Development of peptide-based ^{19}F MRI agents and BPTF-bromodomain inhibitors

A DISSERTATION

SUBMITTED TO THE FACULTY OF

UNIVERSITY OF MINNESOTA

BY

Steven E. Kirberger

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

ADVISOR:

Dr. William C. K. Pomerantz

July 2019

Acknowledgements

I would like to thank all of my family, friends, and collaborators for helping and supporting me throughout my graduate career. In particular:

Will Pomerantz – For guidance, and helping me develop into a professional scientist

Jorden Johnson – For endless support and encouragement at every turn

Peter Ycas – For teamwork and a sounding board for ideas

Andrew Urick – For teaching and advice

Abstract

Molecular imaging is the process of using targeted probes to detect abnormalities at the molecular level by observing interactions to specific biomarkers. Magnetic resonance imaging (MRI) presents an interesting avenue with respect to development of probes for the early detection of disease. In particular, ^{19}F MRI shows promise for this development as the fluorine nucleus possesses many similar characteristics as the conventionally used proton but has the distinction of a lack of background signal found natively in biological systems. State of the art ^{19}F MRI agents involve the use of perfluorinated compounds that often suffer from stability issues, bioaccumulation, as well as persistence within the environment. The first part of this dissertation describes the design and optimization of a peptide-based ^{19}F MRI agent. These peptide scaffolds show promise for future use as ^{19}F MRI probes due to their high signal, water solubility, and facile degradation in vivo to prevent bioaccumulation. The resultant byproducts have also been shown to be environmentally benign. This work is the focus of Chapters 2 and 3 of this document.

In a second project, the development of a small molecule inhibitor of an epigenetic protein target is described. AU1, the first reported small molecule inhibitor of the bromodomain of a protein called BPTF, was discovered in the Pomerantz lab in 2015. As BPTF is a relatively understudied protein, there exists a need to improve the potency of AU1 as a probe for the various functions of its bromodomain, as it has been implicated in numerous diseases including: pancreatic cancer, melanoma, colorectal cancer, hepatocellular carcinoma, breast cancer, bladder cancer, and lung cancer. Structural analogs have been developed and analyzed in an attempt to improve upon AU1 in terms of its potency, solubility, and reduction of potential off-target binding. The work described in Chapter 4 of this document shows progress toward these goals, and the development of our ^{19}F NMR assays for the analysis of protein ligands. A collaborative effort involving the (*S*)-enantiomer of AU1 is briefly described in Chapter 1.

In a third section, the collaborative work between our lab and that of Ratmir Derda to develop peptide auxiliaries to improve therapeutic life-time in vivo is described. Chapter

5 details the use of ^{19}F NMR to analyze the binding strength and location of numerous fluorinated peptides designed in the Derda lab.

Table of Contents

Acknowledgements i

Abstract ii

List of Figures vii

List of Tables x

Chapter 1: Introduction 1

1.1 ^{19}F MRI in biomedicine 1

1.2 Molecular imaging 2

1.3 Magnetic resonance imaging 3

1.3.1 ^{19}F MRI 4

1.4 Perfluorocarbon compounds for ^{19}F MRI 5

1.5 ^{19}F MRI agents based on non-PFC scaffolds 7

1.6 Solid-phase peptide synthesis 8

1.7 The “Disordered Symmetry” Hypothesis 10

1.8 Epidermal Growth Factor Receptor (EGFR) as a Breast Cancer Molecular Imaging Target. 12

1.9 Peptides and proteins as EGFR selective targeting agents 14

1.10 Epigenetics 17

1.11 Bromodomains 21

1.12 Bromodomain and Planthomeo domain containing Transcription Factor (BPTF) 24

1.13 Protein-Observed Fluorine NMR (PrOF NMR) 25

1.14 Small molecules as protein-protein interaction inhibitors 28

1.15 AU1 discovery as first BPTF bromodomain inhibitor 29

1.16 Preface to this Dissertation 31

Chapter 2: Synthesis of Intrinsically Disordered Fluorinated Peptides for Modular Design of High-Signal ^{19}F MRI Agents

..... 33

2.1 Introduction 33

2.2 Results and Discussion 35

2.3 Conclusion 48

2.4 Experimental 49

Chapter 3: Toward a 2nd-generation peptide-based ^{19}F MRI probe

..... 76

3.1 Introduction 76

3.2 Alternative ^{19}F-labeled amino acids 77

3.2.1 Trifluoromethyl methionine (**4A**) 78

3.2.2 Perfluorotertbutyl alkoxynorvaline (**3B**) 81

3.2.3 Bis(trifluoromethyl) benzyl lysine (**3C**) 83

3.3 Development of biomarker selective targeting agents 84

3.3.1 Fibronectin 85

3.3.2 EGFR-targeting peptides 87

3.4 Examining the charge of ^{19}F imaging peptides 89

3.4.1 Determining cell-surface adhesion 90

3.4.2 Reduction in peptide charge 91

3.5 Conclusions and future outlook 97

3.6 Experimental 98

Chapter 4: Selectivity, ligand deconstruction, and cellular activity analysis of a BPTF bromodomain inhibitor 139

4.1 Introduction 139

4.2 PrOF NMR identification of the active enantiomer of ***rac-1*** 141

4.3 X-ray crystal structure of 5FW BPTF supports a lack of structure perturbation from fluorine for PrOF NMR experiments. 142

