Chemical Tools To Probe The Role Of Bromodomains In The Parasite *Trypanosoma cruzi*



Thesis submitted for the degree of:

Doctor of Philosophy in Organic Chemistry

CORENTINE MARIE CÉCILE LAURIN

St Hilda's College, University of Oxford

Trinity Term 2019

Under the supervision of:

Professor Stuart J. Conway (Oxford)

Dr Philip Humphreys (GSK)

Jon Seal (GSK)

BBSRC bioscience for the future

gsk

Department of Chemistry

DECLARATION OF AUTHORSHIP

Name: Corentine Marie Cécile Laurin Candidate number: 1002913

College: St Hilda's college Supervisor: Professor Stuart Conway

Title of thesis: Chemical Tools To Probe The Role Of Bromodomains In The Parasite *Trypanosoma cruzi*

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism.	☑
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 dissertation 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/dissertation/extended essay/assignment/project/other submitted work. (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: ……

Abstract

Chagas disease is a chronic infection caused by the parasite *Trypanosoma cruzi* (*T. cruzi*). The development of the parasite in the human heart muscle or digestive tract causes severe damage, resulting in organ failure. The current treatments, benznidazole and nifurtimox, have a limited efficacy, show significant side effects, and increasing resistance. Post-translational modifications (PTMs) play a crucial role in the life cycle of protozoan parasites. Bromodomains (BDs) are readers of lysine acetylation, one of the PTMs found on histone tails. The function of BDs remains uncharacterised in *T. cruzi*. The life cycle of the parasite is complex and is thought to rely heavily on epigenetic control and therefore our main interest is to validate the *T. cruzi* BDs as therapeutic targets.

Optimisation of expression and purification protocols of the protein BD-containing factors (BDFs) in *E. coli* (Chapter 1) was followed by the development of a waterLOGSY-based assay to screen acetyl-lysine mimic containing compounds (Figure 0.1). The binding of the hits was further characterised by ^{1}H NMR ligand-observed protein titration to obtain K_D. Excitingly, the human BD and kinase ligand, BI-2536 (**8**), was shown to bind to *Tc*BDF3 with a K_D value between 25 and 55 µM. The binding was confirmed to be 14.6±3.0 µM (n=3) using ITC. BI-2536 (**8**) is to date the highest affinity ligand identified and was chosen as a lead structure for further development (Chapter 2).

Based on this initial work, eight functional probes were designed and synthesised to enable the development of higher-throughput assays for *T. cruzi* BDs (Chapter 3). The ^{19}F NMR reporter molecules, photo-activatable diazirines, and fluorescent probes were assessed against *Tc*BDF3 to establish novel assays (Chapter 4). Firstly, the fluorinated probes were used as ^{19}F NMR reporter molecules to observe their binding and displacement to and from *Tc*BDF3. Secondly, a protein LCMS assay was established to characterise binding and subsequent cross-linking of the diazirine probes. Finally, fluorescent probes were validated by fluorescence polarisation. Satisfactorily, one probe binds to *Tc*BDF3, with a K_D value of 3 µM. Displacement of the fluorescent probe by the parent compound BI-2536 (**8**) results in a dose-dependent response, however, a full

curve and an IC_{50} value could not be obtained due to assay interference at high ligand concentrations. Further validation is required, using higher affinity ligands to show that the fluorescence polarisation displacement can generate IC_{50} values.

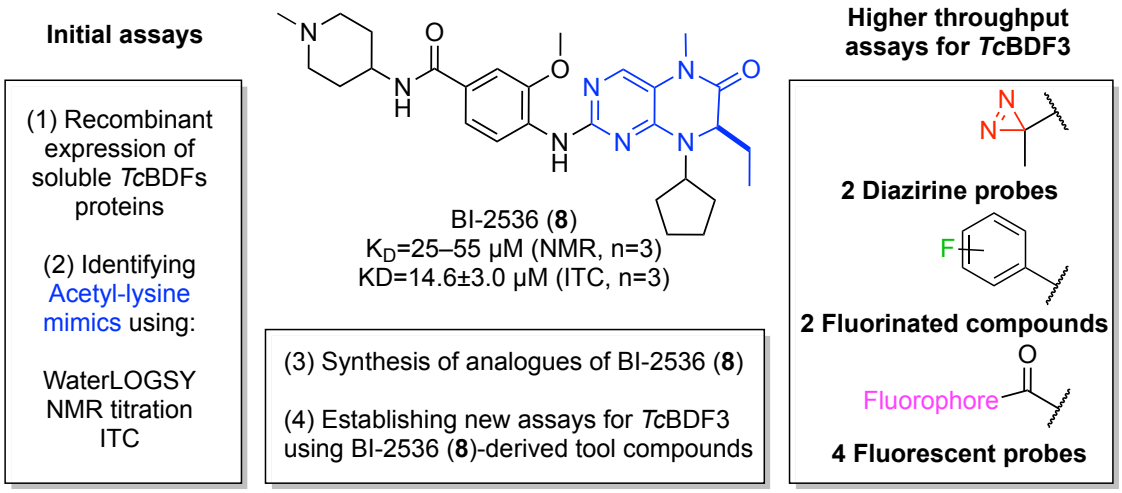


Figure 0.1: The development from initial into higher throughput assays through analogues of BI-2536 (**8**).

With GSK collaborators, 556 diazirines of the GSK PhaBits library were screened against *Tc*BDF3 and 5-1, identifying one and six new binding fragments, respectively. Additionally, a DNA-encoded library of *ca.* 1.9 trillion encodings were screened against *Tc*BDF2 and 5-1, giving 82 and 134 hit clusters, respectively, to be investigated for binding. The low hit rates for both screening campaigns highlights that the *T. cruzi* BDs are challenging targets.

In this work we report the first production of soluble *T. cruzi* BD proteins and their study using NMR assays such as waterLOGSY and NMR titration. Critically, this initial work led to the development of functional probes to provide a tool box to study *Tc*BDF3 at higher throughput with diverse and orthogonal assays. The novel biophysical assays established here can be used to screen efficiently for binding compounds and opens up the possibility of rapid SAR. Identified ligands will enable validation of the bromodomains of *T. cruzi* as targets for Chagas disease for the first time. This will have a wide range of implications in the use of epigenetics targets for treatments of parasites more broadly.

Table of contents

Abstract ..... **i**

Acknowledgments ..... **vii**

Abbreviations ..... **x**

Chapter 1. **Introduction** **1**

1.1. CHAGAS DISEASE 1
1.1.1. Chagas disease: patients, current treatment, and resistance *1*
1.1.1.1. Disease progression stages 2
1.1.1.2. Current treatments for Chagas disease 2
1.1.2. T. cruzi life cycle *3*
1.2. EPIGENETICS 5
1.2.1. Post-translational modifications (PTMs), and their associated reader, writer and eraser proteins. *5*
1.2.2. Bromodomains: reader domains for the acetylation of lysine residues *6*
1.2.2.1. BET BDs 8
1.2.2.2. Non-BET BDs 11
1.2.3. BD protein recombinant expression and biophysical assays *12*
1.2.3.1. Recombinant expression 12
1.2.3.2. Thermal Shift 13
1.2.3.3. WaterLOGSY 14
1.2.3.4. Fluorescence polarisation 15
1.2.3.5. Isothermal titration calorimetry 16
1.2.4. BDs as epigenetics targets for drug development *17*
1.3. EPIGENETICS OF *T. CRUZI* 18
1.3.1. Known PTMs in T. cruzi *18*
1.3.2. T. cruzi BD protein functions: TcBDFs *19*
1.3.3. T. cruzi HDACs and HATS *21*
1.3.4. Epigenetic targets as a strategy against protozoan parasites *21*
1.3.4.1. Susceptibility of *T. cruzi* to OXFBD02 (**7**) 21
1.3.4.2. Non KAc-mimicking *Tc*BDF3 binders 22
1.3.4.3. I-BET151 (**5**) inhibition of the *T. brucei* BD *Tb*BD2 23
1.4. PROJECT AIMS 24

Chapter 2. **Protein expression and structure predictions** **26**

2.1. INTRODUCTION AND AIMS 26
2.2. PROTEIN EXPRESSION 27
2.2.1. H_6-TcBDF2 expression *28*
2.2.2. H_6-TcBDF5 expression optimisation and large scale expression *30*
2.2.2.1. Small scale expression trials for H_6-*Tc*BDF5-1 and H_6-*Tc*BDF5-2 30
2.2.2.2. Large scale expression of *Tc*BDF5-1 31
2.2.3. H_6-TcBDF3 expression *33*
2.2.3.1. H_6-*Tc*BDF3(32..148aa) expression and purification optimisation 33
2.2.3.2. His-tag cleavage with recombinant TEV protease 37
2.2.3.3. Crystallisation trials for *Tc*BDF3 42
2.2.4. N to F TcBDFs mutants *45*
2.2.5. DNA assembly for the TcBDF2, 3 and 5-1 N to F mutants *46*
2.2.5.1. Expression trials for the *Tc*BDF2, 3 and 5-1 N to F mutants 46
2.2.6. Protein expression, purification, and crystallisation summary *48*
2.3. PROTEIN STRUCTURE PREDICTIONS 49
2.3.1. Protein folding prediction and circular dichroism *49*
2.3.2. Homology models *50*
2.4. SUMMARY AND CONCLUSIONS 52

Chapter 3. **Initial hits and assays** **53**

3.1. INTRODUCTION AND AIMS 53
3.2. SYNTHESIS OF LIGANDS FOR SCREENING AGAINST THE *Tc*BDF BDS 54
3.2.1. Synthesis of known acetyl lysine mimics *55*
3.2.1.1. **FR1:** Benzimidazolones fragments 55
3.2.1.2. **FR2**: Synthesis of dimethylisoxazole ligands 61
3.2.1.3. **FR3:** Pyridazolone synthesis 64
3.2.1.4. **FR4**: Synthesis of the I-BET726 core 64
3.2.1.5. **FR5:** Dihydrobenzoxazepine core synthesis 66
3.2.1.6. **FR6**: Synthesis of BI-2536 (**8**) acetyl-lysine mimic 67
3.2.2. Synthesis conclusions *70*
3.3. SCREENING OF FRAGMENT AND KNOWN ACETYL LYSINE MIMIC SCREEN 71
3.3.1. Screening aims *71*
3.3.2. Thermal shift *71*
3.3.3. WaterLOGSY *73*
3.3.3.1. **FR1** Benzimidazolone derivatives against TcBDF3 75
3.3.3.2. **FR3, FR4** and **FR5** do not bind to *Tc*BDF3 79
3.3.3.3. The dihydropteridineone **FR6** fragments binds to *Tc*BDF3 by waterLOGSY 80
3.3.3.4. ATAD2 binders against TcBDF2 81
3.3.4. ^{1}H ligand observed protein titration *82*
3.3.4.1. Establishing ^{1}H ligand observed protein titration for *Tc*BDF3 82
3.3.4.2. Binding affinities of I-BET151 (**5**), BI-2536 (**8**) and OXFBD02 (**7**) to *Tc*BDF3 by NMR titration 83
*3.3.5. Isothermal calorimetry of the TcBDF3/ BI-2536 (**8**) interaction* *86*
3.3.5.1. The binding of BI-2536 (**8**) to *Tc*BDF3 by ITC 86
3.3.5.2. The binding of BI-6727 (**9**) to *Tc*BDF3 by ITC 87
*3.3.6. Investigation of the binding of the N-methylated dihydropteridinone core of BI-2536 (**8**), **(R)-92**, to TcBDF3* *88*
3.4. DNA-ENCODED LIBRARY 89
3.5. SUMMARY AND CONCLUSIONS 91

Chapter 4. **Design and synthesis of tool compounds** **92**

4.1. INTRODUCTION AND AIMS 92
4.2. PREDICTION OF THE MODE OF BINDING OF BI-2536 (**8**) 94
*4.2.1. Docking of BI-2536 (**8**) into the TcBDF3 homology model* *94*
4.2.2. Potential point of modifications *95*
4.3. DESIGN OF TOOL COMPOUNDS 96
4.3.1. Target ^{19}F NMR probes *96*
4.3.2. Target diazirine molecules *98*
4.3.3. Target fluorescent probes *100*
4.4. SYNTHESIS OF COMMON INTERMEDIATES FOR THE TOOL COMPOUNDS 103
*4.4.1. Intermediate **A**: Synthesis of the chloropyrimidine fragment **(R)-92*** *105*
*4.4.2. Intermediate **B**: the aminopyrimidine intermediate **(R)-138*** *105*
*4.4.3. Intermediate **C**: the amino-piperidine intermediate **(R)-139*** *108*
4.4.3.1. First approach: convergent synthesis of Intermediate **C**: **(*R*)-139** (path A) 108
4.4.3.2. Second approach: the sequential synthesis of Intermediate **C**, **(*R*)-139** (path B) 109
4.5. SYNTHESIS OF ^{19}F NMR TOOL COMPOUNDS 111
4.6. SYNTHESIS OF DIAZIRINES TOOL COMPOUNDS 112
4.7. SYNTHESIS OF THE FLUORESCENT TOOL COMPOUNDS 113
*4.7.1. Synthesis **(R)-134** and **(R)-135** by amide coupling* *113*
*4.7.2. Synthesis of **(R)-136** and **(R)-137*** *114*
4.8. SYNTHESIS OF CONTROL COMPOUNDS AND THEIR VALIDATION BY ITC 117
*4.8.1. N-methylated dihydropteridinone control compounds **(R)-155**–**(R)-157*** *117*
*4.8.2. Literature control compounds **15** and **16*** *118*
*4.8.3. Validation of the BI-2536 (**8**) analogues and control compounds by ITC* *119*
4.9. SUMMARY AND CONCLUSIONS 121

Chapter 5. **New biophysical assays using the *Tc*BDF3 tool compounds** **123**

5.1. INTRODUCTION AND AIMS 123

5.2. ^{19}F NMR ASSAYS 123
5.2.1. ^{19}F NMR probes for TcBDF3 ..... *124*
5.2.1.1. The binding of **112**, **(*R*)-130**, and **(*R*)-131** to *Tc*BDF3 124
5.2.1.2. Displacement of ^{19}F NMR reporter molecules to detect *Tc*BDF3 binding 127
5.2.2. Fluorine probes for bromodomains from other parasites *129*
5.3. PHOTO-CROSSLINKING ASSAYS WITH DIAZIRINES PROBES 131
5.3.1. Diazirine tool compounds *132*
5.3.1.1. Cross-linking to *Tc*BDF3 133
5.3.1.2. Cross-linking to *Tc*BDF5-1 136
5.3.2. Screening the GSK PhaBits library: a diazirine fragment library to identify new T. cruzi bromodomain binders *138*
5.3.2.1. PhaBits screen on *Tc*BDF3 139
5.3.2.2. PhaBits screen on *Tc*BDF5-1 143
5.4. FLUORESCENCE POLARISATION ASSAY 146
5.4.1. Characterisation of the FP probes fluorescence intensity and fluorescence polarisation 147
5.4.2. Binding affinities of the FP probes to TcBDF3 and other BD proteins *149*
5.4.2.1. The binding of FP probes to *Tc*BDF3 149
5.4.2.2. The binding of the FP probes to BRD4(1) and BRD4(2) 150
5.4.2.3. The binding of **(*R*)-135** to PLK1 151
5.4.2.4. **(*R*)-135** as a fluorescence probe for *Tc*BDF3 151
5.4.3. Displacement assay set up for TcBDF3 ..... *152*
5.4.3.1. Optimisation using the positive control BI-2536 (**8**) 152
5.4.3.2. Investigation of the displacement of **(*R*)-135** by other ligands for *Tc*BDF3 154
5.4.4. Displacement assay set up for BRD4(1), BRD4(2) and PLK1 ..... *158*
*5.4.5. Using the **(R)-135** for the study of BD proteins from other parasites* *159*
5.4.5.1. FP assays for the *Schistosoma manson*i parasite BD proteins 159
5.4.5.2. FP assays for the *Leishmania donovani* parasite BD proteins 162
5.5. CONCLUSIONS 164

Chapter 6. **Conclusions and future work** **165**

Chapter 7. **Experimental** **168**

7.1. SYNTHETIC METHOD AND MATERIALS 168
7.2. SYNTHESIS 174
7.3. BIOLOGICAL METHOD AND MATERIALS 263
7.3.1. DNA cloning *264*
7.3.1.1. DNA amplification – polymerase chain reaction (PCR) 264
7.3.1.2. DNA purification – agarose gel electrophoresis 264
7.3.2. Protein expression and purification *266*
7.3.2.1. Bacterial starter cultures 267
7.3.2.2. Large-scale culture and protein expression 267
7.3.2.3. Small-scale protein expression trials 267
7.3.2.4. Protein purification 268
7.3.2.5. SDS-PAGE 270
7.3.2.6. Western blotting 271
7.3.2.7. Circular dichroism (CD) 271
7.3.3. Protein crystallisation screens *272*
7.3.3.1. TEV cleavage 272
7.3.3.2. Protein crystallisation trials 273
7.3.4. Biophysical assays protocols *274*
7.3.4.1. Differential scanning fluorimetry 274
7.3.4.2. Water-Ligand Observed *via* Gradient Spectroscopy (WaterLOGSY) 274
7.3.4.3. Ligand-observed proton NMR 275
7.3.4.4. Binding assay based on ^{19}F NMR 276
7.3.4.5. Isothermal titration calorimetry (ITC) 276
7.3.4.6. Fluorescence polarisation (FP) 277
7.3.4.7. Photo-crosslinking mass spectrometry assay 279

References: **281**

Appendices **292**

APPENDIX A. PLASMIDS AND PROTEIN CONSTRUCTS 293

A.1. Plasmid maps 293
A.2. Amino acid sequences 299
A.3. Cloning 303
A.4. Protein expression 309

APPENDIX B. BIOPHYSICAL ASSAYS 319

B.1. Thermal shift 319
B.2. WaterLOGSY spectra 325
B.3. Ligand-observed proton NMR titration 369
B.4. Isothermal calorimetry titration curves 384
B.5. ^{19}F NMR assays 395
B.6. Fluorescence polarisation 397
B.7. PhaBits LCMS spectra examples 402

APPENDIX C. BENZIMIDAZOLONE RING FORMATION MECHANISM 404

APPENDIX D. CRYSTALLISATION TRIALS BUFFERS 405

D.1. Compositions of the commercially available broad screen for crystallisation 405
D.2. Optimisation tray for E5 condition from PEG ION broad screen 419
D.3. Optimisation tray for E8 condition from Crystal screen broad screen 420

SPECTRA FOR NOVEL COMPOUNDS 421

Acknowledgments

Firstly, I am sincerely thankful for my academic supervisor, Prof. Stuart Conway. Thank you Stuart for having me in the group, trusting my judgement, giving advice and feedback when needed. Thank you also for the group days out, Christmas dinner and dessert wine that have glued the group together into a team. I guess you were right, if research was easy, it would have been done before. And yes, in retrospect, I would choose to be supervised by you again.

Thank you to my GSK supervisors Dr Philip Humphreys and Jon Seal. I am grateful to you both for the warm welcome and help during my time at GSK. Thank you to Emma and Jake for the PhaBits collaboration, and Felix and Raquel for including me in the BDs in parasite network. Emma, many thanks for all the help with all the photo-crosslinking experiments. A huge thanks to Gang Yao, Heather O'Keefe and Chris Phelps for screening a whole DNA-encoded library against two of my targets.

I am extremely grateful to my proof-readers. Thank you to Sarah, Amélie, Matthias, Solange, Naomi and Ben for catching my mistakes and always having my back.

I've been privileged to work alongside some amazing humans during my time in the Conway group. First, to the wise post-docs: Ewen, Anthony, Maze, James, Debs and Pietro. Thank you for everything you taught me, being my out-of-hours-lab buddies and being always ready for a pint. Ewen, thank you for everything from being the tough Glasgow demonstrator, the patient lab buddy, counsellor and the enthusiastic drinker! Anthony, thank you for teaching me everything protein and plasmid related, I couldn't have done it without you! Maze, I've been incredibly lucky to have met you and learn from you how to design new assays. I miss our trouble shooting discussions! James and Pietro thank you for never letting the lab feeling sad or empty (nor quiet)! Debs, you have been the best - thank you for you undying enthusiasm and support in tough times.

To Catherine and Larissa, I have started this journey with you and I'm so glad I had you both to share the ups and downs. Catherine, thank you for all the shared gin, the corridor dances and always (!) being there. Larissa, I'll always remember the Milo-Tea-Biscuits venting sessions – our love of food got us through it all!

Amy, Mus, and Joe, I don't think I could have survived the PhD without you. The highlight of our Final year will always be the cocktails by the pool, the lazy river and the crazy golf. Thank you Mus for being my official email proof-reader (+ a fierce sticker war opponent). Thank you Amy for the G11 banter (I LOVE CHEMISTRY/not-my-problem). Thank you Joe for the coffee breaks, runs and walks in the park to stay as "sane" as possible. Also thanks for being my *T. cruzi* buddy, with all the MD, homology and misery.



Figure 1: ACS Orlando 2019

A big thank you to Sarah and Amelie who both taught me so much when I joined the group, you have been amazing desk/fumehood buddies. To Jess and Holly, G11 would not have been the same without you – It's been a pleasure to work with you guys. Adam, keep your un-extinguishable enthusiasm and the dance moves!

We have had many fantastic part 2 students over the years. To cite a few: Fern (whom I started it all with), Leila and Michelle (it's been a pleasure to supervise you in the lab - I'm proud of how well you both did!). Thank you also to Katie, Jessie, Jesse, Thanh, Marguarida, Kathrin, Elliot (smash), James, Oli... And a special thanks to Andrew who did some great modelling-docking-MD on *Tc*BDF3!

To the lower-ground-crew, thank you for having made me a honorary member of the CJS group. Thank you to Jos, Patrick, Jurgen, Thomas, Aiste, Anka, Leah and for all the help

with protein purification and not-so-lucky crystals adventures. Special thanks to Dr Mike McDonough and Dr Adam Hardy for the help with crystallisation and equipment.

I'm very grateful to everyone in the NMR and MS services, and especially to Prof. Tim Claridge and Dr Nader Amin for the constant help with NMR related assays. Thanks also to everyone in stores and the workshop for running everything smoothly.

To Dr Marcus Lee (Sanger; Plasmodium testing) and Dr Nathaniel Jones (University of York; Leishmania BD proteins) thank you for the fruitful collaborations. I hope to be lucky enough to have collaborators like you again in the future!

Many thanks to Naomi, Si, Abby, Nat, and Sarah for all the dancing. Thank you to the Hildabeasts: Miranda, Irene, Hannah (for café baba) and Kristina, Susan and Mpho (for soldiering on and using our meal credits). Thank you also to Henry, for the support and the help moving out of Oxford! A big thank you to Dr Alina Tirla, Dr Ben Stenton and the rest of the Dorfstasse team for being one of the reason to do such a thing as a PhD. Atea you have the fire it takes to do anything you decide, and I'm so lucky to be your friend.

Ben, you've been the best. Thank you for the ever-lasting support and all the happy unforgettable moments; from late take-aways to Michelin-star restaurants. I've been more than lucky to have found such an amazing person to get through all this with.

A special thanks to my grandparents, Panou et Manou, that have always believed in me and never doubted I could follow the steps of my distant illustrious relative Dr Antoine Germain Labarraque. Last but not least, I am extremely grateful to my parents and siblings who have offered constant support, have listened to my speeches on epigenetics and cut out science articles from newspapers for me. Merci du fond du coeur

Abbreviations

° C	degree Celsius
2 D	two-dimensional
4-DMAP	4-dimethylaminopyridine
Å	Angstrom
ß-ME	ß-mercaptoethanol
Ac	acetyl
AlphaScreen™	Amplified luminescent proximity homogeneous assay screen
ATAD2	ATPase AAA domain-containing protein 2
aq.	aqueous
BET	bromodomain and extra terminal
BLAST	basic local alignment search tool
BCP	bromodomain containing protein
BD	bromodomain
Boc	*tert*-butoxycarbonyl
bp	base pair
BRPF1	bromodomain and PHD finger-containing
BSA	Bovine serum albumin
conc.	concentrated
CDI	1,1'-carbonyldiimine
cAMP	cyclic adenosine monophosphate
CHAPS	(3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate)
COSY	correlation spectroscopy
CREBBP	cAMP response element-binding protein binding protein
CPMG	Carr-Purcell-Meiboom-Gill
CV	column volume
d	day
DIPEA	*N,N'*-diisopropylethylamine
DEPT	distortionless enhancement by polarisation transfer

DMI	dimethylisoxazole
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DSF	differential scanning fluorimetry
ee	enantiomeric excess
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ELT	encoded library technology
eq	equivalent
Et	ethyl
FBDD	fragment-based drug discovery
Fmoc	fluorenylmethyloxycarbonyl
FP	fluorescence polarisation
FPLC	ÄKTA fast protein liquid chromatography
g	gram(s)
h	hour(s)
H_6	hexahistidine tag
HAT	histone acetyl transferase
HATU	1-[bis(dimethylamino)methylene]-*1H*-1,2,3-triazolo[4,5-*b*] pyridinium 3-oxide hexafluorophosphate
HDAC	histones deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMBC	heteronuclear multiple-bond correlation spectroscopy
HPLC	high-performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single-quantum correlation spectroscopy
Hz	Hertz
IC_{50}	half maximal inhibitory concentration
IMAC	immobilized metal affinity chromatography
IPA	isopropanol

IPTG	isopropyl β-d-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
IR	infrared
J	coupling constant
K_D	dissociation constant
L	litre
LE	ligand efficiency
LCMS	liquid chromatography mass spectrometry
lit.	literature
LRMS	low resolution mass spectrometry
M	molar
Me	methyl
MeCN	acetonitrile
MHz	megaHertz
MST	MicroScale Electrophoresis
min	minute(s)
mL	millilitre(s)
mol	mole
m.p.	melting point
MS	mass spectrometry
NMR	nuclear magnetic resonance
nOe	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
O/N	overnight
OD_{600}	optical density at 600 nm
P300/CBP	cAMP response element-binding (CREB) binding protein
PBS	phosphate-buffered saline
PCAF	P300/CBP-associated factor
PCR	polymerase chain reaction
PDB	protein data bank
Ph	phenyl

pH	potential of hydrogen
PLK1	Polo-like kinase 1
PPI	protein-protein interaction
ppm	parts per million
PTM	post-translational modification
R_f	retention factor
RT	room temperature
s	second(s)
sat.	saturated
SAR	structure-activity relationship
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SFI	solubility forecast index
SD	standard deviation
TBS	tris-buffered saline
TBTA	Tris[(1-benzyl-*1H*-1,2,3-triazol-4-yl)methyl]amine
T. cruzi	*Trypanosoma cruzi*
*Tc*BDF	*Trypanosoma cruzi* bromodomain factor
TEMED	Tetramethylethylenediamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
Tm	protein melting temperature
ΔTm	thermal shift
TMS	tetramethylsilane
UV	ultraviolet
v/v	volume per unit volume (1% = 1 mL/100 mL)
w/v	weight per unit volume (1% = 1 g/100 mL)
waterLOGSY	water-ligand observed via gradient spectroscopy
WT	wild-type
XL_{50}	half maximal cross-linking concentration

Chapter 1. Introduction

1.1. Chagas disease

Chagas disease is a chronic parasitic infection caused by the *Trypanosoma cruzi* (*T. cruzi*) parasite, and is described as a tropical neglected disease.[1] This parasite is transmitted to mammals by hematophagous triatominae insects. Carlos Chagas (1879-1934) first described the parasite transmission cycle and its symptoms in 1909.[2] Despite having been described 110 years ago, the diagnosis and treatment of Chagas disease are still a major challenge.[3]

1.1.1. Chagas disease: patients, current treatment, and resistance

Chagas disease affects 6–7 million people worldwide and kills around 10 thousand per year.[3] It is endemic to South America, but is found worldwide due to population migration. The disease spreads through the bites of triatominae insects, known as kissing bugs, that are infected by the *T. cruzi* parasite. The parasite develops in the insect's digestive tract and is transmitted through its faeces. Transmission can also occur congenitally, through consumption of contaminated food, blood transfusion or organ transplantation.[1] Additionally, there have been reports of sexual transmissions.[4]

Wild animals, including possums and armadillos, are the main reservoirs for this parasite, rendering the eradication of the disease impossible.[5] However, prevention of further spread of the disease through the use of insecticides, and bed nets, is effective.[1] Additionally, the improvement of housing quality and food storage also impacts disease infection rates.[1]

One major challenge is the effective diagnosis of the disease due to the complex biology of the parasite. Additionally, many patients are infected as children, and if the initial infection phase went unnoticed or was misdiagnosed, then they are not aware of their condition as teenagers or adults.

1.1.1.1. Disease progression stages

The initial infection phase in human patients lasts 1-2 months and the chronic phase persists for the patient's lifespan.[1] The acute phase is commonly asymptotic or appears as a self-limiting febrile illness; the symptoms resolve spontaneously in 90% of the infected patients. Those symptoms include fever, malaise, enlargement of the spleen/liver/lymph nodes, and inflamed sign of entry around the insect vector's bite. Although 60% of the patients never develop clinically apparent disease, the remaining 40% develop a determinate form of chronic disease (cardiac, digestive or cardiodigestive) 10 to 30 years after infection.[1] The progressive development of the parasite in the heart muscle in chronic Chagas disease leads to myocardiopathy, that is the impairment of the contractile and dilation function of the heart (Figure 1.1).[1,6] If untreated, this condition leads to heart failure.[6]

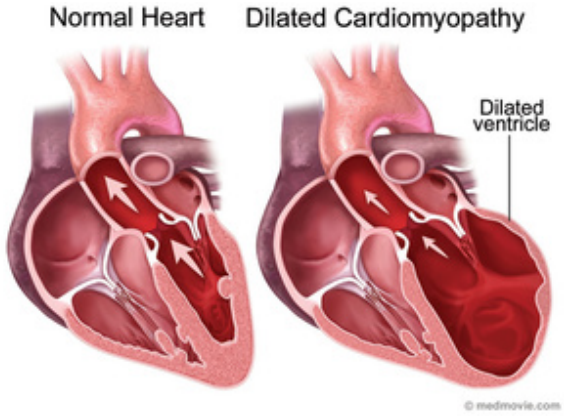
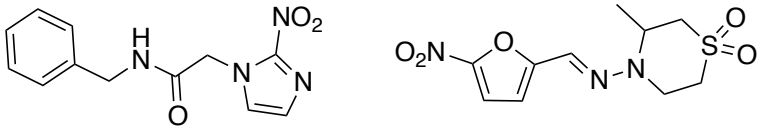


Figure 1.1: Schematic representation of the enlargement of the heart muscle, characteristic of cardiomyopathy.[7]

1.1.1.2. Current treatments for Chagas disease

Only two drugs are currently available for the treatment of Chagas Disease: benznidazole (**1**) and nifurtimox (**2**) (Figure 1.2).[8] Both drugs contain a nitro group, which is thought to generate free radicals and/or reactive electrophiles after metabolism by nitroreductases in the parasite. Unlike humans, the *T. cruzi* parasite has very few enzymatic defences against free radicals rendering it more susceptible to damage.[9] The use of benznidazole and nifurtumox is effective during the acute phase of this disease, if given soon after the

onset of the infection. However, the treatment efficacy diminishes with disease progression, and neither drugs are curative during the chronic phase.[1]



Benznidazole (**1**) Nifurtimox (**2**)

Figure 1.2: The structures of the treatments for Chagas disease: benznidazole (**1**) and nifurtimox (**2**), first registered in the 1960s (Bayer) and 1970s (Hoffmann-la Roche), respectively.[8-9]

Benznidazole (**1**) and nifurtimox (**2**) have significant side effects, with adverse effects occurring in around 40% of treated patients. These effects are a consequence of the mechanism of actions of these drugs, as the generation of free radical and metabolic electrophiles is responsible for a high systemic toxicity.[9] Benznidazole (**1**) and nifurtimox (**2**) cannot be taken by pregnant women or people with liver or kidney failure. Furthermore, nifurtimox (**2**) is contraindicated for patients suffering from neurological or psychiatric disorders. Moreover, resistance to the treatments has first been recorded in 1972, with different *T. cruzi* strains display varying levels of resistance.[10-11]

Despite the scientific community's efforts to harness the differences in metabolism, biosynthesis pathways and enzymatic functions between humans and *T. cruzi*,[9] no new trypanocidal treatments have been approved over the past 40 years.[3]

1.1.2. *T. cruzi* life cycle

T. cruzi is a kinetoplastid protozoan parasite, a flagellated single cell eukaryotic organism, that goes through at least four distinct development stages (Figure 1.3).[12-13] This class of organisms also includes the *Trypanosoma brucei* and *Leishmania* parasites. Together they are referred to as the "TriTryps".

In triatominae insects, the parasite differentiates into epimastigotes in the environment of the mid-gut. This life cycle stage is replicative and proliferative. The epimastigostes anchor themselves to the insect gut wall through their flagellum, and transition to their next life cycle stage: metacyclic trypomastigotes.[13] Once formed, these can detach

themselves from the insect gut wall and, through contamination *via* the insect bite wound, infect mammalian cells.

In mammalian hosts, infective *T. cruzi* metacyclic trypomastigotes differentiate into amastigotes intracellularly. This replicative form eventually converts into trypomastigotes, bursting the host cell. Trypomastigotes, released in the bloodstream, infect new host cells, or are transmitted to hematophagous triatominae insects.[14-15]

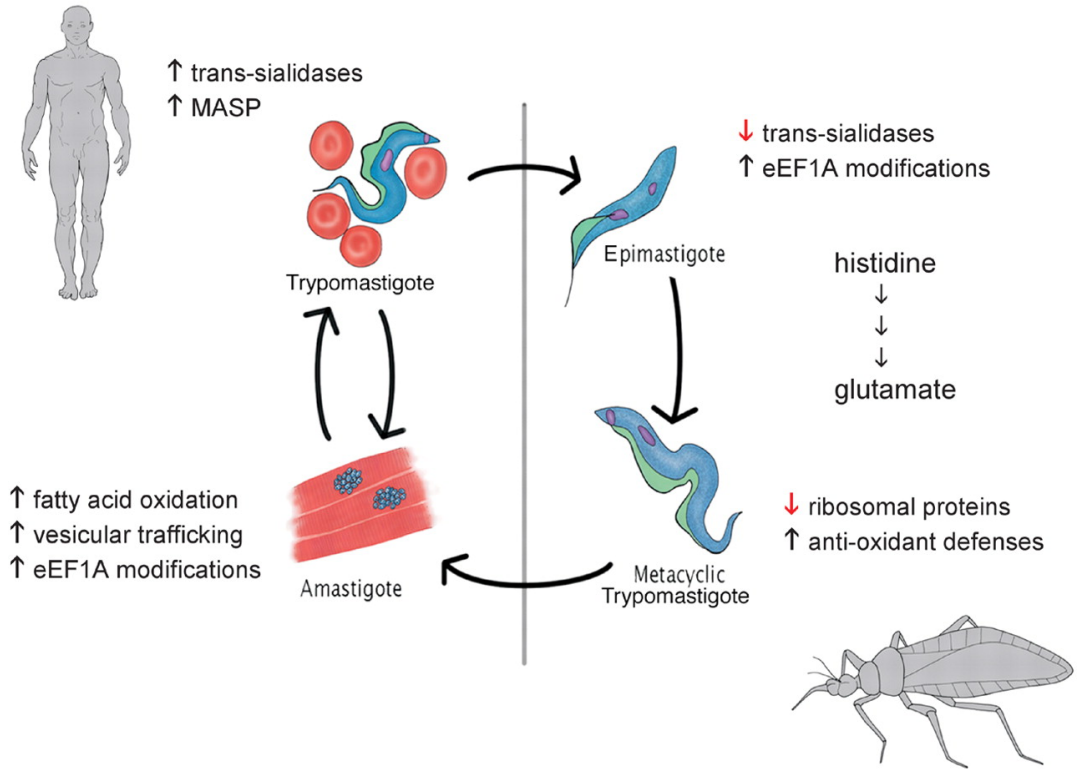


Figure 1.3: Life cycle of *T. cruzi.* In mammalian hosts: *T. cruzi* trypomastigotes circulate in the blood of infected hosts, enter host cells and convert to amastigote forms to replicate. Triatomine bug vectors become infected by ingesting trypomastigotes during the course of a blood meal on infected mammalian hosts. In the insect gut, trypomastigotes convert into epimastigotes, which differentiates into metacyclic trypomastigotes. These initiate new infection in mammals contact with the infected feces. (Reprinted with permission from Atwood, J. A.; *et al*, [15] *Science* **2005,** *309*, 473-476; Copyright 2005 The American Association for the Advancement of Science.)

Although the four distinct development forms of *T. cruzi* share the same DNA, they have distinct phenotypes, including morphology, functions, and metabolism.[16] Switching between distinct phenotypes is a therefore a consequence of differences in gene expression rather than changes of the genomic information. The change of gene expression, that does not arise from changes of the DNA, is controlled, at least in parts, by epigenetic mechanisms.

1.2. Epigenetics

1.2.1. Post-translational modifications (PTMs), and their associated reader, writer and eraser proteins

Epigenetics is the study of the heritable changes in phenotype, which do not result from changes in DNA sequence.[17-18] Changes in structures of proteins, or DNA, that occur after their biosynthesis are referred to as post-translational modifications (PTMs). Epigenetic PTMS include methylation, phosphorylation, lipidation, sumoylation, and acetylation of amino acid residues, or DNA base pair methylation.[19] These modifications are introduced and removed enzymatically by writers and erasers, and recognised by reader proteins.[19]

Histones are proteins in eukaryotic cells which package and order DNA into chromosomes. The histone tails protrude out of the histone/DNA complex and can be modified by acetylation, methylation, phosphorylation and ubiquitination of specific amino-acid. PTMs on histones tails influence how histones and DNA are packed, through tightening or loosening of the nucleosomes thus affecting the accessibility to the transcription factors (Figure 1.4). The acetylation of lysine (KAc) on histones tails is a key PTM,[20-21] which affects gene expression in a wide range of species.

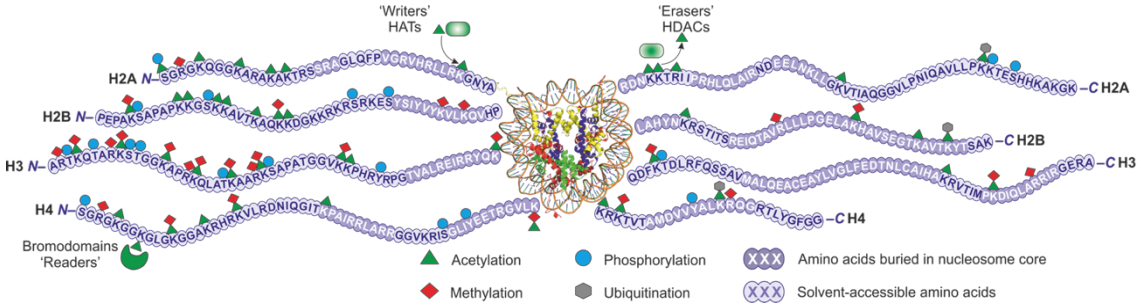


Figure 1.4: PTMs on histone tails are introduced, recognised and removed by "writers", "readers" and "erasers", respectively. Reprinted from Jennings, L. E *et al.*[22] *Future Med. Chem.* **2014,** *6*, 179–204. Copyright 2014 Future Science Group.

PTMs play a central role in gene expression. In particular, the differential expression of genes leads to varying phenotypes from a single genome, which allows cell differentiation and specialisation. The control of these processes by combinatorial

display of simultaneous PTMs on histone tails is referred to as the "histone code".[23] This can be illustrated by the fact that a caterpillar and the resulting butterfly share the same DNA, but their gene expression varies leading to two distinct phenotypes.

The histone acetyl-transferases (HATs) and the histone deacetylases (HDACs) are responsible for the enzymatic acetylation and deacetylation of lysines residues, respectively. Substrates of HATs include DNA-binding proteins (e.g. histones), non-nuclear proteins (e.g. tubulin) and shuttle proteins.[24]. Bromodomains (BDs) are responsible for the recognition of KAc.

1.2.2. Bromodomains: reader domains for the acetylation of lysine residues

BDs are readers for the KAc PTM, and were first identified from the *brahma* gene of drosophila by Tamkun *et al*.[25] BDs are part of large multi-domain BD-containing proteins (BCPs) that participate in protein-protein interactions (PPIs). [26] BDs are druggable targets that affect human diseases including cancer, inflammation, and diabetes.[27] There are 61 BDs in a total of 46 BCPs encoded for in the human genome, and they have been classified into 8 BD subfamilies (Figure 1.5).[28]

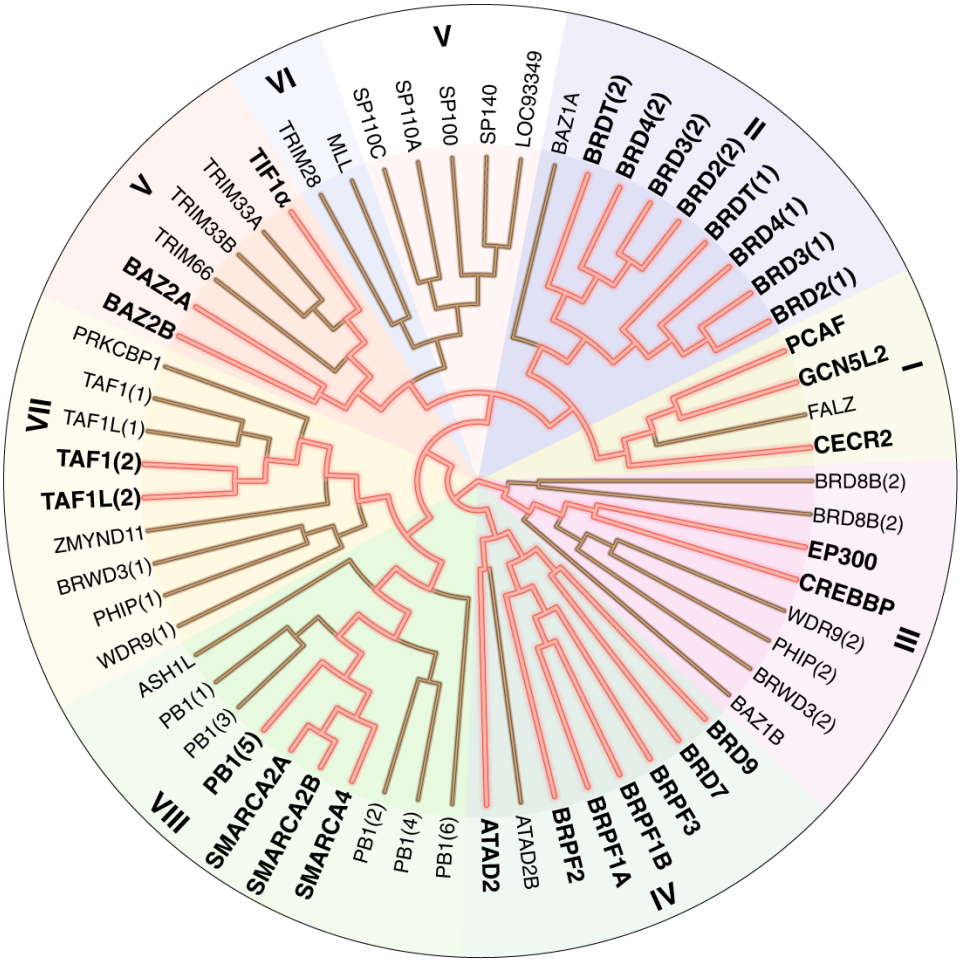


Figure 1.5: Phylogenetic tree of the 61 human BD proteins. BD family members for which probes have been developed are shown in bold and highlighted by dark red lines. Reprinted and adapted from Wu, Q;. *et al*. [29] *Nature Communications* **2019,** *10*, 1915. Copyright 2014 Springer.

The human BDs share a common fold, consisting of four anti-parallel α-helices (αZ, αA, αA, αC) and two loops (AZ, BC). Most of these BDs have a conserved asparagine and tyrosine residue (Figure 1.6),[19, 26] which recognise KAc through hydrogen bonding. BCPs commonly contain other epigenetic domains in close proximity to the BD. Such domains include enzymatic writers and erasers, as HATs and HDACs, other BDs or various protein-interaction mediating domain, such as bromo-adjacent homology (BAH) domains.[30]

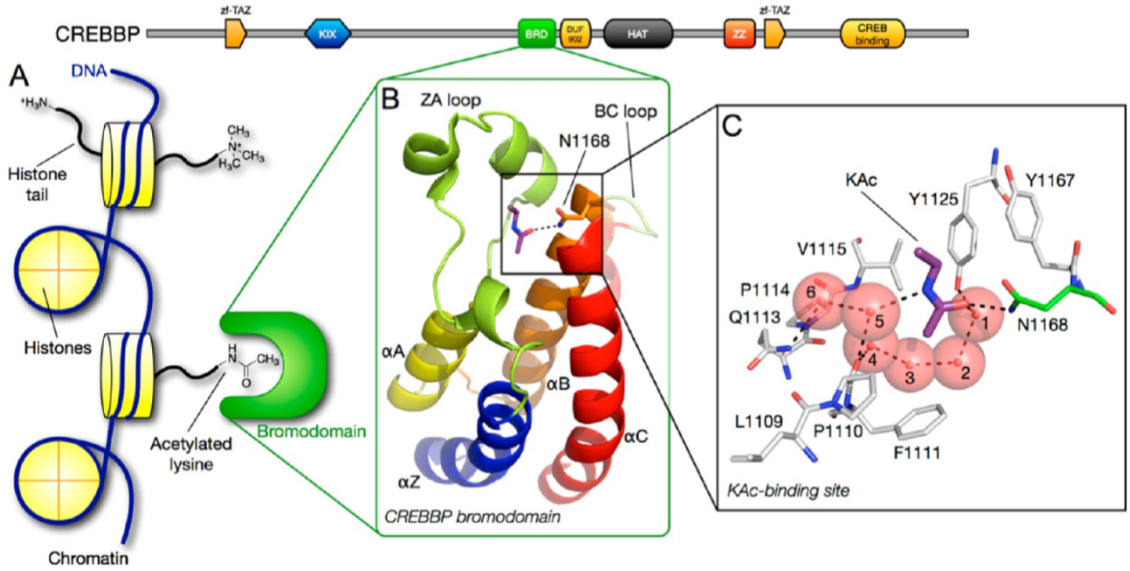


Figure 1.6: (A) Histones tails display multiple PTMs, KAcs are "read" by BCPs. (B) X-Ray crystal structures of the human cAMP response element-binding protein (CREB) binding protein (CREBBP) (PDB code: 3P1C). The bromodomain fold comprises four helices αZ (blue), αA (yellow), αB (orange), and αC (red) and two loop regions known as the ZA loop and the BC loop (light green). (C) The KAc residue is recognised in a well-defined binding pocket through direct interaction with N1168 and through the structure water molecule 1 to Y1125. Reprinted with permission from Hewings, D. S.; *et al.*[19] *J. Med. Chem.* **2012,** *55*, 9393-9413. Copyright 2012 American Chemical Society.

1.2.2.1. BET BDs

The bromodomain and extra C-terminal domain (BET) are tandem BD proteins, which have attracted significant interest in the field of epigenetics. BET BDs (BRD2, BRD3, BRD4 and BRDT) share a high sequence conservation as well as the general structure of two adjacent BD domains (BD1 and BD2, referred to as BRD4(1) and BRD4(2), respectively). One main feature of the BET BDs is their potential to simultaneously recognise two adjacent KAc PTMs on histone tails. Indeed, the combined BD1 and BD2 affinity is higher when multiple acetylation are present within the span of a few amino-acids.[27]

The most studied BET BD is BRD4(1), which was shown early on to be involved in cancer and inflammation. For instance, BRD4 has been shown to bind transcription factors during mitosis.[31] The acetylated histone tail peptide or mimicking ligands most often bind through the ZA-channel, which is lined with the hydrophobic residues. In particular, tryptophan 86 (W), proline 87 (P) and phenylalanine 88 (F) form what is known as the WPF shelf, which can be exploited for ligand affinity.

There has been a rapid expansion of examples of small molecule KAc-mimicking ligands, which inhibit the PPI between BET BDs and histone tails (Figure 1.7). Useful ligands include the pan-BET bromodomain ligands (+)JQ1 (**3**)[32] and I-BET762 (**4**),[33-34] both based on a triazolodiazepine scaffold (Figure 1.7). (+)JQ1 (**3**) is one of the first examples of inhibition of the BET BDs. Filippakopoulos and coworkers[32] chose to target BRD4 in the case of its fusion with nuclear protein in testis (NUT) forms an oncoprotein. The targeting of BRD4-NUT results in the stop of proliferation and in the cell differentiation in human squamous carcinoma. The BRD4-NUT oncogene is a chromosomal translocation and its RNA silencing has similar effects than treatment with (+)JQ1 (**3**) on this aggressive form of cancer.[26, 32, 34]

Other BET BD ligands include: I-BET151 (**5**),[35] I-BET726 (**6**),[36] and the dimethyl isoxazole-containing OXFBD02 (**7**)[37] (Table 1.1). Interestingly, the serine/threonine-specific Polo-like kinase 1 (PLK1) kinase inhibitors BI-2536 (**8**) and BI-6727 (**9**) were later reported to show strong inhibition of BET BD proteins.[38-39] The inhibition of several targets, polypharmacology, has the potential to limit resistance, therefore, these highly potent dual kinase/bromodomain inhibitors are particularly attractive for use as anticancer agents.

Table 1.1: Summary of the KAc mimic and affinities of BET BDs ligands.

BET ligand	KAc mimic	Protein target	Affinity
(+)JQ1 (**3**)[32]	triazolodiazepine	BRD4(1)	K_D=50 nM (ITC)
I-BET762 (**4**)[33-34]	triazolodiazepine	BRD2(1)	K_D=46 nM (ITC)
I-BET151 (**5**)[35]	dimethyl isoxazole	BRD4(1)	IC_{50}=794 nM (FP)
I-BET726 (**6**)[36]	dihydroquinolinone	BRD4(1,2)	K_D=4.4 nM (ITC)
OXFBD02 (**7**)[37]	dimethyl isoxazole	BRD4(1)	IC_{50}=382 nM (AlphaScreen)
BI-2536 (**8**) [38-39]	dihydropteridinone	BRD4(1)	K_D=37 nM (ITC)
BI-6727 (**9**)[38-39]	dihydropteridinone	BRD4(1)	K_D=79 nM (ITC)

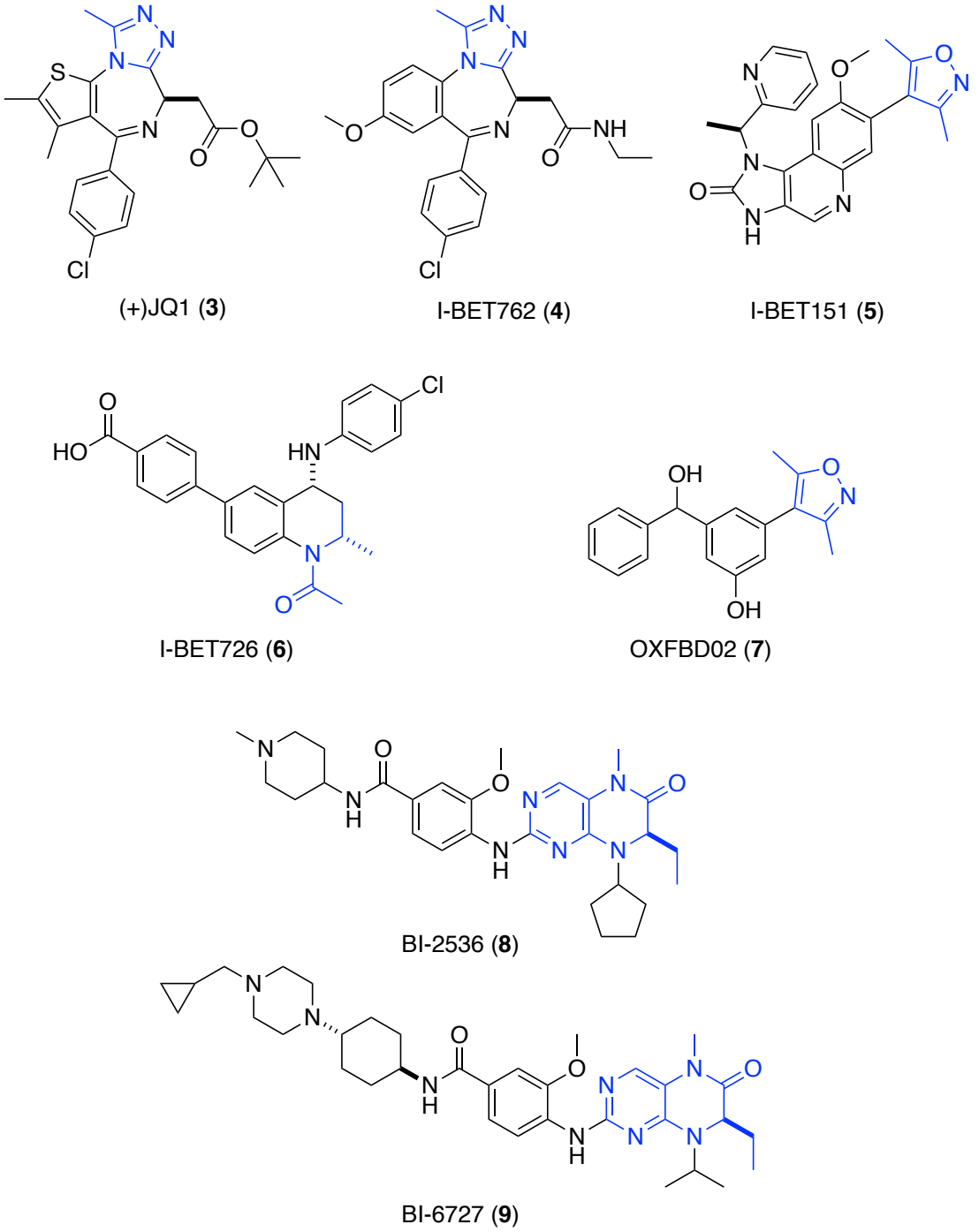


Figure 1.7: The structures of the BET BD inhibitors (+)JQ1 (**3**),[32] I-BET762 (**4**),[34] I-BET151 (**5**),[35] I-BET726 (**6**),[36] and OXFBD02 (**7**)[37]. BI-2536 (**8**) and BI-6727 (**9**) are potent dual PLK1/BRD4, showing inhibition of PLK kinases and BET BD proteins.[38-39] The KAc mimics are highlighted in blue.

BDs share a conserved general structure, however, substantial variation is found in the ZA loop, as well as the N- and C-termini. Despite a high degree of homology between BDs, small differences can be exploited to develop ligands that are selective for specific classes of BDs.

1.2.2.2. Non-BET BDs

Although other BDs have been studied less extensively than BET BDs, they present some interesting features including regulating transcription factors. For example, the inhibition of CREBBP a large protein complex which regulates p53 (Figure 1.6), has been of high interest. This has led to the development of CREBBP BD inhibitors, including I-CBP112 (**10**)[40] and **11**[41] (Figure 1.8).

Similarly, p300/CBP-associated factor (PCAF) is a transcriptional cofactor which contains a BD, a HAT, and a E3 (ubiquitin ligase enzyme 3) ligase ubiquitination domain. PCAF shares a high homology of both structure and function with the general control nonderepressible 5 (GCN5). PCAF and GCN5 are involved in oncologic pathways, hence their selective BD inhibition provides valuable information on these BCPs.[42] **12** is an example of a PCAF/GCN5 chemical probe with a 1800-fold selectivity over the BET BDs.[42] The high selectivity over BET BDs was sought to avoid any interference from the well-characterised BET phenotype during biological characterisation of the inhibition of PCAF and GCN5.

Interestingly, KAc mimics can be tailored to gain selectivity for one BD, or for a sub-family of BDs, over the rest of the human BD proteins. For example, GSK6853 (**13**) binds bromodomain and PHD finger-containing family 1 (BRPF1) with a K_D of 10 nM,[43] and **14** binds BRPF1B and Tripartite motif 24 protein (TRIM24) with a K_D of 137 nM and 222 nM, respectively.[44] Both ligands are based on a benzimidazolone KAc mimic, and show high selectivity over the rest of the human BDs.

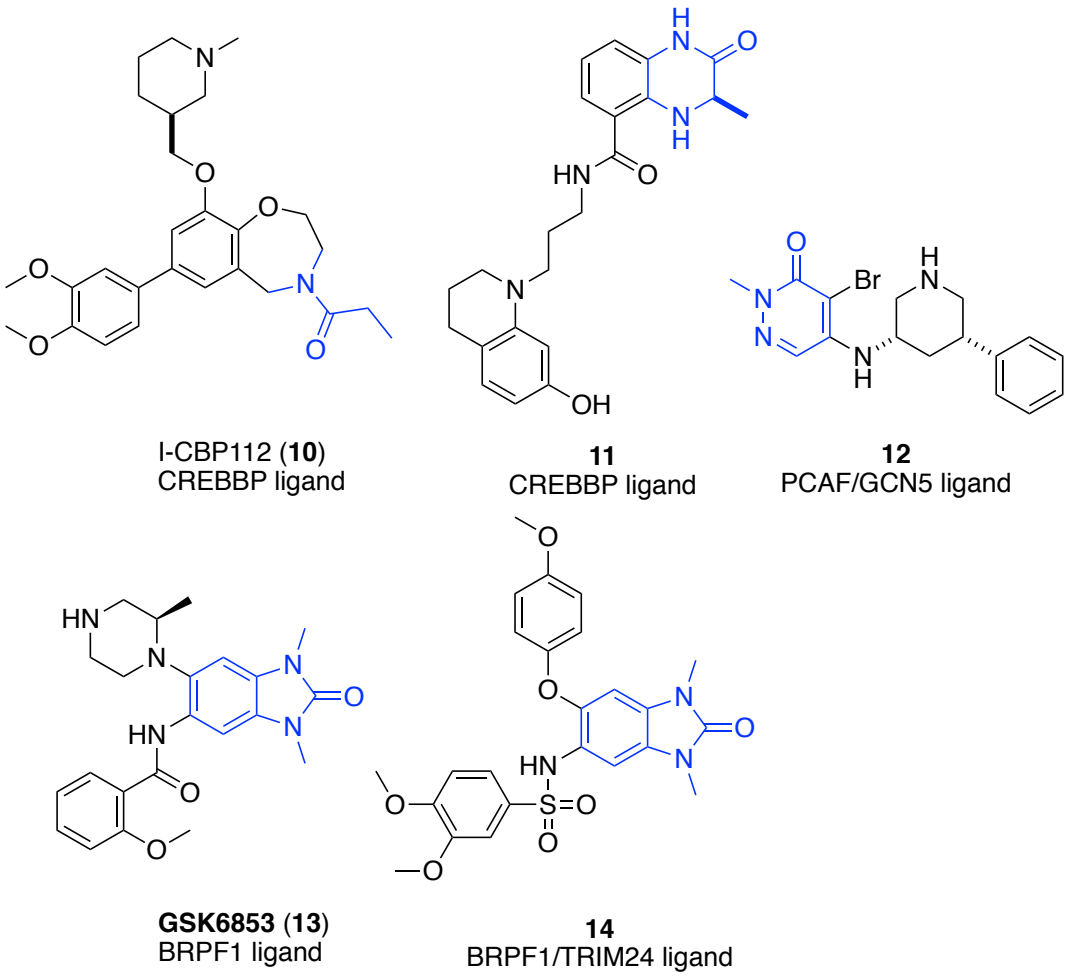


Figure 1.8: The structures of non-BET inhibitors: I-CBP112 (**10**), **11**, **12**, GSK6853 (**13**) and **14**.[40-44]

1.2.3. BD protein recombinant expression and biophysical assays

1.2.3.1. Recombinant expression

Despite being part of multi-domain BCPs, BD proteins can be expressed recombinantly as single domains. Using traditional cloning strategies, BD-encoding genes and vectors containing antibacterial-resistance genes can be combined into plasmids that can be efficiently expressed in bacterial culture.[45] The recombinant single domains are often flanked by a cleavable purification tags, such as a hexa-histidine tag (Histag). Recombinant expression gives BDs in suitable quantities for the use in biophysical assays, which can be used to screen large compound libraries. Additionally, access to soluble human BDs allows for protein crystallisation. Indeed, over the past 20 years, a general effort has led to crystal and NMR-based structures for many human BD proteins.[30, 45]

In addition to their recombinant expression, human BD proteins have been amenable to diverse biophysical or biochemical assays including thermal shift,[46] Alpha Screen,[32] waterLOGSY,[47] fluorescence polarisation,[33, 48] and isothermal calorimetry.[49] The techniques used in this work are described in the following section.

1.2.3.2. Thermal Shift

The thermal shift assay, also referred to as differential scanning fluorometry (DSF), is a rapid screening method to identify small molecule ligands that bind and stabilise or destabilise proteins.[50-53] The melting temperature of a protein (T_m) is measured by the fluorescence increase of a hydrophobic dye.[52] A hydrophobic dye, such as Sypro Orange® (the exact structure of which is not disclosed by the manufacturer) is quenched in aqueous buffer. As the protein denatures, the dye binds the newly exposed hydrophobic patches and the fluorescence intensity (FI) increases (Figure 1.9).[50] The melting temperature is then calculated by a simple fitting of the transition curve. The stabilisation of the protein by a ligand is observed by the shift of the melting temperature (ΔT_m).[50] One of the drawbacks of this widely used assay is that it is highly dependent on physicochemical properties of the ligand and the protein. Additionally, hydrophobic compounds might bind the fluorescent dye directly and interfere with the assay.

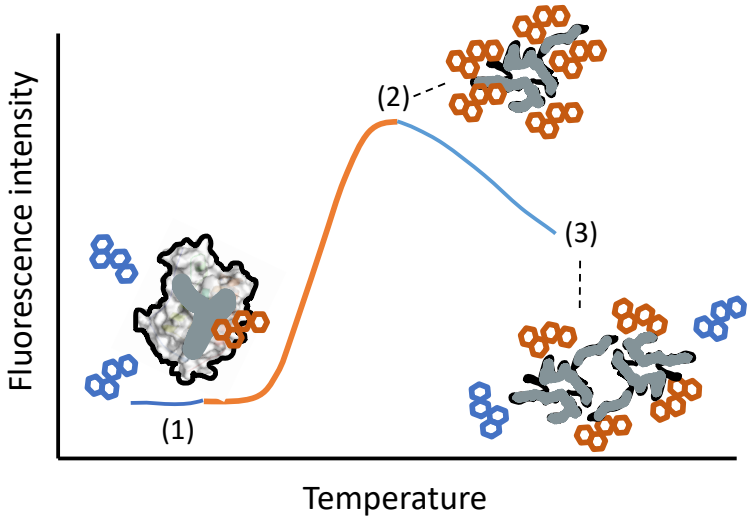


Figure 1.9: Schematic representation of fluorescence intensity versus temperature for the unfolding of protein in the presence of Sypro Orange® dye (blue in aqueous environments, orange in hydrophobic environments) for the thermal shift assay. (1) The dye's fluorescence is quenched in aqueous environment (blue). (2) the fluorescence intensity (FI) goes up as the protein unfolds and hydrophobic patches (in grey) are exposed (3) The FI goes back down as the protein aggregates and hydrophobic patches become inaccessible for the dye.[50]

1.2.3.3. WaterLOGSY

Water-ligand observed *via* gradient spectroscopy (waterLOGSY) is a NMR-based assay used to detect weak binders by observing changes in the sign of the nuclear Overhauser effect (nOe).[47] nOe is the change of nuclear spin population in non-irradiated nuclei caused by T_1 relaxation of irradiated neighbouring nuclei by dipole-dipole mechanism. The waterLOGSY pulse sequence excites the bulk water molecules, supresses the water peak and measures the changes in T_1 relaxation. This pulse sequence allows for detection of binding through the change of nOe signs between free and bound ligands. nOe transferred directly from water protons to ligands is positive (due to fast molecular motion), and freely rotating compound appear as negative peaks in the waterLOGSY spectrum (Figure 1.10). The nOe is negative when is it transferred through the protein (due to slow molecular motion), and the binding compound appears as positive peaks in the waterLOGSY spectrum. Therefore, binders can be identified by their change of nOe sign in the waterLOGSY assay. This technique relies on fast ligand-protein exchange, for the nOe to be transferred to a high enough portion of the ligand population to be measurable.[47, 54]

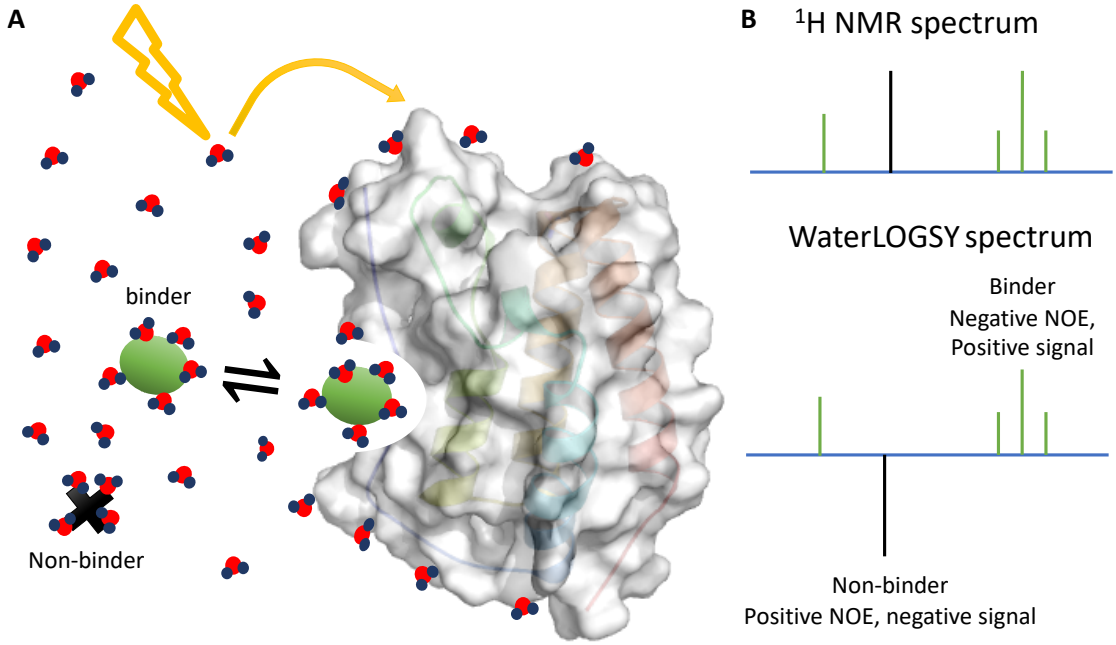


Figure 1.10: Schematic representation of WaterLOGSY assay. (A) The bulk water is excited by the pulse sequence (yellow). nOe is transferred from the water protons to the ligand results either directly (positive nOe) or via the protein bound (negative nOe). (B) The ^{1}H NMR (top) and the corresponding waterLOGSY spectra (bottom) for a binding compound (green) in the presence of a non-binding reference (black).

One important factor is that if a compound has poor solubility and forms aggregates in solution, it will have negative nOe and positive signals in the waterLOGSY spectrum. Therefore, to eliminate false positives, it is important to verify the suitable solubility of the ligand first, as a compound-only sample, before measuring the protein/ligand sample. Overall, WaterLOGSY is suitable to identify weak ligands (1–1000 μM); however, it does not provide information on the binding affinity to the protein, has solubility constraints for the compounds tested and is relatively low throughput.

1.2.3.4. Fluorescence polarisation

Fluorescence polarisation (FP), also referred to as fluorescence anisotropy, is a medium to high throughput competition assay that detects binding of a ligand by displacement of a FP probe bound to the protein. Following excitation by polarised light, the free fast tumbling fluorophore emits depolarised light (Figure 1.11A) whereas the fluorescent probe bound to the protein tumbles much slower, and emits polarised light (Figure 1.11B). Therefore, the displacement of the fluorescent probe to the protein is accompanied by a decrease in the polarisation of the emitted light. The transition between the fully bound (high polarisation) and the fully free (low polarisation) states can be used to determine IC_{50} values for the competitor molecules

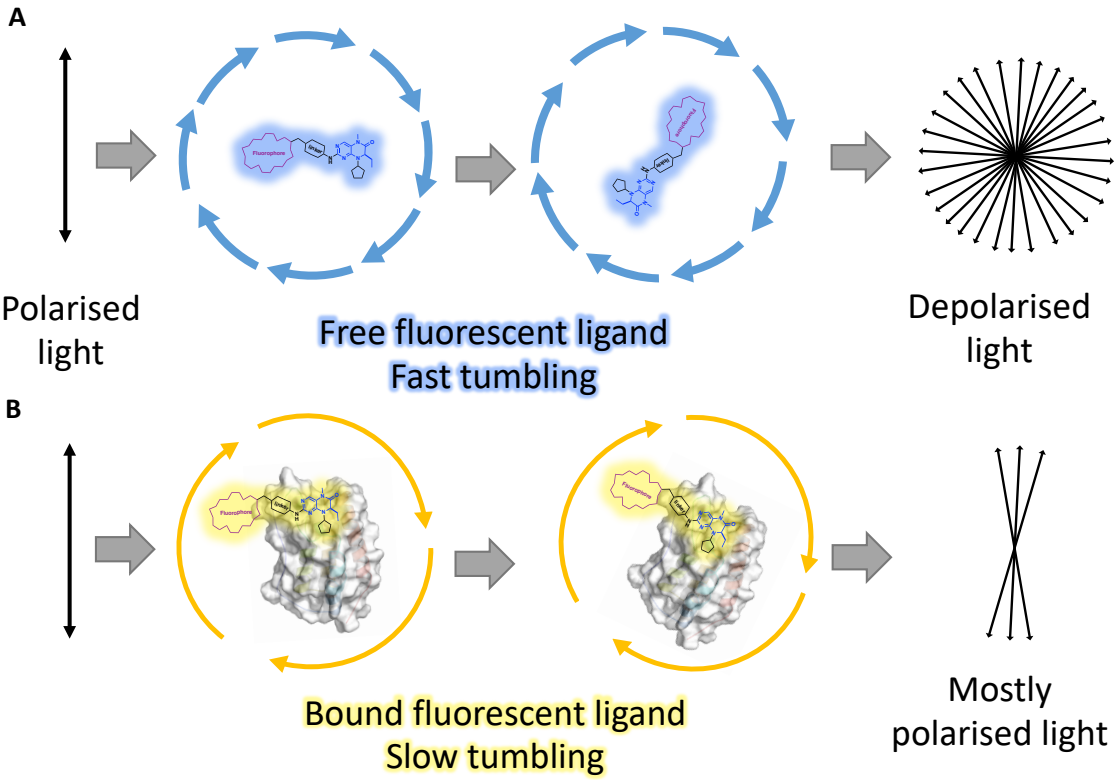


Figure 1.11: Schematic representation of the use of fluorescence polarisation (FP) to detect binding to a protein. The incident polarised light excites the fluorophore, the polarisation state of the emitted light depends on the tumbling rate of the fluorescent ligand: (A) fast tumbling of the free ligand (blue) results in high depolarisation. (B) Slow tumbling of the fluorescent ligand (yellow), arising from binding to a protein, results in a greater amount of polarisation for the emitted light.

1.2.3.5. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is sometimes referred to as "the gold standard" for binding assays. It provides the binding equilibrium constant K_D, the enthalpy ΔH, the Gibbs energy ΔG, and the molarity of the binding process.[55-56] ITC detects the heat changes upon binding, by measuring the power required to keep the sample cell at the same temperature as the reference cell (Figure 1.12). ITC is, therefore, a direct thermodynamic measurement of the interaction. This technique is typically used to determine the K_D of an interaction already identified by a high throughput screening assay, as it can detect medium to tight binders, with K_D values ranging from *ca*. 1 nM to 100 µM. One of the drawbacks of this assay is that high protein concentrations are required to characterise K_D values higher than 20 µM. Additionally, complex binding events with, for example, stepwise binding do not provide clear heat outputs.

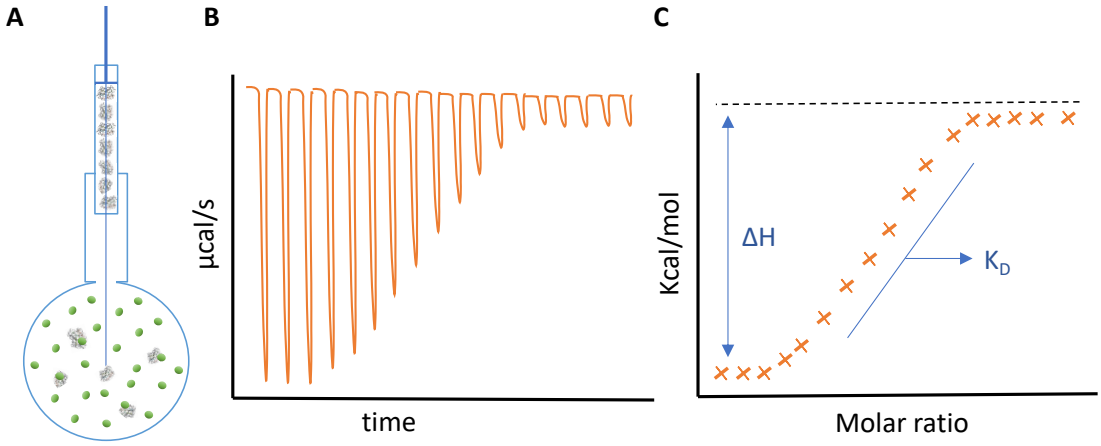


Figure 1.12: ITC determines thermodynamic properties including: stoichiometry of the interaction (n), affinity constant (K_D), change in enthalpy (ΔH), and change in entropy (ΔS). (A) The ITC sample cell and the titrating syringe setup (B) Measured heat given to sample over the time of the titration to maintain sample cell temperature the same as the reference cell. (C) The peak integration of the power input can be plotted against the titration molar ratio to enable the determination of the thermodynamic parameters.

1.2.4. BDs as epigenetics targets for drug development

Over the past 20 years, there has been an exponential increase in publications on the design and optimisation of molecules targeting epigenetic targets.[23] BDs have received a particular attention as they are the main reader domains of KAc. One key factor that allowed for such a rapid growing interest in the targeting of BDs is the ability to recombinantly express the isolated domains and perform a wide variety of biophysical assays.[30, 45] Furthermore, correlation has been shown between cellular activity of BD inhibitors and their binding affinities determined by biophysical techniques.[36, 43, 49, 57-58]

BDs are druggable targets and there are currently several structurally diverse BD inhibitors in Phase I and Phase II clinical trials.[7] The progress of BD inhibitors into the clinic has been rapid despite BD-KAc interaction being a PPI, which has been traditionally considered challenging in medicinal chemistry.[59] BDs play a fundamental role in regulating transcription in humans, hence we hypothesise that BCPs will play equally important roles in other organisms.

1.3. Epigenetics of *T. cruzi*

A whole-organism proteomics analysis of the *T. cruzi* parasite has highlighted the change between carbohydrate to lipid-based metabolism when the parasite transitions from infective trypomastigotes to reproductive amastigotes.[15] Such drastic changes in metabolism require careful epigenetic control during differentiation to various life-form stages. Similarly, proteins associated with protein folding, detoxification, and transcription have been shown to be expressed differentially in the four life cycle stages of the parasite. Furthermore, genes encoding for epigenetics writers, readers and erasers have been identified.[15] For example, 36 histone proteins have been identified to be encoded in the *T. cruzi* genome.[15] Overall, there is increasing evidence that epigenetic targets play a crucial role in the parasitic life cycle of protozoan parasites.[60-61]

1.3.1. Known PTMs in *T. cruzi*

The histone proteins in trypanosomes (*T. cruzi* and *T. brucei*) are the most divergent from the rest of eukaryotic organisms, as they show variations of the sequence, lengths and size relative to their counterparts.[62] There have been multiple reports of PTMs in *T. cruzi*, including the methylation, phosphorylation, and acetylation of histone tails.[14] In particular, the parasite histone H1 is phosphorylated, H2A and H4 show several sites of acetylation, whereas H2B and H3 are preferentially methylated.[60] The *T. cruzi* histone H4 can be acetylated on lysines 4, 10, and 14 (Figure 1.13), with each lysine acetylation associated with a different role.[60-61] K4 acetylation occurs preferentially in dense chromatin regions in proliferative situations, whereas K10 and K14 acetylation is distributed in transition regions between different packing of the chromatin and are associated with DNA repair mechanisms.[61]

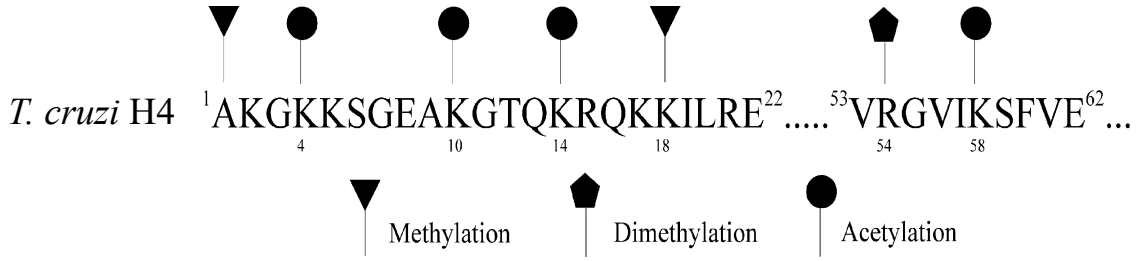


Figure 1.13: Diagram showing the major PTMs found in *T. cruzi* histone H4. Reprinted with permission of Da Cunha *et al*.[60] Copyright 2006 Elsevier .

Interestingly, kinetoplastid parasites (*T. cruzi*, *T. brucei* and *Leishmania donovani*) lack promoter-driven transcription regulation.[14] Therefore, epigenetic mechanism have been hypothesised to be the major transcriptional regulation for these parasites.[14] Abundant evidence of the essential role of epigenetic mechanisms in the life-cycle of parasitic protozoa presents an attractive opportunity in the treatment of Chagas disease. In particular, disrupting the parasite histone code could lead to major misregulation of essential gene expression.[62]

1.3.2. *T. cruzi* BD protein functions: *Tc*BDFs

To date, it has been shown that the *T. cruzi* genome encodes for five sequences which contain a bromodomain: *Tc*BDF1-5.[12, 14, 63, 64] *Tc*BDF5 is a tandem BD, with two adjacent BDs, BD1 and BD2, arranged with a similar spacing as the human BET BD proteins. The identified *Tc*BDFs are thought to be involved in PPIs.[12]

- *Tc*BDF1

*Tc*BDF1 is expressed, to different levels, during each of the life cycle stages of *T. cruzi* and is located in the glycosome.[64] The glycosome is a trypanosomatid-specific organelle that contains glycolytic enzymes. The TriTryp parasites change their energetic metabolism during the different life cycle stages. Hence, the turnover of the glucosome has a vital role in the adaptation of the parasite to nutritional changes. This suggests that *Tc*BDF1 plays a vital adaptive role for *T. cruzi*.[64] Overexpression of wild-type *Tc*BDF1 is detrimental for epimastigotes, but is beneficial for trypomastigotes and amastigotes as it enhances their infectivity and replication, respectively.[64]

- *Tc*BDF2

*Tc*BDF2 has been shown to bind acetylated histone tails H4 and H2a, and peptides derived from the H4 sequence, such as K10Ac-H4 mimic AKGKKSGEAKGTQK(Ac)RQ (Figure 1.13).[63] *Tc*BDF2 is situated in the nucleus and its levels have been observed to be up-regulated following DNA damage by UV radiation.[63] This observation suggests that *Tc*BDF2 is involved in DNA repair mechanisms. During the course of this work, two crystals structures were deposited into the protein data bank (PDB: 6NIM and 6NP7, Figure 1.14A), however, there is not yet any publications associated with these structures.

- *Tc*BDF3

*Tc*BDF3 binds acetylated lysine residues on α-tubulin, one of the components of the cytoskeleton of *T. cruzi* and forms microtubules.[12] *Tc*BDF3 is expressed through all life stages of the parasite and is an exclusively non-nuclear BCP situated in the cytoplasm. Interestingly, *Tc*BDF3 is thought to contribute to the formation and the continued function of flagella in epimastigotes. [12, 65]

- *Tc*BDF4

*Tc*BDF4 has been identified in the *T. cruzi* genome, however, to date there are no reports about the function, expression, or inhibition of this protein.

- *Tc*BDF5

*Tc*BDF5 is tandem BD containing two BDs: *Tc*BDF5-1 and *Tc*BDF5-2. *Tc*BDF5 has been identified in the *T. cruzi* genome, however, to date there are no reports about the function or inhibition of this protein.[65] During the course of this work, a crystal structure for *Tc*BDF5-1 was deposited into the protein data bank (PDB: 6NEY, Figure 1.14B), however, there is no publication associated with this structure.

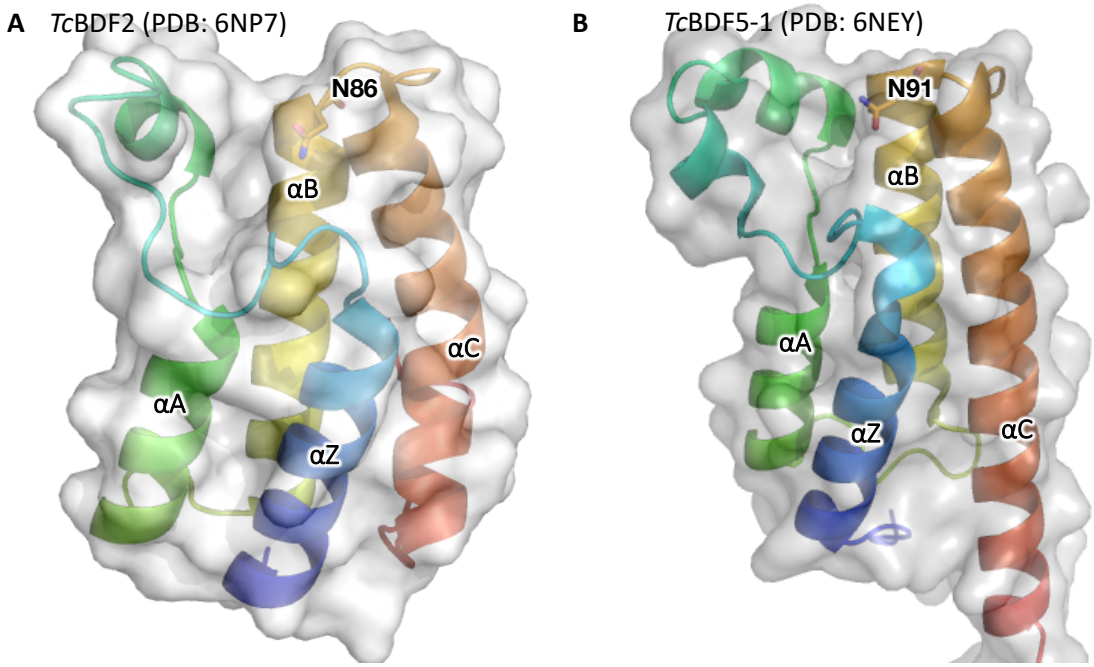


Figure 1.14: The apo-crystal structures of (A) *Tc*BDF2 (PDB: 6NP7) and (B) *Tc*BDF5-1 (PDB: 6NEY) published by Structural Genomics Consortium (SGC) Toronto in January and March 2019, respectively. *Tc*BDF2 and *Tc*BDF5-1 show left-handed bundle of four alpha helices structures and conserved asparagine residues.

1.3.3. *T. cruzi* HDACs and HATS

In addition to BDs, the *T. cruzi* genome encodes genes for regulatory epigenetics enzymes such as HDACs.[15] For example, there are two NAD^+ dependent HDACs in *T. cruzi*, the sirtuins *Tc*Sir2rp1 and *Tc*Sir2rp3. Interestingly, *Tc*Sir2rp3 can be inhibited by salermide, a drug targeting human sirtuins, suggesting that targeting *T. cruzi* HDACs could be used for the treatment of Chagas disease.[66] Similarly, targeting the other classes of HDACs in *T. cruzi* has also shown promising results. For example, the HDAC class III activator resveratrol shows anti-parasitic effects.[16] Furthermore, the *T. brucei* homologues have been shown to be inhibited by several distinct motifs.[67]

1.3.4. Epigenetic targets as a strategy against protozoan parasites

1.3.4.1. Susceptibility of *T. cruzi* to OXFBD02 (**7**)

Preliminary data from Prof. Ivone Carvalho's research group (University of Sao Paulo, unpublished results), showed that OXFBD02 (**7**) has some trypanocidal activity assay on *T. cruzi* in a phenotypic assay (Figure 1.15). This assay determined that OXFBD02 (**7**) has

IC_{50}=18.8 µM, which is 3-fold drop compared to benznidazole (**1**) which has a IC_{50}=5.74 µM in the same assay. This phenotypic assay does not provide information on the potential mode of action for the killing effect of OXFBD02 (**7**).

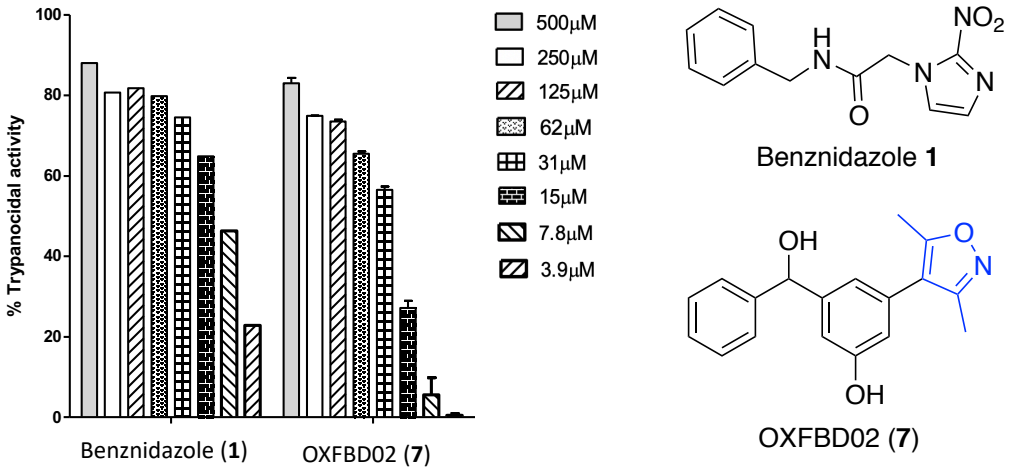
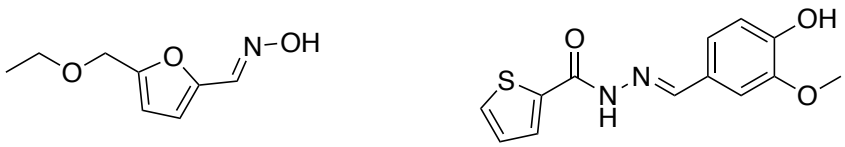


Figure 1.15: Preliminary results of the trypanocidal phenotypic activity of OXFBD02 (**7**) compared to benznidazole. (**1**) Unpublished data from the Carvalho research group (University of Sao Paulo).

1.3.4.2. Non KAc-mimicking *Tc*BDF3 binders

It has been recently shown that overexpression of *Tc*BDF3 protects the parasite from the human bromodomain inhibitors I-BET151 (**5**) and (+)JQ1 (**3**).[68] Additionally, during the course of this work, Serra and co-workers published two studies identifying new *Tc*BDF3 binders (Figure 1.16).[69-70] The first compound, **15**, was isolated from a chemically engineered extract, which was obtained by reacting a herbal crude extracts with hydroxyl amine.[69] The second compound, **16**, was identified through dynamic combinatorial chemistry.[70] These compounds are both of small size, and are not obvious KAc mimics. These compounds were reported by the Serra research group to bind *Tc*BDF3 with K_D values below 2 µM in a tryptophan fluorescence assay (**15**: K_D= 0.29 µM, **16**: K_D= 1.7 µM).[69-70] As these compounds were only tested with one biophysical assay,[69-70] **15** and **16** were synthesised and subjected to further screening in an attempt to verify these leads. The validation experiments using waterLOGSY and ITC did not detect binding (Chapter 3). The compounds were also found to be unstable in water and have limited aqueous solubility.



15 **16**

Figure 1.16: The structures of proposed *Tc*BDF3 ligands, **15** and **16**, published by the Serra research group within the course of this work.[69-70]

1.3.4.3. I-BET151 (**5**) inhibition of the *T. brucei* BD *Tb*BD2

I-BET151 (**5**) binds to *T. brucei* BD2 and has been observed to prevent the parasite from evading the host immune system (see the crystal structure of *Tb*BDF2, Figure 1.17).[71] Interestingly, the I-BET151 (**5**) ligand adopts a different binding mode when binding the protozoan BD compared to the human protein. Indeed, the ligand binds N86 of *Tb*BDF2 through its benzimidazolone moiety, whereas it binds N140 in human BRD4(1) through its dimethylisoxazole (DMI) group. This observation suggests that protozoan parasite BD proteins can be targeted, and selectivity over human protein might be possible.

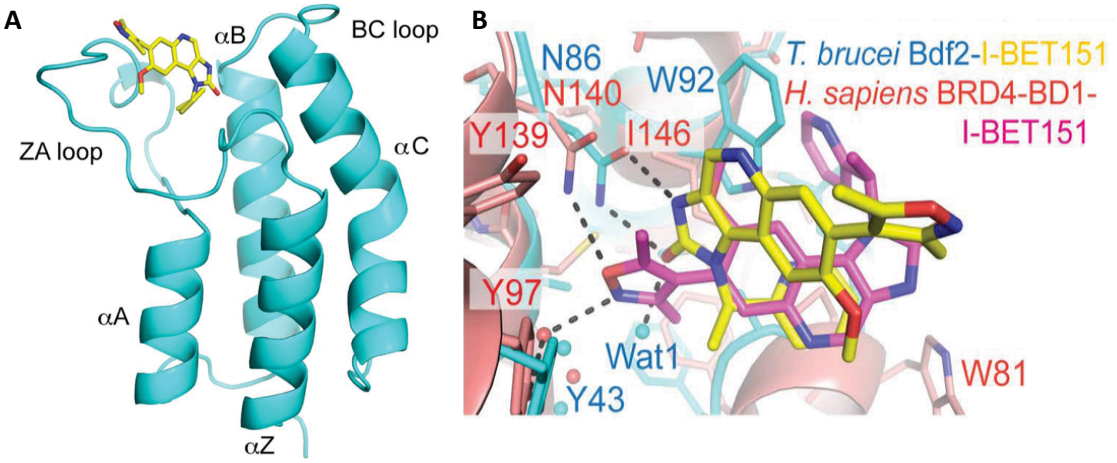


Figure 1.17: (A) *Tb*BDF2 in complex with I-BET151 (**5**) (PDB code 4PKL) (B) *Tb*BDF2 (cyan) and I-BET151 (**5**) (yellow) superimposed BRD4(1) (red) co-crystalised I-BET151 (**5**) (magenta). Adapted with permission from Schulz *et al.*[71] Copyright 2015 Open access under the Creative Commons Attribution License.

Overall, the high homology between the protozoan parasites increases our interest in benzimidazolone and dimethylisoxazole containing molecules for inhibition of the *T. cruzi* BDs. I-BET151 (**5**) could be used as a lead compound and optimised for selectivity of protozoan over human BDs, due to the stark difference in binding modes between *Tb*BDF2 and human BRD4(1).

1.4. Project aims

The life cycle of *T. cruzi* is highly complex and has the potential to be significantly impacted by disruptions of its epigenetic regulation. There is increasing evidence that homologues of the human regulatory epigenetic mechanisms exist in protozoan parasites, and that the parasites high reliance on phenotypic alterations throughout their life cycle makes those attractive targets.

The function of BDs in non-human organisms is poorly understood, however, given the fundamental role played by BDs in regulating transcription in humans, we hypothesise that BCPs will play equally important roles in other organisms. Specifically, we hypothesise that the inhibition of BD proteins in *T. cruzi* will lead to a change in phenotype which might be detrimental to the parasite viability. Consequently, the aim of this project is to design, synthesise and optimise small molecule ligands to probe the function of BCPs from the parasite *T. cruzi*.

To achieve these aims, several challenges needed to be overcome considering that at the start of this work: (1) there were no literature-reported soluble recombinant proteins constructs for the *T. cruzi* BD proteins (2) there were no crystal structures of the *T. cruzi* BD proteins (3) there were limited reports of biophysical assays for these proteins (4) there is, as yet, no evidence that these specific proteins bind KAc-mimicking ligands and behave in an analogous way to the human BDs.

Priority (1): Access to soluble proteins is required for the setup of biophysical assays and is the foundation step for this work. Therefore, the recombinant expression of soluble *T. cruzi* BD protein constructs is the first key priority.

Priority (2): Structural information is extremely valuable when designing and optimising ligands, thus efforts should be made towards crystallisation of the recombinant *T. cruzi* BD protein constructs.

Priority (3): Access to orthogonal reliable biophysical assays is essential to screen for ligands to ensure the early elimination of false positives. Consequently, several orthogonal assays should be established and validated for the study of the *T. cruzi* BD proteins.

Priority (4): First, a limited library of selected compounds should be screened following the hypothesis that the *T. cruzi* BD proteins have the potential to bind KAc mimics. In case of low hit rates, unbiased HTS screens should be trialled through collaborations to identify alternative ligands.

Working through those priorities will allow for the design and optimisation of ligands to probe the role of the *T. cruzi* BCPs. The elucidation of the function of the *T. cruzi* BD proteins will allow determination of whether their inhibition can alter the *T. cruzi* phenotype, and whether it can be used as a strategy to develop novel treatments for Chagas disease.

Chapter 2. Protein expression and structure predictions

2.1. Introduction and Aims

The *T. cruzi* parasite relies on switching between distinct phenotypes through epigenetic mechanisms to survive in humans.[14, 62] The two life forms of *T. cruzi* in humans have different functions and require different metabolism to survive either the muscle, digestive tract or blood. BDs have been identified in *T. cruzi*, and although they haven't been fully characterised, they are thought to bind acetyl lysine residues on the parasite histone tails or α-tubulin.[12, 63] To demonstrate that targeting the *T. cruzi* BDs has the potential to offer new approaches for the treatment of Chagas disease, it was first necessary to confirm their binding to KAc-mimicking moieties. Following the success of fragment-based drug discovery (FBDD) for the study of human BD protein, we chose to use this approach to study the *T. cruzi* BDs. Additionally, BD inhibitors, that bind to human BDs with the highest sequence identity to the studied *T. cruzi* BDs, could be compiled into a focused screening library. Combining these two approaches was thought to provide us valuable information on whether the *Tc*BDFs bind KAc mimics. However, this strategy relies on having access to soluble recombinant proteins that can be expressed and purified from *E. coli* on scale. Additionally, structural information is highly valuable for FBDD, however, there were no available crystal structures of any of the *Tc*BDFs at the start of this work. Structures of *Tc*BDF2 and *Tc*BDF5-1 became available in March 2019 (Section 1.3.2). To date, the literature only described the use of full length *T. cruzi* BD proteins.[68-70] These proteins are not soluble and have to be purified from inclusion bodies, which is a lengthy low-scaling procedure. Critically, this procedure relies on refolding the protein with decreasing amount of urea, bringing the obvious concern about whether the "active" folding state of such proteins is accessible using this method.

The aim of the work detailed in this chapter was to express soluble *Tc*BDFs proteins and attempt their crystallisation. This work would enable the use biophysical assays to assess the binding of *Tc*BDFs to KAc mimics. As such, this work is the foundation of the remainder of the thesis, which focuses on the development of tool compounds for the *Tc*BDFs proteins.

2.2. Protein expression

Our first focus was to express and purify the *T. cruzi* bromodomains in *E. coli*. The BCPs *Tc*BDF1 to 4 have been described briefly in the literature, and the tandem *Tc*BDF5 BCP was identified by Dr Anthony Chan (Conway group, University of Oxford) using a BLAST search. For each of the five BCPs, the BD-encoding regions were identified through a BLAST search by Dr Anthony Chan (Figure 2.1, unpublished data). He then designed and cloned a number of plasmids with truncated *Tc*BDF1 to 5 constructs each containing the BD of interest, with varying start and end amino acids (Appendix A2). The constructs were then expressed in *E. coli*, and purified by sequentially passing the proteins through an ion-metal affinity column (IMAC) and a size exclusion column (SEC) to identify constructs yielding soluble proteins. Insoluble proteins are found mostly in the pellet removed during the initial preparation steps, and also as the first elution peaks in the SEC profile. Through this workflow, Dr Anthony Chan identified soluble constructs for *Tc*BDF2 and *Tc*BDF3: H_6-*Tc*BDF2(1..113aa) and H_6-*Tc*BDF3(32..148aa). H_6- denotes the N terminal His-tag (hexahistidine-tag), and the numbers in parentheses denote the start and end amino acids relative to the full-length construct.

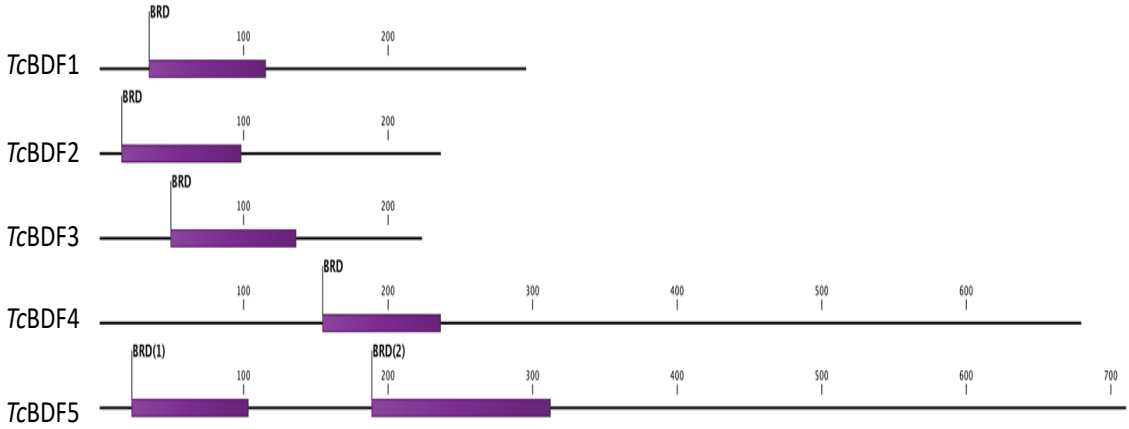


Figure 2.1: The BD-encoding regions of the BCPs *Tc*BDF1 to 5 identified by a BLAST search using CLC Main Workbench v7 (Dr Anthony Chan).

In this Section, optimisation of the large-scale expression of the soluble constructs of *Tc*BDF2 and *Tc*BDF3 will be detailed. Efforts towards finding soluble constructs for *Tc*BDF5-1 and *Tc*BDF5-2, as well as N to F mutants of *Tc*BDF2, *Tc*BDF3 and *Tc*BDF5-1 will be discussed. Finally, the cleavage of the hexahistidine-tag and the screening of

crystallisation conditions for the soluble *T. cruzi* BDs, in particular of *Tc*BDF3, will be described.

2.2.1. H_6-*Tc*BDF2 expression

The plasmid design, cloning, and the subsequent expression trials for the H_6-*Tc*BDF2 constructs were optimised by Dr Anthony Chan. The truncated *Tc*BDF2 construct with the amino acids 1 to 113, H_6-*Tc*BDF2(1..113aa), was identified as a suitable construct through large scale expression. The H_6-*Tc*BDF2(1..113aa) constructs produced soluble protein following expression in *E. coli* BL21 cells and fast protein liquid chromatography (FPLC) purification. These conditions were reproducible, and were used to produce high-quality protein for crystallisation trials and for use in biophysical assays (Figure 2.2).

Crystallisation of human BD proteins is typically done following purification, using a buffer identified through screening of a variety of salt concentrations, pH and small molecules such as PEG. We applied this strategy using commercial buffer screens, which each is a set of 96 different buffers provided to efficiently screen widely different conditions (Appendix D). Crystallisation of H_6-*Tc*BDF2(1..113aa) was attempted using three commercial buffer screen on the apo protein and with the protein in the presence of a 20-fold excess of either OXFBD02 (**7**) or I-BET151 (**5**) (Figure 2.3). Aggregation was observed upon plating, indicating that the concentrations in use are likely to be close to the soluble/crystal interphase. These conditions did not, however, yield protein crystals. Cleaving the His-tag[72] before attempting crystallisation might increase chances of obtaining a protein crystal, as this tag is known to be disordered and in some cases to prevent crystallisation. This option was not explored for *Tc*BDF2(1..113aa) as crystallisation efforts focused on the *Tc*BDF3 constructs (Section 2.2.3.3). Additionally, two *Tc*BDF3 structures became available on the PDB (PDB: 6NIM and 6PN7) early 2019.

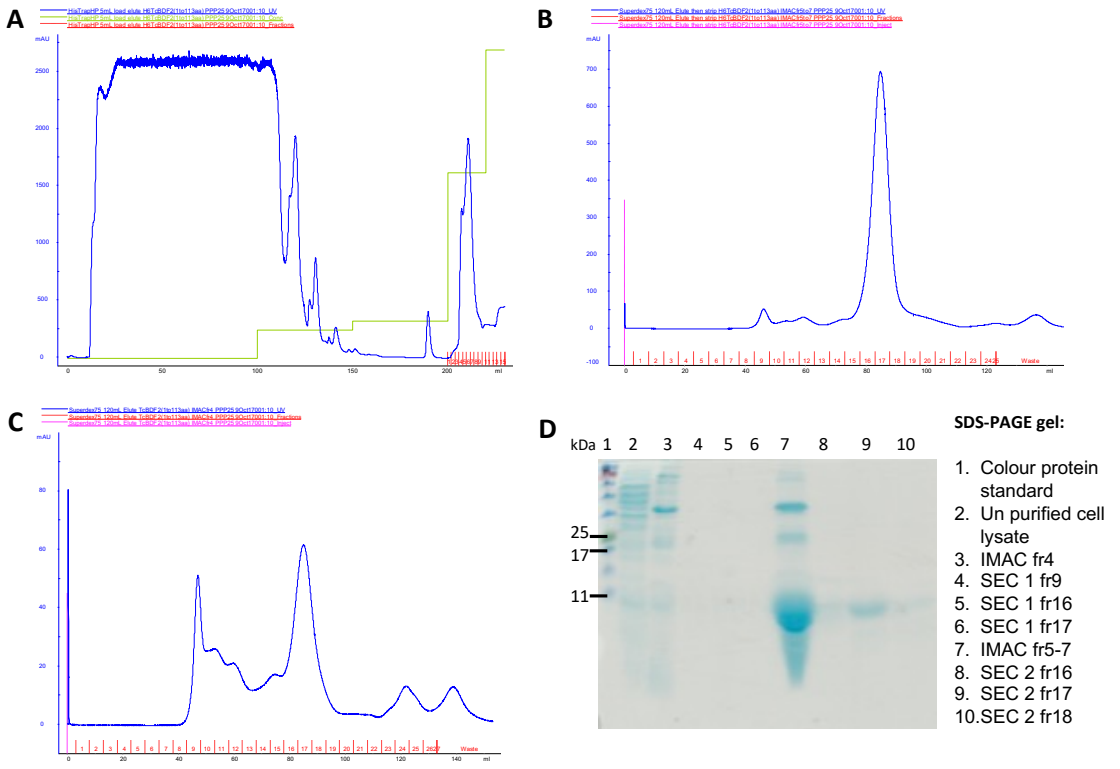


Figure 2.2: H_6-*Tc*BDF2(1..113aa) expression for crystallisation trials: (A). IMAC profile. (B). SEC profile of IMAC fractions 5-7 (SEC 2) showing soluble protein. (C). SEC profile of IMAC fraction 4 (SEC 1) showing aggregation. (D). SDS-PAGE analysis of purified H_6-*Tc*BDF2(1..113aa); Lane 1: Colour protein standard, lane 2: unpurified cell lysate, lane 3: IMAC fraction 4, lanes 4-6: SEC 1 fr 9 and 16-17 showing aggregation, lane 7: IMAC fr5-7, lanes 8-10: SEC 2 fr16-18 showing purified H_6-*Tc*BDF2(1..113aa). The pure fraction SEC 2 fr 17 was concentrated to 7.2 mg/mL and used for crystallisation trials.

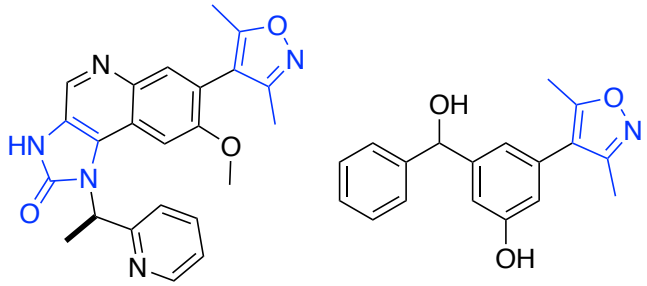


Figure 2.3: The structures of I-BET151 (**5**) and OXFBD02 (**7**).

Table 2.1: Summary of the conditions screened for the crystallisation trials of H_6-*Tc*BDF2(1..113aa). (RT: room temperature, CR: cold room). P denotes the protein concentration, C denotes the compound concentration

Commercial buffer screen	No ligand (P = 502 µM)	OXFBD02 (**7**) (P/C = 100 µM/2 mM)	I-BET151 (5) (P/C = 100 µM/2 mM)
JCSG+	2 RT, 2 CR	1 RT, 1 CR	1 RT, 1 CR
Xal screen	1 RT, 1 CR	–	–
PACT premier	1 RT, 1 CR	–	–

2.2.2. H_6-*Tc*BDF5 expression optimisation and large scale expression

2.2.2.1. Small scale expression trials for H_6-*Tc*BDF5-1 and H_6-*Tc*BDF5-2

To efficiently explore the expression levels and solubility of two set of 9 constructs encoding for the each of the *Tc*BDF5 bromodomains (Figure 2.4), small scale expression was performed using 10 mL cultures in Falcon tubes. The cultures were grown until they reached an OD_{600} of 0.5–0.7, then induced with 0.1 mM IPTG and shaken at 16 °C for 18 h. Small scale lysis was achieved by treatment with BugBuster® detergent, and the samples were analysed using SDS-PAGE (Figure 2.5).

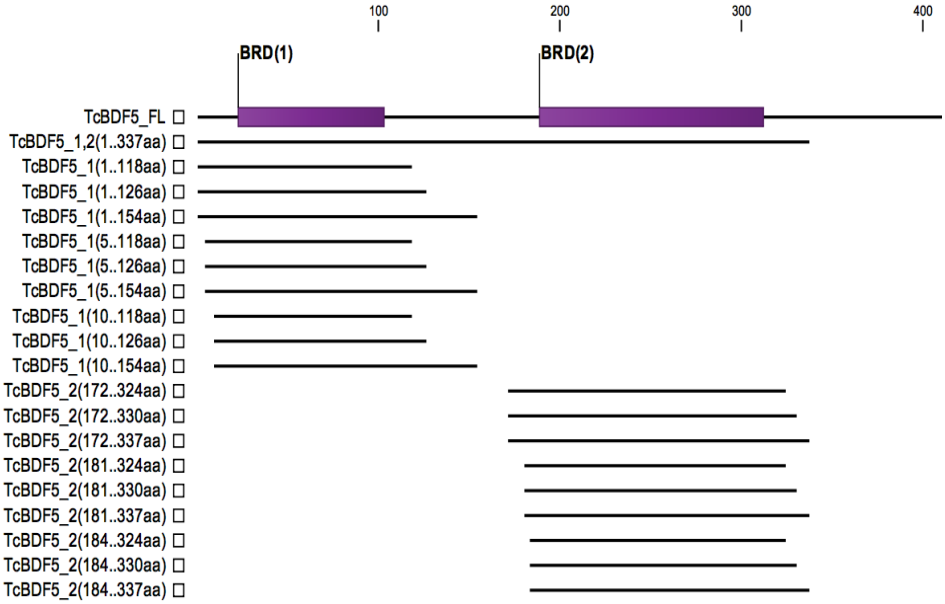


Figure 2.4: The *Tc*BDF5 constructs (residue 1–400 shown) designed and produced by Dr Anthony Chan for expression trials. H_6-*Tc*BDF5 (1..118aa) (Dr Anthony Chan) and H_6-*Tc*BDF5 (1..126aa) were identified as a soluble constructs.

The nine H_6-*Tc*BDF5-1 constructs were shown to be soluble using SDS-PAGE gel. In particular, H_6-*Tc*BDF5-1(1..126aa), H_6-*Tc*BDF5-1(5..118aa) and H_6-*Tc*BDF5-1(5..126aa) had high expression levels, and the majority of the protein was observed in the soluble fractions (Figure 2.5). The large scale expression of H_6-*Tc*BDF5-1 is discussed in Section 2.2.2.2. Unfortunately, no soluble H_6-*Tc*BDF5-2 constructs were identified, hence solubilising tags were investigated. The expression of H_6-*Tc*BDF5-2 as part of fusion protein with a SUMO tag is detailed in Appendix A4.

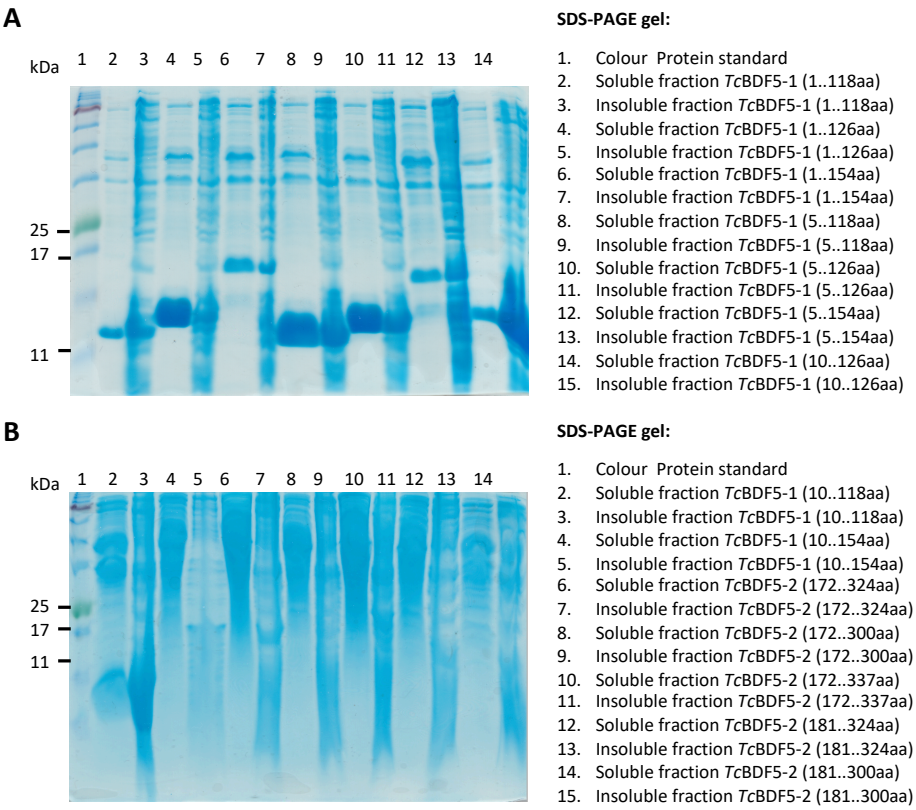


Figure 2.5: Small scale expression trial for *Tc*BDF5: SDS-PAGE (16% Tris/Gly/SDS gel) analysis of H_6-*Tc*BDF5-1 (A) and H_6-*Tc*BDF5-2 (B) constructs. Lane 1:Colour protein standard, lane 2 to 15: Comparison of the expression levels of the various constructs in the supernatant (soluble fraction) or the pellet (insoluble fraction). Cells were lysed with BugBuster®.

2.2.2.2. Large scale expression of *Tc*BDF5-1

In agreement with the SDS-PAGE analysis of the expression trials, large scale expression of H_6-*Tc*BDF5(1..126aa) showed superior yields compared to H_6-*Tc*BDF5(1..118aa), and could be concentrated to higher concentrations without aggregation. H_6-*Tc*BDF5(1..126aa) could be concentrated up to 1868 µM *vs* 1064 µM for H_6-*Tc*BDF5(1..118aa). H_6-*Tc*BDF5-1(1..126aa) was therefore used when *Tc*BDF5-1 was required for biophysical assays. H_6-*Tc*BDF5(1..118aa) was used for crystallisation trials, as it produced purer protein by SDS-PAGE (Table 2.2).

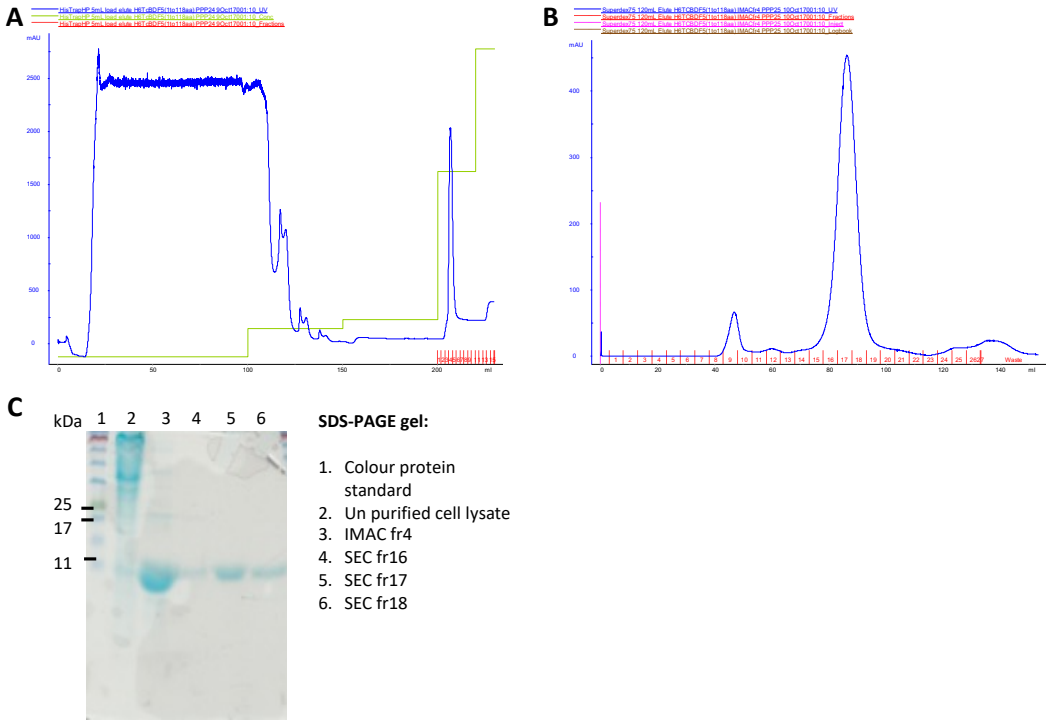


Figure 2.6: Purification of H_6-*Tc*BDF5(1..118aa) (A). IMAC profile (B) SEC profile of IMAC fraction 4 (C). SDS-PAGE gel analysis of purified H_6-*Tc*BDF5(1..118aa); Lane 1: Colour protein standard, lane 2: unpurified cell lysate, lane 3: IMAC fraction 4, lane 4-6: SEC1 fr 16-18.

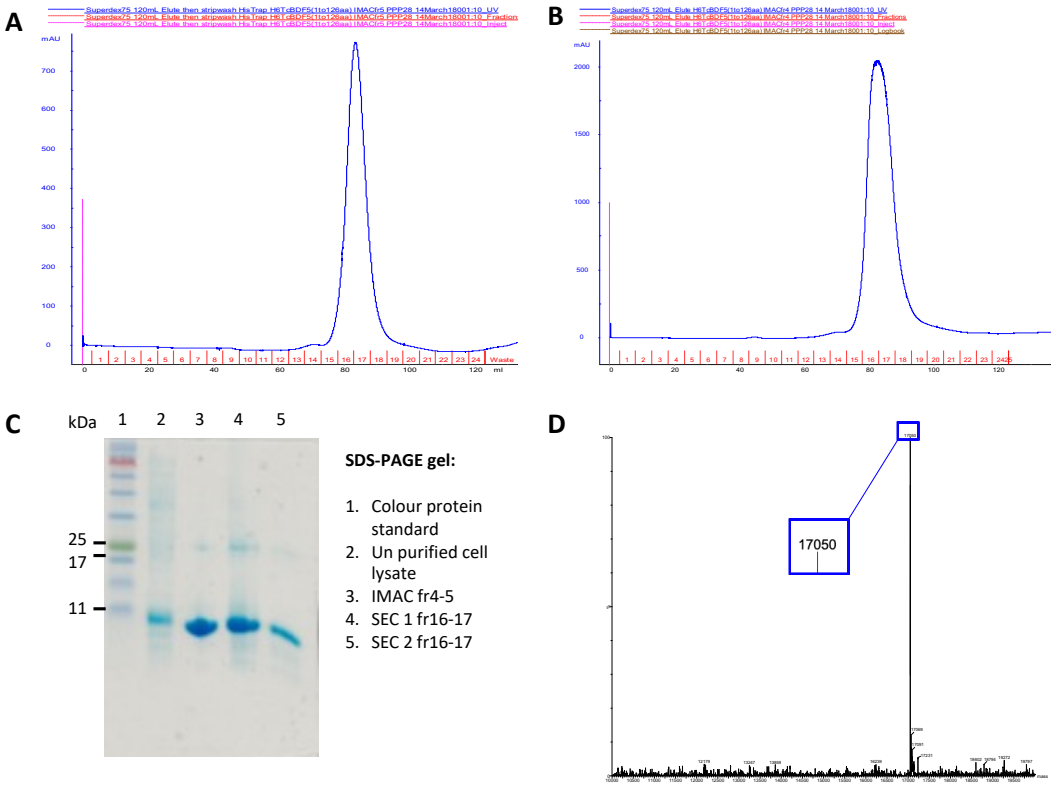


Figure 2.7: Purification of H_6-*Tc*BDF5(1..126aa) for biophysical assays. (A). SEC profile of IMAC fraction 4 (SEC 2). (B). SEC profile of IMAC fraction 5 (SEC 2). (C). SDS-PAGE gel analysis of purified H_6-*Tc*BDF5(1..126aa); Lane 1: Colour protein standard, lane : unpurified cell lysate, lane 3: IMAC fraction 4 to 5, lane 4: SEC1 fr 16-17, lane 5: SEC 2 fr16-17. D. Spectra from protein LCMS on H_6-*Tc*BDF5(1..126aa) (expected mass = 17052.02 Da).

Crystallisation of H_6-*Tc*BDF5(1..118aa) was attempted using commercial buffer screens (Table 2.2) on the apo-protein, and with the protein in the presence of a 20-fold excess of either I-BET151 (**5**) or OXFBD02 (**7**). The conditions screened did not yield protein crystals. Cleaving the His-tag before attempting crystallisation might increase chances of obtaining a protein crystal, however, our priority was set on screening compounds (Chapter 2) and producing a soluble H_6-*Tc*BDF5-2 protein construct (Appendix A4). Additionally, a *Tc*BDF5-1 structure became available on the PDB (PDB: 6NEY) early 2019.

Table 2.2: Summary of the conditions screened for the crystallisation trials of H_6-*Tc*BDF5(1..118aa). (RT: room temperature, CR: cold room). P denotes the protein concentration, C denotes the compound concentration

Commercial buffer screen	No ligand (P = 1064 µM)	OXFBD02 (**7**) (P/C = 100 µM/2 mM)	I-BET151 (5) (P/C = 100 µM/2 mM)
JCSG+	2 RT, 2 CR	1 RT, 1 CR	1 RT, 1 CR
Xal screen	1 RT, 1 CR	–	–
PACT premier	1 RT, 1 CR	–	–

2.2.3. H_6-*Tc*BDF3 expression

2.2.3.1. H_6-*Tc*BDF3(32..148aa) expression and purification optimisation

Following Dr Anthony Chan's initial work, the expression of H_6-*Tc*BDF3(32..148aa) was first performed in 2× TY media: following induction with 0.1 mM IPTG at $OD_{600} \geq 1$, the shakers were then cooled to 16 °C and the cultures were shaken for 18 h before pelleting. This IPTG concentration, and post-induction temperature, ensure slow protein production by the *E. coli* and increase the chances of protein folding, leading to soluble protein. These expression conditions were therefore employed for expressing H_6-*Tc*BDF3(32..148aa) and yielded pellets of medium size (≈ 8 g per 2 L culture). The FPLC-purification gave a low yield of H_6-*Tc*BDF3(32..148aa) (condition 1: 1.5 mg/mL, 0.44 mg of protein per 1 L culture, Table 2.3). Despite small amounts of aggregation being observed, the protein obtained after purification, was of high purity (Appendix A4).

As assay development and crystallisation trials require large amounts of pure soluble protein, it was important to express H_6-*Tc*BDF3(32..148aa) efficiently on large scale. The

first stage of optimisation included pre-cooling the shakers before the induction with IPTG. The cooling was started when the cultures were measured to have an OD_{600} of 0.5 – 0.6 or 0.9 – 1.0. The cultures were left to continue growing in the cooling shakers, and IPTG was added at T ≤ 20 °C. Using these optimised expression condition, a high concentration purification was tested by purifying 2 pellets together (each pellet corresponds to 2 L culture in 2× TY media) to determine if H_6-*Tc*BDF3(32..148aa) is stable to FPLC purification at high concentrations. Concentration of the IMAC fractions prior to the SEC gel filtration did not give excessive aggregation. This approach had not been previously tried as some proteins experience aggregation during spin concentration, but is key for efficient FPLC purification when many IMAC fractions are collected. The SEC filtration yielded highly concentrated purified protein (Figure 2.8). This result shows that the protein is stable to high concentrations, and is likely to tolerate FPLC purification following further expression optimisation. The cell lysis preceding the FPLC purification was improved by using Benzonase® nuclease (Novagen) in addition to polyethyleneimne (PEI) and the SigmaFAST™ Protease Inhibitor cocktail tablet.

These conditions gave H_6-*Tc*BDF3(32..148aa) at high purity and significantly higher concentrations of protein (condition 2: 8.5–10.6 mg/mL, 14.5 mg of protein per 1 L culture, Table 2.3) when purifying 2 pellets together. The protein did not show increased aggregation at the higher concentrations, and so, was suitable for further expression optimisation.

Table 2.3: Summary of the optimisation of the expression H_6-*Tc*BDF3(32..148aa) in *E. coli*.

	Final IPTG concentration	Shaker pre-cooled	OD_{600}	Yield (protein (mg)/ 1 L culture)	Pellet mass	Comments
1	0.1 mM	No	0.9 – 1.0	0.44 mg	8 g	Large scale
2	0.1 mM	Yes	0.9 – 1.0	14.5 mg	2 × 10 g	
3	0.1 mM	Yes	0.5 – 0.6	N/A	N/A	Small scale, analysed by SDS-PAGE.
4	0.5 mM	Yes	0.5 – 0.6	N/A	N/A	
5	1.0 mM	Yes	0.5 – 0.6	N/A	N/A	
6	0.5 mM	Yes	0.5 – 0.6	19 mg	40 g	Large scale (n=2)
7	0.5 mM	Yes	0.5 – 0.6	22 mg	39 g	

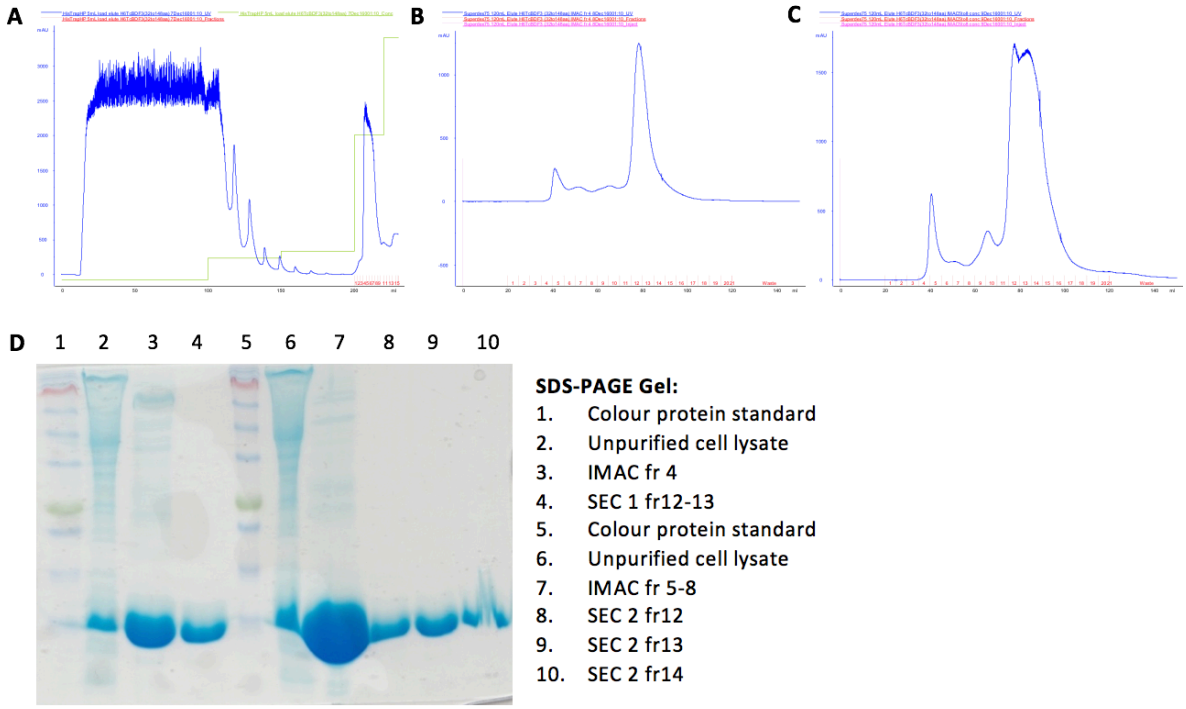


Figure 2.8: Purification of H_6-*Tc*BDF3(32..148aa) at high concentration. A. IMAC purification profile B. SEC profile of IMAC fraction 4 C. SEC profile of IMAC fraction 5 to 8 (concentrated prior to injection) D. SDS-PAGE (16% Tris/Gly/SDS gel) analysis of purified H_6-TcBDF3(32..148aa); Lane 1 and 5: Colour protein standard, broad range (New England Biolabs, Inc.), lane 2 and 6: unpurified cell lysate, lane 3: IMAC fraction 4, lane 4: SEC1 fr 12-13, lane 7 IMAC fraction 5 to 8, lane 8-10: SEC2 fr12 to 14.* Lane 10 distortion due to gel imperfection.

The H_6-*Tc*BDF3(32..148aa) expression optimisation then focused on the IPTG concentration. Three IPTG concentrations were investigated (0.1, 0.5 and 1.0 mM) on small scale, by comparing the H_6-*Tc*BDF3(32..148aa) expression levels by SDS-PAGE gel following small scale lysis with BugBuster® detergent (conditions 3–5 Table 2.3, Figure 2.9). The soluble fractions contained similar amounts of protein, however, the total expression levels, including the insoluble fractions, was higher when using 0.5 mM of IPTG. The low amounts of protein in the soluble fractions was thought to be representative of the poorer efficiency of the cell lysis when using detergents instead of sonication. The optimised IPTG concentration was confirmed by large scale expression and purification, which yielded 76–88 mg of H_6-*Tc*BDF3(32..148aa) from 4 L cultures (condition 7 and 8: 1–2.50 mM, 19–22 mg of protein per 1 L culture, Table 2.3).

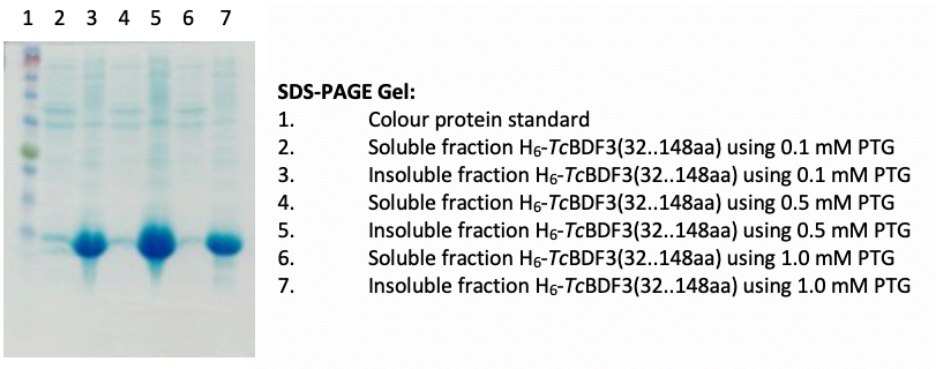


Figure 2.9: SDS-PAGE gel analysis of H_6-*Tc*BDF3(32..148aa). Lane 1: Colour protein standard, lane 2 to 7: Comparison of the expression levels using 0.1, 0.5 and 1.0 mM IPTG in the supernatant (soluble fraction) or the pellet (insoluble fraction) following lysis with BugBuster®.

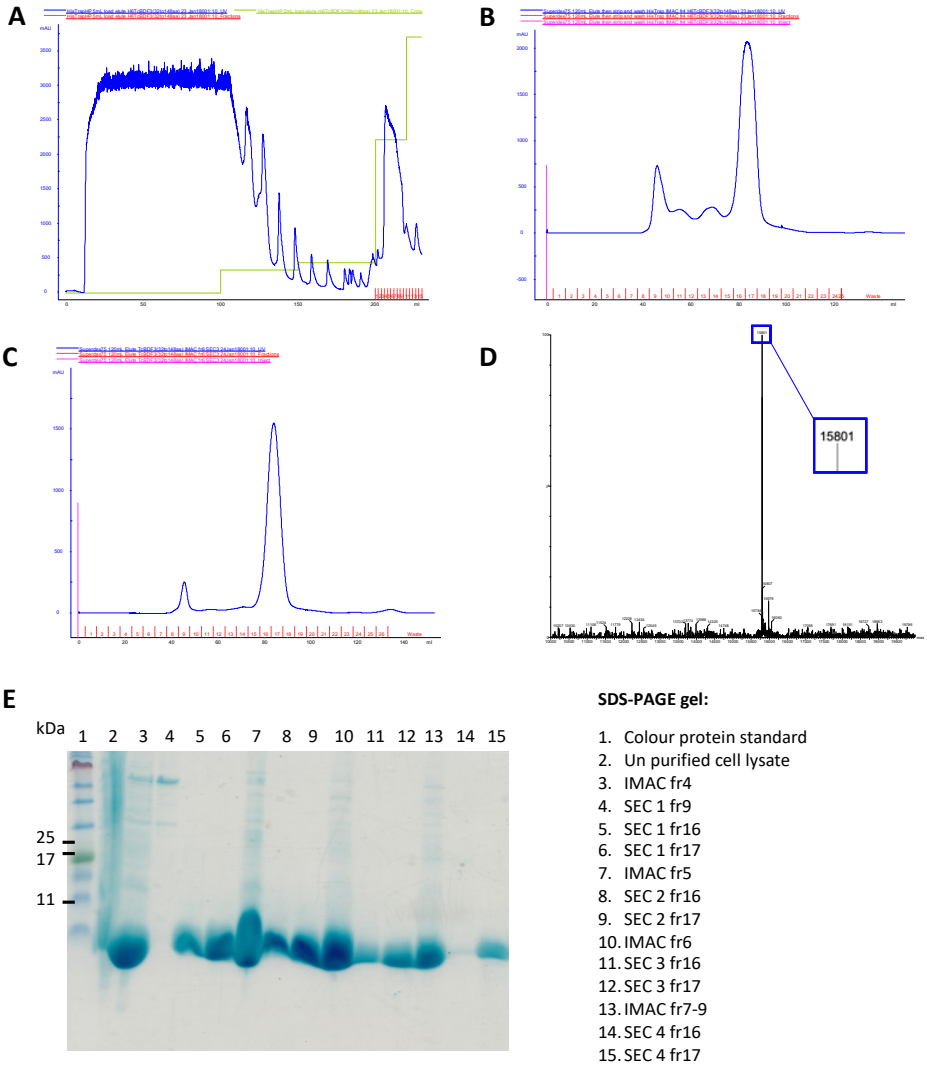


Figure 2.10: Purification of H_6-*Tc*BDF3(32..148aa) following the expression optimisation. A. IMAC purification profile B. SEC profile of IMAC fraction 4 (SEC1) C. SEC profile of IMAC fraction 5 (SEC2, representative of SEC3 and SEC4) D. Protein LCMS mass spectrometry spectrum, expected mass = 15803.7 Da E. SDS-PAGE (16% Tris/Gly/SDS gel) analysis of purified H_6-TcBDF3(32..148aa); Lane 1: Colour protein standard, broad range (New England Biolabs, Inc.), lane 2: unpurified cell lysate, lane 3, 7, 10 and 13: IMAC fraction 4 to 9, lane 4-6, 8-9, 11-12, 14-15: SEC fractions from SEC1-4.

The optimised expression and purification conditions gave an 50-fold increase of protein yield compared to the initial conditions (Table 2.3, Figure 2.10). H_6-*Tc*BDF3(32..148aa) showed high stability and was stored at concentration up to 2.50 mM (39.6 mg/mL).

2.2.3.2. His-tag cleavage with recombinant TEV protease

Small protein tags tend to be flexible and can disrupt protein crystallisation, therefore, cleaving the tag can be advantageous for crystallisation trials.[45] One method to facilitate cleavage of the tag is to introduce a cleavage sequence, such as a tobacco etch virus (TEV) cleaving site.[73] TEV protease recognises the linear peptidic sequence with general form of E-X_{aa}- X_{aa}-Y-X_{aa}-Q-(G/S) where diverse amino acids are tolerated at the X_{aa} position. The sequence that gives the most efficient cleavage is ENLYFQS. The H_6-*Tc*BDF3(32..148aa) plasmid (Appendix A4), includes a TEV cleavage site situated between the His-tag and the protein sequence (Figure 2.11). Work towards identifying suitable cleavage conditions was started by Dr Anthony Chan (preliminary work). rTEV protease produced with standard protein expression and purification techniques, was incubated with fresh IMAC fractions of H_6-*Tc*BDF3(32..148aa). As His-tags bind the nickel ions of the IMAC column, cleaved proteins are expected not to bind to the column and elute with the solvent front. After the incubation of H_6-*Tc*BDF3(32..148aa) with rTEV, the sample was passed through the IMAC column, unfortunately no protein eluted with the solvent front, suggesting that no cleavage had occurred.

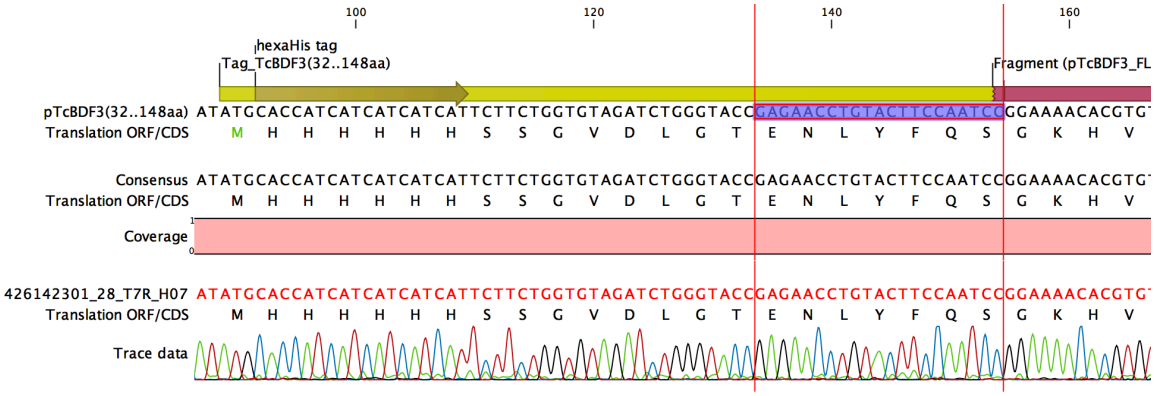


Figure 2.11: Partial plasmid sequencing results showing the His tag and the TEV protease cleaving sequence (ENLYFQS, highlighted). Figure generated with CLC Main Workbench v7.2.

Optimisation of the cleavage condition was performed using small scale screening and analysis to minimise the amount of protein, protease, and time required. Using small reaction volumes, cleavage was monitored by SDS-PAGE analysis, combined with western blot (WB) using an anti-His-tag antibody, and protein mass spectrometry (MS).

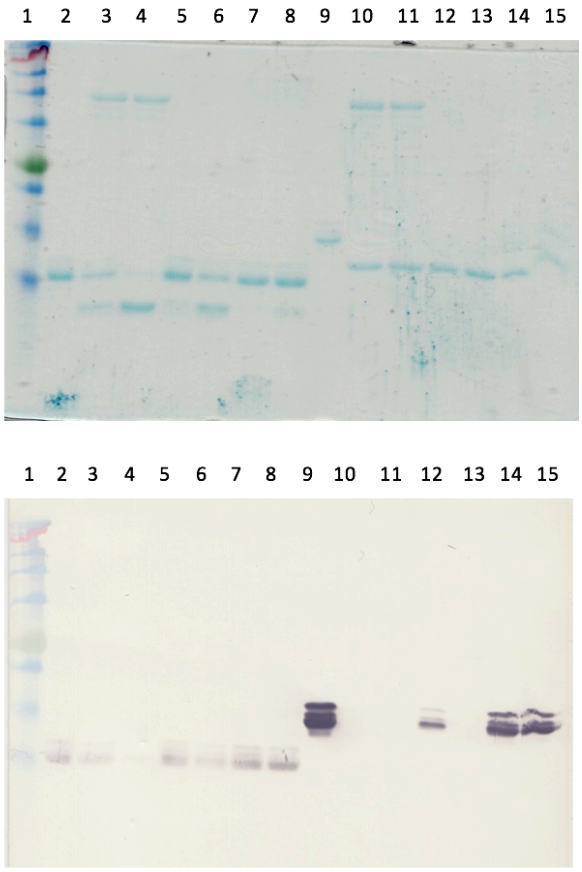
Initial optimisation was conducted using commercially available TEV proteases (Figure 2.12, Table 2.4, and Table 2.5) on H_6-*Tc*BDF3(32..148aa) and H_6-BRD4(1). The cleavage of the His-tag of H_6-BRD4(1) by rTEV is well studied[34, 45] and so this human BD protein was used as a positive control for the commercially available proteases. The general observation was that H_6-*Tc*BDF3(32..148aa) is harder to cleave that H_6-BRD4(1) as it requires higher ratios of TEV protease, longer cleavage time, and is cleaved the most efficiently at room temperature.

Table 2.4: Screening conditions for cleavage of His-tag using commercially available recombinant TEV (rTEV) proteases. Conditions indicated for cleavage of 20 µg of protein, incubating 1–20 h at 4 °C or RT.

	TEV1	TEV2	TEV3
rTEV supplier	Promega®	RnD BioTechne®	InsightBio
Amount of TEV	10 U or 20 U	0.1 µg or 0.4 µg	10 U or 20 U
Cleavage buffer	50 mM HEPES, 0.5 mM EDTA, 1 mM DTT	50 mM Tris, 0.5 mM EDTA, 1 mM DTT	50 mM Tris, 0.5 mM EDTA, 1 mM DTT

Table 2.5: Results of the condition screen for cleavage of His-tag: Full cleavage indicates absence of signal under the anti-his-tag WB (Figure 2.12), as well as absence of full sequence by MS (Appendix A4). Ratios shown are intensity ratios by MS (uncleaved:cleaved). PC: partial cleavage, FC: full cleavage, NC: no cleavage.

	TEV1		TEV2		TEV3	
	H_6-*Tc*BDF3	H_6-BRD4(1)	H_6-*Tc*BDF3	H_6-BRD4(1)	H_6-*Tc*BDF3	H_6-BRD4(1)
Condition 1: 10 U (TEV1 and 3); 0.1 µg (TEV2), 4 °C	PC (1:1)	-	PC (2:1)	-	NC	-
Condition 2: 20 U (TEV1 and 3); 0.4 µg (TEV2), 4 °C	PC	FC	PC	PC	NC	PC
Condition 3: 20 U (TEV1 and 3); 0.4 µg (TEV2), RT	FC	FC	PC (1:5)	FC	PC (2:1)	PC



SDS-PAGE Gel and Western Blot using an Anti-His tag antibody for immunodetection:

1. Colour protein standard
2. H_6-*Tc*BDF3(32..148aa)
3. H_6-*Tc*BDF3(32..148aa) with TEV1 at 4 °C
4. H_6-*Tc*BDF3(32..148aa) with TEV1 at RT
5. H_6-*Tc*BDF3(32..148aa) with TEV2 at 4 °C
6. H_6-*Tc*BDF3(32..148aa) with TEV2 at RT
7. H_6-*Tc*BDF3(32..148aa) with TEV3 at 4 °C
8. H_6-*Tc*BDF3(32..148aa) with TEV3 at RT
9. H_6-BRD4(1)
10. H_6-BRD4(1) with TEV1 at 4 °C
11. H_6-BRD4(1) with TEV1 at RT
12. H_6-BRD4(1) with TEV2 at 4 °C
13. H_6-BRD4(1) with TEV2 at RT
14. H_6-BRD4(1) with TEV3 at 4 °C
15. H_6-BRD4(1) with TEV3 at RT

Figure 2.12: SDS-PAGE gel analysis and western blot for the cleavage conditions screens using 20 U or 0.4 µg of TEV protease for 20 µg of protein (conditions 2 and 3, Table 2.4 and Table 2.5). Aliquots were taken after 20 h incubation at 4 °C or room temperature. Comparison of the SDS-PAGE gel and the WB indicate full cleavage of H_6-BRD4(1) by TEV1 at both temperature, and by TEV2 at room temperature. The temperature dependence is stronger for of H_6-*Tc*BDF3(32..148aa): full cleavage is observed at room temperature with TEV1, whereas only partial cleavage is obtained at 4 °C.

Due to the large amount of commercially-available TEV protease required for the large scale cleavage of H_6-*Tc*BDF3(32..148aa), it was decided to investigate alternative proteases that could be expressed and purified in-house. Two alternative proteases pRK793 and pTH24 (kindly provided by the group of Prof. Christopher J. Schofield (University of Oxford) and by the group of Prof. Manfred Jung (University of Freiburg), respectively) were expressed and purified in-house. (pTH24 was purified by Dr Matthias Schiedel). Both proteases displayed similar activities when cleaving H_6-BRD4(1) in large scale cleavage (Dr Matthias Schiedel, unpublished result).

The cleavage of the His-tag of H_6-*Tc*BDF3(32..148aa) was first attempted on a small scale: The cleavage at 4 °C or RT was monitored at 1, 4 and 20 h using SDS-PAGE gel electrophoresis and protein MS. TEV protease:protein ratios of 1:2, 1:4 or 1:10 were tested. The difference between the ratios is more significant at 4 °C (1:2>1:4>1:10 for cleaving efficiency), and for both temperatures longer incubation time give higher amounts of cleaved protein (Figure 2.13).

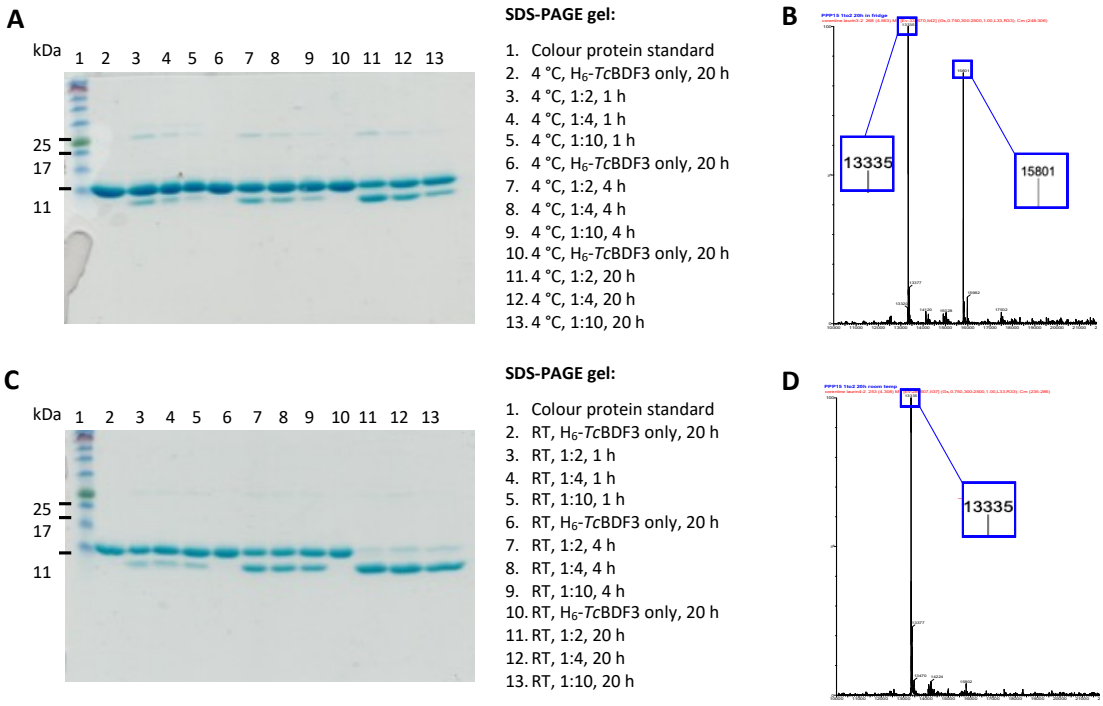


Figure 2.13: Small scale optimization of the H_6-*Tc*BDF3(32..148aa) His-tag cleavage. (A) and (C) SDS-PAGE (16% Tris/Gly/SDS gel) analysis (B) and (D) protein mass spectrometry for 4 °C and RT cleavage using 1:2 TEV protease to protein ratio.

Following these results, large scale cleavage was performed using 1:2 protease:protein ratio at room temperature for 20 h. Cleaved *Tc*BDF3(32..148aa), suitable for crystallisation trials, was produced in high concentrations (17.2 to 23.4 mg/mL) with high purity (Figure 2.14).

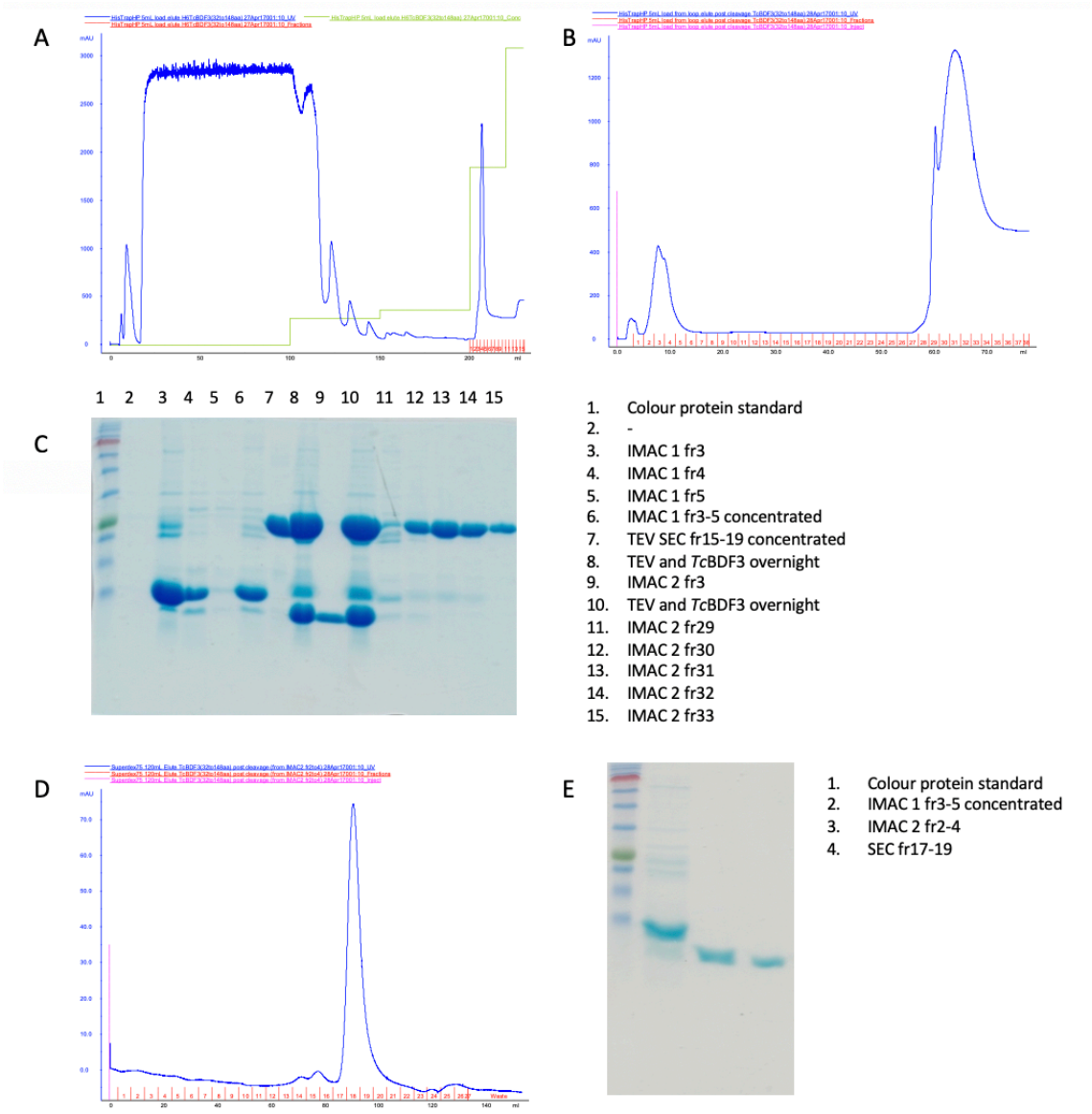


Figure 2.14: Large scale purification and cleavage of H_6-*Tc*BDF3(32..148aa) using a 4 to 1 protein to TEV protease ratio: (A). IMAC profile of H_6-*Tc*BDF3(32..148aa). (B). IMAC 2 profile of *Tc*BDF3(32..148aa) (post cleavage) showing cleaved protein eluting in the solvent front (C). SDS-PAGE gel analysis of purified H_6-TcBDF3(32..148aa) and cleaved *Tc*BDF3(32..148aa); Lane 1: Colour protein standard, lane 3, to 6: IMAC 1 fraction 3 to 5, lane 7: H_6-rTEV SEC fr15-19 (pTH26, Appendic A), lane 8 and 10: H_6-*Tc*BDF3(32..148aa) and H_6-rTEV incubated for 18 h, lane 9: Post cleavage IMAC 2 fr3 showing cleaved *Tc*BDF3(32..148aa), lane 11-15: Post cleavage IMAC 2 fr29 to 33 showing H_6-rTEV. (D). SEC profile of the cleaved protein. (E). SDS-PAGE gel analysis of purified and cleaved *Tc*BDF3(32..148aa); Lane 1: Colour protein standard, lane 2: IMAC 1 fraction 3 to 5, lane 3: Post cleavage IMAC 2 fr2-4, lane 4: post cleavage SEC fr 17-19.

2.2.3.3. Crystallisation trials for *Tc*BDF3

2.2.3.3.a. H_6-*Tc*BDF3 (32..148aa) crystallisation trials

Crystallisation of H_6-*Tc*BDF3(32..148aa) was attempted using commercial broad screen (Table 2.6) on the apo-protein. Over the course of this work, 5 distinct commercially available buffer screens were trialled using concentrations of H_6-*Tc*BDF3(32..148aa) ranging from 10 to 30 mg/mL (*ca*. 700 to 1900 µM). The crystallisation plates were stored either at RT (20–24 °C) or in a cold room (8–12 °C). None of the tested conditions yielded any visible crystals.

Addition of ligands might help stabilise the construct and induce crystallisation. The closely related *Trypanosoma brucei* BD, *Tb*DBF2, has been co-crystallised with I-BET151 (**5**) (PDB: 4PLK). The *Tc*BDF3(32..148aa) homology model was designed using the structure of the BCP PP-BRD20 from *Tb*427tmp (PDB: 5C8G) co-crystallised with dual BRD4/PLK1 inhibitor BI-2536 (**8**) (Section 2.3). Consequently, OXFBD02 (**7**), I-BET151 (**5**) and BI-2536 (**8**) were chosen to screened for co-crystallisation with *Tc*BDF3 (Figure 2.15). Co-crystallisation was attempted by using ligands dissolved in buffer (150 mM HEPES, pH 7.6). Although OXFBD02 (**7**) and I-BET151 (**5**) were insoluble in aqueous buffer, BI-2536 (**8**) was dissolved in buffer as a 25 mM stock (Table 2.6, condition 6). The BI-2536 (**8**) aqueous stock was mixed with protein to obtain a 1:25 protein:ligand mixture. A second attempt using high DMSO stocks (100 mM) of ligands instead of aqueous stock solutions was performed with 1:20 protein:ligand mixtures using the JCSG+ commercially available buffer screen. A large excess of ligand was chosen, as the ligands were predicted to be weak binders for *Tc*BDF3. To achieve these ratios, low concentrations of protein were employed, which might be more dilute that the ideal crystallising conditions.

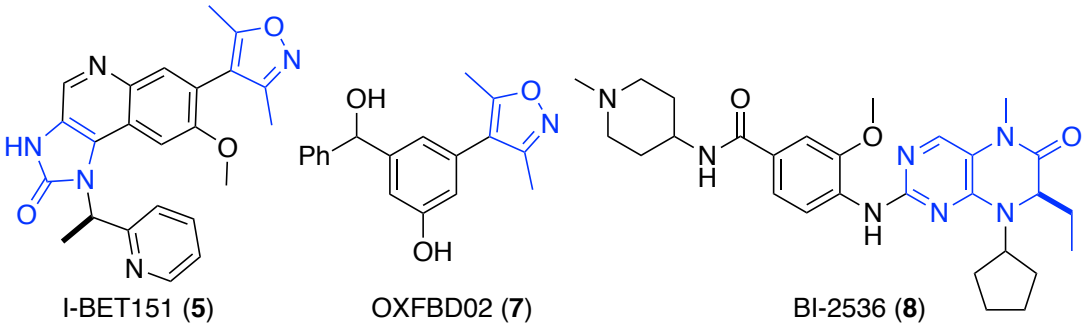


Figure 2.15: The structures of I-BET151 (**5**), OXFBD02 (**7**), and BI-2536 (**8**).

Table 2.6: Summary of the crystallisation conditions trialled for H_6-*Tc*BDF3 (32..248aa) and *Tc*BDF3 (32..248aa). The commercial screen compositions are details in Appendix D. The protein was buffer exchanged into a salt-free buffer (150 mM HEPES, pH 7.6) to limit formation of salt. * denotes protein used in 50 mM HEPES, 500 mM NaCl, pH 7.6 and ↑ denotes conditions that yielded crystals.

	Commercial screen	H_6-*Tc*BDF3 (32..148aa)	*Tc*BDF3 (32..148aa)	α-tubulin peptide	OXFBD02 (7)	I-BET151 (5)	BI-2536 (8)	Final % DMSO	RT plates	CR plates
1	JCSG+	–	4.9 mg/mL	–	–	–	–	–	1	–
2		30 mg/mL *	–	–	–	–	–	–	1	1
3		–	17.2 mg/mL	–	–	–	–	–	1	1
4		20.4 mg/mL	–	–	–	–	–	–	1	1
5		10.2 mg/mL	–	–	–	–	–	–	1	1
6		5.8 mg/mL	–	–	–	–	1:25	0%	1	–
7↑		1.6 mg/mL	–	–	1:20	–	–	5%	1	1
8↑		1.6 mg/mL	–	–	–	1:20	–	5%	1	–
9		1.6 mg/mL	–	–	–	–	1:20	5%	1	1
10		–	23.4 mg/mL	–	–	–	1:1.5	2%	1	1
11	PACT Premier	30 mg/mL *	–	–	–	–	–	–	1	1
12		–	17.2 mg/mL	–	–	–	–	–	1	1
13		20.4 mg/mL	–	–	–	–	–	–	1	1
14		10.2 mg/mL	–	–	–	–	–	–	1	1
15		5.8 mg/mL	–	–	–	–	1:24	–	1	–
16	Salt reaction broad screen	30 mg/mL *	–	–	–	–	–	–	1	1
17		–	17.2 mg/mL	–	–	–	–	–	1	1
18		10.2 mg/mL	–	–	–	–	–	–	1	1
19	Wizard	–	23.4 mg/mL	–	–	–	–	–	2	1
20	PEG Ion	30 mg/mL *	–	–	–	–	–	–	1	1
21		–	17.2 mg/mL	–	–	–	–	–	1	1
22		–	17.2 mg/mL	1:5					1	1
23		–	23.4 mg/mL	–	–	–	1:1.5	2%	1	1
24↑↑		–	23.4 mg/mL	–	–	–	–	–	1	–
25		–	12.5 mg/mL	–	–	–	1:3	2%	1	1
26	Crystal screen	20.4 mg/mL	–	–	–	–	–	–	1	1
27		10.2 mg/mL	–	–	–	–	–	–	1	1
28		5.8 mg/mL	–	–	–	–	1:24	–	1	–
29↑		–	23.4 mg/mL	–	–	–	1:1.5	2%	1	1
30		–	12.5 mg/mL	–	–	–	1:3	2%	1	1
31	Morpheous	–	23.4 mg/mL	–	–	–	–	–	1	–
32	Morpheus II	–	23.4 mg/mL	–	–	–	–	–	1	–

Co-crystallisation attempts with OXFBD02 (**7**) and I-BET151 (**5**) at room temperature each yielded a crystal (Table 2.6, conditions 7 and 8). Crystal 1 formed in 0.14 M $CaCl_2$, 0.1 M sodium acetate, 14% *v/v* ^{i}PrOH, pH 4.6 and crystal 2 formed in 0.2 M $(NH_4)_2HPO_4$, 0.1 M Tris, 50% *v/v* MPD, pH 8.5. Salt crystals often appear to have strong polarisation under polarised light and they have a "sharp" break under pressure rather than having a "soft" break. This is known as the crush test. The two crystals showed minor polarisation under

polarised light, which could be indicative of protein crystals. However, crystal 1 appeared sharp under the crush test, whereas crystal 2 appeared soft. The crystals were collected and cryoprotected by Dr Michael McDonough (University of Oxford) prior to X-Ray shooting (Biochemistry, University of Oxford). Crystal 1 had a strong diffraction pattern, which confirmed that it is a salt crystal. Crystal 2 did not show any diffraction. The absence of diffraction indicates that crystal 2 was not salt, and that it could be disordered protein or other buffer components.

2.2.3.3.b. *Tc*BDF3 (32..148aa) crystallisation trials

Protein tags such as His-tags are useful for purification purposes but are thought to prevent crystallisation, as they tend to be disordered.[74] Using the optimised expression and TEV cleavage conditions (section 2.2.3.1–2.2.3.3), the cleaved protein was produced in high quantity and purity. The crystallisation of apo *Tc*BDF3(32..148aa) was attempted using 6 different commercial broad screen (Table 2.6). Co-crystallisation was also attempted using high DMSO concentration of BI-2536 (**8**) (100 mM). This was conducted using 1:1.5 or 1:3 protein:ligand ratio, by keeping the final DMSO content down to 2% and maximising the protein concentrations (Table 2.6).

In total, three *Tc*BDF3 (32..148aa) crystals were produced at RT (Table 2.6, conditions 24 and 29). Two cube-shaped crystals (15 × 15 µm) were found for the apo-protein (crystal 3 and 4: condition 29, RT, well E5). These crystals formed in 0.2 M sodium malonate, 20% *v/v* 3350 PEG at pH 7.0. One rod-shaped crystal (155 × 15 µm) of *Tc*BDF3(32..148aa) formed in the presence of BI-2536 (**8**) (crystal 5: condition 24, RT, well E8) in 1.5 M NaCl with 10% *v/v* EtOH. Unfortunately, none of the 3 crystals diffracted.

Optimisation of conditions 24 and 29 was attempted by screening varying amounts of PEG and sodium malonate, or EtOH and NaCl, respectively. Two rod-shaped crystals formed from conditions based on the optimisation of condition 29: Crystal 6 formed in 2.0 M NaCl and 9% *v/v* EtOH, and crystal 7 formed in 1.7 M NaCl and 11% *v/v* EtOH.

While harvesting the crystals, crystal 6 was stuck to the bottom of the well and could not be collected. Crystal 7 was collected and cryoprotected, however, it did not afford a diffraction pattern.

Overall, a wide variety of crystallisation conditions were investigated for both H_6-*Tc*BDF3(32..148aa) and *Tc*BDF3(32..148aa), including in both the presence or absence of ligands. Out of the 6 harvestable crystals, 1 was identified as salt and the rest did not show diffraction. The absence of diffraction indicates that they were not salt, and that they could be potentially disordered protein (polymorphic aggregate) or other buffer components.

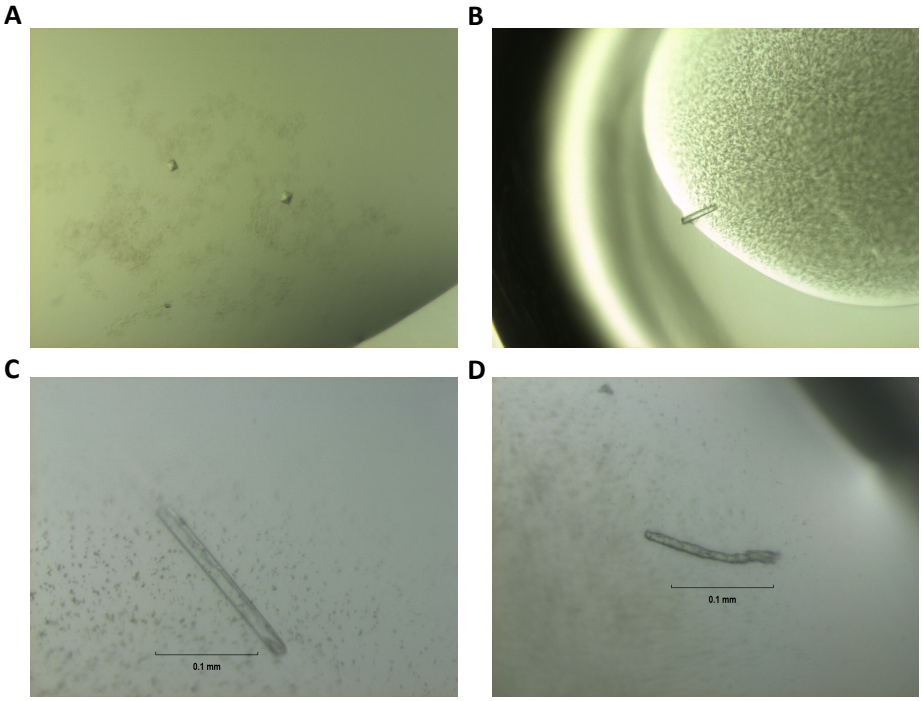


Figure 2.16: Photographs of the *Tc*BDF3 (32..148aa) crystals. (A) Crystals 3 and 4 (16× magnification) (B) Crystal 5 (5× magnification) (C) Crystal 6: Condition 29 (E8 well) optimised to 2.0 M NaCl and 9% *v/v* EtOH (D) Crystal 7: Condition 29 (E8 well) optimised to 1.7 M NaCl and 11% *v/v* EtOH.

2.2.4. N to F *Tc*BDFs mutants

Acetylated lysine residues are recognised by BDs through a hydrogen bond to their conserved asparagine (N) in the binding site (See Chapter 1). The single point mutation of this key N residue can provide a useful negative control protein construct. The mutated

protein is expected to adopt a structure close to the wild-type (WT) but not bind to BDs ligands. Retained binding to a mutated BD is an efficient way to rule out non-specific binders, for example lipophilic ligands, and can guide the choice of a favoured structure for further ligand or tool compound development. In human bromodomains, mutations to alanine (A) or phenylalanine (F) are commonly used.[75-76] For this work the N to F mutation was chosen, as it has been shown to completely eradicate binding to BD ligands, by increasing the bulk in the binding site as well as removing the hydrogen bond donor. Protein sequence analysis with a BLAST[77] search using CLC Main workbench v7.2 allowed for the identification of the asparagine residue thought to recognise KAc (Table 2.7).

Table 2.7: Summary of the position of the asparagine residue to be mutated to phenylalanine.

Protein construct	Position of the Asp residue	Mutation of interest
H_6-*Tc*BDF2(1..113aa)	86	N86F
H_6-*Tc*BDF3(32..148aa)	125	N125F
H_6-*Tc*BDF3(1..118aa)	91	N91F

2.2.5. DNA assembly for the *Tc*BDF2, 3 and 5-1 N to F mutants

A 2-fragment assembly of the linearised pNIC vector with the gene of interest obtained by custom DNA synthesis (Sigma Aldrich) was performed (Appendix A3). The synthesised genes fragments were dissolved in autoclaved MQ H_2O then combined with linearized pNIC vector using a Gibson assembly (NEBbuilder HiFi DNA assembly) kit. The assembly product was transformed into NEB 5α cells and grown on Kanamycin agar plates. Selected colonies were grown as previously described and the subsequent Miniprep DNA purification gave plasmid DNA at concentrations between 90 and 143 ng/µL. Sanger sequencing confirmed assembly.

2.2.5.1. Expression trials for the *Tc*BDF2, 3 and 5-1 N to F mutants

To determine if the H_6-*Tc*BDF2(N86F) and H_6-*Tc*BDF3(N125F) mutants can be overexpressed in *E. coli* in a similar manner to the wild type protein, an expression trial was performed. The transformation into BL21 (DE3) and the starter culture growth were performed as described previously. The 10 mL cultures were cooled to 16 or 18 °C, and

induced with 0.1, 0.5 or 1.0 mM IPTG. After 18 h, the cells were pelleted and lysed with Bugbuster® detergent and the cell pellet and soluble fractions were compared by SDS-PAGE gel analysis (Figure 2.17). No protein was observed in the supernatant for either protein, and only a small amount of overexpression is seen in the cell lysate suggesting that the *Tc*BDF2(N86F) and *Tc*BDF3(N125F) mutants were not soluble.

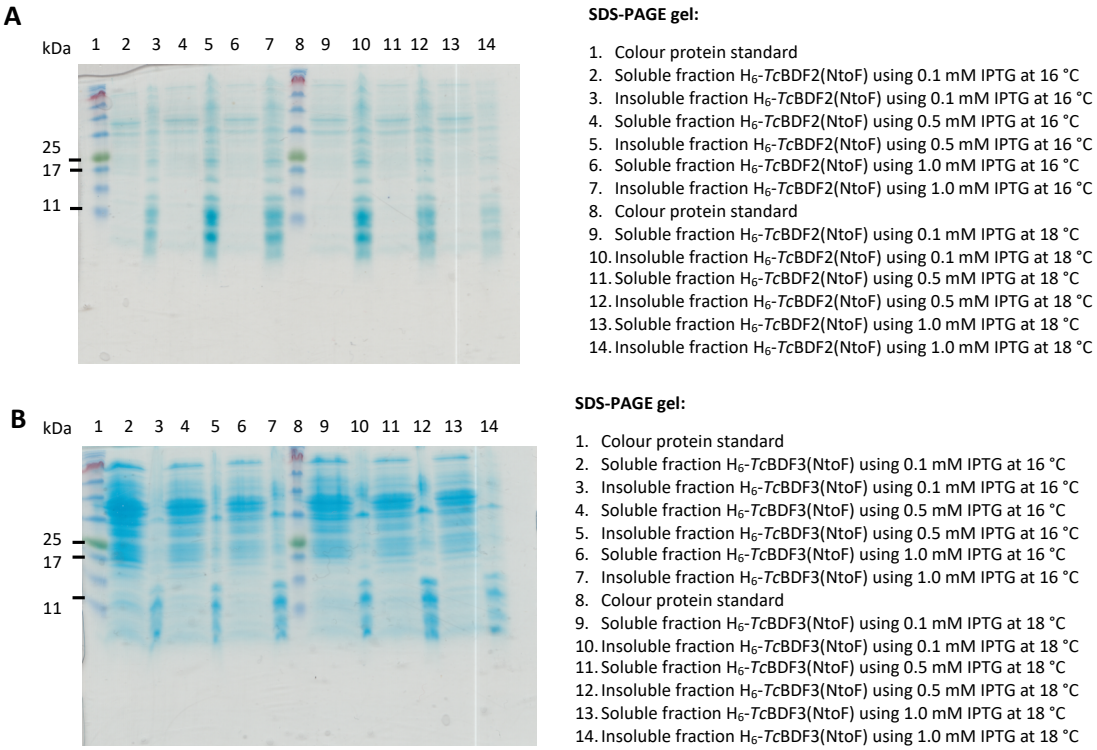


Figure 2.17: SDS-PAGE gel for the (A) H_6-*Tc*BDF2(N86F) and (B) H_6-*Tc*BDF3(N125F) expression trials, using 0.1, 0.5 or 1.0 mM IPTG, with post induction temperatures of 16 or 18 °C. No protein is seen the supernatant, and only a small amount of overexpression is seen in the cell lysate

To order to check if the results observed by SDS-PAGE were linked to lysis ability of Bugbuster®, the expression was repeated on a large scale and the cell lysate was lysed using sonication. This time the standard BD expression conditions (0.1 mM IPTG, at 16 °C) were used for the mutants H_6-*Tc*BDF2(N86F), H_6-*Tc*BDF3(N125F) and H_6-*Tc*BDF5-1(N91F). The wild-type H_6-*Tc*BDF3(32..148aa) was expressed using its corresponding optimised conditions (0.5 mM IPTG, at 16 °C) and was used as a control for the lysis. The same results were observed, showing that the NtoF *Tc*BDF mutants are not predicted to be soluble during a large scale expression and purification.

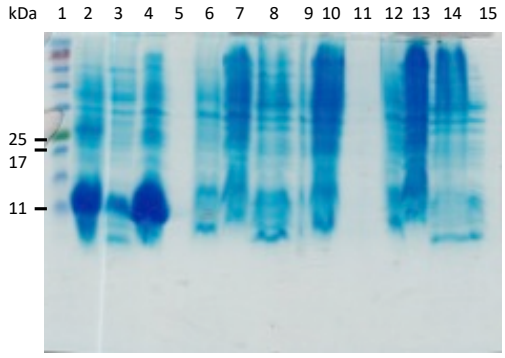


Figure 2.18: SDS-PAGE gel for the H_6-*Tc*BDF3 and the N to F mutants H_6-*Tc*BDF2(N86F), H_6-*Tc*BDF3(N125F) and H_6-*Tc*BDF5-1(N91F) expression trials. As shown previously, H_6-*Tc*BDF3(32..148aa) is overexpressed and can be found in the soluble fraction in addition to the whole cell and insoluble fraction. No overexpressed protein is observed for either of the 3 mutated proteins.

The N to F mutants of the studied proteins did not yield soluble protein constructs under the trialled conditions. Optimisation attempts were unsuccessful, which is thought to be inherent of the protein structure rather than DNA assembly. The single point mutation might affect the expression or the efficient protein folding. Mutation to alanine is also commonly employed to give negative control constructs for BDs. Therefore, N to A mutants could be explored as an alternative as this mutation might be less impactful on solubility.

2.2.6. Protein expression, purification, and crystallisation summary

The expression and purification of the novel H_6-*Tc*BDF2(1..113aa) and H_6-*Tc*BDF3(32..148aa) were improved to yield high amounts of protein. The cleavage of the His-tag of H_6-*Tc*BDF3(32..148aa) was optimised. The newly identified tandem *Tc*BDF5 BD was studied, and using small scale expression trials the H_6-*Tc*BDF5-1(1..118aa) and H_6-*Tc*BDF5-1(1..126aa) constructs were shown to produce soluble protein. The initial *Tc*BDF5-2 constructs did not express. The fusion of *Tc*BDF5-2 with a SUMO solubilising protein afforded overexpression but the resulting protein did not appear to be soluble by SDS-PAGE gel (Appendix A4). N to F mutants, designed to provide negative controls for any potential binders, did not yield soluble protein. Overall, the protein production of the *Tc*BDF2, 3 and 5-1 was optimised to provide large amounts (10–80 mg protein/batch) of soluble protein for biophysical assays.

Crystallisation screens of the *Tc*BDF proteins yielded 6 non-salt crystals for *Tc*BDF3. However, none of the crystals showed any diffraction pattern. BDs are part of BCPs, and the work detailed in this work suggest that the *Tc*BDF2, 3 and 5-1 proteins do not readily crystallise as truncated proteins.

2.3. Protein structure predictions

2.3.1. Protein folding prediction and circular dichroism

In the absence of crystal structures, the protein sequence can be analysed to predict whether it could produce secondary structure, such as α-helices and β-sheets. The structure of BD proteins is characterised by a left-handed bundle of four α-helices linked with two flexible loops. CLC Drug discovery workbench uses BLAST search to identify similarities to proteins from the protein data bank (PDB) to predict secondary structure. This analysis indicated the presence of the expected α-helices for the protein constructs used within this work (Figure 2.19).

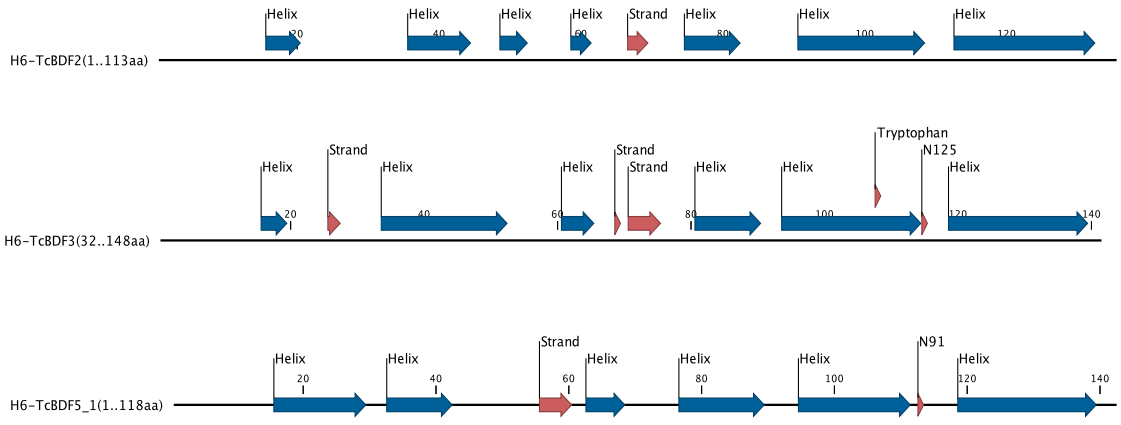


Figure 2.19: Prediction of the 2D structure of H_6-*Tc*BDF2(1..113aa), of H_6-*Tc*BDF3(32..148aa) and of H_6-*Tc*BDF5-1(1..118aa) constructs by sequence analysis generated using CLC Drug discovery workbench 7.

Circular dichroism was used to compare the absorption spectra of H_6-*Tc*BDF2, H_6-*Tc*BDF3 and H_6-*Tc*BDF5-1 with H_6-BRD4(1): the low angle observed at 220 nm is characteristic of α-helices (Figure 2.20), and provides evidence of folding.

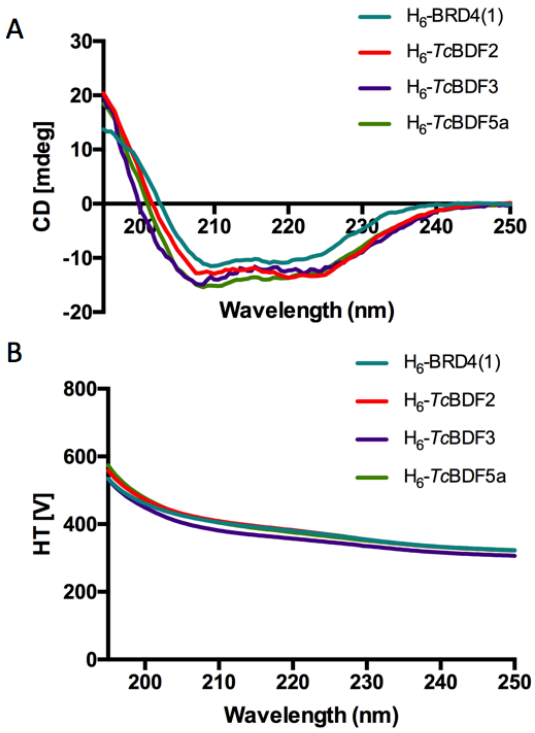


Figure 2.20: A. Overlay of the CD spectra of H_6-*Tc*BDF2(1..113aa), H_6-*Tc*BDF23(32..148aa), H_6-*Tc*BDF5-1(1..118aa) with H_6-BRD4(1) using C = 100 µg/mL. B. Overlay of the photomultiplier tube voltage spectra of H_6-*Tc*BDF2(1..113aa), H_6-*Tc*BDF23(32..148aa), H_6-*Tc*BDF5-1(1..118aa) with H_6-BRD4(1) using C = 100 µg/mL, showing HT ≤ 500 V. Image generated using Prism.

2.3.2. Homology models

Due to the lack of success with protein crystallisation, it was decided to create a homology model for *Tc*BDF2, *Tc*BDF3 and *Tc*BDF5-1. In collaboration with Joseph Bluck (University of Oxford), a BLAST search was performed on the H_6-*Tc*BDF3 (32..248aa) to identify a suitable template to create a homology model. The crystal structure of PP-BRD20 from *Tb*427tmp (temporary annotation for *Tb*BDF2), complexed with BI-2536 (**8**) (PDB: 5C8G), was found to be an ideal template, with a sequence identity of 65%. The structure was deposited in 2015 by Jiang *et al*. from the SGC, however, the associated paper has not been published. The homology model was built and molecular dynamics (MD) studies showed high flexibility of the ZA loop (Figure 2.21). Similarly, BLAST searches were performed on *Tc*BDF2 and *Tc*BDF5-1 to identify templates and homology models were built and validated with MD (Joseph Bluck).

A

Protein	Template Protein	PDB code	Sequence Identity
*Tc*BDF2	*Tb*BRD2	5CZG	65%
*Tc*BDF3	*Tb*427F	5C8G	65%
*Tc*BDF3	LdBPK(1)	5TCM	54%

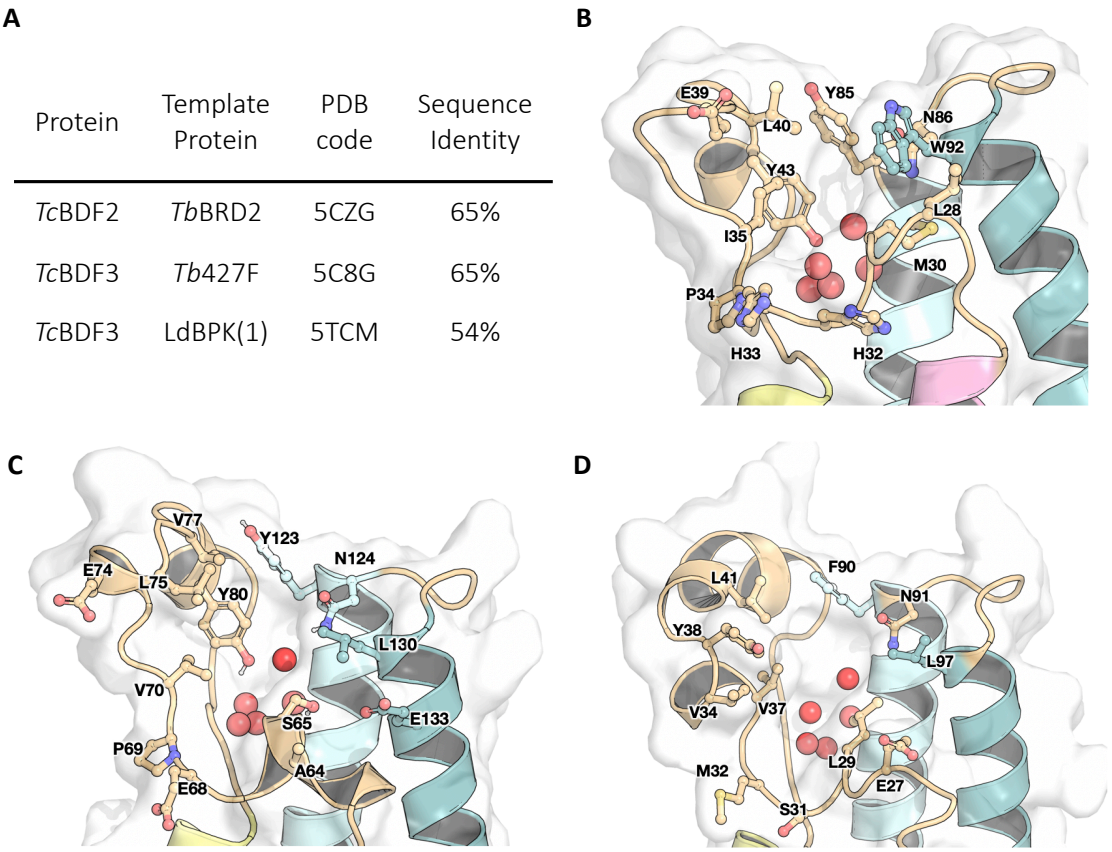


Figure 2.21: Homology model of *Tc*BDF2, *Tc*BDF3 and *Tc*BDF5-1. (A) Summary of the sequence identity of the *T. cruzi* BD proteins with their identified templates (B) The *Tc*BDF2 homology model showing the conserved asparagine residue N86. (C) The *Tc*BDF3 homology model showing the conserved asparagine residue N124. (D) The *Tc*BDF5-1 homology model showing the conserved asparagine residue N91. Image generated using Pymol by Joseph Bluck.

The homology models predict *Tc*BDF2, *Tc*BDF3 and *Tc*BDF5-1 to have four anti-parallel helices and two loops as expected. The conserved asparagine residue, expected to recognise KAc through hydrogen bonding is present in the three models (N86 for *Tc*BDF2; N124 for *Tc*BDF3; N91 for *Tc*BDF5-1). The MD simulations (Joseph Bluck and Andrew Boscek, University of Oxford) predict that the structures are highly flexible. Additionally, the ZA-channel is thought to be wider and shallower than the human BET BDs. For example, the *Tc*BDF3 L75 residue is analogous to L91 in BRD4(1); however, W81 is replaced by the much smaller alanine A64. Overall, the predicted high flexibility and wider ZA-channels suggest that the *Tc*BDF proteins might be challenging targets.

The *Tc*BDF2 and *Tc*BDF5-1 homology models were later suported by the structures deposited in the PDB by SGC Toronto (*Tc*BDF2 PDB: 6NIM, 6NP7; *Tc*BDF5-1 PDB: 6NEY).

2.4. Summary and conclusions

The optimisation of the expression and purification of truncated constructs of *T. cruzi* BD proteins has led to soluble folded proteins, as shown by circular dichroism. This work is unprecedented, and has produced proteins that can be purified from the soluble fraction of the cell lysate, rather than the inclusion bodies as described in the literature. Obtaining to the soluble and folded *T. cruzi* BD protein constructs opens the possibility of assay development and compound screening. Crystallisation screens of the *Tc*BDF proteins did not yield diffracting protein crystals, but homology models were produced to predict their structures. These models could be used as a predictive or rationalising tools as part of this work.

The work detailed in this Chapter is the foundation of the remainder of this thesis, as obtaining soluble protein constructs is an essential requirement for the development of biophysical assays.

Chapter 3. Initial hits and assays

3.1. Introduction and Aims

Having developed reproducible and high-yielding conditions for the expression and purification of H_6-*Tc*BDF2(1..113aa), H_6-*Tc*BDF3(32..248aa) and H_6-*Tc*BDF5-1(1..118aa) (Chapter 2), the next step was to screen a range of compounds to identify potential binders. The protein constructs are predicted to bind acetylated lysine residues on their endogenous binding partners. The putative endogenous binders, acetylated histone tails or acetylated α-tubulin for *Tc*BDF2 and *Tc*BDF3, respectively, are described in the literature[12,63] but are not fully characterised. In particular, there is no binding affinity available for the proposed protein-protein interactions (PPI).

To identify ligands for the *T. cruzi* BDs, we decided to adopt a fragment-based drug discovery (FBDD) strategy. FBDD is well described and offers the possibility of probing a higher amount of the chemical space compared to using large molecules.[78-79] Fragments, complying with the Rule of Three[80] (MW ≤ 300 Da, ≤ 3 hydrogen bond donors, ≤ 3 hydrogen bond acceptors, and logP ≤ 3), are designed and included in a fragment library. Such a library is screened using biophysical assays, and the identified hits are validated and then elaborated into lead compounds. This approach aims at identifying "atom efficient" binders that form high-quality interactions with the protein, as demonstrated by high ligand efficiency (LE).[81] The LE is defined as the Gibbs free energy ΔG divided by the numbers of non-hydrogen atoms. In particular, this approach has been successful for the study of BDs.[82-86]

The aim of the work detailed in this chapter was to synthesise a variety of KAc mimicking fragments, which are known to bind to human BD proteins, and screen them against the *T. cruzi* BD proteins. This work will determine whether the *T. cruzi* BDs bind KAc-mimicking motifs, as predicted, and identify fragments as starting point for further elaboration into higher affinity ligands.

3.2. Synthesis of ligands for screening against the *Tc*BDF BDs

A BLAST search of *Tc*BDF2, *Tc*BDF3 and *Tc*BDF5-1 has been used previously to determine to which human BD proteins they share the largest sequence identity. *Tc*BDF2 shows the highest sequence identity with BRPF1 and ATAD2, while *Tc*BDF3 shows the highest identity with CREBBP and the second BD of the BET BD family (Table 3.1).

Table 3.1: Summary of sequence identity of *Tc*BDF2 and *Tc*BDF3 to human BD proteins.

Protein construct	Human BD protein	sequence identity
H_6-*Tc*BDF2(1..113aa)	BRPF1	33%
	ATAD2	35%
H_6-*Tc*BDF3(32..148aa)	CREBBP	31%
	BRD2(2)	39%
	BRD3(2)	38%
	BRD4(2)	30%

The target fragments are shown in Figure 3.1. Structures containing the **FR1** fragment have been shown to bind BRPF1 and those containing **FR3** bind PCAF. **FR2-**, **FR4-** and **FR6-**based ligands show tight binding to BRD4. The dihydrobenzoxazepine motif **FR5** binds CREBBP. This section describes the synthesis of fragments **FR1**–**6**. The assessment of the binding of each of these fragments, alongside other larger bromodomains ligands, to the *T. cruzi* BDs will be described in Section 3.3. **FR1**–**6** were selected as they cover a range of different common human KAc mimics.

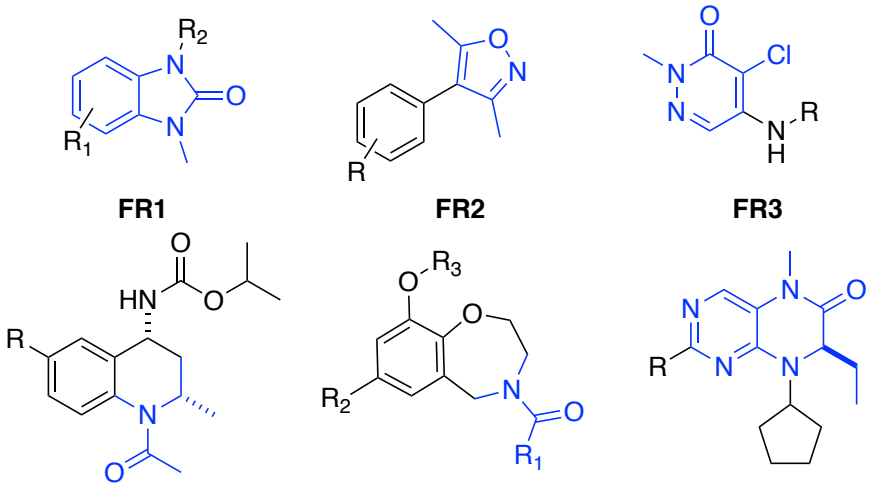


Figure 3.1: Structures of the acetylated-lysine mimicking fragments **FR1**–**6**. Their synthesis is described in Section 3.2. The KAc-mimicking moieties are highlighted in blue.

3.2.1. Synthesis of known acetyl lysine mimics

3.2.1.1. **FR1:** Benzimidazolones fragments

The predicted homology between the *Tc*BDF2 with BRPF1, has been used as starting point to identify suitable KAc mimics for binding to the *T. cruzi* BDs. The benzimidazolone core, **FR1**, is a small and versatile acetyl-lysine mimic. For example, the GSK6853 (**13**) ligand binds BRPF1 with a K_D of 10 nM,[43] and **14** binds TRIM24 and BRPF1 with a K_D of 137 nM and 222 nM, respectively (Figure 3.2).[44]

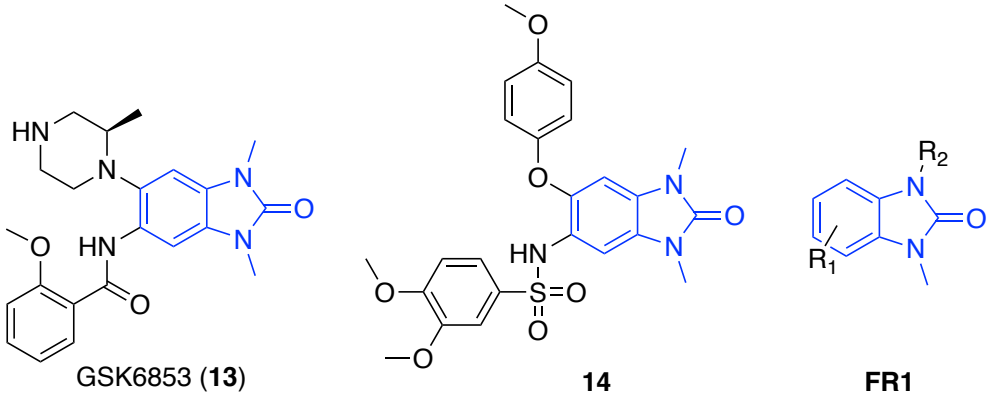
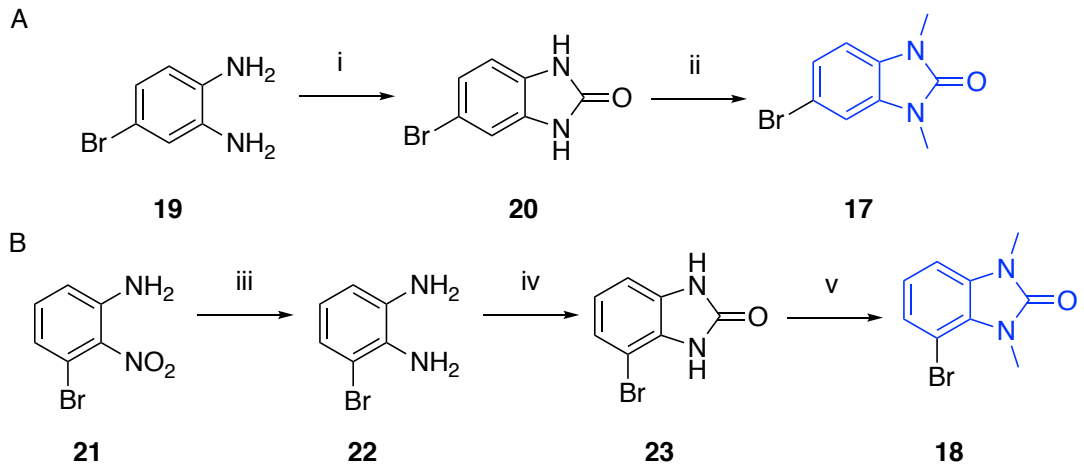


Figure 3.2: The structures of the BRPF1 ligands GSK6853 (**13**) and **14**, and the benzimidazolone fragment **FR1**. The KAc-mimicking moieties are highlighted in blue.

The general synthetic approach employed in this work is illustrated by the synthesis of **17** and **18** (Scheme 3.1), which were formed by benzimidazolone ring formation followed by methylation. The use of CDI to form the benzimidazolone ring gave high yields: 83 and 91% for **20** and **23**, respectively. Methylation with MeI afforded the target fragments **17** and **18** in 91 and 90% yield, respectively.

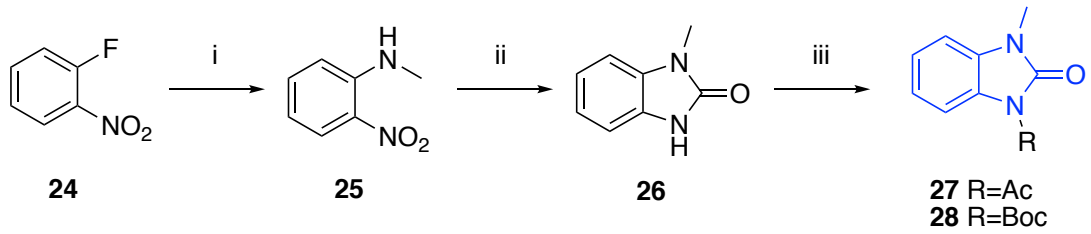


Scheme 3.1: Synthesis of the **17** and **18** fragments. *Reagents and conditions*: (A) (i) CDI, 1,4-dioxane, 40 °C, 1 h, 83%. (ii) MeI, K_2CO_3, DMF, RT, 6 h, 97%. (B) (iii) Fe, NH_4Cl, EtOH:H_20 3:1, 90 °C, 16 h, 74%. (iv) CDI, 1,4-dioxane, 60 °C, 2 h, 91%. (v) MeI, K_2CO_3, DMF, RT, 2 h, 90%. The KAc-mimicking moieties are highlighted in blue.

3.2.1.1.a. Unsymmetrical substitution of the benzimidazolone's nitrogen atoms

To gain insight into the binding interactions of the benzimidazolone fragments to the *T. cruzi* BDs, structure activity relationship (SAR) of the unsymmetrical substitution of the benzimidazolone's nitrogen atoms was explored.

First, an S_NAr reaction with methylamine on 2-fluoronitrobenzene (**24**) gave **25**, and was followed by reduction using palladium on carbon (Scheme 3.2). The resulting dianiline is prone to oxidation (especially on electron rich substrates). Therefore, hydrogenation with palladium on carbon was preferred to reduction with Zn and NH_4Cl in DMF, as it offers a shorter reaction time, and the work up is a filtration instead of an extraction step. Treatment of the dianiline **25** with triphosgene gave the benzimidazolone ring **26** in 36%, compared to 16% with CDI. The difference in yield arises predominantly from the ease of purification. The by-product formed during reaction with triphosgene is phosgene, which can react further, and chlorine (Appendix C). Remaining phosgene and trisphogene is quenched by addition with water and base, to give CO_2 and additional chlorine ions. The imidazole formed as a by-product when using CDI is removed by further washing and recrystalisation steps, leading to lower yields. Further reaction of **26** with acetic or Boc anhydrides gave **27** and **28** in 82 and 96% yields, respectively.



Scheme 3.2: Synthesis of benzimidazole fragments with different N-substituents. *Reagents and conditions*: (i) $MeNH_2 \cdot HCl$, DIPEA, NMP, 80 °C, 18 h, 26%. (ii) (a) H_2, Pd/C, EtOH, 1-3 h (b) Tripshogene, Et_3N, THF, 0 °C then RT, 16 h, 36%. (iii) R_2O, DMAP, Et_3N, CH_2Cl_2, RT, 3-5 h, R=Ac **27**: 82%, R=Boc **28**: 96%. The KAc-mimicking moieties are highlighted in blue.

3.2.1.1.b. Benzyl and pyridyl substitution of the benzimidazolone's nitrogen atoms

The BET ligand I-BET151 (**5**) (Figure 3.3) binds to human BDs through its dimethylisoxazole (DMI) headgroup. The crystal structure of I-BET151 (**5**) with *T. brucei* shows hydrogen bonding of the KAc-binding asparagine residue to the carbonyl of the benzimidazolone and to the nitrogen of the picolyl substituent (Figure 3.3).[71] Hence, the inverted binding of this ligand into the BD of this closely-related parasite poses the question of the binding mode it could adopt if binding to the *T. cruzi* BDs. Therefore, the comparison of the benzyl and 2-, 3- and 4-picolyl substituents on the benzimidazolone **FR1** core is of interest (Scheme 3.3). Synthesis of variations of the DMI core **FR2** are described in Section 3.2.1.2.

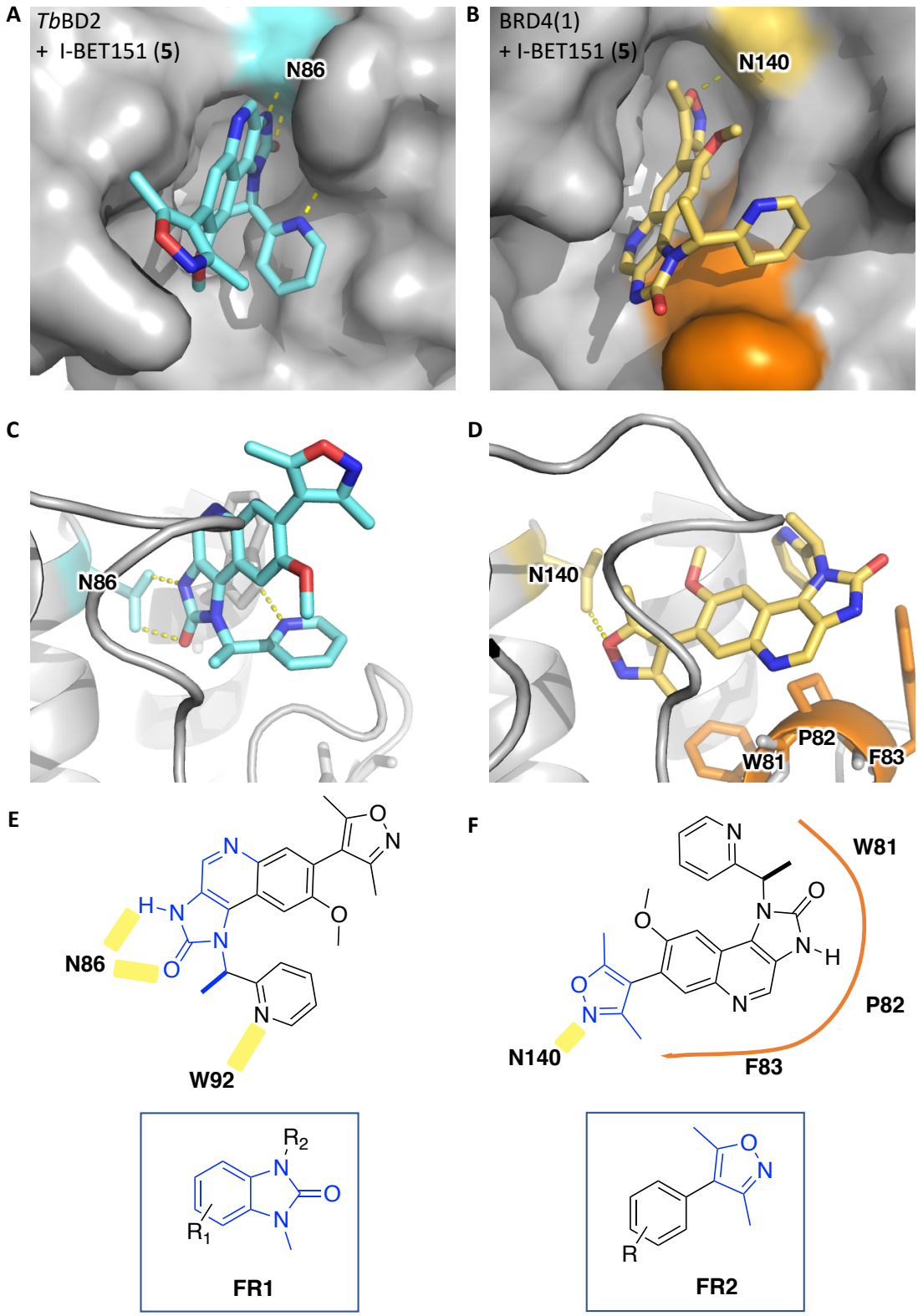
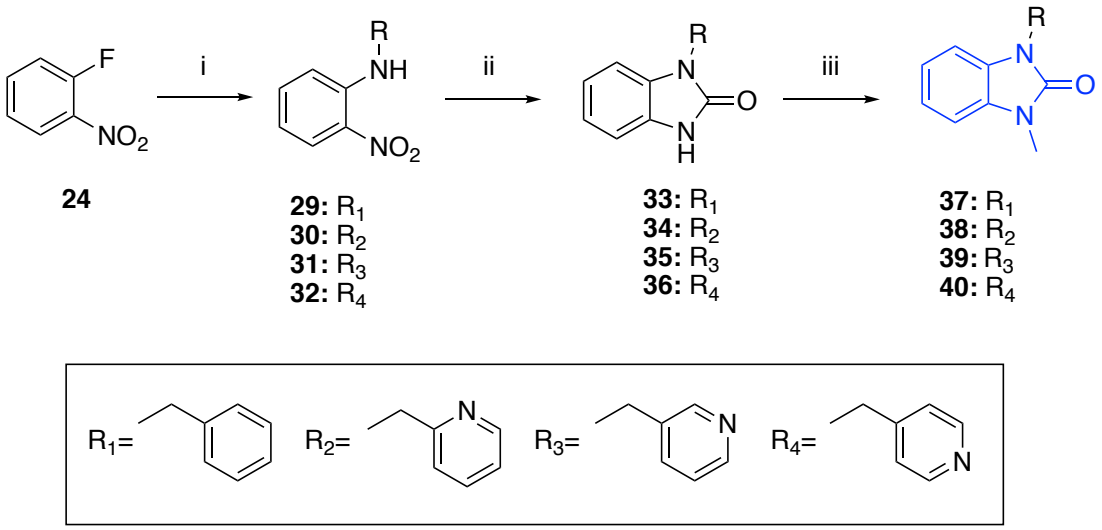


Figure 3.3: Comparison of the binding mode of I-BET151 (**5**) in *Trypanosoma brucei*, *Tb*BD2 (PDB: 4PLK, cyan), and in the human BET BD BRD4(1) (PDB: 3ZYU, yellow). In *Tb*BD2, the ligand binds to Asn86 through its benzimidazolone moiety (A, C, E) and shows additional hydrogen bonding to Trp92 through its picolyl moiety. In BRD4(1), I-BET151 (**5**) binds Asn140 through the DMI, while the rest of the molecule interacts with the WPF shelf. The acetyl lysine mimic are highlighted in blue, the key hydrogen bonding in yellow and the target fragments **FR1** and **FR2** are shown. Image generated using Pymol™ v 1.8.7.0.

Starting with 2-fluoronitrobenzene (**24**), the appropriate benzylic and picolyl amines were reacted using a S_NAr reaction, in NMP, to give the nitro-amine intermediates **29** to **32** (Scheme 3.3). Reduction with palladium on carbon was followed by ring closure using triphosgene to give **33** to **36**. Methylation of **33**-**36** gave the target fragments **37**-**40**. The synthesis of the picolyl analogues gave overall poorer yields than the benzyl counterpart. These differences arose mostly from challenging purifications due to higher numbers of side products during the methylation step. These side-product were thought to arise from methylation of the pyridine moiety. Despite low yields of 11–28%, the target fragments **37**-**40** were obtained in sufficient amounts for screening.

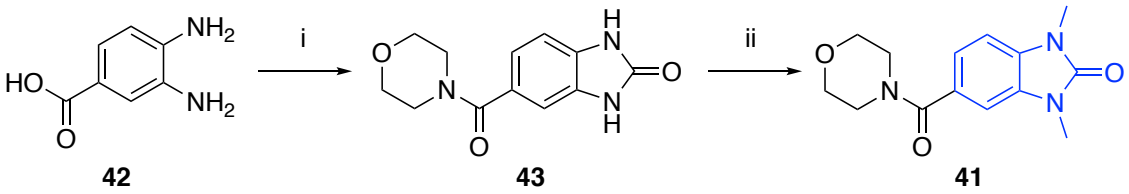


Scheme 3.3: Synthesis of benzyl and picolyl substituted benzimidazolones. *Reagents and conditions*: (i) RNH_2, NMP, 80 °C, 2–16 h; **29**: 99%, **30**: 98%, **31**: 75%, and **32**: 30%. (ii) (a) H_2, Pd/C, EtOH, 1–3 h (b) Triphosgene, Et_3N, THF, 0 °C then RT, 16 h; **33**: 46%, **34**: 76%, **35**: 25%, and **36**: 22%. (iii) MeI, K_2CO_3, DMF, RT, 18 h; **37**: 98%, **38**: 28%, **39**: 28%, **40**: 11%. The KAc-mimicking moieties are highlighted in blue.

3.2.1.1.c. Benzene substitution of benzimidazolone fragments

The SAR of the benzene substitution pattern was also investigated. A series of benzimidazolone fragments, containing amide and amine substituents around the phenyl ring, were targeted for synthesis. The morpholine substitution was chosen for its suitable solubility and its potential for hydrogen bonding. **41** was synthesised by reacting **42** with CDI, forming the closed benzimidazolone core and an imidazole activated amide (Scheme 3.4). Addition of morpholine to this intermediate formed **43.** Methylation of **43**

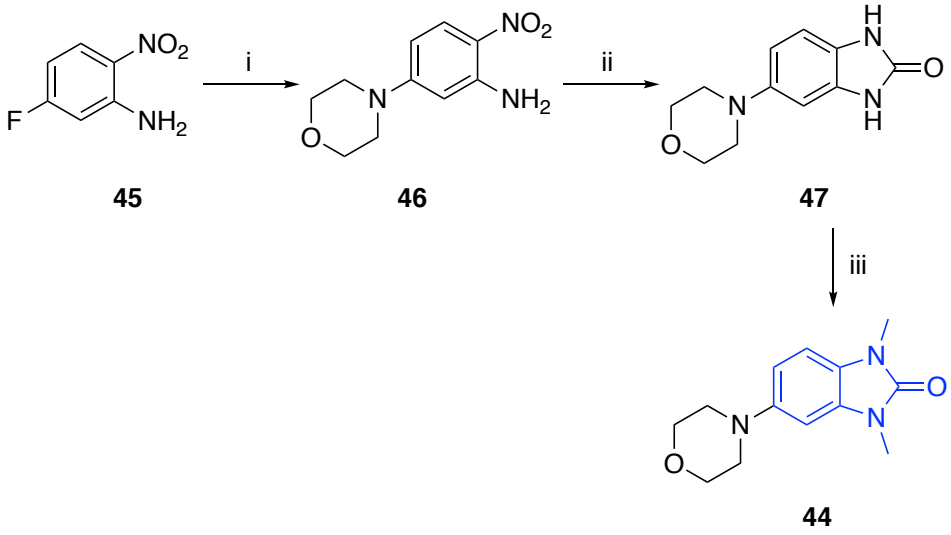
using methyl iodide and K_2CO_3 yielded **41**, which was purified using column chromatography followed by a hexane/water wash.



42 **43** **41**

Scheme 3.4: Synthesis of morpholine amide substituted benzimidazolone fragment **41**. *Reagents and conditions*: (i) CDI, THF or CH_2Cl_2, RT, 3h then morpholine, RT, 1 h, 24% (ii) MeI, K_2CO_3, DMF, 60 °C, 18 h, 36%. The KAc-mimicking moiety is highlighted in blue.

SAR studies are the most useful when matched pairs of compounds are compared. Hence, **44**, the corresponding amine version of **41** was also of interest (Scheme 3.5). The morpholine substituent was introduced using a S_NAr reaction with **45** to give **46**. Reduction of **46** with palladium on carbon and reaction with CDI afford **47** in 69% yield. Methylation using MeI and NaH gave the target **44** fragment in a 72% yield.



45 **46** **47** **44**

Scheme 3.5: Synthesis of morpholine substituted benzimidazolone fragment **44**. *Reagents and conditions*: (i) Morpholine, K_2CO_3, NMP, 100 °C, 3 h, 82% (ii) (a) H_2, Pd/C, EtOH, 3 h (b) CDI, CH_2Cl_2, RT, 69% (iii) MeI, NaH, DMF, 0 °C, 2 h, 72%. The KAc-mimicking moiety is highlighted in blue.

3.2.1.2. **FR2**: Synthesis of dimethylisoxazole ligands

Dimethyl isoxazoles (DMIs) can act as KAc mimetics and are found in several potent BD ligands.[37,46] In particular, the DMI motif is found in the BET BD ligands I-BET151 (**5**) and OXFBD02 (**7**) that bind to BRD4(1). Therefore, we decided to synthesise several fragments with the general structure of **FR2**.

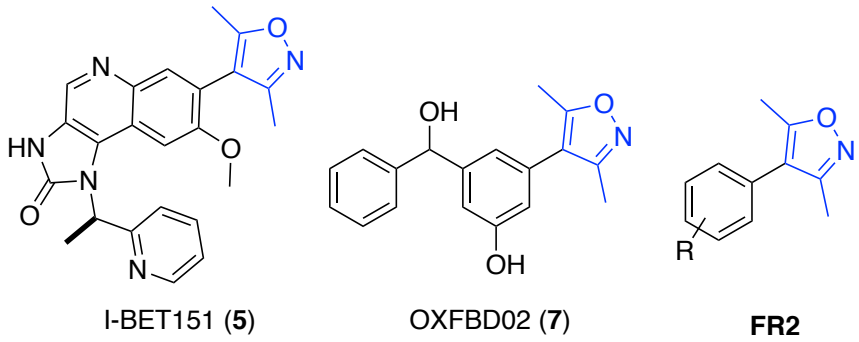
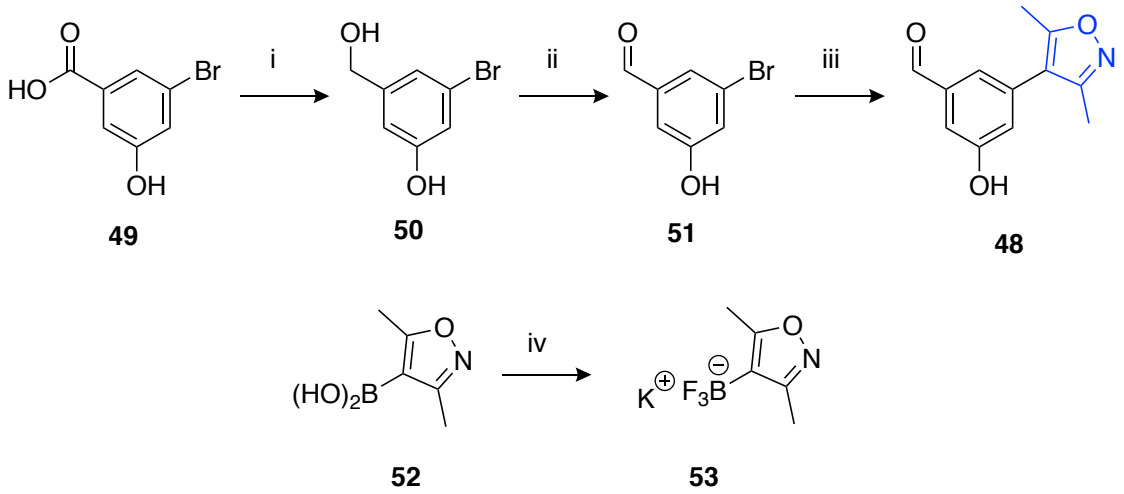


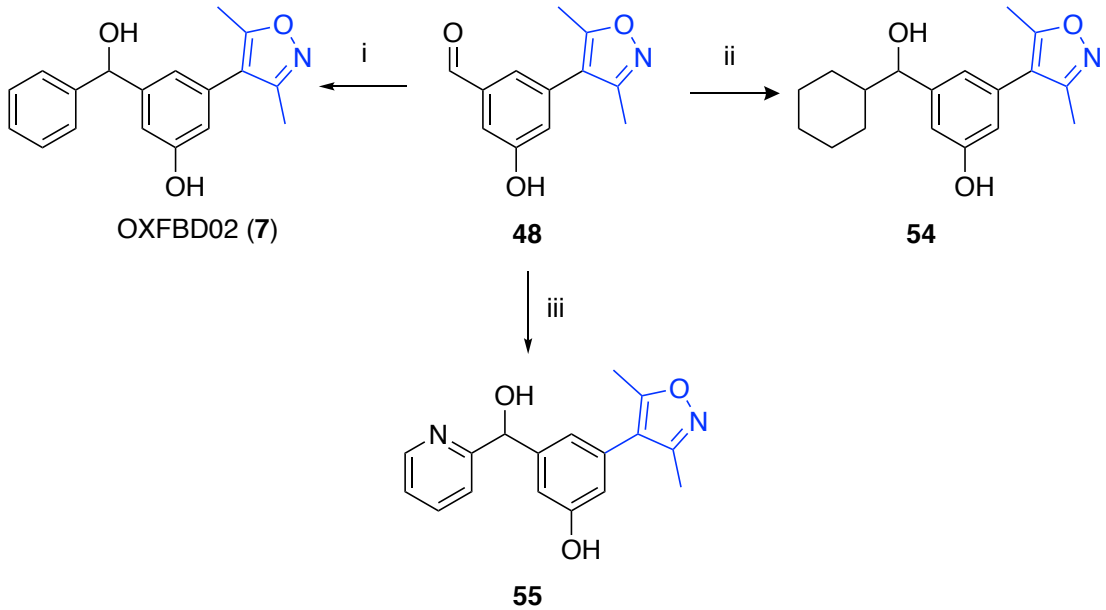
Figure 3.4: Structures of the DMI-based BET BD ligands I-BET151 (**5**) and OXFBD02 (**7**), and of the target **FR2** fragment. The KAc-mimicking moieties are highlighted in blue.

To synthesise the aldehyde **48**, borane reduction of the carboxylic acid **49** yielded **50**, which was then oxidized to the benzaldehyde **51** using MnO_2 in CH_2Cl_2 at 35 °C for 2 days (Scheme 3.6). Suzuki coupling of **51** with trifluoroborate **53** afforded the key intermediate **48**. Molander and coworkers[87] have shown that direct coupling of aryl bromides such as **51**, with boronic acids, such as **52,** is very low yielding due to rapid hydrodeborylation of the heterocycleboronic acids. Therefore, the trifluoroborate **53** was synthesised from the corresponding boronic acid **52** and used to ensure efficient coupling with **51**.[88]



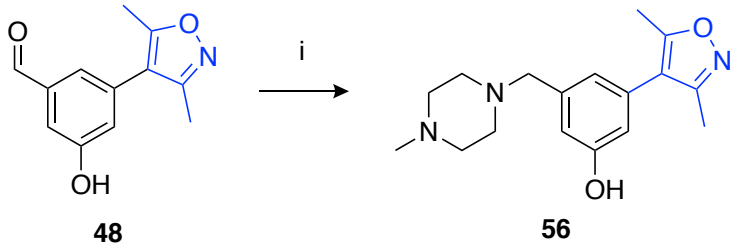
Scheme 3.6: Synthesis of the intermediate **48**. *Reagents and conditions*: (i) BH_3 (1M in THF), THF, 0 °C then RT, 20 h, 82% (ii) MnO_2, CH_2Cl_2, 35 °C, 2 d, 84% (iii) **53**, $Pd(OAc)_2$, RuPhos, Na_2CO_3, EtOH, reflux, 3.5 h, 65% (iv) KHF_2, MeOH, H_2O, 0 °C then RT, 15 mins, 80 %.

The DMI derivatives **7**, **54**, and **55** were obtained from a common intermediate **48** (Scheme 3.6) either by treatment with a Grignard reagent or by reductive amination. OXFBD02 (**7**) was synthesised by reaction of intermediate **48** with commercially available Grignard reagents (Scheme 3.7). The cyclohexyl Grignard reagents, used for the synthesis of **54**, was freshly made using magnesium turnings. Following the procedure by Jennings *et al.*,[58, 89] the 3-pyridyl Grignard reagent was prepared by magnesium/halogen exchange, rather than using commercially-available reagents, leading to **55** in 48% yield.



Scheme 3.7: Synthesis of BRD4(1) ligands **7**, **54** and **55** using Grignard reagents. *Reagents and conditions*: (i) PhMgBr (1 M in THF), THF, 0 °C then RT, 2 h, 48% (ii) Cyclohexanebromine, Mg, I_2, THF, RT, 2 h, 74%. (iii) (a) ^{i}PrMgCl (1.3 M in THF), 3-bromopyridine, THF, RT, 2 h (b) **48**, THF, RT, 4 h, 48%.

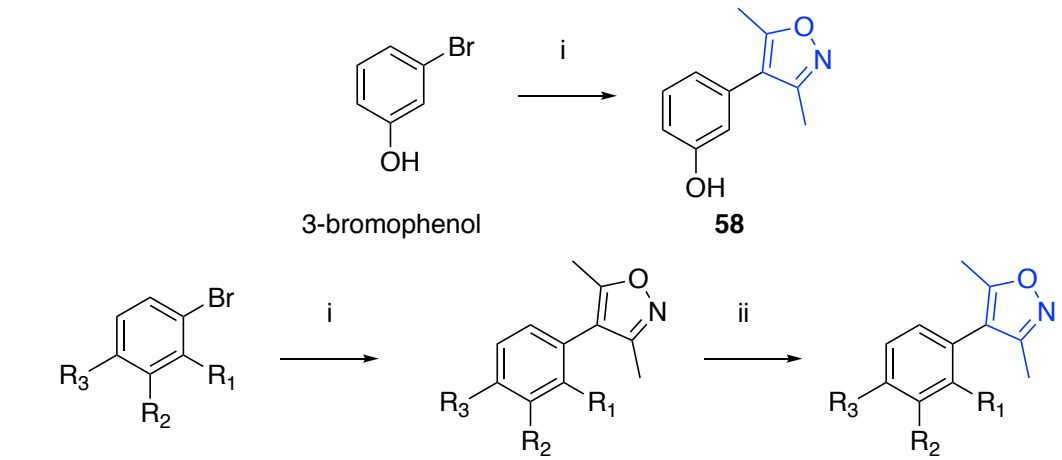
Aliphatic cyclic amine tails are also accessible though this intermediate. For example, **56** was obtained from intermediate **48** by reductive amination (Scheme 3.8).



48 **56**

Scheme 3.8: Synthesis of **56** from intermediate **48** by reductive amination. *Reagents and conditions*: (i) (a) 1-methylpiperazine, AcOH, EtOH RT, 30 mins (b) $NaBH_3CN$, RT, 20 h, 64%.

In addition to the synthesis of OXFBD02 (**7**) and close analogues (Scheme 3.7 and Scheme 3.8), smaller dimethylisoxazole fragments, with variations of the phenolic vector, were synthesised (Scheme3.9). Initially, a one-step procedure was explored for the direct coupling to 3-bromophenol and the DMI trifluoroborate reagent **53** to afford **58**. However, due to the long reaction time an alternative two-step procedure was investigated. This two-step procedure employed a cross coupling of 3-bromoanisole and **53** followed by demethylation with BBr_3 in dichloromethane. This improved the yield from 58% to 87% for the 3-(dimethylisoxazole)phenol (**58**) and this procedure also successfully yielded the 2- and 4- regiosomers **57** and **59**.



60: R_1=OCH_3, R_2=R_3=H
61: R_1=R_3=H, R_2=OCH_3,
62: R_1=R_2=H, R_3=OCH_3

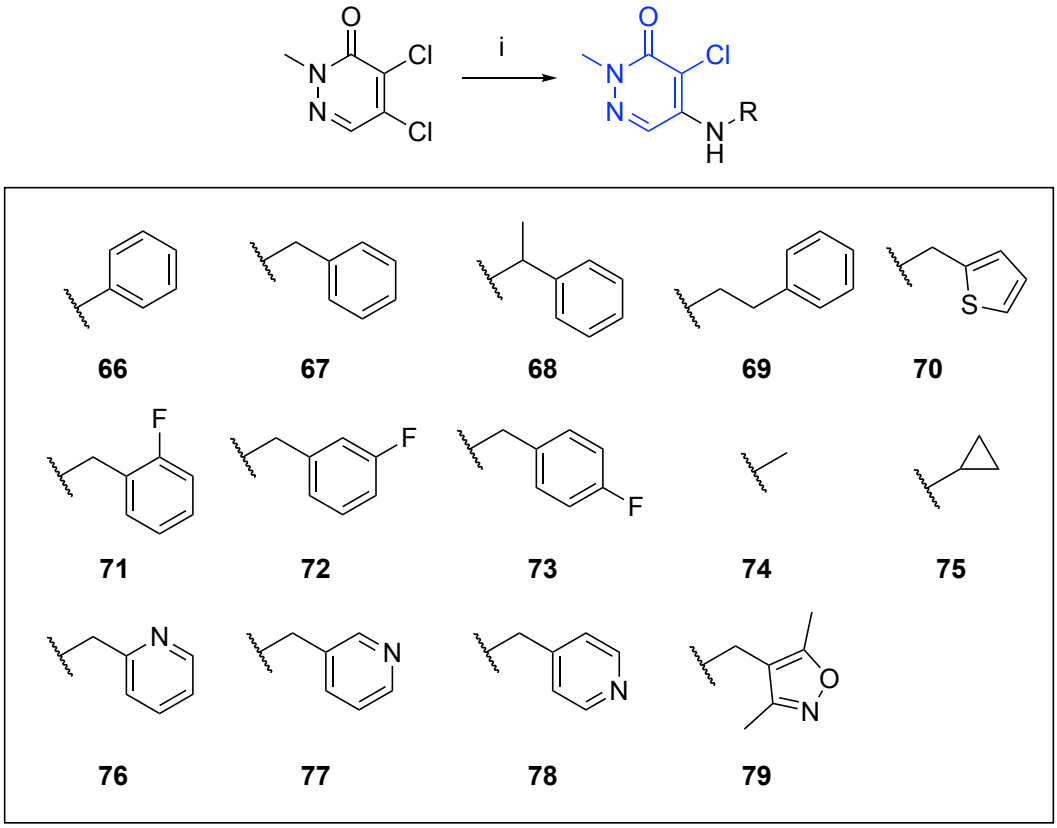
63: R_1=OCH_3, R_2=R_3=H
64: R_1=R_3=H, R_2=OCH_3
65: R_1=R_2=H, R_3=OCH_3

57: R_1=OH, R_2=R_3=H
58: R_1=R_3=H, R_2=OH
59: R_1=R_2=H, R_3=OH

Scheme3.9: Synthesis of phenolic dimethylisoxazole analogues *via* cross coupling and subsequent demethylation. *Reagents and conditions*: (i) **53**, $Pd(OAc)_2$, RuPhos, Na_2CO_3, EtOH, 80 °C: **58**: 48 h, 58%; **63**: 16 h, 42%; **64**: 16 h, 89%; **65**: 16 h, 72%. (ii) BBr_3 (1 M in CH_2Cl_2), CH_2Cl_2, 0 °C to RT, 30 mins: **57**: 80%; **58**: 98%; **59**: 99%.

3.2.1.3. **FR3:** Pyridazolone synthesis

The predicted homology between the *Tc*BDF3 with classical human BDs, such as BRD4 and CREBBP, and PCAF and BRPF1 for *Tc*BDF2 was used as a starting point to identify KAc mimics suitable for binding to the *T. cruzi* proteins. Therefore, variations of the recently published[42] pyridazolone motif that binds to PCAF were synthesised (Scheme 3.10). The S_NAr reaction occurred preferentially at the 5-chloro over the 4-chloro position for large amines, but no regioselectivity was observed when using methyl- or cyclopropylamine. In all cases it was possible to separate the regioisomers using silica chromatography.

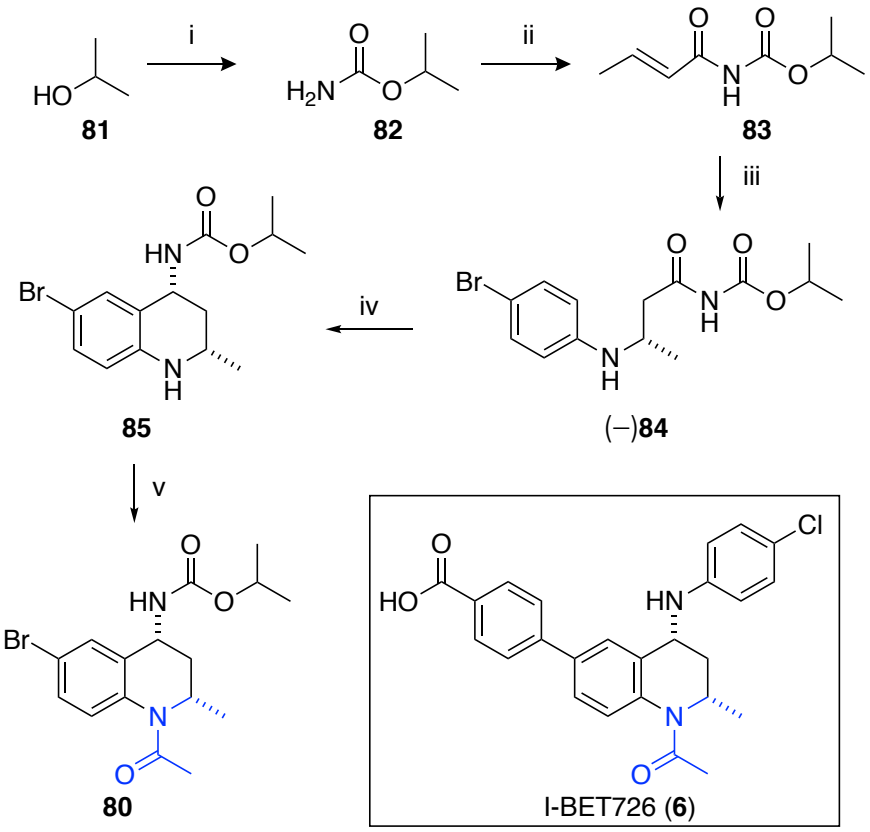


Scheme 3.10: S_NAr reactions to synthesise the pyridazolone analogues. *Reagents and conditions*: (i) RNH_2, NMP, 80 °C, 7–18 h. **66**: 60%, **67**: 80%, **68**: 67%, **69**: 84%, **70**: 63%, **71**: 75%, **72**: 70%, **73**: 62%,**74**: 16%, **75**: 9%, **76**: 46%, **77**: 23%, **78**: 46%, **79**: 55%.

3.2.1.4. **FR4**: Synthesis of the I-BET726 core

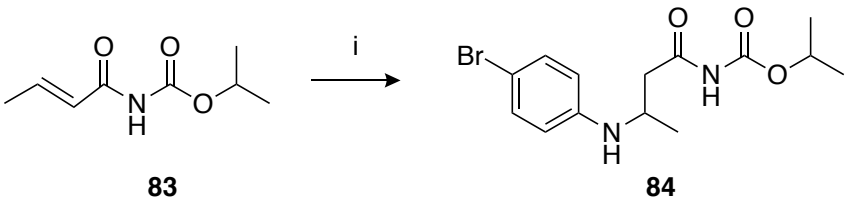
The tetrahydroquinoline fragment is the KAc-mimicking core seen in the BET ligand I-BET726 (**6**). **80** was synthesised following the procedure of Gosmini *et al.*[36] (Scheme 3.11). The isopropyl carbamate **82** was synthesised in 89% yield by treating sodium cyanate and

iso-propylalcohol **81** with TFA. **82** was then reacted with crotonyl chloride and LiO^{t}Bu to yield the key carbamate intermediate **83**. The palladium-catalysed aza-Michael type reaction with 4-bromoaniline gave (–)**84** in 64% yield. The double bond geometry on **83** is important as it determines the enantioselectivity of the stereogenic centre formed. Cyclisation of (–)**84** using $MgCl_2 \cdot 6H_2O$ under reductive conditions give the **85**. Low temperatures (< –15 °C) and short reaction times (< 1.5 h) are critical to achieve the cyclisation in a good yield and ensure a straight forward purification step.[36] Acetylation yielded the I-BET726 (**6**) core **80** in good yields.



Scheme 3.11: Synthesis of **80**. *Reagents and conditions*: (i) NaOCN, TFA, THF, toluene, 25 °C, 4 h, 89% (ii) chrotonyl chloride, LiO^{t}Bu (1 M in THF), THF, –20 °C, 4 h, 39% (iii) 4-Bromoaniline, (*R*)-BINAP(TfO)$_2$Pd(H$_2$O)$_2$, toluene, RT, 20 h, 80%. (iv) $NaBH_4$, $MgCl_2 \cdot 6H_2O$, EtOH, –5 to 0°C, 1h, 80% (v) AcCl, pyridine, CH_2Cl_2, 88%.

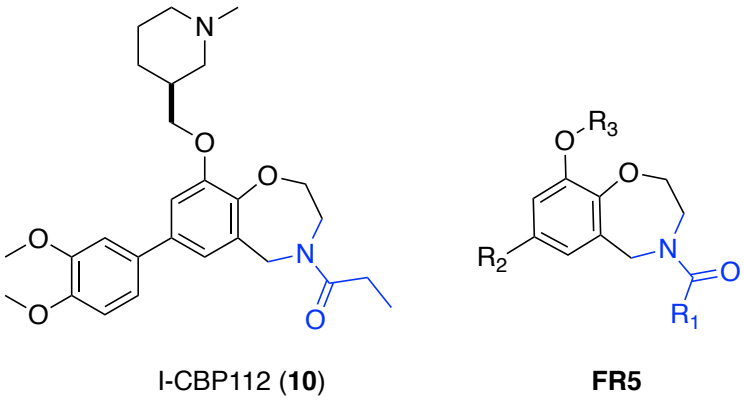
We also decided to synthesise racemic **84** to provide a reference for analytical chiral HPLC. Unfortunately, Pd(II) catalysis using $PdCl_2(MeCN)_2$ was unsuccessful, suggesting that the absence of the bulky chiral phosphine ligand is detrimental. An alternative route, based on a procedure from Roy *et al.*[90] employed $MnCl_2 \cdot 6H_2O$. Regular addition of small proportions of the weak Lewis acid over 8 days gave **84** in 53% yield (Scheme 3.12).



Scheme 3.12: Racemic synthesis of **84**. Reagents and conditions: (i) $MnCl_2 \cdot 6H_2O$, 4-bromoaniline, MeOH:H_2O 1:1, 65 °C, 8 d, 53%.

