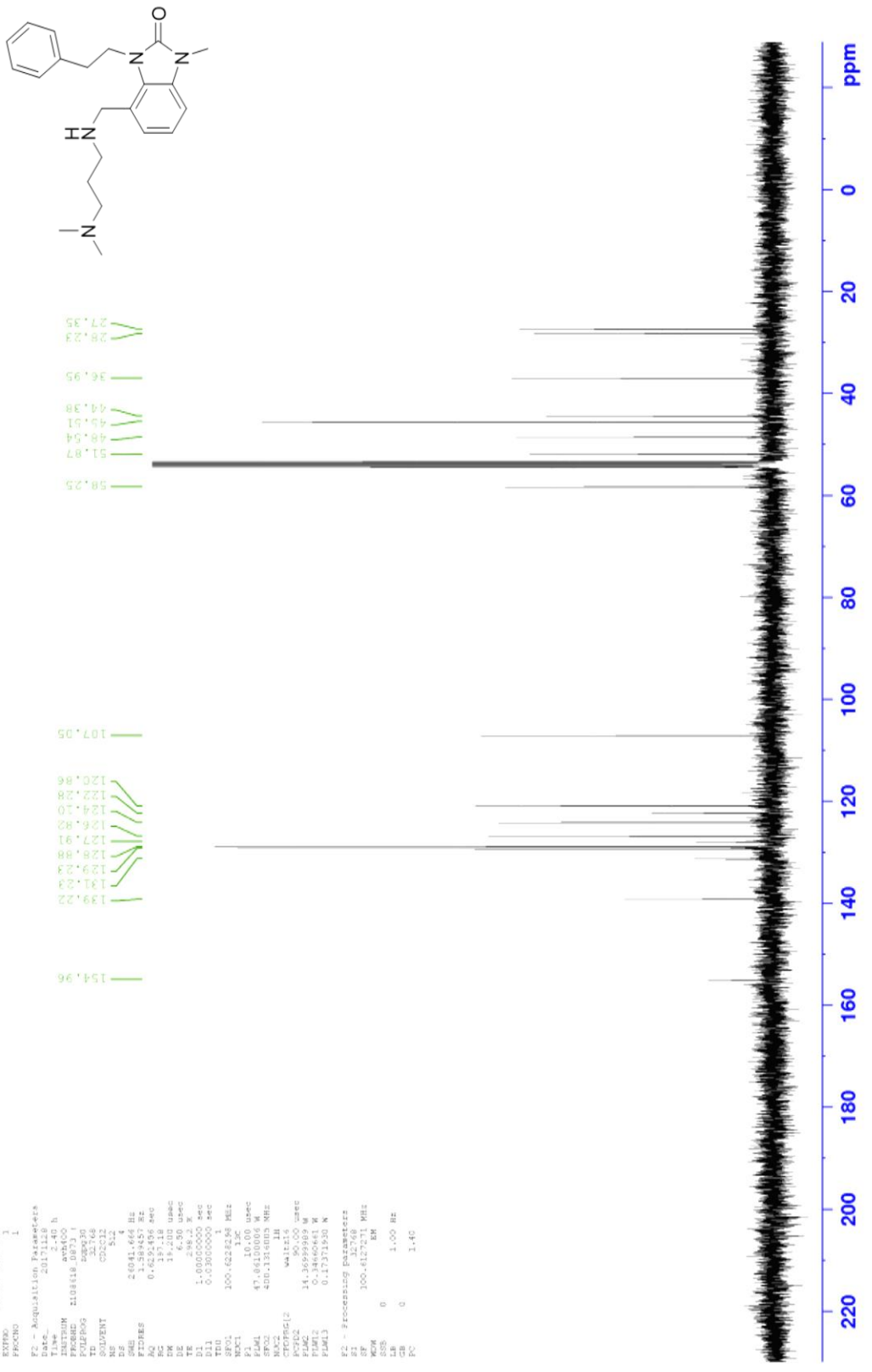
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(3-phenylpropyl)-1,3-dihydro-2*H*-benzimidazol-2-one (116)



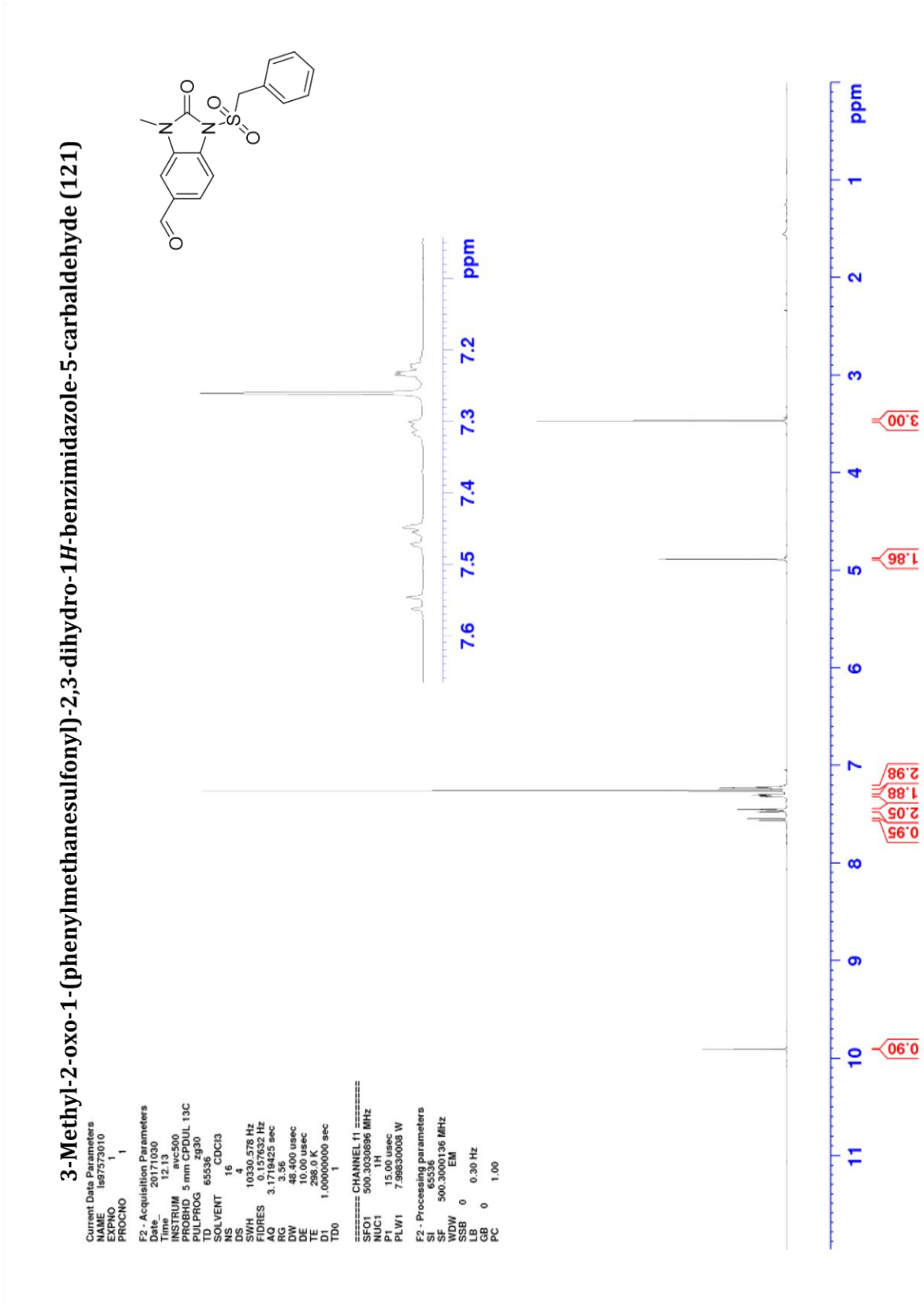
4-{[3-(Dimethylamino)propyl]amino}methyl)-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (117)



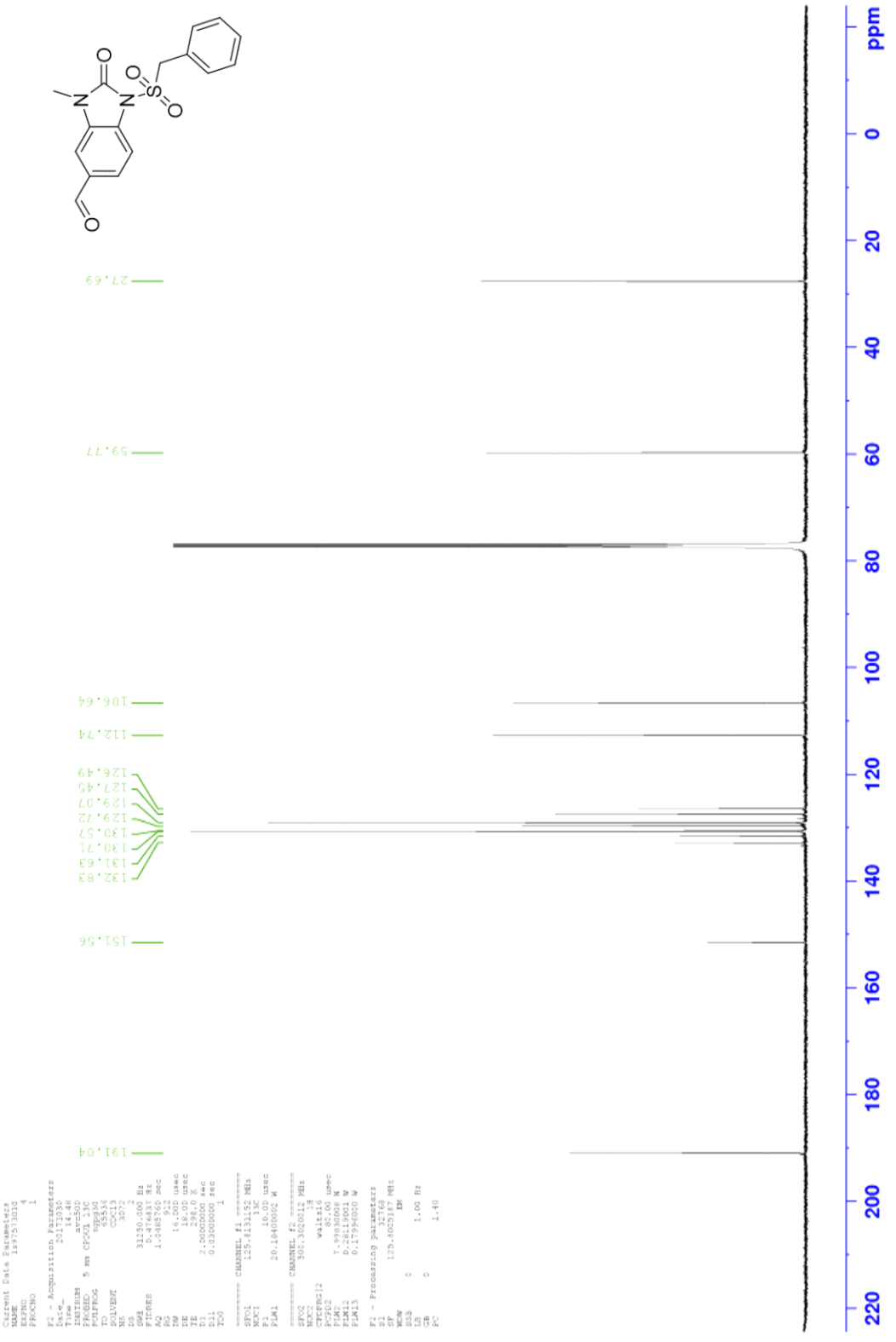
4-{[[3-(Dimethylamino)propyl]amino}methyl)-1-methyl-3-(2-phenylethyl)-1,3-dihydro-2*H*-benzimidazol-2-one (117)



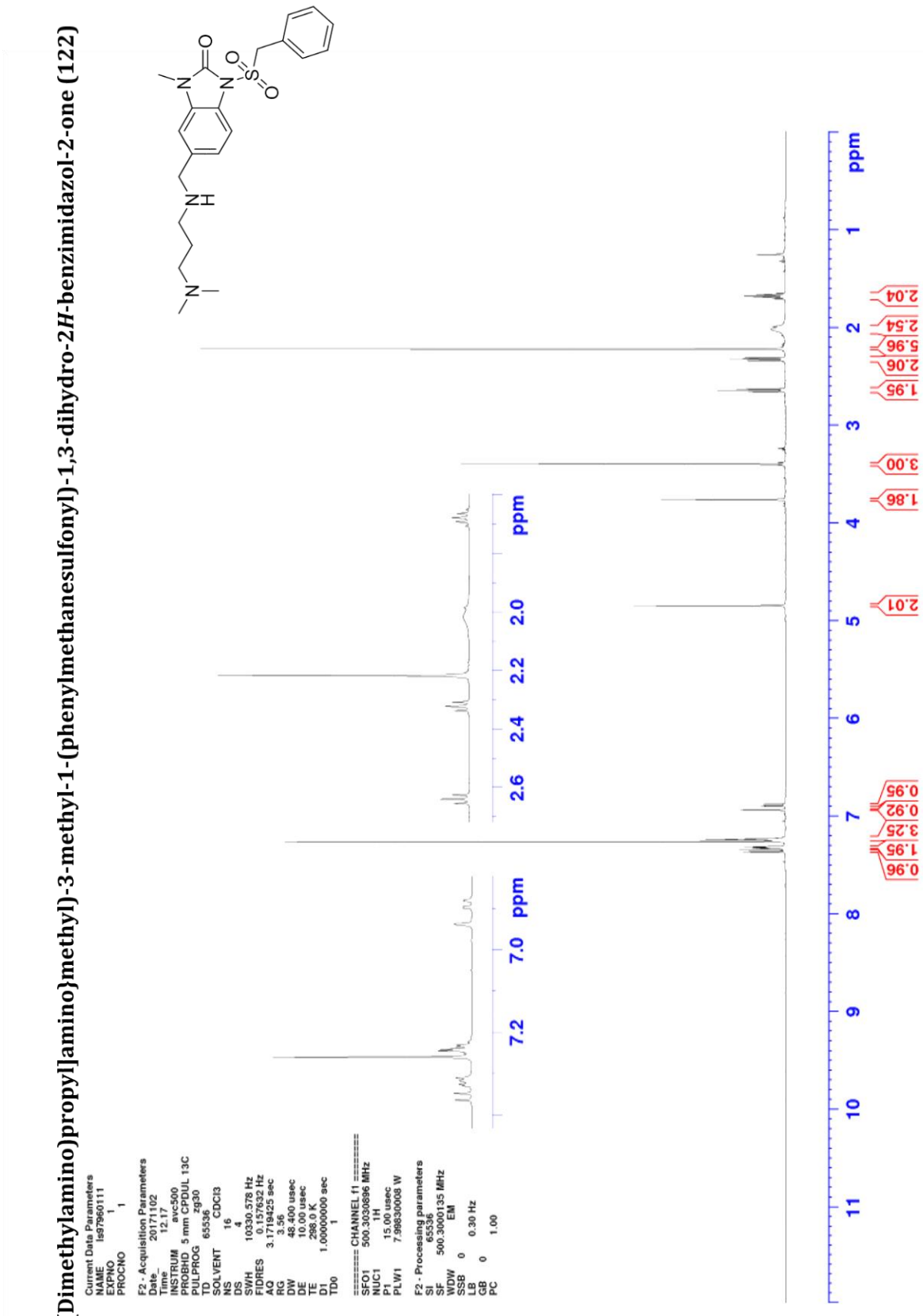
3-Methyl-2-oxo-1-(phenylmethanesulfonyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (121)



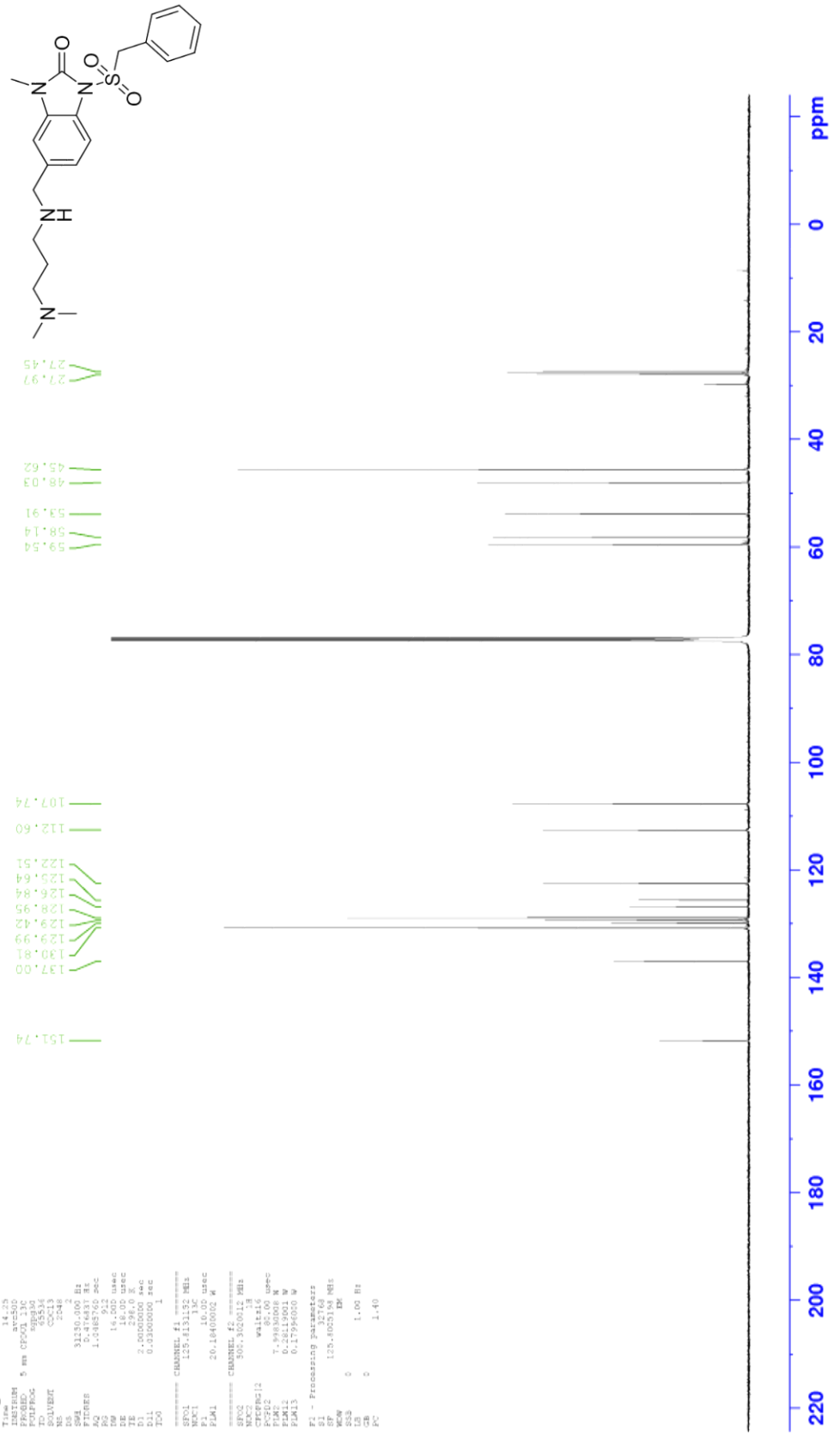
3-Methyl-2-oxo-1-(phenylmethanesulfonyl)-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (121)



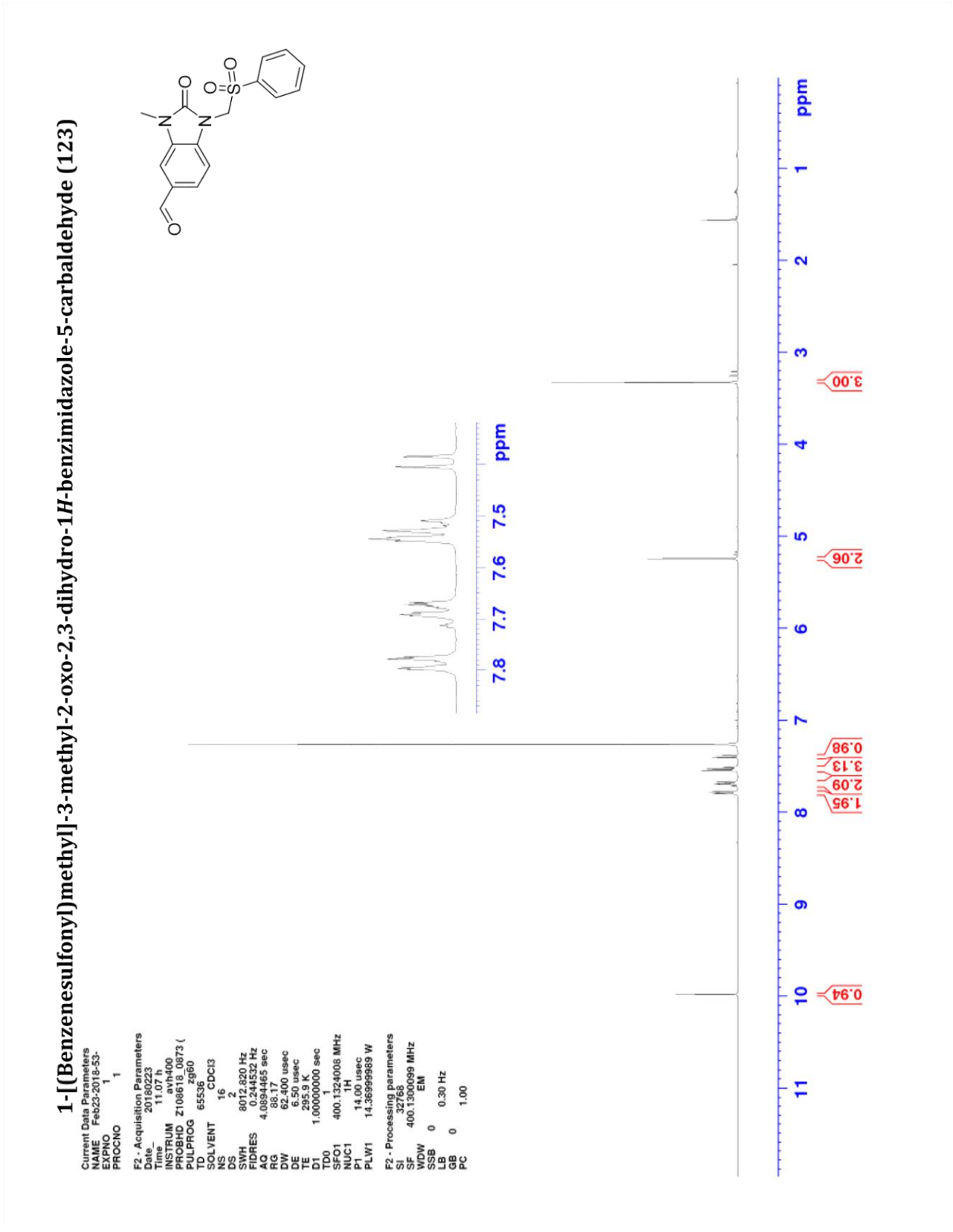
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(phenylmethanesulfonyl)-1,3-dihydro-2*H*-benzimidazol-2-one (122)



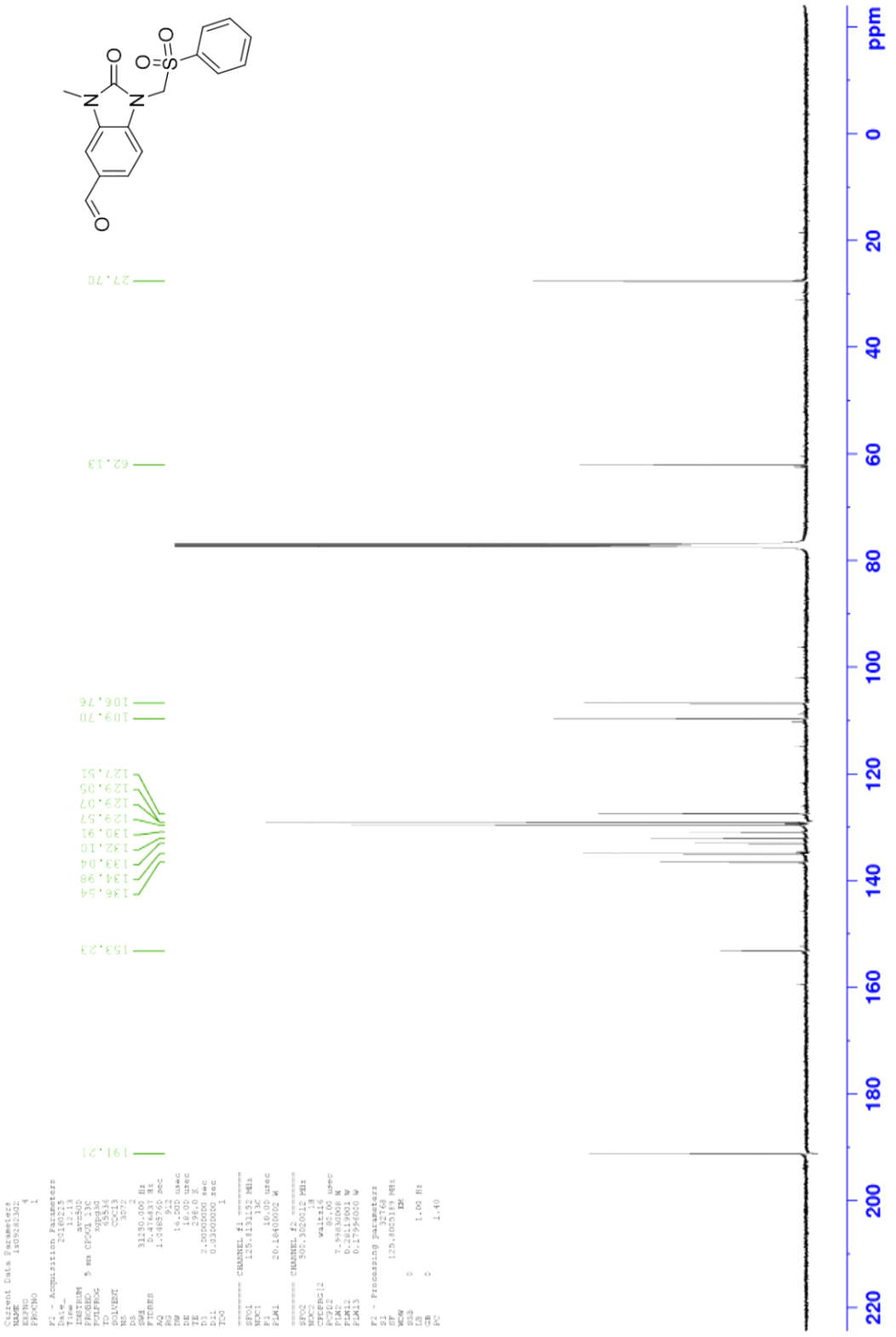
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-(phenylmethanesulfonyl)-1,3-dihydro-2*H*-benzimidazol-2-one (122)



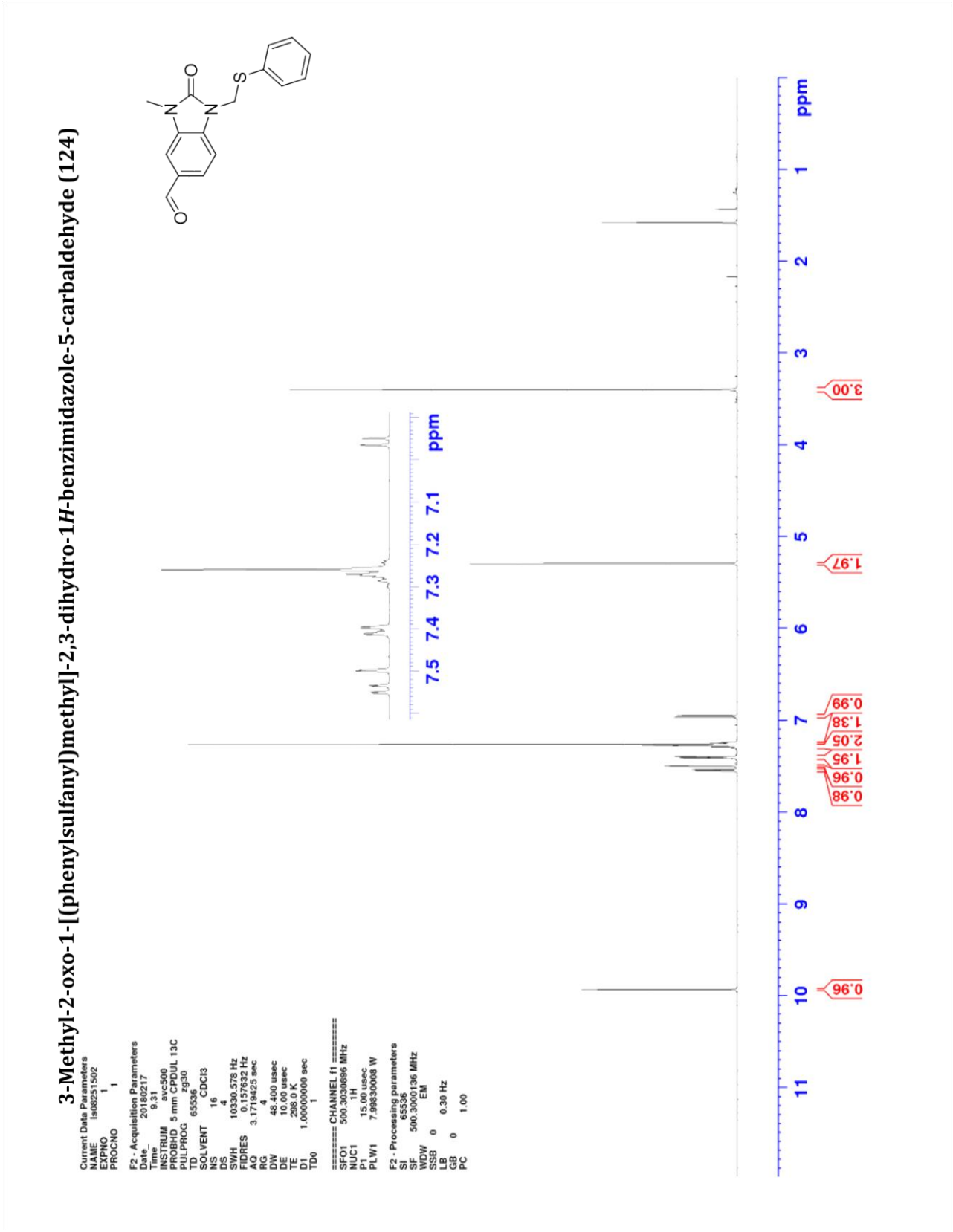
1-[(Benzenesulfonyl)methyl]-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (123)



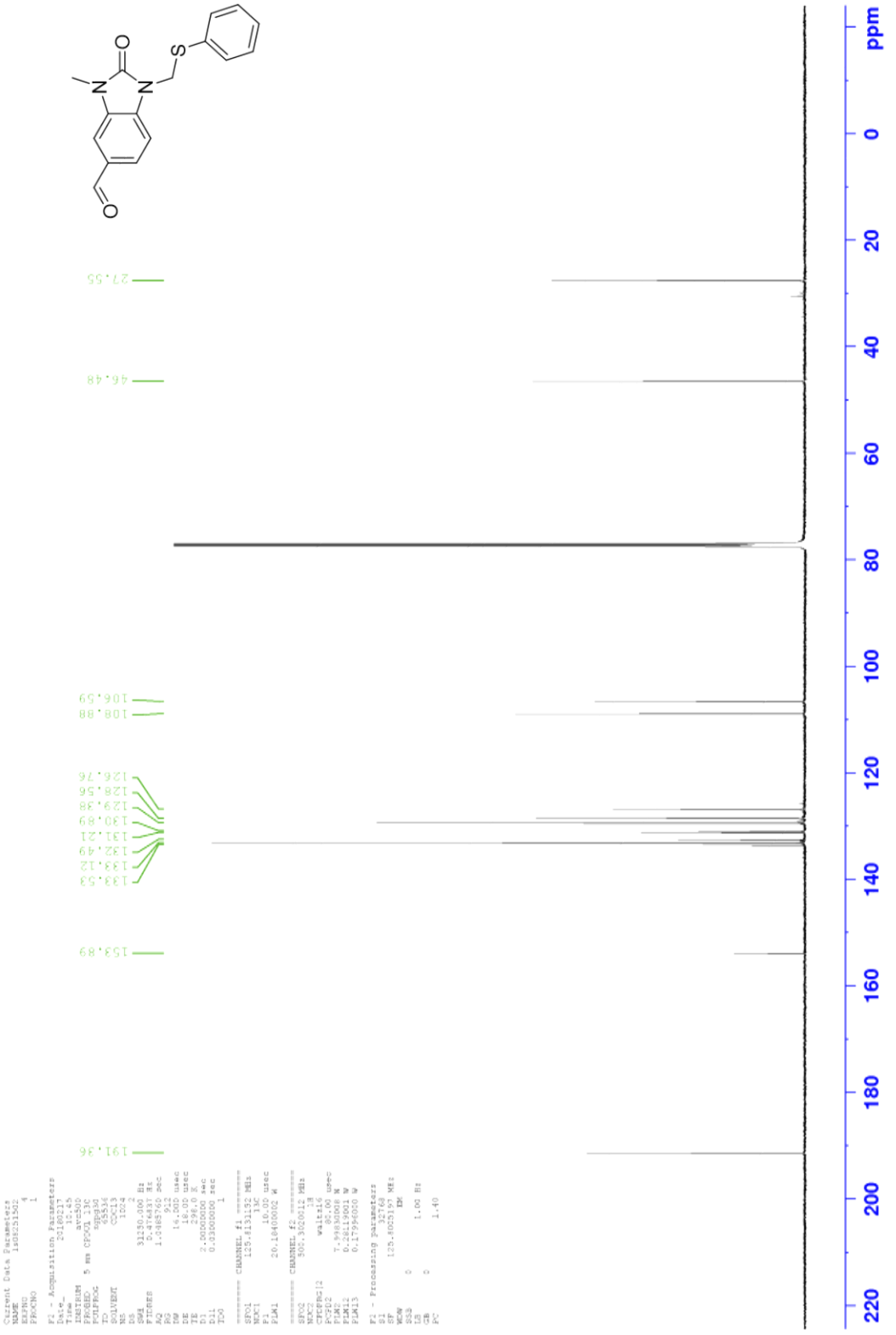
1-[(Benzenesulfonyl)methyl]-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (123)



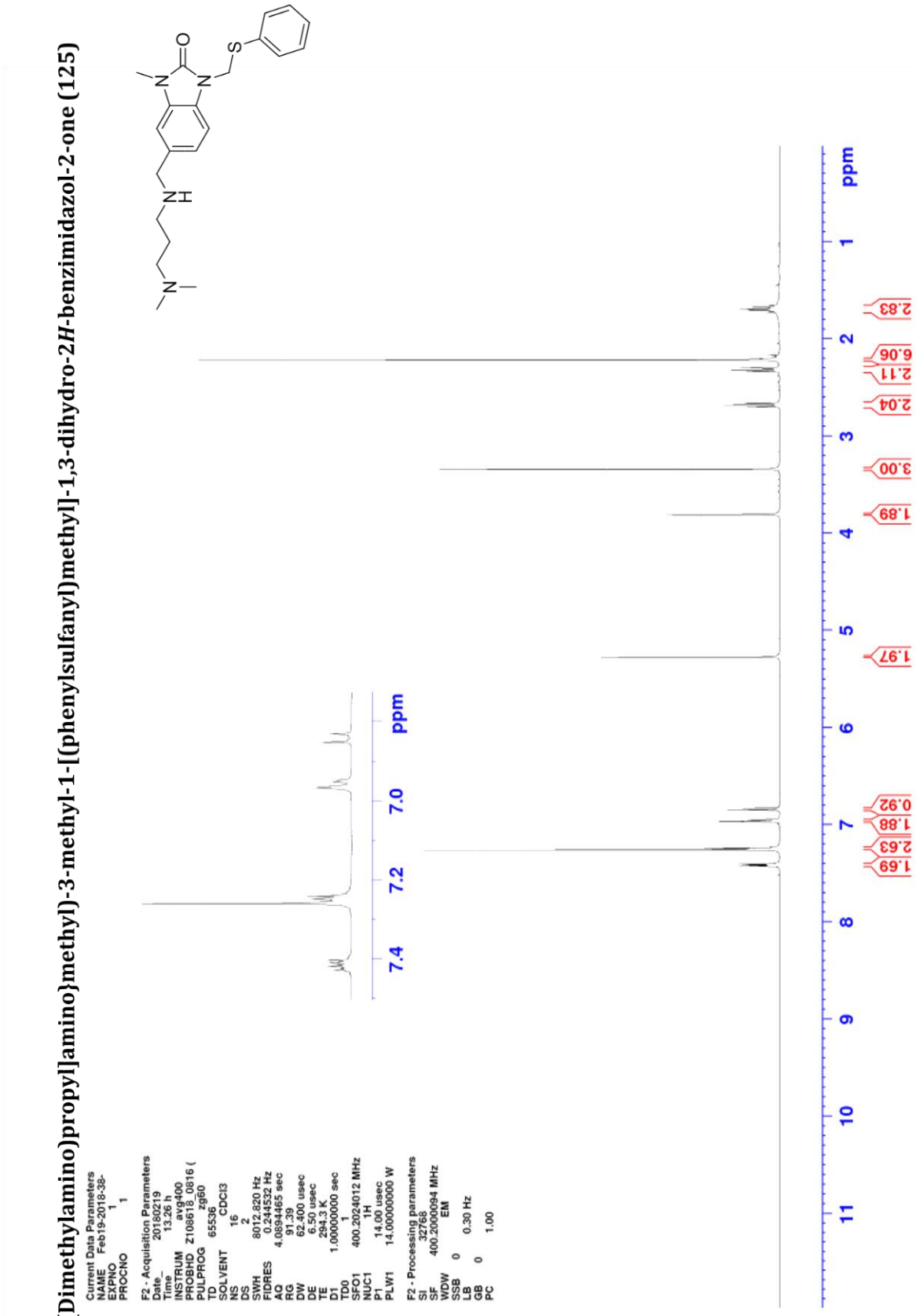
3-Methyl-2-oxo-1-[(phenylsulfanyl)methyl]-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (124)



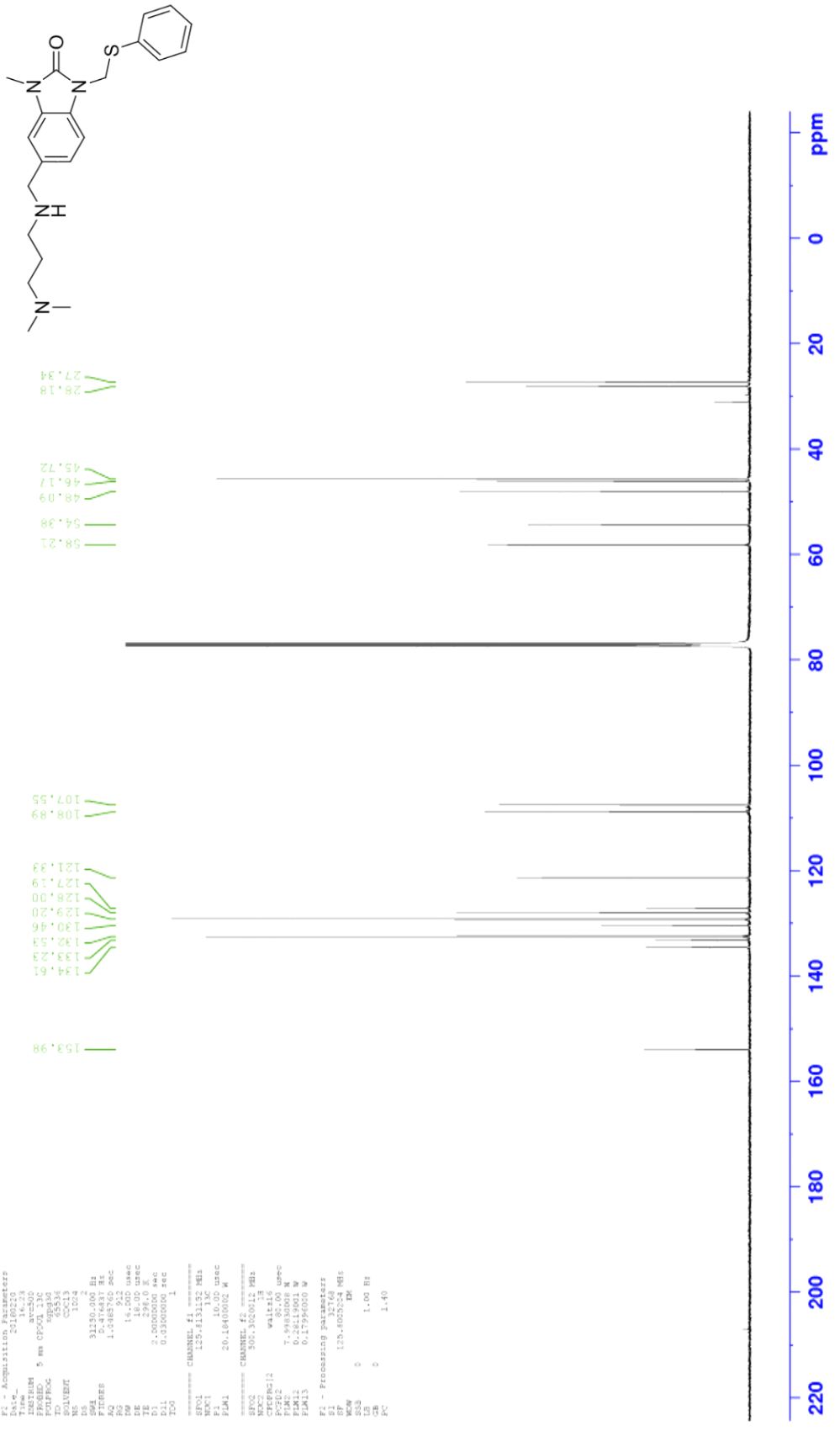
3-Methyl-2-oxo-1-[(phenylsulfanyl)methyl]-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (124)



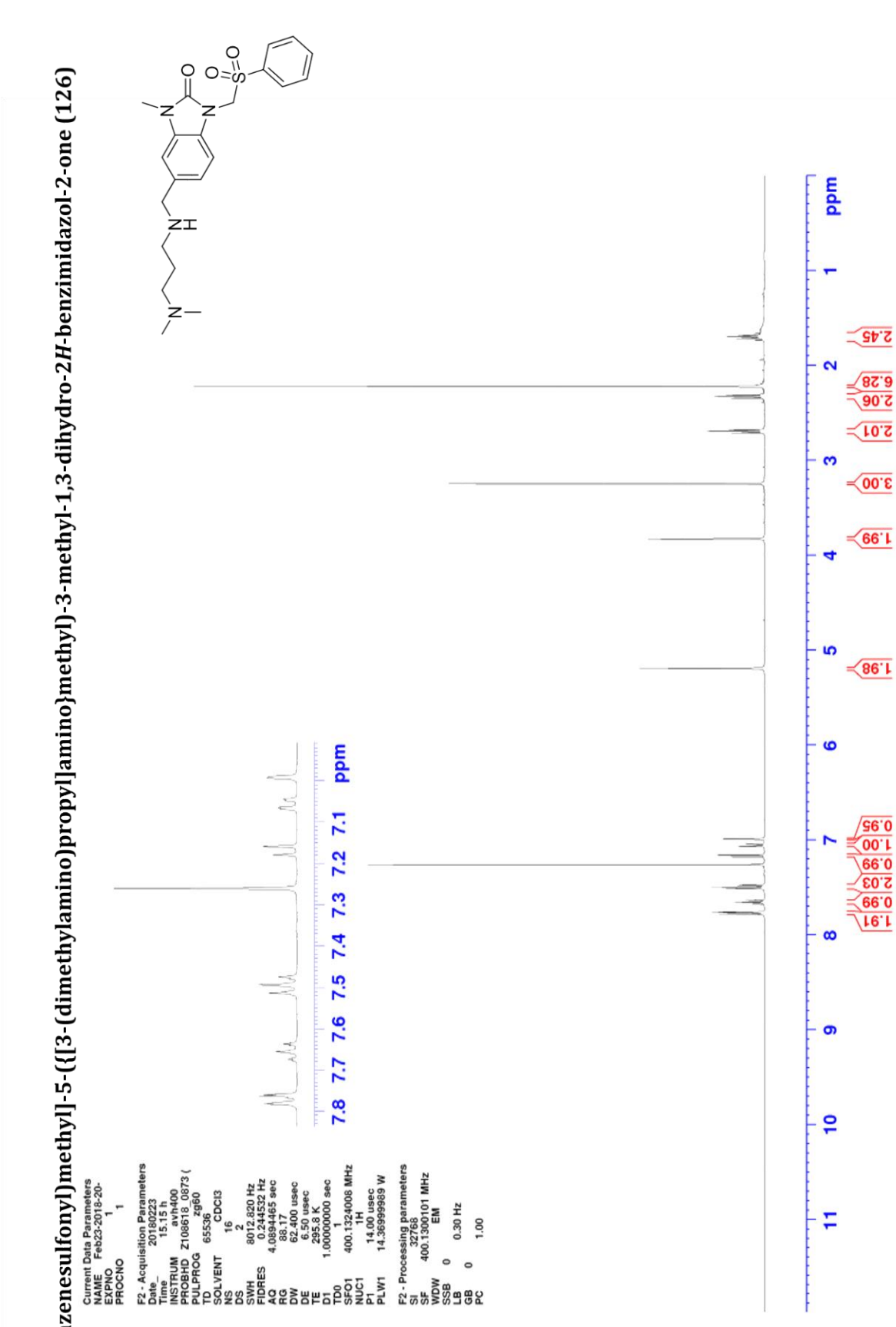
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-[(phenylsulfanyl)methyl]-1,3-dihydro-2*H*-benzimidazol-2-one (125)



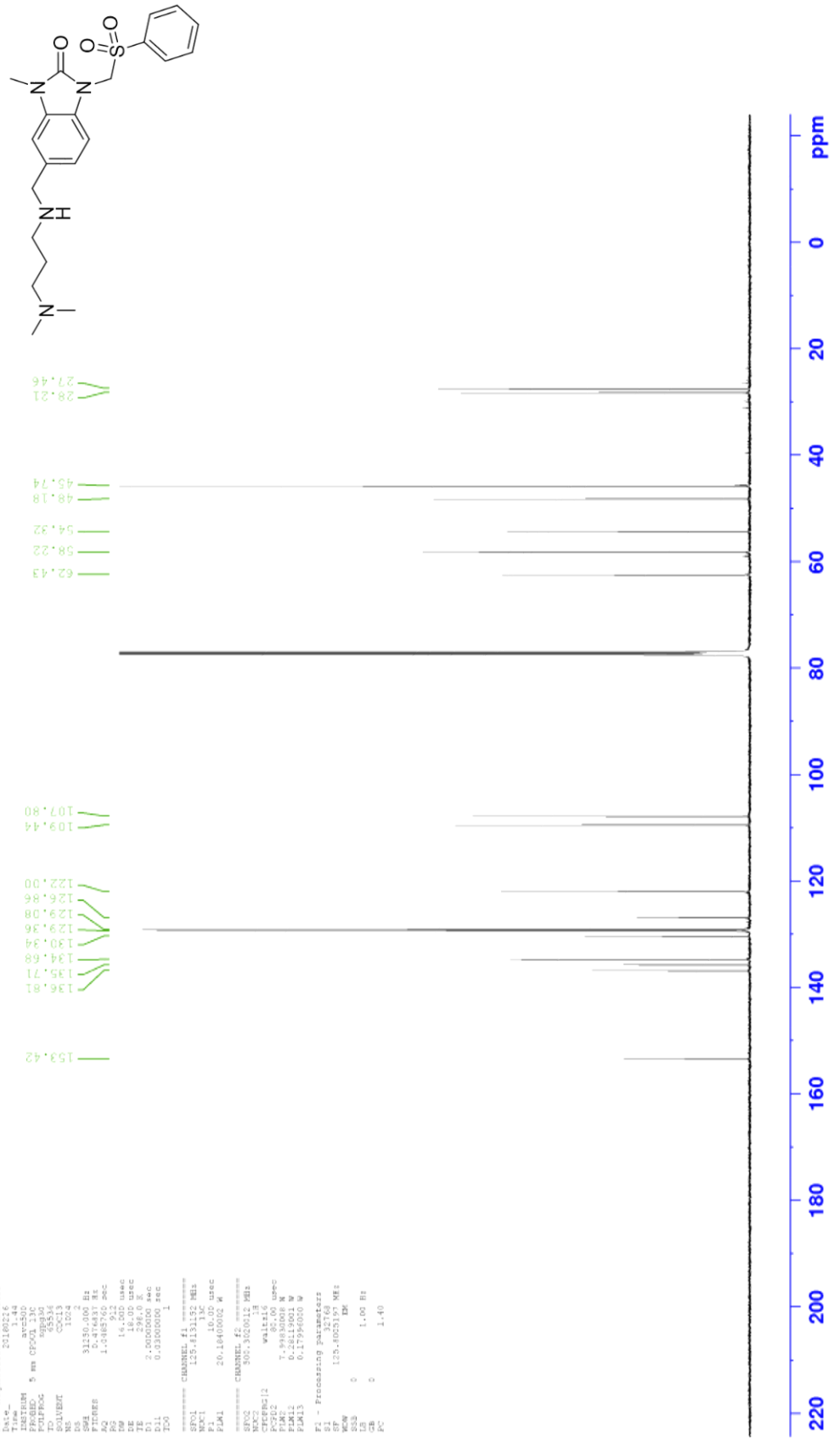
5-({[3-(Dimethylamino)propyl]amino}methyl)-3-methyl-1-[(phenylsulfanyl)methyl]-1,3-dihydro-2*H*-benzimidazol-2-one (125)



