Guanidinium mimetics: synthesis and application as next-generation analogues in cell penetrating peptides

Jack Robertson

Thesis submitted for the degree of Doctor of Philosophy

29th January 2021

Academic Supervisor: Prof. G. A. Burley

Industrial Supervisor: Dr. K. L. Jones

Declaration of Authenticity

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:  (Jack Robertson)

Date: 29th January 2021.

Table of Contents

Acknowledgements 6
Abbreviations 7
Abstract 10
Chapter 1: Introduction 12
1.1 The ever-present problem of cellular uptake 13
1.2 The molecular architecture of the cell membrane 13
1.2.1 Phospholipids 13
1.2.2 Cell membrane proteins 14
1.2.3 Glycoproteins and glycolipids 15
1.2.4 Cholesterol 16
1.3 Mechanisms of cell uptake 17
1.3.1 Passive and facilitated diffusion 17
1.3.2 Active transport mechanisms of cell uptake 18
1.3.3 Endocytosis 19
1.3.4 Permeability of synthetic small molecules 22
1.4 Drug delivery platforms 24
1.4.1 Nanoparticles 24
1.4.2 Liposomes 25
1.4.3 Cell penetrating peptides 26
1.5 Peptides and peptidomimetics as probes and therapeutics 27
1.6 Structure and function of cell penetrating peptides 30
1.6.1 Early developments of CPPs 31
1.6.2 Interactions of CPPs and cell surface species 31
1.6.3 low arginine content CPPs 32
1.6.4 Methods of cell uptake for CPPs 35
1.6.5 Clinical utility of CPPs 37
1.6.6 Cell-penetrating peptidomimetics 38
1.6.7 Drawbacks of existing CPPs 40
1.7 The role of the guanidinium group to enhance cell uptake 41
1.7.1 Bioisosteres of the guanidine group 42
1.7.2 Amidines as isosteres of guanidine 42
1.8 Hypothesis 44
1.9 Aims and objectives of this thesis 44
Chapter 2: Toxicity and uptake of low arginine-containing CPPs 46
2.1 Introduction 47
2.1.1 Synthesis of peptides 47

2.1.2 Monitoring cell permeability of peptides 55
2.2 Aims of Chapter 2 60
2.3 Results and discussion 61
2.3.1 Design and synthesis of CPPs 61
2.3.2 Cell-viability with fluorescently-labelled CPPs 67
2.3.3 Uptake of fluorescently labelled peptides 69
2.3.4 Sub-cellular localisation of CPPs 75
2.4 Summary of Chapter 2 78
2.5 Experimental 79
2.5.1 Peptide synthesis 79
2.5.2 Analytical HPLC 79
2.5.3 Purification of products 79
2.5.4 Mass spectrometry 80
2.5.5 Peptide characterisations 80
2.5.6 Cell culture 84
2.5.7 Flow cytometry 84
2.5.8 Cell viability alamarBlue assay 84
2.5.9 Confocal microscopy 85
Chapter 3: Development of robust Cu-catalysed *N*-arylation of amidines 86
3.1 Introduction 87
3.1.1 *N*-Arylated amidines 87
3.1.2 Transition metal-mediated carbon-nitrogen bond forming reactions 87
3.1.3 Buchwald-Hartwig amination 88
3.1.4 Cu-catalysed Ullmann couplings 91
3.1.5 Chan-Evans-Lam amination 94
3.2 Aims of Chapter 3 99
3.3 Results and discussion 100
3.3.1 Optimisation of a Chan-Lam amidination 100
3.3.2 Mechanistic analysis of the Chan-Lam amidination 104
3.3.3 Major *N*-arylated tautomer formed from the Chan-Lam amidination 110
3.3.4 Substrate scope of the Chan-Lam amidination 112
3.3.5 Chan-Lam amidination of Pentamidine 115
3.3.6 Anti-trypanosome activity of *N*-arylated pentamidine analogues 117
3.4 Summary of Chapter 3 120
3.5 Experimental 121
3.5.1 General information 121
3.5.2 Analytical HPLC 121
3.5.3 Purification of products 121

3.5.4 Analysis of products 122
3.5.5 Sample preparation for EPR spectroscopy 122
3.5.6 EPR measurements 123
3.5.7 Synthetic procedures 123
3.5.8 Reaction optimisation 124
3.5.9 Cu source/benzamidine eq trials 126
3.5.10 Representative HPLC trace (Table 3.10, Entry 21) 127
3.5.11 EPR results 127
3.5.12 UV-vis experiments 128
3.5.13 Precipitate formation in reactions 129
3.5.14 Compound characterisations 129
3.5.15 Cu complex experimental procedures 144
3.5.16 *Trypanosoma* experiments 144
Chapter 4: Synthesis, toxicity and uptake of arginine mimetic CPPs 146
4.1 Introduction 147
4.1.1 Protecting groups for the guanidino group for use in SPPS 147
4.1.2 Synthesis of amidines as bioisosteres 149
4.2 Aims of Chapter 4 154
4.3 Results and discussion 155
4.3.1 Alkyl amidine isosteres 155
4.3.2 Aryl amidine isosteres 157
4.3.3 Peptidomimetics of TP peptides 166
4.3.4 Viability of peptidomimetics 169
4.3.5 Cell uptake of peptidomimetics 170
4.3.6 Sub-cellular localisation of peptidomimetics 173
4.4 Summary of Chapter 4 175
4.5 Experimental 176
4.5.1 General information 176
4.5.2 Peptide synthesis 176
4.5.3 Analytical HPLC 176
4.5.4 LCMS Analysis 177
4.5.5 Purification of products 177
4.5.6 Analysis of products 178
4.5.7 Synthetic procedures 178
4.5.8 General peptide synthesis protocol 180
4.5.9 Peptide characterisation 181
4.5.10 Cell culture 186

4.5.11 Flow cytometry 186
4.5.12 Cell viability alamarBlue assay 186
4.5.13 Confocal microscopy 186
Chapter 5: Conclusions and future directions 188
5.1 Conclusions and future directions of the project 189
Chapter 6: References 193
Appendix 206

Acknowledgements

First and foremost, I would like to thank Prof. Glenn Burley for giving me the opportunity to carry out a PhD in his research group and his continued support throughout the PhD. I would also like to thank Dr. Katherine Jones for her supervision, input into the ongoing direction of the project and many useful discussions. Furthermore, I would like to thank Louise Young, Stuart Woods and in particular Aisya Gany for training and guidance in cell culture and various biological aspects of this project. I would also thank Allan Watson for useful input and discussions regarding the Chan-Lam aspect of this project.

I would like to thank the Burley group at the University of Strathclyde for their guidance and support for the duration of my project. In particular I would like to thank all current members of the Burley group for creating an enjoyable and supportive working environment.

Specifically, I would like to thank previous group member Steven Pauff for his help and guidance during my first two years in the group. I would also like to thank Fergus McWhinnie for his continued help throughout and for proof reading the various drafts of this thesis many times. I would also thank Emma Campbell for her support and putting up with sharing a lab with me for the final years of my time at Strathclyde. In addition, I'd like to thank Fred Peschke for keeping the lab a 'paradise', Imma Abate for appreciating the difference between shortbread and longbread and Jamie Withers for providing chorizo stuffed olives. I would also like to thank Iain McKean for being a fantastic friend throughout my time in the group and of course for taking me on as an honorary Jambo! I would also like to extend my gratitude to Andrea Taladriz-Sender for being an amazing friend and for going above and beyond to help me both in and out of the lab throughout my time in Glasgow.

Finally, thank you to my family for their help and support through the past four years.

Abbreviations

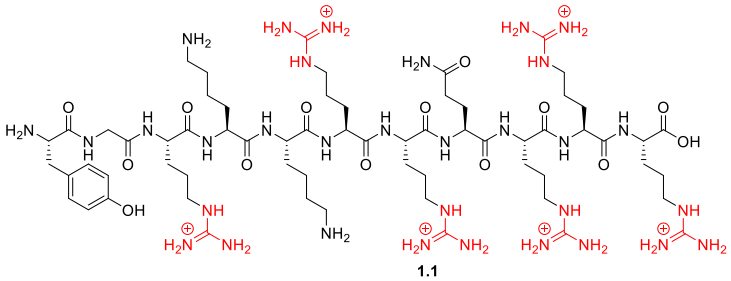
Ac_2O	Acetic anhydride
AcOH	Acetic acid
ADP	Adenosine diphosphate
Akt	Protein Kinase B
API	Active pharmaceutical ingredient
ATP	Adenosine triphosphate
BINAP	2,2′-bis(diphenylphosphino)-1,1′-binaphthyl
Boc	*tert*-butyloxycarbonyl
Boc_2O	Di-*tert*-butyl dicarbonate
BOP	Benzotriazol-1-yloxytris(dimethylamino)phosphonium
BSA	Bovine serum albumin
CaCo-2	Cancer coli-2
Calc.	Calculated
Cbz	Carboxybenzyl
CHO	Chinese hamster ovary cells
CME	Clathrin mediated endocytosis
CPP	Cell penetrating peptide
CPZ	Chlorpromazine
Dap	2,3,-Diaminopropionic acid
DCC	*N*,*N*′-Dicyclohexylcarbodiimide
DCE	Dichloroethane
DCM	Dichloromethane
DIC	*N*,*N*′-Diisopropylcarbodiimide
DIPEA	Diisopropylethylamine
DMD	Duchenne muscular dystrophy
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPPF	1,1'-Bis(diphenylphosphino)ferrocene
DtBPF	1,1'-Bis(di-*tert*-butylphosphino)ferrocene
EDC	1-Ethyl-3(3-dimethylaminopropyl)carbodiimide
EPR	Electron paramagnetic resonance
ESI	Electrospray ionisation
FBS	Foetal bovine serum
FITC	Fluorescein isothiocyante

Fmoc	Fluorenylmethyloxycarbonyl
FSC	Forward scattered
GAG	Glycosyaminoglycan
GFP	Green fluorescent protein
GLUT1	Glucose Transporter 1
GTP	Guanosine triphosphate
h	hours
H2	Histamine H2 receptor
H358	Human Caucasian bronchioalveolar carcinoma
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxide
HBTU	2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium
HCTU	O-(1*H*-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium
HEK293T	Human embryonic kidney cells
HeLa	Helacyton gartleri cells
HIV	Human immunodeficiency virus
HMPA	Hexamethylphosphoramide
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	Hydroxybenzotriazole
HPLC	High pressure liquid chromatography
HRMS	High resolution mass spectrometry
HS	Heperan Sulphate
IR	Infra-red
JNK	c-Jun-*N*-terminal kinase
LCMS	Liquid chromatography mass spectrometry
LDL	Low density lipoprotein
LiHMDS	Lithium hexamethyldisilazide
LogP	Partition coefficient logarithm
LTQ	Linear trap quadropole
MALDI-TOF	Matrix-assisted laser desorption/ionization time of flight
Mbs	*p*-Methoxybenzenesulfonyl
MCF-7	Michigan cancer Foundation-**7**
MDCK	Madin-Darby Canine Kidney
MFI	Median fluorescence intensity
min	minutes
Mtr	4-Methoxy-2,3,6-trimethylphenylsulfonyl
Mts	Mesityl-2-sulfonyl
Mtt	4-Methyltrityl

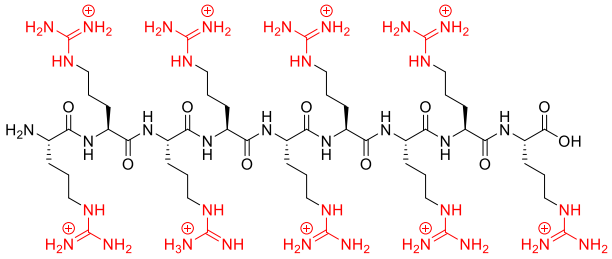
NADP	Nicotinamide adenine dinucleotide phosphate
nm	Nanometres
NMR	Nuclear magnetic resonance
Nys	Nystatin
PAMPA	Parallel artificial membrane permeability assay
Pbf	2,2,4,6,7-Pentamethyl-2,3-dihydrobenzofuran-5-sulfonyl
PBS	Phosphate buffered saline solution
PEG	Polyethylene glycol
P_i	Inorganic phosphate
PI3K	Phosphoinositide 3-kinase
PIP-2	Phosphatidylinositol 4, 5-biphosphate
PKB	Protein kinase B
PLC	Phospholipase C
PMB	*p*-Methoxybenzyl
Pmc	2,2,5,7,8-Pentamethylchroman-6-sulfonyl
PMO	Phophoramidate morpholino oligonucleotide
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
PyAOP	7-Azabenzotriazol-1-yloxy)tripyrrolidinophosphonium
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium
RP	Reverse phase
SAR	Structure activity relationship
SiO_2	Silicon dioxide
S_NAr	Nucleophilic aromatic substitution
SPPS	Solid phase peptide synthesis
SSC	Side scattered
TAMRA	Tetramethylrhodamine
Tbb_B48	Trypanosoma brucei brucei B48 deficient
Tbb_wt	Trypanosoma brucei brucei wild type
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TIPS	Triisopropylsilane
TLC	Thin layer chromatography
Trt	Trityl
U2OS	Human osteosarcoma
UV-vis	Ultra-violet visible
VWD	Variable wavelength detector
Wort	Wortmannin

Abstract

Cell penetrating peptides (CPPs) are short peptide sequences, usually between five and thirty amino acids long, that are able to traverse cell membranes. Since they were first discovered in 1989, CPPs have been widely used to deliver various molecular cargoes into cells, from small molecule therapeutics through to larger biologics such as oligonucleotides and proteins. A molecular hallmark of most CPPs is the incorporation of multiple arginine residues in their primary sequence (*e.g.* **1.1** and **1.2**).



1.1



1.2

Although there has been extensive research and optimisation of arginine-rich CPPs, there remain multiple drawbacks in their utility as drug delivery agents. The most significant and limiting is their the toxicity *in vivo*, associated with the high basicity gained from multiple arginine residues within the primary sequence of these peptides. Therefore, the development of guanidine functional group analogues to mimic the arginine side-chain, while both maintaining the hydrogen bond profile and mitigating the associated toxicity could provide a new generation of delivery vehicles for valuable molecular payloads.

This thesis describes the development of robust synthetic routes towards the preparation and validation of amidines as bio-isosteres of the guanidinium functional group present in arginine.

Chapter one introduces the importance of the guanidinium group in the structure and function of commonly used CPPs. It will also address how they interact with the cell membrane architecture so as to be internalised into the cell.

Chapter two describes the design rationale and synthesis of CPPs containing minimal arginine residues and compares their cellular permeability with common arginine-rich CPPs. The mechanism of cell uptake, viability, compartmentalisation and important regions of these peptides is also evaluated.

Analysis of cell uptake of a suite of CPPs demonstrated that certain low arginine-containing peptides are more cell permeable than arginine-rich CPPs.

Chapter three describes the development of mild Chan-Evans-Lam methodology to directly arylate amidines. The reaction mechanism was investigated and new off-cycle Cu-amidine species were identified. Mono *N*-arylation of amidines was demonstrated under mild conditions on wide variety of substrates including the synthesis of novel analogues of the clinically approved therapeutic pentamidine.

Chapter four describes the synthesis of amidine containing amino acids building blocks for solid phase synthesis and the incorporation of arginine mimetics into CPPs. The toxicity profile, cell uptake and subcellular localisation reveal replacing arginine with amidines enhances uptake of peptide within cells without any adverse toxic effects or perturbation of subcellular localisation.

Finally, Chapter five summarises the work of this thesis and suggests future directions for the development of less toxic and less basic CPPs.

Chapter 1: Introduction

1.1 The ever-present problem of cellular uptake

The ability of a therapeutic to efficiently cross the cell membrane is paramount if the biological target is intracellular. Understanding the processes that influence cellular uptake of compounds is an essential consideration in medicinal and biological chemistry. Designing molecules that can negate the barrier posed by the cell membrane will allow development of new drug delivery systems, provide cellular access for non-permeable therapeutics and provide a greater understanding of cell membrane chemical biology.

In recent years many macromolecule chemical probes and therapeutics have been developed that require cellular delivery.[1,2] Molecules such as peptides and oligonucleotides, which themselves are often not cell permeable, many peptides and oligonucleotides suffer from poor bioavailability and low enzymatic stability.[3,4] Creating species that can reliably deliver agents such as peptides and oligonucleotides would increase the scope of therapeutics and breath of probes that can be used to study biological processes.

1.2 The molecular architecture of the cell membrane

The cellular membrane of eukaryotic cells is a semi-permeable barrier that is essential for the structure and integrity of the cell, and ultimately controls what species can and cannot enter the cell. The 'fluid mosaic' model of the plasma membrane was first established in 1972 by Singer *et al.* and describes the cell membrane as a mixture of four fundamental components: phospholipids; proteins; glycoproteins/ glycolipids and cholesterols.[5] Each of these components are able to freely move throughout the plane of the membrane, interacting both with each other and the greater cellular environment.[5]

1.2.1 Phospholipids

Phospholipids are amphiphilic in nature with a polar, negatively charged phosphate group attached to two hydrophobic hydrocarbon chains *via* a glycerol moiety. The polar phosphate group is often attached to a positively charged choline or ethanolamine moiety and referred to collectively as the head with the hydrocarbon chains known as the tail. Individual lipid molecules self-assemble in a bilayer ensuring the hydrophilic heads are orientated toward the aqueous environment while the membrane is held together by hydrophobic interactions of the phospholipid tails (Fig. 1.1). This bilayer is the fundamental barrier of the cell and as such acts as the major obstacle to entering the cell for extra-cellular species.

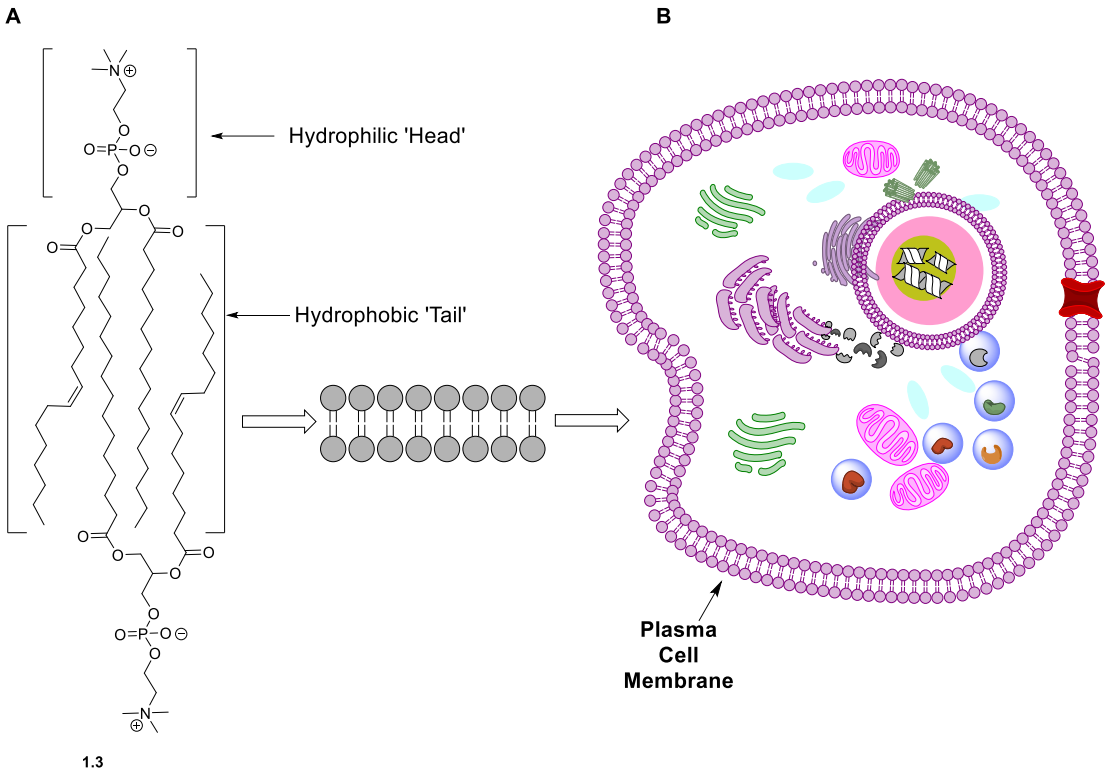


Figure 1.1. (A) Chemical structure a typical phospholipid (**1.3**) and how it is arranged in the lipid bilayer. (B) A figure of a typical eukaryote with its cell membrane.

1.2.2 Cell membrane proteins

Cell membrane proteins are involved in various processes, including acting as transporters for species to cross the membrane and as receptors for the activation of further cellular function. One of the most common examples of an integral transmembrane protein is the Na^+/K^+ ATPase enzyme that allows the cell to import K^+ ions while also exporting Na^+ across the membrane.[6] Peripheral proteins are not involved in the transport of molecules across the membrane but do play pivotal roles, such as triggering further cellular signalling pathways. An example is the Phospholipase C (PLC) enzyme family that degrade a relatively scarce phospholipid (PIP-2), where the subsequent species are later recycled into other important molecules such as extracellular signalling species (Fig. 1.2).[7]

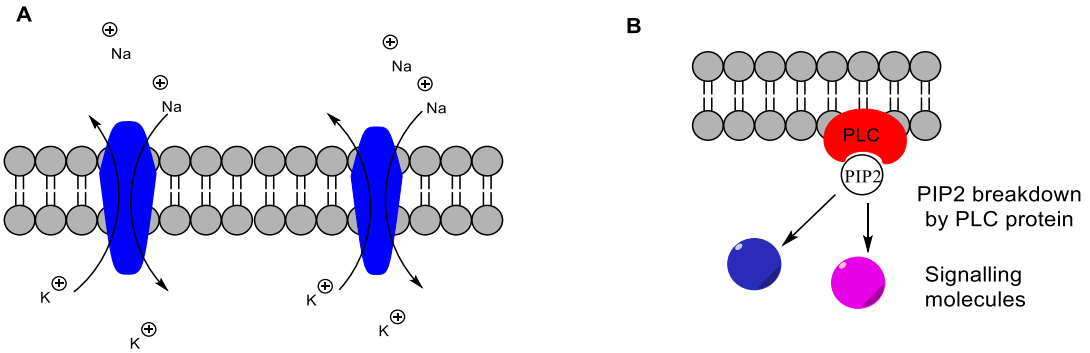


Figure 1.2. (A) Integral proteins such as Na^+/K^+ ATPase. (B) Peripheral PLC protein bound to the phospholipid heads

1.2.3 Glycoproteins and glycolipids

Glycoproteins and glycolipids are both incorporated at the cell surface and decorated with polysaccharides of up to 60 monomer units in length.[8] Both glycoproteins and glycolipids are predominantly concentrated at the extracellular portion of the plasma membrane with a smaller portion contained intracellularly. There are two major classes of cell surface oligosaccharides, *N*-glycosidic and *O*-glycosidic structures (Fig 1.3A). *N*-glycosidic structures are linked to the cell membrane protein through the amide of an asparagine residue, while the *O*-glycosidic units are linked *via* a hydroxyl group of a serine moiety.[9] An important class of polysaccharides at the cell surface are the glycosylaminoglycans (GAGs). GAGs are linear polysaccharides made up of repeating disaccharide units, where the disaccharide is composed of an amino sugar and a uronic acid, for example chondroitin sulphate (**1.4**, Fig. 1.3B).[10] GAGs can be either *N*-glycosidic or *O*-glycosidic depending on the specific polysaccharide and protein pairing. Surface attached GAGs also provide a site of recognition for exogenous proteins at the cell surface, generally binding through electrostatics involving the negatively charged moieties of the GAG and positively charged amino acid residues in the protein.[8]

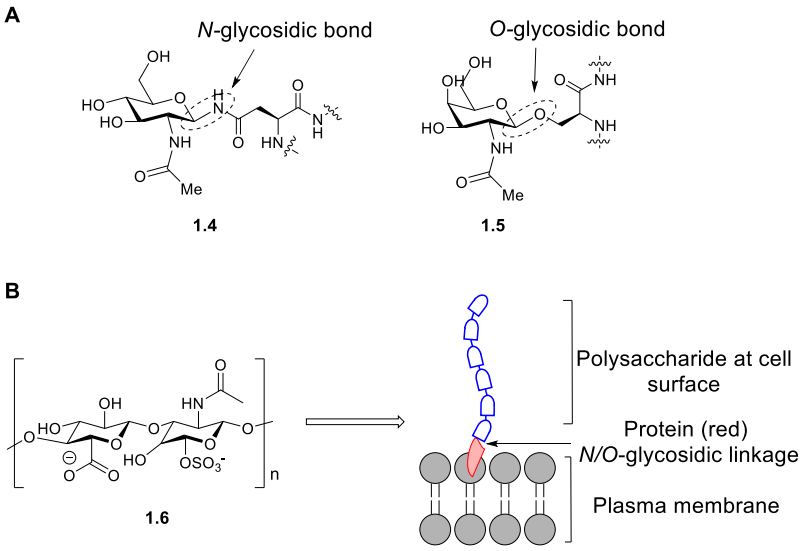


Figure 1.3. (A) *N*-glycosidic **1.4** and *O*-glycosidic **1.5** structures formed within glycoproteins. (B) The repeating disaccharide unit of chondroitin sulphate **1.6** and the representation of the polysaccharide as a glycoprotein

Similar to glycoproteins, glycolipids are categorised into two main groups; glycerol-glycolipids and sphingo-glycolipids. These names derive from to the oligosaccharide being attached to either a glycerol group or a sphingolipid (Fig. 1.4A). The hydrophobic nature of the lipid tails allows the glycolipid to be held at the membrane surface. Glycolipids are found almost exclusively on the cell surface/membranes of organelles and are very rarely found in the cytosol, with only ~5% of glycolipids not membrane bound.[11] Just like the carbohydrates of glycoproteins, the carbohydrates of glycolipids govern their main functions and often interact with each other at the cell surface. These carbohydrate-carbohydrate interactions can occur either within the same membrane or between two different cells. One such example of carbohydrate-carbohydrate interactions is the sphingo-glycolipid dependent

cell-to-cell adhesion process, which occurs due to trans-membrane electrostatic interactions between the two glycolipids and a cation such as Ca^{2+}(Fig. 1.4B).[12]

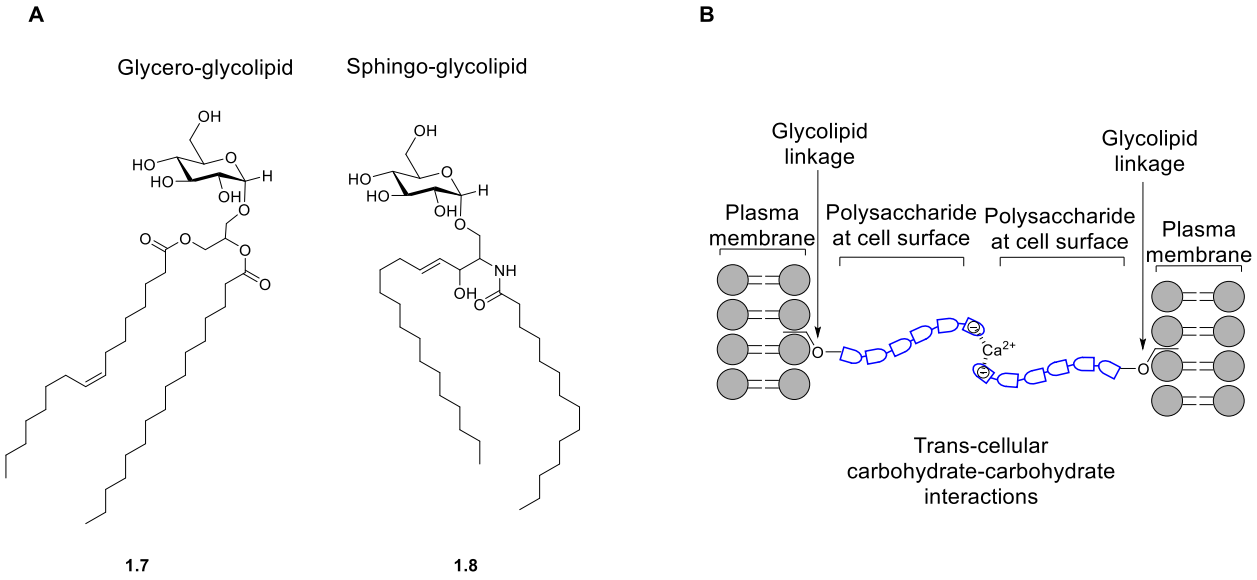


Figure 1.4. (A) Structures of glycero-glycolipid **1.7** and a sphingo-glycolipid **1.8**. (B) Trans-cellular carbohydrate-carbohydrate interactions between glycolipids.

1.2.4 Cholesterol

The final abundant species found in the plasma membrane are different cholesterol molecules. Cholesterol is very lipophilic due to its sterol ring arrangement and alkyl tail. It generally embeds into the lipidic inner region of the bilayer with its single hydroxy group interacting with the facial phosphates and solvent exposed region (Fig. 1.5A). The presence of cholesterol makes crossing the membrane more difficult for small hydrophilic molecules like ions or monosaccharides.[13] Cholesterol principally controls the fluidity of membrane structure. At elevated temperatures, cholesterol disrupts interactions between lipid chains, therefore making the membrane less fluid and therefore less permeable to external species. However, at lower temperatures this disruption maintains membrane fluidity, thus preventing the plasma membrane from freezing. The rigid structure of cholesterol also stabilises the *trans*-gauche isomerisations between the alkyl chains of phospholipids and further stabilises the bilayer (Fig. 1.5B).[13,14]

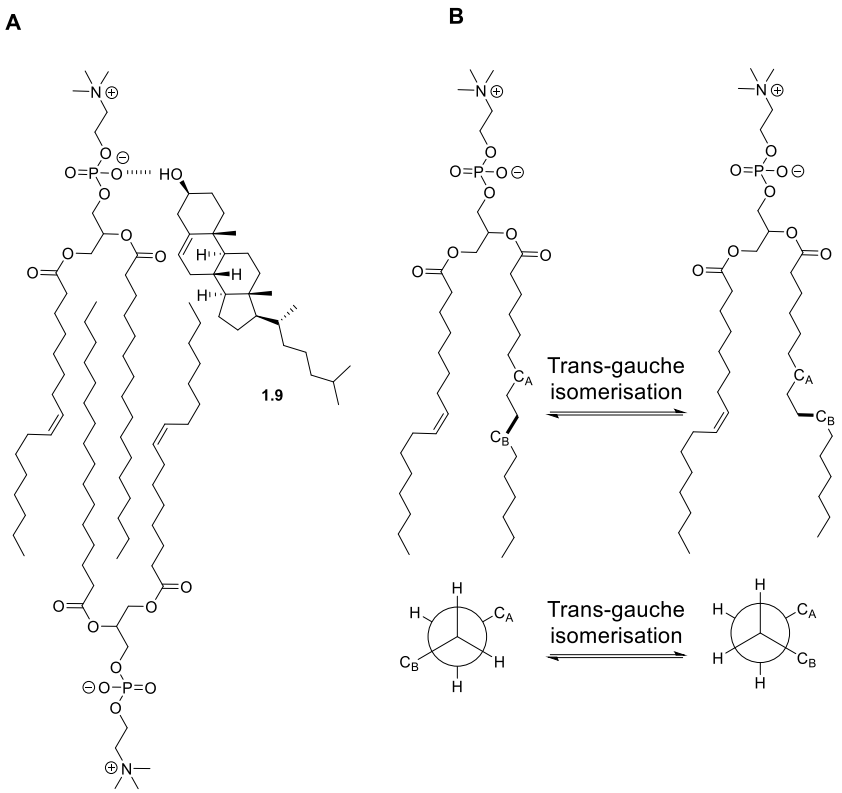


Figure 1.5. (A) Schematic representation of how a cholesterol molecule (**1.9**) is positioned within a phospholipid bilayer. (B) *Trans*-gauche isomerisations that are restricted by cholesterol species within the cell membrane.

1.3 Mechanisms of cell uptake

The cell membrane both protects and holds the cell together. The organelles of the cell are contained either within the cytoplasm or the nucleus, meaning that important biological matter must pass through the membrane to affect cellular function. There are four main ways that species cross the membrane in eukaryotic cells: passive diffusion, facilitated diffusion, active transport and endocytosis.

1.3.1 Passive and facilitated diffusion

Passive diffusion requires the movement of species from a high concentration environment to that of lower concentration across what is known as a concentration gradient.[15,16] Gases, such CO_2 or O_2, enter cells *via* this route-where the molecules move from the aqueous exterior of the cell to the hydrophobic inner region of the bilayer (Fig. 1.6A). The hydrophobicity of a specific molecule therefore directly determines its rate of diffusion into the cell.[15] This process only applies to small uncharged molecules that will not interact with the cell exterior to any great extent.[17]

Molecules that are significantly larger or more polar, require the aid of a transmembrane protein to enter the cytoplasm.[17] This process is called facilitated diffusion and, again, will move the species across the membrane down a concentration gradient in an entropically favourable manner. Channel proteins are the key proteins for facilitated diffusion as they form an aqueous pore through the bilayer in which high-polarity, water soluble species can enter the cell and avoid the lipophilic bilayer interior. An

example is the GLUT1 glucose transporter that allows the small polar molecule glucose to cross the cell membrane (Fig. 1.6B). The GLUT1 protein specifically fits the size and polarity of the glucose molecule stopping cell invasion from other small molecules.[18]

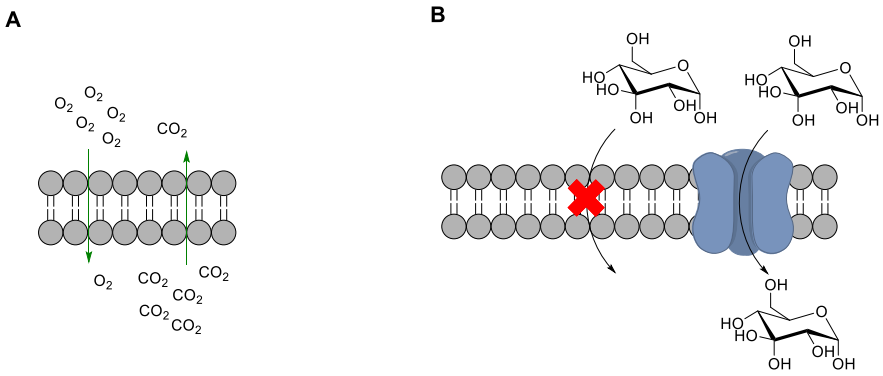


Figure 1.6. (A) Passive diffusion of O_2 and CO_2 across a plasma membrane down a concentration gradient. (B) Facilitated diffusion of glucose via a transmembrane protein.

1.3.2 Active transport mechanisms of cell uptake

Another main method of entry into cells is active transport and, like facilitated diffusion, requires the aid of transmembrane proteins. However, it requires the thermodynamically unfavourable movement against the concentration gradient. The most common way to overcome this energetic barrier is the hydrolysis of ATP to ADP and P_i. The transmembrane protein will contain an ATP binding site and where ATP is available will bind and release the energy required to drive species against the gradient. This is known as primary active transport as it is a single movement of a species *via* the same protein into the cell (Fig. 1.7A). Secondary active transport occurs when a gradient has been established by primary active transport and this causes a second species to be moved against its concentration gradient (Fig.1.7B).[19] A major class of active transport proteins are P-type ATPases, such as Na^+/K^+ ATPase and Ca^{2+} ATPase.[20] These proteins typically catalyse cation uptake and efflux *via* primary active transport.

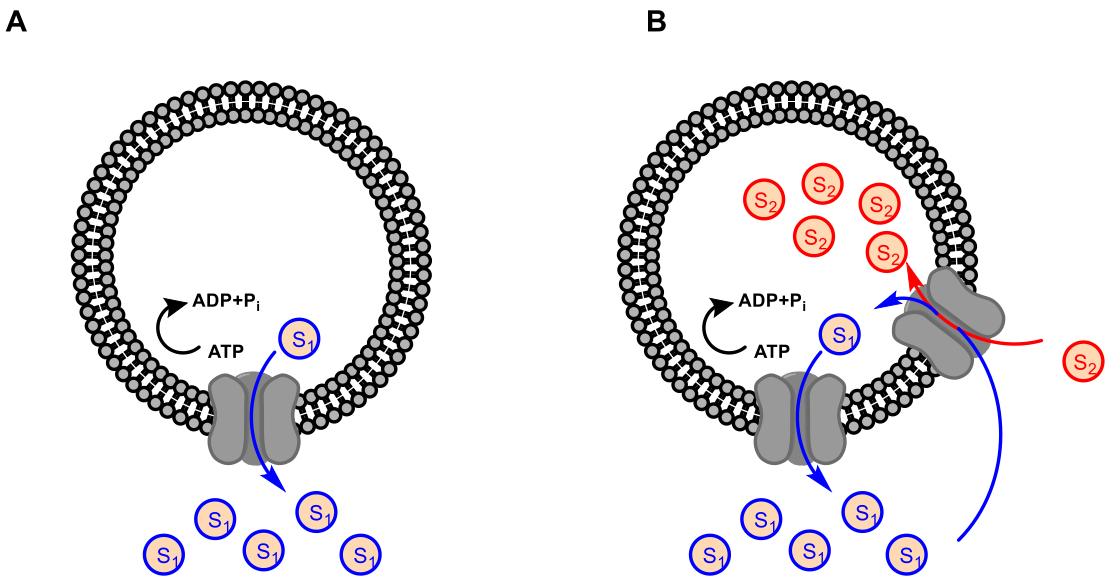


Figure 1.7. (A) Schematic representation of primary active transport. (B) Secondary active transport.

1.3.3 Endocytosis

The entry processes discussed above are primarily relevant for smaller molecules, while endocytotic uptake is usually concerned with macromolecules such as proteins and viruses. In endocytosis a molecule is engulfed and effectively packaged by the plasma membrane, forming a vesicle that then progresses through the bilayer to access the cytoplasm. The three main variants of endocytosis discussed here are clathrin-mediated endocytosis, endocytosis *via* caveolae and macropinocytosis.

Clathrin-mediated endocytosis (CME) involves pits in the cell membrane that are coated with the protein clathrin on the cells' interior. The process begins with a substrate binding to an external receptor protein at the cell surface (Fig. 1.8). This substrate-receptor complex coalesces with a clathrin coated pit. The pit encircles the substrate-receptor complex with the aid of the GTPase enzyme dynamin. The pit detaches from the plasma membrane to form a clathrin coated vesicle. The vesicle subsequently loses its clathrin coating and fuses with small vesicles known as early endosomes. These early endosomes typically have a pH of 6 compared to the pH 7 of extracellular fluid. This fusion causes the dissociation of receptor and substrate, with the receptor protein being recycled back to the plasma membrane. The early endosomes then mature into late endosomes *via* a carrier vesicle. Meanwhile, transporter vesicles from the Golgi deliver acid hydrolases to the late endosome. The late endosome further matures into a lysosome once it has received all degradation enzymes. Here the pH is further reduced from pH 6 to 4 with the aid of H^+ pumps on lysosome surfaces. The substrate may now be degraded by the hydrolases and its resulting building blocks released or, if resistant to the enzymes, the species may enter cytoplasm upon lysosomal breakdown.[21,22]A well-documented example of this form of endocytosis is the entry of low-density lipoproteins (LDL) into the cell and the subsequent release of cholesterol (Fig. 1.8).

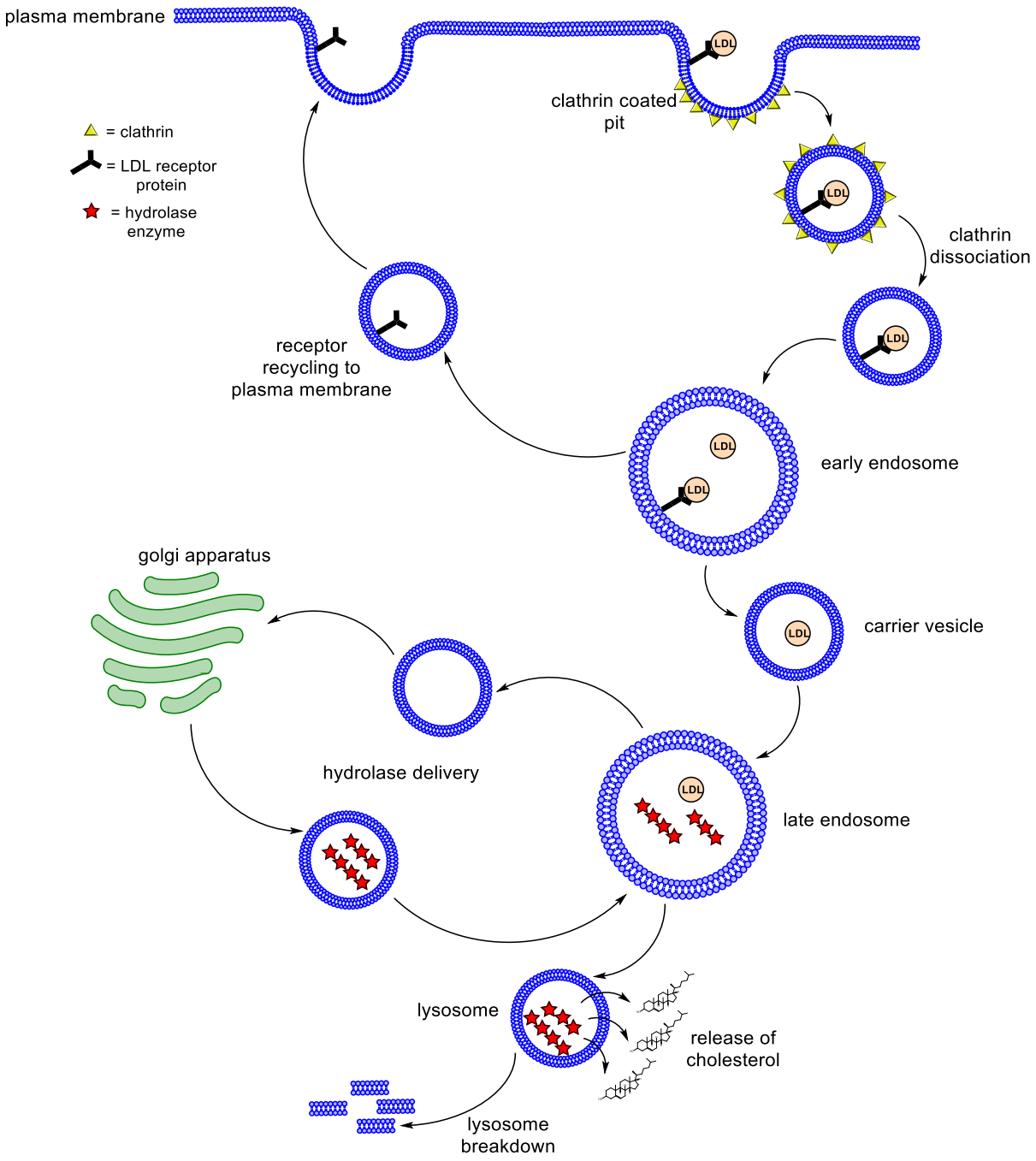


Figure 1.8. The clathrin-mediated endocytosis of LDL resulting in the release of cholesterol into the cytoplasm.

Another pathway of endocytosis is *via* caveolae. Caveolae are small enclaves at the cell membrane that are prearranged by transmembrane caveolin proteins and contain a high density of glycolipids and cholesterol.[23] The main difference to CME is that a specific cell surface receptor is not always required to induce the internalisation of a compound.[24] The exact process remains unclear but it is thought that upon binding to the caveolins a cascade of signalling events causes the recruitment of dynamin which facilitates the formation of a vesicle that can detach from the membrane, just as in CME (Fig. 1.8). Here the route is similar to that previously described with CME, where the caveolins are recycled back the membrane to reform caveolae. The vesicles then fuse to an early endosome (or a vesicle known as a

cavesome) that eventually matures to a lysosome which can release the remains of the substrate (Fig. 1.9).[23,24] Many larger species including viruses can enter cells *via* caveolin mediated endocytosis, one example is the simian virus 40 that upon release accumulates in the smooth endoplasmic reticulum.[25]

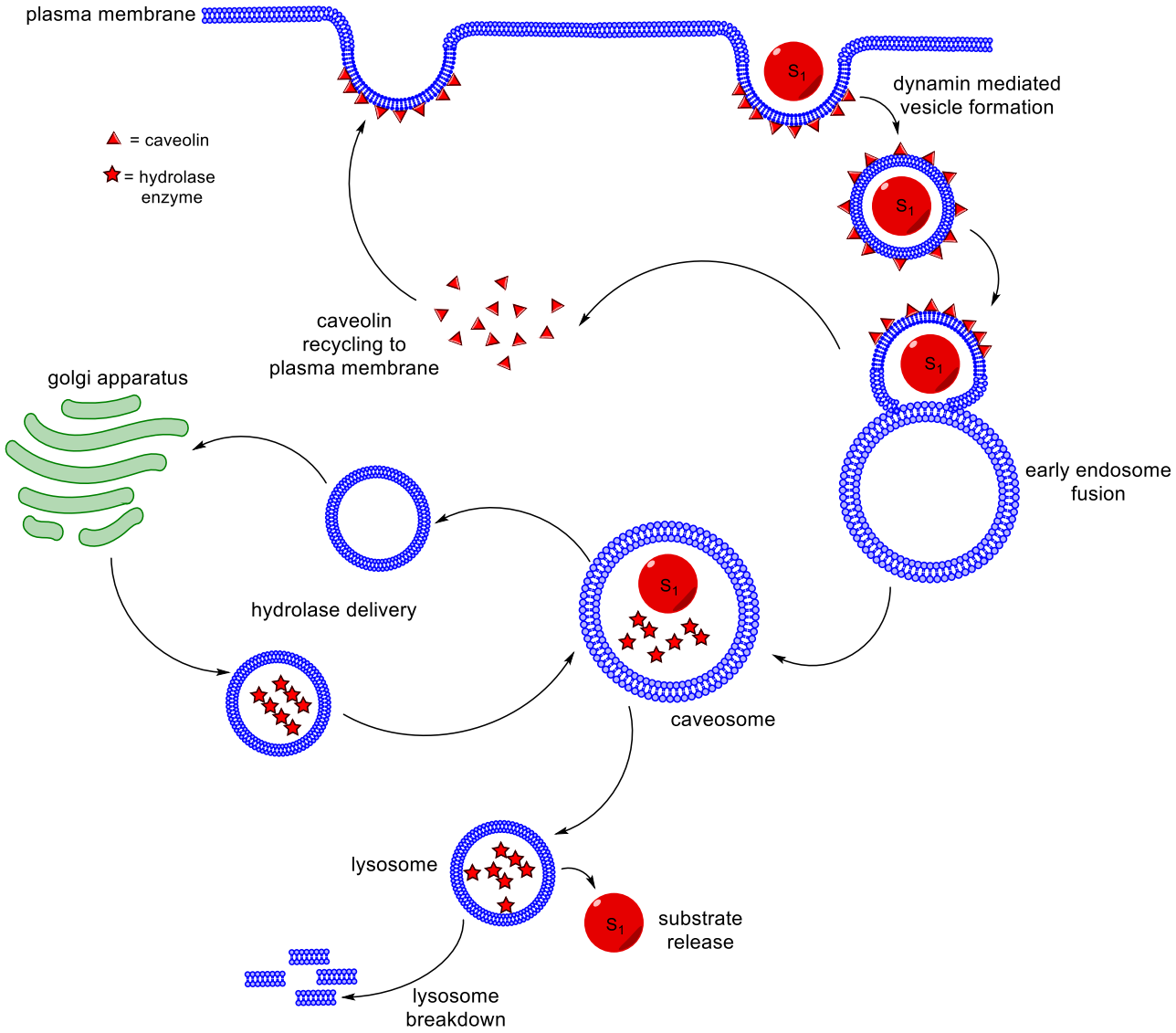


Figure 1.9. Caveolin mediated endocytotic uptake and release of substrate S_1.

Unlike either CME or caveolae mediated, macropinocytosis does not need any specific proteins or receptors to induce its internalisation (Fig. 1.10). It begins with a rearrangement of the cell membrane caused by cytoskeletal actin, this results in the formation of membrane 'ruffles'. These ruffles in the membrane are then able to fold back onto themselves to form a small vesicle that surrounds the substrate. The vesicle traps a lot of extracellular fluid and solute as well as any biomolecules, making this a non-specific mode of endocytosis. The vesicle then enters the cytoplasm and undergoes tubulation before fusing to early endosomes or being recycled back to the surface. Due to its non-specific nature, species taken up by macropinocytosis include bacteria such as Brucella and the HIV-1 virus.[26] The three modes of endocytosis vary as CME typically requires specific binding to a receptor, caveolin requires

species association to a caveolae whereas macropinocytosis can engulf a wider variety of species due to its non-specific nature.

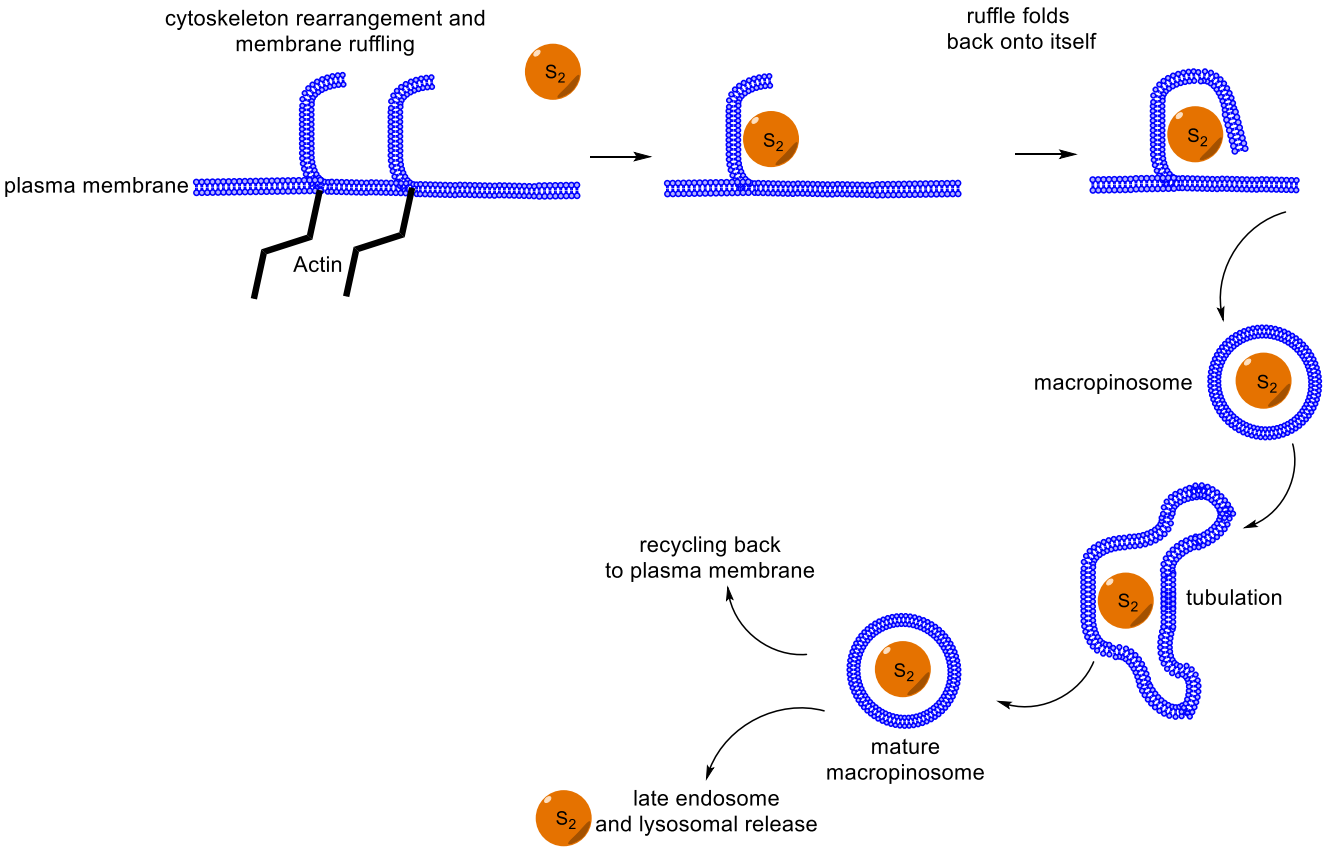


Figure 1.10. The process of macropinocytosis where the plasma membrane actin cytoskeleton cause 'ruffles' to induce internalisation of substrate S_2.

1.3.4 Permeability of synthetic small molecules

The processes discussed above demonstrate the variety of naturally occurring systems designed for cell uptake of different specific species. However, these processes must be adapted and effectively hijacked to ensure efficient incorporation and delivery of a synthetic species such as a therapeutic or probe. The majority of small molecule therapeutics enter cells *via* passive diffusion in order to directly access the cytoplasm without being enclosed in a vesicle or endosome structure. However, these molecules must be polar enough to be water soluble but lipophilic enough to pass through the lipid portion of the plasma membrane. The set of rules to overcome this fundamental dichotomy were designed by Christopher Lipinski in 1997 (Table 1.1).[27]

Table 1.1 Lipinski's Rules of 5 for cell permeability.[27]

Compound Characteristic	**Ideal value**
Molecular weight	≤ 500 g mol^{-1}
Lipophilicity	Log P ≤ 5
Total hydrogen-bond donors	≤ 5
Total hydrogen-bond acceptors	≤ 10

The values were obtained from the analysis of known orally bioavailable compounds and represent a generalised consensus. The rules also predict the cell permeability of a compound since to become orally bioavailable the compound must also be cell permeable. First, the molecular weight limit was proposed because as the weight of a molecule increases the ability of it to passively cross the membrane decreases. The lipophilicity is defined by a compounds partition coefficient (*P*) in polar and non-polar solvent. Molecules with a log*P* value ≤5 are expected to be lipophilic enough to cross the membrane without becoming entrapped within it. The lipid membrane is not totally inert however, with work by sanderson *et al.* demonstrating lipidation of drug like molecules.[28] Compounds which are notably hydrophilic however, those with very low log*P* values are deemed impermeable to cell membranes. The presence of ≤5 hydrogen bond donors and ≤10 hydrogen bond acceptors are deemed a maximum for a molecule to be able to shed its hydration shell to cross the bilayer and passively enter the cell. Multiple associated water molecules will cause the substrate to become trapped within the lipid portion of the cell membrane. Drugs such as the proton pump inhibitor omeprazole are Lipinski compliant and both cell permeable and orally bioavailable.[29] However, there are multiple examples of non-lipinski compliant therapeutics such as anti-cancer drug doxorubicin and cyclosporine A which is used as a therapeutic for various conditions including as an immunosuppressant. Despite being much larger molecules with a significant number of H-bond acceptors these compounds are both cell permeable with cyclosporin A also showing good oral bioavailability in some cases. This highlights that therapeutics can take on a wide variety of characteristics beyond Lipinski's rule of 5.[30,31]

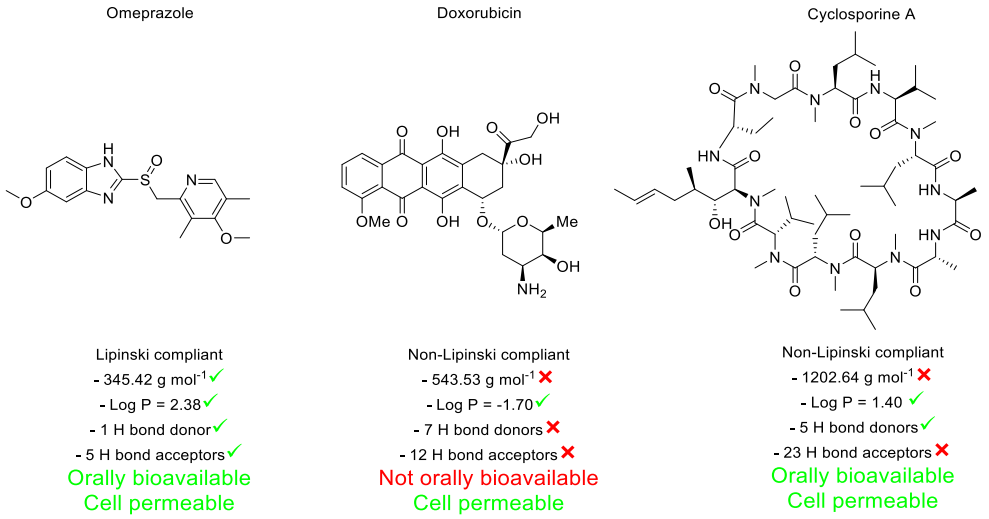


Figure 1.11. Examples of therapeutic agents that are both compliant and non-compliant with Lipinski's rule of 5.

1.4 Drug delivery platforms

1.4.1 Nanoparticles

Instead of relying on passive diffusion for the passage of small molecule therapeutics, synthetic systems can be used to actively deliver small molecules and macromolecules to cells. This strategy has been pursued with a variety of delivery systems and cell lines. One of the prominent species used for drug delivery are nanoparticles. Nanoparticles are structures defined as being a particle that has a diameter between 1 and 1000 nm.[32] The material that makes up a nanoparticle can vary widely from metal colloidal structures like gold to proteinaceous materials. Indeed, one of the first reported delivery method used human serum albumin microspheres to entrap and then release doxorubicin.[2,33] The natural product paclitaxel has been encapsulated in chitosan nanoparticles for anti-cancer applications that allow specific delivery of paclitaxel to certain cell types (Fig. 1.12).[34] The mechanism of uptake of these species is not defined by one process, with uptake occurring by both endocytosis and passive mechanisms.[35] An example where the uptake mechanism is determined by the structure of the nanoparticle is given by Jewell *et al.*[1] They showed gold nanoparticles could deliver large oligonucleotides into cells *via* a direct mechanism or by endocytosis by simply changing the ligands on the outside of the nanoparticle.

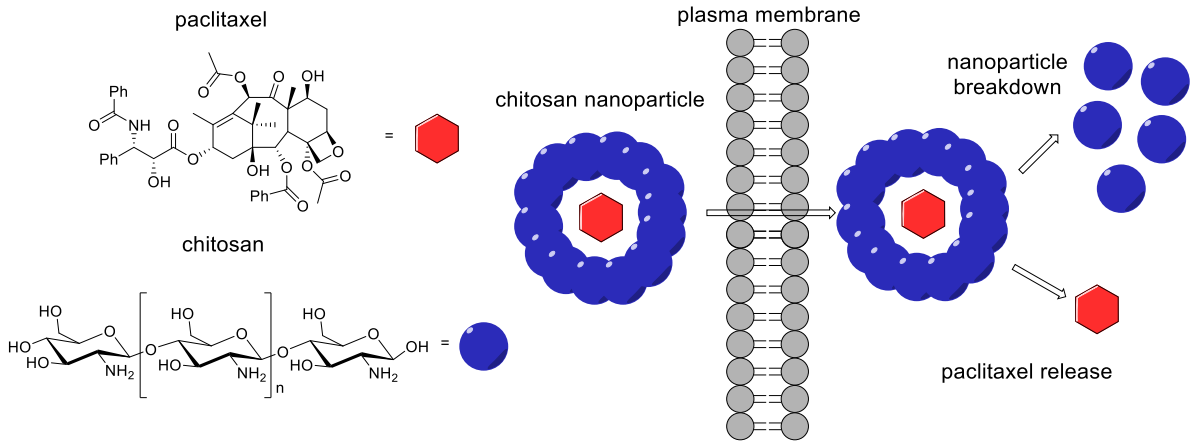


Figure 1.12. Nanoparticle encapsulation and drug delivery of paclitaxel using chitosan.

1.4.2 Liposomes

Liposomes are vesicular structures that are bilayers, and are typically composed of phospholipids, which will intrinsically self-assemble due to their amphipathic nature. Various lipids are used to create liposomes including natural phospholipids like phosphatidylserine **1.10** and synthetic lipids like distearoylphosphatidylcholine **1.11** (Fig. 1.13A).[36] One major example of a liposome delivery system is the clinically approved Doxil® (Fig. 1.13B). Doxil® encases the anti-cancer agent doxorubicin in a lipid sphere which can be targeted to specific cell types using peptides or polymers. Here the Doxil® lipids used are also decorated with polyethylene glycol (PEG) to reduce the metabolic clearance.[37,38] The cell uptake mechanism of liposomes has been observed to be predominantly by endocytotic pathways with the liposome binding to certain cell surface receptors that trigger endocytosis.[39]

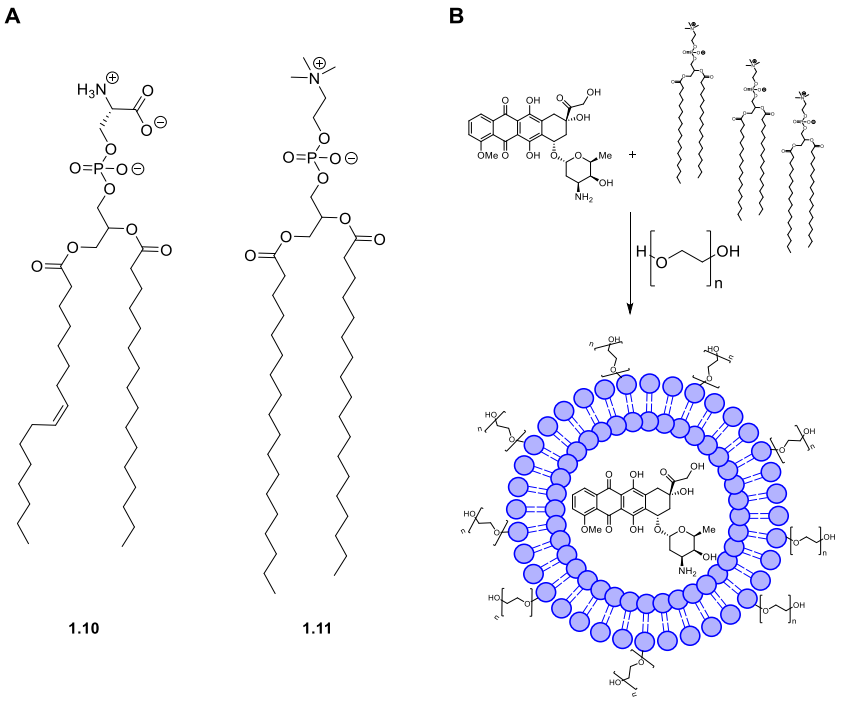


Figure 1.13. (A) Structure of common lipids used for liposome formation Phosphatidylserine **1.10** and distearoylphophatidylcholine **1.11**. (B) Representation of Doxil® with doxorubicin loaded into a liposome.

1.4.3 Cell penetrating peptides

CPPs are short peptides, typically between 5 to 30 amino acids in length, that are intrinsically highly cell permeable. CPPs have been demonstrated as useful delivery vehicles for various cargoes from small fluorophores to large oligonucleotides.[40] The first CPPs were highly cationic and derived from naturally occurring proteins such as Tat (48-60) from the transactivation domain of the HIV-1 virus and pVec from vascular endothelial cadherin protein (Fig. 1.14A & B).[41–43] The negative charges on the cell surface, such as sulphates and phosphates from glycoproteins and phospholipids, interact with the cationic peptides to initiate their internalisation which will be discussed in more detail in *section 1.6.2*.[40]

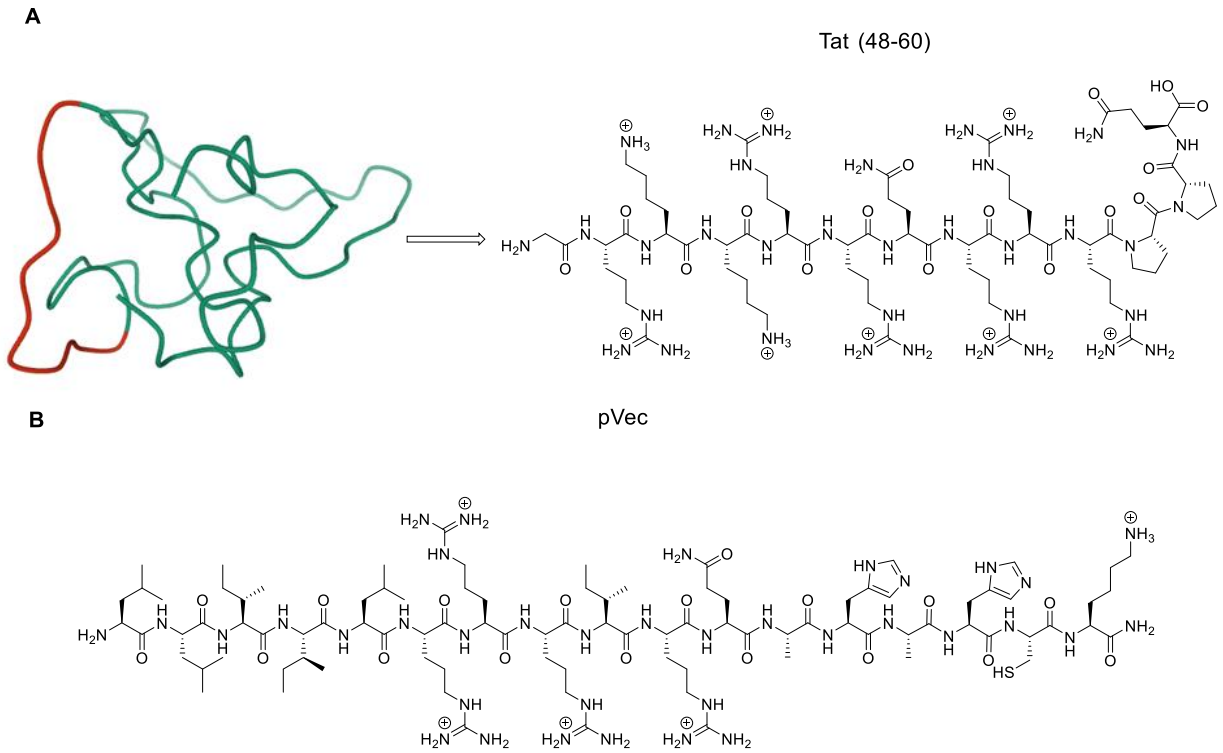


Figure 1.14. (A) Tat protein with CPP residues highlighted (red) and primary structure of the CPP section of Tat (PDB code: 1TIV_1) (48-60).[44] (B) Structure of pVec CPP.

1.5 Peptides and peptidomimetics as probes and therapeutics

As well as CPPs, peptides in general have been developed as therapeutic agents and probes. This is because peptides are surrogates for full-length proteins which are natural ligands for numerous biological targets in protein-protein interactions. Various therapeutics and biological probes have been designed based upon peptide chemistry. The variety of side-chains and intrinsic stereogenicity present in amino acids allows for the easy construction of diverse species for most biological targets. The major drawback of naturally occurring peptides is that they suffer from inherent poor physiochemical stability as well as a low plasma half-life.[3,45] To get around this, a number of chemical modifications have been developed to enhance their stability. Leuprorelin acetate **1.12** (Fig. 1.15) is one example of a medicinal peptide used to treat endometriosis, where the *N*-terminus has been modified to a lactam. Furthermore, the *C*-terminus has been changed to an ethylated amide and the stereocentre in leucine has been changed from an L- to a D-leucine to improve protease stability and oral bioavailability.[46]

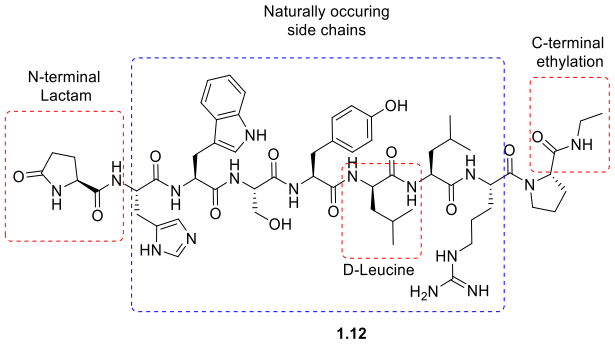


Figure 1.15. Peptide structure of leuprorelin acetate **1.12**, highlighting the 3 non-natural modifications.

These modified peptidic structures are generally loosely referred to as peptidomimetics. Peptidomimetics are peptides that mimic the interactions of a particular residue or section of a peptide but do not have the same chemical structure.[47] This term comes from the idea of bioisosterism, which is employed in small molecule drug discovery. Bioisosteres are substituents or groups with subtly different physiochemical properties that can produce the same biological response in a compound. Bioisosteres are designed to enhance the biophysical properties of a molecule without reducing the biological effect.[48] Common examples of these are the change from a hydrogen atom to a fluorine atom as the fluorine is similar in size to the proton but not susceptible to oxidation by enzymes like cytochrome P450.[49] An early example of using bioisosterism in drug development was the change from a CH_3 to a CF_3 in picornovirus inhibitors (Fig. 1.16A) which resulted in the enhanced metabolic stability.[50] Another bioisostere often employed is a tetrazole in the place of carboxylic acid. The tetrazole proton has been measured to have a similar pK_a to a carboxylic acid (5.08 tetrazole *vs.* 4.64 carboxylic acid).[51] For example, the angiotensin II antagonist losartan employs a tetrazole in a place of a carboxylic acid (Fig. 1.16B).[52]

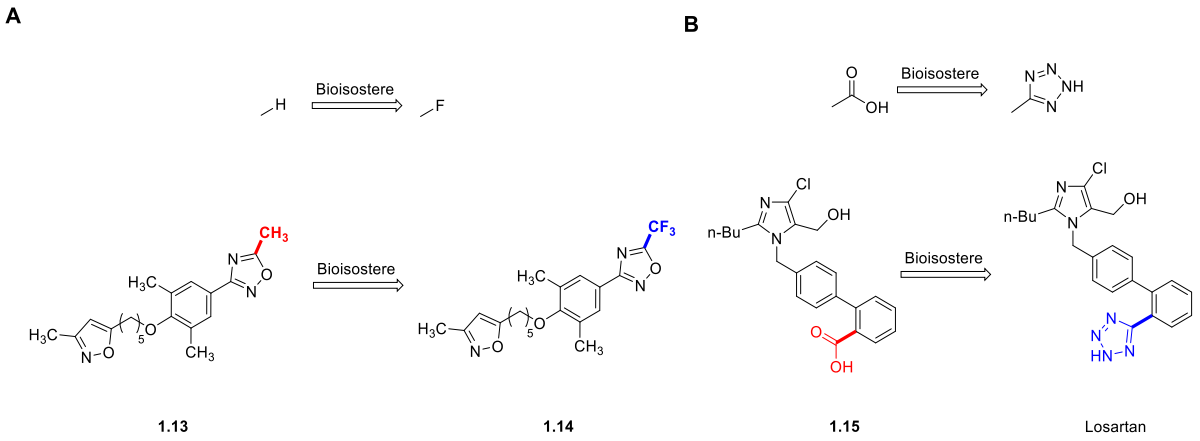


Figure 1.16. (A) A Fluorine as a bioisostere of hydrogen used in the picornovirus inhibitor **1.11**. (B) Tetrazole as a bioisostere of the carboxylic acid found in the angotensin II antagonist losartan.

The modifications in leuprorelin **1.12** are examples of specifically peptidic bioisosteres as each modification mimics the natural alternative. The *N*-terminal lactam and *C*-terminal ethyl amide mimic peptide bonds, while the D-leucine improves the protease stability of the peptide compared to the natural L-leucine.

To employ peptidomimetics the importance of each residue in the peptide must first be investigated. A typical method to probe the dependency of a primary amino acid sequence is known as alanine scanning. In this process each residue of the peptide is sequentially changed to alanine and the peptides function compared to the unaltered form.[53] This is the first step in defining a compounds full structure-activity relationship (SAR) in which the important residues for activity are located. Alanine is used as it contains a relatively small and unfunctionalized side group (methyl). It also does not incorporate any extra backbone flexibility unlike the addition of glycine.

Mimetics of the peptide backbone are commonly used to block the action of peptidases on peptides. This ensures the peptide has a long enough half-life to reach its biological target. There are many isosteres considered for the amide bond in peptides, including the use of simple esters, thioamides and fluoroalkenes.[54] These species have similar electronic configurations to amides and can mimic both resonance forms of the amide (Fig. 1.17A). A common set of backbone mimetics are α-alkylated amino acids, which include variations like dimethylated glycine (2-aminoisobutyric acid),[55] cycloalkane glycine derivatives,[56] α-methyl cysteine and α-methyl threonine.[57] Mimetics like these restrict the φ and ψ dihedral angles and therefore can constrain or disrupt any secondary structure of the peptide. For example cyclohexyl glycine has been used to control the helicity of some peptides (Fig. 1.17B).[56] The use of cycles in the backbone, like diketopiperazine **1.24**, has enhanced protease stability when incorporated into short peptides and mimics both the amide bonds and α-carbons.[58] Further mimetics to the backbone are *N*-methyl amides and aza-amino acids. *N*-methyl amides have been observed to enhance protease stability, cell permeability and bioavailability of peptides whilst maintaining the amide backbone structure.[59] Aza amino acids incorporate a nitrogen atom in place of the α-carbon, removing the stereocentre, but have also demonstrated a global enhancement in peptide stability whilst

maintaining biological activity.[60,61] Both strategies have been employed successfully to PKB/Akt inhibitors.[61]

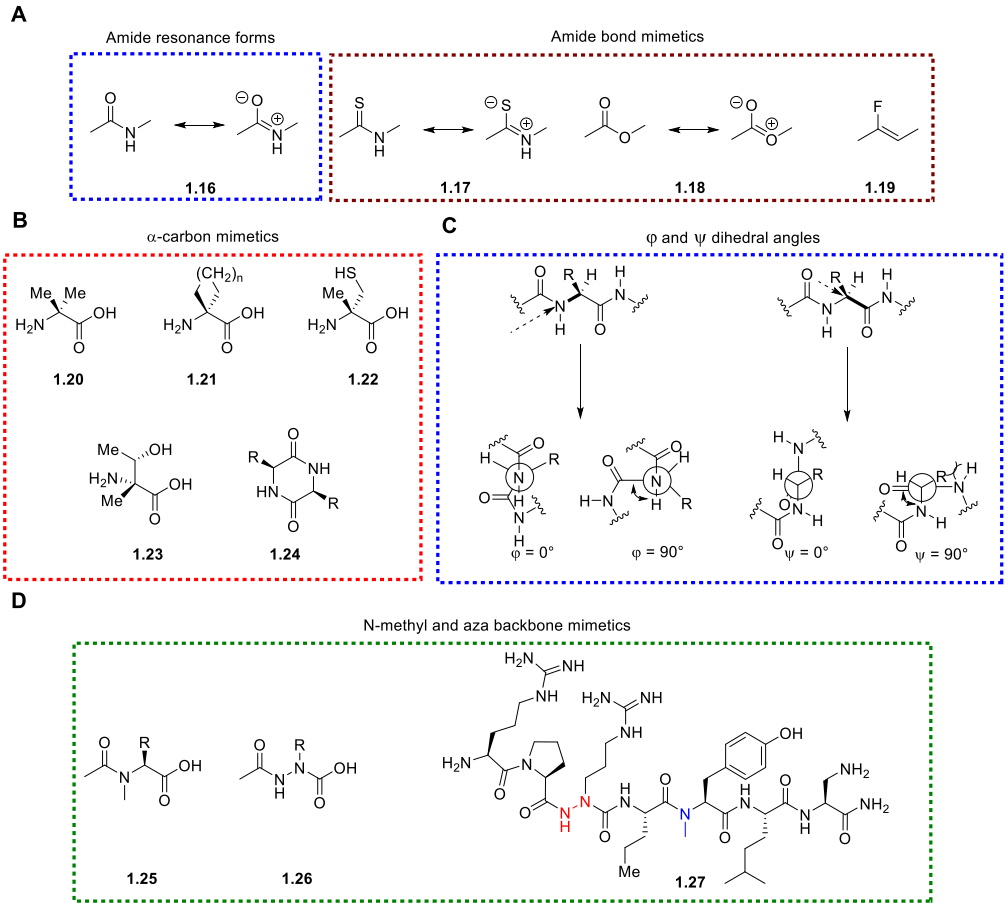


Figure 1.17. (A) Amide bond mimetics **1.16-1.19**. (B) α-carbon substitution mimetics **1.20-1.24**. (C) Dihedral angles and rotations within peptide backbones. (D) *N*-methyl and aza backbone mimetics **1.25-1.26** and PKB/Akt inhibitor **1.27**.

1.6 Structure and function of cell penetrating peptides

Peptidomimetic design often attempts to enhance cell permeability of the sequence as most peptides are not cell permeable whilst maintaining intrinsic biological activity. In contrast, CPPs aim to be predominantly cell penetrative without themselves possessing a therapeutic function. One of the first peptides that demonstrated cellular permeability was a derivative from the Tat protein, a protein synthesised by the human-immunodeficiency virus (HIV).[41] The highly basic 13 amino acid sequence of the Tat protein (Tat 48-60) was shown to be the key in the translocation of the whole 86 residue Tat protein.[42] Another peptide exhibiting a similar ability was the Antp peptide, a 16 amino acid peptide derived from the third helix of the antennapedia homodomain which is a transcription factor from the *drosphilia* genus.[62]

1.6.1 Early developments of CPPs

Since the determination of these CPPs several more have been designed and synthesised with the aim of understanding their capabilities to deliver permeability and ability to deliver species across cellular membranes. Initial studies showed that the Tat peptide was able to improve iron oxide nanoparticle uptake in mouse lymphocytes[63] and both Tat and Antp were shown to improve the delivery of oligonucleotides to HeLa cells.[64] Many other CPPs were developed on the back of the Tat/Antp success (Table 1.2) but the specific mode of entry of CPPs in general is still up for debate.

Table 1.2. A selection of synthetic and biologically isolated CPPs.

Name of CPP	**Source of CPP**	**Primary sequence**
Tat (48-60)	Tat protein	GRKKRRQRRRPPQ[42]
Antp	Antennapedia homodomain	RQIKIWGQNRRMKWKK[62]
MAP	Synthetic	KLALKLALKALKAALKLA[65]
R_9/R_{12}	Synthetic	R_9-R_{12}[66]
H5WYG	Synthetic	GLFHAIAHFIHGGWHGLIHGWYG[67]
PVEC	Cadherin	LLIILRRRIRKQAHAHSK[43]

1.6.2 Interactions of CPPs and cell surface species

Each of the peptides (Table 1.2) contains multiple Arg, Lys or His residues and it was reasonable to assume the highly positive and basic nature of these peptides facilitates their cellular permeability. The phospholipid bilayer is awash with negative charges from both phosphate heads and various polysaccharides. The first port of call for many of these peptides would be electrostatic interactions between the oppositely charged groups. The role of the positive charges was demonstrated by alanine scans of both the Tat 48-60 and Antp peptides.[68,69] Replacement of positive residues in both CPP sequnces by alanine resulted in a decrease in cellular uptake and so reinforcing the idea that electrostatic interactions are needed for cell uptake of CPPs.

The development of polyarginine CPPs came from these initial alanine scans into whole peptide charge but also from studies that showed Arg itself had greater penetrating power than lysine or histidine, hence why no polylysine or polyhistidine CPPs have been extensively pursued. This phenomenon was demonstrated in a study by Mitchell *et al.* where peptides of 6 arginine residues (either D- or L- isomers) or more were superior to the homologous lysine or histidine peptides in terms of cell uptake in Jurkat cells.[70] The mechanism for this increased uptake is attributed to the guanidinium group of the arginine side chain. The guanidinium group has the ability to form bidentate hydrogen bonds with phosphate groups at the cell surface, whereas lysine can only form a single hydrogen bond interaction (Fig. 1.18) meaning weaker bonding to the cell surface for lysine compared to arginine. These arginine-hydrogen

bond complexes can then facilitate diffusion of the peptide into the cell membrane.[71]

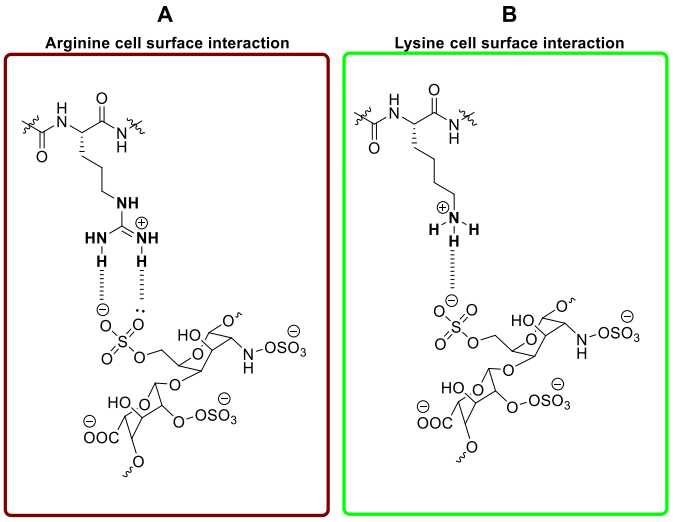


Figure 1.18. (A) Bidentate hydrogen bonding of arginine to a cell surface glycoprotein. (B) Single hydrogen bond of lysine to a cell surface glycoprotein.

1.6.3 low arginine content CPPs

Despite the apparent importance of arginine, there are still many different CPPs that are not dependent on arginine or indeed any positive charge for their permeability. A key example is the pVec peptide. Analysis of the structure-activity relationship of pVec showed that replacement of any of the first 5 amino acid residues (LLIIL) with alanine showed a distinct decrease in cellular uptake of around 50-75%.[72] Alterations to the arginine residues in fact showed an average increase in cellular uptake and clearly showed the importance of hydrophobic residues, and in particular leucine and isoleucine in the pVec CPP.

Along with this finding, the tryptophan residue has shown to impart good cell penetration for certain CPPs, owing to its relatively high aromaticity. Tryptophan has been demonstrated to aid the uptake of arginine-rich CPPs. Tryptophan residues were added to the octaarginine CPP with the CPP RW_{mix} (sequence = RWRRWRRWRRWR) and RWR (sequence = RRRRWWWWRRRR) both having the greatest uptake in this study.[73] The rationale behind tryptophans effect on uptake is due to its favourable interactions with glycosylaminoglycans (GAGs) at the cell surface. The ability of tryptophan to form π-cation interactions with arginine residues in CPPs can increase the pK_a of the guanidinium and lead to a more basic guanidinium cation (Fig. 1.19A). This leads to a greater interaction with negative sulphates of the GAGs. The formation of β strands with tryptophan rich CPPs, when in complexation with GAGs, has also been suggested to lead to a more stable CPP-cell surface binding conformation.[74]

A further characteristic of CPPs that has shown direct effects on cell permeability is the spacing between positive residues in the peptide backbone. Peptides developed with RWR and RW_{mix} motifs demonstrated this, but it has also been established with the so called $(RXR)_4$ peptides

(X = 6-amino hexanoic acid). Rothbard *et al.* displayed the effects of glycine, 4-aminobutyric acid and 6-amino hexanoic acid between arginine residues (Fig. 1.19B).[75] They confirmed that the longer the carbon link between each arginine, the higher is the cellular uptake. In addition, the RXR peptide had a more efficient uptake than the R_{10} polyarginine. The X carbon spacer was determined to be the maximum distance for spacing as an increase to 8-amino caprylic acid showed a decrease in cellular uptake. The rationale behind these results is first, that the greater flexibility in the peptide backbone allows for less conformationally restrained interactions between positively charged arginine side chains and negative groups on the cell surface. Secondly, is the idea that the non-natural spacing residues make degradation by proteases more difficult, though the tests from Rothbard *et al.* only contained trace amounts of serum in their assays. This spacing has been employed in CPPs such as the Pip CPPs (Fig. 1.19C).[76,77]

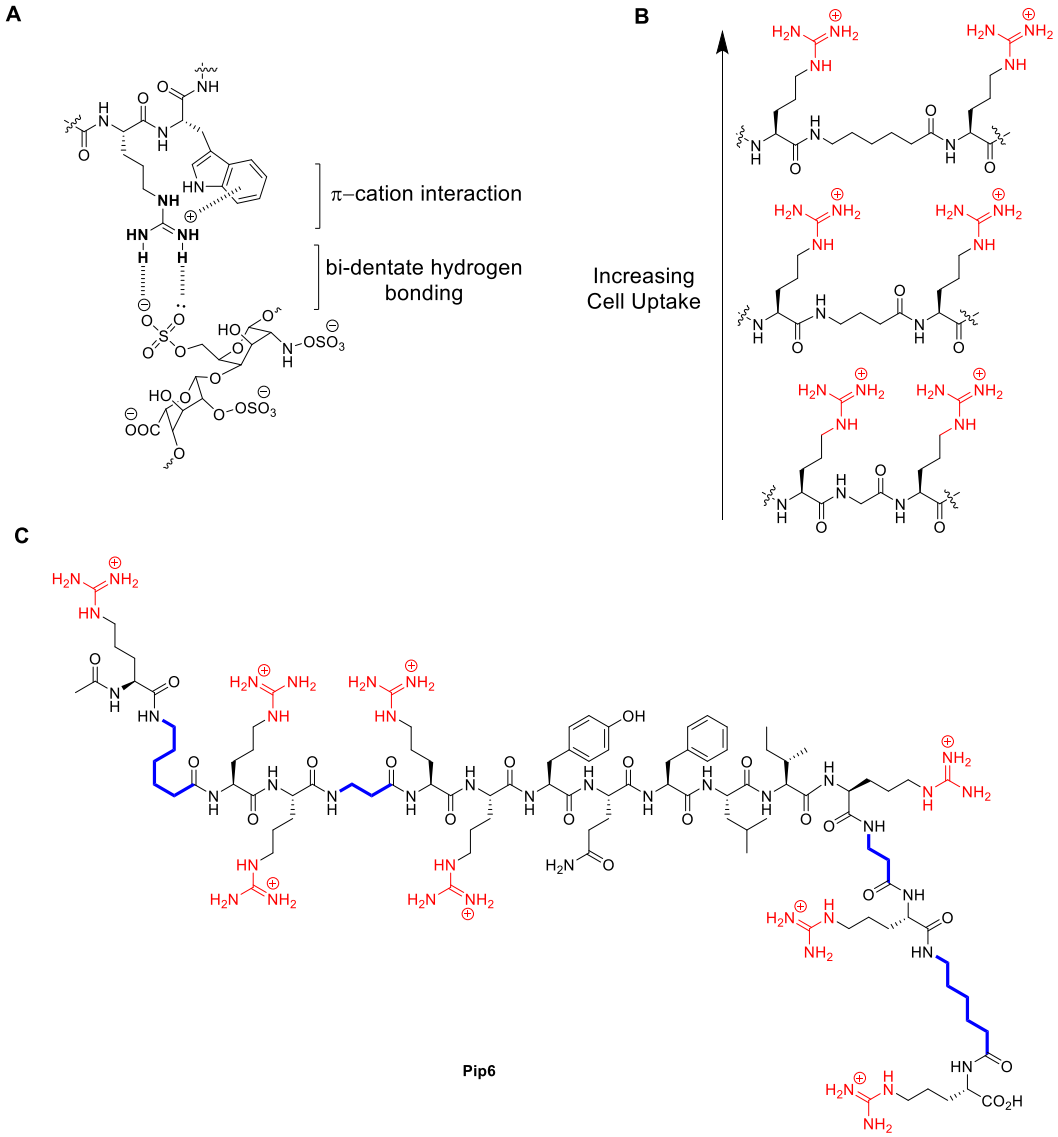


Figure 1.19. (A) π – cation interaction between tryptophan and arginine residues at the cell surface. (B) Increasing distance between arginine residues enhances cell uptake. (C) Pip6 peptide with β-alanine and hexanoic acid spacers highlighted in blue.

The Pip CPPs have been further developed to include specific β-alanine spacers and used for cellular delivery of phosphorodiamidate morpholino oligonucleotides (PMOs) to treat Duchenne muscular dystrophy (DMD). These studies by Wood *et al.* and Gait *et al.* demonstrate the ability of CPPs to carry large molecular weight cargoes across membranes with the PMO used having a molecular weight of 8582 g mol^{-1}, four times the molecular weight of the CPP.[76,78] Other relatively hydrophobic peptides have shown great promise in cell uptake. For example, TP-1 and TP-2 (Fig. 1.20) were synthesised and tested by *Marks et al.*[79] These two peptides showed the capability to directly translocate across CHO cell membranes at room temperature. The R_9 peptide used in this study showed an inability for direct translocation at room temperature with only an ability for uptake by endocytosis, demonstrating the selective nature of uptake of CPPs to specific cell types. A truncated form of TP-1 **1.28** (Fig. 1.20), was later developed and was shown to be as highly permeable to HeLa and HEK293T cell lines as R_9.[80] Interestingly, the mode of entry was determined to be direct translocation in HEK cells but endocytosis in HeLa cells. This observation shows that the uptake mechanism is dependent on the specific cell type, and therefore the cellular surface as well as peptide structure. The high uptake by any mechanism is a promising finding, however no structural relationship has been established for the high permeability of this peptide. As well as this a truncated form of TP-2 **1.29** (Fig. 1.20) has shown cell uptake in CHO cells where aggregation at the cell surface was the key initiating factor for direct translocation.[81]

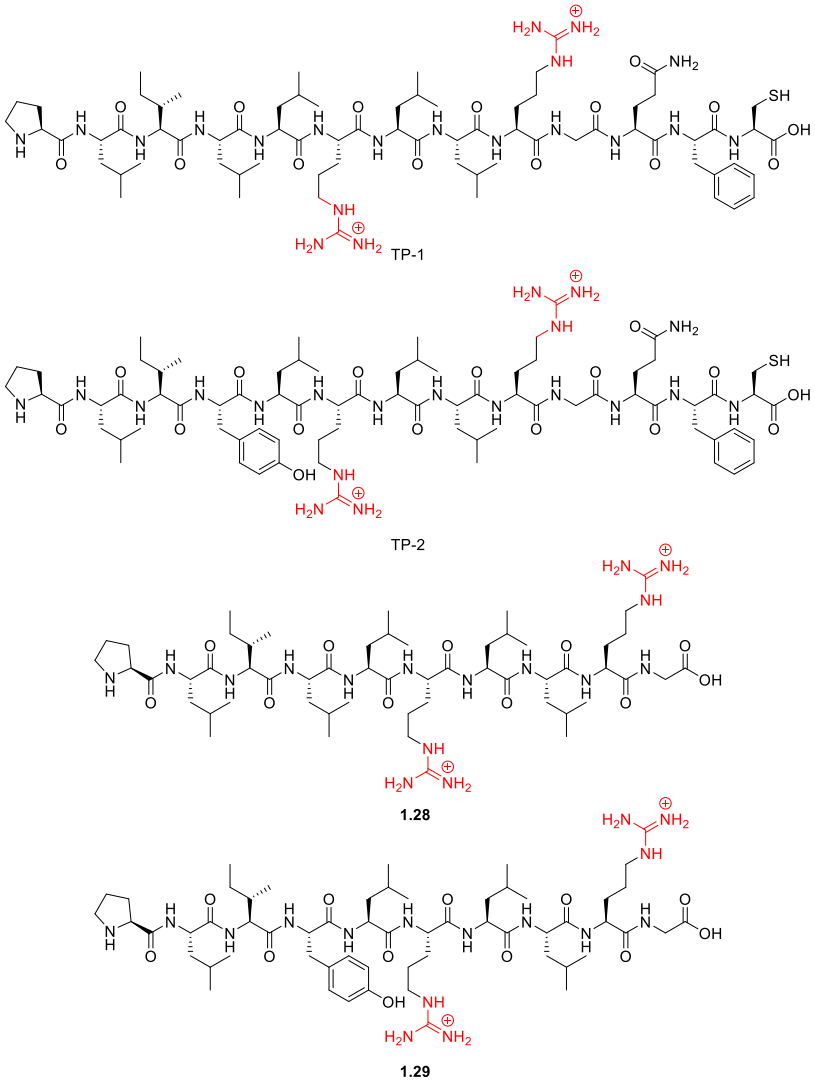


Figure 1.20. Structures of low arginine-containing CPPs TP-1, TP-2, **1.28** and **1.29**

1.6.4 Methods of cell uptake for CPPs

There is clearly a high variability in the structure activity relationships of CPPs, though many do rely on arginine as their key residue for permeability. The other key factor to consider is the effect of the specific cell surface and how different cell lines exhibit different entry routes to either the same or very closely related CPPs (Table 1.3).

Table 1.3. Selected CPPs and their differing entry routes into cells of different cell lines

CPP	Primary sequence	Cell line	Uptake mechanism
TP-1	PLILLRLLRGQFC	CHO	Direct translocation[79]
TP-2	PLIYLRLLRGQFC	CHO	Direct translocation[79]
1.28	PLILLRLLRG	HeLa	Endocytosis[80]
1.28	PLILLRLLRG	HEK	Direct translocation[80]
1.29	PLIYLRLLRG	CHO	Direct translocation[81]
pVEC	LLIILRRRIRKQAHAHSK	BMC	Endocytosis[72]
Tat	GRKKRRQRRRPPQ	HeLa	Direct translocation[42]
ANTP	RQIKIWGQNRRMKWKK	HeLa	Direct translocation[41]

Direct translocation is similar to passive diffusion as the molecule is delivered directly to the cytoplasm (unlike endocytosis pathways) but the specific processes that occur are still disputed and there are three hypotheses for how direct translocation takes place.[40,42] The first possible mechanism uses inverted micelle formation, where the positively charged CPP binds to the cell surface and the cell membrane forms a vesicle surrounding the peptide before the membrane reorganises itself and releases the peptide into the cytosol. The second hypothesis is centred around the formation of a pore in the membrane. The peptide transiently inserts into the membrane forming a pore by which the peptide can pass through and enter the cytosol. The final hypothesis focusses on a process known as adaptive translocation where the peptide remains bound to phospholipids at the phosphate heads and the hydrophobic tails allow for movement through the lipophilic centre of the membrane (Fig. 1.21). The other defining entry method is endocytosis by which the formation of a vesicle at the cell surface surrounds the peptides and carries them past the membrane. Unlike inverted micelle formation, this vesicle continues into the cytoplasm before the eventual endosomal release (*section 1.3.3*).

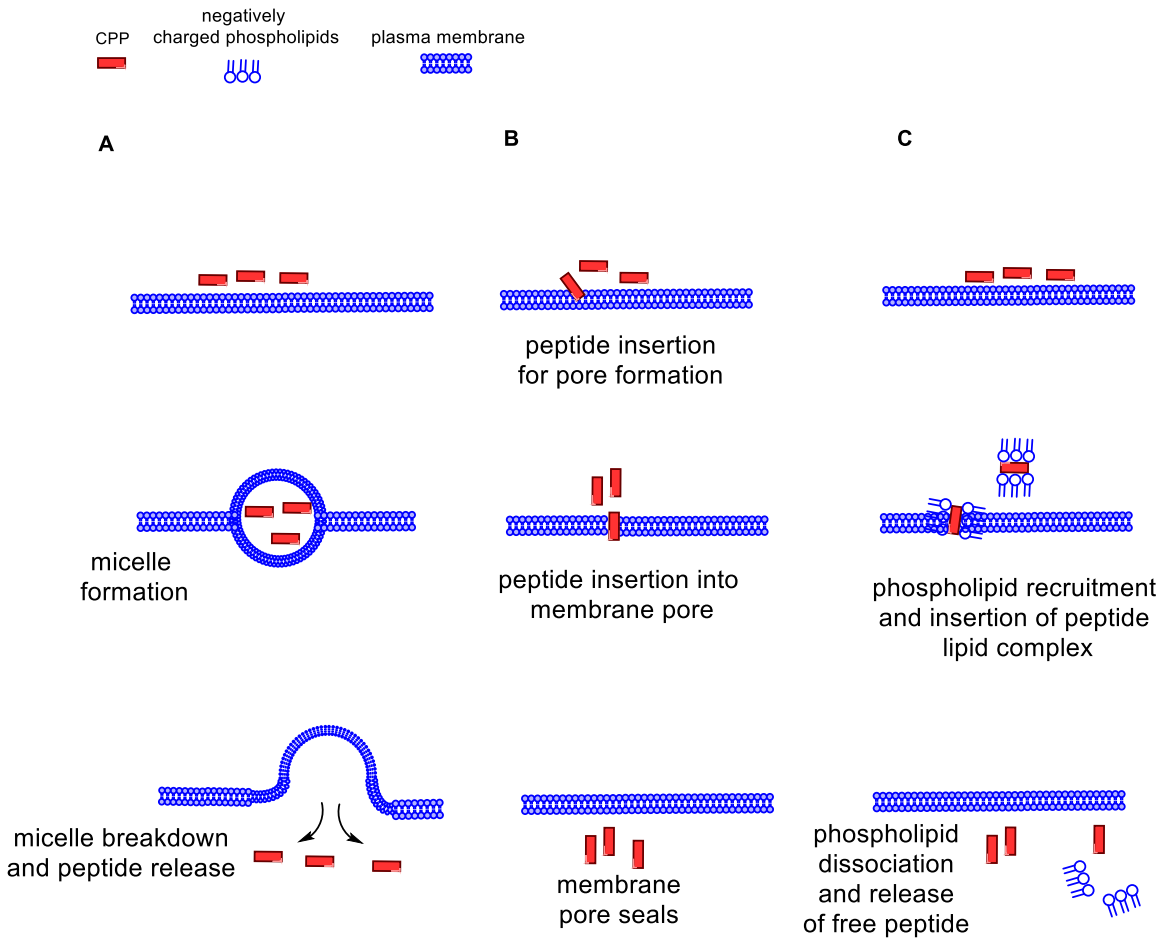


Figure 1.21. The 3 proposed methods of direct translocation by CPPs. (A) Inverted micelle (B) Pore formation (C) Adaptive translocation.[40]

1.6.5 Clinical utility of CPPs

CPPs have progressed from their use in fundamental scientific investigations to an attractive delivery vector for *in vivo* applications, with several progressing to the clinical trial stage.[82] An arginine rich CPP-PMO conjugate (named PPMO-B) was evaluated for its potency, biodistribution and toxicity in mice for the treatment of duchene muscular dystrophy. This CPP-PMO demonstrated enhanced potency in the cardiac, diaphragm and quadricep tissues but also enhanced target protein expression in body-wide tissue.[83] This lack of cell specificity is a concern for using this CPP for any alternative therapy as the wide distribution of the CPP-PMO reduces its therapeutic efficacy. However, the protease stability of these peptides was enhanced by the incorporation of the 6-aminohexanoic acid and β-alanine non-natural amino acids. XG-102 is a Tat-conjugated D-peptide that has passed phase I and II clinical trials as a c-Jun-N-terminal Kinase inhibitor.[84] The phase II trial showed only mild adverse effects (fatigue and headache) in the subjects and promising pharmacokinetic data with a plasma half-life of between 22-39 min.[84] Despite the perceived success of this Tat conjugate other *in vivo* studies of cationic CPPs have proven to again lack tissue specificity. A Tat-β-galactosidase conjugate was synthesised to study organ specificity and this conjugate demonstrated increased distribution in liver, lung and bowel

tissue.[85] Another study by Her *et al.* highlighted some specificity issues where an arginine-rich CPP-Luciferase conjugate was shown to enhance body-wide cell uptake in less than 2 h in mice.[86]

A RXR_4-PMO which is currently being used in the clinic to treat Duchene muscular dystrophy demonstrated increased uptake in liver, spleen, kidney and lungs. This RXR_4-PMO conjugate did show high protease stability, attributed to the 6-aminohexanoic acid groups, with no sign of degradation until 6 h.[87] It is clear some arginine-rich CPPs suffer from a lack of specificity, which could lead to interactions with undesired cell receptors, proteins etc. and so leading to toxicity problems. In particular the clinically arginine-rich CPPs discussed above accumulate significantly in the liver which can lead to hepatotoxicity.[85] Though the incorporation of non-natural amino acid spacers did enhance the *in vivo* stability of some CPPs the problems associated with the lack of specificity means no CPP has been approved for therapeutic use. Further studies are required to fully understand the nature of CPP specificity and toxicity and therefore develop a robust therapeutic.

1.6.6 Cell-penetrating peptidomimetics

One example of using cell-penetrating peptidomimetics as an alternative to cell-penetrating peptides has been explored based on the Tat peptide. Here, both lysine and arginine in the Tat peptide were substituted with isosteres containing ethylenediamine moieties to afford peptide **1.30** (Fig 1.22A).[88] Improved cytosolic delivery was observed in various analogues of Tat in a HEK293T cell line along with no associated cytotoxicity with these peptides. Incorporation of a cyclic dipeptide species into a polylysine peptide **1.31** (Fig. 1.22A) enhanced protease stability with this CPP demonstrating 96% stability over 48 h in human blood serum.[58] The uptake was comparable to the uptake of the established polyarginine R_8 peptide and the toxicity was lower than the R_8 CPP. Looking specifically at arginine, a dipeptide mimetic **1.32** has been developed by Schmuck *et al*. to enhance the bidentate hydrogen bonding to a tridentate interaction (Fig. 1.22B).[89] CPPs bearing tridentate mimetics of arginine were able to deliver the protein avidin to HeLa cells specifically by GAG initiated endocytosis without any toxicity issues. Cyclisation also increases peptide protease stability and has been employed with great success to Tat by Hackenberger *et al.* (Fig. 1.22A).[90] Cyclic Tat **1.33** was able to deliver large cargoes such as GFP and nanobodies specifically to the nucleolus.[90,91] Cyclic versions of CPPs have also been developed by Pei *et al*. such as peptide **1.34** which also incorporates a naphthalene amino acid as a bioisostere of tryptophan to mimic the π system of the tryptophan indole group (Fig. 1.22A).[92,93]

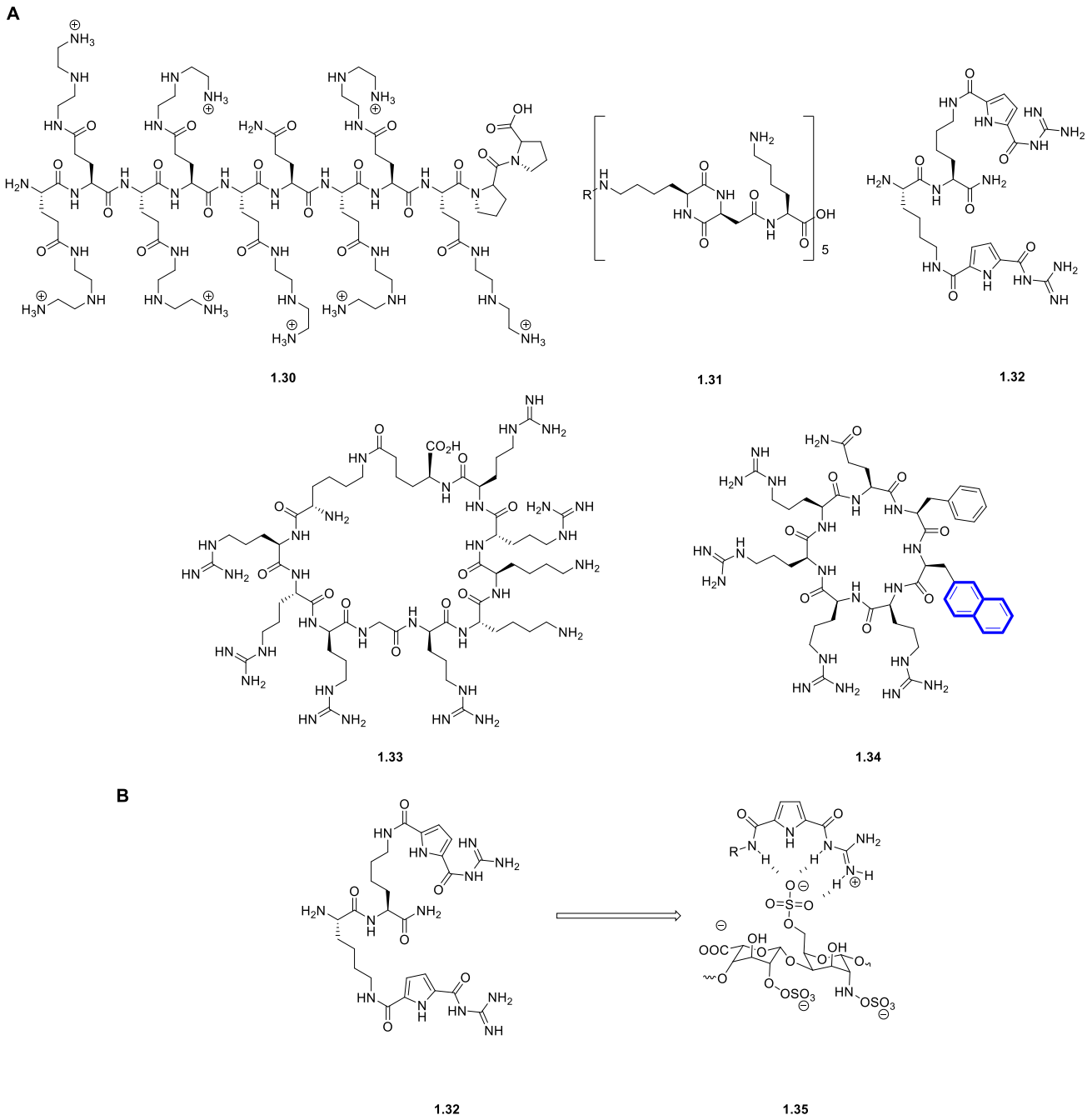


Figure 1.22. (A) Peptidomimetic ethylene diamine bearing Tat **1.30**, cyclic dipeptide species **1.31**, dipeptide arginine mimetic **1.32**, Cyclic Tat **1.33**, Cyclic CPP with tryptophan bioisostere highlighted in blue **1.34**. (B) Tridentate hydrogen bonding from **1.32** to the cell surface, **1.35**.

1.6.7 Drawbacks of existing CPPs

Despite the successes of some CPPs there are still many drawbacks to the currently used CPPs, in particular arginine-rich CPPs such as Tat. A consistent issue with these CPPs is the instability of the peptide, with natural amino acids readily cleaved by proteases. The serine protease trypsin exclusively cleaves amide bonds of cationic residues like lysine or arginine due to a key aspartate moiety in the binding pocket that binds cationic residues (Fig. 1.23).[94–96] This enzyme has limited the application of many linear CPPs as any arginine residues are often cleaved away.

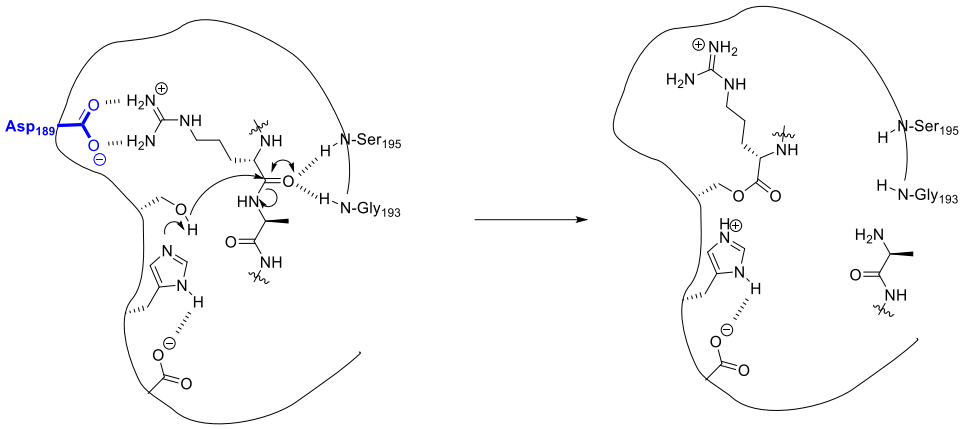


Figure 1.23. The first step in the enzymatic cleavage of a peptide using trypsin highlighting the key Asp_{189} residue required for binding.

Looking at the mechanisms of uptake for CPPs it can be seen that many species are taken up *via* endocytosis mechanisms, however many CPPs can become entrapped in endosomes and not be able to deliver a therapeutic or probe to its target at relevant concentration. It remains unclear if release of a CPP from an endosome relies on a single specific mechanism, though due to the variation in CPP structure a single mechanism would be unlikely.[97] Arginine-rich CPPs suffer greatly from this problem. For example, work by Ferrer *et al.* highlighted that a polyarginine conjugated to a cystic fibrosis transmembrane regulator peptide would become entrapped in endosomes in U2OS cells.[98] Laser irradiation was required to allow cytoplasmic and nuclear distribution of the CPPs.[98] Tat itself has been confirmed to suffer from endosomal entrapment, however work by Pellois *et al.* has demonstrated that a dimer of Tat, dfTat, (Fig. 1.24) can cause endosomal leakage due to its larger arginine density.[99,100]

Though high arginine density aids dfTat and endosomal leakage, it has been associated with various toxic side effects *in vivo*.[101,102] The toxicity of arginine rich CPPs has been attributed to both their high cationic charge and the high basicity of the peptides. Work by Gao *et al*. demonstrated that Tat can lead to both hepatoxicity and cardiotoxicity in MCF-7 tumour bearing mice.[103–105] In addition, many CPPs show a bias towards particular cell lines leading to inefficient cargo delivery and unwanted off-target toxicity. This issue remains prominent for the majority of CPPs due to the fact that cell membranes have the same fundamental architecture. To overcome this problem certain targeting groups have been conjugated to CPPs such antigen specific binders and extracellular protein binding peptides.[106]

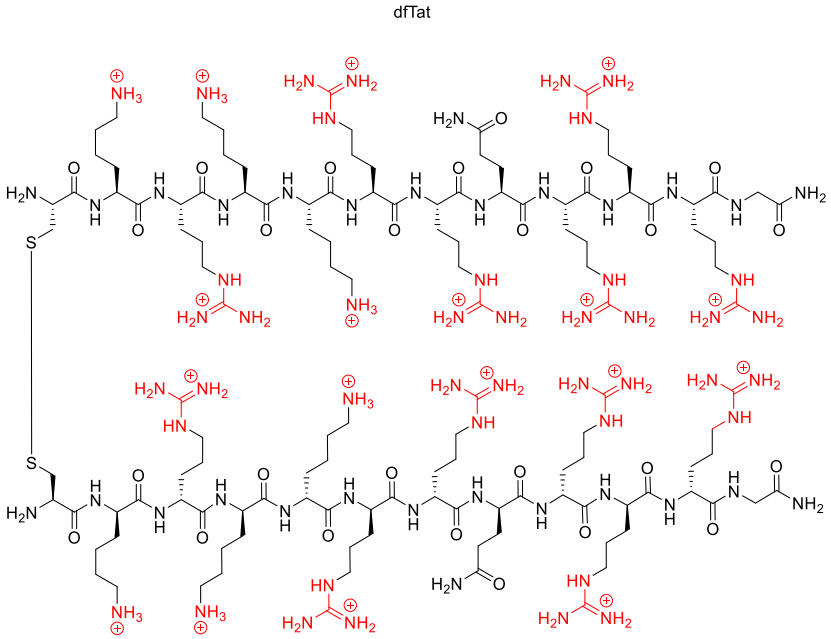


Figure 1.24. Structure of dfTat which causes endosomal leakage.

1.7 The role of the guanidinium group to enhance cell uptake

Arginine contains a guanidinium group in its side-chain which is pivotal for the uptake of many CPPs but can also lead to toxicity issues for arginine rich CPPs. The guanidine group is strongly basic with a pK_{aH} of between 13-14, and upon protonation the subsequent positive charge can be delocalised across the 3 nitrogen atoms. Each nitrogen atom carries a third of the positive charge where the protonation occurs at the imino nitrogen (Fig. 1.25A).[107,108] The guanidinium structure allows CPPs to form bidentate hydrogen bonds and an ion-pair with anionic species on the cell surface. The hydrogen bonds between a CPP and cell surface sulphate from a glycoprotein are formed between the electron rich oxygen atoms in the sulphate and the σ^* anti-bonding orbitals of the N-H bonds in the guanidinium (Fig. 1.25B). The ion-pair forms due to the electrostatic attraction between the negatively charged sulphate and the positively charged guanidinium group. This combination of interactions means the binding between an arginine rich CPP and the cell surface can have three binding interactions per guanidinium group (two hydrogen bonds and one ion-pair, Fig. 1.25C).

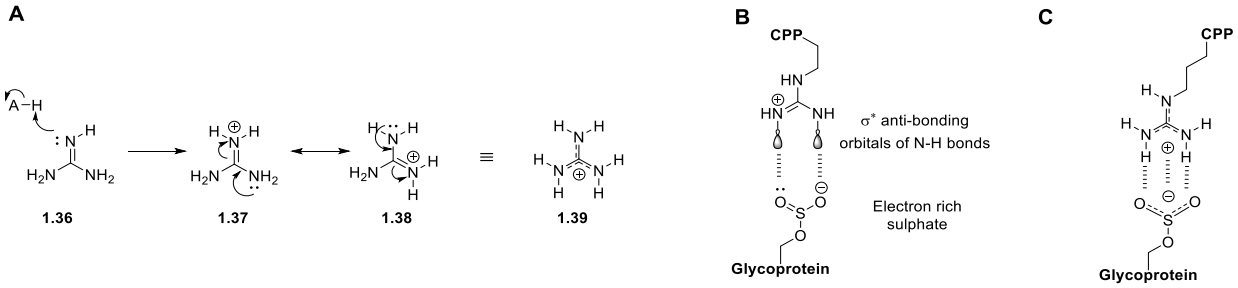


Figure 1.25. (A) Protonation of the sp^2 nitrogen in guanidine **1.36** and resonance forms of the guanidinium cation **1.37**, **1.38** and **1.39**. (B) Hydrogen bonding between the electron rich sulphate of a glycoprotein and anti-bonding orbitals of N-H bonds. (C) The ion-pair and hydrogen bonding complex formed between the guanidinium and a sulphate group.

1.7.1 Bioisosteres of the guanidine group

As discussed in *Section 1.5* peptidomimetics and bioisosteres mimic the interactions of specific groups to elicit a biological response. The guanidine group has been subject to many bioisosteric replacements in both drugs and peptides.[109] Examples of guanidine bioisosteres include cyclic structures in the arginine side-chain such as racemates **1.40** and **1.41**, where the guanidinium is held in either a 5- or 6-membered ring (Fig. 1.26).[110] To reduce the basicity of the guanidine group, and therefore its associated toxicity, a carbonyl group can be introduced adjacent to the guanidine where the pK_{aH} is reduced to 7-8 from 13-14. The use of acetylguanidine linkage has been applied to histamine H2 antagonists **1.42** and **1.43** along with the racemic non-natural amino acid, 5-keto arginine **1.44** (Fig. 1.26).[111–113]

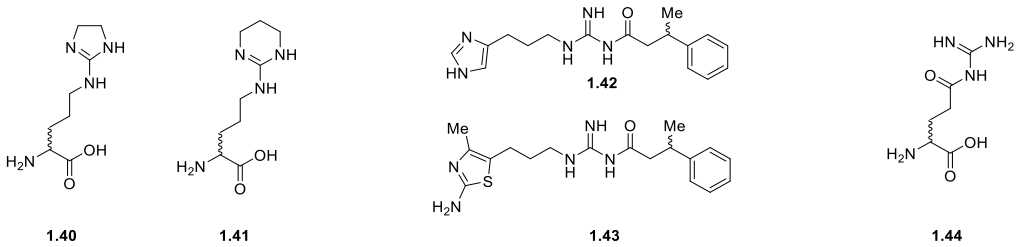


Figure 1.26. Guanidine peptidomimetics. Cyclic guanidine arginine analogues **1.40** and **1.41**. Histamine H2 antagonists with alpha carbonyl guanidines **1.42** and **1.43**. Arginine peptidomimetic 5-keto arginine **1.44**.

1.7.2 Amidines as isosteres of guanidine

Amidines are analogous structures to guanidines with one less nitrogen atom. They are similar to an amide where the carbonyl is replaced with an imino-type carbon-nitrogen bond. Amidines are less basic than guanidines with pK_{aH} values typically in the range of 11-12.[114] As with the guanidine group protonation again occurs at the sp^2 nitrogen to form the amidinium cation (Fig. 1.27A). The delocalised cationic charge is shared across both nitrogen atoms, analogous to a carboxylate anion.

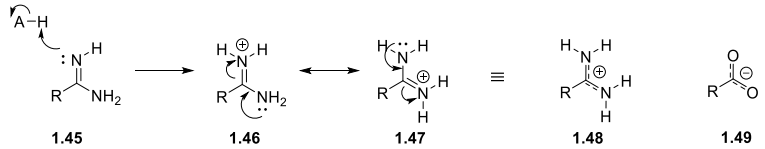


Figure 1.27. Protonation of the sp^2 nitrogen **1.45** with formation of the amidinium cations **1.46**, **1.47** and **1.48** and the analogous carboxylate anion **1.49**.

Due to their matching electrospacial properties to the guanidine group amidines have been employed as bioisosteres for arginine. Amidines are commonly employed as serine protease antagonists. The amidine group is able to form bidentate hydrogen bonds and an ion-pair with carboxylate moieties in the serine protease binding pocket (Fig. 1.28A).[115] There are many therapeutics utilising the amidine motif such as the pro-drug dabigatran and work by Clement *et al.* further investigated amidine containing serine protease inhibitors like amidine **1.50** (Fig. 1.28B).[116,117] Finally amidines have been developed as DNA minor groove binders in species such as pentamidine and distamycin A (Fig. 1.28B) but have seen limited use in peptides.[118]

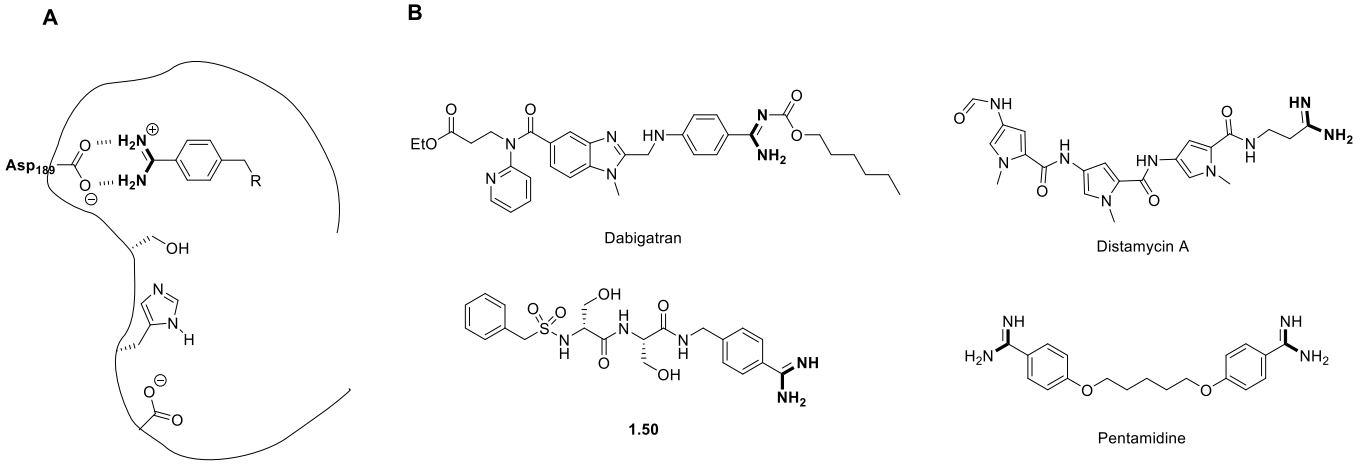


Figure 1.28. (A) Bidentate hydrogen bonding formed between a carboxylate and an amidine in serine protease antagonists. (B) Amidine containing drugs Dabigatran, Distamycin A, Pentamidine and serine protease inhibitor **1.50**.

In summary, the cell membrane provides a barrier for the access to the cytosol. To deliver synthetic molecules to the cell Lipinski's rules are often used to design species that are both orally bioavailable and cell permeable. Peptides also offer an alternative to other types of drug delivery systems through the use of CPPs. CPPs are often rich in arginine residues which allows for binding interactions with the cell membrane but also lead to problems such as toxicity from high basicity. Because of this no CPPs are currently utilised in therapeutic approaches. The importance of the guanidinium groups in CPPs is well established and although many types of bioisosteres have been employed in medicinal compounds, the use of a guanidine mimetic placed in the side chain of a CPP has not been investigated so far.

1.8 Hypothesis

The hypotheses to be tested in this thesis is that non-arginine rich CPPs have as high a cell uptake as arginine rich CPPs such as Tat and Pip6, and that replacement of the guanidinium group with an amidine group will maintain the same level of cell uptake in these CPPs. The assumption is that amidinium groups will form bidentate hydrogen bonds and ion-pairs with cell surface anionic species such as glycoproteins, but will have a higher hydrophobicity (Fig. 1.29). Further it is hypothesised that the replacement will produce a reduction in the toxicity generally associated with the guanidinium-rich CPPs, since amidinium groups are less basic than guanidinium groups.

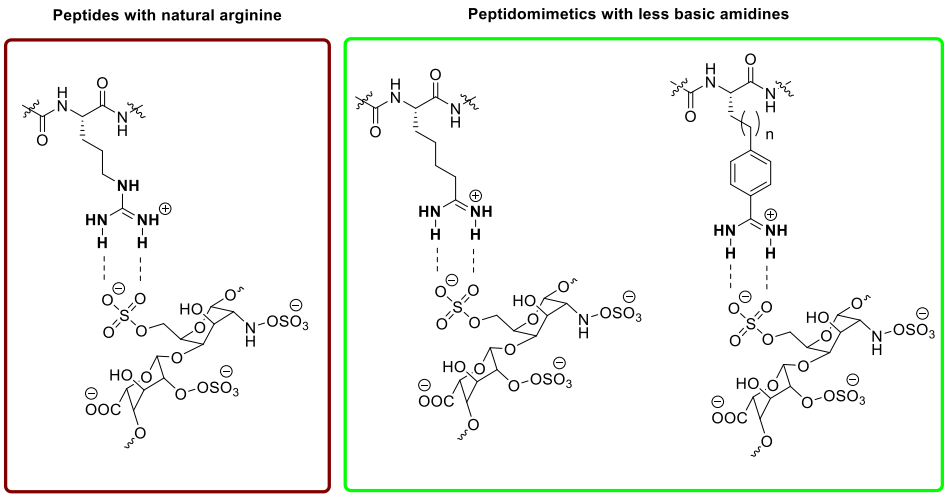


Figure 1.29. Use of amidines as mimetics of the guanidinium group to maintain CPP cell uptake and reduce basicity associated toxicity.

1.9 Aims and objectives of this thesis

The specific aims of this project are to:

(i) Synthesise and subsequently analyse the cell uptake of low arginine-containing CPPs.

(ii) Establish robust routes to prepare amidine-containing building blocks (**1.51** - **1.52**) suitable for use in solid phase peptide synthesis of CPPs.

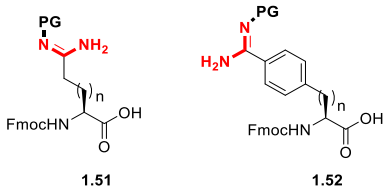


Figure 1.30. General structures of amidine containing building blocks **1.51** and **1.52**.

(iii) Develop mild synthetic methodology to directly arylate amidines for potential application as CPP bioisosteres (Fig. 1.31).

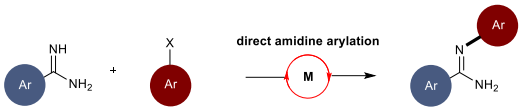


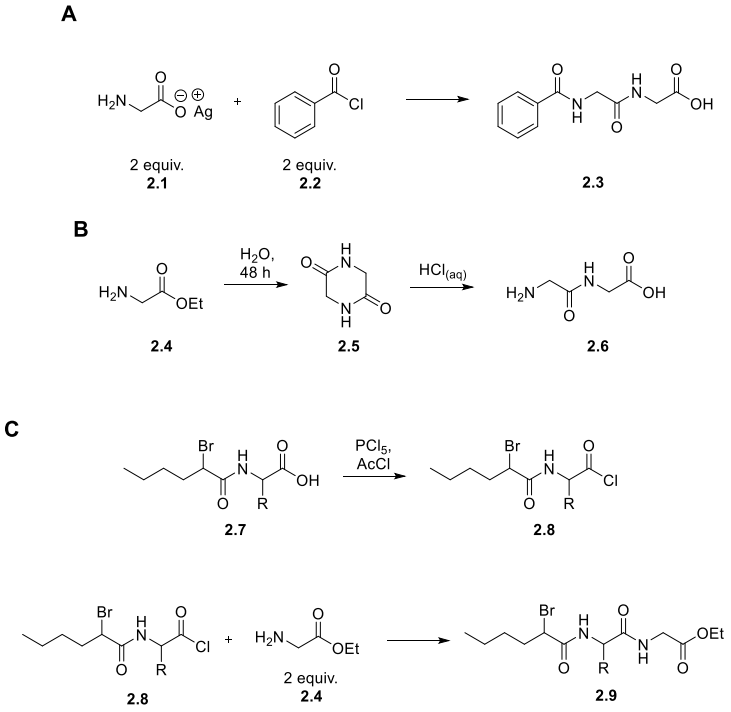
Figure 1.31. Direct arylation of amidines to be investigated.

Chapter 2: Toxicity and uptake of low arginine-containing CPPs

2.1 Introduction

2.1.1 Synthesis of peptides

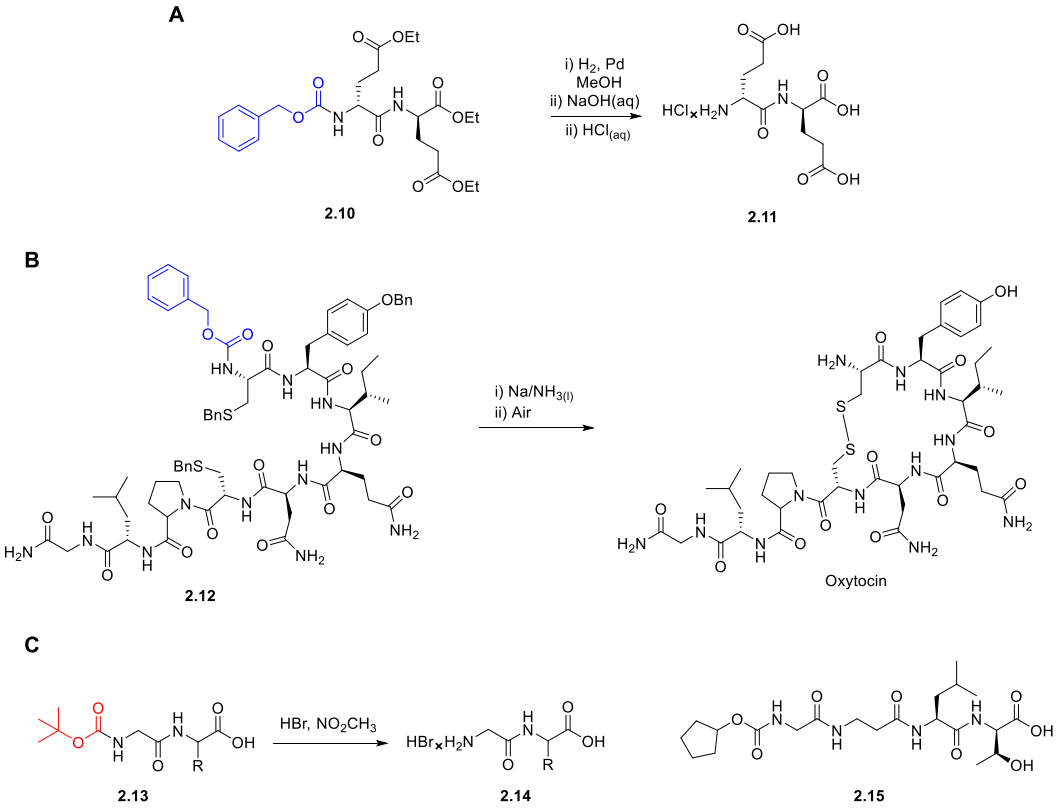
Peptides are highly desirable due to their biological importance and broad applications as both therapeutics and probes, including CPPs.[45] The synthesis of peptides today is typically carried out on a solid support in what is known as solid phase peptide synthesis (SPPS).[119,120] The first synthetically produced peptide was reported by Curtius, where the synthesis was completed in solution using glycine salt **2.1** to synthesise an *N*-protected glycine-glycine dipeptide **2.3** (Scheme 2.1A).[121] Building on this, Fischer *et al*. reported the first synthesis of glycine-glycine dipeptide **2.6** with an unprotected *N*-terminus that was synthesised *via* a piperazine intermediate **2.5** (Scheme 2.1B).[122] Fischer then reported the synthesis of peptides up to 18 amino acids in length when using an acyl chloride coupling strategy, though the peptides were limited to glycine and leucine amino acids only (Scheme 2.1C).[123,124]



Scheme 2.1. (A) Dipeptide synthesis from silver salts reported by Curtius.[121] (B) Synthesis of dipeptide with an unprotected *N*-terminus reported by Fischer *et al*.[122] (C) Peptide synthesis using a chlorination strategy reported by Fischer.[123,124]

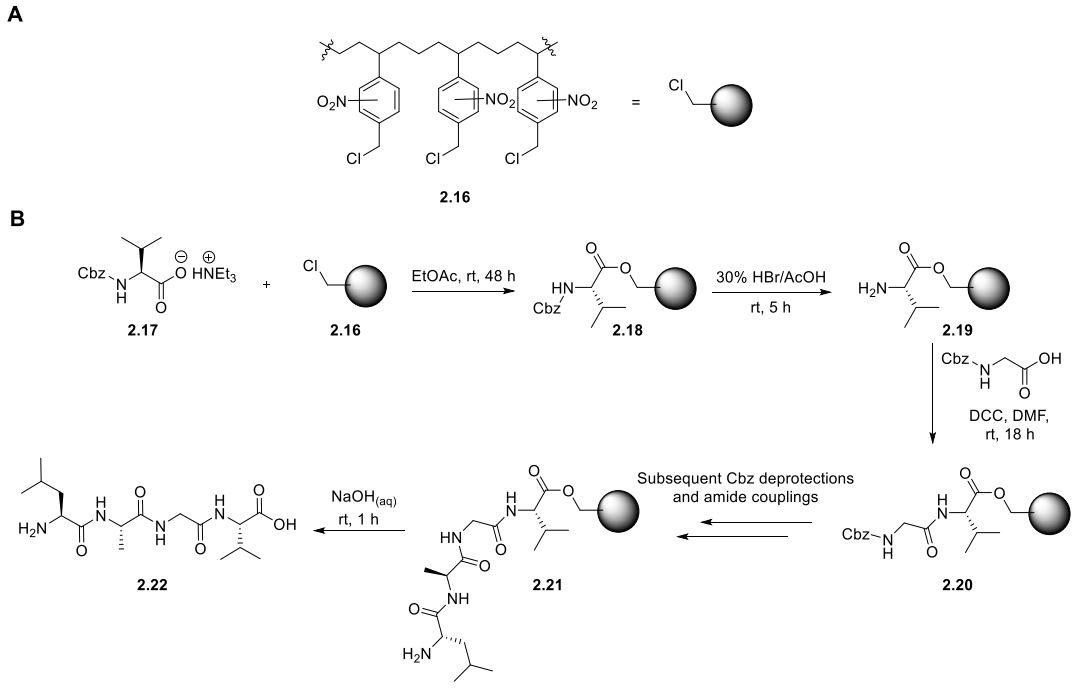
The methods developed by Curtius and Fischer remained limited in scope to peptides containing glycine and leucine and also containing difficult to remove protecting groups.[121–124] To synthesise unprotected *N*-terminal amino acids and allow for easy sequential amino acid couplings, Zervas *et al*. reported the use of the Cbz protecting group which can be removed by hydrogenolysis (Scheme 2.2A).[125] The Cbz group still could not be employed on all amino acid residues as sulphur containing methionine and cysteine poisoned the Pd catalyst thus halting the hydrogenolysis.[125] The Cbz method of protection and deprotection was utilised by Du Vigneaud *et al.* for the first synthesis of the peptide oxytocin, though

the final deprotection was carried out using a Birch reduction due to the presence of cysteine residues (Scheme 2.2B).[126] The next development in peptide synthesis, by Albertson *et al*., was the utilisation of the Boc protecting group at the α-amino position on an amino acid (Scheme 2.2C).[127] The Boc groups acid lability meant it could be used orthogonally to Cbz in the synthesis of peptides and Albertson *et al.* were able to synthesise tetrapeptide **2.15** using a Boc protection strategy (Scheme 2.2).[127]



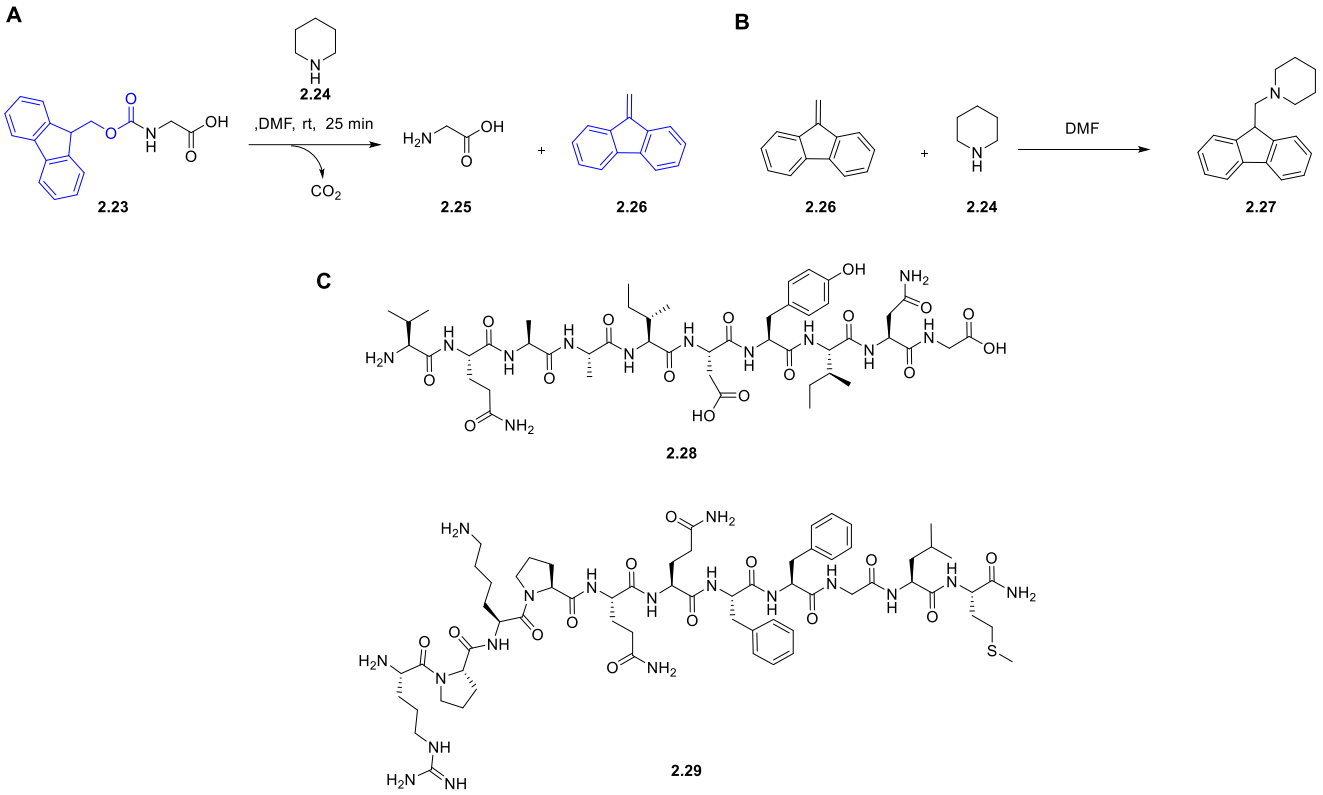
Scheme 2.2. (A) Cbz deprotection by hydrogenolysis. (B) Cbz deprotection used in the synthesis of oxytocin. (C) Boc deprotection strategy used by Albertson *et al* for the synthesis of **2.15**.[127]

Though the synthesis of more complex peptides had been achieved in solution the purification and isolation of multiple intermediates made the process both time-consuming and low yielding. Merrifield *et al.* demonstrated that peptide synthesis could be carried out on a solid support without the need to isolate intermediate species and demonstrated the entire peptide could be cleaved from the solid support in a single step (Scheme 2.3).[128] The Merrifield synthesis used excess reagents to ensure successful amide formations and any side products or unreacted reagents could be easily removed by filtration. The original Merrifield synthesis utilised chloromethylated nitropolystyrene polymer **2.16** as the solid support to synthesise tetrapeptide **2.22** (Scheme 2.3A & B).[128]



Scheme 2.3. (A) Chloromethyl nitropolystyrene resin **2.16** used by Merrifield.[128] (B) SPPS conducted by Merrifield for the synthesis of tetrapeptide **2.22**.[128]

Further work by Merrifield demonstrated that a Boc protection strategy could be employed as well as Cbz meaning a non-nitrated solid support could also be used to access a nonapeptide.[129] A different amine protection strategy was later described by Carpino *et al*. using the base labile Fmoc group on a variety of amino acids (Scheme 2.4A).[130,131] The reactivity of the liberated fulvene **2.26** meant the Fmoc group was less suitable for solution phase peptide synthesis, however a solid phase approach does not exhibit the same limitations (Scheme 2.4B). The ability to wash away side products from the resins allowed for the use of Fmoc chemistry in SPPS by Atherton *et al.* where peptides **2.28** and **2.29** were synthesised on polydimethylacrylamide resins (Scheme 2.4C).[132,133]



Scheme 2.4. Fmoc protecting group cleaved using piperidine **2.24**. (B) Side reaction between fulvene **2.26** and piperidine **2.24**. (C) Peptides **2.28** and **2.29** synthesised on solid support using Fmoc protection chemistry by Atherton *et al.*[133]

The advent of Fmoc chemistry in SPPS was developed along with orthogonal acid sensitive resins upon which peptides could be grown. Resins are made of cross-linked polymer networks where the initial attachment of the first amino acid occurs at the *C*-terminus. Solid phase approaches, however, require sufficient penetration of reagents which is dependent on good swelling of the resin increasing the diffusion rate of reagents in the expanded resin network (Fig. 2.1).[134,135] For polystyrene resins it was reported by Merrifield that swelling was far greater in DCM than DMF for the free resin but that DMF induced swelling was greater with increasing length of the attached peptide chain.[136] This observation is advantageous for Fmoc chemistry as DCM is not compatible with piperidine due to its cross reactivity.[137] Therefore, DMF is generally used for coupling and deprotection steps on solid support and DCM for initial resin swellings.