3.2.1.5. **FR5:** Dihydrobenzoxazepine core synthesis

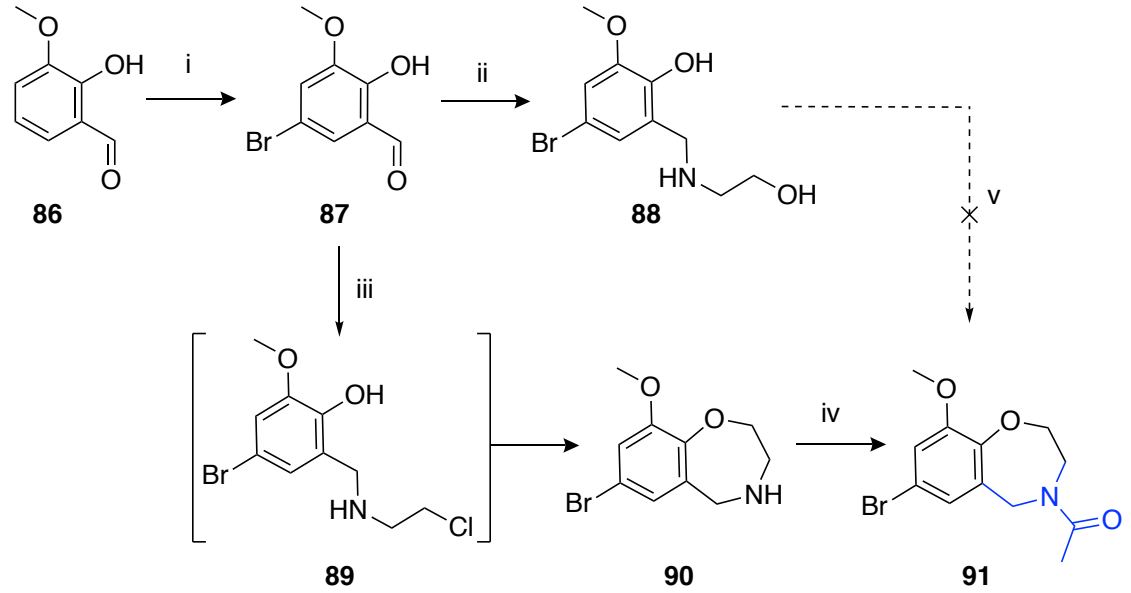
Picaud *et al.*[91] reported the structure of I-CBP112 (**10**) which has a dihydrobenzoxazepine core and binds to CREBBP with a dissociation constant (K_D) of 151 nM (Figure 3.5). Its structure comprises an acetylated cyclic amine fused to a benzene ring, similar to the I-BET726 (**6**) core. A synthetic route to I-CBP112 (**10**) had not been published when this work was started.



I-CBP112 (**10**) **FR5**

Figure 3.5: Structure of the CREBBP inhibitor **I-CBP112** and its core fragment **FR5**.

It was decided to synthesise **91**, a fragment of the general structure of **FR5** as it consisted of the KAc-mimicking core of I-CBP112 (**10**) and functional groups ready for further functionalisation. Bromination of *o*-vanillin (**86**) to afford **87**, followed by reductive amination with 2-aminoethanol gave **88** in good yields. However, acetylation of the NH group of **88** followed by cyclisation to **91** was not successful due to competing acetylation of the alcohol groups. The reductive amination of **87** with 2-chloroethyl amine was therefore explored as an alternative route to the desired dihydrobenzoxazepine core. **89** cyclised readily to **90**, which was then acetylated to give the dihydrobenzoxazepine core **91** in 60% yield (Scheme 3.13).



Scheme 3.13: The synthetic route to the acetylated dihydrobenzooxazepine core **91**. *Reagents and conditions*: (i) KBr, Br_2, acetic acid, 120 °C then RT, 16 h, 37% (ii) 2-Aminoethanol, DCE, 3 Å molecular sieves, RT, 1 h then $NaBH_4$, MeOH, 0 °C to RT, 4 h, 95% (iii) 2-Chloroethyl amine, DCE, 3 Å molecular sieves, RT, 3 h then NEt_3, RT, 1 h, then $NaBH(OAc)_3$, RT, 22 h, 55%. (iv) Ac_2O, K_2CO_3, CH_2Cl_2, RT, 4.5 h, 60%. (v) Ac_2O, K_2CO_3, CH_2Cl_2, RT.

3.2.1.6. **FR6**: Synthesis of BI-2536 (**8**) acetyl-lysine mimic

The PLK1 kinase inhibitor BI-2536 (**8**) binds BRD4(1) with a K_D of 37 nM.[38-39, 92] Its structure comprises a *N*-methylated dihydropteridinone core, which hydrogen bonds to Asp140 in BRD4(1). The amino-pyrimidine moiety of the molecule is the hinge-binding region that binds with the human PLK1 kinase (Figure 3.6). The *N*-methylated dihydropteridinone core **FR6** is an acetylated-lysine mimic and the synthesis of **(*R*)-92** is described below.

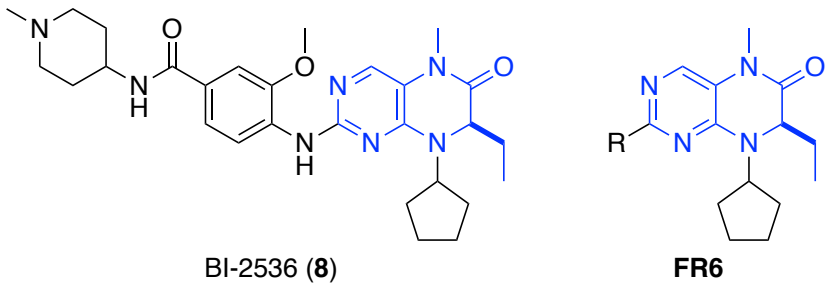
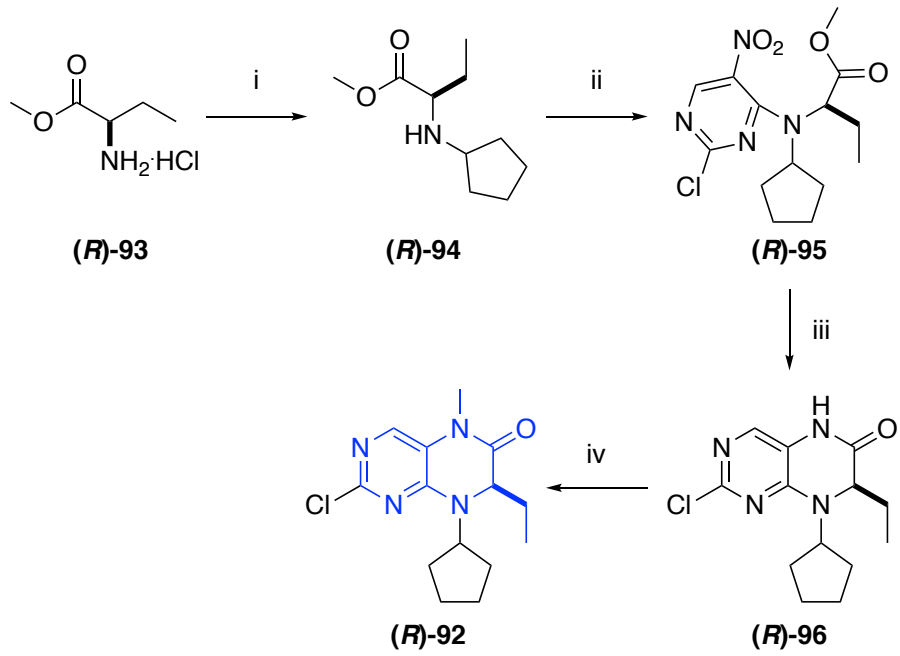


Figure 3.6: The structures of the dual PLK1/BRD4(1) BI-2536 (**8**) and its core **FR6**.

Synthesis of **(*R*)-92** started with a reductive amination of cyclopentanone with **(*R*)-93** (Scheme 3.14). The literature conditions, KOAc in DCE, gave low yields, as likely they are

not basic enough to neutralise the HCl salt of the amino acid. Therefore, trimethylamine in DCE was used to increase the pH above 7, allowing the reaction to proceed. In addition, $Ga(OTf)_3$ was used as a Lewis acid to promote to imine formation, as it has been shown to be effective when reacting hindered systems.[93] The following S_NAr reaction from **(*R*)-94** to **(*R*)-95** required optimisation, as the literature conditions, K_2CO_3 in acetone, gave low yields (Table 3.2). Using these original conditions, the reaction did not reach completion, as judged by TLC analysis, despite increasing the reaction time to 7 days. Increasing the temperature resulted in more side products, and did not result in full consumption of the starting materials. Despite these changes to the reaction conditions, the maximum yield obtained with K_2CO_3 in acetone was 19%. An alternative to the K_2CO_3/acetone conditions involved switching from an S_NAr to a Buchwald-Hartwig cross-coupling reaction using $Pd(OAc)_2$ and XPhos. Full consumption of the limiting reagent was achieved; however, more side products were formed than the first room temperature S_NAr conditions and only 29% yield was obtained for the cross-coupling route. Further literature search showed that Watts *et al*.[94] used $NaHCO_3$ in cyclohexane as an alternative to the original S_NAr conditions (Table 1). These conditions gave yields ranging between 51 and 80%.



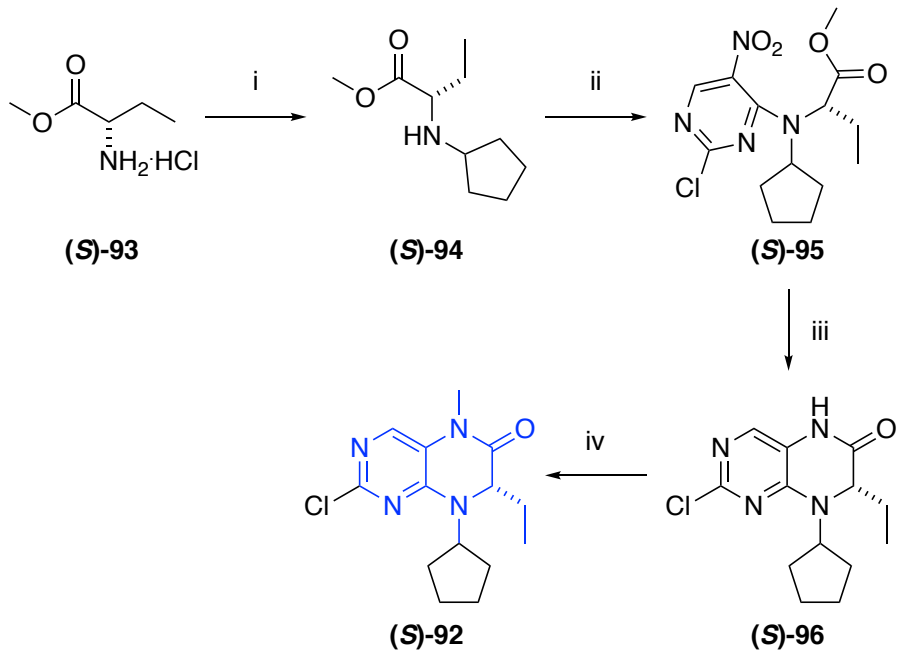
Scheme 3.14: Synthesis of the BI-2536 (**8**) core: **(*R*)-92**. *Reagents and conditions*: (i) NEt_3, $Ga(OTf)_3$, DCE, RT, 16 h, 65–78% (n=2). (ii) 2,4-dichloro-5-nitropyrimidine (a) K_2CO_3, acetone, RT, 1 d, 15-20% (n=5) or $Pd(OAc)_2$, XPhos, Cs_2CO_3, dioxane, 80 °C, 18 h, 20–29% (n=2) or $NaHCO_3$, cyclohexane, RT then 60 °C, 3 d, 51–80% (n=3). (iii) Glacial acetic acid, 70 °C then Fe (activated), 70 °C, 1h, then 100 °C, 4h, 61-81% (n=3). (iv) MeI, DMF then NaH, –10 °C then RT, 1h, 98% or K_2CO_3, MeI, DMF, RT, 3 d, 45%.

Table 3.2: Optimisation of S_NAr reaction to the key intermediate **(*R*)-95**.

Conditions	Temperature, time	Mechanism	Full reagent consumption (TLC)?	Isolated yield
K_2CO_3, acetone	0 °C then RT, 16 h to 7 d	S_NAr	No	19%
K_2CO_3, acetone	0 °C then 80 or 120 °C, 8 h to 2 d	S_NAr	No	< 15%
$Pd(OAc)_2$, XPhos, Cs_2CO_3, dioxane	80 °C, 18 h	Buchwald Hartwig coupling	Yes	29%
$NaHCO_3$, cyclohexane	RT then 60 °C, 3 d	S_NAr	No	51–80%

Reduction of **(*R*)-95**, using iron in glacial acetic acid, gave the cyclised product **(*R*)-96**. The initial reduction of the nitro group is rapid at 70 °C, however, the cyclisation requires 4 h at 100 °C to proceed to completion. Indeed, the yield was only 35% when left for 2 h at 100 °C, but reached between 61 and 81% when left for 4 h. Finally, methylation of **(*R*)-96** using MeI and NaH in DMF gave the truncated core of BI-2536 (**8**), **(*R*)-92**, in 90% yield. Addition of the MeI prior to the base prevents racemisation at the α-position of the lactam ring. Racemisation did occur if the intermediate was left in the presence of NaH before addition of MeI. Some racemisation (<10%) was observed when the NaH addition was conducted above –10 °C. Full racemisation was observed during repeats carried out by Michelle Keller (University of Oxford).[95] She subsequently optimised the methylation by using K_2CO_3 instead of NaH, and increasing the reaction times to 3 days. These conditions led to no detectable racemisation, so these conditions were used in following repeat and scale up reactions.

The (*S*)-enantiomer of **(*R*)-92** was synthesised using the same synthetic approach, as shown in Scheme 3.15.



Scheme 3.15: Synthesis of the **(*S*)-92** core. *Reagents and conditions*: (i) NEt_3, $Ga(OTf)_3$, DCE, RT, 16 h, 92%. (ii) 2,4-dichloro-5-nitropyrimidine, $NaHCO_3$, cyclohexane, RT then 60 °C, 3 d, 74%. (iii) Glacial acetic acid, 70 °C then Fe (activated), 70 °C, 1h, then 100 °C, 4h, 72%. (iv) MeI, DMF then NaH, –10 °C then RT, 1h, 90%.

3.2.2. Synthesis conclusions

A variety of fragments have been synthesised. **FR1** to **FR6** cover a broad range of known acetylated-lysine mimics used in the inhibition of human BDs. The screening of the fragments described here will provide information on if the *T. cruzi* BD proteins bind KAc mimicking motifs as expected.

3.3. Screening of Fragment and known acetyl lysine mimic screen

3.3.1. Screening aims

The initial strategy to study the *T. cruzi* BDs was to determine if they bind acetylated-lysine mimicking fragments analogously to human BD proteins. This section describes the set up and use of biophysical assays to detect binding. Due to the lack of data on the affinity of binding of the *T. cruzi* BDs to their endogenous binders, thermal shift and WaterLOGSY were chosen as initial assays, since neither of these assays require the displacement of a known binder. The combination of these assays can detect compounds binding with a wide range of affinities. Screening focused on identifying binders to *Tc*BDF3, as a primary priority, due to its higher similarity to CREBBP and the BET BD proteins. Although, thermal shift assay set up and screening was done in parallel for *Tc*BDF2 and *Tc*BDF3, the discussion will focus on *Tc*BDF3. The screening results against *Tc*BDF2 and *Tc*BDF5-1 can be found in Appendices B–C.

3.3.2. Thermal shift

The thermal denaturation of a folded protein can be monitored using hydrophobic dyes, for example SYPRO Orange®. The dye's fluorescence is quenched in aqueous buffer, and is augmented when binding to hydrophobic protein regions revealed during protein denaturation. The transition of the dye's fluorescence can be plotted and the point at which the gradient is steepest is defined as the protein melting point, T_m. The change in melting points, ΔT_m, upon addition of compounds can be used to identify ligands that stabilise or destabilise the protein's structure. Binders are typically defined by $\Delta T_m > 2$ °C. The thermal shift assay does not require the displacement of a known ligand. The magnitude of the shift is a result of enthalpic and entropic contributions to binding, hence different affinities might produce identical ΔT_m values. As a result, while compounds can be compared within structural series, series cannot be compared between themselves. Additionally, hydrophobic compounds can bind to the SYPRO Orange® dye directly and interfere with the assay. All thermal stability shift assays were carried out in full collaboration with F. Woodhouse (University of Oxford).

Compounds were screened in triplicate, but significant discrepancies between runs were observed. Following these results, the screening conditions were optimised to minimise errors (Appendix B1). Despite using optimised assay conditions, the screening of a large diverse library of compound, showed overall discrepancies between repeated runs, and large error values within triplicates (Appendix B1). For example, the ΔT_m for OXFBD02 (**7**) against *Tc*BDF3 varied from 0.5 to 2.5 °C and in the cases of ΔT_m > 2 °C the error within the triplicate was within the range of the observed shift. Other examples including **97**–**99** also show large standard errors and differences between repeats.

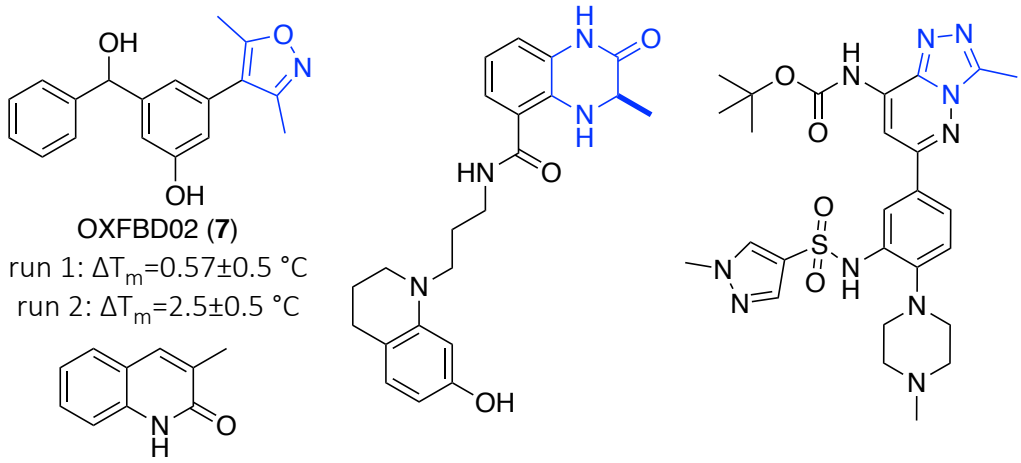


Figure 3.7: Selected ligands showing non-zero thermal shifts, ΔT_m, and their associated standard deviations on repeated runs. * denotes compounds synthesised F. Woodhouse (University of Oxford), ** denotes compounds synthesised by Dr T. Rooney (University of Oxford)and *** denotes compounds synthesised by Dr B. Wilson (University of Oxford).

Overall, compounds that showed ΔT_m > 2–3 °C were mostly compounds with poor solubility, some of which required careful mixing and sonication to solubilise in the assay buffer. Poorly soluble compounds may crash out in the assay buffer over time forming large lipophilic aggregates. The SYPRO Orange® dye is likely to bind these aggregates, hence, solubility issues are thought to interfere with the assay. Overall, the thermal shift data was not considered significant enough to guide choices of acetyl-lysine mimics to optimise for the binding of the *T. cruzi* BD proteins. The observed melting of the protein is indicative of unfolding, which can provide complementary proof of folding to circular dichroism data. So, even if the assay gave ambiguous results when screening for compounds, it provided additional proof of folding for *Tc*BDF2 and *Tc*BDF3.

3.3.3. WaterLOGSY

The potential hits identified using thermal shift were further screened using a water-ligand observed *via* gradient spectroscopy (waterLOGSY) assay. WaterLOGSY can be used to detect binding of small molecules to purified proteins and is orthogonal to fluorescence-based assays. In waterLOGSY experiments magnetisation is transferred from the bulk water to the ligand. The assay relies on the changes in the nOe effect upon binding of compounds to protein, which results in the change from fast to slow tumbling. nOe is the change of spin population in non-irradiated nuclei caused by T_1 relaxation of radiated neighbouring nuclei by a dipole-dipole mechanism. nOe transferred directly from water protons to ligand is positive (due to fast molecular motion), whereas it is negative when is it transferred through the protein (due to slow molecular motion). Hence, ligands can be identified by the change of nOe sign in the waterLOGSY assay. This technique relies on fast ligand-protein exchange, in order for the nOe to be transferred to several ligands to be measurable.[47] WaterLOGSY is suitable to identify weak (µM to mM) ligands; however, it does not provide information on the binding affinity to the protein.

The KAc mimicking fragments **FR1**-**FR6** (Synthesis described in Section 3.2.1) were screened using waterLOGSY, alongside larger known human bromodomain ligands (synthesised by Tim Rooney, David Hewings,[96] Laura Jennings,[89] Fern Woodhouse,[97] Amy Scorah, or bought from commercial suppliers). First, a compound-only sample (without protein) was run to provide a water-supressed ^{1}H NMR spectrum. This was followed by running a waterLOGSY ^{1}H sequence. Negative ligand resonances in the compound only waterLOGSY spectrum indicated suitable solubility. Such compounds were then screened in the presence of the protein. Positive peaks in the presence in the protein indicate binding, as shown for the OXFBD02 (**7**) (Figure 3.8). No changes in the peak sign upon addition protein indicate that no binding was detected under the screening conditions, as shown for the I-BET726 (**6**) core **80** (Figure 3.9). WaterLOGSY relies on fast exchange of the binding event, hence tight binding (nM) with slow off-rate might appear as false negative. This might be the case for large ligands but would be surprising for fragments (MW < 300).

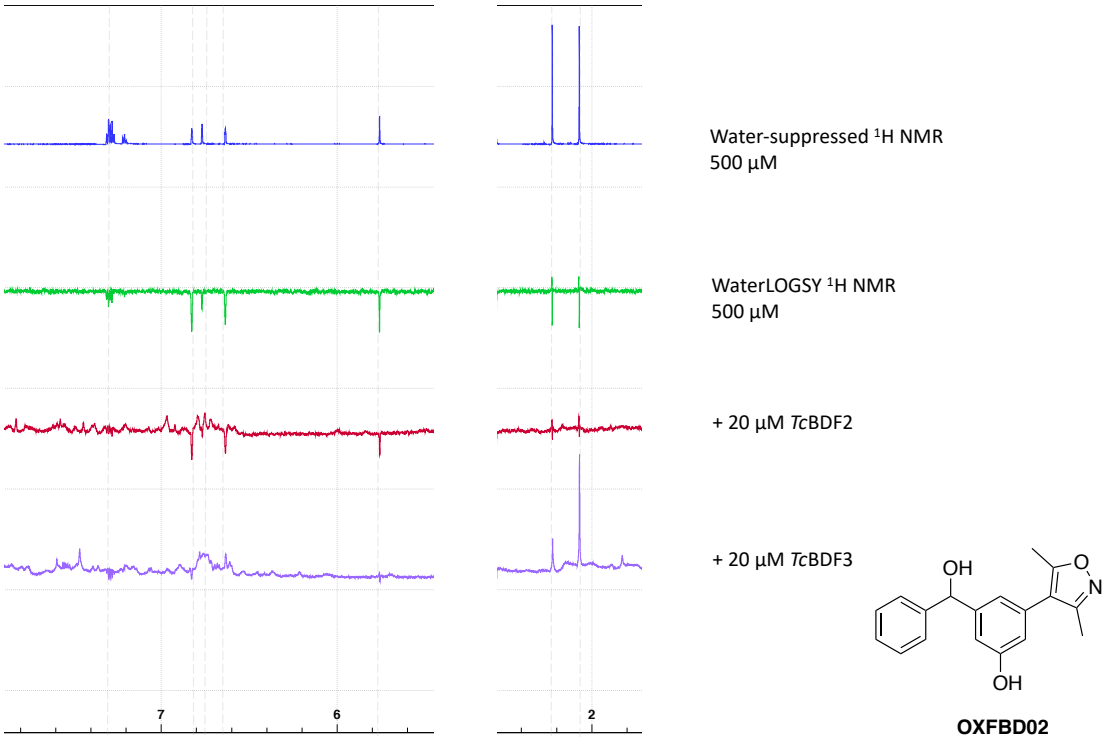


Figure 3.8: Partial spectra stack showing the WaterLOGSY signal for OXFBD02 (**7**). The water-supressed ^{1}H spectrum is shown in **blue**, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation. The waterLOGSY spectrum in the presence of *Tc*BDF2 (red) and *Tc*BDF3 (purple) display inversion of the ligand peaks, indicating binding.

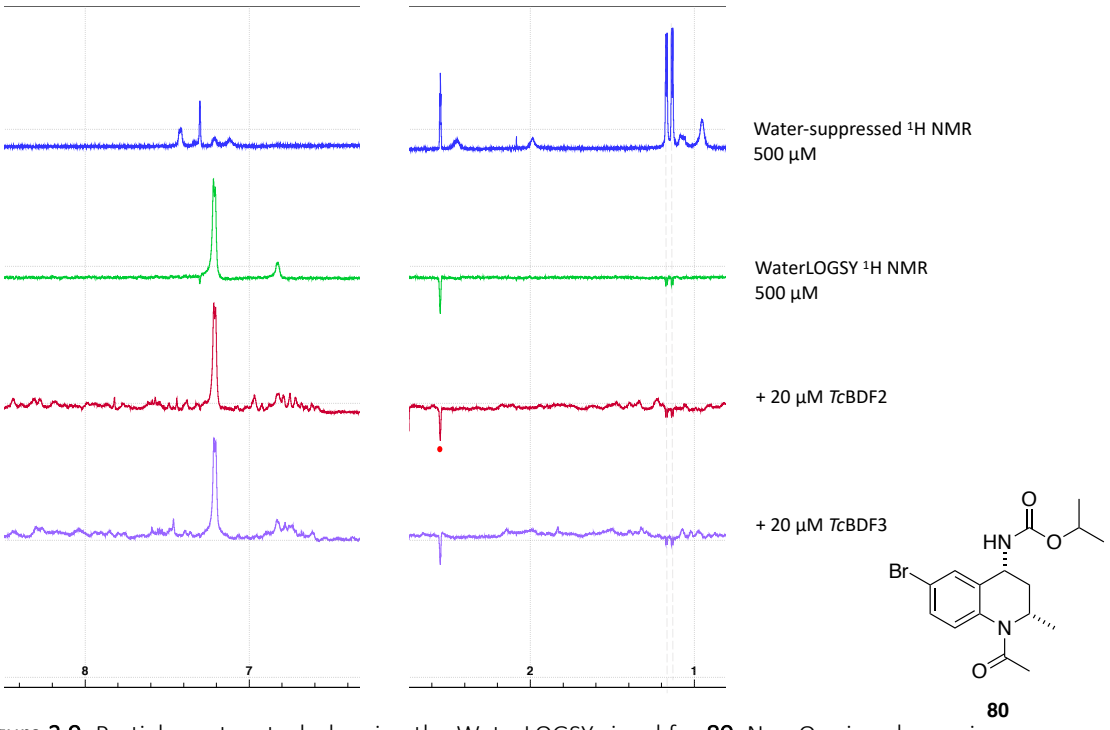


Figure 3.9: Partial spectra stack showing the WaterLOGSY signal for **80**. No nOe sign change is seen upon addition of either protein hence, **80** does not show detectable binding to *Tc*BDF3 or *Tc*BDF2. Same colour scheme as Figure 3.8.

Positive peaks in the compound only waterLOGSY spectrum are indicative that the studied small molecules display large-molecule like NOE. Such compounds form aggregates in solution despite appearing as clear solutions in the NMR tubes. When positive peaks in the compound only waterLOGSY spectrum were observed, the concentration of the ligand was reduced by half and the new sample was run to check for solubility. This process was repeated until a suitable concentration was found. Compounds forming soluble aggregates at concentration at or below 125–50 μM (lowest detectable concentrations) could not be screened using waterLOGSY since binding to the protein cannot be distinguished from aggregation behaviours using this technique.

The selected compounds were screened using waterLOGSY. Results are summarised in Table 3.3–Table 3.6, for the spectra stacks see Appendix B2.

3.3.3.1. **FR1** Benzimidazolone derivatives against TcBDF3

The results for the benzimidazolone **FR1** head group (Table 3.3) show that two regio-isomers of the brominated core **17** and **18**, and **100** bind to *Tc*BDF3 (Figure 3.10). Unsymmetrical substitution of the di-amide for **37**–**40** resulted in the loss of solubility. Polar substitution on the benzyl ring, such as the morpholine group in **41** and **44**, the benzylic alcohol in **101**, or the carboxyl group **102** were not tolerated. Overall, the binding of the benzimidazolone **FR1** fragment to *Tc*BDF3 shows a tight SAR, with no modifications of the dimethyl pattern nor polar benzene substitution being tolerated. Critically, direct aryl substitution resulted in poor solubility, during synthesis and testing. Such analogues would be predicted to be unsuitable for waterLOGSY testing. Therefore, the benzimidazolone **FR1** KAc mimic does not seem to be well suited for binding to *Tc*BDF3 and was not considered further.

Table 3.3: WaterLOGSY results for the **FR1** benzimidazolone series indicated as heat map: red for binding, orange for ambiguous binding and green for non-binding compounds. High to low compound concentrations is colour-coded dark to light blue for soluble compounds and grey for insoluble compounds.

Compound	MW	LogD7.4	SFI	WL solubility	*Tc*BDF3
17	241.09	2.00	4.00	Soluble at 500 µM	Binds
18	241.09	2.00	4.00	Soluble at 500 µM	Binds
27	190.2	0.64	2.64	Soluble at 500 µM	No binding
28	248.28	2.49	4.49	Soluble at 500 µM	No binding
33	225.25	1.78	3.78	Soluble at 500 µM	No binding
34	224.26	2.91	4.91	Soluble at 500 µM	No binding
35	225.25	1.7	3.7	Soluble (*)	No binding
36	225.25	1.7	3.7	Soluble (*)	No binding
37	238.28	2.96	4.96	Not soluble to 125 µM	-
38	239.28	1.82	3.82	Soluble (*)	No binding
39	239.28	1.74	3.74	Soluble (*)	No binding
40	239.28	1.74	3.74	Soluble at 500 µM	Maybe binds
44	247.3	1.31	3.31	Soluble (*)	No binding
47	219.24	1.03	3.03	Soluble (*)	No binding
100	228.25	1.19	4.19	Soluble at 250 µM	Binds
101	242.28	1.31	4.31	Soluble at 250 µM	Maybe binds
102	206.2	-2.28	-0.28	Soluble (*)	No binding
103	164.16	0.38	2.38	Soluble (*)	No binding
104	238.29	2.88	5.88	Not soluble to 125 µM	-
105	346.41	-0.5	2.5	Soluble at 250 µM	No binding
106	388.44	0.93	3.93	Soluble at 250 µM	No binding
107	162.19	1.24	3.24	Soluble at 500 µM	No binding

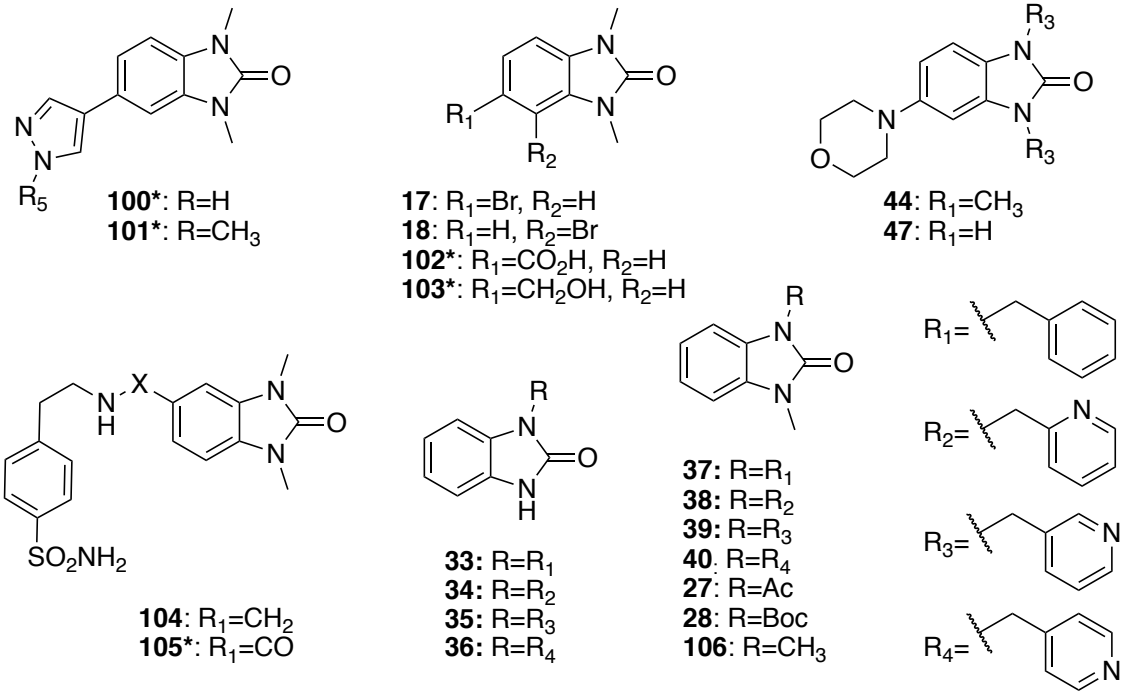


Figure 3.10: Structures of the **FR1** series compounds screened by waterLOGSY. * denotes compounds synthesised by A. Scorah (University of Oxford).

3.3.3.2. **FR2** series: DMIs bind to *Tc*BDF2, *Tc*BDF3 and *Tc*DF5-1

The DMI series looked promising by waterLOGSY: the known BRD4 ligands OXFBD02 (**7**) and I-BET151 (**5**), and the three phenolic regioisomers **57**, **58** and **59** show binding (Table 3.4, Figure 3.11). **113** and **114** did not bind showing that modifications of the dimethyl pattern are not tolerated. Analogues of I-BET151 (**5**) were also tested (provided by Dr Philip Humphreys, GSK). The enantiomer of I-BET151 (**5**), **115**, provide a useful negative control. OXFBD02 (**7**), **57**, and **58** show binding to *Tc*BDF2 and *Tc*BDF5-1, in addition to *Tc*BDF3 (Appendix B2), showing that the DMI series is a promising acetyl-lysine mimic for the *T. cruzi* bromodomains.

Table 3.4: Overall table for WaterLOGSY results on the **FR2** series. Same colour scheme as Table 3.3. Same colour scheme as Table 3.3.

Compound	MW	LogD7.4	SFI	WL solubility	WL against T3
5	415.45	2.69	7.69	soluble at 500 µM	Binds
7	295.34	3.02	6.02	Soluble at 500 µM	Binds
54	301.39	3.41	5.41	Not soluble to x µM	-
57	189.21	2.01	4.01	Soluble at 250 µM	Binds
58	189.21	2.01	4.01	Soluble at 500 µM	Binds
59	189.21	2.01	4.01	Soluble at 250 µM	Binds
63	203.24	3.33	5.33	Not soluble to 62.5 µM	–
64	203.24	3.33	5.33	Soluble at 62.5 µM	Binds
65	203.24	3.33	5.33	Not soluble to 125 µM	–
108	322.36	3.63	5.62	Soluble at 125 µM	No binding
109	288.35	1.7	3.7	Soluble at 125 µM	Maybe binds
110	272.35	1.28	3.28	soluble at 500 µM	Binds
111	325.36	3.80	6.80	Soluble at 250 µM	Maybe binds
112	313.33	4.00	6.00	Soluble at 125 µM	Binds
113	311.33	2.65	4.65	Soluble at 125 µM	No binding
114	233.26	2.92	4.92	Soluble at 125 µM	No binding
115	415.45	2.69	7.69	Soluble at 125 µM	No binding
116	415.45	2.69	7.69	Not soluble to 125 µM	–
117	401.42	2.08	7.08	Not soluble to 125 µM	–
118	400.43	3.27	8.37	Soluble at 125 µM	Binds
119	404.40	3.34	8.34	Not soluble to 125 µM	–
120	141.12	1.06	2.06	Soluble at 250 µM	No binding
121	220.02	1.93	2.93	Soluble at 250 µM	No binding
122	234.05	2.07	3.05	Soluble at 250 µM	No binding
123	173.22	2.32	4.32	Not soluble to 125 µM	–
124	257.21	3.98	5.98	Not soluble to 62.5 µM	–
125	207.21	2.96	4.96	Not soluble to 62.5 µM	–

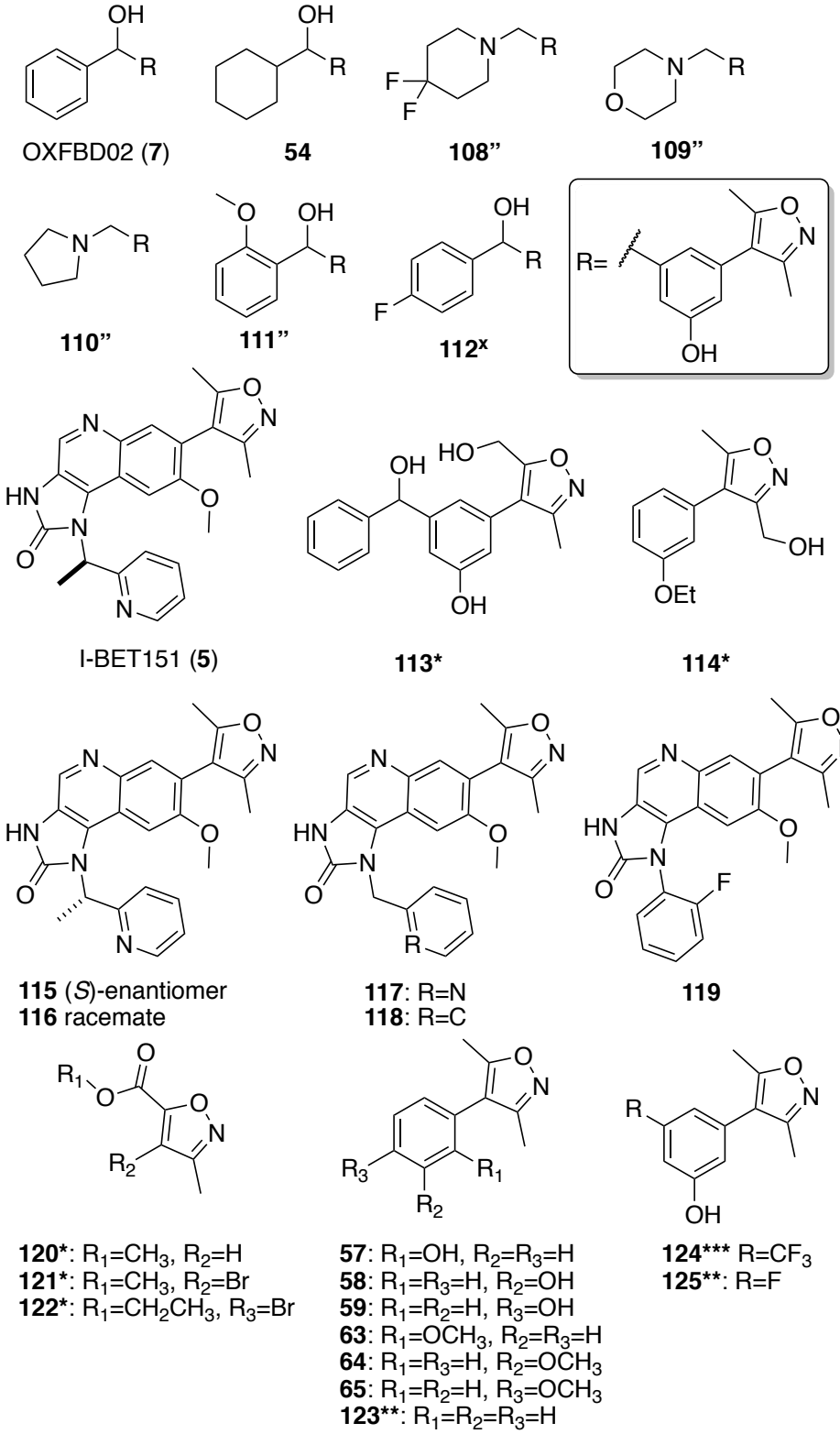


Figure 3.11: Structures of the FR2 series compounds screened by waterLOGSY. " denotes compounds synthesised by Dr L. Jennings (University of Oxford), * denotes compounds synthesised by Dr D. Hewings (University of Oxford), ** denotes compounds synthesised by F. Woodhouse (University of Oxford), *** denotes compounds synthesised by L. Tan (University of Oxford), and ˣ denotes compounds synthesised by R. Berkeley (University of Oxford). The I-BET151 (**5**) analogues were provided by Dr Philip Humphreys (GSK).

3.3.3.3. **FR3, FR4** and **FR5** do not bind to *Tc*BDF3

While **67** showed binding when it was originally screened; none of its analogues showed binding to *Tc*BDF3 (Table 3.5, Figure 3.12). Small modifications (fluorination, methylation) were not to tolerated. Due to the number of **FR3** fragments, this compound series was screened on separate days, and the repeat of the waterLOGSY spectra for **67** did not show binding, suggesting that it was a false positive. Overall, the assessment of the binding of **FR3**–**FR5** by waterLOGSY showed that theses KAc mimics do not bind to the *Tc*BDF3.

Table 3.5: Overall table for WaterLOGSY results on the **FR3**, **FR4** and **FR5** series. Same colour scheme as Table 3.3.

Compound	MW	LogD7.4	SFI	solubility	*Tc*BDF3
66	235.36	2.04	4.04	Soluble at 125 µM	No binding
67	249.7	1.6	3.6	Soluble at 250 µM	False positive
68	263.72	1.92	3.92	Soluble at 125 µM	No binding
69	263.73	1.89	3.89	Soluble at 125 µM	No binding
70	255.72	1.93	3.93	Soluble at 125 µM	No binding
71	267.68	1.72	3.72	Soluble at 125 µM	No binding
72	267.68	1.72	3.72	Soluble at 125 µM	No binding
73	267.68	1.72	3.72	Soluble at 125 µM	No binding
74	173.6	-0.12	0.88	Soluble at 125 µM	No binding
75	213.67	0.66	1.66	Soluble at 250 µM	–
76	250.69	0.47	2.47	Soluble at 125 µM	No binding
77	250.69	0.39	2.39	Soluble at 125 µM	No binding
79	263.73	1.89	3.89	Soluble at 125 µM	No binding
80	369.26	2.55	3.55	Soluble at 500 µM	No binding
85	327.22	2.92	3.92	Not soluble to 125 µM	–
90	258.12	1.78	2.78	Soluble at 125 µM	No binding
91	300.15	1.42	2.42	Soluble at 125 µM	No binding
127	290.36	1.78	2.78	Not soluble to 125 µM	–

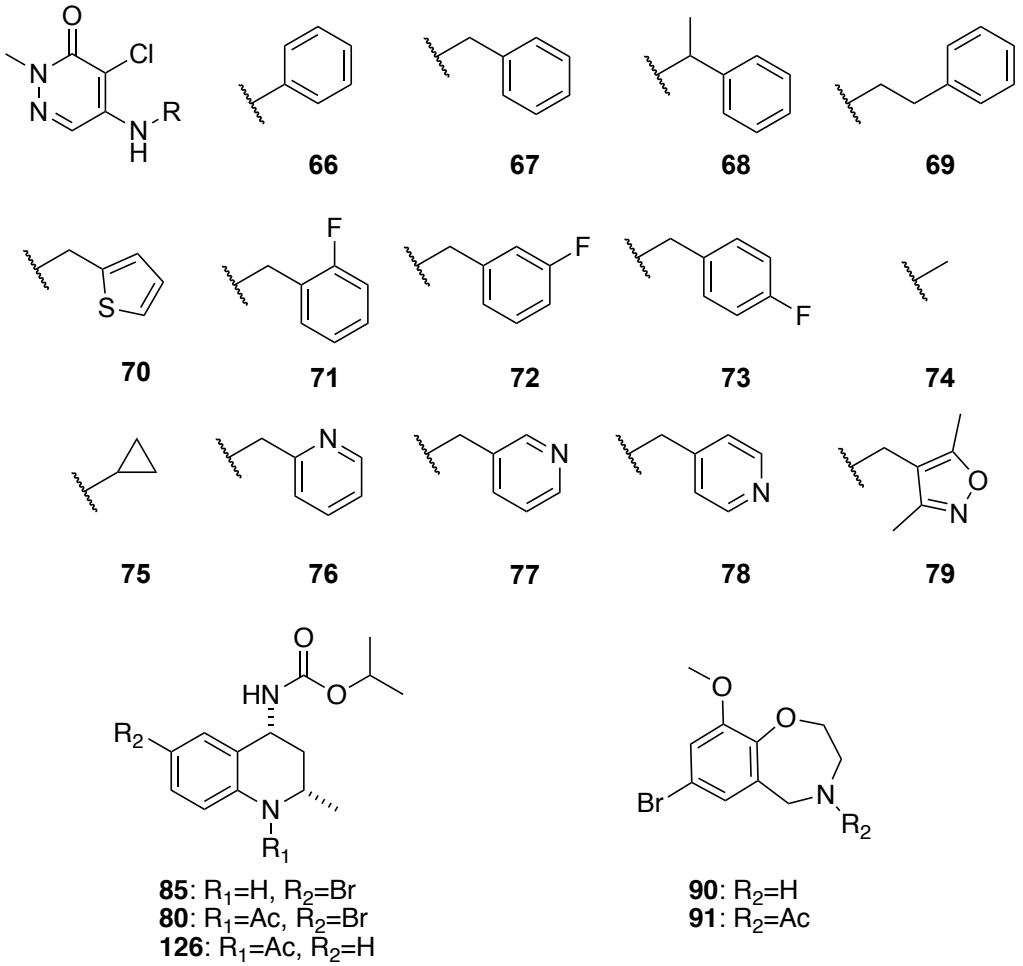


Figure 3.12: The **FR3**, **FR4** and **FR5** series tested by waterLOGSY.

3.3.3.4. The dihydropteridineone **FR6** fragments binds to *Tc*BDF3 by waterLOGSY

Compounds containing the dihydropteridineone KAc-mimic were screened by waterLOGSY (Table 3.6, Figure 3.13). BI-2536 (**8**) and its core fragment (*R*)-**92**, showed binding. The non-methylated core **96**, and the enantiomer of the core, (*S*)-**92**, did not bind to *Tc*BDF3, providing additional information about the predicted mode of binding. Based on these results, this KAc-mimic was predicted to have an analogous binding mode to BRD4(1): hydrogen bonding to the carbonyl group, places the methyl pointing into the "methyl binding" pocket, and the (*R*)-ethyl group to fit into a small hydrophobic pocket. Consequently, the dihydropteridineone series provide an attractive KAc mimic, with a well-established synthetic route,[94, 98] and with a negative control. Unfortunately, BI-6727 (**9**) displayed soluble aggregates down to 62.5 μM, and so was not suitable for testing by waterLOGSY.

Table 3.6: Overall table for WaterLOGSY results on the **FR6** dihydropteridineone KAc mimic. Same colour scheme as Table 3.3.

Compound	MW	LogD7.4	SFI	solubility	*Tc*BDF3
BI-2536 (**8**)	521.66	1.42	3.42	Soluble at 125 µM	Binds
(*R*)-**92**	294.78	2.34	3.34	Soluble at 50 µM	Binds
(S)-**92**	294.78	2.34	3.34	Soluble at 250 µM	No binding
(*R*)-**96**	280.75	2.28	3.28	Soluble at 50 µM	No binding
BI-6727 (**9**)	618.83	3.41	5.41	Not soluble to 62.5 µM	–

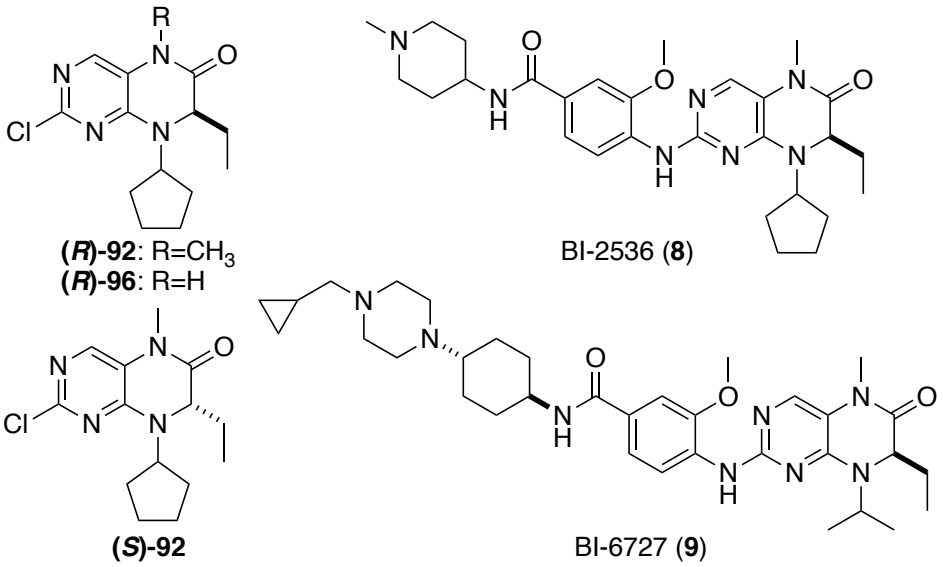


Figure 3.13: The dihydropteridineone **FR6** series tested by waterLOGSY.

3.3.3.5. ATAD2 binders against *Tc*BDF2

Due to the predicted similarity of *Tc*BDF2 to the human bromodomain ATAD2 we were also interested in screening novel ATAD2 binders synthesised by Dr Dan Foley (Nelson research group, University of Leeds). The compounds provided by the Nelson group included aliphatic bi- and tricyclic structures with free hydroxyl or amine groups. These highly polar compounds all exhibited high solubility and were found to be suitable for waterLOGSY screening. Unfortunately, no binding (no signal inversion) was observed for any of the 11 compounds tested (Figure 3.14 and Appendix B2).

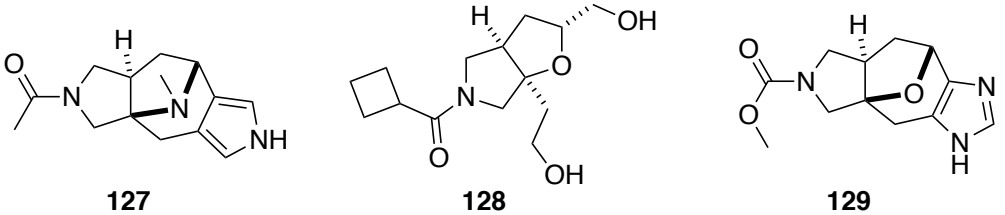


Figure 3.14: Representative examples of ATAD2 binders structures screened against *Tc*BDF2 by WaterLOGSY (synthesised by Dr Dan Foley, University of Leeds).

3.3.4. ^{1}H ligand observed protein titration

WaterLOGSY can detect binding of interactions with a wide range of affinities (µM to mM) but is not quantitative. Further ligand development requires ranking and prioritisation of the hits according to their affinities to the *T. cruzi* BDs. Therefore, quantification of the binding interactions was the next priority. Since the use of NMR-based techniques was successful to detect binding to the *T. cruzi* BDs, titration experiments monitored by NMR were attractive. The local chemical environment around a ligand in solution and a ligand bound to a protein are different, which can lead to differences in the chemical shift or the width of the ligand resonance when monitored by NMR. During the protein titration, the changes in chemical shift or integral of the ligand are thought to arise from a weighted average of the two extreme cases (fully bound or unbound). Monitoring of these changes over the course of the titration was used to measure binding affinities to *Tc*BDF3. This approach was also used on BRD4(1) (Appendix B3).

3.3.4.1. Establishing ^{1}H ligand observed protein titration for *Tc*BDF3

H_6-*Tc*BDF3 was titrated into fixed amounts of ligand to obtain the equilibrium constant K_D of the binding of representative compounds. Chemical shift changes and peak broadening of the ligand signals were observed in the ^{1}H spectra, whereas peaks arising from impurities (in buffer, D_2O or DMSO) were unchanged (Figure 3.15). The observed change of chemical shift $\Delta\partial_{obs}$ appeared to be indicative of binding, and offered a reliable way to measure the binding affinity. The change of broadening could also be tracked. However, only small changes were observed, and due to the high protein concentration required to saturate the ligand, those changes were hidden by the appearance of protein peaks. Therefore, the observed change of chemical shift $\Delta\partial_{obs}$ was used preferentially to the changes of peak width or integral for the protein titration experiments using *Tc*BDF3.

^{1}H NMR modified with a Carr-Purcell-Meiboom-Gill (CPMG) sequence[99-100] allows for the reduction or total removal or broad peaks, allowing clear visualisation of sharp ligand peaks. In the titration experiments performed on *Tc*BDF3 the peaks of interest displayed

broadening in addition to the changes in chemical shifts and, consequently, were supressed in the CPMG ^{1}H NMR. Therefore, the changes in chemical shift were measured on the unmodified ^{1}H spectra.

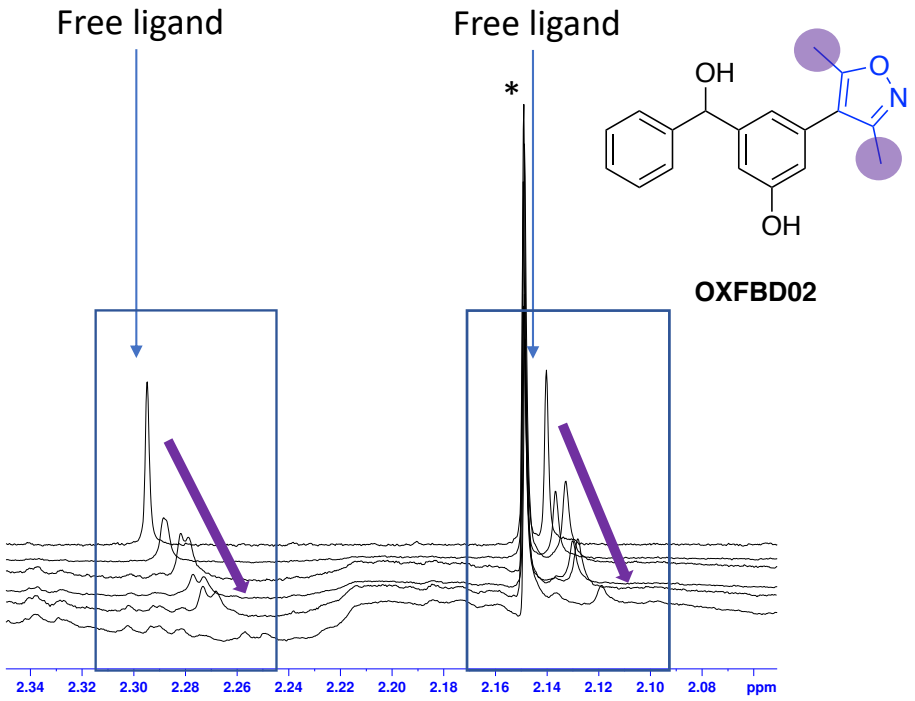


Figure 3.15: Representative example of the change of ligand chemical shift $\Delta\partial_{obs}$observed during titration using *Tc*BDF3. Spectra stack of the ^{1}H ligand-observed protein titration of *Tc*BDF3 into OXFBD02 (**7**), shows the free OXFBD02 (**7**) ligand (top spectra) and the ligand in the presence of increasing amounts of *Tc*BDF3. The chemical shift change $\Delta\partial_{obs}$ is highlighted by purple arrows. Impurities (*) do not display changes in resonance shifts or peak widths. The OXFBD02 (**7**) ligand is racemic and displays a split behaviour for one of the observed methyl groups. This behaviour is more significant when titrating *Tc*BDF5-1 (Appendix B.3.2).

3.3.4.2. Binding affinities of I-BET151 (**5**), BI-2536 (**8**) and OXFBD02 (**7**) to *Tc*BDF3 by NMR titration

Monitoring the changes of peak chemical shifts is less commonly used for ligand-observed experiments, so the principle of the analysis is detailed here. The following analysis has been adapted from known methods.[101] During the titration the P + L ⇌ PL equilibrium is studied. The total concentrations of ligand and protein $[L]_t$ and $[P]_t$ are known. $[L]_t$ is equal to the free ligand concentration [L] and the PL complex concentration [PL]. Similarly, $[P]_t$ is equal to the free protein concentration [P] and the PL complex concentration [PL]. Therefore, $[L]_t$ and $[P]_t$ can be defined using equations 1.1 and 1.2:

$$[L]_t = [L] + [PL] \quad (1.1)$$

$$[P]_t = [P] + [PL] \quad (1.2)$$

The change in chemical shift observed $\Delta\partial_{obs}$ is the weighted average of the shifts in the free and bound states (equation 1.3). f_f and f_b describe the fraction of total ligand free and bound respectively ($f_f + f_b = 1$). ∂_f and ∂_b denote the chemical shift of the free and bound ligand, respectively.

$$\Delta\partial_{obs} = f_f\partial_f \times f_b\partial_b \quad (1.3)$$

The maximum change of shift, $\Delta\partial_{max}$, is observed when the whole ligand population is bound to the protein. The protein-ligand complex can be described as the product of the total ligand concentration and the fraction of the bound ligand (equation 1.4).

$$[PL] = [L]_t \times f_b \quad (1.4)$$

The fraction bound can be described in terms of the ratio between the observed and maximum chemical shift change. Finally, these can be described as a quadratic function in terms of $[L]_t$, $[P]_t$ and K_D:

$$[PL] = [L]_t \frac{\Delta\partial_{obs}}{\Delta\partial_{max}} = \frac{[P]_t + [L]_t + K_D - \sqrt{([P]_t + [L]_t + K_D)^2 - 4[P]_t[L]_t}}{2} \quad (1.5)$$

Prism© was used to fit the data to predict $\Delta\partial_{max}$, then solve equation 1.5 for the processed data. Two approaches were compared to determine the K_D: one keeping $[L]_t$ constant, and one allowing for its fitting in addition to the fitting for the K_D value. This analysis gave the following K_D values for the protein titration into I-BET151 (**5**), BI-2536 (**8**), and OXFBD02 (**7**) (Table 3.7, Figure 3.16).

Table 3.7: Summary tables of the K_D values determined using Prism, with the total ligand concentration set as a fixed value or a variable. Titration of *Tc*DF3 into 100 µM of I-BET151 (**5**), BI-2536 (**8**), and OXFBD02 (**7**).

	I-BET151 (**5**)	BI-2536 (**8**)	OXFBD02 (**7**)
Non-linear fit with $[L]_t = 100\ \mu M$ kept constant.	312 µM	55 µM	508 µM
Non-linear fit with $[L]_t$ as a variable.	263 µM	26 µM	445 µM

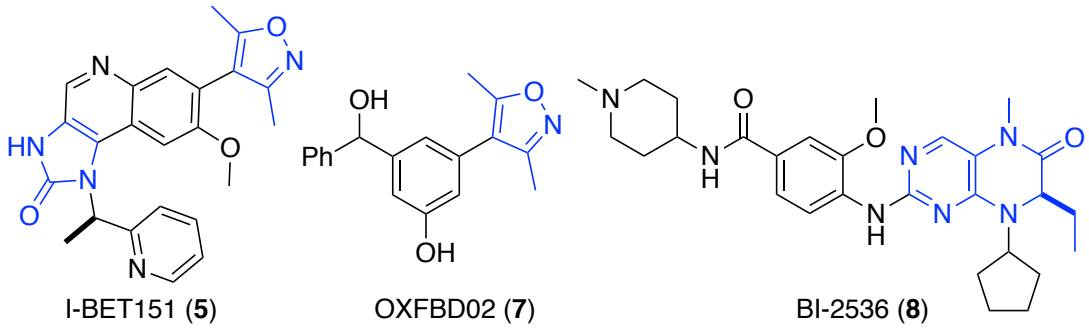


Figure 3.16: The structures of I-BET151 (**5**), BI-2536 (**8**), and OXFBD02 (**7**).

The dissociation constant, K_D, describe the equilibrium between the complex of a large and a small component and their dissociated forms. Here, protein has been titrated into a constant concentration of ligand, and the change of chemical shift has been used to obtain the amount of the bound complex [PL]. The K_D obtained using titrations into different constant ligand concentration should agree, within error. The protein titration into I-BET151 (**5**), OXFBD02 (**7**) and BI-2536 (**8**) were repeated using 50 µM as the ligand concentration to verify the reproducibility of the experiments. If determined and analysed correctly, the K_D should match even when using different [L] as they describe a constant that should be independent of the experiment conditions. Pleasingly, the K_D obtained through each separate experiments are in the same range for BI-2536 (**8**) and OXFBD02 (**7**) (Table 3.8). There seems to be a higher discrepancy in the analysis for I-BET151 (**5**), which arises from a poorer data fitting during the titration into 50 µM of ligand. This is highlighted by a lower R^2 value: 0.94 vs 0.99 (for the full analysis, see Appendix B3).

Table 3.8: Summary tables of the K_D values determined using Prism, with the total ligand concentration set as a fixed value or a variable. Titration of *Tc*DF3 into 50 µM of **5**, **8**, and **7**.

Fit parameters	I-BET151 (**5**)	BI-2536 (**8**)	OXFBD02 (**7**)
Non-linear fit with $[L]_t = 50\ \mu M$ kept constant.	76 µM	45 µM	455 µM
Non-linear fit with $[L]_t$ as a variable.	133 µM	27 µM	418 µM

The ^{1}H ligand-observed protein titration experiment established here provides K_D ranges to characterise the binding to *Tc*BDF3. This allowed for the first ranking of the identified hits for this protein. OXFBD02 (**7**), I-BET151 (**5**) and BI-2536 (**8**) bind to *Tc*BDF3 with K_D ranging from the triple to double digit micromolar allowing for the prioritisation of the

hits. BI-2536 (**8**), with an estimated K_D of 25–50 µM, was identified as the highest affinity ligand to *Tc*BDF3. Using the NMR titration, OXFBD02 (**7**) was also shown to bind to *Tc*BDF5-1 with a K_D range of 150–350 µM (Appendix B3).

3.3.5. Isothermal calorimetry of the TcBDF3/ BI-2536 (**8**) interaction

To identify any false positives and provide more insight into the binding interaction of *Tc*BDF3 to BI-2536 (**8**), isothermal titration calorimetry (ITC) was chosen as an orthogonal assay.

3.3.5.1. The binding of BI-2536 (**8**) to *Tc*BDF3 by ITC

To determine the most suitable screening conditions, ITC studies of the *Tc*BDF3/ BI-2536 (**8**) interaction were carried out using three set of protein to ligand ratios (Table 3.9). These conditions were predicted using the Malvern PeaQITC prediction software to detect binding ranging from 1–100 µM affinity. Although the binding interaction was detected in each case, fitting of the ITC trace for condition A was of lower quality than for the other two conditions. In particular, condition A gave a larger error in the K_D values (K_D=36.7±25 µM), as well as a low N number with a large error (N=0.067±0.433). ITC condition B and C gave higher quality ITC traces with lower errors in the K_D. The N number when using 50 µM of the BI-2536 (**8**) ligand (condition C) was lower than the one obtained when using 40 µM (condition B), which could be explained by the limited solubility of BI-2536 (**8**). BD–ligand interactions are expected to be 1:1 complexes, hence condition B, which has the lowest error in the K_D value and a N number of 0.724±0.029 was chosen as the preferred ITC screening condition.

Table 3.9: Summary of the *Tc*BDF3:**8** ratios used for ITC. The ITC Traces are shown in Appendix B4.

ITC screening conditions	BI-2536 (**8**) concentration	*Tc*BDF3 concentration	K_D (ITC)	N number
A	30 µM	300 µM	36.7±25 µM	0.067±0.433
B	40 µM	400 µM	13.0±3.29 µM	0.724±0.029
C	50 µM	500 µM	25.8±4.64 µM	0.421±0.037

ITC experiments using condition B were repeated in triplicate (Appendix B4), confirming the binding of BI-2536 (**8**) to *Tc*BDF3 with K_D=14.6±3.0 µM (ITC, n=3). These results corelate well with the K_D ranges obtained by the NMR titration experiments (K_D=27–45 µM (NMR, n=2)). The dihydropteridinone-based acetyl-lysine mimic, present in BI-2536 (**8**), was shown to be the most suitable fragment to bind to *Tc*BDF3 amongst the various fragments and ligands screened as part of this work.

3.3.5.2. The binding of BI-6727 (**9**) to *Tc*BDF3 by ITC

BI-6727 (**9**) is an analogue of BI-2536 (**8**) that shows increased affinity to PLK1, compared to BI-2536 (**8**), and is another example of a dihydropteridinone-based BD inhibitor (Figure 3.17).[102] In **9** an isopropyl group replaces the cyclopentyl substituent of the head group of **8**, and it contains an alternative solubilising tail. Due to solubility limitations, this compound could not be tested using waterLOGSY. Using ITC, BI-6727 (**9**) was screened using the optimal conditions determined for BI-2536 (**8**) (condition B, Table 3.9). These experiments confirm the binding of BI-6727 (**9**) to *Tc*BDF3, with K_D value of 57.1±4.9 µM (ITC, n=3). Overall BI-6727 (**9**) binds *Tc*BDF3 with to 2- to 3-fold weaker affinity than BI-2536 (**8**) (Table 3.10). These results confirm the binding of dihydropteridinone-based acetyl-lysine mimics, and suggests the core or tail of BI-2536 (**8**) can be modified, while retaining binding to *Tc*BDF3.

Table 3.10: Summary of the affinities of BI-2536 (**8**) and BI-6727 (**9**) for *Tc*BDF3.

Technique	BI-2536 (**8**)	BI-6727 (**9**)
waterLOGSY	Binding detected	Not soluble
ITC	K_D=14.6±3.0 µM (n=3)	K_D=57.1±4.9 µM (n=3).

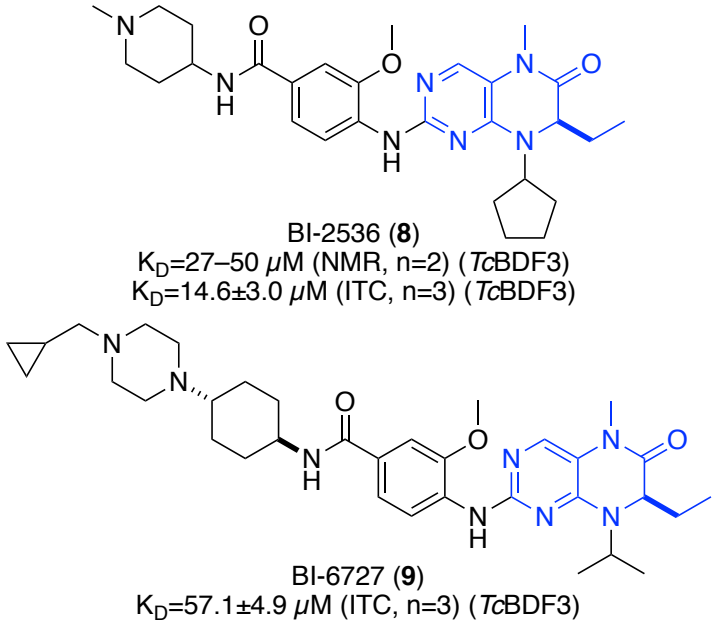


Figure 3.17: Structure of the dihydropteridinone-based dual PLK1/BRD4(1) inhibitors, BI-2536 (**8**) and BI-6727 (**9**), and their binding affinities to *Tc*BDF3. The ITC Traces are shown in Appendix B4.

3.3.6. Investigation of the binding of the *N*-methylated dihydropteridinone core of BI-2536 (**8**), **(*R*)-92**, to *Tc*BDF3

(*R*)-92 binds to *Tc*BDF3, whereas its enantiomer **(*S*)-92** does not show detectable binding by waterLOGSY (Section 3.3.3.4). This indicates that the core interacts directly with the protein, and lowers the chances of a potential inverted binding mode, where the tail of BI-2536 (**8**) would be buried into the binding site. The core *N*-methylation is also essential for binding by waterLOGSY. Indeed, **(*R*)-96** does not show binding to *Tc*BDF3 by waterLOGSY (Figure 3.18). To further prove that binding of BI-2536 (**8**) is a result of the binding of its KAc mimicking head group to the BD, the binding of **(*R*)-92** was assessed using the NMR titration experiment. Analysis of the titration of *Tc*BDF3 into **(*R*)-92** gave a K_D of 335–365 µM, showing a 10-fold drop of affinity as compared to BI-2536 (**8**). The core fragment is half the molecular weight of the ligand, therefore a drop in affinity was expected. Interestingly, the **(*R*)-92** has a higher LE of 0.24 compared to BI-2536 (**8**), which has a LE value of 0.16. The comparison of the LE values shows that the core itself is a good starting point for binding to *Tc*BDF3 (Table 3.11). Overall, BI-2536 (**8**) and **(*R*)-92** have been identified using waterLOGSY, NMR titration, and ITC as lead structures for further development towards the study of *Tc*BDF3.

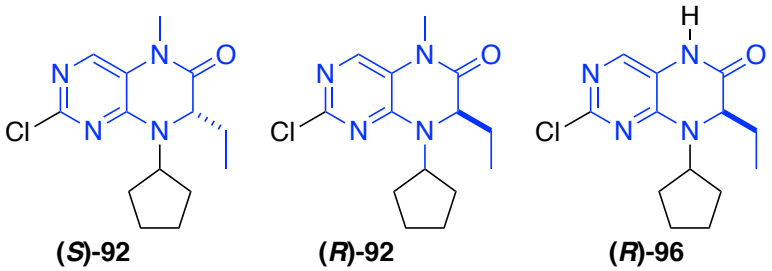


Figure 3.18: Structures of the negative control molecules for the hypothesised mode of binding. The enantiomer KAc mimic **(*S*)-92**, and the unmethylated core **(*R*)-96** did not bind *Tc*BDF3. The KAc mimicking moiety is highlighted in blue.

Table 3.11: Comparison of the binding affinity and LE for BI-2536 (**8**) and **(*R*)-92**.

Ligand	K_D (NMR)	Ligand efficency (LE)	K_D (ITC)
BI-2536 **(8)**	27–45 µM	0.16	14.6±3.0 µM
(*R*)-92	335–365 µM	0.24	–

3.4. DNA-encoded library

The initial strategy to identify *Tc*BDFs binders was to screen known KAc-mimicking fragments and BD ligands. This focused screening allowed the rapid identification of weak affinity ligands through waterLOGSY, however, the hit rates from this biased library cannot be used to assess the target tractability of the *Tc*BDFs proteins. To gain further insight into the ligandability of these proteins so we decided to use an encoded library technology (ELT) to screen a large unbiased library of DNA-encoded compounds. This strategy became available through a collaboration with GSK.

The DNA-encoded library technology uses chimeric structures containing single or double stranded DNA attached to small-molecules (Figure 3.19). Each portion of DNA is defined as encoding for a specific building block, forming a "DNA barcode" unique for each encoded molecule. [103] A library of encodings is screened against an immobilised protein target and the non-binding clusters are removed using wash steps. The retained binders can be eluted separately prior to PCR amplification and DNA sequencing.[103] Selection filters are applied to remove ambiguous hits and the clusters hits are then selected for off-DNA synthesis. This technology is high-throughput and can identify new binders with a known point of attachment, in the case of the initial hits: DNA. Therefore the identified hits are commonly amenable to functionalisation with fluorophores or tags.

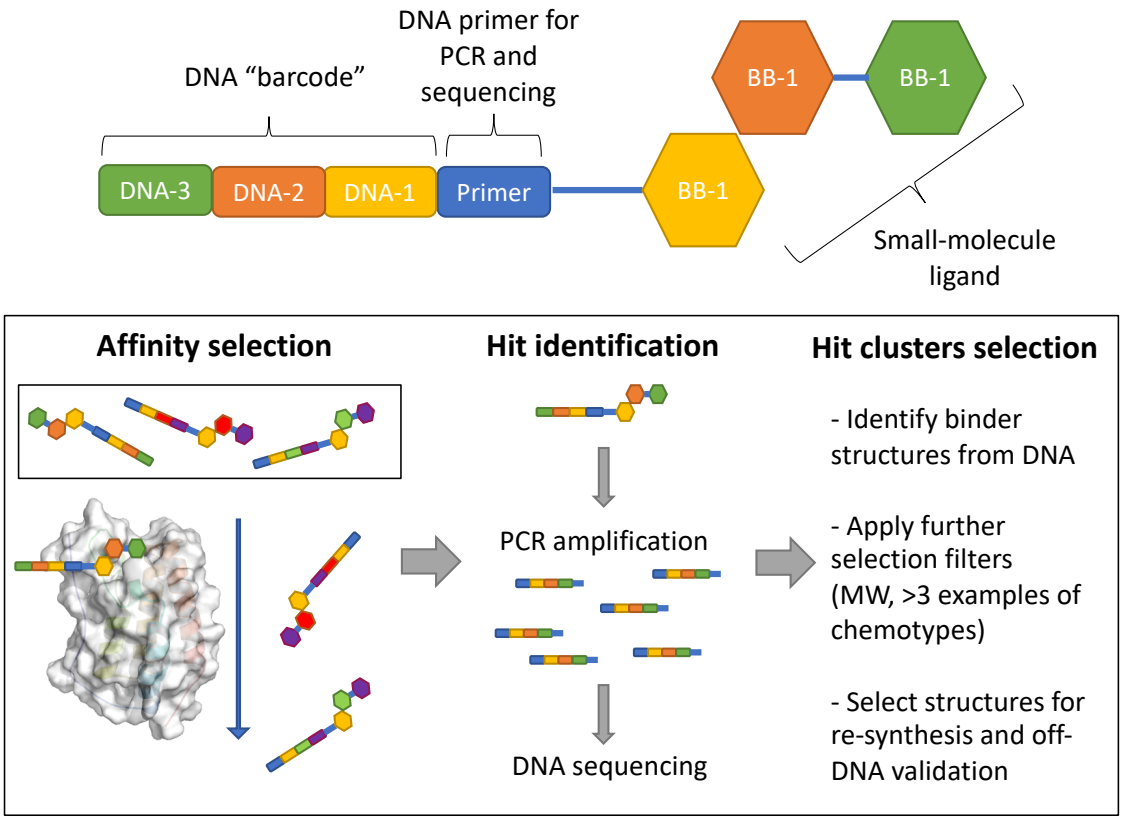


Figure 3.19: DNA-encoded libraries rely on the use of chimeric molecules formed of DNA and small molecules that be synthesised into large compounds library for high-throughput screening. The DNA fragments encode for specific building blocks (BB), and therefore form a "barcode" for each small molecule. The library is screened against the protein target using affinity selection: binders are retained by the immobilised target and non-bindings encodings are removed during wash steps. Hits are identified by PCR amplification and DNA sequencing. Further selection can be done using filters then structures are synthesis off-DNA for validation in orthogonal assays.[103]

*Tc*BDF2 and *Tc*BDF5-1 were selected to be screened using this technology as they both are thought to bind acetylated histone tails. The soluble protein constructs produced as part of this work were used as such as they possess a His-tag amenable for ELT screening. The screen was performed by Gang Yao, Heather O'Keefe and Chris Phelps (GSK Cambridge, MA, USA). The filters applied to select cluster hits included having over 3 copies of each chemical motif, over 3 chemical motifs types per cluster, and a molecular weight between 200 and 550 Da. A total of 82 hit clusters were identified to *Tc*BDF2 out of 1.9 trillion unique encodings screened, however, none of the hit clusters were selective. Similarly, 134 hits clusters were found to bind *Tc*BDF5-1, of which 8 were specific to *Tc*BDF5-1. These hit rates are relatively low, for instance, when this library was screened against one of the *Leishmania donovani* BDs, 470 specific clusters were identified. Unfortunately, the structures of the cluster hits have not yet been released, and therefore, cannot be shown in this thesis.

3.5. Summary and conclusions

Fragments, covering a variety of known acetylated-lysine mimics used in the inhibition of human BDs, have been synthesised. The synthesised fragments and larger BDs known ligands were screened by thermal shift and waterLOGSY to determine if the *T. cruzi* BD proteins bind KAc-mimicking motifs as expected. Overall, the thermal shift data was not considered significant enough to guide choices of acetyl-lysine mimics; However, waterLOGSY data obtained on the *T. cruzi* BD proteins has highlighted several favoured KAc mimicking fragments. Although, waterLOGSY can detect binding of interactions with a wide range of affinities, it is not quantitative. Therefore, an NMR-based titration experiment was established. This is the first report of using NMR-based titration to study the *T. cruzi* BD proteins.

The ^{1}H ligand-observed protein titration experiment allowed the ranking and prioritising of the identified hits for *Tc*BDF3. OXFBD02 (**7**), I-BET151 (**5**) and BI-2536 (**8**) bound to *Tc*BDF3 with K_D ranging from 27 to 455 µM. These interactions are at least 100-fold weaker than the binding of these ligands to BRD4(1). This drop in affinity can be explained by the low sequence identity (only 30%) between *Tc*BDF3 and BRD4(1). Additionally, these ligands interact with "WPF shelf" in the BET BDs, but these residues are "ASF" in *Tc*BDF3.

BI-2536 (**8**) was identified as the highest affinity ligand to *Tc*BDF3 using NMR titration, with K_D values ranging from 25 to 50 µM. BI-2536 (**8**) was also shown to bind *Tc*BDF3 by ITC, with a K_D value of 14.6±3.0 µM (ITC, n=3), which corelates well with the K_D ranges obtained by the NMR titration experiments. Additionally, **(*R*)-92**, which contains the KAc mimic of the BI-2536 (**8**) ligand, gave a K_D of 335–365 µM by NMR titration. Overall, these result show that BI-2536 (**8**) is the highest affinity *Tc*BDF3 binder identified to date, and that its binding to the BD protein arises from its KAc mimicking moiety. Consequently, BI-2536 (**8**) was chosen as a lead for further development (Chapter 4).

Chapter 4. Design and synthesis of tool compounds

4.1. Introduction and aims

Screening of KAc-mimicking compounds using waterLOGSY enabled the identification of binders to the *T. cruzi* BDs. These results confirm that these proteins bind to KAc-mimicking fragments in an analogous way to human BDs. The screening efforts focused on *Tc*BDF3 due to its higher sequence identity (≥ 30%) to the human BD CREBBP, BRD2(2), BRD3(2), and BRD4(2) (Table 3.1). The proton ligand-observed protein NMR titration identified the dual PLK1/BRD4 ligand BI-2536 (**8**) as the highest affinity *Tc*BDF3 ligand, with a K_D of 27–55 μM. Although these titration experiments were a turning point in this work, this NMR assay is low throughput, requiring high amounts of protein and time on the 700 MHz cryoprobe NMR spectrometer. ITC further characterised the binding of BI-2536 (**8**) to *Tc*BDF3 (Section 3.3.5), however is it relatively low throughput. Therefore, we decided to investigate higher throughput screening assays for *Tc*BDF3.

The development of biophysical assays for novel non-enzymatic targets is a common topic in medicinal chemistry and chemical biology.[104] One approach relies on the use of the endogenous partner, or a mimic of the endogenous ligand, to set up displacement assays. For example, displacement of the biotinylated acetylated H4 histone peptide is commonly used for characterisation of human BET BD ligands by Alpha Screen.[57] In the absence of a known endogenous partner, other approaches can be employed. One alternative approach includes the modification of a known binder into a tool compound, such as a fluorescent or diazirine probe. For example, Zhang *et al*.[48] reported the use of a fluorescein labelled triazolodiazepine derived BD ligand to study the inhibition of BRD4(1).

*Tc*BDF3 is thought to bind KAc residues on parasite α-tubulin.[12] However, this interaction is not fully characterised. In particular, there is no available binding affinity, or published sequence for an acetylated α-tubulin mimicking peptide. Consequently, the development of novel assays for *Tc*BDF3 using the displacement of the endogenous partner or its mimics are not readily available. The work detailed in Chapter 2 enabled the

identification and validation of three *Tc*BDF3 ligands. In particular using several orthogonal assays, BI-2536 (**8**) has been shown to have the highest affinity for *Tc*BDF3. Importantly, its KAc-containing core, **(*R*)-92** was shown to be key for binding using waterLOGSY and NMR titration (Figure 4.1). BI-2536 (**8**) was chosen as a lead structure for development of higher throughout biophysical assays for *Tc*BDF3, as it contains a validated KAc mimicking headgroup and a tail that could be altered to attach additional functional groups. In addition, this lead structure has reasonable synthetical tractability and does not contain known assay interference moieties, such as PAINS.[105] We therefore decided to synthesise three categories of tool compounds based on the BI-2536 (**8**) ligand. These target compounds are a ^{19}F NMR reporter molecule, a diazirine photo-activatable probe and a fluorescently-labelled probe.

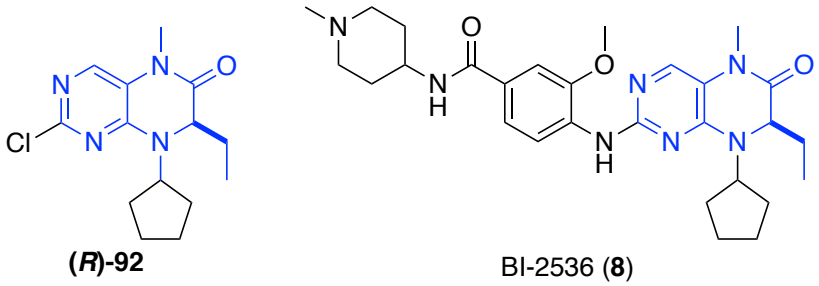


Figure 4.1: The structures of the dihydropteridinone-based KAc mimic **(*R*)-92** and BI-2536 (**8**). The KAc mimicking-moiety is shown in blue.

The aim of the work described in this chapter was to design appropriate modifications of the BI-2536 (**8**) ligand to provide tool compound for mass spectrometry-, fluorescent- or ^{19}F NMR-based biophysical assays (Section 4.2–4.3). The synthesis of common intermediates is described in Section 4.4, and the assembly of the tool compounds is detailed in Section 4.5–4.7. The synthesis of control compounds, designed to be used to validate the assays, is detailed in Section 4.8. The development and set up of the novel assays will be described in Chapter 5.

4.2. Prediction of the mode of binding of BI-2536 (**8**)

4.2.1. Docking of BI-2536 (**8**) into the *Tc*BDF3 homology model

To determine points of modification of the BI-2536 (**8**) ligand, its binding mode was first studied in published BD co-crystal structures: human BRD4(1) (PDB: 4O74)[38-39, 92] and PP-BRD20, a BD-containing protein from *T. brucei* (PDB: 5C8G; PP-BRD20 from *Tb*427tmp). PP-BRD20 was shown in a BLAST search to have 65% sequence identity to *Tc*BDF3. In both proteins BI-2536 (**8**) binds in the ZA-channel of the bromodomain, and forms a hydrogen bond to the conserved asparagine residue (N140 in BRD4(1); N100 in PP-BRD20) (Figure 4.2). The dihydropteridinone *N*-methyl group points towards the binding pocket, whereas the *N*-cyclopentyl group points out of the pocket.

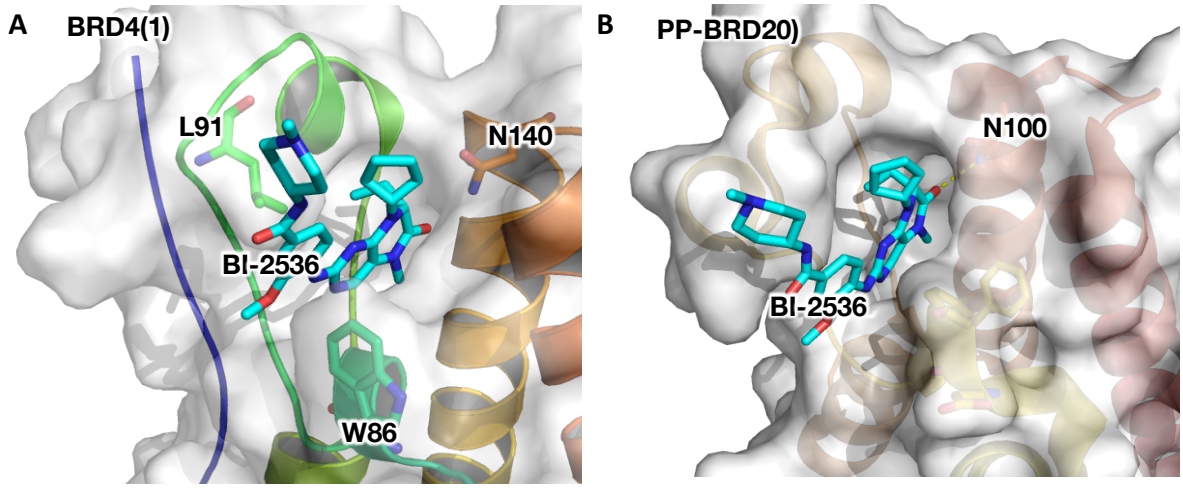


Figure 4.2: The binding mode of BI-2536 (**8**) to (A) BRD4(1) (PDB: 4O74) and (B) PP-BRD20 (PDB: 5C8G). Image generated using PyMol.

The homology model of *Tc*BDF3, built by Joseph Bluck (Chapter 2), was used to dock BI-2536 (**8**). This docking predicts BI-2536 (**8**) binding to *Tc*BDF3 with an orientation analogous to its binding to BRD4(1) and PP-BRD20 (Figure 4.3). The ligand is anchored to the protein *via* hydrogen bonding from the KAc-mimic carbonyl to the binding site asparagine residue (N124 in *Tc*BDF3).

The homology model of *Tc*BDF3 predicts a wider ZA-channel than in BRD4(1). Critically, the methoxy-aromatic group of the ligand is held more tightly in BRD4(1). The BI-2536

(**8**) ligand fills the BRD4(1) ZA-channel, and forms interactions with residues on both sides of the channel (Figure 4.2). The *Tc*BDF3 L75 residue is analogous to L91 in BRD4(1); however, W81 is replaced by the much smaller alanine A64 (Figure 4.3). The distance between the two adjacent residues is much larger than in BRD4(1), and BI-2536 (**8**) does not fill the ZA channel efficiently in *Tc*BDF3. Indeed, the high flexibility of *Tc*BDF3, and its large ZA-channel, lacking bulky aromatic groups makes this protein a challenging target. These findings might account for the difference in binding affinities of BI-2536 (**8**) to BRD4(1) compared to *Tc*BDF3.

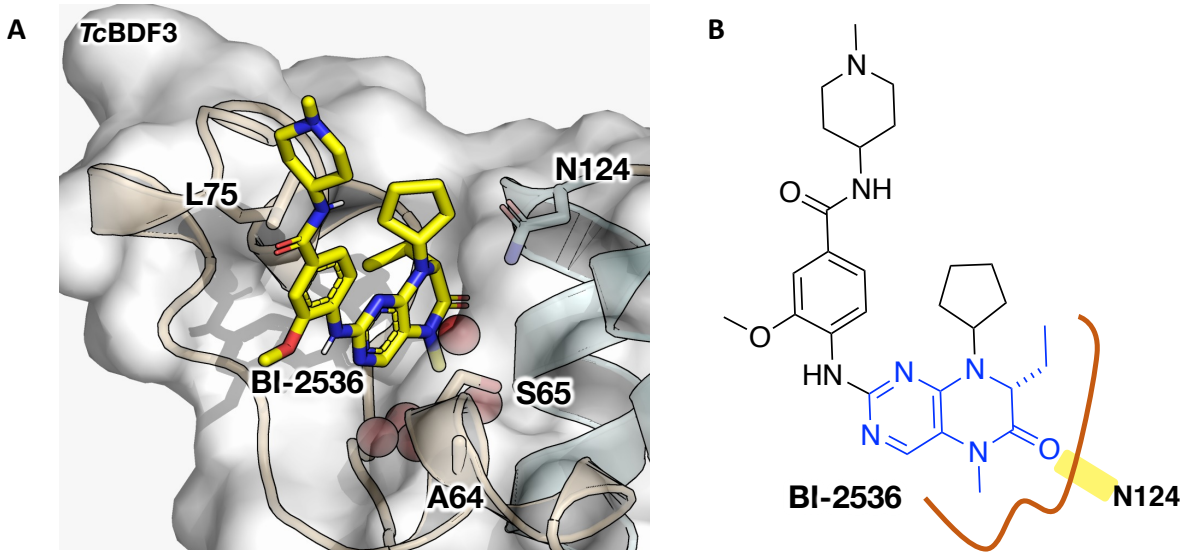


Figure 4.3: (A) Docking of BI-2536 (**8**) into the homology model of *Tc*BDF3 by Joseph Bluck (B) The predicted mode of binding with the hydrogen bonding to N124 represented in yellow. Image generated by Joseph Bluck using PyMol.

4.2.2. Potential point of modifications

The homology model allowed us to predict the binding mode of this ligand to *Tc*BDF3 and suggested reasons why higher affinity ligands have not been identified up to this point. Additionally, the homology model suggests that the piperidine solubilising group of BI-2536 (**8**) is solvent exposed. Using this binding mode model, it was hypothesised that modifications of the tail of the ligand could be made without significant loss of binding affinity. The waterLOGSY and NMR titration data (see Section 3.3.3–3.3.4) also showed that **(*R*)-92**, the *N*-methylated dihydropteridinone core of BI-2536 (**8**), binds to *Tc*BDF3. Direct attachment of tool compounds features onto the core might also be

possible, as **(*R*)-92** has an higher ligand efficiency than BI-2536 (**8**). These modifications are predicted to results in lower affinity, as **(*R*)-92** has a K_D 10-fold lower than BI-2536 (**8**); however, they might be advantageous when the functionality needs to be closer to the binding site, for example, for cross-linking purposes.

Overall, BI-2536 (**8**) was validated as a lead structure by using orthogonal binding assays and was docked into the *Tc*BDF3 homology model. These validation experiments suggest a mode of binding through the KAc mimic, and highlight two potential point of modifications for the design of tool compounds: the piperidine tail and the aniline attachment point.

4.3. Design of tool compounds

4.3.1. Target ^{19}F NMR probes

The first category of tool compound investigated was ^{19}F NMR reporter molecules. These molecules were designed to bind to *Tc*BDF3 and be displaced by other ligands. Using ^{19}F NMR to track the binding of a molecule can be an advantageous alternative to the ^{1}H NMR waterLOGSY or titration experiment, as ^{19}F has a high isotopic abundance, but low chemical abundance. Organic molecules and proteins contain high amounts of protons, but very few fluorine atoms. Consequently, the measurements are fast (30 mins, *vs* 1–4 h for waterLOGSY) and no protein background is observed in the ^{19}F NMR spectra.

The fluorinated ligand has one ^{19}F resonance per fluorine environments with a defined chemical shift and peak width assuming the resonances do not overlap. Upon titration of the protein, the ligand binds to the protein changing the environment of the ^{19}F atom(s) (Figure 4.4). In the case of fast exchange, the resulting peak is a weighted average of both states of the ligand. Bound ^{19}F NMR reporter molecules can be displaced by the addition of other *Tc*BDF3 binders, and the ^{19}F resonances restored to the free ligand state. This allows for the identification of non-fluorinated binders.

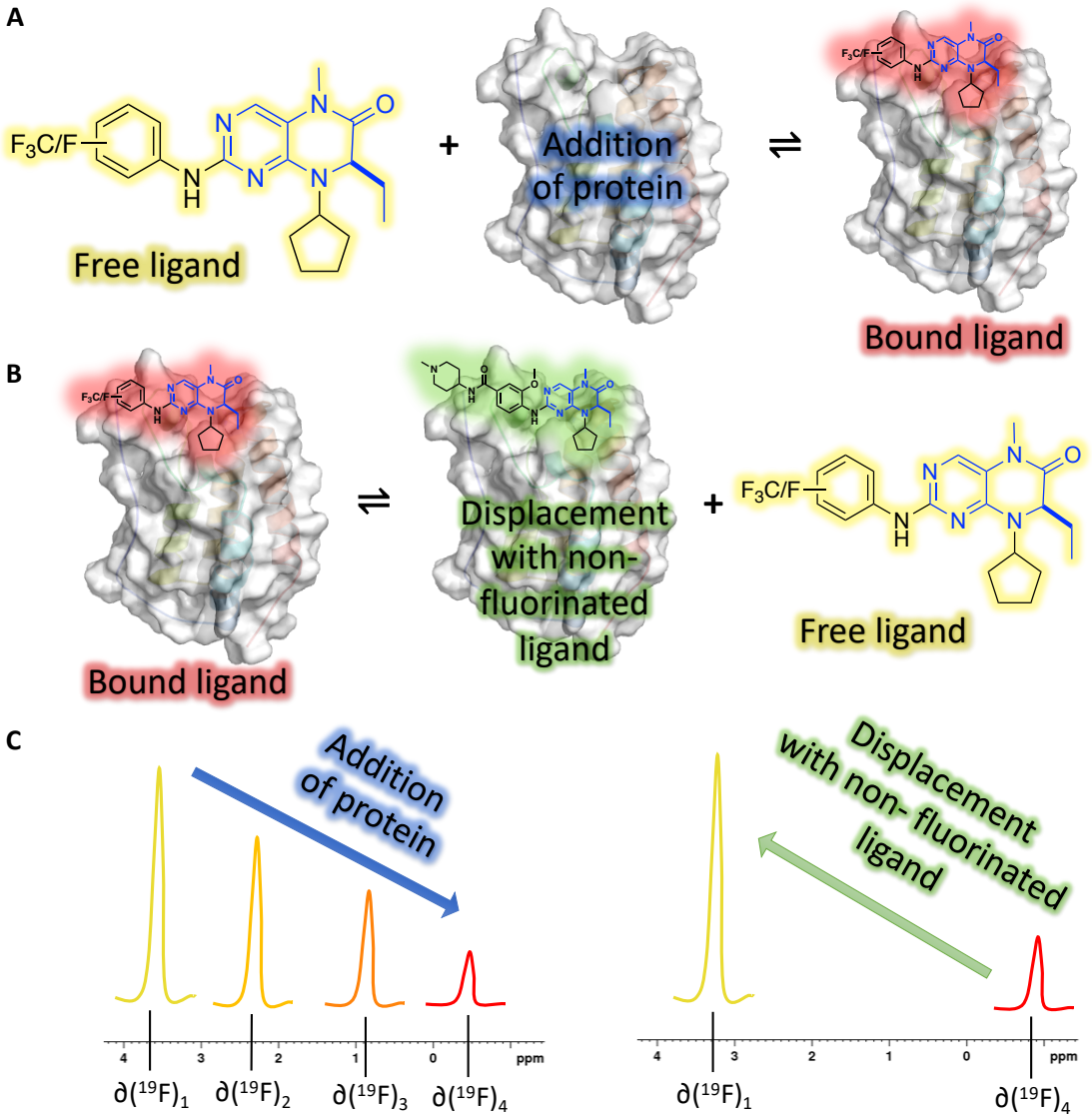


Figure 4.4: Schematic representation of (A) binding of a fluorinated ligand (free ligand: yellow, bound ligand: red) upon addition of protein (blue) and (B) displacement of the bound ligand by a competitor (green) recovering the free ligand (yellow). (C) The addition of protein (blue) can induce changes in chemical shift and in width of the ^{19}F resonance (depicted for the case of fast exchange), which is recovered upon addition of a competitor (green).

The position of the fluorine atom can potentially affect its sensitivity to the binding event. The piperidine ring in BI-2536 (**8**) is thought to be solvent exposed when binding, so modification on this group was predicted to result in minor changes upon binding of BI-2536 (**8**). Therefore, the target ^{19}F NMR reporter molecules **(*R*)-130** and **(*R*)-131** were designed to display fluorine atoms on the aromatic ring directly attached to the KAc mimicking core (Figure 4.5).

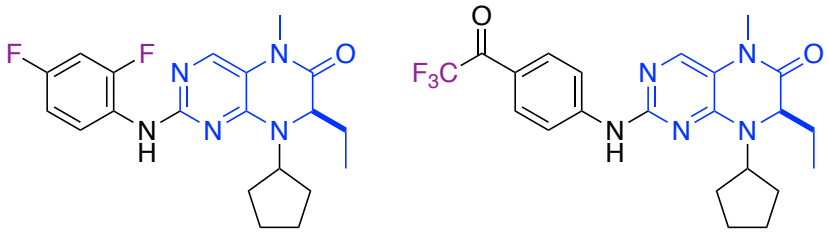


Figure 4.5: Structures of the target ^{19}F reporter molecules **(*R*)-130** and **(*R*)-131**. The KAc mimicking-moiety is shown in blue and the fluorine atoms are shown in purple.

4.3.2. Target diazirine molecules

The second types of target tool compounds investigated were photo-activatable diazirine probes, that could cross-link to the protein upon irradiation. Diazirines are propene-like organic molecules containing a strained double bond between two nitrogen atoms. Irradiation of diazirines liberates nitrogen gas and form a reactive carbene, which can insert into C-H, O-H and N-H bonds (Figure 4.6).[106] The cross-linking event can be detected by mass spectrometry and thus provide a tool to develop fragment screening assays. Given suitable cross-linking ability of the probes, competition assays can be used to screen for fragment binding the bromodomain and prevention of photo-crosslinking. To crosslink to the protein a diazirine probe must efficiently bind the target protein, and the formed carbene must be in proximity of the protein surface. Hence the crosslinking yield is a result of the binding affinity and the cross-linking efficiency, and is not a direct measure of either event (see Chapter 5).

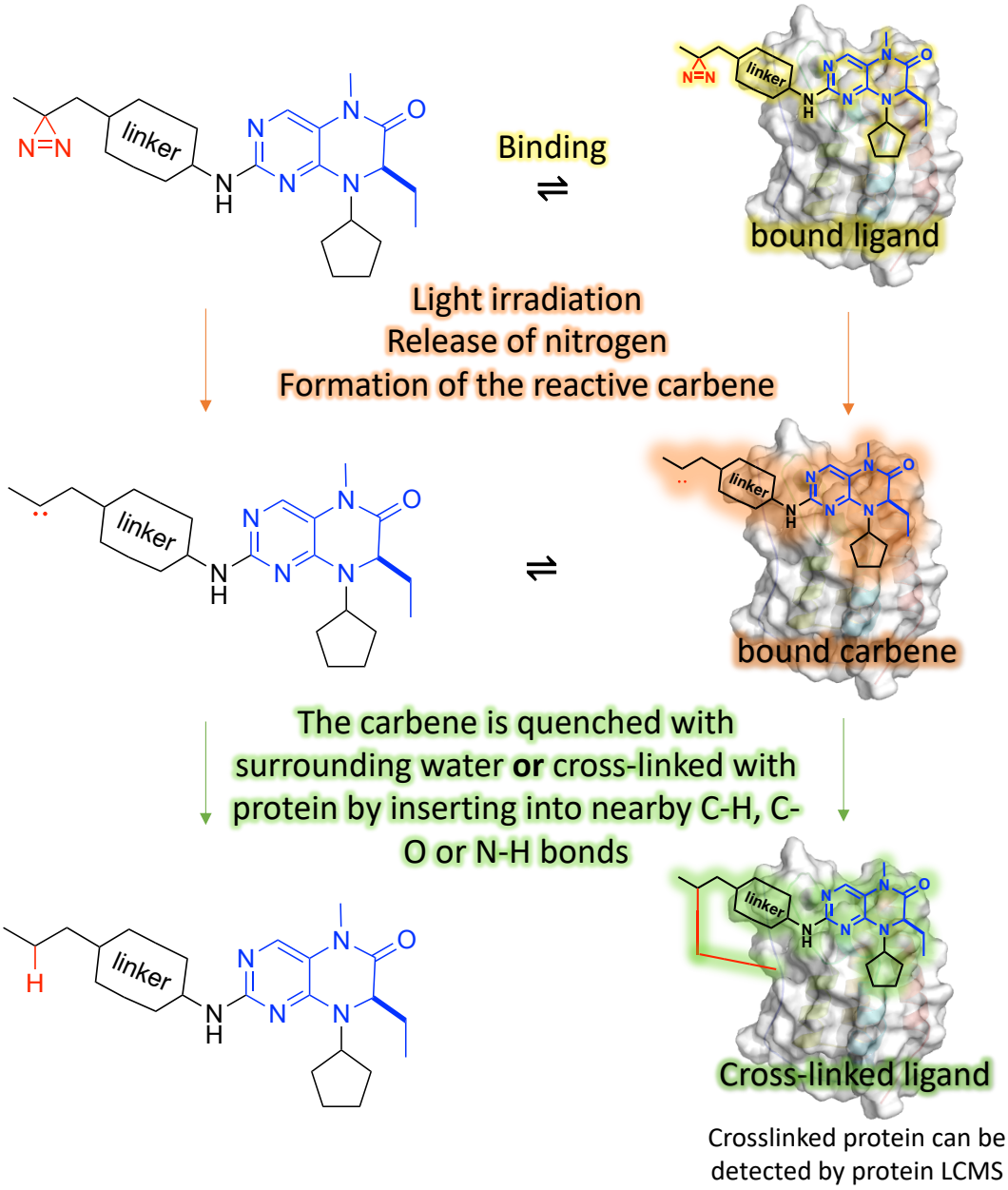
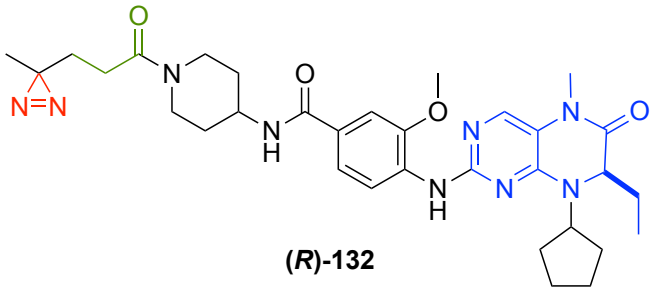


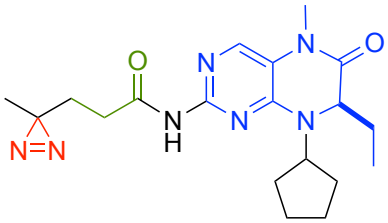
Figure 4.6: Schematic representation of the binding of the diazirine tool compound to the protein and its conversion to a carbene upon light irradiation. The carbene is quenched by reaction of nearby water, or if bound and sufficiently close to suitable residues, it can crosslink to the protein.

Two diazirine target molecules were designed (Figure 4.7). The first target, **(*R*)-132**, comprises BI-2536 (**8**) and an aliphatic diazirine onto the piperidine group. This target was hypothesised to bind *Tc*BDF3 efficiently, as it conserves most of the lead structure. On the other hand, this compound may have a low cross-linking yield due to the distance of the diazirine to the protein surface, as predicted using the homology model (Section 4.2.2). The second target molecule, **(*R*)-133**, consists of the ligand efficient core of BI-2535 (**8**), **(*R*)-92** (LE=0.24) and a short linker to an aliphatic diazirine. This tool

compound is predicted to have a low affinity than **(*R*)-130**, but a higher cross-linking efficiency as the formed carbene is likely to be closer to the protein surface.



(*R*)-132



(*R*)-133

Figure 4.7: The structures of the target photoreactive diazirine tool compounds, **(*R*)-132** and **(*R*)-133**. The diazirine is highlighted in red, the short aliphatic linker in green and the KAc-mimicking core in blue.

4.3.3. Target fluorescent probes

The third type of tool compounds designed to study *Tc*BDF3 were fluorescent probes for use in a fluorescence polarisation (FP) assay. The FP probes consist of a fluorophore, a linker, and the scaffold of BI-2536 (**8**). Following excitation by polarised light, the free fast tumbling fluorophore emits non-polarised light, whereas the fluorescent probe bound to the protein tumbles much slower perturbing the light less, and emits polarised light (Figure 4.8). The linker and fluorophore identity are key as they determine whether binding will be retained to *Tc*BDF3, and if the fluorophore will emit measurable polarised light when bound to the protein. These variables are hard to predict, and it might take several rounds of design and synthesis to obtain a suitable FP probe. Numerous commercially available fluorophores such as the AlexFluor® (AF) fluorophores are provided as *N*-succinamide (NHS) activated esters and can be conveniently appended to ligands bearing a free amine, using DIPEA in DMF at RT. The TAMRA fluorophore is an example of an affordable fluorophore available with an azide functional group. The piperidine ring of BI-2536 (**8**) was chosen as a point of attachment as it is predicted to be

solvent exposed. Overall, four different fluorescent ligands **(*R*)-134**–**(*R*)-137** were designed by varying the linker and commercially available fluorophores (Figure 4.9).

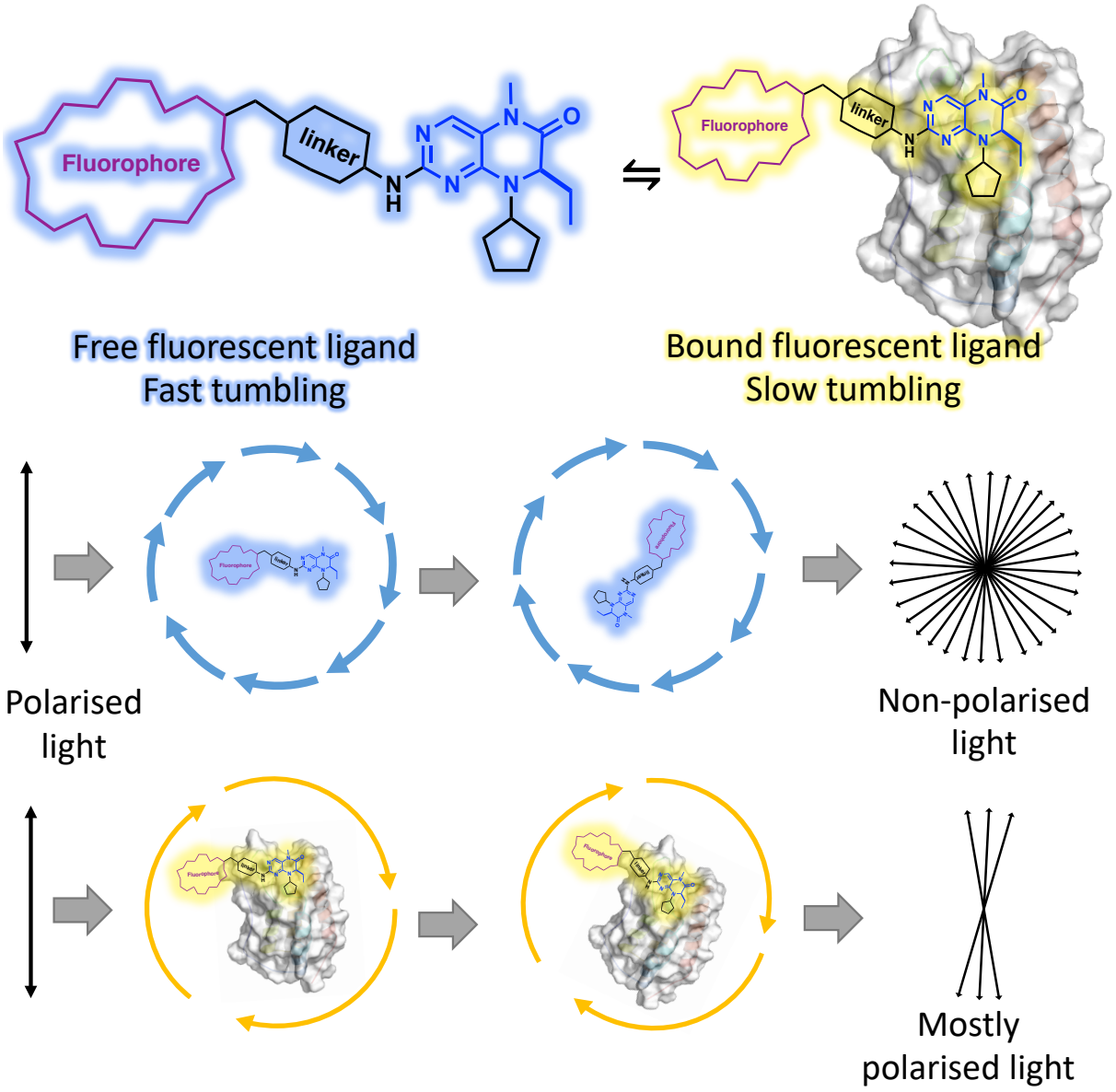


Figure 4.8: Schematic representation of the use of change of fluorescence polarisation (FP) to detect binding to a protein. The incident polarised light excites the fluorophore, the polarisation state of the emitted light depends on the tumbling rate of the fluorescent ligand: fast tumbling of the free ligand (blue) results in non-polarised light. Slow tumbling of the fluorescent ligand, arising from binding to a protein, results in a greater amount of polarisation for the emitted light. The FP probe contain a fluorophore, a linker, and the scaffold of a known binder.

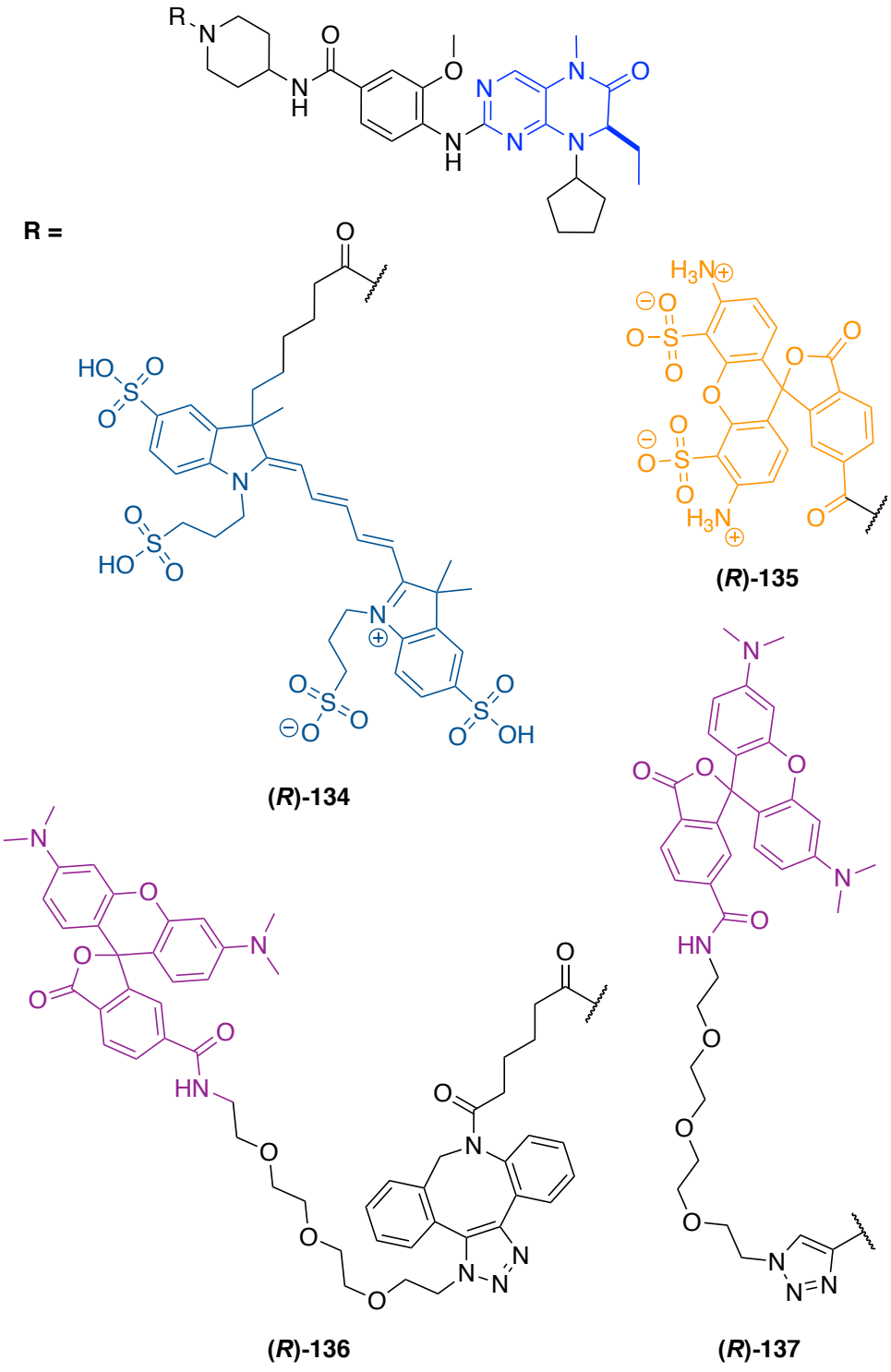


Figure 4.9: The structures of the four target fluorescent probes for *Tc*BDF3: **(*R*)-134**, **(*R*)-135**, **(*R*)-136**, and **(*R*)-137**. The KAc mimicking-core is shown in blue and the fluorophore are represented with colour corresponding to their emission wavelengths.

Overall, three categories of tool compounds were designed. If the tool compounds are shown to bind *Tc*BDF3 in their respective biophysical assays, they could be used to monitor their displacement by competing ligands. These three categories of tool compounds are appealing as the assays they would provide require less protein, and are higher throughput than the initial NMR-based assays (Chapter 2).

4.4. Synthesis of common intermediates for the tool compounds

The target tool compounds, two ^{19}F reporter molecules, 2 diazirines, and 4 fluorescent probes all contain the *N*-methylated dihydropteridinone core of BI-2536 (**8**). The piperidine tail or the aniline bridge were chosen as points of attachments. **A**–**C** were designed as key intermediates in the synthesis the target tool compounds (Figure 4.10). The target ^{19}F reporter molecules, diazirines and fluorescent probes described in Section 4.3 were synthesised from **A**–**C** through alkylation, amide formation, S_NAr, or palladium catalysed cross-coupling reactions. The synthesis of the key intermediate **A**–**C** is be described below, and the subsequent synthesis of the target tool compounds is be detailed in Section 4.5–4.7.

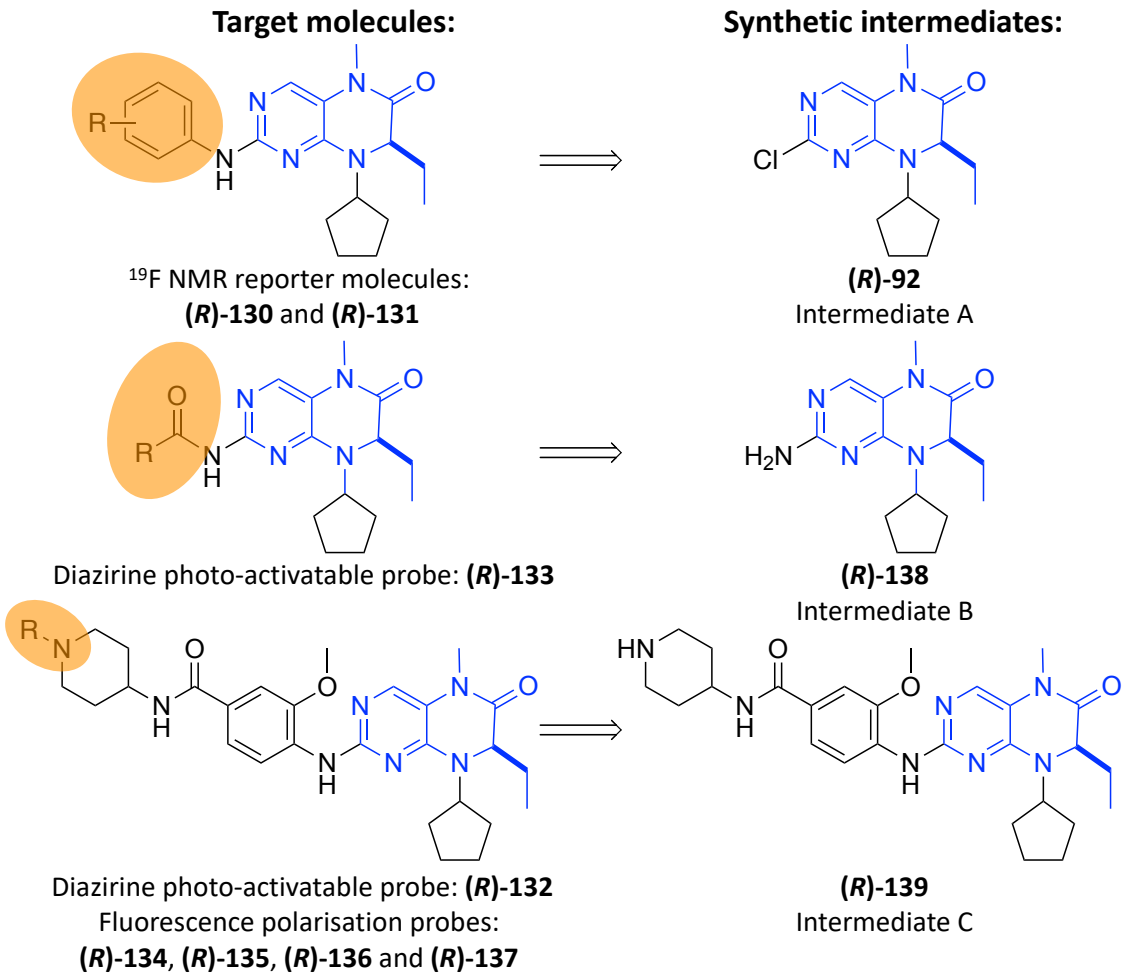
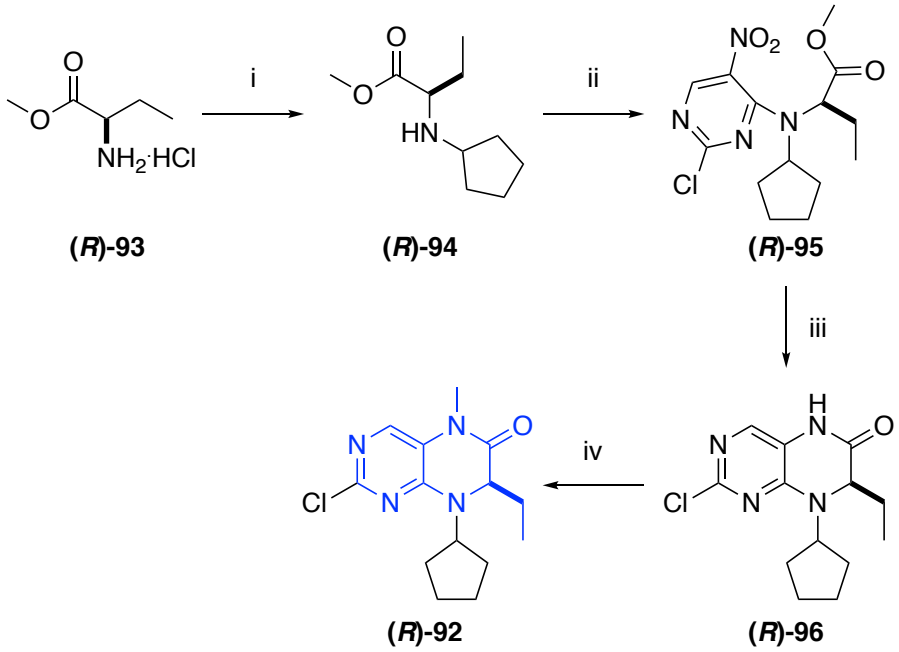


Figure 4.10: The general structures of the three categories of tool compound targets and the key synthetic intermediates **A**–**C**. Intermediate **A** is the BI-2536 (**8**) core tested in Chapter 2, it offers the possibility to attach aniline substituents *via* S_NAr or palladium-catalysed cross-coupling reactions. The closely related intermediate **B**, which is the aminopyrimidine analogue of intermediate **A**, can be used in amide couplings. Intermediate **C** is the intermediate closest in structure to BI-2536 (**8**) and offers the possibility to make tool compounds *via* alkylation or amide couplings.

4.4.1. Intermediate **A**: Synthesis of the chloropyrimidine fragment **(*R*)-92**

(*R*)-92 was synthesised (Scheme 4.1) as described in Chapter 3, Section 3.2.1.6.



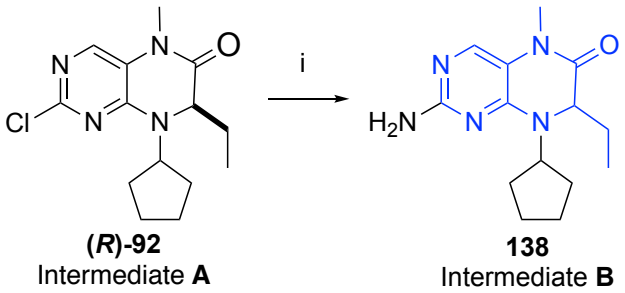
Scheme 4.1: Synthesis of the BI-2536 (**8**) **(*R*)-92** core (Intermediate **A**). *Reagents and conditions*: (i) NEt_3, $Ga(OTf)_3$, DCE, RT, 16 h, 65–78% (n=2). (ii) 2,4-Dichloro-5-nitropyrimidine, $NaHCO_3$, cyclohexane, RT then 60 °C, 3 d, 51–80% (n=3). (iii) Glacial acetic acid, 70 °C then Fe (activated), 70 °C, 1h, then 100 °C, 4h, 61–81% (n=3). (iv) (a) MeI, DMF, –10 °C, 5 mins (b) NaH, –10 °C (c) RT, 1h, 85–98% (n=6) or K_2CO_3, MeI, DMF, RT, 3 d, 45% (n=1).

4.4.2. Intermediate **B**: the aminopyrimidine intermediate **(*R*)-138**

Direct amination of **(*R*)-92** was explored in an attempt to synthesise **(*R*)-138** in one step from the previously described synthetic intermediates (Table 4.1, Scheme 4.2). This route was attractive as it involved the use of intermediate **A**, **(*R*)-92**, which can be synthesised on large scale. Direct amination of **(*R*)-92** with NH_4OH/IPA (Table 4.1, condition A) or 7 M NH_3 in MeOH (Table 4.1, condition B) on a small scale was unsuccessful at temperatures below 150 °C, and required microwave radiation to proceed. After 4 h at 150 °C, 20% conversion was obtained using NH_4OH/IPA versus only 6% for 7 M NH_3 in MeOH as determined by LCMS. Condition A was repeated on a larger scale, and **138** was obtained in 80% yield after 16 h at 160–170 °C (Table 4.1, condition C). The harsh reaction conditions were suspected to be responsible for the racemisation. Analysis by $[\alpha]_D$ and chiral HPLC confirmed that the product had indeed fully racemised.

Table 4.1: Summary of the S_NAr reaction conditions for the direct amination of **(*R*)-92**.

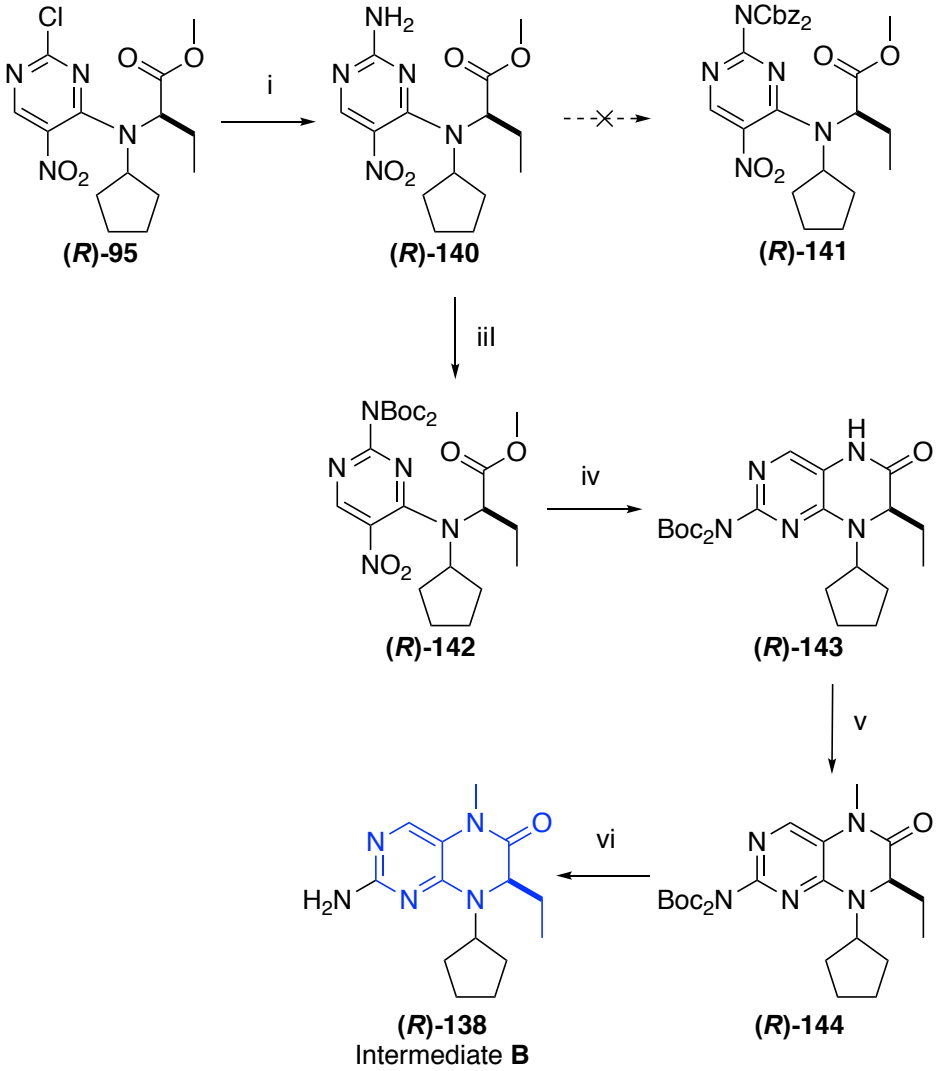
S_NAr condition	reagents	T (°C)	time	Conversion by LCMS (SM%:product%)	Isolated yield
A	NH_4OH/IPA 1:1 (100 eq.)	100 °C	1 h	–	–
		150 °C	5h	76:20	–
		170 °C	17 h (+200 eq. NH_4OH)	57:38	–
		170 °C	23 h	5:90	–
B	7 M NH_3 in MeOH (50 eq)	100 °C	1 h	100:0	–
		150 °C	5 h	94:6	–
C	NH_4OH/IPA 1:1 (100 eq.)	170 °C	6 h	–	–
		160 °C	12 h	–	–
		170 °C	16 h	0:96	80%



Scheme 4.2: Direct amination of **(*R*)-92** to intermediate **B**, **130**, leads to full racemisation of the stereogenic centre. *Reagents and conditions*: (i) NH_4OH/IPA = 4:1, 160–170 °C, 6 h, MW, 80% (n=2).

Direct conversion of **(*R*)-92** to **(*R*)-138** required elevated temperatures, long reaction times under microwave radiation, and excess of ammonium hydroxide leading to racemisation (Table 4.1). Therefore, an alternative synthetic strategy was needed. It was hypothesised that the S_NAr reaction using NH_4OH would be facile in presence of a nitro group.[107-108] The reaction was therefore attempted on the uncyclised **(*R*)-95** intermediate. Indeed, the reaction with ammonium hydroxide afforded **(*R*)-140** in 58–74% yield (n=2), without the need for high temperatures or microwave radiation. We planned to proceed with the reduction of the nitro group, cyclisation, and methylation using an analogous synthetic strategy to the synthesis of intermediate **A**, **(*R*)-92** (Scheme 4.1). To avoid regioselectivity problems during methylation, we decided to protect the aminopyrimidine moiety prior to reduction. The carboxybenzyl (Cbz) protecting group was chosen as it is known to be stable to reduction conditions, including Fe in acetic acid.[109] However, the protected intermediate **(*R*)-141** was unstable, and could not be purified from the degradation products (Scheme 4.3). Therefore, a *tert*-

butyloxycarbonyl (Boc) protecting group was investigated next. Protection of **(*R*)-140** with Boc_2O gave **(*R*)-142** in 68–80% yield (n=2) (Scheme 4.3). Reduction of the nitro group was carried out with Zn and NH_4Cl as an alternative to Fe and acetic acid, which was expected to lead to the deprotection of the Boc group.[109] Methylation of **(*R*)-143** with K_2CO_3 and MeI was quenched before reaching completion, as TLC analysis showed increasing formation of side-product formation. This reaction gave **(*R*)-144** in 28% and the reaction was not optimised further as enough material was obtained for the subsequent steps. Boc deprotection with TFA was very rapid and no side products were observed, so intermediate **B**, **(*R*)-138**, was used without further purification to make the diazirine tool compound **(*R*)-133** (see Section 4.6).



Scheme 4.3:The synthesis of Intermediate **B**, **(*R*)-138**, using a protecting group strategy to avoid racemisation. *Reagents and conditions*: (i) NH_4OH, THF, RT, 6 h, 58–74% (n=4). (ii) Boc_2O, DMAP, DMF, RT, 2 h, 68–80% (n=2). (iii) Zn, NH_4Cl, DMF, RT, 18 h, 49–82% (n=2) (iv) MeI, K_2CO_3, DMF, RT, 24 h, 28% (n=1). (v) TFA, CH_2Cl_2, RT, 20 mins, >99% (n=1).

4.4.3. Intermediate **C**: the amino-piperidine intermediate **(*R*)-139**

The retrosynthesis of intermediate **C**, **(*R*)-139** (Figure 4.11), shows two strategies: a convergent approach (path A) which involves the separate synthesis of the tail **146** before its addition to the core **(*R*)-92**, or a stepwise built onto the **(*R*)-92** fragment (path B).

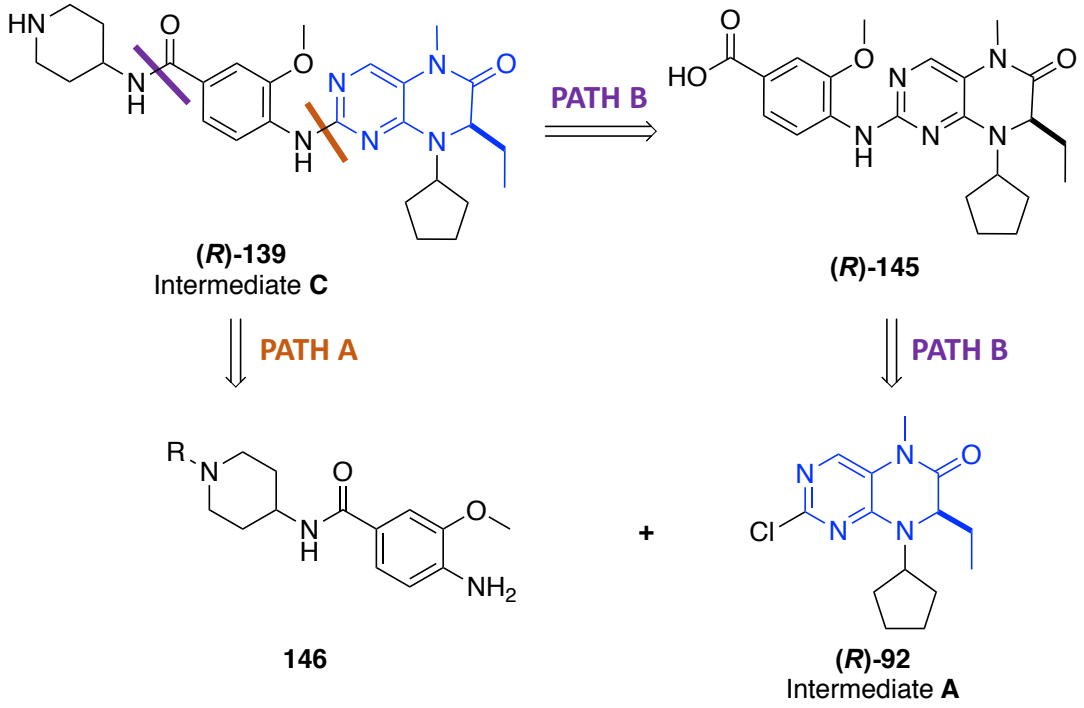
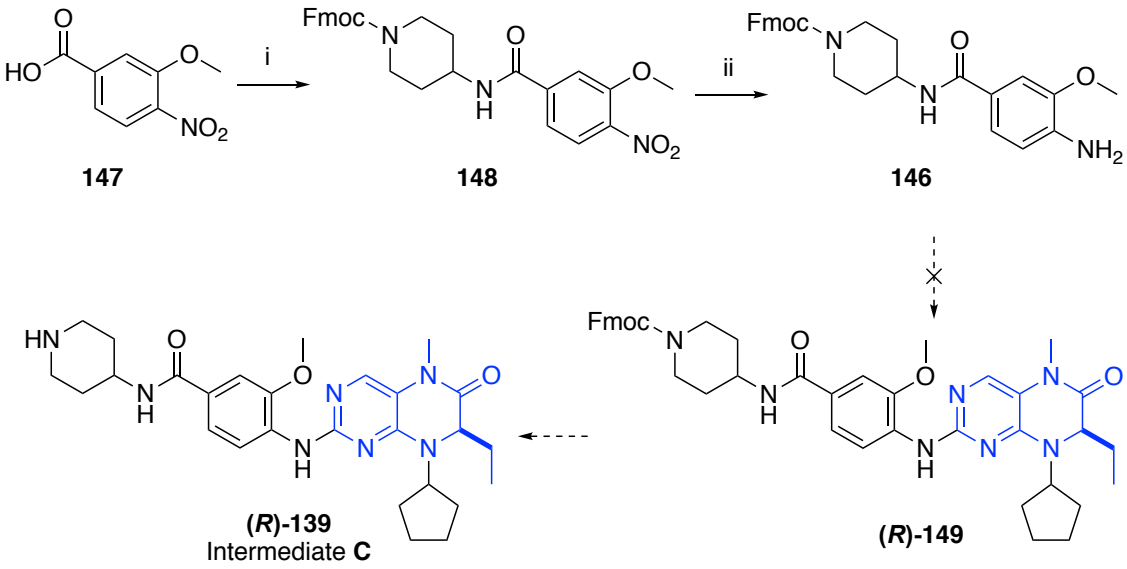


Figure 4.11: The retrosynthesis of the **(*R*)-139**, intermediate **C**, showing the possible convergent (path A, orange) or stepwise (path B, purple) synthetic routes starting from **(*R*)-92**, intermediate **A**.

4.4.3.1. First approach: convergent synthesis of Intermediate **C**: **(*R*)-139** (path A)

To make **(*R*)-139**, a *N*-protected version of the tail of the ligand **146** was first synthesised (Scheme 4.4). Amide coupling of Fmoc-protected 4-amino-piperidine and 3-methoxy-4-nitrobenzoic acid **147** afforded **148** in 99% yield. In this reaction DIPEA was used, rather than Et_3N, which avoided partial Fmoc-deprotection.[109] Next, the synthetic strategy involved reducing the nitro group, then couple the aniline intermediate to **(*R*)-92**, and finally, deprotecting **(*R*)-149** to obtain **(*R*)-139**.

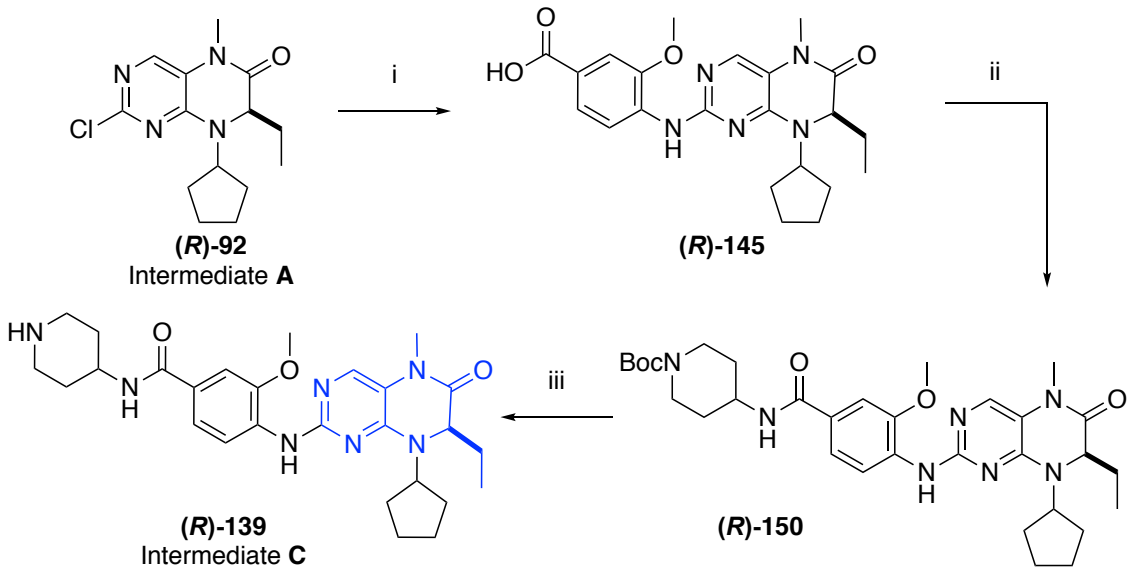
It was found, however, that **146** was insoluble likely due to the Fmoc group, preventing its effective purification and isolation. Importantly, this intermediate could not be used in the subsequent S_NAr with the **(*R*)-92** core, as it could not be solubilised in acetic acid despite using elevated temperatures. The second synthetic strategy involving a stepwise approach (Figure 4.11 path B) was therefore investigated.



Scheme 4.4: Towards the synthesis of derivatives of Intermediate C, **(*R*)-139**, using a convergent approach. The low solubility of **146** prevented its coupling with **(*R*)-92**. *Reagents and conditions*: (i). HATU, CH_2Cl_2, DIPEA, RT, 18 h, >99% (n=2). (ii). Pd/C, $H_{2(g)}$, toluene, RT, 1 h, 70% (n=1).

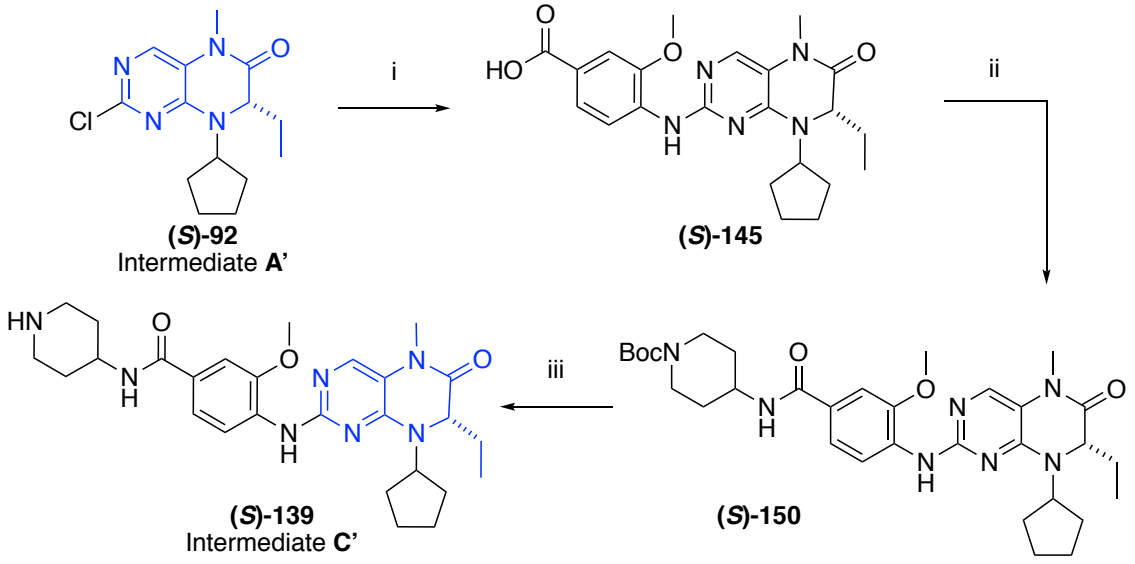
4.4.3.2. Second approach: the sequential synthesis of Intermediate C, (*R*)-139 (path B)

Synthesis of **(*R*)-139** was achieved *via* a stepwise approach (Scheme 4.5). An S_NAr reaction of 4-amino-3-methoxybenzoic acid and the **(*R*)-92** chloropyrimidine, was followed by an amide coupling with 1-Boc-4-aminopiperidine (Scheme 4.5) to give **(*R*)-150**. The **(*R*)-145** intermediate could not be efficiently purified due to its poor solubility.[98] Hence **(*R*)-145** was used without further purification in the amide coupling, and **(*R*)-150** was obtained in a 53% yield over 2 steps. Finally, Boc deprotection of **(*R*)-150** with TFA yielded the key Intermediate **C**, **(*R*)-139**, intermediate in 31–76% yield.



Scheme 4.5: Synthesis of the key secondary amine Intermediate **C**, **(*R*)-139**. *Reagents and conditions*: (i). 4-Amino-3-methoxybenzoic acid, HCl, H_2O/EtOH 5:1, 100 °C, 24 h (ii). HBTU, DIPA, 1-Boc-4-aminopiperidine, RT, 4 h, 50–53% (over 2 steps) (n=2) (iii). TFA, CH_2Cl_2, 0 °C, 1 h, 36–71% (n=3).

Overall, the secondary amine **(*R*)-139** (Intermediate **C**) was synthesised from Intermediate **A** in 3 steps. This key intermediate will be used for the synthesis of the diazirine (**(*R*)-132**) and the fluorescent tool compounds (**(*R*)-134**–**(*R*)-137**). The (*S*)-enantiomer of Intermediate **C**, **(*S*)-139**, will provide a negative control for the assays set up. Therefore, the secondary amine **(*S*)-139** was synthesised from **(*S*)-92** using the identical synthetic route to **(*R*)-139** (Scheme 4.6).

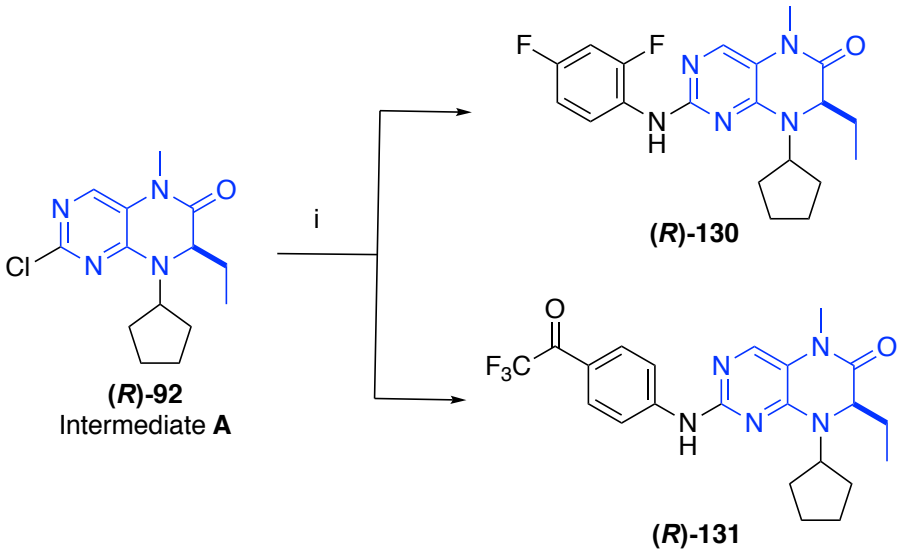


Scheme 4.6: Synthesis of the enantiomer of key secondary amine intermediate **(*S*)-139**, Intermediate **C'**. *Reagents and conditions*: (i). 4-Amino-3-methoxybenzoic acid, HCl, H_2O/EtOH 5:1, 100 °C, 24 h. (ii). HBTU, DIPA, 1-Boc-4-aminopiperidine, RT, 4 h, 72% (over 2 steps). (iii). TFA, CH_2Cl_2, 0 °C, 1 h, 57%, n=1.

Overall, intermediates **A**–**C** (**(*R*)-92**, **(*R*)-138**, and **(*R*)-139**, respectively) were synthesised. These intermediates were used to furnish the target tool compounds (**(*R*)-130** and **(*R*)-131**, Section 4.5; **(*R*)-132** and **(*R*)-133**, Section 4.6; **(*R*)-134**–**(*R*)-137**, Section 4.7) and control compounds (Section 4.8) as described below.

4.5. Synthesis of ^{19}F NMR tool compounds

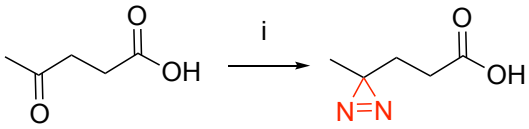
Intermediate **A**, **(*R*)-92** was further functionalised under S_NAr conditions. These conditions could not be used for electron rich anilines, as they would be fully protonated when using HCl in H_2O/EtOH conditions. However, electron poor amines, in particular fluorinated anilines, could be used. As such, the ^{19}F NMR reporter tool compounds, **(*R*)-130** and **(*R*)-131** were synthesised from **(*R*)-92** in 99 and 28% yield, respectively (Scheme 4.7).



Scheme 4.7: Synthesis of the ^{19}F NMR reporter target tool compounds **(*R*)-130** and **(*R*)-131**. *Reagents and conditions*: (i) HCl, EtOH/H_2O, 100 °C, **(*R*)-130:** 2-4-difluorononaniline, >99% (n=1); **(*R*)-131:** (trifluoroacetyl)aniline, 28% (n=1).

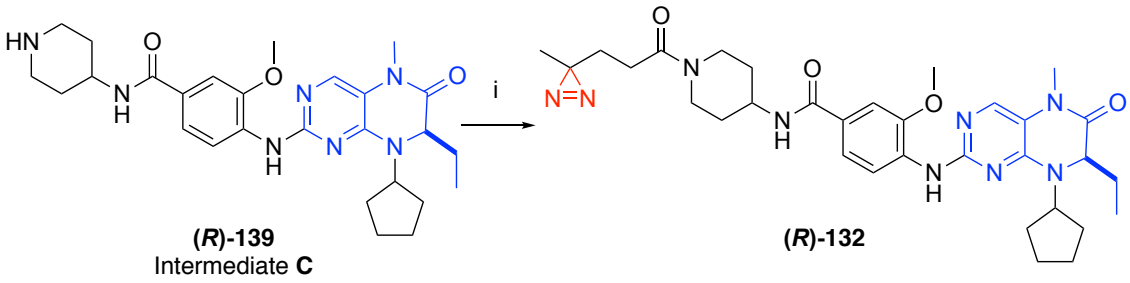
4.6. Synthesis of diazirines tool compounds

Formation of diazirines from aliphatic ketones is well documented,[110-112] and the synthetic conditions are tolerant of other functional groups, for example carboxylic groups. Hence, **151**, was chosen as the diazirine-containing moiety for the synthesis of **(*R*)-132** and **(*R*)-133**. The diazirine-containing portion of the tool compounds was prepared from levunilic acid (**152**) (Scheme 4.8). Formation of the presumed imine was carried out using ammonia in MeOH for the initial imine formation, then treatment with hydroxylamine-*O*-sulfonic acid afforded the diaziridine-containing intermediate. Subsequent oxidisation of the presumed diaziridine using iodine afforded the diazirine **151**. Multiple purification steps were necessary to remove traces of iodine, and **151** was isolated with a 22% yield.



Scheme 4.8: Synthesis of aliphatic diazirine intermediate **151**. *Reagents and conditions*: (i) (a). 1.7 M NH_3 in MeOH, 0 °C, 3 h. (b). NH_2OSO_3, RT, 48 h. (c). NEt_3, I_2, MeOH, 0 °C, 3 h, 22% (over 3 steps) (n=1).

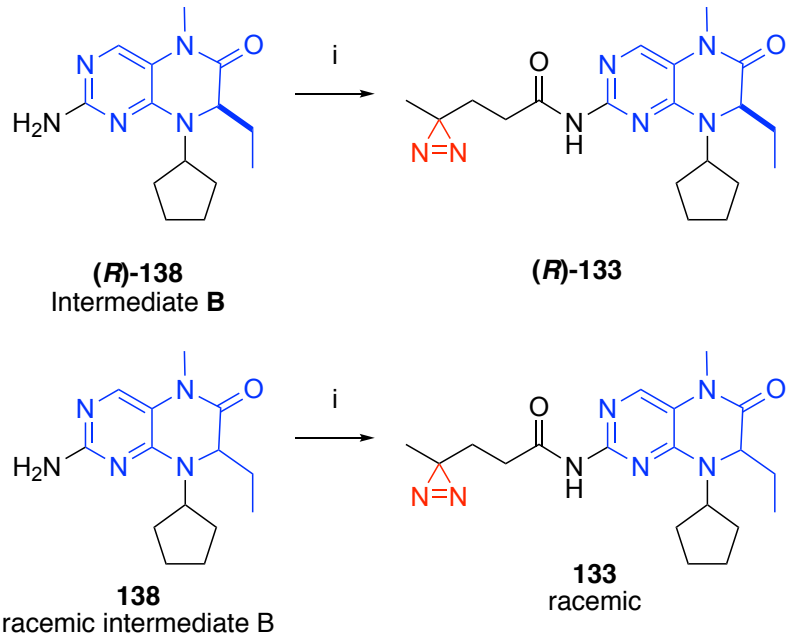
The diazirine-containing linker **151** was combined with intermediate **(*R*)-139** *via* an amide coupling to give the first diazirine target **(*R*)-132** in 53% yield (Scheme 4.9).



Scheme 4.9: Synthesis of diazirine-containing tool compound **(*R*)-132** from intermediate **C**, **(*R*)-139**. *Reagents and conditions*: (i) HATU, DIPEA, CH_2Cl_2, RT, 3 h, 53% (n=1).

The second diazirine-containing target, **(*R*)-133** was designed with a shorter distance between the diazirine photo-activatable group and the acetyl-lysine mimicking head group. The diazirine-containing tool compound **(*R*)-133**, and its racemate, were

synthesised from intermediate **B** *via* an amide coupling using HATU and DIPEA in dichloromethane (Scheme 4.10).



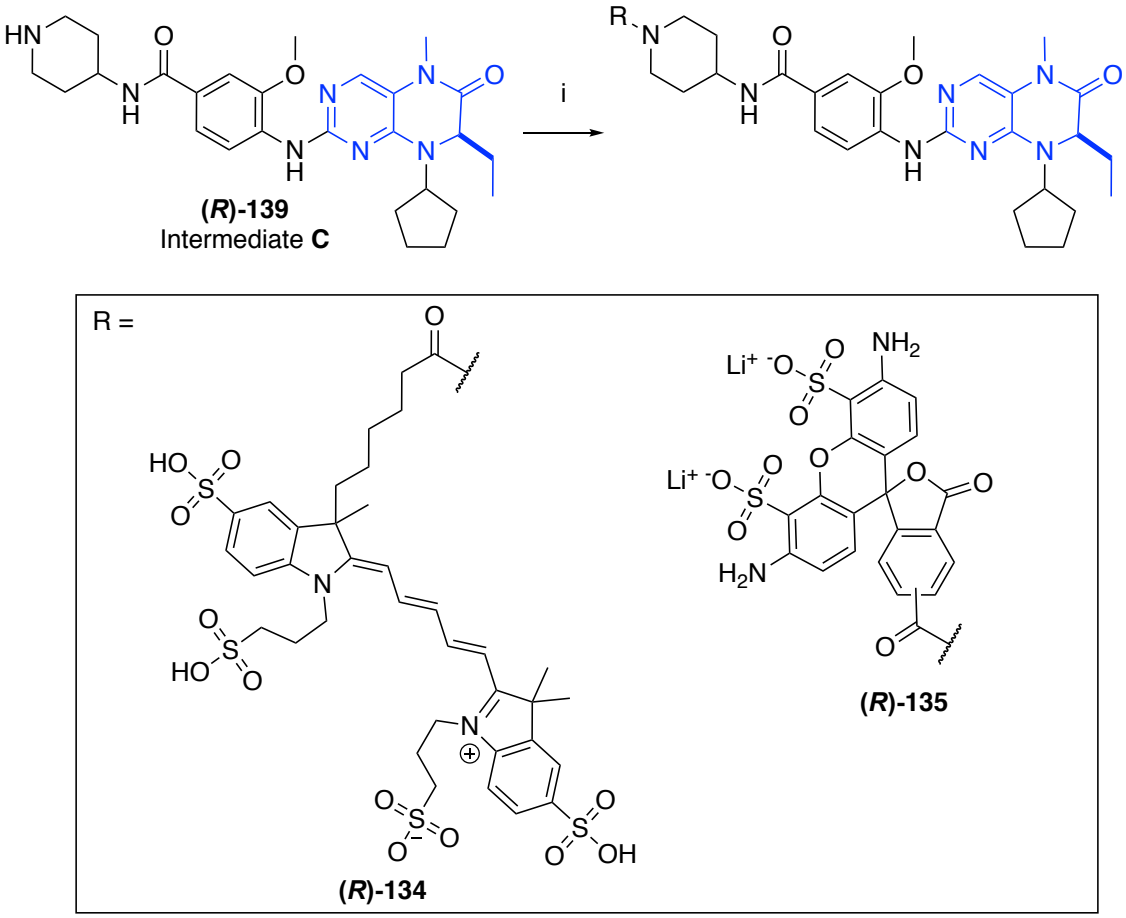
Scheme 4.10: Synthesis of second diazirine-containing tool compound target **(*R*)-133**. *Reagents and conditions*: (i) **151**, HATU, DIPEA, CH_2Cl_2, RT, with **(*R*)-138**: 18 h, 12% (n=1); with **138**: 3 h, 57% (n=1).

4.7. Synthesis of the fluorescent tool compounds

Two approaches were employed to produce the fluorophores **(*R*)-134**–**(*R*)-137**: (1) direct amide coupling of the fluorophores to the secondary amine **(*R*)-139**, or (2) installation of an alkyne group followed by click chemistry with an azide-modified fluorophore.

4.7.1. Synthesis **(*R*)-134** and **(*R*)-135** by amide coupling

The AF647 and AF488 (2.5 mg scale) were linked to the BI-2536 (**8**) skeleton, **(*R*)-139**, *via* an amide coupling, yielding **(*R*)-134** and **(*R*)-135** in 34 and 61% yield, respectively (Scheme 4.11). Reaction mixtures were purified using mass-directed assisted purification (MDAP), a preparative HPLC system, suitable for 1–50 mg scale purification. As the AlexaFluor® 488 fluorophore is provided as a mixture of 3- and 4- regiosomers, **(*R*)-135** was obtained at a mixture two regioisomers.

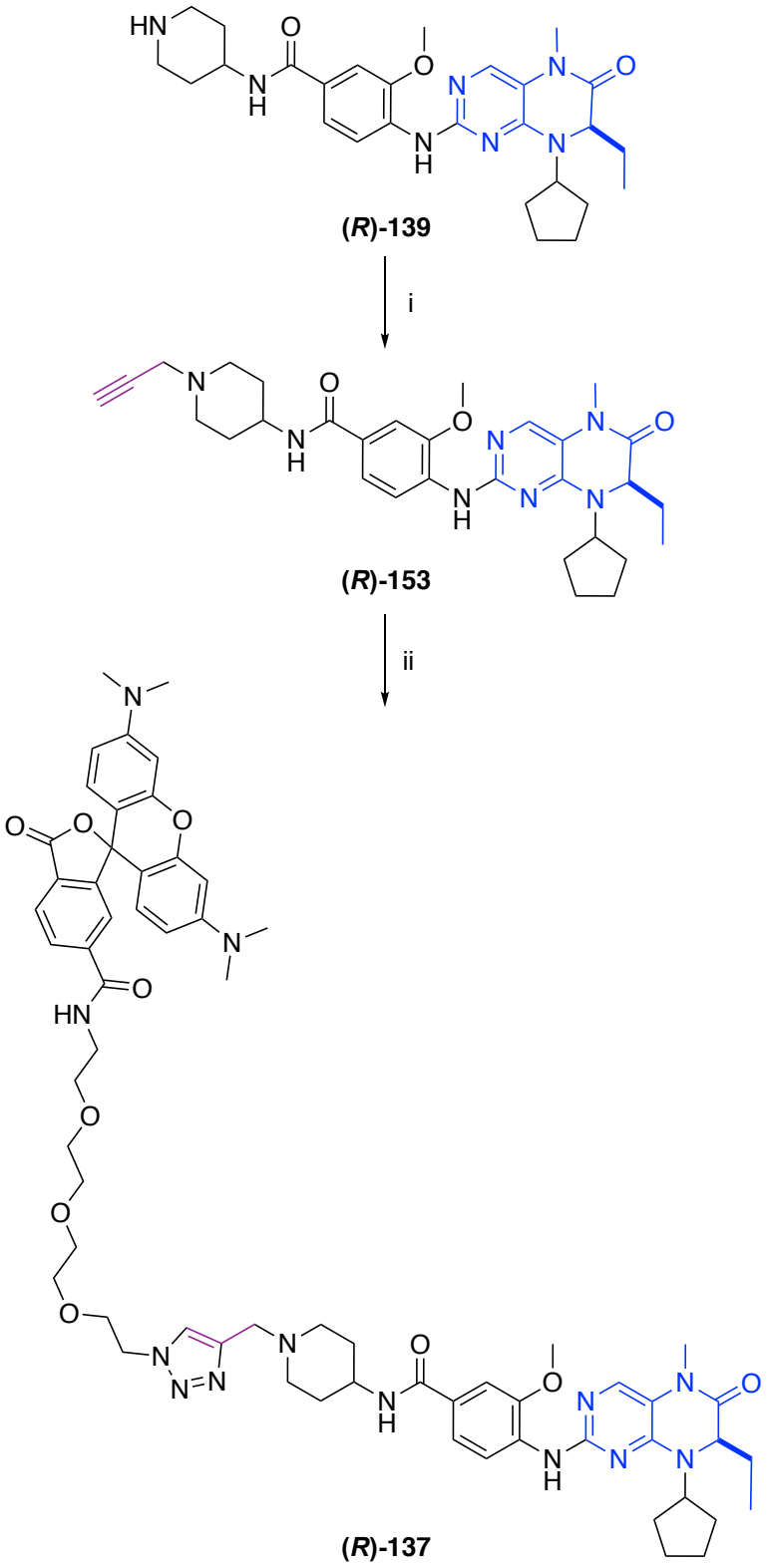


Scheme 4.11: Synthesis of the **(*R*)-134** and **(*R*)-135** fluorophores. *Reagents and conditions*: (i) R-NHS, DIPEA, DMF, RT, 18 h, RT, **(*R*)-134**: R=AF647: 34% (n=1); **(*R*)-134**: R=AF488: 61% (n=1). **(*R*)-135** was obtained at a mixture of 3- and 4-regioisomers.

4.7.2. Synthesis of **(*R*)-136** and **(*R*)-137**

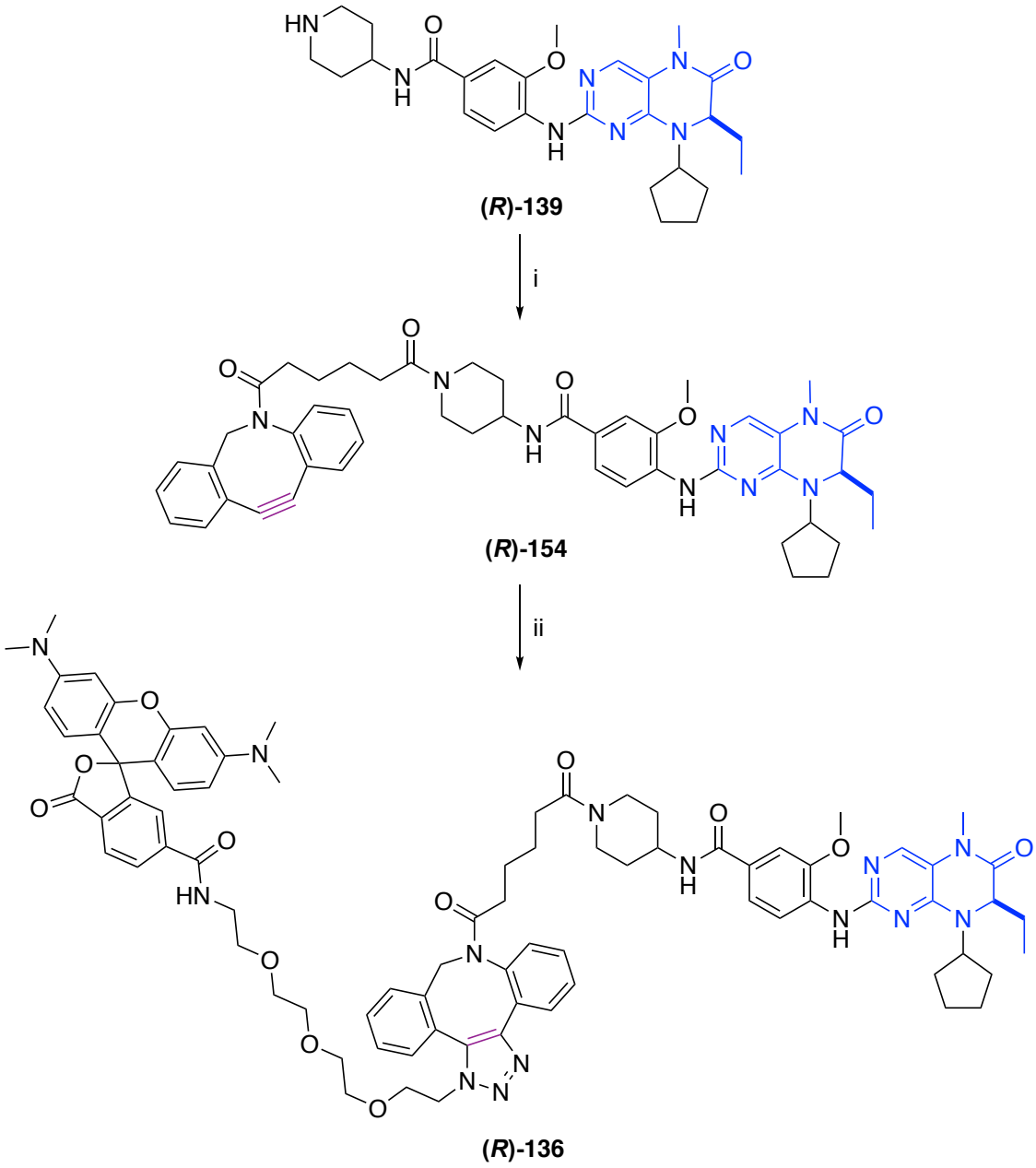
Alkylation of **(*R*)-139** using propargyl bromide and K_2CO_3 gave **(*R*)-153** in 22% yield (Scheme 4.12). The reaction was optimised by reducing the number of equivalents of base from 3 to 1.2, and by decreasing the temperature from room temperature to 0 °C. These changes limited over-alkylation and the product was isolated in a 42% yield. The alkyne **(*R*)-153** was reacted with an azide-functionalised TAMRA fluorophore, using Tris[(1-benzyl-*1H*-1,2,3-triazol-4-yl)methyl]amine (TBTA), $CuSO_4$ and sodium ascorbate in a mixture of ^{t}BuOH/H_2O/DMSO. The Click reaction gave fluorescent tool compound **(*R*)-137** in 37% yield. The TAMRA fluorophore was chosen as it is an affordable commercially available fluorophore, however, as it is a mixture of regioisomers. Consequently, **(*R*)-137**

was also obtained a mixture of two major isomers: the 1,4-disubstituted triazole with the 3- of 4-regioisomer of the fluorescein-derived fluorophore.



Scheme 4.12: Synthesis of **(*R*)-137**. *Reagents and conditions*: (i). Propargyl bromide, K_2CO_3, DMF, 0 °C, 2 h, 42%. (ii). TAMRA-PEG-N_3, TBTA, $CuSO_4$, sodium ascorbate, ^{t}BuOH/H_2O/DMSO = 2:2:1, 2 h, 37% (n=1). The **(*R*)-137** fluorophore is expected to be a mixture of two major isomers: the 1,4-disubstituted triazole with the 3- of 4-regioisomer of the fluorescein-derived fluorophore.

Next, to explore a variety of linkers, the dibenzocyclooctyne-acid (DBCO)[113] was coupled to **(*R*)-139**, to give **(*R*)-154** in 81% yield (Scheme 4.13). The strained alkyne of the DBCO linker was used in a copper-free click reaction to give the final target fluorescent tool compound **(*R*)-136**. Unlike to the copper-catalysed click reaction, the copper-free azide-alkyne addition does not exhibit regioselectivity,[114] and so a total of 4 regioisomers are predicted to be obtained following this transformation: **(*R*)-136** is a mixture of the 1,4- or 1,5-disubstituted triazole with the 3- of 4-regioisomer of the TAMRA fluorophore.

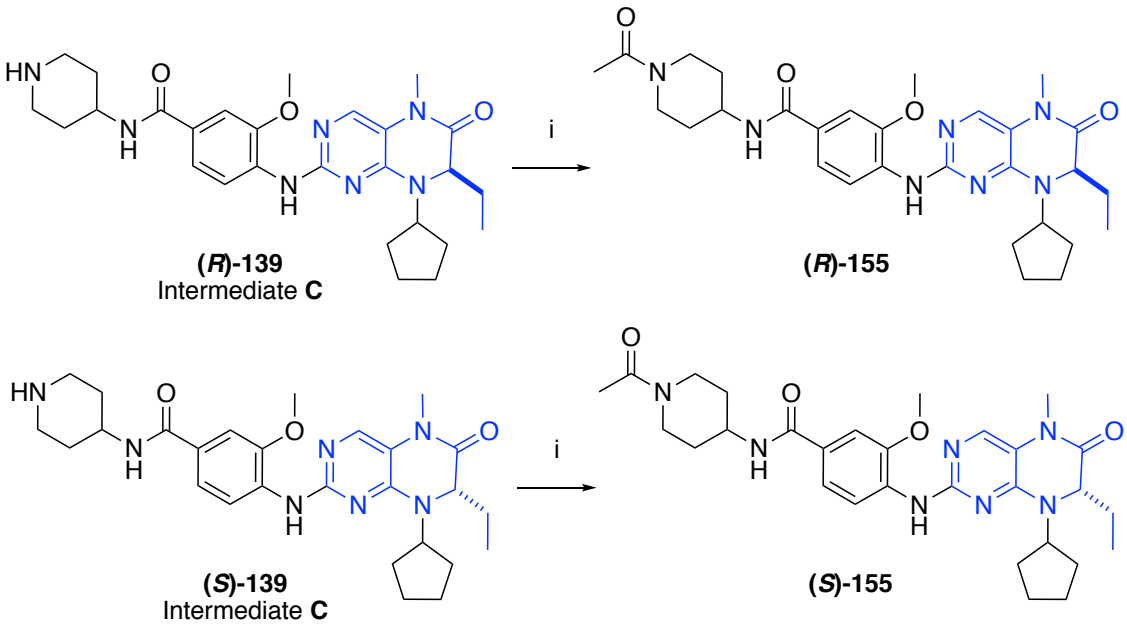


Scheme 4.13: Synthesis of **(*R*)-136**. *Reagents and conditions*: (i) DBCO-acid, HATU, DIPEA, CH_2Cl_2, RT, 1.5 h, 81% (n=1). (ii) TAMRA-PEG-N_3, MeCN/H_2O = 1:1, RT, 2 h, 34% (n=1). The **(*R*)-136** fluorophore is expected to be a mixture of four isomers: the 1,4- or 1,5-disubstituted triazole with the 3- of 4-regioisomer of the fluorescein-derived fluorophore.

4.8. Synthesis of control compounds and their validation by ITC

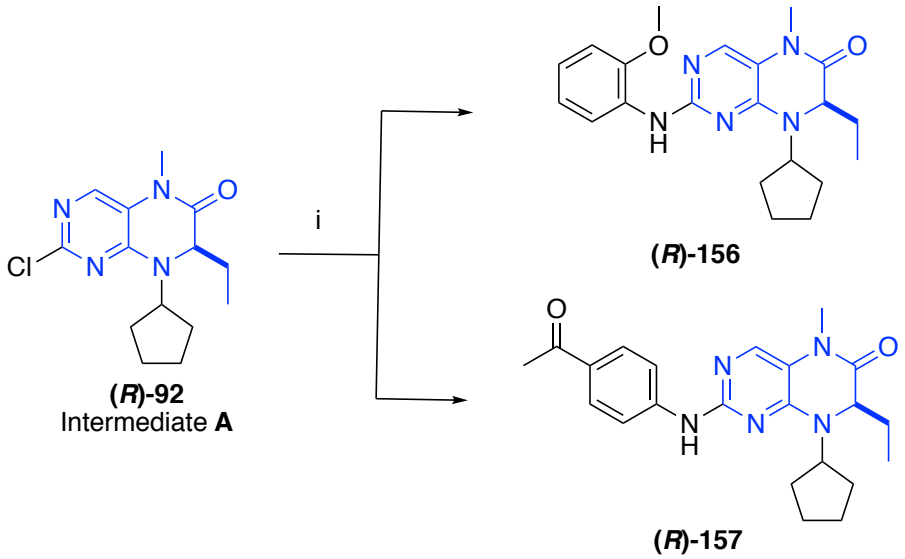
4.8.1. *N*-methylated dihydropteridinone control compounds (*R*)-155–(*R*)-157

The synthesis of key intermediates **A**–**C** (Section 4.4) enabled the synthesis of the target tool compounds (Section 4.5–4.7), but can also be used to provide analogues of BI-2536 (**8**). The first type of targets were (*R*)-**155** and (*S*)-**155**, which could be used to check using ITC whether the acetylation of BI-2536 (**8**) is tolerated for the binding to *Tc*BDF3. The analogues therefore provide a control for the amide linking strategy used for the diazirine (*R*)-**132** and fluorescent probes (*R*)-**134**–(*R*)-**136**. These control compounds. (*R*)-**155** and (*S*)-**155** were synthesised by acetylation of the corresponding (*R*)- and (*S*)-enantiomers of intermediate **C**, **139**, in 25–40% yield (Scheme 4.14).



Scheme 4.14: Synthesis of the control compounds (*R*)-**155** and (*S*)-**155** from (*R*)-**139** and (*S*)-**139**, respectively. *Reagents and conditions*: (i). AcCl, CH_2Cl_2, 0 °C, 1 h, (*R*)-**155**: 40% (n=1), (*S*)-**155**, 25% (n=1).

The second type of analogues were truncated versions of BI-2536 (**8**). These molecules were synthesised from (*R*)-**92** by Buchwald-Hartwig couplings or S_NAr to give (*R*)-**156** and (*R*)-**157** (Scheme 4.15).

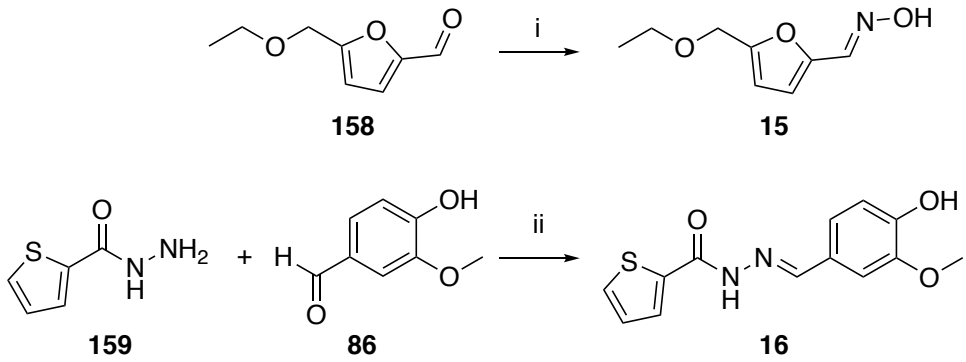


Scheme 4.15: The synthesis of the control compounds **(*R*)-156** and **(*R*)-157** from intermediate A, **(*R*)-92**. These simple analogues of BI-2536 (**8**) could be used to check how much the binding of BI-2536 (**8**) to *Tc*BDF3 is affected by truncation of its tail. *Reagents and conditions*: (i) 2-Methoxyaniline, XPhos, $Pd_2(dba)_3$, K_2CO_3, ^tBuOH, 100 °C, 20 h, 31% (n=1) (ii) 4-Aminoacetophenone, HCl, EtOH/H_2O, 100 °C, 44 h, 64% (n=1).

4.8.2. Literature control compounds **15** and **16**

During the course of this work, Serra and co-workers published two studies identifying new *Tc*BDF3 binders.[69-70] The first compound, **15**, was isolated from a chemically engineered extract, which was obtained by reacting a herbal crude extracts with hydroxyl amine.[69] The second compound, **16**, was identified through dynamic combinatorial chemistry.[70] These compounds are both of low molecular weight, and are not obvious KAc mimics. They were characterised by Serra by an intrinsic tryptophan fluorescent assay; however, the results were not corroborated using any orthogonal assays. In an early report from Serra and co-workers this assay was used to determine a K_D value of 27.1 µM for the binding of I-BET151 (**5**) and *Tc*BDF3.[68] These results report an affinity 10-fold higher than that independently determined by the NMR titration, as described in Section 3.3.4.1. It is important to note that the Serra group used the full-length construct of *Tc*BDF3, which is insoluble when expressed recombinantly and purified from the inclusion bodies. Consequently, we were interested in synthesising and validating **15** and **16** in the *Tc*BDF3 assays established as part of this work.

Synthesis of **15** and **16** followed the synthetic routes detailed in the recent publications from the Serra and co-workers (Scheme 4.16).[69-70] **158** was treated with hydroxylamine to obtain **15** in 14% yield. Purification of **15** was challenging due to its tendency to isomerise to the (*Z*)-oxime. Isomerisation was observed when the product was stored neat or as a DMSO stock. In particular, over 25% isomerisation was observed within a week as a DMSO stock, which raised concerns about the stability of **15** in buffer. The second proposed *Tc*BDF3 binder, **16** was synthesised from **159** and *o*-vanillin **86** in 70% yield. This compound was originally isolated from a combinatorial library and could be hydrolysed back to starting material, however, its low solubility in water allowed for efficient product isolation.



Scheme 4.16: Synthesis of the *Tc*BDF3 binders, **15** and **16**, reported by the Serra group.[69-70] *Reagents and conditions:* (i) Hydroxylamine hydrochloride, EtOH, 0 °C, 1.5 h, 14% (n=1); (ii) TFA, MeOH, RT, 4 h then H_2O, RT, 18 h, 70% (n=1).

4.8.3. Validation of the BI-2536 (**8**) analogues and control compounds by ITC

Validation the binding of the compounds synthesised in Section 4.7 was required to determine if they could be used during the set-up of novel assays using the tool compounds (Chapter 4). Indeed, **(*R*)-155** and **(*S*)-155**, provide information on the amide linkage on the piperidine group. **(*R*)-156** and **(*R*)-157** are truncated versions of BI-2536 (**8**) and were expected to retain some binding to *Tc*BDF3. In addition, literature compounds **15** and **16** are described to have K_D values below 2 µM in a tryptophan fluorescence assay (**15**: K_D=0.29 µM, **16**: K_D=1.7 µM).[68-70] Therefore, ITC was chosen to assess the binding of these compounds, as the work previously described in Chapter 3, characterising the binding of BI-6727 (**9**) to *Tc*BDF3, had shown detection of binding up to 50–75 µM. It is

therefore expected that this assay should detect the expected K_D values of all the control compounds: **15**, **16** and **(*R*)-155**–**(*R*)-156** (Figure 4.12).

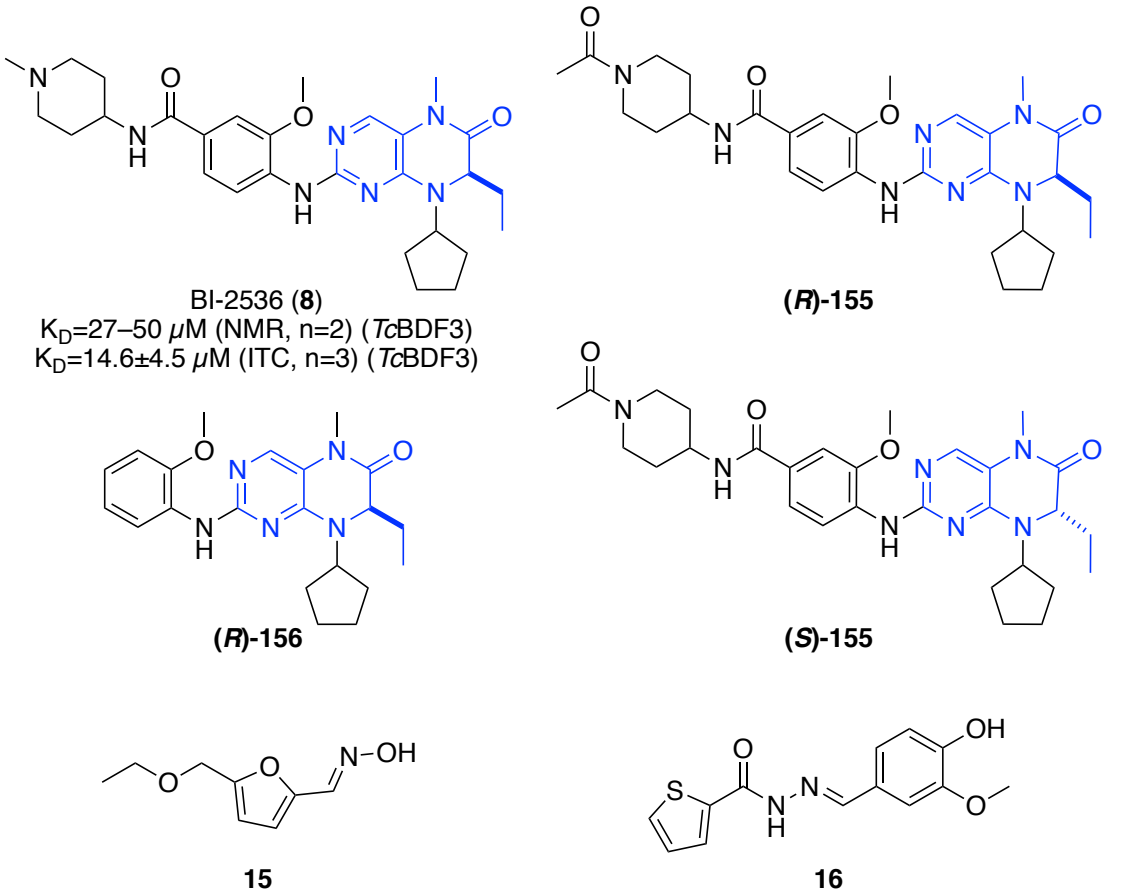


Figure 4.12: The structures of BI-2536 (**8**) and other control compounds. **155**, and **(*R*)-156** are closely related to BI-2536 (**8**), and so were expected to retain binding to *Tc*BDF3. **(*S*)-155** is not expected to bind, as the (*S*)-ethyl group is thought to be detrimental to binding as shown by docking in the homology model (Joseph Bluck). The literature compounds **15** and **16** are described to have K_D values below 2 μM.

Acetylated analogue **(*R*)-155** was shown to bind to *Tc*BDF3 with a K_D of 15.9±8.44 μM, which matches wells the affinity BI-2536 (**8**) (Table 4.2, Figure 4.12). These results suggest that the amide linkage to the piperidine group is tolerated. The opposite enantiomer, **(*S*)-155**, does not show binding and can therefore be used as a negative control. Moreover, the truncated version of BI-2536 (**8**), **(*R*)-156**, also shows binding, with a 2-fold decrease in affinity. Interestingly, neither **15** nor **16** showed any detectable binding by ITC. The two sets of ligand:protein ratios employed were predicted to detect binding in the range of 0.0001–200 μM. This raises the question of whether these non-KAc mimic are false positives identified by the intrinsic tryptophan fluorescence assays used by the Serra research group.[68-70]

Table 4.2: Summary of the K_D values determined by ITC for the compounds synthesised in Section 4.8. The literature compounds, **15** and **16**, reported by Serra and co-workers[69-70] did not show binding to *Tc*BDF3 by ITC, using concentrations ratios to detect tight to low binding affinities. The ITC traces can be found in Appendix B4.

Ligand	[ligand]	[*Tc*BDF3]	K_D (ITC)	K_D (orthogonal assay)	comments
BI-2536 (**8**)	40 μM	400 μM	14.6±4.5 μM	27–45 μM (NMR)	n=3
BI-6727 (**9**)	40 μM	400 μM	57.1±70.5 μM	–	n=3,
(*R*)-155	40 μM	400 μM	15.9±8.44 μM	–	n=1
(S)-155	40 μM	400 μM	No binding	–	n=1
(*R*)-156	40 μM	400 μM	44.7±2.25 μM	–	n=2
15	20 μM*	160	No binding	0.29 μM (Trp fluorescence)[69]	n=2
	50 μM**	500 μM	No binding		n=1
16	20 μM*	160 μM	No binding	1.7 μM (Trp fluorescence)[70]	n=2
	50 μM**	500 μM	No binding		n=1

* ITC conditions designed to detect binding in the range of 0.1–1000 nM;** ITC conditions designed to detect binding in the range of 0.1–200 μM.

4.9. Summary and conclusions

Following expression of the *T. cruzi* BD proteins, *Tc*BDF2, *Tc*BDF3, and *Tc*BDF5-1, waterLOGSY was used to identify KAc mimicking containing binders. The proton ligand observed protein titration described in Chapter 3 (Section 3.3.4), was used to determine K_D ranges to rank and prioritise compounds. Of particular interest was the dual PLK1/BRD4 ligand BI-2536 (**8**) which showed the tightest binding to *Tc*BDF3, with a K_D of 27–45 μM.

It became obvious at this point in the project that there was a need for higher throughput assays. The work described in this Chapter confirmed the binding of the *N*-methylated dihydropteridinone core by NMR titration, and its K_D of 335–365 μM gave further evidence that the binding of BI-2536 (**8**) arises from its KAc mimicking headgroup. Therefore, ^{19}F NMR, diazirine, and fluorescent were designed in order to set up orthogonal and higher throughput assays. The *Tc*BDF3 homology model was used to hypothesise the likely mode of binding and to determine potential points of modification of the BI-2536 (**8**) ligand. Then, three common synthetic intermediates were identified for the synthesis of the tool compounds. Following synthesis of the intermediates, the three categories of tool compounds and control compounds were synthesised.

The synthesis of tool compounds for *Tc*BDF3 opens up the possibility to set up higher throughput assays, that are orthogonal to the initial waterLOGSY and ^{1}H NMR titration experiments. This will allow for the screening of ligands using the displacement of the tool compounds monitored by ^{19}F NMR, protein LCMS or fluorescent plate reader. Additionally, the synthetic routes to the key intermediates and target tool compounds can be used to synthesise analogues of BI-2536 (**8**) as control compounds or to study the SAR.

Chapter 5. New biophysical assays using the *Tc*BDF3 tool compounds

5.1. Introduction and aims

The aim of the work in this chapter was to develop biophysical assays to study *Tc*BDF3 binding using the tool compounds described in Chapter 4. In particular, medium- to high-throughput displacement-based assays were sought. These were needed to provide alternatives to the low-throughput, high protein requirements, of the waterLOGSY and ^{1}H NMR titration experiments (Chapter 3).

5.2. ^{19}F NMR assays

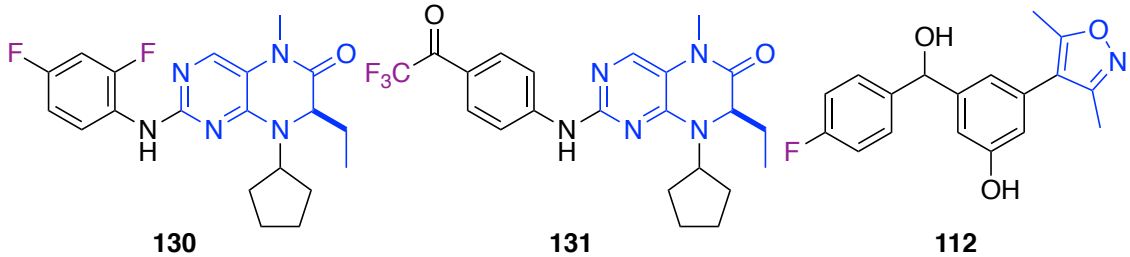
^{19}F-containing molecules have the potential to be used as reporters to detect ligand binding, using ^{19}F NMR, which has shorter experiment times than ^{1}H NMR. The main advantages of this approach are that fluorine atoms do not commonly occur in proteins and that ^{19}F has a high natural abundance. Therefore, tracking the binding of fluorinated ligand to a protein will not suffer from increasing protein background signal, which could mask the studied resonances. The displacement of a well-characterised fluorinated ligand implies at least partial overlap of the binding mode of the reporter molecule and the competitor, giving more information compared to waterLOGSY. Additionally, this technique requires less protein than ITC, and, depending on the concentration used, would only detect binders of equal or lower than K_D values than the reporter molecules. This latter point implies that weak binders (K_D > 500 μM), detectable by waterLOGSY, would be not show displacement unless used in a large excess, and so could be filtered out. Consequently it was of interest to characterise the binding of the ^{19}F NMR reporter molecules **(*R*)-130** and **(*R*)-131** to the *Tc*BDFs.

5.2.1. ^{19}F NMR probes for *Tc*BDF3

5.2.1.1. The binding of **112**, **(*R*)-130**, and **(*R*)-131** to *Tc*BDF3

OXFBD02 (**7**) binds to *Tc*BDF2, 3 and 5-1 in waterLOGSY experiments. BI-2536 (**8**) has only been confirmed to bind *Tc*BDF3. Consequently, the use of a DMI-based fluorinated ligand, such as **112**,[58] would offer the possibility to study of all three proteins (Figure 5.1). First, using the method established by Dr Anthony Chan (University of Oxford), the binding of **112** was assessed by ^{19}F NMR. The change in chemical shift and peak width shows that the ^{19}F resonance experienced a change in environment, which indicates binding to H_6-*Tc*BDF3 and H_6-*Tc*BDF5-1 (Figure 5.2). The addition of H_6-*Tc*BDF2 doesn't induce changes in the ^{19}F NMR spectrum of **112**, even when using a 1.2:1 protein:ligand ratio. Therefore, **112** does not bind H_6-*Tc*BDF2 with a measurable affinity.

Interestingly, the ^{19}F NMR resonances of **112** splits into two symmetrical peak, around its original chemical shift bound to *Tc*BDF3 (Figure 5.2B) or *Tc*BDF5-1 (Figure 5.2C). The change of peak shape is likely to arise from the binding of the two enantiomers of **112**, that experience distinct environments, but bind the protein with similar binding affinities. This behaviour was previously observed during the ^{1}H ligand-observed protein titration experiment to determine the K_D of OXFBD02 (**7**) to *Tc*BDF3 (Chapter 3) and *Tc*BDF5-1 (Appendix B3).



130 **131** **112**

Figure 5.1: The structures of the ^{19}F NMR tool compounds, **(*R*)-130**, **(*R*)-131**, and **112**. **112** was synthesised by R. Berkeley or C. Evans, following the procedure of Jennings *et al.*[58] The KAc mimic is shown in blue and the fluorine atoms are shown in purple.

This initial work using **112** was followed by optimisation of the ^{19}F NMR reference by switching from trifluoroacetone to potassium fluoride (Appendix B5).

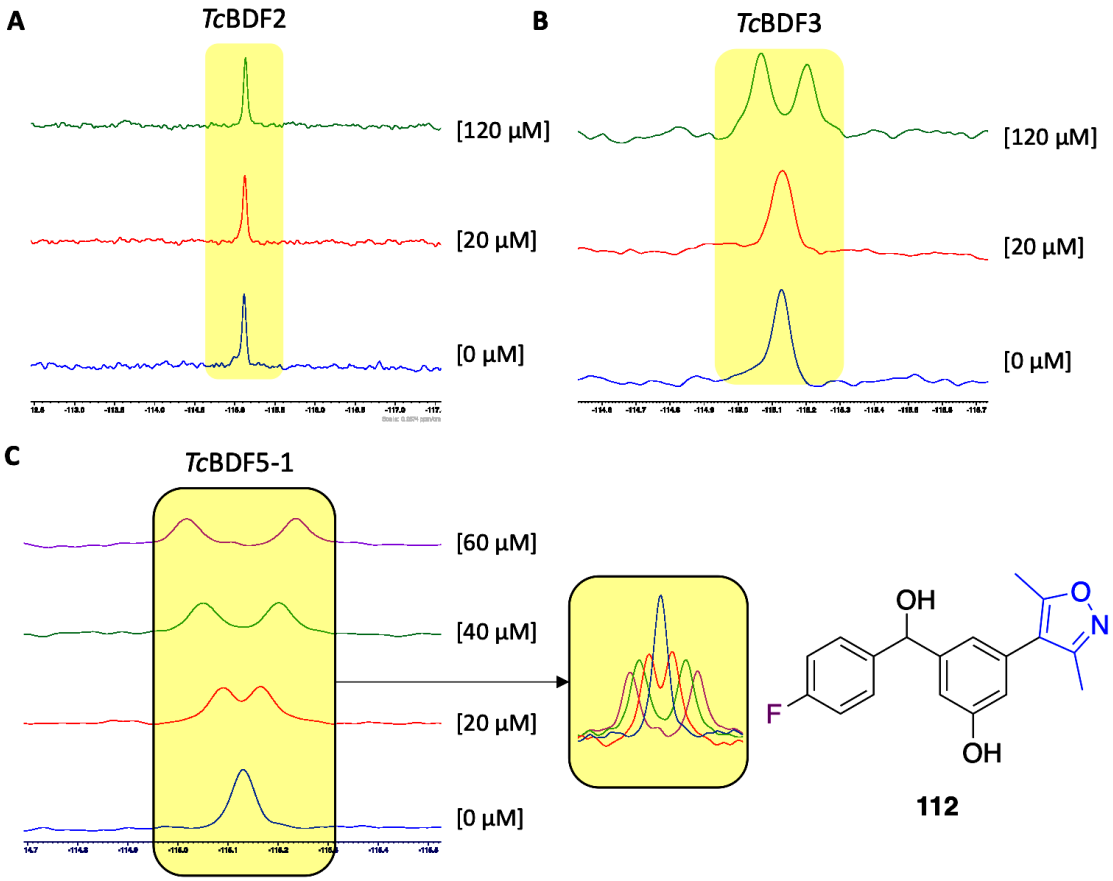


Figure 5.2: The binding of **112** to *Tc*BDF2, *Tc*BDF3 and *Tc*BDF5-1 monitored by ^{19}F NMR, using trifluoroacetone as a reference. Compound only spectrum is shown in blue. (A) No significant change is seen for the addition of 20 (red) or 100 µM (green) of *Tc*BDF2 (B) Addition of *Tc*BDF3 at 20 (red) or 120 µM (green) show a gradual split of the **112** ^{19}F NMR resonance into two peaks of equal intensities. This indicates binding. The shape is thought to arise from distinct binding environments of the 2 enantiomers (C) Similarly, addition of *Tc*BDF5-1 at 20 (red), 40 (green) or 60 µM (purple) shows a gradual split of the **112** peak. **112** was synthesised by Dr L. Jennings, R. Berkeley, A. Hicks or Dr C. Evans (University of Oxford), following the procedure of Jennings *et al*.[58]

Following the use of the DMI-based ^{19}F NMR reporter molecule, **112**, it was of interest to determine if the fluorinated tool compounds, **(*R*)-130** and **(*R*)-131**, would offer advantageous binding to *Tc*BDF3. For example, if **(*R*)-130** and **(*R*)-131** have higher affinity or clearer signals. Therefore, their binding was characterised by ITC titration with *Tc*BDF3. Pleasingly, **(*R*)-130** and **(*R*)-131** showed binding to *Tc*BDF3 with K_D values of 62.4±96.7 µM and 70.8±145 µM, respectively (Table 5.1). These binding affinities are comparable affinities to BI-6727 (**9**), and only a 4-fold decrease compared to BI-2536 (**8**). The next step was to determine if the binding could also be detected by ^{19}F NMR.

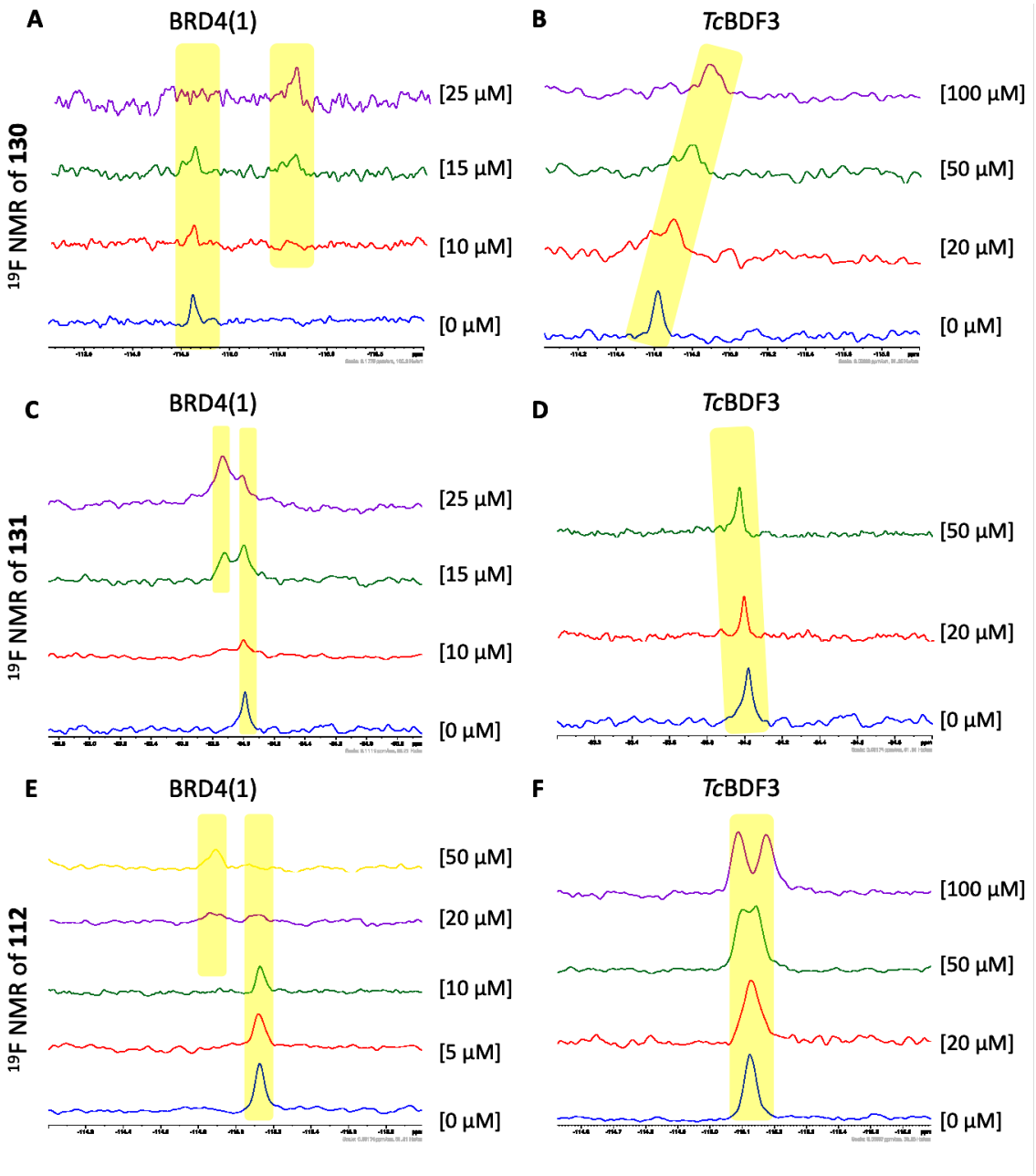


Figure 5.3: The overlay of 564 MHz ^{19}F NMR spectra showing the changes (highlighted in yellow) of width and chemical shift of fluorine resonances of **(*R*)-130**, **(*R*)-131**, and **112** were tested in the presence of BD proteins: BRD4(1) and *Tc*BDF3, in HEPES buffer The spectra have been calibrated to the resonances of 100 µM of KF, the compound only spectra are shown in blue. (A) **(*R*)-130** titrated with BRD4(1) at 10 µM (red), 15 µM (green) and 25 µM (purple) shows the appearance of a distinct new peak. (B) **(*R*)-130** titrated with *Tc*BDF3 at 20 (red), 50 (green) and 100 µM (purple) shows a gradual change in chemical shift. (C) **(*R*)-131** titrated with BRD4(1) at 10 (red), 15 (green) and 25 µM (purple) shows the appearance of an overlapping new peak. (D) **(*R*)-131** titrated with *Tc*BDF3 at 20 µM (red), and 50 µM (green) shows a gradual change in chemical shift. (E) **112** titrated with BRD4(1) at 5µM (red), 10 µM (green), 20 µM (purple), and 50 µM (yellow) shows a gradual change in chemical shift. (F) **112** titrated with *Tc*BDF3 at 20 µM (red), 50 µM (green) and 100 µM (purple) show a gradual spilt of the resonance into two peaks of equal intensities. The shape is thought to arise from distinct binding environments of the 2 enantiomers, which bind with similar K_D values.

Table 5.1: Summary of the ITC titration data for the ^{19}F NMR tool compounds, **130** and **131**, and the K_D values of BI-2536 (**8**) and BI-6727 (**9**) (Chapter 3) for comparison. Values are given with the standard deviation between triplicates (*), or in the case of n=1, the fitting error (**) is shown. The ITC traces are shown in Appendix B4.

Ligand	[ligand]	[*Tc*BDF3]	K_D (ITC)	comments
BI-2536 (**8**)	40 µM	400 µM	14.6±3.0 µM	n=3*
BI-6727 (**9**)	40 µM	400 µM	57.1±4.9 µM	n=3*
(*R*)-**130**	40 µM	400 µM	62.4±96.7 µM	n=1**
(*R*)-**131**	40 µM	400 µM	70.8±145 µM	n=1**

(*R*)-**130**, (*R*)-**131**, and **112** were tested using ^{19}F NMR in the presence of BRD4(1), *Tc*BDF2, *Tc*BDF3, and *Tc*BDF5-1. The three ligands showed binding to BRD4(1) and *Tc*BDF3 (Figure 5.3), and **112** also showed binding to *Tc*BDF5-1. (*R*)-**130** and (*R*)-**131** did not show significant binding to *Tc*BDF2 nor to *Tc*BDF5-1 (Appendix B5).

5.2.1.2. Displacement of ^{19}F NMR reporter molecules to detect *Tc*BDF3 binding

The ^{19}F NMR reporter molecules (*R*)-**130** and **112** showed more pronounced changes during the titration with *Tc*BDF3, and so were selected for evaluation in displacement experiments with non-fluorinated ligands. The displacement of **112** from *Tc*BDF3 by BI-2536 (**8**) was demonstrated by recovering a single peak rather than a split peak. An analogue of BI-2536 (**8**), **160** (Figure 5.4), designed as part of this work, and synthesised by Michelle Keller (unpublished data) was also validated using this technique.[95] The treatment of the pre-incubated mixture of **112** and *Tc*BDF3 with **160** gives a broad peak, resembling two overlapping peak, rather than a distinct split peak as is usually observed at these concentration of **112** and *Tc*BDF3. This work shows that **160** binds *Tc*BDF3, which could not be confirmed by waterLOGSY due to limited solubility. The same experiment carried out with (*R*)-**130** did not show as clear result, with only a small change in the peak width.

Following the ^{19}F NMR results, the binding of **160** to *Tc*BDF3 was confirmed by ITC, giving a K_D value of 18.6±8.4 µM (ITC, n=2) (Appendix B4). These results shows that (*R*)-**130** is less suited than **112** to characterise the displacement for *Tc*BDF3, as it did not clearly detect the binding of either BI-2536 (**8**) or **160**. **112** is expected to bind less tightly to *Tc*BDF3 than (*R*)-**130**, as it is a close analogue of OXFBD02 (**7**) which binds *Tc*BDF3 with a

K_D of 418–508 µM (NMR, n=2). This implies that it is more easily displaced. The characteristic shape of the fluorine resonance of **112** bound to *Tc*BDF3 and its expected lower K_D make **112** a useful a ^{19}F NMR reporter molecule for the study of both *Tc*BDF3 and *Tc*BDF5-1.

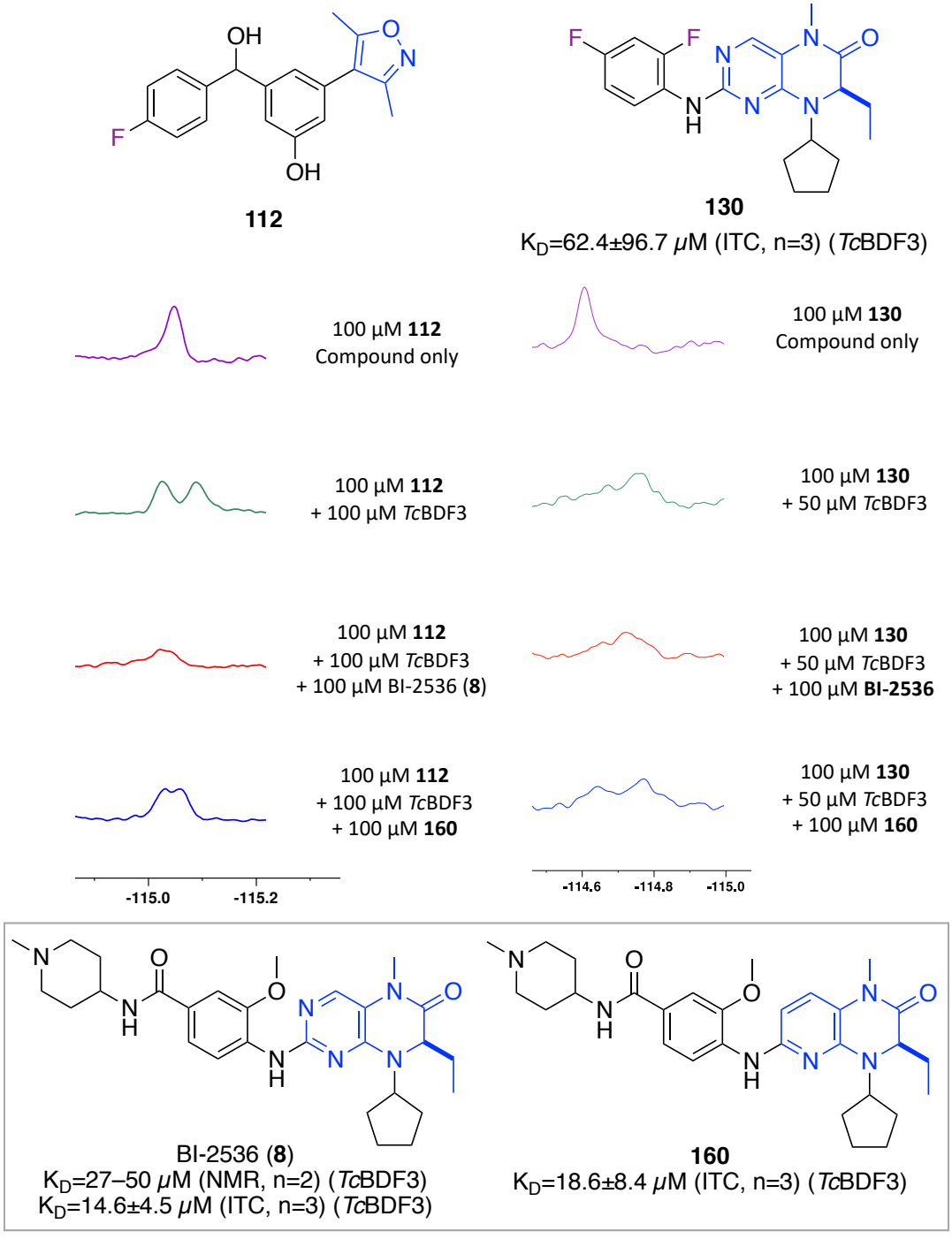


Figure 5.4:The displacement of **112** and **(*R*)-130** from *Tc*BDF3 by BI-2536 (**8**) and **160** monitored by ^{19}F NMR. The addition of *Tc*BDF3 to **112** results in a split peak (or clear doublet) (green). The displacement of **112** by BI-2536 (**8**) (red) recovers a single peak and **160** (blue) shows a transition between a single peak and a split peak (doublet). The displacement of **(*R*)-130** by BI-2536 (**8**) (red) and **160** (blue) gives only small change. **160** was synthesised by Michelle Keller (University of Oxford).

5.2.2. Fluorine probes for bromodomains from other parasites

Leishmania donovani is the main cause of visceral leishmaniasis, the most severe type of leishmaniasis.[3] The systemic infection causes the enlargement of vital organs, and if untreated, this disease is fatal. Several BD-containing proteins have been identified in the *L. donovani* parasite, of which three (*Ld*BD2, *Ld*BD5-1 and *Ld*BD5-2) have been expressed recombinantly by Dr Nathaniel Jones (Wilkinson group, University of York). The sequence identity between the BDs homologues from *L. donovani* and *T. cruzi* ranges between 42 and 53%. Additionally, *Ld*BD5-1 has been co-crystallised with the BI-2536 (**8**) ligand (PDB: 5TCM). Taken together, these data suggest that the tool compounds and biophysical assays developed for the *T. cruzi* BDs might also be useful for the study of these proteins. Therefore, it was decided to screen the fluorinated ligand **112**, **(*R*)-130**, and **(*R*)-131** against *Ld*BD2, *Ld*BD5-1 and *Ld*BD5-2 using ^{19}F NMR. Width and chemical shift differences were observed for **(*R*)-130**, upon addition of the *L. donovani* BDs (Figure 5.5). Minimal changes were observed for **(*R*)-131** and no changes for **112** (Appendix B5).

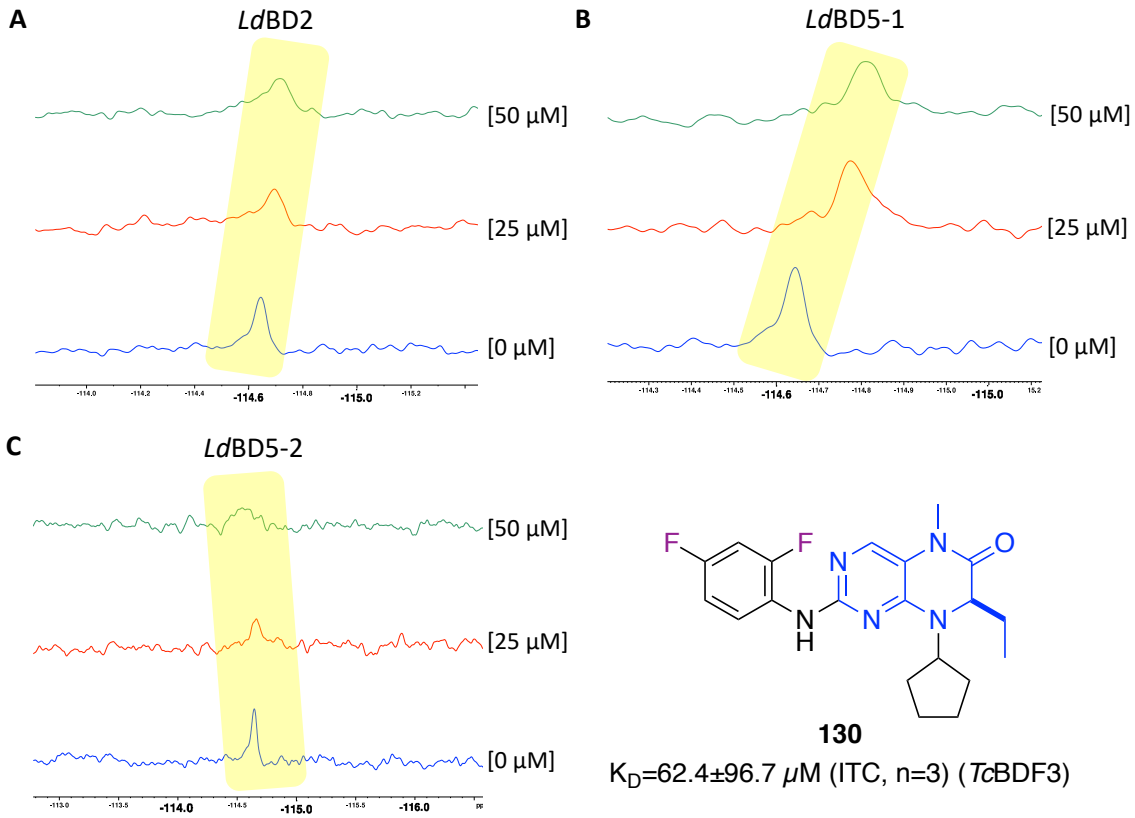


Figure 5.5: The overlay of ^{19}F NMR spectra showing the changes (highlighted in yellow) of width and chemical shift of fluorine resonances of **(*R*)-130** in the presence of (A) *Ld*BD2, (B) *Ld*BD5-1 and (C) *Ld*BD5-2. The observed changes indicated binding of the fluorinated probes to the proteins. Compound-only spectrum in blue, spectrum with 25 or 50 µM protein in red and green, respectively.

Overall the fluorinated ligands **112**, **(*R*)-130**, and **(*R*)-131** can be used to study the binding to various BD proteins. **112** in particular has shown binding to *Tc*BDF2, *Tc*BDF3 and *Tc*BDF5-1 and can be displaced in competition experiments.

5.3. Photo-crosslinking assays with diazirines probes

Photo-crosslinking of diazirines can be seen as a way to freeze the equilibrium of the binding event (ligand/protein). The diazirine probes bind to the protein (equilibrium state), then the diazirine forms a carbene upon light irradiation. The carbene is quenched by water or cross-links irreversibly to the protein by inserting into nearby C-H, C-O or N-H bonds (Figure 5.6).[106]

$$\mathbf{L + P} \overset{(1)}{\rightleftharpoons} \mathbf{L \cdot P} \overset{(2)}{\rightleftharpoons} \mathbf{L^{*} \cdot P} \overset{(3)}{\rightleftharpoons} \mathbf{L{-}P}$$

Figure 5.6: The photo-crosslinking of diazirines to a protein. (1) the binding of the diazirine ligand (L) to the protein (P) depends on the binding affinity, and is characterised by the equilibrium constant K_D (2) The bound diazirine ligand can be photo-activated by light irradiation forming a carbene (L*) (3) The bound carbene crosslinks to the protein to give a protein adduct of mass M(protein)+M(Diazirine ligand)-M(nitrogen)+M(hydrogen) than can de detect by protein LCMS. In steps (1) and (2) the ligand can dissociate from the protein, however, following step (3) the ligand is irreversibly cross-linked to the protein.

The expected LCMS outputs are the following: no cross-linking, high to low amounts of cross-linking (mass of protein + mass of diazirine tool compounds – mass of nitrogen + mass of hydrogen), or multiple cross-linking events (Figure 5.7). The latter case is expected for non-specific binders, for example lipophilic compounds, that bind the protein surface by hydrophobic contact, rather than the binding site through polar contacts.

The first priority was to determine whether if **(*R*)-132** and **133** cross-link to the proteins, then to characterise the minimum amounts of probe required to obtain the maximum cross-linking yield. Following this, the focus was to determine if the cross-linking event could be prevented by the addition of a known binder prior to light irradiation.

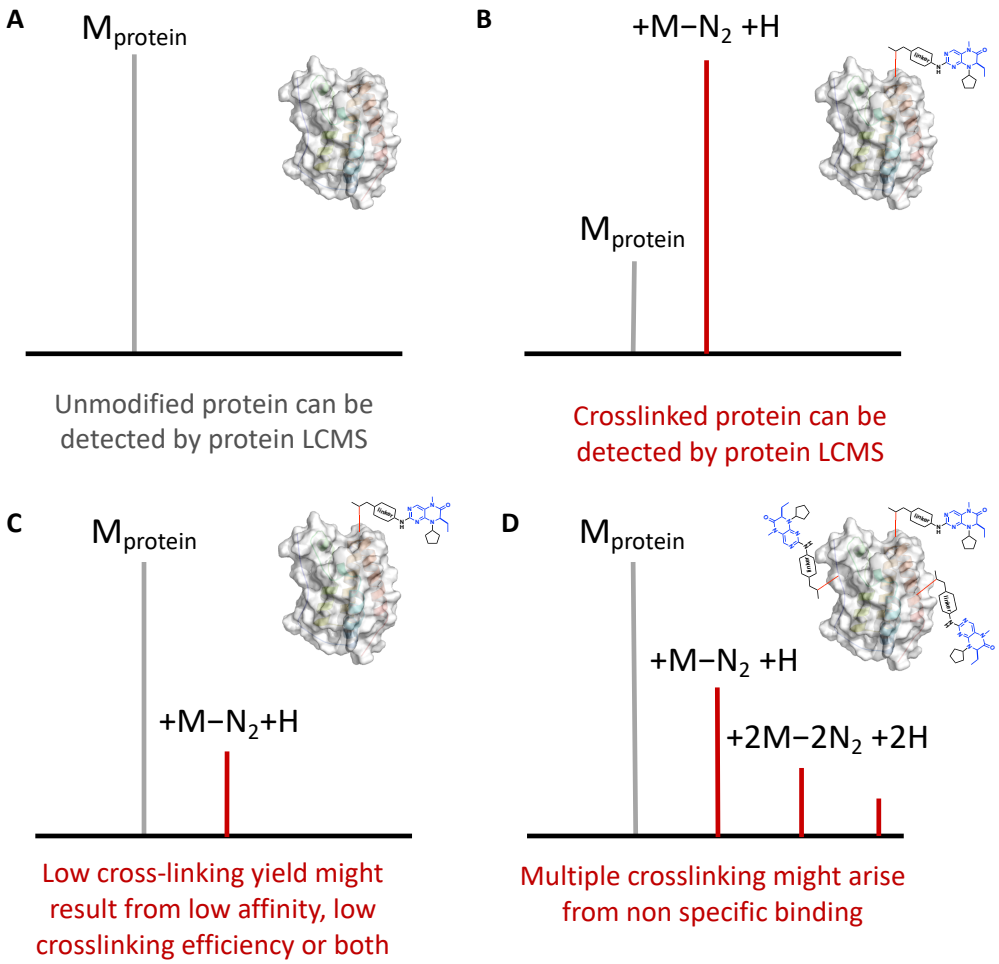


Figure 5.7: Schematic representation of the expected output of the photo-crosslinking of diazirine tool compounds to *Tc*BDF3. (A) The ligand does not bind, no cross linking is detected. (B) The ligand binds and cross-links to the protein efficiently. (C) The ligand binds and cross-links to the protein, the observed cross-linking yield arises from the binding affinity and the cross-linking efficiency, therefore, lower cross-linking yield could arise from lower affinity, lower cross-linking efficiency or both (D) Multiple cross-linking events are observed, the ligand is identified as a non-specific binder.

5.3.1. Diazirine tool compounds

The diazirine tool compounds **(*R*)-132** and **133** (Figure 5.8) were evaluated using mass spectrometry (MS) along a diazirine analogue of OXFBD02 (**7**), **161**. The DMI diazirine probe **161** was synthesised by Dr Giulia Stazi.

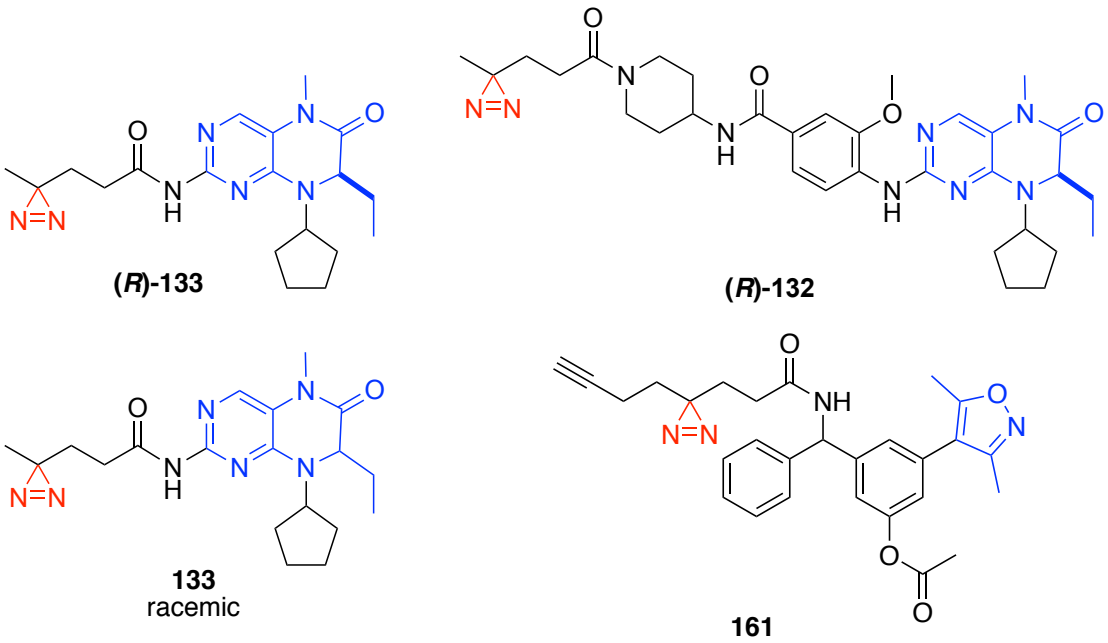


Figure 5.8: The structures of the diazirine tool compounds **133**, (*R*)-**133**, (*R*)-**132**, and **161**. The KAc-mimicking core is highlighted in blue and the diazirine group in red. **161** was synthesised by Dr Giulia Stazi.

A diazirine probe must bind the target protein and the formed carbene must be in proximity of the protein surface to efficiently cross-link to the protein. The crosslinking yield is therefore a result of the binding affinity and the cross-linking efficiency, and is not a direct measure of either events. When the protein target is saturated with diazirine probes then the crosslinking yield should be at its maximum, provided that there is no unspecific binding and that the probe concentration is the only variable. To rank diazirine probes we propose to characterise them according to the concentration at which they reach 50% of their maximum cross-linking effect: XL_{50}. This calculation requires the probe to reach a cross-linking yield plateau from a single cross-linking event.

5.3.1.1. Cross-linking to *Tc*BDF3

The XL_{50} values of the diazirine tool compounds (*R*)-**132**, **133**, and **161** were determined by titrating the compounds into a fixed amount of *Tc*BDF3, irradiating all samples for exactly 10 mins. For the three diazirine probes the cross-linking yield increased with increasing amount of the probes and then reached a plateau. Using the rates at which the diazirine probes reached their respective maximum cross-linking yields, the XL_{50}

values were extrapolated (Figure 5.10). **(*R*)-132**, **133**, and **161** were shown to have XL_{50} values between 8 and 16 µM (Table 5.2).

At the highest tested concentration, **133** and **161** gave maximum cross-linking yield of 17%, whereas **(*R*)-132** yielded 8% of cross-linked protein. The differences in the maximum cross-linking yield values is partially cause by the differences in cross-linking abilities of the probes. This confirms that the extended **(*R*)-132** diazirine probe cross-links less efficiently to the protein than **133**. This is likely to arise from the position of the diazirine, which is thought to be solvent exposed in the case of **(*R*)-132**.

Table 5.2: The diaizirine tool compound XL_{50} values for their binding and cross-linking to *Tc*BDF3. This calculation assumes that the cross-linking ability of the diazirine probe is constant regardless at which concentration they are used. The XL_{50} values are given as the median and standard deviation for each triplicate experiments.

Diazirine probe	XL_{50} from cross-linking titration	Maximum cross-linking yield
(*R*)-132	8.2±1.9 µM	8%
133	14.5±3.9 µM	17%
161	15.2±2.2 µM	17%

The three diazirine tool compounds were shown to have comparable XL_{50} values. However, the lower maximum cross-linking yield observed for **(*R*)-132** could lead to higher errors when measuring the prevention of its cross-linking to *Tc*BDF3. Therefore, **133** and **161** are more suited to be used as tool compounds for *Tc*BDF3.

Using the ^{1}H NMR ligand-observed protein titration (Chapter 1B), the binding of **133** diazirine tool compound to *Tc*BDF3 was measured. Using NMR, the **133** diazirine tool compound was shown to bind to *Tc*BDF3 with a K_D range of 164–198 µM. This suggests the binding affinity to be 10-fold higher than the XL_{50} value extrapolated through the cross-linking data. Previous work on BI-2536 (**8**) has shown that the K_D value estimated by NMR titration correlated with the value measured by ITC (Chapter 3). This confirms that the XL_{50} value does not correlate with the binding affinity as the cross-linking yields are a function of both the binding affinity and the cross-linking efficiency of the diazirine tool compound. Therefore, the XL_{50} values can be used in addition to the maximum cross-

linking yield value to select diazirine for further characterisation, however, should not be used as a unique criteria.

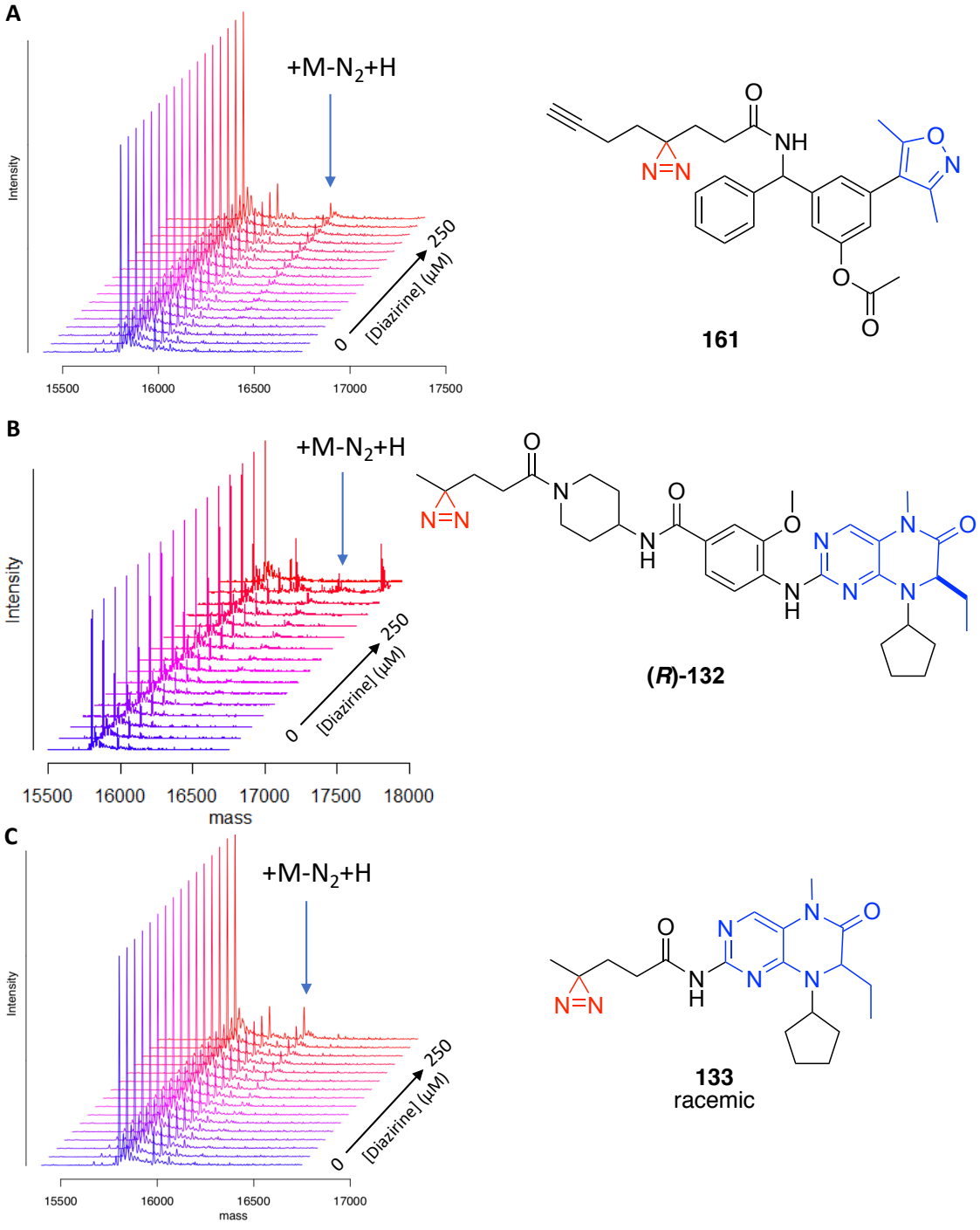


Figure 5.9: The mass-spectrometry spectra stacks for the cross-linking titration of the diazirine probes to *Tc*BDF3. The MS spectra corresponding to 0 to 250 µM of the probe are stacked from low to high concentration into the page, as indicated by the grey arrow. The cross-linking peak, corresponding to M(protein)+M(probe)–M(nitrogen), is highlighted by a blue arrow. (A) The DMI diazirine probe **161** shows the cross-linking yield to *Tc*BFF3 (B) The **(*R*)-132** probe low levels of cross-linking to *Tc*BDF3 (one of the MS runs failed) (C) **133** shows cross-linking to *Tc*BDF3.

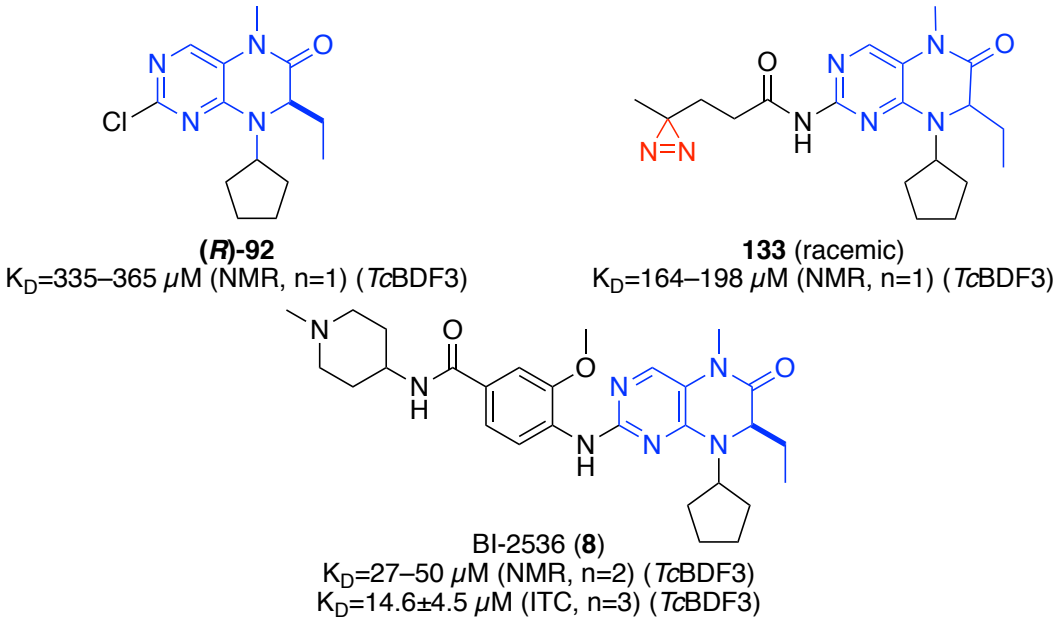


Figure 5.10: The comparison of the binding affinities of **(*R*)-92**, **133** and BI-2536 (**8**) to *Tc*BDF3.

Overall, **133** and **161** were shown to cross-link to *Tc*BDF3, and could be used to study the prevention of their cross-linking by competitor molecules. For example, *Tc*BDF3 binders identified through other techniques could be incubated with the protein and the diazine probes prior to light irradiation. The decrease in cross-linking of the probe would suggest at least partial co-localisation of the binding of the probe and the competitor molecule. **133** and **161** are the first photo-activatable tool compounds reported for *Tc*BDF3. **(*R*)-133** is predicted to have a higher maximum cross-linking yield than its racemate, however, due to time constraints has not yet been tested. This future work, planned to take place in GSK Stevenage, will determine if the single enantiomer **(*R*)-133** offers an advantage over its racemate for the use in cross-linking experiments.

5.3.1.2. Cross-linking to *Tc*BDF5-1

Similar to *Tc*BDF3, the XL_{50} values of the diazirine tool compounds were determined by titrating the tool compounds, **133**, **(*R*)-132**, and **161**, into a fixed amount of *Tc*BDF5-1. Here there is a 20-fold difference in XL_{50} between **161** and **(*R*)-132** and the shorter diazirine **133** (Table 5.3). The DMI diazirine probe **161** has the lowest XL_{50} value and highest cross-linking yield, confirming that the DMI moiety is a more efficient KAc mimic for binding *Tc*BDF5-1 (Figure 5.11).

Table 5.3: The diazirine tool compound pseudo-K_D values for their binding and cross-linking to *Tc*BDF5-1. This calculation assumes that the cross-linking ability of the diazirine probe is constant regardless at which concentration they are used.

Diazirine probe	XL_{50} values from cross-linking titration	Maximum cross-linking yield
161	9.7±1.2 µM	24%
133	202.3±68 µM	7%
(*R*)-132	12.5±3.4 µM	11%

Overall, the DMI diazirine probe **161** cross-linked to *Tc*BDF5-1 in a dose-dependent manner, with a suitable maximum cross-linking yield. This probe, originally designed for the activity-based protein profiling for the study of BRD4(1), was shown to also be an effective photo-activatable tool compound for *Tc*BDF5-1.