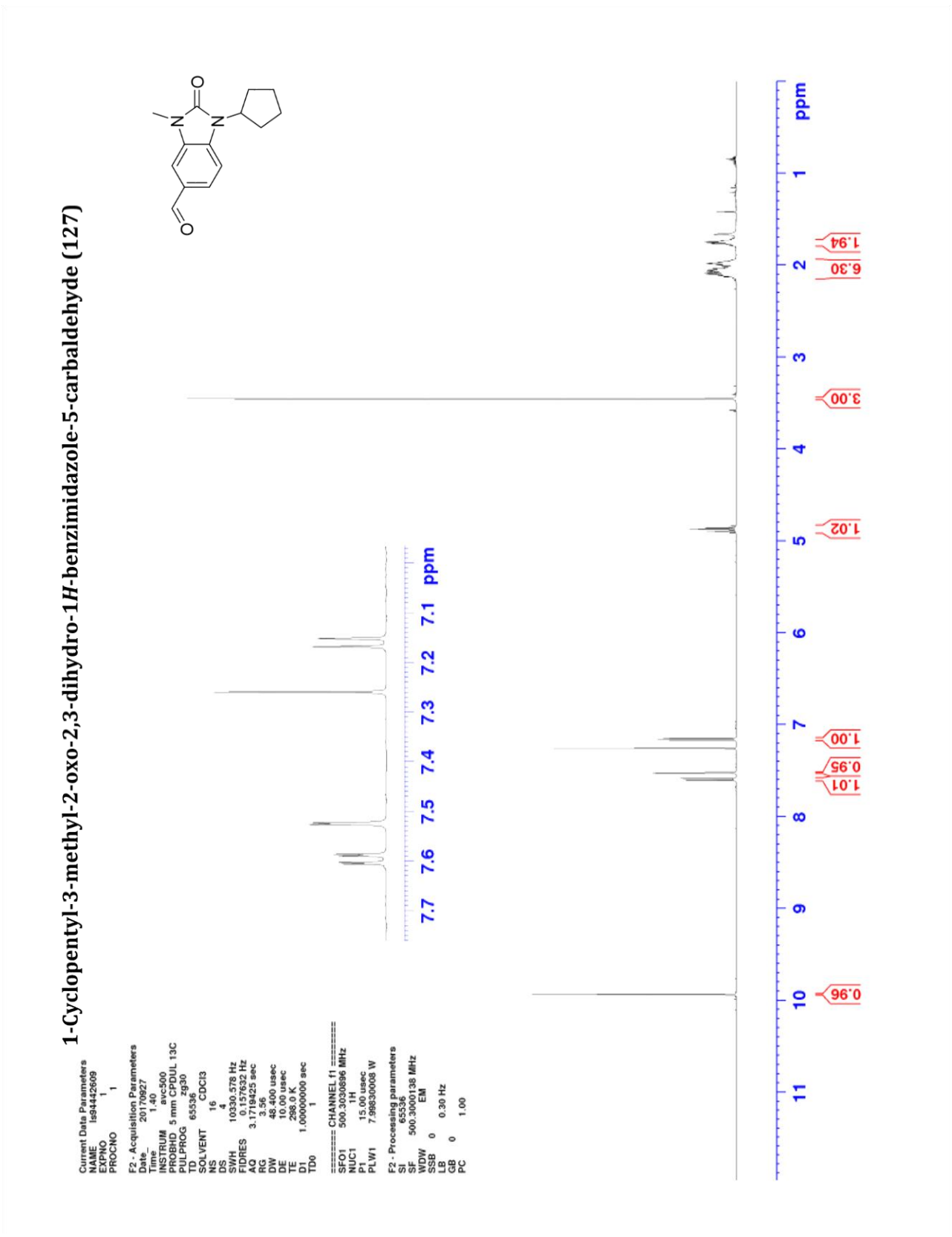
1-[(Benzenesulfonyl)methyl]-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (126)



1-[(Benzenesulfonyl)methyl]-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (126)



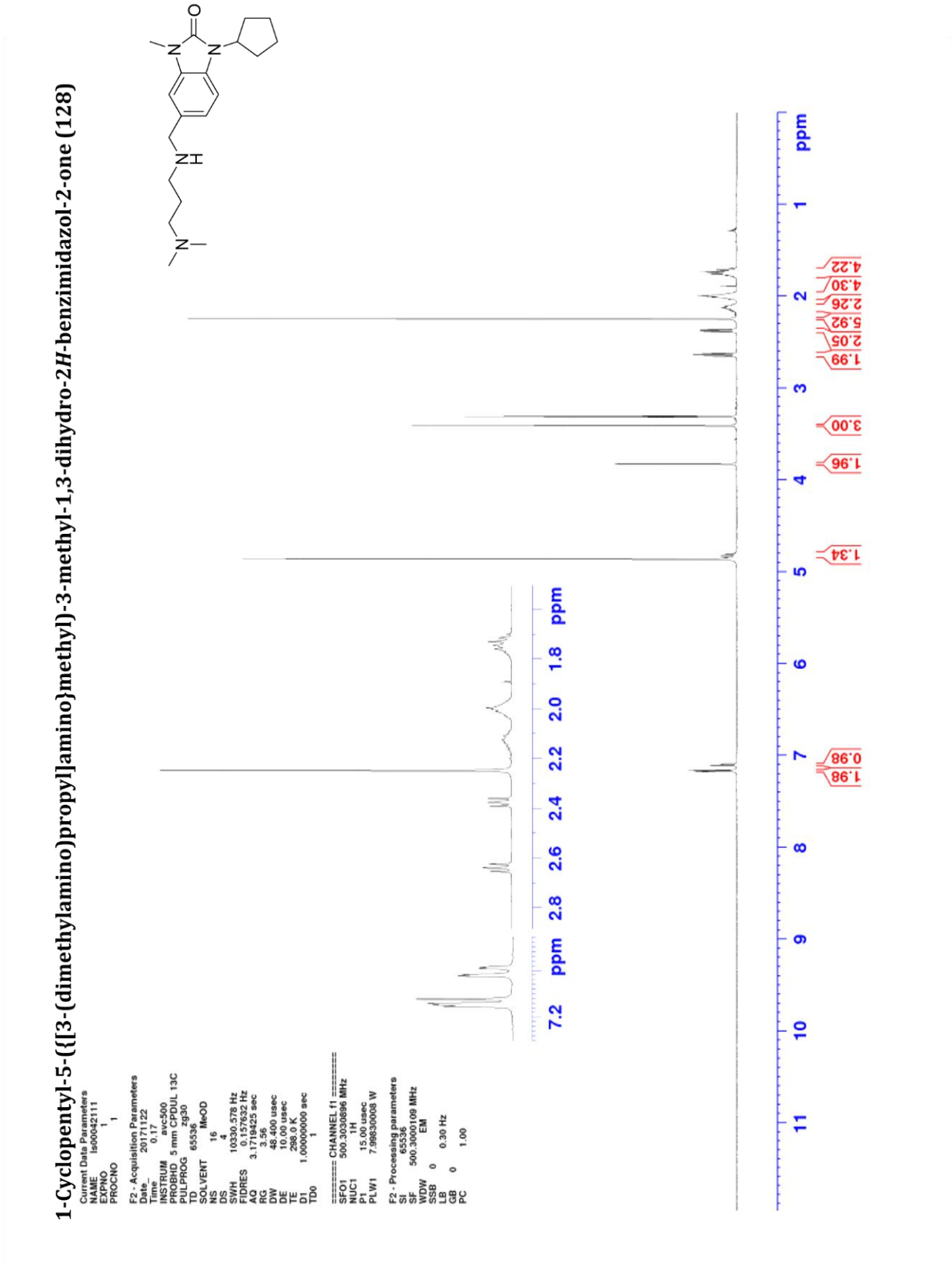
1-Cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (127)



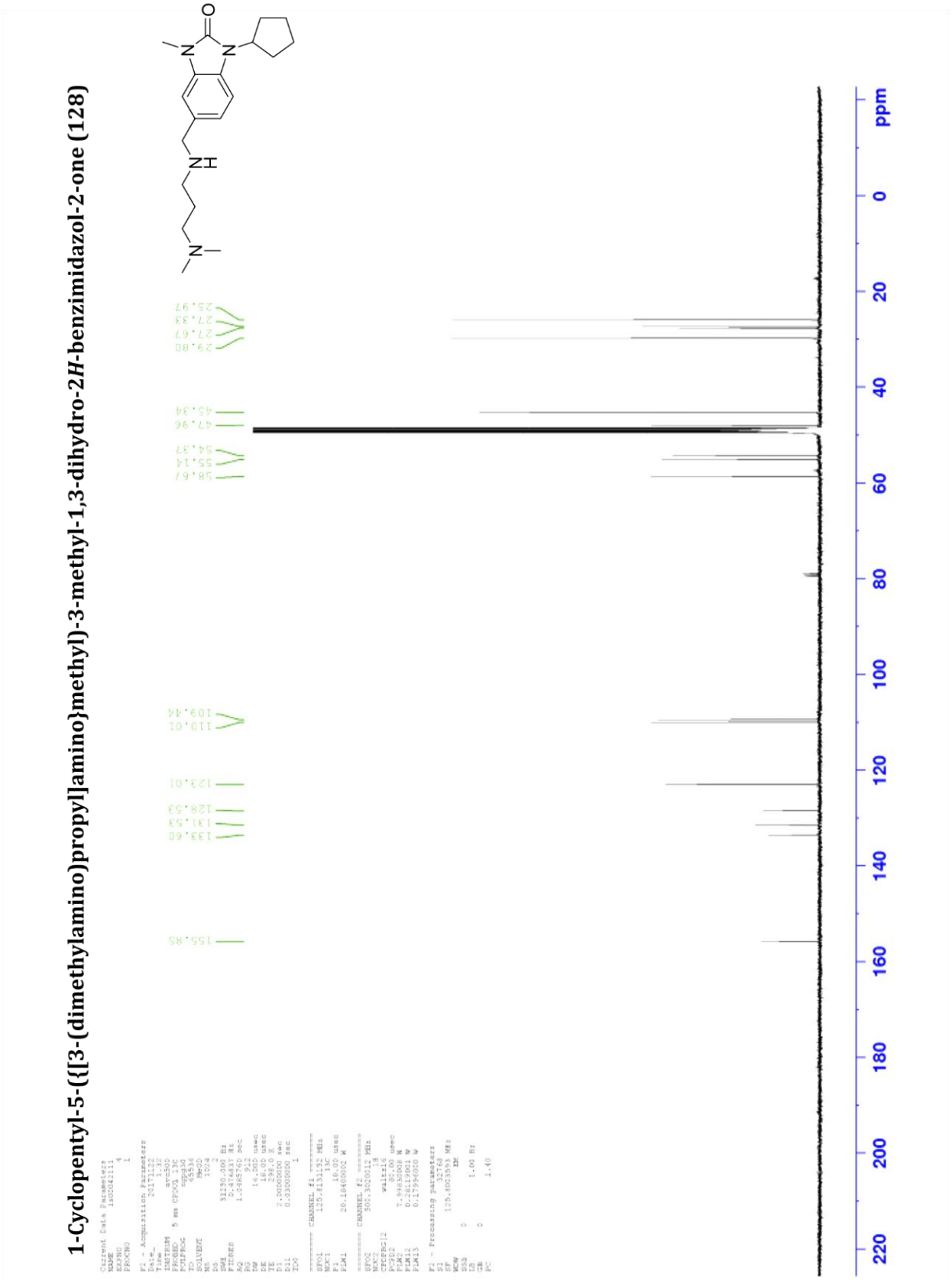
1-Cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (127)



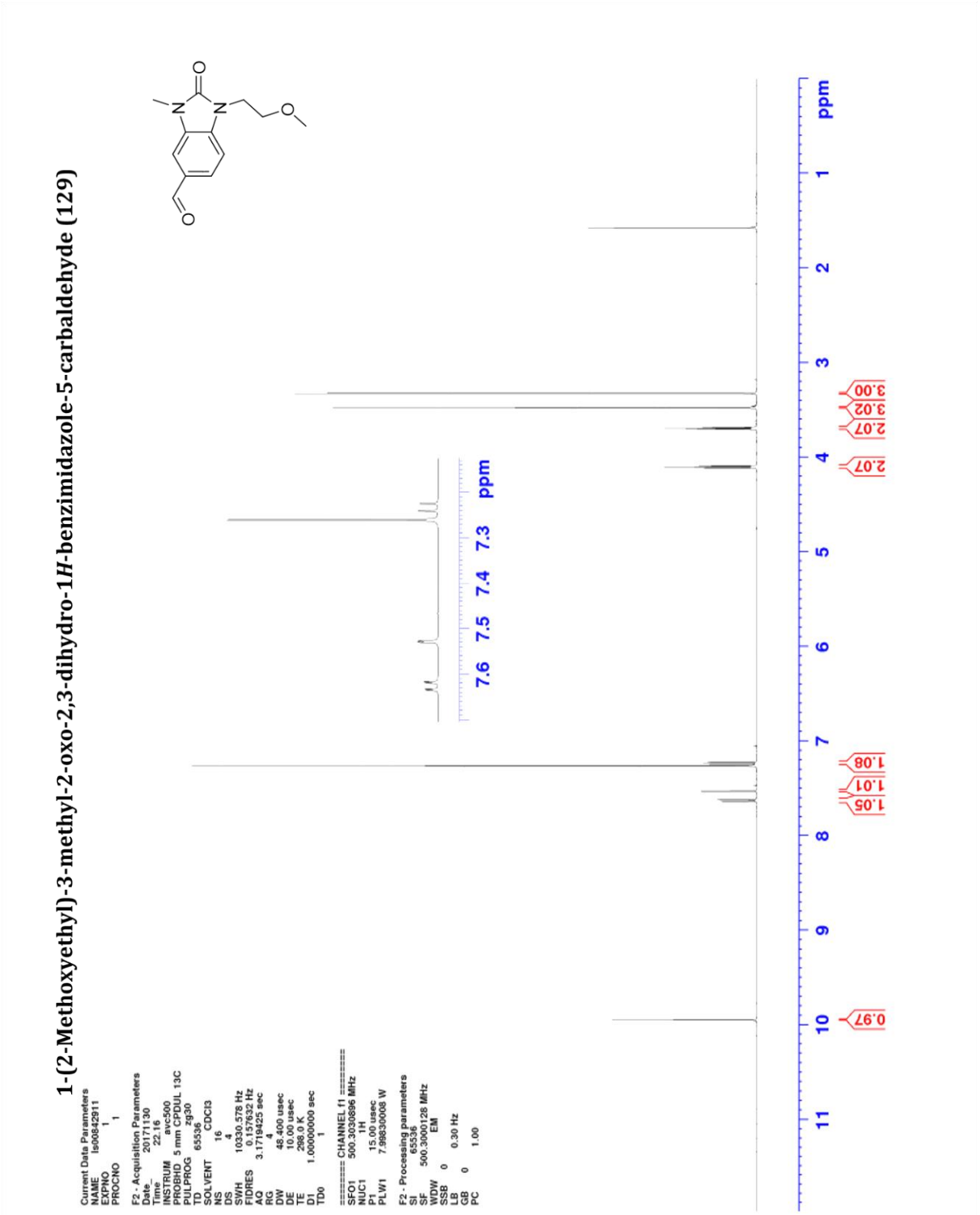
1-Cyclopentyl-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (128)



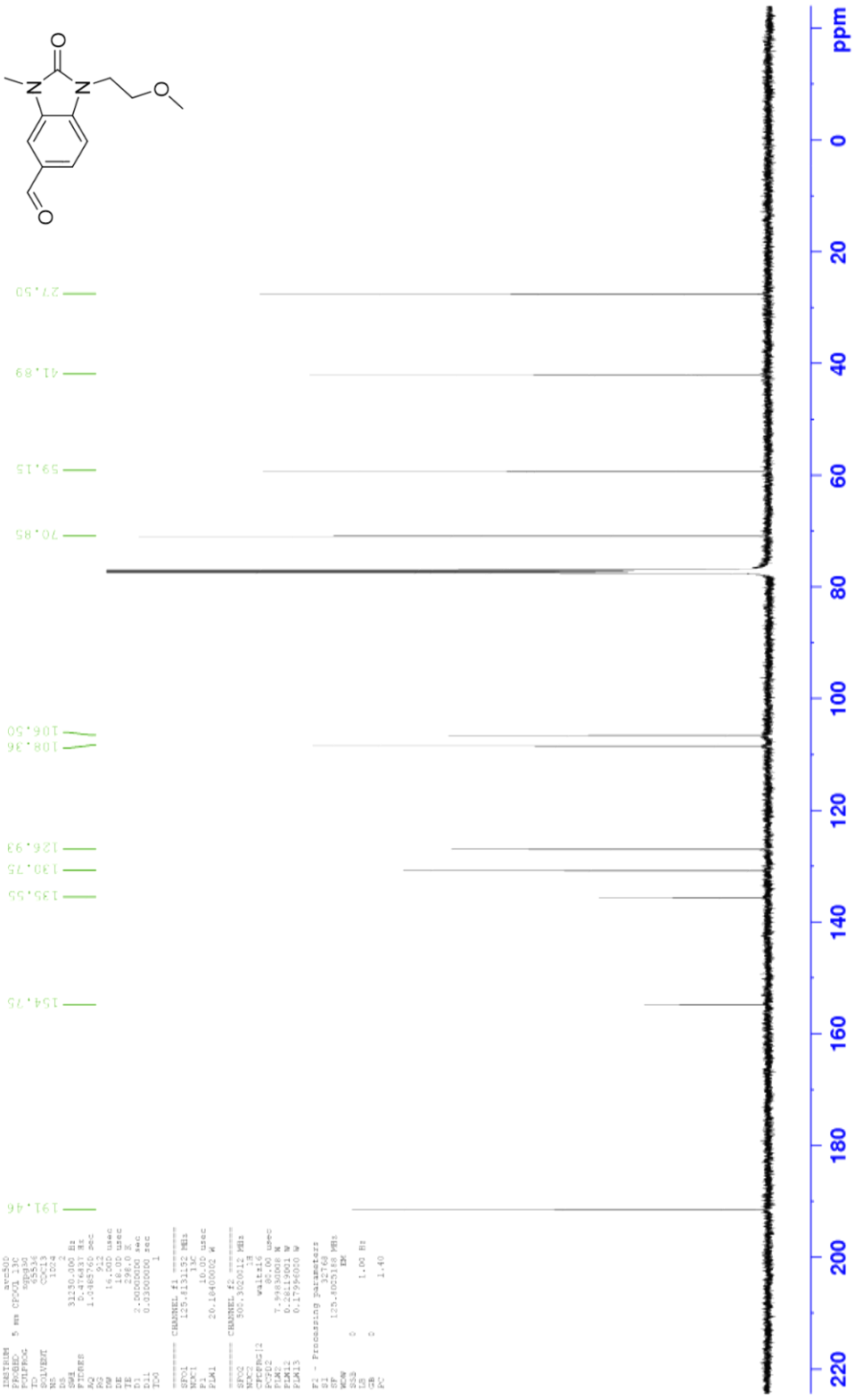
1-Cyclopentyl-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (128)



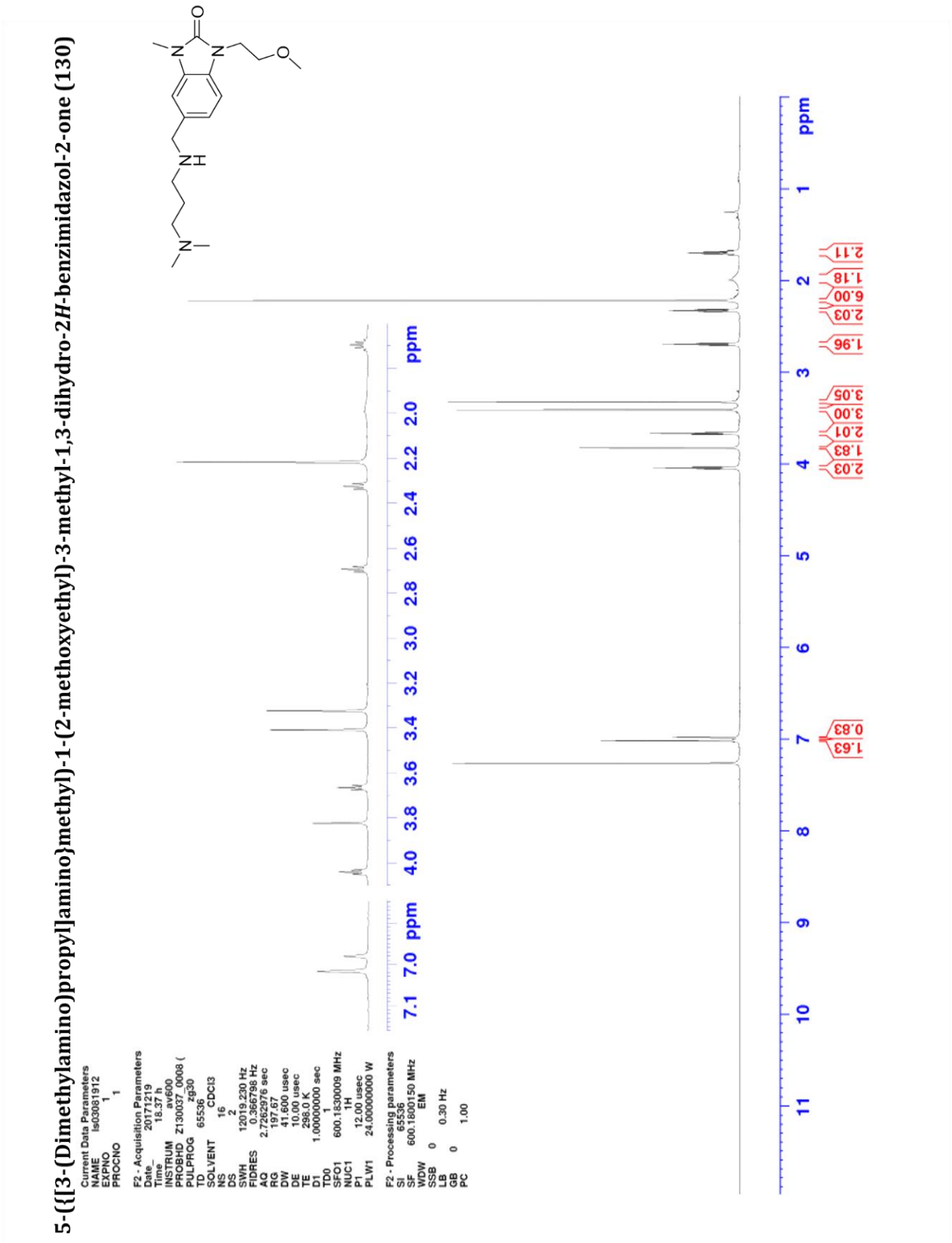
1-(2-Methoxyethyl)-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (129)



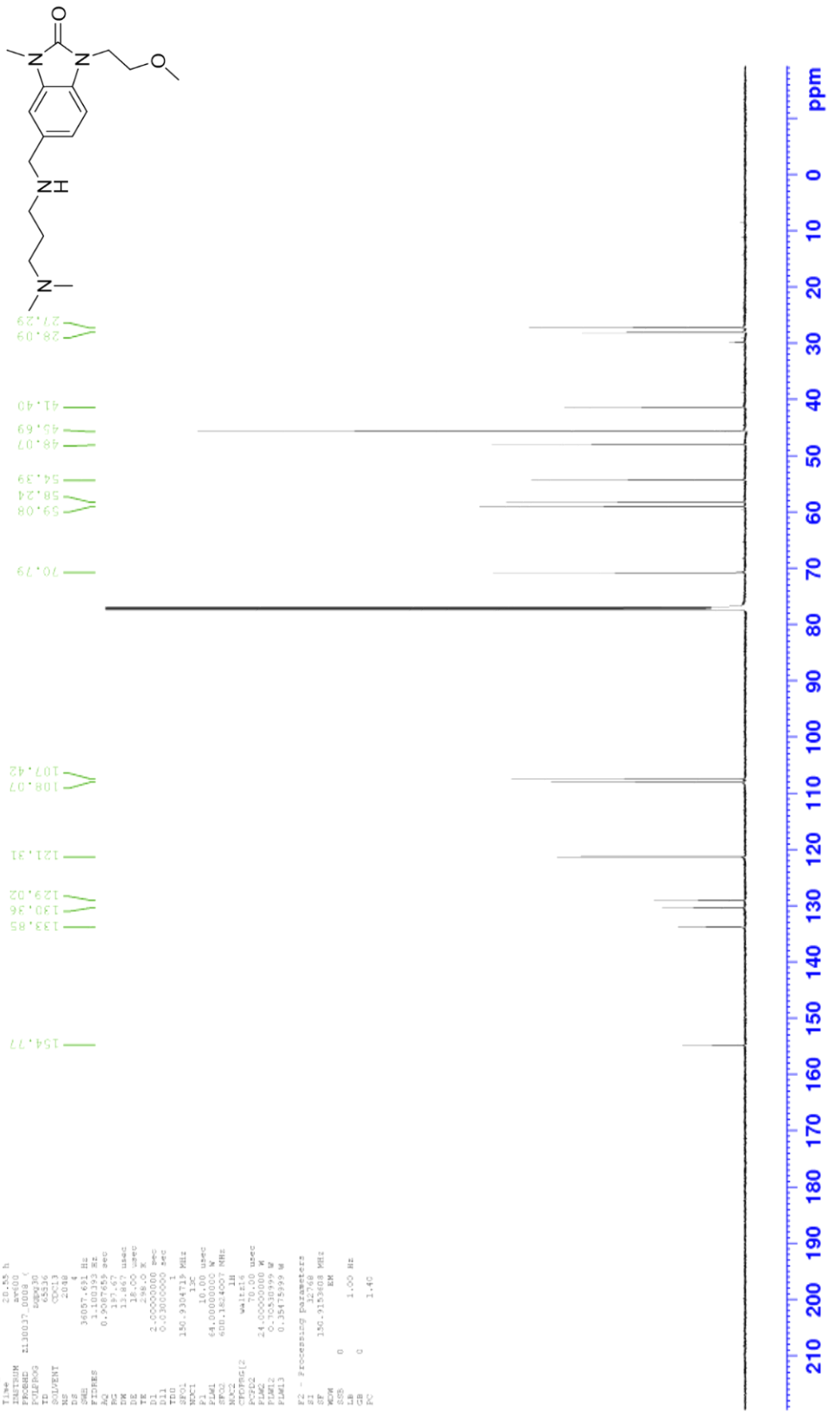
1-(2-Methoxyethyl)-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (129)



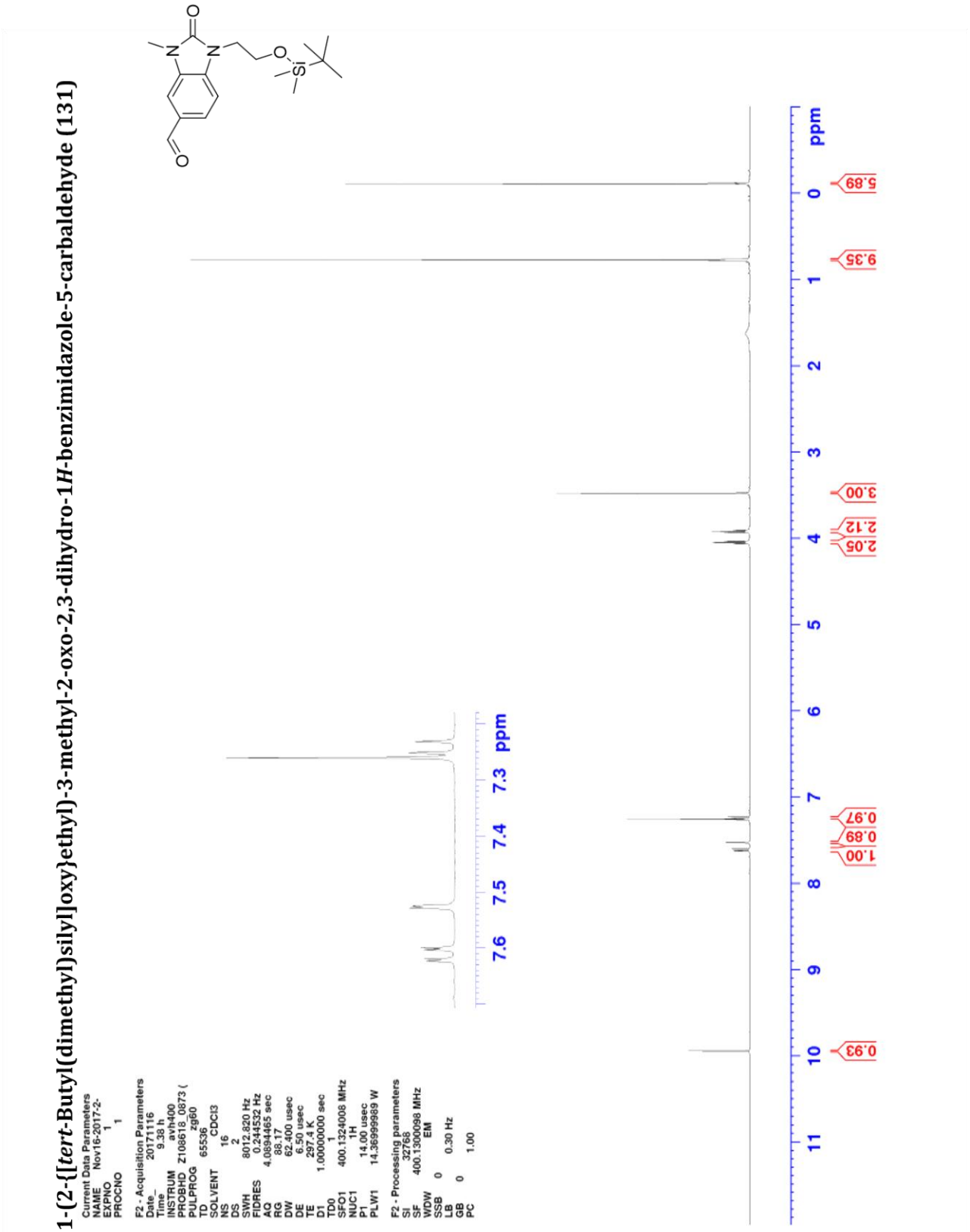
5-({[3-(Dimethylamino)propyl]amino}methyl)-1-(2-methoxyethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (130)



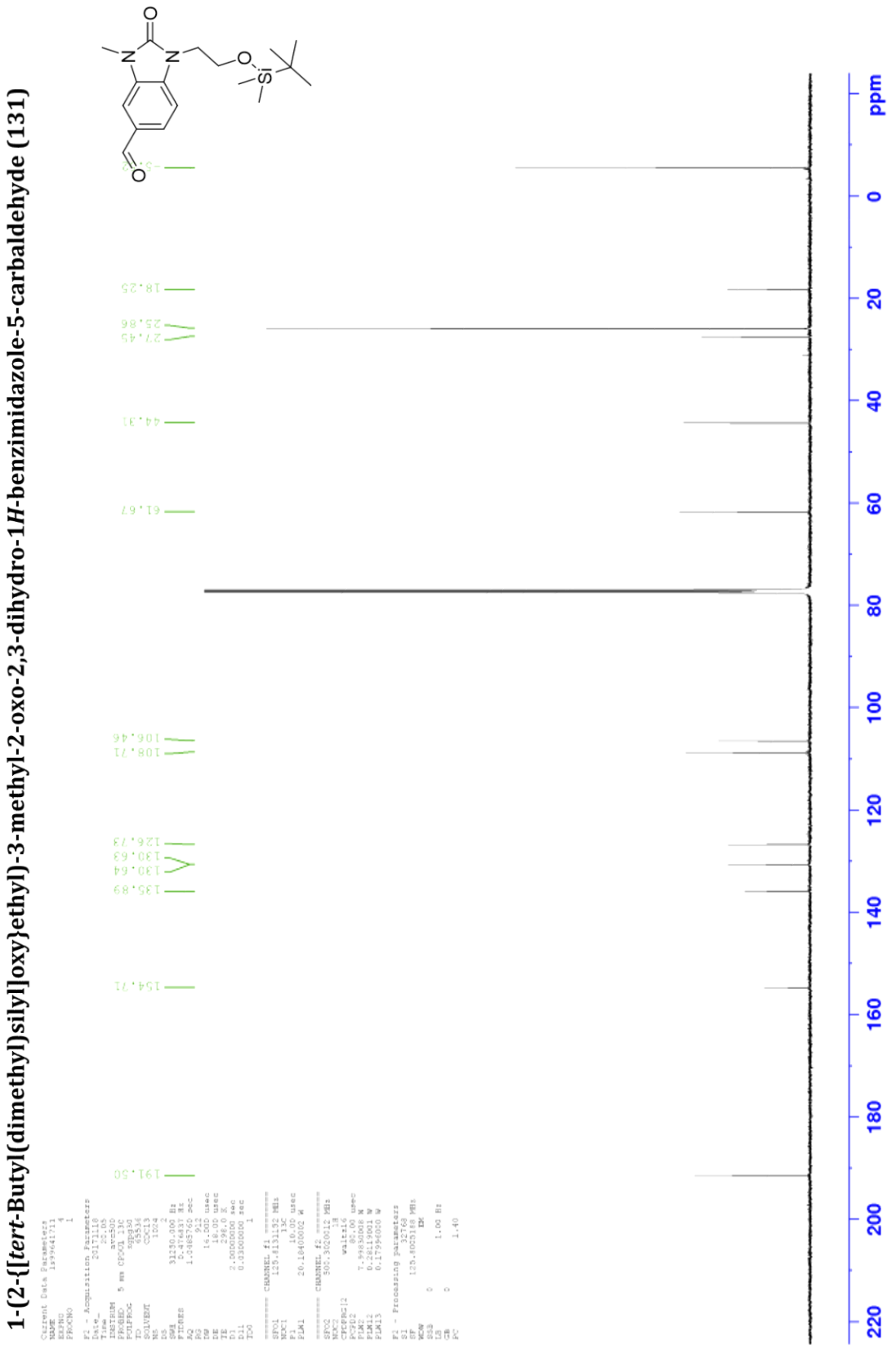
5-({[3-(Dimethylamino)propyl]amino}methyl)-1-(2-methoxyethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (130)



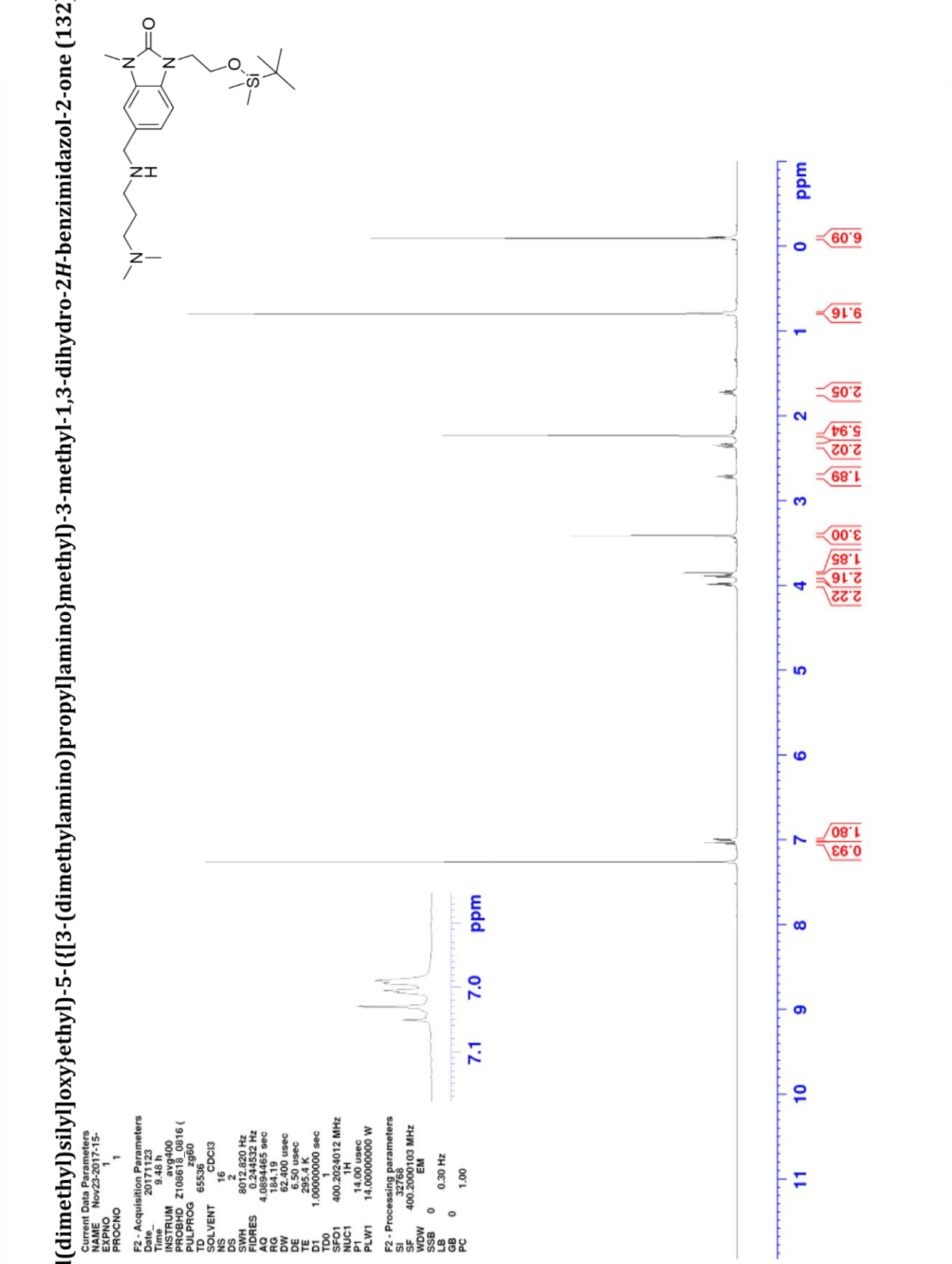
1-(2-{[*tert*-Butyl(dimethyl)silyl]oxy}ethyl)-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (131)



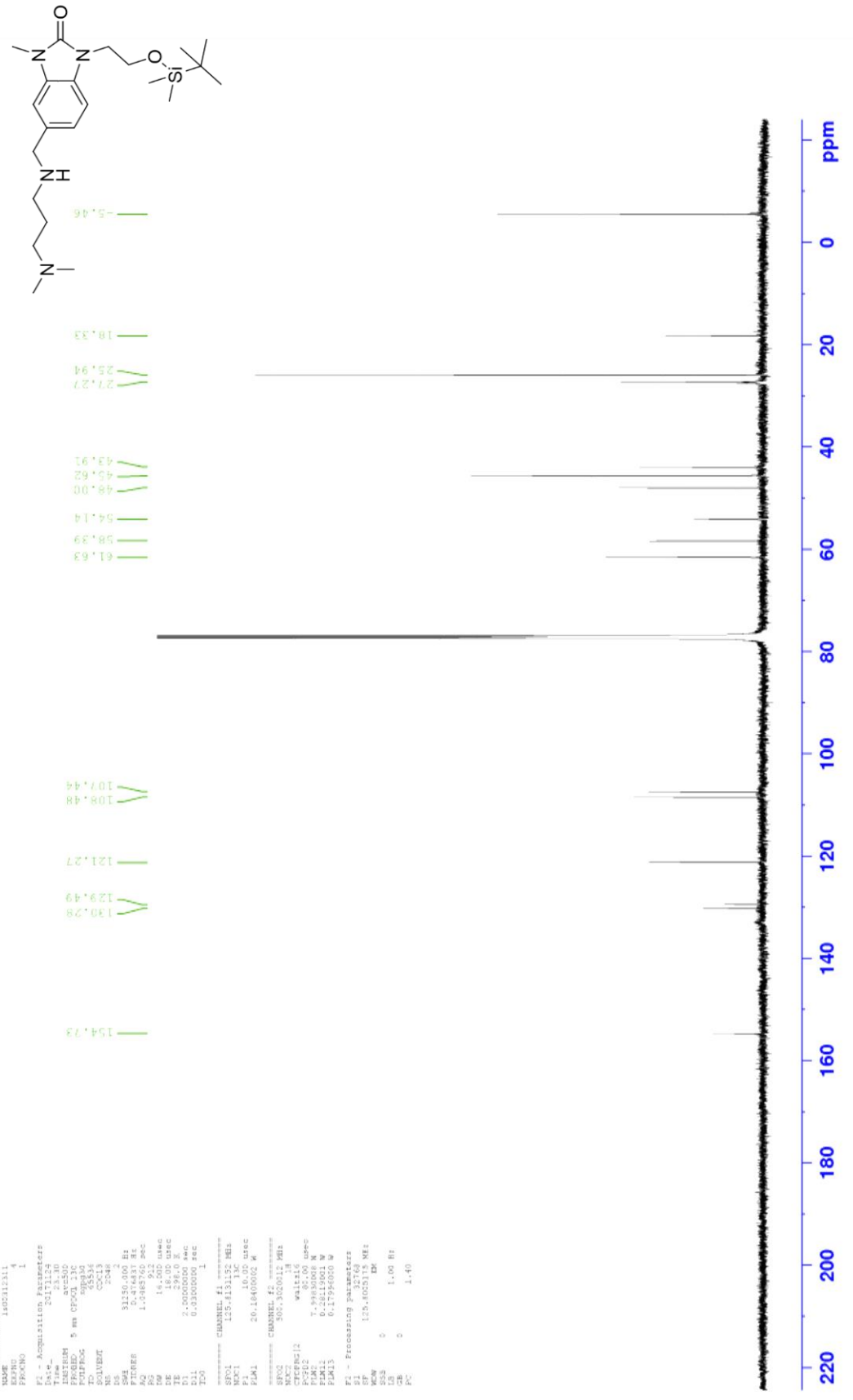
1-(2-{[*tert*-Butyl(dimethyl)silyl]oxy}ethyl)-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carbaldehyde (131)



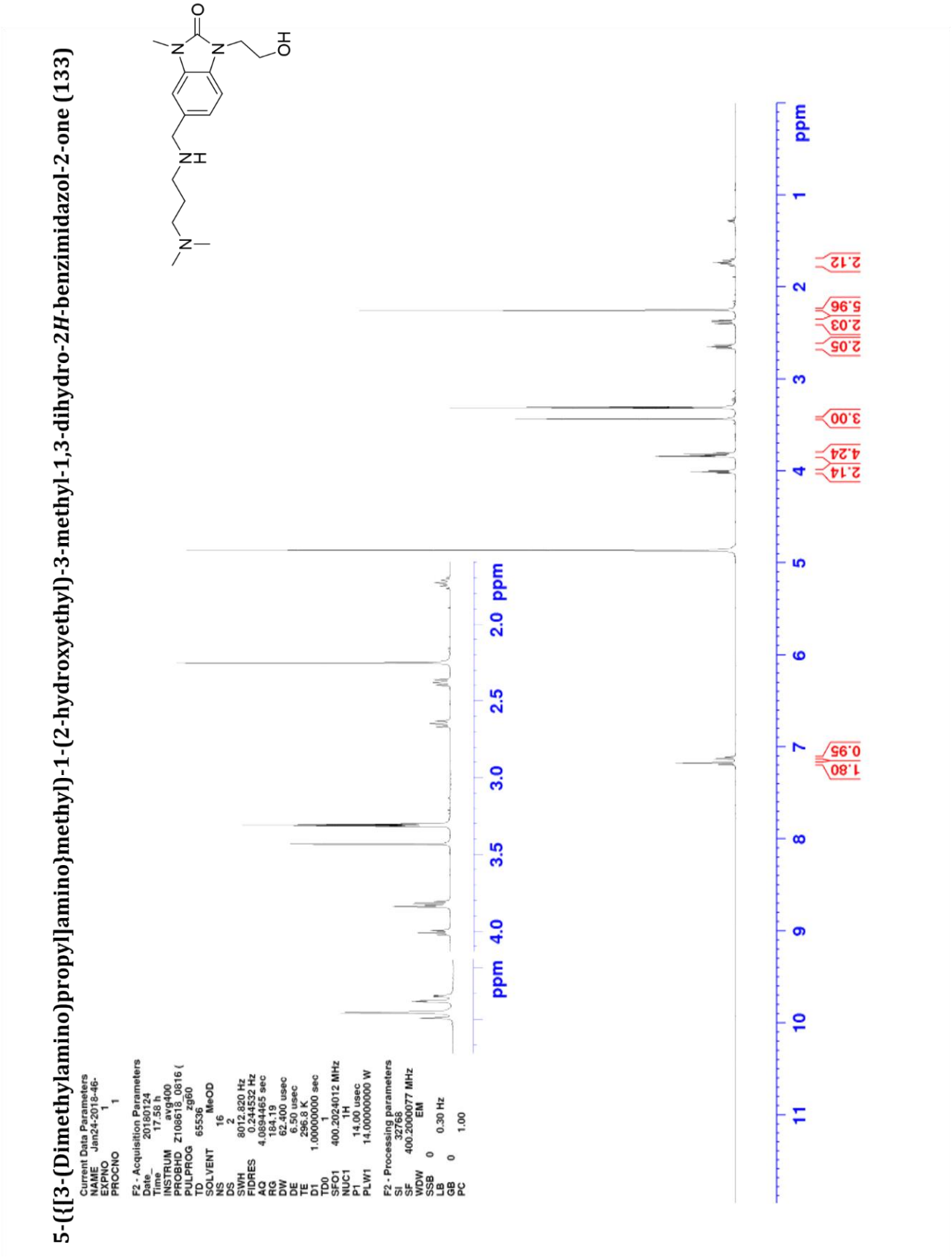
1-(2-{[*tert*-Butyl(dimethyl)silyl]oxy}ethyl)-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (132)