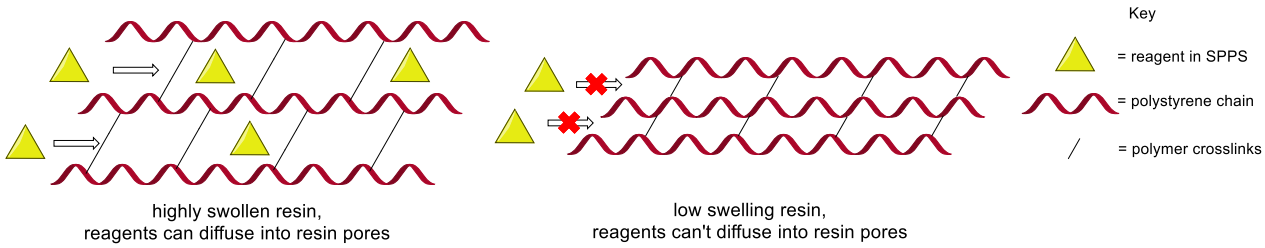
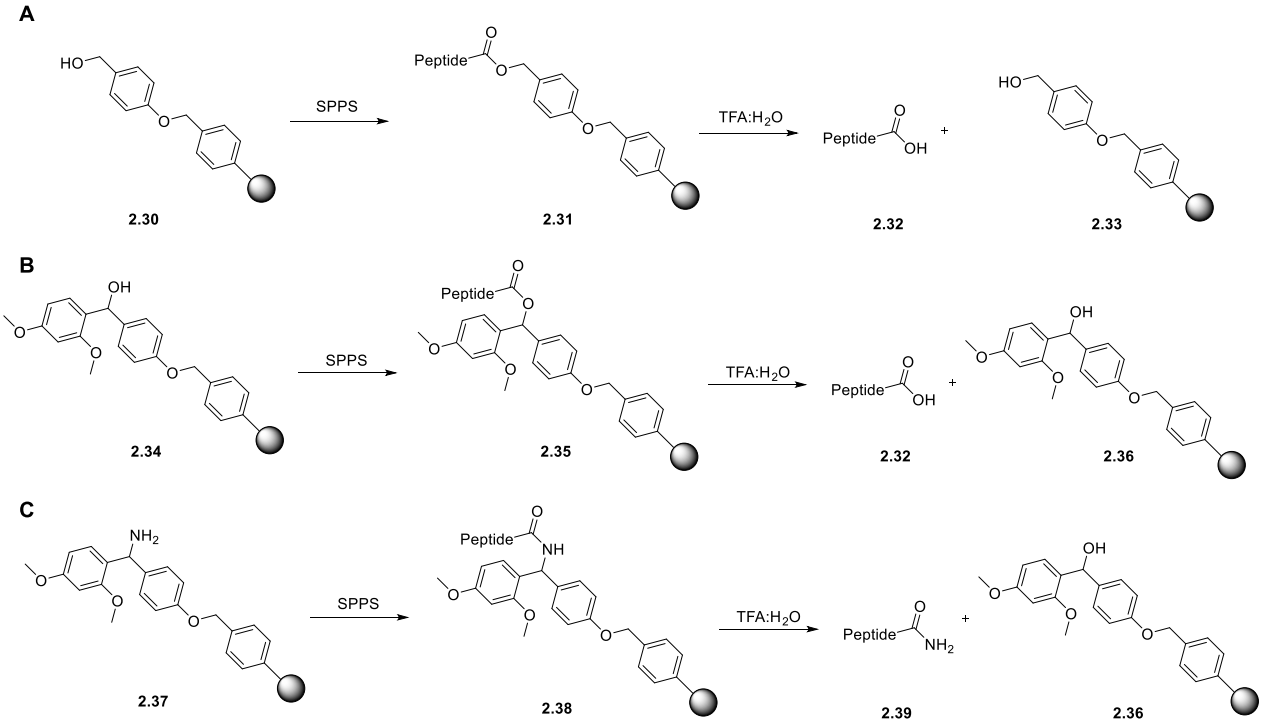


Figure 2.1. Effect of swelling on resin pores and how it improves access of reagents.

An acid sensitive resin was developed by Wang *et al.* for the synthesis of carboxylic acids at the *C*-terminus. The Wang resin contains a benzyl alcohol functionality which forms ester **2.31** upon coupling with an activated amino acid, at which point subsequent couplings can occur while the growing peptide chain is anchored to the resin (Scheme 2.5A).[138] The free acid of the peptide can be liberated from the wang resin using acids such as TFA, leaving a cationic resin that can be quenched with nucleophilic species such as water, phenol or various thiols to prevent unwanted side reactions.[139] Work by Rink *et al.* demonstrated that two further acid sensitive resins, rink acid linker **2.34** and rink amide linker **2.37**, could anchor peptides through either an ester or amide linkage (Scheme 2.5B & C).[140] As with Wang resins, Rink linkers are acid sensitive and can deliver carboxylic acids or primary amides at the *C*-terminus and the resulting cationic resin can be quenched with addition of nucleophiles.[140,141]

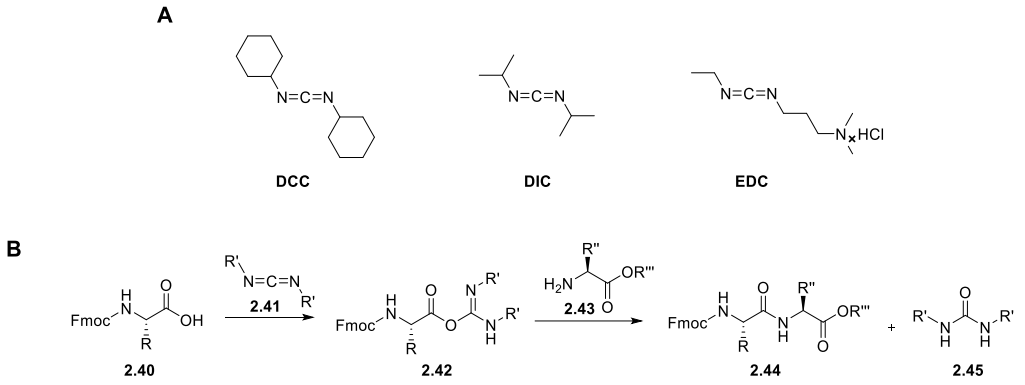


Scheme 2.5. (A) Wang resin linker. (B) Rink acid resin linker. (C) Rink amide resin linker.[138,140]

Extensive research into appropriate coupling agents that can be employed using Fmoc protecting groups and acid sensitive solid supports has been carried out previously by various groups. To date there are

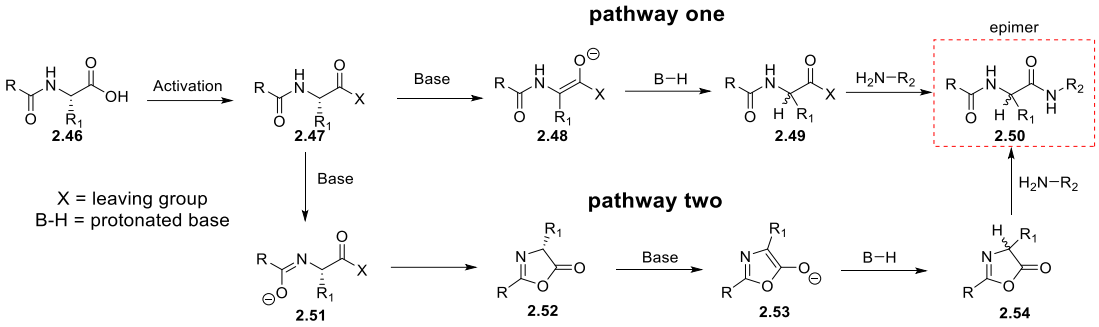
three main classes of amide coupling agents employed in SPPS: carbodiimides, uronium salts and phosphonium salts.[119,120]

The carbodiimides typically used in peptide synthesis are DCC, DIC and EDC (Scheme 2.6A). Amide bond formation with carbodiimides begins with nucleophilic attack of carboxylic acid **2.40** to form activated *O*-acylisourea **2.42**. Species **2.42** then undergoes nucleophilic attack from an amine group that releases the desired amide bond product **2.44** and urea side product **2.45** (Scheme 2.6B).[142,143] DCC was also the first carbodiimide to employed in 1955 by Hess *et al.* in peptide synthesis.[144] Due to the formation of highly a insoluble cyclohexylurea by-product further carbodiimides were sought after for use in SPPS. The carbodiimindes DIC and EDC were developed by Sheenhan *et al.* and Chen *et al.* for to overcome the formation or cyclohexyl urea when using DCC.[145,146]



Scheme 2.6. (A) Structures of carbodiimides DCC, DIC and EDC. (B) General mechanism for a carbodiimide mediated amide coupling using Fmoc amino acids.

A major problem with amide couplings using α-amino acids is the epimerisation of the α-carbon stereocentre.[147] The mechanism of epimerisation can occur *via* two base catalysed pathways (Scheme 2.7).[147,148] In both pathways the initial activation of the amino acid takes place to form an activated species **2.47**. In pathway one, deprotonation occurs at the α-carbon to form enolate **2.48**. Subsequent protonation can occur at either face of the enolate resulting in a loss chiral purity and species **2.49**. The final step in pathway one is the nucleophilic substitution and amide formation to form epimer **2.50**. In pathway two deprotonation occurs on the amide bond to form imidate **2.51**. A 5-*exo-trig* cyclisation then occurs to form oxazolone **2.52**. Deprotonation of **2.52** forms enolate **2.53** which can be protonated from both faces to form oxazolone **2.54**. Final nucleophilic attack by an amine then returns the epimer **2.50**.



Scheme 2.7. The two base catalysed epimerisation pathways that can occur during amide couplings.

The epimerisation problem is prominent for carbodiimide mediated amide couplings due to the high reactivity of *O*-acylisourea.[149] The addition of triazoles such as HOBt and HOAt has been used to reduce the level of epimerisation when using carbodiimides. The less reactive triazole activated ester **2.55** formed reduced the ability to form oxazolones compared *O*-acylisoureas intermediate **2.56** (Fig 2.3).[149–151]

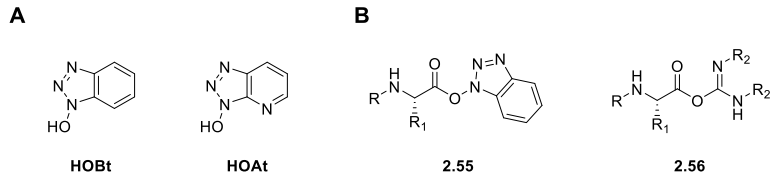


Figure 2.2. (A) Structures of triazole additives HOBt and HOAt. (B) Triazole activated ester **2.55** and O-acylisourea **2.56**.

Two other forms of amide coupling agents, phosphonium and guanidinium salts, take advantage of triazole activated esters but are often formed *in situ*, without the addition of extra triazole reagents (Fig. 2.3). Phosphonium salts agents contain a triazole moiety bound the phosphorus *via* an oxygen atom and were first used by Kenner *et al.* before the development of BOP by Castro *et al.* in 1975.[152,153] The BOP reagent was found to produce the toxic by-product, hexmethylphosphorotriamide (HMPA), during activation and this led to the development of other agents such as PyBOP and PyAOP where the dimethylamino groups are replaced with pyrrolidine species.[154,155] Guanidinium salt agents contain a guanidinium group and an *N*-oxide unlike the phosphonium group in BOP, PyBOP or PyAOP. HBTU was the first guanidinium agent to be developed by Dourtoglu *et al.* and its structure confirmed by Abdelmoty *et al.* in 1994.[156,157] On the back of HBTU other guanidinium agents were developed including HATU and HCTU. Currently, HATU has been reported as returning the highest coupling yields and lowest levels of epimerisation compared to either PyBOP or HBTU.[151,158,159]

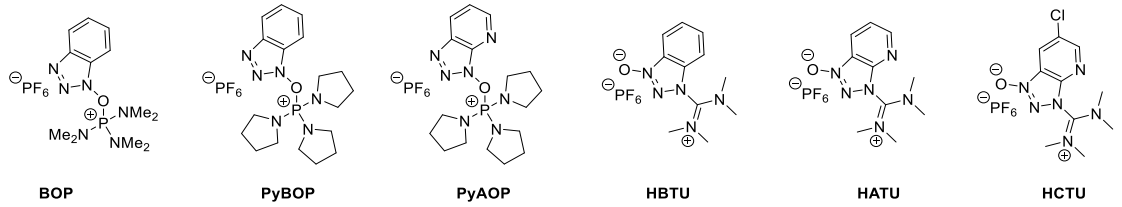
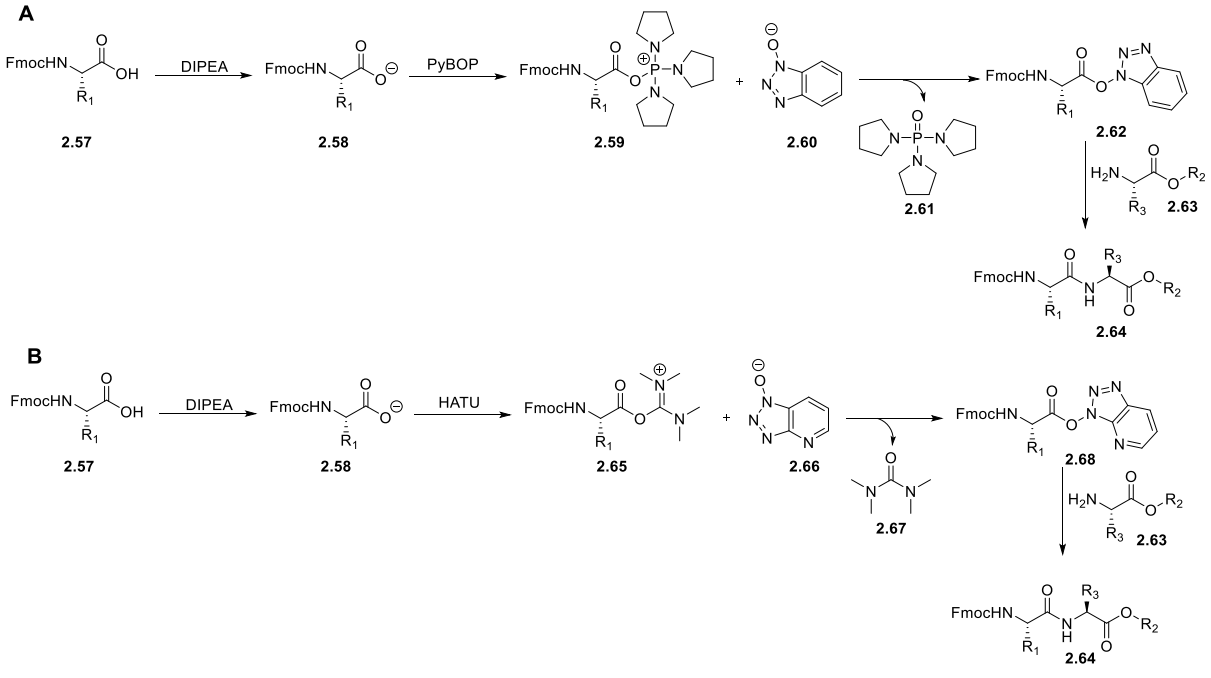


Figure 2.3. Structures of phosphonium and guanidinium salt amide coupling reagents.

The conditions for amide formation when using either phosphonium of uronium salts typically require and excess of base, often a tertiary amine such as DIPEA (Scheme 2.8A). When using phosphonium agents, such as PyBOP, carboxylate **2.58** attacks the electrophilic phosphonium in PyBOP to form phosphonium **2.59** and release triazole **2.60**. Subsequently **2.60** and **2.59** react to form activated ester **2.62** and release phosphine oxide side product **2.61**. Activated ester **2.62** can then undergo nucleophilic substitution after addition of unprotected amine **2.63** to form the amide product **2.64**. This has become the accepted mechanism when using phosphonium agents such as PyBOP.[120,154,160]

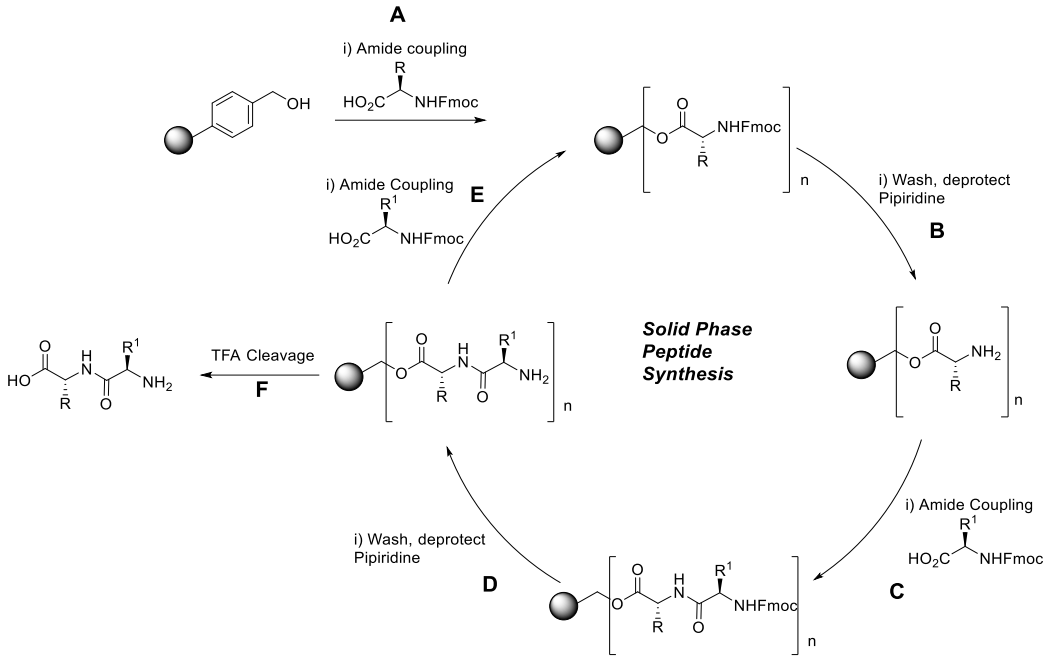
Guanidinium salts, such as HATU, function in a similar way to phosphonium agents where excess base is required to generate carboxylate **2.58** (Scheme 2.8B). Carboxylate **2.58** then attacks the guanidinium species to form *O*-acylisourea **2.65** and triazole **2.66**. Both **2.65** and **2.66** then react to form activated ester **2.68** and release urea **2.67** as a side product. Upon addition of an amine the activated ester **2.68** can again undergo nucleophilic substitution to afford the amide product **2.64**.[119,120,161]



Scheme 2.8. (A) Mechanism for amide formation using PyBOP. (B). Mechanism for amide formation using HATU.

To date both the phosphonium and guanidinium agents are extensively used in SPPS for the synthesis of a variety of peptides. Combined with a Fmoc protection strategy, a general protocol for SPPS has

become well-established (Scheme 2.9).[119,120] The protocol begins with an initial resin swelling and coupling of the first amino-acid to the solid support before coupling side products are washed away. The resin bound amino acid is then deprotected and Fmoc side products washed from the resin. The next amino acid building block is then activated and added to the resin to be coupled to the resin bound amino acid. Repetitions of deprotections and amide couplings are performed until the desired peptide sequence has synthesised on the solid support. The full peptide is finally cleaved from the resin under acidic conditions, typically using a cleavage cocktail containing TFA.



Scheme 2.9. Solid phase peptide synthesis using a wang resin. (A) Resin loading with first amino acid building block. (B) Fmoc deprotection. (C) Amide coupling of next amino acid. (D) Fmoc deprotection of new elongated peptide chain. (E) Amide coupling of next amino acid building block. (F) Cleavage of peptide from the resin using TFA.

2.1.2 Monitoring cell permeability of peptides

Upon synthesis of peptides the evaluation of their biological applications will often require them to be cell permeable. There are numerous ways to assess cell permeability and they include: synthetic model membranes; monolayer assays; cell lysate analysis and live cell fluorescence assays.[97] Each of these methods have advantages and disadvantages over each other and have widely been applied to the field of CPPs.

2.1.2.1 Synthetic model membranes

One of the first model membrane assays developed was PAMPA by Kansy *et al*.[162] The membranes formed for PAMPA make use of a lipid in an alkane solution with Kansy *et al.* using 10% egg-lecithin (which contains several different phospholipids including POPC) in dodecane (Fig. 2.4A). The assay involves two compartments, a donor and acceptor, separated by the synthetic lipid membrane.

Compounds are added to the donor well and after an incubation period the acceptor well contents are quantified, typically by UV-vis or mass spectrometry (Fig. 2.4B).[163,164] Many PAMPA methodologies have been developed to mimic naturally occurring membranes, including the blood brain barrier and gastrointestinal tract.[165] The principle advantage of the PAMPA is its quick preparation and high throughput capacity, with many assays performed in a 96-well plate format.

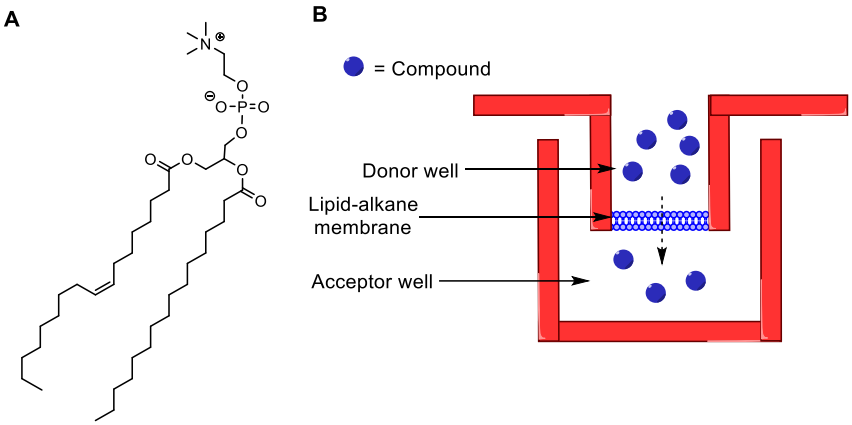


Figure 2.4. (A) Structure of POPC, a constituent of egg-lecithin. (B) Diffusion of a compound through a synthetic lipid alkane membrane in a PAMPA.

Another use of synthetic membranes is by forming unilamellar vesicles. Using lipids such as POPC, vesicles can be formed and uptake of a fluorescent probe, for example in work by Almeida *et al.*, can be monitored upon treatment with a CPP then the uptake quantified (Fig 2.5).[166] The major drawback of this method is it requires the peptide to form pores in the membrane to allow the fluorophore to be released and fluorescence to be measured. Similar protocols have been used looking at permeability of cyclic CPPs in multilamellar vesicles to determine the direct translocation capacity.[167] Despite the ease of using both PAMPA and vesicle assays, they remain limited to monitoring permeability by passive diffusion only. As active processes such as endocytosis are not possible, the permeability of many biomolecules cannot be assessed by PAMPA and vesicular methods.

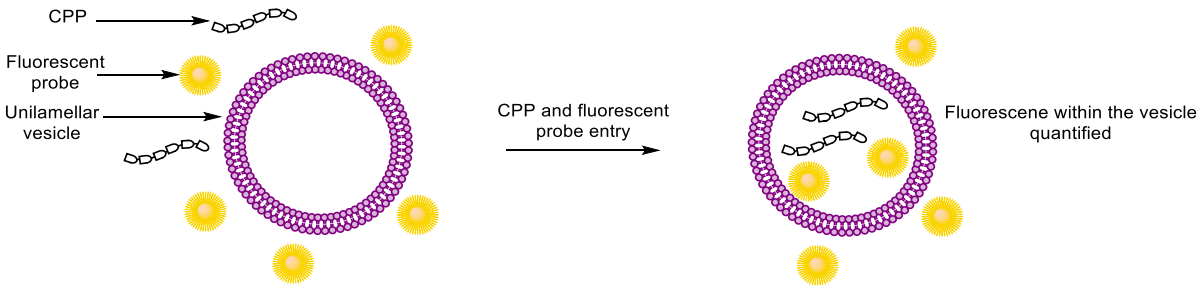


Figure 2.5. Fluorescene based permeability assay using synthetic a unilamellar vesicle.

2.1.2.2 Monolayer assays

Often employed to assess permeability through a single layer of cells, the most common monolayer assays are CaCo-2 and MDMK. CaCo-2 cells are used to mimic enterocytes and therefore the

environment of the small intestine, whereas MDMK cells are nephrocytes and mimic the renal environment.[168,169] The monolayer assays have a similar set up to the PAMPA but with the key difference being the use of live cells as the membrane barrier within each well (Fig. 2.6). Compounds are added to a donor well where they can pass through the monolayer of cells and be analysed from the acceptor well, by LCMS or HPLC techniques.[170] The monolayer assays offer an advantage over synthetic membranes as active methods of uptake can occur, therefore not only direct translocation of compounds is monitored. However, a compound must pass through the cell to the acceptor well so species which are retained or exhibit slow transport through the cellular environment will not be registered as permeable. Though both monolayer and synthetic membrane assays can allow for a high throughput screen of compounds neither technique can be used to observe sub-cellular localisation.

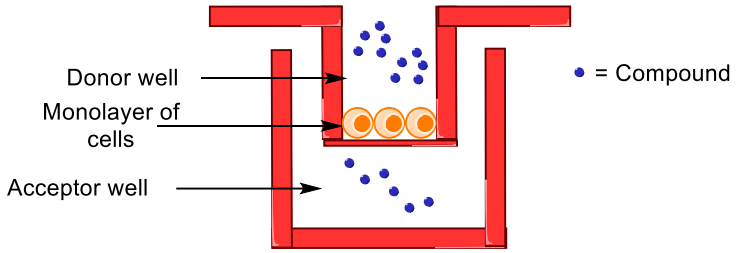


Figure 2.6. Representation of a monolayer assay like the CaCo-2 or MDMK assays.

2.1.2.3 Cell lysate analysis

The analysis of cell lysate is a technique that allows for the quantification of total cell uptake of a compound. A compound is incubated with live cells before the cells are washed then lysed for the species within the cell to be quantified (Fig 2.7). Lysate analysis was first used to identify the uptake of a peptide substrate by Oehlke *et al.* where cells were incubated with a peptide then washed with a diazo aniline which would react with extracellular peptide. The cell lysate was then analysed by HPLC or MALDI-TOF-MS and unmodified peptide detected was deemed to have been internalised.[65] An alternative to HPLC quantification is the use of MALDI-TOF-MS, work by Ekman *et al.* and Burlina *et al.* used MALDI-TOF-MS to accurately quantify the amount of peptide within a cell.[171,172] The method requires an isotopically labelled form of the peptide which allows accurate quantification from the semi-quantitative MALDI spectrum. Cell lysate based assays have the advantage over synthetic membranes as using live cells allows uptake by both active and passive pathways, therefore detection does not require a compound to passively cross membranes. The analysis of cell lysate, however, is often complicated by the vast amount of species within the cell. Therefore, the methods generally require multiple washing and fractionating steps to separate cell species from the molecule of interest. The use of HPLC and mass spectrometry also requires synthesis of authentic standards which reduces the throughput since this requires synthetic effort.

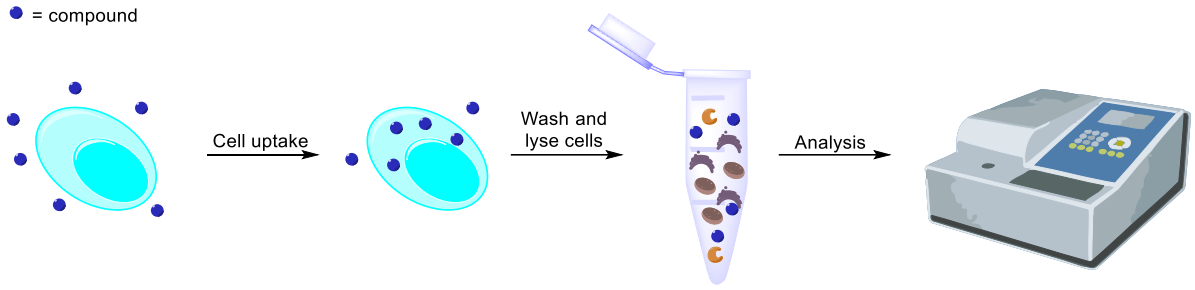


Figure 2.7. Protocol that quantifies cell uptake of a compound from cell lysate.

2.1.2.4 Fluorescence based cell uptake assays.

The use of fluorophores to monitor and quantify uptake in live cells is one of the most commonly used methods, particularly in the field of CPPs. The two main analysis techniques, typically used in tandem, are fluorescence microscopy and flow cytometry. Microscopy is a non-destructive technique that allows for the imaging of a monolayer of live cells and the subsequent quantification of a fluorophore of interest. It can often be difficult to distinguish cell surface bound and internalised compound by images of monolayers and thus accurately quantify cell penetration.

Flow cytometry analysis requires a suspension of cells within a liquid which are taken up and focussed through a nozzle into a flow chamber one cell at a time (Fig. 2.8).[173] A laser pulse hits the cells as they pass through the chamber and the transmitted light is measured in either forward or sideways direction, as well as at the specific wavelengths of the attached fluorophore. Forward scattered light (FSC) is detected along a path in line with the laser, while side scattered light (SSC) is measured perpendicular to the incident light.[174] The FSC is directly proportional to the diameter of the cell due to the singular diffractions of light around the cell. The SSC detects the granularity or internal complexity of the cell, as species within the cell cause light to refract or reflect.[174] A particular population of cells within a certain range of FSC and SSC can then be measured for their level of fluorescence to quantify the uptake of a fluorophore.[97] For microscopy and flow cytometry experiments, to ensure any non-permeable

compound is not analysed, the cells are incubated with the fluorescently labelled compound before being washed repeatedly with combinations of trypsin and buffered salt solutions.[97]

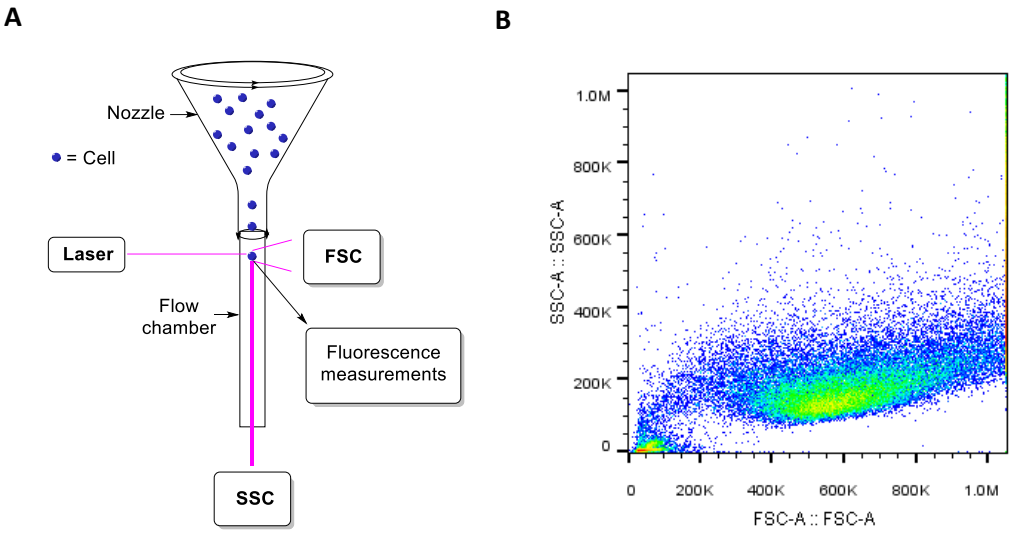


Figure 2.8.(A) Representation of flow cytometry. (B) FSC and SSC plot of cells used for flow cytometry analysis.

The major caveat of fluorescence-based techniques is that the compound must be labelled with a fluorophore, thus changing the structure and chemical properties of the molecule of interest. However, since the overall aim of a CPP is to deliver species to a cell the addition of a fluorophore can demonstrate the peptides ability to deliver a cargo.

2.2 Aims of Chapter 2

Many of the current CPPs rely on arginine residues to facilitate their efficient cell uptake, though it has been reported that there are toxicity concerns over using arginine rich CPPs.[102,105,175] Two CPPs that are effectively arginine deficient are **TP-1** and **TP-2**. These peptides have been previously investigated and found to proceed *via* a direct translocating mechanism in CHO cells but require endocytotic uptake in HeLa cells.[79,80] Their specific mechanism of uptake in human cell lines is unknown but they could offer a less toxic alternative to current arginine rich CPPs. Furthermore, the importance of the arginine residues in both uptake and retention within a cell has not been assessed for either TP CPPs.

The specific aims of this chapter are to:

(i) Prepare fluorescently-labelled **TP-1**, **TP-2** and arginine-rich CPPs for cell uptake analysis.

(ii) Determine the toxicity of this peptide series.

(iii) Investigate the specific mechanism of uptake for the TP CPPs using flow cytometry and compare these with arginine-rich CPPs.

(iv) Identify any differences in subcellular localisation within the CPP series.

2.3 Results and discussion

2.3.1 Design and synthesis of CPPs

Both **TP-1** and **TP-2** CPPs contain only two arginine residues, each separated by two leucine residues to give an RLLR motif (Fig 2.9). To understand the importance of this motif upon a peptides permeability two peptides were designed that might possess cell penetrating properties, **2.68** and **2.69** (**RLLRLLR** and **RLLRRLLR**), based on the common repeating unit (Fig. 2.9). Peptides **2.68** and **2.69** contain three and four arginine residues respectively, which may allow for cell uptake based on the arginine content alone.

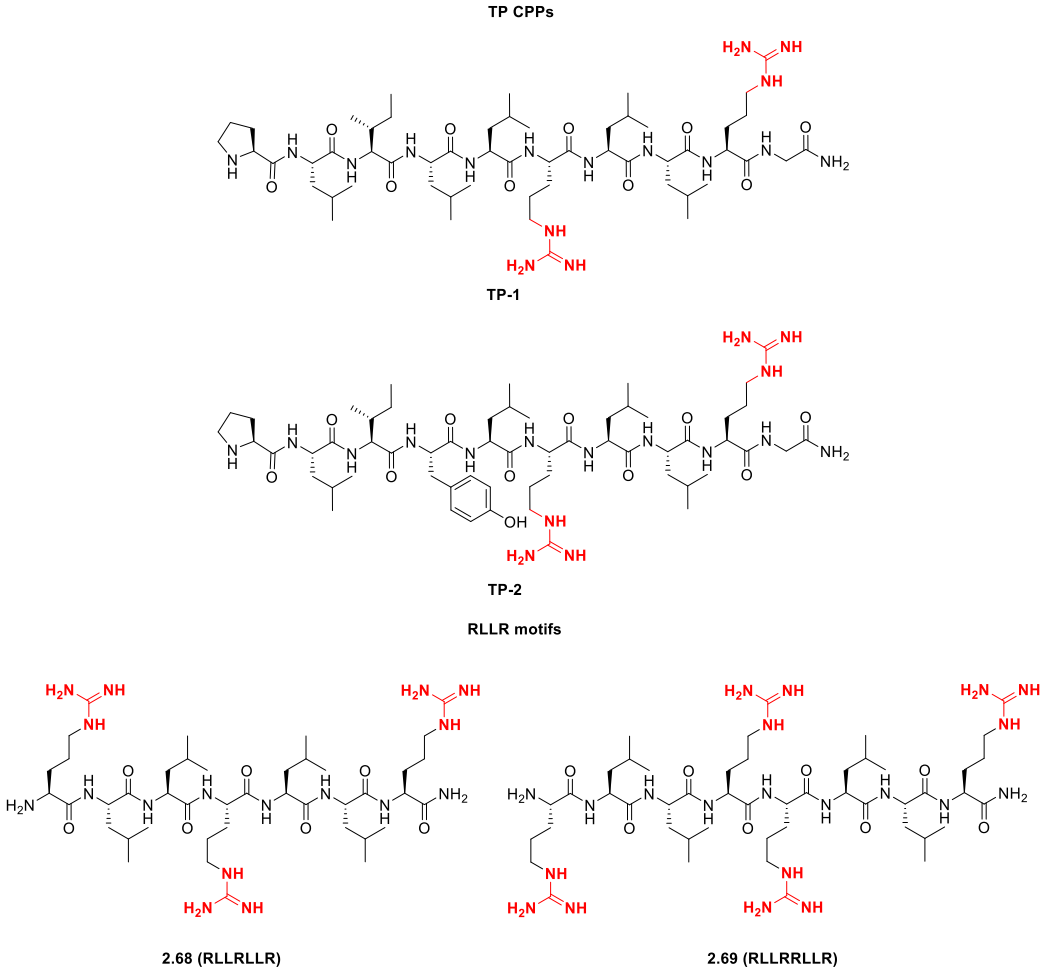


Figure 2.9. Structures of TP CPPs and RLLR motif peptides **2.68** and **2.69**. Arginine residues highlighted in red.

The TP CPPs also contain lipophilic residues at the *N*-terminus (PLILL and PLIYL) which may allow for insertion into a lipid membrane as reported by Wimley *et al.* for **TP-2**.[176] Amphipathicity of CPPs has been highlighted as a property that can enhance cellular uptake due to a balance in lipophilicity and hydrophilicity allowing binding and insertion into plasma membranes.[176] The amphipathic nature of the TP CPPs and the RLLR motifs could also be important for their mechanism of uptake. The helical wheels of the TP CPPs highlights both arginine groups on the same side of the helix (Fig 2.10A). Both

2.68 and **2.69** contain an arginine-rich, hydrophilic helical face balanced with a hydrophobic side of the helix (Fig 2.10B).

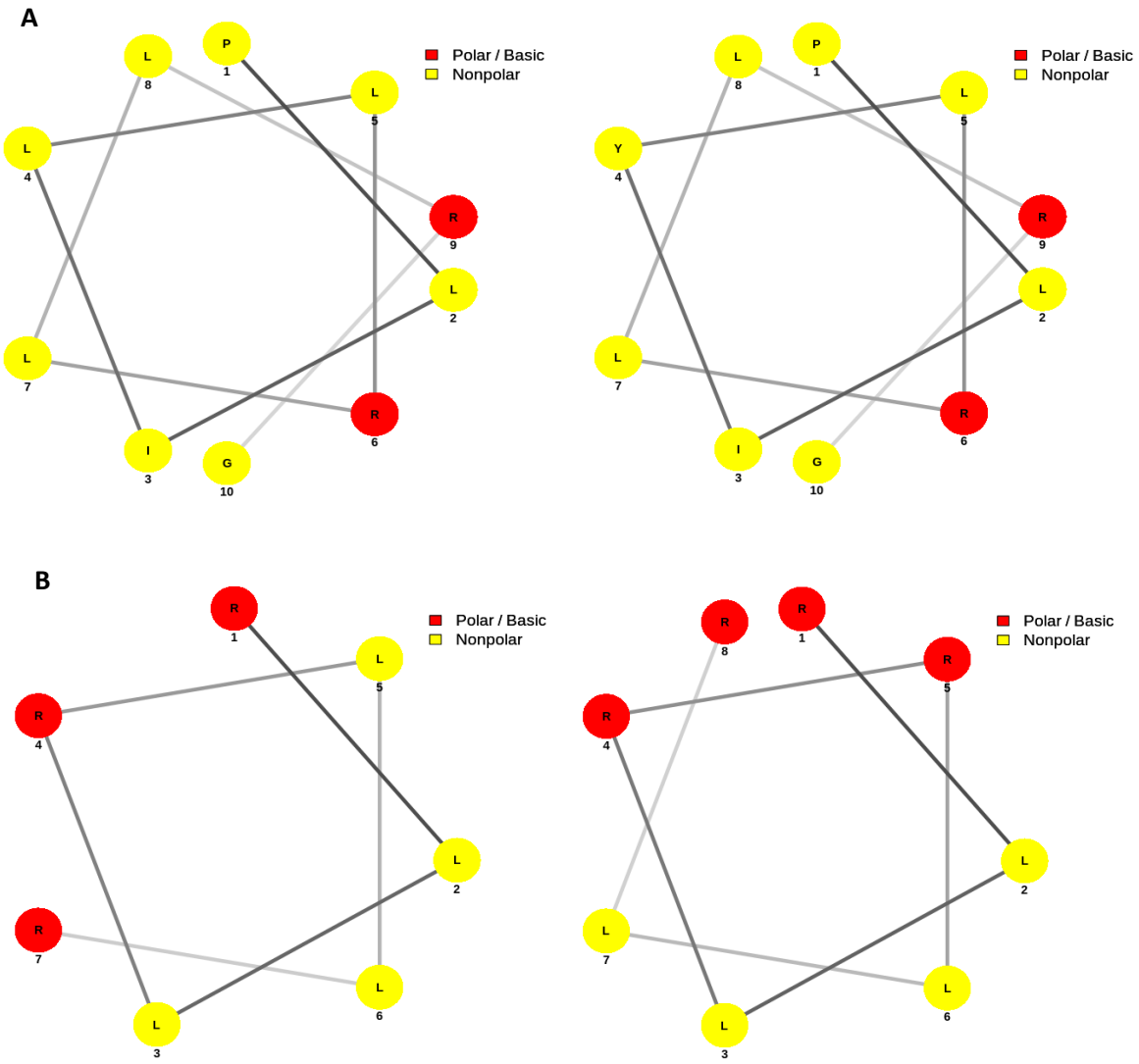


Figure 2.10. (A) Helical wheels of TP-1 and TP-2. (B) Helical wheels of **2.68** and **2.69**. Nonpolar residues highlighted in yellow and polar residues highlighted in red.

The ability of arginine to bind to the anionic species of the cell membrane could allow **2.68** and **2.69** to interact with cell surface. The hydrophobic leucine region of both peptides would aid in insertion of the lipid bilayer *via* hydrophobic interactions.

Of the techniques discussed in *section 2.1.2*, fluorescence monitoring of labelled peptides by flow cytometry allows for quantification of cellular uptake of CPPs in live cells. The method has many advantages as firstly, there is no need for the synthesis of additional standards required for mass spectrometry techniques. Secondly, no destructive techniques are required to analyse the cell interior unlike cell lysate experiments. Finally, using live cells allows for all possible energy dependant uptake

mechanisms to be investigated, unlike artificial membrane assays like PAMPA or CaCo-2. Based on these advantages fluorescent CPPs were to be synthesised for flow cytometry analysis.

The specific fluorophore used will inevitably change the properties of the CPP. Several fluorogenic species have been used to track and identify the TP peptides, including TAMRA **2.70**, carboxyfluorescein **2.71** and coumarin **2.72** (Fig 2.11).[79–81] Both **2.70** and **2.72** were attached to the TP peptides using a cysteine at the *C*-terminus to react with a maleimide on the fluorophore. When using **2.71** the CPP was coupled directly to the *N*-terminus forming an amide bond. Both maleimides **2.70** and **2.72** require the addition of a cysteine at the *C*-terminus for conjugation, whereas **2.71** could be conjugated directly to the peptide. Dye **2.70** exists as zwitterion at physiological pH whereas **2.71** is anionic. This fundamental change in overall charge has been shown to directly affect the permeability of the peptide conjugates which was reported to be greater using TAMRA based dyes than fluorescein based dyes.[177]

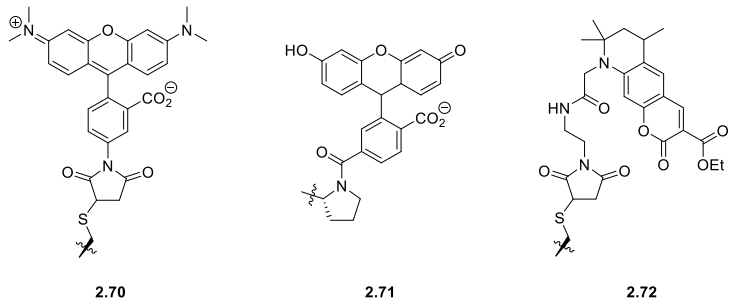
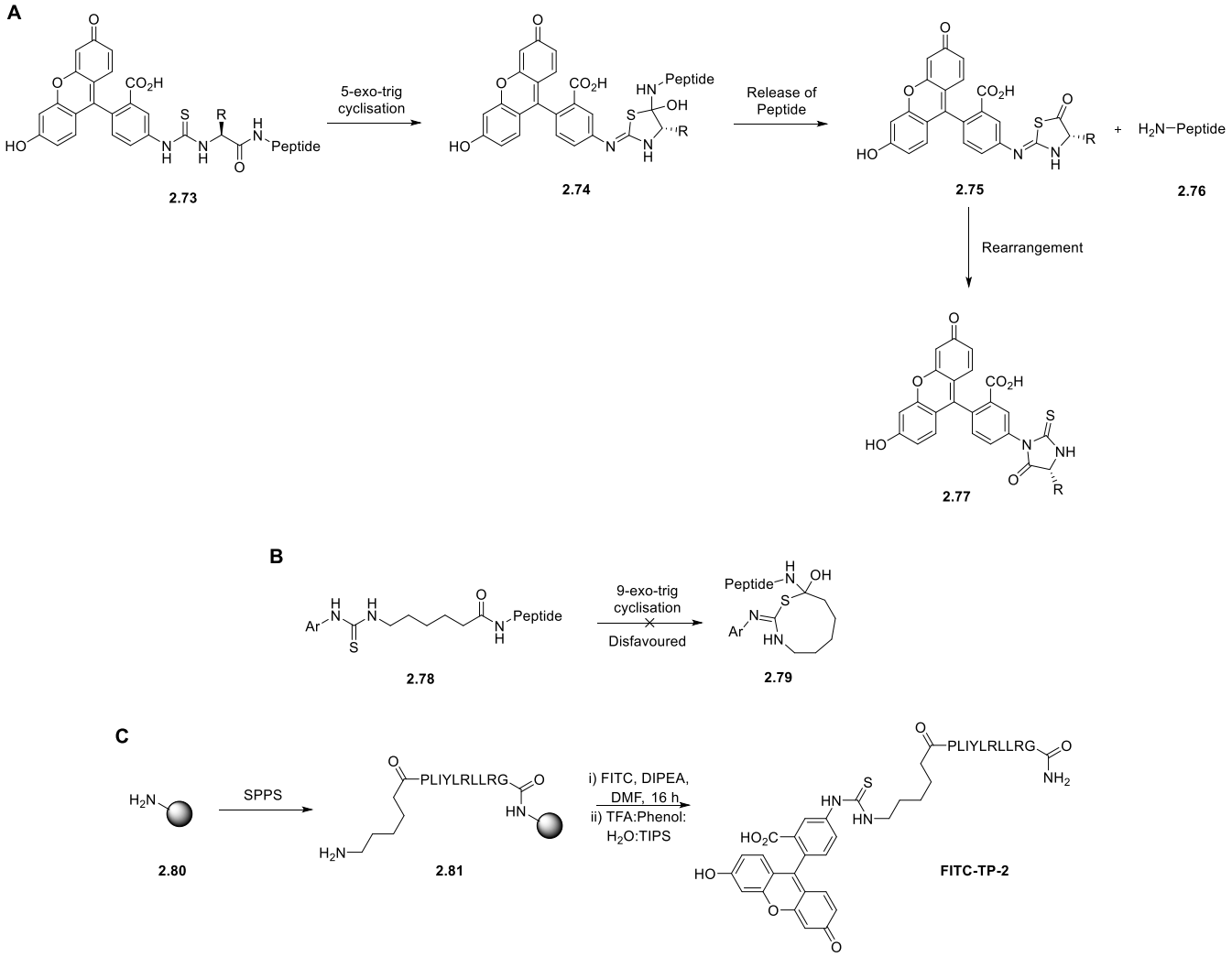


Figure 2.11. Structures of dye-CPP attachments previously used for TP CPPs.

Based on the increased difficulty of delivering fluorescein over TAMRA dyes, fluorescein isothiocyanate (FITC) was selected as a relevant and challenging marker to label CPPs for uptake studies. An isothiocyanate forms a thiourea linkage at the *N*-terminus of a corresponding peptide. However, it has been shown that thioureas attached directly to the *N*-terminal amine in a peptide can undergo an Edman degradation (Scheme 2.10A). The sulphur of thiourea **2.73** can attack the amide carbonyl of the adjacent amino acid residue in a 5-*exo*-trig cyclisation. The resulting thiazolinol **2.74** forms thiazolinone derivative **2.75** and subsequently releases the peptide **2.76** with a free amine at the *N*-terminus. Thiazolinone **2.75** finally undergoes a rearrangement to afford the thiohydatoin **2.77**. To avoid the Edman degradation during FITC labelling of a CPP, a spacer group can be used to stop the cyclisation step. Amino hexanoic acid (Ahx) has been widely used to this end and ensures that the Edman cyclisation would be the highly disfavoured 9-*exo*-trig and thus significantly reduces the loss of dye from the peptide.



Scheme 2.10. (A) Edman degradation of fluorescein thiourea peptide. (B) Ahx spacer used to prevent Edman degradation. (C) Synthesis route to FITC labelled TP-2.

The fluorescently-labelled CPPs were synthesised using a standard Fmoc strategy SPPS with HATU as the amino acid activating agent (*e.g.* scheme 2.10C). A rink amide resin was used in order to return an amide at the *C*-terminus for all CPPs. In addition, Ahx was used as a linker between the final residue on the peptide and the FITC fluorophore to reduce Edman degradation.

To accompany the TP peptides and RLLR motifs two arginine-rich CPPs, **Tat** and **Pip6**, were synthesised with analogous fluorescein labels to compare the effect of arginine content on toxicity and uptake of the peptides (Fig. 2.12).

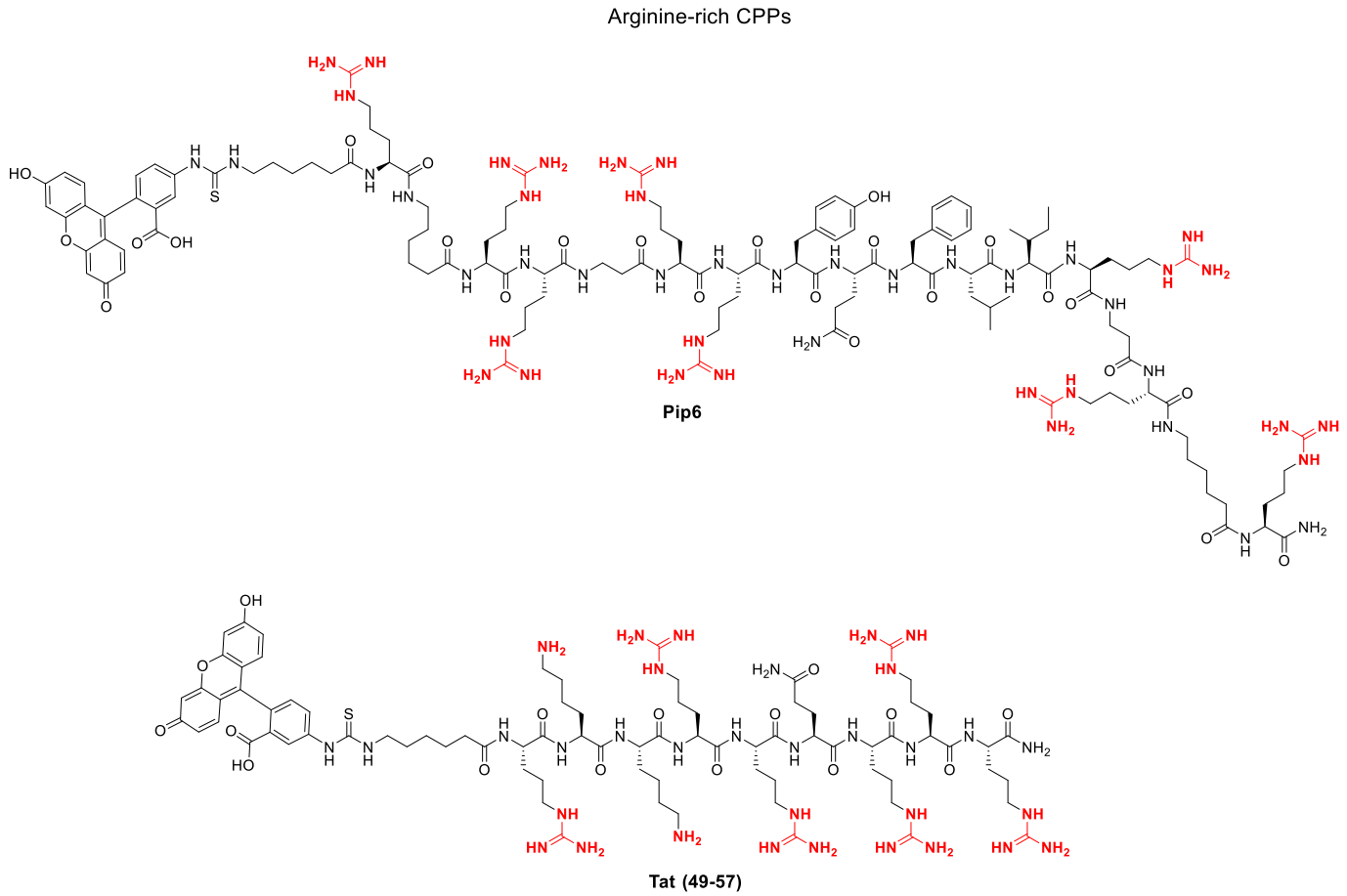


Figure 2.12. Fluorescently labelled arginine-rich CPPs, **Pip6** and **Tat**, that were synthesised using SPPS.

Pip6 is a CPP that has previously been used to delivery PMO cargos for the treatment of duchenne muscular dystrophy and so would provide a robust comparison to the TP and RLLR CPPs. Both labelled TP CPPs were isolated in overall yields of 5% and 7% after purification. The shorter labelled RLLR motif peptides, **2.80** and **2.81**, were isolated in 23% and 25% overall yield after purification (Fig 2.13, HPLC traces in appendix *chapter 2*).

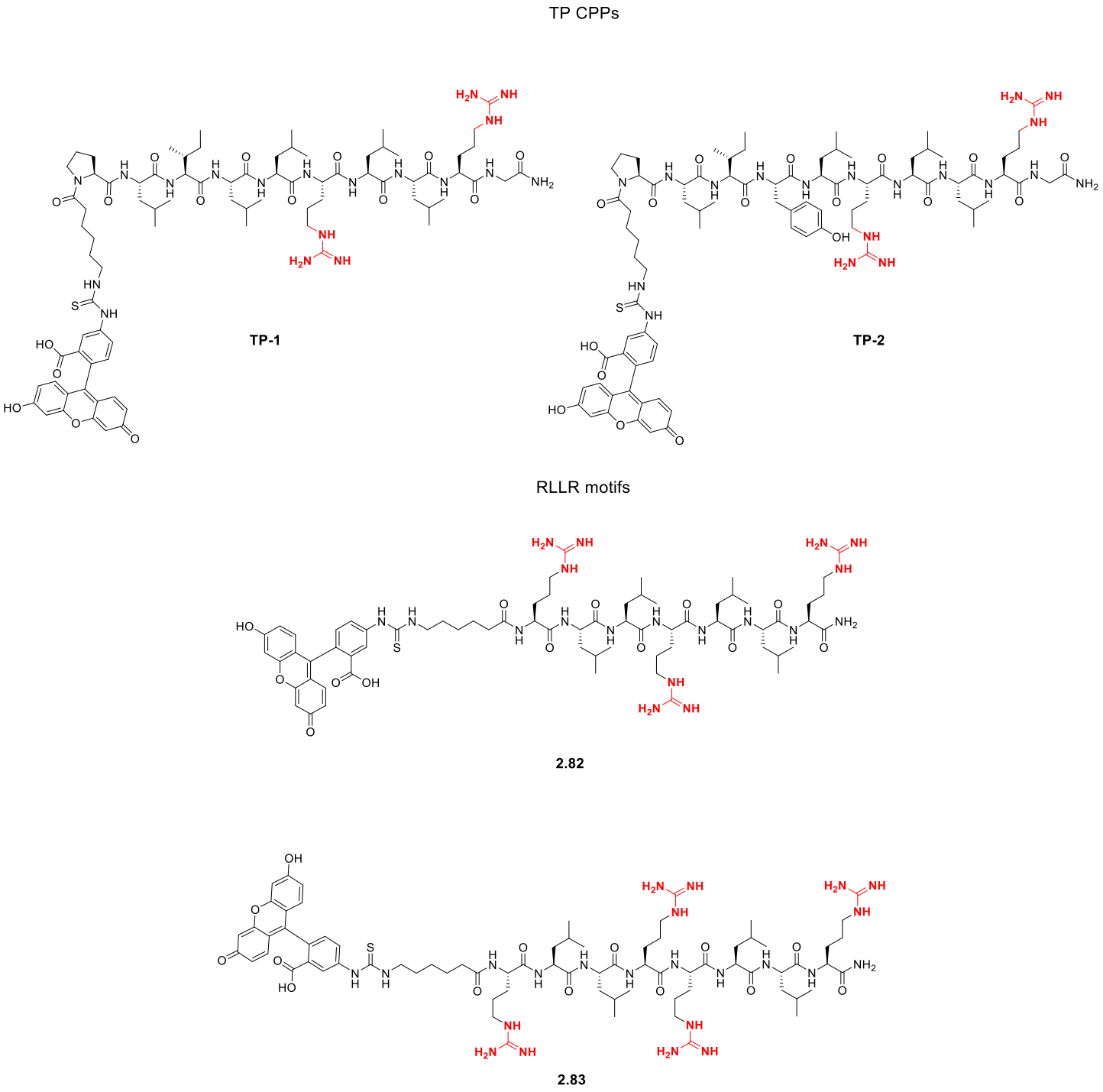


Figure 2.13. FITC labelled TP CPPs and RLLR motif peptides to be assessed for the toxicity and cell uptake.

2.3.2 Cell-viability with fluorescently-labelled CPPs

To test the hypothesis that a high arginine content in CPPs is associated with higher toxicity, the viability of several relevant cell lines was assessed using an alamarBlue assay with exposure to each of the labelled CPPs. AlamarBlue is permeable to mammalian cells but not fluorescent, however it is reduced by the metabolic environment of a living cell (NADPH and NADH can both reduce alamarBlue) to its fluorescent derivative.[178] The viability was first assessed in HeLa cells (Fig. 2.14A & B). Both **TP-1** and **TP-2**, containing only two arginine residues each, showed near 100% viability at 30 µM and close to 90% viability at 100 µM. Along with the TP peptides the first RLLR motif **2.82** displayed a similar viability profile in HeLa cells with slightly lower viability at 100 µM at near 80%. The common CPP **Tat**, which contains six arginine residues and an overall positive charge of 8, only showed a reduction in viability at 100 µM near 75%. The **Pip6** CPP, containing eight arginine residues, appeared to be the most toxic peptide with a cell viability of near 10% at 30 µM and less than 10% at 100 µM. Finally, the second RLLR motif **2.83** returned a higher cell viability than **Pip6** at 30 µM, but cell viability reduces to similar levels at 100 µM peptide.

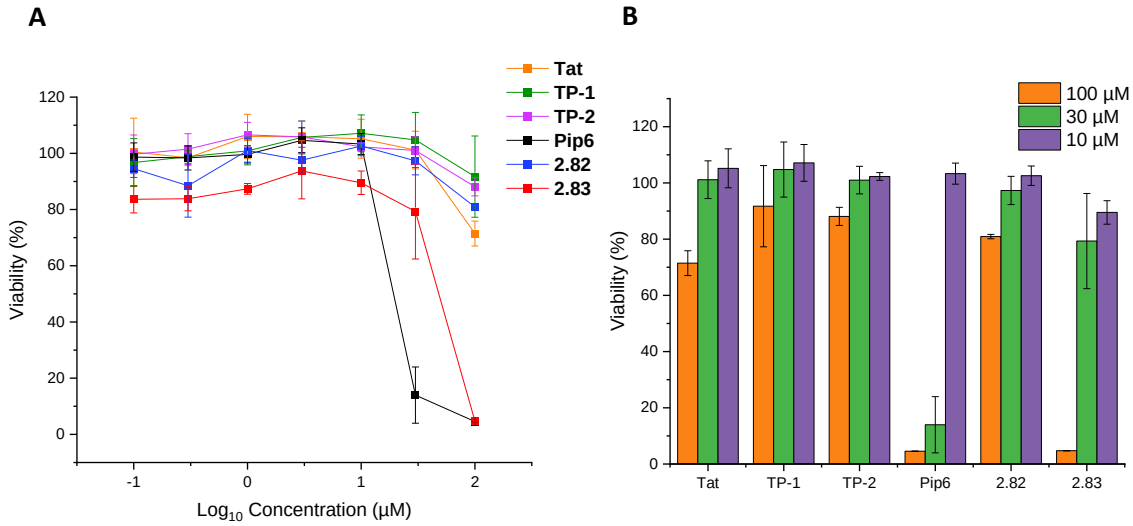


Figure 2.14. (A) Viability of Hela cells treated with CPPs using an alamarBlue assay after 24 h incubation at 37 °C. (B) Viability of Hela cells treated with CPPs at 100 µM, 30 µM and 10 µM.

To assess if the toxicity profiles of each peptide were consistent between different cell lines, viability was quantified of the osteoblast cell line U2OS using the same CPP treatments (Fig. 2.15A & B). As with HeLa cells, U2OS cells exhibited high viabilities near 100% at 30 µM and similarly at 100 µM when incubated with both TP peptides. Cells were viable at around 90% with both **Tat** and **2.82** even at 100 µM concentration. Similar to HeLa cells, U2OS cells had significantly lower viabilities with **Pip6** and **2.83** at 100 µM with cells only being around 10% viable with **Pip6**. At 30 µM **Pip6** showed a marginal improvement to 30% viability, though cells with **2.83** improved to 90% viable at 30 µM.

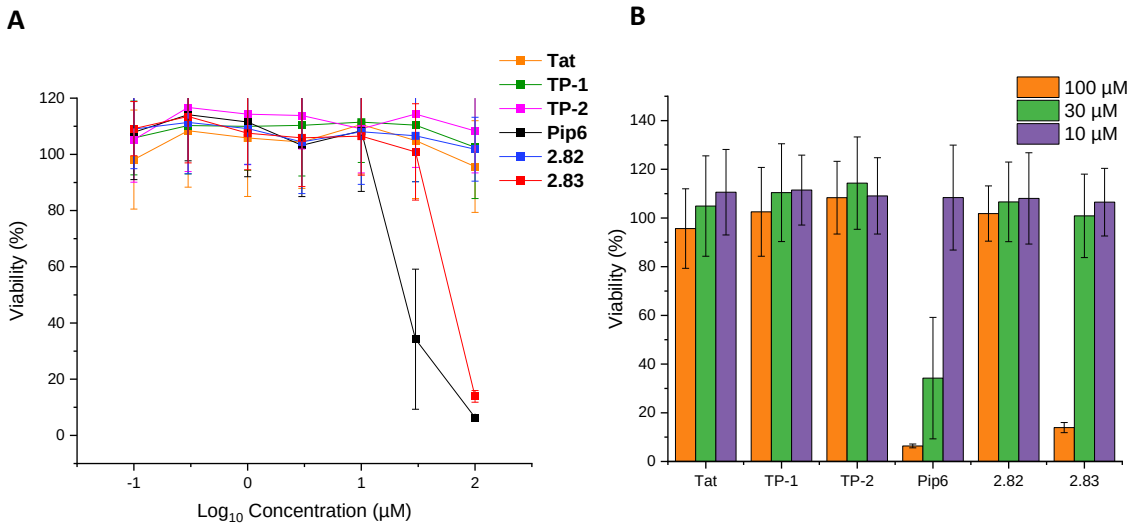


Figure 2.15. (A) Viability of U2OS cells treated with CPPs using an alamarBlue assay after 24 h incubation at 37 °C. (B) Viability of U2OS cells treated with CPPs at 100 µM, 30 µM and 10 µM.

From the viability experiments it can be seen that the most toxic peptides are **Pip6** and **2.83**, which appears to show that high charge and arginine content is not the only prevailing factor in CPP toxicity. This is demonstrated most notably by **Tat** as it was less toxic than **2.83** despite carrying a high charge and three more arginine residues. The toxicity profiles of **Pip6** and **2.83** are similar across both HeLa and U2OS cell lines with both having near 100% viability at 10 µM concentration. This suggests that any toxicity is more likely to occur *via* a general necrotic mechanism since an apoptotic mechanism would cause a dose response change in viability, not an increase in 80% viability from 30 µM to 10 µM, as observed in HeLa cells.

2.3.3 Uptake of fluorescently labelled peptides

Flow cytometry was utilised in order to quantify the level of uptake of each FITC labelled CPP. Based on the viability studies, a 5 µM concentration was selected for the uptake experiments in order to minimise any toxicity to each cell line. The first experiment conducted was ran over 4 h in HeLa cells to compare the uptake of both **TP-1/TP-2** against the arginine-rich CPPs **Pip6** and **Tat** (Fig. 2.16 A & B). Intriguingly, both TP peptides returned higher median fluorescence intensities (MFI) than both **Tat** and **Pip6** over 4 h, with **TP-2** returning the highest MFI value. This suggests that both TP peptides, in particular **TP-2**, are taken up and remain within the cell at higher concentrations than both the arginine rich CPPs after 4 h. Furthermore, both RLLR motif peptides **2.82** and **2.83** returned relatively low MFI values compared to all other CPPs, suggesting somewhat low cellular penetration for both peptides. This implies the RLLR unit within both TP peptides is not the principle contributor to their cell permeability. Finally, the low uptake of **2.82** and **2.83** also highlights that increasing arginine content within a CPP does guarantee improved uptake and cellular retention, as both TP CPPs had higher MFIs.

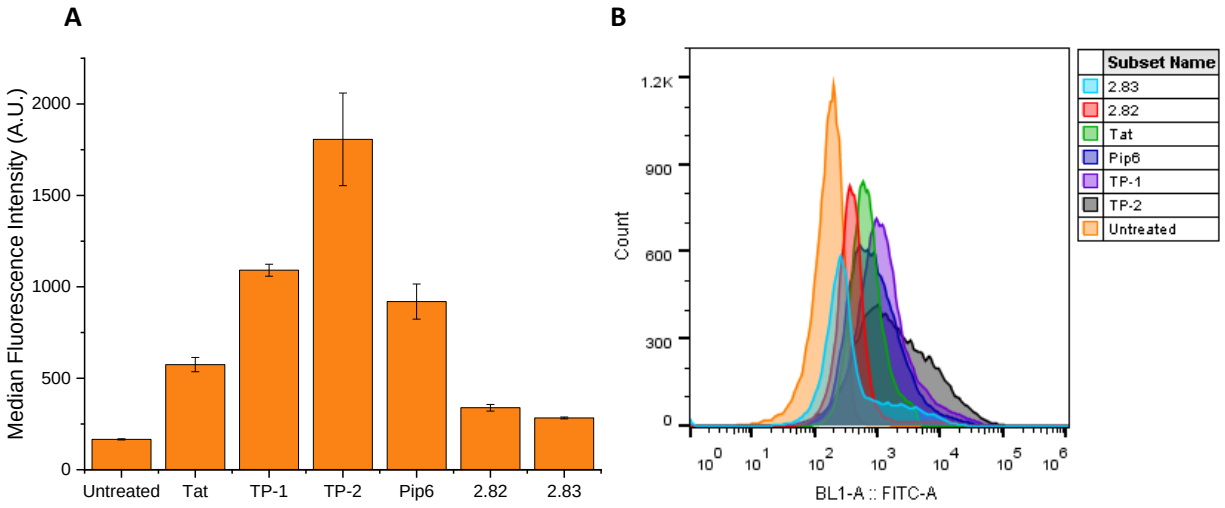


Figure 2.16. (A) Comparison of uptake of FITC labelled CPPs at 5 µM after incubation at 37 °C for 4 h in HeLa cells. Error bars are mean ± SD of triplicate experiments. (B) Overlay of histograms of CPPs after 4 h incubation at 37 °C with HeLa cells.

To further understand the mechanism of cellular uptake of the TP and arginine rich CPPs, the experiments were performed over 1 h (Fig 2.17 A & B). In this shorter time period both arginine-rich CPPs gave significantly higher MFI than over 4 h with **Pip6** showing over a 4-fold decrease in fluorescence between 1 h and 4 h. In contrast, both TP peptides showed relatively similar fluorescence values between 1 h and 4 h with a small but significant decrease over time observed with **TP-1** whereas a small increase over time was observed with **TP-2**. The reduction in MFI over time for both **Tat** and **Pip6** indicates that both CPPs are initially taken up at a higher rate than both TP peptides, but are either exported from the cell or degraded at a significantly higher rate. The less pronounced changes over time

seen with TP peptides suggests that they are potentially less susceptible to degradation and export from the cell but are not taken up in as great a quantity as the **Pip6** peptide over 1 h.

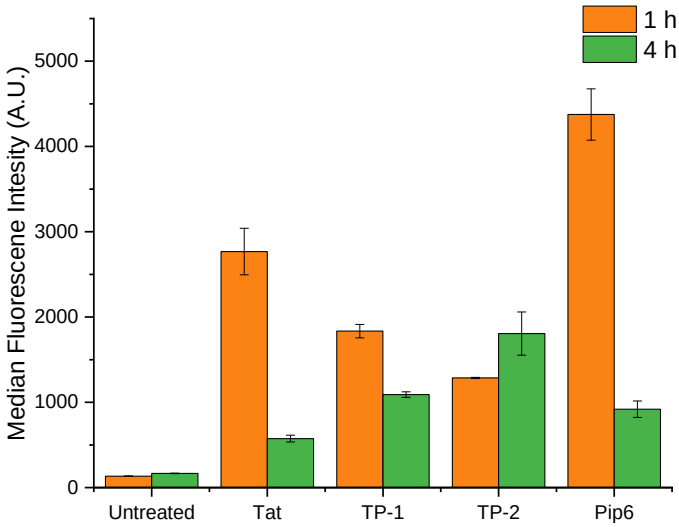


Figure 2.17. Comparison in uptake of CPPs after incubation at 37 °C for 1 h and 4 h with HeLa cells. Error bars are mean ± SD of triplicate experiments.

As these peptides could be entering cells by either a direct mechanism or *via* an endocytotic mechanism (mechanisms discussed in *chapter 1 sections 1.3.3* and *1.6.4*) the incubation of peptides with cells was conducted at 4 °C to depress all energy dependant processes within the cell (Fig 2.18 A, B & C). First looking at the TP peptides, there is reduction in MFI of both peptides with around a 50% reduction in fluorescence for **TP-1** and a near an 80% reduction with **TP-2** (Fig 2.18A). These results highlight that the uptake of both TP peptides in HeLa cells is *via* an energy dependant mechanism, such as endocytosis. This uptake of **TP-1** has been identified as endocytosis in HeLa cells previously, though no specific mechanism was established.[80] When comparing all CPPs a significant increase in MFI is observed for both **Tat** and **Pip6** upon a decrease in incubation temperature (Fig 2.18B). The increase in MFI for both arginine rich CPPs first highlights that the uptake of both peptides is *via* an energy independent mechanism such as direct translocation. Secondly, as the low temperature will likely disable any proteolytic and exocytotic mechanisms this could prevent the degradation and expulsion of each peptide, hence a much higher intracellular concentration of CPP is maintained within the cell over a 4 h time period.

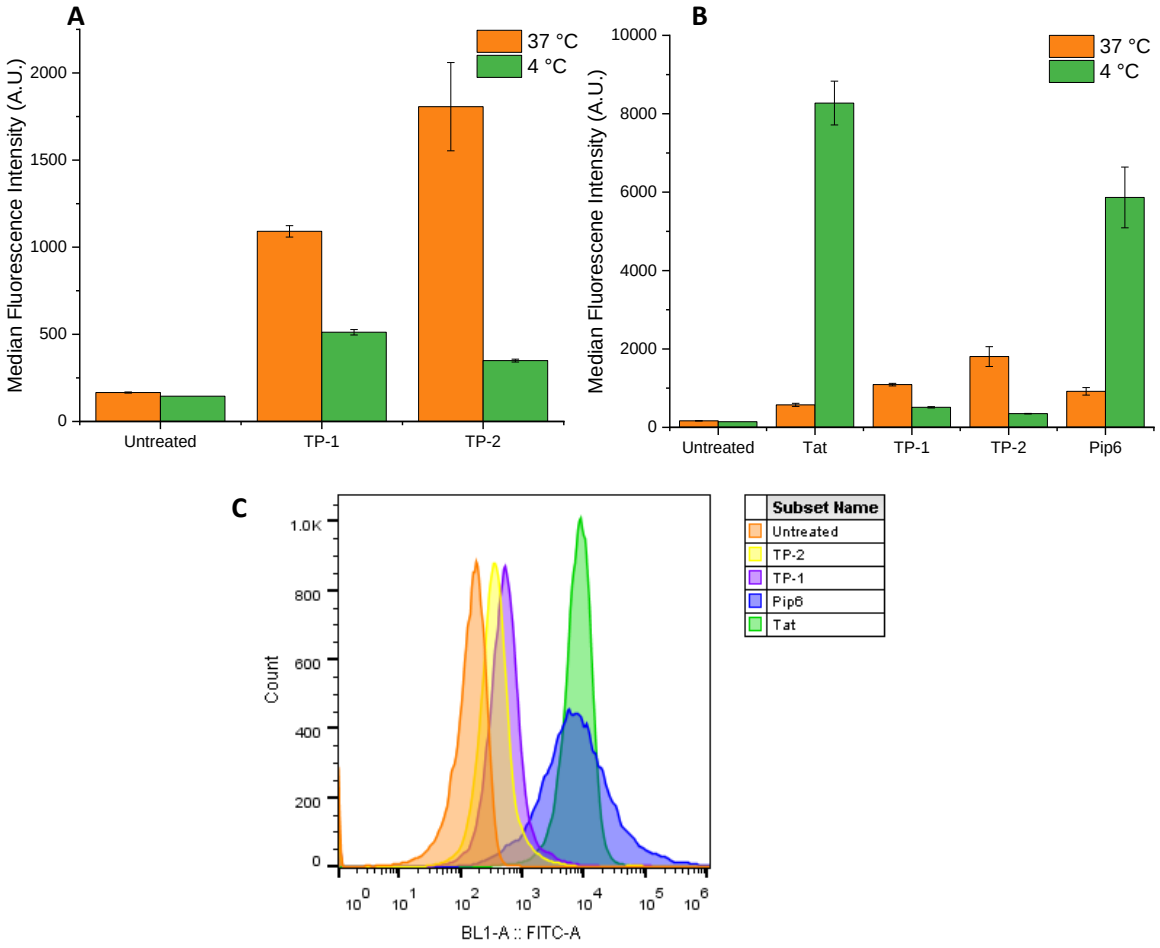


Figure 2.18. Comparative analyses of cell uptake of (A) TP peptides after incubation at 37 °C and 4 °C for 4 h, and (B) TP peptides with **Tat** and **Pip6** after incubation at 37 °C and 4 °C for 4 h. Error bars are mean ± SD of triplicate experiments. (C) Overlay of histograms of CPPs after 4 h incubation at 4 °C with HeLa cells.

From the 4 °C experiments it was demonstrated that both TP peptides are taken up *via* an energy dependant process, such as endocytosis. To investigate the specific mechanisms, three endocytosis inhibitors were investigated: Chlorpromazine (CPZ), Nystatin (Nys) and Wortmannin (Wort) (Fig 2.19). Chlorpromazine causes assembly of adaptor proteins and clathrin at endosomal membranes, which in-turn leads to depletion of clathrin at the cell surface. This effectively stops both clathrin-mediated endocytosis and endosomal recycling to the cell surface.[179] Nystatin is a polyene that binds to cholesterol within the plasma membrane and causes disassembly of caveolin at the cell surface.[180] Finally, Wortmannin inhibits the PI3 kinases (PI3K) which has been demonstrated to affect macropinosome closure and actin reorganisation to cause membrane ruffling.[181] Each of these inhibitors were successfully employed in a H358 cell line by Arora *et al.* to identify the endocytotic mechanism of uptake for peptidomimetics.[182]

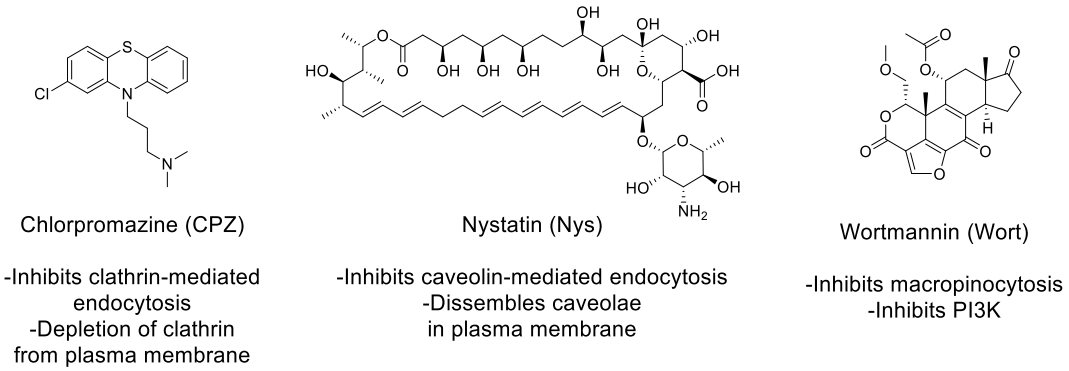


Figure 2.19. Structures of endocytosis inhibitor Chlorpromazine, Nystatin and Wortmannin.

To investigate clathrin-mediated endocytosis cells were pre-incubated with chlorpromazine before the addition of each CPP and then analysed by flow cytometry after 4 h incubation (Fig 2.20A & B). When comparing the MFI values of the TP peptides at 4 °C, 37 °C and with CPZ, a small increase in fluorescence was observed with **TP-1** upon addition of CPZ. Increases in incorporation upon inhibition of clathrin-mediated endocytosis have been reported in various cell lines, where upregulation of other endocytotic pathways can occur, which could explain subtle fluorescent increases upon treatment of CPZ.[183] For **TP-2** there is a minor reduction but not at the same level as with 4 °C incubation. As both TP peptide MFI values were not reduced significantly, and in the case of **TP-1** a small increase, this implies that the mechanism of uptake for the TP CPPs is not *via* clathrin-mediated endocytosis. For the **Pip6** peptide CPZ incubation returned an increase in MFI compared to 37 °C alone, similar to the increase seen at 4 °C. Taking into account the previous results and significant increase when suppressing energy dependant mechanisms, potentially inhibition of clathrin endosome recycling could minimise the expulsion of the **Pip6** peptide but that clathrin-mediated endocytosis does not play a role in uptake. With the **Tat** peptide a similar MFI value was observed with CPZ and without at 37 °C indicating that clathrin endosome recycling does not play a role in either uptake or expulsion of this CPP.

Caveolin-mediated endocytosis was probed by pre-incubating the cells with Nystatin before the addition of FITC CPP. Looking at the TP peptides, neither showed a reduction in fluorescence with the caveolin uptake pathway inhibited. The slight increase observed in fluorescence signals for both the TP CPPs in the nystatin experiment can be attributed to Nystatin upregulation of non-caveolae dependant endocytosis, as reported by Lu *et al.* for the uptake of the endostatin protein.[184] Both arginine rich CPP experiments, **Tat** and **Pip6**, showed no significant change in fluorescence upon incubation with nystatin compared to the standard 37 °C. The minimal fluorescence change observed for **Tat** and **Pip6** upon nystatin treatment further confirms uptake and cargo expulsion for these CPPs does not proceed *via* the caveole-mediated endocytosis pathway.

The final inhibitor used was wortmannin, to inhibit macropinocytosis. **TP-1**, as with CPZ and nystatin, showed no reduction in fluorescence compared to the 37 °C. The fact that no fluorescence reduction

occurs for **TP-1** with wortmannin signifies that macropinocytosis is not a major pathway for its uptake into cells of **TP-1**. Looking at **TP-2**, there was a noticeable reduction in fluorescence upon incubation with wortmannin, though not to the same extent as with 4 °C incubation. The reduction in **TP-2** fluorescence indicates that macropinocytosis may play a role in the uptake of **TP-2** though it is not the major pathway of cell uptake. For **Tat**, like with CPZ and nystatin, no significant change in fluorescence was observed confirming the uptake of **Tat** to be *via* a direct mechanism in the HeLa cell line. Wortmannin incubation with **Pip6** demonstrated a 2-fold increase in fluorescence which could signify macropinosome recycling is another, though less significant, pathway for expulsion of the **Pip6** peptide.

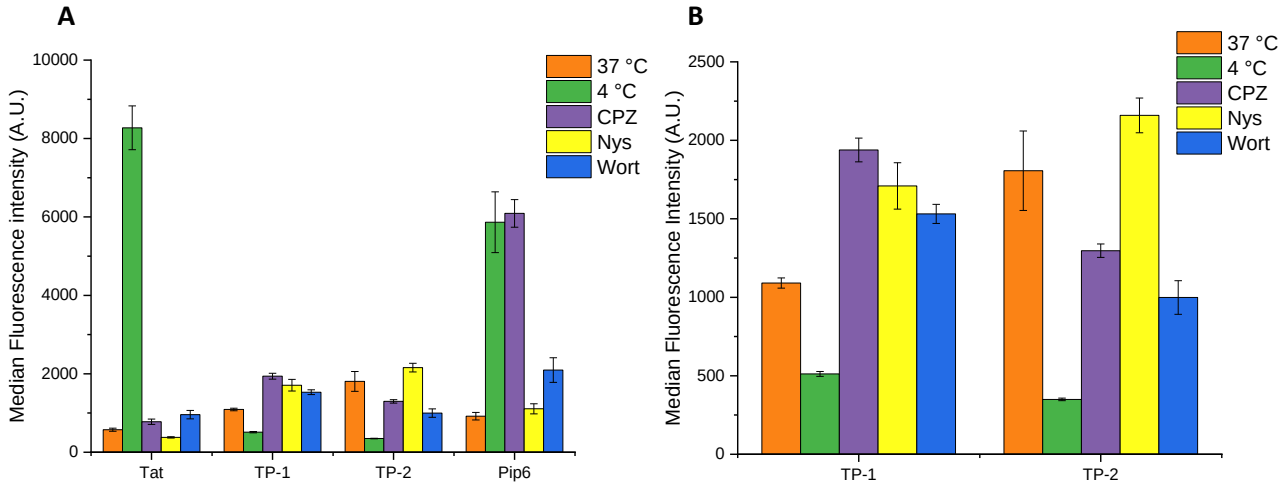


Figure 2.20. Comparative analyses of (A) CPPs after incubation at 37 °C, 4 °C and with CPZ, Nystatin and Wortmannin for 4 h in HeLa cells, and (B) TP CPPs with endocytotic inhibitors. Error bars are mean ± SD of triplicate experiments.

The U2OS cell line was assessed in parallel to quantify the penetration of each of the CPPs and to identify any differences in uptake profiles across a different cell line. First, looking at a comparison between RLLR motifs **2.82** and **2.83** and the TP peptides, it can be seen that both RLLR motif peptides are taken up more poorly than both TP peptides in U2OS cells after 4 h (Fig. 2.21A). The low uptake of the RLLR peptides is consistent with HeLa cell observation, again highlighting that uptake of both TP CPPs is not solely dependent on the RLLR motif within each peptide. Next, to compare the TP peptides with **Tat** and **Pip6** both 1 h and 4 h incubation experiments were carried out. It can be seen that over 4 h both TP peptides and **Pip6** return similar MFI values to each other whereas the **Tat** CPP has the lowest concentration in U2OS cells the least after 4 h (Fig 2.21B). Upon incubation over 1 h a similar result to HeLa cells observed, where both arginine-rich CPPs show a significantly higher incorporation. The higher fluorescence after 1 h for both **Tat** and **Pip6** indicates that, like in HeLa cells, both arginine-rich CPPs are potentially degraded and expelled from U2OS cells between 1 h and 4 h incubation. The TP peptides return similar MFI values after 1 h as with 4 h suggesting they are not expelled significantly from U2OS cells after a further 3 h incubation.

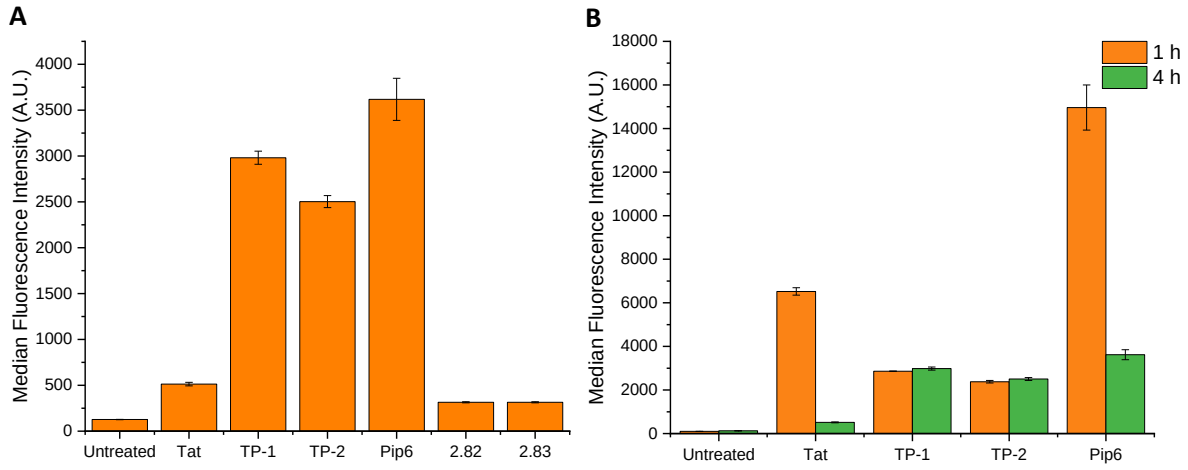


Figure 2.21. Comparative analyses of (A) TP, arginine-rich and RLLR CPPs after incubation at 37 °C for 4 h in U2OS cells, and (B) CPPs after incubation at 37 °C for 1 h and 4 h with U2OS cells. Error bars are mean ± SD of triplicate experiments.

Assessing the mechanisms of uptake of the TP peptides and arginine-rich CPPs was conducted again with 4 °C, CPZ, nystatin and wortmannin experiments (Fig 2.22). When comparing fluorescence values of the TP peptides at 4 °C, 37 °C and with CPZ it can be seen that both TP values were not reduced upon the addition of CPZ but were by incubation at 4 °C. A reduction in MFI of near 80% for both TP peptides at 4 °C implies that, like HeLa cells, an energy dependant mechanism is responsible for their cell uptake. As no reduction in fluorescence was observed with CPZ, a clathrin-mediated endocytotic mechanism can be ruled out for the uptake of TP CPPs in U2OS cells. The nystatin experiments show an increase in fluorescence for TP peptides, which again suggests potential upregulation of alternative endocytosis pathways but no uptake *via* caveolin-mediated endocytosis. Similarly, wortmannin incubation experiments showed only a small fluorescence reduction, indicating that macropinocytosis is not a major uptake pathway for either TP peptides in U2OS cells.

For the **Pip6** peptide, 4 °C and CPZ incubation returned a significant increase in MFI compared to 37 °C alone. Like HeLa cells, this could suggest that inhibition of clathrin endosome recycling reduces expulsion of the **Pip6** peptide but the uptake of the CPP occurs most likely *via* a direct translocation mechanism. With the **Tat** peptide, unlike HeLa cells, a 4-fold increase in MFI was observed with CPZ, though a 20-fold increase is seen after 4 °C. The changes seen with **Tat** suggest clathrin endosome recycling may play a role in expulsion of this CPP, though the major pathway of degradation and expulsion is not clathrin-mediated. Both the nystatin and wortmannin experiments returned similar fluorescence levels to 37 °C for the arginine-rich CPPs, implying that neither caveolae-mediated endocytosis or macropinocytosis play a key role in the uptake or expulsion of **Tat** and **Pip6** in U2OS cells.

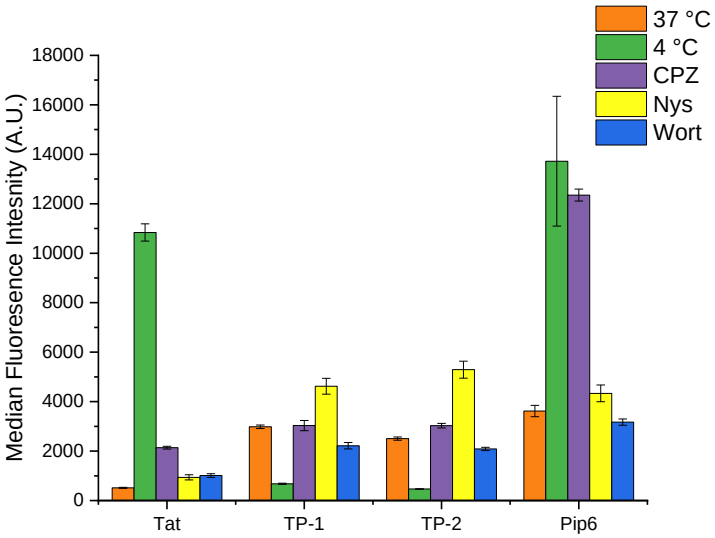


Figure 2.22. Comparison in uptake of CPPs after incubation at 37 °C, 4 °C and with CPZ, Nystatin and Wortmannin for 4 h with U2OS cells. Error bars are mean ± SD of triplicate experiments.

When looking at histograms of U2OS cells treated with TP peptides, distinct cell populations with higher MFI were observed (Fig 2.23A). A discrete second population is not formed of a distinct size or shape as the cells from this population display a random distribution across forward and side scatter variance (Fig 2.23B). This could indicate a small amount of cells uptake the CPPs at higher levels.

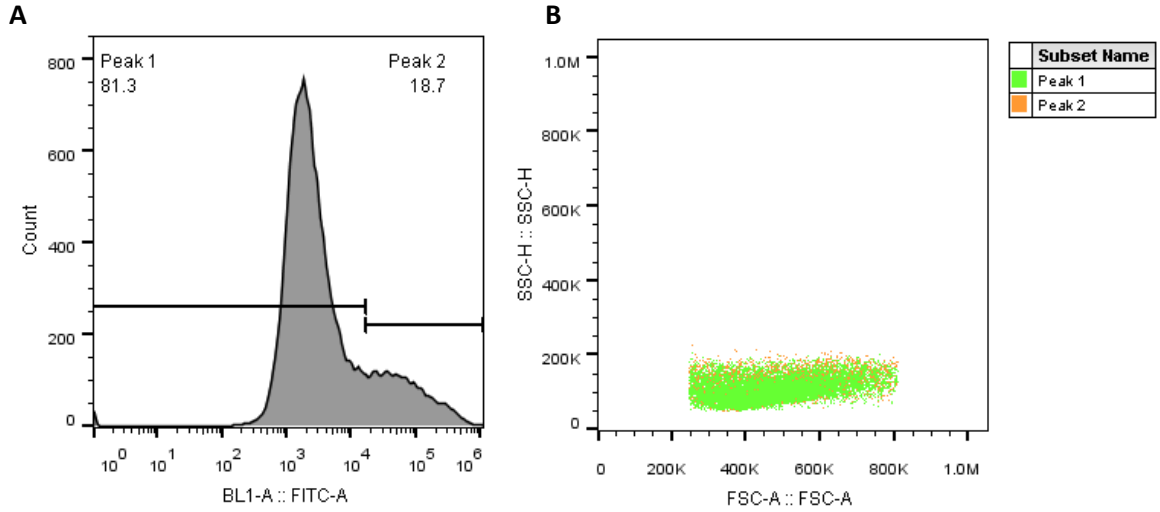


Figure 2.23. (A) Histogram of TP-2 treated U2OS cells after 4 h at 37 °C. (B) Overlay of cell populations from peaks 1 and 2 showing both have similar forward and side scatter distributions.

Overall similar trends in uptake for each CPP were observed across both cell lines, indicating a non-cell line specific mechanism of uptake for each CPP.

2.3.4 Sub-cellular localisation of CPPs

In order to investigate if there were any differences between the TP peptide and arginine rich CPPs in localisation within cells, confocal microscopy was used to image HeLa cells after treatment with fluorescein labelled CPPs. In addition to the CPPs, the HeLa cells were treated with the nuclear stain

Hoescht 33258 (Fig. 2.24). Hoescht 33258 is a bis-benzimide that binds to the minor groove of DNA and is fluorescent with an excitation/emission of $\lambda_{ex/em}$ = 365/465 nm.[185] The fluorescence characteristics of Hoescht 33258 mean it does not overlap with fluorescein and so allows identification of the nucleus as well as the subcellular localisation of the peptide. The use of a nuclear dye also allows for the discrimination between the nucleus and the cytoplasm of the cell, with any overlay of the two different colours indicating nuclear localisation.

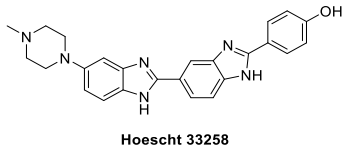


Figure 2.214. Structure of Hoescht 33258.

Looking first at the arginine-rich CPPs, **Tat** and **Pip6**, after 4 h incubation it can be seen that **Tat** is generally nuclear excluded with no specific overlay with the nucleus (Fig 2.25A). In addition to being excluded from the nucleus the peptide appears as small round puncta throughout the cytoplasm. Round puncta in cells has often been associated with endosomes, however it is unlikely to be endosomal entrapment from cell uptake as flow cytometry suggested a non-energy dependant uptake mechanism for **Tat**.[97,186] Puncta in HeLa cells have been previously observed with **Tat**.[105,187] A potential explanation for the appearance of puncta is that these are vesicles associated with the expulsion of **Tat** from the cell. **Pip6** appeared to have some overlay with the Hoescht stain suggesting that **Pip6** was localised within the nucleus along with some general cytoplasm location (Fig 2.25B). The small amount of nuclear localisation of **Pip6** could be associated with its high cationic charge and thus strong electrostatic interactions with anionic DNA allow it to be retained in the nucleus to some degree.

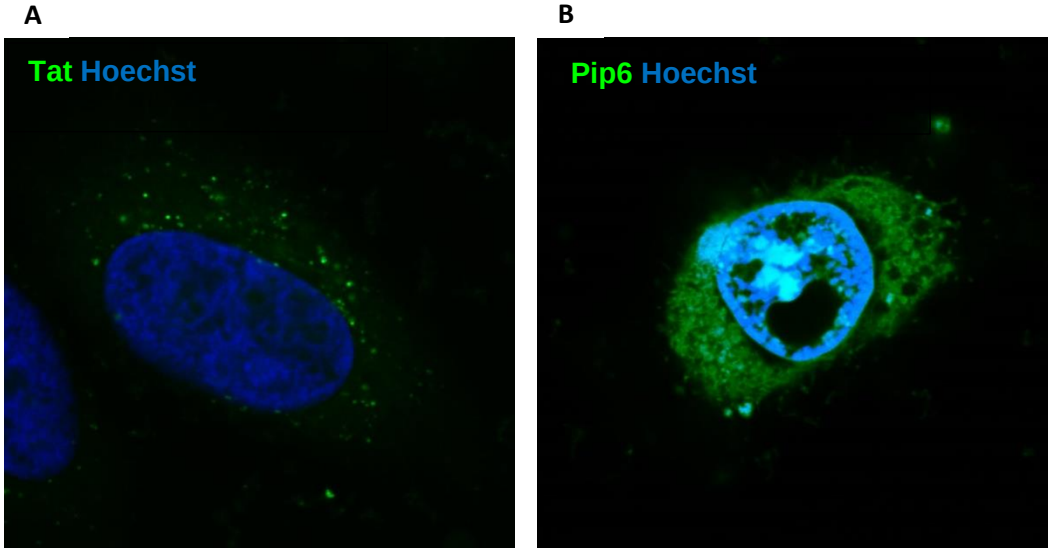


Figure 2.25. HeLa cells treated with (A) **Tat** and Hoechst 33258. (B) **Pip6** and Hoescht 33258.

Looking at both the TP peptides it can first be noted that there is no clear nuclear overlay with either **TP-1** or **TP-2** (Fig. 2.26A & B). However, there does appear to be some green fluorescence in the nucleus, though this is not as bright as the rest of the cytoplasm (Fig. 2.26C). The images indicate that both TP peptides a present throughout the cytoplasm and do not appear as any puncta, therefore suggesting the TP CPPs can escape any potential vesicle they use during uptake. They also appear to permeate the nucleus though are not present in the same level as in the cytoplasm.

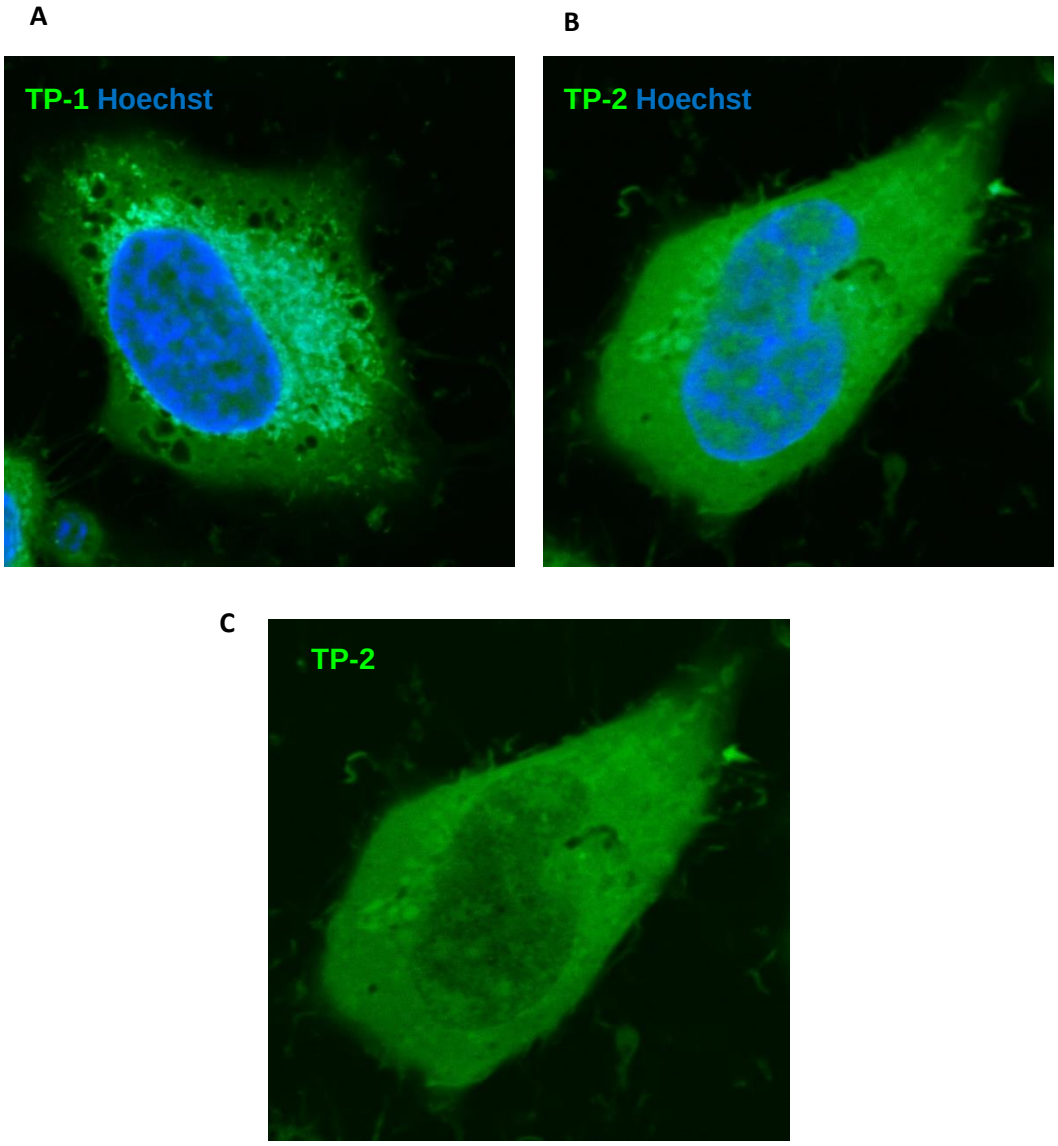


Figure 2.26. HeLa cell treated with (A) **TP-1** and Hoescht 33258. (B) **TP-2** and Hoescht 33258. (C) **TP-2** showing lower concentration in the nucleus.

2.4 Summary of Chapter 2

Fluorescein labelled CPPs were synthesised by SPPS to investigate their cell uptake and viability in both HeLa and U2OS cells. Viability was first investigated across the two different cell lines with both arginine-rich and low arginine containing CPPs. Across both HeLa and U2OS cell lines, the established arginine rich CPP **Pip6** showed greater toxicity than both low arginine containing CPPs, **TP-1** and **TP-2**, with less than 30% viable cells at 30 μM peptide.

The uptake of the CPPs was then assessed at a non-lethal concentration of peptide. The RLLR motif contained within both TP peptides was assessed for its cell permeability through the analysis of peptides **2.82** and **2.83**. It was identified in both cell lines that simple repeating units of RLLR is not sufficient to achieve good cell permeability. The RLLR motif within the TP peptides is not the sole factor that confers their cell permeability.

Across both cell lines it was observed that both TP peptides had similar, or even higher, concentrations to **Pip6** or **Tat** within cells after 4 h incubation. Upon further timed experiments it was identified that both arginine-rich CPPs are taken up at greater levels than the TP CPPs after 1 h. This appears to indicate that **Pip6** and **Tat** are degraded and/or expelled from the cell significantly more than the TP CPPs after 4 h.

Further uptake inhibition experiments identified that both TP CPPs are taken up *via* an energy-dependant mechanism, through the major pathway is not one of the three main endocytosis mechanisms (clathrin-mediated endocytosis, caveolae-mediated endocytosis and macropinocytosis). The arginine-rich CPPs enter cells *via* an energy independent mechanism. The expulsion of **Pip6** observed over time could be dependent on clathrin endosome recycling as inhibition of this process gave an increased level of **Pip6** cargo within both cell lines.

The TP CPPs also appear to localise in the cytoplasm of HeLa cells and thus escape any endosomes involved in their cell uptake.

.

2.5 Experimental

2.5.1 Peptide synthesis

Peptide synthesis was completed on an automated Tribute® peptide synthesiser with an IntelliSynth UV-monitoring system and feedback control system. Rink amide resin (100-200 mesh), Fmoc-Phe-OH, Fmoc-Gly-OH, Fmoc-Tyr(O^{t}Bu)-OH, Fmoc-Pro-OH, Fmoc-Leu-OH, Fmoc-Ile-OH, Fmoc-Gln(Trt)-OH, Fmoc-Arg(Pbf)-OH and Fmoc-εAhx-OH were purchased from Merck Millipore and used without further purification. Fluorescein 5-isothiocyanate (FITC) was purchased from sigma-aldrich and used without further purification.

2.5.2 Analytical HPLC

RP-HPLC was carried out using Aeris 3.6 µm RP 250 × 4.6 mm widepore XB C18 column using a DIONEX 3000 series HPLC equipped with a VWD3400 photodiode array detector. Purifications were performed using water (0.1% TFA) as Solvent A and acetonitrile (0.1% TFA) as Solvent B and were run at a flow rate of 1.0 mL/min. For HPLC traces see appendix *chapter 2*.

Analytical RP-HPLC method A:

Absorbance detection was set to 220 nm.

Table 2.1. Gradient used for analytical RP-HPLC method A

Time (min)	Solvent A	Solvent B
0	95%	5%
5	95%	5%
50	10%	90%
55	10%	90%
60	95%	5%

2.5.3 Purification of products

Semi-preparative reversed-phase HPLC purification was carried out on a kinetex 5 µm RP 150 x 21.2 mm XB C18 column using a DIONEX 3000 series HPLC equipped with a VWD3400 variable wavelength detector. Purifications were performed using water (0.1% TFA) as Solvent A and acetonitrile (0.1% TFA) as Solvent B and were run at a flow rate of 12.0 mL/min.

RP-HPLC method A:

Absorbance detection was set to 220 nm.

Table 2.2. Gradient used for RP-HPLC method A.

Time (min)	Solvent A	Solvent B
0	95%	5%
5	95%	5%
45	10%	90%
50	10%	90%
53	95%	5%

2.5.4 Mass spectrometry

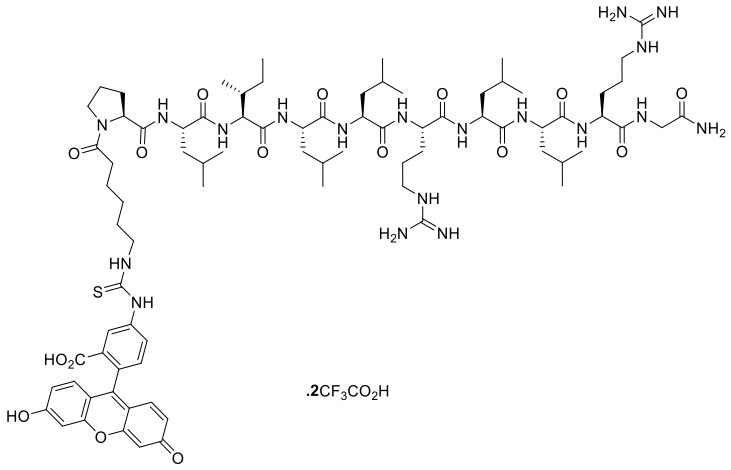
High-resolution mass spectra were recorded on a Bruker Micro TOF II at the university of Edinburgh.

2.5.5 Peptide characterisations

General peptide synthesis protocol A

On a Tribute® solid phase peptide synthesiser, Rink amide resin (0.15 mmol) was swelled in DCM (5.00 mL) for 30 min. Two 10 min agitations with 20% (v/v) solution of piperidine in DMF (5.00 mL) was used to deprotect terminal Fmoc groups. Activation of Fmoc protected amino acids (0.45 mmol, 3.00 equiv) was achieved using HATU (0.38 mmol, 2.50 eq) and 0.5 M DIPEA in DMF (5.00 mL). The reaction mixture of the activated esters of Fmoc amino acids were then added sequentially to the resin and agitated for 20 min at room temperature. Upon final coupling of appropriate residue the resin was washed with DCM (5.00 mL) and dried under N_2. The resin was then removed from the automated synthesiser and manually swelled in DCM (5.00 mL) for 30 min. After swelling the resin was washed with DMF (4 × 2.00 mL) and two 10 min agitations with 20% (v/v) solution of piperidine in DMF (5.00 mL) was used to deprotect the terminal Fmoc group. Separately in a 5.00 mL glass vial FITC (117 mg, 0.30 mmol) was dissolved in DMF (1.50 mL) and DIPEA (183 µL, 1.05 mmol) was added to form a bright red solution. The red FITC solution was then added to the resin, protected from light and agitated for 16 h at room temperature. After the 16 h agitation the resin was washed with DMF (4 × 2.00 mL) before being washed with MeOH (4 × 2.00 mL) and finally DCM (4 × 2.00 mL). The peptide was then cleaved from the resin using a mixture of TFA:phenol:water:TIPS (90:5:5:2, 3.00 mL) and agitated for 4 h. The cleavage solution was then filtered from the resin and the peptide precipitated using cold Et_2O (30 mL, -20˚C) to afford an orange precipitate. The solid was washed with Et_2O (3 × 10 mL) to remove excess TFA and the resulting orange solid was purified by RP-HPLC method A.

FITC – XPLILLRLLRG (**TP-1**)

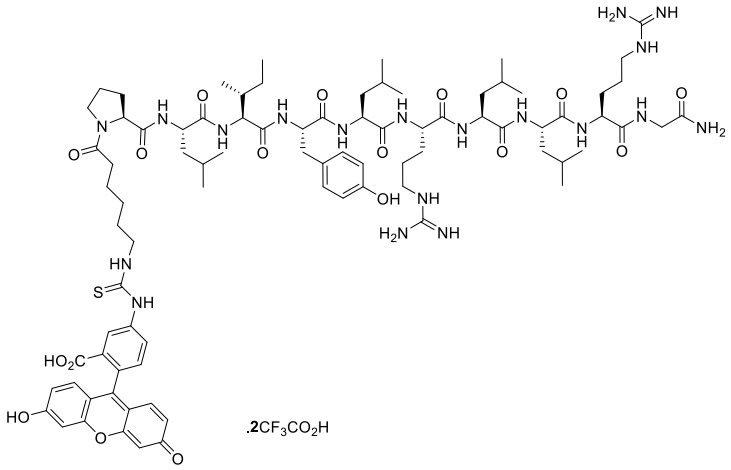


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (14 mg, 5 %).

HRMS: $C_{82}H_{127}N_{19}O_{16}S^{2+}$ Theoretical: 832.9709 Observed: 832.9710.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 31.3 min, 96%.

FITC – XPLIYLRLLRG (**TP-2**)

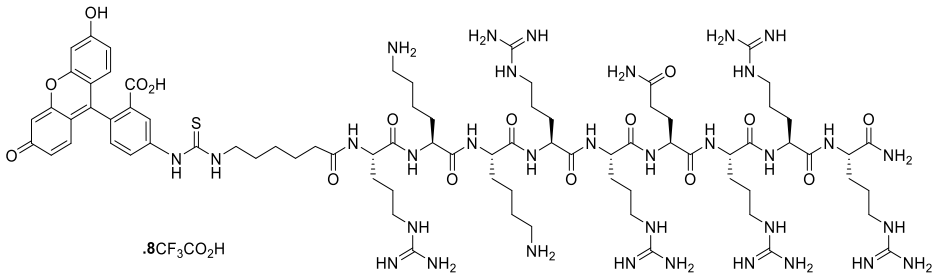


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (19 mg, 7 %).

HRMS: $C_{85}H_{125}N_{19}O_{17}S^{2+}$ Theoretical: 857.9605 Observed: 857.9672.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 28.4 min, 97%.

FITC – XRKKRRQRRR (**Tat**)

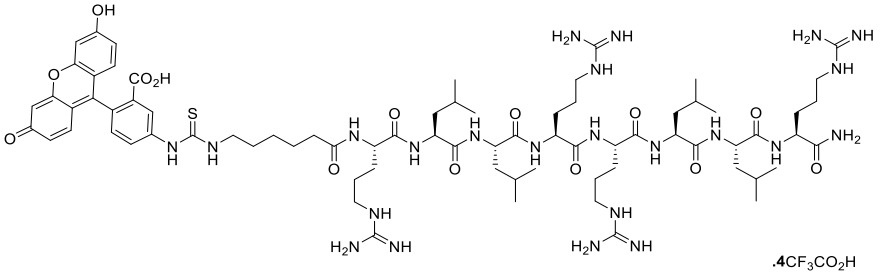


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (22 mg, 6%).

HRMS: $C_{80}H_{133}N_{33}O_{16}S^{4+}$ Theoretical: 658.3235 Observed: 658.3228

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 17.4 min, 100%.

FITC – XRLLRRLLR (**2.83**)

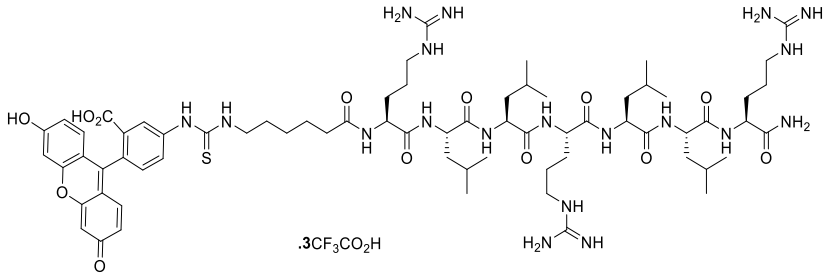


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (70 mg, 23%).

HRMS: $C_{75}H_{119}N_{23}O_{14}S^{2+}$ Theoretical: 798.9508 Observed: 798.9549.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 24.3 min, 98%.

FITC – XRLLRLLR (**2.82**)

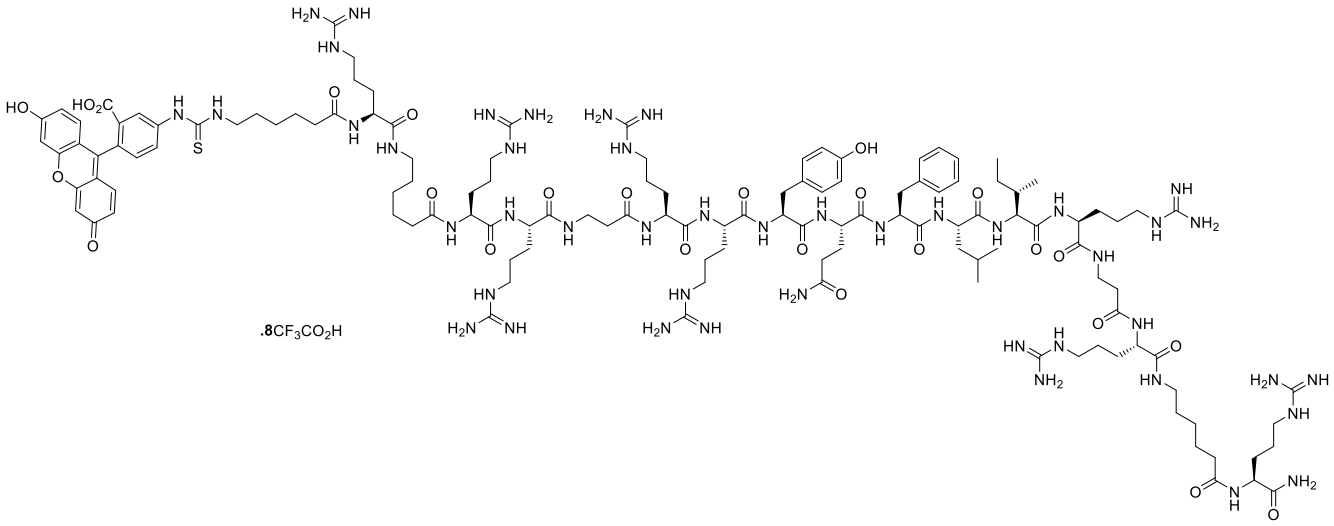


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (68 mg, 25%).

HRMS: $C_{69}H_{107}N_{19}O_{13}S^{2+}$ Theoretical: 720.9003 Observed: 720.9060.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 24.6 min, 99%.

FITC – XRXRRBRRYQFLIRBRXR (**Pip6**)



Using general peptide synthesis protocol A on 0.20 mmol scale of resin. Rink amide resin (307 mg, 0.20 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (20 mg, 3%).

HRMS: $C_{128}H_{206}N_{45}O_{25}S.CF_3CO_2H^{5+}$ Theoretical: 584.1176 Observed: 584.1211.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 19.7 min, 99%.

2.5.6 Cell culture

HeLa and U2OS cells were maintained in medium consisting of Dulbecco's modified Eagle's medium (DMEM), 10% FBS, 1% L-Glutamine and 1% penicillin/streptomycin. Cells were cultured in a humidified incubator at 37 °C with 5% CO_2 atmosphere. Cells were cultured in a humidified incubator at 37 °C with 5% CO_2 and passaged twice a week at confluence > 80%. BSA, FBS, L-Glutamine, penicillin/streptomycin and TrypLee Xpress were purchased sterile from sigma-aldrich or Thermo fischer scientific and used without further treatment.

2.5.7 Flow cytometry

HeLa or U2OS cells were cultured in 24 well plates (4.5×10^5 cells per well) for 24 h. On the day of the experiment, the cells were incubated with 5 μM FITC-peptide in DMEM with 10% FBS at 37 °C or 4 °C for required time. The media was then removed and wells were washed with PBS (2×500 μL per well) and then detached from the plate with TrypLee Xpress (500 μL per well). PBS containing 2% bovine serum albumin (500 μL) was added to each well and cells transferred to 5 mL polystyrene FACS tubes. Cells were centrifuged and washed with PBS containing 2% bovine serum albumin (2×500 μL) before being suspended in PBS containing 2% bovine serum albumin (1.00 mL each). Cells were finally analysed on a Thermo fischer Attune NxT flow cytometer with laser excitation/emission set at 490 nm/519 nm and a flow rate of 200 μL min^{-1}. Flow cytometry data was analysed using FlowJo software with gating set at a range of 51K-285K side scatter and 251K-815K for forward scatter variance for every sample. Median fluorescence intensity values were calculated from the gating range using FlowJo.

For inhibitor experiments cells were incubated with inhibitor (CPZ 10 μM, Nys 50 μM, Wortmannin 0.2 μM) in DMEM with 10% FBS at 37 °C for 1 h before the media was replaced with fresh DMEM with 10% FBS and cells were incubated with inhibitor again (CPZ 10 μM, Nys 50 μM, Wortmannin 0.2 μM) and 5 μM FITC-peptide for 4 h.

2.5.8 Cell viability alamarBlue assay

HeLa or U2OS cells were cultured in 96 well plates (2.6×10^3 cells per well) for 24 h. On the day of the experiment, the cells were incubated with 100 μM, 30 μM, 10 μM, 3 μM, 1 μM, 0.3 μM, 0.1 μM, FITC-peptide in DMEM with 10% FBS (100 uL total well volume) at 37 °C for 24 h. AlamarBlue® was then added to each well (0.5mM, 10 μL) and incubated for a further 6 h for HeLa cells and 24 h for U2OS cells. Cells were then analysed using a Hldex plate reader with laser excitation/emission set at 560 nm/595 nm. Fluorescence of untreated cells were used as a 100% viable control and viabilities were calculated as a percentage of the control fluorescence values using OriginPro 2019b software.

2.5.9 Confocal microscopy

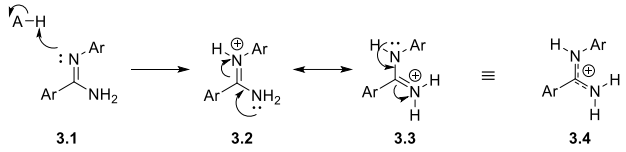
HeLa cells were cultured in 8 well plates ibidi plates (3.5×10^4 cells per well) for 24 h. On the day of the experiment, the cells were incubated with 5 μM FITC-peptide in DMEM with 10% FBS (200 uL total well volume) at 37 °C for 4 h. After 3.5 h incubation, Hoescht 33258 was added at a concentration of 5.6 μM and cells incubated at 37 °C for the final 0.5 h. Cells were removed from the incubator and media removed. Each well was then washed with serum and phenol red free media (2×200 μL) then incubated with 4% formaldehyde in PBS (200 μL) for 20 min at room temperature. After incubation the 4% formaldehyde solution was removed and PBS (200 μL) added to each well. The cells were then imaged using a Leica S8 confocal microscope. Excitation/emission was set at 352 nm/461 nm for Hoescht 33258 imaging and 495 nm/519 nm for FITC labelled peptides. Images were processed using ImageJ software.

Chapter 3: Development of robust Cu-catalysed *N*-arylation of amidines

3.1 Introduction

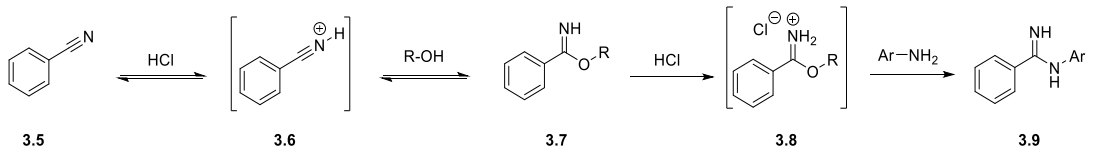
3.1.1 *N*-Arylated amidines

Amidines are important functional groups that are used extensively throughout the pharmaceutical industry and are key basic functionalities in therapeutics including antithrombic, analgesic and antiviral species.[118,188,189] Furthermore, amidines are key precursors in the synthesis of common pharmacophores like quinazolinones and benzimidazoles.[190,191] *N*-substituted amidines are where one of the nitrogen atoms is functionalised to form an imine-like arrangement. The *N*-substituted amidine can often be protonated like a free amidine due to a similar pK_{aH}. Tautomerisation of *N*-substituted amidines can occur between substitution on the imino nitrogen like structure **3.1** to substitution on the amino nitrogen in structure **3.3** (Scheme 3.1).



Scheme 3.1. *N*-Substituted amidine protonation and tautomerization.

Traditional synthetic routes to prepare *N*-substituted amidines have focused on the Pinner reaction which utilises a pre-activated nitrile and an aniline as a nucleophile (Scheme 3.2). However, this approach is limited by either the highly acidic conditions needed for nitrile activation, the extreme temperatures required or very long reaction times.[192,193] The variation and primarily the reactivity of coupling partner functional groups is limited, as many acid and base sensitive protecting groups cannot be employed on the nitrile reagent. The mechanism of the pinner reaction begins with activation of nitrile **3.5** with either a bronsted acid or a lewis acid to form activated nirile **3.6**. Nucleophilc attack of an alcohol onto the activated nitrile **3.6** forms imidate **3.7**. Activation of imidate **3.7** forms **3.8** followed by nucleophilic attack by an aniline will form the corresponding *N*-arylated amidine **3.9**.



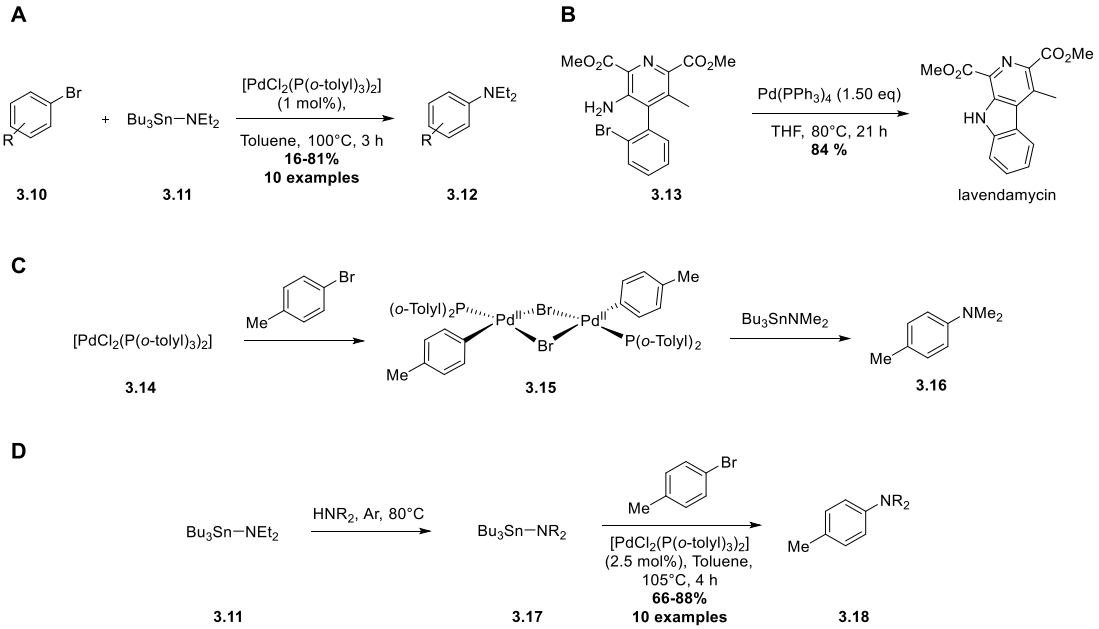
Scheme 3.2. Bronsted acid mediated mechanism of the Pinner reaction to synthesise *N*-substituted amidines.

3.1.2 Transition metal-mediated carbon-nitrogen bond forming reactions

Carbon-nitrogen bond formation is a particularly important transformation due the abundance of the carbon-nitrogen bonds in natural products, therapeutics and bioactive molecules. Direct carbon-nitrogen bond formation has been extensively explored where the use of a transition metal, such as Pd or Cu, will mediate and catalyse the bond formation.[194,195]

3.1.3 Buchwald-Hartwig amination

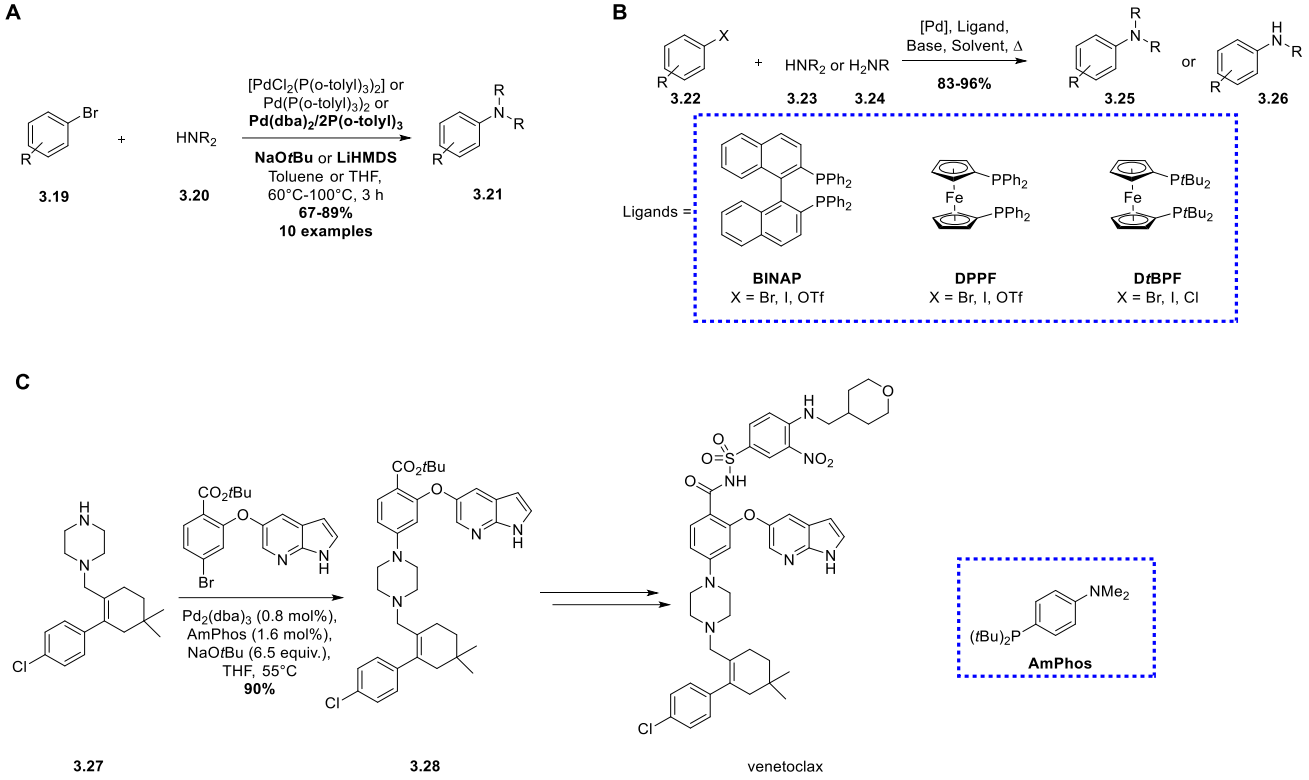
Pd-catalysed carbon-nitrogen bond formation was first reported by Migita *et al*. where aryl bromides were reacted with aminostannane nucleophiles in the presence of a Pd catalyst (Scheme 3.3A).[196] Building from this work, Panek *et al.* utilised Pd for an intramolecular bond formation in the synthesis of lavendamycin (Scheme 3.3B).[197] The lavendamycin synthesis, however, required stoichiometric amounts of Pd as when using catalytic quantities only starting materials were isolated. A report published by Hartwig *et al.* identified key Pd^0 d^{10} and Pd^{II} d^8 systems as the active catalytic species in the reaction (Scheme 3.3C).[198] A further report by Buchwald *et al*. demonstrated that aminostannane nucleophiles can be used with improved yields and scope compared to the original Migita publication (Scheme 3.3D).[199]



Scheme 3.3 (A) Pd catalysed carbon-nitrogen bond formation conducted by Migita *et al*.[196] (B) Pd mediated ring closure synthesis of lavendamycin.[197] (C) Key active Pd species **3.15** isolated by Hartwig *et al*.[198] (D) Synthesis of aryl amines reported by Buchwald *et al.*[199]

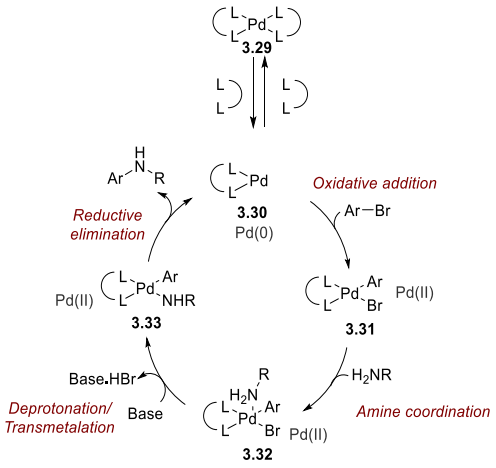
Further development of Pd-mediated carbon-nitrogen bond forming led to publications by both Buchwald and Hartwig *et al.* where the need for toxic aminostannanes was removed. Using alternative catalysts to those used previously (**3.14**) a new strategy using sterically demanding bases, such as NaO*t*Bu or LiHMDS, with free amines completed the amination without the need for stannanes (Scheme 3.4A).[200,201] Further development of phosphine ligands such as BINAP, DPPF and D*t*BPF enhances the scope of the Buchwald-Hartwig reaction to primary amines and can be paired with aryl chloride, iodide and triflate partners (Scheme 3.4B).[202–206] The advancement in ligand design demonstrated the robustness of the Buchwald-Hartwig process and is now widely exploited in various academic and industrial settings. For example, the synthesis of active pharmaceutical ingredient (API)

3.28 is obtained through a Buchwald-Hartwig amination. API **3.28** is a key precursor for the anti-cancer agent venetoclax (Scheme 3.4C).[207]



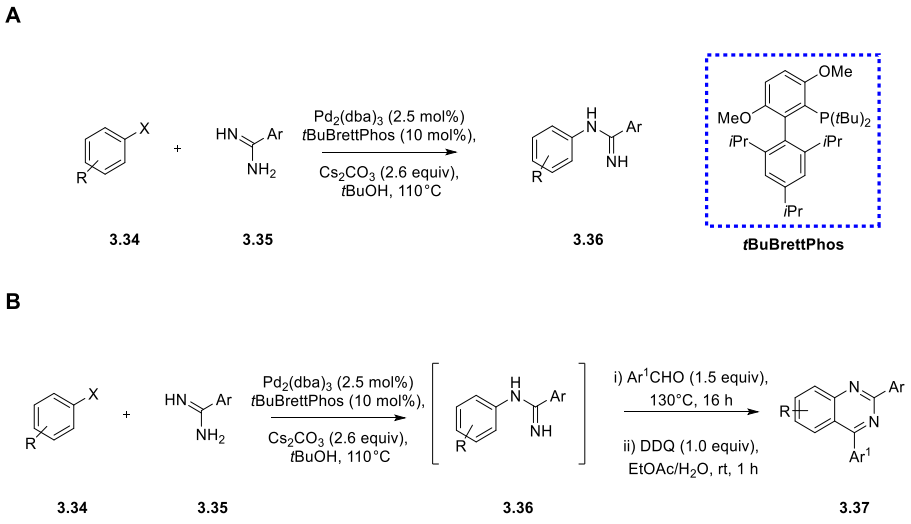
Scheme 3.4. (A) Use of bulky bases allowing access to stannane-free aminations.[200,201] (B) Ligands developed allowing use of primary amines, aryl iodides and aryl chlorides.[202–206] (C) Synthesis of venetoclax *via* a Buchwald-Hartwig amination.[207]

With the advent of bis-phosphine ligands as alternatives to catalysts like **3.14**, several reaction intermediates were isolated to elucidate a viable catalytic cycle (Scheme 3.5).[208] The cycle begins with the ligand dissociation to form the active Pd^0 species **3.30**. Oxidative addition then takes place with the aryl halide and to form the Pd^{II} intermediate **3.31**. The amine species then coordinates to the Pd^{II} centre forming the five-coordinate Pd complex **3.32** followed by deprotonation/transmetalation to form Pd^{II} species **3.33**. The cycle is then completed by reductive elimination to release the aryl amine product and reform the Pd^0 species **3.30**.