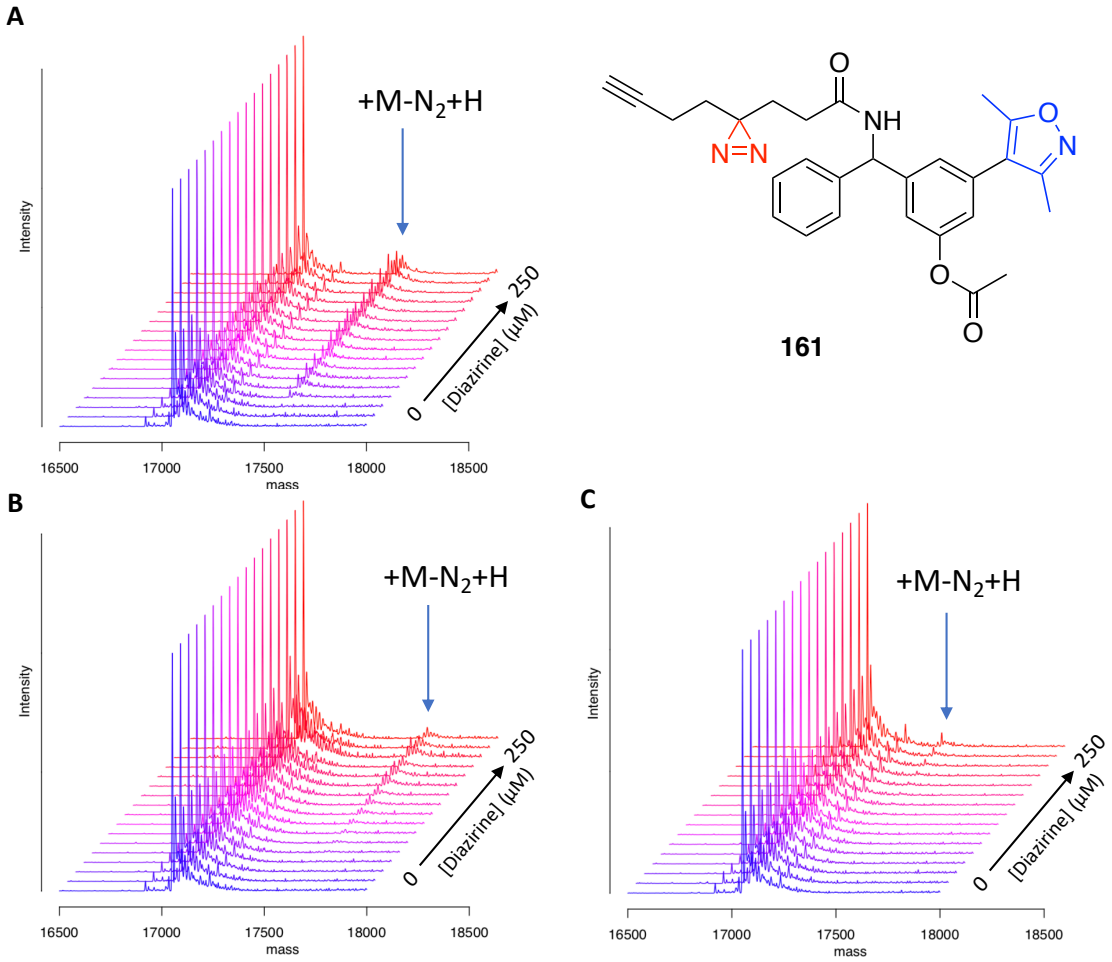


Figure 5.11: The mass-spectrometry spectra stacks for the cross-linking titration of the diazirine probes to *Tc*BDF5-1. The MS spectra corresponding to 0 to 250 µM of the probe are stacked from low to high concentration into the page, as indicated by the grey arrow. The cross-linking peak corresponding to M(protein)+M(probe)–M(nitrogen), is highlighted by a blue arrow. (A) The DMI diazirine probe **161** shows the highest cross-linking yield and lowest apparent K_D value for *Tc*BFF5-1 (B) The **(*R*)-132** probe shows gradual cross-linking to *Tc*BDF5-1 (C) **133** only shows cross-linking to *Tc*BDF5-1 at concentrations above 60 µM.

The cross-linking of diazirines to proteins is an appealing method, as protein LCMS runs and their analysis are well established, have the potential to be high throughput, and allow for detection of a wide range of binding affinities. Since the binding and cross-linking efficiencies are distinct, the cross-linking yield, which is the assay's output, is a function of both events. A tight binder, with a solvent exposed diazirine, will most likely be quenched by solvent rather than cross-link to the protein, despite binding efficiently to the protein. On the other hand, a weak ligand with a well-placed diazirine is expected to produce a carbene in the proximity of the protein surface and lead to an apparent high cross-linking efficiency. This implies that the cross-linking yield itself cannot be used to distinguish between different binding affinities. Although the two different factors can mask each other's effect, the diazirine probes are expected to reach a maximum cross-linking yield when the protein is saturated by the ligand. Therefore, the diazirine probes can be ranked by their XL_{50} values, extrapolated through the titration of the probe into protein. The diazirine tool compounds that lead to their maximum cross-linking yield when present in low amounts have a lower XL_{50} values than diazirines that required to be present in higher quantities to obtain their cross-linking maximum yield.

5.3.2. Screening the GSK PhaBits library: a diazirine fragment library to identify new *T. cruzi* bromodomain binders

In collaboration with Emma Grant (GSK Stevenage, University of Strathclyde) and Dr Jacob Bush (GSK Stevenage), the GSK PhaBits library was screened against *Tc*BDF3 and *Tc*BDF5-1. The PhaBits library is a collection of diazirine containing fragments, which was designed to identify new binding fragments to be used for challenging protein targets. As part of her PhD work, Emma Grant created the library by selecting over 550 compounds and modifying them with a diazirine linker to provide a varied library of photo-activatable cross-linkers (GSK Stevenage and the University of Strathclyde, unpublished results). This library was validated against BRD4(1) prior to this collaborative work.

The screening strategy was as follow: (1) screen the 528 PhaBits compounds and identify cross-linking fragments (2) validate the fragments by preventing the cross-linking events with a known binder (3) titrate the validated cross-linker, identify their maximum cross-linking yields and their XL_{50} values (Figure 5.12).

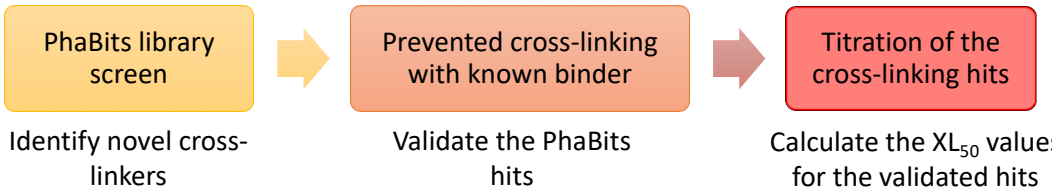


Figure 5.12: Summary of the screening strategy employed for the PhaBits library against *Tc*BDF3 and *Tc*BDF5-1 to identify new *T cruzi* BD binders.

The criteria for identifying hits from the PhaBits screen was the detection of a single cross-linking event, with a cross-linking yield over 1.5%. The following prevented cross-linking validation is a key step. The pre-incubation of the protein with a known binder prior to irradiation with the cross-linking hits, is predicted to reduce the cross-linking yield if the PhaBits hit. The drop in cross-linking yield is characteristic of specific binders that co-localise at least partially when binding the protein.

5.3.2.1. PhaBits screen on *Tc*BDF3

In collaboration with Emma Grant, the 556 PhaBits compounds were screened by protein LCMS. The compounds were incubated with the protein on ice, then irradiated at 302 nm for 10 mins prior to the LCMS runs. The library was screened in 384-well plates. The data were deconvoluted, then analysed manually.

Overall, 220 fragments were found to modify *Tc*BDF3, of which 99 met the minimum threshold of cross-linking (Table 5.4). Over a hundred compounds were shown to oxidise the protein, which indicates that a radical chain reaction occurred. Representative spectra illustrating the cross-linking results obtained when screening the PhaBits library against *Tc*BDF3 can be found in Appendix B7.

Table 5.4: Summary of the outcome of the PhaBits library screen against *Tc*BDF3. The cross-linking yield, calculated by the peak intensity, was used to characterise the diazirine compounds. Diazirines are considered "hits" when cross-linking yields above 1.5% are detected. Diazirines leading to radical chain reactions or multiple cross-linking events result from non-specific binding. Ambiguous results and failed MS runs are also reported.

Cross-linking Results	Number of fragment hits
5% or more cross-linking	58
1.5 to 5% cross-linking	41
protein +16 (indicating a radical chain reaction resulting in oxidation of the protein)	112
multiple cross-linking events	9
MS run failed	10
ambigious	6

68 of the PhaBits hits were further validated by co-incubation with BI-2536 (**8**): 14 hits showed full decrease of cross-linking to *Tc*BDF3, 8 partially prevented cross-linking, and 21 had no effect. The hits, which cross-linking was prevented by BI-2536 (**8**), were therefore identified as binding at least partially in the binding pocket, and were further characterised. Finally, 9 of the validated PhaBits hits (Figure 5.13) were ranked by their cross-linking yields and pseudo-K_D values, extracted from the titration of increasing amounts of the diazirine into fixed amount of *Tc*BDF3 (Table 5.5). Five of the tested diazirines, **163**, **165**, **166**, **167**, and **169**, did not show a plateauing behaviour, and so their pseudo-K_D values could not reliably be determined. The four other hits **162**, **164**, **168**, and **170**, show pseudo-K_D values from 157 to 1094 µM. This validation work was performed by Emma Grant at GSK Stevenage.

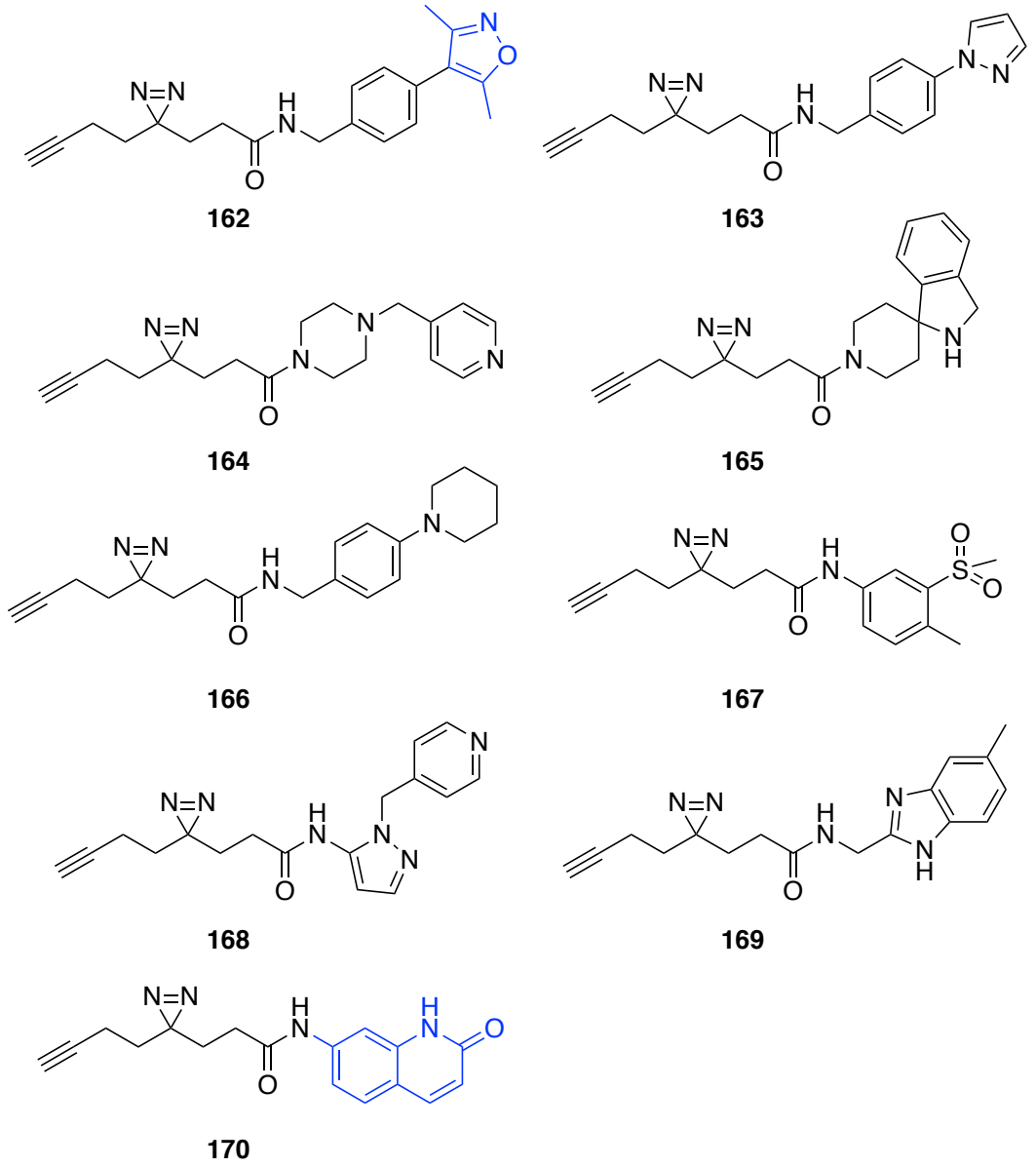


Figure 5.13: The structures of the nine novel hits for *Tc*BDF3 from screening the PhaBits library. (Validation through cross-linking prevention by pre-incubating protein with known binder was performed by Emma Grant, GSK Stevenage). The known KAc-mimicking moieties are highlighted in blue.

Table 5.5: Summary of the calculated maximum cross-linking yields and pseudo-K_D values obtained for the nine PhaBits hits for *Tc*BDF3. This XL_{50} calculation assumes that the cross-linking ability of the diazirine probe is constant regardless at which concentration they are used. N/D denotes that the data or its fitting was ambiguous and the values could not be determined.

	162	163	164	165	166	167	168	169	170
Maximum cross-linking yield	24%	5%	12%	1%	2%	3%	9%	14%	22%
XL_{50} (μM)	1094	N/D	292.3	N/D	N/D	N/D	157.1	N/D	391.7
R^2	0.964	N/D	0.979	N/D	N/D	N/D	0.976	N/D	0.972

The step-wise validation of the initial PhaBits hit has identified four new binders for *Tc*BDF3: **162**, **164**, **169** and **170** (Figure 5.13). **162**, contain a dimethyl isoxazole moiety, which have been previously shown to bind *Tc*BDF3 by waterLOGSY. The other hits, **164** and **168**, are novel structures, both containing a 4-pyridinyl group. The last hit, **170**, contains a quinolinone fragment seen in other BD ligands. Following the identification of these novel *Tc*BDF3 binders using the PhaBits library, it was important to validate further the binding using an orthogonal assay. To do so, **162**, **164** and **168** were tested by waterLOGSY. The fourth hit, **170**, was not available in enough quantity to be tested further. The compounds were not soluble enough at 125 µM, and so had to be tested at 62.5 or 31.25 µM. **162** showed inversion of the two methyl peaks indicating binding, however, **164** and **168** did not show binding.

WaterLOGSY can detect a large range of binding affinity, with affinities from low milli- to low micromolar typically detected. This technique relies on fast exchange, and therefore cannot detect the binding of higher affinity ligands. The tested compound following the PhaBits screen are fragments, therefore are unlikely to be high affinity ligand appearing as false negatives. To confirm that **164** and **168** are indeed not tight binders undetectable by waterLOGSY, ITC titration was run on the three compounds. None of the tested compounds showed binding, confirming that **162** is a weak ligand and that **164** and **168** are not binding with detectable affinities (K_D > mM).

Overall, 556 diazirine fragments were tested against *Tc*BDF3 using the PhaBits diazirine library. Four hits showed a dose-dependent cross-linking, and a decreased cross-linking yield when pre-incubating the protein with a known binder. Of these, three validated hits, **162**, **164** and **168**, were tested as non-diazirine fragment using waterLOGSY and ITC. This approach identified **162** as a hit for *Tc*BDF3, and showed that the binding of the other two Phabits hits, **164** and **168**, was not detectable. The detection of the previously-identified DMI KAc mimic confirms the initial fragment work completed as part of this work (Chapter 1), and illustrates well the challenge of identifying high-affinity ligands for *Tc*BDF3.

5.3.2.2. PhaBits screen on *Tc*BDF5-1

The 556-compound Phabits library was screened by LCMS following irradiation with *Tc*BDF5-1 to identify novel cross-linkers. 235 fragments modified the protein, of which 170 met the minimum threshold of cross-linking (Table 5.6). 17 compounds showed multiple cross-linking events, some of which also showed this behaviour when screened with *Tc*BDF3 (structures not released), highlighting the non-specificity of those compounds.

Table 5.6: Summary of the outcome of the PhaBits library screen against *Tc*BDF5-1. The cross-linking yield, calculated by the peak intensity, was used to characterise the diazirines compounds. Diazirines are considered "hits" when cross-linking yields above 1.5% are detected. Diazirines leading to radical chain reactions or multiple cross-linking events result from non-specific binding. Ambiguous results and failed MS runs are also reported.

Cross-linking results	Number of fragment hits
5% or more cross-linking	89
1.5 to 5% cross-linking	81
protein +16 (indicating a radical chain reaction resulting in oxidation of the protein	48
multiple cross-linking events	17
MS run failed	2
ambigious	15

The hits were further validated by screening them against *Tc*BDF5-1 in the presence of OXFBD02 (**7**), a *Tc*BDF5-1 ligand with a K_D range of 150–350 µM (Appendix B3). The cross-linking data gave the same result for every well: an unusual multiple cross-linking pattern. After further investigation, it was shown that OXFBD02 (**7**) crosslinks directly to the protein. This prevents its use as a control compound. Consequently, the validation experiment was repeated with BI-2536 (**8**), a *Tc*BDF3 binder. Due to the expected low binding affinity (> mM) of this ligand to *Tc*BDF5-1, fragments binding tighter than BI-2536 (**8**) were not expected to be displaced during the incubation, and so not prevent cross-linking. The absence of a suitable known binder, compatible with light irradiation, meant that binders might not appear as false negatives. As DMI have been shown to bind *Tc*BDF5-1 using waterLOGSY, the DMI-containing PhaBits hits were included in the next validation step despite not showing prevented cross-linking using the BI-2536 (**8**).

Overall, 28 hits showed decreased cross-linking in the presence of BI-2536 (**8**). These compounds, and 2 additional DMI compounds hits (**172** and **173**), were further characterised by titrating increasing amounts of the PhaBits hits into a fixed amount of *Tc*BDF5-1 to determine their maximum cross-linking yield. 13 hits showed a maximum cross-linking yield above 5%, with 4 hits with yields above 20% and 3 above 35%. Overall, six novel hits (**171**–**176**) were identified for *Tc*BDF5-1, of which, two contained a DMI moiety (Figure 5.14).

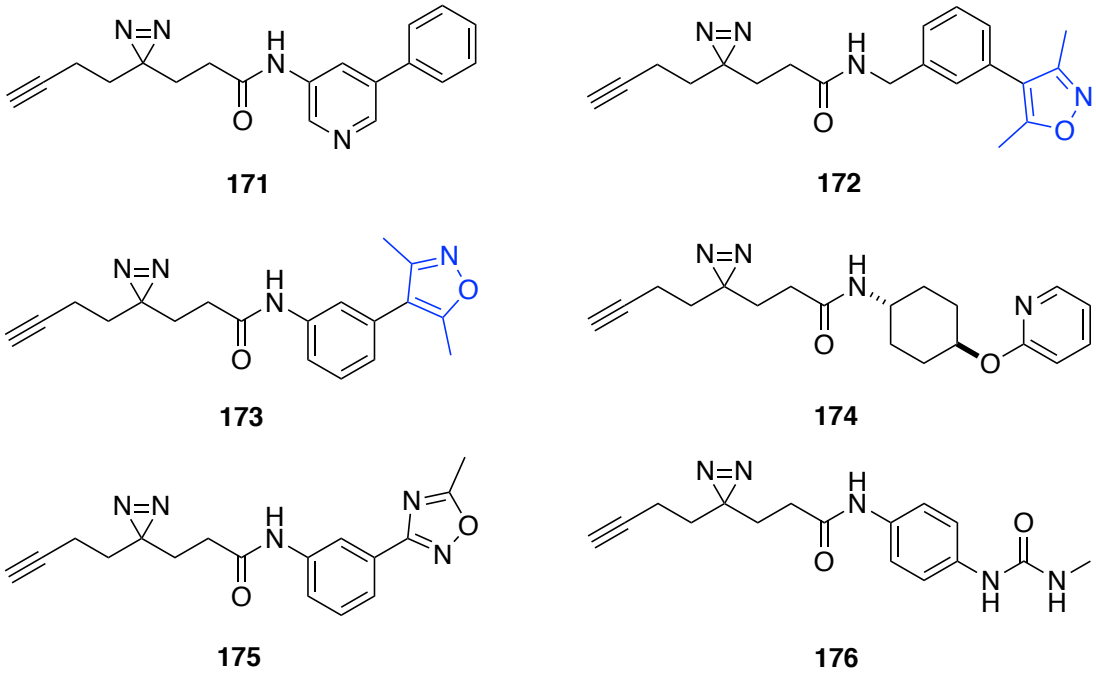


Figure 5.14: The structures of the PhaBits hits identified for *Tc*BDF5-1. The known KAc-mimicking moieties are highlighted in blue.

Table 5.7: Summary of the calculated maximum cross-linking yields and XL_{50} values obtained for the six PhaBits hits for *Tc*BDF5-1. This XL_{50} calculation assumes that the cross-linking ability of the diazirine probe is constant regardless at which concentration they are used. N/D denotes that the data or fitting was ambiguous and the values could not be determined.

Compound	**171**	**172**	**173**	**174**	**175**	**176**
Maximum cross-linking yield	41%	41%	35%	26%	13%	22%
XL_{50} (μM)	692.8 μM	134.4 μM	25.92 μM	82.07 μM	N/D	206.5 μM
R^2	0.9104	0.919	0.9507	0.9769	N/D	0.9771

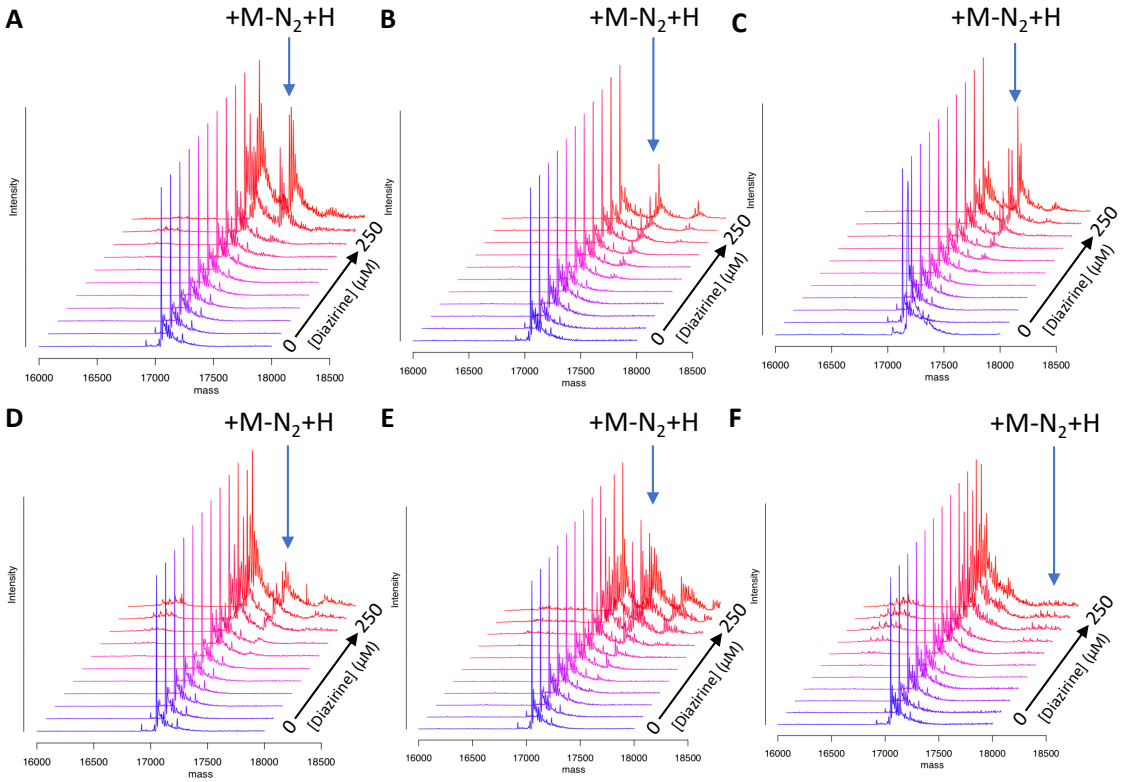


Figure 5.15: Overlay of the MS spectra for the titration experiments of the PhaBits hits. The ratio of the peaks integrals was used to extrapolated pseudo-K_D values. (A) **171** (B) **172** (C) **173** (D) **174** (E) **175** (F) **176**. The MS spectra corresponding to 0 to 250 µM of the probe are stacked from low to high concentration into the page.

Overall, the use of diazirines has allowed development of tool compounds for *Tc*BDF3, and enabled identification of new fragments that bind *Tc*BDF3 and *Tc*BDF5-1 using the PhaBits library. The diazirine tool compounds **(*R*)-132** and **133** cross-link to *Tc*BDF3. **133** binds to *Tc*BDF3 with a K_D range of 164–198 µM (n=1, NMR) and can reached up to 17% of cross-linking yield. The OXFBD02 (**7**)-based diazirine **161** is a suitable photo-activatable probe for both *Tc*BDF3 and *Tc*BDF5-1. The PhaBits fragment diazirine library, designed to identify fragments for challenging targets, was screened against both *Tc*BDF3 and *Tc*BDF5-1. A step-wise validation of the initial PhaBits hit has identified three new binders for *Tc*BDF3 and six fragments for *Tc*BDF5-1.

5.4. Fluorescence polarisation assay

The fluorescent probes designed for targeting *Tc*BDF3 consist of a fluorophore, a linker and the scaffold of the BI-2536 (**8**) ligand and could be used in a fluorescence polarisation (FP) assay. The binding of the fluorescent probe to the protein is accompanied by an increase in the emitted polarised light (Figure 5.16A). The change in FP can be conveniently measured using a plate reader fitted with the appropriate wavelength filters. The fluorescence polarisation reaches a plateau when all the FP probe is bound. Consequently, the titration of the protein into a fixed amount of the probe can be used to determine the affinity of the probe to the protein (Figure 5.16B).

Choosing a concentration of protein for which a FP probe is saturated (plateaued), gives a protein:FP probe ratio that can be used for displacement experiments. Increasing amounts of a competitor molecules will gradually displace the bound FP probe, which will be monitored by a decrease of polarisation of the emitted light. The transition between the fully bound (high polarisation) and the fully free (low polarisation) states can be used to determine IC_{50} values for the competitor molecules (Figure 5.16C). The advantages of this assay are the increased throughput and the low protein requirements.

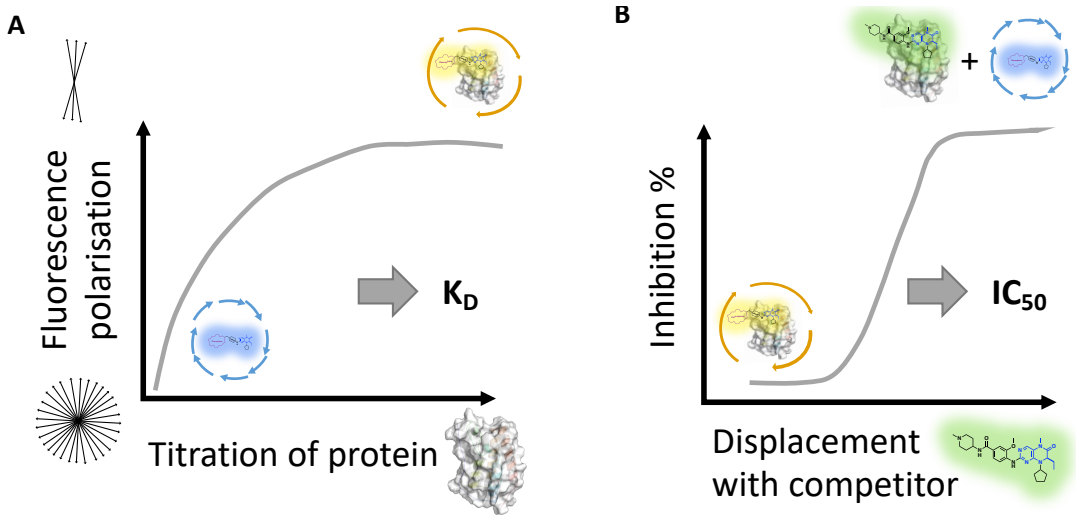


Figure 5.16: Schematic representation of the use of change of fluorescence polarisation (FP) to detect binding to a protein. (A) the titration curve obtained for the binding of the fluorescent probe to the protein: an increase in the polarised emitted light reaches a plateau when all the FP probe is bound, allowing for the determination of a K_D value. (B) The displacement of the bound fluorescent probe can be monitored by FP and used to determine IC_{50} values for the competitor molecules.

5.4.1. Characterisation of the FP probes fluorescence intensity and fluorescence polarisation

The fluorescent probes **(*R*)-134**, **(*R*)-135**, **(*R*)-136** and **(*R*)-137** (Figure 5.17) were first characterised by their fluorescence intensity (FI) and polarisation (FP), to determine the appropriate concentration of FP probe needed to get a stable polarisation signal and a suitable fluorescent intensity. The FI needs to be high enough to avoid large measurement errors. The concentration at which the probe emits a stable FP signal, can be used for the next experiments, for which this value will corresponds to the signal of the fully unbound probe.

The characterisation of the FP probes was achieved by serial dilutions of each probe in buffer into a 384-well plate, and measurement of the FI and FP using a plate reader. The FI and FP determined for fluorescent probes **(*R*)-134**, **(*R*)-135**, **(*R*)-136** and **(*R*)-137** show a linear FI *vs* probe concentration, and all displays high fluorescence intensities at concentrations above 10 nM (Appendix B6). Additionally, **(*R*)-134**, **(*R*)-135**, **(*R*)-136** and **(*R*)-137** have a stable FP from a concentration of 20 nM upwards. Together these findings show that the suitable probe concentrations are 20–40 nM.

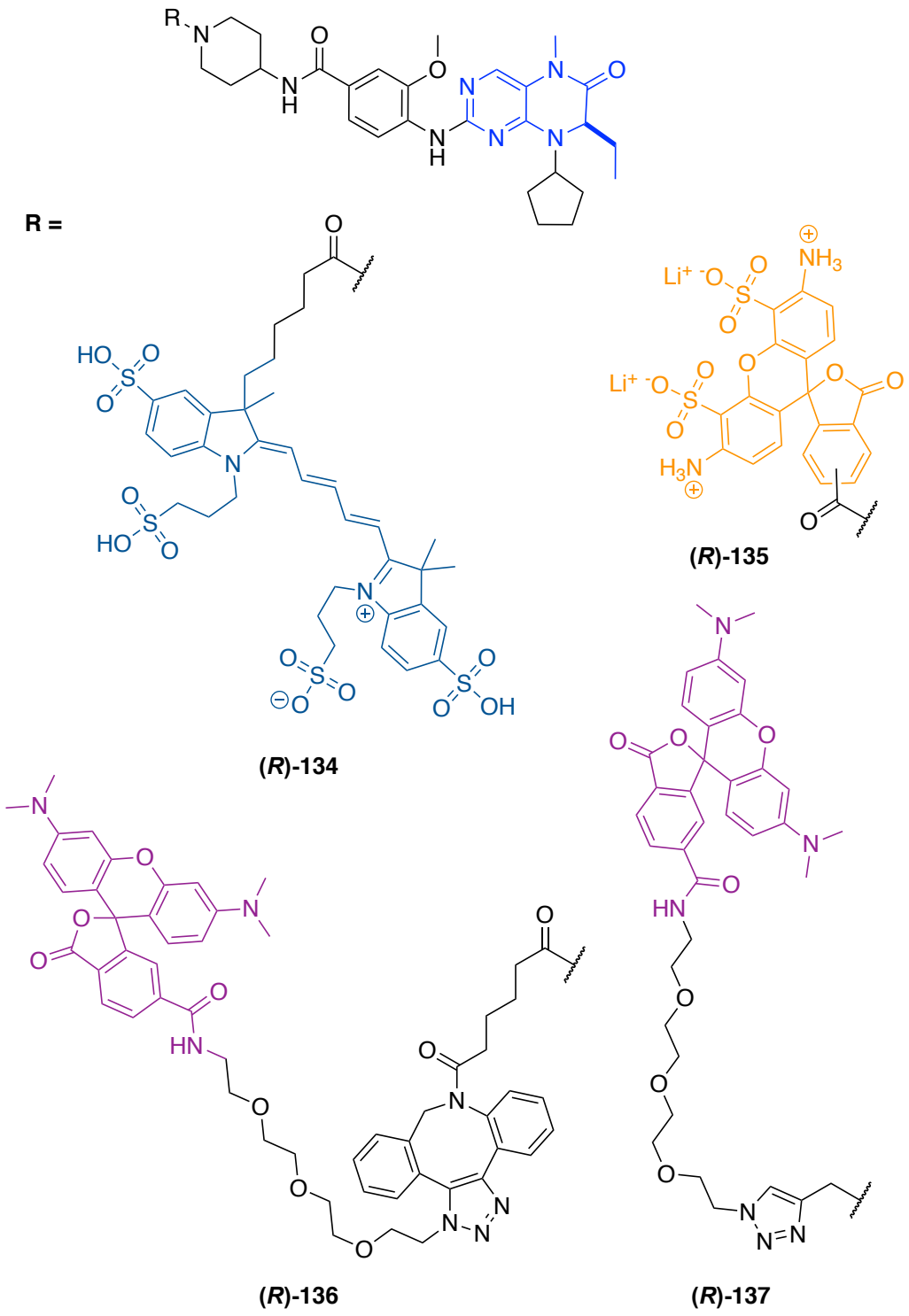


Figure 5.17: The structures of the four fluorescent tool compounds, **(*R*)-134**, **(*R*)-135**, **(*R*)-136** and **(*R*)-137**, designed to bind *Tc*BDF3.

5.4.2. Binding affinities of the FP probes to *Tc*BDF3 and other BD proteins

Binding of the fluorescent probe to a protein is accompanied by an increase in the emitted polarised light, as it tumbles more slowly when bound to the protein than when in solution. Therefore, on the condition that the chosen linker and fluorophore do not interfere with the binding, an increase in fluorescence polarisation of the FP probes is expected upon titration with the protein. A high value of polarisation for the bound probe is desirable, since the difference in values between the bound and free FP probe defines the assay window. In addition, the lowest IC_{50} that can be determined is half of the protein concentration. Therefore, the next priority was to determine the binding affinities of the FP probes, **(*R*)-134**, **(*R*)-135**, **(*R*)-136** and **(*R*)-137**, against *Tc*BDF3 and other BD proteins to identify suitable FP probes for displacement experiments.

5.4.2.1. The binding of FP probes to *Tc*BDF3

First, the binding of **(*R*)-134**, **(*R*)-135**, **(*R*)-136** and **(*R*)-137** to *Tc*BDF3 and BRD4(1) was assessed by fluorescence polarisation, by titrating the protein into 20–40 nM of the FP probes.

(*R*)-134, **(*R*)-136**, and **(*R*)-137** did not show binding to *Tc*BDF3 (Figure 5.18, A, C and D). The lack of binding could arise from the detrimental placement of the linker or the fluorophore, or from the "propeller effect". In the case of a flexible linker with many free rotating bonds, the fluorophore might still tumble fast even if the ligand portion of the probe is bound to the protein. This is known as the "propeller effect",[115-116] which might be avoided by using rigid linkers. Pleasingly, **(*R*)-135** showed binding to *Tc*BDF3 with a K_D value of 3.9±1.2 μM (n=2) (Figure 5.18, B). The **(*R*)-135** tool compound is the first *Tc*BDF3 binder with a K_D less than 10 μM. Additionally, its polarisation ranges between *ca.* 20 and 80 mP, giving a suitable assay window. Overall, the **(*R*)-135** tool compound is the first reported fluorescent probe for the *T. cruzi* bromodomains proteins and is the highest affinity *Tc*BDF3 ligand identified to date.

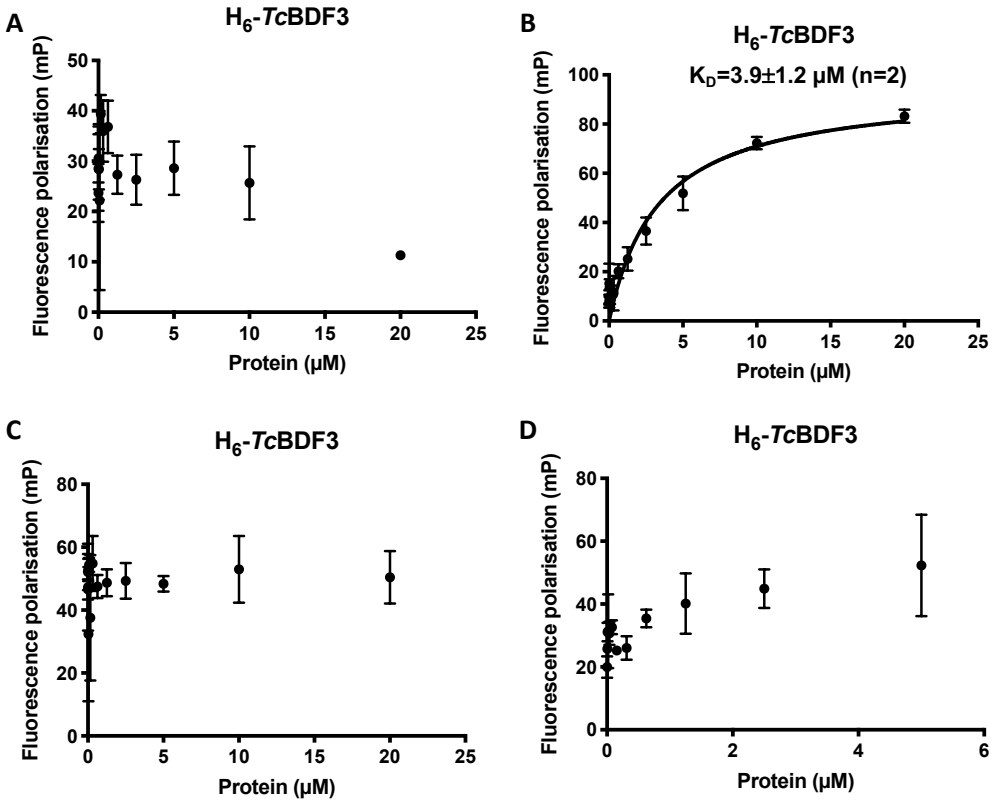


Figure 5.18: The fluorescence polarisation observed for the (A) (*R*)-134, (B) (*R*)-135, (C) (*R*)-136 and (D) (*R*)-137 fluorescent tool compounds observed when titrating *Tc*BDF3. (*R*)-134, (*R*)-136, and (*R*)-137 do not show binding to *Tc*BDF3. (*R*)-135 shows binding to *Tc*BDF3 with K_D=3.9±1.2 μM (n=2).

5.4.2.2. The binding of the FP probes to BRD4(1) and BRD4(2)

The four fluorescent probes were expected to retain a strong binding to BRD4 and PLK1. Therefore, we determined whether if these compounds bind to BRD4(1) and (2). Using the identical method as for *Tc*BDF3 (Section 5.4.2.1), the human bromodomains were titrated into fixed amount of fluorescent probes and the binding was monitored by FP. (*R*)-134 and (*R*)-136 did not show binding to either of the BRD4 domains, however, both (*R*)-135 and (*R*)-137 showed high-affinity binding to the studied proteins (Table 5.8, Appendix B6).

(*R*)-135 binds to BRD4(1) and BRD4(2) with K_D values of 11.1±2.3 nM and 20.1±1.9 nM, respectively. (*R*)-137 binds to BRD4(1) with a K_D=9.6±1.8 nM and BRD4(2) with a K_D=28.3±8.3 nM..

Table 5.8: Summary of the binding of the tool compounds, **(*R*)-134**, **(*R*)-135**, **(*R*)-136**, and **(*R*)-137**, to *Tc*BDF3, BRD4(1), and BRD4(2).

FP tool compound	*Tc*BDF3 K_D value (FP)	BRD4(1) K_D value (FP)	BRD4(2) K_D value (FP)
(*R*)-134	*Not binding*	*Not binding*	*Not binding*
(*R*)-135	3.9±1.2 µM (n=2)	9.9±2.5 nM (n=3)	14.9±1.7 nM (n=2)
(*R*)-136	*Not binding*	*Not binding*	*Not binding*
(*R*)-137	*Not binding*	9.6±1.8 nM (n=1)	28.3±8.3 nM (n=1)

5.4.2.3. The binding of **(*R*)-135** to PLK1

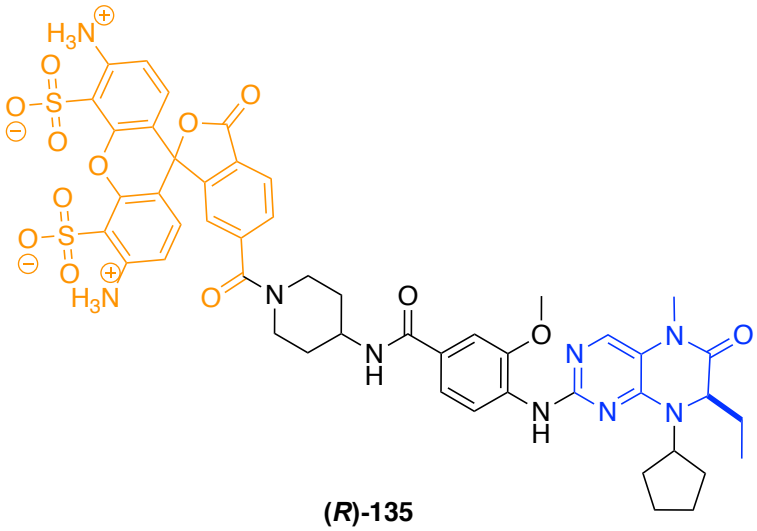
It was also of interest to assess the binding of the fluorescent tool compounds, **(*R*)-135**, to additional proteins. The tool compounds were designed using the structure of the BI-2536 (**8**) BRD4/PLK1 ligand so assessing the binding to PLK1 was also of interest. Therefore, the binding of the fluorescent tool compounds **(*R*)-135** to CREBBP and PLK1 was assessed by FP titration. **(*R*)-135** showed binding to CREBBP, with a K_D of 1.25±0.26 µM and also bound to PLK1, with a K_D of 5.5±1.5 nM (n=2) (Appendix B6).

5.4.2.4. **(*R*)-135** as a fluorescence probe for *Tc*BDF3

The fluorescent tool compound **(*R*)-135** is the first example of a FP probe for *Tc*BDF3 (Figure 5.19). The binding of **(*R*)-135** was shown to bind BRD4(1), BRD4(2), and PLK1, in addition to its binding to *Tc*BDF3 by FP titration (Table 5.9). The 2 to 3-fold increase in binding compared to the BI-2536 (**8**) ligand could be a consequence of the binding of the fluorophore moiety to hydrophobic patches near the binding site.

Table 5.9: Summary of the K_D values determined for **(*R*)-135** to the *Tc*BDF3, *Tc*BDF5 BD1, BRD4 BD1, BRD4 BD1 bromodomains and PLK1. The estimated protein concentration suitable for FP displacement assay is based on the lowest protein concentration for which the probe FP is close to its maximum value, ie. The plateau in the titration curve.

Protein	**(*R*)-135** K_D value	Estimated protein concentration suitable for FP displacement assay
*Tc*BDF3	3.9±1.2 µM (n=2)	5–10 µM
*Tc*BDF2	*Not binding*	–
*Tc*BDF5-1	*Not binding*	–
BRD4(1)	9.9±2.5 nM (n=3)	0.1–0.2 µM
BRD4(2)	14.9±1.7 nM (n=2)	0.1–0.2 µM
PLK1	5.5±1.5 nM (n=2)	0.05–0.1 µM



(*R*)-135
K_D=3.9±1.2 μM (FP, n=2) (*Tc*BDF3)
K_D=9.9±2.5 nM (FP, n=3) (BRD4-1)
K_D=14.9±1.7 nM (FP, n=2) (BRD4-2)
K_D=5.5±1.5 nM (FP, n=2) (PLK1)

Figure 5.19: The fluorescent tool compound **(*R*)-135**.

5.4.3. Displacement assay set up for *Tc*BDF3

5.4.3.1. Optimisation using the positive control BI-2536 (**8**)

(*R*)-135 displayed its maximum fluorescence polarisation at concentrations of *Tc*BDF3 at or above 10 µM. Therefore, the displacement of the **(*R*)-135** FP probe from *Tc*BDF3 was first attempted using 10 µM of protein in triplicate with 20 µL total volume per well: partial displacement was observed, however, the data generated using these conditions showed errors within the triplicates and lack of dose-response (Figure 5.20, A). To improve the data, the displacement conditions were optimised. First, the assay volume was changed from 20 to 40 µL and the FP probe concentration was increased from 40 to 100 nM. These changes reduced the observed standard error of the measurement of the triplicate well. Further changes included switching from 0.25 to 0.5% final DMSO content, allowing to screen BI-2536 (**8**) at higher concentrations. These conditions allowed to reach 35% displacement of the FP tool compounds by the ligand (Figure 5.20, B).

The 3-fold increase of binding of the FP probe **(*R*)-135** with respect to the unmodified BI-2536 (**8**) ligand predicts that the displacement of the fluorescent tool compound will require an excess of the BI-2536 (**8**) ligand. Moreover, MD simulations (Joseph Bluck) has

shown that **(*R*)-135** is likely to bind the protein through its fluorophore portion in addition to its binding to the KAc binding pocket. Together these findings rationalise why only partial displacement is observed. In particular, the displaced FP probe might still bind the protein through its fluorophore portion, and is likely to displace the competitor molecule.

The final optimisation of the displacement was the switch from a "classical displacement" to a "prevented binding" set up. The first case assumes the pre-incubation of the FP probe and the protein prior to the addition of the competitor molecule, here BI-2536 (**8**). The prevented binding set up describes the pre-incubation of the competitor compound with the protein of interest prior to the addition of the fluorescent reporter molecule. This second setup was hypothesised to be beneficial for the study of **(*R*)-135** and BI-2536 (**8**), as it might minimise the effect of the non-specific binding of the hydrophobic moiety, which renders the FP probe **(*R*)-135** less displaceable. Indeed, using a prevented binding approach the **(*R*)-135** was be displaced up to 50% (Figure 5.20, C).

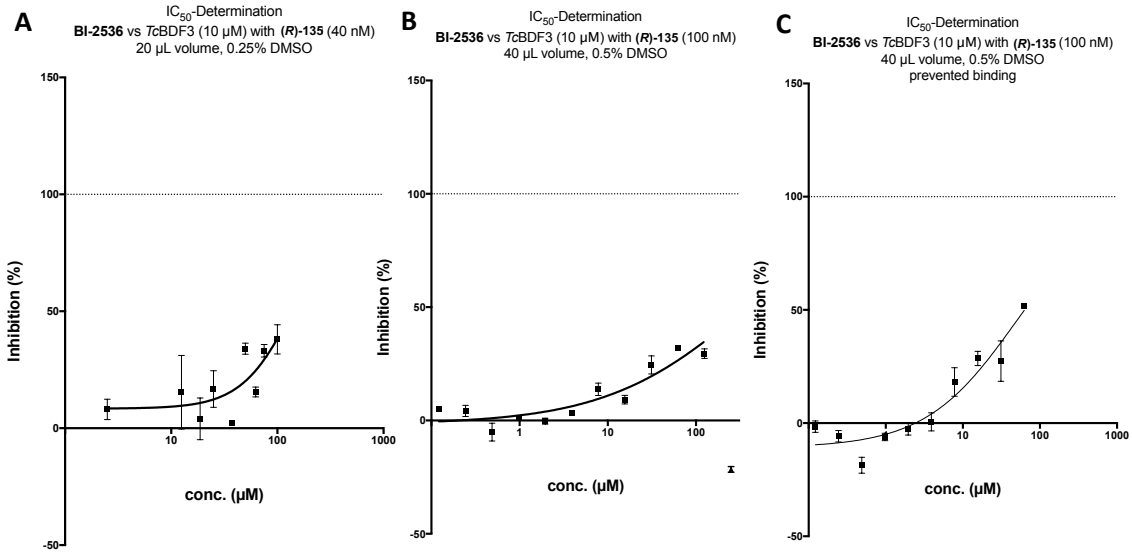


Figure 5.20: The optimisation of the inhibition curve of **(*R*)-135** and *Tc*BDF3 in the presence of the BI-2536 (**8**) ligand. (A) The initial displacement conditions detected potential inhibition but large errors were observed under the employed conditions (B) The optimised FP probe concentration and assay volume gave higher quality data. The displacement setup, with pre-incubation of the FP probe and the protein prior to addition of the competitor, gives a maximum of ~35% inhibition. (C) The prevented binding, with pre-incubation of the competitor compounds and the protein prior to addition of the FP probe, gives an improved 50% inhibition.

Despite using the optimised screening conditions, a full inhibition curve, and therefore an IC_{50} value, could not be obtained due to assay interference when using the BI-2536 (**8**) ligand at concentrations above 100 µM. Therefore, the displacement of the **(*R*)-135** tool compound by BI-2536 (**8**) did not afford quantitative data; however, the partial displacement shows that FP inhibition can be detected, and is predicted to yield full inhibition curves for higher affinity ligands. Overall, the FP tool compound **(*R*)-135** has been used to set up a novel displacement-based assay for *Tc*BDF3 and will detect binding of ligand that bind *Tc*BDF3 with equal or higher affinities to the current positive control, BI-2536 (**8**), or alternatively, equipotent ligands which do not interfere with the assay at concentration at or above 100 µM.

5.4.3.2. Investigation of the displacement of **(*R*)-135** by other ligands for *Tc*BDF3

5.4.3.2.a. Using fluorescence polarisation to detect the binding of BI-6727 (**9**)

BI-6727 (**9**) binds to *Tc*BDF3 with a K_D=57.1±70.5 µM (ITC, n=3) (Chapter 3). BI-6727 (**9**) partially displaced **(*R*)-135** using the conditions optimised using BI-2536 (**8**), and reach a maximum of 35–40% inhibition in the FP assay (Figure 5.21). These finding show that this novel FP assay can detect *Tc*BDF3 binding, up to 50 µM affinities. This is additional evidence that the tool compound **(*R*)-135** can be used to detect binding of ligands that bind *Tc*BDF3 with equal or higher affinities to BI-2536 (**8**).

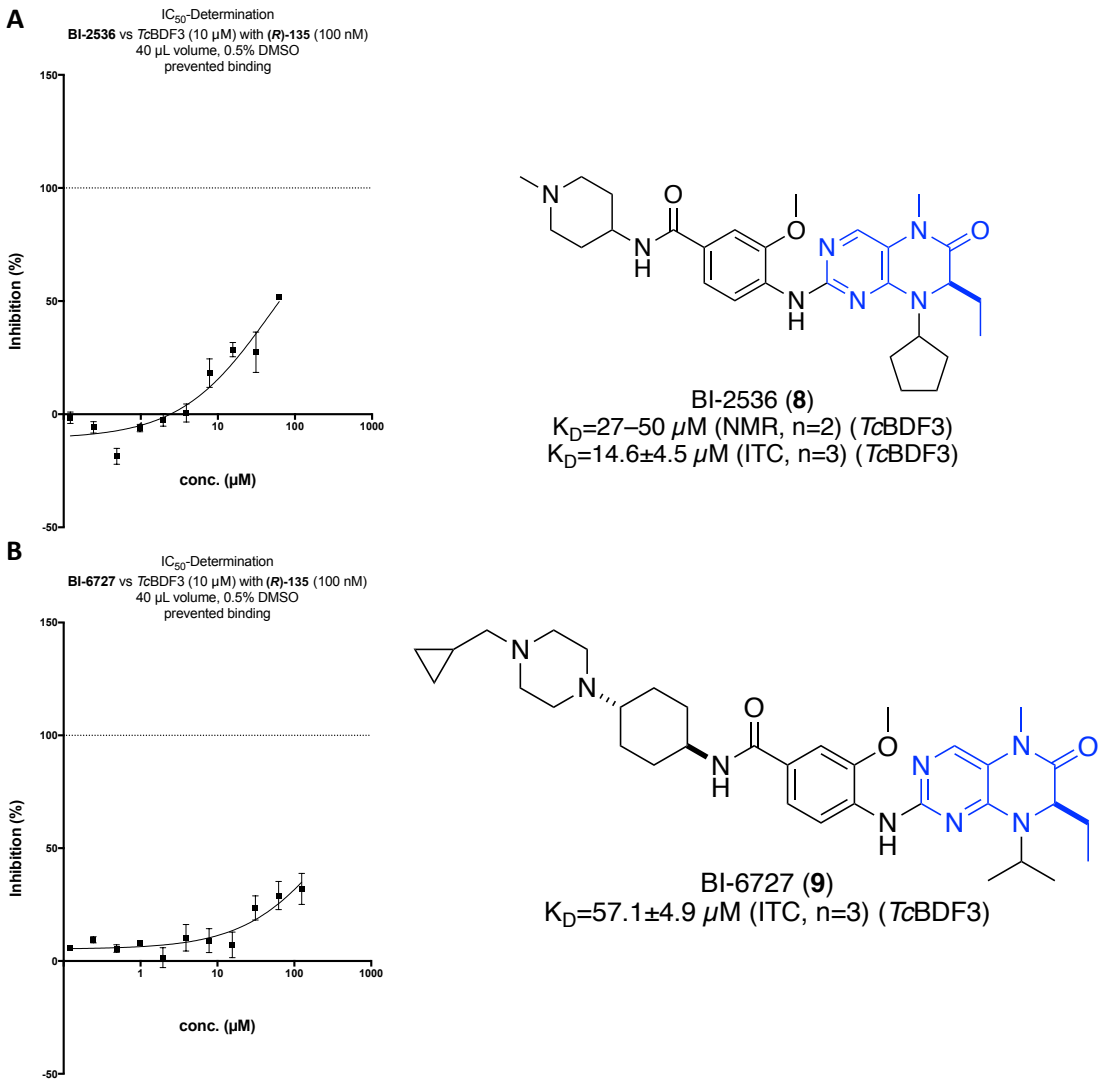


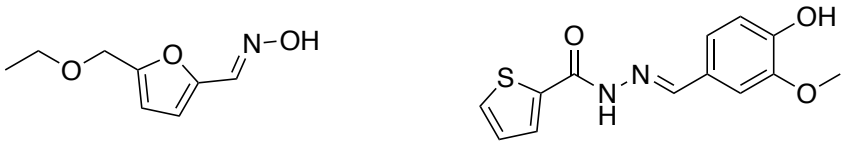
Figure 5.21: Comparison of the inhibition of the interaction of the FP probe **(*R*)-135** and *Tc*BDF3 by BI-2536 (**8**) or BI-6727 (**9**). (A) The prevented binding using BI-2536 (**8**) showed a maximum inhibition of 50%. (B) the prevented binding of **(*R*)-135** to *Tc*BDF3 was also be detected when using BI-6727 (**9**), and showed a maximum inhibition of 35–40%. The KAc mimics are highlighted in blue.

5.4.3.2.b. Using FP to detect the binding of OXFBD02 (**7**)

FP displacement was next attempted with OXFBD02 (**7**), which is estimated to have a 10-fold weaker binding than BI-2536 (**8**) by ^{1}H NMR ligand observed protein titration (Chapter 3). Using the tool compound **(*R*)-135** no binding was observed by FP (Appendix B6). This is additional evidence that **(*R*)-135** can be used to detect binding of ligands that bind *Tc*BDF3 with affinities lower or equal to BI-2536 (**8**). However, cannot detect weak ligand with K_D values above 100 µM.

5.4.3.2.c. Using FP to detect the binding of **15** and **16**

Following the detection of the binding of BI-6727 (**9**) to *Tc*BDF3 through FP, other compounds were considered for screening. Of particular interest were the recently published *Tc*BDF3 binding fragments **15** and **16**.[69-70] These compounds were reported by the Serra research group to bind *Tc*BDF3 with K_D values below 2 µM in a tryptophan fluorescence assay (**15**: K_D= 0.29 µM (Trp fluorescence), **16**: K_D= 1.7 µM (Trp fluorescence)).[69-70] It is important to note that the Serra group used the full-length construct of *Tc*BDF3, which is insoluble when expressed recombinantly. In addition to the concern about the folding state of the protein construct used in these publications, only one biophysical assay is used to determine compound affinity.



15 **16**

Figure 5.22: The structures of the *Tc*BDF3 literature compounds, **15** and **16**, published by the Serra research group within the course of this work.[69-70]

To assess the binding of **15** and **16**, ITC titration experiments were performed using the soluble H_6-*Tc*BDF3(32..148aa) construct (Chapter 4). However, no binding was detected. ITC has been used to detect the binding of BI-2536 (**8**) and other analogues, and should detect the binding of **15** and **16** if their binding affinities are below 2 µM. To further verify these findings, the FP displacement assays was utilised to access the displacement of **(*R*)-135** by **15**. Despite the reported K_D values below 0.5 µM for **15**, no FP inhibition could be detected in neither the displacement nor the prevented binding set up (Figure 5.23).

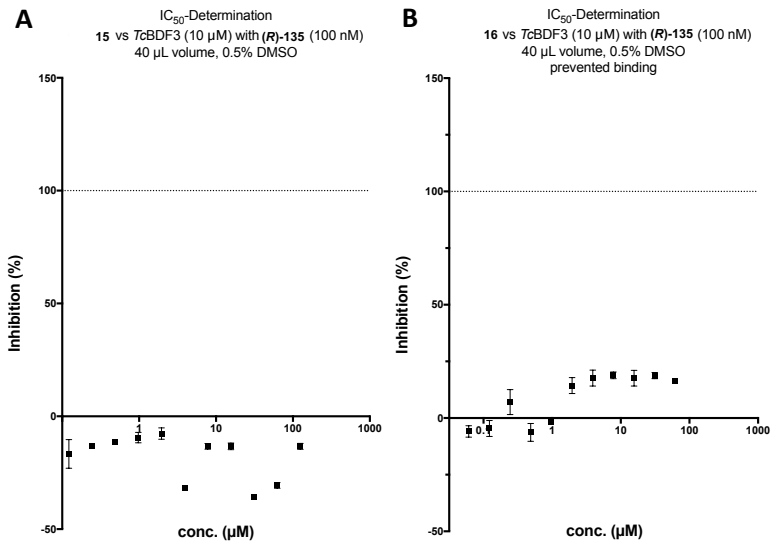


Figure 5.23: The FP inhibition assay do not detect inhibition of the interaction between **(*R*)-135** and *Tc*BDF3 by the literature ligand **15** in neither the (A) displacement nor the (B) prevented binding assay set up.

In the case of BI-2536 (**8**) and its analogues, the binding to *Tc*BDF3 has been detected by several orthogonal assays: waterLOGSY, ^{1}H NMR ligand observed protein titration, ITC and FP. The results of these assays have shown good correlation. For example, the binding of BI-2536 (**8**) has been shown through all these techniques, and the binding of the weaker ligand OXFBD02 (**7**) has been shown by the first two. Therefore, to further build evidence of the lack of measurable binding of the literature compounds **15** and **16**, they were also screened by waterLOGSY. **15** did not show binding when screened by waterLOGSY. **16** did not show suitable solubility down to 75 µM. The low solubility had been expected as this compound precipitated with the addition of water as a purification method during its synthesis (see section 4.8.2).

Table 5.10: Summary of the binding of the proposed *Tc*BDF3 binders **15** and **16** using orthogonal assays, using BI-2536 (**8**) as a comparison.

Technique	BI-2536 (**8**)	**15**	**16**
waterLOGSY	Binding detected	No binding detected	Not soluble
ITC	K_D=14.6±3.0 µM (n=3)	No binding detected (n=2)	No binding detected (n=2)
FP using **(*R*)-135**	Partial displacement (up to 50% inhibition)	No displacement	N/D
Tryptophan fluorescence assays	N/D	K_D= 0.29 µM[69]	K_D= 1.7 µM[70]

Overall, multiple assays, set up and validated as part of this work, have been used to show that **15** and **16** do not show binding to *Tc*BDF3 with measurable binding affinities, contrary to the reported findings of the Serra research group.[69-70]

5.4.4. Displacement assay set up for BRD4(1), BRD4(2) and PLK1

(*R*)-135 shows high affinity for BRD4(1), BRD4(2) and PLK1 (Section 5.4.2). The set-up of a FP displacement assay for these proteins would allow for the screening of binders for *Tc*BDF3 and the assessment of their selectivity over human protein simultaneously in one assay set up in a 384-well plate. Therefore, the next priority was the validate the use of the FP probe **(*R*)-135** for BRD4(1), BRD4(2) and PLK1. This was completed through the displacement of the known dual PLK1/BRD4 ligands BI-2536 (**8**) and BI-6727 (**9**) and the BRD4 ligand OXFBD02 (**7**).

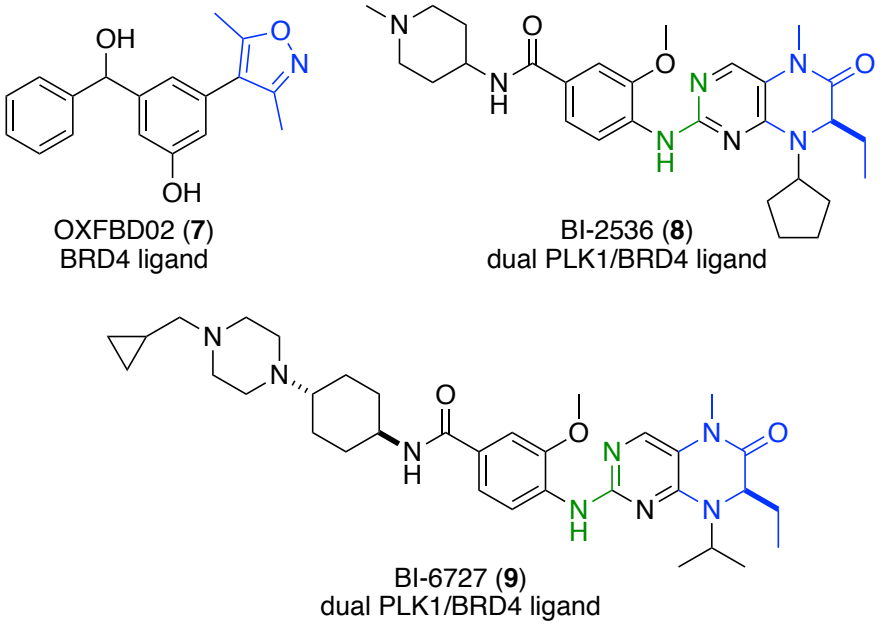


Figure 5.24: The structures of the BRD4 ligand OXFBD02 (**7**), and the dual PLK1/BRD4 ligands BI-2536 (**8**) and BI-6727 (**9**). The KAc mimic are highlighted in blue and the kinase hinge binding regions are highlighted in green.

As expected, the dual PLK1/BRD4 ligands BI-2536 (**8**) and BI-6727 (**9**) showed efficient displacement of the FP probe **(*R*)-135** from BRD4 and PLK1 proteins. The BRD4 ligand OXFBD02 (**7**), showed displacement for both BRD4 domains but not PLK1 (Table 5.11). IC_{50} values are highly assay dependent as they depend on the binding affinity of the displaced molecules.

Table 5.11: Summary of the IC_{50} values obtained using FP for the displacement of **(*R*)-135** from BRD4(1), BRD4(2) and PLK1 by BI-2536 (**8**), BI-6727 (**9**) or OXFBD02 (**7**). AS denotes AlphaScreen.

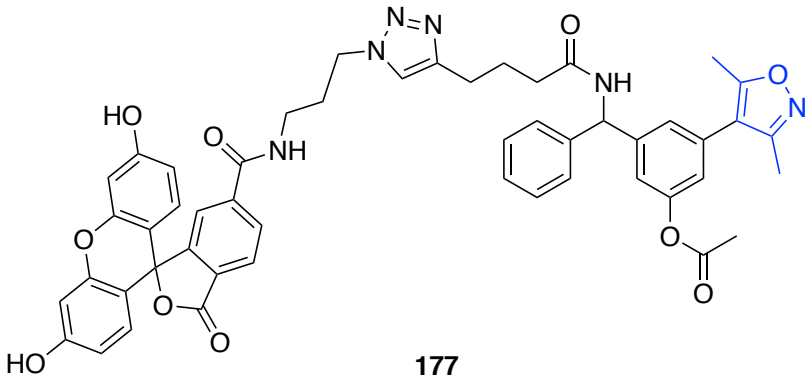
	FP assay with **(*R*)-135**			Literature values	
IC_{50} values	BRD4(1)	BRD4(2)	PLK1	BRD4(1)	PLK1
BI-2536 (**8**)	169 nM	262 nM	45 nM	IC_{50}= 25 nM (AS)[39]	IC_{50}=0.83 nM[92]
BI-6727 (**9**)	632 nM	722 nM	98 nM	K_D=79 nM (ITC)[38-39]	IC_{50}=0.87 nM[102]
OXFBD02 (**7**)	2773 nM	1019 nM	–	IC_{50}=382 nM (AS)[37]	–

5.4.5. Using the **(*R*)-135** for the study of BD proteins from other parasites

Similar to *T. cruzi*, the *Schistosoma mansoni (S. mansoni)* and *Leishmania donovani* (*L. donovani*) parasites are the causes of neglected diseases. [3, 117] The study of the *S. mansoni* and *L. donovani* BD proteins could offer additional insight into the effect of the inhibition of specific epigenetic mechanisms in parasites.[67] Therefore, it was of interest to determine if the tool compounds establish as part of this work could be used to study the *S. mansoni* and *L. donovani* BD proteins.

5.4.5.1. FP assays for the *Schistosoma mansoni* parasite BD proteins

The *S. mansoni* parasite causes schistosomiasis, which is a life-threatening disease that affects 210 million people worldwide. The disease is spread by contact with contaminated water. Schistosomiasis is treated using the drug praziquantel; however, resistance has been observed.[117] Work in the group (Prof. Stuart Conway and Dr Matthias Schiedel, University of Oxford)) has shown that the genome of *S. mansoni*[118] encodes for BD-containing proteins, including *Sm*BRD3, which is a tandem BD protein with over 30% sequence identity with human BRD3. A DMI-based FP probe **177** (Figure 5.25: unpublished work, Dr Matthias Schiedel) binds *Sm*BRD3(1) and the dual construct *Sm*BRD3(1,2) with K_D values of 2.23 µM and 495 nM, respectively. The FP probe **177** did not show binding to the second domain of *Sm*BRD3. *Tc*BDF3 has a 30% sequence identity to the second domain of the human BET proteins, and so the fluorescent probes developed for this protein as part of this work were thought to potentially offer new tool compounds to study *Sm*BRD3(2).



177
*Sm*BRD3-1 FP probe

Figure 5.25: The structure of *Sm*BRD3-1 FP probe **177** designed and synthesised by Dr Matthias Schiedel. The FP probe **177** does not bind *Sm*BRD3-2.

The fluorescent probe **(*R*)-135** showed high-affinity binding to *Sm*BRD3(2) (K_D=374.5±23.5 nM) and an even higher affinity to the tandem construct *Sm*BRD3(1,2) (K_D=38.0±5.2 nM). This was followed by development of a displacement assay of **(*R*)-135** for *Sm*BRD3(2) to interrogate its suitability to screen for potential competitors (Dr Matthias Schiedel). The ligands I-BET726 (**6**) and BI-2536 (**8**) were chosen as they were shown to bind to *Sm*BRD3(2) by ITC (K_D of 3.81±0.58 µM and 1.85±0.36 µM, respectively). Using 40 nM of **(*R*)-135** and a protein concentration of 1 µM, the ligands were tested by pre-incubating the FP probe and *Sm*BRD3(2). Both ligands displaced **(*R*)-135**, with IC_{50} values of 10.7±1.14 µM and 5.33±0.47 µM, respectively (Figure 5.26). The IC_{50} values obtained for the FP assays with **(*R*)-135** are in good agreement with the ITC K_D values (Table 5.12).

Table 5.12: Comparison of the K_D and IC_{50} values obtained for *Sm*BRD3(2) by ITC or FP using **(*R*)-135**. Experiments performed by Dr Matthias Schiedel.

Compounds	ITC assay (K_D)	FP assay (IC_{50})
BI-2536 (**8**)	3.81 µM ± 0.576 µM	10.7 µM ± 1.14 µM
I-BET726 (**6**)	1.85 µM ± 0.361 µM	5.33 µM ± 0.47 µM

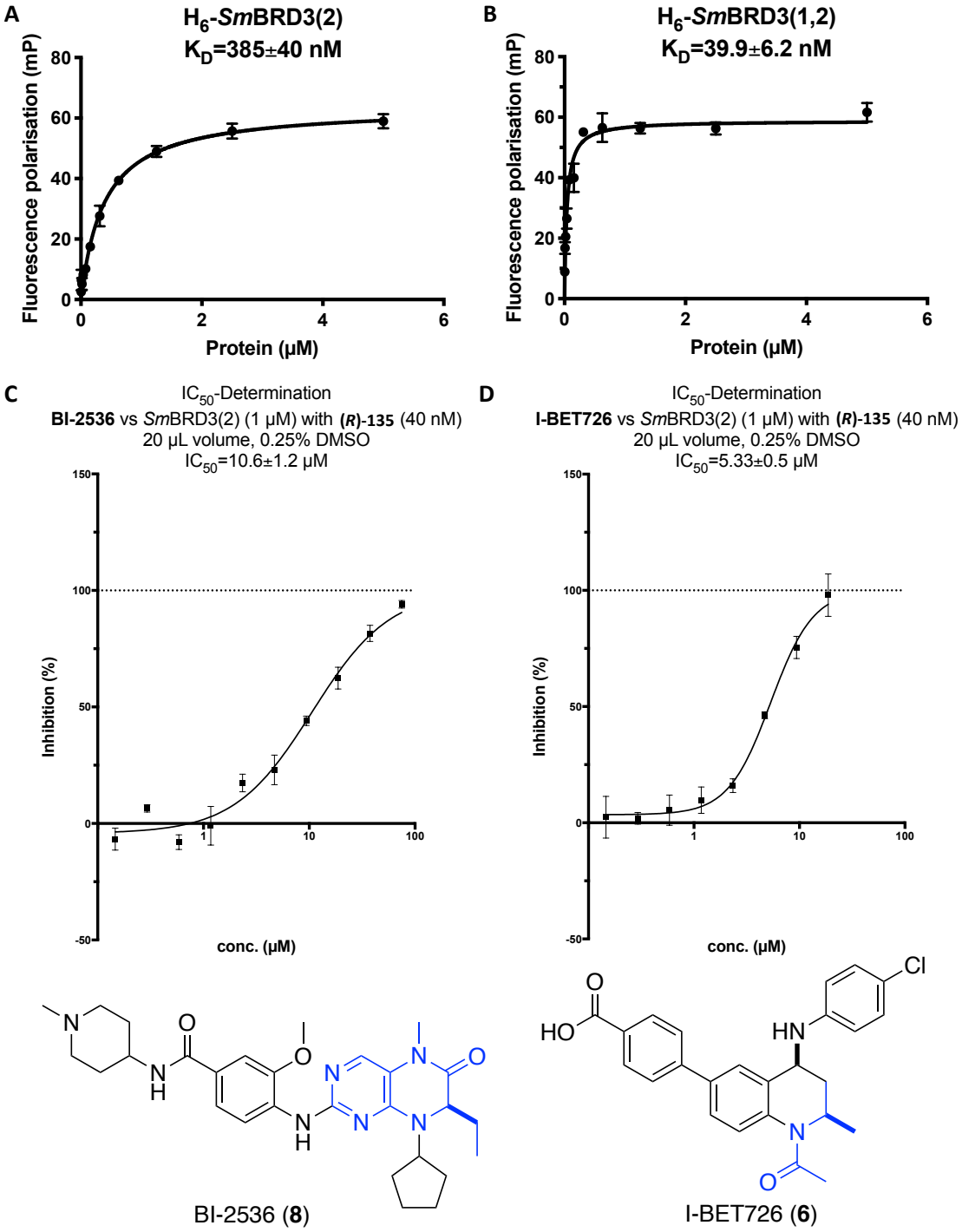


Figure 5.26: The use of **(*R*)-135** for the study of *Sm*BRD3(1,2) and *Sm*BRD3(2) by FP. (A) the binding curve of **(*R*)-135** to *Sm*BRD3(1,2) and (B) *Sm*BRD3(2). The displacement of **(*R*)-135** from *Sm*BRD3-2 by (C) BI-2536 (**8**) and (D) I-BET726 (**6**). Displacement experiments performed by Dr Matthias Schiedel (University of Oxford).

In conclusion, **(*R*)-135**, originally designed for the study of *Tc*BDF3, was also established as a screening tool for *Sm*BRD3(2). This has provided a complimentary assay to the one established for *Sm*BRD3(1) by Dr Matthias Schiedel (University of Oxford) using **177**.

5.4.5.2. FP assays for the *Leishmania donovani* parasite BD proteins

As described in Section 5.2.2, *Ld*BD2, *Ld*BD5(1) and *Ld*BD5(2) from the homologous protozoan parasite *Leishmania donovani* were provided by Dr Nathaniel Jones from the Wilkinson group (University of York). Interestingly, there is a 53% sequence identity between the *Tc*BDF5-1 and *Ld*BD5(1), 49% between *Tc*BDF2 and *Ld*BD2, and 42% between *Tc*BDF5-2 and *Ld*BD5(2). Additionally, *Ld*BD5(1) has been co-crystallised with the BI-2536 (**8**) ligand (PDB: 5TCM). Taken together, these data suggest that the tool compounds and biophysical assays developed for the *T. cruzi* BDs might also be useful for the study of these proteins. Therefore, it was decided to screen the **(*R*)-135** tool compound for binding to *Ld*BD2, *Ld*BD5(1) and *Ld*BD5(2).

The titration of the proteins into **(*R*)-135** was monitored by the increase in FP (Figure 5.27). *Ld*BD2 and *Ld*BD5-1 do not show binding; however, *Ld*BD5-2 binds the FP with a K_D of 2.4±1.1 µM (n=3) and shows a large difference (> 100 mP) in FP of the free and bound probe, giving a suitable assay window.

The titrations were followed by attempting to obtain an IC_{50} values for the displacement of the FP probe **(*R*)-135** with the BI-2536 (**8**) ligand. No inhibition was detected, neither in the displacement setup (with pre-incubation of the FP probe and the protein), nor the prevented binding setup (with pre-incubation of the competitor ligand and the protein).

The direct displacement of the fluorescent probe from *Ld*BD5-2 seems to be challenging, and attempts at optimising the inhibition experiments were not successful (Appendix B6). Further optimisation was prevented by the absence of binding data of BI-2536 (**8**) to the studied *Ld*BD proteins. This was also complicated by the absence of a structurally different positive control which could have been used for the displacement assay.

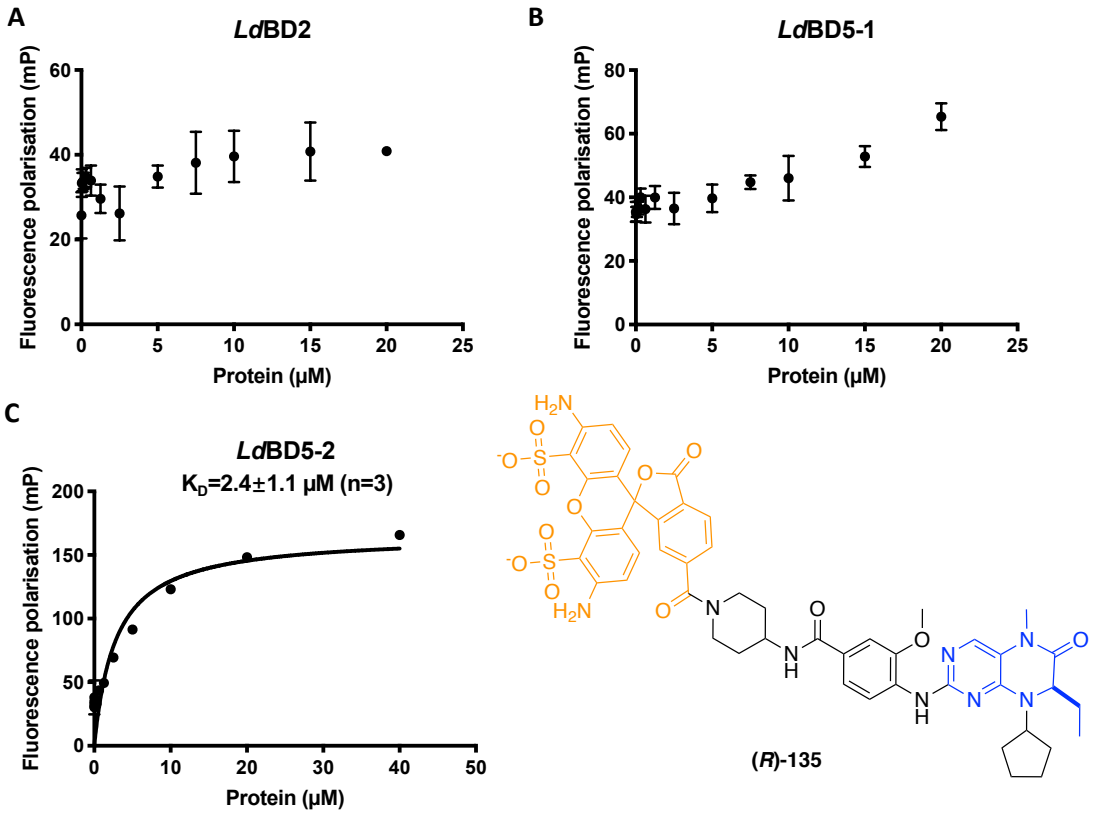


Figure 5.27: The binding curves for **(*R*)-135** titrated with (A) *Ld*BD2, (B) *Ld*BD5-1 and (C-E) *Ld*BD5-2 observed by fluorescence polarisation. *Ld*BD5-2 shows binding to the FP **(*R*)-135** with a K_D values of 2.4±1.1 μM (FP, n=3) a large difference (> 100 mP) in FP of the free and bound probe, giving a suitable assay window.

Overall, the fluorescent tool compound **(*R*)-135** has been shown to binding *Ld*BD5-2 and provides the first FP probe for this protein. The setup of the displacement assay was not successful, and should be repeated when well-characterised *Ld*BD5-2 binders have been identified and can be used as reliable positive controls for the assay set-up.

5.5. Conclusions

^{19}F NMR reporter molecules, diazirine, and fluorescent probes were designed and synthesised to set up orthogonal and higher throughput assays (Chapter 4).

This chapter described the development of novel assays for *Tc*BDF3 based on the displacement of the tool compounds monitored by ^{19}F NMR, protein LCMS or fluorescence polarisation. The ^{19}F NMR reporter molecules, **(*R*)-130**, **(*R*)-131** and **112**, have been used to detect the binding of *Tc*BDF3, giving an alternative to waterLOGSY, with shorter experiments times and less constraining solubility requirements. The **133** and **161** diazirines have been shown to cross-link to *Tc*BDF3 in a dose-dependent manner, and are the first reported photo-activatable molecules designed for *Tc*BDF3. The PhaBits diazirine library was screened and several novel fragments have been identified for *Tc*BDF3 and *Tc*BDF5-1. Finally, the fluorescent probe **(*R*)-135** was used to set up the first FP assay for *Tc*BDF3, giving the first medium throughput assay for this protein. Only a partial displacement curve was obtained for BI-2536 (**8**), used as a positive control; however, the assay could be used to detect binder with higher affinities. **(*R*)-135** was also shown to bind *Sm*BRD3-2 and *Ld*BD5-2, additionally giving novel assays for BDs from different parasite species.

Following this work, there is a now a varied selection of tool compounds for the study of *Tc*BDF3, providing orthogonal biophysical assays to study *Tc*BDF3 and other parasite BD proteins efficiently. This provides a considerable improvement on the initial assays, with higher throughput, lowered protein usage and the use of readily available machines. Taken together, these results suggest that the newly established biophysical and biochemical assays for *Tc*BDF3 will allow for the faster screening and characterisation of binders. This opens up the possibility of rapid SAR work, as well as continued improved of the techniques established here.

Chapter 6. Conclusions and future work

The *T. cruzi* parasite relies on switching between distinct phenotypes through epigenetic mechanisms to thrive in humans. This led to the hypothesis that interference with epigenetic mechanisms would be detrimental to parasite survival in humans. The aim of this project was to design, synthesise and optimise small molecule ligands to probe the function of BCPs from the parasite *T. cruzi*.

Soluble *T. cruzi* BD proteins were expressed recombinantly (Chapter 2). Although crystallisation attempts were not successful, the optimisation of the protein production yielded high purity protein suitable for biophysical assays for *Tc*BDF2, *Tc*BDF3, and *Tc*BDF5-1. Screening using waterLOGSY showed that the *T. cruzi* BDs do indeed bind KAc-mimicking molecules (Chapter 3). The use of waterLOGSY and ^{1}H NMR ligand-observed protein titrations has identified BI-2536 (**8**) as the highest affinity ligand to bind *Tc*BDF3, with a K_D value of 27–55 μM by NMR titration and of 14.6±4.5 μM (n=3) by ITC. Therefore, BI-2536 (**8**) was chosen as the lead structure for the synthesis of tool compounds. Eight functional tool compounds for *Tc*BDF3 were designed and synthesised to provide new reporter molecules for this *T. cruzi* BDs: two ^{19}F NMR reporter molecules, two photo-crosslinking diazirines and four fluorescent probes (Chapter 4).

Finally, novel assays for *Tc*BDF3 were developed using the displacement of the tool compounds monitored by ^{19}F NMR, protein LCMS or fluorescence polarisation. The ^{19}F NMR reporter molecules, **(*R*)-130**, **(*R*)-131** and **112**, have been used to detect the binding of *Tc*BDF3, giving an alternative to waterLOGSY with shorter experiments times and less constraining solubility requirements. The **133** and **161** diazirines have been shown to cross-link to *Tc*BDF3 in a dose-dependent manner and are the first reported photo-activatable molecules designed for *Tc*BDF3. Finally, the fluorescent probe **(*R*)-135** was used to set up the first FP assay for *Tc*BDF3. Only a partial displacement curve was obtained for BI-2536 (**8**); however, we hypothesise that the assay can be used to detect binders with higher affinities. **(*R*)-135** was also shown to bind *Sm*BRD3-2 and *Ld*BD5-2, additionally giving novel assays for BDs from different parasite species.

Considerable improvements on assay throughput has been achieved using the tool compounds, however, further validation and optimisation is required. The FP assay with (*R*)-**135** has the highest throughput but higher affinity ligands need to be identified to provide full displacement curves and show that this assay could be used to provide IC_{50} values. Future work should focus on screening analogues of BI-2536 (**8**) and other BD ligands using the assays established during this work to identify higher affinity ligands. These ligands will be essential to further develop the assays, generate SAR data, and could be used in co-crystallisation trials for *Tc*BDF3.

The orthogonal assays used in this work demonstrated that several KAc-mimicking ligands bind the *T. cruzi* BD proteins analogously to their human homologues. Several differences, however, have been observed. Firstly, the homology models of *Tc*BDF2, *Tc*BDF3, and *Tc*BDF5-1 suggest higher flexibility and shallower binding sites compared to human BCPs. Additionally, *Tc*BDF2 and *Tc*BDF3 did not produce reliable data using thermal shift which is commonly used for BDs, suggesting that these proteins are more challenging targets than their human counterparts.

A PhaBits diazirine library was screened and several novel fragments have been identified for *Tc*BDF3 and *Tc*BDF5-1. Additionally, a DNA-encoded library was screened and clusters binding *Tc*BDF2 and *Tc*BDF5-1 have been discovered. However, these two screening campaigns gave overall low hit rates which confirms that *T. cruzi* BD proteins are less readily ligandable compared to human BD proteins, such as BRD4(1).

Following the work detailed in this thesis, there is a now a varied selection of orthogonal biophysical and cbiochemical assays for the study of *Tc*BDF3 and other parasite BD proteins. This work provides a significant improvement on the initial assays, with higher throughput, lowered protein usage and the use of readily available machines. Altogether, these results suggest that the newly established biophysical assays for *Tc*BDF3 will allow for the faster screening and characterisation of binders. This opens up the possibility for rapid SAR work, as well as continued improvement of the techniques established here.

Future work should focus on SAR to achieve higher affinity towards *Tc*BDF3 and selectivity over human proteins. This work would allow for the identification of selective ligands and control compounds to be used in phenotypic screening to determine whether specific BD inhibition in *T. cruzi* leads to trypanocidal activity. Further work could also include the design and synthesis of a *Tc*BDF3 selective pull-down probe to study target engagement in the parasite. Further elucidation of the function of the *T. cruzi* BD proteins would determine if their inhibition can alter the *T. cruzi* phenotype and if it can be used as a strategy to develop novel treatments for Chagas disease.

Chapter 7. Experimental

7.1. Synthetic method and materials

All reactions under an inert atmosphere were carried out using flame-dried glassware. Solutions were added *via* syringe or cannula. Anhydrous solvents were dried by passing through a column of activated basic alumnia according to Grubbs' procedure,[119] or obtained from an MBRAUN Solvent Purification System 5, and stored over activated 3 Å molecular sieves under argon. When stated, THF was distilled from sodium and benzophenone. Solvents were degassed using a flow of argon and sonication for a minimum of 15 mins. Celite® refers to Celite® 545 filter aid, treated with sodium carbonate, flux-calcined (Sigma Aldrich).

Reagents were obtained from Acros Organics, Alfa Aesar, Apollo Scientific, Fisher Scientific, Fluka, Fluorochem, Merck or Sigma Aldrich and were used without further purification unless otherwise stated.

NMR data was processed using MestReNova® 10.0 or Bruker TopSpin® v3.5 or v4.0 software. Spectra were assigned using COSY, HSQC and HMBC experiments as necessary.

^{1}H NMR spectra were recorded on Bruker AVIII HD 400 (400 MHz) spectrometer or Bruker AVII 500 (500 MHz) spectrometer. Chemical shifts are reported as δ part per million (ppm), to the nearest 0.01 ppm, relative to the solvent reference peak: $CDCl_3$ (δ_H = 7.26 ppm), CD_3OD (δ_H = 3.31 ppm), acetone-D_6 (δ_H = 2.05 ppm), DMSO-D_6 (δ_H = 2.50 ppm).[120] The multiplicity of each signal is indicated by: s (singlet), d (doublet), t (triplet), dd (doublet of doublets), q (quartet), sp (septet) or m (multiplet). Identical proton coupling constants (*J*) are averaged in each spectrum and are reported to the nearest 0.1 Hz. Spectra were assignment using COSY, HSQC, and HMBC experiments as necessary.

^{13}C NMR spectra were recorded on Bruker AVIII HD 400 (101 MHz) spectrometer or Bruker AVII 500 (126 MHz) spectrometer. Chemical shifts are reported as δ_C part per million (ppm) relative to the solvent reference peak to the nearest 0.1 ppm, unless peaks

are indistinguishable, in which case they are quoted to the nearest 0.01 ppm. Residual protic solvent was used as the internal reference: $CDCl_3$ (δ_C = 77.2 ppm), CD_3OD (δ_C = 49.0 ppm), acetone-d_6 (δ_C = 29.8 ppm), DMSO-d_6 (δ_C = 39.5 ppm).[120] The signals in the ^{13}C NMR spectra were labelled as C, CH, CH_2 or CH_3 according to the DEPT spectrum. The multiplicity of each signal is singlet unless otherwise stated.

^{11}B NMR spectra were recorded on Bruker AVII 500 (160 MHz) spectrometer. Chemical shifts are reported as δ_B part per million (ppm) to the nearest 0.1 ppm.

^{19}F NMR spectra were recorded on Bruker AVIII HD 400 (377 MHz) spectrometer or on Bruker AVII 500 (470 MHz) spectrometer. Chemical shifts are reported as δ_F part per million (ppm), to the nearest 0.1 ppm.

Thin layer chromatography (TLC) was performed on normal phase Merck silica gel 60 F_{245} aluminium-supported thin layer chromatography sheets. Spots were visualised by either absorption under UV light (254 nm) or by thermal development after dipping into either an alkaline solution of $KMnO_4$, or an ethanolic solution of ninhydrin. Reaction progress was monitored at appropriate times by TLC analysis.

Infrared spectroscopy (IR) were recorded on a Bruker Tensor 27 FT-IR spectrometer from a thin film deposited onto a diamond ATR module. Only selected maximum absorbances $\tilde{\nu}_{max}$ of the most intense peaks are reported (cm^{-1}): br = broad, w = weak, m = medium, s = strong.

Column chromatography was performed either using Geduran Silicagel 60 (40–63 µm) under a positive pressure of compressed nitrogen, or on a Biotage SP1 automated column chromatography system using KP-Sil® SNAP Flash Silica Cartridges where stated, or on a Teledyne ISCO Combiflash Rf automated chromatography system using Redi*Sep*® Rf and Redi*Sep* Rf Gold® silica cartridges.

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).

Melting points on recrystallised samples were determined using either a) a Leica Galen III hot stage microscope or b) a Griffin capillary tube melting point apparatus and are uncorrected. The solvents of crystallisation are shown in parentheses.

Specific optical rotation $[\alpha]_D^T$ were measured using a PerkinElmer Model 241 polarimeter, in cells with a path length of 1 dm at 25 °C and a light source maintained at 589 nm. The concentration (*c*) is expressed in g/100 mL. $[\alpha]_D^T$values are reported in implied units of 10^{-1} °cm^2g^{-1}.

In vacuo refers to the removal of solvent on a Buchi™ rotary evaporator under reduced pressure with a water bath at 40 °C.

Brine refers to a saturated aqueous solution of sodium chloride.

Petroleum ether refers to the fraction boiling between 40–60 °C.

Analytical HPLC was carried out on a PerkinElmer Flexar system with a Binary LC Pump and UV/Vis LC Detector. For determination of enantiomeric purity, a ChiralPak AD-H column (5 µm, 4.6 × 150 mm), ChiralPak AS-H column (5 µm, 4.6 × 150 mm), or a ChiralPak OG column (10 µm, 4.6 × 150 mm) was employed with an isocratic method of 20 to 60 mins, with flow rates between 0.5 and 1.2 mL/min. Samples were injected in a mixture of heptane and IPA or CH_2Cl_2. For determination of compound purity, a reversed phase (RP) a Dionex Acclaim® 120 column (C18, 5 µm, 12 Å, 4.6 × 150 mm) was employed, with a flow rate 1 mL/min and detection at 254 nm. Samples were injected in H_2O, H_2O and MeCN, DMSO or methanol. The gradients employed for RP are as described below:
Method A: 10-minute gradient of 95% H_2O/5% MeCN (Solvent A) to 95% MeCN/5% H_2O (Solvent B)

Method B: 10-minute gradient of 95% H_2O/5% MeCN/0.1% TFA (Solvent A) to 95% MeCN/5% H_2O/0.1% TFA (Solvent B)

Liquid chromatography mass spectrometry (LCMS)

For determination of compound purity and mass by LCMS, an Acquity UPLC CSH C18 column (50 mm x 2.1 mm, i.d. 1.7 µm packing diameter) was employed at 40 °C. The UV detection was a summed signal from wavelength of 210 nm to 350 nm, and the MS detection was conducted using Waters QDA spectrometer with alternate-scan positive and negative electrospray (100 to 1000 AMU; scan frequency: 5 Hz). *m/z* values are reported in Daltons and followed by the corresponding purity determined by the LC peak intensity percentage abundance in parentheses. The 2 mins methods for LCMS were either "high pH" or "formic acid" as describe below:

- "high pH"

A = 10 mM ammonium bicarbonate in H_2O adjusted to pH 10 with ammonia solution.
B = Acetonitrile.

The gradient employed was:

Time (min)	Flow rate (mL/min)	%A	%B
0	1	100	0
0.05	1	100	0
1.5	1	3	97
1.9	1	3	97
2	1	100	0

- "Formic acid"

A = 0.1% *v/v* Solution of formic acid in H_2O.
B = 0.1% *v/v* Solution of formic acid in acetonitrile.

The gradient employed was:

Time (min)	Flow rate (mL/min)	%A	%B
0	1	97	3
1.5	1	3	97
1.9	1	3	97
2.0	1	98	2

Mass-directed assisted purification (MDAP)

For the purification of small amounts of compounds, MDAP HPLC was conducted on a Xselect CSH C18 column (150 mm x 30 mm i.d. 5 µm packing diameter) at ambient temperature. The UV detection was a summed signal from wavelength of 210 nm to 350 nm, and the MS detection was conducted using a Waters QDA spectrometer with alternate scan positive/negative electrospray (scan range : 100 to 1000 AMU; scan frequency : 4.3 Hz). The samples (up to 100 mg) were injected in 1 mL of DMSO. The 15 mins methods for MDAP PLC were either "MDAP method A", "MDAP method B" or "MDAP method C" as described below:

A = 0.1% *v/v* solution of formic acid in H_2O.
B = 0.1% *v/v* solution of formic acid in acetonitrile.

- "MDAP Method A"; formic acid focused preparative open access LC/MS

The gradient employed was:

Time (min)	Flow rate (mL/min)	%A	%B
0	40	100	0
1	40	100	0
10	40	75	25
10.5	40	1	99
15	40	1	99

- "MDAP Method B"; formic acid focused preparative open access LC/MS gradient

The gradient employed was:

Time (min)	Flow rate (mL/min)	%A	%B
0	40	85	15
1	40	85	15
10	40	45	55
10.5	40	1	99
15	40	1	99

- "MDAP Method C"; formic acid focused preparative open access LC/MS gradient

The gradient employed was:

Time (min)	Flow rate (mL/min)	%A	%B
0	40	70	30
1	40	70	30
10	40	15	85
10.5	40	1	99
15	40	1	99

High resolution liquid chromatography mass spectrometry (HR-LCMS)

For the determination of high resolution mass spectra by LCMS, an Acquity UPLC CSH C18 column (100 mm × 2.1 mm i.d. 1.7 µm packing diameter) at 50 °C was employed. The UV detection was a summed signal from wavelength of 210 nm to 350 nm, and the MS detection was acquired using a Waters XEVO G2-XS QTof spectromer with positive electrospray ionisation. The HRMS are reported by their retention times (rt) in mins, followed by the corresponding purity determined by the LC peak intensity percentage abundance in parentheses, then the detected *m/z* values in Daltons, followed by the relative *m/z* peak intensity in parentheses.

A = 0.1% *v/v* solution of Formic Acid in H_2O.

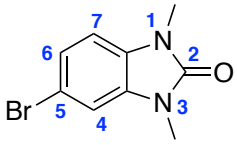
B = 0.1% *v/v* solution of Formic Acid in Acetonitrile.

The gradient employed was:

Time (min)	Flow Rate (mL/min)	%A	%B
0	0.8	97	3
8.5	0.8	0	100
9.0	0.8	0	100
9.5	0.8	97	3
10	0.8	97	3

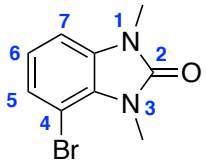
7.2. Synthesis

5-Bromo-1,3-dimethyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (17)



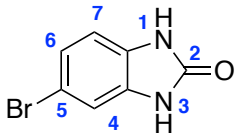
MeI (400 µL, 6.37 mmol, 3.00 eq.) was added to a solution of 5-bromo-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**20**) (450 mg, 2.12 mmol, 1.00 eq.) and K_2CO_3 (1.76 g, 12.7 mmol, 6.00 eq.) in DMF (4 mL), and the reaction mixture was stirred at room temperature for 6 h. H_2O (30 mL) was added and the mixture was extracted with EtOAc (3 × 50 mL). The organic components were washed with LiCl (10% *w/v*, 6 × 100 mL) then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo* to afford 5-bromo-1,3-dimethyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**17**) (497 mg, 97%) as a light brown solid: R_f 0.85 (CH_2Cl_2:MeOH = 9:1); m.p. > 300 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3065 (w), 2980 (w), 1697 (s) (C=O), 1505 (m), 1446 (m), 1240 (w), 1134 (w), 942(w), 857 (w), 788, 680(w); ^{1}H NMR (400 MHz, DMSO-D_6) 7.22 (1 H, dd, *J* 8.2, 1.8, C^7-H), 7.11 (1 H, d, *J* 1.8, C^4-H), 6.83 (1 H, d, *J* 8.2, C^6-H), 3.405 (3 H, s, N-CH_3) 3.395 (3 H, s, N-CH_3); ^{13}C NMR (101 MHz, DMSO-D_6) δ 154.2 (C^2), 131.6 (C^{3a}), 129.4 (C^{7a}), 123.7 (C^4), 113.2 (C^5), 111.0 (C^7), 109.7 (C^6), 27.58 (NCH_3), 27.52 (NCH_3); LRMS *m/z* (ES^+) 241 ($[M^{81}Br+H]^+$, 100%), 243 ($[M^{79}Br+H]^+$, 70%); HRMS *m/z* (ES^+) found 242.99514, $C_9H_9{}^{81}BrN_2O$ requires 242.99506. The data are in good agreement with the literature.[97]

4-Bromo-1,3-dimethyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (18)



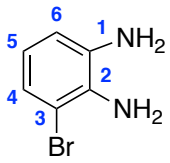
MeI (705 µL, 11.3 mol, 3.0 eq.) was added dropwise to a solution of 4-bromo-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**23**) (1.0 g, 3.77 mmol, 1.0 eq.) and K_2CO_3 (3.13 g, 22.7 mol, 6.0 eq.) in DMF (1.5 mL). The reaction mixture was stirred at room temperature for 2 h. H_2O (30 mL) was added, the precipitate was collected by filtration and dried under vacuum, affording 4-bromo-1,3-dimethyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**18**) (825 mg, 90%) as brown solid: R_f 0.85 (CH_2Cl_2:MeOH = 9:1); m.p. > 300 °C (H_2O); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2946 (w) (CH), 1698 (s) (C=O), 1491 (m), 1461 (m), 1331 (w), 1256 (w), 1087 (m), 1029 (m), 758 (m), 740 (m), 723 (m); ^{1}H NMR (400 MHz, DMSO-D_6) 7.23-7.15 (2 H, m, C^5-H and C^7-H), 7.00 (1 H, dd, *J* 7.9, 7.9, C^6-H), 3.61 (3 H, s, N^1-CH_3), 3.33 (3 H, s, N^3-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 154.4 (C^2), 132.3 (C^{7a}), 127.4 (C^{3a}), 125.9 (C^5), 122.8 (C^6), 107.8 (C^7), 100.9 (C^4), 29.9 (N^1-CH_3), 27.7 (N^3-CH_3); LRMS *m/z* (ES^+) 242 ([M^{81}Br+H]$^+$, 100%), 244 ([M^{79}Br+H]$^+$, 70%), 263 [M^{79}Br+Na]$^+$, 20%), 265 [M^{81}Br+Na]$^+$, 40%); Analytical HPLC (RP, method A) rt 8.67 mins, purity 99%. The data are in good agreement with the literature.[97]

5-Bromo-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (20)



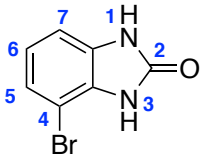
Carbonyldiimidazole (1.95 g, 12.1 mmol, 1.5 eq.) was added to a solution of 5-bromobenzene-1,2-diamine (1.50 g, 8.02 mmol, 1.0 eq.) in 1,4-dioxane (15 mL). The solution was stirred at 40 °C for 1 h. H_2O (10 mL) was added, the precipitate was collected by filtration and washed with H_2O (20 mL), then CH_2Cl_2 (20 mL) and petroleum ether (20 mL). Further precipitate from the aqueous washes was collected by filtration, the solids were combined, and dried *in vacuo* to afford 5-bromo-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**20**) (1.42 g, 83%) as a brown solid: R_f 0.12 (EtOH); m.p. > 300 °C (H_2O); ^{1}H (400 MHz, DMSO-D_6) δ 10.77 (2 H, s, 2 × NH), 7.08 (1 H, dd, *J* 8.2, 1.9, C^6-H), 7.05 (1 H, d, *J* 1.9, C^4-H), 6.87 (1 H, d, *J* 8.2, C^7-H). LRMS *m/z* (ES^+) 215 [M^{81}Br+H]$^+$, 100%), 213 [M^{79}Br+H]$^+$, 70%). The data are in good agreement with the literature.[97]

3-Bromobenzene-1,2-diamine (22)



Iron (0) powder (2.3 g, 41.5 mol, 3.0 eq.), which had been activated with concentrated HCl for 30 mins, then filtered, and washed with H_2O and Et_2O, was added to a solution of 3-bromo-2-nitroaniline (3.0 g, 13.8 mol, 1.0 eq.) in EtOH:H_2O (3:1, 40 mL). NH_4Cl (4.43 g, 82.9 mol, 6.0 eq.) was added portion wise at 0 °C, and the reaction mixture stirred at 90 °C for 16 h. The mixture was cooled to room temperature, and filtered through Celite®, washed with H_2O (160 mL) and EtOH (80 mL). The volatile components were removed *in vacuo* and the mixture extracted with EtOAc (2 × 250 mL). The combined organic components were washed with brine (2 × 150 mL), dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. Purification using 25 g SNAP column on Biotage chromatography system, eluting with CH_2Cl_2:MeOH (1:0 to 9:1), afforded 3-bromobenzene-1,2-diamine (**22**) (1.89 g, 74%) as a light brown solid: R_f 0.61 (CH_2Cl_2:MeOH = 9:1); m.p. 44–46 °C (MeOH) [lit.[121] 52–54 °C]; ^{1}H NMR (400 MHz, $CDCl_3$) 6.97 (1 H, dd, *J* 7.9, 1.4, C^4-H), 6.65 (1 H, dd, *J* 7.7, 1.4, C^6-H), 6.56 (1 H, dd, *J* 7.9, 7.9, C^5-H), 3.81 (2 H, s, NH_2), 3.46 (2 H, NH_2); LRMS *m/z* (ES^+) 189 ([M^{79}Br+H]$^+$, 100%), 187 ([M^{81}Br +H]$^+$, 90%). The data are in good agreement with the literature.[122]

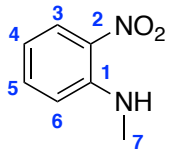
4-Bromo-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (23)



Carbonyldiimidazole (156 mg, 0.96 mmol, 1.2 eq.) was added to a solution of 3-bromobenzene-1,2-diamine (150 mg, 0.80 mmol, 1.0 eq.) in dioxane (3 mL). The solution was stirred at 60 °C for 2 h. H_2O (10 mL) was added, the precipitate was collected by filtration and washed with H_2O (20 mL) to afford 4-bromo-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**23**) (154 mg, 91%) as a pale yellow solid: R_f 0.22 (EtOH); m.p. > 300 °C (H_2O); ^{1}H δ (400 MHz, DMSO-D_6) 11.04 (1 H, s, NH), 10.92 (1 H, s, NH), 7.10 (1 H,

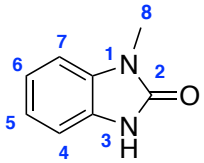
dd, *J* 7.8, 1.3; C^5-H), 6.96–6.80 (2 H, m, C^6-H and C^7-H). LRMS *m/z* (ES^-) 214 [$M^{81}Br$–H]$^-$, 100%), 212 [$M^{79}Br$–H]$^-$, 60%). The data are in good agreement with the literature.[97]

N-Methyl-2-nitroaniline (25)



2-Fluoronitroaniline (1.00 g, 7.08 mol, 1.00 eq.) and $MeNH_2$·HCl (1.43 g, 21.3 mol, 3.00 eq.) were dissolved in NMP (12 mL). DIPEA (6.1 mL, 35.4 mol, 5.00 eq.) was added, the vial was sealed then stirred at 80 °C for 17 h. H_2O (50 mL) was added, and the mixture was extracted with EtOAc (2 × 200 mL). The combined organic components were washed with LiCl (10% *w/v*, 6 × 80 mL) and brine (2 × 100 mL) then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification of the oil by silica gel chromatography (Biotage; SNAP 10 g), eluting with petroleum ether and CH_2Cl_2 (1:0 to 3:1), afforded *N*-methyl-2-nitroaniline (281 mg, 26% yield) as an orange solid: R_f 0.34 (petroleum ether:CH_2Cl_2 = 4:1); m.p. 26–28 °C (CH_2Cl_2) [lit.[123] 34 –35 °C (MeCN)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.18 (1 H, dd, *J* 8.5, 1.3, C^3-H), 8.05 (1 H, s, N-H), 7.47 (1 H, ddd, *J* 8.7, 7.4, 1.3, C^5-H), 6.84 (1 H, dd, *J* 8.7, 1.3, C^6-H), 6.66 (1 H, ddd, *J* 8.5, 7.4, 1.3, C^4-H), 3.03 (3 H, d, *J* 5.1, C^7-H_3); The compound did not ionise under ESI or ES. The data are in good agreement with the literature.[97, 123]

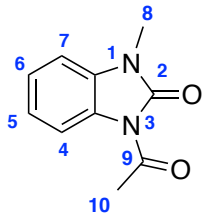
1-Methyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (26)



A solution of *N*-methyl-2-nitroaniline (**25**) (500 mg, 3.28 mmol, 1.0 eq.) and 10% Pd/C (175 mg, 0.164 mmol, 0.050 eq.) in EtOH (8.0 mL) was stirred under H_2-atmosphere for 2 h at room temperature. The reaction mixture was filtered through Celite® and concentrated *in vacuo* to afford *N*-methylbenzene-1,2-diamine. The intermediate was dissolved in dry THF (4 mL), and Et_3N (920 µL, 6.56 mmol, 2.0 eq.) was added. Triphosgene (777 mg, 2.62 mmol, 0.80 eq.) in THF (2 mL) was added dropwise to the

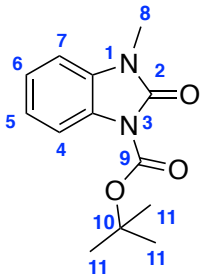
reaction mixture at 0 °C. Precipitate appeared and the reaction mixture was diluted with THF (2 mL). The mixture was left to warm slowly to room temperature and stirred for 18 h. Saturated aq. $NaHCO_3$ (50 mL) was added dropwise, then EtOAc (30 mL) was added and the mixture was stirred for 20 mins. The layers were separated then the aqueous was extracted with EtOAc (2 × 20 mL). The organic components were washed with sat. aq. $NaHCO_3$ (100 mL), brine (100 mL), and H_2O (100 mL), then dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Recrystallisation from the minimum amount of boiling THF afforded 1-methyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**26**) (177 mg, 36% yield) as light brown solid: R_f 0.59 (EtOH); m.p. 166–168 °C (THF) [lit.[124] 187–190 °C (EtOH)]; ^{1}H NMR (400 MHz, DMSO-D6) δ 10.80 (1 H, s, N^3-H), 7.09–7.05 (1 H, m, C^7-H), 7.03–6.94 (3 H, m, C^4-H, C^5-H and C^6-H), 3.27 (3 H, s, C^8-H_3); LRMS *m/z* (ES^-) 248 ([M–H]$^-$, 100 %). The data are in good agreement with the literature.[125]

1-Acetyl-3-methyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (27)



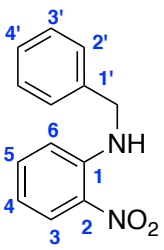
1-Methyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**26**) (50 mg, 0.34 mmol, 1.0 eq.) and DMAP (10 mg, 0.07 mmol, 0.20 eq.) were dissolved in CH_2Cl_2 (1 mL). Et_3N (60 µL, 0.40 mmol, 1.2 eq.) and Ac_2O (90 µL, 0.37 mmol, 1.1 eq.) were added, then the reaction mixture was stirred at room temperature for 1 h. The reaction was diluted with CH_2Cl_2 (5 mL) and washed with H_2O (5 mL), 1 M aq. HCl (5 mL), H_2O (5 mL), sat. aq. $NaHCO_3$ (5 mL), then dried over anhydrous Na_2SO_4. Filtration followed by concentration *in vacuo*. afforded 1-acetyl-3-methyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**27**) (81 mg, 96%) as a colourless solid: R_f 0.90 (CH_2Cl_2:MeOH = 20:1); m.p. 118–119 °C (EtOAc) [lit.[126] 122–123 °C (IPA)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.22–8.14 (1 H, m, C^7-H), 7.26 (1 H, ddd, *J* 7.9, 7.8, 1.3. C^5-H), 7.15 (1 H, ddd, *J* 7.9, 7.8, 1.3, C^6-H), 7.01–6.88 (1 H, m, C^4-H), 3.39 (3 H, s, C^8-H_3), 2.75 (3 H, s, C^{10}-H_3); LRMS *m/z* (ES^+) 271 ([M+Na]$^+$, 100 %); Analytical HPLC (RP, method A) rt 6.11 mins, purity 96%. The data are in good agreement with the literature.[126]

tert-Butyl 3-methyl-2-oxo-2,3-dihydro-1*H*-benzo[*d*]imidazole-1-carboxylate (28)



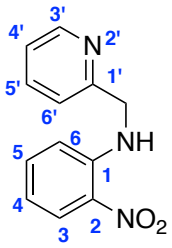
1-Methyl-1,3-dihydro-2*H*-benzo[d]imidazol-2-one (**26**) (50 mg, 0.34 mmol, 1.0 eq.) and DMAP (2.0 mg, 0.020 mmol, 0.050 eq.) were dissolved in THF (1 mL). Et_3N (50 µL, 0.37 eq, 1.1 eq.) was added then the reaction mixture was cooled to 0 °C. Boc_2O (90 µL, 0.37 mmol, 1.1 eq.) was added and the reaction mixture was stirred at room temperature for 3 h. The volatile components were removed *in vacuo* then H_2O (2.0 mL) was added and the mixture was extracted with CH_2Cl_2 (2 × 5 mL). The combined organic components were washed with 1 M aq. HCl (5.0 mL) and H_2O (5.0 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification of the oil by silica gel chromatography, eluting with CH_2Cl_2 and MeOH (1:0 to 9:1), afforded *tert*-butyl 3-methyl-2-oxo-2,3-dihydro-1*H*-benzo[*d*]imidazole-1-carboxylate (**28**) (81 mg, 96%) as a colourless solid: R_f 0.73 (EtOAc); m.p. 62–64 °C (MeOH); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2934 (br, w), 2360 (w) (N-H), 1789 (s) (C=O), 1749 (s) (C=O), 1494 (m), 1327 (s), 1156 (m), 1088 (w), 747 (m); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.85–7.79 (1 H, m, C^7-H), 7.21 (1 H, ddd, *J* 7.8, 7.7, 1.2, C^6-H), 7.12 (1 H, ddd, *J* 7.8, 7.7, 1.2, C^5-H), 6.9 8–6.91 (1 H, m, C^4-H), 3.38 (3 H, s, C^8-H_3), 1.67 (9 H, s, C^{11}-H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 148.9 (C^2), 130.0 (C^9), 126.2 (C^{7a}), 123.9 (C^{3a}), 122.1 (C^7), 114.3 (C^5 or C^6), 107.3 (C^4), 84.6 (C^{10}), 28.1 (C^{11}), 27.0 (C^8); LRMS *m/z* (ES^+) 271 ([M+Na]$^+$, 100 %); HRMS *m/z* (ESI^+) found 271.1051, $C_{13}H_{15}N_2O_3$ requires [M+H]$^+$ 271.1053; Analytical HPLC (RP, method A) rt 9.41 mins, purity > 99%.

N-Benzyl-2-nitroaniline (29)



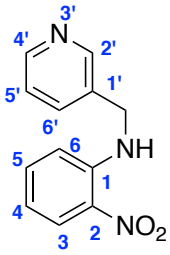
2-Fluoronitroaniline (**24**) (1.00 g, 7.08 mol, 1.00 eq.) was dissolved in NMP (8 mL). Benzyl amine (1.54 mL, 14.2 mol, 2.00 eq.) was added, and the reaction mixture was stirred at 80 °C for 3 h. H_2O (15 mL) was added, and the mixture was extracted with EtOAc (2 × 40 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 6 × 60 mL) then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo* to afford *N*-benzyl-2-nitroaniline (**29**) (1.62 g, > 99%) as an orange solid: R_f 0.28 (petroleum ether/CH_2Cl_2= 3:1); m.p. 70–72 °C (EtOAc) [lit.[127] 72 °C (EtOAc)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.42 (1 H, s, N-H), 8.20 (1 H, dd, *J* 8.5, 1.6, C^3-H), 7.45–7.27 (6 H, m, Ph-H and C^5-H), 6.82 (1 H, dd, *J* 8.7, 1.2, C^6-H), 6.67 (1 H, ddd, *J* 8.5, 6.9, 1.2, C^4-H), 4.56 (2 H, d, *J* 5.7, $C^{1'a}$-H_2); HRMS *m/z* (ESI$^+$) found 229.0975, $C_{13}H_{12}N_2O_2$ requires $[M+H]^+$ 229.0972. The data are in good agreement with the literature.[127,128]

2-Nitro-*N*-(pyridin-2-ylmethyl)aniline (30)



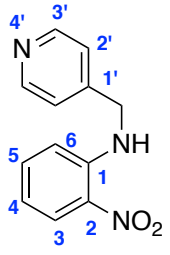
2-Fluoronitroaniline (**24**) (1.00 g, 7.08 mol, 1.00 eq.) was dissolved in NMP (8 mL). 2-(Aminomethyl)pyridine (2.92 mL, 28.32 mol, 4.00 eq.) was added, and the reaction mixture was stirred at 80 °C for 2 h. Sat. aq. $NaHCO_3$ (20 mL) was added, and the mixture was extracted with EtOAc (2 × 50 mL). The combined organic components were washed with LiCl (10% *v/w*, 5 × 60 mL) then dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo* to afford *N*-piridin-2'-yl-2-nitroaniline (**30**) (1.66 g, 98%) as a yellow solid: R_f 0.06 (petroleum ether:EtOAc = 5:1); m.p. 86–88 °C (EtOAc) [lit.[129] 92–93 °C (EtOH)]; ^{1}H NMR (400 MHz, $CDCl_3$) 8.88 (1 H, s, NH), 8.64 (1 H, dt, *J* 4.9, 1.2, $C^{3'}$-H), 8.21 (1 H, dd, *J* 8.5, 1.6, C^3-H), 7.68 (1 H, ddd, *J* 7.5, 1.8, $C^{5'}$-H), 7.40 (1 H, ddd, *J* 8.5, 7.0, 1.4, C^5-H), 7.32 (1 H, d, *J* 7.8, C^6-H), 7.24 (1 H, dd, *J* 7.5, 4.9, $C^{4'}$-H), 6.82 (1 H, dd, *J* 8.7, 1.2, $C^{6'}$-H), 6.68 (1 H, ddd, *J* 8.5, 7.0, 1.4, C^4-H), 4.68 (2 H, d, *J* 5.5, $C^{1'a}$-H_2); Analytical HPLC (RP, method A) rt 9.20 mins, purity > 99%. The compound could not be detected by mass spectrometry. The data are in good agreement with the literature.[129]

2-Nitro-*N*-(pyridin-3-ylmethyl)aniline (31)



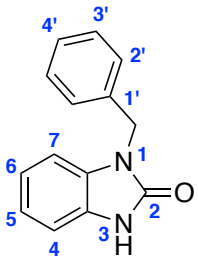
2-Fluoronitroaniline (**24**) (1.0 g, 7.1 mol, 1.0 eq.) was dissolved in NMP (6.0 mL). 3-(Aminomethyl)pyridine (2.9 mL, 28.32 mol, 4.0 eq.) was added, and the reaction mixture was stirred at room temperature for 5 h. H_2O (20 mL) was added, and the mixture was extracted with EtOAc (3 × 150 mL). The combined organic components were washed with LiCl (10% *v/w*, 5 × 60 mL) then dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. The crude mixture (1.6 g) was filtered through silica, washing with EtOAc, then crystallised with minimum amounts of boiling $CHCl_3$ to afford 2-nitro-*N*-(pyridin-3-ylmethyl)aniline (**31**) (1.2 g, 75%) as bright orange crystals: R_f 0.51 (petroleum ether:EtOAc = 3:1); m.p. 80–82 °C ($CHCl_3$); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3331 (w), 1568 (w), 1507 (m), 1233 (m), 1156 (w), 1045 (w), 774 (w), 744 (w); ^{1}H NMR (500 MHz, $CDCl_3$) 8.64 (1 H, dd, *J* 2.1, 0.8, $C^{2'}$-H), 8.61 (1 H, dd, *J* 4.8, 1.6, $C^{4'}$-H), 8.40 (1 H, t, *J* 5.8, NH), 8.20 (1 H, dd, *J* 8.7, 1.6, C^6-H), 7.67 (1 H, dd, *J* 7.9, 2.1, $C^{6'}$-H), 7.40 (1 H, ddd, *J* 8.5, 7.0, 1.6, C^5-H), 7.30 (1 H, ddd, *J* 7.9, 4.8, 0.8, $C^{5'}$-H), 6.78 (1 H, dd, *J* 8.7, 1.2, C^3-H), 6.70 (1 H, ddd, *J* 8.5, 7.0, 1.2, C^4-H), 4.58 (2 H, d, *J* 5.8, $C^{1'a}$-H_2). ^{13}C NMR (126 MHz, $CDCl_3$) δ 149.4 ($C^{4'}$), 149.0 ($C^{2'}$), 144.9 ($C^{1'}$), 136.5 (C^5), 134.8 ($C^{6'}$), 133.1 (C^1), 132.7 (C^2), 127.1 (C^3), 123.9 (C^6), 116.4 (C^4), 114.0 ($C^{5'}$), 44.8 (C^{1a}); LRMS *m/z* (ES^+) 230 ($[M+H]^+$, 100%); HRMS *m/z* (ESI^+) found 230.0925, $C_{12}H_{11}N_2O_2$ requires $[M+H]^+$ 230.0924; Analytical HPLC (RP, method A) rt 8.57 mins, purity > 99%.

2-Nitro-*N*-(pyridin-4-ylmethyl)aniline (32)



2-Fluoronitroaniline (**24**) (1.00 g, 7.08 mol, 1.00 eq.) was dissolved in NMP (6 mL). 4-(Aminomethyl)pyridine (1.45 mL, 14.2 mol, 2.00 eq.) was added, and the reaction mixture was stirred at 80 °C for 1 h. H_2O (80 mL) was added, and the mixture was extracted with EtOAc (2 × 100 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 3 × 80 mL) then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo* to afford 2-nitro-*N*-(pyridin-4-ylmethyl)aniline (**32**) (1.19 g, 74%) as an orange solid: R_f 0.30 (EtOAc); m.p. 66–68°C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3387 (w), 2360 (m), 1615 (m), 1503 (m), 1416 (m), 1354 (m), 1282 (m), 1249 (m), 1037 (m), 984 (m), 861 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.61–8.57 (2 H, m, $C^{3'}$-H), 8.49 (1 H, br t, *J* 6.0, NH), 8.22 (1 H, dd, *J* 8.5, 1.6, C^3-H), 7.37 (1 H, ddd, *J* 8.5, 7.0, 1.6, C^5-H), 7.28–7.24 (2 H, m, $C^{2'}$-H), 6.72 (1 H, ddd, *J* 8.5, 7.0, 1.2, C^4-H), 6.65 (1 H, dd, *J* 8.5, 1.2, C^6-H), 4.60 (2 H, d, *J* 6.0, CH_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 150.5 ($C^{3'}$), 146.9 ($C^{1'}$), 144.9 (C^1), 136.5 (C^5), 132.8 (C^2), 127.1 (C^3), 121.8 ($C^{2'}$), 116.5 (C^4), 114.1 (C^6), 46.1 ($C^{1'a}$ -H_2); LRMS *m/z* (ES^+) 230 ($[M+H]^+$, 100%); HRMS *m/z* (ESI^+) found 230.0925 (100%), $C_{12}H_{11}N_2O_2$ requires $[M+H]^+$, 230.0924; Analytical HPLC (RP, method A) rt 6.05 mins, purity 97%.

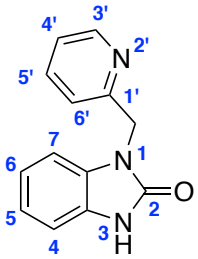
1-Benzyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (33)



A solution of *N*-benzyl-2-nitroaniline (**29**) (500 mg, 2.52 mmol, 1.0 eq.) and 10% Pd/C (134 mg, 0.13 mmol, 0.050 eq.) in EtOH (25.0 mL) was stirred under H_2-atmosphere for 30 mins at room temperature. The reaction mixture was filtered through Celite® and concentrated *in vacuo*. The intermediate was dissolved in dry THF (5 mL), then Et_3N (700 µL, 6.56 mmol, 2.0 eq.) was added. Triphosgene (260 mg, 0.88 mmol, 0.35 eq.) was added portionwise to the reaction mixture at 0 °C, then the reaction mixture was stirred at 0 °C for 1.5 h. Sat. aq. $NaHCO_3$ (5 mL) then 1 M NaOH (15 mL) were added dropwise at 0 °C. The THF was removed *in vacuo*, the precipitate was collected by filtration and washed with H_2O (2 × 30 mL) to afford 1-benzyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**33**) (259 mg, 24%) as an off-white solid: R_f 0.37 (petroleum ether:EtOAc = 5:1); m.p.

190–192 °C (THF) [lit.[130] 198–200 °C (EtOH)]; ^{1}H NMR (400 MHz, DMSO-D_6) δ 10.96 (1 H, s, NH), 7.37–7.23 (5 H, m, Ph-H), 7.04–6.91 (4 H, m, C^4-H to C^7-H), 5.00 (2 H, s, C-H_2); LRMS *m/z* (ES^-) 223 ([M–H]$^-$, 100%).; Analytical HPLC (RP, method A) rt 8.35 mins, purity > 99%. The data are in good agreement with the literature.[130]

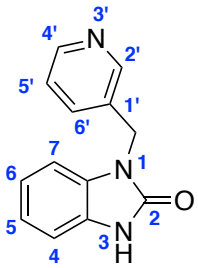
1-(Pyridin-2-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (34)



A solution of 2-nitro-*N*-(pyridin-2-ylmethyl)aniline (**30**) (500 mg, 2.18 mmol, 1.0 eq.) and 10% Pd/C (116 mg, 0.11 mmol, 0.050 eq.) in EtOH (25.0 mL) was stirred under an H_2-atmosphere for 30 mins at room temperature. The reaction mixture was filtered through Celite® and concentrated *in vacuo*. The intermediate was dissolved in dry THF (10 mL), then Et_3N (600 µL, 4.36 mmol, 2.0 eq.) was added. Triphosgene (230 mg, 0.76 mmol, 0.35 eq.) was added portion wise to the reaction mixture at 0 °C, then the reaction mixture was stirred at 0 °C for 1.5 h. Saturated aq. $NaHCO_3$ (5 mL) then 1 M NaOH (15 mL) were added dropwise at 0 °C and stirred for 20 mins. The THF was removed *in vacuo*, the precipitate was collected by filtration and washed with H_2O (2 × 30 mL). The filtrate was concentrated *in vacuo* and purified by silica gel chromatography, eluting with petroleum ether:EtOAc (1:0 to 0:1). The purified solid and the filtered solid were combined and dried *in vacuo* to afford 1-(pyridin-2-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**34**) (372 mg, 76%) as an off-white solid: R_f 0.26 (EtOAc); m.p. 178–179 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3145 (w), 1687 (s) (C=O), 1395 (m), 1150 (w), 748 (m), 686 (m), 641 (m); ^{1}H NMR (500 MHz, $CDCl_3$) 9.12 (1 H, s, N-H), 8.59 (1 H, ddd, *J* 4.9, 1.8, 1.0, $C^{3'}$-H), 7.62 (1 H, ddd, *J* 7.7, 7.7, 1.8, $C^{5'}$-H), 7.24–7.16 (2 H, m, $C^{6'}$-H and $C^{4'}$-H), 7.11–7.08 (1 H, m, C^7-H), 7.06 (1 H, ddd, *J* 7.6, 7.5 1.4, C^5-H or C^6-H), 7.02 (1 H, ddd, *J* 7.6, 7.5, 1.5, C^5-H or C^6-H), 6.95 (1 H, dd, *J* 7.6, 1.2, C^4-H), 5.23 (2 H, s, $C^{1'a}$-H_2); ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.1 ($C^{1'}$), 155.2 (C^2=O), 149.5 ($C^{3'}$), 137.1 ($C^{5'}$), 130.2 (C^{7a}), 127.7 (C^{3a}), 122.7 ($C^{6'}$), 121.9 (C^5 or C^6), 121.62 (C^5 or C^6), 121.55 (C^7), 109.5 ($C^{4'}$), 108.8 (C^4), 46.5 (C^{1a}); LRMS *m/z* (ES^-)

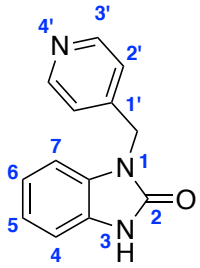
224 ([M–H]$^-$, 100%); HRMS *m/z* (ESI$^-$) found 224.0829, $C_{13}H_{11}N_3O$ requires [M–H]$^-$ 224.08294; Analytical HPLC (RP, method A) rt 6.72 mins, purity 99%.

1-(Pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (35)



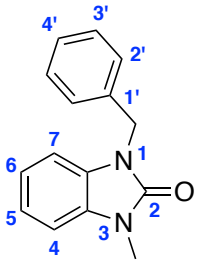
A solution of 2-Nitro-*N*-(pyridin-3-ylmethyl)aniline (**31**) (500 mg, 2.18 mmol, 1.0 eq.) and 10% Pd/C (116 mg, 0.11 mmol, 0.050 eq.) in EtOH (20.0 mL) was stirred under H_2-atmosphere for 40 mins at room temperature. The reaction mixture was filtered through Celite® and concentrated *in vacuo*. The intermediate was dissolved in dry THF (15 mL), then pyridine (450 µL, 4.4 mmol, 2.0 eq.) was added. Triphosgene (230 mg, 0.76 mmol, 0.35 eq.) was added dropwise portion wise to the reaction mixture at 0 °C, then the reaction mixture was stirred at 0 °C for 3 h. Saturated aq. $NaHCO_3$ (2 mL) then 1 M NaOH (4 mL) were added dropwise at 0 °C, then the reaction mixture was warmed to room temperature and stirred for 2 h. The THF was removed *in vacuo*, and the precipitate was removed by filtration. The filtrate was lyophilised and triturated with MeCN Purification by silica gel chromatography, using a 10 g SNAP column on a Biotage purification system, eluting with CH_2Cl_2:MeOH (1:0 to 20:1) afforded 1-(pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**35**) (125 mg, 25%) as an off-white solid: R_f 0.20 (EtOAc); m.p. 140–144 °C (MeOH); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3016 (w), 2970 (w), 1738 (s) (C=O), 1366 (w), 1228 (w), 1216 (w), 869 (w), 639 (w); ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.58 (1 H, s, N-H), 8.64–8.53 (2 H, m, C$^{3'}$-H), 7.25–7.20 (2 H, m, C$^{2'}$-H), 7.18 –7.14 (1 H, m, C^7-H), 7.10 (1 H, ddd, *J* 7.7, 1.4, 1.1, C^4-H), 7.04 (1 H, ddd, *J* 7.7, 7.7, 1.4, C^6-H), 6.81 (1 H, ddd, *J* 7.7, 7.7, 1.1, C^5-H), 5.13 (2 H, s, C$^{1'a}$-H$_2$); ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.6 (C^2), 149.3 (C$^{2'}$ or C$^{4'}$), 148.9 (C$^{2'}$ or C$^{4'}$), 135.3 (C$^{1'}$), 132.0 (C^{7a}), 129.7 (C^{3a}), 128.0 (C$^{6'}$), 123.8 (C^7), 122.1 (C$^{5'}$), 121.6 (C^5), 109.8 (C^6), 108.2 (C^4), 42.1 (C^{1a}); LRMS *m/z* (ES$^+$) 226 ([M+H]$^+$, 100%); HRMS *m/z* (ESI$^+$) found 226.0975, $C_{13}H_{11}N_3O$ requires [M+H]$^+$ 226.0975; Analytical HPLC (RP, method A) rt 6.52 mins, purity 99%.

1-(Pyridin-4-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (36)



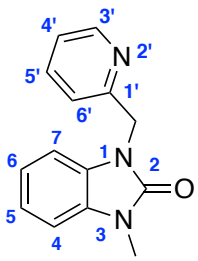
A solution of 2-nitro-*N*-(pyridin-4-ylmethyl)aniline (**32**) (0.5 g, 2.2 mmol, 1.0 eq.) and 10% Pd/C (0.12 mg, 0.11 mmol, 0.050 eq.) in EtOH (25 mL) was stirred under H_2-atmosphere for 40 mins at room temperature. The reaction mixture was filtered through Celite® and concentrated *in vacuo*. The intermediate was dissolved in dry THF (10 mL), then pyridine (0.45 mL, 4.4 mmol, 2.0 eq.) was added. Triphosgene (0.24 g, 0.78 mmol, 0.35 eq.) was added portion wise to the reaction mixture at 0 °C, then the reaction mixture was stirred at 0 °C for 3 h. Saturated aq. $NaHCO_3$ (2 mL) then 1 M NaOH (4 mL) were added dropwise at 0 °C, and the warmed to room temperature and stirred for 2 h. The THF was removed *in vacuo*, and the precipitate was removed by filtration. The filtrate was lyophilised and purified by silica gel chromatography, using a 10 g SNAP column on a Biotage purification system, eluting with CH_2Cl_2:MeOH (1:0 to 20:1). The product was collected and purified further by column chromatography, eluting with petroleum ether/EtOAc (1:0 to 0:1) to afford 1-(pyridin-4-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**36**) (94 mg, 22%) as a colourless solid: R_f 0.22 (EtOAc); m.p. 140–142 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2980 (w), 1690 (C=O), 1484 (m), 1395 (w), 1156 (m), 1004 (m), 750 (m), 727 (m); ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.58 (1 H, s, NH), 8.64–8.53 (2 H, m, $C^{3'}$-H), 7.25–7.20 (2 H, m, $C^{2'}$-H), 7.18–7.12 (1 H, m, C^7-H), 7.10 (1 H, ddd, *J* 7.7, 1.4, 1.1, C^4-H), 7.04 (1 H, ddd, *J* 7.7, 7.7, 1.4, C^6-H), 6.81 (1 H, ddd, *J* 7.7, 7.7, 1.1, C^5-H), 5.13 (2 H, s, $C^{1'a}$-H_2); ^{13}C NMR (101 MHz, $CDCl_3$) δ 155.5 (C^2), 150.3 ($C^{3'}$), 145.1 ($C^{1'}$), 129.8 (C^{3a}), 127.9 (C^{7a}), 122.2 (C^5 or C^6), 122.0 ($C^{2'}$), 121.7 (C^5 or C^6), 109.9 (C^4 or C^7), 108.2 (C^4 or C^7), 43.4 ($C^{1'a}$-H_2); LRMS *m/z* (ES^+) 226 ($[M+H]^+$, 100%); HRMS *m/z* (ESI^-) found 224.08284, $C_{13}H_{11}N_3O$ requires $[M–H]^-$ 224.08294; Analytical HPLC (RP, method A) rt 6.51 mins, purity > 99%;

1-Benzyl-3-methyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (37)



MeI (140 µL, 2.22 mmol, 10.0 eq.) was added dropwise to a solution of 1-benzyl-3-methyl-1,3-dihydro-2*H*-benzo[d]imidazol-2-one (**33**) (50 mg, 0.22 mmol, 1.0 eq.) and K_2CO_3 (125 mg, 0.89 mol, 4.0 eq.) in DMF (1.0 mL). The reaction mixture was stirred at room temperature for 16 h. H_2O (30 mL) was added, and the mixture was extracted with EtOAc (2 × 50 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 2 × 150 mL), then dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo* to afford 1-benzyl-3-methyl-1,3-dihydro-2*H*-benzo[d]imidazol-2-one (**37**) (52 mg, 98%) as colourless amorphous solid: R_f 0.35 (petroleum ether:EtOAc = 5:1); m.p. 72–74°C (CH_2CL_2); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2931 (w) (CH), 1702 (s) (C=O), 1497 (m), 1401 (w), 1250 (w), 1020 (w), 798 (m), 698 (m); ^{1}H NMR (700 MHz, DMSO-D_6) 7.31 (4 H, m, Ph-H), 7.25 (1 H, m, $C^{4'}$-H), 7.16 (1 H, dd, *J* 7.7, 1.2, C^7-H), 7.09 (1 H, dd, *J* 7.7, 1.2, C^4-H), 7.05 (1 H, ddd, *J* 7.7, 7.6, 1.2, C^6-H), 7.00 (1 H, ddd, *J* 7.7, 7.6, 1.2, C^5-H), 5.04 (2 H, s, C^8-H_2), 3.36 (3 H, s, N-CH_3); ^{13}C NMR (176 MHz, DMSO-D_6) δ 154.2 (C^2), 137.5 ($C^{1'}$), 130.2 (C^{7a}), 129.11 (C^{3a}), 129.08 ($C^{3'}$), 127.9 ($C^{4'}$), 127.8 ($C^{2'}$), 121.6 (C^6), 121.4 (C^5), 108.6 (C^7), 108.4 (C^4), 44.14 (C^8-H_2), 27.5 (CH_3); LRMS *m/z* (ES^+) 239 ($[M+H]^+$, 60%), 261 ($[M+Na]^+$, 100%); HRMS *m/z* (ESI^+) found 239.11795, $C_{15}H_{14}N_2O$ requires ($[M+H]^+$, 239.11789; Analytical HPLC (RP, method A) rt 9.23 mins, purity 99%.

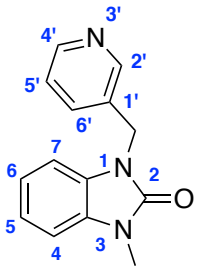
1-Methyl-3-(pyridin-2-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (38)



MeI (140 µL, 2.22 mmol, 10 eq.) was added dropwise to a solution of 1-(pyridin-2-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**34**) (50.0 mg, 0.221 mmol, 1.0 eq.)

and K_2CO_3 (123 mg, 0.81 mol, 4.0 eq.) in DMF (1.0 mL). The reaction mixture was stirred at room temperature for 16 h. H_2O (5 mL) was added, and the mixture was extracted with EtOAc (5 × 50 mL). The combined organic components were washed with LiCl (10% *w/v*, 2 × 50 mL), then dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. Purification by column chromatography, eluting with petroleum ether:EtOAc (1:1 to 0:1) afforded 1-methyl-3-(pyridin-2-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**38**) (14 mg, 28%) as an amorphous brown residue: R_f 0.52 (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3478 (w), 2957 (w), 1684 (s) (C=O), 1508 (w), 1448 (m), 1220 (w), 1118 (w), 969 (w);^{1}H NMR (400 MHz, $CDCl_3$) δ 8.57 (1 H, ddd, *J* 4.8, 1.9, 1.0, C$^{3'}$-H), 7.60 (1 H, dd, *J* 7.7, 7.7, 1.8, C$^{5'}$-H), 7.22–7.15 (2 H, m, C$^{6'}$-H and C$^{4'}$-H), 7.09 (1 H, ddd, *J* 7.6, 7.5, 1.3, C^6-H), 7.05–6.93 (3 H, m, C^4-H, C^5-H, C^7-H), 5.21 (2 H, s, C$^{1'a}$-H$_2$), 3.48 (3 H, s, N-CH$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ 156.3 (C$^{1'}$), 155.5 (C^2=O), 149.4 (C$^{3'}$), 136.9 (C$^{5'}$), 130.1 (C^{7a}), 129.2 (C^{3a}), 122.6 (C$^{6'}$), 121.7 (C^5 or C^6), 121.5 (C^5 or C^6), 121.4 (C^7), 108.5 (C$^{4'}$), 107.4 (C^4), 46.9 (C^{1a}), 27.3 (N-CH$_3$); LRMS *m/z* (ES$^+$) 240 ([M+H]$^+$, 100%); HRMS *m/z* (ESI$^+$) found 240.11304, $C_{14}H_{13}N_3O$ requires [M+H]$^+$ 240.11314; Analytical HPLC (RP, method A) rt 7.41, purity 98%.

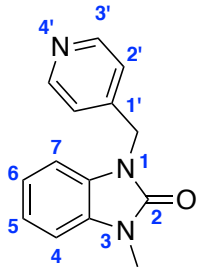
1-Methyl-3-(pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (39)



MeI (20 µL, 0.26 mmol, 3.0 eq.) was added dropwise to a solution of 1-(pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**35**) (20 mg, 0.087 mmol, 1.0 eq.) and K_2CO_3 (72 mg, 0.56 mol, 6.0 eq.) in DMF (1.0 mL). The reaction mixture was stirred at room temperature for 20 h. H_2O (10 mL) was added, and the mixture was extracted with EtOAc (3 × 20 mL). The combined organic components were washed with LiCl (10% *w/v*, 5 × 50 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, elution with CH_2Cl_2:MeOH (slow gradient from 1:0 to 20:1) afforded 1-methyl-3-(pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**39**) (6 mg, 28%) as colourless solid: R_f 0.52 (CH_2Cl_2:MeOH = 20:1); m.p. 104–105°C (MeOH); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3467 (w), 2980 (w), 1701 (s) (C=O), 1498 (m), 1401 (w), 1289

(w), 1127 (w), 751 (m), 633 (m); ^{1}H NMR (500 MHz, $CDCl_3$) 8.65 (1 H, s, $C^{2'}$-H), 8.56–8.50 (1 H, m, $C^{4'}$-H), 7.69–7.62 (1 H, m, $C^{6'}$-H), 7.24 (1 H, d, *J* 7.7, $C^{5'}$-H), 7.10 (1 H, ddd, *J* 7.7, 7.7, 1.2, C^{6}-H), 7.04 (1 H, ddd, *J* 7.7, 7.7, 1.2, C^{5}-H), 7.00 (1 H, dd, *J* 7.7, 1.2, C^{7}-H), 6.92–6.86 (1 H, m, C^{4}-H), 5.10 (2 H, s, $C^{1'a}$-H_2), 3.47 (3 H, s, N-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.5 (C^{2}), 149.3 ($C^{2'}$ or $C^{4'}$), 149.0 ($C^{2'}$ or $C^{4'}$), 135.4 ($C^{6'}$), 132.1 ($C^{1'}$), 130.1 (C^{7a}), 128.7 (C^{3a}), 123.7 ($C^{5'}$), 121.7 (C^{5} or C^{6}), 121.5 (C^{5} or C^{6}), 107.9 (C^{4} or C^{7}), 107.7 (C^{4} or C^{7}), 42.4 ($C^{1'a}$), 27.3 (N-CH_3); LRMS *m/z* (ES^+) 240 ($[M+H]^+$, 100%); HRMS *m/z* (ESI^+) found 240.11316, $C_{14}H_{13}N_3O$ requires $[M+H]^+$ 240.11314; Analytical HPLC (RP, method A) rt 7.16 mins, purity > 99%.

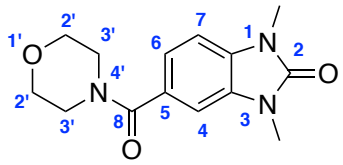
1-Methyl-3-(pyridin-4'-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (40)



MeI (20 µL, 0.26 mmol, 3.0 eq.) was added dropwise to a solution of 1-(pyridin-4-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**36**) (20 mg, 0.087 mmol, 1.0 eq.) and K_2CO_3 (72 mg, 0.56 mol, 6.0 eq.). The reaction mixture was stirred at room temperature for 20 h. H_2O (10 mL) was added, and the mixture was extracted with EtOAc (3 × 20 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 5 × 50 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with CH_2Cl_2:MeOH (1:0 to 20:1) afforded 1-methyl-3-(pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**40**) (3 mg, 11%) as colourless oil: R_f 0.58 (CH_2Cl_2:MeOH = 20:1); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3015 (w), 2970 (w), 1738 (s) (C=O), 1435 (w), 1366 (m), 1228 (m), 1216 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.62–8.48 (2 H, m, $C^{3'}$-H), 7.20 (2 H, d, *J* 5.0, $C^{2'}$-H), 7.13 (1 H, dd, *J* 7.6, 1.3, C^{7}-H), 7.04 (2 H, 2 × t, *J* 7.6, C^{5}-H and C^{6}-H), 6.80 (1 H, dd, *J* 7.6, 1.3, C^{4}-H), 5.09 (2 H, s, CH_2), 3.49 (3 H, s, N-CH_3). ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.5 (C^{2}), 150.1 ($C^{3'}$), 145.5 ($C^{1'}$), 130.1 (C^{3a}), 128.8 (C^{7a}), 122.1 ($C^{2'}$), 121.9 (C^{5} or C^{6}), 121.5 (C^{5} or C^{6}), 107.9 (C^{4} or C^{7}), 107.8 (C^{4} or C^{7}), 43.8 ($C^{1'a}$-H_2), 27.4 (CH_3). LRMS *m/z* (ES^+) 240 ($[M+H]^+$, 100%); HRMS *m/z* (ESI^+) found 240.11327,

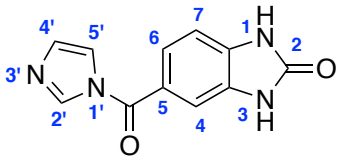
$C_{14}H_{13}N_3O$ requires $[M+H]^+$ 240.11314; Analytical HPLC (RP, method A) rt 6.22 mins, purity 94%.

1,3-Dimethyl-5-(morpholine-4-carbonyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (41)



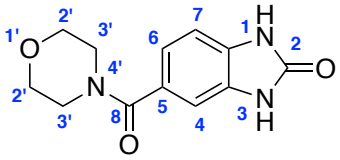
5-(Morpholine-4-carbonyl)-1,3-dihydro-2*H*-benzo[d]imidazol-2-one (**43**) (15 mg, 0.06 mmol, 1.0 eq.) was dissolved in DMF (1 mL). K_2CO_3 (33 mg, 0.24 mmol, 4.0 eq.) and MeI (20 µL, 0.32 mmol, 5.3 eq.) were added and the reaction mixture was stirred at 60 °C for 15 h. MeI (20 µL, 0.32 mmol, 5.3 eq.) was added and the reaction mixture was stirred at 60 °C for a further 4 h. H_2O (10 mL) was added and the mixture was extracted with EtOAc (40 mL). The organic components were washed with 0.5 M aq. LiCl (6 × 15 mL) then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification of the oil by silica gel chromatography, eluting with CH_2Cl_2:MeOH (1:0 to 20:1) afforded the product and traces of grease. The product was dissolved in acetonitrile and washed with hexane (3 × 20 mL), then dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo* to afford 1,3-dimethyl-5-(morpholine-4-carbonyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**41**) (5.9 mg, 36%): R_f 0.66 (CH_2Cl_2/MeOH = 20:1); m.p. 169–170 °C (MeCN); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2921 (w), 2361 (w), 1710 (s) (C=O), 1626 (s) (C=O), 1509 (m), 1477 (m), 1300 (w), 1071 (m), 823 (m), 763 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.14 (1 H, dd, *J* 7.9, 1.2, C^6-H), 7.11 (1 H, d, *J* 1.2, C^4-H), 6.96 (1 H, d, *J* 7.9, C^7-H), 3.90–3.48 (8 H, br m, $C^{2'}$-H_2 and $C^{3'}$-H_2), 3.38 (3 H, s, NCH_3), 3.37 (3 H, s, NCH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.6 (C^{5a}), 154.7 (C^2), 131.3 (C^{3a}), 130.1 (C^{7a}), 128.1 (C^5), 120.6 (C^6), 107.1 (C^4), 106.7 (C^7), 66.9 ($C^{3'}$), 48.6 ($C^{2'}$), 43.0 ($C^{2'}$), 27.41 (NCH_3), 27.37 (NCH_3); LRMS *m/z* 276 (ES^+) ($[M+H]^+$, 50%); HRMS *m/z* (ESI^+) found 276.1343, $C_{14}H_{18}N_3O_3$ requires $[M+H]^+$ 276.1343; Analytical HPLC (RP, method A) rt 5.68 mins, purity 94%.

5'-(1*H*-Imidazole-1'-carbonyl)-benzimidazol-2-one (43a)



3,4-Diaminobenzoic acid (1.0 g, 6.4 mmol, 1.0 eq.) and *N,N'*-carbonyldiimidazole (2.2 g, 12.8 mmol, 2.1 eq.) were dissolved in dry CH_2Cl_2 (8 mL) and stirred at room temperature for 3 h. EtOAc (10 mL) was added and the reaction mixture was stirred for 30 mins, then was cooled to 0 °C. The precipitate was collected by filtration, washed with EtOAc (2 × 20 mL) and H_2O (5 × 20 mL), then dried *in vacuo* to afford 5'-(1*H*-imidazole-1'-carbonyl)-benzimidazol-2-one (**43a**) (398 mg, 25%) as a light brown solid. The intermediate was used in following reactions without further purification: R_f 0.00 (EtOAc or EtOH); m.p. > 200 °C (decomposition) (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3135 (w), 2982 (w), 2361 (w), 1725 (s) (C=O), 1698 (s) (C=O), 1473 (m), 1371 (s), 1247 (m), 1020 (w), 875 (w), 710 (w); ^{1}H NMR (400 MHz, DMSO-D_6) δ 11.21 (1 H, s, N-H), 11.00 (1 H, s, N-H), 8.22 (1 H, s, $C^{2'}$-H), 7.69 (1 H, d, J 1.5, $C^{5'}$-H), 7.48 (1 H, dd, J 8.2, 1.6, C^6-H), 7.35 (1 H, d, J 1.6, C^4-H), 7.15 (1 H, d, J 1.5, $C^{4'}$-H), 7.12 (1 H, d, J 8.2, C^7-H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.39 (C^{5a}), 155.8 (C^2), 138.94 ($C^{2'}$), 135.6 (C^{3a}), 135.1 (C^{7a}), 130.5 ($C^{4'}$), 124.8 (C^6), 123.8 (C^5), 119.3 ($C^{5'}$), 110.6 (C^4), 108.8 (C^7); LRMS *m/z* (ES^-) 227 ($[M-H]^-$, 100%); HRMS *m/z* (ES^-) found 227.0575, $C_{11}H_8N_4O_2$ requires $[M-H]^-$ 227.0574; Isolated intermediate unstable on RP HPLC.

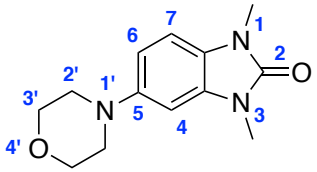
5-(Morpholine-4-carbonyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (43)



3,4-Diaminobenzoic acid (500 mg, 3.28 mmol, 1.00 eq.) and *N,N'*-carbonyldiimidazole (1.11 g, 6.90 mmol, 2.10 eq.) were dissolved in dry THF (4 mL) and stirred at room temperature for 3 h. Morpholine (325 µL, 3.71 mmol, 1.13 eq.) was added to the thick slurry and the reaction mixture was stirred at room temperature for 1 h. The clear solution precipitated upon addition of EtOAc (10 mL), H_2O (20 mL), and citric acid (0.5 M, 50 mL). CH_2Cl_2 (300 mL) and $CHCl_3$ (100 mL) were added and the brown solid was filtered to yield 5-(morpholine-4-carbonyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**43**) (192 mg,

24%) as a brown solid: R_f 0.37 (EtOH); m.p. > 300 °C ($CHCl_3$); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3091 (w), 2860 (w), 1690 (s) (C=O), 1623 (m) (C=O), 1583 (s), 1431 (m), 1232 (w) , 1070 (w); ^{1}H NMR (400 MHz, DMSO-D_6) δ 10.81 (1 H, s, N-H), 10.74 (1 H, s, N-H), 7.01 (1 H, dd, *J* 8.0, 1.5, C^6-H), 6.99–6.82 (2 H, m, C^4-H and C^7-H), 3.58 (4 H, br s, $C^{2'}$-H_2), 3.49 (4 H, br s, $C^{3'}$-H_2); ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.9 (C^8), 155.8 (C^2), 131.4 (C^{3a}), 129.9 (C^5), 128.0 (C^{7a}), 120.6 (C^6), 108.4 (C^4 or C^7), 108.2 (C^4 or C^7), 66.6 ($C^{2'}$ and $C^{3'}$); LRMS *m/z* (ES^-) 246 ([M–H]$^-$, 100%); HRMS *m/z* (ESI^-) found 246.0882, $C_{12}H_{12}N_3O_3$ requires [M–H]$^-$ 246.0873; Analytical HPLC (RP, method A) rt 4.80 mins, purity > 99%.

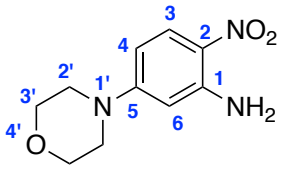
1,3-Dimethyl-5-morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (44)



To a solution of 5-morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**47**) (100 mg, 0.45 mmol, 1.0 eq.) in DMF (2 mL), NaH (60% in mineral oil) (38 mg, 0.88 mmol, 2.2 eq.) was added at 0 °C. The cloudy solution was stirred for 40 mins at 0 °C, followed by addition of MeI (55 µL, 0.88 mmol, 2.2 eq.). The reaction mixture was stirred for a further 1.5 h until TLC analysis show full consumption of the starting material. Brine (20 mL) and sat. $NaHCO_3$ (10 mL) were added, and the mixture was extracted with EtOAc (3 × 60 mL). The combined organic components were washed with sat. $NaHCO_3$ (20 mL) and aq. LiCl (10% *w/v*, 3 × 60 mL), then dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with CH_2Cl_2 and MeOH (slow gradient from 1:0 to 98:2), afforded 1,3-dimethyl-5-morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**44**) (82 mg, 72%) as a colourless solid: R_f 0.61 (CH_2Cl_2:MeOH = 20:1); m.p. 120–122 °C (MeOH); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2961 (w), 2857 (w), 1683 (s) (C=O), 1509 (m), 1326 (m), 1147 (m), 1068 (w), 969 (w), 789 (w), 626 (w); ^{1}H NMR (400 MHz, DMSO-D_6) 6.98 (1 H, d, *J* 8.5, C^6-H), 6.81 (1 H, d, *J* 2.3, C^4-H), 6.66 (1 H, dd, *J* 8.5, 2.3, C^7-H), 3.77–3.68 (4 H, m, $C^{3'}$-H_2), 3.29 (3 H, s, N-CH_3), 3.27 (3 H, s, N-CH_3), 3.09–3.00 (4 H, m, $C^{2'}$-H_2); ^{13}C NMR (101 MHz, DMSO-D_6) δ 154.1 (C^2), 147.0 (C^5), 130.4 (C^{3a}), 123.5 (C^{7a}), 108.8 (C^6), 107.7 (C^7), 97.0 (C^4), 66.2 ($C^{3'}$), 50.2 ($C^{2'}$), 26.9 (N-CH_3), 26.8 (N-CH_3); LRMS

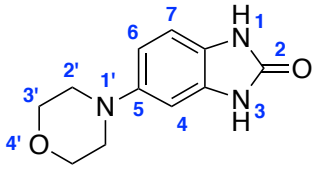
m/z (ES^+) 248 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) 248.13935 found, $C_{13}H_{17}N_3O_2$ requires $[M+H]^+$ 248.13935; Analytical HPLC (RP, method A) rt 6.39 mins, purity > 99%.

5-Morpholino-2-nitroaniline (46)



2-Amino-4-fluoronitrobenzene (**45**) (507 mg, 2.98 mmol, 1.0 eq.) was dissolved in dry NMP (4 mL). Morpholine (400 µL, 4.57 mmol, 1.58 eq.) was added, and the reaction mixture was stirred at 100 °C for 3 h. H_2O (20 mL) and Et_2O (50 mL) were added. The precipitate was collected by filtration, and dried under vacuum. The filtrate was extracted with EtOAc (100 mL). The organic components were washed with aq. LiCl (10% *w/v*, 5 × 60 mL), dried over $MgSO_4$, and concentrated *in vacuo* then crystallised from EtOAc/Et_2O 1:1. The combined solids were dried under vacuum to afford 5-morpholino-2-nitroaniline (**46**) as an off-white solid (530 mg, 82%): R_f 0.16 (petroleum ether:EtOAc = 2:1); m.p. 183–184 °C (EtOAc:Et_2O = 1:1) [lit.[131] 184–186 °C (MeOH)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04 (1 H, d, *J* 9.7, C^3-H), 6.27 (1 H, dd, *J* 9.7, 2.6, C^4-H), 6.14 (2 H, s, N-H_2), 5.95 (1 H, d, *J* 2.6, C^6-H), 3.88–3.76 (4 H, m, $C^{3'}$-H_2), 3.34–3.26 (4 H, m, $C^{2'}$-H_2). LRMS *m/z* (ES^+) 246 ($[M+Na]^+$, 30%). The data are in good agreement with the literature.[131]

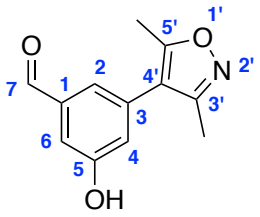
5-Morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (47)



A solution of 5-morpholino-2-nitroaniline (**46**) (0.50 g, 2.2 mmol, 1.0 eq.) and 10% Pd/C (0.13 g, 0.12 mmol, 0.050 eq.) in EtOH (50 mL) was stirred under H_2-atmosphere for 2 h at room temperature. The reaction mixture was filtered through Celite® and concentrated *in vacuo* to afford 4-morpholinobenzene-1,2-diamine. The intermediate was dissolved in dry 1,4-dioxane (10 mL), then Et_3N (0.600 mL, 4.6 mmol, 2.0 eq.) was added. Triphosgene (0.27 mg, 0.89 mmol, 0.40 eq.) was added to the solution at 0 °C, then the reaction mixture was stirred at 0 °C for 2 h. Saturated aq. $NaHCO_3$ (5 mL) then

1 M NaOH (15 mL) were added dropwise at 0 °C, then the mixture was left to warm to room temperature with vigorous stirring for 15 mins. The 1,4-dioxane was removed *in vacuo*, and the precipitate was collected by filtration. Crystallisation from the minimum amount of EtOH:H_2O 1:1 yielded 5-morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (**47**) (380 mg, 69%) as a brown solid: R_f 0.75 (EtOH); m.p.: 268–272 °C (decomposition) (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2957 (br) (NH), 2846, 1709 (s) (C=O), 1511, 1337 (w), 1263, 1127, 858, 743; ^{1}H NMR (500 MHz, $CDCl_3$) 10.42 (1 H, br s, N-H), δ 10.29 (1 H, br s, N-H), 6.68 (1 H, d, J 8.4, C^7-H), 6.55 (1 H, dd, J 8.4, 2.3, C^6-H), 6.52 (1 H, d, J 2.3, C^4-H), 3.74–3.70 (4 H, m, $C^{3'}$-H_2), 2.98–2.94 (4 H, m, $C^{2'}$-H_2); ^{13}C NMR (101 MHz, $CDCl_3$) δ 156.1 (C^2), 146.8 (C^5), 130.9 (C^{3a}), 123.9 (C^{7a}), 109.5 (C^7), 109.1 (C^6), 98.0 (C^4), 66.7 ($C^{3'}$), 50.8 ($C^{2'}$); LRMS *m/z* (ES^-) 218 ([M–H]$^-$, 100%); HRMS *m/z* (ES^+) found 220.1079, $C_{11}H_8N_4O_2$ requires [M+H]$^+$ 220.1080; Analytical HPLC (RP, method A) rt 5.18 mins, purity 98%.

3-(3',5'-Dimethyl-1',2'-oxazol-4'-yl)-5-hydroxybenzaldehyde (48)



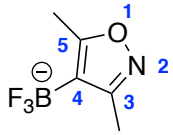
3-Bromo-5-hydroxybenzoic acid (**49**) (3.0 g, 14 mmol, 1.0 eq.) was dissolved in dry THF (120 mL) and cooled to 0 °C. Using a dropping funnel, BH_3 (1 M in THF, 40 mL, 40 mmol, 3.0 eq.) was added over 45 mins. The reaction mixture was then warmed to room temperature and stirred for 20 h. TLC analysis indicated complete consumption of starting material. The reaction mixture was quenched by the addition of conc. HCl solution (6 mL) at 0°C and the volatile components were removed *in vacuo*. The resulting mixture was diluted with H_2O (90 mL) and extracted with EtOAc (3 × 80 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (3 × 200 mL), H_2O (100 mL), and brine (200 mL) then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether : EtOAc (3:1 to 1:1), afforded 3-bromo-5-hydroxybenzyl alcohol (**50**) (2.31 g, 82%) as a colourless solid: R_f 0.37 (petroleum ether:EtOAc = 1:1); m.p. 194–196 °C ($CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.08 (1 H, s, C^4-H), 6.93 (1 H, s, C^2-H), 6.79 (1 H, s, C^6-H), 4.03 (2

H, s, C^7-H_2); LRMS *m/z* (ES^-) 201 ($[M^{79}Br–H]^-$, 100%), 203 ($[M^{81}Br–H]^-$, 80%). The data are in good agreement with the literature.[37]

3-Bromo-5-hydroxybenzyl alcohol (**50**) (2.20 g, 10.8 mmol, 1.0 eq.) was dissolved in CH_2Cl_2 (80 mL) and acetone (2 mL). MnO_2 (6.38 g, 73.4 mmol, 6.8 eq.) was added and the reaction mixture was heated under reflux for 2 d. The reaction mixture was filtered through Celite®. The solvents were removed *in vacuo* to give 3-bromo-5-hydroxybenzaldehyde (**51**) (1.81 g, 84%) as a colourless solid: R_f 0.76 (petroleum ether:EtOAc = 1:1); m.p. 137–139 °C ($CHCl_3$) [lit.[37] 137–139 °C ($CHCl_3$)]; ^{1}H NMR (400 MHz, acetone-D_6) δ 9.93 (1 H, s, C^7(O)-H), 9.31 (1 H, s, C^5-OH), 7.56 (1 H, dd, *J* 1.5, 1.5, C^4-H), 7.38–7.30 (2 H, m, C^2-H and C^6-H); LRMS *m/z* (ES^-) 199 ($[M^{79}Br–H]^-$, 100%), 201 ($[M^{81}Br–H]^-$, 80%). The data are in good agreement with the literature.[37]

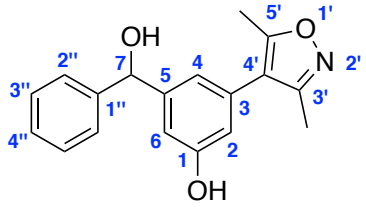
To a solution of 3-bromo-5-hydroxybenzaldehyde (**51**) (700 mg, 3.48 mmol, 1.0 eq.), potassium (3,5-dimethylisoxazol-4-yl)trifluoroborate (**53**) (778 mg, 3.83 mmol, 1.1 eq.), RuPhos (196 mg, 0.42 mmol, 0.12 eq.) and Na_2CO_3 (737 mg, 6.96 mmol, 2.0 eq.) in degassed EtOH (35 mL) was added $Pd(OAc)_2$ (47 mg, 0.21 mmol, 0.06 eq.). The reaction mixture was heated under reflux for 3.5 h, then was filtered through Celite®, and concentrated *in vacuo*. The dark brown residue was suspended in CH_2Cl_2 and sonicated. The solid was collected by filtration. The filtrate was collected, concentrated *in vacuo* and then purified by silica gel chromatography, eluting with petroleum ether and EtOAc (gradient from 5:1 to 3:1). The solids were combined to afford 3-(3',5'-dimethyl-1',2'-oxazol-4'-yl)-5-hydroxybenzaldehyde (**48**) (640 mg, 85%) as a yellow solid: R_f 0.33 (petroleum ether:EtOAc = 3:2); m.p. 192–194 °C (CH_2Cl_2) [lit.[37] 184–186 °C (EtOAc)]; ^{1}H NMR (400 MHz, DMSO-D_6) δ 10.25 (1 H, s, C^7(O)-H), 10.03 (1 H, s, C^5-OH), 7.44 (1 H, dd, *J* 1.5, 1.5, C^2-H), 7.32 (1 H, dd, *J* 2.1, 1.5, C^4-H), 7.15 (1 H, dd, *J* 2.1, 2.1, C^6-H), 2.49 (3 H, s, $C^{5'}$-CH_3), 2.31 (3 H, s, $C^{3'}$-CH_3); LRMS *m/z* (ES^-) 216 ($[M–H]^-$, 100%). The data are in good agreement with the literature.[37]

Potassium (3,5-dimethylisoxazol-4-yl)trifluoroborate (53)



KHF_2 (4.2 g, 53 mmol, 3.0 eq.) was added to a suspension of (3,5-dimethylisoxazol-4-yl)boronic acid (**52**) (2.5 g, 18 mmol, 1.0 eq.) in MeOH (5 mL) at 0°C. H_2O (12 mL) was then added dropwise. The reaction mixture was then warmed to room temperature and stirred for 10 mins. The solvents were removed *in vacuo*. The crude solid was purified by Soxhlet extraction for 18 h with acetone (200 mL). The extraction solvent was concentrated *in vacuo*, and the residue was dissolved in a minimal amount of boiling acetone (250 mL). Precipitation was achieved by addition of cold Et_2O (200 mL). The colourless solid was collected by filtration. The precipitation protocol was repeated for the filtrate concentrated *in vacuo* (using 100 mL of cold Et_2O). The combined solids were dried under vacuum to give the potassium (3,5-dimethylisoxazol-4-yl)trifluoroborate (**53**) (2.9 g, 81%) as a powdery colourless solid: m.p. 342–345 °C (Et_2O) [lit.[87] >200 °C]; ^{1}H NMR (400 MHz, DMSO-D_6) δ 2.05 (3 H, s, C^{5}-CH_3), 2.15 (3 H, s, C^{3}-CH_3); ^{11}B NMR (160 MHz, DMSO-D_6) δ 2.32 (1 B, q, J_{BF} 48.1); ^{19}F NMR (470 MHz, DMSO-D_6) δ –134.51 (3 F, 1:1:1:1 q, J_{BF} 47.0); LRMS: *m/z* (ES^+) 204 ([M+H]$^+$, 100%). The data are in good agreement with the literature.[87]

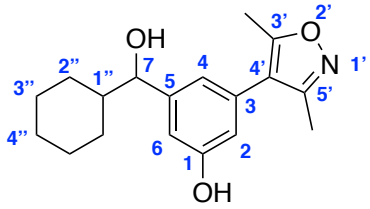
3-(3,5-Dimethylisoxazol-4-yl)-5-[hydroxyl(phenyl)methyl]phenol (7)



Following the procedure of Hewings *et al.*,[37] PhMgBr (1 M in THF, 0.69 mL, 0.69 mmol, 3.0 eq.) was added dropwise to a solution of 3-(3',5'-dimethyl-1',2'-oxazol-4'-yl)-5-hydroxybenzaldehyde (**48**) (100 mg, 0.46 mmol, 1.0 eq.) in distilled THF at 0 °C. The reaction mixture was quenched by addition of 1 M aq. HCl (1.5 mL) at 0 °C. The volatile components were removed *in vacuo* then sat. aq. NH_4Cl (0.5 mL) was added. The residue was diluted with H_2O (25 mL) and extracted Et_2O (3 × 15 mL), the combined organic components were washed with brine (3 × 50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification of the oil by silica gel chromatography, eluting with

petroleum ether and EtOAc (gradient from 1:0 to 1:1), afforded 3-(3,5-dimethylisoxazol-4-yl)-5-[hydroxyl(phenyl)methyl]phenol (**7**) as an off-white crystalline solid (87 mg, 64%): R_f 0.45 (petroleum ether/EtOAc = 1:1); m.p. 186–187 °C (MeCN) [lit.[37] 187–188 °C ($CHCl_3$)]; ^{1}H NMR (400 MHz, acetone-D_6) δ 8.31 (1 H, s, C^1-OH), 7.32 (2 H, d, *J* 7.5, $C^{2''}$-H), 7.18 (2 H, dd, *J* 7.5, 7.5, $C^{3''}$-H), 7.08 (1 H, dd, *J* 7.5, 7.5, $C^{4''}$-H), 6.82–6.74 (2 H, m, C^4-H and C^2-H), 6.57–6.51 (1 H, m, C^6-H), 5.68 (1 H, d, *J* 4.0, C^7-H), 4.76 (1 H, d, *J* 4.0, C^7-OH), 2.23 (3 H, s, $C^{5'}$-CH_3), 2.05 (3 H, s, $C^{3'}$-CH_3); LRMS *m/z* (ES^+) 296 ($[M+H]^+$, 100%). The data are in good agreement with literature.[37]

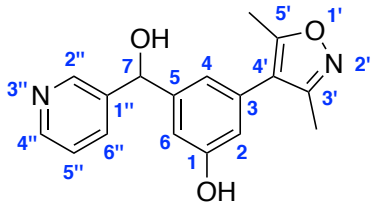
3-(3,5-Dimethylisoxazol-4-yl)-5-(hydroxy(cyclohexyl)methyl)phenol (54)



Following the procedure of Jennings *et al.*,[58] cyclohexanebromide (0.39 mL, 3.1 mmol, 6.8 eq.) in dry THF (1.0 mL) was added dropwise to a flask containing magnesium turnings (69 mg, 2.8 mmol, 6.1 eq.) and a crystal of iodine in dry THF (1.0 mL). Following initiation, the reaction solution was stirred for 1 h, then 3-(3,5-dimethylisoxazol-4-yl)-5-hydroxybenzaldehyde (**48**) (0.10 g, 0.46 mmol, 1.0 eq.) in dry THF (2.0 mL) was added dropwise. The reaction solution was stirred at rt for 2 h then the reaction was quenched with sat. aq. NH_4Cl (15 mL), and the aqueous phase was extracted with EtOAc (3 × 50 mL). The combined organic components were washed with brine (50 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with EtOAc in petroleum ether (1:0 to 1:4), afforded 3-(cyclohexyl(hydroxy)methyl)-5-(3,5-dimethylisoxazol-4-yl)phenol (**54**) (103 mg, 74%) as a colourless solidified foam: R_f 0.51 (petroleum ether:EtOAc = 1:1); m.p. 84–86 °C ($CHCl_3$) [lit.[58] 84–86 °C ($CHCl_3$)]; ^{1}H NMR (500 MHz,$CDCl_3$) δ 6.81 (1 H, dd, *J* 2.5, 1.5, C^4-H), 6.73 (1 H, dd, *J* 1.5, 1.5, C^2-H), 6.64 (1 H, dd, *J* 2.5, 1.5, C^6-H), 5.27 (1 H, s, OH), 4.37 (1 H, dd, *J* 7.0, 2.7, C^7-H), 2.41 (3 H, s, $C^{5'}$-CH_3), 2.27 (3 H, s, $C^{3'}$-CH_3), 1.99–1.93 (1 H, m, 1.6, $C^{1''}$-H), 1.91 (1 H, br d, *J* 2.7, OH), 1.81–1.74 (1 H, m, $C^{2''}$-H_a), 1.73–1.61 (3 H, m, $C^{2''}$-H_a, $C^{2''}$-H_a), 1.49–1.41 (1 H, m, $C^{4''}$-H_a), 1.30–0.90 (5 H, m, $C^{2''}$-H_b, $C^{2''}$-H_b, $C^{4''}$-H_b); LRMS *m/z* (ES^-) 300 ($[M–H]^-$, 100%); HRMS *m/z* (ES^+) found 302.17507 (100%), 303.17851 (30%); $C_{18}H_{23}NO_3$

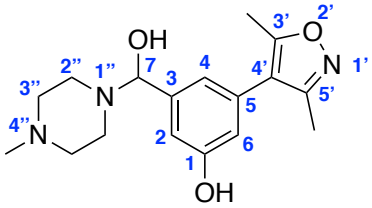
requires $[M+H]^+$ 302.17507; Analytical HPLC (RP, method A) rt 9.11 mins, purity 99%. The data are in good agreement with literature.[58]

3-(3,5-Dimethylisoxazol-4-yl)-5-(hydroxy(pyridin-3-yl)methyl)phenol (55)



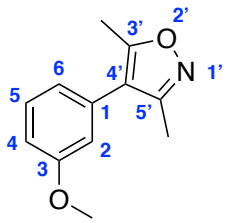
Following the procedure of Jennings *et al.*,[58] 1.3 M solution of isopropyl magnesium chloride lithium chloride complex in THF (2.00 mL, 378 mg, 2 .60 mmol, 6.0 eq.) was added a dropwise to a solution of 3-bromopyridine (215 µL, 2.22 mmol, 5.1 eq.) in anhydrous THF (2.0 mL) at room temperature. The reaction mixture was stirred for 2 h, after which time a solution of 3-(3',5'-dimethyl-1',2'-oxazol-4'-yl)-5-hydroxybenzaldehyde (**48**) (94 mg, 0.433 mmol, 1.0 eq.) in anhydrous THF (2 mL) was added dropwise, and the reaction solution stirred at room temperature for 4 h. The reaction was quenched with sat. aq. NH_4Cl (15 mL),and the aqueous phase was extracted with EtOAc (3 × 30 mL). The combined organic components were washed with H_2O (30 mL), brine (30 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc, (1:0 to 0:1) then with and EtOAc:EtOH (1:0:0 to 8:2) was followed by purification by silica gel chromatography with CH_2Cl_2:MeOH (1:0 to 9:1), to afford 3-(3,5-dimethylisoxazol-4-yl)-5-(hydroxy(pyridin-3-yl)methyl)phenol (**55**) (62 mg, 48%) as a colourless solid: R_f 0.46 (MeOH:CH_2Cl_2 = 1:9); m.p. 175–177 °C (EtOAc); ^{1}H NMR (400 MHz, DMSO-D6) δ 9.56 (1 H, s, C^1-OH), 8.61 (1 H, dd, *J* 2.3, 0.9, $C^{2''}$-H), 8.43 (1 H, dd, *J* 4.8, 1.7, $C^{6''}$-H), 7.75 (1 H, dddd, *J* 7.9, 2.3, 1.7, 0.9, $C^{4''}$-H), 7.34 (1 H, ddd, *J* 7.9, 4.8, 0.9, $C^{5''}$-H), 6.86–6.78 (2 H, m, C^4-H and C^6-H), 6.60 (1 H, dd, *J* 2.3, 1.6, C^2-H), 6.06 (1 H, d, *J* 4.1, C^7-H), 5.74 (1 H, d, *J* 4.1, C^7-OH), 2.36 (3 H, s, $C^{5'}$-CH_3), 2.18 (3 H, s, $C^{3'}$-CH_3); LRMS *m/z* (ES^+) 295 ($[M+H]^+$, 100%); Analytical HPLC (RP, method B) rt 2.05 mins, purity 99%. The data are in good agreement with literature.[58]

3-(3,5-Dimethylisoxazol-4-yl)-5-((4-methylpiperazin-1-yl)methyl)phenol (56)



3-(3,5-Dimethylisoxazol-4-yl)-5-hydroxybenzaldehyde (**48**) (50 mg, 0.23 mmol, 1.0 eq.) was added to a solution of 1-methylpiperazine (30 mg, 0.30 mmol, 1.3 eq.) in EtOH (2.5 mL) and the reaction mixture was stirred at room temperature for 5 mins. AcOH (glacial) was added to the suspension until pH 4 was reached and all solids were dissolved. The reaction mixture was stirred at room temperature for 25 mins. $NaCNBH_3$ (12 mg, 0.17 mmol, 0.7 eq.) was added as a solid and the reaction mixture was stirred for 20 h. The solvents were removed *in vaccuo* and the residue was loaded onto Celite. Purification by silica gel, eluting with CH_2Cl_2:MeOH (1.0 to 4:1), afforded 3-(3,5-dimethylisoxazol-4-yl)-5-((4-methylpiperazin-1-yl)methyl)phenol (**56**) (45 mg, 64%) as an amorphous colourless solid: R_f 0.33 (CH_2Cl_2:MeOH = 5:1); ^{1}H NMR (500 MHz, CD_3OD) δ 6.81 (1 H, dd, *J* 2.4, 1.5, C^4-H), 6.77 (1 H, dd, *J* 1.5, 1.5, C^6-H), 6.66 (1 H, dd, *J* 2.4, 1.5, C^2-H), 3.59 (2 H, s, $C^{2''}$-H or $C^{3''}$-H), 3.35 (2 H, s, $C^{2''}$-H or $C^{3''}$-H), 2.92 (2 H, s, $C^{2''}$-H or $C^{3''}$-H), 2.67 (2 H, s, $C^{2''}$-H or $C^{3''}$-H), 2.59 (3 H, s, NCH_3), 2.40 (3 H, s, $C^{5'}$-CH_3), 2.24 (3 H, s, $C^{3'}$-CH_3); LRMS *m/z* (ES^+) 302 ($[M+H]^+$, 100%). The data are in good agreement with the literature.[58]

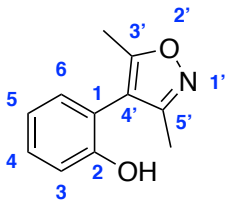
4-(3-Methoxyphenyl)-3,5-dimethylisoxazole (64)



Potassium (3,5-dimethylisoxazol-4-yl)trifluoroborate (**53**) (179 mg, 0.882 mmol, 1.10 eq.), Na_2CO_3 (255 mg, 2.406 mmol, 3.00 eq.), RuPhos (75 mg, 0.16 mmol, 0.20 eq.), and $Pd(OAc)_2$ (18 mg, 0.080 mmol, 0.10 eq.) were placed in a microwave vial. The vial was evacuated under vacuum for 5 mins then refilled with argon. 3-Bromoanisole (101 µL, 0.802 mmol, 1.00 eq.) and EtOH (4.0 mL) were added, the vial was sealed and stirred at 80 °C for 16 h. RuPhos (75 mg, 0.16 mmol, 0.20 eq.), and $Pd(OAc)_2$ (18 mg,

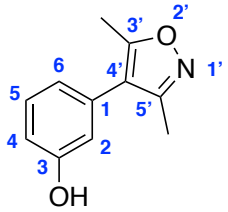
0.080 mmol, 0.10 eq.) were added, and the reaction mixture was stirred at 80 °C for 2 h. The mixture was filtered over a plug of silica, washing with EtOAc (4 × 40 mL) then concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 1:1), afforded 4-(3-methoxyphenyl)-3,5-dimethylisoxazole (145 mg, 89%) as a light yellow oil: R_f 0.43 (petroleum ether:EtOAc = 5:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.35 (1 H, ddd, *J* 8.2, 7.6, 0.4, C^5-H), 6.90 (1 H, ddd, *J* 8.2, 2.6, 1.0, C^6-H), 6.84 (1 H, ddd, *J* 7.6, 1.6, 1.0, C^4-H), 6.78 (1 H, ddd, *J* 2.6, 1.6, 0.4, C^2-H), 3.84 (3 H, s, O-CH_3), 2.41 (3 H, $C^{5'}$-CH_3), 2.28 (3 H, s, $C^{3'}$-CH_3); LRMS *m/z* (ES^+) 204 ($[M+Na]^+$, 100%); Analytical HPLC (RP, method A) rt 14.4 mins, purity > 99%. The data are in good agreement with the literature.[132,133]

2-(3,5-Dimethylisoxazol-4-yl)phenol (57)



BBr_3 (1 M in CH_2Cl_2, 0.75 mL, 0.75 mmol, 6.0 eq.) was added dropwise, over 15 min under vigorous stirring, to a solution of 4-(2-methoxyphenyl)-3,5-dimethylisoxazole (**63**) (25 mg, 0.12 mmol, 1.0 eq.) in dry CH_2Cl_2 (1 mL) at 0 °C. The reaction mixture was allowed to warm to room temperature and stirred for an additional 18 h. The reaction mixture was poured into a cooled sat. aq. $NaHCO_3$ solution (20 mL) then extracted with EtOAc (3 × 20 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (50 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was dissolved in MeOH and concentrated *in vacuo*, three successive times. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 1:4), afforded 2-(3,5-dimethylisoxazol-4-yl)phenol (**57**) (18 mg, 80%) as an off-white solid: R_f 0.20 (petroleum ether:EtOAc = 5:1); m.p. 120–122 °C (EtOAc) [lit.[134] 61–62 °C (petroleum ether)]; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.30 (1 H, ddd, *J* 8.2, 7.3, 1.8, C^4-H), 7.07 (1 H, dd, *J* 7.6, 1.7, C^6-H), 7.02–6.94 (2 H, m, C^3-H and C^5-H), 5.71 (1 H, s, OH), 2.33 (3 H, s, CH_3), 2.17 (3 H, s, CH_3); LRMS *m/z* (ES^+) 190 ($[M+H]^+$, 100%). The ^{1}H NMR data are in good agreement with the literature.[134]

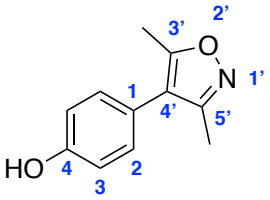
4-(3-Hydroxyphenyl)-3,5-dimethylisoxazole (58)



Method 1: BBr_3 (1 M in CH_2Cl_2, 750 µL, 0.75 mmol, 6.0 eq.) was added dropwise, over 15 min under strong stirring, to 4-(3-Methoxyphenyl)-3,5-dimethylisoxazole (**64**) (25 mg, 0.12 mmol, 1.0 eq.) in dry CH_2Cl_2 (1 mL) at 0 °C. The reaction mixture was allowed to warm to room temperature and stirred for an additional 30 mins. The reaction mixture was poured into a cooled sat. aq. $NaHCO_3$ solution (25 mL) then extracted with EtOAc (2 × 20 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (50 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was dissolved in MeOH and concentrated *in vacuo* three successive times, to afford 4-(3-hydroxyphenyl)-3,5-dimethylisoxazole (**58**) (23 mg, 98%) as a colourless solid: R_f 0.73 (petroleum ether:EtOAc = 1:1); m.p. 114–116°C (MeOH) [lit.[133] 102–106 °C (EtOAc)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.3–7.27 (1 H, m, C^5-H), 6.88 (1 H, ddd, *J* 8.2, 2.5, 1.0, C^6-H), 6.82–6.76 (2 H, m, C^2-H and C^4-H), 2.41 (3 H, s, $C^{3'}$-H_3), 2.28 (3 H, s, $C^{5'}$-H_3); LRMS *m/z* (ES^-) 188 ($[M–Na]^-$, 100%). The data are in good agreement with the literature.[96,133]

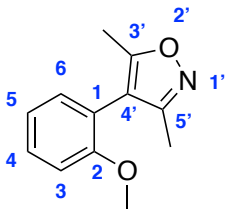
Method 2: Potassium (3,5-dimethylisoxazol-4-yl)trifluoroborate (**53**) (129 mg, 0.64 mmol, 1.1 eq.), Na_2CO_3 (129 mg, 1.2 mmol, 2.1 eq.), RuPhos (33 mg, 0.069 mmol, 0.20 eq.), and $Pd(OAc)_2$ (7.8 mg, 0.080 mmol, 0.10 eq.) were placed in a microwave vial. The vial was evacuated under vacuum for 5 mins then refilled with argon. 3-Bromophenol (**61**) (100 mg, 0.58 mmol, 1.0 eq.) and EtOH (3 mL) were added, the vial was sealed and stirred at 80 °C for 48 h. The mixture was filtered over a plug of silica, washing with EtOAc (3 × 30 mL) then concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 1:1), afforded 4-(3-hydroxyphenyl)-3,5-dimethylisoxazole (**58**) (63 mg, 58%) as a colourless solid. The data are identical as the ones obtained method 1.

4-(3,5-Dimethylisoxazol-4-yl)phenol (59)



BBr_3 (1 M in CH_2Cl_2, 0.75 mL, 0.75 mmol, 6.0 eq.) was added dropwise, over 15 min under vigorous stirring, to a solution of 4-(4-methoxyphenyl)-3,5-dimethylisoxazole (**65**) (25 mg, 0.12 mmol, 1.0 eq.) in dry CH_2Cl_2 (1 mL) at 0 °C. The reaction mixture was allowed to warm to room temperature and stirred for an additional 30 mins. The reaction mixture was poured into a cooled sat. aq. $NaHCO_3$ solution (15 mL) then extracted with EtOAc (2 × 20 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (2 × 40 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was dissolved in MeOH and concentrated *in vacuo*, three successive times, to afford 4-(3,5-dimethylisoxazol-4-yl)phenol (**59**) (23 mg, 98%) as a colourless solid: R_f 0.25 (petroleum ether:EtOAc = 5:1); m.p. 148–150 °C (MeOH); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.18–7.07 (2 H, m, C^2-H), 6.97–6.88 (2 H, m, C^3-H), 5.27 (1 H, s, O-H), 2.38 (3 H, s, $C^{5'}$-CH_3), 2.25 (3 H, s, $C^{3'}$-CH_3); LRMS *m/z* (ES^-) 288 ($[M–H]^-$, 100%); Analytical HPLC (RP, method B) rt 6.94 mins, purity > 99%. The data are in good agreement with the literature.[135]

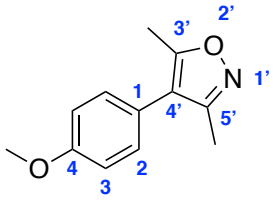
4-(2-Methoxyphenyl)-3,5-dimethylisoxazole (63)



Potassium (3,5-dimethylisoxazol-4-yl)trifluoroborate (**53**) (179 mg, 0.882 mmol, 1.10 eq.), Na_2CO_3 (255 mg, 2.41 mmol, 3.00 eq.), RuPhos (75 mg, 0.16 mmol, 0.20 eq.), and $Pd(OAc)_2$ (18 mg, 0.080 mmol, 0.10 eq.) were placed in a microwave vial. The vial was evacuated under vaccum for 5 mins then refilled with argon. 2-Bromoanisole (100 µL, 0.802 mmol, 1.00 eq.) in degassed EtOH (4 mL) was added, the vial was sealed and stirred at 80 °C for 16 h. RuPhos (75 mg, 0.16 mmol, 0.20 eq.), and $Pd(OAc)_2$ (18 mg, 0.080 mmol, 0.10 eq.) were added, and the reaction mixture was stirred at 80 °C for 3 h. The mixture was filtered over a plug of silica, washing with EtOAc (4 × 40 mL) then concentrated *in*

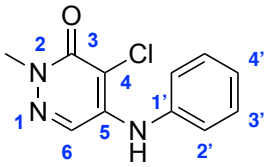
vacuo. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 1:1), afforded 4-(2-methoxyphenyl)-3,5-dimethylisoxazole (**63**) (68 mg, 42%) as a colourless oil: R_f 0.52 (petroleum ether:EtOAc = 5:1); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.36 (1 H, ddd, *J* 8.3, 7.5, 1.8, C^5-H), 7.12 (1 H, dd, *J* 7.6, 1.8, C^3-H), 7.01 (1 H, ddd, *J* 7.6, 7.5, 1.1, C^4-H), 6.98 (1 H, dd, *J* 8.3, 1.1, C^5-H), 3.80 (3 H, s, OCH_3), 2.30 (3 H, s, $C^{3'}$-CH_3), 2.17 (3 H, s, $C^{5'}$-CH_3); LRMS *m/z* (ES^+) 203 ($[M+H]^+$, 100%); Analytical HPLC (RP, method A) rt 14.4 mins, purity > 99%. The data are in good agreement with the literature.[136]

4-(4-Methoxyphenyl)-3,5-dimethylisoxazole (65)



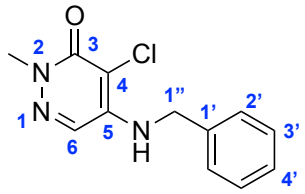
Potassium (3,5-dimethylisoxazol-4-yl)trifluoroborate (**53**) (180 mg, 0.88 mmol, 1.1 eq.), Na_2CO_3 (255 mg, 2.4 mmol, 3.0 eq.), RuPhos (75 mg, 0.16 mmol, 0.20 eq.), and $Pd(OAc)_2$ (18 mg, 0.080 mmol, 0.10 eq.) were placed in a microwave vial. The vial was evacuated *in vacuo* for 5 mins then refilled with argon. 4-Bromoanisole (100 µL, 0.80 mmol, 1.0 eq.) and EtOH (4 mL) were added, the vial was sealed and stirred at 80 °C for 16 h. The mixture was filtered over a plug of silica, washing with EtOAc (5 × 20 mL) then concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 1:1), afforded 4-(4-methoxyphenyl)-3,5-dimethylisoxazole (120 mg, 72%) as a colourless solid: R_f 0.70 (petroleum ether:EtOAc = 5:2); m.p. 63–65°C (CH_2Cl_2) [lit.[137] 66–66 °C (EtOAc)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.21–7.15 (2 H, m, C^2-H), 7.01–6.89 (2 H, m, C^3-H), 3.85 (3 H, s, OCH_3), 2.38 (3 H, s, $C^{5'}$-CH_3), 2.25 (3 H, s, $C^{3'}$-CH_3); LRMS *m/z* (ES^+) 204 ($[M+H]^+$, 40%); Analytical HPLC (RP, method A) rt 2.96 mins, purity > 99%. The data are in good agreement with the literature.[137,138]

4-Chloro-2-methyl-5-(phenylamino)pyridazin-3(2*H*)-one (66)



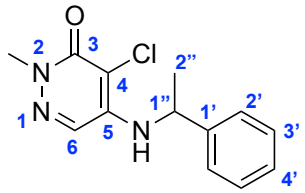
A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (180 mg, 1.00 mmol, 1.0 eq.), aniline (455 µL, 5.00 mmol, 5.00 eq.) in NMP (0.50 mL) was stirred at 120 °C for 16 h. Aniline (455 µL, 5.00 mmol, 5.00 eq.) was added and the reaction mixture heated to 120 °C under microwave irradiation for 2 h. Aniline (455 µL, 5.00 mmol, 5.00 eq.) was added and the reaction mixture heated to 120 °C under microwave irradiation for 1 h. The reaction mixture was cooled to room temperature and diluted with H_2O (20 mL) and 1 M HCl (20 mL), then extracted with EtOAc (2 × 60 mL). The combined organic components were washed with 1 M HCl (80 mL) and aq. LiCl (10% *w/v*, 6 × 80 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 0:1) afforded 4-chloro-2-methyl-5-(phenylamino)pyridazin-3(2*H*)-one (143 mg, 60%) as a purple solid. A portion of the product (16 mg) was stirred with decolourising charcoal in CH_2Cl_2 (3 mL) for 2 h at room temperature, filtered through Celite® and concentrated *in vacuo* to give 4-chloro-2-methyl-5-(phenylamino)pyridazin-3(2*H*)-one (**66**) (14 mg, 93% recovery) as a colourless solid: R_f 0.35 (petroleum ether:EtOAc = 1:1); m.p. 154–156 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3255 (w), 1632 (w), 1678 (C=O), 1526 (m), 1494 (m), 1345 (w), 875 (m), 749 (m); ^{1}H NMR (400 MHz, DMSO-D_6) δ 8.75 (1 H, s, C^6-H), 7.65 (1 H, s, N-H), 7.45–7.36 (2 H, m, $C^{3'}$-H), 7.29–7.24 (2 H, m, $C^{2'}$-H), 7.24–7.15 (2 H, m, $C^{4'}$-H); ^{13}C NMR (101 MHz, DMSO-D_6) δ 157.5 (C=O), 142.8 (C^6), 139.0 ($C^{1'}$), 129.8 ($C^{3'}$), 128.1 (C^4), 125.4 ($C^{4'}$), 124.0 ($C^{2'}$), 109.0 (C^4), 40.1 (N-CH_3); LRMS *m/z* (ES^-) 234 ($[M–H]^-$, 100%); HRMS *m/z* (ES^-) found 234.04393 (100%), $C_{11}H_{10}ON_3Cl$ requires $[M^{35}Cl–H]^-$ 234.04396; Analytical HPLC (RP, method A) rt 7.67 mins, purity 99%.

5-(Benzylamino)-4-chloro-2-methylpyridazin-3(2*H*)-one (67)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (90 mg, 0.50 mmol, 1.0 eq.) and $BnNH_2$ (110 µL, 1.0 mmol, 2.0 eq.) in NMP (1 mL) was stirred at room temperature for 16 h. TLC analysis showed the presence of remaining starting material. Additional $BnNH_2$ (110 µL, 1.0 mmol, 2.0 eq.) was added and the solution was stirred at 80 °C for 3 h. The reaction mixture was cooled to room temperature and diluted with H_2O (2 mL), then extracted with EtOAc (2 × 60 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 4 × 60 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 1:4), afforded 5-(benzylamino)-4-chloro-2-methylpyridazin-3(2*H*)-one (**67**) (98.8 mg, 80%) as a colourless solid: R_f 0.34 (petroleum ether:EtOAc = 1:1); m.p. 164–166 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3283 (w), 1644 (w), 1604 (s) (C=O), 1519 (m), 1313 (w), 1005 (w), 875 (m), 764 (w), 695 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.50 (1 H, s, C^6-H), 7.41–7.36 (2 H, m, $C^{3'}$-H), 7.36–7.32 (1 H, m, $C^{4'}$-H), 7.32–7.28 (2 H, m, $C^{2'}$-H), 5.10 (1 H, s, NH), 4.55 (2 H, d, *J* 6.0, $C^{1''}$-H_2), 3.75 (3 H, s, N^2-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 157.9 (C^3), 144.1 (C^5), 136.9 ($C^{1'}$), 129.3 ($C^{3'}$), 128.3 ($C^{4'}$), 127.0 ($C^{2'}$), 125.5 (C^6), 108.3(C^4), 47.3 ($C^{1''}$), 40.4 (CH_3); LRMS *m/z* (ES^-) 248 ($[M–H]^-$, 100%); HRMS *m/z* (ES^+) found 250.07393 (100%), $C_{12}H_{12}ClN_3O$ requires $[M^{35}Cl+H]^+$ 250.07417; Analytical HPLC (RP, method A) rt 7.75 mins, purity > 99%.

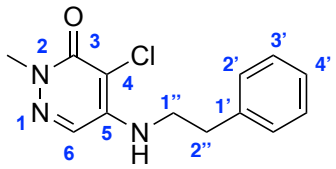
4-Chloro-2-methyl-5-((1-phenylethyl)amino)pyridazin-3(2*H*)-one (68)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (180 mg, 1.00 mmol, 1.0 eq.), 1-phenylethylamine (200 µL, 1.50 mmol, 1.5 eq.) and pyridine (160 µL, 2.00 mmol, 2.0 eq.) in NMP (1.5 mL) was stirred at 80 °C for 4 h. 1-Phenylethylamine (200 µL, 1.50 mmol,

1.5 eq.) and pyridine (160 µL, 2.00 mmol, 2.0 eq.) were added and the solution was stirred at 40 °C for 16 h. The reaction mixture was cooled to room temperature and diluted with H_2O (20 mL), then extracted with EtOAc (3 × 30 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 4 × 60 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 1:1), afforded 4-chloro-2-methyl-5-((1-phenylethyl)amino)pyridazin-3(2*H*)-one (**68**) (175 mg, 67%) as a colourless solid: R_f 0.35 (petroleum ether:EtOAc = 1:1); m.p. 106–108 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3264 (w), 1633 (s) (C=O), 1434 (w), 1347 (w), 849 (m), 705 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.27–7.22 (2 H, m, C$^{2'}$-H), 7.19 (1 H, s, C^6-H), 7.19–7.14 (3 H, m, C$^{3'}$-H and C$^{4'}$-H), 4.84 (1 H, d, *J* 6.2, N-H), 4.58 (1 H, dq, *J* 6.2, 6.7, C$^{1''}$-H), 3.57 (3 H, s, N^2-CH_3), 1.51 (3 H, d, *J* 6.7, C$^{2''}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 157.8 (C^3), 143.4 (C^5), 142.8 (C$^{1'}$), 129.3 (C$^{3'}$), 127.9 (C^6), 126.0 (C$^{4'}$), 125.4 (C$^{2'}$), 108.2 (C^4), 53.3 (C$^{1''}$), 40.2 (N^2-CH_3), 25.0 (C$^{2''}$-H_3); LRMS *m/z* (ES$^+$) 264 ([M+H]$^+$, 100%), 286 ([M+Na]$^+$, 70%); HRMS *m/z* (ES$^+$) found 264.08975 (100%), $C_{13}H_{15}ClN_3O$ requires [M^{35}Cl+H]$^+$ 264.008982; Analytical HPLC (RP, method A) rt 8.33 mins, purity > 99%.

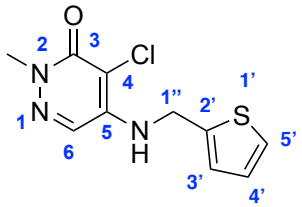
4-Chloro-2-methyl-5-(phenethylamino)pyridazin-3(2*H*)-one (69)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (0.18 g, 1.0 mmol, 1.0 eq.) and 2-phenylethan-1-amine (0.40 mL, 2.0 mmol, 2.0 eq.) in NMP (1.5 mL) was stirred at 80 °C for 6 h. The reaction mixture was cooled to room temperature and diluted with H_2O (25 mL), then extracted with EtOAc (2 × 40 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 4 × 30 mL), dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 0:1), afforded 4-chloro-2-methyl-5-(phenethylamino)pyridazin-3(2*H*)-one (**69**) (224 mg, 84%) as a colourless solid: R_f 0.46 (EtOAc); m.p. 140–142 °C (CH_2Cl_2); υ_{max} (neat)/cm^{-1} 3301 (w), 2360 (w), 1601 (s) (C=O), 1521 (m), 1449 (m), 1321 (w), 873 (m), 737 (m), 694 (m); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.50 (1 H, s, C^6-H), 7.37–7.31 (2 H, m, C$^{3'}$-H), 7.31–7.26 (1 H, m,

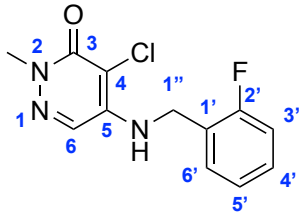
$C^{4'}$-H), 7.22–7.17 (2 H, m, $C^{2'}$-H), 4.69 (1 H, s, NH), 3.75 (3 H, s, N^2-CH_3), 3.58 (3 H, td, *J* 6.9, 6.0, $C^{1''}$-H), 2.94 (2 H, t, *J* 6.9, $C^{2''}$-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 157.8 (C^3), 144.0 (C^5), 137.45($C^{1'}$), 129.1 ($C^{3'}$), 128.8 ($C^{2'}$), 127.2 ($C^{4'}$), 125.2 (C^6), 107.9 (C^4), 44.5 ($C^{1''}$), 40.3 (N^2-CH_3), 36.3 ($C^{2''}$); LRMS *m/z* (ES^+) 286 ([$M^{35}Cl$+Na]$^+$, 100%), 288 ([$M^{37}Cl$+Na]$^+$, 30%); HRMS *m/z* (ES^+) found 264.08972 (100%), $C_{13}H_{14}ClN_3O$ requires [$M^{35}Cl$+H]$^+$ 264.08982; Analytical HPLC (RP, method A) rt 8.13 mins, purity 99%.

4-Chloro-2-methyl-5-((thiophen-2-ylmethyl)amino)pyridazin-3(2*H*)-one (70)



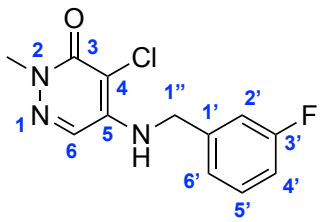
A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (90 mg, 0.50 mmol, 1.0 eq.), thiophen-2-ylmethanamine (77 µL, 0.75 mmol, 1.5 eq.) and DIPEA (0.25 mL, 1.5 mmol, 3.0 eq.) in NMP (0.75 mL) was stirred at 80 °C for 18 h. The reaction mixture was cooled to room temperature and diluted with H_2O (20 mL), then extracted with EtOAc (2 × 40 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 6 × 30 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 0:1) afforded 4-chloro-2-methyl-5-((thiophen-2-ylmethyl)amino)pyridazin-3(2*H*)-one (**70**) (82 mg, 63%) as an off-white solid: R_f 0.42 (petroleum ether:EtOAc = 1:1); m.p. 138–140 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 32926 (w), 3081 (w), 1588 (s) (C=O), 1513 (m), 1451 (w), 1212 (w), 997 (m), 867 (m), 772 (m), 659 (m); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.59 (1 H, s, C^6-H), 7.28 (1 H, dd, *J* 5.1, 1.2, $C^{5'}$-H), 7.03 – 6.97 (2 H, m, $C^{3'}$-H, $C^{4'}$-H), 5.16 (1 H, s, N-H), 4.71 (2 H, d, *J* 6.5, $C^{1''}$-H_2), 3.75 (3 H, s, N^2-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 157.5 (C^3), 142.8 (C^5), 139.0 ($C^{2'}$), 129.9 ($C^{5'}$), 128.1 (C^6), 125.4 ($C^{3'}$ or $C^{4'}$), 124.0 ($C^{3'}$ or $C^{4'}$), 109.0 (C^4), 42.5 ($C^{1''}$), 40.1 (N^2-CH_3). LRMS *m/z* (ES^-) 254 ([M–H]$^-$, 100%); HRMS *m/z* (ES^-) found 254.01618 (100%), $C_{10}H_{10}ON_3{}^{35}Cl^{32}S$ requires [$M^{35}C$–H]$^-$ 254.01603; Analytical HPLC (RP, method A) rt 7.44 mins, purity 97%.

4-Chloro-5-((2-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (71)



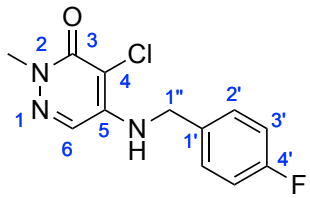
A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (90 mg, 0.50 mmol, 1.0 eq.), 2-(fluorophenyl)methanamine (85 µL, 0.75 mmol, 1.5 eq.) and DIPEA (0.25 mL, 1.5 mmol, 3.0 eq.) in NMP (0.75 mL) was stirred at 80 °C for 18 h. The reaction mixture was cooled to room temperature and diluted with H_2O (20 mL), then extracted with EtOAc (2 × 35 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 5 × 30 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 0:1) afforded 4-chloro-5-((2-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (**71**) (101 mg, 75%) as an off-white solid: R_f 0.21 (petroleum ether:EtOAc = 1:1); m.p. 182–184 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3282 (CH), 1632 (w), 1606 (C=O), 1455 (w), 1223 (w), 872, 759, 742; ^{1}H NMR (500 MHz, acetone-D_6) δ 7.70 (1 H, s, C^6-H), 7.46–7.38 (1 H, m, C$^{4'}$-H or C$^{6'}$-H), 7.38–7.32 (1 H, m, C$^{4'}$-H or C$^{6'}$-H), 7.23–7.12 (2 H, m, C$^{3'}$-H, C$^{5'}$-H), 6.45 (1 H, br s, NH), 4.77 (2 H, d, *J* 6.3, C$^{1''}$-H_2), 3.60 (3 H, s, N^2-CH_3); ^{13}C NMR (126 MHz, acetone-D_6) δ 161.3 (d, J_{CF} 239, C$^{2'}$), 157.7 (C^3=O), 145.3 (C^5), 130.34 (d, J_{CF} 8.2, C$^{1'}$), 129.8 (d, J_{CF} 4.4, C$^{4'}$ or C$^{6'}$), 126.5 (d, J_{CF} 14.7, C$^{4'}$ or C$^{6'}$), 126.3 (C^6), 125.51 (d, J_{CF} 3.5, C$^{5'}$), 116.3 (d, J_{CF} 21.4, C$^{3'}$), 107.4 (C^4), 40.9 (d, J_{CF} 4.8, C$^{1''}$), 39.9 (NCH_3); ^{19}F NMR (377 MHz, acetone-D_6) δ –120.0 — –120.1 (1 F, m); LRMS *m/z* (ES$^-$) 266 ([M–H]$^-$, 100%); HRMS *m/z* (ES$^+$) found 268.06469 (100%), $C_{12}H_{11}ClFN_3O$ requires [M^{35}Cl+H]$^+$ 268.06474; Analytical HPLC (RP, method A) rt 7.89 mins, purity > 99%.

4-Chloro-5-((3-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (72)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (90 mg, 0.50 mmol, 1.0 eq.), 3-(fluorophenyl)methanamine (85 µL, 0.75 mmol, 1.5 eq.) and and DIPEA (0.25 mL, 1.5 mmol, 3.0 eq.) in NMP (0.75 mL) was stirred at 80 °C for 18 h. The reaction mixture was cooled to room temperature and diluted with H_2O (20 mL), then extracted with EtOAc (2 × 35 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 5 × 30 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 0:1) afforded 4-chloro-5-((3-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (**72**) (98 mg, 70%) as a colourless solid: R_f 0.15 (petroleum ether:EtOAc = 1:1); m.p. 130–132 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3299 (CH), 2979 (w), 1631 (w), 1606 (s) (C=O), 1449 (w), 1249 (w), 872 (m), 768 (m), 768 (m); ^{1}H NMR (500 MHz, acetone-D_6) δ 7.63 (1 H, s, C^6-H), 7.41 (1 H, td, *J* 8.0, 6.0, $C^{5'}$-H), 7.27–7.24 (1 H, m, $C^{4'}$-H or $C^{6'}$-H), 7.17 (1 H, ddd, *J* 9.9, 2.1, 2.1, $C^{4'}$-H or $C^{6'}$-H), 7.08–7.01 (1 H, m, $C^{2'}$-H), 6.62 (1 H, br s, NH), 4.75 (2 H, d, *J* 6.3, $C^{1''}$-H_2), 3.59 (3 H, s, N^2-CH_3); ^{13}C NMR (126 MHz, , acetone-D_6) δ 164.4–163.2 (m, $C^{3'}$), 156.8 (C^3), 144.5 (C^5), 141.98 (d, J_{CF} 6.4, $C^{1'}$), 130.61 (d, J_{CF} 8.2, $C^{5'}$), 125.5 (C^6), 122.63 (d, J_{CF} 2.8, $C^{6'}$), 114.02 (d, J_{CF} 21.4, $C^{2'}$), 113.55 (d, J_{CF} 22.2, $C^{4'}$), 106.6 (C^4), 45.4 ($C^{1''}$), 39.0 (NCH_3); ^{19}F NMR (377 MHz, acetone-D_6) δ –114.4 — –114.6 (1 F, m); LRMS *m/z* (ES^-) 266 ([M–H]$^-$, 100%); HRMS *m/z* (ES^+) found 268.06462 (100%), $C_{12}H_{11}ClFN_3O$ requires [M^{35}Cl+H]$^+$ 268.06474; Analytical HPLC (RP, method A, 300 nm) rt 7.91 mins, purity > 99%.

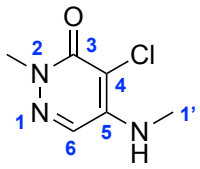
4-Chloro-5-((4-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (73)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (180 mg, 1.00 mmol, 1.0 eq.), 4-(fluorobenzyl)methanamine (172 µL, 1.50 mmol, 1.5 eq.) and pyridine (160 µL, 2.00 mmol, 2.0 eq.) in NMP (1.5 mL) was stirred at 80 °C for 18 h. 4-Fluorobenzylamine (172 µL, 1.50 mmol, 1.5 eq.) and pyridine (160 µL, 2.00 mmol, 2.0 eq.) were added and the solution was stirred at 80 °C for 16 h. The reaction mixture was cooled to room temperature and diluted with H_2O (25 mL), then extracted with EtOAc (3 × 25 mL). The

combined organic components were washed with aq. LiCl (10% *w/v*, 4 × 40 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 10:8), afforded 4-chloro-5-((4-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (**73**) (165 mg, 62%) as a colourless solid: R_f 0.28 (petroleum ether:EtOAc = 1:1); m.p. 200–202 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3282 (w), 1605 (s) (C=O), 1505 (m), 1159 (m), 876 (m), 815 (m), 656 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.47 (1 H, s, C^6-H), 7.31–7.26 (2 H, m, $C^{2'}$-H), 7.12–7.05 (2 H, m, $C^{3'}$-H), 5.06 (1 H, br t, *J* 5.2, N-H), 4.52 (2 H, dd, *J* 6.0, 0.9, $C^{2'}$-H_2), 3.75 (3 H, s, N^2-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) 162.5 (d, J_{CF} 247, $C^{4'}$), 157.7 (C^3=O), 143.8 (C^5), 132.51 (d, J_{CF} 3.3, $C^{1'}$), 128.60 (d, J_{CF} 8.4, $C^{2'}$), 125.2 (C^6), 116.2 (d, J_{CF} 21.5, $C^{3'}$), 108.4 (C^4), 46.5 ($C^{1''}$), 40.3 (CH_3) ^{19}F NMR (377 MHz, $CDCl_3$) δ –113.6 — –113.8 (1 F, m); LRMS *m/z* (ES^-) 266 ($[M–H]^-$, 100%); HRMS *m/z* (ES^+) found 268.06476 (100%), $C_{12}H_{11}ClFN_3O$ requires $[M^{35}Cl+H]^+$ 268.06474; Analytical HPLC (RP, method A) rt 7.95 mins, purity 99%.

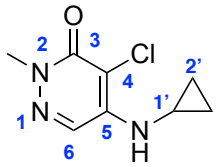
4-Chloro-2-methyl-5-(methylamino)pyridazin-3(2*H*)-one (74)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (90 mg, 0.50 mmol, 1.0 eq.), $MeNH_2$·HCl (101 mg, 1.5 mmol, 3.0 eq.) and DIPEA (44 µL, 2.5 mmol, 5.0 eq.) in NMP (1 mL) was stirred at 60 °C for 16 h. TLC analysis showed the presence of remaining starting material. Additional $MeNH_2$·HCl (101 mg, 1.5 mmol, 3.0 eq.) and DIPEA (44 µL, 2.5 mmol, 5.0 eq.) were added and the solution was stirred at 80 °C for 3 h. The reaction mixture was cooled to room temperature and diluted with H_2O (2 mL), then extracted with EtOAc (2 × 60 mL). The combined organic components were washed with LiCl (10% *w/v*, 4 × 60 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 10:7), recovered a portion of the starting material (20 mg, 23% recovery) and afforded 4-chloro-2-methyl-5-(methylamino)pyridazin-3(2*H*)-one (**74**) (14 mg, 16%) as a colourless solid: R_f 0.29 (petroleum ether:EtOAc = 1:1); m.p.

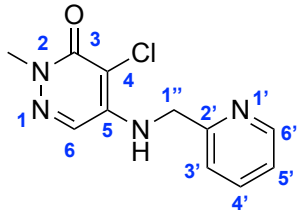
140–142 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3307 (w), 2922 (w), 2851 (w), 1605 (s) (C=O), 1562 (w), 1421 (w), 1036 (w), 962 (w), 773 (m), 665; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.48 (1 H, s, C^6-H), 5.88 (1 H, s, N-H), 3.72 (3 H, s, N^2-CH_3), 2.17 (3 H, s, $C^{1'}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.4 (C^3), 140.3 (C^5), 139.6 (C^6), 106.2 (C^4), 40.0 ($C^{1'}$), 31.6 (N^2-CH_3); LRMS *m/z* (ES^+) 174 ([M+H]$^+$, 100%); HRMS *m/z* (ES^+) found 174.0430 (100%), $C_6H_8ClN_3O$ requires [M^{35}Cl+H]$^+$ 174.0429; Analytical HPLC (RP, method A) rt 7.01 mins, purity 98%.

4-Chloro-5-(cyclopropylamino)-2-methylpyridazin-3(2*H*)-one (75)



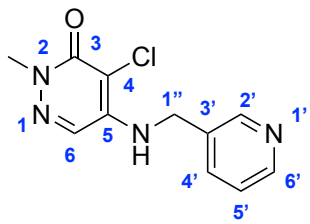
A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (180 mg, 1.00 mmol, 1.0 eq.) and cyclopropanamine (208 µL, 3.00 mmol, 3.0 eq.) in NMP (1.5 mL) was stirred at 80 °C for 7 h. The reaction mixture was cooled to room temperature and diluted with H_2O (25 mL), then extracted with EtOAc (2 × 40 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 4 × 30 mL), then dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with eluting with petroleum ether:EtOAc (1:0 to 0:1), followed by purification by silica gel chromatography, eluting with petroleum ether:EtOAc (1:0 to 2:1), afforded 4-chloro-5-(cyclopropylamino)-2-methylpyridazin-3(2*H*)-one (**75**) (18 mg, 9%) as a yellow amorphous residue: R_f 0.72 (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3277 (w), 2924 (w), 1620 (s) (C=O), 1591 (m), 1544 (m), 1490 (m), 1372 (w), 919 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.51 (1 H, s, C^6-H), 6.00 (1 H, br s, NH), 3.71 (3 H, s, N^5-CH_3), 3.09 (1 H, dtt, *J* 6.9, 3.5, 3.5, $C^{1'}$-H), 0.89–0.79 (2 H, m, $C^{2'}$-H_a), 0.74–0.63 (2 H, m, $C^{2'}$-H_b); ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.2 (C^3), 140.3 (C^5), 139.6 (C^6), 107.1 (C^4), 40.0 (N^5-CH_3), 26.3 ($C^{1'}$), 10.2 ($C^{2'}$); LRMS *m/z* (ES^+) 200 ([M^{35}Cl+H]$^+$, 100%); HRMS *m/z* (ES^+) found 200.0587 (100%), $C_8H_{10}ClON_3$ requires [M^{35}Cl+H]$^+$ 200.0585; Analytical HPLC (RP, method A) rt 8.35 mins, purity > 99%.

4-Chloro-2-methyl-5-((pyridin-2-ylmethyl)amino)pyridazin-3(2*H*)-one (76)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (180 mg, 1.00 mmol, 1.0 eq.) and 2-(aminomethyl)pyridine (515 µL, 5.00 mmol, 5.0 eq.) in NMP (1.5 mL) was stirred at 80 °C for 7 h. The reaction mixture was cooled to room temperature and diluted with H_2O (20 mL), then extracted with EtOAc (2 × 30 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 4 × 40 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with Petroluem ether:EtOAc (1:0 to 0:1), afforded 4-chloro-2-methyl-5-((pyridin-2-ylmethyl)amino)pyridazin-3(2*H*)-one (**76**) (116 mg, 46%) as a colourless solid: R_f 0.22 (EtOAc); m.p. 130–138 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3328 (w), 2060 (w), 2160 (w), 1615 (C=O), 1521 (m), 1442 (m), 996 (m), 872 (m), 837 (w), 776 (m); ^{1}H (500 MHz, $CDCl_3$) δ 8.63 (1 H, ddd, *J* 5.0, 1.8, 1.0, C$^{6'}$-H), 7.73 (1 H, ddd, *J* 7.7, 7.7, 1.8, C$^{4'}$-H), 7.59 (1 H, s, C^6-H), 7.32–7.26 (2 H, m, C$^{3'}$-H and C$^{5'}$-H), 6.13 (1 H, br t, *J* 5.3, NH), 4.65 (2 H, d, *J* 5.3, C$^{1''}$-H), 3.77 (3 H, s, CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 157.9 (C^3), 155.2 (C$^{2'}$), 149.5 (C$^{6'}$), 143.8 (C^5), 137.1 (C$^{4'}$), 125.7 (C^6), 122.9 (C$^{3'}$ or C$^{5'}$), 121.4 (C$^{3'}$ or C$^{5'}$), 108.2 (C^4), 47.5 (C$^{1''}$), 40.3 (CH_3); LRMS *m/z* (ES$^+$) 273 ([M+Na]$^+$, 100%); HRMS *m/z* (ES$^+$) found 273.05127 (100%), $C_{11}H_{11}ON_4Cl$ requires [M^{35}Cl +Na]$^+$ 273.05136; Analytical HPLC (RP, method A) rt 5.97 mins, purity 99%.

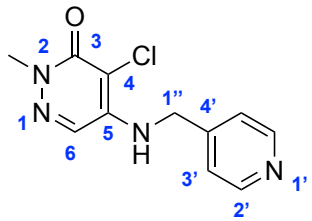
4-Chloro-2-methyl-5-((pyridin-3-ylmethyl)amino)pyridazin-3(2*H*)-one (77)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (180 mg, 1.00 mmol, 1.0 eq.) and 3-(aminomethyl)pyridine (509 µL, 5.00 mmol, 5.0 eq.) in NMP (1.5 mL) was stirred at 80 °C for 7 h. The reaction mixture was cooled to room temperature and diluted with H_2O (20 mL), then extracted with EtOAc (2 × 35 mL). The combined organic components

were washed with aq. LiCl (10% *w/v*, 4 × 40 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with EtOAc:EtOH (1:0 to 10:2), afforded 4-chloro-2-methyl-5-((pyridin-3-ylmethyl)amino)pyridazin-3(2*H*)-one (**77**) (58 mg, 23%) as a colourless solid: R_f 0.28 (EtOAc); m.p. 86–88 °C (EtOH); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3292 (CH), 2980 (w), 2160 (w), 1592 (C=O), 1542 (m), 1479 (m), 1424 (m), 933 (m), 897(m) , 712 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.59 (1 H, dd, *J* 2.3, 1.6, $C^{2'}$-H), 8.57 (1 H, dd, *J* 4.8, 1.6, $C^{6'}$-H), 7.65 (1 H, dddd, *J* 7.9, 2.3, 1.6, 0.8, $C^{4'}$-H), 7.51 (1 H, s, C^6-H), 7.31 (1 H, ddd, *J* 7.9, 4.8, 0.8, $C^{5'}$-H), 6.20 (1 H, br s, NH), 4.98 (2 H, d, *J* 6.8, $C^{1''}$-H_2), 3.76 (3 H, s CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.3 (C^3), 149.2 ($C^{2'}$ or $C^{6'}$), 148.9 ($C^{2'}$ or $C^{6'}$), 139.1 (C^6), 134.8 ($C^{3'}$ and $C^{4'}$), 134.3 (C^5), 123.7 ($C^{5'}$), 107.7 (C^4), 45.3 ($C^{1''}$), 40.1 (N^2-CH_3); LRMS *m/z* (ES^-) 249 ($[M–H]^-$, 100%); HRMS *m/z* (ESI^+) found 251.0696, $C_{11}H_{11}N_4OCl$ requires $[M^{35}Cl+H]^+$ 251.0694; Analytical HPLC (RP, method A) rt 6.97 mins, purity 99%.

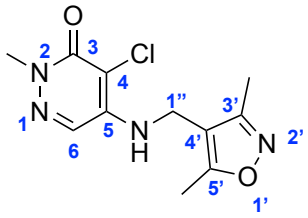
4-Chloro-2-methyl-5-((pyridin-4-ylmethyl)amino)pyridazin-3(2*H*)-one (78)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (180 mg, 1.00 mmol, 1.0 eq.), 4-(aminomethyl)pyridine (508 µL, 5.00 mmol, 5.0 eq.) in NMP (1.5 mL) were stirred at 80 °C for 7 h. The reaction mixture was cooled to room temperature and diluted with H_2O (20 mL), then extracted with EtOAc (2 × 35 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 4 × 40 mL), then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with EtOAc:EtOH (1:0 to 10:3), afforded 4-chloro-2-methyl-5-((pyridin-4-ylmethyl)amino)pyridazin-3(2*H*)-one (**78**) (33 mg, 13%) as a colourless solid: R_f 0.10 (EtOAc); m.p. 202–204 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3216 (w), 1598 (C=O), 1519 (m), 1402 (m), 1231 (w), 972 (m), 874 (m), 720 (m); ^{1}H (500 MHz, $CDCl_3$) δ 8.61–8.50 (2 H, m, $C^{2'}$-H), 7.48 (1 H, s, C^6-H), 7.22–7.14 (2 H, m, $C^{3'}$-H), 6.30 (1 H, br t, *J* 7.3, N-H), 4.94 (2 H, d, *J* 7.3, $C^{1''}$-H_2), 3.75 (3 H, s, N^2-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.3 (C^3), 150.3 ($C^{2'}$), 148.3 ($C^{4'}$), 139.20 (C^5), 139.19 (C^6), 121.6 ($C^{3'}$), 107.9 (C^4), 46.5

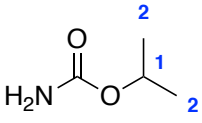
($C^{1''}$), 40.3 (N^2-CH_3); LRMS *m/z* (ES^+) 251 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) found 251.0696 (100%), $C_{11}H_{11}ON_4Cl$ requires $[M^{35}Cl+H]^+$ 251.0694; Analytical HPLC (RP, method B) rt 5.56 mins, purity 95%.

4-Chloro-5-(((3',5'-dimethylisoxazol-4'-yl)methyl)amino)-2-methylpyridazin-3(2*H*)-one (79)



A solution of 4,5-dichloro-2-methylpyridazin-3(2*H*)-one (180 mg, 1.00 mmol, 1.0 eq.), and (3,5-dimethylisoxazol-4-yl)methanamine (380 mg, 3.0 mmol, 3.0 eq.) in NMP (1.5 mL) was stirred at 80 °C for 6 h. The reaction mixture was cooled to room temperature and diluted with H_2O (30 mL), then extracted with EtOAc (2 × 50 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 4 × 50 mL), dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. The crude was first purified using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 0:1). Then, purification by silica gel chromatography, eluting with petroleum ether:EtOAc (1:0 to 0:1), afforded 4-chloro-5-(((3',5'-dimethylisoxazol-4'-yl)methyl)amino)-2-methylpyridazin-3(2*H*)-one (**79**) (148 mg, 55%) as a colourless amorphous residue: R_f 0.40 (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3291 (w), 2360 (w), 1607 (s) (C=O), 1446 (m), 997 (m), 872 (m); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 (1 H, s, C^6-H), 4.52 (1 H, br s, NH), 4.25 (2 H, d, *J* 5.1, $C^{1''}$-H), 3.78 (3 H, s, N^2-CH_3), 2.42 (3 H, s, CH_3), 2.28 (3 H, s, CH_3); LRMS *m/z* (ES^+) 269 ($[M^{35}Cl+H]^+$, 100%), HRMS *m/z* (ES^+) found 269.0800 (100%), $C_{11}H_{13}ClO_2N_4$ requires $[M^{35}Cl+H]^+$ 269.0800; Analytical HPLC (RP, method A) rt 6.27 mins, purity > 99%.

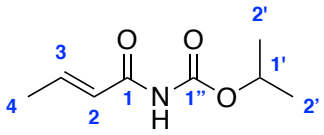
Isopropyl carbamate (82)



TFA (20 mL, 30 g, 0.3 mol, 2.0 eq.) was added dropwise to a solution of ^{i}PrOH (10 mL, 7.8 g, 0.13 mol, 1.0 eq.) and NaOCN (12.6 g, 0.26 mol, 1.5 eq.) in toluene (40 mL) at 0 °C. The

biphasic reaction mixture was stirred for 4 h. H_2O (100 mL) was added and the mixture was extracted with EtOAc (2 × 150 mL). The combined organic components were washed with brine (75 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude material (10.7 g) was dissolved in hot EtOAc (15 mL) and hexane (200 mL, added in small portion) over a water bath. When the solids were completely dissolved in the hot solvents the mixture was cooled to room temperature then placed in the freezer overnight. The needle-like crystals were washed with successive addition and removal of hexane by pipetting, then the crystals were collected by filtration to afford isopropyl carbamate (**82**) as colourless crystals (8.25 g, 60%): m.p. 90–92 °C (EtOAc and hexane) [lit.[139] 88–90 °C (Et_2O)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.89 (1 H, hept, *J* 6.3, C^1-H), 4.67 (s, 2H, NH_2), 1.24 (6 H, d, *J* 6.3, C^2-H_3); LRMS *m/z* (ES^+) 126 ([M+Na]$^+$, 100%). The data are in good agreement with literature.[139]

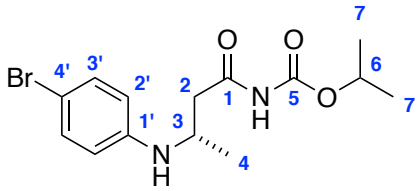
Isopropyl (*E*)-but-2-enoylcarbamate (83)



In a flame-dried pear-shaped 2-neck flask, fitted with a thermometer, isopropylcarbamate (500 mg, 4.84 mmol, 1.00 eq.) was dissolved in distilled THF (5 mL). Crotonyl chloride (520 µL, 570 mg, 5.43 mmol, 1.12 eq.) was added. The reaction mixture was cooled to approximately –40 °C then LiO^tBu (1 M in THF, 10.9 mL, 10.9 mmol, 2.25 eq.) was added dropwise over an hour. During the addition the reaction mixture was kept below –30 °C using an unsaturated dry ice/acetone bath. The reaction mixture was left to warm to 0 °C and stirred for 30 mins. A mixture of Et_2O and H_2O (25 mL and 20 mL, respectively) was added followed by vigorous stirring until the solution reached room temperature. The layers were separated, then the aqueous layer was extracted with Et_2O (15 mL). The combined organic components were washed with brine (15 mL) and H_2O (15 mL) then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether and EtOAc (gradient from 1:0 to 1:1) with 0.5% of Et_3N, afforded isopropyl (*E*)-but-2-enoylcarbamate (**83**) (322 mg, 39%) as a colourless solid: R_f 0.70 (petroleum ether:EtOAc = 2:1); mp: 72–74 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3270 (w), 2982 (w), 1752 (s) (C=O), 1649

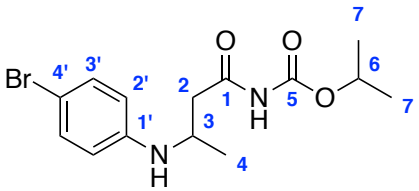
(m), 1506 (s), 1102 (s), 938 (s); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.27 (1 H, br s, NH), 7.13 (1 H, dq, *J* 15.3, 6.9, C^{3}-H), 6.87 (1 H, dq, *J* 15.3, 1.7, C^{2}-H), 4.98 (1 H, hept, *J* 6.2, $C^{1'}$-H), 1.94 (3 H, dd, *J* 6.9, 1.7, C^{4}-H_3), 1.29 (6 H, d, *J* 6.3, $C^{2'}$-H_3 and $C^{3'}$-H_3); ^{13}C NMR (126 MHz, DMSO-D_6) δ 166.2 (C^{1}), 151.5 ($C^{1''}$), 146.4 (C^{3}), 123.1 (C^{2}), 70.4 ($C^{1'}$), 21.9 ($C^{2'}$ and $C^{3'}$), 18.5 (C^{4}). LRMS: *m/z* (ES^+) 172 ($[M+H]^+$, 35%, 194 ($[M+Na]^+$, 40%), .365 ($[2M+Na]^+$, 100%); HRMS *m/z* (ESI^+) found 172.0968, $C_8H_{14}NO_3$ requires $[M+H]^+$ 172.0968. The ^{1}H NMR data are in good agreement with literature.[36]

Isopropyl (*S*)-(3-((4'-bromophenyl)amino)butanoyl)carbamate ((–)84)



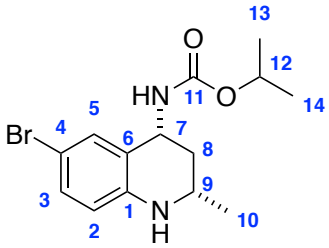
((*R*)-BINAP)ditriflatebis(acetonitrile)palladium(II) (93 mg, 0.088 mmol, 0.05 eq.) was added to a solution of isopropyl (*E*)-but-2-enoylcarbamate (300 mg, 1.75 mmol, 1.00 eq.) in dry toluene (8.0 mL). A red gummy ball formed and the colourless solution turned yellow. The mixture was stirred for 30 min prior to the addition of 4-bromoaniline (453 mg, 2.63 mmol, 1.50 eq.). After this, solids dissolved completely and the mixture was stirred for 20 h. The solvents were removed *in vacuo*. Purification of the yellow oil by silica gel chromatography, eluting with CH_2Cl_2, afforded isopropyl (*S*)-(3-((4'-bromophenyl)amino)butanoyl)carbamate (**(–)84**) (530 mg, 88%): R_f 0.59 (CH_2Cl_2); $[\alpha]_D^{25}$: –16.14 (*c* 1.0, $CHCl_3$); m.p. 128–130 °C (EtOAc) [lit.[140] 131 °C (EtOAc)]; $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3346 (m), 3262 (w), 2980 (s), 1747 (s) (C=O), 1504 (s), 1260 (m), 1103 (m), 825 (m); ^{1}H NMR (500 MHz, DMSO-D_6) δ 10.48 (1 H, s, N-H), 7.27–7.06 (2 H, m, $C^{3'}$-H), 6.61–6.34 (2 H, m, $C^{2'}$-H), 5.67 (1 H, d, *J* 8.7, N-H), 4.85 (1 H, hept, *J* 6.3, C^{6}-H), 3.86–3.65 (1 H, m, C^{3}-H), 2.68 (1 H, dd, *J* 16.0, 5.6, C^{2}-H_a), 2.45 (1 H, dd, *J* 16.0, 5.9, $C^{2'}$-H_b), 1.22 (6 H, d, *J* 6.3, C^{7}-H_3), 1.12 (3 H, d, *J* 6.4, C^{4}-H_3); ^{13}C NMR (126 MHz, DMSO-D_6) δ 171.1 (C^{1}), 150.9 (C^{5}), 146.9 ($C^{1'}$), 131.4 ($C^{3'}$). 114.3 ($C^{2'}$), 106.1 ($C^{4'}$), 68.5 (C^{6}-H), 44.5 (C^{3}), 42.7 (C^{2}), 21.6 (C^{7}-H_3), 20.1 (C^{4}-H_3); LRMS: *m/z* (ESI^+) 345 ($[M^{81}Br+H]^+$, 100%), 344 ($[M^{80}Br+H]^+$, 80%). HRMS *m/z* (ESI^+) found 343.0651, $C_{14}H_{19}{}^{79}BrN_2O_3$ requires $[M+H]^+$ 343.0652; Analytical HPLC (RP, method A) rt 11.96 mins, 97%. Chiral HPLC (heptane:IPA = 80:20, 0.5 mL/min) rt 4.70 mins, > 99%. The ^{1}H NMR data are in good agreement with literature.[36]

Isopropyl (3-((4-bromophenyl)amino)butanoyl)carbamate (84)



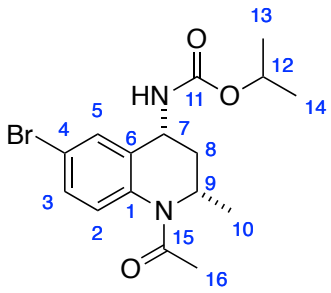
Isopropyl (*E*)-but-2-enoylcarbamate (100 mg, 0.58 mmol, 1.00 eq.), 4-bromoaniline (105 mg, 0.61 mmol, 1.05 eq.) and $MnCl_2{\cdot}6H_2O$ (12 mg, 0.06 mmol, 0.1 eq.) were dissolved in aq. MeOH (50% *v/v*, 1.0 mL) and heated at 65 °C for 2h. Additional portions of $MnCl_2{\cdot}6H_2O$ (12 mg, 0.06 mmol, 0.1 eq.) were added daily, and the mixture was stirred at 65 °C for 8 d. The volatile components were removed *in vacuo*, the residue was dissolved in H_2O (6 mL), then extracted with EtOAc (3 × 10 mL). The combined organic components were washed with H_2O (5 mL) and brine (2 × 5 mL) then dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc (3:1 to 1:1), afforded isopropyl (3-((4-bromophenyl)amino)butanoyl)carbamate (**84**) as a colourless solid (105 mg, 53%): R_f 0.36 (petroleum ether:EtOAc = 2:1); m.p. 125–126 °C (EtOAc) [lit.[140] 163 °C (EtOAc)]; LRMS *m/z* (ESI$^+$) 343 ([M^{81}Br +H]$^+$, 100%), 345 ([M^{79}Br +H]$^+$, 82%), HRMS *m/z* (ESI$^+$) found 343.0653, $C_{14}H_{19}{}^{79}BrN_2O_3$ requires $[M+H]^+$ 343.0652; Analytical HPLC (RP, method A) rt 11.23 mins, 99%. Chiral HPLC (heptane:IPA = 80:20, 0.5 mL/min) rt 6.61 mins, 52.6% and 13.73 mins, 47.4%. The spectroscopic data are identical to those obtained for **(–)-84**.