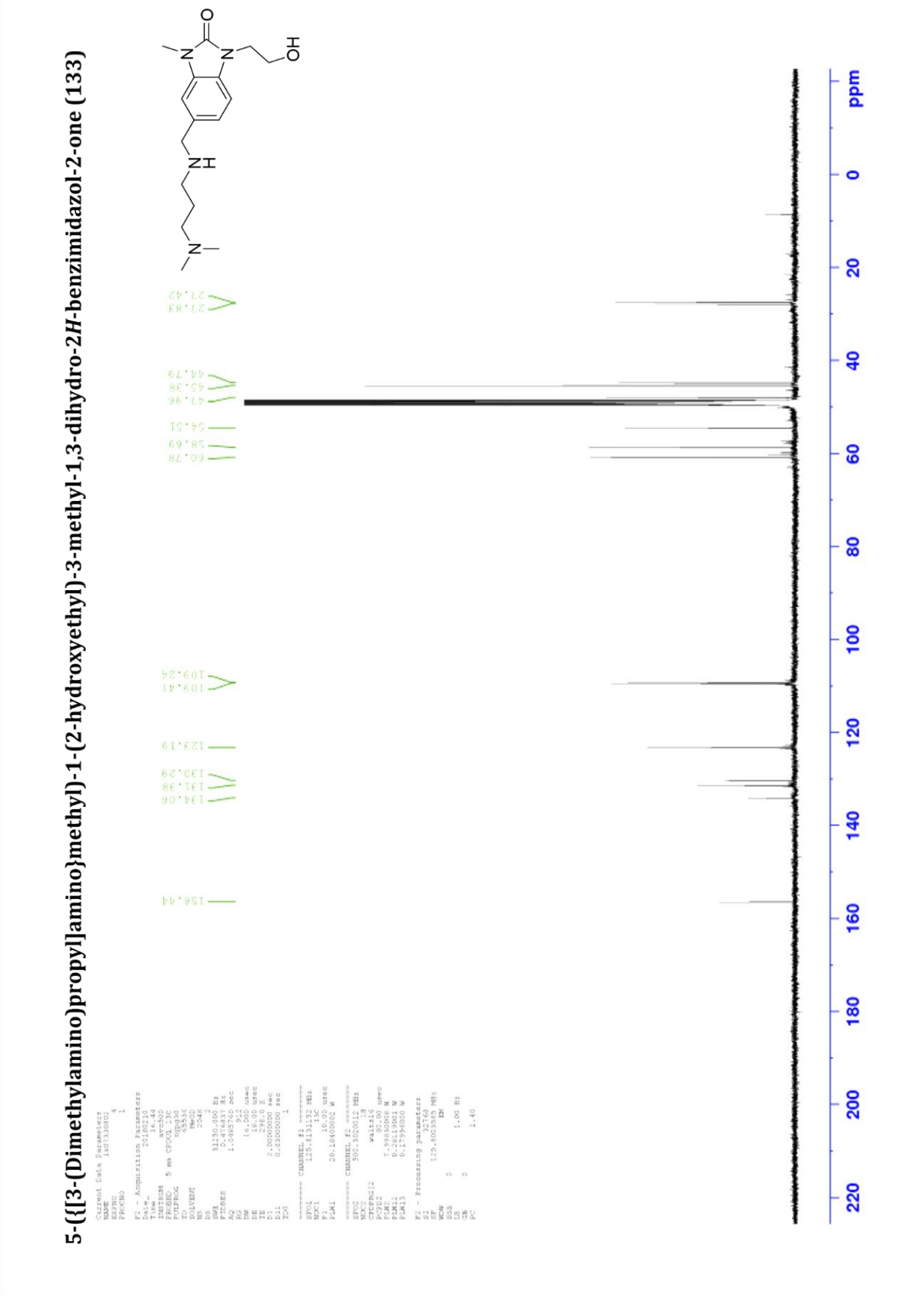
1-(2-{[*tert*-Butyl(dimethyl)silyl]oxy}ethyl)-5-({[3-(dimethylamino)propyl]amino}methyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (132)



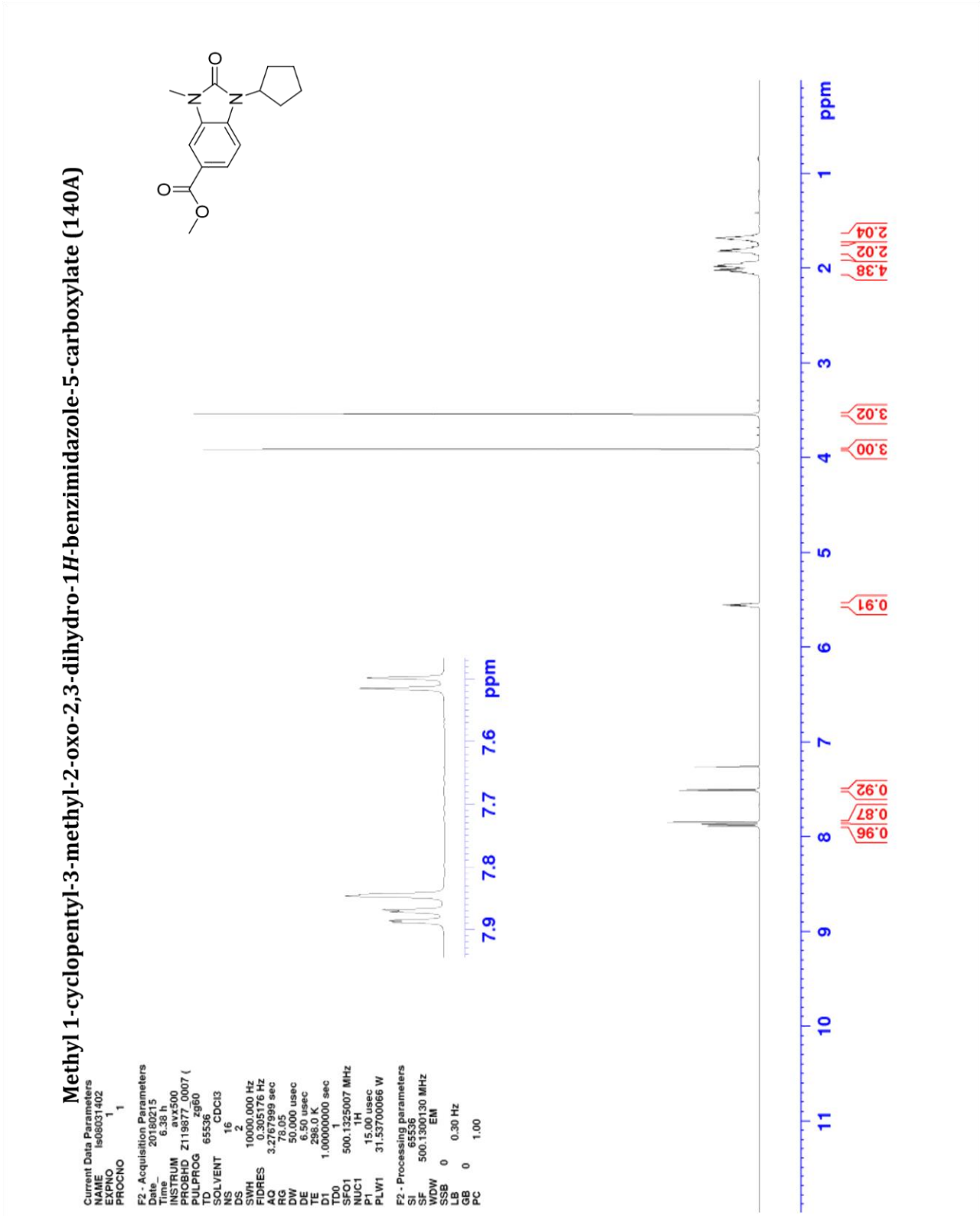
5-({[3-(Dimethylamino)propyl]amino}methyl)-1-(2-hydroxyethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (133)



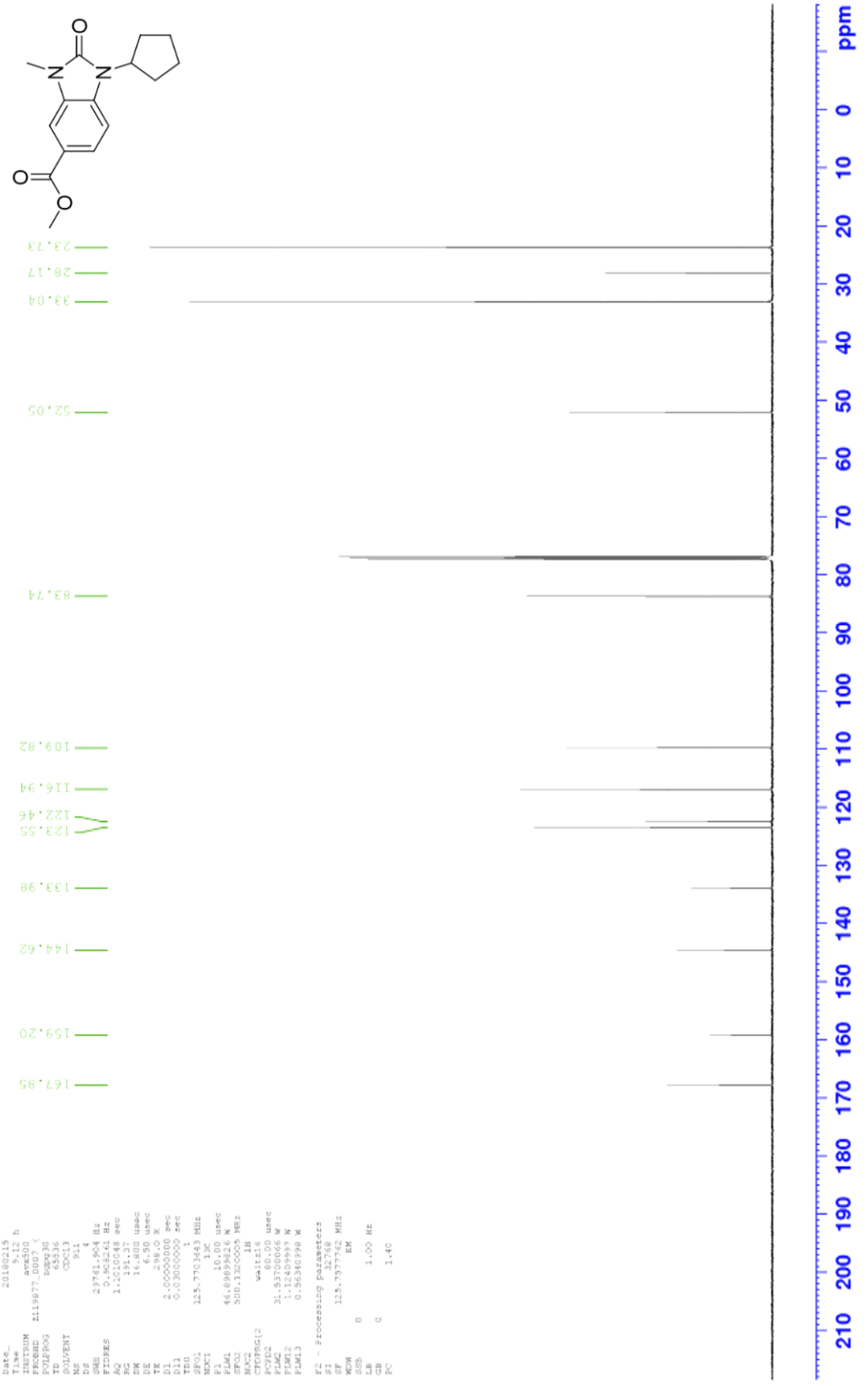
5-({[3-(Dimethylamino)propyl]amino}methyl)-1-(2-hydroxyethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (133)



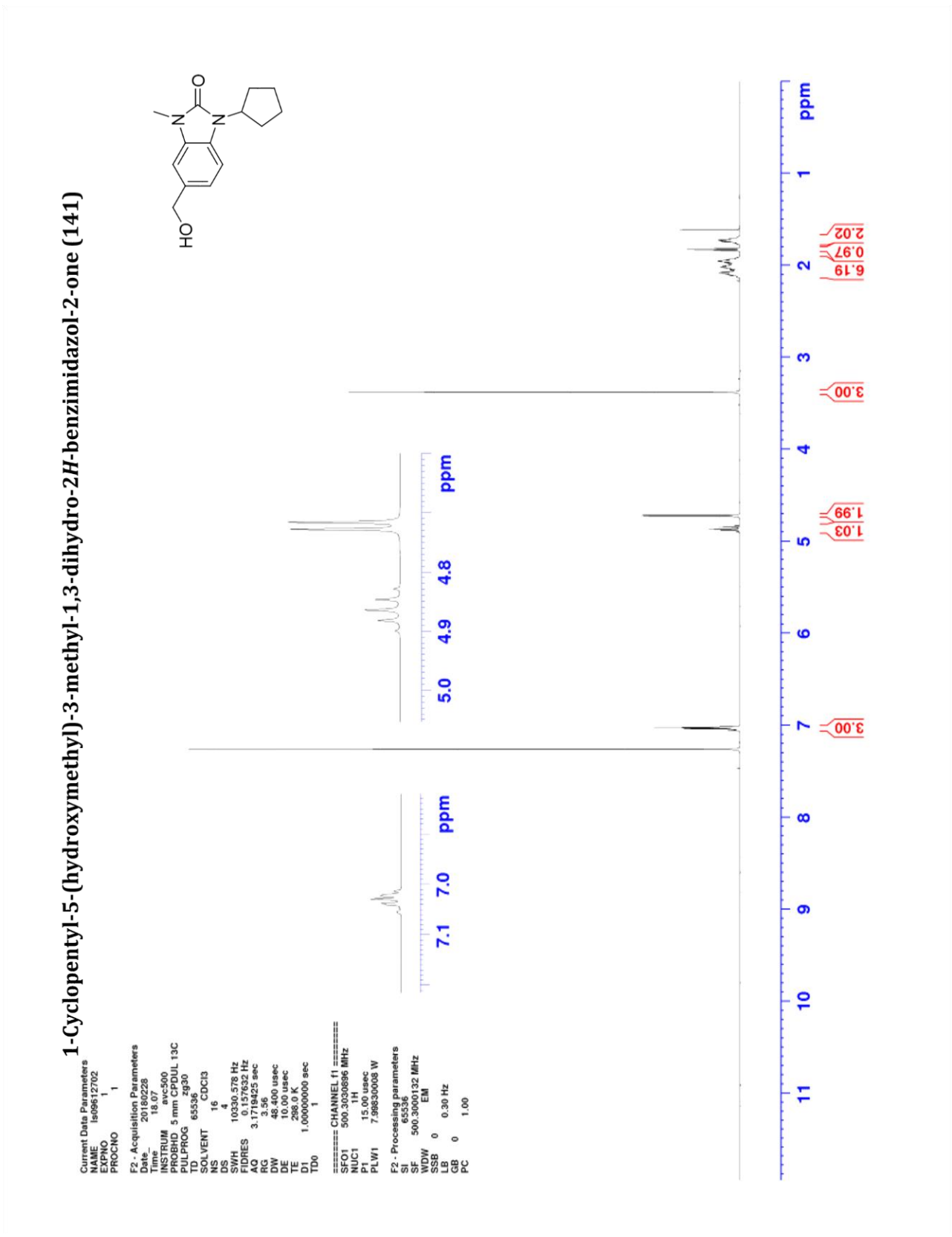
Methyl 1-cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (140A)



Methyl 1-cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazole-5-carboxylate (140A)



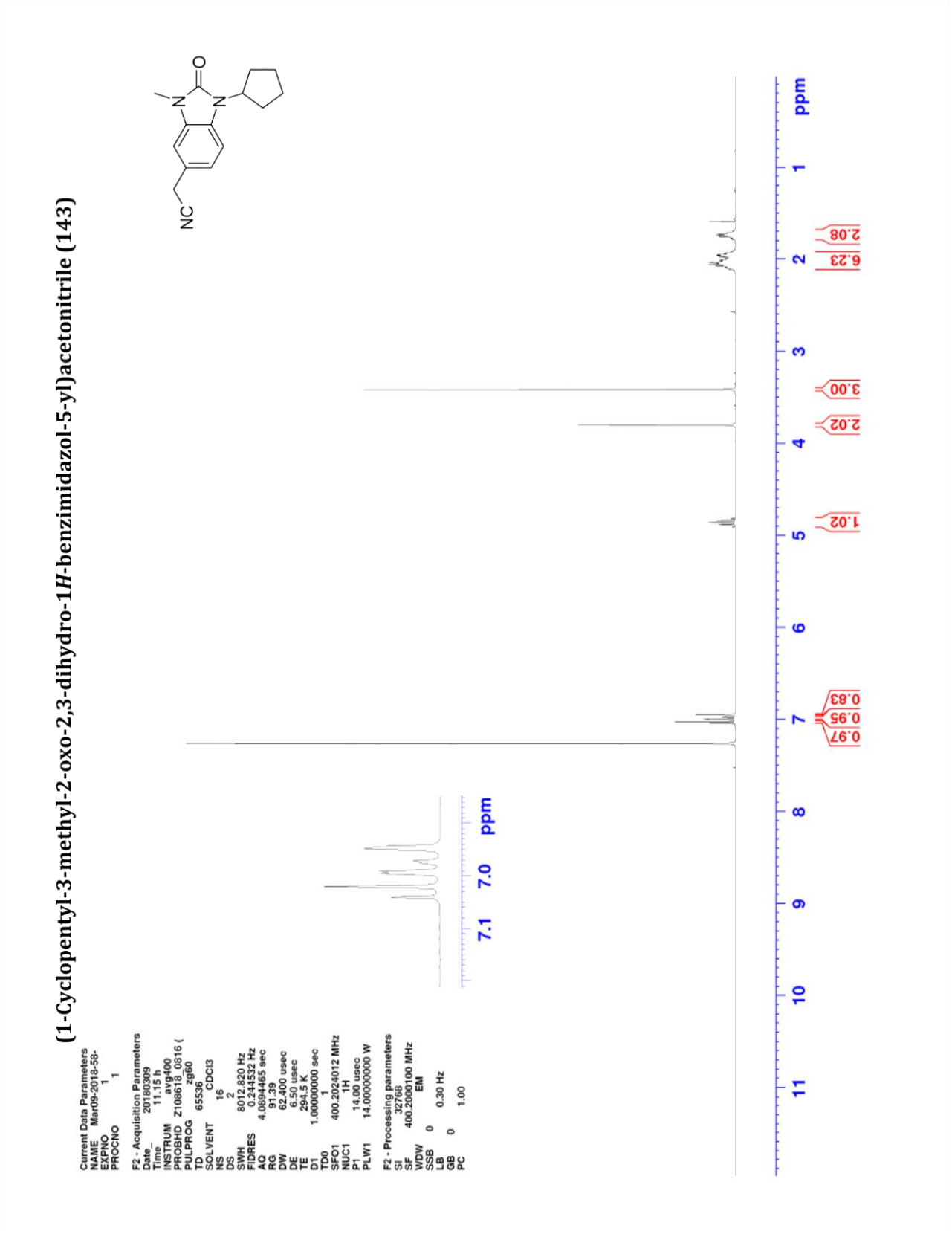
1-Cyclopentyl-5-(hydroxymethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (141)



1-Cyclopentyl-5-(hydroxymethyl)-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (141)



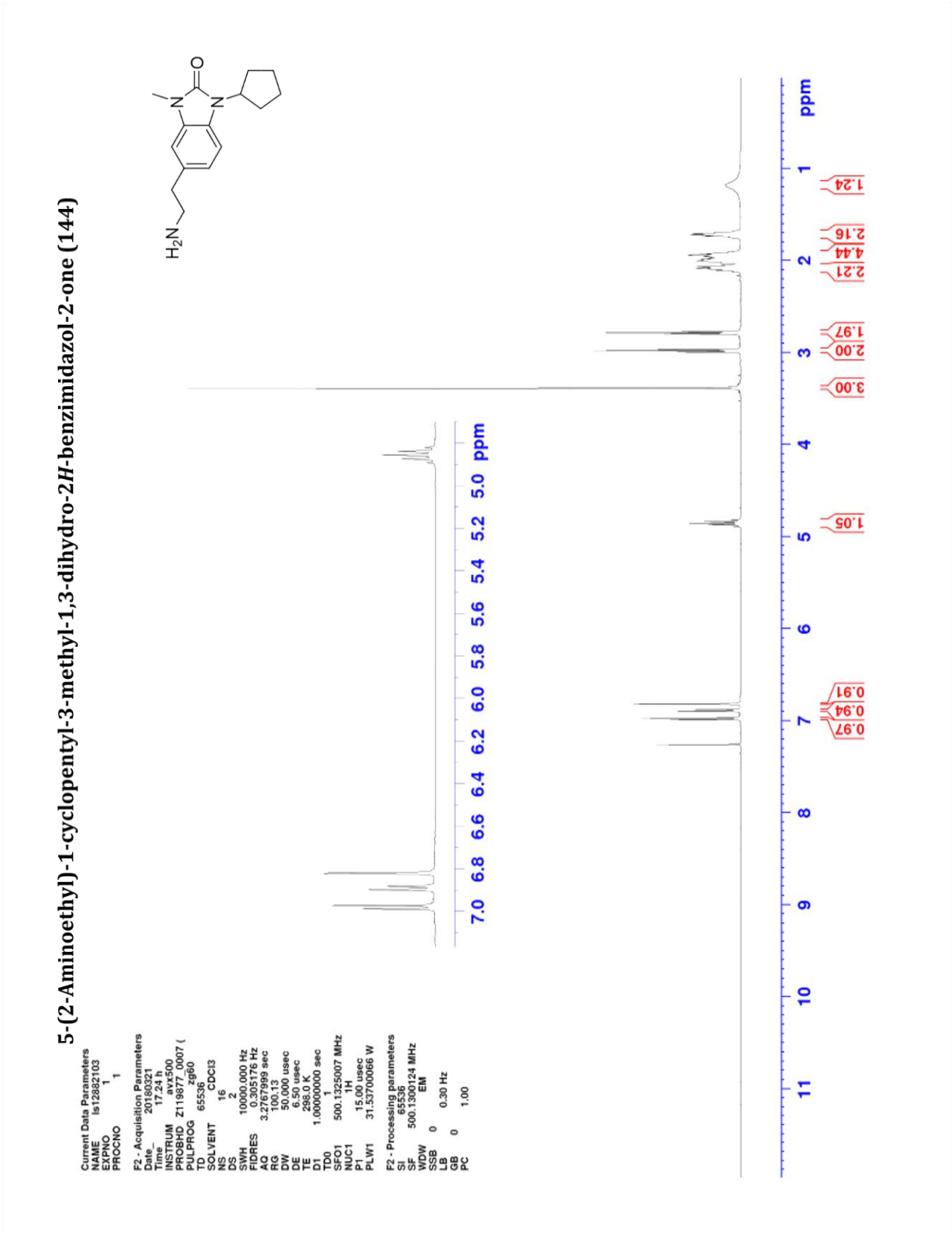
(1-Cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)acetonitrile (143)



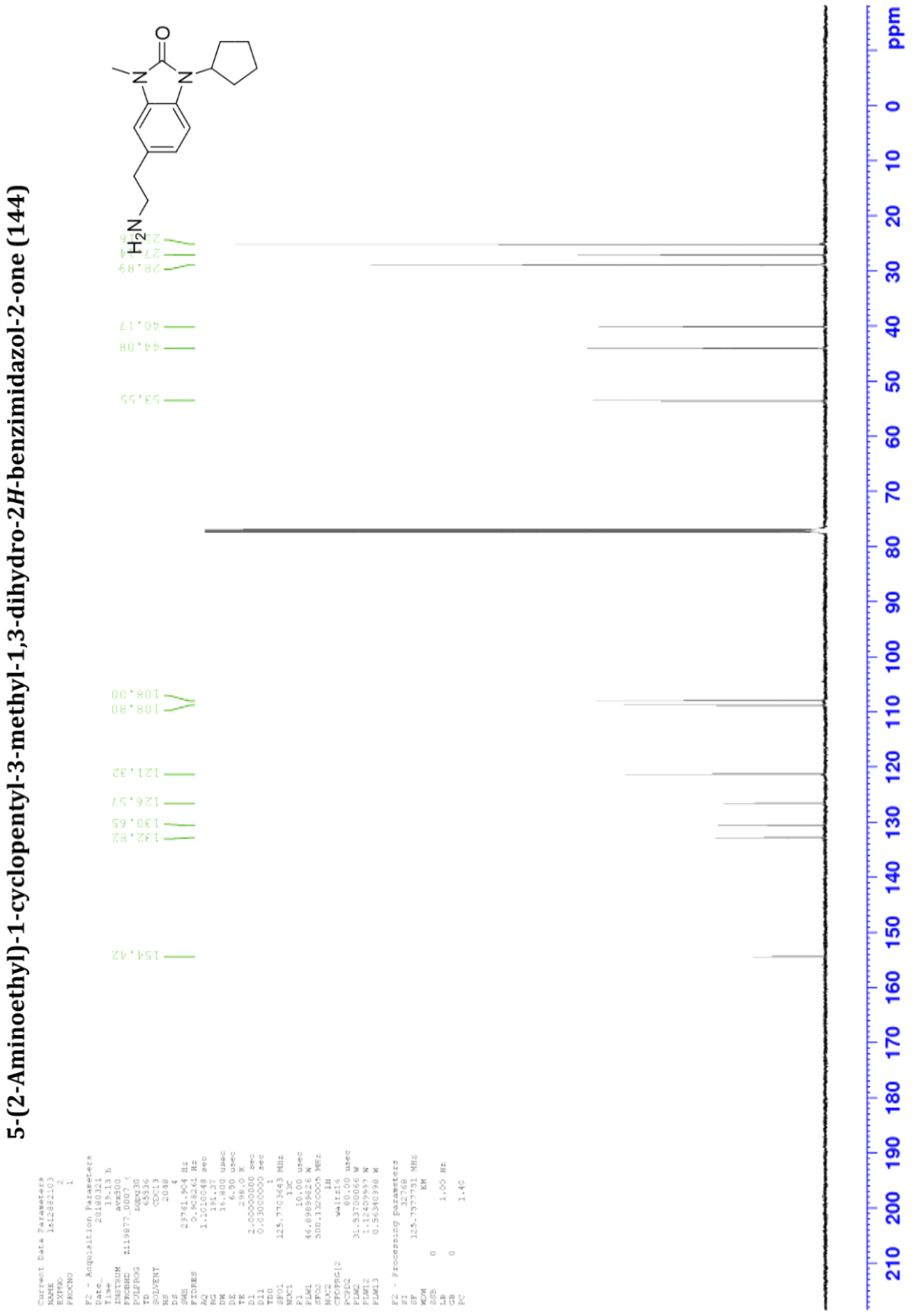
(1-Cyclopentyl-3-methyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)acetonitrile (143)

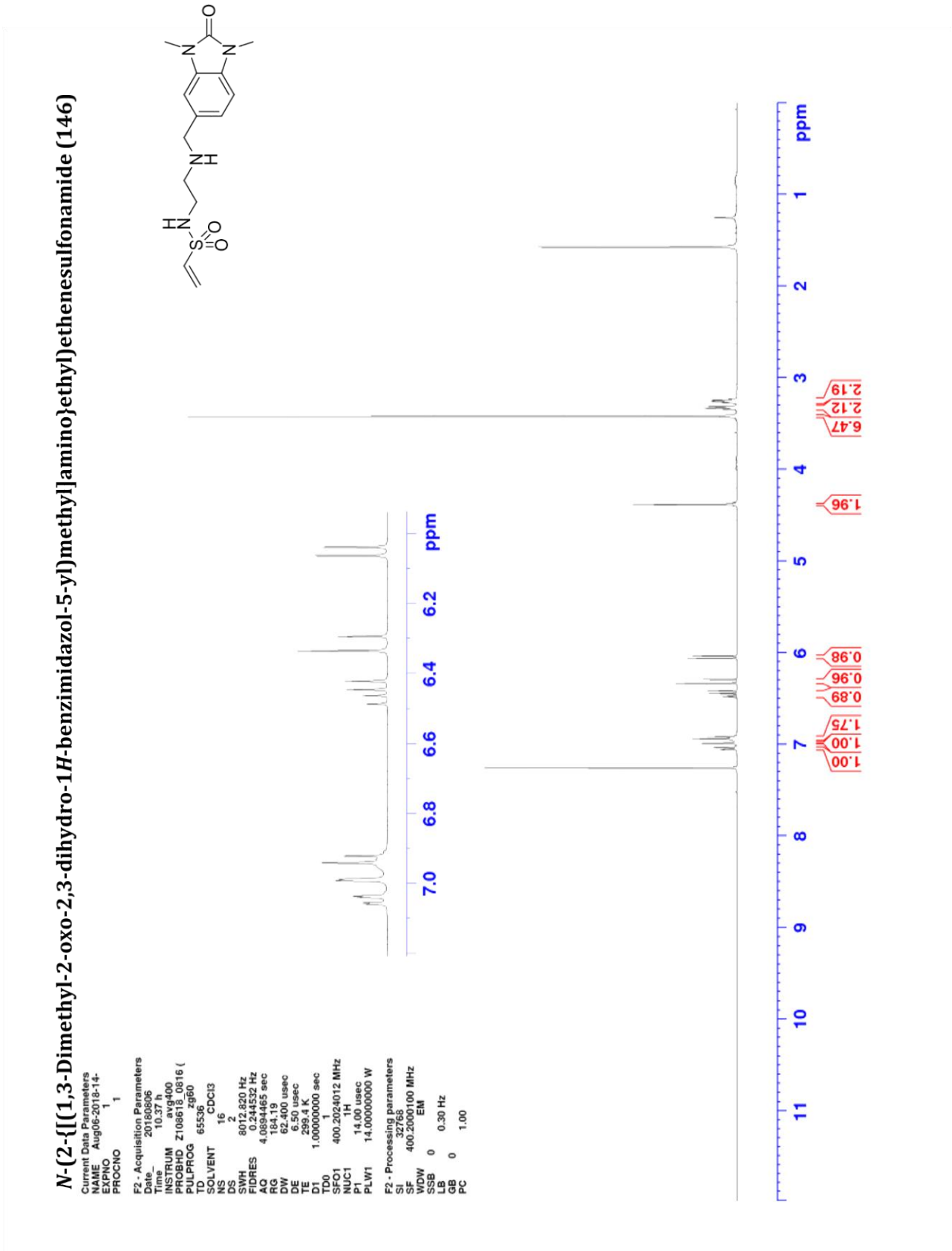


5-(2-Aminoethyl)-1-cyclopentyl-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (144)

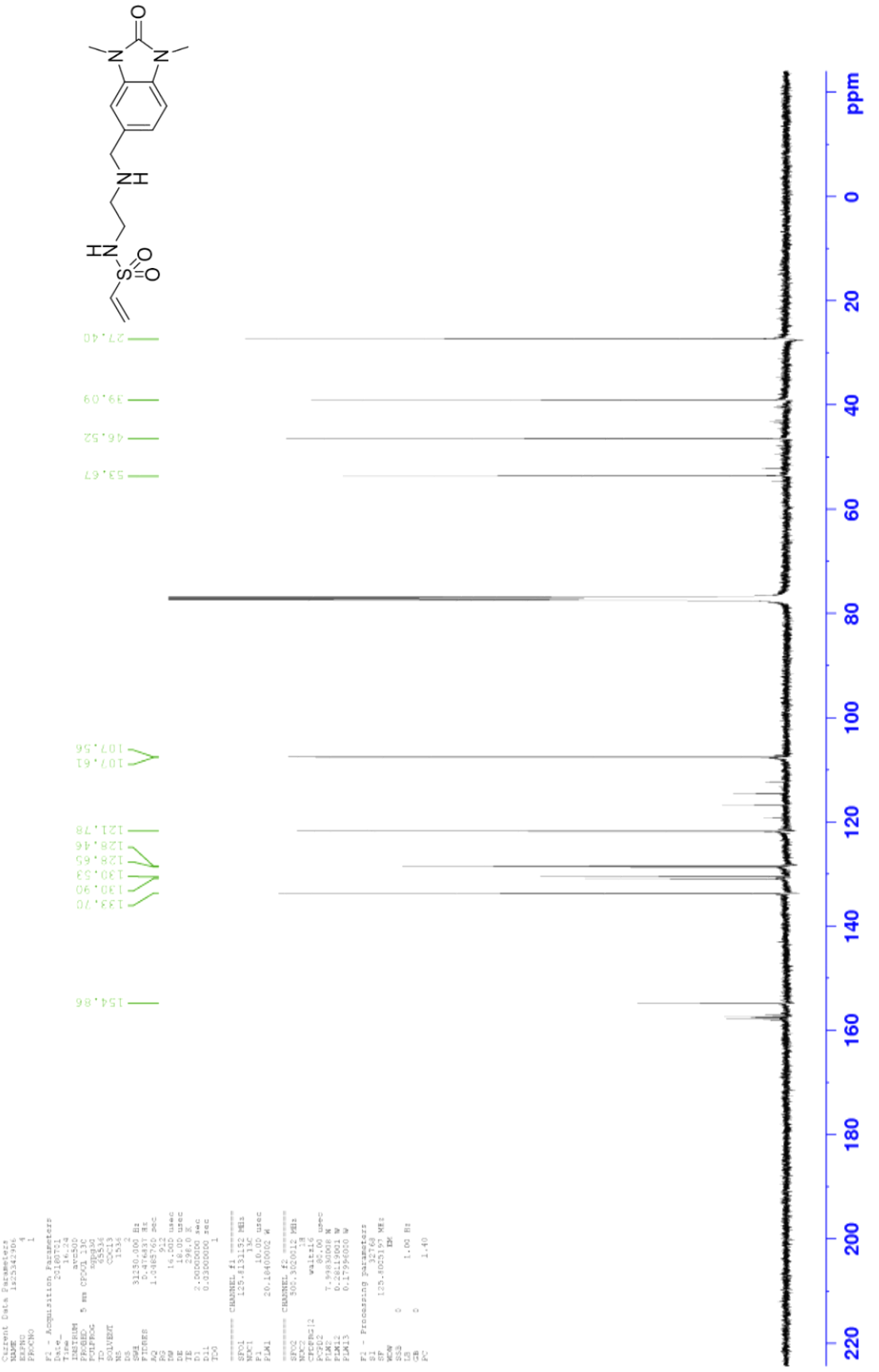


5-(2-Aminoethyl)-1-cyclopentyl-3-methyl-1,3-dihydro-2*H*-benzimidazol-2-one (144)

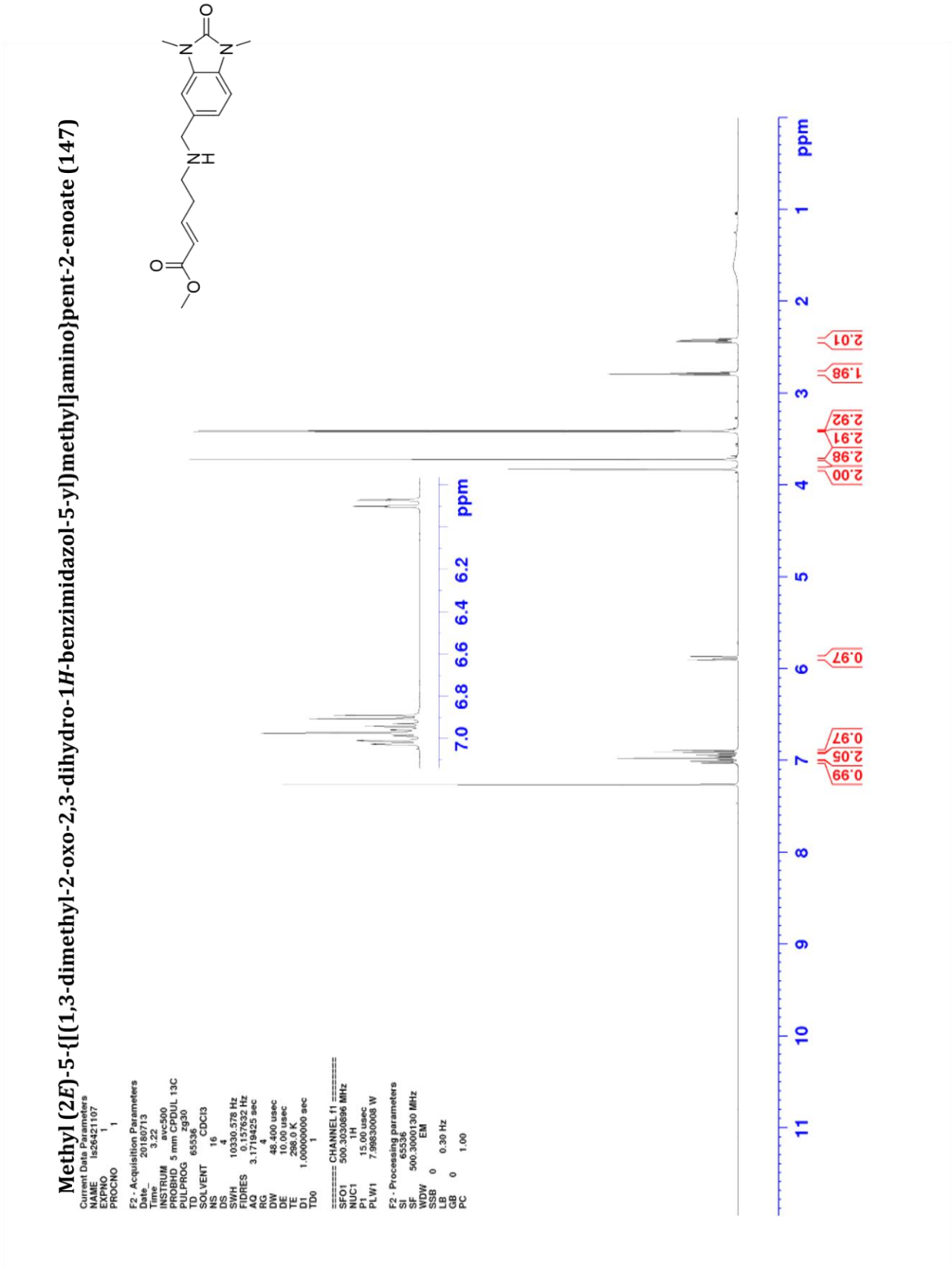



N-(2-{[(1,3-Dimethyl-2-oxo-2,3-dihydro-1H-benzimidazol-5-yl)methyl]amino}ethyl)ethenesulfonamide (146)
Current Data Parameters
NAME Aug06-2018-14-
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20180806
Time 10.37 h
INSTRUM avg400
PROBHD Z108618_0816 (
PULPROG zg60
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8012.820 Hz
FIDRES 0.244532 Hz
AQ 4.0894465 sec
RG 184.19
DW 62.400 usec
DE 6.50 usec
TE 299.4 K
D1 1.00000000 sec
TD0 1
SFO1 400.2024012 MHz
NUC1 1H
P1 14.00 usec
PLW1 14.00000000 W
F2 - Processing parameters
SI 32768
SF 400.2000100 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.0
6.8
6.6
6.4
6.2
ppm
1.96
6.47
2.12
2.19
1.00
1.00
1.75
0.89
0.96
0.98
ppm

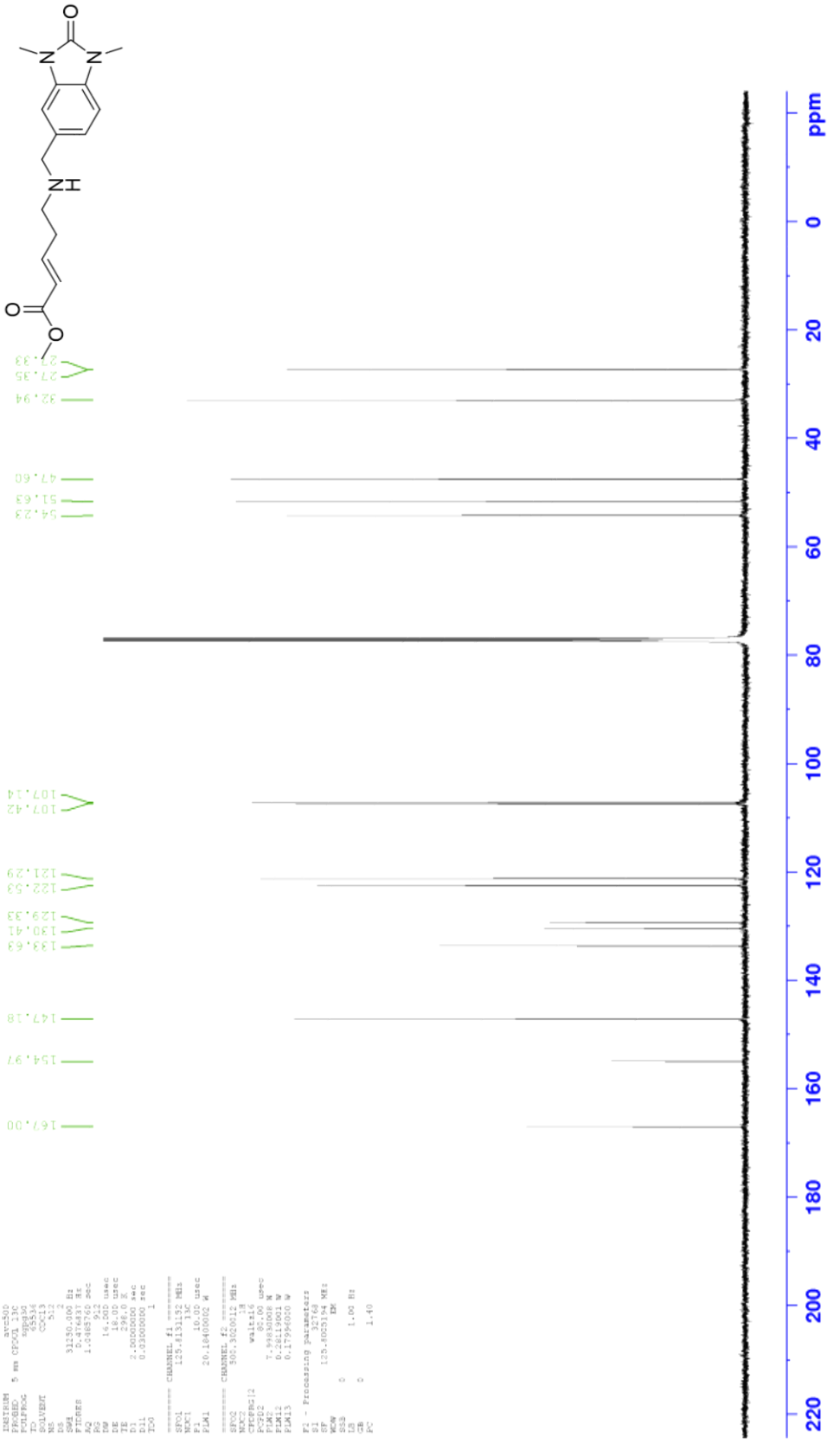
***N*-(2-{[(1,3-Dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}ethyl)ethenesulfonamide (146)**

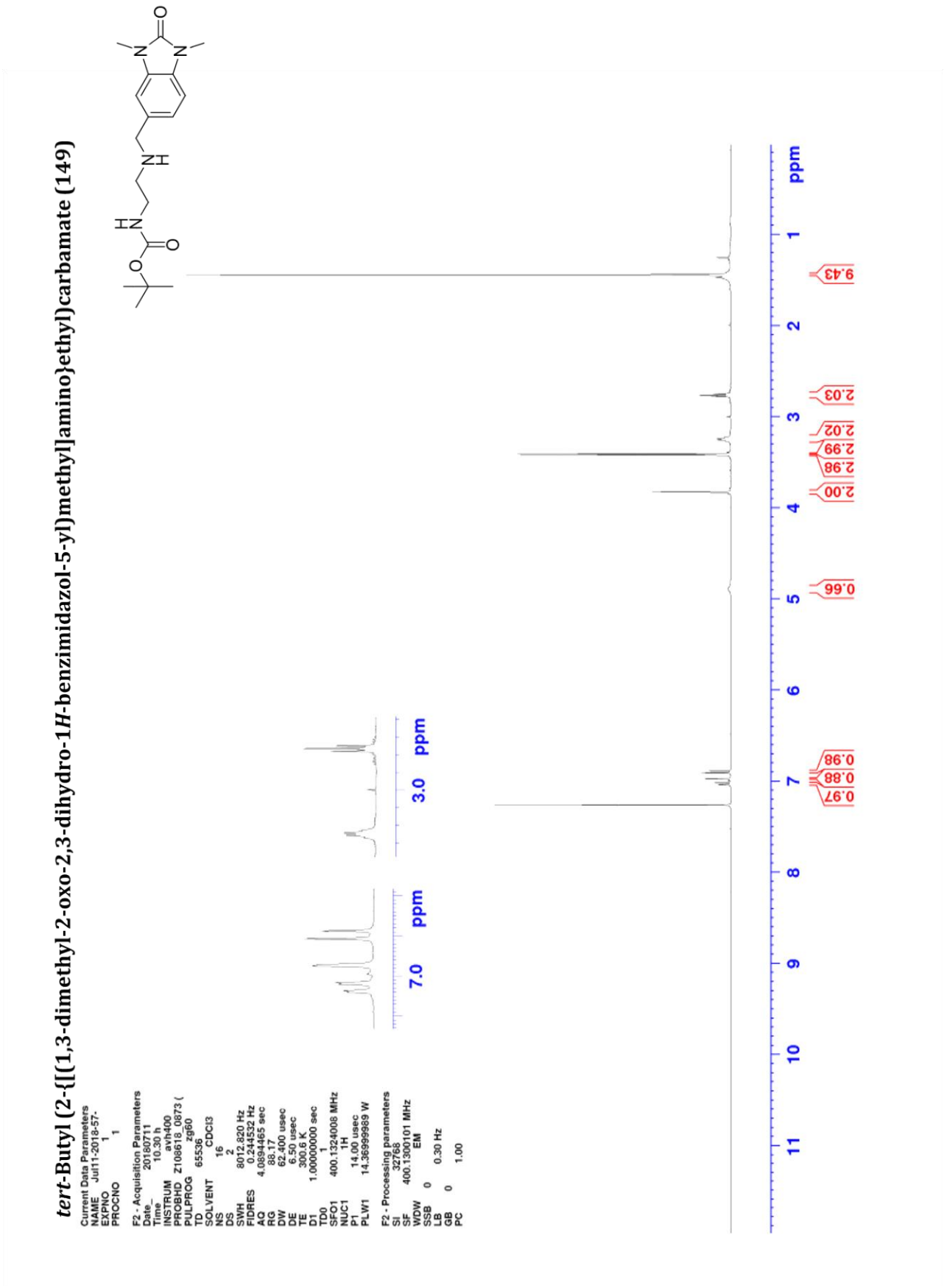


Methyl (2*E*)-5-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}pent-2-enoate (147)