Scheme 3.5. General catalytic cycle of the Buchwald-Hartwig reaction using bidentate phosphine ligands.[208]

The mechanism and robustness of the Buchwald-Hartwig amination has become well established. However, there is only one report, by Buchwald *et al.*, for the direct arylation of amidines using Pd catalysed methods.[209] In this synthetic route of *N*-arylated amidines were obtained using $Pd_2(dba)_3$ and the monodentate phosphine ligand *t*BuBrettPhos. The *N*-aryl amidine **3.36** was then subjected to a condensation reaction to form a quinazolinone **3.37**. The Buchwald-Hartwig amidination and condensation were then combined in a one-pot process for the synthesis of a variety of quinazolinones (Scheme 3.6).

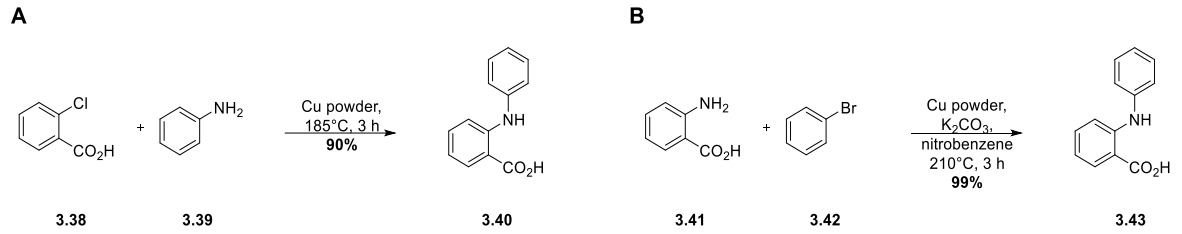


Scheme 3.6. (A) Conditions for Buchwald-Hartwig amidination. (B) One-pot Buchwald-Hartwig amidination and condensation for the synthesis of quinazolinones.[209]

This remains the only example of a Buchwald-Hartwig amidination and it has certain drawbacks, most notably the need for significantly elevated temperatures. Furthermore, as with most Buchwald-Hartwig reactions there requires a maintenance of a strict inert atmosphere and addition of, often expensive, ligands.

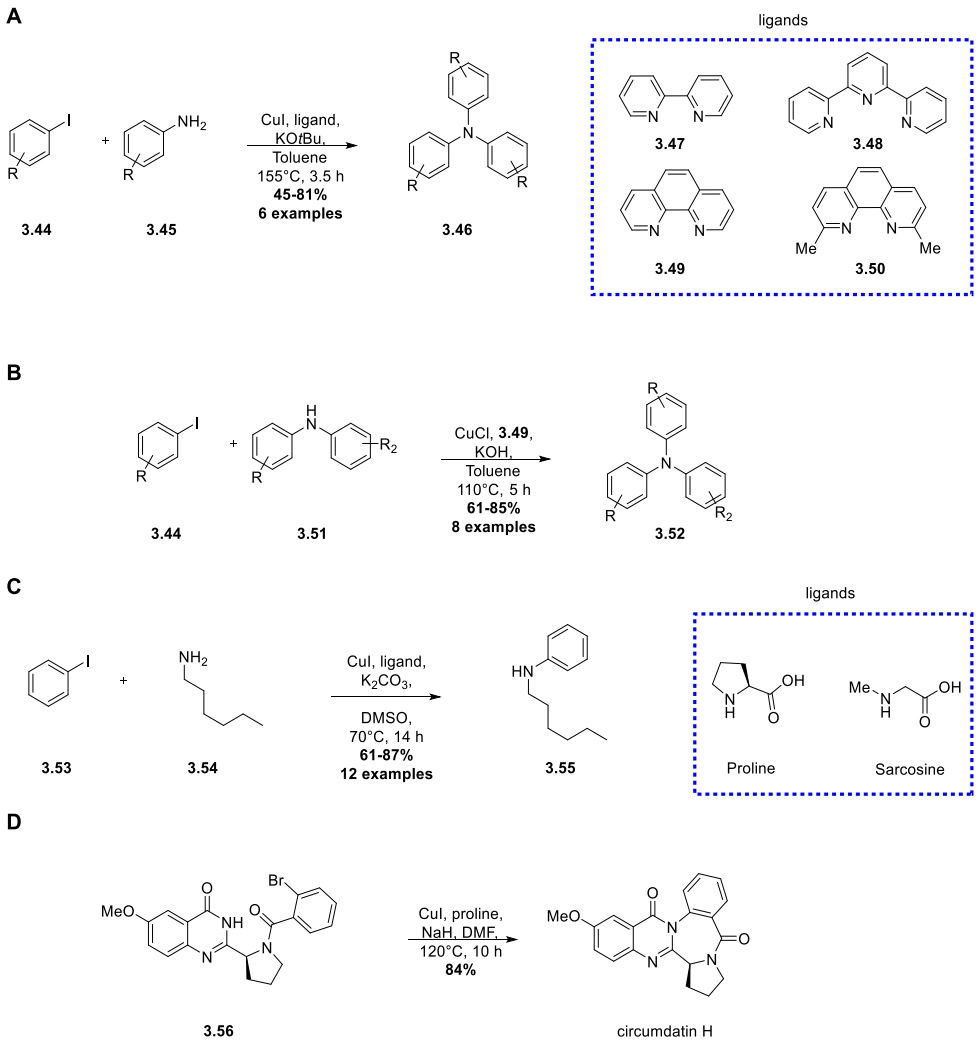
3.1.4 Cu-catalysed Ullmann couplings

Cu-catalysed processes for carbon-nitrogen bond formation have become an attractive alternative to Pd-catalysed processes due to the abundance and low cost of Cu. The first examples of this came in the early 20th century with the pioneering works by Ullmann and Goldberg.[210,211] Using Cu powder, secondary anilines **3.40** and **3.43** are formed from a reaction between primary anilines **3.39** or **3.41** and aryl bromide/chloride **3.42** or **3.38** (Scheme 3.7A & B).



Scheme 3.7. (A) Cu catalysed amination by Ullmann. (B) Cu catalysed amination by Goldberg.[210,211]

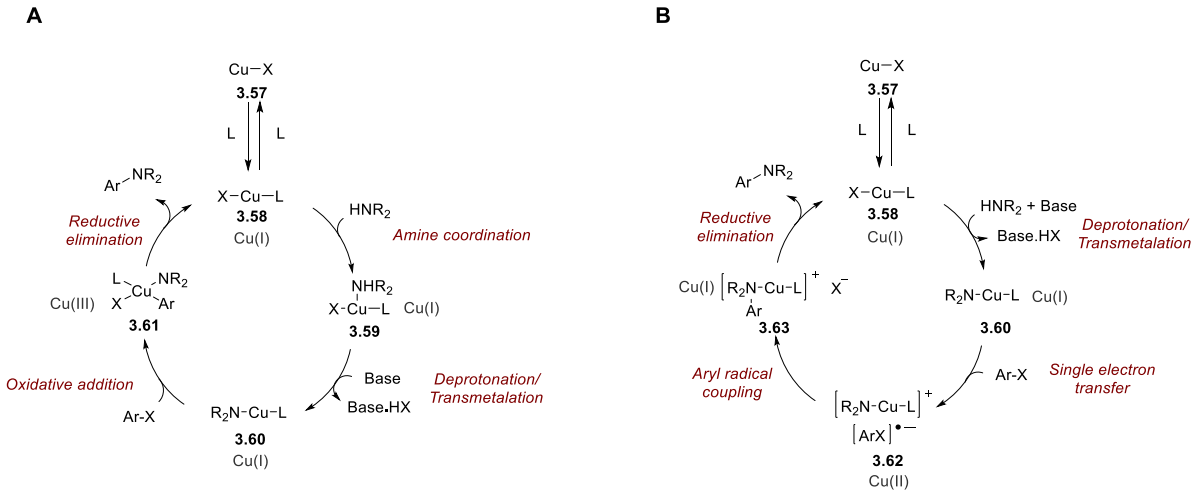
Although Ullmann aminations are a step-efficient route to form carbon-nitrogen products, excessive temperatures are a limiting factor for the wider scope. The development of different Cu catalysts and ligands does allow for milder reaction conditions and broader substrate scope. The use of CuI with bidenate, nitrogen bearing ligands **3.47**-**3.50** is one strategy used to prepare arylated anilines in the presence of KO*t*Bu (Scheme 3.8A).[212] Ligands such as 1,10-phenanthroline **3.49** have been utilised for the successful monoarylation and diarylation of anilines with aryl iodides in the presence of KOH and CuCl (Scheme 3.8B).[213] Amino acids like proline and sarcosine have also been used as ligands in the Ullmann reaction, enhancing the scope to both alkyl and aryl amines (Scheme 3.8C).[214] Due to these advances, Ullmann reactions have been utilised in the synthesis of the natural product circumdatin H.[215] An amide, using proline ligands, was coupled with aryl bromide to produce this important mitochondrially targeted alkaloid (Scheme 3.8D).



Scheme 3.8. (A) CuI catalysed amination utilising nitrogen bearing ligands.[212] (B) CuCl mediated amination using Phenanthroline **3.49**.[213] (C) Amino acid ligand mediated Ullmann reaction.[214] (D) Synthesis of circumdatin H using an Ullmann reaction.[215]

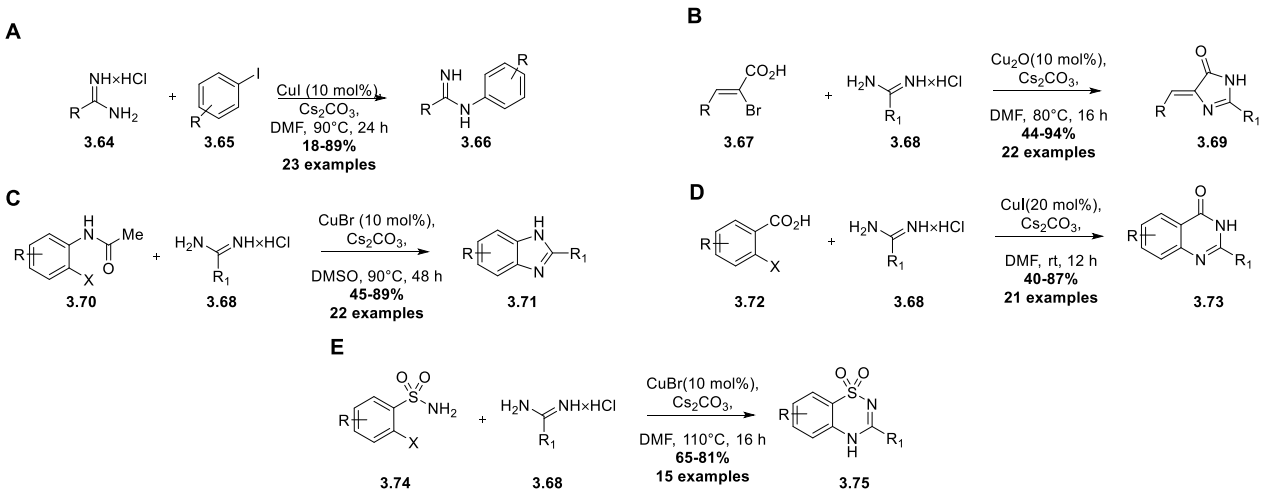
The precise mechanism of the Ullmann reaction remains less well defined than the analogous Buchwald-Hartwig amination. The general mechanism would appear similar to Pd-catalysed processes with amine coordination to the metal centre, activation of the aryl-halogen bond and subsequent carbon-nitrogen bond formation (Scheme 3.9A). This has been supported by work conducted by Buchwald *et al.* where a general trend of aryl halide reactivity (ArI > ArBr > ArCl >> ArF) suggested activation of the aryl-halogen bond and selectivity for *N*-arylation over *O*-arylation in aminoalcohols supported amine coordination over other nucleophiles.[216,217] The proposed oxidative/elimination mechanism begins with ligand association to a Cu^{I} centre to form species **3.58**. Subsequent amine coordination and transmetalation leads to the amine bearing Cu^{I} species **3.60**. Here oxidative addition takes place with an aryl halide to form the Cu^{III} ccomplex **3.61** before reductive elimination releases the *N*-arylated product and subsequently regenerates a catalytically active Cu^{I} species.

Alongside this work, EPR and computational experiments have been used to identify the presence of Cu^{II} centres in some Ullmann reactions.[218,219] This led to an alternative mechanism being proposed involving radical formation from a single electron transfer (Scheme 3.9B). The mechanism is similar to the oxidative addition/reductive elimination process initially proposed, with amine coordination and transmetalation affording the common Cu^{I} species **3.60**. A single electron transfer then takes places from **3.60** to the aryl halide to form a radical anion and Cu^{II} complex **3.62**. Coupling of the aryl radical and Cu produces cationic Cu^{I} species **3.63** before reductive elimination yields the *N*-arylated product and regenerates Cu^{I} centre **3.58**. Although the mechanism of Ullmann aminations remains under debate the oxidative addition/reductive elimination mechanism is the most widely accepted with recent work by Davies *et al.* highlighting the importance of ligands on the oxidative reaction kinetics.[220]



Scheme 3.9. Proposed mechanisms for the Ullmann cross coupling. (A) Oxidative addition/reductive elimination mechanism.[216,217] (B) Single electron transfer mechanism.[218]

The Ullmann reaction has also been employed on amidine substrates, both directly and in many domino processes to access other structures. Antilla *et al.* developed an efficient mono *N*-arylation of amidines with a broad scope of various aryl iodides and amidines (Scheme 3.10A).[221] Vinyl bromides and amidines have been reacted using Ullmann chemistry in a one-pot process for the synthesis of imidazolones, which are core structures for many fluorescent proteins (Scheme 3.10B).[222] Another report utilising the amidine-Ullmann process reported access to benzimidazoles by coupling amidines with *ortho*-haloacetanilides (Scheme 3.10C).[223] As mentioned in *section 3.1.1*, quinazolinones are desirable pharmacophores and these have been accessed through Ullmann chemistry with amidines as reported by Fu *et al*. (Scheme 3.10D).[190] Further work by Fu *et al.* demonstrated the application of amidine-Ullmann chemistry in the synthesis of benzothiadiazine pharmacophores (Scheme 3.10E).[224]

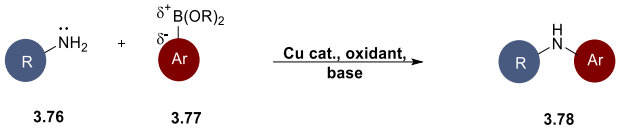


Scheme 3.10. Examples of Ullmann chemistry using amidine substrates. (A) Mono *N*-arylation of amidines.[221] (B) Synthesis of imidazolones.[222] (C) Benzimidazole synthesis.[223] (D) Quinazolinones.[190] (E) Benzothiadiazines.[224]

Arylation of amidines has been explored in a variety of scenarios in Ullmann chemistry.[221–224] Despite this, the Ullmann amination has similar drawbacks to the Buchwald-Hartwig reaction. They include elevated reaction temperatures, the need for an inert atmosphere as well as the use of high-boiling solvents like DMSO or DMF. These conditions are relatively harsh and will limit the scope and scalability of the reaction.

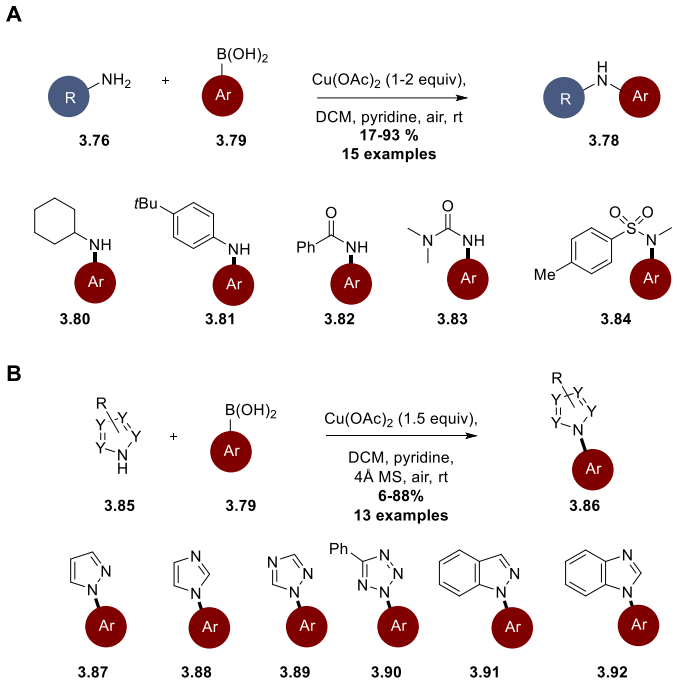
3.1.5 Chan-Evans-Lam amination

The key distinction of the Chan-Evans-Lam reaction relative to the Buchwald-Hartwig and Ullmann reactions is it involves nucleophile-nucleophile coupling (Scheme 3.11). The reaction is Cu catalysed and couples an amine with an organoboron species and is facilitated by an external oxidant which is not required in electroneutral processes such as Ullmann or Buchwald-Hartwig aminations.



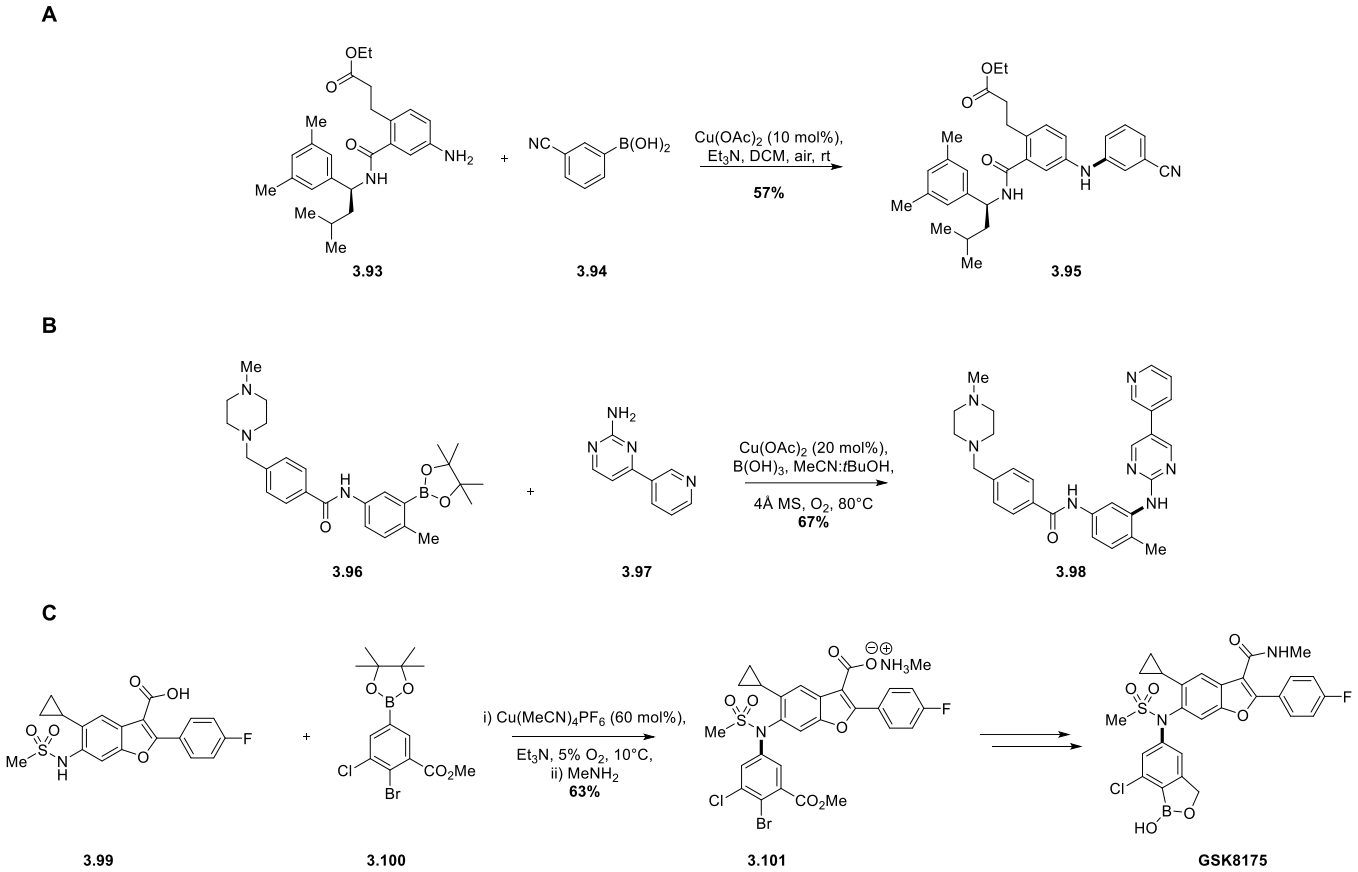
Scheme 3.11. General scheme for the Chan-Evan-Lam amination

The first report of the Chan-Evans-Lam reaction was in 1998 and reported coupling of a variety of *N*-nucleophiles with aryl boronic acids. The reactions were conducted under air at room temperature using $Cu(OAc)_2$, pyridine and DCM. These were comparatively milder than established Buchwald and Ullmann transformations. The reaction also demonstrated the potential breadth of the possible substrate scope with a great number of *N*-nucleophiles giving favourable yields, including alkyl amines, anilines, amides and urea coumpounds (Scheme 3.12A).[225] The reaction scope was further enhanced with work by Lam *et al.* demonstrating the reaction can successfully couple a selection of *N*-nucleophilic heterocycles again at room temperature under air (Scheme 3.12B).[226]



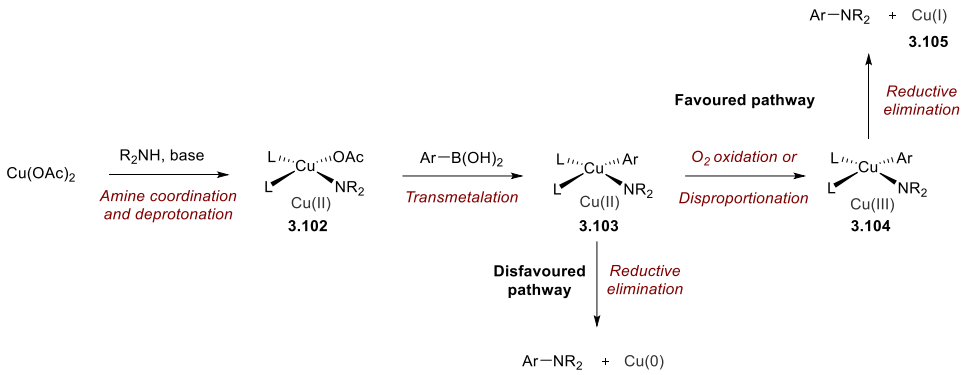
Scheme 3.12. (A) Chan-Lam amination scope published by Chan et al. in 1998.[225] (B) Heterocycles subjected to the Chan-Lam amination.[226]

The Chan-Lam reaction was further adapted in the late 2000s to allow the use of sub-stoichiometric amounts of Cu, making it more comparable to other metal cross-coupling reactions.[227] Utilising O_2 from air as the reaction oxidant has directly allowed the formation of carbon-nitrogen coupling products where a strong oxidant would be incompatible. For example, catalytic Chan-Lam chemistry has been utilised in medicinal chemistry for the synthesis of EP3 receptor antagonist **3.95** (Scheme 3.13A).[228] Work by Watson *et al.* has established the Chan-Lam reaction for the synthesis of the tyrosine kinase inhibitor imatinib **3.98** under catalytic conditions (Scheme 3.13B).[229] Further work by Watson *et al.* demonstrated the scalability of the Chan-Lam amination by synthesising compound **3.101**, which is an intermediate in the synthesis of the API NS5B en-route to the preparation of polymerase inhibitor GSK8175 (Scheme 3.13C).[230] The Chan-Lam reaction to synthesise **3.101** was conducted on a 15 g scale under mild catalytic conditions.



Scheme 3.13. (A) Catalytic Chan-Lam reaction for the synthesis of EP3 inhibitor **3.95**.[228] (B) Chan-Lam synthesis of inhibitor **3.98**.[229] (C) 15 g scale Chan-Lam synthesis of **3.101** towards the API GSK8175.[230]

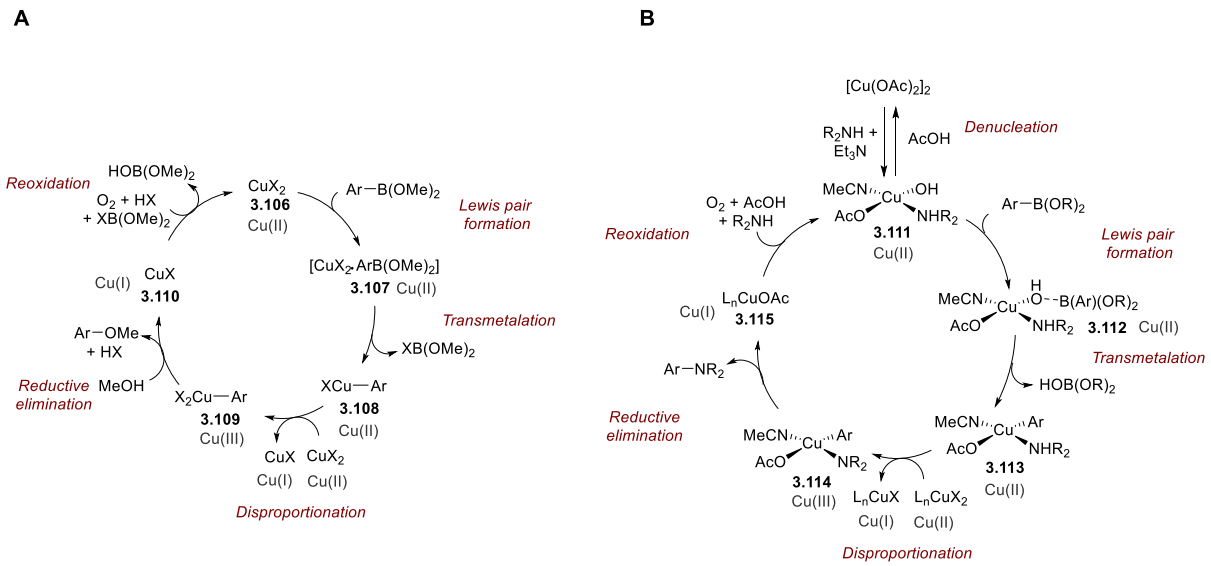
Lam and co-workers were the first to propose a mechanism for carbon-nitrogen bond formation using the Chan-Lam reaction (Scheme 3.14).[231,232] This built upon the carbon-oxygen bond formation processes investigated by Evans *et al*.[233] The mechanism begins with coordination of an amine nucleophile to $Cu(OAc)_2$ followed by deprotonation to afford the Cu^{II} species **3.102**. Subsequent transmetalation affords the aryl-bearing Cu^{II} complex **3.103**. Complex **3.103** either undergoes reductive elimination to release the aryl amine product and Cu^0 or be oxidised to the Cu^{III} complex **3.104**. Finally, reductive elimination of complex **3.104** affords aryl amine product and Cu^I species **3.105**. The favoured pathway was suggested to be *via* oxidation of **3.103** as only trace amount of Cu^0 was detected in the reaction mixtures, hence disfavouring reductive elimination of **3.103**. The formation of **3.104** was proposed to occur either by O_2 oxidation or by disproportionation.



Scheme 3.11. Chan-Lam amination mechanism proposed by Lam and co-workers.[231,232]

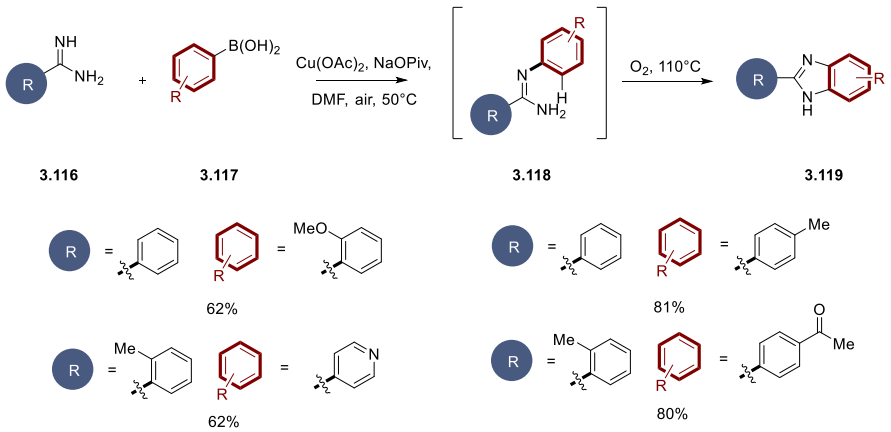
The mechanism proposed by Lam *et al.* did not follow a catalytic cycle and therefore the catalytic mechanism was further investigated by Stahl *et al*. Using the Chan-Lam etherication process as a model through a combination of EPR experiments and O_2 uptake kinetic studies a catalytic mechanism was proposed.[227] First, the reaction proceeds without O_2, supporting the idea of a disproportionation pathway to oxidise Cu^{II} to Cu^{III}. However, O_2 uptake studies highlighted that Cu^{I} solutions are readily oxidised to Cu^{II} in the presence of O_2. Finally, EPR analysis highlighted that the majority of the Cu present in the reaction was Cu^{II}-lacking strong electron donor ligands, suggesting the rate-limiting step of the reaction was transmetalation.[227] Based on these observations a catalytic mechanism was proposed (Scheme 3.15A).[227,234] The mechanism starts with the Cu^{II} complex **3.106** forming a Lewis pair with the organoboron species resulting in the formation of complex **3.107**. Then, transmetalation of **3.107** occurs to deliver Cu^{II} complex **3.108** and subsequent disproportionation yields Cu^{III} species **3.109**. Reductive elimination then takes place to form the Cu^{I} complex **3.110** before reoxidation in the presence of O_2 takes place to reform the initial Cu^{II} centre.

Building on this work by Stahl, Watson *et al.* fully established the catalytic cycle for the Chan-Lam amination process through a combination of EPR, crystallography and mass spectrometry techniques (Scheme 3.15B).[229] Using both exemplar aryl and alkyl amines with an aryl boronic acid, key Cu intermediates were observed with mass spectrometry and off-cycle species were isolated and identified by X-ray crystallography. EPR analysis also confirmed that the $Cu(OAc)_2$ dimer undergoes denucleation to a Cu^{II} monomer upon treatment with either alkyl or aryl amines. The mechanism begins with the denucleation of $[Cu(OAc)_2]_2$ to form mononuclear Cu^{II} complex **3.111** before a formation of a Lewis-pair with the organoboron species. Lewis-pair **3.112** subsequently undergoes transmetalation to yield the aryl Cu^{II} centre in **3.113**. Disproportionation occurs to produce the Cu^{III} complex **3.114** before reductive elimination affords aryl amine product and Cu^{I} complex **3.115**. Finally, oxidation of **3.115** occurs utilising O_2 to complete the catalytic cycle.



Scheme 3.15 (A) Stahl's etherication Chan-Lam mechanism.[235] (B) Watson's Chan-Lam amination mechanism.[229]

With the mechanism of the Chan-Lam reaction fully established by Stahl and Watson the application of the Chan-Lam reaction has greatly expanded in recent years.[236] Neuville *et al*. applied the reaction to amidines in a sequential process to synthesise benzimidazoles (Scheme 3.16).[237] Unlike many Chan-Lam reactions, the conditions used required elevated temperatures and the use of DMF as the solvent in the Chan-lam step of the reaction. The scope within the paper was limited to only four examples of *N*-aryl amidines and poorly explored the changes in the reaction mechanism when using amidines. There remains no route or guidance on controlling Chan-Lam reactivity for mono *N*-arylation of amidine substrates.



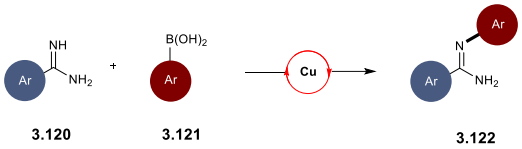
Scheme 3.16. Sequential Chan-Lam and C-H activation synthesis of benzimidazoles from amidines with the four mono-aryl amidines isolated.[237]

3.2 Aims of Chapter 3

Current transition metal-mediated methodology for the synthesis of mono *N*-arylated amidines have several drawbacks. In both Buchwald-Hartwig and Ullmann reactions elevated temperatures; addition of often expensive ligands; use of high boiling solvents and the need for an inert atmosphere are required to obtain reasonable product conversion. The Chan-Lam amination is an attractive alternative to Buchwald-Hartwig, Ullmann or Pinner reactions for synthesis of mono *N*-arylated amidines due to its inherent milder reaction conditions and accessibility of a range of boronic acids. The Chan-Lam reaction can generally be performed at lower temperatures under air without additional ligands in easy-to-handle solvents. However, the use Chan-Lam conditions with amidines is relatively unexplored where currently only the work by Neuville *et al*. have investigated Chan-Lam amidination.[237]

The specific aims of this chapter are to:

(i) Optimise an efficient Chan-Lam reaction for the synthesis of mono *N*-arylated amidines (Scheme 3.17).



Scheme 3.17. Proposed Chan-Lam amidination

(ii) Investigate any deviation in the mechanism of *N*-arylation of amidines relative to the existing Chan-Lam mechanism.

(iii) Establish the scope and limitataions of Chan-Lam amidination with a range of boronic acids.

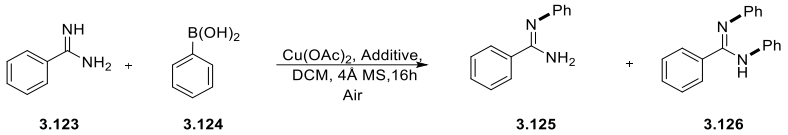
(iv) Apply Chan-Lam amidination conditions to a biologically relevant amidine and investigate any differences in biological efficacy between the mono *N*-aryl amidine and the parent free amidine.

3.3 Results and discussion

3.3.1 Optimisation of a Chan-Lam amidination

The reaction was optimised using the model substrates of benzamidine **3.123** and phenyl boronic acid **3.124** which were used to assess the production of *N*-phenylbenzimidamide **3.125** and *N*,*N*-diphenylbenzimidamide **3.126**. As the amidine has two nitrogen atoms for arylation it was deemed necessary to assess the quantity of bi-arylation observed during the initial reaction optimisation (Table 3.1).

Table 3.1. Initial amidine Chan-Lam conditions tested.



3.123 **3.124** **3.125** **3.126**

Entry	**3.123:3.124**	Additive	$Cu(OAc)_2$ (mol%)	Rxn temp	**3.125** (%)[(a)]	**3.126** (%)[(a)]
1	1.0:2.0	Et_3N	100	rt	8	74
2	1.0:2.0	$B(OH)_3$	100	rt	26	5
3	2.0:1.0	$B(OH)_3$	100	rt	40	13
4	2.0:1.0	Et_3N	100	rt	59	10

(a) Isolated yields.

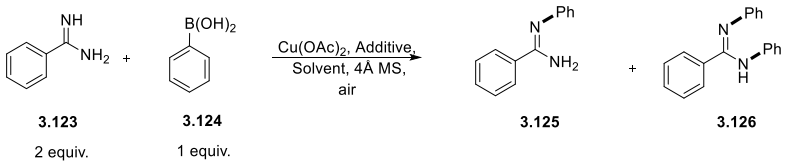
As a starting point to design effective conditions, initial conditions reported by Chan *et al*. were assessed (Table 3.1, Entry 1).[225] Application of Chan's conditions, using two equivalents of **3.124** and one equivalent of **3.123**, formed the monoarylated species **3.125** in 8% and the bi-arylated species **3.126** as the major product in 74%. In an effort to reduce the over reactivity observed, the additive was changed from Et_3N to $B(OH)_3$ (Table 3.1, Entry 2). In the Chan-Lam mechanism established by Watson *et al.* $B(OH)_3$ was used to trap unwanted acetate biproducts in the reaction that can cause catalyst inhibition.[229] Using $B(OH)_3$ and an excess of **3.124** (Table 3.1, Entry 2) offered a significant reduction in bi-arylation to only 5% and a bias to mono-arylation with a 26% yield recovered. The remaining mass balance of the reaction could possibly be attributed to degradation of boronic acid by protodeborination and oxidative deborination.[229,238] Further detail in the Watson mechanism (Fig. 3.4B) details deborination events are mediated by Cu^I species. To overcome Cu^I deborination the addition of excess nucleophile promotes the oxidation of Cu^I to Cu^{II} and therefore reduces deborination.[229,236]

Applying an excess of amidine (Table 3.1, Entry 3) produced **3.125** in a 40% isolated yield, with bi-arylated **3.126** isolated in a 13% yield. A final change to Et_3N in combination with an excess of amidine (Table 3.1, Entry 4) returned a good yield of 59% for **3.125**. The reactions highlighted in Table 3.1

suggest that the excess amidine can reduce both over reaction to the biarylated **3.126** and boronic acid degradation by deborination events.

Working from the improved method for mono-arylation further optimisation was carried out maintaining the 2:1 stoichiometry of amidine to boronic acid (Table 3.2). The optimisation would now attempt to address the effects of changing solvent, additive, reaction time and temperature.

Table 3.2. Reaction optimisation for the synthesis of amidine **3.125**.



3.123 2 equiv. **3.124** 1 equiv. **3.125** **3.126**

Entry	Solvent	Rxn time	Additive	$Cu(OAc)_2$ (mol%)	Rxn temp (°C)	**3.125** (%)[a]	**3.126** (%)[a]
1	CH_2Cl_2	16h	$B(OH)_3$	100	rt	45	11
2	CH_2Cl_2	16h	Et_3N	100	rt	58	14
3	CH_2Cl_2	16h	K_2CO_3	100	rt	40	8
4	MeOH	16h	K_2CO_3	100	rt	59	2
5	MeOH	16h	K_2CO_3	50	rt	58	2
6	MeOH	16h	K_2CO_3	100	80	36	<1
7	***i*-PrOH**	**2h**	**K_2CO_3**	**100**	**rt**	**76 (83)[b]**	**5**
8	*i*-PrOH	16h	Et_3N	100	rt	70	6
9	*i*-PrOH	16h	$B(OH)_3$	100	rt	38	14
10	*i*-PrOH	16h	K_2CO_3	50	rt	7	0
11	*i*-PrOH	2h	K_2CO_3	50	rt	2	0

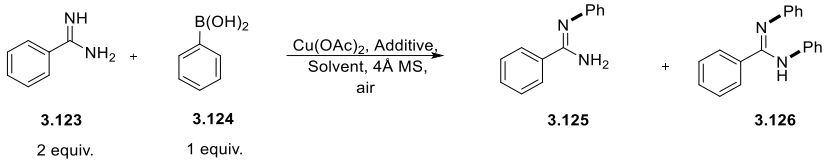
(a) Reaction completed on a 0.20 mmol scale and conversion calculated by HPLC based upon standard concentration curves. (b) Isolated yield on 1.00 mmol scale.

Repeating reaction conditions from Table **3.1** with the use of Et_3N and $B(OH)_3$ in combination with CH_2Cl_2 resulted in conversions of 45% and 58% to **3.125** (Table 3.2, Entries 1 and 2). However, biarylated product was still present in over 10% with both Et_3N and $B(OH)_3$ entries. Inorganic bases are less common in Chan-Lam reactions but selecting K_2CO_3, which has a similar pK_{aH} to Et_3N (pK_{aH} ~ 10) brought a small reduction in biaryltion to 8% along with a reduction in **3.125** to 40%, though it is noted that K_2CO_3 is poorly soluble in CH_2Cl_2 (Table 3.2, Entry 3). Interestingly, changing the solvent to MeOH and still using K_2CO_3 reduced the production of **3.126** to 2% (Table 3.2, Entry 4). Using this

combination of solvent and base allowed a substoichiometric level of $Cu(OAc)_2$ to be used with 50 mol% $Cu(OAc)_2$ yielding similar monoarylation conversion of 58% and maintained a low level of biarylation (Table 3.2, Entry 5). In an attempt to increase conversion to **3.125** the reaction temperature was increased to 80 °C, however this gave a significant reduction in monoarylated product (Table 3.2, Entry 6).

Another change of solvent to *i*-PrOH allowed for **3.125** to be isolated in a yield of 83% after only 2 h when using K_2CO_3 as the base (Table 3.2, Entry 7). A change in additive to Et_3N with *i*-PrOH produced a conversion of 70% for **3.125** and using $B(OH)_3$ returned only a conversion of 38% mono-arylation (Table 3.2, Entries 8 and 9). Using 50 mol% $Cu(OAc)_2$ in *i*-PrOH resulted in 10-fold decreases in conversion to **3.125** after both 16 h and 2 h (Table 3.2, Entries 10 and 11). It was noted that a significant amount of purple precipitate formed within 5 minutes under these conditions, compared to the typical blue/green suspensions generally observed. The reaction could be performed using substoichiometric levels of $Cu(OAc)_2$ in MeOH with K_2CO_3 and return a reasonable conversion of 58% (Table 3.2, Entry 5). Though using *i*-PrOH with K_2CO_3 only returned 7% **3.125** over the same time period. Using 100 mol% of $Cu(OAc)_2$ with K_2CO_3 in *i*-PrOH provided the best yield of **3.125**. A further optimisation was undertaken to identify conditions that would produce a similar yield of **3.125** but using substoichiometric amounts of $Cu(OAc)_2$ (Table 3.3).

Table 3.3. Optimisation of a substoichiometric $Cu(OAc)_2$ Chan-Lam amidination.



3.123 2 equiv. **3.124** 1 equiv. **3.125** **3.126**

Entry	Solvent	Rxn time	Additive	$Cu(OAc)_2$ (mol%)	Rxn temp (°C)	**3.125** (%)[a]	**3.126** (%)[a]
1	*i*-PrOH	2h	K_2CO_3	50	rt	2	--
2	DMF	24h	NaOPiv	20	50	81[c]	2
3	DMF	2h	NaOPiv	20	50	54[c]	--
4	*i*-PrOH	2h	K_2CO_3	50	50	56	--
5	*i*-PrOH	2h	K_2CO_3	20	50	62	--
6	*i*-PrOH	2h	NaOPiv	50	50	68	2
7	*i*-PrOH	2h	NaOPiv	20	50	47	--
8	*i*-PrOH	2h	NaOPiv	20	rt	14	--
9	***i*-PrOH**	**24h**	**K_2CO_3**	**20**	**50**	**76 (81)[b]**	**--**
10	*i*-PrOH	24h	K_2CO_3	10	50	64	--

(a) Conversion calculated by HPLC based upon standard concentration curves. (b) Isolated yield on 1.00 mmol scale. (c) Reaction performed using 1.0 equiv **3.123** and 1.2 equiv **3.124**. (--) Yield too low to be determined.

Using the conditions from the first step of the benzimidazole synthesis by Neuville *et al.*, where 20 mol% $Cu(OAc)_2$ was used, afforded 81% of the mono-arylated species **3.125** (Table 3.3, Entry 2).[237] These conditions gave comparable conversions to the optimised *i*-PrOH conditions from Table 3.2 (Entry 7). The main parameters changed in the Neuville conditions are the reaction solvent (DMF), additive (NaOPiv), temperature (50 °C) and time (24 h). To investigate the increased reaction time in the Neuville conditions the reaction was repeated at 2 h and this resulted in a lower **3.125** conversion of 54% (Table 3.3, Entry 3), suggesting a lower reaction rate than the optimised *i*-PrOH reaction. To examine the influence of the reaction temperature, the reaction was performed at 50 °C, using *i*-PrOH, K_2CO_3 and either 50 mol% or 20 mol% $Cu(OAc)_2$ (Table 3.3, Entries 4 & 5). The increase in temperature returned conversion of 56% mono-arylation using 50 mol% $Cu(OAc)_2$ and 62% using 20 mol% $Cu(OAc)_2$.

The time and temperature experiments identified firstly, that the Neuville conditions require 24 h for reaction completion as reducing the reaction time to 2 h reduced the conversion to **3.125** by 30%.

Secondly, the experiments identified that a combination of *i*-PrOH and K_2CO_3 produced good conversions of **3.125** using substoichiometric amounts of $Cu(OAc)_2$, but the reaction requires heating to 50 °C, suggesting elevated temperature overcomes the formation of an inactive species formed at room temperature. To scrutinise the reaction additive, NaOPiv was used in combination with *i*-PrOH. Using 50 mol% $Cu(OAc)_2$ with *i*-PrOH and NaOPiv formed **3.125** in 68% conversion when the reaction was heated to 50 °C (Table 3.3, Entry 6). A further reduction to 20 mol% returned a conversion to the monoarylated product of 47% (Table 3.3, Entry 7). Using 20 mol% $Cu(OAc)_2$ with *i*-PrOH and NaOPiv returned a low conversion of 14% for **3.125** when the reaction was performed at room temperature (Table 3.3, Entry 8). The NaOPiv experiments suggest that the additive is a less important factor than reaction temperature for improved conversion to **3.125** as reasonable conversions were only observed when heating the reaction (Table 3.3, Entries 4-7). Finally, **3.125** could be isolated in a yield of 81% when using *i*-PrOH with K_2CO_3, 20 mol% $Cu(OAc)_2$ at 50 °C for 24 h (Table 3.3, Entry 9). Finally performing the reaction with 10 mol% $Cu(OAc)_2$ saw a reduction to 64% conversion, indicating 20 mol% to be the optimum $Cu(OAc)_2$ loading (Table 3.3, Entry 10). The final two experiments demonstrate the importance of temperature and time in order to perform the amidination catalytically. From the optimisation tables two sets of mild conditions have been identified for Chan-Lam amidination (Table 3.2, Entry 7 & Table 3.3, Entry 9).

3.3.2 Mechanistic analysis of the Chan-Lam amidination

Though the mechanism of the Chan-Lam amination was established by Watson *et al*. the change from an amine to an amidine in the reaction would likely significantly effect and even alter the reaction pathway.[229] The greater basicity of the amidine would affect its binding to the Cu centre and the inorganic bases K_2CO_3 and NaOPiv could also act as potential ligands. As discussed in *Section 3.1.5* $Cu(OAc)_2$ exists as a dimer and must undergo denucleation for the Chan-Lam reaction to proceed. The $Cu(OAc)_2$ dimer exists in a 'paddlewheel'-like structure with two square pyramidal Cu centres and bridging acetate ligands (Fig. 3.1, **3.127**).[239] A dimeric Cu-piperidine species and tetrameric Cu-piperidine species were isolated and identified as key off-cycle Cu species in the Chan-Lam amination using piperidine (Fig 3.1, **3.128** and **3.129**).[229,240] The species can then denucleate to afford active mononuclear species and hence demonstrate catalytic competence.

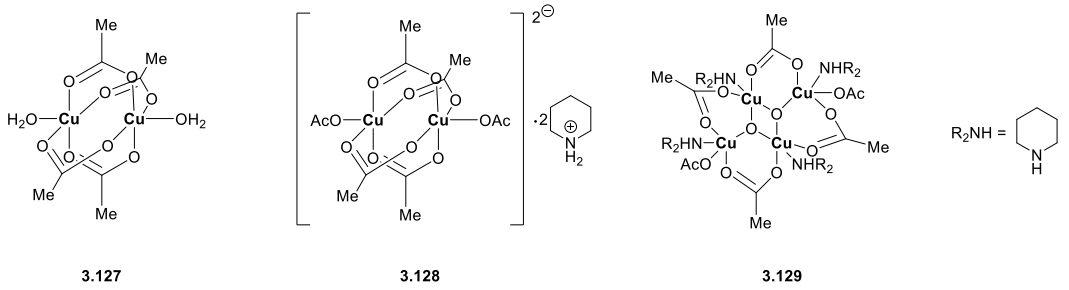


Figure 3.1. $Cu(OAc)_2$ paddlewheel structure **3.127**, Cu-piperidine dimer **3.128**, Cu-piperidine tetramer **3.129**.

As an amidine is analogous to a carboxylate group, where the oxygen atoms are replaced with nitrogen atoms it was envisaged that possible amidine analogues of complexes **3.127** and **3.128** may be reaction species (Fig. 3.2, **3.130** and **3.131**). In the amination process the amine de-nucleates the $Cu(OAc)_2$ and it was postulated that square pyramidal mononuclear amidine Cu complexes may then be formed after denucleation (Fig 3.2, **3.132**). It was postulated that the anions of K_2CO_3 and NaOPiv could form complexes with Cu (Fig. 3.2, **3.133** and **3.134**).

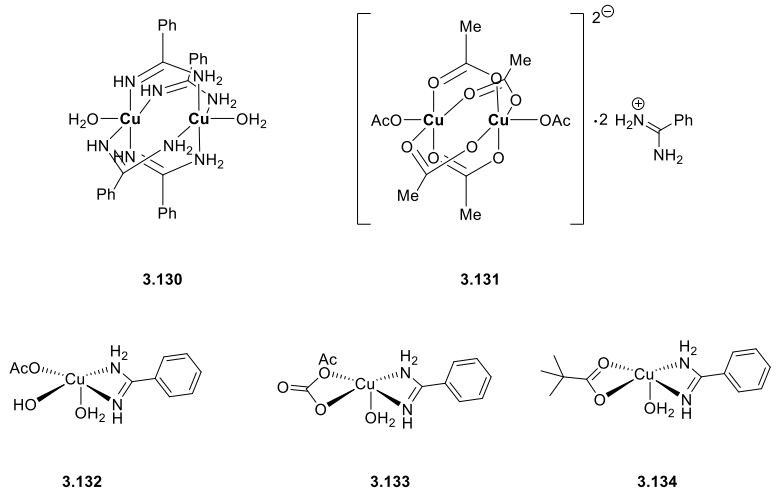
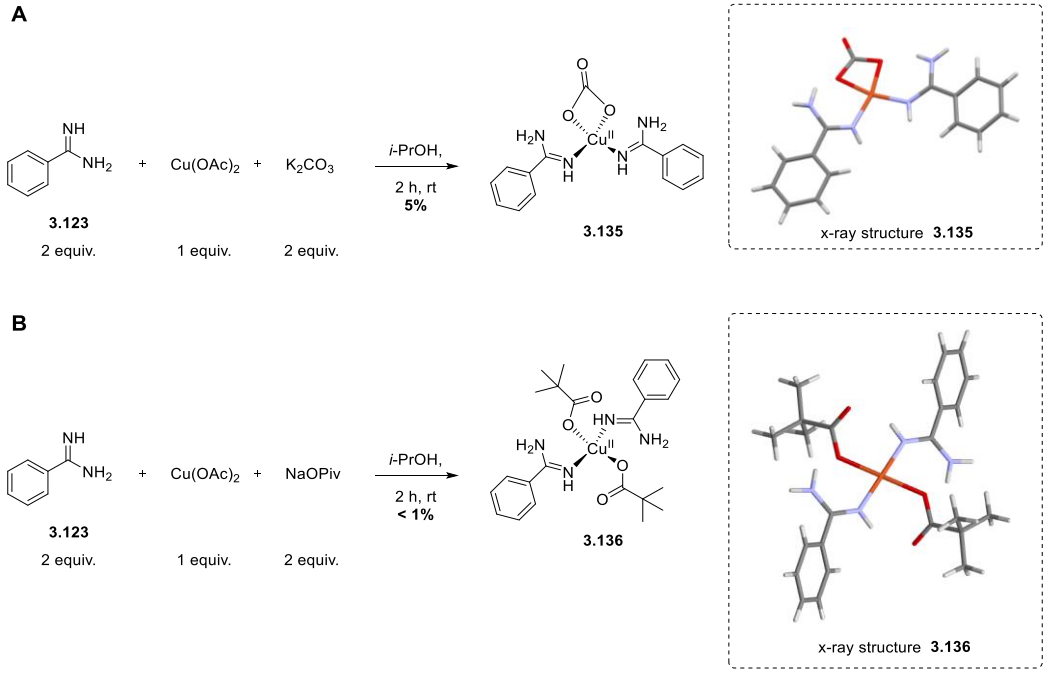


Figure 3.2. Proposed amidine-Cu species present in the Chan-Lam amidination.

To isolate any Cu intermediates formed prior to transmetalation in the Chan-Lam amidination the reaction was performed without any boronic acid (Scheme 3.18A). Performing the reaction without boronic acid resulted in the formation of a distinct purple precipitate. From the reaction optimisation one key observation was that a low conversion to product at substoichiometric levels of $Cu(OAc)_2$ coincided with the formation a purple precipitate (*Section 3.5.13*). It was postulated that a non-catalytic complex formed as the purple precipitate may provide an explanation for the low conversions at room temperature using catalytic amounts of $Cu(OAc)_2$. Attempts to recrystallise the precipitate first with MeOH, then *i*-PrOH failed. A change to MeCN afforded needle-like purple crystals. Analysis by x-ray crystallography confirmed the crystals to be a C_2 symmetric, square planar, mononuclear Cu^{II} complex with a carbonate ligand and two benzamidine ligands (Scheme 3.18A, **3.135**). Performing an analogous process using NaOPiv instead of K_2CO_3 and recrystalising from MeCN afforded a small amount of pale

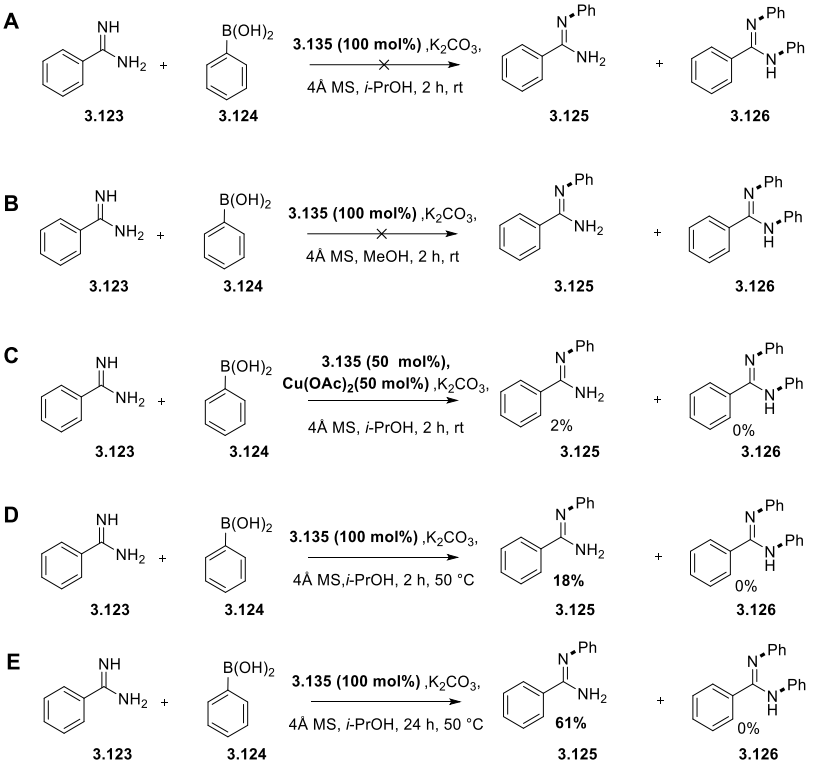
green crystals. X-ray crystallographic analysis revealed the structure to be another C_2 symmetric, square planar, mononuclear Cu^{II} complex with two pivalate ligands and two benzamidine ligands (Scheme 3.18B, **3.136**).



Scheme 3.12. (A) Synthesis of Cu complex **3.135** and crystal structure. (B) Synthesis of Cu complex **3.136** and crystal structure.

Both **3.135** and **3.136** have similar geometry but the carbonate in **3.135** acts as a bidentate ligand whereas in **3.136** pivalate acts as a monodentate ligand. To investigate the optimised conditions, complex **3.135** was used in several control experiments. Initially, it was hypothesised that complex **3.135** could be an intermediate prior to transmetalation in the Chan-Lam cycle. However, complex **3.135** in combination with either *i*-PrOH or MeOH as the reaction solvent produced no monoarylated or biarylated product (Scheme 3.19A & B). The lack of turnover suggests that **3.135** is not a reaction intermediate but potentially an off-cycle species. In an attempt to demonstrate activity of **3.135**, the complex was used in tandem with $Cu(OAc)_2$ (Scheme 3.19C). The hypothesis was that 100 mol% of Cu species (50 mol% **3.135** and 50 mol% $Cu(OAc)_2$) would provide comparable conversion to optimised conditions using 100 mol% $Cu(OAc)_2$ alone (Table 3.2, Entry 7). Only 2% conversion to mono-arylated amidine was observed with the tandem $Cu(OAc)_2$:**3.135** reaction. The results of the three experiments with **3.135** suggest that instead of a reaction intermediate **3.135** is a parasitic reaction species that cannot be activated with the addition Cu. As **3.135** was isolated without any boronic acid in the reaction mixture it is likely that **3.135** is an off-cycle side product, the formation of which is in competition with the transmetalation step. To examine the effect of temperature on the formation of **3.135** the amidination was completed at 50 °C for 2 h and 24 h (Scheme 3.19D and 3.19E). Interestingly, after 2 h at 50 °C a conversion of 18% to monoarlyated **3.125** was observed and after 24 h reaction time

at 50 °C 61% conversion to **3.125** was seen. The temperature experiments demonstrate that **3.135** is in equilibrium with active catalytic species and its formation is not an irreversible process.



Scheme 3.19. (A) Complex **3.135** experiment in *i*-PrOH at room temperature. (B) Complex **3.135** experiment in MeOH at room temperature. (C) Tandem $Cu(OAc)_2$: **3.135** experiment. (D) **3.135** temperature experiment for 2 h. (E) **3.135** temperature experiment for 24 h.

Since the isolation of **3.135** showed the importance of Cu valency, UV-vis analysis was then used to investigate the oxidation state of Cu during the reaction. UV-vis spectroscopy has been used previously to monitor the formation and disappearance of Cu^{II} species in Chan-Lam reactions where Cu^{II} species can be followed by measuring the absorbance at 650-700 nm.[229,230] Therefore, a standard amidination reaction was performed and using UV-vis analysis of reaction aliquots the loss of Cu^{II} species could be followed (Fig 3.3B).

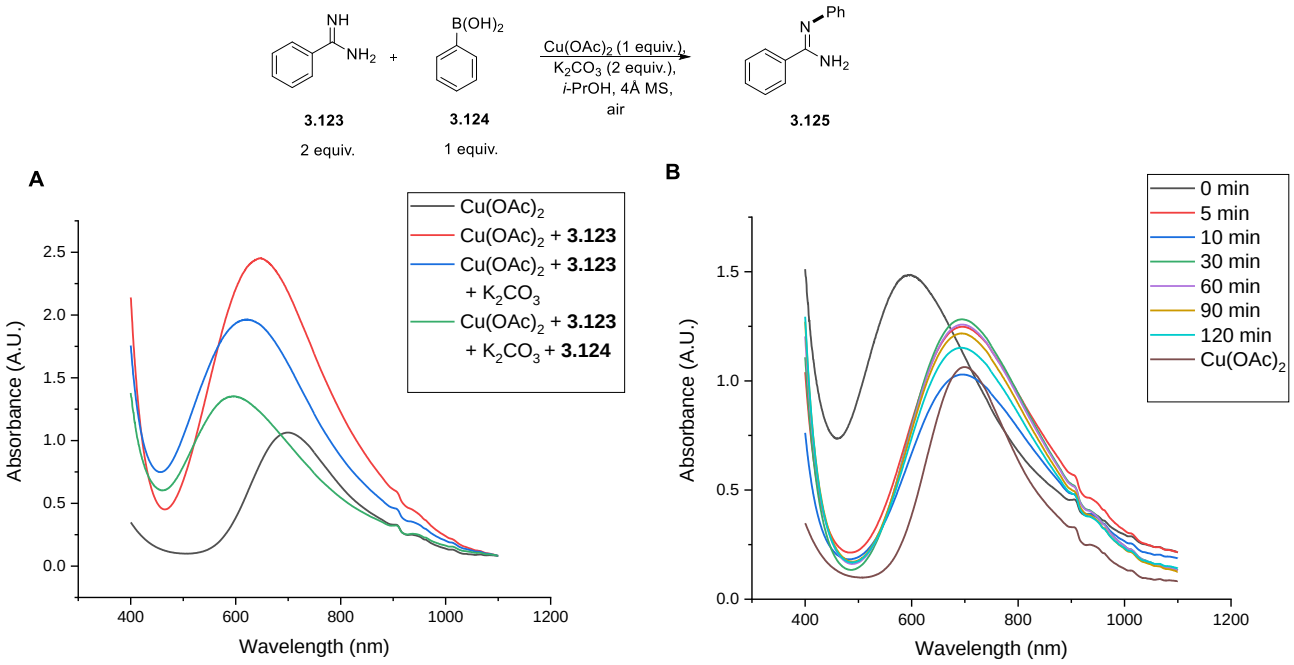


Figure 3.3. (A) UV-analysis of $Cu(OAc)_2$ with the different reaction reagents. (B) UV-vis analysis of the Chan-Lam amidination over 2 h.

Initial UV-vis analysis of $Cu(OAc)_2$ alone shows a signal at 700 nm confirming a Cu^{II} centre (Fig. 3.3A, black signal). Upon addition of amidine **3.123** there is a distinct shift to 650 nm, which suggests a possible change in ligand environment but not oxidation state (Fig. 3.7A, red signal). Further addition of K_2CO_3 and boronic acid **3.124** also showed shifts to 630 nm and 600 nm, again suggesting a change in Cu environment but not oxidation state. This phenomena of changing λ_{max} with changing Cu^{II} environment has been reported by both Becker *et al.* and Faller *et al.*[241,242]

The standard amidination reaction was monitored showing an initial peak at 600 nm at 0 min, like the control experiment, but a distinct shift to 700 nm is observed after 5 min reaction time (Fig 3.7B, black signal). Subsequent time points after 5 min show fluctuations in absorbance but no clear trend showing complete consumption of Cu^{II} species (Fig. 3.7, 5 min to 120 min signals). The UV-vis data suggests after 5 min there remains a continuous amount of Cu^{II} species throughout the reaction. It also suggests that the Cu species present immediately in the reaction is different to $Cu(OAc)_2$ and the other species in the reaction. The shift in wavelength could be attributed to the change in ligands on the Cu centre; the $Cu(OAc)_2$ denucleation or both processes.

To further investigate the Cu^{II} species in the reaction mixture EPR analysis can give semi-quantitative data as to the oxidation present. As discussed in *Section 3.1.5* $Cu(OAc)_2$ exists as a dimeric species that must undergo denucleation to give the active mononuclear Cu^{II} species.[227,243] EPR spectroscopy detects unpaired electrons in reaction mixtures and therefore can be used to characterise the Cu^{II} d^9 system. The $Cu(OAc)_2$ denucleation process has been previously monitored using EPR.[229,230,234] The formation of

EPR active mononuclear Cu^{II} can be monitored as it is paramagnetic whereas dinuclear $Cu(OAc)_2$ is not.

As complex **3.135** formed as a mononuclear species it was necessary to clarify which reagent in the reaction was causing the $Cu(OAc)_2$ denucleation. Analysis of the EPR spectra showed no observable signals of paramagnetic $Cu(OAc)_2$ in the presence of K_2CO_3 or **3.124** (Fig. 3.4A). This suggests that the dinuclear $Cu(OAc)_2$ species remains intact and neither the boronic acid or carbonate causes any denucleation. It has been reported that inorganic bases like K_3PO_4 can induce denucleation of $Cu(OAc)_2$ and it is therefore interesting that the oxygen bearing carbonate counterion has no effect on the $Cu(OAc)_2$ paddlewheel.[230] However, there is a clear EPR signal upon the addition of benzamidine **3.123** indicating the formation of a paramagnetic d^9 Cu(II) centre (Fig. 3.4B). A signal was also observed, to a lesser extent, upon addition of all reaction reagents (Fig. 3.4B). EPR analysis shows that the amidine alone causes the initial denucleation in the Chan-Lam reaction cycle. This amidine dependant denucleation suggests that amidine-Cu binding must occur first before carbonate binding on route to formation the of complex mononuclear **3.135**.

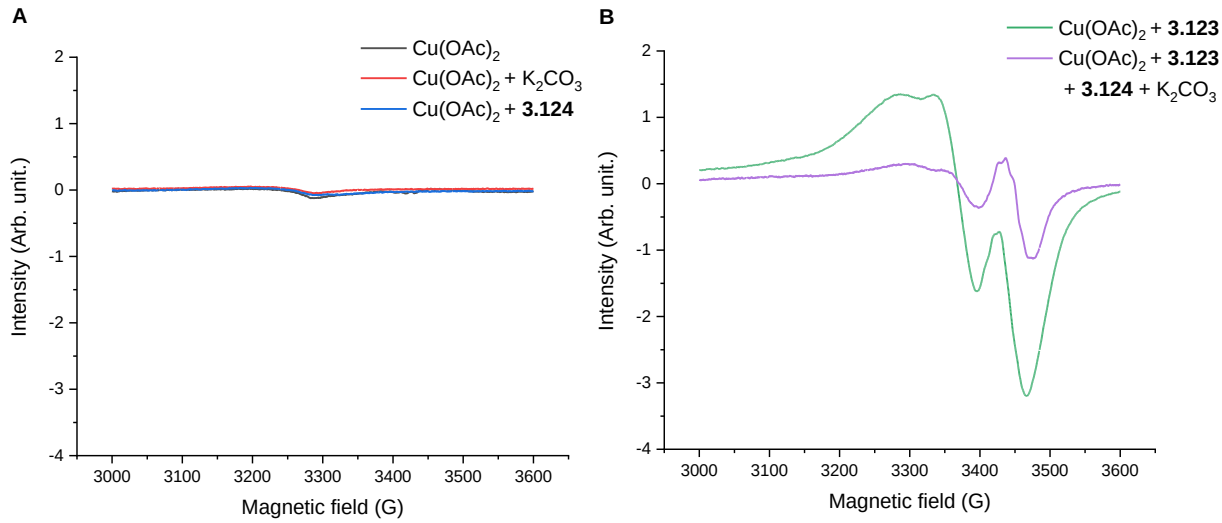
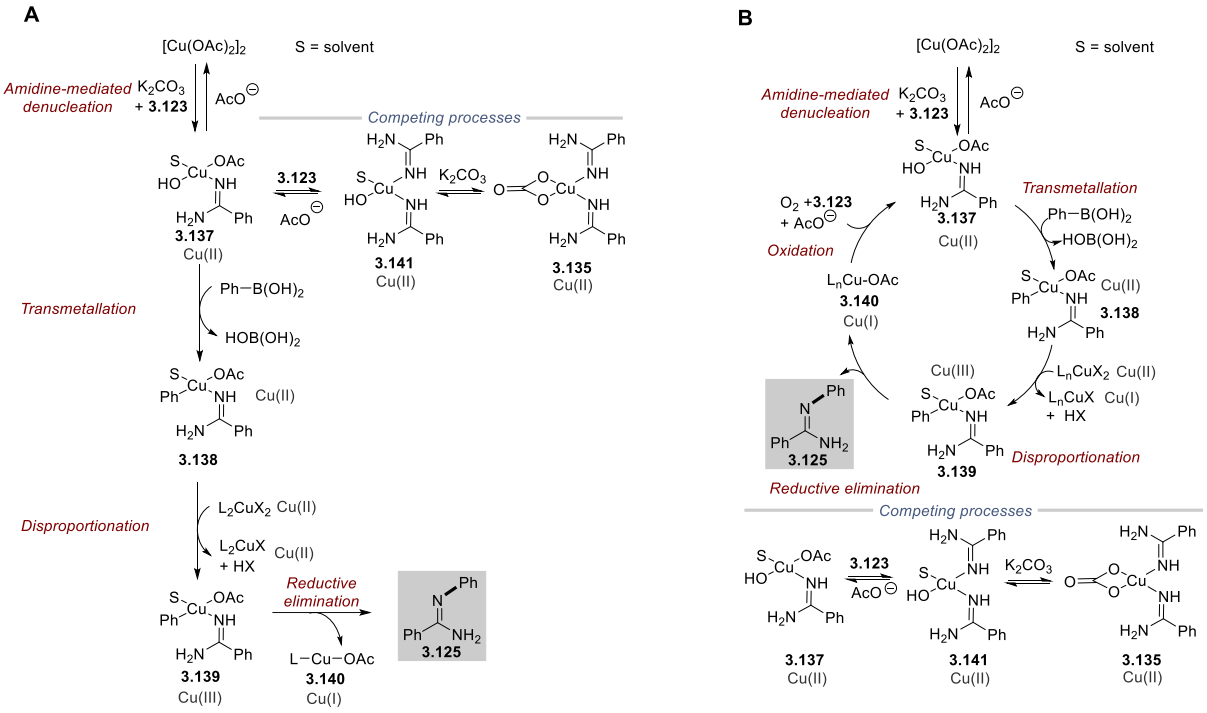


Figure 3.4. (A) EPR spectrum of $Cu(OAc)_2$ (black), $Cu(OAc)_2$ with K_2CO_3 (red) and $Cu(OAc)_2$ with $PhB(OH)_2$ (blue). (B) EPR spectra of $Cu(OAc)_2$ with amidine **3.123** (green) and the reaction mixture (purple).

Based on the results of UV-vis spectroscopy analysis, EPR spectroscopy analysis and control reactions using complex **3.135**, two mechanisms can be postulated. The first proposed mechanism concerns the use of stoichiometric $Cu(OAc)_2$ at room temperature (Scheme 3.20A). The proposed mechanism begins with an amidine-mediated denucleation from dimeric $Cu(OAc)_2$ to an active Cu^{II} amidine complex **3.137**, as suggested by EPR analysis. Complex **3.137** can then undergo transmetalation with boronic acid to give an aryl Cu^{II} species **3.138**. Analogous to the Watson and Stahl mechanisms a disproportionation step could then occur to yield a Cu^{III} complex **3.139**.[229,234] Complex **3.139** can undergo reductive elimination to liberate the monoarylated amidine **3.125**. As the reaction uses

stoichiometric amounts of $Cu(OAc)_2$ the Cu^I complex **3.140** may not undergo the final reoxidation step back to an active Cu^{II} species. Due to the isolation of **3.135** and its inability to facilitate any arylation at room temperature, a competitive reaction pathway is proposed where a second amidine binds to **3.137** forming **3.141** before transmetalation with K_2CO_3 to produce the Cu^{II} complex **3.135**. The competitive reaction pathway is favoured at sub-stoichiometric levels of $Cu(OAc)_2$ when the reaction is performed at room temperature, thus reducing the extent of transmetalation with boronic acid and reducing the conversion to **3.125**.

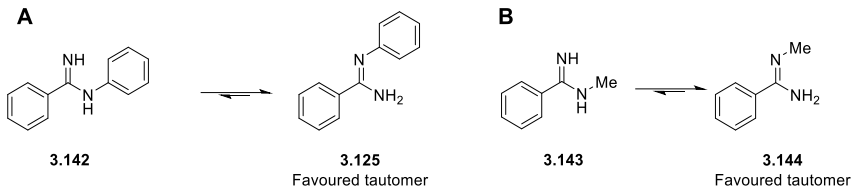
The second mechanism concerns the use of catalytic amounts of $Cu(OAc)_2$ at 50 °C (Scheme 3.20B). As with the non-catalytic mechanism amidine-mediated denucleation affords **3.137** before subsequent transmetalation, disproportionation and reductive elimination afford Cu^I **3.140** and amidine **3.125**. The key difference in the catalytic mechanism is the oxidation of **3.140** under air and reformation of **3.137** through the addition of **3.123** and acetate. The equilibrium established around the competitive formation of the pathway to form **3.135** is shifted to favour the formation of active species **3.137**.



Scheme 3.20.(A) Proposed mechanism for Chan-Lam amidination using stoichiometric $Cu(OAc)_2$. (B) Proposed mechanism for Chan-Lam amidination using catalytic $Cu(OAc)_2$.

3.3.3 Major *N*-arylated tautomer formed from the Chan-Lam amidination

To fully assign the structure of monosubstituted amidine products from the Chan-Lam amidination, and prior to a reaction scope being carried out, the identity of the major tautomer needed to be confirmed. The tautomerisation of *N*-substituted amidines has been reported by Clement *et al.*[244] Using ^{15}N NMR experiments the favoured tautomer was identified as substitution of the amidine at the imino nitrogen (Schemes 3.21A & B)



3.142 3.125 Favoured tautomer 3.143 3.144 Favoured tautomer

Scheme 3.21. Favoured tautomers of *N*-aryl amidines (A) and *N*-alkyl amidines (B) identified by Clement *et al*.[244]

Model substrate **3.125** was confirmed to have same tautomerisation pattern as reported first using variable temperature NMR experiments (Fig. 3.5).[244] The ^{1}H NMR of **3.125** at 300 K shows a broad singlet at around 5 ppm integrating to two protons, suggesting a single nitrogen-hydrogen bond environment and tautomer **3.125**. Upon cooling to 220 K the signal at 5 ppm begins to split into separate broad singlets, suggesting the formation of two separate nitrogen-hydrogen environments and tautomer **3.142**. The NMR experiments thus suggest that the product of the Chan-Lam amidination is favoured to tautomer **3.125** as reported by Clement *et al* and the amino substituted tautomer **3.142** can only be favoured at 220 K.[244]

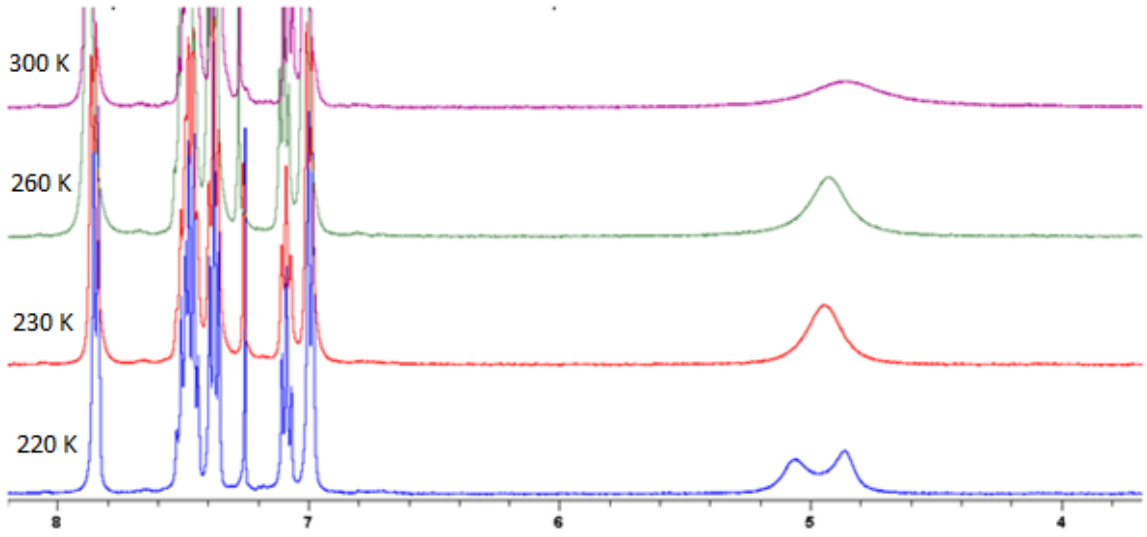
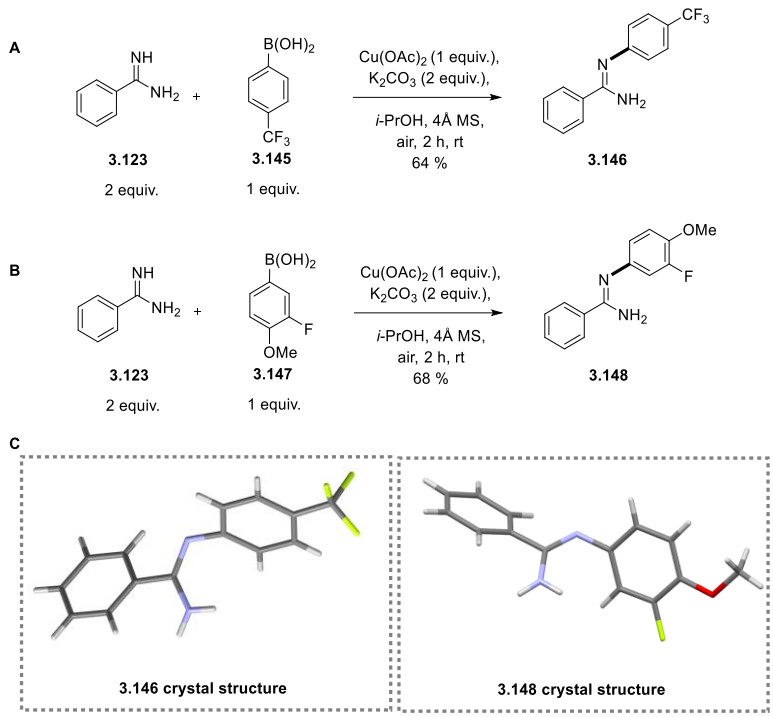


Figure 3.5. ^{1}H NMR spectra of **3.125** ran at decreasing temperatures. All NMR ran in $CDCl_3$.

In addition to the NMR experiments, two different substrates were synthesised using the Chan-Lam amidination (Scheme 3.22A & B). Using boronic acids **3.145** and **3.147** afforded substrates **3.145** and **3.148** in good yields of 64% and 68%. Both these substrates were able to be crystallised in high enough purity to be analysed by X-ray crystallography. The structures confirmed that substitution occurs at the imino nitrogen of the amidine and the imino bond lies in a *Z* configuration (Scheme 3.22C).



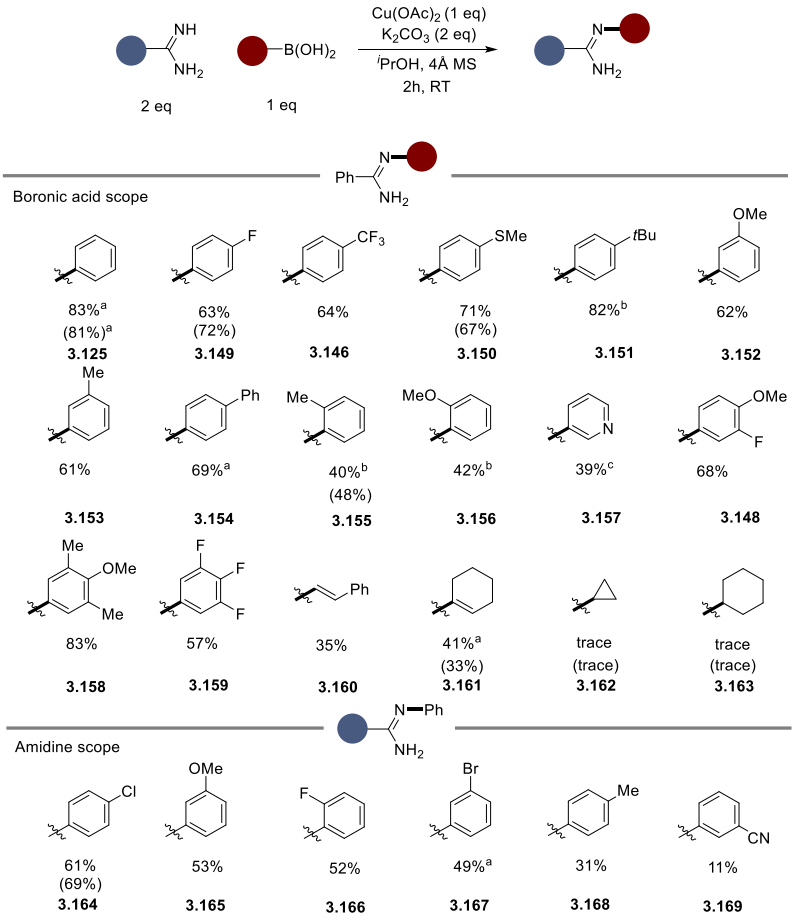
Scheme 3.22. (A) Chan-Lam amidination using (4-(trifluoromethyl) phenyl)boronic acid **3.145**. (B) Chan-Lam amidination using (3-fluoro-4-methoxyphenyl)boronic acid **3.147**. (C) Crystal structures of *N*-arylated amidine products **3.146** and **3.148**.

3.3.4 Substrate scope of the Chan-Lam amidination

To now assess the scope of the optimised conditions developed, a variety of boronic acids and amidine coupling partners were explored (Scheme 3.23). Using simple monosubstituted, electron-deficient boronic acids substrates, performed well **3.149** and **3.146** with isolated yields of 63% and 64%, respectively. More electron-rich boronic acids also gave good to excellent yields **3.150** to **3.154** ranging from 61% to 82% when either meta- or para-substituted. A switch to ortho-substituted boronic acids saw a return of more moderate yields for amidines **3.155** and **3.156**. The reduction in yield using ortho-substituted aryl boronic acids has been reported previously and can be attributed to the greater steric effect caused by ortho-substitution.[229,245] Turning to heterocyclic boronic acids, 3-pyridylboronic acid performed moderately well to yield **3.157**. Utilising a disubstituted boronic acid gave a good yield of 68% for **3.148**. Pleasingly, the use of tri-substituted boronic acids gave good to excellent yields of 83% for (4-methoxy-3,5-dimethylphenyl)boronic acid **3.158** and 57% for the highly electron-deficient (3,4,5-trifluorophenyl)boronic acid **3.159**. Some alkenyl species were also tolerated providing moderate yields of 35% and 41% of the corresponding *N*-arylated products **3.160** and **3.161**. Chan-Lam chemistry using alkyl boronic acids has so far been limited to only a few examples using cyclopropylboronic acids.[246,247] For the amidine system presented here a switch to cyclopropyl or cyclohexylboronic acids gave only trace amounts of products **3.162** and **3.163**. Furthermore, the appearance of a purple precipitate during these reactions indicated the formation of complex **3.135**. Complex **3.135** was confirmed was confirmed by a matching X-ray crystal structure after recrystallisation of the purple

precipitate isolated from the reaction mixture of **3.163**. The complex formation is likely due to the competing reaction pathway (Fig. 3.9A) dominating because of the poor transmetalation of the alkyl-boron species. Finally, several different aryl amidines were tolerated (**3.164** to **3.168**) in moderate to good yields of 31% to 61%. Only nitrile containing **3.169** returned a comparably low yield of 11%.

To assess the substoichiometric reaction selection of examples from the scope were repeated using 20 mol% $Cu(OAc)_2$ (Scheme 3.23, isolated yields in parentheses). Under the catalytic conditions, model substrate **3.125** was isolated in an excellent yield of 81%. Using an electron-deficient boronic acid afforded **3.149** in a yield of 72%. The electron-rich substrate **3.150** could be isolated in a yield of 67%, comparative to a yield of 71% isolated using stoichiometric conditions. Ortho-substituted **3.155** and alkene **3.161** again afforded a lower yield than other aryl groups with isolated yields of 48% and 33%, respectively. Unfortunately, no improvement was seen when alkyl boronic acids were used, with both alkyl-modified products **3.162** and **3.163** only detected in trace amounts. The poor performance of alkyl boronic acids suggests that the increased temperature and time only shift the equilibrium away from **3.135** but do not overcome the inherent transmetalation issues with alkyl boronic acids. Finally, amidine **3.164** was isolated in a good yield of 69%, an improvement on the stoichiometric conditions.



Scheme 3.23. Reaction scope of the optimised Chan-Lam amidination. Isolated yields on 0.40 mmol scale. (a) 1.00 mmol scale. (b) 0.20 mmol scale. (c) Isolated as the corresponding TFA salt. Yields quoted in parentheses are for catalytic conditions: $Cu(OAc)_2$ (20 mol%), K_2CO_3 (2 equiv.), 4 Å MS, *i*-PrOH, 50 °C, 24 h. No yield reported indicates no reaction attempted for catalytic conditions.

Both ortho-substituted boronic acids and heteroaryl boronic acids were inferior substrates compared to other aryl boronic acids. The lower yield of **3.156** is also observed in the work by Neuville *et al.* (Scheme 3.17), though still higher than the 48% observed here using the developed *i*-PrOH conditions. In addition, alkyl boronic acids afforded no product using either stoichiometric or catalytic conditions which is likely attributed to their poorer reactivity compared to aryl boronic acids.[236,247]

3.3.5 Chan-Lam amidination of Pentamidine

Pentamidine **3.170** is a DNA minor-groove binder that is used in the clinic against human African trypanosomiasis and binds to an AATT region of DNA with hydrogen bonds to A and C residues (Fig. 3.6B).[118,248] Synthesis of pentamidine and its analogues is often performed using pinner chemistry such as work by Tidwell *et al*.[249] Pinner chemistry will typically deliver symmetric analogues of pentamidine where both amidine moieties have been substituted (Fig. 3.6C).[249,250] Therefore the developed monosubstituted Chan-Lam chemistry was applied to the **3.170** scaffold to produce non-symmetric variants.

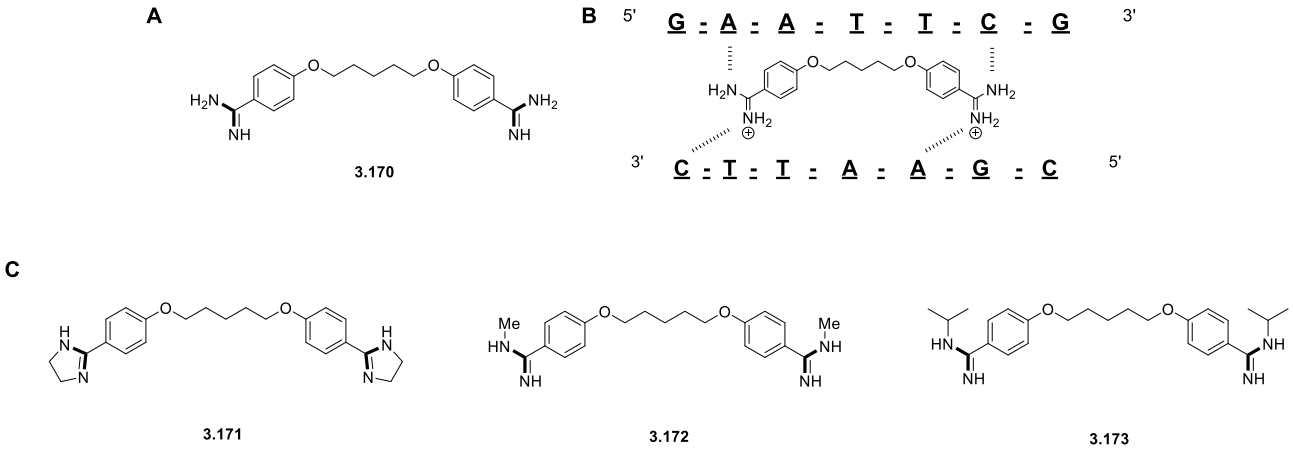
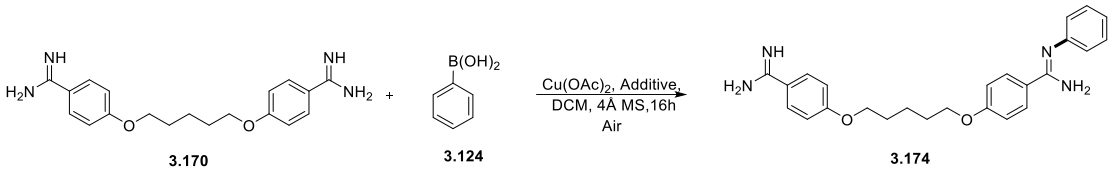


Figure 3.6. (A) Structure of pentamidine **3.170**. (B) Schematic representing how pentamidine interacts with the minor grove of DNA. (C) Structures of previously synthesised pentamidine analogues by Tidwell *et al*.

The investigation of Chan-Lam amidination on **3.170** was initially trialled using the slightly higher yielding stoichiometric conditions. With these conditions a modest 12% yield of the monoarylated diamidine **3.174** was achieved (Table 3.4, Entry 1). Due to the low amount of product formed the ratio of **3.170**:**3.124** was reduced to 1.0:1.0 which afforded the same yield of 12% (Table 3.4, Entry 2). The 1.0:1.0 ratio reaction identified that excess **3.170** does not lead to a higher yield in the pentamidine amidination reaction. Along with this, significant amounts of **3.170** was observed by HPLC after 16 h while no **3.124** was observed, suggesting unreacted amidine and decomposition of **3.124**. Maintaining the same ratio but changing the additive to the organic base Et_3N again returned the same yield of 12% (Table 3.4, Entry 3). Of particular note is the initial three reactions with **3.170** exhibited very low solubility in *i*-PrOH and this could be a potential reason for large amounts of unreacted **3.170**. To enhance solubility of **3.170** MeOH was used in combination with Et_3N which afforded **3.174** in an improved yield of 24% (Table 3.4, Entry 4). Although the change in solvent clearly improved the yield of **3.174**, significant amounts of starting material **3.170** were still observed during HPLC purification. Increasing the reaction temperature to 50 °C did not improve the yield of **3.174** and only returned 20% of **3.174** (Table 3.4, Entry 5). To increase the reactivity of **3.170** an excess of boronic acid **3.124** was used but this reaction only gave a 17% yield of **3.174** (Table 3.4, Entry 6). Finally, to compare the

conditions trialled here the amidination conditions reported by Neuville *et al.* were used but only afforded **3.174** in a yield of 19% (Table 3.4, Entry 7).[237]

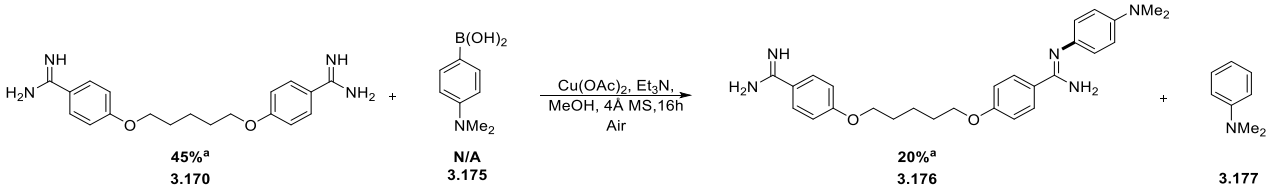
Table 3.4. Chan-Lam amidination reactions using diamidine **3.170**.



Entry	**3.170:3.124**	Solvent	Additive	$Cu(OAc)_2$ (mol%)	Rxn temp	**3.174** (%)[(a)]
1	2.0:1.0	*i*-PrOH	K_2CO_3	100	rt	12
2	1.0:1.0	*i*-PrOH	K_2CO_3	100	rt	12
3	1.0:1.0	*i*-PrOH	Et_3N	100	rt	12
4	**1.0:1.0**	**MeOH**	**Et_3N**	**100**	**rt**	**24**
5	1.0:1.0	MeOH	Et_3N	100	50 °C	20
6	1.0:2.0	MeOH	Et_3N	100	rt	17
7	1.0:1.2	DMF	NaOPiv	20	50 °C	19

(a) Isolated yields.

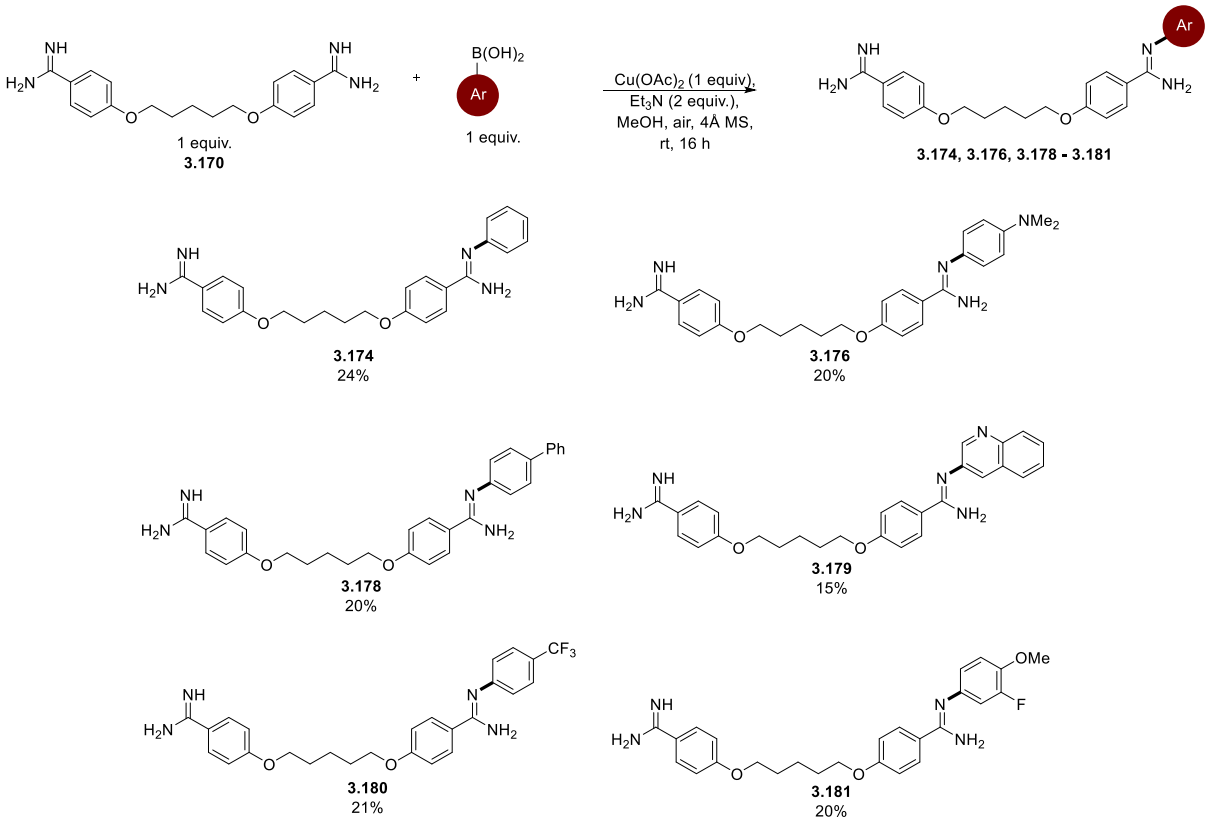
In summary the optimised amidination conditions for diamidine **3.170** (table 3.4, entry 4) returned 24% of the desired product **3.174**. These conditions were applied to boronic acid **3.175** where 45% starting material **3.170** was isolated with monoarylated product **3.176** isolated in a yield of 20%. The protodeborination product **3.177** was observed in LCMS analysis of the reaction but could not be isolated (Scheme 3.24). This experiment suggests that protodeborination of the boronic acid could be the underlying reason for low conversion of diamidine **3.170** to monosubstituted product.



Scheme 3.24. Chan-Lam amidination of **3.170** using boronic acid **3.175**. (a) Isolated yield.

To further probe the *N*-arylation of diamidines using the change in solvent and base a series of **3.170** analogues were synthesised using the optimised conditions (Scheme 3.25). The use of MeOH and Et_3N allowed access to six pentamidine analogues in 15-24% isolated yields. All of the major products formed were monosubstituted leading to six unsymmetrical pentamidine analogues. Using electron-rich boronic acids returned analogues **3.176** and **3.178** in 20% yields, slightly lower than the 24% obtained

for **3.174**. Using a heterocyclic boronic acid returned analogue **3.179** in a yield of 15%. Diamidine **3.180** was afforded in a yield of 21% using an electron-deficient boronic acid and a multi-substituted boronic acid was used to synthesise **3.181** in a yield of 20%. The scope demonstrates that singly substituted pentamidine analogues can be accessed using Chan-Lam amidination chemistry utilising both electron-rich and electron-deficient aryl boronic acids.



Scheme 3.25. Unsymmetrical pentamidine analogues synthesised using the Chan-Lam amidination reaction. Isolated yield on 0.20 mmol scale quoted.

3.3.6 Anti-trypanosome activity of *N*-arylated pentamidine analogues

To investigate the bioactivity of the monosubstituted pentamidine analogues synthesised analogues the compounds were tested against a model of human African trypanosomiasis (assays conducted by Dr Federica Giordani, University of Glasgow). Though pentamidine **3.170** is clinically approved for treatment against human African trypanosomiasis a resistant stain has developed.[118] The pentamidine analogues **3.174** and **3.179-3.181** were analysed for their potency and EC_{50} values were determined against both the wildtype and pentamidine-resistant trypanosomiasis. As a comparison, pentamidine **3.170**, diminazene **3.182** and pentamidine isethionate **3.183** were used as positive controls in the assays (Fig. 3.7A). The viability of trypanosomes was assessed using an alamarBlue assay. As discussed in *section 2.3.1* alamarBlue **3.184** is permeable to cells, in this case trypanosomes, but not fluorescent however it is reduced by the metabolic environment of a living cell (NADPH and NADH can both

reduce **3.184**) to the highly fluorescent **3.185** (Fig 3.7B).[178] The fluorescent output of live trypanosomes is high whereas dead trypanosomes, which cannot reduce **3.184** give a low fluorescent readout.

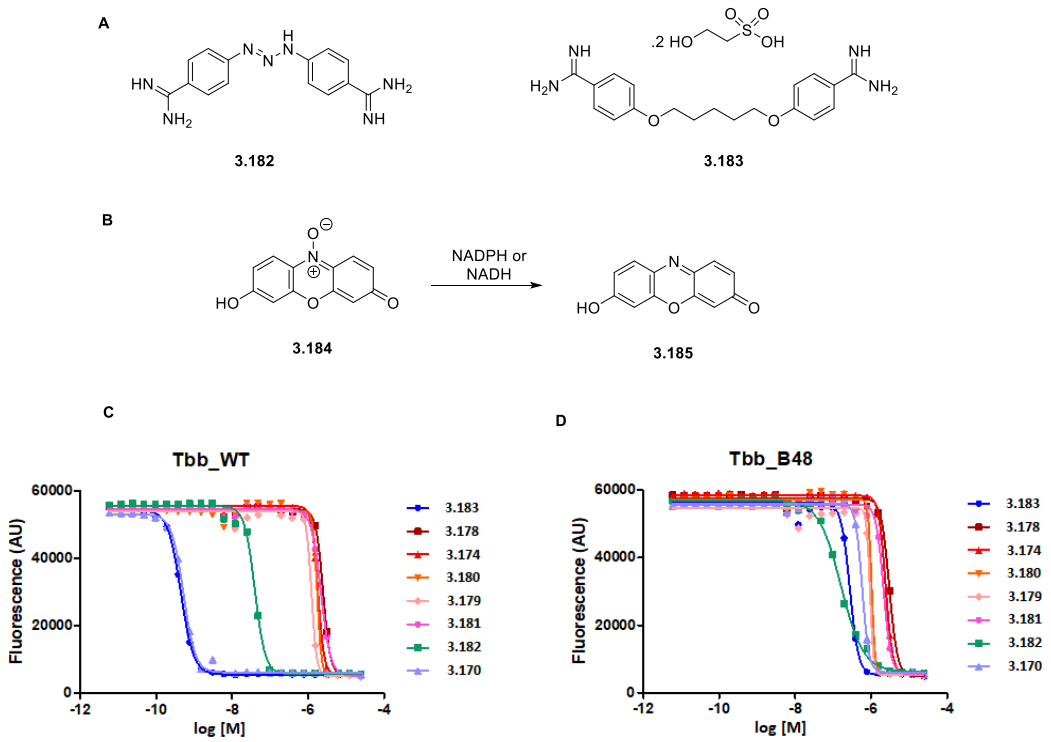
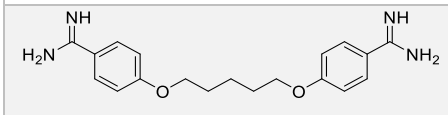
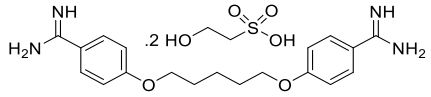
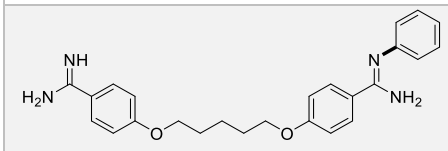
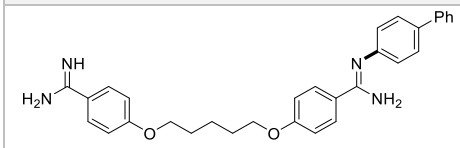
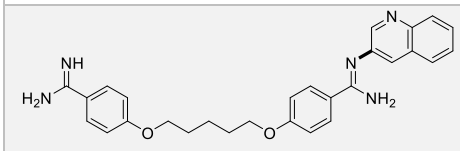
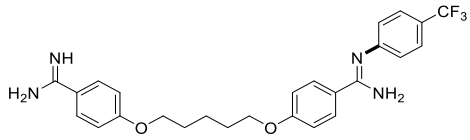
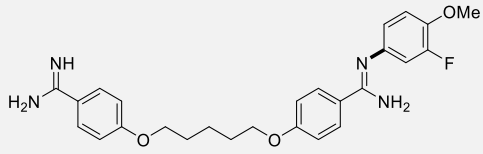


Figure 3.7. (A) Structures of diminazene **3.182** and pentamidine isethionate **3.183**. (B) Reduction of Alamar Blue **3.184** to its fluorescent analogue **3.185**. (C) Viability of wildtype trypanosomes against pentamidine and synthesised analogues. (D) Viability of pentamidine-resistant trypanosomes against pentamidine and synthesised analogues.

The experiments against the wild-type trypanosomes (Tbb_WT) confirmed both pentamidines **3.170** and **3.183** had high anti-trypanosome activity with EC_{50} values of 0.58 nM and 0.50 nM (Fig. 3.14C, Table 3.5). Unfortunately, all of the analogues synthesised showed significantly lower activity against the wild-type trypanosomes with **3.180** showing the lowest EC_{50} of 1.57 µM (Fig 3.7C, Table 3.5). Looking at the pentamidine-resistant trypanosomes it can be seen that **3.170** and **3.183** return significantly poorer EC_{50} values of 594 nM and 237 nM (Fig, 3.7D, Table 3.5). As with the wild-type cell line, all analogues showed lower activity than unsubstituted pentamidines in the resistant strain of trypanosomes (Tbb_B48, Fig 3.7D). The most active analogue was **3.179** returning an EC_{50} value of 1.04 µM (Fig. 3.7D, Table 3.5). Although each analogue had lower activity than unsubstituted **3.170** it can be noted that there was no significant reduction in their activity between both strains of trypansomes. The poorer EC_{50} values could be attributed to either poor DNA binding affinity or low cell permeability and could be assessed by DNA melting analysis.

Table 3.5. EC_{50} values of pentamidine and analogues against Tbb_WT and Tbb_B48 trypanosoma cell lines.

Compound	**EC_{50} Tbb_WT (nM)**	**EC_{50} Tbb_B48 (nM)**
	0.580	593.7
	0.500	236.9
	1847	2263
	1688	1032
	2409	2405
	1572	1040
	2320	1799

3.4 Summary of Chapter 3

This chapter has disclosed a facile method for the preparation of mono-arylated amidines using mild conditions with a variety of readily available coupling partners. The reaction can be performed using either catalytic or stoichiometric amounts of $Cu(OAc)_2$ but the catalytic reactions require longer reaction times and heating to 50 °C. The stoichiometric reaction proceeds under air at room temperature without the need for the addition of Cu-stabilising ligands or significantly elevated temperatures, unlike analogous cross coupling reactions such as Ullmann or Buchwald-Hartwig processes.[209,221–224]

Mechanistic studies showed that an amidine denucleates $Cu(OAc)_2$, followed by transmetalation with a boronic acid. Of particular note is the transmetalation step is in competition with the formation of **3.135**. Complex **3.135** is an off-cycle mononuclear Cu^{II} centre that is unlike previous multi nuclear species isolated in Chan-Lam reactions with amine nucleophiles.[229,240] Furthermore, **3.135** is catalytically competent at elevated temperatures meaning its formation is not always detrimental to reaction progress. The reaction can tolerate a variety of boronic acid coupling partners, including alkenyl boronic acids with a significantly broader scope of the reaction is demonstrated than previously reported by Neuville *et al.* (Scheme 3.16). However, alkyl boronic acids are not tolerated under either stoichiometric or catalytic conditions most likely due to their poor transmetalation capability. A variety of aryl amidine substrates also return mono *N*-arylated products in modest to good yields. Finally, the synthesis of six new pentamidine analogues is highlighted. These high value, unsymmetrical analogues have not been produced previously using standard Pinner chemistry, which is prone to over addition to the symmetric analogues. Each analogue shows mild activity against trypanosomes but still lower than the unsubstituted, therapeutic, pentamidine.

3.5 Experimental

3.5.1 General information

All reagents and solvents were obtained from commercial suppliers and were used without further purification. Pentamidine free base was synthesised from pentamidine isethionate by previously reported methods.[251] All glassware was oven-dried (150 °C) before use. All reactions were carried out under air. Reactions were monitored by thin layer chromatography (TLC) using Merck silica plates coated with fluorescent indicator UV254. TLC plates were analysed using 254/365 nm UV light or developed using potassium permanganate solution.

3.5.2 Analytical HPLC

RP-HPLC was carried out using an Aeris 3.6 µm RP 250 x 4.6 mm widepore XB C18 column using a DIONEX 3000 series HPLC system equipped with a VWD3400 photodiode array detector. Analysis was performed using the method outlined below.

The absorbance was detected at 254 nm. Solvent A: Water (0.1 % TFA) Solvent B: Acetonitrile (0.1 % TFA). Flow rate 1.0 mL/min

Table 3.6. Analytical HPLC method A.

Time (min)	Solvent A	Solvent B
0	99 %	1 %
5	99 %	1%
30	50 %	50 %
35	50 %	50 %
40	99 %	1 %
45	99 %	1 %

3.5.3 Purification of products

Normal-phase flash chromatography was carried out using ZEOprep 60 HYD 40-63 µm silica gel. Semi-preparative reversed-phase HPLC purification was carried out using a Phenomenex Clarity 5 µm RP 250 x 10.00 mm XB C18 column using a DIONEX 3000 series HPLC system equipped with a VWD3400 variable wavelength detector. Purifications were performed using 2 different systems using water (0.1% TFA) as Solvent A and acetonitrile (0.1% TFA) as Solvent B. The flow rate was set to 12.0 mL/min. RP-HPLC method **A**: The absorbance was detected at 254 nm. Solvent A: Water. Solvent B: Acetonitrile.

Table 3.7. RP-HPLC method A.

Time (min)	Solvent A	Solvent B
0	99 %	1 %
5	99 %	1%
45	50 %	50 %
50	50 %	50 %
55	99 %	1 %
60	99 %	1 %

RP-HPLC method **B**: The absorbance was detected at 254 nm. Solvent A: Water (0.1 % TFA). Solvent B : Acetonitrile (0.1 % TFA).

Table 3.8. RP-HPLC method B**.**

Time (min)	Solvent A	Solvent B
0	99 %	1 %
5	99 %	1%
25	50 %	50 %
30	50 %	50 %
35	10 %	90 %
40	10 %	90 %
45	99 %	1 %
50	99 %	1 %

3.5.4 Analysis of products

Fourier-Transform Infra-Red (FTIR) spectra were obtained on a Shimadzu IRAffinity-1 spectrometer. ^{19}F NMR spectra were obtained on a Bruker AV 400 spectrometer at 376 MHz. ^{1}H and ^{13}C NMR spectra were obtained on either a Bruker AV 400 at 400 MHz and 125 MHz, respectively, or Bruker DRX 500 at 500 MHz and 126 MHz, respectively. Chemical shifts are reported in ppm and coupling constants are reported in Hz. High-resolution mass spectra were recorded on an LTQ Orbitrap XL 1 mass spectrometer at the EPSRC UK National Mass Spectrometry Facility (Swansea) and a Thermo Exactive Orbitrap mass spectrometer at the University of St Andrews.

3.5.5 Sample preparation for EPR spectroscopy

To 10 mL glass vials was added $Cu(OAc)_2$ (36.0 mg, 0.20 mmol) in *i*-PrOH (5.20 mL) to give a solution of 38.5 mM $Cu(OAc)_2$. Appropriate additive quantities were used as detailed in Table 3.9.

Table 3.9. Samples prepared for EPR analysis**.**

Sample	Additive(s)	Mass, mmol
1	N/A	N/A
2	benzamidine	48.0 mg, 0.40 mmol
3	K_2CO_3	55.0 mg, 0.40 mmol
4	$PhB(OH)_2$	24.0 mg, 0.20 mmol
5	benzamidine, K_2CO_3 and $PhB(OH)_2$	48.0 mg 0.40 mmol, 55.0 mg 0.40 mmol, 24.0 mg 0.20 mmol

3.5.6 EPR measurements

The EPR spectra were obtained using a Bruker EMX 10/12 spectrometer operating at ~9.5 GHz with 100 kHz modulation. Measurements were performed using an ELEXSYS Super High Sensitivity Probehead (Bruker ER4122SHQE). The EPR spectra were recorded at 293 K using 10 mW microwave power, 2000 G field sweep centred at 3300 G with 1024 points resolution. A time constant and conversion time of 40.96 ms was used with a modulation amplitude of 0.8 mT and a microwave frequency of 9.766 GHz. EPR measurements carried out by Dr Bela Bode, University of St. Andrews.

3.5.7 Synthetic procedures

Chan-Lam procedure A

To an oven-dried microwave vial was added the corresponding amidine (1.00 mmol), boronic acid (2.00 mmol), $Cu(OAc)_2$ (181 mg, 1.00 mmol), NEt_3 (278 µL, 2.00 mmol) and activated 4Å molecular sieves (1.00 g). Dichloromethane (3.25 mL) was then added to the solid mixture, sealed under air and stirred at room temperature for 16 h. The resulting suspension was filtered and the precipitate was washed with MeOH (3.00 mL). The resulting solution was concentrated *in vacuo* and the residue purified by silica gel chromatography.

Chan-Lam procedure B

To an oven-dried microwave vial was added the corresponding amidine (0.40 mmol), boronic acid (0.20 mmol), $Cu(OAc)_2$ (36.0 mg, 0.20 mmol), K_2CO_3 (55.0 mg, 0.40 mmol) and activated 4Å molecular sieves (200 mg). Isopropanol (0.65 mL) was then added to the solid mixture, sealed under air and stirred at room temperature for 2 h. The resulting suspension was filtered and the precipitate was washed with MeOH (5 × 1 mL). The resulting solution was concentrated *in vacuo* and the residue purified by either silica gel chromatography or RP-HPLC.

Chan-Lam procedure C

To an oven-dried microwave vial was added the corresponding amidine (0.40 mmol), boronic acid (0.20 mmol), $Cu(OAc)_2$ (7.2 mg, 0.04 mmol), K_2CO_3 (55 mg, 0.40 mmol) and activated 4Å molecular sieves (200 mg). Isopropanol (0.65 mL) was then added to the solid mixture, sealed under air and stirred in a sand bath at 50 °C for 24 h. The resulting suspension was filtered and the precipitate was washed with MeOH (5 × 1 mL). The resulting solution was concentrated *in vacuo* and the residue purified by either silica gel chromatography or RP-HPLC.

Chan-Lam procedure D

To an oven dried microwave vial was added pentamidine free base (68.0 mg, 0.20 mmol), boronic acid (0.20 mmol), $Cu(OAc)_2$ (36.0 mg, 0.20 mmol) and activated 4Å molecular sieves (200 mg). MeOH

(0.65 mL) was then added to the solid mixture before final addition of Et_3N (56 µL, 0.40 mmol), sealed under air and stirred at room temperature for 16 h. The resulting suspension was filtered, and the precipitate was washed with MeOH (5 × 1.50 mL). The resulting solution was purified by semi-preparative RP-HPLC and lyophilised to afford the amidine product.

3.5.8 Reaction optimisation

To an oven dried 5 mL microwave vial was added the corresponding amidine (0.40 mmol), phenyl boronic acid (0.20 mmol), $Cu(OAc)_2$ (as per Table 3.10), base or additive (0.40 mmol) and activated 4Å molecular sieves (200 mg). Solvent (0.65 mL) was then added to the solid mixture, sealed under air and stirred at room temperature for a defined time period. A 50 µL aliquot of the reaction was removed and diluted with 950 µL of MeOH. The resulting sample was then analysed by HPLC using analytical HPLC method A. Conversion was calculated using standard concentration curves of *N*'-phenylbenzimidamide (**3.125**) TFA salt and *N*,*N*'-diphenylbenzimidamide free base (**3.126**).

Table 3.10. Chan-Lam amidination optimisation of conditions.

Entry	**Solvent**	**Reaction time**	**Additive**	**$Cu(OAc)_2$ eq**	**Reaction temp**	**3.125 conv.[a]**	**3.126 conv.[a]**
1	DCM	16 h	$B(OH)_3$	1.0	rt	45%	11%
2	DCM	16 h	Et_3N	1.0	rt	58% (59%)[b]	14%
3	MeOH	16 h	K_2CO_3	1.0	rt	59%	2%
4	MeOH	16 h	K_2CO_3	0.5	rt	58%	2%
5	MeOH	16 h	K_2CO_3	1.0	80°C	36%	<1%
6	DMSO	16 h	Et_3N	1.0	rt	47%	2%
7	DMSO	16 h	K_2CO_3	1.0	rt	59%	<1%
8	MeCN	16 h	K_2CO_3	1.0	rt	52%	8%
9	*i*-PrOH	16 h	$B(OH)_3$	1.0	rt	38%	14%
10	***i*-PrOH**	**16 h**	**K_2CO_3**	**1.0**	**rt**	**70% (72%)[b]**	**4%**
11	*i*-PrOH	16 h	Et_3N	1.0	rt	70%	6%
12	*i*-PrOH	16 h	K_2CO_3	0.5	rt	7%	0%
13	*i*-PrOH	16 h	K_2CO_3	1.0	80°C	61%	0%
14	*i*-PrOH	16 h	K_2CO_3	1.2	rt	62%	4%
15	*i*-PrOH	16 h	K_2CO_3	1.5	rt	60%	4%
16	*i*-PrOH	16 h	K_2CO_3	2.0	rt	59%	4%
17	*i*-PrOH	16 h	K_2CO_3	1.0	0-10°C	63%	5%
18	*i*-PrOH	16 h	Pyridine	1.0	rt	59%	3%

19	*i*-PrOH	8 h	K_2CO_3	1.0	rt	63%	4%
20	*i*-PrOH	4 h	K_2CO_3	1.0	rt	63%	3%
21	***i*-PrOH**	**2 h**	**K_2CO_3**	**1.0**	**rt**	**76% (83%)[b]**	**5%**
22	*i*-PrOH	1 h	K_2CO_3	1.0	rt	58%	3%
23	*i*-PrOH	0.5 h	K_2CO_3	1.0	rt	60%	3%
24	*i*-PrOH	48 h	K_2CO_3	1.0	rt	69%	7%
25	*i*-PrOH	2 h	K_2CO_3	0.5	rt	2%	0%
26	*i*-PrOH	2 h	K_2CO_3 + MnO_2 (3 eq.)	0.5	rt	2%	0%
27	*i*-PrOH	2 h	None	1.0	rt	3%	0%
28	*i*-PrOH:MeOH (3:1)	16 h	K_2CO_3	1.0	rt	63%	<1%
29	*i*-PrOH:MeOH (1:1)	16 h	K_2CO_3	1.0	rt	52%	<1%
30	^{t}BuOH	16 h	K_2CO_3	1.0	rt	49%	10%
31	H_2O	16 h	K_2CO_3	1.0	rt	<1%	0%
32	DMF	2 h	K_2CO_3	1.0	rt	31%	2%
33	DMF	16 h	K_2CO_3	1.0	rt	55%	<1%
34	DCE	2 h	K_2CO_3	1.0	rt	40%	3%
35	MeOH	2 h	Et_3N	1.0	rt	51%	<1%
36	*i*-PrOH	2 h	K_2CO_3	0.5	50°C	56%	--
37	*i*-PrOH	2 h	K_2CO_3	0.2	50°C	62%	--
38	*i*-PrOH	2 h	K_2CO_3	0.1	50°C	18%	--
39	*i*-PrOH	24 h	K_2CO_3	0.5	50°C	76%	2%
40	***i*-PrOH**	**24 h**	**K_2CO_3**	**0.2**	**50°C**	**81%[b]**	**--**
41	*i*-PrOH	24 h	K_2CO_3	0.1	50°C	64%	--
42	DMF	24 h	NaOPiv	0.5	50°C	72%[c]	7%[c]
43	DMF	24 h	NaOPiv	0.2	50°C	81%[c]	2%[c]
44	DMF	2 h	NaOPiv	0.2	50°C	54%[c]	--
45	*i*-PrOH	2 h	NaOPiv	1.0	rt	66%	4%
46	*i*-PrOH	2 h	NaOPiv	0.5	rt	64%	2%
47	*i*-PrOH	2 h	NaOPiv	0.2	rt	14%	--
48	*i*-PrOH	2 h	NaOPiv	1.0	50°C	71%	5%
49	*i*-PrOH	2 h	NaOPiv	0.5	50°C	68%	2%
50	*i*-PrOH	2 h	NaOPiv	0.2	50°C	47%	--
51	*i*-PrOH	24 h	NaOPiv	1.0	50°C	63%	5%
52	*i*-PrOH	24 h	NaOPiv	0.5	50°C	69%	3%
53	*i*-PrOH	24 h	NaOPiv	0.2	50°C	47%	--
54	DCM	16h	K_2CO_3	100	rt	40	8
55	MeOH	16h	K_2CO_3	100	rt	59	2

a) Conversion calculated using standard calibration curve b) Isolated yield on 1.00 mmol scale c) 0.25 M reaction concentration.

3.5.9 Cu source/benzamidine eq trials

To an oven dried 5 mL microwave vial was added the benzamidine (as per Table 3.11), phenyl boronic acid (0.20 mmol), Cu source (0.20 mmol), base or additive (0.40 mmol) and activated 4Å molecular sieves (200 mg). Solvent (0.65 mL) was then added to the solid mixture, sealed under air and stirred at room temperature for desired time. A 50 µL aliquot of the reaction was removed and diluted with 950 µL of MeOH. The resulting sample was then analysed by HPLC using analytical HPLC method A. Conversion was calculated using standard concentration curves of *N*'-phenylbenzimidamide (**3.125**) TFA salt and *N*,*N*'-diphenylbenzimidamide (**3.126**).

Table 3.11. Reaction optimisation using altering Cu source and eq of benzamidine.

Entry	**Solvent**	**Reaction time**	**Additive**	**Cu source**	**3.123 eq.**	**3.125 conv.[a]**	**3.126 conv.[a]**
1	MeOH	16 h	K_2CO_3	$Cu(OAc)_2$	1.5	49%	2%
2	MeOH	16 h	K_2CO_3	$Cu(OAc)_2$	1.2	43%	2%
3	DCM	16 h	$B(OH)_3$	$Cu(OAc)_2$	2.0	26%[b]	5%[b]
4	*i*-PrOH	16 h	None	$Cu_2(OH)_2CO_3$	2.0	0%	0%
5	*i*-PrOH	16 h	K_2CO_3	$Cu(OAc)_2$	1.5	63%	5%
6	*i*-PrOH	16 h	K_2CO_3	$Cu_2(OH)_2CO_3$	2.0	<1%	0%
7	*i*-PrOH	16 h	K_2CO_3	$Cu(OAc)_2$	1.0	46%	6%
8	*i*-PrOH	2 h	K_2CO_3	$[Cu(MeCN)]PF_6$ (0.5 eq)	2.0	5%	0%
9	*i*-PrOH	16 h	K_2CO_3	$[Cu(MeCN)]PF_6$	2.0	35%	3%
10	*i*-PrOH	2 h	K_2CO_3	CuOAc	2.0	2%	0%
11	MeOH	2 h	K_2CO_3	CuOAc	2.0	14%	0%
12	*i*-PrOH	2 h	K_2CO_3 + KOAc (1 eq)	CuOAc	2.0	<1%	0%
13	*i*-PrOH	2 h	K_2CO_3	Complex **3.135**	2.0	2%	0%
14	*i*-PrOH	2 h	K_2CO_3	Complex **3.135 (**0.5 eq), $Cu(OAc)_2$ (0.5 eq)	2.0	2%	0%

a) Conversion calculated using standard calibration curve b) Isolated yield on 1.00 mmol scale

3.5.10 Representative HPLC trace (Table 3.10, Entry 21)

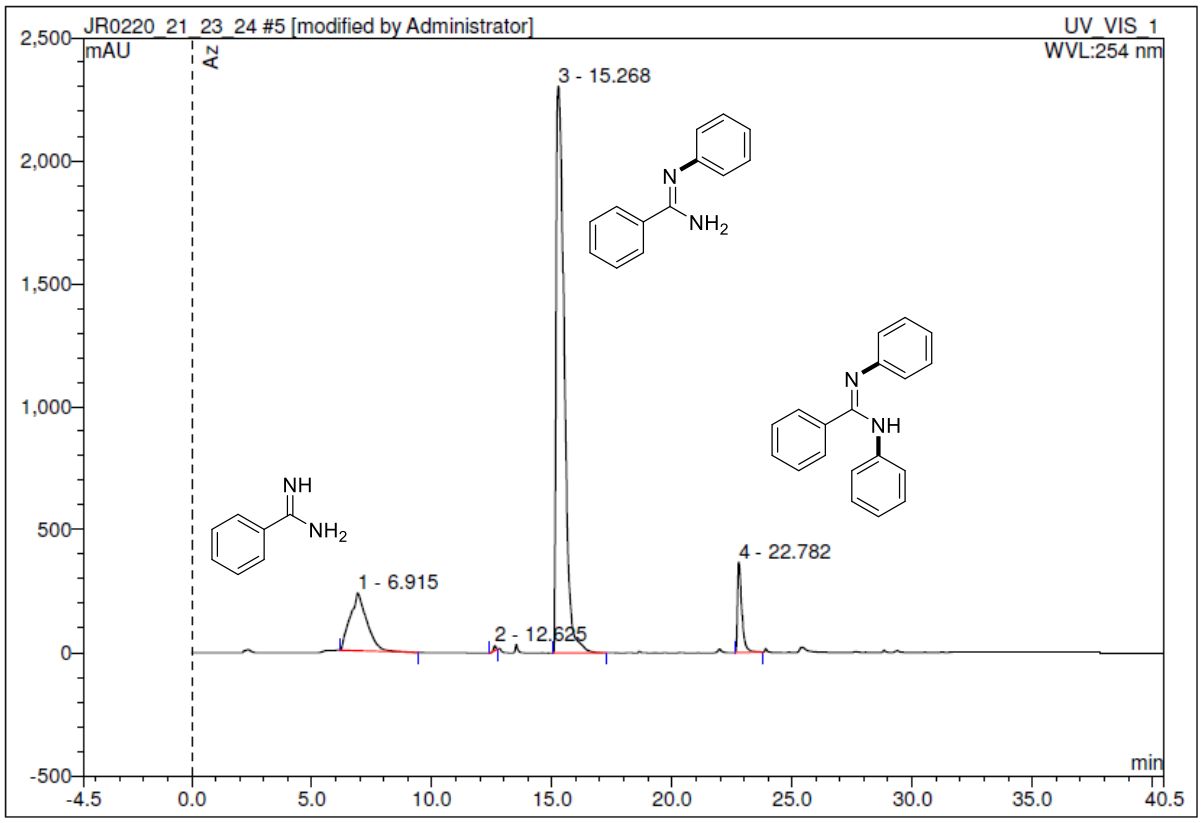


Figure 3.8. HPLC trace showing starting material **3.123** (7.2 min), major product **3.125** (15.4 min) and minor product **3.126** (22.9 min).

3.5.11 EPR results

Results detailed in Figure 3.16 illustrate de-nucleation of the $Cu(OAc)_2$ dimer in the presence of benzamidine (green) to give a mononuclear Cu^{II} signal. $Cu(OAc)_2$ gives no signal (black) and addition of K_2CO_3 (red) or $PhB(OH)_2$ (blue) provides no de-nucleation. Mononuclear Cu^{II} signal again observed in the presence of amidine with K_2CO_3 and phenyl boronic acid (purple). Note precipitate formation for the sample containing all reagents shows a less intense signal (purple).

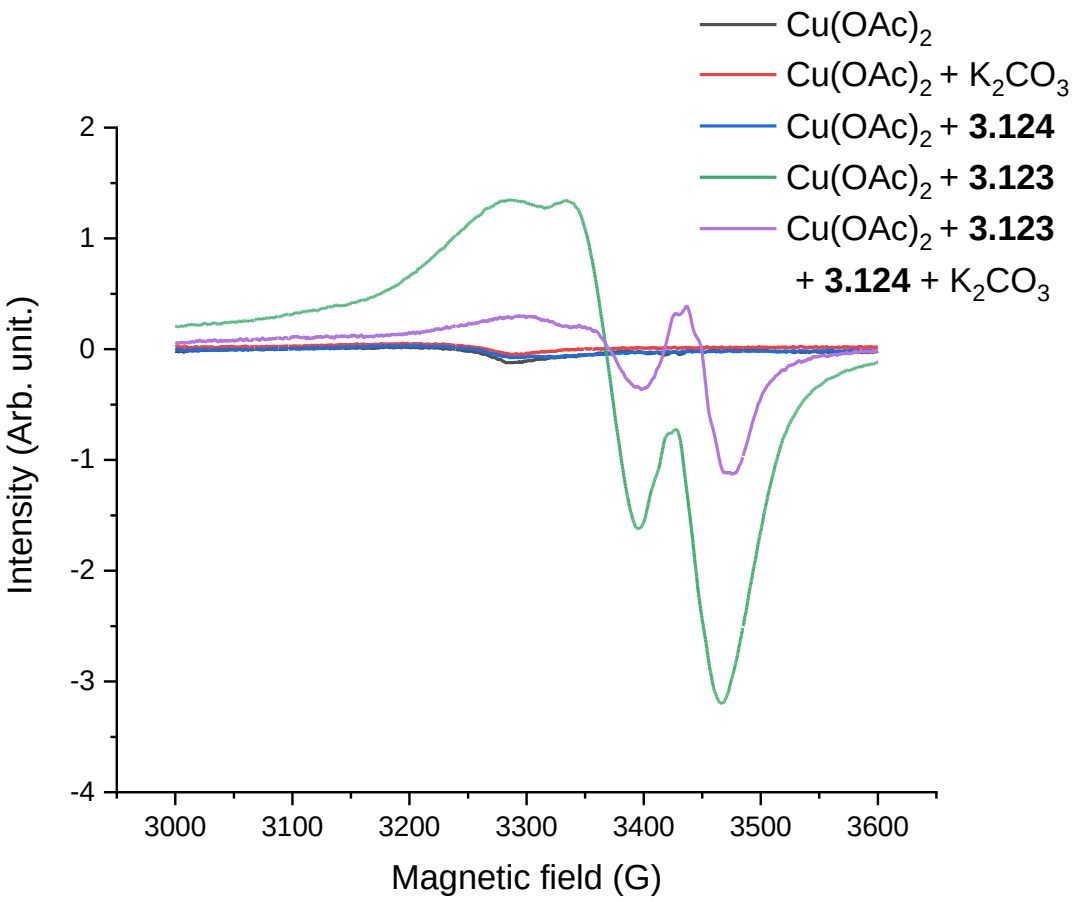


Figure 3.9. EPR spectra showing the de-nucleation of $Cu(OAc)_2$ with the different substrate combinations.

3.5.12 UV-vis experiments

A standard amidination reaction between benzamidine **3.123** and boronic acid **3.124** was performed using general Chan-Lam procedure B on a 1.00 mmol scale of boronic acid **3.124.** Aliquots of the reaction mixture (50 µL) were taken at the following reaction timepoints: 0 min, 5 min, 10 min, 30 min, 60 min, 90 min, 120 min. Each aliquot was diluted with *i*-PrOH (350 µL) for analysis. Analyses were performed on a Shimadzu UV spectrometer UV 1800 equipped with a thermo-controlled cell holder possessing a cell path length of 1 cm and the wavelength range was set to 300 nm - 1100 nm.

3.5.13 Precipitate formation in reactions

Images of precipitate formed during substoichiometric reactions discussed in *section 3.3.1*.

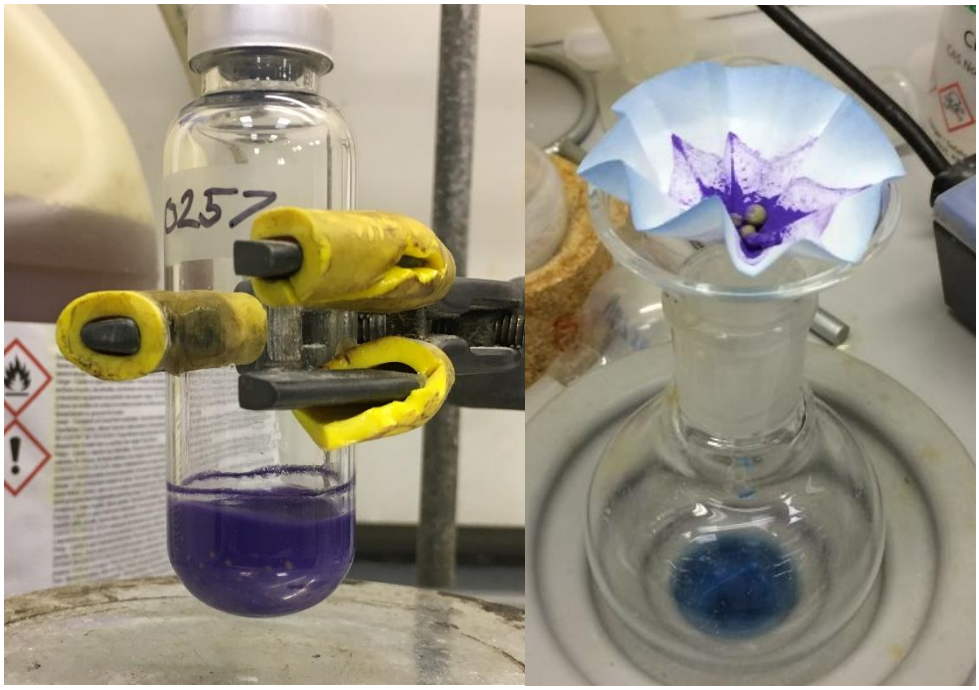
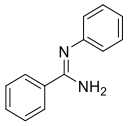


Figure 3.10. Purple precipitate formed during reactions using substoichiometric copper.

3.5.14 Compound characterisations

(*Z*)-*N*'-phenylbenzimidamide (**3.125**)



Chan-Lam procedure A (1.00 mmol scale boronic acid). Purification was carried out using a gradient of 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a white powder (8 mg, 9 %).

Chan-Lam procedure B (1.00 mmol scale boronic acid). Purification was carried out using a gradient of 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a white powder (162 mg, 83 %).

Chan-Lam procedure C (1.00 mmol scale boronic acid). Purification was carried out using a gradient of 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (159 mg, 81 %).

mp 108-112 °C.

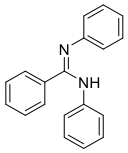
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.87 (d, 2H, *J* = 6.9 Hz, N=C_q-C_q-C*H* × 2), 7.43-7.51 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 7.35-7.39 (m, 2H, N-C_q-CH-C*H* × 2), 7.05-7.09 (m, 1H, N-C_q-CH-CH-C*H*), 7.00-7.02 (m, 2H, N-C_q-C*H* × 2), 4.85 (br s, 2H, NH_2).

$^{13}C\{^{1}H\}$ NMR (125 MHz, $CDCl_3$): δ/ppm: 154.2 (C_q), 151.1(C_q), 136.4 (C_q), 130.5 (*C*-H), 129.7 (2 × *C*-H), 128.4 (2 × *C*-H), 127.5 (2 × *C*-H), 122.3 (*C*-H), 122.0 (2 × *C*-H).

IR (neat): **ν_{max}/cm^{-1}**: 3565 (N-H stretch), 3344 (N-H stretch), 3051 (C-H stretch), 3040 (C-H stretch), 1615 (C=N stretch), 1589 (C=C stretch), 1569 (C=C stretch).

HRMS (ESI): Calc. for $C_{13}H_{13}N_2^+$. Theoretical: 197.1073 Observed: 197.1072.

(*Z*)-*N*,*N'*-diphenylbenzimidamide (**3.126**)



Chan-Lam procedure A (1.00 mmol scale boronic acid). Purification was carried out using a gradient of 0% to 60% Petroleum ether (40-60) in EtOAc with 1% Et_3N modifier afforded a colourless solid (201 mg, 74 %).

Chan-Lam procedure B (1.00 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60):EtOAc with 1% Et_3N modifier, affording a colourless solid (8 mg, 3%).

mp 140-143 °C.

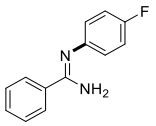
^{1}H NMR (500 MHz, DMSO-d6): δ/ppm: 9.18 (br s, 1H, N*H*), 7.88 (d, 2H, *J* = 7.3 Hz, N=C_q-C_q-C*H* × 2), 7.26-7.33 (m, 7H, 7 × C*H*) 7.96-7.06 (m, 3H, 3 ×C*H*), 6.73-6.77 (m, 1H, N-C_q-CH-CH-C*H*), 6.58 (d, 2H, *J* = 7.1 Hz, N-C_q-C*H* × 2).