Isopropyl ((2*S*,4*R*)-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate ((–)85)



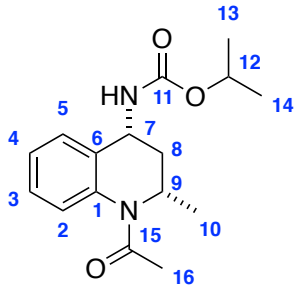
Isopropyl (*S*)-(3-((4-bromophenyl)amino)butanoyl)carbamate (**(–)84**) (25.0 mg, 0.07 mmol, 1.00 eq.) was dissolved in EtOH (1 mL, denatured with 0.5 % toluene). The reaction mixture was cooled to –10 °C, $NaBH_4$ (3.0 mg, 0.05 mmol, 0.75 eq.) was added and the reaction mixture was stirred for 30 mins. $MgCl_2{\cdot}6H_2O$ (16.3 mg, 0.08 mmol, 1.10 eq.) in 0.1 mL of H_2O was added dropwise, keeping the temperature below –10 °C. The reaction mixture was stirred at –10 to –5 °C for 1 h. The reaction mixture was quenched by the dropwise addition of citric acid (69 mg) in 1 M aq. HCl (1.5 mL), then was diluted with CH_2Cl_2 (1.5 mL). The biphasic mixture was stirred vigorously for 1 hour. The phases were separated and the aqueous phase was extracted with CH_2Cl_2 (2 × 2 mL). The organic components were washed with H_2O (5 mL) and brine (5 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification of the oil by silica gel chromatography, eluting with petroleum ether:EtOAc (1:0 to 1:1), afforded isopropyl ((2*S*,4*R*)-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (**(–)85**) as a colourless solid (16 mg, 68%): R_f 0.57 (petroleum ether:EtOAc 1:1); $[\alpha]_D^{25}$: –14.9 (*c* 1.00, $CHCl_3$); mp: 148–150 °C (EtOAc) [lit.[140] 167 °C (EtOAc)]; $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3276 (w), 2979 (m), 1682 (s) (C=O), 1537 (s), 1487 (s), 1250 (m), 1156 (m), 1060 (m), 798 (m); ^{1}H NMR (500 MHz, DMSO-D_6) δ 7.39 (1 H, d, *J* 9.1, 1H, N^{11a}-H), 7.03 (1 H, dd, *J* 8.6, 2.4, C^3-H), 6.95 (1 H, d, *J* 2.4, C^5-H), 6.42 (1 H, d, *J* 8.6, C^2-H), 5.89 (1 H, s, N^{9a}-H), 4.83 (1 H, hept, *J* 6.3, C^{12}-H), 4.74 (1 H, ddd, *J* 12.0, 9.1, 5.5, C^7-H), 3.44 (1 H, dqd, *J* 12.0, 6.2, 2.2, C^9-H), 1.91 (1 H, ddd, *J* 12.0, 5.5, 2.2, C^8-H_a), 1.44 (1 H, ddd, *J* 12.0 , 12.0, 12.0 C^8-H_b), 1.24 (3 H, d, *J* 6.3, C^{13}-H_3 or C^{14}-H_3), 1.21 (3 H, d, *J* 6.3, C^{13}-H_3 or C^{14}-H_3), 1.13 (3 H, d, *J* 6.2, C^{10}-H_3); ^{13}C NMR (126 MHz, DMSO-D_6) δ 156.30 (C^{11}), 144.7 (C^6), 129.8 (C^3), 128.5 (C^5), 124.0 (C^1), 115.2 (C^2), 105.8 (C^4), 66.9 (C^{12}), 47.2 (C^7), 46.0 (C^9), 36.6 (C^8), 22.1 (C^{13} and C^{14}), 21.7 (C^{10}); LRMS *m/z* (ES^+) 224 ($[M^{79}Br+3CH_3CN+2H]^{2+}$, 100%), 226 ($[M^{81}Br+3CH_3CN+2H]^{2+}$, 90%), 349 ($[M^{79}Br+Na]^+$, 45%), 351 ($[M^{81}Br+Na]^+$, 41%); Analytical HPLC (RP, method A) rt 12.28 mins, 96%. The ^{1}H NMR data are in good agreement with literature.[36]

Isopropyl ((2*S*,4*R*)-1-acetyl-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate ((+)80)



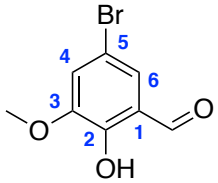
In a flame-dried flask under argon, isopropyl ((2*S*,4*R*)-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (**(–)85**) (38.2 mg, 0.12 mmol, 1.00 eq.) was dissolved in dry CH_2Cl_2 (1 mL). Pyridine (28 µL, 0.35 mmol, 3.00 eq.) in dry CH_2Cl_2 (0.1 mL) then AcCl (13 µL,0.18 mmol, 1.50 eq.) in dry CH_2Cl_2 (0.1 mL) were added. The solution was stirred at room temperature for 4 h. H_2O (5 mL) was added and the mixture was extracted with EtOAc (2 × 8 mL). The organic components were washed with sat. aq $NaHCO_3$ (5 mL), followed by H_2O (5 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification of the crude solid by silica gel chromatography, eluting with petroleum ether and EtOAc (1:0 to 1:1), afforded isopropyl ((2*S*,4*R*)-1-acetyl-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (**(+)80**) as a colourless solid (38 mg, 88%): R_f 0.17 (petroleum ether:EtOAc = 2:1); $[\alpha]_D^{25}$: +298.4 (*c* 1.00, $CHCl_3$) [lit.[140] $[\alpha]_D^{23.1}$: +251.1 (*c* 0.61, EtOH), lit.[36] $[\alpha]_D^{20}$: +281.1 (*c* 0.508, EtOH),]; m.p. 141–143 °C (EtOAc) [lit.[140] 163 °C (EtOAc)]; $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3303 (w), 2979 (w), 2361 (m), 1704 (s) (C=O), 1644 (s) (C=O), 1533 (m), 1255 (m), 964 (m); ^{1}H NMR (500 MHz, DMSO-D_6) δ 7.64 (1 H, d, *J* 8.6 Hz, N-H), 7.47 (1 H, dd, *J* 8.4, 2.2, C^3-H), 7.30 (1 H, d, *J* 8.4, C^2-H), 7.21 (1 H, d, *J* 2.2, C^5-H), 4.83 (1 H, hept, *J* 6.3, C^{12}-H), 4.63– 4.58 (1 H, m, C^9-H), 4.37 (1 H, ddd, *J* 12.2, 8.6, 4.3, C^7-H), 2.44 (1 H, ddd, *J* 12.2, 8.4, 4.3, C^8-H_a), 2.05 (3 H, s, C^{16}-H_3), 1.25 (3 H, d, *J* 6.3, C^{13}-H_3 or C^{14}-H_3), 1.22 (3 H, d, *J* 6.3, C^{13}-H_3 or C^{14}-H_3), 1.21–1.16 (1 H, m, C^8-H_b), 1.01 (3 H, d, *J* 6.3 Hz, C^{10}-H_3); ^{13}C NMR (101 MHz, DMSO-D_6) δ 168.4 (C^{15}), 155.7 (C^{11}), 139.2 (C^6), 135.5 (C^1),129.6 (C^3), 128.2 (C^2), 125.5 (C^5), 117.9 (C^4), 67.3 (C^{12}), 47.0 (C^9), 46.7 (C^7), 22.6 (C^{16}), 22.0 (C^{13} and C^{14}), 21.3 (C^{10}); LRMS *m/z* (ES^+) 369 ([M^{79}Br+H]$^+$, 90%), 371 ([M^{81}Br+H]$^+$, 100%), 391 ([M^{79}Br+Na]$^+$, 40%), 393 ([M^{81}Br+Na]$^+$, 47%); HRMS *m/z* (ESI^+) found 391.0624, $C_{16}H_{21}BrN_2O_3$ requires [M^{79}Br+Na]$^+$ 391.0628; Analytical HPLC (RP, method A) rt 12.23 mins, 99%. The ^{1}H NMR data are in good agreement with literature.[36]

Isopropyl ((2*S*,4*R*)-1-acetyl-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (126)



A solution of isopropyl ((2*S*,4*R*)-1-acetyl-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (15 mg, 0.041 mmol, 1.0 eq.) , 10% Pd/C (2 mg, 0.02 mmol, 0.4 eq.) in EtOH (2.0 mL) was stirred under H_2-atmosphere for 18 h at room temperature. The reaction mixture was filtered through Celite® and concentrated *in vacuo* to afford isopropyl ((2*S*,4*R*)-1-acetyl-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (**126**) (8 mg, 65%) as an amorphous solid: R_f 0.15 (petroleum ether:EtOAc = 2:1); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3305 (w), 2979 (s), 2931 (s), 2360 (w), 1697 (s) (C=O), 1637 (s) (C=O), 1536 (m), 1340 (s), 1256 (w), 1112 (m), 762 (m); ^{1}H NMR (400 MHz, CD_2Cl_2) δ 7.33–7.20 (3 H, m, Ar-CH), 7.14 (1 H, br d, *J* 7.5, Ar-CH), 4.92 (1 H, hept, *J* 6.2, C^{12}-H), 4.95–4.82 (1 H, m, N-H), 4.82–4.75 (1 H, m, C^9-H), 4.60 (1 H, ddd, *J* 12.8, 8.7, 4.5, C^8-H_a), 2.56 (1 H, ddd, *J* 13.0, 8.7, 4.2, C^7-H), 2.07 (3 H, s, C^{16}-H_3), 1.33–1.22 (6 H, m, C^{13}-H_3 and C^{14}-H_3), 1.22–1.12 (1 H, m, C^8-H_b) 1.11 (3 H, d, *J* 6.4, C^{10}-H_3); ^{13}C NMR (101 MHz, CD_2Cl_2) δ 169.2 (C^{15}), 155.7 (C^{11}), 136.7 (C^1), 136.4 (C^6), 127.0 (C^3 or C^5), 126.1 (C^3 or C^5), 125.7 (C^2), 122.4 (C^4), 68.4 (C^{12}), 47.1 (C^7), 46.9 (C^9), 41.0 (C^8), 22.7 (C^{16}), 21.9 (C^{13} and C^{14}), 21.0 (C^{10}); LRMS *m/z* (ESI$^+$) 313 ([M+Na]$^+$, 100%); HRMS *m/z* (ESI$^+$) found 313.1523, $C_{16}H_{22}N_2O_3$ requires [M+Na]$^+$ 313.1523.

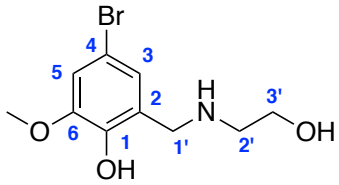
5-Bromo-2-hydroxy-3-methoxybenzaldehyde (87)



2-Hydroxy-3-methoxybenzaldehyde (**86**) (1.00 g, 6.57 mol, 1.00 eq.) and KBr (1.56 g, 13.2 mol, 2.00 eq.) were dissolved in 80% aq. acetic acid (30 mL). Br_2 (0.36 mL, 6.93 mol, 1.05 eq.) in glacial acetic acid (9 mL) was added dropwise with a Pasteur pipette over 2 h. Precipitation was observed at the end of the addition. The reaction mixture was heated

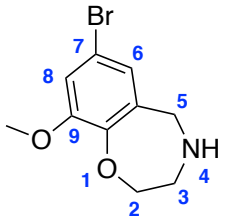
at 115 °C for 15 mins, and stirred at room temperature for 16 h. The precipitate was collected by filtration and crystallised from hot aq. EtOH (50% *v/v*, 100 mL) to afford 5-bromo-2-hydroxy-3-methoxybenzaldehyde (**87**) (55.7 mg, 37%) as a yellow needles: R_f 0.63 (petroleum ether:EtOAc = 5:1); m.p. 120–122 °C (H_2O:EtOH = 1:1) [lit.[141] 119–121 °C (H_2O:EtOH = 3:2)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.03 (1 H, s, C(O)H), 9.88 (1 H, s, C^2-OH), 7.34 (1 H, d, *J* 2.2, C^6-H), 7.20 (1 H, d, *J* 2.2, C^4-H), 3.94 (3 H, s, CH_3); LRMS *m/z* (ES^-) 229 ($[M^{81}Br–H]^-$, 80 %), 231 ($[M^{79}Br–H]^-$, 100 %). The data are in good agreement with literature.[141]

4-Bromo-2-(((2-hydroxyethyl)amino)methyl)-6-methoxyphenol (88)



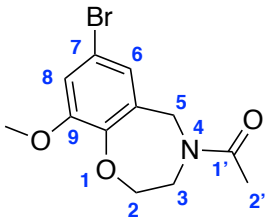
5-Bromo-2-hydroxy-3-methoxybenzaldehyde (**87**) (250 mg, 1.08 mmol, 1.00 eq.) was dissolved in dried 1,2-dichloroethane (2 mL) in a flame-dried flask under argon. Ethanolamine (80 µL, 1.30 mmol, 1.20 eq.) was added and the imine intermediate precipitated. $NaBH(OAc)_3$ (182 mg, 0.85 mmol, 0.80 eq.) was added, and the mixture was stirred for 1 h. No reduction of the imine was observed. MeOH was added and the solvents were removed *in vacuo*. The residue was dissolved in MeOH (10 mL), and $NaBH_4$ (61.3 mg, 1.62 mmol, 1.5 eq.) was added at 0 °C. The mixture was left to warm to room temperature and stirred for 4 h. 1 M aq. HCl was added until pH 7–8 was obtained. The mixture was extracted with CH_2Cl_2 (3 × 10 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (2 × 10 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*, to afford 4-bromo-2-(((2-hydroxyethyl)amino)methyl)-6-methoxyphenol (**88**) (283.3 mg, 95%) as a colourless solid: R_f 0.31 (CH_2Cl_2:EtOH:NH_4OH = 50:8:1); m.p. 148–150 °C (CH_2Cl_2) [lit.[142] 153–154°C (MeOH)]; $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3160 (w), 2933 (w), 2829 (w), 1480 (m), 1443 (m), 1212 (m), 1084 (m), 966 (m), 761 (m); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.89 (1 H, d, *J* 2.0, C^3-*H*), 6.77 (1 H, d, *J* 2.0, C^5-*H*), 3.99 (2 H, s, $C^{1'}$-H_2), 3.85 (3 H, s, O-C*H*$_3$), 3.78 (2 H, d, *J* 4.7, $C^{3'}$-H_2), 2.80 (2 H, d, *J* 4.7, $C^{2'}$-H_2), 1.25 (1 H, s, O*H*); LRMS *m/z* (ES^-): 274 ($[M^{79}Br–H]^-$, 100%), 276 ($[M^{81}Br–H]^-$, 90%). The data are in good agreement with the literature.[142]

7-Bromo-9-methoxy-2,3,4,5-tetrahydrobenzo[*f*][1,4]oxazepane (90)



5-Bromo-2-hydroxy-3-methoxybenzaldehyde (**88**) (360 mg, 1.29 mmol, 1.00 eq.) was dissolved in dry 1,2-dichloroethane (8.0 mL) in a flame dried flask with 3 Å molecular sieves. 2-Chloroethyl amine hydrochloride (180 mg, 1.55 mmol, 1.20 eq.) was added and the reaction mixture was stirred at room temperature for 3 h. Et_3N (360 µL, 2.58 mmol, 2.00 eq.) was added and the reaction mixture was added for 1 h. $NaBH(OAc)_3$ (920 mg, 4.32 mmol, 4.00 eq.) was added, and the mixture was stirred at room temperature for 22 h. The mixture was diluted with aq. sat. $NaHCO_3$ (50 mL) and extracted with EtOAc (2 × 100 mL). The combined organic components were washed with H_2O (50 mL) and sat. aq. $NaHCO_3$ (50 mL), dried over anhydrous $MgSO_4$, filtered, and concentrated *in vacuo*, giving 7-bromo-9-methoxy-2,3,4,5-tetrahydrobenzo[*f*][1,4]oxazepane (**90**) (284 mg, 75%) as a purple solid: R_f 0.48 (petroleum ether:EtOAc = 3:1); m.p. 156–158 °C (EtOAc) [lit.[142] 119–120 °C (CH_2Cl_2)]; $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3160 (w), 2933 (w), 2829 (w), 1480 (s), 1245 (m), 1227 (s), 1084 (m); ^1H (500 MHz, $CDCl_3$) δ 7.00 (1 H, d, *J* 2.2, C^6-H), 6.94 (1 H, d, *J* 2.2, C^8-H), 6.19 (1 H, s, N-H), 4.43 (2 H, s, C^5-H_2), 4.37–4.26 (2 H, m, C^2-H_2), 3.89 (3 H, s, O-CH_3), 3.57–3.46 (2 H, m, C^3-H_2). LRMS *m/z* (ES^+) 302 ([M^{79}Br+2Na–H]$^+$, 100%), 304 ([M^{81}Br+2Na–H]$^+$, 90%); Analytical HPLC (RP, method A) rt 9.87 mins, purity 98%. The ^1H NMR data are in good agreement with the literature.[142]

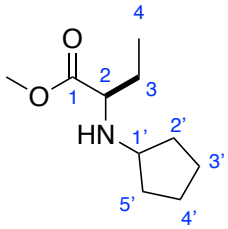
1-(7-Bromo-9-methoxy-2,3-dihydrobenzo[*f*][1,4]oxazepin-4(5*H*)-yl)ethan-1-one (91)



Ac_2O (36 µL, 0.38 mmol, 2.0 eq.) was added to a solution of 7-bromo-9-methoxy-2,3,4,5-tetrahydrobenzo[*f*][1,4]oxazepane (**90**) (50 mg, 0.19 mmol, 1.0 eq.) and K_2CO_3 (105 mg, 0.76 mmol, 4.0 eq.) in CH_2Cl_2 (2.0 mL). The reaction mixture was stirred at room temperature for 4.5 h, then diluted with H_2O (10 mL) and extracted with EtOAc (2 ×

20 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (20 mL), H_2O (20 mL), dried over $MgSO_4$, filtered, and concentrated in vacuo. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (95:5 to 3:7), afforded 1-(7-bromo-9-methoxy-2,3-dihydrobenzo[*f*][1,4]oxazepin-4(5*H*)-yl)ethan-1-one (**91**) (35 mg, 60%) as a colourless oil: R_f 0.41 (petroleum ether:EtOAc = 2:1); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2980 (w), 1744 (s) (C=O), 1482 (m), 1169 (s), 1035 (m), 851 (m), 760 (m); ^1H (500 MHz, $CDCl_3$) δ 7.07 (1 H, d, *J* 2.2, C^6-H), 7.04 (1 H, d, *J* 2.2, C^8-H), 4.34 (2 H, s, C^5-H_2), 4.30–4.19 (2 H, m, C^3-H_2), 3.82 (3 H, s, O-CH_3), 3.41–3.29 (2 H, m, C^2-H_2), 2.32 (3 H, s, $C^{2'}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.1 ($C^{1'}$), 158.1 (C^{9a}), 152.4 (C^9), 138.1 (C^{6a}), 124.6 (C^6), 119.5, (C^7), 116.0 (C^8), 62.0 (C^2), 56.5 (O-CH_3), 43.8 (C^5), 43.6 (C^3), 20.6 ($C^{2'}$); LRMS *m/z* (ES^+) 301 ([$M^{79}Br$+H]$^+$, 100%), 303 ([$M^{81}Br$+H]$^+$, 85%); Analytical HPLC (RP, method A) rt 8.33 mins, purity > 99%. The data are in good agreement with the literature.[142]

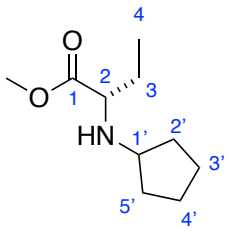
Methyl (*R*)-2-(cyclopentylamino)butanoate ((*R*)-94)



Methyl *(R)*-2-aminobutanoate hydrochloride (1.25 g, 8.14 mol, 1.2 eq.) and $Ga(OTf)_2$ (175 mg, 0.339 mol, 0.05 eq.) were dissolved in dry 1,2-dichloroethane (20 mL) in the presence of 3 Å molecular sieves. Et_3N (950 µL, 6.78 mol, 1.0 eq.) and cyclopentanone (600 µL, 6.78 mol, 1.0 eq.) were added and the reaction mixture was stirred at room temperature for 2 h. $NaBH(OAc)_3$ (2.87 g, 13.6 mol, 2.0 eq.) was added and the reaction mixture was stirred at room temperature for 18 h. The reaction mixture was basified with Et_3N (3 mL), then filtered through a pad of Celite®, and washed with CH_2Cl_2 (250 mL, with 1% Et_3N). Sat. aq. K_2CO_3 (250 mL) was added to the filtrate and the layers were separated. The aqueous layer was further extracted with CH_2Cl_2 (100 mL). The organic components were combined, washed with sat. aq. K_2CO_3 (2 × 150 mL), dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. MeOH (200 mL) was added and removed *in vacuo* to remove the excess borane species. The mixture was suspended in $CHCl_3$ (50 mL) and the

insoluble impurities were removed by filtration to afford methyl (*R*)-2-(cyclopentylamino)butanoate (**(*R*)-94**) (977 mg, 78%) as light yellow oil: R_f 0.08 (petroleum ether:EtOAc = 20:1) (not UV active, monitored with ninhydrin stain); $[\alpha]_D^{25}$: +25.3 (*c* 1.0, MeOH) [lit.[94] $[\alpha]_D^{21.4}$: +11.8 (*c* 1.0, MeOH)]; $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2951 (br), 1737 (s, C=O), 1460 (w), 1098 (w), 801; ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.72 (3 H, s, O-CH_3), 3.21 (1 H, dd, *J* 6.6, 6.6, C^2-H), 2.97 (1 H, dddd, *J* 6.7, 6.7, 6.6, 6.6, C$^{1'}$-H), 2.07 (1 H, br s, N-H), 1.94–1.71 (2 H, m, C$^{2'}$-H$_a$, C$^{5'}$-H$_a$), 1.71–1.58 (4 H, m, C^3-H$_2$, C$^{3'}$-H$_a$, C$^{4'}$-H$_a$), 1.57–1.39 (2 H, m, C$^{2'}$-H$_b$, C$^{5'}$-H$_b$), 1.36–1.22 (2 H, m, C$^{3'}$-H$_b$, C$^{4'}$-H$_b$), 0.85 (3 H, dd, *J* 7.5, 7.5, C^4-H$_3$); LRMS *m/z* (ES$^+$) 186 ([M^{35}Cl+H]$^+$, 100%). The data are in good agreement with the literature.[94,98]

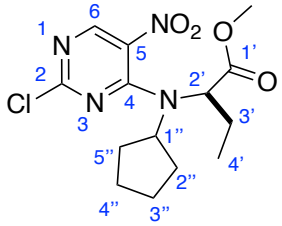
Methyl (*S*)-2-(cyclopentylamino)butanoate ((*S*)-94)



Methyl *(S)*-2-aminobutanoate hydrochloride (1.25 g, 8.14 mol, 1.2 eq.) and $Ga(OTf)_2$ (175 mg, 0.339 mol, 0.05 eq.) were dissolved in dry 1,2-DCE (25 mL) in the presence of 3 Å molecular sieves. Et_3N (950 µL, 6.78 mol, 1.0 eq.) and cyclopentanone (600 µL, 6.78 mol, 1.0 eq.) were added and the reaction mixture was stirred at room temperature for 2 h. $NaHB(OAc)_3$ (2.87 g, 13.6 mol, 2.0 eq.) was added and the reaction mixture was stirred at room temperature for 18 h. The reaction mixture was basified with Et_3N (3 mL) then filtered through a pad of Celite®, and washed with CH_2Cl_2 (200 mL, with 1% Et_3N) and CH_2Cl_2 (100 mL). Sat. aq. K_2CO_3 (200 mL) was added to the filtrate and the layers were separated. The aqueous layer was further extracted with CH_2Cl_2 (100 mL). The organic components were combined, washed with sat. aq. K_2CO_3 (2 × 150 mL), dried over anhydrous Na_2SO_4, filtered, and concentrated *in vacuo*. MeOH (200 mL) was added and removed *in vacuo* to remove the excess borane species. The mixture was suspended in $CHCl_3$ (50 mL) and the insoluble impurities were removed by filtration, washing with $CHCl_3$ (3 × 50 mL). The solvents was removed *in* vacuo to afford methyl (*S*)-2-(cyclopentylamino)butanoate (**(*S*)-94**) (1.16 g, 92%) as light yellow oil: R_f 0.08 (petroleum

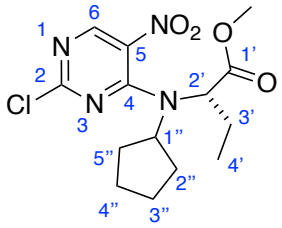
ether:EtOAc = 20:1) (not UV active, monitored with ninhydrin stain); $[\alpha]_D^{25}$: –14.2 (*c* 1.0, MeOH) [lit.[94] $[\alpha]_D^{21.4}$: –12.0 (*c* 1.0, MeOH)]; LRMS *m/z* (ES^+) 186 ($[M^{35}Cl+H]^+$, 100%). The spectrometry data are identical to enantiomer **(*R*)-94**, are in good agreement with the literature.[94,98]

Methyl (*R*)-2-((2-chloro-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate ((*R*)-95)



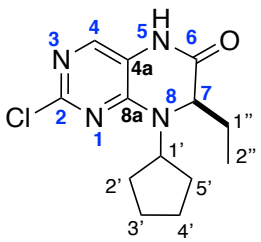
Methyl (*R*)-2-(cyclopentylamino)butanoate (124 mg, 0.667 mmol, 1.0 eq.) was dissolved in cyclohexane (7 mL). $NaHCO_3$ (224 mg, 2.67 mmol, 4.0 eq.) was added, then the reaction mixture was stirred at room temperature for 30 mins. 2,4-Dichloro-5-nitropyrimidine (150 mg, 0.773 mmol, 1.1 eq.) was added and the reaction mixture was stirred at 60 °C for 3 days. H_2O (30 mL) was added, and the mixture was extracted with EtOAc (2 × 40 mL). The organic components were combined, washed with H_2O (20 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 9:1, slow gradient; then 9:1 to 5:1, fast gradient), afforded methyl (*R*)-2-((2-chloro-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate (**(*R*)-95**) (117 mg, 51%) as a light orange solid: R_f 0.75 (petroleum ether:EtOAc = 1:1; $[\alpha]_D^{25}$: +292 (*c* 1.0, $CHCl_3$) [lit.[94] $[\alpha]_D^{21.4}$: +228.3 (*c* 1.0, MeOH)]; m.p. 96–98 °C (EtOAc) [lit.[98] 111 °C]; $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2971 (w), 1740 (m), 1579 (s), 1347 (m), 1201 (m); ^1H NMR (400 MHz, $CDCl_3$) 8.66 (1 H, s, C^6-H), 3.75 (4 H, s, OCH_3 and $C^{2'}$-H), 3.57 (1 H, dd, *J* 8.2, 8.4, $C^{1''}$-H), 2.40 (1 H, dqd, *J* 14.9, 7.5, 6.1, $C^{3'}$-H_a), 2.26–2.14 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 2.06 (1 H, ddd, *J* 15.5, 14.4, 7.7, $C^{3'}$-H_b), 1.97–1.85 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 1.88–1.43 (6 H, m, $C^{2''}$-H_b, $C^{5''}$-H_b, $C^{3''}$-H_2 and $C^{4''}$-H_2), 1.04 (3 H, dd, *J* 7.5, 7.5, $C^{4'}$-H_3); LRMS *m/z* (ES^+) 343 ($[M^{35}Cl+H]^+$, 100%). The data are in good agreement with the literature.[94,98]

Methyl (*S*)-2-((2-chloro-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate ((*S*)-95)



Methyl (*S*)-2-(cyclopentylamino)butanoate (2.1 g, 10 mmol, 1.0 eq.) was dissolved in cyclohexane (65 mL). $NaHCO_3$ (8.2 g, 97 mmol, 9.5 eq.) was added, then the reaction mixture was stirred at room temperature for 30 mins. 2,4-Dichloro-5-nitropyrimidine (2.2 g,11 mmol, 1.1 eq.) was added and the reaction mixture was stirred at 60 °C for 3 days. H_2O (200 mL) was added, and the mixture was extracted with EtOAc (2 × 200 mL). The organic components were combined, washed with H_2O (20 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography (Combi*flash*Rf+, 80 g column), eluting with cyclohexane:EtOAc (1:0 to 5:1), afforded methyl (*S*)-2-((2-chloro-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate (**(*S*)-95**) (2.8 g, 74%) as a light orange solid: R_f 0.75 (petroleum ether:EtOAc = 1:1; lit.[94] $[\alpha]_D^{21.4}$: –229.9 (*c* 1.0, MeOH); m.p. 96–98 °C (EtOAc). LCMS (formic acid) rt 1.28 min (94%), *m/z* (ES^+) 343 ($[M^{35}Cl+H]^+$, 100%). The spectroscopic data are identical to the enantiomer **(*R*)-95** and in good agreement with the literature.[94,98]

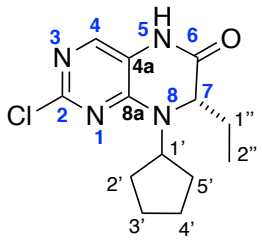
(*R*)-2-Chloro-8-cyclopentyl-7-ethyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-96)



Methyl (*R*)-2-((2-chloro-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate (**(*R*)-95**) (60 mg, 0.18 mmol, 1.0 eq.) was dissolved in glacial AcOH (0.7 mL) and heated to 70 °C. Fe (50 mg, 0.87 mmol, 2.0 eq.) was added in 2 portions, and the solution was stirred at 70 °C for 2 h. The temperature was increased to 100 °C and the reaction mixture was stirred for a further 3 h. The mixture was allowed to cool to room temperature, then was filtered over Celite®, washing with MeOH (4 × 50 mL). The filtrate was concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system,

eluting with petroleum ether:EtOAc (20:1 to 10:7), afforded (*R*)-2-chloro-8-cyclopentyl-7-ethyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-96**) (31 mg, 62%) as a colourless solid: R_f 0.45 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: –111 (*c* 1.0, MeOH) [lit.[94] $[\alpha]_D^{21.4}$: –96.2 (*c* 1.0, MeOH)]; $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2968 (br), 1692 (s, C=O), 1603, 1363, 1193, 767 (w); m.p. 168–170°C (EtOAc) [lit.[98] 173 °C]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.42 (1 H, s, NH), 7.61 (1 H, s, C^4-H), 4.33 (1 H, dddd, *J* 8.2, 8.2, 8.2, 8.2, C$^{1'}$-H), 4.20 (1 H, dd, *J* 7.6, 3.6, C^7-H), 2.12–2.03 (1 H, m, C$^{2'}$-Ha or C$^{5'}$-Ha), 2.00–1.70 (7 H, m, C$^{2'}$-H$_a$ or C$^{5'}$-H$_a$, C$^{2'}$-H$_b$, C$^{5'}$-H$_b$, C$^{3'}$-H$_a$, C$^{4'}$-H$_a$ and C$^{1'''}$-H$_2$), 1.71–1.60 (2 H, m, C$^{3'}$-H$_b$ or C$^{4'}$-H$_b$), 0.93 (3 H, dd, *J* 7.5, 7.5, C$^{2''}$-H$_3$). LRMS *m/z* (ES$^+$) 279 ([M^{35}Cl–H]$^-$, 100%). Chiral HPLC (AD-H, Hexane:IPA = 97:3, 1.8 mL/min, 20 mins) major rt 11.24 mins, >99%, minor peak not seen. The data are in good agreement with the literature.[94,98]

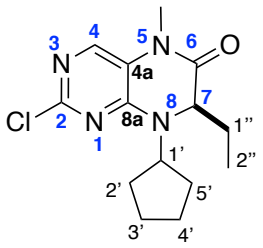
(*S*)-2-Chloro-8-cyclopentyl-7-ethyl-7,8-dihydropteridin-6(5*H*)-one ((*S*)-96)



Methyl (*S*)-2-((2-chloro-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate (**(*S*)-95**) (2.26 g, 6.59 mmol, 1.0 eq.) was dissolved in glacial AcOH (35 mL) and heated to 70 °C. Fe (1.84 g, 33.0 mmol, 5.0 eq.) was added, and the solution was stirred at 70 °C for 1 h. The temperature was increased to 100 °C and the reaction mixture was stirred for a further 2 h. The mixture was allowed to cool to room temperature, then was filtered over Celite®, washing with MeOH (3 × 150 mL). The filtrate was concentrated *in vacuo*. Purification by silica gel chromatography (Combi*flash*Rf+, 120 g column), eluting with cyclohexane:EtOAc (1:0 to 0:1), afforded (*S*)-2-chloro-8-cyclopentyl-7-ethyl-7,8-dihydropteridin-6(5*H*)-one (**(*S*)-96**) (1.4 g, 72%) as a colourless solid: R_f 0.45 (petroleum ether:EtOAc = 1:1); lit.[94] $[\alpha]_D^{21.4}$: +95.3 (*c* 1.0, MeOH); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2968 (br), 1692 (s, C=O), 1603, 1363, 1193, 767 (w); m.p. 168–170°C (EtOAc); LCMS (formic acid) rt 1.02 min (>99%), *m/z* (ES$^+$) 281 ([M^{35}Cl+H]$^+$, 100%); Chiral HPLC (AD-H, Hexane:IPA = 97:3, 1.8 mL/min, 20 mins) major rt 9.26 mins, >99%, minor peak not seen. The

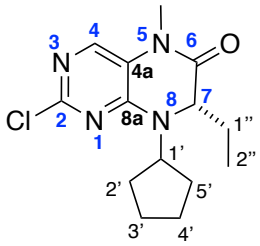
spectroscopic data are identical to the enantiomer **(*R*)-96** and in good agreement with the literature.[94,98]

(*R*)-2-Chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-92)



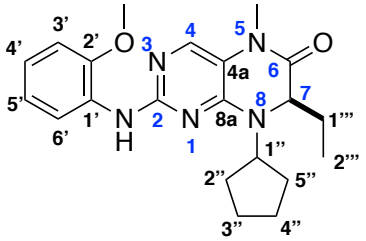
MeI (10 µL, 0.14 mmol, 2.0 eq.) was added to a solution of (*R*)-2-chloro-8-cyclopentyl-7-ethyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-96**) (20 mg, 0.071 mmol, 1.0 eq.) in dry DMF (0.5 mL). The reaction mixture was cooled to –10 °C using a NaCl and ice bath, then NaH (3.5 mg, 0.14 mmol, 2.0 eq.) was added. The reaction mixture was stirred at –10 °C for 1 h, then was warmed to room temperature and stirred for 1 h. The mixture was quenched with the addition of crushed ice, then was diluted with H_2O (20 mL) and extracted with EtOAc (2 × 35 mL). The combined organic components were washed with aq. LiCl (10% *w/v*, 6 × 25 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc (1:0 to 0:1), afforded (*R*)-2-chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-92**) (21 mg, 98%) as a colourless solid: R_f 0.56 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: –80.0 (*c* 1.0, MeOH) [lit.[94] $[\alpha]_D^{21.4}$: –88.3 (*c* 1.0, MeOH)]; $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2968 (w), 1682 (s) (C=O), 1578 (m), 1409 (m), 1185 (m), 1058 (m); m.p. 82–84 °C (EtOAc) [lit.[98] 94 °C (EtOAc)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.66 (1 H, s, C^4-H), 4.41–4.28 (1 H, m, C$^{1'}$-H), 4.24 (1 H, dd, *J* 7.5, 3.6, C^7-H), 3.32 (3 H, s, N-CH_3), 2.13–2.02 (1 H, m, C$^{2'}$-H_a or C$^{5'}$-H_a), 2.02–1.80 (4 H, m, C$^{2'}$-H_a or C$^{5'}$-H_a, C$^{1''}$-H_a, and C$^{3'}$-H_2 or C$^{4'}$-H_2), 1.80–1.60 (5 H, m, C$^{2'}$-H_b, C$^{5'}$-H_b, C$^{1''}$-H_b, and C$^{3'}$-H_2 or C$^{4'}$-H_2), 0.86 (3 H, dd, *J* 7.5, 7.5, C$^{2''}$-H_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.9 (C^6), 154.1 (C^{8a}), 152.5 (C^2), 138.1 (C^4), 121.4 (C^{4a}), 60.9 (C^7), 59.6 (C$^{1'}$), 29.6 (C$^{2'}$ or C$^{5'}$), 29.4 (C$^{2'}$ or C$^{5'}$), 28.4 (NCH_3), 27.5 (C$^{1''}$), 24.2 (C$^{3'}$ or C$^{4'}$), 23.9 (C$^{3'}$ or C$^{4'}$), 9.2 (C$^{2''}$); LRMS *m/z* (ES$^+$) 295 ([M^{35}Cl+H]$^+$, 100%); Analytical HPLC (RP, method A) rt 10.17 mins, purity > 99%. Chiral HPLC (AD-H, Hexane:IPA = 90:10, 1.0 mL/min, 20 mins) minor rt 7.74 mins, 2.5%, major rt 10.18 mins, 97.5%. The ^{1}H spectroscopic data are in good agreement with the literature.[94,98]

(*S*)-2-Chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*S*)-92)



MeI (540 µL, 8.55 mmol, 2.0 eq.) was added to a solution of (*S*)-2-chloro-8-cyclopentyl-7-ethyl-7,8-dihydropteridin-6(5*H*)-one (**(*S*)-96**) (1.20 g, 4.27 mmol, 1.0 eq.) in dry DMF (15 mL). The reaction mixture was cooled to 0 °C, then NaH (350 mg, 8.55 mmol, 2.0 eq.) was added. The reaction mixture was stirred at 0 °C for 30 mins, then was warmed to room temperature and stirred for 1 h. The mixture was quenched with the addition of crushed ice, then was diluted with brine (150 mL) and extracted with EtOAc (2 × 150 mL). The combined organic components were washed with sat. aq. LiCl (5 × 100 mL), dried over an hydrophobic frit, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with cyclohexane:EtOAc (1:0 to 0:1), afforded (*S*)-2-chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*S*)-92**) (1.3 g, 90%) as a colourless solid: R_f 0.56 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: +28.9 (*c* 1.0, MeOH) [lit.[94] $[\alpha]_D^{21.4}$: +93.5 (*c* 1.0, MeOH)]; m.p. 82–84 °C (EtOAc); LCMS (formic acid) rt 1.02 min (99%), *m/z* (ES$^+$) 295 ([M^{35}Cl+H]$^+$, 100%); Chiral HPLC (AD-H, Hexane:IPA = 90:10, 1.0 mL/min, 20 mins) major rt 7.27 mins, 92.2%, minor rt 9.74 mins, 6.8%. The spectroscopic data are identical to the enantiomer **(*R*)-92** and in good agreement with the literature.[94,98]

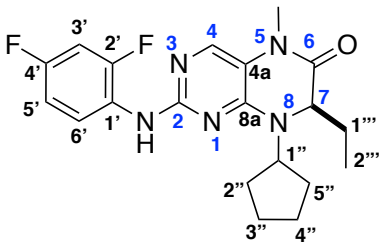
(*R*)-8-Cyclopentyl-7-ethyl-2-((2'-methoxyphenyl)amino)-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-156)



XPhos (5.0 mg, 0.010 mmol, 0.15 eq.) and $Pd_2(dba)_3$ (3.0 mg, 0.0033 mmol, 0.05 eq.) were dried *in vacuo* for 5 mins, then placed under an atmosphere of argon and dissolved in degassed ^{t}BuOH (300 µL). In a microwave vial, (*R*)-2-chloro-8-cyclopentyl-7-ethyl-5-

methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-92**) (20 mg, 0.068 mmol, 1.0 eq.), 2-methoxyaniline (8.3 mg, 0.068 mmol, 1.0 eq.) and K_2CO_3 (38 mg, 0.27 mmol, 4.0 eq.) were dried under vacuum for 5 mins, then placed under an atmosphere of Argon and dissolved in degassed ^{t}BuOH (200 µL). The palladium complex solution was added to the microwave vial, which was sealed, then the reaction mixture was stirred at 100 °C for 20 h. The reaction mixture was filtered through Celite®, washing through with CH_2Cl_2 (6 × 30 mL) then concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc (1:0 to 1:5), afforded (*R*)-8-cyclopentyl-7-ethyl-2-((2'-methoxyphenyl)amino)-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-156**) (8 mg, 31%) as a light brown residue: R_f 0.12 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: –479.6 (*c* 0.25, $CHCl_3$); υ_{max} (thin film)/cm^{-1} 3428 (w), 2961 (w), 1672 (s) (C=O) , 1500 (w), 1435 (s), 1244 (m), 1027 (m), 745 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.43 (1 H, dd, *J* 7.6, 2.0, C$^{6'}$-H), 7.66 (1 H, s, C^4-H), 7.44 (1 H, br s, N-H), 6.99–6.90 (2 H, m, C$^{4'}$-H and C$^{5'}$-H), 6.88 (1 H, dd, *J* 7.7, 1.9, C$^{3'}$-H), 4.56–4.43 (1 H, m, C$^{1''}$-H), 4.21 (1 H, dd, *J* 7.9, 3.7, C^7-H), 3.90 (3 H, s, OCH_3), 3.32 (3 H, s, NCH_3), 2.19–2.07 (1 H, m, C$^{2''}$-H$_a$ or C$^{5''}$-H$_a$), 2.07–1.90 (1 H, m, C$^{2''}$-H$_a$ or C$^{5''}$-H$_a$), 1.90–1.51 (8 H, m, C$^{2''}$-H$_b$, C$^{5''}$-H$_b$, C$^{1'''}$-H$_2$, C$^{3''}$-H$_2$, C$^{4''}$-H$_2$), 0.87 (3 H, dd, *J* 7.5, 7.5, C$^{2'''}$-H$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ 163.8 (C^6), 155.7 (C^{8a}), 152.3 (C^2), 147.6 (C$^{2'}$), 138.1 (C^4), 130.0 (C$^{1'}$), 120.9 (C$^{4'}$ or C$^{5'}$), 120.7 (C$^{4'}$ or C$^{5'}$), 117.9 (C$^{6'}$), 115.8 (C^{4a}), 109.9 (C$^{3'}$), 60.0 (C^7), 58.4 (C$^{1''}$), 55.7 (OCH_3), 29.7 (C$^{2''}$ or C$^{5''}$), 29.2 (C$^{2''}$ or C$^{5''}$), 28.2 (NCH_3), 27.1 (C$^{1'''}$), 23.5 (C$^{3''}$ or C$^{4''}$), 23.1 (C$^{3''}$ or C$^{4''}$), 9.2 (C$^{2'''}$-H$_3$); LRMS *m/z* (ES$^+$) 382 ([M+H]$^+$, 100%); HRMS *m/z* (ESI$^+$) found 382.2239, $C_{21}H_{27}N_5O_2$ requires [M+H]$^+$ 382.2238; Analytical HPLC (RP, method B) rt 8.63 mins, purity 98%.

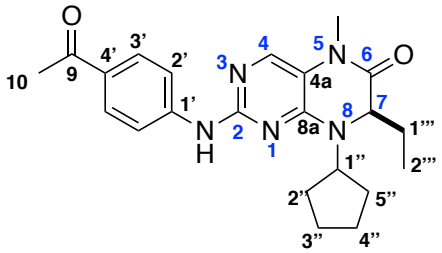
(7*R*)-8-Cyclopentyl-2-((2',4'-difluorophenyl)amino)-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-130)



(R)-2-Chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-92**) (20 mg, 0.068 mmol, 1.0 eq.) was suspended in H_2O (0.8 mL) and EtOH (0.2 mL). 2-4-

difluoronaniline (12 µL, 0.14 mmol, 2.0 eq.) and 20 µL conc. HCl were added. The resulting solution was stirred under reflux conditions for 18 h, then sat. aq. $NaHCO_3$ (20 mL) was added. The mixture was extracted with CH_2Cl_2 (2 × 40 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (20 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether and EtOAc (1:0 to 0:1), afforded (7*R*)-8-cyclopentyl-2-((2',4'-difluorophenyl)amino)-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-130**) as an amorphous brown solid (26 mg, > 99%): R_f 0.79 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: –57.8 (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3053 (w), 1673 (s, C=O) , 1974 (w), 1506 (m), 1428 (s), 1279 (m), 845 (m), 778 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.37–8.24 (1 H, m, $C^{6'}$-H), 7.63 (1 H, s, C^4-H), 6.93–6.82 (2 H, m, $C^{3'}$-H and $C^{5'}$-H), 6.82–6.71 (1 H, m, N-H), 4.47–4.32 (1 H, m, $C^{1''}$-H), 4.20 (1 H, dd, *J* 7.8, 3.7, C^7-H), 3.31 (3 H, s, NCH_3), 2.13–2.02 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 2.02–1.89 (1 H, m, and $C^{2''}$-H_a or $C^{5''}$-H_a), 1.90–1.74 (4 H, m, $C^{1'''}$-H_a, $C^{2''}$-H_b or $C^{5''}$-H_b, and $C^{3''}$-H_2 or $C^{4''}$-H_2), 1.74–1.54 (4 H, m, $C^{1'''}$-H_b, $C^{2''}$-H_b or $C^{5''}$-H_b, and $C^{3''}$-H_2 or $C^{4''}$-H_2), 0.87 (3 H, dd, *J* 7.5, 7.5, $C^{2'''}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 163.7 (C^6), 157.0 (dd, J_{CF} 244.3, 11.6, $C^{2'}$-F), 152.4 (dd, J_{CF} 244.3, 11.6, $C^{4'}$-F), 155.3 (C^2 or C^{8a}), 152.3 (C^2 or C^{8a}), 137.8 (C^4), 125.1 (dd, J_{CF} 3.2, 3.4, $C^{6'}$), 121.3 (dd, J_{CF} 8.6, 2.4, $C^{1'}$), 116.4 (C^{4a}), 110.5 (dd, J_{CF} 21.4, 3.6, $C^{5'}$), 103.4 (dd, J_{CF} 26.6, 26.6, $C^{3'}$), 60.2 (C^7), 58.6 ($C^{1''}$), 29.6 ($C^{2''}$ or $C^{5''}$), 29.2 ($C^{2''}$ or $C^{5''}$), 28.2 (N-CH_3), 27.2 ($C^{1'''}$), 23.5 ($C^{3''}$ or $C^{4''}$), 23.2 ($C^{3''}$ or $C^{4''}$), 9.2 ($C^{2'''}$); ^{19}F NMR (471 MHz, $CDCl_3$) δ – 126.8 (1 F, br d, J_{CF} 2.9), – 119.0 (1 F, br s); LRMS *m/z* (ES^+) 388 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) found 388.19406 (100%), $C_{20}H_{23}ON_5F_2$ requires $[M+H]^+$ 388.19434; Analytical HPLC (RP, method B) rt 7.53 mins, purity 99%.

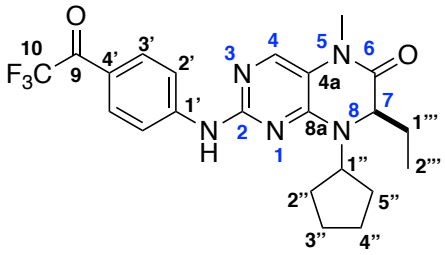
(7*R*)-2-((4'-Acetylphenyl)amino)-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-157)



(*R*)-2-Chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-92**) (60 mg, 0.21 mmol, 1.0 eq.) was suspended in H_2O (2.4 mL) and EtOH (0.6 mL). 4-

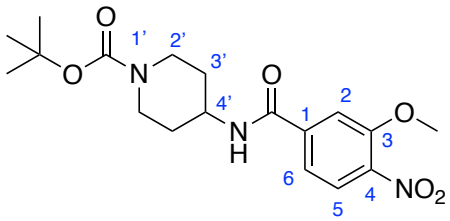
Aminoacetophenone (56 mg, 0.41 mmol, 2.0 eq.) and conc. HCl (80 µL) were added. The resulting solution was stirred under reflux for 20 h. Additional 4-aminoacetophenone (56 mg, 0.41 mmol, 2.0 eq.) was added, and the reaction mixture was stirred under reflux for 24 h. The precipitated mixture was diluted with H_2O:EtOH (5:1, 5 mL) and was extracted with EtOAc (50 mL). The pH of the aqueous layer was adjusted to 5–6 using sat. K_2CO_3 and 1 M HCl, then further extracted with EtOAc (2 × 100 mL). The combined organic components were washed with H_2O (2 × 50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether and EtOAc (slow gradient from 1:0 to 1:1, then fast from 1:1 to 0:1), afforded (*7R*)-2-((4-acetylphenyl)amino)-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-157**) (51 mg, 64%) as light yellow solid: R_f 0.35 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: –91.4 (*c* 1.0, $CHCl_3$); m.p. 115–116 °C (EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3308 (w), 2942 (w), 1668 (s) (C=O), 1602 (s) (C=O), 1566 (m), 1486 (m), 1438 (s), 1273 (m), 1238 (m), 1176 (m), 839 (m); ^{1}H NMR (700 MHz, $CDCl_3$) δ 7.96–7.87 (2 H, m, $C^{3'}$-H), 7.70 (1 H, s, C^4-H), 7.74–7.65 (2 H, m, $C^{2'}$-H), 4.56–4.40 (1 H, m, $C^{1''}$-H), 4.23 (1 H, dd, *J* 7.9, 3.7, C^7-H), 3.32 (3 H, s, N-CH_3), 2.56 (3 H, s, C^{10}-H_3), 2.18–2.10 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 2.02–1.95 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 1.91–1.63 (8 H, m, $C^{3''}$-H_2, $C^{4''}$-H_2, $C^{1'''}$-H_2, $C^{2''}$-H_b, $C^{5''}$-H_b), 0.87 (3 H, dd, *J* 7.5, 7.5, $C^{2'''}$-H_3); ^{13}C (176 MHz, $CDCl_3$) δ 196.7 ($C^{4'a}$), 163.7 (C^6), 154.8 (C^{8a}), 152.3 (C^2), 144.9 ($C^{1'}$), 137.7 (C^4), 130.3 ($C^{4'}$), 129.9 ($C^{3'}$), 117.1 ($C^{2'}$), 116.7 (C^{4a}), 60.0 ($C^{1''}$), 58.9 (C^7), 31.5 ($C^{1'''}$), 29.7 ($C^{2''}$ or $C^{5''}$), 29.2 ($C^{2''}$ or $C^{5''}$), 28.1 (NCH_3), 27.2 ($C^{3'}$), 26.3 (C^{10}), 23.5 ($C^{3''}$ or $C^{4''}$), 23.1 ($C^{3''}$ or $C^{4''}$), 9.1 ($C^{2'''}$); LRMS *m/z* (ES^+) 394 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) found 394.22323 (100%), $C_{22}H_{27}N_5O_2$ requires $[M+H]^+$ 394.22375; Analytical HPLC (RP, method B) rt 6.11 mins, purity 98%.

(*7R*)-8-Cyclopentyl-7-ethyl-5-methyl-2-((4'-(2,2,2-trifluoroacetyl)phenyl)amino)-7,8-dihydropteridin-6(5*H*)-one ((*R*)-131)



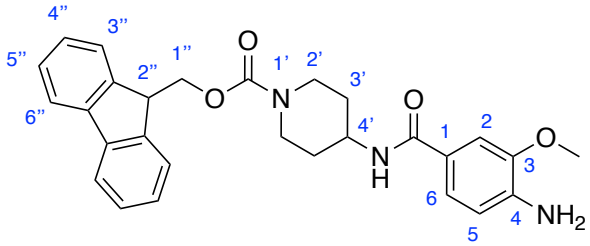
(Trifluoroacetyl)aniline (77 mg, 0.41 mmol, 2.0 eq.) and conc. HCl (60 µL) were added to a suspension of (*R*)-2-chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-92**) (60 mg, 0.21 mmol, 1.0 eq.) in H_2O (2.4 mL) and EtOH (0.6 mL). The resulting solution was stirred under reflux for 50 h. The precipitated mixture was suspended in H_2O (10 mL), the pH was adjusted to pH 5–6 using sat. K_2CO_3, then was fully solubilised using 0.5 M citric acid (30 mL) and H_2O (10 mL). The mixture was extracted with CH_2Cl_2 (2 × 125 mL). The combined organic components were washed with 0.5 M citric acid (2 × 50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether and EtOAc (slow gradient from 1:0 to 1:1, then fast gradient from 1:1 to 0:1), afforded *(7R)*-8-cyclopentyl-7-ethyl-5-methyl-2-((4-(2,2,2-trifluoroacetyl)phenyl)amino)-7,8-dihydropteridin-6(5H)-one (**(*R*)-131**) (26 mg, 28%) as bright yellow solid: R_f 0.31 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: –66.0 (c 1.0, $CHCl_3$); m.p. 183–185 °C (EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3267 (w), 2966 (w), 1656 (s) (C=O), 1601 (s) (C=O), 1584 (s), 1424 (s), 1340 (m), 1238 (m), 1162 (w), 1141 (m), 935 (m), 765 (m); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.08–7.97 (2 H, m, $C^{3'}$-H), 7.82–7.73 (2 H, m, $C^{2'}$-H), 7.69 (1 H, s, C^4-H), 7.30 (1 H, br s, NH), 4.55–4.39 (1 H, m, $C^{1''}$-H), 4.25 (1 H, dd, *J* 7.8, 3.7, $C^{1'''}$-H), 3.34 (3 H, s, N-CH_3), 2.23–2.08 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 2.08–1.95 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 1.95–1.52 (8 H, m, $C^{2''}$-H_b, $C^{5''}$-H_b, $C^{1'''}$-H_2, $C^{3''}$-H_2, $C^{5''}$-H_2), 0.88 (3 H, dd, *J* 7.6, 7.6, $C^{2'''}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 178.8 (q, J_{CF} 34.2, C^9), 163.6 (C^6), 154.3 (C^2 or C^{8a}), 152.39 (C^2 or C^{8a}), 147.2 (C^{4a}), 137.54 (C^4), 132.2 (q, J_{CF} 2.2, $C^{3'}$), 122.6 ($C^{2''}$), 117.26 ($C^{2'}$), 117.17 (q, J_{CF} 290.0, CF_3), 60.2 (C^7), 58.8 ($C^{1''}$), 29.8 ($C^{2''}$ or $C^{5''}$), 29.3 ($C^{2''}$ or $C^{5''}$), 28.4 (N-CH_3), 27.4 ($C^{1'''}$), 23.6 ($C^{3''}$ or $C^{4''}$), 23.3 ($C^{3''}$ or $C^{4''}$), 9.3 ($C^{2'''}$); ^{19}F NMR (471 MHz, $CDCl_3$) δ –70.8 (3 F, br s, CF_3); LRMS *m/z* (ES^+) 448 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) found 448.19522 (100%), $C_{22}H_{24}F_3N_5O_2$ requires $[M+H]^+$ 448.19549; Analytical HPLC (RP, method B) rt 6.47 mins, purity 98%.

tert-Butyl 4-(3-methoxy-4-nitrobenzamido)piperidine-1-carboxylate (146a)



HBTU (190 mg, 0.499 mmol, 1.20 eq.) and dry NEt_3 (100 µL, 0.733 mmol, 1.80 eq.) were added to a solution of 3-methoxy-4-nitrobenzoic acid (82.0 mg, 0.416 mmol, 1.00 eq.) in dry CH_2Cl_2 (2.0 mL). The resulting suspension was stirred at room temperature for 1 h. 1-Boc-4-aminopiperidine (100 mg, 0.499 mmol, 1.20 eq.) was added and the reaction mixture was stirred at room temperature for 24 h. The reaction mixture was diluted with CH_2Cl_2 (50 mL), and was washed with 1 M HCl (30 mL). The aqueous layer was extracted with CH_2Cl_2 (25 mL), the combined organic components were washed with H_2O (2 × 50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (10:1 to 0:1), afforded *tert*-butyl 4-(3-methoxy-4-nitrobenzamido)piperidine-1-carboxylate (**146a**) (148 mg, 94%) as a colourless solid: R_f 0.68 (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3251 (w), 2947 (w), 1674 (s) (C=O), 1634 (s) (C=O), 1536 (s), 1418 (m), 1235 (m), 1141 (m), 839 (m); m.p. 166–168 °C (EtOAc); ^{1}H NMR (400 MHz, CD_3OD) δ 7.84 (1 H, d, *J* 8.3, C^5-H), 7.68 (1 H, d, *J* 1.6, C^2-H), 7.49 (1 H, dd, *J* 8.3, 1.6, C^6-H), 4.17–4.04 (3 H, m, $C^{4'}$-H and $C^{2'}$-H_a), 4.01 (3 H, s, O-CH_3), 2.95 (2 H, br s, $C^{2'}$-H_b), 2.00–1.92 (2 H, m, $C^{3'}$-H_a), 1.47 (11 H, s, $C^{3'}$-H_b and 3 × CH_3); ^{13}C NMR (101 MHz, CD_3OD) δ 166.1 (C^{1a}=O), 155.1 ($C^{1'a}$=O), 152.2 (C^3), 141.4 (C^1), 139.5 (C^4), 124.7 (C^5), 118.9 (C^6), 112.7 (C^2), 79.8 (CH-$(CH_3)_3$), 55.9 (OCH_3), 47.7 ($C^{4'}$), 37.5 ($C^{2'}$), 31.1 ($C^{3'}$), 27.3 (3 × CH_3). LRMS *m/z* (ES^+) 378 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) found 402.16333 (100%), $C_{18}H_{25}O_6N_3$ requires $[M+Na]^+$ 402.16356.

(9*H*-Fluoren-9-yl)methyl 4-(4-amino-3-methoxybenzamido)piperidine-1-carboxylate (146)

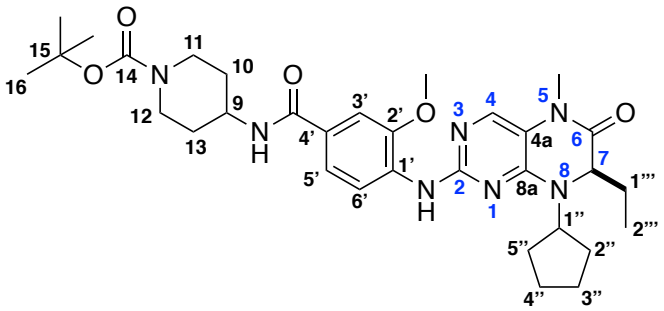


HBTU (211 mg, 0.557 mmol, 1.20 eq.) and DIPEA (125 µL, 0.696 mmol, 1.50 eq.) were added to a solution of 3-methoxy-4-nitrobenzoic acid (91.5 mg, 0.464 mmol, 1.00 eq.) in dry CH_2Cl_2 (2.0 mL). The resulting suspension was stirred at room temperature for 1 h. A solution of 1-Fmoc-4-aminopiperidine (200 mg, 0.557 mmol, 1.20 eq.) and DIPEA (125 µL, 0.696 mmol, 1.50 eq.) in dry CH_2Cl_2 (2.0 mL) was added and the reaction mixture was stirred at room temperature for 18 h. The reaction mixture was diluted with 1 M HCl (50 mL), and was extracted with CH_2Cl_2 (2 × 100 mL). The combined organic components were washed with H_2O (50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was suspended in $CHCl_3$/CH_2Cl_2/EtOAc/EtOH/toluene (2:2:1:1:2, 50 mL) and the insoluble impurities were removed by filtration. The filtrate was concentrated *in vacuo* to give presumed (9*H*-fluoren-9-yl)methyl 4-(3-methoxy-4-nitrobenzamido)piperidine-1-carboxylate (**148**) as a colourless solid (240 mg, > 99%); R_f 0.58 (petroleum ether:EtOAc = 1:1). The intermediate was used without further purification in the next step.

(9*H*-Fluoren-9-yl)methyl 4-(3-methoxy-4-nitrobenzamido)piperidine-1-carboxylate (**148**) (100 mg, 0.199 mmol, 1.00 eq.) was suspended in toluene (25 mL) and MeOH (2.0 mL), and the resulting fine suspension was purged with a flow of $N_{2(g)}$ for 10 mins prior to addition of 10% Pd/C (42.0 mg, 0.0398 mmol, 0.200 eq.). The reaction mixture was stirred under H_2-atmosphere for 1 h at room temperature, then filtered through Celite® with MeOH (4 × 60 mL), and concentrated *in vacuo*. Purification using a SNAP 10 g column on the Biotage chromatography system, eluting with petroleum ether:EtOAc (1:0 to 0:1), afforded (9*H*-fluoren-9-yl)methyl 4-(4-amino-3-methoxybenzamido)piperidine-1-carboxylate (**146**) (66 mg, 70%) as a coulourless solid: R_f 0.31 (petroleum ether:EtOAc = 1:1); m.p. 144–146 °C (toluene); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3303 (w), 2938 (w), 1693 (s) (C=O), 1616 (s) (C=O), 1535 (m), 1443 (m), 1322 (m), 1223 (s), 1184 (m), 777 (m), 738 (m); ^1H

NMR (400 MHz, $CDCl_3$) δ 7.77 (2 H, dt, *J* 7.5, 0.9, $C^{6''}$-H), 7.58 (2 H, dt, *J* 7.5, 0.9, 0.9, $C^{3''}$-H), 7.40 (2 H, ddd, *J* 7.5, 7.5, 0.9, $C^{4''}$-H), 7.35 (1 H, d, *J* 1.9, C^{2}-H), 7.32 (2 H, ddd, *J* 7.5, 7.5, 1.9, $C^{5''}$-H), 7.10 (1 H, dd, *J* 8.1, 1.9, C^{6}-H), 6.67 (1 H, d, *J* 8.1, C^{5}-H), 5.80 (1 H, d, *J* 7.8, NH), 4.46 (2 H, d, *J* 6.6, $C^{1''}$-H_2), 4.25 (1 H, t, *J* 6.6, $C^{2''}$-H_3), 4.21–3.95 (4 H, m, $C^{4'}$-H and NH_2), 3.92 (3 H, s, O-CH_3), 2.97 (2 H, br s, $C^{2'}$-H_a), 2.01 (2 H, br d, *J* 13.0, $C^{2'}$-H_b), 1.45–1.20 (4 H, m, $C^{3'}$-H_2); ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.7 (C^{1a}=O), 155.2 ($C^{1'a}$=O), 146.8 (C^{3}), 144.1 (C^{4} or $C^{2''a}$), 141.4 (C^{4} or $C^{2''a}$), 139.8 ($C^{6''a}$), 127.7 ($C^{5''}$ or $C^{6''}$), 127.0 ($C^{5''}$ or $C^{6''}$), 124.9 ($C^{4''}$), 124.1 (C^{1}), 120.0 ($C^{3''}$), 119.5 (C^{6}), 113.2 (C^{5}), 109.9 (C^{2}), 67.2 ($C^{1''}$), 55.7 (OCH_3), 47.5 ($C^{2''}$), 46.9 ($C^{4'}$), 43.1 ($C^{2'}$), 32.2 ($C^{3'}$); LRMS *m/z* (ES^+) 472 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) found 472.22311 (100%), $C_{28}H_{30}O_4N_3$ requires $[M+H]^+$ 472.22308; Analytical HPLC (RP, method B) rt 7.51 mins, purity 97%.

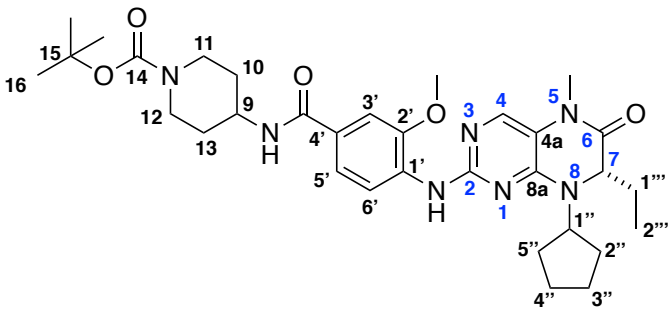
tert-Butyl (_7R_)-4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxybenzamido)piperidine-1-carboxylate ((_R_)-150)



Conc. HCl (0.51 mL, 16.6 mmol, 10.0 eq.) was added to a suspension of (*R*)-2-chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-92**) (700 mg, 1.66 mmol, 1.00 eq.) and 4-amino-3-methoxybenzoic acid (422 mg, 2.49 mmol, 1.50 eq.) in H_2O (4.8 mL) and EtOH (1.2 mL). The mixture was stirred at 100 °C for 20 h, then the reaction mixture was concentrated *in vacuo*. The residue was dissolved in DMF (12 mL), then diisopropylamine (0.47 mL, 3.32 mmol, 2.00 eq.) and HBTU (1.01 g, 2.66 mmol, 1.60 eq.) were added. The reaction was stirred at room temperature for 2 h, then *tert*-butyl 4-aminopiperidine-1-carboxylate (570 mg, 2.83 mmol, 1.70 eq.) was added. After 2 h, HBTU (316 mg, 0.831 mmol, 0.50 eq.) was added, and the mixture was stirred at room temperature for 18 h. Diisopropylamine (1.19 mL, 8.31 mmol, 5.00 eq.) and *tert*-butyl 4-aminopiperidine-1-carboxylate (333 mg, 1.66 mmol, 1.00 eq.) were added, and the reaction mixture was stirred for 2 h, then diluted with H_2O (100 mL) and extracted

with EtOAc (2 × 150 mL). The combined organic components were washed with sat. aq. LiCl (4 × 100 mL), filtered over an hydrophobic frit and concentrated *in vacuo*. Purification by silica gel chromatography (Combi*flashRf*+, 80 g column), eluting with CH_2Cl_2:MeOH (1:0 to 10:1) afforded *tert*-butyl (*7R*)-4''''-(4'-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxybenzamido)piperidine-1''''-carboxylate (**(*R*)-150**) (736 mg, 64%) as a pale brown oil. A small amount of *tert*-butyl (*7R*)-4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxybenzamido)piperidine-1-carboxylate **(*R*)-150** (50 mg) was further purified by MDAP (method D), while the rest was used as such in the following reactions. The mass-directed assisted purification (MDAP) preparative HPLC fractions were collected and concentrated *in vacuo* to give *tert*-butyl (7*R*)-4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidine-1-carboxylate (**(*R*)-150**) (26 mg, 51% recovery) as a colourless solid: R_f 0.55 (CH_2Cl_2:MeOH:NH_4OH = 90:9:1); $[\alpha]_D^{25}$: –44.9 (c 1.0, MeOH) [lit.[143] $[\alpha]_D^{22}$: –112 (*c* 0.2, $CHCl_3$)]; m.p. 88–90 °C (CH_2Cl_2) [lit.[143] 132 °C (EtOAc)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.51 (1 H, d, *J* 8.4, $C^{5'}$-H), 7.66 (1 H, s, C^4-H), 7.41 (1 H, d, *J* 1.9, $C^{3'}$-H), 7.23 (1 H, dd, *J* 8.5, 1.9, $C^{6'}$-H), 6.00 (1 H, d, *J* 7.8, N^{9a}-H), 4.56–4.40 (1 H, m, $C^{1''}$-H), 4.21 (1 H, dd, *J* 7.9, 3.7, C^7-H), 4.16–4.02 (1 H, m, C^9-H), 3.96 (3 H, s, O-CH_3), 3.31 (3 H, s, N-CH_3), 2.92 (2 H, t, *J* 12.6, C^{11}-H_a, C^{12}-H_a), 2.21–2.07 (3 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a, C^{10}-H_a, C^{13}-H_a), 2.08–1.92 (3 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a, C^{11}-H_b, C^{12}-H_b), 1.92–1.60 (8 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a, $C^{2''}$-H_b, $C^{5''}$-H_b, $C^{3''}$-H, $C^{4''}$-H_2, $C^{1'''}$-H_2), 1.46 (11 H, s, C^{16}-H_3, C^{10}-H_b, C^{13}-H_b), 0.87 (3 H, dd, *J* 7.5, 7.5, $C^{2'''}$-H_3); LCMS (high pH) rt 1.32 mins (97%), *m/z* (ES^+) 609 ($[M+H]^+$, 100%); Chiral HPLC (IB-N5 (250 × 4.6 mm, 5 µm), Hexane:EtOH = 75:25 with 0.2% *v/v* isopropylamine, 1.0 mL/min, 25 mins) major rt 14.9 mins, 92.9%, minor rt 17.9 mins, 7.1%. The ^{1}H NMR data are in good agreement with the literature.[143]

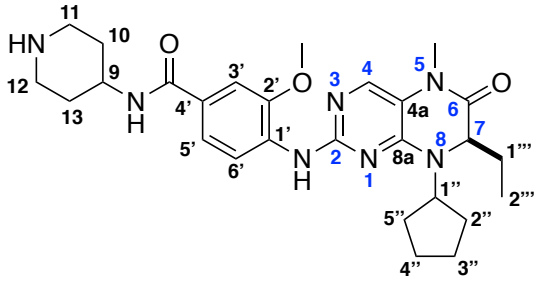
tert-Butyl (*7S*)-4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxybenzamido)piperidine-1-carboxylate ((*S*)-150)



Conc. HCl (0.51 mL, 16.6 mmol, 10.0 eq.) was added to a suspension of (*S*)-2-chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*S*)-92**) (700 mg, 1.66 mmol, 1.00 eq.) and 4-amino-3-methoxybenzoic acid (420 mg, 2.49 mmol, 1.50 eq.) in H_2O (4.8 mL) and EtOH (1.2 mL). The mixture was stirred at 100 °C for 20 h, then the reaction mixture was concentrated *in vacuo*. The residue was dissolved in DMF (12 mL), then diisopropylamine (0.47 mL, 3.32 mmol, 2.00 eq.) and HBTU (1.01 g, 2.66 mmol, 1.60 eq.) were added. The reaction was stirred at room temperature for 2 h, then *tert*-butyl 4-aminopiperidine-1-carboxylate (566 mg, 2.83 mmol, 1.70 eq.) was added. After 2 h, no conversion was observed. HBTU (316 mg, 0.831 mmol, 0.50 eq.) was added, and the mixture was stirred at room temperature for 18 h. Diisopropylamine (1.19 mL, 8.31 mmol, 5.00 eq.) and tert-butyl 4-aminopiperidine-1-carboxylate (334 mg, 1.66 mmol, 1.00 eq.) were added, and the reaction mixture was stirred for 2 h, then diluted with H_2O (100 mL) and extracted with EtOAc (2 × 150 mL). The combined organic components were washed with sat. aq. LiCl (4 × 100 mL), filtered over an hydrophobic frit and concentrated *in vacuo*. Purification by silica gel chromatography (Combi*flashRf*+, 80 g column), eluting with CH_2Cl_2 and MeOH (gradient from 1:0 to 10:1) afforded *tert*-butyl (*7S*)-4-(4'-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxybenzamido)piperidine-1-carboxylate (**(*S*)-150**) (1.10 g, 72%) as a pale brown oil. A small amount of *tert*-butyl (*7S*)-4-(4'-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxybenzamido)piperidine-1-carboxylate **(*S*)-150** (50 mg) was further purified by MDAP (method D), while the rest was used as such in the following reactions. The MDAP fractions were collected and concentrated *in vacuo* to give *tert*-butyl (*7S*)-4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidine-1-

carboxylate (**(*S*)-150**) (22 mg, 44% recovery) as a colourless solid: R_f 0.55 (CH_2Cl_2:MeOH:NH_4OH = 90:9:1); $[\alpha]_D^{25}$: +72.2 (c 1.0, MeOH); LCMS (high pH) rt 1.32 mins (97%), *m/z* (ES^+) 609 ($[M+H]^+$, 100%); Chiral HPLC (IB-N5 (250 × 4.6 mm, 5 µm), Hexane:EtOH = 75:25 with 0.2% *v/v* isopropylamine, 1.0 mL/min, 25 mins) minor rt 14.9 mins, 12.0%, major rt 17.9 mins, 88.0%. The spectroscopic data are identical to the enantiomer **(*R*)-150** and are in good agreement with the literature.[143]

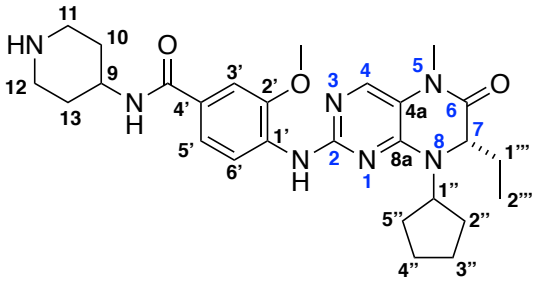
(7*R*)-4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxy-*N*-(piperidin-4-yl)benzamide ((*R*)-139)



Trifluoroacetic acid (1.52 ml, 19.7 mmol, 20.0 eq.) was added to a solution of (*R*)-2-chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-92**) (600 mg, 0.987 mmol, 1.00 eq.) in CH_2Cl_2 (6.0 mL) at 0 °C. The reaction mixture was stirred for 1 h at 0 °C, then the reaction mixture was warmed to room temperature, and stirred for 1 h. The reaction mixture was concentrated *in vacuo*, then the residue was dissolved in MeOH (10 mL) and K_2CO_3 (*ca*. 2 g) was added at 0 °C. The mixture was stirred vigorously for 10 mins, then diluted with H_2O (50 mL) and extracted with EtOAc (2 × 50 mL). The combined organic components were washed with H_2O (25 mL), then filtered over an hydrophobic frit and concentrated *in vacuo*. The residue was taken in minimal amount of MeOH, then loaded onto a 10 g SCX column, washed with MeOH (5 x 20 mL), then eluted with 2M NH_3 in MeOH (5 × 10 mL) using a fish tank. The collected fractions were concentrated *in vacuo*, then purified by silica gel chromatography (Combi*flashRf*+, 40 g column), eluting with CH_2Cl_2:2 M NH_3 in MeOH (1:0 to 8:2) to afford (7*R*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(piperidin-4-yl)benzamide (**(*R*)-139**) (354 mg, 70%) as a beige solid R_f 0.17 (CH_2Cl_2:MeOH:NH_4OH = 90:9:1); $[\alpha]_D^{25}$: –92.1 (*c* 1.0, MeOH) [lit.[143] $[\alpha]_D^{22}$: –86 (*c* 0.2, $CHCl_3$)]; m.p. 164–166 °C (MeOH); [lit.[143] 138 °C (EtOAc)] ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.54 (1 H, d, *J* 8.4, $C^{5'}$-H), 7.68 (1 H, s, C^4-H), 7.59 (1 H, s, N^{2a}-H), 7.42 (1 H, d, *J* 2.0, $C^{2'}$-H), 7.23 (1 H, dd, *J* 8.4, 2.0, $C^{6'}$-H),

5.95 (1 H, d, *J* 8.0, N^{9a}-H), 4.52 (1 H, p, *J* 8.0, $C^{1''}$-H), 4.21 (1 H, dd, *J* 7.9, 3.8, C^{7}-H), 4.16–4.04 (1 H, m, C^{9}-H), 3.97 (3 H, s, O-CH_3), 3.33 (3 H, s, N-CH_3), 3.11 (2 H, dt, *J* 12.8, 3.8, C^{11}-H_a, C^{12}-H_a), 2.77 (2 H, td, *J* 12.8, 2.5, C^{11}-H_b, C^{12}-H_b), 2.23–1.92 (3 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a, C^{10}-H_b, C^{13}-H_b), 1.92–1.58 (8 H, m, $C^{4''}$-H_2, $C^{1'''}$-H_2, $C^{2''}$-H_b, $C^{5''}$-H_b, $C^{3''}$-H_b, and $C^{2''}$-H_a or $C^{5''}$-H_a), 1.52–1.35 (2 H, m, C^{10}-H_a, C^{13}-H_a), 0.88 (3 H, dd, *J* 7.5, 7.5, $C^{2''}$-H_3); LCMS (high pH) rt 1.15 mins (> 99%), m/z (ES^+) 508 ($[M+H]^+$, 100%); HR-LCMS (formic acid) rt 1.92 mins (95%), *m/z* (ES^+) found 508.3037 (100%), $C_{27}H_{37}N_7O_3$ requires $[M+H]^+$ 508.3031. Chiral HPLC (IE5 (250 × 4.6 mm, 5 µm), Hexane:IPA = 50:50 with 0.2% *v/v* isopropylamine, 1.0 mL/min, 30 mins) minor rt 19.0 mins, 9.0%, major rt 21.4 mins, 90.0%. The data are in good agreement with the literature.[143]

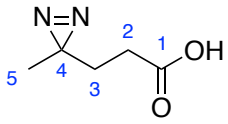
(*7S*)-4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxy-N-(piperidin-4-yl)benzamide ((*S*)-139)



Trifluoroacetic acid (1.52 ml, 19.7 mmol, 20.0 eq.) was added to a solution of (*S*)-2-chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*S*)-92**) (600 mg, 0.987 mmol, 1.00 eq.) in CH_2Cl_2 (6.0 mL) at 0 °C. The reaction mixture was stirred for 1 h at 0 °C, then the reaction mixture was warmed to room temperature, and stirred for 1 h. The reaction mixture was concentrated *in vacuo*, then the residue was dissolved in MeOH (10 mL) and K_2CO_3 (*ca.* 2 g) was added at 0 °C. The mixture was stirred vigorously for 10 mins, then diluted with H_2O (50 mL) and extracted with EtOAc (2 × 50 mL). The combined organic components were washed with H_2O (25 mL), then filtered over an hydrophobic frit, and concentrated *in vacuo*. The residue was taken in minimal amount of MeOH, then loaded onto a 10 g SCX column, washed with MeOH (5 x 20 mL), then eluted with 2M NH_3 in MeOH (5 × 10 mL) using a fish tank. The collected fractions were concentrated *in vacuo*, then purified by silica gel chromatography (Combi*flash*Rf+, 40 g column), eluting with CH_2Cl_2:2 M NH_3 in MeOH (1:0 to 8:2) to afford (*7S*)-4-((8-

cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxy-N-(piperidin-4-yl)benzamide (**(*S*)-139**) (299 mg, 57%) as a pale beige solid: R_f 0.17 (CH_2Cl_2:MeOH:NH_4OH = 90:9:1); $[\alpha]_D^{25}$: +91.1 (*c* 1.0, MeOH); m.p. 165–166 °C (MeOH); LCMS (high pH) rt 1.01 mins (98%), *m/z* (ES^+) 509 ($[M+H]^+$, 100%); Chiral HPLC (IE5 (250 × 4.6 mm, 5 µm), Hexane:IPA = 50:50 with 0.2% *v/v* isopropylamine, 1.0 mL/min, 30 mins) major rt 19.0 mins, 90.4%, minor rt 21.4 mins, 9.6%. The spectroscopic data are identical to the enantiomer **(*R*)-139** and are in good agreement with the literature.[143]

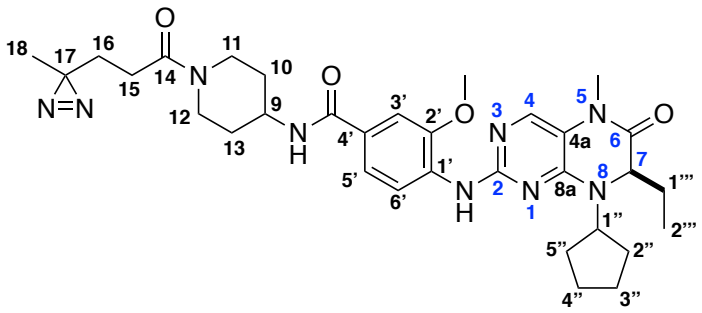
3-(3-Methyl-3*H*-diazirin-3-yl)propanoic acid (151)



4-Oxopentanoic acid (levulinic acid) (4.0 g, 34 mmol, 1.0 eq.) and 7 M NH_3 in MeOH (45 mL, 3.2 mmol, 9.1 eq.) were stirred at 0 °C for 3 h, then hydroxylamine-*O*-sulfonic acid (5.8 g, 51 mmol, 1.5 eq.) was added portionwise and the mixture was left to warm to room temperature then stirred for 42 h. Nitrogen was bubbled through the reaction mixture for 1 h and the mixture was filtered. The retentate was washed with MeOH (2 × 20 mL) and the filtrate was concentrated *in vacuo*. The residue was dissolved in MeOH (25 mL), triethylamine (14 mL, 1.0 mol, 3.0 eq.) was added and the reaction mixture was cooled to 0 °C. Iodine (8.7 g, 34 mmol, 1.0 eq.) was added portionwise at 0 °C under nitrogen in the dark with vigorous stirring until the dark colour remained. 2 M aq. HCl (*ca*. 30 mL) and aq. sodium thiosulfate (10% *w/v*, 150 mL) were added, and the mixture was extracted with EtOAc (3 × 60 mL). The combined organic components were washed with aq. sodium thiosulfate (10% *w/v*, 2 × 100 mL), then filtered over an hydrophobic frit and concentrated *in vacuo*. The crude was filtered over silica, flushing with CH_2Cl_2:MeOH (9:1, *ca*. 4 CV), then purified by silica gel chromatography (Combi*flashRf*+, 80 g column), eluting with CH_2Cl_2:MeOH (1:0 to 9:1) to afford 3-(3-methyl-3*H*-diazirin-3-yl)propanoic acid (**151**) (1.02 g, 22%) as a colourless liquid. R_f 0.05 (CH_2Cl_2:MeOH = 9:1); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2980 (w), 2666 (w), 1708 (s) (C=O), 1437 (m), 1387 (m), 1223 (m), 918 (w); ^1H (600 MHz, $CDCl_3$) 2.25 (2 H, t, *J* 6.5, C^2-H), 1.73 (2 H, t, *J* 6.5, C^3-H), 1.06 (3 H, s, C^5-H_3); ^{13}C (176 MHz, $CDCl_3$) δ 178.0 (C^1), 29.4 (C^3), 28.4 (C^2), 25.0 (C^4), 19.7 (C^5); LCMS (formic acid) rt 0.36 mins (20%), *m/z* (ES^-) 127 ($[M-H]^-$, 100%). The compound is not UV active,

so LCMS or HPLC could not be used to determine purity. Purity by ^{1}H NMR is estimated at > 90%. The data are in good agreement with the literature.[110]

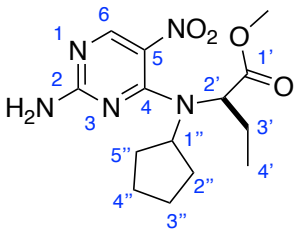
(*R*)-4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(1-(3-(3-methyl-3*H*-diazirin-3-yl)propanoyl)piperidin-4-yl)benzamide ((*R*)-132)



HATU (17.8 mg, 0.047 mmol, 1.2 eq.) and DIPEA (14 µL, 0.078 mmol, 2.0 eq.) were added to a solution of 3-(3-methyl-3*H*-diazirin-3-yl)propanoic acid (**151**) (5.0 mg, 0.039 mmol, 1.0 eq.) in CH_2Cl_2 (0.4 mL). The reaction was stirred for 1 h at room temperature, then (*R*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(piperidin-4-yl)benzamide (**(*R*)-139**) (24 mg, 0.047 mmol, 1.2 eq.) was added, and the mixture was stirred for 1.5 h. The volatile components were removed under a flow of nitrogen, then the residue was dissolved in DMSO (0.9 mL) and purified by MDAP (method C). The solvents were removed under a flow of nitrogen to afford (*R*)-4-((8 -cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(1-(3-(3-methyl-3*H*-diazirin-3-yl)propanoyl)piperidin-4-yl)benzamide (**(*R*)-132**) (14 mg, 53%) as an amorphous colourless solid: R_f 0.46 (EtOAc); $[\alpha]_D^{25}$: –48.2 (*c* 0.41, MeOH); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3339 (w), 2967 (w), 1674 (br, s) (C=O) 1630 (s) (C=O), 1502 (s), 1381 (m), 1322 (m), 1245 (m), 1035 (m), 953 (w), 757 (w); ^{1}H NMR (700 MHz, $CDCl_3$) 8.54 (1 H, d, *J* 8.4, C$^{6'}$-H), 7.68 (1 H, s, C^{4}-H), 7.65 (1 H, s, N^{2a}-H), 7.42 (1 H, d, *J* 2.0, C$^{3'}$-H), 7.23 (1 H, dd, *J* 8.5, 2.0, C$^{5'}$-H), 5.95 (1 H, d, *J* 7.7, N^{9a}-H), 4.60 (1 H, m, C^{11}-H$_{eq}$), 4.51 (1 H, tt, *J* 8.0, 8.0, C$^{1''}$-H), 4.25–4.16 (2 H, m, C^{7}-H and C^{9}-H), 3.98 (3 H, s, O-CH$_3$), 3.80 (1 H, m, C^{12}-H$_{eq.}$), 3.33 (3 H, s, N^{5}-CH$_3$), 3.22–3.14 (1 H, m, C^{12}-H$_{ax}$), 2.81–2.72 (1 H, m, C^{11}-H$_{ax}$), 2.23–1.92 (6 H, m, C^{15}-H$_2$, C$^{1'''}$-H$_a$, C$^{5''}$-H$_a$, C^{10}-H$_{eq}$ and C^{13}-H$_{eq}$), 1.91–1.75 (6 H, m, C^{16}-H$_2$, C$^{1'''}$-H$_b$, C$^{2''}$-H$_a$, C$^{3''}$-H$_a$, C$^{4''}$-H$_a$), 1.75–1.65 (4 H, m, C$^{2''}$-H$_b$, C$^{3''}$-H$_b$, C$^{4''}$-H$_b$, C$^{5''}$-H$_b$), 1.51–1.37 (2 H, m, C^{10}-H$_{ax}$ and C^{13}-H$_{ax}$), 1.06 (3 H, s, C^{18}-H$_3$), 0.88 (3 H, dd, *J* 7.5, 7.5, C$^{2'''}$-H$_3$); ^{13}C NMR (176 MHz, $CDCl_3$) δ 169.6 (C^{14}), 166.7 (C$^{4'a}$), 163.7 (C^{6}), 155.0 (C^{2}), 152.3 (C^{8a}), 147.3 (C$^{2'}$), 137.7 (C^{4}),

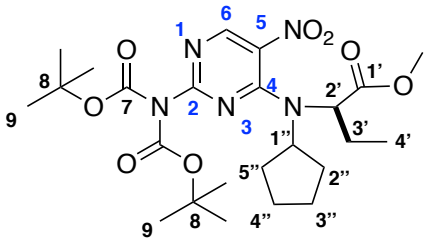
133.4 ($C^{1'}$), 126.1 ($C^{4'}$), 118.9 ($C^{5'}$), 116.5 (C^{4a}), 116.0 ($C^{6'}$), 109.1 ($C^{3'}$), 59.9 (C^{7}), 58.4 ($C^{1''}$), 56.0 (O-CH_3), 47.1 (C^{9}), 44.4 (C^{12}), 41.0 (C^{11}), 32.9 (C^{13}), 31.9 (C^{10}), 29.7 ($C^{2''}$ or $C^{5''}$), 29.6 ($C^{2''}$ or $C^{5''}$), 29.3 ($C^{1'''}$), 27.2 (C^{16}), 27.1 (C^{15}), 25.5 (C^{17}), 23.6 ($C^{3''}$ or $C^{4''}$), 23.2 ($C^{3''}$ or $C^{4''}$), 20.1 (C^{18}), 9.2 ($C^{2'''}$); LCMS (formic acid) rt 0.82 mins (95%), *m/z* (ES^+) 618 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) found 618.3486 (100%), $C_{32}H_{43}N_9O_4$ $[M+H]^+$ requires 618.3511.

Methyl (*R*)-2-((2-amino-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate ((*R*)-140)



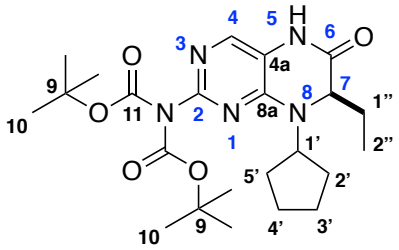
NH_4OH (400 µL) was added once every hour to a stirring solution of methyl (*R*)-2-((2-chloro-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate (**(*R*)-95**) (200 mg, 0.583 mmol, 1.0 eq.) in THF (2.0 mL) at room temperature for a total of 4 h. H_2O (50 mL) was added, and the mixture was extracted with CH_2Cl_2 (2 × 50 mL). The organic components were combined, washed with H_2O (50 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc (1:0 to 0:1), afforded methyl (*R*)-2-((2-amino-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate (**(*R*)-140**) (138 mg, 73%) as a yellow solid: R_f 0.32 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: +578 (*c* 1.0, MeOH); m.p. 154–156 °C (EtOAc); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 3335 (w), 2969 (w), 2360 (w), 1723 (s) (C=O), 1653 (m), 1579 (m), 1547 (m), 1338 (m), 1008 (w), 734 (m); ^{1}H NMR (500 MHz, $CDCl_3$) δ (1 H, s, C^6-H), 5.27–5.00 (2 H, m, NH_2), 3.67 (3 H, s, OCH_3), 3.61 (1 H, dd, *J* 8.0, 6.0, $C^{2'}$-H), 3.58–3.50 (1 H, m, $C^{1''}$-H), 2.33 (1 H, dtd, *J* 14.7, 7.6, 6.0, $C^{3'}$-H_a), 2.28–2.20 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 2.00 (1 H, dt, *J* 14.7, 7.6, $C^{3'}$-H_a), 1.91–1.82 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 1.81–1.41 (6 H, m, $C^{2''}$-H_b, $C^{5''}$-H_b, $C^{3''}$-H_2 and $C^{4''}$-H_2), 1.03 (3 H, dd, *J* 7.5, 7.5, $C^{4'}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 172.7 (C=O), 160.2 (C^2), 158.6 (C^4), 155.7 (C^6), 125.5 (C^5), 64.0 ($C^{2'}$), 59.1 ($C^{1''}$), 52.1 (OCH_3), 30.6 ($C^{2''}$ or $C^{5''}$), 27.7 ($C^{2''}$ or $C^{5''}$), 23.0 ($C^{3'}$), 22.84 ($C^{3''}$ or $C^{4''}$), 22.80 ($C^{3''}$ or $C^{4''}$), 11.7 ($C^{4'}$); LRMS *m/z* (ES^+) 324 ($[M+NH]^+$, 100%); HRMS *m/z* (ES^+) found 324.16597 (100%), $C_{14}H_{21}N_5O_4$ $[M+H]^+$ requires 324.16663; Analytical HPLC (RP, method B) rt 6.71 mins, purity 99%.

Methyl (*R*)-2-((2-(bis(*tert*-butoxycarbonyl)amino)-5-nitropyrimidin-4-yl)(cyclopentyl)amino) butanoate ((*R*)-142)



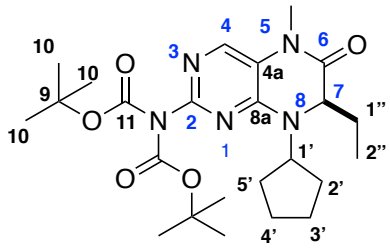
Boc_2O (730 mg, 3.34 mmol, 3.60 eq.), methyl (*R*)-2-((2-amino-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate (**(*R*)-140**) (300 mg, 0.928 mmol, 1.00 eq.) were dissolved in DMF (1.0 mL). DMAP (340 mg, 2.78 mmol, 3.00 eq.) was added and the reaction mixture was stirred for 1.5 h. The mixture was diluted with H_2O (100 mL) and extracted with EtOAc (2 × 50 mL). The organic components were combined, washed with aq. LiCl (10% *w/v*, 4 × 50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc (1:0 to 3:1), afforded methyl (*R*)-2-((2-(bis(*tert*-butoxycarbonyl)amino)-5-nitropyrimidin-4-yl)(cyclopentyl)amino) butanoate (**(*R*)-142**) (330 mg, 68%) as a pale yellow solid: R_f 0.58 (petroleum ether:EtOAc = 3:1); $[\alpha]_D^{25}$: +177 (*c* 1.0, MeOH); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2978 (w), 2360 (m), 1799 (s) (C=O), 11742 (br, s) (C=O), 1586 (m), 1282 (m), 1157 (m), 1118 (w), 858 (w), 733 (w); m.p. 147–149 °C ($CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.79 (1 H, s, C^6-H), 3.68 (4 H, s, $C^{2'}$-H and O-CH_3), 3.59 (1 H, dddd, *J* 8.3, 8.3, 8.3, 8.3, $C^{1''}$-H), 2.34 (1 H, dqd, *J* 15.0, 7.5, 5.7, $C^{3'}$-H_a), 2.23–2.13 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 2.13–2.01 (1 H, m, $C^{3'}$-H_b), 1.99–1.88 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 1.81–1.52 (6 H, m, $C^{2''}$-H_b, $C^{5''}$-H_b, $C^{3''}$-H_2 and $C^{4''}$-H_2), 1.48 (18 H, s, C^9-H_3), 1.00 (3 H, dd, *J* 7.5, 7.5, $C^{4'}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.0 ($C^{1'}$), 156.7 (C^6), 156.5 (C^2), 155.1 (C^4), 150.2 (C^7), 130.1 (C^5), 84.0 (C^8), 63.9 ($C^{1''}$), 60.4 ($C^{2'}$), 52.4 (O-CH_3), 30.4 ($C^{2''}$ or $C^{5''}$), 28.0 (C^9), 27.6 ($C^{2''}$ or $C^{5''}$), 23.2 ($C^{3''}$ and $C^{4''}$), 23.1 ($C^{3'}$), 11.7 ($C^{4'}$); LRMS *m/z* (ES^+) 324 ([M–2Boc+3H]$^+$, 100%), 546 ([M+Na]$^+$, 20%); HRMS *m/z* (ES^+) found 524.2710 (100%), $C_{24}H_{38}N_5O_8$ [M+H]$^+$ requires 524.2715; Analytical HPLC (RP, method A) rt 11.7 mins, purity 97%.

tert-Butyl (*R*)-(*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate ((*R*)-143)



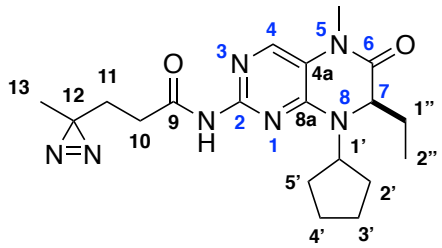
NH_4Cl (0.85 mg, 15 mmol, 25 eq.) and zinc powder (1.0 g, 15 mmol, 25 eq.) were added to a solution of methyl (*R*)-2-((2-(bis(*tert*-butoxycarbonyl)amino)-5-nitropyrimidin-4-yl)(cyclopentyl)amino) butanoate (**(*R*)-142**) (0.32 g, 0.61 mmol, 1.0 eq.) in DMF (6.0 mL). The reaction was stirred for 22 h at room temperature, then the reaction mixture was diluted with EtOAc (50 mL), filtered over a microfiber filter, rinsing with EtOAc (2 × 50 mL). The filtrate was diluted with H_2O (100 mL) and extracted with EtOAc (50 mL). The organic components were combined, washed with aq. LiCl (10% *w/v*, 5 × 50 mL), dried on $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc (9:1 to 0:1), afforded *tert*-butyl (*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate (**(*R*)-143**) (230 mg, 82%) as a colourless solid: R_f 0.53 (EtOAc); $[\alpha]_D^{25}$: –128 (*c* 1.0, MeOH); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2974 (w), 2360, 1791 (s) (C=O), 1696 (br, s) (C=O), 1608 (m), 1415 (m), 1285 (m), 1104 (m), 920 (w), 734 (w); m.p. 187–189 °C (EtOAc); ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.00 (1 H, s, N^5-H), 7.74 (1 H, s, C^4-H), 4.24 (1 H, dd, *J* 7.0, 3.4, C^7-H), 4.17–4.06 (1 H, m, $C^{1'}$-H), 2.09–1.71 (8 H, m, $C^{1''}$-H_2, $C^{2'}$-H_2, $C^{5'}$-H_2, $C^{3'}$-H_b and $C^{4'}$-H_b), 1.70–1.51 (2 H, m, $C^{3'}$-H_b and $C^{4'}$-H_b), 1.43 (18 H, s, C^{10}-H_3), 0.90 (3 H, dd, *J* 7.5, 7.5, $C^{2''}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.1 (C^6), 152.8 (C^2), 151.8 (C^{8a}), 151.0 (C^{11}), 138.6 (C^4), 117.4 (C^{4a}), 82.9 (C^9), 61.9 (C^7), 60.4 ($C^{1'}$), 29.3 ($C^{2'}$ and $C^{5'}$), 28.1 (C^{10}), 27.2 ($C^{1''}$), 24.24 ($C^{3'}$ or $C^{4'}$), 24.21 ($C^{3'}$ or $C^{4'}$), 8.6 ($C^{2''}$). LRMS *m/z* (ES^+) 288 ([M–Boc–^{t}Bu+3H]$^+$, 100%), 484 ([M+Na]$^+$, 15%); HRMS *m/z* (ES^+) found 462.2709 (100%), $C_{23}H_{35}N_5O_5$ requires [M+H]$^+$ 462.2711; Analytical HPLC (RP, method A) rt 8.75 mins, purity 96%.

tert-Butyl (*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate ((*R*)-144)



MeI (140 µL, 2.22 mmol, 5.1 eq.) was added dropwise to a solution of *tert*-butyl (*R*)-(*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate (**(*R*)-143**) (200 mg, 0.433 mmol, 1.0 eq.) and K_2CO_3 (310 mg, 2.22 mmol, 5.1 eq.) in dry DMF (6.0 mL). The reaction mixture was stirred at room temperature for 24 h, then was diluted with H_2O (50 mL)and extracted with EtOAc (3 × 50 mL). The organic components were combined, washed with aq. LiCl (10% *w/v*, 5 × 50 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc (3:1 to 0:1), afforded *tert*-butyl (*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate (**(*R*)-144**) (58 mg, 28%) as a pale yellow oil: R_f 0.26 (EtOAc); $[\alpha]_D^{25}$: −38 (*c* 1.0, $CHCl_3$); $\tilde{\nu}_{max}$ (neat)/cm^{-1} 2975 (w), 2360 (m), 1791 (s) (C=O), 1750 (s) (C=O), 1683 (br, s) (C=O), 1586 (m), 1425 (m), 1368 (m), 1279 (m), 1162 (m), 1101 (m), 854 (w); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.81 (1 H, s, C^4-H), 4.28 (1 H, dd, *J* 7.1, 3.4, C^7-H), 4.11 (1 H, dddd, *J* 8.4, 8.4, 8.4, 8.4, $C^{1'}$-H), 3.35 (3 H, s, N^5-CH_3), 2.06–1.79 (7 H, m, $C^{1''}$-H_a, $C^{3'}$-H_a, $C^{4'}$-H_a, $C^{2'}$-H_2 and $C^{5'}$-H_2), 1.74 (1 H, dt, *J* 14.6, 7.4, $C^{1''}$-H_b), 1.62–1.51 (2 H, m, $C^{3'}$-H_b and $C^{4'}$-H_b), 1.44 (18 H, s, 6 × C^{10}-H_3), 0.82 (3 H, dd, *J* 7.5, 7.5, $C^{2''}$-H_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 164.5 (C^6=O), 152.5 (C^2), 152.3 (C^{8a}), 151.0 (C^{11}=O), 137.6 (C^4), 120.6 (C^{4a}), 82.9 (C^9), 61.8 ($C^{1''}$), 60.2 ($C^{1'}$), 29.32 ($C^{2'}$ or $C^{5'}$), 29.30 ($C^{2'}$ or $C^{5'}$), 28.3 (N^5-CH_3), 28.1 (C^{10}), 27.4 ($C^{1''}$), 24.2 ($C^{3'}$ and $C^{4'}$), 8.8 ($C^{2''}$); LRMS *m/z* (ES^+) 498 ($[M+Na]^+$, 100%); HRMS *m/z* (ES^+) found 462.2709 (100%), $C_{23}H_{35}N_5O_5$ requires $[M+H]^+$ 462.2711; Analytical HPLC (RP, method A) rt 5.2 mins, purity 88%.

(*R*)-*N*-(8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)-3-(3-methyl-3*H*-diazirin-3-yl)propenamide ((*R*)-133)

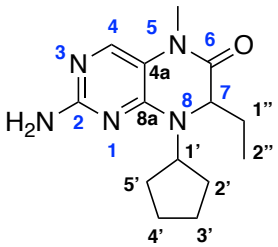


Trifluoroacetic acid (0.50 mL, 6.5 mmol, 66 eq) was added dropwise to a solution of *tert*-butyl (*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate (**(*R*)-144**) (47 mg, 0.098 mmol, 1 eq.) in CH_2Cl_2 (1 mL). The reaction mixture was stirred at room temperature for 20 mins, then the solvents were removed *in vacuo* to afford (*R*)-2-amino-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-138**) (38 mg, >99%) as a yellow oil: R_f 0.30 (EtOAc); Chiral HPLC (IG5 (250 × 4.6 mm, 5 µm), MeOH 0.2% *v/v* isopropylamine, 1.0 mL/mins, 15 mins) minor rt 6.7 mins, 0.2%, major rt 8.2 mins, 99.8%. The crude title compound was used immediately in the following reaction without further purification.

HATU (35 mg, 0.092 mmol, 1.2 eq.) and DIPEA (32 µL, 0.19 mmol, 2.4 eq.) were added to a solution of 3-(3-methyl-3*H*-diazirin-3-yl)propanoic acid (**151**) (10 mg, 0.018 mmol, 1.0 eq.) in CH_2Cl_2 (0.5 mL). The reaction mixture was stirred for 30 mins at room temperature, then crude (*R*)-2-amino-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-138**) (36 mg, 0.092 mmol, 1.2w eq.) and DIPEA (32 µL, 0.19 mmol, 2.4 eq.) in CH_2Cl_2 (0.5 mL) were added. DIPEA (32 µL, 0.19 mmol, 2.4 eq.) was added to obtain a pH of 8–9. The reaction was stirred for 12 h. HATU (140 mg, 0.37 mmol, 4.8 eq.) and DIPEA (128 µL, 0.74 mmol, 9.6 eq.) were added to a solution of 3-(3-methyl-3*H*-diazirin-3-yl)propanoic acid (**151**) (30 mg, 0.023 mmol, 3.0 eq.) in CH_2Cl_2 (0.5 mL), the mixture was stirred for 30 mins then added dropwise to the reaction mixture. Stirring at room temperature for 5 h was followed by addition of HATU (35 mg, 0.092 mmol, 1.2 eq.) and DIPEA (32 µL, 0.19 mmol, 2.4 eq.), and further stirring for 2 h. The reaction mixture was diluted with H_2O (40 mL), then extracted with CH_2Cl_2 (3 × 30 mL). The combined organic components were washed with brine (30 mL) and H_2O (30 mL), then concentrated *in vacuo*. Purification by silica gel chromatography, eluting with petroleum ether:EtOAc (4:1 to 0:1), was followed

by purification by preparative TLC (elution with EtOAc, silica extraction with EtOH) to afford (*R*)-*N*-(8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)-3-(3-methyl-3*H*-diazirin-3-yl)propenamide (**(*R*)-133**) (4.5 mg, 12%) as an amorphous corlourless solid: R_f 0.20 (petroleum ether:EtOAc = 1:1); $[\alpha]_D^{25}$: –17.4 (*c* 0.125, $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2959 (w), 2359 (w), 1671 (s) (C=O), 1601 (m), 1488 (s), 1431 (m), 846 (m); ^{1}H NMR (500 MHz, DMSO-D_6) δ 10.02 (1 H, s, N^5-H), 7.86 (1 H, s, C^4-H), 4.39–4.08 (2 H, m, C^7-H and $C^{1'}$-H), 3.25 (3 H, s, N^5-CH_3), 2.38 (2 H, br t, *J* 7.6, C^{10}-H), 2.07–1.70 (7 H, m, $C^{2'}$-H_2, $C^{5'}$-H_2, $C^{1''}$-H_a, $C^{3'}$-H_a, and $C^{4'}$-H_a), 1.70–1.40 (5 H, m, C^{11}-H_2, $C^{1''}$-H_b, $C^{3'}$-H_b, and $C^{4'}$-H_b), 1.02 (3 H, s, C^{13}-H_3), 0.75 (3 H, dd, *J* 7.5, 7.5, $C^{2''}$-H_3); ^{13}C NMR (126 MHz, DMSO-D_6) δ 170.5 (C^9), 163.7 (C^6), 152.4 (C^2), 151.6 (C^{8a}), 138.4 (C^4), 118.4 (C^{4a}), 60.5 (C^7), 59.1 ($C^{1'}$), 31.1 (C^{10}), 29.6 (C^{11}), 28.9 ($C^{2'}$ or $C^{5'}$), 28.7 ($C^{2'}$ or $C^{5'}$), 28.2 (N^5-CH_3), 27.0 ($C^{1''}$), 26.3 (C^{12}), 23.7 ($C^{3'}$ or $C^{4'}$), 23.5 ($C^{3'}$ or $C^{4'}$), 19.9 (C^{13}), 9.3 ($C^{2''}$); LRMS *m/z* (ES^+) 358 ($[M–N_2+H]^+$, 100%), 386 ($[M+H]^+$, 20%); HRMS *m/z* (ES^+) found 386.2301 (100%), $C_{19}H_{27}N_7O_2$ requires $[M+H]^+$ 386.2299; Analytical HPLC (RP, method B) rt 6.69 mins, purity 95%. Chiral HPLC (IG5 (250 × 4.6 mm, 5 µm), MeOH with 0.2% *v/v* isopropylamine, 1.0 mL/mins, 10 mins) minor rt 7.3 mins, 0.2%, major rt 8.6 mins, 99.8%.

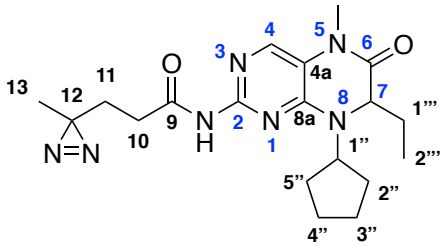
2-Amino-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (138)



(*R*)-2-Chloro-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (**(*R*)-92**) (0.15 g, 0.51 mmol, 1.0 eq.) was dissolved in isopropanol (1.2 mL) and ammonium hydroxide (2.6 mL, 70 mmol, 250 eq.), then submitted to microwave irradiation for 6 h at 170 °C, 6 h at 160 °C and then 4 h at 170 °C. H_2O (150 mL) was added and the mixture was extracted with EtOAc (150 mL). The organic components were filtered over an hydrophobic frit. Purification by silica gel chromatography (Combiflash*Rf*+, 24 g column), eluting with cyclohexane:EtOAc (1:0 to 0:1) and then with EtOH:Et_3N (9:1, 2 CV), afforded 2-amino-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5H)-one (**138**) (112 mg, 80%) as a colourless oil: R_f 0.30 (EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3339 (w), 2970 (w), 2873

(w), 1665 (s) (C=O), 1459 (s), 1247 (m), 954 (m); ^{1}H NMR (600 MHz, DMSO-D_6) δ 7.58 (1 H, s, C^{4}-H), 5.87 (2 H, s, NH_2), 4.36–4.27 (1 H, m, $C^{1'}$-H), 4.07 (1 H, dd, *J* 7.9, 3.7, C^{7}-H), 3.17 (3 H, s, N^{5}-CH_3), 1.93–1.83 (2 H, m, $C^{2'}$-H_a, $C^{5'}$-H_a), 1.79–1.65 (5 H, m, $C^{2'}$-H_b, $C^{5'}$-H_b ,$C^{3'}$-H_a, $C^{4'}$-H_a and $C^{1''}$-H_a), 1.62–1.48 (3 H, m, $C^{3'}$-H_b, $C^{4'}$-H_b and $C^{1''}$-H_b), 0.75 (3 H, dd, *J* 7.5, 7.5, $C^{2''}$-H_3); ^{13}C NMR (151 MHz, DMSO-D_6) δ 162.7 (C^{6}), 159.1 (C^{2}), 151.7 (C^{8a}), 139.2 (C^{4}), 113.8 (C^{4a}), 59.6 (C^{7}), 57.5 ($C^{1'}$), 28.9 ($C^{2'}$ or $C^{5'}$), 28.7 ($C^{2'}$ or $C^{5'}$), 27.6 (N^{5}-CH_3), 26.1 ($C^{1''}$), 23.4 ($C^{3'}$ or $C^{4'}$), 23.1 ($C^{3'}$ or $C^{4'}$), 9.0 ($C^{2''}$). LCMS (formic acid) rt 0.52 mins (96%), *m/z* (ES^{+}) 277 ($[M+H]^{+}$, 100%).

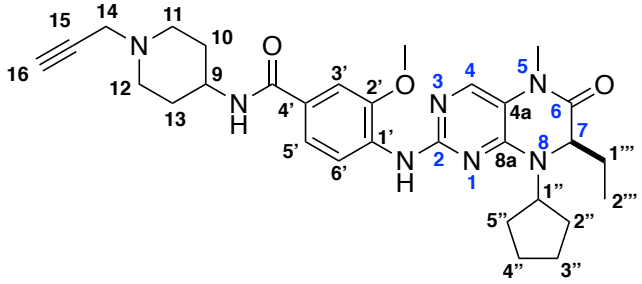
_N_-(8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)-3-(3-methyl-3*H*-diazirin-3-yl)propenamide (133)



HATU (72 mg, 0.19 mmol, 1.2 eq.) and DIPEA (0.055 mL, 0.31 mmol, 2.0 eq.) were added to a solution of 3-(3-methyl-3*H*-diazirin-3-yl)propanoic acid (**151**) (20 mg, 0.16 mmol, 1.0 eq.) in CH_2Cl_2 (1.0 mL). The reaction mixture was stirred for 1.5 h at room temperature, then 2-amino-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5H)-one (**138**) (52 mg, 0.19 mmol, 1.2 eq.) in CH_2Cl_2 (0.5 mL).was added. The reaction mixture was stirred for 2.5 h, then HATU (72 mg, 0.19 mmol, 1.2 eq.) and DIPEA (0.055 mL, 0.31 mmol, 2.0 eq.) were added. The reaction was stirred for 2 h, then left to stand for 48 h. Evaporation was observed, hence CH_2Cl_2 (0.5 ml) and DMF (0.25 mL) were added to re-solubilise the mixture. HATU (72 mg, 0.19 mmol, 1.2 eq.) and DIPEA (0.055 mL, 0.31 mmol, 2.0 eq.) were added and the RM was stirred at room temperature for 18 h. The reaction mixture was diluted with H_2O (30 mL), then extracted with EtOAc (2 × 30 mL). The combined organic components were washed with sat. aq. LiCl (50 mL), then concentrated *in vacuo*. The residue was dissolved in MeOH:DMSO (1:1, 1 mL) and purified by MDAP (method C). The fractions were combined and concentrated under a flow of nitrogen in the dark to afford *N*-(8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-

tetrahydropteridin-2-yl)-3-(3-methyl-3*H*-diazirin-3-yl)propenamide (**133**) (36 mg, 57%) as an off-white powdery solid: R_f 0.20 (petroleum ether:EtOAc = 1:1); LCMS (high pH) rt 1.08 mins (100%), *m/z* (ES^+) 387 ($[M+H]^+$, 100%); HR-LCMS (formic acid) rt 2.91 mins (100%), *m/z* (ES^+) found 386.2311 (100%), $C_{19}H_{27}N_7O_2$ requires $[M+H]^+$ 386.2299. The spectroscopic data are identical to the enantiomer **(*R*)-133**.

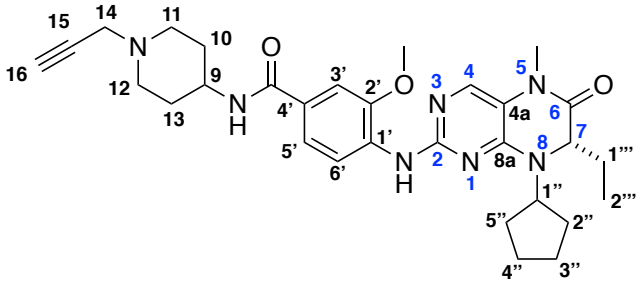
(*R*)-4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(1-(prop-2-yn-1-yl)piperidin-4-yl)benzamide ((*R*)-153)



Propargyl bromide solution (80% in toluene) (280 µL, 0.248 mmol, 1.05 eq.) and K_2CO_3 (39 mg, 0.284 mmol, 1.20 eq.) were added to a solution of (*7R*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-N-(piperidin-4-yl)benzamide (**(*R*)-139**) (120 mg, 0.236 mmol, 1.00 eq.) in DMF (1.5 mL) at 0 °C. The solution was stirred for 2 h at 0 °C, diluted with H_2O, then using sat aq. K_2CO_3, the pH was adjusted with to 10-11, and the mixture was extracted with EtOAc (3 × 80 mL). The combined organic components were washed with sat. aq. LiCl (4 x 50 mL) (the pH of each portion was adjusted with *ca.* 0.2 mL of sat aq. K_2CO_3), filtered over an hydrophobic frit and concentrated *in vacuo*. Purification by silica gel chromatography (Combi*flashRf*+, 24 g column), eluting with CH_2Cl_2:[20% of 2 M NH_3 in MeOH into CH_2Cl_2] (1:0 to 9:1) afforded (*R*)-4'-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxy-*N*-(1-(prop-2-yn-1-yl)piperidin-4-yl)benzamide (**(*R*)-153**) (57 mg, 42%) as a beige solid: R_f 0.16 (EtOAc:EtOH = 99:1); m.p. 103–105 °C (MeOH); $[\alpha]_D^{25}$: –70.7 (*c* 0.5, MeOH); υ_{max} (thin film)/cm^{-1} 3298 (br), 2979 (br), 1665 (s, br) (C=O), 1594, 1500, 1472, 1437, 1215; ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.53 (1 H, d, *J* 8.4, $C^{6'}$-H), 7.67 (1 H, s, C^4-H), 7.58 (1 H, s, N^{2a}-H), 7.41 (1 H, d, *J* 1.9, $C^{3'}$-H), 7.23 (1 H, dd, *J* 8.5, 1.9, $C^{5'}$-H), 5.99 (1 H, d, *J* 7.9, N^{9a}-H), 4.54–4.43 (1 H, m, $C^{1''}$-H), 4.20 (1 H, dd, *J* 7.9, 3.7, C^7-H), 4.05–3.96 (1 H, m, C^9-H), 3.96 (3 H, s, OCH_3), 3.34 – 3.28 (5 H, m, C^{14}-H_2 and NCH_3), 2.89 (2 H, dt, *J* 11.9, 3.8, C^{11}-

$H_{eq,}$ C^{12}- H_{eq}), 2.41 (2 H, td, *J* 11.6, 2.6, C^{11}-H_{ax}, C^{12}-H_{ax}), 2.26 (1 H, t, *J* 2.4, C^{16}-H), 2.19–2.10 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 2.10–2.05 (2 H, m, C^{10}-$H_{eq,}$ C^{13}- H_{eq}), 2.05–1.94 (1 H, m, $C^{2''}$-H_a or $C^{5''}$-H_a), 1.94–1.64 (8 H, m, $C^{1'''}$-$H_{2,}$ $C^{3''}$-H_2, $C^{4''}$-H_2, $C^{2''}$-H_b, $C^{5''}$-H_b), 1.64–1.55 (2 H, m, C^{10}-H_{ax} and C^{13}-H_{ax}), 0.86 (3 H, t, *J* 7.5, $C^{2'''}$-H_3). ^{13}C NMR (151 MHz, $CDCl_3$) δ 166.6 ($C^{4'a}$=O), 163.7 (C^6=O), 155.0 (C^2), 152.3 (C^{8a}), 147.2 ($C^{2'}$), 137.9 (C^4), 133.2 ($C^{4'}$), 126.5 ($C^{1'}$), 118.9 ($C^{5'}$), 116.4 (C^{4a}), 115.9 ($C^{6'}$), 109.1 ($C^{3'}$), 78.8 (C^{15}), 73.1 (C^{16}), 59.9 (C^7), 58.3 ($C^{1''}$), 55.9 (OCH_3), 51.1 (C^{11}), 47.0 (C^{14}), 46.5 (C^9), 38.6 (C^{12}), 32.2 (C^{10}), 29.7 ($C^{2''}$ or $C^{5''}$), 29.3 ($C^{2''}$ or $C^{5''}$), 28.1 (N^5-CH_3), 27.1 ($C^{3''}$ or $C^{4''}$), 23.6 ($C^{3''}$ or $C^{4''}$), 23.2 ($C^{1'''}$), 9.2 ($C^{2'''}$); LCMS (high pH) rt 1.13 mins (100%), *m/z* (ES^+) 547 ($[M+H]^+$, 100%). Chiral HPLC (IB-N5 (250 × 4.6 mm, 5 µm), Hexane:EtOH = 65:35 with 0.2% *v/v* isopropylamine, 1.0 mL/min, 20 mins) major rt 12.7 mins, 86.6%, minor rt 14.6 mins, 13.4%.

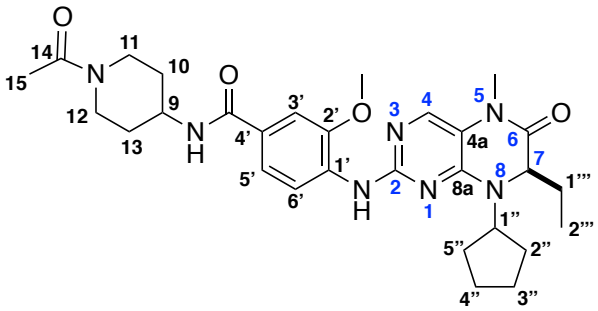
(7*S*)-4'-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxy-*N*-(1-(prop-2-yn-1-yl)piperidin-4-yl)benzamide ((*S*)-153)



Propargyl bromide solution (80% in toluene) (280 µL, 0.248 mmol, 1.05 eq.) and K_2CO_3 (39 mg, 0.284 mmol, 1.20 eq.) were added to a solution of (*S*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-N-(piperidin-4-yl)benzamide (**(*S*)-139**) (120 mg, 0.236 mmol, 1.00 eq.) in DMF (1.5 mL) at 0 °C. The solution was stirred for 2 h at 0 °C, diluted with H_2O, then using sat aq. K_2CO_3, the pH was adjusted with to 10–11, and the mixture was extracted with EtOAc (3 × 80 mL). The combined organic components were washed with sat. aq. LiCl (4 x 50 mL) (each portion was pH modified with *ca.* 0.2 mL of sat aq. K_2CO_3), filtered over an hydrophobic frit and concentrated *in vacuo*. Purification by silica gel chromatography (dry loaded on Isolute HM-N; Combi*flashRf*+, 24 g column), eluting with CH_2Cl_2:[20% of 2M NH_3 in MeOH into CH_2Cl_2] (1:0 to 9:1) afforded (*S*)-4'-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxy-*N*-(1-(prop-2-yn-1-yl)piperidin-4-

yl)benzamide (**(*S*)-153**) (47 mg, 35%) as a beige solid: R_f 0.16 (EtOAc:EtOH = 99:1); m.p. 103–105 °C (MeOH); $[\alpha]_D^{25}$: +73.4 (c 0.5, MeOH); LCMS (high pH) rt 1.13 mins (96%), *m/z* (ES^+) 547 ($[M+H]^+$, 100%); Chiral HPLC (IB-N5 (250 × 4.6 mm, 5 µm), Hexane:EtOH = 65:35 with 0.2% *v/v* isopropylamine, 1.0 mL/mins, 20 mins) minor rt 12.7 mins, 9.9%, major rt 14.6 mins, 90.1%. The spectroscopic data are identical to the enantiomer **(*R*)-153**.

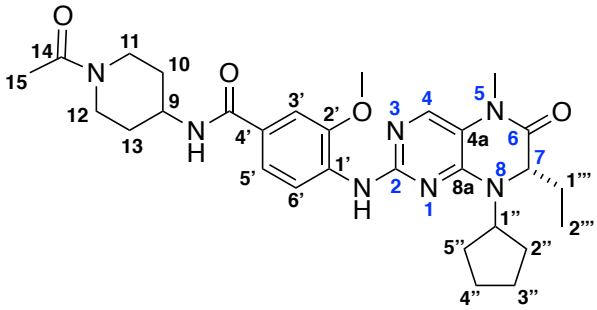
(*R*)-*N*-(1-Acetylpiperidin-4-yl)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamide ((*R*)-155)



AcCl (5 µL, 0.032 mmol, 1.1 eq.) and NEt_3 (8 µL, 0.059 mmol, 2.0 eq.) were added to a solution of (*R*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(piperidin-4-yl)benzamide (**(*R*)-139**) (15 mg, 0.029 mmol, 1.0 eq.) in CH_2Cl_2 (300 µL) at 0 °C. The reaction mixture was stirred for 1 h, then diluted with H_2O (15 mL), and extracted with CH_2Cl_2 (3 × 10 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (30 mL) and brine (30 mL), then dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with $CHCl_3$:MeOH:Et_3N (100:0:1 to 95:5:1), was followed by a MeCN:hexane grease wash to afford (*R*)-*N*-(1-acetylpiperidin-4-yl)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamide (**(*R*)-155**) (6 mg, 40%) as a colourless solid: R_f 0.28 ($CHCl_3$:MeOH:Et_3N = 95:5:1); $[\alpha]_D^{25}$: –46.3 (*c* 0.68, MeOH); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.54 (1 H, d, *J* 8.4, $C^{6'}$-H), 7.75–7.58 (2 H, m, C^4-H and N^{2a}-H), 7.42 (1 H, d, *J* 1.9, $C^{3'}$-H), 7.24–7.20 (1 H, m, $C^{5'}$-H), 5.99 (1 H, d, *J* 7.8, N^{9a}-H), 4.68–4.56 (1 H, m, C^{11}-H_{ax}), 4.51 (1 H, dddd, *J* 8.0, 8.0, 8.0, 8.0, $C^{1''}$-H), 4.26–4.13 (2 H, m, C^7-H and C^9-H), 3.98 (3 H, s, OCH_3), 3.91–3.78 (1 H, m, C^{11}-H_{ax}), 3.32 (3 H, s, NCH_3), 3.23 (1 H, ddd, *J* 14.1, 12.0, 2.8, C^{12}-$H_{ax.}$), 2.87–2.77 (1 H, m, C^{11}-H_{eq}), 2.25–1.92 (8 H, m, C^{12}-H_{eq}, $C^{2''}$-H_a, $C^{5''}$-H_a, C^{10}-H_{ax}, C^{13}-H_{ax}, C^{15}-H_3), 1.91–1.64 (8 H, m, $C^{2''}$-H_b, $C^{5''}$-H_b, $C^{1'''}$-H_2, $C^{3''}$-H_2, $C^{4''}$-H_2), 1.49–1.35 (2 H, m, C^{10}-H_{eq}, C^{13}-$H_{eq.}$), 0.87 (3 H, dd, *J* 7.5, 7.5, $C^{2'''}$-H_3); ^{13}C NMR (126 MHz,

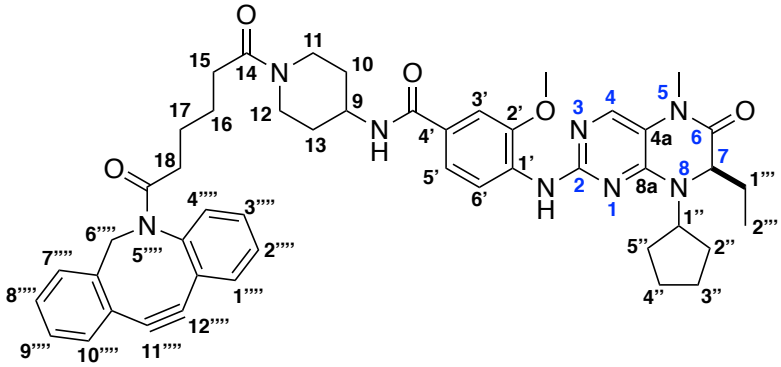
$CDCl_3$) δ 168.9 (C^{14}), 166.7 ($C^{4'a}$=O), 163.7 (C^6), 155.0 (C^2 or C^{8a}), 152.3 (C^2 or C^{8a}), 147.2 ($C^{2'}$), 137.7 ($C^{4'}$), 133.4 ($C^{4'}$), 126.1 ($C^{1'}$), 118.9 ($C^{5'}$), 116.4 (C^{4a}), 116.0 ($C^{6'}$), 109.1 ($C^{3'}$), 59.9 (C^7), 58.4 ($C^{1''}$), 56.0 (O-CH_3), 47.1 (C^9), 45.4 (C^{11}), 40.6 (C^{12}), 32.9 (C^{10} or C^{13}), 31.9 (C^{10} or C^{13}), 29.8 ($C^{2''}$ or $C^{5''}$), 29.3 ($C^{2''}$ or $C^{5''}$), 28.2 (N-CH_3), 27.2 ($C^{1'''}$), 23.6 ($C^{3''}$ or $C^{4''}$), 23.2 ($C^{3''}$ or $C^{4''}$), 21.5 (C^{15}), 9.2 ($C^{2'''}$), 8.6; LRMS *m/z* (ES^+) 550 ($[M+H]^+$, 40%); HRMS *m/z* (ES^+) found 550.31395 (100%), $C_{29}H_{39}O_4N_7$ requires $[M+H]^+$ 550.31363; Analytical HPLC (RP, method A) rt 2.05 mins, purity > 99%.

(*S*)-*N*-(1-acetylpiperidin-4-yl)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamide ((*S*)-155)



AcCl (5 µL, 0.032 mmol, 1.1 eq.) and NEt_3 (8 µL, 0.059 mmol, 2.0 eq.) were added to a solution of (*S*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(piperidin-4-yl)benzamide (**(*S*)-139**) (15 mg, 0.029 mmol, 1.0 eq.) in CH_2Cl_2 (300 µL) at 0 °C. The reaction mixture was stirred for 1 h, then diluted with H_2O (15 mL), and extracted with CH_2Cl_2 (3 × 10 mL). The combined organic components were washed with sat. aq. $NaHCO_3$ (30 mL) and brine (30 mL), then dried on $MgSO_4$, filtered, and concentrated *in vacuo*. Purification by silica gel chromatography, eluting with $CHCl_3$:MeOH (1:0 to 95:5), was followed by a MeCN:hexane grease wash to afford (*S*)-*N*-(1-acetylpiperidin-4-yl)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamide (**(*S*)-155**) (4 mg, 25%) as colourless solid: R_f 0.28 ($CHCl_3$:MeOH:Et_3N = 95:5:1); $[\alpha]_D^{25}$: +50.6 (*c* 0.38, MeOH); LRMS *m/z* (ES^+) 550 ($[M+H]^+$, 40%); HRMS *m/z* (ES^+) found 550.31273 (100%), $C_{29}H_{39}O_4N_7$ requires $[M+H]^+$ 550.31363; Analytical HPLC (RP, method A) rt 2.40 mins, purity 99%. The spectroscopic data are identical to the enantiomer **(*R*)-139**.

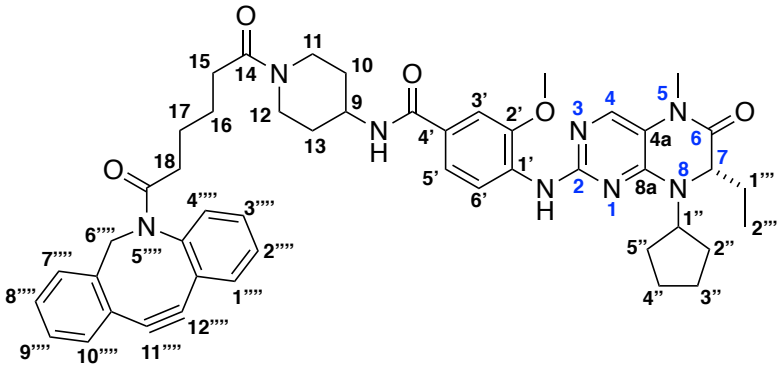
4-{[(7*R*)-8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl]amino}-*N*-{1-[6-(11,12-didehydrodibenzo[*b*,*f*]azocin-5(6*H*)-yl)-6-oxohexanoyl]piperidin-4-yl}-3-methoxybenzamide ((*R*)-154)



HATU (23 mg, 0.059 mmol, 1.2 eq.) and DIPEA (0.026 mL, 0.15 mmol, 3.0 eq.) were added to a solution of dibenzocyclooctyne-acid (20 mg, 0.059 mmol, 1.2 eq.) in CH_2Cl_2 (0.4 mL) at room temperature. The reaction mixture was stirred at room temperature for 1 h prior to addition of (*R*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-N-(piperidin-4-yl)benzamide (**(*R*)-139**) (25 mg, 0.049 mmol, 1.0 eq.). The reaction mixture was stirred for 1.5 h, then purification by silica gel chromatography (Combi*flash*Rf+, 4 g column), eluting with CH_2Cl_2:MeOH (1:0 to 9:1) afforded 4-{[(7*R*)-8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl]amino}-*N*-{1-[6-(11,12-didehydrodibenzo[*b*,*f*]azocin-5(6*H*)-yl)-6-oxohexanoyl]piperidin-4-yl}-3-methoxybenzamide (**(*R*)-154**) (25 mg, 76%) as an amorphous off-white solid: R_f 0.15 (EtOAc:EtOH = 99:1); $[\alpha]_D^{25}$: –35.6 (*c* 1.0, MeOH); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3282 (w), 2980 (m), 1634 (br, s) (C=O), 1610 (s) (C=O), 1605 (s) (C=O), 1471 (m), 1322 (m), 1251 (m), 753 (m); ^{1}H NMR (600 MHz, $CDCl_3$) δ 8.54 (1 H, dd, *J* 8.4, 3.3, $C^{6'}$-H), 7.69–7.57 (3 H, m, C^4-H, $C^{1''''}$-H or $C^{10''''}$-H, NH), 7.43 (1 H, dd, *J* 6.8, 1.9, $C^{3'}$-H), 7.26 (8 H, s, $C^{5'}$-H, $C^{1''''}$-H or $C^{10''''}$-H, $C^{2''''}$-H to $C^{4''''}$-H, $C^{7''''}$-H to $C^{9''''}$-H), 6.26–6.11 (1 H, m, NH), 5.13 (1 H, d, *J* 13.9, $C^{6''''}$-H_a), 4.56–4.44 (2 H, m, $C^{1''}$-H, C^{11}-H_a or C^{12}-H_a), 4.20 (1 H, dd, *J* 7.9, 3.7, C^7-H), 4.18–4.10 (1 H, m, C^9-H), 3.98–3.93 (3 H, m, OCH_3), 3.73–3.56 (2 H, m, $C^{6''''}$-H_b, C^{11}-H_a or C^{12}-H_a), 3.31 (3 H, s, NCH_3), 3.12 –2.89 (1 H, m, C^{11}-H_b or C^{12}-H_b), 2.73–2.57 (1 H, m, C^{11}-H_b or C^{12}-H_b), 2.27–2.18 (1 H, m, C^{15}-H_a or C^{18}-H_a), 2.18–1.76 (12 H, m, $C^{3''}$-H_a, $C^{4''}$-H_a, $C^{2''}$-H_a, $C^{5''}$-H_a, $C^{1'''}$-H_a, C^{15}-H_a, C^{16}-H_a, C^{17}-H_a, C^{18}-H_a, C^{16}-H_b or C^{17}-H_b, $C^{2''}$-H_b or $C^{5''}$-H_b, $C^{3''}$-H_b or $C^{4''}$-H_b), 1.76–1.63 (3 H, m, $C^{1'''}$-H_b, $C^{2''}$-H_b or $C^{5''}$-H_b, $C^{3''}$-H_b or $C^{4''}$-H_b), 1.55–1.20 (3 H, m, C^{15}-H_b, C^{18}-H_b, C^{16}-H_b or C^{17}-H_b), 0.86 (3 H, dd, *J* 7.5, $C^{2'''}$-H_3).

^{13}C NMR (151 MHz, $CDCl_3$) δ 173.17*, 173.15* (C^{18a}=O), 171.2 ($C^{4'a}$=O), 166.73, 166.70* ($C1^4$=O), 163.9 (C^6=O), 155.2 (C^2), 152.4 (C^{8a}), 151.9 ($C^{4''''a}$ or $C^{6''''a}$), 148.3 ($C^{4''''a}$ or $C^{6''''a}$), 147.3 ($C^{2'}$), 138.0 (C^4), 133.5 ($C^{1'}$), 132.4 ($C^{1''''}$ or $C^{10''''}$), 129.1 ($C^{4'}$), 128.5 ($C^{1''''}$ or $C^{3''''}$/$C^{4''''}$/$C^{7''''}$/$C^{8''''}$/$C^{10''''}$), 128.4 ($C^{3''''}$ or $C^{4''''}$/$C^{7''''}$/$C^{8''''}$), 128.16*, 128.13* ($C^{3''''}$ or $C^{4''''}$/$C^{7''''}$/$C^{8''''}$), 127.83*, 127.78* ($C^{3''''}$ or $C^{4''''}$/$C^{7''''}$/$C^{8''''}$), 127.2 ($C^{3''''}$ or $C^{4''''}$/$C^{7''''}$/$C^{8''''}$), 126.24*, 126.22* ($C^{10''''a}$), 125.5 ($C^{5'}$), 123.2 ($C^{2''''}$ or $C^{9''''}$), 122.7 ($C^{2''''}$ or $C^{9''''}$), 119.31*, 119.24* ($C^{11''''}$ or $C^{12''''}$), 116.5 (C^{4a}), 116.0 ($C^{6'}$), 115.1 ($C^{1''''a}$), 109.2 ($C^{3'}$), 108.0 ($C^{11''''}$ or $C^{12''''}$), 60.0 (C^7), 58.5 ($C^{1''}$), 56.1 (OCH_3), 55.5 ($C^{6''''}$), 47.23*, 47.17* (C^9), 44.68*, 44.64* (C^{11} or C^{12}), 40.91*, 40.85* (C^{11} or C^{12}), 34.63*, 34.61* (C^{15} or C^{18}), 33.12*, 33.03* (C^{15} or C^{18}), 32.96*, 32.90* (C^{10} or C^{13}), 32.09*, 32.02* (C^{10} or C^{13}), 29.9 ($C^{2''}$ or $C^{5''}$), 29.4 ($C^{2''}$ or $C^{5''}$), 28.3 (NCH_3), 27.3 ($C^{1'''}$), 25.09*, 25.02* (C^{16} or C^{17}), 24.77*, 24.65* (C^{16} or C^{17}), 23.7 ($C^{3''}$ or $C^{4''}$), 23.3 ($C^{3''}$ or $C^{4''}$), 9.3 ($C^{2'''}$); LCMS (high pH) rt 1.37 mins (94%), *m/z* (ES^+) 823 ($[M+H]^+$, 100%); HRMS *m/z* (ES^+) found 823.4274 (100%), $C_{48}H_{54}N_8O_5$ requires $[M+H]^+$ 823.4290. Chiral HPLC (Whelk-O1 (*R,R*) (250 × 4.6 mm, 5 µm), MeOH:MeCN = 50:50 with 0.1% *v/v* isopropylamine, 1.0 mL/mins, 25 mins) rt 15.1 mins, 52.8%, rt 17.6 mins, 47.2%. * Differences between two major rotamers can be seen for some of the resonances in the ^{13}C NMR. The peak showing split resonances of equal intensity are reported as the two observed chemical shifts.

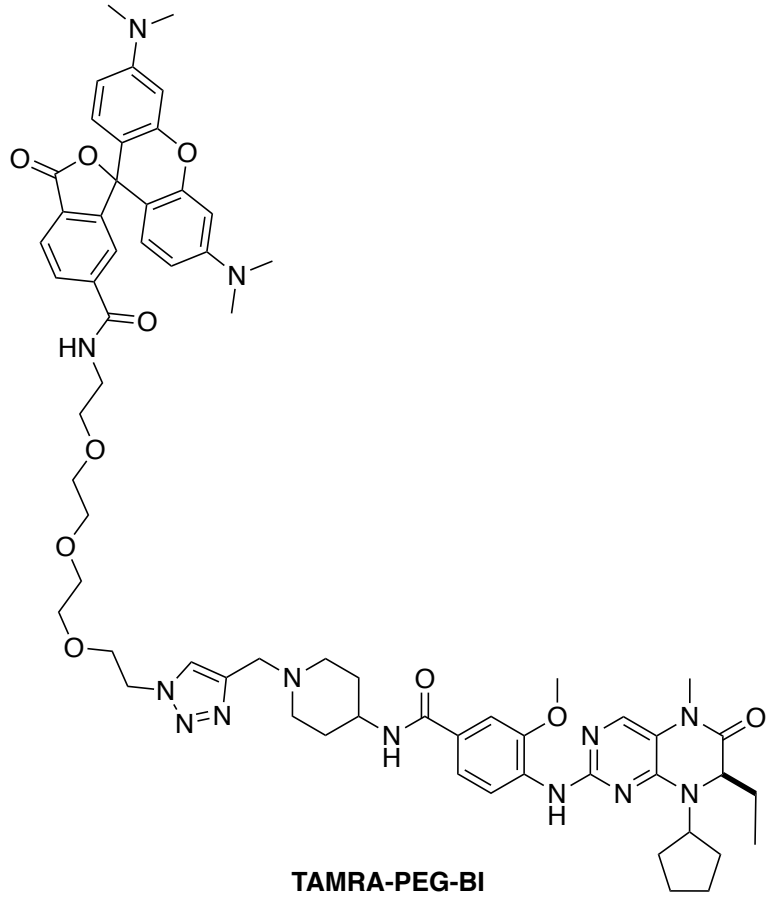
4'-{[(7*S*)-8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl]amino}-*N*-{1-[6-(11,12-didehydrodibenzo[*b,f*]azocin-5(6*H*)-yl)-6-oxohexanoyl]piperidin-4-yl}-3'-methoxybenzamide ((*S*)-154)



HATU (23 mg, 0.059 mmol, 1.2 eq.) and DIPEA (0.026 mL, 0.15 mmol, 3.0 eq.) were added to a solution of dibenzocyclooctyne-acid (20 mg, 0.059 mmol, 1.2 eq.) in CH_2Cl_2 (0.4 mL) at room temperature. The reaction mixture was stirred at room temperature for 1 h prior

to addition of (*S*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-N-(piperidin-4-yl)benzamide (**(*S*)-139**) (25 mg, 0.049 mmol, 1.0 eq.). The reaction mixture was stirred for 1.5 h, then purification by silica gel chromatography (Combi*flashRf*+, 4 g column), eluting with CH_2Cl_2:MeOH (1:0 to 9:1) afforded 4-{[(*S*)-8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl]amino}-*N*-{1-[6-(11,12-didehydrodibenzo[*b,f*]azocin-5(6*H*)-yl)-6-oxohexanoyl]piperidin-4-yl}-3-methoxybenzamide (**(*R*)-154**) (24 mg, 75%) as an amorphous off white solid: R_f 0.15 (EtOAc:EtOH = 99:1); $[\alpha]_D^{25}$: +28.4 (c 0.5, MeOH); LCMS (high pH) rt 1.37 mins (88%), *m/z* (ES^+) 824 ($[M+H]^+$, 100%); Chiral HPLC (Whelk-O1 (*R,R*) (250 × 4.6 mm, 5 µm), MeOH:MeCN = 50:50 with 0.1% *v/v* isopropylamine, 1.0 mL/mins, 25 mins) rt 15.1 mins, 54.0%, rt 17.6 mins, 46.0%. The spectroscopic data are identical to the enantiomer **(*R*)-139**.

(*R*)-N-(2-(2-(2-(2-(4-((4-(4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidin-1-yl)methyl)-1H-1,2,3-triazol-1-yl)ethoxy)ethoxy)ethoxy)ethyl)-3',6'-bis(dimethylamino)-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthene]-6-carboxamide ((*R*)-137)



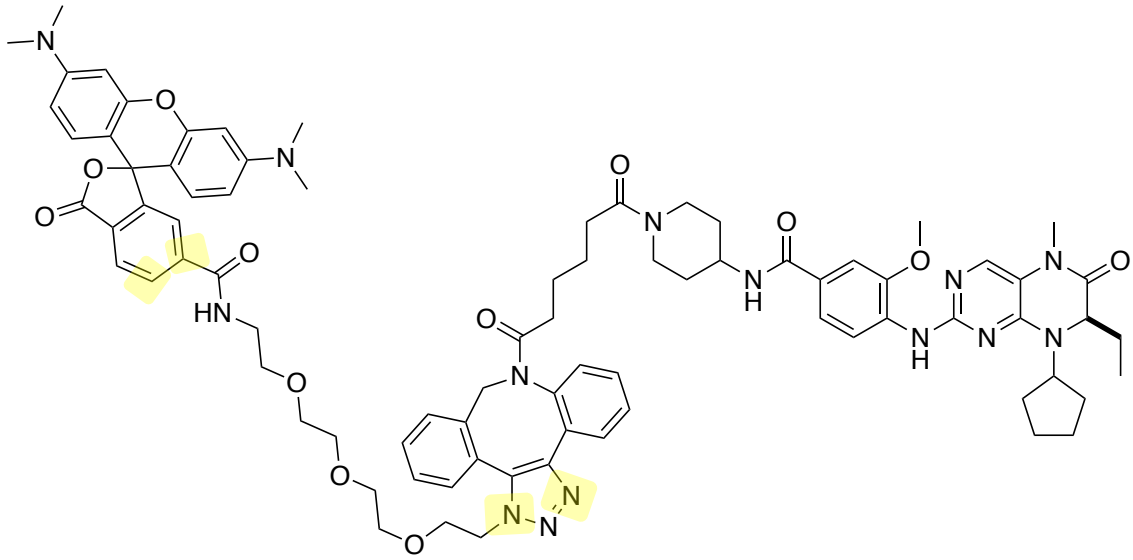
TAMRA-PEG-BI

(*R*)-4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(1-(prop-2-yn-1-yl)piperidin-4-yl)benzamide (**(*R*)-153**) (7.3 mg, 0.013 mmol, 1.2 eq.) and TAMRA-PEG-azide (7.0 mg, 0.011 mmol, 1.0 eq.) were dissolved in H_2O (300 µL) and ^{t}BuOH (300 µL). Tris(benzyltriazolylmethyl)amine (TBTA) (0.59 mg, 1.1 µmol, 0.10 eq.) in DMSO (0.150 mL) was added, followed by 0.1 M aq. copper(II) sulfate (0.011 mL, 1.1 µmol, 0.10 eq.) and 0.1 M aq. sodium ascorbate (0.022 mL, 2.2 µmol, 0.20 eq.) in that order. The reaction mixture was stirred at room temperature in the dark for 1 h. The crude was diluted to 1 mL with H_2O and purified by MDAP (method C). Collected fractions were frozen in tared vials and lyophilised. The fractions were combined to afford (*R*)-*N*-(2-(2-(2-(2-(4-((4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidin-1-yl)methyl)-1*H*-1,2,3-triazol-1-yl)ethoxy)ethoxy)ethoxy)ethyl)-3',6'-bis(dimethylamino)-3-oxo-3H-

spiro[isobenzofuran-1,9'-xanthene]-6-carboxamide (**(*R*)-137**) (5 mg, 37%) as an amorphous bright pink/red solid. LCMS (high pH) rt: 1.01 mins (92%), *m/z* (ES^+) 393 $[M+3H]^{3+}$, 30%), 589 $[M+2H]^{2+}$, 100%); HR-LCMS (formic acid) rt 2.93 mins (91%), *m/z* (ES^+) 1176.5957, $C_{63}H_{77}N_{13}O_{10}$ requires $[M+H]^+$ 1176.5995.

Note: The TAMRA-PEG-azide starting material is a mixture of 5- and 6- regioisomers. The regioselectivity of the copper click chemistry is expected to be approximately 9:1 according to the literature, hence a mixture of 2 main regioisomers, and 2 minor is expected. Partial separation of the regioisomers is seen in the LCMS trace.

(*R*)-*N*-(2-(2-(2-(2-(8-(6-(4-(4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidin-1-yl)-6-oxohexanoyl)-8,9-dihydro-1H-dibenzo[*b,f*][1,2,3]triazolo[4,5-d]azocin-[1/3]-yl)ethoxy)ethoxy)ethoxy)ethyl)-3',6'-bis(dimethylamino)-3-oxo-3*H*-spiro[isobenzofuran-1,9'-xanthene]-[5/6]-carboxamide ((*R*)-136)

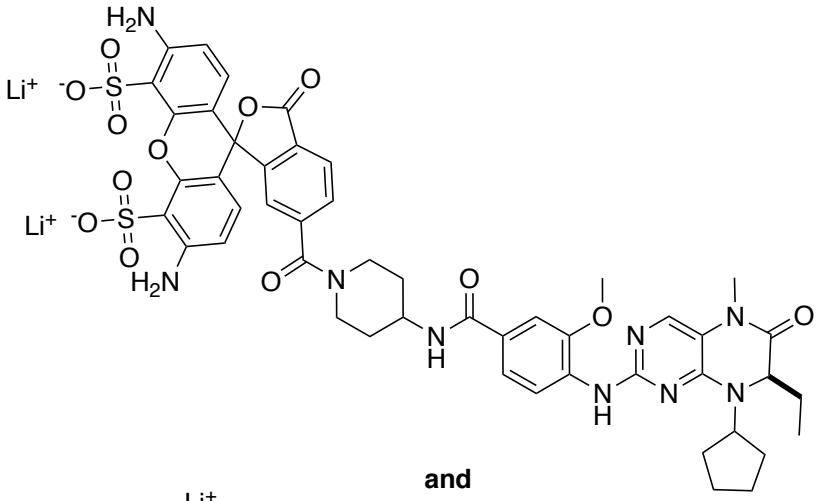


4-{[(7*R*)-8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl]amino}-*N*-{1-[6-(11,12-didehydrodibenzo[*b,f*]azocin-5(6*H*)-yl)-6-oxohexanoyl]piperidin-4-yl}-3-methoxybenzamide (**(*R*)-154**) (10 mg, 0.012 mmol, 1.2 eq.) was dissolved in acetonitrile (600 µL) and H_2O (300 µL), then added to TAMRA-PEG-azide (6.6 mg, 0.010 mmol, 1.0 eq.). The bright red mixture was stirred at room temperature in the dark for 2 h. The crude was diluted to 1 mL with acetonitrile and purified by MDAP, high pH method C (extended method). Fractions collected by automation and directly from the waste were

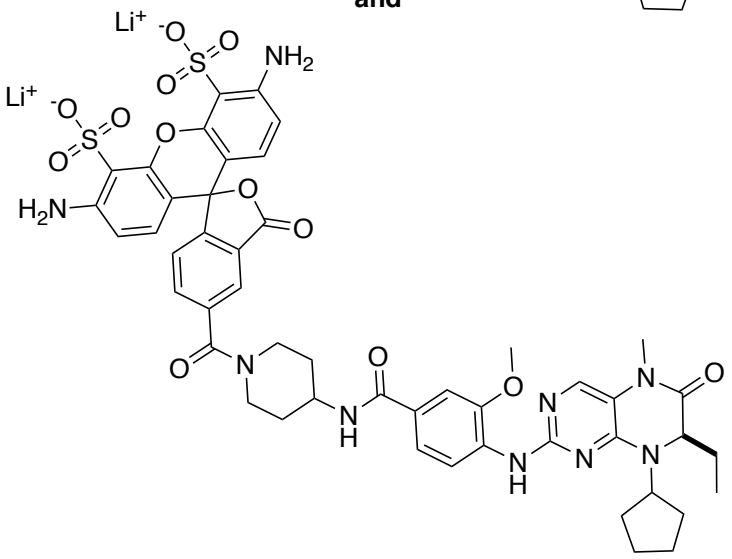
frozen at —80 °C in tared vials and then lyophilised. The fractions were combined to afford (*R*)-*N*-(2-(2-(2-(2-(8-(6-(4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidin-1-yl)-6-oxohexanoyl)-8,9-dihydro-1*H*-dibenzo [b,f][1,2,3]triazolo[4,5-d]azocin-1-yl)ethoxy)ethoxy)ethoxy)ethyl)-3',6'-bis(dimethylamino)-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthene]-6-carboxamide (**(*R*)-136**) (6.7 mg, 34%) as an amorphous bright pink/red solid. LCMS (high pH) rt 1.15 mins (40%), *m/z* (ES^+) 728 $[M+2H]^{2+}$, 60%), rt 1.16 mins (44%), *m/z* (ES^+) 728 $[M+2H]^{2+}$, 65%); HR-LCMS (formic acid) rt 4.05 mins (42%), *m/z* (ES^+) 727.3581, $C_{81}H_{92}N_{14}O_{12}$ requires $[M+2H]^{2+}$ 727.3588. Manual HR-LCMS peak integration used to detect 1453.7068, $C_{81}H_{92}N_{14}O_{12}$ requires $[M+H]^+$ 1453.7097. Two peaks are seen in the LCMS due to partial separation of the 5- and 6- regioisomers in the 10 mins LC method.

Note: The TAMRA-PEG-azide is a mixture of 5- and 6- regioisomers. The regioselectivity of the copper free click chemistry is approximately 1:1 according to the literature, hence a mixture of 4 regioisomers is expected. Partial separation of the regioisomers might be seen in the LCMS trace. The 4 different vectors are highlighted on the structures in yellow.

(*R*)-3',6'-diamino-[5/6]-(4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidine-1-carbonyl)-3-oxo-3*H*-spiro[isobenzofuran-1,9'-xanthene]-4',5'-disulfonate, dilithium salt ((*R*)-135)



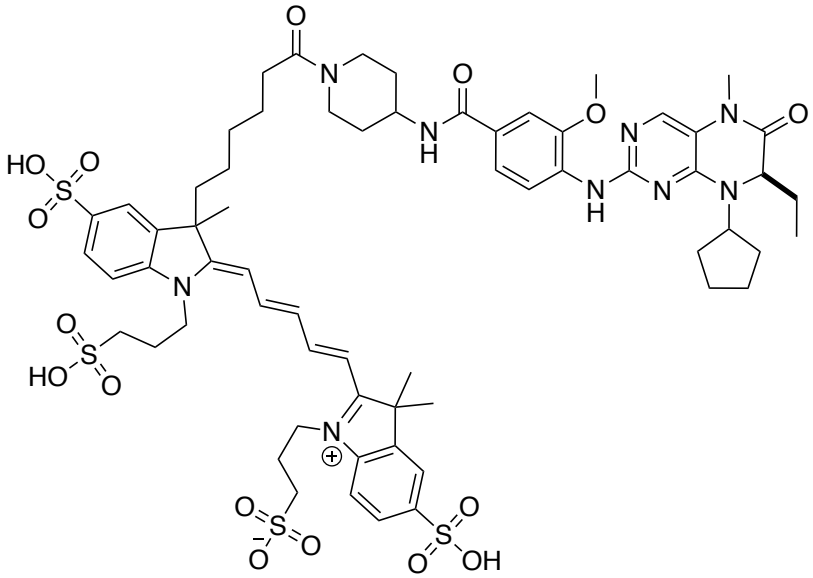
and



(*R*)-4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(piperidin-4-yl)benzamide ((*R*)-139) (2.9 mg, 5.8 µmol, 1.5 eq.) in DMSO (0.4 mL) was combined with Alexa Fluor™ 488 NHS Ester (3',6'-diamino-[5/6]-(((2,5-dioxopyrrolidin-1-yl)oxy)carbonyl)-3-oxo-3*H*-spiro[isobenzofuran-1,9'-xanthene]-4',5'-disulfonate) (2.5 mg, 3.9 µmol, 1.0 eq.) in 0.1 mL DMSO. DIPEA (4.1 µL, 23 µmol, 6.0 eq.) was added and the mixture was stirred in the dark for 18 h. PyBOP (2.0 mg, 3.9 µmol, 1.0 eq.) was added and the reaction mixture was stirred at room temperature for 30 mins in the dark. The crude material was diluted with of DMSO (1 mL) and purified by MDAP (method B). Collected fractions were frozen in tared vials and lyophilised. The fractions were combined to afford (*R*)-3',6'-diamino-6-(4-(4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidine-

1-carbonyl)-3-oxo-3*H*-spiro[isobenzofuran-1,9'-xanthene]-4',5'-disulfonate, dilithium salt (**(*R*)-135**) (1.5 mg, 34%) as an amorphous bright yellow solid: LCMS (high pH) rt: 0.74 mins (100%), *m/z* (ES^+) 513.3 $[M+2H]^{2+}$, 45%), 521.8 $[M+H+Na]^{2+}$, 100%); HR-LCMS (formic acid) rt: 3.54 mins (70%), *m/z* (ES^+) 1024.2938, $C_{48}H_{49}N_9O_{13}S_3$ requires $[M+H]^+$ 1024.2970; Two peaks seen in the LCMS due to partial separation of the 5- and 6-regioisomers in the 10 mins LC method. LCMS consistent with product with purity > 90%.

3-(2-((1*E*,3*E*)-5-((*Z*)-3-(6-(4-(4-(((R)-8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidin-1-yl)-6-oxohexyl)-3-methyl-5-sulfo-1-(3-sulfopropyl)indolin-2-ylidene)penta-1,3-dien-1-yl)-3,3-dimethyl-5-sulfo-3H-indol-1-ium-1-yl)propane-1-sulfonate, tetraammonium salt ((*R*)-134)

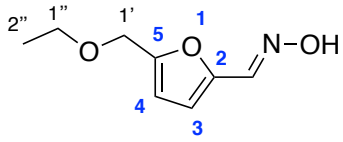


AF647-BI

Alexa Fluor™ 647 NHS Ester ((3-(2-((1*E*,3*E*)-5-((*Z*)-3-(6-((2,5-dioxopyrrolidin-1-yl)oxy)-6-oxohexyl)-3-methyl-5-sulfo-1-(3-sulfopropyl)indolin-2-ylidene)penta-1,3-dien-1-yl)-3,3-dimethyl-5-sulfo-3H-indol-1-ium-1-yl)propyl)(hydroxy)(oxo)-λ6-sulfanolate) (2.5 mg, 2.6 µmol, 1.0 eq.) in DMSO (200 µL) was added to a solution of (*R*)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(piperidin-4-yl)benzamide (**(*R*)-139**) (2.0 mg, 3.9 µmol, 1.5 eq.) in DMSO (300 µL). DIPEA (2.8 µL, 0.016 mmol, 6.0 eq.) was added and the reaction mixture was stirred at room temperature in the dark for 18 h. PyBOP (1.5 mg, 2.9 µmol, 1.1 eq.) was added and the mixture was stirred in the dark at room temperature for 1.5 h. The crude was diluted to 1 mL of DMSO

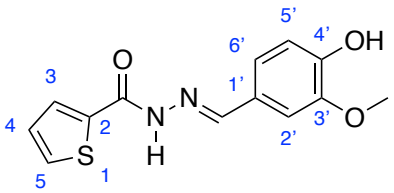
and purified by MDAP, (method A). Collected fractions were frozen in tared vials (each fraction split into 2 vials) and lyophilised to give 3-(2-((1*E*,3*E*)-5-((*Z*)-3-(6-(4-(4-(((*R*)-8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamido)piperidin-1-yl)-6-oxohexyl)-3-methyl-5-sulfo-1-(3-sulfopropyl)indolin-2-ylidene)penta-1,3-dien-1-yl)-3,3-dimethyl-5-sulfo-3*H*-indol-1-ium-1-yl)propane-1-sulfonate, tetraammonium salt (**(*R*)-134**) (2.6 mg, 62%) as an amorphous bright blue solid: LCMS (high pH) rt 0.65 mins (100%), *m/z* (ES^+) 449 $[M+2H]^{2+}$, 100%), 673 $[M+3H]^{3+}$, 50%); HRMS *m/z* (ES^+) 1348.4712 (100%)., $C_{63}H_{82}N_9O_{16}S_4$ requires M^+ 1348.4841.

(E)-5-(Ethoxymethyl)furan-2-carbaldehyde oxime (15)



Hydroxylamine hydrochloride (270 mg, 3.89 mmol, 1.20 eq.) was added to a solution of 5-(ethoxymethyl)furan-2-carbaldehyde (455 µL, 3.25 mmol, 1.00 eq.) in EtOH (50 mL) at 0 °C. The reaction mixture was stirred at 0 °C for 1.5 h, then Na_2SO_4 (*ca.* 2 g) was added, then the volatiles were removed under a flow of nitrogen, keeping the mixture cold with an ice bath. The residue was suspended in EtOH (20 mL) and filtered, washing with additional EtOH (3 × 20 mL). The filtrate was concentrated using a flow of nitrogen, then the crude material was stored at –20 °C overnight. Purification by silica gel chromatography, eluting hexane and EtOAc (20 CV at 19:1, then a gradient over 6 CV from 19:1 to 8:2), afforded (*E*)-5-(ethoxymethyl)furan-2-carbaldehyde oxime (**15**) (74 mg, 14%) as a colourless oil: R_f 0.22 (hexane:EtOAc = 8:2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.24 (1 H, s, O-H), 7.97 (1 H, s, -C(H)=N-), 6.58 (1 H, d, *J* 3.4, C^3-H), 6.41–6.34 (1 H, m, C^4-H), 4.47 (2 H, s, $C^{1'}$-H_2), 3.56 (2 H, q, *J* 7.0 $C^{1''}$-H_2), 1.23 (3 H, t, *J* 7.0, $C^{2''}$-H_3). LRMS m/z (ES^+) 192 ($[M+Na]^+$, 100%); Analytical HPLC (RP, method A) rt 7.95 mins, purity 95%. The data are in good agreement with the literature.[69] Storage neat at –20 °C due to partial isomerisation to the (*Z*)-isomer when stored in DMSO, MeOH or $CHCl_3$.

(*E*)-*N*'-(4-Hydroxy-3-methoxybenzylidene)thiophene-2-carbohydrazide (16)



Trifluoroacetic acid (10 µL, 0.13 mmol, 0.03 eq) was added to a solution of 4-hydroxy-3-methoxybenzaldehyde (0.71 g, 4.7 mmol, 1.0 eq) and thiophene carboxylic acid hydrazine (0.76 g, 5.35 mmol, 1.2 eq) in MeOH (20 mL). The reaction mixture was stirred for 4 h at RT. H_2O (20 mL) was added, the suspension was left to settle for overnight, then the precipitate was collected by filtration to afford (*E*)-*N*'-(4-hydroxy-3-methoxybenzylidene)thiophene-2-carbohydrazide (**16**) (896 mg, 70%) as a colourless solid: R_f 0.35 (petroleum ether:EtOAc = 1:1); mp 206–208°C (MeOH) [lit.[144] 206 °C (EtOH)]; ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.67 (1 H, s, N-H), 9.55 (1 H, s, O-H), 8.32 (1 H, s, –C(H)=N), 8.06 – 7.91 (2 H, m, $C^{2'}$-H or $C^{6'}$-H), 7.91 – 7.78 (1 H, m, $C^{2'}$-H or $C^{6'}$-H), 7.43–7.25 (1 H, m, C^3-H or C^4-H or C^5-H), 7.21 (1 H, dd, *J* 5.0, 3.8, C^3-H or C^4-H or C^5-H), 7.19–7.03 (1 H, m, C^3-H or C^4-H or C^5-H), 6.85 (1 H, d, *J* 8.1, $C^{5'}$-H), 3.84 (3 H, br s, OCH_3). LRMS *m/z* (ES^+) 299 ($[M+Na]^+$, 100%), 277 ($[M+H]^+$, 40%); Analytical HPLC (RP, method A) rt 4.74 mins, purity 95%. The data are in good agreement with the literature.[70, 144]

7.3. Biological method and materials

Biological materials were obtained from the following suppliers:

Material	Supplier
Bacto agar	Fisher Scientific
SDS-PAGE reagents	Bio-Rad and Sigma
IPTG, kanamycin and ampiccilin	Apollo Scientific
HEPES free acid > 99.5 %	Apollo Scientific
96-well PCR	Perkin Elmer
Thermowell Sealing Tape	Corning
Coomasie blue	Fisher Scientific
β-mercaptoethanol	Sigma Aldrich
organic solvents (EtOH, MeOH)	Fisher Scientific and Sigma Aldrich
DMSO	Fisher Scientific
10× TBS and TTBS	ALFA AESAR
CHAPS	Apollo Scientific
BSA	Apollo Scientific
EDTA	Sigma Aldrich
Imidazole	Sigma Aldrich

Purified (Milli-Q) water was obtained from a Millipore Elix® Reverse Osmosis system which was further purified by a Millipore Milli-Q® Synthesis system with a 0.22 µm filter on the outlet. Reagent quantities given as weight per volume (*w/v*) correspond to g/mL unless otherwise stated.

pH measurements were performed using a VWR pH 100 pH meter or a Jenway pH meter 3305, with an Aldrich glass/calomel combination electrode. Calibration was carried out between pH 4.0–7.0 or 7.0–10.0 immediately before use, with buffer solutions of phthalate (pH 4.0), phosphate (pH 7.0) and borate (pH 10.0) from Fischer Scientific. Electrodes were stored in a 4 M aqueous potassium chloride solution.

Centrifugation: Protein samples of up to 2.0 mL were centrifuged in a Beckman Microfuge 22R Centrifuge (14000 rpm) at 4 °C. Cell growth media and cell lysates were centrifuged in a Beckman Avanti J-25 (8500 rpm: JA-10 rotor; 15000 rpm: JA-16.250 rotor and 24000 rpm: JA-25.50 rotor, respectively).

Incubation of cells was performed in a New Brunswick G-25 Shaker Incubator at 37 °C and 180 rpm. Optical density at 600 nm was measured with a Novaspec III visible spectrophotometer.

7.3.1. DNA cloning

DNA fragments were purchased from ThermoFisher Life Technologies through GeneArt® gene synthesis and diluted to 20 ng/µL.

7.3.1.1. DNA amplification – polymerase chain reaction (PCR)

DNA fragments were amplified using Q5® High-Fidelity DNA Polymerase (New England Biolabs) following the recommended procedure. The reaction 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 [Tm of primer + 3] °C for 10 seconds, extension at 72 °C for 30 seconds/kb, then final extension at 72 °C for 2 minutes. The primer melting temperatures were calculated using NEB Tm calculator.

7.3.1.2. 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.

7.3.1.2.a. DNA assembly

Cloning was performed using a one-step isothermal *in vitro* recombination reaction consisting of T5 exonuclease, Phusion DNA polymerase and Taq DNA ligase. Linear vectors (~0.03 pmol) and gene inserts (~0.1 pmol) with complementary regions were incubated at 50 °C for 60 min in a 1:3 ratio with Gibson Assembly master mix (NEB): 20 µL reaction volume, 10 µL 2 × Gibson assembly (GA) master mix. 2 µL of the reaction mixture was used directly to transform NEB 5α (#C2987) competent cells. Competent cells were thawed on ice and aliquoted to 20 µL before 1.0 to 2.0 µL of plasmid DNA was added. The tubes were left on ice for 10 to 30 min before placement into a water bath at 42 °C ('heat shocked') for 30 s. The tubes were returned to ice for 5 mins, diluted with 500 µL 2× TY and incubated at 37 °C for 1 h. 10 to 100 µL of the incubated mixture was plated onto agar plates containing the appropriate antibiotic and incubated overnight.

7.3.1.2.b. Plasmid DNA isolation - MiniPrep

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 forward sequence primer was used. DNA sequence was analysed on CLC Main Workbench (v7.0). In the case of poor sequencing data towards the end of the sequencing, addition sequencing was done using the T7 promoter reverse primer.

7.3.2. Protein expression and purification

Media and Growth Conditions: *Escherichia coli* (*E. coli*) strains were cultured in 2× TY medium (16 g/L Oxoid™ Tryptone, 10 g/L Oxoid™ Yeast Extract, 5 g/L NaCl) at 37 °C and 180 rpm. When solid medium was required, Bacto agar swas added to a final concentration of 1.5% (*w/v*). Standard sterile practices were followed using an open flame, on sterilised surfaces with aq. EtOH (70% *v/v*). Media and equipment were sterilised by autoclaving at 121 °C (*Thermo Live Sciences* MAT90 LEI) for 15 mins. Solutions of IPTG and antibiotics were sterilised by filtering through 0.22 µm filters (Minisart®, Sartorius Stedim). Final concentrations of kanamycin and ampicillin were 50 and 100 µg/mL respectively, unless indicated otherwise.

E. coli bacterial strains of the following genotypes were used:

- NEB 5α: for plasmid amplification purposes.
- BL21 (DE3): for protein expression. DE3 denotes a chromosomal copy of the T7 RNA polymerase gene, transcribing genes under T7 promoter.

Competent cells were prepared according to the published $CaCl_2$-mediated transformation of *E. coli* protocol. Competent cells were thawed on ice and aliquoted to 20 µL before 1.0–2.0 µL of plasmid DNA was added. The tubes were left on ice for 10 to 30 min before placement into a water bath at 42 °C ('heat shocked') for 30 s. The tubes were returned to ice for 5 mins, diluted with 500 µL 2× TY and incubated at 37 °C for 1 h. 10–100 µL of the incubated mixture was plated onto agar plates containing the appropriate antibiotic and incubated overnight. Using the $CaCl_2$ method, plasmid DNA was transformed into *E. coli* BL21(DE3) cells for protein purification and NEB 5-α cells for plasmid amplification purposes.

Protein expressions of truncated H_6-*Tc*BDF2(1..113aa) and H_6-*Tc*BDF3(32..148aa) were first optimised by Dr Anthony Chan. The proteins were stored at –80 °C in aliquots in buffer (50 mM HEPES, 500 mM NaCl, pH 7.6) following flash freezing in liquid nitrogen. When required, aliquots were thawed on ice prior to use. Once thawed, sampled were not refrozen.

7.3.2.1. Bacterial starter cultures

Following transformation of the plasmid DNA into *E. coli* (see above), colonies were picked from the agar plates, and transfected into 10–100 mL of 2× TY media with the appropriate antibiotic. The starter culture were incubated at 37 °C at 180 rpm for 18 h. The starter cultures were used for the inoculation of fresh 2× TY media. Glycerol stocks (25% *v/v*), made as a 1:1 mix of starter culture and sterile filtered aq. glycerol (50% *v/v*), were flash frozen with liquid nitrogen, and stored at –80 °C. When required, glycerol were used to transfect 10–100 mL of 2× TY media to produce fresh starter cultures as described.

7.3.2.2. Large-scale culture and protein expression

Large-scale recombinant protein expression was performed using 2 L Turnair® flasks containing 1 L of 2× TY medium containing the appropriate antibiotic. Flasks were inoculated with 10 mL of the bacterial starter culture and incubated at 37 °C at 180 rpm. When the cultures reached an optical density OD_{600} of 0.6–1.0, the shaker temperature was equilibrated to 16 °C. The cultures were induced with IPTG (0.1–0.5 mM) then incubated for 16–20 h. After this time, cultures were harvested by centrifugation, and the resulting bacterial pellet was stored at –80 °C.

7.3.2.3. Small-scale protein expression trials

The plasmid DNA was transformed into *E. coli* BL21(DE3) cells, using the $CaCl_2$ method and spread on 2× TY Bacto agar plates containing kanamycin. One colony was incubated in 10 mL 2× TY with kanamycin at 37 °C for 16–20 h, then a portion of the overnight starter culture (100 µL) was subcultured in 2× TY (10 mL) with kanamycin for 3h. When the bacterial culture reached an optical density OD_{600} of 0.5– 1.2, IPTG was added to culture to obtain a final IPTG concentration of 0.1 mM unless otherwise stated. The proteins were expressed at 16 °C for 16–20 h. 4 mL of each culture was centrifuged for 1 mins, and the supernatant was discarded. The cells were lysed using 300 µL BugBuster® Protein Extraction Reagent (EMD Millipore Corp.)

(vortexed for 30 s, 3 times over 30 mins, incubating on ice) then centrifuged for 1 min. The supernatant, also referred to as the soluble fractions, was collected. Aliquots (40 µL) were mixed with 2× Laemmli dye (40 µL, containing 10% ß-ME) . The insoluble pellet was solubilised using MQ water (150 µL) and 2× Laemmli dye (150 µL, containing 10% ß-ME). The samples were heated at 100°C for 10 mins prior to loading of the SDS-PAGE gels. SDS-PAGE gel (16% resolving gel) were run (200 mV, 60–90 mins) then visualised using Coomassie blue stain.

7.3.2.4. Protein purification

The buffers used for protein purification are detailed in Table 7.1. Cell pellets were resuspended in extraction buffer (100 mL per 20 g cell mass), 25 Units (U) of Benzoase was added, then the cells were lysed by sonication using thirty 5 sec bursts interrupted by 5 to 10 sec pauses. Polyethyleneimine (PEI, 0.15% *v/v*) was added, and the cells were incubated on ice for 15 mins. The lysate was clarified by centrifugation and the supernatant decanted, filtered through a 0.45 µm filter, and applied to a purification column.

Table 7.1: Buffer compositions for protein purification

	IMAC buffers (pH 7.6)					SEC buffer (pH 7.6)
Reagents	Binding	Extraction	Wash	Elution	Strip	SEC elution
HEPES	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM
NaCl	0.5 M	0.5 M	0.5 M	0.5 M	0.5 M	0.5 M
Imidazole	5 mM	5 mM	30–60 mM	500 mM	–	–
β-ME	10 mM	–	10 mM	10 mM	–	–
EDTA	–	–	–	–	100 mM	–
Sigma*Fast* protease inhibitor tablet	–	1	–	–	–	–

Immobilised metal affinity chromatography (IMAC) was performed with a HisTrap™ (6 mL) column (GE Healthcare) charged with 4 CV of 100 mM $NiSO_4$ and washed with 10 CV of binding buffer prior to loading the cell lysate at 1 mL/min. The column was then washed with approximately 20 CV binding buffer. When all residual products of bacterial fermentation were eluted from the column by binding buffer, the protein of interest was eluted with 10 CV gradient of 0 to 100% of elution buffer. Eluted protein was collected and fractionated; fractions containing the highest levels of pure protein (as determined by the FPLC UV trace) were concentrated using 5 or 10K MWCO spin concentrators (GE Healthcare) in a Beckmann Allegra™ 21R centrifuge.

Further purification was achieved by size exclusion chromatography (SEC) gel filtration. Concentrated fractions (<2 mL) were loaded on a 120 or 300 mL Superdex 75 size exclusion chromatography column (Amersham). Proteins were eluted with SEC buffer (1.1 CV). The fractions containing the highest levels of pure protein (as determined by SDS-PAGE gels) were collected and concentrated (to 50–1500 µM) as described before.

Protein concentration was estimated by measuring the absorbance at 280 nm using a Nanodrop® ND-1000 spectrophotometer (Nanodrop® Technologies Inc.) with the 'Protein A280' program module according to the manufacturer's instructions. Protein molecular weight and molar extinction coefficient were calculated using ProtParam on the ExPASy Bioinformatics Resource Portal. Protein mass spectrometry was carried out on a Waters LCT Premier mass spectrometer diluted to 0.02–0.1 mg/mL with gel filtration buffer.

7.3.2.5. SDS-PAGE

Polyacrylamide gel electrophoresis (PAGE) was used to determine protein purity. Gels were prepared using 70 mm × 100 mm glass plates, with 0.75 mm spacers. Where possible, gels were cast (Table 7.2), and run in duplicates.

Table 7.2: Gel composition (quantities for two gels)

Reagent	16% Resolving gel	4% Stacking gel
Milli-Q water	2.1 mL	2.9 mL
1.5 m Tris-HCl, pH 8.8	2.5 mL	–
0.5 m Tris-HCl, pH 6.8	–	1.3 mL
30% acrylamide/bis	5.3 mL	0.7 mL
10% *w/v* SDS	100 µL	50 µL
TEMED	8 µL	5 µL
10% *w/v* APS	100 µL	50 µL

TEMED and ammonium persulfate (APS) were added just prior to pouring the gels. The resolving gel was cast, then covered with a thin layer of methanol to ensure a level surface, which was removed before casting the stacking gel. Gels were run on a mini-PROTEAN Tetra Electrophoresis System (BioRad) at constant potential (Table **7.3**). Typically, the gel was run at 100 V for 10 to 15 mins, then at 200 V for the remaining of the run to ensure neat stacking. For denaturing PAGE analysis, samples were prepared by mixing with sample loading buffer (2× or 4×, containing 10% ß-ME) and incubated at 100 °C for 10 mins. All gels were loaded with a molecular weight ladder: colour protein standard, broad range (NEB). Following electrophoresis, gels were stained for 15 mins with InstantBlue™ Coomasie® stain, then de-stained until satisfactory contrast was obtained.

Table 7.3: Buffer compositions used for SDS-PAGE gels

Buffer	Reagent	Amount required per litre
SDS-PAGE running buffer	Tris-HCl	30 g
	glycine	144 g
	SDS	10 g
SDS-PAGE stain	Coomasie® brilliant blue	5 g
	acetic acid	100 mL
	methanol	500 mL
SDS-PAGE destaining solution	methanol	500 mL
	acetic acid	100 mL

7.3.2.6. Western blotting

Proteins in the electrophoresed gels were transferred onto PVDF membranes (Immun-Blot® PDVF membrane sandwiches, BioRad) using by a MiniTrans Blot® transfer system. Transfers were run for 1 h at 100 V (constant) and 350 mA. The membrane was washed with tris-buffered saline solution (TBS, 50 mM Tris-HCl, pH 7.6, 150 mM NaCl, 3 × 5 mins) then blocked for 1 h using dried non-fat milk powder (Blotting grade blocker, BioRad, 5% *w/v*) in tris-buffered saline solution. The membranes were washed with TBS-Tween (0.1% Tween, 3 × 5 mins), then incubated in blocking solution with H_6 epitope tag antibody (HRP conjugate HIS.H8, 1:5000 dilution) for 1 h at room temperature. The membranes were washed with TBS-Tween (0.1% Tween, 4 × 5 mins), then with TBS (2 × 5 mins) prior to visualisation using either colorimetric detection (Opti-4CN Substrate Kit, Bio-Rad) or chemiluminescence (Clarity™ Western ECL Substrate).

7.3.2.7. Circular dichroism (CD)

Circular dichroism measurements were acquired using an Jasco J-815 Spectropolarimeter (Biochemistry department, University of Oxford). All experiments were performed in a 1 mm path length cuvette using 100 μg/mL protein in 10 mM sodium phosphate buffer (pH 7.6).

7.3.3. Protein crystallisation screens

7.3.3.1. TEV cleavage

Removal of the Histag was achieved through cleavage of H_6-*Tc*BDF3 with TEV protease at room temperature or 4 °C for 2 to 20 h. Cleavage buffers were either (A) 50 mM TrisMaz base, 0.5 mM EDTA, 1 MM DTT at pH 8.0 or (B) HEPES 50 mM, NaCl 150 mM at pH 7.6. Cleavage was monitored by protein LCMS and SDS-PAGE gel followed by either Coomassie blue staining or Western Blotting and visualisation. To optimise the cleavage, different ratios of TEV protease to protein were investigated. Commercially available TEV proteases were bought from Promega®, RD BioTechne or InsightBio. The plasmid for pRK793 was provided by the group of Prof. Christopher J. Schofield as a glycerol stock of transfected *E. coli* BL21 (DE3) cells. The plasmid for pTH24 was provided by the group of Prof. Manfred Jung (expression and expression of the pTH24 construct was completed by Dr Matthias Schiedel).

For large scale studies, the TEV proteases were expressed in 2× TY medium containing a final concentration of 100 μg/mL of Ampicillin as described above. The TEV proteases were expressed in 2 to 4 L cultures and purified by IMAC and SEC columns as described above using the buffers described in Table 7.4

Table 7.4: Buffer compositions for TEV protein purification and cleavage

	IMAC buffers (pH 7.6)					SEC buffer (pH 7.6)
Reagents	Binding	Extraction	Wash	Elution	Strip	SEC elution
HEPES	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM
NaCl	0.5 M	0.5 M	0.5 M	0.5 M	0.5 M	0.5 M
Imidazole	5 mM	5 mM	30 to 60 mM	500 mM	–	–
TCEP	0.5 mM	–	0.5 mM	0.5 mM	–	0.5 mM
EDTA	–	–	–	–	100 mM	–
Sigma*Fast* protease inhibitor tablet	–	1	–	–	–	–

7.3.3.2. Protein crystallisation trials

The Phoenix crystallography system by Art Robbins Instruments was used to dispense the protein in 96-well IntelliPlates, with each buffer well accompanied by 3 protein and buffer drop to yield 3 protein concentrations per well. Commercial screens of buffer conditions were tested. Crystal fishing was done by Dr Michael McDonough (University of Oxford), and the crystal X-ray shooting was done in house (Chemistry department, University of Oxford) or at Diamond Light Source (Harwell Science and Innovation Campus). Plates were checked twice a week in the 2 week following set up, then once a week for a month, then every month for a further 2 months. Visual checks were done using a Leica M205-C microscope. Potential crystals were checked under polarised light to identify salt crystals.

7.3.4. Biophysical assays protocols

7.3.4.1. Differential scanning fluorimetry

The differential scanning fluorimetry (DSF) or thermal stability assays were carried out following the protocol described by Niesen *et al.*[50] and data was analysed as described previously.[50-51] Thermal melting experiments were carried out using the Mx3005p real-time PCR machine (Agilent) (Biochemistry department, University of Oxford) and employing a protein concentration of 2, 5 or 8 µM and an inhibitor concentration of 50 or 150 µM. Buffer conditions were 50 mM buffer, pH 5.5–10, 500 mM NaCl and 1:1000 to 1:250 dilution of SYPRO® Orange (Invitrogen, CA). The 96-well PCR plates were sealed and centrifuged for 2 min at 1000 rpm at 25 °C before measurement. Temperature was raised with a step of 1 °C per 30 s from 25 °C to 85 °C, and fluorescence readings were taken at each interval. Excitation and emission filters for the SYPRO® Orange dye were set to 465 and 590 nm, respectively. The observed temperature shift, ΔT_m, for each inhibitor was recorded as the difference between the transition midpoints of sample and reference wells containing protein without ligand in the same plate. All thermal stability shift assays were carried out in full collaboration with F. Woodhouse.

7.3.4.2. Water-Ligand Observed *via* Gradient Spectroscopy (WaterLOGSY)

Typical procedure for NMR WaterLOGSY: Protein in HEPES containing storage buffers were exchanged into Invitrogen 10 mM Na/K phosphate buffer (0.9% NaCl, pH 7.6) using GE Healthcare PD spintrap G-25 columns, eluting on two pre-equilibrated columns successively for each protein. Protein concentrations in the phosphate buffer were calculated from by measuring the absorbance at 280 nm using a Nanodrop® ND-1000 spectrophotometer (Nanodrop® Technologies Inc.) with the 'Protein A280' program module according to the manufacturer's instructions. Samples were prepared to a total sample volume of 160 µL phosphate buffer, with 500 µM compound (unless otherwise stated) from 50 mM DMSO-D6 stock solutions, 10% D_2O for signal locking and 20 µM protein (unless otherwise stated) for protein samples. After loading into 3 mm Hildenberg 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. For each compound, a water-supressed ^{1}H NMR spectrum was recorded as a reference, followed by a waterLOGSY spectrum. When the waterLOGSY of the compound showed suitable solubility, waterLOGSY spectra in the presence of the proteins of interest were recorded [H_6-*Tc*BDF2 (1..113aa), H_6-*Tc*BDF3 (32..148aa), H_6-*Tc*BDF5-1 (1..113aa)]. Training by Prof. Timothy Claridge and Dr Larissa See.

7.3.4.3. Ligand-observed proton NMR

Protein in HEPES containing storage buffers were exchanged into Invitrogen 10 mM Na/K phosphate buffer (0.9% NaCl, pH 7.6) using dialysis. Protein concentrations in the phosphate buffer were calculated from by measuring the absorbance at 280 nm using a NanoDrop Lite spectrophotometer (Nanodrop® Technologies Inc.) by using the predicted protein absorbance. Samples were prepared to a total sample volume of 160 µL phosphate buffer, with 50 µM compound from 5 mM DMSO-D_6 stock solutions, 10% D_2O for signal locking and the appropriate amount of protein. After loading into 3 mm Hildenberg 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 an inverse TCI cryoprobe or TXI probe using a standard excitation-sculpting water suppression pulse sequence (332 or 64 scans using the TCI cryoprobe; 256 or 512 scans using the TCI probe). For each compound, a compound-only sample was recorded as a reference, followed by samples with protein. Each spectrum was processed with Bruker Topsin® 3.2, the data was recorded on Microsoft Excel®, then the K_D was calculated using the "one site–specific biding; non-linear regression" function using Graphpad Prism®. In Topspin, spectra were phase corrected and calibrated to the residual DMSO peak, then local baseline corrections were performed ("absf n" function).

- *Tc*BDF3 or *Tc*BDF5

The chemical shifts of the peaks of interest and the DMSO reference peaks were measured. Using the evolution of the $\left|\partial_{peak} - \partial_{DMSO}\right|$ distance over the titration, the change in chemical shift of the peak of interest with increasing protein concentration was plotted for each compound.

- *Hm*BRD4(1)

The integral of the peaks of interest and the DMSO reference peaks were calculated in each spectrum for the determined chemical shift range (constant ranges were used for each studied compound). Using the ratio of integrals, the integral decrease of the peak of interest with increasing protein concentration was plotted for each compound.

7.3.4.4. Binding assay based on ^{19}F NMR

Typical procedure for ^{19}F NMR binding assay: The protein samples were thawed on ice and centrifuged. The protein concentration was determined by measuring the absorbance at 280 nm using a Nanodrop® ND-1000 spectrophotometer (Nanodrop® Technologies Inc.) with the 'Protein A280' program module according to the manufacturer's instructions. Samples were prepared to a total sample volume of either 500 or 160 µL phosphate buffer, with 100 µM compound (unless otherwise stated) from 10–20 mM DMSO-D_6 stock solutions, 10% D_2O for signal locking, 5 µL of a fluorinated compound reference in buffer, and protein (0–150 µM) for protein samples. The ^{19}F reference solution in buffer was either (a) 0.01 % (*v/v*) solution of 1,1,1-trifluoroacetone or (b) 10 mM KF. After loading into either 5 or 3 mm Hildenberg NMR tubes for Bruker Match™ system, samples were briefly centrifuged in a manual benchtop rotor. All samples were incubated at room temperature for a minimum of 10 min prior to their measurement. Spectra were recorded using a Bruker AVIII 600 MHz spectrometer with Prodigy N2 broadband cryoprobe (128–384 scans). Experiments on *Sm*BD3 were performed with Dr Matthias Schiedel. Initial experiments on *Tc*BDF5 were performed by Dr Anthony Chan

7.3.4.5. Isothermal titration calorimetry (ITC)

All calorimetric experiments were performed on a MicroCal PEAQ-ITC (Biochemistry department, University of Oxford) or MicroCal PEAQ-ITC Automated (Malvern) (Chemistry department, University of Oxford) and analysed with the MicroCal ORIGIN software package using a single binding site model. The first data point was excluded from the analysis. The ITC buffer was 50 mM HEPES, 150 mM NaCl at pH 7.4. Proteins were dialysed at 4 °C overnight in a Slide-A-Lyzer™ MINI Dialysis Device (2000 MWCO; Thermo Scientific Life Technologies) into

the assay buffer supplemented with DMSO (0.2% final concentration). Protein concentrations were determined by measuring the absorbance at 280 nm using a Nanodrop® ND-1000 spectrophotometer (Nanodrop® Technologies Inc.) with the 'Protein A280' program module according to the manufacturer's instructions (tyrosine ε_{280} = 1280 $M^{-1}cm^{-1}$). Protein samples were diluted to the required concentration using dialysis buffer. DMSO stock solutions of the compounds were diluted with the ITC buffer (containing no DMSO) to the appropriate concentration to give compound samples (0.2% final DMSO concentration). A typical ITC experiment required 160 µL of protein sample (syringue) and 420 µL of compound sample (cell). The cell was stirred at 750 rpm, with reference power set to 4 or 5 µcal/sec and temperature held at 298 °K. After an initial delay of 60 s, 19 × 2 µL injections (first injection 0.4 µL) were performed with a spacing of 180 s. Heats of dilution were measured under the same conditions and subtracted for analysis.

7.3.4.6. Fluorescence polarisation (FP)

All fluorescence polarisation experiments were performed using the BMG LABTECH CLARIOstar® or PHERAstar® *FS* microplate readers, processed using the BMG LABTECH Mars software and analysed using Graphpad Prism®. BSA (0.1% (*w/w*)) was added freshly to the FP assay buffer (25 mM HEPES, 100 mM NaCl, CHAPS (0.05%), pH = 7.6) before each experiment. The fluorescent intensity and polarization measurements were performed using black 384-well non-binding microplate (CORNING) in triplicate (20 or 40 µL). Training by Dr Matthias Schiedel.

7.3.4.6.a. Characterisation of the FP probe binding

Total fluorescence intensity:

Serial dilutions of the fluorescently-labelled probe (ranging from 1.25 to 640 nM) were made in assay buffer and dispensed in triplicate (20 µL). Blank controls contained only buffer (20 µL). Fluorescence intensities were recorded with a PHERAstar *FS*® or CLARIOSstar® plate reader (BMG LABTECH). For the fluorescein-labelled probes the following settings were used: excitation filter – FP 485 nm, emission filter 1 – FP p-pol

520 nm, emission filter 2 – FP s-pol 520 nm. Total fluorescence intensities were calculated using the formula $2 \times I_P + I_S$. The resulting values were plotted against the respective probe concentrations using GraphPad Prism and fitted with linear regression.

Binding curves and K_D determination:

A 2× concentrated serial dilution series (ranging from 0.0048 to 10 µM (final concentrations)) of the respective protein were made in buffer. 10 µL of the 2× protein dilutions and 10 µL of a 20–80 nM solution of the fluorescently-labelled probe (final assay concentration is 10–40 nM, according to the probe used) were dispensed in a black 384-well non-binding microplate (CORNING) resulting in a final assay volume of 20 µL per well. Fluorescence polarization (FP) was measured from these wells after incubation for 30 min at 25 °C with an PHERAstar FS plate reader (BMG LABTECH) using following settings: excitation filter – FP 485 nm, emission filter 1 – FP p-pol 520 nm, emission filter 2 – FP s-pol 520 nm. Fluorescence polarization (P) values were determined using the formula $\frac{(I_s - G \times I_p)}{(I_s + G \times I_p)} \times 1000\ (mP)$, where I_s is the fluorescence intensity measured in the s-plane and I_P is the fluorescence intensity measured in the p-plane, both blank corrected. G is the grating factor, which is an instrument specific constant. Finally, P was plotted to a binding curve with GraphPad Prism by using a one-site binding (hyperbola) equation:

$$Y = \frac{B_{max} \times [R]}{K_D + [R]}$$

where B_{max} is the maximal binding.

7.3.4.6.b. Ligand screening by FP

Fluorescence polarization-based competition assay:

IC_{50} determination experiments were performed in low-volume black 384 well plates, with a final volume of either 20 or 40 µL , as described below.

FP total assay volume	Protein mix (2 × concentration)	Probe mix (4 × concentration)	Inhibitor mix (4 × concentration)
20 µL	10 µL	5 µL	5 µL
40 µL	20 µL	10 µL	10 µL

Freshly-thawed protein aliquots were diluted with the appropriate amounts of FP buffer to yield a 2 × concentration protein mix. The FP probe, stored as a 10 mM stock in DMSO at –80 °C, was diluted using FP buffer to an 10 µM intermediate mix, which was diluted further to yield the desired 4 × probe mix. Serial dilutions (1:1) of the inhibitor DMSO stock solutions into DMSO were used to give the desired 4× inhibitor mixes. These dilutions contained 4% DMSO (final assay concentration: 1% DMSO), unless otherwise stated.

Positive controls (P_{pos}, complete binding of the fluorescent probe) were performed in triplicate and contained 10 µL of the 4 × solution of fluorescence-labelled probe (40–100 nM final concentration), 10 µL assay buffer containing an appropriate DMSO content, and 20 µL of the 2× concentrated solution of the respective protein. Negative controls (P_{neg}, unbound fluorescent probe) were performed in triplicate and contained only the fluorescently labelled probe (40–100 nM final concentration) in 40 µL of assay buffer.

The assay was set up either as a displacement or a prevented binding experiment by varying the order of addition of the assay components:

- For the displacement experiments, the probe and the protein were mixed and incubated on ice for 15 mins before the addition to the inhibitor mix.
- For the prevented binding experiments, the inhibitor and the protein were mixed and incubated on ice for 15 mins, and the probe was added just before the FP measurement was performed.

Inhibition values (I) were calculated using the following equation:

$$I = 100 \times \left(1 - \left(\frac{P_I - P_{neg}}{P_{pos} - P_{neg}}\right)\right)\%$$

7.3.4.7. Photo-crosslinking mass spectrometry assay

Protein aliquots were thawed at 0 °C and the protein of interest was diluted to 10 mL at 1 µM in 10% MQ PBS. Using an Labcyte Echo® 555 Liquid Handler 150 nL of each well in the parent plate (containing the PhaBits library at 20 mM DMSO stock solutions) was

transferred into a Greiner 384 low volume plate (white) to prepare the probe plate. The probe plate was placed on ice. 15 µL of protein stock solution (1 µM) was dispensed into wells A1–14 through to P1–14. The plate was left on ice for 15 minutes for incubation. The plate was then irradiated at 302 nm for 10 mins. The plate was centrifuged (1000 rpm, 1 min) to remove any bubbles. The plate was then sampled by LCMS-TOF mass spectrometry (see acquisition methods in the Supplementary Information).

Data Manipulation:

LCMS-TOF intact mass spectrometry afforded a peak for the unlabelled protein, and a peak for the labelled protein (corresponding to $M_{protein}$+probe–N_2). The csv. files of deconvoluted spectra results were analysed using R Studio software to generate pdf files of the spectra.

References:

1. Rassi Jr, A.; Rassi, A.; Marin-Neto, J. A., Chagas disease. *The Lancet* 2010, *375*, 1388-1402.
2. Chagas, C., Nova Tripanozomiaze Humana: Estudos Sobre a Morfolojia e o Ciclo Evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., Ajente Etiolojico de Nova Entidade Morbida do Homem. *Memórias do Instituto Oswaldo Cruz* 1909, *1*, 159-218.
3. *Research Priorities for Chagas Disease, Human African Trypanosomiasis and Leishmaniasis*. World Health Organisation: WHO Technical Report Series, 2012.
4. Gomes, C.; Almeida, A. B.; Rosa, A. C.; Araujo, P. F.; Teixeira, A. R. L., American Trypanosomiasis and Chagas Disease: Sexual Transmission. *International Journal of Infectious Diseases* 2019, *81*, 81-84.
5. Yeo, M.; Acosta, N.; Llewellyn, M.; Sánchez, H.; Adamson, S.; Miles, G. A. J.; López, E.; González, N.; Patterson, J. S.; Gaunt, M. W.; Arias, A. R. d.; Miles, M. A., Origins of Chagas Disease: Didelphis Species are Natural Hosts of *Trypanosoma cruzi* I and Armadillos hosts of *Trypanosoma cruzi* II, including Hybrids. *Int. J. Parasitol.* 2005, *35*, 225-233.
6. Bonney, K. M.; Luthringer, D. J.; Kim, S. A.; Garg, N. J.; Engman, D. M., Pathology and Pathogenesis of Chagas Heart Disease. *Annual Review of Pathology: Mechanisms of Disease* 2019, *14*, 421-447.
7. https://medmovie.com
8. Sgambatti de Andrade, A. L. S.; Zicker, F.; de Oliveira, R. M.; Almeida e Silva, S.; Luquetti, A.; Travassos, L. R.; Almeida, I. C.; de Andrade, S. S.; Guimarães de Andrade, J.; Martelli, C. M. T., Randomised Trial of Efficacy of Benznidazole in Treatment of Early *Trypanosoma cruzi* Infection. *The Lancet* 1996, *348*, 1407-1413.
9. Maya, J. D.; Cassels, B. K.; Iturriaga-Vásquez, P.; Ferreira, J.; Faúndez, M.; Galanti, N.; Ferreira, A.; Morello, A., Mode of Action of Natural and Synthetic Drugs against *Trypanosoma cruzi* and their Interaction with the Mammalian Host. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 2007, *146*, 601-620.
10. Filardi, L. S.; Brener, Z., Susceptibility and Natural Resistance of *Trypanosoma cruzi* Strains to Drugs Used Clinically in Chagas Disease. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1987, *81*, 755-759.
11. Brener, Z.; Costa, C. A. G.; Chiari, C., Differences in the Susceptibility of *Trypanosoma cruzi* Strains to Active Chemotherapeutic Agents. *Revista do Instituto de Medicina Tropical de Sao Paulo* 1976, *18*, 450-455.
12. Alonso, V. L.; Villanova, G. V.; Ritagliatib, C.; Mottac, M. C. M.; Cribba, P.; Serra, E. C., *Trypanosoma cruzi* Bromodomain Factor 3 Binds Acetylated α-Tubulin and Concentrates in the Flagellum during Metacyclogenesis. *Eukaryot. Cell* 2014, *13*, 822-831.
13. Tyler, K. M.; Engman, D. M., The life Cycle of *Trypanosoma cruzi* Revisited. *Int. J. Parasitol.* 2001, *31*, 472-481.
14. Alonso, V. L.; Serra, E. C., Lysine Acetylation: Elucidating the Components of an Emerging Global Signaling Pathway in Trypanosomes. *J. Biomed. and Biotech.* 2012, *2012*, 1-16.

15. Atwood, J. A.; Weatherly, D. B.; Minning, T. A.; Bundy, B.; Cavola, C.; Opperdoes, F. R.; Orlando, R.; Tarleton, R. L., The *Trypanosoma cruzi* Proteome. *Science* 2005, *309*, 473-476.
16. Campo, V. A., Comparative Effects of Histone Deacetylases Inhibitors and Resveratrol on *Trypanosoma cruzi* Replication, Differentiation, Infectivity and Gene Expression. *International Journal for Parasitology: Drugs and Drug Resistance* 2017, *7*, 23-33.
17. Waddington, C. H., Preliminary Notes on the Development of the Wings in Normal and Mutant Strains of Drosophila. *Proc. Natl. Acad. Sci. USA* 1939, *25*, 299-307.
18. Holliday, R., The Inheritance of Epigenetic Defects. *Science* 1987, *238*, 163-170.
19. Hewings, D. S.; Rooney, T. P. C.; Jennings, L. E.; Hay, D. A.; Schofield, C. J.; Brennan, P. E.; Knapp, S.; Conway, S. J., Progress in the Development and Application of Small Molecule Inhibitors of Bromodomain–Acetyl-lysine Interactions. *J. Med. Chem.* 2012, *55*, 9393-9413.
20. Allfrey, V. G.; Mirsky, A. E., Acetylation and Methylation of Histones and their Possible Role in the Regulation of RNA Synthesis. *Proc. Natl. Acad. Sci.* 1964, *51*, 786-794.
21. Gershey, E. L.; Vidali, G.; Allfrey, V. G., Chemical Studies of Histone Acetylation. The Occurrence of ε-N-Acetyllysine in the f2a1 Histone. *J. Biol. Chem.* 1968, *243*, 5018-5022.
22. Jennings, L. E.; Measures, A. R.; Wilson, B. G.; Conway, S. J., Phenotypic Screening and Fragment-Based Approaches to the Discovery of Small-Molecule Bromodomain Ligands. *Future Med. Chem.* 2014, *6*, 179–204.
23. Schiedel, M.; Moroglu, M.; Ascough, D. M. H.; Chamberlain, A. E. R.; Kamps, J. J. A. G.; Sekirnik, A. R.; Conway, S. J., Chemical epigenetics: the impact of chemical- and chemical biology techniques on bromodomain target validation. *Angew. Chem. Int. Ed.* 2019, *58*, 17930-17952.
24. Kouzarides, T., Acetylation: a Regulatory Modification to Rival Phosphorylation? *The EMBO Journal* 2000, *19*, 1176-1179.
25. Tamkun, J. W.; Deuring, R.; Scott, M. P.; Kissinger, M.; Pattatucci, A. M.; Kaufman, T. C.; Kennison, J. A., Brahma: A Regulator of Drosophila Homeotic Genes Structurally Related to the Yeast Transcriptional Activator SNF2SWI2. *Cell* 1992, *68*, 561-572.
26. Roberto Sanchez, J. M., Ming-Ming Zhou, The Bromodomain: From Epigenome Reader to Druggable Target. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 2014, *1839*, 676–685.
27. Shi, J.; Vakoc, C. R., The Mechanisms behind the Therapeutic Activity of BET Bromodomain Inhibition. *Mol. Cell* 2014, *54*, 728-736.
28. Filippakopoulos, P.; Knapp, S., The bromodomain interaction module. *FEBS Lett.* 2012, *586*, 2692-2704.
29. Wu, Q.; Heidenreich, D.; Zhou, S.; Ackloo, S.; Krämer, A.; Nakka, K.; Lima-Fernandes, E.; Deblois, G.; Duan, S.; Vellanki, R. N.; Li, F.; Vedadi, M.; Dilworth, J.; Lupien, M.; Brennan, P. E.; Arrowsmith, C. H.; Müller, S.; Fedorov, O.; Filippakopoulos, P.; Knapp, S., A chemical toolbox for the study of bromodomains and epigenetic signaling. *Nature Communications* 2019, *10*, 1915.
30. Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.; Lambert, J.-P.; Barsyte-Lovejoy, D.; Felletar, I.; Volkmer, R.; Müller, S.; Pawson, T.; Gingras, A.-C.; Arrowsmith, Cheryl H.; Knapp, S., Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell* 2012, *149*, 214-231.

31. Dey, A.; Nishiyama, A.; Karpova, T.; McNally, J.; Ozato, K., Brd4 Marks Select Genes on Mitotic Chromatin and Directs Postmitotic Transcription. *Molecular Biology of the Cell* 2009, *20*, 4899-4909.
32. Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E., Selective Inhibition of BET Bromodomains. *Nature* 2010, *468*, 1067-1073.
33. Chung, C.-w.; Coste, H.; White, J. H.; Mirguet, O.; Wilde, J.; Gosmini, R. L.; Delves, C.; Magny, S. M.; Woodward, R.; Hughes, S. A.; Boursier, E. V.; Flynn, H.; Bouillot, A. M.; Bamborough, P.; Brusq, J.-M. G.; Gellibert, F. J.; Jones, E. J.; Riou, A. M.; Homes, P.; Martin, S. L.; Uings, I. J.; Toum, J.; Clément, C. A.; Boullay, A.-B.; Grimley, R. L.; Blandel, F. M.; Prinjha, R. K.; Lee, K.; Kirilovsky, J.; Nicodeme, E., Discovery and Characterization of Small Molecule Inhibitors of the BET Family Bromodomains. *J. Med. Chem.* 2011, *54*, 3827-3838.
34. Mirguet, O.; Gosmini, R.; Toum, J.; Clément, C. A.; Barnathan, M.; Brusq, J.-M.; Mordaunt, J. E.; Grimes, R. M.; Crowe, M.; Pineau, O.; Ajakane, M.; Daugan, A.; Jeffrey, P.; Cutler, L.; Haynes, A. C.; Smithers, N. N.; Chung, C.-w.; Bamborough, P.; Uings, I. J.; Lewis, A.; Witherington, J.; Parr, N.; Prinjha, R. K.; Nicodème, E., Discovery of Epigenetic Regulator I-BET762: Lead Optimization to Afford a Clinical Candidate Inhibitor of the BET Bromodomains. *J. Med. Chem.* 2013, *56*, 7501-7515.
35. Seal, J.; Lamotte, Y.; Donche, F.; Bouillot, A.; Mirguet, O.; Gellibert, F.; Nicodeme, E.; Krysa, G.; Kirilovsky, J.; Beinke, S.; McCleary, S.; Rioja, I.; Bamborough, P.; Chung, C.-W.; Gordon, L.; Lewis, T.; Walker, A. L.; Cutler, L.; Lugo, D.; Wilson, D. M.; Witherington, J.; Lee, K.; Prinjha, R. K., Identification of a novel series of BET family bromodomain inhibitors: Binding mode and profile of I-BET151 (GSK1210151A). *Bioorg. Med. Chem. Lett.* 2012, *22*, 2968-2972.
36. Gosmini, R.; Nguyen, V. L.; Toum, J.; Simon, C.; Brusq, J.-M. G.; Krysa, G.; Mirguet, O.; Riou-Eymard, A. M.; Boursier, E. V.; Trottet, L.; Bamborough, P.; Clark, H.; Chung, C.-w.; Cutler, L.; Demont, E. H.; Kaur, R.; Lewis, A. J.; Schilling, M. B.; Soden, P. E.; Taylor, S.; Walker, A. L.; Walker, M. D.; Prinjha, R. K.; Nicodème, E., The Discovery of I-BET726 (GSK1324726A), a Potent Tetrahydroquinoline ApoA1 Up-Regulator and Selective BET Bromodomain Inhibitor. *J. Med. Chem.* 2014, *57*, 8111-8131.
37. Hewings, D. S.; Fedorov, O.; Filippakopoulos, P.; Martin, S.; Picaud, S.; Tumber, A.; Wells, C.; Olcina, M. M.; Freeman, K.; Gill, A.; Ritchie, A. J.; Sheppard, D. W.; Russell, A. J.; Hammond, E. M.; Knapp, S.; Brennan, P. E.; Conway, S. J., Optimization of 3,5-Dimethylisoxazole Derivatives as Potent Bromodomain Ligands. *J. Med. Chem.* 2013, *56*, 3217-3227.
38. Ciceri, P.; Müller, S.; O'Mahony, A.; Fedorov, O.; Filippakopoulos, P.; Hunt, J. P.; Lasater, E. A.; Pallares, G.; Picaud, S.; Wells, C.; Martin, S.; Wodicka, L. M.; Shah, N. P.; Treiber, D. K.; Knapp, S., Dual Kinase-Bromodomain Inhibitors for Rationally Designed Polypharmacology. *Nature Chemical Biology* 2014, *10*, 305-312.
39. Ember, S. W. J.; Zhu, J.-Y.; Olesen, S. H.; Martin, M. P.; Becker, A.; Berndt, N.; Georg, G. I.; Schönbrunn, E., Acetyl-lysine Binding Site of Bromodomain-Containing Protein 4 (BRD4) Interacts with Diverse Kinase Inhibitors. *ACS Chemical Biology* 2014, *9*, 1160-1171.

40. Picaud, S.; Fedorov, O.; Thanasopoulou, A.; Leonards, K.; Jones, K.; Meier, J.; Olzscha, H.; Monteiro, O.; Martin, S.; Philpott, M.; Tumber, A.; Filippakopoulos, P.; Yapp, C.; Wells, C.; Che, K. H.; Bannister, A.; Robson, S.; Kumar, U.; Parr, N.; Lee, K.; Lugo, D.; Jeffrey, P.; Taylor, S.; Vecellio, M. L.; Bountra, C.; Brennan, P. E.; Mahony, A.; Velichko, S.; Müller, S.; Hay, D.; Daniels, D. L.; Urh, M.; La Thangue, N. B.; Kouzarides, T.; Prinjha, R.; Schwaller, J.; Knapp, S., Generation of a Selective Small Molecule Inhibitor of the CBP/p300 Bromodomain for Leukemia Therapy. *Cancer Res.* 2015, *75*, 5106-5119.
41. Rooney, T. P. C.; Filippakopoulos, P.; Fedorov, O.; Picaud, S.; Cortopassi, W. A.; Hay, D. A.; Martin, S.; Tumber, A.; Rogers, C. M.; Philpott, M.; Wang, M.; Thompson, A. L.; Heightman, T. D.; Pryde, D. C.; Cook, A.; Paton, R. S.; Müller, S.; Knapp, S.; Brennan, P. E.; Conway, S. J., A Series of Potent CREBBP Bromodomain Ligands Reveals an Induced-Fit Pocket Stabilized by a Cation–π Interaction. *Angew. Chem. Int. Ed.* 2014, *53*, 6126-6130.
42. Humphreys, P. G.; Bamborough, P.; Chung, C.-w.; Craggs, P. D.; Gordon, L.; Grandi, P.; Hayhow, T. G.; Hussain, J.; Jones, K. L.; Lindon, M.; Michon, A.-M.; Renaux, J. F.; Suckling, C. J.; Tough, D. F.; Prinjha, R. K., Discovery of a Potent, Cell Penetrant, and Selective p300/CBP-Associated Factor (PCAF)/General Control Nonderepressible 5 (GCN5) Bromodomain Chemical Probe. *J. Med. Chem.* 2017, *60*, 695-709.
43. Bamborough, P.; Barnett, H. A.; Becher, I.; Bird, M. J.; Chung, C.-w.; Craggs, P. D.; Demont, E. H.; Diallo, H.; Fallon, D. J.; Gordon, L. J.; Grandi, P.; Hobbs, C. I.; Hooper-Greenhill, E.; Jones, E. J.; Law, R. P.; Le Gall, A.; Lugo, D.; Michon, A.-M.; Mitchell, D. J.; Prinjha, R. K.; Sheppard, R. J.; Watson, A. J. B.; Watson, R. J., GSK6853, a Chemical Probe for Inhibition of the BRPF1 Bromodomain. *ACS Medicinal Chemistry Letters* 2016, *7*, 552-557.
44. Bennett, J.; Fedorov, O.; Tallant, C.; Monteiro, O.; Meier, J.; Gamble, V.; Savitsky, P.; Nunez-Alonso, G. A.; Haendler, B.; Rogers, C.; Brennan, P. E.; Müller, S.; Knapp, S., Discovery of a Chemical Tool Inhibitor Targeting the Bromodomains of TRIM24 and BRPF. *J. Med. Chem.* 2016, *59*, 1642-1647.
45. Savitsky, P.; Bray, J.; Cooper, C. D. O.; Marsden, B. D.; Mahajan, P.; Burgess-Brown, N. A.; Gileadi, O., High-throughput Production of Human Proteins for Crystallization: The SGC Experience. *Journal of Structural Biology* 2010, *172*, 3-13.
46. Hewings, D. S.; Wang, M.; Philpott, M.; Fedorov, O.; Uttarkar, S.; Filippakopoulos, P.; Picaud, S.; Vuppusetty, C.; Marsden, B.; Knapp, S.; Conway, S. J.; Heightman, T. D., 3,5-Dimethylisoxazoles Act As Acetyl-lysine-mimetic Bromodomain Ligands. *J. Med. Chem.* 2011, *54*, 6761-6770.
47. Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M., Identification of Compounds with Binding Affinity to Proteins *via* Magnetization Transfer from Bulk Water*. *J. Biomol. NMR* 2000, *18*, 65-68.
48. Zhang, G.; Liu, R.; Zhong, Y.; Plotnikov, A. N.; Zhang, W.; Zeng, L.; Rusinova, E.; Gerona-Nevarro, G.; Moshkina, N.; Joshua, J.; Chuang, P. Y.; Ohlmeyer, M.; He, J. C.; Zhou, M.-M., Down-regulation of NF-κB Transcriptional Activity in HIV-associated Kidney Disease by BRD4 Inhibition. *Journal of Biological Chemistry* 2012, *287*, 28840-28851.
49. Nicodeme, E.; Jeffrey, K. L.; Schaefer, U.; Beinke, S.; Dewell, S.; Chung, C.-w.; Chandwani, R.; Marazzi, I.; Wilson, P.; Coste, H.; White, J.; Kirilovsky, J.; Rice, C. M.; Lora, J. M.; Prinjha, R. K.; Lee, K.; Tarakhovsky, A., Suppression of inflammation by a synthetic histone mimic. *Nature* 2010, *468*, 1119-1123.

50. Niesen, F. H.; Berglund, H.; Vedadi, M., The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions that Promote Protein Stability. *Nat. Protocols* 2007, *2*, 2212-2221.
51. Bullock, A. N.; Debreczeni, J. É.; Fedorov, O. Y.; Nelson, A.; Marsden, B. D.; Knapp, S., Structural Basis of Inhibitor Specificity of the Human Protooncogene Proviral Insertion Site in Moloney Murine Leukemia Virus (PIM-1) Kinase. *J. Med. Chem.* 2005, *48*, 7604-7614.
52. Ericsson, U. B.; Hallberg, B. M.; DeTitta, G. T.; Dekker, N.; Nordlund, P., Thermofluor-based High-Throughput Stability Optimization of Proteins for Structural Studies. *Anal. Chem.* 2006, *357*, 289-298.
53. Lavinder, J. J.; Hari, S. B.; Sullivan, B. J.; Magliery, T. J., High-Throughput Thermal Scanning: A General, Rapid Dye-Binding Thermal Shift Screen for Protein Engineering. *JACS* 2009, *131*, 3794-3795.
54. Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B., WaterLOGSY as a Method for Primary NMR Screening: Practical Aspects and Range of Applicability. *J. Biomol. NMR* 2001, *21*, 349-359.
55. Freyer, M. W.; Lewis, E. A., Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. In *Methods in Cell Biology*, Academic Press: 2008; Vol. Volume 84, pp 79-113.
56. Pierce, M. M.; Raman, C. S.; Nall, B. T., Isothermal Titration Calorimetry of Protein–Protein Interactions. *Methods* 1999, *19*, 213-221.
57. Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E., Selective Inhibition of BET Bromodomains. *Nature* 2010, *468*, 1067-1073.
58. Jennings, L. E.; Schiedel, M.; Hewings, D. S.; Picaud, S.; Laurin, C. M. C.; Bruno, P. A.; Bluck, J. P.; Scorah, A. R.; See, L.; Reynolds, J. K.; Moroglu, M.; Mistry, I. N.; Hicks, A.; Guzanov, P.; Clayton, J.; Evans, C. N. G.; Stazi, G.; Biggin, P. C.; Mapp, A. K.; Hammond, E. M.; Humphreys, P. G.; Filippakopoulos, P.; Conway, S. J., BET Bromodomain Ligands: Probing the WPF Shelf to Improve BRD4 Bromodomain Affinity and Metabolic Stability. *Bioorg. Med. Chem.* 2018, *26*, 2937-2957.
59. Filippakopoulos, P.; Knapp, S., Targeting Bromodomains: Epigenetic Readers of Lysine Acetylation. *Nat Rev Drug Discov* 2014, *13*, 337-356.
60. Da Cunha, J. P. C.; Nakayasu, E. S.; de Almeida, I. C.; Schenkman, S., Post-Translational Modifications of *Trypanosoma cruzi* Histone H4. *Mol. Biochem. Parasit.* 2006, *150*, 268-277.
61. Nardelli, S.; da Cunha, J.; Motta, M.; Schenkman, S., Distinct acetylation of *Trypanosoma cruzi* histone H4 during cell cycle, parasite differentiation, and after DNA damage. *Chromosoma* 2009, *118*, 487-499.
62. Sullivan Jr, W. J.; Naguleswaran, A.; Angel, S. O., Histones and Histone Modifications in Protozoan Parasites. *Cellular Microbiology* 2006, *8*, 1850-1861.
63. Villanova, G. V.; Nardelli, S. C.; Cribb, P.; Magdaleno, A.; Silber, A. M.; Motta, M. C. M.; Schenkman, S.; Serra, E., *Trypanosoma cruzi B*romodomain Factor 2 (BDF2) Binds to Acetylated Histones and is Accumulated After UV Irradiation. *Int. J. Parasitol.* 2009, *39*, 665-673.

64. Ritagliati, C.; Villanova, G. V.; Alonso, V. L.; Zuma, A. A.; Cribb, P.; Motta, M. Cristina M.; Serra, E. C., Glycosomal Bromodomain Factor 1 from *Trypanosoma cruzi* Enhances Trypomastigote Cell Infection and Intracellular Amastigote Growth. *Biochem. J.* 2015, *473*, 73-85.
65. Jeffers, V.; Yang, C.; Huang, S.; Sullivan, W. J., Jr., Bromodomains in Protozoan Parasites: Evolution, Function, and Opportunities for Drug Development. *Microbiol Mol Biol Rev* 2017, *81*, e00047-00016.
66. Moretti, N. S.; da Silva Augusto, L.; Clemente, T. M.; Antunes, R. P. P.; Yoshida, N.; Torrecilhas, A. C.; Cano, M. I. N.; Schenkman, S., Characterization of *Trypanosoma cruzi* Sirtuins as Possible Drug Targets for Chagas Disease. *Antimicrobial Agents and Chemotherapy* 2015, *59*, 4669-4679.
67. Hailu, G. S.; Robaa, D.; Forgione, M.; Sippl, W.; Rotili, D.; Mai, A., Lysine Deacetylase Inhibitors in Parasites: Past, Present, and Future Perspectives. *J. Med. Chem.* 2017, *60*, 4780-4804.
68. Alonso, V. L.; Ritagliati, C.; Cribb, P.; Cricco, J. A.; Serra, E. C., Overexpression of Bromodomain Factor 3 in *Trypanosoma cruzi* (TcBDF3) Affects Differentiation of the Parasite and Protects it Against Bromodomain Inhibitors. *FEBS* 2016, *283*, 2051-2066.
69. Ramallo, I. A.; Alonso, V. L.; Rua, F.; Serra, E.; Furlan, R. L. E., A Bioactive *Trypanosoma cruzi* Bromodomain Inhibitor from Chemically Engineered Extracts. *ACS Combinatorial Science* 2018, *20*, 220-228.
70. García, P.; Alonso, V. L.; Serra, E.; Escalante, A. M.; Furlan, R. L. E., Discovery of a Biologically Active Bromodomain Inhibitor by Target-Directed Dynamic Combinatorial Chemistry. *ACS Medicinal Chemistry Letters* 2018, *9*, 1002-1006.
71. Schulz, D.; Mugnier, M. R.; Paulsen, E.-M.; Kim, H.-S.; Chung, C.-w. W.; Tough, D. F.; Rioja, I.; Prinjha, R. K.; Papavasiliou, F. N.; Debler, E. W., Bromodomain Proteins Contribute to Maintenance of Bloodstream Form Stage Identity in the African Trypanosome. *PLoS Biol* 2015, *13*, e1002316.
72. Smith, M. C.; Furman, T. C.; Ingolia, T. D.; Pidgeon, C., Chelating Peptide-Immobilized Metal Ion Affinity Chromatography. A New Concept in Affinity Chromatography for Recombinant Proteins. *Journal of Biological Chemistry* 1988, *263*, 7211-7215.
73. Arnau, J.; Lauritzen, C.; Petersen, G. E.; Pedersen, J., Current Strategies for the Use of Affinity Tags and Tag Removal for the Purification of Recombinant Proteins. *Protein Expression and Purification* 2006, *48*, 1-13.
74. Carson, M.; Johnson, D. H.; McDonald, H.; Brouillette, C.; DeLucas, L. J., His-tag Impact on Structure. *Acta Crystallographica Section D* 2007, *63*, 295-301.
75. Philpott, M.; Rogers, C. M.; Yapp, C.; Wells, C.; Lambert, J.-P.; Strain-Damerell, C.; Burgess-Brown, N. A.; Gingras, A.-C.; Knapp, S.; Müller, S., Assessing Cellular Efficacy of Bromodomain Inhibitors Using Fluorescence Recovery After Photobleaching. *Epigenetics & Chromatin* 2014, *7*, 14-26.
76. Measures, A. Probing Tandem Epigenetic Reader Domains. DPhil thesis, University of Oxford, Department of Chemistry, 2016.
77. Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J., Basic local alignment search tool. *Journal of Molecular Biology* 1990, *215*, 403-410.
78. Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C., Fragment-Based Approaches in Drug Discovery and Chemical Biology. *Biochemistry* 2012, *51*, 4990-5003.

79. Erlanson, D. A.; McDowell, R. S.; O'Brien, T., Fragment-Based Drug Discovery. *J. Med. Chem.* 2004, *47*, 3463-3482.
80. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H., A 'Rule of Three' for Fragment-Based Lead Discovery? *Drug Discovery Today* 2003, *8*, 876-877.
81. Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A., The Maximal Affinity of Ligands. *Proceedings of the National Academy of Sciences* 1999, *96*, 9997-10002.
82. Chung, C.-w.; Dean, A. W.; Woolven, J. M.; Bamborough, P., Fragment-Based Discovery of Bromodomain Inhibitors Part 1: Inhibitor Binding Modes and Implications for Lead Discovery. *J. Med. Chem.* 2012, *55*, 576-586.
83. Bamborough, P.; Diallo, H.; Goodacre, J. D.; Gordon, L.; Lewis, A.; Seal, J. T.; Wilson, D. M.; Woodrow, M. D.; Chung, C.-w., Fragment-Based Discovery of Bromodomain Inhibitors Part 2: Optimization of Phenylisoxazole Sulfonamides. *J. Med. Chem.* 2012, *55*, 587-596.
84. Demont, E. H.; Chung, C.-w.; Furze, R. C.; Grandi, P.; Michon, A.-M.; Wellaway, C.; Barrett, N.; Bridges, A. M.; Craggs, P. D.; Diallo, H.; Dixon, D. P.; Douault, C.; Emmons, A. J.; Jones, E. J.; Karamshi, B. V.; Locke, K.; Mitchell, D. J.; Mouzon, B. H.; Prinjha, R. K.; Roberts, A. D.; Sheppard, R. J.; Watson, R. J.; Bamborough, P., Fragment-Based Discovery of Low-Micromolar ATAD2 Bromodomain Inhibitors. *J. Med. Chem.* 2015, *58*, 5649-5673.
85. Unzue, A.; Xu, M.; Dong, J.; Wiedmer, L.; Spiliotopoulos, D.; Caflisch, A.; Nevado, C., Fragment-Based Design of Selective Nanomolar Ligands of the CREBBP Bromodomain. *J. Med. Chem.* 2016, *59*, 1350-1356.
86. Zhao, L.; Cao, D.; Chen, T.; Wang, Y.; Miao, Z.; Xu, Y.; Chen, W.; Wang, X.; Li, Y.; Du, Z.; Xiong, B.; Li, J.; Xu, C.; Zhang, N.; He, J.; Shen, J., Fragment-Based Drug Discovery of 2-Thiazolidinones as Inhibitors of the Histone Reader BRD4 Bromodomain. *J. Med. Chem.* 2013, *56*, 3833-3851.
87. Molander, G. A.; Canturk, B.; Kennedy, L. E., Scope of the Suzuki–Miyaura Cross-Coupling Reactions of Potassium Heteroaryltrifluoroborates. *J. Org. Chem.* 2009, *74*, 973-980.
88. Lennox, A. J. J.; Lloyd-Jones, G. C., Preparation of Organotrifluoroborate Salts: Precipitation-Driven Equilibrium under Non-Etching Conditions. *Angew. Chem. Int. Ed.* 2012, *51*, 9385-9388.
89. Jennings, L. E. Development of Ligands to Target Bromodomain-Histone Interactions. DPhil thesis, University of Oxford, Department of Chemistry, 2015.
90. A. Roy; D. Kundu; S. K. Kundu; A. Majee; Hajra, A., Manganese (II) Chloride-Catalyzed Conjugated Addition of Amines to Electron Deficient Alkenes in Methanol-Water Medium. *The Open Catalysis Journal* 2010, *3*, 34-39.
91. Picaud, S.; Fedorov, O.; Thanasopoulou, A.; Leonards, K.; Jones, K.; Meier, J.; Olzscha, H.; Monteiro, O.; Martin, S.; Philpott, M.; Tumber, A.; Filippakopoulos, P.; Yapp, C.; Wells, C.; Che, K. H.; Bannister, A.; Robson, S.; Kumar, U.; Parr, N.; Lee, K.; Lugo, D.; Jeffrey, P.; Taylor, S.; Vecellio, M. L.; Bountra, C.; Brennan, P. E.; O'Mahony, A.; Velichko, S.; Müller, S.; Hay, D.; Daniels, D. L.; Urh, M.; La Thangue, N. B.; Kouzarides, T.; Prinjha, R.; Schwaller, J.; Knapp, S., Generation of a Selective Small Molecule Inhibitor of the CBP/p300 Bromodomain for Leukemia Therapy. *Cancer Res.* 2015, *75*, 5106-5119.
92. Steegmaier, M.; Hoffmann, M.; Baum, A.; Lénárt, P.; Petronczki, M.; Krššák, M.; Gürtler, U.; Garin-Chesa, P.; Lieb, S.; Quant, J.; Grauert, M.; Adolf, G. R.; Kraut, N.; Peters, J.-M.; Rettig, W. J., BI-2536, a Potent and Selective Inhibitor of Polo-like Kinase 1, Inhibits Tumor Growth In Vivo. *Current Biology* 2007, *17*, 316-322.

93. Murray, R. G.; Whitehead, D. M.; Le Strat, F.; Conway, S. J., Facile One-Pot Wynthesis of 5-Substituted Hydantoins. *Organic & Biomolecular Chemistry* 2008, *6*, 988-991.
94. Watts, E.; Heidenreich, D.; Tucker, E.; Raab, M.; Strebhardt, K.; Chesler, L.; Knapp, S.; Bellenie, B.; Hoelder, S., Designing Dual Inhibitors of Anaplastic Lymphoma Kinase (ALK) and Bromodomain-4 (BRD4) by Tuning Kinase Selectivity. *J. Med. Chem.* 2019, *62*, 2618-2637.
95. Keller, M. Development of Selective Ligands for Non-Human Bromodomains. Master's thesis, University of Oxford, Department of Chemistry, 2019.
96. Hewings, D. S. Developing Inhibitors of Bromodomain-Histone Interactions. DPhil thesis, University of Oxford, Department of Chemistry, 2014.
97. Woodhouse, F. Fragment-Based Studies of Acetyl-Lysine Mimics for the Inhibition of Non-Human Bromodomains. Master's thesis, University of Oxford, Department of Chemistry, 2016.
98. Chen, L.; Yap, J. L.; Yoshioka, M.; Lanning, M. E.; Fountain, R. N.; Raje, M.; Scheenstra, J. A.; Strovel, J. W.; Fletcher, S., BRD4 Structure–Activity Relationships of Dual PLK1 Kinase/BRD4 Bromodomain Inhibitor BI-2536. *ACS Medicinal Chemistry Letters* 2015, *6*, 764-769.
99. Carr, H. Y.; Purcell, E. M., Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Physical Review* 1954, *94*, 630-638.
100. Meiboom, S.; Gill, D., Modified Spin - Echo Method for Measuring Nuclear Relaxation Times. *Review of Scientific Instruments* 1958, *29*, 688-691.
101. Fielding, L., NMR Methods for the Determination of Protein- Ligand Dissociation Constants. *Current Topics in Medicinal Chemistry* 2003, *3*, 39-53.
102. Rudolph, D.; Steegmaier, M.; Hoffmann, M.; Grauert, M.; Baum, A.; Quant, J.; Haslinger, C.; Garin-Chesa, P.; Adolf, G. R., BI 6727, A Polo-like Kinase Inhibitor with Improved Pharmacokinetic Profile and Broad Antitumor Activity. *Clinical Cancer Research* 2009, *15*, 3094-3102.
103. Yao, G.; Lu, X.; Phelps, C.; Evindar, G., Chapter 7 DNA-encoded Library Technology (ELT). In *Chemical and Biological Synthesis: Enabling Approaches for Understanding Biology*, The Royal Society of Chemistry: 2018; pp 153-183.
104. Berg, T., Modulation of Protein–Protein Interactions with Small Organic Molecules. *Angew. Chem. Int. Ed.* 2003, *42*, 2462-2481.
105. Baell, J.; Walters, M. A., Chemistry: Chemical Con Artists Foil Drug Discovery. *Nature* 2014, *513*, 481–483.
106. Hill, J. R.; Robertson, A. A. B., Fishing for Drug Targets: A Focus on Diazirine Photoaffinity Probe Synthesis. *J. Med. Chem.* 2018, *61*, 6945-6963.
107. Burger, M. T.; Han, W.; Lan, J.; Nishiguchi, G.; Bellamacina, C.; Lindval, M.; Atallah, G.; Ding, Y.; Mathur, M.; McBride, C.; Beans, E. L.; Muller, K.; Tamez, V.; Zhang, Y.; Huh, K.; Feucht, P.; Zavorotinskaya, T.; Dai, Y.; Holash, J.; Castillo, J.; Langowski, J.; Wang, Y.; Chen, M. Y.; Garcia, P. D., Structure Guided Optimization, in Vitro Activity, and in Vivo Activity of Pan-PIM Kinase Inhibitors. *ACS Medicinal Chemistry Letters* 2013, *4*, 1193-1197.
108. Lin, S.; Wang, C.; Ji, M.; Wu, D.; Lv, Y.; Sheng, L.; Han, F.; Dong, Y.; Zhang, K.; Yang, Y.; Li, Y.; Chen, X.; Xu, H., Discovery of New Thienopyrimidine Derivatives as Potent and Orally Efficacious Phosphoinositide 3-Kinase Inhibitors. *Bioorg. Med. Chem.* 2018, *26*, 637-646.

109. Wuts;, P. G. M.; Greene, T. W., *Protection for the Amino Group. Greene's Protective Groups in Organic Synthesis*. P. G. Wuts (Ed): Wiley Online Books, 2014.
110. Kambe, T.; Correia, B. E.; Niphakis, M. J.; Cravatt, B. F., Mapping the Protein Interaction Landscape for Fully Functionalized Small-Molecule Probes in Human Cells. *JACS* 2014, *136*, 10777-10782.
111. Ahad, A. M.; Jensen, S. M.; Jewett, J. C., A Traceless Staudinger Reagent To Deliver Diazirines. *Org. Lett.* 2013, *15*, 5060-5063.
112. Wang, L.; Ishida, A.; Hashidoko, Y.; Hashimoto, M., Dehydrogenation of the NH−NH Bond Triggered by Potassium tert-Butoxide in Liquid Ammonia. *Angew. Chem. Int. Ed.* 2017, *56*, 870-873.
113. Debets, M. F.; van Berkel, S. S.; Schoffelen, S.; Rutjes, F. P. J. T.; van Hest, J. C. M.; van Delft, F. L., Aza-dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition. *Chemical Communications* 2010, *46*, 97-99.
114. Pickens, C. J.; Johnson, S. N.; Pressnall, M. M.; Leon, M. A.; Berkland, C. J., Practical Considerations, Challenges, and Limitations of Bioconjugation via Azide–Alkyne Cycloaddition. *Bioconjugate Chemistry* 2018, *29*, 686-701.
115. Huang, X.; Aulabaugh, A., Application of Fluorescence Polarization in HTS Assays. In *High Throughput Screening: Methods and Protocols, Second Edition*, Janzen, W. P.; Bernasconi, P., Eds. Humana Press: Totowa, NJ, 2009; pp 127-143.
116. Jolley, M. E., Fluorescence Polarization Assays for the Detection of Proteases and Their Inhibitors. *Journal of Biomolecular Screening* 1996, *1*, 33-38.
117. Wang, W.; Wang, L.; Liang, Y.-S., Susceptibility or resistance of praziquantel in human schistosomiasis: a review. *Parasitology Research* 2012, *111*, 1871-1877.
118. Protasio, A. V.; Tsai, I. J.; Babbage, A.; Nichol, S.; Hunt, M.; Aslett, M. A.; De Silva, N.; Velarde, G. S.; Anderson, T. J. C.; Clark, R. C.; Davidson, C.; Dillon, G. P.; Holroyd, N. E.; LoVerde, P. T.; Lloyd, C.; McQuillan, J.; Oliveira, G.; Otto, T. D.; Parker-Manuel, S. J.; Quail, M. A.; Wilson, R. A.; Zerlotini, A.; Dunne, D. W.; Berriman, M., A Systematically Improved High Quality Genome and Transcriptome of the Human Blood Fluke Schistosoma mansoni. *PLOS Neglected Tropical Diseases* 2012, *6*, e1455.
119. Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J., Safe and Convenient Procedure for Solvent Purification. *Organometallics* 1996, *15*, 1518-1520.
120. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A., NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* 1997, *62*, 7512-7515.
121. Sunder, S.; Peet, N. P., Nitro group participation in the formation of heterocycles. 1979, *16*, 33-37.
122. Cook, J.; Zusi, F. C.; Hill, M. D.; Fang, H.; Pearce, B.; Park, H.; Gallagher, L.; McDonald, I. M.; Bristow, L.; Macor, J. E.; Olson, R. E., Design and synthesis of a novel series of (1 ′ S,2R,4 ′ S)-3H-4 ′ -azaspiro[benzo[4,5]imidazo[2,1-b]oxazole-2,2 ′ -bicyclo[2.2.2]octanes] with high affinity for the α7 neuronal nicotinic receptor. *Bioorg. Med. Chem. Lett.* 2017, *27*, 5002-5005.
123. Patil, V. V.; Shankarling, G. S., Steric-Hindrance-Induced Regio- and Chemoselective Oxidation of Aromatic Amines. *J. Org. Chem.* 2015, *80*, 7876-7883.
124. Fife, T. H.; Hutchins, J. E. C.; Wang, M. S., Highly efficient intramolecular nucleophilic reactions. Cyclization of p-nitrophenyl N-(2-mercaptophenyl)-N-methylcarbamate and phenyl N-(2-aminophenyl)-N-methylcarbamate. *JACS* 1975, *97*, 5878-5882.

125. Ando, K.; Kobayashi, T.; Uchida, N., Practical Methylenation Reaction for Aldehydes and Ketones Using New Julia-Type Reagents. *Org. Lett.* 2015, *17*, 2554-2557.
126. Shiokawa, Y.; Ohki, S., Studies on Benzimidazoles and Related Compounds. IV. Reactivity of 2-Azido-1-methylbenzimidazole. *Chemical & Pharmaceutical Bulletin* 1971, *19*, 401-408.
127. Likhar, P. R.; Arundhathi, R.; Kantam, M. L.; Prathima, P. S., Amination of Alcohols Catalyzed by Copper-Aluminium Hydrotalcite: A Green Synthesis of Amines. *Eur. J. Org. Chem.* 2009, *2009*, 5383-5389.
128. Isley, N. A.; Linstadt, R. T. H.; Kelly, S. M.; Gallou, F.; Lipshutz, B. H., Nucleophilic Aromatic Substitution Reactions in Water Enabled by Micellar Catalysis. *Org. Lett.* 2015, *17*, 4734-4737.
129. Brouwer, C.; Jenko, K. J.; Zoghbi, S. S.; Morse, C. L.; Innis, R. B.; Pike, V. W., Translocator protein ligands based on N-methyl-(quinolin-4-yl)oxypropanamides with properties suitable for PET radioligand development. *Eur. J. Med. Chem.* 2016, *124*, 677-688.
130. Meanwell, N. A.; Sit, S. Y.; Gao, J.; Wong, H. S.; Gao, Q.; Laurent, D. R. S.; Balasubramanian, N., Regiospecific Functionalization of 1,3-Dihydro-2H-benzimidazol-2-one and Structurally Related Cyclic Urea Derivatives. *J. Org. Chem.* 1995, *60*, 1565-1582.
131. Yan, W.; Wang, X.; Dai, Y.; Zhao, B.; Yang, X.; Fan, J.; Gao, Y.; Meng, F.; Wang, Y.; Luo, C.; Ai, J.; Geng, M.; Duan, W., Discovery of 3-(5′-Substituted)-Benzimidazole-5-(1-(3,5-dichloropyridin-4-yl)ethoxy)-1*H*-indazoles as Potent Fibroblast Growth Factor Receptor Inhibitors: Design, Synthesis, and Biological Evaluation. *J. Med. Chem.* 2016, *59*, 6690-6708.
132. Piller, F. M.; Metzger, A.; Schade, M. A.; Haag, B. A.; Gavryushin, A.; Knochel, P., Preparation of Polyfunctional Arylmagnesium, Arylzinc, and Benzylic Zinc Reagents by Using Magnesium in the Presence of LiCl. *Chemistry – A European Journal* 2009, *15*, 7192-7202.
133. Tan, L. Developing ligands for non-human bromodomains. Master's thesis, University of Oxford, Department of Chemistry, 2018.
134. Rashid, S. O.; Almadhhi, S. S.; Berrisford, D. J.; Raftery, J.; Vitorica-Yrezabal, I.; Whitehead, G.; Quayle, P., Radical Truce-Smiles reactions on an isoxazole template: Scope and limitations. *Tetrahedron* 2019, *75*, 2413-2430.
135. Molloy, J. J.; Law, R. P.; Fyfe, J. W. B.; Seath, C. P.; Hirst, D. J.; Watson, A. J. B., A modular synthesis of functionalised phenols enabled by controlled boron speciation. *Organic & Biomolecular Chemistry* 2015, *13*, 3093-3102.
136. M. Best, W.; L. Ghisalberti, E.; Powell, M., Simple Synthesis of Symmetrical 4-Substituted 3,5-Dialkylisoxazoles. *Journal of Chemical Research, Synopses* 1998, 388-389.
137. Zheng, Y.; Yang, C.; Zhang-Negrerie, D.; Du, Y.; Zhao, K., One-pot synthesis of isoxazoles from enaminones: an application of Fe(II) as the catalyst for ring expansion of 2H-azirine intermediates. *Tetrahedron Lett.* 2013, *54*, 6157-6160.
138. Markovic, T.; Murray, P. R. D.; Rocke, B. N.; Shavnya, A.; Blakemore, D. C.; Willis, M. C., Heterocyclic Allylsulfones as Latent Heteroaryl Nucleophiles in Palladium-Catalyzed Cross-Coupling Reactions. *JACS* 2018, *140*, 15916-15923.
139. Sardarian, A. R.; Inaloo, I. D., 4-Dodecylbenzenesulfonic acid (DBSA) promoted solvent-free diversity-oriented synthesis of primary carbamates, S-thiocarbamates and ureas. *RSC Advances* 2015, *5*, 76626-76641.

140. Shadrick, W. R.; Slavish, P. J.; Chai, S. C.; Waddell, B.; Connelly, M.; Low, J. A.; Tallant, C.; Young, B. M.; Bharatham, N.; Knapp, S.; Boyd, V. A.; Morfouace, M.; Roussel, M. F.; Chen, T.; Lee, R. E.; Kiplin Guy, R.; Shelat, A. A.; Potter, P. M., Exploiting a water network to achieve enthalpy-driven, bromodomain-selective BET inhibitors. *Bioorg. Med. Chem.* 2018, *26*, 25-36.
141. Pilkington, L. I.; Barker, D., Asymmetric Synthesis and CD Investigation of the 1,4-Benzodioxane Lignans Eusiderins A, B, C, G, L, and M. *J. Org. Chem.* 2012, *77*, 8156-8166.
142. Popp, T. A.; Tallant, C.; Rogers, C.; Fedorov, O.; Brennan, P. E.; Müller, S.; Knapp, S.; Bracher, F., Development of Selective CBP/P300 Benzoxazepine Bromodomain Inhibitors. *J. Med. Chem.* 2016, *59*, 8889-8912.
143. Scharow, A.; Knappe, D.; Reindl, W.; Hoffmann, R.; Berg, T., Development of Bifunctional Inhibitors of Polo-Like Kinase 1 with Low-Nanomolar Activities Against the Polo-Box Domain. 2016, *17*, 759-767.
144. Modi, C. K.; Trivedi, P. M., Zeolite-Y encapsulated nanohybrid materials: synthesis, spectroscopic characterization, and catalytic significance. *Journal of Coordination Chemistry* 2012, *65*, 525-538.