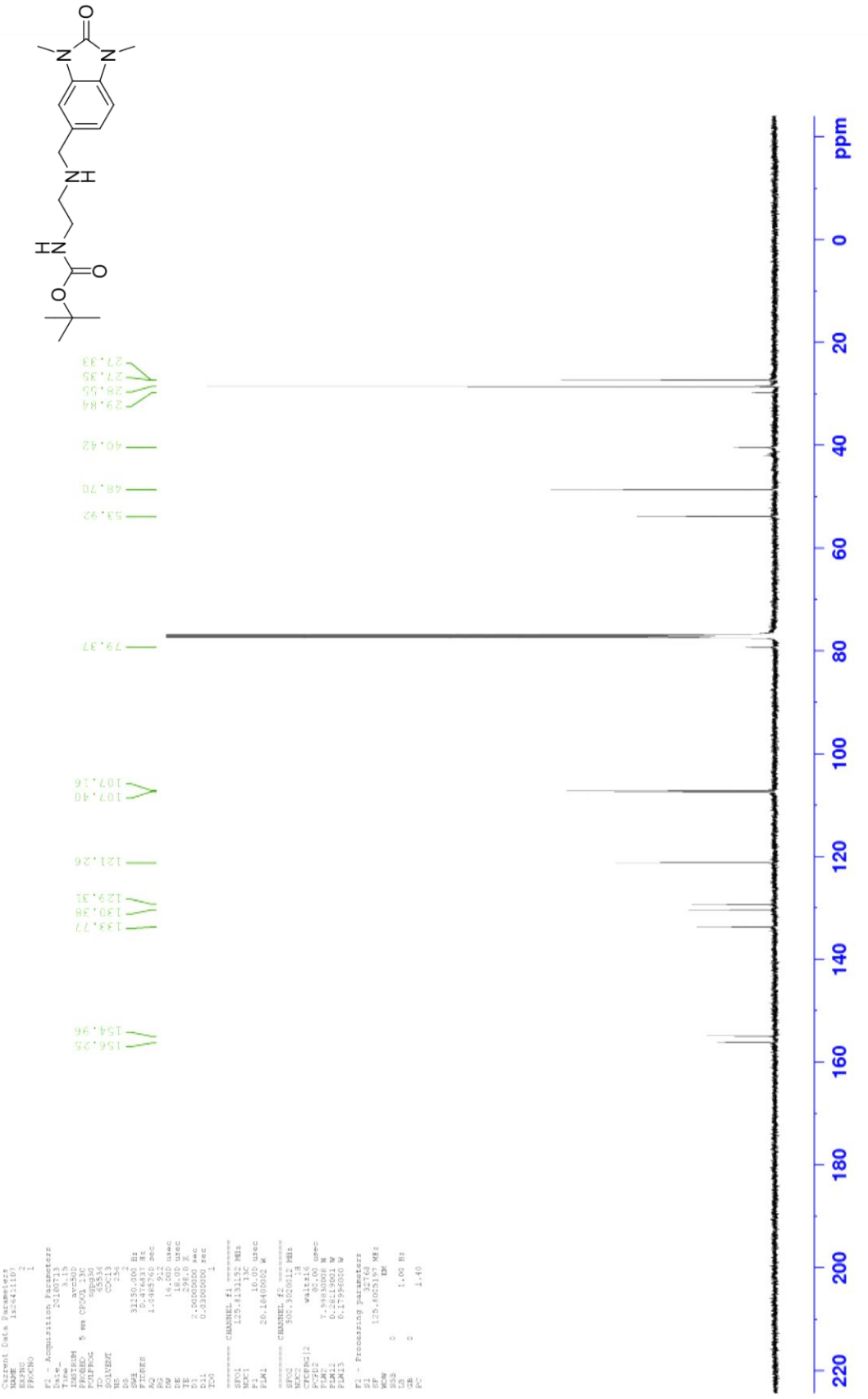


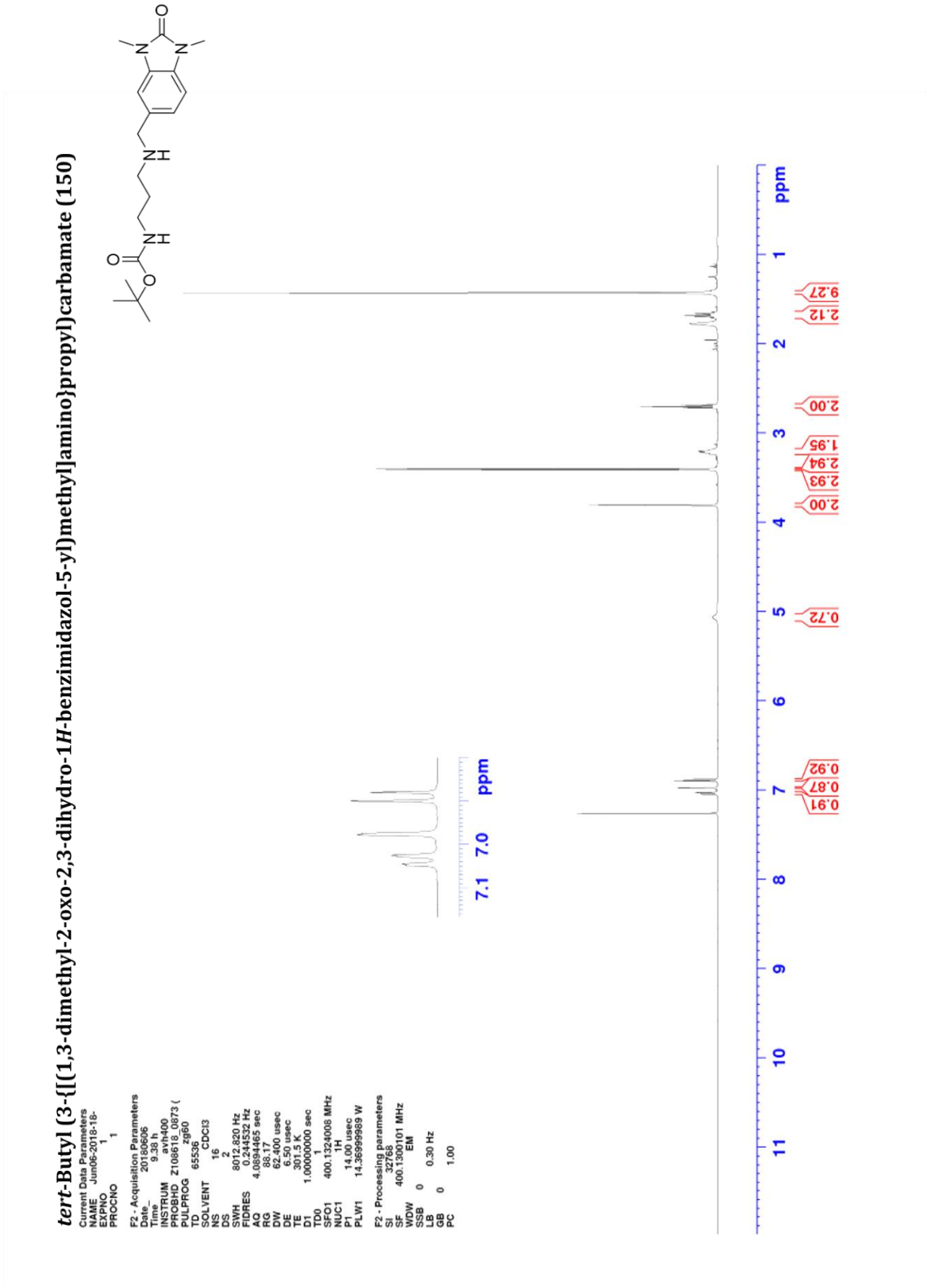
Methyl (2*E*)-5-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}pent-2-enoate (147)



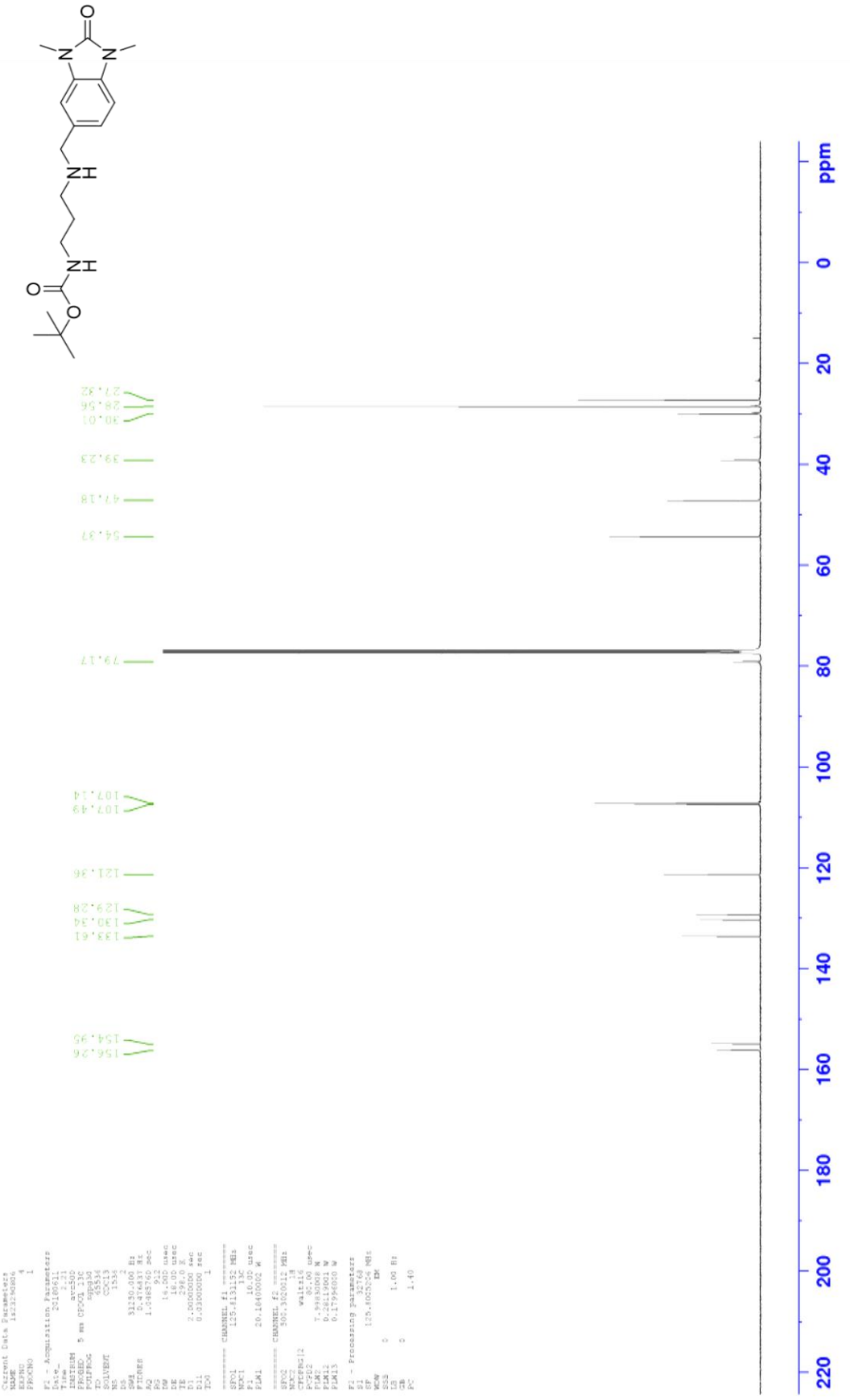

tert-Butyl (2-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1H-benzimidazol-5-yl)methyl]amino}ethyl)carbamate (149)
Current Data Parameters
NAME Jul11-2018-57-
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20180711
Time 10.30 h
INSTRUM avh400
PROBHD Z108618_0873 (
PULPROG zg60
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8012.820 Hz
FIDRES 0.244532 Hz
AQ 4.0894465 sec
RG 88.17
DW 62.400 usec
DE 6.50 usec
TE 300.6 K
D1 1.00000000 sec
TD0 1
SFO1 400.1324008 MHz
NUC1 1H
P1 14.00 usec
PLW1 14.36999989 W
F2 - Processing parameters
SI 32768
SF 400.1300101 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.0 ppm
3.0 ppm
9.43
2.03
2.02
2.99
2.98
2.00
0.66
0.98
0.88
0.97
ppm

***tert*-Butyl (2-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}ethyl)carbamate (149)**

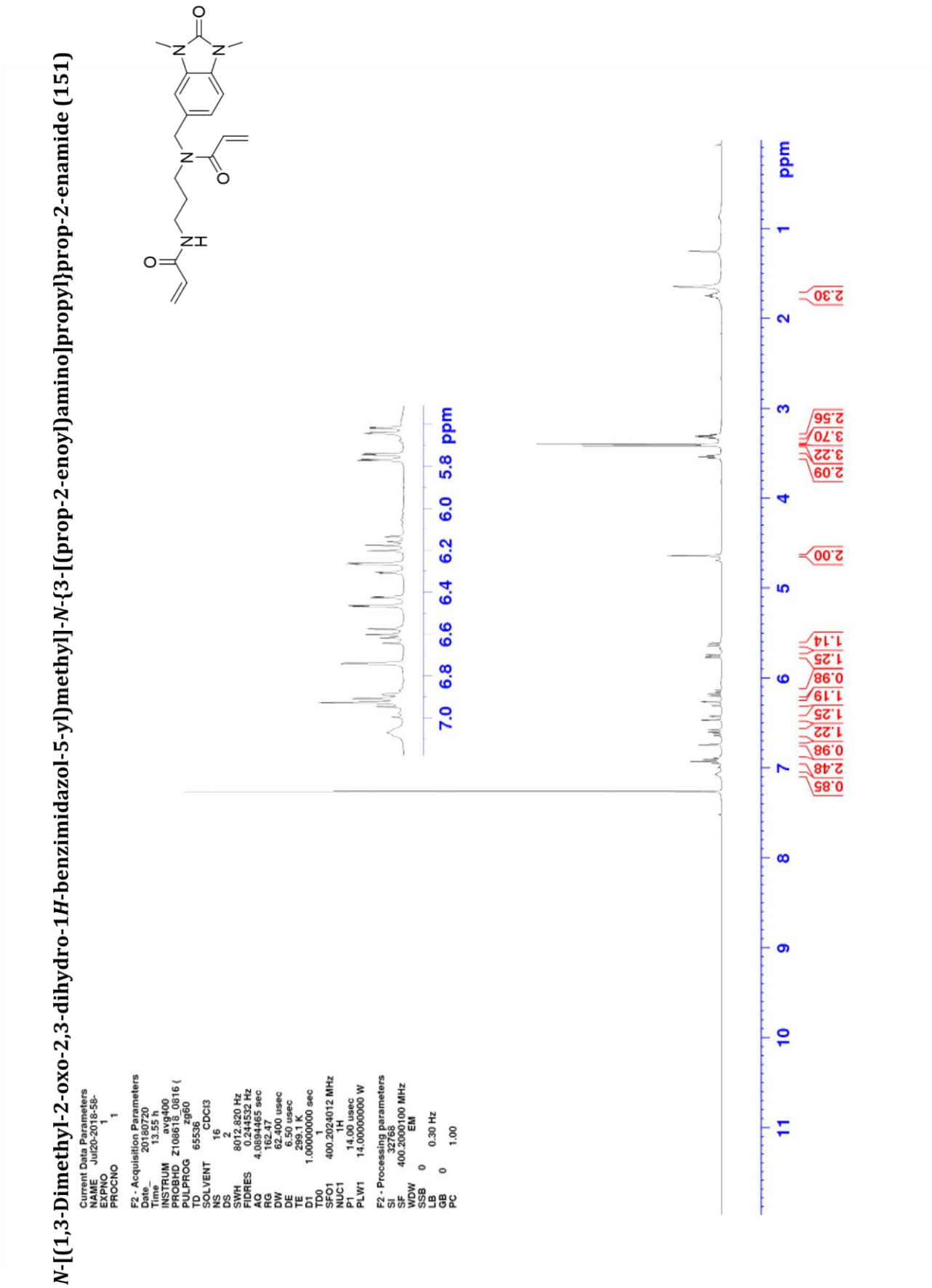



tert-Butyl (3-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1H-benzimidazol-5-yl)methyl]amino}propyl)carbamate (150)
Current Data Parameters
NAME Jun06-2018-18-
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20180606
Time 9.38 h
INSTRUM avh400
PROBHD Z108618_0873 (
PULPROG zg60
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8012.820 Hz
FIDRES 0.244532 Hz
AQ 4.0894465 sec
RG 88.17
DW 62.400 usec
DE 6.50 usec
TE 301.5 K
D1 1.00000000 sec
TD0 1
SFO1 400.1324008 MHz
NUC1 1H
P1 14.00 usec
PLW1 14.36999989 W
F2 - Processing parameters
SI 32768
SF 400.1300101 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.1 7.0 ppm
0.91
0.87
0.92
0.72
2.00
2.93
2.94
1.95
2.00
2.12
9.27
11 10 9 8 7 6 5 4 3 2 1 ppm

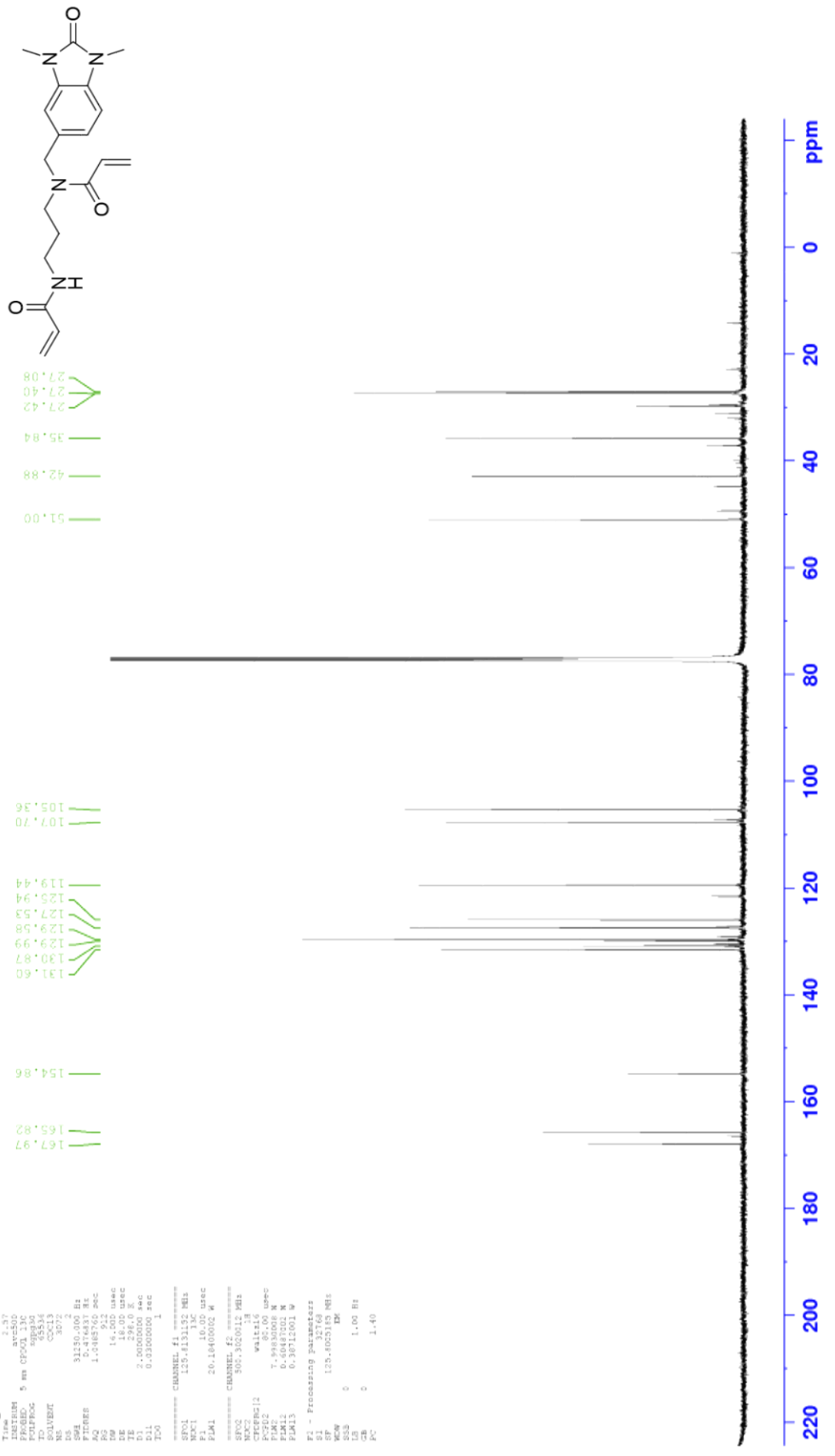
***tert*-Butyl (3-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}propyl)carbamate (150)**



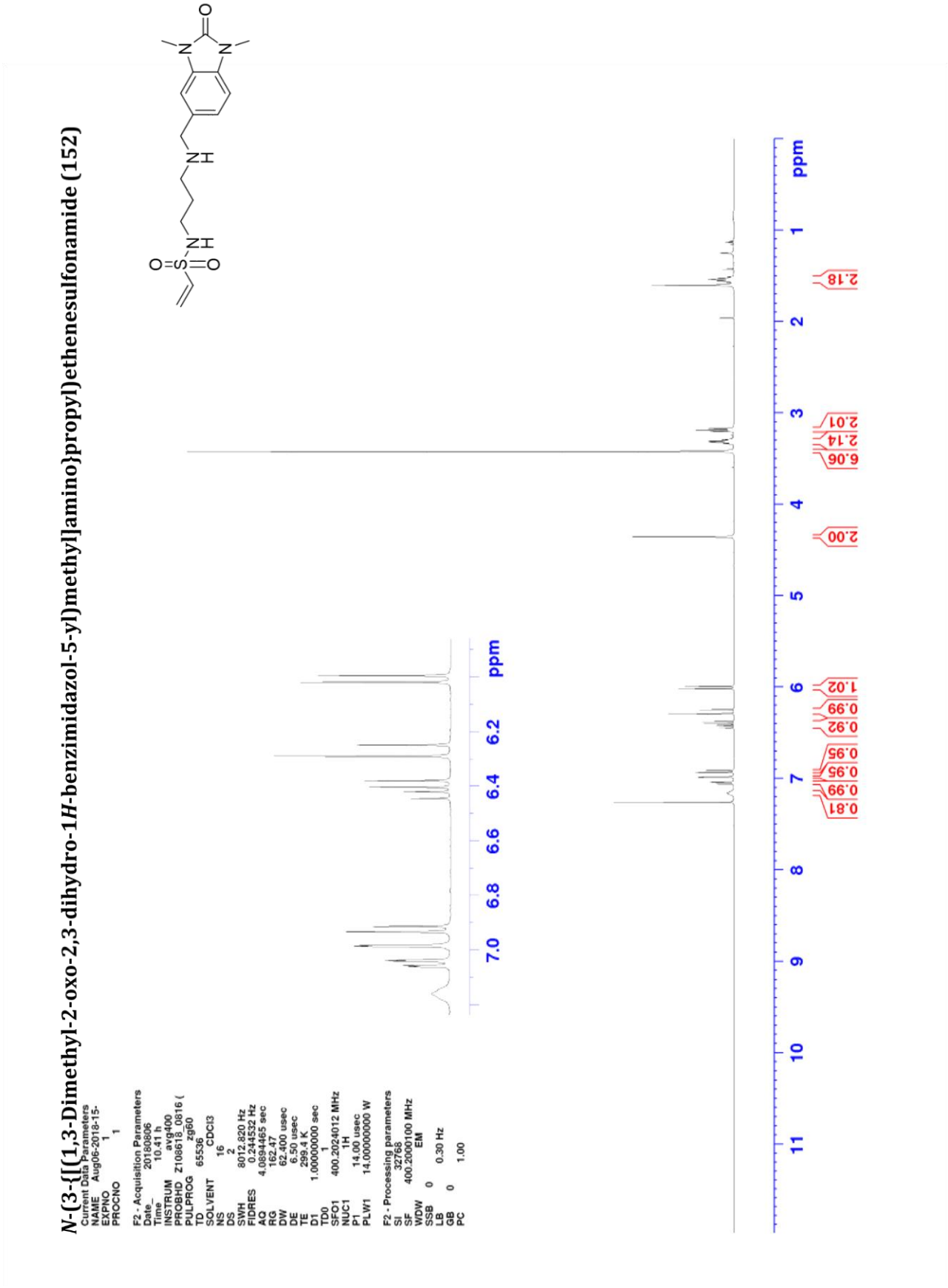
_N_-[(1,3-Dimethyl-2-oxo-2,3-dihydro-1_H_-benzimidazol-5-yl)methyl]-_N_-{3-[(prop-2-enoyl)amino]propyl}prop-2-enamide (151)



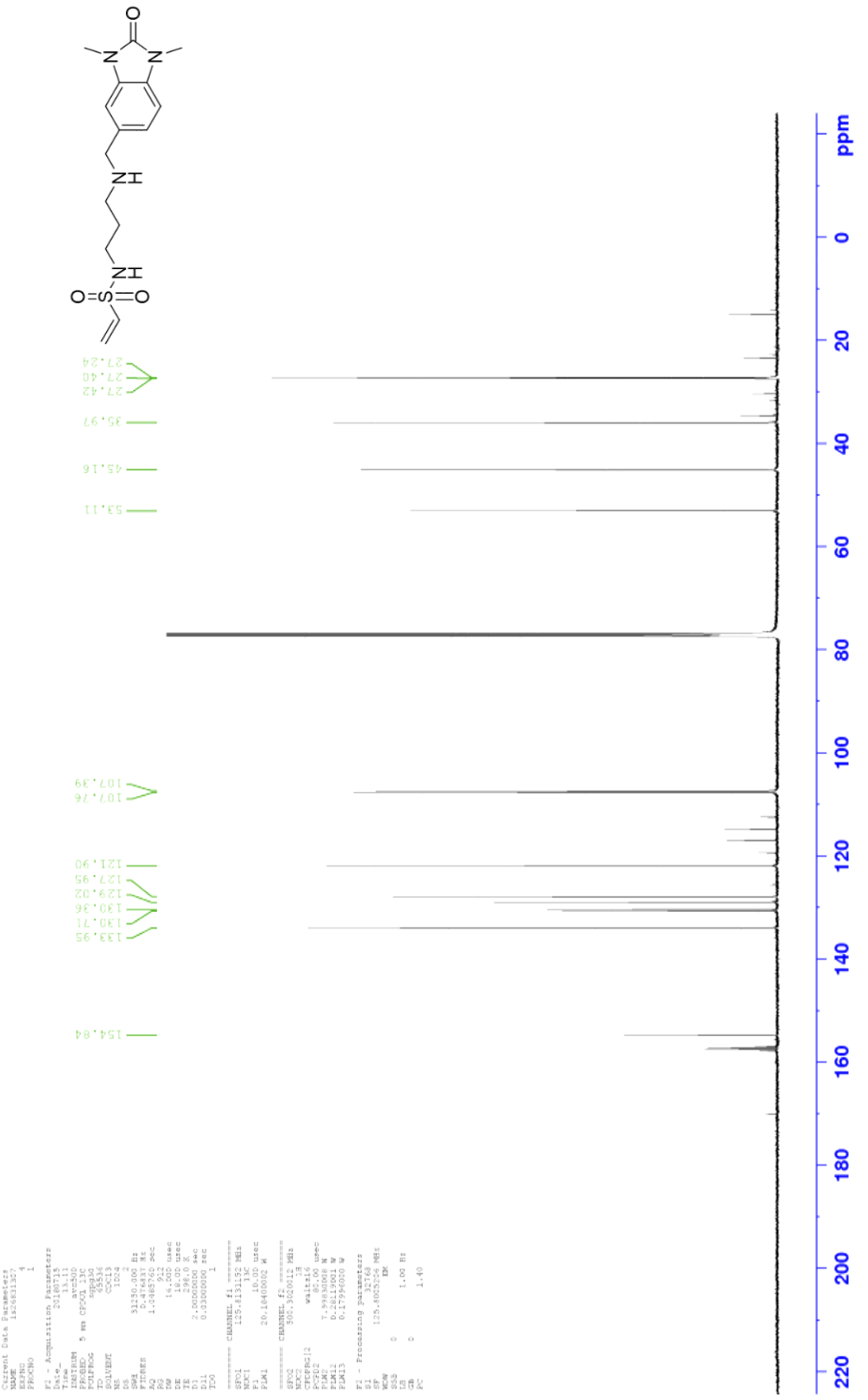
_N_-[(1,3-Dimethyl-2-oxo-2,3-dihydro-1_H_-benzimidazol-5-yl)methyl]-_N_-{3-[(prop-2-enoyl)amino]propyl}prop-2-enamide (151)

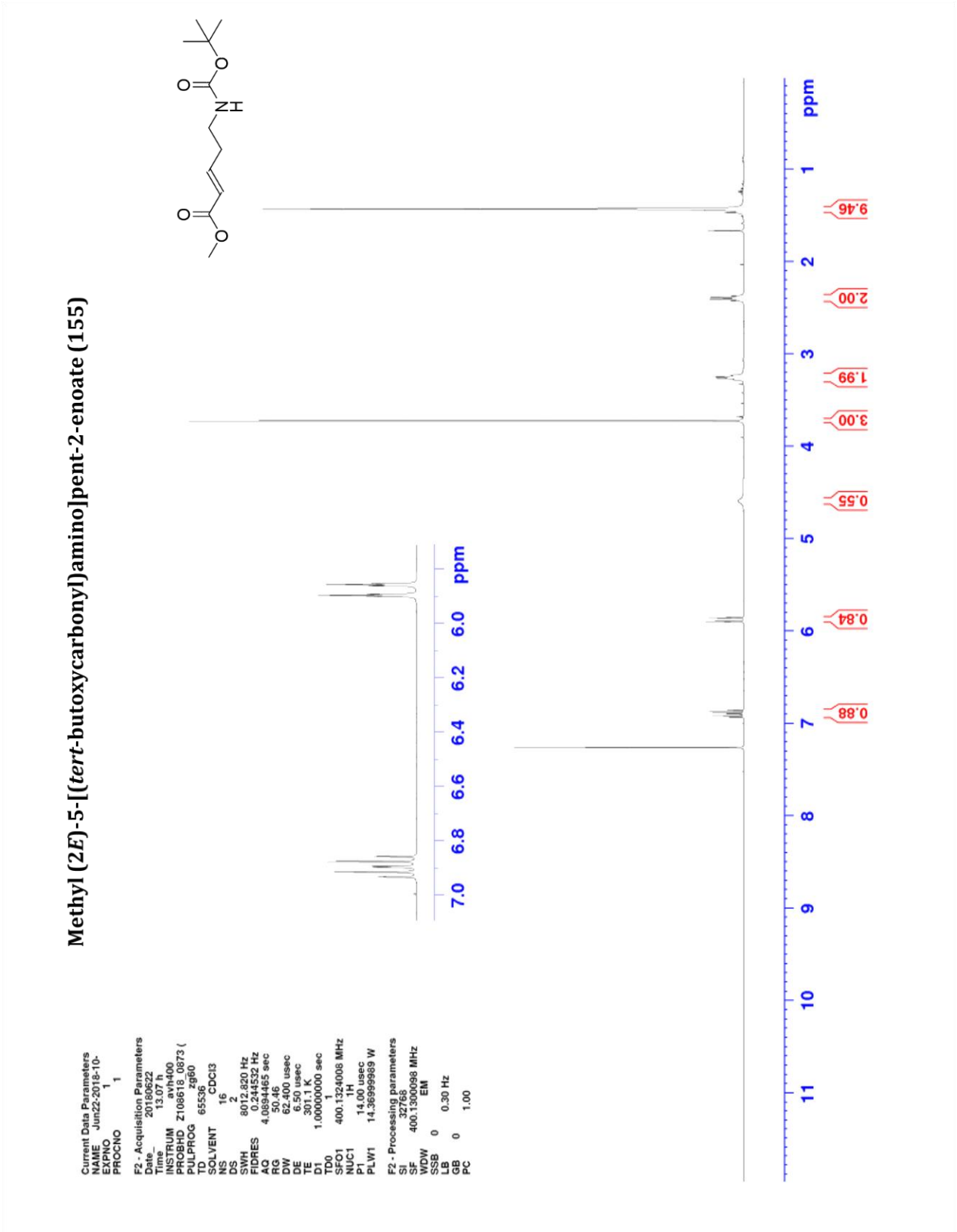


N-(3-{[(1,3-Dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}propyl)ethenesulfonamide (152)

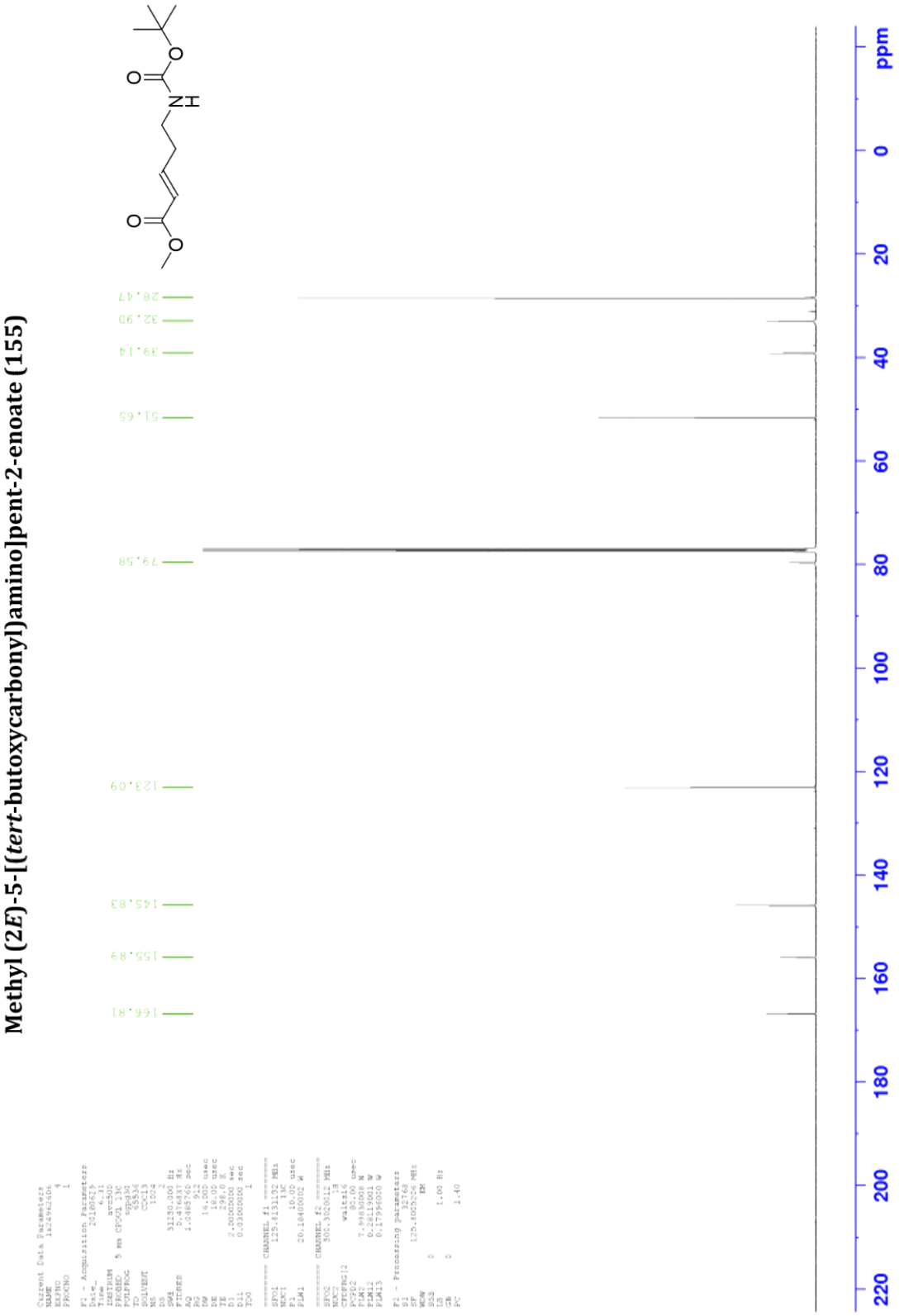


***N*-(3-{[(1,3-Dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}propyl)ethenesulfonamide (152)**

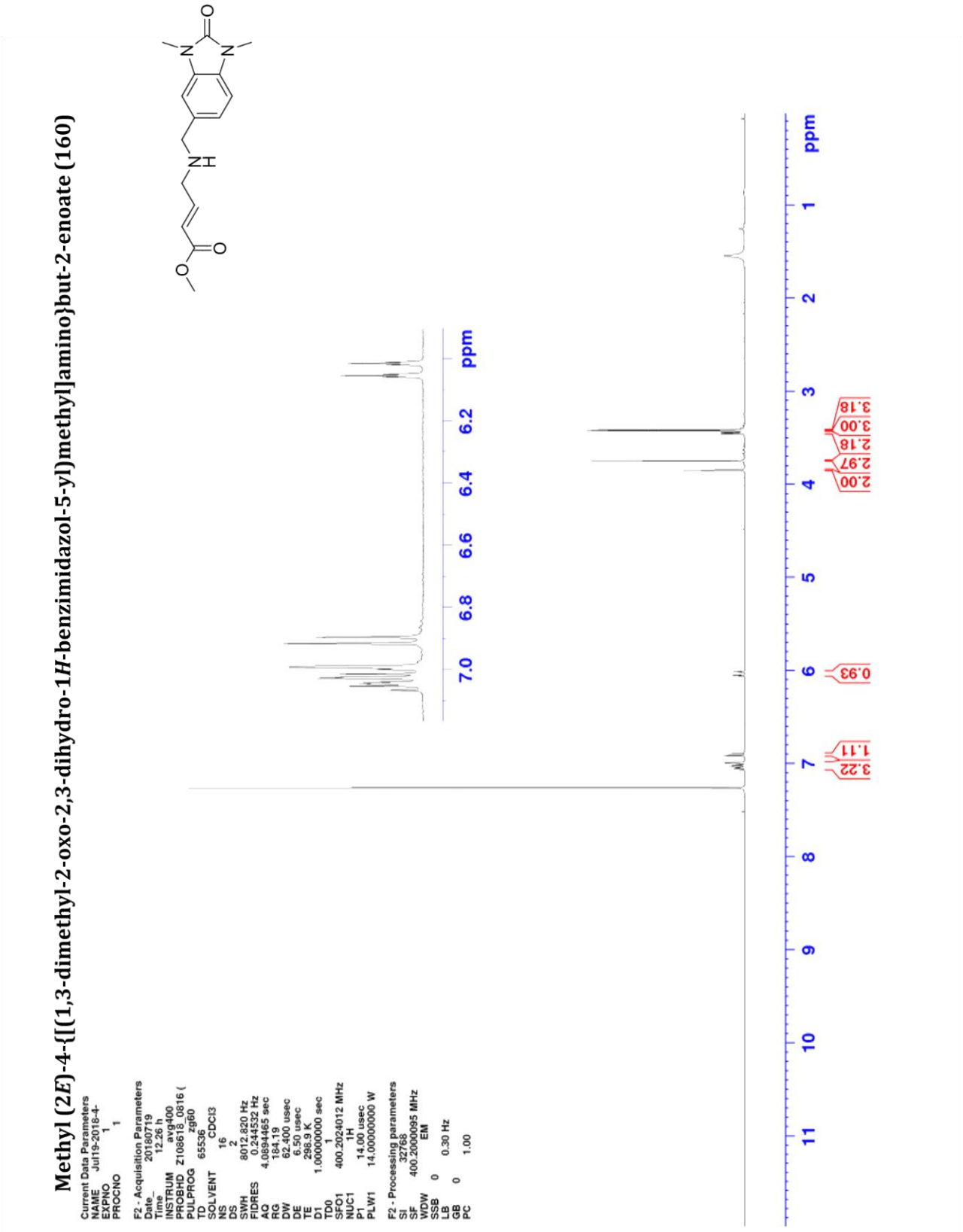



Methyl (2E)-5-[(tert-butoxycarbonyl)amino]pent-2-enoate (155)
Current Data Parameters
NAME Jun22-2018-10-
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20180622
Time 13.07 h
INSTRUM avh400
PROBHD Z108618_0873 (
PULPROG zg60
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8012.820 Hz
FIDRES 0.244532 Hz
AQ 4.0894465 sec
RG 50.46
DW 62.400 usec
DE 6.50 usec
TE 301.1 K
D1 1.00000000 sec
TD0 1
SFO1 400.1324006 MHz
NUC1 1H
P1 14.00 usec
PLW1 14.36999989 W
F2 - Processing parameters
SI 32768
SF 400.1300098 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
O
O
O
O
NH
7.0
6.8
6.6
6.4
6.2
6.0
ppm
11
10
9
8
7
6
5
4
3
2
1
ppm
9.46
2.00
1.99
3.00
0.55
0.84
0.88

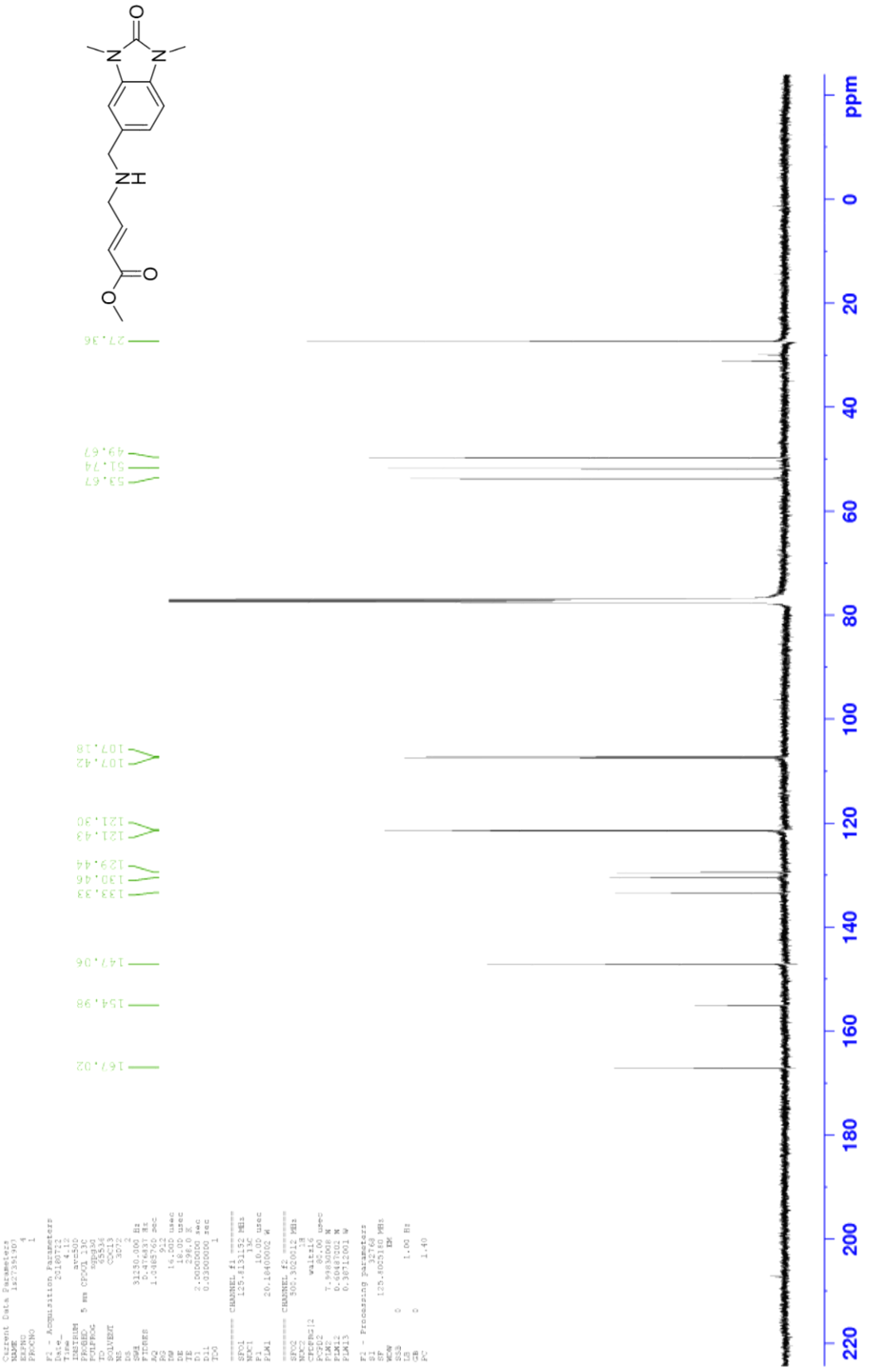
Methyl (2*E*)-5-[(*tert*-butoxycarbonyl)amino]pent-2-enoate (155)



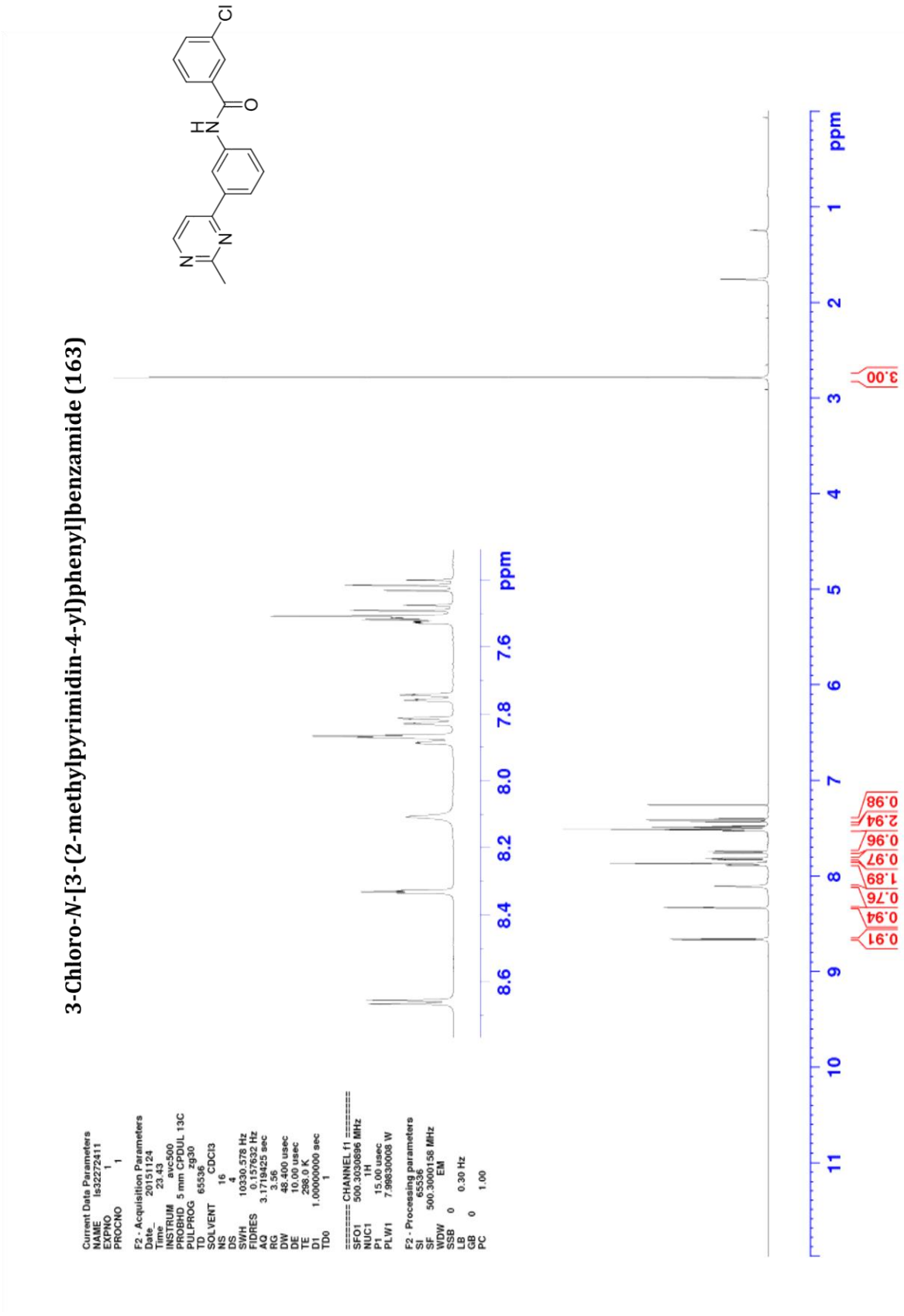
Methyl (2*E*)-4-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}but-2-enoate (160)



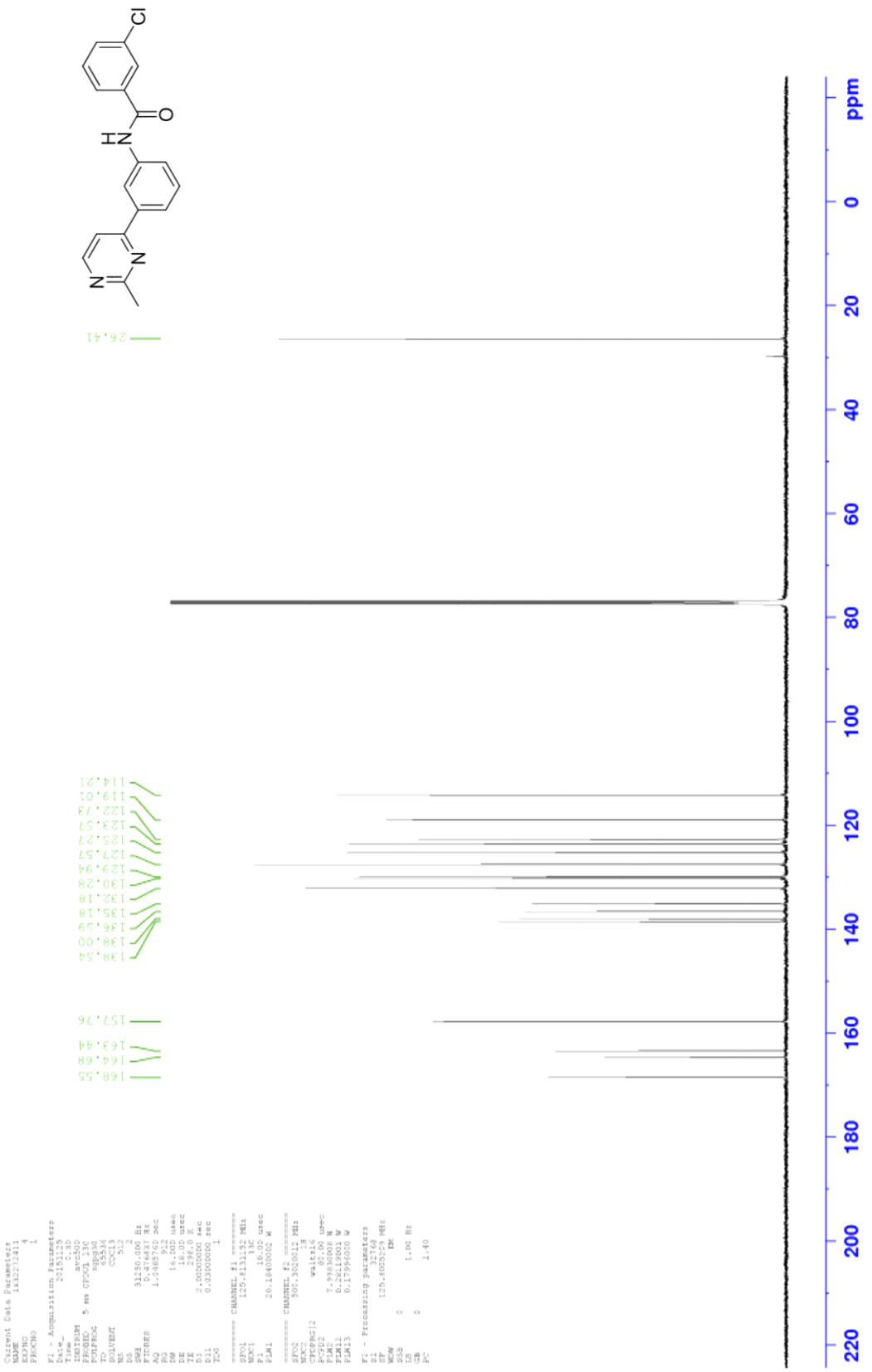
Methyl (2*E*)-4-{[(1,3-dimethyl-2-oxo-2,3-dihydro-1*H*-benzimidazol-5-yl)methyl]amino}but-2-enoate (160)