$^{13}C\{^{1}H\}$ NMR (125 MHz, DMSO-d6): δ/ppm: 154.2 (C_q), 150.6 (C_q), 141.3 (C_q), 134.7 (C_q), 128.9 (*C*-H), 128.9 (*C*-H), 128.8 (2 × *C*-H), 128.3 (2 × *C*-H), 127.9 (2 × *C*-H), 122.2 (2 × *C*-H), 121.8 (2 × *C*-H), 120.9 (*C*-H), 119.5 (2 × *C*-H).

IR (neat): **ν_{max}/cm^{-1}**: 3290 (N-H stretch), 3053 (C-H stretch), 3025 (C-H stretch), 1627 (C=N stretch), 1587 (C=C stretch), 1578 (C=C stretch), 1528 (C=C stretch).

HRMS (ESI): Calc. for $C_{19}H_{17}N_2^+$. Theoretical: 273.1386 Observed: 273.1387.

(*Z*)-*N'*-(4-fluorophenyl)benzimidamide (**3.149**)



Chan-Lam procedure B (0.20 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a white powder (27 mg, 63%)

Chan-Lam procedure C (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (61 mg, 71%).

mp 113-117 °C.

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.84 (br s, 2H, N=C_q-C_q-C*H* × 2), 7.43-7.50 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 7.03-7.06 (m, 2H, F-C_q-C*H* × 2), 6.92-6.95 (m, 2H, F-C_q-CH-C*H* × 2), 4.84 (br s, 2H, N*H*$_2$).

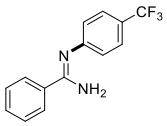
^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 158.6 (d, *J* = 241.4 Hz, *C*-F), 154.9 (C_q), 144.9 (C_q), 135.0 (C_q), 130.2 (*C*-H), 128.1 (2 × *C*-H), 126.3 (2 × *C*-H), 122.3 (d, *J* = 7.9 Hz, 2 × *C*-H), 115.6 (d, *J* = 22.2 Hz, 2 × *C*-H).

^{19}F{^{1}H} NMR (470 MHz, $CDCl_3$): δ/ppm: -121.3.

IR (neat): **ν_{max} /cm^{-1}**: 3465 (N-H stretch), 3335 (N-H stretch), 3056 (C-H stretch), 1613 (C=N stretch), 1600 (C=C stretch), 1569 (C=C stretch), 1498 (C=C stretch).

HRMS (ESI): Calc. for $C_{13}H_{12}N_2F^+$. Theoretical: 215.0979 Observed: 215.0978.

(*Z*)-*N*'-(4-(trifluoromethyl)phenyl)benzimidamide (**3.146**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (68 mg, 64%).

mp 143-145 °C.

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.86 (br s, 2H, N=C_q-C_q-C*H* × 2), 7.61 (d, 2H, *J* = 8.2 Hz, F_3C-C_q- C*H* × 2), 7.44-7.53 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 7.08 (d, 2H, *J* = 8.2, F_3C-C_q-CH-C*H* × 2), 4.86 (br s, 2H, N*H*$_2$).

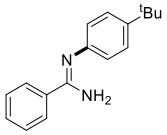
^{13}C{^{1}H} NMR (101 MHz, $CDCl_3$): δ/ppm: 155.1 (C_q), 152.1 (C_q), 134.4 (C_q), 130.6 (*C*-H), 128.2 (2 × *C*-H), 126.4 (2 × *C*-H), 126.2 (q, *J* = 3.5 Hz, 2 × *C*-H), 124.7 (q, *J* = 32.9, C_q), 124.0 (q, *J* = 270.6 Hz, *C*F_3), 121.5 (2 × *C*-H).

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ/ppm: -61.8.

IR (neat): **ν_{max} /cm^{-1}**: 3463 (N-H stretch), 3330 (N-H stretch), 3049 (C-H stretch), 1619 (C=N stretch), 1599 (C=C stretch), 1573 (C=C stretch).

HRMS (ESI): Calc. for $C_{14}H_{11}N_2F_3^+$. Theoretical: 265.0947 Observed: 265.0947.

(*Z*)-*N*'-(4-(tert-butyl)phenyl)benzimidamide (**3.151**)



Chan-Lam procedure B (0.20 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% Petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (41 mg, 82%).

mp 153-155 °C.

^{1}H NMR (500MHz, $CDCl_3$): δ/ppm: 7.86 (d, 2H, *J* = 5.0 Hz, N=C_q-C_q-C*H* × 2), 7.42-7.49 (m, 3H,

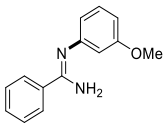
N=C_q-C_q-CH-C*H* × 2-C*H*) 7.35 (d, 2H, *J* = 8.5 Hz, ^{t}Bu-C_q- C*H* × 2), 6.92 (d, 2H, *J* = 8.5 Hz, ^{t}Bu-C_q-CH-C*H* × 2), 4.92 (br s, 2H, NH_2), 1.32 (s, 9H, CH_3 × 3).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ /ppm: 154.9 (C_q), 145.5 (C_q), 135.1 (C_q), 130.1 (*C*-H), 128.1 (2 × *C*-H), 126.4 (2 × *C*-H), 125.8 (2 × *C*-H), 122.3 (C_q), 120.7 (2 × *C*-H), 33.8 (C_q), 30.9 (3 × *C*H_3).

IR (neat): **ν_{max}/cm^{-1}**: 3439 (N-H stretch), 3121 (N-H stretch), 3054 (C-H stretch), 2898 (C-H alkyl stretch), 2863 (C-H alkyl stretch), 1632 (C=N stretch), 1597 (C=C stretch), 1571 (C=C stretch).

HRMS (ESI): Calc. for $C_{17}H_{21}N_2^+$. Theoretical : 253.1699 Observed : 253.1701.

(*Z*)-*N*'-(3-methoxyphenyl)benzimidamide (**3.152**)



Chan-Lam procedure B (0.20 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (28 mg, 62%).

mp 108-112 °C.

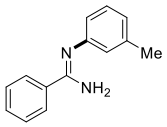
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.86 (d, 2H, *J* = 5.4 Hz, N=C_q-C_q-C*H* × 2), 7.43-7.50 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*) 7.24-7.27 (m, 1H, , N-C_q-CH-C*H*), 6.63 (dd, 1H, *J* = 8.4 Hz, 2.3 Hz, N-C_q-CH-CH-C*H*), 6.56-6.59 (m, 2H, N-C_q-C*H* × 2,) 4.89 (br s, 2H, NH_2), 3.80 (s, 3H, CH_3).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 160.3 (C_q), 154.6 (C_q), 150.4 (C_q), 135.11 (C_q), 130.2 (*C*-H), 129.8 (*C*-H), 128.1 (2 × *C*-H), 126.3 (2 × *C*-H), 113.4 (*C*-H), 108.6 (*C*-H), 106.7 (*C*-H), 54.7 (O-*C*H_3).

IR (neat): **ν_{max}/cm^{-1}**: 3432 (N-H stretch), 3279 (N-H stretch), 3095 (C-H stretch), 2934 (C-H alkyl stretch), 2831 (C-H alkyl stretch), 1636 (C=N stretch), 1587 (C=C stretch), 1573 (C=C stretch).

HRMS (ESI): Calc. for $C_{14}H_{15}N_2O^+$. Theoretical: 227.1178 Observed: 227.1179.

(*Z*)-*N*'-(*m*-tolyl)benzimidamide (**3.153**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (51 mg, 61%).

mp 98-102 °C.

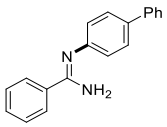
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.86 (d, 2H, *J* = 7.0 Hz, N=C_q-C_q-C*H* × 2), 7.42-7.49 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*) 7.22-7.25 (m, 1H, N-C_q-CH-C*H*), 6.88 (d, 1H, *J* = 7.5 Hz, N-C_q-CH-CH-C*H*), 6.83 (s, 1H, N-C_q-C*H*-C_q-CH_3), 6.79 (d, 1H, *J* = 7.9 Hz, N-C_q-C*H*-CH) 4.84 (br s, 2H, NH_2), 2.35 (s, 3H, CH_3).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm 154.4 (C_q), 148.9 (C_q), 138.8 (C_q), 135.3 (C_q), 130.1 (*C*-H), 128.8 (*C*-H), 128.0 (2 × *C*-H), 126.3 (2 × *C*-H), 123.4 (*C*-H), 121.8 (*C*-H), 118.1 (*C*-H), 21.0 (*C*H_3).

IR (neat): **ν_{max}/cm^{-1}**: 3434 (N-H stretch), 3285 (N-H stretch), 3127 (C-H stretch), 1625 (C=N stretch), 1604 (C=C stretch), 1591 (C=C stretch), 1571 (C=C stretch).

HRMS (ESI): Calc. for $C_{14}H_{15}N_2^+$. Theoretical: 211.1230 Observed: 211.1224.

(*Z*)-*N*'-([1,1'-biphenyl]-4-yl)benzimidamide (**3.154**)



Chan-Lam procedure B (1.00 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (188 mg, 69%).

mp 173-177 °C.

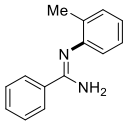
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.88 (d, 2H, *J* = 5.8 Hz, N=C_q-C_q-C*H* × 2), 7.60-7.62 (m, 4H, 4 × C*H*), 7.42-7.51 (m, 5H, 5 × C*H*), 7.31-7.34 (m, 1H, N-C_q-CH-CH-C_q-CH-CH-C*H*), 7.07 (d, 2H, *J* = 8.4 Hz, 2 × C*H*), 4.91 (br s, 2H, NH_2).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 155.0 (C_q), 148.9 (C_q), 141.0 (C_q), 135.9 (C_q), 135.7 (C_q), 130.7 (*C*-H), 128.8 (2 × *C*-H), 128.6 (2 × *C*-H), 128.2 (2 × *C*-H), 126.9 (*C*-H), 126.8 (2 × *C*-H), 126.7 (2 × *C*-H), 122.1 (2 × *C*-H).

IR (neat): **ν_{max} /cm^{-1}**: 3463 (N-H stretch), 3344 (N-H stretch), 3056 (C-H stretch), 3028 (C-H stretch), 1610 (C=N stretch), 1589 (C=C stretch), 1569 (C=C stretch).

HRMS (ESI): Calc. for $C_{19}H_{17}N_2^+$. Theoretical: 273.1386 Observed: 273.1386.

(*Z*)-*N*'-(*o*-tolyl)benzimidamide (**3.155**)



Chan-Lam procedure B (0.20 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (17 mg, 40%).

Chan-Lam procedure C (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-hite powder (40 mg, 48%).

mp 115-118 °C.

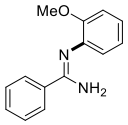
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.89 (br s, 2H, N=C_q-C_q-C*H* × 2), 7.43-7.50 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 7.22 (d, 1H, *J* = 7.4, N-C_q-C_q-C*H*), 7.16-7.19 (m, 1H, N-C_q-CH-C*H*), 6.97-7.00 (m, 1H, N-C_q-C_q-CH-C*H*), 6.88 (d, 1H, *J* = 7.4 Hz, N-C_q-C*H*), 4.71 (br s, 2H, NH_2), 2.22 (s, 3H, CH_3).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 153.0 (C_q), 147.0 (C_q), 134.7 (C_q), 129.8 (*C*-H), 129.7 (*C*-H), 128.7 (C_q), 127.6 (2 × *C*-H), 125.9 (2 × *C*-H), 125.8 (*C*-H), 122.3 (*C*-H), 120.2 (*C*-H), 16.7 (*C*H_3).

IR (neat): **ν_{max}/cm^{-1}**: 3436 (N-H stretch), 3162 (N-H stretch), 3056 (C-H stretch), 2917(C-H alkyl stretch), 1630 (C=N stretch), 1595 (C=C stretch), 1563 (C=N stretch).

HRMS (ESI): Calc. for $C_{14}H_{15}N_2^+$. Theoretical: 211.1230 Observed: 211.1229.

(*Z*)-*N*'-(2-methoxyphenyl)benzimidamide (**3.156**)



Chan-Lam procedure B (0.20 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (39 mg, 42%).

mp 109-113 °C.

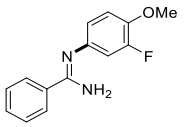
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.86 (d, 2H, *J* = 6.2 Hz, N=C_q-C_q-C*H* × 2), 7.42-7.50 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 7.19 (br s, 1H, N-C_q-C*H*), 7.03-7.08 (m, 1H, N-C_q-C_q-CH-C*H*), 6.94-6.99 (m, 2H, N-C_q-CH-C*H*, N-C_q-C_q-C*H*), 4.72 (br s, 2H, NH_2), 3.83 (s, 3H, O-CH_3).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 154.0 (C_q), 151.3 (C_q), 139.8 (C_q), 136.3(C_q), 130.4 (*C*-H), 128.4 (2 × *C*-H), 127.6 (2 × *C*-H), 123.2 (*C*-H), 122.9 (*C*-H), 121.5 (*C*-H), 112.8 (*C*-H), 55.8 (O-*C*H_3).

IR (neat) ν_{max} (neat): 3434 (N-H stretch), 3294 (N-H stretch), 3054 (C-H stretch), 2826 (C-H alkyl stretch), 1640 (C=N stretch), 1569 (C=C stretch).

HRMS (ESI): $C_{14}H_{15}N_2O^+$. Theoretical: 227.1179 Observed: 227.1171.

(*Z*)-*N*'-(3-fluoro-4-methoxyphenyl)benzimidamide (**3.148**)



Chan-Lam procedure B (0.20 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (67 mg, 68%).

mp 119-123 °C.

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.83 (br s, 2H, N=C_q-C_q-C*H* × 2), 7.43-7.50 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 6.95 (m, 1H, N-C_q-C*H*-CH), 6.78 (d, 1H, *J* = 12.4 Hz, N-C_q-C*H*-CF), 6.70 (d, 1H, *J* = 8.5 Hz N-C_q-CH-C*H*), 4.91 (br s, 2H, NH_2) 3.88 (s, 3H, O-CH_3).

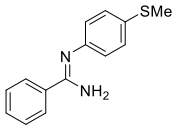
^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 155.7 (C_q), 153.0 (d, *J* = 246.8 Hz, C-F), 143.5 (d, *J* = 10.8, C_q), 143.0 (C_q), 135.5 (C_q), 130.7 (*C*-H), 128.6 (2 × *C*-H), 126.8 (2 × *C*-H), 117.1 (*C*-H), 114.7 (d, *J* = 2.1Hz, *C*-H), 110.2 (d, *J* = 18.9 Hz, *C*-H), 56.7 (O-*C*H_3).

^{19}F{^{1}H} NMR (376 MHz, $CDCl_3$): δ/ppm: -133.5.

IR (neat): **ν_{max}/cm^{-1}**: 3458 (N-H stretch), 3287 (N-H stretch), 3135 (C-H stretch), 3017 (C-H stretch), 2842 (C-H alkyl stretch), 1627 (C=N stretch), 1593 (C=C stretch), 1569 (C=C stretch).

HRMS (ESI): Calc. for $C_{14}H_{14}N_2FO^+$. Theoretical: 245.1085 Observed: 245.1078.

(*Z*)-*N*'-(4-(methylthio)phenyl)benzimidamide (**3.150**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 40% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (69 mg, 71%).

Chan-Lam procedure C (0.04 mmol scale boronic acid). Purification was carried out using a gradient 0% to 40% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (65 mg, 67%).

mp 123-127 °C.

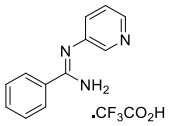
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.82 (d, 2H, *J* = 6.8, N=C_q-C_q-C*H* × 2), 7.42-7.49 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 7.28 (d, 2H, *J* = 8.4 Hz, S-C_q-C*H* × 2), 6.93 (d, 2H, *J* = 8.4 Hz, S-C_q-CH-C*H* × 2), 4.85 (br s, 2H, NH_2) 2.48 (s, 3H, S-CH_3).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ 155.3 (C_q), 147.2 (C_q), 135.5 (C_q), 131.9 (C_q), 130.7 (*C*-H), 129.0 (2 × *C*-H), 128.6 (2 × *C*-H), 126.8 (2 × *C*-H), 122.4 (2 × *C*-H), 17.0 (S-*C*H_3).

IR (neat): **ν_{max}/cm^{-1}**: 3465 (N-H stretch), 3324 (N-H stretch), 2915 (C-H alkyl stretch), 1621 (C=N stretch), 1600 (C=C stretch), 1589 (C=C stretch), 1558 9C=C stretch).

HRMS (ESI): Calc. for $C_{14}H_{15}N_2S^+$. Theoretical: 243.0950 Observed: 243.0945.

(*Z*)-*N*'-(pyridin-3-yl)benzimidamide trifluoroacetate (**3.157**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using RP-HPLC method A, affording a white amorphous solid (47 mg, 38%).

^{1}H NMR (500 MHz, DMSO-d_6): δ/ppm: 8.19 (s, 1H, N=C*H*-CH), 8.13, (s, 1H, N=C*H*-C_q), 7.97 (d, 2H, *J* = 6.9 Hz, N=C_q-C_q-C*H* × 2), 7.45-7.49 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 7.26-7.49 (m, 2H, N-C_q-C*H*-C*H*), 6.59 (br s, 2H, NH_2).

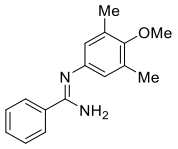
^{13}C{^{1}H} NMR (125 MHz, DMSO-d_6): δ/ppm: 155.2 (C_q), 146.5 (*C*-H), 143.6 (C_q), 142.8 (*C*-H), 135.5 (C_q), 130.3 (*C*-H), 128.7 (*C*-H), 128.0 (2 × *C*-H), 127.1 (2 × *C*-H), 123.9 (*C*-H).

^{19}F{^{1}H} NMR (470 MHz, DMSO-d_6): δ/ppm: -73.5.

IR (neat): **ν_{max}/cm^{-1}**: 3376 (N-H stretch), 3318 (N-H stretch), 3185 (C-H stretch), 1647 (C=N stretch), 1610 (C=N stretch), 1573 (C=C stretch).

HRMS (ESI): Calc. for $C_{12}H_{12}N_3^+$. Theoretical: 198.1026 Observed: 198.1022.

(*Z*)-*N*'-(4-methoxy-3,5-dimethylphenyl)benzimidamide (**3.158**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 30% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder. (85 mg, 83%).

mp 140-143 °C.

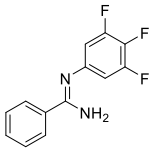
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.84 (d, 2H, *J* = 6.9 Hz, N=C_q-C_q-C*H* × 2), 7.41-7.47 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 6.64 (s, 2H, N-C_q-C*H* × 2), 4.86 (br s, 2H, N*H*$_2$), 3.72 (s, 3H, O-C*H*$_3$), 2.27 (s, 6H, C_q-C*H*$_3$ × 2).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 154.2 (C_q), 152.3 (C_q), 144.7 (C_q), 135.5 (C_q), 131.3 (C_q), 130.0 (*C*-H), 128.0 (2 × *C*-H), 126.2 (2 × *C*-H), 121.0 (2 × *C*-H), 59.3 (O-*C*H$_3$), 15.7 (2 × *C*H$_3$).

IR (neat) ν_{max}/cm^{-1}: 3413 (N-H stretch), 3341 (N-H stretch), 3207 (C-H stretch), 2930 (C-H alkyl stretch), 2921 (C-H alkyl stretch), 2854 (C-H alkyl stretch), 1638 (C=N stretch), 1569 (C=C stretch).

HRMS (ESI): Calc. for $C_{16}H_{19}N_2O^+$. Theoretical: 255.1492 Observed: 255.1487.

(*Z*)-*N*'-(3,4,5-trifluorophenyl)benzimidamide (**3.159**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 30% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (57 mg, 57%).

mp 108-110 °C.

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.79 (br s, 2H, N=C_q-C_q-C*H* × 2), 7.50-7.53 (m, 1H, N=C_q-C_q-CH-CH-C*H*), 7.44-7.47 (m, 2H, N=C_q-C_q-CH-C*H* × 2), 6.58-6.62 (m, 2H, N-C_q-C*H* × 2), 5.99 (br s, 2H, N*H*$_2$).

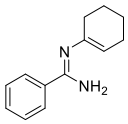
^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 154.5 (C_q), 150.9 (ddd, *J* = 248.4 Hz, 10.5 Hz, 4.9 Hz, 2 × *C*-F), 144.6 (ddd, *J* = 13.4 Hz, 9.1 Hz, 3.6 Hz, *C*-F), 135.9 (t, *J* = 15.8 Hz, C_q), 134.0 (C_q), 130.1 (*C*-H), 127.7 (2 × *C*-H), 125.7 (2 × *C*-H), 104.8 (dd, *J* = 16.7 Hz, 4.9 Hz, 2 × *C*-H).

^{19}F{^{1}H} NMR (470 MHz, $CDCl_3$): δ/ppm: -133.9 (d, *J* = 19.2 Hz), -168.2 (t, *J* = 19.2 Hz).

IR (neat): **ν_{max}/cm^{-1}**: 3426 (N-H stretch), 3277 (N-H stretch), 3116 (C-H stretch), 1636 (C=N stretch) 1574 (C=C stretch), 1517 (C=C stretch).

HRMS (ESI): Calc. for $C_{13}H_{10}F_3N_2^+$. Theoretical: 251.0791 Observed: 251.0782.

(*Z*)-*N*'-(cyclohex-1-en-1-yl)benzimidamide (**3.161**)



Chan-Lam procedure B (1.0 mmol scale boronic acid). Purification was carried out using a gradient 0% to 30% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (83 mg, 41%).

Chan-Lam procedure C (1.00 mmol scale boronic acid). Purification was carried out using a gradient 0% to 30% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a white powder (66 mg, 33%).

mp 103-105 °C.

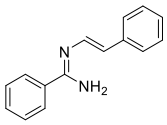
^{1}H NMR (500 MHz,$CDCl_3$): δ/ppm: 7.76 (d, 2H, *J* = 7.1 Hz, N=C_q-C_q-C*H* × 2), 7.37-7.44 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 5.14 (br s, 1H, N-C_q=C*H*), 5.00 (br s, 2H, N*H*$_2$), 2.10-2.13 (m, 4H, N-C_q=CH-C*H*$_2$, N-C_q-C*H*$_2$), 1.72-1.77 (m, 2H, N-C_q-CH_2-C*H*$_2$), 1.60-1.65 (m, 2H, N-C_q-CH_2-CH_2-C*H*$_2$).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 153.2 (C_q), 144.0 (C_q), 135.7 (C_q), 129.7 (*C*-H), 127.9 (2 × *C*-H), 126.1 (2 × *C*-H), 109.5 (*C*-H), 27.5 (*C*H_2), 23.9 (*C*H_2), 22.6 (*C*H_2), 22.0 (*C*H_2).

IR (neat): **ν_{max}/cm^{-1}**: 3458 (N-H stretch), 3292 (N-H stretch), 3025 (C-H stretch), 2928 (C-H alkyl stretch), 2909 (C-H alkyl stretch), 2850 (C-H alkyl stretch), 2831 (C-H alkyl stretch), 1651 (C=N stretch), 1574 (C=C stretch).

HRMS (ESI): Calc. for $C_{13}H_{17}N_2^+$. Theoretical: 201.1386 Observed: 201.1381.

(*Z*)-*N*'-((*E*)-styryl)benzimidamide (**3.160**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a pale-yellow powder (31 mg, 35 %).

mp 117-121°C

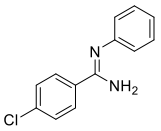
^{1}H NMR (500 MHz, DMSO-d6): δ/ppm: 7.91 (d, 2H, *J* = 6.9 Hz, N=C_q-C_q-C*H* × 2), 7.86 (d, 1H, *J* = 13.4 Hz, N-CH=C*H*-C_q), 7.44-7.46 (m, 2H, N-CH=CH-C_q-C*H* × 2), 7.38-7.43 (m, 3H, N=C_q-C_q-CH-C*H* × 2-C*H*), 7.24-7.27 (m, 2H, N-CH=CH-C_q-CH-C*H* × 2), 7.07-7.13 (m, 3H, N*H*$_2$, CH=CH-C_q-CH-CH-C*H*), 6.32 (d, 1H, *J* = 13.4 Hz, N-C*H*=CH-C_q).

^{13}C{^{1}H} NMR (125 MHz, DMSO-d6): δ/ppm: 155.6 (C_q), 138.9 (C_q), 136.6 (C_q), 136.0 (*C*-H), 130.4 (*C*-H), 128.9 (2 × *C*-H), 128.5 (2 × *C*-H), 127.4 (2 × *C*-H), 126.0 (*C*-H), 125.9 (2 × *C*-H), 121.2 (*C*-H).

IR (neat): **ν_{max}/cm^{-1}**: 3430 (N-H stretch), 3285 (N-H stretch), 3054 (C-H stretch), 3019 (C-H stretch), 1627 (C=N stretch), 1587 (C=C stretch), 1548 (C=C stretch).

HRMS (ESI): Calc. for $C_{15}H_{15}N_2^+$. Theoretical: 223.1230 Observed: 223.1225.

(*Z*)-4-chloro-*N*'-phenylbenzimidamide (**3.164**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 40% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (56 mg, 61%).

Chan-Lam procedure C (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a white powder (63 mg, 69 %).

mp 129-132 °C.

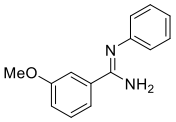
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.78 (d, 2H, *J* = 7.6 Hz, N=C_q-C_q-C*H* × 2), 7.41 (m, 2H, Cl-C_q-C*H* × 2) 7.34-7.37 (m, 2H, N-C_q-CH-C*H* × 2), 7.06-7.09 (m, 1H, N-C_q-CH-CH-C*H*), 6.96 (d, 2H, *J* = 7.4 Hz, N-C_q-C*H* × 2) 4.80 (br s, 2H, N*H*$_2$).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 154.3 (C_q), 149.1 (C_q), 136.8 (C_q), 134.0 (C_q), 129.6 (2 × *C*-H), 128.8 (2 × *C*-H), 128.3 (*C*-H), 123.4 (2 × *C*-H), 121.6 (2 × *C*-H).

IR (neat): **ν_{max}/cm^{-1}**: 3456 (N-H stretch), 3304 (N-H stretch), 3067 (C-H stretch), 3026 (C-H stretch), 1621 (C=N stretch), 1591 (C=C stretch), 1582 (C=C stretch), 1558 (C=C stretch).

HRMS (ESI): $C_{13}H_{12}ClN_2^+$. Theoretical: 231.0684 Observed: 231.0677.

(*Z*)-3-methoxy-*N*'-phenylbenzimidamide (**3.165**)



Using HCl salt of amidine following Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (48 mg, 53%).

mp 104-107 °C.

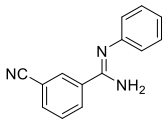
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.43 (br s, 1H, N=C_q-C_q-C*H*-CH), 7.32-7.37 (m, 4H, 4 × C*H*), 7.05-7.09 (m, 1H, N-C_q-CH-CH-C*H*), 6.98-7.04 (m, 3H, 3 × C*H*), 5.06 (br s, 2H, N*H*$_2$), 3.86 (s, 3H, O-C*H*$_3$).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 159.3 (C_q), 154.8 (C_q), 148.2 (C_q), 136.3 (C_q), 129.1 (2 × *C*-H), 129.0 (*C*-H), 122.8 (2 × *C*-H), 121.3 (*C*-H), 118.5 (*C*-H), 116.6 (*C*-H), 111.6 (*C*-H), 55.0 (O-*C*H$_3$).

IR (neat): **ν_{max}/cm^{-1}**: 3465 (N-H stretch), 3335 (N-H stretch), 3056 (C-H stretch), 2922 (C-H alkyl stretch), 2850 (C-H alkyl stretch), 2829 (C-H alkyl stretch), 1621 (C=N stretch), 1600 (C=C stretch), 1589 (C=C stretch), 1574 (C=C stretch).

HRMS (ESI): Calc. for $C_{14}H_{15}N_2O^+$. Theoretical: 227.1179 Observed: 227.1226.

(*Z*)-3-cyano-*N*'-phenylbenzimidamide (**3.169**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a white powder (10 mg, 11%).

mp 105-109 °C.

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 8.19 (br s, 1H, N≡C-C_q-C*H*-C_q), 8.10 (d, 1H, *J* = 7.6 Hz, N=C_q-C_q-C*H*-CH), 7.76 (m, 1H, N≡C-C_q-C*H*-CH) 7.55-7.59 (m, 1H, NC-C_q-CH-C*H*), 7.35-7.39 (m, 2H, N-C_q-CH-C*H* × 2), 7.09-7.12 (m, 1H, N-C_q-CH-CH-C*H*), 6.98 (d, 2H, *J* = 7.4 Hz, , N-C_q-C*H* × 2), 4.98 (br s, 2H, N*H*$_2$).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 152.6 (C_q), 148.0 (C_q), 136.2 (C_q), 133.5 (*C*-H), 130.7 (*C*-H), 130.3 (*C*-H), 129.2 (2 × *C*-H), 129.0 (*C*-H), 123.4 (2 × *C*-H), 121.0 (*C*-H), 117.7(*C*≡N), 112.5 (C_q).

IR (neat): **ν_{max}/cm^{-1}**: 3426 (N-H stretch), 3253 (N-H stretch), 3051 (C-H stretch), 3028 (C-H stretch), 2227 (C≡N stretch), 1640 (C=N stretch), 1619 (C=N stretch), 1589 (C=C stretch), 1578 (C=C stretch).

HRMS (ESI): Calc. for $C_{14}H_{12}N_3^+$. Theoretical: 222.1026 Observed: 222.1020.

(*Z*)-2-fluoro-*N*'-phenylbenzimidamide (**3.166**)

Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (45 mg, 52%).

mp 95-99 °C.

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 8.09 (br s, 1H, N=C_q-C_q-C*H*), 7.40-7.46 (m, 1H, F-C_q-C*H*), 7.33-7.37 (m, 2H, N-C_q-CH-C*H* × 2), 7.22-7.26 (m, 1H, F-C_q-CH-C*H*), 7.09-7.14 (m, 1H, F-C_q-CH-CH-C*H*), 7.05-7.08 (m, 1H, N-C_q-CH-CH-C*H*), 6.98 (d, 2H, *J* = 7.4 Hz, N-C_q-C*H* × 2), 5.16 (br s, 2H, N*H*$_2$).

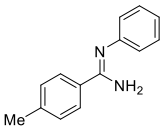
^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 160.1 (d, *J* = 248.1 Hz, *C*-F) 151.4 (C_q), 148.1 (C_q), 131.6 (d, *J* = 9.1 Hz, *C*-H), 130.6 (d, *J* = 1.3 Hz, *C*-H), 129.1 (2 × *C*-H), 124.1 (d, *J* = 3.4 Hz, *C*-H), 122.7 (2 × *C*-H), 122.2 (d, *J* = 10.1 Hz, C_q), 121.2 (*C*-H), 115.6 (d, *J* = 23.6 Hz, *C*-H).

^{19}F{^{1}H} NMR (470 MHz, $CDCl_3$): δ/ppm: -115.2.

IR (neat): **ν_{max}/cm^{-1}**: 3428 (N-H stretch), 3277 (N-H stretch), 3028 (C-H stretch), 1632 (C=N stretch), 1619 (C=N stretch), 1606 (C=C stretch), 1589 (C=C stretch).

HRMS (ESI): Calc. for $C_{13}H_{12}FN_2^+$. Theoretical: 215.0979 Observed: 215.0970.

(*Z*)-4-methyl-*N*'-phenylbenzimidamide (**3.168**)



Chan-Lam procedure B (0.40 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording a colourless solid (26 mg, 31%).

mp 130-133 °C.

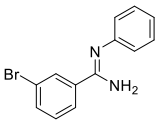
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 7.72 (d, 2H, *J* = 6.6 Hz, N=C_q-C_q-C*H* × 2), 7.32-7.35 (m, 2H, N-C_q-CH-C*H* × 2) 7.24 (m, 2H, H_3C-C_q-C*H* × 2), 7.05-7.08 (m, 1H, N-C_q-CH-CH-C*H*), 6.97-6.99 (d, 2H, *J* = 7.4 Hz, N-C_q-C*H* × 2), 5.01 (br s, 2H, N*H*$_2$).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 155.2 (C_q), 148.2 (C_q), 140.6 (C_q), 131.8 (C_q), 129.0 (2 × *C*-H), 128.8 (*C*-H), 126.4 (2 × *C*-H), 122.8 (2 × *C*-H), 121.4 (2 × *C*-H), 20.9 (*C*H_3).

IR (neat): **ν_{max}/cm^{-1}**: 3441 (N-H stretch), 3292 (N-H stretch), 3067 (C-H stretch), 3051 (C-H stretch), 2915 (C-H alkyl stretch), 1627 (C=N stretch), 1587 (C=C stretch), 1563 (C=C stretch), 1578 (C=C stretch).

HRMS (ESI): Calc. for $C_{14}H_{15}N_2^+$. Theoretical: 211.1230 Observed: 211.1224.

(*Z*)-3-bromo-*N*'-phenylbenzimidamide (**3.167**)



Chan-Lam procedure B using corresponding amidine HCl salt (1.00 mmol scale boronic acid). Purification was carried out using a gradient 0% to 50% petroleum ether (40-60) in EtOAc with 1% Et_3N modifier, affording an off-white powder (133 mg, 49 %).

mp 103-106 °C.

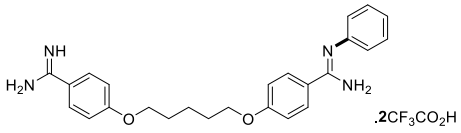
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 8.03 (br s, 1H, Br-C_q-C*H*-C_q), 7.73 (d, 1H, *J* = 7.2 Hz, N=C_q-C_q-C*H*-CH), 7.56 (d, 1H, *J* = 7.9 Hz, Br-C_q-C*H*-CH), 7.34-7.37 (m, 2H, N-C_q-CH-C*H* × 2), 7.27-7.30 (m, 1H, Br-C_q-CH-C*H*), 7.06-7.09 (m, 1H, N-C_q-CH-CH-C*H*), 6.96 (2H, d, *J* = 7.2 Hz, N-C_q-C*H* × 2), 4.88 (2H, br s, N*H*$_2$).

^{13}C{H}NMR (125 MHz, $CDCl_3$): δ/ppm: 153.5 (C_q), 149.2 (C_q), 137.9 (C_q), 133.5 (*C*-H), 130.2 (*C*-H), 130.0 (*C*-H), 129.6 (2 × *C*-H), 125.4, 123.3 (*C*-H), 122.7 (C_q), 121.5 (2 × *C*-H).

IR (neat): **ν_{max}/cm^{-1}**: 3434 (N-H stretch), 3276 (N-H stretch), 3110 (N-H stretch), 3056 (C-H stretch), 1634 (C=N stretch), 1604 (C=C stretch), 1586 (C=C stretch), 1558 (C=C stretch).

HRMS (ESI): Calc. for $C_{13}H_{12}BrN_2^+$. Theoretical: 275.0178 Observed: 275.0177.

(*Z*)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-*N*'-phenylbenzimidamide (**3.174**)



Using Chan-Lam procedure C (0.20 mmol scale boronic acid). Purification was carried out by semi-preparative RP-HPLC using method B, affording a white amorphous solid (31 mg, 24%).

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 11.21 (br s, 1H, CF_3-CO_2H), 9.65 (br s, 1H, CF_3-CO_2H), 9.13 (br s, 2H, NH_2), 8.84 (br s, 3H, 3 × N*H*), 7.89 (d, 2H, *J* = 8.6 Hz, C_q-N=C_q-C_q-C*H* × 2) 7.82 (d, 2H, *J* = 8.8 Hz, HN=C_q-C_q-C*H* × 2) 7.56-7.60 (m, 2H, N-C_q-C*H*× 2), 7.45-7.49 (m, 3H, N-C_q-CH-C*H*× 2-C*H*) 7.21 (d, 2H, *J* = 8.6 Hz, O-C_q-C*H* × 2), 7.17 (d, 2H, *J* = 8.8 Hz, O-C_q-C*H* × 2), 4.12-4.17 (m, 4H, O-CH_2 × 2), 1.81-1.88 (m, 4H, O-CH_2-CH_2 × 2), 1.57-1.65 (m, 2H, O-CH_2-CH_2-CH_2).

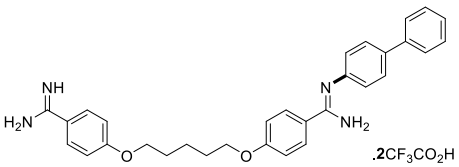
^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 164.7 (C_q), 163.1 (C_q), 163.0 (C_q), 162.3 (C_q), 134.8 (C_q), 130.8 (2 × *C*-H), 130.1 (2 × *C*-H), 129.9 (2 × *C*-H), 128.2 (*C*-H), 125.4 (2 × *C*-H), 120.0 (C_q), 119.4 (C_q), 114.7 (4 × *C*-H), 68.1 (*C*H_2), 68.0 (*C*H_2), 28.2 (2 × *C*H_2), 22.0 (*C*H_2).

^{19}F{^{1}H} NMR (470 MHz, $CDCl_3$): δ/ppm: -73.5.

IR (neat): **ν_{max}/cm^{-1}**:3326 (N-H Stretch), 3105 (CO_2H stretch), 2947 (CO_2H stretch), 1667 (C=O stretch), 1610 (C=N stretch), 1495 (C=C stretch).

HRMS (ESI): Calc. for $C_{25}H_{29}N_4O_2^+$. Theoretical: 417.2285 Observed: 417.2272.

(*Z*)-*N*'-([1,1'-biphenyl]-4-yl)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)benzimidamide (**3.178**)



Using Chan-Lam procedure C (0.20 mmol scale boronic acid). Purification was carried out by semi-preparative RP-HPLC using method B, affording a white amorphous solid (29 mg, 20%).

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 11.27 (br s, 1H, CF_3-CO_2H), 9.70 (br s, 1H, CF_3-CO_2H), 9.13 (s, 2H, NH_2) 8.95 (br s, 1H, N*H*), 8.86 (s, 2H, NH_2) 7.87-7.92 (m, 4H, 4 × C*H*) 7.82 (d, 2H, *J* = 9.0 Hz, HN=C_q-C_q-C*H* × 2), 7.74 (d, 2H, *J* = 7.4 Hz, C_q-C_q-C*H* × 2), 7.51-7.56 (m, 4H, 4 × C*H*), 7.42-7.44 (m, 1H, C_q-C_q-CH-CH-C*H*), 7.23 (d, 2H, *J* = 8.7 Hz, O-C_q-C*H* × 2), 7.18 (d, 2H, *J* = 9.0 Hz, O-C_q-C*H* × 2), 4.13-4.18 (m, 4H, O-CH_2 × 2), 1.82-1.89 (m, 4H, O-CH_2-CH_2 × 2), 1.59-1.65 (m, 2H, O-CH_2-CH_2-CH_2).

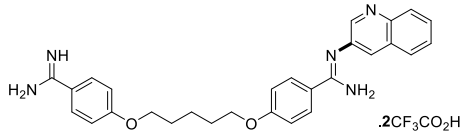
^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 164.7 (C_q), 163.1 (C_q), 163.0 (C_q), 162.4 (C_q), 139.8 (C_q), 139.0 (C_q), 134.2 (C_q), 130.9 (2 × *C*-H), 130.2 (2 × *C*-H), 129.1 (2 × *C*-H), 128.1 (2 × *C*-H), 127.9 (*C*-H), 126.7 (2 × *C*-H), 125.9 (2 × *C*-H), 120.0 (C_q), 119.4 (C_q), 114.8 (4 × *C*-H), 68.1 (*C*H_2), 68.0 (*C*H_2), 28.2 (2 × *C*H_2), 22.1 (*C*H_2).

^{19}F{^{1}H} NMR (470 MHz, $CDCl_3$): δ/ppm: -73.5.

IR (neat): **ν_{max}/cm^{-1}**: 3320 (N-H stretch), 3125 (CO_2H stretch), 2945 (CO_2H stretch), 1671 (C=O stretch), 1610 (C=N stretch), 1493 (C=C stretch).

HRMS (ESI): Calc. for $C_{31}H_{33}N_4O_2^+$. Theoretical: 493.2598 Observed: 493.2590.

(*Z*)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-*N*'-(quinolin-3-yl)benzimidamide (**3.179**)



Using Chan-Lam procedure C (0.20 mmol scale boronic acid). Purification was carried out by semi-preparative RP-HPLC using method B, affording a white amorphous solid (21 mg, 15%).

^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 11.6 (br s, 1H, CF_3-CO_2H), 9.83 (br s, 1H, CF_3-CO_2H), 9.14 (br s, 3H, 3 × N*H*), 8.98 (br s, 1H, N-C_q-C*H*-N), 8.93 (br s, 2H, NH_2), 8.52 (br s, 1H, N-C_q-C*H*-C_q), 8.09-8.14 (m, 2H, N-C_q-CH-C_q-C*H*-C*H*), 7.97 (d, 2H, *J* = 8.5 Hz, C_q-N=C_q-C_q-C*H* × 2), 7.86-7.90 (m, 1H, N-C_q-CH-N-C_q-C*H*), 7.83 (d, 2H, *J* = 8.9 Hz, HN=C_q-C_q-C*H* × 2), 7.71-7.74 (m, 1H, N-C_q-CH-N-C_q-CH-C*H*), 7.25 (d, 2H, *J* = 8.5 Hz, O-C_q-C*H* × 2), 7.17 (d, 2H, *J* = 8.9 Hz, O-C_q-C*H* × 2), 4.13-4.19 (m, 4H, O-CH_2 × 2), 1.82-1.90 (m, 4H, O-CH_2-CH_2 × 2), 1.58-1.66 (m, 2H, O-CH_2-CH_2-CH_2).

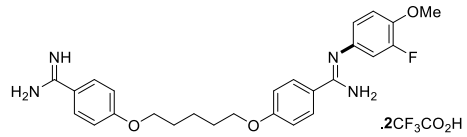
^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 164.6 (C_q), 163.4 (C_q), 163.3 (C_q), 163.0 (C_q), 158.1 (C_q), 157.8 (C_q), 148.2 (*C*-H), 146.6 (C_q), 132.5 (*C*-H), 130.9 (2 × *C*-H), 130.4 (*C*-H), 130.2 (2 × *C*-H), 128.8 (*C*-H), 128.4 (*C*-H), 127.6 (*C*-H), 119.8 (C_q), 119.4 (C_q), 114.8 (2 × *C*-H), 114.7 (2 × *C*-H), 68.1 (*C*H_2), 68.0 (*C*H_2), 28.2 (2 × *C*H_2), 22.1 (*C*H_2).

^{19}F{^{1}H} NMR (470 MHz, $CDCl_3$): δ/ppm: -73.7.

IR (neat): **ν_{max}/cm^{-1}**: 3326 (N-H stretch), 3093 (CO_2H stretch), 2943 (CO_2H stretch), 1667 (C=O stretch), 1608 (C=N stretch), 1493 (C=C stretch).

HRMS (ESI): Calc. for $C_{28}H_{30}N_5O_2^+$. Theoretical: 468.2394 Observed: 468.2385.

(*Z*)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-*N*'-(3-fluoro-4-methoxyphenyl)benzimidamide (**3.181**)



Using Chan-Lam procedure C (0.20 mmol scale boronic acid). Purification was carried out by semi-preparative RP-HPLC using method B, affording a white amorphous solid (27 mg, 20%).

^{1}H NMR (500 MHz, DMSO-*d6*): δ/ppm: 11.14, (br s, 1H, CF_3-CO_2H), 9.62 (br s, 1H, CF_3-CO_2H) 9.13 (br s, 2H, NH_2), 8.90 (br s, 2H, NH_2), 8.75 (br s, 1H, N*H*), 7.87 (d, 2H, *J* = 8.7 Hz, C_q-N=C_q-C_q-C*H* × 2), 7.82 (d, 2H, *J* = 8.9 Hz, HN=C_q-C_q-C*H* × 2), 7.45 (d, 1H, *J* = 10.7 Hz, N-C_q-C*H*-CF), 7.33-7.38 (m, 1H, N-C_q-C*H*-CH), 7.16-7.26 (m, 5H, 5 × C*H*), 4.12-4.16 (m, 4H, O-CH_2 × 2), 3.91 (s, 3H, CH_3), 1.83-1.86 (m, 4H, O-CH_2-CH_2 × 2), 1.57-1.65 (m, 2H, O-CH_2-CH_2-CH_2).

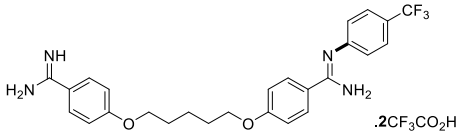
^{13}C{^{1}H} NMR (125 MHz, DMSO-*d6*): δ/ppm: 165.2 (C_q), 163.6 (C_q), 163.5 (C_q), 163.3 (C_q), 151.9 (d , *J* = 248.9 Hz, *C*-F), 147.6 (C_q), 131.2 (2 × *C*-H), 130.6 (2 × *C*-H), 127.6 (C_q), 123.2 (*C*-H), 120.4 (C_q), 119.9 (C_q), 115.3 (2 × *C*-H), 115.2 (2 × *C*-H), 115.0 (*C*-H) 114.8 (d, *J* = 20.0 Hz, *C*-H), 68.6 (CH_2), 68.5 (CH_2), 56.8 (CH_3), 28.7 (CH_2), 22.6 (CH_2).

^{19}F{^{1}H} NMR (470 MHz, DMSO-*d6*): δ/ppm: -73.7, -132.5.

IR (neat): **ν_{max}/cm^{-1}**: 3330 (N-H stretch), 3095 (CO_2H stretch), 2943 (CO_2H stretch), 2874 (CO_2H stretch), 1669 (C=O stretch), 1610 (C=N stretch), 1521 (C=C stretch), 1495 (C=C stretch).

HRMS (ESI): Calc. for $C_{26}H_{30}FN_4O_3^+$. Theoretical: 465.2296 Observed: 465.2289.

(*Z*)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-*N*'-(4-(trifluoromethyl)phenyl)benzimidamide (**3.180**)



Using Chan-Lam procedure C (0.20 mmol scale boronic acid). Purification was carried out by semi-preparative RP-HPLC using method B, affording a white amorphous solid (30 mg, 21%).

^{1}H NMR (500 MHz DMSO-*d6*): δ/ppm: 11.43 (br s, 1H, CF_3-CO_2H), 9.82 (br s, 1H, CF_3-CO_2H), 9.12 (br s, 3H, 3 × N*H*), 8.81 (br s, 2H, NH_2), 7.94 (d, 2H, *J* = 8.5 Hz, F_3C-C_q-CH-C*H* × 2), 7.89 (d, 2H, *J* = 8.7 Hz, C_q-N=C_q-C_q-C*H* × 2), 7.81 (d, 2H, *J* = 8.9 Hz, HN=C_q-C_q-C*H* × 2), 7.67 (d, 2H, *J* = 7.8 Hz, F_3C-C_q-C*H* × 2), 7.22 (d, 2H, *J* = 8.9 Hz, O-C_q-C*H* × 2), 7.16 (d, 2H, *J* = 8.7 Hz, O-C_q-C*H* × 2), 4.12-4.16 (m, 4H, O-CH_2 × 2), 1.81-1.87 (m, 4H, O-CH_2-CH_2 × 2), 1.57-1.63 (m, 2H, O-CH_2-CH_2-CH_2).

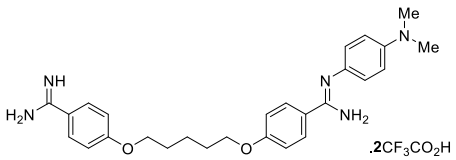
^{13}C{^{1}H} NMR (125 MHz, DMSO-*d6*): δ/ppm: 164.6 (C_q), 163.3 (C_q), 163.0 (C_q), 162.7 (C_q), 139.0 (C_q), 131.1 (2 × C*H*), 130.2 (2 × C*H*), 128.0 (q , *J* = 31.2 Hz, C_q) 127.1 (q, *J* = 3.5 Hz, 2 × C*H*), 126.1 (2 × C*H*), 123.9 (q, *J* = 272.1 Hz, *C*F_3), 119.9 (C_q), 119.4 (C_q), 114.8 (4 × C*H*), 68.0 (CH_2), 68.1 (CH_2), 28.2 (2 × CH_2), 22.0 (CH_2).

^{19}F{^{1}H} NMR (470 MHz, DMSO-*d6*)**:** δ/ppm: -60.9, -73.6.

IR (neat): **ν_{max}/cm^{-1}**: 3328 (N-H stretch), 3105 (CO_2H stretch), 2947 (CO_2H stretch), 2874 (CO_2H stretch), 1667 (C=O stretch), 1608 (C=N stretch), 1495 (C=C stretch).

HRMS (ESI): $C_{26}H_{28}F_3N_4O_2^+$ Theoretical: 485.2159 Observed: 485.2149.

(*Z*)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-N'-(4-(dimethylamino)phenyl)benzimidamide (**3.176**)



Using Chan-Lam procedure C (0.20 mmol scale boronic acid). Purification was carried out by semi-preparative RP-HPLC using method B, affording a pale green amorphous solid (27 mg, 20%).

^{1}H NMR (500 MHz DMSO-*d6*): δ/ppm: 10.90 (br s, 1H, CF_3-CO_2H), 9.45 (br s, 1H, CF_3-CO_2H), 9.12 (br s, 2H, NH_2), 8.78 (br s, 2H, NH_2), 8.49 (br s, 1H, N*H*), 7.86 (d, 2H, *J* = 8.9 Hz, C_q-N=C_q-C_q-C*H* × 2), 7.81 (d, 2H, *J* = 8.9 Hz, HN=C_q-C_q-C*H* × 2), 7.23 (d, 2H, *J* = 9.0 Hz, N-C_q-C*H* × 2), 7.16-7.24 (m, 4H, O-C_q-C*H* × 4), 6.85 (d, 2H, *J* = 8.7 Hz, $(H_3C)_2$N-C_q-C*H* × 2), 4.12-4.16 (m, 4H, O-CH_2 × 2), 2.97 (s, 6H, N(CH_3)$_2$), 1.81-1.87 (m, 4H, O-CH_2-CH_2 × 2), 1.59-1.64 (m, 2H, O-CH_2-CH_2-CH_2).

^{13}C{^{1}H} NMR (125 MHz, DMSO-*d6*): δ/ppm: 164.5 (C_q), 163.1 (C_q), 162.9 (C_q), 162.3 (C_q), 150.0 (C_q), 130.5 (2 × C*H*), 130.2 (2 × C*H*), 126.6 (2 × C*H*),), 122.7 (C_q), 120.1(C_q) 119.4 (C_q), 114.8 (2 × C*H*) 114.7 (2 × C*H*), 112.8 (2 × C*H*), 68.0 (CH_2), 68.1 (CH_2), 39.70 (N(*C*H_3)$_2$, Observed from a HSQC experiment), 28.2 (2 × CH_2), 22.1 (CH_2).

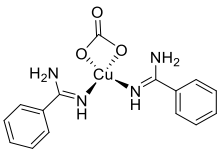
^{19}F{^{1}H} NMR (470 MHz, DMSO-*d6*): δ/ppm: -73.7.

IR (neat): **ν_{max}/cm^{-1}**: 3327 (N-H stretch), 3105 (CO_2H stretch), 2947 (CO_2H stretch),

2874 (CO_2H stretch), 1647 (C=O stretch), 1608 (C=N stretch), 1592 (C=C stretch), 1490 (C=C stretch). **HRMS (ESI)**: Calc. for $C_{27}H_{34}N_5O_2^+$. Theoretical: 460.2707 Observed: 460.2712.

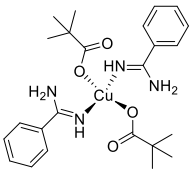
3.5.15 Cu complex experimental procedures

Complex **3.135**



To an oven dried 100 mL round-bottomed flask was added was added benzamidine (1.00 g, 8.32 mmol), $Cu(OAc)_2$ (785 mg, 4.32 mmol) and K_2CO_3 (1.15 g, 8.32 mmol). Isopropanol (18.1 mL) was then added to the solid mixture, sealed under air, and stirred at room temperature for 2 h. The resulting purple suspension was filtered, and the purple precipitate was dried under high vacuum to yield a purple powder-like solid (2.01 g). This purple powder was then recrystallised from MeCN and subjected to a hot filtration to afford deep purple crystals (80 mg, 5%). X-ray crystallography data in *appendix Chapter 3*. X-ray crystallography carried out by Dr Alan Kennedy, University of Strathclyde.

Complex **3.136**



To an oven dried 100 mL round-bottomed flask was added was added benzamidine (100 mg, 0.83 mmol), $Cu(OAc)_2$ (76 mg, 0.42 mmol) and K_2CO_3 (103 mg, 0.83 mmol). Isopropanol (18.1 mL) was then added to the solid mixture, sealed under air, and stirred at room temperature for 2 h. The resulting blue-green suspension was filtered, and the green precipitate was dried under high vacuum to yield a green powder-like solid (107 mg). This green powder was then recrystallised from MeCN and subjected to a hot filtration to afford small pale green solid (2 mg, <1%). X-ray crystallography data in *appendix Chapter 3*. X-ray crystallography carried out by Dr Alan Kennedy, University of Strathclyde.

3.5.16 *Trypanosoma* experiments

Activity against trypanosomes was assessed using a modified alamarBlue protocol.[252] Bloodstream form *Trypanosoma brucei* Lister 427 wild type and derived diamidine-resistant *tbat1*$^{-/-3}$ and B48 lines were cultivated in HMI-11 medium (Gibco) added with 10% FBS (Gibco) at 37 °C in a humidified, 5% CO_2 environment. For the assay, cells from a mid-log phase culture were seeded into 96-well plates at a density of 2×10^4 cells/ml (200 µL/well) and exposed to serial dilutions of test compounds for 48 h. After addition of 20 µL of a 0.5 mM alamarBlue solution (Sigma-Aldrich), cells were further incubated for 24 h to allow reduction of the dye by living trypanosomes. Output was measured using a FLUOstar

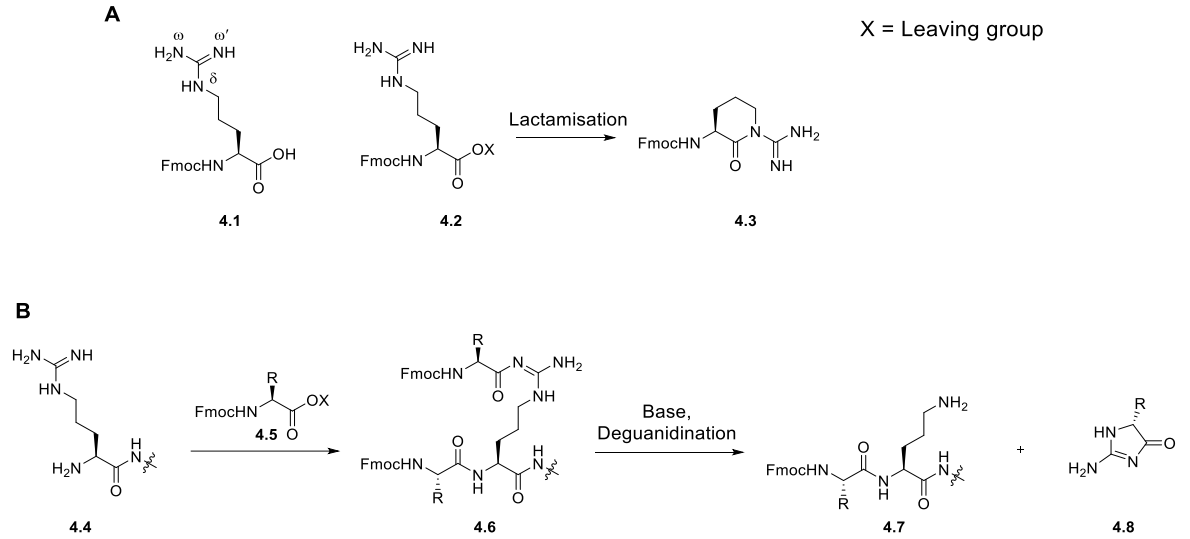
OPTIMA fluorimeter (BMG Labtech), with excitation and emission wavelengths set to 544 nm and 590 nm. EC_{50} values were determined by nonlinear regression analysis using GraphPad Prism 5. Experiments were carried out on three independent occasions. Untreated cells and cells treated with pentamidine were used as negative and positive controls, respectively.

Chapter 4: Synthesis, toxicity and uptake of arginine mimetic CPPs

4.1 Introduction

4.1.1 Protecting groups for the guanidino group for use in SPPS

As discussed in *chapter 2*, Fmoc protected building blocks are the standard reagents for SPPS and arginine is no exception. The major issue for arginine reagents is the guanidino groups potential for cross reactivity. The major side reactions observed with the guanidino group occur upon activation of the carboxylic acid group, when δ-lactam formation and deguanidination processes can occur (Scheme 4.1A & B).[253,254]



Scheme 4.1. (A) Structure of an Fmoc-Arg building block and the δ lactam formation process. (B) Deguanidation of the guandino group from arginine during SPPS.

The δ lactam formation occurs when the activated ester **4.2** reacts with the δ nitrogen of the guanidino group in a 6-*exo*-trig cyclisation resulting in lactam **4.3**. The deguanidation process occurs when first, unprotected guanidine **4.4** is acetylated with activated ester **4.5** in the amide coupling step of SPPS to form **4.6**. The Fmoc deprotection step unveils free amine groups from **4.6** and a 5-*endo*-trig cyclisation occurs at the guanidino carbon. The cyclisation releases cyclic species **4.8** and an ornithine side chain on peptide **4.7**. Both lactamisation and deguanidation result from unprotected nitrogen atoms in the guanidine group, δ for the lactamisation and ω′ for the deguanidation. The protection of the guanidine needs to be sufficient to block both processes whilst still achieving efficient amide couplings required for SPPS.

A fully protected guanidine group, $N^{\delta\omega\omega'}$, would supress both lactamisation and deguanidation processes however, $N^{\delta\omega}$ protected species were reported to supress lactamisation but full protection of the guanidine proved synthetically challenging.[254] The first protecting groups employed by Jaeger then Rink *et al.* were $N^{\delta\omega}$ Boc or Adoc groups, which did not prove successful at supressing the deguanidation (Fig. 4.1).[253,255] $N^{\omega\omega'}$ Boc protected arginine groups like **4.10** were able to supress

deguanidation but subsequently suffered from δ lactamisations.[256] In addition bis-Boc protected arginine species were found to couple less efficiently at the desired N^{α} position.[256,257]

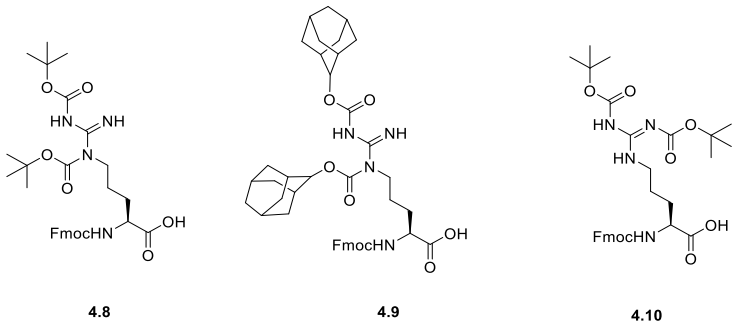


Figure 4.1. $N^{\delta\omega}$ Boc/Adoc and $N^{\omega\omega'}$ Boc protected arginines previously used in peptide synthesis.

Alternatives to the carbamate based protecting groups are the N^{ω}-sulphonyl guandino protecting groups. Tosylate groups were first reported by Fischer *et al.* and subsequent work by Fujino *et al.* led to the synthesis of the Mbs protected arginine which could be used in the synthesis of a tetrapeptide and a 9-mer peptide (Fig. 4.2).[258,259] Furthermore, sulphonyl groups Mts and Mtr were developed by Yajima *et al.* and Masahiko *et al.* Though an improvement on Tosylate and Mbs groups extensive reaction times or heating were required for efficient deprotections, especially with multiple arginine residues (Fig. 4.2).[260,261] A more labile group, Pmc, was first reported by Ramage *et al.* where increasing alkyl substitution of the benzene ring was deemed important for increasing acid lability.[262,263] Subsequent work by Carpino *et al.* demonstrating a smaller 5-membered ring attached to the benzene functionality, Pbf, further increased the acid lability of the sulphonyl group. (Fig. 4.2).[264]

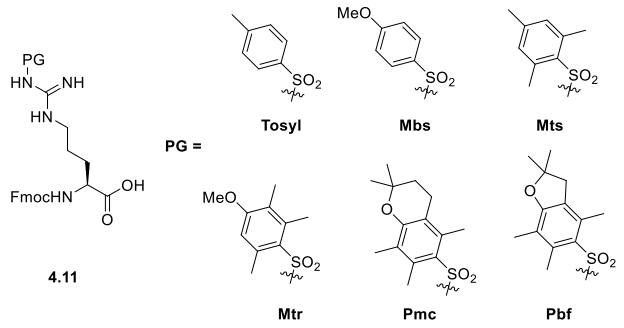
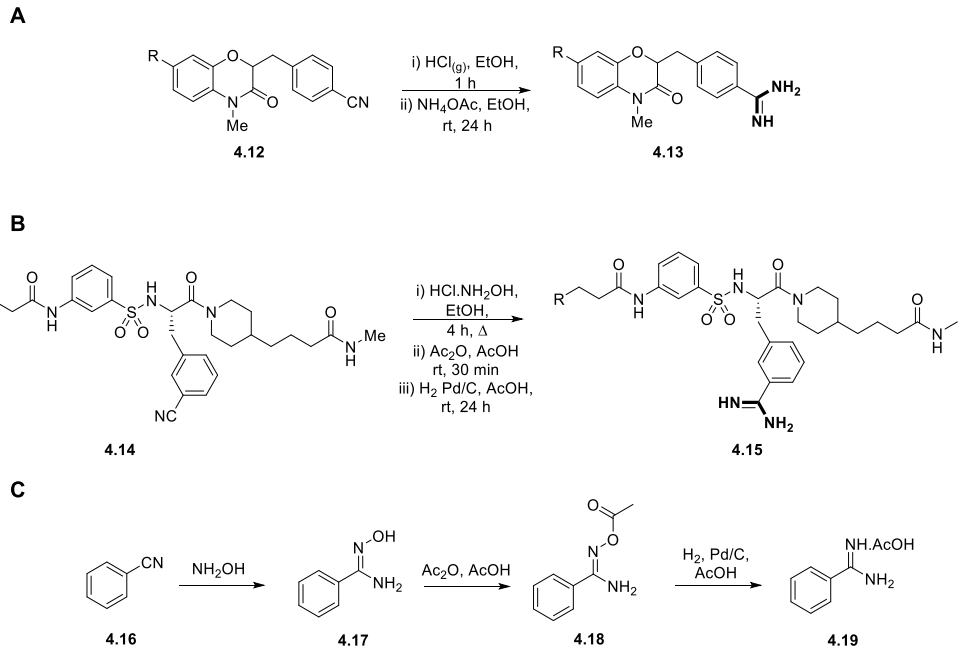


Figure 4.2. Sulphonyl-guanidino protecting groups that have been used in peptide synthesis.

The Pbf group has become the established choice for guanidino protection for three main reasons. Firstly, the ease of synthesis compared to bis protected options. Secondly, the Pbf group's higher acid lability in comparison with other sulphonyl protecting groups. Finally, the Pbf group's ability to supress deguanidation is a final advantage though δ lactam formation remains a persistent issue.[254,257]

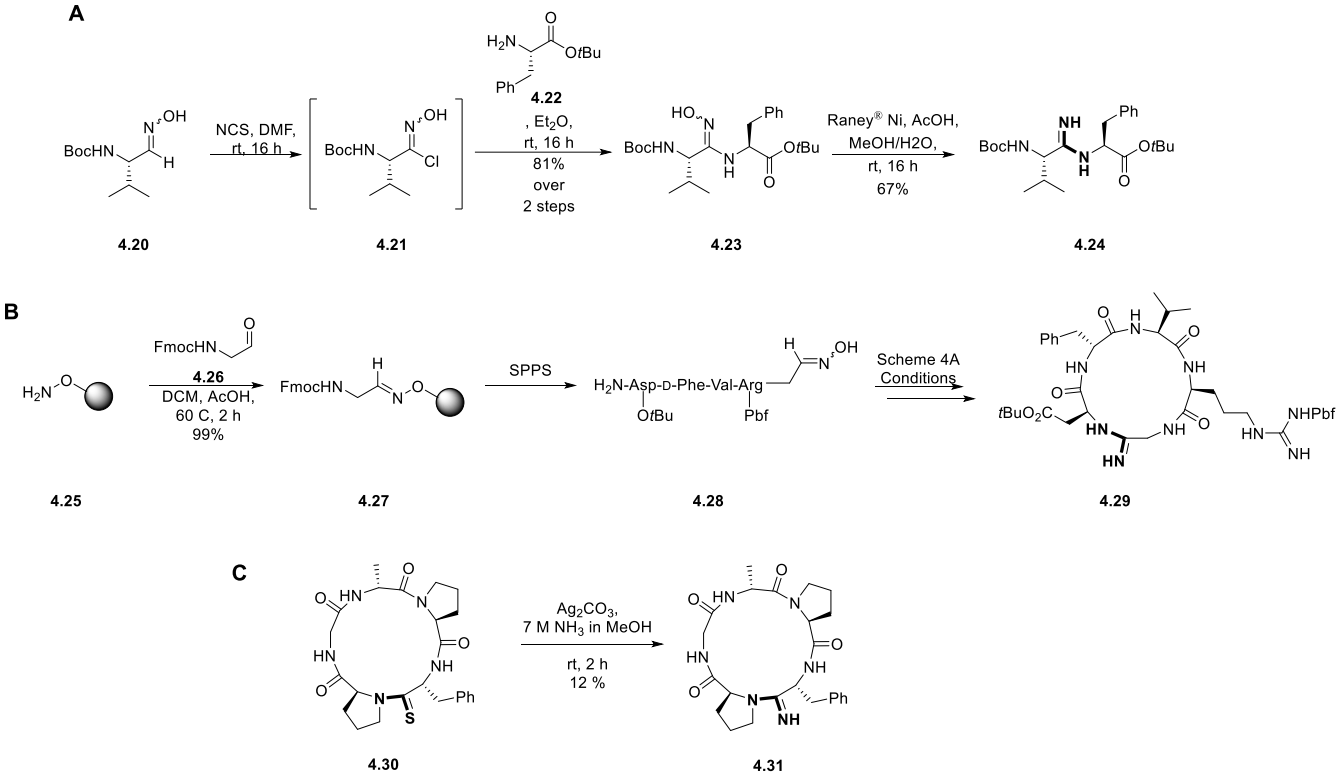
4.1.2 Synthesis of amidines as bioisosteres

As discussed in *chapter 1*, amidines have been employed as bioisosteres of guanidines in serine protease and thrombin inhibitors.[117,188] Pinner reactions are employed for the synthesis of amidines in the case of compound **4.13**, whereas inhibitor **4.15** is synthesised *via* an amidoxime through reduction of an acetylated intermediate (Scheme 4.2A, B & C).[265,266]



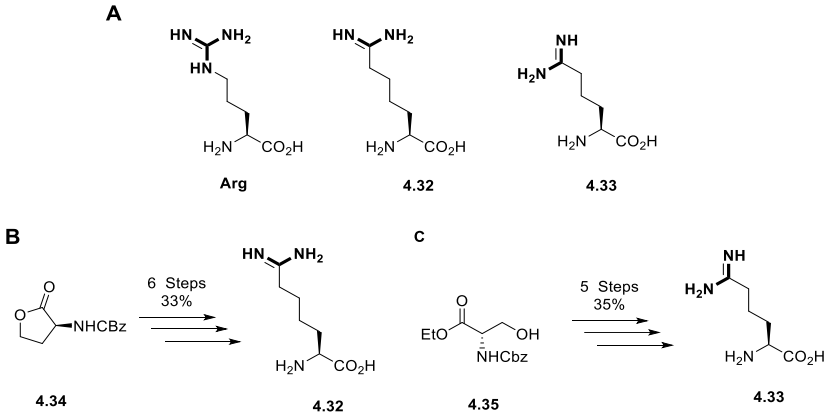
Scheme 4.2. (A) Synthesis of amidine serine protease inhibitor **4.13**. (B) Synthesis of amidine inhibitor **4.15**. (C) Amidine formation from nitrile **4.16** via amidoxime intermediate **4.17**.

Both **4.13** and **4.15** are just two of many amidine structures used as guanidine mimetics in Lipinski compliant drug molecules.[115] Amidine-containing peptides are less common and more synthetically challenging to make, due to Pinner type chemistry not being tolerated by SPPS techniques. The use of amidines as peptide bond bioisoteres was reported by Fujii *et al.*, where like **4.15**, they were formed using an amidoxime intermediate (Scheme 4.3A).[267] Staring from aldoxime **4.20**, chlorination using NCS takes places to form **4.21**. Substitution of the chloride with amino acid **4.22** returned amidoxime **4.23** which was finally reduced with Raney® Ni to yield amidine **4.24**. This amidoxime synthesis was utilised on resin to synthesise amidine containing peptide **4.26** (Scheme 4.3B). More recently work by Yudin *et al.* utilised a silver mediated amidination of a thioamide to synthesise amidine peptide **4.28** (Scheme 4.3C).[268]



Scheme 4.3. (A) Amidine peptide bond synthesis reported by Fujii *et al.*[267] (B) Synthesis of amidine containing peptide **4.29**. (C) Silver mediated amidine peptide synthesis.

The use of amidines as direct comparisons to arginine has not been widely explored in peptide synthesis. The synthesis of the amidine versions of arginine, L-indospicine **4.32**, was first reported by Hagerty *et al*. in 1971 then further developed by Stuehr then Bence *et al.*[269–271] An improved total synthesis of **4.32** and L-norindospicine **4.33** were recently reported by De Voss *et al.*[272] Starting from simple and commercially available precursors, the syntheses reported by De Voss were a total of six steps for **4.32** in an overall yield of 33% and a total of five steps for **4.33** in an overall yield of 35% (Scheme 4.4B & C).



Scheme 4.4. (A) Structures of arginine, L-indospicine and L-norindospicine. (B) Starting material and overall yield in the synthesis of **4.32** by De Voss *et al.*[272] (C) Starting material and overall yield in the synthesis of **4.33**.

Despite the success in the synthesis of amidines **4.32** and **4.33**, neither have been employed as building blocks for SPPS. Incorporating orthogonal base and acid labile protecting groups at the N^{α} and amidine moieties in both **4.32** and **4.33** would prove difficult as in both syntheses the amidine is formed using Pinner chemistry. Both amidine amino acids would require less stringent protection than arginine as they lack the N^{δ} nitrogen, thus eliminating the possibility of δ lactam formation.

Similar to **4.32** and **4.33** are amidine natural products **L-NIL** and **L-NIO**, both containing a methyl group in substitution of the N^{ω} of arginine (Fig. 4.3). Both have been reported as guanidine mimetics, acting as inhibitors of nitric oxide synthases.[273,274] Neither has been synthesised for use in SPPS though a tetrazole analogue of **L-NIL**, **4.36**, has been prepared as a prodrug to inhibit nitric oxide synthase.[275]

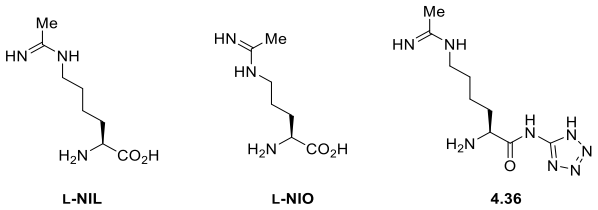


Figure 4.3. Structures of L-NIL, L-NIO and prodrug **4.36**.

Synthesis of a series of protected amidines, analogous to the amidine moiety in **L-NIO** or **L-NIL**, was reported by Madalengoitia *et al*. using isothiocyanate starting materials (Fig. 4.4).[276] Using Pmc isothiocyante **4.37**, a Grignard reaction then affords thioamide **4.38**. Mukaiyamas reagent **4.39** is then used as a thiophile to activate the thioamide to nucleophilic attack of an amine. Subsequent displacement of sulphur by the amine affords amidine **4.40**. Final Pmc deprotection was achieved using TFA:TMSOTf to afford the amidinium product **4.41**. The scope of the synthesis tolerated both alkyl and aromatic Grignard reagents as well as a variety of amines in the amidine formation reaction. The amidination reaction proceeded in moderate to excellent yields ranging from 41% for amidine **4.48** to 91% for amidine **4.42**. Attempts to use bulky Grignard partners such as *tert*-butyl were unsuccessful, as was the use of $NaBH_4$ on **4.37** in an effort to synthesise a thioformamide equivalent of **4.38**. In addition to the Grignard reaction limitations, it was noted that electron deficient amines such as *p*-nitroaniline afforded no amidine product. Furthermore, a benzyl thioamide **4.38** could not react with any amine to form an amidine.

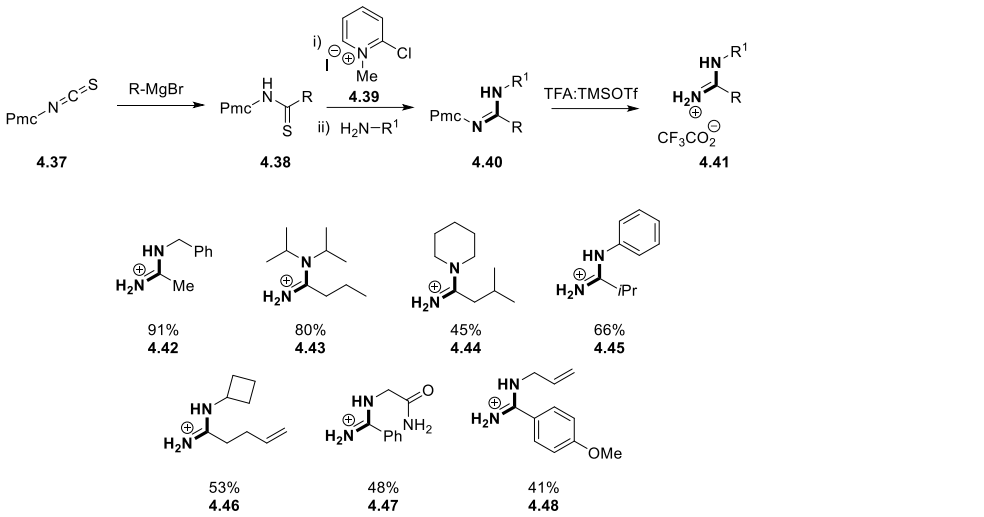
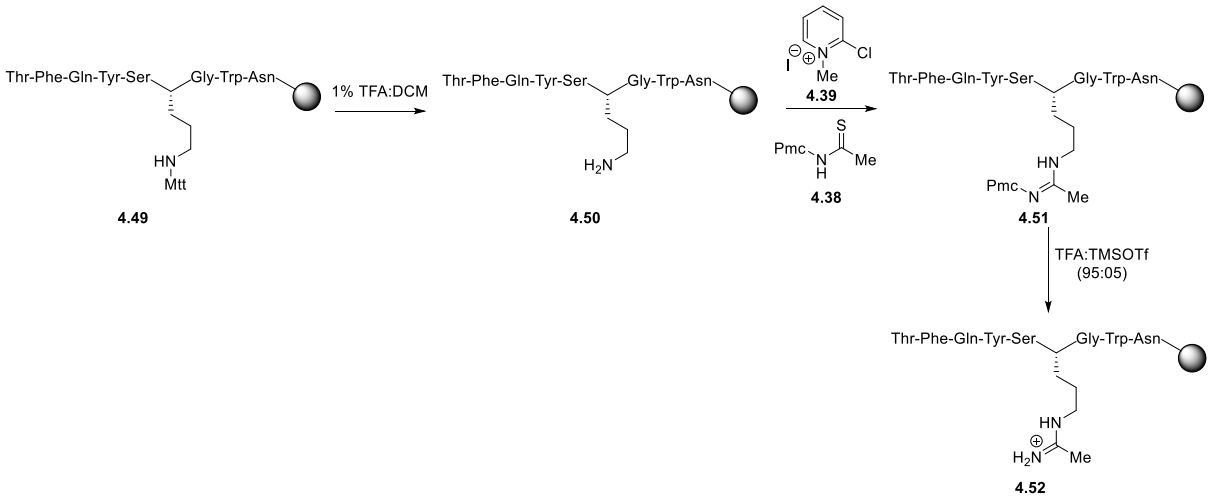


Figure 4.4. Synthesis and scope of amidines synthesised *via* isothiocyanates by Madalengoita *et al*.[276] Yields are for the amidine formation reaction to afford Pmc protected products **4.40**.

Building on this work, Madelengoita *et al.* applied this methodology to peptide synthesis.[277] Using similar methodology Pmc thioamides were synthesised then coupled to an ornithine side chain on resin to afford the amidine product (Scheme 4.5).



Scheme 4.5. On resin amidine peptide synthesis using Pmc protected thioamides.

Peptide **4.49** is prepared using standard SPPS with Mtt protected ornithine present, the Mtt can then be deprotected using 1% TFA in DCM to afford the free amine **4.50**. The amidination reaction can then take place, using the same conditions as Scheme 4.5, that affords the amidine containing peptide **4.51**. Standard TFA cleavage conditions can then afford the free amidine peptide **4.52**. This methodology was then used in a stepwise manner to synthesise peptide **4.53** containing three different amidine side chains (Fig. 4.5).

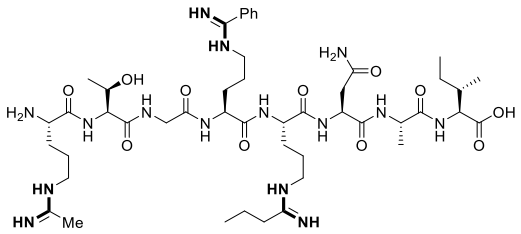


Figure 4.5. Structure of amidine containing peptide **4.53** synthesised by Madelengoita *et al.*

Limitations of this synthesis include using a benzyl thioamide **4.39**, as with solution phase conditions in Fig. 4.4, that does not afford any amidine product. It is also noted that all amidine peptides were observed through crude LCMS analysis and were not isolated.

Amidines offer a potential alternative to arginine residues for application in cell penetrating peptides though synthesis of amidine peptides remains limited to the examples discussed above. In addition to this, no aryl amidine has been employed within a peptide, of any kind, to date.

4.2 Aims of Chapter 4

As discussed in chapter two, many CPPs rely on arginine for their uptake. However as highlighted in *chapter 2* there are specific toxicity issues associated with some arginine rich CPPs like **Pip6** across multiple cell lines. Both low arginine-containing CPPs, **TP-1** and **TP-2**, offer an alternative to **Pip6** based on their lower toxicity and comparable uptake and retainment with cells. These peptides have been previously investigated, though the specific structure activity relationship surrounding the two arginine residues in each peptide is not yet understood. The use of amidines and other guanidine mimetics would provide new information on the functioning of these peptides.

The specific aims of this chapter are to:

(i) Develop protocols for the synthesis of amidine-containing building blocks for SPPS such as **4.54** and **4.55**.

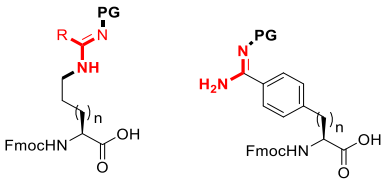


Figure 4.6. Structure of amidine containing amino acids for use in SPPS.

(ii) Synthesise a panel of TP peptide analogues with a range of arginine mimetics to investigate arginine importance with both CPPs.

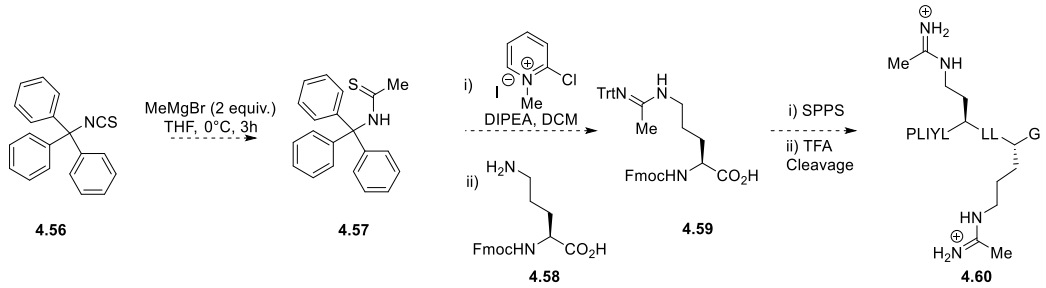
(iii) Identify any differences in toxicity and uptake between the varying TP mimetic CPPs. Highlighting the role, if any, of arginine in the uptake of TP peptides.

(iv) Investigate any subcellular localisation difference between arginine rich CPPs and TP mimetic peptides.

4.3 Results and discussion

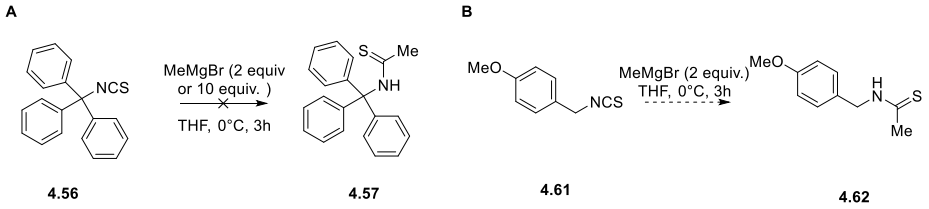
4.3.1 Alkyl amidine isosteres

Based on the literature methods reported, the synthesis of amidines *via* thioamides was the route to be investigated for the synthesis of an amidine N^{α}-Fmoc amino acid. In order to synthesise an amidine like **4.54**, commercially available isothiocyantes were to be used to reduce the amount of steps that were used by Madalengoita *et al*. in the synthesis of Pmc isothiocyanate **4.37**. [276] The trityl group was proposed as useful protecting group for an amidine in SPPS. The trityl group is an acid labile but base stable group often used in SPPS and would protect the amidine from acetylation though not from δ lactamisation. The initial Grignard reaction would use methyl magnesium bromide, chosen to mimic **L-NIL** and offer the least steric bulk around the amidine moiety. Subsequent amidination would be attempted using thioamide **4.57** and Fmoc-Orn **4.58** to provide amidine building block **4.59** that could be used in SPPS (Scheme **4.6**).



Scheme 4.6. Proposed synthesis of an alkyl amidine building block for use in SPPS.

The initial Grignard reaction was unsuccessful using analogous conditions used by Madalengoita *et al.*, with mainly starting material observed according to LCMS (Scheme 4.7A).[276] Increasing the number of equivalents of the Grignard reagent from 2.00 to 10.00, resulted in the formation of a complex mixture of unidentifiable side-products. Based on this a change to a smaller isothiocyanate was proposed, with *p*-methoxy benzyl (PMB) isothiocyante **4.61** to be used (Scheme 4.7B). The PMB group could be used in SPPS as it is base stable and can be removed either by oxidative cleavage or with strongly acidic conditions, similar to cleavage conditions in SPPS.[278,279]



Scheme 4.7. (A) Unsuccessful Grignard reaction with trityl **4.56**. (B) Proposed PMB isothiocyanate Grignard reaction.

The Grignard reaction using **4.61** proceeded with good conversion to thioamide **4.62** however, upon acidic work-up, an increase in side-product, **4.63**, was observed (Fig. 4.7A & B). The side-product had

a *m/z* value of 377, suggesting the formation of **4.63**. The formation of **4.63** could be attributed to product **4.62** reacting with starting-material **4.61** (Fig. 4.7C).

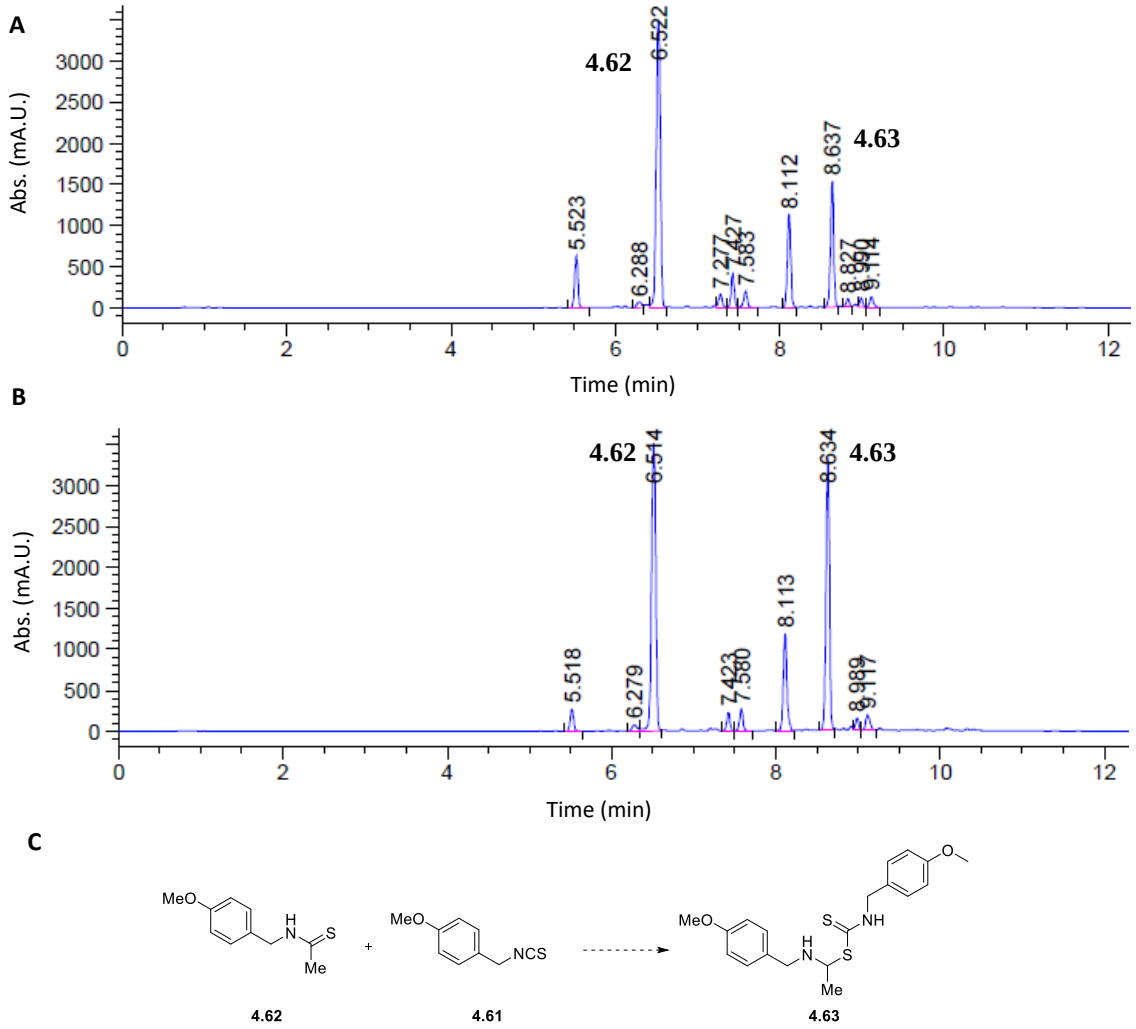
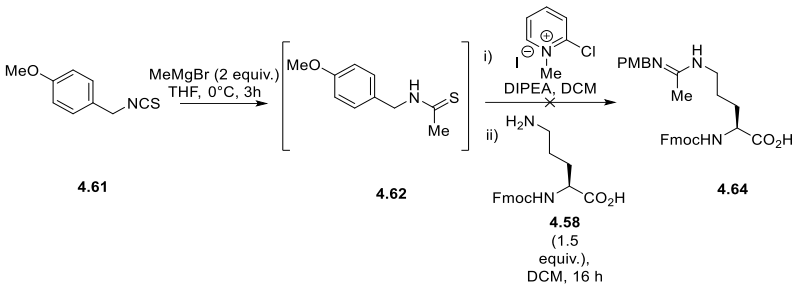


Figure 4.7. (A) Chromatogram of crude Grignard reaction mixture. (B) Chromatogram of reaction after acidic work-up. (C) Possible structure and formation of side product **4.63**.

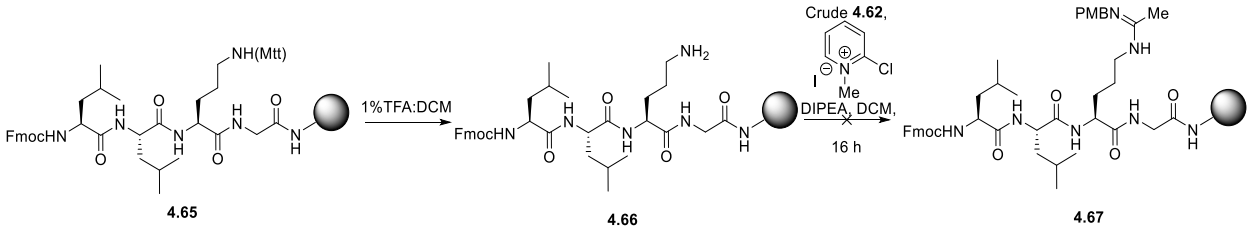
In an attempt to reduce the amounts of **4.61** present upon formation of the product **4.62** the addition of reagents was reversed to add the isothiocyante to the Grignard solution. Unfortunately, this still did not reduce the amount of **4.63** significantly after work-up of the reaction. In addition to the side product formation, product **4.62** proved difficult to isolate from the reaction mixture. Upon purification and evaporation, the resulting oil consisted of an inseparable mixture of **4.62**, **4.63** and another unidentifiable side-product.

Based on the difficulty of isolation of **4.62**, the crude material was carried forward without purification in an attempt to prepare amidine **4.64** in a sufficient yield for use in SPPS (Scheme 4.8). Unfortunately, the crude material did not form any amidine **4.64** upon reaction with ornithine **4.58**, though no evidence of thioamide **4.62** was observed.



Scheme 4.8. Attempted synthesis of PMB protected amidine **4.64**.

As the synthesis of an Fmoc amino acid proved unsuccessful, an on-resin amidine peptide synthesis was attempted, analogous to the on-resin approach by Madalengoita *et al.* (Scheme 4.9).[277] A tetrapeptide sequence containing an Orn(Mtt) group was synthesised, the Mtt group was deprotected using 1% TFA in DCM to yield the free Orn amine **4.66**. The amidination was attempted using crude thioamide **4.62** and Mukaiyama reagent. Upon cleaving some peptide from the resin, only free amine **4.66** was observed by LCMS with no amidine **4.67** present.



Scheme 4.9. Attempted on-resin synthesis of an alkyl amidine containing peptide.

As on-resin amidine synthesis proved unsuccessful using crude thioamide **4.62**, the aryl amidine **4.55** was seen as a viable alternative to the alkyl amidine building block **4.64**.

4.3.2 Aryl amidine isosteres

Aryl amidine groups may offer an additional advantage to the alkyl amidine due to the aromatic system forming π-anion interactions with negatively charged groups at the cell surface (Fig. 4.8). This interaction is similar to the tryptophan-based interactions discussed in *chapter 1 section 1.6.2*. In peptide synthesis the aryl amidine also reduces the possibility of side reactions as δ lactamisation is not possible, so protection from acetylation of the amidine is the main purpose any protecting group.

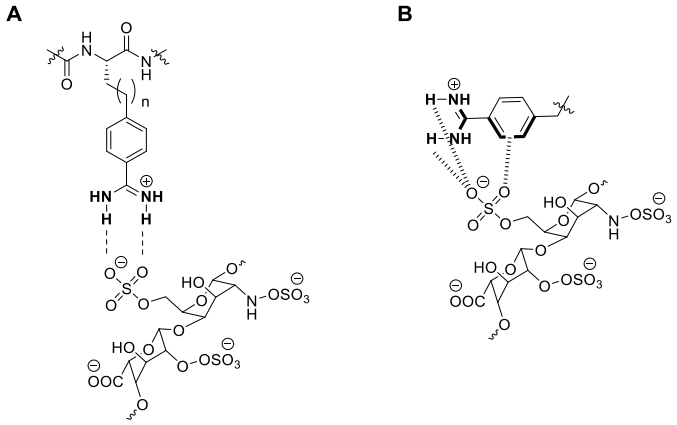
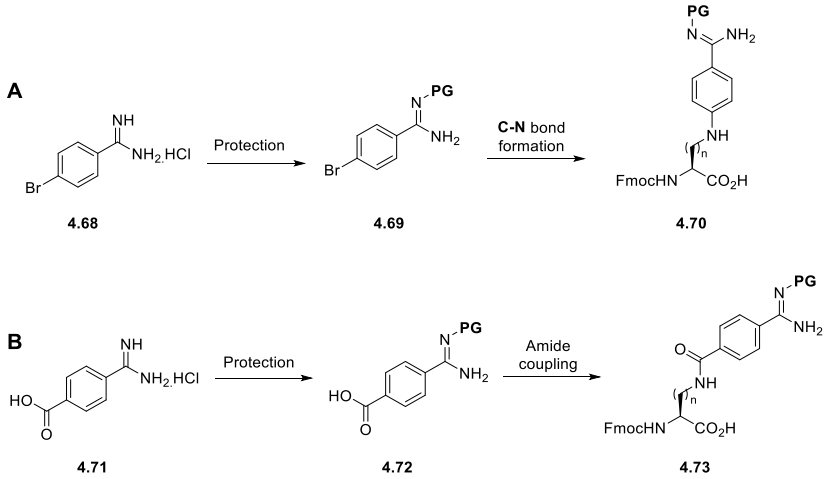


Figure 4.8. (A) Bidentate hydrogen bonding from an aryl amidine to the cell surface. (B) Hydrogen bonding and π-anion interactions between an aryl amidine and the cell surface.

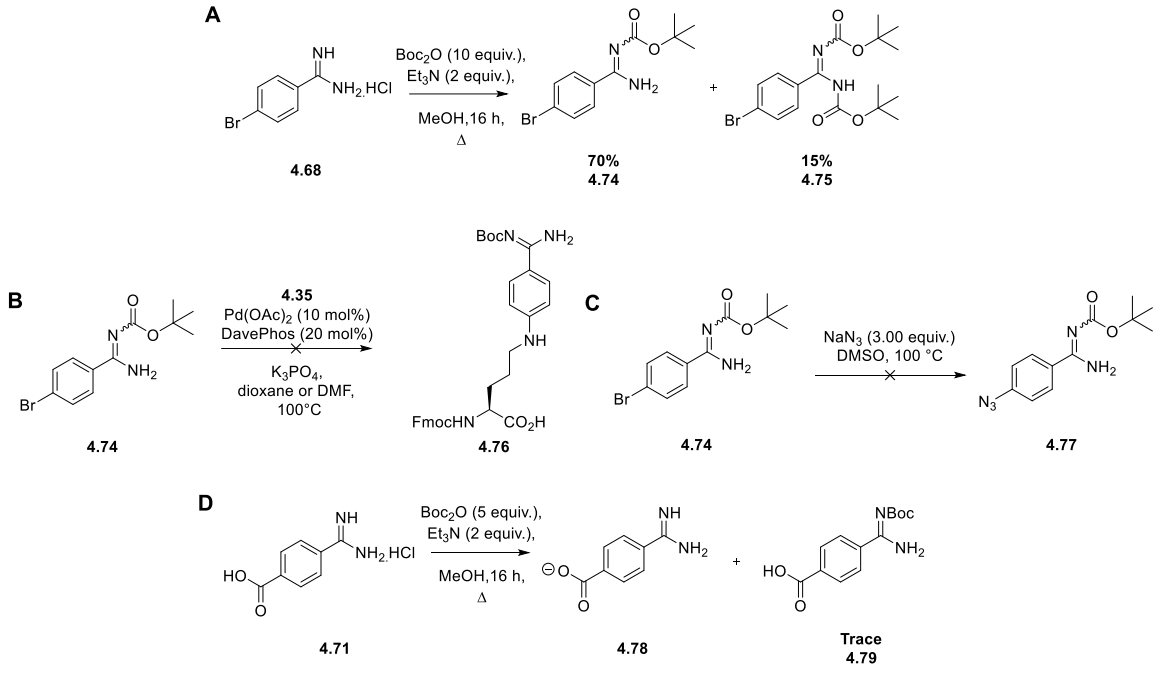
Using commercially available aryl amidines, two proposed synthetic routes were designed (Scheme 4.10 A & B). Starting from aryl bromide **4.68** in route A, the amidine can be protected to form compound **4.69** before an amination reaction to form amino acid **4.70** that can be used in SPPS. Route B starts from an aryl carboxylic acid **4.71** and the amidine can then undergo a protection to form **4.72**. An amide coupling reaction can then be undertaken to form **4.73**, an Fmoc amino acid ready for use in SPPS.



Scheme 4.10. (A) Proposed synthesis of amidine amino acid **4.70** starting from aryl bromide **4.68**. (B) Amidine amino acid synthesis starting from carboxylic acid **4.71**.

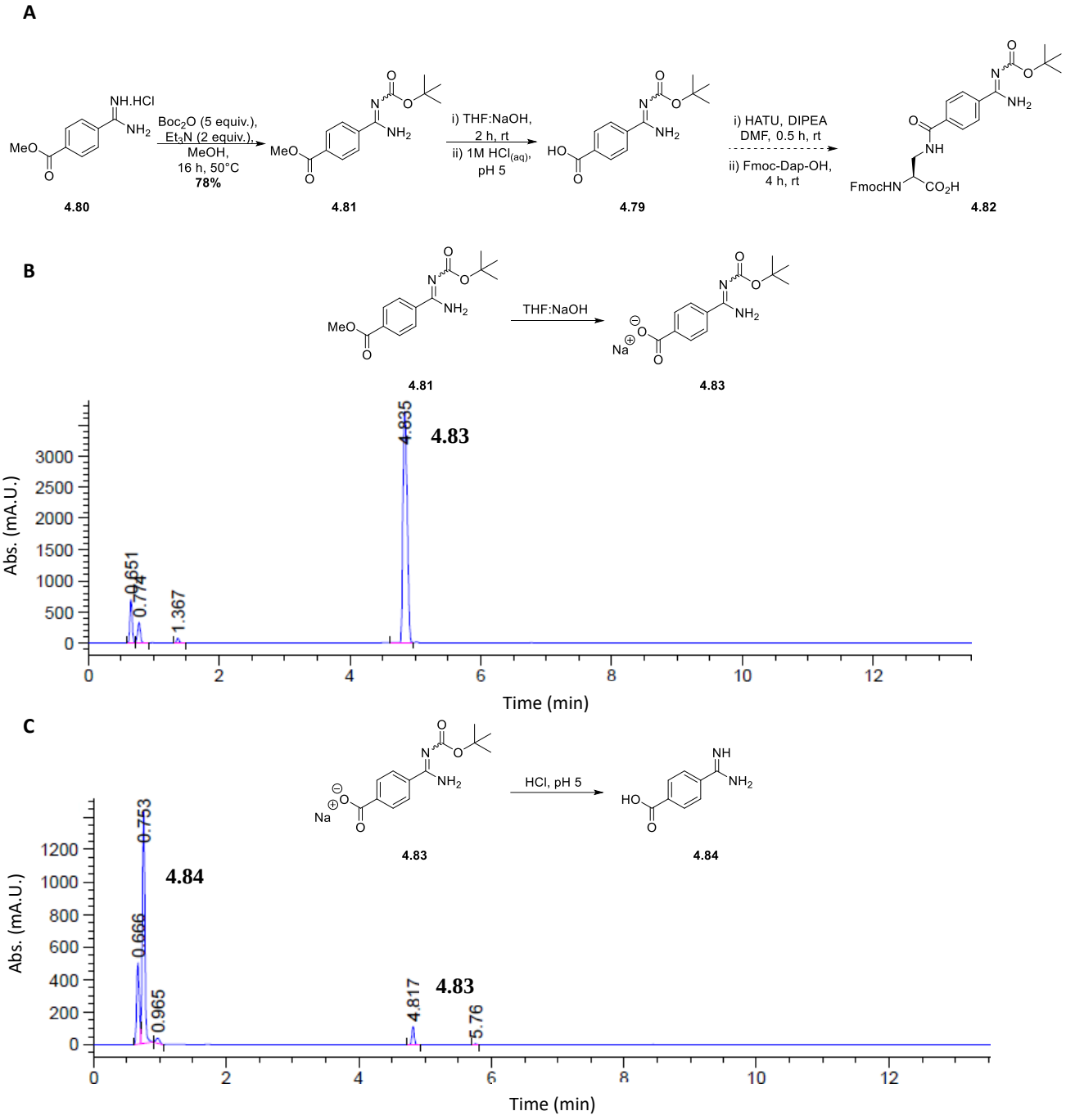
Starting from **4.68** a Boc protection was carried out using Boc_2O to form the mono-protected amidine **4.74**. Despite using a large excess of Boc_2O (10.00 equiv.) the mono-protected amidine remained the major product with a 70% yield and only 15% di-protected **4.75** (Scheme 4.11A). The mono-protected amidine would still offer sufficient protection against acetylation of the amidine and thus was carried forward. Subsequently a Buchwald-Hartwig amination was attempted using Orn **4.58** (Scheme 4.11B). The amination reaction proved unsuccessful in both dioxane and DMF, in both cases returning mainly starting materials. Due to the failure of the amination of **4.74**, an azidation reaction was performed using

NaN_3 which have previously been reported using aryl nitrile species (Scheme 4.11C).[280] The azide formation was also unsuccessful, returning only starting material **4.74** which suggests the amidine group was not sufficiently electron withdrawing to facilitate the S_NAr type reaction. Following route B, the initial Boc protection of **4.71** was unsuccessful as upon addition of base to the reaction a large quantity of precipitate formed, which was most likely due to formation of carboxylate **4.78** which was highly insoluble is most organic solvents. This issue of carboxylate formation meant only trace amounts of Boc protected **4.79** were observed (Scheme 4.11D).



Scheme 4.11. (A) Boc protection reaction of amidine **4.68**. (B) Buchwald-Hartwig reaction with amidine **4.74**. (C) Attempted azide formation using **4.74**. (D) Unsuccessful Boc protection using carboxylic acid **4.71**.

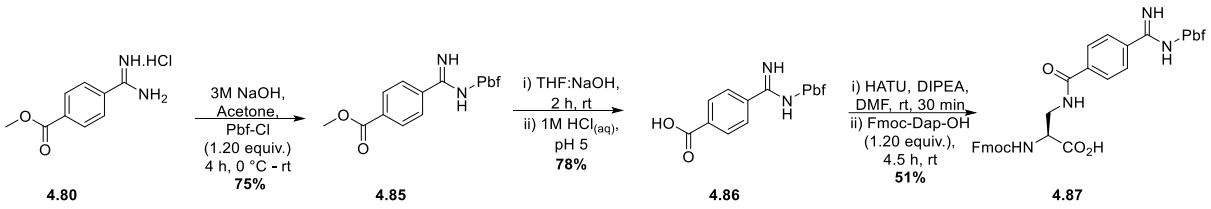
Due to the likely insoluble carboxylate **4.78** forming during Boc protections, ester **4.80** was used in the reaction to prevent this (Scheme 4.12A). The protection reaction returned a yield of 78% for the mono-protected amidine **4.81**. To afford the free acid **4.79** a saponification reaction was performed using NaOH, which appeared to proceed to near full conversion after 2 h reaction time (Scheme 4.12B). Unfortunately, acidification to pH 5 resulted in the deprotection of the amidine group affording mainly the free amidine, free acid product **4.84 (**Scheme 4.12C). These results highlight the enhanced acid lability of the Boc group when protecting an amidine compared to an amine, which would usually require acidification below pH 4 to remove the Boc group.



Scheme 4.12.(A) Synthetic route to Boc protected amidine amino acid **4.82**. (B) Chromatogram of saponification before acidification. (C) Chromatogram of saponification reaction after acidification to pH 5.

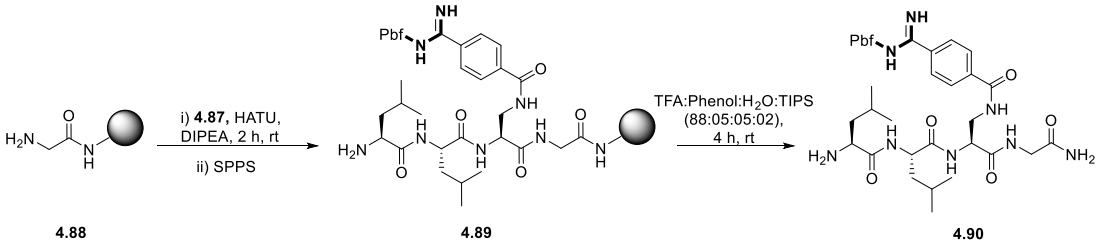
Due to instability of the Boc protecting group to even weakly acidic solutions, the protecting group strategy for the amidine was changed to the common guanidino protecting group, Pbf. The Pbf group was introduced to amidine **4.80** using Pbf chloride and conditions reported by Ramage *et al*. for the analogous Pmc group (Scheme 4.13). Initially, the Pbf protection afforded amidine **4.85** in a yield of 53%. The Pbf reaction also afforded a significant amount of sulphonic acid byproduct and thus a reduction in equivalents of Pbf chloride from 1.60 to 1.20 equivalents was used to afford **4.85** in an improved yield of 75%. The subsequent saponification reaction afforded amidine **4.86** in a yield of 78%

without any deprotection issues upon acidic work-up. Finally, amidine amino acid **4.87** was isolated in a yield of 51% after a HATU mediated amide coupling reaction with Fmoc-Dap-OH, an amine version of the amino acid serine. Utilising Pbf protection proved more reliable than the Boc strategy, affording the Fmoc protected amino acid **4.87** in an overall yield of 30% over 3 steps.



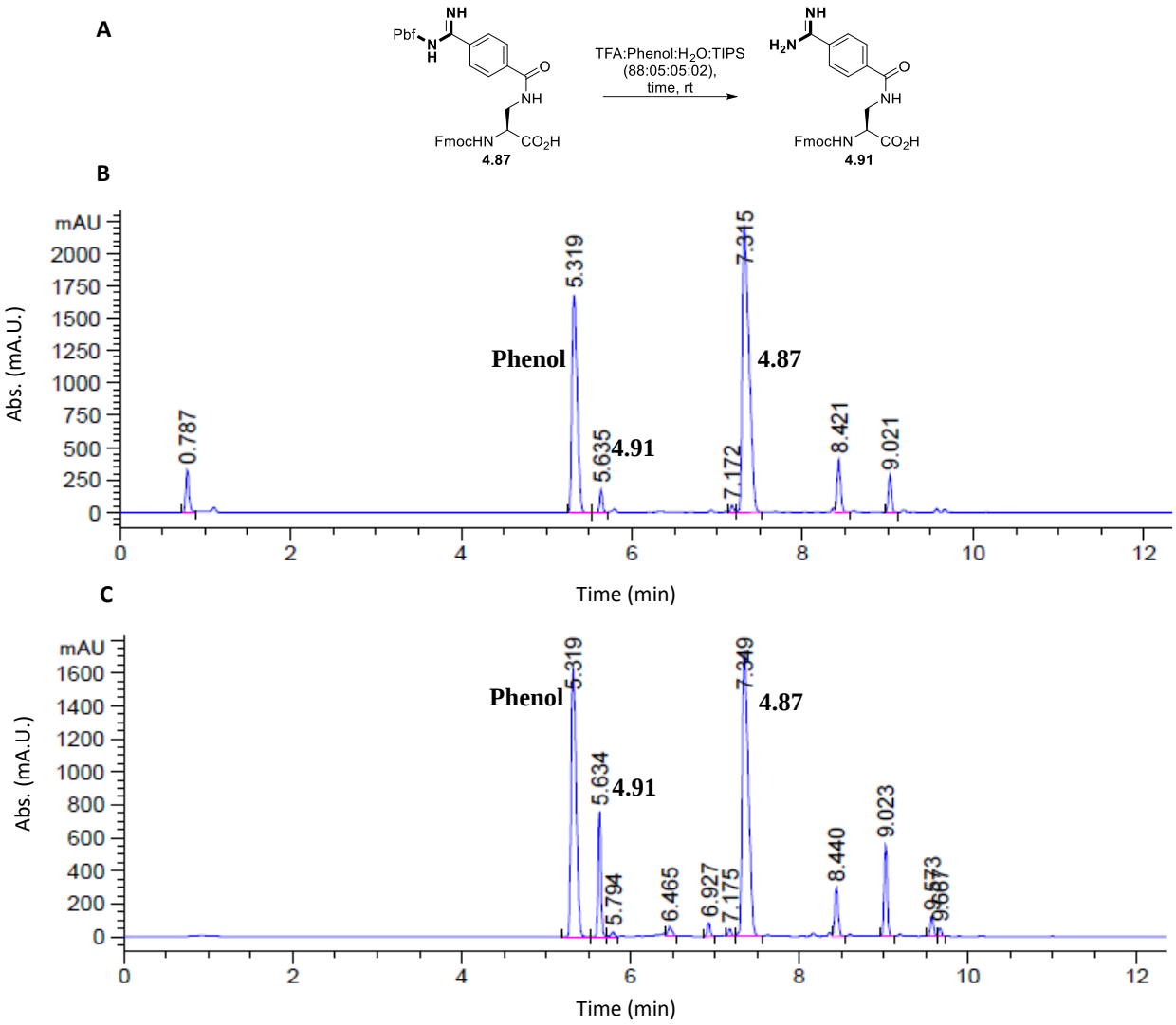
Scheme 4.13. Successful synthetic route to amidine amino acid **4.87**.

With amidine **4.87** in hand, it was envisioned that amidine versions of the TP peptides could be synthesised. To trial **4.87** in SPPS the first four residues in the TP peptide were synthesised, replacing the arginine with **4.87** in the synthesis (Scheme 4.14). The amide coupling reaction with **4.87** was successful using a HATU/DIPEA approach as the tetrapeptide **4.90** was observed through LCMS analysis after TFA cleavage from the resin. Unfortunately, after 4 h TFA incubation the Pbf group remained on the amidine, with none of the free amidine version of **4.90 observed**. This highlights the enhanced acid stability of the Pbf group on an amidine group, which could be attributed to the lower pK_a of the amidinium compared to the guanidinium group formed during deprotection process.



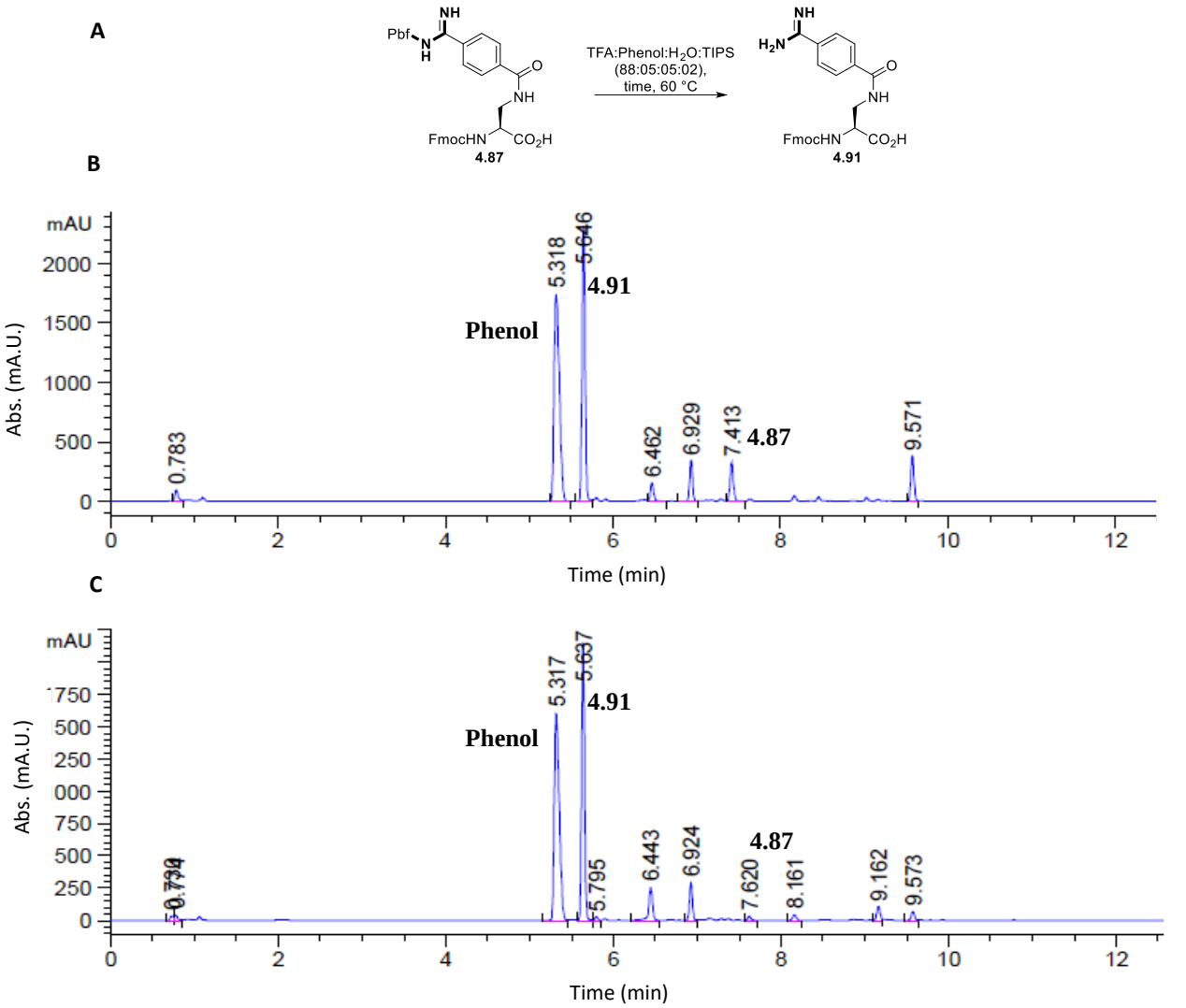
Scheme 4.14. SPPS using amidine building block to afford Pbf protected peptide **4.90**.

To investigate different Pbf deprotection strategies amidine **4.87** was subjected to different conditions that could then be applied to peptide cleavage. Starting with reagent B (TFA:Phenol:H_2O:TIPS, 88:05:05:02), used in all previous peptide synthesis discussed, no conversion to the free amidine **4.91** can be observed after 4 h at rt. Reaction after 24 h at rt using reagent B resulted in small conversion to **4.91**, though the major product remained the Pbf protected starting material (Scheme 4.15A, B & C).



Scheme 4.15.(A) Pbf deprotection reaction using reagent B at rt. (B) Chromatogram of reaction after 4 h. (C) Chromatogram of reaction after 24 h.

In order to overcome the acid stability of the Pbf protected amidine the reaction was attempted again but heated to 60 °C (Scheme 4.16A, B & C). After 4 h at 60 °C a significant improvement in deprotection is observed with near 90% conversion to the free amidine **4.91**, with no **4.87** remaining after 24 h. These temperature change experiments show the Pbf is labile in acid but requires heat in order to overcome the energy barrier to successful deprotection.



Scheme 4.16. (A) Pbf deprotection using reagent B at 60 °C. (B) Chromatogram of 60 °C reaction after 4 h. (C) Chromatogram of 60 °C reaction after 24 h.

With successful deprotection conditions in hand the synthesis of fluorescently labelled TP-1 amidine, **TP-3**, and TP-2 amidine, **TP-4**, could be attempted (Fig. 4.9).

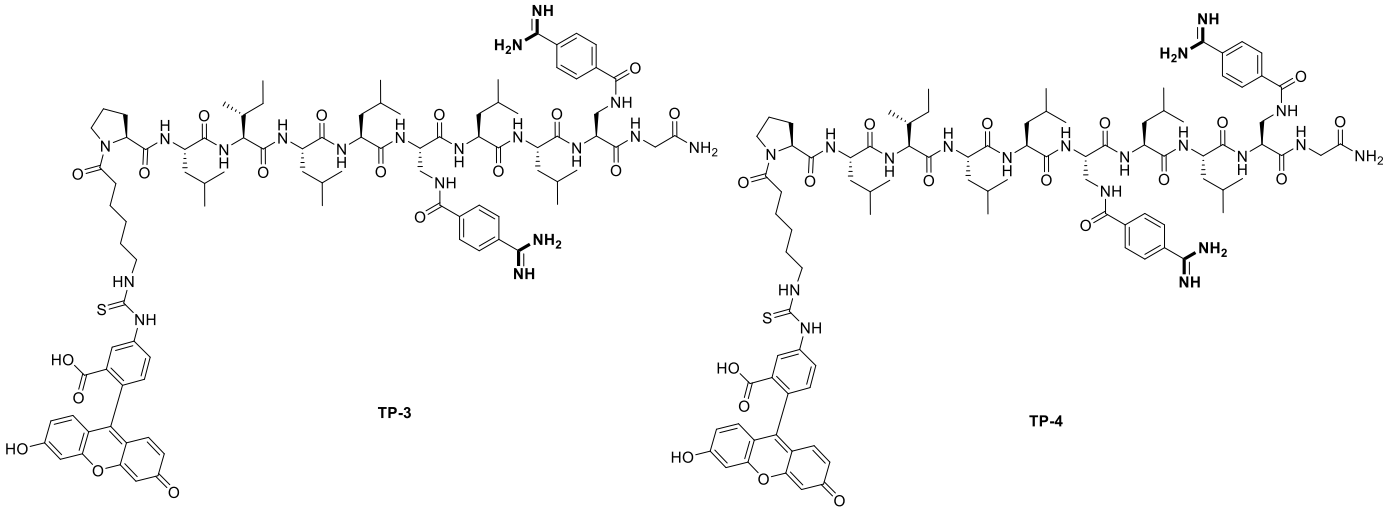


Figure 4.9. Structures of the amidine TP peptides **TP-3** and **TP-4**.

To begin with, the synthesis of each peptide was conducted on an automated synthesiser using HATU/DIPEA as the activating agents and 2 h reaction time for each **4.87** amide coupling step. Under these conditions the isolated yield for both **TP-3** and **TP-4** after full synthesis and cleavage at 60 °C were less than 0.5%. During automated SPPS, the poor coupling efficiency for the **TP-3** resulted in a significant reduction in absorbance, and therefore amount of Fmoc group present after the second amide coupling of **4.87** position 5 (Fig. 4.10, L-6). This indicates that the amide coupling efficiency for **4.87** in position 5 is lower than the previous positions and the subsequent amide couplings all have a similarly low coupling efficiency (Fig. 4.10, L-7 to X-11).

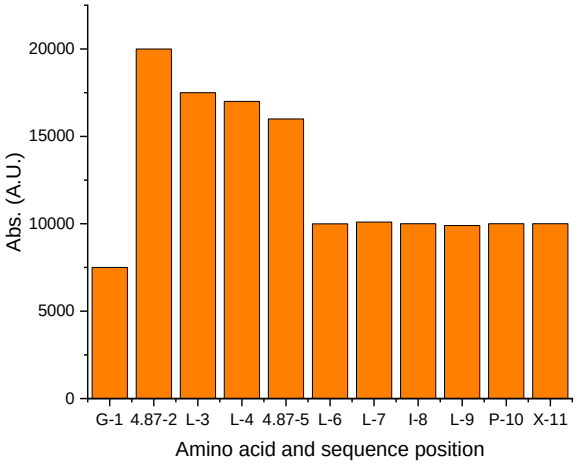
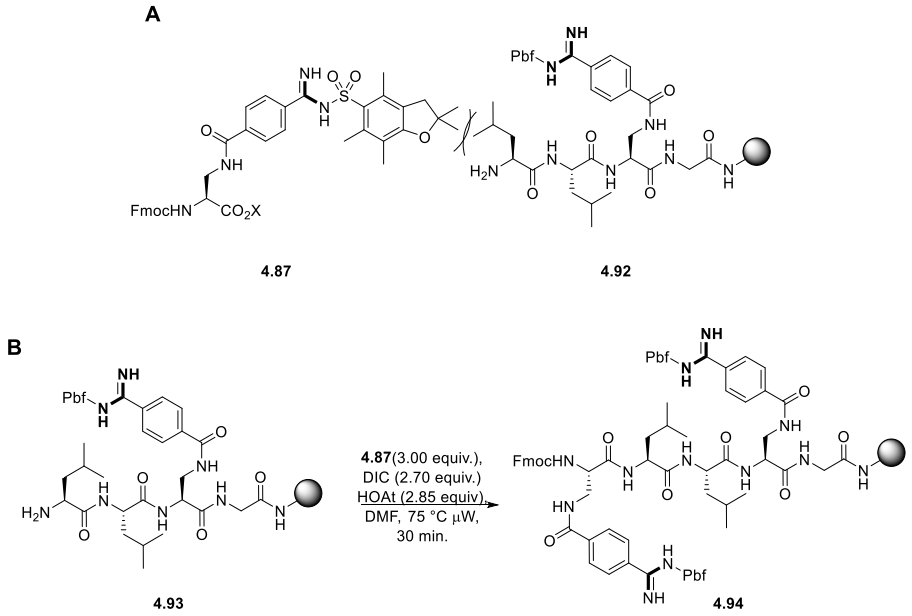


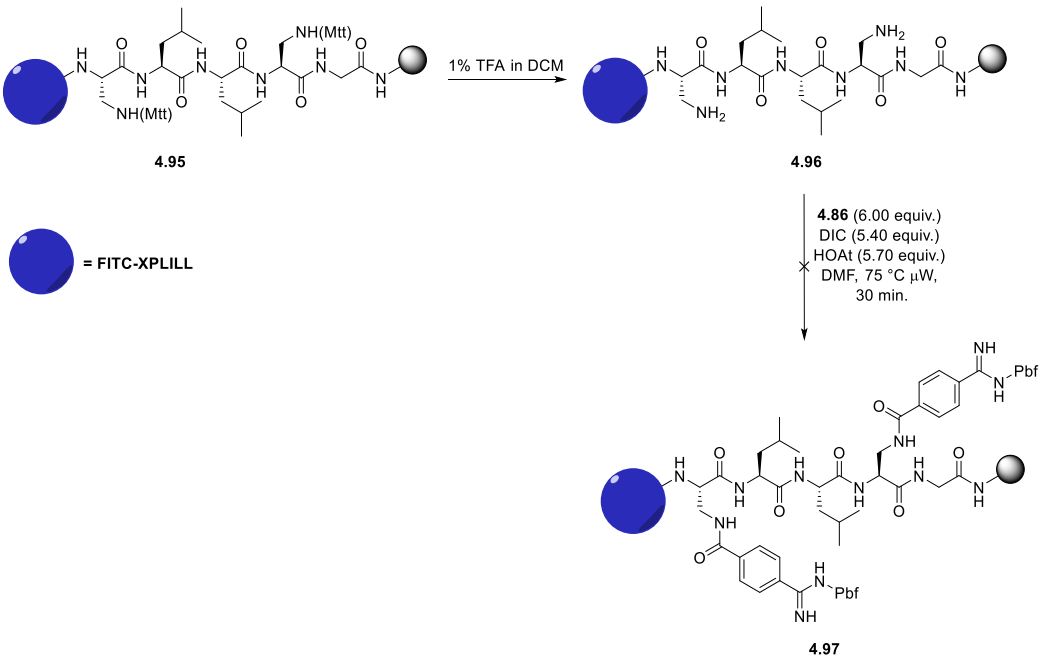
Figure 4.10. Absorbance reading for Fmoc deprotection steps during **TP-3** synthesis. G-1 value of 7500 is due to no Fmoc present on the rink amide resin used.

The lower coupling efficiency of **4.87** in position 5 is possibly attributed to steric clash from the iso-propyl group of Leu in position 4 and the Pbf group of activated **4.87** (Scheme 4.17A). Due to the apparent low coupling efficiency of **4.87** in position 5, the synthesis was attempted again but using microwave assisted amide couplings with DIC/HOAt for each **4.87** coupling step (Scheme 4.17B).



Scheme 4.17. (A) Steric clash between Pbf groups during coupling of **4.87** at position 5 in the **TP-3** and **TP-4** peptide synthesis. (B) DIC and microwave assisted amide coupling used for the synthesis of amidine containing CPPs.

The DIC and microwave conditions were used for both **4.87** amide couplings and for Leu at position 6. Pleasingly the application of these conditions allowed for the synthesis of **TP-4** in an overall isolated yield of 4%. Unfortunately, the same application to **TP-3** did not improve the isolated yield when compared to standard HATU conditions, with similar low Fmoc deprotection absorbances observed from L-6 to X-11. In order to overcome the coupling issues with **TP-4**, a side chain functionalisation was to be attempted making using of Mtt protecting groups (Scheme 4.18).



Scheme 4.18. Attempted side chain functionalisation of Dap(Mtt) groups to form amidine peptide **4.97**.

The full SPPS of a Dap(Mtt) protected peptide, **4.95**, was completed before both Mtt groups were removed under mildly acidic conditions to afford **4.96**. Microwave assisted amide couplings were then attempted on both Dap group side chains using amidine **4.86**. Unfortunately, these couplings proved unsuccessful under the DIC mediated conditions. The Mtt deprotection conditions were confirmed to be successful using a trinitrobenzene sulphonic acid test, meaning the coupling reactions were unsuccessful but not due to insufficient deprotections of Dap side chains.

4.3.3 Peptidomimetics of TP peptides

To further investigate the role of the arginine groups in the TP peptides further, a series of mimetic and non-mimetic interactions were to be investigated (Fig 4.11). Looking at the arginine cell surface interaction proposed for cell uptake, a replacement to lysine was first employed as this would maintain the positive charge but remove the second hydrogen bond formed through the guanidinium species on arginine. Next, the use of urea containing amino acid citrulline, the urea is isoelectronic to the guanidine group and has the possibility of forming two hydrogen bonds through the NH's of the urea moiety. Citrulline would not carry a positive charge at physiological pH so would aid in determining the importance of a net positive charge in cell uptake of the TP peptides. The use of glutamic acid groups would provide evidence about negatively charged TP CPPs and should provide a reduction in cell uptake if positive charge plays any role in cell uptake. The use D-Arg would maintain the same cell surface interaction profile as the standard TP CPPs but would provide some information on the stability of the arginine groups in both peptides and its effect on cell uptake. Finally, the use of **TP-4** would determine if an aryl amidine is a suitable arginine mimetic, having both a positive charge and bidentate hydrogen bonding capability, to maintain or improve cell uptake for the peptide.

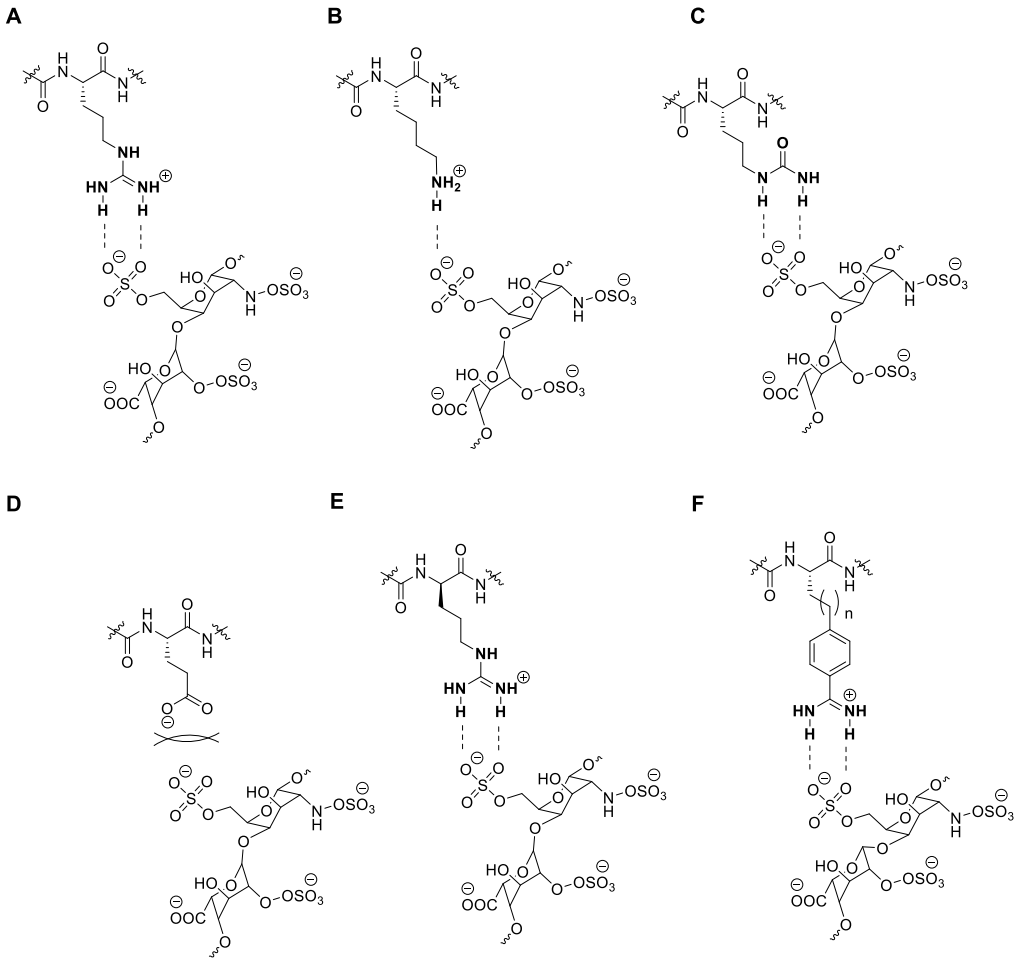


Figure 4.11. (A) Arginine cell surface interaction. (B) Lysine cell surface interaction. (C) Citrulline cell surface interaction. (D) Glutamic acid cell surface interaction. (E) D-Arginine cell surface interaction. (F) Aryl amidine cell surface interaction.

To investigate the interactions outlined in Fig. 11 both arginine residues in the TP peptides were changed to the appropriate mimetic. Using standard SPPS the relevant **TP-1** analogues were synthesised with fluorescein labels at the *N*-terminus in order to quantify uptake by flow cytometry (Fig. 4.12). In total five **TP-1** based CPPs were synthesised. These are the original arginine peptide **TP-1**, the lysine analogue **TP-5**; the Citrulline analogue **TP-6**; the D-Arg analogue **TP-7** and the glutamic acid analogue **TP-8**.

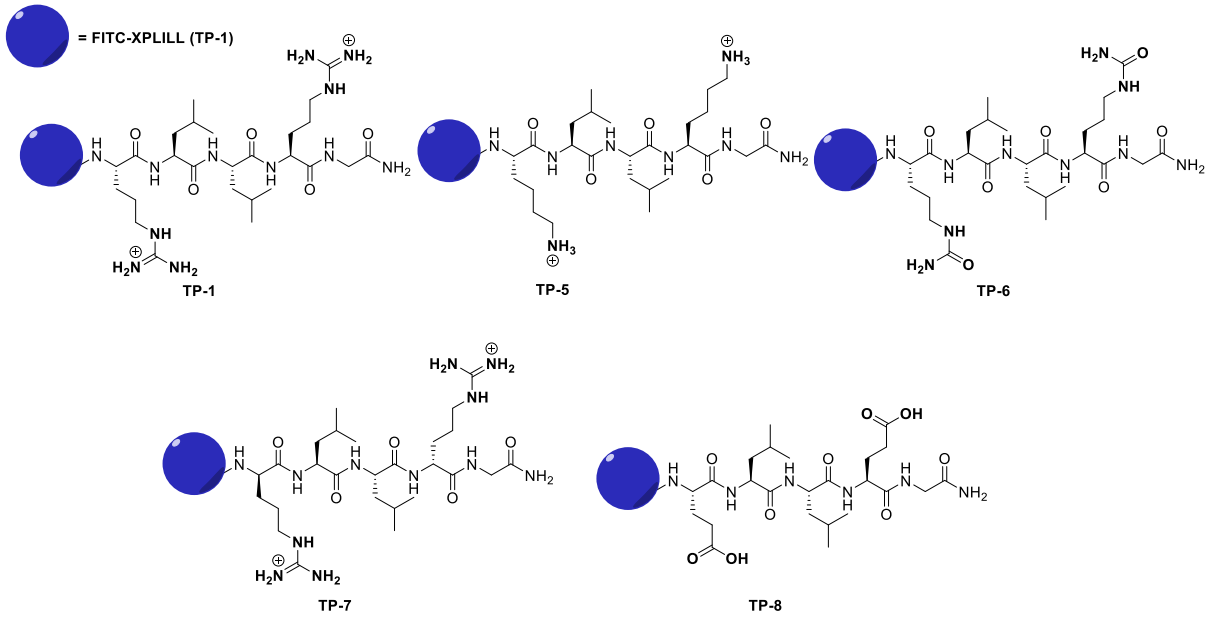


Figure 4.12. Structures of **TP-1** and the analogues synthesised by SPPS.

The same changes to structures were applied to the **TP-2** CPP but with the addition of the amidine containing CPP **TP-4**. All fluorescein labelled **TP-2** analogues were successfully synthesised by SPPS to afford six different peptides for biological analysis: **TP-2**, **TP-9**, **TP-10**, **TP-11**, **TP-12** and **TP-4** (Fig. 4.13).

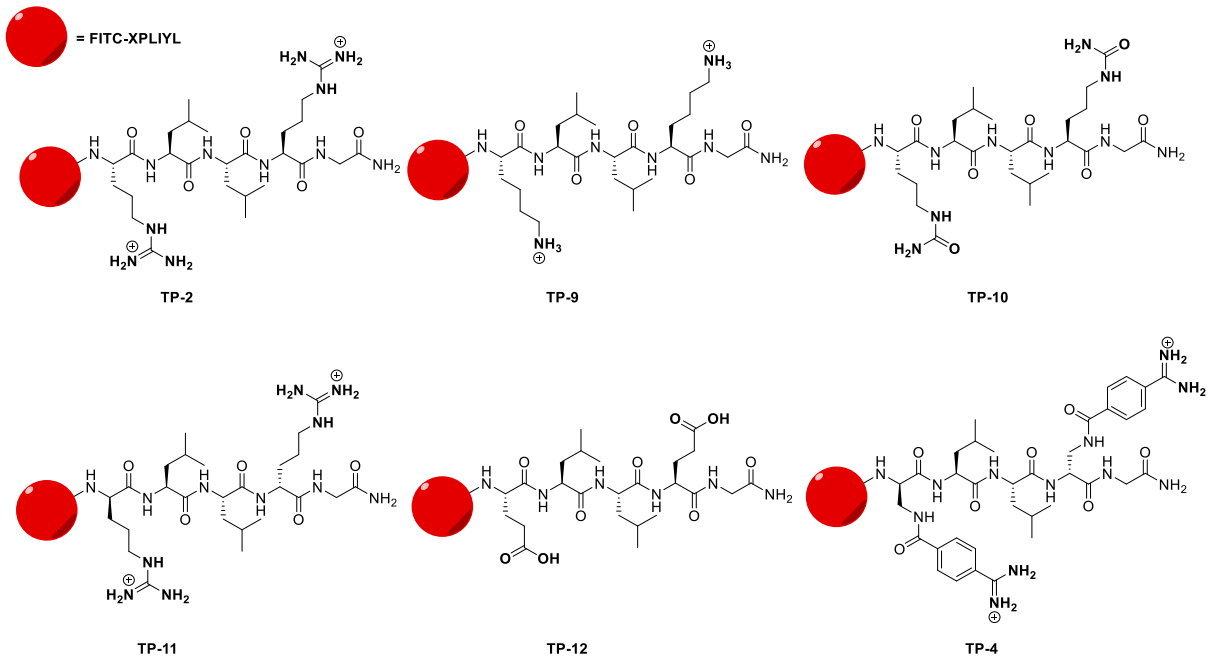


Figure 4.13. Structures of **TP-2** and analogues synthesised by SPPS.