Appendices

Appendix A. Plasmids and protein constructs 293
A.1. Plasmid maps 293
A.2. Amino acid sequences 299
A.3. Cloning 303
A.4. Protein expression 309
Appendix B. Biophysical assays 319
B.1. Thermal shift 319
B.2. WaterLOGSY spectra 325
B.3. Ligand-observed proton NMR titration 369
B.4. Isothermal calorimetry titration curves 384
B.5. ^{19}F NMR assays 395
B.6. Fluorescence polarisation 397
B.7. PhaBits LCMS spectra examples 402
Appendix C. Benzimidazolone ring formation mechanism 404
Appendix D. Crystallisation trials buffers 405
D.1. Compositions of the commercially available broad screen for 405
D.2. Optimisation tray for E5 condition from PEG ION broad screen: 419
D.3. Optimisation tray for E8 condition from Crystal screen broad screen 420
Spectra for novel compounds 421

Appendix A. Plasmids and protein constructs

A.1. Plasmid maps

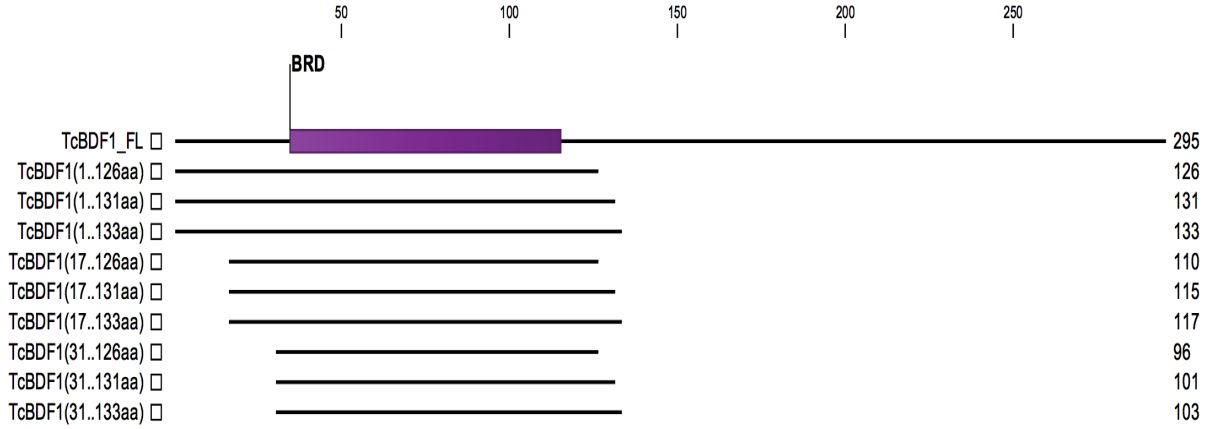


Figure A.1: The various TcBDF1 constructs designed and produced by Dr Anthony Chan for expression trials. No TcBDF1 protein construct was found to be soluble.

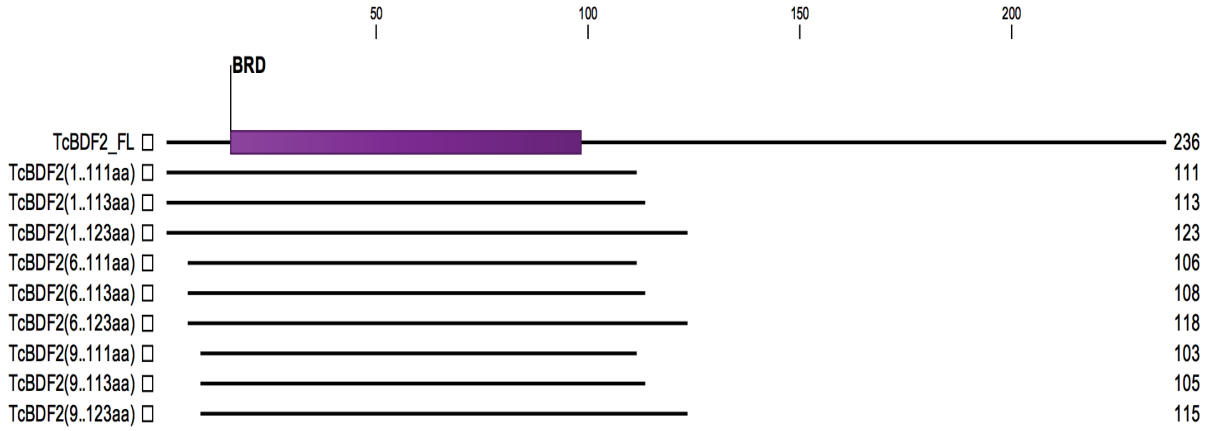


Figure A.2: The various TcBDF2 constructs designed and produced by Dr Anthony Chan for expression trials. H_6-TcBDF2(1..113aa) was identified as a soluble construct.

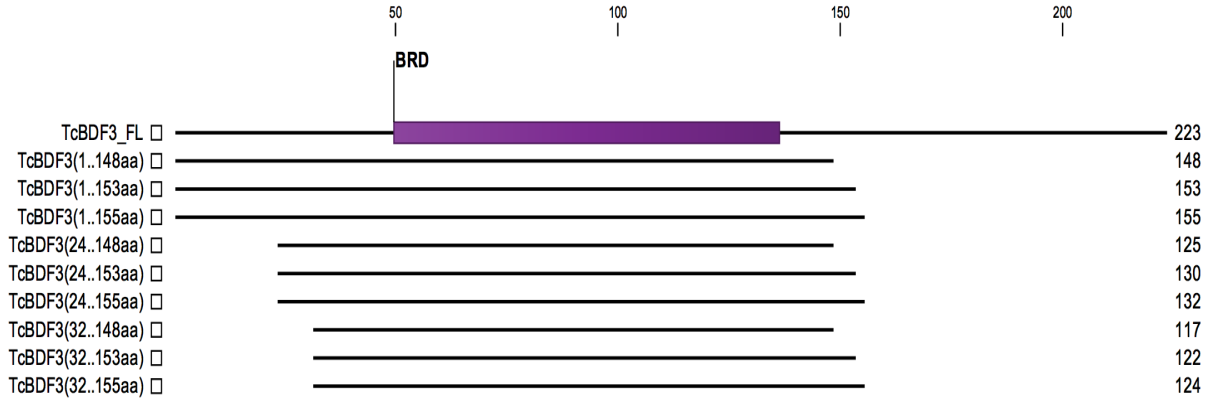


Figure A.3: The various TcBDF3 constructs designed and produced by Dr Anthony Chan for expression trials. H_6-TcBDF3(32..148aa) was identified as a soluble construct.

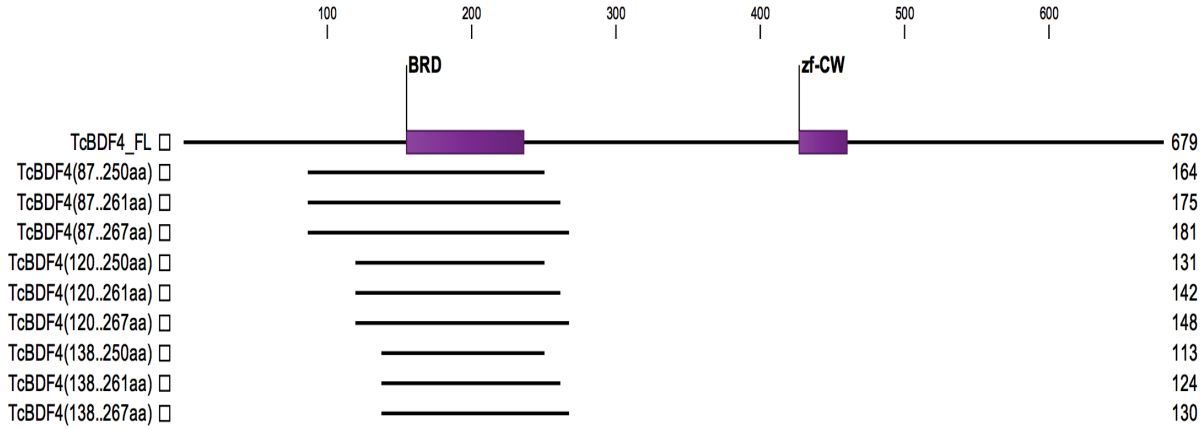


Figure A.4: The various TcBDF4 constructs designed and produced by Dr Anthony Chan for expression trials. No TcBDF1 protein construct was found to be soluble.

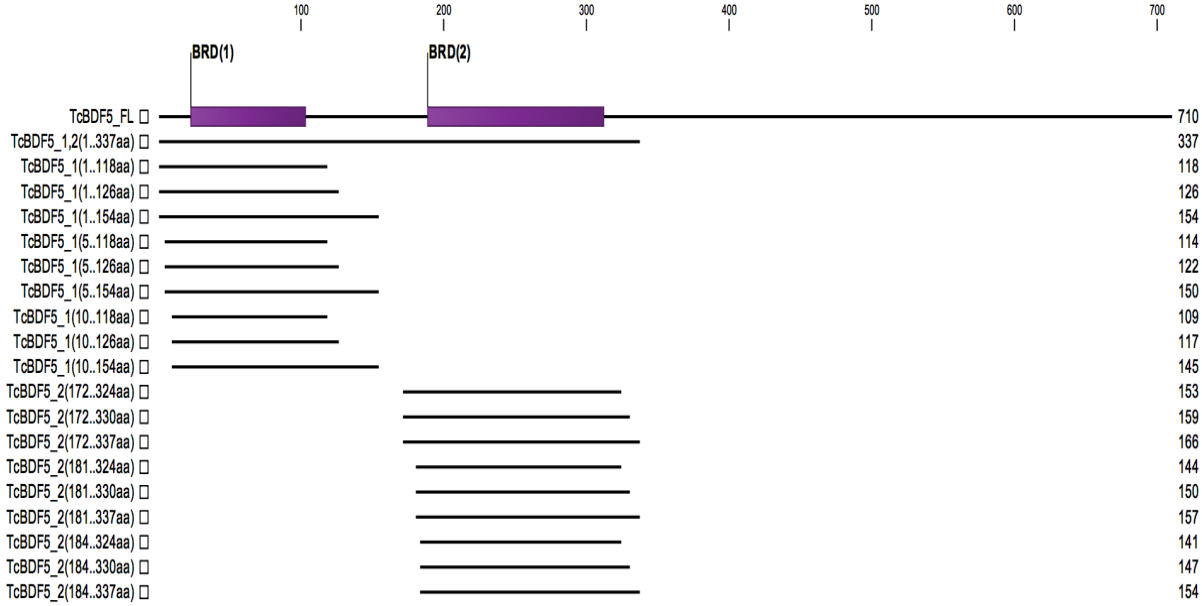


Figure A.5: The various TcBDF5 constructs designed and produced by Dr Anthony Chan for expression trials. H_6-TcBDF5 (1..118aa) (Dr Anthony K. Chan) and H_6-TcBDF5 (1..126aa) were identified as a soluble constructs.

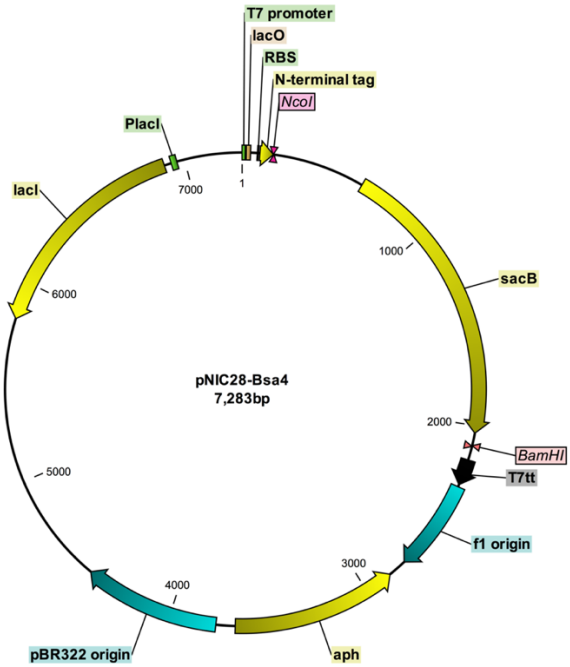


Figure A.6: A plasmid map depicting the cloning and expression vector, pNIC28-Bsa4. The plasmid shown here is the pNIC28-Bsa4 vector (Addgene plasmid # 26103)[4], which was drawn to scale. All of the other expression plasmids used in this study share the same vector backbone as that of the pNIC28-Bsa4 vector. The vector contains in-frame hexa-histidine codons right after the initiation start (ATG) for protein purification by immobilized metal affinity chromatography (IMAC). *T7* promoter is used for transcribing target gene, and *T7tt* is the *T7* transcriptional terminator for terminating gene transcription. The *Nco*I and *Bam*HI restriction sites used to linearize the vector for Gibson assembly[5] are shown. The *sacB* gene allows negative selection of plasmid without gene insert on 5% sucrose. Protein expression can be initiated by addition of isopropyl β-d-1-thiogalactopyranoside (IPTG), which removes transcriptional repression exerted by the LacI repressor (encoded by the constitutively expressed *lacI* gene). The locations of the *lacO* (*lac* operator, the site bound by the LacI repressor) and TEV site (protease cutting site for removal of the hexa-histidine purification tag in the protein product by TEV protease with a recognition sequence of "ENLYFQ/S", where the slash mark indicates the site of cleavage) are indicated. DNA replication origins, pBR322 origin (for double-stranded DNA replication of plasmid in *E. coli*) and f1 (for single stranded replication in presence of a helper phage), are shown. The kanamycin selectable marker gene, *aph*, is labelled, which encodes aminoglycoside 3'-phosphotransferase. The plasmid map was generated and annotated by using CLC Main Workbench v7.9.1 (QIAGEN) by Dr Anthony Chan.

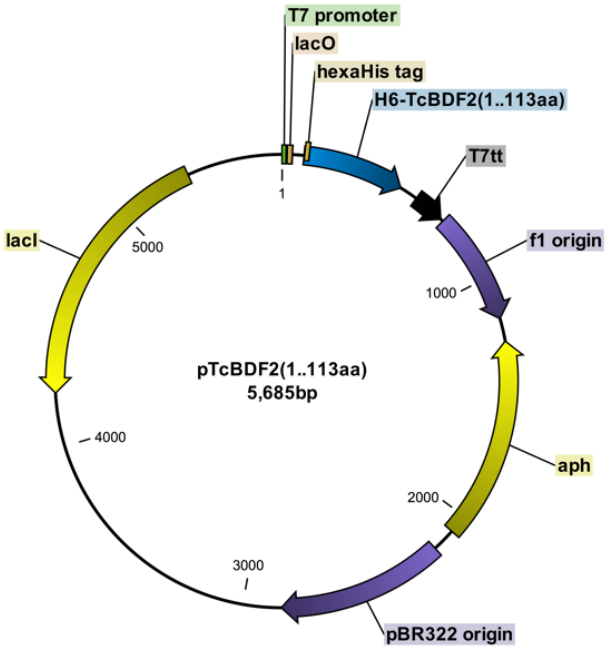


Figure A.7: Plasmid map of H_6-*Tc*BDF2(1..113aa) showing the key parts of the constructs, including the promoters, the operators and the TEV cleaving site. This plasmid map is representative of the plasmids used in this work. The *T7* promoter is used for transcribing target gene, and *T7tt* is the *T7* transcriptional terminator for terminating gene transcription. Protein expression can be initiated by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG), which removes transcriptional repression exerted by the LacI repressor (encoded by the constitutively expressed lacI gene). The lac operator (lacO) is the site bound by the LacI repressor. The kanamycin selectable marker gene, *aph*, encodes aminoglycoside 3'-phosphotransferase. The plasmid map was generated and annotated by using CLC Main Workbench v7.9.1.

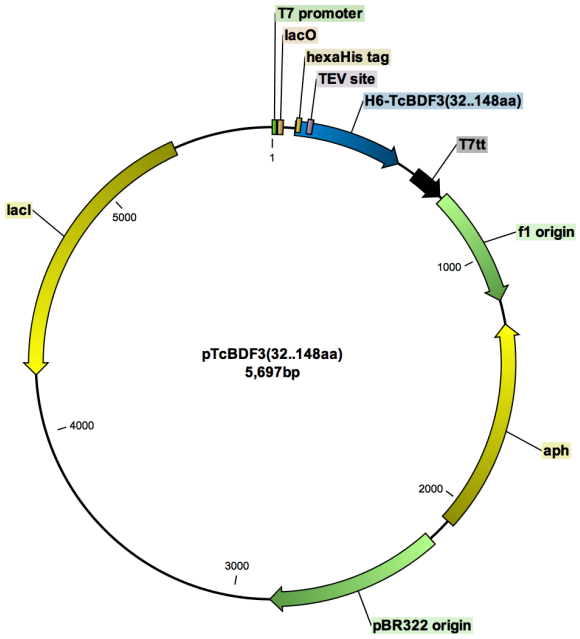


Figure A.8: Plasmid map of H_6-*Tc*BDF3(32..148aa) showing the key parts of the constructs, including the promoters, the operators and the TEV cleaving site.

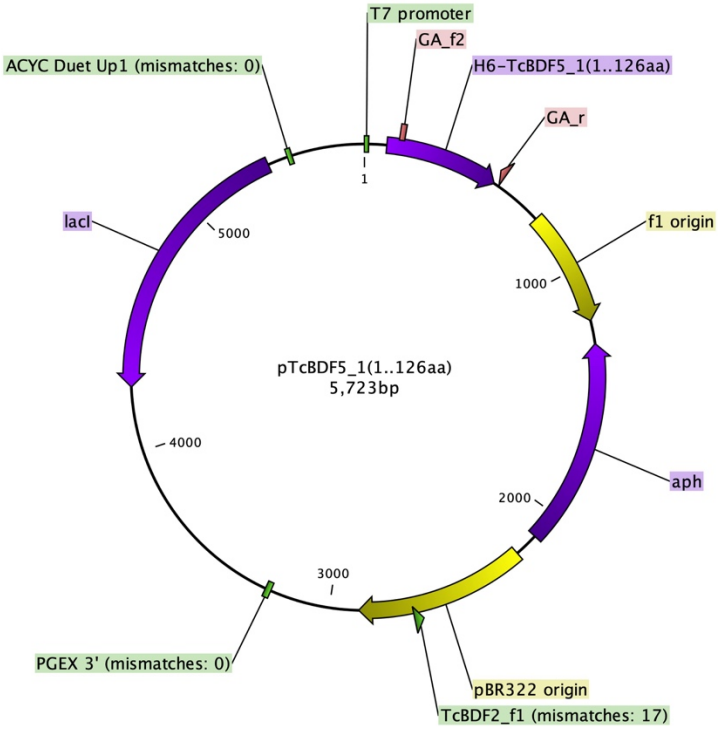


Figure A.9: Plasmid map of H_6- *Tc*BDF5(1..126aa) showing the key parts of the constructs, including the promoters, the operators and the TEV cleaving site.

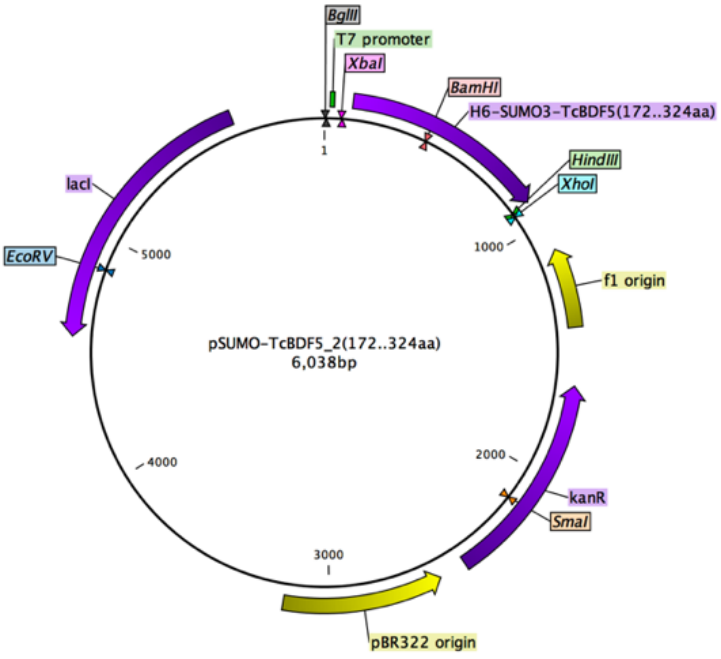


Figure A.10: Plasmid map of H_6-SUMO-*Tc*BDF5(172..324aa) showing the key parts of the constructs, including the promoters, the operators and the TEV cleaving site.

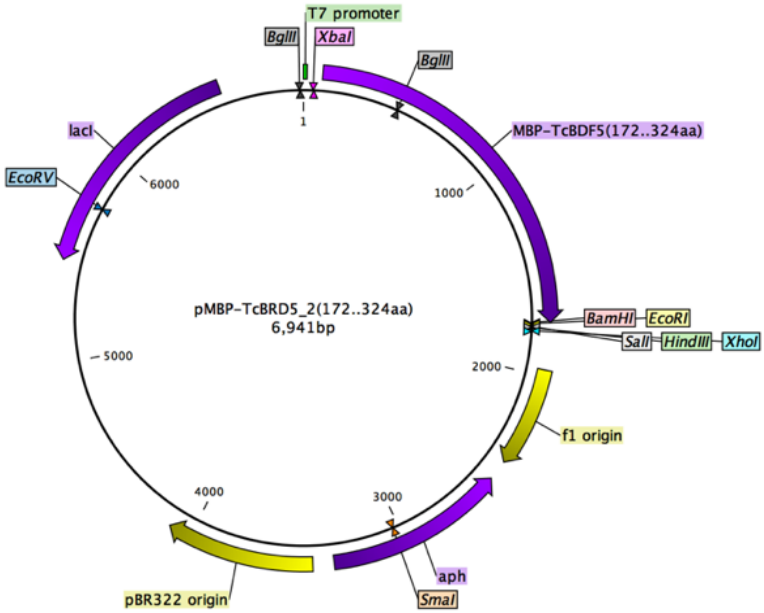


Figure A.11: Plasmid map of H_6-MBP-*Tc*BDF5(172..324aa) showing the key parts of the constructs, including the promoters, the operators and the TEV cleaving site.

A.2. Amino acid sequences

The sequences for the proteins purified as part of this work are detailed below. Sequences are given in the conventional amino acid 1-letter code. The following sequence features are highlighted: Histag (grey), the TEV cleaving site (yellow), N to F point mutation (turquoise).

A.2.1. *Trypanosoma cruzi* BD proteins

H_6-*Tc*BFDF2(1..113aa): (MW: 15398.3 Da, ε_{280} = 20065 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSMGKRGREGTLDKPRCLAFVHQLWDKDKLKMFHHPISAAELP
DYHKVINYPVDLSTIRQGIESGKYDSDADVQNAVAQMIANALEYNAKGTEWHQQALSFRNIYLDV
ARQCGLS

H_6-*Tc*BFDF2(1..113aa)N86F (MW: 15431.4 Da, ε_{280} = 20065 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSMGKRGREGTLDKPRCLAFVHQLWDKDKLKMFHHPISAAELP
DYHKVINYPVDLSTIRQGIESGKYDSDADVQNAVAQMIANALEYFAKGTEWHQQALSFRNIYLDVA
RQCGLS

H_6-*Tc*BFDF3(32..148aa): (MW: 15803.7 Da, ε_{280} = 17545 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSGKHVFRSDLSTVPEVRDIYDCLYKLYAEESASASFREPVNALELG
VFNYYEVVTEPMSLRTVLDRIAEGGHYSQASQVLADVEKIWSNCEKYNGADSALVKEAKKCQGILT
RLRERL

*Tc*BFDF3(32..148aa) (cleaved): (MW: 13338.1 Da, ε_{280} = 16055 $M^{-1}cm^{-1}$)

SGKHVFRSDLSTVPEVRDIYDCLYKLYAEESASASFREPVNALELGVFNYYEVVTEPMSLRTVLDRIAE
GGHYSQASQVLADVEKIWSNCEKYNGADSALVKEAKKCQGILTRLRERL

H_6-*Tc*BFDF3(32..148aa)N124F (MW: 15836.8 Da, ε_{280} = 17545 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSGKHVFRSDLSTVPEVRDIYDCLYKLYAEESASASFREPVNALELG
VFNYYEVVTEPMSLRTVLDRIAEGGHYSQASQVLADVEKIWSNCEKYFGADSALVKEAKKCQGILTR
LRERL

H_6-*Tc*BFDF5-1(1..118aa): (MW: 16226.2 Da, ε_{280} = 12045 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSMEERRQKFYPEEELVALVRSLDRPQDEGLFSMDVLVVYPHLE
QEYTRVCPKRCDLATAAEKAANEAYSYDVNLTALREDIKLMVNNCYRFNGTKGPLANIAERFEAFAK
EQIDAYVTK

H_6-*Tc*BFDF5-1(1..126aa): (MW: 17052 Da, ε_{280} = 17052.2 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSMEERRQKFYPEEELVALVRSLDRPQDEGLFSMDVLVVYPHLE
QEYTRVCPKRCDLATAAEKAANEAYSYDVNLTALREDIKLMVNNCYRFNGTKGPLANIAERFEAFAK
EQIDAYVTKKAGGRRLS

H_6-*Tc*BFDF5-1(1..118aa)N91F (MW: 16259.3 Da, ε_{280} = 12045 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSMEERRQKFYPEEELVALVRSLDRPQDEGLFSMDVLVVYPHLE
QEYTRVCPKRCDLATAAEKAANEAYSYDVNLTALREDIKLMVNNCYRFFGTKGPLANIAERFEAFAK
EQIDAYVTK

H_6-SUMO-*Tc*BFDF5-2(172..324aa) (MW: 29503.1 Da, ε_{280} = 16055 $M^{-1}cm^{-1}$)

MKHHHHHHPMSDYDIPTTENLYFQGAMGNDHINLKVAGQDGSVVQFKIKRHTPLSKLMKAYCER
QGLSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGGSRSGEMDGLSATSAEMVRLVD
SLNRREDQGAFAVDVAEKYPELKKSYEEVCPVKMNLTILRERAASGYYLGIPQPTTTSKNPEPQTQSK
RRQLLASDASPSSVFGSTTAESLVLLRNDVELMVSNCVRFNSNVQSWVELARSFQSFAHKKIDDF

H_6-MBP-*Tc*BFDF5-2(172..324aa) (MW: 61281.9 Da, ε_{280} = 79425 $M^{-1}cm^{-1}$)

MHHHHHHKTEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGP
DIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNP
PKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAG
LTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKP
FVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENA
QKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGIEGRG
ENLYFQSRSGEMDGLSATSAEMVRLVDSLNRREDQGAFAVDVAEKYPELKKSYEEVCPVKMNLTIL
RERAASGYYLGIPQPTTTSKNPEPQTQSKRRQLLASDASPSSVFGSTTAESLVLLRNDVELMVSNCV
RFNSNVQSWVELARSFQSFAHKKIDDF

A.2.2. Human BD proteins

H_6-BRD4(1) (Addgene #38942) (MW: 17549.1 Da, ε_{280} = 28545 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSMNPPPPETSNPNKPKRQTNQLQYLLRVVLKTLWKHQFAWPF
QQPVDAVKLNLPDYYKIIKTPMDMGTIKKRLENNYYWNAQECIQDFNTMFTNCYIYNKPGDDIVL
MAEALEKLFLQKINELPTEE

H_6-PLK1 (MW: 16673.0 Da, ε_{280} = 22265 $M^{-1}cm^{-1}$) (Gift from Dr Fiona Sorrell, SGC)

MHHHHHHSSGVDLGTENLYFQSMAPAAAPPAKEIPEVLVDPRSRRRYVRGRFLGKGGFAKCFEIS
DADTKEVFAGKIVPKSLLLKPHQREKMSMEISIHRSLAHQHVVGFHGFFEDNDFVFVVLELCRRRSL
LELHKRRKALTEPEARYYLRQIVLGCQYLHRNRVIHRDLKLGNLFLNEDLEVKIGDFGLATKVEYDGE
RKKTLCGTPNYIAPEVLSKKGHSFEVDVWSIGCIMYTLLVGKPPFETSCLKETYLRIKKNEYSIPKHINP
VAASLIQKMLQTDPTARPTINELLNDEFFTSGYIPARLPITCLTIPPRFSIAPS

H_6-CREBBP (Addgene #38977) (MW: 16673.0 Da, ε_{280} = 26930 $M^{-1}cm^{-1}$)

MHHHHHHSSGVDLGTENLYFQSMRKKIFKPEELRQALMPTLEALYRQDPESLPFRQPVDPQLLGIP
DYFDIVKNPMDLSTIKRKLDTGQYQEPWQYVDDVWLMFNNAWLYNRKTSRVYKFCSKLAEVFEQ
EIDPVMQSLG

A.2.3. *Leishmania donovani* BD proteins

Provided by Dr Nathaniel Jones (Prof. Anthony Wilkinson's research group, Biochemistry department, University of York). Protein production by Dr Nathaniel Jones and the Protein Production Facility of the University of York Technology Facility; plasmid from SGC Toronto.

H_6-*Ld*BDF2 (MW: 16614.7 Da, ε_{280} = 19940 $M^{-1}cm^{-1}$)

MHHHHHHSSGRENLYFQGMDVSKRPREEFHKEQCLSFVKKLWAADTLAMFHYPVSATEVP
GYYDVVDTPMDLSTIRKNIEQGKYRTDTEVENDVVLMLSNALDFNEKGSQWHDLAKQLKK
RYLTLAQESGLSFDADQAFIPTK

H_6-*Ld*BDF5-1 (MW: 17931.2 Da, ε_{280} = 16055 $M^{-1}cm^{-1}$)

MHHHHHHSSGRENLYFQGPKLYNEADVAALVRSLDRAEDHHIFAVDVLETYPYLAESYTK
VCPRRCDLATAAQKALEGAYSYDLRLEGLKADIALMASNCVAYNGPTSAYAETAAKFERY
ALEQIDAFVLEHNGGC

H_6-*Ld*BDF5-2 (MW: 15190.96 Da, ε_{280} = 15150 $M^{-1}cm^{-1}$)

MHHHHHHSSGRENLYFQGAAPPSTREMVQLVDSLNRREDGGAFSVDVAEAYPDLRDSYRK
ICPRPMNLILMRQRAKEGYYTSGSATVYGDTVAASLTRLREDIELLVRNCITFNVKVESW
VTLARSFQAFAHRRVDDFVLRHAAFLRGTTMGAEVYE

A.3. Cloning

A.3.1. Cloning approaches for one point mutations

The first approach was a 3-fragment assembly: the linearised pNIC vector with 2 overlapping fragments (Figure A.12A). Both gene fragments were designed to have one overlapping region binding to the linearised vector and 1 region to overlap each other. The overlap between the two gene fragments was designed to overlap on the point of mutation, with primers designed to encode for F rather than N. The fragments were designed to be amplified from the WT plasmid, with the chosen primers, then purified by DNA agarose gel before the 3 fragment Gibson DNA assembly. This approach was not successful due to the small size of the two gene encoding fragments which could not be seen on the DNA gel.

The second approach was a 2-fragment assembly of the linearised pNIC vector with the gene of interest obtained by custom DNA synthesis (Sigma Aldrich) (Figure A.12B). The synthesised genes fragments were dissolved in autoclaved MQ H_2O then combined with linearized pNIC vector using a Gibson assembly (NEBbuilder HiFi DNA assembly) kit. The assembly product was transformed into NEB 5α cells and grown on Kanamycin agar plates.

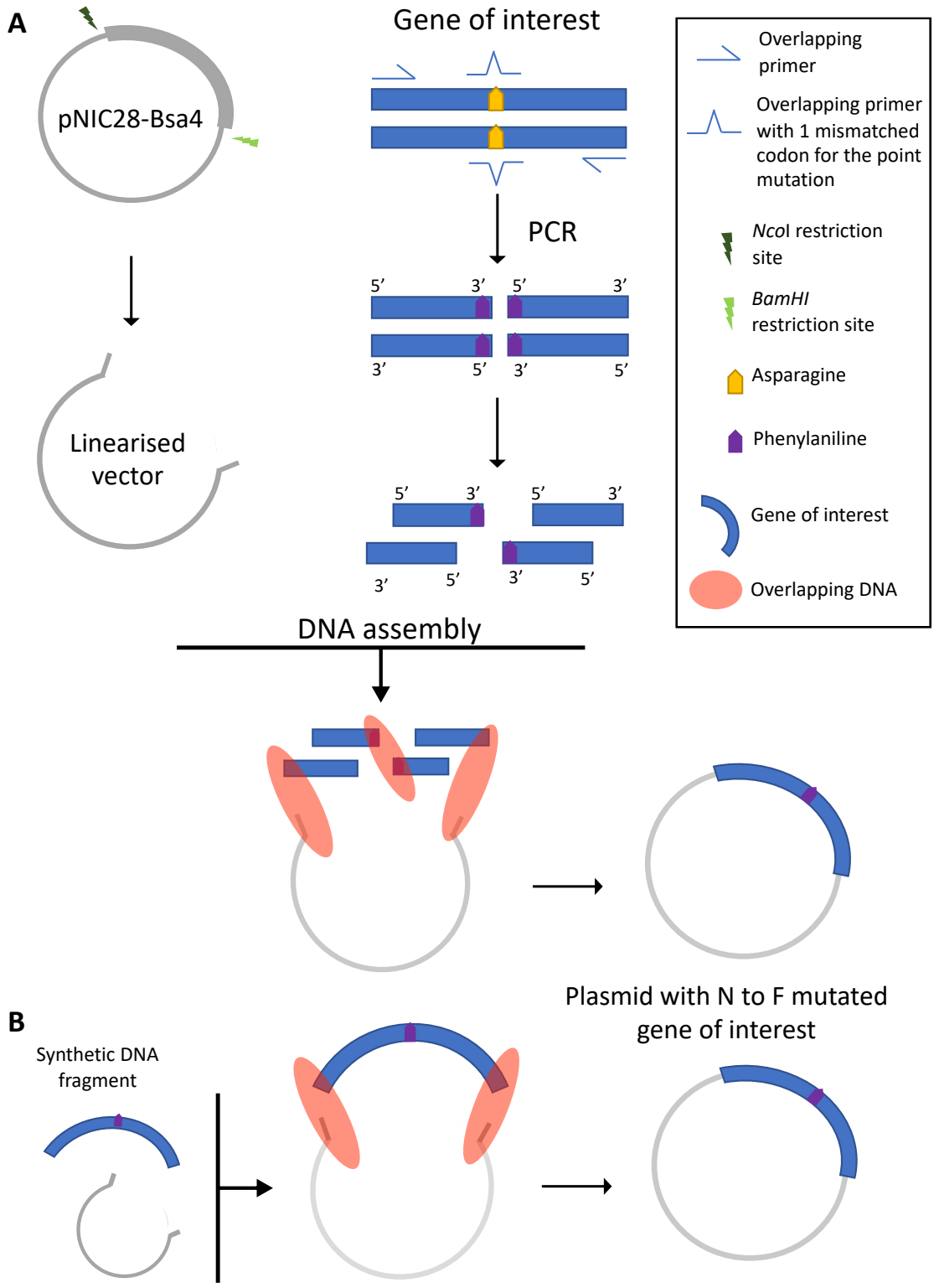


Figure A.12: Schematic summary of the two approaches used for single point mutation. (A) 3-fragment assembly relies on the use of a primer binding the region to be mutated with 1 mismatch corresponding to the mutation (B) 2-fragment assembly of the linearised vector and the synthetic DNA fragment.

A.3.2. Cloning of pTcBDF3(32..148aa)-N124F:

Vector: pNIC28-Bsa4 (7,283 bp) → *Nco*I-HF + *Bam*HI-HF (CutSmart) → 5,337 bp

PCR-1: pTcBDF3(32..148aa) + GA_f2 + TcBDF3_N124F_r → p1 (305 bp)

PCR-2: pTcBDF3(32..148aa) + TcBDF3_N124F_f + GA_r2 → p2 (110 bp)

Gibson assembly: Linearized vector + p1 + p2 → **pTcBDF3(32..148aa)-N124F** (5,697 bp).

Sequencing primer(s): T7F, T7R

A.3.3. Cloning of pTcBDF3(32..148aa)-N124A:

Vector: pNIC28-Bsa4 (7,283 bp) → *Nco*I-HF + *Bam*HI-HF (CutSmart) → 5,337 bp

PCR-1: pTcBDF3(32..148aa) + GA_f2 + TcBDF3_N124A_r → p3 (305 bp)

PCR-2: pTcBDF3(32..148aa) + TcBDF3_N124A_f + GA_r2 → p4 (110 bp)

Gibson assembly: Linearized vector + p3 + p4 → **pTcBDF3(32..148aa)-N124A** (5,697 bp).

Sequencing primer(s): T7F, T7R

Table A.1: Oligo primers used in cloning and sequencing *TcBDF3* gene constructs.

Primer	Orientation	Sequence (5' to 3')[1]	Tm (°C)[2]
T7F	Forward	TAATACGACTCACTATAGGG	55
T7R	Reverse	GCTAGTTATTGCTCAGCGG	61
GA_f2	Forward	AGAACCTGTACTTCCAATCC	61
TcBDF3_N124F_r	Reverse	GGCGCCAAAATACTTCTC	59
TcBDF3_N124F_f	Forward	GAGAAGTATTTTGGCGCC	59
TcBDF3_N124A_r	Reverse	GGCGCCCGCATACTTCTC	59
TcBDF3_N124A_f	Forward	GAGAAGTATGCGGGCGCC	59
GA_r2	Reverse	CGGAGCTCGAATTCG	60

[1]Regions for initial annealing to DNA template are underlined.

[2]Melting temperatures (Tm) were predicted by NEB Tm Calculator (http://tmcalculator.neb.com/#!/) using the settings for Q5 High-Fidelity DNA Polymerase. Only the underlined sequences were used to calculate the Tm.

*Tc*BDF2 N86F fragment	AGAACCTGTACTTCCAATCCATGGGGAAGCGTGGGCGTGAGGGGACTTTGGACAAG CCGCGCTGCCTTGCTTTTGTACATCAGCTTTGGGACAAGGACAAGCTCAAGATGTTTC ACCACCCAATCAGCGCTGCGGAGCTGCCGGACTACCACAAGGTCATAAATTATCCCGT TGACCTGTCGACAATTCGTCAGGGCATTGAGTCCGGAAAGTACGACAGCGATGCGGA TGTGCAGAACGCCGTTGCGCAGATGATTGCCAACGCACTGGAGTACTTCGCAAAGGG GACTGAGTGGCACCAACAGGCGCTGTCGTTCCGAAACATTTACTTGGACGTTGCAAG GCAATGCGGGCTGAGTTGATAAGGATCCGAATTCGAGCTCCG
*Tc*BDF3 N125F fragment	AGAACCTGTACTTCCAATCCGGAAAACACGTGTTCCGTAGTGATCTCAGTACCGTCCC GGAGGTTCGTGACATCTACGACTGCCTGTACAAACTTTACGCGGAGGAGAGTGCATC GGCGTCCTTTCGGGAGCCTGTAAACGCACTGGAATTGGGTGTTTTCAACTACTATGAG GTTGTGACGGAGCCCATGTCGCTGCGTACCGTGCTAGACCGCATTGCCGAGGGGGGT CACTACTCTCAAGCTTCTCAGGTGTTGGCAGATGTGGAGAAGATTTGGAGCAACTGC GAGAAGTATTTTGGCGCCGACTCCGCTTTGGTTAAAGAGGCCAAAAAATGCCAGGGG ATACTGACAAGGCTGCGGGAGCGTCTTTGATAAGGATCCGAATTCGAGCTCCG
*Tc*BDF5_1 N91F fragment	AGAACCTGTACTTCCAATCCATGGAAGAACGTCGCCAGAAATTCTATCCGGAAGAAG AACTGGTTGCACTGGTTCGTAGCCTGGATCGTCCGCAGGATGAAGGTCTGTTTAGCAT GGATGTTCTGGTTGTTTATCCGCATCTGGAACAAGAATATACCCGTGTTTGTCCGAAA CGTTGTGATCTGGCAACCGCAGCAGAAAAAGCAGCAAATGAAGCATATAGCTATGAT GTGAATCTGACCGCACTGCGCGAAGATATCAAACTGATGGTGAATAATTGCTATCGCT TCTTCGGCACCAAAGGTCCGCTGGCAAATATTGCAGAACGTTTTGAAGCATTCGCCAA AGAACAAATTGATGCCTACGTTACCAAATGATAAGGATCCGAATTCGAGCTCCG
GA_f2	AGAACCTGTACTTCCAATCC
GA-r2	CGGAGCTCGAATTCG

A.3.4. Cloning of pSUMO3-TcBDF5_2(172..324aa):

Vector: pETM11-SUMO3GFP (6,290 bp) → *Hind*III + *Bam*HI-HF (CutSmart) → linearized vector (5,567 bp)

PCR: pH6-TcBDF_5_2(172..324aa) (5,804 bp) + GA_f_sumo + GA_f_sumo → fragment (502 bp)

Gibson assembly: Linearized vector + fragment → **pSUMO-TcBDF_5_2(172..324aa)** (6,038 bp)

Sequencing primer(s): T7F or T7R

Primer design by Dr Matthias Schiedel.

Table A.2: Oligo primers employed in this study.

Sequencing Primer	Orientation	Sequence (5' to 3') [1]	Tm (°C) [2]
T7F [3]	Forward	TAATACGACTCACTATAGGG	55
T7R [3]	Reverse	GCTAGTTATTGCTCAGCGG	61
Cloning Primer	**Orientation**	**Sequence (5' to 3') [1]**	**Tm (°C) [2]**
GA_f_sumo	Forward	AGCAACAGACCGGTGG*ATCC*CGTAGCGGTG	61
GA_r_sumo	Reverse	AGTGCGGCCGCAAGCTTTTATCAAAAATCATCGATTTTTTTGTGTG	61

[1] Regions for initial annealing to DNA templates are underlined.

[2] Melting temperatures (Tm) are predicted by NEB Tm Calculator (http://tmcalculator.neb.com/ - !/) using the settings for Q5 High-Fidelity DNA Polymerase.

[3] These DNA sequencing primers are available at Source BioScience (http://www.sourcebioscience.com).

A.3.5. Cloning of MBP-TcBDF5(172..324aa)

Vector: pNIC28-Bsa4 (7,283 bp) → *Nde*I + *Bam*HI-HF (CutSmart) → linearized vector (5,271bp)

PCR-1: pRK793 (Addgene #8827) + H6MBP_GAf + H6MBP_r → pa (1,239 bp)

PCR-2: pTcBDF5(172..324aa)(5,804 bp) + MBPTcBDF5_f + GA_r → pb (510 bp)

PCR-3: pTcBDF5(172..337aa)(5,843 bp) + MBPTcBDF5_f + TcBDF5_2r1 → pb (510 bp)

PCR-4: pTcBDF5_1,2(1..337aa)(6,356 bp) + MBPTcBDF5_f 2+ TcBDF5_2r1 → pb (510 bp)

Method-1 – Gibson assembly: Linearized vector + pa + pb → **pH_6MBP-TcBDF5(172..324aa)** (6,941 bp)

Sequencing primer(s): T7F, MBP_f, T7R; Primer design by Dr Matthias Schiedel.

Table A.3: Oligo primers employed in this study.

Sequencing Primer	Orientation	Sequence (5' to 3') [1]	Tm (°C) [2]
T7F [3]	Forward	TAATACGACTCACTATAGGG	55
T7R [3]	Reverse	GCTAGTTATTGCTCAGCGG	61
MBP_f	Forward	GTCTGACCTTCCTGGTT	61
Cloning Primer	**Orientation**	**Sequence (5' to 3') [1]**	**Tm (°C) [2]**
H6MBP_GAf	Forward	TTAACTTTAAGAAGGAGATATACATATGCACCATCATCATCATCATAAAACTGAAGAAGGTAAACTGGT	(69), 61
H6MBP_r	Reverse	TCACCGCTACGGGATTGAAAATAAAGATTTTCTCCC	(55), 62
MBPTcBDF5_f	Forward	GGGAGAAAATCTTTATTTTCAA*TCCCGTAGCGGT*	(55), 61
MBPTcBDF5_f2	Forward	GGGAGAAAATCTTTATTTTCAATC*CCGTAGCGGTGAAATG*	(59), 60
GA_r	Reverse	CGGAGCTCGAATTCG	60
TcBDF5_2r1	Reverse	CGGAGCTCGAATTCGGATCCTTATCAAAAATCATCGATTTTTTTGTGTGC	60

[1] Regions for initial annealing to DNA templates are underlined.

[2] Melting temperatures (Tm) are predicted by NEB Tm Calculator (http://tmcalculator.neb.com/ - !/) using the settings for Q5 High-Fidelity DNA Polymerase.

[3] These DNA sequencing primers are available at Source BioScience (http://www.sourcebioscience.com).

A.4. Protein expression

A.4.1. Towards soluble fusion proteins of H_6-*Tc*BDF5-2

The small scale expression trials did not identify any soluble H_6-*Tc*BDF5-2 constructs. This result suggests that the constructs encoding for truncated proteins with varying start and end amino acid, compared to the full length protein, are either not folded during the expression in *E. coli*, or are folded correctly after expression but aggregate over time. BDs exist as multi-domain BCPs in a cellular context, therefore, producing a soluble single domain protein construct can be challenging. There is no guarantee that truncating a soluble multi-domain BCP to a single domain construct will produce a soluble protein when expressed recombinantly in *E. coli*. We next explored is the addition SUMO to create a soluble fusion protein. The fused domains can favour the correct folding of the studied protein, and prevent protein aggregation.[9] The plasmid *p*SUMO-*Tc*BDF5(172..324aa) (Figure A.10), designed to code for the fusion proteins of *Tc*BDF5-2 with SUMO (with Dr Matthias Schiedel), was cloned and purified.

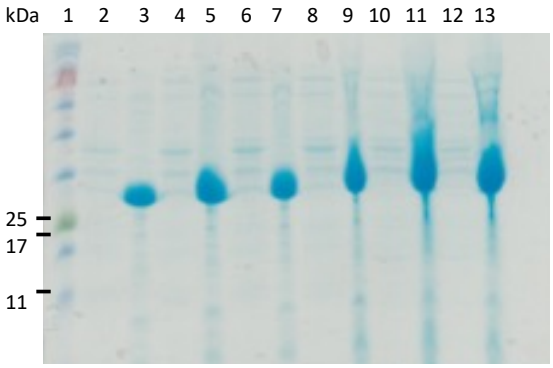
To produce the plasmid, a 3-fragment DNA assembly approach was employed.[10] The linearised vector was combined with a DNA fragment encoding for the solubilising protein and with the fragment encoding for *Tc*BDF5-2. The assembled plasmid was purified and then transformed into NEB-5α cells. Sanger sequencing confirmed successful DNA assembly.

Small-scale expression of pSUMO-*Tc*BDF5-2 was optimised by comparing the expression levels by SDS-PAGE gel when different IPTG amounts and post induction temperatures (Table A.4). Large amounts of IPTG will corresponds to a very large activation of the gene expression, which could lead to rapid protein production unfolded protein. Inversely, low amounts of IPTG might not induce the expression efficiently and lead to low amounts of protein being produced. The temperature post-induction can also affect the folding of the protein, which lower temperatures giving a slower protein production, which might be necessary for the folding of the protein.

Table A.4: Summary of the variations of IPTG and post induction temperatures used for the small-scale expression of pSUMO-*Tc*BDF5-2.

Post induction temperature	IPTG concentration			
16 °C	0.05 mM	0.1 mM	0.5 mM	1.0 mM
37 °C	–	–	0.5 mM	1.0 mM

Clear expression bands were observed for all conditions, which was a clear improvement from the *Tc*BDF5-2 constructs trialled previously. However, the protein was only seen in the insoluble fractions (Figure A.10). This observation indicate that the SUMO-*Tc*BDF5-2 would also be likely to aggregate during large scale expression and purification. To overcome, an alternative solubilising protein, such as MBP, could be explored. However, this was not pursued as we decided to focus our efforts on crystallisation and screening of *Tc*BDF3.



SDS-PAGE gel:

1. Colour protein standard
2. Soluble fraction H_6-SUMO-*Tc*BDF5-2 with 1.0 mM IPTG, 16 °C
3. Insoluble fraction H_6-SUMO-*Tc*BDF5-2 with 1.0 mM IPTG, 16 °C
4. Soluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.5 mM IPTG, 16 °C
5. Insoluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.5 mM IPTG, 16 °C
6. Soluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.1 mM IPTG, 16 °C
7. Insoluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.1 mM IPTG, 16 °C
8. Soluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.05 mM IPTG, 16 °C
9. Insoluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.05 mM IPTG, 16 °C
10. Soluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.5 mM IPTG, 37 °C
11. Insoluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.5 mM IPTG, 37 °C
12. Soluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.1 mM IPTG, 37 °C
13. Insoluble fraction H_6-SUMO-*Tc*BDF5-2 with 0.1 mM IPTG, 37 °C

Figure A.13: SDS-PAGE (16% Tris/Gly/SDS gel) analysis of H_6-SUMO-*Tc*BDF5-2 expression trial. Lane 1: Colour protein standard, broad range (New England Biolabs, Inc.), lane 2 to 13: Comparison of the expression levels in the supernatant (soluble fraction) or the pellet (insoluble fraction) following lysis with BugBuster®

A.4.2. Expression of H6-TcBDF2(1..113aa) and H6-TcBDF3(32..148aa)

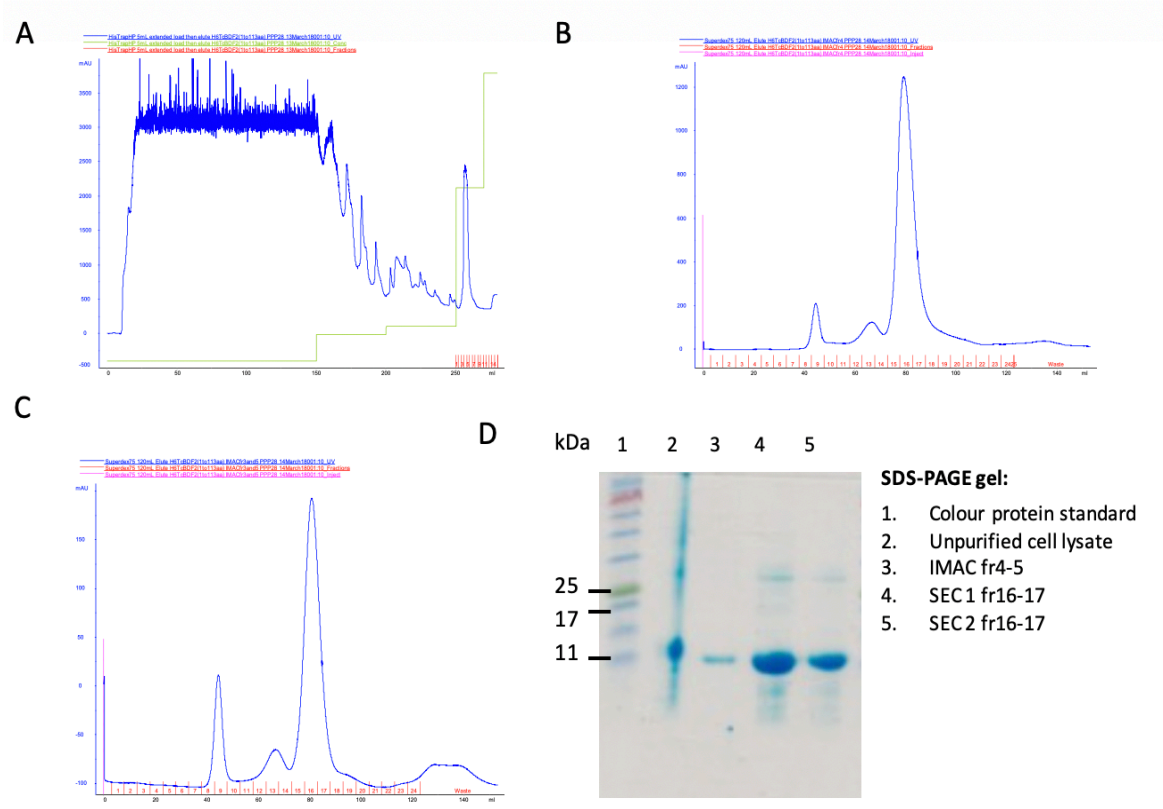


Figure A.14: H_6-*Tc*BDF2(1..113aa) expression for biophysical assays: A. IMAC profile. B. SEC profile of IMAC fr4 (SEC 1). C. SEC profile of IMAC fr5 (SEC 2). C. SDS-PAGE (16% Tris/Gly/SDS gel) analysis of purified H_6-*Tc*BDF2(1..113aa); Lane 1: Colour protein standard, broad range (New Englad Biolabs, Inc.), lane : unpurified cell lysate, lane 3: IMAC fraction 4 to 5, lane 4: SEC1 fr 16-17, lane 5: SEC 2 fr16-17.

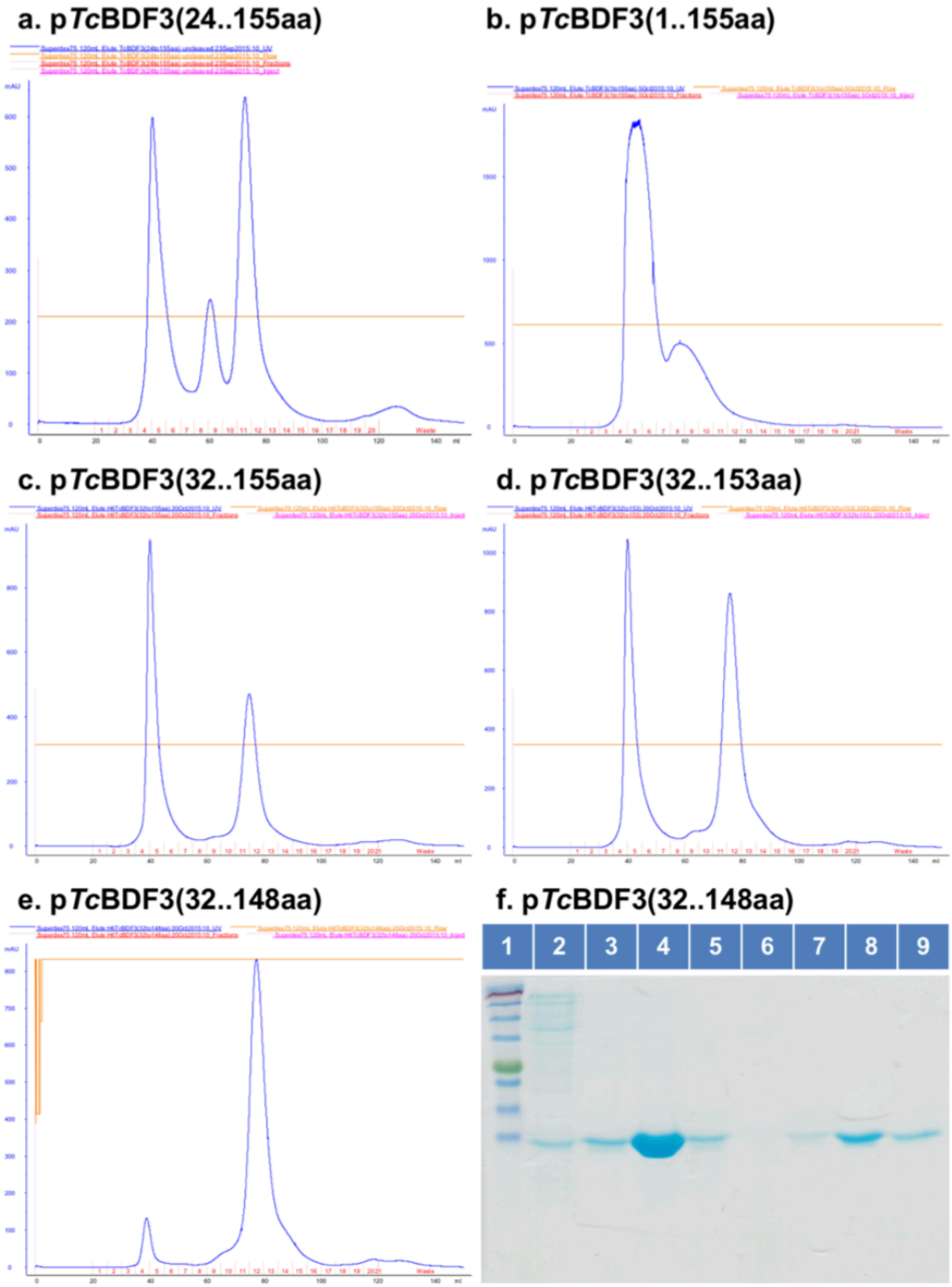


Figure A.15: Purification of various *Tc*BDF3 constructs by Dr Anthony K. Chan: a-e: SEC profiles, showing aggregation (first eluting) peak and folded protein (last eluting) peaks of the studied protein constructs: H_6-*Tc*BDF3(32..148aa) shows minimum aggregation in the SEC profile. f. SDS-PAGE (16% Tris/Gly/SDS gel) analysis of purified H_6-*Tc*BDF3(32..148aa).

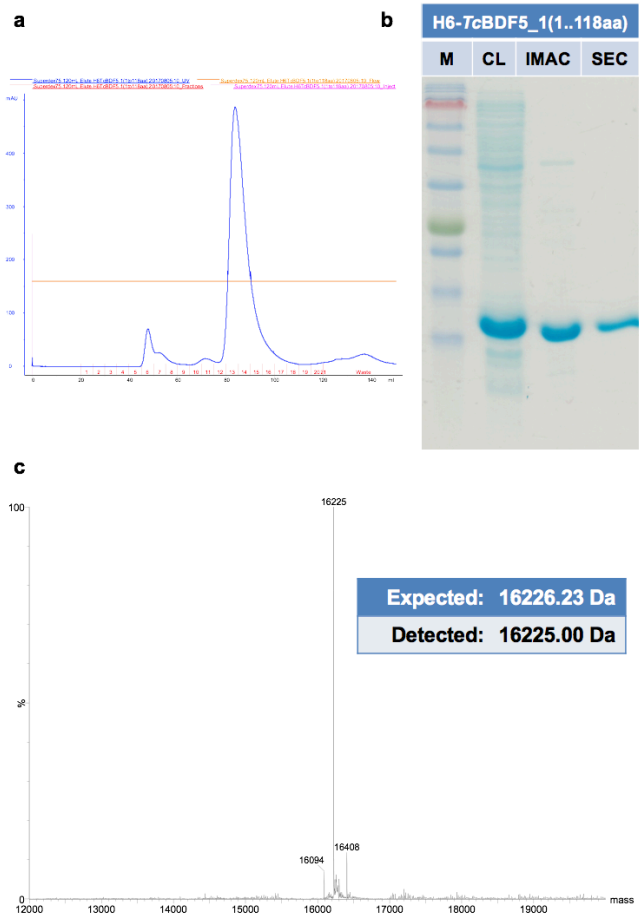


Figure A.16: Expression and purification of H_6-*Tc*BDF5a (1..118aa) by Dr Anthony Chan: A. SEC profile. B. SDS-PAGE (16% Tris/Gly/SDS gel) analysis of purified H_6-*Tc*BDF5(1..118aa): M= colour protein standard, broad range (New England Biolabs, Inc.), CL= unpurified cell lysate.

A.4.3. TEV cleavage of H6-TcBDF3(32..148aa)

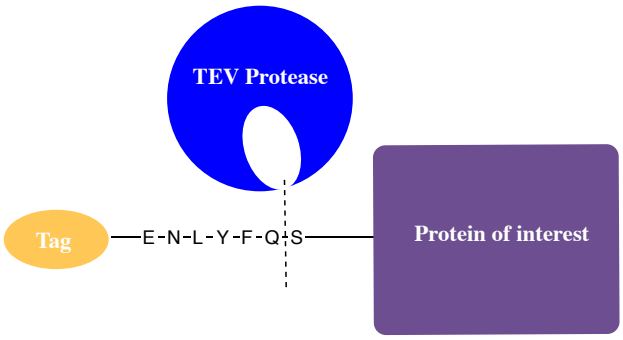


Figure A.17: Schematic explanation of the TEV protease cleaving process. The enzyme cleaves the sequence between the glutamine and the serine residues.

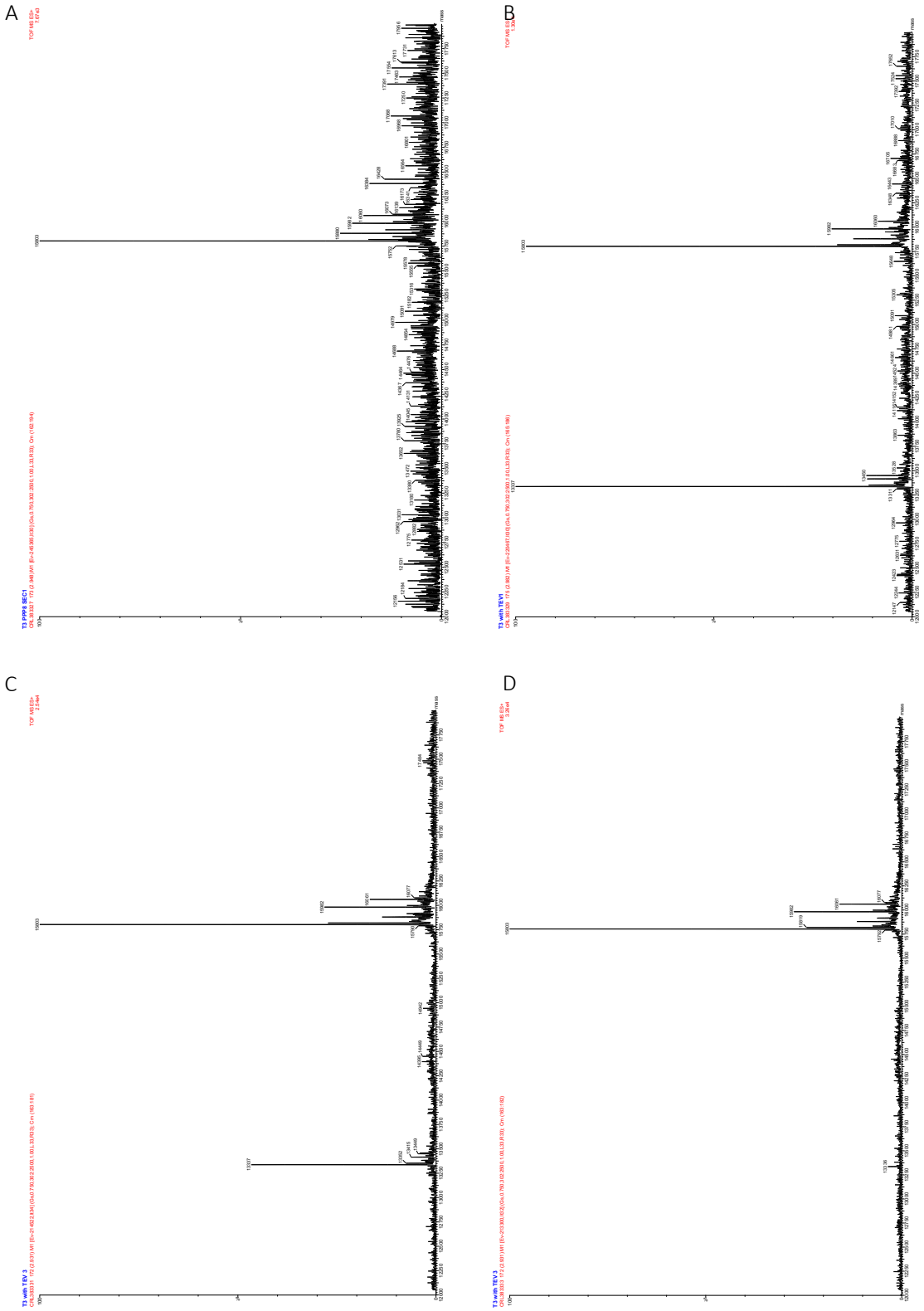


Figure A.18: Mass Spectrometry spectra of A. H_6-TcBDF3(32..148aa) in Tris buffer, and H_6-*Tc*BDF3(32..148aa) (20 µg) incubated 20h at 4 °C with B. TEV1 (10 U), C. TEV2 (0.1 µg) or D. TEV3 (10 U).

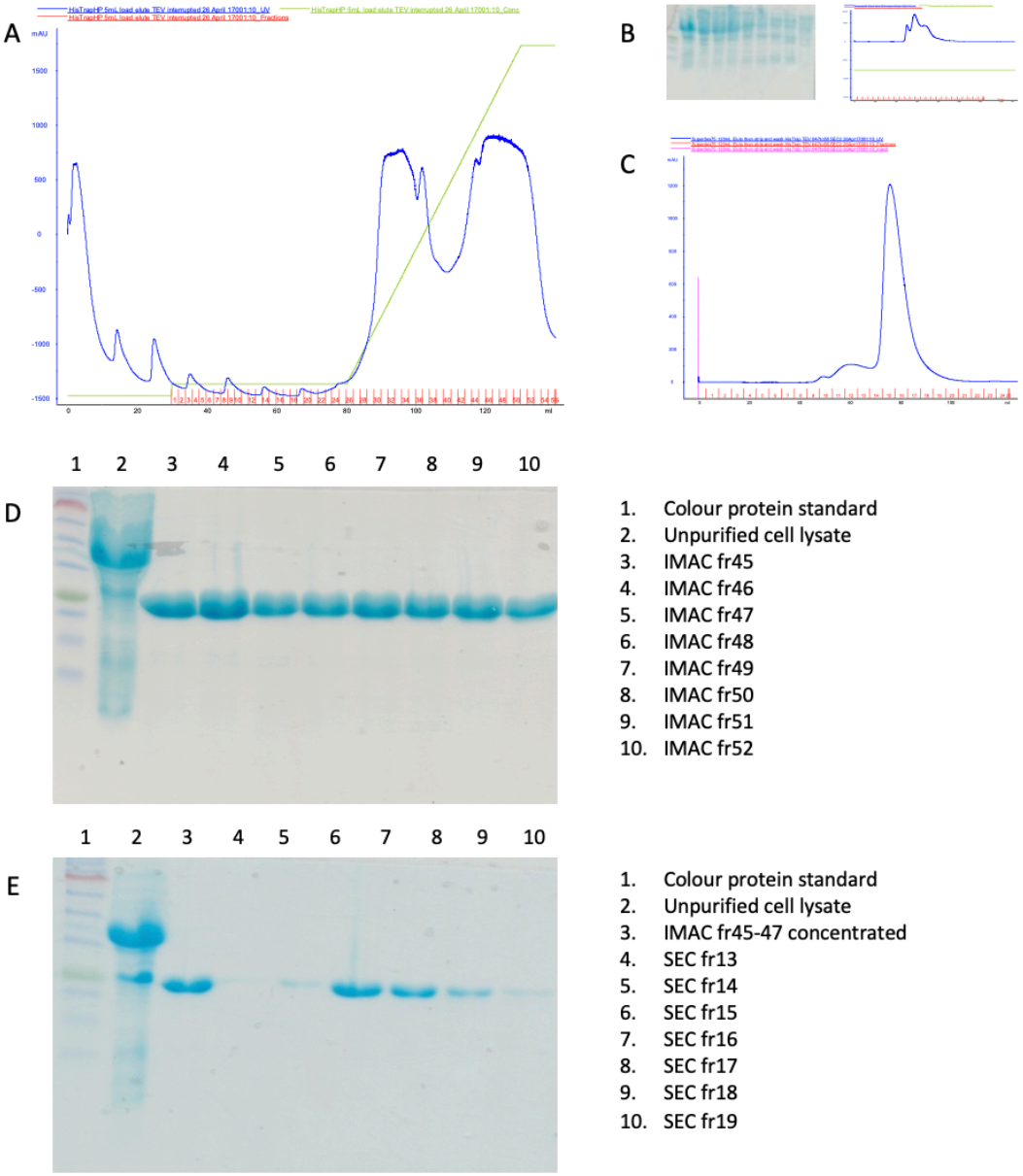
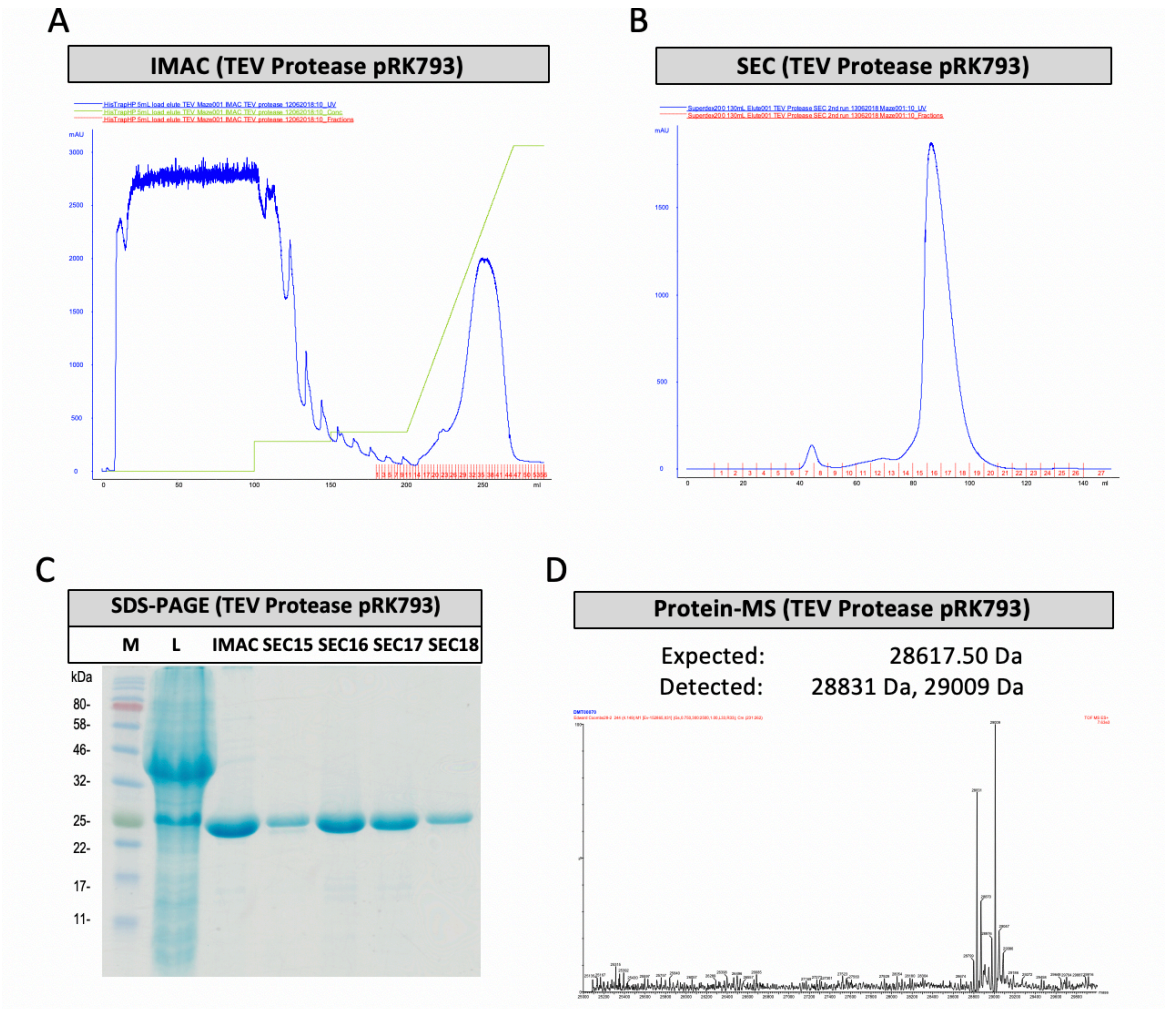


Figure A.19: Purification of H_6-rTEV (pTH26, gift from Dr Illaria Pettinati, Schofield group, Oxford). A. IMAC purification profile B. Protein aggregation in IMAC fr 30-40 observed by SDS-PAGE gel (16% Tris/Gly/SDS gel) and SEC profile. C. SEC profile of IMAC fraction 45-52. D. SDS-PAGE (16% Tris/Gly/SDS gel) analysis of purified H_6-TEV; Lane 1: Colour protein standard, broad range (New England Biolabs, Inc.), lane 2: unpurified cell lysate, lane 3 to 10: IMAC fraction 45 to 52. E. SDS-PAGE (16% Tris/Gly/SDS gel) analysis of purified H_6-TEV; Lane 1: Colour protein standard, broad range (New England Biolabs, Inc.), lane 2: unpurified cell lysate, lane 3: IMAC fraction 47-50 concentrated, lane 4-10 SEC fr 13 to 19.



gure A.20: Purification of TEV Protease (pRK793). (A) IMAC profile, (B) SEC profile, (C) SDS-PAGE gel (M = marker, L = Lysate, IMAC = collected eluate of immobilized metal ion affinity chromatography, SEC = collected eluate of size exclusion chromatography, CL = cell lysate), and (D) spectrum of protein mass spectrometry. Purification performed by Dr Matthias Schiedel.

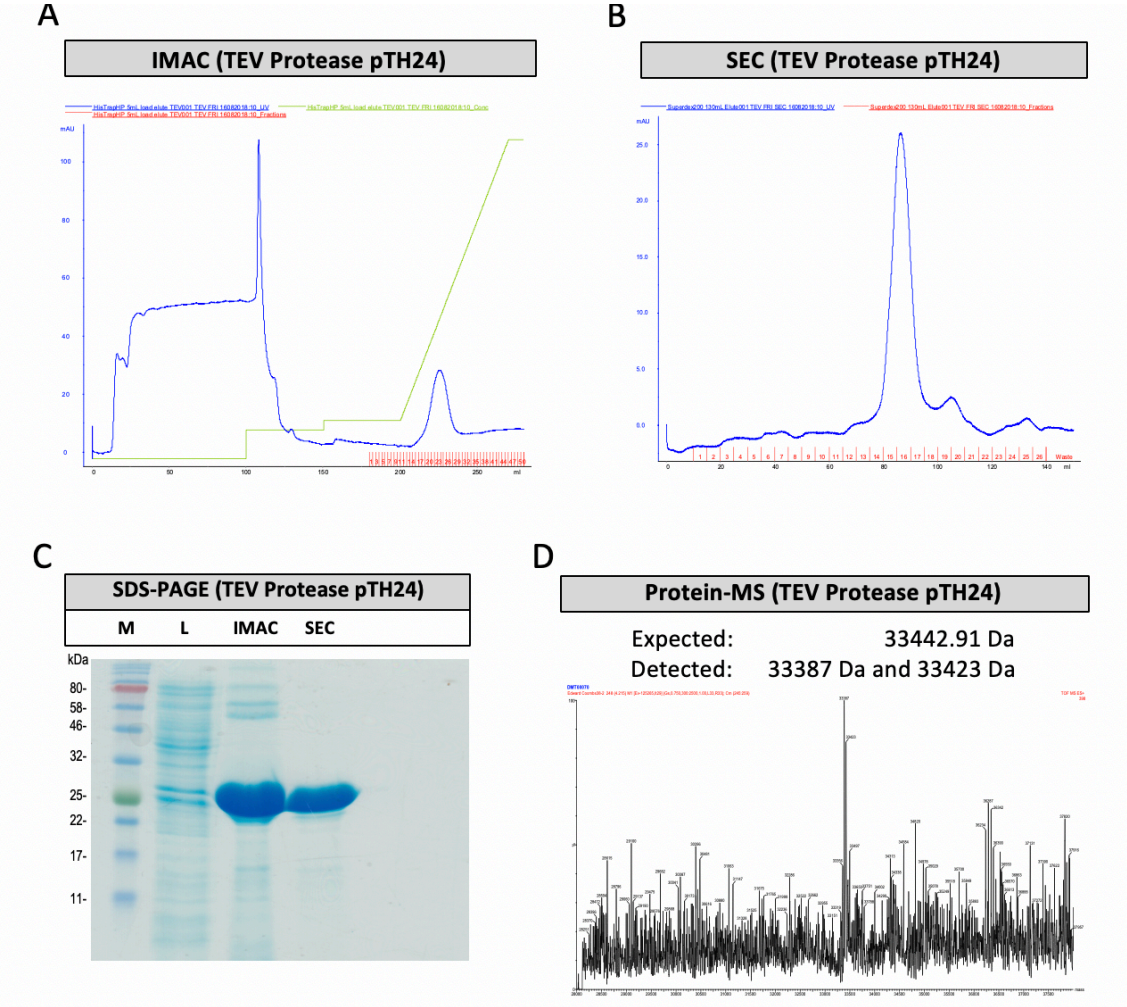


Figure A.21: Purification of TEV Protease (pTH24). (A) IMAC profile, (B) SEC profile, (C) SDS-PAGE gel (M = marker, L = Lysate, IMAC = collected eluate of immobilized metal ion affinity chromatography, SEC = collected eluate of size exclusion chromatography, CL = cell lysate), and (D) spectrum of protein mass spectrometry. Purification performed by Dr Matthias Schiedel.

A.4.4. Crystallisation trials

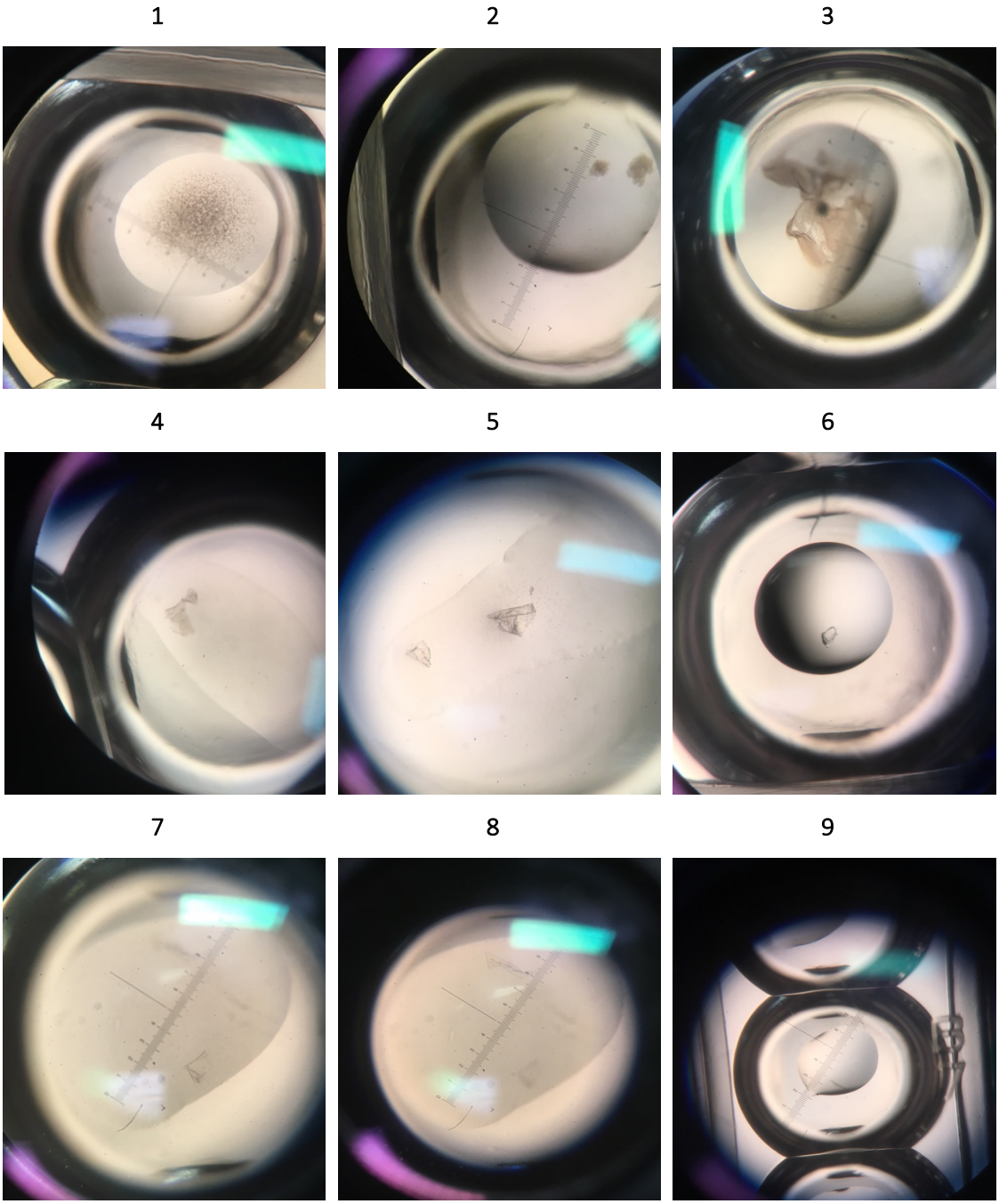


Figure A.22: *Tc*BDF3 crystallisation trials: Pictures of selected examples of aggregation (1-3), and crystal formation (salt and/or protein) (4-9).

Appendix B. Biophysical assays

B.1. Thermal shift

B.1.1. Thermal shift assay optimisation

The assay conditions, especially the buffer composition and the ratio of dye to protein can influence protein melting temperature. Ideally, the melting temperature of the apo protein should be above 35 °C to prevent partial denaturation at room temperature.[31] To identify suitable assay conditions, the buffer stability of the proteins was determined. The melting temperatures T_m of H_6-*Tc*BDF2 and H_6-*Tc*BDF3 was measured in triplicate for 16 different buffer and pH combinations (Figure B.1). The differential scanning thermometry (DSF) assay was performed with the standard conditions described by Niesen *et al.*:[31] Protein concentration of 2 µM and a 1:1000 SYPRO Orange® dye dilution. TAPS at pH 8.5 and HEPES at pH 7.5 were identified as appropriate buffers as they display T_m above 35 °C and small errors for both proteins.

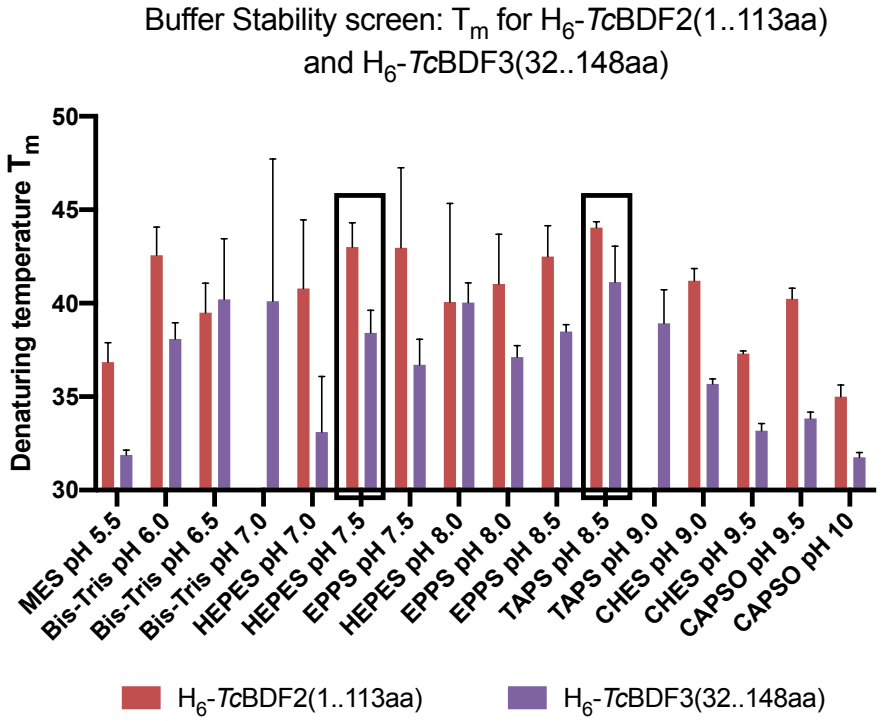
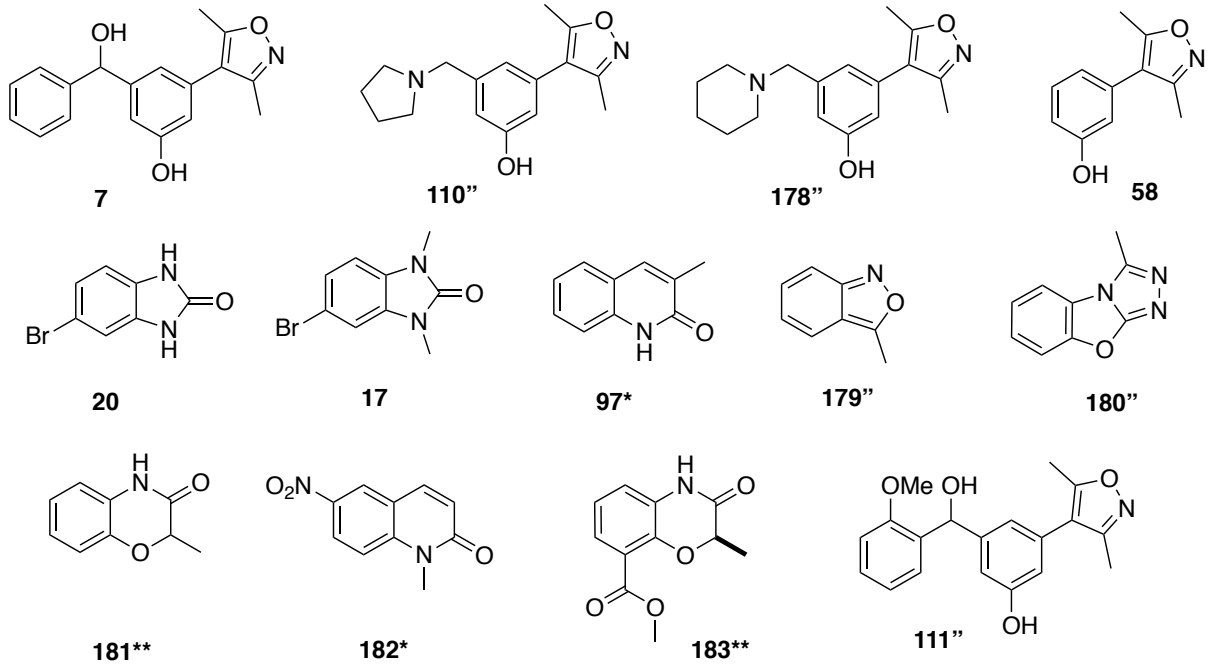


Figure B.1: The melting temperature of H_6-*Tc*BDF2(1..113aa) and H_6-*Tc*BDF3(32..148aa) in the various screened buffers

The first attempt to screen compounds by thermal shifts ΔT_m was not successful. The measured fluorescence intensity was low, especially for H_6-*Tc*BDF2 for which the curve fitting wasn't possible. Although, data processing was possible for H_6-*Tc*BDF3, and some non-zero shifts are observed, a significant error within triplicates is observed, due to the low fluorescence intensity, rending the thermal shifts ambiguous (Figure B.2).



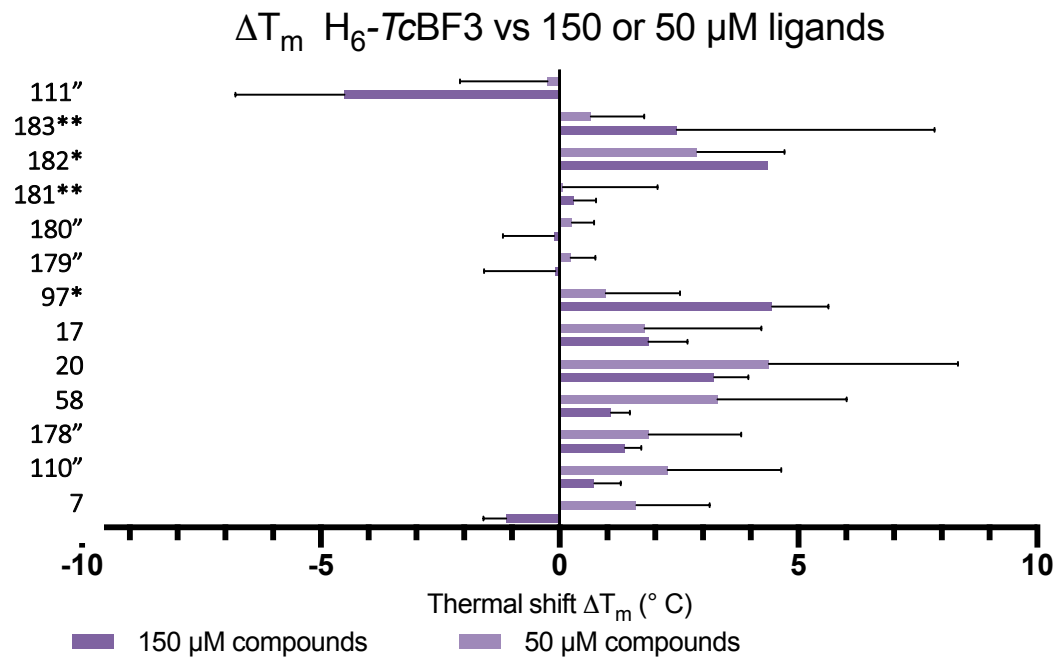


Figure B.2: The thermal shift data for the screening of 13 compounds against H_6-*Tc*BDF3. The large error bars arise from the low fluorescence intensities observed. The data are ambiguous and demonstrate that the assay requires further optimisation. " denotes compouds synthesised by Dr L. Jennings, * denotes compouds synthesised by F. Woodhouse, ** denotes compouds synthesised by Dr M. Brand.

The low fluorescence intensity observed during the first compound screen led to poor curve fitting and large standard deviations. The fluorescence of the dye can be affected by several factors: the protein to dye ratio, the lamp, the PCR plate and to a lesser extent the buffer.[31] The screening conditions used were 2 µM of protein, 1:1000 dye dilution (a 5x dye concentration) using black 96-well plates. To optimise the fluorescence reading several ratio of dye to protein were tested (Table B.1). The fluorescence intensity of the H_6-*Tc*BDF2 and H_6-*Tc*BDF3 proteins were recorded by DSF for all conditions in triplicate on both a black 96-well PCR plate and on a white 96-well PCR plate.

Table B.1: Summary of the optimisation of the protein:dye ratio in white or black PCR plates.

Conditions	Protein	Sypro® orange
1	2 µM	5×
2	2 µM	10×
3	5 µM	10×
4	5 µM	15×
5	8 µM	15×
6	8 µM	20×

The fluorescence intensity maxima for both H_6-*Tc*BDF2 and H_6-*Tc*BDF3 show that the best screening conditions are 5 µM of protein with 10× of dye (condition 3, Table B.1) on white PCR plate. Typically, black PCR plates are used to diminish cross-talk between wells; however, the fluorescence intensity with black PCR plates was too low. High concentration of dye can lead to destabilisation of the protein and lower of the protein melting temperatures T_m. Since the melting temperatures T_m for both for H_6-*Tc*BDF2 and for H_6-*Tc*BDF3 are above 40 °C (45.14 and 41,93 °C respectively, Figure B.3) using the optimal conditions (condition 3, on white 96-well plate, Table B.1), these conditions were chosen for subsequent screening of compounds.

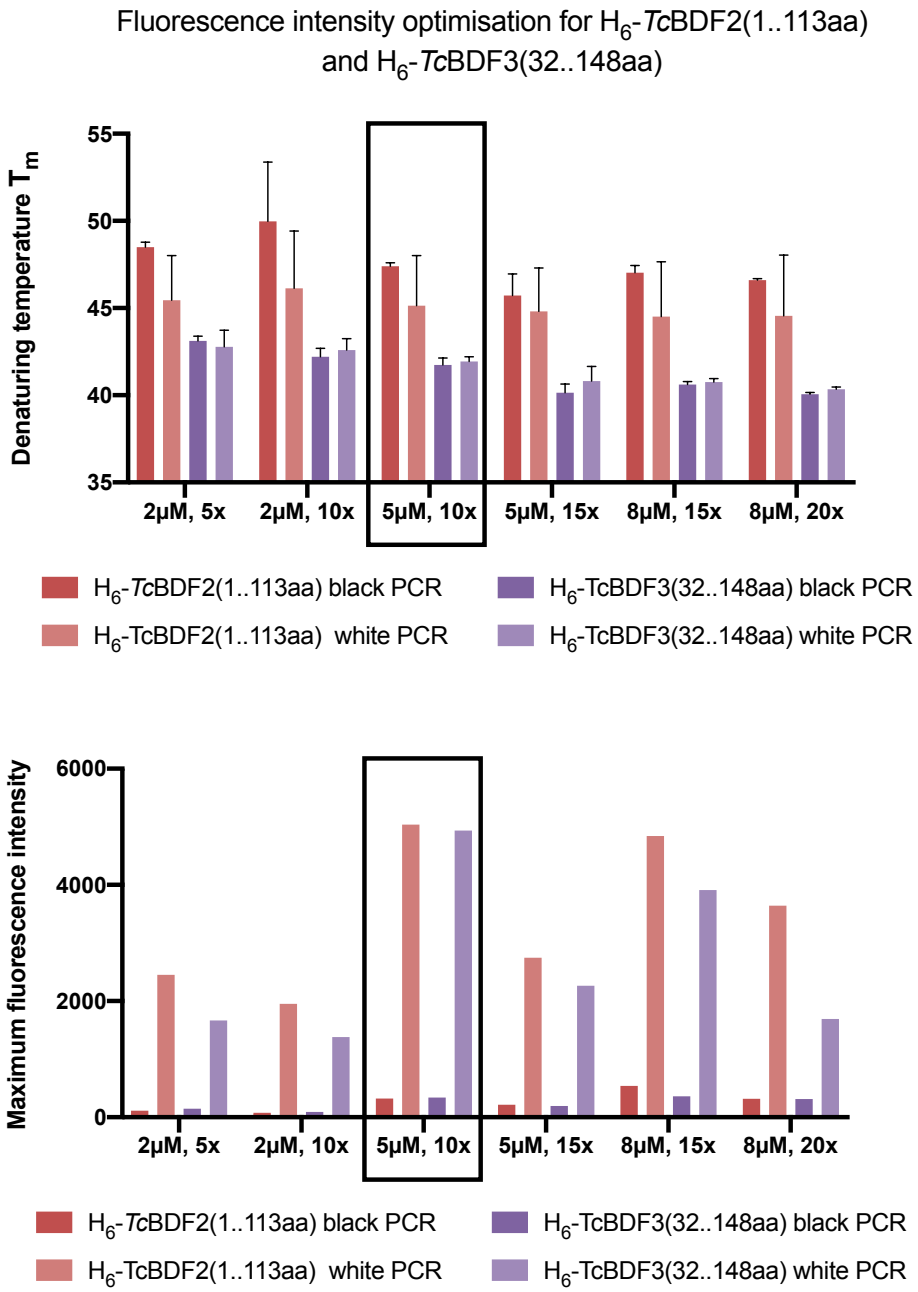


Figure B.3: Fluorescence optimisation screen for DSF: maximum fluorescence intensity and denaturing temperatures given for the various protein to dye ratios in white or black 96-well PCR plates.

B.1.2. Thermal shift screen results

A focused library of compound was created by combining compounds synthesised as part of Chapter 1B and compounds synthesised for human BD proteins by Dr David Hewing, Dr Laura Jennings, Dr Brian Wilson, Dr Timothy Rooney, Dr Michael Brand and Fern Woodhouse. The compounds were screened against *Tc*BDF2 and *Tc*BDF3 using the optimised thermal shift screening conditions.

Table B.2: Combined results of the DSF compound screens using 150 µM compound, 8 µM of protein and 10× Sypro™ Orange. Thermal shifts, ΔT_m, are given as a difference to DMSO control well; averages between runs are not given due to the disruptions observed between runs. Heat map shows relative ΔT_m, red indicates large ΔT_m, and green indicates small ΔT_m. The standard deviations (SD) are given for each triplicates, heat map shows relative SD, light grey indicates small SD, and dark grey indicates large SD. " denotes compounds synthesised by Dr L. Jennings, * denotes compounds synthesised by F. Woodhouse, ** denotes compounds synthesised by Dr T. Rooney, ° denotes compounds synthesised by Dr M. Brand, "' denotes compounds synthesised by C.A.F.

		*Tc*BF2		*Tc*BF3	
ID	MW	SD	ΔT_m	SD	ΔT_m
85	327.2	0.09	0.43	2.66	1.74
80	369.3	0.12	0.51	1.75	1.65
5	415.2	0.07	-0.54	0.42	-0.16
3	457.0	0.27	-0.05	0.15	-1.00
107*	162.2	0.63	-0.04	1.86	1.03
104	238.3	0.33	-0.59	1.46	-0.44
184*	134.1	0.18	0.28	1.70	0.74
185*	240.3	0.39	0.48	0.52	1.12
17	241.1	0.47	0.14	0.64	0.67
		0.33	0.41	0.39	-0.38
18	241.1	0.20	0.90	0.58	1.25
		0.46	0.42	0.01	0.44
54	213.0	0.27	-0.05	2.05	-0.25
		0.19	0.35	0.65	0.37
57	213.0	0.92	-0.36	2.70	0.48
186°	211.2	0.58	0.31	0.69	2.49
		0.25	0.58	0.21	0.51
187°	211.2	0.05	0.66	2.12	3.82
		0.08	0.57	0.23	0.45
189**	173.2	0.17	-0.32	1.22	-1.32
97*	159.2	0.02	1.40	1.26	2.10
		0.23	0.99	0.14	-0.70
188**	175.2	0.25	1.79	0.67	0.71
110"	272.4	0.35	1.06	0.52	4.59
		0.05	0.21	0.32	0.40
		0.19	0.96	1.72	-0.32
178"	286.2	0.09	0.84	0.12	1.72
		0.12	0.18	0.23	0.91
		0.09	-0.04	0.11	0.08
111"	325.4	0.29	-0.85	8.03	1.49
190"	251.4	0.11	0.27	0.21	0.18
191"	296.3	0.11	0.07	0.11	0.08
56	189.2	0.30	0.10	1.30	-0.80
		0.21	0.20	0.13	0.17
123*	173.2	0.56	-0.13	0.83	0.09
		0.08	0.03	1.29	-0.31
		0.13	0.17	0.68	0.70
108	322.4	0.29	0.11	0.45	-0.34
192	188.0	0.48	-0.29	0.71	0.28
7	521.6	1.17	-0.63	N/D	N/D
		1.13	3.12		
		0.4	3.83		

		*Tc*BF2		*Tc*BF3	
ID	MW	SD	ΔT_m	SD	ΔT_m
193	172.2	0.48	-0.45	2.74	3.62
		0.36	-0.01	0.22	0.58
194	267.3	0.56	0.19	0.30	2.30
		0.19	0.05	0.11	0.64
182	204.2	0.96	0.82	0.48	0.40
1	260.3	0.07	0.29	0.28	0.17
2	287.3	0.31	-0.06	0.64	0.13
195"	230.3	0.07	0.22	0.21	0.38
		0.15	0.32	0.19	0.33
196"	214.2	0.18	0.46	0.28	0.04
		0.17	0.30	0.42	0.54
197"	179.2	0.26	0.27	0.35	0.15
		0.13	0.49	0.21	0.39
198"	148.2	0.05	0.27	0.17	0.36
179"	133.2	0.36	-0.02	0.14	0.07
		0.08	0.27	0.54	0.38
180"	173.2	0.36	0.27	0.16	0.10
199"'	180.2	0.12	0.29	0.35	0.52
200"'	190.2	0.13	0.19	0.37	0.41
201°	223.2	0.17	0.16	0.29	0.31
202°	296.3	0.57	0.87	2.22	5.22
		0.19	0.39	0.25	0.62
98**	394.5	0.79	1.18	2.12	4.06
		0.27	1.52	0.13	1.97
204**	457.4	0.64	0.99	1.47	3.24
		0.37	0.87	0.28	0.87
183**	221.2	0.10	0.41	0.48	0.60
181**	163.2	0.28	0.16	0.18	-0.05
205'	438.5	0.32	1.50	1.93	2.22
		0.16	2.43	0.37	-0.03
203'	283.7	0.19	-0.01	0.96	1.42
		0.03	0.23	0.03	0.37
		0.21	0.14	0.19	0.84
99'	582.7	0.11	0.49	2.12	4.99
		0.12	1.02	0.62	1.21
206'	494.6	0.39	0.48	0.57	0.97
		0.06	-0.42	0.25	-0.33
207'	589.7	0.47	0.18	0.45	1.44
		0.12	-0.62	0.29	-0.32
67	249.69	1.2	0.77	N/D	N/D
		1.3	2.13		

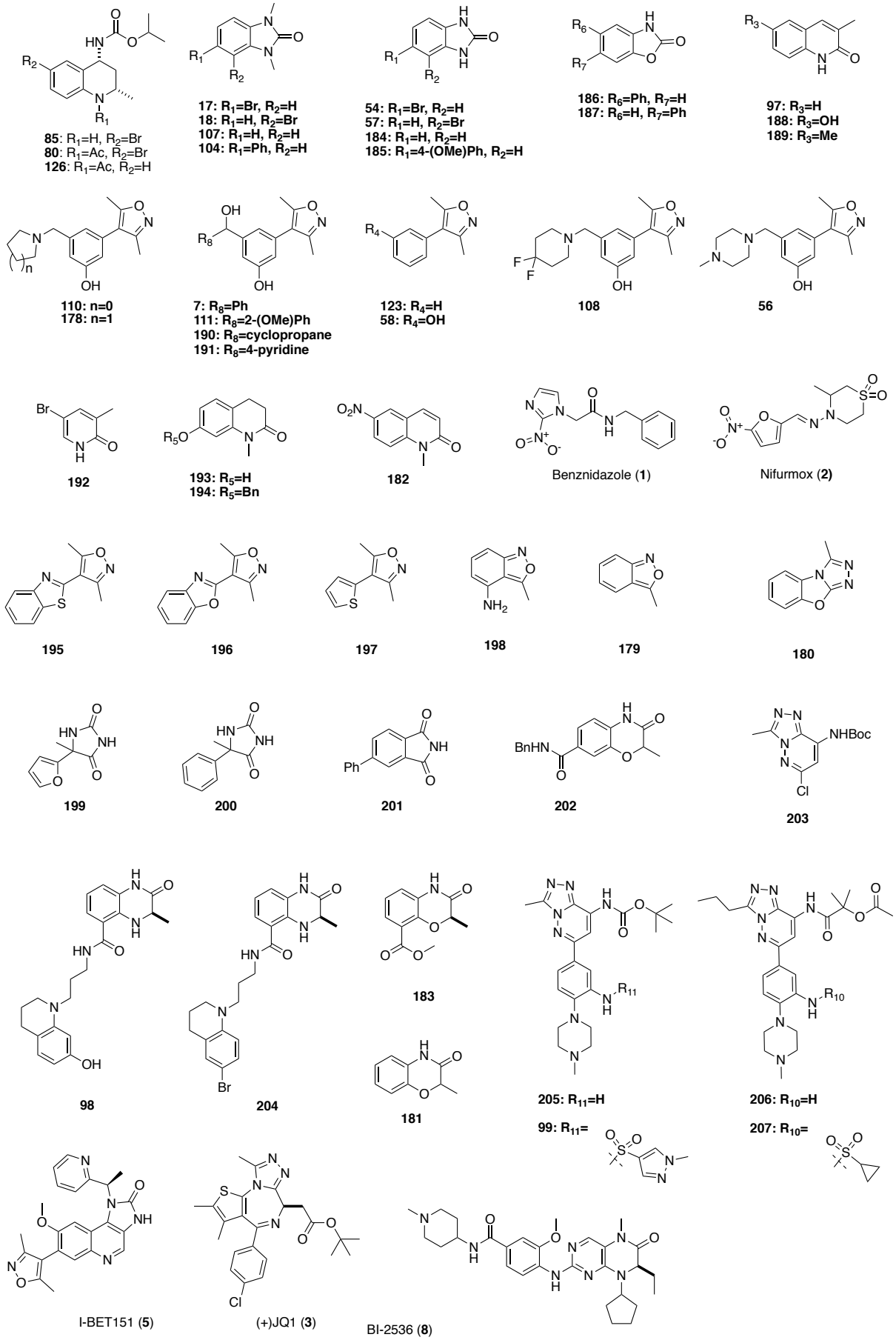


Figure B.4: Structures of the compounds screened using the DSF assay.

B.2. WaterLOGSY spectra

B.2.1. Spectra against *Tc*BDF2, *Tc*BDF3, *Tc*BDF5-1 and *Sm*BRD3(1,2)

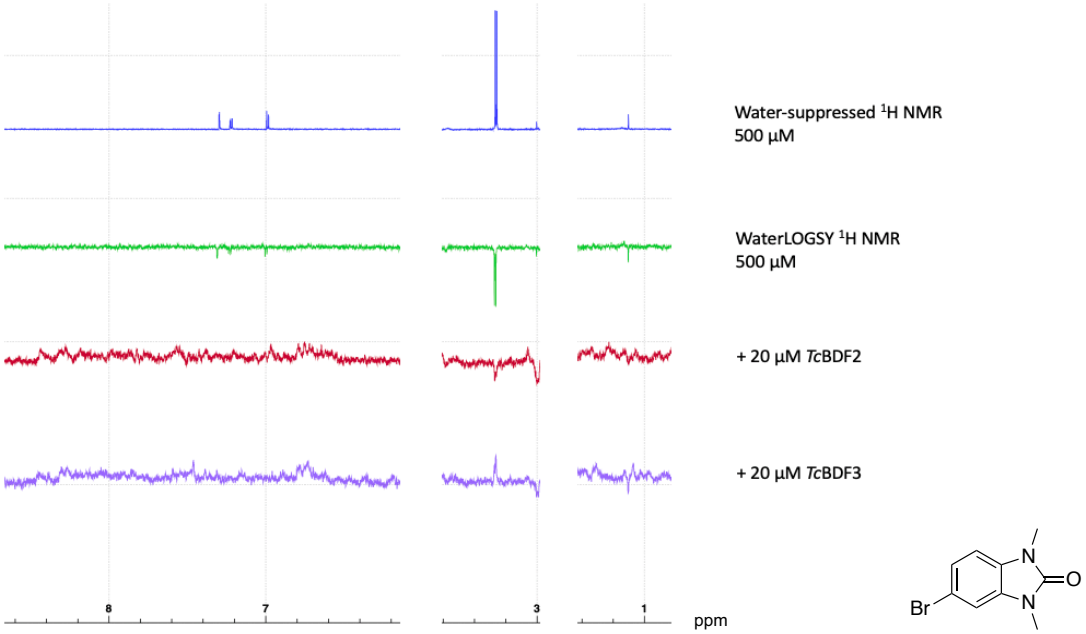


Figure B.5: Partial waterLOGSY spectra of **17**. The water suppressed ^{1}H spectrum is shown in blue, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. Disappearance of the ligand peaks on addition of the *Tc*BDF2 (red), shows potential weak binding to *Tc*BDF2. The waterLOGSY spectrum in the presence or 20 µM *Tc*BDF3 (purple) display partial inversion of the ligand peaks, indicating potential binding.

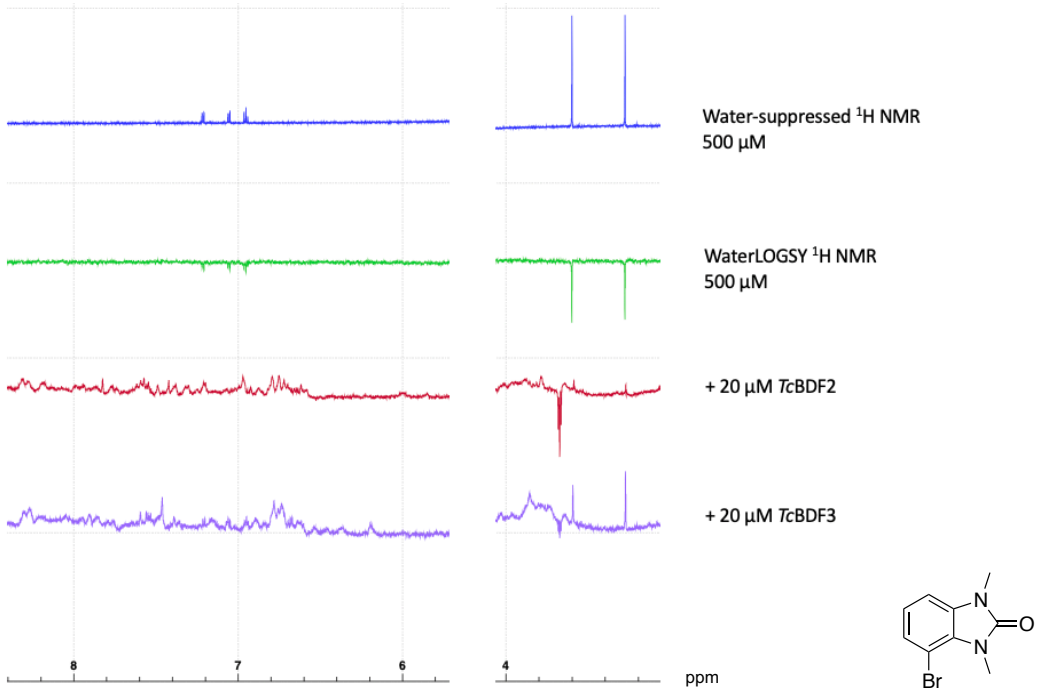


Figure B.6: Partial waterLOGSY spectra of **18**. The water suppressed ^{1}H spectrum is shown in blue, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. The waterLOGSY spectrum in the presence 20 µM of *Tc*BDF2 (red) or 20 µM *Tc*BDF3 (purple) display inversion of the ligand peaks, indicating binding to both proteins.

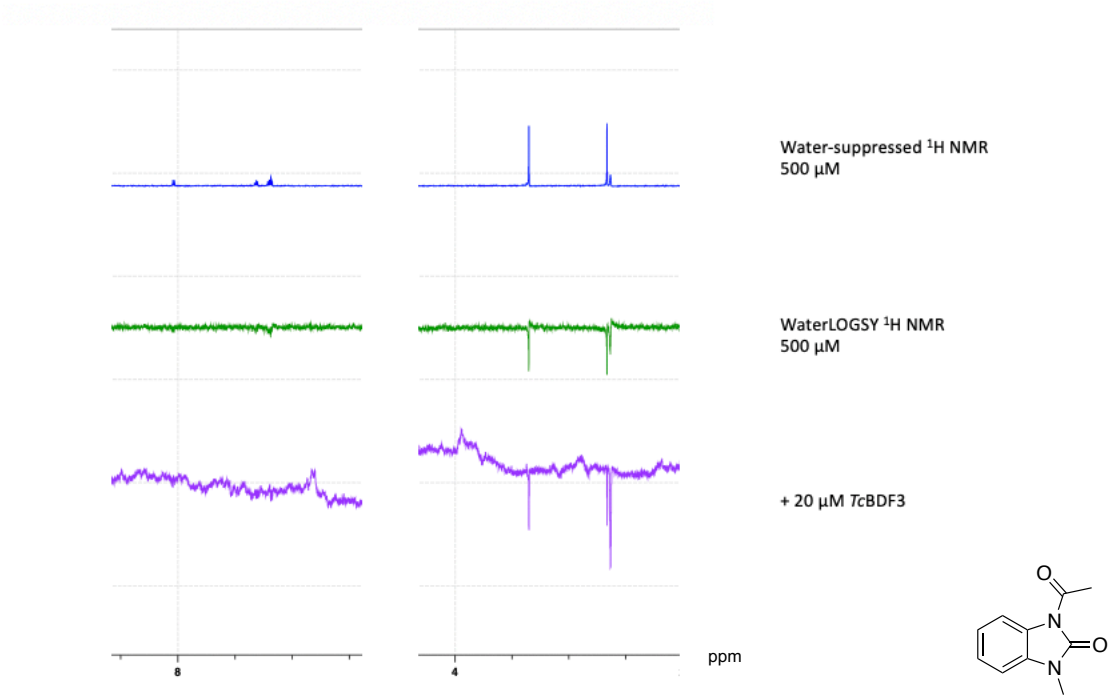


Figure B.7: Partial waterLOGSY spectra of **27**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

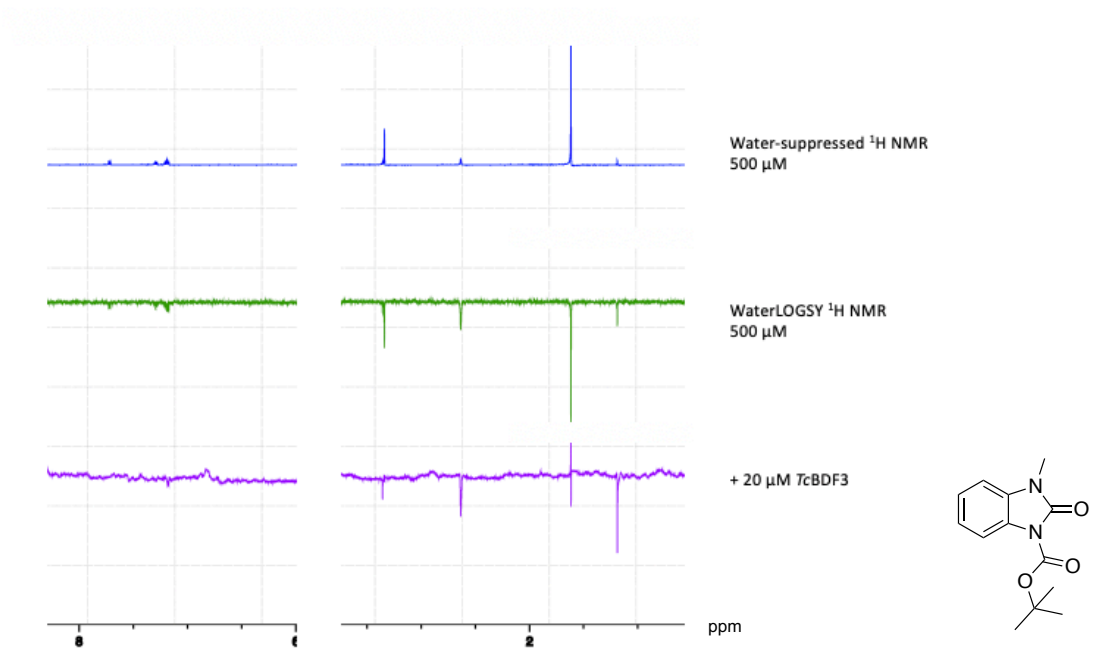


Figure B.8: Partial waterLOGSY spectra of **28**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

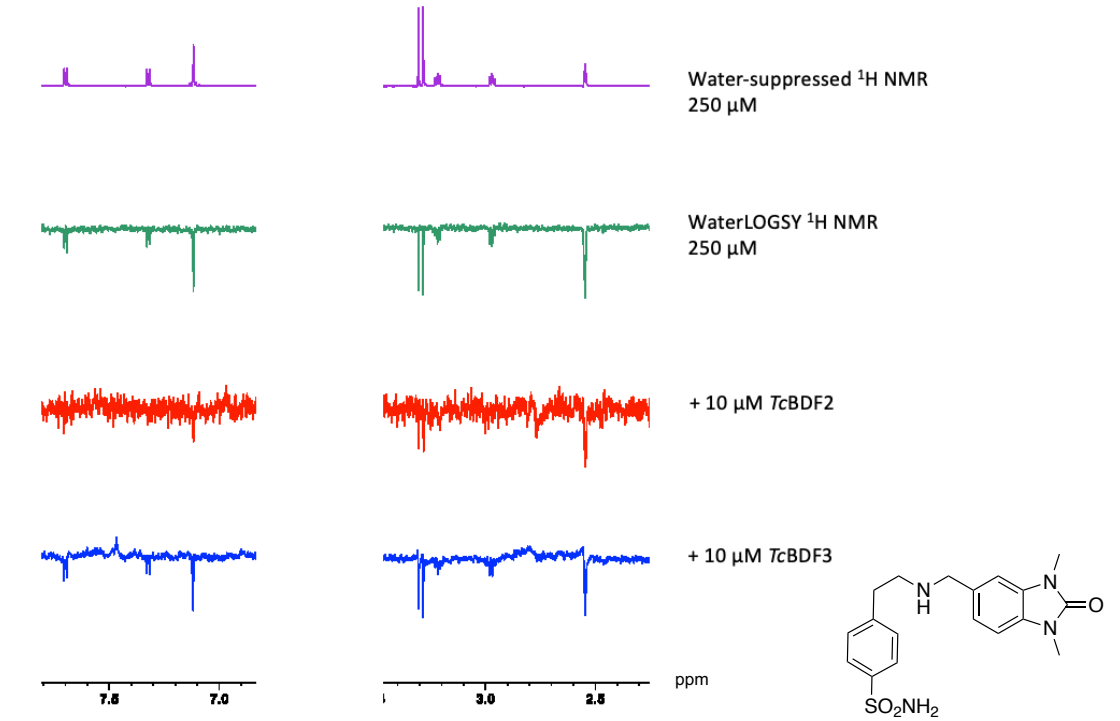


Figure B.9: Partial waterLOGSY spectra of **105**. The water suppressed ^{1}H spectrum is shown in purple, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. No inversion of the ligand peak is detected in the presence of 10 µM of TcBDF2 (red) or 10 µM of TcBDF3 (blue) indicating the ligand does not bind.

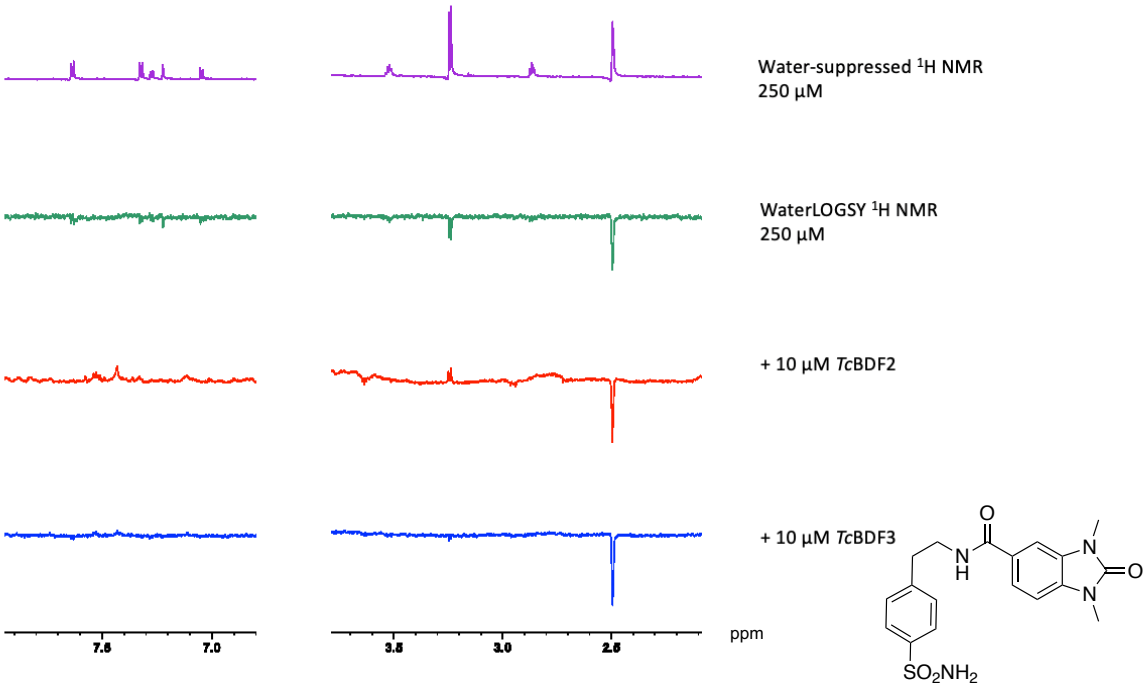


Figure B.10: Partial waterLOGSY spectra of **106**. The water suppressed ^{1}H spectrum is shown in purple, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. No inversion of the ligand peak is detected in the presence of 10 µM of TcBDF2 (red) or 10 µM of TcBDF3 (blue) indicating the ligand does not bind. **106** was synthesised by Amy Scorah.

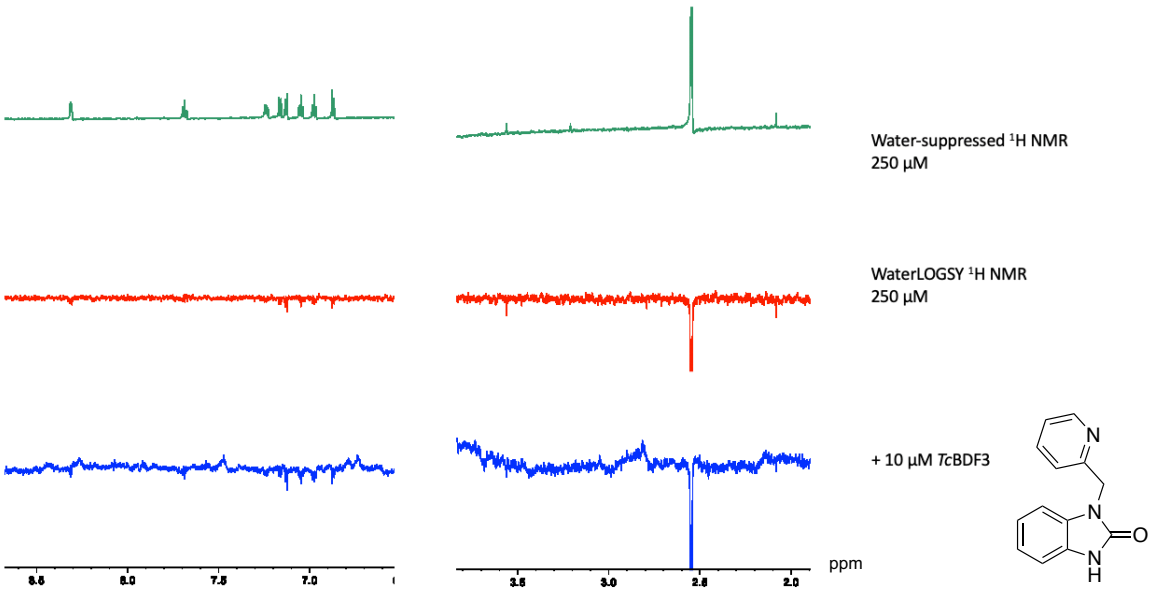


Figure B.11: Partial waterLOGSY spectra of **34**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind.

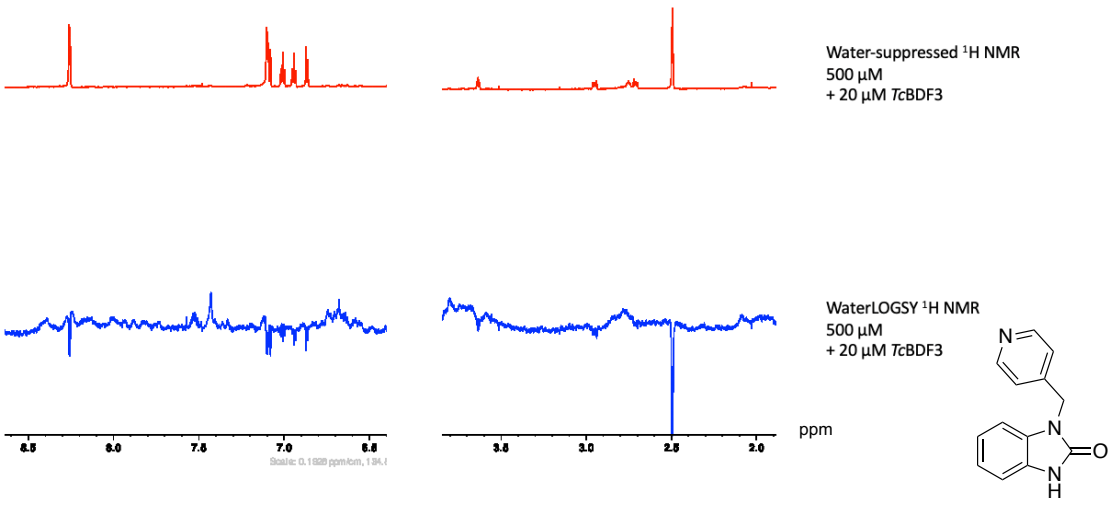


Figure B.12: Partial waterLOGSY spectra of **36**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

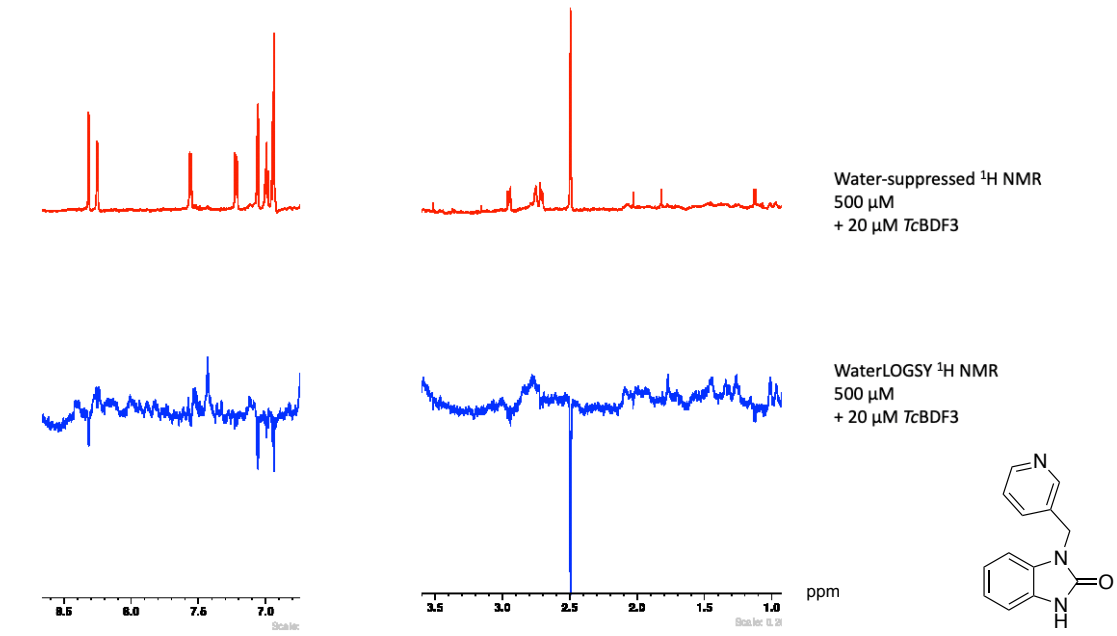


Figure B.13: Partial waterLOGSY spectra of **35**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

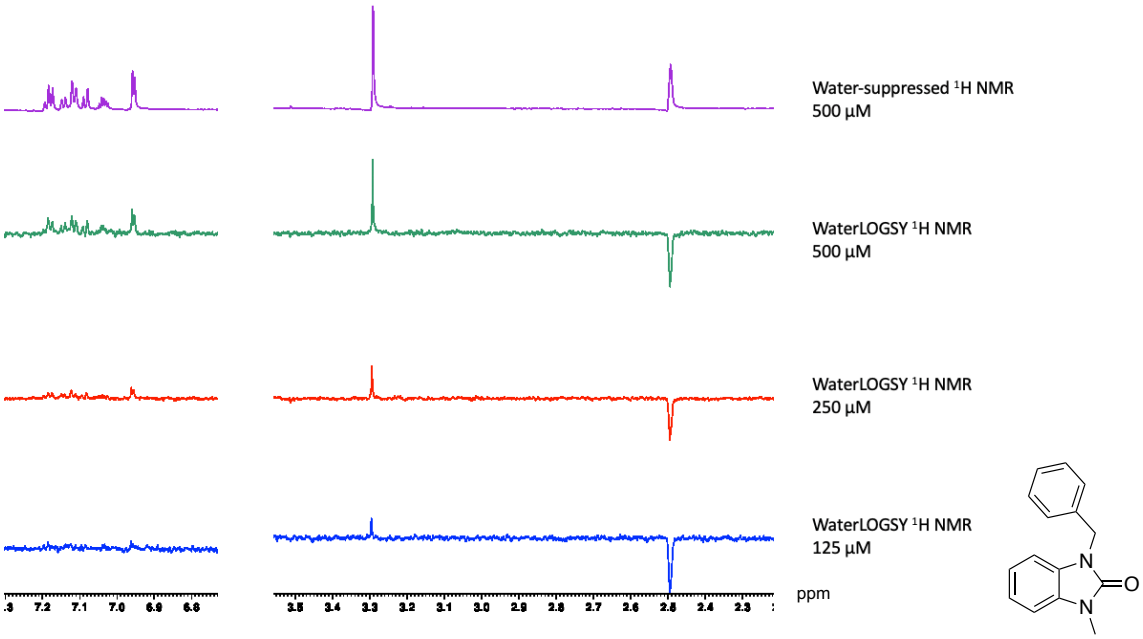


Figure B.14: Partial waterLOGSY spectra of **37**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 500 µM (green), 250 µM (red) and 125 µM (blue). The water-suppressed ^{1}H NMR spectrum (purple) is given as reference.

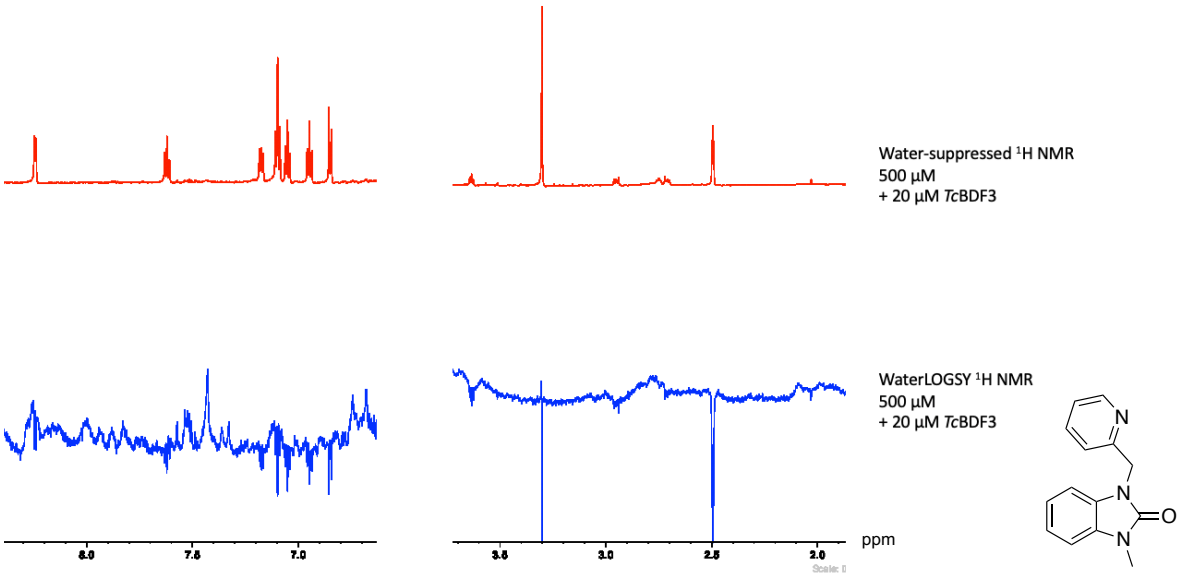


Figure B.15: Partial waterLOGSY spectra of **38**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

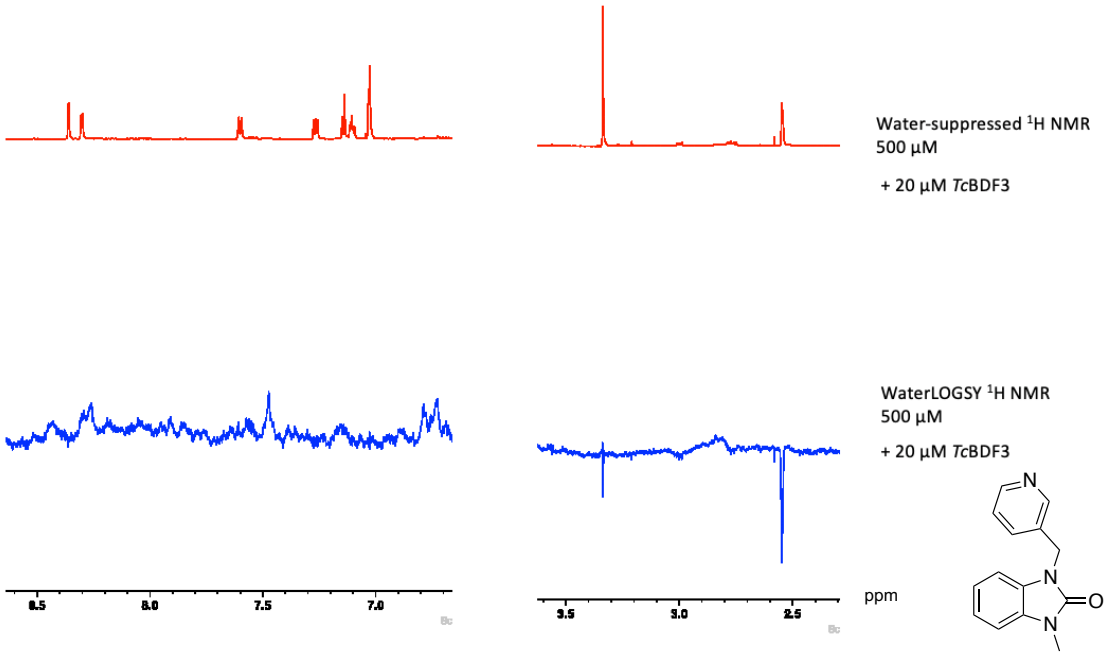


Figure B.16: Partial waterLOGSY spectra of **39**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

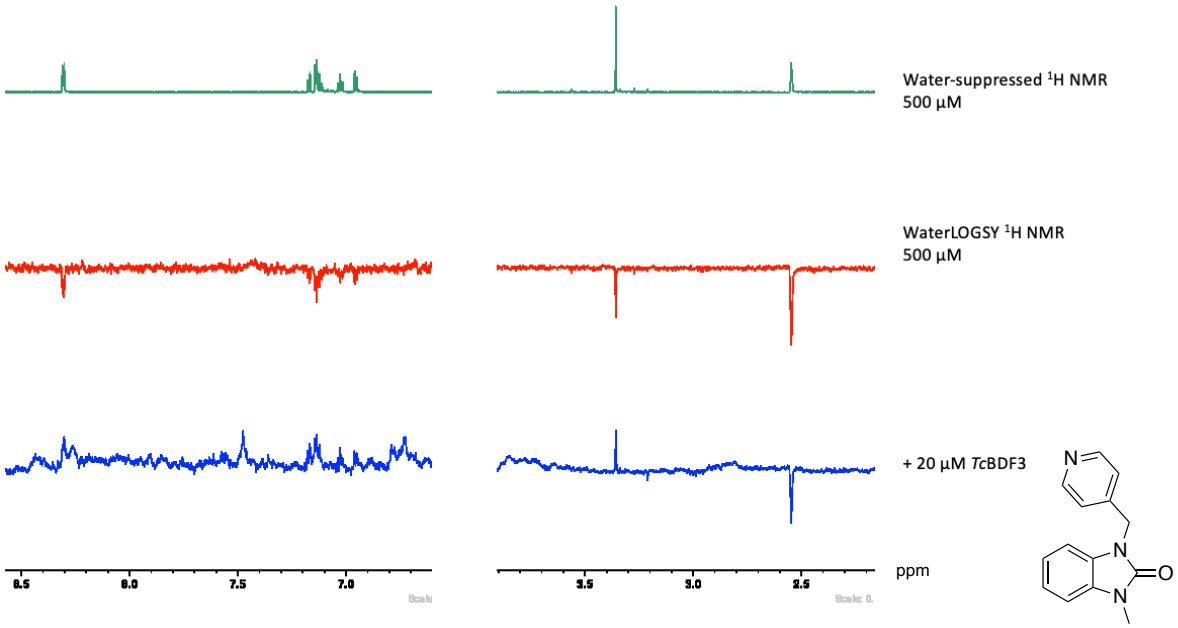


Figure B.17: Partial waterLOGSY spectra of **40**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

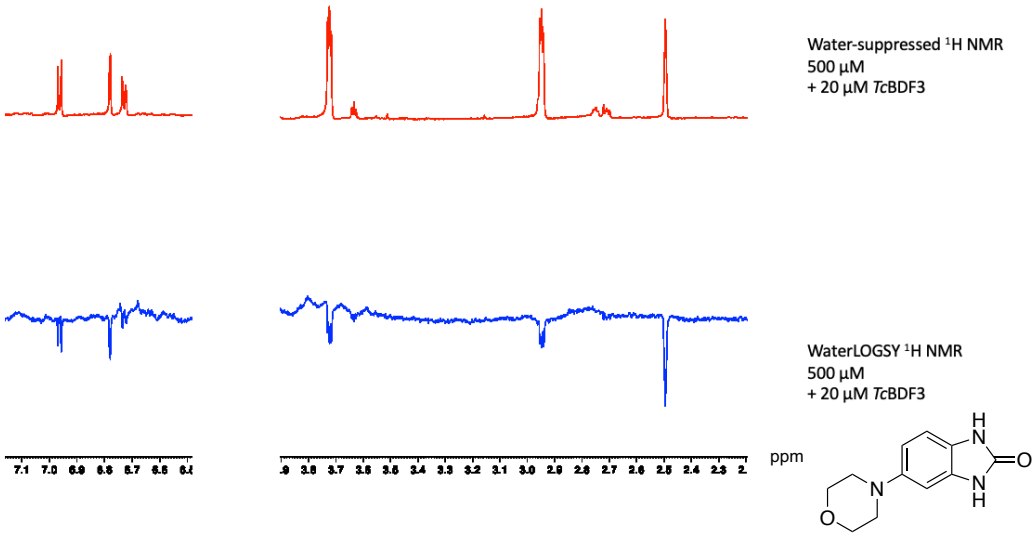


Figure B.18: Partial waterLOGSY spectra of **47**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

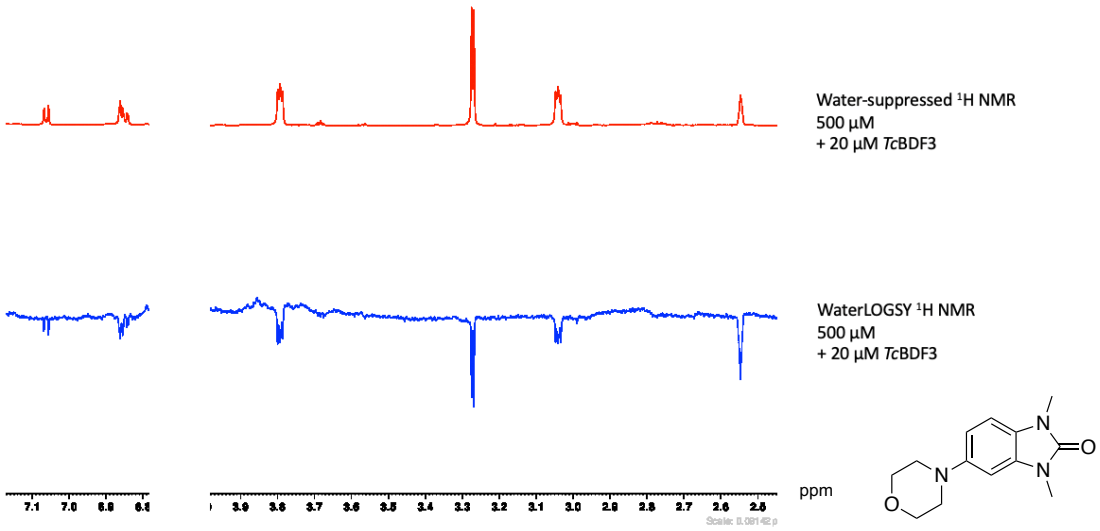


Figure B.19: Partial waterLOGSY spectra of **44**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

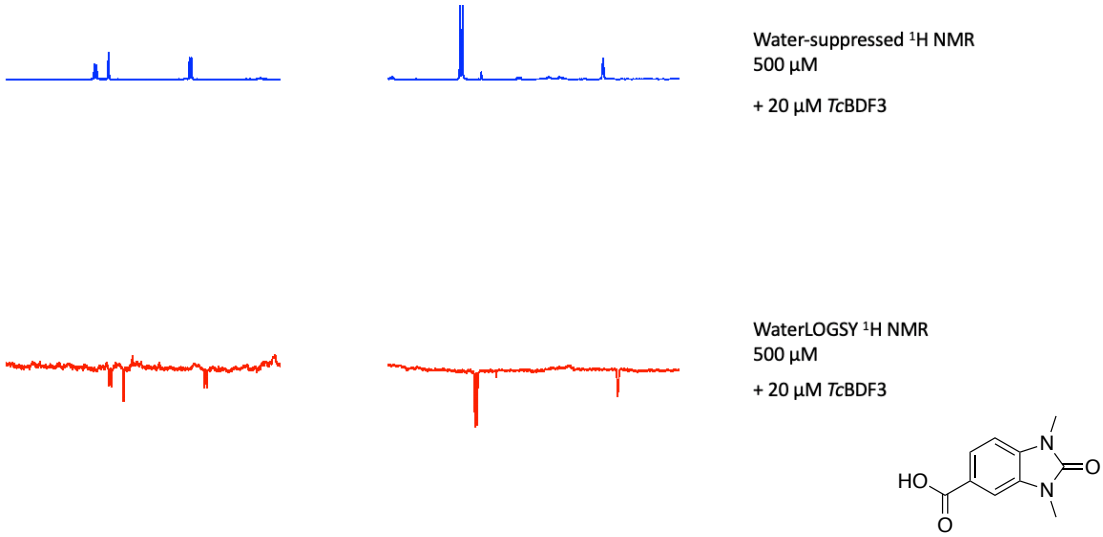


Figure B.20: Partial waterLOGSY spectra of **102**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference. **102** was synthesised by Amy Scorah.

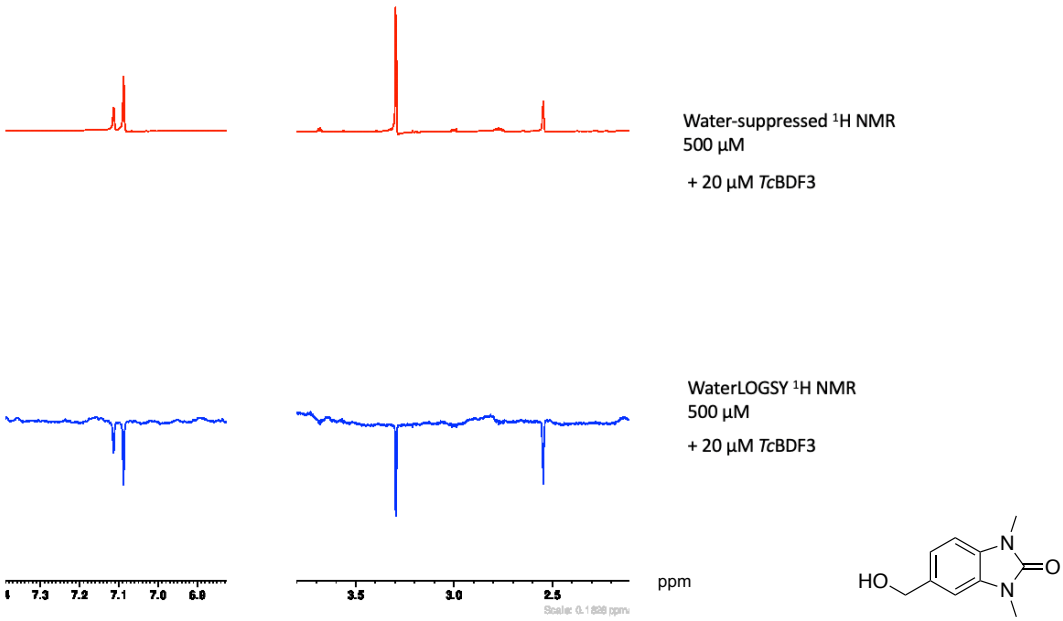


Figure B.21: Partial waterLOGSY spectra of **103**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference. **103** was synthesised by Amy Scorah.

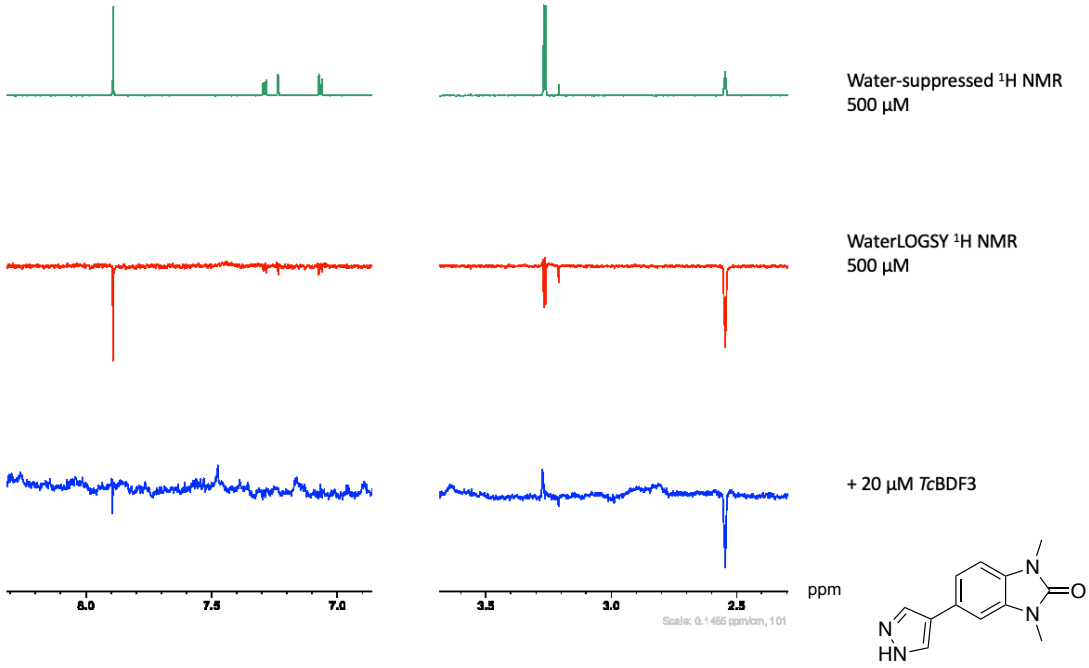


Figure B.22: Partial waterLOGSY spectra of **100**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. Partial inversion of the ligand aliphatic peaks is detected in the presence of 20 µM of TcBDF3 (blue), however, the aromatics peaks are not inverted. This ambiguous result might indicate binding. **100** was synthesised by Amy Scorah.

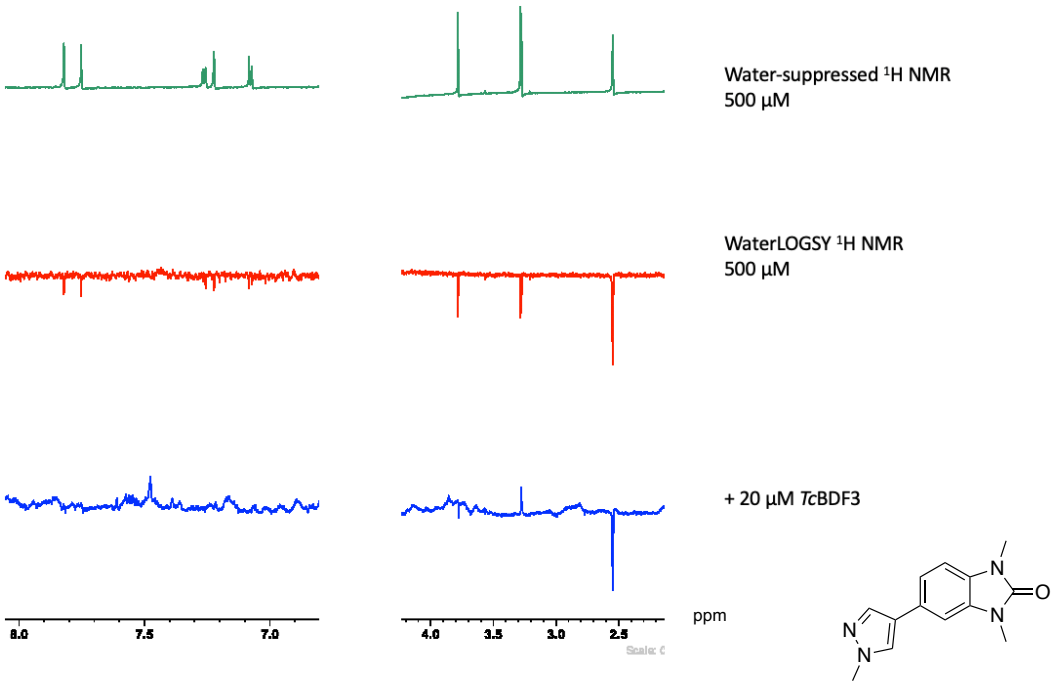


Figure B.23: Partial waterLOGSY spectra of **101**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. Partial inversion of the ligand aliphatic peaks is detected in the presence of 20 µM of TcBDF3 (blue), however, the aromatics peaks are not inverted. This ambiguous result might indicate binding. **101** was synthesised by Amy Scorah (University of Oxford).

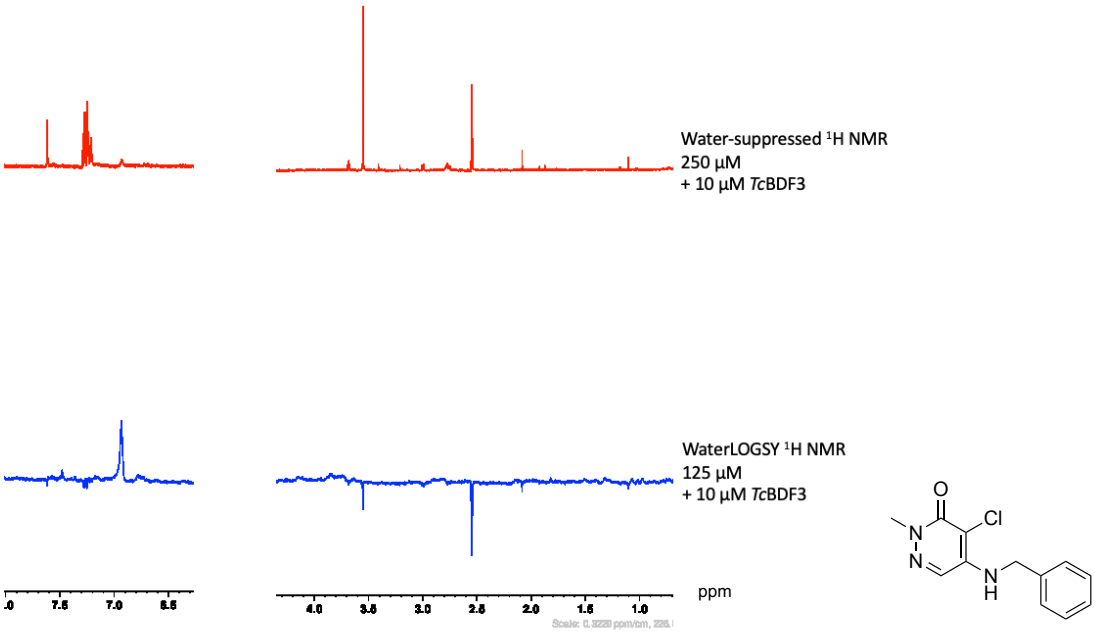


Figure B.24: Partial waterLOGSY spectra of **67**. No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 10 µM protein (red) is given as a reference.

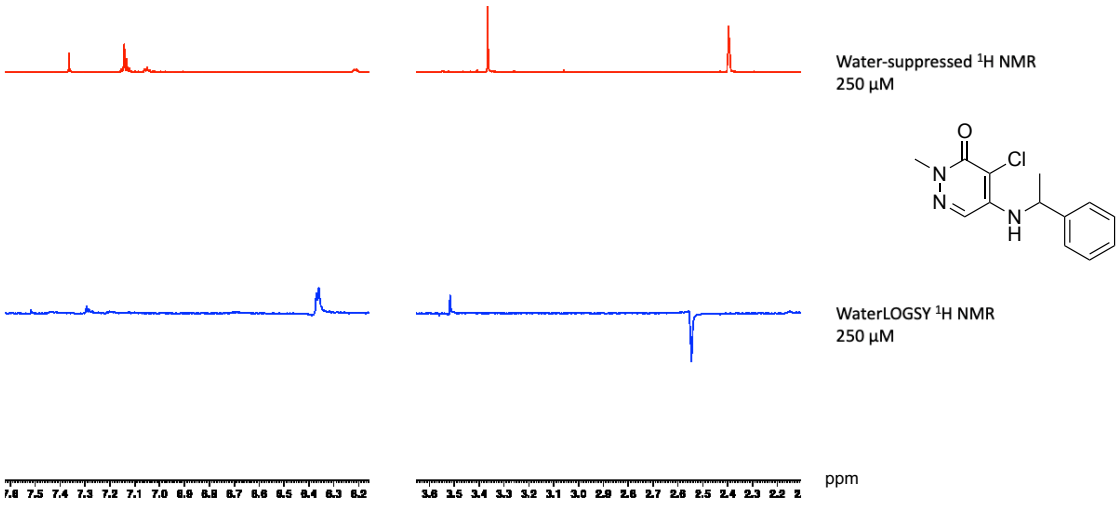


Figure B.25: Partial waterLOGSY spectra of **68**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 250 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference.

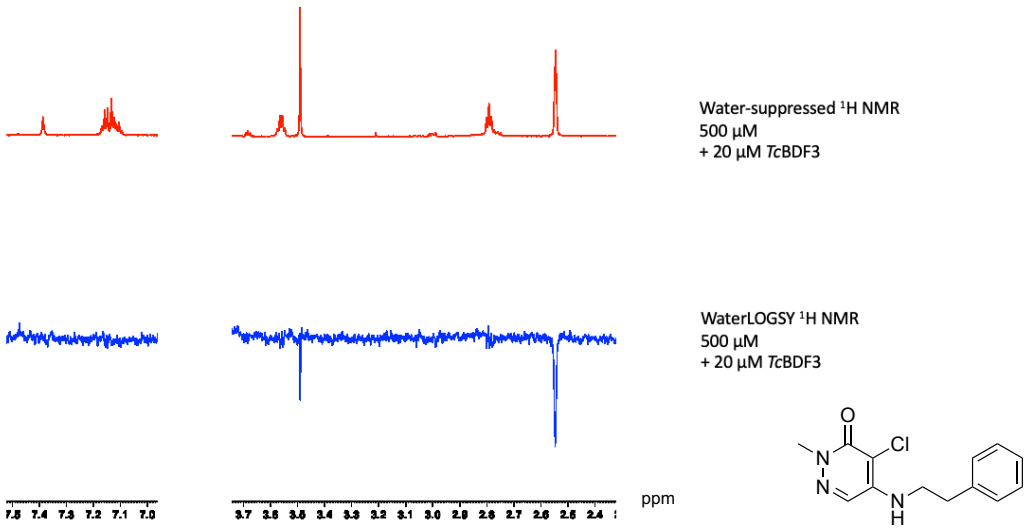


Figure B.26: Partial waterLOGSY spectra of **69**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

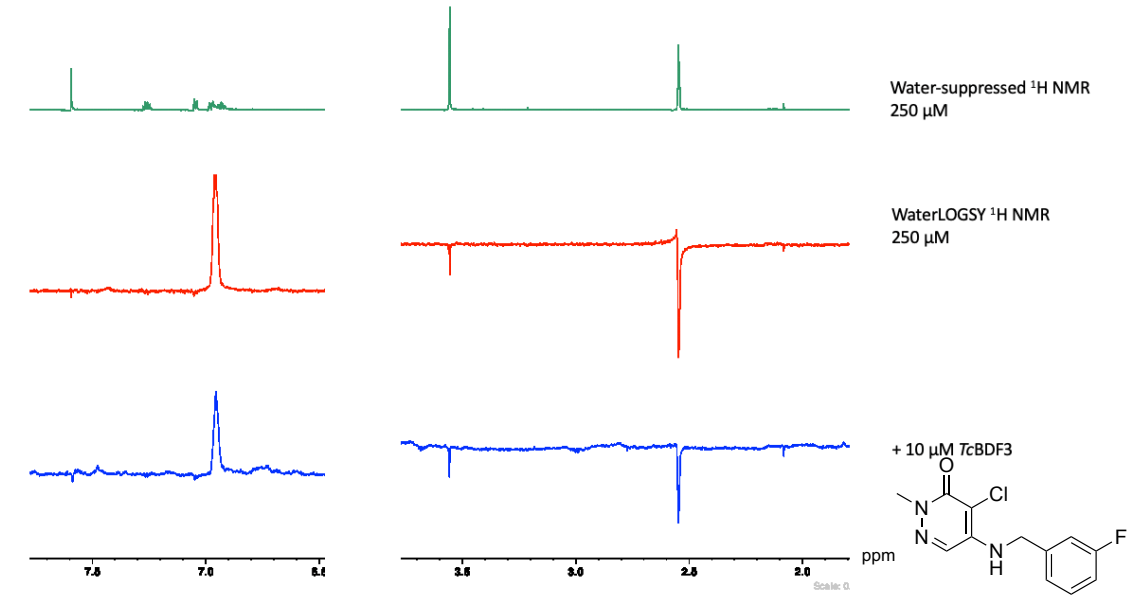


Figure B.27: Partial waterLOGSY spectra of 72. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. No clear inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind. The large positive signal in the 2 waterLOGSY spectra corresponds to an exchangeable proton.

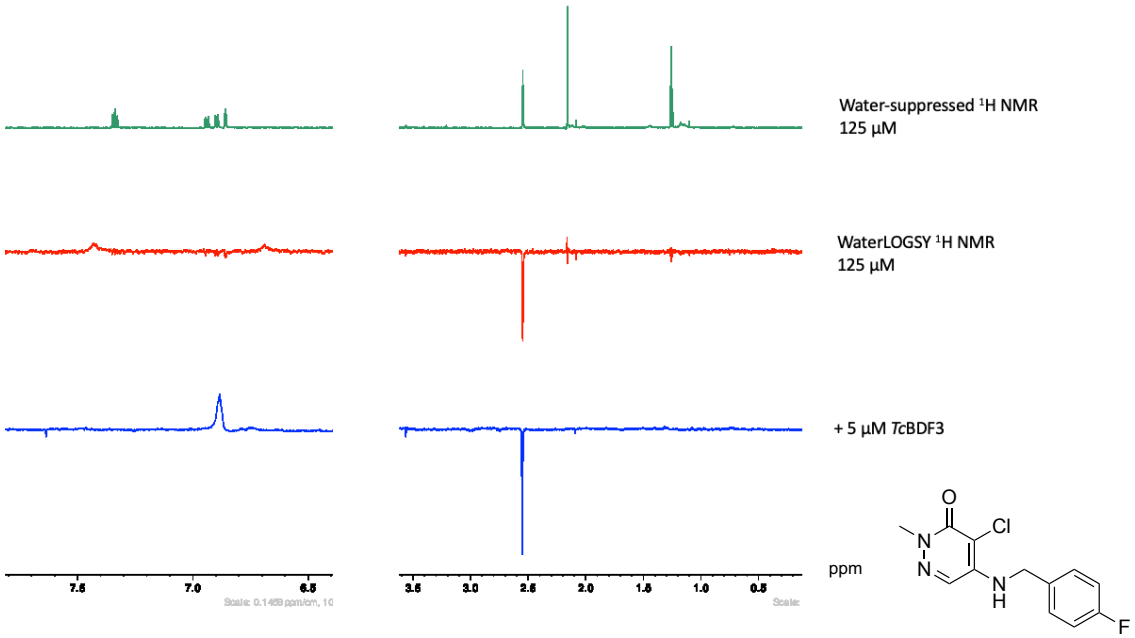


Figure B.28: Partial waterLOGSY spectra of 73. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

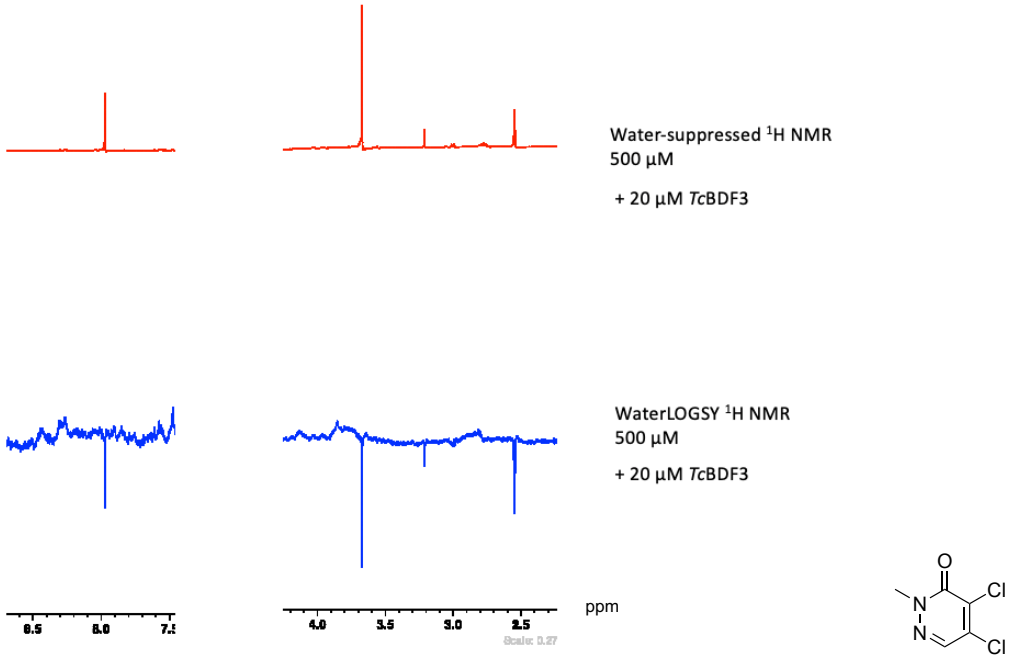


Figure B.29: Partial waterLOGSY spectra of **74a**. No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 10 µM protein (red) is given as a reference.

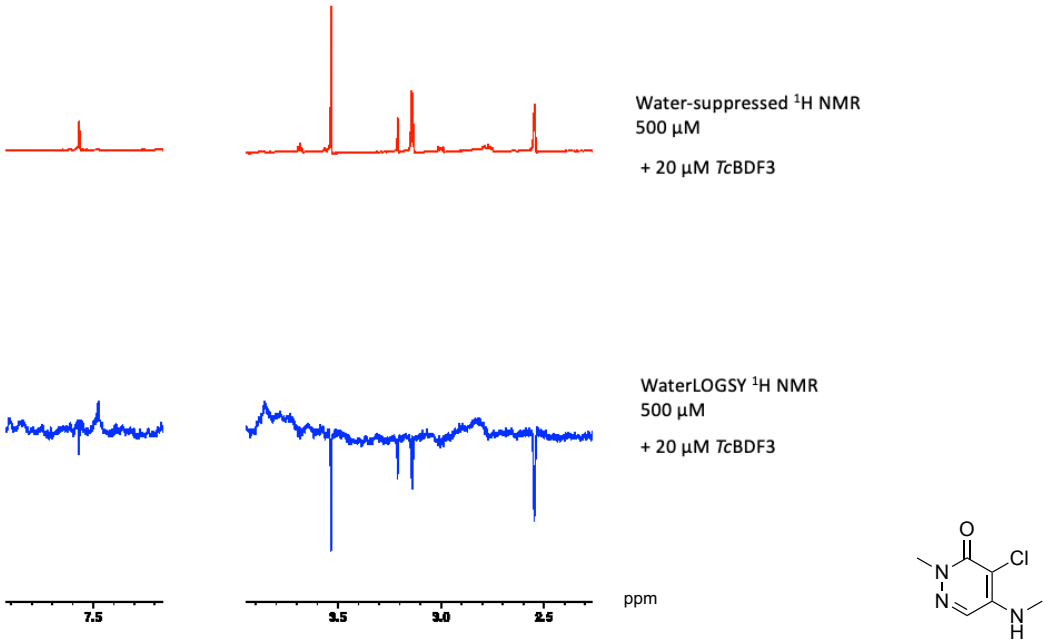


Figure B.30: Partial waterLOGSY spectra of **74**. No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 10 µM protein (red) is given as a reference.

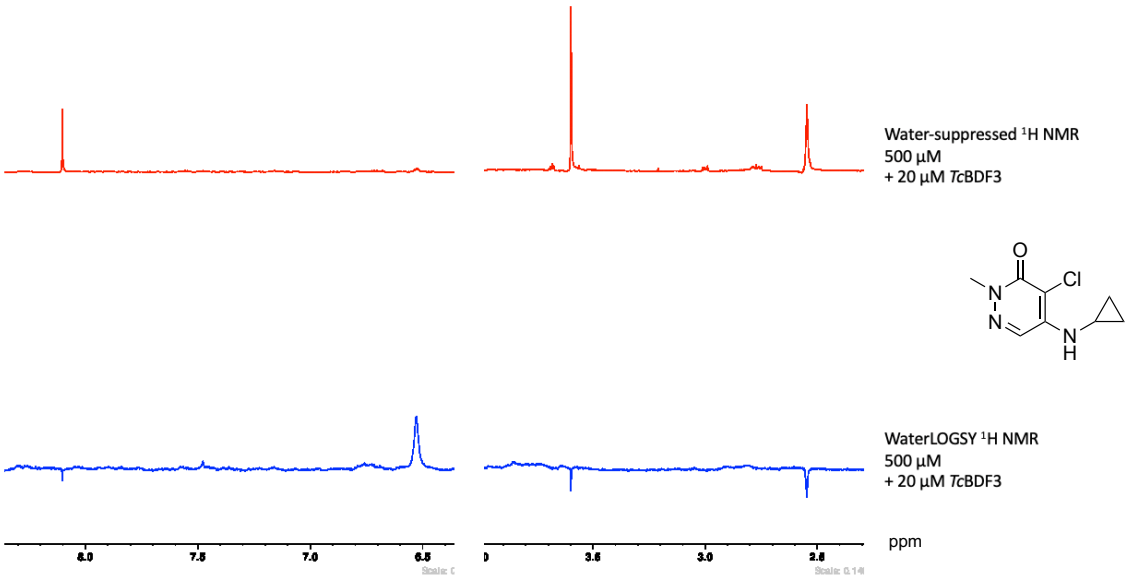


Figure B.31: Partial waterLOGSY spectra of **75**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

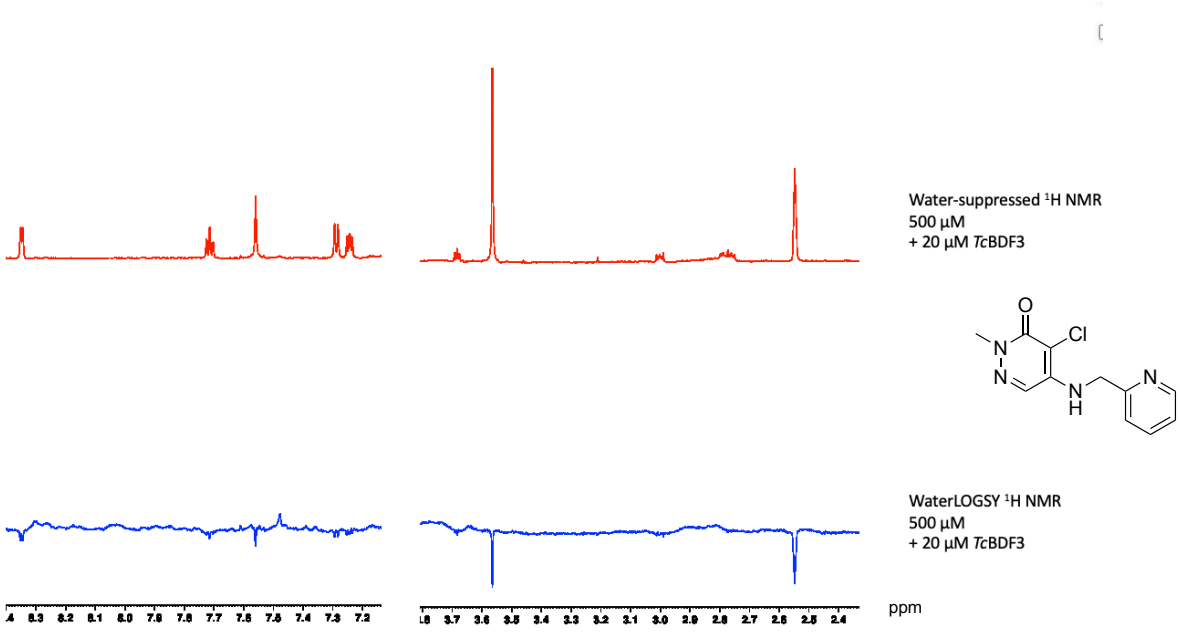


Figure B.32: Partial waterLOGSY spectra of **76**. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

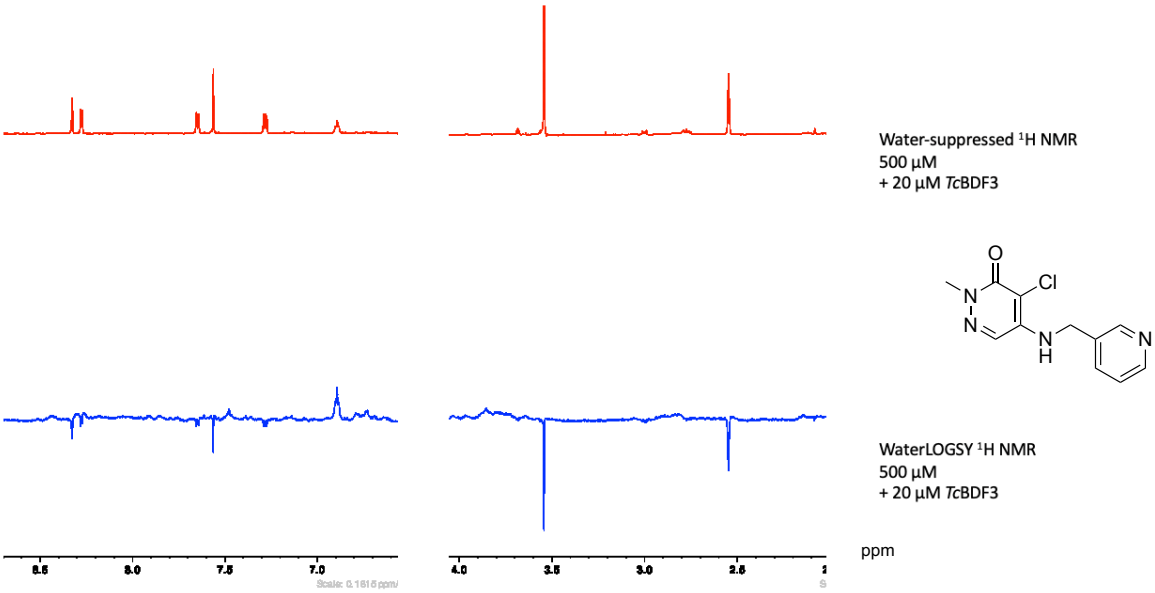


Figure B.33: Partial waterLOGSY spectra of 77. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

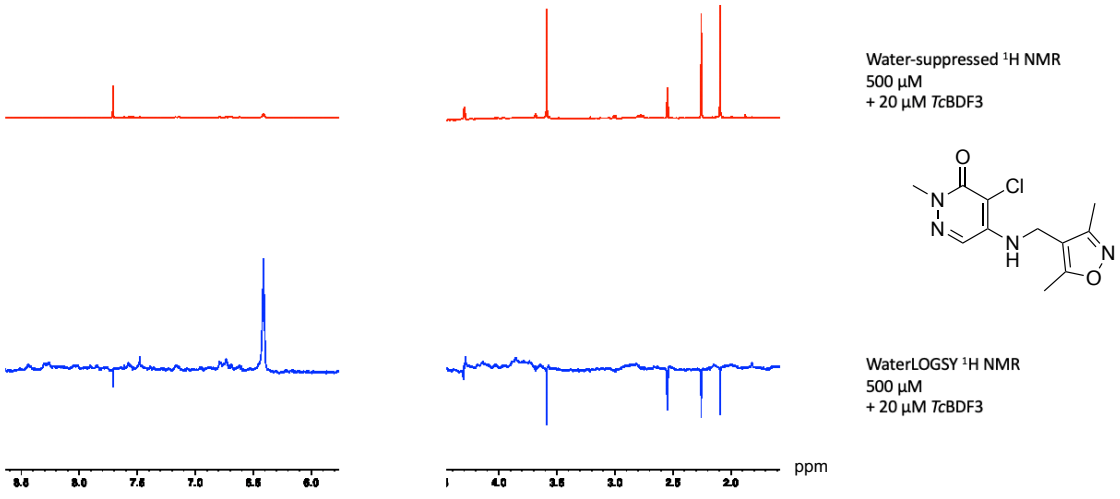


Figure B.34: Partial waterLOGSY spectra of 79. No clear inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference.

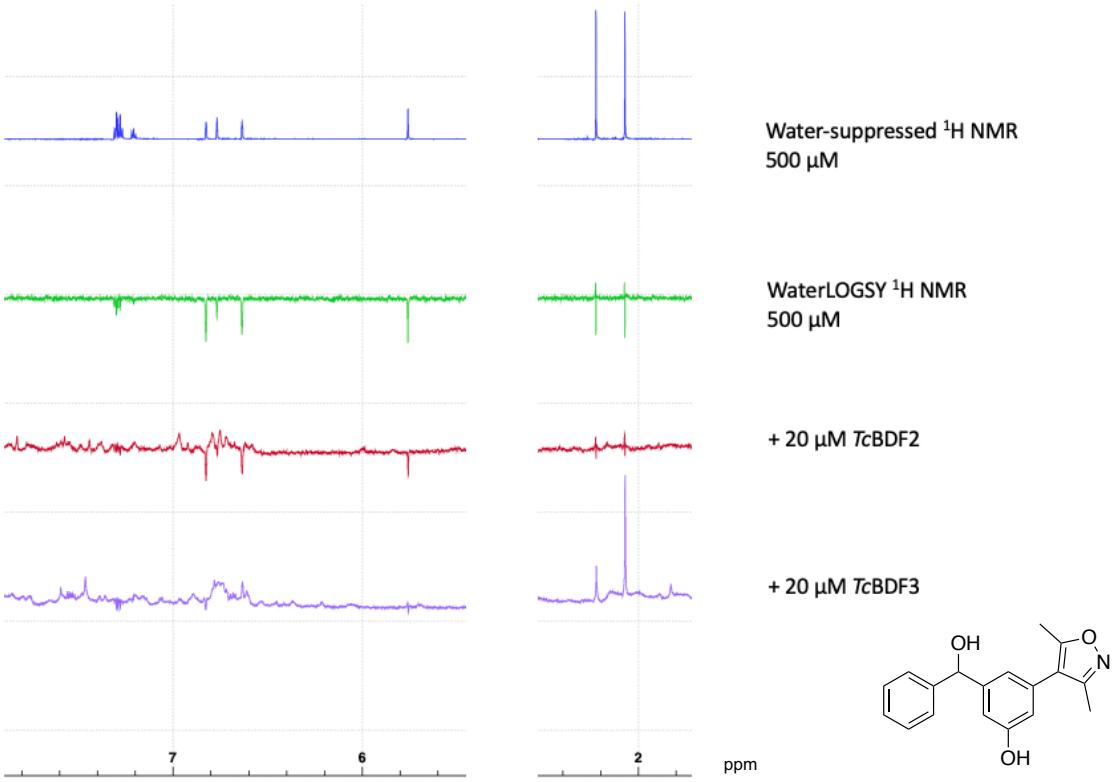


Figure B.35: Partial waterLOGSY spectra of OXFBD02 (**7**). The water suppressed ^{1}H spectrum is shown in blue, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. The waterLOGSY spectrum in the presence of 20 µM of *Tc*BDF2 (red) or 20 µM *Tc*BDF3 (purple) display inversion of the ligand peaks, indicating binding to both proteins.

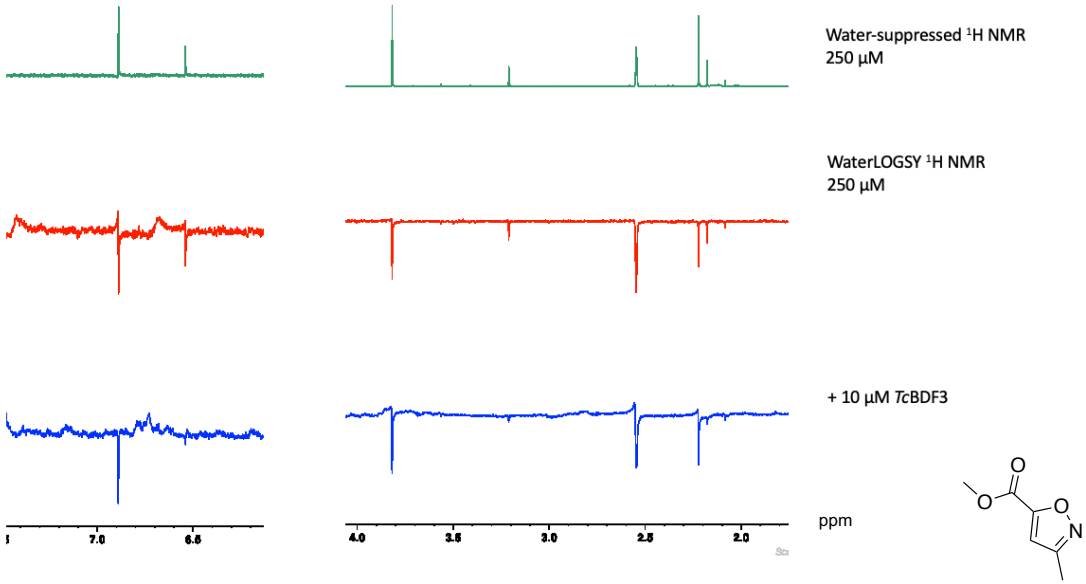


Figure B.36: Partial waterLOGSY spectra of **120**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. **120** was synthesised by Dr David Hewings (University of Oxford).

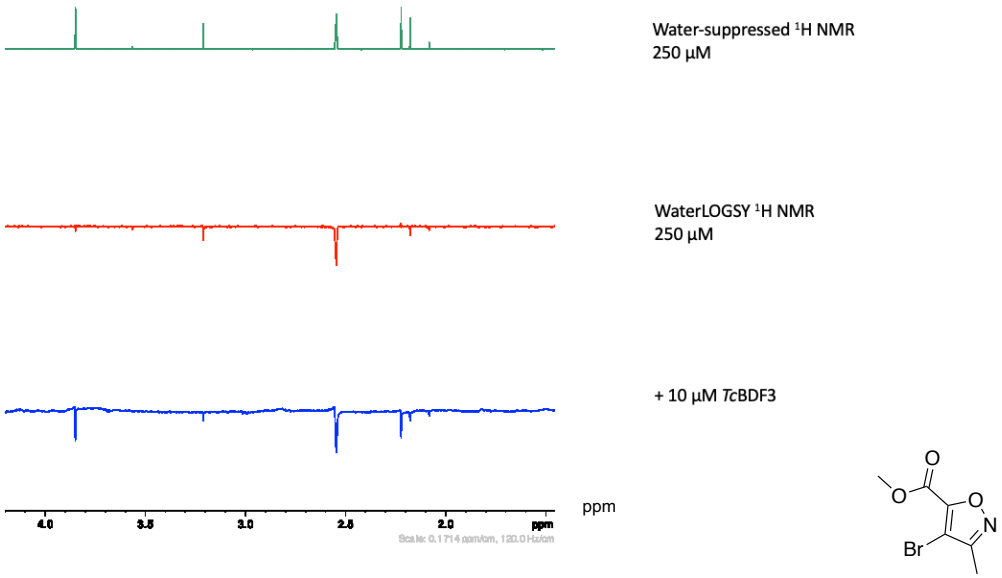


Figure B.37: Partial waterLOGSY spectra of **121**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. **121** was synthesised by Dr David Hewings (University of Oxford).

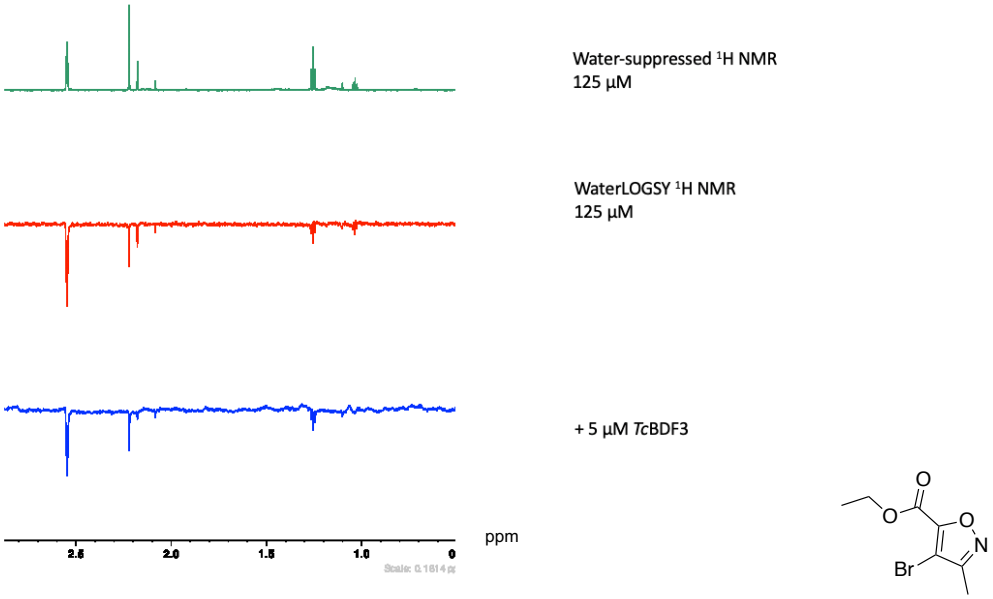


Figure B.38: Partial waterLOGSY spectra of **122**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No clear inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind. **122** was synthesised by Dr David Hewings (University of Oxford).

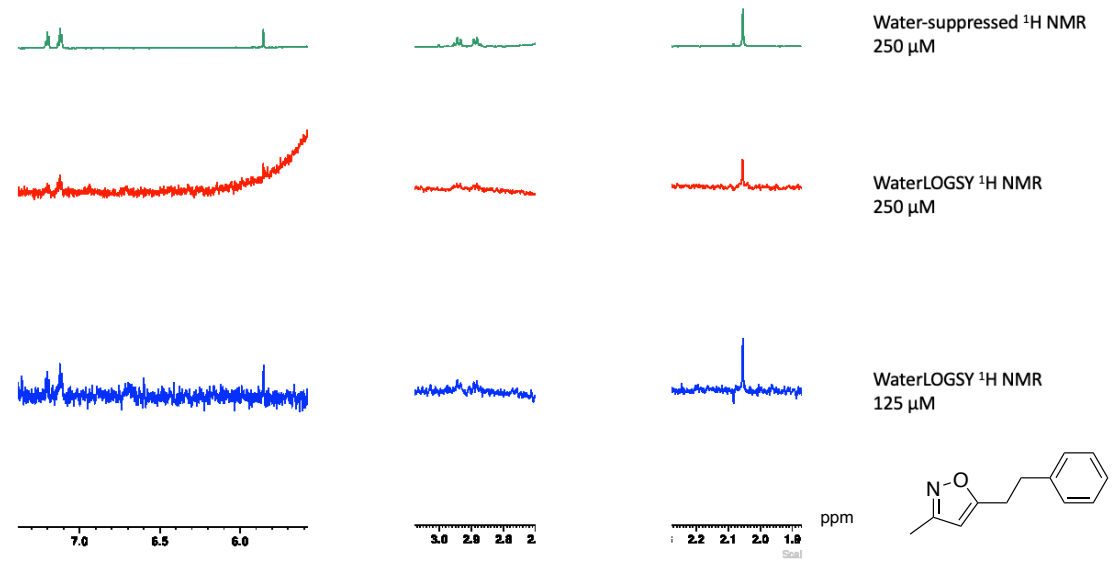


Figure B.39: Partial waterLOGSY spectra of 208. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 250 µM (red) and 125 µM (blue). The water-suppressed ^{1}H NMR spectrum (green) is given as reference. 208 was synthesised by Dr David Hewings (University of Oxford).

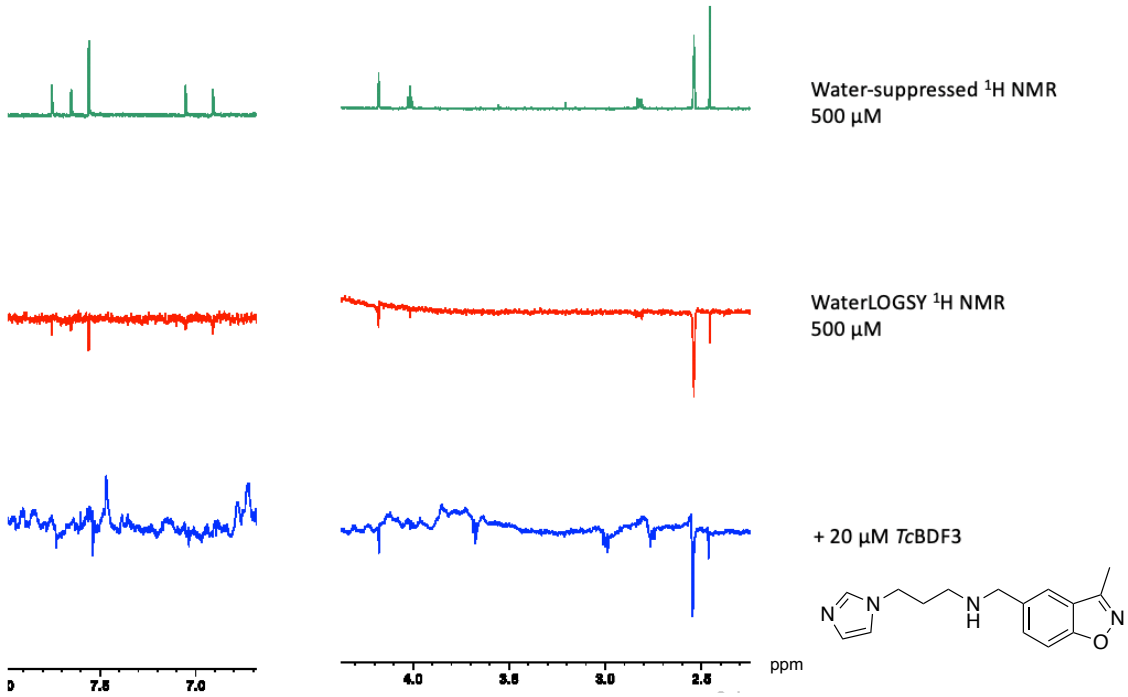


Figure B.40: Partial waterLOGSY spectra of 209. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. No inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. 209 was synthesised by Amy Scorah (University of Oxford).

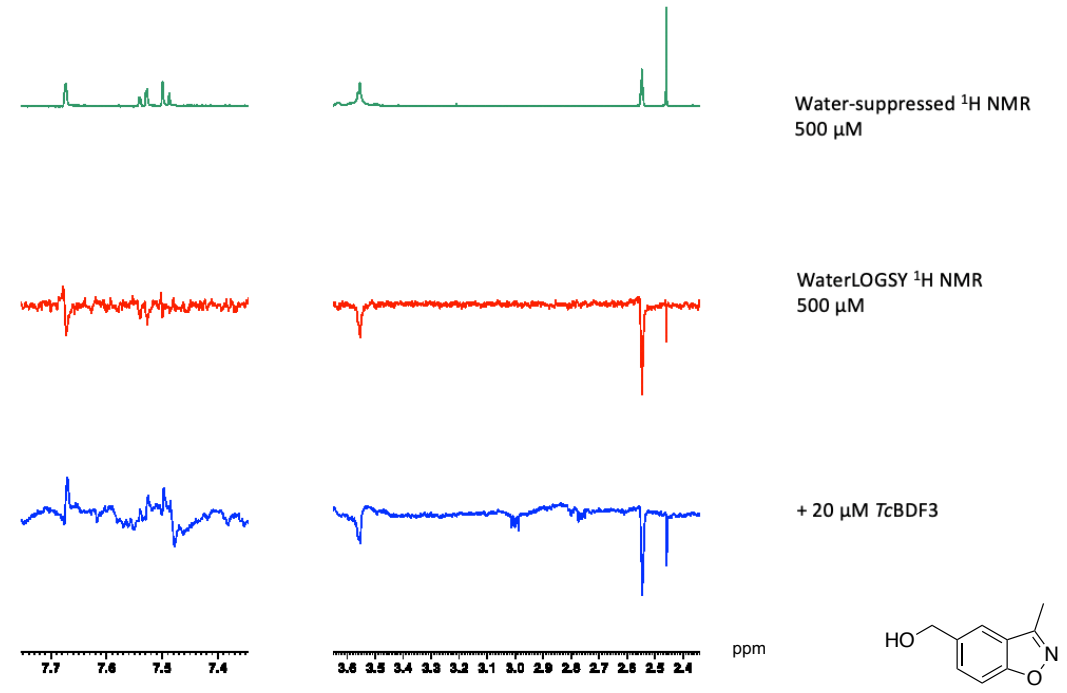


Figure B.41: Partial waterLOGSY spectra of **210**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. No inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. **210** was synthesised by Amy Scorah (University of Oxford).

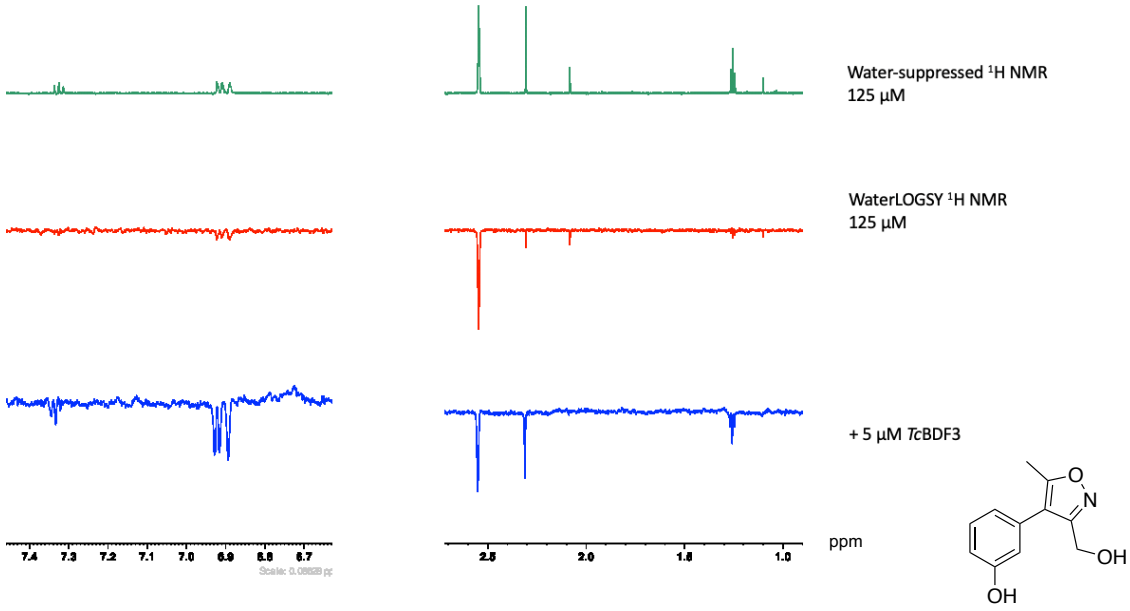


Figure B.42: Partial waterLOGSY spectra of **114**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No clear inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind. **114** was synthesised by Dr David Hewings (University of Oxford).

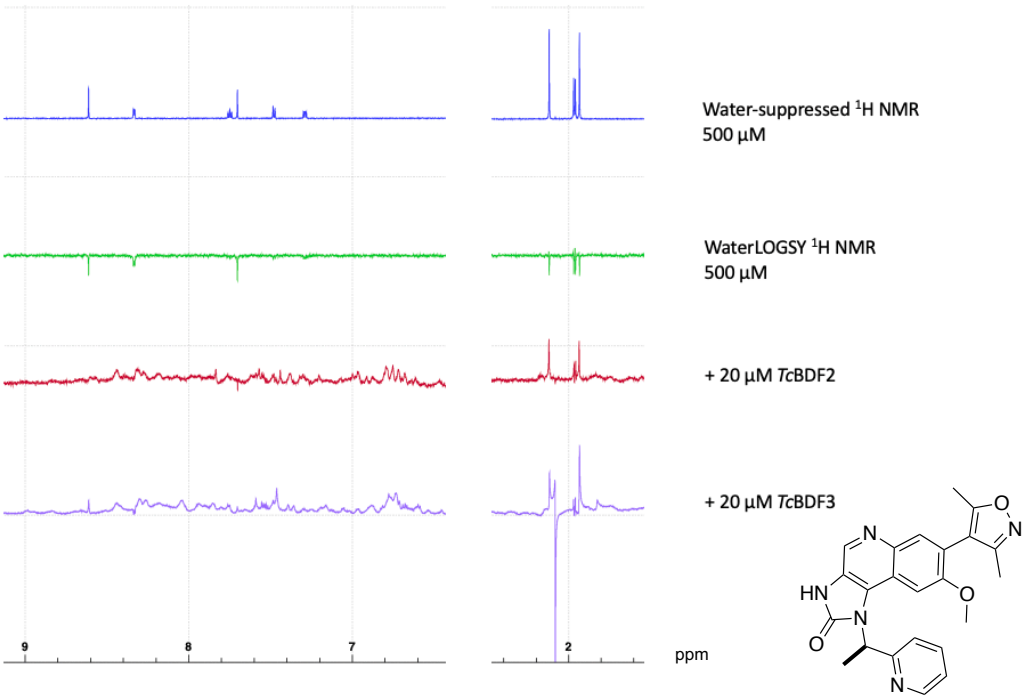


Figure B.43: Partial waterLOGSY spectra of I-BET151 (**5**). The water suppressed ^{1}H spectrum is shown in blue, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. The waterLOGSY spectrum in the presence of 20 µM of *Tc*BDF2 (red) or 5 µM of *Tc*BDF3 (purple) display inversion of the ligand peaks, indicating binding to both proteins.

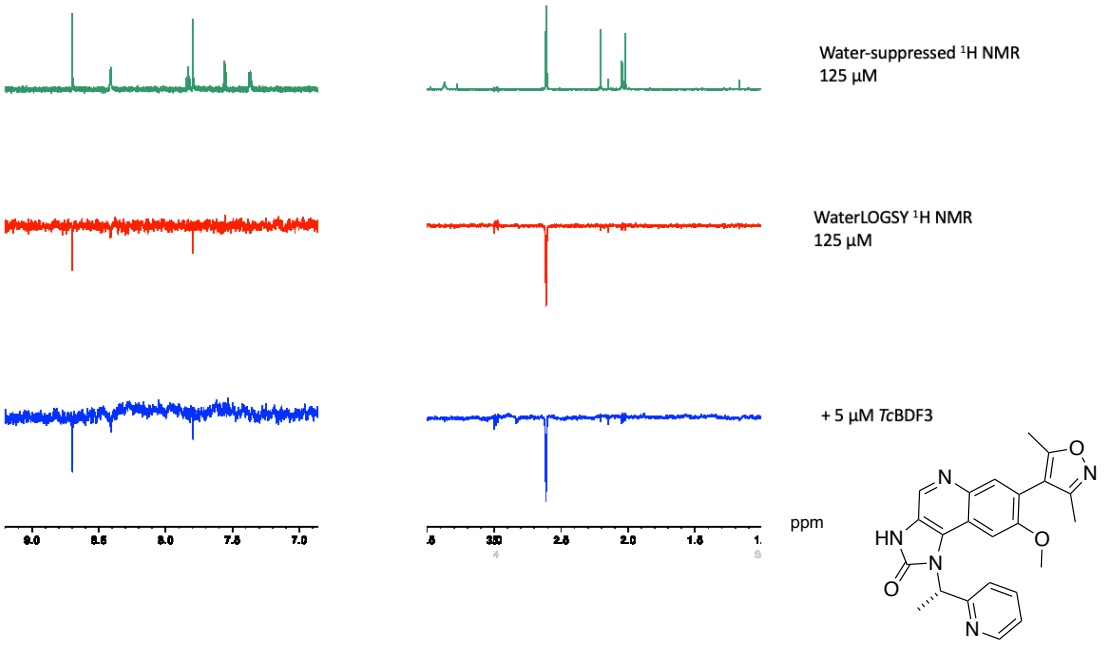


Figure B.44: Partial waterLOGSY spectra of **115**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No clear inversion of the ligand peak is detected in the presence of 5 µM of *Tc*BDF3 (blue) indicating the ligand does not bind.

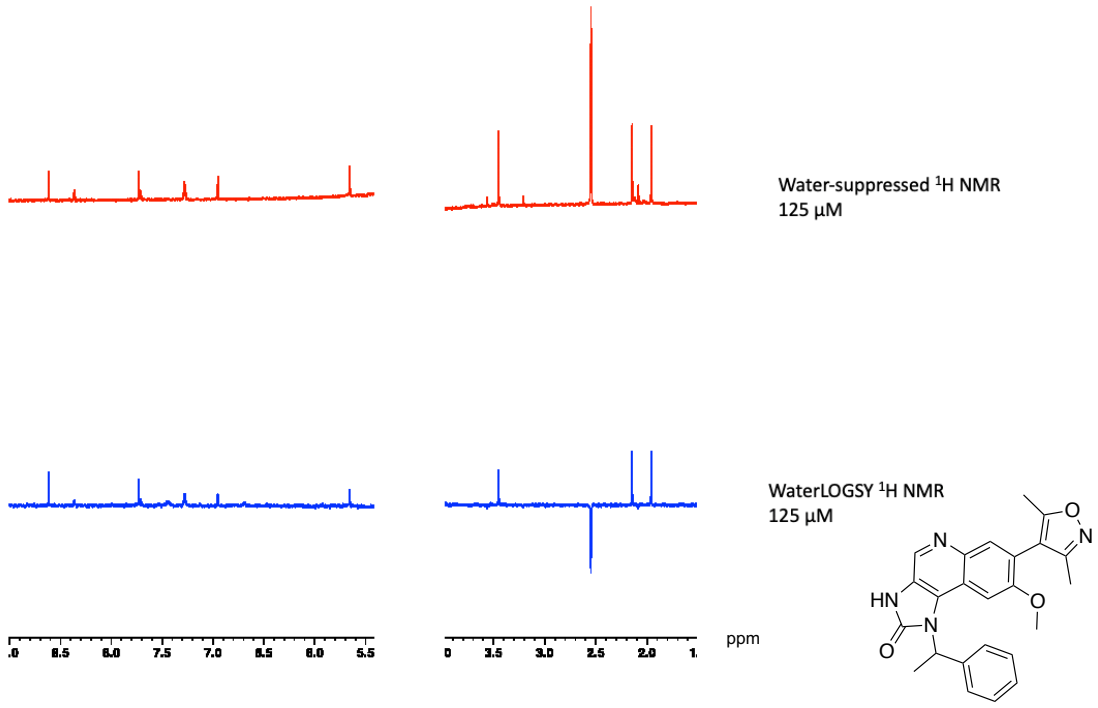


Figure B.45: Partial waterLOGSY spectra of **116**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 125 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference.

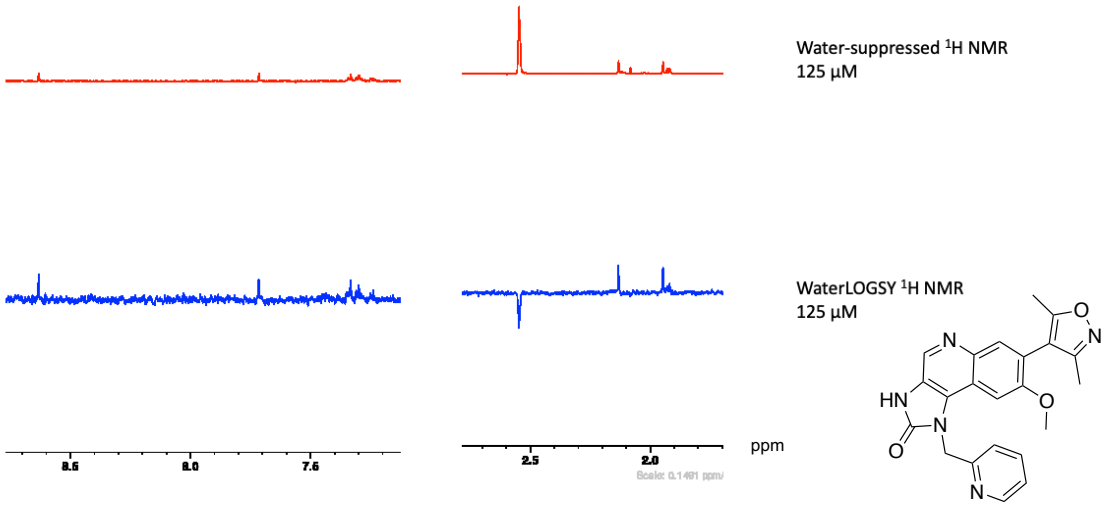


Figure B.46: Partial waterLOGSY spectra of **117**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 125 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference.

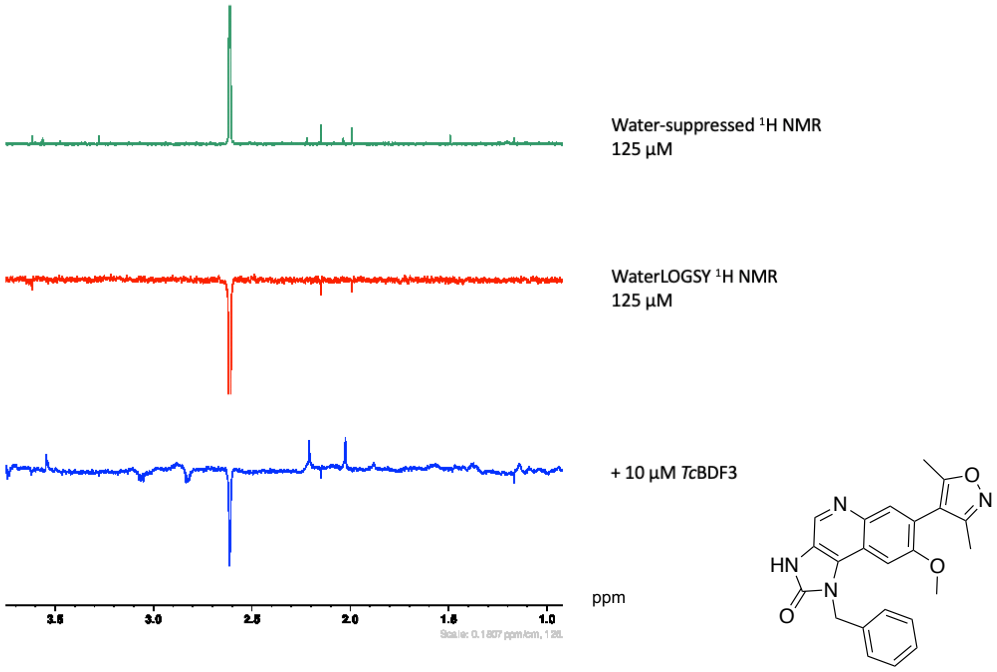


Figure B.47: Partial waterLOGSY spectra of **118**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. The waterLOGSY spectrum in the presence of 10 µM of *Tc*BDF3 (blue) display inversion of the ligand peaks, indicating binding.

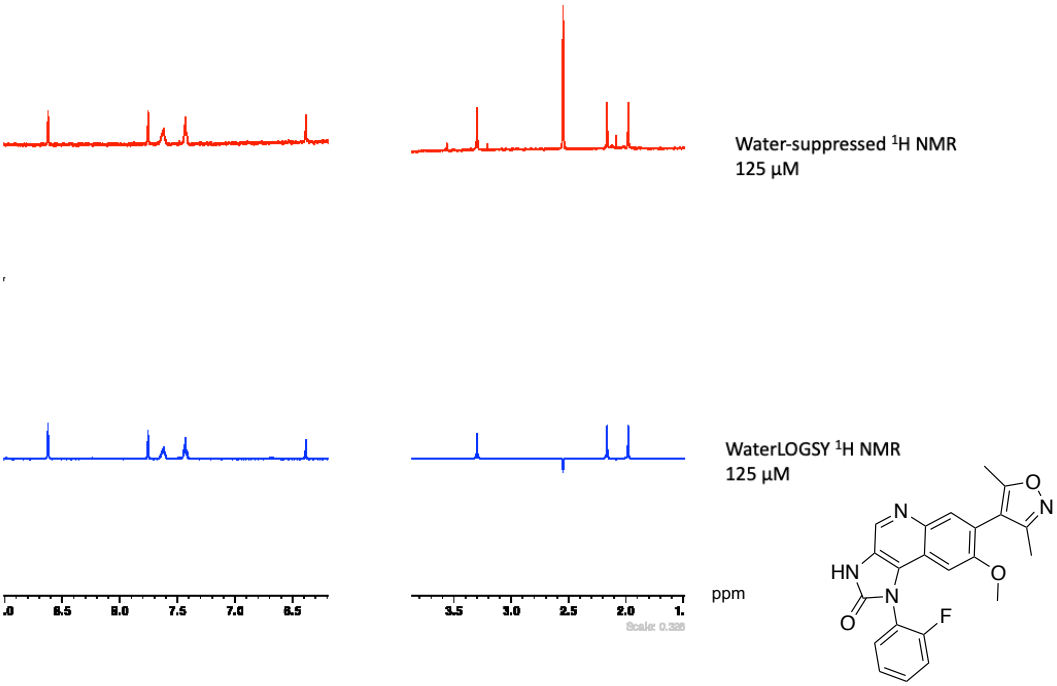


Figure B.48: Partial waterLOGSY spectra of **119**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 125 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference.

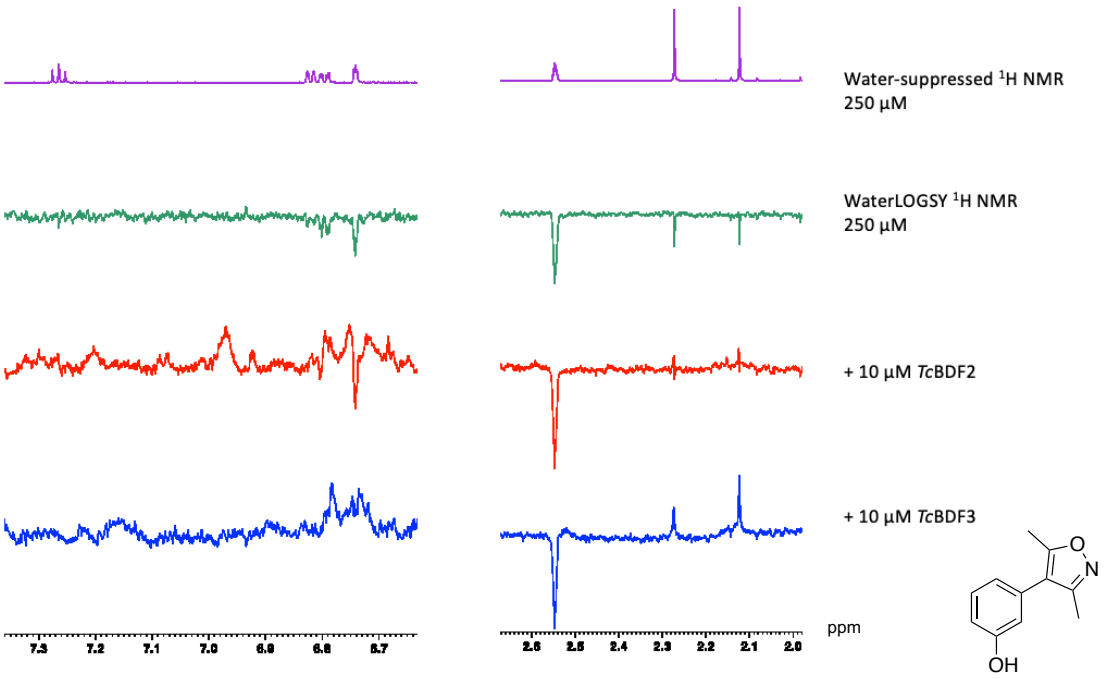


Figure B.49: Partial waterLOGSY spectra of **58**. The water suppressed ^{1}H spectrum is shown in purple, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. The waterLOGSY spectrum in the presence of 10 µM of *Tc*BDF2 (red) or 10 µM of *Tc*BDF3 (blue) display inversion of the ligand peaks, indicating binding to both proteins.

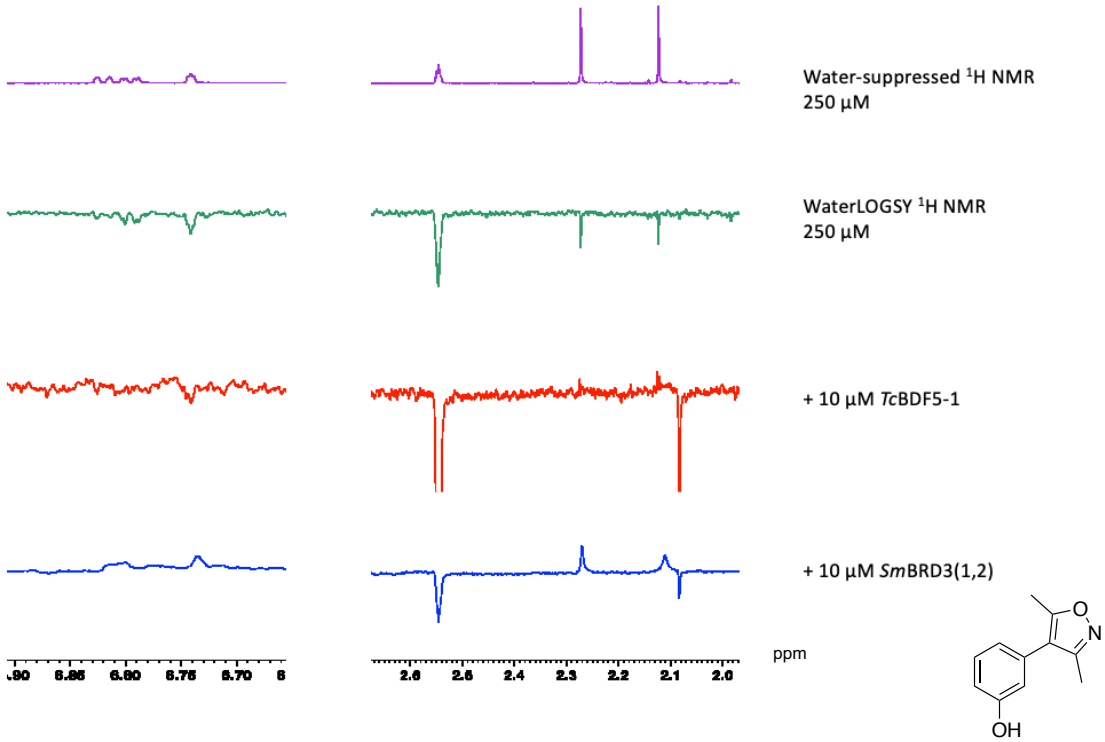


Figure B.50: Partial waterLOGSY spectra of **58**. The water suppressed ^{1}H spectrum is shown in purple, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. The waterLOGSY spectrum in the presence of 10 µM of *Tc*BDF5-1 (red) or 10 µM of *Sm*BRD3(1,2) (blue) display inversion of the ligand peaks, indicating binding to both proteins.

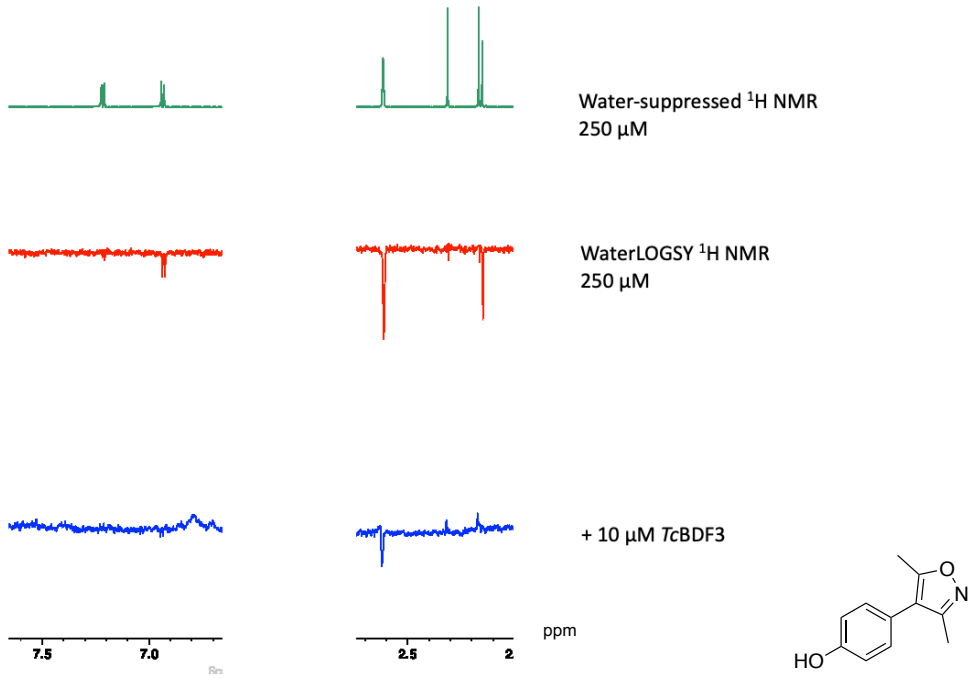


Figure B.51: Partial waterLOGSY spectra of **59**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. The waterLOGSY spectrum in the presence of 10 µM of *Tc*BDF3 (blue) display inversion of the ligand peaks, indicating binding.

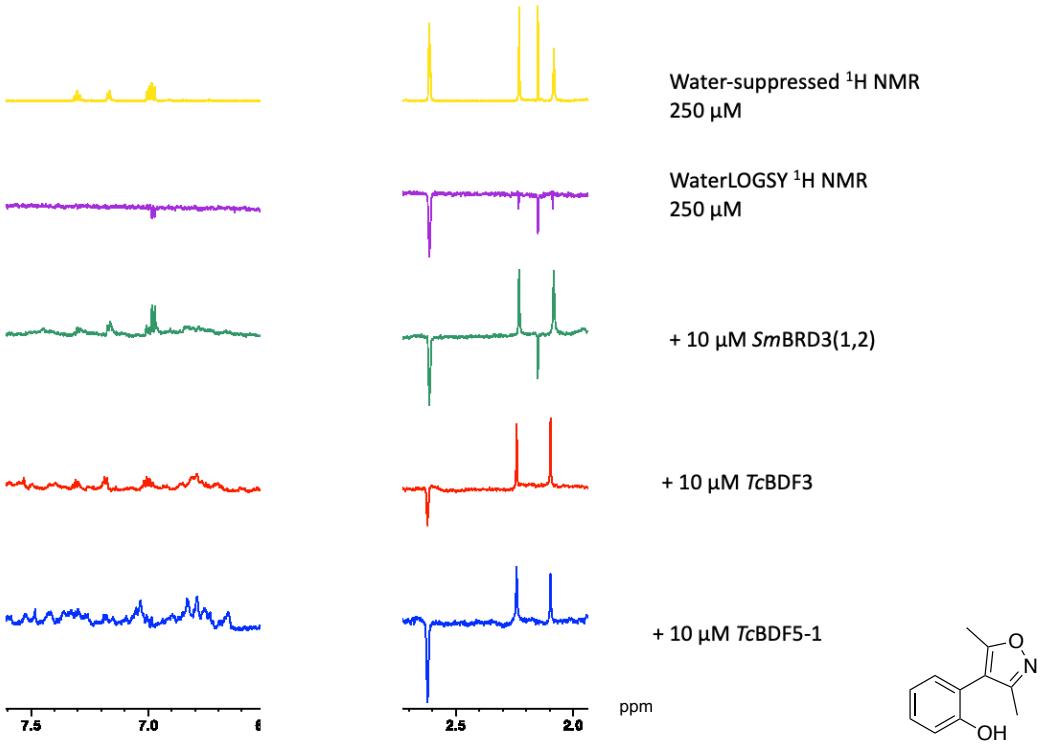


Figure B.52: Partial waterLOGSY spectra of **57**. The water suppressed ^{1}H spectrum is shown in yellow, the waterLOGSY NMR of the compound only (purple) shows inversion of the compounds peaks indicating suitable solubilisation at 250 µM. The waterLOGSY spectrum in the presence of 10 µM of *Sm*BRD3(1,2) (green), 10 µM of *Tc*BDF2 (red) or 5 µM of *Tc*BDF5-1 (blue) display inversion of the ligand peaks, indicating binding to all three proteins.

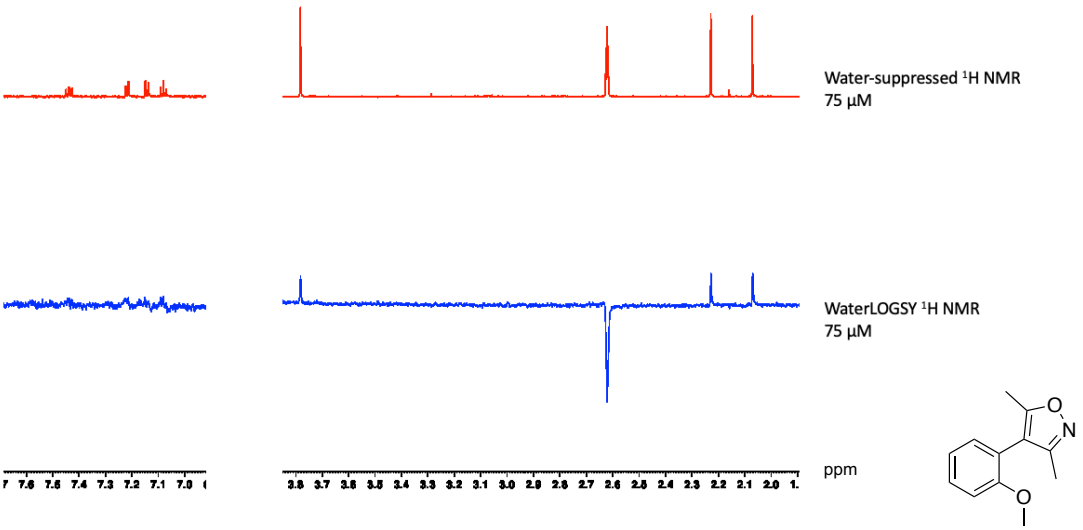


Figure B.53: Partial waterLOGSY spectra of **63**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 75 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference.

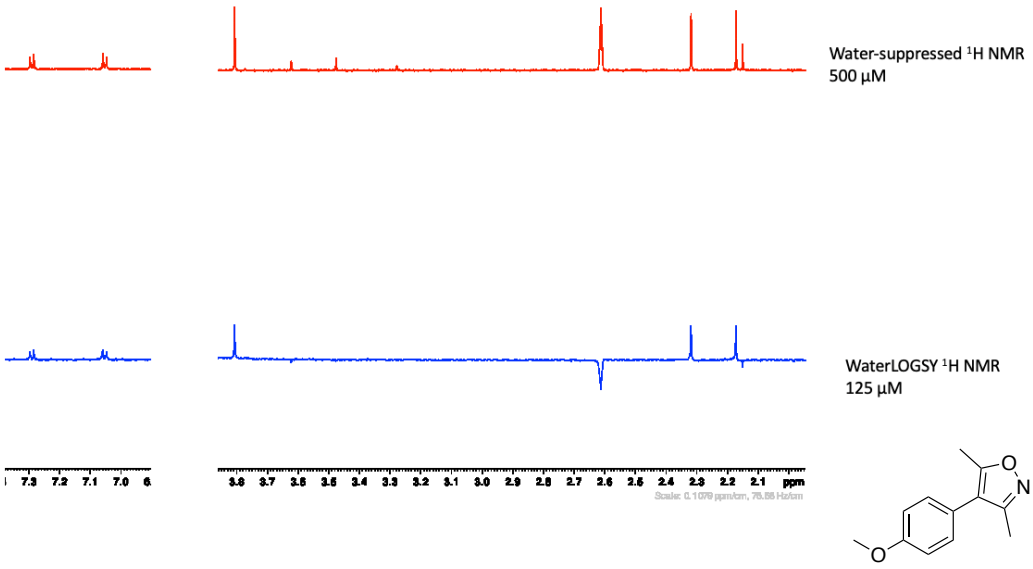


Figure B.54: Partial waterLOGSY spectra of **65**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 125 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference.

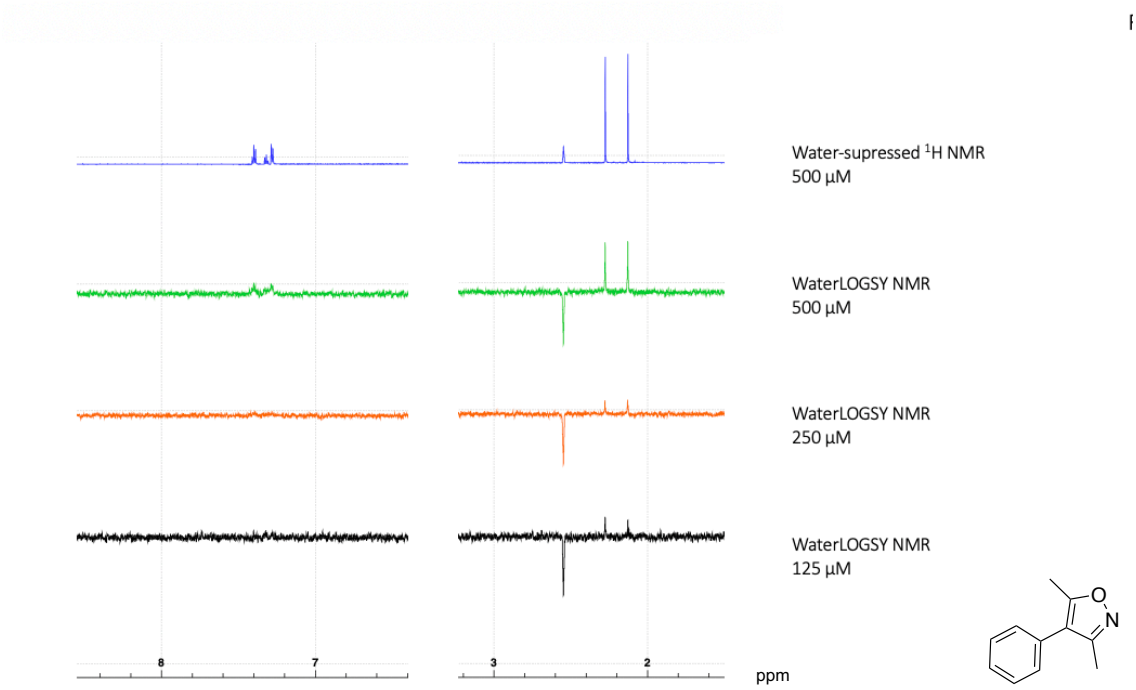


Figure B.55: Partial waterLOGSY spectra of **123**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 500 µM (green), 250 µM (orange) and 125 µM (black). The water-suppressed ^{1}H NMR spectrum (purple) is given as reference. **123** was synthesised by Fern Woodhouse (University of Oxford).

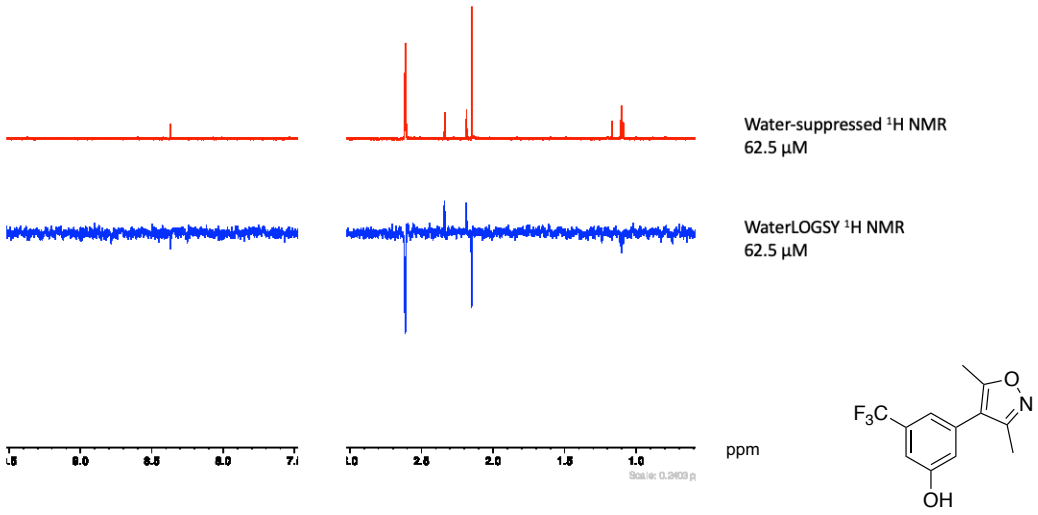


Figure B.56: Partial waterLOGSY spectra of **124**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 62.5 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference. **124** was synthesised by Leila Tan (University of Oxford).

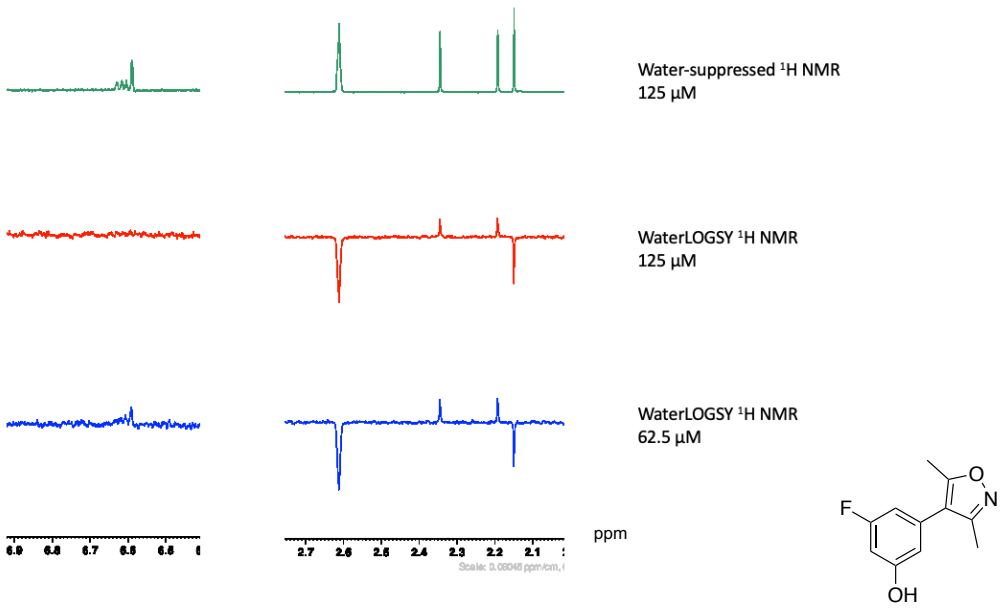


Figure B.57: Partial waterLOGSY spectra of **125**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 125 µM (red) and 62.5 µM (blue). The water-suppressed ^{1}H NMR spectrum (green) is given as reference. **125** was synthesised by Leila Tan (University of Oxford).

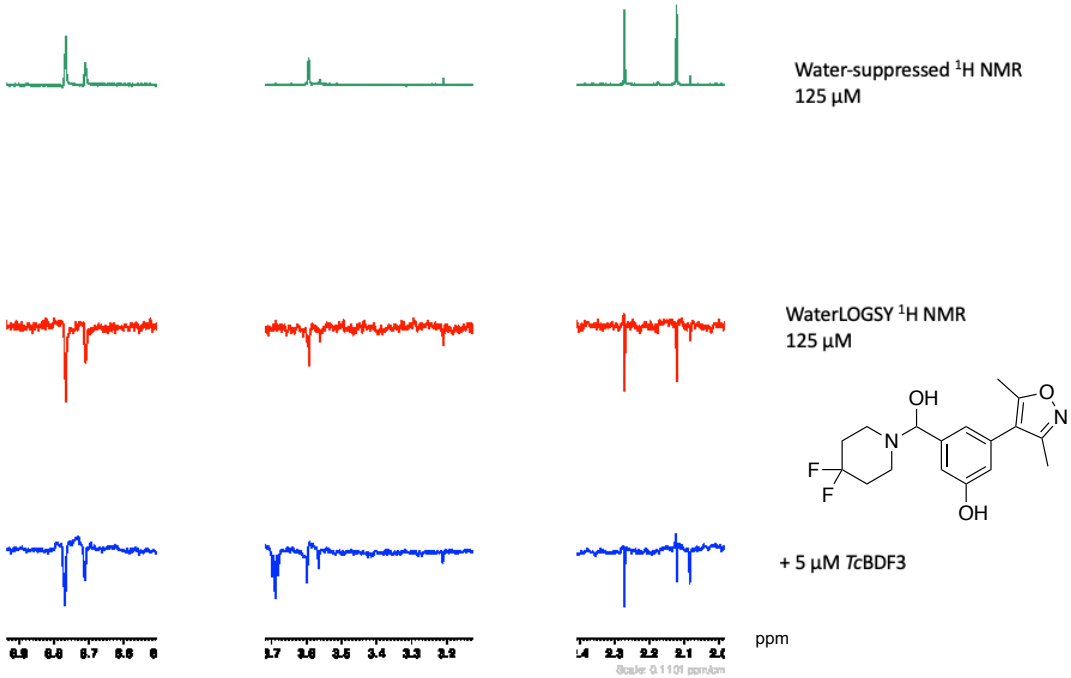


Figure B.58: Partial waterLOGSY spectra of **108**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind. **108** was synthesised by Dr Laura Jennings (University of Oxford).