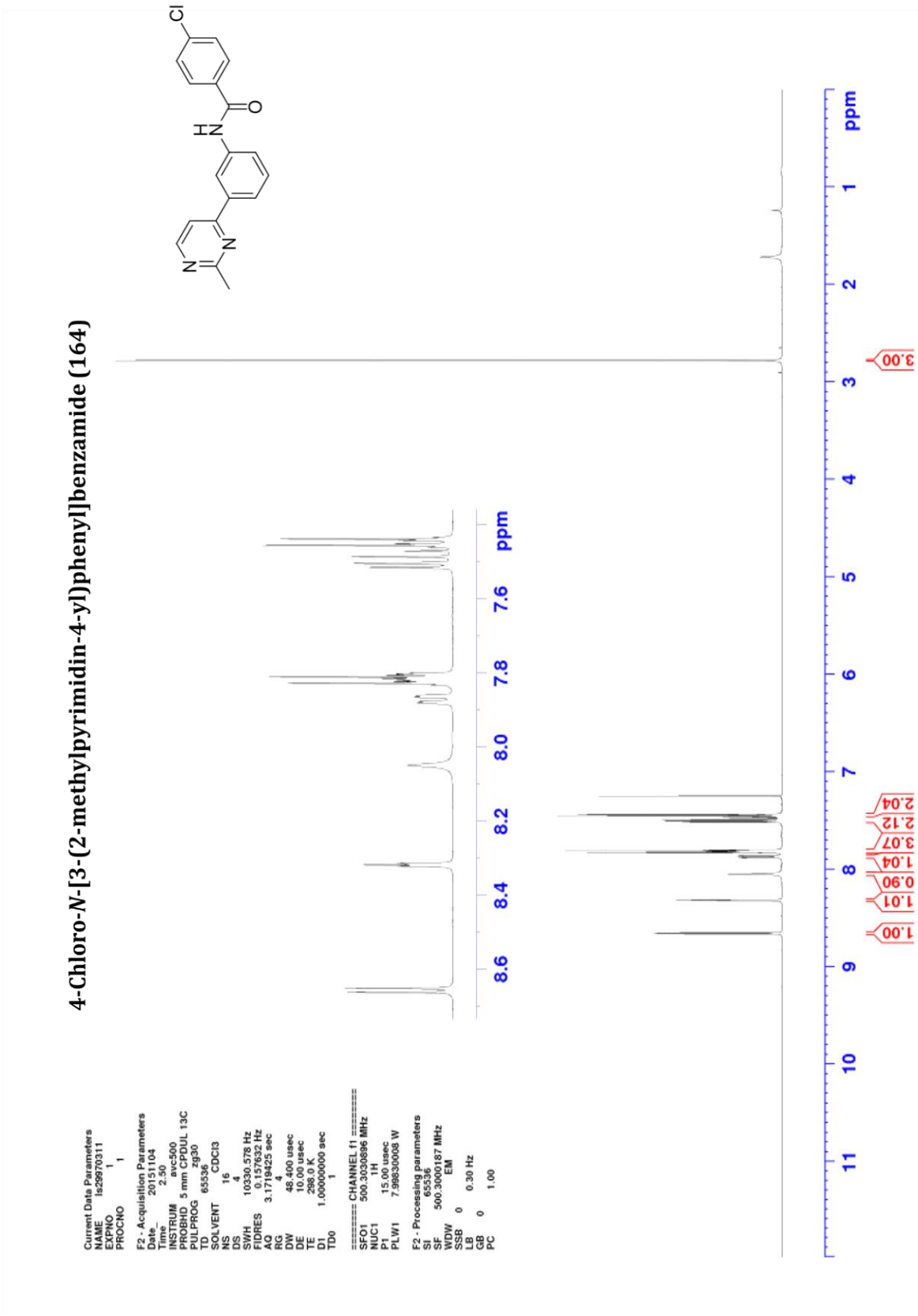
3-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (163)



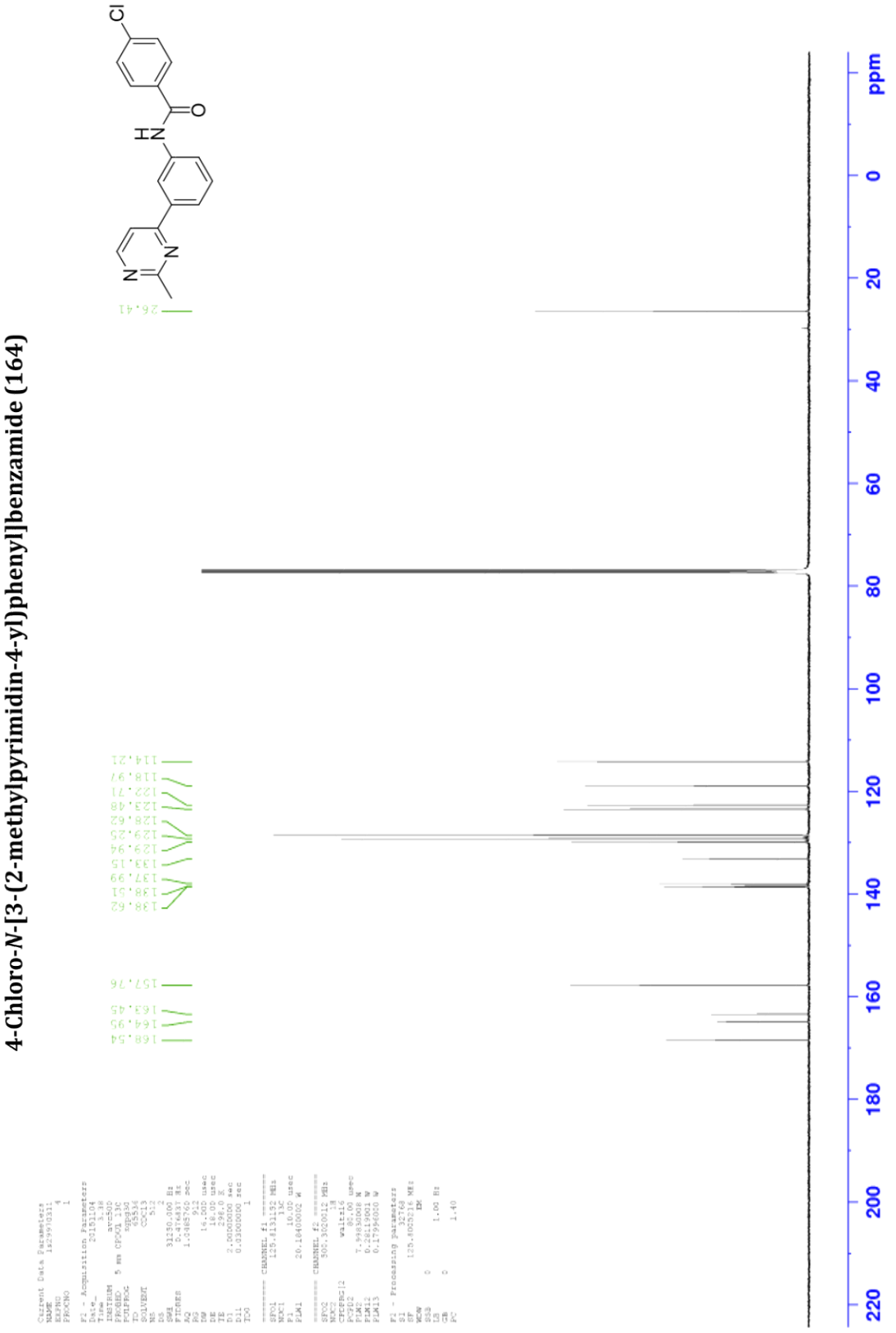
3-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (163)

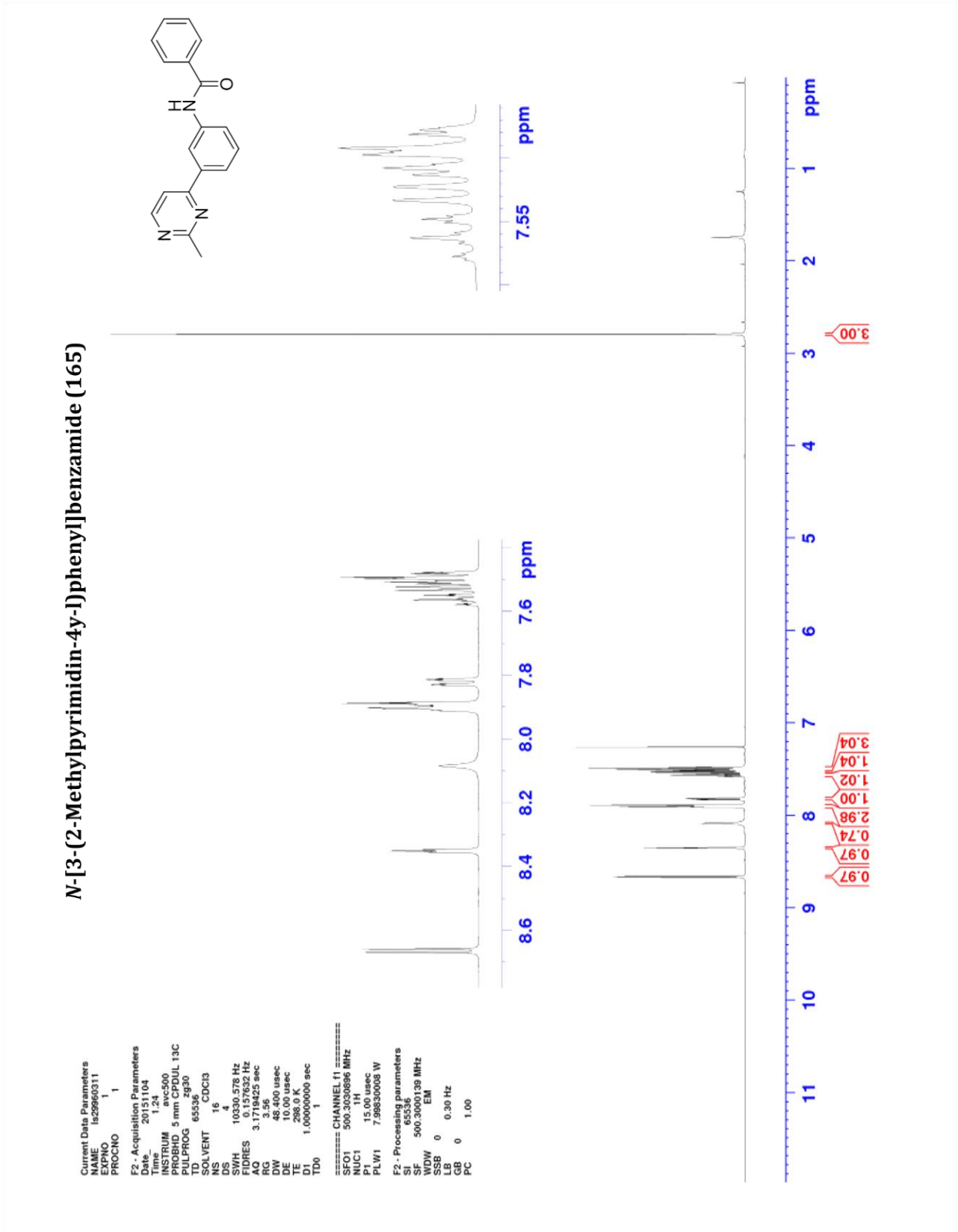


4-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (164)

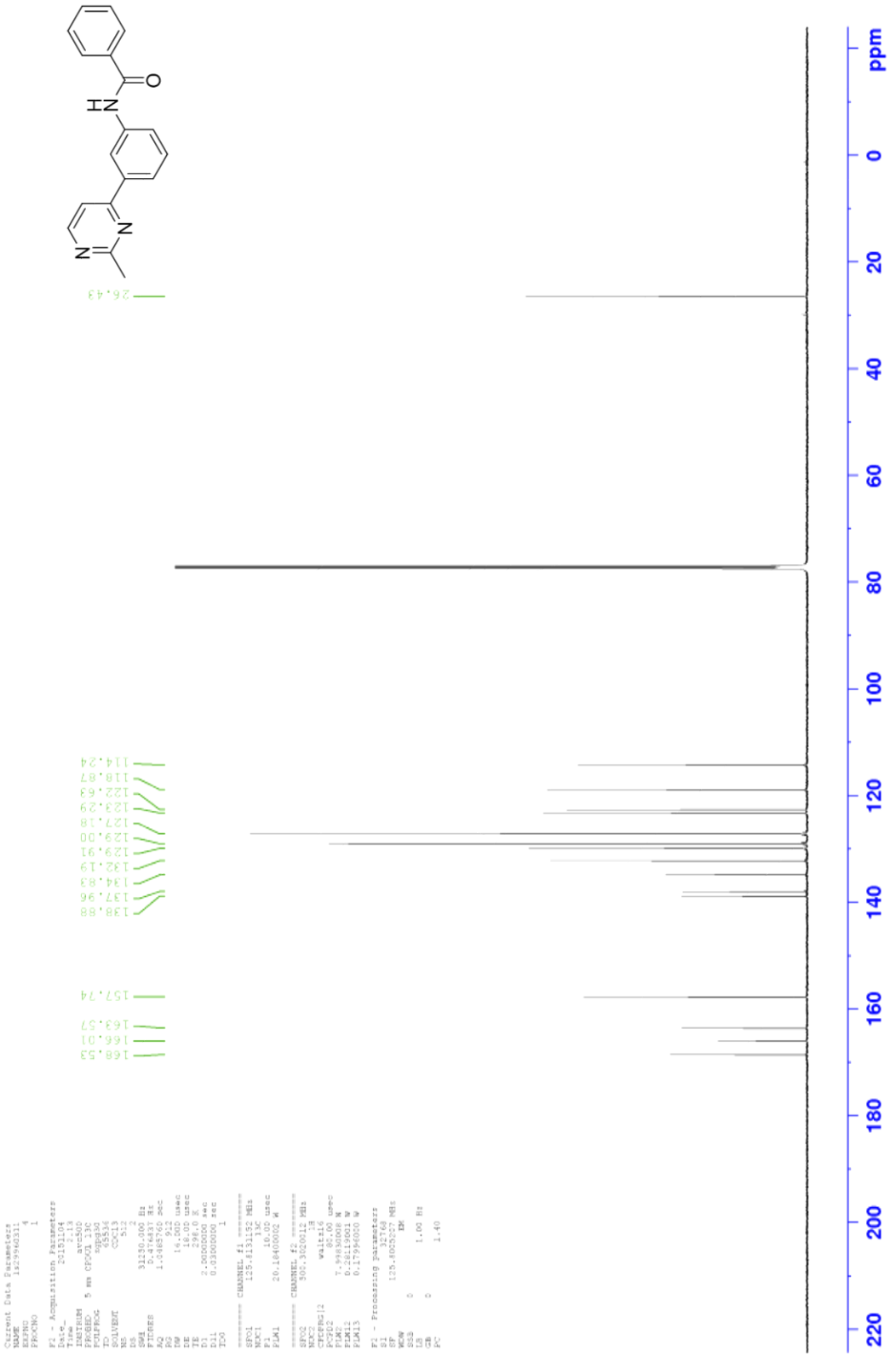


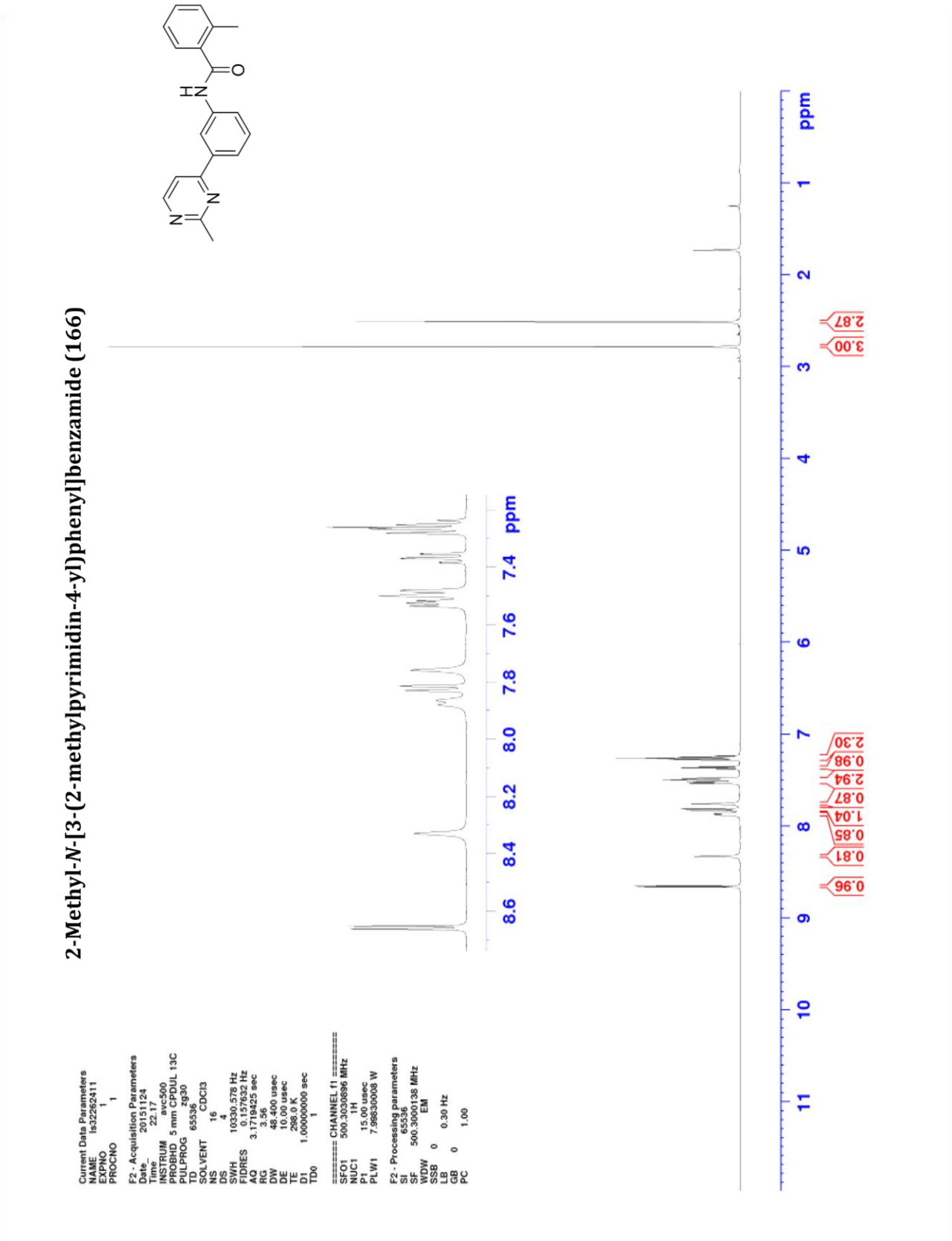
4-Chloro-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (164)



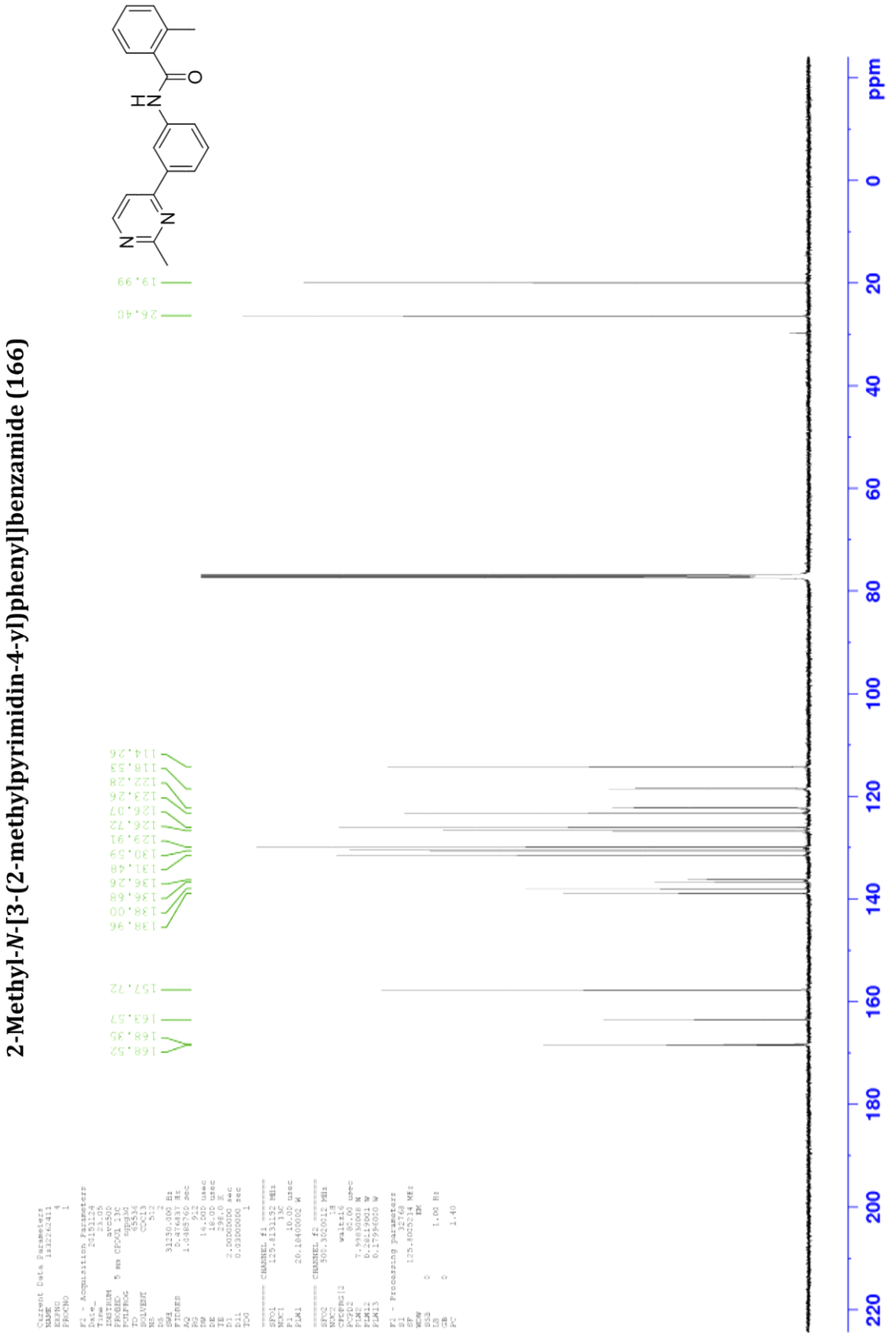

N-[3-(2-Methylpyrimidin-4y-l)phenyl]benzamide (165)
Current Data Parameters
NAME ls29960311
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20151104
Time 1.24
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.1719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000139 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
8.6
8.4
8.2
8.0
7.8
7.6 ppm
7.55 ppm
11
10
9
8
7
6
5
4
3
2
1 ppm
3.00
0.97
0.97
0.74
2.98
1.00
1.02
1.04
3.04

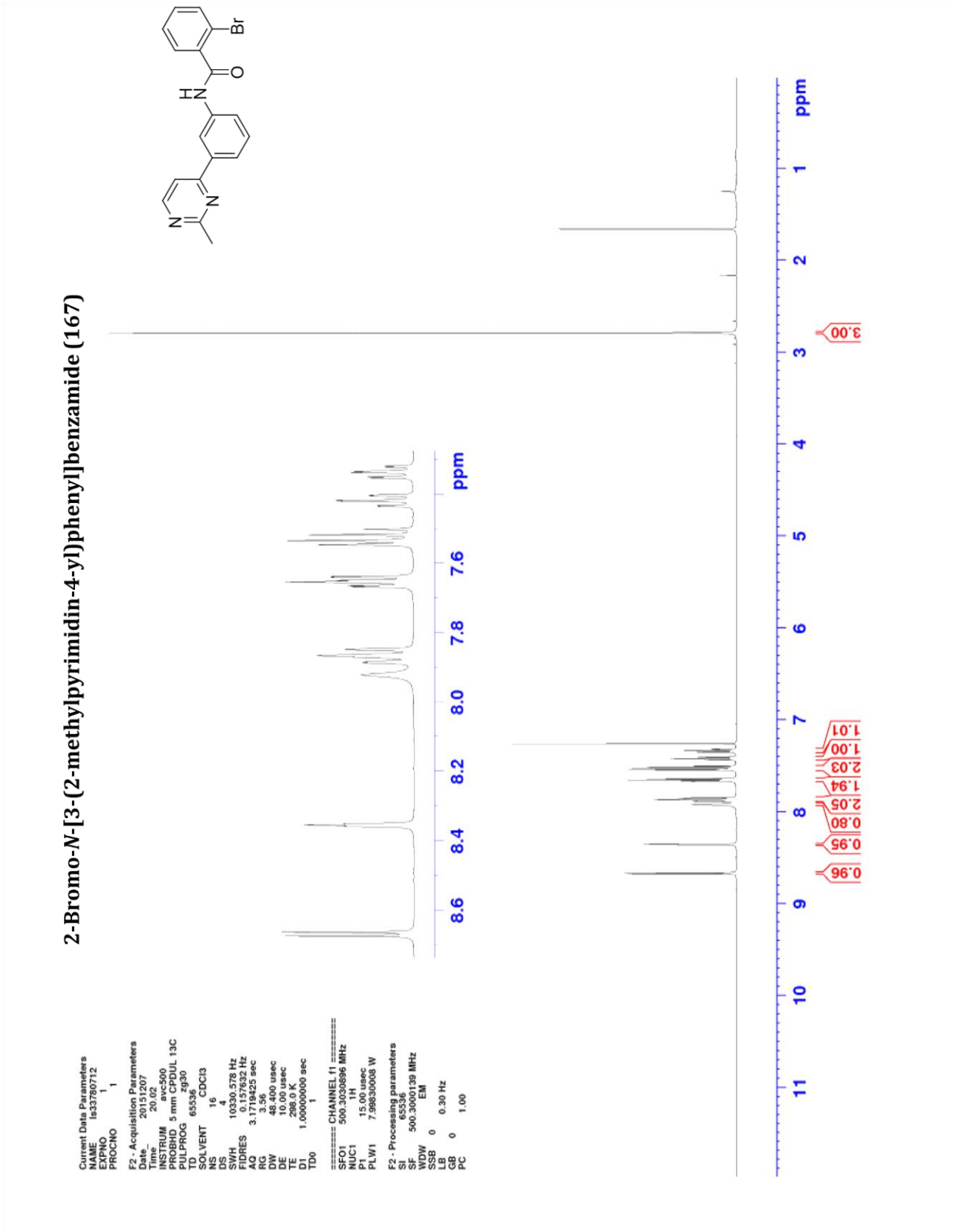
N-[3-(2-Methylpyrimidin-4-yl)phenyl]benzamide (165)



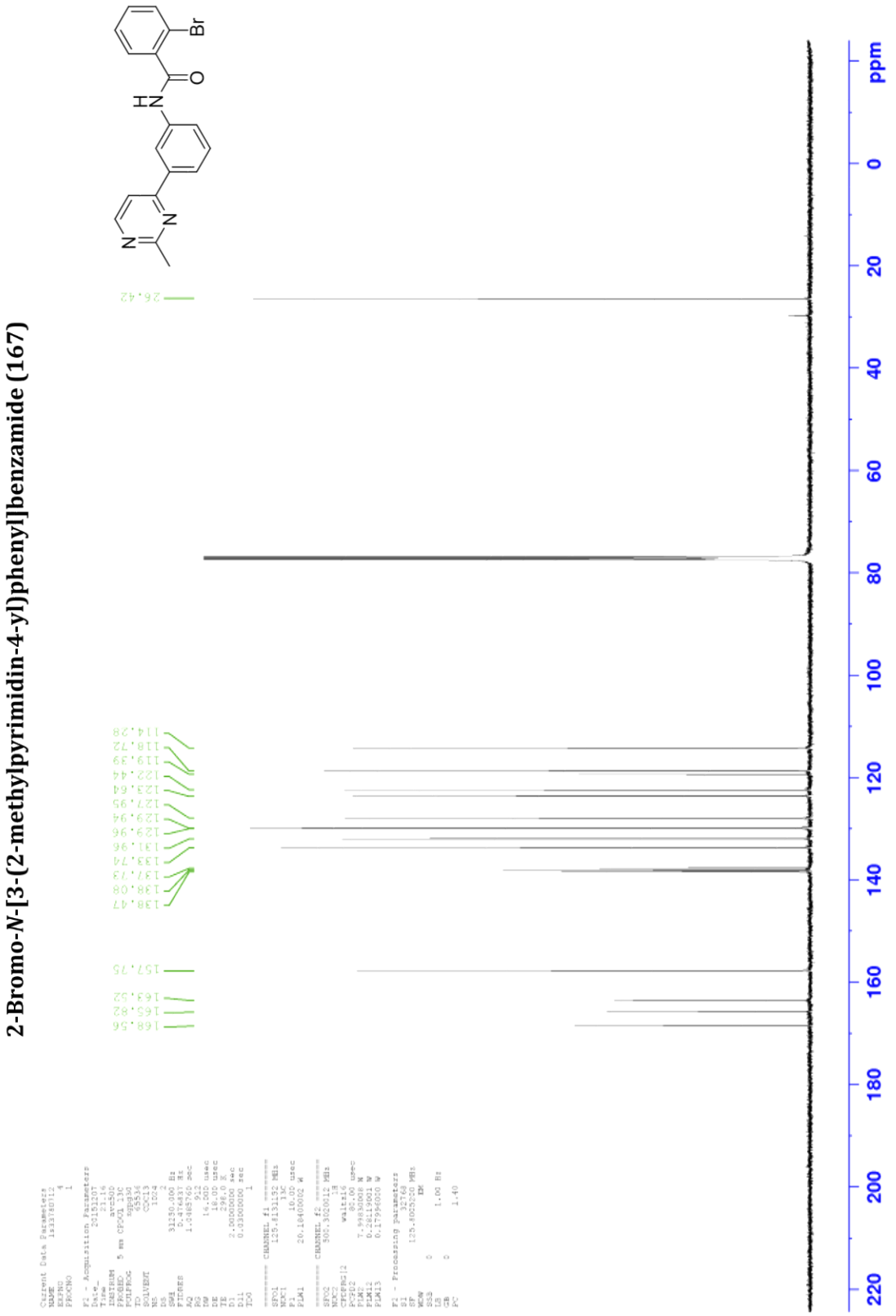

2-Methyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (166)
Current Data Parameters
NAME ls32262411
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20151124
Time 22.17
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000138 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
8.6
8.4
8.2
8.0
7.8
7.6
7.4 ppm
11
10
9
8
7
6
5
4
3
2
1 ppm
0.96
0.81
0.85
1.04
0.87
2.94
0.98
2.30
3.00
2.87

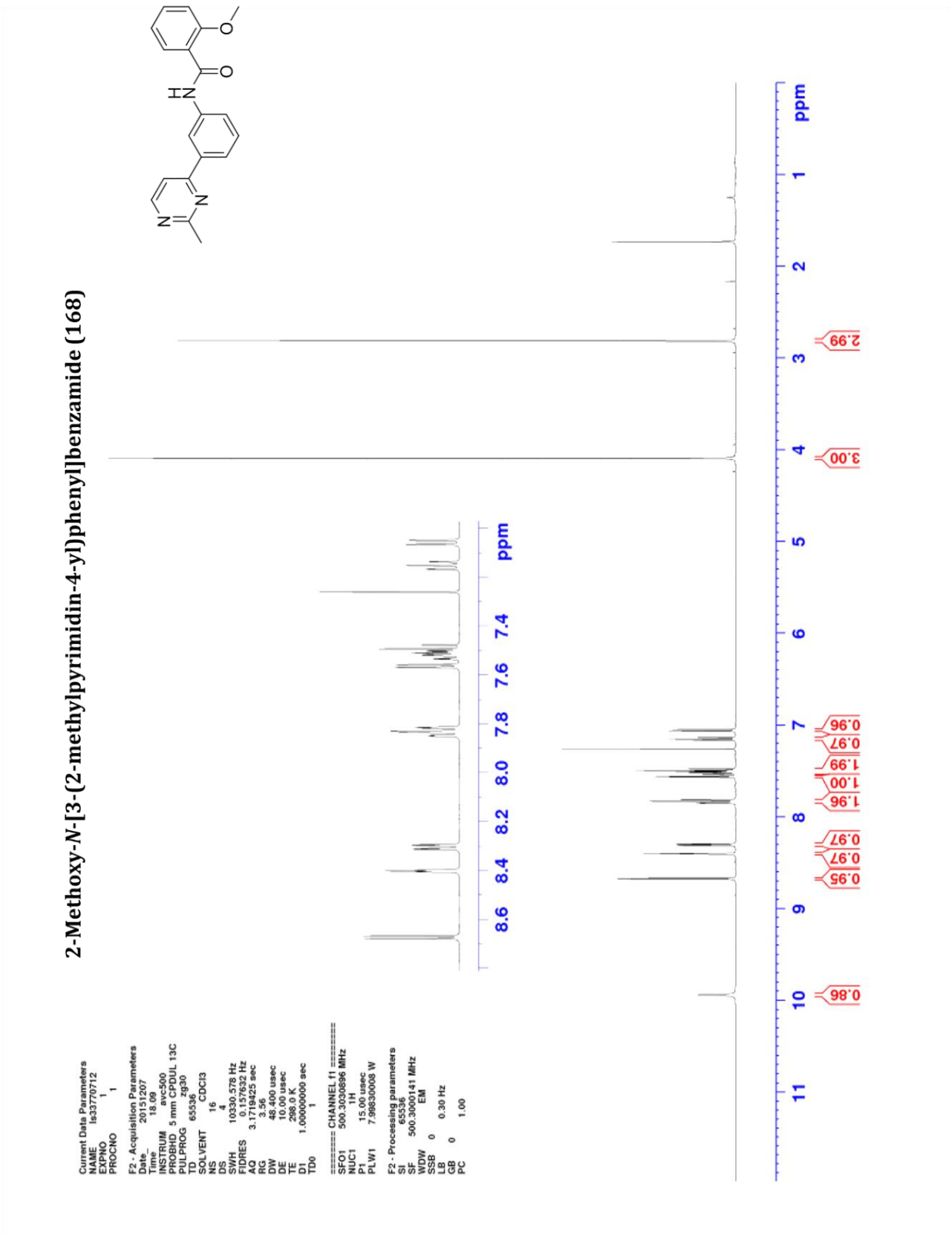
2-Methyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (166)



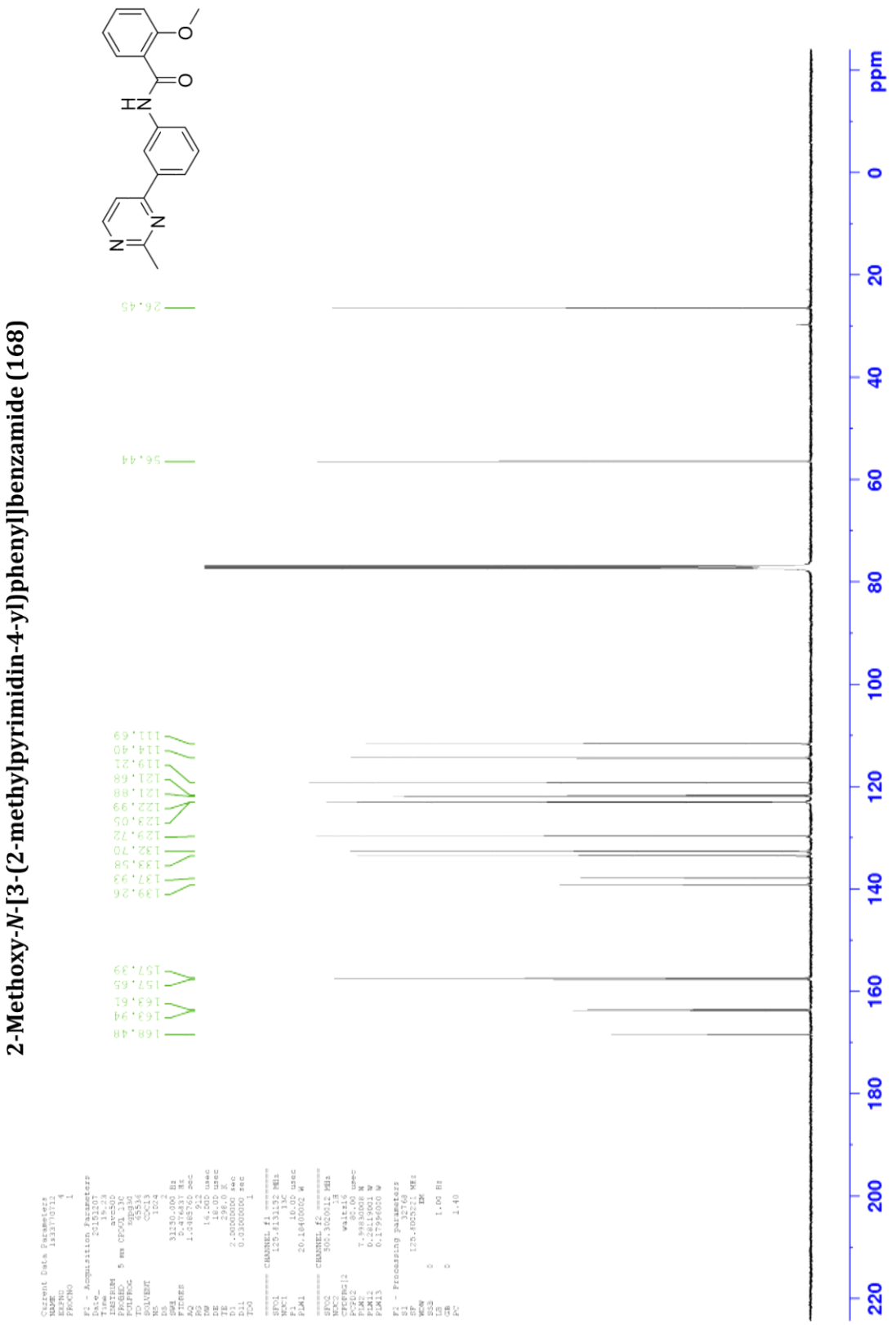

2-Bromo-N-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (167)
Current Data Parameters
NAME ls33780712
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20151207
Time 20.02
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000139 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
Br
O
H
N
N
N
8.6
8.4
8.2
8.0
7.8
7.6
ppm
11
10
9
8
7
6
5
4
3
2
1
ppm
0.96
0.95
0.80
2.05
1.94
2.03
1.00
1.01
3.00

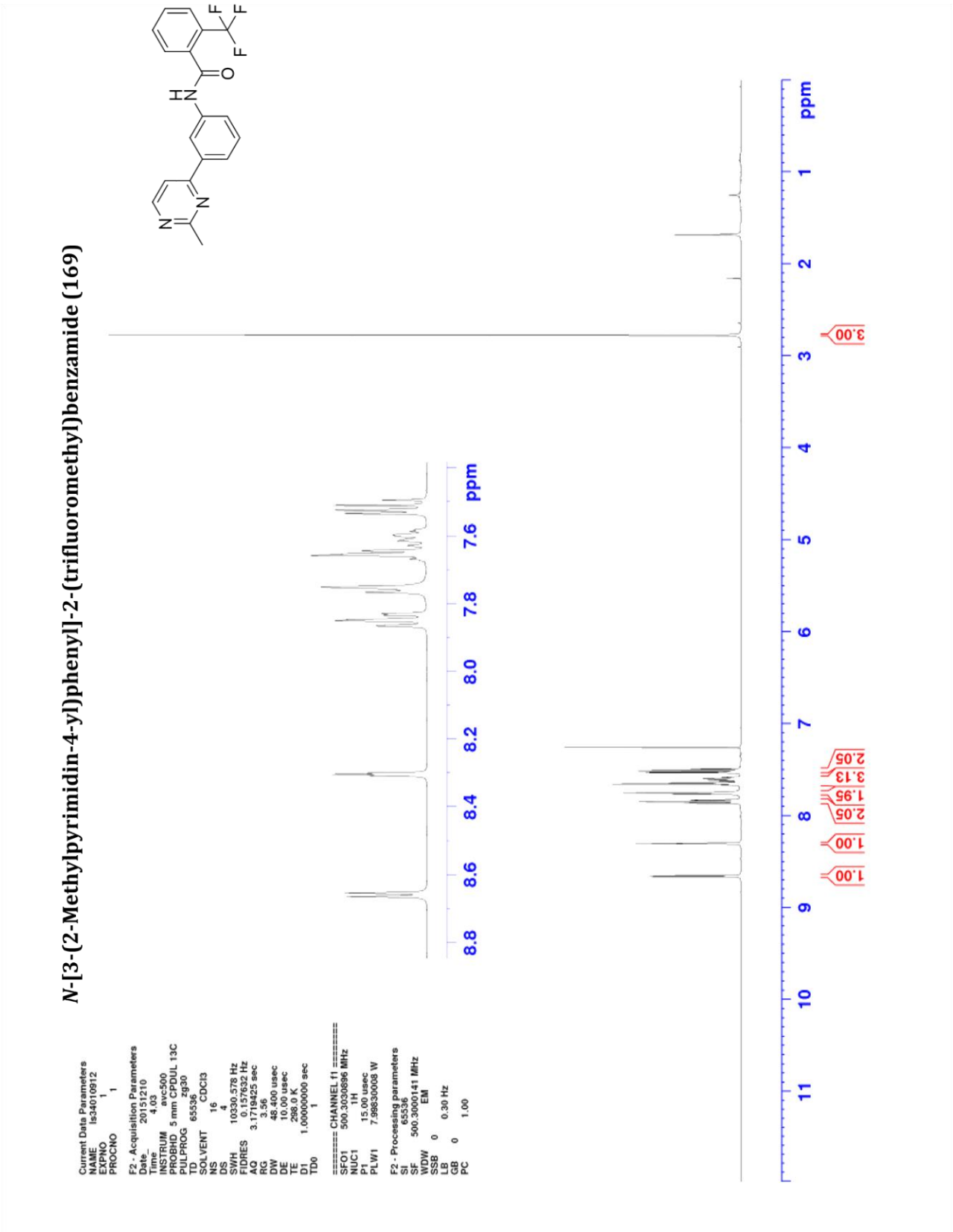
2-Bromo-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (167)



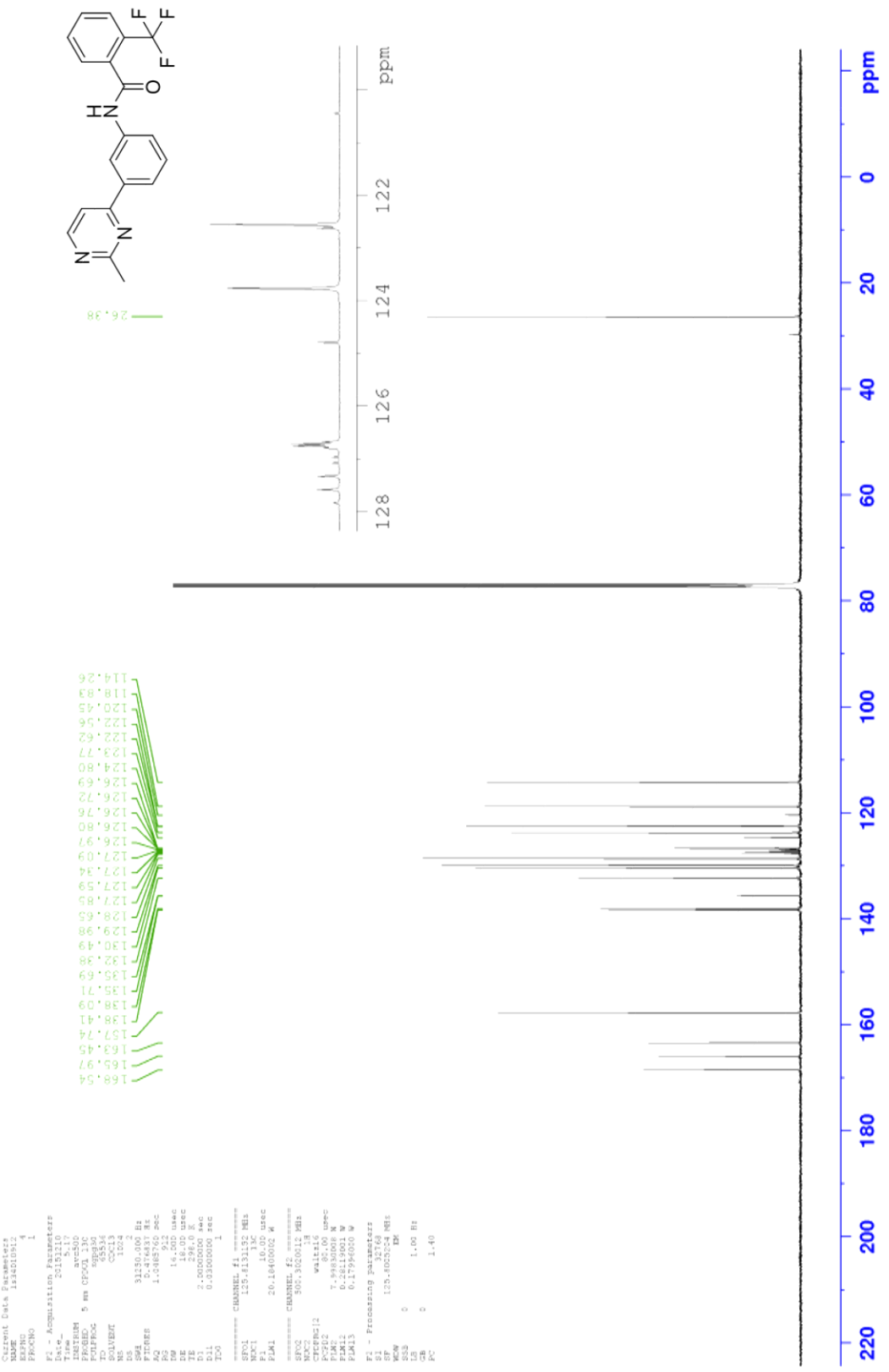

2-Methoxy-N-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (168)
Current Data Parameters
NAME ls33770712
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20151207
Time 18.09
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000141 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
ppm
2.99
3.00
0.96
0.97
1.99
1.00
1.96
0.97
0.97
0.95
0.86

2-Methoxy-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]benzamide (168)




N-[3-(2-Methylpyrimidin-4-yl)phenyl]-2-(trifluoromethyl)benzamide (169)
Current Data Parameters
NAME ls34010912
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20151210
Time 4.03
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000141 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
8.8
8.6
8.4
8.2
8.0
7.8
7.6 ppm
11
10
9
8
7
6
5
4
3
2
1
ppm
3.00
1.00
1.00
2.05
1.95
3.13
2.05