4.3.4 Viability of peptidomimetics

Viability of HeLa cells was first analysed with the **TP-1** analogues using an alamarBlue assay, as discussed in chapter 2 section 2.3.2. The changes to the arginine groups in **TP-1** appeared to show no reduction in HeLa cell viability. The changes to lysine, citrulline, glutamic acid and D-arginine resulted in greater than 95% viability even at high concentrations of 100 µM peptide (Fig. 4.14A & B). In addition to this each **TP-1** analogue has a distinct viability advantage to **Pip6**, as at 30 µM **Pip6** returns only a 15% viability of HeLa cells.

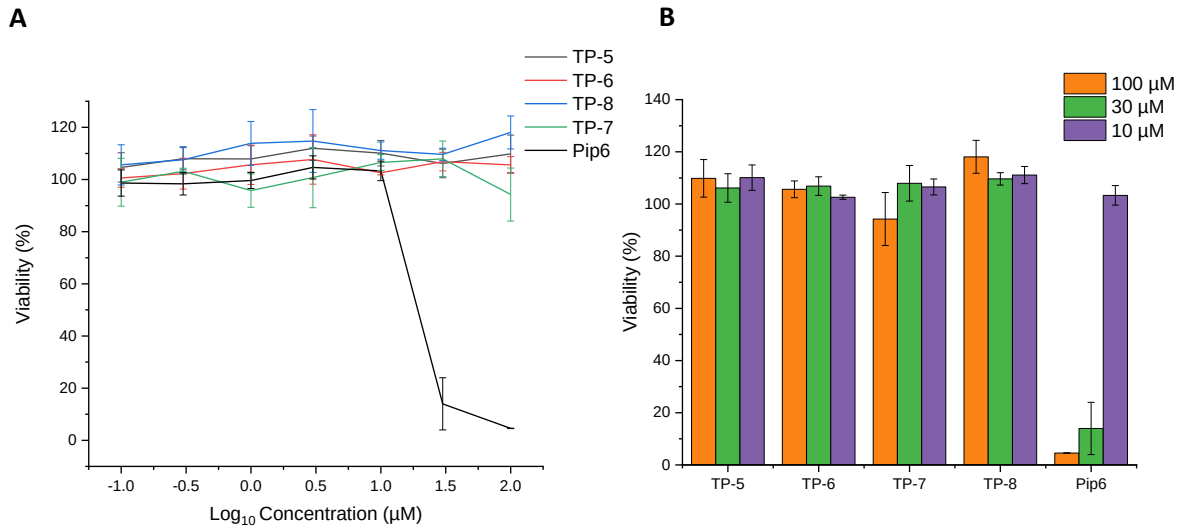


Figure 4.14. (A) Viability of HeLa cells after 24 h incubation at 37 °C with **TP-1** analogues. (B) Viability of HeLa cells with analogues at 100 µM, 30 µM and 10 µM.

Viability of HeLa cells with **TP-2** analogues returned similar results to the **TP-1** analogues, where all peptides gave a high HeLa cell viability even at 100 µM (Fig. 4.15A & B). The introduction of amidine group with the **TP-4** peptide also achieved high HeLa cell viability. The high viability of HeLa cells with the **TP-2** analogues, like the **TP-1** analogues, reinforces a significant advantage of these peptides over **Pip6** with regards to toxicity.

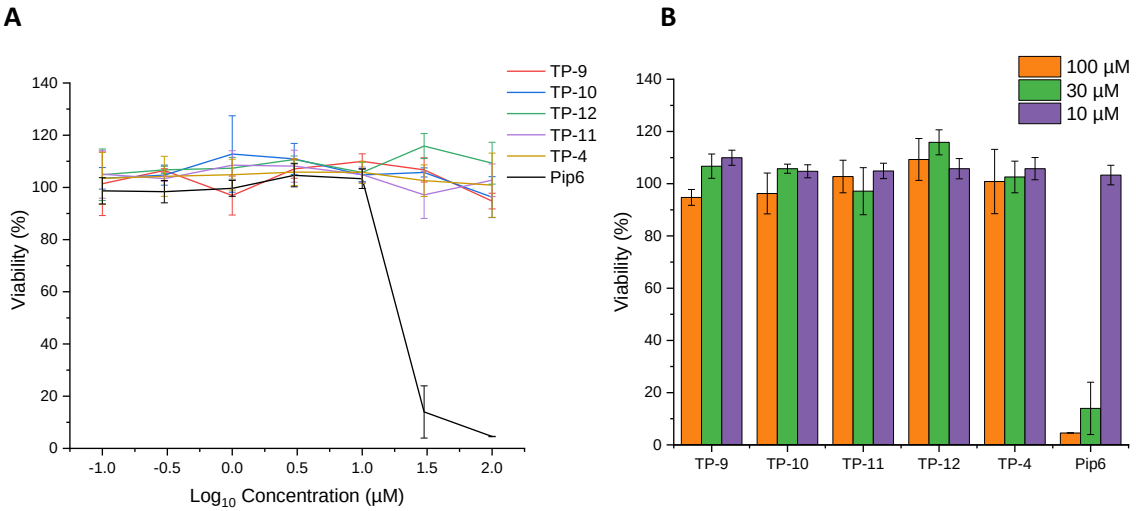


Figure 4.15. (A) Viability of HeLa cells after 24 h incubation at 37 °C with **TP-2** analogues. (B) Viability of HeLa cells with **TP-2** analogues at 100 μM, 30 μM and 10 μM.

4.3.5 Cell uptake of peptidomimetics

The **TP-1** analogues were first analysed for their uptake in HeLa cells after 4 h incubation at 37 °C at 5 μM concentration (Fig. 4.16A & B). First it can be noticed that **TP-5** shows no reduction in fluorescence output compared to **TP-1** and that **TP-6** exhibits only a small reduction in fluorescence compared to **TP-1**. The **TP-6** peptide also showed a similar fluorescence to **Pip6** after 4 h, highlighting that a neutral CPP can be retained in HeLa cells at similar levels to an arginine rich CPP over 4 h. The similar fluorescence intensity observed between **TP-1** and **TP-5** indicated that the cationic charge, without bidentate hydrogen bonding, is sufficient for uptake in the **TP-1** CPP. The negatively charged **TP-8** returned the lowest level of fluorescence of all peptides, highlighting that the negative charge in this peptide reduces the amount of cell uptake. Finally, a change to D-Arg, with **TP-7**, showed a clear increase in fluorescence in comparison to both **TP-1** and **TP-5**. The improvement in uptake for **TP-7** could be attributed to an enhanced protease stability and therefore lower levels of clearance of the CPP. These results indicate first that arginine is not essential for uptake of the **TP-1** CPP and that a cationic charge is more essential than specifically arginine itself.

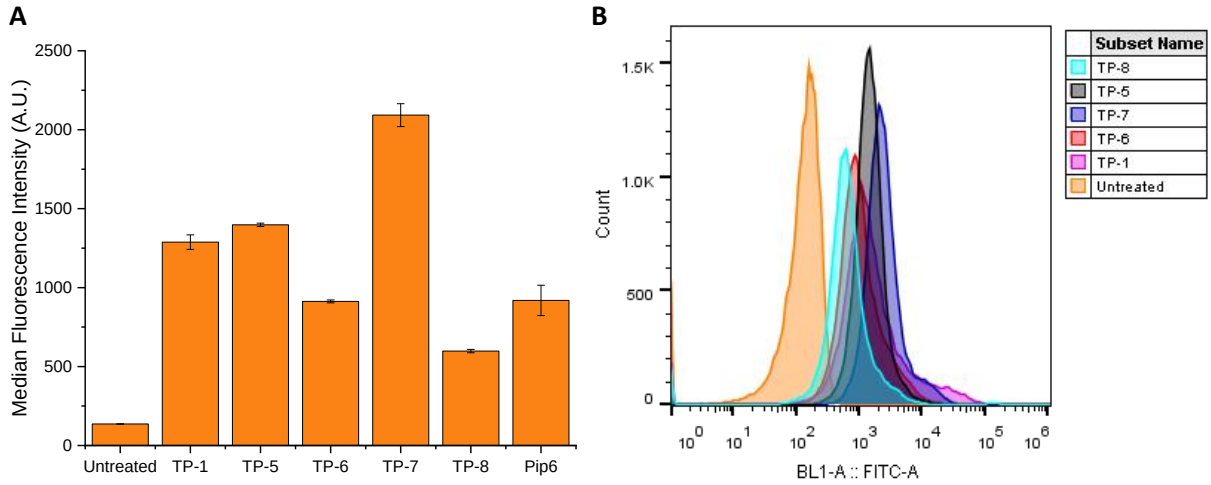


Figure 4.16. (A) Comparison in uptake of **TP-1** analogues after incubation at 37 °C for 4 h with HeLa cells. Error bars are mean ± SD of triplicate experiments. (B) Overlay of histograms of HeLa cells after 4 h incubation with **TP-1** analogues.

Similar observations were seen with the **TP-2** analogues as with the **TP-1** analogues (Fig. 4.17A & B). Again, a change from arginine to lysine **TP-9** had no real effect on fluorescence implying the **TP-2** CPP, like **TP-1**, only requires cationic charge at the arginine positions and not the additional hydrogen bonding profile offered by the guanidinium groups of arginine. A small reduction in fluorescence was observed with **TP-10** and a further reduction was observed with the anionic **TP-12**. The reductions in fluorescence with **TP-10** and **TP-12** imply that a change from cationic **TP-2** to a neutral, then negatively charged peptide, results in lower uptake of the CPP within the cell. **TP-11** returned a higher level of fluorescence than **TP-2** again suggesting that an enhanced stability enhances the quantity of CPP with the cell. Lastly, the **TP-4** peptide gave close to a 4-fold greater fluorescence than **TP-2**. The **TP-4** fluorescence was significantly higher than any other CPP tested over a 4 h time frame. The higher level of fluorescence of **TP-4** indicates that the use of aryl amidines enhances uptake and stability of the TP CPPs within the cell, as well as outperforming arginine rich CPPs such as **Pip6**.

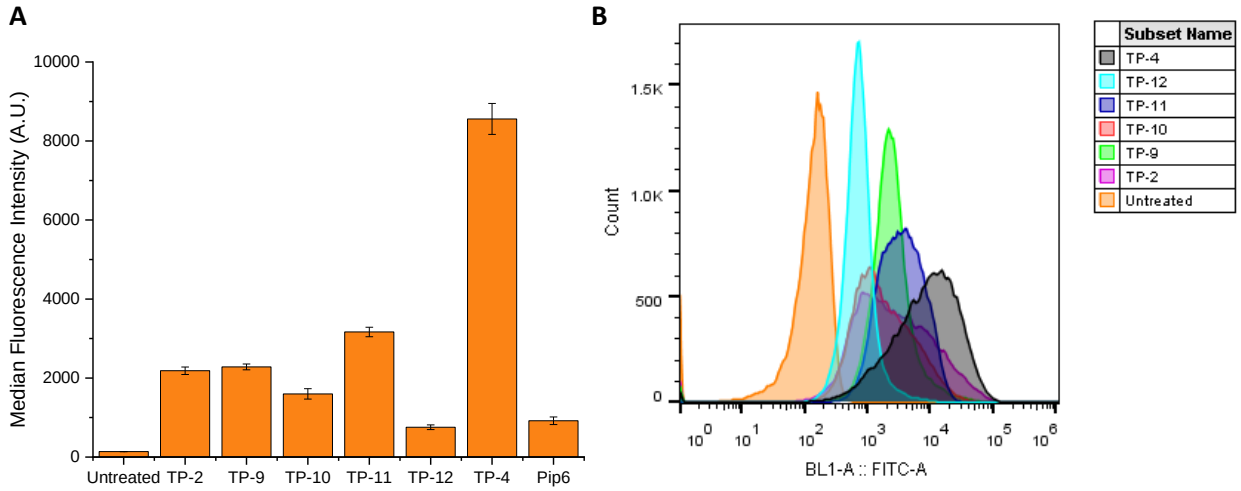


Figure 4.17. (A) Comparison in uptake of **TP-2** analogues after incubation at 37 °C for 4 h with HeLa cells. Error bars are mean ± SD of triplicate experiments. (B) Overlay of histograms of HeLa cells after 4 h incubation with **TP-2** analogues.

To further investigate the cell uptake profile of **TP-4** both timed and temperature dependant experiments were conducted (Fig. 4.18A & B). First, looking at the cell uptake of both **Pip6** and **TP-4** in HeLa cells over 1 h, it can be seen that for **Pip6** there is a reduction in fluorescence over time (Fig 4.18A). With **TP-4** the reverse is true with a near 8-fold increase in fluorescence from 1 h to 4 h incubation time. The comparison of both CPPs highlights that the uptake of **Pip6** is faster than **TP-4** but overall, after 4 h a greater amount of cargo is delivered within the cell by the **TP-4** CPP. To investigate the type of uptake undertaken by **TP-4** a 4 °C incubation experiment was undertaken in order to shut down any active process within the cell, including endocytotic uptake (Fig 4.18B). Upon incubation of cells at 4 °C there is a significant reduction in fluorescence of **TP-4**, indicating low levels of internalised peptide after 4 h. This reduced fluorescence contrasts with **Pip6** which shows an increased level of fluorescence at 4 °C than at 37 °C, which could be attributed to a reduction in protease activity and exocytotic mechanisms as discussed in chapter 2. The uptake of **TP-4** can be attributed to an energy dependant pathway, potentially a clathrin-caveolae independent endocytosis mechanism like **TP-2**. The greater level of fluorescence observed for **TP-4** could also indicate the aryl amidine groups enhanced protease stability and potentially binding to a specific cell surface receptor that can initiate an endocytotic uptake mechanism.

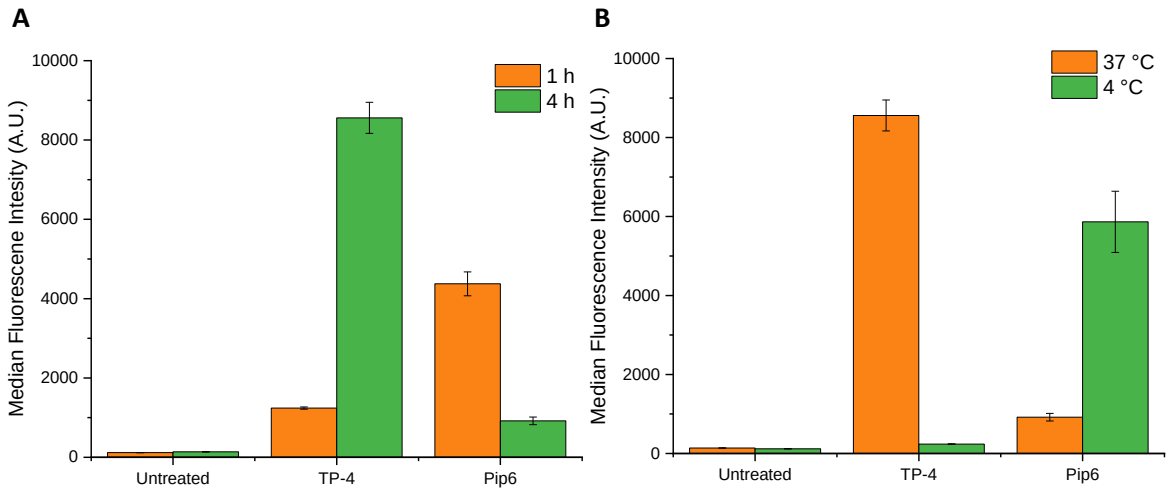


Figure 4.18. (A) Fluorescence of HeLa cells after treatment with **TP-4** and **Pip6** over 1 h and 4 h at 37 °C. (B) Fluorescence of HeLa cells after treatment with**TP-4** and **Pip6** over 4 h at 37 °C and 4 °C.

4.3.6 Sub-cellular localisation of peptidomimetics

Confocal microscopy was used to investigate any specific sub-cellular localisation of the CPPs. Using Hoechst 33258 to stain the nucleus, any nuclear localisation or exclusion could be observed with the fluorescein labelled CPPs synthesised. First, comparing the D-Arg analogue **TP-11** to natural arginine peptide **TP-2** and to arginine-rich **Pip6** (Fig. 4.19A, B & C). The **Pip6** CPP appears to show some overlap with Hoechst 33258 indicating some nuclear localisation. Both **TP-2** and **TP-11** appear to show little overlap with Hoescht dye and generally located throughout the cell cytoplasm. The similar localisation of **TP-2** and **TP-11** highlights that inverting the stereochemistry has no real effect on the peptide cellular localisation once taken up.

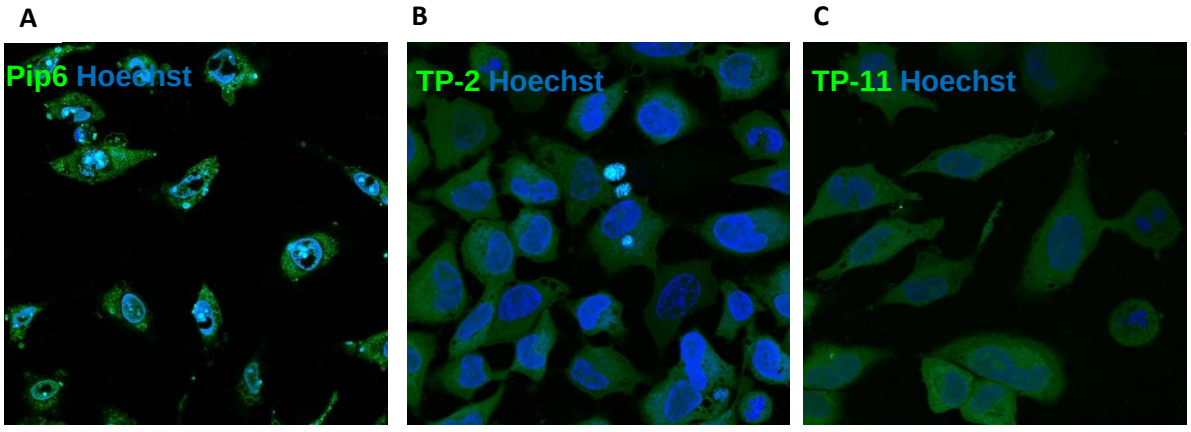


Figure 4.19. (A) Fluorescent Image of HeLa cells after treatment with Hoechst 33258 and **Pip6**. (B) Fluorescent Image of HeLa cells after treatment with Hoechst 33258 and **TP-2**. (C) Fluorescent Image of HeLa cells after treatment with Hoechst 33258 and **TP-11**.

Looking at **TP-4** in comparison to both **Pip6** and **TP-2** localisation was observed freely throughout the cytoplasm (Fig. 4.20A, B & C). A similar observation of localisation throughout the cytoplasm was observed with **TP-4** and **TP-2** highlights that the amidine group does not assert any change in cellular localisation when compared to the natural arginine. In addition, as **TP-4** appears generally throughout the cytoplasm, it could indicate the CPP can escape endosomes within 4 h incubation time.

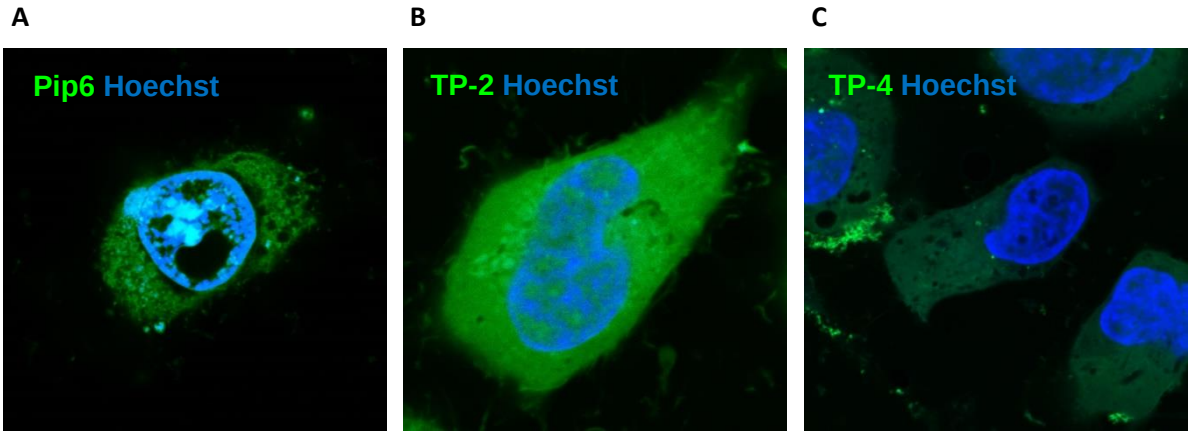


Figure 4.20. (A) Fluorescent Image of a HeLa cell after treatment with Hoechst 33258 and **Pip6**. (B) Fluorescent Image of a HeLa cell after treatment with Hoechst 33258 and **TP-2**. (C) Fluorescent Image of HeLa cells after treatment with Hoechst 33258 and **TP-4**.

Overall, it appears that changes to the arginine residues in the **TP-2** peptide have little effect on the localisation of peptide but can maintain or enhance the overall cell uptake. The aryl amidine groups maintain a distinctly different cell uptake profile in the TP peptides compared to arginine-rich CPPs like **Pip6**.

4.4 Summary of Chapter 4

In this chapter the amidine group was investigated as guanidine mimetic for application in CPPs. Initially, alkyl amidine building blocks could not be synthesised for application in SPPS. Fortunately, an aryl amidine building block was prepared in a three-step synthesis with an overall yield of 30%. The Pbf protecting group of the amidine was found to be less labile than a Pbf protected guanidine and required heating to fully liberate the free amidine side chain. Furthermore, incorporation of the aryl amidine in a peptide chain required harsher conditions than the standard arginine building blocks, with microwave assisted heating used to overcome potential steric blocking effects of the Pbf protected aryl amidine.

A series of fluorescein labelled mimetic CPPs, based on the TP peptides, were synthesised by SPPS. These peptides included lysine, citrulline, glutamic acid, D-arginine and aryl amidine mimetics. Each of the mimetic peptides was shown to relatively non-toxic to HeLa cells, with viability of the cells generally above 95% for each mimetic CPP.

The mimetic CPPs allowed for a greater understanding of important residues in both the TP peptides with regards to cell uptake. As changes from arginine to lysine in both TP peptides returned similar levels of uptake, it can be determined that cationic charge at the arginine positions is one of the factors for uptake and additional hydrogen bonding is not required. The citrulline and glutamic acid CPPs further reinforce the need for cationic charge, as both peptides saw a reduction in fluorescence compared to arginine and lysine based TP peptides upon flow cytometry analysis. The change to D-arginine enhanced the level of uptake of the TP peptides and could be attributed to a greater level of protease stability.

The aryl amidine **TP-2** analogue, **TP-4**, returned the greatest level of fluorescence after 4 h incubation of any CPP analysed in HeLa cells. Upon further inspection it was found that **TP-4** was taken up *via* an energy dependant mechanism and was also taken up into cells slower than standard arginine rich CPP **Pip6**. The enhanced uptake could be attributed to the greater protease stability of the aryl amidine groups and their ability to strongly interact with an appropriate species on the cell surface that mediates cell uptake. Both aryl amidines and D-arginine changes to the **TP-2** peptide did not alter the CPPs subcellular localisation, with both peptides freely located in the cytoplasm which indicates that the aryl amidine may be able to efficiently escape endosomes within 4 h after uptake.

This work highlights that arginine residues are not essential for cell uptake of the TP CPPs and that an aryl amidine peptide, **TP-4**, is a suitable peptidomimetic for the original **TP-2** CPP.

4.5 Experimental

4.5.1 General information

All reagents and solvents were obtained from commercial suppliers and were used without further purification. All reactions were carried out under air unless otherwise stated. Reactions were monitored by thin layer chromatography (TLC) using Merck silica plates coated with fluorescent indicator UV254. TLC plates were analysed using 254/365 nm UV light or developed using potassium permanganate solution.

4.5.2 Peptide synthesis

Peptide synthesis was completed on an automated Tribute® peptide synthesiser with an IntelliSynth UV-monitoring system and feedback control system. Rink amide resin (100-200 mesh), Fmoc-Phe-OH, Fmoc-Gly-OH, Fmoc-Tyr(O^{t}Bu)-OH, Fmoc-Pro-OH, Fmoc-Leu-OH, Fmoc-Ile-OH, Fmoc-Gln(Trt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-D-Arg(Pbf)-OH, Fmoc-Cit-OH, Fmoc-Lys(Boc)-OH and Fmoc-εAhx-OH were purchased from Merck Millipore or Fluorochem and used without further purification. Fluoresecein 5-isothiocyanate (FITC) was purchased from sigma-aldrich and used without further purification.

4.5.3 Analytical HPLC

RP-HPLC was carried out using Aeris 3.6 µm RP 250 × 4.6 mm widepore XB C18 column using a DIONEX 3000 series HPLC equipped with a VWD3400 photodiode array detector. Purifications were performed using water (0.1% TFA) as Solvent A and acetonitrile (0.1% TFA) as Solvent B and were run at a flow rate of 1.0 mL/min. For analytical HPLC traces see appendix *chapter 4*.

Analytical RP-HPLC method A:

Absorbance detection was set to 220 nm.

Table 4.1. Gradient used for analytical RP-HPLC method A.

Time (min)	Solvent A	Solvent B
0	95%	5%
5	95%	5%
50	10%	90%
55	10%	90%
60	95%	5%

4.5.4 LCMS Analysis

LCMS was carried out using a Zorbax 3.6 µm RP 34 mm × 150 mm C18 column using a Agilent quadrupole 1200 series. Chromatography were performed using 5 mM ammonium acetate in water as Solvent A and 5 mM ammonium acetate in acetonitrile as Solvent B and were run at a flow rate of 1.0 mL/min.

LC: method:

Absorbance detection was set to 254 nm.

Table 4.2. Gradient used for analytical LCMS method.

Time (min)	Solvent A	Solvent B
0	95%	5%
1.48	95%	5%
8.5	0%	100%
13.5	0%	100%
16.5	95%	5%
18	95%	5%

4.5.5 Purification of products

Semi-preparative reversed-phase HPLC purification was carried out on a kinetex 5 µm RP 150 × 21.2 mm XB C18 column using a DIONEX 3000 series HPLC equipped with a VWD3400 variable wavelength detector. Purifications were performed using water (0.1% TFA) as Solvent A and acetonitrile (0.1% TFA) as Solvent B and were run at a flow rate of 12.0 mL/min.

RP-HPLC method A:

Absorbance detection was set to 220 nm.

Table 4.3. Gradient used for RP-HPLC method A.

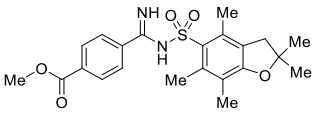
Time (min)	Solvent A	Solvent B
0	95%	5%
5	95%	5%
45	10%	90%
50	10%	90%
53	95%	5%

4.5.6 Analysis of products

Fourier-Transform Infra-Red (FTIR) spectra were obtained on a Shimadzu IRAffinity-1 spectrometer. ^{19}F NMR spectra were obtained on a Bruker AV 400 spectrometer at 376 MHz. ^{1}H and ^{13}C NMR spectra were obtained on either a Bruker AV 400 at 400 MHz and 125 MHz, respectively, or Bruker DRX 500 at 500 MHz and 126 MHz, respectively. Chemical shifts are reported in ppm and coupling constants are reported in Hz. High-resolution mass spectra were recorded on a Bruker microTOF II mass spectrometer at the University of Edinburgh.

4.5.7 Synthetic procedures

Methyl 4-(*N*-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)carbamimidoyl)benzoate (**4.85**)



Methyl 4-carbamimidoylbenzoate hydrochloride (2.00 g, 9.32 mmol) was suspended in acetone (40.0 mL) at 0 °C before the addition of 3 M NaOH solution (6.00 mL, 18.0 mmol) to form a clear solution. Separately Pbf chloride (3.23 g, 11.2 mmol) was dissolved in acetone (13.5 mL) then added dropwise to the reaction mixture. The resulting pale-yellow solution was stirred at 0 °C for 2 h until a white suspension formed, the reaction was then stirred for a further 2 h at rt. The reaction was subsequently acidified to pH 6 with 3 M HCl solution and the suspension filtered. The solid was washed with water (15 mL) then acetone (15 mL). The filtrate was then extracted with EtOAc (3 × 50 mL) and the combined organic layers were then washed with brine (100 mL) before being dried over $MgSO_4$. The organic solution was then concentrated *in vacuo* and the resulting residue purified by SiO_2 chromatography in gradient of 0% to 40% EtOAc in petroleum ether (40-60) to afford a colourless solid (3.03 g, 75%).

mp 400 °C degraded.

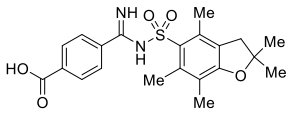
^{1}H NMR (500 MHz, $CDCl_3$): δ/ppm: 8.21 (br s, 1H, N*H*), 8.05 (d, 2H, *J* = 8.5 Hz, H_3CO-C(O)-C_q-C*H* × 2), 7.82 (d, 2H, , *J* = 8.5 Hz, H_3CO-C(O)-C_q-CH-C*H* × 2), 6.22 (br s, 1H, N*H*), 3.92 (s, 3H, C_q-C(O)-OC*H*$_3$), 2.96 (s, 2H, O-$C_q(CH_3)_2$-C*H*$_2$), 2.62 (s, 3H, $S(O)_2$-C_q-C_q-(C*H*$_3$)-C_q-(CH_3)),), 2.55 (s, 3H, $S(O)_2$-C_q-C_q-(C*H*$_3$)-C_q), 2.10 (s, 3H, $S(O)_2$-C_q-$C_q(CH_3)$-C_q-C*H*$_3$), 1.46 (s, 6H, C_q-C*H*$_3$ × 2).

^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$): δ/ppm: 166.0 (C_q), 160.2 (C_q), 159.6 (C_q), 139.3 (C_q), 137.9 (C_q), 133.5 (C_q), 133.2 (C_q), 130.9 (C_q), 129.9 (2 × *C*-H), 127.3 (2 × *C*-H), 124.9 (C_q), 117.8 (C_q), 86.7 (C_q), 52.4 (*C*H$_3$), 43.1 (*C*H$_2$), 28.6 (2 × *C*H$_3$), 19.2 (*C*H$_3$), 17.9 (*C*H$_3$), 12.4 (*C*H$_3$).

IR (neat): **ν_{max} /cm^{-1}**: 3450 (N-H stretch), 3334 (N-H stretch), 2973 (C-H stretch), 2929 (C-H stretch), 1703 (C=O stretch), 1628 (C=N stretch), 1578 (C=C stretch), 1533 (C=C stretch).

HRMS (ESI): Calc. for $C_{22}H_{27}N_2O_5S^+$. Theoretical: 431.1635 Observed: 431.1637.

4-(*N*-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)carbamimidoyl)benzoic acid (**4.86**)



Amidine **4.85** (2.50 g, 5.81 mmol) was dissolved in THF (29.1 mL) and then 1M NaOH solution (29.1 mL, 29.1 mmol) was added dropwise, and the resulting solution was stirred at rt for 2 h. The THF was removed from the reaction *in vacuo* and then the reaction was diluted with water (25 mL). The aqueous solution was then acidified to pH 5 with 1 M HCl solution to form an off-white gum-like precipitate. The suspension was filtered and the solid washed with water (25 mL). The white solid was then dried under high vacuum before being purified by SiO_2 chromatography in gradient of 0% to 50% EtOAc in petroleum ether (40-60) with 1% AcOH modifier to afford a colourless solid (1.90 g, 78%).

mp 400 °C degraded.

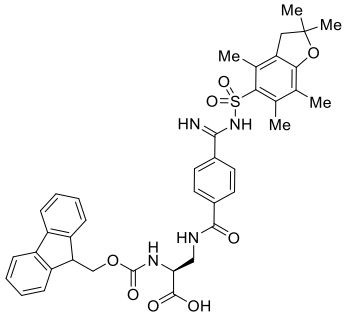
^{1}H NMR (500 MHz, DMSO-*d6*): δ/ppm: 13.23 (br s, 1H, CO_2H), 8.98 (br s, 1H, N*H*), 8.06 (br s, 1H, N*H*), 8.00 (d, 2H, *J* = 8.7 Hz, HO_2C-C_q-C*H* × 2), 7.91 (d, 2H, *J* = 8.7 Hz, HO_2C-C_q-CH-C*H* × 2), 3.00 (s, 2H, O-$C_q(CH_3)_2$-CH_2), 2.52 (s, 3H, $S(O)_2$-C_q-C_q-(CH_3)-C_q-(CH_3)), 2.46 (s, 3H, $S(O)_2$-C_q-C_q-(CH_3)-C_q), 2.04 (s, 3H, $S(O)_2$-C_q-$C_q(CH_3)$-C_q-CH_3), 1.43 (s, 6H, C_q-CH_3 × 2).

^{13}C{^{1}H} NMR (125 MHz, DMSO-*d6*): δ/ppm: 167.1 (C_q), 160.9 (C_q), 158.8 (C_q), 138.5 (C_q), 138.0 (C_q), 134.2 (C_q), 132.7 (C_q), 132.5 (C_q), 129.8 (2 × *C*-H), 128.4 (2 × *C*-H), 125.3 (C_q), 117.2 (C_q), 87.2 (C_q), 42.7 (*C*H_2), 28.7 (2 × *C*H_3), 19.3 (*C*H_3), 18.1 (*C*H_3), 12.7 (*C*H_3).

IR (neat): **ν_{max} /cm^{-1}**: 3358 (N-H stretch), 3223 (N-H stretch), 3091 (CO_2H stretch), 2972 (C-H stretch), 2937 (C-H stretch), 1710 (C=O stretch), 1670 (C=N stretch), 1589 (C=C stretch), 1533 (C=C stretch).

HRMS (ESI): Calc. for $C_{21}H_{25}N_2O_5S^+$. Theoretical: 417.1479 Observed: 417.1474.

(*S*)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(*N*-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)carbamimidoyl)benzamido)propanoic acid (**4.87**)



Amidine **4.86** (1.10 g, 2.64 mmol) and HATU (1.00 g, 2.64 mmol) were dissolved in DMF (26.4 mL) and then DIPEA (920 µL, 5.28 mmol) was added dropwise and the resulting yellow solution was stirred at rt for 30 min. The reaction became a dark orange colour to which Fmoc-Dap-OH (1.15 g, 3.17 mmol) was added portion wise and the reaction mixture was stirred at rt for 4.5 h. EtOAc (80 mL) and water (80 mL) were added and the organic layer separated. The aqueous layer was then extracted with EtOAc (2 × 80 mL). The combined organic layers were then washed with brine (150 mL), dried over $MgSO_4$ then filtered, before being concentrated *in vacuo*. The resulting residue was purified by SiO_2

chromatography in a gradient of 0% to 4% MeOH in DCM with 1% AcOH modifier to afford a colourless solid (905 mg, 51%).

mp 113-117 °C.

$[\alpha]_D$ = -16.86 (c = 0.011 g mL^{-1}, MeOH)

^{1}H NMR (500 MHz, CD_3OD): δ/ppm: 7.76 (d, 2H, *J* = 8.6 Hz, HN=C_q-C_q-CH-C*H* × 2), 7.72 (d, 2H, *J* = 8.6 Hz, HN=C_q-C_q-C*H* × 2), 7.62 (dd, 2H, *J* = 7.5, 3.3 Hz, CO_2-CH_2-CH-C_q-CH-CH-CH-C*H* × 2), 7.50 (d, 2H, *J* = 7.5 Hz, CO_2-CH_2-CH-C_q-C*H* × 2), 7.22 (2H, m, CO_2-CH_2-CH-C_q-CH-CH-C*H* × 2), 7.12 (2H, m, CO_2-CH_2-CH-C_q-CH-C*H* × 2), 4.41 (1H, m, HO_2C-C*H*), 4.19 (2H, m, CO_2-CH_2), 4.06 (1H, t, *J* = 7.1 Hz, CO_2-CH_2-C*H*) 3.74 (1H, m, HO_2C-CH-C*H*), 3.63 (1H, m, HO_2C-CH-C*H*), 2.91 (s, 2H, O-C_q(CH_3)$_2$-CH_2), 2.49 (s, 3H, S(O)$_2$-C_q-C_q-(CH_3)- C_q-(CH_3)), 2.44 (s, 3H, S(O)$_2$-C_q-C_q-(CH_3)-C_q), 1.99 (s, 3H, S(O)$_2$-C_q-C_q(CH_3)-C_q-CH_3), 1.35 (s, 6H, C_q-CH_3 × 2).

^{13}C{^{1}H} NMR (125 MHz, CD_3OD): δ/ppm: 172.1 (C_q), 168.3 (C_q), 161.6 (C_q), 159.2 (C_q), 157.2 (C_q), 143.8 (C_q), 143.7 (C_q), 141.1 (2 × C_q), 138.7 (C_q), 137.4 (C_q), 136.6 (C_q), 132.9 (C_q), 131.3 (C_q), 127.6 (2 × *C*-H), 127.4 (2 × *C*-H), 127.2 (2 × *C*-H), 126.7 (2 × *C*-H), 125.0 (C_q), 124.8 (2 × *C*-H), 119.5 (2 × *C*-H), 117.3 (C_q), 86.6 (C_q), 66.7 (*C*H_2), 53.8 (*C*H), 46.8 (*C*H), 42.4 (*C*H_2), 40.9 (*C*H_2), 27.3 (2 × *C*H_3), 18.0 (*C*H_3), 16.9 (*C*H_3), 11.0 (*C*H_3).

IR (neat): **ν_{max} /cm^{-1}**: 3394 (N-H stretch), 3311 (N-H stretch), 3115 (CO_2H stretch), 2972 (C-H stretch), 2927 (C-H stretch), 1716 (C=O stretch), 1705 (C=O stretch), 1627 (C=N stretch), 1575 (C=C stretch), 1527 (C=C stretch).

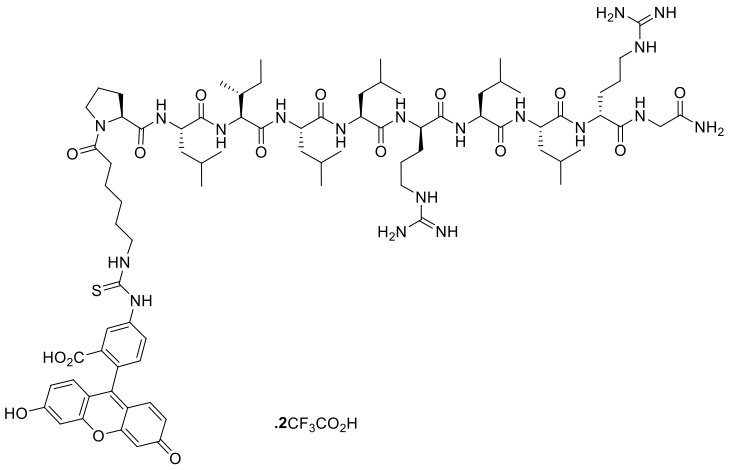
HRMS (ESI): Calc. for $C_{39}H_{41}N_4O_8S^+$. Theoretical: 725.2640 Observed: 725.2624.

4.5.8 General peptide synthesis protocol

On a Tribute® solid phase peptide synthesiser Rink amide resin (0.15 mmol) was swelled in DCM (5.00 mL) for 30 min. Two 10 min agitations with 20% (v/v) solution of piperidine in DMF (5.00 mL) was used to deprotect terminal Fmoc groups. Activation of Fmoc protected amino acids (0.45 mmol, 3.00 equiv) was achieved using HATU (0.38 mmol, 2.50 eq) and 0.5 M DIPEA in DMF (5.00 mL). The reaction mixture of the activated esters of Fmoc amino acids were then added sequentially to the resin and agitated for 20 min at room temperature. Upon final coupling of appropriate residue, the resin was washed with DCM (5.00 mL) and dried under N_2. The resin was then removed from the automated synthesiser and manually swelled in DCM (5.00 mL) for 30 min. After swelling the resin was washed with DMF (4 × 2.00 mL) and two 10 min agitations with 20% (v/v) solution of piperidine in DMF (5.00 mL) was used to deprotect the terminal Fmoc group. Separately in a 5.00 mL glass vial FITC (117 mg, 0.30 mmol) was dissolved in DMF (1.50 mL) and DIPEA (183 µL, 1.05 mmol) was added to form a bright red solution. The red FITC solution was then added to the resin, protected from light and agitated for 16 h at room temperature. After the 16 h agitation the resin was washed with DMF (4 × 2.00 mL) before being washed with MeOH (4 × 2.00 mL) and finally DCM (4 × 2.00 mL). The peptide was then cleaved from the resin using a mixture of TFA:phenol:water:TIPS (90:5:5:2, 3.00 mL) and agitated for 4 h. The cleavage solution was then filtered from the resin and the peptide precipitated using cold Et_2O (30 mL, -20˚C) to afford an orange precipitate. The solid was washed with Et_2O (3 × 10 mL) to remove excess TFA and the resulting orange solid was purified by RP-HPLC method A.

4.5.9 Peptide characterisation

FITC – XPLILLrLLrG **(TP-7)**

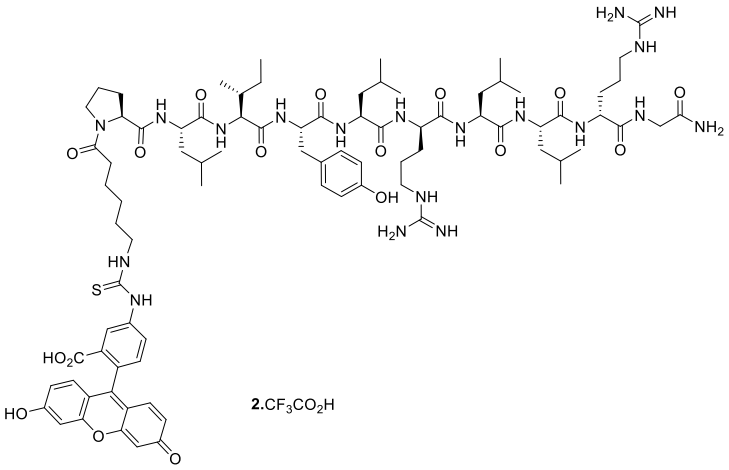


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (10 mg, 5%).

HRMS: $C_{82}H_{127}N_{19}O_{16}S^{2+}$. Theoretical: 832.9709 Observed: 832.9711.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 30.1 min, 100%.

FITC – XPLIYLrLLrG **(TP-11)**

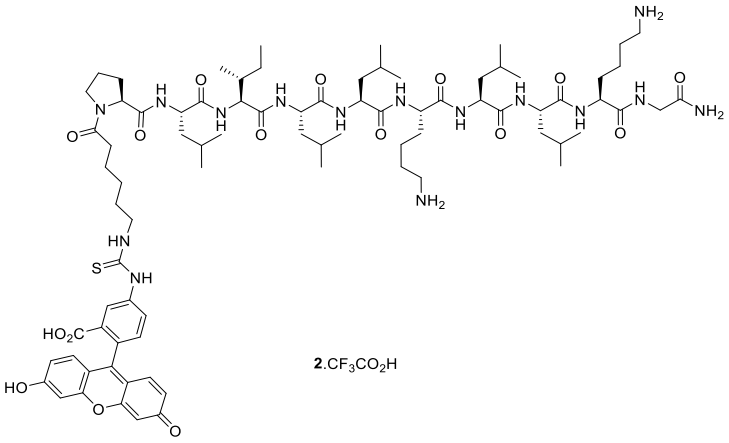


Using general peptide synthesis protocol A on a 0.10 mmol scale of resin. Rink amide resin (192 mg, 0.10 mmol, 0.52 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (12 mg, 6%).

HRMS: $C_{85}H_{125}N_{19}O_{17}S^{2+}$. Theoretical: 857.9605 Observed: 857.9600.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 28.6 min, 97%.

FITC – XPLILLKLLKG (**TP-5**)

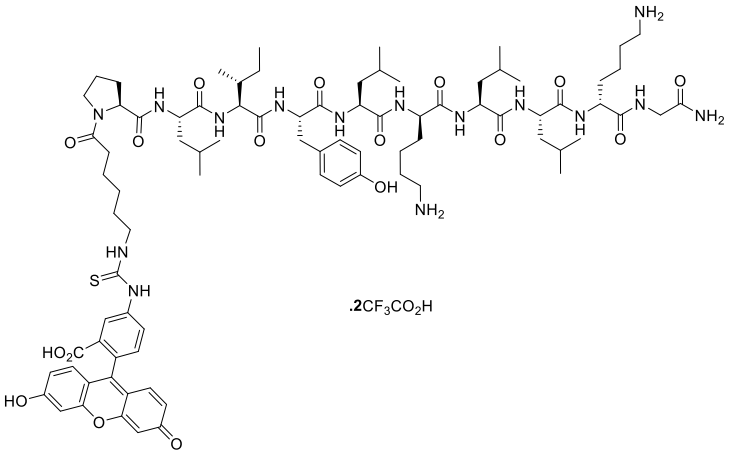


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (21 mg, 8%).

HRMS: $C_{82}H_{127}N_{15}O_{16}S^{2+}$. Theoretical: 804.9648 Observed: 804.9647.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 30.6 min, 97%.

FITC – PLIYLKLLKG (**TP-9**)

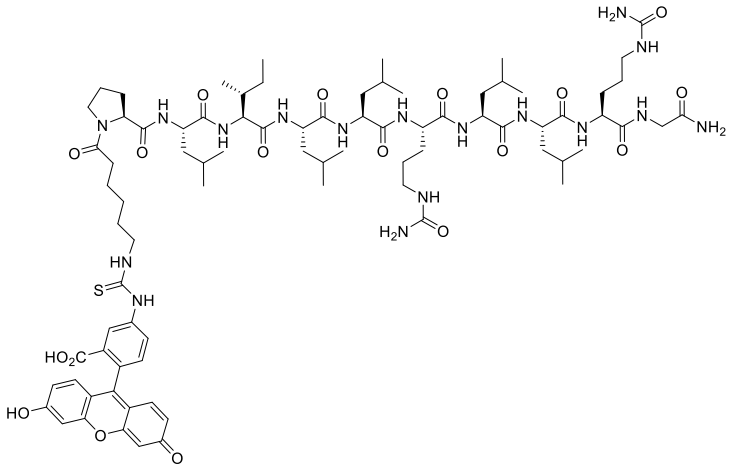


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (17 mg, 6%).

HRMS: $C_{85}H_{125}N_{15}O_{17}S^{2+}$. Theoretical: 829.9544 Observed: 829.9544.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 29.0 min, 97%.

FITC – PLILLCitLLCitG (**TP-6**)

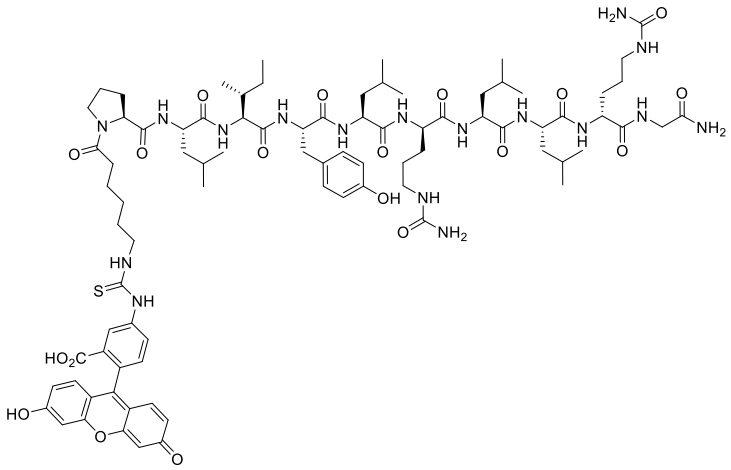


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (7 mg, 3%).

HRMS: $C_{82}H_{125}N_{17}O_{18}S^{2+}$. Theoretical: 833.9549 Observed: 833.9548.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 x 21.2 mm XB C18 column) R_t = 33.7 min, 98%.

FITC –XPLIYLCitLLCitG (**TP-10**)

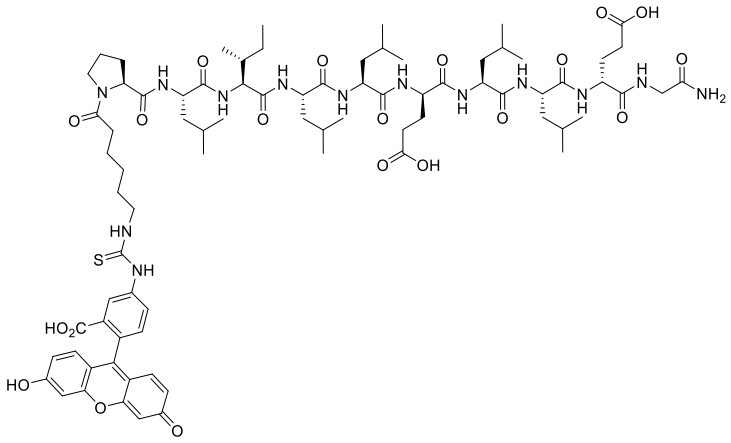


Using general peptide synthesis protocol A. Rink amide resin (231 mg, 0.15 mmol, 0.65 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (5 mg, 2%).

HRMS: $C_{82}H_{123}N_{17}O_{19}S^{2+}$. Theoretical: 858.9445 Observed: 858.9444.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 30.3 min, 100%.

FITC – XPLILLELLEG (**TP-8**)

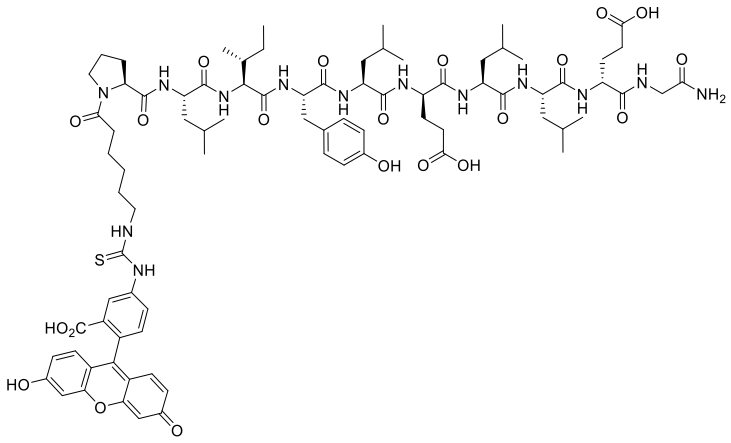


Using general peptide synthesis protocol A on a 0.075 mmol scale of resin. Rink amide resin (144 mg, 0.075 mmol, 0.52 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (9 mg, 7%).

HRMS: $C_{80}H_{117}N_{13}O_{20}S^{2+}$. Theoretical: 805.9124 Observed: 805.9100.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 34.6 min, 99%.

FITC –XPLIYLELLEG (**TP-12**)

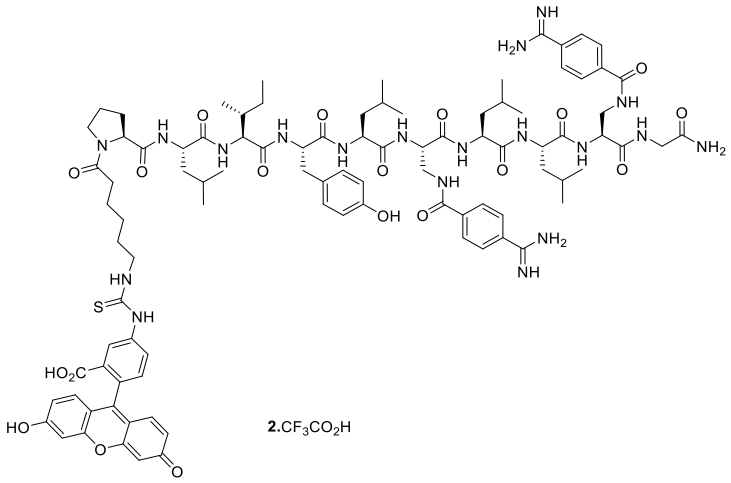


Using general peptide synthesis protocol A on a 0.75 mmol scale of resin. Rink amide resin (144 mg, 0.075 mmol, 0.52 mmol g^{-1}). Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (8 mg, 6%).

HRMS: $C_{83}H_{115}N_{13}O_{21}S^{2+}$. Theoretical: 830.9020 Observed: 830.9004.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 µm RP 150 × 21.2 mm XB C18 column) R_t = 31.1 min, 98%.

FITC – XPLIYLAmLLAmG (**TP-4**)



On a Tribute® solid phase peptide synthesiser Rink amide chem matrix resin (50 mg, 0.026 mmol, 0.52 mmol g^{-1}) was swelled in DCM (2.00 mL) for 30 min. Two 10 min agitations with 20% (v/v) solution of piperidine in DMF (4.00 mL) was used to deprotect terminal Fmoc groups. Activation of Fmoc protected amino acids (0.078 mmol, 3.00 equiv) was achieved using HATU (0.065 mmol, 2.50 eq) and 0.5 M DIPEA in DMF (2.00 mL). Activation of Fmoc Amidine **4.87** and Fmoc-Leu-OH position 6 was achieved with DIC (11 µL, 0.070 mmol, 2.70 equiv.) and HOAt (10 mg, 0.074 mmol, 2.85 equiv.) in DMF (2.00 mL). The reaction mixture of the activated esters of Fmoc amino acids, except amidine **4.87** and Fmoc-Leu-OH position 6, were then added sequentially to the resin and agitated for 20 min at room temperature. Amidine **4.87** and Fmoc-Leu-OH position 6 activation solutions were added to the resin and heated to 75 °C for 30 min. Upon final coupling of appropriate residue the resin was washed with DCM (2.00 mL) and dried under N_2. The resin was then removed from the automated synthesiser and manually swelled in DCM (2.00 mL) for 30 min. After swelling the resin was washed with DMF (4 × 2.00 mL) and two 10 min agitations with 20% (v/v) solution of piperidine in DMF (2.00 mL) was used to deprotect the terminal Fmoc group. Separately in a 5.00 mL glass vial FITC (28 mg, 0.078 mmol) was dissolved in DMF (1.00 mL) and DIPEA (32 µL, 0.182 mmol) was added to form a bright red solution. The red FITC solution was then added to the resin, protected from light and agitated for 16 h at room temperature. After the 16 h agitation the resin was washed with DMF (4 × 2.00 mL) before being washed with MeOH (4 × 2.00 mL) and finally DCM (4 × 2.00 mL). The peptide was then cleaved from the resin using a mixture of TFA:phenol:water:TIPS (90:5:5:2, 3.00 mL) and agitated for 4 h at 60 °C. The cleavage solution was then filtered from the resin and the peptide precipitated using cold Et_2O (30 mL, -20˚C) to afford an orange precipitate. The solid was washed with Et_2O (3 × 10 mL) to remove excess TFA and the resulting orange solid was purified by RP-HPLC method A. Appropriate HPLC fractions were lyophilised to afford a powder like orange solid (2 mg, 4%).

HRMS: $C_{95}H_{125}N_{19}O_{19}S^{2+}$. Theoretical: 933.9554 Observed: 933.9534.

RP-HPLC (Analytical RP-HPLC method A, Kinetex 5 μm RP 150 × 21.2 mm XB C18 column) R_t = 28.2 min, 98%.

4.5.10 Cell culture

HeLa cells were maintained in medium consisting of Dulbecco's modified Eagle's medium (DMEM), 10% FBS, 1% L-Glutamine and 1% penicillin/streptomycin. Cells were cultured in a humidified incubator at 37 °C with 5% CO_2 atmosphere. Cells were cultured in a humidified incubator at 37 °C with 5% CO_2 and passaged twice a week at confluence > 80%. BSA, FBS, L-Glutamine, penicillin/streptomycin and TrypLee Xpress were purchased sterile from sigma-aldrich or Thermo fischer scientific and used without further treatment.

4.5.11 Flow cytometry

HeLa cells were cultured in 24 well plates (4.5×10^5 cells per well) for 24 h. On the day of the experiment, the cells were incubated with 5 μM FITC-peptide in DMEM with 10% FBS at 37 °C or 4 °C for required time. The media was then removed and wells were washed with PBS (2 × 500 μL per well) and then detached from the plate with TrypLee Xpress (500 μL per well). PBS containing 2% bovine serum albumin (500 μL) was added to each well and cells transferred to 5 mL polystyrene FACS tubes. Cells were centrifuged and washed with PBS containing 2% bovine serum albumin (2 × 500 μL) before being suspended in PBS containing 2% bovine serum albumin (1.00 mL each). Cells were finally analysed on a Thermo fischer Attune NxT flow cytometer with laser excitation/emission set at 490 nm/519 nm and a flow rate of 200 μL min^{-1}. Flow cytometry data was analysed using FlowJo software with gating set at a range of 51K-285K side scatter and 251K-815K for forward scatter variance for every sample. Median fluorescence intensity values were calculated from the gating range using FlowJo.

4.5.12 Cell viability alamarBlue assay

HeLa cells were cultured in 96 well plates (2.6×10^3 cells per well) for 24 h. On the day of the experiment, the cells were incubated with 100 μM, 30 μM, 10 μM, 3 μM, 1 μM, 0.3 μM, 0.1 μM, FITC-peptide in DMEM with 10% FBS (100 uL total well volume) at 37 °C for 24 h. AlamarBlue® was then added to each well (0.5 mM, 10 μL) and incubated for a further 6 h for HeLa cells and 24 h for U2OS cells. Cells were then analysed using a Hldex plate reader with laser excitation/emission set at 560 nm/595 nm. Fluorescence of untreated cells were used as a 100% viable control and viabilities were calculated as a percentage of the control fluorescence values using OriginPro 2019b software.

4.5.13 Confocal microscopy

HeLa cells were cultured in 8 well plates ibidi plates (3.5×10^4 cells per well) for 24 h. On the day of the experiment, the cells were incubated with 5 μM FITC-peptide in DMEM with 10% FBS (200 uL

total well volume) at 37 °C for 4 h. After 3.5 h incubation, Hoescht 33258 was added at a concentration of 5.6 μM and cells incubated at 37 °C for the final 0.5 h. Cells were removed from the incubator and media removed. Each well was then washed with serum and phenol red free media (2 × 200 μL) then incubated with 4% formaldehyde in PBS (200 μL) for 20 min at room temperature. After incubation the 4% formaldehyde solution was removed and PBS (200 μL) added to each well. The cells were then imaged using Leica S8 confocal microscope. Excitation/emission was set at 352 nm/461 nm for Hoescht 33258 imaging and 495 nm/519 nm for FITC labelled peptides. Images were processed using ImageJ software.

Chapter 5: Conclusions and future directions

5.1 Conclusions and future directions of the project

This thesis describes the preparation and biological evaluation of fluorescently labelled CPPs in order to compare levels of toxicity and uptake between arginine rich CPPs and low arginine containing CPPs. In addition, investigations into the specific mechanisms of cell uptake and sub-cellular localisation were undertaken for both arginine rich and low arginine containing CPPs. Synthetic methodology to directly functionalise amidines was reported for potential use in CPPs and other biologically relevant compounds. Finally, the synthesis and biological evaluation of peptidomimetics were described including the use of amidines as guanidine mimetics.

The initially hypothesised advantages to low arginine containing CPPs and amidine mimetics were:

i) Lower toxicity than highly basic, arginine rich CPPs.
ii) Similar levels of uptake relative to standard arginine-rich CPPs.
iii) Amidine mimetics maintain or even enhance cell uptake of CPPs through a similar cell surface binding profile.
iv) The physiochemical properties of the amidine group could be modulated through development of mild synthetic methodology.

The overall outcomes of this work were that low arginine containing CPPs **TP-1** and **TP-2** were less toxic than arginine rich CPPs **Tat** and **Pip6**. **Pip6** showed significantly higher toxicity than **TP-1** and **TP-2** across two different cell lines at concentrations of 30 μM and higher, suggesting a general necrotic mechanism of toxicity. The levels of uptake of both TP CPPs were similar to **Pip6** and greater that **Tat** after 4 h incubation time. Further investigation found that both arginine rich CPPs were initially taken up into cells at higher quantities but suffered from a form of expulsion from cells that was not observed with **TP-1** or **TP-2**. The mechanism of uptake for each TP CPP was found to be *via* an energy dependant pathway in both HeLa and U2OS cells, though none of the three major classes of endocytosis (clathrin mediated, caveolae mediated and macropinocytosis) appeared to majorly contribute to uptake. Furthermore, both **TP-1** and **TP-2** appeared to be generally nuclear excluded within HeLa cells, with both peptides appeared throughout the cells cytoplasm.

Synthetic methodology to mono-arylate aryl amidines was developed using the Chan-Evan-Lam reaction. Two sets of conditions were developed, one using stoichiometric $Cu(OAc)_2$ that could be carried out at room temperature within 2 h, and a second that used only 20 mol% $Cu(OAc)_2$ but required heating to 50 °C and a reaction time of 24 h. Mechanistic investigation identified that the amidine was the sole contributor to $Cu(OAc)_2$ denucleation. An amidine-Cu^{II} species was identified and found to not be active but in equilibrium with active Cu species in the main catalytic cycle. The Chan-Lam amidination was successfully applied to a variety of aryl amidines and boronic acids coupling partners, including some alkenyl boronic acids, though alkyl boronic acids proved unsuccessful in the reaction.

Further investigation into the scope of the reaction could be employed in the future through the use of alkyl amidines and also identification of deborination products from low yielding reactions. The reaction was also used to make a series of mono-arylated pentamidine analogues though each analogue did not show any therapeutic advantage over the free pentamidine. The low activity against trypanosomes could be attributed to the *N*-arylated amidine group reducing the compounds ability to bind to the adenosine uptake protein, which is essential for the uptake of pentamidine by trypanosoma.[281]

A series of peptidomimetic TP CPPs were prepared, including an aryl amidine peptidomimetic **TP-4**. All peptidomimetics were found to be relatively non-toxic to HeLa cells, even up to 100 μM concentration. The peptidomimetics identified that cationic charge, and not the guanidinium group, were required to maintain uptake of both TP CPPs. Replacement of the guanidine groups with aryl amidines in **TP-2** increased the amount of peptide within HeLa cells after 4 h, showing that **TP-4** is a reliable **TP-2** peptidomimetic. Further investigation found that **TP-4** was taken up *via* an energy-dependant mechanism, likely using a similar pathway used for **TP-2**. The sub cellular localisation of **TP-4** was also found to be similar to **TP-2** with general dispersion throughout the cytoplasm.

Overall low arginine CPPs, including arginine mimetics CPPs, generally had lower toxicity profiles than arginine rich CPPs. Furthermore, the use of amidine mimetics were able to improve the uptake of a low arginine containing CPP without changing its toxicity or cellular localisation profiles.

Due to the success of aryl amidine mimetics in the **TP-4** CPP and development of the Chan-Lam amidination, this work could provide the basis for the development of new delivery vehicles and therapeutically relevant *N*-aryl amidines. Further developments of this work could include (Fig. 5.1):

i) Application of **TP-4** to deliver of PMO cargos that have previously been delivered by CPPs such as **Pip6**.
ii) Investigation into the wider therapeutic application of monoaryl pentamidine analogues
iii) Synthesis of mono-arylated amidines by the Chan-Lam amidination for use in SPPS to synthesise more amidine containing CPPs.
iv) Use of aryl amidines as arginine mimetics in other peptide systems such as mitochondrial penetrating peptides, **5.2**.
v) Further synthesis of less basic arginine bioisosteres which could include α-keto guanidine **5.3** and α-keto amidine **5.4**.

Specifically, the series of pentamidine analogues that showed relatively low activity against trypanosoma brucei brucei (*chapter 3*) could be further investigated for their effectiveness against leishmania trypanosomes, as pentamidine resistance in leishmania was reported as early as 2002.[282] Furthermore pentamidine has been reported to reverse RNA splicing defects associated with myotonic

dystrophy.[283] Monoarylated analogues could be used to investigate the reversal of this splice switching and the compounds would likely not suffer from the uptake issue that gave poor EC_{50} results in the trypanosoma.

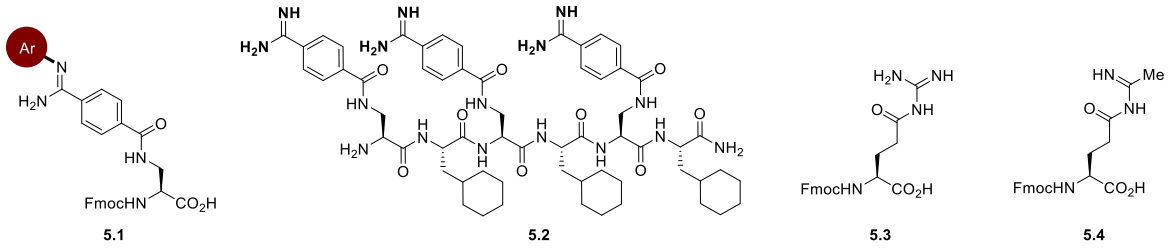
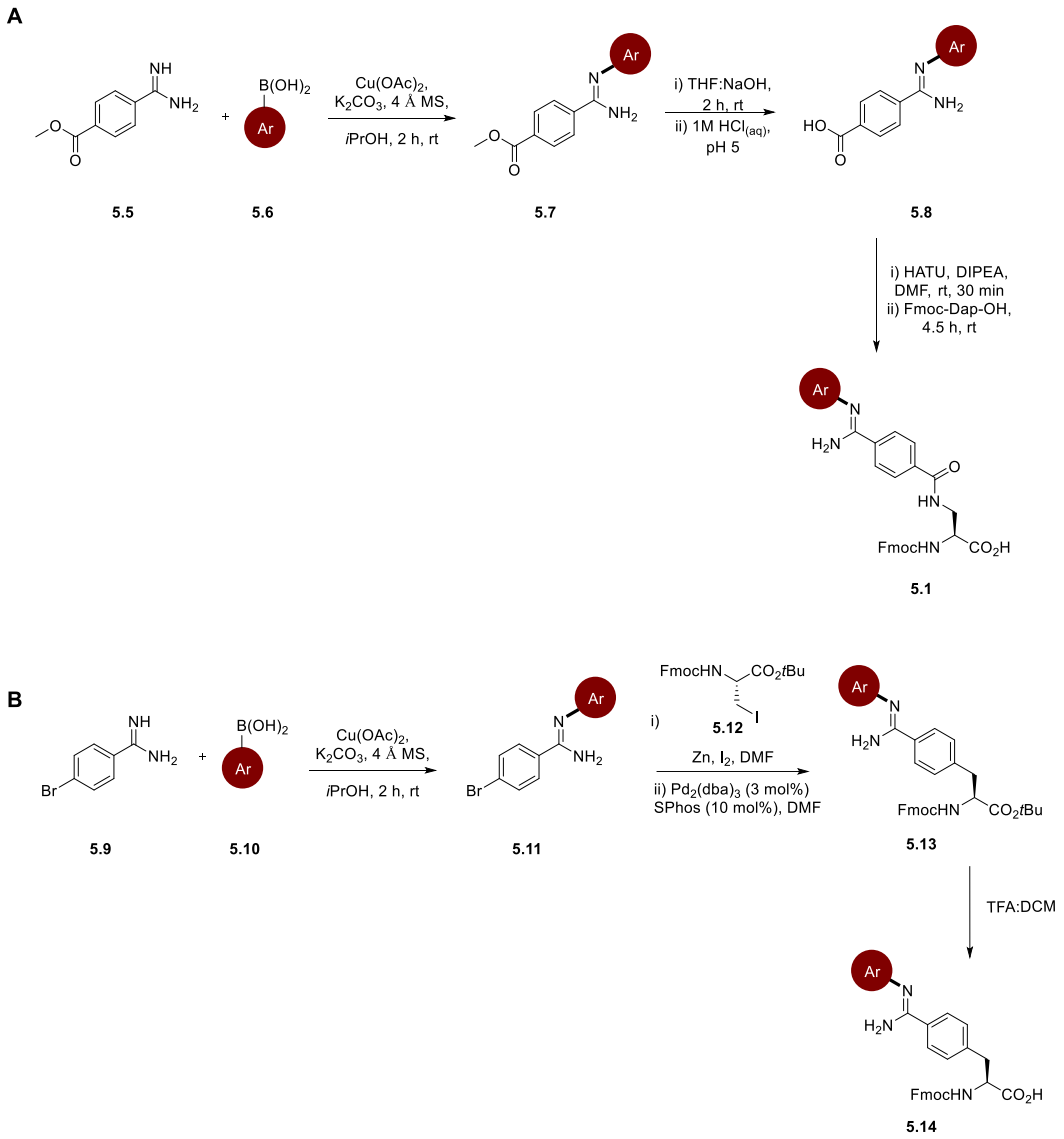


Figure 5.1. Potential future compounds that could be used as bioisosteres and peptidomimetics. Mono-aryl amidine **5.1**, mitochondrial penetrating peptide mimetic **5.2** and α- keto mimetics **5.3** and **5.4**.

The *N*-arylation of amidines will allow for the modulation of lipophilicity of amidine containing peptides. To synthesise building block **5.1** a similar synthesis can be employed as discussed in *chapter 4* but employing the Chan-Lam amidination reaction in place of a Pbf protecting step (Scheme 5.1A). In addition to this synthesis Negishi couplings could be utilised to synthesise aryl amidine amino acids, similar to previous reports for the synthesis of aryl amino acids (Scheme 5.1B).[284] Negishi couplings have been reported by Cobb *et al.* for the synthesis of fmoc amino acid building blocks that bear an aryl group on the side chain.[285] To undertake this synthesis the Chan-Lam amidination could be employed on an aryl halide amidine **5.9** to afford **5.11** before a Negishi coupling using iodo amino acid **5.12** could yield the amidine amino acid **5.13**. A subsequent acidic deprotection of **5.13** would give rise to building block **5.14** that could be utilised in SPPS. The Negishi approach allows for a carbon-carbon linkage at the side chain of the amino acid and therefore a more flexible side chain than the amide in **5.1**.

Finally, the investigation of the serum stability of all amidine containing peptides could be carried out. Enhanced stability to proteases could be a significant factor in the increased uptake observed with **TP-4** against its natural arginine analogue **TP-2** in *chapter 4*.



Scheme 5.1. (A) Amide coupling route to *N*-aryl amidine amino acid **5.1**. (B) Negishi coupling route to *N*-aryl amidine amino acid **5.14**.

The use of amidines as arginine mimetics in CPPs and the wider peptide field could yield novel peptides with advantageous protease stability and toxicity profiles. New aryl amidine amino acids could be synthesised with relative ease using Chan-Lam amidinations and previously reported Negishi chemistry. Due to the importance of the arginine residue in CPPs, further development of new arginine mimetics alongside amidines could open up a variety of opportunities in the peptide drug delivery field.

Chapter 6: References

(1) Jewell, C. M.; Jung, J. M.; Atukorale, P. U.; Carney, R. P.; Stellacci, F.; Irvine, D. J. *Angew. Chemie - Int. Ed.* **2011**, *50* (51), 12312–12315.

(2) Patra, J. K.; Das, G.; Fraceto, L. F.; Campos, E. V. R.; Rodriguez-Torres, M. D. P.; Acosta-Torres, L. S.; Diaz-Torres, L. A.; Grillo, R.; Swamy, M. K.; Sharma, S.; Habtemariam, S.; Shin, H. S. *J. Nanobiotechnology* **2018**, *16* (1),1–33..

(3) Lau, J. L.; Dunn, M. K. *Bioorganic Med. Chem.* **2018**, *26* (10), 2700–2707.

(4) Goodchild, J. *Methods Mol. Biol.* **2011**, *764*, 1–15.

(5) S. J. Singer and Garth L. Nicolson. *Science (80).* **1972**, *175*, 720–731.

(6) Morth, J. P.; Pedersen, B. P.; Toustrup-Jensen, M. S.; Sørensen, T. L.-M.; Petersen, J.; Andersen, J. P.; Vilsen, B.; Nissen, P. C. *Nature* **2007**, *450* (7172), 1043–1049.

(7) Schmidt, M.; Oude Weernink, P. A.; vom Dorp, F.; Stope, M. B.; Jakobs, K. H. *Adv. Mol. Cell Biol.* **2004**, *33*, 431–450.

(8) Kjeuin, L.; Lindahl, U. Proteoglycans: Structure and Interactions. *Annu. Rev. Biochem.* **1991**, *60*, 443–475.

(9) Gahmberg, C. G. Elsevier/North-Holland Biomedical Press, 1981; Vol. 1.

(10) Mikami, T.; Kitagawa, H. *Biochim. Biophys. Acta - Gen. Subj.* **2013**, *1830* (10), 4719–4733.

(11) Gillard, B. K.; Thurmon, L. T.; Marcus, D. M. *Glycobiology* **1993**, *3* (1), 57–67.

(12) Boggs, J. M.; Menikh, A.; Rangaraj, G. *Biophys. J.* **2000**, *78* (2), 874–885.

(13) Yang, S.-T.; Kreutzberger, A. J. B.; Lee, J.; Kiessling, V.; Tamm, L. K. *Chem. Phys. Lipids* **2016**, *199*, 136–143.

(14) Assemblies, B. G.; Yellin, N.; Levin, I. W. *Biochemistry* **1972**, *16* (4), 642–647.

(15) Harvey Lodish, Arnold Berk, S Lawrence Zipursky, Paul Matsudaira, David Baltimore and J. D. *Molecular Cell Biology, 4th Edition*, 4th ed.; W.H Freeman and Co Ltd: New York, 2000.

(16) Borbás, E.; Sinkó, B.; Tsinman, O.; Tsinman, K.; Kiserdei, É.; Démuth, B.; Balogh, A.; Bodák, B.; Domokos, A.; Dargó, G.; Balogh, G. T.; Nagy, Z. K. *Mol. Pharm.* **2016**, *13* (11), 3816–3826.

(17) Alberts B, Johnson A, L. J. *Molecular Biology of the Cell*, 4th ed.; Garland Science: New York, 2002.

(18) Mueckler, M.; Caruso, C.; Baldwin, S. a; Panico, M.; Blench, I.; Morris, H. R.; Allard, W. J.; Lienhard, G. E.; Lodish, H. F. *Science* **1985**, *229* (4717), 941–945.

(19) David L Nelson, M. M. C. *Lehninger Principles of Biochemistry*, 4th ed.; W.H Freeman and Co Ltd: New York, 2005.

(20) Palmgren, M. G.; Nissen, P. P-Type ATPases. *Annu. Rev. Biophys.* **2011**, *40* (1), 243–266.

(21) Cooper, G. M.; Hausman, R. E. *The Cell A Molecular Approach*, 4th ed.; ASM Press: Washington D.C., 2007.

(22) Lodish, H.; Berk, A. *Molecular Cell Biology*, 5th ed.; W.H Freeman and Co Ltd: New York, 2003.

(23) Pelkmans, L.; Bürli, T.; Zerial, M.; Helenius, A. *Cell* **2004**, *118* (6), 767–780..

(24) Hansen, C. G.; Nichols, B. J. *Trends Cell Biol.* **2010**, *20* (4), 177–186.

(25) Pelkmans, L.; Kartenbeck, J.; Helenius, A. *Nat. Cell Biol.* **2001**, *3* (5), 473–483.

(26) Lim, J. P.; Gleeson, P. A. *Immunol. Cell Biol.* **2011**, *89* (8), 836–843. https://doi.org/10.1038/icb.2011.20.

(27) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25.

(28) Britt, H. M.; García-Herrero, C. A.; Denny, P. W.; Mosely, J. A.; Sanderson, J. M. *Chem. Sci.* **2019**, *10* (3), 674–680.

(29) Vaz-da-Silva, M.; Loureiro, A. I.; Nunes, T.; Maia, J.; Tavares, S.; Falcão, A.; Silveira, P.; Almeida, L.; Soares-da-Silva, P. B. *Clin. Drug Investig.* **2005**, *25* (6), 391–399..

(30) Richard L. Momparber, Myron Karon, Stuart E. Siegel. *Cancer Res.* **1976**, *36*, 2891–2895.

(31) Satoshi Matsuda, S. K. *Immunopharmacology* **2000**, *47*, 119–125.

(32) Vert, M.; Hellwich, K.-H.; Hodge, P.; Schué, F.; Kubisa, P.; Rinaudo, M.; Hess, M.; Doi, Y. *Pure Appl. Chem.* **2012**, *84* (2), 377–410.

(33) Widder, k. J.; Flouret, G.; Senye, A. *J. Pharm. Sci.* **1979**, *68* (1), 79–82.

(34) Gupta, U.; Sharma, S.; Khan, I.; Gothwal, A.; Sharma, A. K.; Singh, Y.; Chourasia, M. K.; Kumar, V. *Int. J. Biol. Macromol.* **2017**, *98*, 810–819.

(35) Nakamura, H.; Watano, S. *KONA Powder Part. J.* **2018**, *2018* (35), 49–65.

(36) Alavi, M.; Karimi, N.; Safaei, M.. *Adv. Pharm. Bull.* **2017**, *7* (1), 3–9.

(37) Malam, Y.; Loizidou, M.; Seifalian, A. M. *Trends Pharmacol. Sci.* **2009**, *30* (11), 592–599.

(38) Bulbake, U.; Doppalapudi, S.; Kommineni, N.; Khan, W. *Pharmaceutics* **2017**, *9* (2), 1–33.

(39) Düzgüneş, N.; Nir, S.. *Adv. Drug Deliv. Rev.* **1999**, *40* (1–2), 3–18.

(40) Bechara, C.; Sagan, S. *FEBS Lett.* **2013**, *587* (12), 1693–1702.

(41) Frankel, A. D.; Pabo, C. O. *Cell* **1988**, *55* (6), 1189–1193.

(42) Eric Vive`s, Priscille Brodin, and B. L. *J. Biol. Chem.* **1997**, *272* (25), 16010–16017.

(43) Elmquist, A.; Lindgren, M.; Bartfai, T.; Langel U; Langel, Ü.; Langel, *Exp. Cell Res.* **2001**, *269* (2), 237–244.

(44) Bayer, P.; Kraft, M.; Ejchart, A.; Westendorp, M.; Frank, R.; Rösch, P. *J. Mol. Biol.* **1995**, *247* (4), 529–535.

(45) Fosgerau, K.; Hoffmann, T. *Drug Discov. Today* **2015**, *20* (1), 122–128.

(46) Crosignani, P. G.; Luciano, A.; Ray, A.; Bergqvist, A *Hum. Reprod.* **2006**, *21* (1), 248–256.

(47) Vagner, J.; Qu, H.; Hruby, V. J. *Opin. Chem. Biol.* **2008**, *12* (3), 292–296..

(48) Meanwell, N. A. *J. Med. Chem.* **2011**, *54* (8), 2529–2591.

(49) Hans-Joachim Bohm, David Banner, Stefanie Bendels, Manfred Kansy, Bernd Kuhn, Klaus M¸ller, Ulrike Obst-Sander, and M. S. *ChemBioChem* **2004**, *5*, 637–643.

(50) Diana, G. D.; Rudewicz, P.; Pevear, D. C.; Nitz, T. J.; Aldous, S. C.; Aldous, D. J.; Robinson, D. T.; Draper, T.; Dutko, F. J.; Aldi, C.; Gendron, G.; Oglesby, R. C.; Volkots, D. L.; Reuman, M.; Bailey, T. R.; Czemiak, R.; Block, T.; Roland, R.; Oppermann, J. *J. Med. Chem.* **1995**, *38* (8), 1355–1371.

(51) Lassalas, P.; Gay, B.; Lasfargeas, C.; James, M. J.; Tran, V.; Vijayendran, K. G.; Brunden, K. R.; Kozlowski, M. C.; Thomas, C. J.; Smith, A. B.; Huryn, D. M.; Ballatore, C. *J. Med. Chem.* **2016**, *59* (7), 3183–3203.

(52) Carini, D. J.; Christ, D. D.; Duncia, J. V.; Pierce, M. E.. *Pharm. Biotechnol.* **1998**, *11*, 29–56.

(53) Morrison, K. L.; Weiss, G. A. *Curr. Opin. Chem. Biol.* **2001**, *5* (3), 302–307.

(54) Choudhary, A.; Raines, R. T. *ChemBioChem* **2011**, *12* (12), 1801–1807.

(55) Isabella L. Karle and Padmanabhan Balaran. *Biochemistry*, **1990**, *29* (29), 6747–6756.

(56) Formaggio, F.; Moretto, V.; Crisma, M.; Toniolo, C.; Kaptein, B.; Broxterman, Q. B. *J. Pept. Res.* **2004**, *63* (2), 161–170.

(57) Goodman, M.; Shao, H. *Pure Appl. Chem.* **1996**, *68* (6), 1303–1308.

(58) Madhu, C.; Voshavar, C.; Rajasekhar, K.; Govindaraju, T. *Org. Biomol. Chem.* **2017**, *15*, 3170–3174.

(59) Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. N-Methylation of Peptides:. *Acc. Chem. Res.* **2008**, *41* (10), 1331–1342.

(60) Chingle, R.; Proulx, C.; Lubell, W. D. *Acc. Chem. Res.* **2017**, *50* (7), 1541–1556.

(61) Tal-Gan, Y.; Freeman, N. S.; Klein, S.; Levitzki, A.; Gilon, C. *Chem. Biol. Drug Des.* **2011**, *78* (5), 887–892.

(62) Derossi, D.; Joliot, A. H.; Chassaing, G.; Prochiantz, A. *J. Biol. Chem.* **1994**, *269* (14), 10444–10450.

(63) Josephson, L.; Tung, C.; Moore, A.; Weissleder, R. *Bioconjug. Chem.* **1999**, *10* (2), 186–191.

(64) Astriab-Fisher, A.; Sergueev, D.; Fisher, M.; Ramsay Shaw, B.; Juliano, R. L.. *Pharm. Res.* **2002**, *19* (6), 744–754.

(65) Oehlke, J.; Scheller, A.; Wiesner, B.; Krause, E.; Beyermann, M.; Klauschenz, E.; Melzig, M.; Bienert, M. *Biochim. Biophys. Acta - Biomembr.* **1998**, *1414* (1–2), 127–139.

(66) Futaki, S.; Suzuki, T.; Ohashi, W.; Yagami, T.; Tanaka, S.; Ueda, K.; Sugiura, Y. *J. Biol. Chem.* **2001**, *276* (8), 5836–5840.

(67) Midoux, P.; Kichler, A.; Maurizot, J.; Monsigny, M. *Bioconjug. Chem.* **1998**, *9* (2), 260–267.

(68) Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (24), 13003–13008.

(69) Fischer, P. M.; Zhelev, N. Z.; Wang, S.; Melville, J. E.; Fåhraeus, R.; Lane, D. P. *J. Pept. Res.* **2000**, *55* (2), 163–172.

(70) Mitchell, D. J.; Kim, D. T.; Steinman, L.; Fathman, C. G.; Rothbard, J. B. P. *J. Pept. Res.* **2000**, *56* (5), 318–325.

(71) Rothbard, J. B.; Jessop, T. C.; Lewis, R. S.; Murray, B. A.; Wendner, P. A. Peptides into Cells. *J. Am. Chem. Soc.* **2004**, *126* (31), 9506–9507.

(72) Elmquist, A.; Hansen, M.; Langel, U. *Biochim. Biophys. Acta* **2006**, *1758* (6), 721–729.

(73) Rydberg, H. a; Matson, M.; Åmand, H. L.; Esbjo, E. K.; Norde, B. *Biochemistry* **2012**, *51* (27), 5531–5539.

(74) Bechara, C.; Pallerla, M.; Zaltsman, Y.; Burlina, F.; Alves, I. D.; Lequin, O.; Sagan, S.. *FASEB J.* **2013**, *27* (2), 738–749.

(75) Rothbard, J. B.; Kreider, E.; Pattabiraman, K.; Pelkey, E. T.; VanDeusen, C. L.; Wright, L.; Wylie, B. L.; Wender, P. A. *J. Med. Chem.* **2002**, *45*, 141–160.

(76) Betts, C.; Saleh, A. F.; Arzumanov, A. a; Hammond, S. M.; Godfrey, C.; Coursindel, T.; Gait, M. J.; Wood, M. J. *Mol. Ther. — Nucleic Acids* **2012**, *1* (8), e38.

(77) Ivanova, G. D.; Arzumanov, A.; Abes, R.; Yin, H.; Wood, M. J. A.; Lebleu, B.; Gait, M. J. *Nucleic Acids Res.* **2008**, *36* (20), 6418–6428.

(78) O'Donovan, L.; Okamoto, I.; Arzumanov, A. a; Williams, D. L.; Deuss, P.; Gait, M. J. Parallel *Nucleic Acid Ther.* **2014**, *00* (0), 1–10.

(79) Marks, J. R.; Placone, J.; Hristova, K.; Wimley, W. C. S. *J. Am. Chem. Soc.* **2011**, *133* (23), 8995–9004.

(80) Schmidt, S.; Adjobo-Hermans, M. J. W.; Kohze, R.; Enderle, T.; Brock, R.; Milletti, F. *Bioconjug. Chem.* **2017**, *28*, 382–389.

(81) Macchi, S.; Signore, G.; Boccardi, C.; Di Rienzo, C.; Beltram, F.; Cardarelli, F. *Sci. Rep.* **2015**, *5*, 16914.

(82) Guidotti, G.; Brambilla, L.; Rossi, D. *Trends Pharmacol. Sci.* **2017**, *38* (4), 406–424.

(83) Jearawiriyapaisarn, N.; Moulton, H. M.; Buckley, B.; Roberts, J.; Sazani, P.; Fucharoen, S.; Iversen, P. L.; Kole, R. *Mol. Ther.* **2008**, *16* (9), 1624–1629.

(84) Deloche, C.; Lopez-Lazaro, L.; Mouz, S.; Perino, J.; Abadie, C.; Combette, J.-M. XG-102. *Pharmacol. Res. Perspect.* **2014**, *2* (1), e00020.

(85) Cai, S.-R.; Xu, G.; Becker-Hapak, M.; Ma, M.; Dowdy, S. F.; McLeod, H. L. *Eur. J. Pharm. Sci.* **2006**, *27* (4), 311–319.

(86) Lee, M. S.; Kwon, E. H.; Choi, H. S.; Kwon, S. H.; Lee, C. H.; Shim, I. sop; Lee, S. K.; Her, S. *Biochem. Biophys. Res. Commun.* **2010**, *394* (2), 348–353.

(87) Amantana, A.; Moulton, H. M.; Cate, M. L.; Reddy, M. T.; Whitehead, T.; Hassinger, J. N.; Youngblood, D. S.; Iversen, P. L. *Bioconjug. Chem.* **2007**, *18* (4), 1325–1331.

(88) Umemoto, T.; Sakamoto, K.; Fukuda, Y.; Adachi, Y.; Tani, A.; Asami, T. A. *Chem. Lett.* **2017**, *46* (6), 889–891.

(89) Li, M.; Mosel, S.; Knauer, S. K.; Schmuck, C. A. *Org. Biomol. Chem.* **2018**, *16* (13), 2312–2317.

(90) Nischan, N.; Herce, H. D.; Natale, F.; Bohlke, N.; Budisa, N.; Cardoso, M. C.; Hackenberger, C. P. R. *Angew. Chemie - Int. Ed.* **2015**, *54* (6), 1950–1953.

(91) Herce, H. D.; Schumacher, D.; Schneider, A. F. L.; Ludwig, A. K.; Mann, F. A.; Fillies, M.; Kasper, M. A.; Reinke, S.; Krause, E.; Leonhardt, H.; Cardoso, M. C.; Hackenberger, C. P. R.. *Nat. Chem.* **2017**, *9* (8), 762–771.

(92) Qian, Z.; Larochelle, J. R.; Jiang, B.; Lian, W.; Hard, R. L.; Selner, N. G.; Luechapanichkul, R.; Barrios, A. M.; Pei, D. *Biochemistry* **2014**, *53* (24), 4034–4046.

(93) Qian, Z.; Martyna, A.; Hard, R. L.; Wang, J.; Appiah-Kubi, G.; Coss, C.; Phelps, M. A.; Rossman, J. S.; Pei, D. *Biochemistry* **2016**, *55* (18), 2601–2612.

(94) Hedstrom, L. *Chem. Rev.* **2002**, *102* (12), 4501–4523.

(95) Olsen, J. V.; Ong, S. E.; Mann, M.. *Mol. Cell. Proteomics* **2004**, *3* (6), 608–614.

(96) Leiros, H. K. S.; Brandsdal, B. O.; Andersen, O. A.; Os, V. Tr . *Protein Sci.* **2004**, 1056–1070.

(97) Deprey, K.; Becker, L.; Kritzer, J.; Plückthun, A. *Bioconjug. Chem.* **2019**, *30* (4), 1006–1027.

(98) Maiolo, J. R.; Ottinger, E. A.; Ferrer, M. *J. Am. Chem. Soc.* **2004**, *126* (47), 15376–15377.

(99) LeCher, J. C.; Nowak, S. J.; McMurry, J. L. *Biomol. Concepts* **2017**, *8* (3–4), 131–141.

(100) Allen, J. K.; Brock, D. J.; Kondow-McConaghy, H. M.; Pellois, J. P. *Biomolecules* **2018**, *8* (3).

(101) Tunnemann Gisela, Ter-Avetisyan Gohar, Robert M. Martin, Stockl Martin, A. H. and C. M. C. *J. Pept. Sci.* **2008**, *14*, 469–476.

(102) Meng, X.; Li, T.; Zhao, Y.; Wu, C. *ACS Chem. Biol.* **2018**.

(103) Jones, S. W.; Christison, R.; Bundell, K.; Voyce, C. J.; Brockbank, S. M. V.; Newham, P.; Lindsay, M. A. *Br. J. Pharmacol.* **2005**, *145* (8), 1093–1102.

(104) Smith, B. A.; Daniels, D. S.; Coplin, A. E.; Jordan, G. E.; McGregor, L. M.; Schepartz, A. *J. Am. Chem. Soc.* **2008**, *130* (10), 2948–2949.

(105) Gao, W.; Yang, X.; Lin, Z.; He, B.; Mei, D.; Wang, D.; Zhang, H.; Zhang, H.; Dai, W.; Wang, X.; Zhang, Q. *J. Control. Release* **2017**, *261*, 174–186..

(106) Martín, I.; Teixidó, M.; Giralt, E. *Building Cell Selectivity into CPP-Mediated Strategies*; 2010; Vol. 3.

(107) Warburton, W. *J. Chem. Soc* **1951**, 2492–2494.

(108) Göbel, M.; Klapötke, T. M. *Chem. Commun.* **2007**, *13* (30), 3180–3182..

(109) Trabocchi, A. *Peptidomimetics in Organic and Medicinal Chemistry: The Art ofTransforming Peptides in Drugs, First Edition.*; 2014; pp 37–73.

(110) Kennedy, K. J.; Simandan, T. L.; Dix, T. A. A. *Synth. Commun.* **1998**, *28* (4), 741–746.

(111) Ghorai, P.; Kraus, A.; Keller, M.; Götte, C.; Igel, P.; Schneider, E.; Schnell, D.; Bernhardt, G.; Dove, S.; Zabel, M.; Elz, S.; Seifert, R.; Buschauer, A. *J. Med. Chem.* **2008**, *51* (22), 7193–7204.

(112) Kraus, A.; Ghorai, P.; Birnkammer, T.; Schnell, D.; Elz, S.; Seifert, R.; Dove, S.; Bernhardt, G.; Buschauer, A. *ChemMedChem* **2009**, *4* (2), 232–240.

(113) Lee, Y.; Marletta, M. A.; Martasek, P.; Roman, L. J.; Masters, B. S. S.; Silverman, R. B. *Bioorganic Med. Chem.* **1999**, *7* (6), 1097–1104.

(114) Clayden, J. Acidity, Basicity and PKa. In *Organic Chemistry*; 2001; pp 181–224.

(115) Peterlin Masic, L. *Curr. Med. Chem.* **2006**, *13* (30), 3627–3648.

(116) Van Ryn, J.; Stangier, J.; Haertter, S.; Liesenfeld, K. H.; Wienen, W.; Feuring, M.; Clemens, A. *Thromb. Haemost.* **2010**, *103* (6), 1116–1127.

(117) Kotthaus, J.; Steinmetzer, T.; Van De Locht, A.; Clement, B. *J. Enzyme Inhib. Med. Chem.* **2011**, *26* (1), 115–122.

(118) Barrett, M. P.; Gemmell, C. G.; Suckling, C. J. *Pharmacol. Ther.* **2013**, *139* (1), 12–23.

(119) Palomo, J. M. *RSC Adv.* **2014**, *4* (62), 32658–32672.

(120) Jaradat, D. M. M. *Amino Acids* **2018**, *50* (1), 39–68.

(121) Curtius, T. Ueber. *J Prakt Chem* **1882**, *26*, 145–208.

(122) Fischer, E.; Forneau, E. *Berichte der Dtsch. Chem. Gesellschaft* **1901**, *34*, 2868–2877.

(123) Fischer, E. *Berichte der Dtsch. Chem. Gesellschaft* **1905**, *38*, 605–619.

(124) Fischer, E. *Berichte der Dtsch. Chem. Gesellschaft* **1907**, *40* (2), 1754–1767.

(125) Bergmann, M.; Zervas, L. *Berichte der Dtsch. Chem. Gesellschaft* **1932**, *65*, 1192–1201.

(126) Du Vigneaud, V.; Ressler, C.; Swan, J. M.; Roberts, C. W.; Katsoyannis, P. G.; Gordon, S. *J. Am. Chem. Soc.* **1953**, *75* (19), 4879–4880.

(127) McKay, F. C.; Albertson, N. F. *J. Am. Chem. Soc.* **1957**, *79* (17), 4686–4690.

(128) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149–2154.

(129) Merrifield, R. B. *Biochemistry* **1964**, *3* (9), 1385–1390.

(130) Carpino, L. A.; Han, G. Y. *J. Am. Chem. Soc.* **1970**, *92* (19), 5748–5749.

(131) Carpino, L. A.; Han, G. Y. *J. Org. Chem.* **1972**, *37* (22), 3404–3409.

(132) Atherton, E.; Fox, H.; Harkiss, D.; Logan, C. J.; Sheppard, R. C.; Williams, B. J. *J. Chem. Soc. Chem. Commun.* **1978**, No. 13, 537–539.

(133) Atherton, E.; Logan, C. J.; Sheppard, R. C. *J Chem Soc Perkin Trans 1* **1981**, *1*, 538–546.

(134) Kent, S. B. H. *Chemical Synthesis of Peptides and Proteins*; 1988; Vol. 57.

(135) Hudson, D. *J. Comb. Chem.* **1999**, *1*, 333–360.

(136) Merrifield, B. *Br. Polym. J.* **1984**, *16* (4), 173–178.

(137) Williams, H. *J. Org. Chem.* **1964**, *29*, 2046–2047.

(138) Wang, S.-S. *J. Am. Chem. Soc.* **1973**, *95* (4), 1328–1333.

(139) King, D. S.; Fields, C. G.; Fields, G. B. *Int. J. Pept. Protein Res.* **1990**, *36* (3), 255–266.

(140) Rink, H. *Tetrahedron Lett.* **1987**, *28* (33), 3787–3790.

(141) Albericio, F.; Kneib-cordonier, N.; Biancalana, S.; Gera, L.; Masada, R. I.; Hudson, D.; Barany, G. *J. Org. Chem.* **1990**, *55*, 3730–3743.

(142) Merrifield, R. B.; Gisin, B. F.; Bach, A. N. *J. Org. Chem.* **1977**, *42*, 1291–1295.

(143) Detar, D. F.; Silverstein, R. *J. Am. Chem. Soc.* **1966**, *88* (5), 1013–1019.

(144) Sheehan, J. C.; Hess, G. P. *J. Am. Chem. Soc.* **1955**, *77* (4), 1067–1068.

(145) Sheehan, J. C.; Boshart, G. L.; Cruickshank, P. A. *J. Org. Chem.* **1961**, *26* (7), 2525–2528.

(146) Benoiton, N. L.; Chen, M. F. *J. Chem. Soc. Chem. Commun.* **1981**, *11*, 543–545.

(147) El-Faham, A.; Albericio, F. *Chem. Rev.* **2011**, *111*, 6557–6602.

(148) Goodman, M.; Stueben, K. C. *J. Org. Chem.* **1962**, *27* (10), 3409–3416.

(149) Pedersen, S. L.; Tofteng, A. P.; Malik, L.; Jensen, K. J.. *Chem. Soc. Rev.* **2012**, *41* (5), 1826–1844.

(150) König, W.; Geiger, R. *Chem. Ber.* **1970**, *103* (3), 788–798.

(151) Carpino, L. A. *J. Am. Chem. Soc.* **1993**, *115* (10), 4397–4398.

(152) Gawne, G.; Kenner, G. W.; Sheppard, R. C. *J. Am. Chem. Soc.* **1969**, *91* (20), 5669–5671.

(153) Castro, B.; Dormoy, J. R.; Evin, G.; Selve, C. *Tetrahedron Lett.* **1975**, *16* (14), 1219–1222.

(154) Coste, J.; Le-Nguyen, D.; Castro, B. *Tetrahedron Lett.* **1990**, *31* (2), 205–208.

(155) Albericio, F.; Cases, M.; Alsina, J.; Triolo, S. A.; Carpino, L. A.; Kates, S. A. *Tetrahedron Lett.* **1997**, *38* (27), 4853–4856.

(156) Dourtoglou, V.; Ziegler, J. C.; Gross, B. *Tetrahedron Lett.* **1978**, *19* (15), 1269–1272.

(157) Abdelmoty, I.; Albericio, F.; Carpino, L. A.; Foxman, B. M.; Kates, S. A. *Lett. Pept. Sci.* **1994**, *1* (2), 57–67.

(158) Hood, C. A.; Fuentes, G.; Patel, H.; Page, K.; Menakuru, M.; Jae H., P. F. *J. Pept. Sci.* **2008**, *14*, 97–101.

(159) Albericio, F.; Bofill, J. M.; El-Faham, A.; Kates, S. A.. *J. Org. Chem.* **1998**, *63* (26), 9678–9683.

(160) Hudson, D.. *J. Org. Chem.* **1988**, *53* (3), 617–624.

(161) Carpino, L. A.; Imazumi, H.; El-Faham, A.; Ferrer, F. J.; Zhang, C.; Lee, Y.; Foxman, B. M.; Henklein, P.; Hanay, C.; Mügge, C.; Wenschuh, H.; Klose, J.; Beyermann, M.; Bienert, M. *Angew. Chemie - Int. Ed.* **2002**, *41* (3), 441–445.

(162) Kansy, M.; Senner, F.; Gubernator, K.. *J. Med. Chem.* **1998**, *41* (7), 1007–1010.

(163) Pye, C. R.; Hewitt, W. M.; Schwochert, J.; Haddad, T. D.; Townsend, C. E.; Etienne, L.; Lao, Y.; Limberakis, C.; Furukawa, A.; Mathiowetz, A. M.; Price, D. A.; Liras, S.; Lokey, R. S. *J. Med. Chem.* **2017**, *60* (5), 1665–1672.

(164) Hewitt, W. M.; Leung, S. S. F.; Pye, C. R.; Ponkey, A. R.; Bednarek, M.; Jacobson, M. P.; Lokey, R. S. *J. Am. Chem. Soc.* **2015**, *137* (2), 715–721.

(165) Faller, B. *Curr. Drug Metab.* **2008**, *9* (9), 886–892.

(166) Wheaten, S. A.; Ablan, F. D. O.; Spaller, B. L.; Trieu, J. M.; Almeida, P. F. *J. Am. Chem. Soc.* **2013**, *135* (44), 16517–16525.

(167) Qian, Z.; Liu, T.; Liu, Y. Y.; Briesewitz, R.; Barrios, A. M.; Jhiang, S. M.; Pei, D. *ACS Chem. Biol.* **2013**, *8* (2), 423–431.

(168) Cho, M. J.; Thompson, D. P.; Cramer, C. T.; Vidmar, T. J.; Scieszka, J. F. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*. 1989, pp 71–77.

(169) Artursson, P.; Karlsson, J.. *Biochem. Biophys. Res. Commun.* **1991**, *175*, 880–885.

(170) Stevenson, C. L.; Augustijns, P. F.; Hendren, R. W. *Int. J. Pharm.* **1999**, *177* (1), 103–115.

(171) Gobom, J.; Kraeuter, K. O.; Persson, R.; Steen, H.; Roepstorff, P.; Ekman, R. *Anal. Chem.* **2000**, *72* (14), 3320–3326.

(172) Burlina, F.; Sagan, S.; Bolbach, G.; Chassaing, G. *Angew. Chemie - Int. Ed.* **2005**, *44* (27), 4244–4247.

(173) Picot, J.; Guerin, C. L.; Le Van Kim, C.; Boulanger, C. M. *Cytotechnology* **2012**, *64*, 109–130.

(174) Givan, A. L. *Methods Mol. Biol. 699*, 1–29.

(175) Li, Q.; Xu, M.; Cui, Y.; Huang, C.; Sun, M. *Pharmacol. Res. Perspect.* **2017**, *5* (5), 1–5.

(176) Cruz, J.; Mihailescu, M.; Wiedman, G.; Herman, K.; Searson, P. C.; Wimley, W. C.; Hristova, K. *Biophys. J.* **2013**, *104* (11), 2419–2428.

(177) Fischer, R.; Waizenegger, T.; Köhler, K.; Brock, R. A. *Biochim. Biophys. Acta - Biomembr.*

2002, *1564* (2), 365–374.

(178) Rampersad, S. N. *Sensors* **2012**, *12* (9), 12347–12360.

(179) Dutta, D.; Williamson, C. D.; Cole, N. B. *PLoS One* **2012**, *7*, 1–9.

(180) Zhu, X. D.; Zhuang, Y.; Ben, J. J.; Qian, L. L.; Huang, H. P.; Bai, H.; Sha, J. H.; He, Z. G.; Chen, Q. *J. Biol. Chem.* **2011**, *286*, 8231–8239.

(181) Christoforidis, S.; Miaczynska, M.; Ashman, K.; Wilm, M.; Zhao, L.; Yip, S. C.; Waterfield, M. D.; Backer, J. M.; Zerial, M.. *Nat. Cell Biol.* **1999**, *1*, 249–252.

(182) Yoo, D. Y.; Barros, S. A.; Brown, G. C.; Rabot, C.; Bar-Sagi, D.; Arora, P. S. *J. Am. Chem. Soc.* **2020**, *142*, 14461–14471.

(183) Vercauteren, D.; Vandenbroucke, R. E.; Jones, A. T.; Rejman, J.; Demeester, J.; De Smedt, S. C.; Sanders, N. N.; Braeckmans, K. *Mol. Ther.* **2010**, *18*, 561–569.

(184) Chen, Y.; Wang, S.; Lu, X.; Zhang, H.; Fu, Y.; Luo, Y. *Blood* **2011**, *117*, 6392–6403.

(185) Weisblum, B.; Haenssler, E. *Chromosoma* **1974**, *46*, 255–260.

(186) Peraro, L.; Kritzer, J. A. *Angew. Chemie - Int. Ed.* **2018**, *57* (37), 11868–11881.

(187) Kim, S. M.; Chae, M. K.; Lee, C.; Yim, M. S.; Bang, J. K.; Ryu, E. K. *Amino Acids* **2014**, *46* (11), 2595–2603.

(188) Janez Ilas, Tihomir Tomasic and D. K. *J. Med. Chem.* **2008**, *51*, 2863–2867.

(189) Renton, P.; Green, B.; Maddaford, S.; Rakhit, S.; Andrews, J. S. *ACS Med. Chem. Lett.* **2012**, *3* (3), 227–231.

(190) Liu, X.; Fu, H.; Jiang, Y.; Zhao, Y. *Angew. Chemie - Int. Ed.* **2009**, *48* (2), 348–351.

(191) Brasche, G.; Buchwald, S. L. *Angew. Chemie Int. Ed.* **2008**, *47*, 1932–1934.

(192) Schaefer, F. C.; Peters, G. A. *J. Org. Chem.* **1961**, *26* (2), 412–418.

(193) Pfaff, D.; Nemecek, G.; Podlech, J. *Beilstein J. Org. Chem.* **2013**, *9*, 1572–1577.

(194) Hartwig, J. F.. *Nature* **2008**, *455* (7211), 314–322.

(195) Bariwal, J.; Van Der Eycken, E. *Chem. Soc. Rev.* **2013**, *42* (24), 9283–9303.

(196) Migita, T. *Chem. Lett.* **1983**, 927–928.

(197) Dale Boger, L.; Panek, J. S. *Tetrahedron Lett.* **1984**, *25* (30), 3175–3178.

(198) Paul, F.; Patt, J.; Hartwig, J. F. *J. Am. Chem. Soc.* **1994**, *116* (13), 5969–5970.

(199) Guram, A. S.; Buchwald, S. L. *J. Am. Chem. Soc.* **1994**, *116* (17), 7901–7902.

(200) Louie, J.; Hartwig, J. F. *Tetrahedron Lett.* **1995**, *36* (21), 3609–3612.

(201) Guram, A. S.; Rennels, R. A.; Buchwald, S. L. *Angew. Chemie Int. Ed. English* **1995**, *34* (12), 1348–1350.

(202) Wolfe, J. P.; Wagaw, S.; Buchwald, S. L. *J. Am. Chem. Soc.* **1996**, *118* (30), 7215–7216.

(203) Wolfe, J. P.; Buchwald, S. L. *J. Org. Chem.* **2000**, *65* (4), 1144–1157.

(204) Driver, M. S.; Hartwig, J. F. *J. Am. Chem. Soc.* **1996**, *118* (30), 7217–7218.

(205) Hamann, B. C.; Hartwig, J. F. *J. Am. Chem. Soc.* **1998**, *120* (29), 7369–7370.

(206) Louie, J.; Driver, M. S.; Hamann, B. C.; Hartwig, J. F. *J. Org. Chem.* **1997**, *62* (5), 1268–1273.

(207) Ku, Y. Y.; Chan, V. S.; Christesen, A.; Grieme, T.; Mulhern, M.; Pu, Y. M.; Wendt, M. D. *J. Org. Chem.* **2019**, *84* (8), 4814–4829.

(208) Dorel, R.; Grugel, C. P.; Haydl, A. M. *Angew. Chemie - Int. Ed.* **2019**, *58* (48), 17118–17129.

(209) Mcgowan, M. A.; Mcavoy, C. Z.; Buchwald, S. L. *Org. Lett.* **2012**, *5* (8), 2010–2013.

(210) Ullmann, F. *Berichte der Dtsch. Chem. Gesellschaft* **1903**, *36*, 2382–2384.

(211) Goldberg, I. *Berichte der Dtsch. Chem. Gesellerschaft* **1906**, *39*, 5–6.

(212) Patil, N. M.; Kelkar, A. A.; Chaudhari, R. V. *Tetrahedron Lett.* **2002**, *43*, 7243–7146.

(213) Goodbrand, H. B.; Hu, N. X. *J. Org. Chem.* **1999**, *64* (2), 670–674.

(214) Ma, D.; Cai, Q.; Zhang, H. *Org. Lett.* **2003**, *5* (14), 2453–2455.

(215) Kshirsagar, U. A.; Argade, N. P. *Org. Lett.* **2010**, *12* (16), 3716–3719.

(216) Altman, R. A.; Hyde, A. M.; Huang, X.; Buchwald, S. L. O. *J. Am. Chem. Soc.* **2008**, *130* (29), 9613–9620.

(217) Shafir, A.; Lichtor, P. A.; Buchwald, S. L. *J. Am. Chem. Soc.* **2007**, *129* (12), 3490–3491.

(218) Jones, G. O.; Liu, P.; Houk, K. N.; Buchwald, S. L. *J. Am. Chem. Soc.* **2010**, *132* (17), 6205–6213.

(219) Yamagishi, T. *Bull. Chem. Soc. Jpn.* **1978**, *51*, 277–282.

(220) Lo, Q. A.; Sale, D.; Braddock, D. C.; Davies, R. P.. *ACS Catal.* **2018**, *8* (1), 101–109.

(221) Cortes-Salva, M.; Garvin, C.; Antilla, J. C. *J. Org. Chem.* **2011**, *76* (5), 1456–1459. https://doi.org/10.1021/jo102235u.

(222) Gong, X.; Yang, H.; Liu, H.; Jiang, Y.; Zhao, Y.; Fu, H. *Org. Lett.* **2010**, *12* (14), 3128–3131.

(223) Yang, D.; Fu, H.; Hu, L.; Jiang, Y.; Zhao, Y. *J. Org. Chem.* **2008**, *73* (19), 7841–7844.

(224) Yang, D.; Liu, H.; Yang, H.; Fu, H.; Hu, L.; Jiang, Y.; Zhao, Y. *Adv. Synth. Catal.* **2009**, *351* (11–12), 1999–2004.

(225) Chan, D. M. T.; Monaco, K. L.; Wang, R. P.; Winters, M. P.. *Tetrahedron Lett.* **1998**, *39* (19), 2933–2936.

(226) Lam, P. Y. S.; Clark, C. G.; Saubern, S.; Adams, J.; Winters, M. P.; Chan, D. M. T.; Combs, A.. *Tetrahedron Lett.* **1998**, *39* (19), 2941–2944.

(227) King, A. E.; Brunold, T. C.; Stahl, S. S.. *J. Am. Chem. Soc.* **2009**, *131* (14), 5044–5045.

(228) Asada, M.; Obitsu, T.; Kinoshita, A.; Nagase, T.; Yoshida, T.; Yamaura, Y.; Takizawa, H.; Yoshikawa, K.; Sato, K.; Narita, M.; Nakai, H.; Toda, M.; Tobe, Y.. *Bioorganic Med. Chem.* **2010**, *18* (9), 3212–3223.

(229) Vantourout, J. C.; Miras, H. N.; Isidro-Llobet, A.; Sproules, S.; Watson, A. J. B. *J. Am. Chem. Soc.* **2017**, *139* (13), 4769–4779.

(230) Vantourout, J. C.; Li, L.; Bendito-Moll, E.; Chabbra, S.; Arrington, K.; Bode, B. E.; Isidro-Llobet, A.; Kowalski, J. A.; Nilson, M. G.; Wheelhouse, K. M. P.; Woodard, J. L.; Xie, S.; Leitch, D. C.; Watson, A. J. B. *ACS Catal.* **2018**, *8* (10), 9560–9566.

(231) Lam, P. Y. S.; Bonne, D.; Vincent, G.; Clark, C. G.; Combs, A. P. *Tetrahedron Lett.* **2003**, *44* (8), 1691–1694.

(232) Lam, P. Y. S.; Vincent, G.; Bonne, D.; Clark, C. G. *Tetrahedron Lett.* **2003**, *44* (26), 4927–4931.

(233) Evans, D. A.; Katz, J. L.; West, T. R. *Tetrahedron Lett.* **1998**, *39* (19), 2937–2940.

(234) King, A. E.; Ryland, B. L.; Brunold, T. C.; Stahl, S. S. *Organometallics* **2012**, *31* (22), 7948–7957.

(235) King, A. E.; Brunold, T. C.; Stahl, S. S.. *J. Am. Chem. Soc.* **2009**, *131* (14), 5044–5045.

(236) West, M. J.; Fyfe, J. W. B.; Vantourout, J. C.; Watson, A. J. B. *Chem. Rev.* **2019**, *119* (24), 12491–12523.

(237) Li, J.; Bénard, S.; Neuville, L.; Zhu, J. *Org. Lett.* **2012**, *14* (23), 5980–5983.

(238) Chu, L.; Qing, F. L. *Org. Lett.* **2010**, *12* (21), 5060–5063.

(239) Van Niekerk, J. N.; Schoening, F. R. *Nature* **1953**, *171*, 36–37.

(240) Klein, J. E. M. N.; Thatcher, R. J.; Whitwood, A. C.; Taylor, R. J. K. *Acta Crystallogr. Sect. C Cryst. Struct. Commun.* **2011**, *67* (4), m108–m110.

(241) Sheinblatt, M.; Becker, E. D.. *J. Biol. Chem.* **1967**, *242* (13), 3159–3163.

(242) Gonzalez, P.; Bossak-Ahmad, K.; Vileno, B.; Wezynfeld, N. E.; El Khoury, Y.; Hellwig, P.; Hureau, C.; Bal, W.; Faller, P.. *Chem. Commun.* **2019**, *55* (56), 8110–8113.

(243) Catterick, J.; Thornton, P. *Adv. Inorg. Chem. Radiochem.* **1977**, *20* (C), 291–362.

(244) Clement, B.; Kämpchen, T. *Chem. Ber.* **1986**, *119* (3), 1101–1104.

(245) Zu, W.; Liu, S.; Jia, X.; Xu, L. *Org. Chem. Front.* **2019**, *6* (9), 1356–1360.

(246) Bénard, S.; Neuville, L.; Zhu, J. *Chem. Commun.* **2010**, *46* (19), 3393–3395.

(247) Derosa, J.; O'Duill, M. L.; Holcomb, M.; Boulous, M. N.; Patman, R. L.; Wang, F.; Tran-Dubé, M.; McAlpine, I.; Engle, K. M. *J. Org. Chem.* **2018**, *83* (7), 3417–3425.

(248) Edwards, K. J.; Jenkins, T. C.; Neidle, S. *Biochemistry* **1992**, *31* (31), 7104–7109.

(249) Bakunova, S. M.; Bakunov, S. A.; Patrick, D. A.; Kumar, E. V. K. S.; Ohemeng, K. A.; Bridges, A. S.; Wenzler, T.; Barszcz, T.; Jones, S. K.; Werbovetz, K. A.; Brun, R.; Tidwell, R. *J. Med. Chem.* **2009**, *52* (7), 2016–2035.

(250) Athri, P.; Wenzler, T.; Tidwell, R.; Bakunova, S. M.; Wilson, W. D. P. *Eur. J. Med. Chem.* **2010**, *45* (12), 6147–6151.

(251) Scala, A.; Piperno, A.; Micale, N.; Mineo, P. G.; Abbadessa, A.; Risoluti, R.; Castelli, G.; Bruno, F.; Vitale, F.; Cascio, A.; Grassi, G. *J. Biomed. Mater. Res. - Part B Appl. Biomater.* **2018**, *106* (8), 2778–2785.

(252) Räz, B.; Iten, M.; Grether-Bühler, Y.; Kaminsky, R.; Brun, R. *Acta Trop.* **1997**, *68* (2), 139–147. https://doi.org/10.1016/S0001-706X(97)00079-X.

(253) Rink, H.; Sieber, P.; Raschdorf, F. *Tetrahedron Lett.* **1984**, *25*, 621–624.

(254) Yang, Y. Peptide Fragmentation/Deletion Side Reactions. In *Side Reactions in Peptide Synthesis*; 2016; pp 1–31.

(255) Jäger, G.; Geiger, R. *Chem. Ber.* **1970**, *103*, 1727–1747.

(256) Verdini, A. S.; Lucietto, P.; Fossati, G.; Giordani, C. *Tetrahedron Lett.* **1992**, *33* (43), 6541–6542.

(257) Alhassan, M.; Kumar, A.; Lopez, J.; Albericio, F.; de la Torre, B. G.. *Int. J. Mol. Sci.* **2020**, *21* (12), 1–12.

(258) Fischer, E. *Berichte der Dtsch. Chem. Gesellerschaft* **1915**, *48*, 93–102.

(259) Nishimura, O.; Fujino, M. *Chem. Pharm. Bull.* **1976**, *24*, 1568–1575.

(260) Haruaki, Y.; Masaharu, T.; Kanaki, J.; Kazuya, M. T. *J.C.S. Chem. Comm.* **1978**, 482–483.

(261) Fujino, M.; Wakimasu, M.; Kitada, C. *Chem. Pharm. Bull.* **1981**, *29*, 2825–2831.

(262) Ramage, R.; Green, J. *Tetrahedron Lett.* **1987**, No. 20, 2287–2290.

(263) Robert Ramage, J. G. and A. J. B. *Tetrahedron* **1991**, *41*, 6353–6370.

(264) Carpino, L. A.; Shroff, H.; Triolo, S. A.; Mansour M. E., E.-S.; Wenschuh, H.; Albericio, F. *Tetrahedron Lett.* **1993**, *34*, 7829–7832.

(265) Schweinitz, A.; Dönnecke, D.; Ludwig, A.; Steinmetzer, P.; Schulze, A.; Kotthaus, J.; Wein, S.; Clement, B.; Steinmetzer, T. *Bioorganic Med. Chem. Lett.* **2009**, *19* (7), 1960–1965.

(266) Judkins, B. D.; Allen, D. G.; Cook, T. A.; Evans, B.; Sdharwala, T. E. *Synth. Commun.* **1996**, *26*, 4351–4367.

(267) Inokuchi, E.; Yamada, A.; Hozumi, K.; Tomita, K.; Oishi, S.; Ohno, H.; Nomizu, M.; Fujii, N. *Org. Biomol. Chem.* **2011**, *9* (9), 3421–3427.

(268) Huh, S.; Appavoo, S. D.; Yudin, A. K. *Org. Lett.* **2020**.

(269) Culvenor, C. C. J.; Foster, M. C.; Hagarty, M. P. *Aust. J. Chem.* **1971**, *24*, 371–375.

(270) Feldman, P. L.; Chi, S.; Sennequier, N. Stuehr. *Bioorganic Med. Chem. Lett.* **1996**, *6*, 111–114.

(271) Bence, A. K.; Crooks, P. A. *Synth. Commun.* **2002**, *32* (13), 2075–2082.

(272) Lang, C. S.; Wong, S. H.; Chow, S.; Challinor, V. L.; Yong, K. W. L.; Fletcher, M. T.; Arthur, D. M.; Ng, J. C.; De Voss, J. J. *Org. Biomol. Chem.* **2016**, *14* (28), 6826–6832.

(273) Moore, W. M.; Webber, R. K.; Jerome, G. M.; Tjoeng, F. S.; Misko, T. P.; Currie, M. G. L-. *J. Med. Chem.* **1994**, *37* (23), 3886–3888.

(274) Eustache, J.; Grob, A.; Lam, C.; Sellier, O.; Schulz, G. *Bioorganic Med. Chem. Lett.* **1998**, *8* (21), 2961–2966.

(275) Hallinan, E. A.; Tsymbalov, S.; Dorn, C. R.; Pitzele, B. S.; Hansen, D. W.; Moore, W. M.; Jerome, G. M.; Connor, J. R.; Branson, L. F.; Widomski, D. L.; Zhang, Y.; Currie, M. G.; Manning, P. T. *J. Med. Chem.* **2002**, *45* (8), 1686–1689.

(276) Flemer, S.; Madalengoitia, J. S. C. *Synthesis.* **2011**, No. 10, 1638–1648.

(277) Flemer, S.; Madalengoitia, J. S. *J. Pept. Sci.* **2012**, *18* (1), 30–36..

(278) Wright, J. A.; Yu, J.; Spencer, J. B. *Tetrahedron Lett.* **2001**, *42* (24), 4033–4036.

(279) Williams, A. L.; Dandepally, S. R.; Kotturi, S. V. A. *Mol. Divers.* **2010**, *14* (4), 697–707.

(280) D'Anna, F.; Marullo, S.; Noto, R. Aryl. *J. Org. Chem.* **2010**, *75* (3), 767–771.

(281) Alghamdi, A. H.; Munday, J. C.; Campagnaro, G. D.; Gurvic, D.; Svensson, F.; Okpara, C. E.; Kumar, A.; Quintana, J.; Abril, M. E. M.; Milić, P.; Watson, L.; Paape, D.; Settimo, L.; Dimitriou, A.; Wielinska, J.; Smart, G.; Anderson, L. F.; Woodley, C. M.; Kelly, S. P. Y.; Ibrahim, H. M. S.; Hulpia, F.; Al-Salabi, M. I.; Eze, A. A.; Sprenger, T.; Teka, I. A.; Gudin,

S.; Weyand, S.; Field, M.; Dardonville, C.; Tidwell, R. R.; Carrington, M.; O'neill, P.; Boykin, D. W.; Zachariae, U.; De Koning, H. P. *Elife* **2020**, *9*, 1–33.

(282) Basselin, M.; Denise, H.; Coombs, G. H.; Barrett, M. P. *Antimicrob. Agents Chemother.* **2002**, *46* (12), 3731–3738.

(283) Warf, M. B.; Nakamori, M.; Matthys, C. M.; Thornton, C. A.; Berglund, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (44), 18551–18556.

(284) Brittain, W. D. G.; Cobb, S. L. *Org. Biomol. Chem.* **2017**, *16* (1), 10–20.

(285) Hudson, A. S.; Caron, L.; Colgin, N.; Cobb, S.. *Amino Acids* **2015**, *47* (4), 779–785.