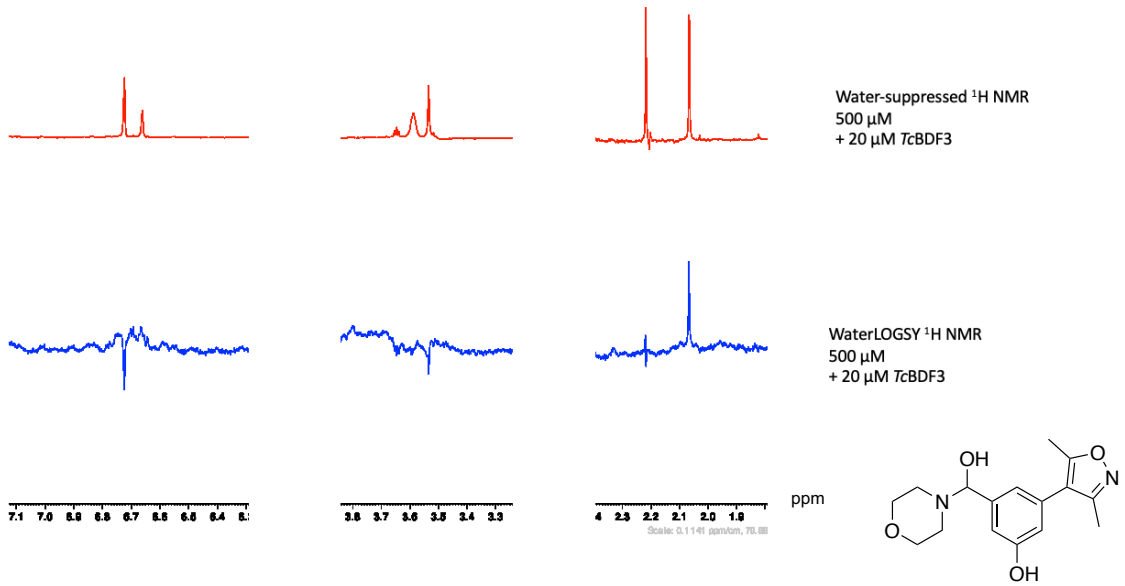


Figure B.59: Partial waterLOGSY spectra of **109**. Partial inversion of the ligand aliphatic peaks is detected in the presence of 20 µM of TcBDF3 (blue), however, the aromatics peaks are not inverted. This ambiguous result might indicate binding. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference. **109** was synthesised by Dr Laura Jennings (University of Oxford).

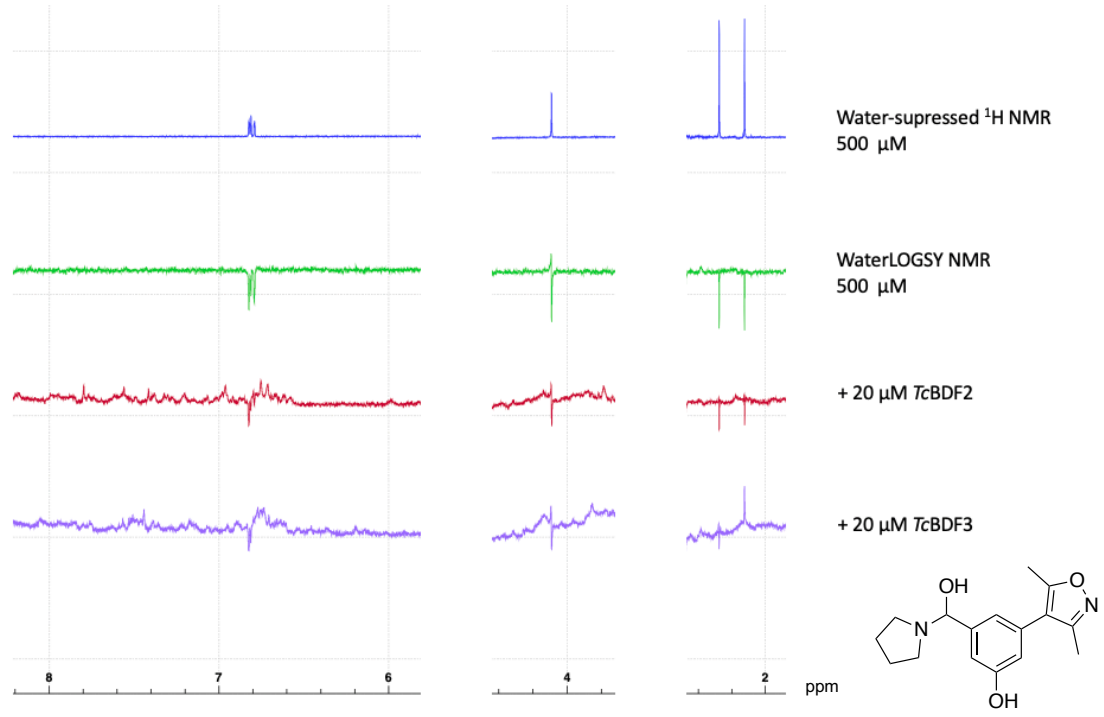


Figure B.60: Partial waterLOGSY spectra of **110**. The water suppressed ^{1}H spectrum is shown in blue, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. The waterLOGSY spectrum in the presence 20 µM *Tc*BDF3 (purple) display partial inversion of the ligand peaks, indicating potential binding. No inversion is seen in the waterLOGSY spectrum in the presence of 20 µM of *Tc*BDF2 (red). **LJ403** was synthesised by Dr Laura Jennings (University of Oxford).

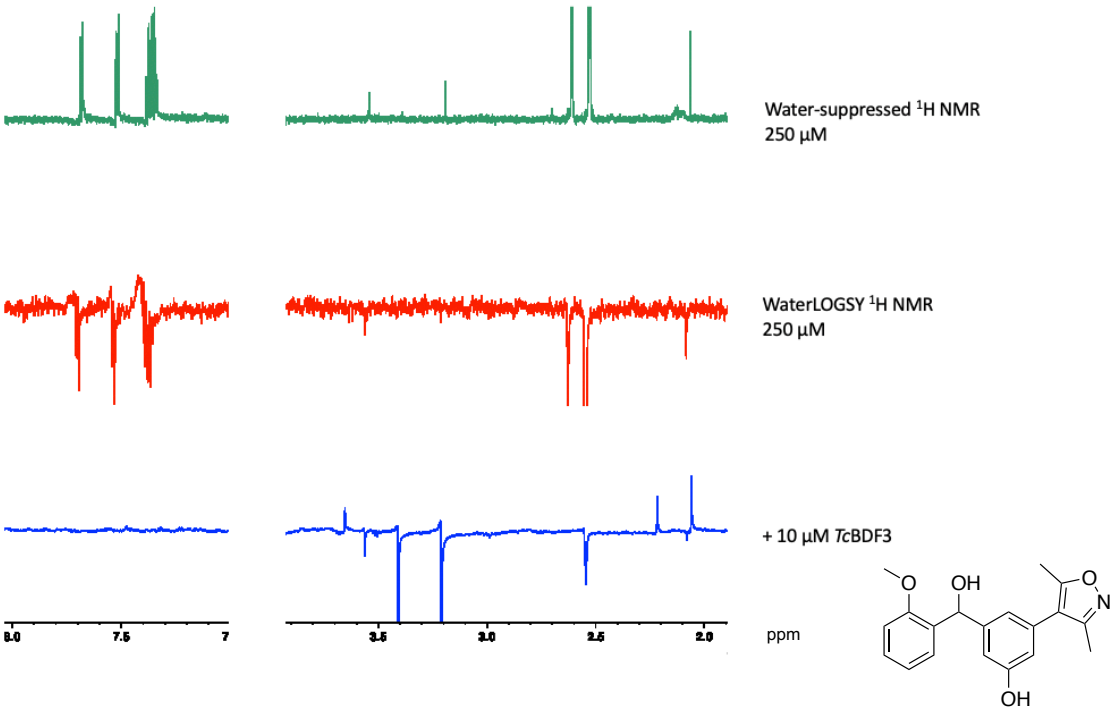


Figure B.61: Partial waterLOGSY spectra of **111**. Partial inversion of the ligand aliphatic peaks is detected in the presence of 20 µM of TcBDF3 (blue), however, the aromatics peaks are not inverted. This ambiguous result might indicate binding. The water suppressed ^{1}H spectrum of 500 µM ligand with 20 µM protein (red) is given as a reference. **111** was synthesised by Ray Berkeley (University of Oxford).

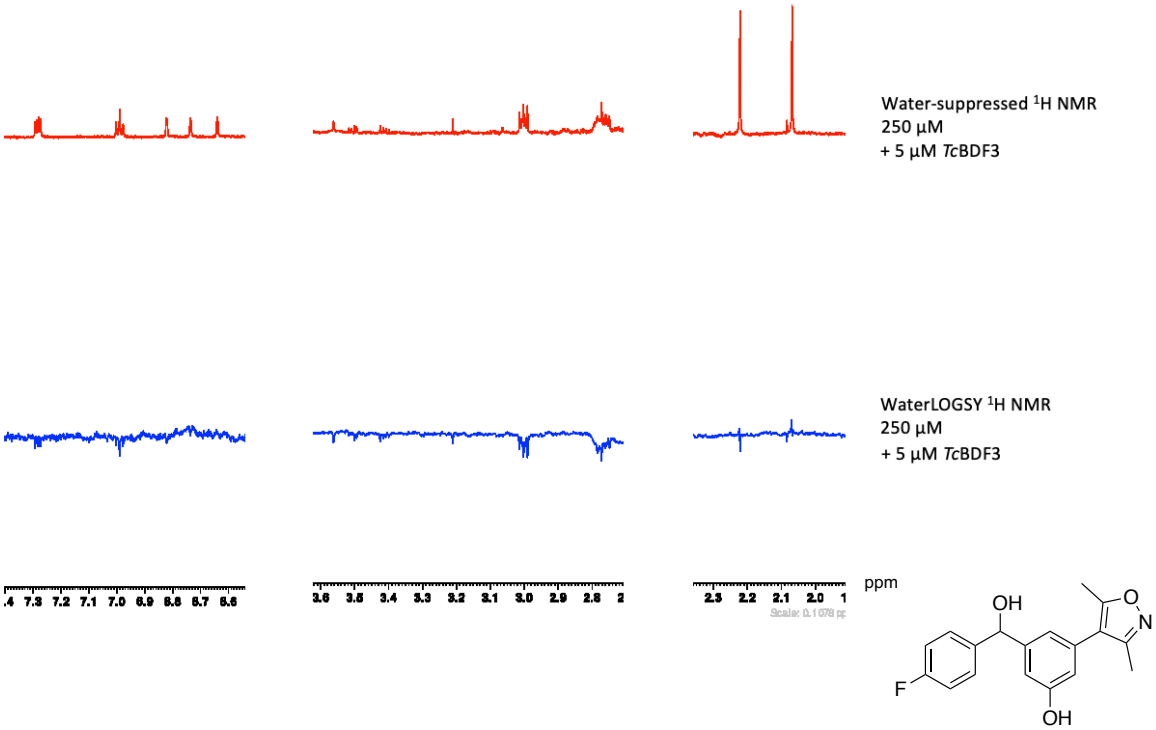


Figure B.62: Partial waterLOGSY spectra of **112**. No clear inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 5 µM protein (red) is given as a reference. **112** was synthesised by Dr Laura Jennings, Amy Hicks, Ray Berkeley, or Dr Charles Evans (University of Oxford).

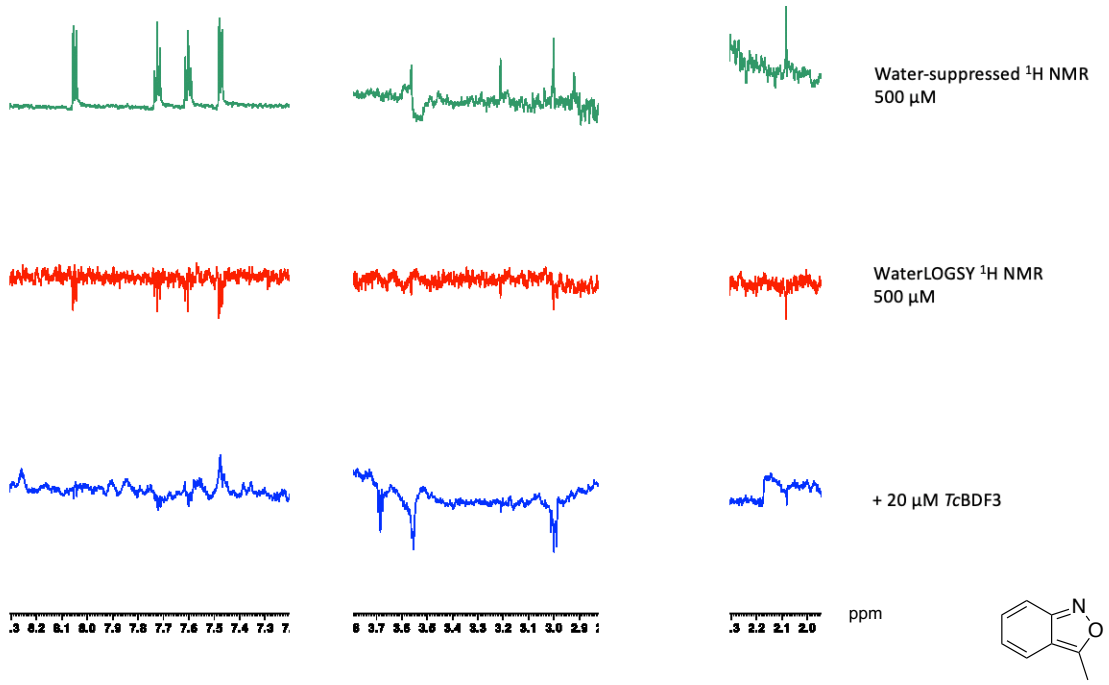


Figure B.63: Partial waterLOGSY spectra of **179**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. No inversion of the ligand peak is detected in the presence of 20 µM of TcBDF3 (blue) indicating the ligand does not bind. **179** was synthesised by Dr Laura Jennings (University of Oxford).

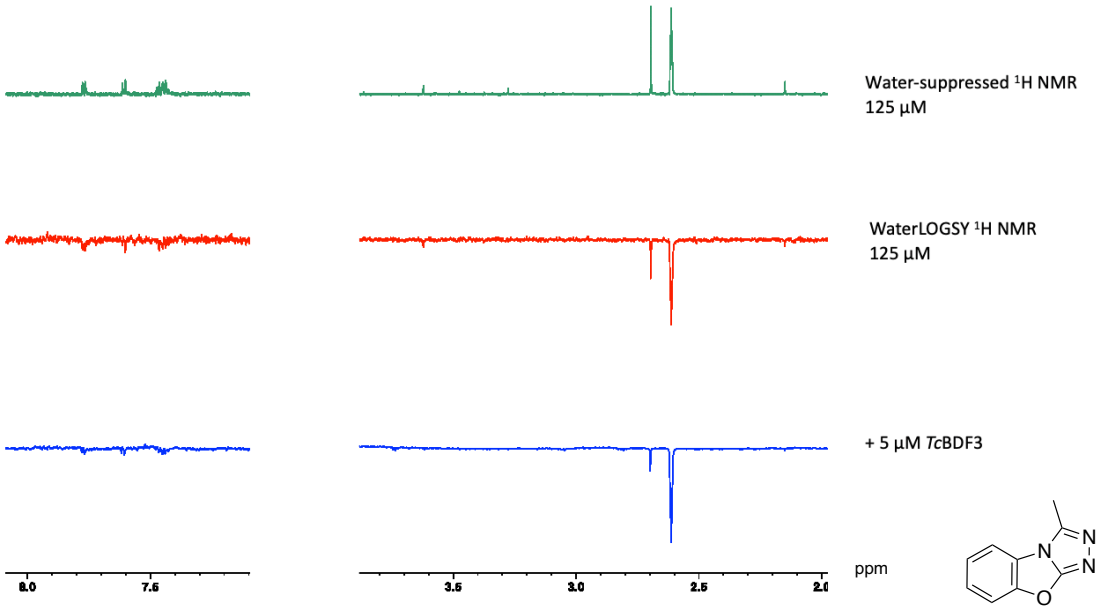


Figure B.64: Partial waterLOGSY spectra of **180**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind. **180** was synthesised by Dr Laura Jennings (University of Oxford).

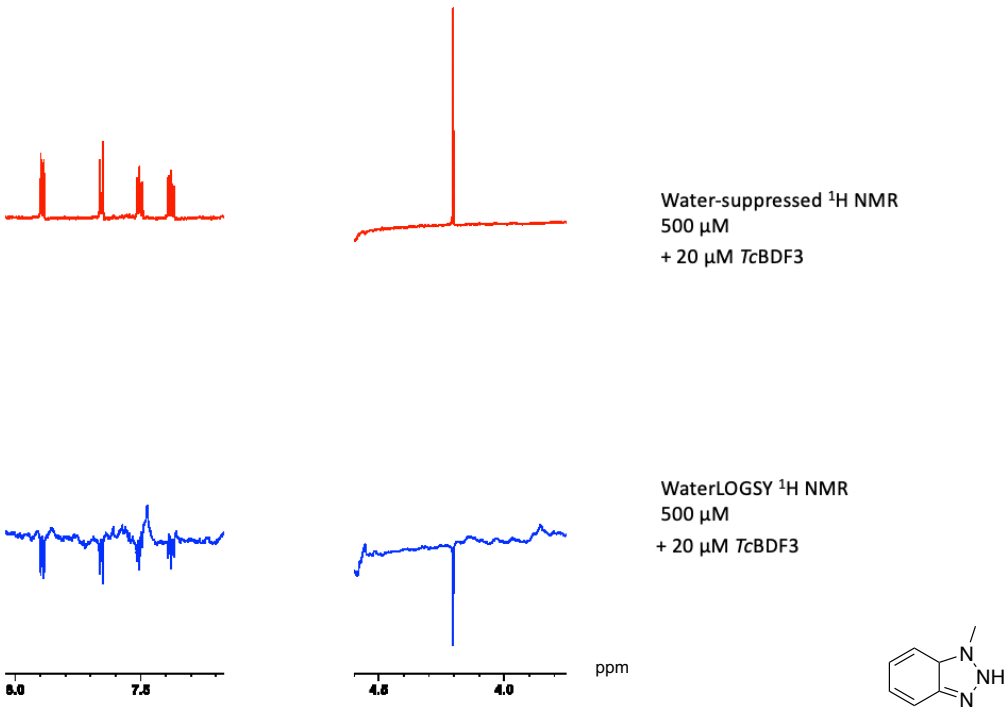


Figure B.65: Partial waterLOGSY spectra of **211**. No clear inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 5 µM protein (red) is given as a reference. **211** was synthesised by Dr Laura Jennings (University of Oxford).

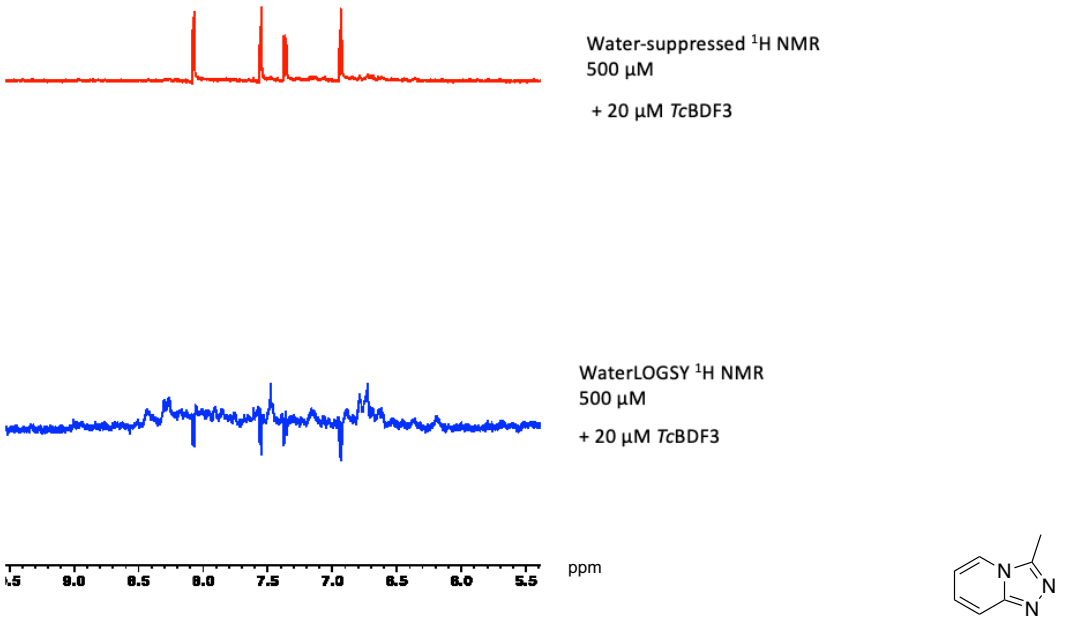


Figure B.66: Partial waterLOGSY spectra of **212** No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 10 µM protein (red) is given as a reference. **212** was synthesised by Dr Laura Jennings (University of Oxford).

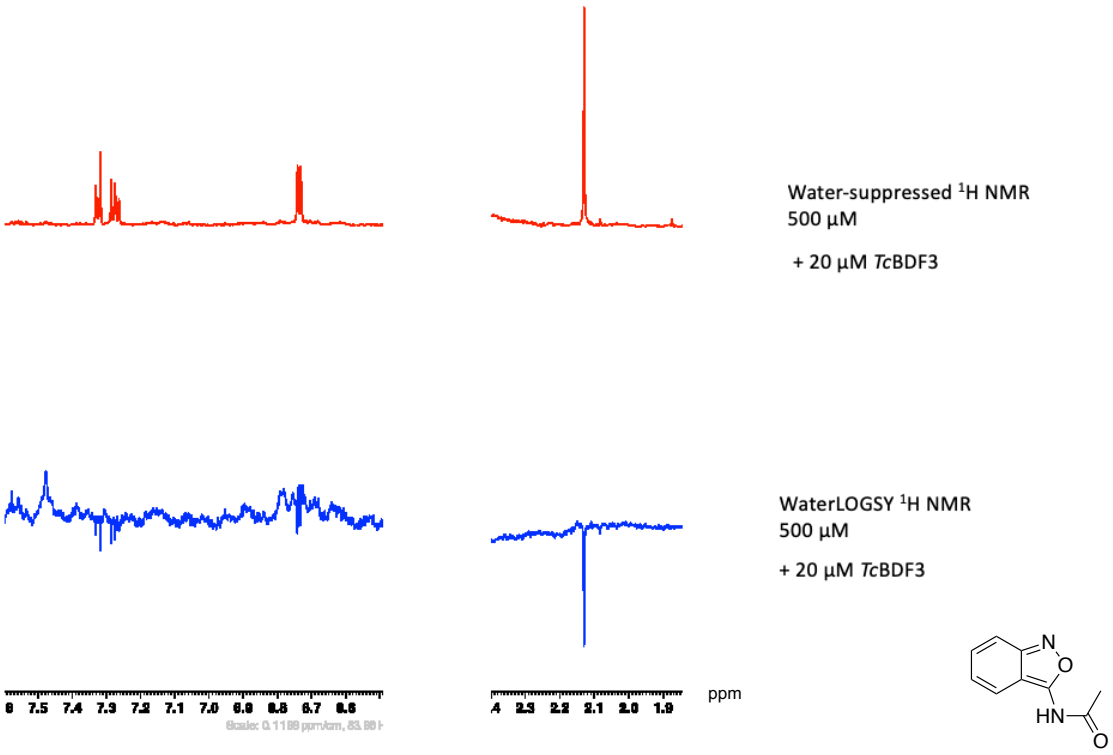


Figure B.67: Partial waterLOGSY spectra of **213** No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 10 µM protein (red) is given as a reference. **213** was synthesised by Dr Laura Jennings (University of Oxford).

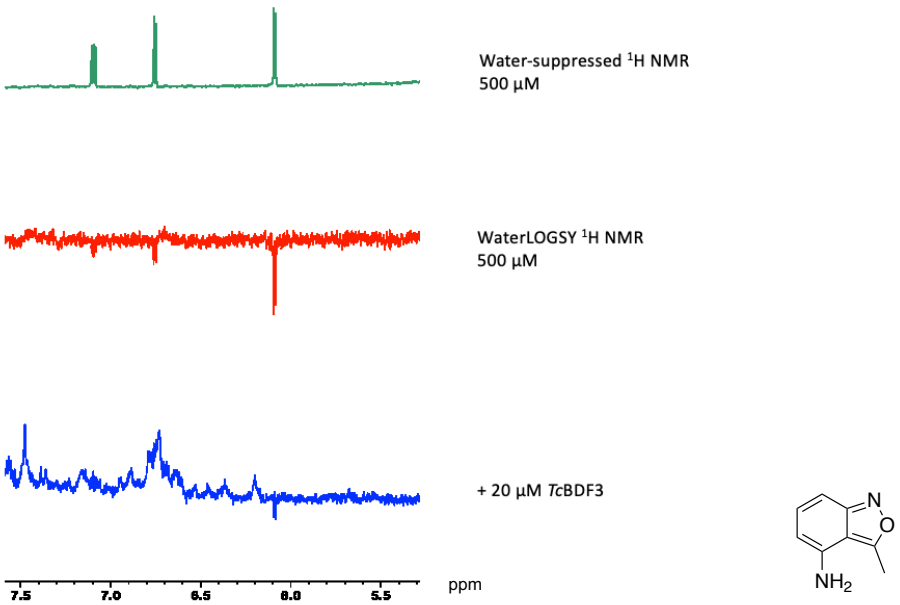


Figure B.68: Partial waterLOGSY spectra of **214** No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 10 µM protein (red) is given as a reference. **214** was synthesised by Dr Laura Jennings (University of Oxford).

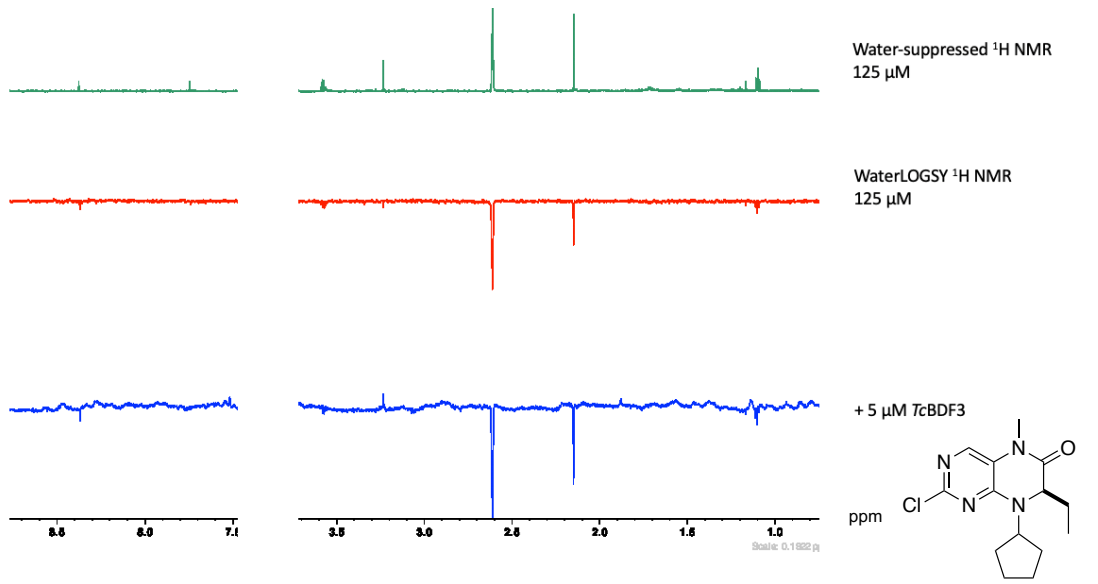


Figure B.69: Partial waterLOGSY spectra of (*R*)-92. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. The waterLOGSY spectrum in the presence of 5 µM of *Tc*BDF3 (blue) display inversion of the ligand peaks, indicating binding.

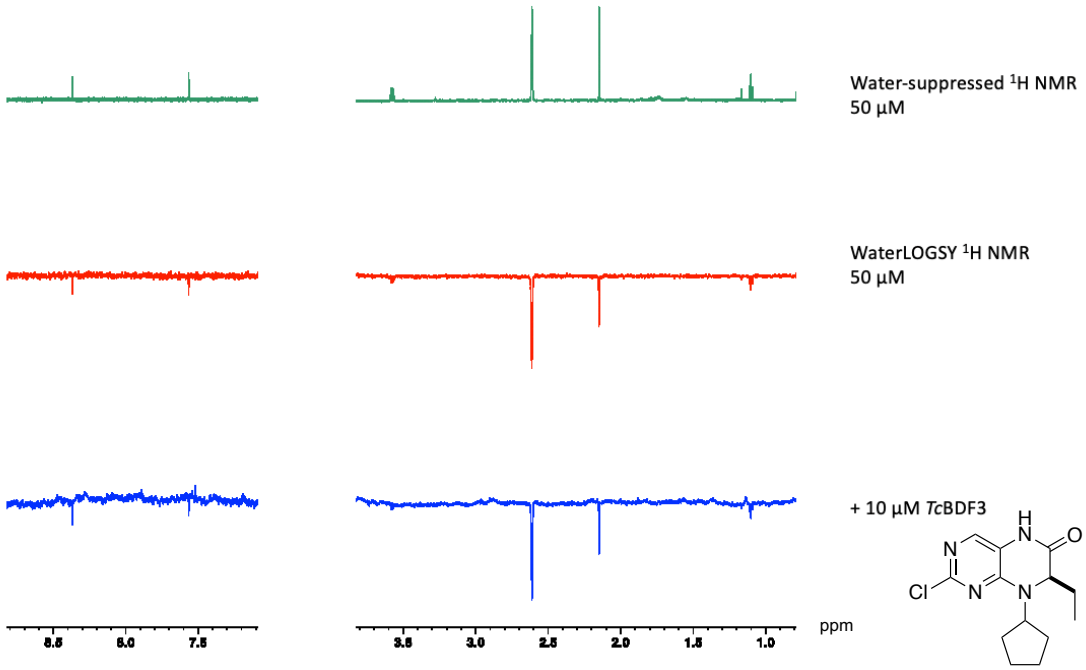


Figure B.70: Partial waterLOGSY spectra of (*R*)-96. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 50 µM. No inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind.

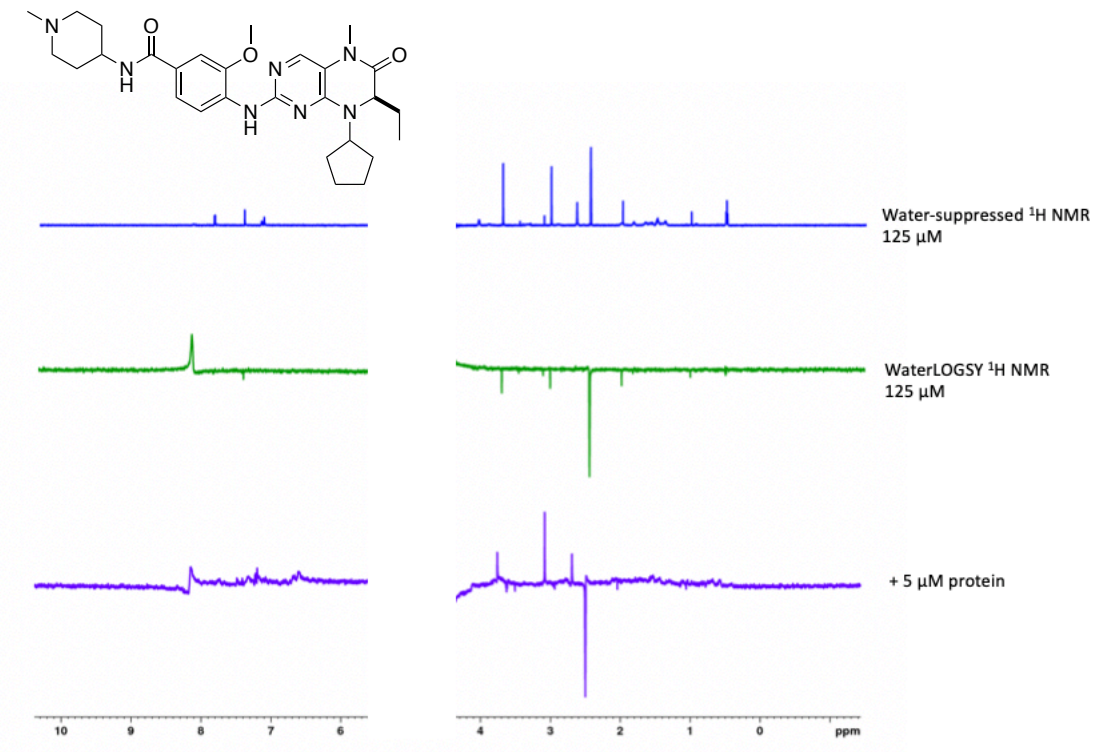


Figure B.71: Partial waterLOGSY spectra of BI-2536 (**8**). The water suppressed ^{1}H spectrum is shown in blue, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. The waterLOGSY spectrum in the presence of 5 µM of *Tc*BDF3 (purple) display inversion of the ligand peaks, indicating binding.

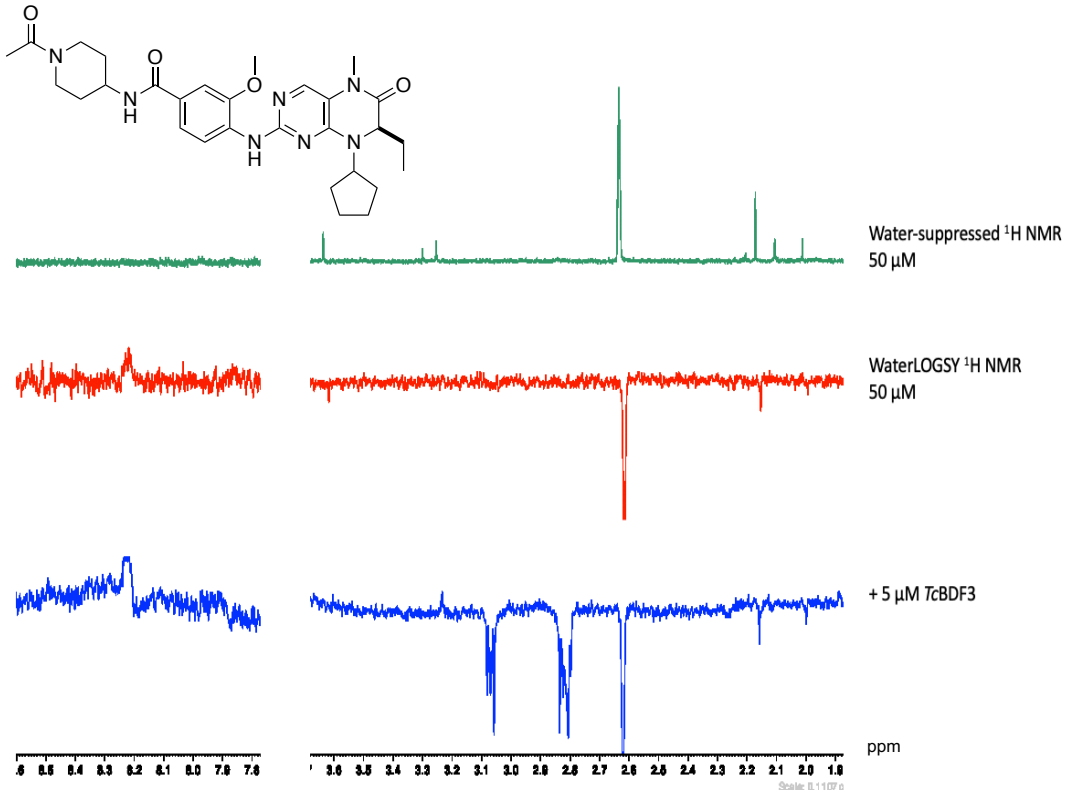


Figure B.72: Partial waterLOGSY spectra of **(*R*)-155**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 50 µM. The waterLOGSY spectrum in the presence of 5 µM of *Tc*BDF3 (blue) display inversion of the ligand peaks, indicating binding.

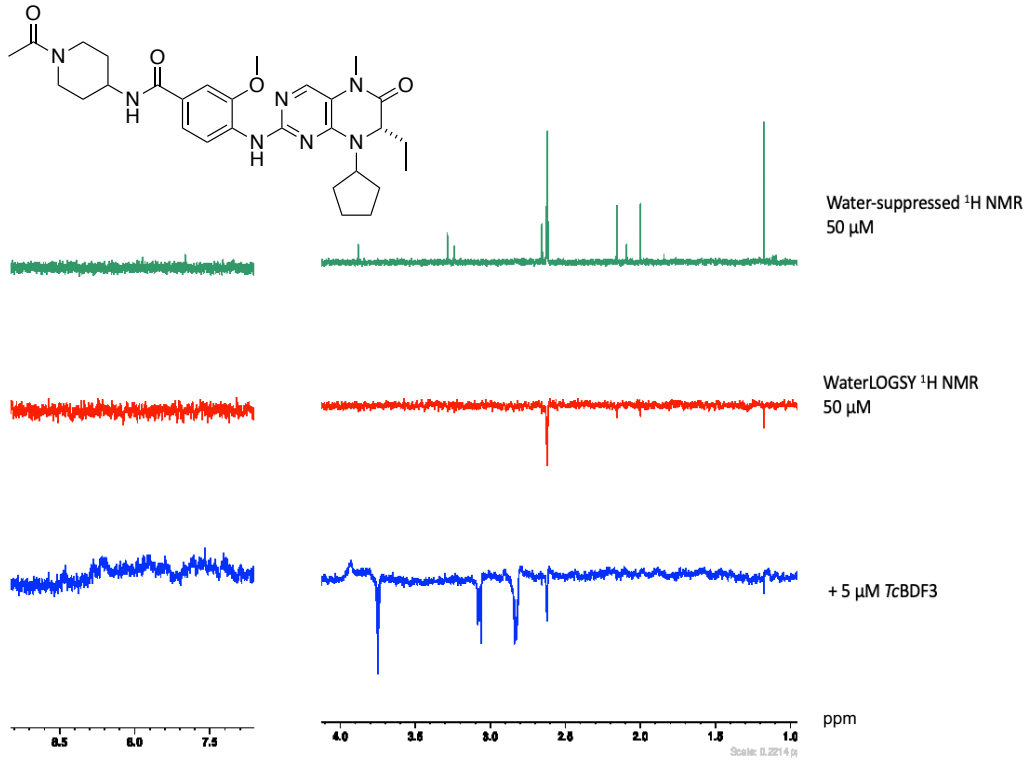


Figure B.73: Partial waterLOGSY spectra of (*S*)-155. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 50 µM. No clear inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

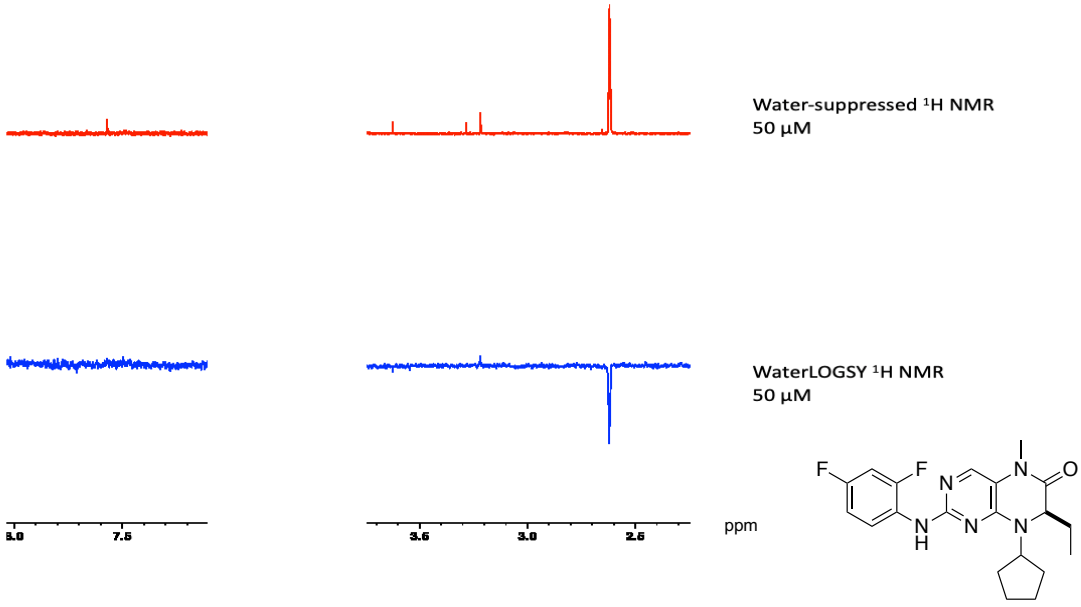


Figure B.74: Partial waterLOGSY spectra of (*R*)-130. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 50 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference.

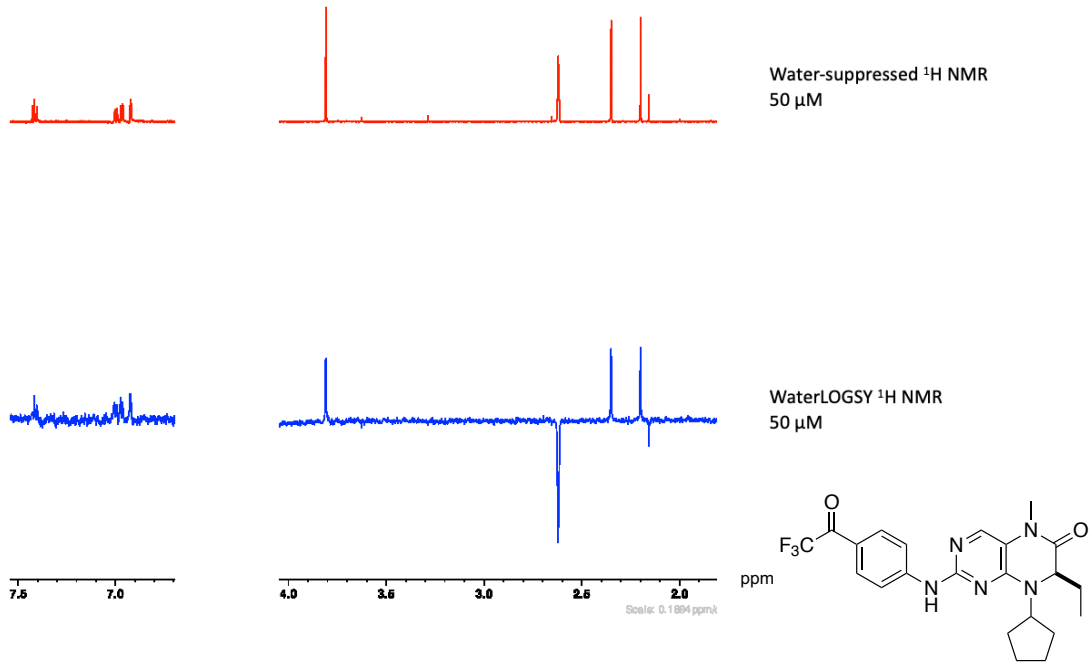


Figure B.75: Partial waterLOGSY spectra of (*R*)-131. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 50 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference.

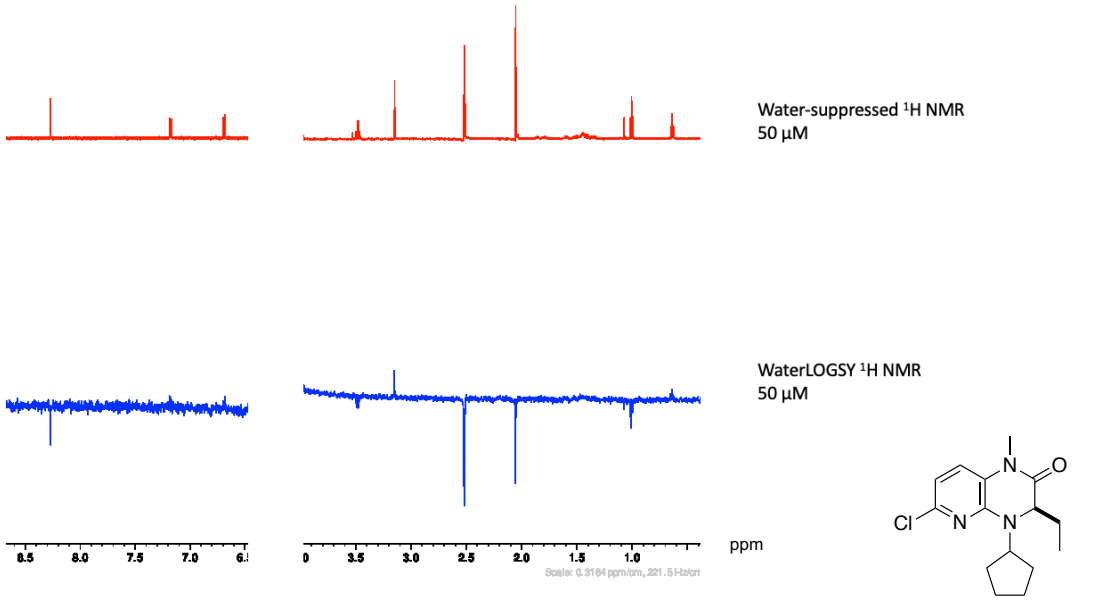


Figure B.76: Partial waterLOGSY spectra of **215**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 62.5 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference. **215** was synthesised by Michelle Keller (University of Oxford).

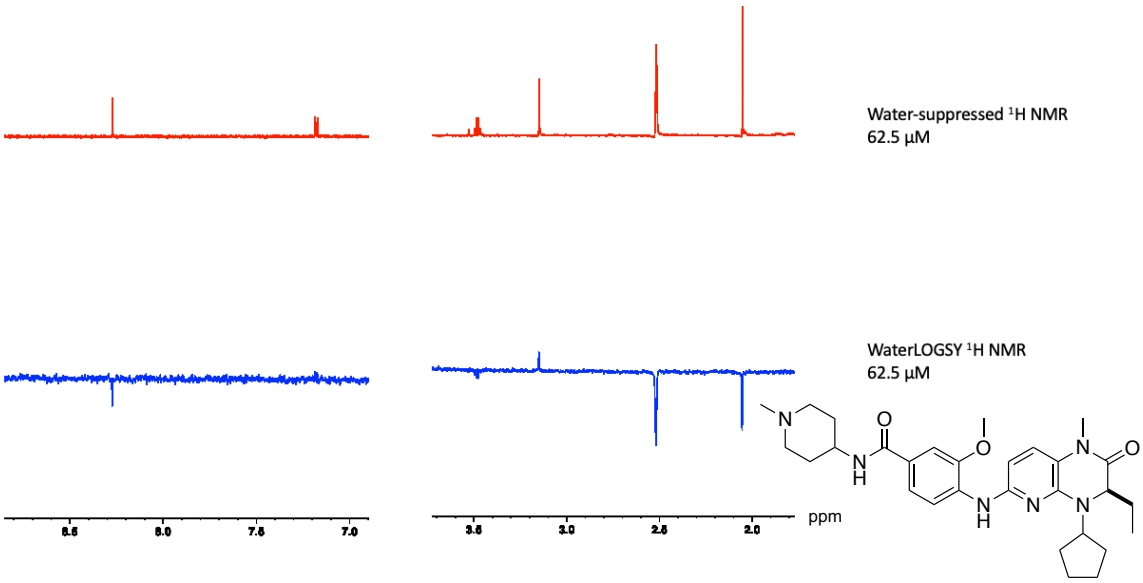


Figure B.77: Partial waterLOGSY spectra of **160**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 62.5 µM (blue). The water-suppressed ^{1}H NMR spectrum (red) is given as reference. **160** was synthesised by Michelle Keller (University of Oxford).

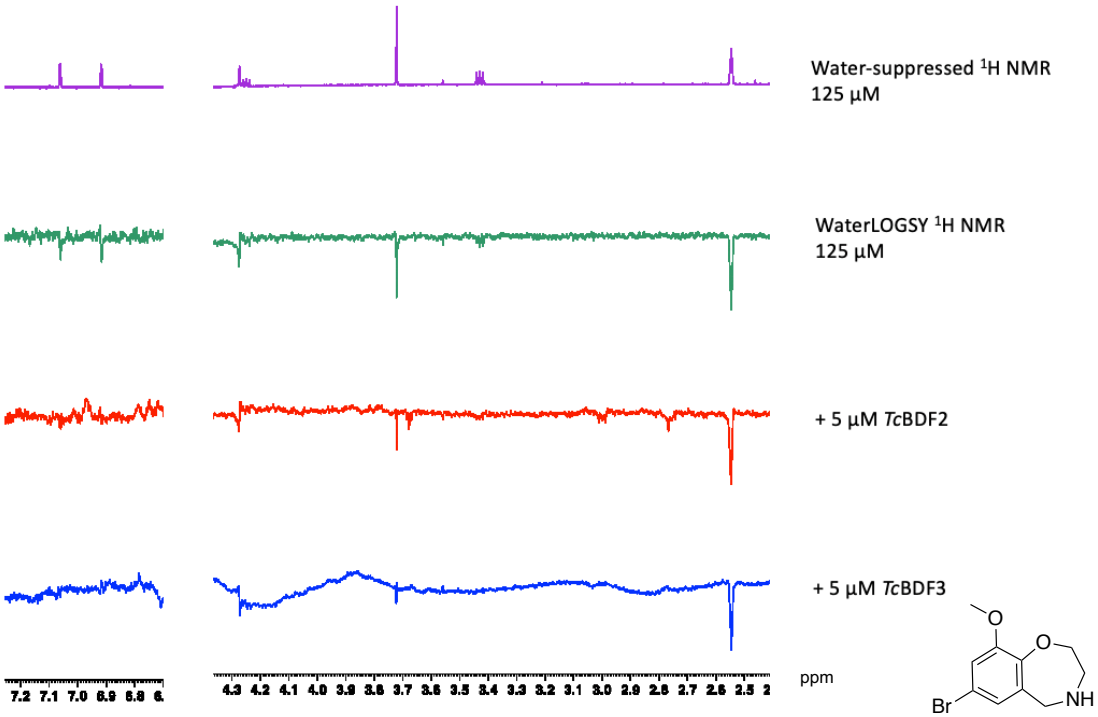


Figure B.78: Partial waterLOGSY spectra of **90**. The water suppressed ^{1}H spectrum is shown in purple, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF2 (red) or 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

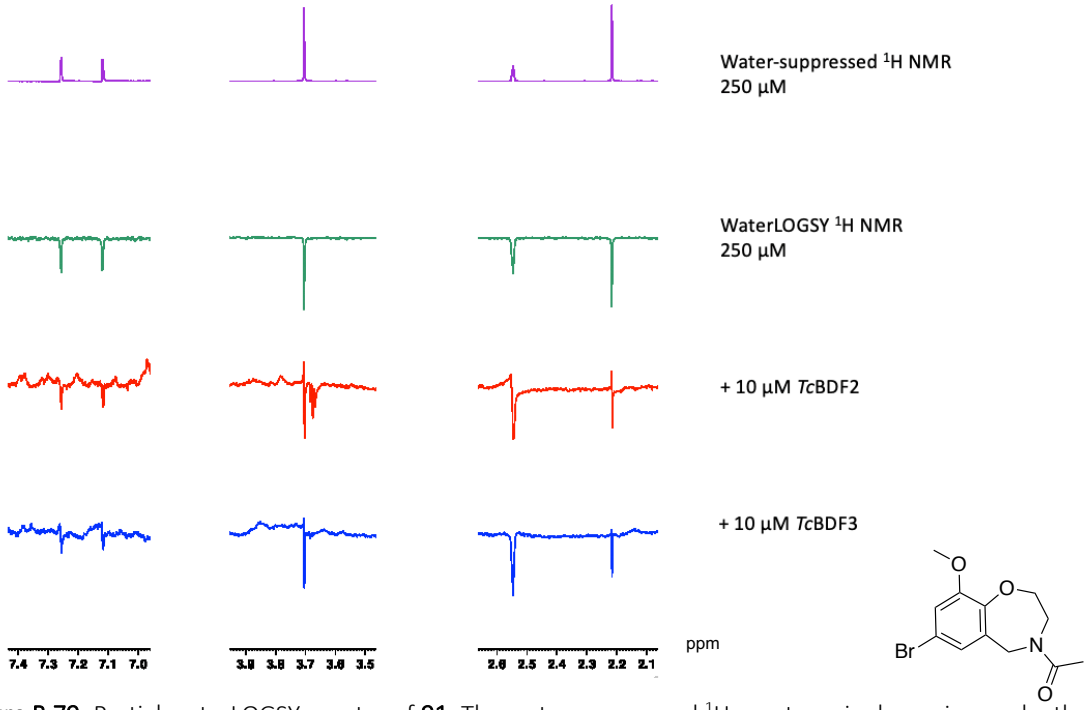


Figure B.79: Partial waterLOGSY spectra of 91. The water suppressed ^{1}H spectrum is shown in purple, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF2 (red) or 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

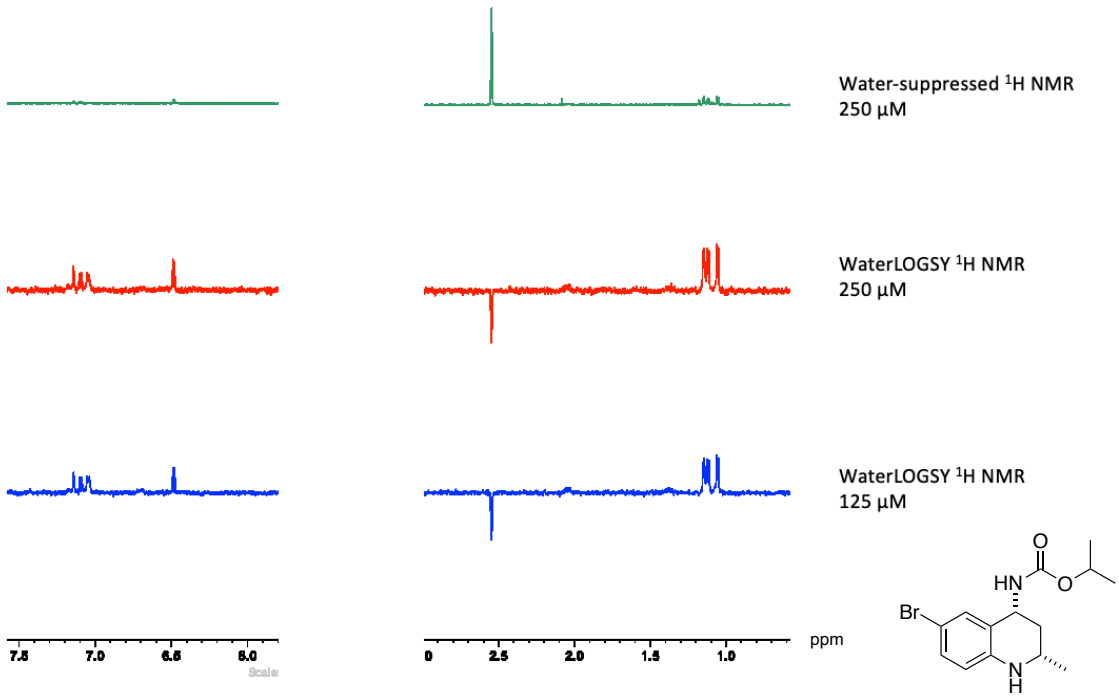


Figure B.80: Partial waterLOGSY spectra of 85. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 500 µM (red) and 250 µM (blue). The water-suppressed ^{1}H NMR spectrum (green) is given as reference.

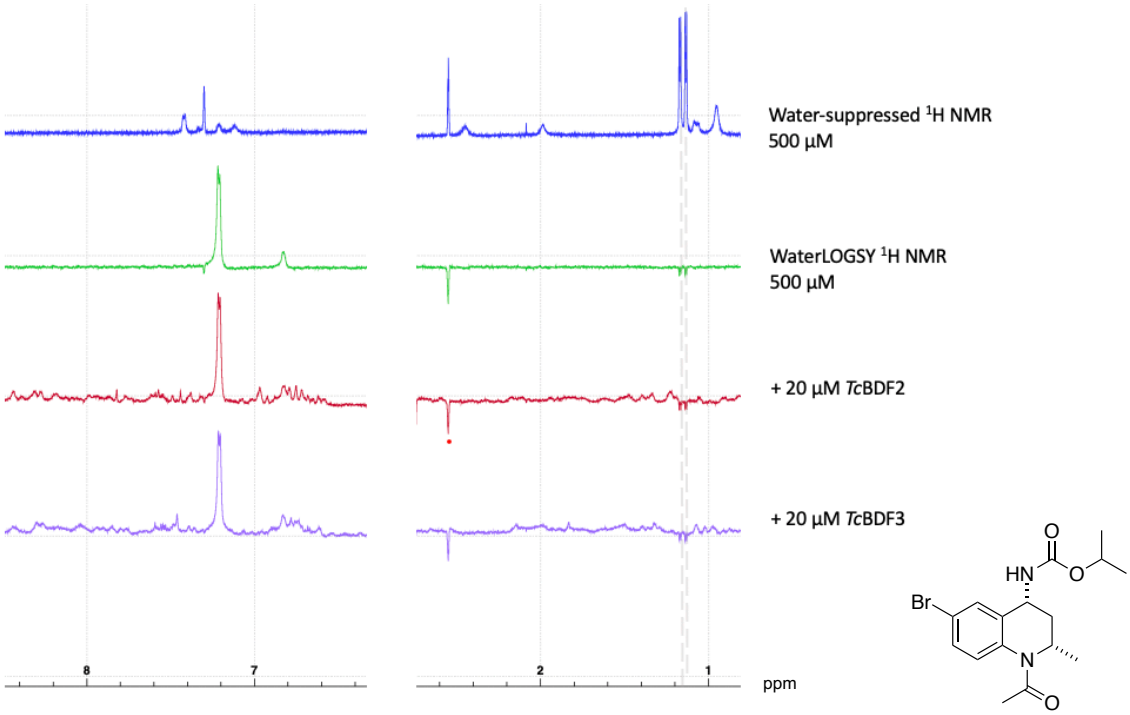


Figure B.81: Partial waterLOGSY spectra of **80**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 125 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

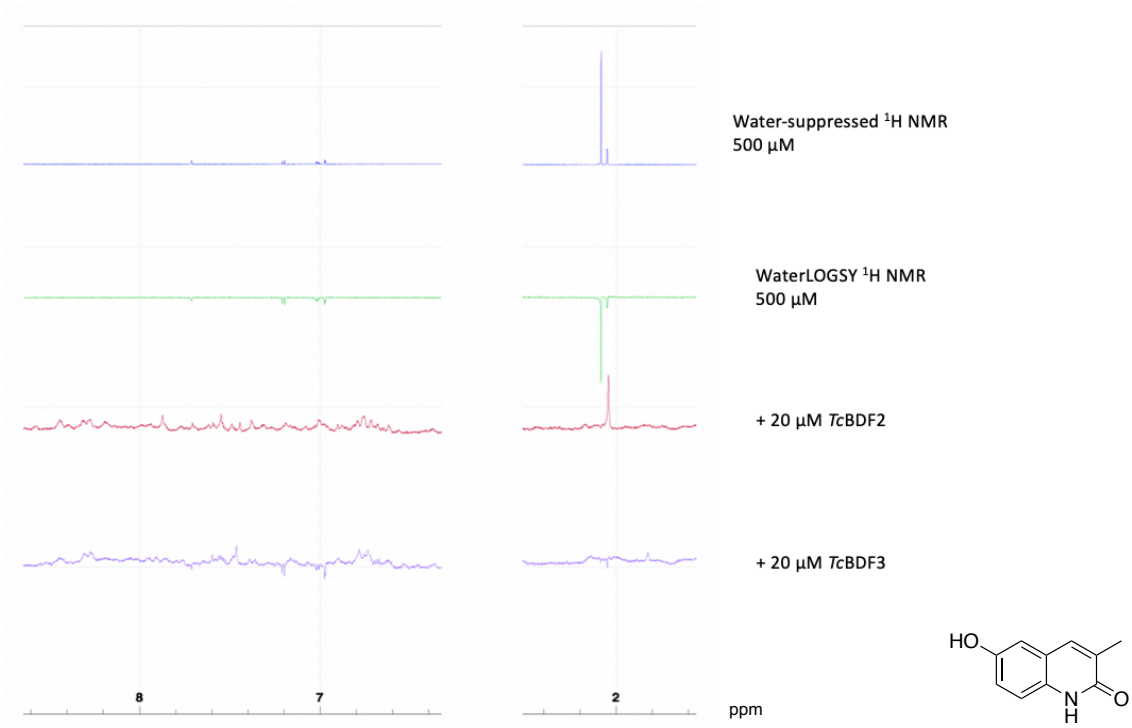


Figure B.82: Partial waterLOGSY spectra of **188.** The water suppressed ^{1}H spectrum is shown in blue, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation at 500 µM. Inversion of the ligand peaks on addition of 20 µM *Tc*BDF2 (red) indicates binding. The peaks are not inverted in the presence of 20 µM TcBDF3 (purple), showing that the binding is selective to *Tc*BDF2. **188** was synthesised by Fern Woodhouse (University of Oxford).

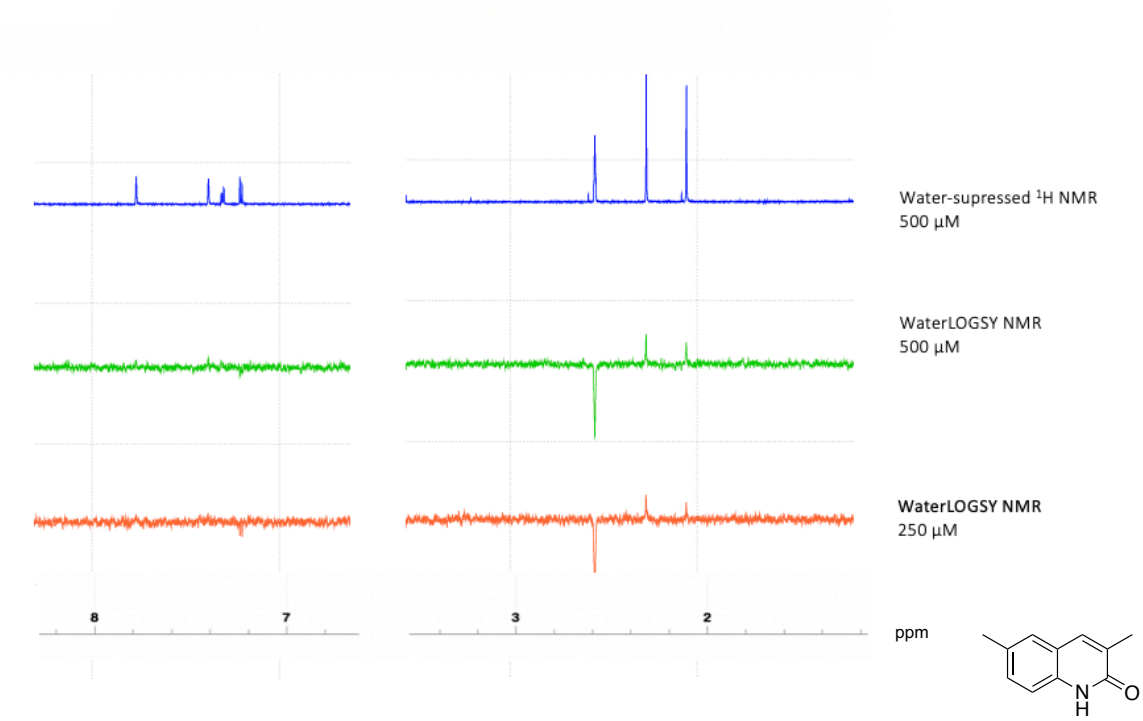


Figure B.83: Partial waterLOGSY spectra of **189**. Positive ligand peaks in the waterLOGSY spectrum indicates aggregation at 500 µM (green) and 250 µM (orange). The water-suppressed ^{1}H NMR spectrum (purple) is given as reference. **189** was synthesised by Fern Woodhouse (University of Oxford).

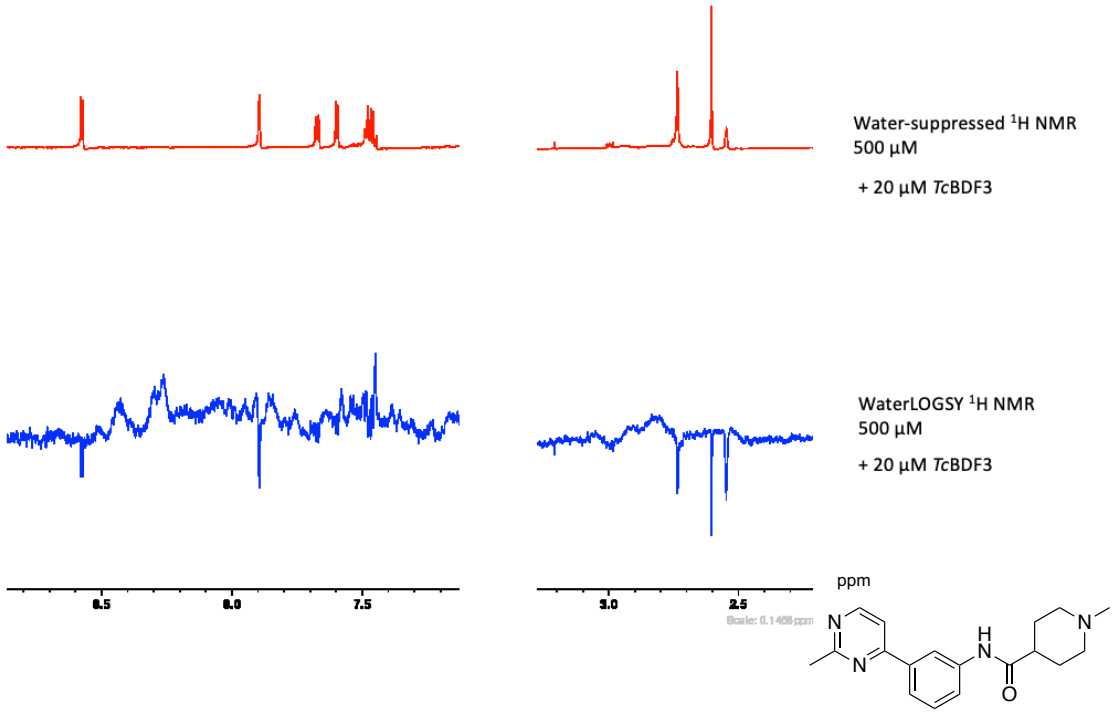


Figure B.84: Partial waterLOGSY spectra of **216** No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 10 µM protein (red) is given as a reference. **216** was synthesised by Dr Larissa See (University of Oxford).

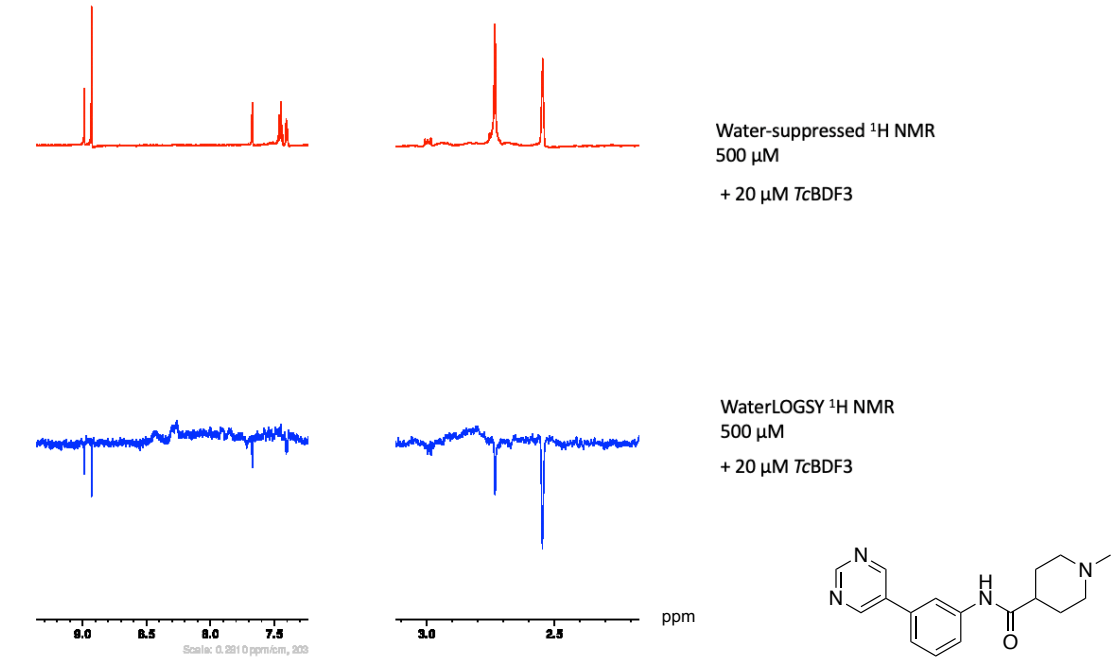


Figure B.85: Partial waterLOGSY spectra of **217** No clear inversion of the ligand peak is detected in the presence of 10 µM of TcBDF3 (blue) indicating the ligand does not bind. The water suppressed ^{1}H spectrum of 250 µM ligand with 10 µM protein (red) is given as a reference. **217** was synthesised by Dr Larissa See (University of Oxford).

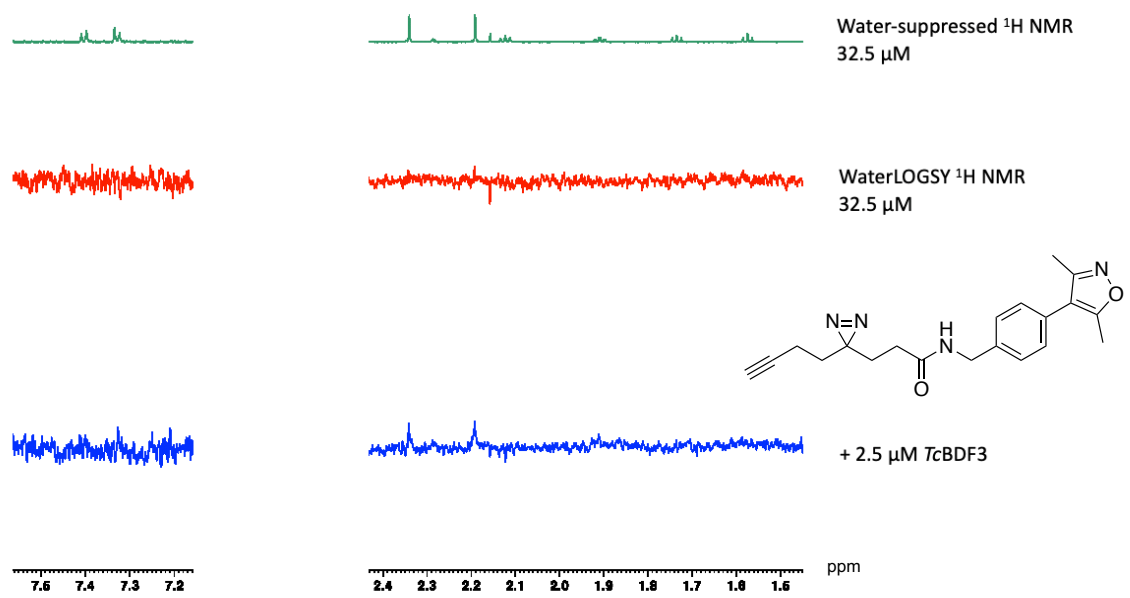


Figure B.86: Partial waterLOGSY spectra of **162**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 32.5 µM. The waterLOGSY spectrum in the presence 2.5 µM *Tc*BDF3 (blue) display inversion of the ligand peaks, indicating binding.

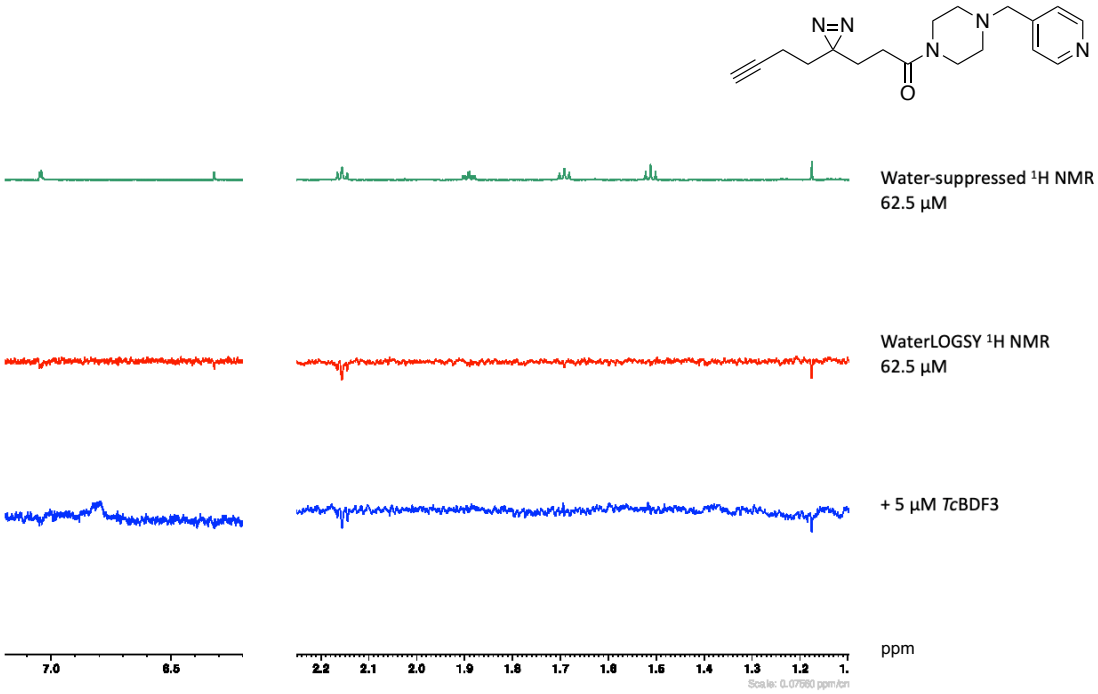


Figure B.87: Partial waterLOGSY spectra of **164**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 62.5 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

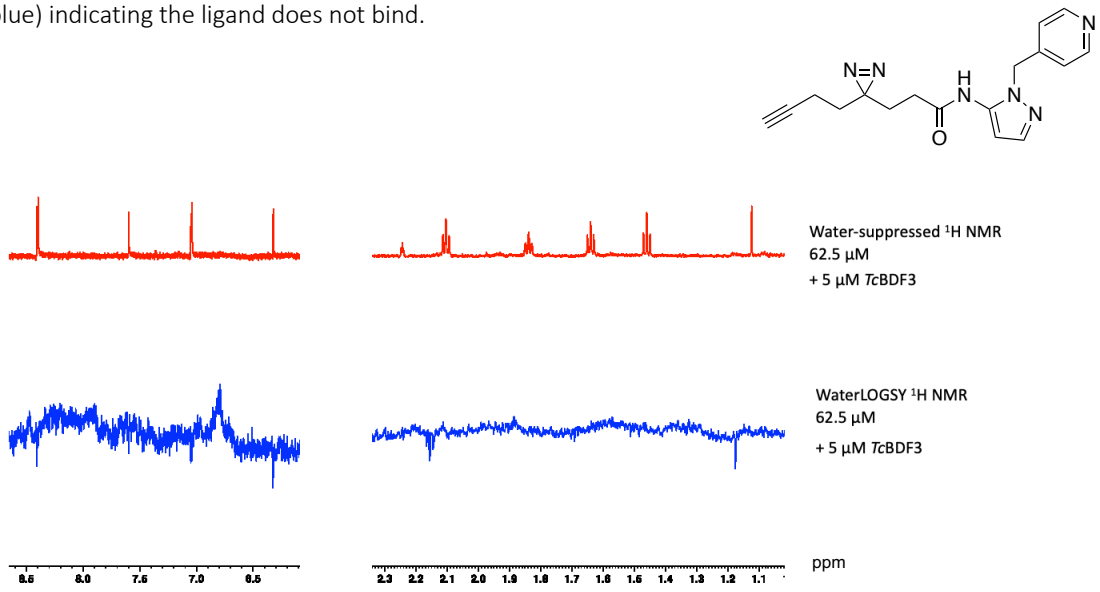


Figure B.88: Partial waterLOGSY spectra of **168**. The water suppressed ^{1}H spectrum is shown in green, the waterLOGSY NMR of the compound only (red) shows inversion of the compounds peaks indicating suitable solubilisation at 62.5 µM. No inversion of the ligand peak is detected in the presence of 5 µM of TcBDF3 (blue) indicating the ligand does not bind.

B.2.2. WaterLOGSY of ATAD2 binders against TcBDF2

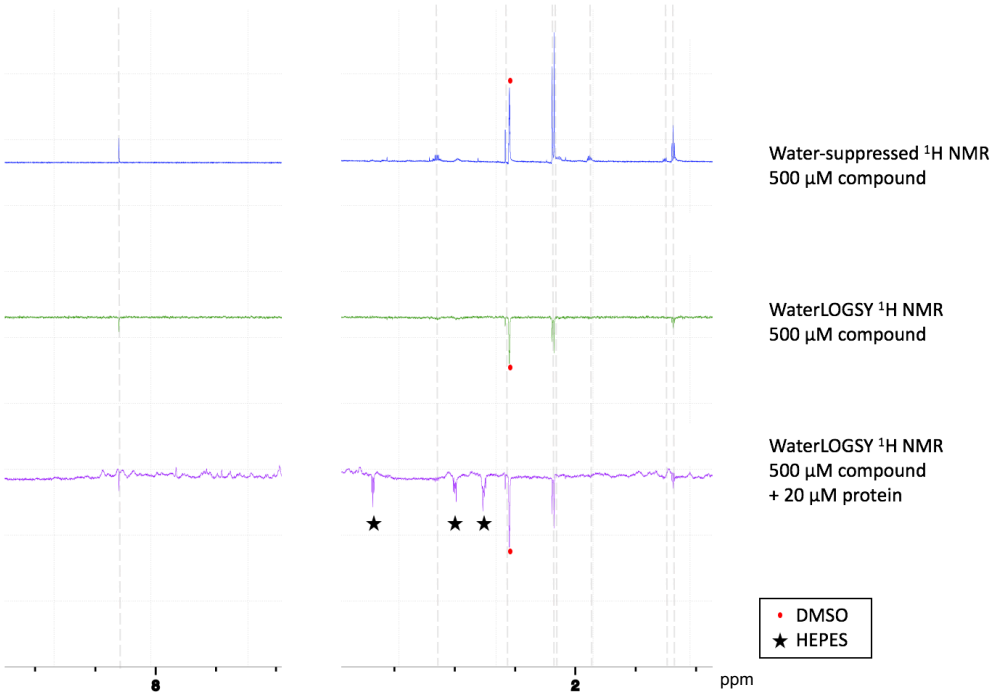


Figure B.89: Partial waterLOGSY spectra of **218** (representative results obtained for the ATAD2 binders against *Tc*BDF2). The water repressed ^{1}H spectra is shown in blue, the waterLOGSY NMR of the compound only (green) shows inversion of the compounds peaks indicating suitable solubilisation. The waterLOGSY spectra in the presence of the protein (purple) does not show inversion of the compound peaks indicating that the compound is not binding.

Table B.3: WaterLOGSY screen of ATAD2 binders against *Tc*BDF2 results. No binding was observed.

Compound	MW	LogD7.4	SFI	solubility	*Tc*BDF2
127	245.32	0.15	1.15	Soluble at 500 µM	No binding
128	269.34	0.89	0.89	Soluble at 500 µM	No binding
129	249.27	0.30	1.30	Soluble at 500 µM	No binding
218	263.34	-2.37	-1.37	Soluble at 500 µM	No binding
219	267.37	-0.11	-0.11	Soluble at 500 µM	No binding
220	219.24	-0.27	0.63	Soluble at 500 µM	No binding
221	235.28	-3.42	-2.42	Soluble at 500 µM	No binding
222	241.28	-0.36	-0.36	Soluble at 500 µM	No binding
223	191.10	-2.8	-1.8	Soluble at 500 µM	No binding
224	261.28	0.59	1.59	Soluble at 500 µM	No binding
225	233.27	0.09	1.09	Soluble at 500 µM	No binding

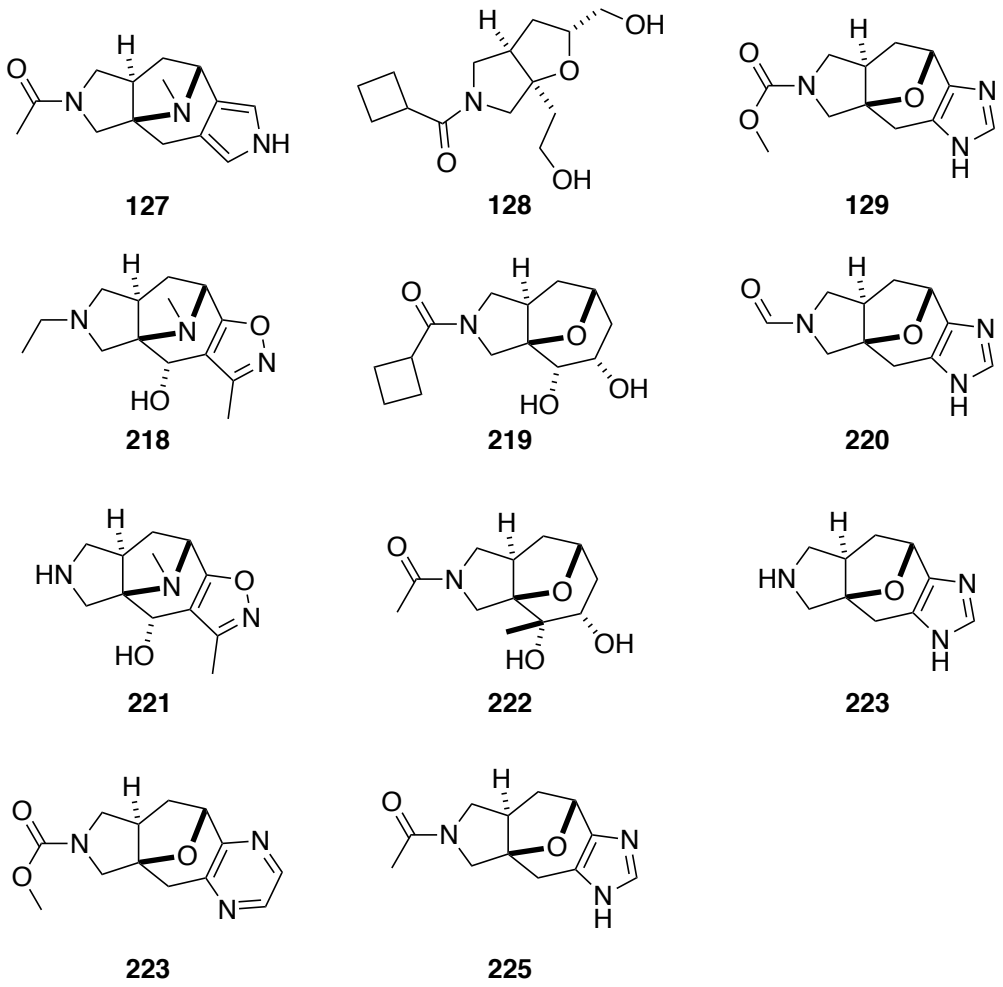


Figure B.90: Structures of ATAD2 binders screened against *Tc*BDF2 by WaterLOGSY (synthesised by Dr Dan Foley, University of Leeds).

B.3. Ligand-observed proton NMR titration

B.3.1. Screening against BRD4(1)

226, **227**, and **228**, are potent CREBBP ligands developed in the group (Dr Michael Brand, Dr James Clayton: unpublished data). These ligands were thought to weakly bind BRD4(1), and so their binding was assessed by ligand-observed proton NMR titration to determine their selectivity and as a proof of concept for this technique. Titration of BRD4(1) into fixed amounts of ligands gave rise to increased signal broadening until full disappearance of the ligand proton resonances (Figure B.91). The rate of the change in the integral over the titration was analysed with Prism to determine the K_D of **226**, **227**, and **228** to BRD4(1) (Table B.4). These experiments show that micromolar range K_D values can be obtained through proton ligand observed protein titrations experiments and provided useful selectivity data for the studied CREBBP ligands.

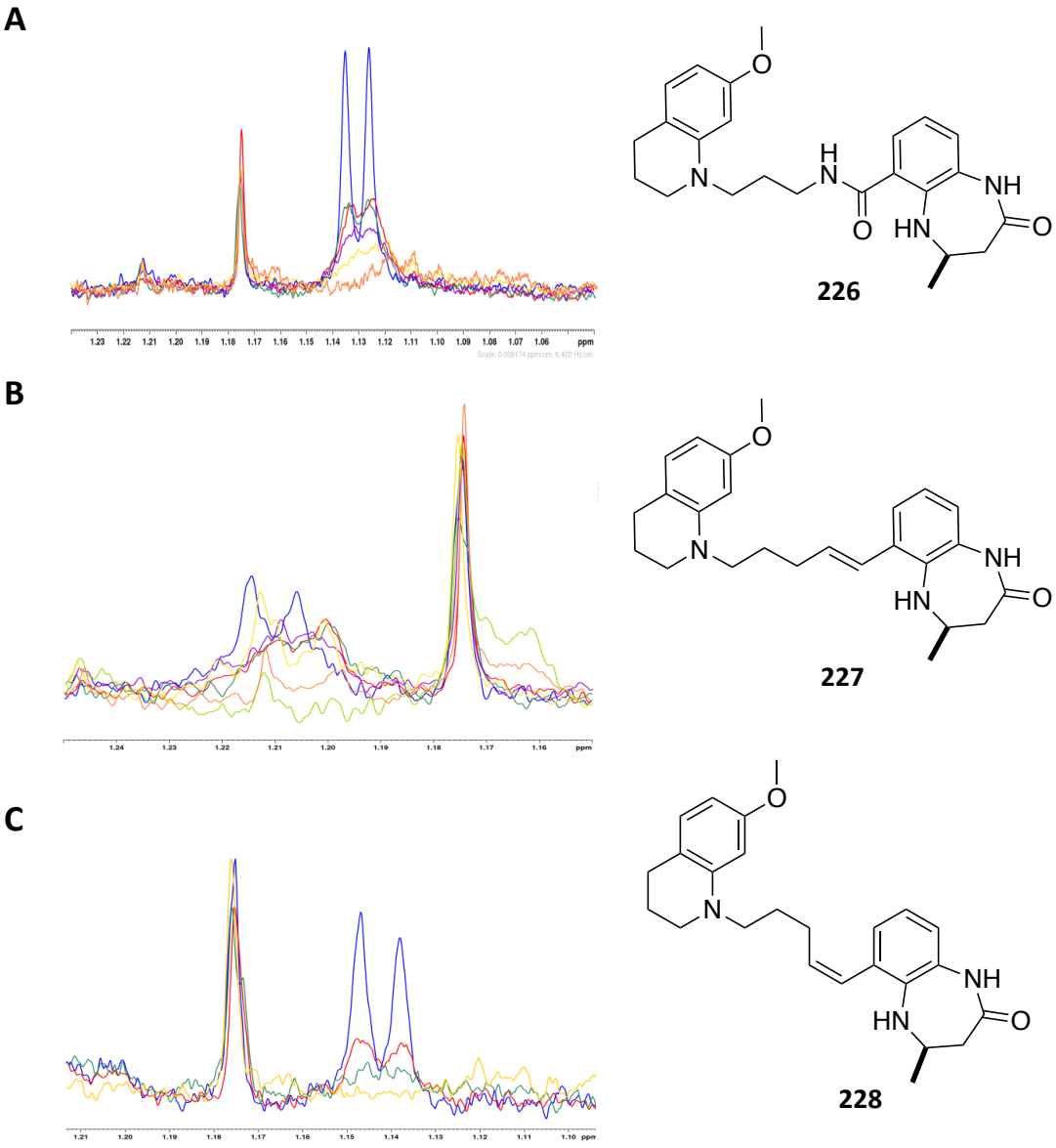
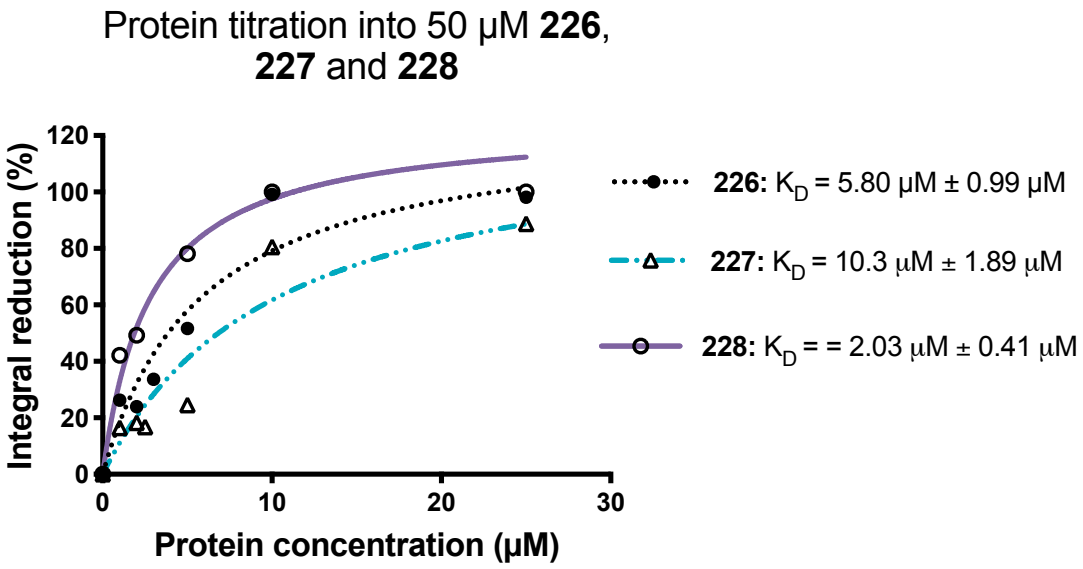


Figure B.91: ^{1}H NMR ligand observed protein titration of CREBBP ligands **226**, **227**, and **228** against BRD4(1). Thee ^{1}H NMR spectra stack show increasing broadening until disappeared upon addition of the protein. (A). Overlay of spectra of the free ligand **226** (blue), and the ligand in the presence of BRD4(1) at 1 µM (red), 2 µM (green), 3 µM (purple), 5 µM (yellow) and 10 µM (orange). (B). Overlay of spectra of the free ligand **227** (blue), and the ligand in the presence of BRD4(1) at 1 µM (red), 2 µM (green), 2.5 µM (purple), 5 µM (yellow), 10 µM (orange) and 25 µM (light green). (C). Overlay of spectra of the free ligand **228** (blue), and the ligand in the presence of BRD4(1) at 1 µM (red), 2 µM (green) and 5 µM (yellow). The ligand concentration is 50 µM in all cases.

Table B.4: Summary of the K_D values of **226**, **227**, and **228** to BRD4(1) determined by ^{1}H NMR ligand observed protein titration.

	226	227	228
K_D	5.8±1 µM	10±2 µM	2.8±0.4 µM
R^2	0.924	0.901	0.966



	226	227	228
One site -- Specific binding			
Best-fit values			
Bmax	= 125.0	= 125.0	= 125.0
Kd	5.796	10.25	2.811
Std. Error			
Kd	0.9869	1.983	0.3735
95% CI (asymptotic)			
Kd	3.381 to 8.211	5.399 to 15.10	1.851 to 3.771
Goodness of Fit			
Degrees of Freedom	6	6	5
R squared	0.9244	0.9014	0.9658
Sum of Squares	657.1	714.3	258.2
Sy.x	10.46	10.91	7.185
Constraints			
Bmax	Bmax = 125	Bmax = 125	Bmax = 125
Number of points			
# of X values	8	8	8

Figure B.92: Ligand-observed proton NMR titration analysis for the titration of 0–25 µM BRD4(1) into 50 µM **226**, **227**, and **228** using Prism v8. The non-linear data fit is shown with the ligand concentration set either as a variable or as a fixed value.

B.3.2. *Tc*BDF5-1 and **OXFBD02** by NMR titration

H_6-*Tc*BDF5-1 was titrated into fixed amounts of **OXFBD02** to obtain the equilibrium constant K_D of the binding. Chemical shift changes and peak broadening of the ligand signals were observed in the ^1H spectra, whereas impurities' peaks stayed constant. The observed change of chemical shift $\Delta\partial_{obs}$ appeared to be indicative of binding, and offered a reliable way to measure the binding affinity.

Interestingly, the tractable OXFBD02 (**7**) peaks appeared as split into two peaks during the titration (Figure B.93). This has previously been observed when titrating OXFBD02 (**7**) with *Tc*BDF3, however, the effect was less significant. The two new peaks have equal intensity, and OXFBD02 (**7**) is a racemate, therefore, it was hypothesised that the each peak corresponded to one of the two enantiomers. To confirm this, (*R*)- and (*S*)-(**7**)were separated by chiral preparative HPLC (GSK Stevenage). Comparison of the racemate and each enantiomer in the presence of the protein showed that each split corresponded to one enantiomer (Figure B.93).

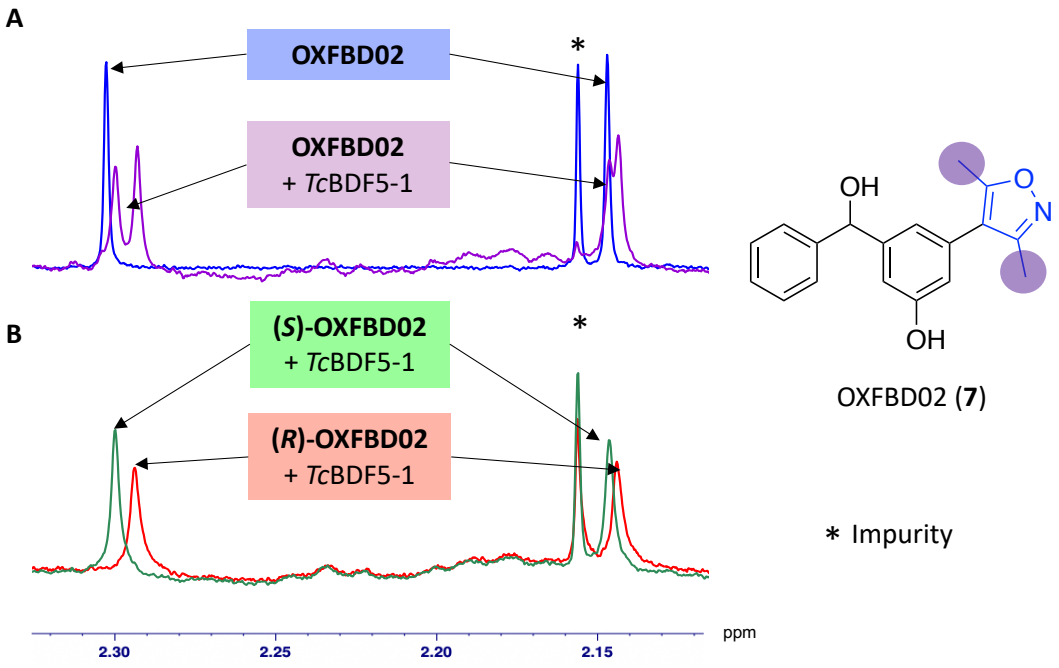
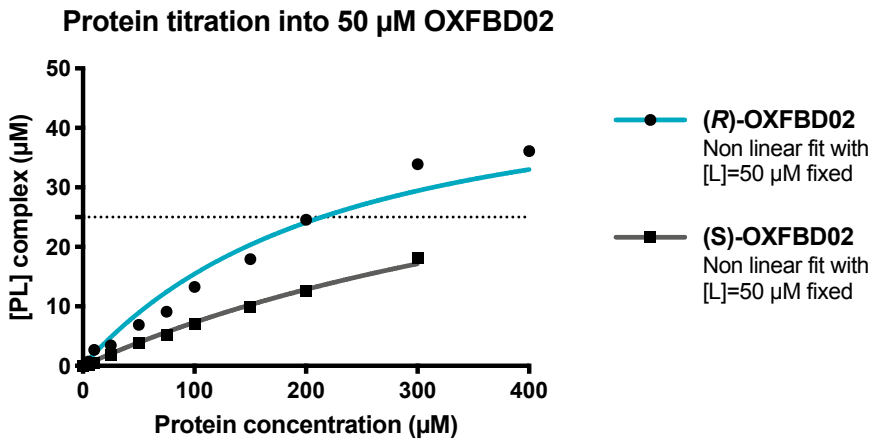


Figure B.93: Overlay of OXFBD02 (**7**), and each of its enantiomers, in the presence or absence of *Tc*BDF5-1. (A) **OXFBD02** (racemate) compound (50 µM) (blue), overlaid with OXFBD02 (**7**) (50 µM) with *Tc*BDF5-1 (50 µM) (purple). (B) (*R*)-OXFBD02 (*R*)-(**7**) (red), (*S*)-OXFBD02 (*S*)-(**7**) (green) with *Tc*BDF5-1 (50 µM).

This observation suggests that the two enantiomers of OXFBD02 (**7**) experience differences when bound to the protein. The total change in chemical shift is linked to the environment of the bound state. The rate at which the resonances reach their maximum change can be analysed to determine the binding constant. Therefore the two peaks per resonance resulting from the titrations were analysed separately to determine whether if distinguishable K_D values were obtain for each enantiomer (Table B.5). In the case of *Tc*BDF3, no significant differences in affinity could be measured, however; the pronounced differences in maximum chemical shift change was also accompanied by a difference in K_D values for *Tc*BDF5-1. The (*R*)- enantiomer of OXFBD02 (**7**) appears to be two-fold tighter than the (*S*)- enantiomer, with K_D values ranging from 189 to 476 µM and 541 to 1175 µM, respectively.

Table B.5: Summary tables of the K_D values determined using Prism, with the total ligand concentration set as a fixed value or a variable. Titration of *Tc*BDF5-1 into 50 µM of OXFBD02 (**7**). The results for TcBDF3, as described in Chapter 3, are given for comparison purposes.

	OXFBD02 (**7**) *vs* H_6-*Tc*BDF5-1 Tracking the (*R*)-enantiomer	OXFBD02 (**7**) *vs* H_6-*Tc*BDF5-1 Tracking the (*S*)-enantiomer	OXFBD02 (**7**) *vs* H_6-*Tc*BDF3
Non-linear fit with $[L]_t = 50\ \mu M$ kept constant.	189 µM	541 µM	455–508 µM (n=2)
Non-linear fit with $[L]_t$ as a variable.	476 µM	1175 µM	418–445 µM (n=2)
Non-linear fit with $[L]_t = 25\ \mu M$ kept constant.	189 µM	541 µM	455–508 µM (n=2)
Non-linear fit with $[L]_t$ as a variable.	476 µM	1175 µM	418–445 µM (n=2)



	(*R*)-OXFBD02	(S)-OXFBD02
NMR Titration (quadratic)		
Best-fit values		
Lt	= 50.00	= 50.00
Kd	189.1	541.1
Std. Error		
Kd	17.00	13.09
95% CI (asymptotic)		
Kd	151.2 to 226.9	511.5 to 570.7
Goodness of Fit		
Degrees of Freedom	10	9
R squared	0.9656	0.9947
Sum of Squares	58.30	1.742
Sy.x	2.415	0.4399
Constraints		
Lt	Lt = 50	Lt = 50
Number of points		
# of X values	11	10
# Y values analyzed	11	10

	(*R*)-OXFBD02	(S)-OXFBD02
NMR Titration (quadratic)		
Best-fit values		
Lt	87.44	93.20
Kd	476.4	1175
Std. Error		
Lt	12.96	11.89
Kd	106.2	176.0
95% CI (asymptotic)		
Lt	58.12 to 116.8	65.78 to 120.6
Kd	236.1 to 716.7	768.7 to 1580
Goodness of Fit		
Degrees of Freedom	9	8
R squared	0.9920	0.9992
Sum of Squares	13.62	0.2505
Sy.x	1.230	0.1769
Number of points		
# of X values	11	10
# Y values analyzed	11	10

Figure B.94: Ligand-observed proton NMR titration analysis for the titration of 0–400 µM *Tc*BDF5-1 into 50 µM OXFBD02 (**7**) (racemic) ligand using Prism v8. The non-linear data fit is shown with the ligand concentration set either as a variable or as a fixed value. The ligand peaks split into two equal intensity peaks that were analysed separately to determine the K_D of each enantiomer.

B.3.3. *Tc*BDF3 NMR titration analysis

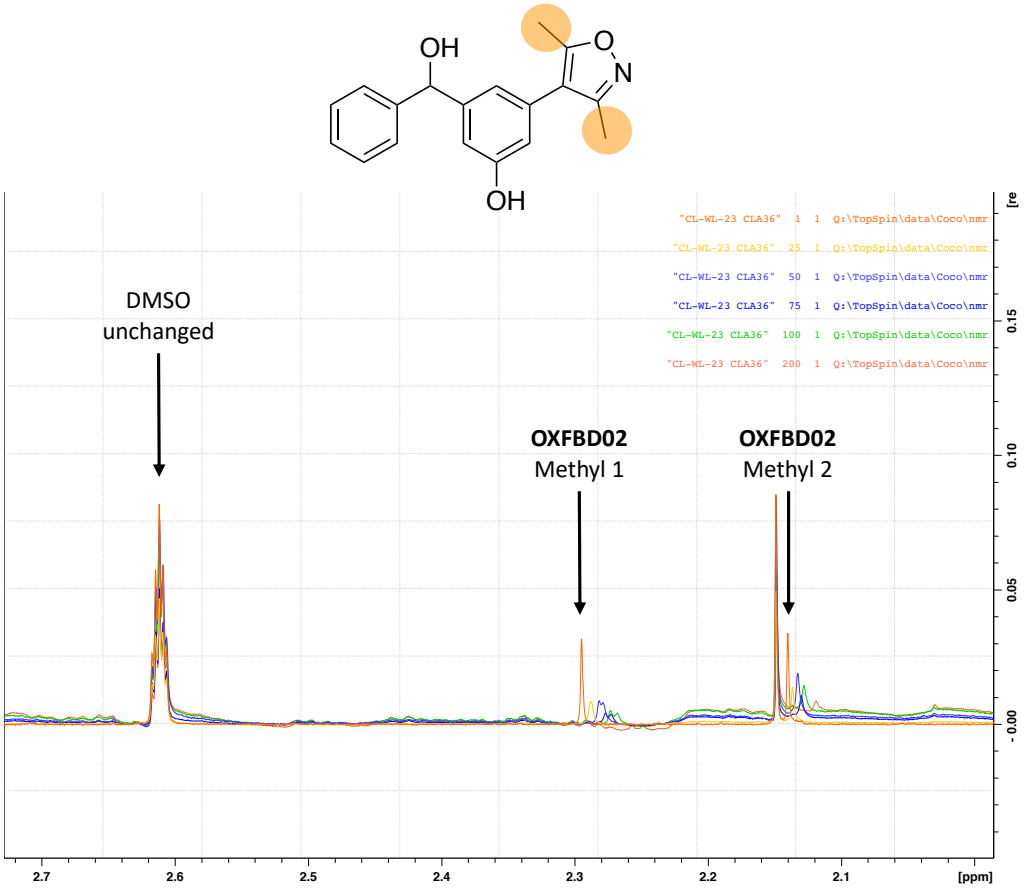
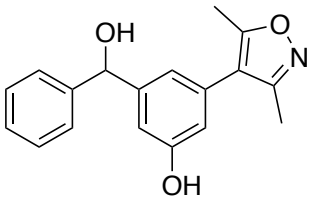
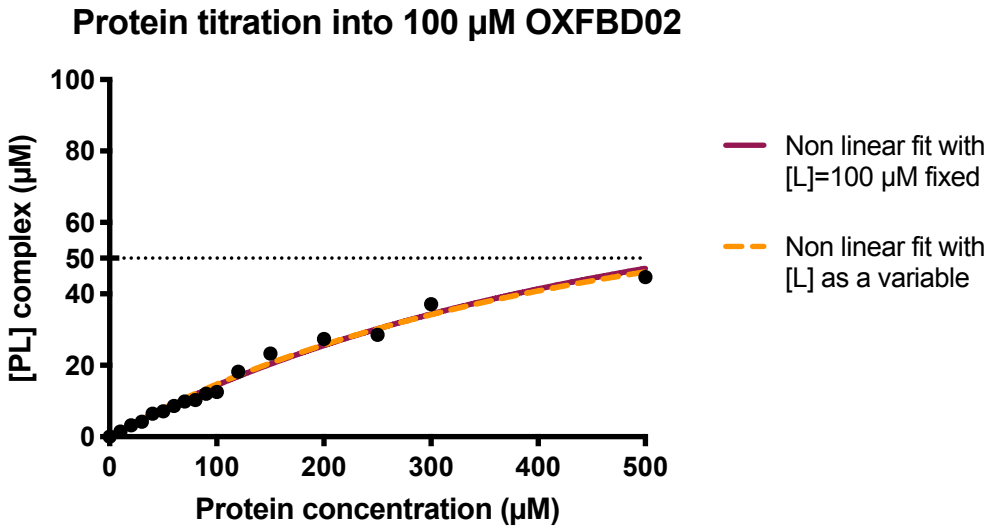


Figure B.95: Partial spectra overlay for the ligand-observed proton NMR of 0–500 µM *Tc*BDF3 into 100 µM OXFBD02 (**7**) ligand.



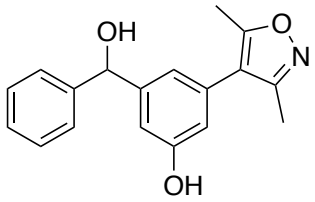
OXFBD02



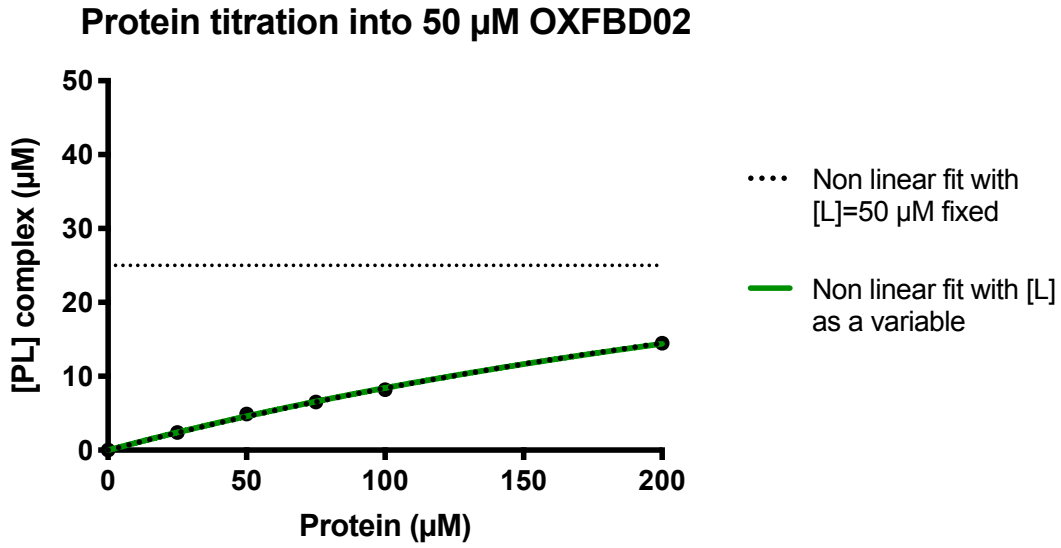
	∂obs
NMR Titration (quadratic) [3]	
Best-fit values	
Lt	= 100.0
Kd	507.9
Std. Error	
Kd	15.73
95% CI (asymptotic)	
Kd	474.5 to 541.2
Goodness of Fit	
Degrees of Freedom	16
R squared	0.9857
Sum of Squares	38.55
Sy.x	1.552
Constraints	
Lt	Lt = 100
Number of points	
# of X values	17

	∂obs
NMR Titration (quadratic) [3]	
Best-fit values	
Lt	91.34
Kd	444.4
Std. Error	
Lt	9.659
Kd	71.77
95% CI (asymptotic)	
Lt	70.76 to 111.9
Kd	291.4 to 597.4
Goodness of Fit	
Degrees of Freedom	15
R squared	0.9863
Sum of Squares	36.78
Sy.x	1.566
Number of points	
# of X values	17

Figure B.96: Ligand-observed proton NMR titration analysis for the titration of 0–500 µM *Tc*BDF3 into 100 µM **OXFBD02** ligand using Prism v8. The non-linear data fit is shown with the ligand concentration set either as a variable or as a fixed value.



OXFBD02



	[PL] (µM)
NMR Titration (quadratic)	
Best-fit values	
Lt	= 50.00
Kd	455.1
Std. Error	
Kd	6.319
95% CI (asymptotic)	
Kd	438.9 to 471.4
Goodness of Fit	
Degrees of Freedom	5
R squared	0.9986
Sum of Squares	0.1764
Sy.x	0.1879
Constraints	
Lt	Lt = 50
Number of points	
# of X values	6

	[PL] (µM)
NMR Titration (quadratic)	
Best-fit values	
Lt	46.89
Kd	418.7
Std. Error	
Lt	5.152
Kd	60.65
95% CI (asymptotic)	
Lt	32.59 to 61.20
Kd	250.3 to 587.1
Goodness of Fit	
Degrees of Freedom	4
R squared	0.9987
Sum of Squares	0.1637
Sy.x	0.2023
Number of points	
# of X values	6
# Y values analyzed	6

Figure B.97: Ligand-observed proton NMR titration analysis for the titration of 0–200 µM *Tc*BDF3 into 50 µM **OXFBD02** ligand using Prism v8. The non-linear data fit is shown with the ligand concentration set either as a variable or as a fixed value.

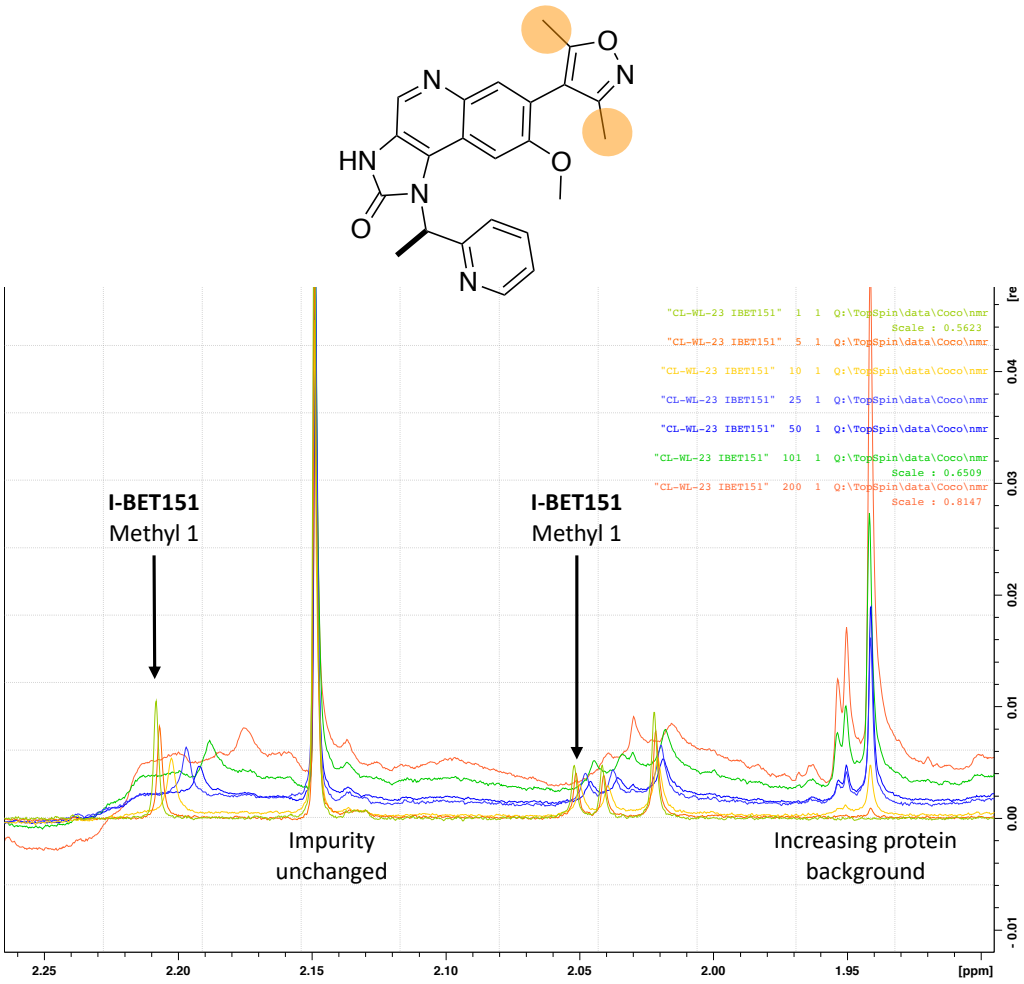
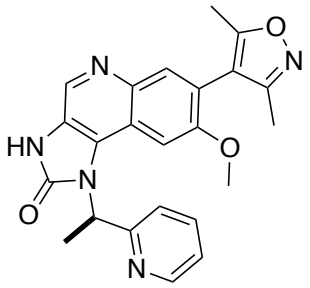
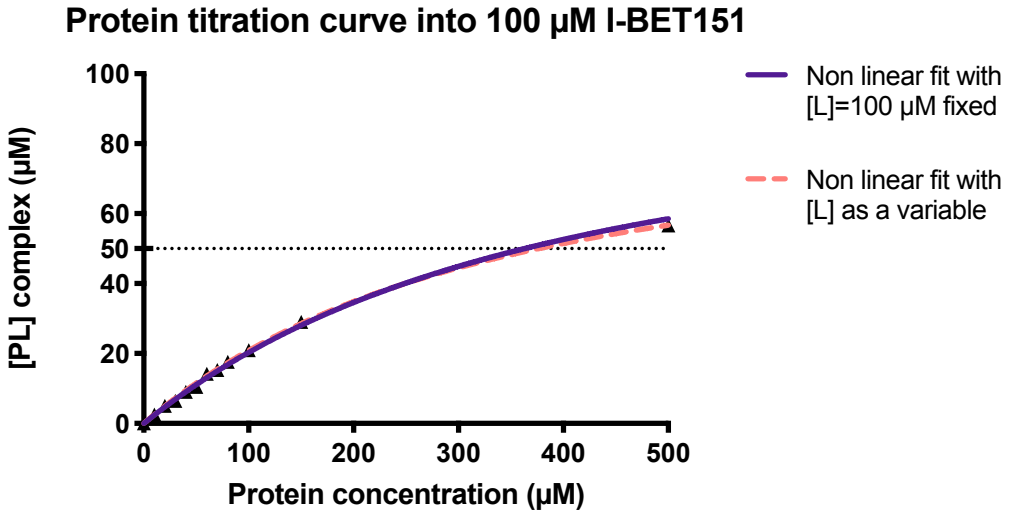


Figure B.98: Partial spectra overlay for the ligand-observed proton NMR of 0–500 µM *Tc*BDF3 into 100 µM I-BET151 (**5**) ligand.



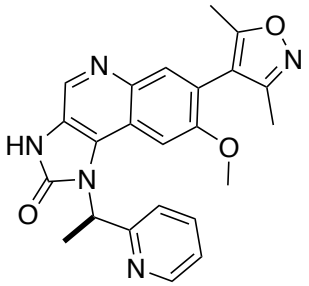
I-BET151



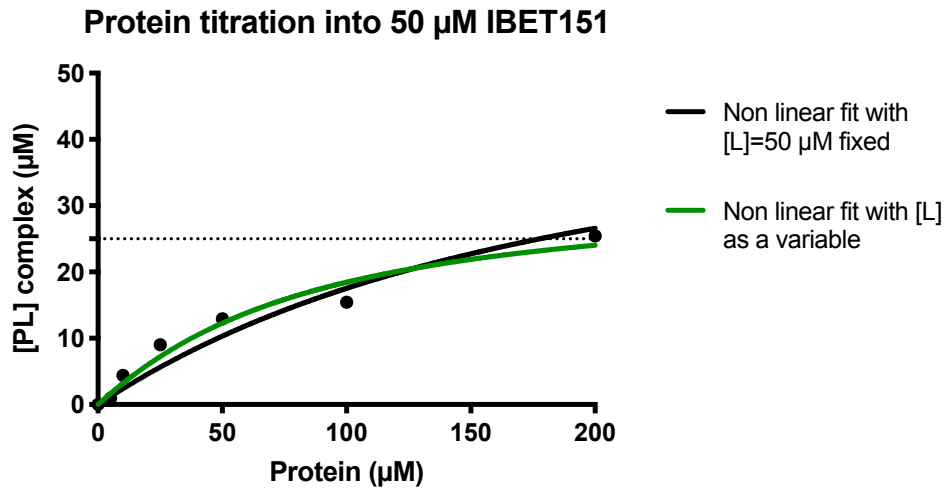
	∂obs
NMR Titration (quadratic) [3]	
Best-fit values	
Lt	= 100.0
Kd	312.9
Std. Error	
Kd	6.948
95% CI (asymptotic)	
Kd	297.6 to 328.2
Goodness of Fit	
Degrees of Freedom	11
R squared	0.9969
Sum of Squares	8.002
Sy.x	0.8529
Constraints	
Lt	Lt = 100
Number of points	
# of X values	12

	∂obs
NMR Titration (quadratic) [3]	
Best-fit values	
Lt	90.39
Kd	263.3
Std. Error	
Lt	1.965
Kd	10.48
95% CI (asymptotic)	
Lt	86.01 to 94.77
Kd	240.0 to 286.7
Goodness of Fit	
Degrees of Freedom	10
R squared	0.9990
Sum of Squares	2.677
Sy.x	0.5174
Number of points	
# of X values	12

Figure B.99: Ligand-observed proton NMR titration analysis for the titration of 0–500 µM *Tc*BDF3 into 100 µM I-BET151 (**5**) ligand using Prism v8. The non-linear data fit is shown with the ligand concentration set either as a variable or as a fixed value.



I-BET151



	[PL] (µM)
NMR Titration (quadratic) [3]	
Best-fit values	
Lt	= 50.00
Kd	152.6
Std. Error	
Kd	19.13
95% CI (asymptotic)	
Kd	105.8 to 199.4
Goodness of Fit	
Degrees of Freedom	6
R squared	0.9421
Sum of Squares	28.34
Sy.x	2.173
Constraints	
Lt	Lt = 50
Number of points	
# of X values	7

	[PL] (µM)
NMR Titration (quadratic) [3]	
Best-fit values	
Lt	32.54
Kd	62.22
Std. Error	
Lt	5.174
Kd	24.76
95% CI (asymptotic)	
Lt	19.24 to 45.84
Kd	-1.439 to 125.9
Goodness of Fit	
Degrees of Freedom	5
R squared	0.9660
Sum of Squares	16.64
Sy.x	1.824
Number of points	
# of X values	7

Figure B.100: Ligand-observed proton NMR titration analysis for the titration of 0–200 µM *Tc*BDF3 into 50 µM I-BET151 (**5**) ligand using Prism v8. The non-linear data fit is shown with the ligand concentration set either as a variable or as a fixed value.

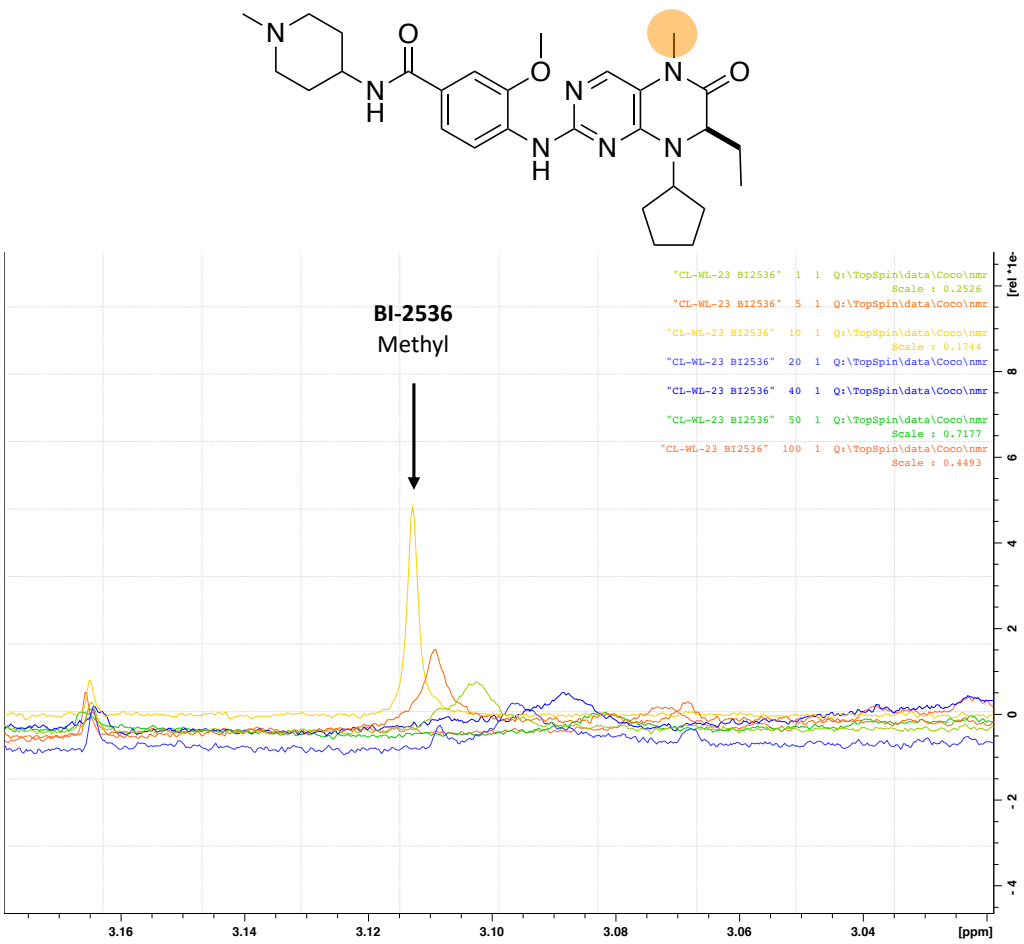
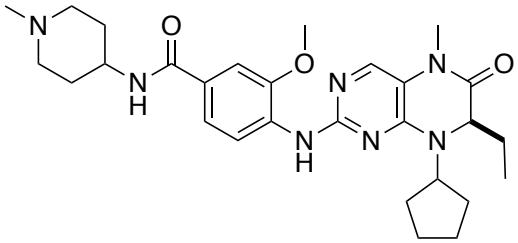
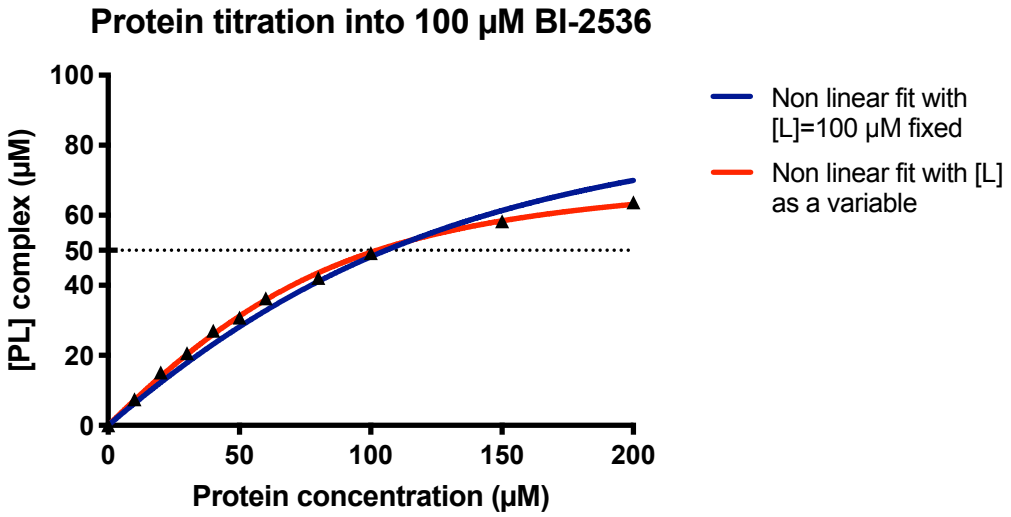


Figure B.101: Partial spectra overlay for the ligand-observed proton NMR of 0–200 µM *Tc*BDF3 into 100 µM BI-2536 (**8**) ligand.



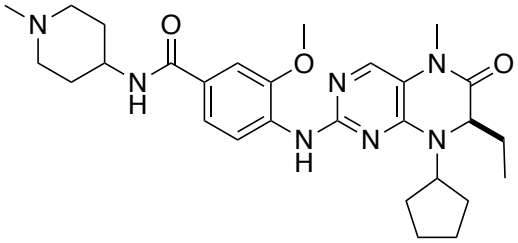
BI-2536



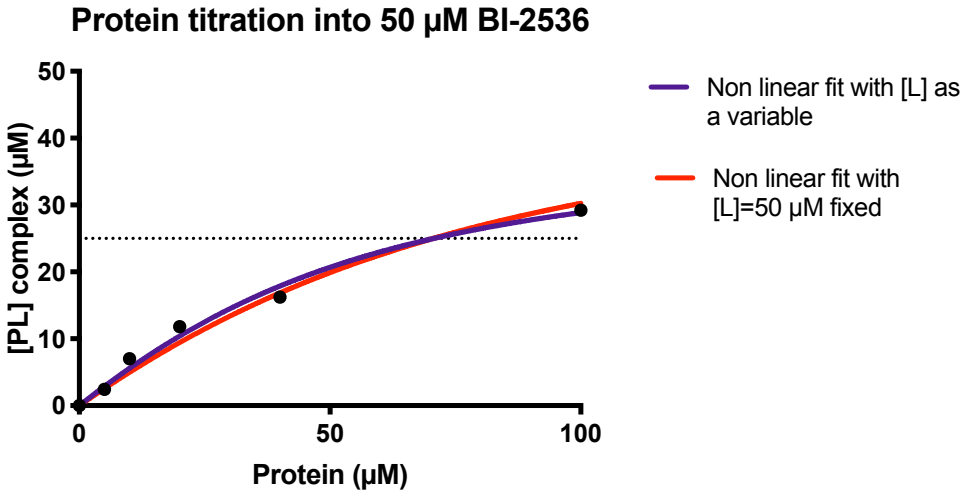
	∂obs
NMR Titration (quadratic) [3]	
Best-fit values	
Lt	= 100.0
Kd	55.81
Std. Error	
Kd	4.501
95% CI (asymptotic)	
Kd	45.79 to 65.84
Goodness of Fit	
Degrees of Freedom	10
R squared	0.9753
Sum of Squares	103.2
Sy.x	3.212
Constraints	
Lt	Lt = 100
Number of points	
# of X values	11

	∂obs
NMR Titration (quadratic) [3]	
Best-fit values	
Lt	75.42
Kd	26.64
Std. Error	
Lt	1.381
Kd	1.604
95% CI (asymptotic)	
Lt	72.29 to 78.54
Kd	23.02 to 30.27
Goodness of Fit	
Degrees of Freedom	9
R squared	0.9987
Sum of Squares	5.420
Sy.x	0.7760
Number of points	
# of X values	11

Figure B.102: Ligand-observed proton NMR titration analysis for the titration of 0–500 µM *Tc*BDF3 into 100 µM **BI-2536** ligand using Prism v8. The non-linear data fit is shown with the ligand concentration set either as a variable or as a fixed value.



BI-2536

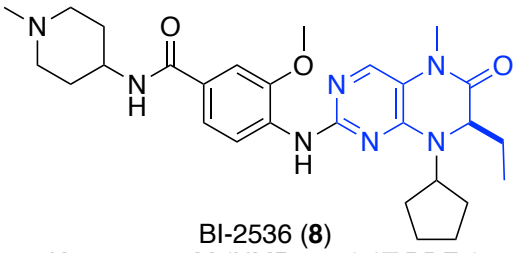


	[PL] (µM)
NMR Titration (quadratic)	
Best-fit values	
Lt	= 50.00
Kd	45.54
Std. Error	
Kd	5.023
95% CI (asymptotic)	
Kd	32.62 to 58.45
Goodness of Fit	
Degrees of Freedom	5
R squared	0.9800
Sum of Squares	11.42
Sy.x	1.511
Constraints	
Lt	Lt = 50
Number of points	
# of X values	6

	[PL] (µM)
NMR Titration (quadratic)	
Best-fit values	
Lt	39.96
Kd	27.33
Std. Error	
Lt	4.882
Kd	8.852
95% CI (asymptotic)	
Lt	26.41 to 53.52
Kd	2.750 to 51.91
Goodness of Fit	
Degrees of Freedom	4
R squared	0.9873
Sum of Squares	7.252
Sy.x	1.346
Number of points	
# of X values	6

Figure B.103: Ligand-observed proton NMR titration analysis for the titration of 0–500 µM *Tc*BDF3 into 100 µM BI-2536 (**8**) ligand using Prism v8. The non-linear data fit is shown with the ligand concentration set either as a variable or as a fixed value.

B.4. Isothermal calorimetry titration curves



BI-2536 (**8**)
K_D=27–50 μM (NMR, n=2) (*Tc*BDF3)
K_D=14.6±3.0 μM (ITC, n=3) (*Tc*BDF3)

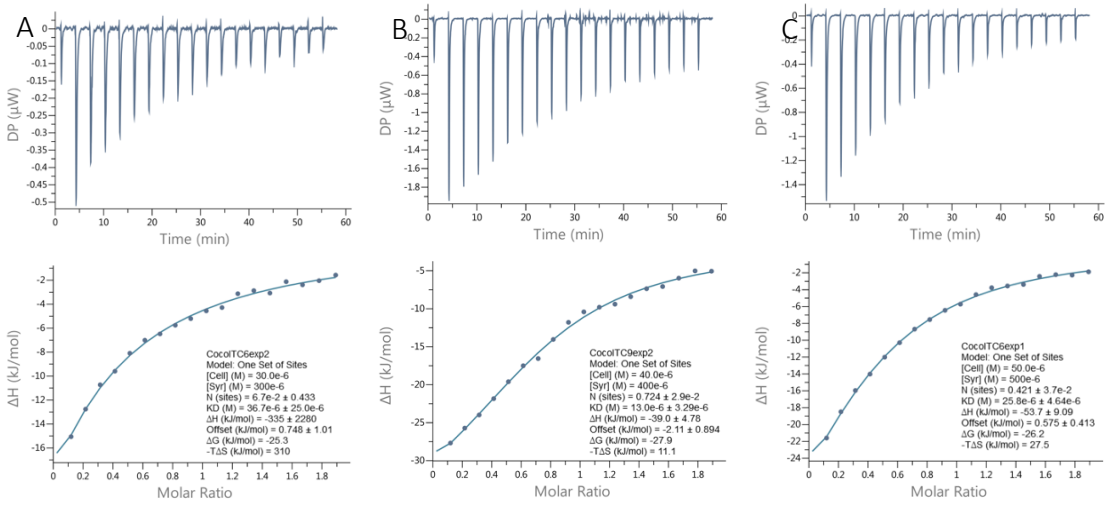


Figure B.104: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) with the dual PLK1/BRD4(1) ligand **BI-2536**: Condition A ligand:protein 30:300 µM, condition B ligand:protein 40:400 µM, condition C ligand:protein 50:3500 µM. These conditions show binding with K_D=36.7±25 µM, 13.0±3.29 µM and 25.8±4.64 µM µM, respectively.

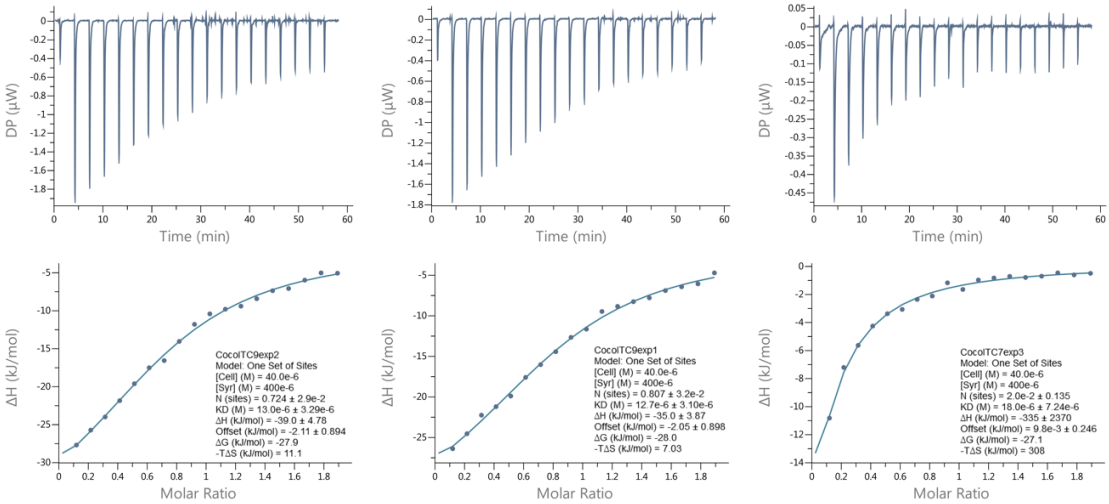
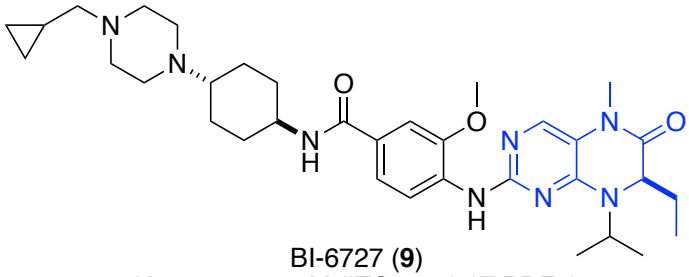


Figure B.105: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with the dual PLK1/BRD4(1) ligand BI-2536 (**8**) (40 µM), showing binding with K_D=12.7±3.10 µM, 13.0±3.29 µM and 18.0±7.24 µM respectively.



BI-6727 (**9**)
K_D=57.1±4.9 μM (ITC, n=3) (*Tc*BDF3)

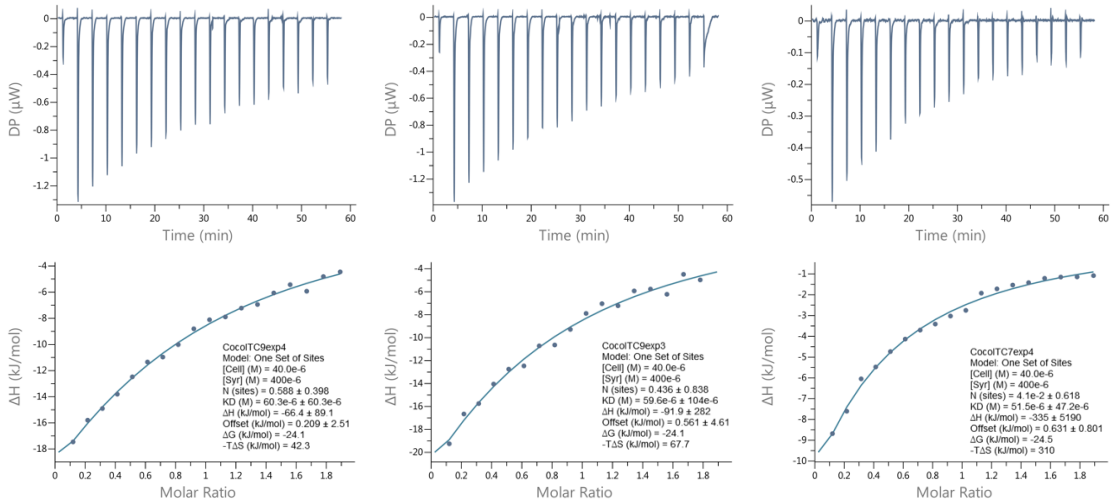
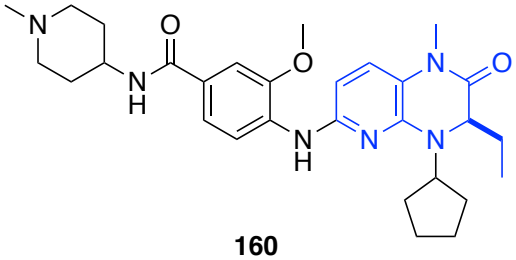


Figure B.106: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with the dual PLK1/BRD4(1) ligand BI-6727 (**9**) (40 µM), showing binding with K_D=59.6±104 µM, 60.3±60.3 µM and 51.5±47.2 µM respectively.

Compound	Compound concentration	Protein concentration	K_D (ITC)
BI-2536 (**8**)	40 µM	400 µM	12.7±3.10 µM
	40 µM	400 µM	13.0±3.29 µM
	40 µM	400 µM	18.0±7.24 µM
	50 µM	500 µM	25.8±4.64 µM
	30 µM	300 µM	36.7±25 µM
BI-6727 (**9**)	40 µM	400 µM	59.6±104 µM
	40 µM	400 µM	60.3±60.3 µM
	40 µM	400 µM	51.5±47.2 µM



160

K_D=18.6±8.4 μM (ITC, n=3) (*Tc*BDF3)

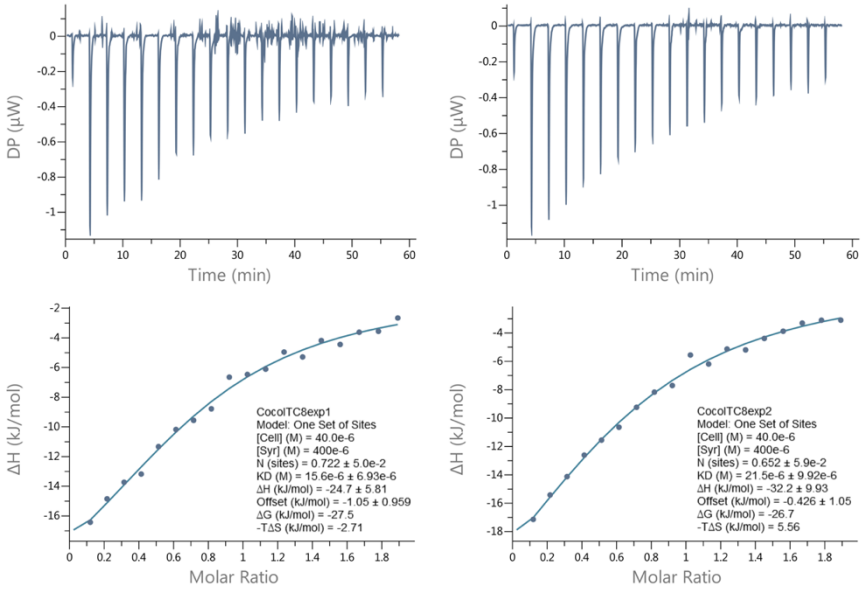


Figure B.107: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **160** (40 µM), showing binding with K_D=15.6±6.93 µM and 21.5±9.92 µM respectively. **160** was synthesised by Michelle Keller.

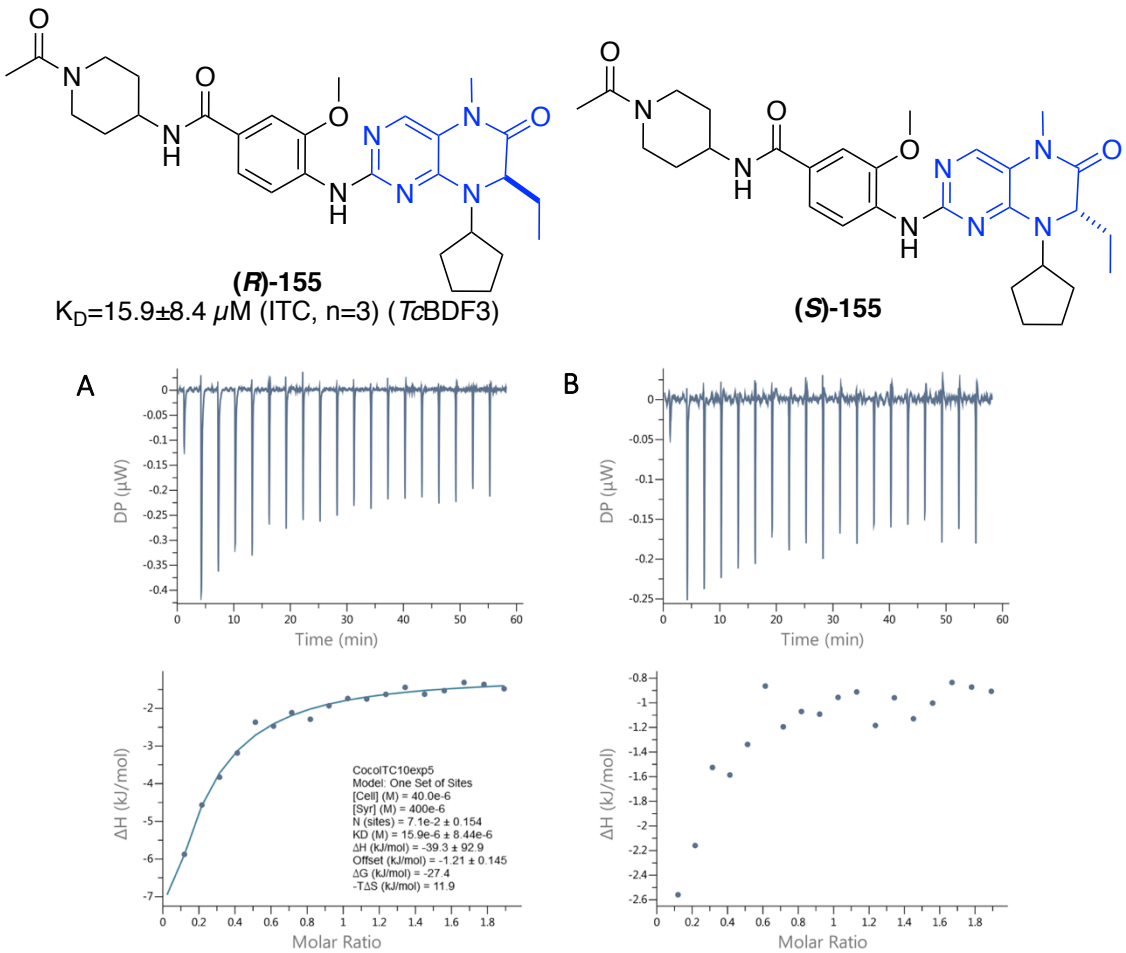


Figure B.108: (A) ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 μM) with **(*R*)-CLD43** (40 μM), showing binding with K_D=15.9±8.44 μM. (B) ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 μM) with **(*S*)-CLD43** (40 μM). No binding is detected.

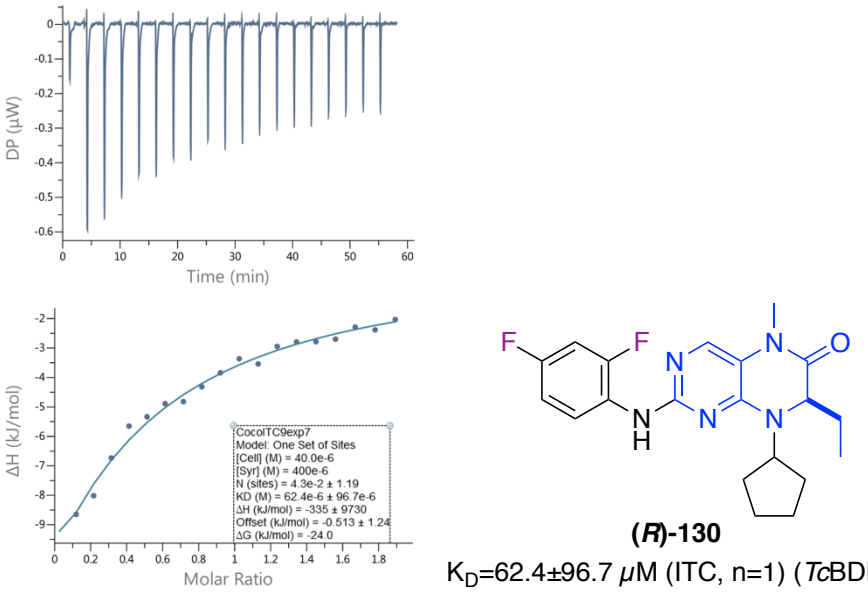


Figure B.109: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 μM) with **(*R*)-CLC65** (40 μM), showing binding with K_D=62.4±96.7 μM.

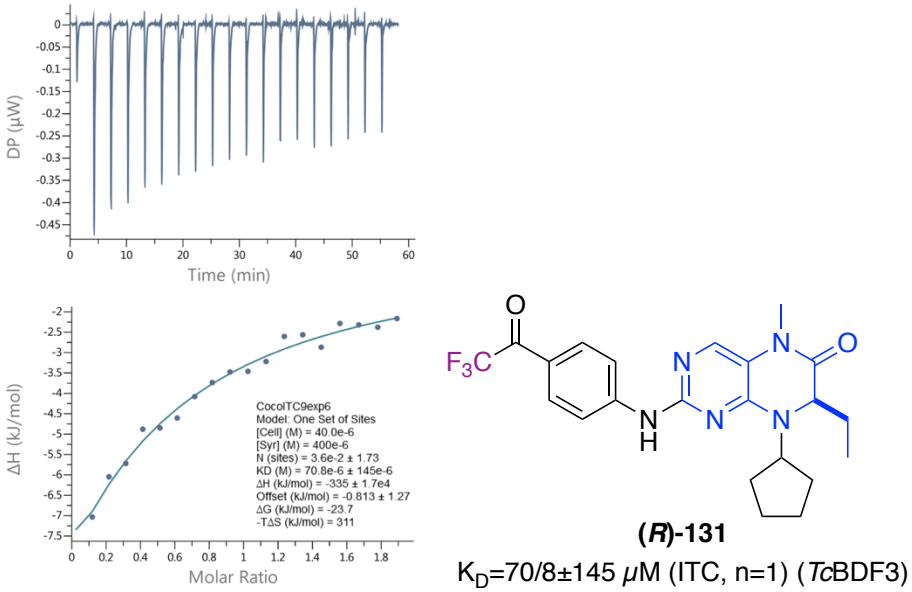


Figure B.110: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **(*R*)-CLD11** (40 µM), showing binding with K_D=70.8±145 µM.

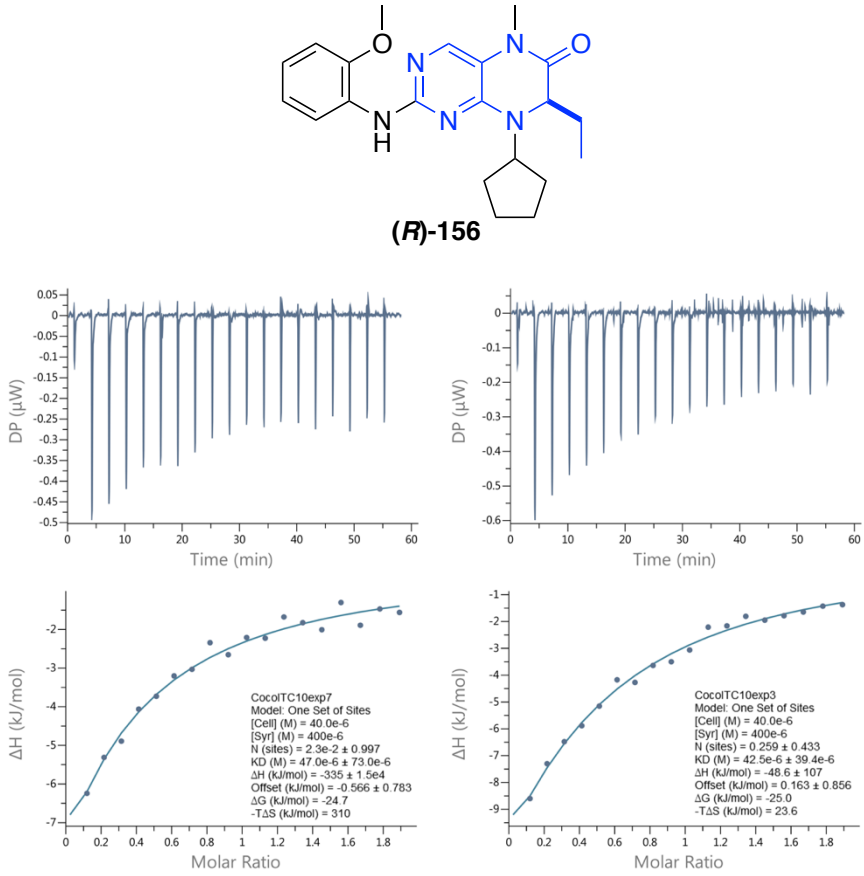


Figure B.111: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **(*R*)-CLC75** (40 µM), showing binding with K_D=47.0±73.0 µM and 42.5±39.4 µM respectively.

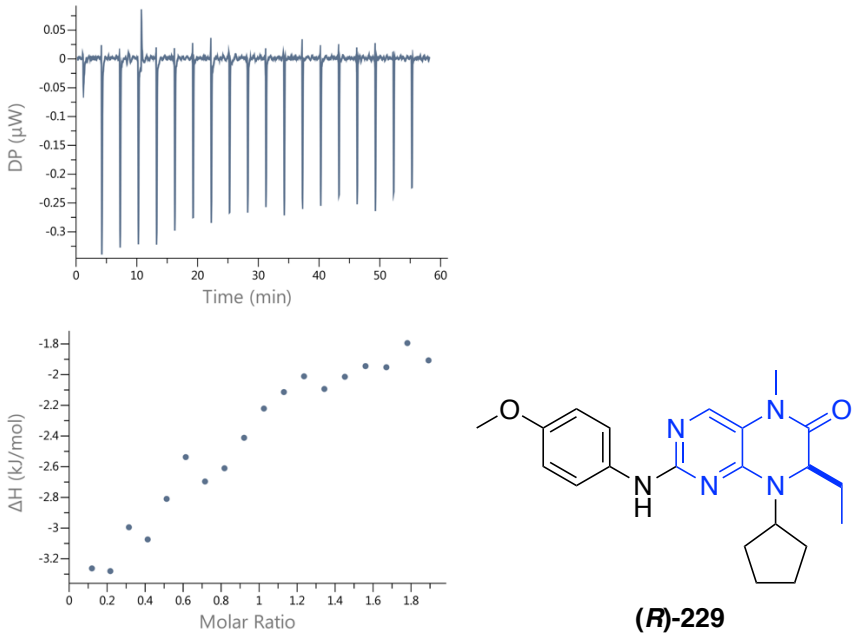


Figure B.112: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **(*R*)-229** (40 µM). No binding is detected..

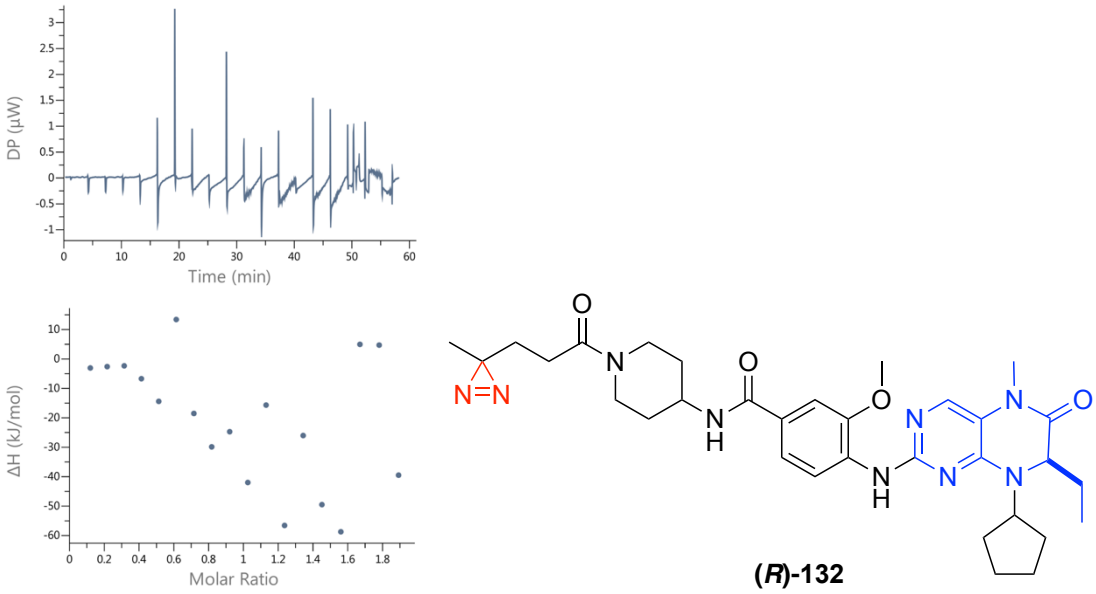


Figure B.113: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **(*S*)-CLN42** (40 µM). No binding is detected. The positive peaks are thought to correspond to air bubble and non-homogenous compound solution due to poor ligand solubility.

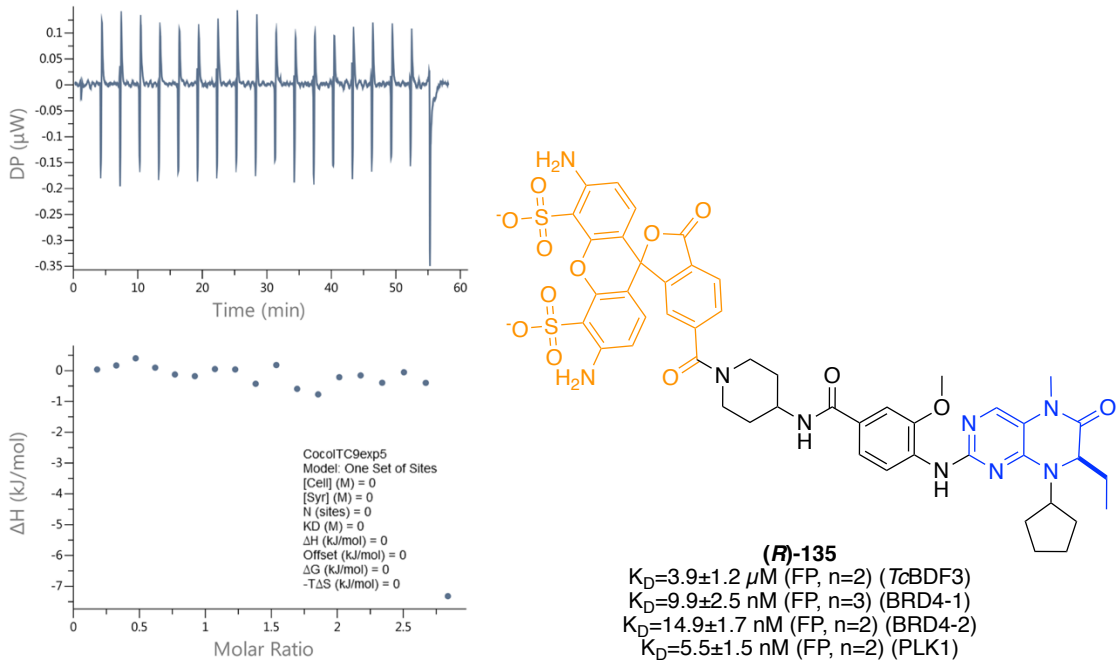


Figure B.114: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **(*R*)-135** (40 µM). No binding is detected. The lack of detection of the binding event, observed by FP have a K_D of 3 µM, is thought to arise from secondary binding from the fluorophore moiety.

ITC titration experiment was using a 300:30 µM of compound:protein ratios to further characterise the binding of **(*R*)-135** to *Tc*BDF3. None of the performed runs showed the expected binding titration curve. The binding affinity detected by FP, K_D=3.9±1.2 µM (FP, n=2), estimates the affinity to be below 5 µM, which should easily be detected under the employed conditions. Docking followed by MD simulation, performed by Joseph Bluck, showed that the hydrophobic fluorophore moiety of the FP probe **(*R*)-135** is predicted to bind to the protein surface. This additional binding event is thought to occur at similar rates to the specific binding of the ligand to the protein, through hydrogen bonding of the KAc mimic to Asp124. ITC detects binding events through the change of heat generated during the titration experiments, hence the simultaneous binding events of the two portions of the fluorescent probe is thought to be the reason why the binding of **(*R*)-135** cannot be detected by ITC.

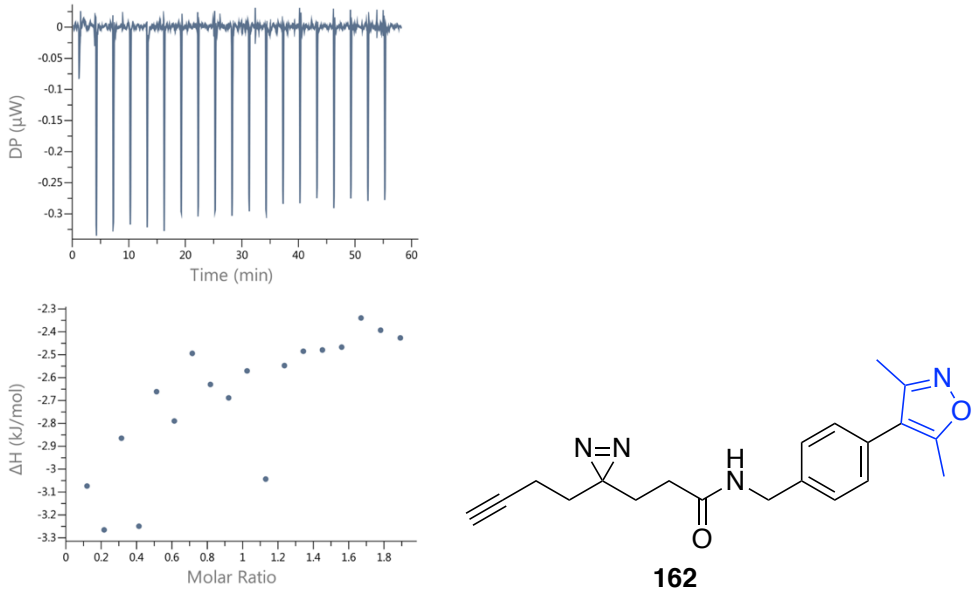


Figure B.115: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **162** (40 µM). No binding is detected.

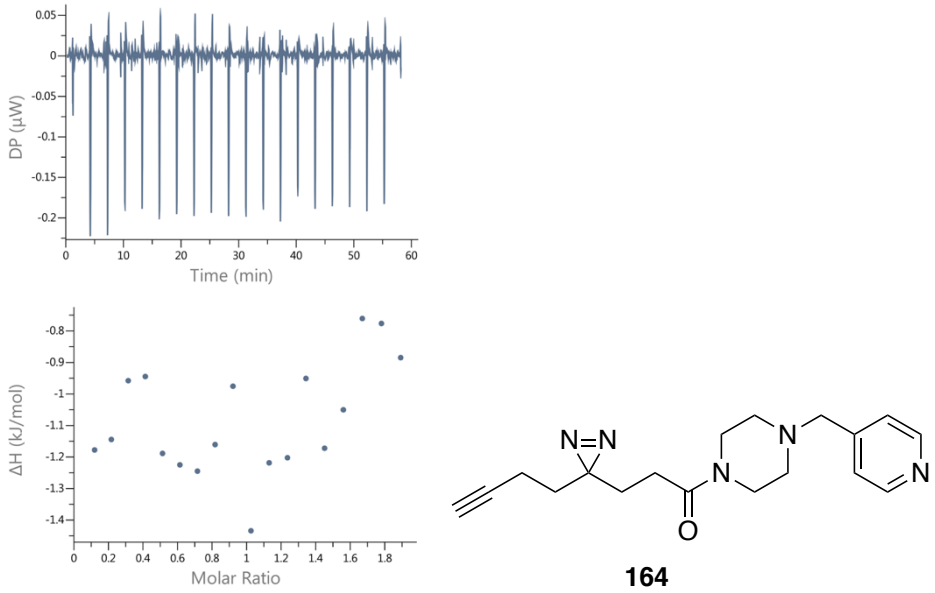
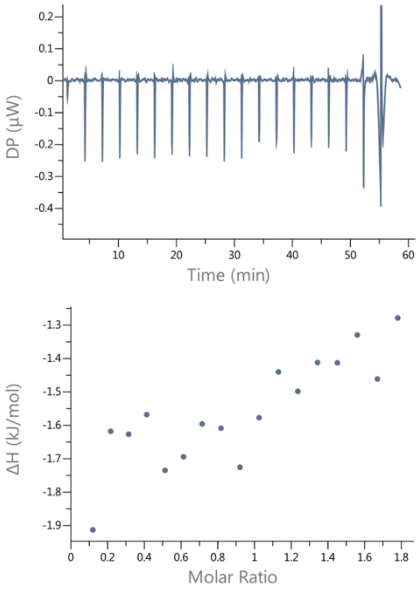
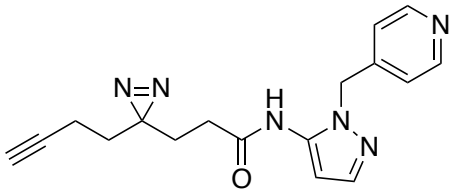


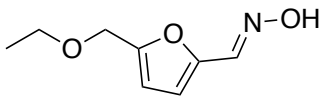
Figure B.116: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **164** (40 µM). No binding is detected.





168

Figure B.117: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) (400 µM) with **168** (40 µM). No binding is detected.



15

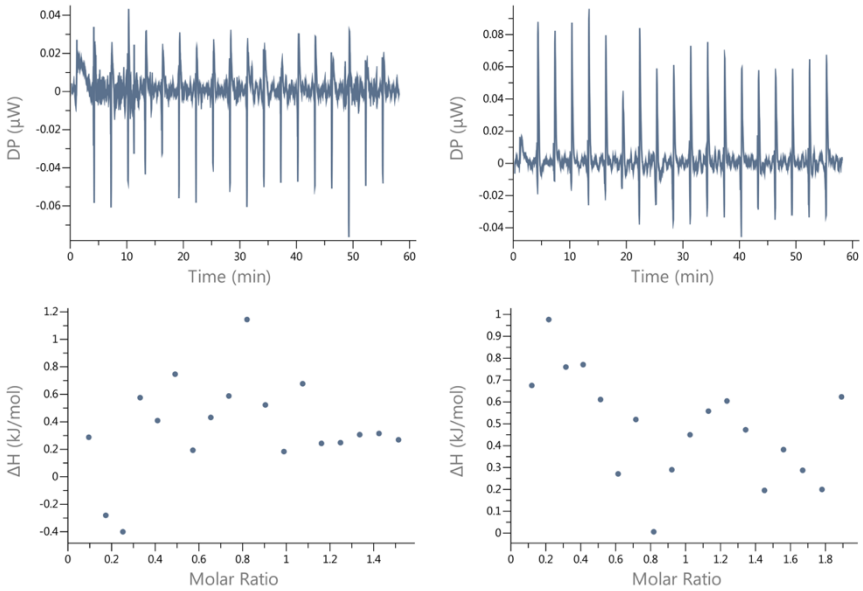


Figure B.118: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) with the proposed *Tc*BDF3 binder **15**. Condition A ligand:protein 20:160 µM, condition B ligand:protein 50:500 µM. No binding is detected for either experimental conditions.

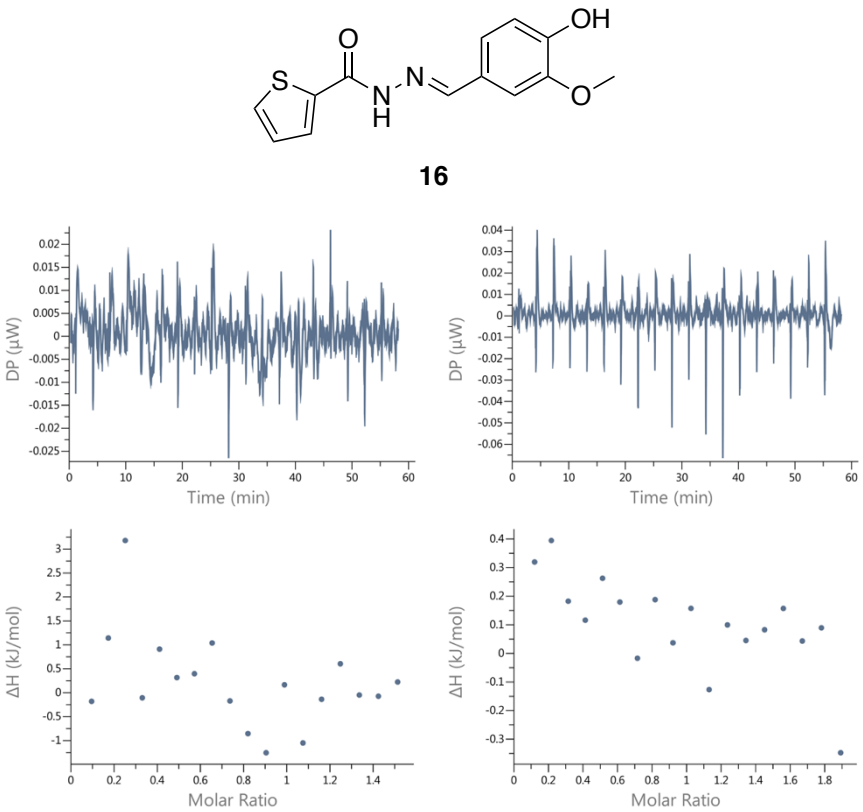


Figure B.119: ITC traces for the titration of H_6-*Tc*BDF3(32..148aa) with the proposed *Tc*BDF3 binder **16**. Condition A ligand:protein 20:160 µM, condition B ligand:protein 50:500 µM. No binding is detected for either experimental conditions.

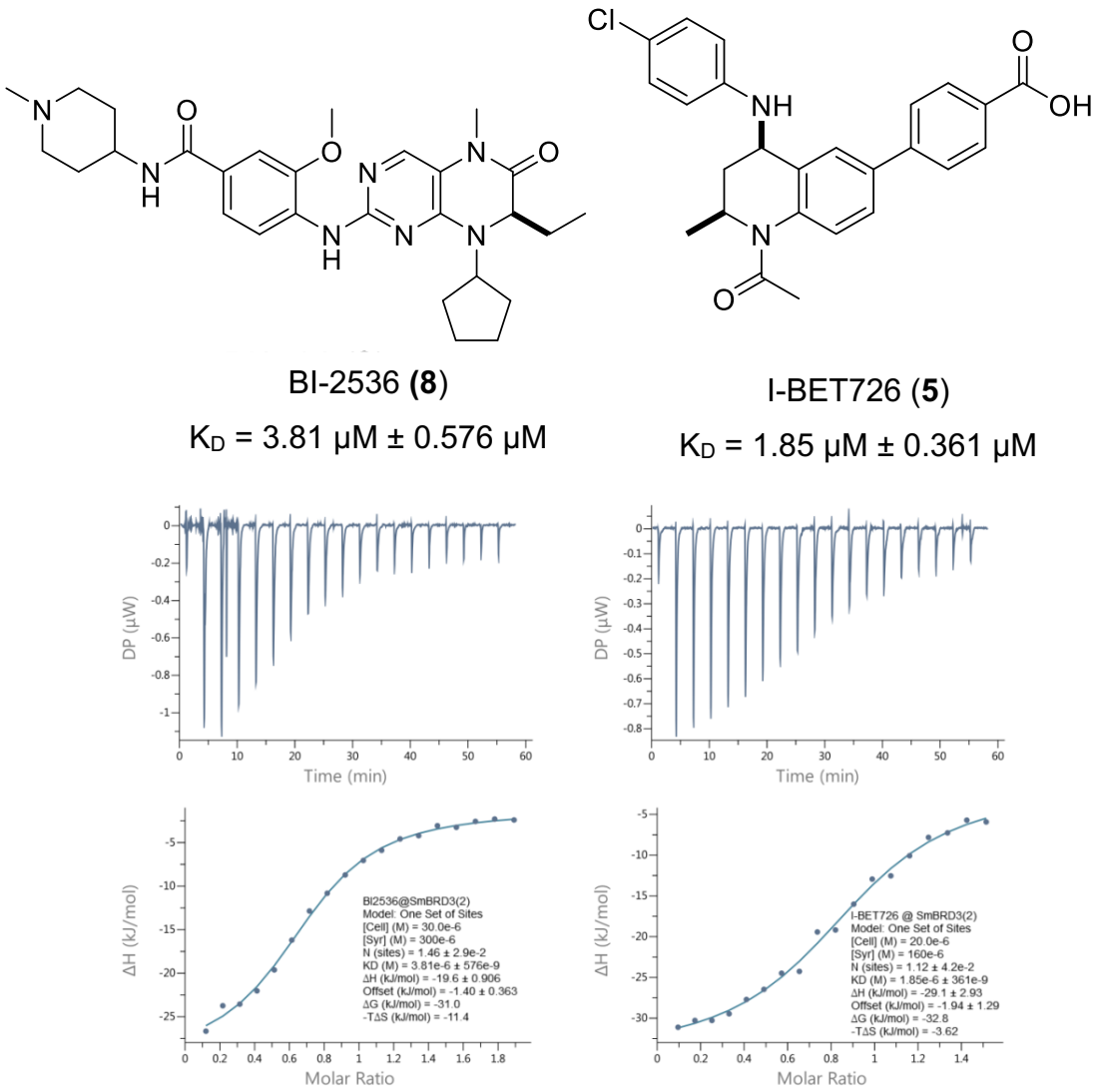


Figure B.120: ITC traces for the titration of H_6-*Sm*BRD3(2) with (A) BI-2536 (**8**) using a ligand:protein 30:300 µM and (B) I-BET726 (**5**) using a ligand:protein 20:160 µM. Experiment performed by Dr Matthias Schiedel.

B.5. ^{19}F NMR assays

B.5.1. Fluorine reference optimisation

The initial work on **112**, done in collaboration with Dr Anthony Chan (University of Oxford), used trifluoroacetone (TFAc) as a ^{19}F NMR reference. This solvent is very volatile, so it cannot be used as a quantitave reference. Additionally, TFAc has three resonnances at –86.33, –86.79 and –84.13 ppm which overlap with **(*R*)-131** trifluoromethyl group at –84.01 ppm. Therefore, efforts were made to find a more suitable ^{19}F NMR reference. First, 1,4-difluorobenzene (DFB) was tested as its resonance at –120.94 does not overlap with any of the tool compounds resonances. Unfortunately, DFB did not show high solubility in buffer and appeared to bind non-specifically to protein: upon addition of *Tc*BDF5-1 to **112** and DFB, both resonances broadened and decreased in integrals (Appendix X). Consequently, DFB could not be used as a fluorine reference. Second, potassium fluoride, which ^{19}F resonance appears at –119.79 ppm, was tested as a reference. KF did not show non-specific binding, and it is suitable a quantitative reference, as the solid salt can be prepared into an accurate stock solution in the assay buffer.

B.5.1. ^{19}F NMR reporter molecules against *Leishmania donovani*

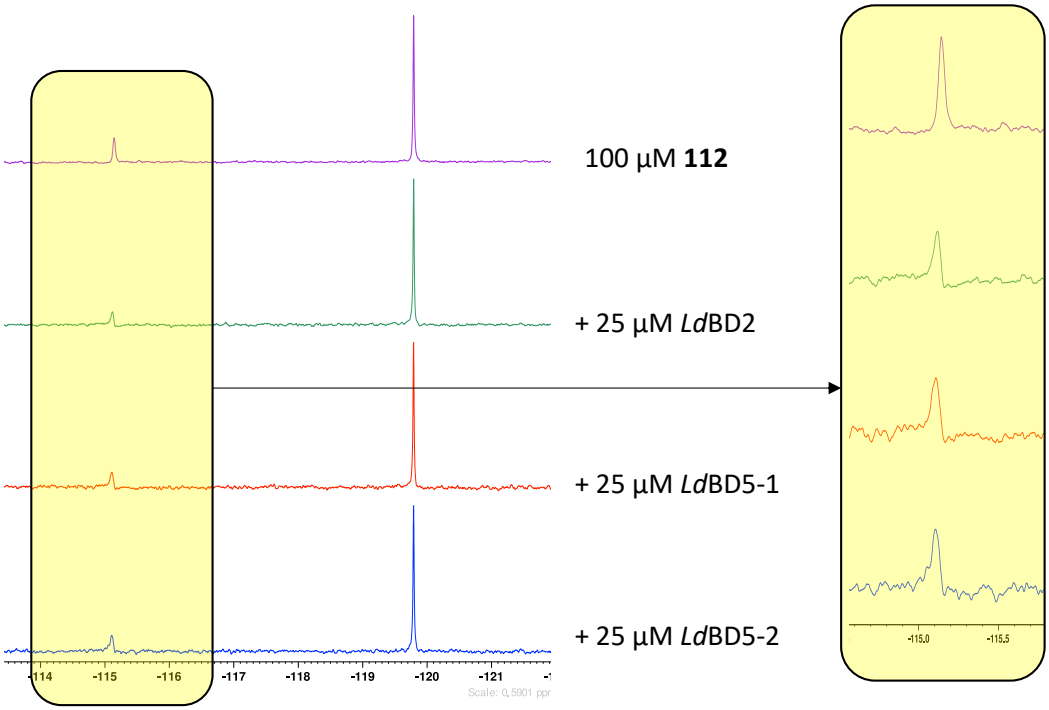


Figure B.121: The binding of **112** to *Ld*BD2, *Ld*BD5-1 and *Ld*BD 5-2 monitored by ^{19}F NMR, using KF as a reference. Compound only spectrum is shown in purple. No significant change is seen for the addition of 25 µM of (A) *Ld*BD2 (green), (B) *Ld*BD5-1 (red) or (C) *Ld*BD 5-2 (blue). **112** was synthesised by R. Berkeley or C. Evans, following the procedure of Jennings *et al*.[1]

B.6. Fluorescence polarisation

B.6.1. FI and FP of (*R*)-134, (*R*)-135, (*R*)-136 and (*R*)-137

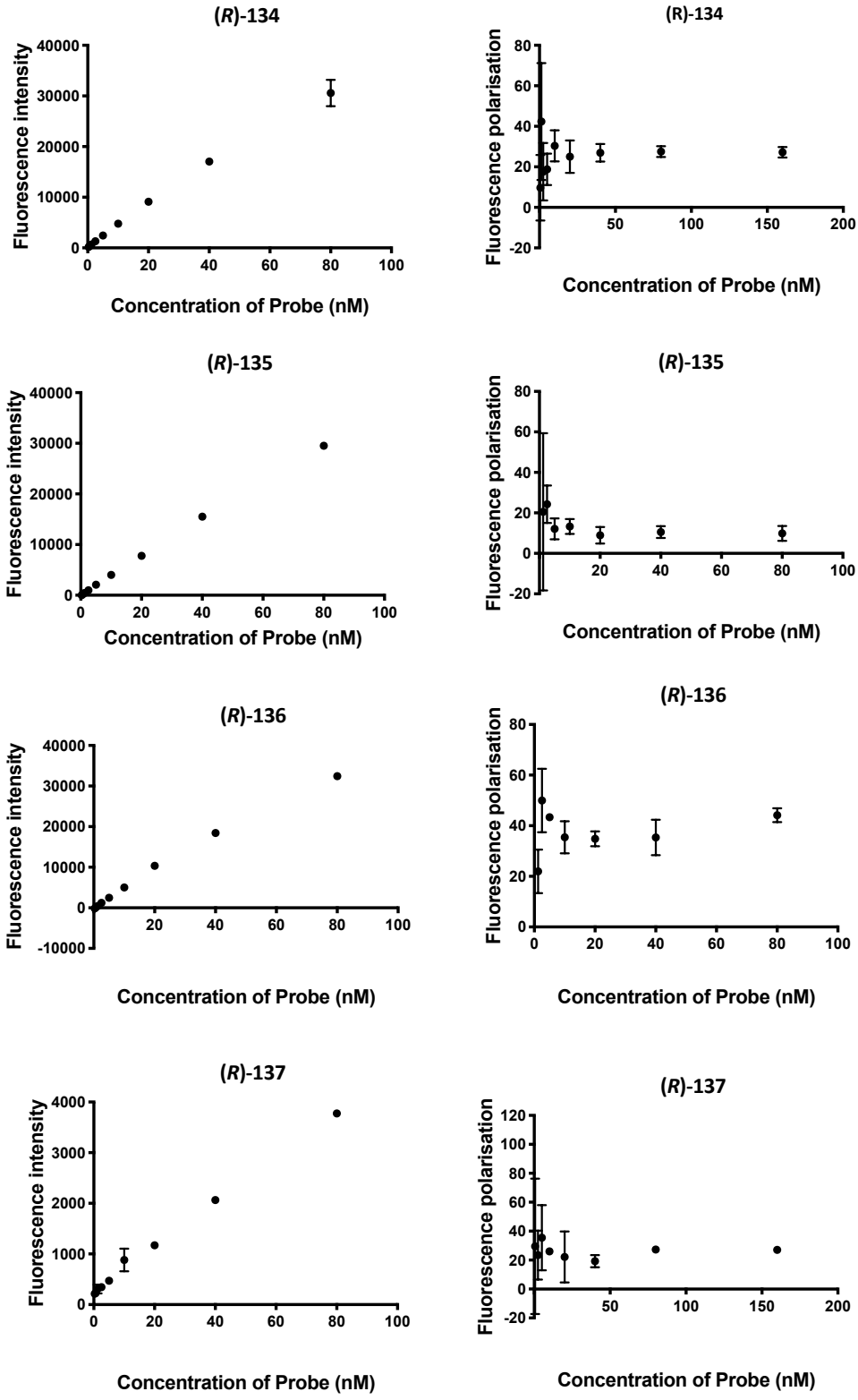


Figure B.122: The fluorescent intensity and polarisation for the fluorescent probes: (*R*)-134, (*R*)-135, (*R*)-136 and (*R*)-137. Left-hand side (A)–(D): ((*R*)-134, (*R*)-135, (*R*)-136 and (*R*)-137 show a linear FI *vs* probe concentration. Right-hand side (A)–(D): (*R*)-134, (*R*)-135, (*R*)-136 and (*R*)-137 show highly variable FP values at low probe concentration, and a stabilised FP at concentrations > 20 nM.

B.6.2. Binding of **(*R*)-135** to BRD4(1), BRD4(2) and PLK1

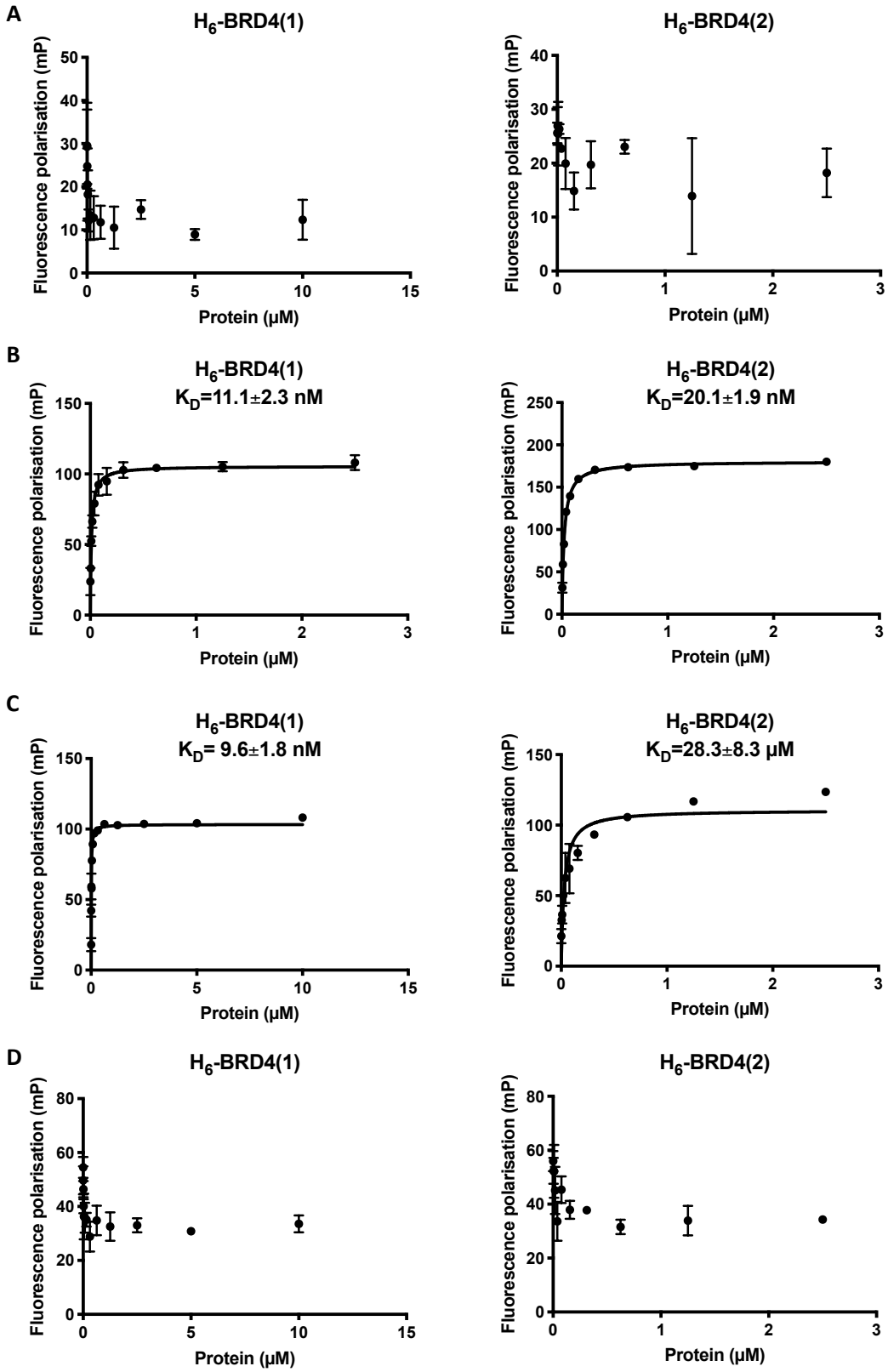


Figure B.123: The fluorescence polarisation observed for the (A) **(*R*)-134**, (B) **(*R*)-135**, (C) **(*R*)-136** and (D) **(*R*)-137** fluorescent tool compounds observed when titrating BRD4(1) and BRD4(2). **(*R*)-134** and **(*R*)-137** do not show binding to either BRD4 domains. **(*R*)-135** shows binding to both BRD4(1) and BRD4(2) with K_D=11.1±2.3 nM and 20.1±1.9 nM, respectively. **(*R*)-136** also shows binding to both protein with K_D=9.6±1.8 nM and 28.3±8.3 nM, respectively.

B.6.3. Displacement of (*R*)-135 from BRD4(1), BRD4(2) and PLK1

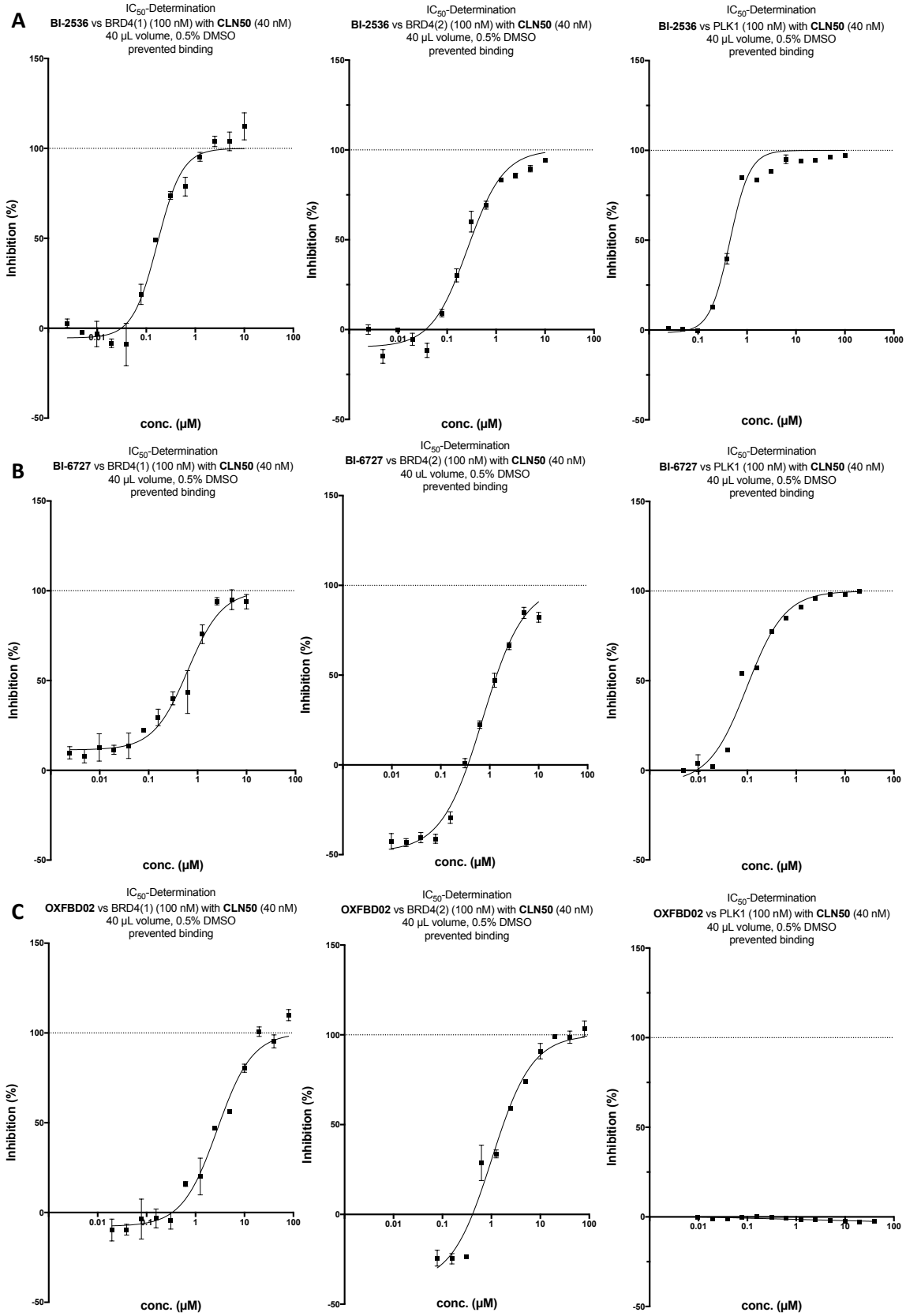


Figure B.124: The FP polarisation displacement of **(*R*)-135** from BRD4(1), BRD4(2) and PLK1 by (A) the dual PLK1/BRD4 ligand BI-2536 (**8**), showing displacement for the three proteins (B) the dual PLK1/BRD4 ligand BI-6727 (**9**), showing displacement for the three proteins and (C) the BRD4 ligand OXFBD02 (**7**), showing displacement for both BRD4 domains but not PLK1 kinase.

B.6.4. Displacement of **(*R*)-135** from *Tc*BDF3 using OXFBD02 (**7**)

To further investigate the assay detection limit, the FP displacement was next attempted with OXFBD02 (**7**). This ligand has been shown to bind to *Tc*BDF3 using waterLOGSY, and is estimated to have a 10-fold weaker binding than BI-2536 (**8**) by ^{1}H NMR ligand observed protein titration. Using the tool compound **(*R*)-135** no binding was observed by FP. (Appendix X).This is additional evidence that **(*R*)-135** can be used to detect binding of ligands that bind *Tc*BDF3 with affinities lower or equal to BI-2536 (**8**). However, cannot detect weak ligand with K_D values above 100 µM.

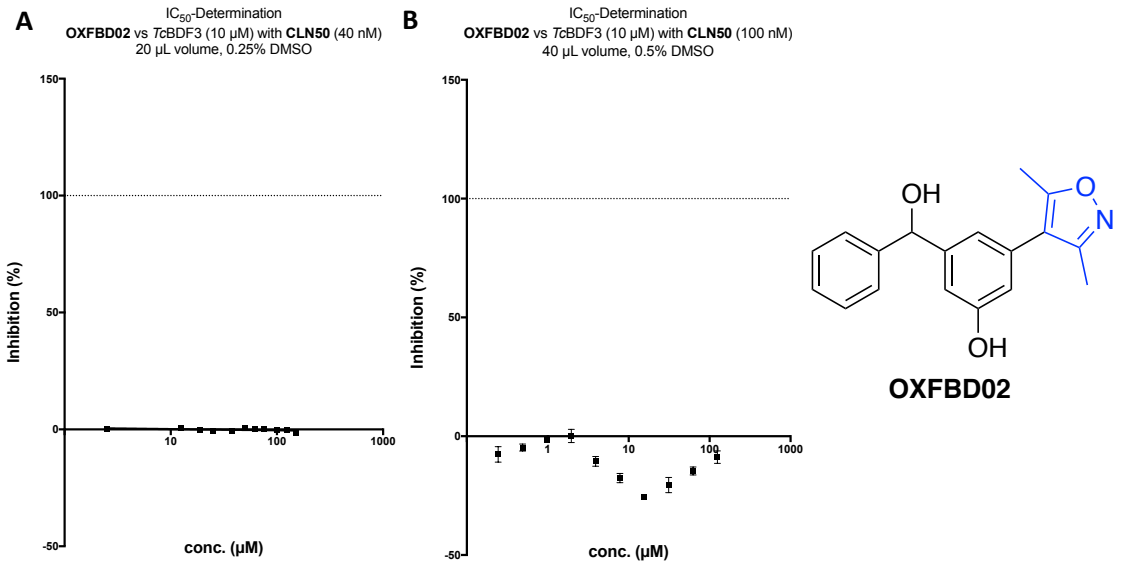


Figure B.125: No displacement of the **(*R*)-135** fluorescent probe from *Tc*BDF3 when using OXFBD02 (**7**). (A) displacement experiment using 40 nM of the FP probe in 20 µL assay volume (B) displacement experiment using 100 nM of the FP probe in 40 µL assay volume.

B.6.5. Displacement of **(*R*)-CLN50** from *Ld*BD5-2 using **BI-2536**

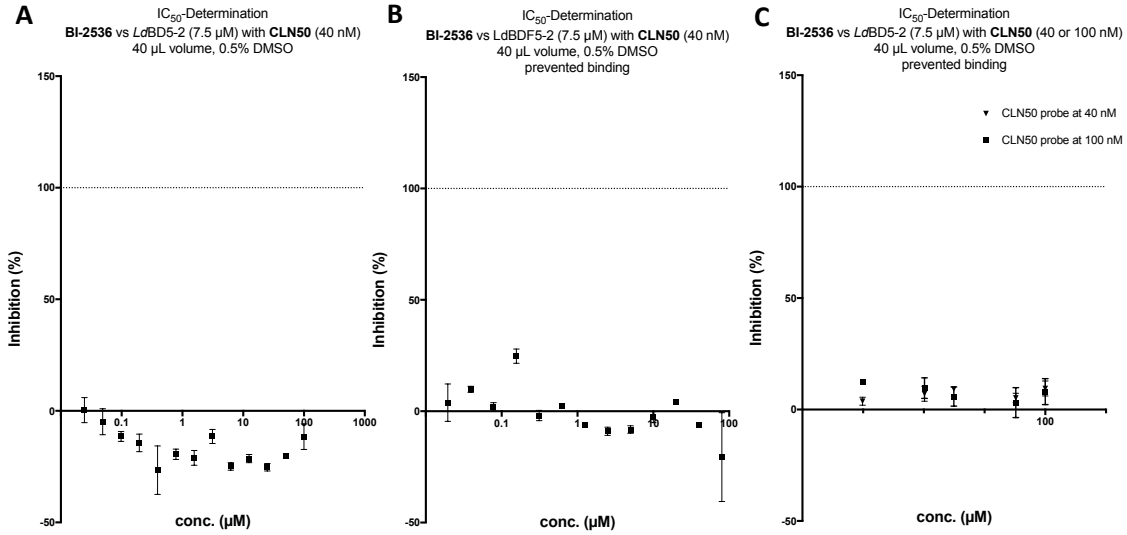


Figure B.126: The FP inhibition assay do not detect inhibition of the interaction between **(*R*)-135** and *Ld*BD5-2 by BI-2536 (**8**) in neither the (A) displacement nor the (B-C) prevented binding assay set up. (C) Increasing the FP probe concentration from 40 to 100 nM did not affect the results.

B.7. PhaBits LCMS spectra examples

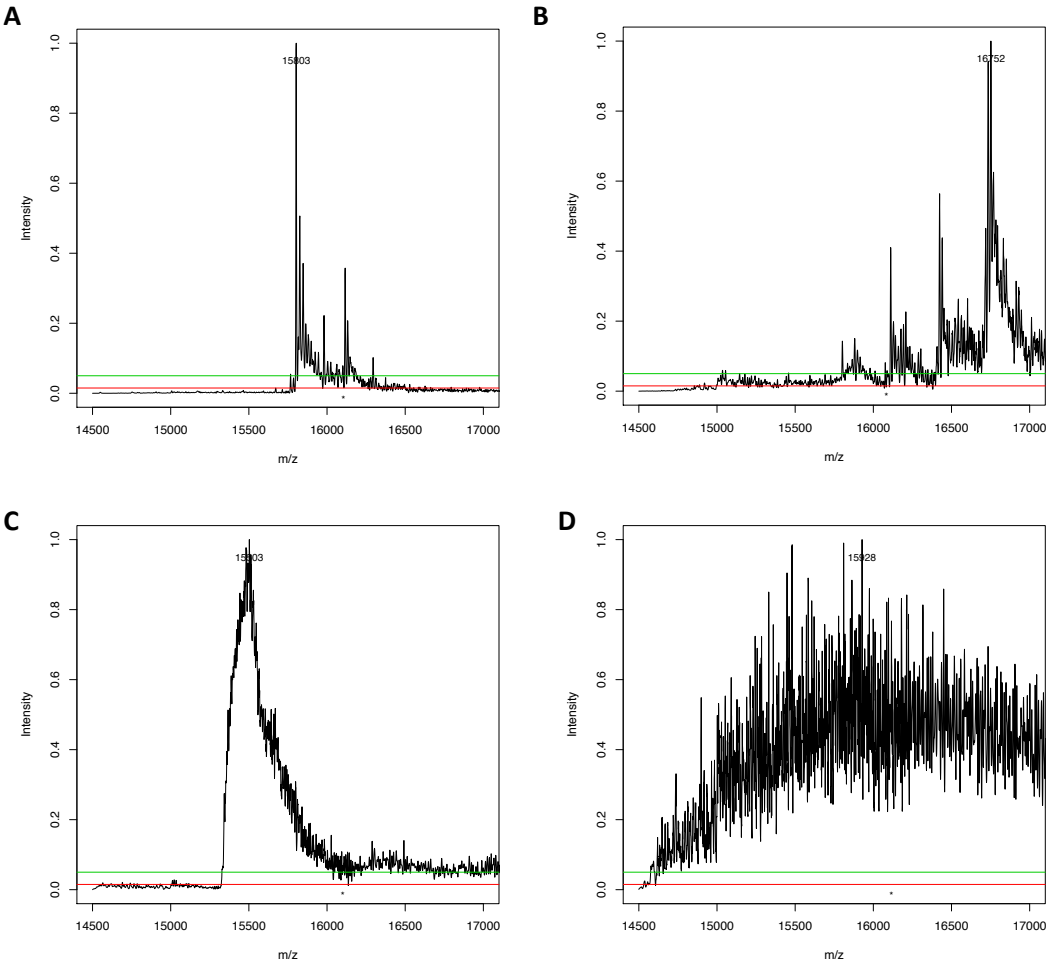


Figure B.127: Selected examples of LCMS spectra obtained for the cross-linking experiments (here: *Tc*BDF3 and PhaBits library). (A) Crosslinking >5% (green horizontal line). The expected crosslinked mass is shown by an asterix above the x axis. (B) multiple cross-linking events (C) Protein oxidation arising from radical reaction (D) failed LCMS run.

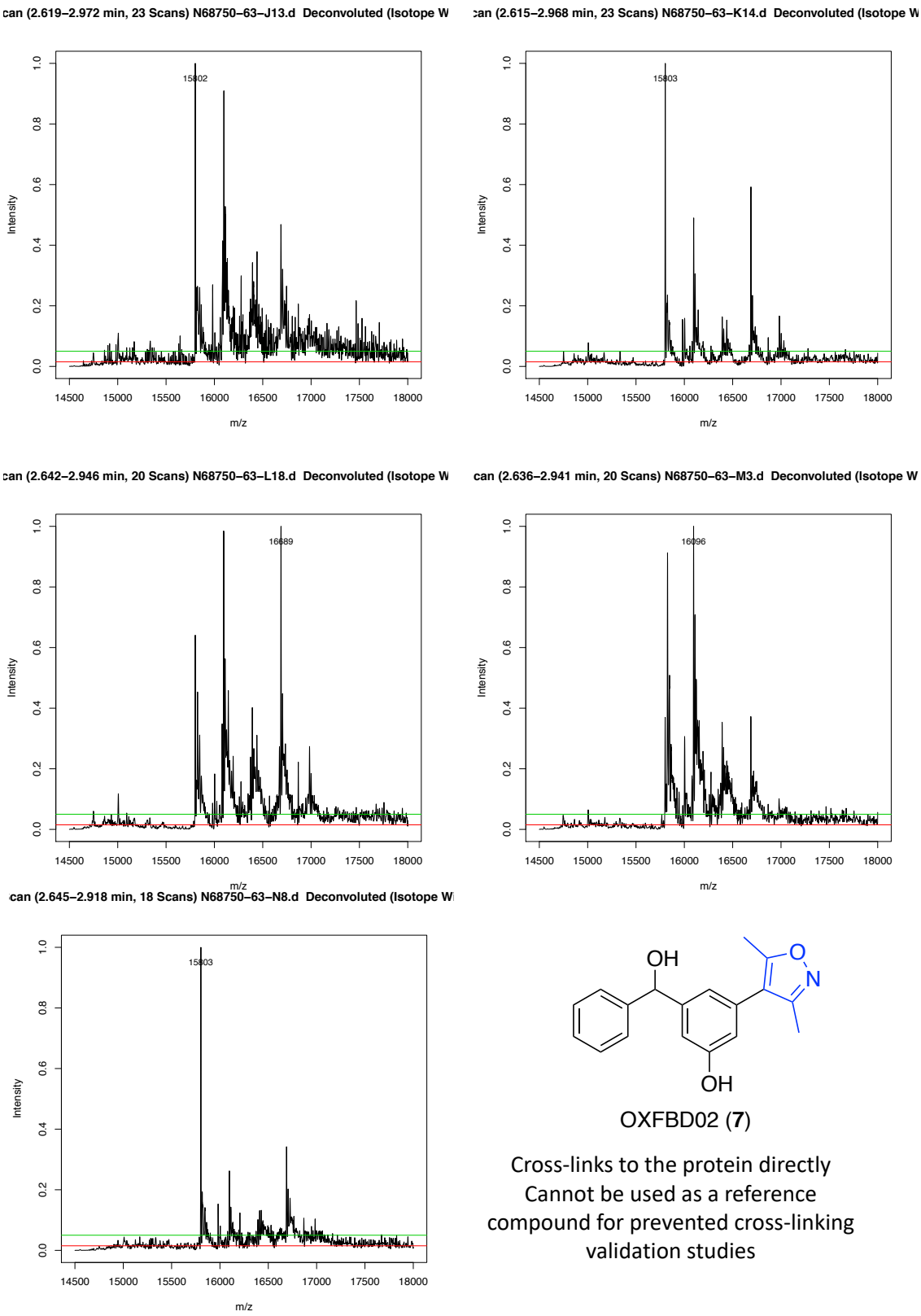
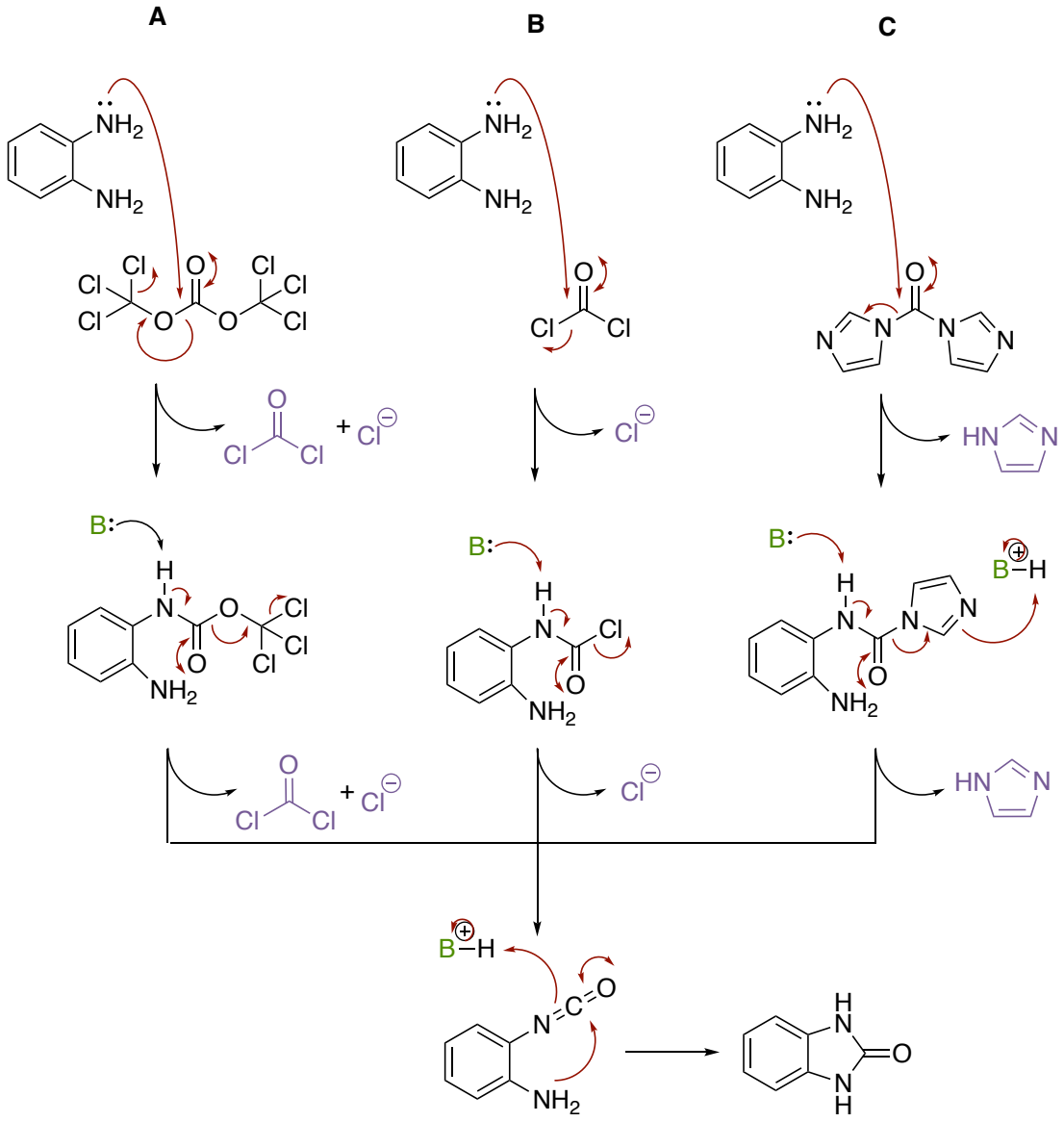


Figure B.128: Selected examples of LCMS spectra obtained for the cross-linking prevention experiments: incubation of *Tc*BDF3 and PhaBits library hits in presence of OXFBD02 (**7**) results in multiple cross-linking events in every well. OXFBD02 (**7**) cross-links directly with the protein and so cannot be used as a reference (known binder) for prevented cross-linking validation experiments.

Appendix C. Benzimidazolone ring formation mechanism



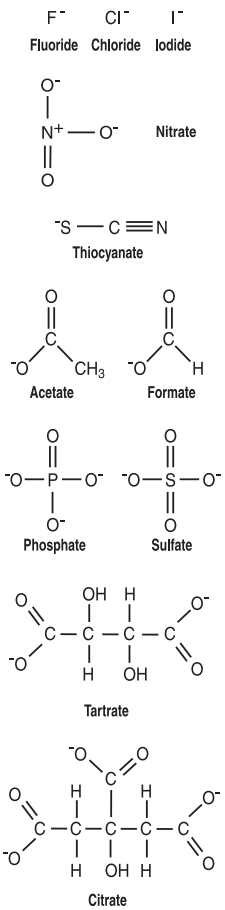
Scheme C.1: The mechanism of the benzimidazolone ring formation from dianilines using trisphosgene (A), phosgene (B) or CDI (C). Phosgene is produced as a side product when using triphosgene, hence both mechanisms (A and B) are involved when triphosgene is used as a reagent.

Appendix D. Crystallisation trials buffers

D.1. Compositions of the commercially available broad screen for

PEG / Ion HT™ HR2-139 Reagent Formulation

Well #	Salt	Well #	Polymer	Well #	pH ◊
1. (A1)	0.2 M Sodium fluoride	1. (A1)	20% w/v Polyethylene glycol 3,350	1. (A1)	7.3
2. (A2)	0.2 M Potassium fluoride	2. (A2)	20% w/v Polyethylene glycol 3,350	2. (A2)	7.3
3. (A3)	0.2 M Ammonium fluoride	3. (A3)	20% w/v Polyethylene glycol 3,350	3. (A3)	6.2
4. (A4)	0.2 M Lithium chloride	4. (A4)	20% w/v Polyethylene glycol 3,350	4. (A4)	6.8
5. (A5)	0.2 M Magnesium chloride hexahydrate	5. (A5)	20% w/v Polyethylene glycol 3,350	5. (A5)	5.9
6. (A6)	0.2 M Sodium chloride	6. (A6)	20% w/v Polyethylene glycol 3,350	6. (A6)	6.9
7. (A7)	0.2 M Calcium chloride dihydrate	7. (A7)	20% w/v Polyethylene glycol 3,350	7. (A7)	5.1
8. (A8)	0.2 M Potassium chloride	8. (A8)	20% w/v Polyethylene glycol 3,350	8. (A8)	7.0
9. (A9)	0.2 M Ammonium chloride	9. (A9)	20% w/v Polyethylene glycol 3,350	9. (A9)	6.3
10. (A10)	0.2 M Sodium iodide	10. (A10)	20% w/v Polyethylene glycol 3,350	10. (A10)	7.0
11. (A11)	0.2 M Potassium iodide	11. (A11)	20% w/v Polyethylene glycol 3,350	11. (A11)	7.0
12. (A12)	0.2 M Ammonium iodide	12. (A12)	20% w/v Polyethylene glycol 3,350	12. (A12)	6.2
13. (B1)	0.2 M Sodium thiocyanate	13. (B1)	20% w/v Polyethylene glycol 3,350	13. (B1)	6.9
14. (B2)	0.2 M Potassium thiocyanate	14. (B2)	20% w/v Polyethylene glycol 3,350	14. (B2)	7.0
15. (B3)	0.2 M Lithium nitrate	15. (B3)	20% w/v Polyethylene glycol 3,350	15. (B3)	7.1
16. (B4)	0.2 M Magnesium nitrate hexahydrate	16. (B4)	20% w/v Polyethylene glycol 3,350	16. (B4)	5.9
17. (B5)	0.2 M Sodium nitrate	17. (B5)	20% w/v Polyethylene glycol 3,350	17. (B5)	6.8
18. (B6)	0.2 M Potassium nitrate	18. (B6)	20% w/v Polyethylene glycol 3,350	18. (B6)	6.8
19. (B7)	0.2 M Ammonium nitrate	19. (B7)	20% w/v Polyethylene glycol 3,350	19. (B7)	6.2
20. (B8)	0.2 M Magnesium formate dihydrate	20. (B8)	20% w/v Polyethylene glycol 3,350	20. (B8)	7.0
21. (B9)	0.2 M Sodium formate	21. (B9)	20% w/v Polyethylene glycol 3,350	21. (B9)	7.2
22. (B10)	0.2 M Potassium formate	22. (B10)	20% w/v Polyethylene glycol 3,350	22. (B10)	7.3
23. (B11)	0.2 M Ammonium formate	23. (B11)	20% w/v Polyethylene glycol 3,350	23. (B11)	6.6
24. (B12)	0.2 M Lithium acetate dihydrate	24. (B12)	20% w/v Polyethylene glycol 3,350	24. (B12)	7.9
25. (C1)	0.2 M Magnesium acetate tetrahydrate	25. (C1)	20% w/v Polyethylene glycol 3,350	25. (C1)	7.9
26. (C2)	0.2 M Zinc acetate dihydrate	26. (C2)	20% w/v Polyethylene glycol 3,350	26. (C2)	6.4
27. (C3)	0.2 M Sodium acetate trihydrate	27. (C3)	20% w/v Polyethylene glycol 3,350	27. (C3)	8.0
28. (C4)	0.2 M Calcium acetate hydrate	28. (C4)	20% w/v Polyethylene glycol 3,350	28. (C4)	7.5
29. (C5)	0.2 M Potassium acetate	29. (C5)	20% w/v Polyethylene glycol 3,350	29. (C5)	8.1
30. (C6)	0.2 M Ammonium acetate	30. (C6)	20% w/v Polyethylene glycol 3,350	30. (C6)	7.1
31. (C7)	0.2 M Lithium sulfate monohydrate	31. (C7)	20% w/v Polyethylene glycol 3,350	31. (C7)	6.0
32. (C8)	0.2 M Magnesium sulfate heptahydrate	32. (C8)	20% w/v Polyethylene glycol 3,350	32. (C8)	6.0
33. (C9)	0.2 M Sodium sulfate decahydrate	33. (C9)	20% w/v Polyethylene glycol 3,350	33. (C9)	6.7
34. (C10)	0.2 M Potassium sulfate	34. (C10)	20% w/v Polyethylene glycol 3,350	34. (C10)	6.8
35. (C11)	0.2 M Ammonium sulfate	35. (C11)	20% w/v Polyethylene glycol 3,350	35. (C11)	6.0
36. (C12)	0.2 M Sodium tartrate dibasic dihydrate	36. (C12)	20% w/v Polyethylene glycol 3,350	36. (C12)	7.3
37. (D1)	0.2 M Potassium sodium tartrate tetrahydrate	37. (D1)	20% w/v Polyethylene glycol 3,350	37. (D1)	7.4
38. (D2)	0.2 M Ammonium tartrate dibasic	38. (D2)	20% w/v Polyethylene glycol 3,350	38. (D2)	6.6
39. (D3)	0.2 M Sodium phosphate monobasic monohydrate	39. (D3)	20% w/v Polyethylene glycol 3,350	39. (D3)	4.7
40. (D4)	0.2 M Sodium phosphate dibasic dihydrate	40. (D4)	20% w/v Polyethylene glycol 3,350	40. (D4)	9.1
41. (D5)	0.2 M Potassium phosphate monobasic	41. (D5)	20% w/v Polyethylene glycol 3,350	41. (D5)	4.8
42. (D6)	0.2 M Potassium phosphate dibasic	42. (D6)	20% w/v Polyethylene glycol 3,350	42. (D6)	9.2
43. (D7)	0.2 M Ammonium phosphate monobasic	43. (D7)	20% w/v Polyethylene glycol 3,350	43. (D7)	4.6
44. (D8)	0.2 M Ammonium phosphate dibasic	44. (D8)	20% w/v Polyethylene glycol 3,350	44. (D8)	8.0
45. (D9)	0.2 M Lithium citrate tribasic tetrahydrate	45. (D9)	20% w/v Polyethylene glycol 3,350	45. (D9)	8.4
46. (D10)	0.2 M Sodium citrate tribasic dihydrate	46. (D10)	20% w/v Polyethylene glycol 3,350	46. (D10)	8.3
47. (D11)	0.2 M Potassium citrate tribasic monohydrate	47. (D11)	20% w/v Polyethylene glycol 3,350	47. (D11)	8.3
48. (D12)	0.2 M Ammonium citrate dibasic	48. (D12)	20% w/v Polyethylene glycol 3,350	48. (D12)	5.1



◊ Measured pH at 25 ° C

PEG/Ion HT contains ninety-six unique reagents. To determine the formulation of each reagent, simply read across the page.

HAMPTON RESEARCH
Solutions for Crystal Growth

Well #	Salt	Well #	Buffer ◊	Well #	Polymer
49. (E1)	0.1 M Sodium malonate pH 4.0	49. (E1)	None	49. (E1)	12% w/v Polyethylene glycol 3,350
50. (E2)	0.2 M Sodium malonate pH 4.0	50. (E2)	None	50. (E2)	20% w/v Polyethylene glycol 3,350
51. (E3)	0.1 M Sodium malonate pH 5.0	51. (E3)	None	51. (E3)	12% w/v Polyethylene glycol 3,350
52. (E4)	0.2 M Sodium malonate pH 5.0	52. (E4)	None	52. (E4)	20% w/v Polyethylene glycol 3,350
53. (E5)	0.1 M Sodium malonate pH 6.0	53. (E5)	None	53. (E5)	12% w/v Polyethylene glycol 3,350
54. (E6)	0.2 M Sodium malonate pH 6.0	54. (E6)	None	54. (E6)	20% w/v Polyethylene glycol 3,350
55. (E7)	0.1 M Sodium malonate pH 7.0	55. (E7)	None	55. (E7)	12% w/v Polyethylene glycol 3,350
56. (E8)	0.2 M Sodium malonate pH 7.0	56. (E8)	None	56. (E8)	20% w/v Polyethylene glycol 3,350
57. (E9)	4% v/v Tacsimate™ pH 4.0	57. (E9)	None	57. (E9)	12% w/v Polyethylene glycol 3,350
58. (E10)	8% v/v Tacsimate™ pH 4.0	58. (E10)	None	58. (E10)	20% w/v Polyethylene glycol 3,350
59. (E11)	4% v/v Tacsimate™ pH 5.0	59. (E11)	None	59. (E11)	12% w/v Polyethylene glycol 3,350
60. (E12)	8% v/v Tacsimate™ pH 5.0	60. (E12)	None	60. (E12)	20% w/v Polyethylene glycol 3,350
61. (F1)	4% v/v Tacsimate™ pH 6.0	61. (F1)	None	61. (F1)	12% w/v Polyethylene glycol 3,350
62. (F2)	8% v/v Tacsimate™ pH 6.0	62. (F2)	None	62. (F2)	20% w/v Polyethylene glycol 3,350
63. (F3)	4% v/v Tacsimate™ pH 7.0	63. (F3)	None	63. (F3)	12% w/v Polyethylene glycol 3,350
64. (F4)	8% v/v Tacsimate™ pH 7.0	64. (F4)	None	64. (F4)	20% w/v Polyethylene glycol 3,350
65. (F5)	4% v/v Tacsimate™ pH 8.0	65. (F5)	None	65. (F5)	12% w/v Polyethylene glycol 3,350
66. (F6)	8% v/v Tacsimate™ pH 8.0	66. (F6)	None	66. (F6)	20% w/v Polyethylene glycol 3,350
67. (F7)	0.1 M Succinic acid pH 7.0	67. (F7)	None	67. (F7)	12% w/v Polyethylene glycol 3,350
68. (F8)	0.2 M Succinic acid pH 7.0	68. (F8)	None	68. (F8)	20% w/v Polyethylene glycol 3,350
69. (F9)	0.1 M Ammonium citrate tribasic pH 7.0	69. (F9)	None	69. (F9)	12% w/v Polyethylene glycol 3,350
70. (F10)	0.2 M Ammonium citrate tribasic pH 7.0	70. (F10)	None	70. (F10)	20% w/v Polyethylene glycol 3,350
71. (F11)	0.1 M DL-Malic acid pH 7.0	71. (F11)	None	71. (F11)	12% w/v Polyethylene glycol 3,350
72. (F12)	0.2 M DL-Malic acid pH 7.0	72. (F12)	None	72. (F12)	20% w/v Polyethylene glycol 3,350
73. (G1)	0.1 M Sodium acetate trihydrate pH 7.0	73. (G1)	None	73. (G1)	12% w/v Polyethylene glycol 3,350
74. (G2)	0.2 M Sodium acetate trihydrate pH 7.0	74. (G2)	None	74. (G2)	20% w/v Polyethylene glycol 3,350
75. (G3)	0.1 M Sodium formate pH 7.0	75. (G3)	None	75. (G3)	12% w/v Polyethylene glycol 3,350
76. (G4)	0.2 M Sodium formate pH 7.0	76. (G4)	None	76. (G4)	20% w/v Polyethylene glycol 3,350
77. (G5)	0.1 M Ammonium tartrate dibasic pH 7.0	77. (G5)	None	77. (G5)	12% w/v Polyethylene glycol 3,350
78. (G6)	0.2 M Ammonium tartrate dibasic pH 7.0	78. (G6)	None	78. (G6)	20% w/v Polyethylene glycol 3,350
79. (G7)	2% v/v Tacsimate™ pH 4.0	79. (G7)	0.1 M Sodium acetate trihydrate pH 4.6	79. (G7)	16% w/v Polyethylene glycol 3,350
80. (G8)	2% v/v Tacsimate™ pH 5.0	80. (G8)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	80. (G8)	16% w/v Polyethylene glycol 3,350
81. (G9)	2% v/v Tacsimate™ pH 6.0	81. (G9)	0.1 M BIS-TRIS pH 6.5	81. (G9)	20% w/v Polyethylene glycol 3,350
82. (G10)	2% v/v Tacsimate™ pH 7.0	82. (G10)	0.1 M HEPES pH 7.5	82. (G10)	20% w/v Polyethylene glycol 3,350
83. (G11)	2% v/v Tacsimate™ pH 8.0	83. (G11)	0.1 M Tris pH 8.5	83. (G11)	16% w/v Polyethylene glycol 3,350
84. (G12)	None	84. (G12)	0.07 M Citric acid, 0.03 M BIS-TRIS propane / pH 3.4	84. (G12)	16% w/v Polyethylene glycol 3,350
85. (H1)	None	85. (H1)	0.06 M Citric acid, 0.04 M BIS-TRIS propane / pH 4.1	85. (H1)	16% w/v Polyethylene glycol 3,350
86. (H2)	None	86. (H2)	0.05 M Citric acid, 0.05 M BIS-TRIS propane / pH 5.0	86. (H2)	16% w/v Polyethylene glycol 3,350
87. (H3)	None	87. (H3)	0.04 M Citric acid, 0.06 M BIS-TRIS propane / pH 6.4	87. (H3)	20% w/v Polyethylene glycol 3,350
88. (H4)	None	88. (H4)	0.03 M Citric acid, 0.07 M BIS-TRIS propane / pH 7.6	88. (H4)	20% w/v Polyethylene glycol 3,350
89. (H5)	None	89. (H5)	0.02 M Citric acid, 0.08 M BIS-TRIS propane / pH 8.8	89. (H5)	16% w/v Polyethylene glycol 3,350
90. (H6)	0.02 M Calcium chloride dihydrate, 0.02 M Cadmium chloride hydrate, 0.02 M Cobalt(II) chloride hexahydrate	90. (H6)	None	90. (H6)	20% w/v Polyethylene glycol 3,350
91. (H7)	0.01 M Magnesium chloride hexahydrate 0.005 M Nickel(II) chloride hexahydrate	91. (H7)	0.1 M HEPES sodium pH 7.0	91. (H7)	15% w/v Polyethylene glycol 3,350
92. (H8)	0.02 M Zinc chloride	92. (H8)	None	92. (H8)	20% w/v Polyethylene glycol 3,350
93. (H9)	0.15 M Cesium chloride	93. (H9)	None	93. (H9)	15% w/v Polyethylene glycol 3,350
94. (H10)	0.2 M Sodium bromide	94. (H10)	None	94. (H10)	20% w/v Polyethylene glycol 3,350
95. (H11)	1% w/v Tryptone, 0.001 M Sodium azide	95. (H11)	0.05 M HEPES sodium pH 7.0	95. (H11)	12% w/v Polyethylene glycol 3,350
96. (H12)	1% w/v Tryptone, 0.001 M Sodium azide	96. (H12)	0.05 M HEPES sodium pH 7.0	96. (H12)	20% w/v Polyethylene glycol 3,350

◊ Buffer pH is that of a 1.0 M stock prior to dilution with other reagent components: pH with HCl or NaOH.

PEG/Ion HT contains ninety-six unique reagents. To determine the formulation of each reagent, simply read across the page.

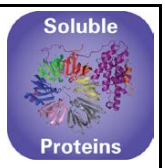
34 Journey
Aliso Viejo, CA 92656-3317 U.S.A.
Tel: (949) 425-1321 • Fax: (949) 425-1611
E-mail: tech@hrmail.com
Website: www.hamptonresearch.com

HAMPTON RESEARCH
Solutions for Crystal Growth

© 1991 - 2018 Hampton Research Corp. all rights reserved
Printed in the United States of America. This guide or parts thereof may not be reproduced in any form without the written permission of the publishers.



Molecular Dimensions



JCSG-*plus* HT-96 Wells A1- D12 MD1–40

Well #	Conc.	Salt	Conc.	Buffer	pH	Conc.	Precipitant
A1	0.2 M	Lithium sulfate	0.1 M	Sodium acetate	4.5	50 % w/v	PEG 400
A2		None	0.1 M	Sodium citrate	5.5	20 % w/v	PEG 3000
A3	0.2 M	Ammonium citrate dibasic		None		20 % w/v	PEG 3350
A4	0.02 M	Calcium chloride dihydrate	0.1 M	Sodium acetate	4.6	30 % v/v	MPD
A5	0.2 M	Magnesium formate dihydrate		None		20 % w/v	PEG 3350
A6	0.2 M	Lithium sulfate	0.1 M	Phosphate/citrate	4.2	20 % w/v	PEG 1000
A7		None	0.1 M	CHES	9.5	20 % w/v	PEG 8000
A8	0.2 M	Ammonium formate		None		20 % w/v	PEG 3350
A9	0.2 M	Ammonium chloride		None		20 % w/v	PEG 3350
A10	0.2 M	Potassium formate		None		20 % w/v	PEG 3350
A11	0.2 M	Ammonium phosphate monobasic	0.1 M	Tris	8.5	50 % v/v	MPD
A12	0.2 M	Potassium nitrate		None		20 % w/v	PEG 3350
B1	0.8 M	Ammonium sulfate	0.1 M	Citrate	4.0		None
B2	0.2 M	Sodium thiocyanate		None		20 % w/v	PEG 3350
B3		None	0.1 M	BICINE	9.0	20 % w/v	PEG 6000
B4		None	0.1 M	HEPES	7.5	10 % w/v	PEG 8000
						8 % v/v	Ethylene glycol
B5		None	0.1 M	Sodium cacodylate	6.5	40 % v/v	MPD
						5 % w/v	PEG 8000
B6		None	0.1 M	Phosphate/citrate	4.2	40 % v/v	Ethanol
						5 % w/v	PEG 1000
B7		None	0.1 M	Sodium acetate	4.6	8 % w/v	PEG 4000
B8	0.2 M	Magnesium chloride hexahydrate	0.1 M	Tris	7.0	10 % w/v	PEG 8000
B9		None	0.1 M	Citrate	5.0	20 % w/v	PEG 6000
B10	0.2 M	Magnesium chloride hexahydrate	0.1 M	Sodium cacodylate	6.5	50 % v/v	PEG 200
B11	1.6 M	Sodium citrate tribasic dihydrate pH 6.5		None			None
B12	0.2 M	Potassium citrate tribasic monohydrate		None		20 % w/v	PEG 3350
C1	0.2 M	Sodium chloride	0.1 M	Phosphate/citrate	4.2	20 % w/v	PEG 8000
C2	1.0 M	Lithium chloride	0.1 M	Citrate	4.0	20 % w/v	PEG 6000
C3	0.2 M	Ammonium nitrate		None		20 % w/v	PEG 3350
C4		None	0.1 M	HEPES	7.0	10 % w/v	PEG 6000
C5	0.8 M	Sodium phosphate monobasic monohydrate	0.1 M	Sodium HEPES	7.5		None
	0.8 M	Potassium phosphate monobasic		None			
C6		None	0.1 M	Phosphate/citrate	4.2	40 % v/v	PEG 300
C7	0.2 M	Zinc acetate dihydrate	0.1 M	Sodium acetate	4.5	10 % w/v	PEG 3000
C8		None	0.1 M	Tris	8.5	20 % v/v	Ethanol
C9		None	0.1 M	Sodium/potassium phosphate	6.2	25 % v/v	1,2-Propandiol
						10 %v/v	Glycerol
C10		None	0.1 M	BICINE	9.0	10 % w/v	PEG 20,000
						2 % v/v	1,4-Dioxane
C11	2.0 M	Ammonium sulfate	0.1 M	Sodium acetate	4.6		None
C12		None		None		10 % w/v	PEG 1000
						10 % w/v	PEG 8000
D1		None		None		24 % w/v	PEG 1500
						20 % v/v	Glycerol
D2	0.2 M	Magnesium chloride hexahydrate	0.1 M	Sodium HEPES	7.5	30 % v/v	PEG 400
D3	0.2 M	Sodium chloride	0.1 M	Sodium/potassium phosphate	6.2	50 % v/v	PEG 200
D4	0.2 M	Lithium sulfate	0.1 M	Sodium acetate	4.5	30 % w/v	PEG 8000
D5		None	0.1 M	HEPES	7.5	70 % v/v	MPD
D6	0.2 M	Magnesium chloride hexahydrate	0.1 M	Tris	8.5	20 % w/v	PEG 8000
D7	0.2 M	Lithium sulfate	0.1 M	Tris	8.5	40 % v/v	PEG 400
D8		None	0.1 M	Tris	8.0	40 % v/v	MPD
D9	0.17 M	Ammonium sulfate		None		25.5 % w/v	PEG 4000
						15 % v/v	Glycerol
D10	0.2 M	Calcium acetate hydrate	0.1 M	Sodium cacodylate	6.5	40 % v/v	PEG 300
D11	0.14 M	Calcium chloride dihydrate	0.07 M	Sodium acetate	4.6	14 % v/v	2-Propanol
						30 % v/v	Glycerol
D12	0.04 M	Potassium phosphate monobasic		None		16 % w/v	PEG 8000
						20 % v/v	Glycerol

Molecular Dimensions

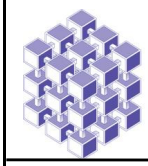
Soluble Proteins

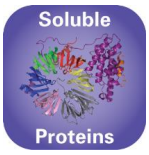
JCSG-*plus* HT-96 Wells E1 – H12 MD1–40

Well #	Conc.	Salt	Conc.	Buffer	pH	Conc.	Precipitant
E1	1.0 M	Sodium citrate tribasic dihydrate	0.1 M	Sodium cacodylate	6.5		None
E2	2.0 M	Ammonium sulfate	0.1 M	Sodium cacodylate	6.5		None
	0.2 M	Sodium chloride					
E3	0.2 M	Sodium chloride	0.1 M	HEPES	7.5	10 % v/v	2-Propanol
E4	1.26 M	Ammonium sulfate	0.1 M	Tris	8.5		None
	0.2 M	Lithium sulfate					
E5		None	0.1 M	CAPS	10.5	40 % v/v	MPD
E6	0.2 M	Zinc acetate dihydrate	0.1 M	Imidazole	8.0	20 % w/v	PEG 3000
E7	0.2 M	Zinc acetate dihydrate	0.1 M	Sodium cacodylate	6.5	10 % v/v	2-Propanol
E8	1.0 M	Ammonium phosphate dibasic	0.1 M	Sodium acetate	4.5		None
E9	1.6 M	Magneisum sulfate heptahydrate	0.1 M	MES	6.5		None
E10		None	0.1 M	BICINE	9.0	10 % w/v	PEG 6000
E11	0.16 M	Calcium acetate hydrate	0.08 M	Sodium cacodylate	6.5	14.4 % w/v	PEG 8000
						20 % v/v	Glycerol
E12		None	0.1 M	Imidazole	8.0	10 % w/v	PEG 8000
F1	0.05 M	Cesium chloride	0.1 M	MES	6.5	30 % v/v	Jeffamine® M-600
F2	3.2 M	Ammonium sulfate	0.1 M	Citrate	5.0		None
F3		None	0.1 M	Tris	8.0	20 % v/v	MPD
F4		None	0.1 M	HEPES	7.5	20 % v/v	Jeffamine® M-600
F5	0.2 M	Magnesium chloride hexahydrate	0.1 M	Tris	8.5	50 % v/v	Ethylene glycol
F6		None	0.1 M	BICINE	9.0	10 % v/v	MPD
F7	0.8 M	Succinic acid pH 7.0		None			None
F8	2.1 M	DL-Malic acid pH 7.0		None			None
F9	2.4 M	Sodium malonate dibasic monohydrate pH 7.0		None			None
F10	1.1 M	Sodium malonate dibasic monohydrate	0.1 M	HEPES	7.0	0.5 % v/v	Jeffamine® ED-2003
F11	1.0 M	Succinic acid	0.1 M	HEPES	7.0	1 % w/v	PEG 2000 MME
F12		None	0.1 M	HEPES	7.0	30 % v/v	Jeffamine® M-600
G1		None	0.1 M	HEPES	7.0	30 % v/v	Jeffamine® ED-2003
G2	0.02 M	Magnesium chloride hexahydrate	0.1 M	HEPES	7.5	22 % w/v	Poly(acrylic acid sodium salt) 5100
G3	0.01 M	Cobalt(II) chloride hexahydrate	0.1 M	Tris	8.5	20 % w/v	Polyvinylpyrrolidone
G4	0.2 M	TMAO	0.1 M	Tris	8.5	20 % w/v	PEG 2000 MME
G5	0.005 M	Cobalt(II) chloride hexahydrate	0.1 M	HEPES	7.5	12 % w/v	PEG 3350
	0.005 M	Cadmium chloride hemi(pentahydrate)					
	0.005 M	Magnesium chloride hexahydrate					
	0.005 M	Nickel(II) chloride hexahydrate					
G6	0.2 M	Sodium malonate dibasic monohydrate		None		20 % w/v	PEG 3350
G7	0.1 M	Succinic acid		None		15 % w/v	PEG 3350
G8	0.15 M	DL-Malic acid		None		20 % w/v	PEG 3350
G9	0.1 M	Potassium thiocyanate		None		30 % w/v	PEG 2000 MME
G10	0.15 M	Potassium bromide		None		30 % w/v	PEG 2000 MME
G11	2.0 M	Ammonium sulfate	0.1 M	BIS-Tris	5.5		None
G12	3.0 M	Sodium chloride	0.1 M	BIS-Tris	5.5		None
H1	0.3 M	Magnesium formate dihydrate	0.1 M	BIS-Tris	5.5		None
H2	1.0 M	Ammonium sulfate	0.1 M	BIS-Tris	5.5	1 % w/v	PEG 3350
H3		None	0.1 M	BIS-Tris	5.5	25 % w/v	PEG 3350
H4	0.2 M	Calcium chloride dihydrate	0.1 M	BIS-Tris	5.5	45 % v/v	MPD
H5	0.2 M	Ammonium acetate	0.1 M	BIS-Tris	5.5	45 % v/v	MPD
H6	0.1 M	Ammonium acetate	0.1 M	BIS-Tris	5.5	17 % w/v	PEG 10,000
H7	0.2 M	Ammonium sulfate	0.1 M	BIS-Tris	5.5	25 % w/v	PEG 3350
H8	0.2 M	Sodium chloride	0.1 M	BIS-Tris	5.5	25 % w/v	PEG 3350
H9	0.2 M	Lithium sulfate	0.1 M	BIS-Tris	5.5	25 % w/v	PEG 3350
H10	0.2 M	Ammonium acetate	0.1 M	BIS-Tris	5.5	25 % w/v	PEG 3350
H11	0.2 M	Magnesium chloride hexahydrate	0.1 M	BIS-Tris	5.5	25 % w/v	PEG 3350
H12	0.2 M	Ammonium acetate	0.1 M	HEPES	7.5	45 % v/v	MPD

Abbreviations: Bis Tris; Bis-(2-hydroxyethyl)imino-tris(hydroxymethyl)methane, **CAPS**; N-Cyclohexyl-3-aminopropanesulfonic acid**, CHES**; 2-(N-Cyclohexylamino)ethane Sulfonic Acid**, HEPES;** 2-(4-(2-Hydroxyethyl)-1-piperazinyl)ethanesulfonic Acid, **Na HEPES**; 2-(4-(2-Hydroxyethyl)-1-piperazinyl)ethanesulfonic Acid Sodium Salt, **MES**; 2-(N-morpholino)ethanesulfonic acid, **MPD;** 2,4-methyl pentanediol, **PEG**; Polyethylene glycol (2K, 6K , 8K and 10K correspond to the molecular weight, in thousands of Daltons, of PEG), **TMAO:** Trimethylamine N-oxide, **Tris**; 2-Amino-2-(hydroxymethyl)propane-1,3-diol.