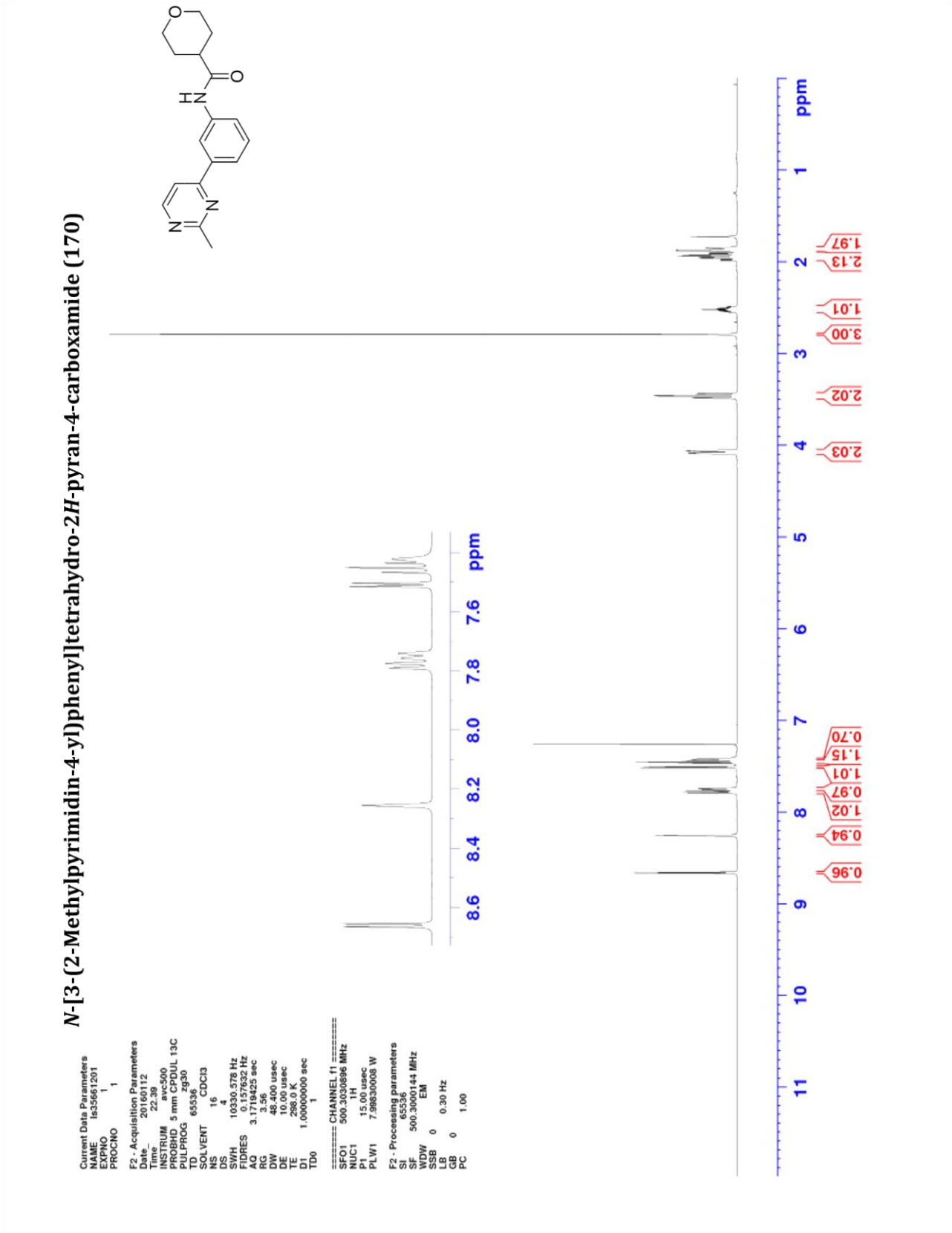
N-[3-(2-Methylpyrimidin-4-yl)phenyl]-2-(trifluoromethyl)benzamide (169)



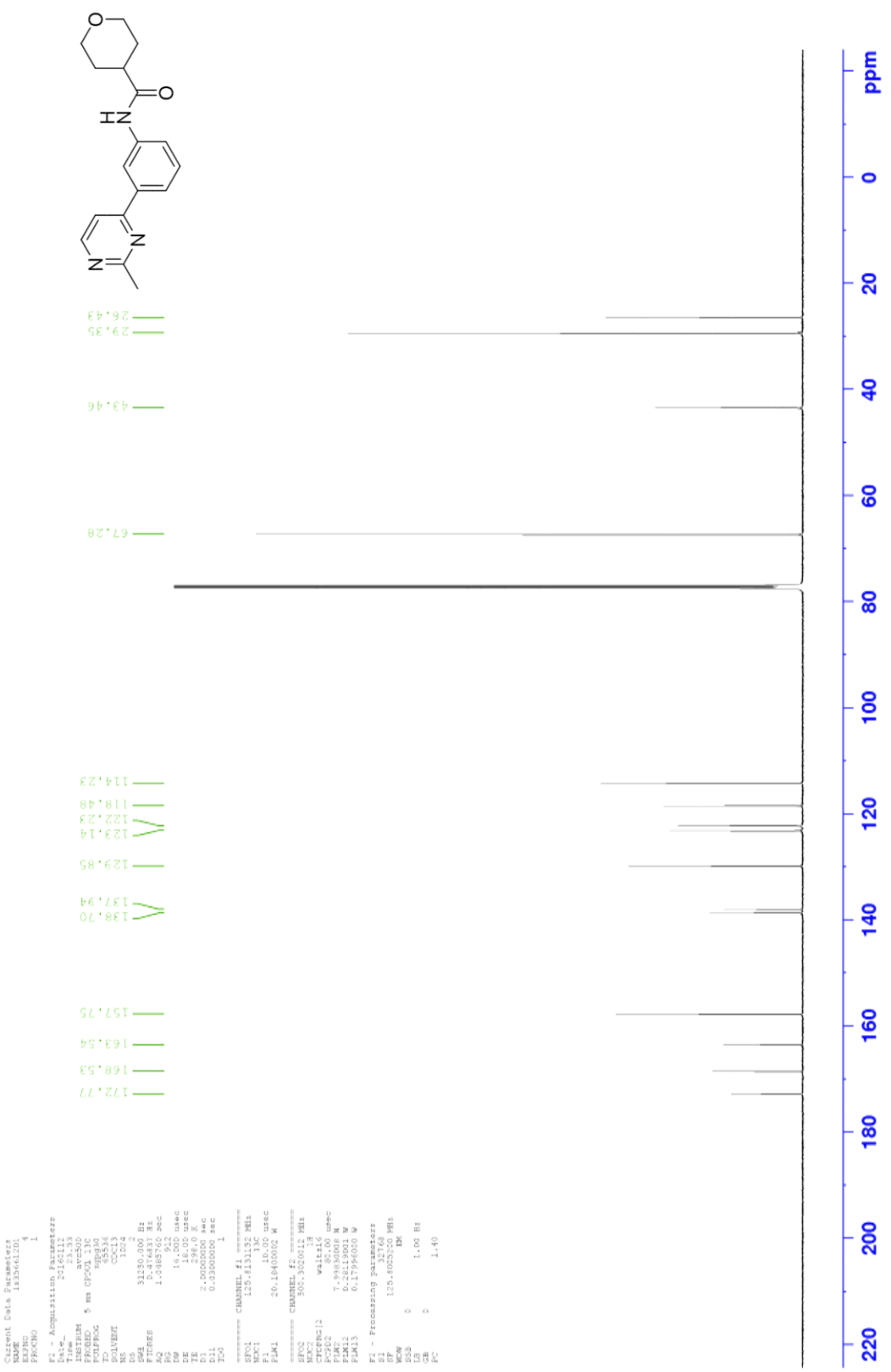
N-[3-(2-Methylpyrimidin-4-yl)phenyl]-2-(trifluoromethyl)benzamide (169)



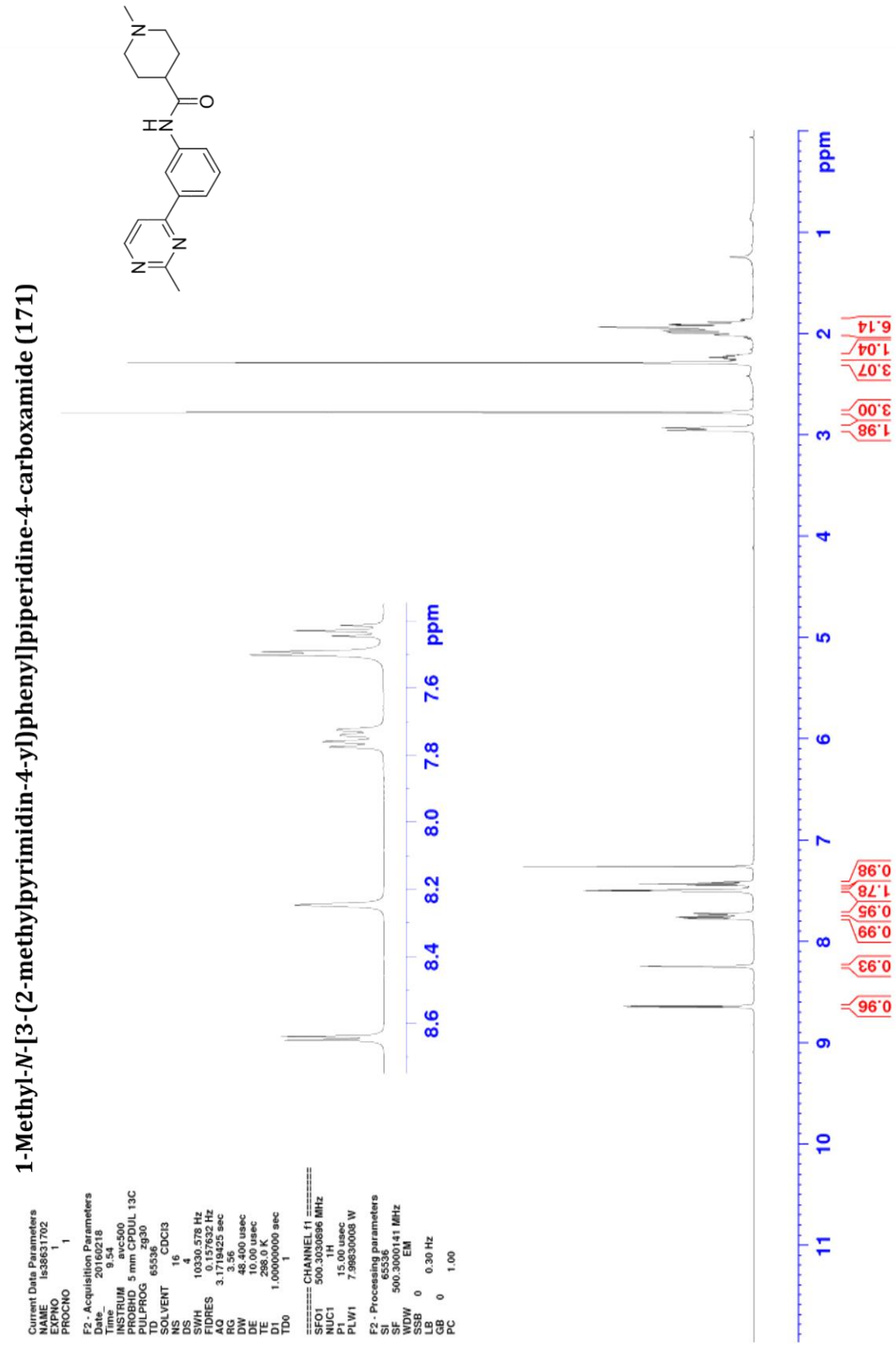
N-[3-(2-Methylpyrimidin-4-yl)phenyl]tetrahydro-2*H*-pyran-4-carboxamide (170)



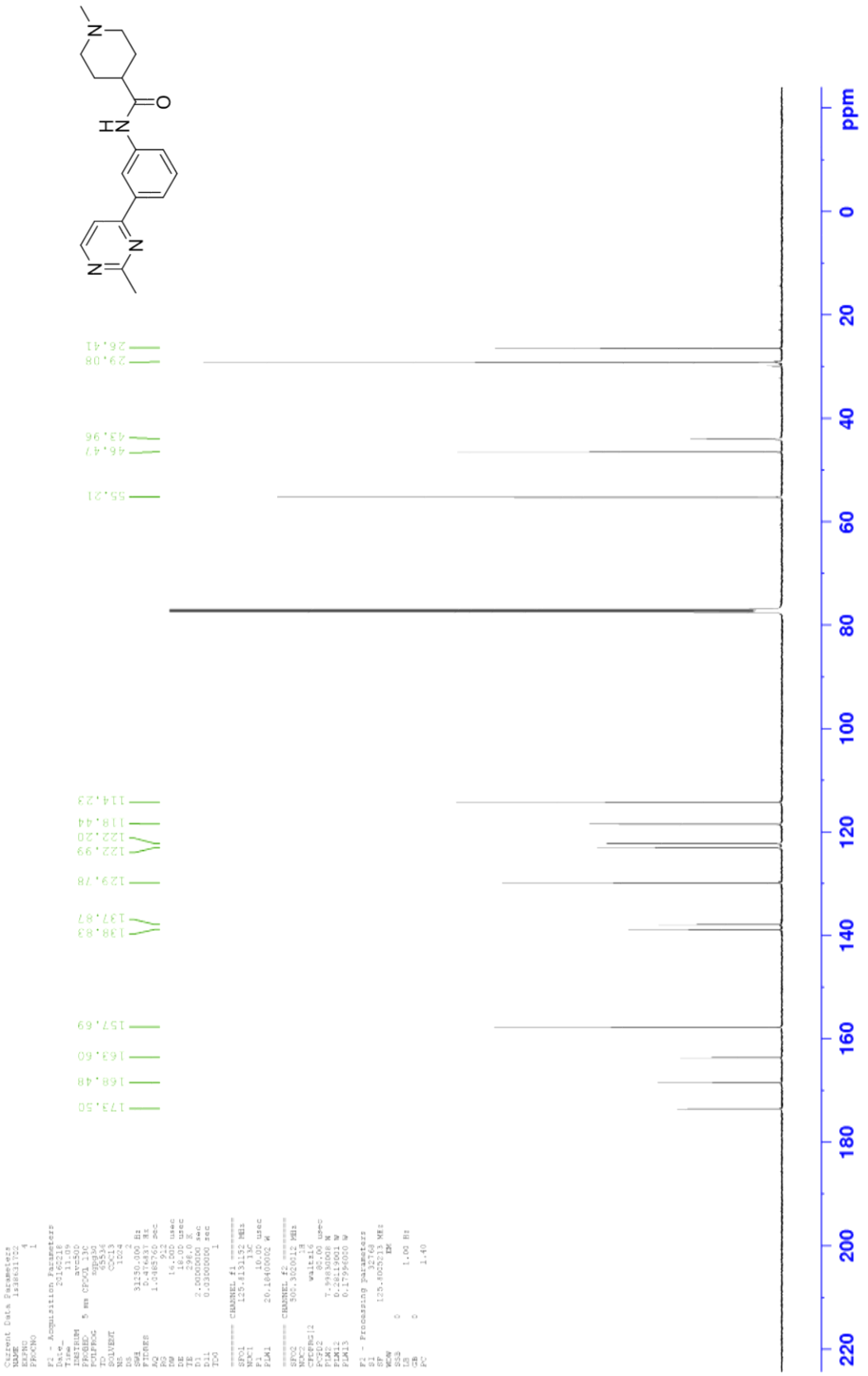
***N*-[3-(2-Methylpyrimidin-4-yl)phenyl]tetrahydro-2*H*-pyran-4-carboxamide (170)**

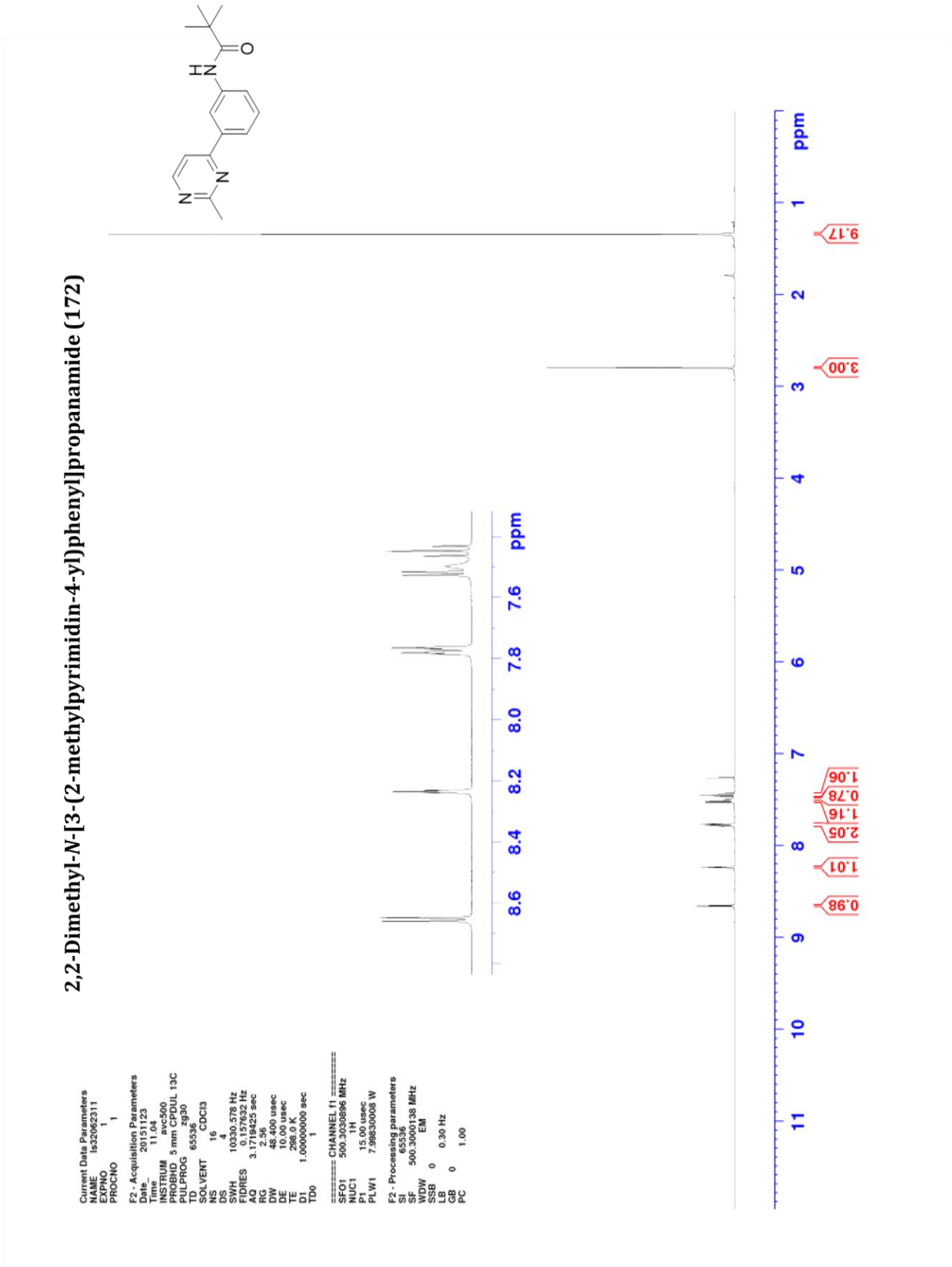


1-Methyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]piperidine-4-carboxamide (171)

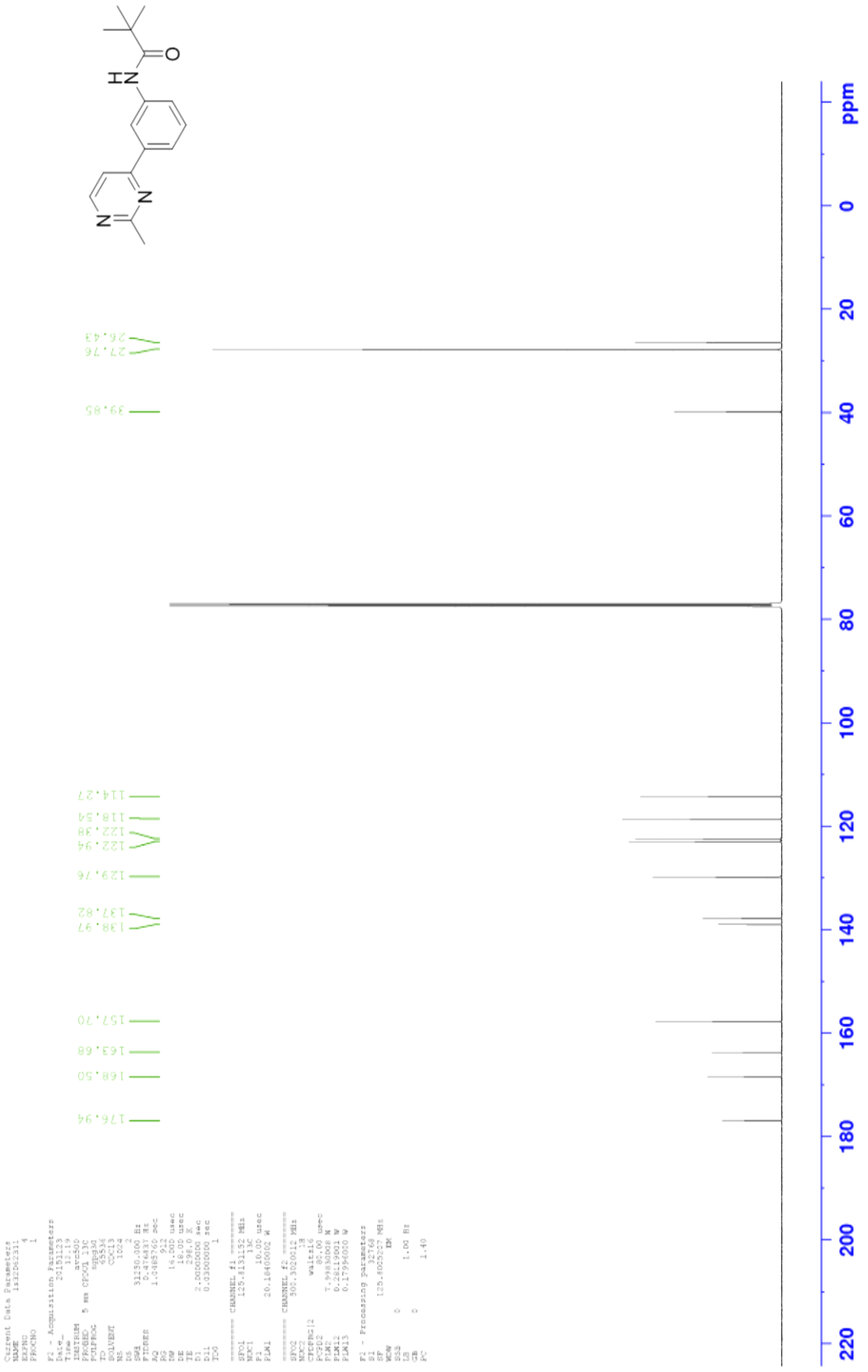


1-Methyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]piperidine-4-carboxamide (171)

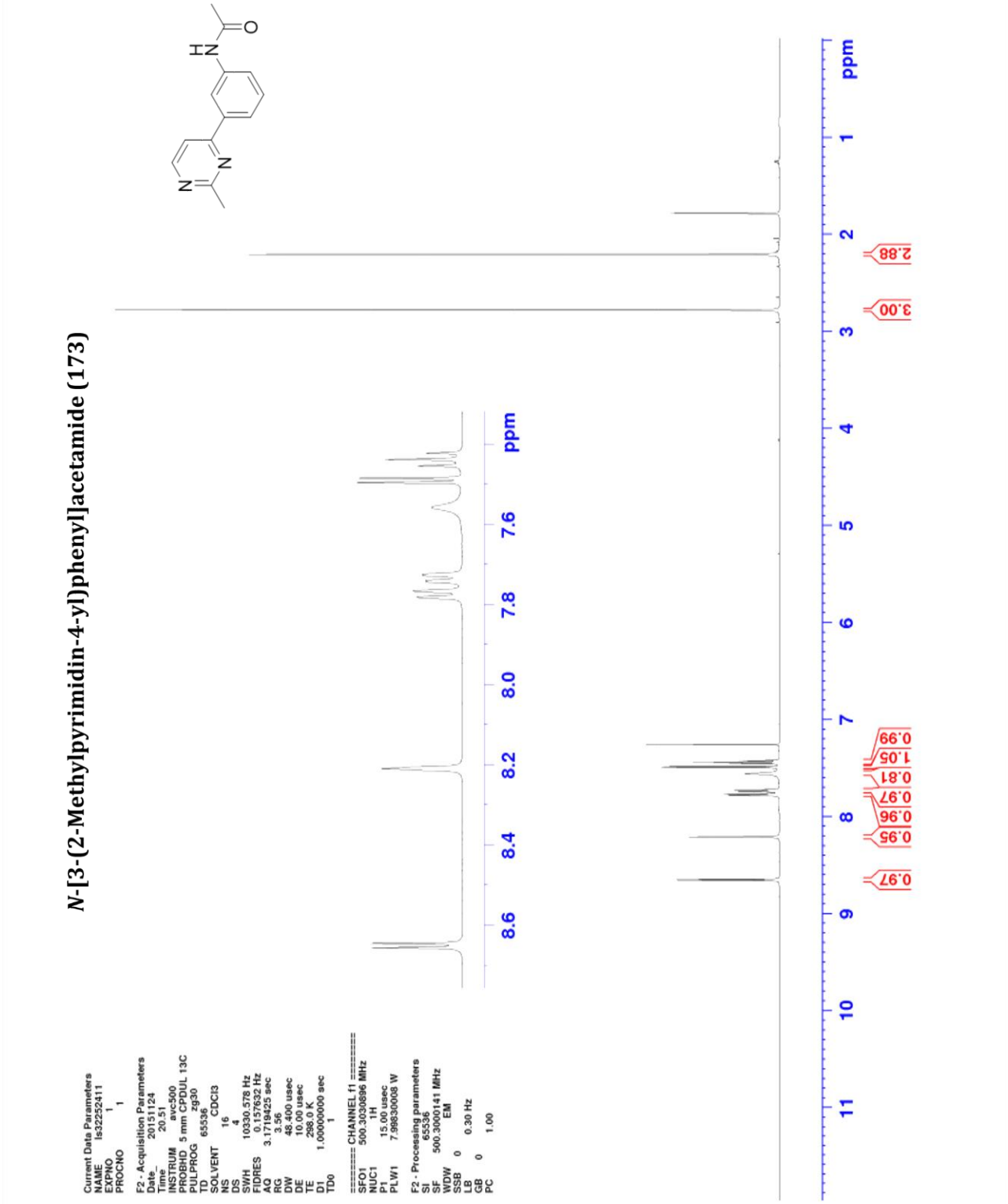



2,2-Dimethyl-N-[3-(2-methylpyrimidin-4-yl)phenyl]propanamide (172)
Current Data Parameters
NAME ls32062311
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20151123
Time 11.04
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 2.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000138 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
N
N
N
H
O
9.17
3.00
2.05
1.16
0.78
1.06
1.01
0.98
ppm
8.6
8.4
8.2
8.0
7.8
7.6
ppm

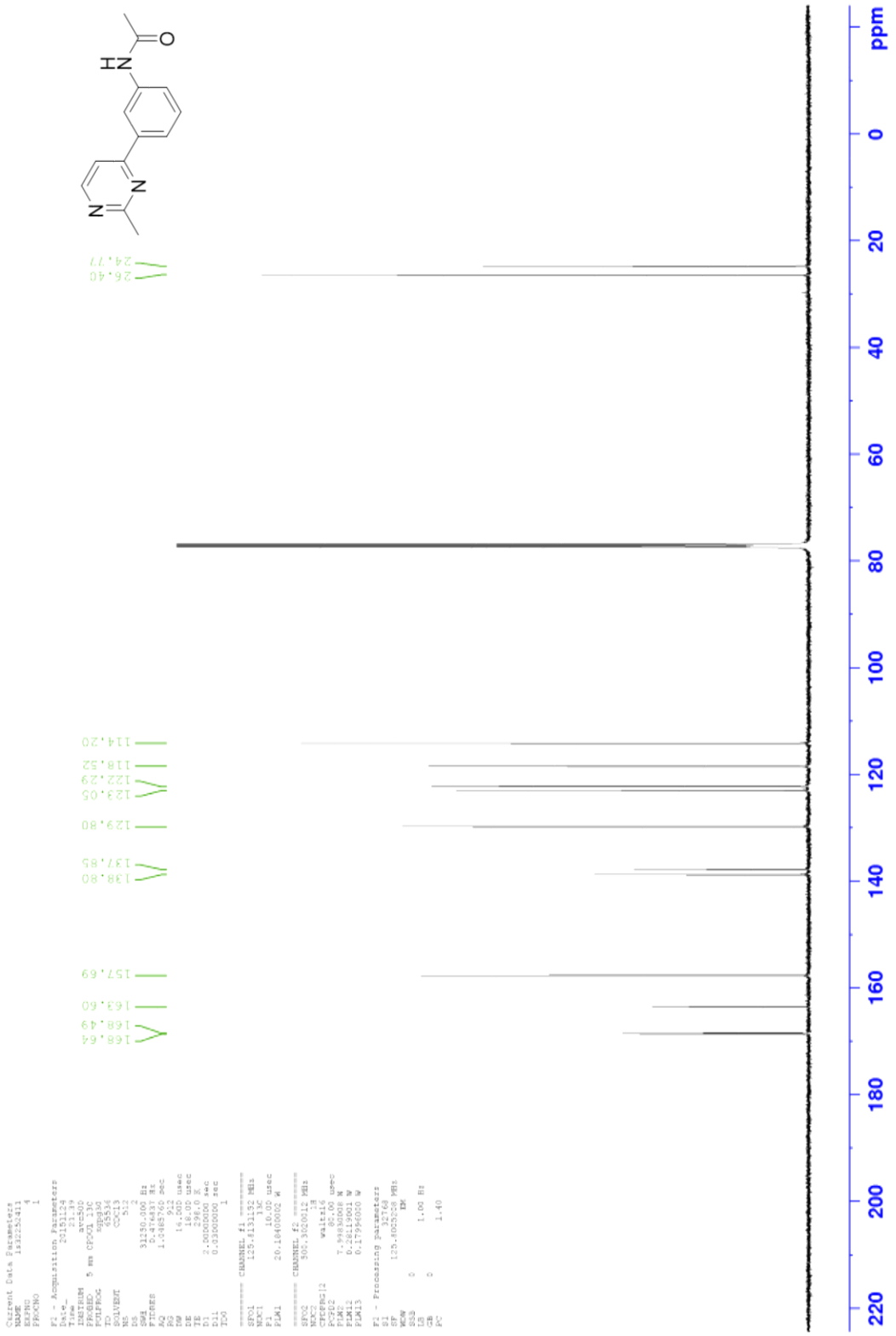
2,2-Dimethyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]propanamide (172)



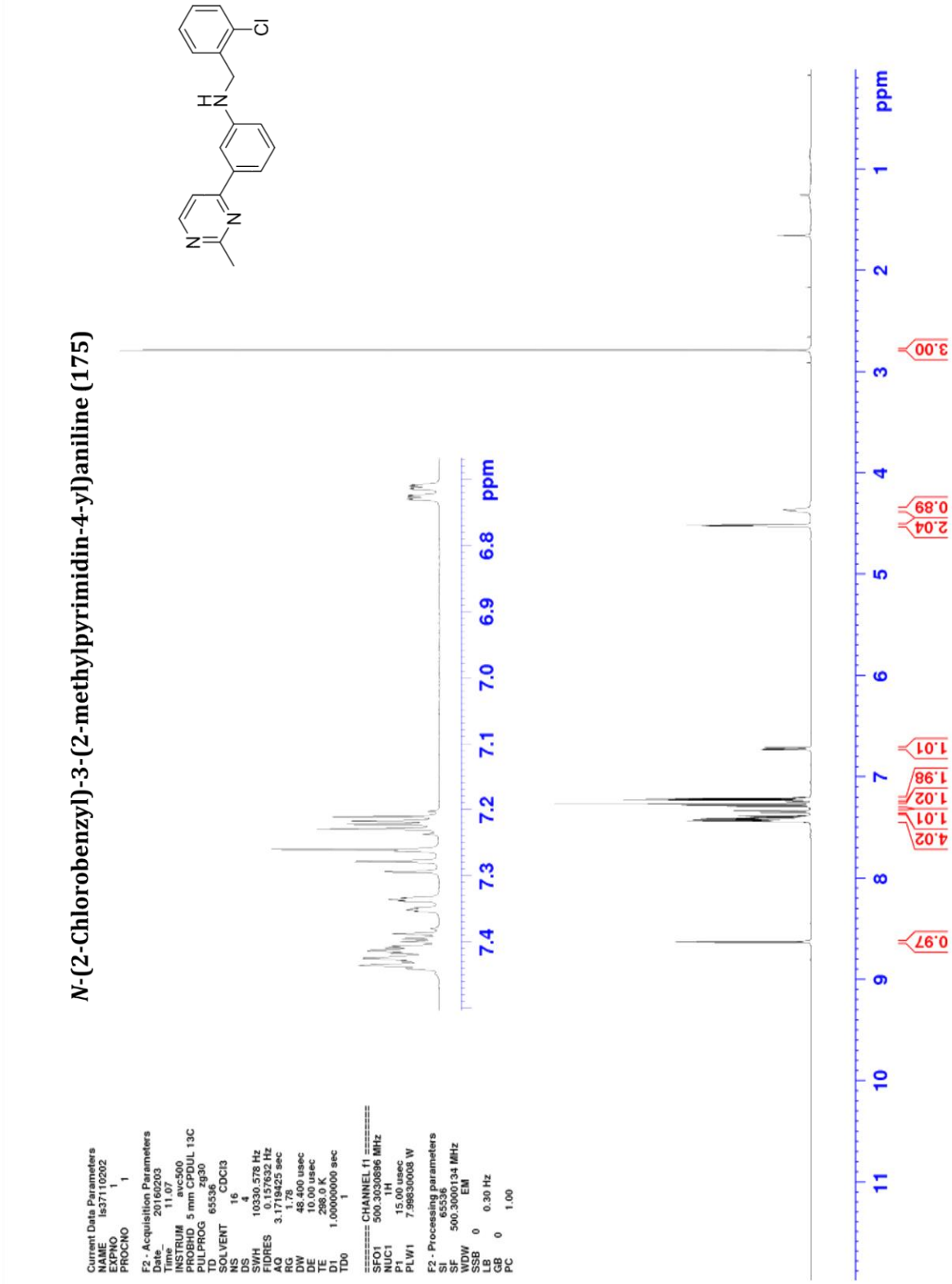
N-[3-(2-Methylpyrimidin-4-yl)phenyl]acetamide (173)



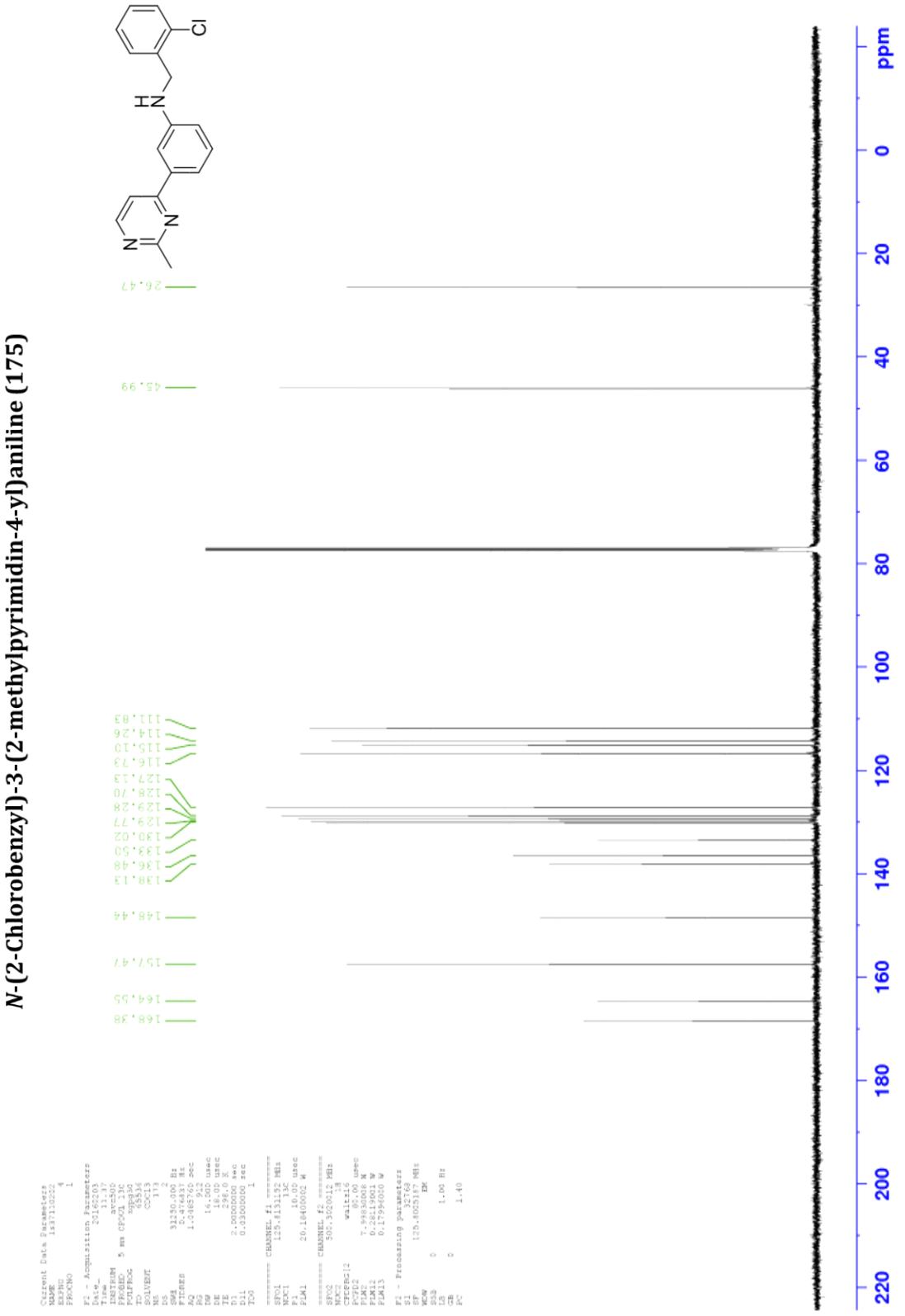
***N*-[3-(2-Methylpyrimidin-4-yl)phenyl]acetamide (173)**



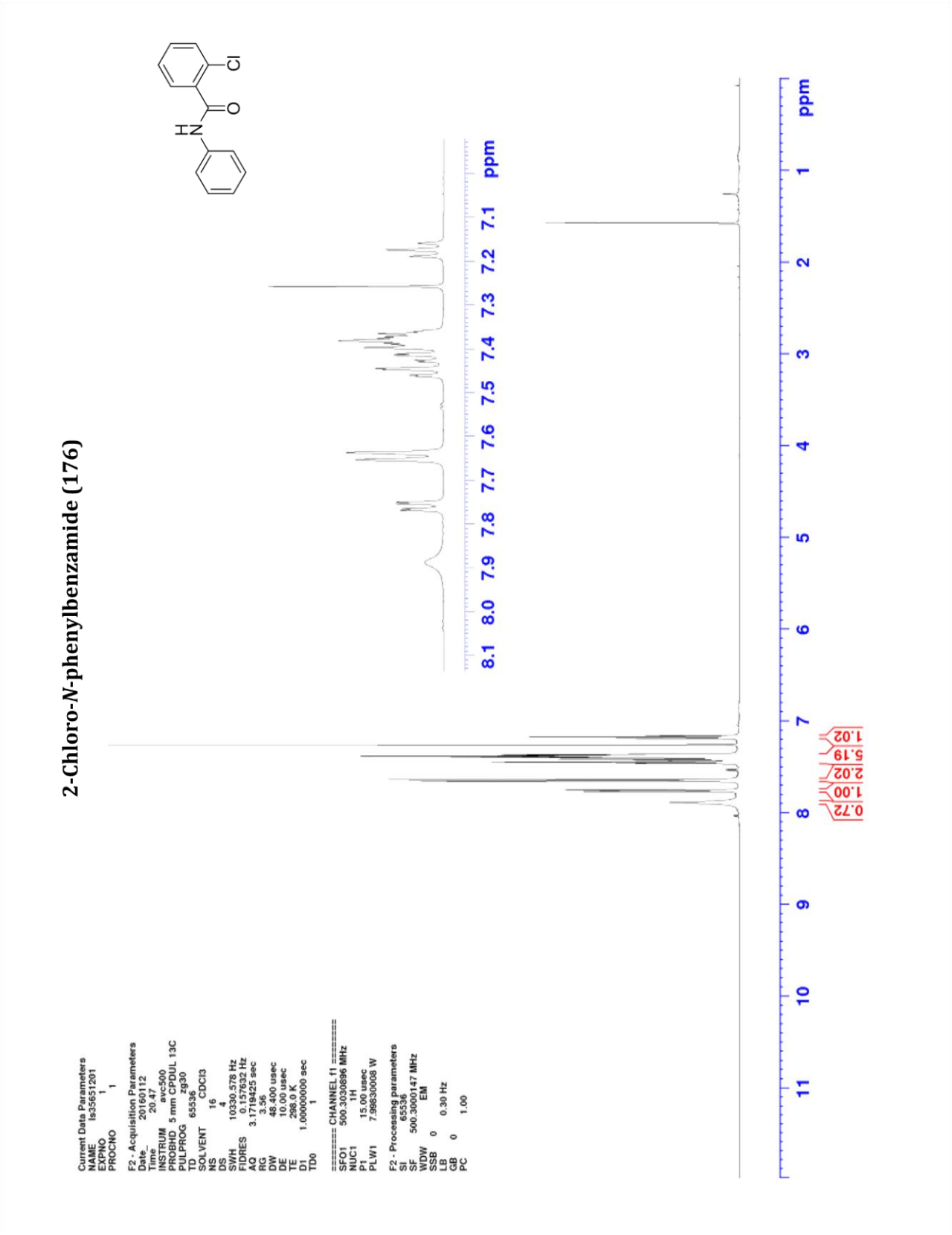
N-(2-Chlorobenzyl)-3-(2-methylpyrimidin-4-yl)aniline (175)



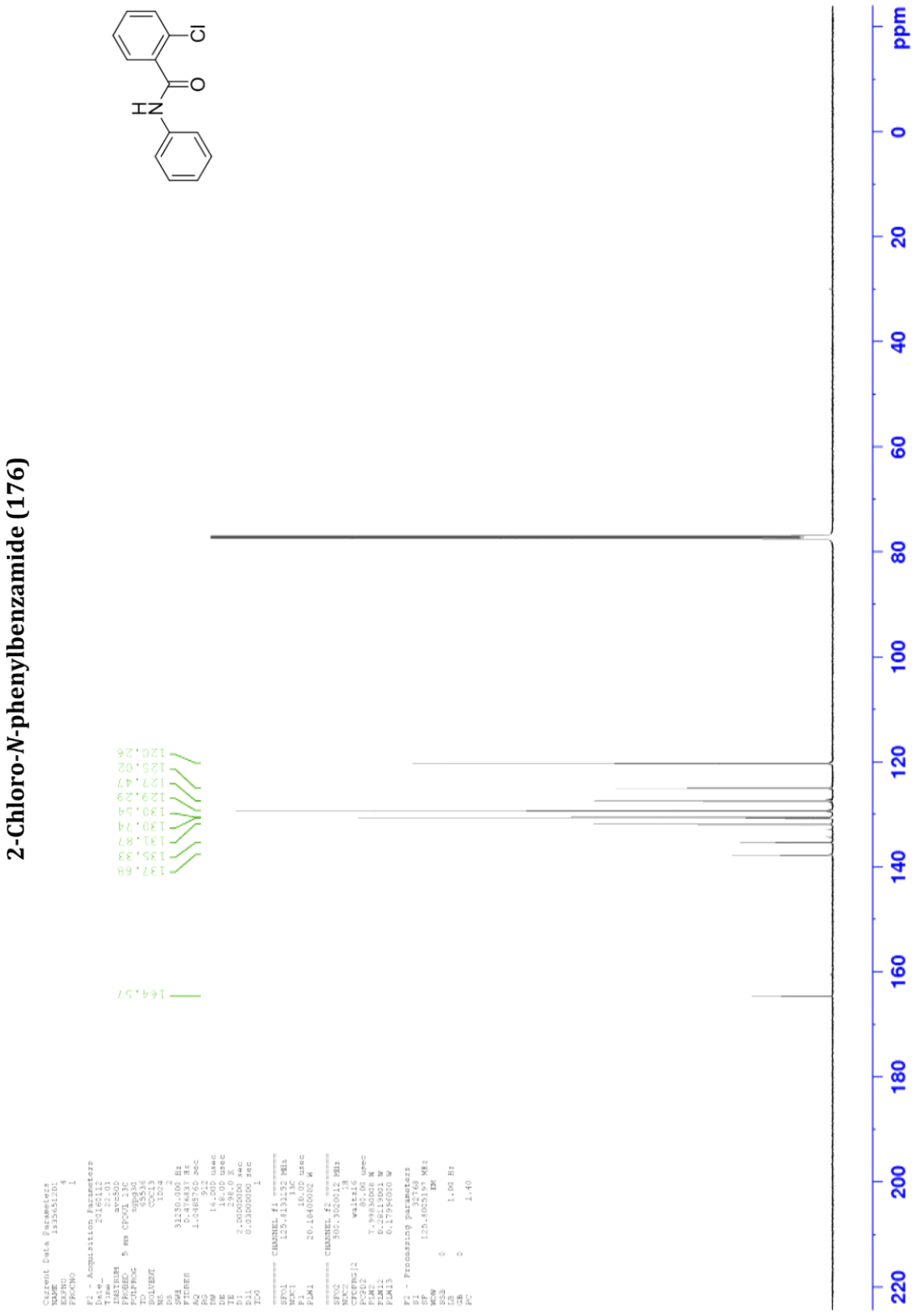
N-(2-Chlorobenzyl)-3-(2-methylpyrimidin-4-yl)aniline (175)