Appendix

See attached document

Appendix

Guanidinium mimetics: synthesis and application as next-generation analogues in cell penetrating peptides

Jack Robertson

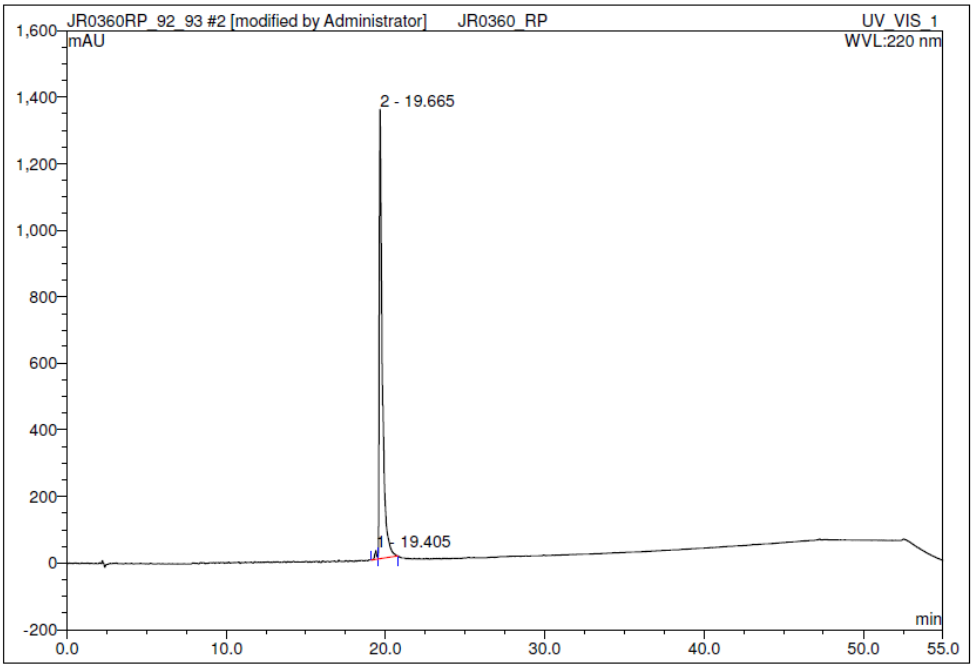
Contents

Chapter 2: HPLC traces, HRMS, Flow cytometry histograms, microscopy images 2
HPLC traces 3
HRMS 6
Flow cytometry histograms 11
HeLa cell experiments 11
U2OS cell experiments 13
Microscope images 15
Chapter 3: NMR spectra and X-ray crystallographic data tables. 16
NMR spectra 17
X-ray crystallography tables 52
Complex **3.135** 52
Complex **3.136** 61
(*Z*)-*N*'-(4-(trifluoromethyl)phenyl)benzimidamide (**3.146**) 70
(*Z*)-*N*'-(3-fluoro-4-methoxyphenyl)benzimidamide (**3.148**) 77
Chapter 4: NMR spectra, HPLC traces, HR mass spectra, Flow cytometry histograms and microscope images. 86
NMR Spectra 87
HPLC traces 90
High resolution mass spectra 95
Flow cytometry 103
Microscope images 104

Chapter 2: HPLC traces, HRMS, Flow cytometry histograms, microscopy images

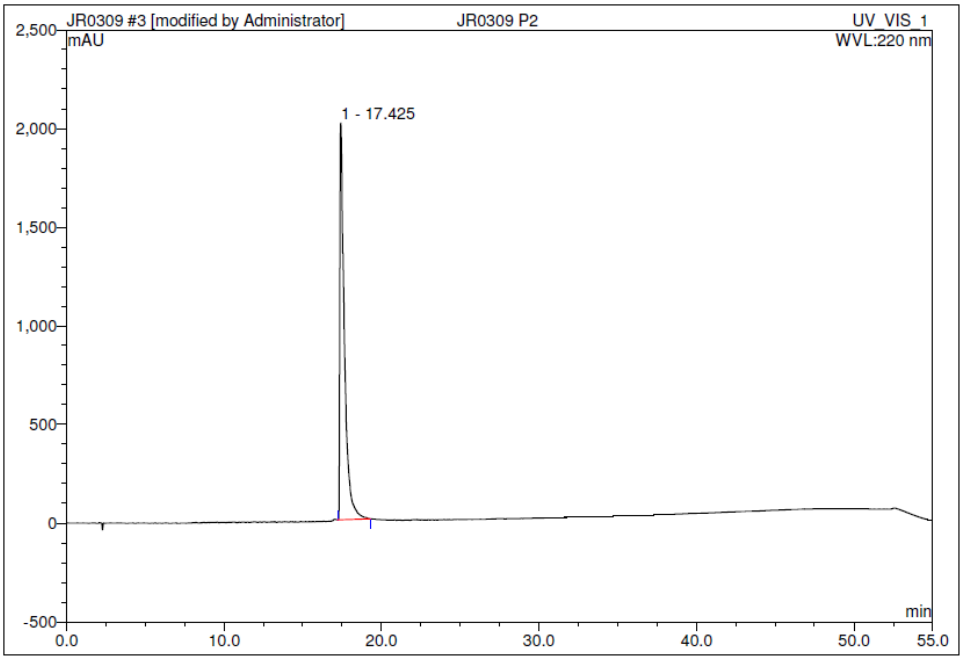
HPLC traces

Pip6



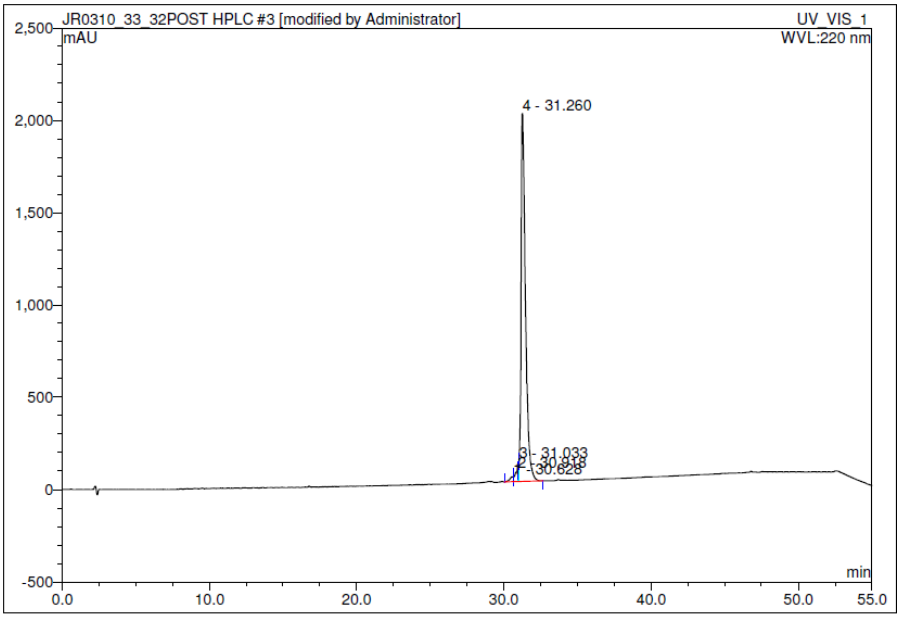
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	19.41	n.a.	26.142	3.626	1.15	n.a.	BM
2	19.67	n.a.	1349.971	311.196	98.85	n.a.	MB*
Total:			1376.114	314.822	100.00	0.000	

Tat



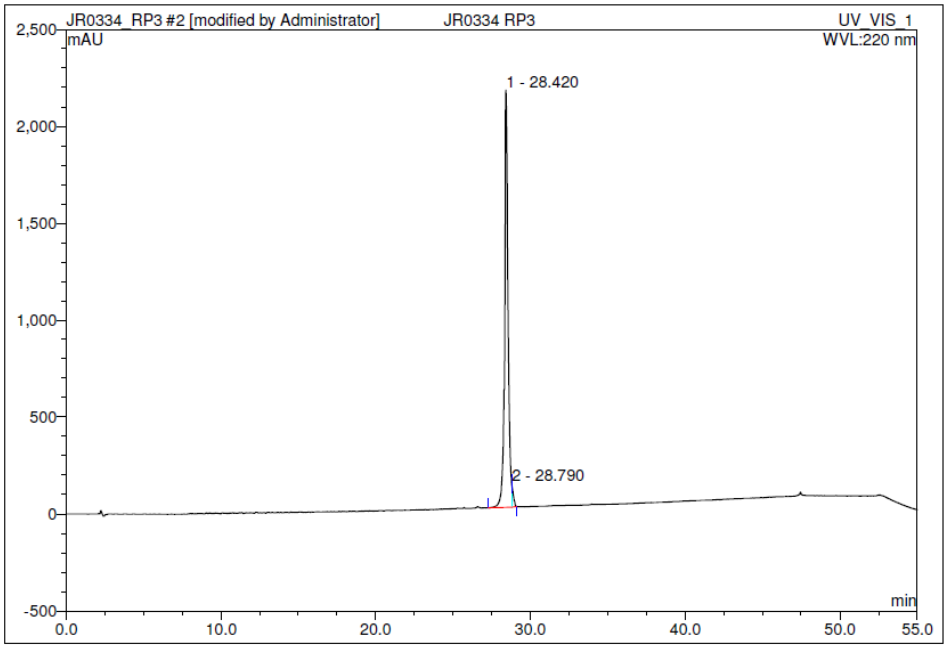
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	17.43	n.a.	2011.392	691.964	100.00	n.a.	BMB
Total:			2011.392	691.964	100.00	0.000	

TP-1



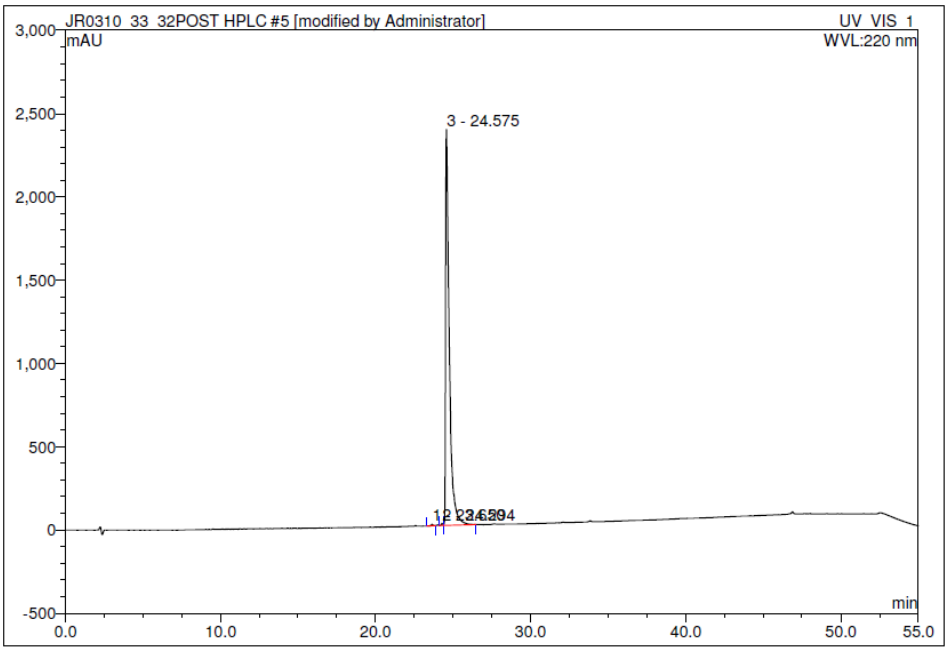
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	30.63	n.a.	25.415	7.762	1.11	n.a.	BM *
2	30.92	n.a.	58.166	9.867	1.41	n.a.	M *
3	31.03	n.a.	109.205	9.733	1.39	n.a.	M *
4	31.26	n.a.	1992.757	673.879	96.10	n.a.	MB*
Total:			2185.542	701.242	100.00	0.000	

TP-2



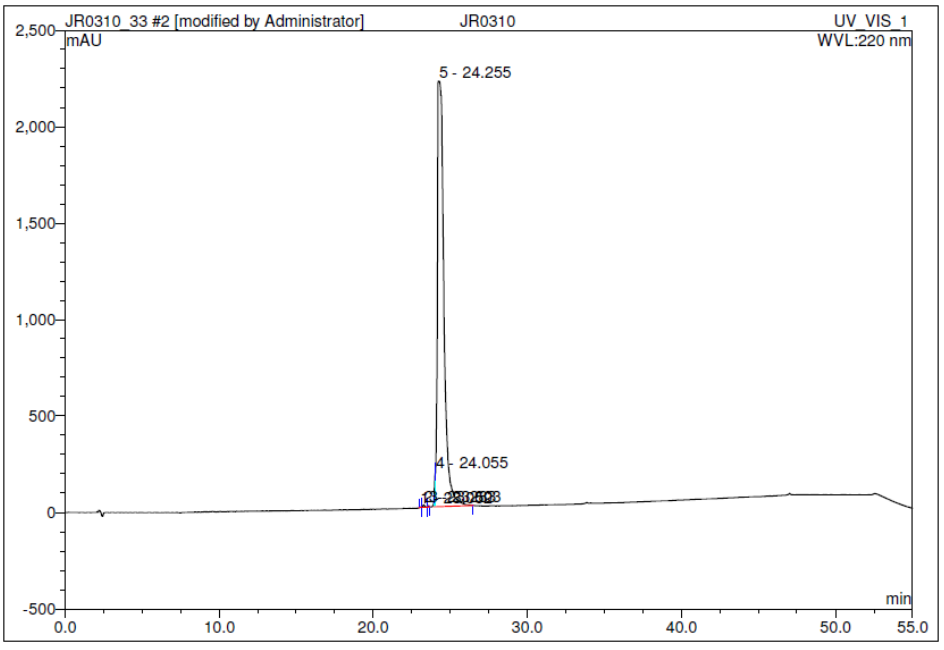
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	28.42	n.a.	2152.139	559.278	97.43	n.a.	BM *
2	28.79	n.a.	120.176	14.725	2.57	n.a.	MB*
Total:			2272.315	574.003	100.00	0.000	

2.82



No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	23.65	n.a.	8.384	1.362	0.21	n.a.	BMB
2	24.29	n.a.	12.612	1.251	0.19	n.a.	BM
3	24.58	n.a.	2376.733	656.948	99.60	n.a.	MB
Total:			2397.729	659.561	100.00	0.000	

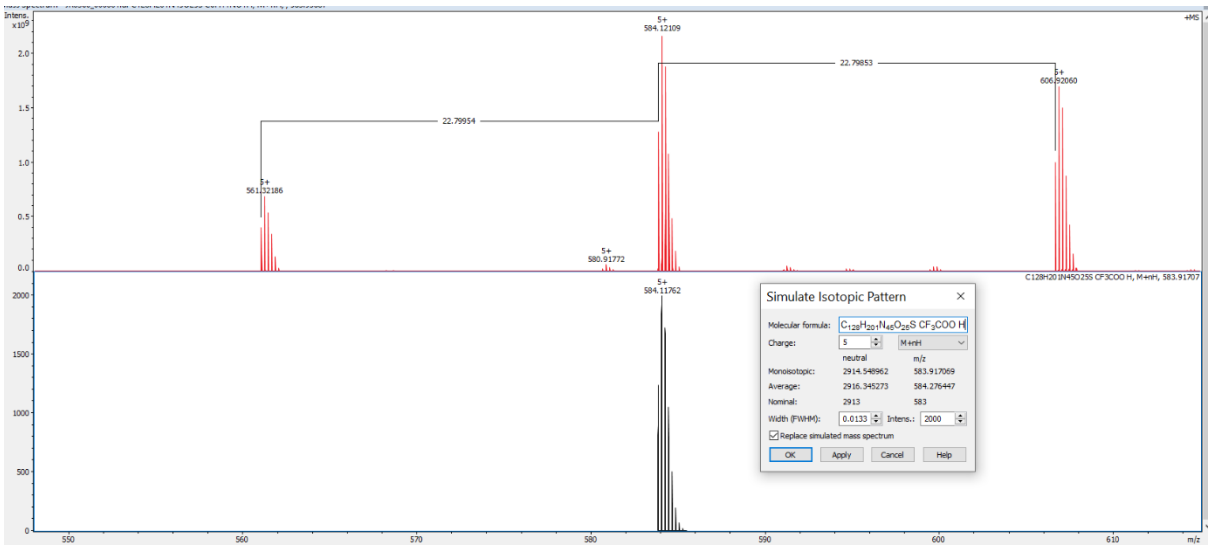
2.83



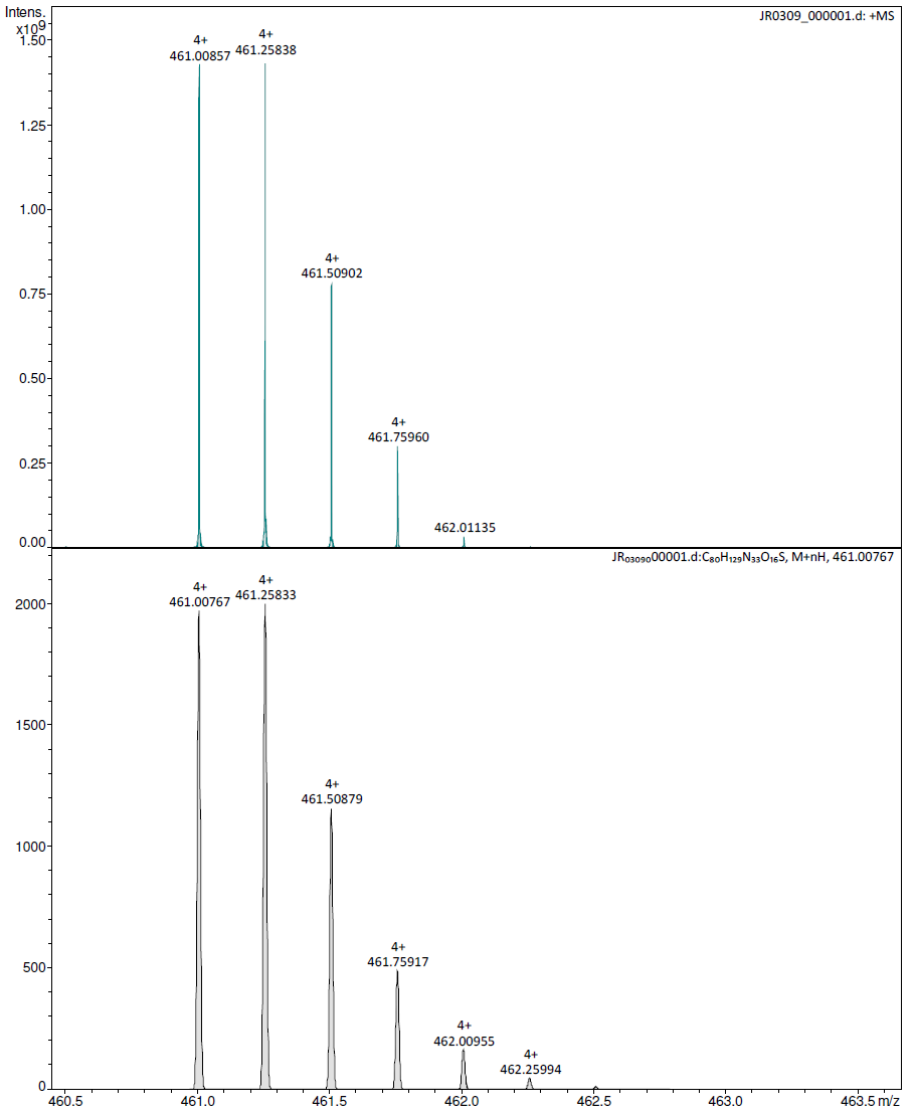
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	23.05	n.a.	3.780	0.300	0.03	n.a.	BMB*
2	23.28	n.a.	12.192	1.825	0.17	n.a.	BMB
3	23.60	n.a.	6.875	0.547	0.05	n.a.	BMB*
4	24.06	n.a.	181.002	14.162	1.28	n.a.	M *
5	24.26	n.a.	2206.976	1088.139	98.48	n.a.	MB*
Total:			2410.826	1104.974	100.00	0.000	

HRMS

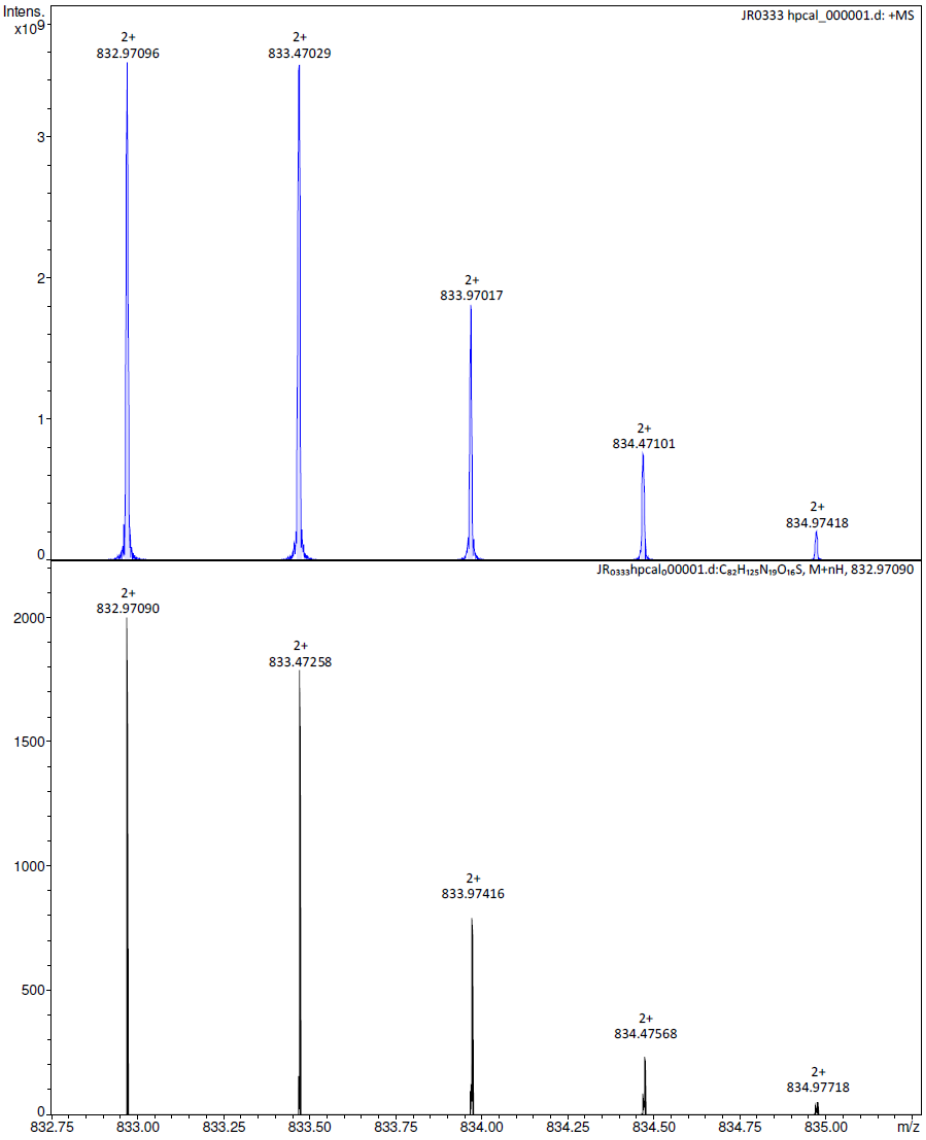
Pip6



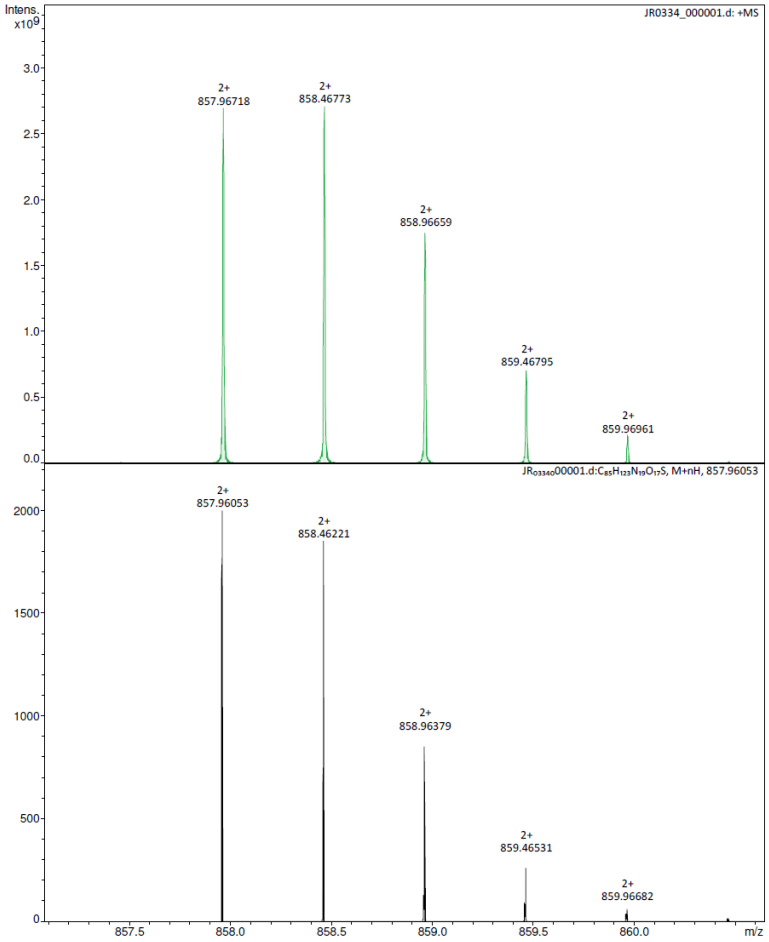
Tat



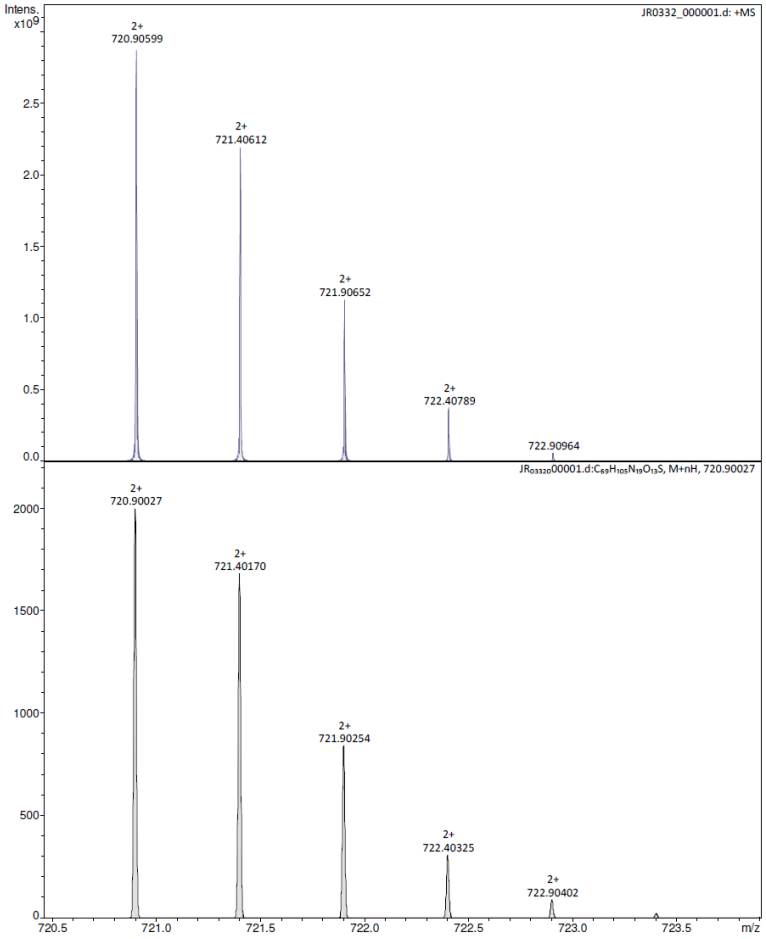
TP-1


Intens.
x10^9
JR0333 hpcal_000001.d: +MS
2+
832.97096
2+
833.47029
2+
833.97017
2+
834.47101
2+
834.97418
JR0333hpcal00001.d:C82H125N19O16S, M+nH, 832.97090
2+
832.97090
2+
833.47258
2+
833.97416
2+
834.47568
2+
834.97718
m/z

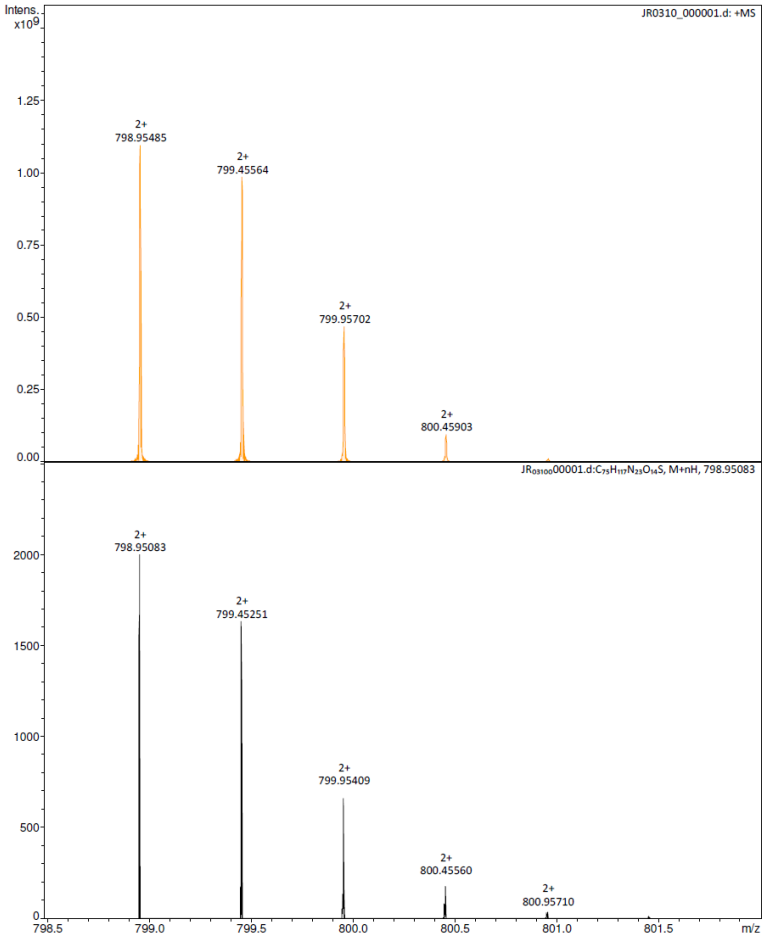
TP-2


Intens.
x10^9
JR0334_000001.d: +MS
2+
857.96718
2+
858.46773
2+
858.96659
2+
859.46795
2+
859.96961
JR03340 00001.d:C85H123N19O17S, M+nH, 857.96053
2+
857.96053
2+
858.46221
2+
858.96379
2+
859.46531
2+
859.96682
m/z

2.82


Intens.
x10^9
JR0332_000001.d: +MS
2+
720.90599
2+
721.40612
2+
721.90652
2+
722.40789
722.90964
JR0332c00001.d:C69H105N19O13S, M+nH, 720.90027
2+
720.90027
2+
721.40170
2+
721.90254
2+
722.40325
2+
722.90402
m/z

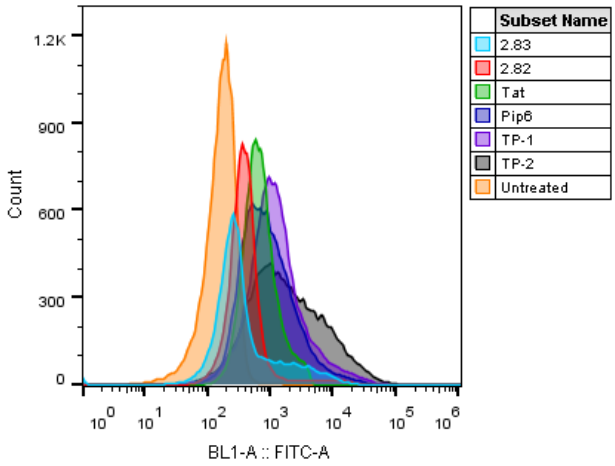
2.83


Intens.
x10^9
JR0310_000001.d: +MS
2+
798.95485
2+
799.45564
2+
799.95702
2+
800.45903
JR03100000001.d:C75H117N23O14S, M+nH, 798.95083
2+
798.95083
2+
799.45251
2+
799.95409
2+
800.45560
2+
800.95710
m/z

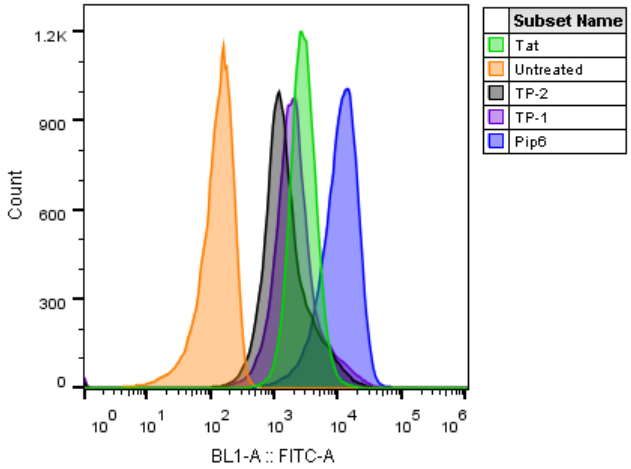
Flow cytometry histograms

HeLa cell experiments

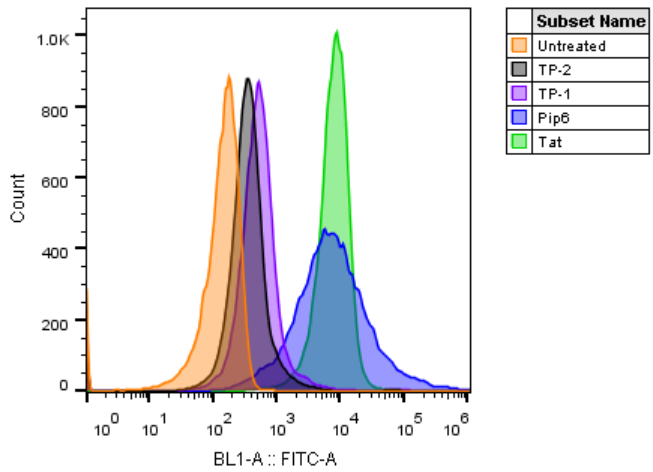
4 h incubation at 37 °C



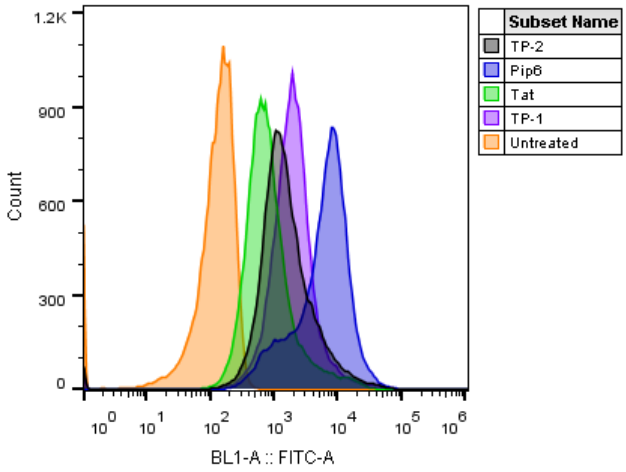
1 h incubation at 37 °C



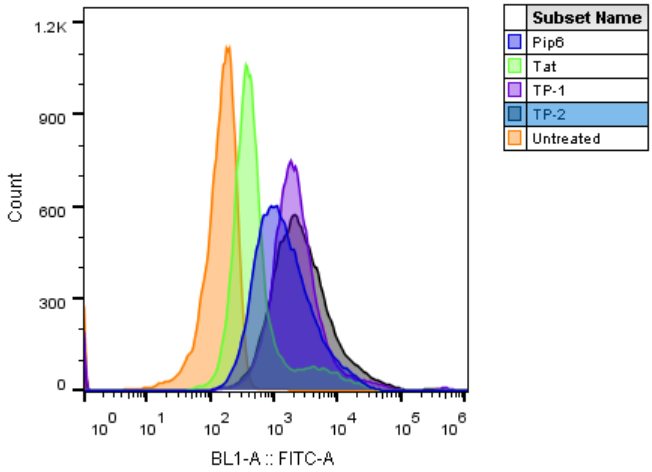
4 h incubation at 4 °C



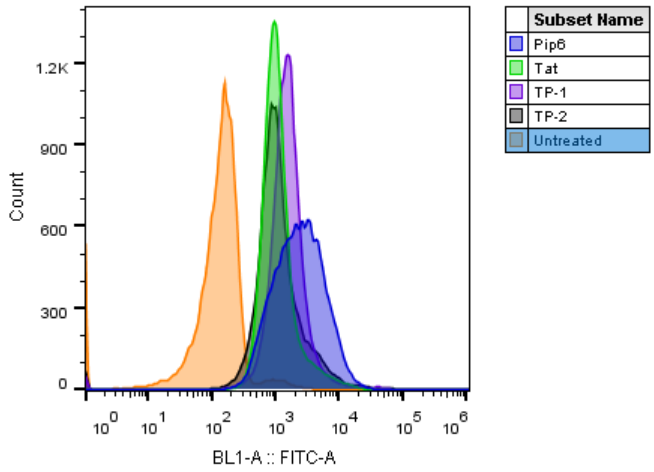
4 h incubation at 37 °C with 10 µM Chlorpromazine



4 h incubation at 37 °C with 50 µM Nystatin

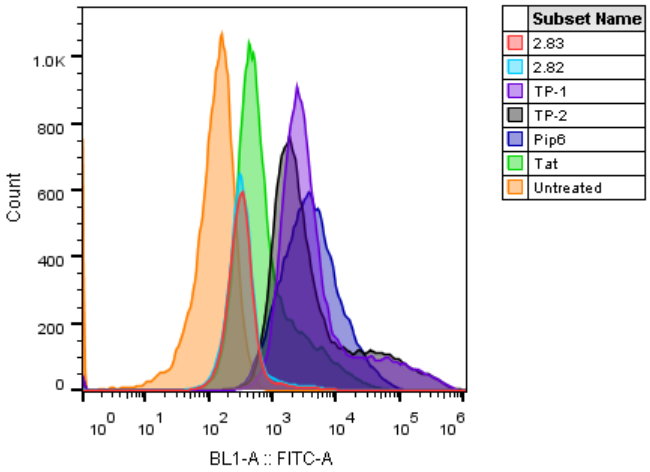


4 h incubation at 37 °C with 0.2 µM Wortmannin

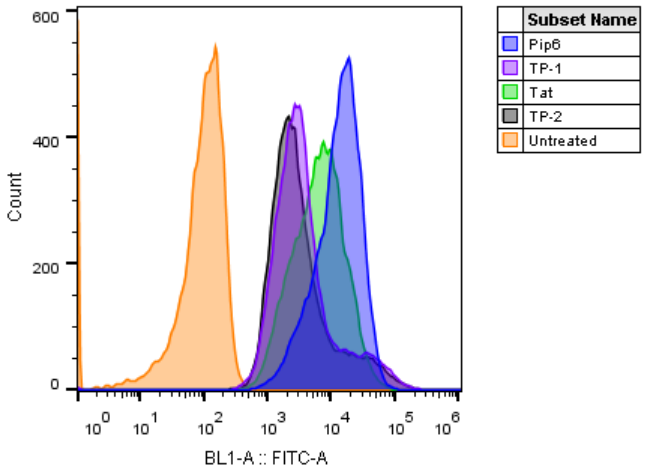


U2OS cell experiments

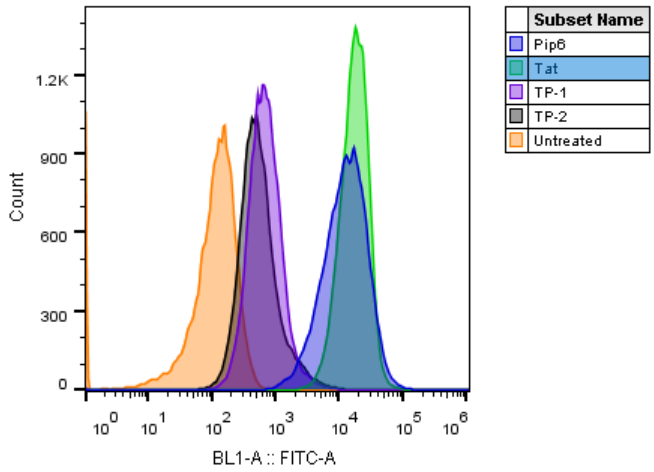
4 h incubation at 37 °C



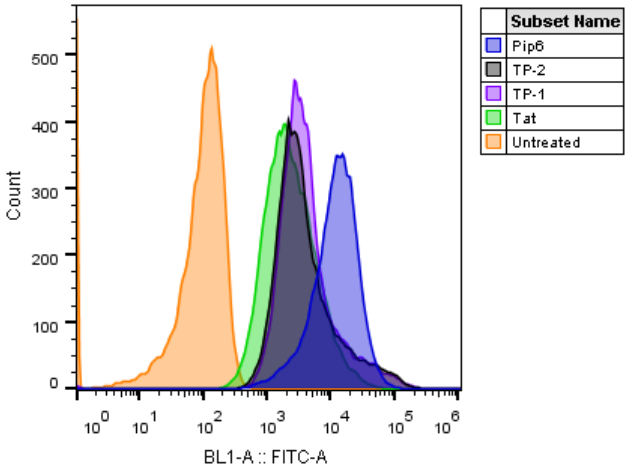
1 h incubation at 37 °C



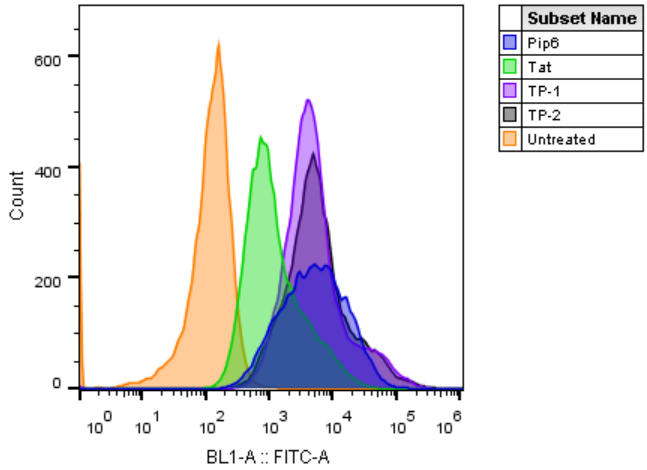
4 h incubation at 4 °C



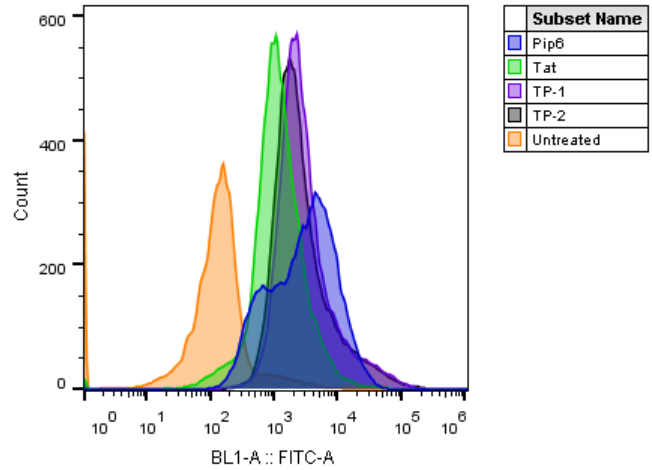
4 h incubation at 37 °C with 10 µM Chlorpromazine



4 h incubation at 37 °C with 50 µM Nystatin

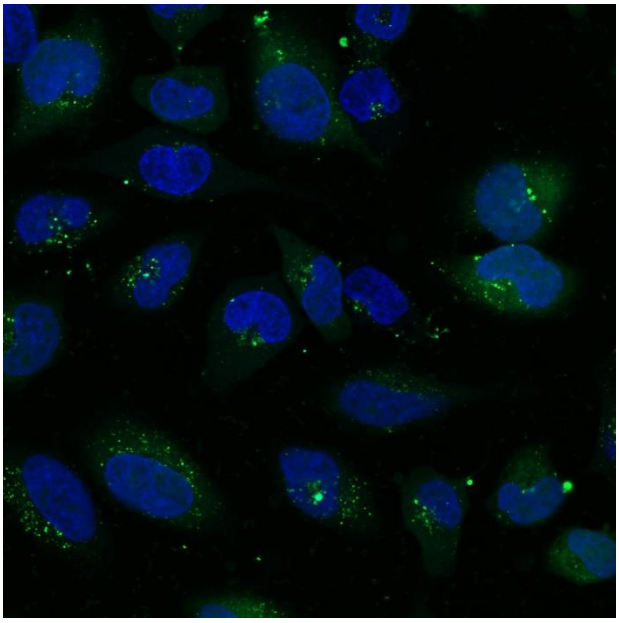


4 h incubation at 37 °C with 0.2 µM Wortmannin

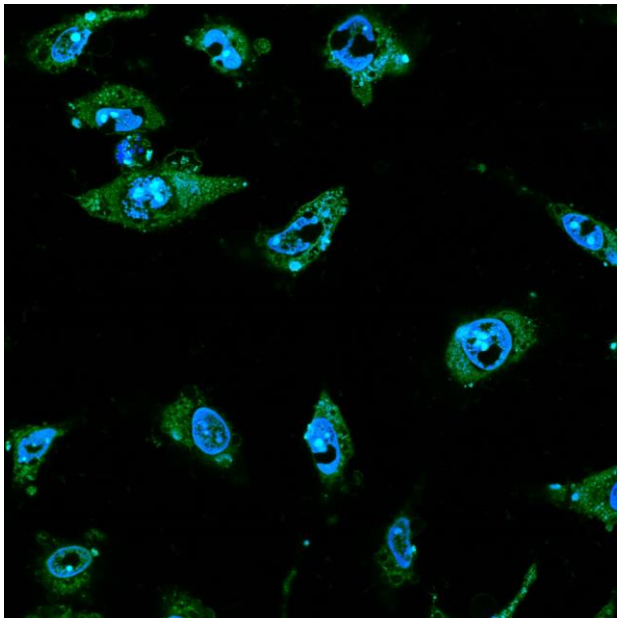


Microscope images

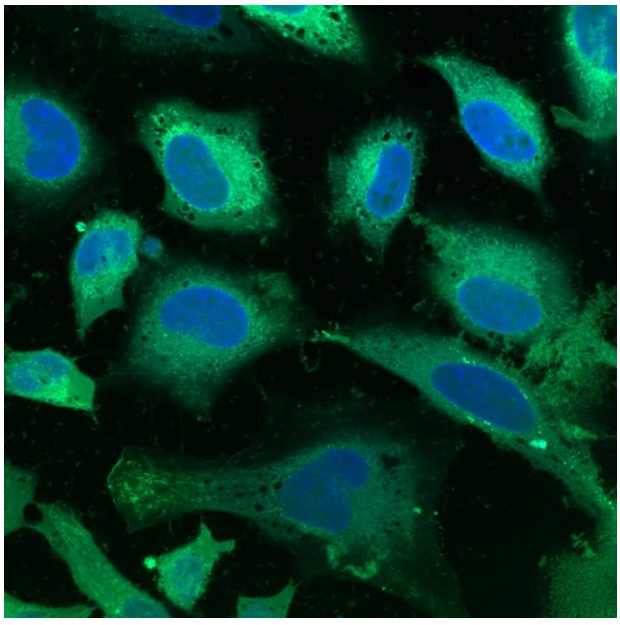
Tat



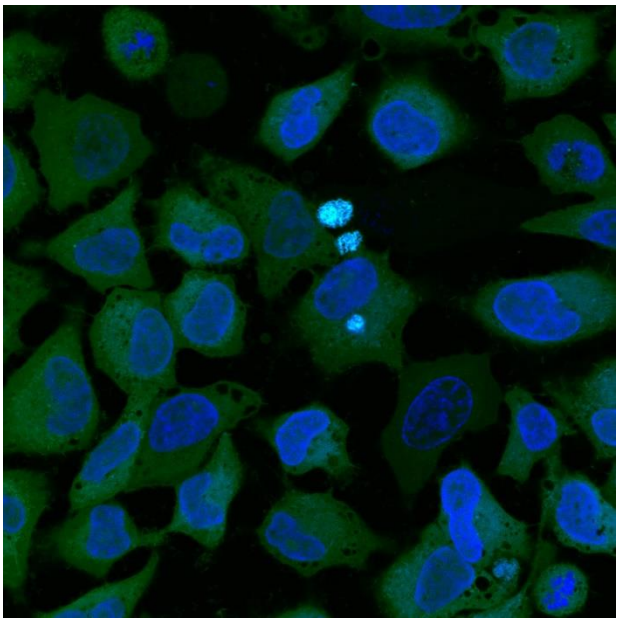
Pip6



TP-1



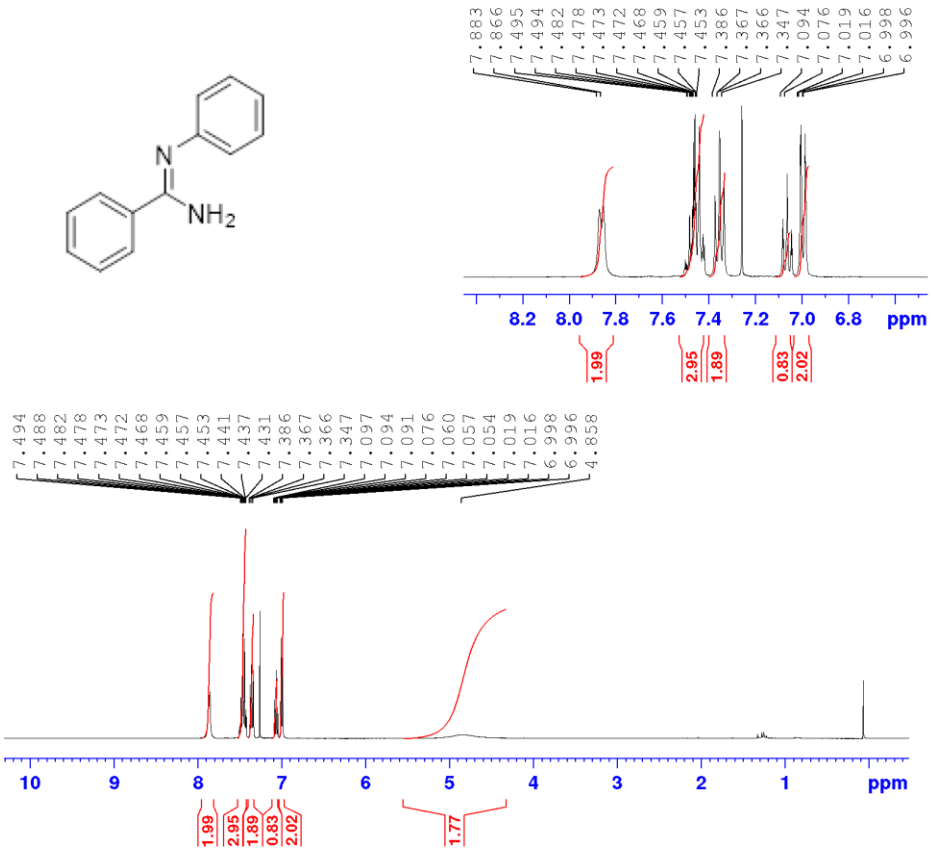
TP-2

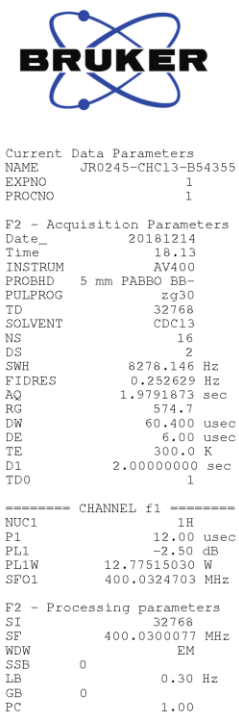


Chapter 3: NMR spectra and X-ray crystallographic data tables.

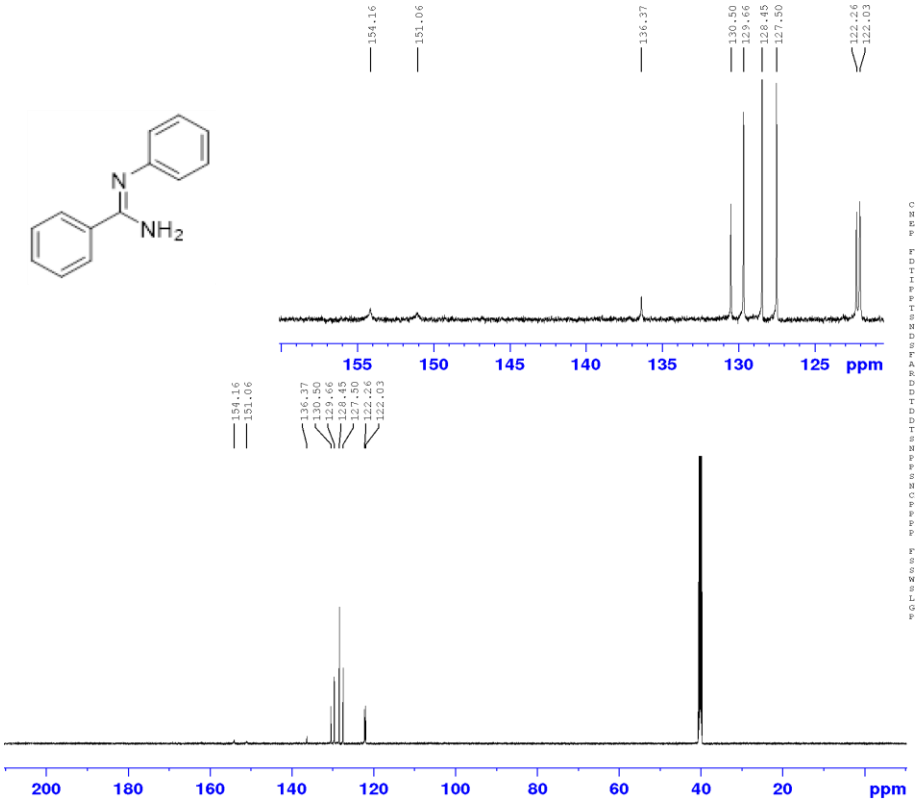
NMR spectra

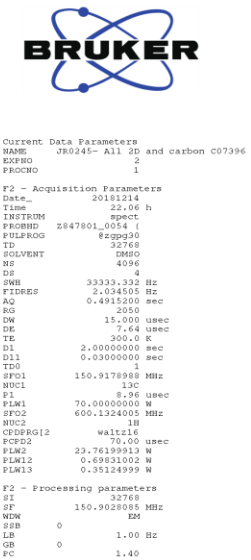
(Z)-*N'*-phenylbenzimidamide (**3.125**) - ^{1}H NMR



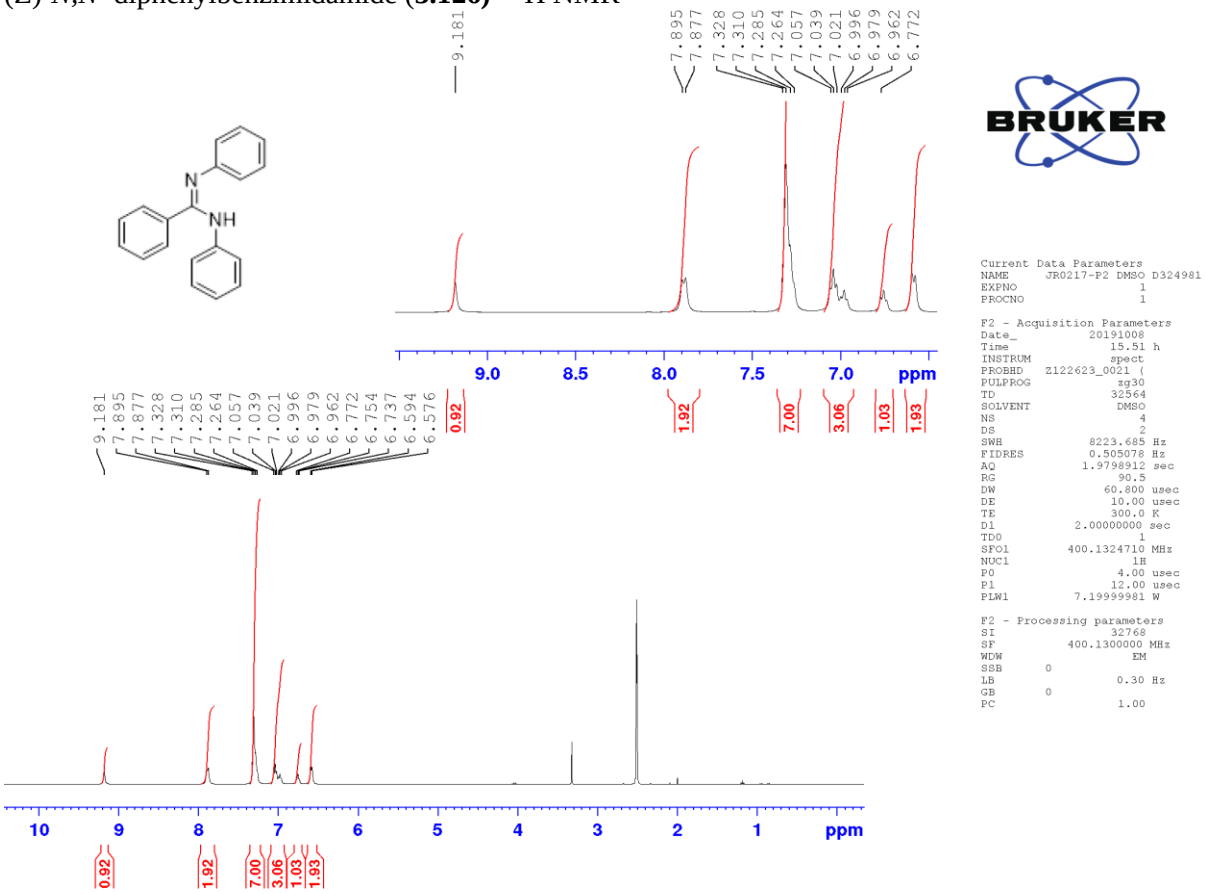


^{13}C NMR





(Z)-*N*,*N'*-diphenylbenzimidamide (**3.126)** - ^{1}H NMR

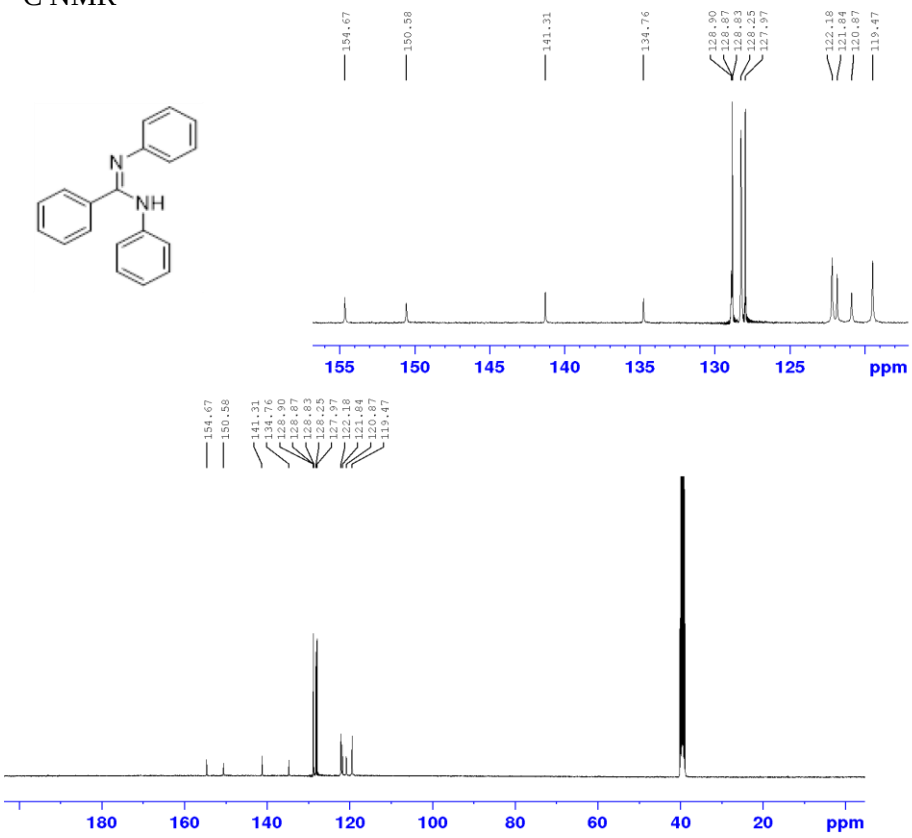


Current Data Parameters
NAME JR0217-P2 DMSO D324981
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20191008
Time 15.51 h
INSTRUM spect
PROBHD Z122623_0021 (
PULPROG zg30
TD 32564
SOLVENT DMSO
NS 4
DS 2
SWH 8223.685 Hz
FIDRES 0.505078 Hz
AQ 1.9798912 sec
RG 90.5
DW 60.800 usec
DE 10.00 usec
TE 300.0 K
D1 2.00000000 sec
TD0 1
SFO1 400.1324710 MHz
NUC1 1H
P0 4.00 usec
P1 12.00 usec
PLW1 7.19999981 W

F2 - Processing parameters
SI 32768
SF 400.1300000 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

^{13}C NMR

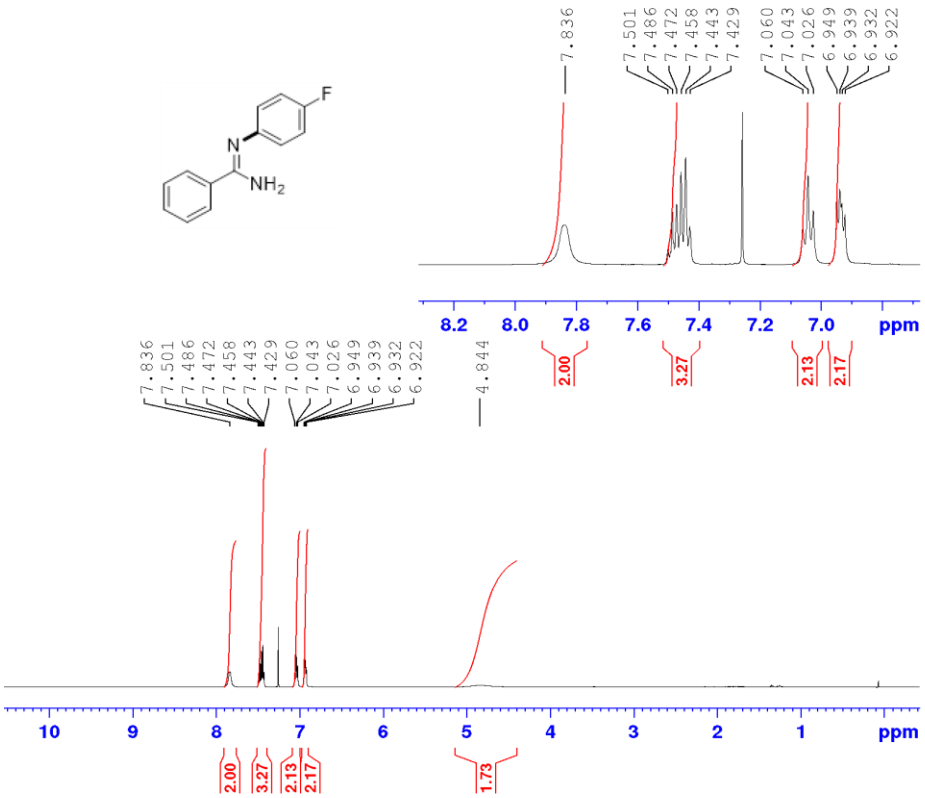


Current Data Parameters
NAME JR0217-P2 Carbon DMSO D325220
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20191013
Time 1.44 h
INSTRUM spect
PROBHD Z122623_0021 (
PULPROG zgpg30
TD 16384
SOLVENT DMSO
NS 8192
DS 4
SWH 25252.525 Hz
FIDRES 3.082584 Hz
AQ 0.3244032 sec
RG 2050
DW 19.800 usec
DE 18.00 usec
TE 300.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
SFO1 100.6228298 MHz
NUC1 13C
P0 3.33 usec
P1 10.00 usec
PLW1 35.00000000 W
SFO2 400.1316005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 90.00 usec
PLW2 7.50000000 W
PLW12 0.16379000 W
PLW13 0.08238400 W

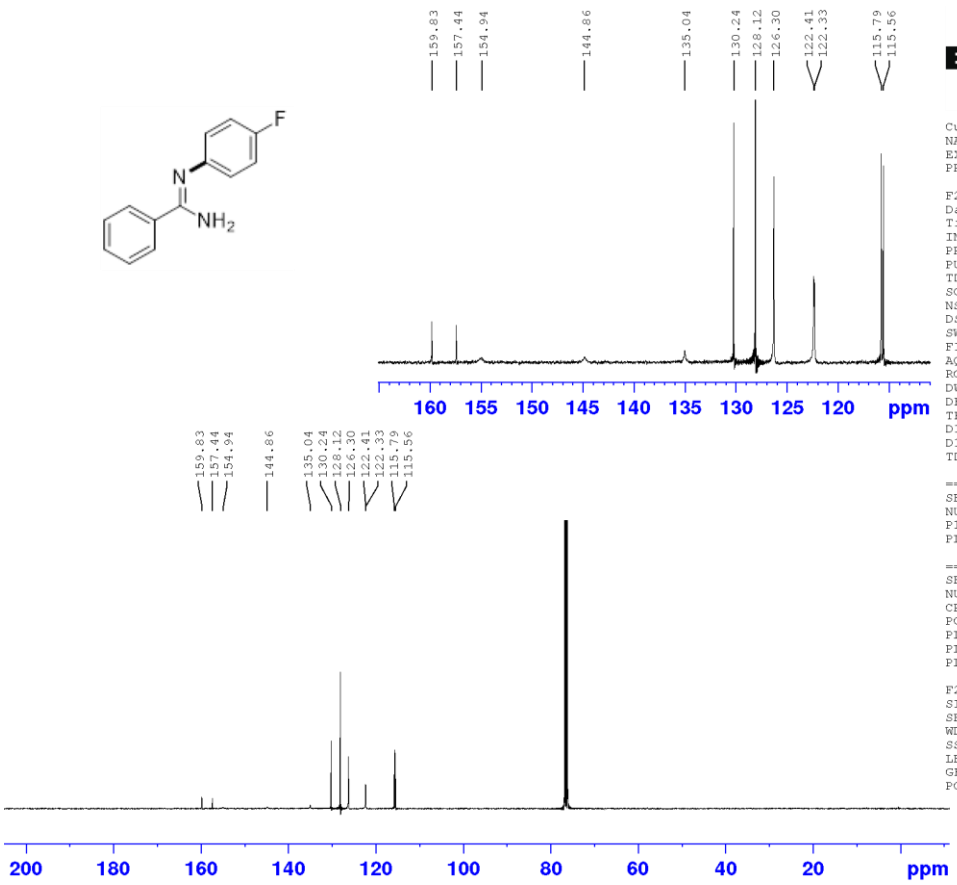
F2 - Processing parameters
SI 32768
SF 100.6128193 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

(Z)-*N*'-(4-fluorophenyl)benzimidamide (**3.149**) - ^{1}H NMR



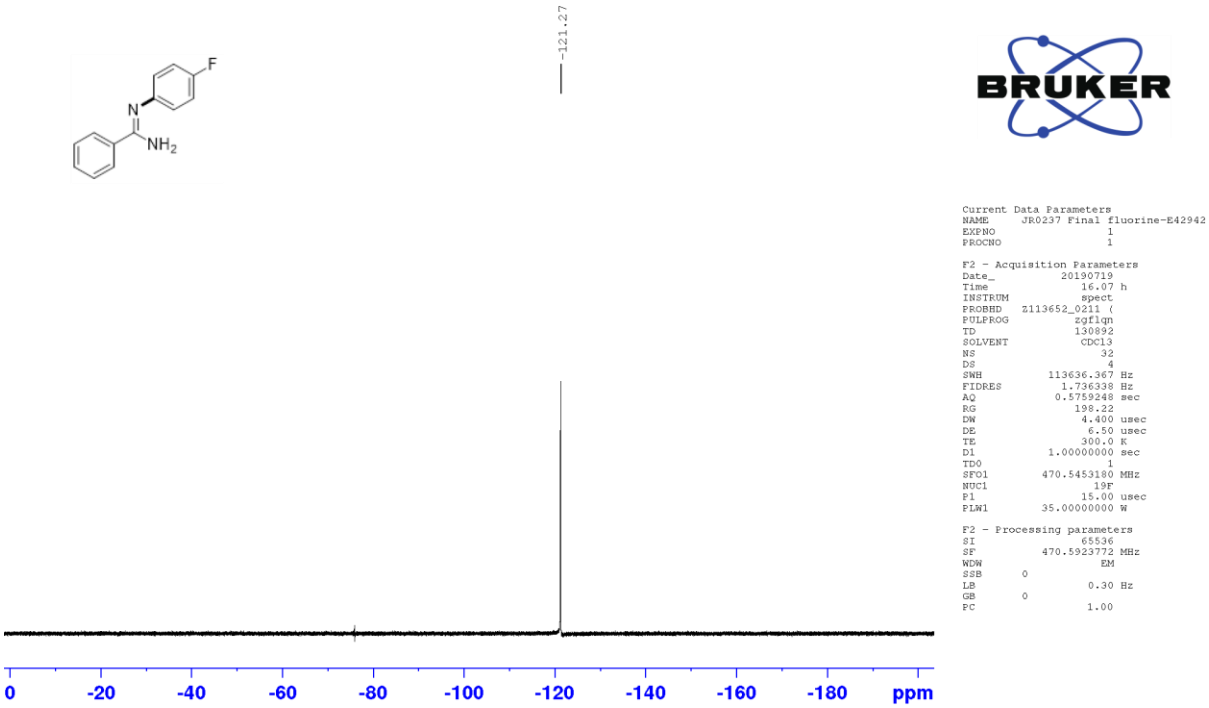
```
Current Data Parameters
NAME      JR0237B-E38730
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190106
Time              19.01 h
INSTRUM             spect
PROBHD    Z113652_0211 (
PULPROG             zg30
TD                 39578
SOLVENT             CDCl3
NS                     16
DS                      2
SWH           10000.000 Hz
FIDRES            0.505331 Hz
AQ                1.9789000 sec
RG                   198.22
DW                 50.000 usec
DE                  6.50 usec
TE                 300.0 K
D1              2.00000000 sec
TD0                    1
SFO1           500.1330885 MHz
NUC1                   1H
P1                 10.00 usec
PLW1           20.00000000 W

F2 - Processing parameters
SI                 65536
SF            500.1300143 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```

^{13}C NMR

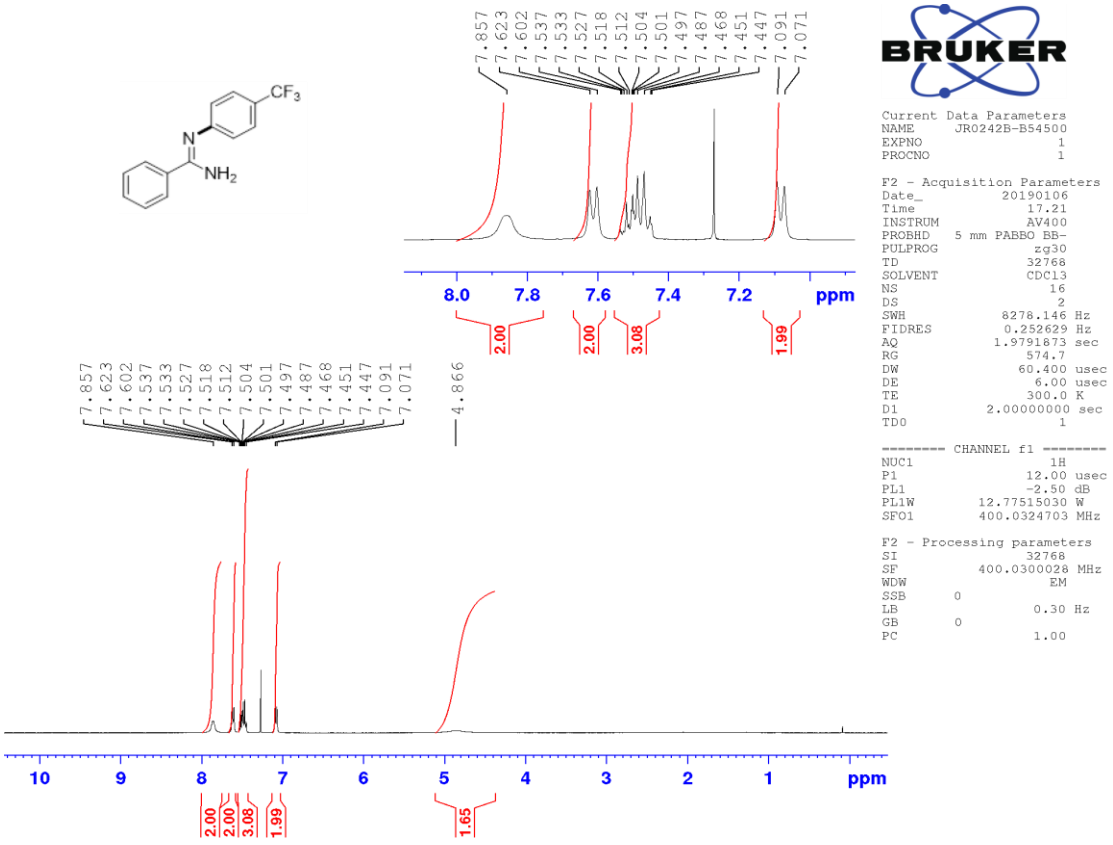


```
Current Data Parameters
NAME      JR0237 Carbon-D311600
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190113
Time               6.48
INSTRUM             spect
PROBHD    5 mm CPPBBO BB
PULPROG            zgpg30
TD                 16384
SOLVENT             CDCl3
NS                  8192
DS                     4
SWH           25252.525 Hz
FIDRES            1.541292 Hz
AQ                0.3244032 sec
RG                  2050
DW                 19.800 usec
DE                 18.00 usec
TE                 300.0 K
D1              2.00000000 sec
D11             0.03000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1          100.6228298 MHz
NUC1                  13C
P1                 10.00 usec
PLW1           35.00000000 W

======== CHANNEL f2 ========
SFO2          400.1316005 MHz
NUC2                   1H
CPDPRG[2          waltz16
PCPD2              90.00 usec
PLW2            7.50000000 W
PLW12           0.16379000 W
PLW13           0.08238400 W

F2 - Processing parameters
SI                 32768
SF            100.6128193 MHz
WDW                   EM
SSB       0
LB                 1.00 Hz
GB        0
PC                 1.40
```

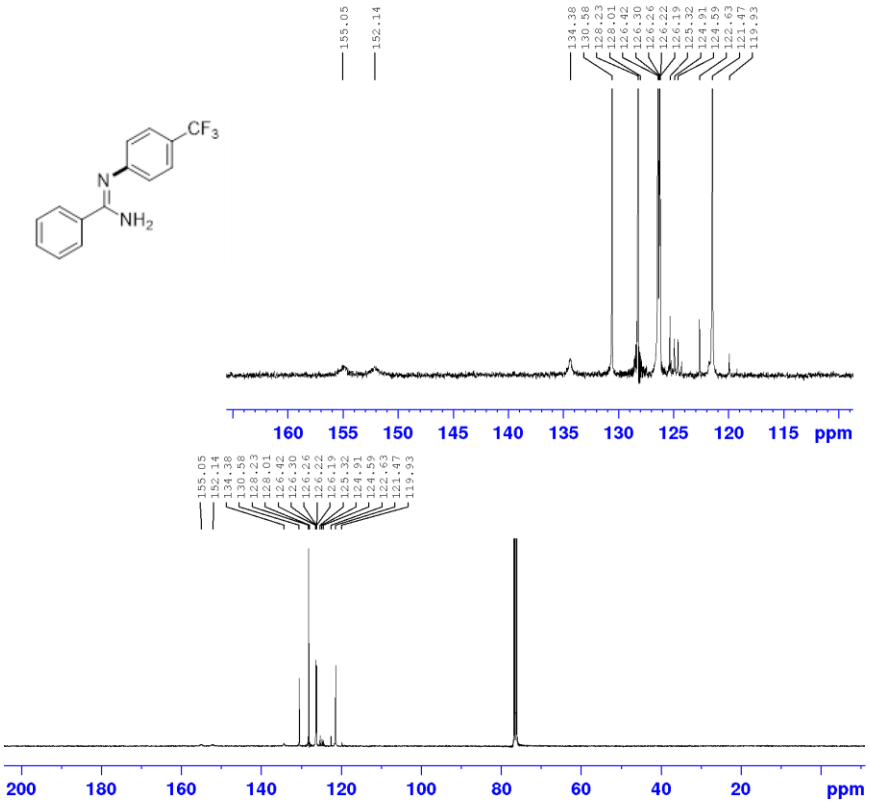
^{19}F NMR



(Z)-*N*'-(4-(trifluoromethyl)phenyl)benzimidamide (**3.146**) - ^{1}H NMR

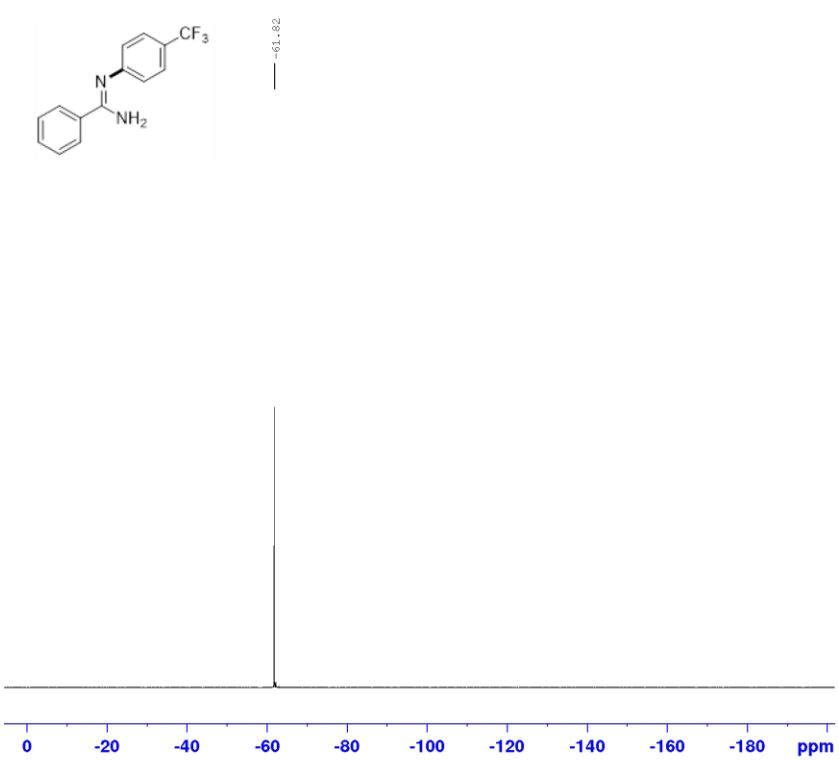


^{13}C NMR



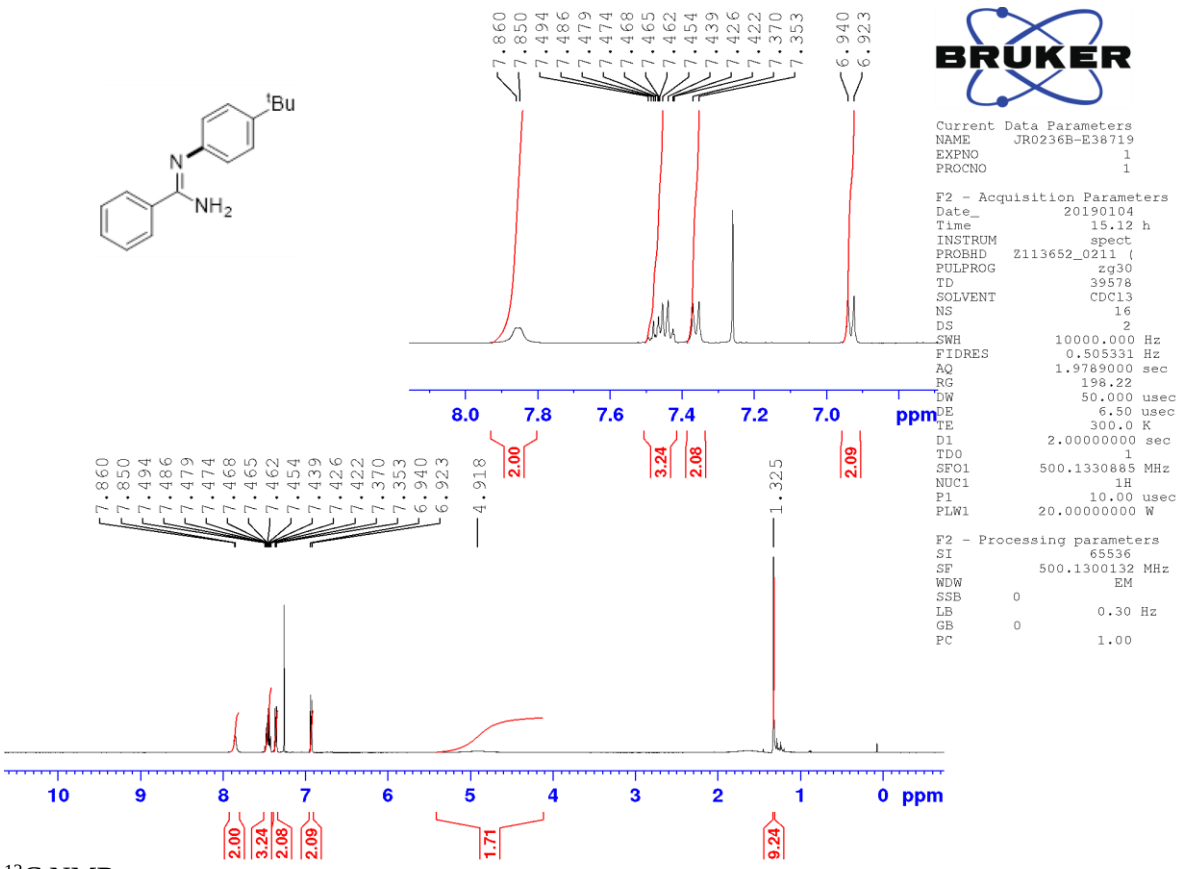

```
Current Data Parameters
NAME      JR0242 Fluorine and carbon -D316978
EXPNO                 2
PROCNO                1

F2 - Acquisition Parameters
Date_          20190414
Time               3.17 h
INSTRUM           spect
PROBHD    Z122623_0021 (
PULPROG          zgpg30
TD                16384
SOLVENT           CDCl3
NS                 8192
DS                    4
SWH           25252.525 Hz
FIDRES         3.082584 Hz
AQ             0.3244032 sec
RG                 2050
DW               19.800 usec
DE               18.00 usec
TE                300.0 K
D1            2.00000000 sec
D11           0.03000000 sec
TD0                   1
SFO1        100.6228298 MHz
NUC1                13C
P0                 3.33 usec
P1                10.00 usec
PLW1         35.00000000 W
SFO2         400.1316005 MHz
NUC2                1H
CPDPRG[2         waltz16
PCPD2             90.00 usec
PLW2          7.50000000 W
PLW12         0.16379000 W
PLW13         0.08238400 W

F2 - Processing parameters
SI                 32768
SF          100.6128193 MHz
WDW                   EM
SSB       0
LB                 1.00 Hz
GB        0
PC                 1.40
```

^{19}F NMR

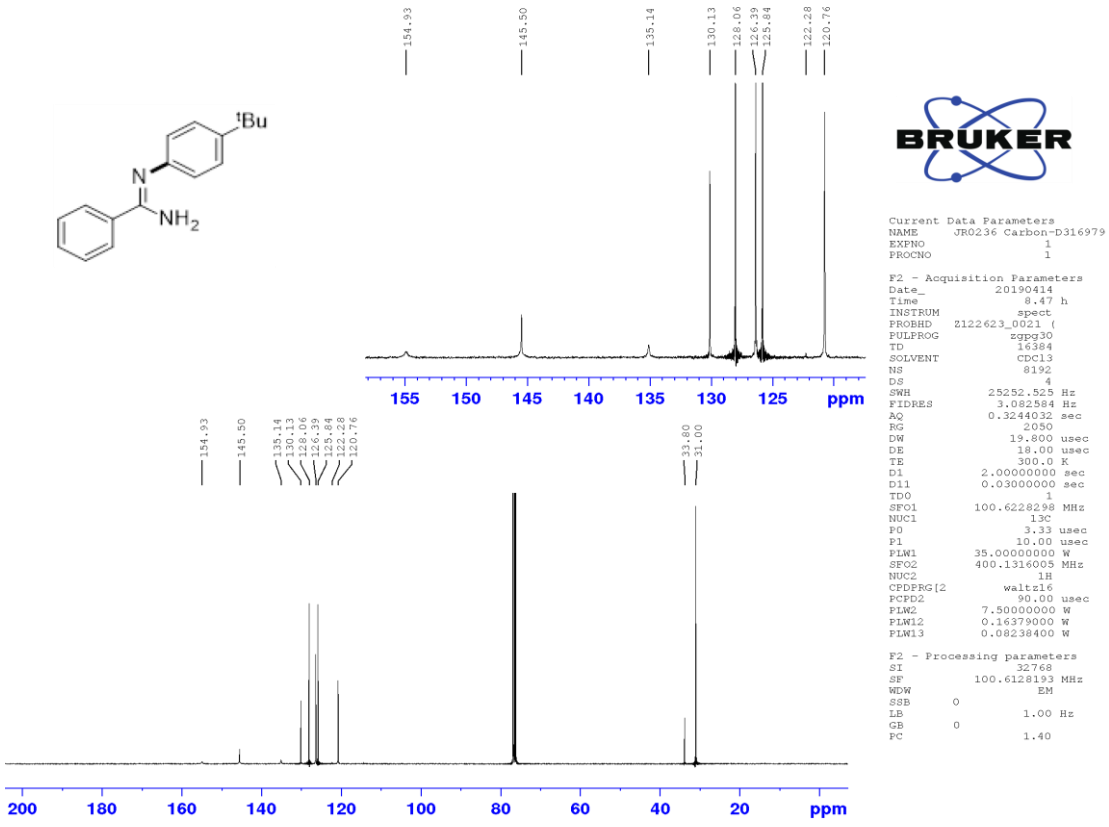



```
Current Data Parameters
NAME      JR0242 Fluorine and carbon -D316978
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_          20190413
Time              14.31 h
INSTRUM           spect
PROBHD    Z122623_0021 (
PULPROG             zg30
TD                261896
SOLVENT            CDCl3
NS                    32
DS                    4
SWH           89285.711 Hz
FIDRES         0.681841 Hz
AQ             1.4666176 sec
RG                 25.4
DW                5.600 usec
DE               18.37 usec
TE                300.0 K
D1            2.00000000 sec
TD0                   1
SFO1        376.4607162 MHz
NUC1                19F
P0                5.00 usec
P1               15.00 usec
PLW1          5.19999981 W

F2 - Processing parameters
SI                262144
SF          376.4983660 MHz
WDW                   EM
SSB       0
LB                 0.50 Hz
GB        0
PC                 1.00
```

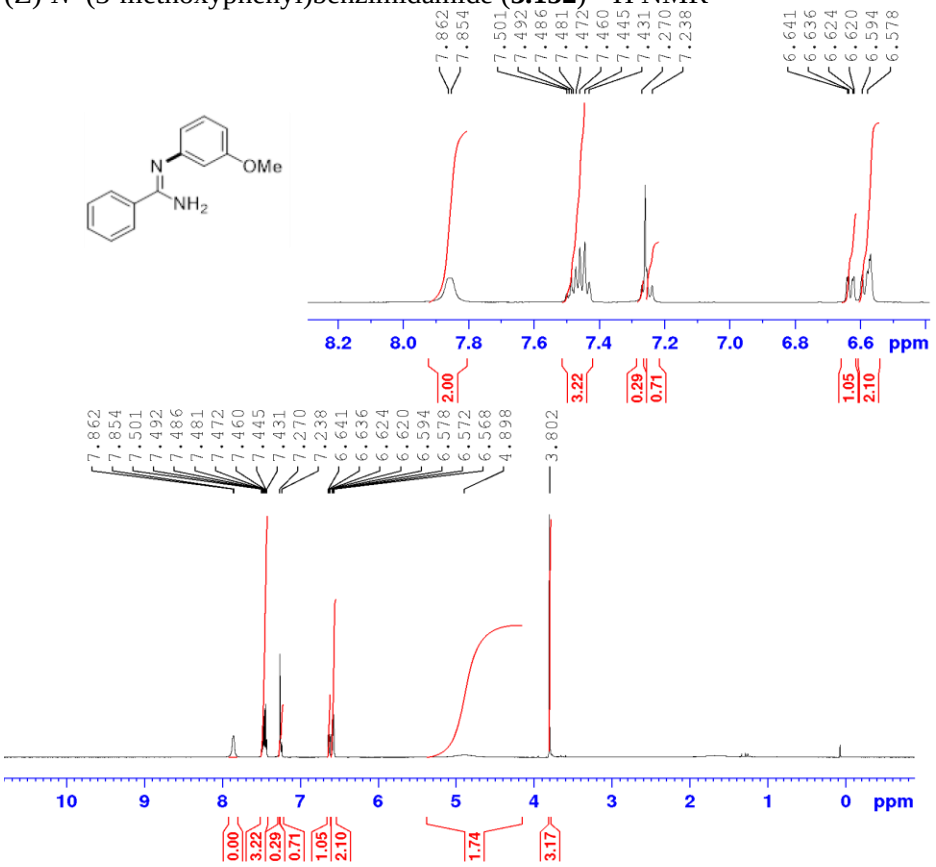
(Z)-*N*'-(4-(tert-butyl)phenyl)benzimidamide (**3.151**) -^{1}H NMR



^{13}C NMR

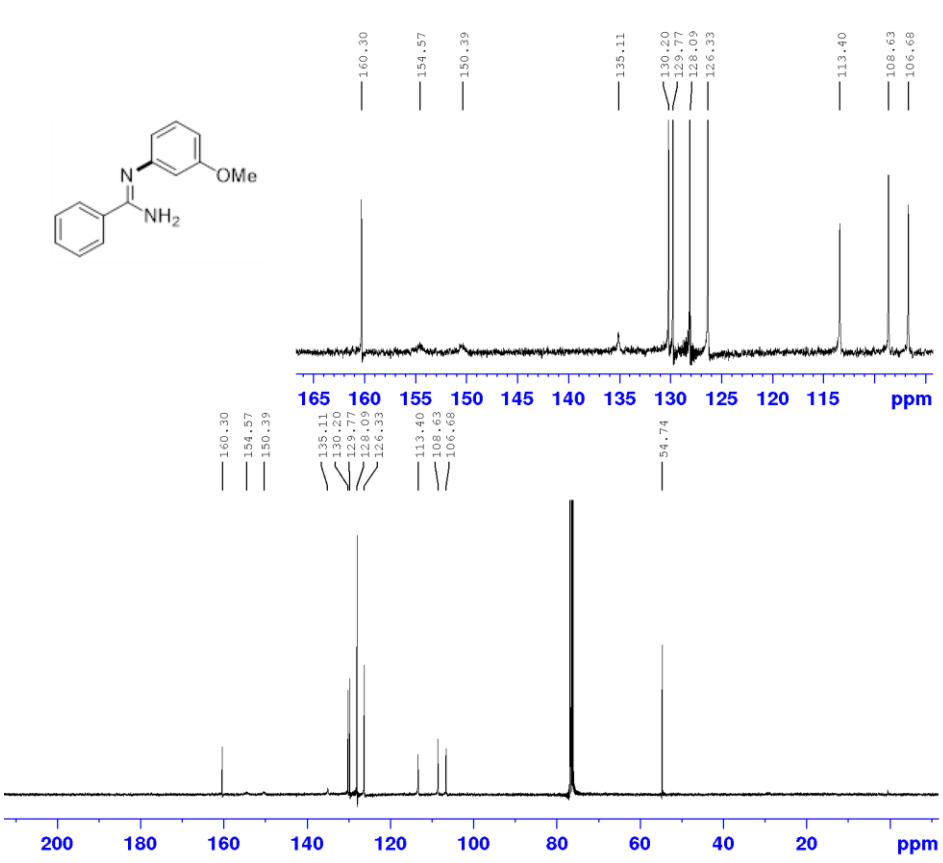


(Z)-*N*'-(3-methoxyphenyl)benzimidamide (**3.152**) -^{1}H NMR



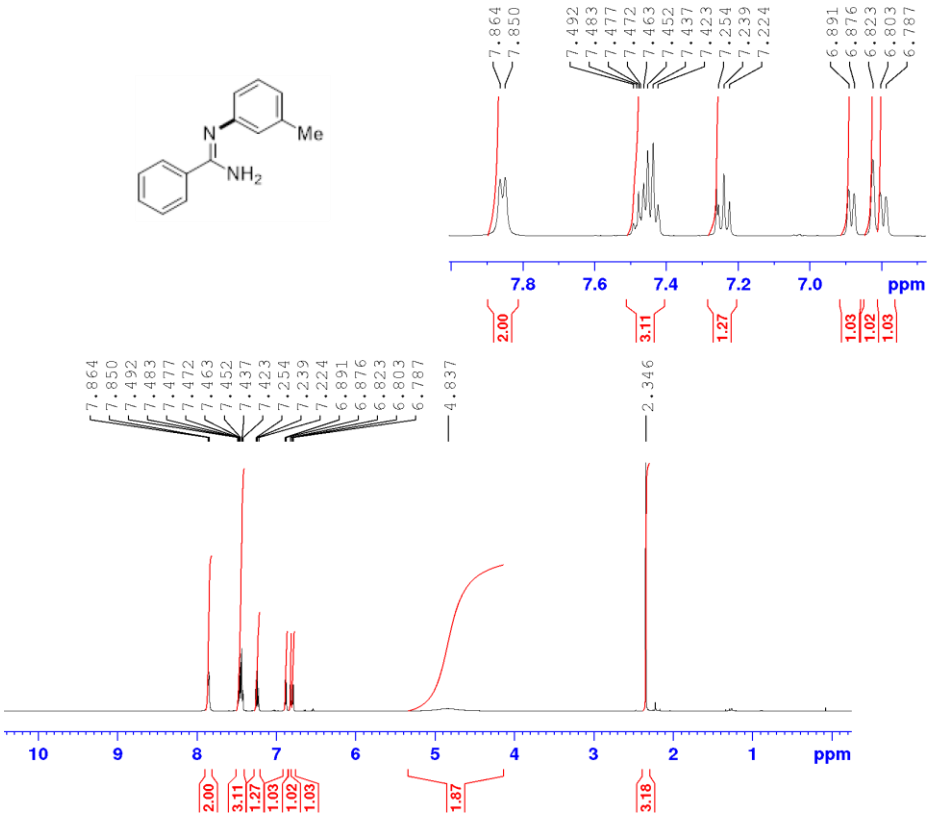

```
Current Data Parameters
NAME      JR0235B-E38729
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190106
Time              16.38 h
INSTRUM           spect
PROBHD    Z113652_0211 (
PULPROG           zg30
TD               39578
SOLVENT           CDCl3
NS                  16
DS                   2
SWH          10000.000 Hz
FIDRES           0.505331 Hz
AQ               1.9789000 sec
RG                 198.22
DW                50.000 usec
DE                 6.50 usec
TE                300.0 K
D1           2.00000000 sec
TD0                  1
SFO1         500.1330885 MHz
NUC1                 1H
P1                10.00 usec
PLW1         20.00000000 W

F2 - Processing parameters
SI                65536
SF           500.1300133 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

^{13}C NMR



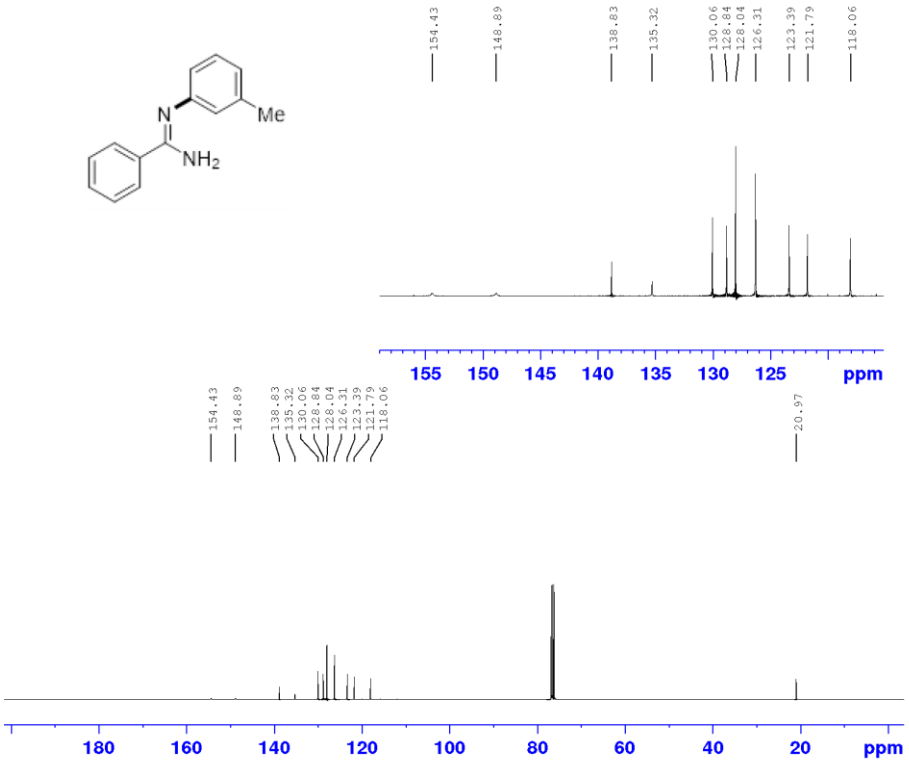

```
Current Data Parameters
NAME      JR0235 Carbon-D311552
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190113
Time              1.16
INSTRUM           spect
PROBHD    5 mm CPPBBO BB
PULPROG          zgpg30
TD                16384
SOLVENT           CDCl3
NS                 8192
DS                    4
SWH          25252.525 Hz
FIDRES           1.541292 Hz
AQ               0.3244032 sec
RG                 2050
DW                19.800 usec
DE                18.00 usec
TE                300.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                  1

======== CHANNEL f1 ========
SFO1         100.6228298 MHz
NUC1                 13C
P1                10.00 usec
PLW1         35.00000000 W

======== CHANNEL f2 ========
SFO2         400.1316005 MHz
NUC2                 1H
CPDPRG[2          waltz16
PCPD2             90.00 usec
PLW2          7.50000000 W
PLW12         0.16379000 W
PLW13         0.08238400 W

F2 - Processing parameters
SI                32768
SF           100.6128193 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

(Z)-*N*'-(m-tolyl)benzimidamide (**3.153**) -^{1}H NMR



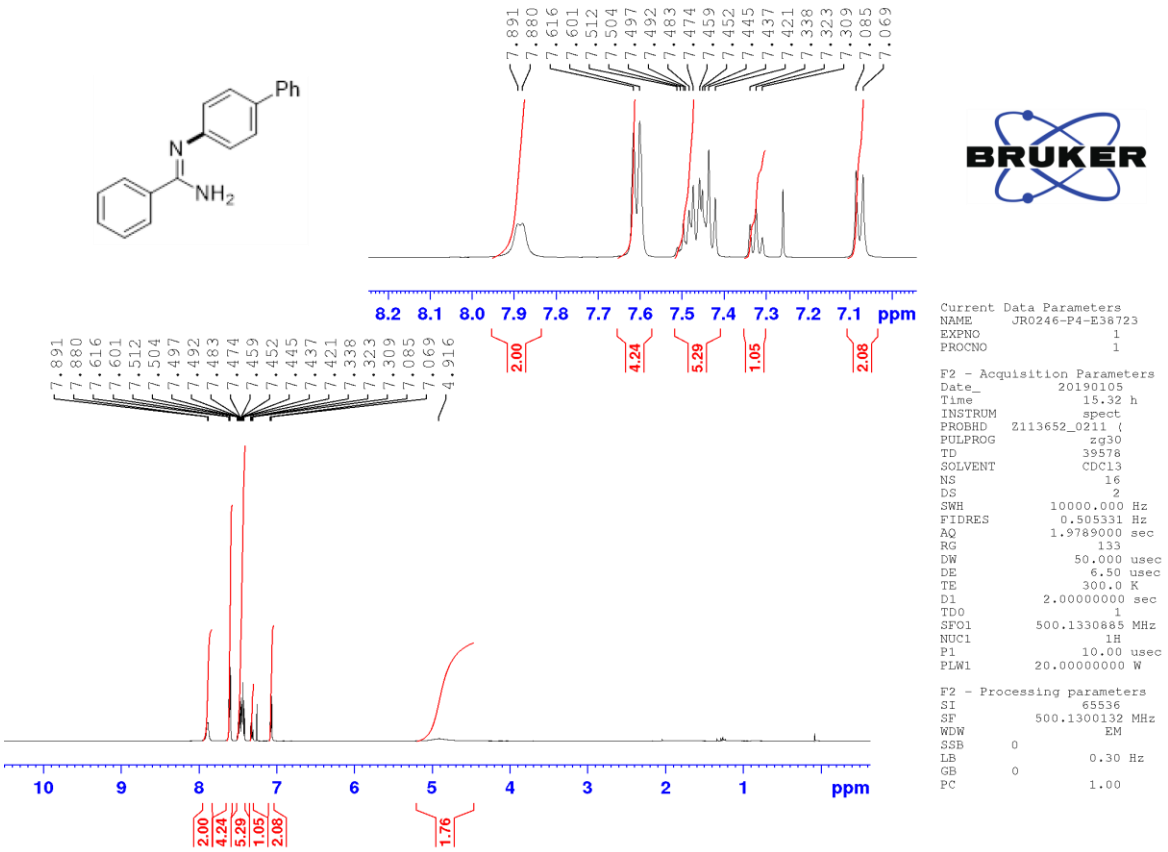
```
Current Data Parameters
NAME      JR0239B-E38731
EXPNO                1
PROCNO               1

F2 - Acquisition Parameters
Date_           20190106
Time              21.25 h
INSTRUM             spect
PROBHD    Z113652_0211 (
PULPROG             zg30
TD                 39578
SOLVENT            CDCl3
NS                    16
DS                     2
SWH           10000.000 Hz
FIDRES            0.505331 Hz
AQ              1.9789000 sec
RG                118.85
DW                50.000 usec
DE                 6.50 usec
TE                300.0 K
D1             2.00000000 sec
TD0                    1
SFO1          500.1330885 MHz
NUC1                  1H
P1                10.00 usec
PLW1          20.00000000 W

F2 - Processing parameters
SI                 65536
SF            500.1300132 MHz
WDW                   EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

^{13}C NMR



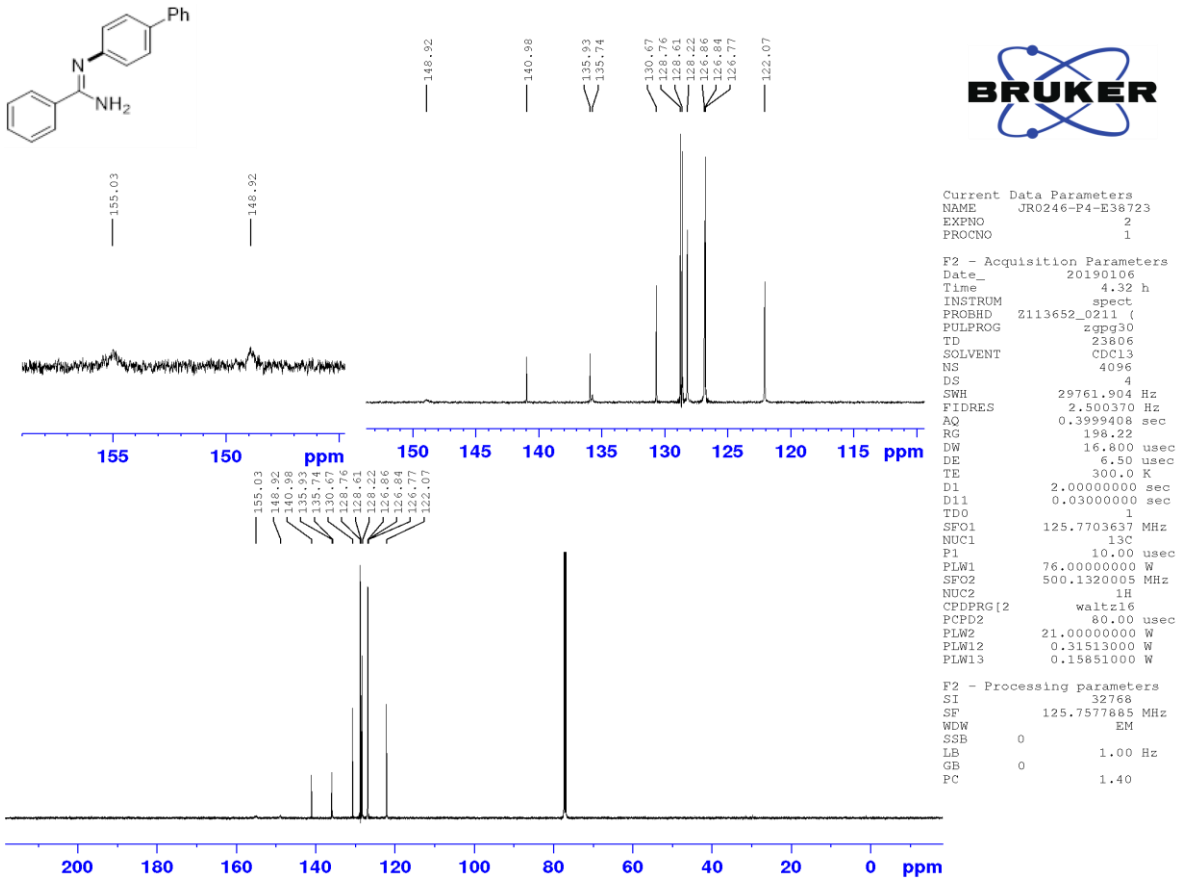
```
Current Data Parameters
NAME      JR0239 Carbon-D311789
EXPNO                1
PROCNO               1

F2 - Acquisition Parameters
Date_           20190116
Time               5.07
INSTRUM             spect
PROBHD   5 mm CPPBBO BB
PULPROG            zgpg30
TD                 16384
SOLVENT            CDCl3
NS                  8192
DS                     4
SWH           25252.525 Hz
FIDRES            1.541292 Hz
AQ              0.3244032 sec
RG                 2050
DW                19.800 usec
DE                18.00 usec
TE                300.0 K
D1             2.00000000 sec
D11            0.03000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1          100.6228298 MHz
NUC1                  13C
P1                10.00 usec
PLW1          35.00000000 W

======== CHANNEL f2 ========
SFO2          400.1316005 MHz
NUC2                  1H
CPDPRG[2          waltz16
PCPD2              90.00 usec
PLW2           7.50000000 W
PLW12          0.16379000 W
PLW13          0.08238400 W

F2 - Processing parameters
SI                 32768
SF            100.6128193 MHz
WDW                   EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

(Z)-*N*'-([1,1'-biphenyl]-4-yl)benzimidamide (**3.154**) - ^{1}H NMR



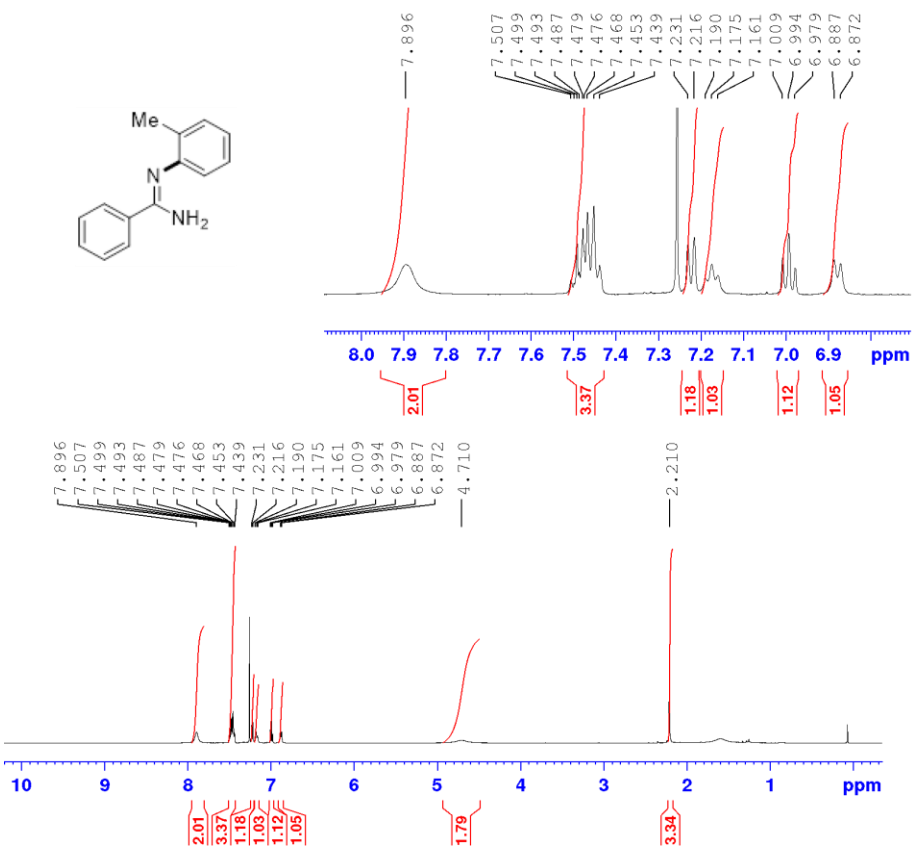

```
Current Data Parameters
NAME      JR0246-P4-E38723
EXPNO                1
PROCNO               1

F2 - Acquisition Parameters
Date_            20190105
Time              15.32 h
INSTRUM            spect
PROBHD   Z113652_0211 (
PULPROG            zg30
TD                39578
SOLVENT           CDCl3
NS                   16
DS                    2
SWH           10000.000 Hz
FIDRES            0.505331 Hz
AQ              1.9789000 sec
RG                  133
DW                50.000 usec
DE                 6.50 usec
TE                300.0 K
D1             2.00000000 sec
TD0                  1
SFO1          500.1330885 MHz
NUC1                 1H
P1                10.00 usec
PLW1          20.00000000 W

F2 - Processing parameters
SI                65536
SF            500.1300132 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

^{13}C NMR



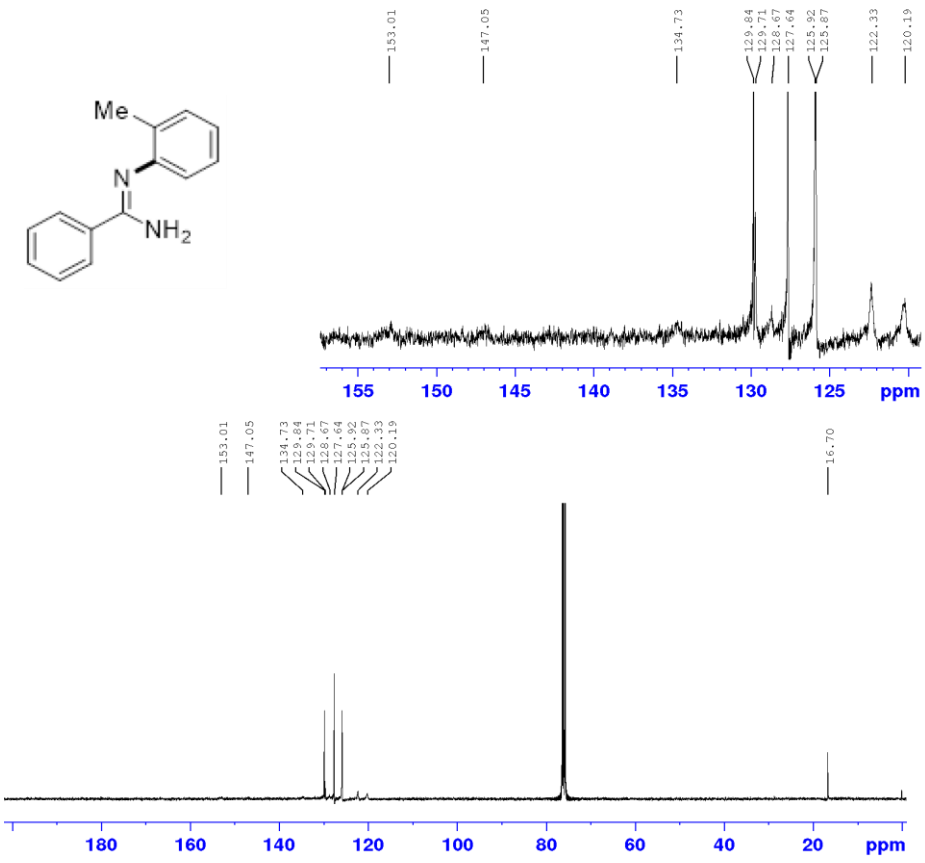

```
Current Data Parameters
NAME      JR0246-P4-E38723
EXPNO                2
PROCNO               1

F2 - Acquisition Parameters
Date_            20190106
Time              4.32 h
INSTRUM            spect
PROBHD   Z113652_0211 (
PULPROG          zgpg30
TD                23806
SOLVENT           CDCl3
NS                 4096
DS                    4
SWH           29761.904 Hz
FIDRES            2.500370 Hz
AQ              0.3999408 sec
RG                198.22
DW                16.800 usec
DE                 6.50 usec
TE                300.0 K
D1             2.00000000 sec
D11            0.03000000 sec
TD0                  1
SFO1          125.7703637 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          21.00000000 W
PLW12          0.31513000 W
PLW13          0.15851000 W

F2 - Processing parameters
SI                32768
SF            125.7577885 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

(Z)-*N*'-(*o*-tolyl)benzimidamide (**3.155**) - ^{1}H NMR



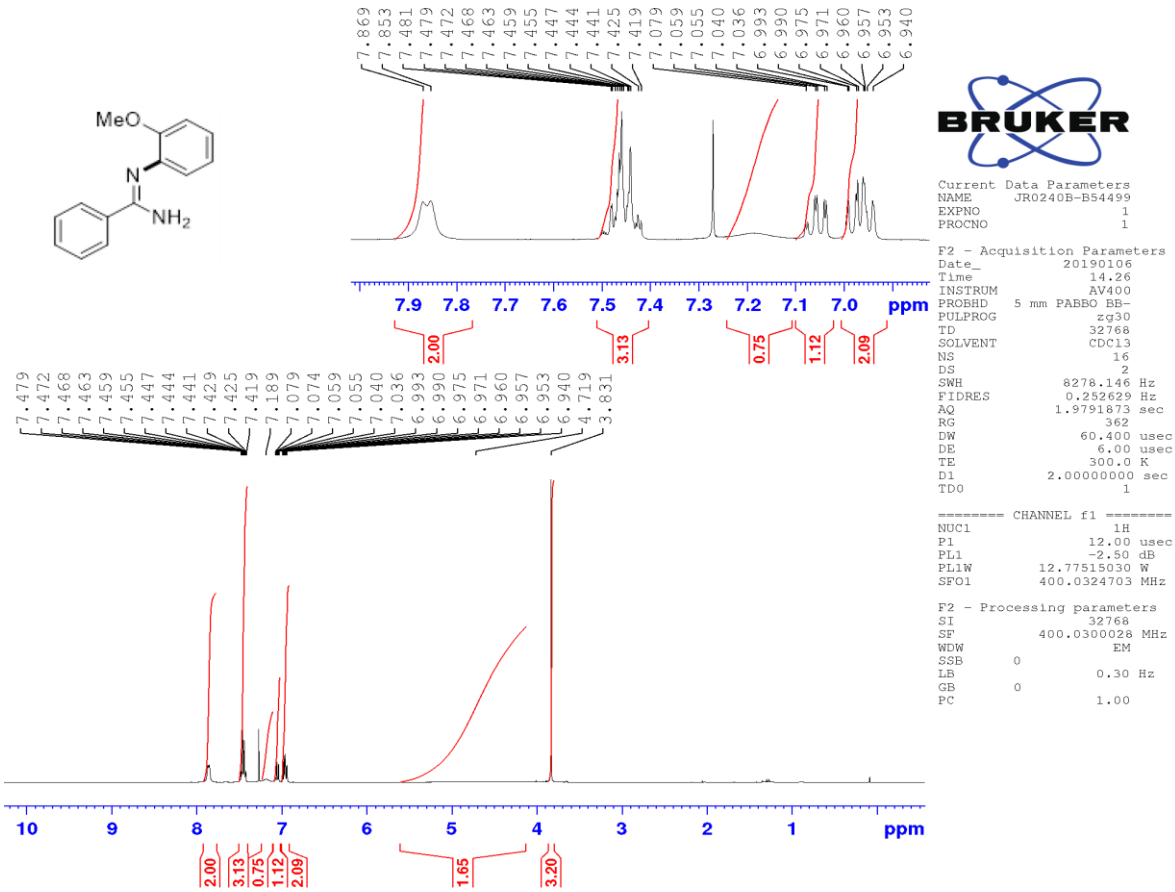

```
Current Data Parameters
NAME      JR0234B-E38728
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190106
Time              14.14 h
INSTRUM             spect
PROBHD    Z113652_0211 (
PULPROG             zg30
TD                 39578
SOLVENT             CDCl3
NS                    16
DS                     2
SWH           10000.000 Hz
FIDRES            0.505331 Hz
AQ              1.9789000 sec
RG                198.22
DW                 50.000 usec
DE                  6.50 usec
TE                300.0 K
D1              2.00000000 sec
TD0                    1
SFO1          500.1330885 MHz
NUC1                  1H
P1                 10.00 usec
PLW1          20.00000000 W

F2 - Processing parameters
SI                 65536
SF            500.1300145 MHz
WDW                   EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

^{13}C NMR




```
Current Data Parameters
NAME      JR0234 Carbon-D311551
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190112
Time               6.18
INSTRUM             spect
PROBHD   5 mm CPPBBO BB
PULPROG           zgpg30
TD                 16384
SOLVENT             CDCl3
NS                  8192
DS                     4
SWH           25252.525 Hz
FIDRES            1.541292 Hz
AQ              0.3244032 sec
RG                 2050
DW                19.800 usec
DE                18.00 usec
TE                300.0 K
D1              2.00000000 sec
D11             0.03000000 sec
TD0                    1

======== CHANNEL f1 ========
SFO1          100.6228298 MHz
NUC1                  13C
P1                 10.00 usec
PLW1          35.00000000 W

======== CHANNEL f2 ========
SFO2          400.1316005 MHz
NUC2                  1H
CPDPRG[2          waltz16
PCPD2              90.00 usec
PLW2           7.50000000 W
PLW12          0.16379000 W
PLW13          0.08238400 W

F2 - Processing parameters
SI                 32768
SF            100.6128661 MHz
WDW                   EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

(Z)-*N*'-(2-methoxyphenyl)benzimidamide (**3.156**) - ^{1}H NMR



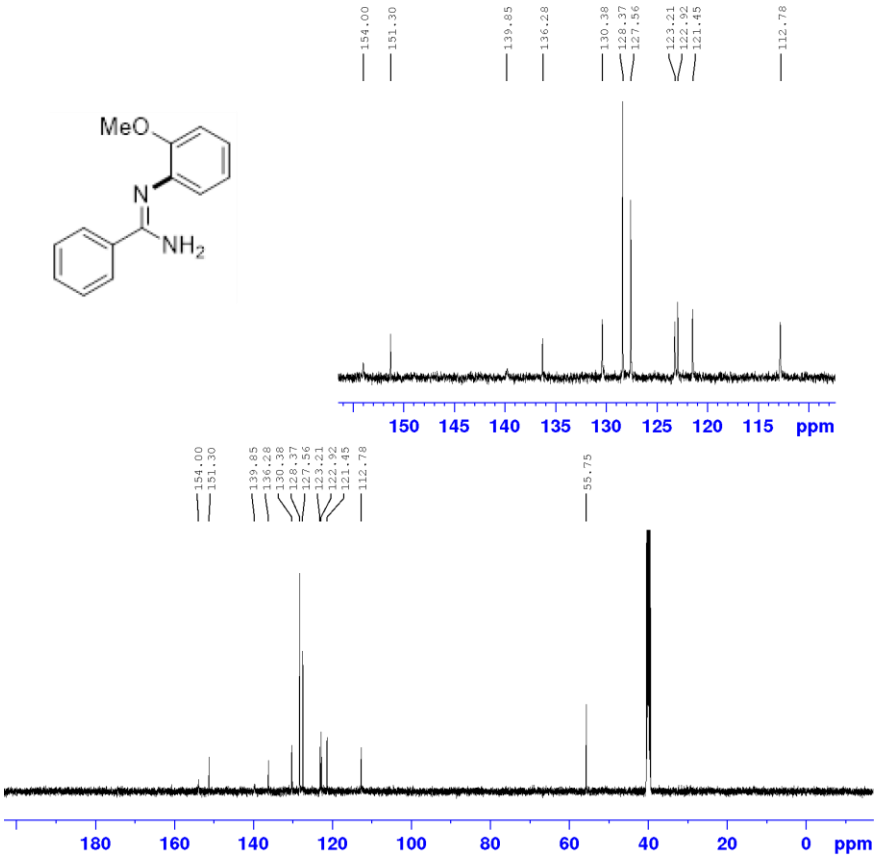
Current Data Parameters
NAME JR0240B-B54499
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190106
Time 14.26
INSTRUM AV400
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 32768
SOLVENT CDCl3
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.252629 Hz
AQ 1.9791873 sec
RG 362
DW 60.400 usec
DE 6.00 usec
TE 300.0 K
D1 2.00000000 sec
TD0 1

======== CHANNEL f1 ========
NUC1 1H
P1 12.00 usec
PL1 -2.50 dB
PL1W 12.77515030 W
SFO1 400.0324703 MHz

F2 - Processing parameters
SI 32768
SF 400.0300028 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

^{13}C NMR

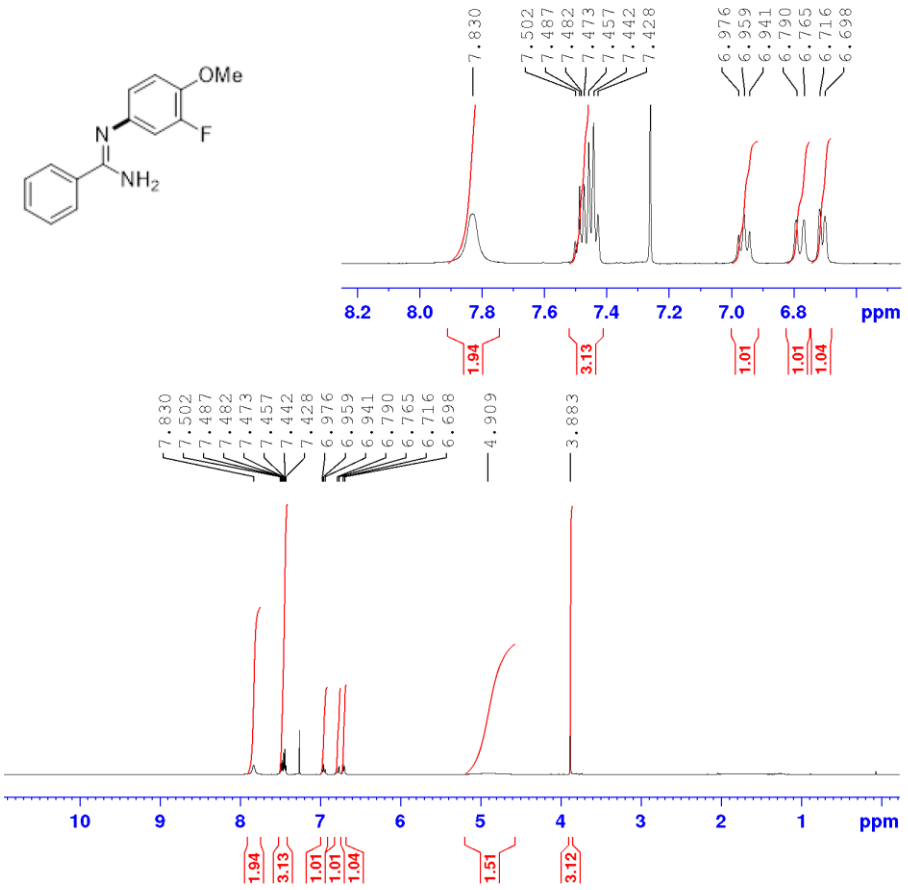


Current Data Parameters
NAME E38199-JR0240
EXPNO 2
PROCNO 1

F2 - Acquisition Parameters
Date_ 20181129
Time 23.16 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zgpg30
TD 23806
SOLVENT DMSO
NS 2048
DS 4
SWH 29761.904 Hz
FIDRES 2.500370 Hz
AQ 0.3999408 sec
RG 198.22
DW 16.800 usec
DE 6.50 usec
TE 300.0 K
D1 0.69999999 sec
D11 0.03000000 sec
TD0 1
SFO1 125.7703637 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 21.00000000 W
PLW12 0.31513000 W
PLW13 0.15851000 W

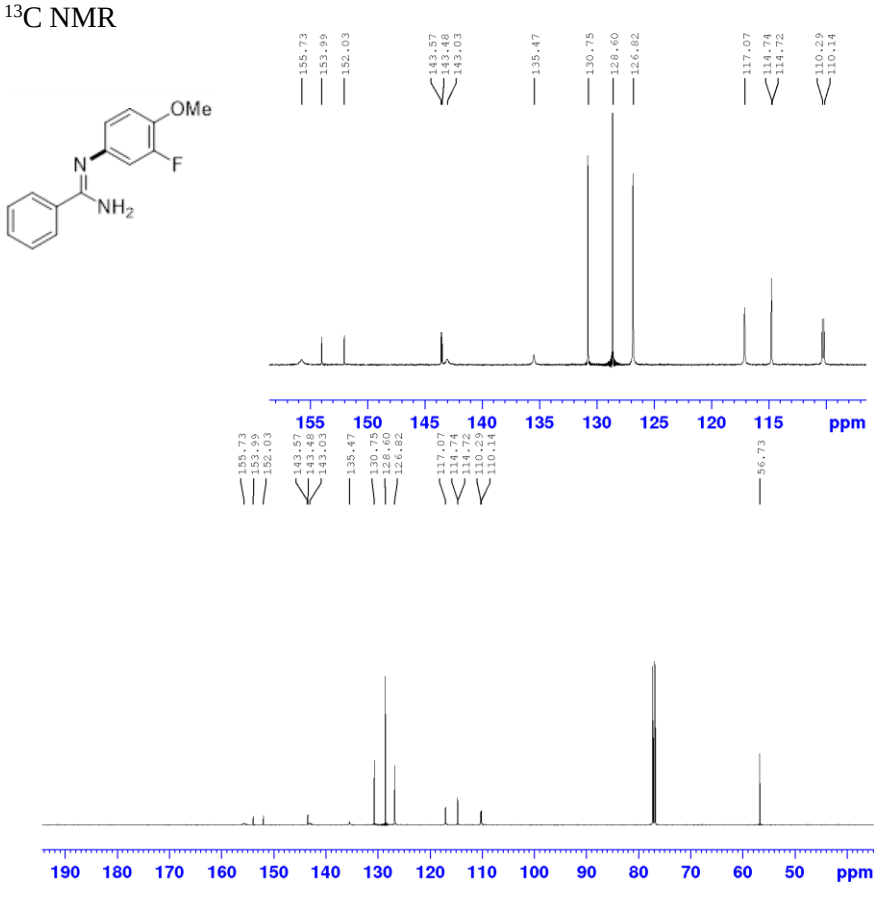
F2 - Processing parameters
SI 32768
SF 125.7577885 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

(*Z*)-*N*'-(3-fluoro-4-methoxyphenyl)benzimidamide (**3.148**) - ^{1}H NMR



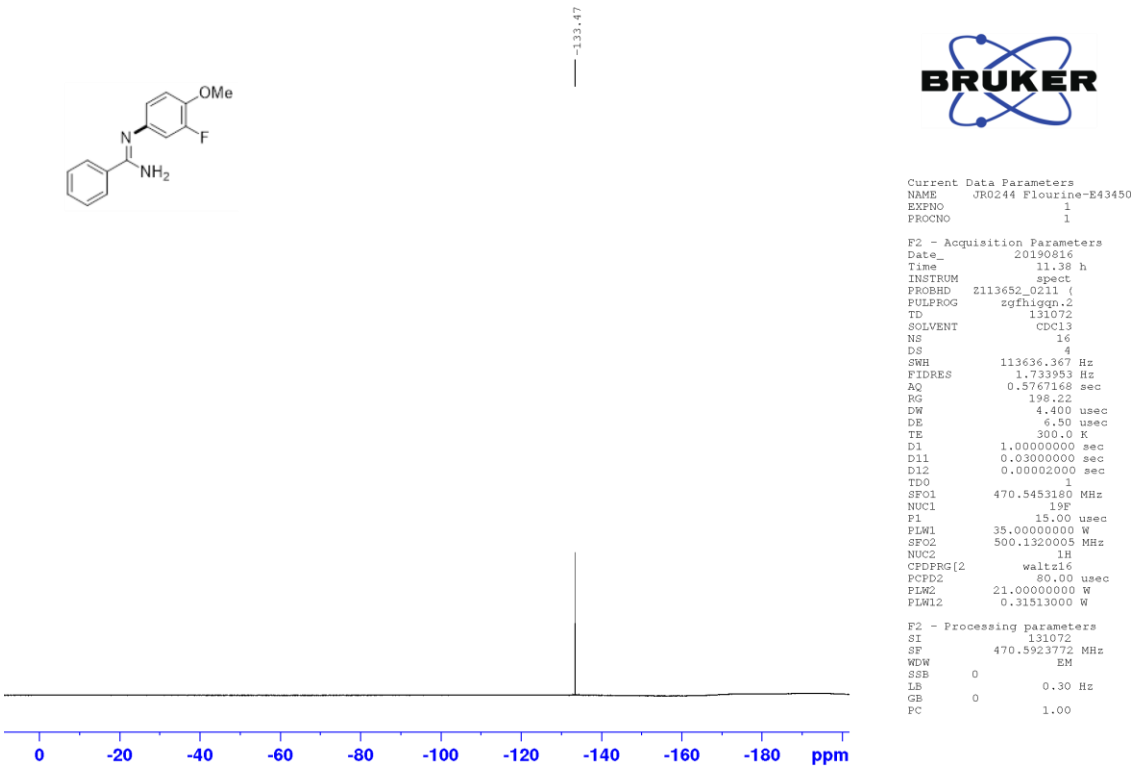
```
Current Data Parameters
NAME      JR0244B-E38720
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190104
Time               15.20 h
INSTRUM             spect
PROBHD    Z113652_0211 (
PULPROG             zg30
TD                 39578
SOLVENT            CDCl3
NS                    16
DS                     2
SWH           10000.000 Hz
FIDRES            0.505331 Hz
AQ              1.9789000 sec
RG                198.22
DW                 50.000 usec
DE                  6.50 usec
TE                300.0 K
D1              2.00000000 sec
TD0                    1
SFO1          500.1330885 MHz
NUC1                  1H
P1                 10.00 usec
PLW1          20.00000000 W

F2 - Processing parameters
SI                 65536
SF            500.1300133 MHz
WDW                   EM
SSB        0
LB                 0.30 Hz
GB         0
PC                 1.00
```

^{13}C NMR



```
Current Data Parameters
NAME      JR0244-Fluorine and Carbon E41184
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190414
Time               1.12 h
INSTRUM             spect
PROBHD    Z113652_0211 (
PULPROG          zgpg30
TD                 23806
SOLVENT            CDCl3
NS                  8192
DS                     4
SWH           29761.904 Hz
FIDRES            2.500370 Hz
AQ              0.3999408 sec
RG                198.22
DW                 16.800 usec
DE                  6.50 usec
TE                300.0 K
D1              2.00000000 sec
D11             0.03000000 sec
TD0                    1
SFO1          125.7703637 MHz
NUC1                  13C
P0                  3.33 usec
P1                 10.00 usec
PLW1          76.00000000 W
SFO2          500.1320005 MHz
NUC2                  1H
CPDPRG[2         waltz16
PCPD2              80.00 usec
PLW2          21.00000000 W
PLW12          0.31513000 W
PLW13          0.15851000 W

F2 - Processing parameters
SI                 32768
SF            125.7577885 MHz
WDW                   EM
SSB        0
LB                 1.00 Hz
GB         0
PC                 1.40
```

^{19}F NMR

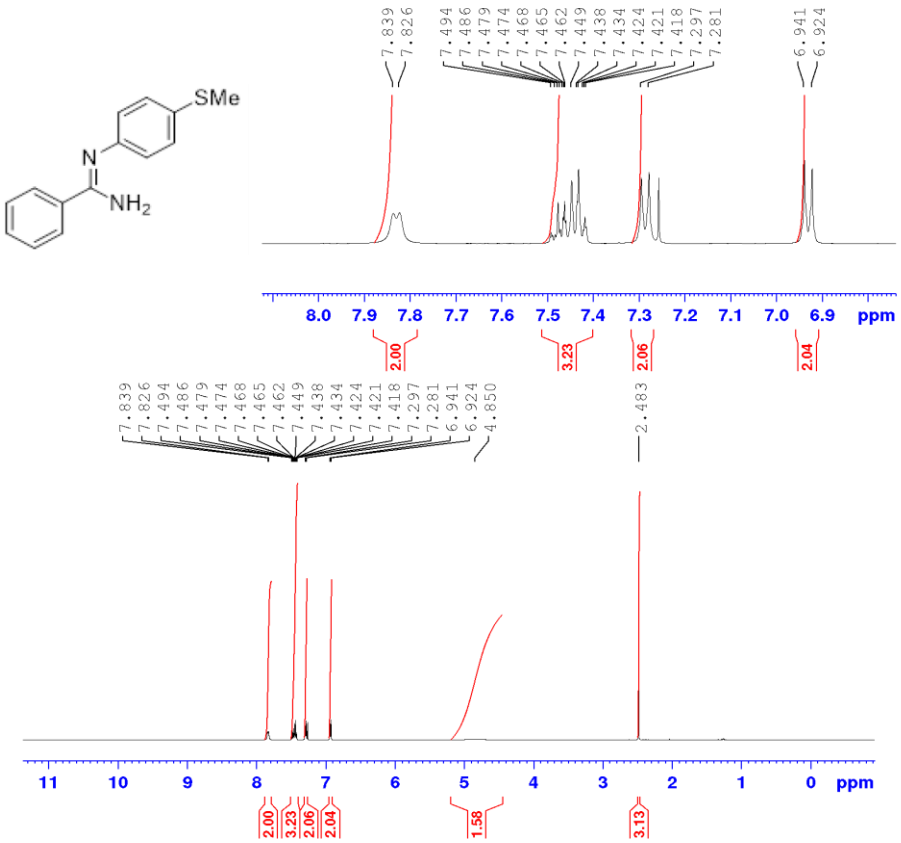


Current Data Parameters
NAME JR0244 Flourine-E43450
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190816
Time 11.38 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zgfhigqn.2
TD 131072
SOLVENT CDCl3
NS 16
DS 4
SWH 113636.367 Hz
FIDRES 1.733953 Hz
AQ 0.5767168 sec
RG 198.22
DW 4.400 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec
D11 0.03000000 sec
D12 0.00002000 sec
TD0 1
SFO1 470.5453180 MHz
NUC1 19F
P1 15.00 usec
PLW1 35.00000000 W
SFO2 500.1320005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 21.00000000 W
PLW12 0.31513000 W

F2 - Processing parameters
SI 131072
SF 470.5923772 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

(Z)-*N*'-(4-(methylthio)phenyl)benzimidamide (**3.150**) - ^{1}H NMR

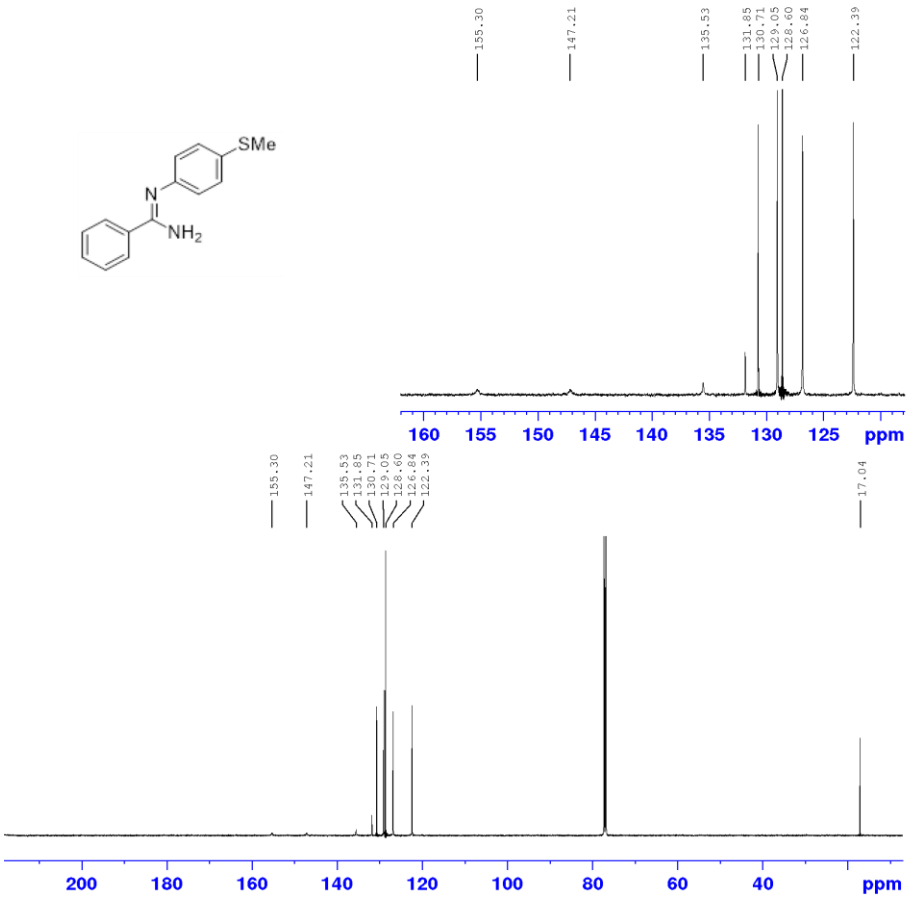


Current Data Parameters
NAME JR0278-E40878
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190329
Time 16.32 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zg30
TD 39578
SOLVENT CDCl3
NS 16
DS 2
SWH 10000.000 Hz
FIDRES 0.505331 Hz
AQ 1.9789000 sec
RG 118.65
DW 50.000 usec
DE 6.50 usec
TE 300.0 K
D1 2.00000000 sec
TD0 1
SFO1 500.1330885 MHz
NUC1 1H
P0 3.33 usec
P1 10.00 usec
PLW1 20.00000000 W

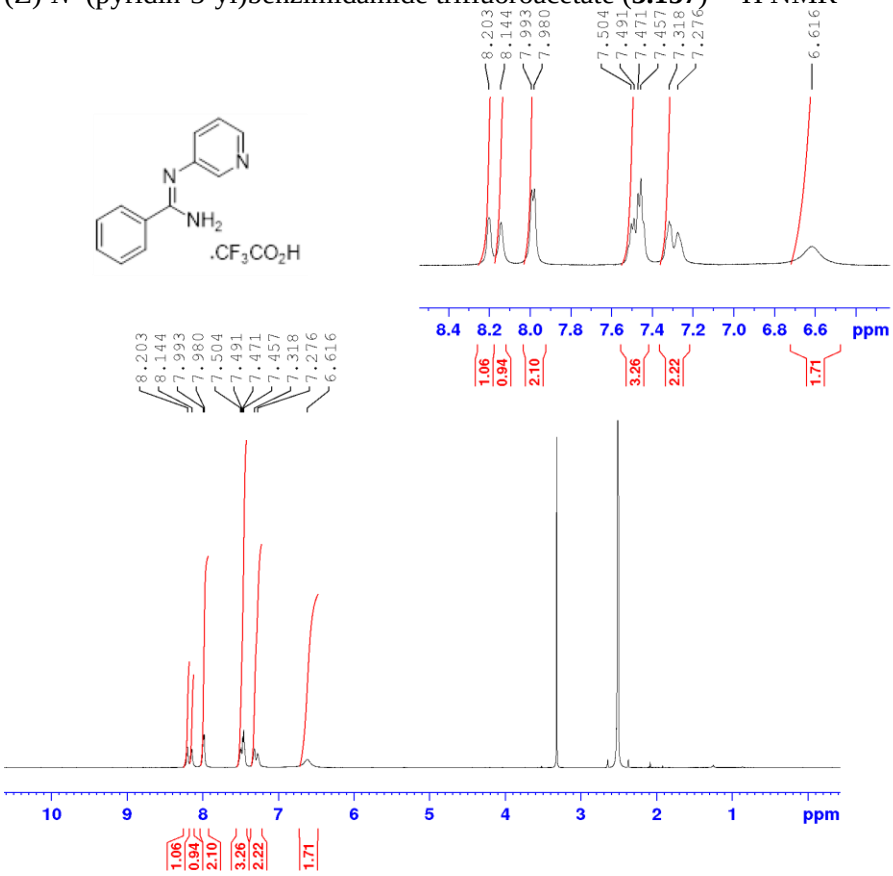
F2 - Processing parameters
SI 65536
SF 500.1300142 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

^{13}C NMR



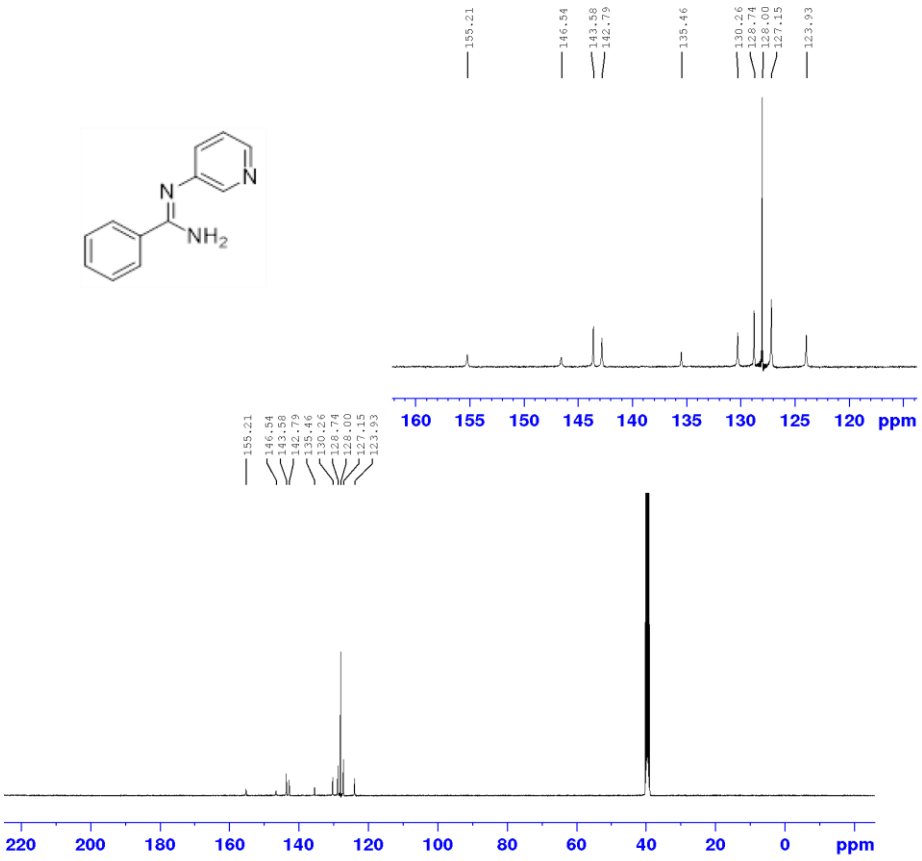
```
Current Data Parameters
NAME       JR0278-E40878
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190330
Time                2.28 h
INSTRUM            spect
PROBHD    Z113652_0211 (
PULPROG          zgpg30
TD                23806
SOLVENT           CDCl3
NS                 8192
DS                    4
SWH           29761.904 Hz
FIDRES         2.500370 Hz
AQ             0.3999408 sec
RG                198.22
DW                16.800 usec
DE                 6.50 usec
TE                299.9 K
D1            2.00000000 sec
D11           0.03000000 sec
TD0                    1
SFO1          125.7703637 MHz
NUC1                 13C
P0                 3.33 usec
P1                10.00 usec
PLW1          76.00000000 W
SFO2          500.1320005 MHz
NUC2                 1H
CPDPRG[2          waltz16
PCPD2              80.00 usec
PLW2          21.00000000 W
PLW12          0.31513000 W
PLW13          0.15851000 W

F2 - Processing parameters
SI                 32768
SF          125.7577885 MHz
WDW                   EM
SSB                    0
LB                 1.00 Hz
GB                     0
PC                 1.40
```

(Z)-*N*'-(pyridin-3-yl)benzimidamide trifluoroacetate (**3.157**) - ^{1}H NMR



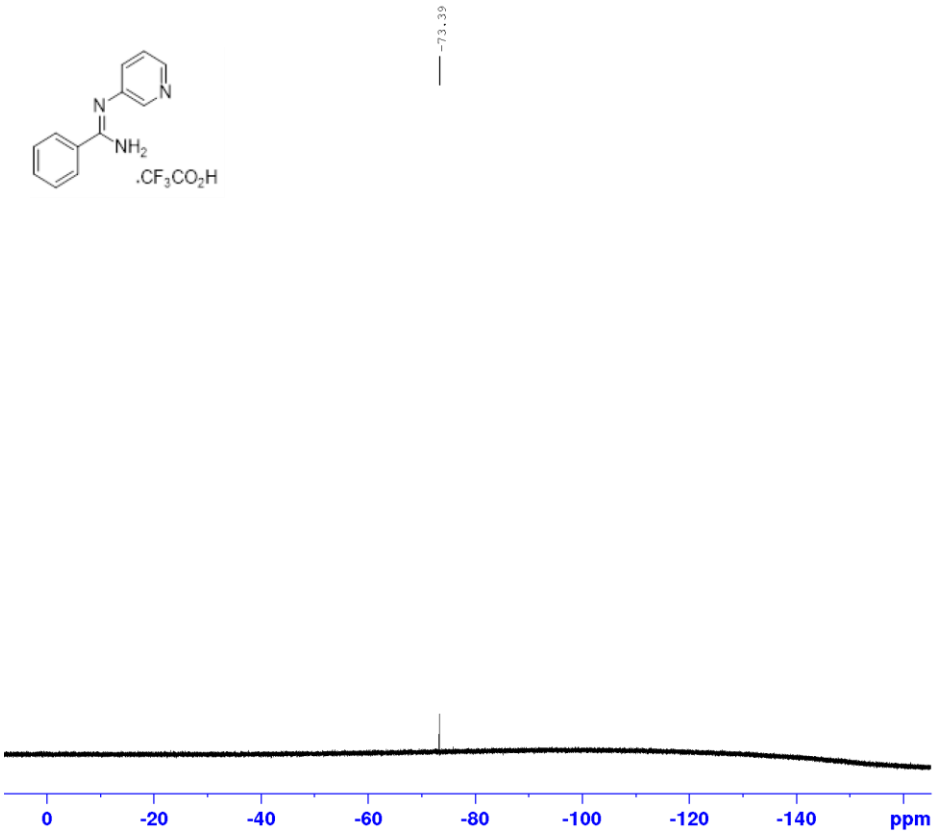
```
Current Data Parameters
NAME      JR0241proton-E41447
EXPNO                       1
PROCNO                      1