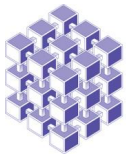
N.B. Jeffamine ED-2001 has been superseded with Jeffamine ED-2003. Polyvinylpyrolidone K15 is called Polyvinylpyrolidine.



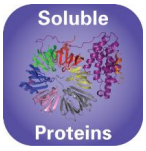


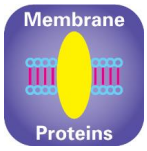
PACT *premier*™ HT-96 / FX-96 **Conditions A1-D12** **MD1-36 / MD1-36-FX**

Well #	Conc.	Salt	Conc.	Buffer	pH	Conc.	Precipitant
A1			0.1 M	SPG	4.0	25 % w/v	PEG 1500
A2			0.1 M	SPG	5.0	25 % w/v	PEG 1500
A3			0.1 M	SPG	6.0	25 % w/v	PEG 1500
A4			0.1 M	SPG	7.0	25 % w/v	PEG 1500
A5			0.1 M	SPG	8.0	25 % w/v	PEG 1500
A6			0.1 M	SPG	9.0	25 % w/v	PEG 1500
A7	0.2 M	Sodium chloride	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
A8	0.2 M	Ammonium chloride	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
A9	0.2 M	Lithium chloride	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
A10	0.2 M	Magnesium chloride hexahydrate	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
A11	0.2 M	Calcium chloride dihydrate	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
A12	0.01 M	Zinc chloride	0.1 M	Sodium acetate	5.0	20 % w/v	PEG 6000
B1			0.1 M	MIB	4.0	25 % w/v	PEG 1500
B2			0.1 M	MIB	5.0	25 % w/v	PEG 1500
B3			0.1 M	MIB	6.0	25 % w/v	PEG 1500
B4			0.1 M	MIB	7.0	25 % w/v	PEG 1500
B5			0.1 M	MIB	8.0	25 % w/v	PEG 1500
B6			0.1 M	MIB	9.0	25 % w/v	PEG 1500
B7	0.2 M	Sodium chloride	0.1 M	MES	6.0	20 % w/v	PEG 6000
B8	0.2 M	Ammonium chloride	0.1 M	MES	6.0	20 % w/v	PEG 6000
B9	0.2 M	Lithium chloride	0.1 M	MES	6.0	20 % w/v	PEG 6000
B10	0.2 M	Magnesium chloride hexahydrate	0.1 M	MES	6.0	20 % w/v	PEG 6000
B11	0.2 M	Calcium chloride dihydrate	0.1 M	MES	6.0	20 % w/v	PEG 6000
B12	0.01 M	Zinc chloride	0.1 M	MES	6.0	20 % w/v	PEG 6000
C1			0.1 M	PCTP	4.0	25 % w/v	PEG 1500
C2			0.1 M	PCTP	5.0	25 % w/v	PEG 1500
C3			0.1 M	PCTP	6.0	25 % w/v	PEG 1500
C4			0.1 M	PCTP	7.0	25 % w/v	PEG 1500
C5			0.1 M	PCTP	8.0	25 % w/v	PEG 1500
C6			0.1 M	PCTP	9.0	25 % w/v	PEG 1500
C7	0.2 M	Sodium chloride	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
C8	0.2 M	Ammonium chloride	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
C9	0.2 M	Lithium chloride	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
C10	0.2 M	Magnesium chloride hexahydrate	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
C11	0.2 M	Calcium chloride dihydrate	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
C12	0.01 M	Zinc chloride	0.1 M	HEPES	7.0	20 % w/v	PEG 6000
D1			0.1 M	MMT	4.0	25 % w/v	PEG 1500
D2			0.1 M	MMT	5.0	25 % w/v	PEG 1500
D3			0.1 M	MMT	6.0	25 % w/v	PEG 1500
D4			0.1 M	MMT	7.0	25 % w/v	PEG 1500
D5			0.1 M	MMT	8.0	25 % w/v	PEG 1500
D6			0.1 M	MMT	9.0	25 % w/v	PEG 1500
D7	0.2 M	Sodium chloride	0.1 M	Tris	8.0	20 % w/v	PEG 6000
D8	0.2 M	Ammonium chloride	0.1 M	Tris	8.0	20 % w/v	PEG 6000
D9	0.2 M	Lithium chloride	0.1 M	Tris	8.0	20 % w/v	PEG 6000
D10	0.2 M	Magnesium chloride hexahydrate	0.1 M	Tris	8.0	20 % w/v	PEG 6000
D11	0.2 M	Calcium chloride dihydrate	0.1 M	Tris	8.0	20 % w/v	PEG 6000
D12	0.002 M	Zinc chloride	0.1 M	Tris	8.0	20 % w/v	PEG 6000



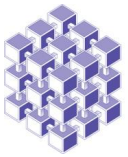
Molecular Dimensions



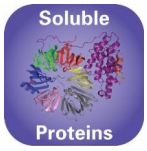


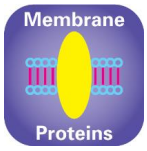
Morpheus® Conditions 1-48 (Box 1) MD1–46

Tube	Conc	Ligands	Conc	Buffer	pH	Conc	Precipitant
1-1	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 1
1-2	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 2
1-3	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 3
1-4	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 4
1-5	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 1
1-6	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 2
1-7	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 3
1-8	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 4
1-9	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
1-10	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
1-11	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
1-12	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4
1-13	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 1
1-14	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 2
1-15	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 3
1-16	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 4
1-17	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 1
1-18	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 2
1-19	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 3
1-20	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 4
1-21	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
1-22	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
1-23	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
1-24	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4
1-25	0.09 M	NPS	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 1
1-26	0.09 M	NPS	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 2
1-27	0.09 M	NPS	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 3
1-28	0.09 M	NPS	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 4
1-29	0.09 M	NPS	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 1
1-30	0.09 M	NPS	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 2
1-31	0.09 M	NPS	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 3
1-32	0.09 M	NPS	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 4
1-33	0.09 M	NPS	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
1-34	0.09 M	NPS	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
1-35	0.09 M	NPS	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
1-36	0.09 M	NPS	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4
1-37	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 1
1-38	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 2
1-39	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 3
1-40	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 4
1-41	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 1
1-42	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 2
1-43	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 3
1-44	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 4
1-45	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
1-46	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
1-47	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
1-48	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4



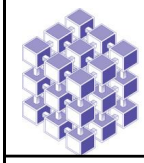
Molecular Dimensions

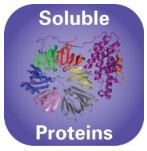




Morpheus® Conditions 1-48 (Box 1) MD1–46

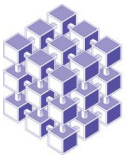
Tube	Conc	Ligands	Conc	Buffer	pH	Conc	Precipitant
1-1	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 1
1-2	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 2
1-3	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 3
1-4	0.06 M	Divalents	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 4
1-5	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 1
1-6	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 2
1-7	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 3
1-8	0.06 M	Divalents	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 4
1-9	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
1-10	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
1-11	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
1-12	0.06 M	Divalents	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4
1-13	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 1
1-14	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 2
1-15	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 3
1-16	0.09 M	Halogens	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 4
1-17	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 1
1-18	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 2
1-19	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 3
1-20	0.09 M	Halogens	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 4
1-21	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
1-22	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
1-23	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
1-24	0.09 M	Halogens	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4
1-25	0.09 M	NPS	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 1
1-26	0.09 M	NPS	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 2
1-27	0.09 M	NPS	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 3
1-28	0.09 M	NPS	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 4
1-29	0.09 M	NPS	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 1
1-30	0.09 M	NPS	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 2
1-31	0.09 M	NPS	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 3
1-32	0.09 M	NPS	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 4
1-33	0.09 M	NPS	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
1-34	0.09 M	NPS	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
1-35	0.09 M	NPS	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
1-36	0.09 M	NPS	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4
1-37	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 1
1-38	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 2
1-39	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 3
1-40	0.12 M	Alcohols	0.1 M	Buffer System 1	6.5	50 % v/v	Precipitant Mix 4
1-41	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 1
1-42	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 2
1-43	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 3
1-44	0.12 M	Alcohols	0.1 M	Buffer System 2	7.5	50 % v/v	Precipitant Mix 4
1-45	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 1
1-46	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 2
1-47	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 3
1-48	0.12 M	Alcohols	0.1 M	Buffer System 3	8.5	50 % v/v	Precipitant Mix 4

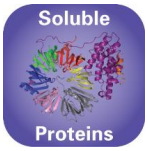


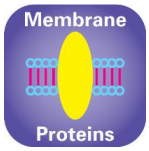


PACT *premier*™ HT-96 / FX-96 **Conditions E1-H12** **MD1-36 / MD1-36-FX**

Well #	Conc.	Salt	Conc.	Buffer	pH	Conc.	Precipitant
E1	0.2 M	Sodium fluoride				20 % w/v	PEG 3350
E2	0.2 M	Sodium bromide				20 % w/v	PEG 3350
E3	0.2 M	Sodium iodide				20 % w/v	PEG 3350
E4	0.2 M	Potassium thiocyanate				20 % w/v	PEG 3350
E5	0.2 M	Sodium nitrate				20 % w/v	PEG 3350
E6	0.2 M	Sodium formate				20 % w/v	PEG 3350
E7	0.2 M	Sodium acetate trihydrate				20 % w/v	PEG 3350
E8	0.2 M	Sodium sulfate				20 % w/v	PEG 3350
E9	0.2 M	Potassium sodium tartrate tetrahydrate				20 % w/v	PEG 3350
E10	0.02 M	Sodium/potassium phosphate				20 % w/v	PEG 3350
E11	0.2 M	Sodium citrate tribasic dihydrate				20 % w/v	PEG 3350
E12	0.2 M	Sodium malonate dibasic monohydrate				20 % w/v	PEG 3350
F1	0.2 M	Sodium fluoride	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F2	0.2 M	Sodium bromide	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F3	0.2 M	Sodium iodide	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F4	0.2 M	Potassium thiocyanate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F5	0.2 M	Sodium nitrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F6	0.2 M	Sodium formate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F7	0.2 M	Sodium acetate trihydrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F8	0.2 M	Sodium sulfate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F9	0.2 M	Potassium sodium tartrate tetrahydrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F10	0.02 M	Sodium/potassium phosphate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F11	0.2 M	Sodium citrate tribasic dihydrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
F12	0.2 M	Sodium malonate dibasic monohydrate	0.1 M	Bis-Tris propane	6.5	20 % w/v	PEG 3350
G1	0.2 M	Sodium fluoride	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G2	0.2 M	Sodium bromide	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G3	0.2 M	Sodium iodide	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G4	0.2 M	Potassium thiocyanate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G5	0.2 M	Sodium nitrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G6	0.2 M	Sodium formate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G7	0.2 M	Sodium acetate trihydrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G8	0.2 M	Sodium sulfate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G9	0.2 M	Potassium sodium tartrate tetrahydrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G10	0.02 M	Sodium/potassium phosphate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G11	0.2 M	Sodium citrate tribasic dihydrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
G12	0.2 M	Sodium malonate dibasic monohydrate	0.1 M	Bis-Tris propane	7.5	20 % w/v	PEG 3350
H1	0.2 M	Sodium fluoride	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H2	0.2 M	Sodium bromide	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H3	0.2 M	Sodium iodide	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H4	0.2 M	Potassium thiocyanate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H5	0.2 M	Sodium nitrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H6	0.2 M	Sodium formate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H7	0.2 M	Sodium acetate trihydrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H8	0.2 M	Sodium sulfate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H9	0.2 M	Potassium sodium tartrate tetrahydrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H10	0.02 M	Sodium/potassium phosphate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H11	0.2 M	Sodium citrate tribasic dihydrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350
H12	0.2 M	Sodium malonate dibasic monohydrate	0.1 M	Bis-Tris propane	8.5	20 % w/v	PEG 3350

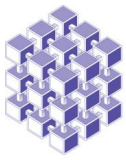


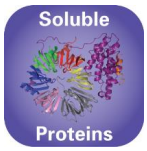


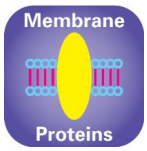


Morpheus II Conditions 1-48 (Box 1) MD1-91

Screen ID	Conc.	Additives (PDB ligands)	Conc.	Buffer	pH	Conc.	Precipitant
1-1	90 mM	LiNaK	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 5
1-2	90 mM	LiNaK	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 6
1-3	90 mM	LiNaK	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 7
1-4	90 mM	LiNaK	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 8
1-5	90 mM	LiNaK	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 5
1-6	90 mM	LiNaK	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 6
1-7	90 mM	LiNaK	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 7
1-8	90 mM	LiNaK	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 8
1-9	90 mM	LiNaK	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 5
1-10	90 mM	LiNaK	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 6
1-11	90 mM	LiNaK	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 7
1-12	90 mM	LiNaK	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 8
1-13	2 mM	Divalents II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 5
1-14	2 mM	Divalents II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 6
1-15	2 mM	Divalents II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 7
1-16	2 mM	Divalents II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 8
1-17	2 mM	Divalents II		none		50 % v/v	Precipitant Mix 5
1-18	2 mM	Divalents II		none		50 % v/v	Precipitant Mix 6
1-19	2 mM	Divalents II		none		50 % v/v	Precipitant Mix 7
1-20	2 mM	Divalents II		none		50 % v/v	Precipitant Mix 8
1-21	2 mM	Divalents II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 5
1-22	2 mM	Divalents II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 6
1-23	2 mM	Divalents II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 7
1-24	2 mM	Divalents II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 8
1-25	4 mM	Alkalis	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 5
1-26	4 mM	Alkalis	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 6
1-27	4 mM	Alkalis	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 7
1-28	4 mM	Alkalis	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 8
1-29	4 mM	Alkalis	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 5
1-30	4 mM	Alkalis	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 6
1-31	4 mM	Alkalis	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 7
1-32	4 mM	Alkalis	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 8
1-33	4 mM	Alkalis	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 5
1-34	4 mM	Alkalis	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 6
1-35	4 mM	Alkalis	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 7
1-36	4 mM	Alkalis	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 8
1-37	2 mM	Oxometalates	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 5
1-38	2 mM	Oxometalates	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 6
1-39	2 mM	Oxometalates	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 7
1-40	2 mM	Oxometalates	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 8
1-41	2 mM	Oxometalates	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 5
1-42	2 mM	Oxometalates	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 6
1-43	2 mM	Oxometalates	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 7
1-44	2 mM	Oxometalates	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 8
1-45	2 mM	Oxometalates	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 5
1-46	2 mM	Oxometalates	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 6
1-47	2 mM	Oxometalates	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 7
1-48	2 mM	Oxometalates	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 8







Morpheus II Conditions 1-48 (Box 2) MD1-91

Screen ID	Conc.	Additives (PDB ligands)	Conc.	Buffer	pH	Conc.	Precipitant
2-1	2 mM	Lanthanides	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 5
2-2	2 mM	Lanthanides	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 6
2-3	2 mM	Lanthanides	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 7
2-4	2 mM	Lanthanides	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 8
2-5	2 mM	Lanthanides	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 5
2-6	2 mM	Lanthanides	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 6
2-7	2 mM	Lanthanides	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 7
2-8	2 mM	Lanthanides	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 8
2-9	2 mM	Lanthanides	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 5
2-10	2 mM	Lanthanides	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 6
2-11	2 mM	Lanthanides	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 7
2-12	2 mM	Lanthanides	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 8
2-13	100 mM	Monosaccharides II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 5
2-14	100 mM	Monosaccharides II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 6
2-15	100 mM	Monosaccharides II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 7
2-16	100 mM	Monosaccharides II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 8
2-17	100 mM	Monosaccharides II	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 5
2-18	100 mM	Monosaccharides II	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 6
2-19	100 mM	Monosaccharides II	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 7
2-20	100 mM	Monosaccharides II	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 8
2-21	100 mM	Monosaccharides II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 5
2-22	100 mM	Monosaccharides II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 6
2-23	100 mM	Monosaccharides II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 7
2-24	100 mM	Monosaccharides II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 8
2-25	100 mM	Amino acids II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 5
2-26	100 mM	Amino acids II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 6
2-27	100 mM	Amino acids II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 7
2-28	100 mM	Amino acids II	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 8
2-29	100 mM	Amino acids II	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 5
2-30	100 mM	Amino acids II	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 6
2-31	100 mM	Amino acids II	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 7
2-32	100 mM	Amino acids II	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 8
2-33	100 mM	Amino acids II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 5
2-34	100 mM	Amino acids II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 6
2-35	100 mM	Amino acids II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 7
2-36	100 mM	Amino acids II	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 8
2-37	40 mM	Polyamines	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 5
2-38	40 mM	Polyamines	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 6
2-39	40 mM	Polyamines	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 7
2-40	40 mM	Polyamines	0.1 M	Buffer System 4	6.5	50 % v/v	Precipitant Mix 8
2-41	40 mM	Polyamines	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 5
2-42	40 mM	Polyamines	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 6
2-43	40 mM	Polyamines	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 7
2-44	40 mM	Polyamines	0.1 M	Buffer System 5	7.5	50 % v/v	Precipitant Mix 8
2-45	40 mM	Polyamines	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 5
2-46	40 mM	Polyamines	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 6
2-47	40 mM	Polyamines	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 7
2-48	40 mM	Polyamines	0.1 M	Buffer System 6	8.5	50 % v/v	Precipitant Mix 8

Tube #	Salt	Buffer ◊	Precipitant
1.	None	0.1 M Citric acid pH 3.5	2.0 M Ammonium sulfate
2.	None	0.1 M Sodium acetate trihydrate pH 4.5	2.0 M Ammonium sulfate
3.	None	0.1 M BIS-TRIS pH 5.5	2.0 M Ammonium sulfate
4.	None	0.1 M BIS-TRIS pH 6.5	2.0 M Ammonium sulfate
5.	None	0.1 M HEPES pH 7.5	2.0 M Ammonium sulfate
6.	None	0.1 M Tris pH 8.5	2.0 M Ammonium sulfate
7.	None	0.1 M Citric acid pH 3.5	3.0 M Sodium chloride
8.	None	0.1 M Sodium acetate trihydrate pH 4.5	3.0 M Sodium chloride
9.	None	0.1 M BIS-TRIS pH 5.5	3.0 M Sodium chloride
10.	None	0.1 M BIS-TRIS pH 6.5	3.0 M Sodium chloride
11.	None	0.1 M HEPES pH 7.5	3.0 M Sodium chloride
12.	None	0.1 M Tris pH 8.5	3.0 M Sodium chloride
13.	None	0.1 M BIS-TRIS pH 5.5	0.3 M Magnesium formate dihydrate
14.	None	0.1 M BIS-TRIS pH 6.5	0.5 M Magnesium formate dihydrate
15.	None	0.1 M HEPES pH 7.5	0.5 M Magnesium formate dihydrate
16.	None	0.1 M Tris pH 8.5	0.3 M Magnesium formate dihydrate
17.	None	None - pH 5.6	1.26 M Sodium phosphate monobasic monohydrate 0.14 M Potassium phosphate dibasic
18.	None	None - pH 6.9	0.49 M Sodium phosphate monobasic monohydrate 0.91 M Potassium phosphate dibasic
19.	None	None - pH 8.2	0.056 M Sodium phosphate monobasic monohydrate 1.344 M Potassium phosphate dibasic
20.	None	0.1 M HEPES pH 7.5	1.4 M Sodium citrate tribasic dihydrate
21.	None	None	1.8 M Ammonium citrate tribasic pH 7.0
22.	None	None	0.8 M Succinic acid pH 7.0
23.	None	None	2.1 M DL-Malic acid pH 7.0
24.	None	None	2.8 M Sodium acetate trihydrate pH 7.0
25.	None	None	3.5 M Sodium formate pH 7.0
26.	None	None	1.1 M Ammonium tartrate dibasic pH 7.0
27.	None	None	2.4 M Sodium malonate pH 7.0
28.	None	None	35% v/v Tacsimate™ pH 7.0
29.	None	None	60% v/v Tacsimate™ pH 7.0
30.	0.1 M Sodium chloride	0.1 M BIS-TRIS pH 6.5	1.5 M Ammonium sulfate
31.	0.8 M Potassium sodium tartrate tetrahydrate	0.1 M Tris pH 8.5	0.5% w/v Polyethylene glycol monomethyl ether 5,000
32.	1.0 M Ammonium sulfate	0.1 M BIS-TRIS pH 5.5	1% w/v Polyethylene glycol 3,350
33.	1.1 M Sodium malonate pH 7.0	0.1 M HEPES pH 7.0	0.5% v/v Jeffamine® ED-2001 pH 7.0
34.	1.0 M Succinic acid pH 7.0	0.1 M HEPES pH 7.0	1% w/v Polyethylene glycol monomethyl ether 2,000
35.	1.0 M Ammonium sulfate	0.1 M HEPES pH 7.0	0.5% w/v Polyethylene glycol 8,000
36.	15% v/v Tacsimate™ pH 7.0	0.1 M HEPES pH 7.0	2% w/v Polyethylene glycol 3,350
37.	None	None	25% w/v Polyethylene glycol 1,500
38.	None	0.1 M HEPES pH 7.0	30% v/v Jeffamine® M-600® pH 7.0
39.	None	0.1 M HEPES pH 7.0	30% v/v Jeffamine® ED-2001 pH 7.0
40.	None	0.1 M Citric acid pH 3.5	25% w/v Polyethylene glycol 3,350
41.	None	0.1 M Sodium acetate trihydrate pH 4.5	25% w/v Polyethylene glycol 3,350
42.	None	0.1 M BIS-TRIS pH 5.5	25% w/v Polyethylene glycol 3,350
43.	None	0.1 M BIS-TRIS pH 6.5	25% w/v Polyethylene glycol 3,350
44.	None	0.1 M HEPES pH 7.5	25% w/v Polyethylene glycol 3,350
45.	None	0.1 M Tris pH 8.5	25% w/v Polyethylene glycol 3,350
46.	None	0.1 M BIS-TRIS pH 6.5	20% w/v Polyethylene glycol monomethyl ether 5,000
47.	None	0.1 M BIS-TRIS pH 6.5	28% w/v Polyethylene glycol monomethyl ether 2,000
48.	0.2 M Calcium chloride dihydrate	0.1 M BIS-TRIS pH 5.5	45% v/v (+/-)-2-Methyl-2,4-pentanediol

◊ Buffer pH is that of a 1.0 M stock prior to dilution with other reagent components: pH with HCl or NaOH.

Index contains ninety-six unique reagents. To determine the formulation of each reagent, simply read across the page.

Tube #	Salt	Tube #	Buffer ◊	Tube #	Precipitant
49.	0.2 M Calcium chloride dihydrate	49.	0.1 M BIS-TRIS pH 6.5	49.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
50.	0.2 M Ammonium acetate	50.	0.1 M BIS-TRIS pH 5.5	50.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
51.	0.2 M Ammonium acetate	51.	0.1 M BIS-TRIS pH 6.5	51.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
52.	0.2 M Ammonium acetate	52.	0.1 M HEPES pH 7.5	52.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
53.	0.2 M Ammonium acetate	53.	0.1 M Tris pH 8.5	53.	45% v/v (+/-)-2-Methyl-2,4-pentanediol
54.	0.05 M Calcium chloride dihydrate	54.	0.1 M BIS-TRIS pH 6.5	54.	30% v/v Polyethylene glycol monomethyl ether 550
55.	0.05 M Magnesium chloride hexahydrate	55.	0.1 M HEPES pH 7.5	55.	30% v/v Polyethylene glycol monomethyl ether 550
56.	0.2 M Potassium chloride	56.	0.05 M HEPES pH 7.5	56.	35% v/v Pentaerythritol propoxylate (5/4 PO/OH)
57.	0.05 M Ammonium sulfate	57.	0.05 M BIS-TRIS pH 6.5	57.	30% v/v Pentaerythritol ethoxylate (15/4 EO/OH)
58.	None	58.	0.1 M BIS-TRIS pH 6.5	58.	45% v/v Polypropylene glycol P 400
59.	0.02 M Magnesium chloride hexahydrate	59.	0.1 M HEPES pH 7.5	59.	22% w/v Poly(acrylic acid sodium salt) 5,100
60.	0.01 M Cobalt(II) chloride hexahydrate	60.	0.1 M Tris pH 8.5	60.	20% w/v Polyvinylpyrrolidone K 15
61.	0.2 M L-Proline	61.	0.1 M HEPES pH 7.5	61.	10% w/v Polyethylene glycol 3,350
62.	0.2 M Trimethylamine N-oxide dihydrate	62.	0.1 M Tris pH 8.5	62.	20% w/v Polyethylene glycol monomethyl ether 2,000
63.	5% v/v Tacsimate™ pH 7.0	63.	0.1 M HEPES pH 7.0	63.	10% w/v Polyethylene glycol monomethyl ether 5,000
64.	0.005 M Cobalt(II) chloride hexahydrate 0.005 M Nickel(II) chloride hexahydrate 0.005 M Cadmium chloride hydrate 0.005 M Magnesium chloride hexahydrate	64.	0.1 M HEPES pH 7.5	64.	12% w/v Polyethylene glycol 3,350
65.	0.1 M Ammonium acetate	65.	0.1 M BIS-TRIS pH 5.5	65.	17% w/v Polyethylene glycol 10,000
66.	0.2 M Ammonium sulfate	66.	0.1 M BIS-TRIS pH 5.5	66.	25% w/v Polyethylene glycol 3,350
67.	0.2 M Ammonium sulfate	67.	0.1 M BIS-TRIS pH 6.5	67.	25% w/v Polyethylene glycol 3,350
68.	0.2 M Ammonium sulfate	68.	0.1 M HEPES pH 7.5	68.	25% w/v Polyethylene glycol 3,350
69.	0.2 M Ammonium sulfate	69.	0.1 M Tris pH 8.5	69.	25% w/v Polyethylene glycol 3,350
70.	0.2 M Sodium chloride	70.	0.1 M BIS-TRIS pH 5.5	70.	25% w/v Polyethylene glycol 3,350
71.	0.2 M Sodium chloride	71.	0.1 M BIS-TRIS pH 6.5	71.	25% w/v Polyethylene glycol 3,350
72.	0.2 M Sodium chloride	72.	0.1 M HEPES pH 7.5	72.	25% w/v Polyethylene glycol 3,350
73.	0.2 M Sodium chloride	73.	0.1 M Tris pH 8.5	73.	25% w/v Polyethylene glycol 3,350
74.	0.2 M Lithium sulfate monohydrate	74.	0.1 M BIS-TRIS pH 5.5	74.	25% w/v Polyethylene glycol 3,350
75.	0.2 M Lithium sulfate monohydrate	75.	0.1 M BIS-TRIS pH 6.5	75.	25% w/v Polyethylene glycol 3,350
76.	0.2 M Lithium sulfate monohydrate	76.	0.1 M HEPES pH 7.5	76.	25% w/v Polyethylene glycol 3,350
77.	0.2 M Lithium sulfate monohydrate	77.	0.1 M Tris pH 8.5	77.	25% w/v Polyethylene glycol 3,350
78.	0.2 M Ammonium acetate	78.	0.1 M BIS-TRIS pH 5.5	78.	25% w/v Polyethylene glycol 3,350
79.	0.2 M Ammonium acetate	79.	0.1 M BIS-TRIS pH 6.5	79.	25% w/v Polyethylene glycol 3,350
80.	0.2 M Ammonium acetate	80.	0.1 M HEPES pH 7.5	80.	25% w/v Polyethylene glycol 3,350
81.	0.2 M Ammonium acetate	81.	0.1 M Tris pH 8.5	81.	25% w/v Polyethylene glycol 3,350
82.	0.2 M Magnesium chloride hexahydrate	82.	0.1 M BIS-TRIS pH 5.5	82.	25% w/v Polyethylene glycol 3,350
83.	0.2 M Magnesium chloride hexahydrate	83.	0.1 M BIS-TRIS pH 6.5	83.	25% w/v Polyethylene glycol 3,350
84.	0.2 M Magnesium chloride hexahydrate	84.	0.1 M HEPES pH 7.5	84.	25% w/v Polyethylene glycol 3,350
85.	0.2 M Magnesium chloride hexahydrate	85.	0.1 M Tris pH 8.5	85.	25% w/v Polyethylene glycol 3,350
86.	0.2 M Potassium sodium tartrate tetrahydrate	86.	None	86.	20% w/v Polyethylene glycol 3,350
87.	0.2 M Sodium malonate pH 7.0	87.	None	87.	20% w/v Polyethylene glycol 3,350
88.	0.2 M Ammonium citrate tribasic pH 7.0	88.	None	88.	20% w/v Polyethylene glycol 3,350
89.	0.1 M Succinic acid pH 7.0	89.	None	89.	15% w/v Polyethylene glycol 3,350
90.	0.2 M Sodium formate	90.	None	90.	20% w/v Polyethylene glycol 3,350
91.	0.15 M DL-Malic acid pH 7.0	91.	None	91.	20% w/v Polyethylene glycol 3,350
92.	0.1 M Magnesium formate dihydrate	92.	None	92.	15% w/v Polyethylene glycol 3,350
93.	0.05 M Zinc acetate dihydrate	93.	None	93.	20% w/v Polyethylene glycol 3,350
94.	0.2 M Sodium citrate tribasic dihydrate	94.	None	94.	20% w/v Polyethylene glycol 3,350
95.	0.1 M Potassium thiocyanate	95.	None	95.	30% w/v Polyethylene glycol monomethyl ether 2,000
96.	0.15 M Potassium bromide	96.	None	96.	30% w/v Polyethylene glycol monomethyl ether 2,000

◊ Buffer pH is that of a 1.0 M stock prior to dilution with other reagent components: pH with HCl or NaOH.

Index contains ninety-six unique reagents. To determine the formulation of each reagent, simply read across the page.

34 Journey
Aliso Viejo, CA 92656-3317 U.S.A.
Tel: (949) 425-1321 • Fax: (949) 425-1611
E-mail: tech@hrmail.com
Website: www.hamptonresearch.com

HAMPTON RESEARCH
Solutions for Crystal Growth

© 1991 - 2016 Hampton Research Corp. all rights reserved
Printed in the United States of America. This guide or parts thereof may not be reproduced in any form without the written permission of the publishers.

Well #	Salt	Well #	Buffer ◊	Well #	Precipitant
1. (A1)	0.02 M Calcium chloride dihydrate	1. (A1)	0.1 M Sodium acetate trihydrate pH 4.6	1. (A1)	30% v/v (+/-)-2-Methyl-2,4-pentanediol
2. (A2)	None	2. (A2)	None	2. (A2)	0.4 M Potassium sodium tartrate tetrahydrate
3. (A3)	None	3. (A3)	None	3. (A3)	0.4 M Ammonium phosphate monobasic
4. (A4)	None	4. (A4)	0.1 M TRIS hydrochloride pH 8.5	4. (A4)	2.0 M Ammonium sulfate
5. (A5)	0.2 M Sodium citrate tribasic dihydrate	5. (A5)	0.1 M HEPES sodium pH 7.5	5. (A5)	30% v/v (+/-)-2-Methyl-2,4-pentanediol
6. (A6)	0.2 M Magnesium chloride hexahydrate	6. (A6)	0.1 M TRIS hydrochloride pH 8.5	6. (A6)	30% w/v Polyethylene glycol 4,000
7. (A7)	None	7. (A7)	0.1 M Sodium cacodylate trihydrate pH 6.5	7. (A7)	1.4 M Sodium acetate trihydrate
8. (A8)	0.2 M Sodium citrate tribasic dihydrate	8. (A8)	0.1 M Sodium cacodylate trihydrate pH 6.5	8. (A8)	30% v/v 2-Propanol
9. (A9)	0.2 M Ammonium acetate	9. (A9)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	9. (A9)	30% w/v Polyethylene glycol 4,000
10. (A10)	0.2 M Ammonium acetate	10. (A10)	0.1 M Sodium acetate trihydrate pH 4.6	10. (A10)	30% w/v Polyethylene glycol 4,000
11. (A11)	None	11. (A11)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	11. (A11)	1.0 M Ammonium phosphate monobasic
12. (A12)	0.2 M Magnesium chloride hexahydrate	12. (A12)	0.1 M HEPES sodium pH 7.5	12. (A12)	30% v/v 2-Propanol
13. (B1)	0.2 M Sodium citrate tribasic dihydrate	13. (B1)	0.1 M TRIS hydrochloride pH 8.5	13. (B1)	30% v/v Polyethylene glycol 400
14. (B2)	0.2 M Calcium chloride dihydrate	14. (B2)	0.1 M HEPES sodium pH 7.5	14. (B2)	28% v/v Polyethylene glycol 400
15. (B3)	0.2 M Ammonium sulfate	15. (B3)	0.1 M Sodium cacodylate trihydrate pH 6.5	15. (B3)	30% w/v Polyethylene glycol 8,000
16. (B4)	None	16. (B4)	0.1 M HEPES sodium pH 7.5	16. (B4)	1.5 M Lithium sulfate monohydrate
17. (B5)	0.2 M Lithium sulfate monohydrate	17. (B5)	0.1 M TRIS hydrochloride pH 8.5	17. (B5)	30% w/v Polyethylene glycol 4,000
18. (B6)	0.2 M Magnesium acetate tetrahydrate	18. (B6)	0.1 M Sodium cacodylate trihydrate pH 6.5	18. (B6)	20% w/v Polyethylene glycol 8,000
19. (B7)	0.2 M Ammonium acetate	19. (B7)	0.1 M TRIS hydrochloride pH 8.5	19. (B7)	30% v/v 2-Propanol
20. (B8)	0.2 M Ammonium sulfate	20. (B8)	0.1 M Sodium acetate trihydrate pH 4.6	20. (B8)	25% w/v Polyethylene glycol 4,000
21. (B9)	0.2 M Magnesium acetate tetrahydrate	21. (B9)	0.1 M Sodium cacodylate trihydrate pH 6.5	21. (B9)	30% v/v (+/-)-2-Methyl-2,4-pentanediol
22. (B10)	0.2 M Sodium acetate trihydrate	22. (B10)	0.1 M TRIS hydrochloride pH 8.5	22. (B10)	30% w/v Polyethylene glycol 4,000
23. (B11)	0.2 M Magnesium chloride hexahydrate	23. (B11)	0.1 M HEPES sodium pH 7.5	23. (B11)	30% v/v Polyethylene glycol 400
24. (B12)	0.2 M Calcium chloride dihydrate	24. (B12)	0.1 M Sodium acetate trihydrate pH 4.6	24. (B12)	20% v/v 2-Propanol
25. (C1)	None	25. (C1)	0.1 M Imidazole pH 6.5	25. (C1)	1.0 M Sodium acetate trihydrate
26. (C2)	0.2 M Ammonium acetate	26. (C2)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	26. (C2)	30% v/v (+/-)-2-Methyl-2,4-pentanediol
27. (C3)	0.2 M Sodium citrate tribasic dihydrate	27. (C3)	0.1 M HEPES sodium pH 7.5	27. (C3)	20% v/v 2-Propanol
28. (C4)	0.2 M Sodium acetate trihydrate	28. (C4)	0.1 M Sodium cacodylate trihydrate pH 6.5	28. (C4)	30% w/v Polyethylene glycol 8,000
29. (C5)	None	29. (C5)	0.1 M HEPES sodium pH 7.5	29. (C5)	0.8 M Potassium sodium tartrate tetrahydrate
30. (C6)	0.2 M Ammonium sulfate	30. (C6)	None	30. (C6)	30% w/v Polyethylene glycol 8,000
31. (C7)	0.2 M Ammonium sulfate	31. (C7)	None	31. (C7)	30% w/v Polyethylene glycol 4,000
32. (C8)	None	32. (C8)	None	32. (C8)	2.0 M Ammonium sulfate
33. (C9)	None	33. (C9)	None	33. (C9)	4.0 M Sodium formate
34. (C10)	None	34. (C10)	0.1 M Sodium acetate trihydrate pH 4.6	34. (C10)	2.0 M Sodium formate
35. (C11)	None	35. (C11)	0.1 M HEPES sodium pH 7.5	35. (C11)	0.8 M Sodium phosphate monobasic monohydrate, 0.8 M Potassium phosphate monobasic
36. (C12)	None	36. (C12)	0.1 M TRIS hydrochloride pH 8.5	36. (C12)	8% w/v Polyethylene glycol 8,000
37. (D1)	None	37. (D1)	0.1 M Sodium acetate trihydrate pH 4.6	37. (D1)	8% w/v Polyethylene glycol 4,000
38. (D2)	None	38. (D2)	0.1 M HEPES sodium pH 7.5	38. (D2)	1.4 M Sodium citrate tribasic dihydrate
39. (D3)	None	39. (D3)	0.1 M HEPES sodium pH 7.5	39. (D3)	2% v/v Polyethylene glycol 400, 2.0 M Ammonium sulfate
40. (D4)	None	40. (D4)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	40. (D4)	20% v/v 2-Propanol, 20% w/v Polyethylene glycol 4,000
41. (D5)	None	41. (D5)	0.1 M HEPES sodium pH 7.5	41. (D5)	10% v/v 2-Propanol, 20% w/v Polyethylene glycol 4,000
42. (D6)	0.05 M Potassium phosphate monobasic	42. (D6)	None	42. (D6)	20% w/v Polyethylene glycol 8,000
43. (D7)	None	43. (D7)	None	43. (D7)	30% w/v Polyethylene glycol 1,500
44. (D8)	None	44. (D8)	None	44. (D8)	0.2 M Magnesium formate dihydrate
45. (D9)	0.2 M Zinc acetate dihydrate	45. (D9)	0.1 M Sodium cacodylate trihydrate pH 6.5	45. (D9)	18% w/v Polyethylene glycol 8,000
46. (D10)	0.2 M Calcium acetate hydrate	46. (D10)	0.1 M Sodium cacodylate trihydrate pH 6.5	46. (D10)	18% w/v Polyethylene glycol 8,000
47. (D11)	None	47. (D11)	0.1 M Sodium acetate trihydrate pH 4.6	47. (D11)	2.0 M Ammonium sulfate
48. (D12)	None	48. (D12)	0.1 M TRIS hydrochloride pH 8.5	48. (D12)	2.0 M Ammonium phosphate monobasic

◊ Buffer pH is that of a 1.0 M stock prior to dilution with other reagent components: pH with HCl or NaOH.

Crystal Screen™ (Deep Well Block) contains forty-eight unique reagents beginning at position A1.
To determine the formulation of each reagent, simply read across the page.

HAMPTON RESEARCH
Solutions for Crystal Growth

Well #	Salt	Well #	Buffer ◊	Well #	Precipitant
49. (E1)	2.0 M Sodium chloride	49. (E1)	None	49. (E1)	10% w/v Polyethylene glycol 6,000
50. (E2)	0.5 M Sodium chloride, 0.01 M Magnesium chloride hexahydrate	50. (E2)	None	50. (E2)	0.01 M Hexadecyltrimethylammonium bromide
51. (E3)	None	51. (E3)	None	51. (E3)	25% v/v Ethylene glycol
52. (E4)	None	52. (E4)	None	52. (E4)	35% v/v 1,4-Dioxane
53. (E5)	2.0 M Ammonium sulfate	53. (E5)	None	53. (E5)	5% v/v 2-Propanol
54. (E6)	None	54. (E6)	None	54. (E6)	1.0 M Imidazole pH 7.0
55. (E7)	None	55. (E7)	None	55. (E7)	10% w/v Polyethylene glycol 1,000, 10% w/v Polyethylene glycol 8,000
56. (E8)	1.5 M Sodium chloride	56. (E8)	None	56. (E8)	10% v/v Ethanol
57. (E9)	None	57. (E9)	0.1 M Sodium acetate trihydrate pH 4.6	57. (E9)	2.0 M Sodium chloride
58. (E10)	0.2 M Sodium chloride	58. (E10)	0.1 M Sodium acetate trihydrate pH 4.6	58. (E10)	30% v/v (+/-)-2-Methyl-2,4-pentanediol
59. (E11)	0.01 M Cobalt(II) chloride hexahydrate	59. (E11)	0.1 M Sodium acetate trihydrate pH 4.6	59. (E11)	1.0 M 1,6-Hexanediol
60. (E12)	0.1 M Cadmium chloride hydrate	60. (E12)	0.1 M Sodium acetate trihydrate pH 4.6	60. (E12)	30% v/v Polyethylene glycol 400
61. (F1)	0.2 M Ammonium sulfate	61. (F1)	0.1 M Sodium acetate trihydrate pH 4.6	61. (F1)	30% w/v Polyethylene glycol monomethyl ether 2,000
62. (F2)	0.2 M Potassium sodium tartrate tetrahydrate	62. (F2)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	62. (F2)	2.0 M Ammonium sulfate
63. (F3)	0.5 M Ammonium sulfate	63. (F3)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	63. (F3)	1.0 M Lithium sulfate monohydrate
64. (F4)	0.5 M Sodium chloride	64. (F4)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	64. (F4)	2% v/v Ethylene imine polymer
65. (F5)	None	65. (F5)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	65. (F5)	35% v/v tert-Butanol
66. (F6)	0.01 M Iron(III) chloride hexahydrate	66. (F6)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	66. (F6)	10% v/v Jeffamine® M-600®
67. (F7)	None	67. (F7)	0.1 M Sodium citrate tribasic dihydrate pH 5.6	67. (F7)	2.5 M 1,6-Hexanediol
68. (F8)	None	68. (F8)	0.1 M MES monohydrate pH 6.5	68. (F8)	1.6 M Magnesium sulfate heptahydrate
69. (F9)	0.1 M Sodium phosphate monobasic monohydrate, 0.1 M Potassium phosphate monobasic	69. (F9)	0.1 M MES monohydrate pH 6.5	69. (F9)	2.0 M Sodium chloride
70. (F10)	None	70. (F10)	0.1 M MES monohydrate pH 6.5	70. (F10)	12% w/v Polyethylene glycol 20,000
71. (F11)	1.6 M Ammonium sulfate	71. (F11)	0.1 M MES monohydrate pH 6.5	71. (F11)	10% v/v 1,4-Dioxane
72. (F12)	0.05 M Cesium chloride	72. (F12)	0.1 M MES monohydrate pH 6.5	72. (F12)	30% v/v Jeffamine® M-600®
73. (G1)	0.01 M Cobalt(II) chloride hexahydrate	73. (G1)	0.1 M MES monohydrate pH 6.5	73. (G1)	1.8 M Ammonium sulfate
74. (G2)	0.2 M Ammonium sulfate	74. (G2)	0.1 M MES monohydrate pH 6.5	74. (G2)	30% w/v Polyethylene glycol monomethyl ether 5,000
75. (G3)	0.01 M Zinc sulfate heptahydrate	75. (G3)	0.1 M MES monohydrate pH 6.5	75. (G3)	25% v/v Polyethylene glycol monomethyl ether 550
76. (G4)	None	76. (G4)	None	76. (G4)	1.6 M Sodium citrate tribasic dihydrate pH 6.5
77. (G5)	0.5 M Ammonium sulfate	77. (G5)	0.1 M HEPES pH 7.5	77. (G5)	30% v/v (+/-)-2-Methyl-2,4-pentanediol
78. (G6)	None	78. (G6)	0.1 M HEPES pH 7.5	78. (G6)	10% w/v Polyethylene glycol 6,000, 5% v/v (+/-)-2-Methyl-2,4-pentanediol
79. (G7)	None	79. (G7)	0.1 M HEPES pH 7.5	79. (G7)	20% v/v Jeffamine® M-600®
80. (G8)	0.1 M Sodium chloride	80. (G8)	0.1 M HEPES pH 7.5	80. (G8)	1.6 M Ammonium sulfate
81. (G9)	None	81. (G9)	0.1 M HEPES pH 7.5	81. (G9)	2.0 M Ammonium formate
82. (G10)	0.05 M Cadmium sulfate hydrate	82. (G10)	0.1 M HEPES pH 7.5	82. (G10)	1.0 M Sodium acetate trihydrate
83. (G11)	None	83. (G11)	0.1 M HEPES pH 7.5	83. (G11)	70% v/v (+/-)-2-Methyl-2,4-pentanediol
84. (G12)	None	84. (G12)	0.1 M HEPES pH 7.5	84. (G12)	4.3 M Sodium chloride
85. (H1)	None	85. (H1)	0.1 M HEPES pH 7.5	85. (H1)	10% w/v Polyethylene glycol 8,000, 8% v/v Ethylene glycol
86. (H2)	None	86. (H2)	0.1 M HEPES pH 7.5	86. (H2)	20% w/v Polyethylene glycol 10,000
87. (H3)	0.2 M Magnesium chloride hexahydrate	87. (H3)	0.1 M Tris pH 8.5	87. (H3)	3.4 M 1,6-Hexanediol
88. (H4)	None	88. (H4)	0.1 M Tris pH 8.5	88. (H4)	25% v/v tert-Butanol
89. (H5)	0.01 M Nickel(II) chloride hexahydrate	89. (H5)	0.1 M Tris pH 8.5	89. (H5)	1.0 M Lithium sulfate monohydrate
90. (H6)	1.5 M Ammonium sulfate	90. (H6)	0.1 M Tris pH 8.5	90. (H6)	12% v/v Glycerol
91. (H7)	0.2 M Ammonium phosphate monobasic	91. (H7)	0.1 M Tris pH 8.5	91. (H7)	50% v/v (+/-)-2-Methyl-2,4-pentanediol
92. (H8)	None	92. (H8)	0.1 M Tris pH 8.5	92. (H8)	20% v/v Ethanol
93. (H9)	0.01 M Nickel(II) chloride hexahydrate	93. (H9)	0.1 M Tris pH 8.5	93. (H9)	20% w/v Polyethylene glycol monomethyl ether 2,000
94. (H10)	0.1 M Sodium chloride	94. (H10)	0.1 M BICINE pH 9.0	94. (H10)	20% v/v Polyethylene glycol monomethyl ether 550
95. (H11)	None	95. (H11)	0.1 M BICINE pH 9.0	95. (H11)	2.0 M Magnesium chloride hexahydrate
96. (H12)	None	96. (H12)	0.1 M BICINE pH 9.0	96. (H12)	2% v/v 1,4-Dioxane, 10% w/v Polyethylene glycol 20,000

◊ Buffer pH is that of a 1.0 M (0.5 M for MES) stock prior to dilution with other reagent components: pH with HCl or NaOH.

Crystal Screen 2™ (Deep Well Block) contains forty-eight unique reagents beginning at position E1.
To determine the formulation of each reagent, simply read across the page.

34 Journey
Aliso Viejo, CA 92656-3317 U.S.A.
Tel: (949) 425-1321 • Fax: (949) 425-1611
E-mail: tech@hrmail.com
Website: www.hamptonresearch.com

HAMPTON RESEARCH
Solutions for Crystal Growth

© 1991-2018 Hampton Research Corp. all rights reserved
Printed in the United States of America. This guide or parts thereof may not be reproduced in any form without the written permission of the publishers.

D.2. Optimisation tray for E5 condition from PEG ION broad screen:

Sample:TcBDF3-BI2536
Concentration:20mg/ml
Tray Number:2
Title:PEG ION Screen E5
Date:Friday, May 17, 2019 2:04 AM

0.07 M Sodium Malonate 7 % PEG 3350	0.07 M Sodium Malonate 8 % PEG 3350	0.07 M Sodium Malonate 9 % PEG 3350	0.07 M Sodium Malonate 10 % PEG 3350	0.07 M Sodium Malonate 11 % PEG 3350	0.07 M Sodium Malonate 12 % PEG 3350	0.07 M Sodium Malonate 13 % PEG 3350	0.07 M Sodium Malonate 14 % PEG 3350	0.07 M Sodium Malonate 15 % PEG 3350	0.07 M Sodium Malonate 16 % PEG 3350	0.07 M Sodium Malonate 17 % PEG 3350	0.07 M Sodium Malonate 18 % PEG 3350
0.08 M Sodium Malonate 7 % PEG 3350	0.08 M Sodium Malonate 8 % PEG 3350	0.08 M Sodium Malonate 9 % PEG 3350	0.08 M Sodium Malonate 10 % PEG 3350	0.08 M Sodium Malonate 11 % PEG 3350	0.08 M Sodium Malonate 12 % PEG 3350	0.08 M Sodium Malonate 13 % PEG 3350	0.08 M Sodium Malonate 14 % PEG 3350	0.08 M Sodium Malonate 15 % PEG 3350	0.08 M Sodium Malonate 16 % PEG 3350	0.08 M Sodium Malonate 17 % PEG 3350	0.08 M Sodium Malonate 18 % PEG 3350
0.09 M Sodium Malonate 7 % PEG 3350	0.09 M Sodium Malonate 8 % PEG 3350	0.09 M Sodium Malonate 9 % PEG 3350	0.09 M Sodium Malonate 10 % PEG 3350	0.09 M Sodium Malonate 11 % PEG 3350	0.09 M Sodium Malonate 12 % PEG 3350	0.09 M Sodium Malonate 13 % PEG 3350	0.09 M Sodium Malonate 14 % PEG 3350	0.09 M Sodium Malonate 15 % PEG 3350	0.09 M Sodium Malonate 16 % PEG 3350	0.09 M Sodium Malonate 17 % PEG 3350	0.09 M Sodium Malonate 18 % PEG 3350
0.10 M Sodium Malonate 7 % PEG 3350	0.10 M Sodium Malonate 8 % PEG 3350	0.10 M Sodium Malonate 9 % PEG 3350	0.10 M Sodium Malonate 10 % PEG 3350	0.10 M Sodium Malonate 11 % PEG 3350	0.10 M Sodium Malonate 12 % PEG 3350	0.10 M Sodium Malonate 13 % PEG 3350	0.10 M Sodium Malonate 14 % PEG 3350	0.10 M Sodium Malonate 15 % PEG 3350	0.10 M Sodium Malonate 16 % PEG 3350	0.10 M Sodium Malonate 17 % PEG 3350	0.10 M Sodium Malonate 18 % PEG 3350
0.11 M Sodium Malonate 7 % PEG 3350	0.11 M Sodium Malonate 8 % PEG 3350	0.11 M Sodium Malonate 9 % PEG 3350	0.11 M Sodium Malonate 10 % PEG 3350	0.11 M Sodium Malonate 11 % PEG 3350	0.11 M Sodium Malonate 12 % PEG 3350	0.11 M Sodium Malonate 13 % PEG 3350	0.11 M Sodium Malonate 14 % PEG 3350	0.11 M Sodium Malonate 15 % PEG 3350	0.11 M Sodium Malonate 16 % PEG 3350	0.11 M Sodium Malonate 17 % PEG 3350	0.11 M Sodium Malonate 18 % PEG 3350
0.12 M Sodium Malonate 7 % PEG 3350	0.12 M Sodium Malonate 8 % PEG 3350	0.12 M Sodium Malonate 9 % PEG 3350	0.12 M Sodium Malonate 10 % PEG 3350	0.12 M Sodium Malonate 11 % PEG 3350	0.12 M Sodium Malonate 12 % PEG 3350	0.12 M Sodium Malonate 13 % PEG 3350	0.12 M Sodium Malonate 14 % PEG 3350	0.12 M Sodium Malonate 15 % PEG 3350	0.12 M Sodium Malonate 16 % PEG 3350	0.12 M Sodium Malonate 17 % PEG 3350	0.12 M Sodium Malonate 18 % PEG 3350
0.13 M Sodium Malonate 7 % PEG 3350	0.13 M Sodium Malonate 8 % PEG 3350	0.13 M Sodium Malonate 9 % PEG 3350	0.13 M Sodium Malonate 10 % PEG 3350	0.13 M Sodium Malonate 11 % PEG 3350	0.13 M Sodium Malonate 12 % PEG 3350	0.13 M Sodium Malonate 13 % PEG 3350	0.13 M Sodium Malonate 14 % PEG 3350	0.13 M Sodium Malonate 15 % PEG 3350	0.13 M Sodium Malonate 16 % PEG 3350	0.13 M Sodium Malonate 17 % PEG 3350	0.13 M Sodium Malonate 18 % PEG 3350
0.14 M Sodium Malonate 7 % PEG 3350	0.14 M Sodium Malonate 8 % PEG 3350	0.14 M Sodium Malonate 9 % PEG 3350	0.14 M Sodium Malonate 10 % PEG 3350	0.14 M Sodium Malonate 11 % PEG 3350	0.14 M Sodium Malonate 12 % PEG 3350	0.14 M Sodium Malonate 13 % PEG 3350	0.14 M Sodium Malonate 14 % PEG 3350	0.14 M Sodium Malonate 15 % PEG 3350	0.14 M Sodium Malonate 16 % PEG 3350	0.14 M Sodium Malonate 17 % PEG 3350	0.14 M Sodium Malonate 18 % PEG 3350

D.3. Optimisation tray for E8 condition from Crystal screen broad screen

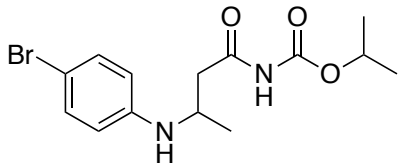
Sample:TcBDF3-BI2536
Concentration:20mg/ml
Tray Number:1
Title:Crystal Screen E8
Date:Friday, May 17, 2019 1:57 AM

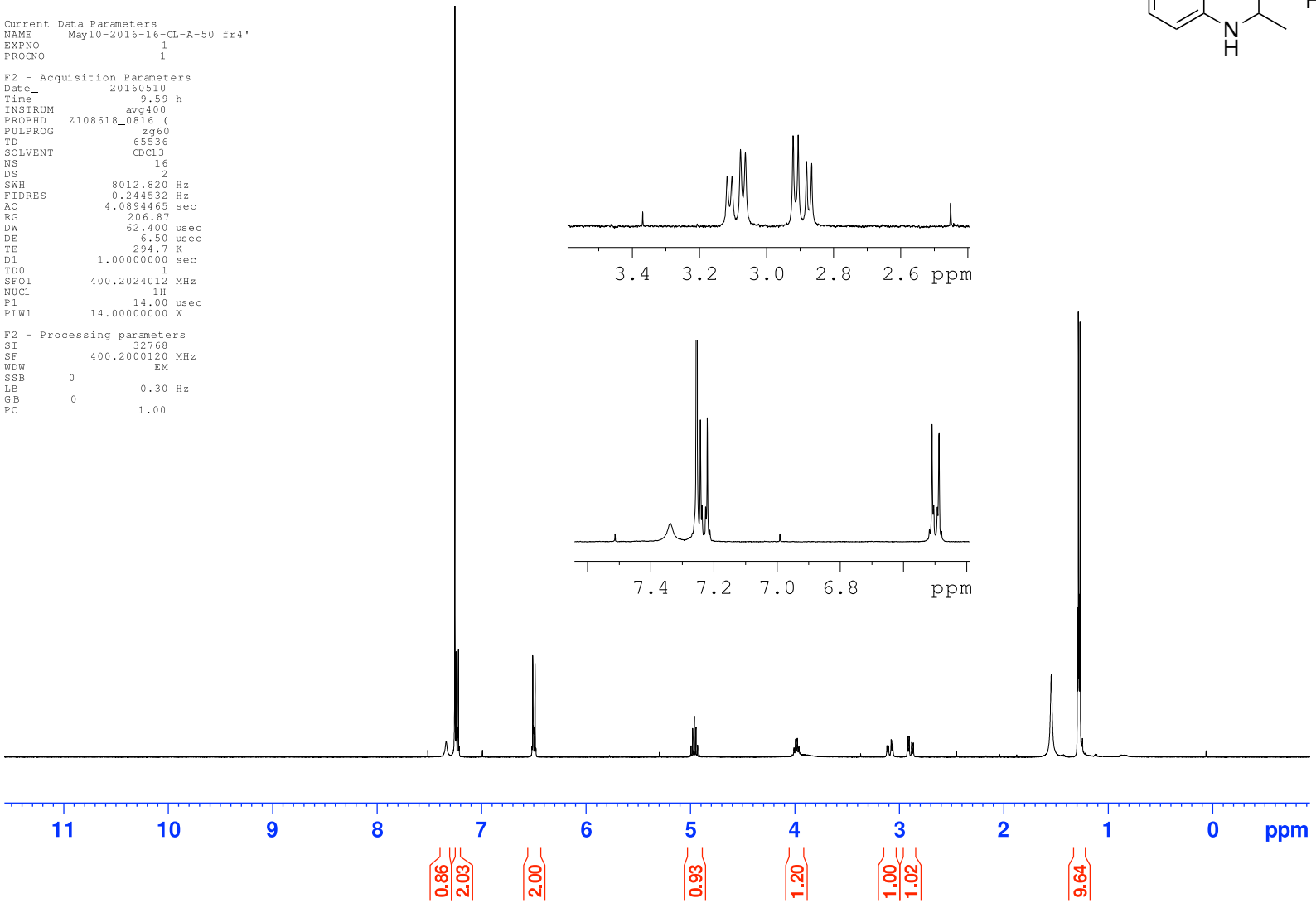
0.9 M NaCl 7 % Ethanol	1.0 M NaCl 7 % Ethanol	1.1 M NaCl 7 % Ethanol	1.2 M NaCl 7 % Ethanol	1.3 M NaCl 7 % Ethanol	1.4 M NaCl 7 % Ethanol	1.5 M NaCl 7 % Ethanol	1.6 M NaCl 7 % Ethanol	1.7 M NaCl 7 % Ethanol	1.8 M NaCl 7 % Ethanol	1.9 M NaCl 7 % Ethanol	2.0 M NaCl 7 % Ethanol
0.9 M NaCl 8 % Ethanol	1.0 M NaCl 8 % Ethanol	1.1 M NaCl 8 % Ethanol	1.2 M NaCl 8 % Ethanol	1.3 M NaCl 8 % Ethanol	1.4 M NaCl 8 % Ethanol	1.5 M NaCl 8 % Ethanol	1.6 M NaCl 8 % Ethanol	1.7 M NaCl 8 % Ethanol	1.8 M NaCl 8 % Ethanol	1.9 M NaCl 8 % Ethanol	2.0 M NaCl 8 % Ethanol
0.9 M NaCl 9 % Ethanol	1.0 M NaCl 9 % Ethanol	1.1 M NaCl 9 % Ethanol	1.2 M NaCl 9 % Ethanol	1.3 M NaCl 9 % Ethanol	1.4 M NaCl 9 % Ethanol	1.5 M NaCl 9 % Ethanol	1.6 M NaCl 9 % Ethanol	1.7 M NaCl 9 % Ethanol	1.8 M NaCl 9 % Ethanol	1.9 M NaCl 9 % Ethanol	2.0 M NaCl 9 % Ethanol
0.9 M NaCl 10 % Ethanol	1.0 M NaCl 10 % Ethanol	1.1 M NaCl 10 % Ethanol	1.2 M NaCl 10 % Ethanol	1.3 M NaCl 10 % Ethanol	1.4 M NaCl 10 % Ethanol	1.5 M NaCl 10 % Ethanol	1.6 M NaCl 10 % Ethanol	1.7 M NaCl 10 % Ethanol	1.8 M NaCl 10 % Ethanol	1.9 M NaCl 10 % Ethanol	2.0 M NaCl 10 % Ethanol
0.9 M NaCl 11 % Ethanol	1.0 M NaCl 11 % Ethanol	1.1 M NaCl 11 % Ethanol	1.2 M NaCl 11 % Ethanol	1.3 M NaCl 11 % Ethanol	1.4 M NaCl 11 % Ethanol	1.5 M NaCl 11 % Ethanol	1.6 M NaCl 11 % Ethanol	1.7 M NaCl 11 % Ethanol	1.8 M NaCl 11 % Ethanol	1.9 M NaCl 11 % Ethanol	2.0 M NaCl 11 % Ethanol
0.9 M NaCl 12 % Ethanol	1.0 M NaCl 12 % Ethanol	1.1 M NaCl 12 % Ethanol	1.2 M NaCl 12 % Ethanol	1.3 M NaCl 12 % Ethanol	1.4 M NaCl 12 % Ethanol	1.5 M NaCl 12 % Ethanol	1.6 M NaCl 12 % Ethanol	1.7 M NaCl 12 % Ethanol	1.8 M NaCl 12 % Ethanol	1.9 M NaCl 12 % Ethanol	2.0 M NaCl 12 % Ethanol
0.9 M NaCl 13 % Ethanol	1.0 M NaCl 13 % Ethanol	1.1 M NaCl 13 % Ethanol	1.2 M NaCl 13 % Ethanol	1.3 M NaCl 13 % Ethanol	1.4 M NaCl 13 % Ethanol	1.5 M NaCl 13 % Ethanol	1.6 M NaCl 13 % Ethanol	1.7 M NaCl 13 % Ethanol	1.8 M NaCl 13 % Ethanol	1.9 M NaCl 13 % Ethanol	2.0 M NaCl 13 % Ethanol
0.9 M NaCl 14 % Ethanol	1.0 M NaCl 14 % Ethanol	1.1 M NaCl 14 % Ethanol	1.2 M NaCl 14 % Ethanol	1.3 M NaCl 14 % Ethanol	1.4 M NaCl 14 % Ethanol	1.5 M NaCl 14 % Ethanol	1.6 M NaCl 14 % Ethanol	1.7 M NaCl 14 % Ethanol	1.8 M NaCl 14 % Ethanol	1.9 M NaCl 14 % Ethanol	2.0 M NaCl 14 % Ethanol

Spectra for novel compounds

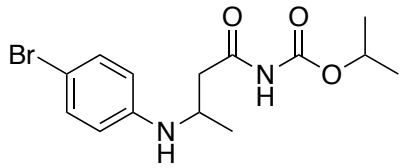
Below are shown the ^{1}H, ^{13}C and ^{19}F of novel compounds.

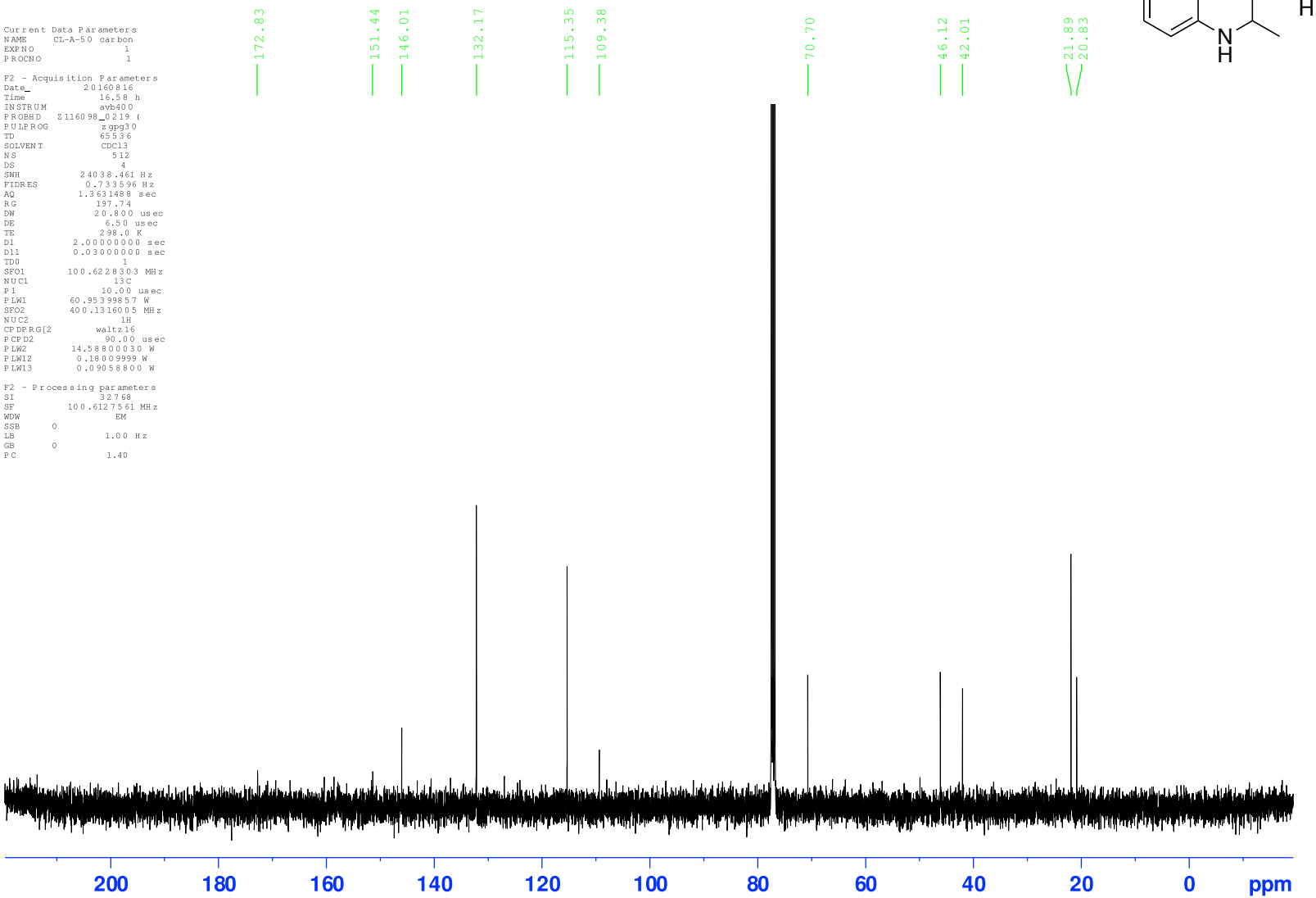
Isopropyl (3-((4-bromophenyl)amino)butanoyl)carbamate (84)





Isopropyl (3-((4-bromophenyl)amino)butanoyl)carbamate (84)



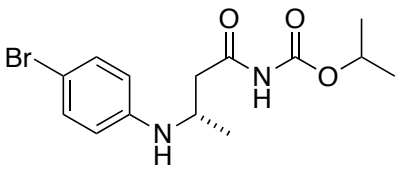


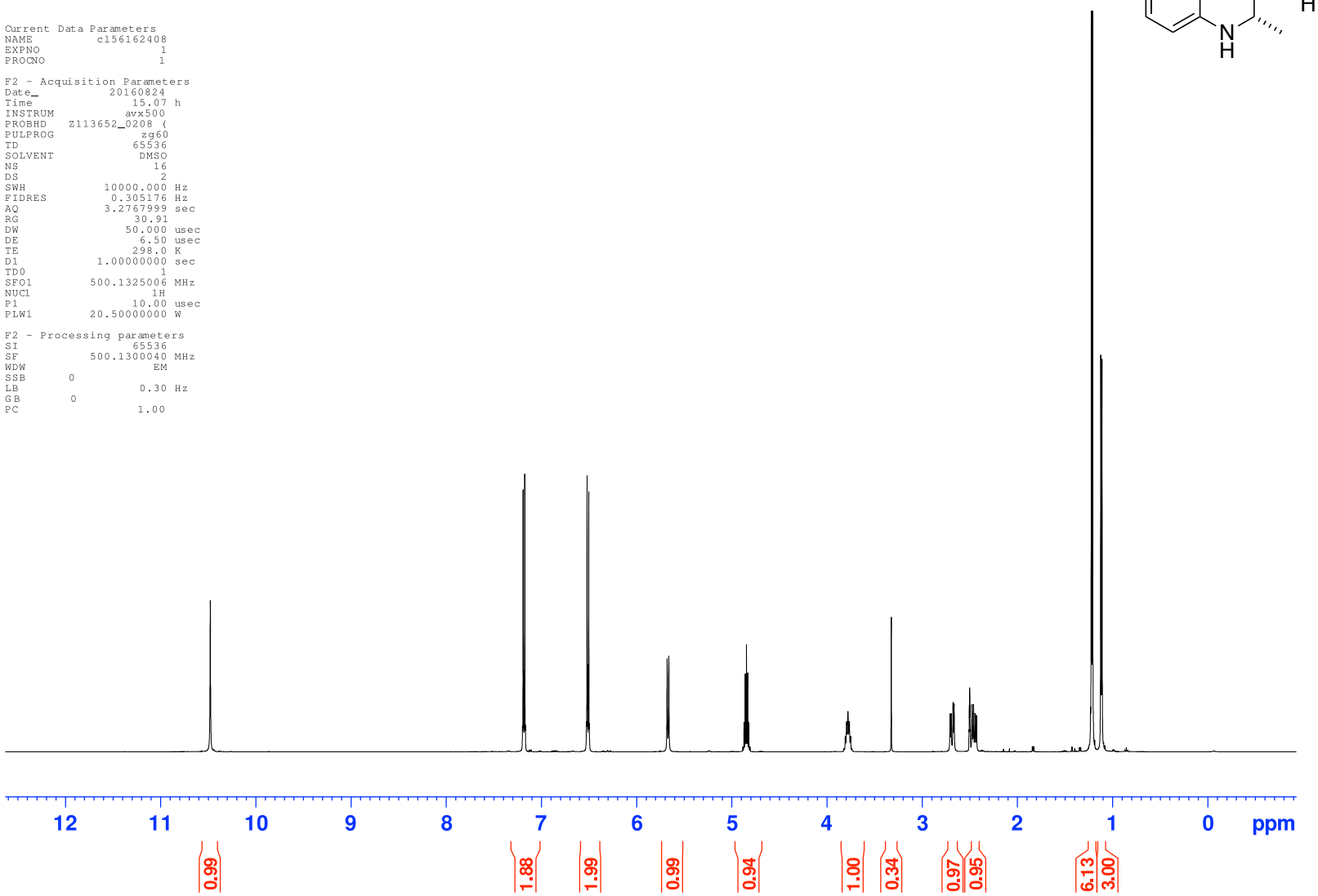
Current Data Parameters
NAME CL-A-50 carbon
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20160816
Time 16.58 h
INSTRUM avb400
PROBHD Z116098_0219 (
PULPROG zgpg30
TD 65536
SOLVENT CDCl3
NS 512
DS 4
SWH 24038.461 Hz
FIDRES 0.733596 Hz
AQ 1.3631488 sec
RG 197.74
DW 20.800 usec
DE 6.50 usec
TE 298.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
SFO1 100.6228303 MHz
NUC1 13C
P1 10.00 usec
PLW1 60.95399857 W
SFO2 400.1316005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 90.00 usec
PLW2 14.58800030 W
PLW12 0.18009999 W
PLW13 0.09058800 W

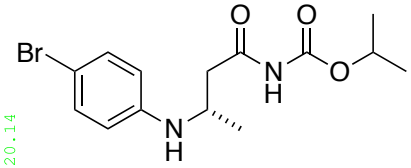
F2 - Processing parameters
SI 32768
SF 100.6127561 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

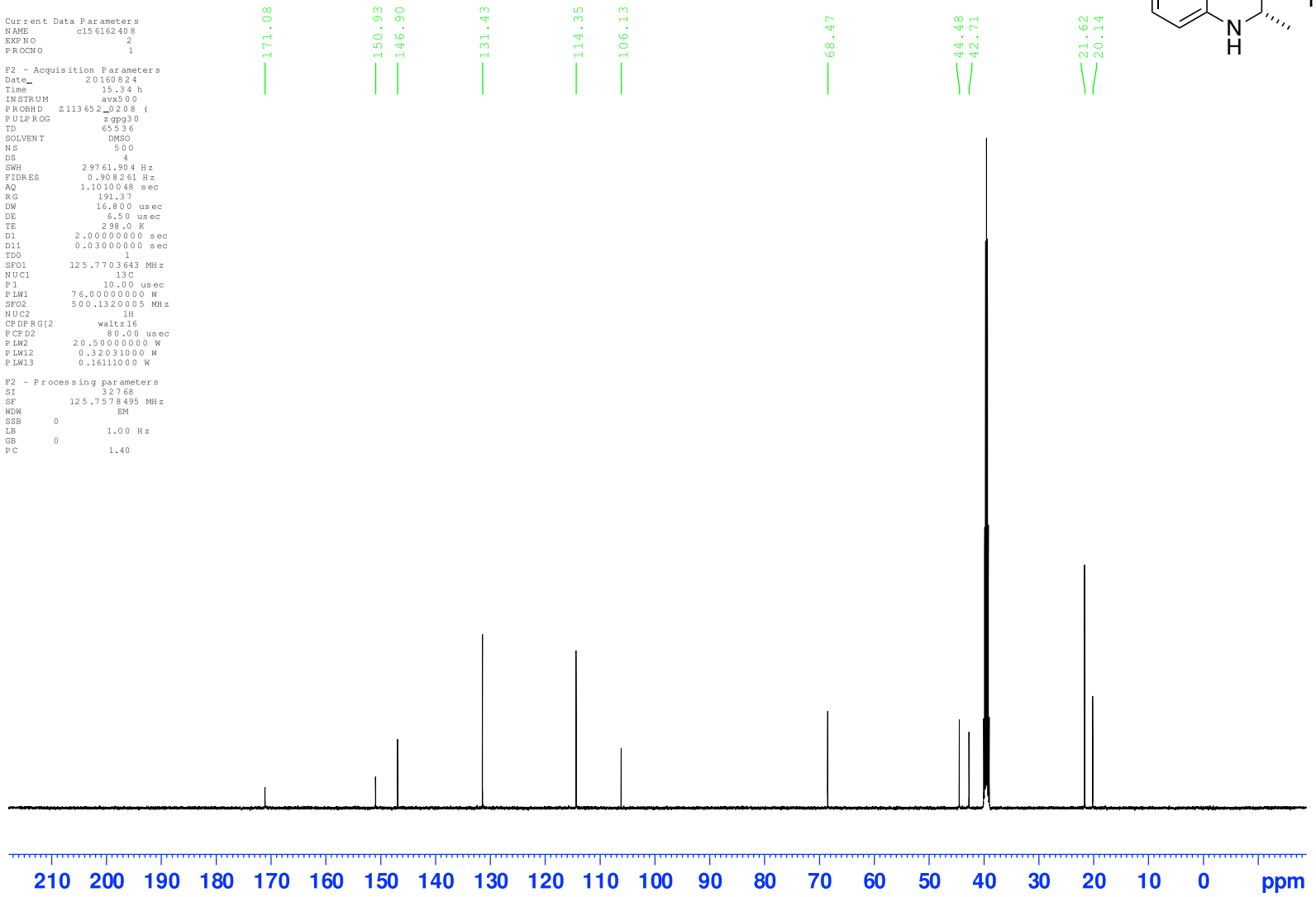
Isopropyl (*S*)-(3-((4'-bromophenyl)amino)butanoyl)carbamate ((*S*)-84)



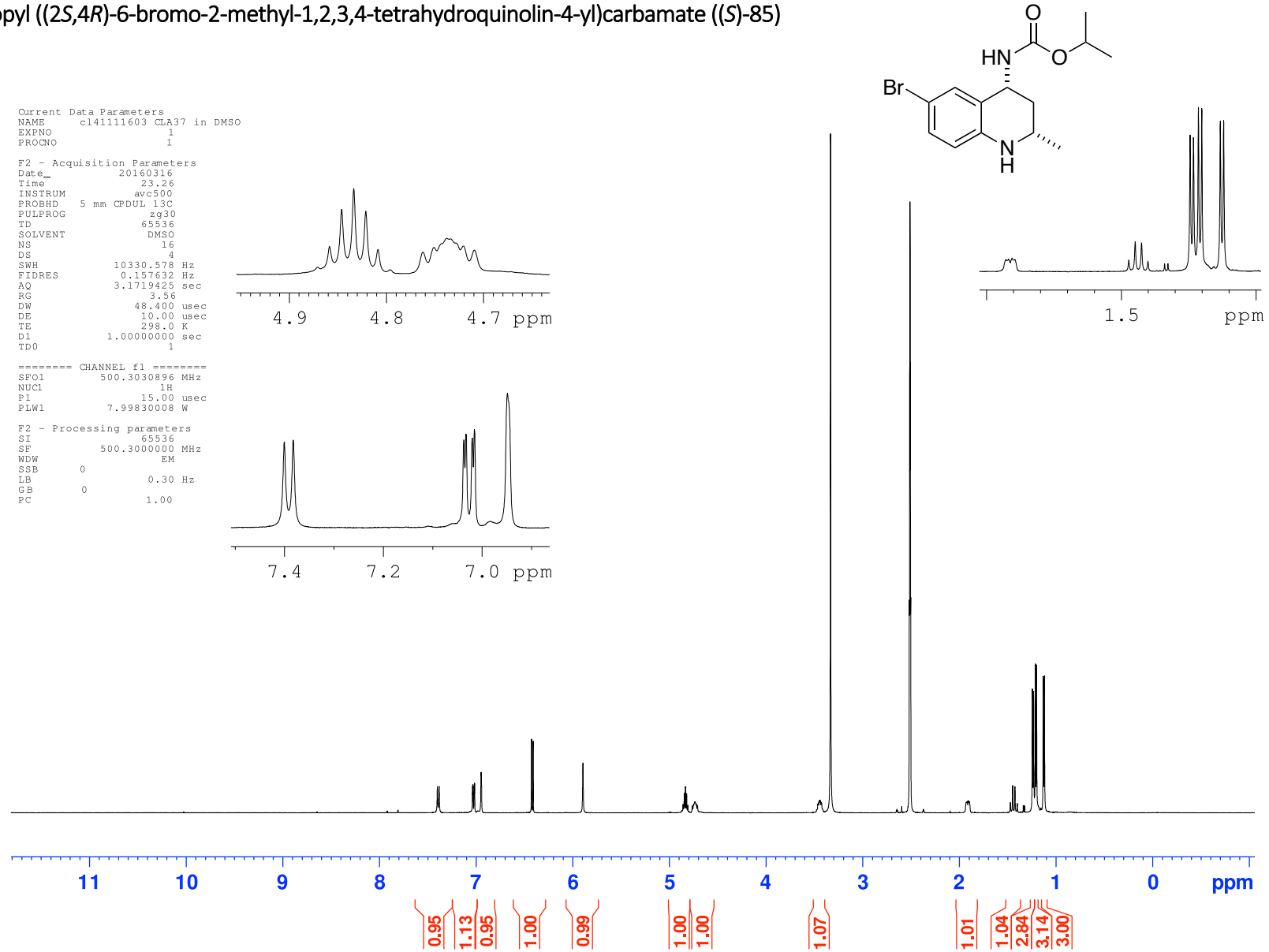


Isopropyl (*S*)-(3-((4'-bromophenyl)amino)butanoyl)carbamate ((*S*)-84)

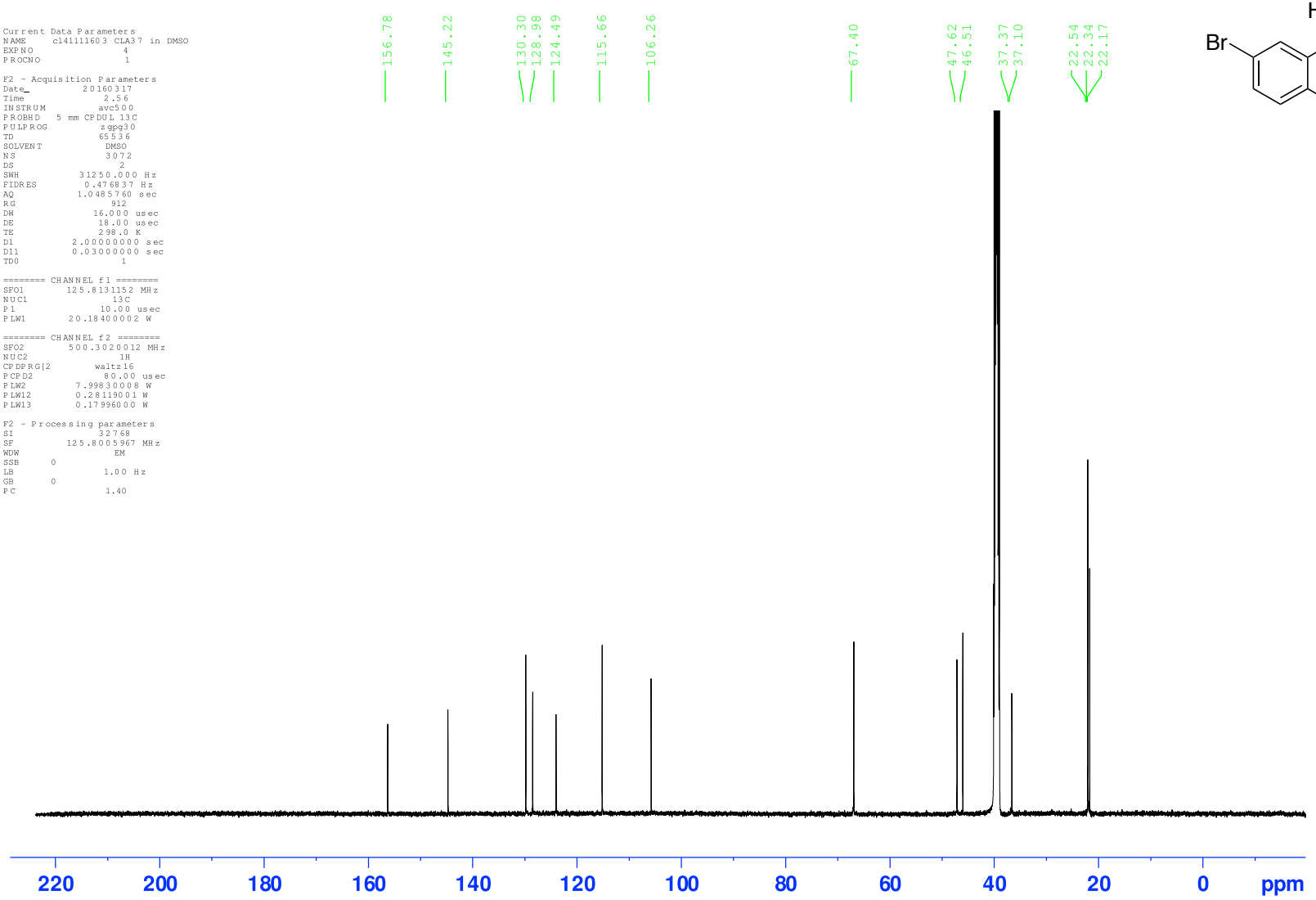




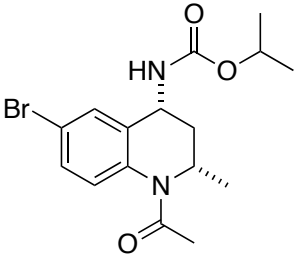
Isopropyl ((2*S*,4*R*)-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate ((*S*)-85)



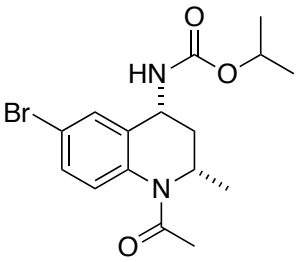
Isopropyl ((2*S*,4*R*)-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate ((*S*)-85)

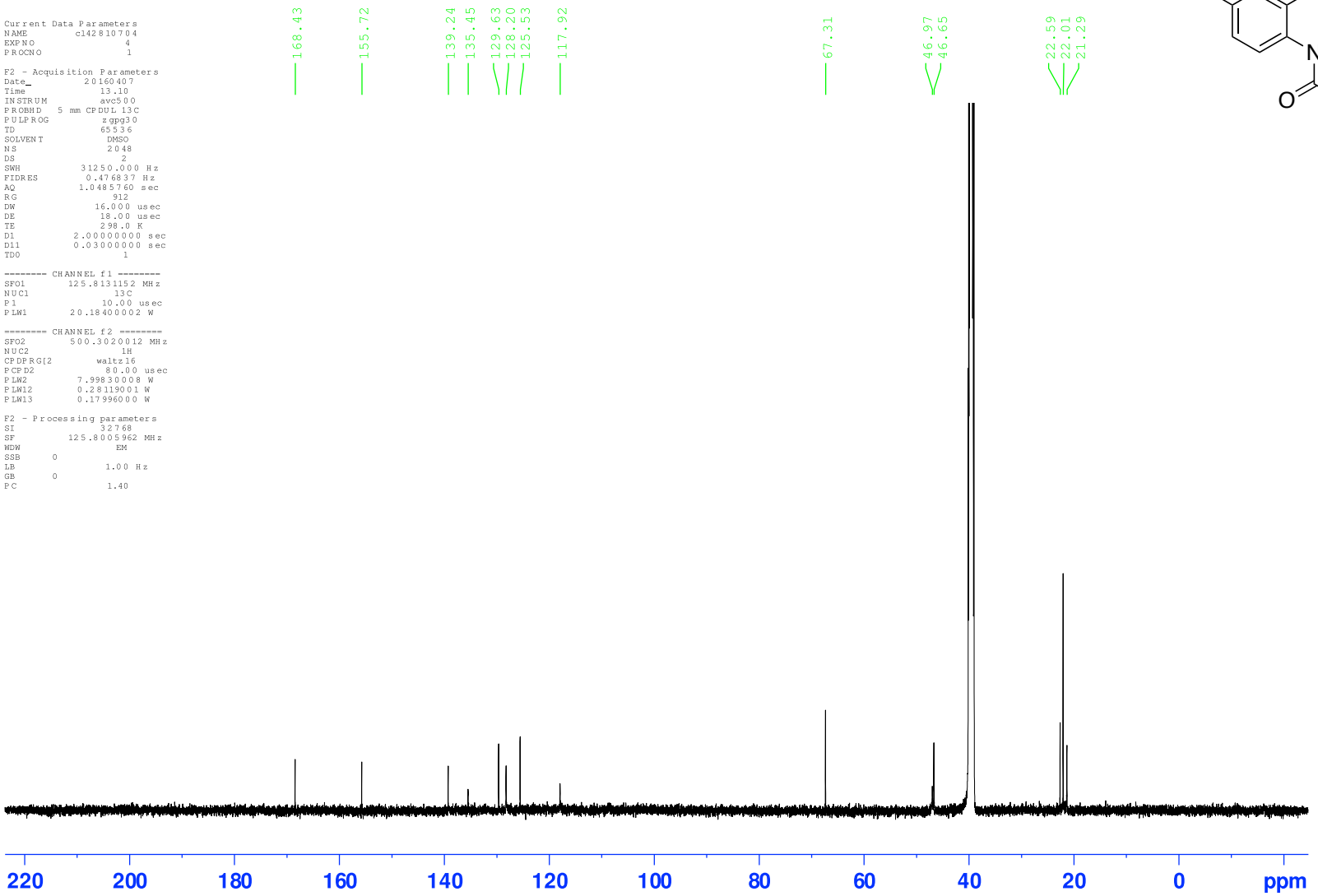


Isopropyl ((2*S*,4*R*)-1-acetyl-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (80)

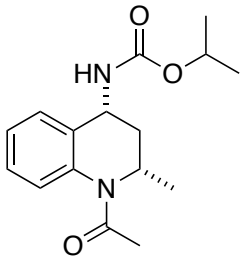


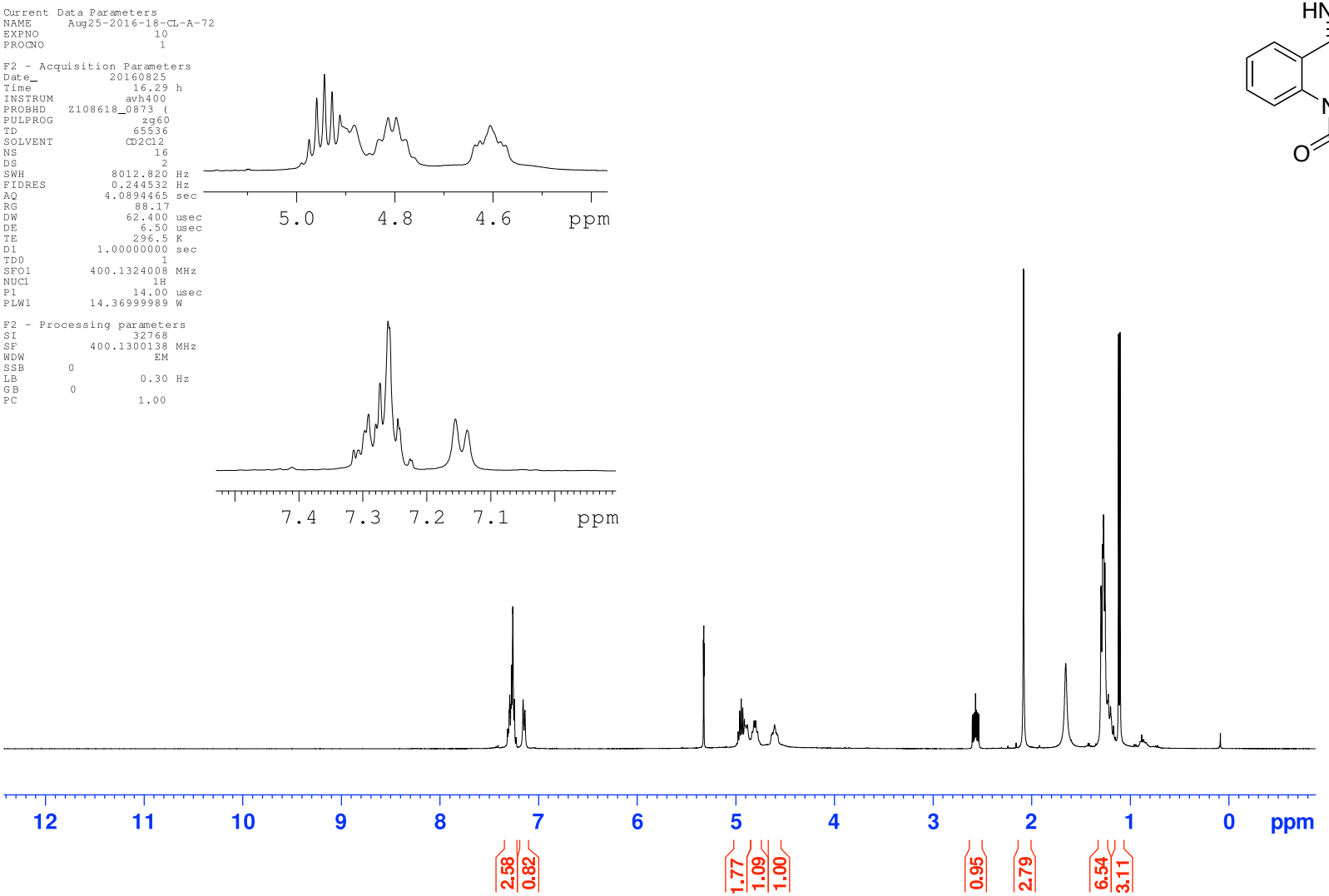
Isopropyl ((2*S*,4*R*)-1-acetyl-6-bromo-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (80)



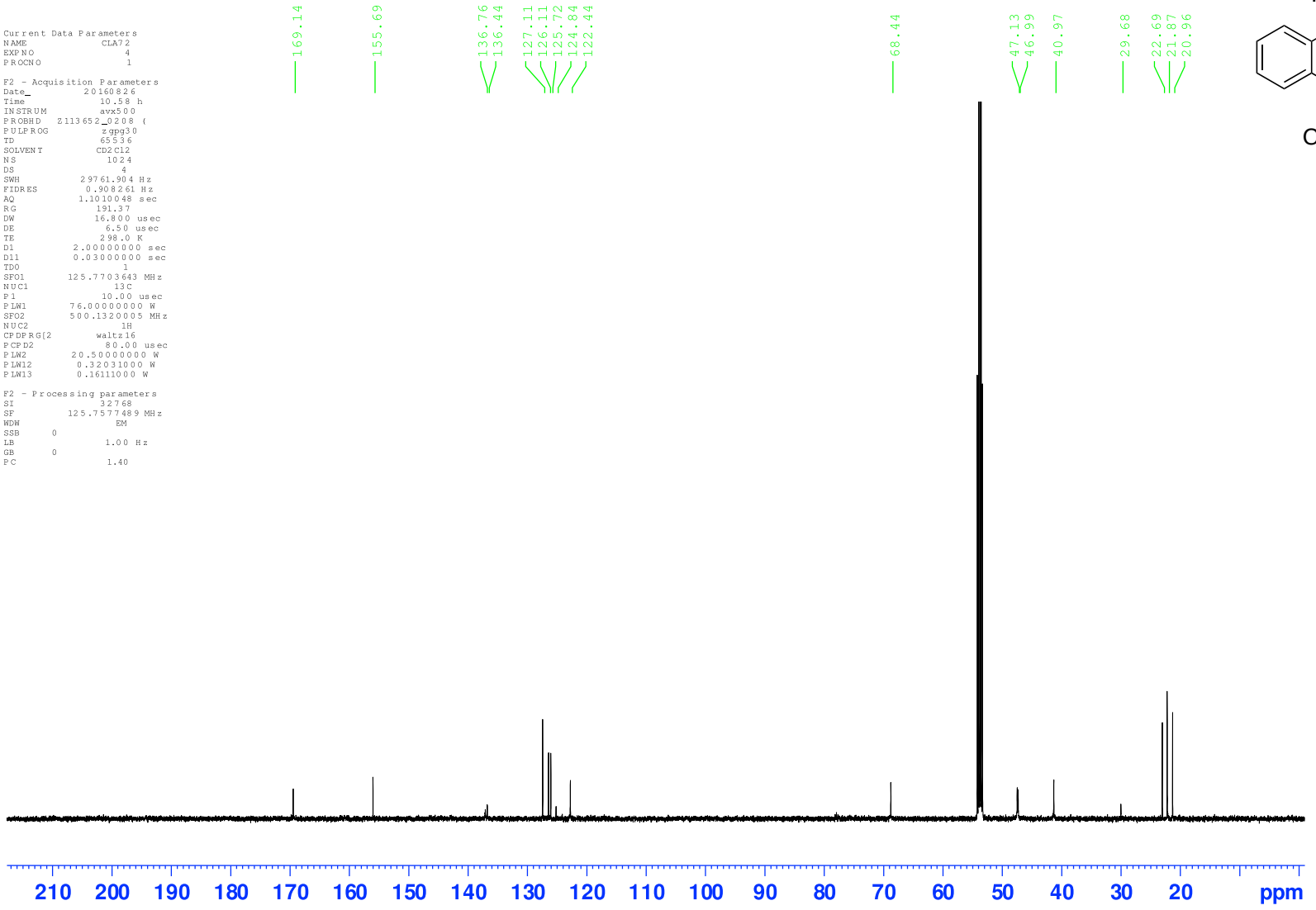


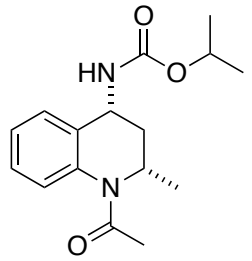
Isopropyl ((2*S*,4*R*)-1-acetyl-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (126)





Isopropyl ((2*S*,4*R*)-1-acetyl-2-methyl-1,2,3,4-tetrahydroquinolin-4-yl)carbamate (126)



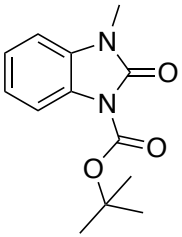


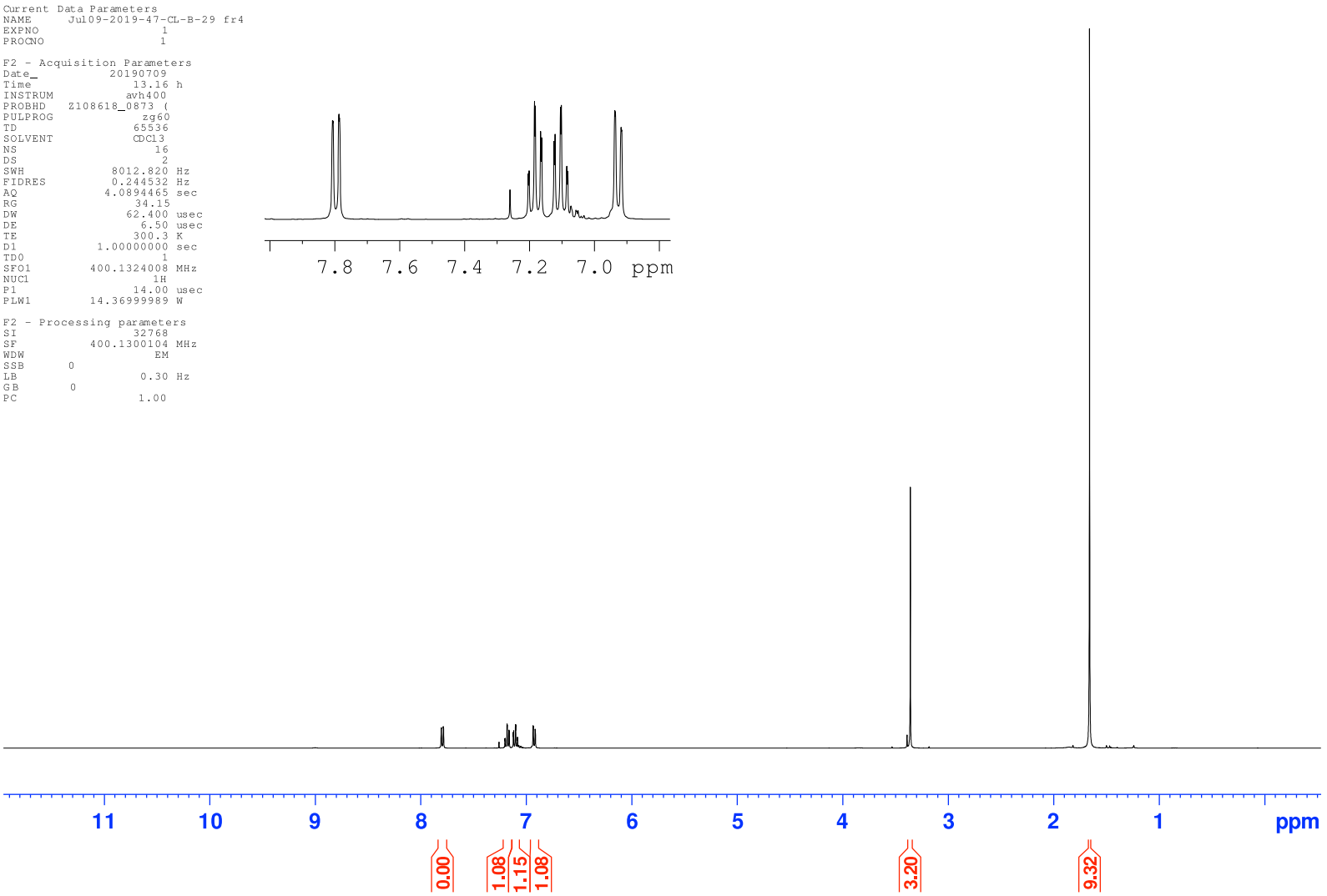

```
Current Data Parameters
NAME                  CLA72
EXPNO                     4
PROCNO                    1

F2 - Acquisition Parameters
Date_           20160826
Time                10.58 h
INSTRUM             avx500
PROBHD   Z113652_0208 (
PULPROG             zgpg30
TD                  65536
SOLVENT             CD2Cl2
NS                    1024
DS                       4
SWH              29761.904 Hz
FIDRES           0.908261 Hz
AQ               1.1010048 sec
RG                  191.37
DW                 16.800 usec
DE                  6.50 usec
TE                  298.0 K
D1            2.00000000 sec
D11           0.03000000 sec
TD0                      1
SFO1         125.7703643 MHz
NUC1                    13C
P1                  10.00 usec
PLW1         76.00000000 W
SFO2         500.1320005 MHz
NUC2                    1H
CPDPRG[2            waltz16
PCPD2                80.00 usec
PLW2         20.50000000 W
PLW12         0.32031000 W
PLW13         0.16111000 W

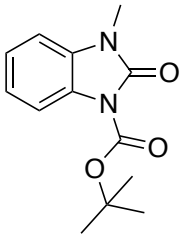
F2 - Processing parameters
SI                   32768
SF           125.7577489 MHz
WDW                     EM
SSB       0
LB                  1.00 Hz
GB        0
PC                  1.40
```

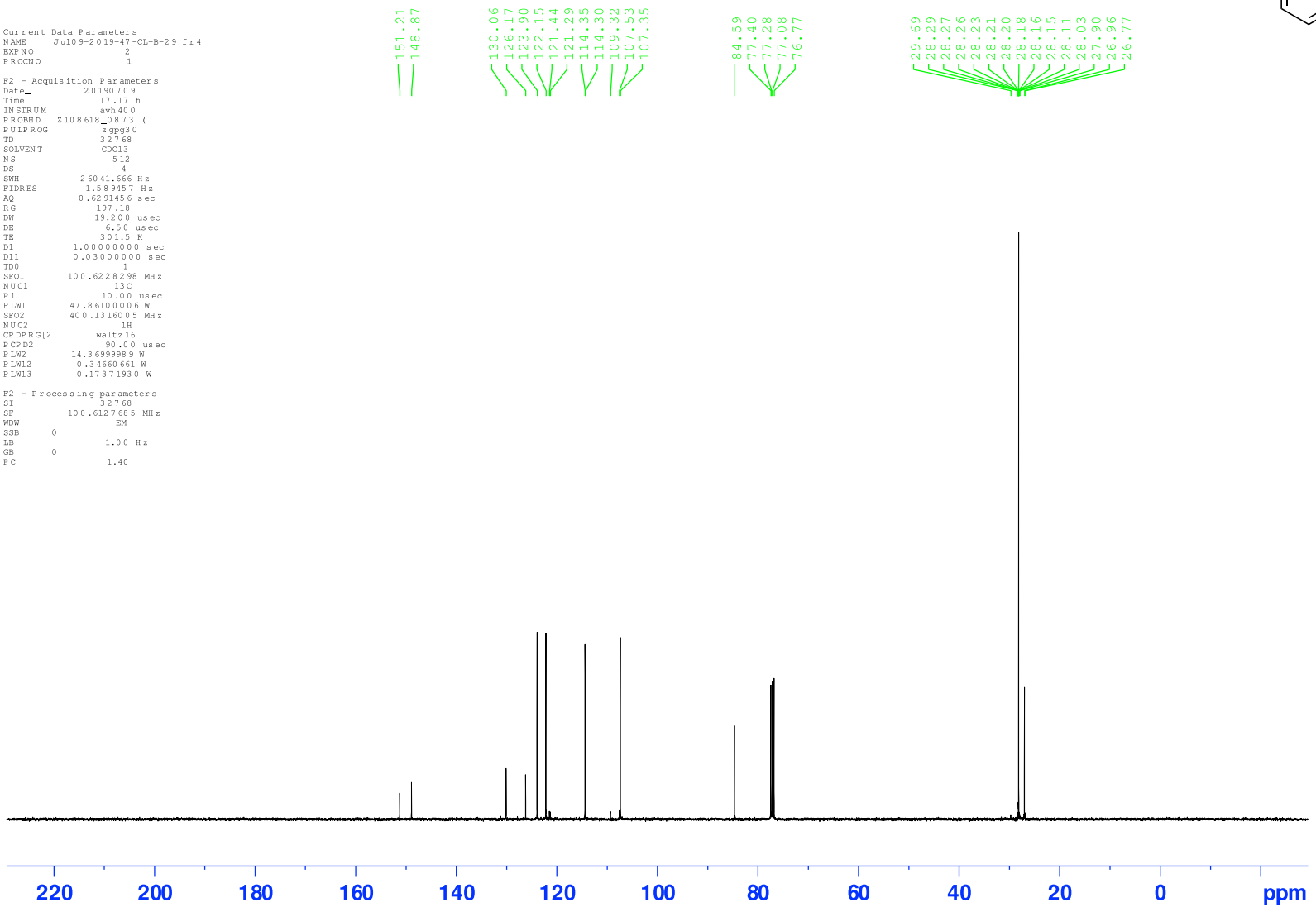
1-Methyl-3-(*tert*-butyloxycarbonyl)benzimidazol-2-one (28)



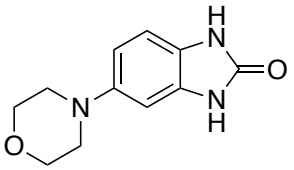


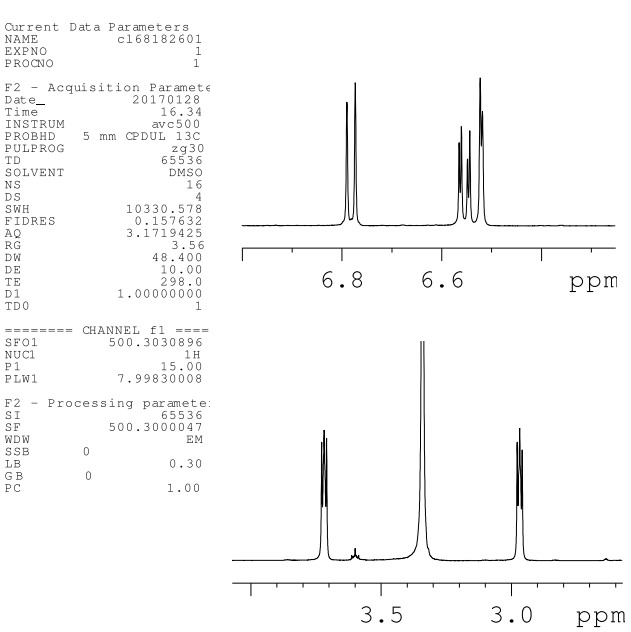
Methyl-3-(*tert*-butyloxycarbonyl)benzimidazol-2-one (28)



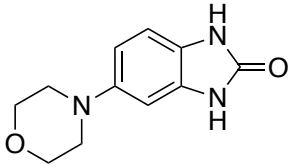


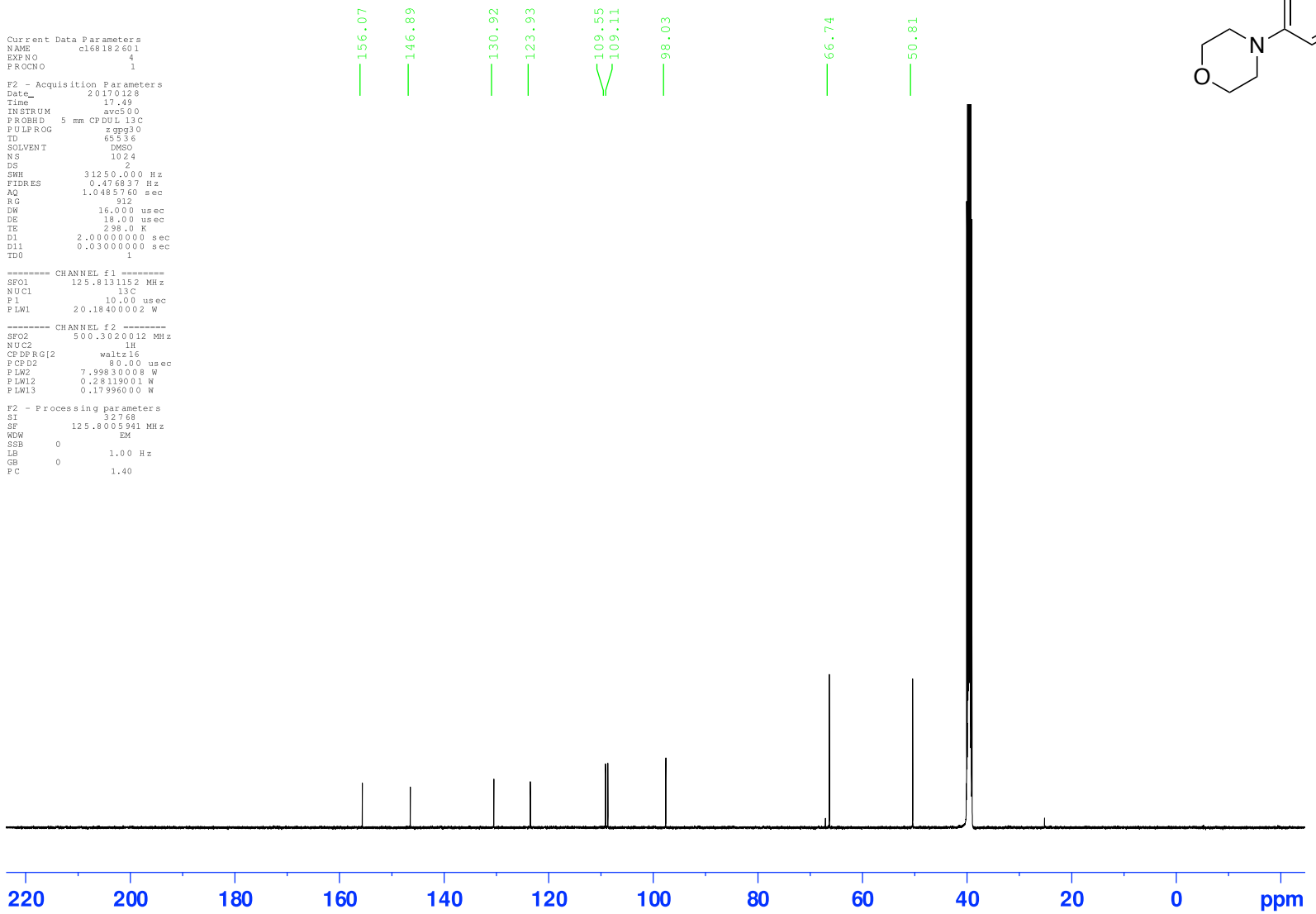
5-Morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (47)



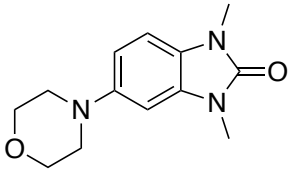


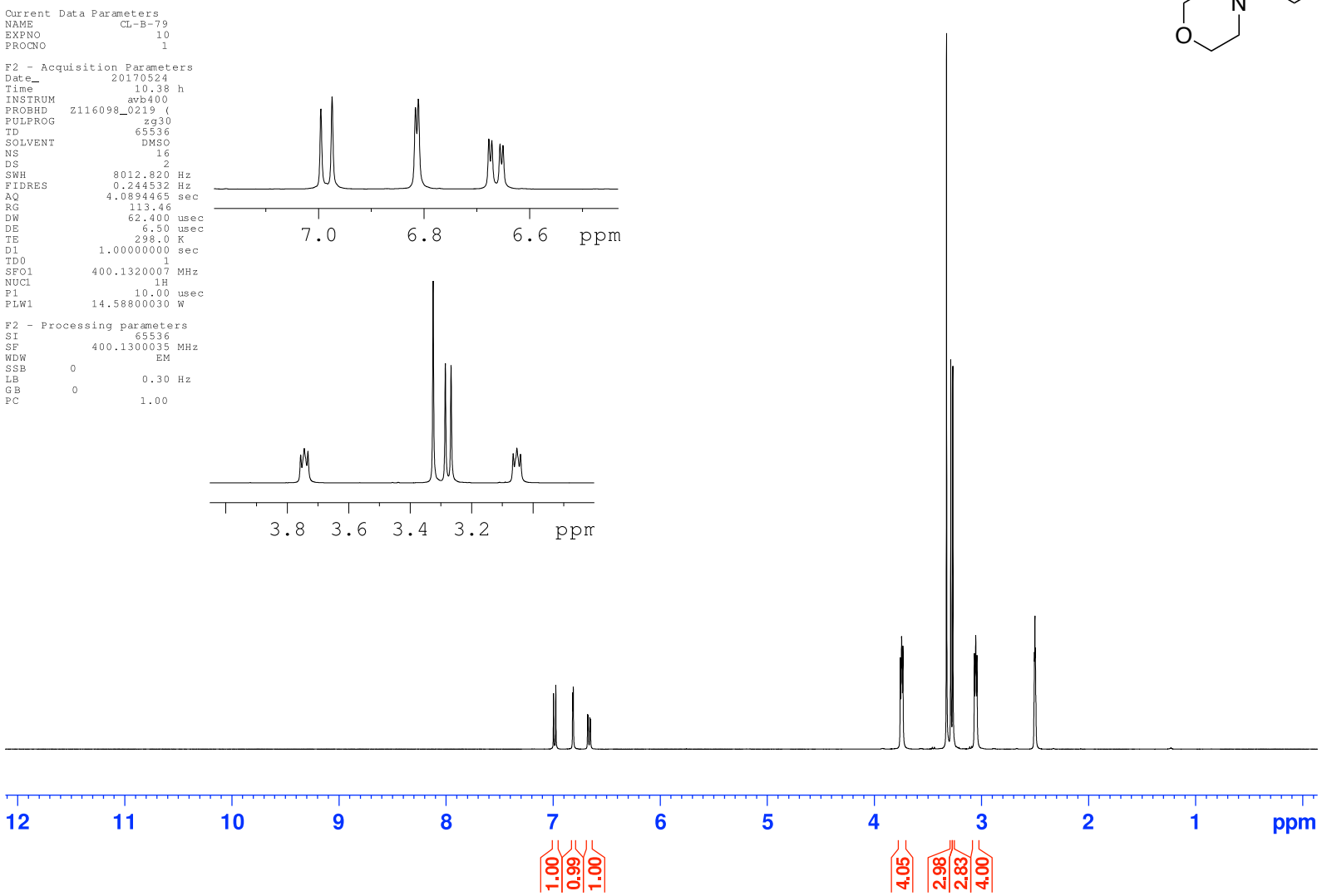
5-Morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (47)



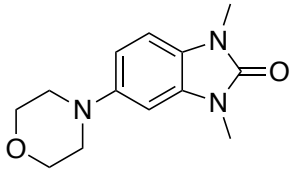


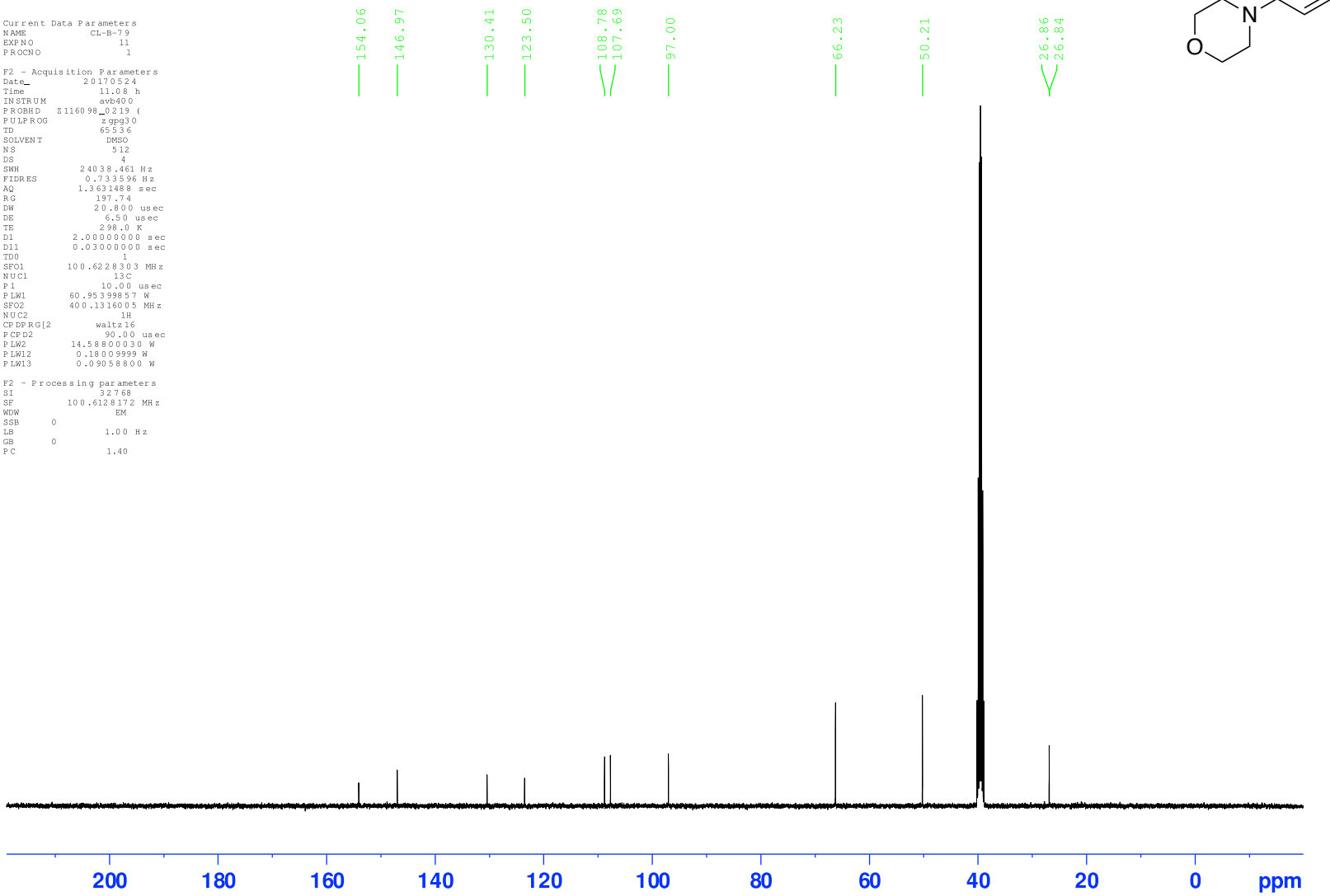
1,3-Dimethyl-5-morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (44)



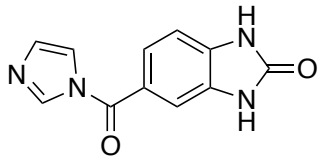


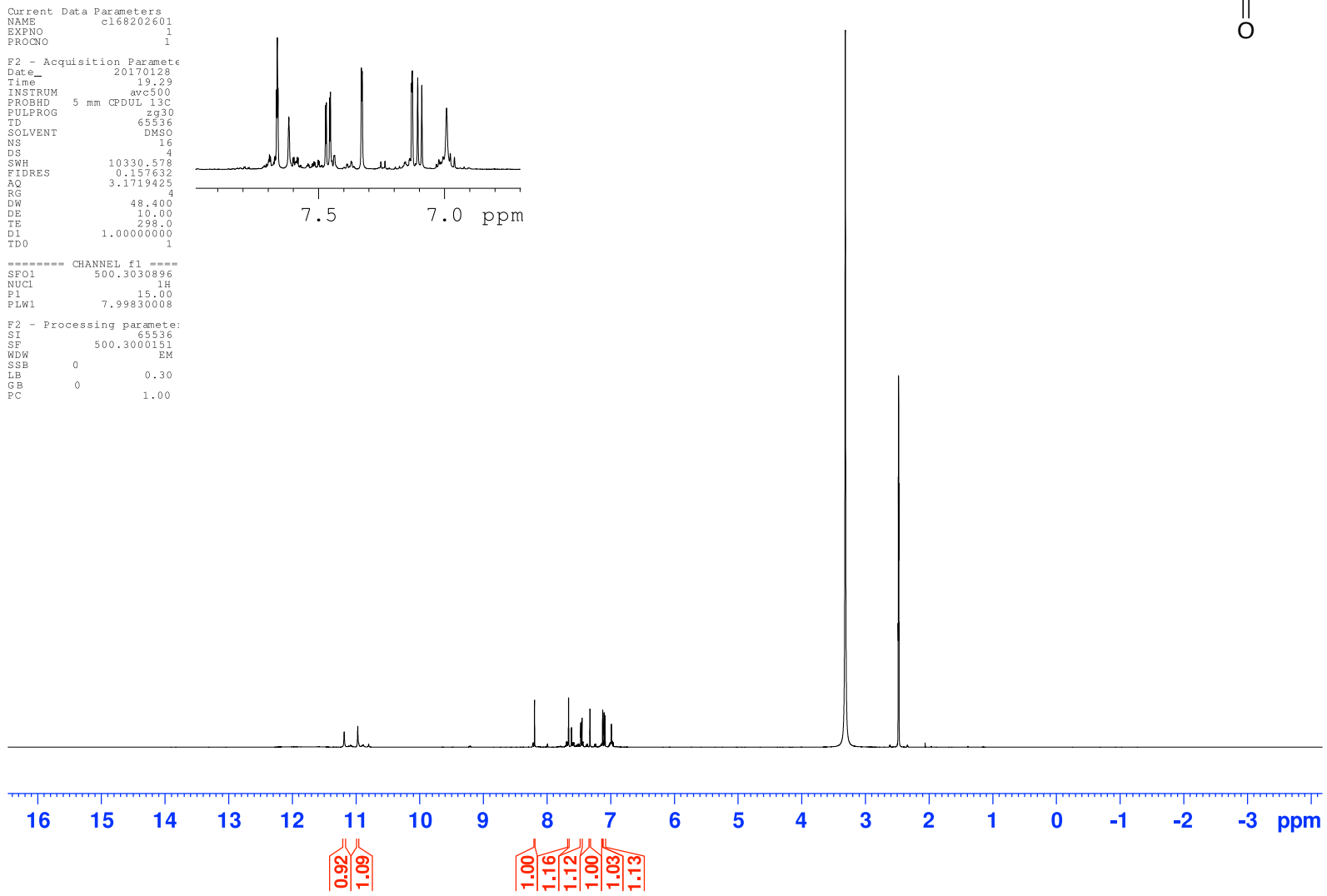
1,3-Dimethyl-5-morpholino-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (44)



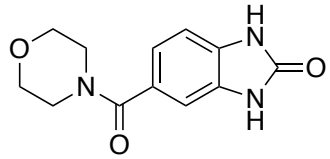


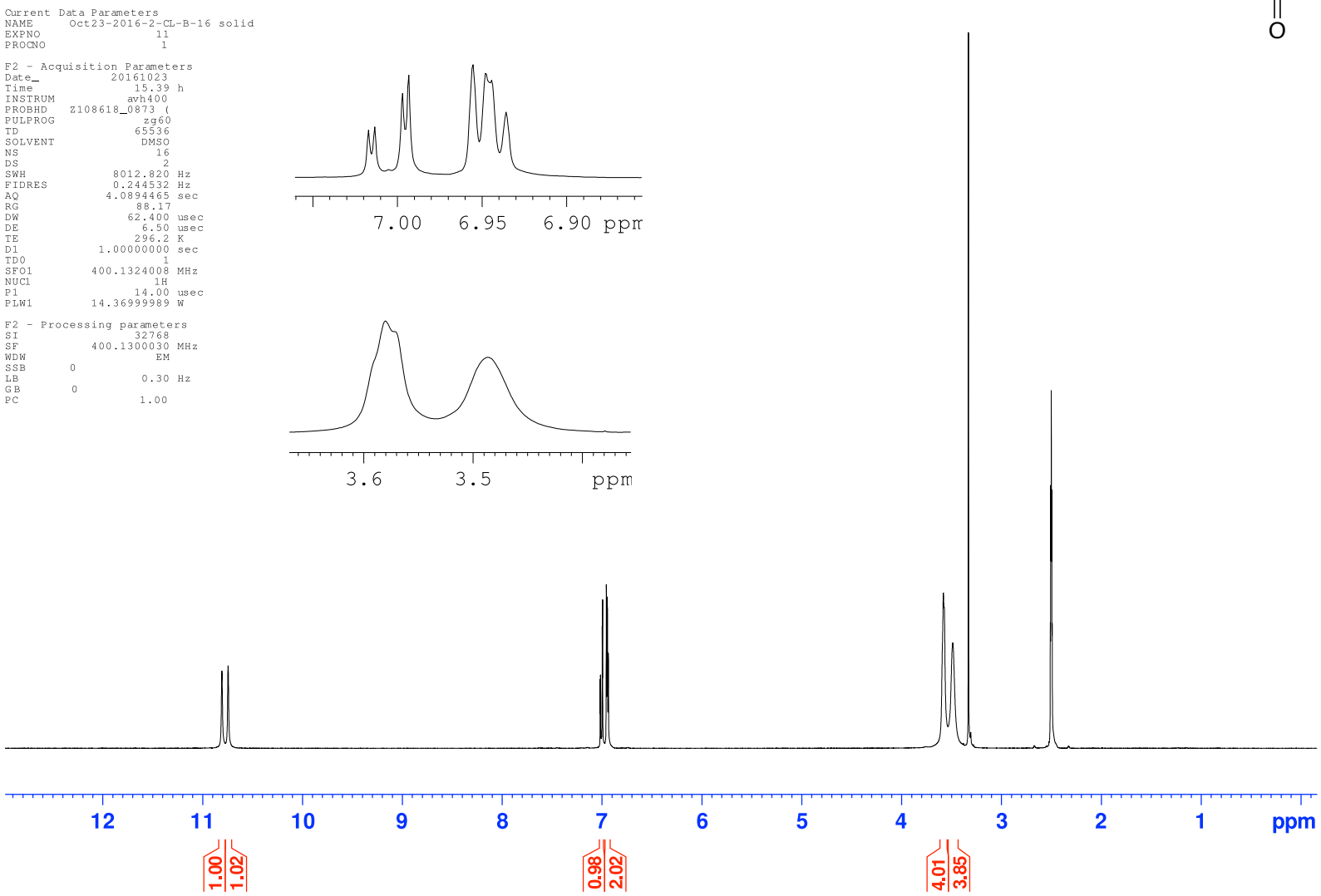
5’-(1*H*-Imidazole-1’-carbonyl)-benzimidazol-2-one (43a)



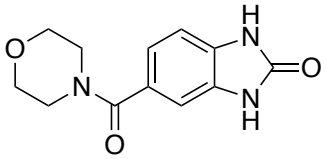


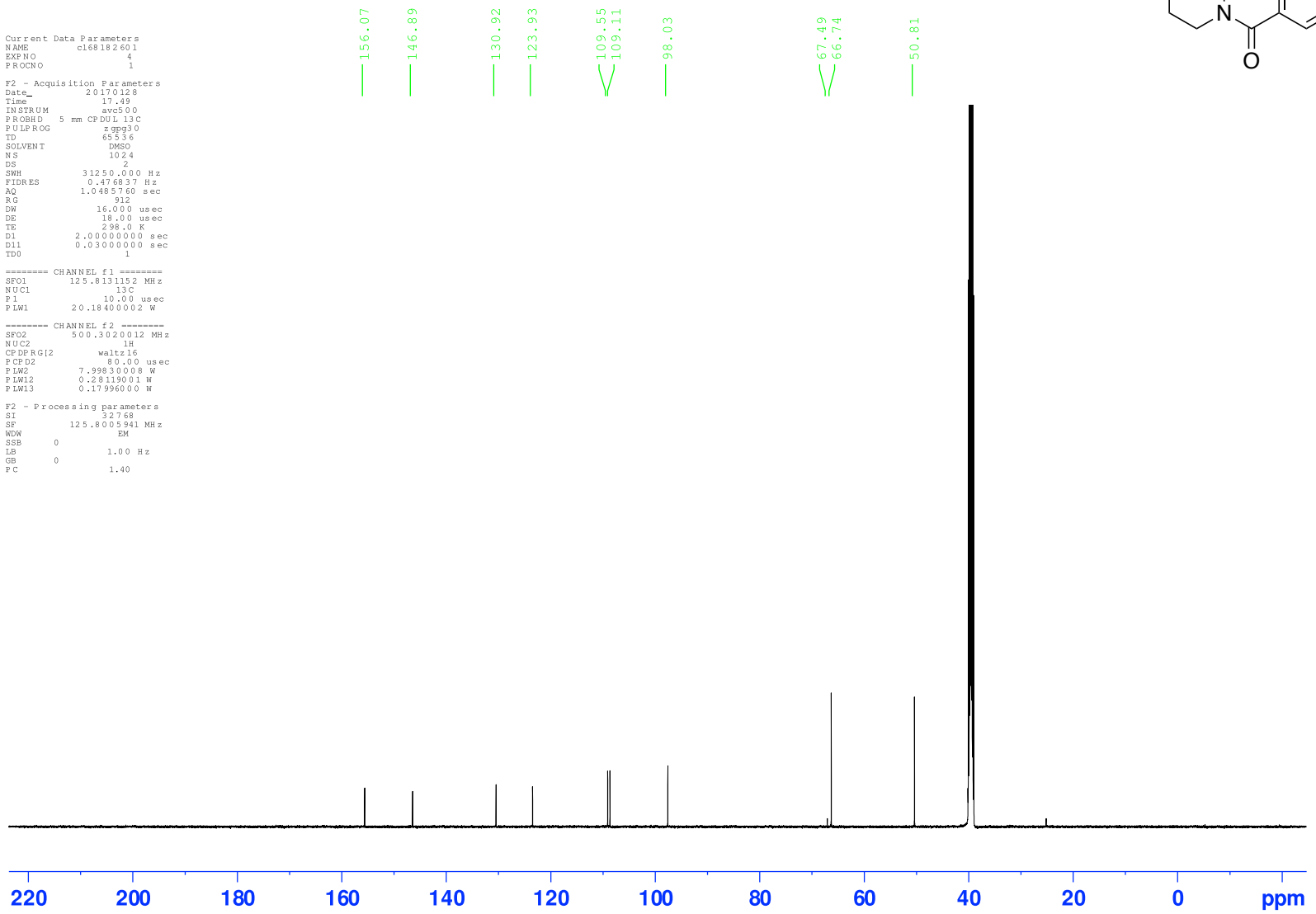
5-(Morpholine-4-carbonyl)-benzimidazol-2-one (43)





5-(Morpholine-4-carbonyl)-benzimidazol-2-one (43)






```
Current Data Parameters
NAME          cl68182601
EXPNO                  4
PROCNO                 1

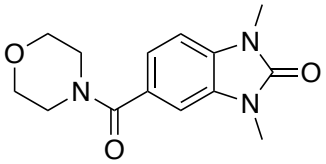
F2 - Acquisition Parameters
Date_          20170128
Time              17.49
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD               65536
SOLVENT           DMSO
NS                1024
DS                   2
SWH           31250.000 Hz
FIDRES        0.476837 Hz
AQ            1.04857760 sec
RG                 912
DW              16.000 usec
DE              18.00 usec
TE              298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                  1

======== CHANNEL f1 ========
SFO1        125.8131152 MHz
NUC1               13C
P1               10.00 usec
PLW1        20.18400002 W

-------- CHANNEL f2 --------
SFO2        500.3020012 MHz
NUC2               1H
CPDPRG[2         waltz16
PCPD2             80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

F2 - Processing parameters
SI                32768
SF          125.8005941 MHz
WDW                EM
SSB      0
LB              1.00 Hz
GB       0
PC              1.40
```

1,3-Dimethyl-5-(morpholine-4-carbonyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (41)

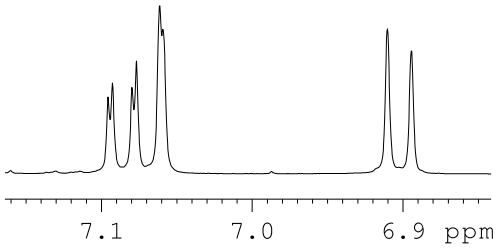


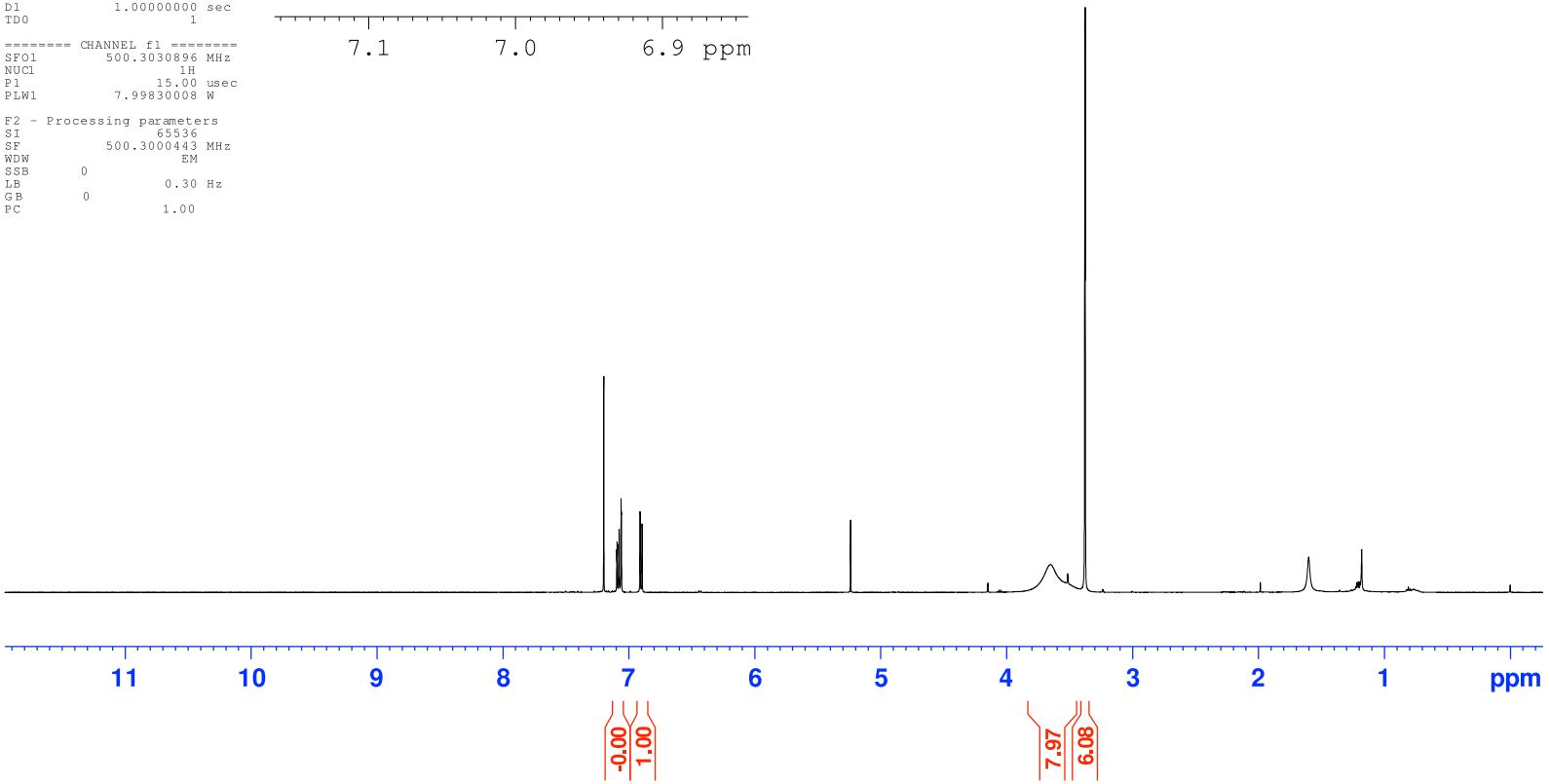
```
Current Data Parameters
NAME      c162060711 CLB17
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_          20161109
Time               0.30
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                285.4 K
D1            1.00000000 sec
TD0                   1

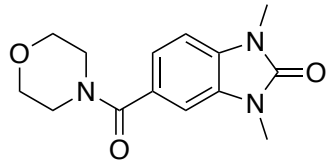
======== CHANNEL f1 ========
SFO1        500.3030896 MHz
NUC1                1H
P1                15.00 usec
PLW1         7.99830008 W

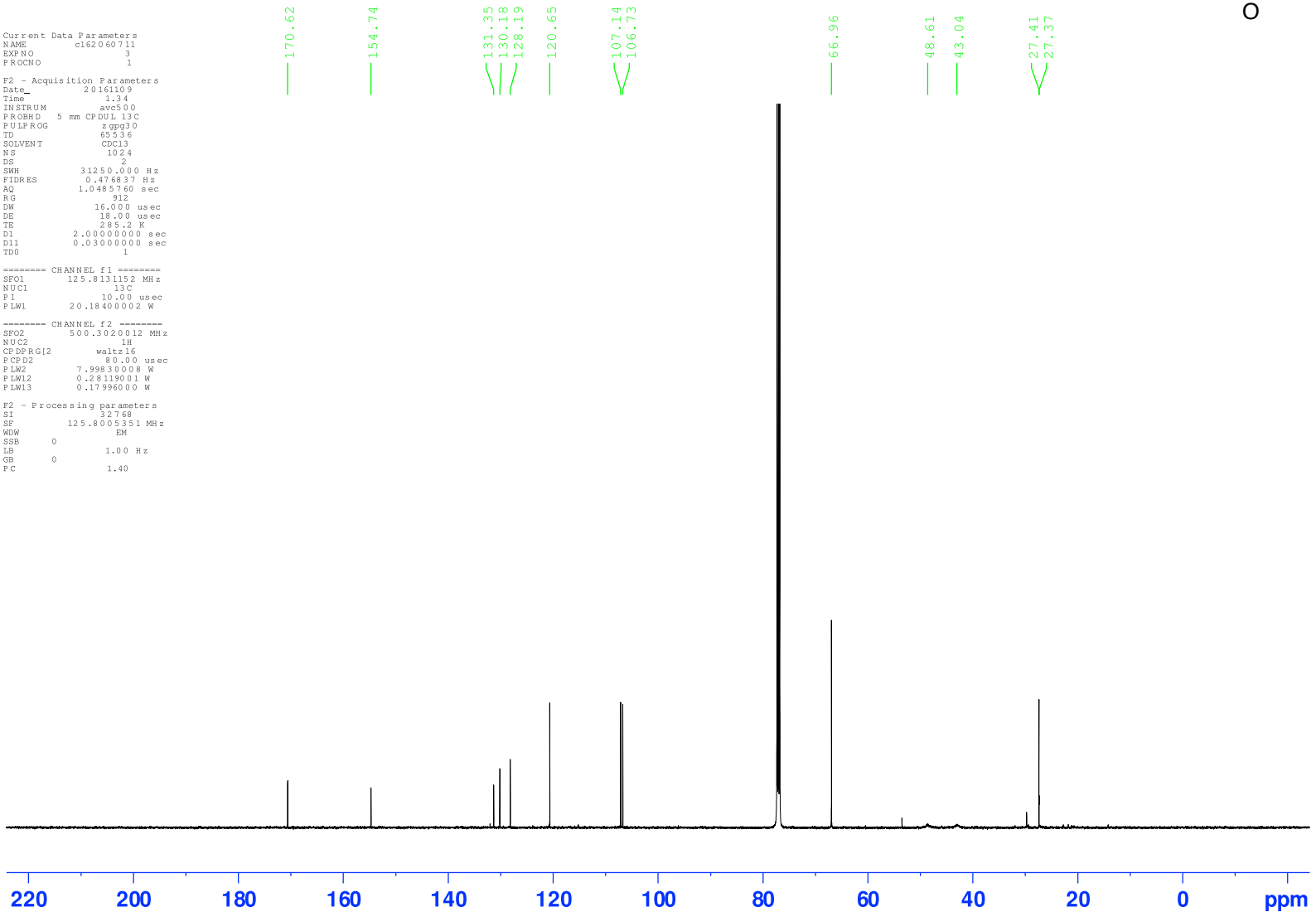
F2 - Processing parameters
SI                65536
SF          500.3000443 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```



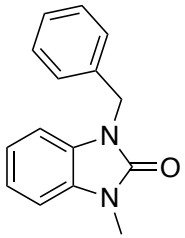


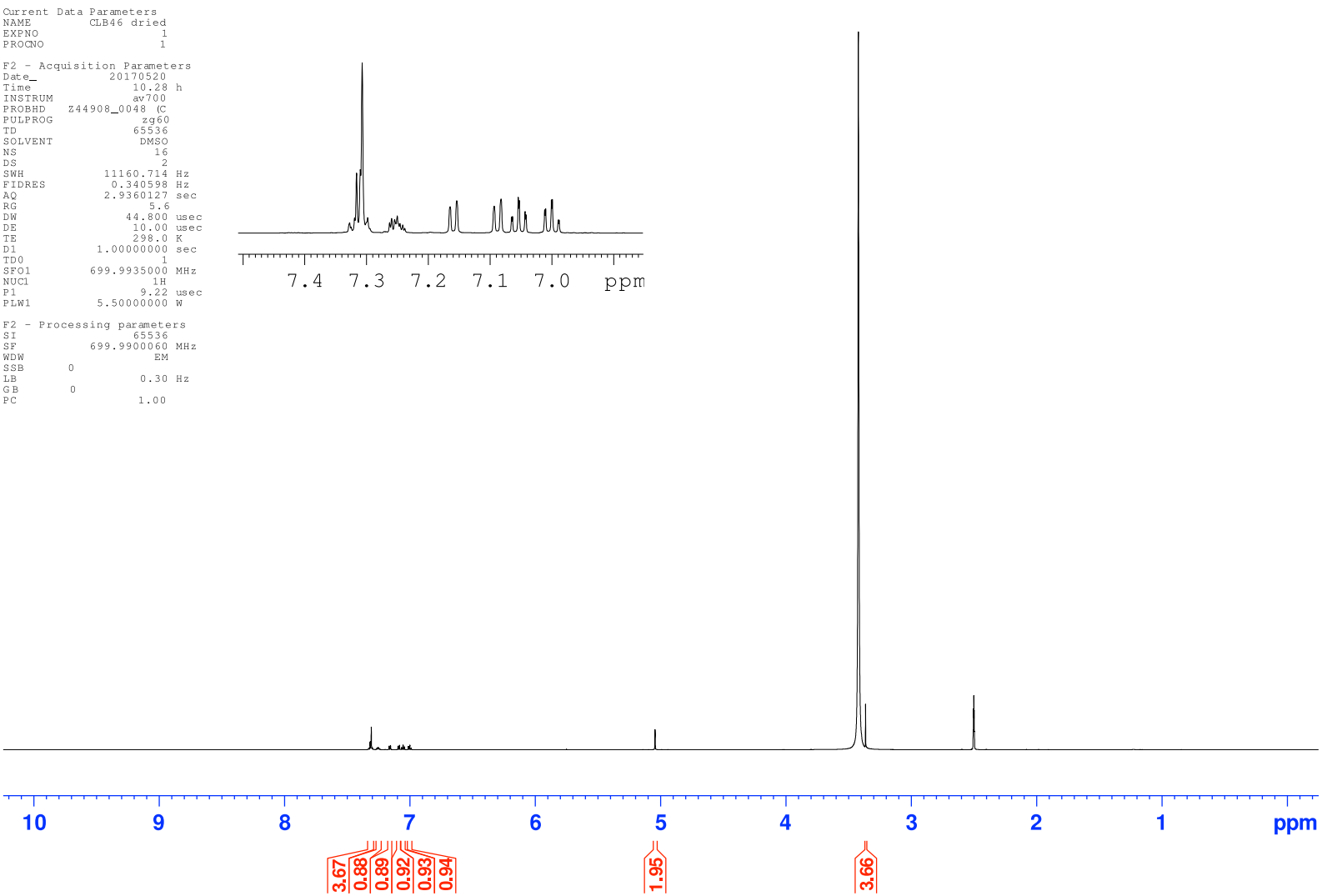
1,3-Dimethyl-5-(morpholine-4-carbonyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (41)



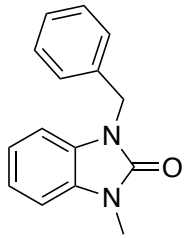


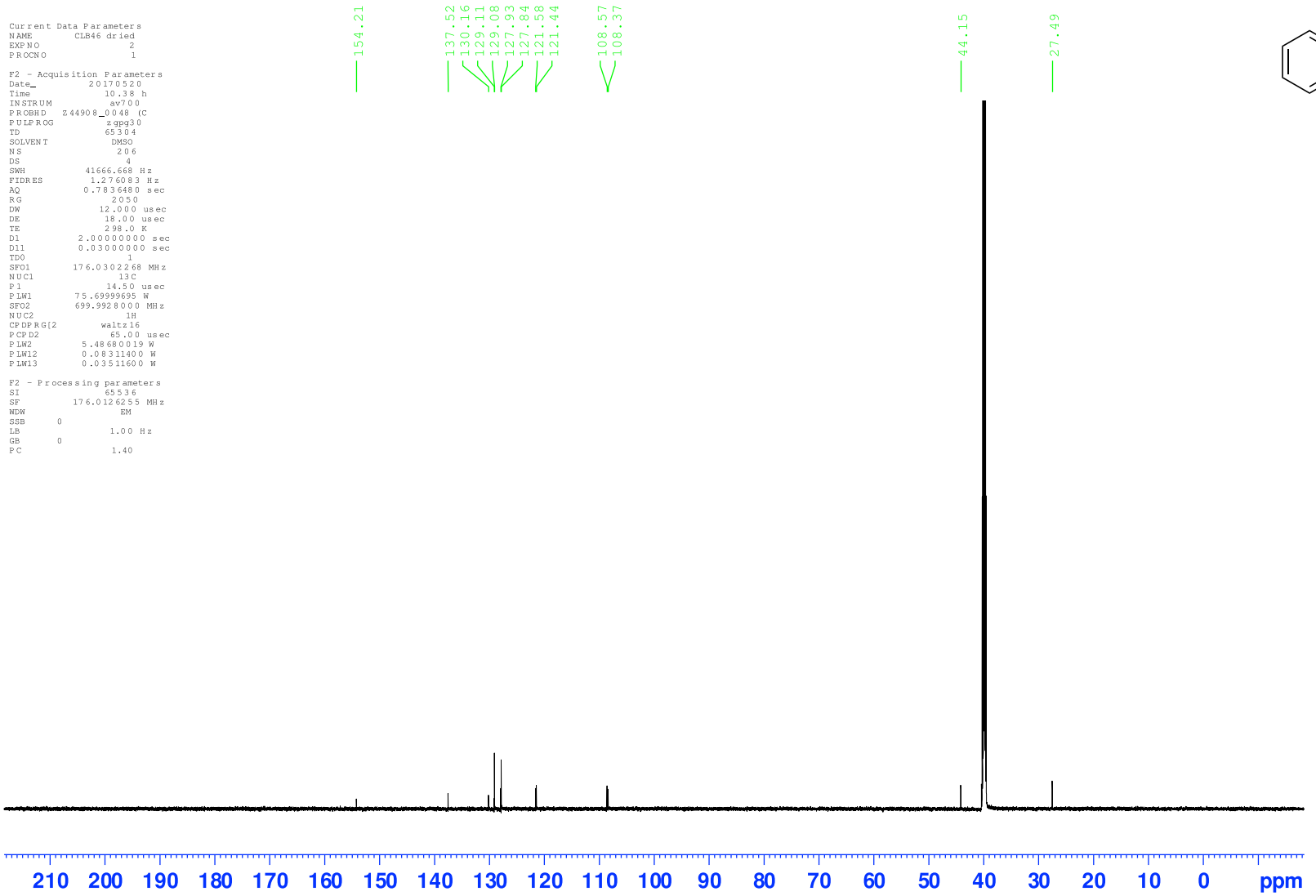
1-Benzyl-3-methyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (37)



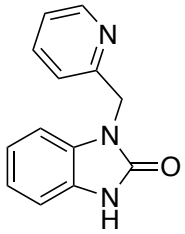


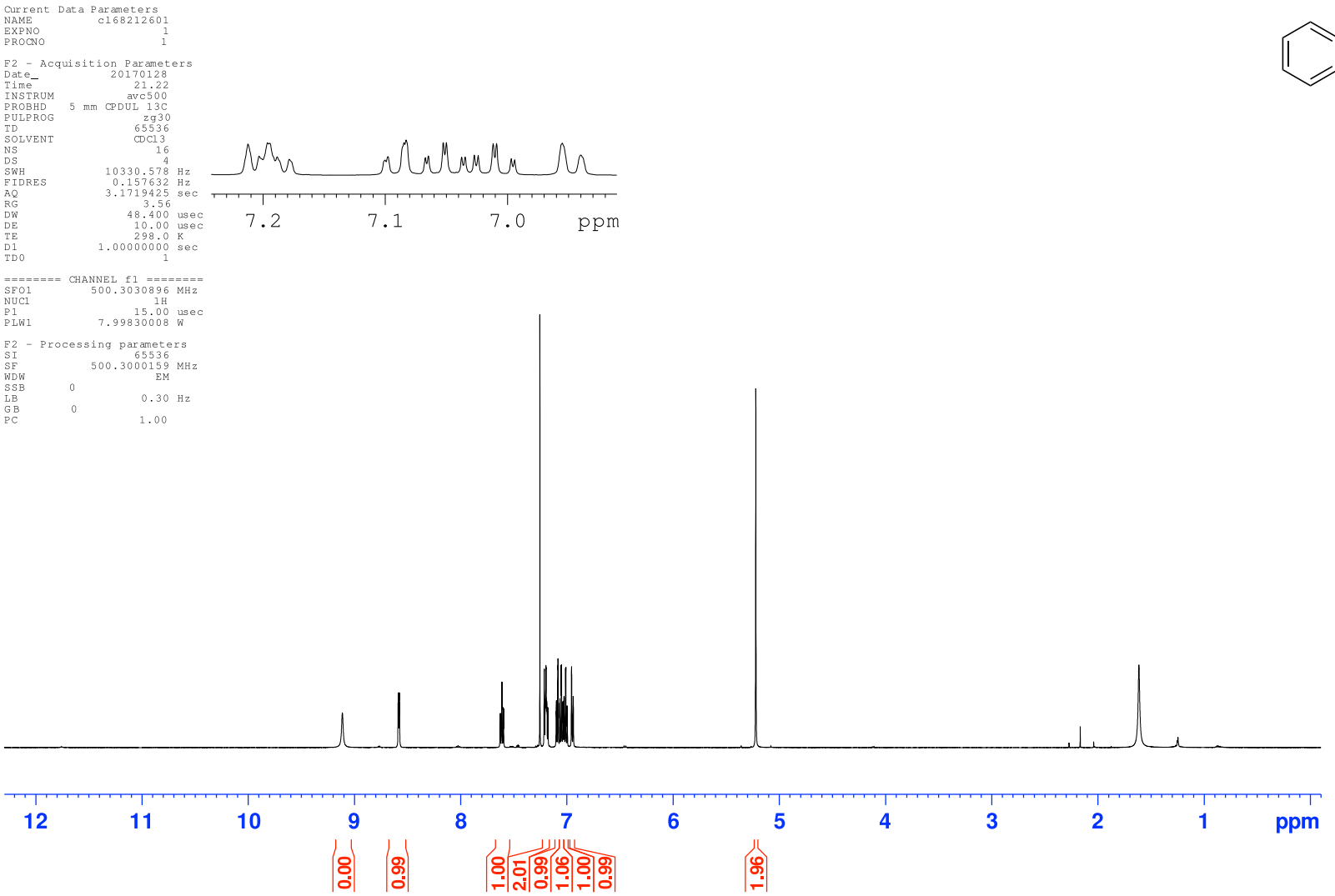
1-Benzyl-3-methyl-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (37)



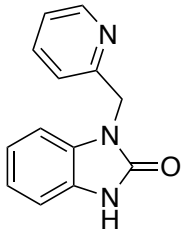


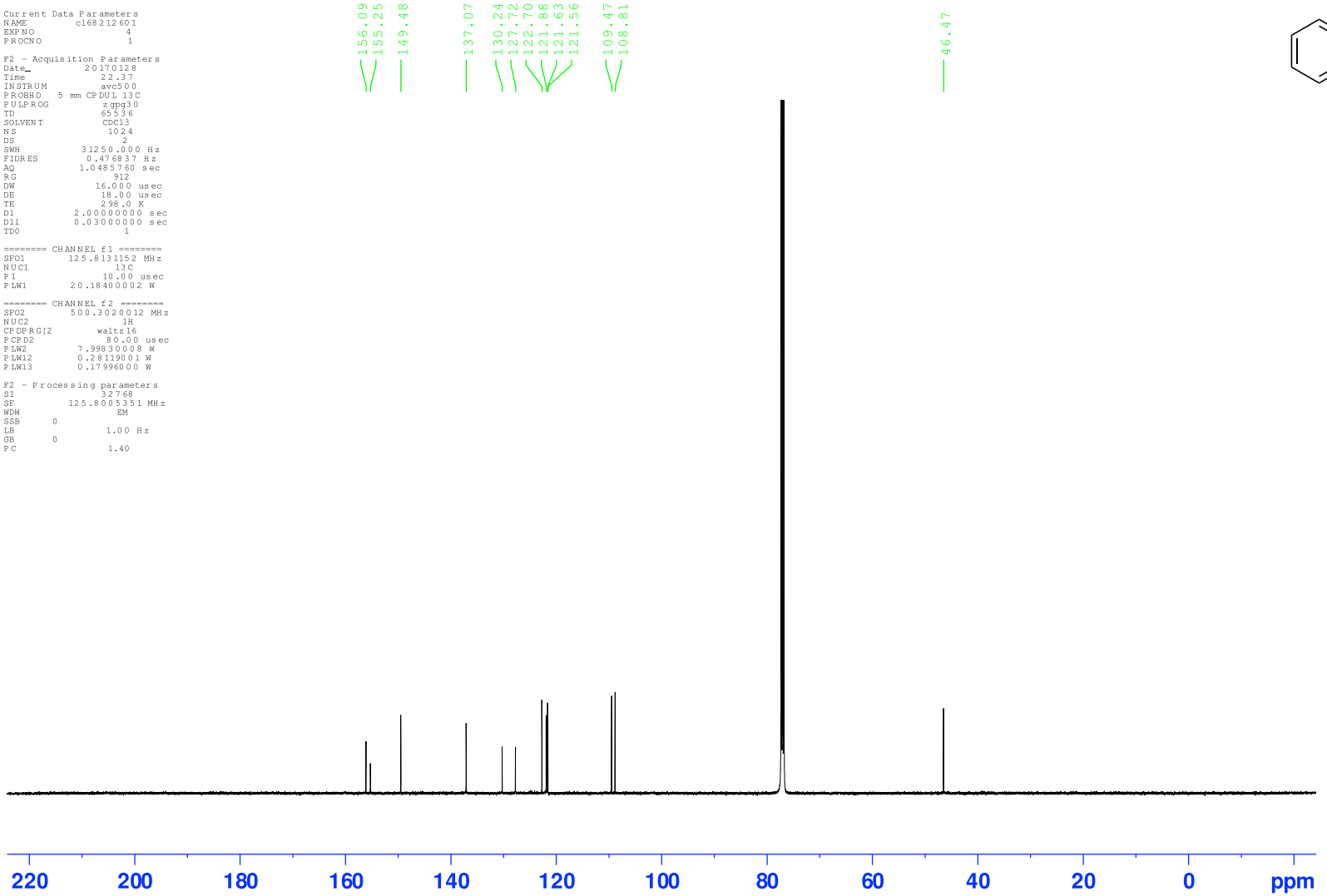
1'-(Pyridin-2'-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (34)



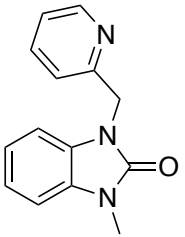


1'-(Pyridin-2'-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (34)





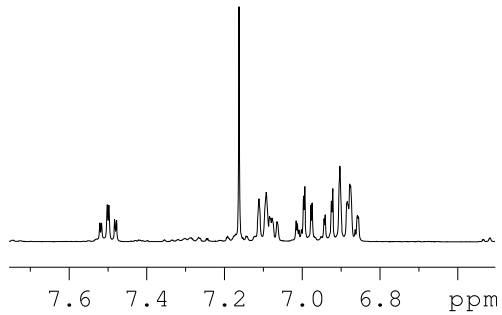
1-Methyl-3-(pyridin-2-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (38)

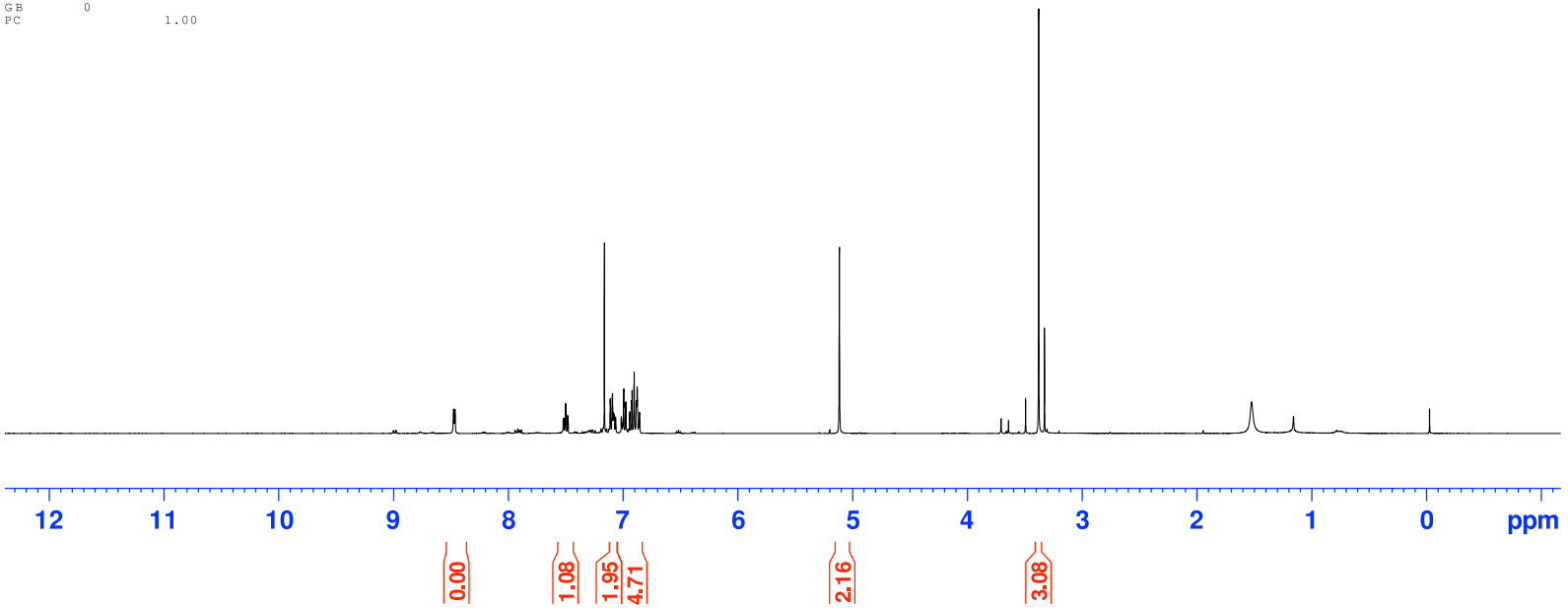



```
Current Data Parameters
NAME      Jul11-2019-38-CL-B-47
EXPNO                         1
PROCNO                        1

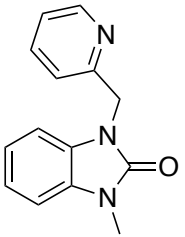
F2 - Acquisition Parameters
Date_              20190711
Time                  17.28 h
INSTRUM               avh400
PROBHD    Z108618_0873 (
PULPROG                 zg60
TD                     65536
SOLVENT                CDCl3
NS                        16
DS                         2
SWH                 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

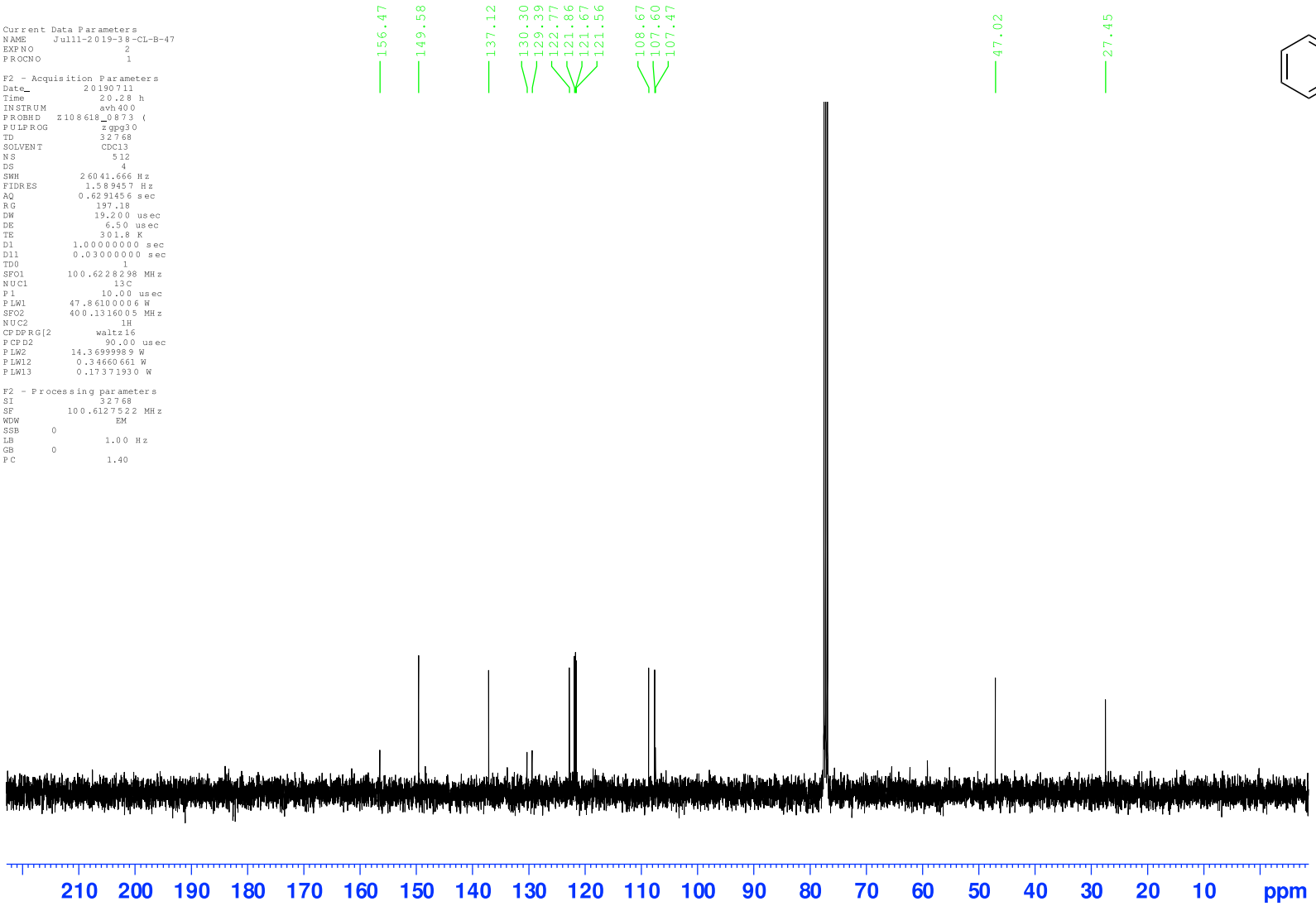
F2 - Processing parameters
SI                     32768
SF               400.1300491 MHz
WDW                       EM
SSB       0
LB                      0.30 Hz
GB        0
PC                      1.00
```



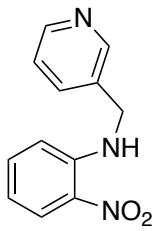


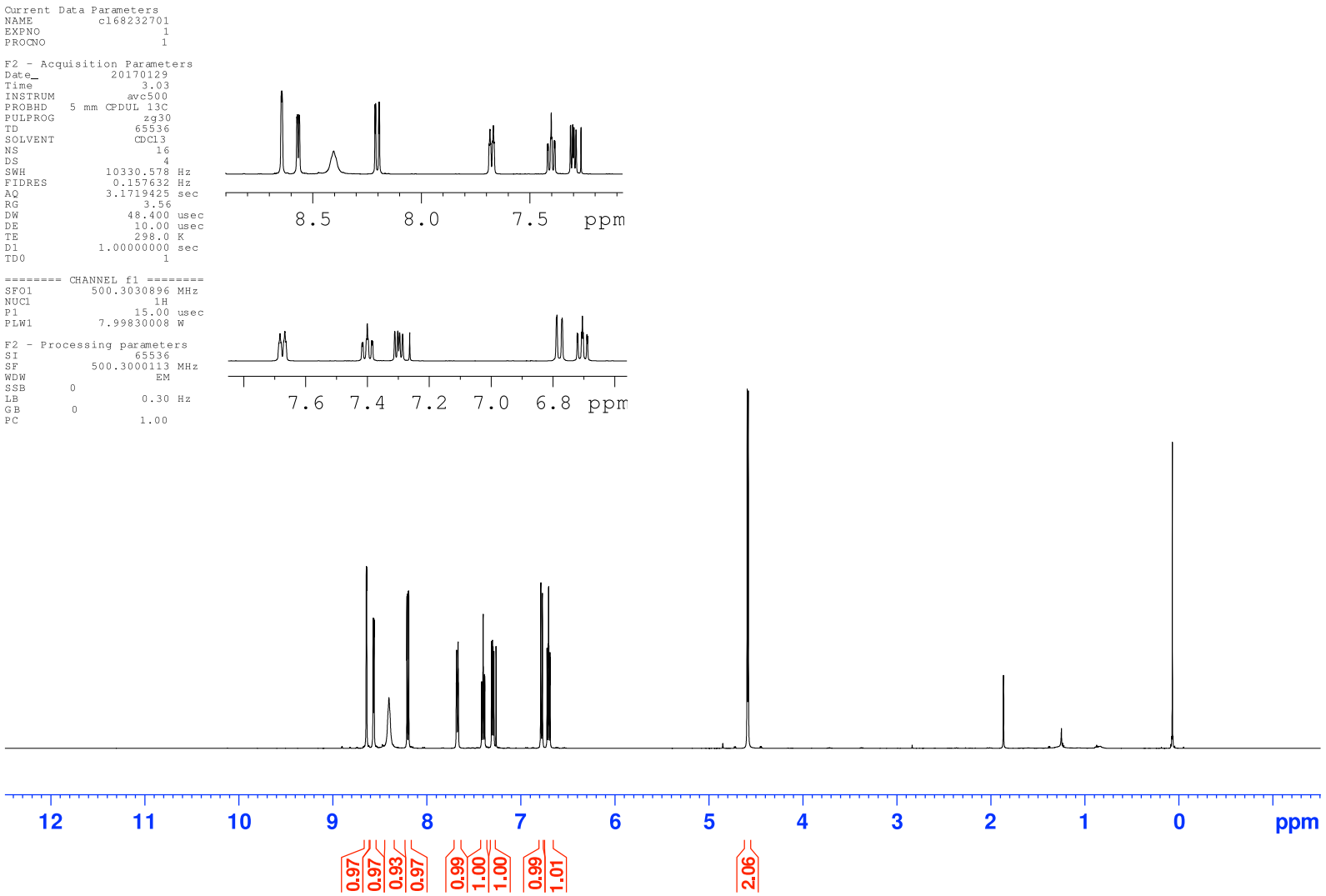
1-Methyl-3-(pyridin-2-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (38)



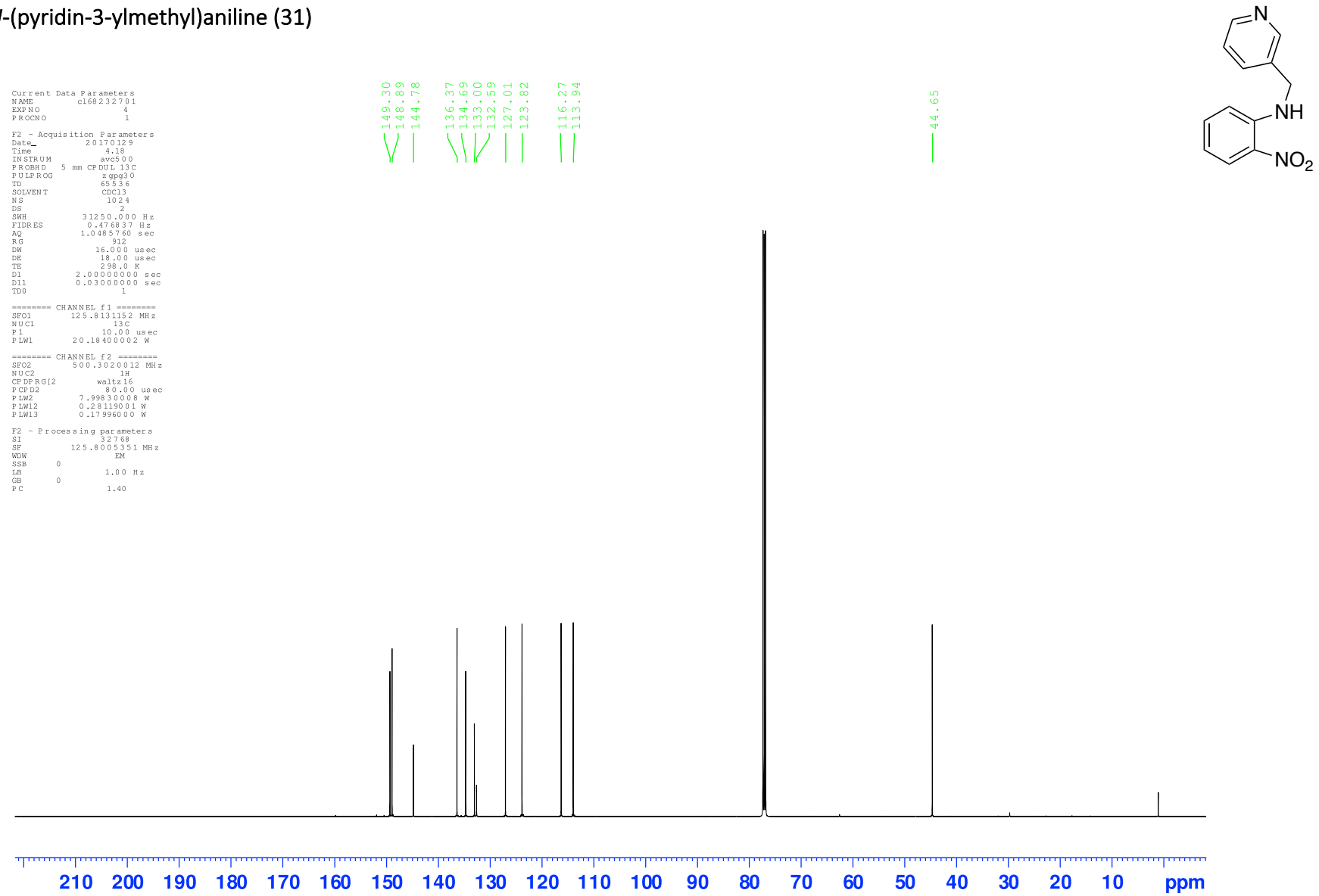


2-Nitro-*N*-(pyridin-3-ylmethyl)aniline (31)

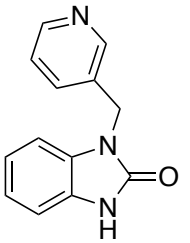


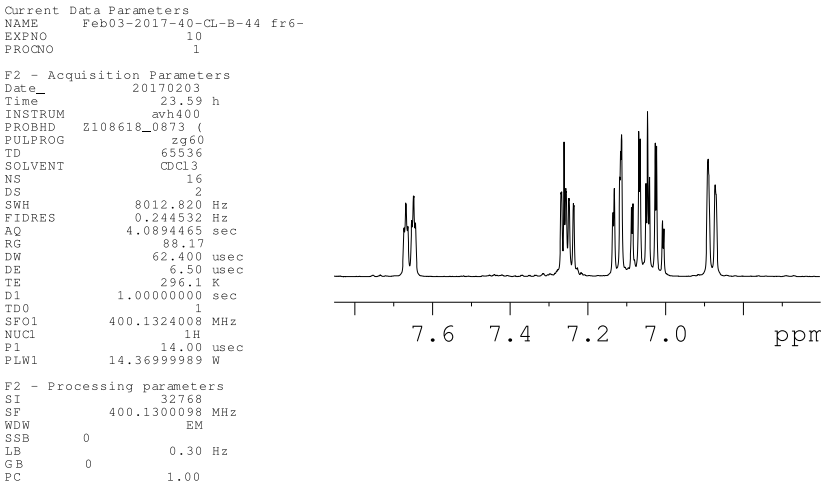


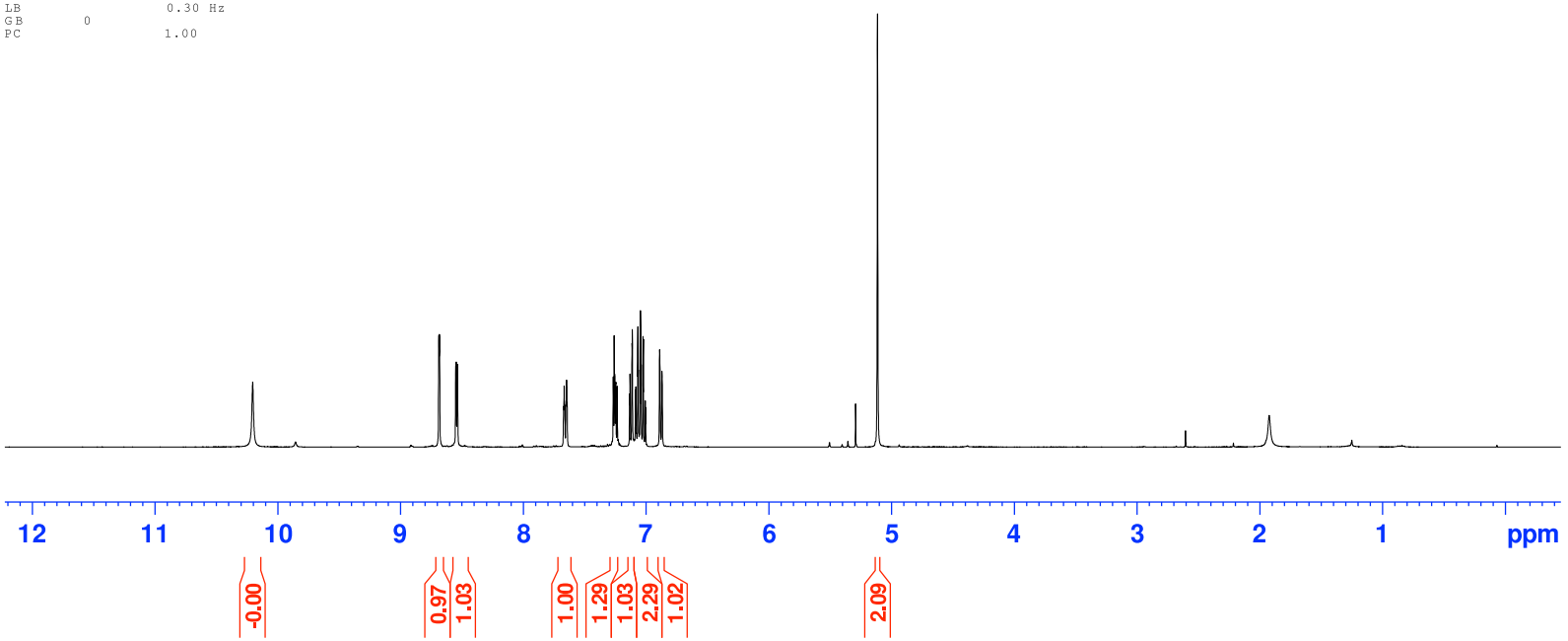
2-Nitro-*N*-(pyridin-3-ylmethyl)aniline (31)



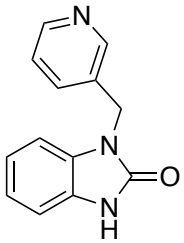
1-(Pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (35)

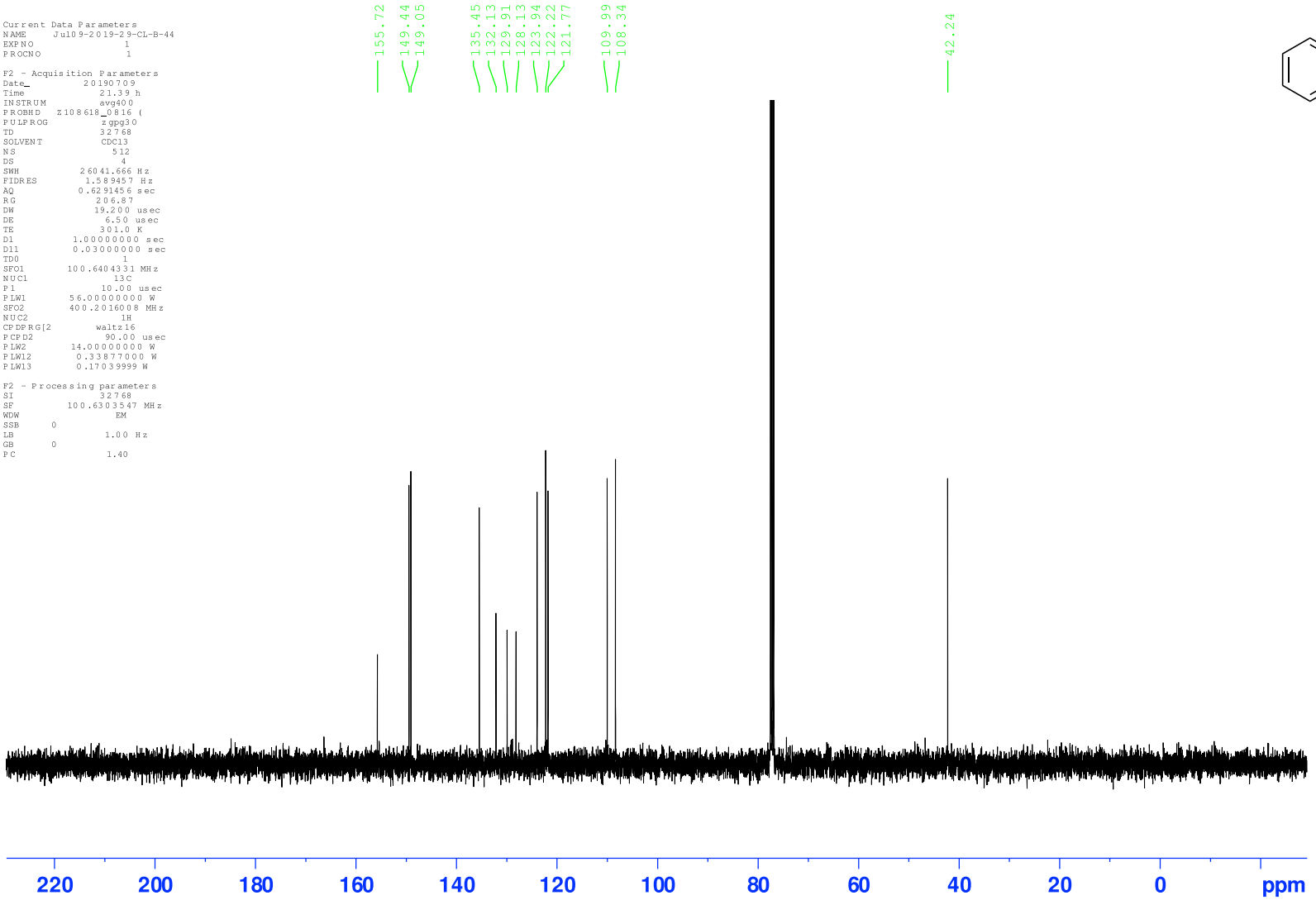






1-(Pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (35)



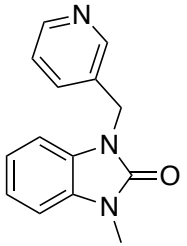


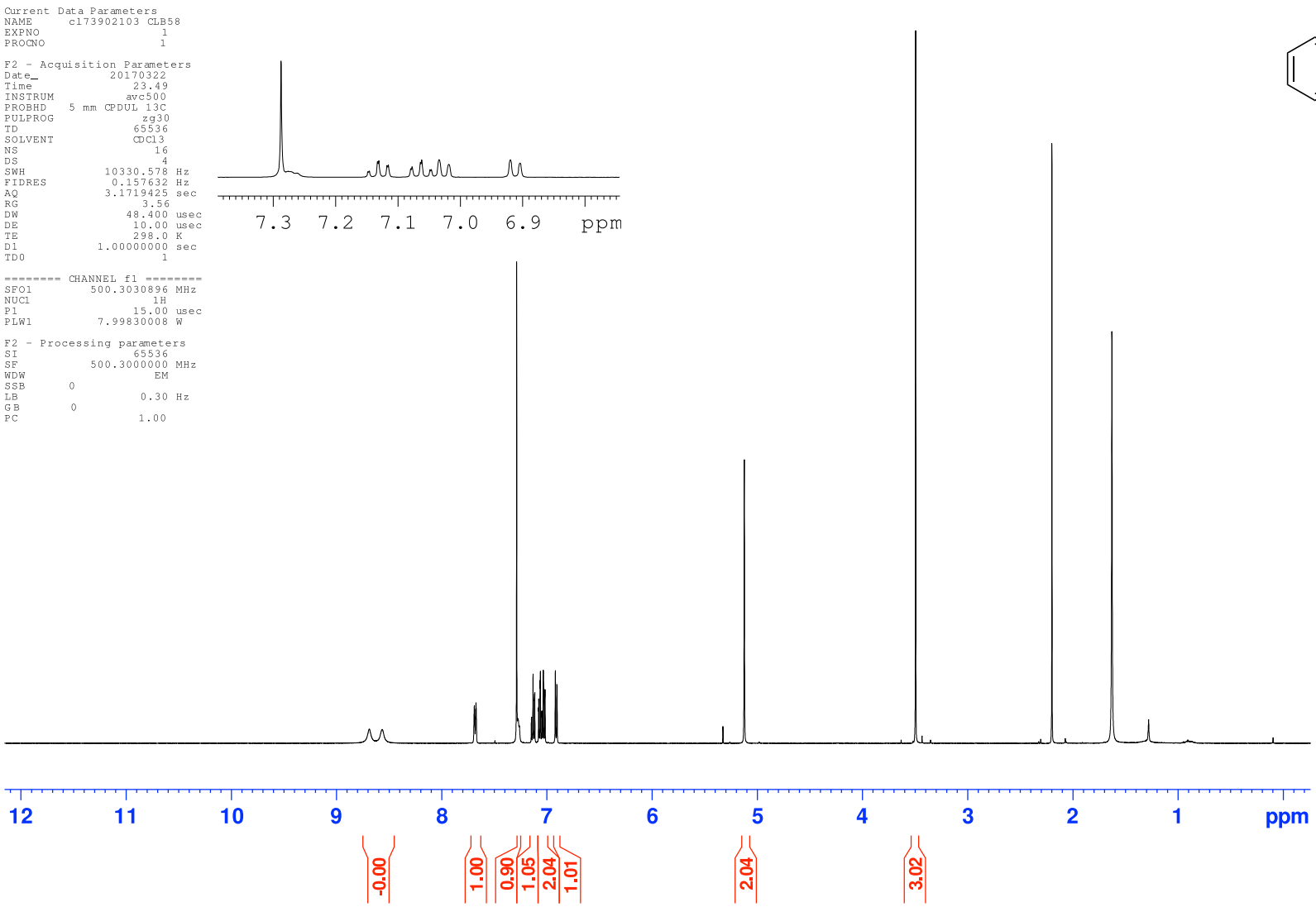

```
Current Data Parameters
NAME      Jul09-2019-2-9-CL-B-44
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_          20190709
Time              21.39 h
INSTRUM           avg400
PROBHD    Z108618_0816 (
PULPROG          zgpg30
TD               32768
SOLVENT          CDCl3
NS                 512
DS                   4
SWH           26041.666 Hz
FIDRES         1.589457 Hz
AQ            0.6291456 sec
RG               206.87
DW              19.200 usec
DE               6.50 usec
TE              301.0 K
D1           1.00000000 sec
D11          0.03000000 sec
TD0                  1
SFO1        100.64043331 MHz
NUC1                13C
P1               10.00 usec
PLW1         56.00000000 W
SFO2        400.20160008 MHz
NUC2                1H
CPDPRG[2         waltz16
PCPD2            90.00 usec
PLW2         14.00000000 W
PLW12         0.33877000 W
PLW13         0.17039999 W

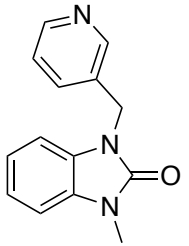
F2 - Processing parameters
SI                 32768
SF          100.6303547 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

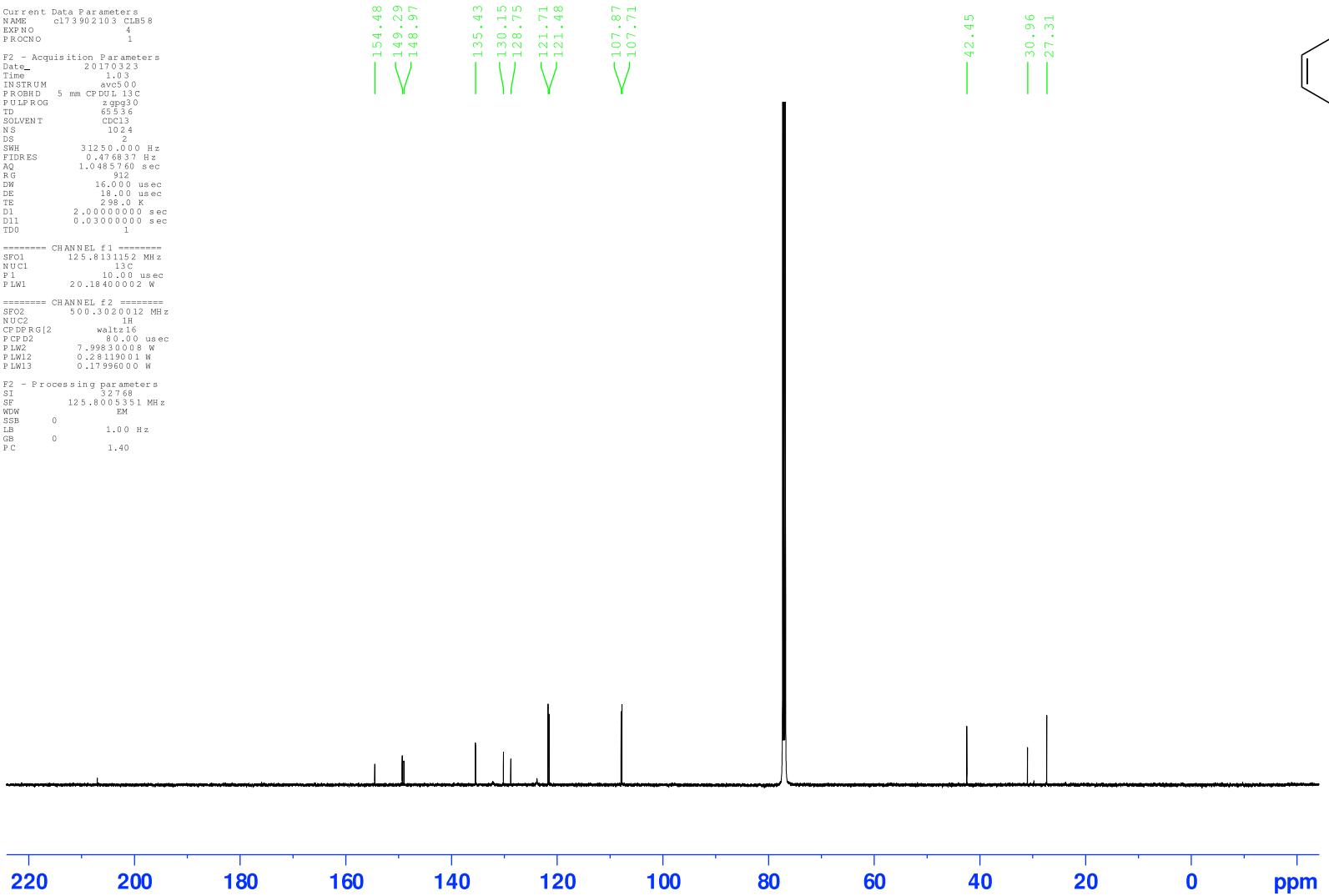
1-Methyl-3-(pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (39)



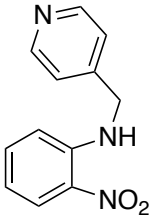


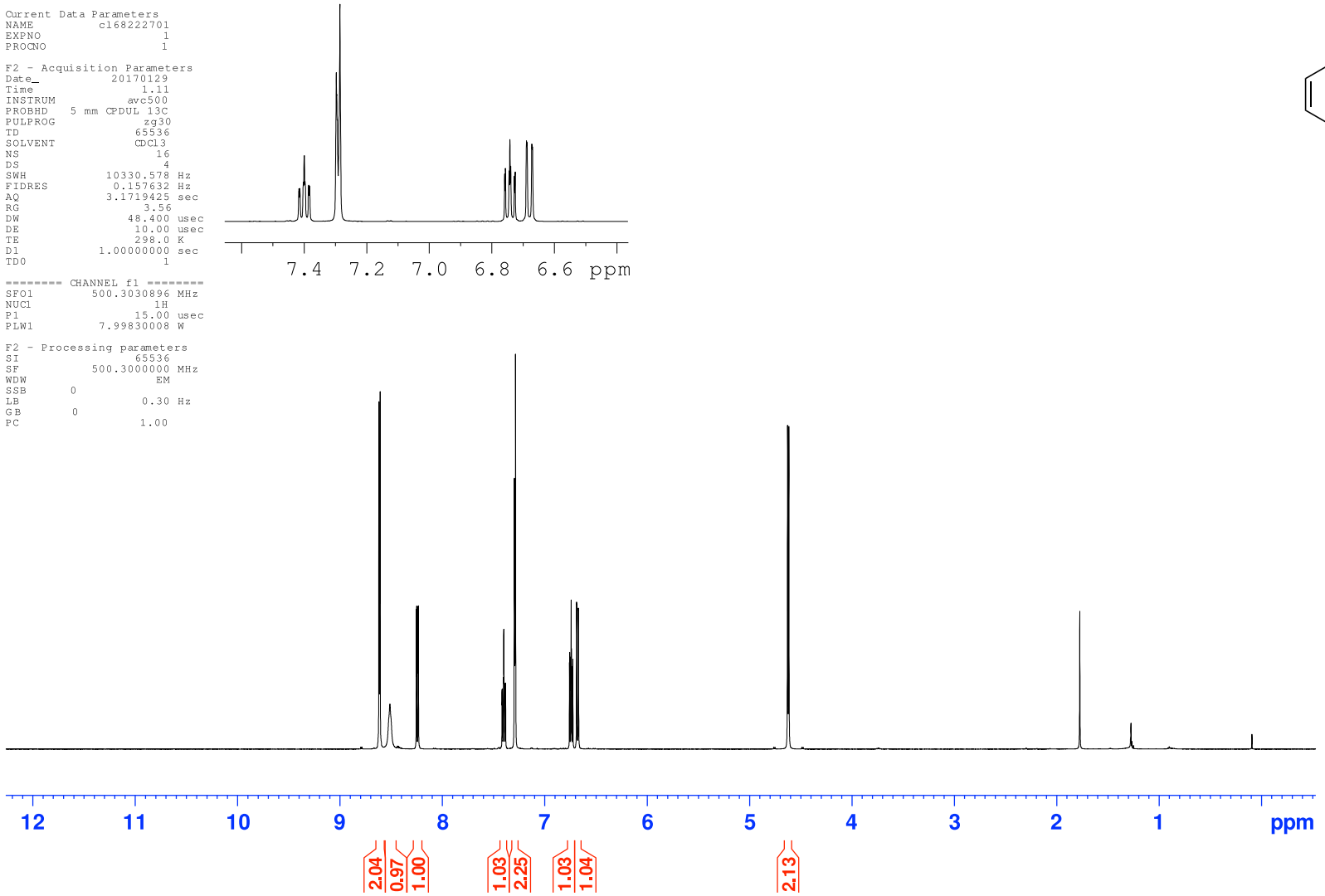
1-Methyl-3-(pyridin-3-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (39)



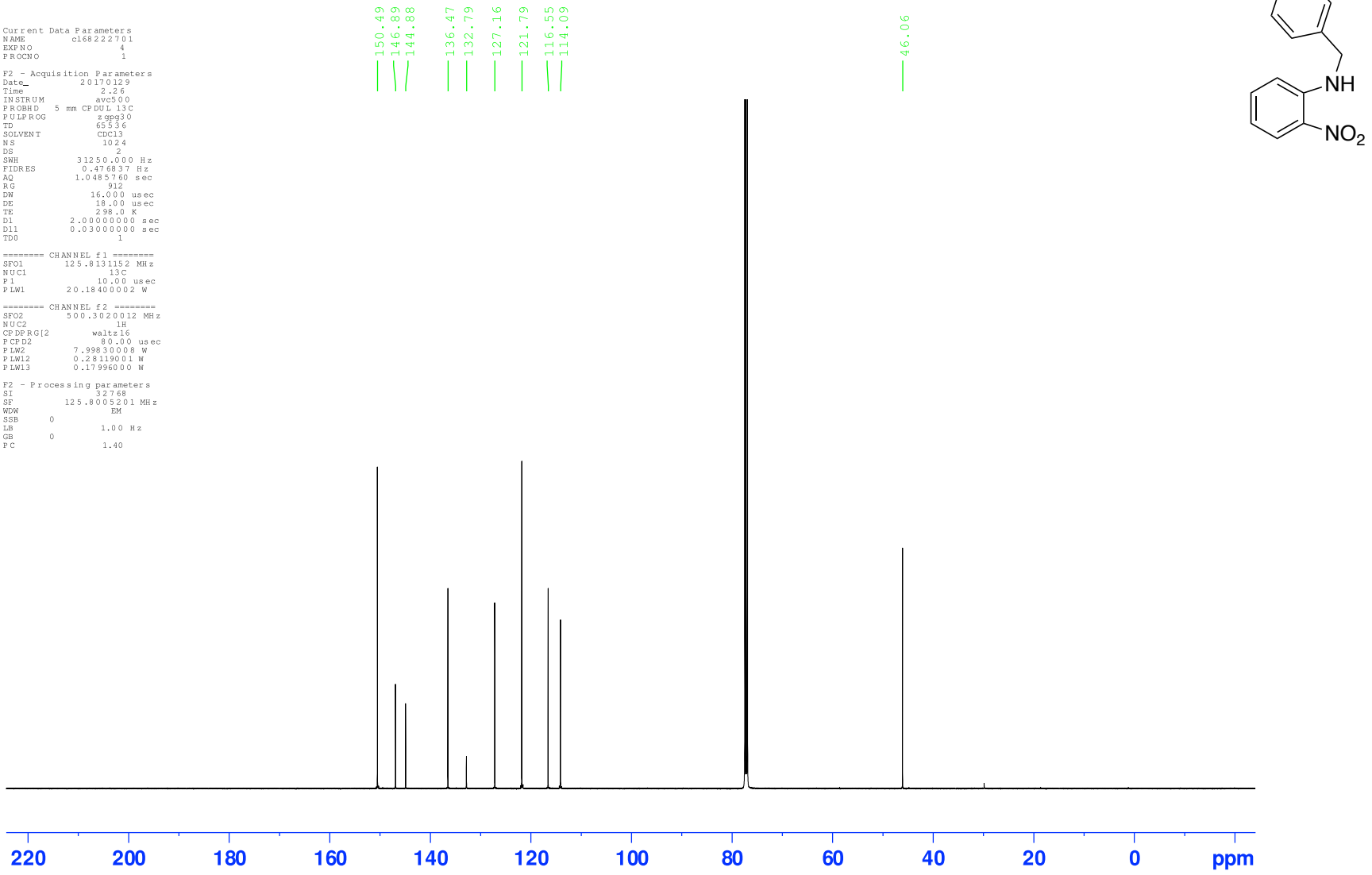


2-Nitro-*N*-(pyridin-4-ylmethyl)aniline (32)





2-Nitro-*N*-(pyridin-4-ylmethyl)aniline (32)




```
Current Data Parameters
NAME          c168222701
EXPNO                  4
PROCNO                 1

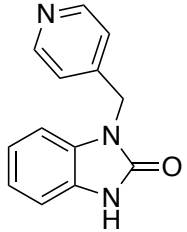
F2 - Acquisition Parameters
Date_          20170129
Time               2.26
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD               65536
SOLVENT          CDCl3
NS                1024
DS                   2
SWH          31250.000 Hz
FIDRES        0.476837 Hz
AQ           1.04857600 sec
RG                 912
DW              16.000 usec
DE              18.00 usec
TE              298.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                  1

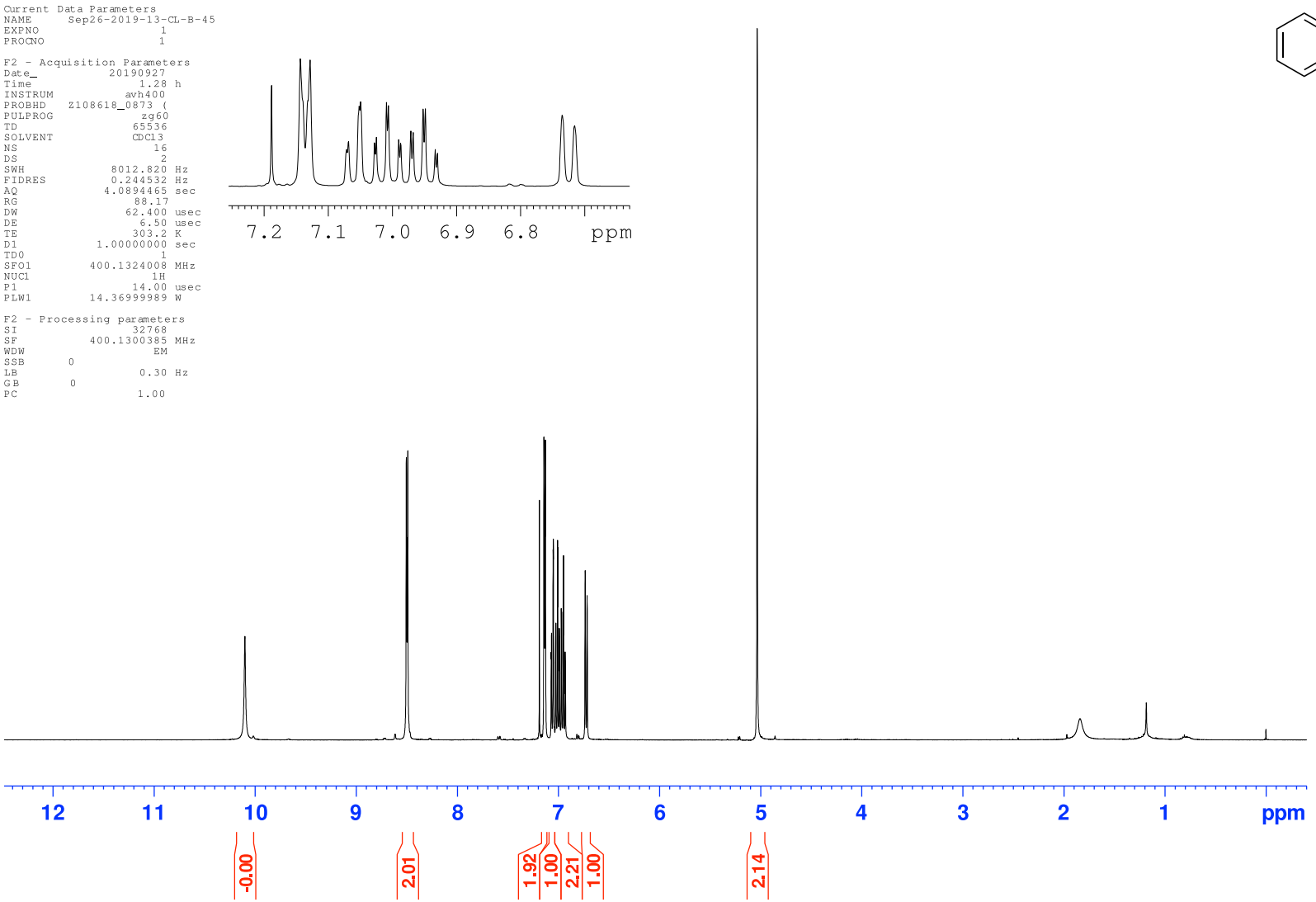
======== CHANNEL f1 ========
SFO1        125.8131152 MHz
NUC1               13C
P1              10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2               1H
CPDPRG[2         waltz16
PCPD2           80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

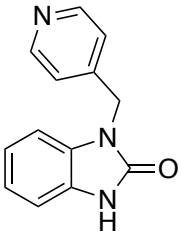
F2 - Processing parameters
SI               32768
SF          125.8005201 MHz
WDW                 EM
SSB      0
LB              1.00 Hz
GB       0
PC              1.40
```

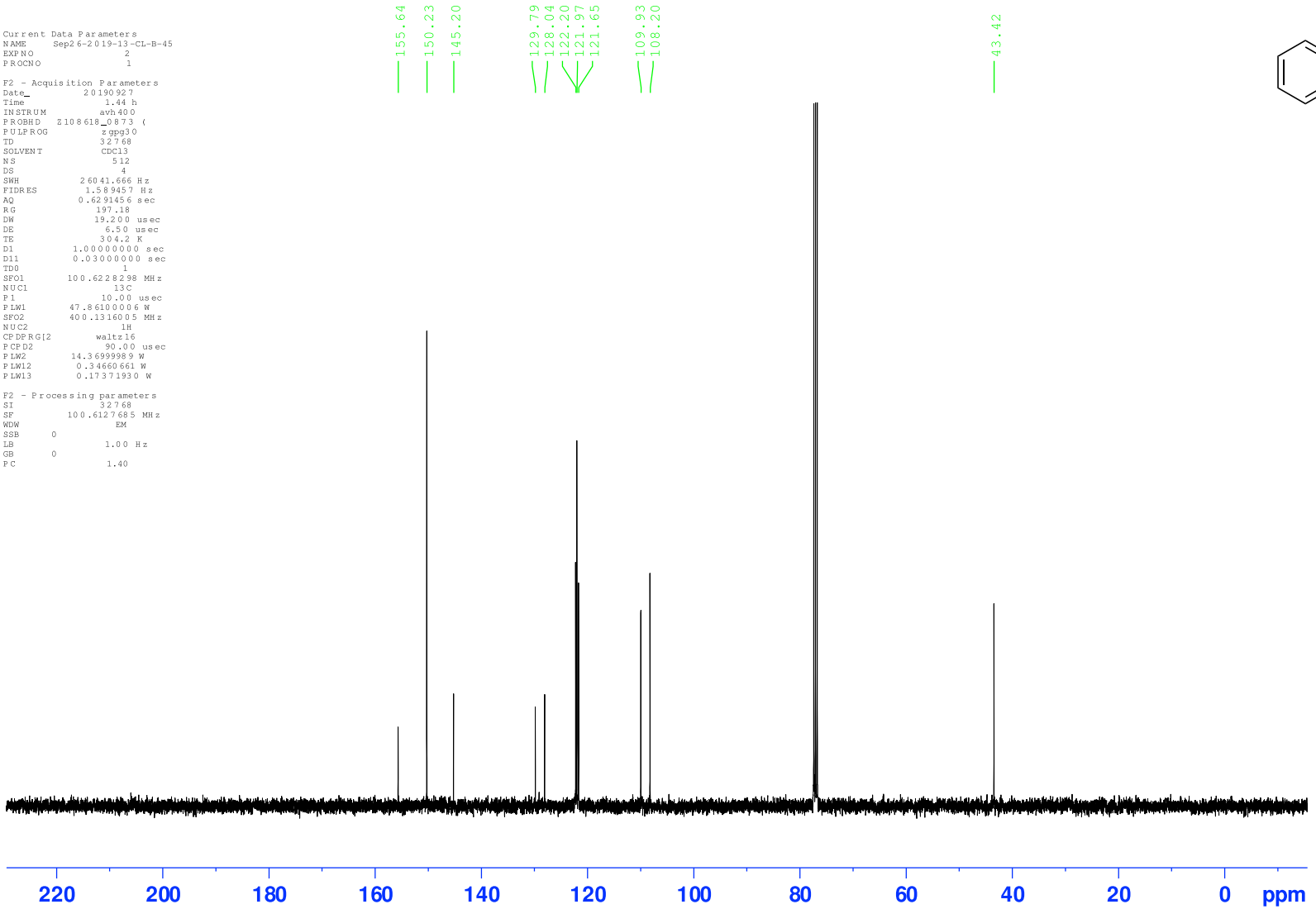
1-(Pyridin-4'-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (36)



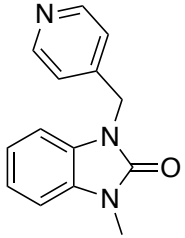


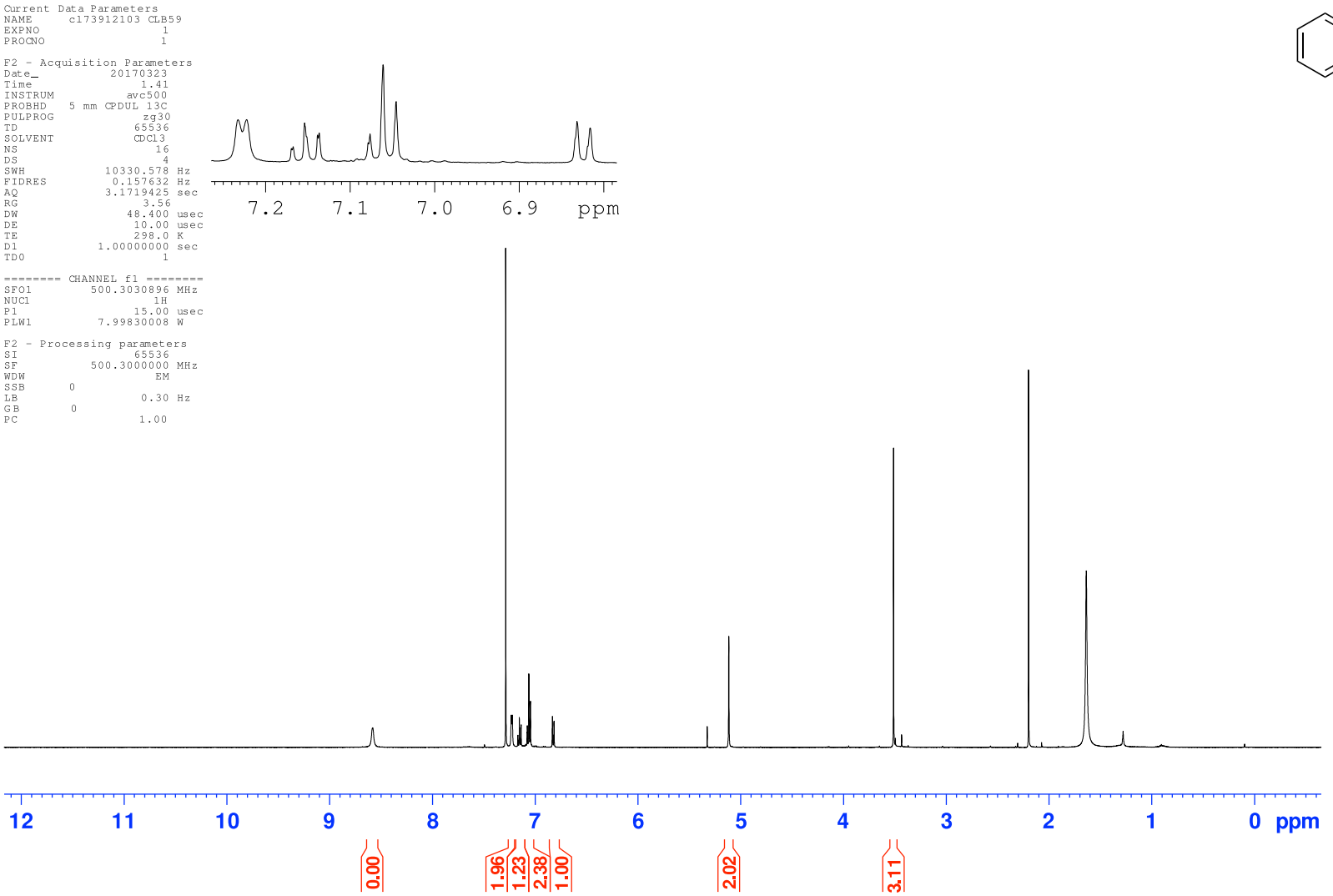
1-(Pyridin-4’-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (36)



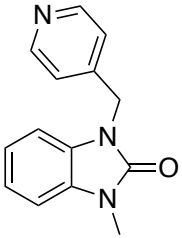


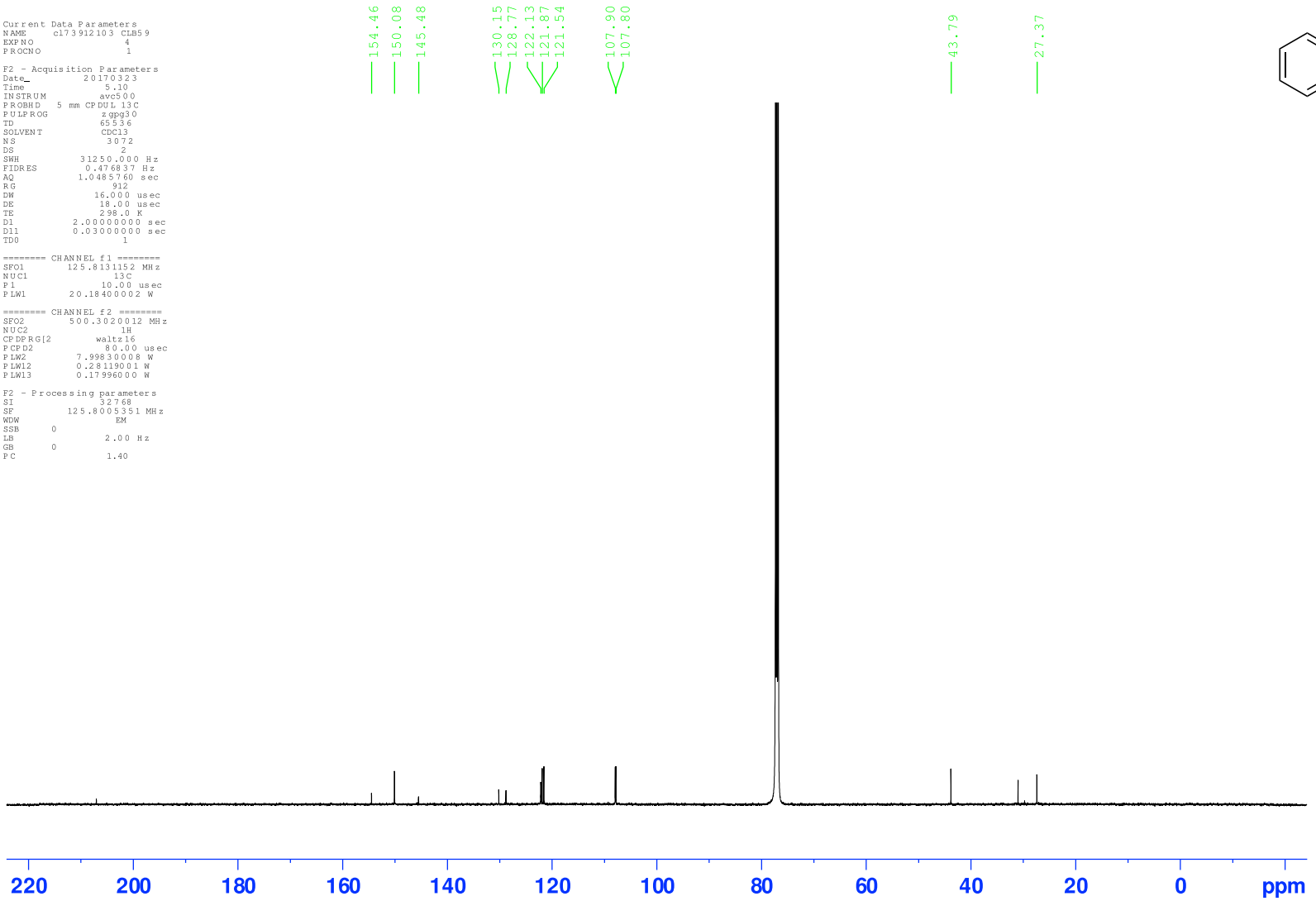
1-Methyl-3-(pyridin-4’-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (40)



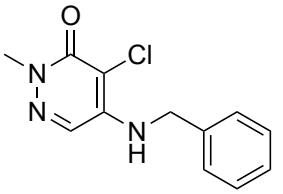


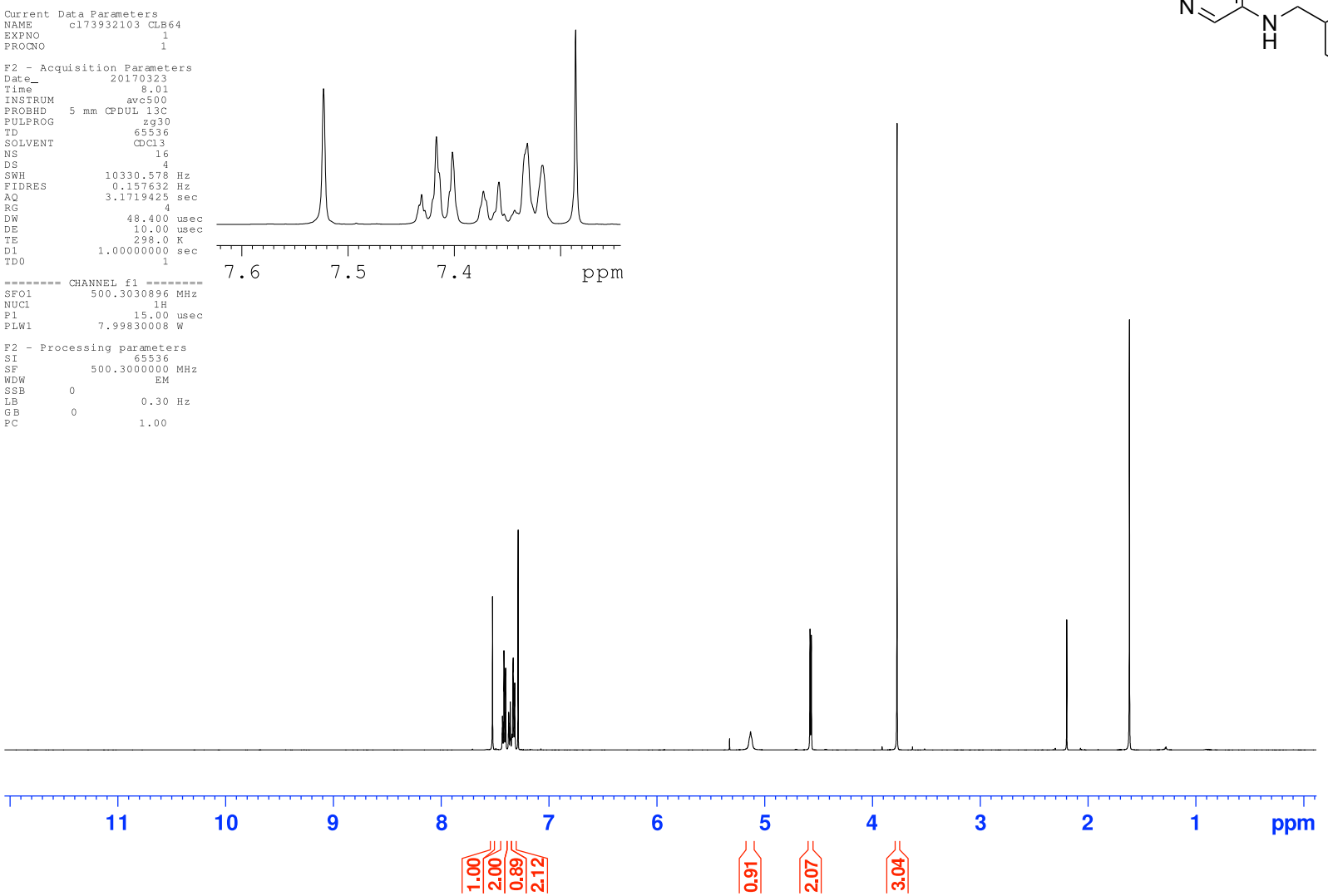
1-Methyl-3-(pyridin-4’-ylmethyl)-1,3-dihydro-2*H*-benzo[*d*]imidazol-2-one (40)



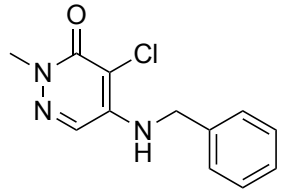


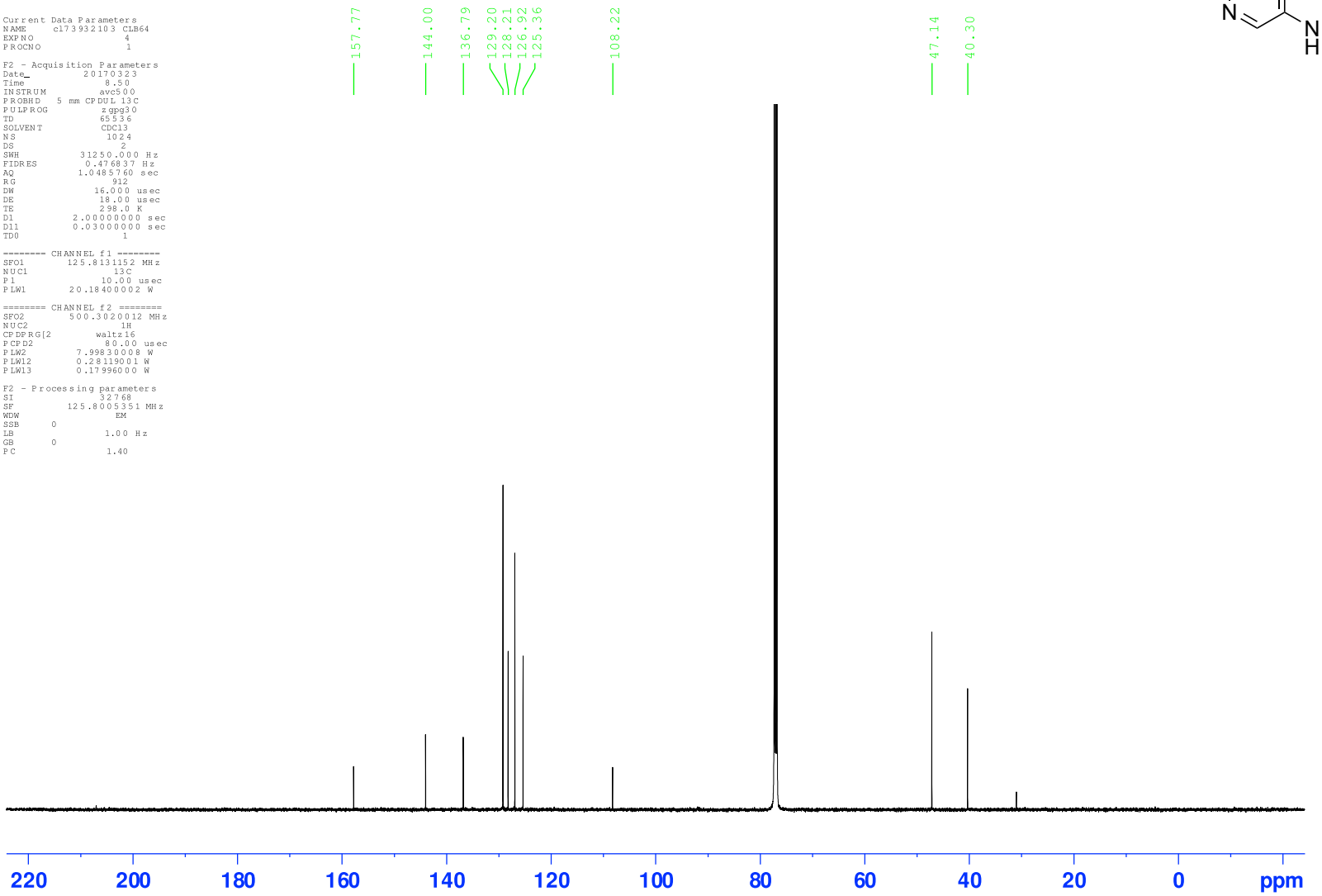
5-(Benzylamino)-4-chloro-2-methylpyridazin-3(2*H*)-one (67)



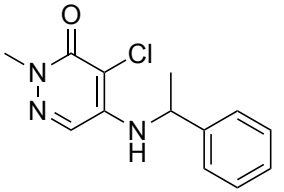


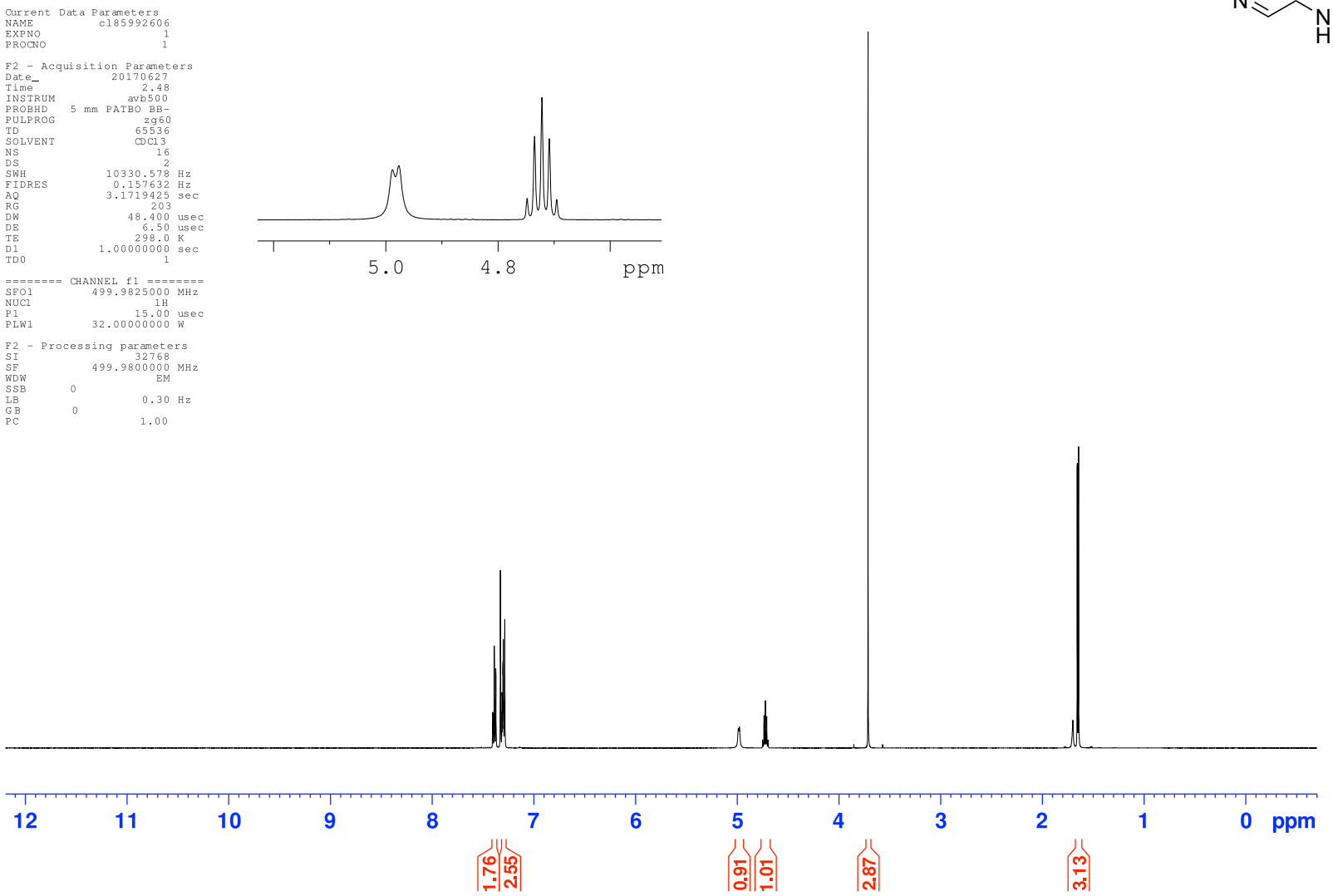
5-(Benzylamino)-4-chloro-2-methylpyridazin-3(2*H*)-one (67)



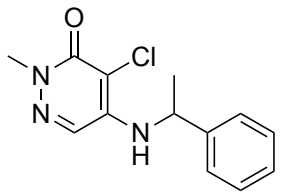


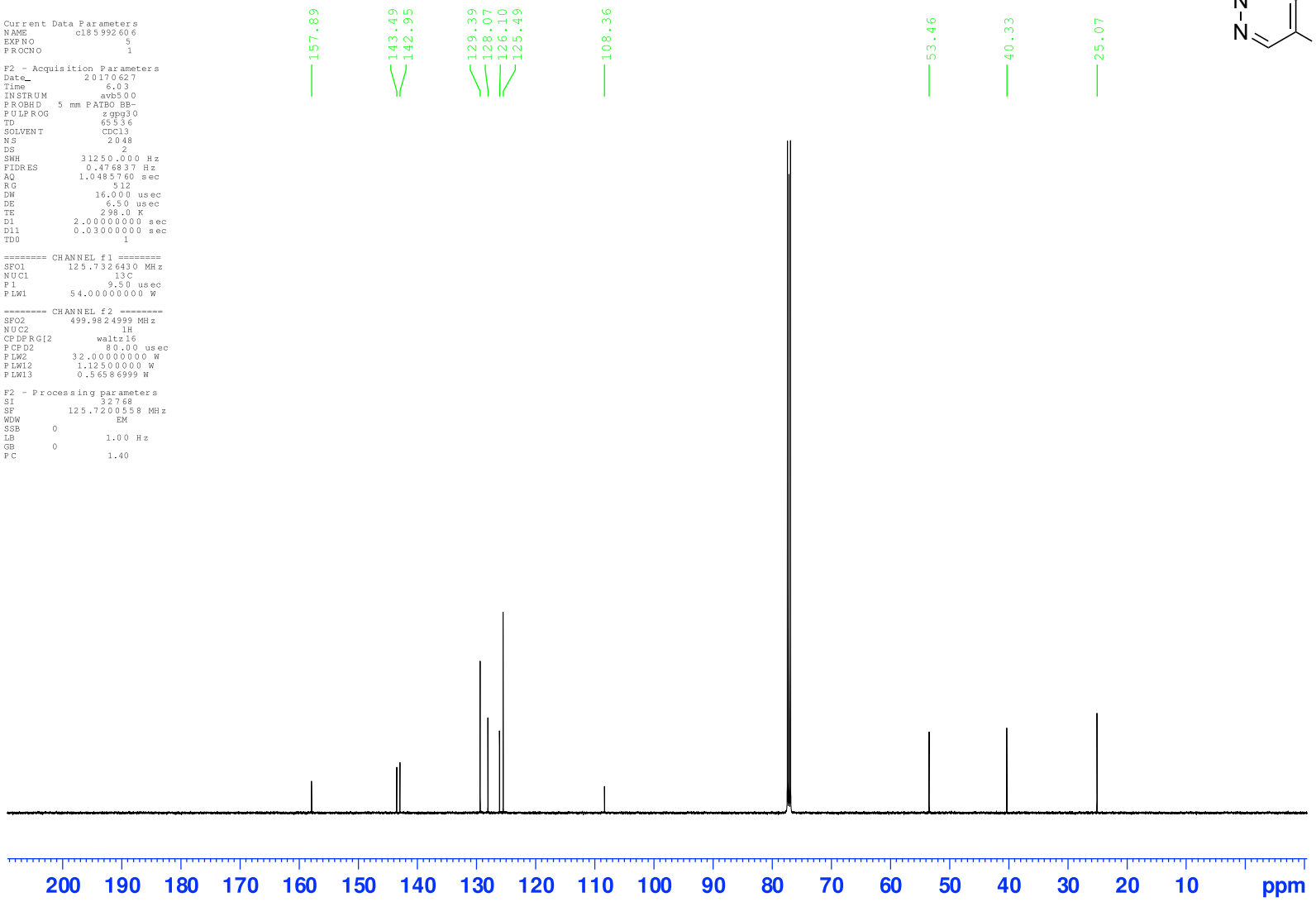
4-Chloro-2-methyl-5-((1-phenylethyl)amino)pyridazin-3(2*H*)-one (68)



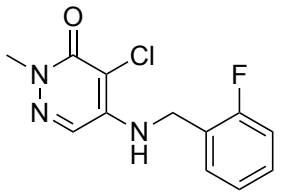


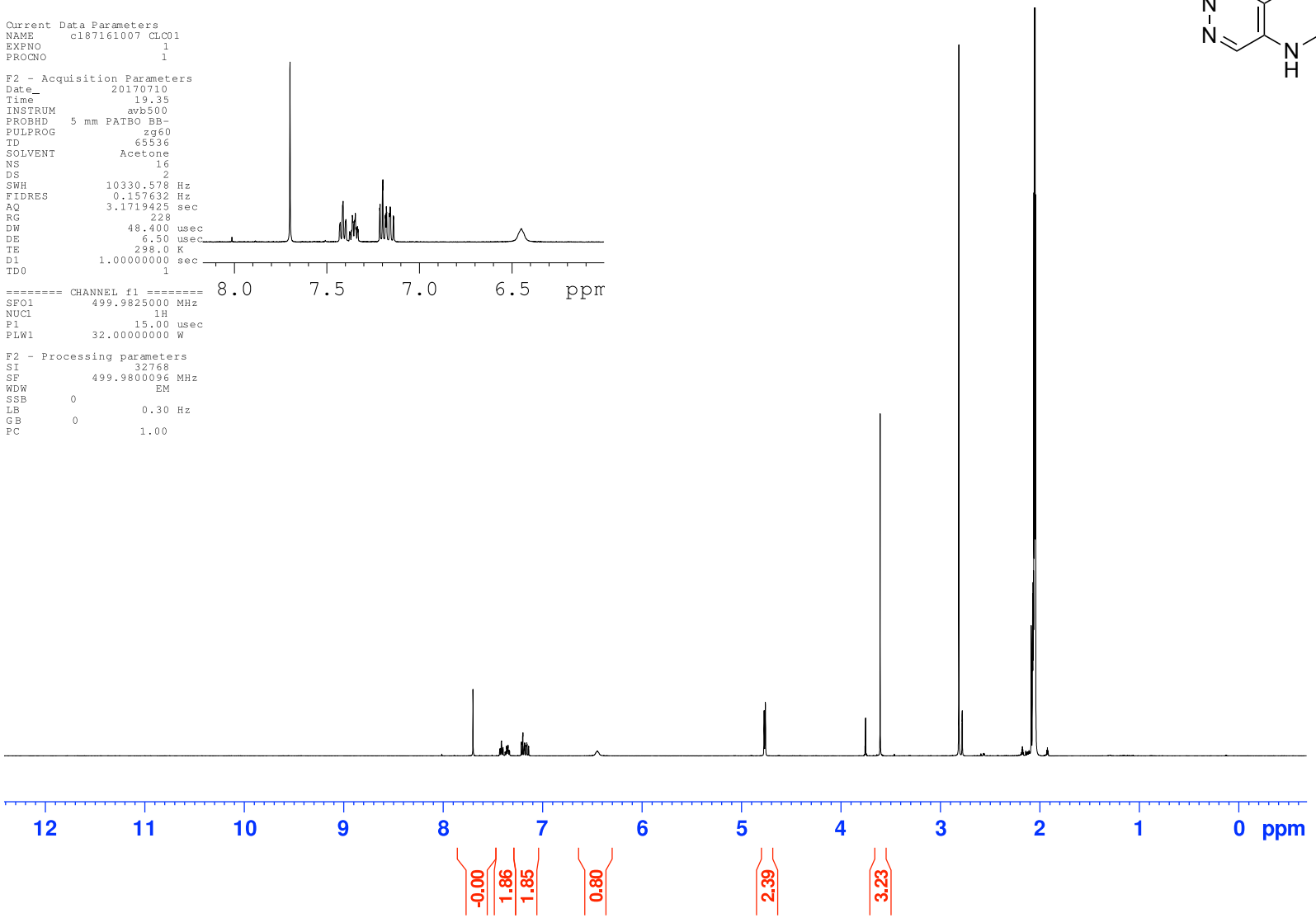
4-Chloro-2-methyl-5-((1-phenylethyl)amino)pyridazin-3(2*H*)-one (68)