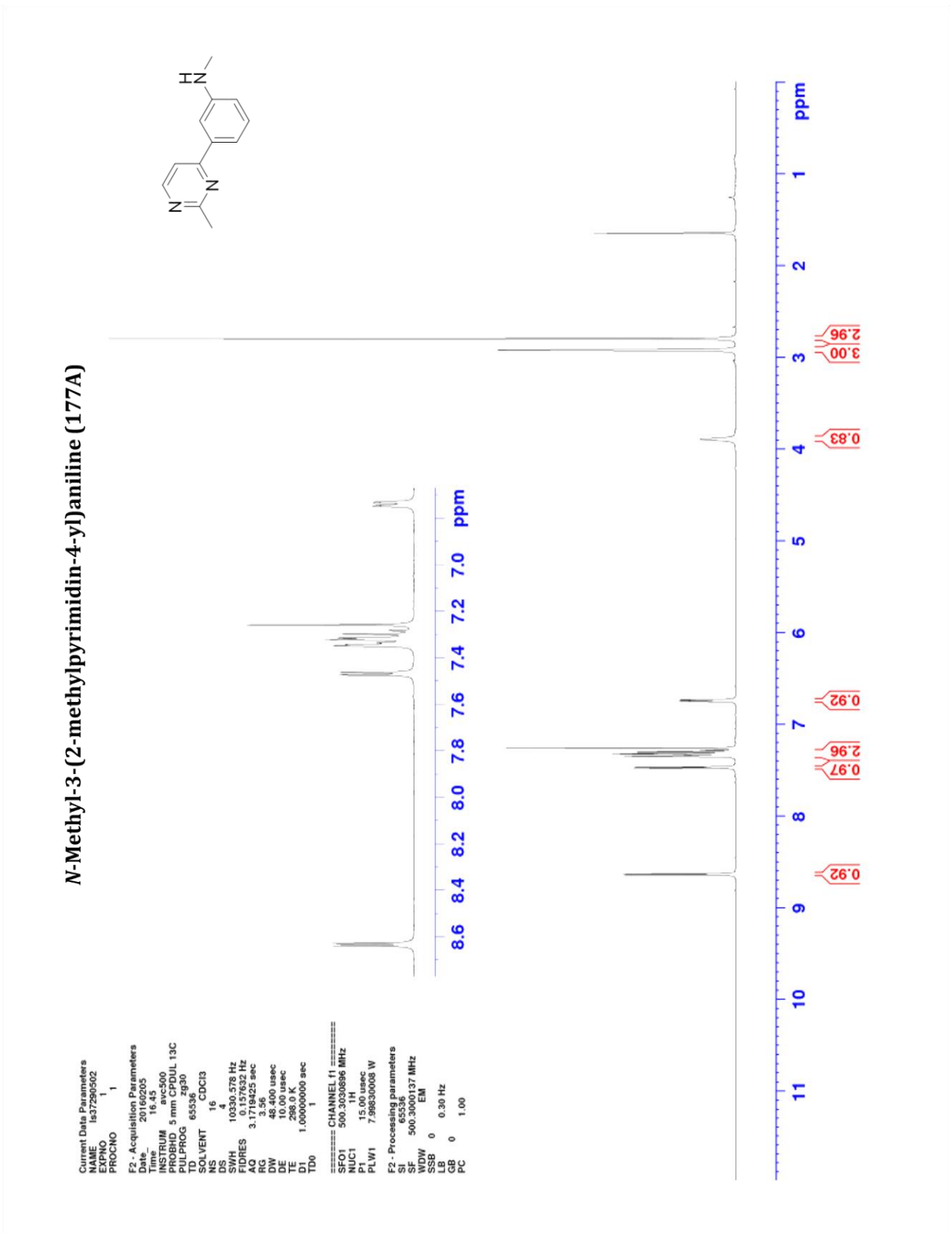


2-Chloro-*N*-phenylbenzamide (176)

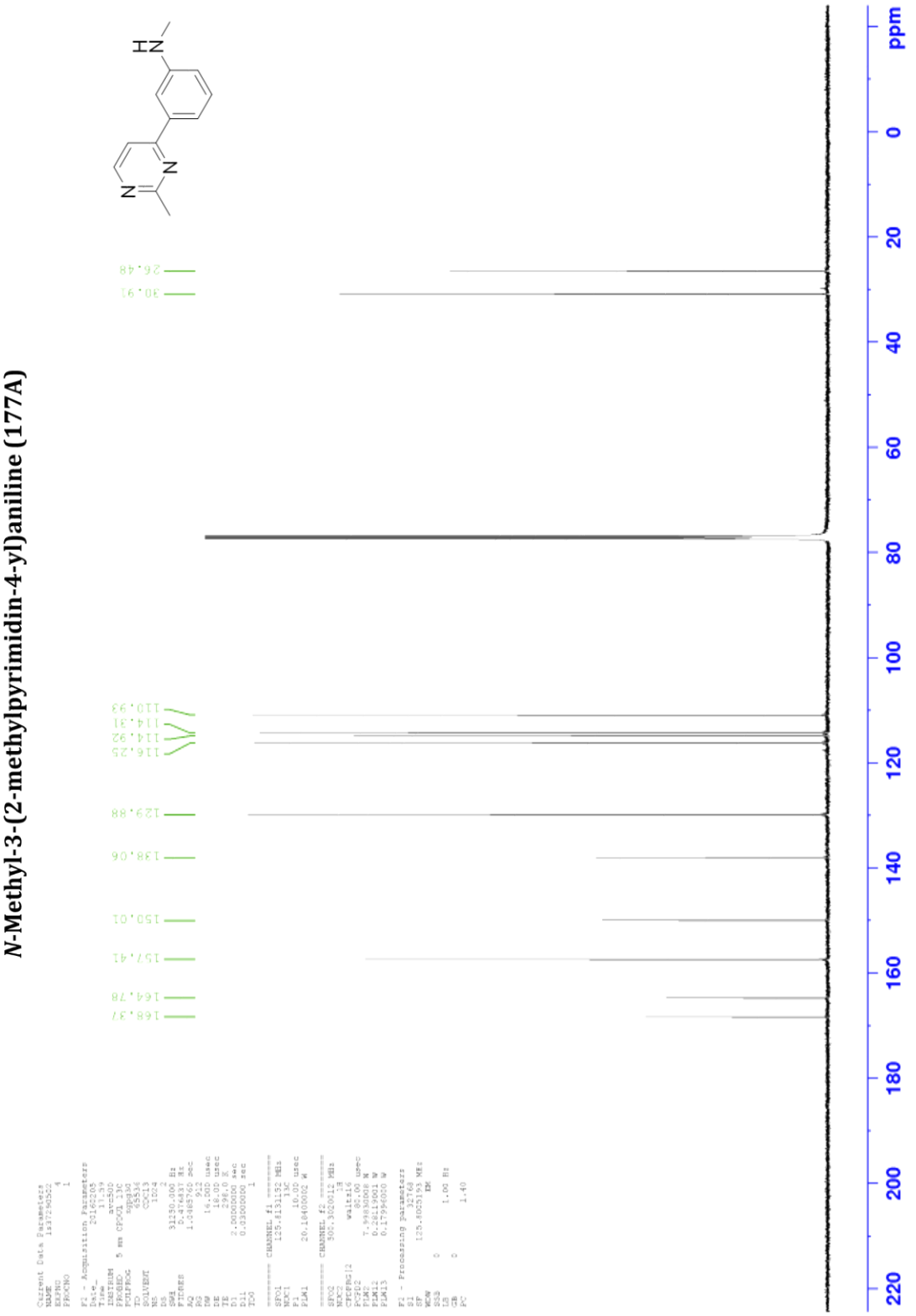


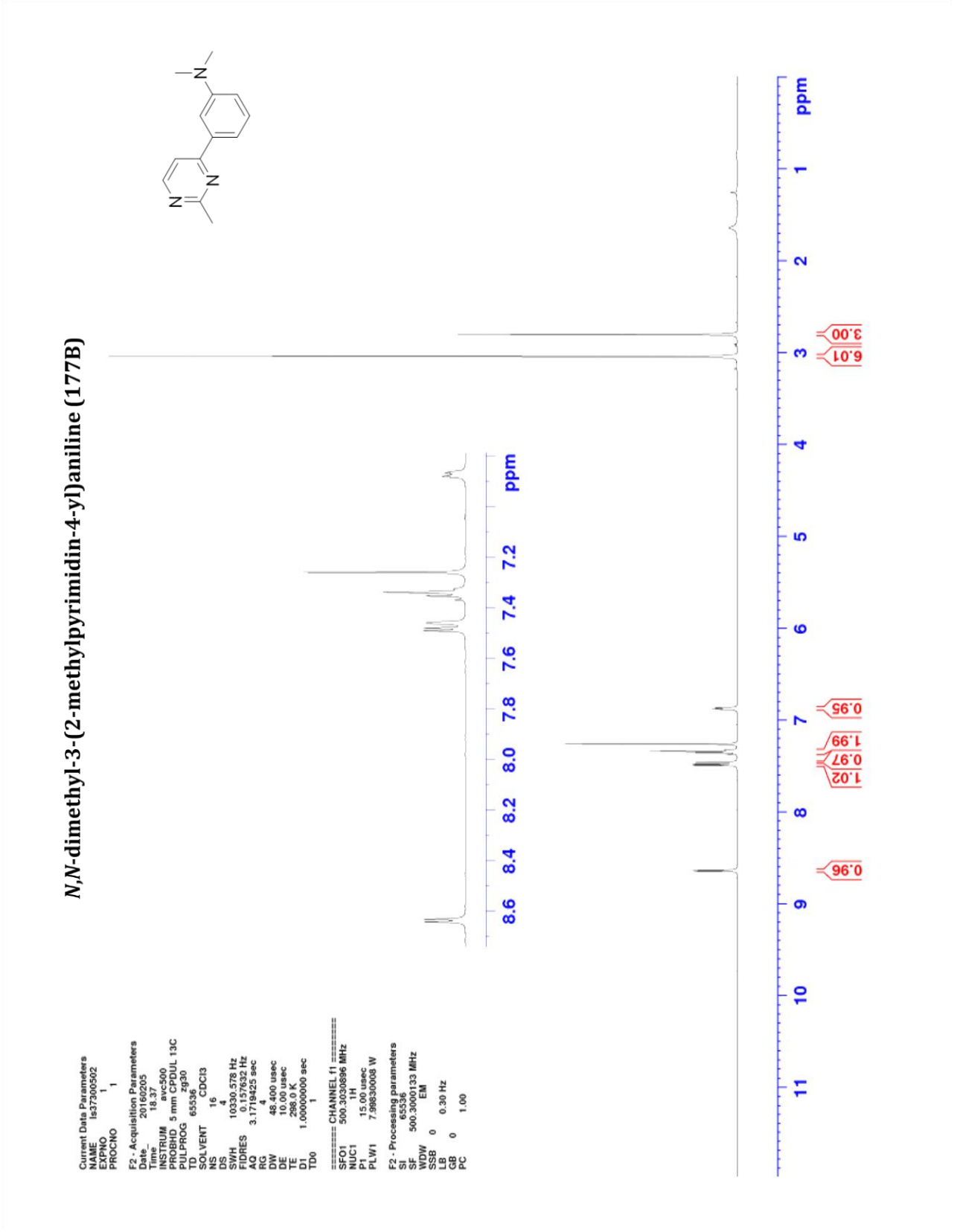
2-Chloro-*N*-phenylbenzamide (176)



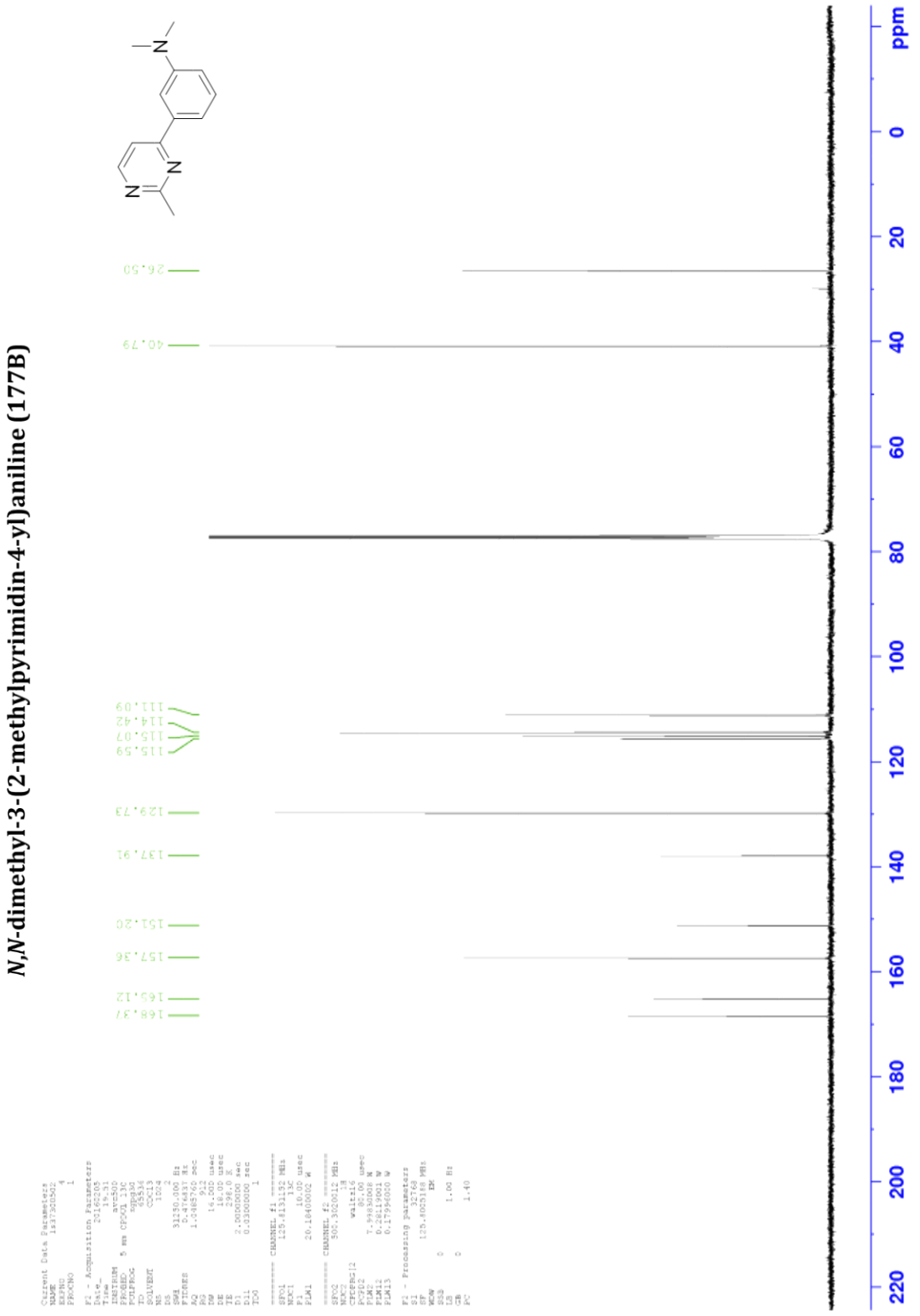

N-Methyl-3-(2-methylpyrimidin-4-yl)aniline (177A)
Current Data Parameters
NAME ls37290502
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20160205
Time 16.45
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000137 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
8.6 8.4 8.2 8.0 7.8 7.6 7.4 7.2 7.0 ppm
0.92
0.97
2.96
0.92
0.83
3.00
2.96
11 10 9 8 7 6 5 4 3 2 1 ppm

N-Methyl-3-(2-methylpyrimidin-4-yl)aniline (177A)

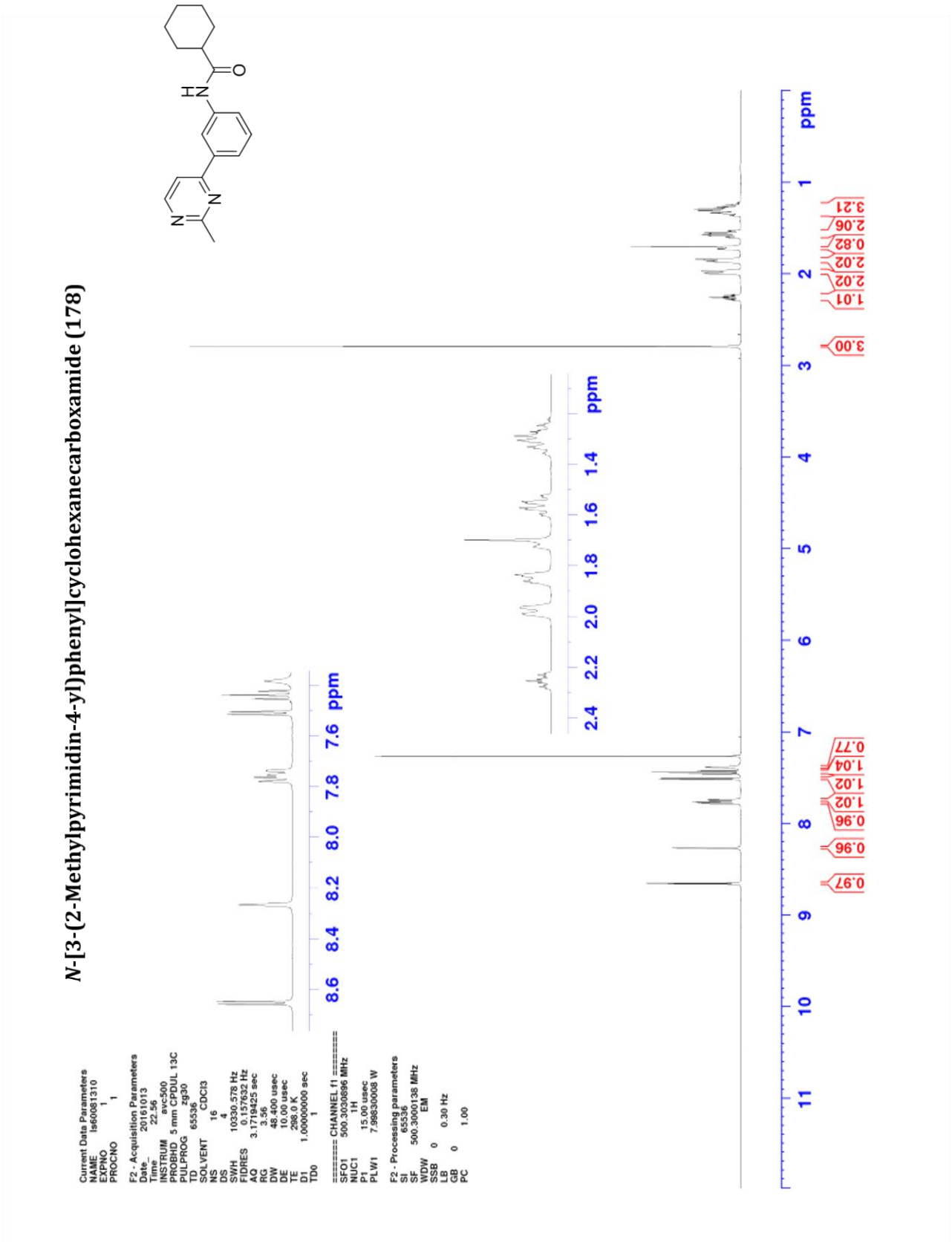



N,N-dimethyl-3-(2-methylpyrimidin-4-yl)aniline (177B)
Current Data Parameters
NAME ls37300502
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20160205
Time 18.37
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 4
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000133 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
8.6
8.4
8.2
8.0
7.8
7.6
7.4
7.2
ppm
0.96
1.02
0.97
1.99
0.95
6.01
3.00
11
10
9
8
7
6
5
4
3
2
1
ppm

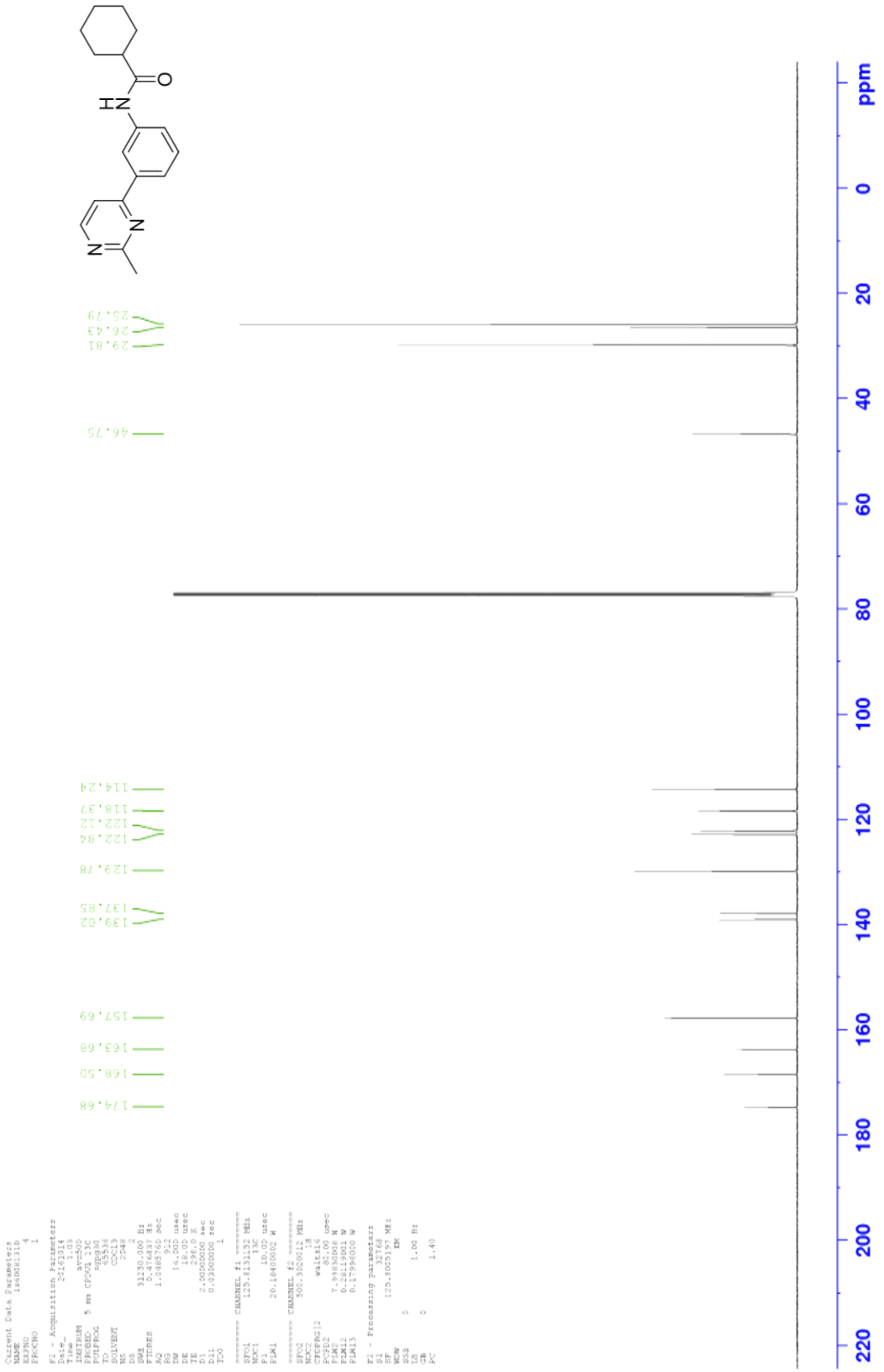
N,N-dimethyl-3-(2-methylpyrimidin-4-yl)aniline (177B)

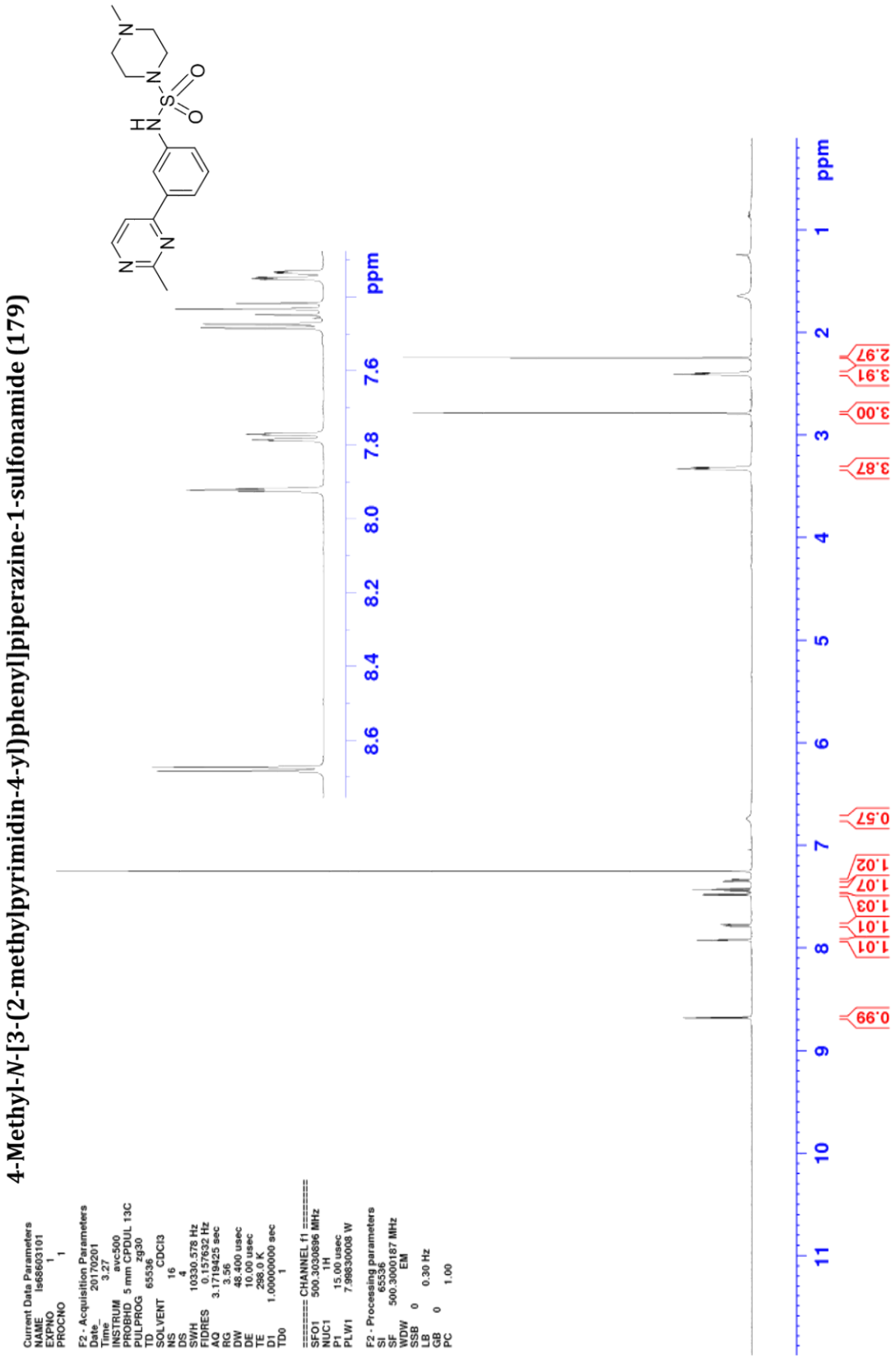


N-[3-(2-Methylpyrimidin-4-yl)phenyl]cyclohexanecarboxamide (178)

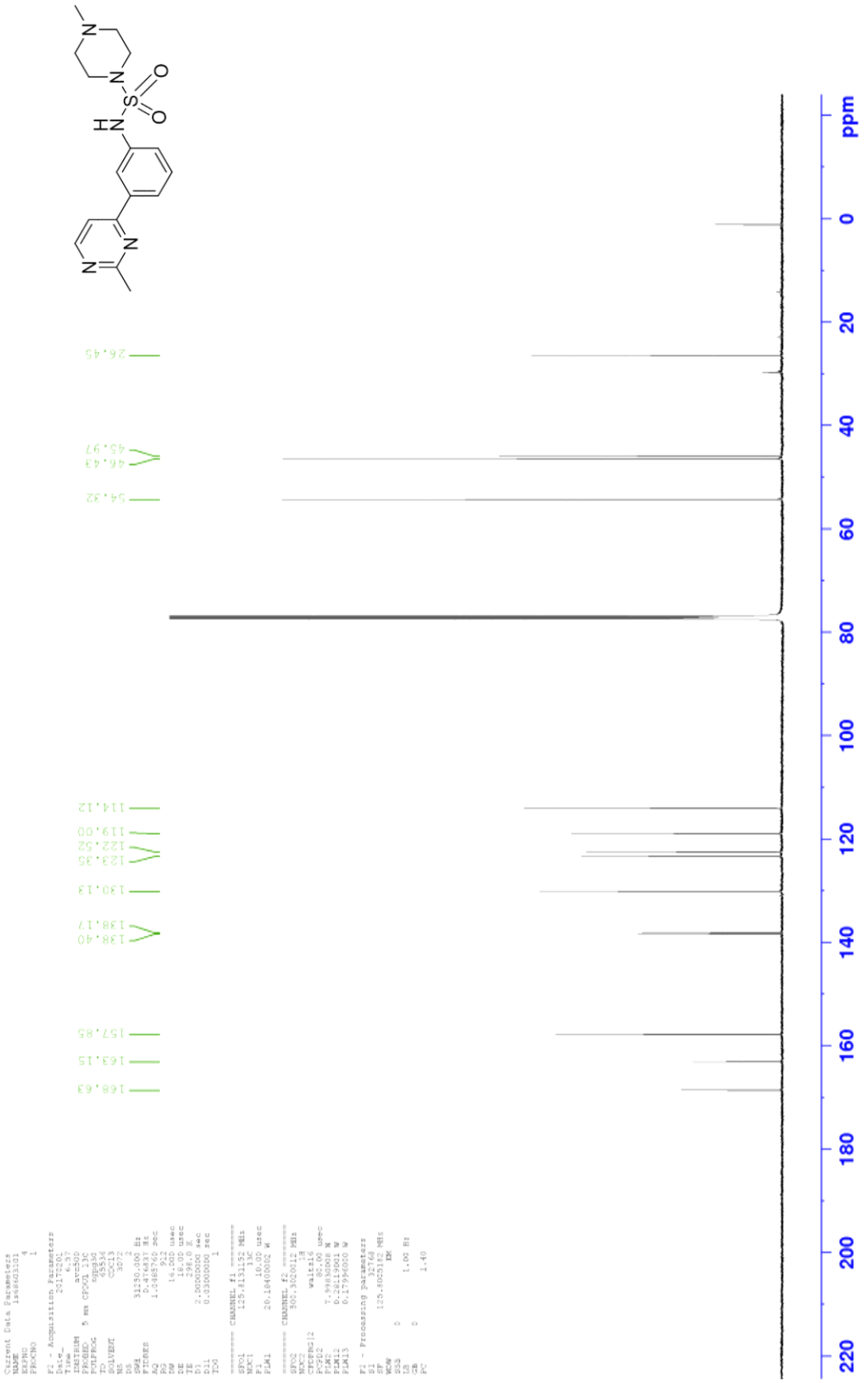


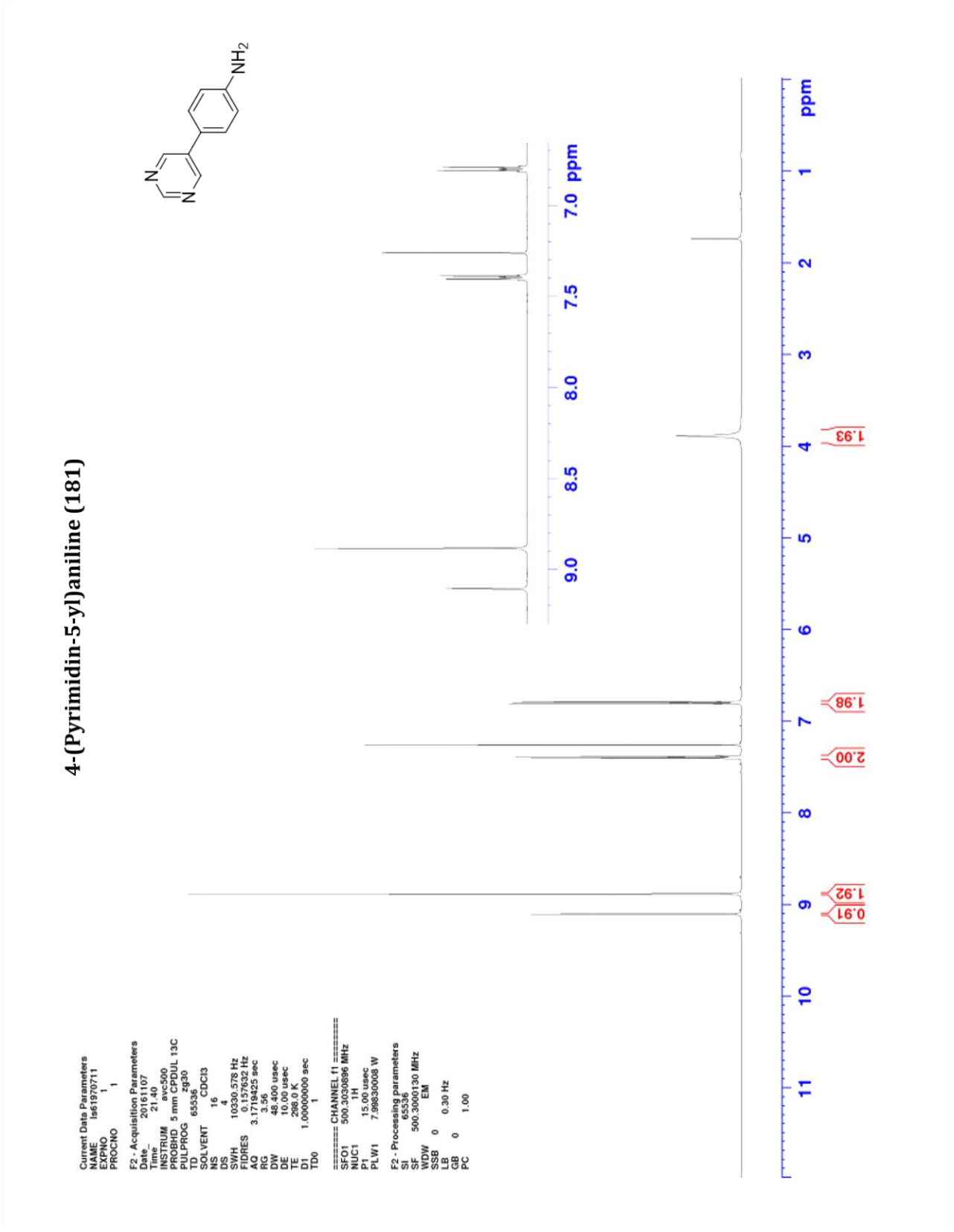
N-[3-(2-Methylpyrimidin-4-yl)phenyl]cyclohexanecarboxamide (178)



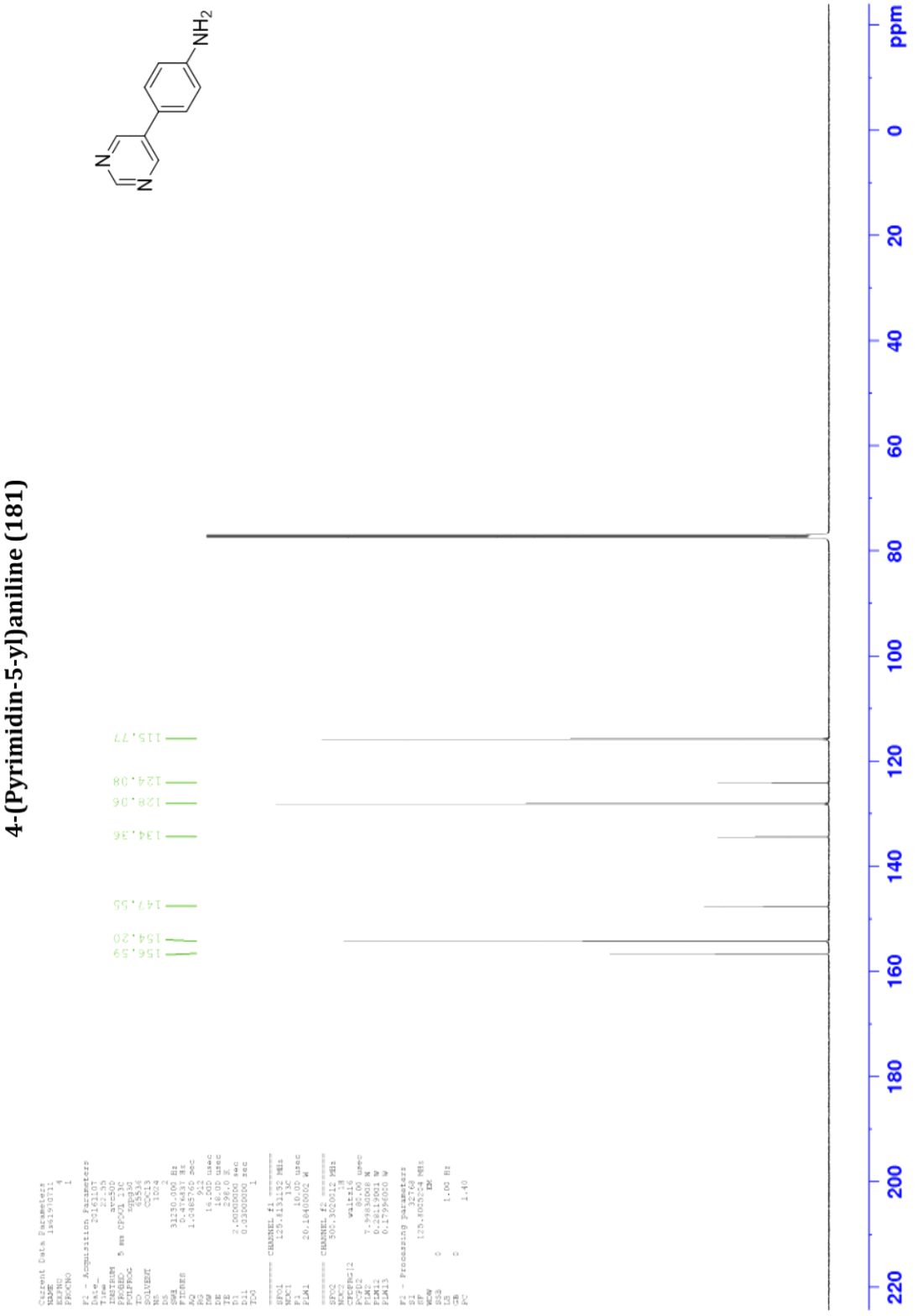

4-Methyl-N-[3-(2-methylpyrimidin-4-yl)phenyl]piperazine-1-sulfonamide (179)
Current Data Parameters
NAME ls68603101
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20170201
Time 3.27
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000187 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
8.6
8.4
8.2
8.0
7.8
7.6
ppm
0.99
1.01
1.01
1.01
1.03
1.07
1.02
0.57
3.87
3.00
3.91
2.97
11
10
9
8
7
6
5
4
3
2
1
ppm

4-Methyl-*N*-[3-(2-methylpyrimidin-4-yl)phenyl]piperazine-1-sulfonamide (179)

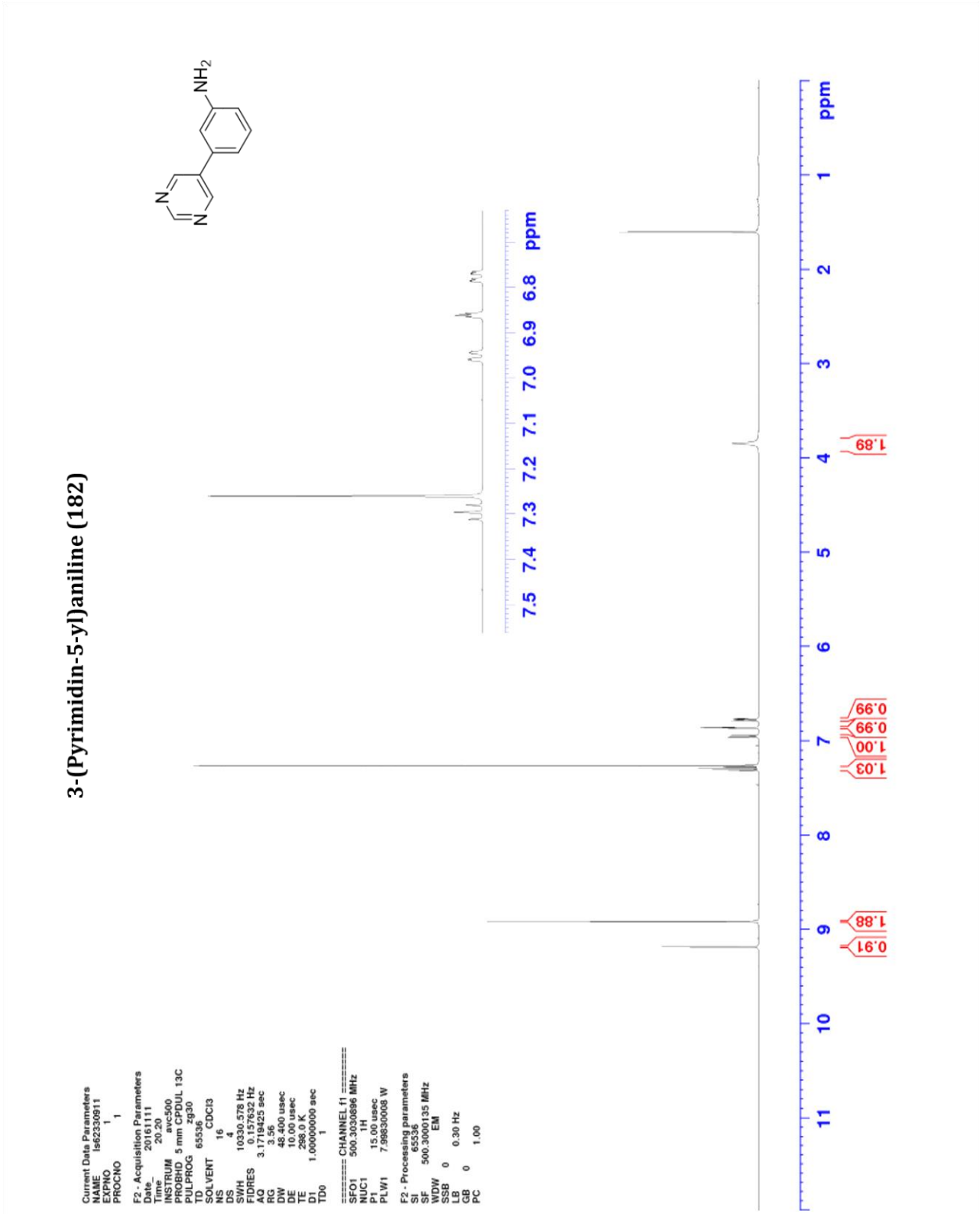



4-(Pyrimidin-5-yl)aniline (181)
NH2
N
N
Current Data Parameters
NAME ls61970711
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20161107
Time 21.40
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.1719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000130 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
9.0
8.5
8.0
7.5
7.0 ppm
11
10
9
8
7
6
5
4
3
2
1 ppm
0.91
1.92
2.00
1.98
1.93

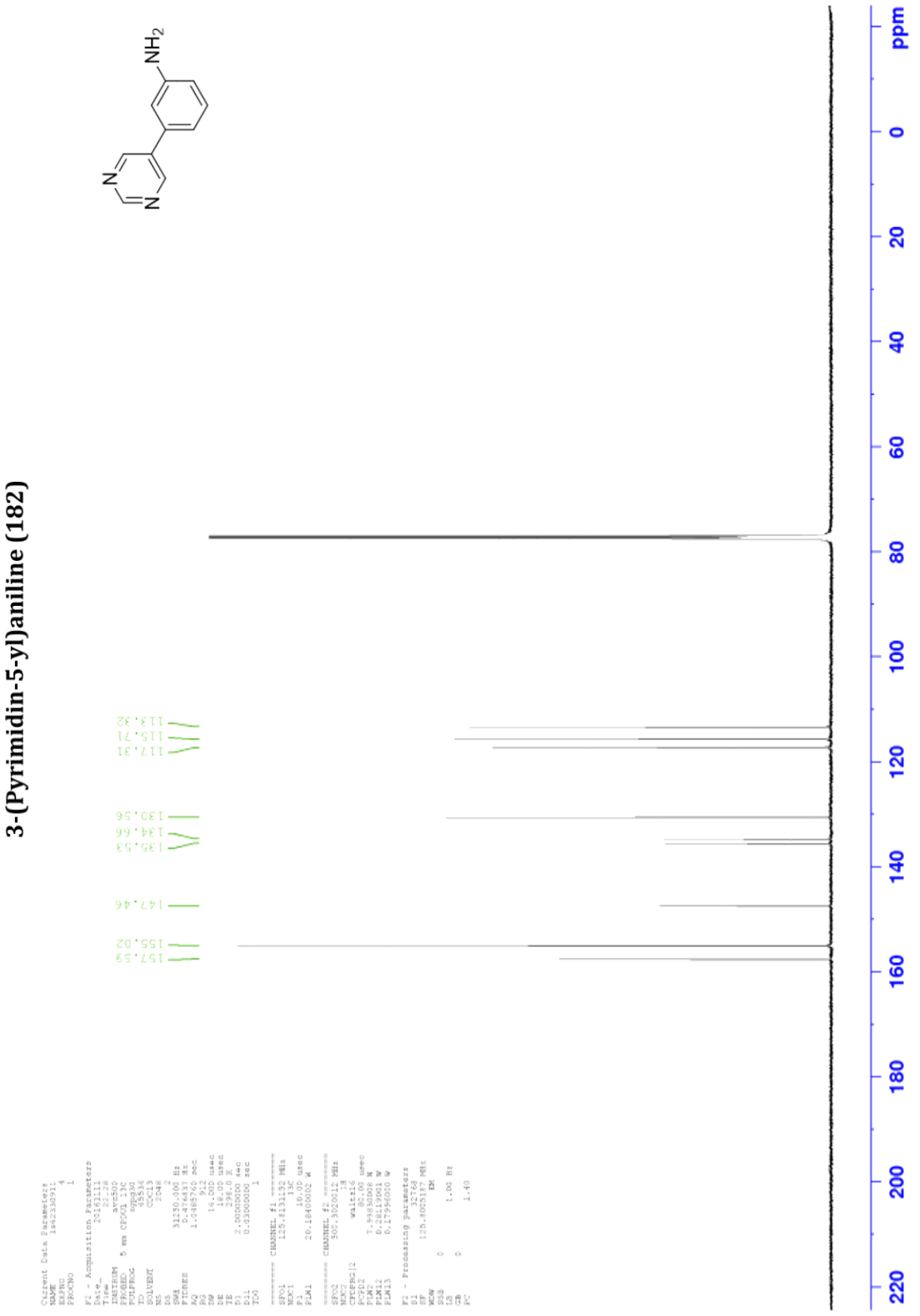
4-(Pyrimidin-5-yl)aniline (181)

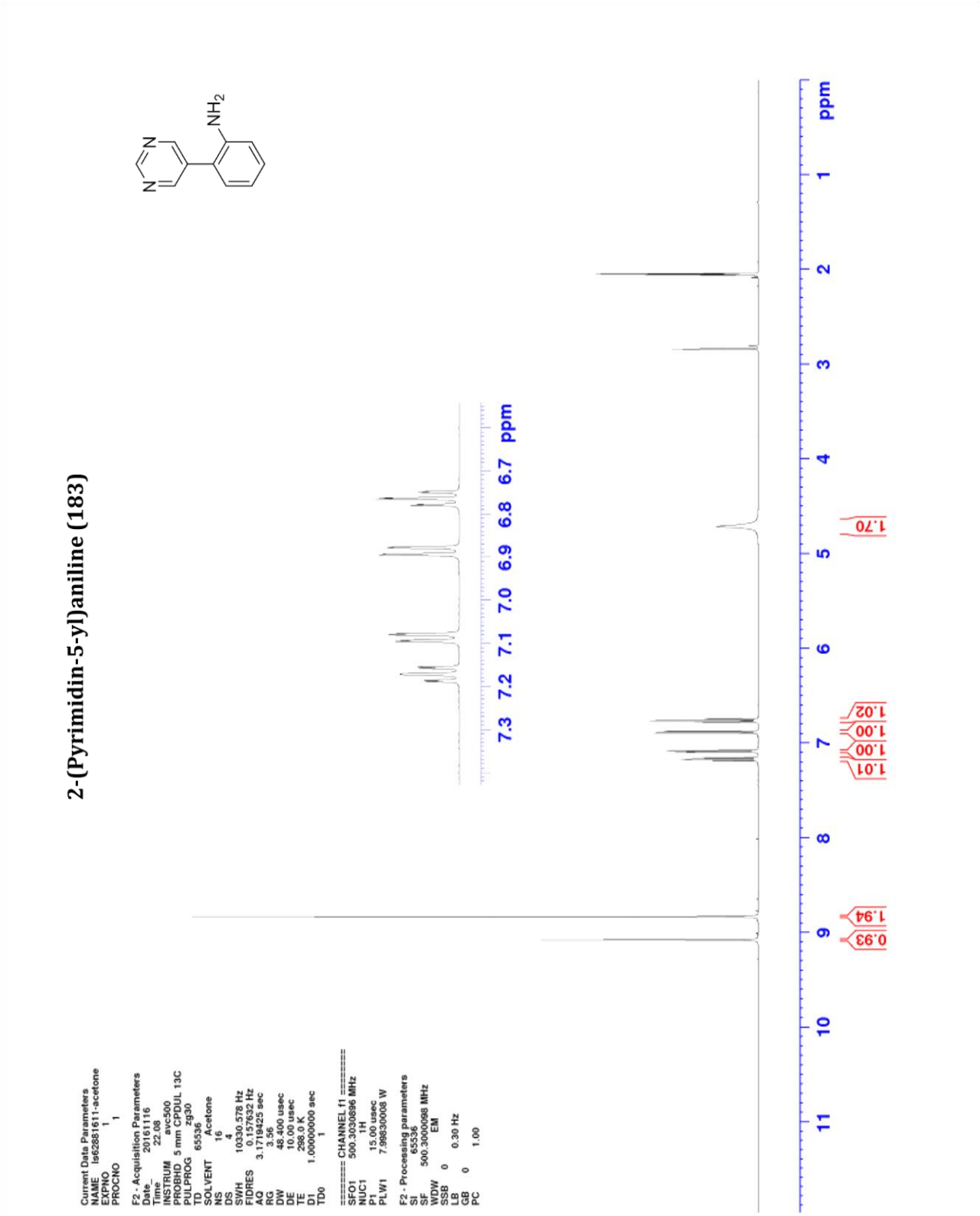


3-(Pyrimidin-5-yl)aniline (182)