F2 - Acquisition Parameters
Date_            20190430
Time                17.25 h
INSTRUM            spect
PROBHD    Z113652_0211 (
PULPROG             zg30
TD                 39578
SOLVENT             DMSO
NS                    16
DS                     2
SWH           10000.000 Hz
FIDRES         0.505331 Hz
AQ             1.9789000 sec
RG                175.12
DW                50.000 usec
DE                 6.50 usec
TE                300.0 K
D1            2.00000000 sec
TD0                    1
SFO1          500.1330885 MHz
NUC1                   1H
P0                 3.33 usec
P1                10.00 usec
PLW1          20.00000000 W

F2 - Processing parameters
SI                 65536
SF          500.1300000 MHz
WDW                   EM
SSB                    0
LB                 0.30 Hz
GB                     0
PC                 1.00
```

^{13}C NMR



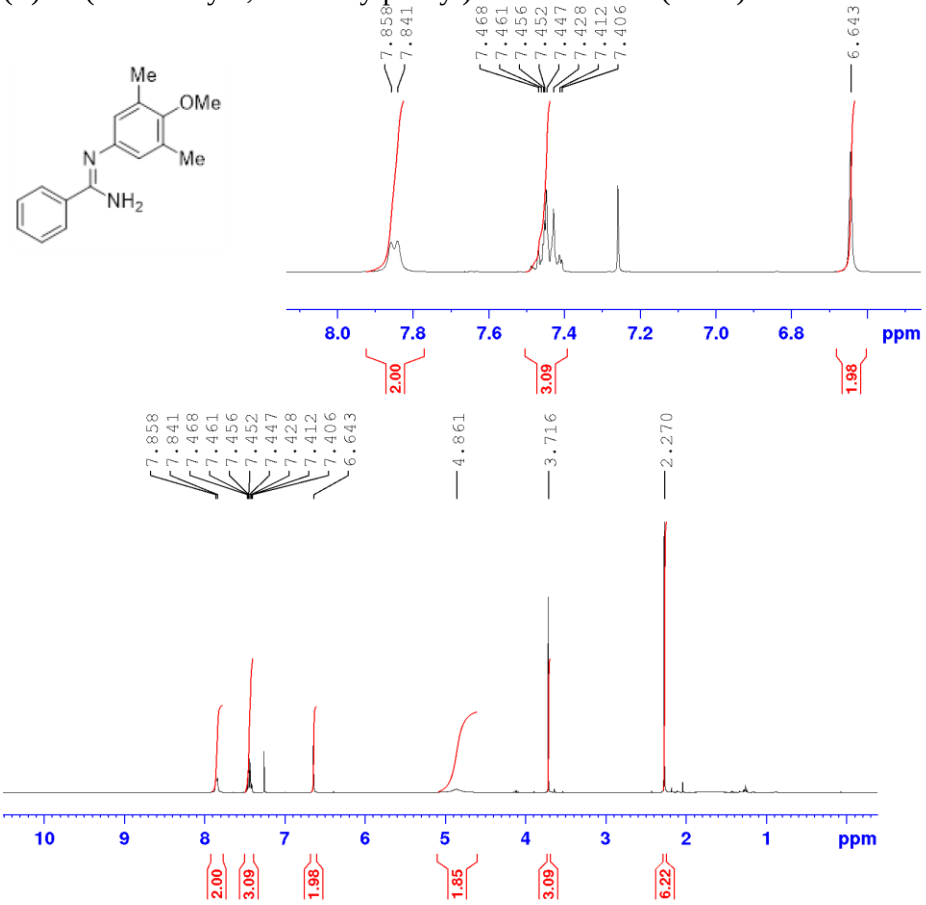

```
Current Data Parameters
NAME       JR0241 Carbon-D316994
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_           20190414
Time              22.27 h
INSTRUM             spect
PROBHD   Z122623_0021 (
PULPROG          zgpg30
TD                16384
SOLVENT            DMSO
NS                 8192
DS                    4
SWH           25252.525 Hz
FIDRES         3.082584 Hz
AQ            0.3244032 sec
RG                 2050
DW               19.800 usec
DE               18.00 usec
TE                300.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1
SFO1        100.6228298 MHz
NUC1                13C
P0                3.33 usec
P1               10.00 usec
PLW1        35.00000000 W
SFO2        400.1316005 MHz
NUC2                1H
CPDPRG[2         waltz16
PCPD2             90.00 usec
PLW2         7.50000000 W
PLW12        0.16379000 W
PLW13        0.08238400 W

F2 - Processing parameters
SI                32768
SF          100.6128193 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

^{19}F NMR



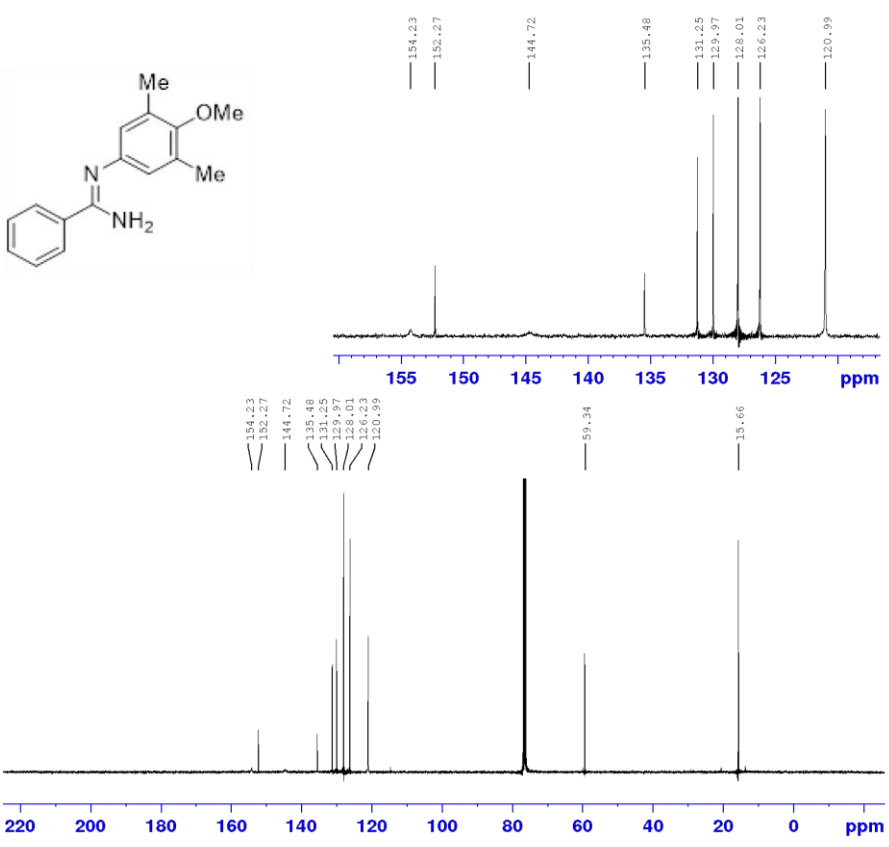

```
Current Data Parameters
NAME       JR0241 Fluorine-E43449
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_           20190816
Time              11.30 h
INSTRUM             spect
PROBHD   Z113652_0211 (
PULPROG       zgfhigqn.2
TD                131072
SOLVENT            DMSO
NS                   16
DS                    4
SWH          113636.367 Hz
FIDRES         1.733953 Hz
AQ            0.5767168 sec
RG                198.22
DW                4.400 usec
DE                6.50 usec
TE                300.0 K
D1           1.00000000 sec
D11          0.03000000 sec
D12          0.00002000 sec
TD0                   1
SFO1        470.5453180 MHz
NUC1                19F
P1               15.00 usec
PLW1        35.00000000 W
SFO2        500.1320005 MHz
NUC2                1H
CPDPRG[2         waltz16
PCPD2             80.00 usec
PLW2        21.00000000 W
PLW12        0.31513000 W

F2 - Processing parameters
SI                131072
SF          470.5923772 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

(Z)-*N*'-(4-methoxy-3,5-dimethylphenyl)benzimidamide (**3.158**) - ^{1}H NMR



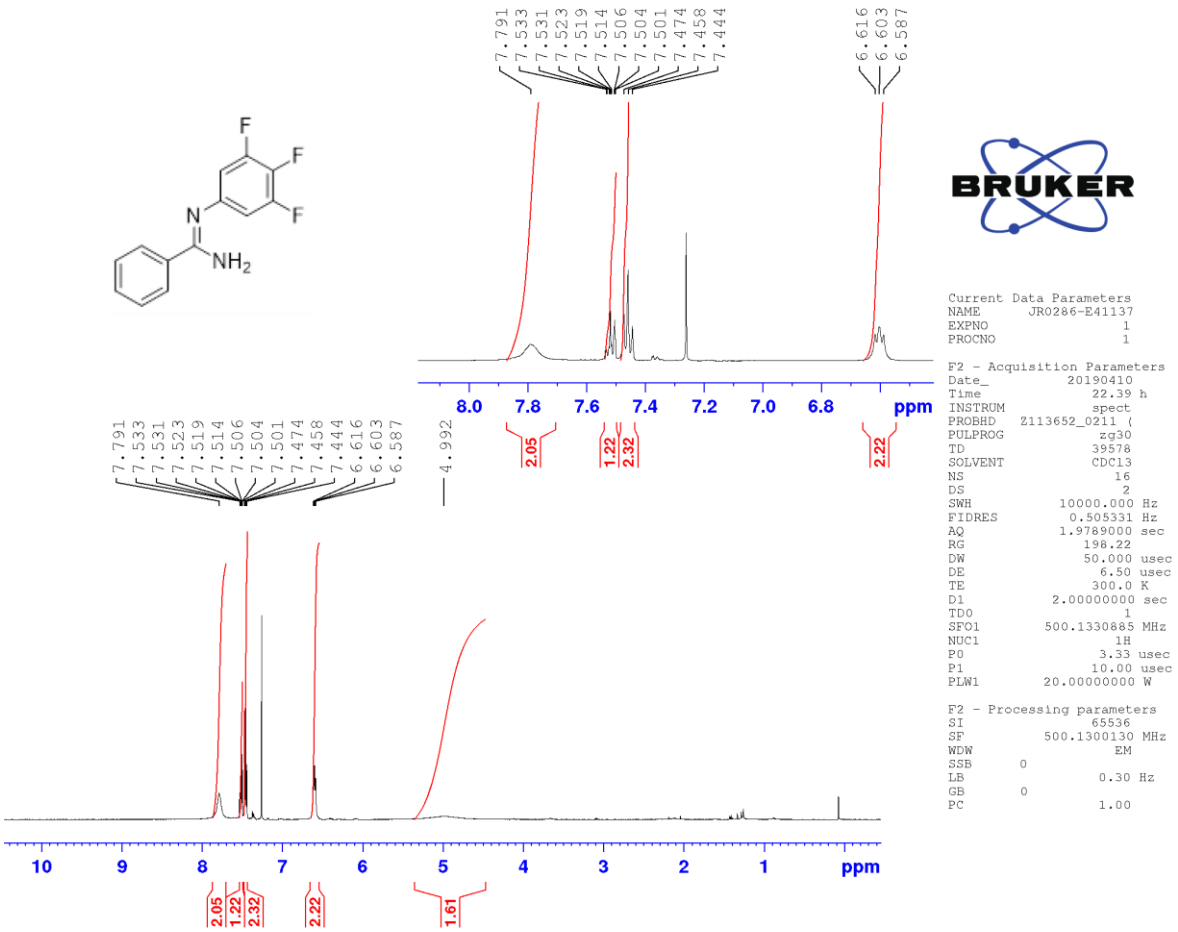

```
Current Data Parameters
NAME                  D321451
EXPNO                       1
PROCNO                      1

F2 - Acquisition Parameters
Date_              20190714
Time                  19.27 h
INSTRUM                 spect
PROBHD    Z122623_0021 (
PULPROG                 zg30
TD                     32564
SOLVENT                CDCl3
NS                          4
DS                          2
SWH                8223.685 Hz
FIDRES             0.505078 Hz
AQ                 1.9798912 sec
RG                        128
DW                   60.800 usec
DE                   10.00 usec
TE                   300.0 K
D1                 2.00000000 sec
TD0                         1
SFO1           400.1324710 MHz
NUC1                       1H
P0                    4.00 usec
P1                   12.00 usec
PLW1            7.19999981 W

F2 - Processing parameters
SI                      32768
SF             400.1300108 MHz
WDW                        EM
SSB          0
LB                    0.30 Hz
GB           0
PC                    1.00
```

^{13}C NMR

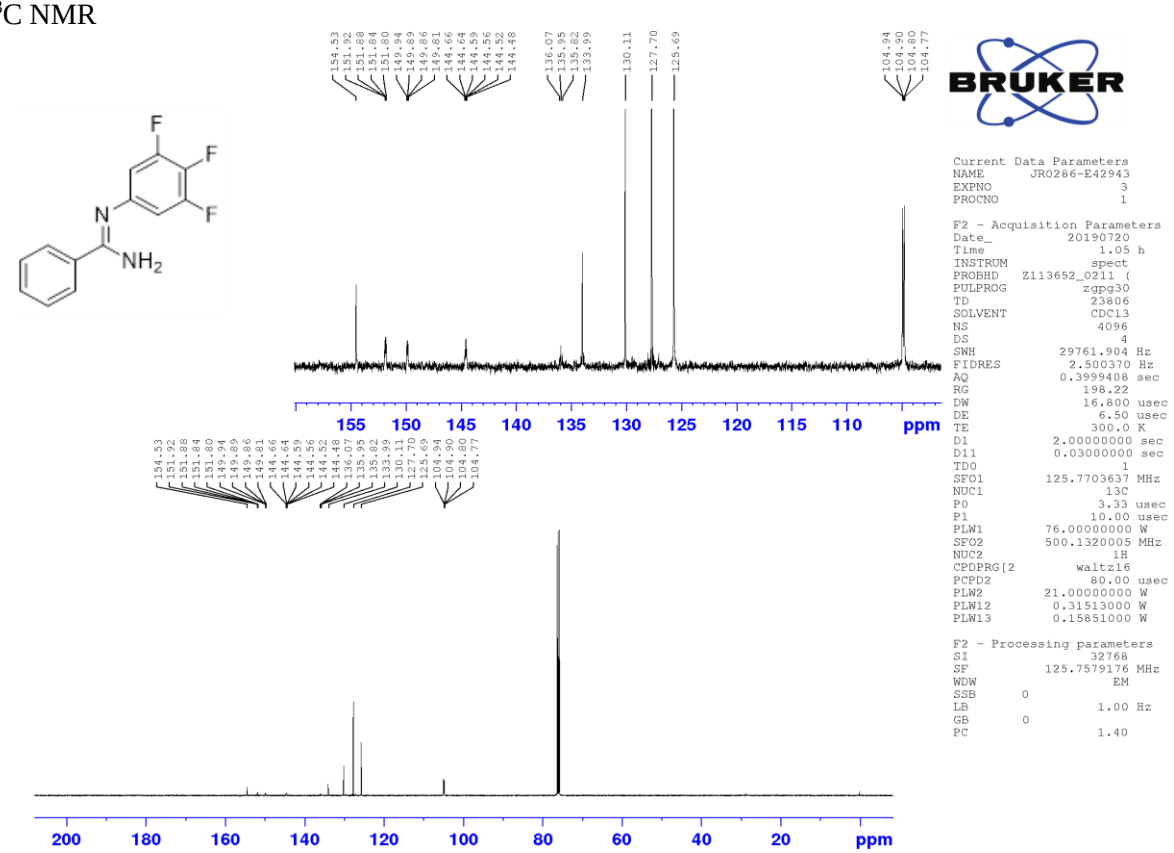



```
Current Data Parameters
NAME                  D321451
EXPNO                       3
PROCNO                      1

F2 - Acquisition Parameters
Date_              20190714
Time                  23.10 h
INSTRUM                 spect
PROBHD    Z122623_0021 (
PULPROG               zgpg30
TD                     16384
SOLVENT                CDCl3
NS                      4096
DS                         4
SWH               25252.525 Hz
FIDRES             3.082584 Hz
AQ                 0.3244032 sec
RG                      2050
DW                   19.800 usec
DE                   18.00 usec
TE                   300.0 K
D1                 2.00000000 sec
D11                0.03000000 sec
TD0                         1
SFO1           100.6228298 MHz
NUC1                      13C
P0                    3.33 usec
P1                   10.00 usec
PLW1           35.00000000 W
SFO2           400.1316005 MHz
NUC2                       1H
CPDPRG[2             waltz16
PCPD2                 90.00 usec
PLW2            7.50000000 W
PLW12           0.16379000 W
PLW13           0.08238400 W

F2 - Processing parameters
SI                      32768
SF             100.6128193 MHz
WDW                        EM
SSB          0
LB                    1.00 Hz
GB           0
PC                    1.40
```

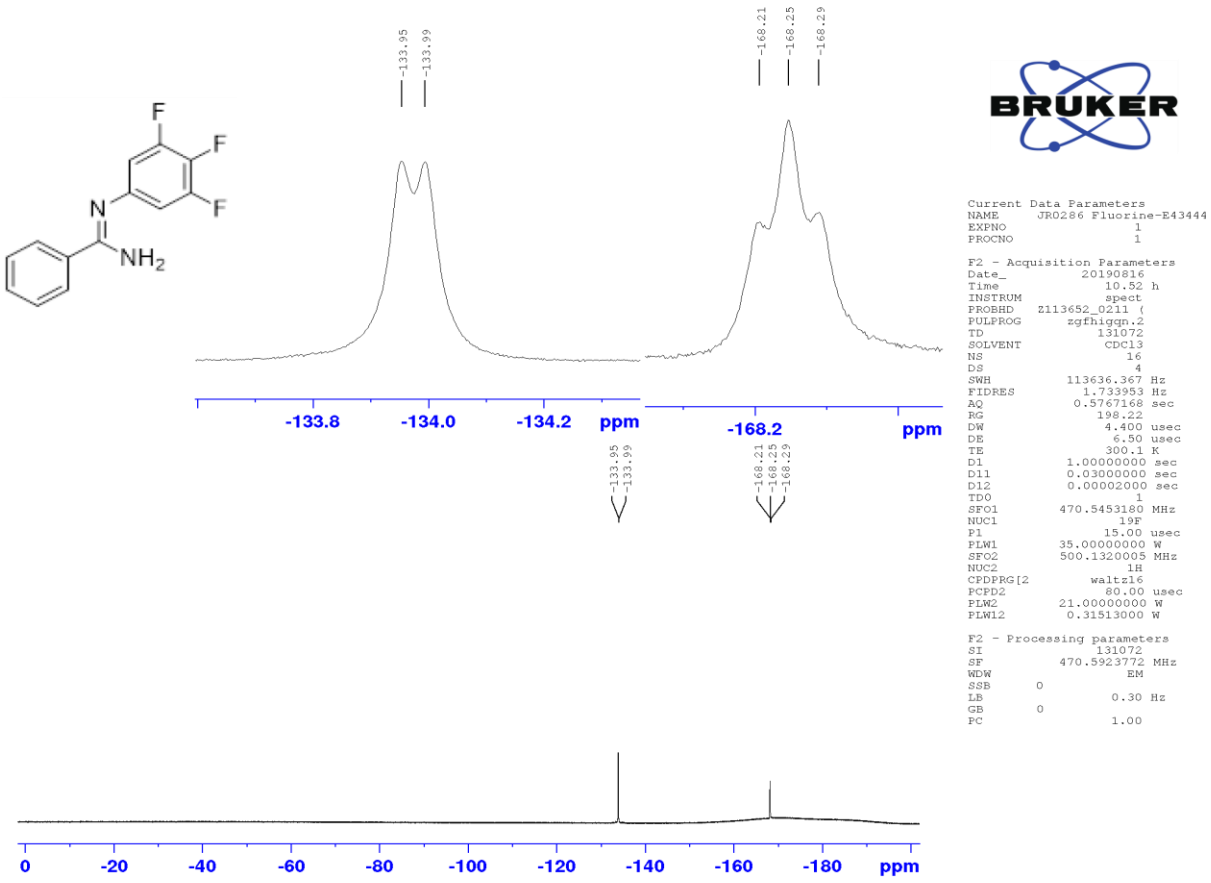
(Z)-*N*'-(3,4,5-trifluorophenyl)benzimidamide (**3.159**) - ^{1}H NMR



^{13}C NMR

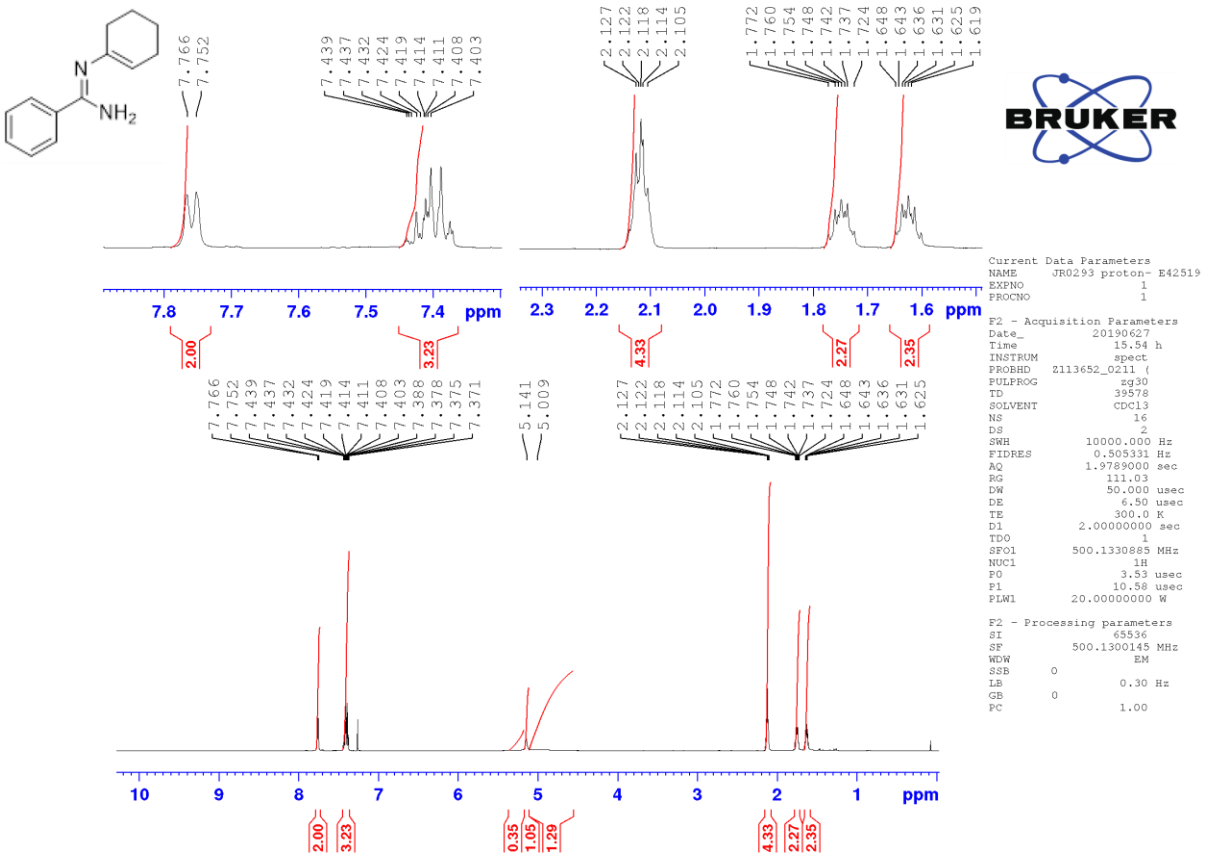


^{19}F NMR



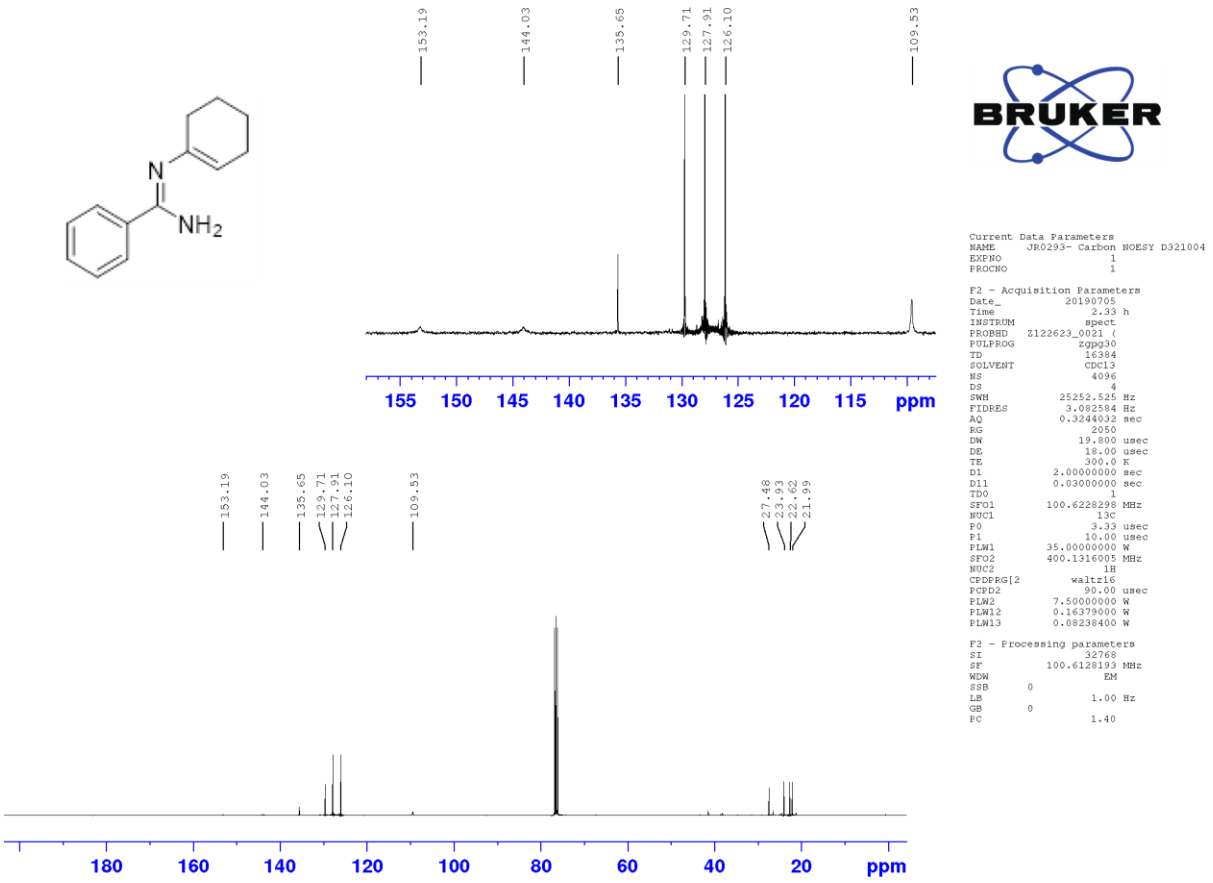

```
Current Data Parameters
NAME       JR0286 Fluorine-E43444
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_               20190816
Time                  10.52 h
INSTRUM                spect
PROBHD     Z113652_0211 (
PULPROG          zgfhigqn.2
TD                    131072
SOLVENT                CDCl3
NS                        16
DS                         4
SWH                113636.367 Hz
FIDRES               1.733953 Hz
AQ                   0.5767168 sec
RG                      198.22
DW                     4.400 usec
DE                     6.50 usec
TE                     300.1 K
D1                  1.00000000 sec
D11                 0.03000000 sec
D12                 0.00002000 sec
TD0                        1
SFO1               470.5453180 MHz
NUC1                      19F
P1                    15.00 usec
PLW1               35.00000000 W
SFO2               500.1320005 MHz
NUC2                      1H
CPDPRG[2               waltz16
PCPD2                 80.00 usec
PLW2               21.00000000 W
PLW12               0.31513000 W

F2 - Processing parameters
SI                    131072
SF                 470.5923772 MHz
WDW                       EM
SSB          0
LB                    0.30 Hz
GB           0
PC                    1.00
```

(*Z*)-*N*'-(cyclohex-1-en-1-yl)benzimidamide (**3.161**) - ^{1}H NMR

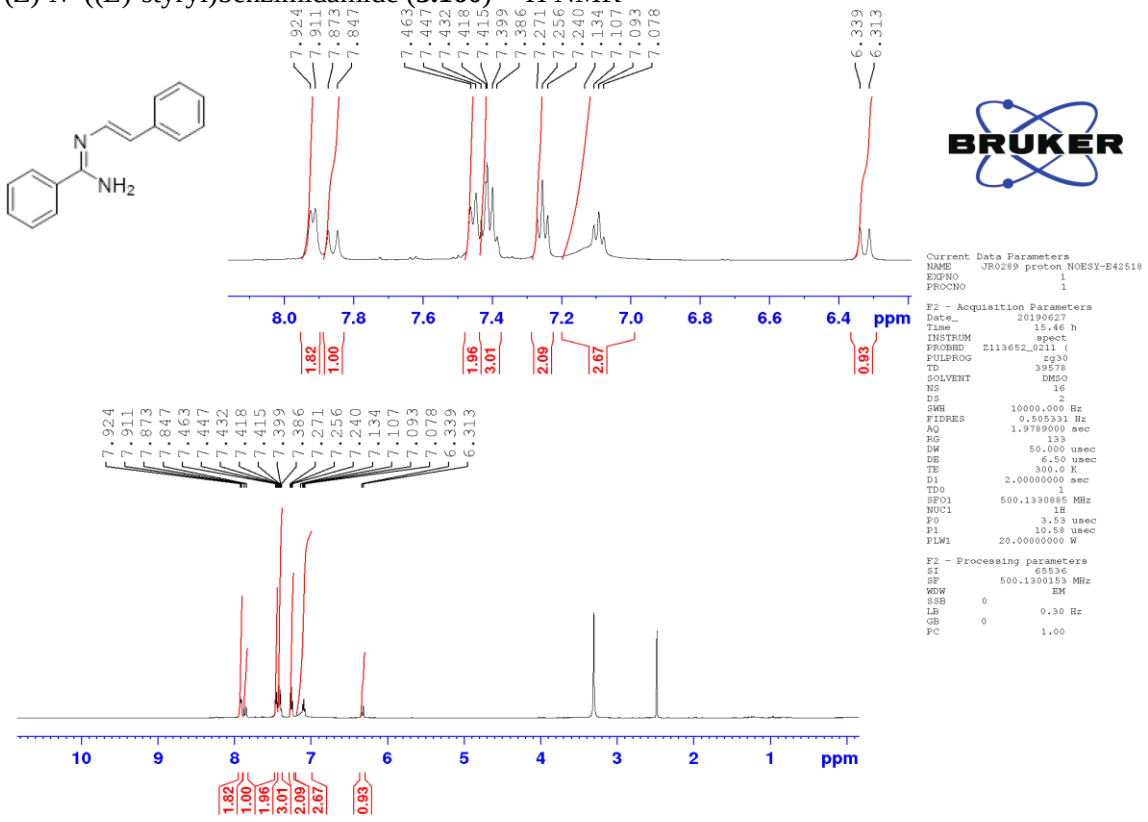



```
Current Data Parameters
NAME       JR0293 proton- E42519
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_               20190627
Time                  15.54 h
INSTRUM                spect
PROBHD     Z113652_0211 (
PULPROG                zg30
TD                     39578
SOLVENT                CDCl3
NS                        16
DS                         2
SWH                10000.000 Hz
FIDRES               0.505331 Hz
AQ                   1.9789000 sec
RG                      111.03
DW                    50.000 usec
DE                     6.50 usec
TE                     300.0 K
D1                  2.00000000 sec
TD0                        1
SFO1               500.1330885 MHz
NUC1                      1H
P0                     3.53 usec
P1                    10.58 usec
PLW1               20.00000000 W

F2 - Processing parameters
SI                     65536
SF                 500.1300145 MHz
WDW                       EM
SSB          0
LB                    0.30 Hz
GB           0
PC                    1.00
```

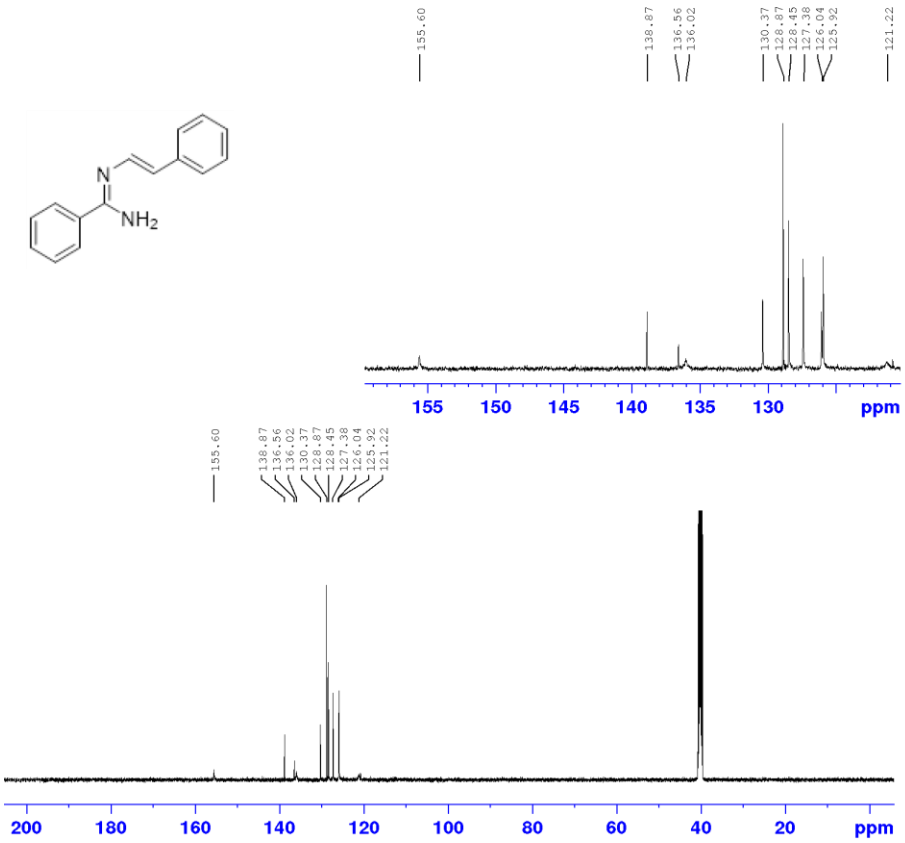
^{13}C NMR



(*Z*)-*N*'-((*E*)-styryl)benzimidamide (**3.160**) - ^{1}H NMR

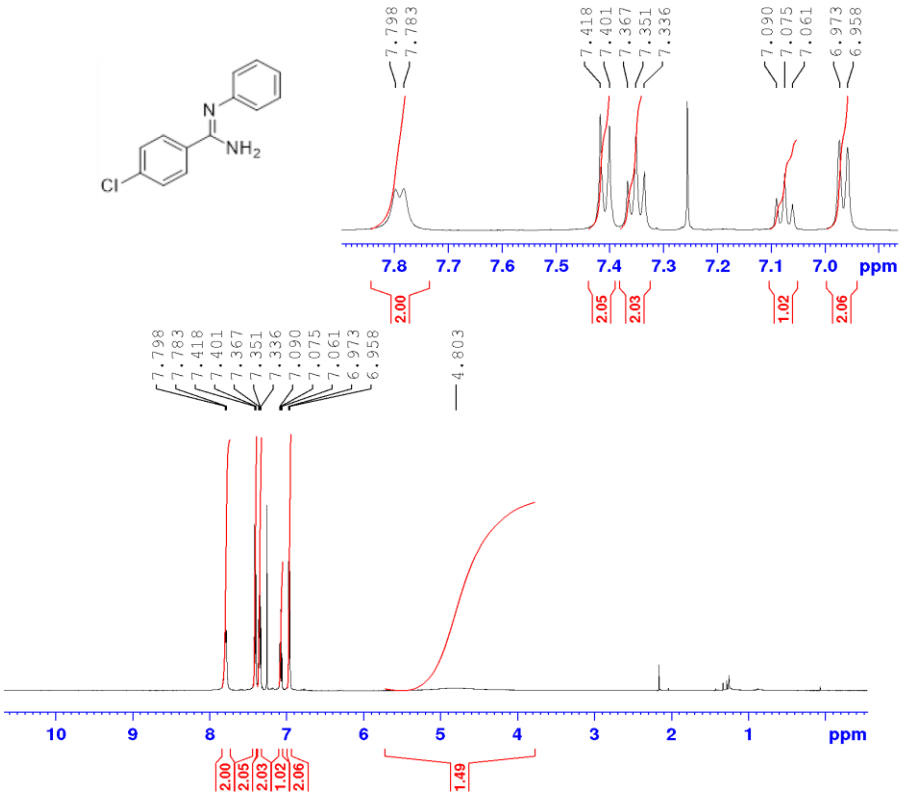


^{13}C NMR



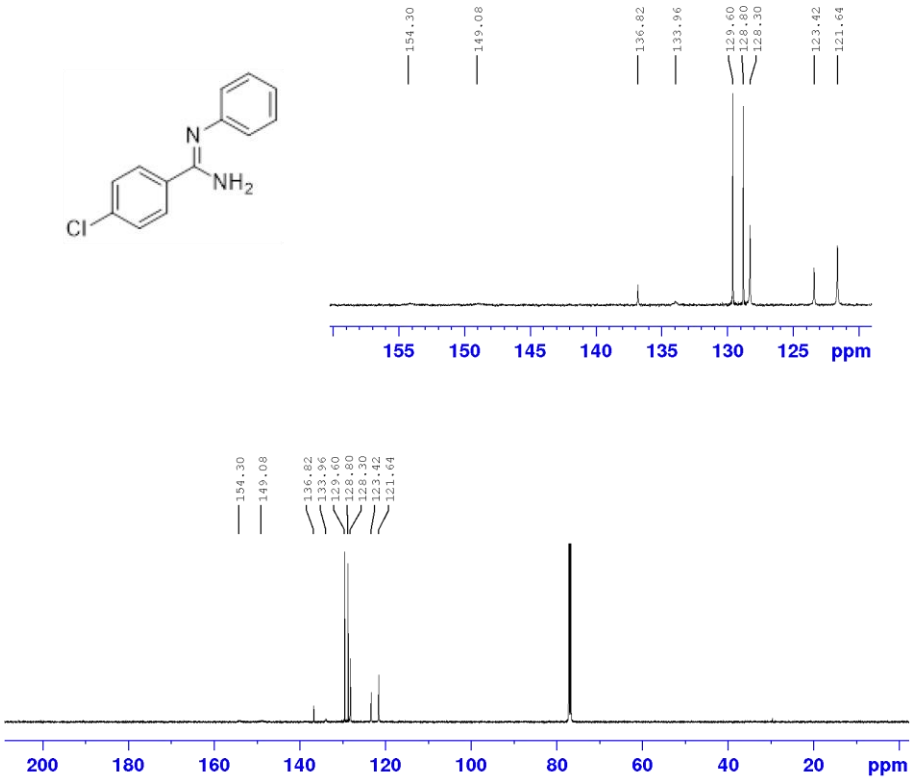

```
Current Data Parameters
NAME      JR0289-Carbon and HSQC E42663
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_          20190706
Time               1.54 h
INSTRUM            spect
PROBHD    Z113652_0211 (
PULPROG          zgpg30
TD                23806
SOLVENT            DMSO
NS                 2048
DS                    4
SWH           29761.904 Hz
FIDRES         2.500370 Hz
AQ             0.3999408 sec
RG                198.22
DW                16.800 usec
DE                 6.50 usec
TE                300.0 K
D1             3.00000000 sec
D11            0.03000000 sec
TD0                   1
SFO1          125.7703637 MHz
NUC1                 13C
P0                 3.33 usec
P1                10.00 usec
PLW1          76.00000000 W
SFO2          500.1320005 MHz
NUC2                 1H
CPDPRG[2          waltz16
PCPD2             80.00 usec
PLW2          21.00000000 W
PLW12          0.31513000 W
PLW13          0.15851000 W

F2 - Processing parameters
SI                32768
SF            125.7577895 MHz
WDW                  EM
SSB                   0
LB                 1.00 Hz
GB                    0
PC                 1.40
```

(Z)-4-chloro-*N*'-phenylbenzimidamide (**3.164**) - ^{1}H NMR



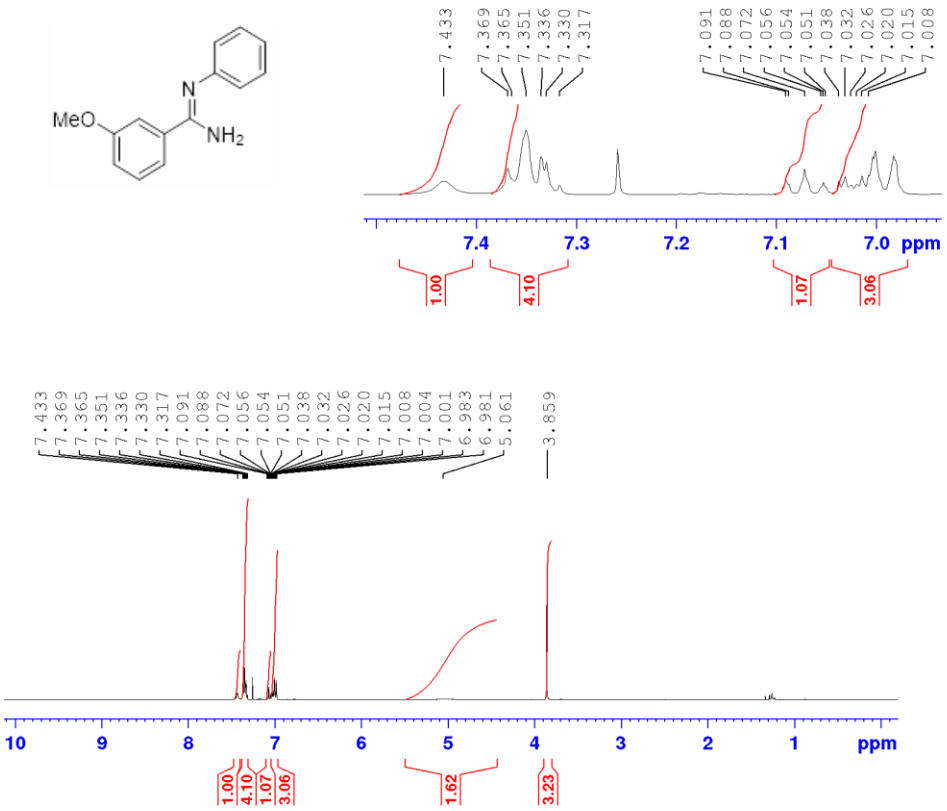

```
Current Data Parameters
NAME          JR0273-E40761
EXPNO                     1
PROCNO                    1

F2 - Acquisition Parameters
Date_              20190322
Time                 17.46 h
INSTRUM              spect
PROBHD    Z113652_0211 (
PULPROG              zg30
TD                  39578
SOLVENT             CDCl3
NS                     16
DS                      2
SWH            10000.000 Hz
FIDRES             0.505331 Hz
AQ             1.9789000 sec
RG                  175.12
DW                 50.000 usec
DE                  6.50 usec
TE                 300.0 K
D1             2.00000000 sec
TD0                     1
SFO1           500.1330885 MHz
NUC1                   1H
P0                  3.33 usec
P1                 10.00 usec
PLW1           20.00000000 W

F2 - Processing parameters
SI                  65536
SF             500.1300152 MHz
WDW                    EM
SSB                     0
LB                   0.30 Hz
GB                      0
PC                   1.00
```

^{13}C NMR



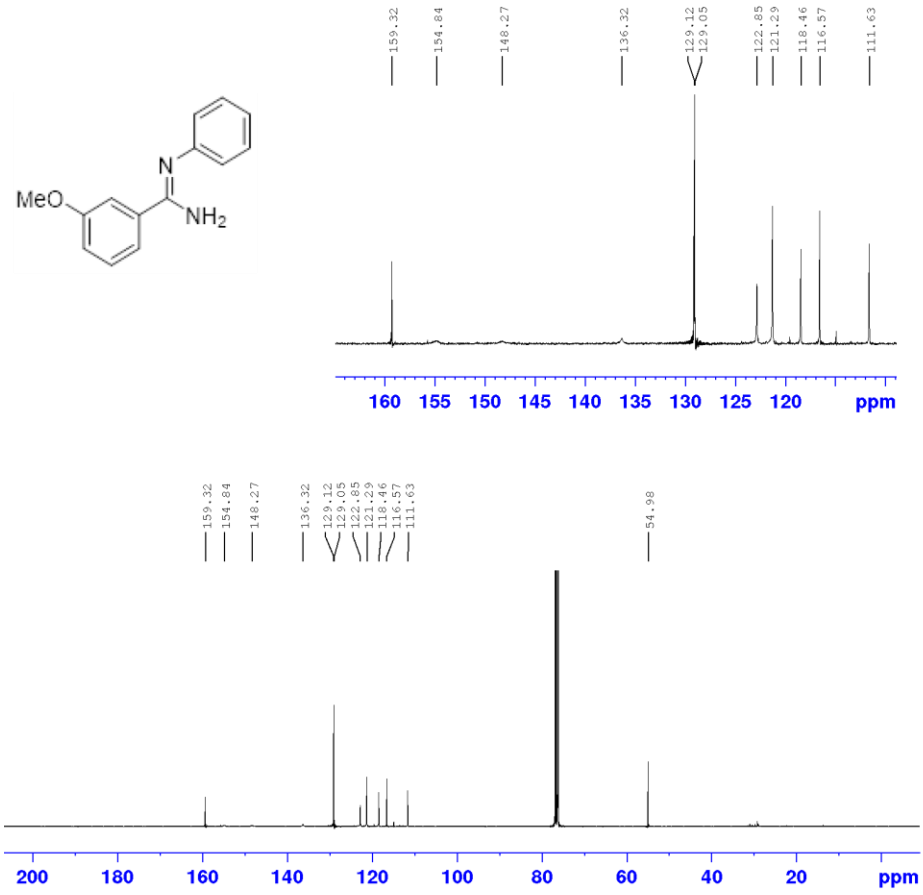

```
Current Data Parameters
NAME       JR0273-E40761
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190323
Time              3.41 h
INSTRUM            spect
PROBHD   Z113652_0211 (
PULPROG          zgpg30
TD                23806
SOLVENT           CDCl3
NS                 8192
DS                    4
SWH           29761.904 Hz
FIDRES         2.500370 Hz
AQ             0.3999408 sec
RG               198.22
DW              16.800 usec
DE               6.50 usec
TE              299.9 K
D1            2.00000000 sec
D11           0.03000000 sec
TD0                   1
SFO1         125.7703637 MHz
NUC1               13C
P0               3.33 usec
P1              10.00 usec
PLW1         76.00000000 W
SFO2         500.1320005 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2            80.00 usec
PLW2         21.00000000 W
PLW12         0.31513000 W
PLW13         0.15851000 W

F2 - Processing parameters
SI                32768
SF           125.7577885 MHz
WDW                  EM
SSB       0
LB              1.00 Hz
GB        0
PC              1.40
```

(*Z*)-3-methoxy-*N*'-phenylbenzimidamide (**3.165**) - ^{1}H NMR



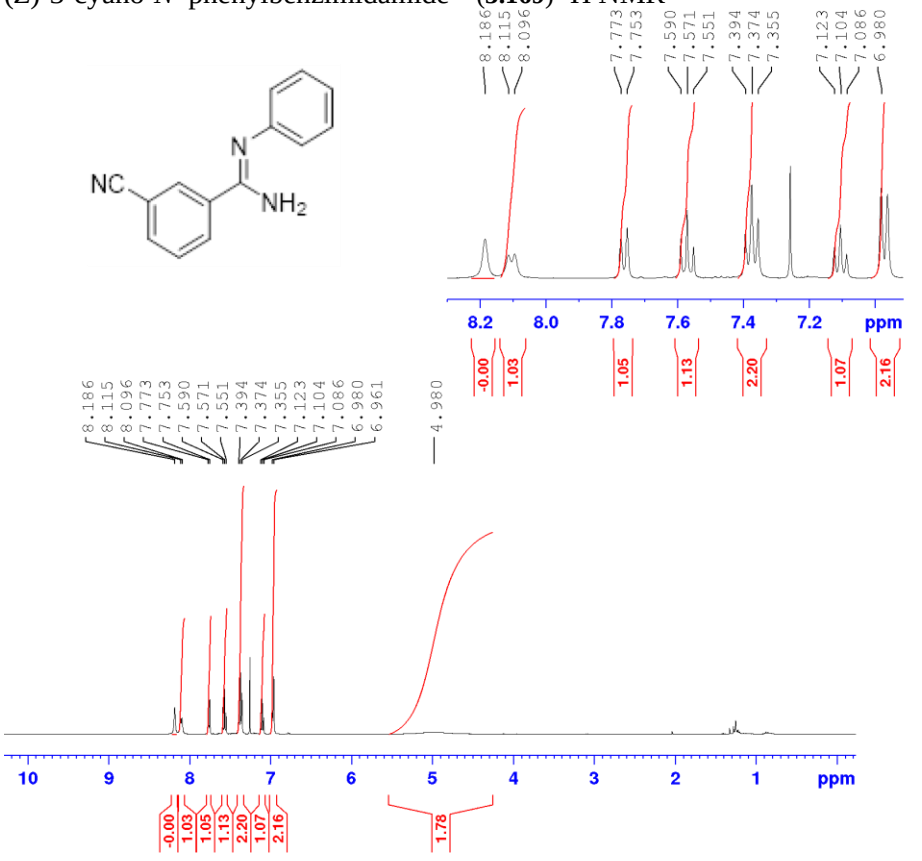

```
Current Data Parameters
NAME       JR0281-D316653
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190406
Time             14.10 h
INSTRUM            spect
PROBHD   Z122623_0021 (
PULPROG            zg30
TD                32564
SOLVENT           CDCl3
NS                    4
DS                    2
SWH           8223.685 Hz
FIDRES         0.505078 Hz
AQ             1.9798912 sec
RG                  101
DW              60.800 usec
DE              10.00 usec
TE              300.0 K
D1            2.00000000 sec
TD0                   1
SFO1         400.1324710 MHz
NUC1               1H
P0               4.00 usec
P1              12.00 usec
PLW1          7.19999981 W

F2 - Processing parameters
SI                32768
SF           400.1300109 MHz
WDW                  EM
SSB       0
LB              0.30 Hz
GB        0
PC              1.00
```

^{13}C NMR



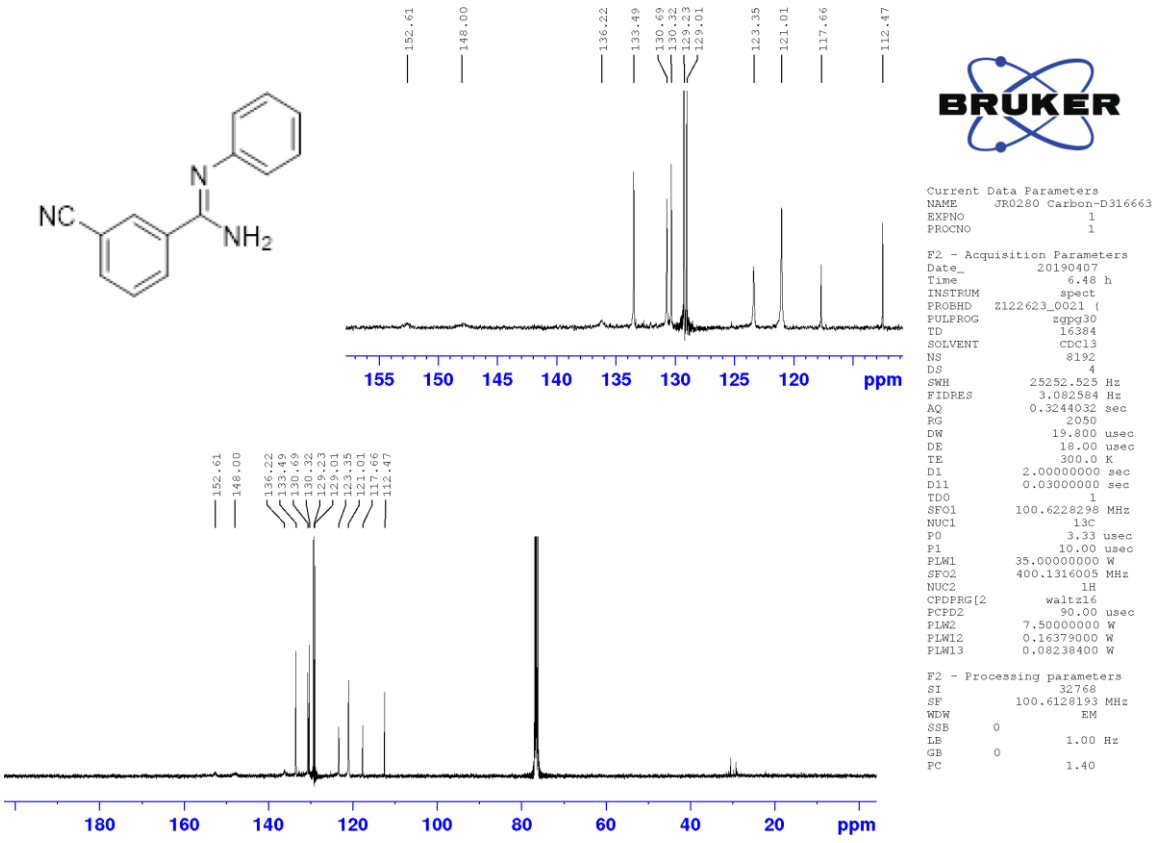

```
Current Data Parameters
NAME      JR0281-D316653
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190407
Time               1.17 h
INSTRUM           spect
PROBHD    Z122623_0021 (
PULPROG          zgpg30
TD                16384
SOLVENT           CDCl3
NS                 8192
DS                    4
SWH           25252.525 Hz
FIDRES         3.082584 Hz
AQ            0.3244032 sec
RG                 2050
DW               19.800 usec
DE               18.00 usec
TE                300.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                   1
SFO1        100.6228298 MHz
NUC1               13C
P0                3.33 usec
P1               10.00 usec
PLW1         35.00000000 W
SFO2        400.1316005 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2            90.00 usec
PLW2          7.50000000 W
PLW12         0.16379000 W
PLW13         0.08238400 W

F2 - Processing parameters
SI                32768
SF           100.6128193 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

(*Z*)-3-cyano-*N*'-phenylbenzimidamide – (**3.169**) ^{1}H NMR



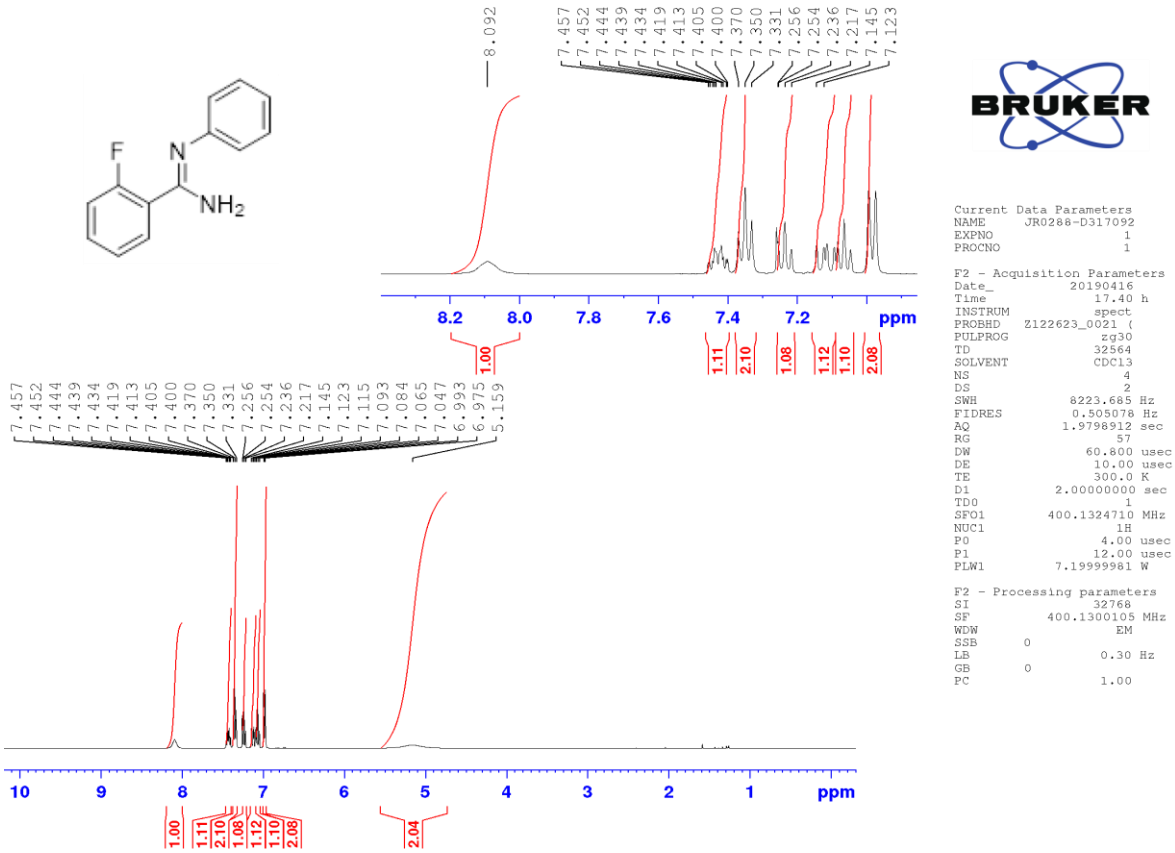

```
Current Data Parameters
NAME      JR0280-D316652
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190406
Time              14.02 h
INSTRUM           spect
PROBHD    Z122623_0021 (
PULPROG           zg30
TD                32564
SOLVENT           CDCl3
NS                    4
DS                    2
SWH            8223.685 Hz
FIDRES         0.505078 Hz
AQ            1.9798912 sec
RG                  128
DW               60.800 usec
DE               10.00 usec
TE                300.0 K
D1           2.00000000 sec
TD0                   1
SFO1        400.1324710 MHz
NUC1               1H
P0                4.00 usec
P1               12.00 usec
PLW1         7.19999981 W

F2 - Processing parameters
SI                32768
SF           400.1300116 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

^{13}C NMR



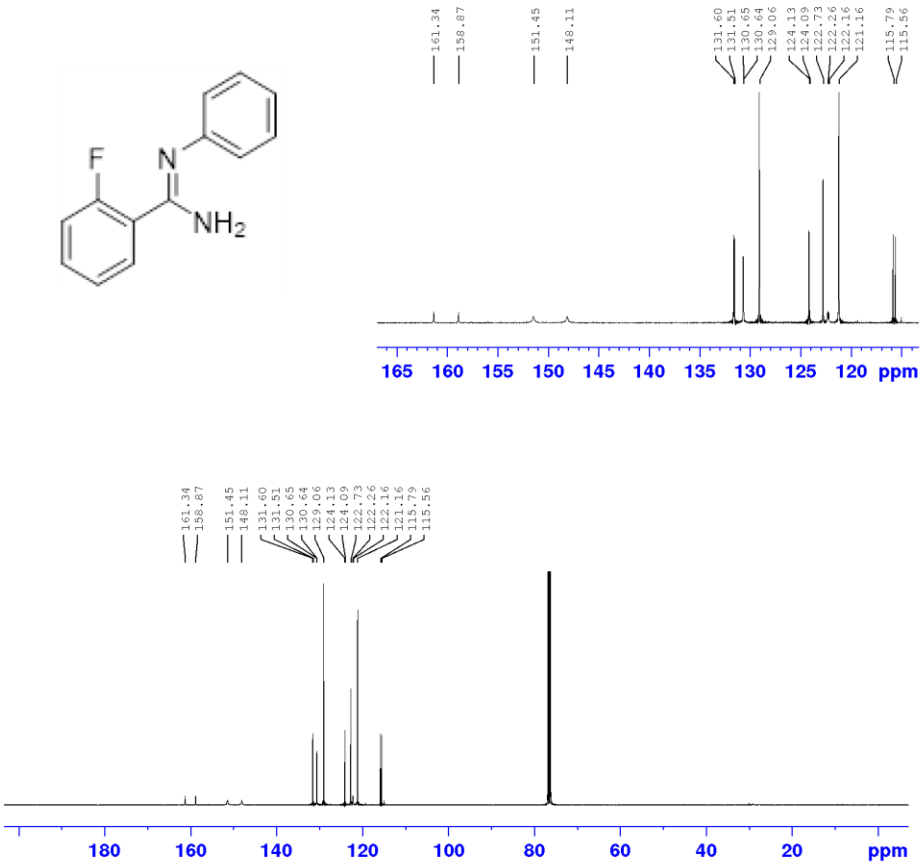

```
Current Data Parameters
NAME      JR0280 Carbon-D316663
EXPNO                        1
PROCNO                       1

F2 - Acquisition Parameters
Date_              20190407
Time                   6.48 h
INSTRUM                spect
PROBHD     Z122623_0021 (
PULPROG             zgpg30
TD                   16384
SOLVENT              CDCl3
NS                    8192
DS                       4
SWH              25252.525 Hz
FIDRES            3.082584 Hz
AQ                0.3244032 sec
RG                     2050
DW                19.800 usec
DE                18.00 usec
TE                 300.0 K
D1              2.00000000 sec
D11             0.03000000 sec
TD0                      1
SFO1           100.6228298 MHz
NUC1                    13C
P0                   3.33 usec
P1                  10.00 usec
PLW1           35.00000000 W
SFO2           400.1316005 MHz
NUC2                    1H
CPDPRG[2            waltz16
PCPD2               90.00 usec
PLW2            7.50000000 W
PLW12           0.16379000 W
PLW13           0.08238400 W

F2 - Processing parameters
SI                   32768
SF            100.6128193 MHz
WDW                     EM
SSB          0
LB                    1.00 Hz
GB           0
PC                    1.40
```

(Z)-2-fluoro-*N*'-phenylbenzimidamide (**3.166**)- ^{1}H NMR



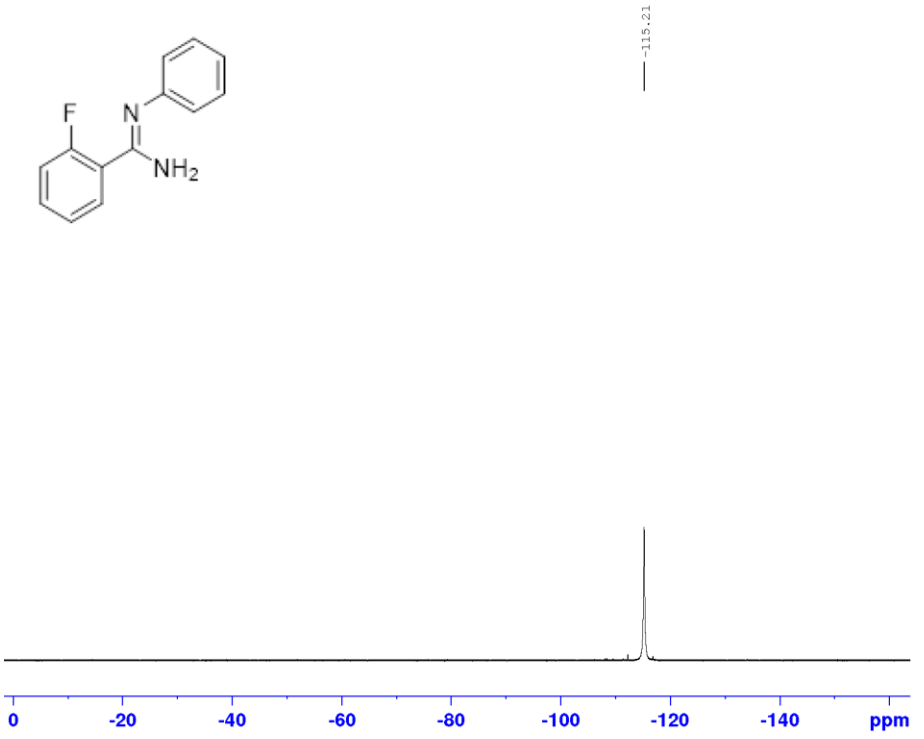

```
Current Data Parameters
NAME      JR0288-D317092
EXPNO                   1
PROCNO                  1

F2 - Acquisition Parameters
Date_            20190416
Time                17.40 h
INSTRUM             spect
PROBHD     Z122623_0021 (
PULPROG             zg30
TD                 32564
SOLVENT            CDCl3
NS                     4
DS                     2
SWH             8223.685 Hz
FIDRES           0.505078 Hz
AQ               1.9798912 sec
RG                    57
DW               60.800 usec
DE               10.00 usec
TE                300.0 K
D1             2.00000000 sec
TD0                    1
SFO1          400.1324710 MHz
NUC1                  1H
P0                4.00 usec
P1               12.00 usec
PLW1           7.19999981 W

F2 - Processing parameters
SI                 32768
SF           400.1300105 MHz
WDW                   EM
SSB          0
LB                0.30 Hz
GB           0
PC                1.00
```

^{13}C NMR



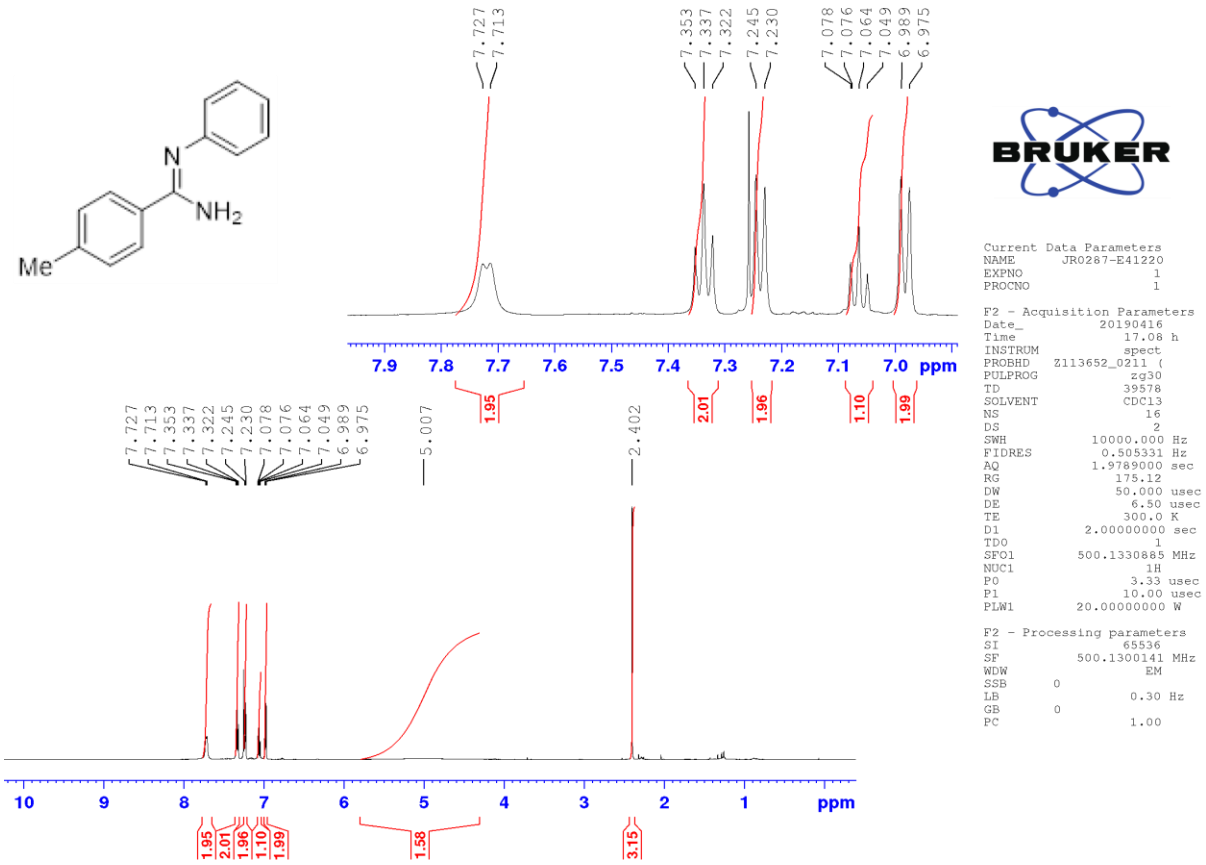

```
Current Data Parameters
NAME      JR0288-D317092
EXPNO                  3
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190417
Time               1.39 h
INSTRUM           spect
PROBHD   Z122623_0021 (
PULPROG          zgpg30
TD                16384
SOLVENT            CDCl3
NS                 8192
DS                    4
SWH           25252.525 Hz
FIDRES         3.082584 Hz
AQ            0.3244032 sec
RG                 2050
DW               19.800 usec
DE               18.00 usec
TE                300.0 K
D1            2.00000000 sec
D11           0.03000000 sec
TD0                   1
SFO1         100.6228298 MHz
NUC1                13C
P0                3.33 usec
P1               10.00 usec
PLW1         35.00000000 W
SFO2         400.1316005 MHz
NUC2                1H
CPDPRG[2         waltz16
PCPD2             90.00 usec
PLW2          7.50000000 W
PLW12         0.16379000 W
PLW13         0.08238400 W

F2 - Processing parameters
SI                32768
SF           100.6128193 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

^{19}F NMR




```
Current Data Parameters
NAME      JR0288-D317092
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190416
Time              17.44 h
INSTRUM           spect
PROBHD   Z122623_0021 (
PULPROG             zg30
TD                261896
SOLVENT            CDCl3
NS                   32
DS                    4
SWH           89285.711 Hz
FIDRES         0.681841 Hz
AQ            1.4666176 sec
RG                   16
DW               5.600 usec
DE               18.37 usec
TE                300.0 K
D1            2.00000000 sec
TD0                   1
SFO1         376.4607162 MHz
NUC1                19F
P0                5.00 usec
P1               15.00 usec
PLW1          5.19999981 W

F2 - Processing parameters
SI               262144
SF           376.4983660 MHz
WDW                  EM
SSB       0
LB                0.50 Hz
GB        0
PC                1.00
```

(*Z*)-4-methyl-*N*'-phenylbenzimidamide - (**3.168**) ^{1}H NMR

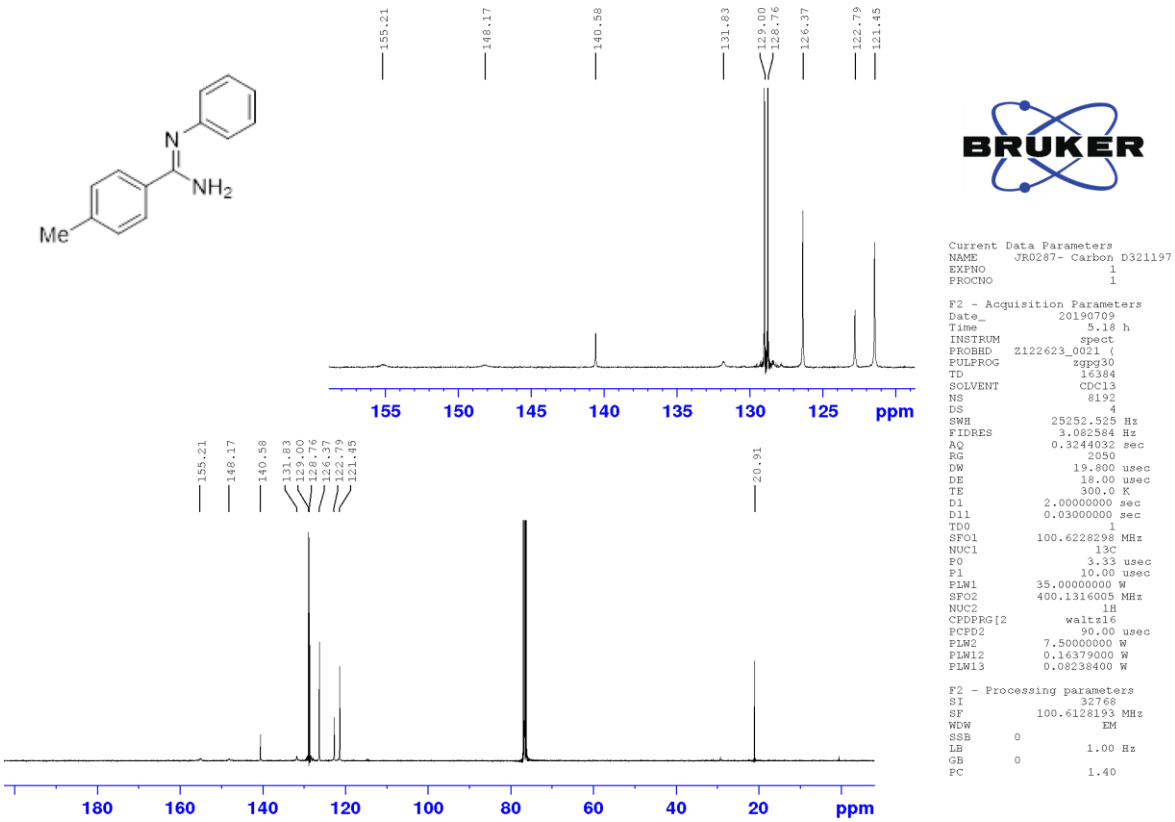


Current Data Parameters
NAME JR0287-E41220
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190416
Time 17.08 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zg30
TD 39578
SOLVENT CDCl3
NS 16
DS 2
SWH 10000.000 Hz
FIDRES 0.505331 Hz
AQ 1.9789000 sec
RG 175.12
DW 50.000 usec
DE 6.50 usec
TE 300.0 K
D1 2.00000000 sec
TD0 1
SFO1 500.1330885 MHz
NUC1 1H
P0 3.33 usec
P1 10.00 usec
PLW1 20.00000000 W

F2 - Processing parameters
SI 65536
SF 500.1300141 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

^{13}C NMR

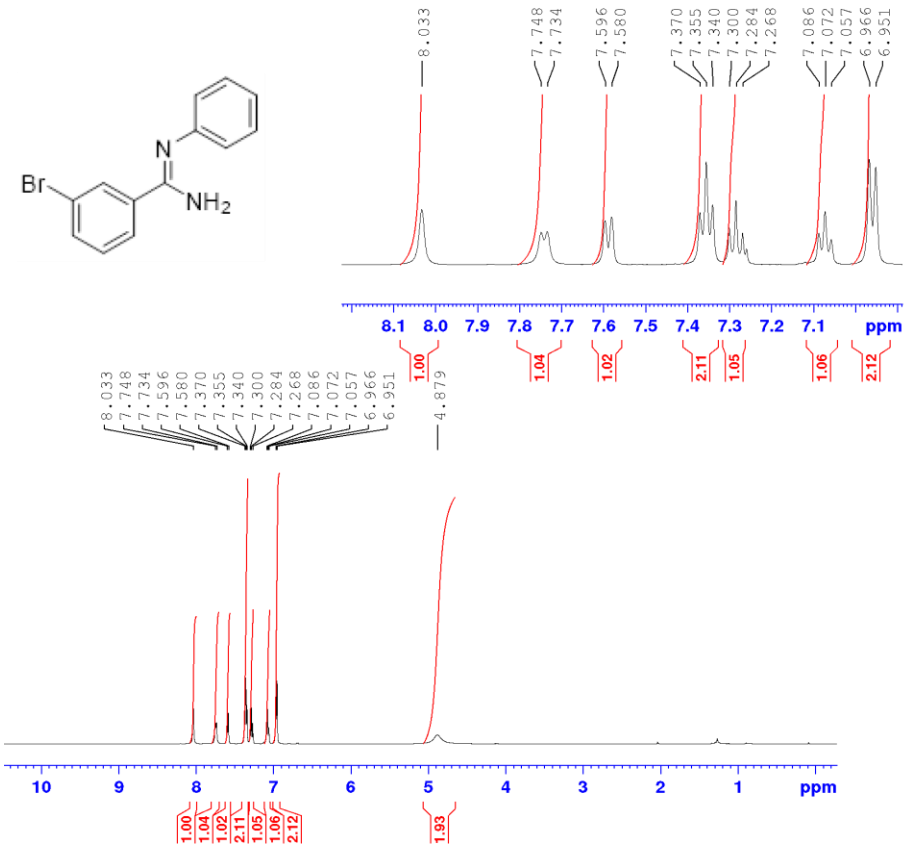


Current Data Parameters
NAME JR0287- Carbon D321197
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190709
Time 5.18 h
INSTRUM spect
PROBHD Z122623_0021 (
PULPROG zgpg30
TD 16384
SOLVENT CDCl3
NS 8192
DS 4
SWH 25252.525 Hz
FIDRES 3.082584 Hz
AQ 0.3244032 sec
RG 2050
DW 19.800 usec
DE 18.00 usec
TE 300.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
SFO1 100.6228298 MHz
NUC1 13C
P0 3.33 usec
P1 10.00 usec
PLW1 35.00000000 W
SFO2 400.1316005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 90.00 usec
PLW2 7.50000000 W
PLW12 0.16379000 W
PLW13 0.08238400 W

F2 - Processing parameters
SI 32768
SF 100.6128193 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

(Z)-3-bromo-*N*'-phenylbenzimidamide – (**3.167**) ^{1}H NMR

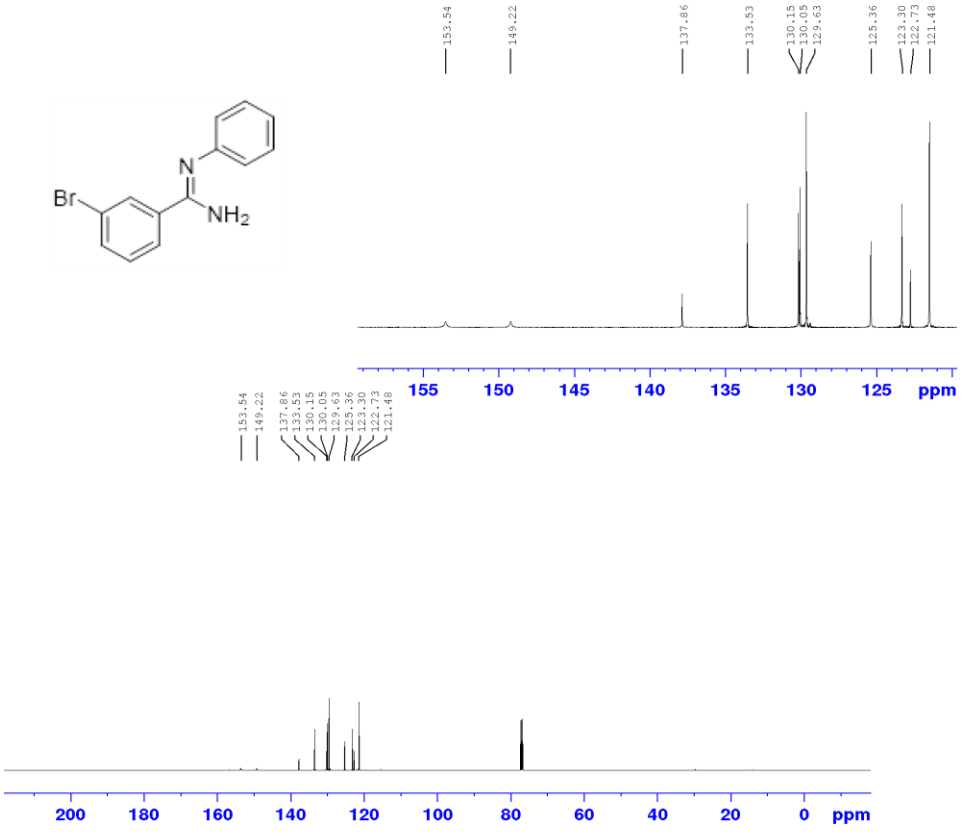


Current Data Parameters
NAME JR0329-E43459
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190816
Time 17.06 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zg30
TD 39578
SOLVENT CDCl3
NS 16
DS 2
SWH 10000.000 Hz
FIDRES 0.505331 Hz
AQ 1.9789000 sec
RG 68.83
DW 50.000 usec
DE 6.50 usec
TE 300.0 K
D1 2.00000000 sec
TD0 1
SFO1 500.1330885 MHz
NUC1 1H
P0 3.53 usec
P1 10.58 usec
PLW1 20.00000000 W

F2 - Processing parameters
SI 65536
SF 500.1300142 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

^{13}C NMR

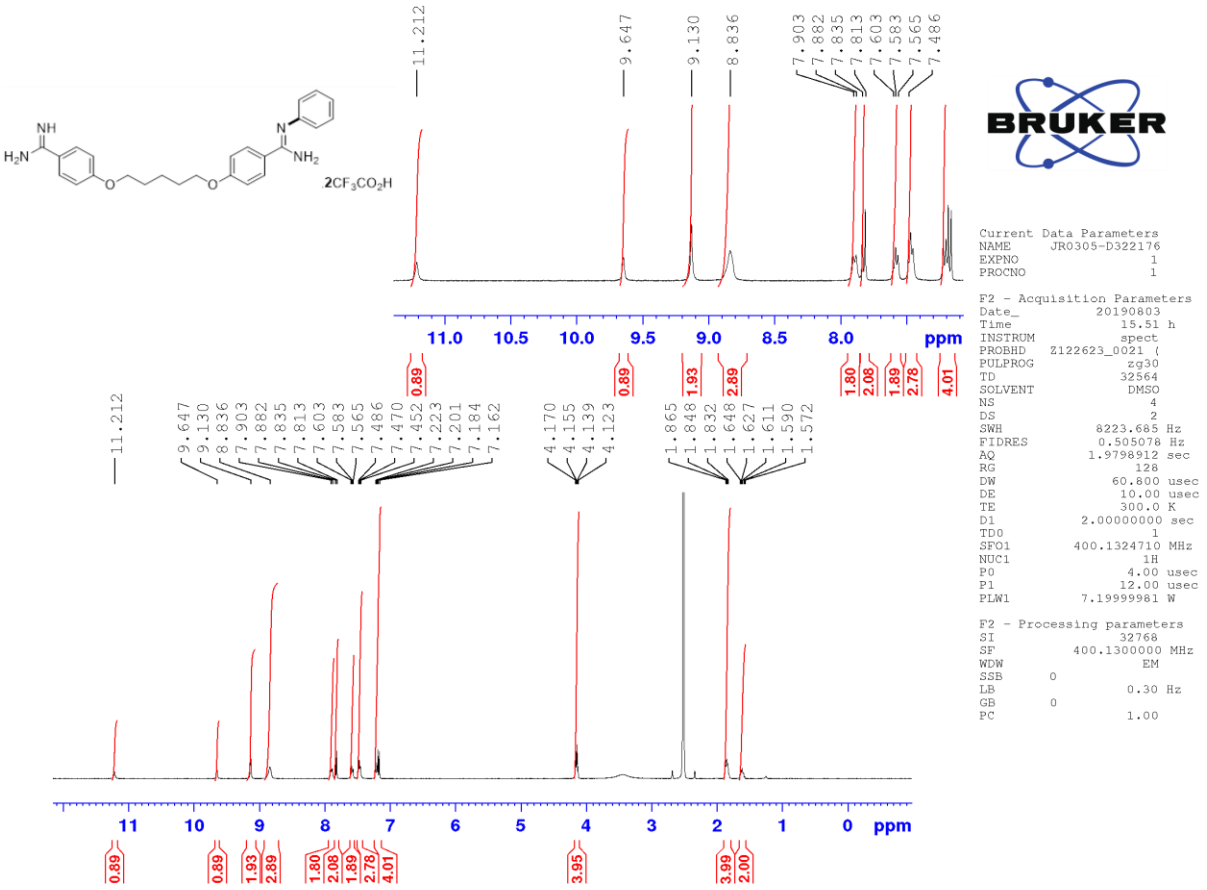


Current Data Parameters
NAME JR0329-E43459
EXPNO 4
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190817
Time 3.52 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zgpg30
TD 23806
SOLVENT CDCl3
NS 8192
DS 4
SWH 29761.904 Hz
FIDRES 2.500370 Hz
AQ 0.3999408 sec
RG 198.22
DW 16.800 usec
DE 6.50 usec
TE 299.9 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
SFO1 125.7703637 MHz
NUC1 13C
P0 3.33 usec
P1 10.00 usec
PLW1 76.00000000 W
SFO2 500.1320005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 21.00000000 W
PLW12 0.31513000 W
PLW13 0.15851000 W

F2 - Processing parameters
SI 32768
SF 125.7577885 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

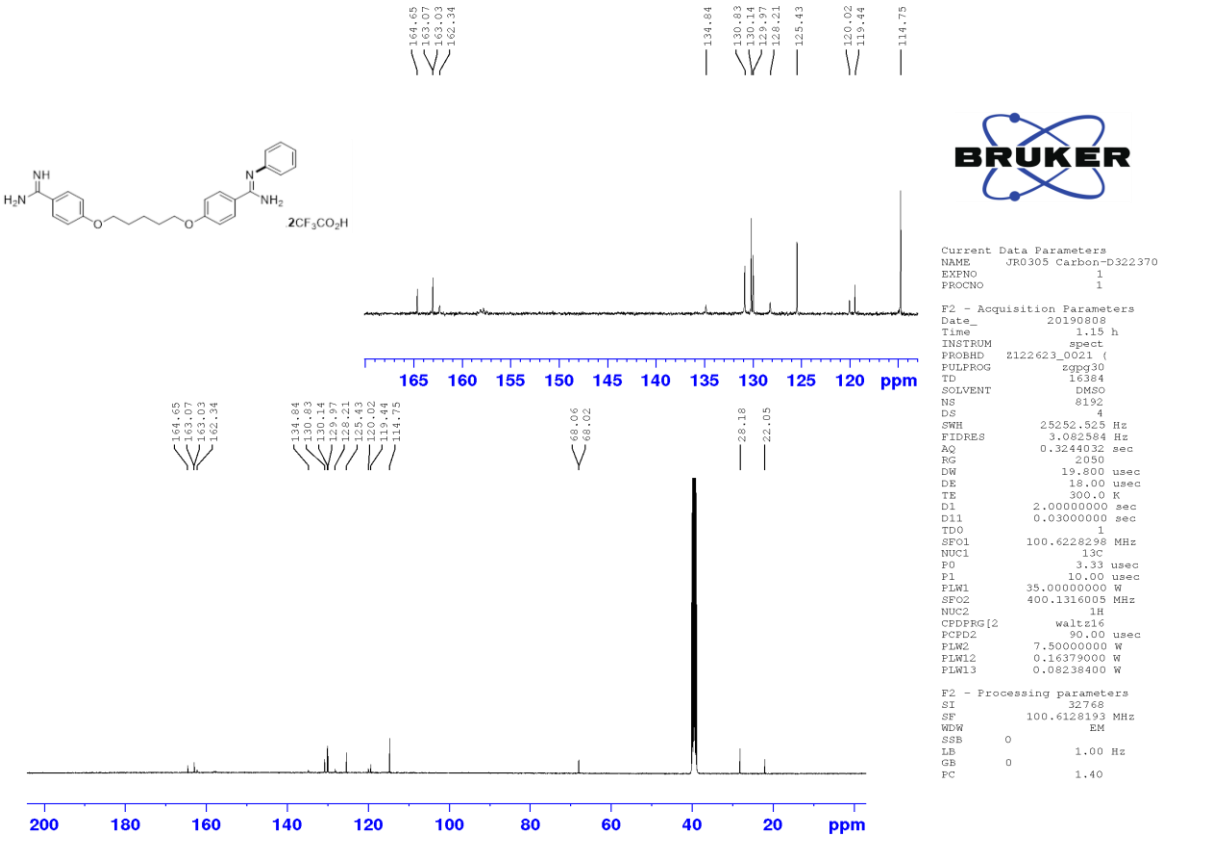
(*Z*)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-*N*'-phenylbenzimidamide (**3.174**) - ^{1}H NMR



BRUKER

```
Current Data Parameters
NAME      JR0305-D322176
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190803
Time               15.51 h
INSTRUM             spect
PROBHD   Z122623_0021 (
PULPROG             zg30
TD                 32564
SOLVENT             DMSO
NS                     4
DS                     2
SWH             8223.685 Hz
FIDRES             0.505078 Hz
AQ              1.9798912 sec
RG                   128
DW                60.800 usec
DE                10.00 usec
TE                300.0 K
D1             2.00000000 sec
TD0                    1
SFO1          400.1324710 MHz
NUC1                  1H
P0                 4.00 usec
P1                12.00 usec
PLW1           7.19999981 W

F2 - Processing parameters
SI                 32768
SF            400.1300000 MHz
WDW                   EM
SSB          0
LB                 0.30 Hz
GB           0
PC                 1.00
```

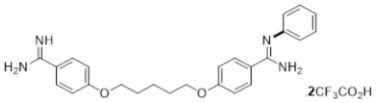
^{13}C NMR

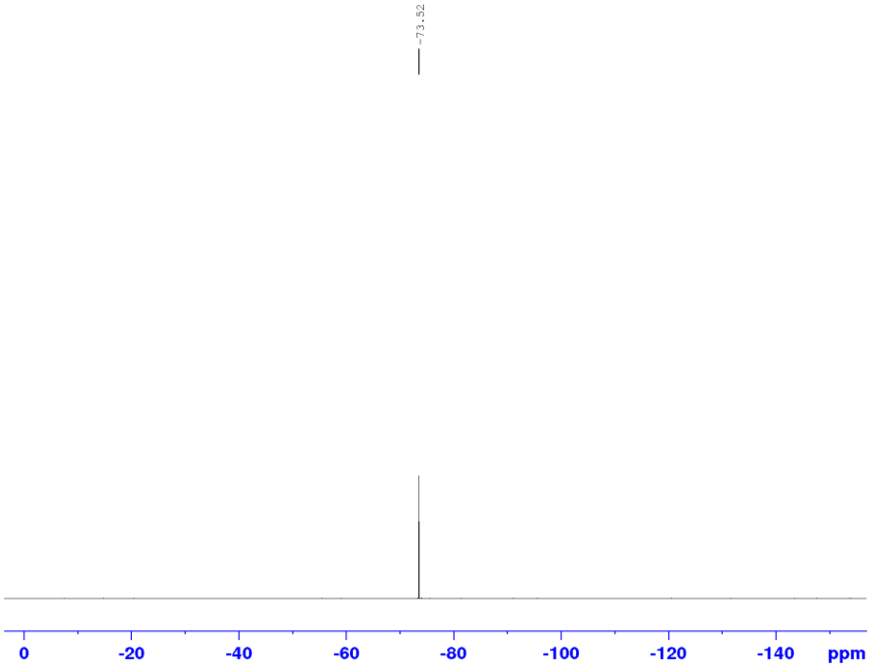


BRUKER

```
Current Data Parameters
NAME      JR0305 Carbon-D322370
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190808
Time               1.15 h
INSTRUM             spect
PROBHD   Z122623_0021 (
PULPROG           zgpg30
TD                 16384
SOLVENT             DMSO
NS                  8192
DS                     4
SWH            25252.525 Hz
FIDRES             3.082584 Hz
AQ              0.3244032 sec
RG                  2050
DW                19.800 usec
DE                18.00 usec
TE                300.0 K
D1             2.00000000 sec
D11            0.03000000 sec
TD0                    1
SFO1          100.6228298 MHz
NUC1                  13C
P0                 3.33 usec
P1                10.00 usec
PLW1          35.00000000 W
SFO2          400.1316005 MHz
NUC2                  1H
CPDPRG[2          waltz16
PCPD2             90.00 usec
PLW2           7.50000000 W
PLW12          0.16379000 W
PLW13          0.08238400 W

F2 - Processing parameters
SI                 32768
SF            100.6128193 MHz
WDW                   EM
SSB          0
LB                 1.00 Hz
GB           0
PC                 1.40
```

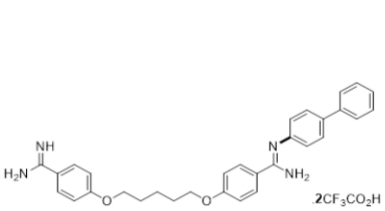
^{19}F NMR

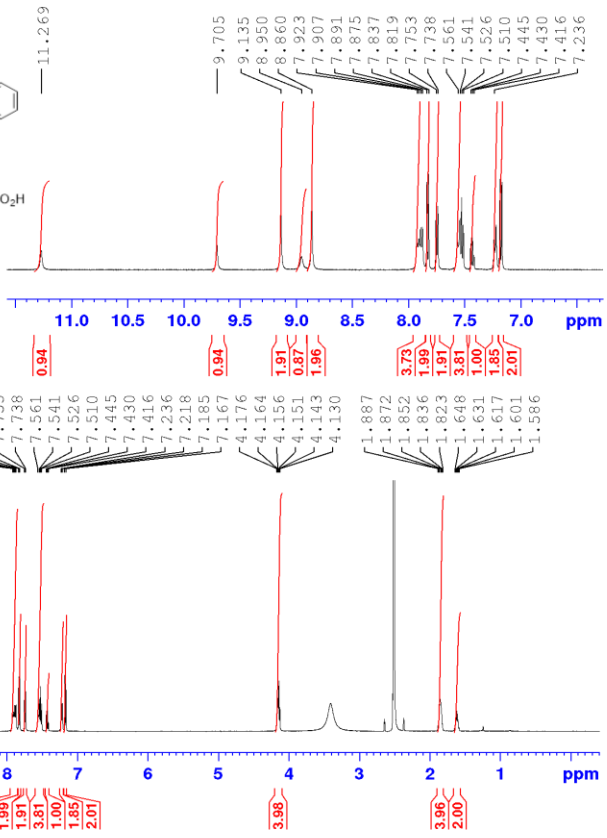





```
Current Data Parameters
NAME      JR0305-D322176
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190803
Time              15.54 h
INSTRUM             spect
PROBHD   Z122623_0021 (
PULPROG             zg30
TD                261896
SOLVENT             DMSO
NS                    32
DS                     4
SWH           89285.711 Hz
FIDRES         0.681841 Hz
AQ            1.4666176 sec
RG                    16
DW                5.600 usec
DE                18.37 usec
TE                300.0 K
D1            2.00000000 sec
TD0                    1
SFO1         376.4607162 MHz
NUC1                 19F
P0                5.00 usec
P1               15.00 usec
PLW1          5.19999981 W

F2 - Processing parameters
SI                  262144
SF           376.4983660 MHz
WDW                    EM
SSB       0
LB                0.50 Hz
GB        0
PC                1.00
```

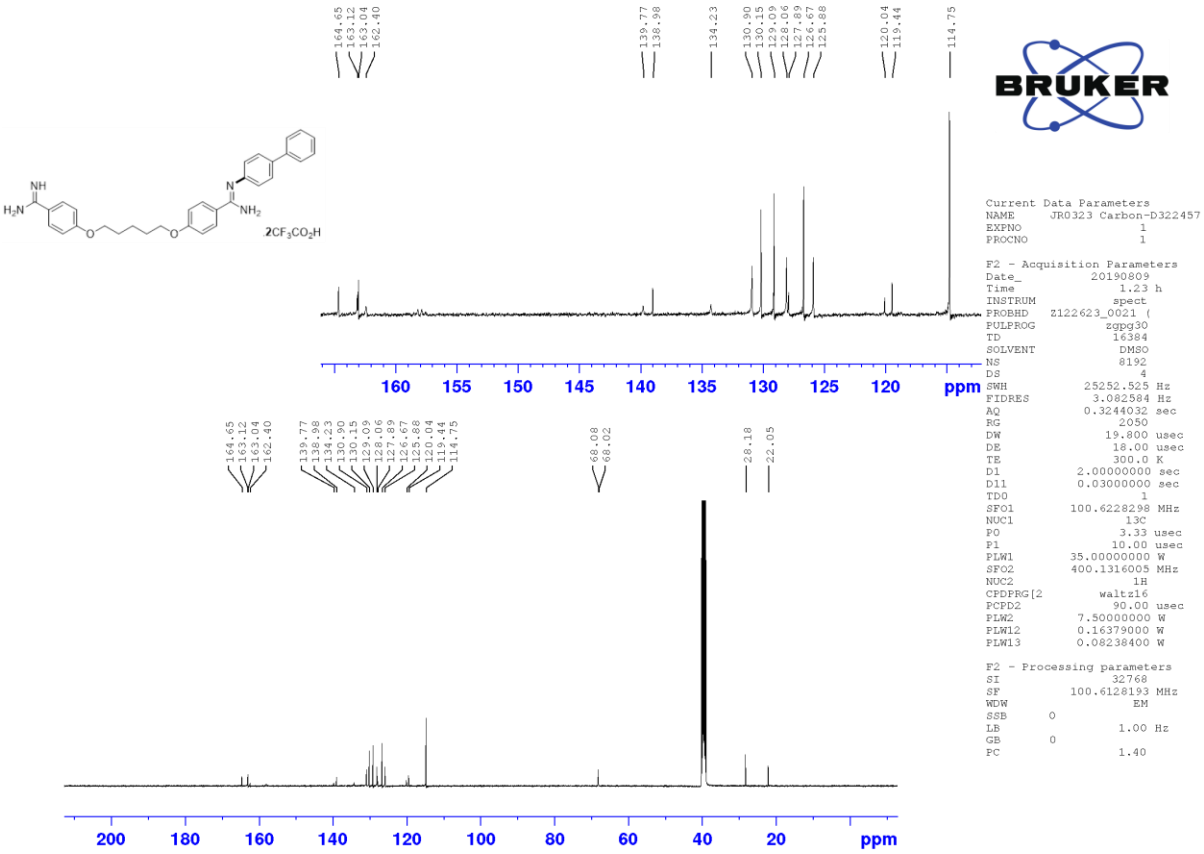
(Z)-*N*'-([1,1'-biphenyl]-4-yl)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)benzimidamide (**3.178**) - ^{1}H NMR





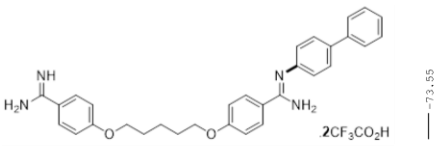

```
Current Data Parameters
NAME      JR0323-E42944
EXPNO                 1
PROCNO                1

F2 - Acquisition Parameters
Date_           20190719
Time              16.23 h
INSTRUM             spect
PROBHD   Z113652_0211 (
PULPROG             zg30
TD                 39578
SOLVENT             DMSO
NS                    16
DS                     2
SWH           10000.000 Hz
FIDRES         0.505331 Hz
AQ            1.9789000 sec
RG                175.12
DW                50.000 usec
DE                 6.50 usec
TE                300.0 K
D1            2.00000000 sec
TD0                    1
SFO1         500.1330885 MHz
NUC1                 1H
P0                3.53 usec
P1               10.58 usec
PLW1         20.00000000 W

F2 - Processing parameters
SI                  65536
SF           500.1300000 MHz
WDW                    EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

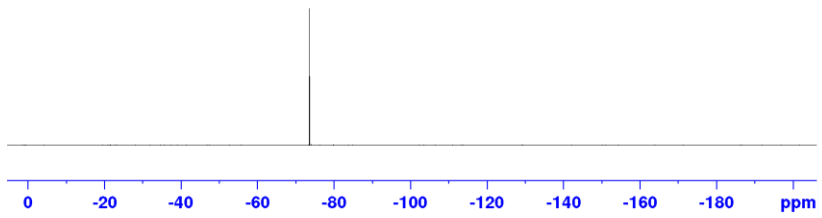
^{13}C NMR




```
Current Data Parameters
NAME      JR0323 Carbon-D322457
EXPNO                1
PROCNO               1

F2 - Acquisition Parameters
Date_           20190809
Time              1.23 h
INSTRUM           spect
PROBHD    Z122623_0021 (
PULPROG          zgpg30
TD               16384
SOLVENT           DMSO
NS                8192
DS                   4
SWH          25252.525 Hz
FIDRES        3.082584 Hz
AQ            0.3244032 sec
RG                2050
DW              19.800 usec
DE              18.00 usec
TE               300.0 K
D1           2.00000000 sec
D11          0.03000000 sec
TD0                  1
SFO1        100.6228298 MHz
NUC1               13C
P0               3.33 usec
P1              10.00 usec
PLW1        35.00000000 W
SFO2        400.1316005 MHz
NUC2               1H
CPDPRG[2        waltz16
PCPD2            90.00 usec
PLW2         7.50000000 W
PLW12        0.16379000 W
PLW13        0.08238400 W

F2 - Processing parameters
SI                32768
SF          100.6128193 MHz
WDW                  EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

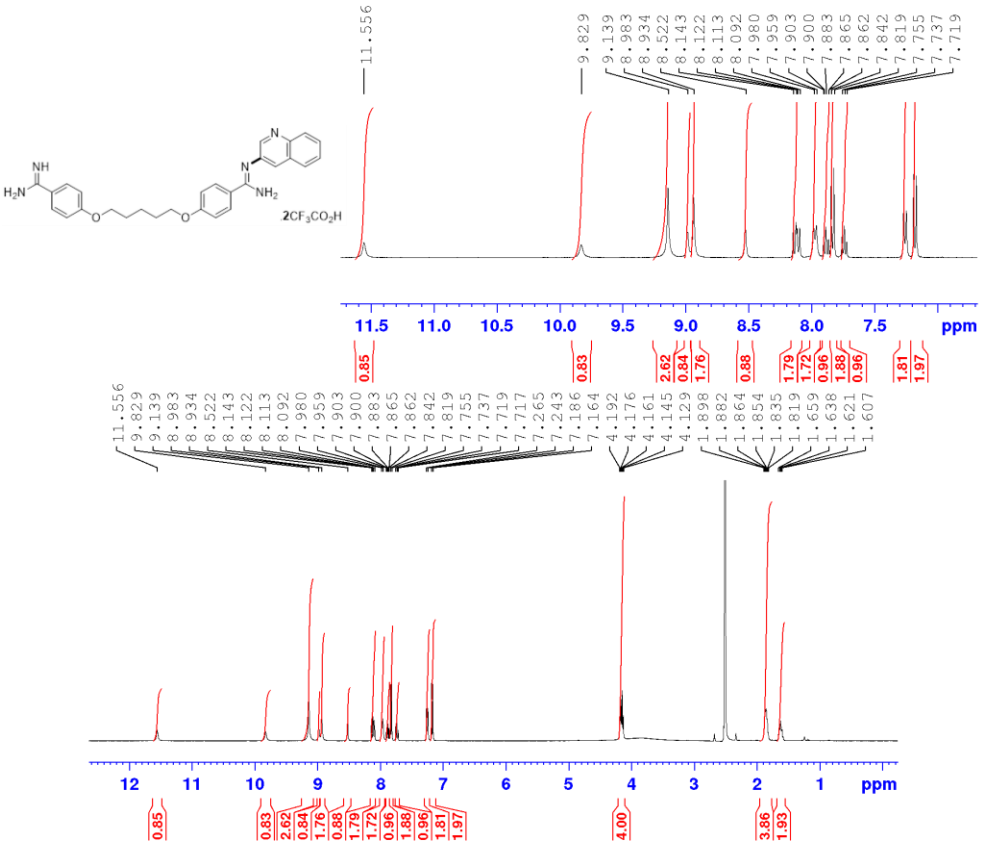
^{19}F NMR






```
Current Data Parameters
NAME      JR0323-E42944
EXPNO                2
PROCNO               1

F2 - Acquisition Parameters
Date_           20190719
Time             16.26 h
INSTRUM           spect
PROBHD    Z113652_0211 (
PULPROG          zgflqn
TD              130892
SOLVENT           DMSO
NS                  32
DS                   4
SWH         113636.367 Hz
FIDRES        1.736338 Hz
AQ            0.5759248 sec
RG              198.22
DW               4.400 usec
DE               6.50 usec
TE               300.0 K
D1           1.00000000 sec
TD0                  1
SFO1        470.5453180 MHz
NUC1               19F
P1              15.00 usec
PLW1        35.00000000 W

F2 - Processing parameters
SI                65536
SF          470.5923772 MHz
WDW                  EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

(*Z*)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-*N*'-(quinolin-3-yl)benzimidamide (**3.179**) - ^{1}H NMR

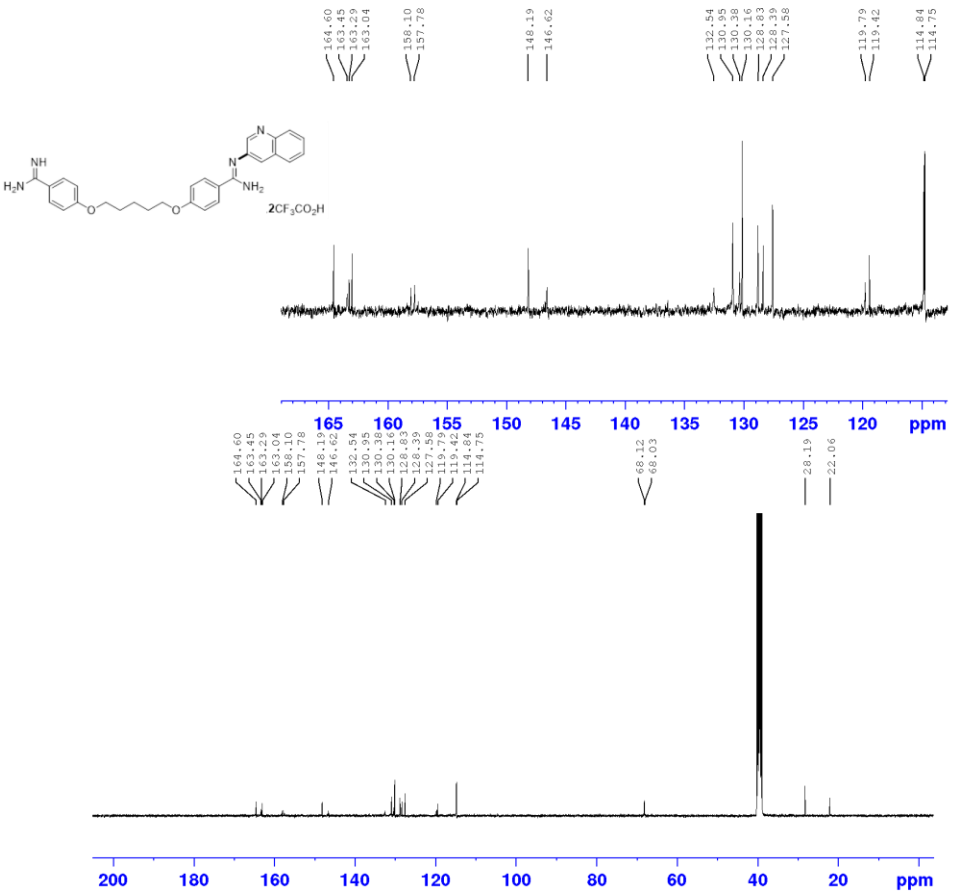


Current Data Parameters
NAME JR0324-D321850
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190723
Time 18.59 h
INSTRUM spect
PROBHD Z122623_0021 (
PULPROG zg30
TD 32564
SOLVENT DMSO
NS 4
DS 2
SWH 8223.685 Hz
FIDRES 0.505078 Hz
AQ 1.9798912 sec
RG 114
DW 60.800 usec
DE 10.00 usec
TE 300.0 K
D1 2.00000000 sec
TD0 1
SFO1 400.1324710 MHz
NUC1 1H
P0 4.00 usec
P1 12.00 usec
PLW1 7.19999981 W

F2 - Processing parameters
SI 32768
SF 400.1300000 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

^{13}C NMR

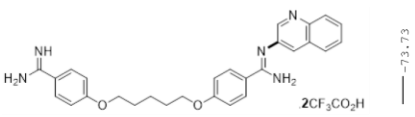


Current Data Parameters
NAME JR0324-D321850
EXPNO 4
PROCNO 1

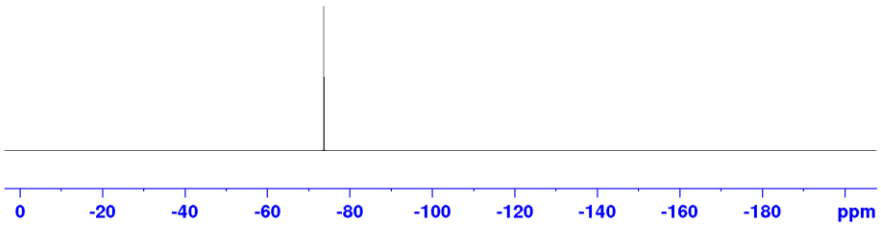
F2 - Acquisition Parameters
Date_ 20190724
Time 2.48 h
INSTRUM spect
PROBHD Z122623_0021 (
PULPROG zgpg30
TD 16384
SOLVENT DMSO
NS 4096
DS 4
SWH 25252.525 Hz
FIDRES 3.082584 Hz
AQ 0.3244032 sec
RG 2050
DW 19.800 usec
DE 18.00 usec
TE 300.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
SFO1 100.6228298 MHz
NUC1 13C
P0 3.33 usec
P1 10.00 usec
PLW1 35.00000000 W
SFO2 400.1316005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 90.00 usec
PLW2 7.50000000 W
PLW12 0.16379000 W
PLW13 0.08238400 W

F2 - Processing parameters
SI 32768
SF 100.6128193 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

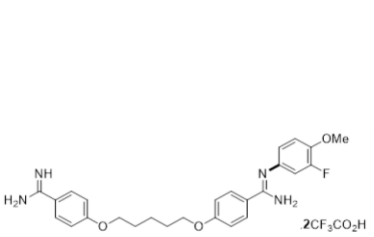
^{19}F NMR

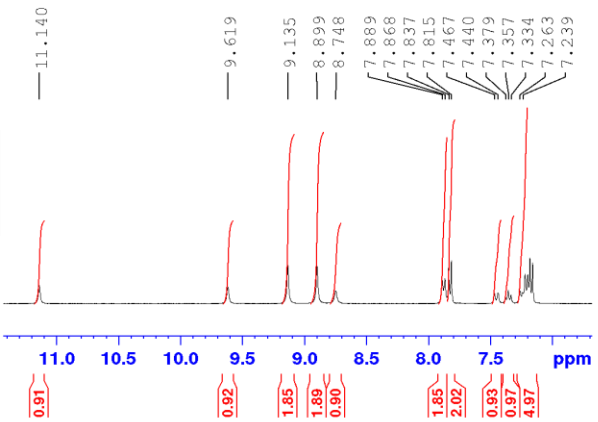


```
Current Data Parameters
NAME      JR0324-D321850
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190723
Time              19.04 h
INSTRUM             spect
PROBHD   Z122623_0021 (
PULPROG             zg30
TD                261896
SOLVENT              DMSO
NS                     32
DS                      4
SWH           89285.711 Hz
FIDRES         0.681841 Hz
AQ            1.4666176 sec
RG                     16
DW                5.600 usec
DE               18.37 usec
TE                300.0 K
D1           2.00000000 sec
TD0                    1
SFO1        376.4607162 MHz
NUC1                 19F
P0                5.00 usec
P1               15.00 usec
PLW1          5.19999981 W

F2 - Processing parameters
SI                262144
SF          376.4983660 MHz
WDW                   EM
SSB      0
LB                0.50 Hz
GB       0
PC                1.00
```

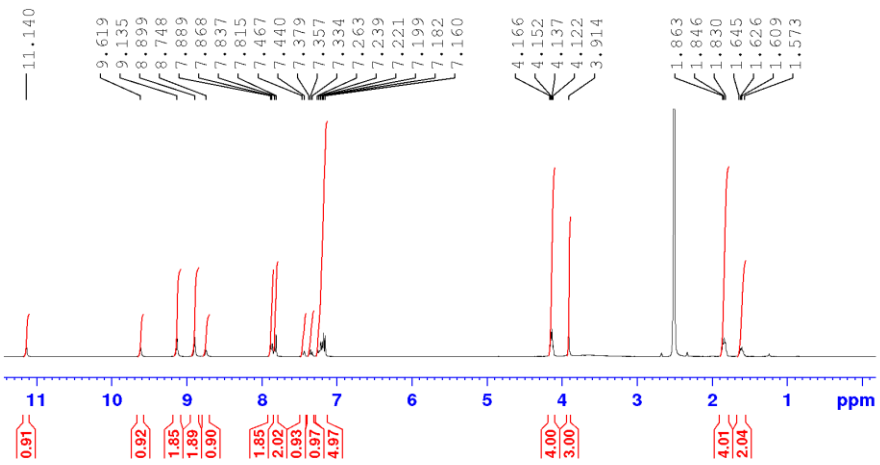


(Z)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-*N*'-(3-fluoro-4-methoxyphenyl)benzimidamide (**3.181**) - ^{1}H NMR

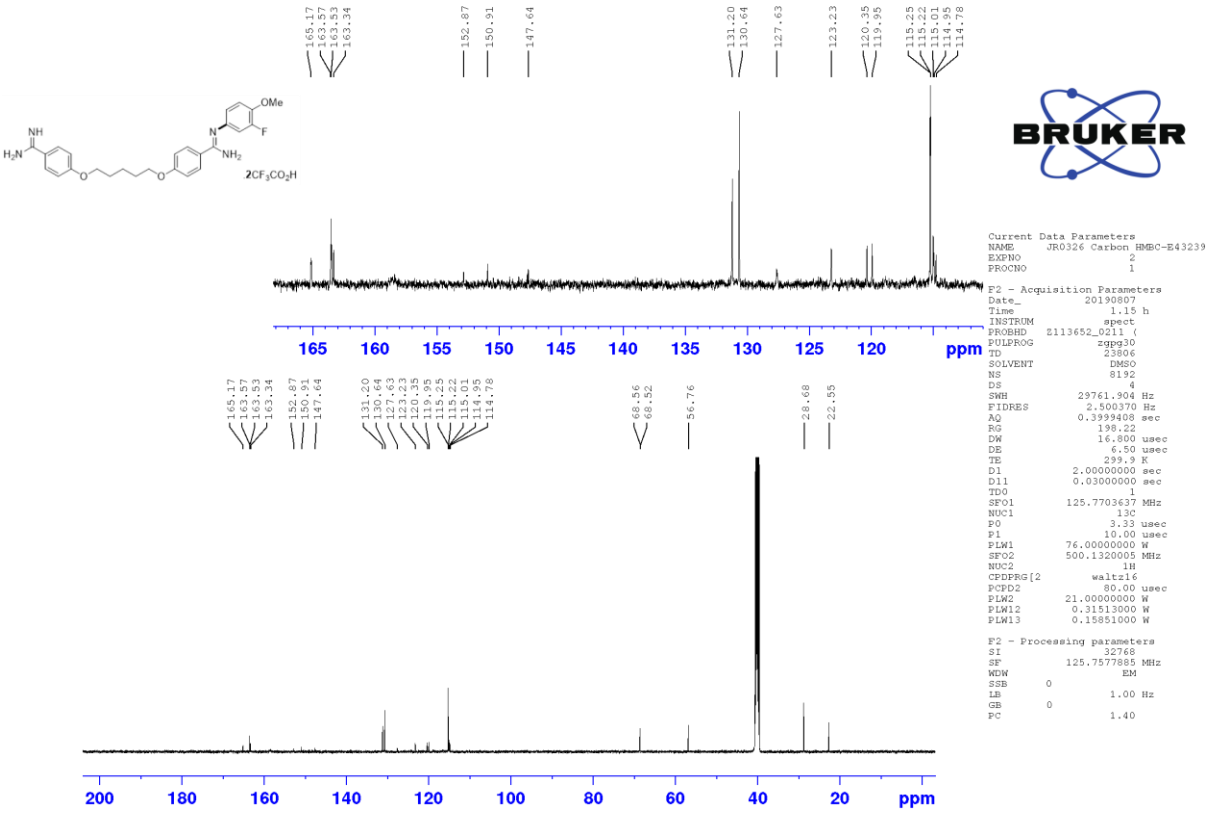




```
Current Data Parameters
NAME      JR0326-D322177
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190803
Time              15.58 h
INSTRUM             spect
PROBHD   Z122623_0021 (
PULPROG             zg30
TD                 32564
SOLVENT              DMSO
NS                     4
DS                     2
SWH           8223.685 Hz
FIDRES        0.505078 Hz
AQ            1.9798912 sec
RG                   114
DW               60.800 usec
DE               10.00 usec
TE                300.0 K
D1           2.00000000 sec
TD0                    1
SFO1        400.1324710 MHz
NUC1                 1H
P0                4.00 usec
P1               12.00 usec
PLW1          7.19999981 W

F2 - Processing parameters
SI                 32768
SF          400.1300000 MHz
WDW                   EM
SSB      0
LB                0.30 Hz
GB       0
PC                1.00
```



^{13}C NMR

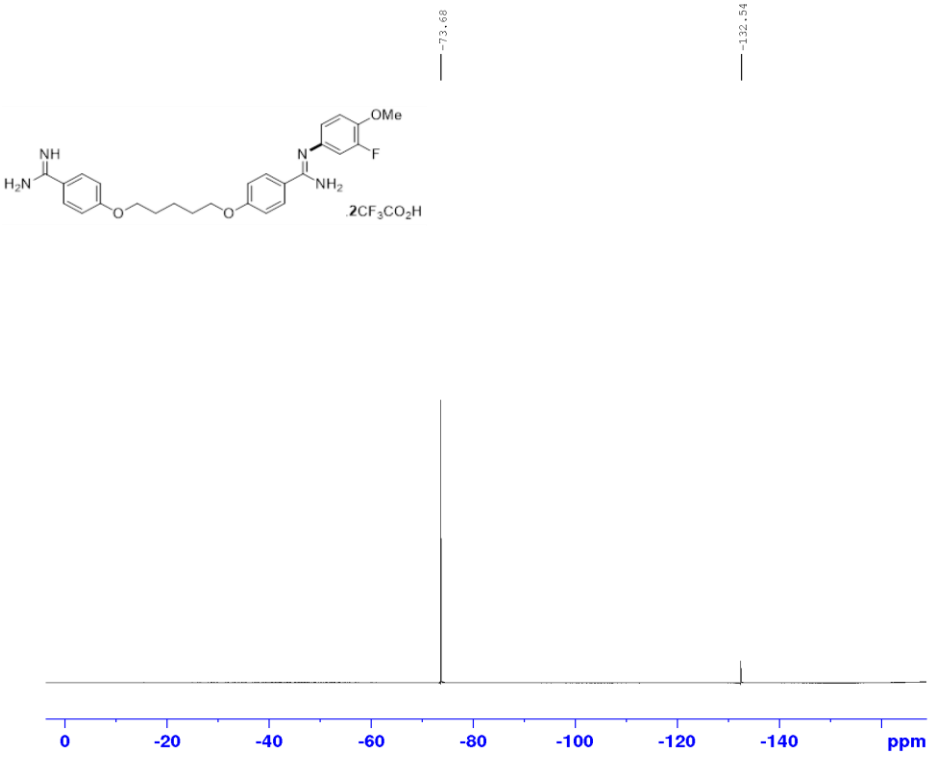


Current Data Parameters
NAME JR0326 Carbon HMBC-E43239
EXPNO 2
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190807
Time 1.15 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zgpg30
TD 23806
SOLVENT DMSO
NS 8192
DS 4
SWH 29761.904 Hz
FIDRES 2.500370 Hz
AQ 0.3999408 sec
RG 198.22
DW 16.800 usec
DE 6.50 usec
TE 289.9 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
SFO1 125.7703637 MHz
NUC1 13C
P0 3.33 usec
P1 10.00 usec
PLW1 76.00000000 W
SFO2 500.1320005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 21.00000000 W
PLW12 0.31513000 W
PLW13 0.15851000 W

F2 - Processing parameters
SI 32768
SF 125.7577885 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

^{19}F NMR

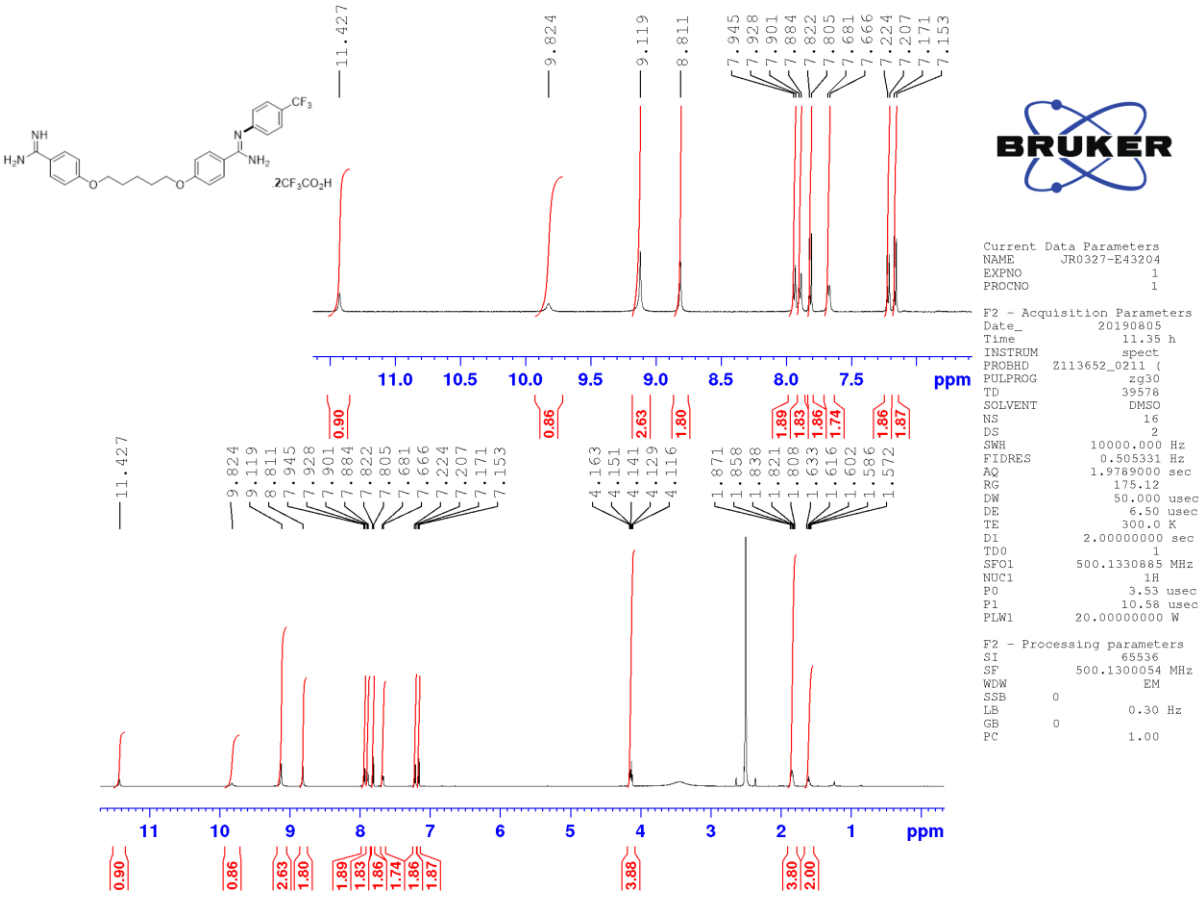


Current Data Parameters
NAME JR0326 Fluorine-E43445
EXPNO 1
PROCNO 1

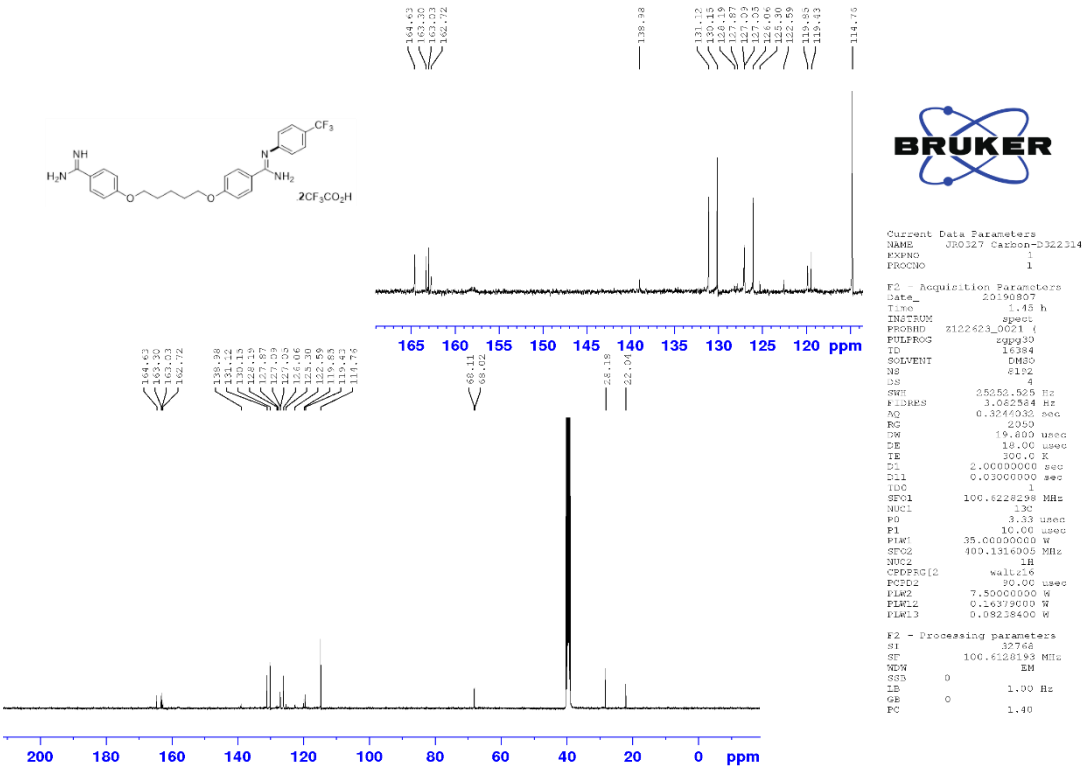
F2 - Acquisition Parameters
Date_ 20190816
Time 11.01 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zgfhigqn.2
TD 131072
SOLVENT DMSO
NS 16
DS 4
SWH 113636.367 Hz
FIDRES 1.733953 Hz
AQ 0.5767168 sec
RG 198.22
DW 4.400 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec
D11 0.03000000 sec
D12 0.00002000 sec
TD0 1
SFO1 470.5453180 MHz
NUC1 19F
P1 15.00 usec
PLW1 35.00000000 W
SFO2 500.1320005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 21.00000000 W
PLW12 0.31513000 W

F2 - Processing parameters
SI 131072
SF 470.5923772 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

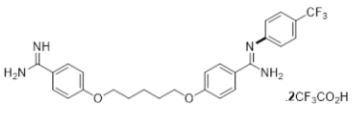
(Z)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-*N*'-(4-(trifluoromethyl)phenyl)benzimidamide (**3.180**) - ^{1}H NMR

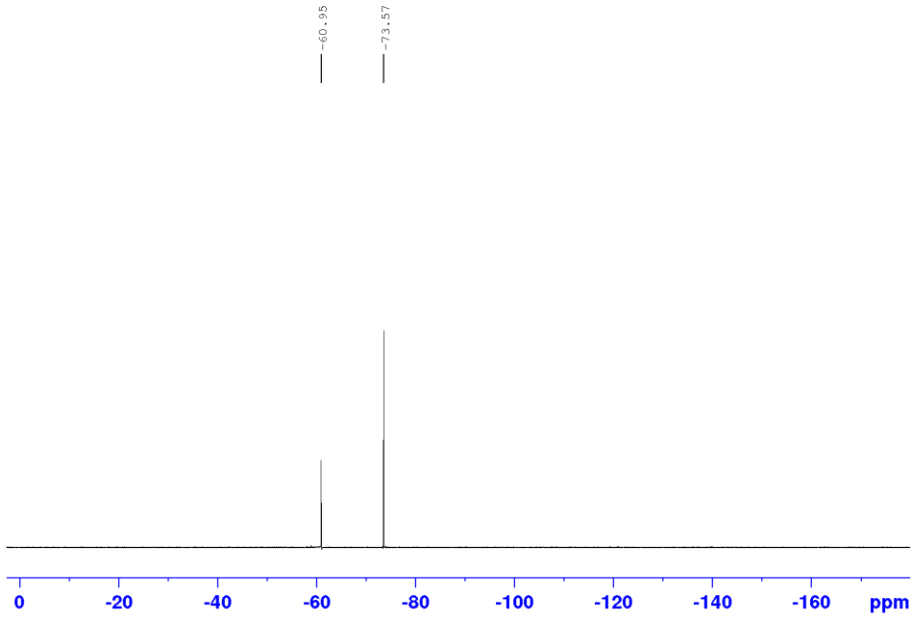


^{13}C NMR



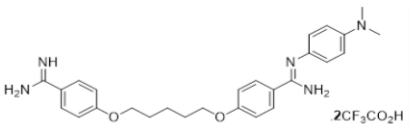
^{19}F NMR





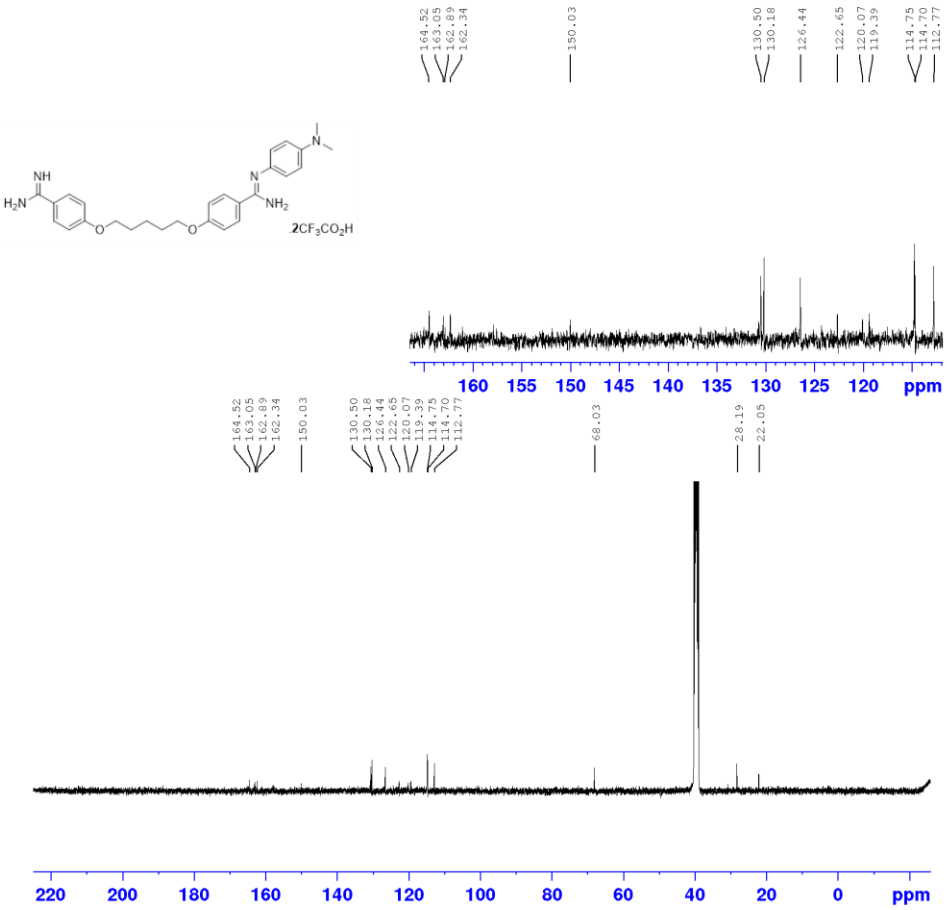

```
Current Data Parameters
NAME       JR0327-E43204
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20190805
Time              11.39 h
INSTRUM            spect
PROBHD    Z113652_0211 (
PULPROG          zgflqn
TD               130892
SOLVENT             DMSO
NS                    32
DS                     4
SWH           113636.367 Hz
FIDRES          1.736338 Hz
AQ              0.5759248 sec
RG                198.22
DW                4.400 usec
DE                6.50 usec
TE                300.0 K
D1             1.00000000 sec
TD0                    1
SFO1          470.5453180 MHz
NUC1                 19F
P1                15.00 usec
PLW1          35.00000000 W

F2 - Processing parameters
SI                 65536
SF           470.5923772 MHz
WDW                   EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