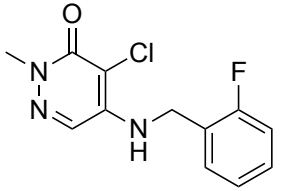


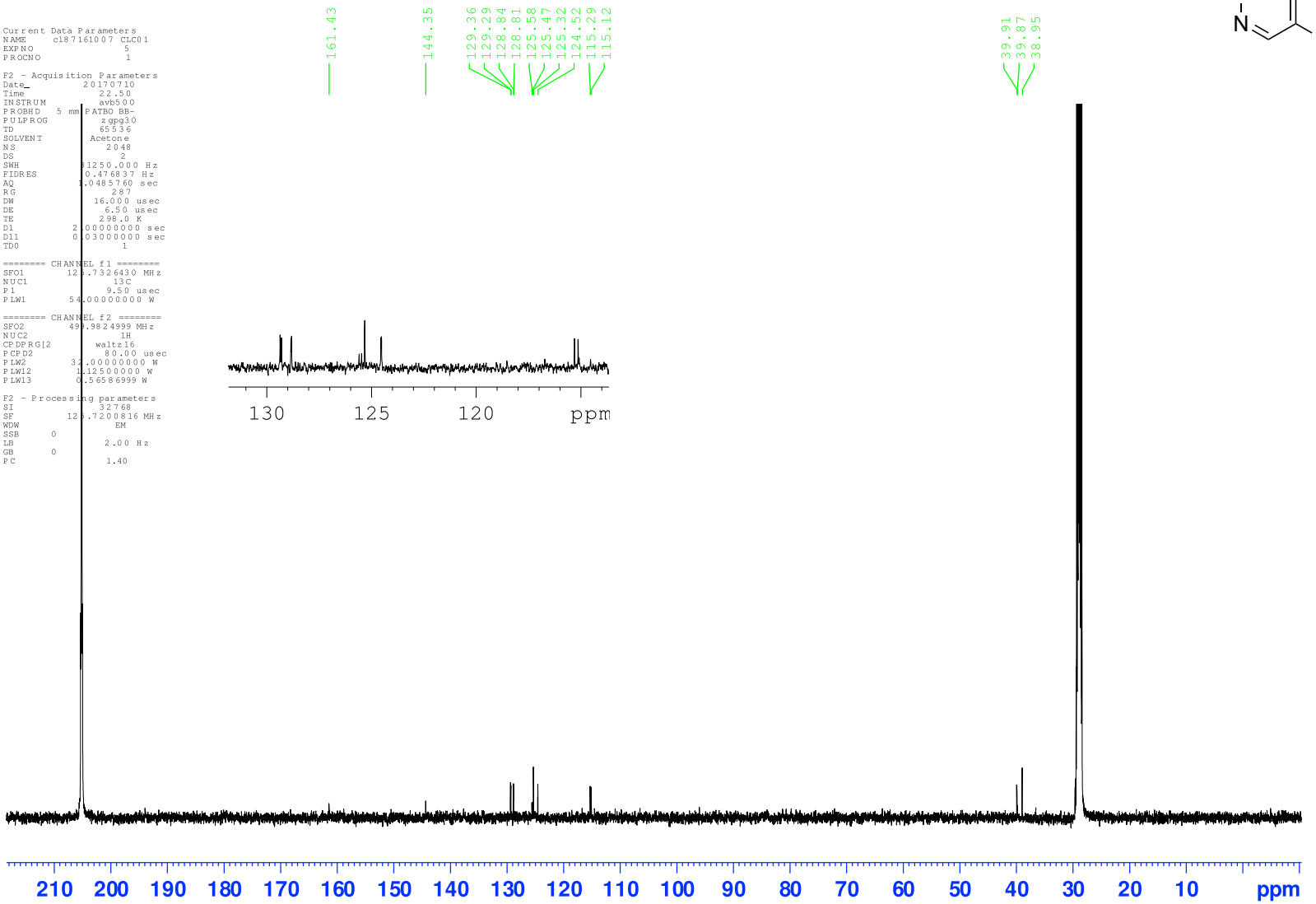
4-Chloro-5-((2-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (71)



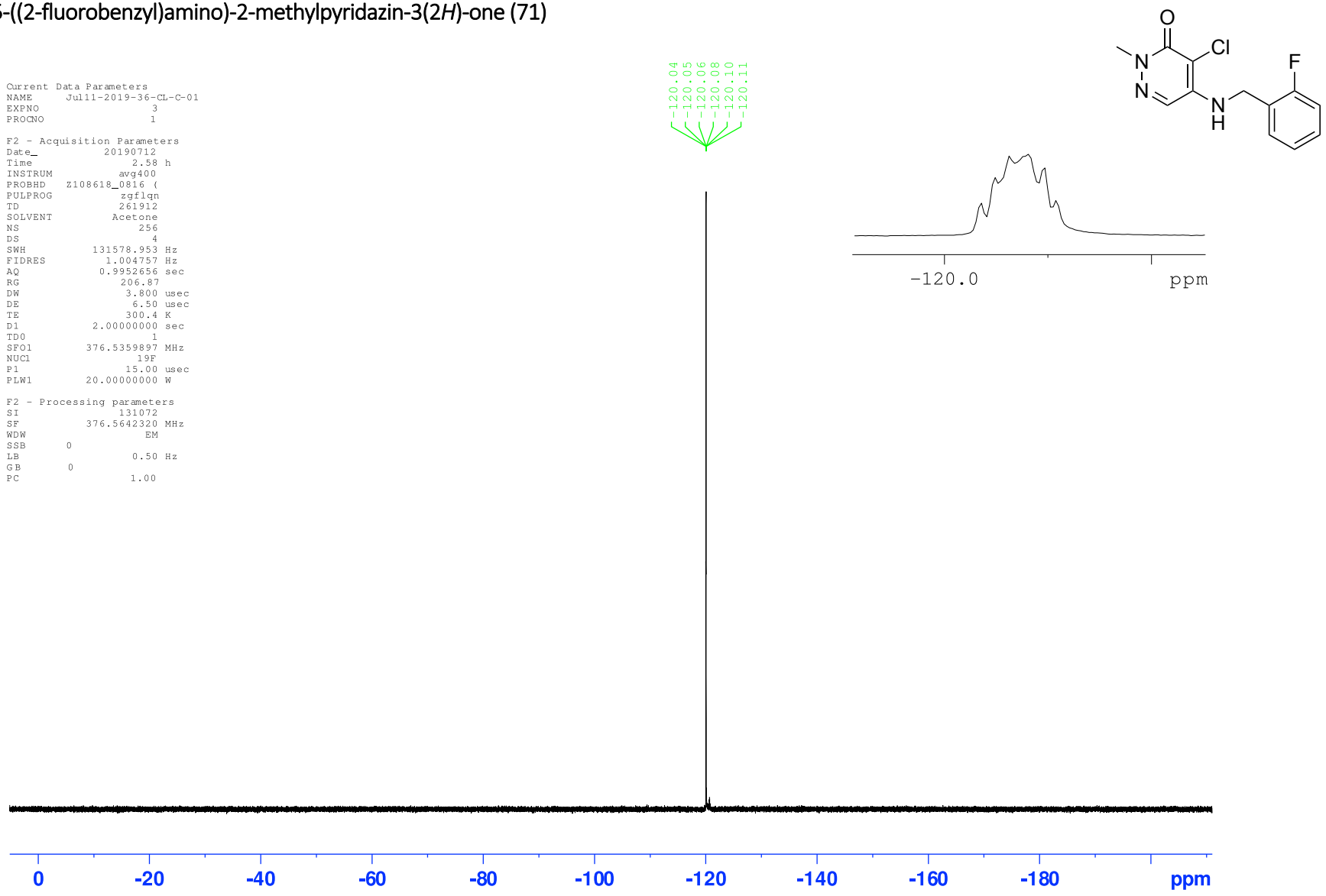


4-Chloro-5-((2-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (71)

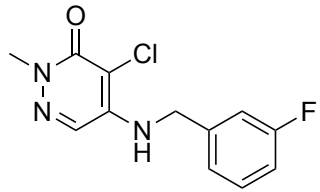


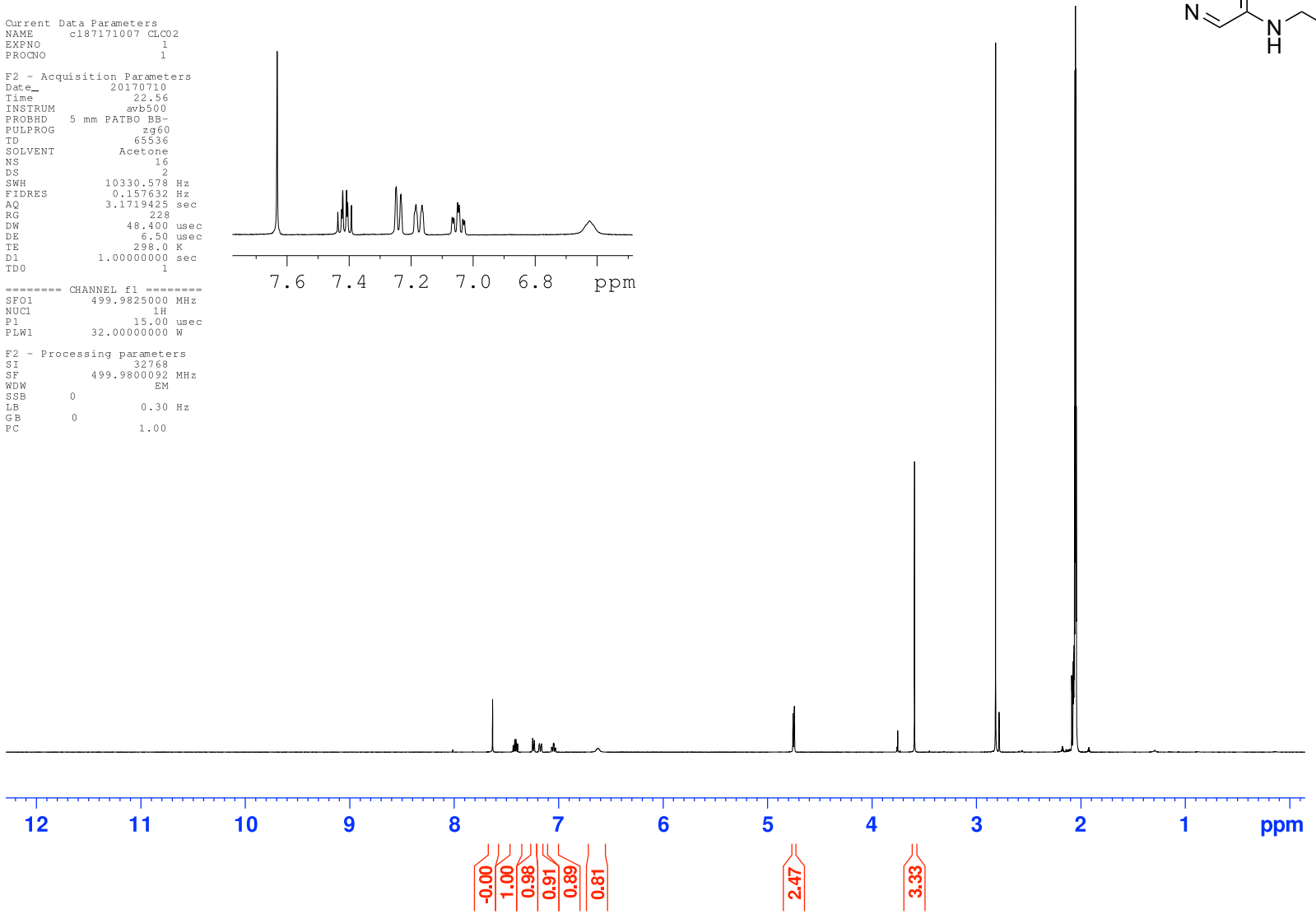


4-Chloro-5-((2-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (71)

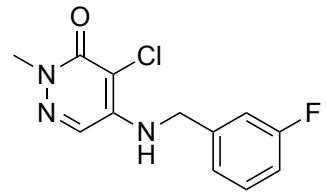


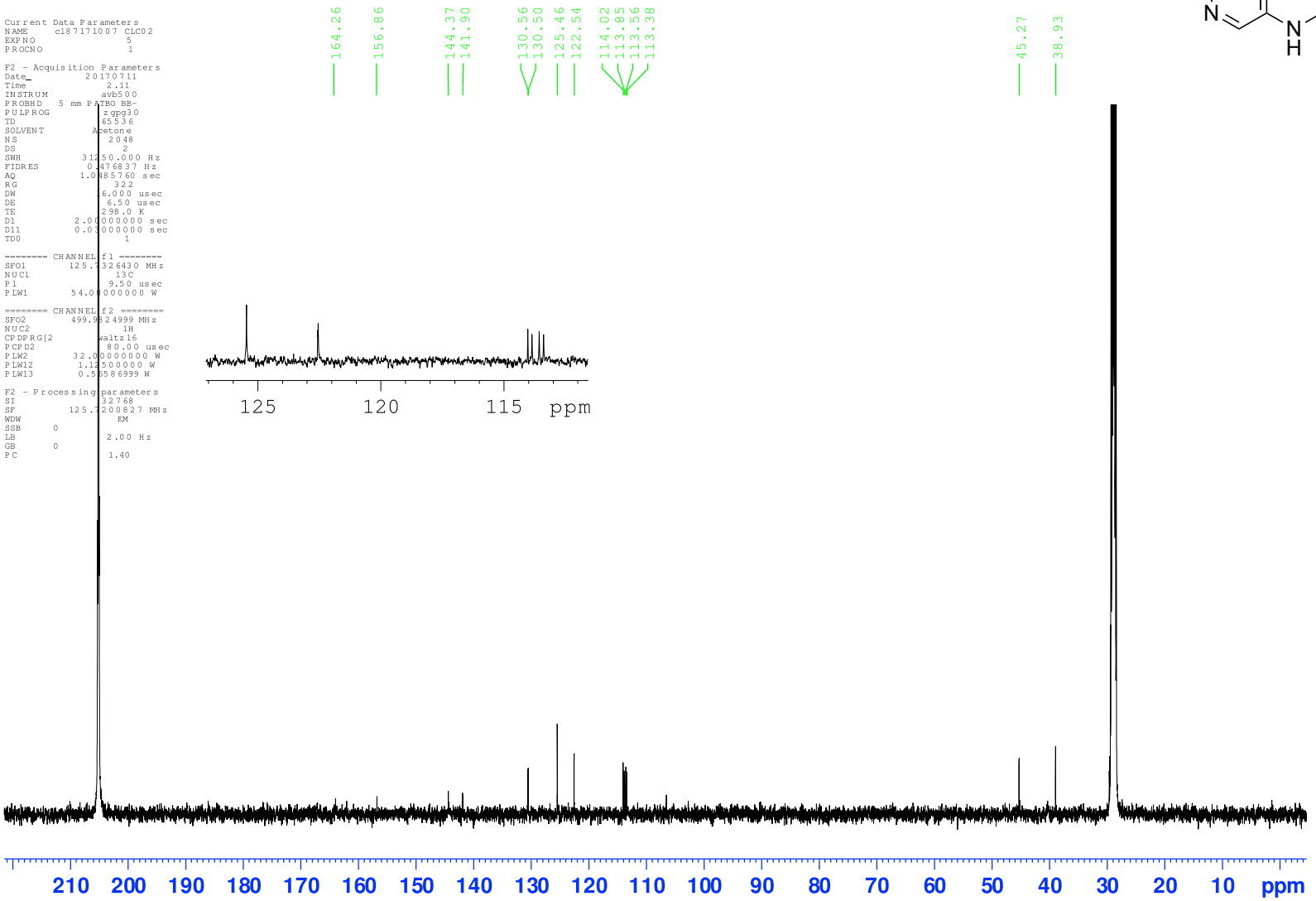
4-Chloro-5-((3-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (72)



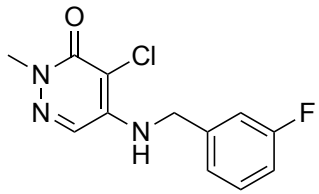


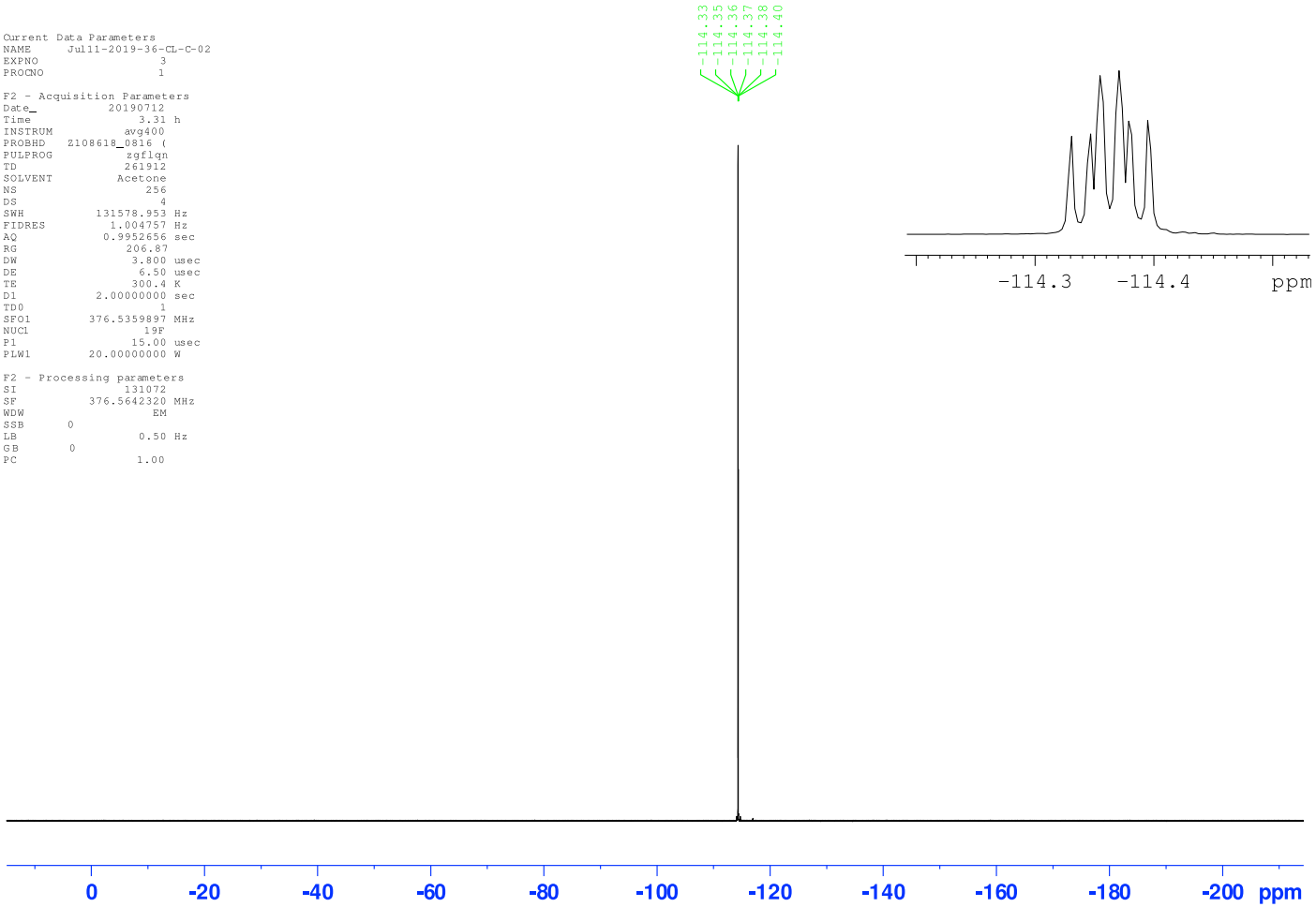
4-Chloro-5-((3-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (72)



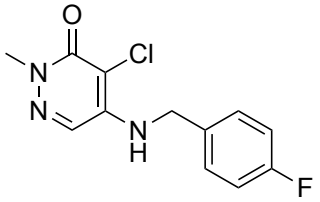


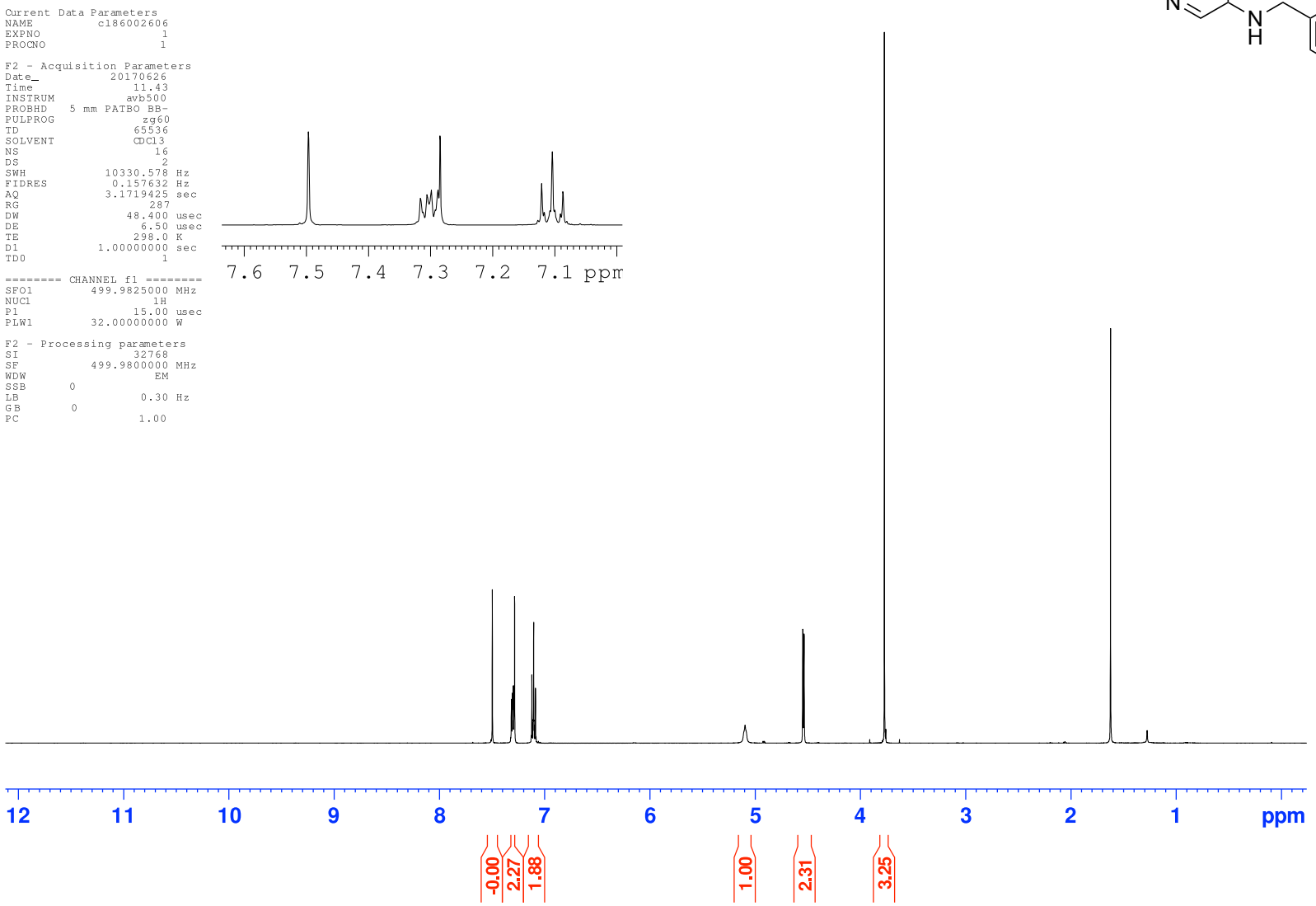
4-Chloro-5-((3-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (72)



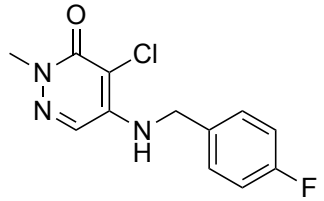


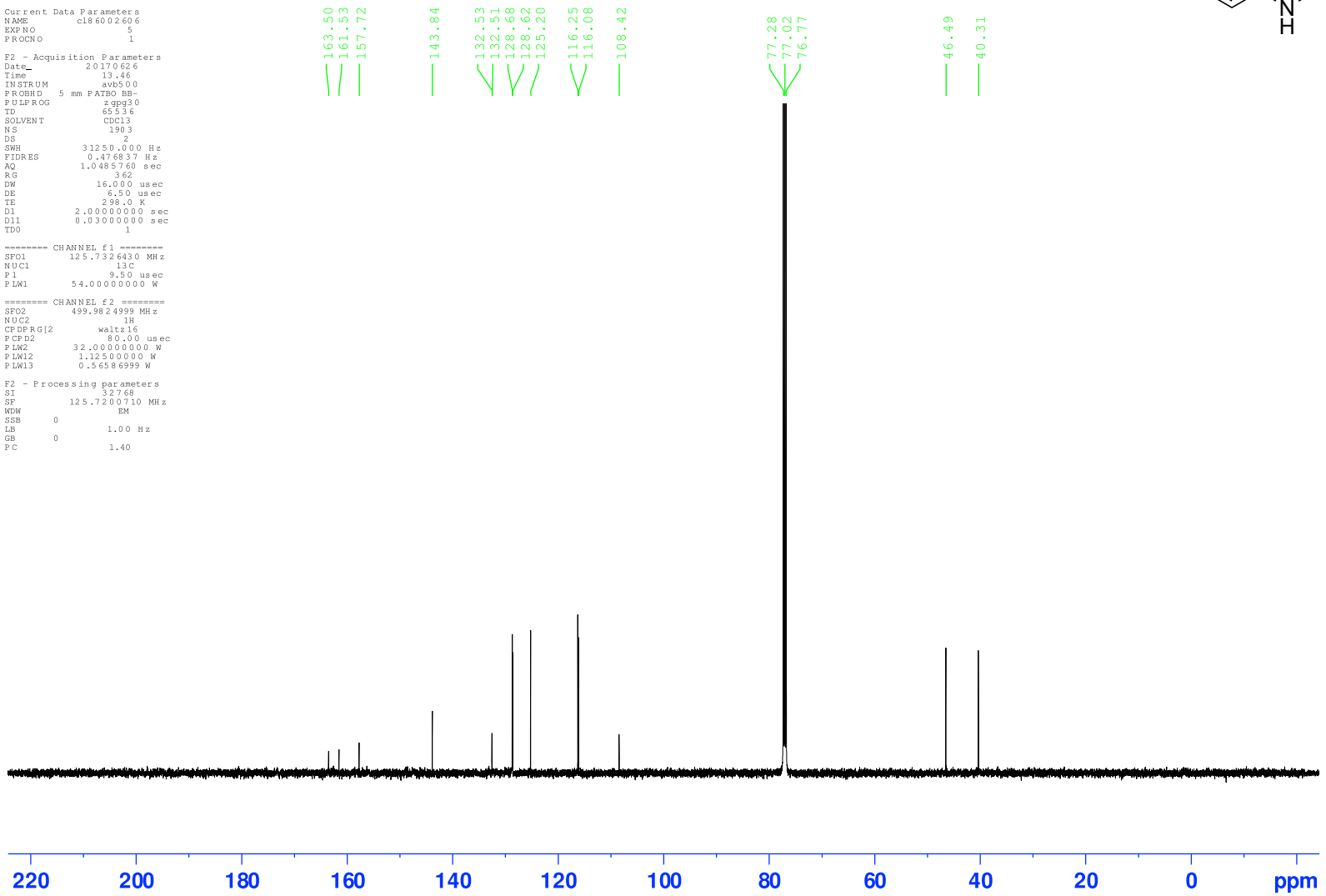
4-Chloro-5-((4-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (73)



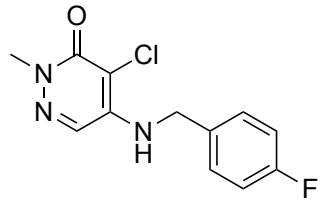


4-Chloro-5-((4-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (73)





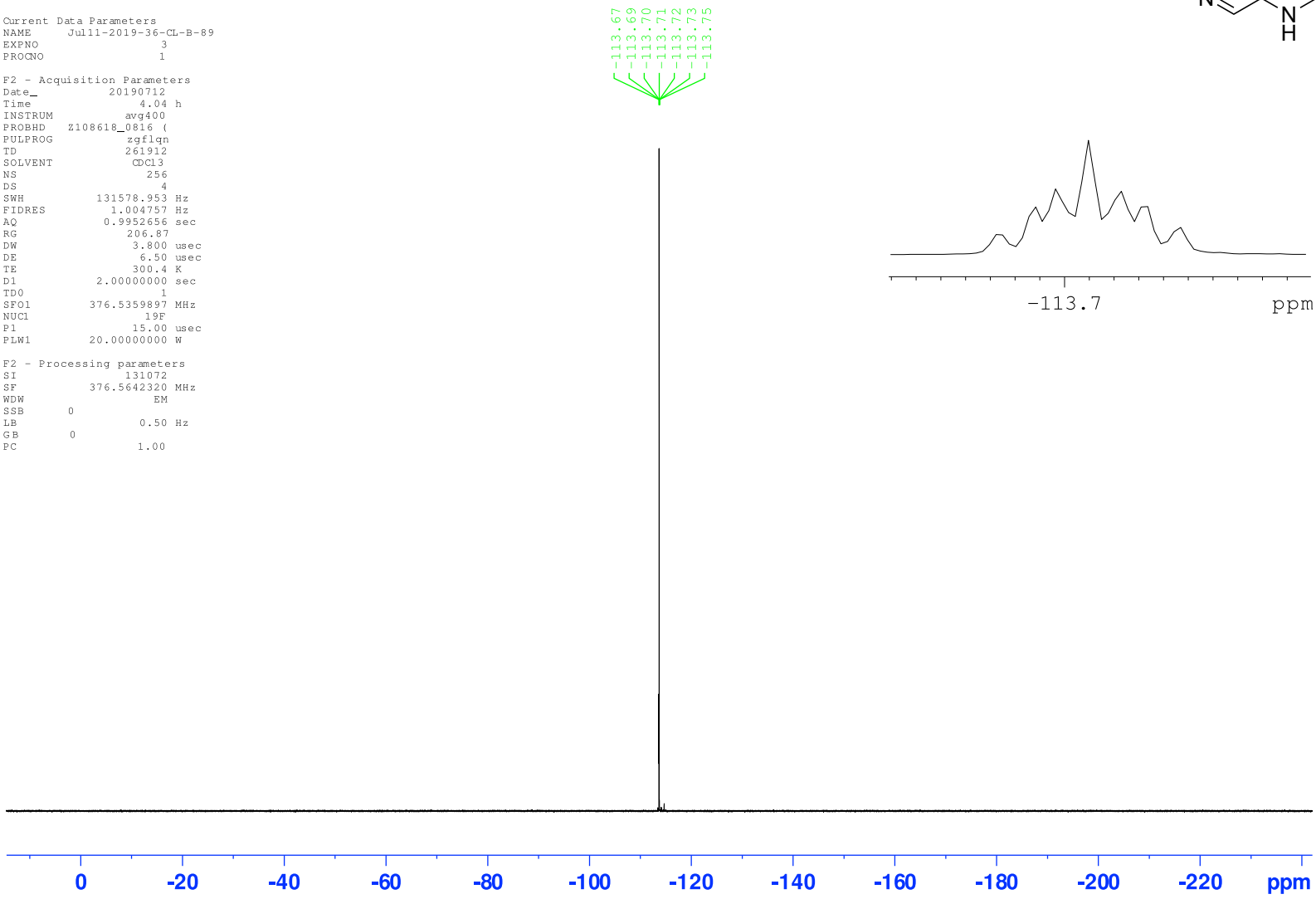
4-Chloro-5-((4-fluorobenzyl)amino)-2-methylpyridazin-3(2*H*)-one (73)



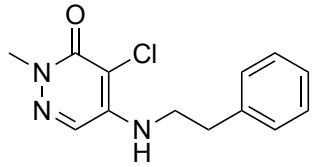

```
Current Data Parameters
NAME      Jul11-2019-36-CL-B-89
EXPNO                        3
PROCNO                       1

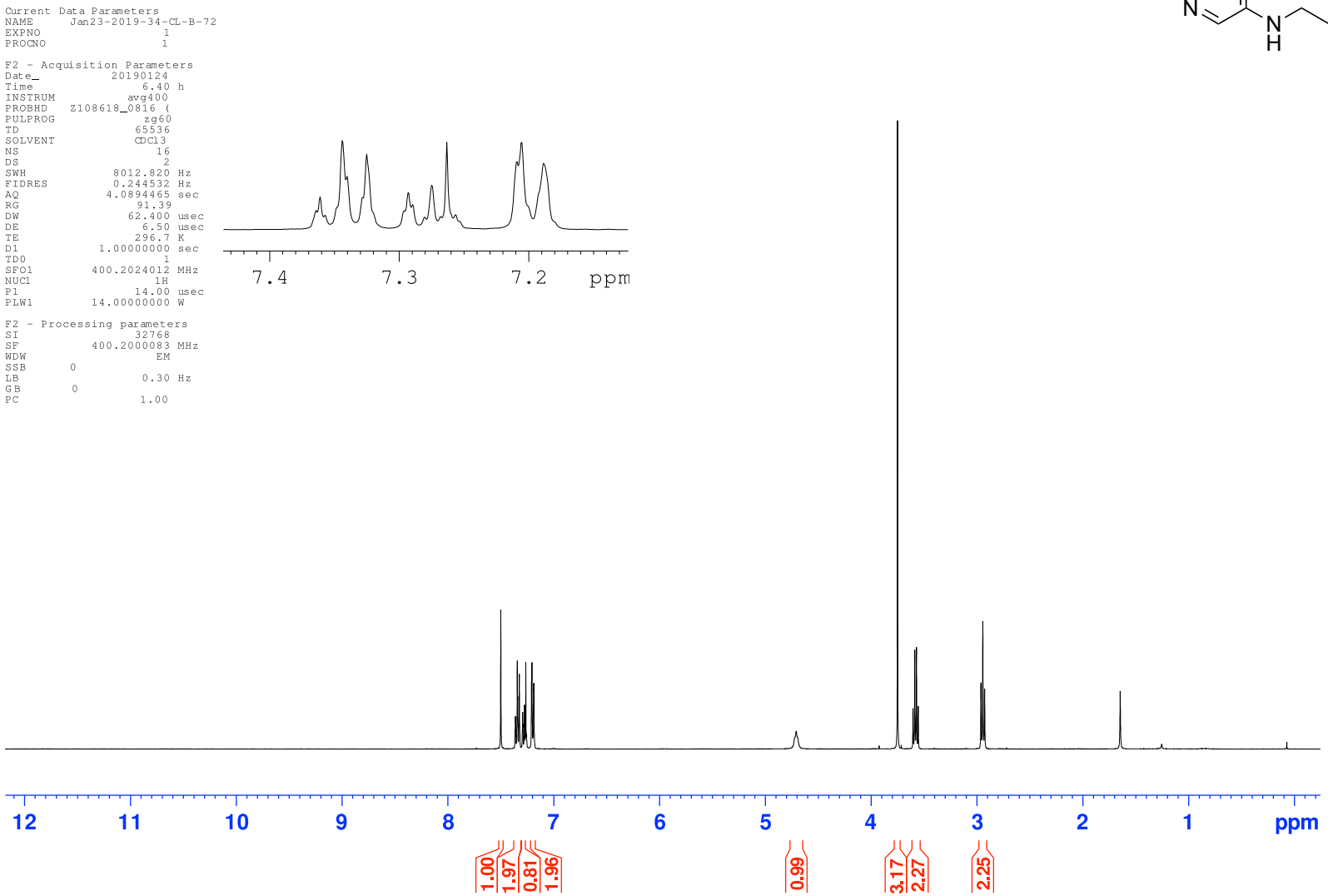
F2 - Acquisition Parameters
Date_              20190712
Time                 4.04 h
INSTRUM            avg400
PROBHD    Z108618_0816 (
PULPROG            zgflqn
TD                 261912
SOLVENT              CDCl3
NS                     256
DS                       4
SWH           131578.953 Hz
FIDRES          1.004757 Hz
AQ              0.9952656 sec
RG                 206.87
DW                  3.800 usec
DE                  6.50 usec
TE                300.4 K
D1            2.00000000 sec
TD0                      1
SFO1         376.5359897 MHz
NUC1                  19F
P1                  15.00 usec
PLW1         20.00000000 W

F2 - Processing parameters
SI                  131072
SF           376.5642320 MHz
WDW                    EM
SSB       0
LB                  0.50 Hz
GB        0
PC                  1.00
```

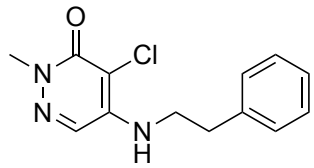


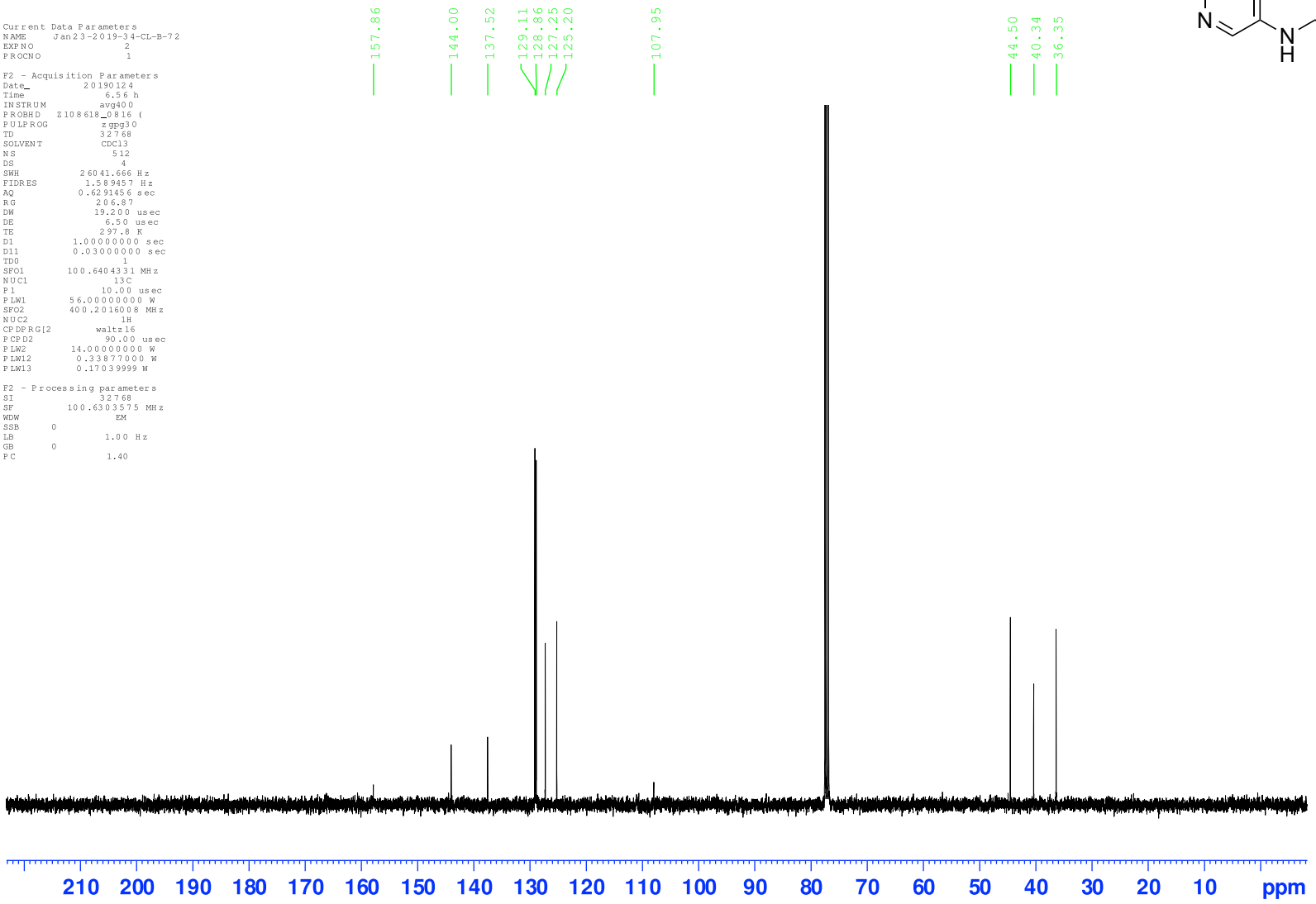
4-Chloro-2-methyl-5-(phenethylamino)pyridazin-3(2*H*)-one (69)



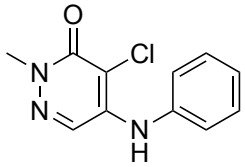


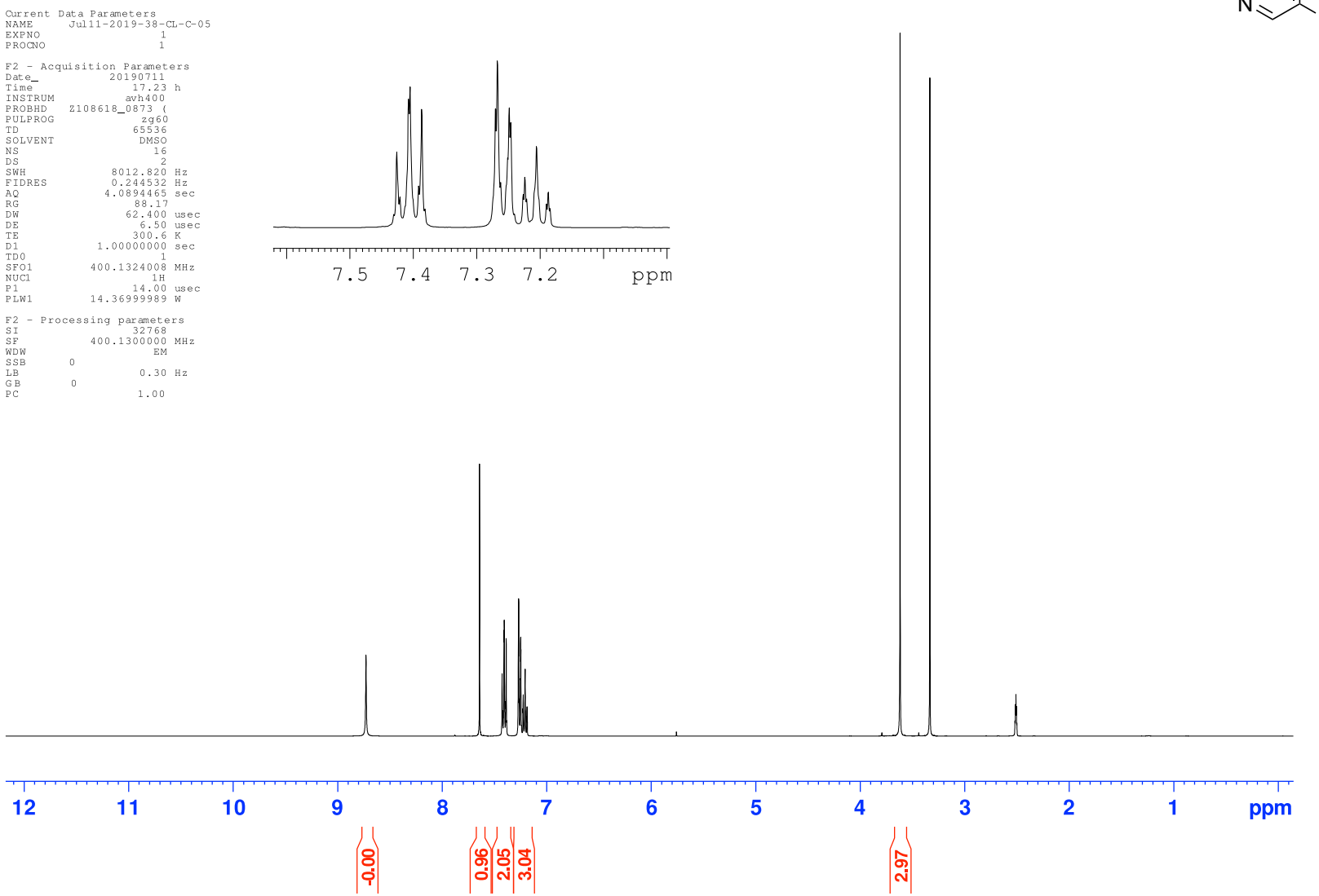
4-Chloro-2-methyl-5-(phenethylamino)pyridazin-3(2*H*)-one (69)



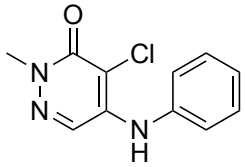


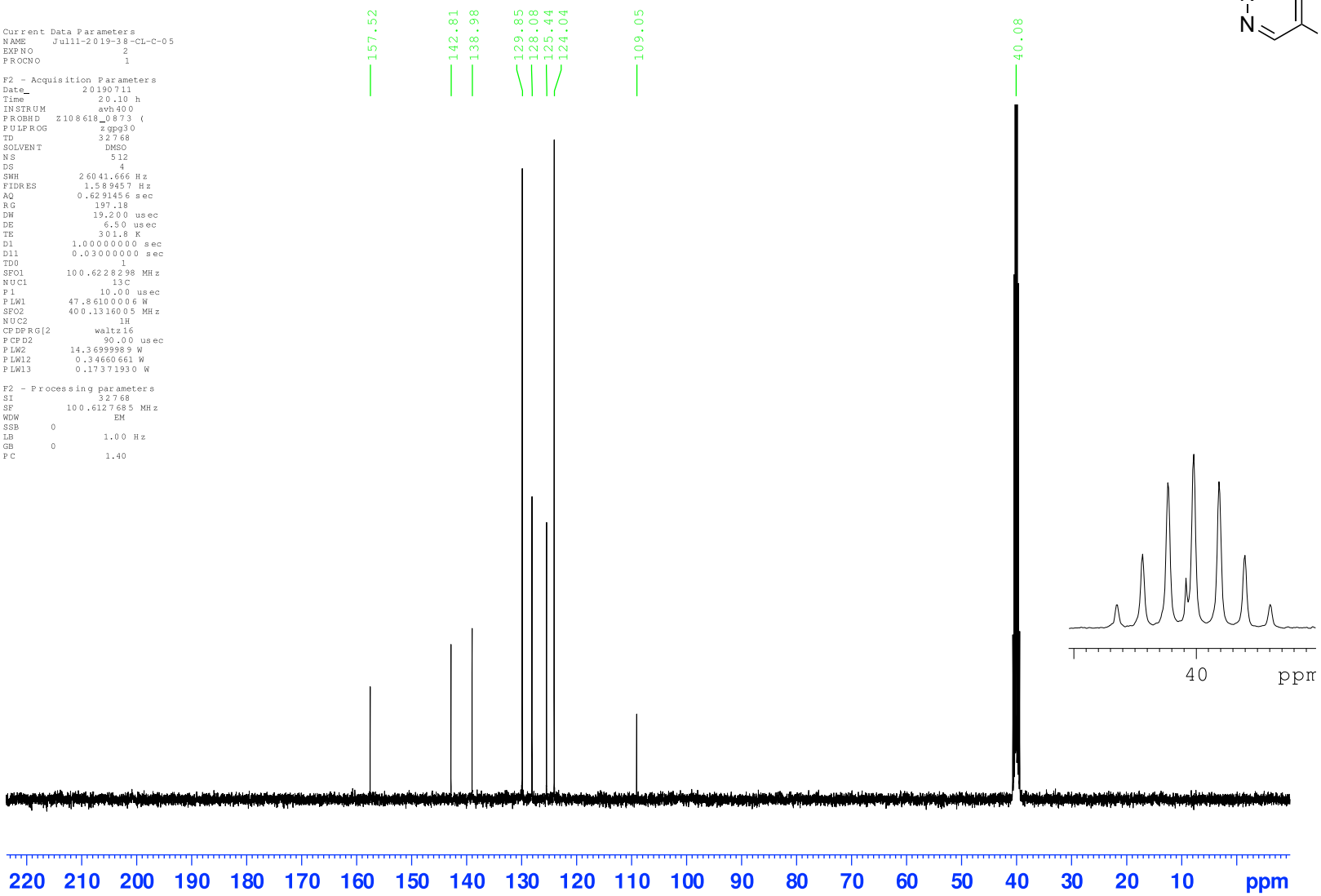
4-Chloro-2-methyl-5-(phenylamino)pyridazin-3(2*H*)-one (66)



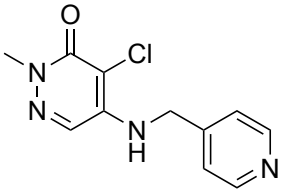


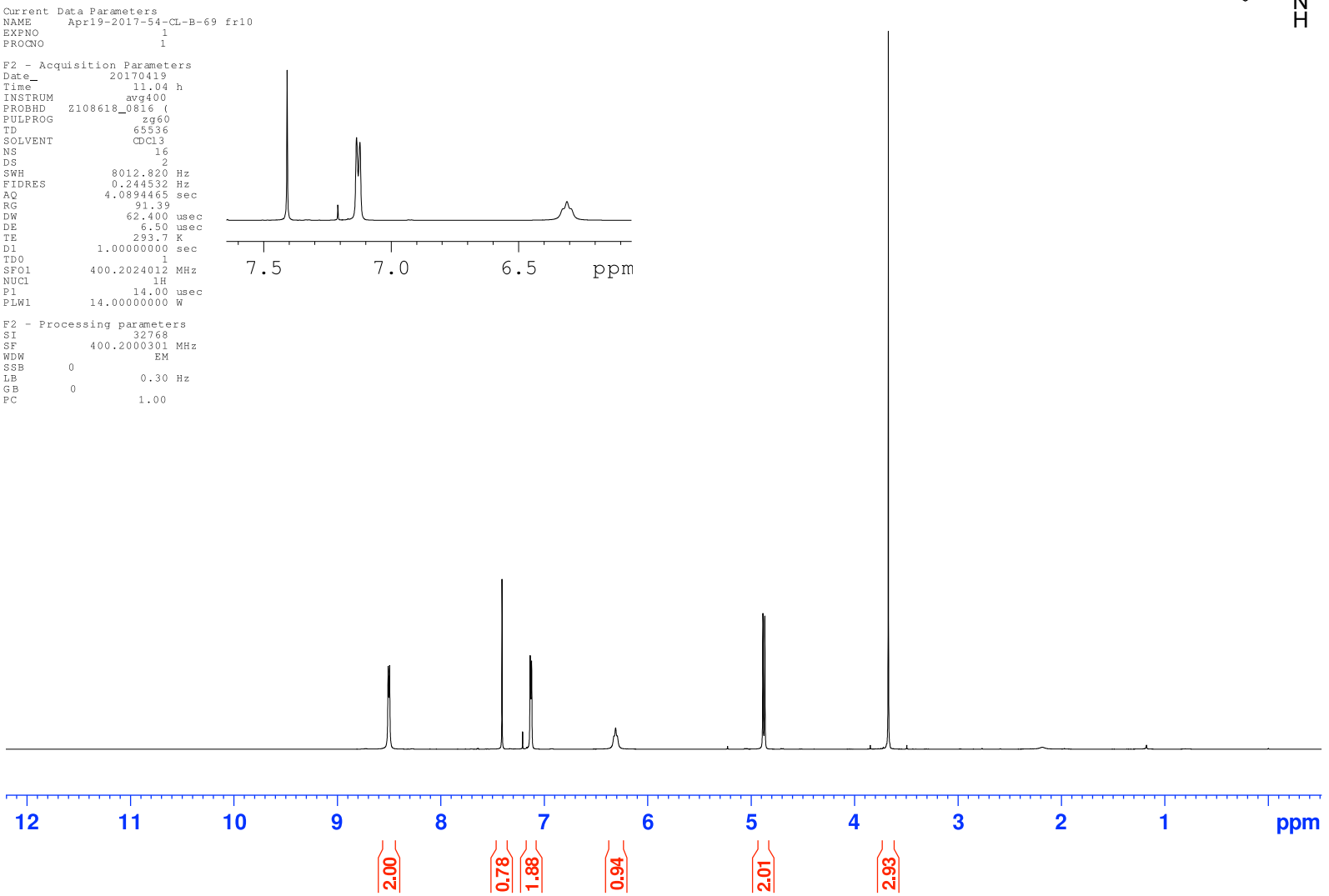
4-Chloro-2-methyl-5-(phenylamino)pyridazin-3(2*H*)-one (66)



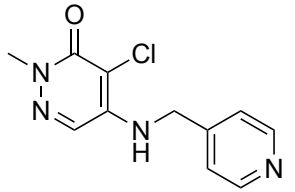


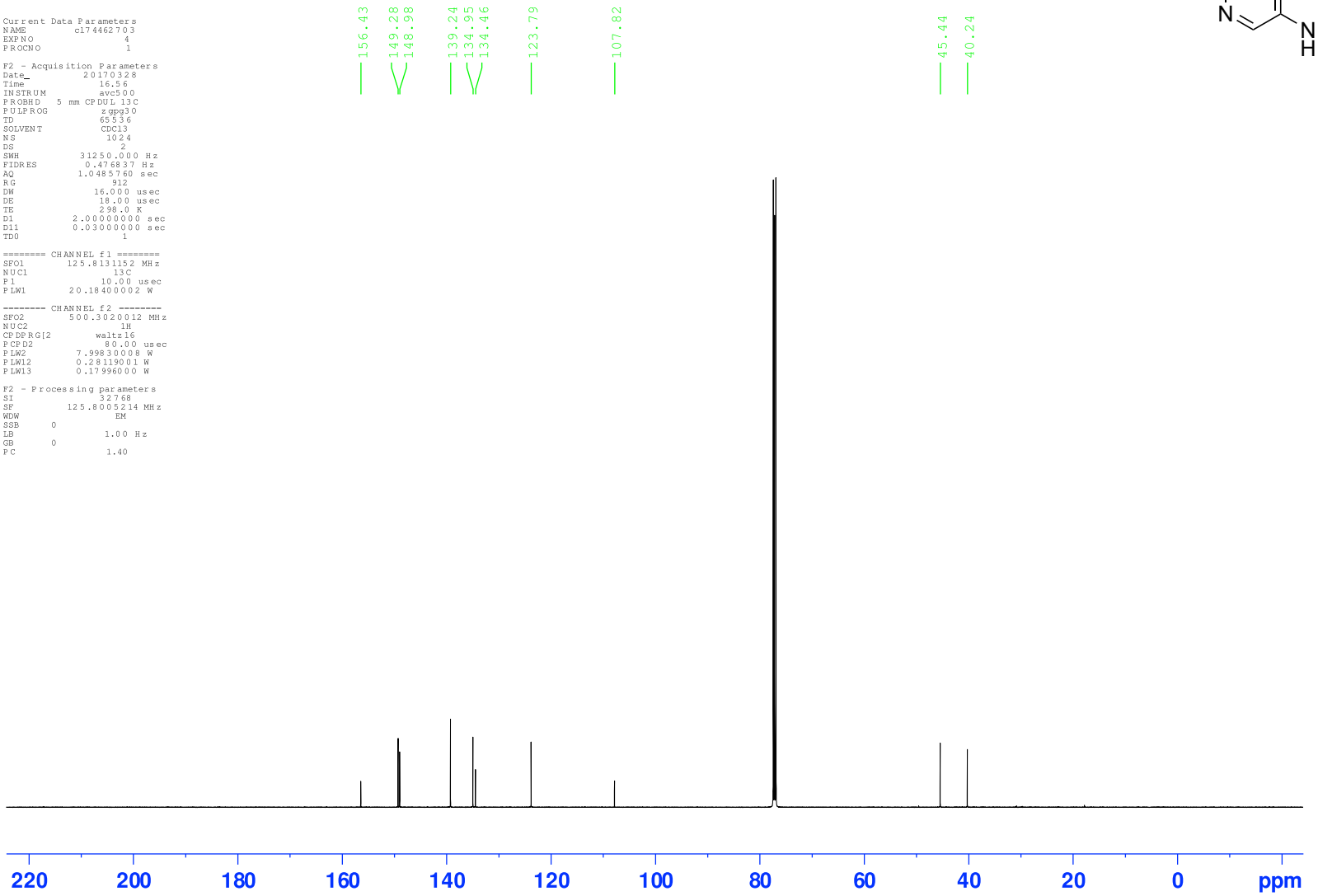
4-Chloro-2-methyl-5-((pyridin-4-ylmethyl)amino)pyridazin-3(2*H*)-one (78)



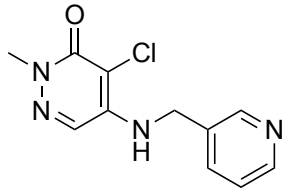


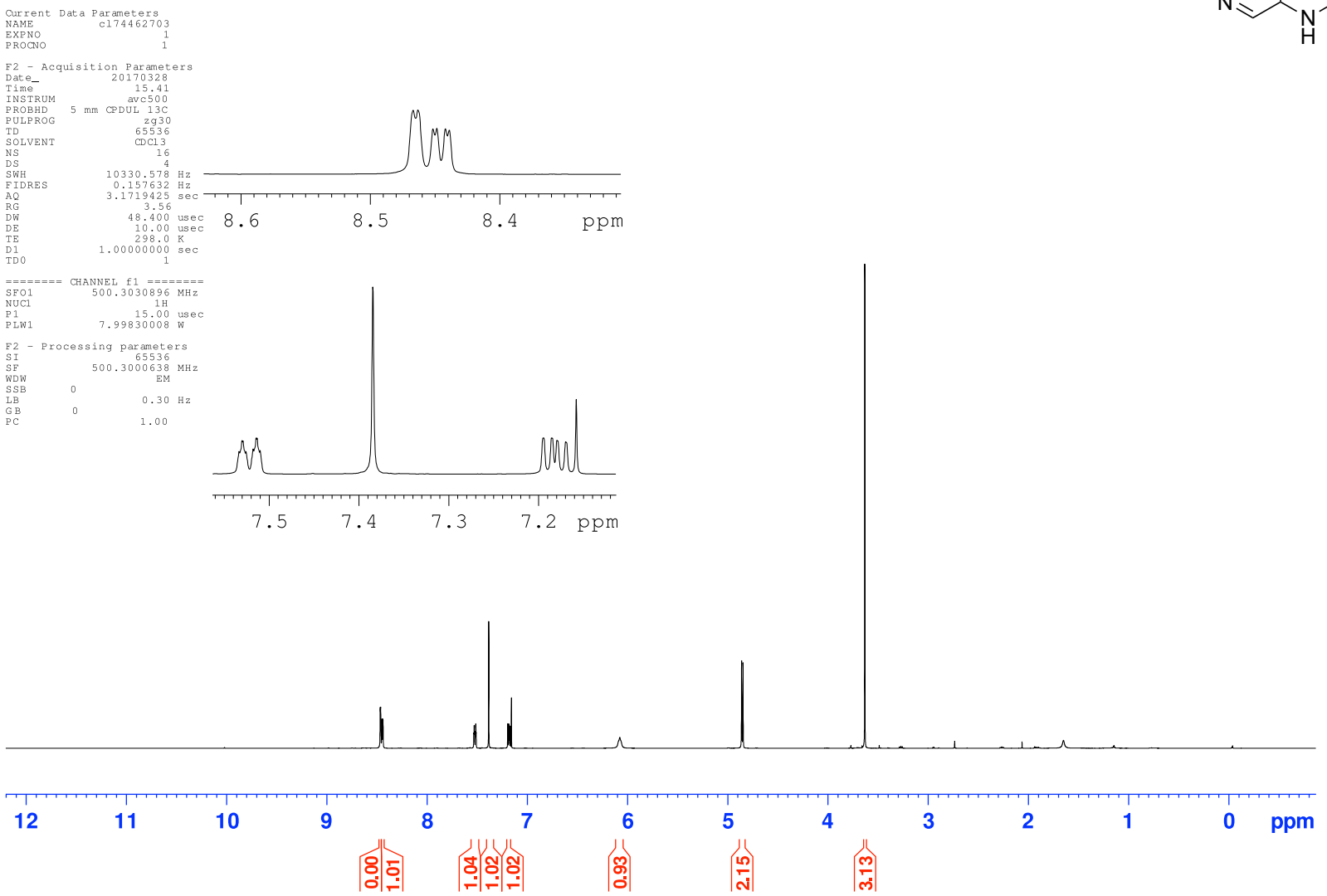
4-Chloro-2-methyl-5-((pyridin-4-ylmethyl)amino)pyridazin-3(2*H*)-one (78)



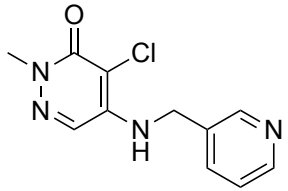


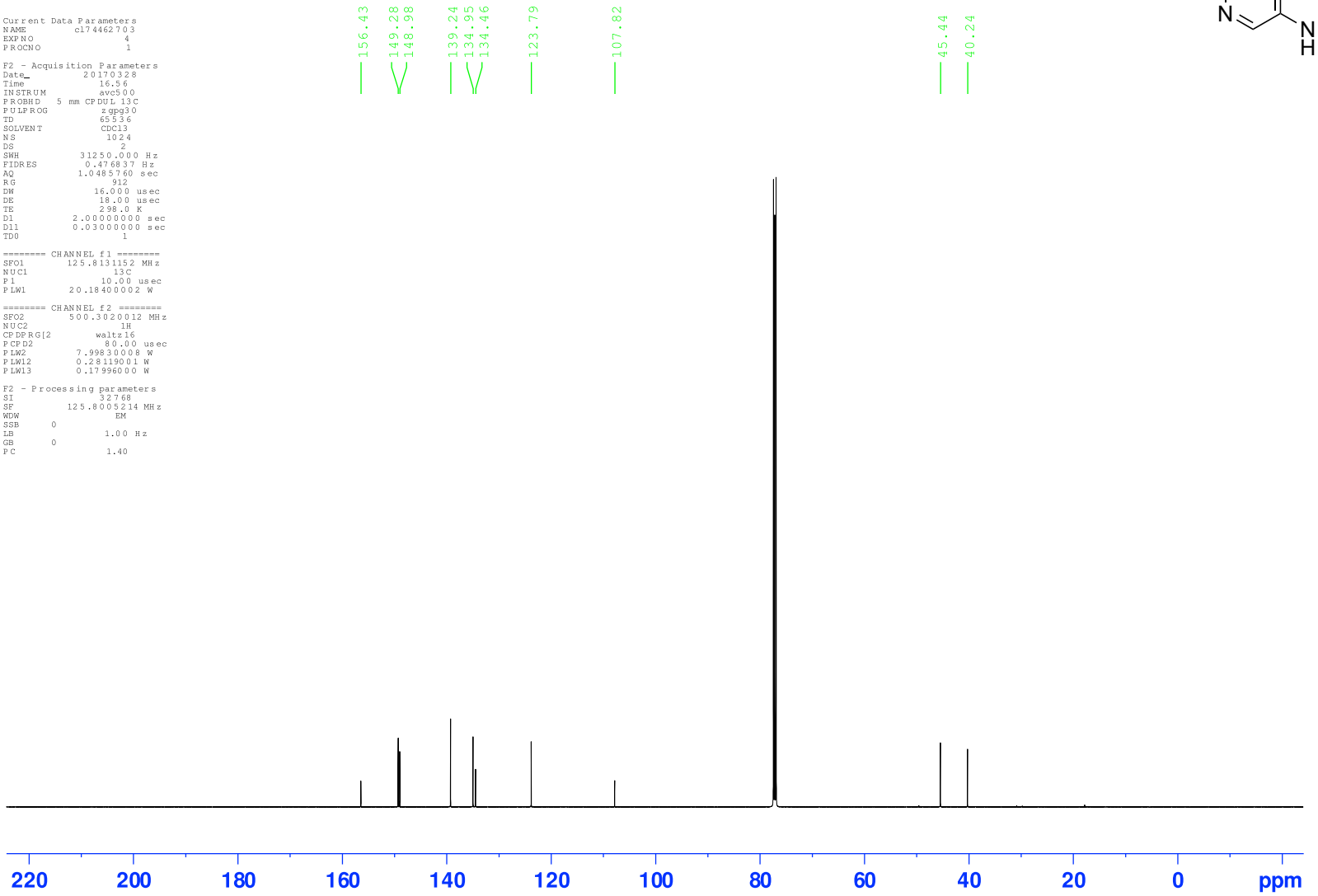
4-Chloro-2-methyl-5-((pyridin-3-ylmethyl)amino)pyridazin-3(2*H*)-one (77)



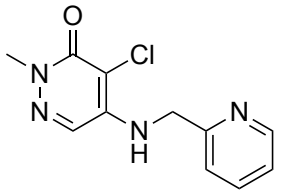


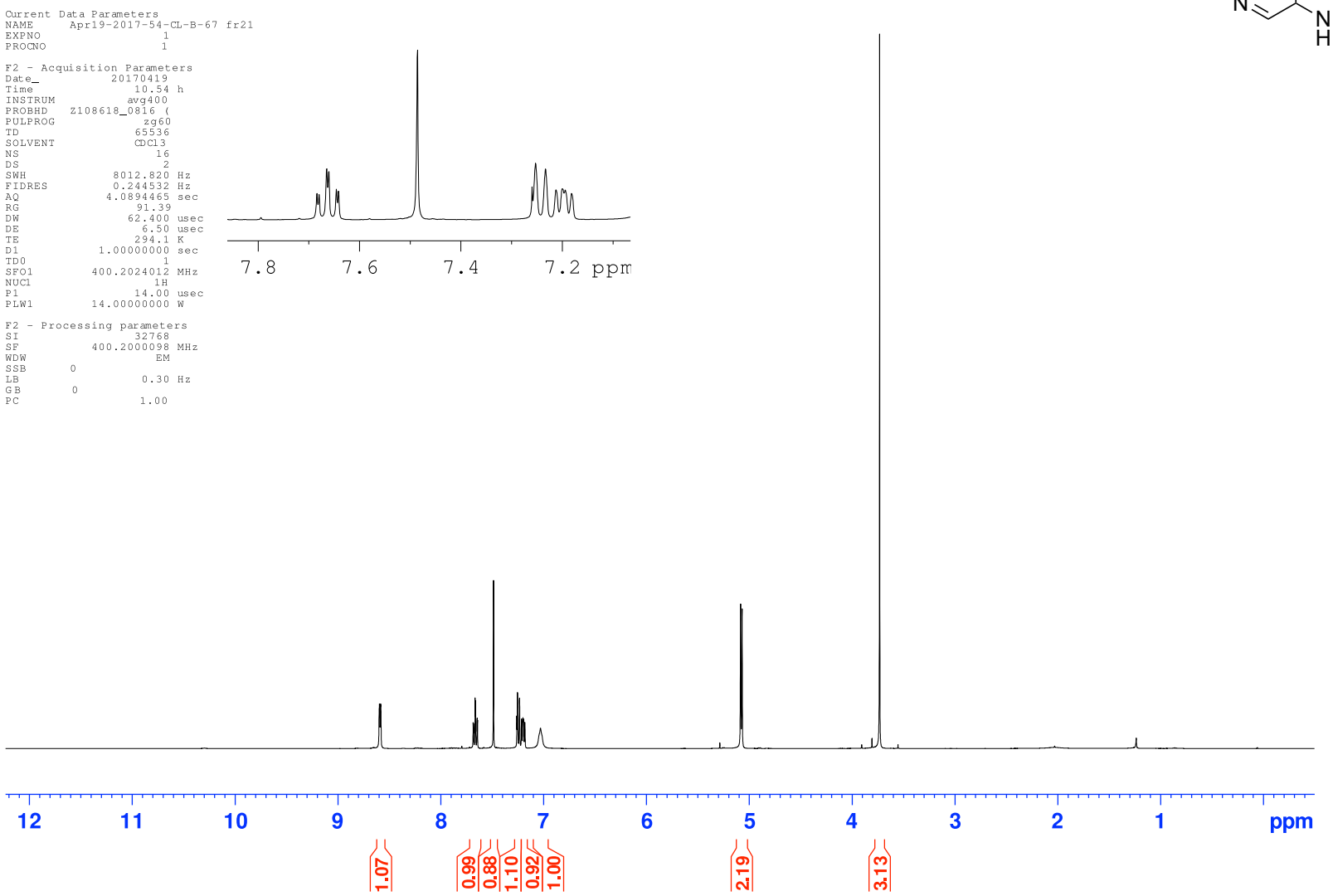
4-Chloro-2-methyl-5-((pyridin-3-ylmethyl)amino)pyridazin-3(2*H*)-one (77)



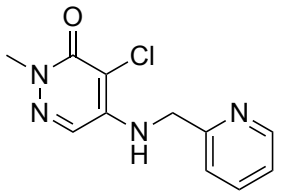


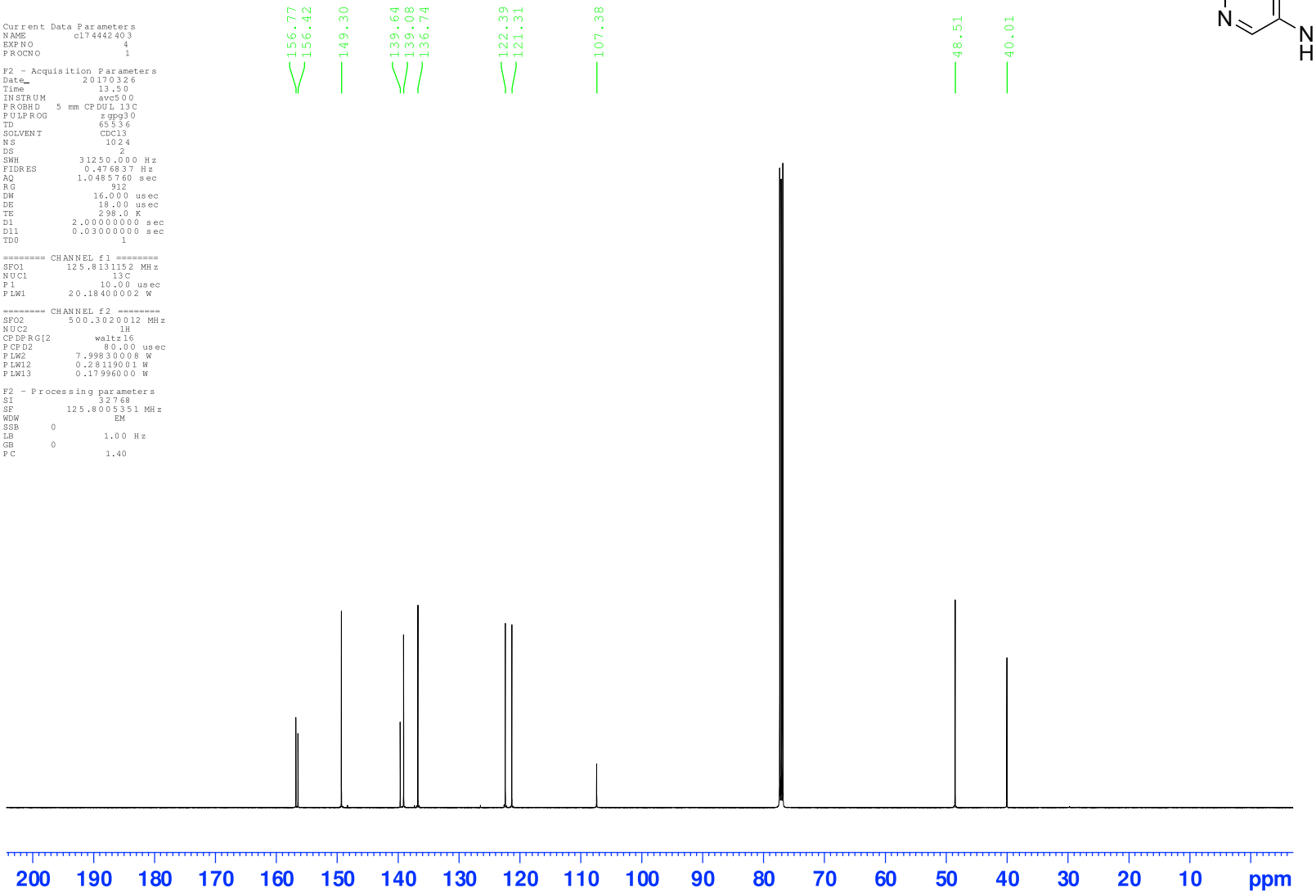
4-Chloro-2-methyl-5-((pyridin-2-ylmethyl)amino)pyridazin-3(2*H*)-one (76)





4-Chloro-2-methyl-5-((pyridin-2-ylmethyl)amino)pyridazin-3(2*H*)-one (76à)






```
Current Data Parameters
NAME          c174442403
EXPNO                  4
PROCNO                 1

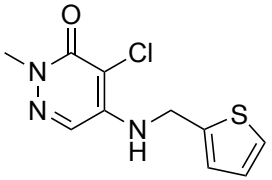
F2 - Acquisition Parameters
Date_          20170326
Time              13.50
INSTRUM          avc500
PROBHD   5 mm CPDUL 13C
PULPROG          zgpg30
TD               65536
SOLVENT          CDCl3
NS                1024
DS                   2
SWH           31250.000 Hz
FIDRES         0.476837 Hz
AQ            1.0485760 sec
RG                 912
DW               16.000 usec
DE               18.00 usec
TE               298.0 K
D1            2.00000000 sec
D11           0.03000000 sec
TD0                  1

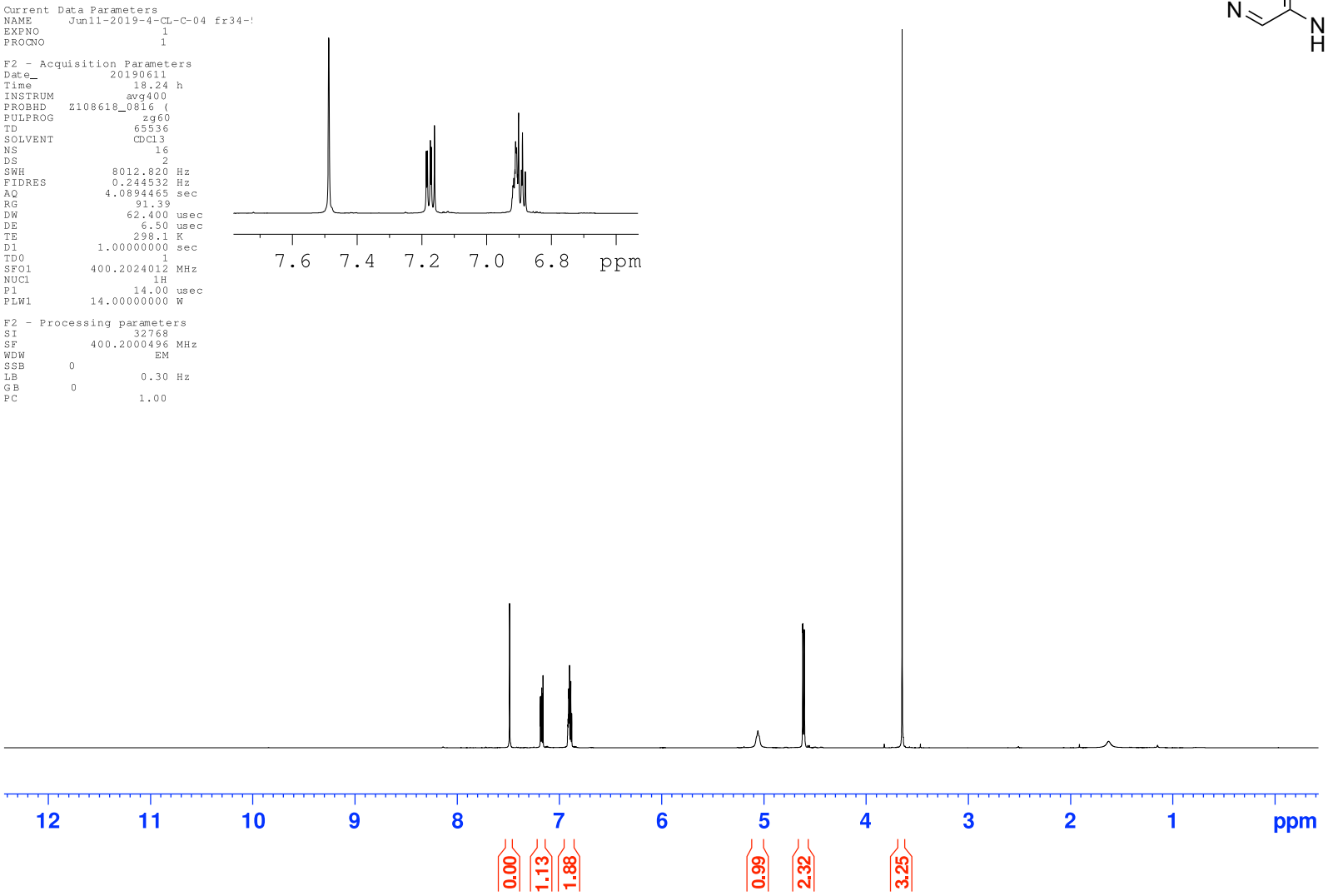
======== CHANNEL f1 ========
SFO1        125.8131152 MHz
NUC1               13C
P1               10.00 usec
PLW1        20.18400002 W

======== CHANNEL f2 ========
SFO2        500.3020012 MHz
NUC2               1H
CPDPRG[2         waltz16
PCPD2             80.00 usec
PLW2         7.99830008 W
PLW12        0.28119001 W
PLW13        0.17996000 W

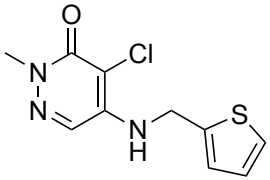
F2 - Processing parameters
SI                32768
SF          125.8005351 MHz
WDW                EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

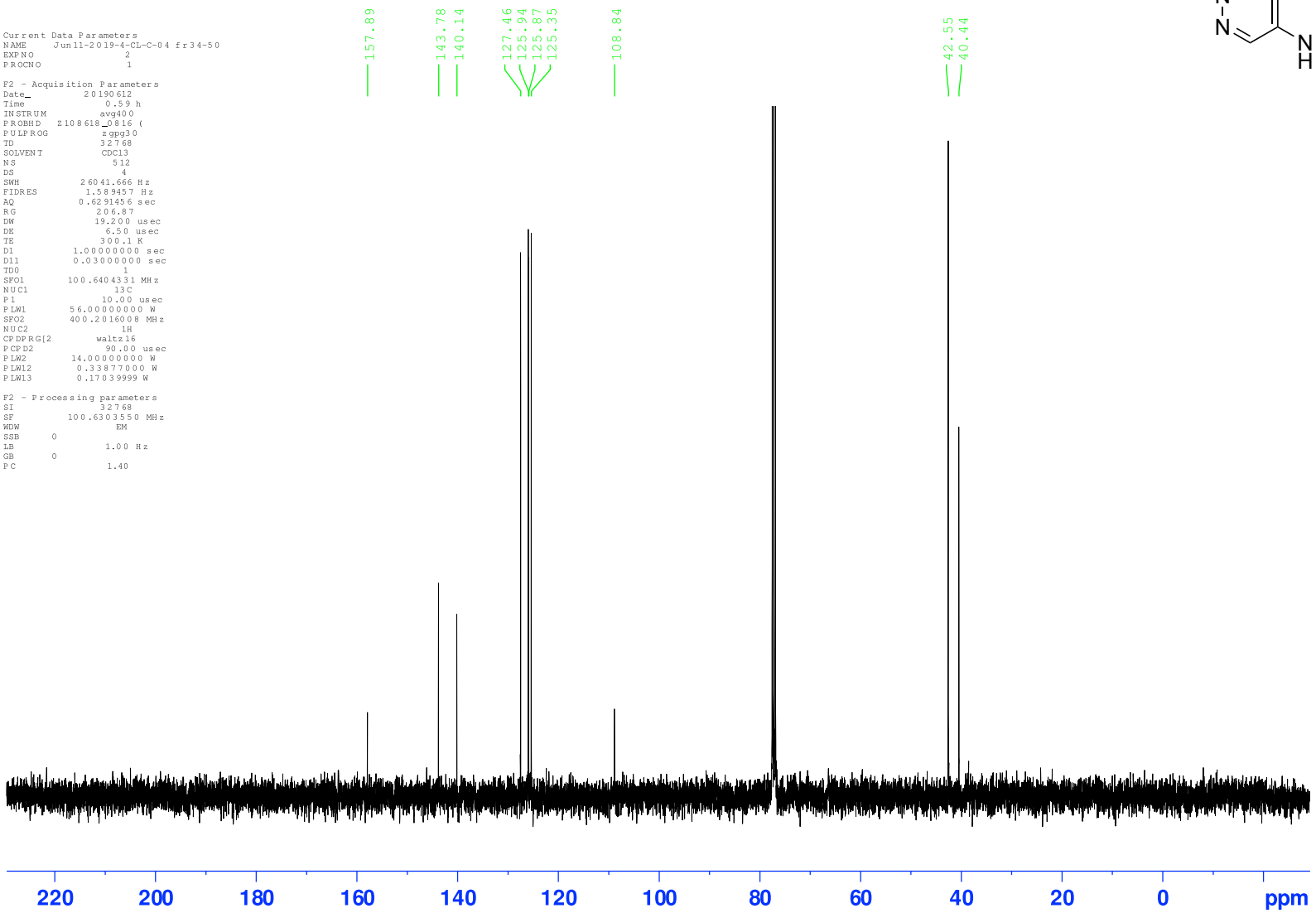
4-Chloro-2-methyl-5-((thiophen-2-ylmethyl)amino)pyridazin-3(2*H*)-one (70)



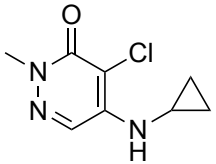


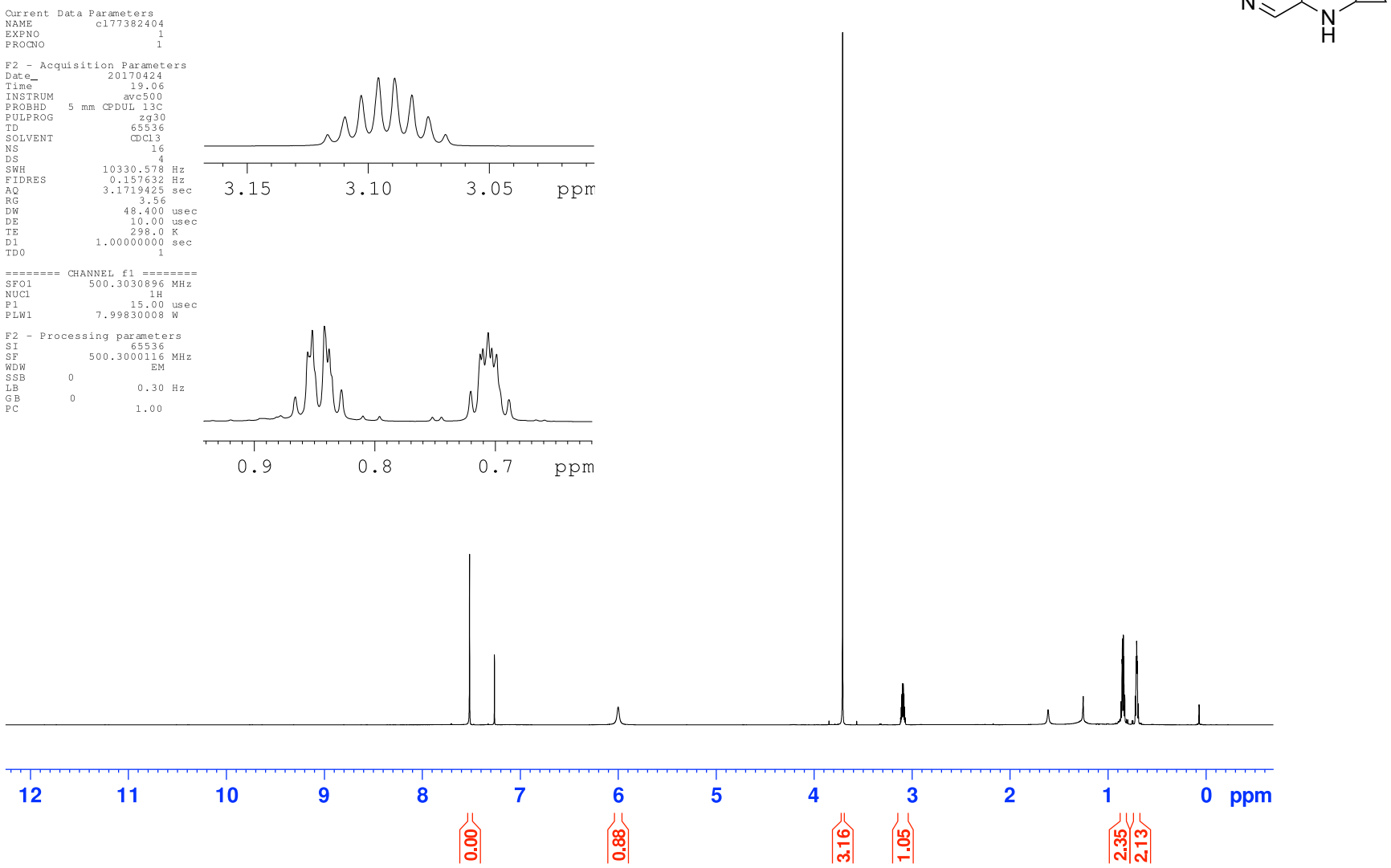
4-Chloro-2-methyl-5-((thiophen-2-ylmethyl)amino)pyridazin-3(2*H*)-one (70)



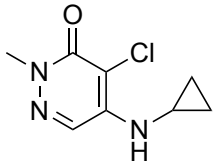


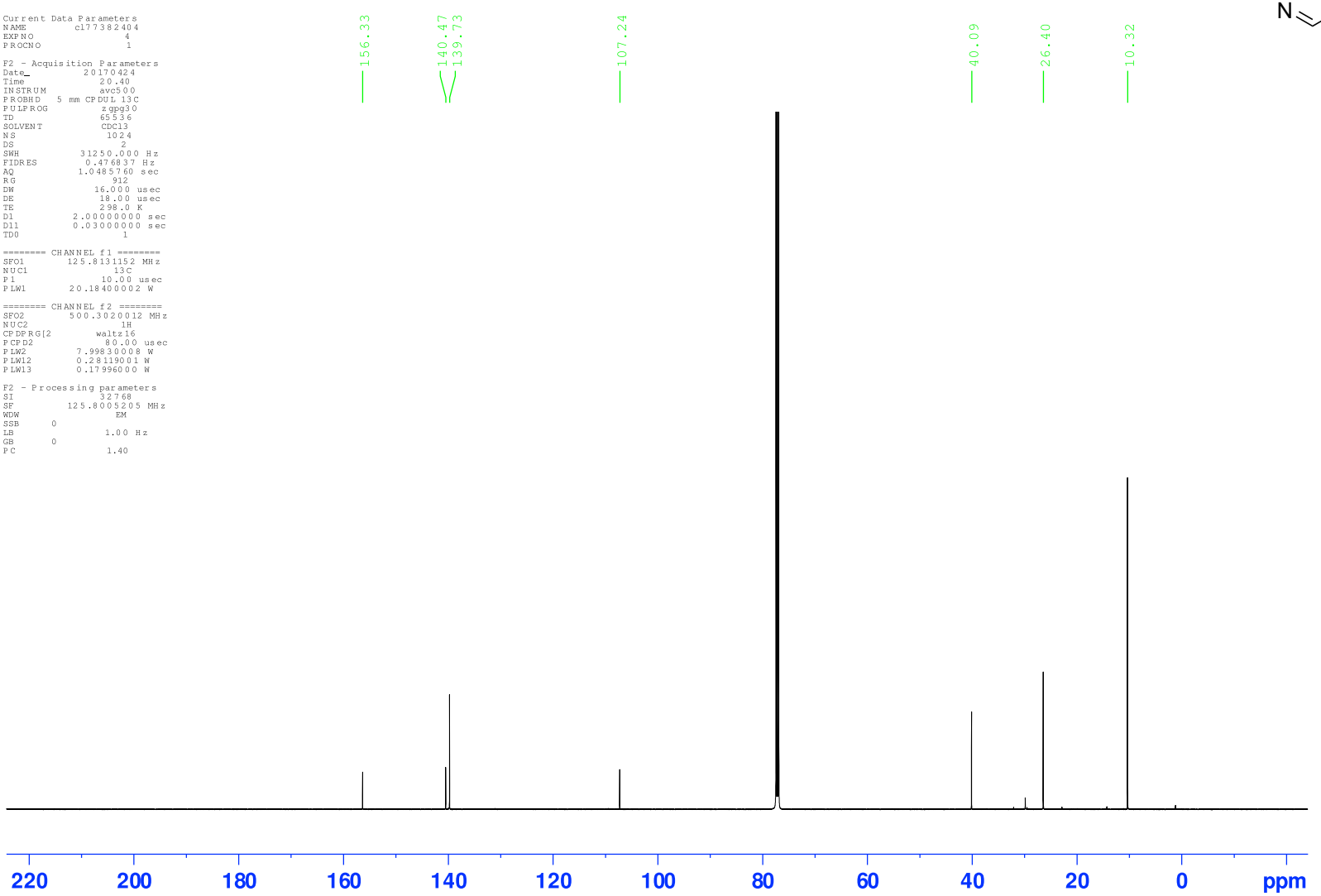
4-Chloro-5-(cyclopropylamino)-2-methylpyridazin-3(2*H*)-one (75)



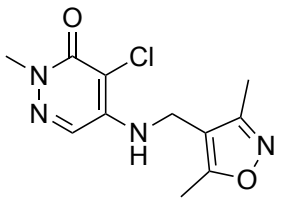


4-Chloro-5-(cyclopropylamino)-2-methylpyridazin-3(2*H*)-one (75)





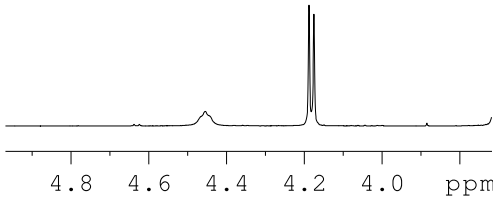
4-Chloro-5-(((3’,5’-dimethylisoxazol-4’-yl)methyl)amino)-2-methylpyridazin-3(2*H*)-one (79)

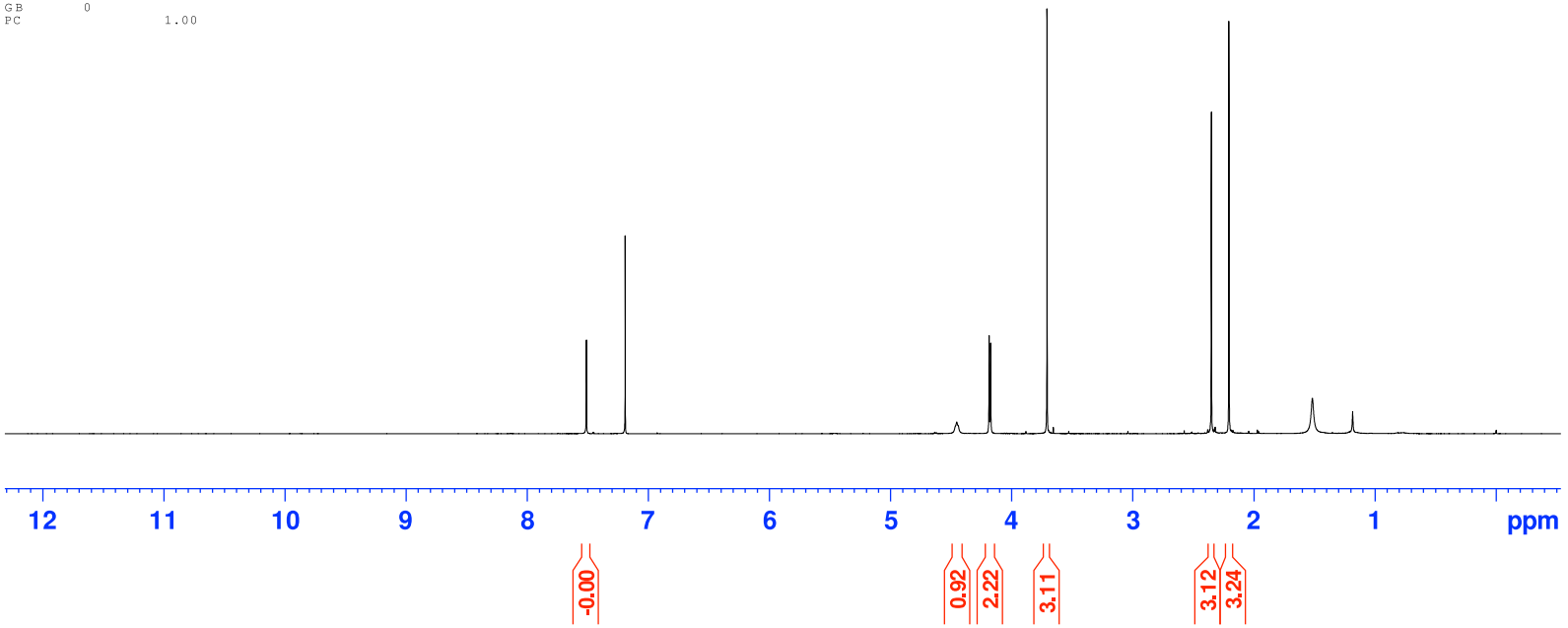



```
Current Data Parameters
NAME      Sep26-2019-13-CL-B-70 fr35
EXPNO                  1
PROCNO                 1

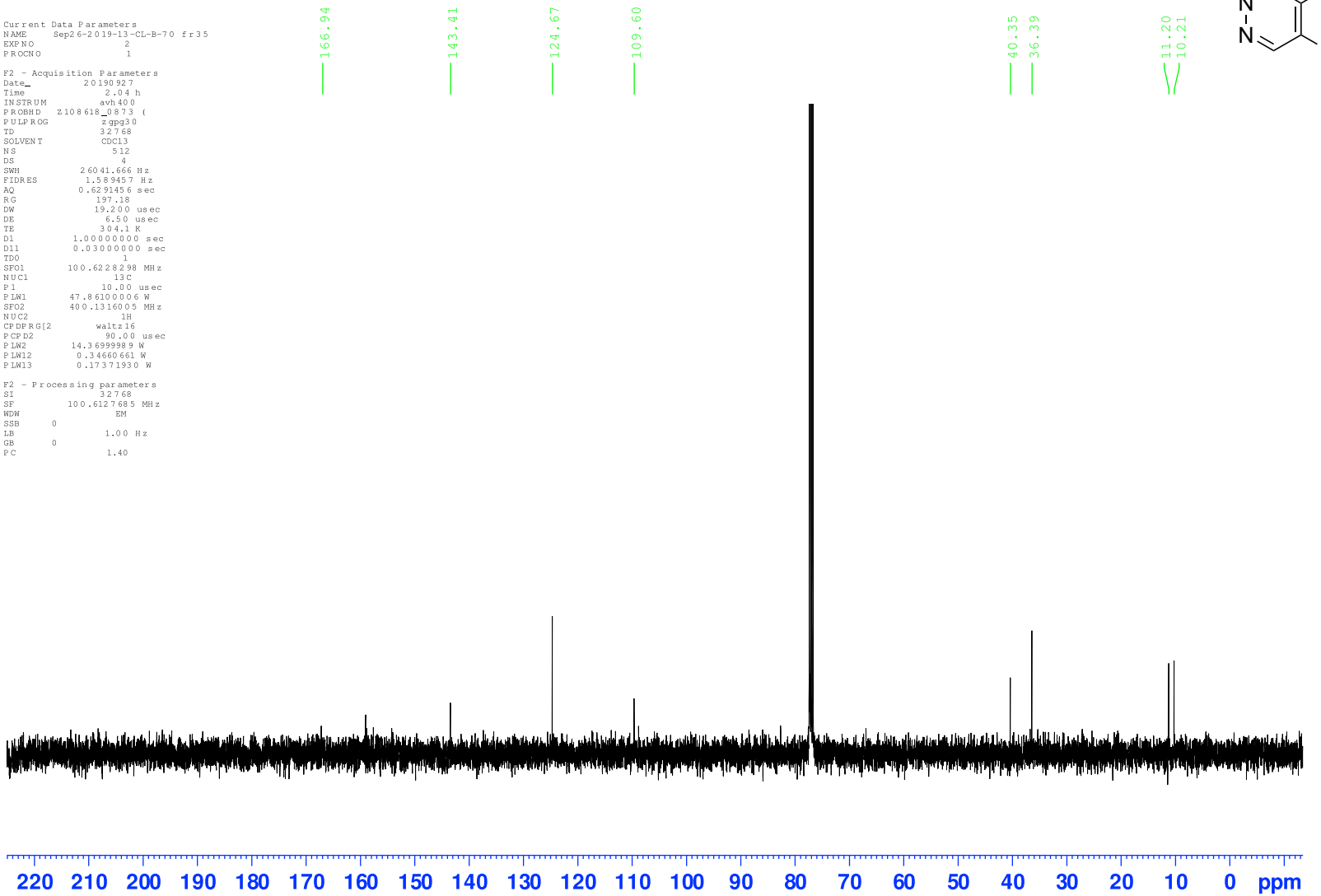
F2 - Acquisition Parameters
Date_           20190927
Time               1.48 h
INSTRUM          avh400
PROBHD    Z108618_0873 (
PULPROG            zg60
TD                65536
SOLVENT           CDCl3
NS                   16
DS                    2
SWH            8012.820 Hz
FIDRES         0.244532 Hz
AQ             4.0894465 sec
RG                88.17
DW                62.400 usec
DE                 6.50 usec
TE                303.2 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           400.1300371 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

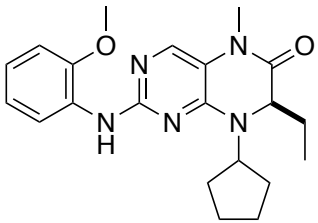


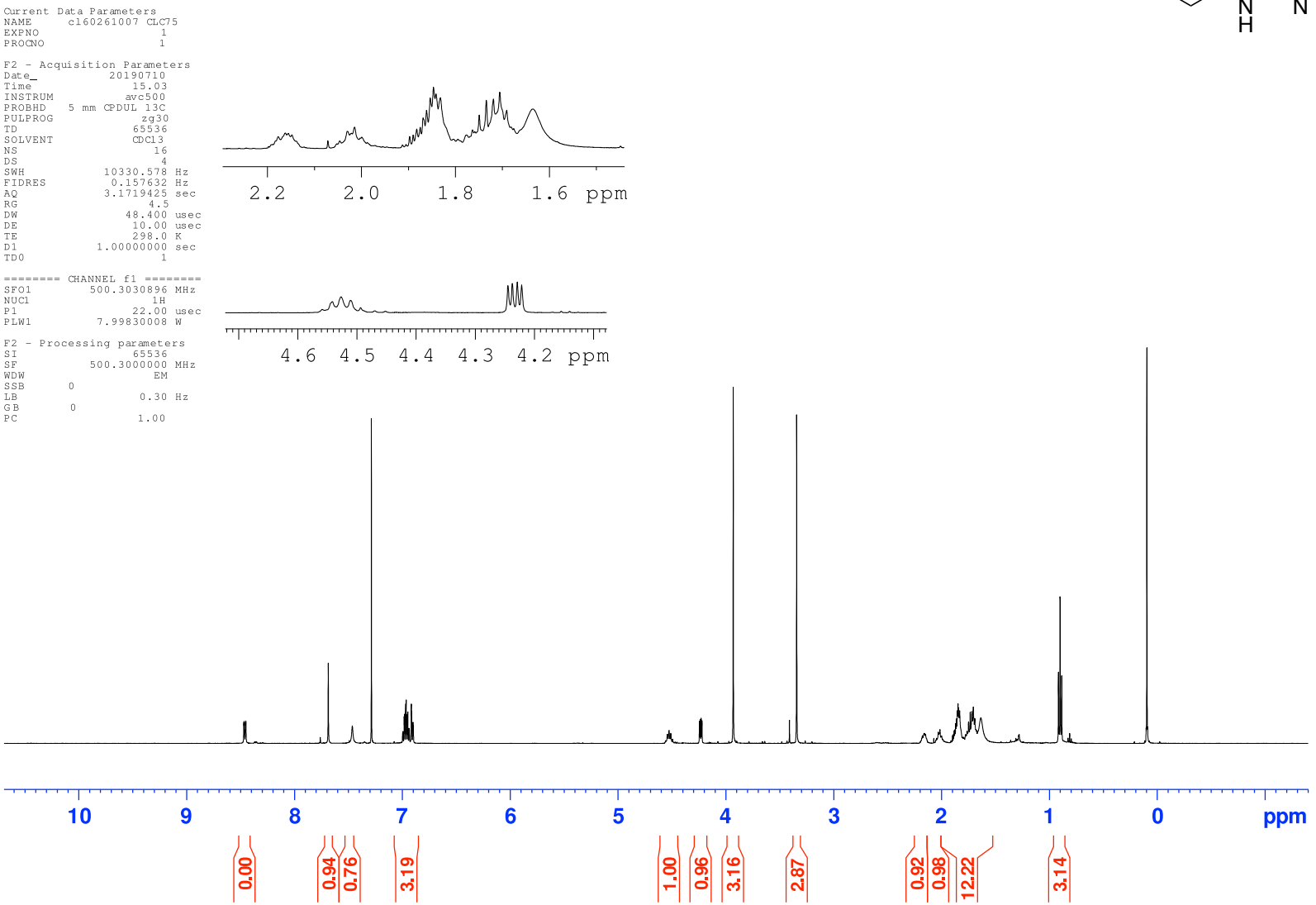


4-Chloro-5-(((3’,5’-dimethylisoxazol-4’-yl)methyl)amino)-2-methylpyridazin-3(2*H*)-one (79)

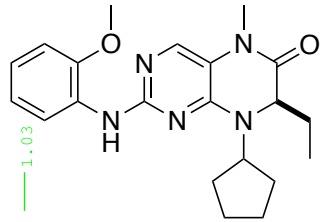


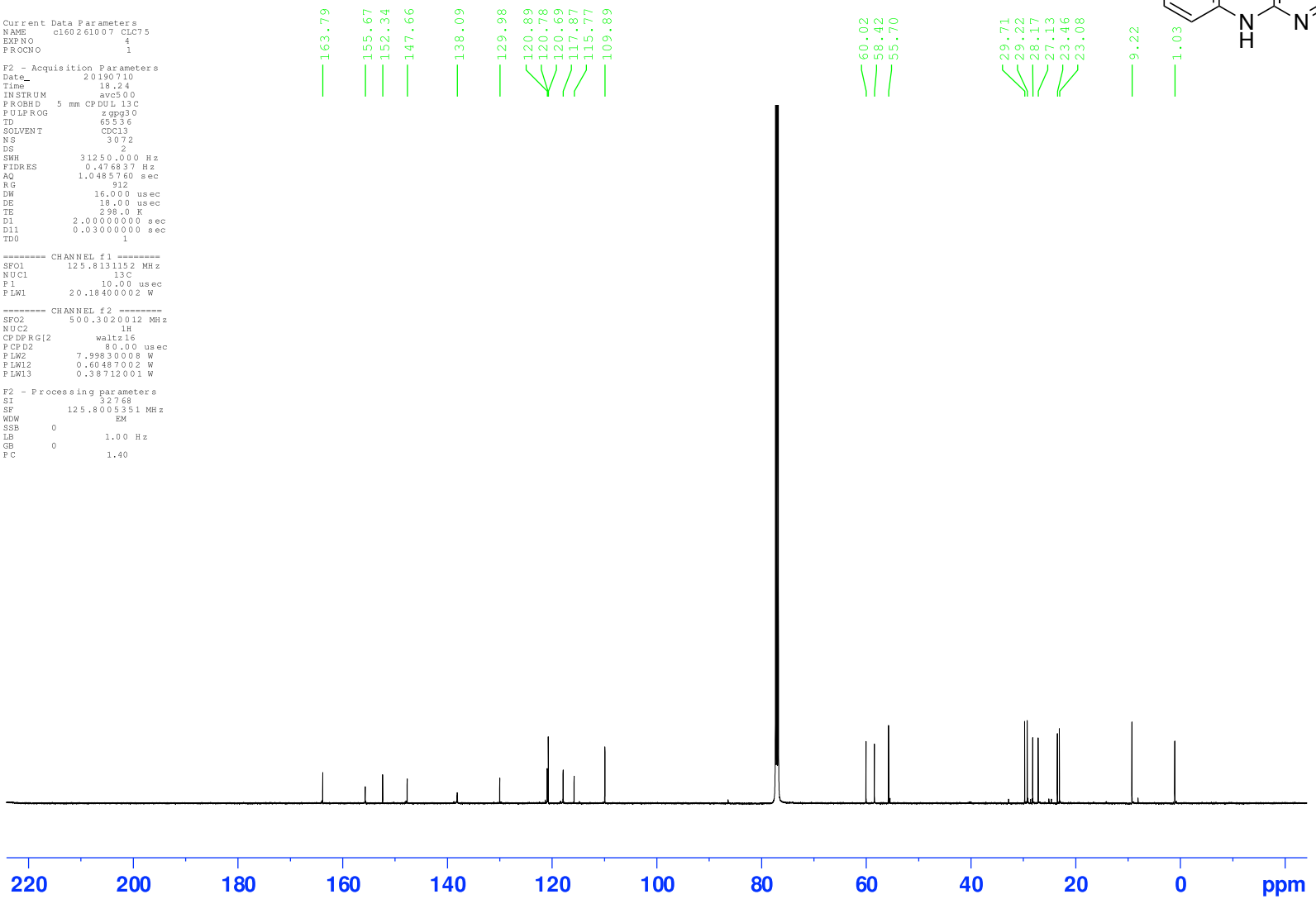
(*R*)-8-Cyclopentyl-7-ethyl-2-((2'-methoxyphenyl)amino)-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-156)



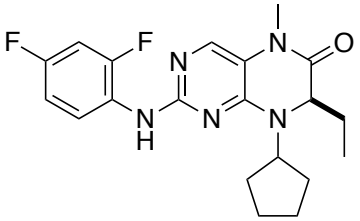


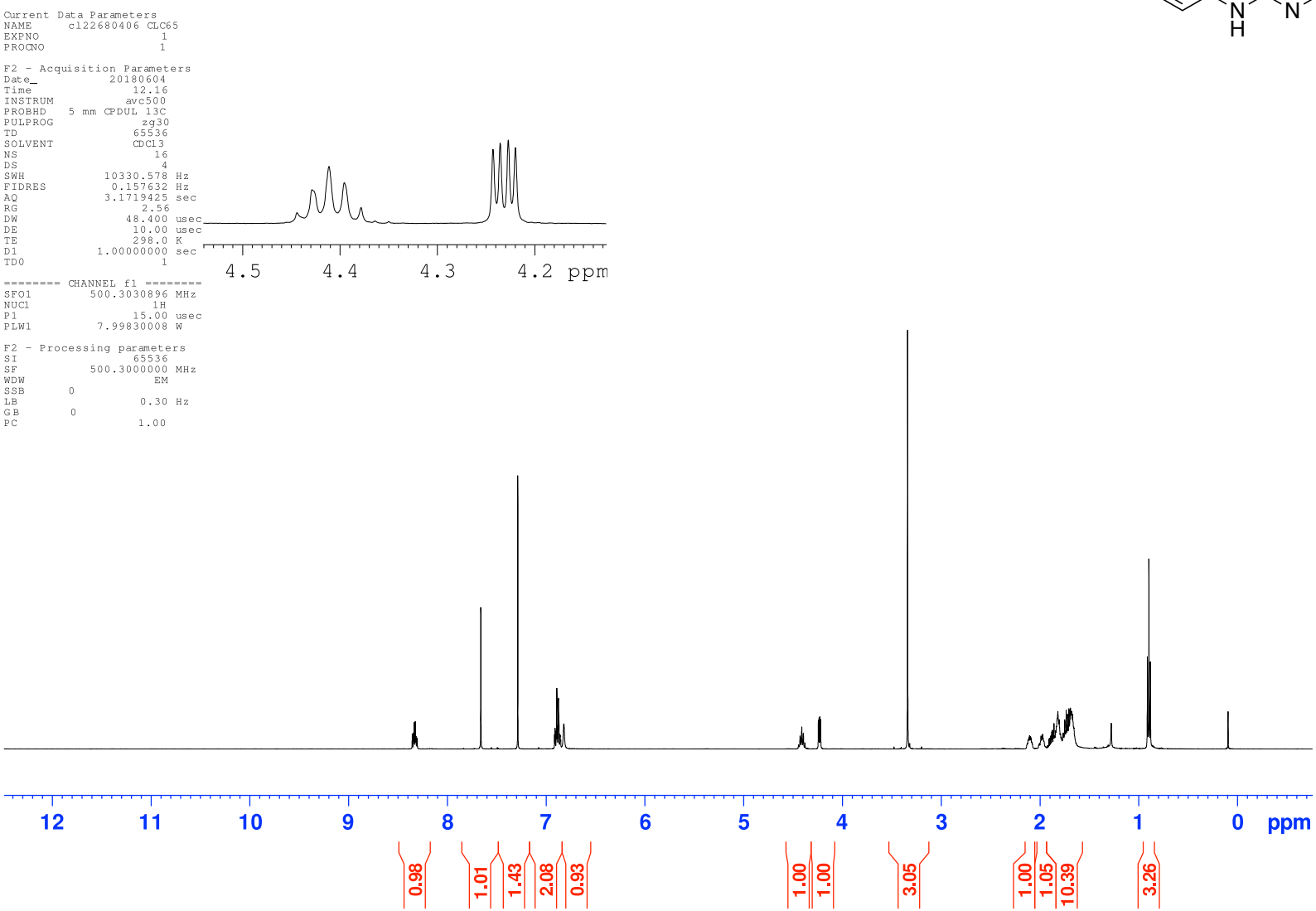
(*R*)-8-Cyclopentyl-7-ethyl-2-((2'-methoxyphenyl)amino)-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-156)



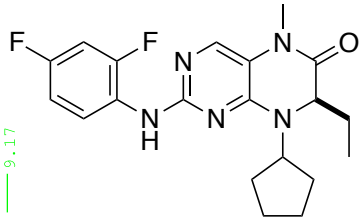


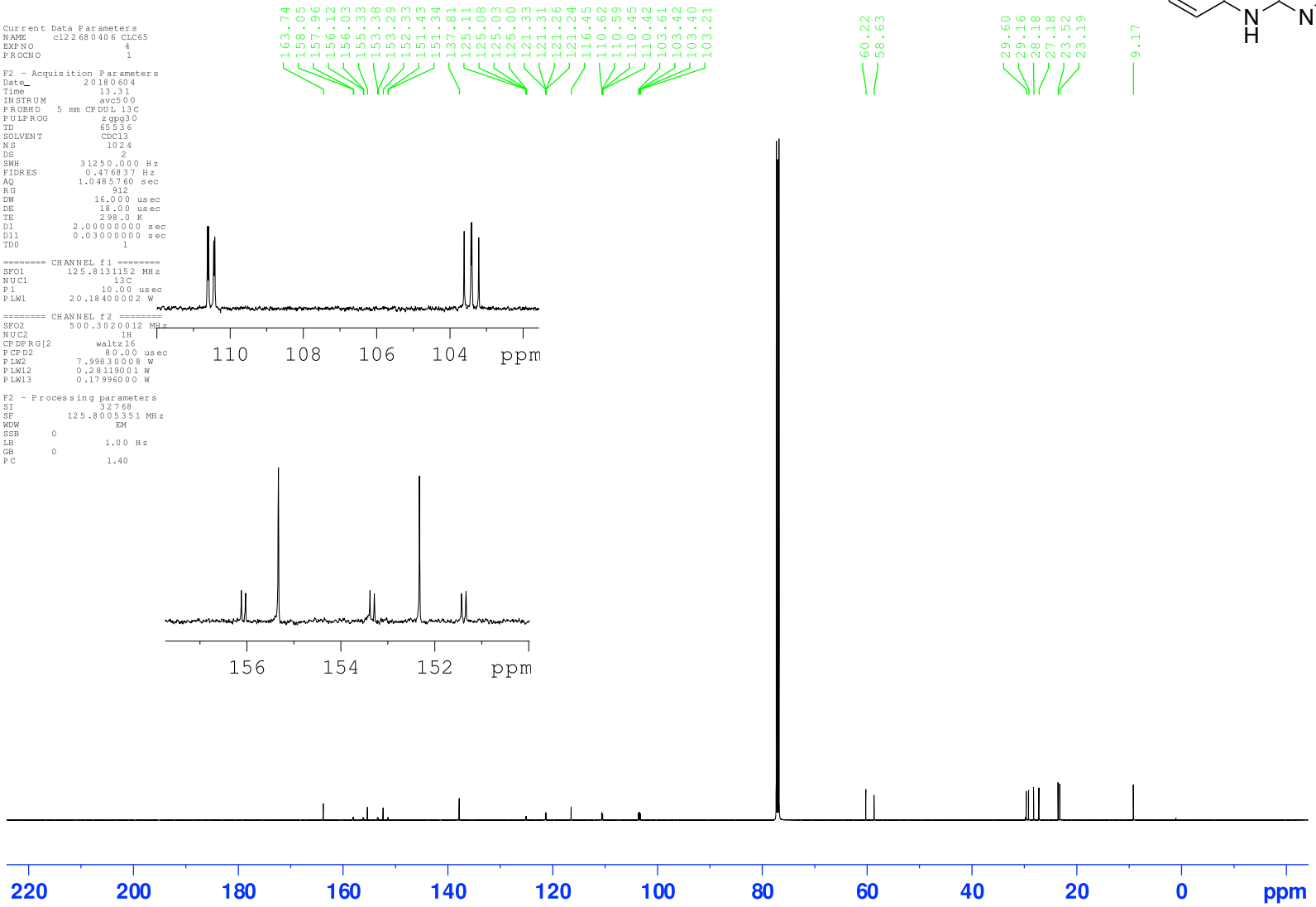
(7*R*)-8-Cyclopentyl-2-((2',4'-difluorophenyl)amino)-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-130)



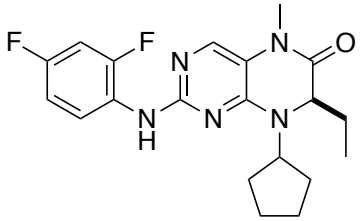


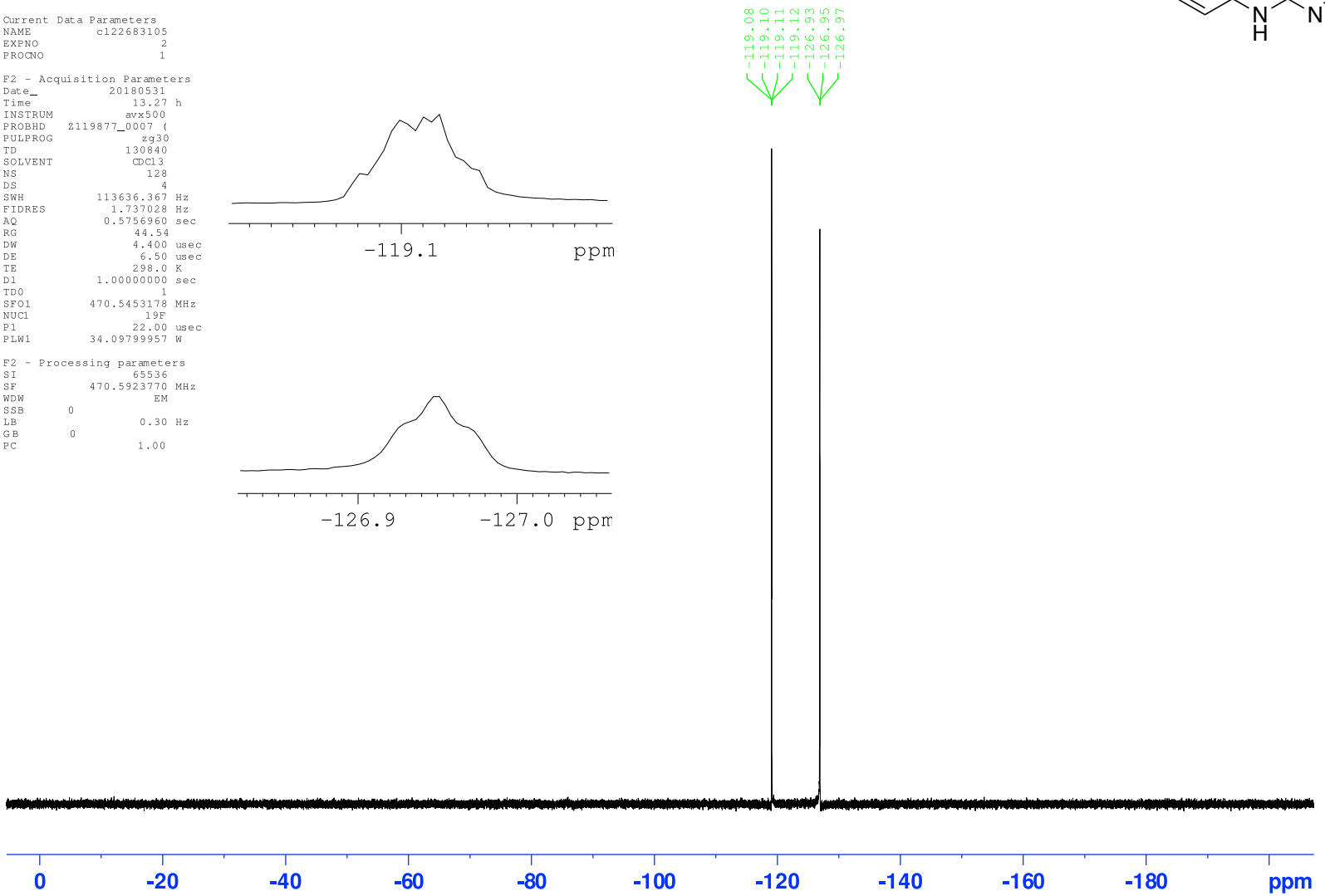
(7*R*)-8-Cyclopentyl-2-((2',4'-difluorophenyl)amino)-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-130)



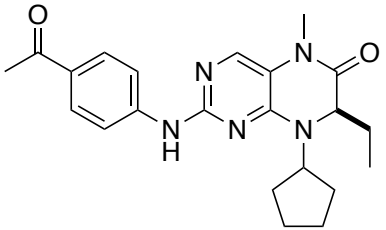


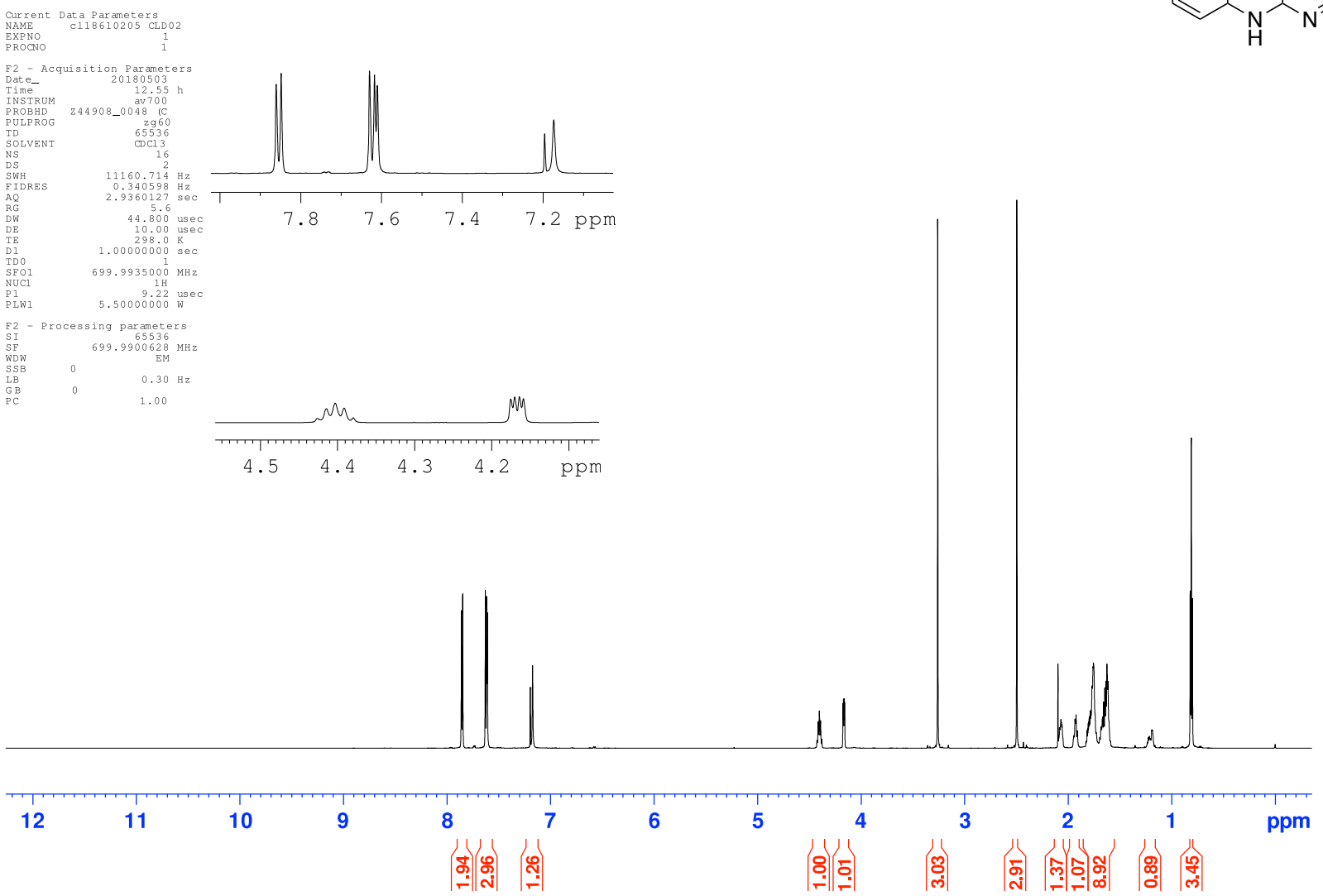
(7*R*)-8-Cyclopentyl-2-((2',4'-difluorophenyl)amino)-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-130)



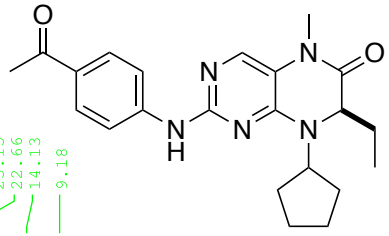


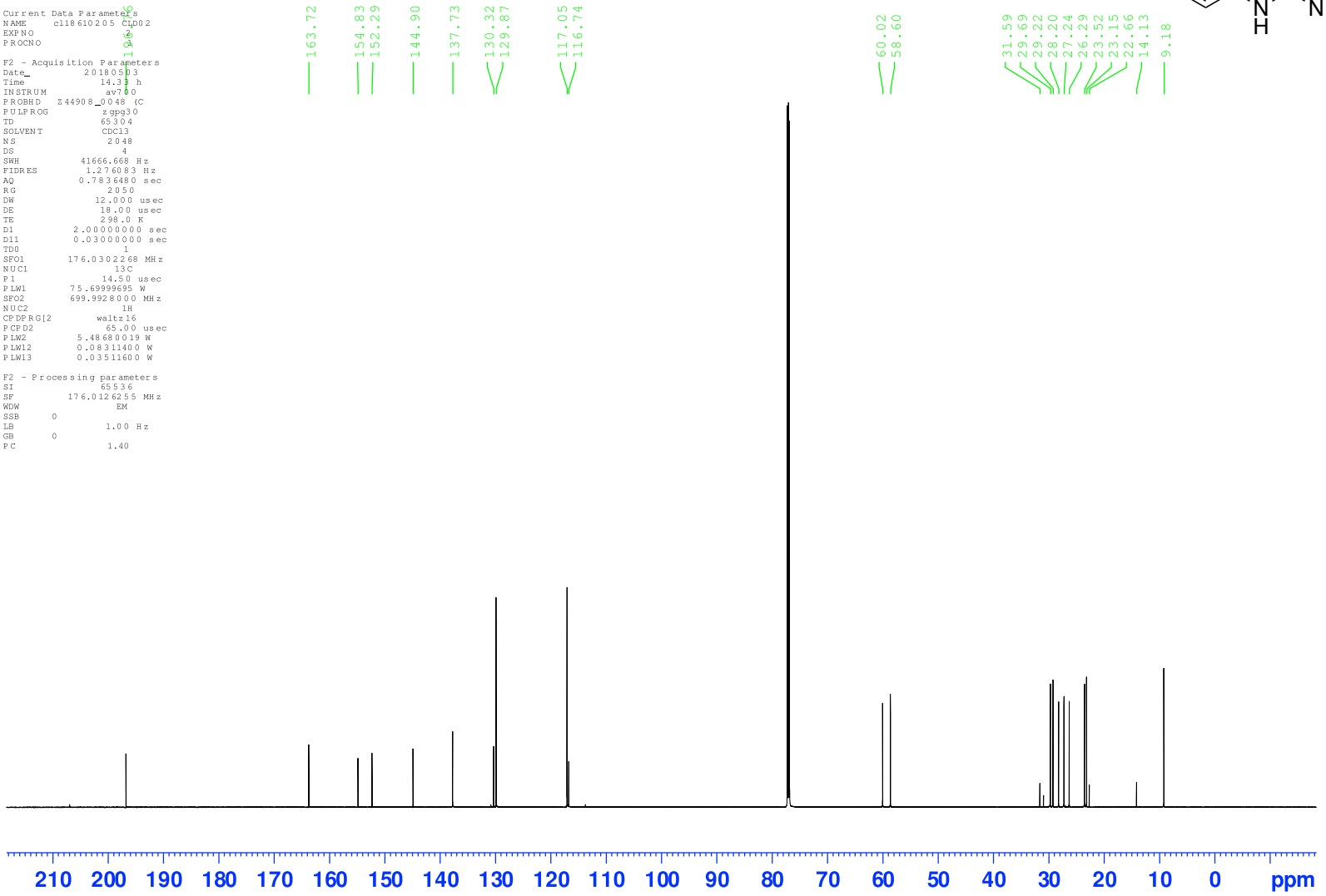
(*7R*)-2-((4'-Acetylphenyl)amino)-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-157)



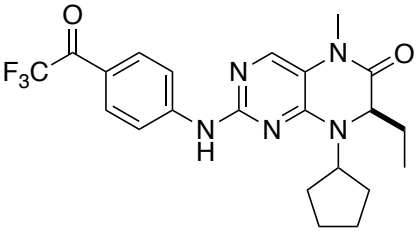


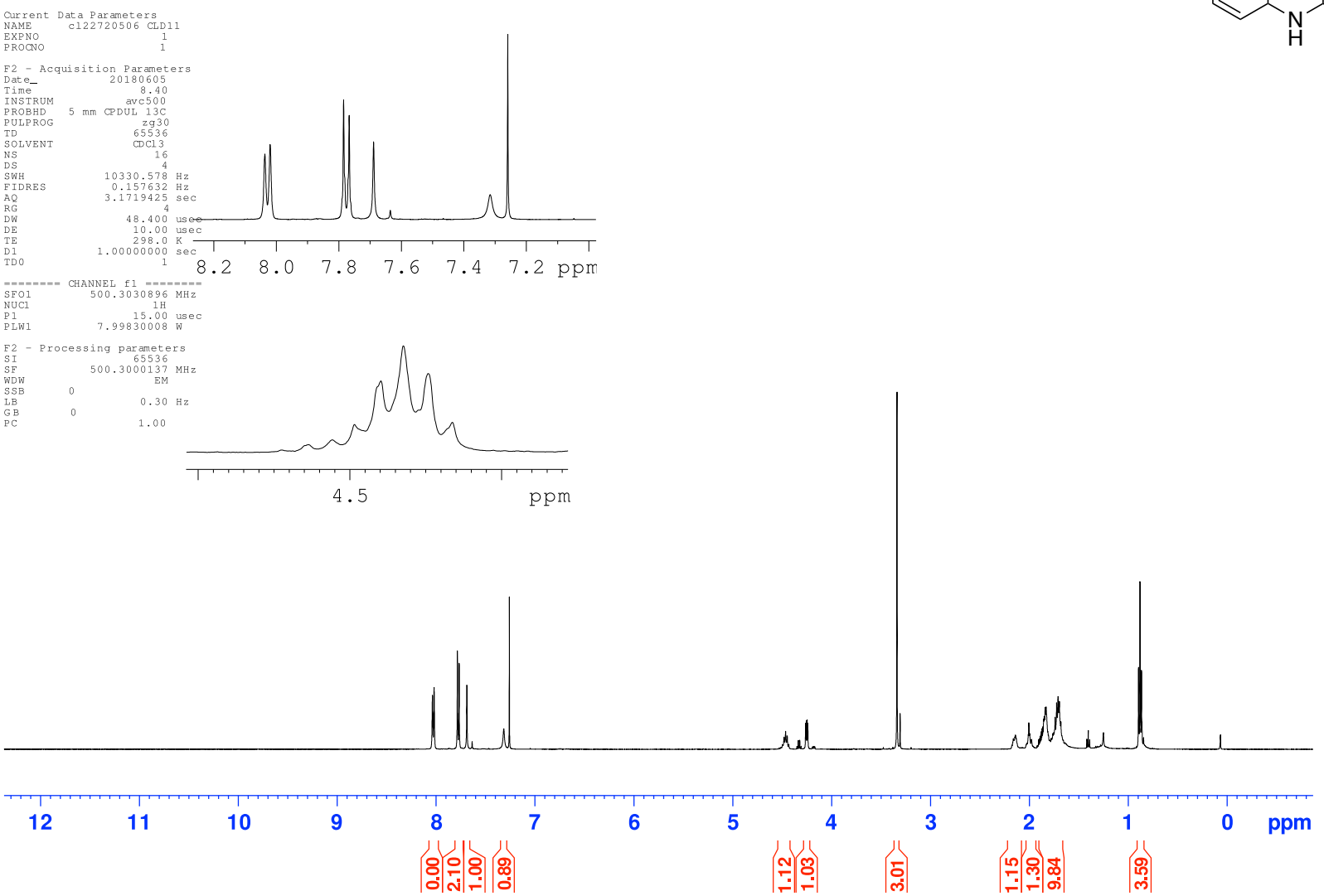
(*7R*)-2-((4'-Acetylphenyl)amino)-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one ((*R*)-157)



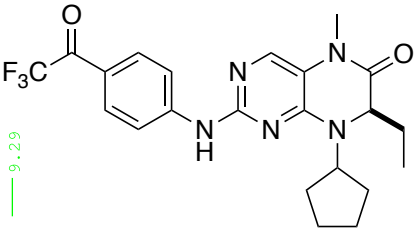


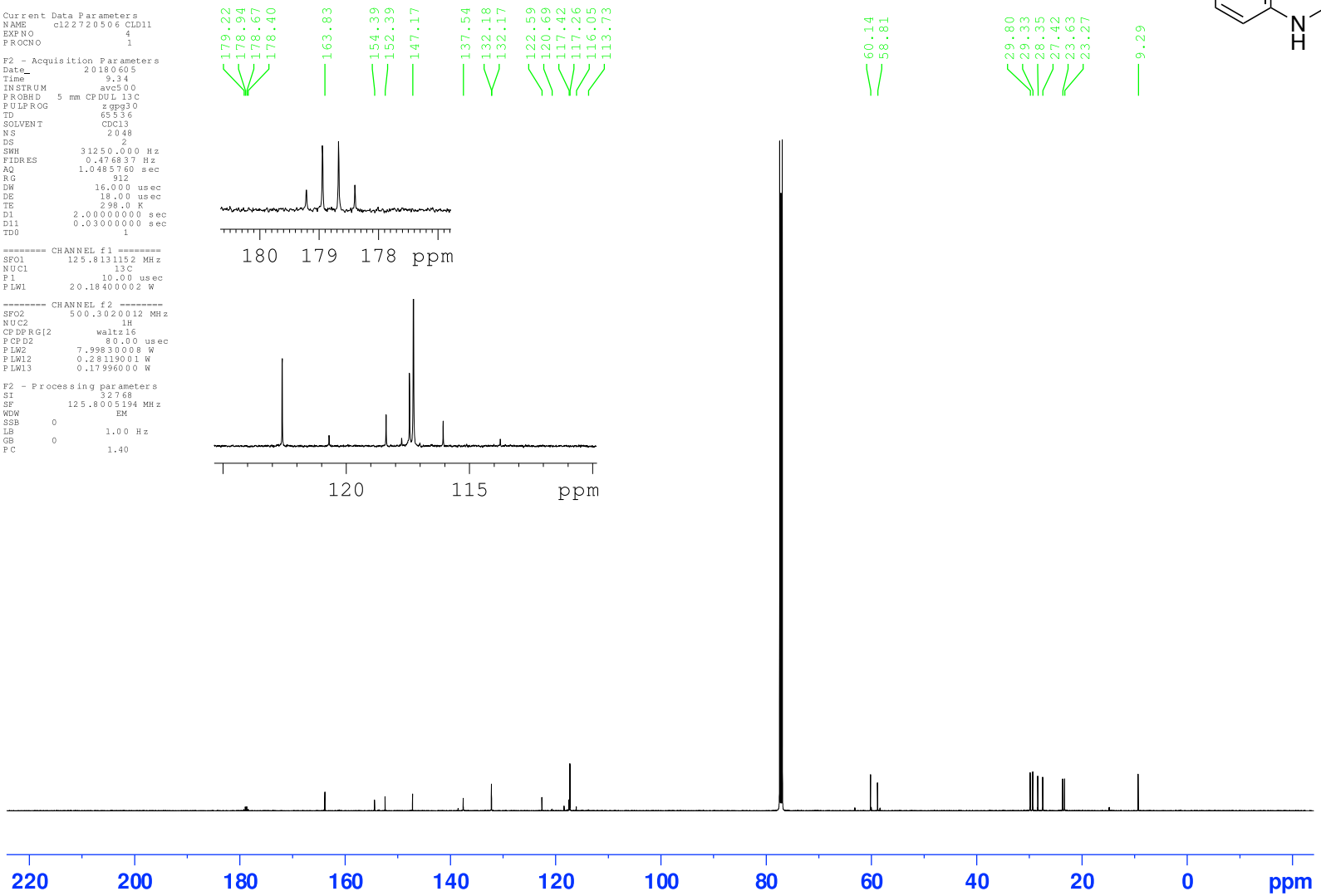
(*7R*)-8-Cyclopentyl-7-ethyl-5-methyl-2-((4'-(2,2,2-trifluoroacetyl)phenyl)amino)-7,8-dihydropteridin-6(5*H*)-one ((*R*)-131)



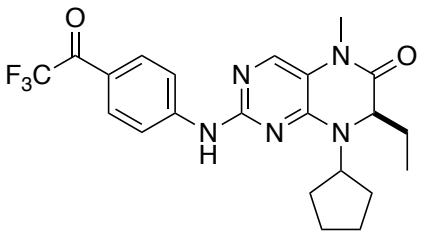


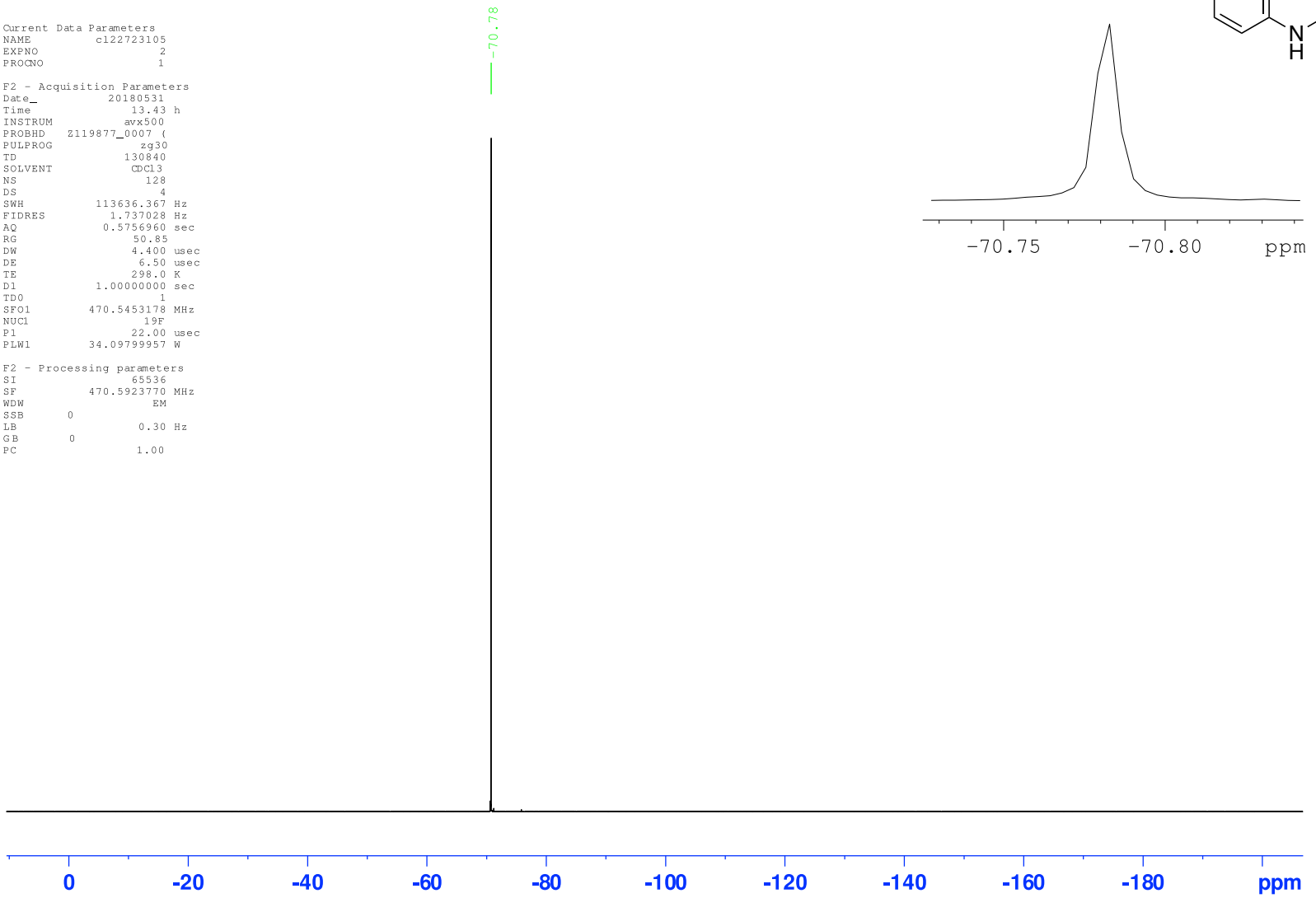
(*7R*)-8-Cyclopentyl-7-ethyl-5-methyl-2-((4'-(2,2,2-trifluoroacetyl)phenyl)amino)-7,8-dihydropteridin-6(5*H*)-one ((*R*)-131)



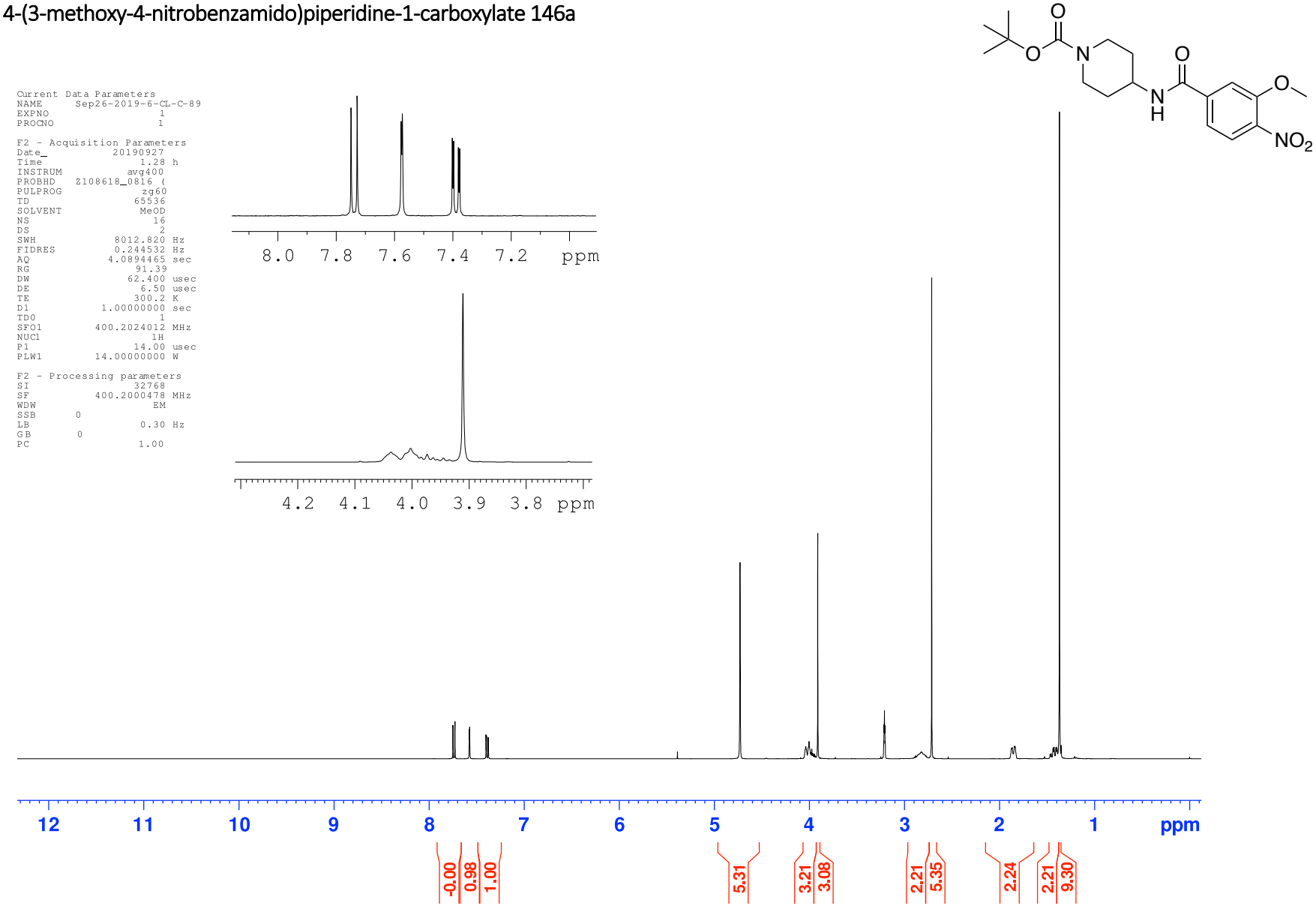


(*7R*)-8-Cyclopentyl-7-ethyl-5-methyl-2-((4'-(2,2,2-trifluoroacetyl)phenyl)amino)-7,8-dihydropteridin-6(5*H*)-one ((*R*)-131)

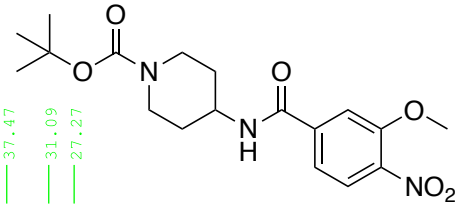


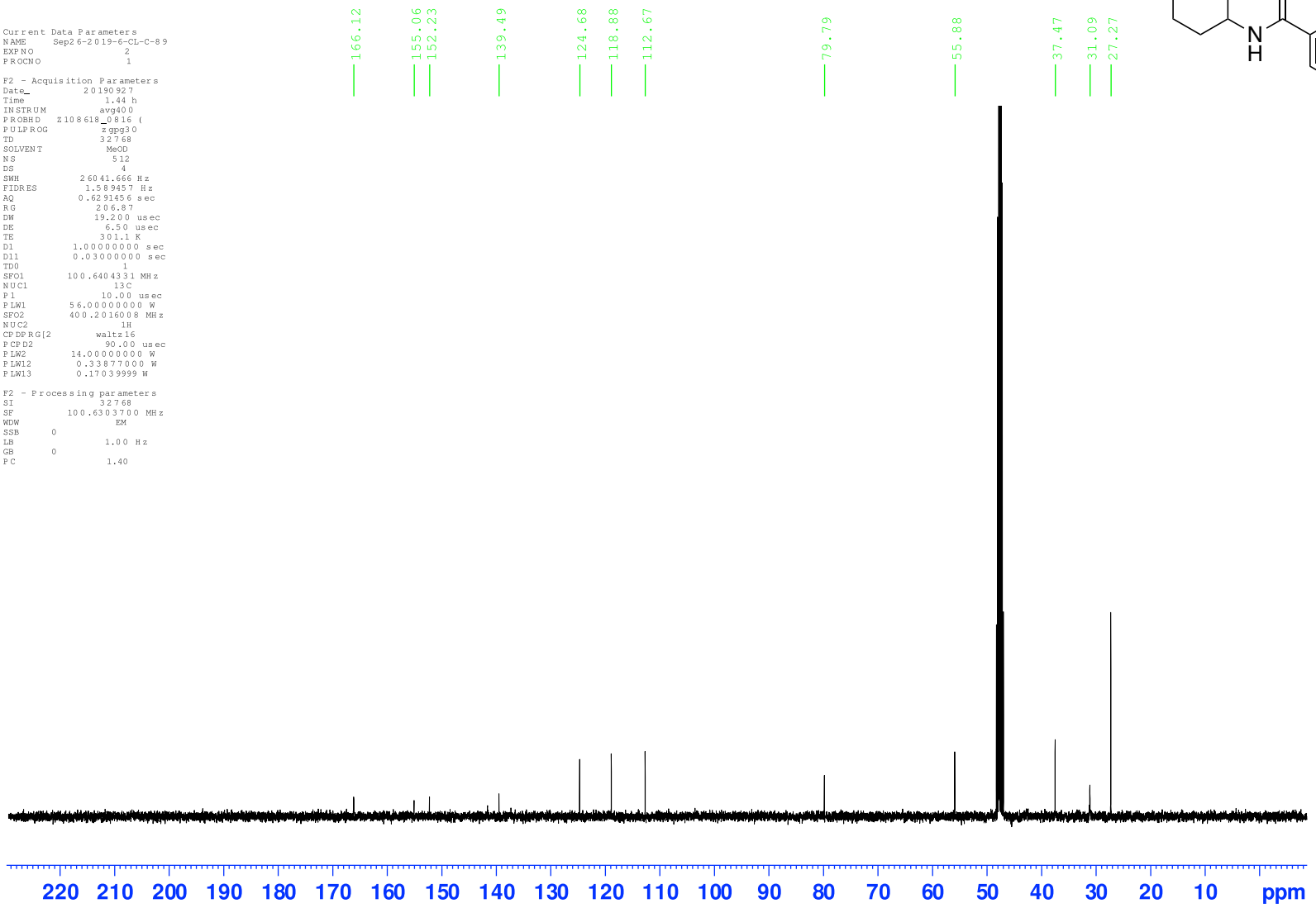


tert-Butyl 4-(3-methoxy-4-nitrobenzamido)piperidine-1-carboxylate 146a

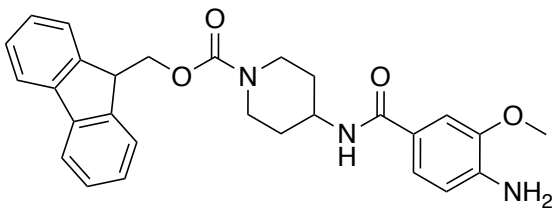


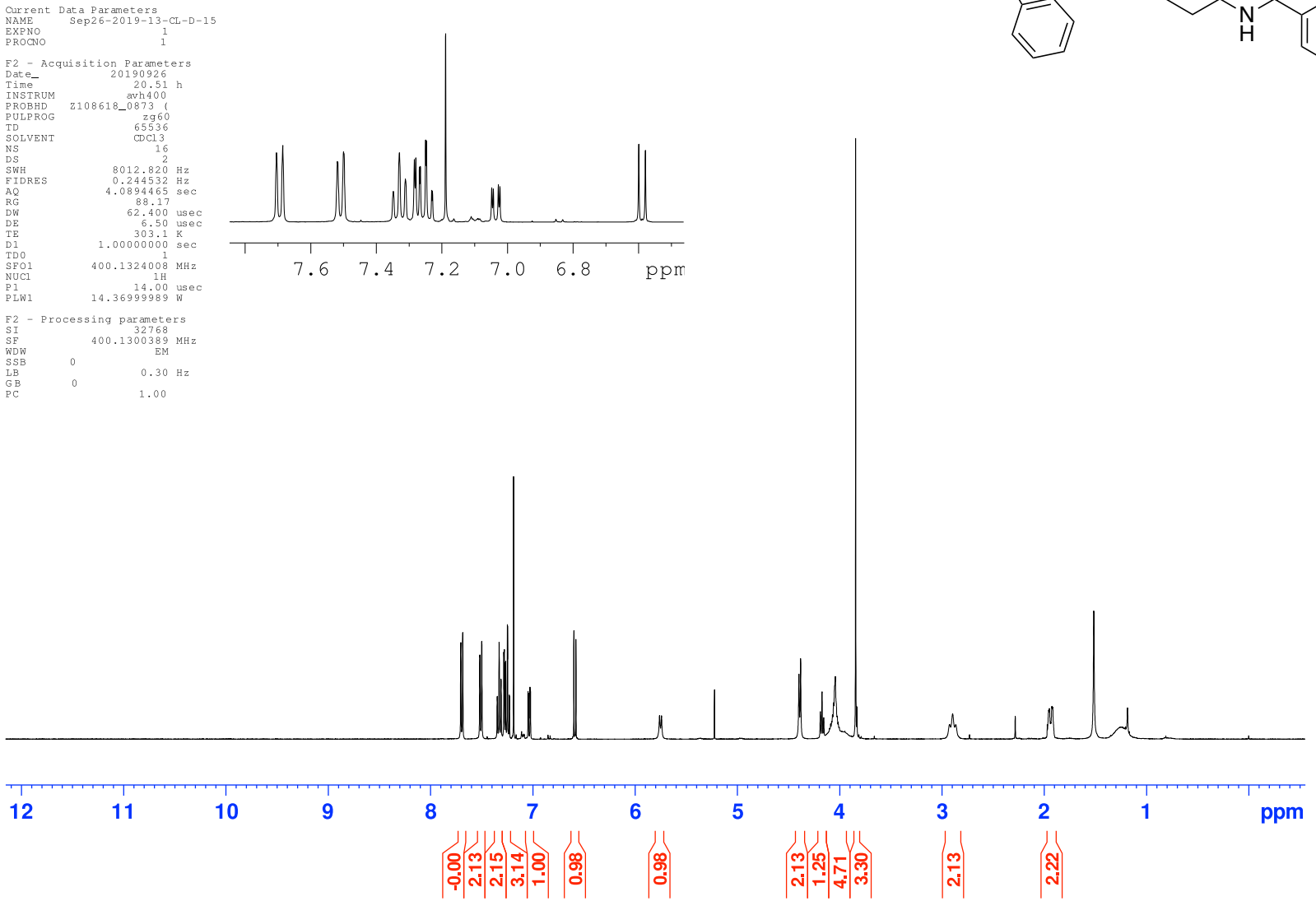
tert-Butyl 4-(3-methoxy-4-nitrobenzamido)piperidine-1-carboxylate 146a



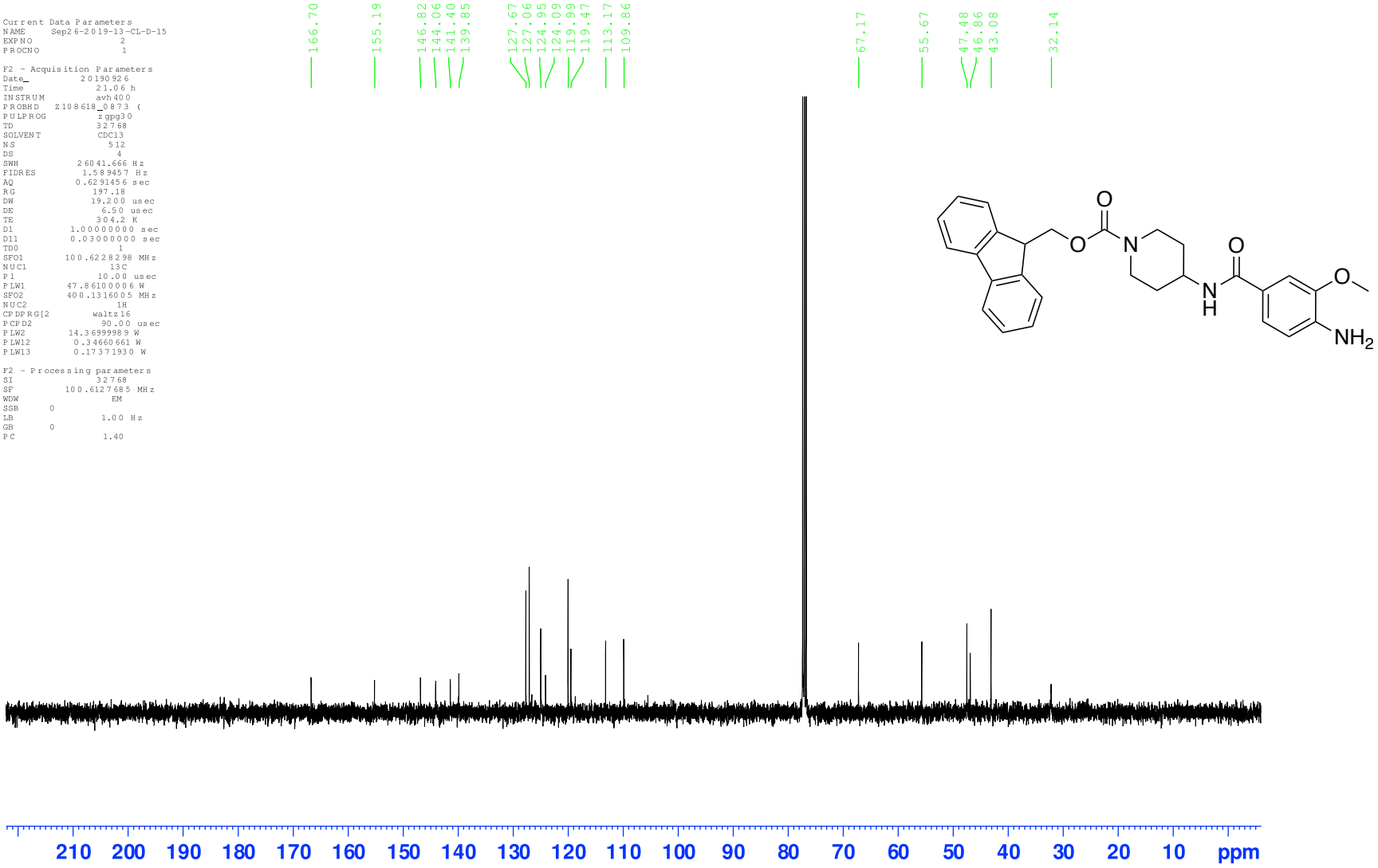


(9*H*-Fluoren-9-yl)methyl 4-(4-amino-3-methoxybenzamido)piperidine-1-carboxylate 146

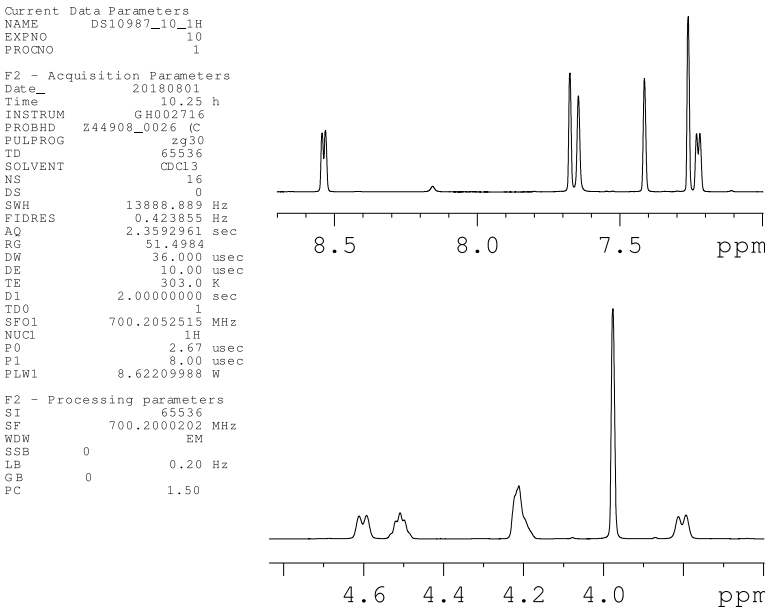


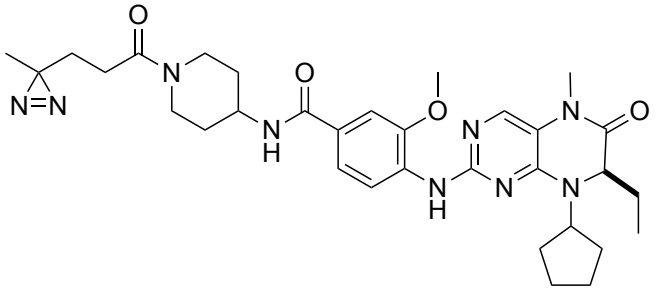


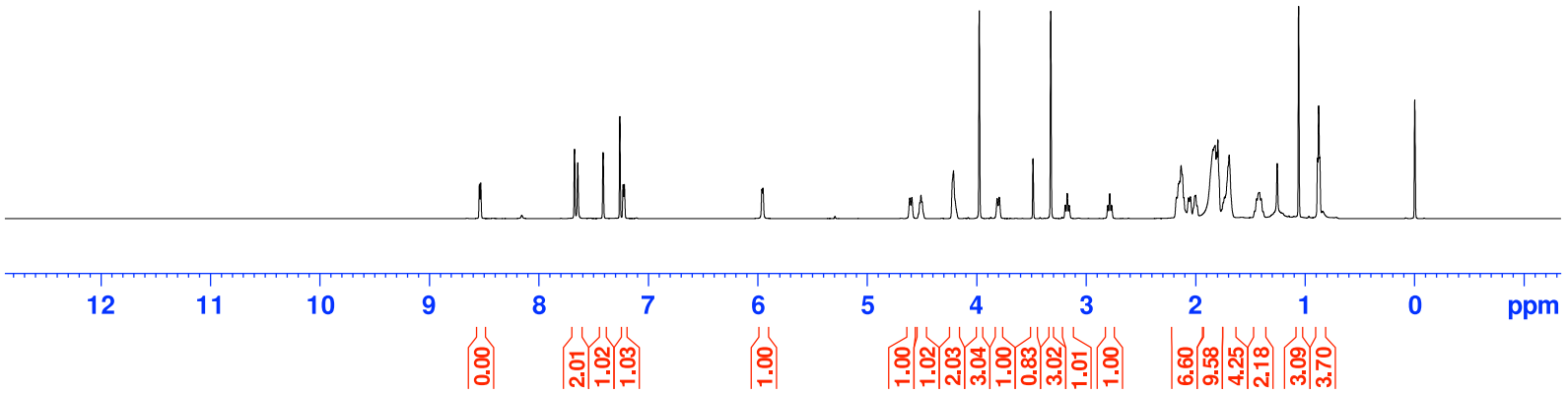
(9*H*-Fluoren-9-yl)methyl 4-(4-amino-3-methoxybenzamido)piperidine-1-carboxylate 146



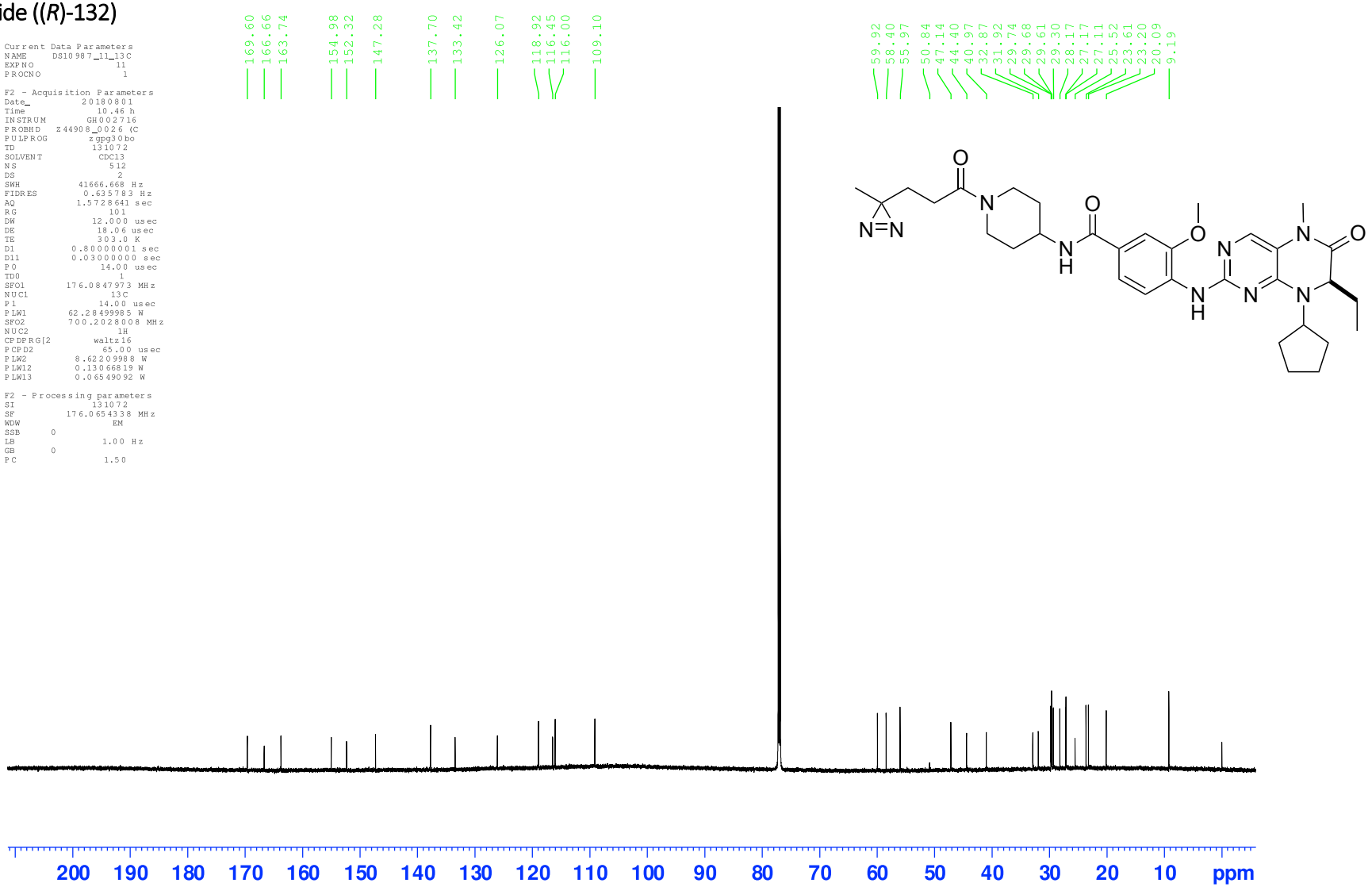
(*R*)-4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(1-(3-(3-methyl-3*H*-diazirin-3-yl)propanoyl)piperidin-4-yl)benzamide ((*R*)-132)



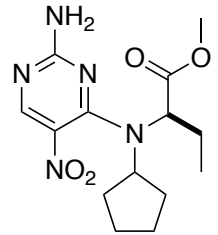


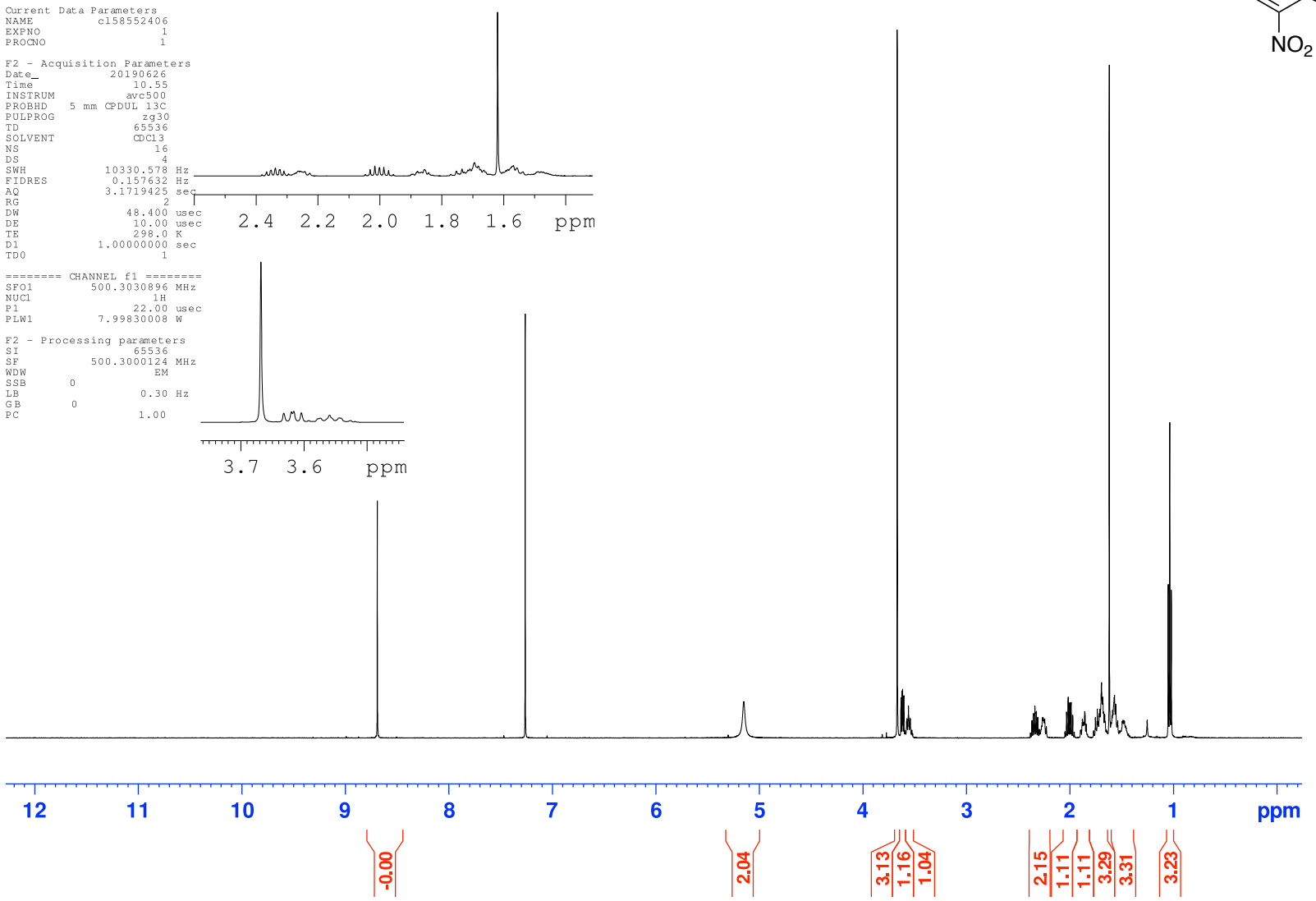


(*R*)-4-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxy-*N*-(1-(3-(3-methyl-3*H*-diazirin-3-yl)propanoyl)piperidin-4-yl)benzamide ((*R*)-132)

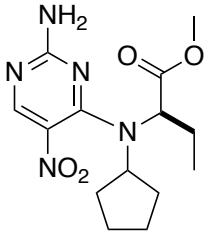


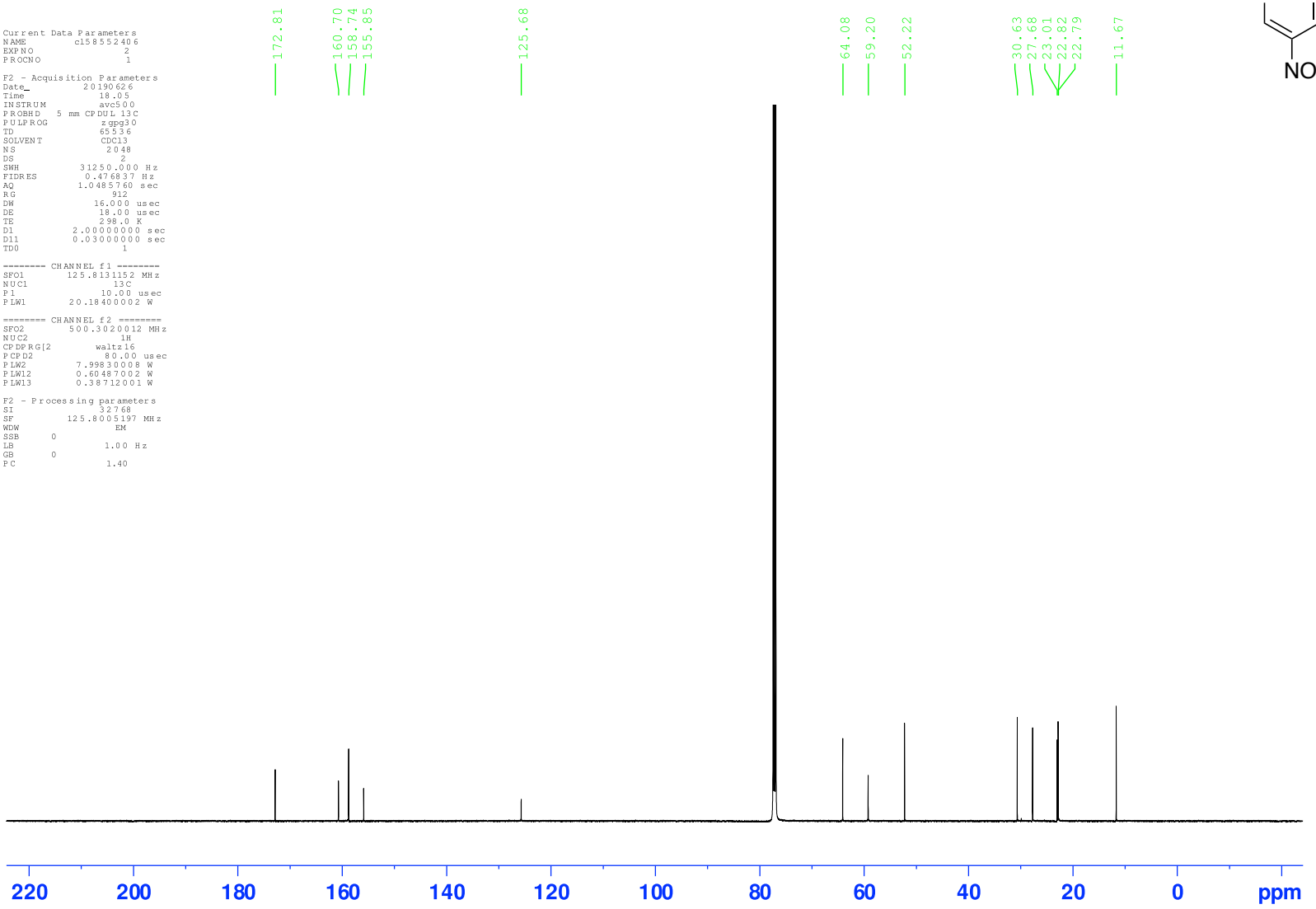
Methyl (*R*)-2-((2-amino-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate ((*R*)-140)



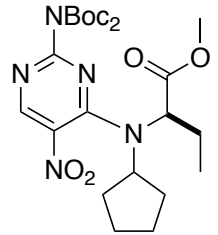


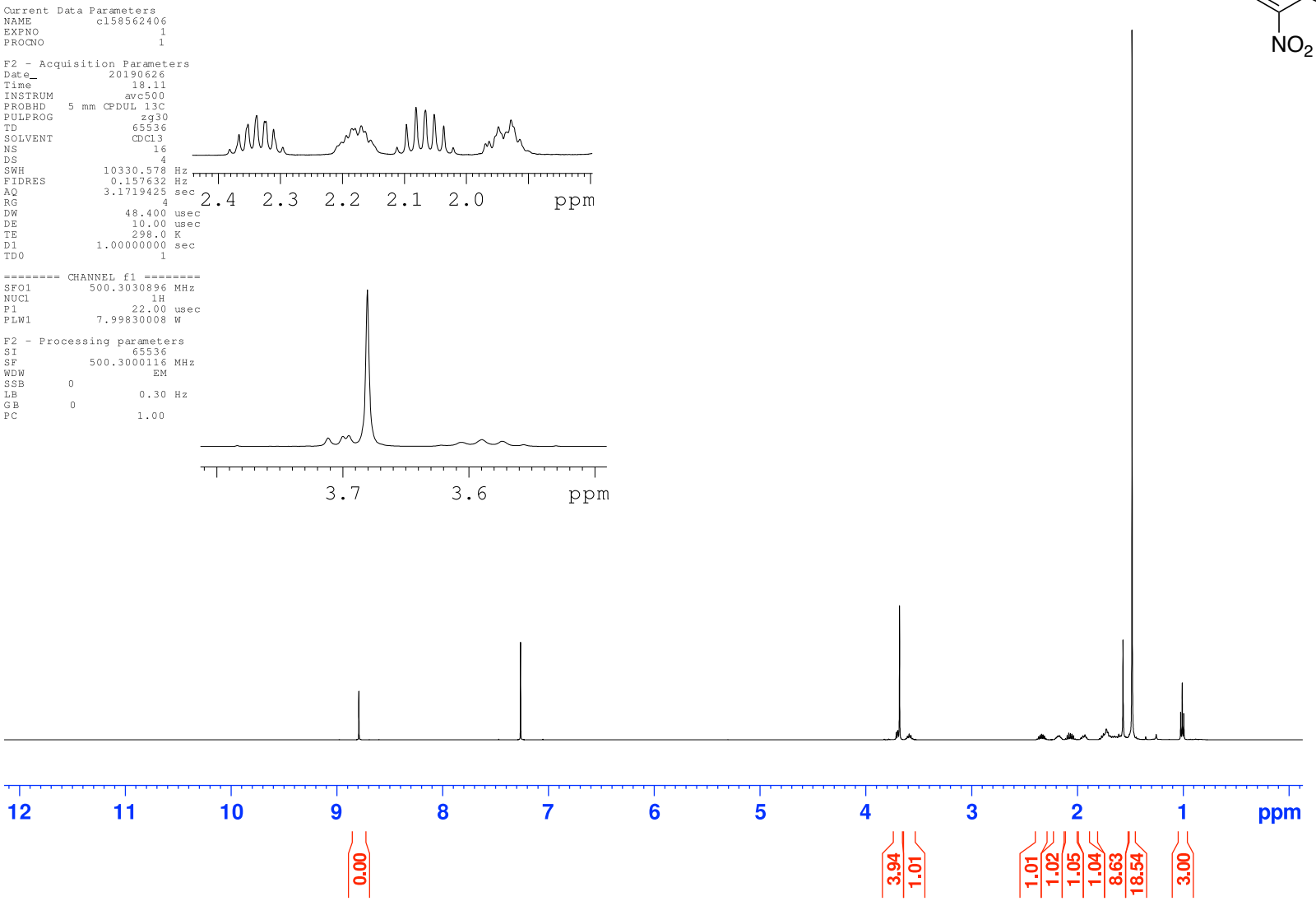
Methyl (*R*)-2-((2-amino-5-nitropyrimidin-4-yl)(cyclopentyl)amino)butanoate ((*R*)-140)



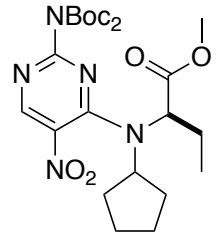


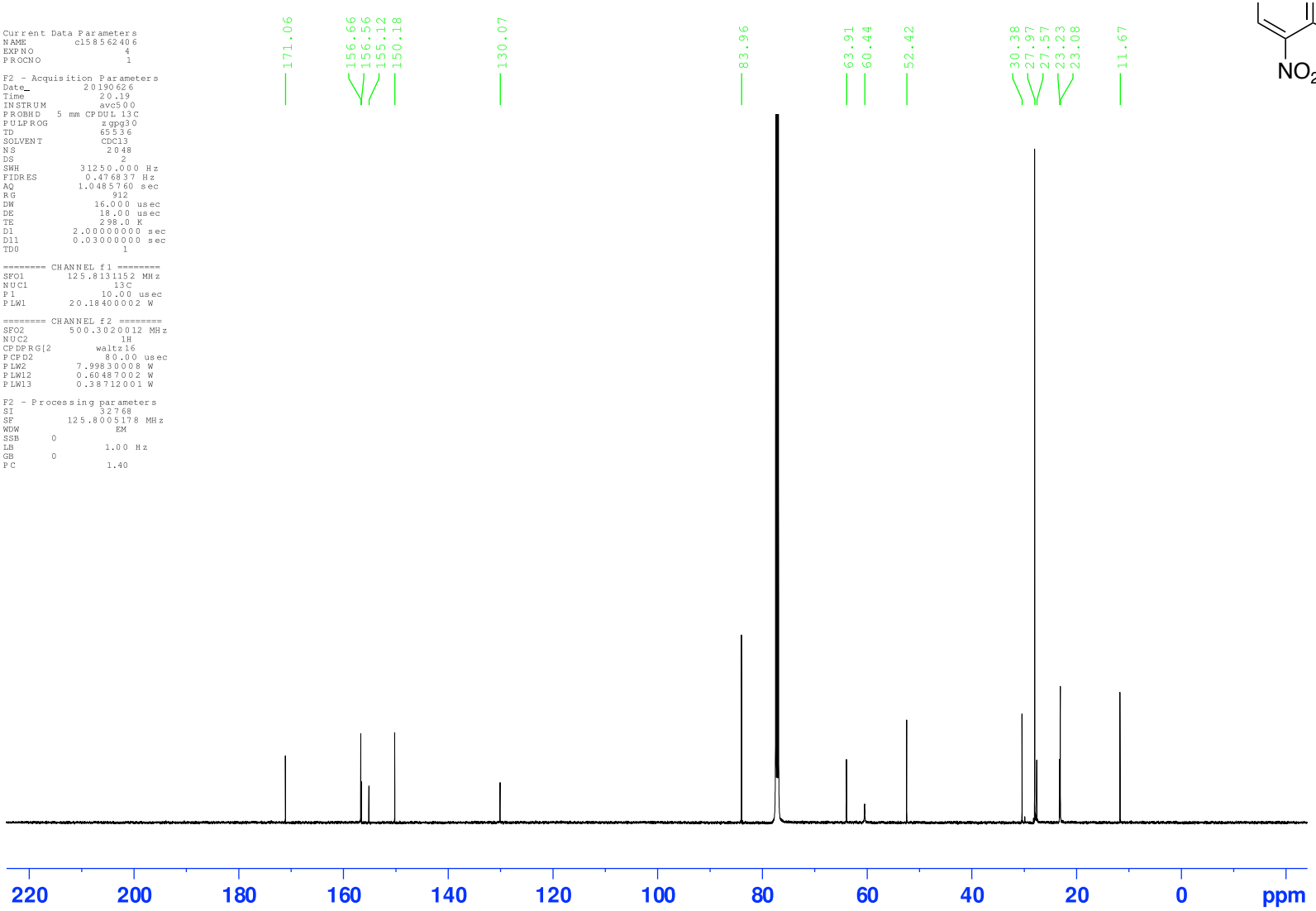
Methyl (*R*)-2-((2-(bis(*tert*-butoxycarbonyl)amino)-5-nitropyrimidin-4-yl)(cyclopentyl)amino) butanoate ((*R*)-142)



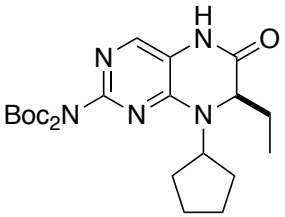


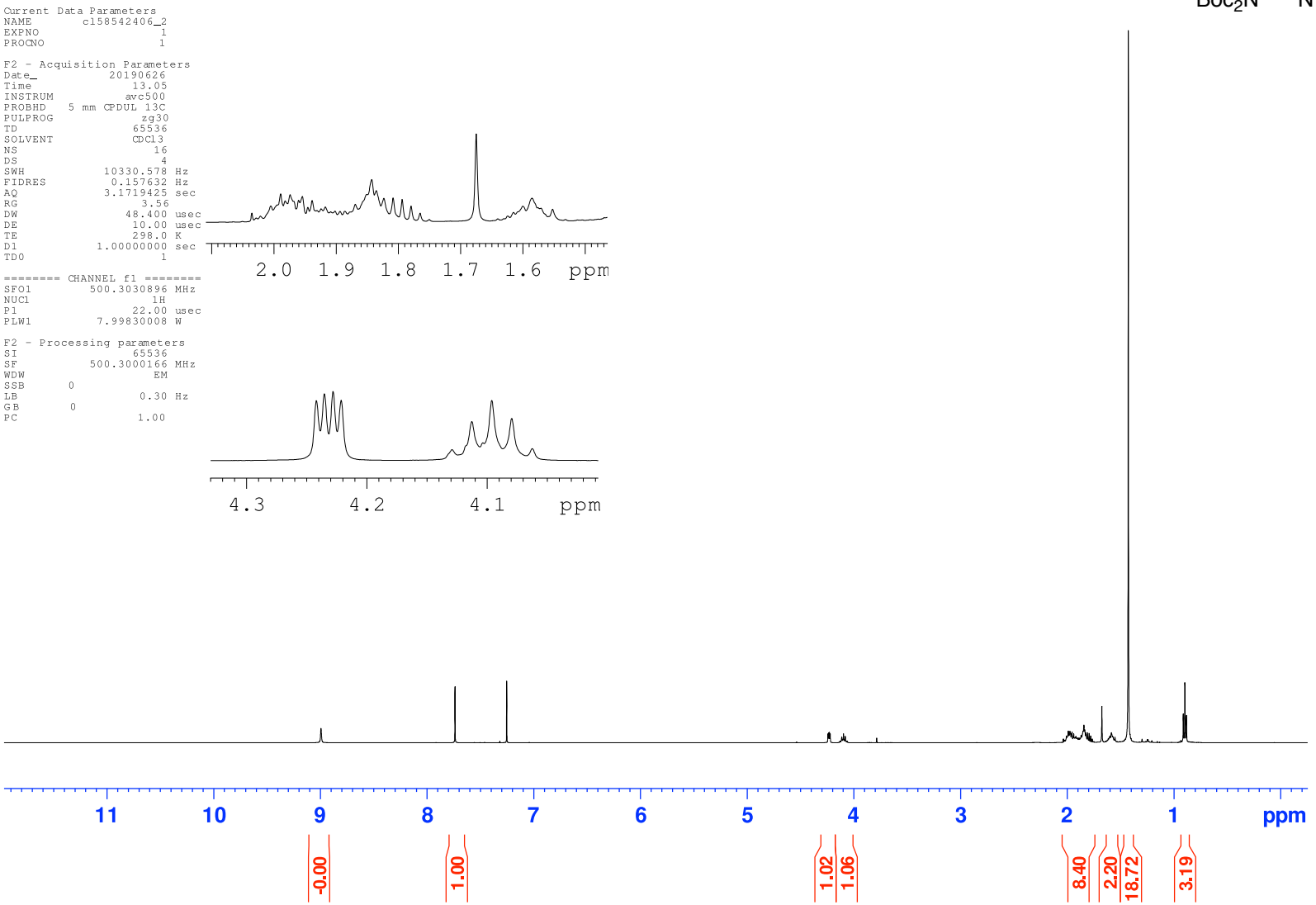
Methyl (*R*)-2-((2-(bis(*tert*-butoxycarbonyl)amino)-5-nitropyrimidin-4-yl)(cyclopentyl)amino) butanoate ((*R*)-142)



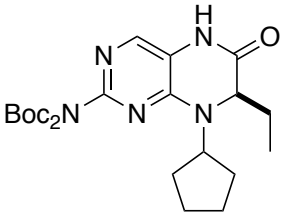


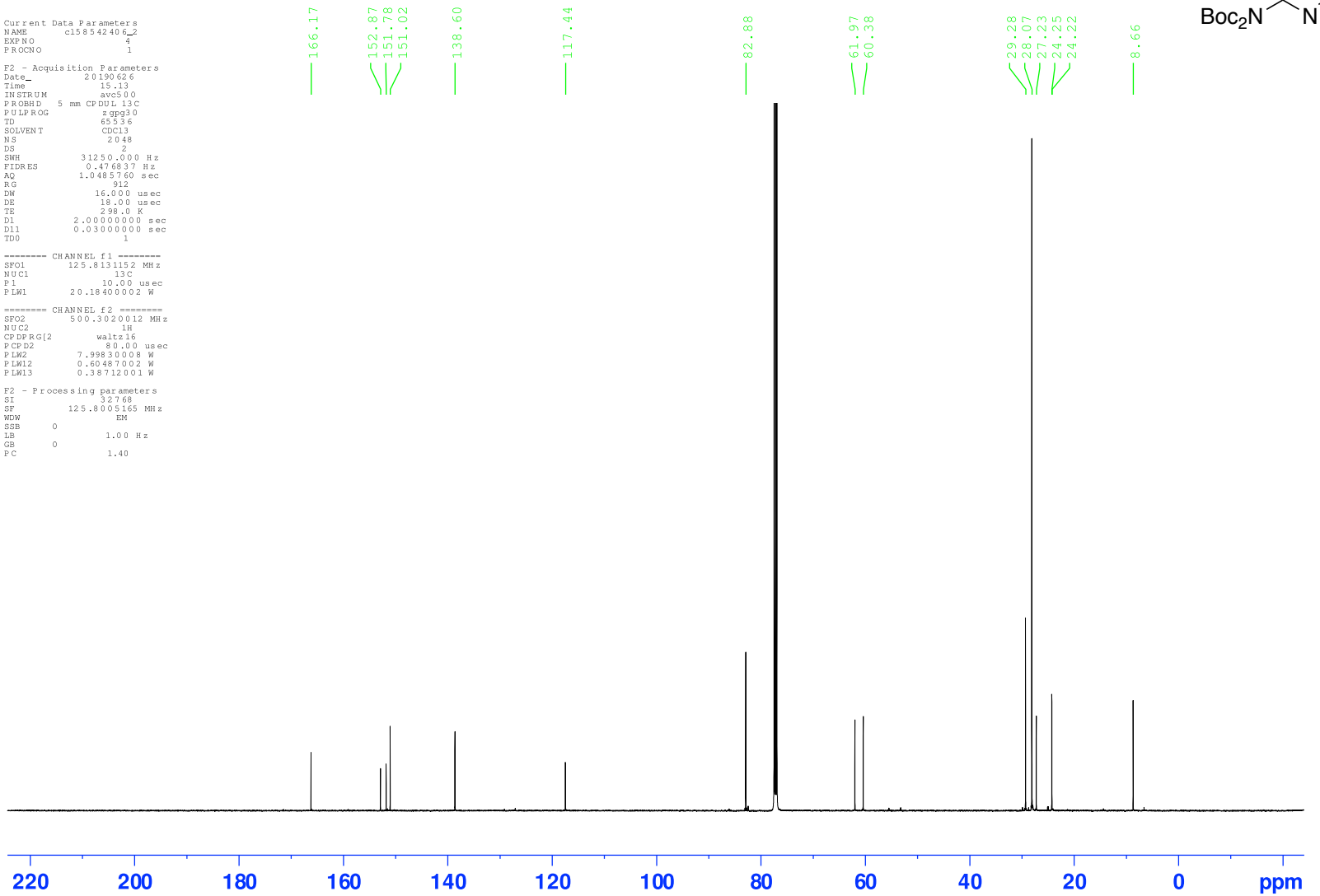
tert-Butyl (*R*)-(*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate ((*R*)-143)



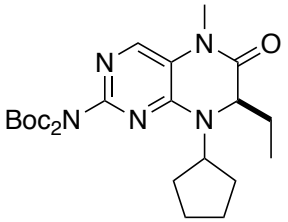


tert-Butyl (*R*)-(*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate ((*R*)-143)

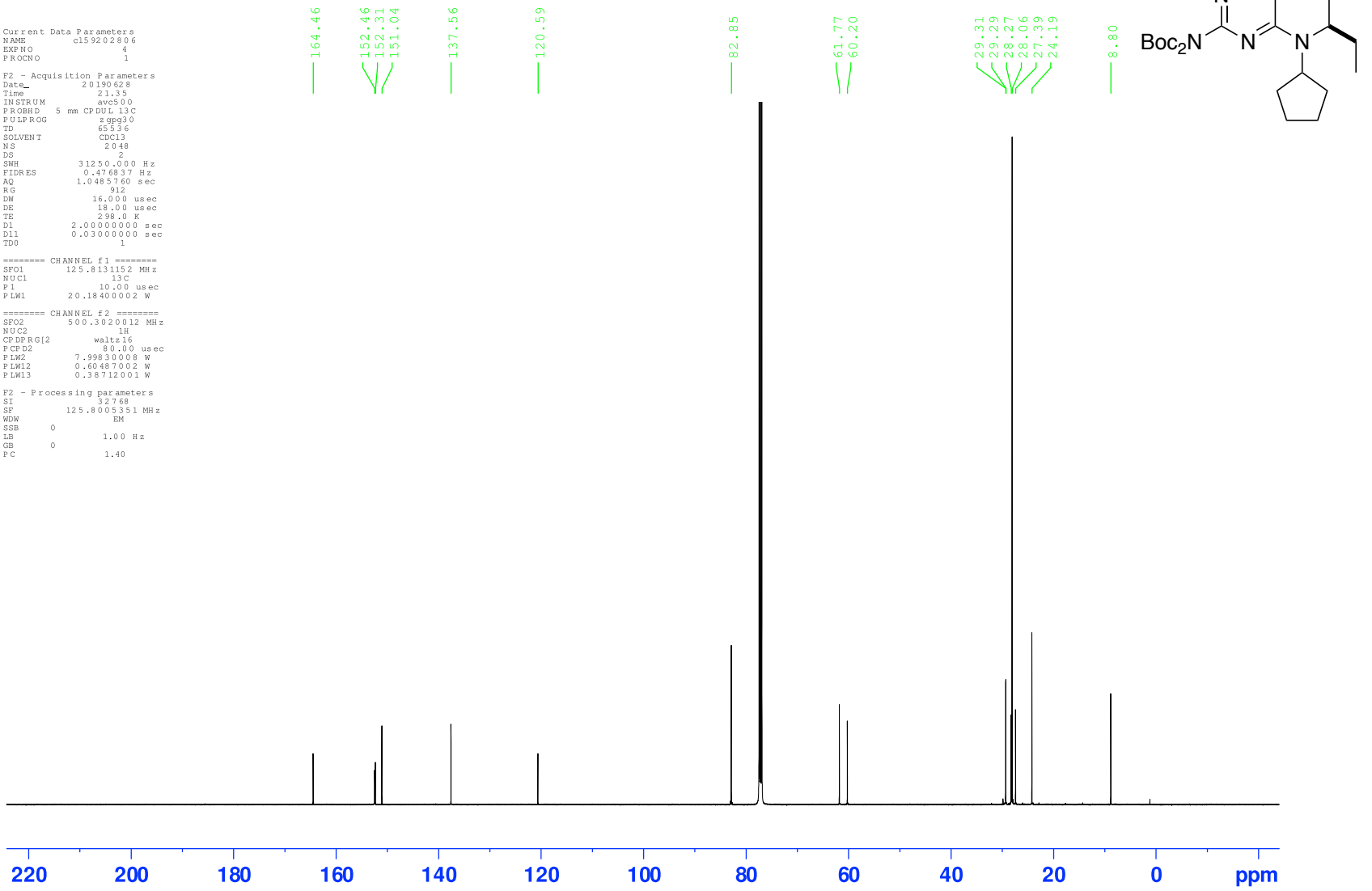




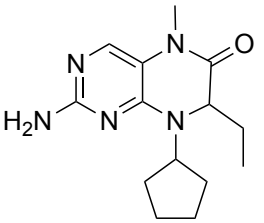
tert-Butyl (*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate ((*R*)-144)

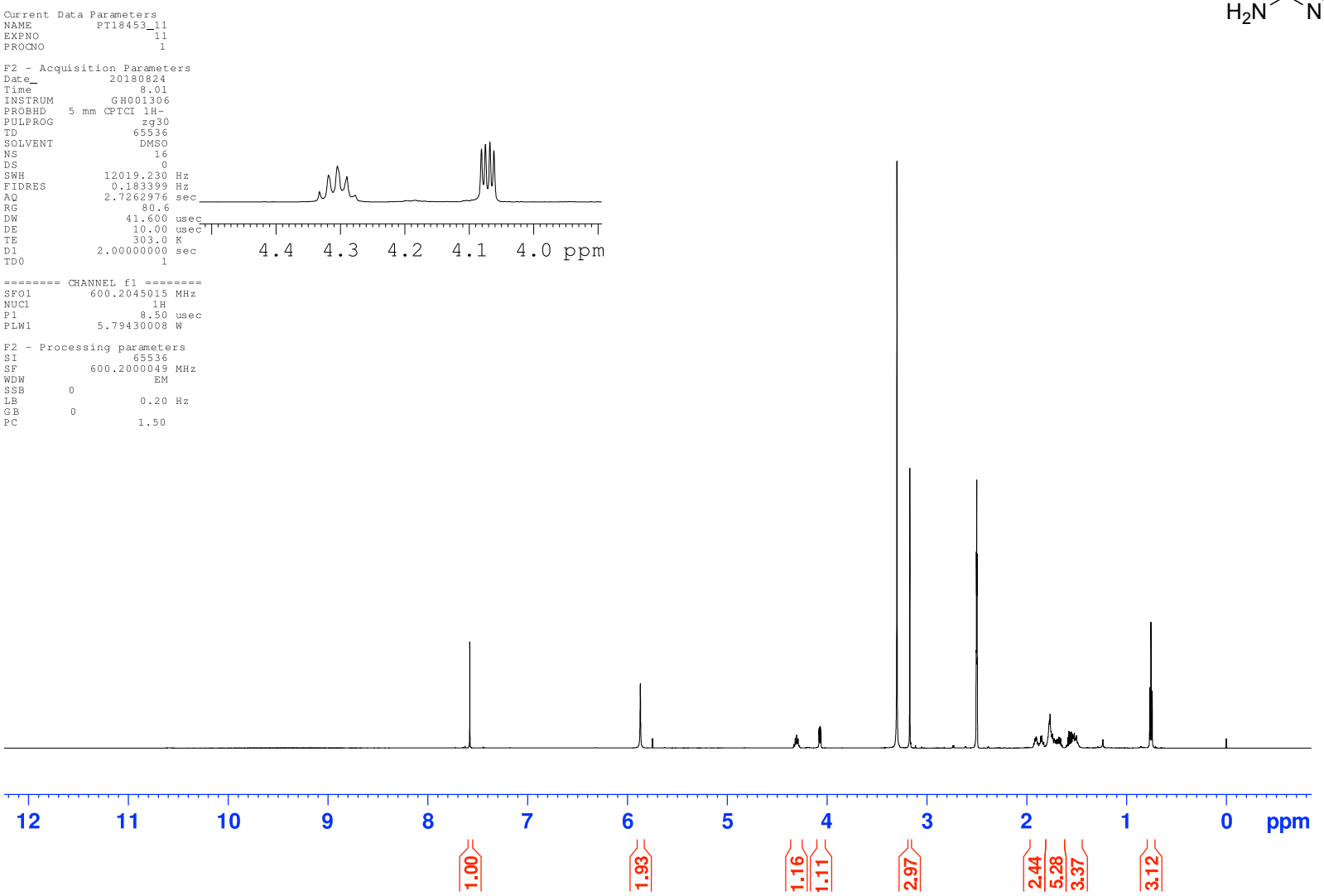


tert-Butyl (*tert*-butoxycarbonyl)(8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)carbamate ((*R*)-144)

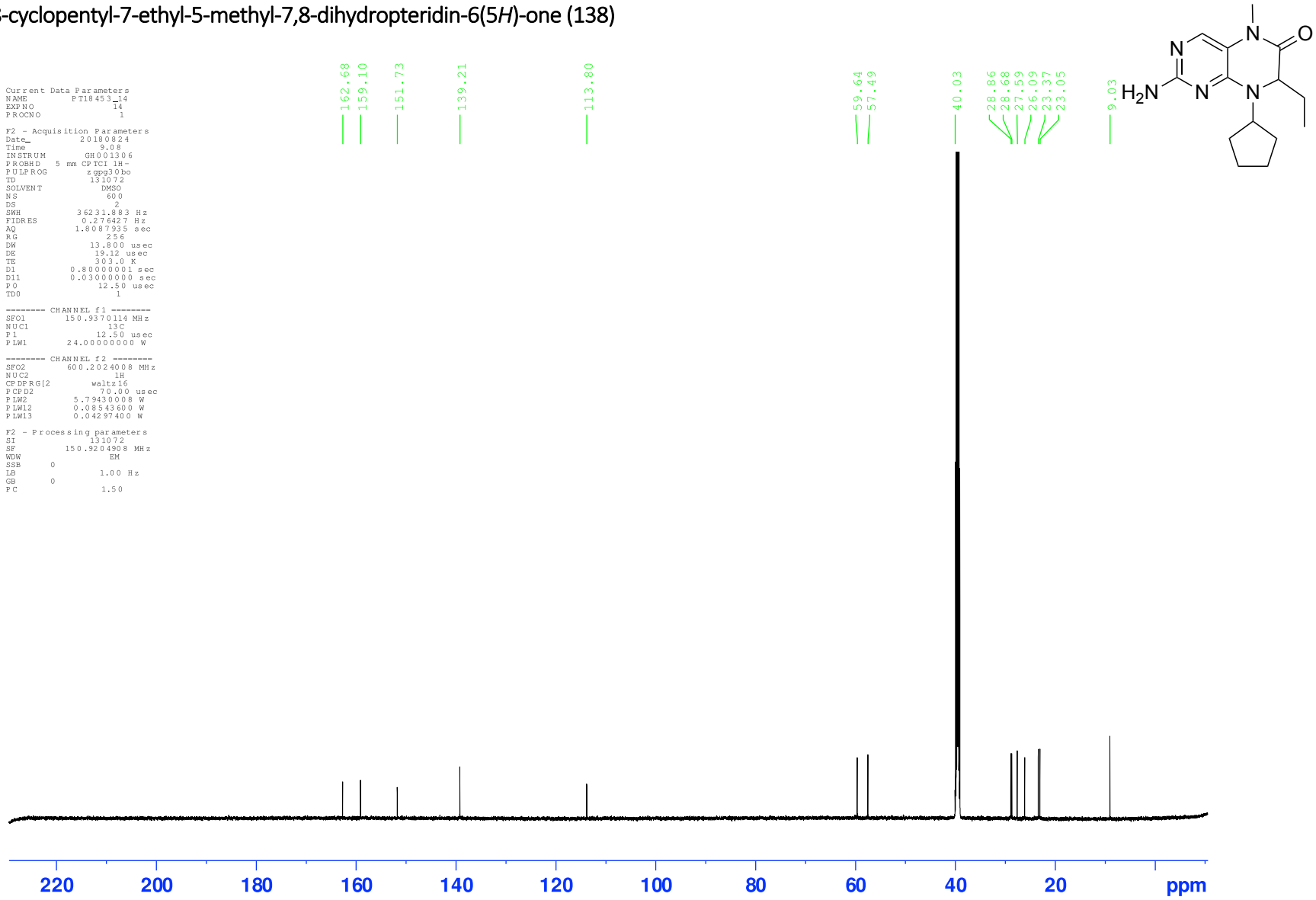


2-Amino-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (138)

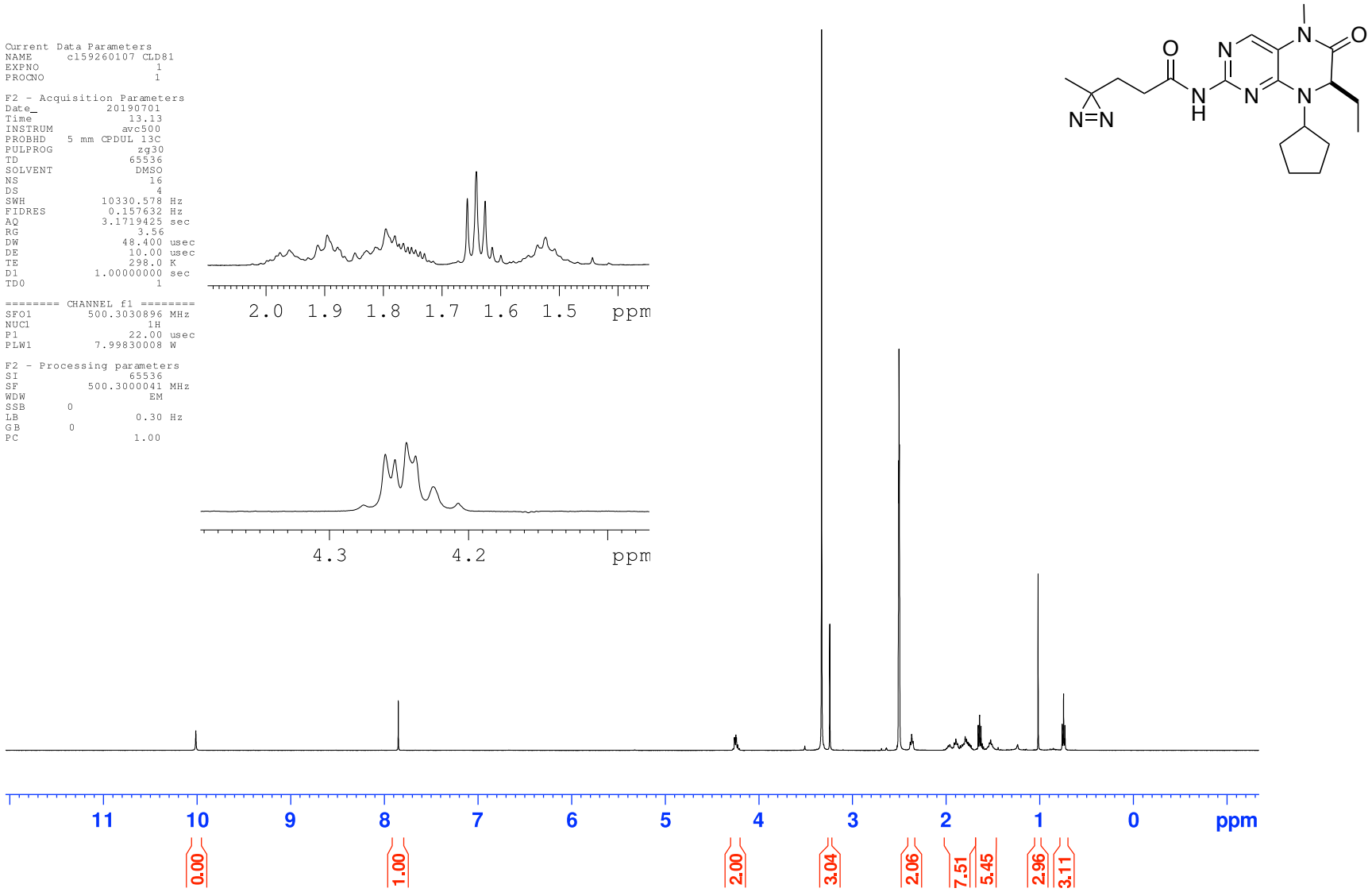




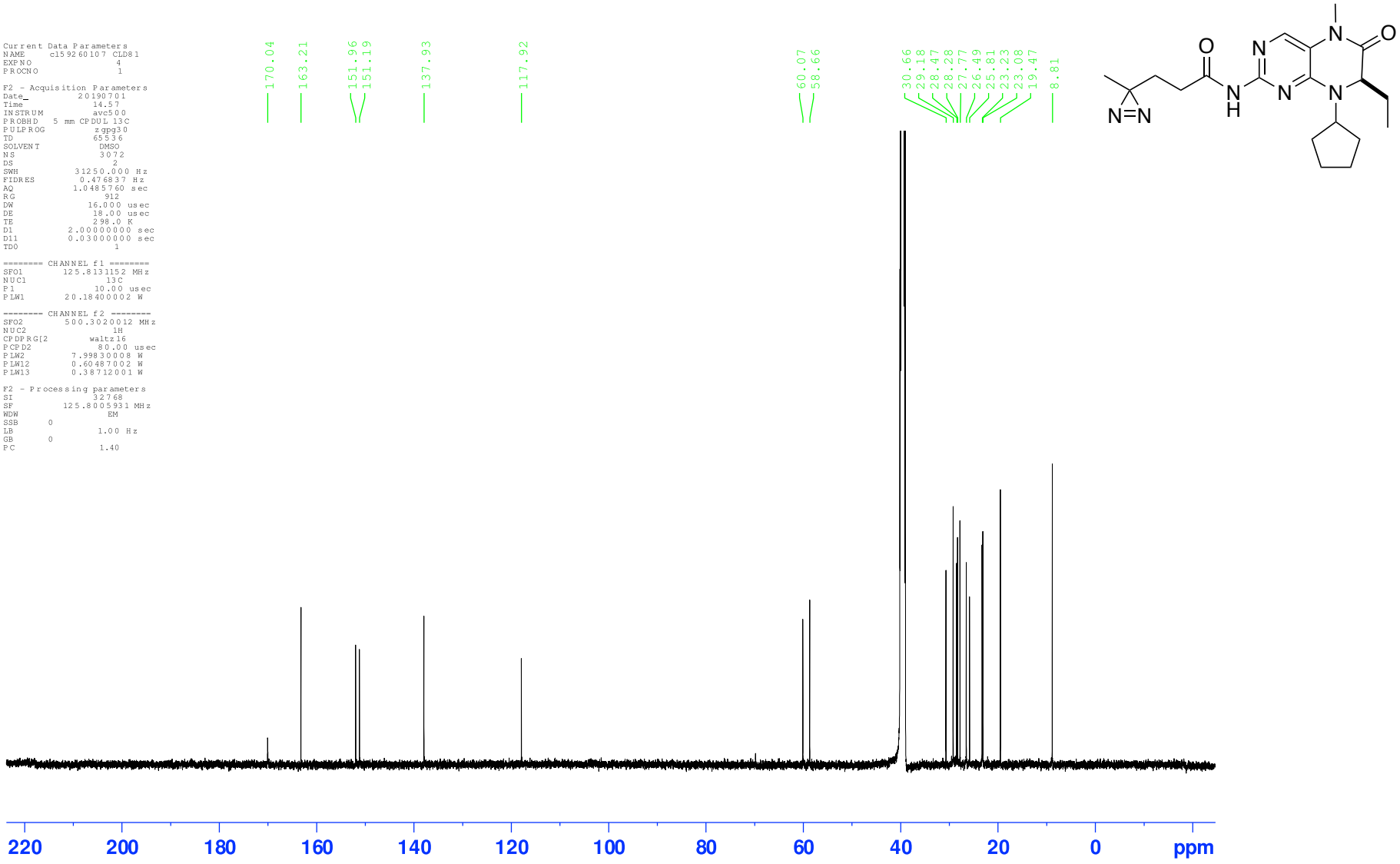
2-Amino-8-cyclopentyl-7-ethyl-5-methyl-7,8-dihydropteridin-6(5*H*)-one (138)



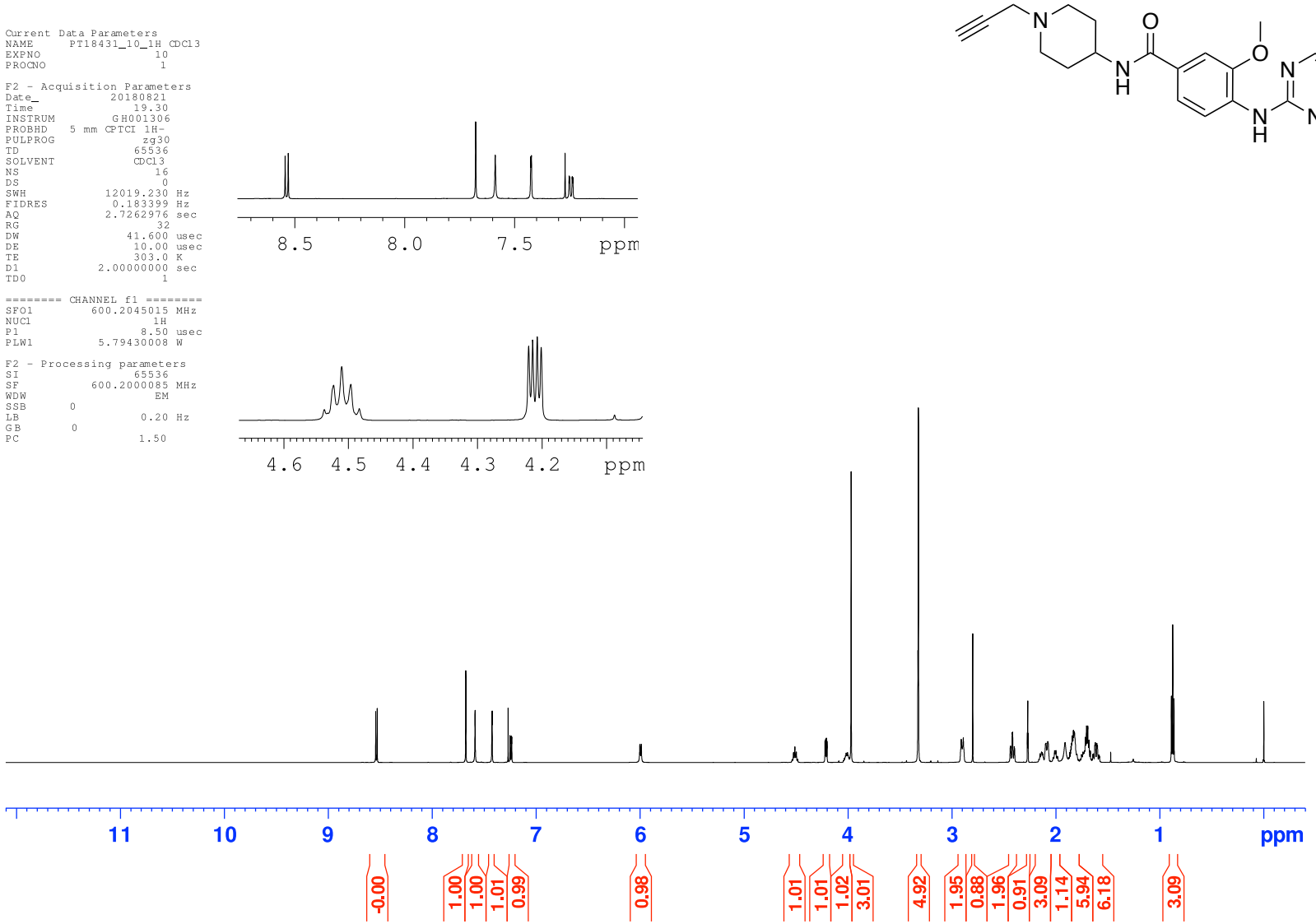
(*R*)-*N*-(8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)-3-(3-methyl-3*H*-diazirin-3-yl)propenamide ((*R*)-133)



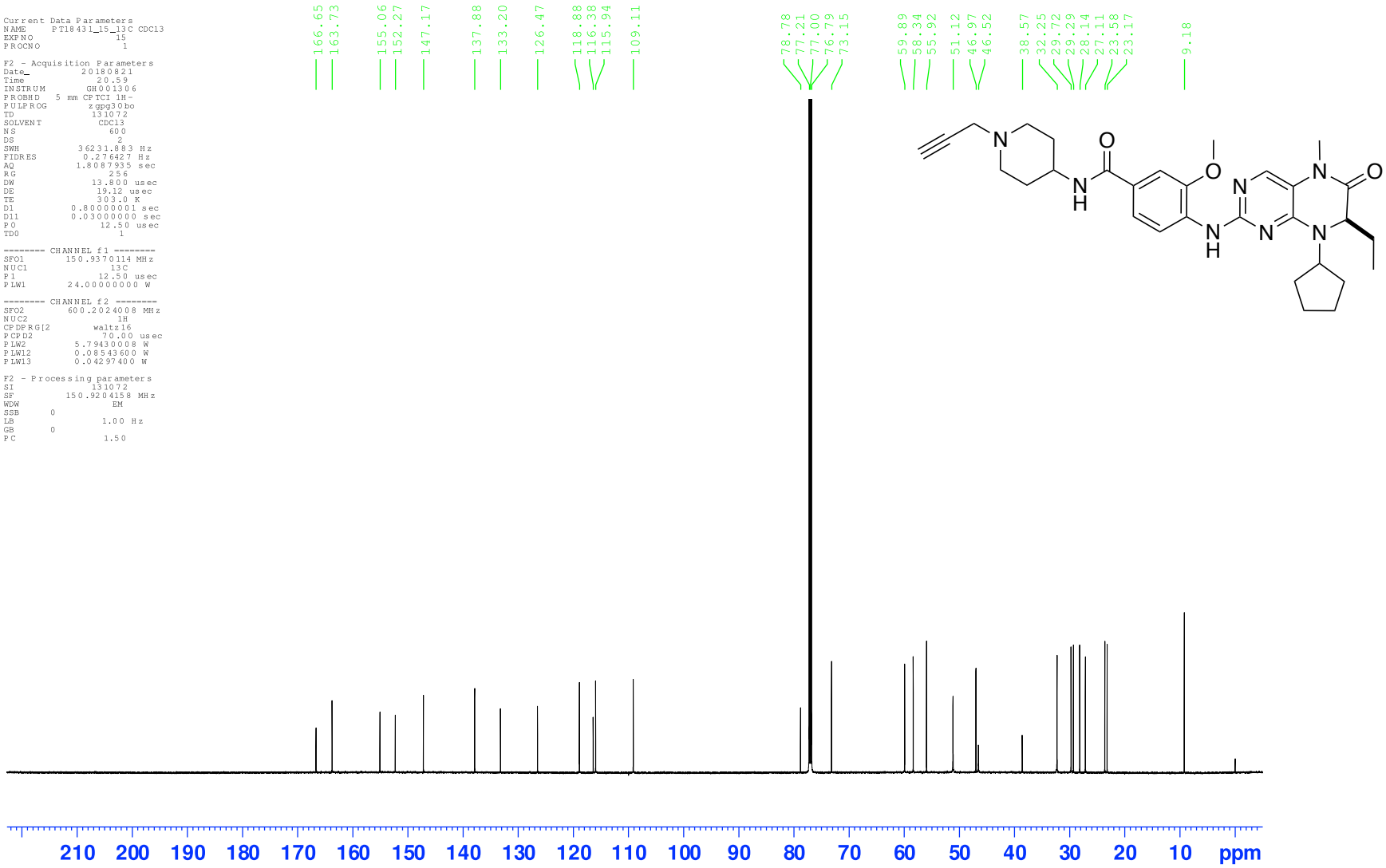
(*R*)-*N*-(8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)-3-(3-methyl-3*H*-diazirin-3-yl)propenamide (((*R*)-133)



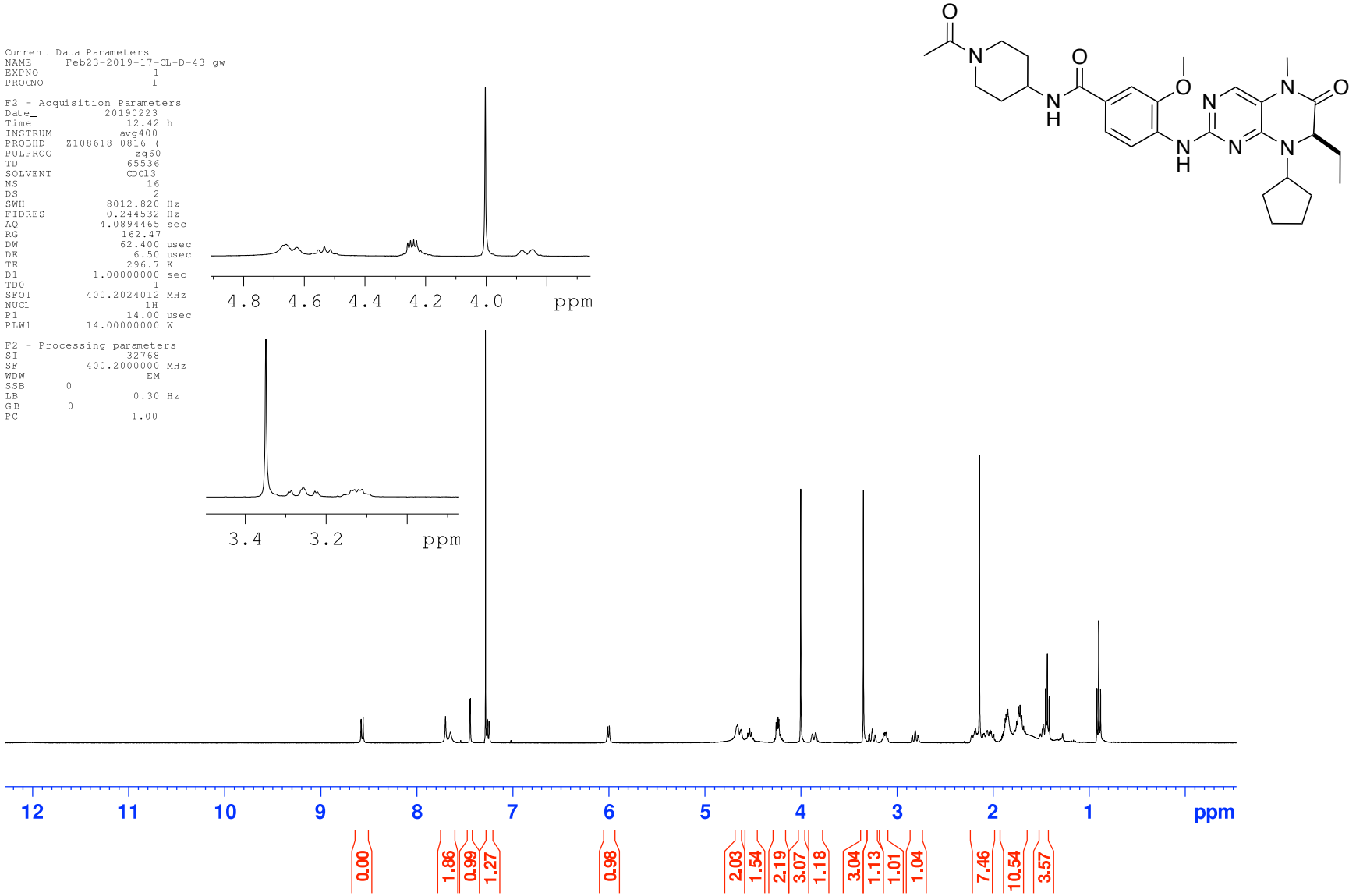
(7*R*)-4'-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxy-*N*-(1-(prop-2-yn-1-yl)piperidin-4-yl)benzamide ((*R*)-153)



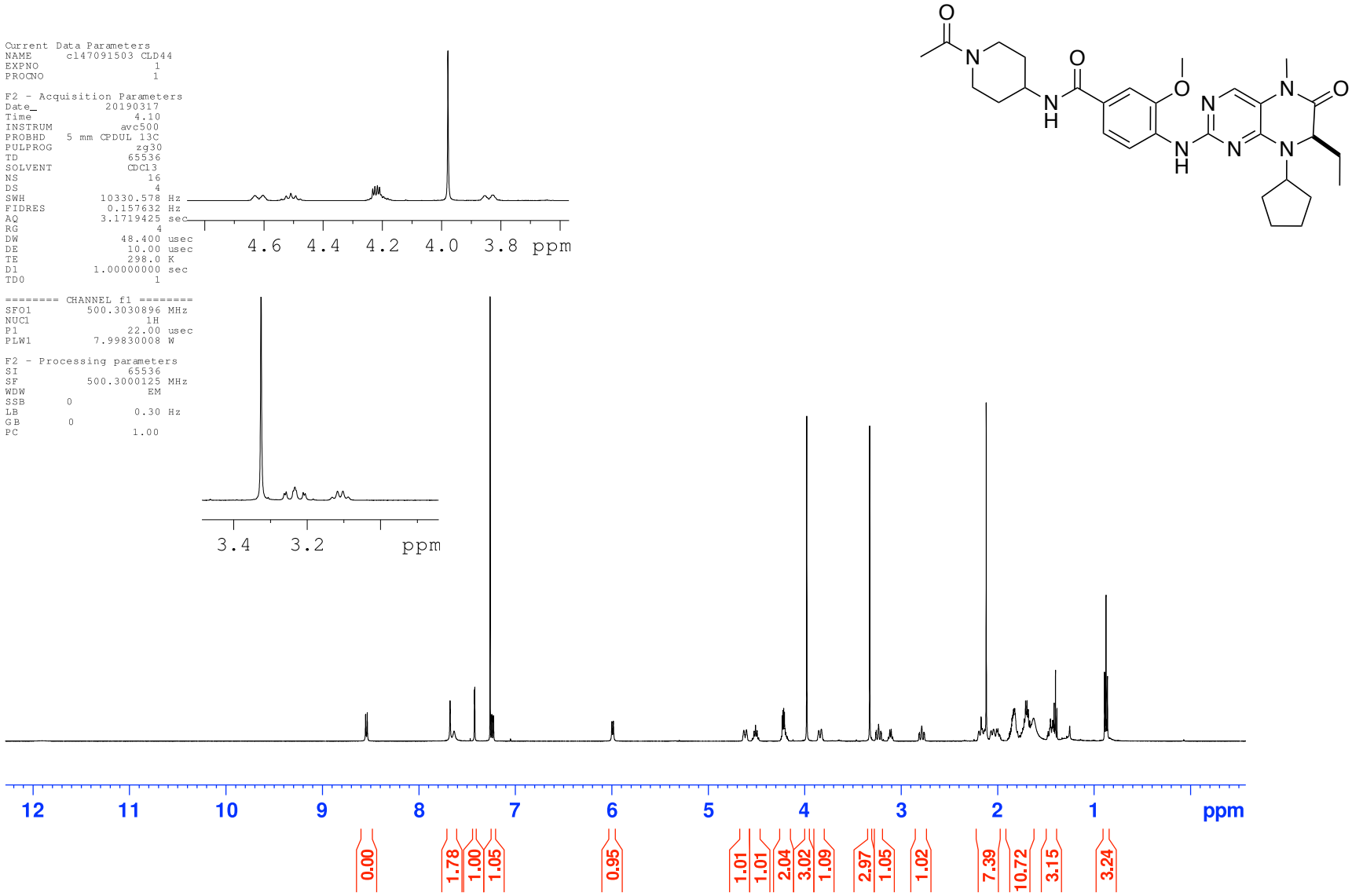
(7*R*)-4'-((8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3'-methoxy-*N*-(1-(prop-2-yn-1-yl)piperidin-4-yl)benzamide ((*R*)-153)



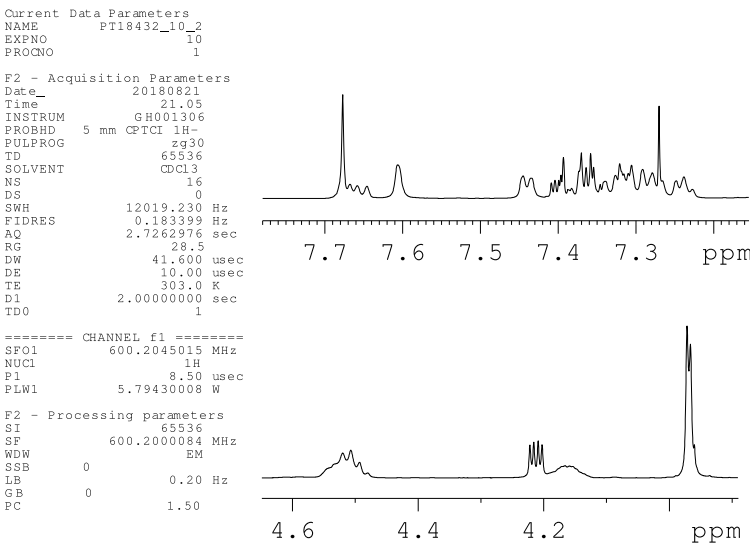
(*R*)-*N*-(1-Acetylpiperidin-4-yl)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamide ((*R*)-155)

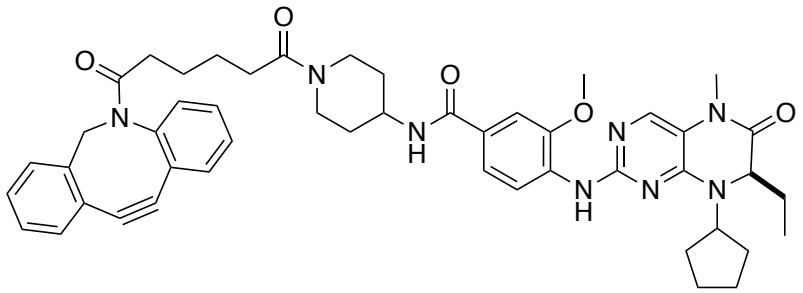


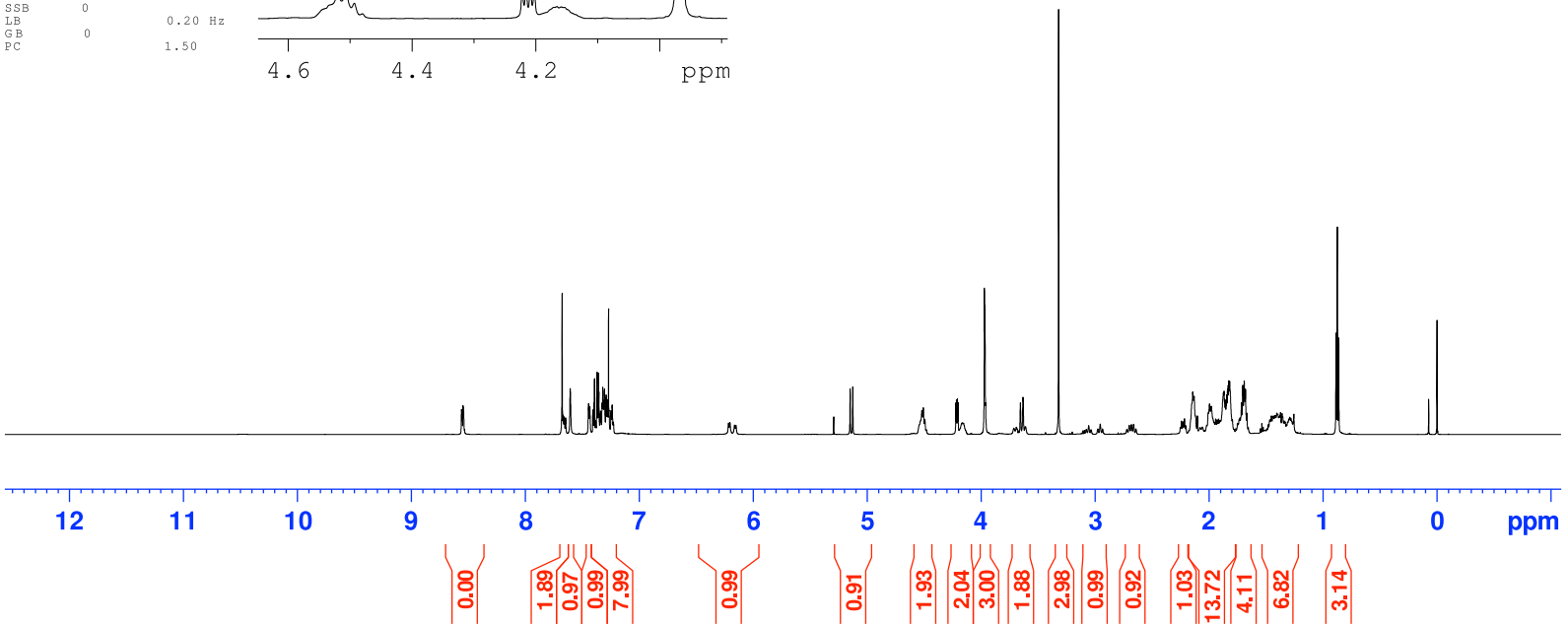
(*R*)-*N*-(1-Acetylpiperidin-4-yl)-4-((8-cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl)amino)-3-methoxybenzamide ((*R*)-155)



4-{[(7*R*)-8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl]amino}-*N*-{1-[6-(11,12-didehydrodibenzo[*b*,*f*]azocin-5(6*H*)-yl)-6-oxohexanoyl]piperidin-4-yl}-3-methoxybenzamide ((*R*)-154)







4-{[(7*R*)-8-Cyclopentyl-7-ethyl-5-methyl-6-oxo-5,6,7,8-tetrahydropteridin-2-yl]amino}-*N*-{1-[6-(11,12-didehydrodibenzo[*b*,*f*]azocin-5(6*H*)-yl)-6-oxohexanoyl]piperidin-4-yl}-3-methoxybenzamide ((*R*)-154)