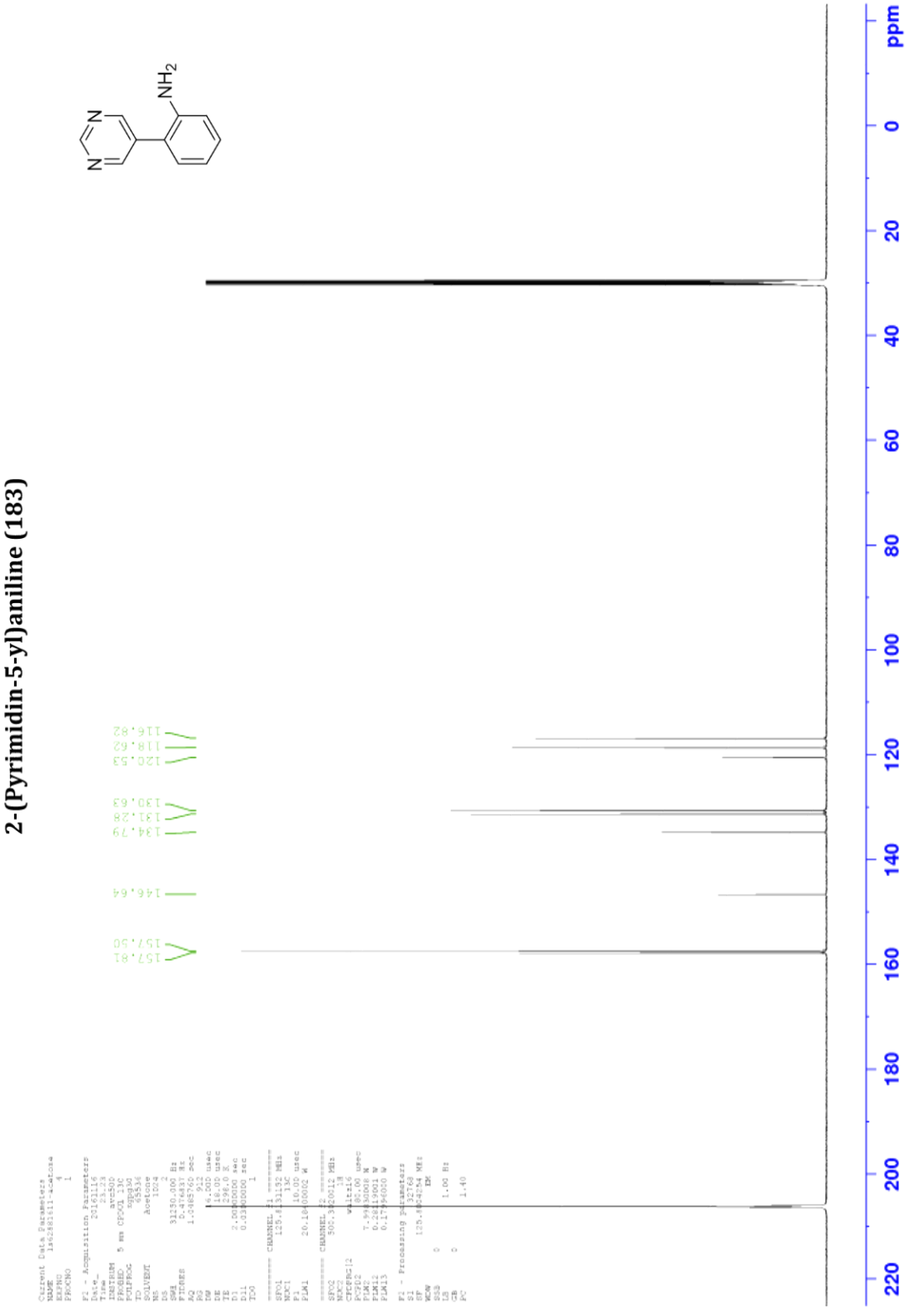


3-(Pyrimidin-5-yl)aniline (182)

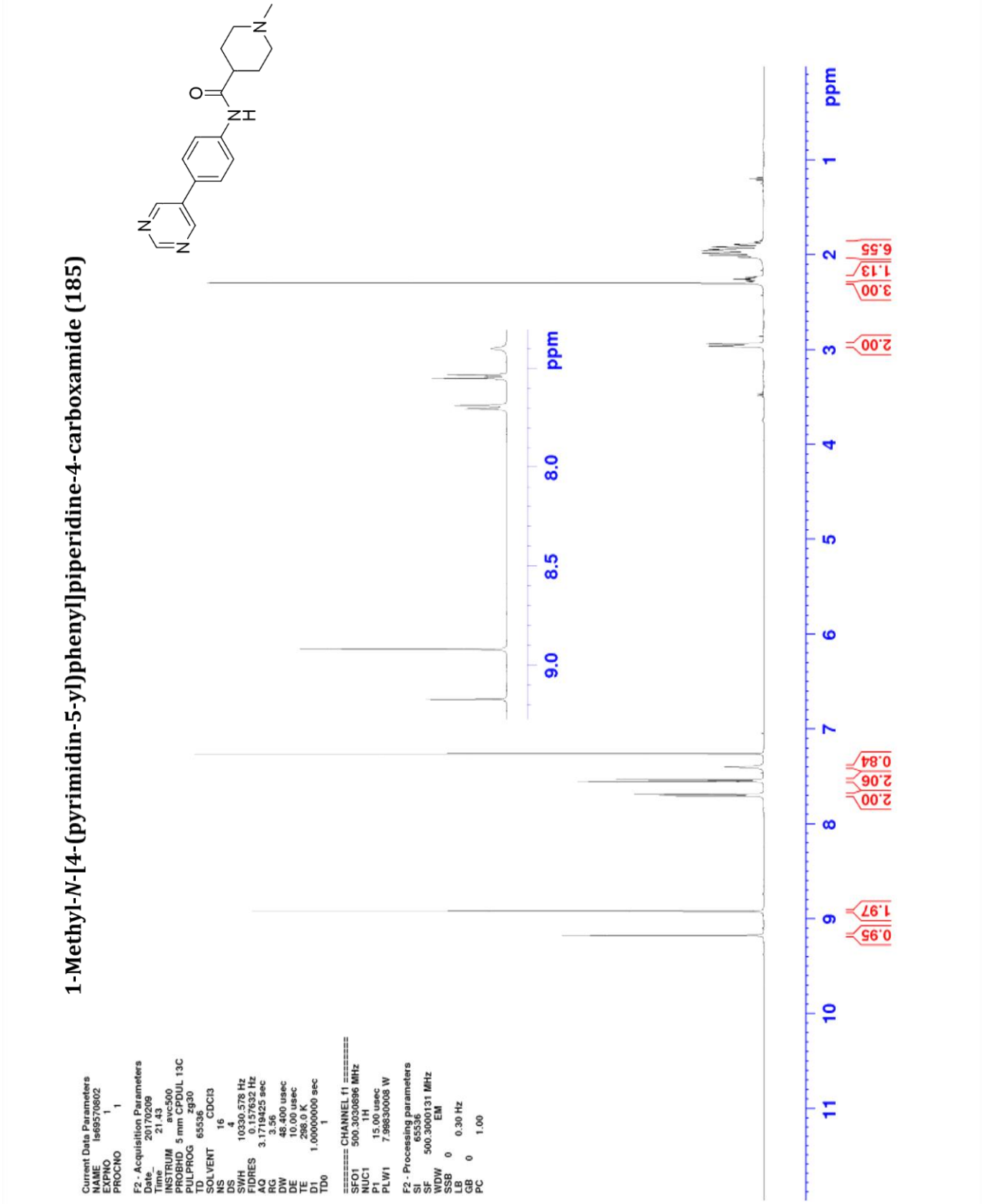



2-(Pyrimidin-5-yl)aniline (183)
NH2
N
N
Current Data Parameters
NAME Is62881611-acetone
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20161116
Time 22.08
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT Acetone
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000098 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
7.3 7.2 7.1 7.0 6.9 6.8 6.7 ppm
0.93
1.94
1.01
1.00
1.00
1.02
1.70
11 10 9 8 7 6 5 4 3 2 1 ppm

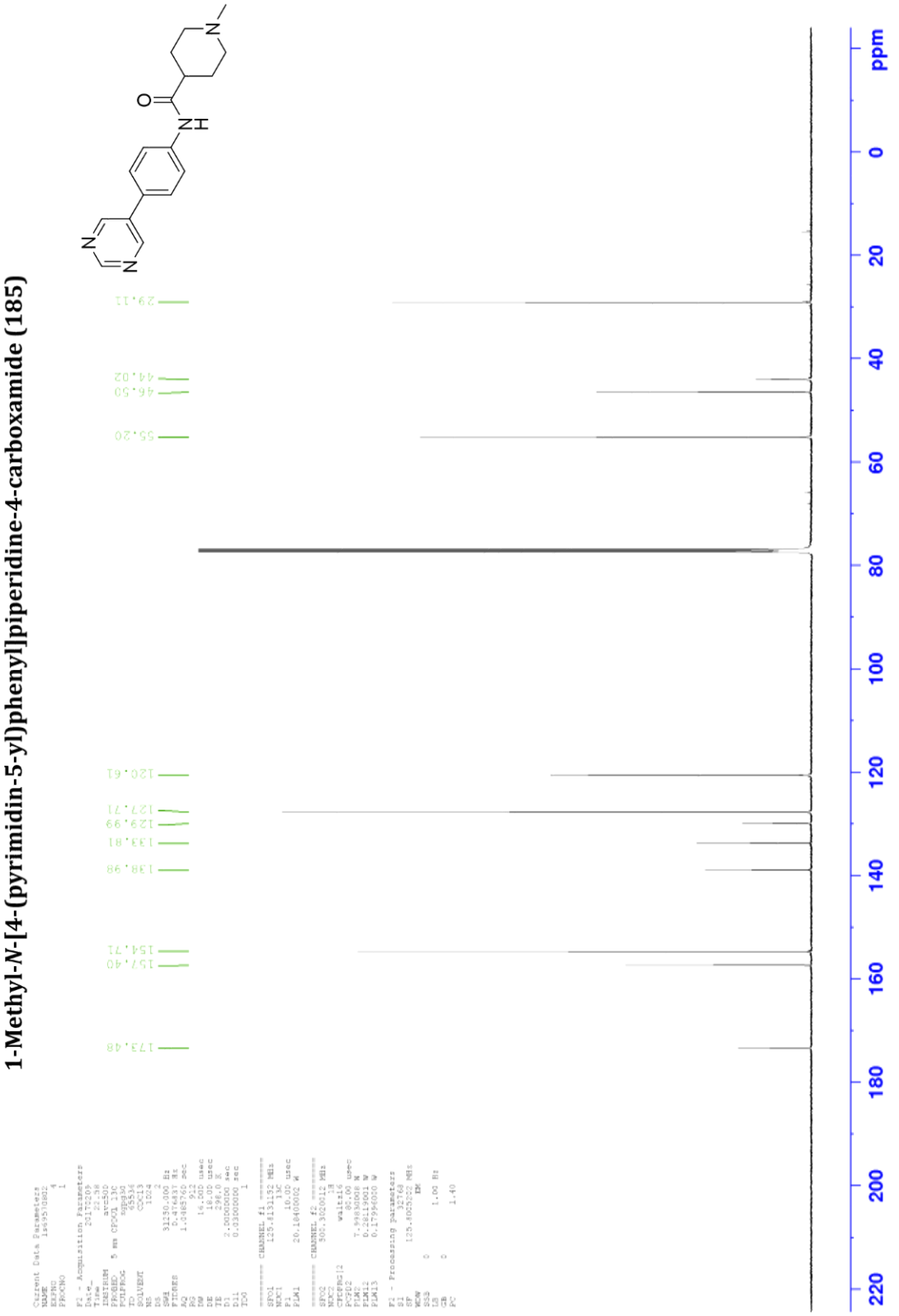
2-(Pyrimidin-5-yl)aniline (183)



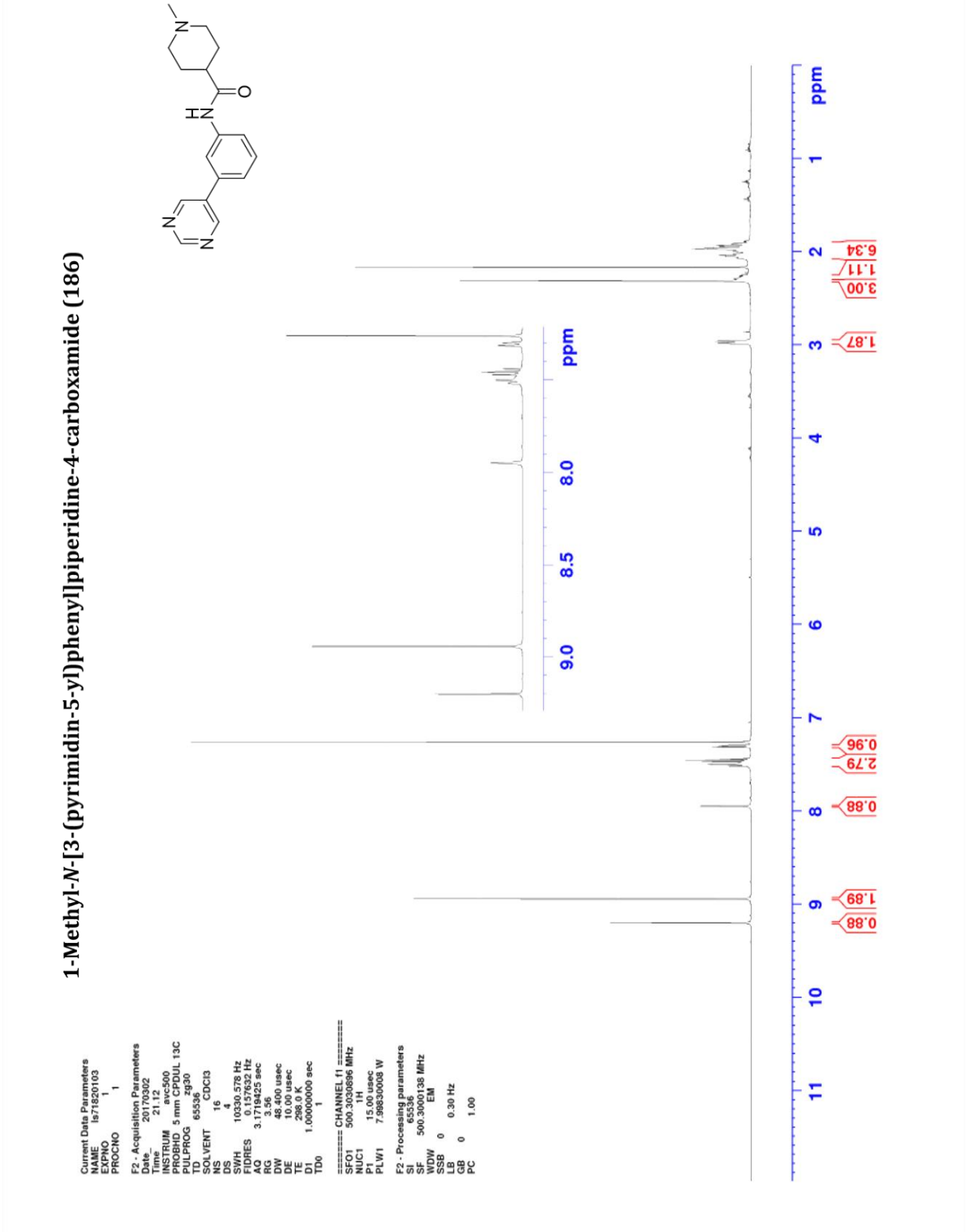
1-Methyl-*N*-[4-(pyrimidin-5-yl)phenyl]piperidine-4-carboxamide (185)



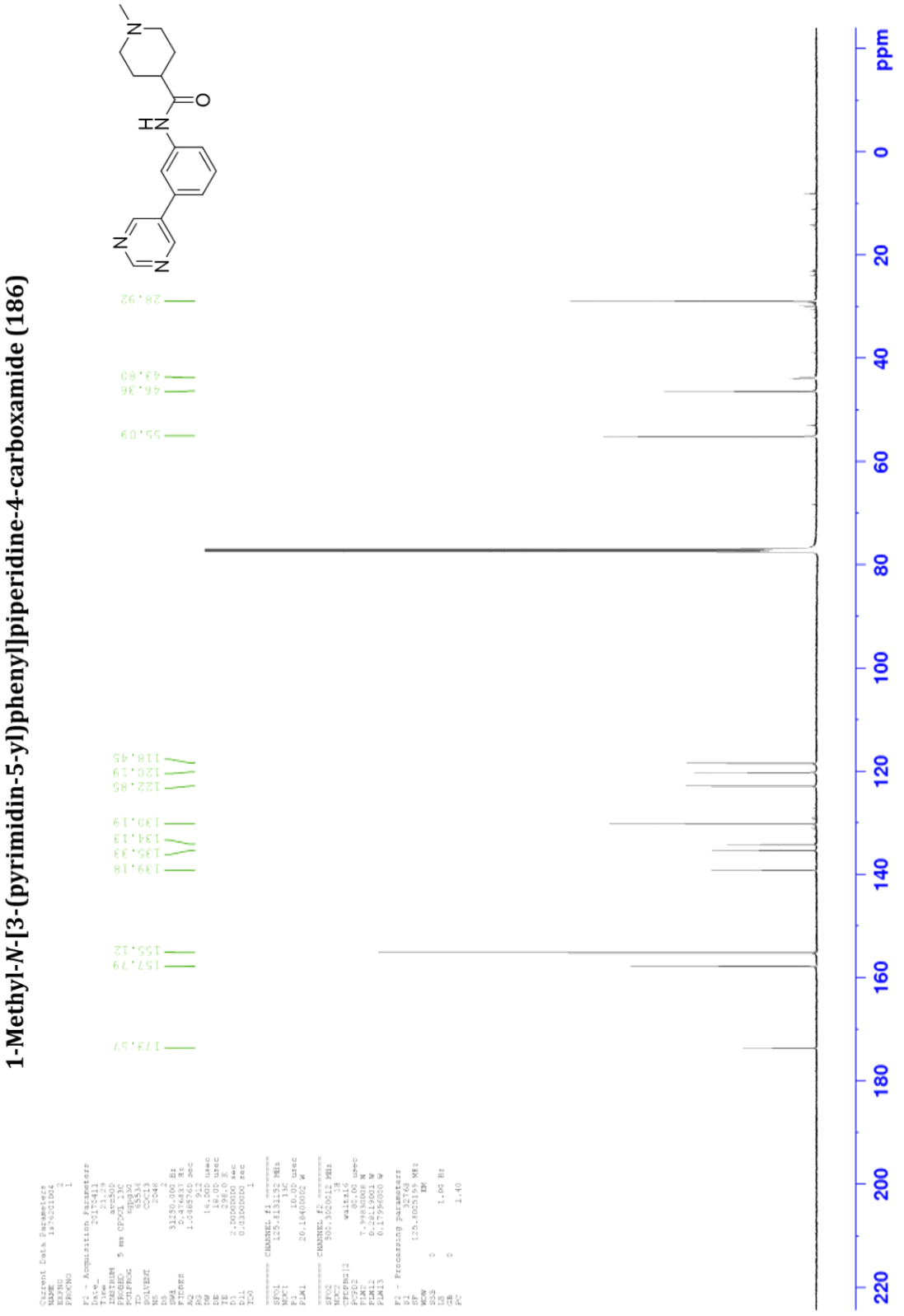
1-Methyl-*N*-[4-(pyrimidin-5-yl)phenyl]piperidine-4-carboxamide (185)



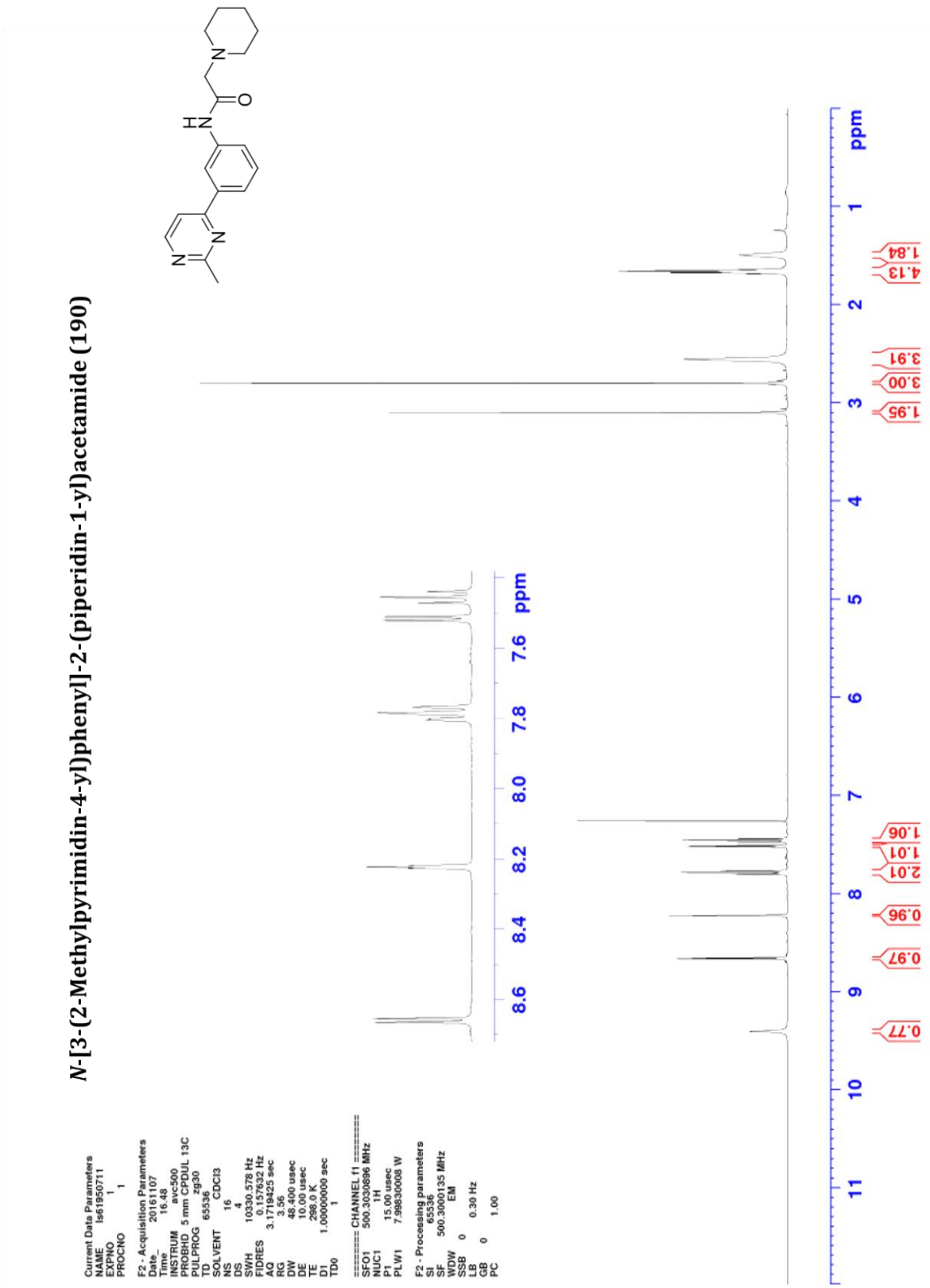
1-Methyl-*N*-[3-(pyrimidin-5-yl)phenyl]piperidine-4-carboxamide (186)



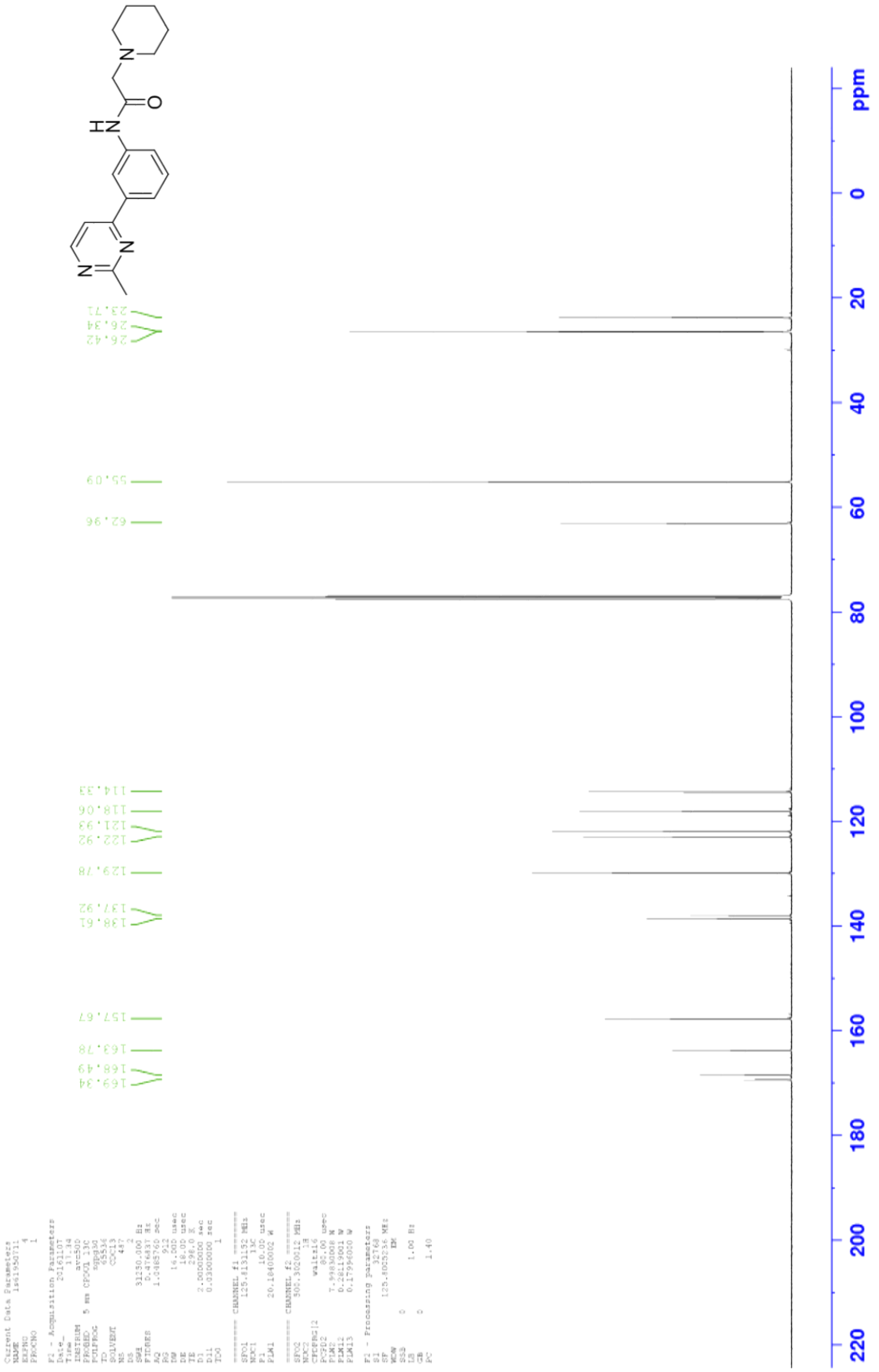
1-Methyl-*N*-[3-(pyrimidin-5-yl)phenyl]piperidine-4-carboxamide (186)



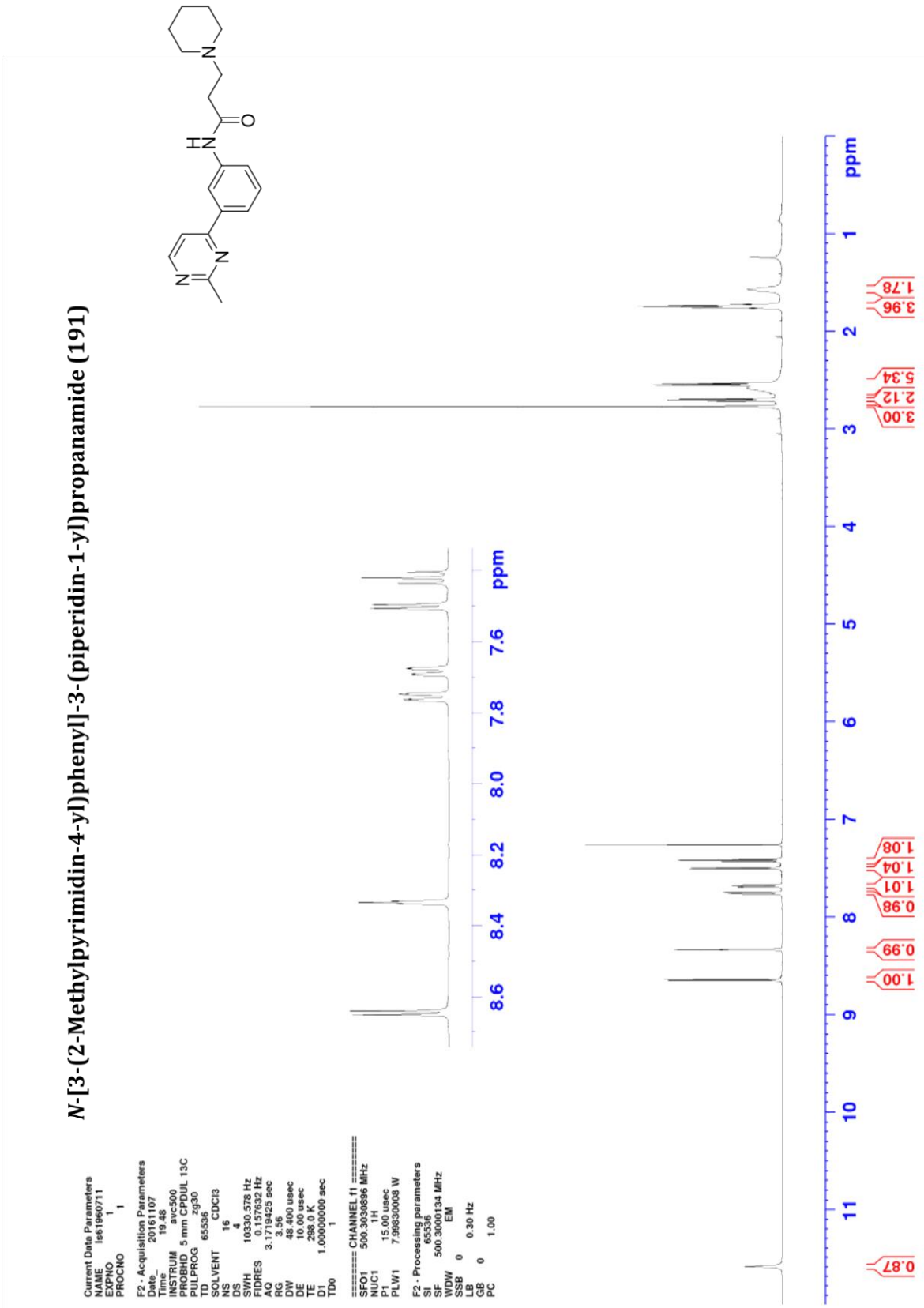
N-[3-(2-Methylpyrimidin-4-yl)phenyl]-2-(piperidin-1-yl)acetamide (190)



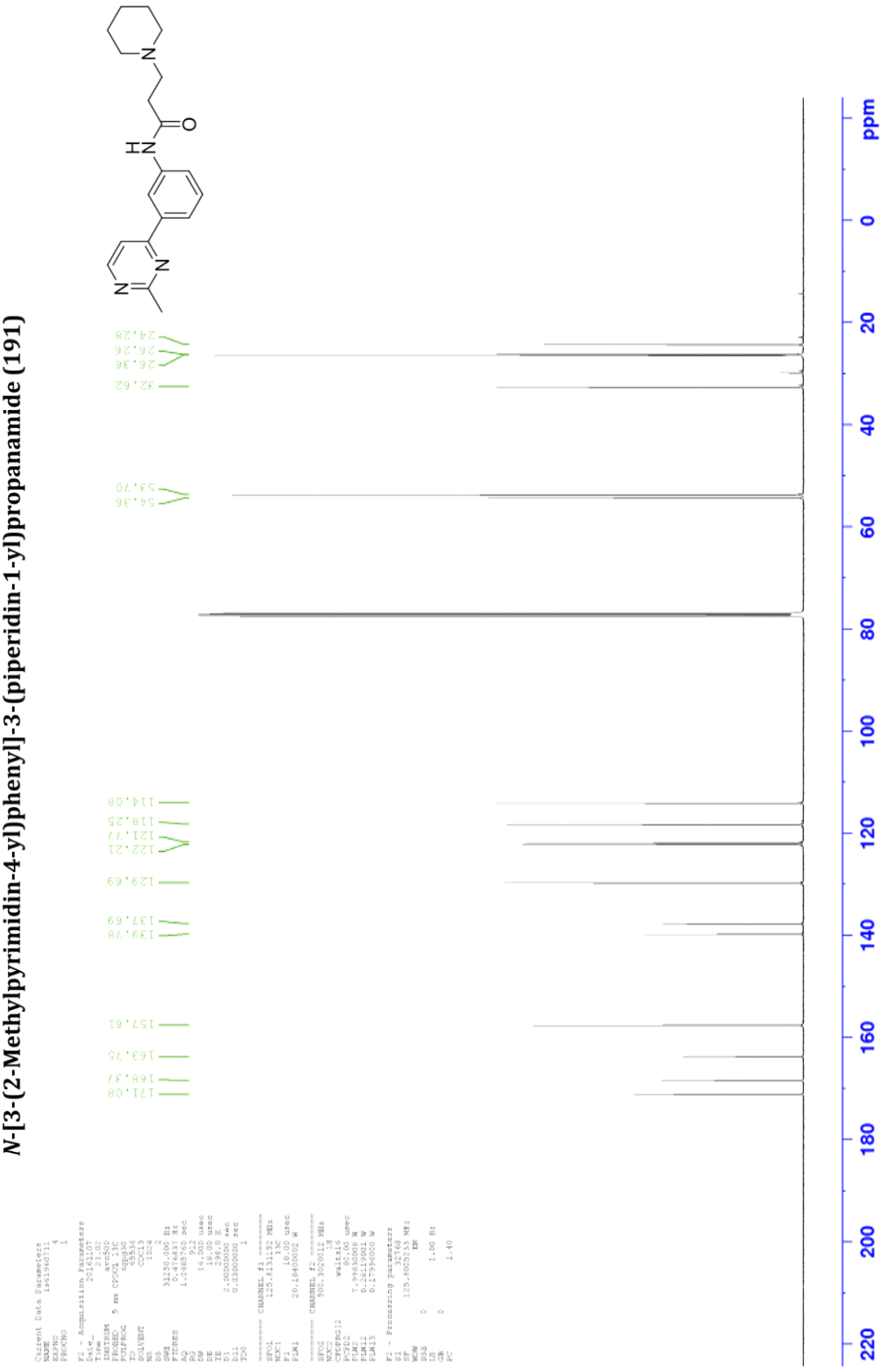
***N*-[3-(2-Methylpyrimidin-4-yl)phenyl]-2-(piperidin-1-yl)acetamide (190)**



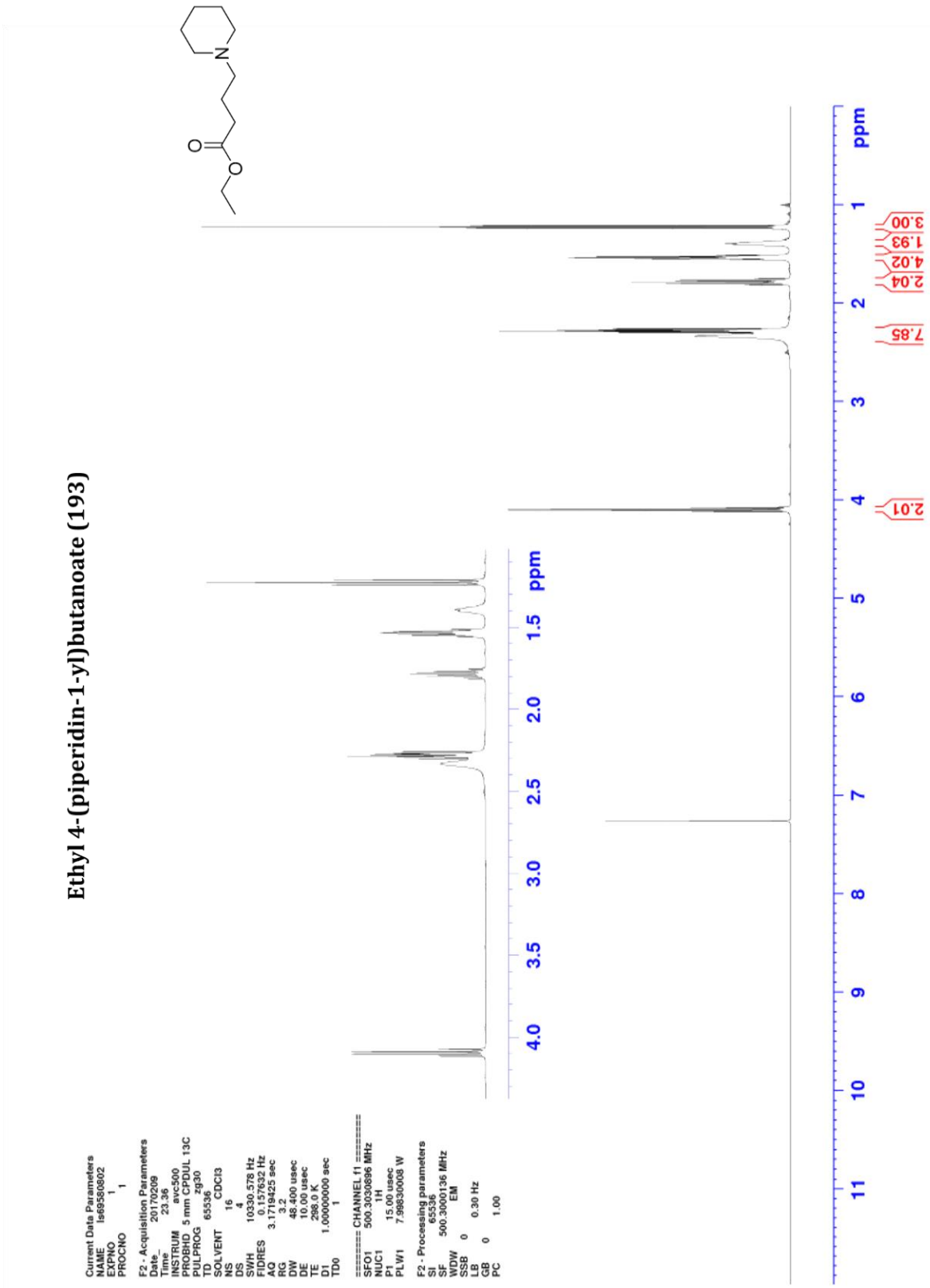
N-[3-(2-Methylpyrimidin-4-yl)phenyl]-3-(piperidin-1-yl)propanamide (191)



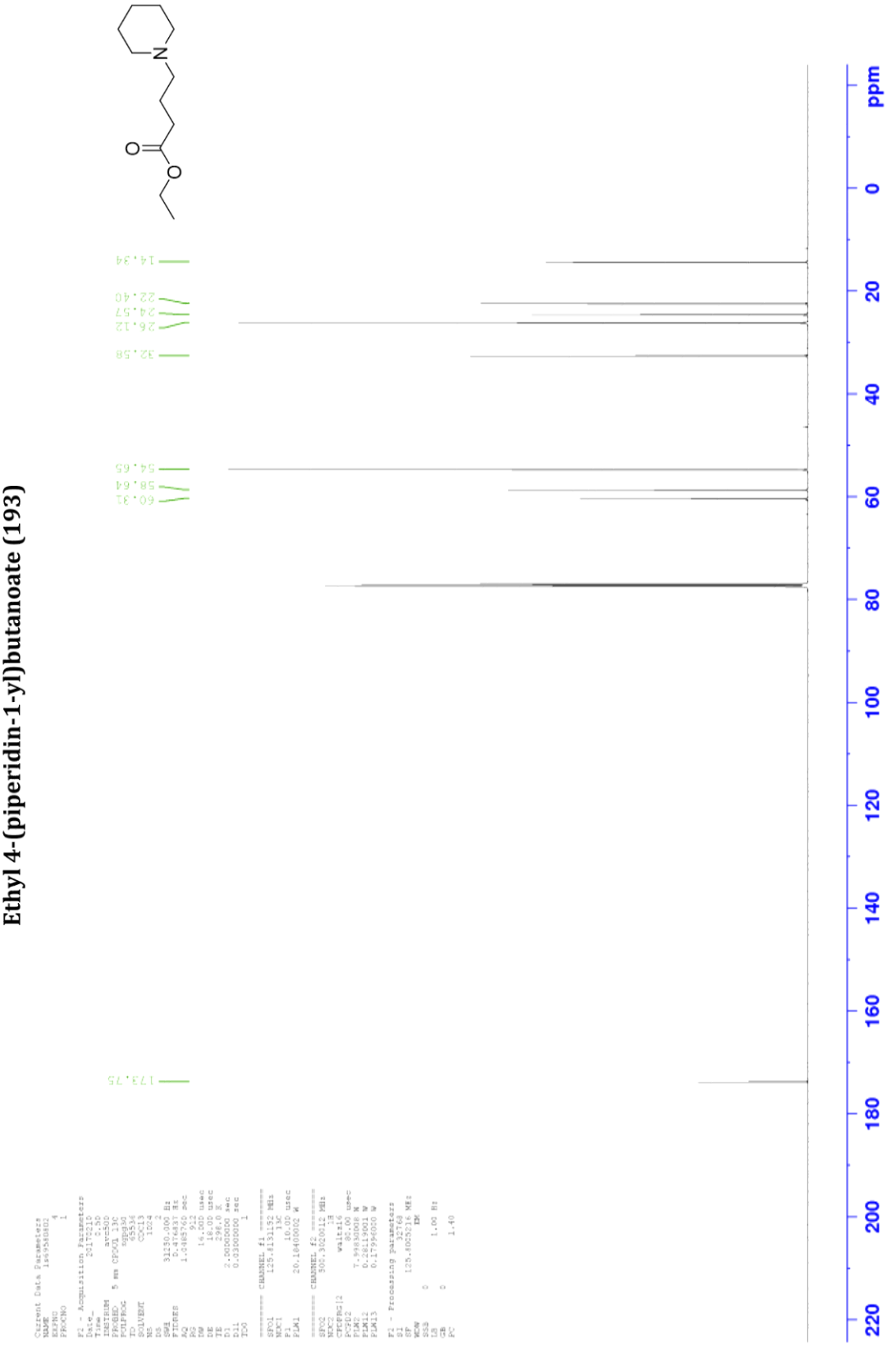
***N*-[3-(2-Methylpyrimidin-4-yl)phenyl]-3-(piperidin-1-yl)propanamide (191)**

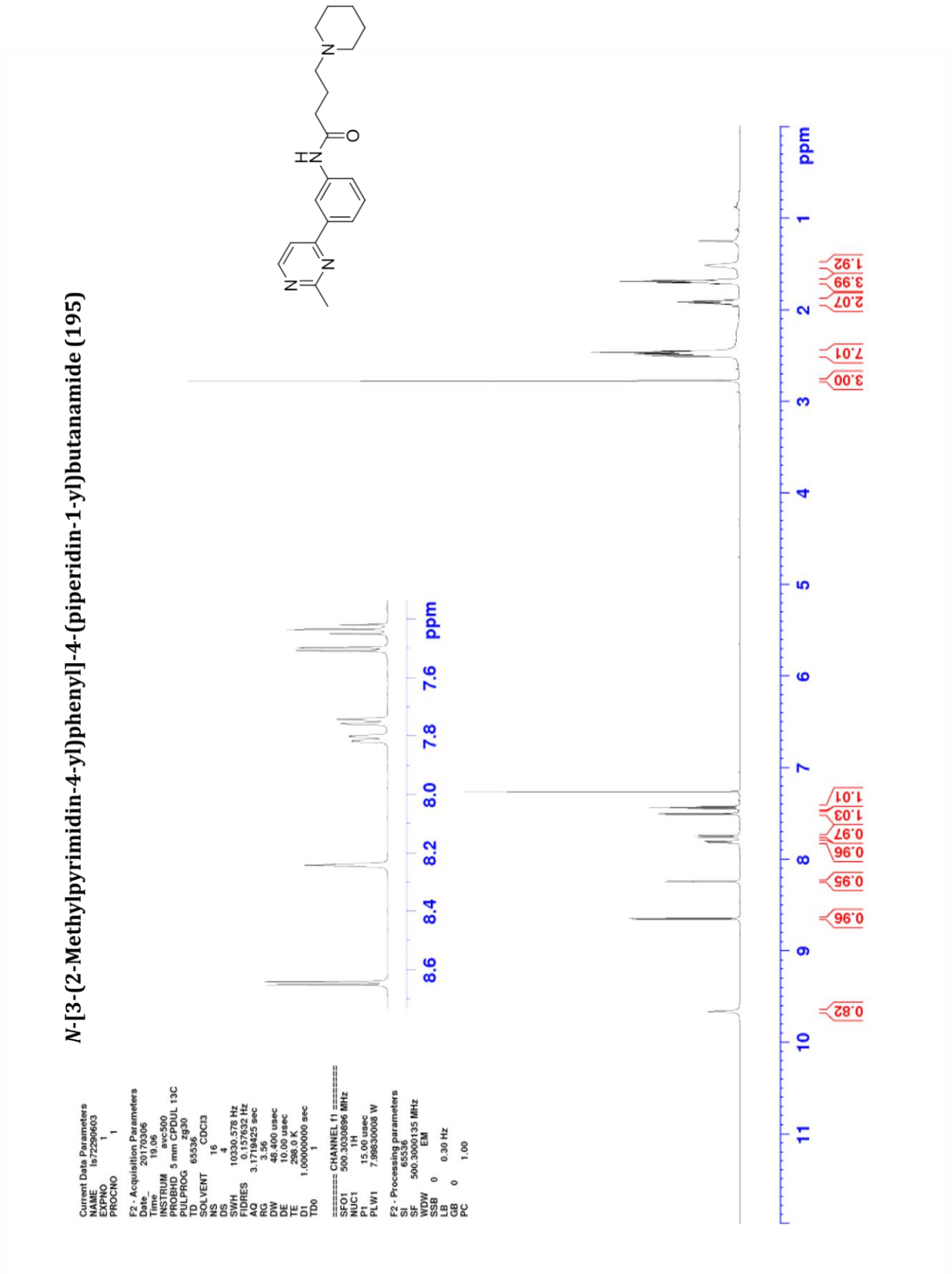


Ethyl 4-(piperidin-1-yl)butanoate (193)



Ethyl 4-(piperidin-1-yl)butanoate (193)




N-[3-(2-Methylpyrimidin-4-yl)phenyl]-4-(piperidin-1-yl)butanamide (195)
Current Data Parameters
NAME ls72290603
EXPNO 1
PROCNO 1
F2 - Acquisition Parameters
Date_ 20170306
Time 19.06
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17194425 sec
RG 3.56
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 15.00 usec
PLW1 7.99830008 W
F2 - Processing parameters
SI 65536
SF 500.3000135 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00
8.6
8.4
8.2
8.0
7.8
7.6
ppm
0.82
0.96
0.95
0.96
0.97
1.03
1.01
3.00
7.01
2.07
3.99
1.92
11
10
9
8
7
6
5
4
3
2
1
ppm

N-[3-(2-Methylpyrimidin-4-yl)phenyl]-4-(piperidin-1-yl)butanamide (195)