(Z)-4-((5-(4-carbamimidoylphenoxy)pentyl)oxy)-N'-(4-(dimethylamino)phenyl)benzimidamide (**3.176**) - ^{1}H NMR




```
Current Data Parameters
NAME       JR0390-D332945
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20200314
Time              17.21 h
INSTRUM            spect
PROBHD    Z122623_0021 (
PULPROG            zg30
TD                32564
SOLVENT             DMSO
NS                    32
DS                     2
SWH            8223.685 Hz
FIDRES         0.505078 Hz
AQ             1.9798912 sec
RG                   114
DW                60.800 usec
DE                10.00 usec
TE                300.0 K
D1             2.00000000 sec
TD0                    1
SFO1          400.1324710 MHz
NUC1                  1H
P0                 4.00 usec
P1                12.00 usec
PLW1           7.19999981 W

F2 - Processing parameters
SI                 32768
SF           400.1300000 MHz
WDW                   EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

^{13}C NMR

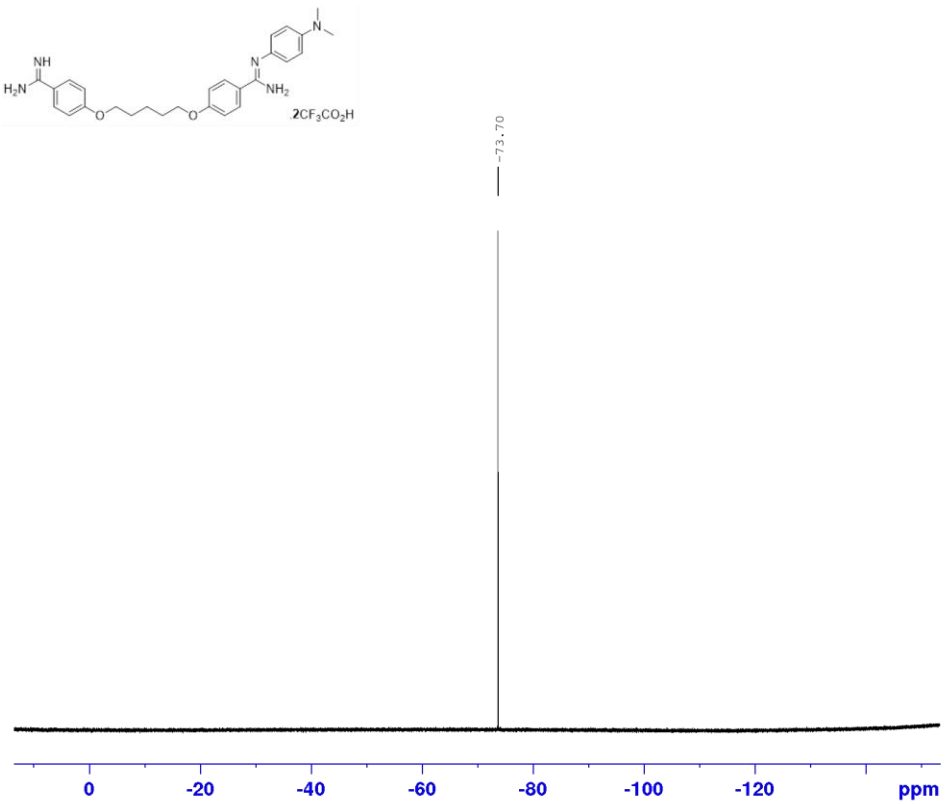


Current Data Parameters
NAME JR0390-D332945
EXPNO 3
PROCNO 1

F2 - Acquisition Parameters
Date_ 20200315
Time 2.12 h
INSTRUM spect
PROBHD Z122623_0021 (
PULPROG zgpg30
TD 16384
SOLVENT DMSO
NS 8192
DS 4
SWH 25252.525 Hz
FIDRES 3.082584 Hz
AQ 0.3244032 sec
RG 2050
DW 19.800 usec
DE 18.00 usec
TE 300.0 K
D1 2.00000000 sec
D11 0.03000000 sec
TD0 1
SFO1 100.6228298 MHz
NUC1 13C
P0 3.33 usec
P1 10.00 usec
PLW1 35.00000000 W
SFO2 400.1316005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 90.00 usec
PLW2 7.50000000 W
PLW12 0.16379000 W
PLW13 0.08238400 W

F2 - Processing parameters
SI 32768
SF 100.6128193 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

^{19}F NMR



Current Data Parameters
NAME JR0390 Fluorine - E47338
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20200807
Time 12.52 h
INSTRUM spect
PROBHD Z113652_0211 (
PULPROG zgfhigqn.2
TD 131072
SOLVENT DMSO
NS 16
DS 4
SWH 113636.367 Hz
FIDRES 1.733953 Hz
AQ 0.5767168 sec
RG 198.22
DW 4.400 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec
D11 0.03000000 sec
D12 0.00002000 sec
TD0 1
SFO1 470.5453180 MHz
NUC1 19F
P1 15.00 usec
PLW1 35.00000000 W
SFO2 500.1320005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 80.00 usec
PLW2 21.00000000 W
PLW12 0.31513000 W

F2 - Processing parameters
SI 131072
SF 470.5923772 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

X-ray crystallography tables

Complex **3.135**

Table 3.1 Crystal data and structure refinement for Complex **3.135**.

Empirical formula	C15 H16 Cu N4 O3	
Formula weight	363.86	
Temperature	123(2) K	
Wavelength	1.54184 Å	
Crystal system	Orthorhombic	
Space group	P 2_1 2_1 2_1	
Unit cell dimensions	a = 9.9702(5) Å	α= 90°.
	b = 11.3577(4) Å	β= 90°.
	c = 13.1156(5) Å	γ = 90°.
Volume	1485.19(11) $Å^3$	
Z	4	
Density (calculated)	1.627 Mg/m^3	
Absorption coefficient	2.271 mm^{-1}	
F(000)	748	
Crystal size	0.40 x 0.08 x 0.03 mm^3	
Theta range for data collection	5.151 to 73.200°.	
Index ranges	-12<=h<=11, -7<=k<=13, -15<=l<=16	
Reflections collected	5091	
Independent reflections	2919 [R(int) = 0.0301]	
Completeness to theta = 70.000°	100.0 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	1.00000 and 0.22234	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	2919 / 4 / 230	
Goodness-of-fit on F^2	1.030	
Final R indices [I>2sigma(I)]	R1 = 0.0337, wR2 = 0.0886	
R indices (all data)	R1 = 0.0360, wR2 = 0.0908	
Absolute structure parameter	-0.05(2)	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.471 and -0.286 $e.Å^{-3}$	

Table 3.2. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$x 10^3) for bur_cu_purple. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
Cu(1)	1414(1)	85(1)	3900(1)	15(1)
O(1)	1043(3)	1082(2)	2701(2)	16(1)
O(2)	1353(3)	-815(2)	2613(2)	17(1)
O(3)	903(3)	187(2)	1176(2)	19(1)
N(1)	1264(4)	1237(2)	4984(2)	18(1)
N(2)	163(4)	2737(3)	4191(3)	21(1)
N(3)	1928(3)	-1217(3)	4780(2)	16(1)
N(4)	618(4)	-2638(3)	4024(3)	24(1)
C(1)	659(4)	2263(3)	5031(3)	15(1)
C(2)	481(3)	2899(3)	6019(3)	15(1)
C(3)	-276(4)	3937(3)	6063(3)	20(1)
C(4)	-420(4)	4547(4)	6973(3)	23(1)
C(5)	177(4)	4121(4)	7853(3)	24(1)
C(6)	928(4)	3100(3)	7816(3)	24(1)
C(7)	1086(4)	2501(3)	6902(3)	20(1)
C(8)	1517(4)	-2306(3)	4715(3)	15(1)
C(9)	2058(4)	-3256(3)	5395(3)	18(1)
C(10)	2296(4)	-3030(4)	6422(3)	21(1)
C(11)	2817(4)	-3913(4)	7029(4)	29(1)
C(12)	3132(4)	-4998(4)	6630(4)	31(1)
C(13)	2909(4)	-5217(4)	5598(4)	29(1)
C(14)	2360(4)	-4345(3)	4991(3)	22(1)
C(15)	1090(3)	154(3)	2117(3)	16(1)

Table 3.3. Bond lengths [Å] and angles [°] for Complex **3.135**.

Cu(1)-N(1)	1.938(3)
Cu(1)-N(3)	1.945(3)
Cu(1)-O(1)	1.973(3)
Cu(1)-O(2)	1.974(2)
Cu(1)-C(15)	2.362(4)
O(1)-C(15)	1.304(4)
O(2)-C(15)	1.305(4)
O(3)-C(15)	1.249(4)
N(1)-C(1)	1.314(5)
N(1)-H(1N)	0.88(5)
N(2)-C(1)	1.322(5)
N(2)-H(2N)	0.874(14)
N(2)-H(3N)	0.884(14)
N(3)-C(8)	1.306(5)
N(3)-H(4N)	0.80(5)
N(4)-C(8)	1.329(5)
N(4)-H(5N)	0.869(14)
N(4)-H(6N)	0.874(14)
C(1)-C(2)	1.494(5)
C(2)-C(7)	1.382(5)
C(2)-C(3)	1.402(5)
C(3)-C(4)	1.388(5)
C(3)-H(3)	0.9500
C(4)-C(5)	1.386(6)
C(4)-H(4)	0.9500
C(5)-C(6)	1.382(6)
C(5)-H(5)	0.9500
C(6)-C(7)	1.387(5)
C(6)-H(6)	0.9500
C(7)-H(7)	0.9500
C(8)-C(9)	1.500(5)
C(9)-C(14)	1.379(5)
C(9)-C(10)	1.391(6)
C(10)-C(11)	1.382(6)
C(10)-H(10)	0.9500
C(11)-C(12)	1.376(7)

C(11)-H(11)	0.9500
C(12)-C(13)	1.393(7)
C(12)-H(12)	0.9500
C(13)-C(14)	1.384(6)
C(13)-H(13)	0.9500
C(14)-H(14)	0.9500

N(1)-Cu(1)-N(3)	95.64(13)
N(1)-Cu(1)-O(1)	100.52(11)
N(3)-Cu(1)-O(1)	163.55(12)
N(1)-Cu(1)-O(2)	167.01(12)
N(3)-Cu(1)-O(2)	97.01(12)
O(1)-Cu(1)-O(2)	67.04(10)
N(1)-Cu(1)-C(15)	133.90(13)
N(3)-Cu(1)-C(15)	130.46(13)
O(1)-Cu(1)-C(15)	33.51(11)
O(2)-Cu(1)-C(15)	33.53(11)
C(15)-O(1)-Cu(1)	89.8(2)
C(15)-O(2)-Cu(1)	89.8(2)
C(1)-N(1)-Cu(1)	132.0(3)
C(1)-N(1)-H(1N)	116(3)
Cu(1)-N(1)-H(1N)	112(3)
C(1)-N(2)-H(2N)	119(4)
C(1)-N(2)-H(3N)	124(4)
H(2N)-N(2)-H(3N)	117(5)
C(8)-N(3)-Cu(1)	126.8(3)
C(8)-N(3)-H(4N)	114(3)
Cu(1)-N(3)-H(4N)	118(3)
C(8)-N(4)-H(5N)	124(3)
C(8)-N(4)-H(6N)	125(4)
H(5N)-N(4)-H(6N)	109(4)
N(1)-C(1)-N(2)	119.6(4)
N(1)-C(1)-C(2)	121.6(3)
N(2)-C(1)-C(2)	118.8(3)
C(7)-C(2)-C(3)	118.4(4)
C(7)-C(2)-C(1)	121.1(3)
C(3)-C(2)-C(1)	120.4(4)
C(4)-C(3)-C(2)	120.8(4)

C(4)-C(3)-H(3)	119.6
C(2)-C(3)-H(3)	119.6
C(5)-C(4)-C(3)	119.9(4)
C(5)-C(4)-H(4)	120.1
C(3)-C(4)-H(4)	120.1
C(6)-C(5)-C(4)	119.7(4)
C(6)-C(5)-H(5)	120.1
C(4)-C(5)-H(5)	120.1
C(5)-C(6)-C(7)	120.3(4)
C(5)-C(6)-H(6)	119.9
C(7)-C(6)-H(6)	119.9
C(2)-C(7)-C(6)	121.0(4)
C(2)-C(7)-H(7)	119.5
C(6)-C(7)-H(7)	119.5
N(3)-C(8)-N(4)	121.6(3)
N(3)-C(8)-C(9)	122.0(3)
N(4)-C(8)-C(9)	116.4(3)
C(14)-C(9)-C(10)	120.1(4)
C(14)-C(9)-C(8)	119.7(4)
C(10)-C(9)-C(8)	120.3(4)
C(11)-C(10)-C(9)	119.2(4)
C(11)-C(10)-H(10)	120.4
C(9)-C(10)-H(10)	120.4
C(12)-C(11)-C(10)	121.1(4)
C(12)-C(11)-H(11)	119.5
C(10)-C(11)-H(11)	119.5
C(11)-C(12)-C(13)	119.5(4)
C(11)-C(12)-H(12)	120.2
C(13)-C(12)-H(12)	120.2
C(14)-C(13)-C(12)	119.7(4)
C(14)-C(13)-H(13)	120.2
C(12)-C(13)-H(13)	120.2
C(9)-C(14)-C(13)	120.4(4)
C(9)-C(14)-H(14)	119.8
C(13)-C(14)-H(14)	119.8
O(3)-C(15)-O(1)	123.4(3)
O(3)-C(15)-O(2)	123.2(3)
O(1)-C(15)-O(2)	113.3(3)

O(3)-C(15)-Cu(1)	179.3(3)
O(1)-C(15)-Cu(1)	56.65(17)
O(2)-C(15)-Cu(1)	56.69(17)

Symmetry transformations used to generate equivalent atoms:

Table 3.4. Anisotropic displacement parameters ($Å^2 x\ 10^3$)for Complex **3.135**. The anisotropic displacement factor exponent takes the form: $-2\pi^2[\ h^2a^{*2}U^{11} + ... + 2\ h\ k\ a^*\ b^*\ U^{12}\ ]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
Cu(1)	20(1)	10(1)	14(1)	0(1)	-2(1)	1(1)
O(1)	24(1)	11(1)	15(1)	1(1)	-1(1)	1(1)
O(2)	25(1)	10(1)	16(1)	-2(1)	-1(1)	2(1)
O(3)	27(1)	15(1)	15(1)	0(1)	1(1)	-3(1)
N(1)	27(2)	12(1)	13(1)	-2(1)	-1(2)	3(1)
N(2)	34(2)	14(2)	15(2)	-1(1)	0(1)	7(1)
N(3)	22(2)	12(1)	16(2)	3(1)	-6(1)	1(1)
N(4)	30(2)	16(2)	26(2)	7(1)	-10(2)	-6(1)
C(1)	18(2)	10(2)	17(2)	1(1)	-1(2)	-4(1)
C(2)	17(2)	12(2)	18(2)	0(1)	2(2)	-4(1)
C(3)	19(2)	18(2)	23(2)	1(2)	-3(2)	5(1)
C(4)	19(2)	24(2)	25(2)	-8(2)	3(2)	5(2)
C(5)	25(2)	26(2)	21(2)	-7(2)	3(2)	-5(2)
C(6)	33(2)	24(2)	16(2)	3(2)	-3(2)	-4(2)
C(7)	28(2)	12(2)	21(2)	0(1)	1(2)	0(1)
C(8)	17(2)	12(2)	16(2)	0(1)	0(2)	3(1)
C(9)	14(2)	16(2)	22(2)	6(2)	-3(2)	-3(1)
C(10)	18(2)	18(2)	26(2)	3(1)	-3(2)	-5(1)
C(11)	23(2)	35(2)	29(2)	12(2)	-10(2)	-9(2)
C(12)	17(2)	28(2)	47(2)	23(2)	-7(2)	-4(2)
C(13)	20(2)	15(2)	53(3)	9(2)	4(2)	2(1)
C(14)	22(2)	16(2)	30(2)	4(2)	6(2)	0(1)
C(15)	15(1)	12(2)	20(2)	1(1)	2(1)	-2(1)

Table 3.5. Hydrogen coordinates (x 10^4) and isotropic displacement parameters (Å^{2}x 10^3) for Complex **3.135**.

	x	y	z	U(eq)
H(1N)	1650(50)	1000(40)	5560(30)	21
H(4N)	2530(50)	-1120(40)	5170(40)	20
H(3)	-696	4227	5463	24
H(4)	-926	5255	6993	27
H(5)	71	4530	8480	29
H(6)	1338	2806	8418	29
H(7)	1617	1807	6883	24
H(10)	2102	-2278	6702	25
H(11)	2960	-3768	7734	35
H(12)	3499	-5594	7054	37
H(13)	3134	-5960	5314	35
H(14)	2190	-4498	4290	27
H(2N)	310(60)	2390(40)	3610(20)	41(16)
H(3N)	-220(60)	3440(30)	4170(40)	49(17)
H(5N)	280(40)	-2170(30)	3570(30)	18(12)
H(6N)	160(50)	-3290(30)	4040(40)	40(15)

Table 3.6. Torsion angles [°] for Complex **3.135** .

Cu(1)-N(1)-C(1)-N(2)	-9.3(6)
Cu(1)-N(1)-C(1)-C(2)	168.8(3)
N(1)-C(1)-C(2)-C(7)	7.9(5)
N(2)-C(1)-C(2)-C(7)	-173.9(3)
N(1)-C(1)-C(2)-C(3)	-174.0(4)
N(2)-C(1)-C(2)-C(3)	4.2(5)
C(7)-C(2)-C(3)-C(4)	-0.3(5)
C(1)-C(2)-C(3)-C(4)	-178.4(3)
C(2)-C(3)-C(4)-C(5)	-0.7(6)
C(3)-C(4)-C(5)-C(6)	0.9(6)
C(4)-C(5)-C(6)-C(7)	0.0(6)
C(3)-C(2)-C(7)-C(6)	1.3(6)
C(1)-C(2)-C(7)-C(6)	179.4(3)
C(5)-C(6)-C(7)-C(2)	-1.1(6)
Cu(1)-N(3)-C(8)-N(4)	-2.7(6)
Cu(1)-N(3)-C(8)-C(9)	175.7(3)
N(3)-C(8)-C(9)-C(14)	-137.6(4)
N(4)-C(8)-C(9)-C(14)	40.9(5)
N(3)-C(8)-C(9)-C(10)	40.3(5)
N(4)-C(8)-C(9)-C(10)	-141.2(4)
C(14)-C(9)-C(10)-C(11)	-0.9(6)
C(8)-C(9)-C(10)-C(11)	-178.9(3)
C(9)-C(10)-C(11)-C(12)	1.5(6)
C(10)-C(11)-C(12)-C(13)	-0.7(6)
C(11)-C(12)-C(13)-C(14)	-0.7(6)
C(10)-C(9)-C(14)-C(13)	-0.5(6)
C(8)-C(9)-C(14)-C(13)	177.4(3)
C(12)-C(13)-C(14)-C(9)	1.3(6)
Cu(1)-O(1)-C(15)-O(3)	179.1(3)
Cu(1)-O(1)-C(15)-O(2)	-1.0(3)
Cu(1)-O(2)-C(15)-O(3)	-179.1(3)
Cu(1)-O(2)-C(15)-O(1)	1.0(3)

Symmetry transformations used to generate equivalent atoms:

Table 3.7. Hydrogen bonds for Complex **3.135** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(3)-H(4N)...O(3)#1	0.80(5)	2.30(5)	3.066(4)	161(4)
N(2)-H(2N)...O(1)	0.874(14)	2.04(3)	2.850(4)	154(5)
N(2)-H(3N)...O(3)#2	0.884(14)	2.146(18)	3.017(4)	168(5)
N(4)-H(5N)...O(2)	0.869(14)	2.26(4)	2.872(4)	128(4)
N(4)-H(6N)...O(3)#3	0.874(14)	2.045(16)	2.911(4)	171(5)

Symmetry transformations used to generate equivalent atoms:

#1 -x+1/2,-y,z+1/2 #2 -x,y+1/2,-z+1/2 #3 -x,y-1/2,-z+1/2

Complex **3.136**

Table 3.8. Crystal data and structure refinement for **3.136**.

Identification code	burley_purp_check	
Empirical formula	C24 H34 Cu N4 O4	
Formula weight	506.09	
Temperature	123(2) K	
Wavelength	1.54184 Å	
Crystal system	Monoclinic	
Space group	P $2_1/c$	
Unit cell dimensions	a = 10.8059(9) Å	α= 90°.
	b = 11.1884(7) Å	β= 109.134(10)°.
	c = 10.8233(11) Å	γ = 90°.
Volume	1236.25(19) $Å^3$	
Z	2	
Density (calculated)	1.360 Mg/m^3	
Absorption coefficient	1.546 mm^{-1}	
F(000)	534	
Crystal size	0.38 x 0.14 x 0.02 mm^3	
Theta range for data collection	4.331 to 67.473°.	
Index ranges	-12<=h<=12, -12<=k<=13, -10<=l<=12	
Reflections collected	5457	
Independent reflections	2072 [R(int) = 0.1456]	
Completeness to theta = 67.474°	93.2 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	1.00000 and 0.33808	

Refinement method	Full-matrix least-squares on F^2
Data / restraints / parameters	2072 / 63 / 188
Goodness-of-fit on F^2	0.971
Final R indices [I>2sigma(I)]	R1 = 0.0895, wR2 = 0.2349
R indices (all data)	R1 = 0.1390, wR2 = 0.2743
Extinction coefficient	n/a
Largest diff. peak and hole	0.753 and -0.807 e.$Å^{-3}$

Table 3.9. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$x 10^3) for **3.136**. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
Cu(1)	0	0	10000	64(1)
N(1)	647(6)	298(5)	8557(6)	58(2)
N(2)	-429(6)	1987(5)	7519(9)	78(2)
C(2)	1153(7)	1114(5)	6677(6)	58(2)
C(7)	1538(7)	2177(6)	6282(6)	64(2)
C(3)	1434(6)	44(6)	6168(7)	60(2)
O(1)	-1658(9)	828(10)	8895(9)	50(2)
O(2)	-2129(10)	1871(9)	10434(10)	57(2)
C(8)	-2482(17)	1446(17)	9329(19)	49(3)
C(4)	2094(8)	53(7)	5273(7)	73(2)
C(1)	430(6)	1113(5)	7627(8)	62(2)
C(6)	2202(9)	2175(7)	5378(7)	74(3)
C(5)	2466(9)	1104(8)	4868(7)	74(2)
C(9)	-3855(15)	1513(12)	8345(14)	59(3)
C(10)	-3800(15)	2190(14)	7126(13)	71(3)
C(11)	-4373(15)	231(13)	7994(18)	75(3)
C(12)	-4679(14)	2181(15)	9036(16)	73(3)
C(12A)	-4261(14)	2445(12)	8065(16)	73(3)
C(10A)	-4138(16)	491(14)	7031(13)	71(3)
C(11A)	-4650(14)	416(15)	8963(16)	75(3)
C(8A)	-2381(15)	1055(15)	9082(18)	49(3)
C(9A)	-3826(14)	1160(11)	8334(14)	59(3)
O(1A)	-1844(8)	46(9)	8967(8)	50(2)
O(2A)	-1783(10)	1894(9)	9733(9)	57(2)

Table 3.10. Bond lengths [Å] and angles [°] for **3.136**

Cu(1)-N(1)#1	1.937(7)
Cu(1)-N(1)	1.937(7)
Cu(1)-O(1A)	1.940(8)
Cu(1)-O(1A)#1	1.940(8)
Cu(1)-O(1)#1	2.023(9)
Cu(1)-O(1)	2.023(9)
N(1)-C(1)	1.321(9)
N(1)-H(1N)	0.878(10)
N(2)-C(1)	1.327(8)
N(2)-H(2N)	0.877(10)
N(2)-H(3N)	0.876(10)
C(2)-C(7)	1.374(10)
C(2)-C(3)	1.393(9)
C(2)-C(1)	1.481(12)
C(7)-C(6)	1.390(12)
C(7)-H(7)	0.9500
C(3)-C(4)	1.378(12)
C(3)-H(3)	0.9500
O(1)-C(8)	1.329(18)
O(2)-C(8)	1.226(19)
C(8)-C(9)	1.519(9)
C(4)-C(5)	1.361(11)
C(4)-H(4)	0.9500
C(6)-C(5)	1.388(11)
C(6)-H(6)	0.9500
C(5)-H(5)	0.9500
C(9)-C(12)	1.532(10)
C(9)-C(10)	1.538(10)
C(9)-C(11)	1.541(10)
C(10)-H(10A)	0.9800
C(10)-H(10B)	0.9800
C(10)-H(10C)	0.9800
C(11)-H(11A)	0.9800
C(11)-H(11B)	0.9800
C(11)-H(11C)	0.9800
C(12)-H(12A)	0.9800

C(12)-H(12B)	0.9800
C(12)-H(12C)	0.9800
C(12A)-C(9A)	1.511(9)
C(12A)-H(12D)	0.9800
C(12A)-H(12E)	0.9800
C(12A)-H(12F)	0.9800
C(10A)-C(9A)	1.533(9)
C(10A)-H(10D)	0.9800
C(10A)-H(10E)	0.9800
C(10A)-H(10F)	0.9800
C(11A)-C(9A)	1.532(10)
C(11A)-H(11D)	0.9800
C(11A)-H(11E)	0.9800
C(11A)-H(11F)	0.9800
C(8A)-O(2A)	1.224(18)
C(8A)-O(1A)	1.294(16)
C(8A)-C(9A)	1.510(9)

N(1)#1-Cu(1)-N(1)	180.0
N(1)#1-Cu(1)-O(1A)	84.0(3)
N(1)-Cu(1)-O(1A)	96.0(3)
N(1)#1-Cu(1)-O(1A)#1	96.0(3)
N(1)-Cu(1)-O(1A)#1	84.0(3)
O(1A)-Cu(1)-O(1A)#1	180.0
N(1)#1-Cu(1)-O(1)#1	85.9(3)
N(1)-Cu(1)-O(1)#1	94.1(3)
N(1)#1-Cu(1)-O(1)	94.1(3)
N(1)-Cu(1)-O(1)	85.9(3)
O(1)#1-Cu(1)-O(1)	180.0
C(1)-N(1)-Cu(1)	135.4(5)
C(1)-N(1)-H(1N)	115(5)
Cu(1)-N(1)-H(1N)	109(5)
C(1)-N(2)-H(2N)	129(5)
C(1)-N(2)-H(3N)	128(6)
H(2N)-N(2)-H(3N)	100(2)
C(7)-C(2)-C(3)	119.6(8)
C(7)-C(2)-C(1)	120.0(6)
C(3)-C(2)-C(1)	120.5(6)

C(2)-C(7)-C(6)	119.8(7)
C(2)-C(7)-H(7)	120.1
C(6)-C(7)-H(7)	120.1
C(4)-C(3)-C(2)	120.1(8)
C(4)-C(3)-H(3)	119.9
C(2)-C(3)-H(3)	119.9
C(8)-O(1)-Cu(1)	126.5(9)
O(2)-C(8)-O(1)	121.1(13)
O(2)-C(8)-C(9)	125.8(14)
O(1)-C(8)-C(9)	113.1(13)
C(5)-C(4)-C(3)	120.6(8)
C(5)-C(4)-H(4)	119.7
C(3)-C(4)-H(4)	119.7
N(1)-C(1)-N(2)	121.4(9)
N(1)-C(1)-C(2)	121.8(6)
N(2)-C(1)-C(2)	116.8(7)
C(5)-C(6)-C(7)	120.1(8)
C(5)-C(6)-H(6)	119.9
C(7)-C(6)-H(6)	119.9
C(4)-C(5)-C(6)	119.8(9)
C(4)-C(5)-H(5)	120.1
C(6)-C(5)-H(5)	120.1
C(8)-C(9)-C(12)	105.8(12)
C(8)-C(9)-C(10)	108.8(13)
C(12)-C(9)-C(10)	111.7(14)
C(8)-C(9)-C(11)	108.6(13)
C(12)-C(9)-C(11)	110.6(14)
C(10)-C(9)-C(11)	111.1(14)
C(9)-C(10)-H(10A)	109.5
C(9)-C(10)-H(10B)	109.5
H(10A)-C(10)-H(10B)	109.5
C(9)-C(10)-H(10C)	109.5
H(10A)-C(10)-H(10C)	109.5
H(10B)-C(10)-H(10C)	109.5
C(9)-C(11)-H(11A)	109.5
C(9)-C(11)-H(11B)	109.5
H(11A)-C(11)-H(11B)	109.5
C(9)-C(11)-H(11C)	109.5

H(11A)-C(11)-H(11C)	109.5
H(11B)-C(11)-H(11C)	109.5
C(9)-C(12)-H(12A)	109.5
C(9)-C(12)-H(12B)	109.5
H(12A)-C(12)-H(12B)	109.5
C(9)-C(12)-H(12C)	109.5
H(12A)-C(12)-H(12C)	109.5
H(12B)-C(12)-H(12C)	109.5
C(9A)-C(12A)-H(12D)	109.5
C(9A)-C(12A)-H(12E)	109.5
H(12D)-C(12A)-H(12E)	109.5
C(9A)-C(12A)-H(12F)	109.5
H(12D)-C(12A)-H(12F)	109.5
H(12E)-C(12A)-H(12F)	109.5
C(9A)-C(10A)-H(10D)	109.5
C(9A)-C(10A)-H(10E)	109.5
H(10D)-C(10A)-H(10E)	109.5
C(9A)-C(10A)-H(10F)	109.5
H(10D)-C(10A)-H(10F)	109.5
H(10E)-C(10A)-H(10F)	109.5
C(9A)-C(11A)-H(11D)	109.5
C(9A)-C(11A)-H(11E)	109.5
H(11D)-C(11A)-H(11E)	109.5
C(9A)-C(11A)-H(11F)	109.5
H(11D)-C(11A)-H(11F)	109.5
H(11E)-C(11A)-H(11F)	109.5
O(2A)-C(8A)-O(1A)	123.7(13)
O(2A)-C(8A)-C(9A)	120.5(12)
O(1A)-C(8A)-C(9A)	115.8(12)
C(8A)-C(9A)-C(12A)	112.3(11)
C(8A)-C(9A)-C(11A)	111.6(13)
C(12A)-C(9A)-C(11A)	114.4(14)
C(8A)-C(9A)-C(10A)	109.4(12)
C(12A)-C(9A)-C(10A)	109.1(12)
C(11A)-C(9A)-C(10A)	99.2(13)
C(8A)-O(1A)-Cu(1)	112.1(9)

Symmetry transformations used to generate equivalent atoms:

#1 -x,-y,-z+2

Table 3.11. Anisotropic displacement parameters ($Å^2$x 10^3)for **3.136**. The anisotropic displacement factor exponent takes the form: $-2\pi^2[h^2a^{*2}U^{11} + ... + 2 h k a^* b^* U^{12}]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
Cu(1)	44(1)	68(1)	58(1)	-28(1)	-13(1)	17(1)
N(1)	50(3)	36(3)	70(3)	-4(2)	-3(3)	16(2)
N(2)	56(4)	27(3)	133(6)	2(3)	3(4)	13(2)
C(2)	54(4)	26(3)	62(4)	-6(3)	-27(3)	7(3)
C(7)	67(4)	39(4)	54(4)	-1(3)	-24(4)	10(3)
C(3)	47(3)	42(4)	64(4)	-9(3)	-20(3)	9(3)
O(1)	37(3)	54(5)	49(3)	5(4)	0(2)	11(4)
O(2)	52(5)	41(3)	60(5)	-15(4)	-6(3)	4(3)
C(8)	40(4)	34(9)	61(6)	6(4)	2(4)	18(5)
C(4)	71(5)	53(5)	63(4)	-12(4)	-20(4)	13(4)
C(1)	43(3)	27(3)	88(5)	-15(3)	-16(3)	5(2)
C(6)	90(6)	47(5)	56(4)	13(3)	-17(4)	7(4)
C(5)	78(5)	71(6)	52(4)	9(3)	-9(4)	11(4)
C(9)	42(3)	47(7)	70(4)	9(4)	-5(3)	10(4)
C(10)	61(6)	69(6)	60(5)	-4(4)	-11(4)	14(5)
C(11)	50(6)	80(6)	82(7)	-7(5)	4(5)	-6(5)
C(12)	53(5)	66(7)	82(7)	-3(5)	-2(5)	5(5)
C(12A)	53(5)	66(7)	82(7)	-3(5)	-2(5)	5(5)
C(10A)	61(6)	69(6)	60(5)	-4(4)	-11(4)	14(5)
C(11A)	50(6)	80(6)	82(7)	-7(5)	4(5)	-6(5)
C(8A)	40(4)	34(9)	61(6)	6(4)	2(4)	18(5)
C(9A)	42(3)	47(7)	70(4)	9(4)	-5(3)	10(4)
O(1A)	37(3)	54(5)	49(3)	5(4)	0(2)	11(4)
O(2A)	52(5)	41(3)	60(5)	-15(4)	-6(3)	4(3)

Table 3.12. Hydrogen coordinates (x 10^4) and isotropic displacement parameters ($Å^2$x 10^3) for **3.136**.

	x	y	z	U(eq)
H(1N)	1270(50)	-210(50)	8600(80)	69
H(2N)	-910(70)	2370(70)	6820(40)	94
H(3N)	-850(70)	2180(80)	8060(60)	94
H(7)	1352	2911	6626	77
H(3)	1171	-694	6438	72
H(4)	2290	-681	4935	87
H(6)	2475	2909	5108	89
H(5)	2906	1104	4238	89
H(10A)	-3236	1756	6732	107
H(10B)	-3445	2993	7380	107
H(10C)	-4685	2253	6492	107
H(11A)	-3764	-218	7669	113
H(11B)	-5235	259	7316	113
H(11C)	-4447	-164	8775	113
H(12A)	-5596	2198	8468	109
H(12B)	-4353	3001	9227	109
H(12C)	-4615	1771	9855	109
H(12D)	-4126	2863	8895	109
H(12E)	-5192	2470	7540	109
H(12F)	-3746	2837	7585	109
H(10D)	-3762	925	6453	107
H(10E)	-5090	434	6617	107
H(10F)	-3762	-314	7191	107
H(11D)	-4361	-419	9030	113
H(11E)	-5575	464	8424	113
H(11F)	-4541	728	9839	113

Table 3.13. Torsion angles [°] for **3.136**.

C(3)-C(2)-C(7)-C(6)	0.1(8)
C(1)-C(2)-C(7)-C(6)	-179.2(5)
C(7)-C(2)-C(3)-C(4)	-0.1(8)
C(1)-C(2)-C(3)-C(4)	179.3(5)
Cu(1)-O(1)-C(8)-O(2)	22(3)
Cu(1)-O(1)-C(8)-C(9)	-156.2(11)
C(2)-C(3)-C(4)-C(5)	-0.6(9)
Cu(1)-N(1)-C(1)-N(2)	-1.3(10)
Cu(1)-N(1)-C(1)-C(2)	177.2(4)
C(7)-C(2)-C(1)-N(1)	-141.4(6)
C(3)-C(2)-C(1)-N(1)	39.3(8)
C(7)-C(2)-C(1)-N(2)	37.2(8)
C(3)-C(2)-C(1)-N(2)	-142.1(6)
C(2)-C(7)-C(6)-C(5)	0.5(9)
C(3)-C(4)-C(5)-C(6)	1.2(10)
C(7)-C(6)-C(5)-C(4)	-1.1(10)
O(2)-C(8)-C(9)-C(12)	-1(3)
O(1)-C(8)-C(9)-C(12)	176.6(16)
O(2)-C(8)-C(9)-C(10)	119(2)
O(1)-C(8)-C(9)-C(10)	-63(2)
O(2)-C(8)-C(9)-C(11)	-120(2)
O(1)-C(8)-C(9)-C(11)	58(2)
O(2A)-C(8A)-C(9A)-C(12A)	22(3)
O(1A)-C(8A)-C(9A)-C(12A)	-157.1(16)
O(2A)-C(8A)-C(9A)-C(11A)	-108(2)
O(1A)-C(8A)-C(9A)-C(11A)	73(2)
O(2A)-C(8A)-C(9A)-C(10A)	143.2(19)
O(1A)-C(8A)-C(9A)-C(10A)	-36(2)
O(2A)-C(8A)-O(1A)-Cu(1)	2(3)
C(9A)-C(8A)-O(1A)-Cu(1)	-178.7(12)

Symmetry transformations used to generate equivalent atoms:

#1 -x,-y,-z+2

(*Z*)-*N*'-(4-(trifluoromethyl)phenyl)benzimidamide (**3.146**)

Table 3.14. Crystal data and structure refinement for **3.146**.

Identification code	bur_aug19	
Empirical formula	C14 H11 F3 N2	
Formula weight	264.25	
Temperature	123(2) K	
Wavelength	1.54184 Å	
Crystal system	Triclinic	
Space group	P -1	
Unit cell dimensions	a = 5.4166(7) Å	α= 76.595(11)°.
	b = 8.4710(11) Å	β= 84.634(11)°.
	c = 13.9435(19) Å	γ = 86.123(11)°.
Volume	618.93(14) $Å^3$	
Z	2	
Density (calculated)	1.418 Mg/m^3	
Absorption coefficient	1.004 mm^{-1}	
F(000)	272	
Crystal size	0.40 x 0.20 x 0.10 mm^3	
Theta range for data collection	5.623 to 69.996°.	
Index ranges	-6<=h<=4, -10<=k<=10, -16<=l<=16	
Reflections collected	3917	
Independent reflections	2344 [R(int) = 0.0344]	
Completeness to theta = 69.996°	99.4 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	1.00000 and 0.30576	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	2344 / 0 / 180	
Goodness-of-fit on F^2	1.044	
Final R indices [I>2sigma(I)]	R1 = 0.0601, wR2 = 0.1651	
R indices (all data)	R1 = 0.0750, wR2 = 0.1850	
Extinction coefficient	n/a	
Largest diff. peak and hole	0.381 and -0.318 $e.Å^{-3}$	

Table 3.15. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$x 10^3) for **3.146**. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
F(1)	6528(5)	960(2)	557(1)	82(1)
F(2)	7395(5)	3427(3)	308(2)	88(1)
F(3)	3805(5)	2860(4)	120(2)	100(1)
N(1)	3415(3)	2582(2)	4737(2)	35(1)
N(2)	-897(4)	2437(3)	4734(2)	45(1)
C(1)	5600(6)	2390(4)	697(2)	52(1)
C(2)	4898(5)	2386(3)	1754(2)	40(1)
C(3)	3361(5)	3641(3)	1995(2)	42(1)
C(4)	2818(4)	3707(3)	2974(2)	39(1)
C(5)	3774(4)	2516(3)	3730(2)	34(1)
C(6)	5291(4)	1251(3)	3478(2)	36(1)
C(7)	5858(4)	1193(3)	2498(2)	39(1)
C(8)	1184(4)	2575(3)	5161(2)	35(1)
C(9)	865(4)	2607(3)	6228(2)	35(1)
C(10)	-1061(4)	1825(3)	6850(2)	39(1)
C(11)	-1262(5)	1781(3)	7854(2)	42(1)
C(12)	468(5)	2525(3)	8248(2)	43(1)
C(13)	2373(5)	3329(3)	7633(2)	42(1)
C(14)	2596(4)	3369(3)	6627(2)	38(1)

Table 3.16. Bond lengths [Å] and angles [°] for **3.146**.

F(1)-C(1)	1.330(3)
F(2)-C(1)	1.344(4)
F(3)-C(1)	1.302(4)
N(1)-C(8)	1.294(3)
N(1)-C(5)	1.413(3)
N(2)-C(8)	1.347(3)
N(2)-H(1N)	0.86(3)
N(2)-H(2N)	0.88(3)
C(1)-C(2)	1.488(4)
C(2)-C(7)	1.382(4)
C(2)-C(3)	1.389(4)
C(3)-C(4)	1.383(4)
C(3)-H(3)	0.9500
C(4)-C(5)	1.391(3)
C(4)-H(4)	0.9500
C(5)-C(6)	1.396(3)
C(6)-C(7)	1.384(3)
C(6)-H(6)	0.9500
C(7)-H(7)	0.9500
C(8)-C(9)	1.489(3)
C(9)-C(10)	1.392(3)
C(9)-C(14)	1.397(3)
C(10)-C(11)	1.386(4)
C(10)-H(10)	0.9500
C(11)-C(12)	1.383(4)
C(11)-H(11)	0.9500
C(12)-C(13)	1.386(4)
C(12)-H(12)	0.9500
C(13)-C(14)	1.389(4)
C(13)-H(13)	0.9500
C(14)-H(14)	0.9500
C(8)-N(1)-C(5)	119.31(19)
C(8)-N(2)-H(1N)	121.6(19)
C(8)-N(2)-H(2N)	118(2)
H(1N)-N(2)-H(2N)	118(3)

F(3)-C(1)-F(1)	108.0(3)
F(3)-C(1)-F(2)	104.2(3)
F(1)-C(1)-F(2)	104.6(3)
F(3)-C(1)-C(2)	114.2(3)
F(1)-C(1)-C(2)	113.3(2)
F(2)-C(1)-C(2)	111.7(2)
C(7)-C(2)-C(3)	119.8(2)
C(7)-C(2)-C(1)	120.7(2)
C(3)-C(2)-C(1)	119.5(2)
C(4)-C(3)-C(2)	120.1(2)
C(4)-C(3)-H(3)	120.0
C(2)-C(3)-H(3)	120.0
C(3)-C(4)-C(5)	120.8(2)
C(3)-C(4)-H(4)	119.6
C(5)-C(4)-H(4)	119.6
C(4)-C(5)-C(6)	118.5(2)
C(4)-C(5)-N(1)	123.1(2)
C(6)-C(5)-N(1)	118.3(2)
C(7)-C(6)-C(5)	120.8(2)
C(7)-C(6)-H(6)	119.6
C(5)-C(6)-H(6)	119.6
C(2)-C(7)-C(6)	120.1(2)
C(2)-C(7)-H(7)	120.0
C(6)-C(7)-H(7)	120.0
N(1)-C(8)-N(2)	125.5(2)
N(1)-C(8)-C(9)	118.2(2)
N(2)-C(8)-C(9)	116.2(2)
C(10)-C(9)-C(14)	119.1(2)
C(10)-C(9)-C(8)	121.3(2)
C(14)-C(9)-C(8)	119.6(2)
C(11)-C(10)-C(9)	120.9(2)
C(11)-C(10)-H(10)	119.6
C(9)-C(10)-H(10)	119.6
C(12)-C(11)-C(10)	119.8(2)
C(12)-C(11)-H(11)	120.1
C(10)-C(11)-H(11)	120.1
C(11)-C(12)-C(13)	119.8(2)
C(11)-C(12)-H(12)	120.1

C(13)-C(12)-H(12)	120.1
C(12)-C(13)-C(14)	120.7(2)
C(12)-C(13)-H(13)	119.6
C(14)-C(13)-H(13)	119.6
C(13)-C(14)-C(9)	119.6(2)
C(13)-C(14)-H(14)	120.2
C(9)-C(14)-H(14)	120.2

Symmetry transformations used to generate equivalent atoms:

Table 3.16. Anisotropic displacement parameters ($Å^2 x\ 10^3$)for bur_aug19. The anisotropic displacement factor exponent takes the form: $-2\pi^2[\ h^2 a^{*2} U^{11} + ... + 2\ h\ k\ a^*\ b^*\ U^{12}\ ]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
F(1)	148(2)	49(1)	49(1)	-20(1)	5(1)	20(1)
F(2)	124(2)	81(2)	58(1)	-21(1)	32(1)	-30(1)
F(3)	105(2)	152(3)	52(1)	-43(1)	-29(1)	42(2)
N(1)	39(1)	30(1)	38(1)	-10(1)	-4(1)	-2(1)
N(2)	36(1)	61(1)	40(1)	-16(1)	-2(1)	-6(1)
C(1)	71(2)	41(1)	44(2)	-12(1)	-6(1)	8(1)
C(2)	47(1)	33(1)	42(1)	-13(1)	-4(1)	-3(1)
C(3)	54(1)	31(1)	38(1)	-4(1)	-6(1)	2(1)
C(4)	44(1)	29(1)	44(1)	-11(1)	-2(1)	3(1)
C(5)	35(1)	30(1)	40(1)	-11(1)	-5(1)	-6(1)
C(6)	40(1)	28(1)	41(1)	-7(1)	-6(1)	-3(1)
C(7)	44(1)	31(1)	45(1)	-13(1)	-3(1)	0(1)
C(8)	38(1)	25(1)	42(1)	-8(1)	-5(1)	-1(1)
C(9)	36(1)	27(1)	42(1)	-11(1)	-4(1)	4(1)
C(10)	41(1)	33(1)	44(1)	-11(1)	-6(1)	-1(1)
C(11)	42(1)	36(1)	46(1)	-7(1)	2(1)	-1(1)
C(12)	50(1)	42(1)	39(1)	-14(1)	-5(1)	9(1)
C(13)	44(1)	39(1)	46(1)	-17(1)	-10(1)	4(1)
C(14)	38(1)	33(1)	43(1)	-11(1)	-3(1)	0(1)

Table 3.17. Hydrogen coordinates (x 10^4) and isotropic displacement parameters ($Å^2$x 10^3) for **3.146**.

	x	y	z	U(eq)
H(3)	2681	4455	1485	50
H(4)	1781	4576	3132	47
H(6)	5941	419	3986	43
H(7)	6909	332	2336	47
H(10)	-2254	1313	6582	46
H(11)	-2584	1241	8270	50
H(12)	351	2485	8937	52
H(13)	3539	3858	7903	50
H(14)	3919	3912	6213	45
H(1N)	-2340(50)	2490(30)	5040(20)	38(7)
H(2N)	-800(50)	2530(40)	4090(30)	47(8)

Table 3.18. Torsion angles [°] for **3.146**.

F(3)-C(1)-C(2)-C(7)	143.4(3)
F(1)-C(1)-C(2)-C(7)	19.2(4)
F(2)-C(1)-C(2)-C(7)	-98.6(3)
F(3)-C(1)-C(2)-C(3)	-39.9(4)
F(1)-C(1)-C(2)-C(3)	-164.0(3)
F(2)-C(1)-C(2)-C(3)	78.1(3)
C(7)-C(2)-C(3)-C(4)	0.8(4)
C(1)-C(2)-C(3)-C(4)	-176.0(2)
C(2)-C(3)-C(4)-C(5)	-0.8(4)
C(3)-C(4)-C(5)-C(6)	-0.1(3)
C(3)-C(4)-C(5)-N(1)	175.6(2)
C(8)-N(1)-C(5)-C(4)	61.6(3)
C(8)-N(1)-C(5)-C(6)	-122.7(2)
C(4)-C(5)-C(6)-C(7)	0.8(3)
N(1)-C(5)-C(6)-C(7)	-175.11(19)
C(3)-C(2)-C(7)-C(6)	-0.1(4)
C(1)-C(2)-C(7)-C(6)	176.7(2)
C(5)-C(6)-C(7)-C(2)	-0.7(3)
C(5)-N(1)-C(8)-N(2)	3.1(3)
C(5)-N(1)-C(8)-C(9)	178.60(19)
N(1)-C(8)-C(9)-C(10)	-147.3(2)
N(2)-C(8)-C(9)-C(10)	28.6(3)
N(1)-C(8)-C(9)-C(14)	29.6(3)
N(2)-C(8)-C(9)-C(14)	-154.5(2)
C(14)-C(9)-C(10)-C(11)	-0.6(4)
C(8)-C(9)-C(10)-C(11)	176.3(2)
C(9)-C(10)-C(11)-C(12)	0.1(4)
C(10)-C(11)-C(12)-C(13)	0.9(4)
C(11)-C(12)-C(13)-C(14)	-1.3(4)
C(12)-C(13)-C(14)-C(9)	0.7(4)
C(10)-C(9)-C(14)-C(13)	0.2(3)
C(8)-C(9)-C(14)-C(13)	-176.7(2)

Symmetry transformations used to generate equivalent atoms:

Table 3.19. Hydrogen bonds for **3.146** [Å and °]

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(2)-H(1N)...N(1)#1	0.86(3)	2.37(3)	3.074(3)	140(2)

Symmetry transformations used to generate equivalent atoms:

#1 x-1,y,z

(*Z*)-*N*'-(3-fluoro-4-methoxyphenyl)benzimidamide (**3.148**)

Table 3.20. Crystal data and structure refinement for **3.148**.

Identification code	burley_jr0244_nmr	
Empirical formula	C14 H13 F N2 O	
Formula weight	244.26	
Temperature	123(2) K	
Wavelength	1.54184 Å	
Crystal system	Monoclinic	
Space group	P 2_1/c	
Unit cell dimensions	a = 13.3307(4) Å	α= 90°.
	b = 10.3332(3) Å	β= 99.032(3)°.
	c = 8.6427(3) Å	γ = 90°.
Volume	1175.76(6) Å^3	
Z	4	
Density (calculated)	1.380 Mg/m^3	
Absorption coefficient	0.822 mm^{-1}	
F(000)	512	
Crystal size	0.4 x 0.25 x 0.2 mm^3	
Theta range for data collection	3.357 to 73.080°.	
Index ranges	-14<=h<=16, -8<=k<=12, -9<=l<=10	
Reflections collected	6721	
Independent reflections	2315 [R(int) = 0.0289]	
Completeness to theta = 70.000°	99.9 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	1.00000 and 0.07105	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	2315 / 0 / 172	
Goodness-of-fit on F^2	1.073	
Final R indices [I>2sigma(I)]	R1 = 0.0446, wR2 = 0.1235	
R indices (all data)	R1 = 0.0511, wR2 = 0.1292	

Extinction coefficient	n/a
Largest diff. peak and hole	0.251 and -0.236 e.Å^{-3}

Table 3.21. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters (Å^{2}x 10^3) for **3.148**. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
F(1)	13896(1)	-508(1)	8081(1)	44(1)
O(1)	14579(1)	1120(1)	6134(1)	34(1)
N(1)	10196(1)	-614(1)	6934(2)	31(1)
N(2)	10616(1)	1588(1)	7469(1)	25(1)
C(1)	10009(1)	600(1)	7415(2)	24(1)
C(2)	8977(1)	818(1)	7820(2)	25(1)
C(3)	8137(1)	193(1)	6976(2)	29(1)
C(4)	7168(1)	458(2)	7283(2)	31(1)
C(5)	7024(1)	1349(2)	8425(2)	33(1)
C(6)	7860(1)	1959(1)	9289(2)	32(1)
C(7)	8830(1)	1695(1)	8996(2)	27(1)
C(8)	11607(1)	1429(1)	7101(2)	25(1)
C(9)	12287(1)	469(1)	7778(2)	29(1)
C(10)	13250(1)	420(1)	7417(2)	30(1)
C(11)	13607(1)	1292(1)	6401(2)	27(1)
C(12)	12941(1)	2251(1)	5756(2)	28(1)
C(13)	11957(1)	2317(1)	6105(2)	27(1)
C(14)	14925(1)	1977(2)	5029(2)	36(1)

Table 3.22. Bond lengths [Å] and angles [°] for **3.148**.

F(1)-C(10)	1.3544(17)
O(1)-C(11)	1.3631(18)
O(1)-C(14)	1.4306(19)
N(1)-C(1)	1.3566(18)
N(1)-H(1N)	0.889(19)
N(1)-H(2N)	0.86(2)
N(2)-C(1)	1.2995(19)
N(2)-C(8)	1.4165(18)
C(1)-C(2)	1.4882(19)
C(2)-C(3)	1.395(2)
C(2)-C(7)	1.398(2)
C(3)-C(4)	1.386(2)
C(3)-H(3)	0.9500
C(4)-C(5)	1.385(2)
C(4)-H(4)	0.9500
C(5)-C(6)	1.392(2)
C(5)-H(5)	0.9500
C(6)-C(7)	1.384(2)
C(6)-H(6)	0.9500
C(7)-H(7)	0.9500
C(8)-C(13)	1.389(2)
C(8)-C(9)	1.407(2)
C(9)-C(10)	1.369(2)
C(9)-H(9)	0.9500
C(10)-C(11)	1.393(2)
C(11)-C(12)	1.388(2)
C(12)-C(13)	1.3928(19)
C(12)-H(12)	0.9500
C(13)-H(13)	0.9500
C(14)-H(14A)	0.9800
C(14)-H(14B)	0.9800
C(14)-H(14C)	0.9800
C(11)-O(1)-C(14)	116.49(12)
C(1)-N(1)-H(1N)	120.0(12)
C(1)-N(1)-H(2N)	122.0(14)

H(1N)-N(1)-H(2N)	116.3(17)
C(1)-N(2)-C(8)	119.88(12)
N(2)-C(1)-N(1)	126.44(14)
N(2)-C(1)-C(2)	117.73(12)
N(1)-C(1)-C(2)	115.71(13)
C(3)-C(2)-C(7)	119.14(13)
C(3)-C(2)-C(1)	120.17(13)
C(7)-C(2)-C(1)	120.63(13)
C(4)-C(3)-C(2)	120.33(14)
C(4)-C(3)-H(3)	119.8
C(2)-C(3)-H(3)	119.8
C(5)-C(4)-C(3)	120.32(14)
C(5)-C(4)-H(4)	119.8
C(3)-C(4)-H(4)	119.8
C(4)-C(5)-C(6)	119.67(14)
C(4)-C(5)-H(5)	120.2
C(6)-C(5)-H(5)	120.2
C(7)-C(6)-C(5)	120.32(14)
C(7)-C(6)-H(6)	119.8
C(5)-C(6)-H(6)	119.8
C(6)-C(7)-C(2)	120.20(14)
C(6)-C(7)-H(7)	119.9
C(2)-C(7)-H(7)	119.9
C(13)-C(8)-C(9)	117.83(13)
C(13)-C(8)-N(2)	118.57(13)
C(9)-C(8)-N(2)	123.45(13)
C(10)-C(9)-C(8)	119.80(14)
C(10)-C(9)-H(9)	120.1
C(8)-C(9)-H(9)	120.1
F(1)-C(10)-C(9)	119.14(14)
F(1)-C(10)-C(11)	117.81(14)
C(9)-C(10)-C(11)	123.05(14)
O(1)-C(11)-C(12)	126.18(14)
O(1)-C(11)-C(10)	116.77(13)
C(12)-C(11)-C(10)	117.05(14)
C(11)-C(12)-C(13)	120.84(14)
C(11)-C(12)-H(12)	119.6
C(13)-C(12)-H(12)	119.6

C(8)-C(13)-C(12)	121.41(13)
C(8)-C(13)-H(13)	119.3
C(12)-C(13)-H(13)	119.3
O(1)-C(14)-H(14A)	109.5
O(1)-C(14)-H(14B)	109.5
H(14A)-C(14)-H(14B)	109.5
O(1)-C(14)-H(14C)	109.5
H(14A)-C(14)-H(14C)	109.5
H(14B)-C(14)-H(14C)	109.5

Symmetry transformations used to generate equivalent atoms:

Table 3.23. Anisotropic displacement parameters ($Å^2 x\ 10^3$) for **3.148**. The anisotropic displacement factor exponent takes the form: $-2\pi^2[\ h^2 a^{*2} U^{11} + ... + 2\ h\ k\ a^*\ b^*\ U^{12}\ ]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
F(1)	34(1)	43(1)	57(1)	21(1)	12(1)	14(1)
O(1)	26(1)	35(1)	42(1)	7(1)	9(1)	1(1)
N(1)	30(1)	21(1)	46(1)	-4(1)	13(1)	-2(1)
N(2)	25(1)	21(1)	30(1)	-1(1)	5(1)	-1(1)
C(1)	28(1)	20(1)	25(1)	1(1)	4(1)	0(1)
C(2)	29(1)	19(1)	29(1)	4(1)	6(1)	0(1)
C(3)	30(1)	24(1)	32(1)	-1(1)	5(1)	-1(1)
C(4)	27(1)	27(1)	38(1)	1(1)	4(1)	-3(1)
C(5)	29(1)	30(1)	42(1)	4(1)	12(1)	3(1)
C(6)	38(1)	24(1)	35(1)	-2(1)	13(1)	0(1)
C(7)	31(1)	21(1)	29(1)	1(1)	7(1)	-3(1)
C(8)	25(1)	22(1)	27(1)	-3(1)	4(1)	-1(1)
C(9)	31(1)	24(1)	32(1)	4(1)	7(1)	0(1)
C(10)	30(1)	25(1)	34(1)	3(1)	3(1)	4(1)
C(11)	24(1)	26(1)	30(1)	-3(1)	4(1)	-2(1)
C(12)	30(1)	24(1)	32(1)	2(1)	5(1)	-4(1)
C(13)	26(1)	21(1)	33(1)	0(1)	2(1)	0(1)
C(14)	29(1)	36(1)	46(1)	5(1)	12(1)	-4(1)

Table 3.24. Hydrogen coordinates (x 10^4) and isotropic displacement parameters ($Å^2$x 10^3) for **3.148**.

	x	y	z	U(eq)
H(3)	8229	-416	6188	34
H(4)	6600	26	6707	37
H(5)	6358	1543	8617	39
H(6)	7764	2561	10085	38
H(7)	9398	2109	9596	32
H(9)	12078	-144	8483	34
H(12)	13159	2871	5067	34
H(13)	11516	2985	5651	32
H(14A)	14476	1907	4020	55
H(14B)	15619	1743	4898	55
H(14C)	14915	2868	5414	55
H(1N)	10786(15)	-794(18)	6620(20)	33(5)
H(2N)	9836(15)	-1260(20)	7130(20)	43(5)

Table 3.25. Torsion angles [°] for **3.148**.

C(8)-N(2)-C(1)-N(1)	-4.5(2)
C(8)-N(2)-C(1)-C(2)	179.70(12)
N(2)-C(1)-C(2)-C(3)	140.81(15)
N(1)-C(1)-C(2)-C(3)	-35.5(2)
N(2)-C(1)-C(2)-C(7)	-36.2(2)
N(1)-C(1)-C(2)-C(7)	147.53(14)
C(7)-C(2)-C(3)-C(4)	1.2(2)
C(1)-C(2)-C(3)-C(4)	-175.84(13)
C(2)-C(3)-C(4)-C(5)	0.3(2)
C(3)-C(4)-C(5)-C(6)	-1.4(2)
C(4)-C(5)-C(6)-C(7)	1.0(2)
C(5)-C(6)-C(7)-C(2)	0.5(2)
C(3)-C(2)-C(7)-C(6)	-1.6(2)
C(1)-C(2)-C(7)-C(6)	175.45(13)
C(1)-N(2)-C(8)-C(13)	133.80(15)
C(1)-N(2)-C(8)-C(9)	-50.8(2)
C(13)-C(8)-C(9)-C(10)	-1.4(2)
N(2)-C(8)-C(9)-C(10)	-176.77(13)
C(8)-C(9)-C(10)-F(1)	179.83(14)
C(8)-C(9)-C(10)-C(11)	0.4(2)
C(14)-O(1)-C(11)-C(12)	-3.0(2)
C(14)-O(1)-C(11)-C(10)	176.96(14)
F(1)-C(10)-C(11)-O(1)	1.3(2)
C(9)-C(10)-C(11)-O(1)	-179.28(14)
F(1)-C(10)-C(11)-C(12)	-178.76(13)
C(9)-C(10)-C(11)-C(12)	0.7(2)
O(1)-C(11)-C(12)-C(13)	179.17(14)
C(10)-C(11)-C(12)-C(13)	-0.8(2)
C(9)-C(8)-C(13)-C(12)	1.3(2)
N(2)-C(8)-C(13)-C(12)	176.90(13)
C(11)-C(12)-C(13)-C(8)	-0.2(2)

Symmetry transformations used to generate equivalent atoms:

Table 3.26. Hydrogen bonds for **3.148** [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
N(1)-H(2N)...N(2)#1	0.86(2)	2.34(2)	3.1579(18)	159.6(18)

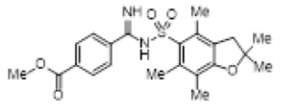
Symmetry transformations used to generate equivalent atoms:

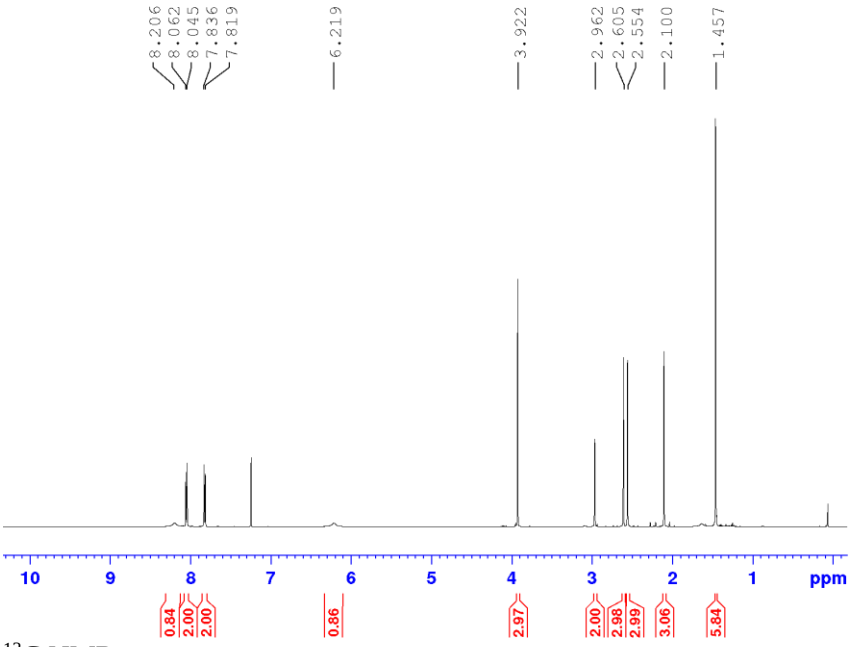
#1 -x+2,y-1/2,-z+3/2

Chapter 4: NMR spectra, HPLC traces, HR mass spectra, Flow cytometry histograms and microscope images.

NMR Spectra

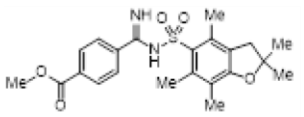
Methyl 4-(*N*-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)carbamimidoyl)benzoate (**4.85**) – ^{1}H NMR

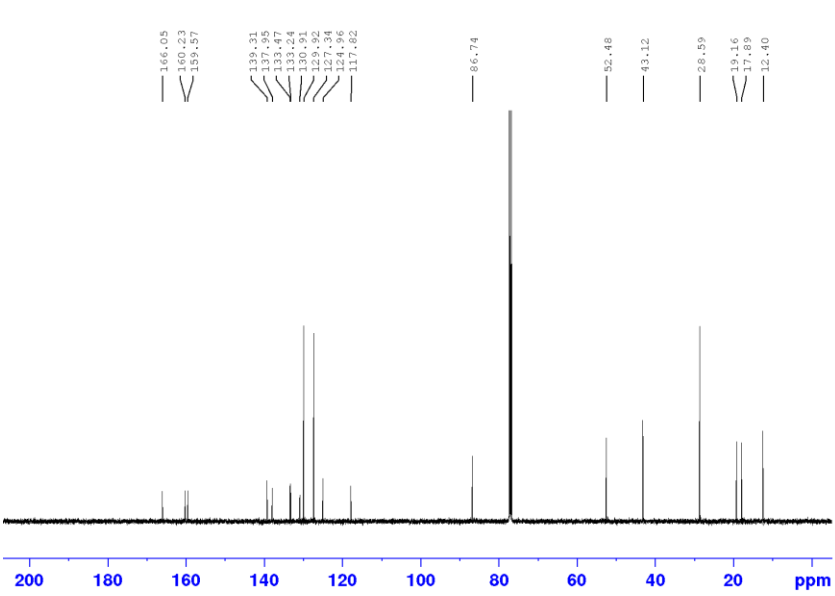





```
Current Data Parameters
NAME       JR0423-E48413
EXPNO                  1
PROCNO                 1

F2 - Acquisition Parameters
Date_           20200924
Time                1.07 h
INSTRUM            spect
PROBHD    Z113652_0211 (
PULPROG             zg30
TD                 39578
SOLVENT            CDCl3
NS                    16
DS                     2
SWH            10000.000 Hz
FIDRES         0.505331 Hz
AQ             1.9789000 sec
RG                   133
DW                50.000 usec
DE                 6.50 usec
TE                300.0 K
D1             2.00000000 sec
TD0                    1
SFO1         500.1330885 MHz
NUC1                  1H
P0                 3.53 usec
P1                10.58 usec
PLW1         20.00000000 W

F2 - Processing parameters
SI                 65536
SF            500.1300183 MHz
WDW                   EM
SSB       0
LB                 0.30 Hz
GB        0
PC                 1.00
```

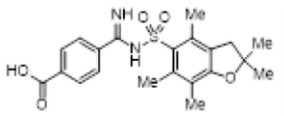
^{13}C NMR

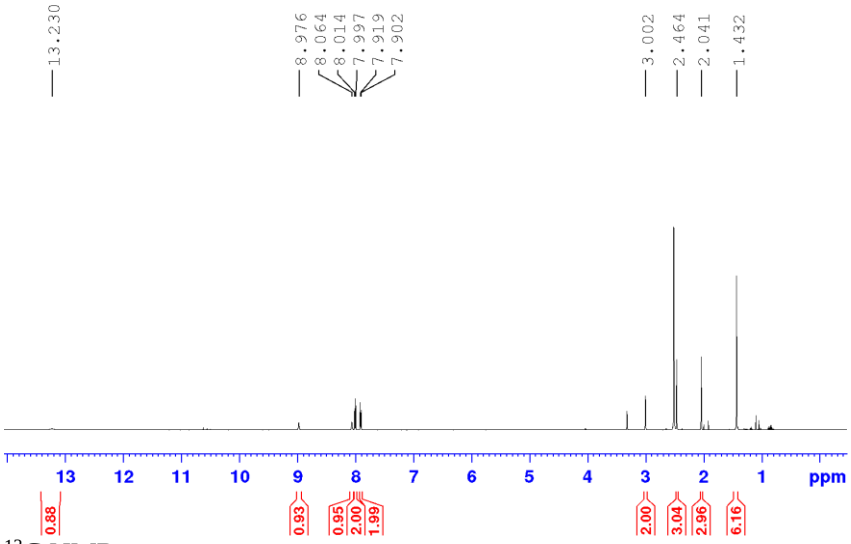





```
Current Data Parameters
NAME       JR0423-E48413
EXPNO                  2
PROCNO                 1

F2 - Acquisition Parameters
Date_           20200924
Time                1.48 h
INSTRUM            spect
PROBHD    Z113652_0211 (
PULPROG           zgpg30
TD                 23806
SOLVENT            CDCl3
NS                  2048
DS                     4
SWH            29761.904 Hz
FIDRES         2.500370 Hz
AQ             0.3999408 sec
RG                 198.22
DW                16.800 usec
DE                 6.50 usec
TE                300.0 K
D1             0.69999999 sec
D11            0.03000000 sec
TD0                    1
SFO1         125.7703637 MHz
NUC1                  13C
P0                 3.33 usec
P1                10.00 usec
PLW1         76.00000000 W
SFO2         500.1320005 MHz
NUC2                  1H
CPDPRG[2          waltz16
PCPD2             80.00 usec
PLW2         21.00000000 W
PLW12         0.31513000 W
PLW13         0.15851000 W

F2 - Processing parameters
SI                 32768
SF           125.7577885 MHz
WDW                   EM
SSB       0
LB                 1.00 Hz
GB        0
PC                 1.40
```

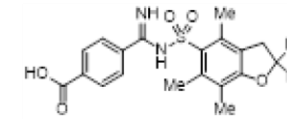
4-(*N*-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)carbamimidoyl)benzoic acid (**4.86**) – ^{1}H NMR

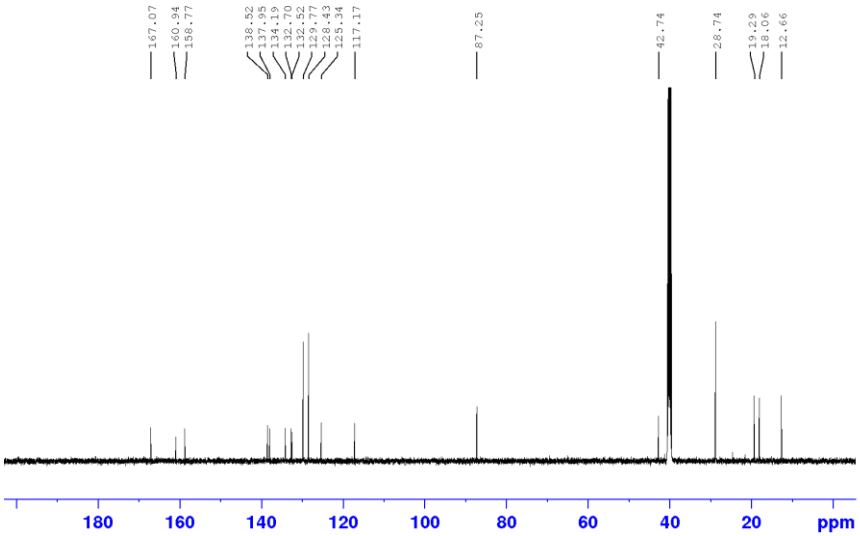





```
Current Data Parameters
NAME      JR0424-P1-E48536
EXPNO                    1
PROCNO                   1

F2 - Acquisition Parameters
Date_           20200929
Time              17.08 h
INSTRUM             spect
PROBHD   Z113652_0211 (
PULPROG             zg30
TD                 39578
SOLVENT             DMSO
NS                    16
DS                     2
SWH           10000.000 Hz
FIDRES            0.505331 Hz
AQ              1.9789000 sec
RG                 157.73
DW                50.000 usec
DE                 6.50 usec
TE                300.0 K
D1             2.00000000 sec
TD0                    1
SFO1          500.1330885 MHz
NUC1                  1H
P0                 3.53 usec
P1                10.58 usec
PLW1          20.00000000 W

F2 - Processing parameters
SI                 65536
SF            500.1300000 MHz
WDW                   EM
SSB       0
LB                0.30 Hz
GB        0
PC                1.00
```

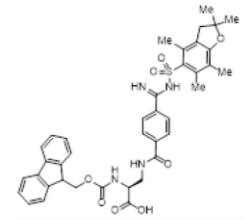
^{13}C NMR

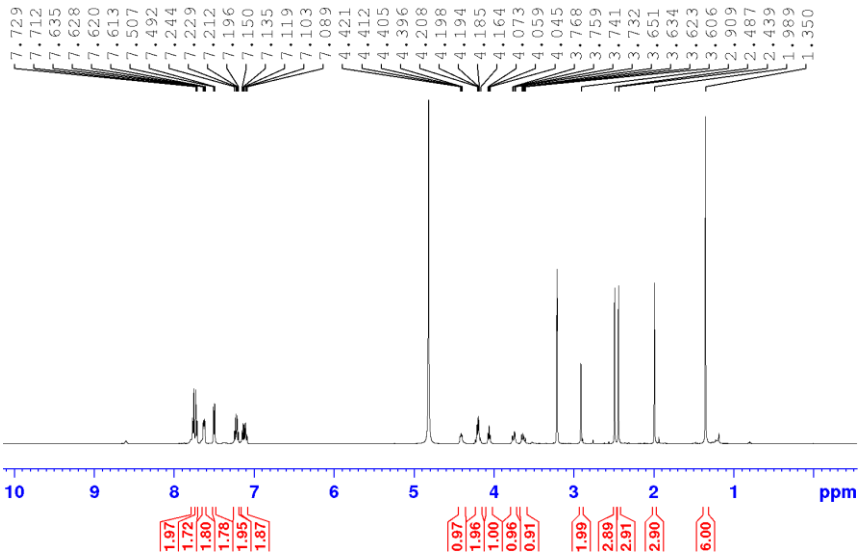





```
Current Data Parameters
NAME      JR0424-P1-E48536
EXPNO                    2
PROCNO                   1

F2 - Acquisition Parameters
Date_           20200929
Time              18.45 h
INSTRUM             spect
PROBHD   Z113652_0211 (
PULPROG           zgpg30
TD                 23806
SOLVENT             DMSO
NS                  2048
DS                     4
SWH           29761.904 Hz
FIDRES            2.500370 Hz
AQ              0.3999408 sec
RG                 198.22
DW                16.800 usec
DE                 6.50 usec
TE                300.0 K
D1             0.69999999 sec
D11            0.03000000 sec
TD0                    1
SFO1          125.7703637 MHz
NUC1                 13C
P0                 3.33 usec
P1                10.00 usec
PLW1          76.00000000 W
SFO2          500.1320005 MHz
NUC2                  1H
CPDPRG[2          waltz16
PCPD2             80.00 usec
PLW2          21.00000000 W
PLW12          0.31513000 W
PLW13          0.15851000 W

F2 - Processing parameters
SI                 32768
SF            125.7577885 MHz
WDW                   EM
SSB       0
LB                1.00 Hz
GB        0
PC                1.40
```

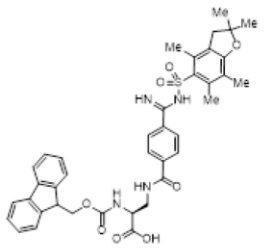
(*S*)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(*N*-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)carbamimidoyl)benzamido)propanoic acid (**4.87**) – ^{1}H NMR

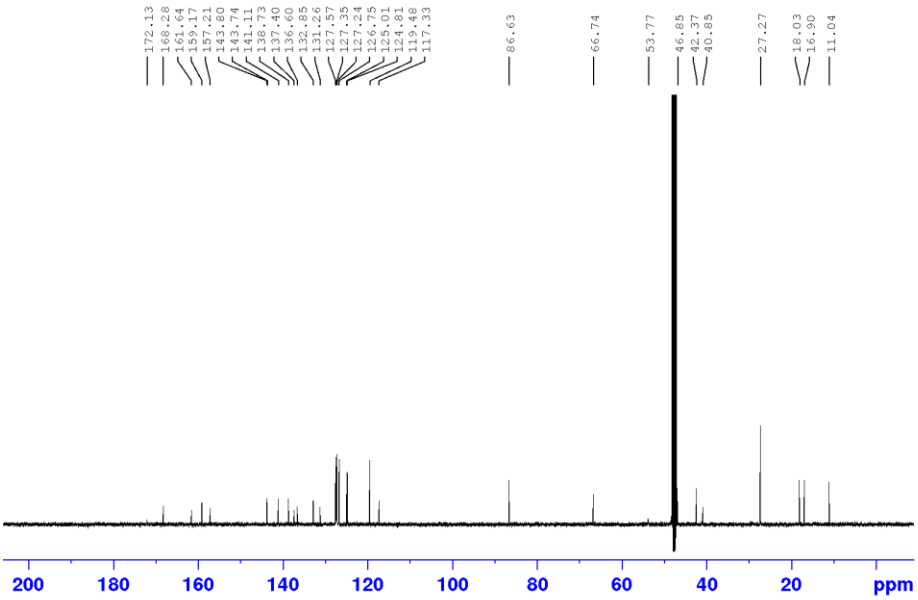





```
Current Data Parameters
NAME      JR0422-P1-E48812
EXPNO                    1
PROCNO                   1

F2 - Acquisition Parameters
Date_              20201009
Time                 16.54 h
INSTRUM              spect
PROBHD    Z113652_0211 (
PULPROG              zg30
TD                  39578
SOLVENT              MeOD
NS                     16
DS                      2
SWH             10000.000 Hz
FIDRES              0.505331 Hz
AQ                1.9789000 sec
RG                   157.73
DW                  50.000 usec
DE                   6.50 usec
TE                  300.0 K
D1               2.00000000 sec
TD0                     1
SFO1            500.1330885 MHz
NUC1                   1H
P0                   3.53 usec
P1                  10.58 usec
PLW1            20.00000000 W

F2 - Processing parameters
SI                  65536
SF              500.1300611 MHz
WDW                    EM
SSB          0
LB                   0.30 Hz
GB           0
PC                   1.00
```

^{13}C NMR

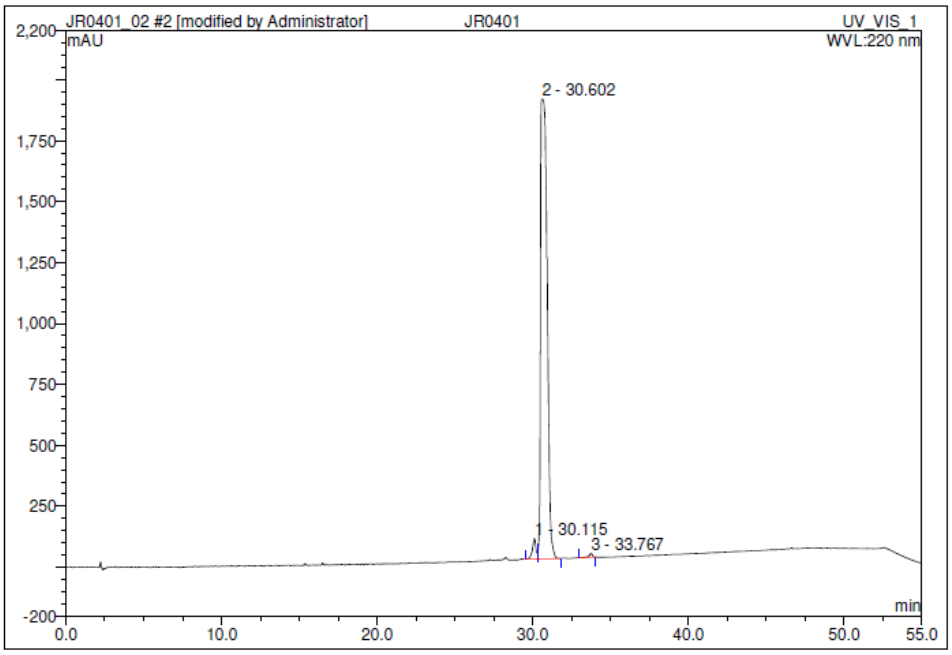





```
Current Data Parameters
NAME      JR0422-P1-E48812
EXPNO                    2
PROCNO                   1

F2 - Acquisition Parameters
Date_              20201009
Time                 21.46 h
INSTRUM              spect
PROBHD    Z113652_0211 (
PULPROG            zgpg30
TD                  23806
SOLVENT              MeOD
NS                   4096
DS                      4
SWH             29761.904 Hz
FIDRES              2.500370 Hz
AQ                0.3999408 sec
RG                   198.22
DW                  16.800 usec
DE                   6.50 usec
TE                  300.0 K
D1               2.00000000 sec
D11              0.03000000 sec
TD0                     1
SFO1            125.7703637 MHz
NUC1                   13C
P0                   3.33 usec
P1                  10.00 usec
PLW1            76.00000000 W
SFO2            500.1320005 MHz
NUC2                   1H
CPDPRG[2            waltz16
PCPD2                80.00 usec
PLW2            21.00000000 W
PLW12            0.31513000 W
PLW13            0.15851000 W

F2 - Processing parameters
SI                  32768
SF              125.7577885 MHz
WDW                    EM
SSB          0
LB                   1.00 Hz
GB           0
PC                   1.40
```

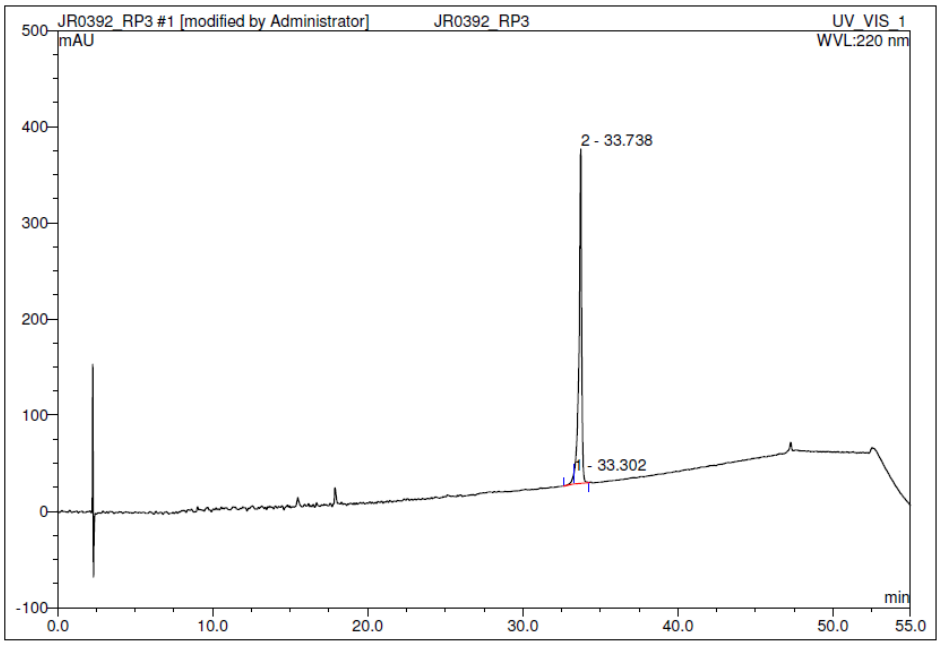
HPLC traces

TP-5



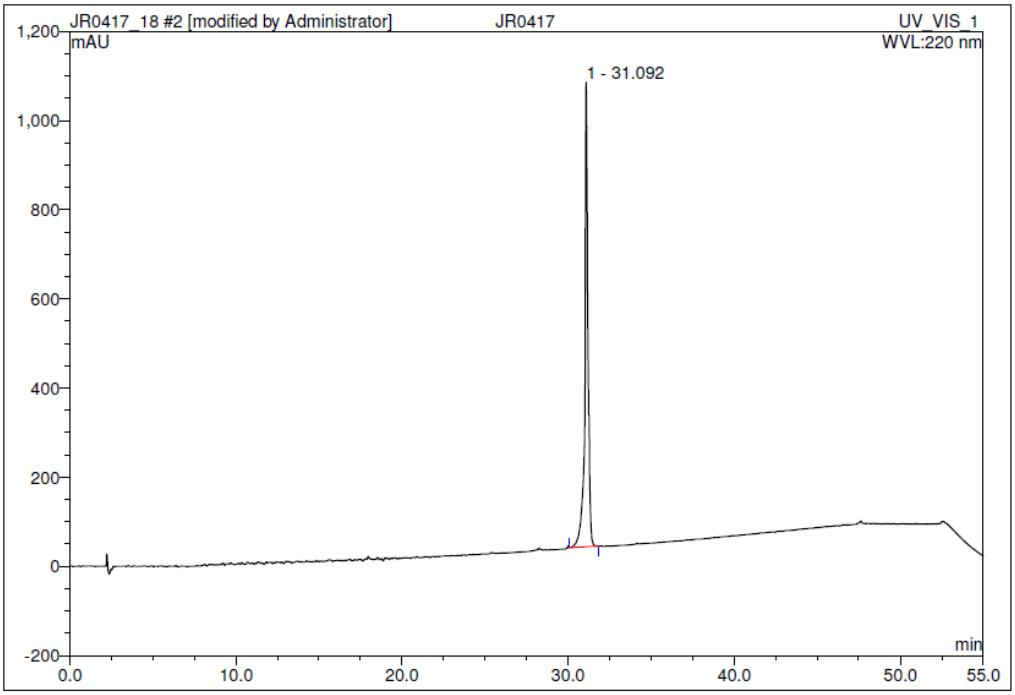
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	30.12	n.a.	84.407	22.538	2.38	n.a.	BM
2	30.60	n.a.	1887.515	919.439	97.20	n.a.	MB
3	33.77	n.a.	16.673	3.912	0.41	n.a.	BMB
Total:			1988.596	945.889	100.00	0.000	

TP-6



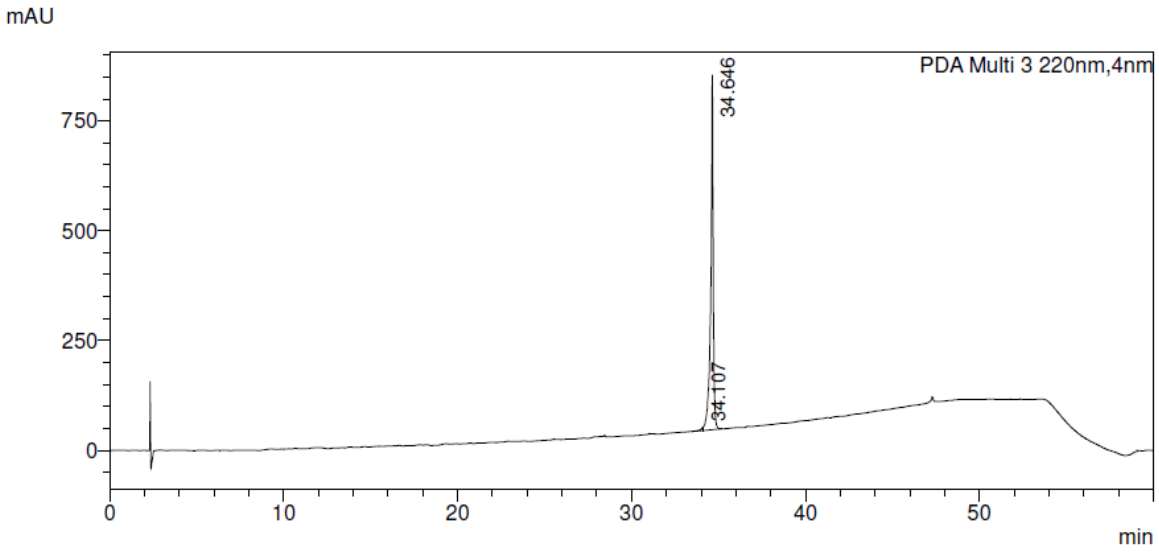
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	33.30	n.a.	10.626	1.423	2.09	n.a.	BM *
2	33.74	n.a.	347.483	66.587	97.91	n.a.	MB*
Total:			358.109	68.010	100.00	0.000	

TP-7



No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	31.09	n.a.	1042.011	230.254	100.00	n.a.	BMB
Total:			1042.011	230.254	100.00	0.000	

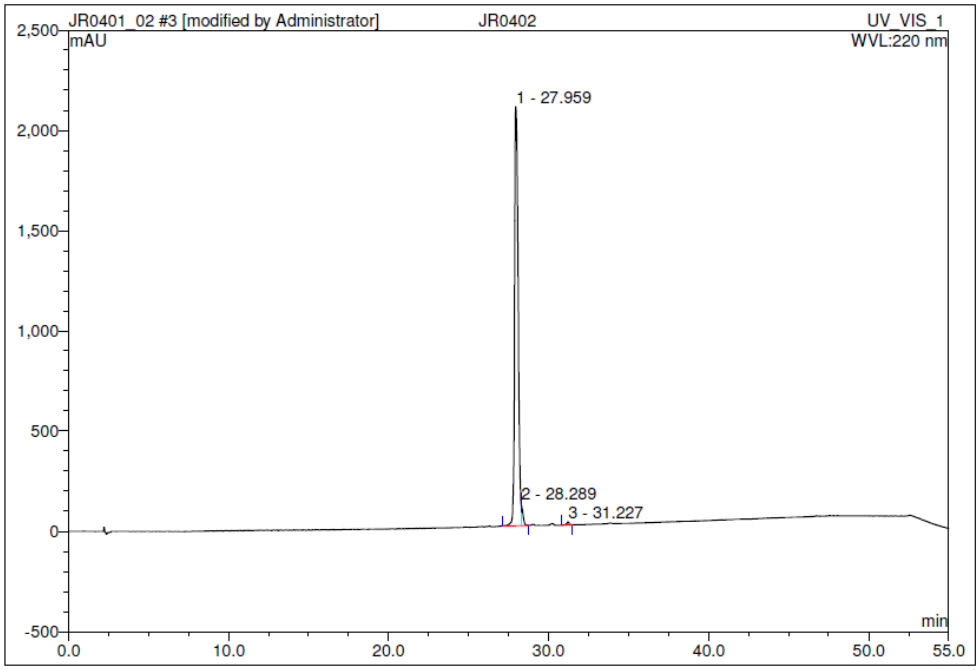
TP-8



PDA Ch3 220nm

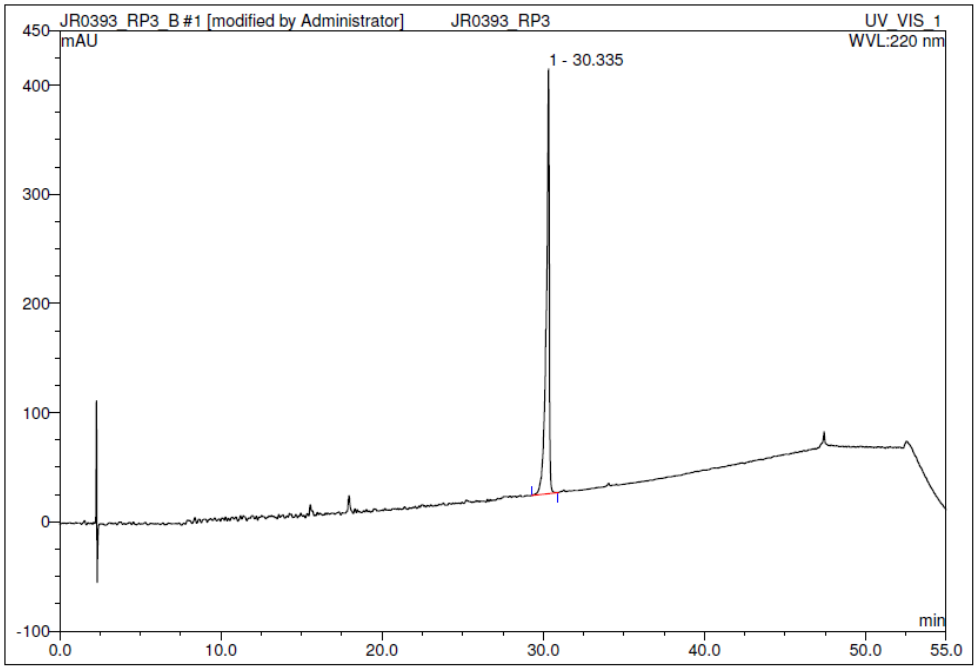
Peak#	Ret. Time	Area	Area%
1	34.107	58400	0.721
2	34.646	8038552	99.279
Total		8096953	100.000

TP-9



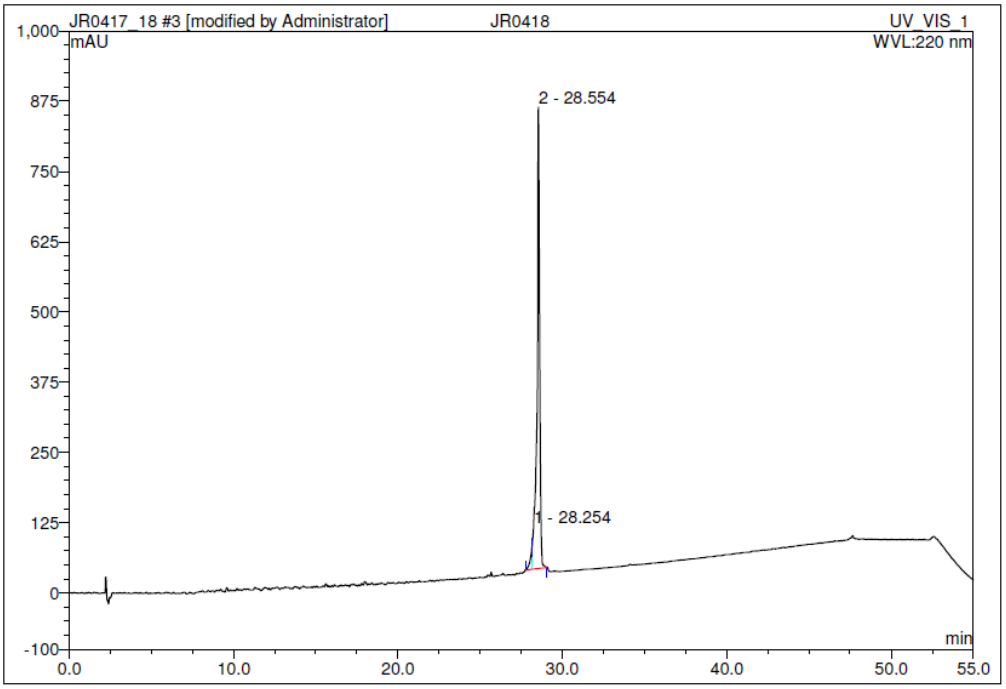
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	27.96	n.a.	2090.558	527.678	97.05	n.a.	BM *
2	28.29	n.a.	112.754	13.566	2.50	n.a.	MB*
3	31.23	n.a.	14.880	2.476	0.46	n.a.	BMB
Total:			2218.192	543.720	100.00	0.000	

TP-10



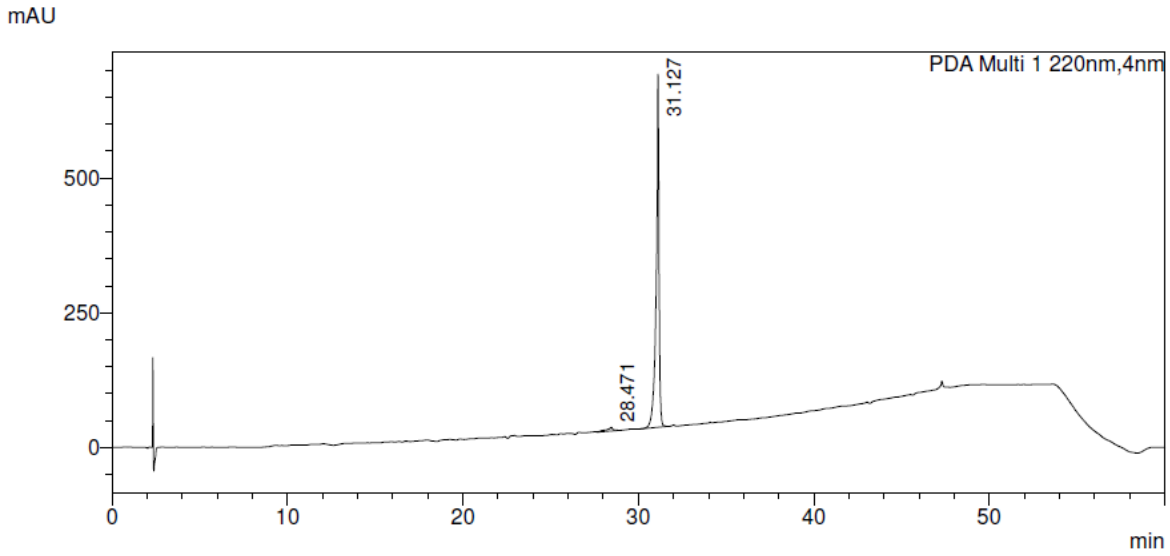
No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	30.34	n.a.	388.774	82.938	100.00	n.a.	BMB
Total:			388.774	82.938	100.00	0.000	

TP-11



No.	Ret.Time min	Peak Name	Height mAU	Area mAU*min	Rel.Area %	Amount	Type
1	28.25	n.a.	75.311	4.702	3.03	n.a.	BM *
2	28.55	n.a.	820.761	150.641	96.97	n.a.	MB*
Total:			896.072	155.343	100.00	0.000	

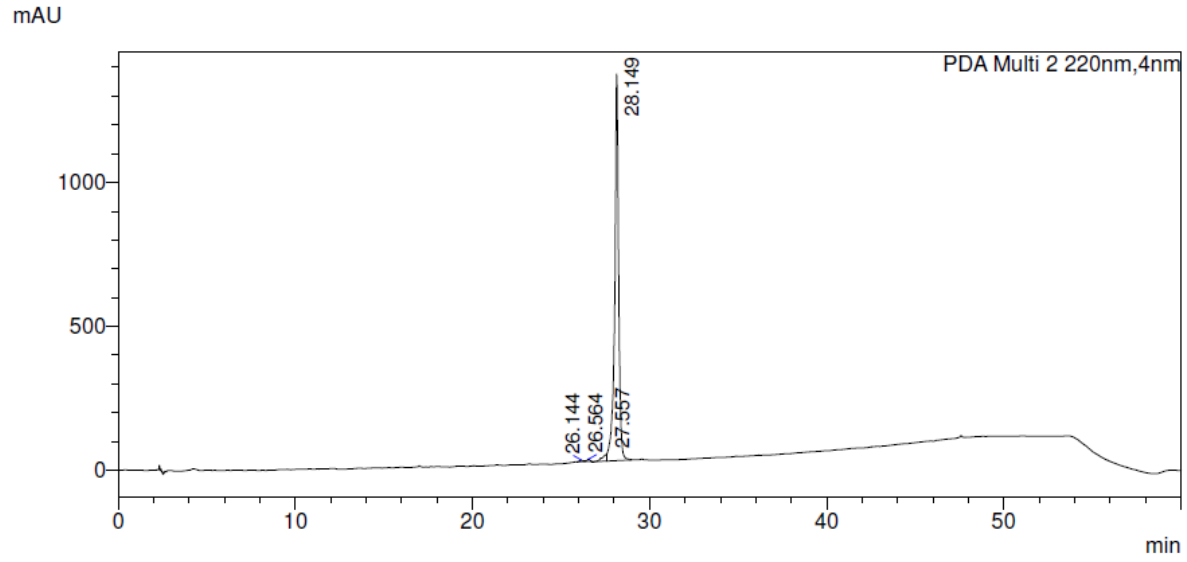
TP-12



PDA Ch1 220nm

Peak#	Ret. Time	Area	Area%
1	28.471	170812	2.314
2	31.127	7212014	97.686
Total		7382826	100.000

TP-4

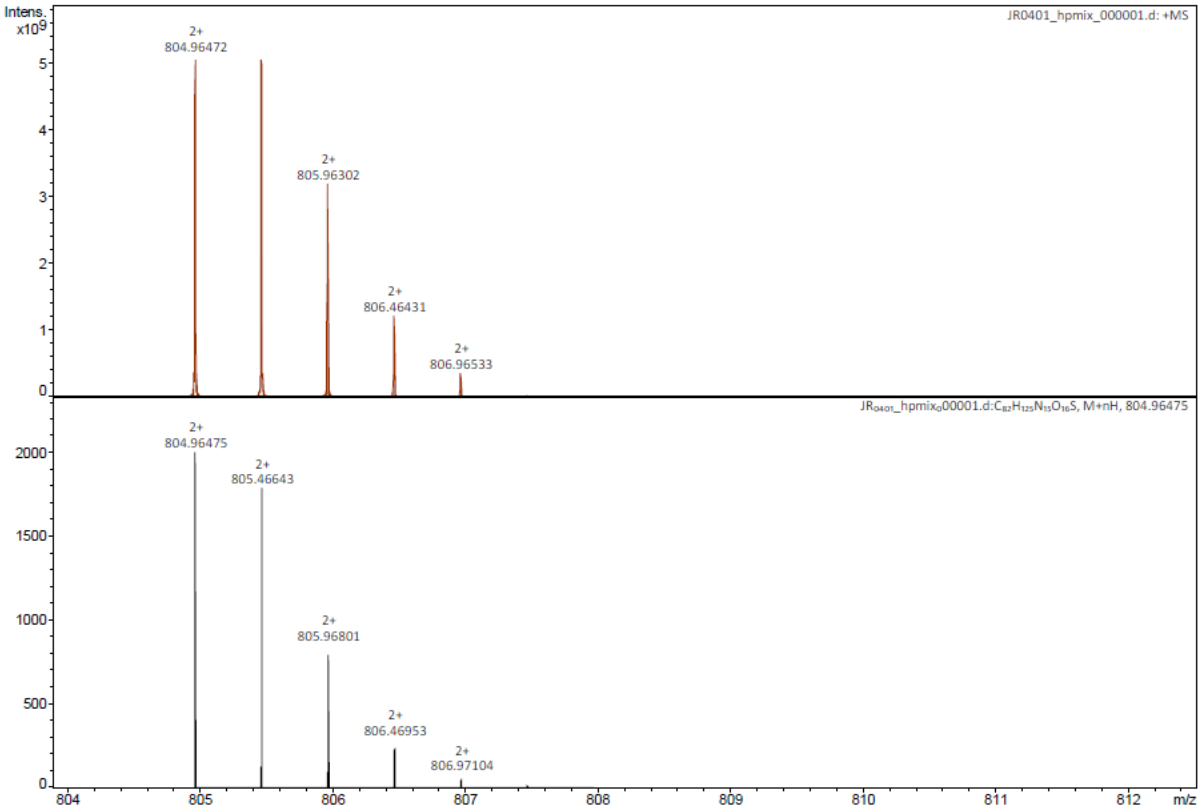


PDA Ch2 220nm

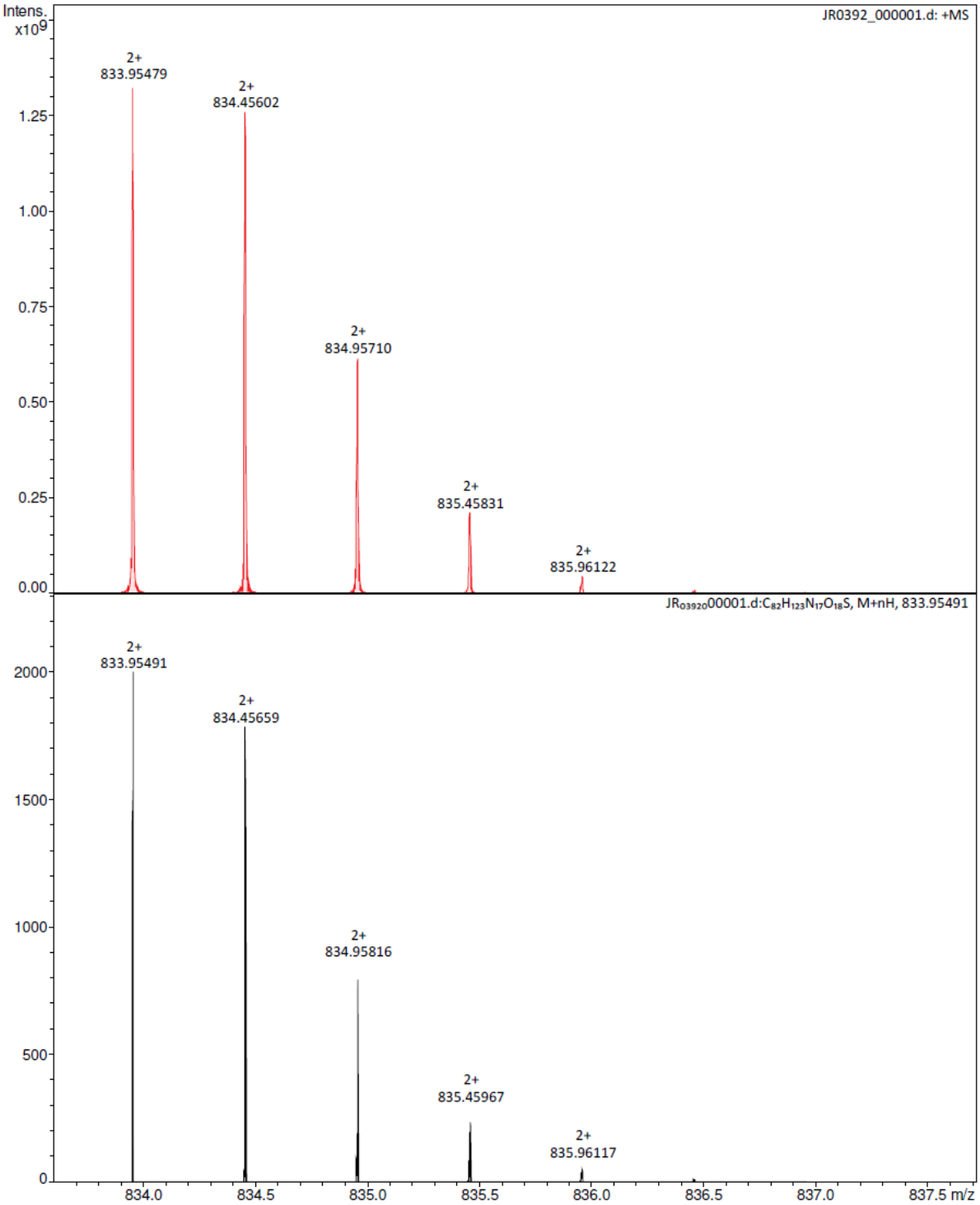
Peak#	Ret. Time	Area	Area%
1	26.144	52118	0.244
2	26.564	92292	0.432
3	27.557	379693	1.777
4	28.149	20843234	97.547
Total		21367338	100.000

High resolution mass spectra

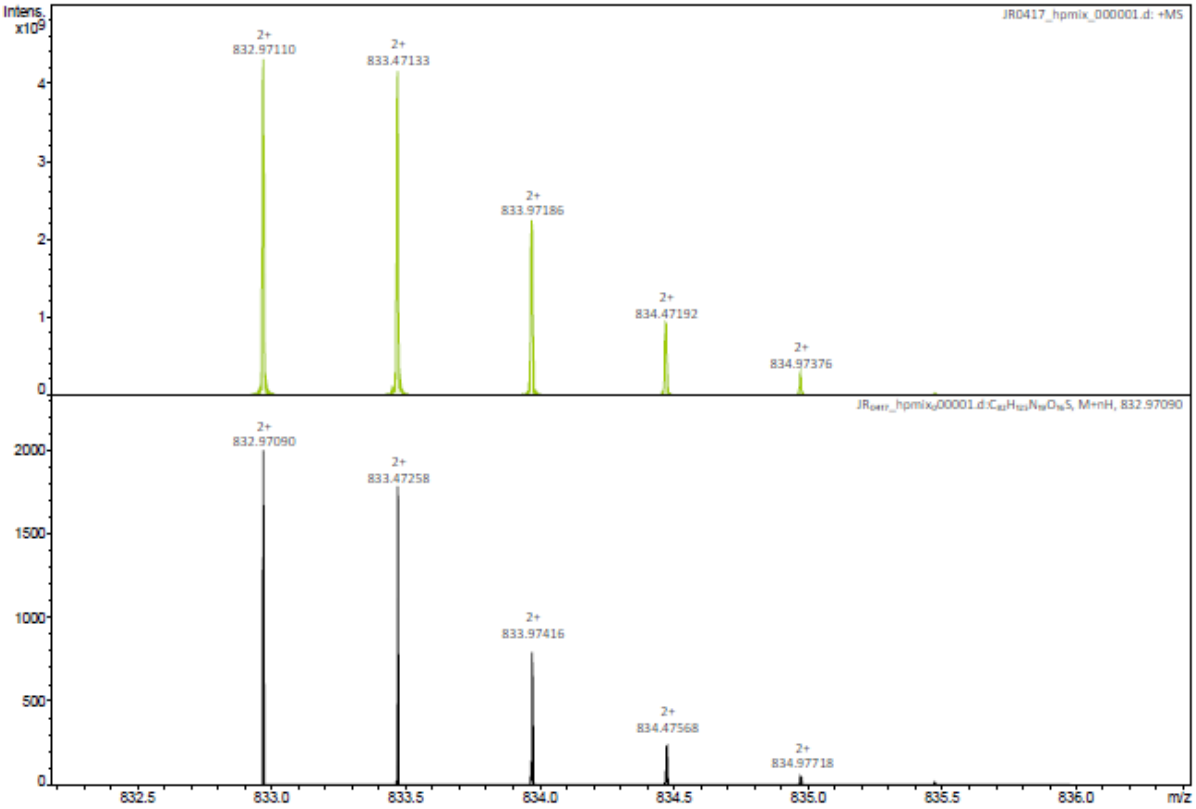
TP-5



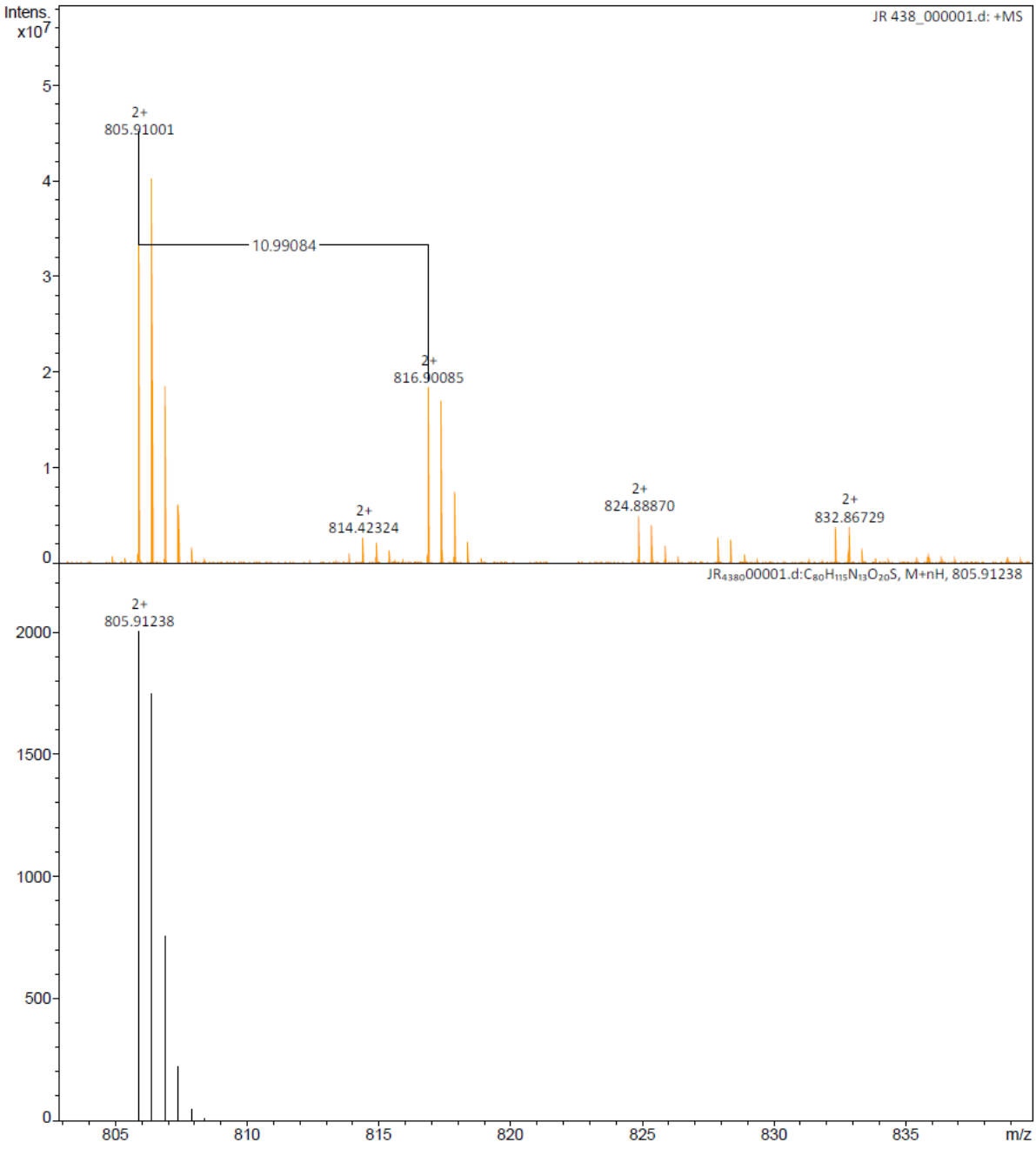
TP-6


Intens.
x10^9
JR0392_000001.d: +MS
2+
833.95479
2+
834.45602
2+
834.95710
2+
835.45831
2+
835.96122
1.25
1.00
0.75
0.50
0.25
0.00
JR03920 00001.d:C82H123N17O18S, M+nH, 833.95491
2+
833.95491
2+
834.45659
2+
834.95816
2+
835.45967
2+
835.96117
2000
1500
1000
500
0
834.0
834.5
835.0
835.5
836.0
836.5
837.0
837.5 m/z

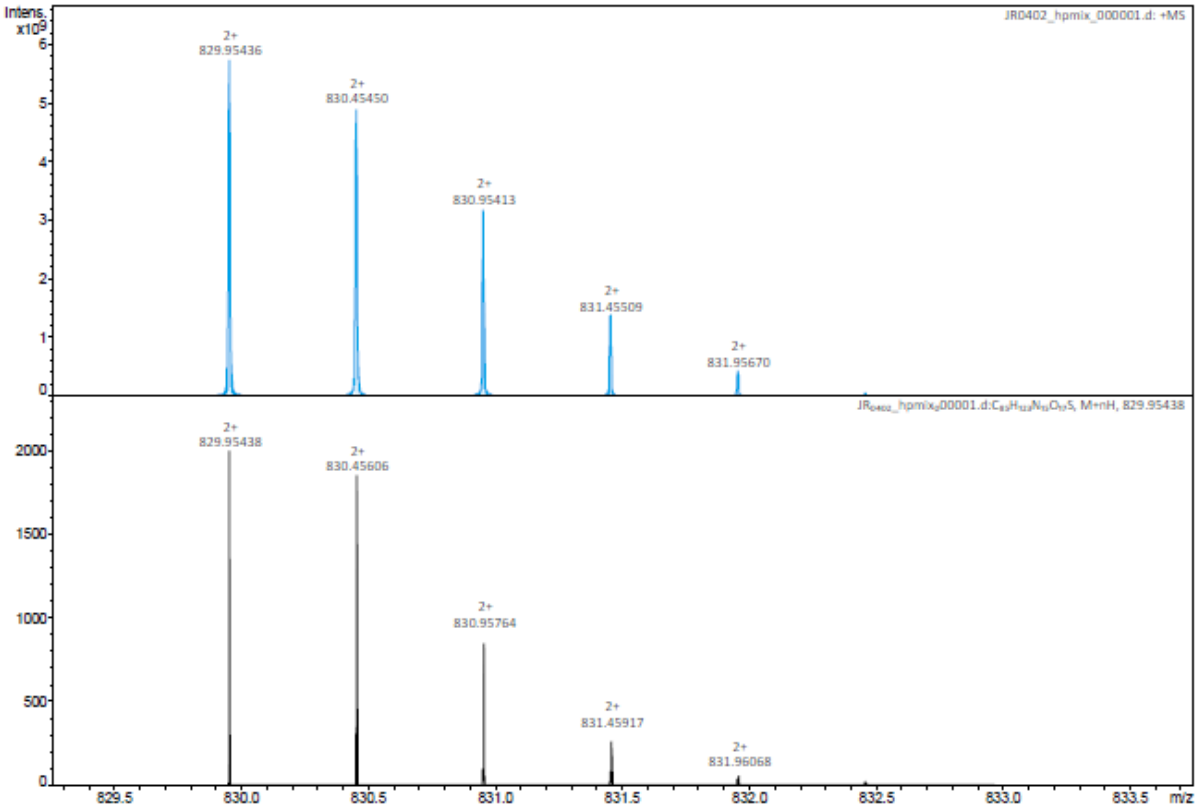
TP-7


Intens.
x10^9
JR0417_hpmix_000001.d: +MS
2+
832.97110
2+
833.47133
2+
833.97186
2+
834.47192
2+
834.97376
JR0417_hpmix_000001.d:C82H103N19O16S, M+nH, 832.97090
2+
832.97090
2+
833.47258
2+
833.97416
2+
834.47568
2+
834.97718
m/z

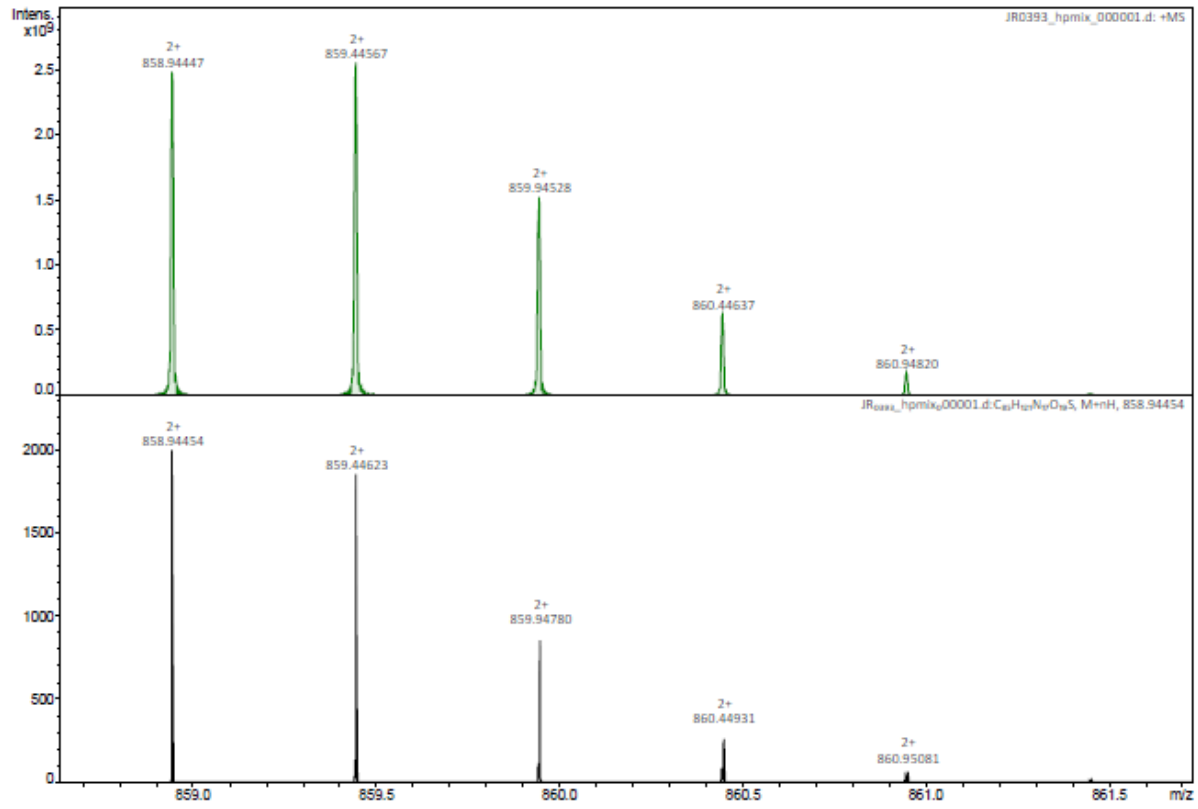
TP-8


Intens.
x10^7
JR 438_000001.d: +MS
2+
805.91001
10.99084
2+
816.90085
2+
814.42324
2+
824.88870
2+
832.86729
JR4380 00001.d:C80H115N13O20S, M+nH, 805.91238
2+
805.91238
m/z

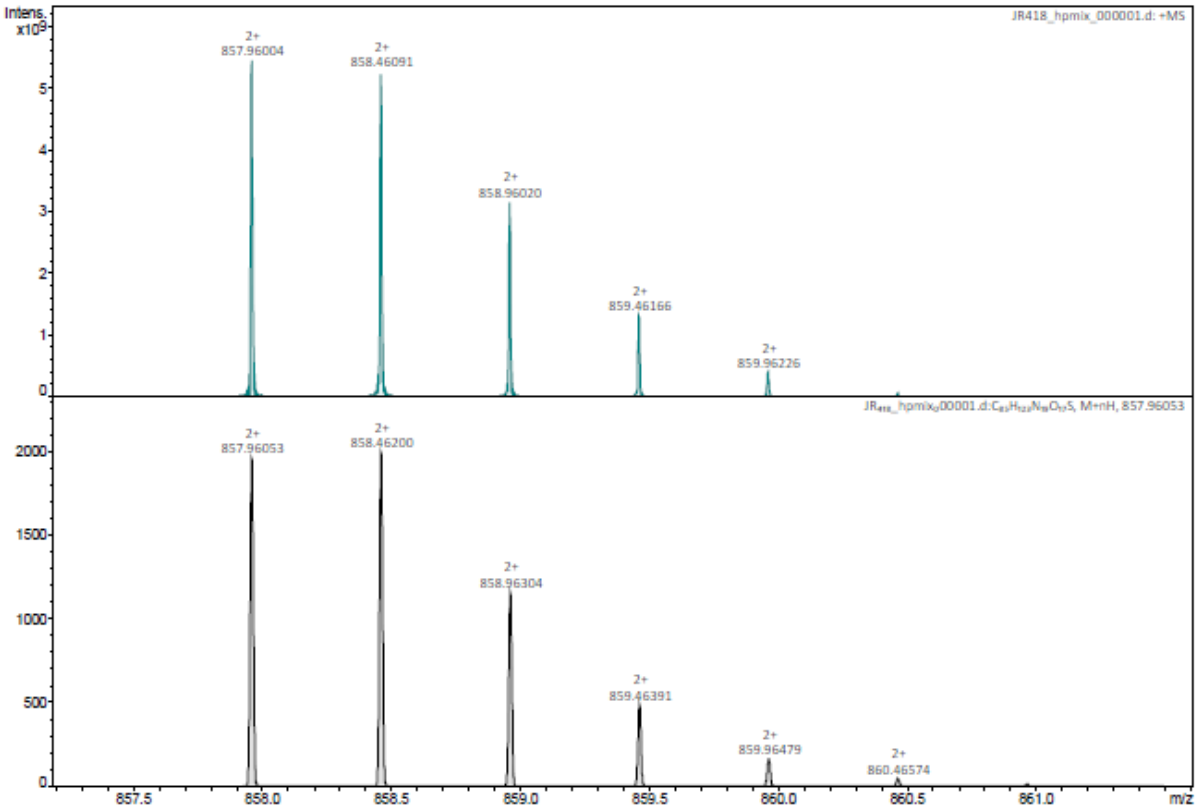
TP-9


Intens.
x10^9
JR0402_hpmix_000001.d: +MS
2+
829.95436
2+
830.45450
2+
830.95413
2+
831.45509
2+
831.95670
JR0402_hpmix_00001.d:C83H132N18O17S, M+nH, 829.95438
2+
829.95438
2+
830.45606
2+
830.95764
2+
831.45917
2+
831.96068
m/z

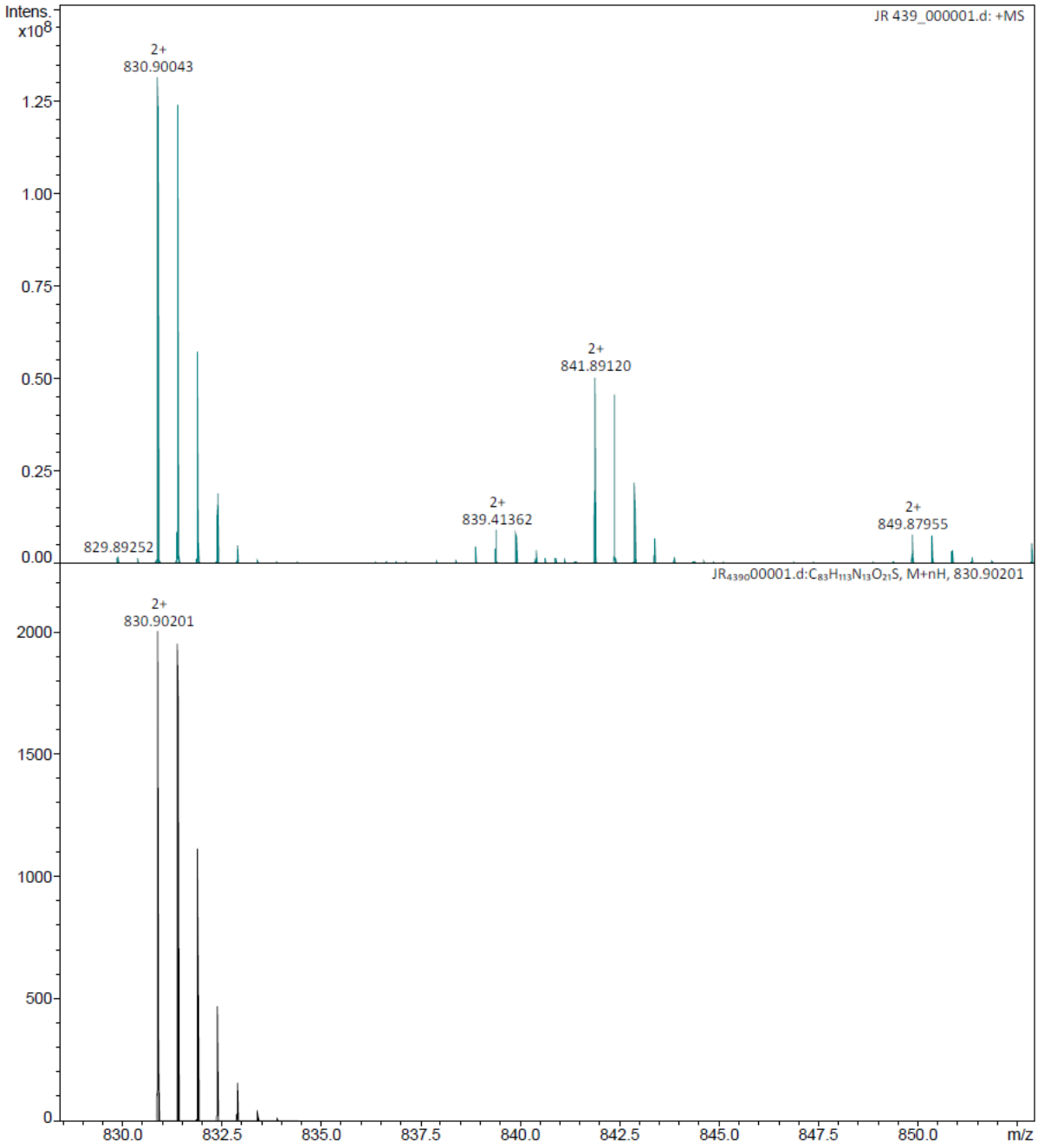
TP-10


Intens.
x10^9
JR0393_hpmix_000001.d: +MS
2+
858.94447
2+
859.44567
2+
859.94528
2+
860.44637
2+
860.94820
JR0393_hpmix_00001.d:C82H132N18O18S, M+nH, 858.94454
2+
858.94454
2+
859.44623
2+
859.94780
2+
860.44931
2+
860.95081
m/z

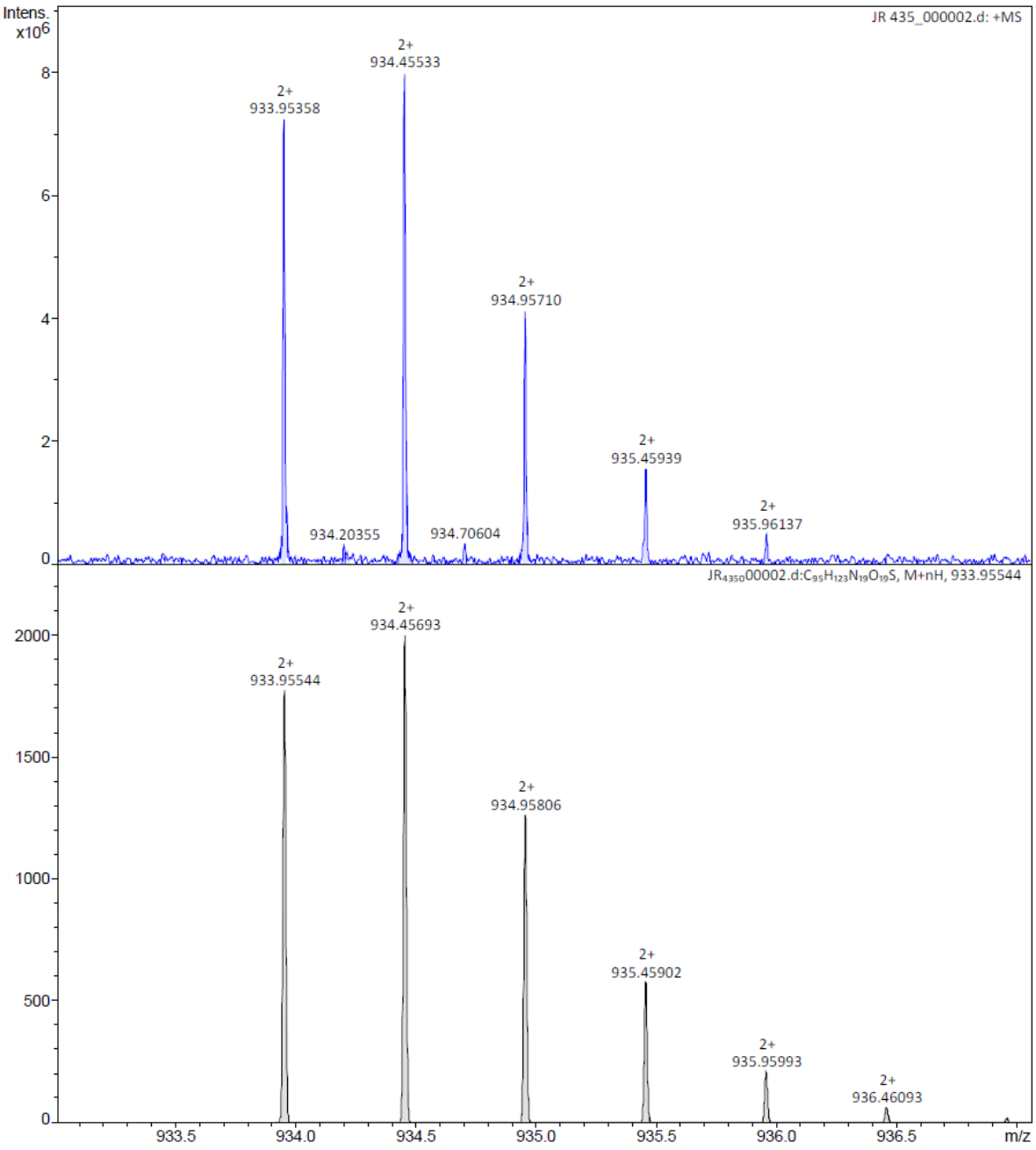
TP-11


Intens.
x10^9
JR418_hpmix_000001.d: +MS
2+
857.96004
2+
858.46091
2+
858.96020
2+
859.46166
2+
859.96226
JR418_hpmix_00001.d:C83H122N18O23S, M+nH, 857.96053
2+
857.96053
2+
858.46200
2+
858.96304
2+
859.46391
2+
859.96479
2+
860.46574
m/z

TP-12

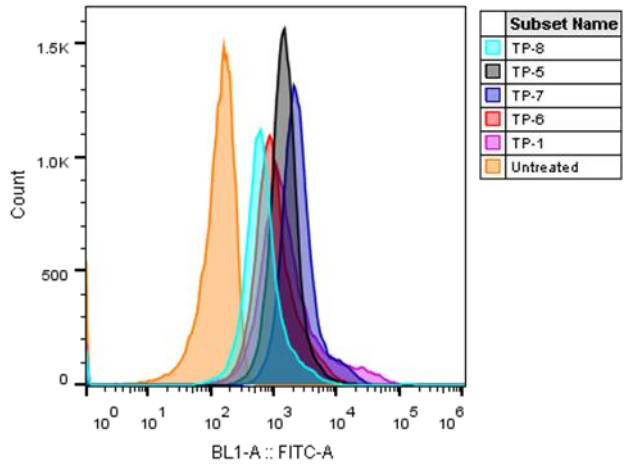

Intens.
x10^8
JR 439_000001.d: +MS
2+
830.90043
1.25
1.00
0.75
0.50
0.25
0.00
2+
841.89120
2+
839.41362
2+
849.87955
829.89252
JR4390 00001.d:C83H113N13O21S, M+nH, 830.90201
2+
830.90201
2000
1500
1000
500
0
830.0
832.5
835.0
837.5
840.0
842.5
845.0
847.5
850.0
m/z

TP-4

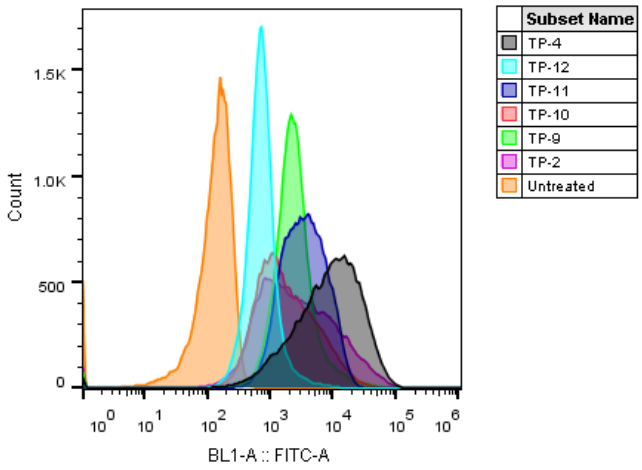

Intens.
x10^6
JR 435_000002.d: +MS
2+
933.95358
2+
934.45533
2+
934.95710
2+
935.45939
2+
935.96137
934.20355
934.70604
JR4350 00002.d:C95H123N19O19S, M+nH, 933.95544
2+
933.95544
2+
934.45693
2+
934.95806
2+
935.45902
2+
935.95993
2+
936.46093
m/z

Flow cytometry

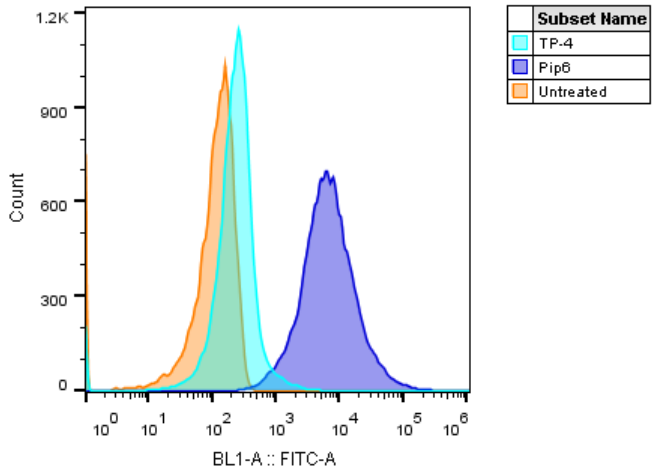
TP-1 analogues - 4 h incubation at 37 °C



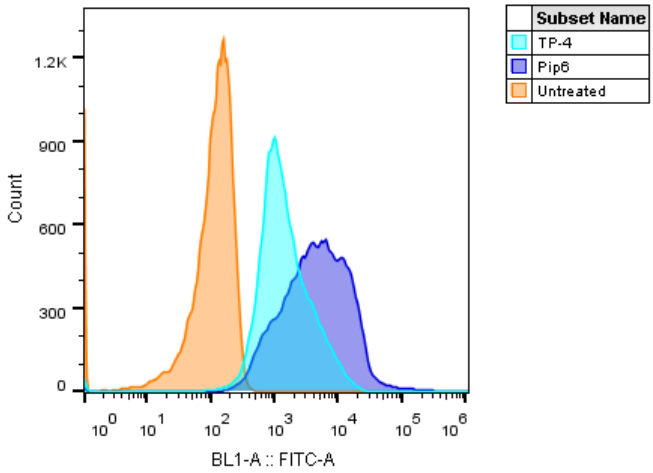
TP-2 analogues - 4 h incubation at 37 °C



TP-4 and **Pip6** - 4 h incubation at 4 °C

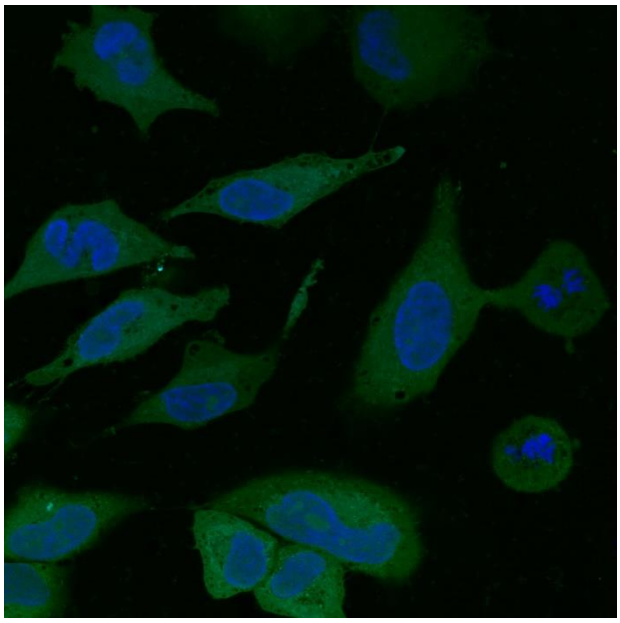


TP-4 and **Pip6 -** 1 h incubation at 37 °C



Microscope images

TP-11



TP-4