4.4 **(*S*)-1** Structure-Activity Relationships. 144

4.5 Additional SAR 151

4.6 Ligand deconstruction of **(*S*)-1** by PrOF NMR 152

4.7 Molecular dynamic simulations of **(*S*)-1** with BPTF 153

4.8 Bromodomain selectivity analysis of **(*S*)-1** 155

4.9 Kinase selectivity analysis of ***rac*-1** stereoisomers 157

4.10 Cancer cell lines sensitivity to genetic altering of *BPTF* gene and ***rac*-1** stereoisomers 158

4.11 Conclusions 161

4.12 Experimental 162

Chapter 5: ^{19}F NMR analysis of Human Serum Albumin-binding peptides 308

5.1 Introduction 308

5.2 Binding affinity ranking of PFS peptides 310

5.3 Competitive binding analysis with known HSA-binding drugs 314

5.4 Comparison between HSA and rat serum albumin (RSA) 316

5.5 Conclusions and future outlook 317

5.6 Experimental 318

References 319

List of Figures

Figure 1-1: General outline for molecular imaging 2
Figure 1-2*:* Poisson-Boltzmann electrostatic map of three perfluorinated compounds 7
Figure 1-3: Select ^{19}F MRI agents based on non-PFC scaffolds 8
Figure 1-4: General schematic for Fmoc solid phase peptide synthesis 9
Figure 1-5: ^{19}F NMR of unfolded IFABP 11
Figure 1-6: Diagram of receptor tyrosine kinase activity 12
Figure 1-7: EGFR and select biological pathways leading to tumorigenesis 13
Figure 1-8: Selected EGFR-targeting peptides 15
Figure 1-9: NMR structure of Fibronectin scaffold with a ^{64}Cu chelating DOTA tag 16
Figure 1-10: General structure of heterochromatin and euchromatin. 18
Figure 1-11: Select post translational modifications 20
Figure 1-12: Structure of the bromodomain of BPTF 21
Figure 1-13: Structures of bromodomain inhibitors 23
Figure 1-14: Phylogenetic tree of bromodomains 23
Figure 1-15: Binding site of the bromodomain of BPTF 24
Figure 1-16: Example PrOF NMR titration spectra 26
Figure 1-17: ^{19}F NMR of 5FW BRD4 bound by (+)-JQ1 28
Figure 1-18: Dual screen of 5FW BPTF and 5FW BRD4 to identify BPTF inhibitor…..30
Figure 1-19: Structure of AU1 and lucierfase assay 31
Figure 2-1: Disordered ^{19}F peptide design constructed by solid-phase peptide synthesis and current state-of-the-art molecules 34
Figure 2-2: Far-UV CD spectra and ^{19}F NMR spectra of fluorinated peptide 36
Figure 2-3: ^{1}H NMR of backbone amide doublets in peptide **1** 37
Figure 2-4: α-helical representation of **1** dissolved in methanol 38
Figure 2-5: Structures of ^{19}F-labeled peptides 39
Figure 2-6: Tryptic digest of **2a** 41
Figure 2-7: Tryptic digest of **2b** 42
Figure 2-8: Tryptic digest of **2c** 42
Figure 2-9: Degradation study of Peptide **2b** subjected to buffer only, pH = 7.1 43
Figure 2-10: Degradation study of Peptide **2b** subjected to buffer only, pH = 8.1 44

Figure 2-11: 19F NMR spectrum of **2a** in E. coli cellular lysate 45
Figure 2-12: ^{19}F MRI phantom image of **3** at various concentrations 46
Figure 2-13: ^{19}F NMR of BSA conjugated to **3** 47
Figure 3-1: Structures of fluorinated amino acids 78
Figure 3-2: Newman projection of a CF_3 moiety attached to a sulfur or carbon center ... 79
Figure 3-3: Synthetic scheme of **3A** 79
Figure 3-4: Sequence of **3-1** and resulting ^{19}F NMR spectrum 80
Figure 3-5: Synthesis of **3B** 81
Figure 3-6: Sequence of **3-2** and corresponding ^{19}F NMR spectrum 82
Figure 3-7: Synthesis of **3C** 83
Figure 3-8: Sequence of **3-3** and its ^{19}F NMR spectrum 84
Figure 3-9: Fibronectin scaffold 85
Figure 3-10: Bioconjugation pair between **3-5** and fibronectin 86
Figure 3-11: Conjugation of EGFR-targeting and ^{19}F imaging peptides 87
Figure 3-12: Flow cytometry analysis of **3-4b** with A431 and SKOV-3 cells 88
Figure 3-13: Flow cytometry analysis of **3-6** with A431 and SKOV-3 cells 89
Figure 3-14: Sequences of **3-7** and **3-8**, used for confocal microscopy experiments 90
Figure 3-15: Confocal microscopy images of HeLa cells with ^{19}F-labeled peptide 91
Figure 3-16: ^{19}F NMR spectra of a series of peptides with reduced lysine content 92
Figure 3-17: Flow cytometry analysis of **3-7** and **3-11** incubated with A431 cells 93
Figure 3-18: Flow cytometry analysis of **3-12** and **3-13** incubated with HCC-38 cells 95
Figure 3-19: Flow cytometry analysis of **3-12** and **3-13** incubated with HEK-293 cells . 95
Figure 3-20: ^{19}F NMR overlay of peptides containing **3C** with lysine replacements 97
Figure 4-1: X-ray crystal structure of 5FW BPTF 140
Figure 4-2: PrOF NMR spectra of AU1 enantiomers 142
Figure 4-3: Electron density and models of W residue of BPTF 143
Figure 4-4: General synthetic scheme for ***rac*-1** 144
Figure 4-5: Pharmacokinetic study of **(*S*)-1** 145
Figure 4-6: Structures of select AU1 analogs 146
Figure 4-7: Sample PrOF NMR titration data of **9** with 5FW BPTF 147
Figure 4-8: PrOF NMR titration of **9** with 5FW BRD4(BD1) 148

Figure 4-9: PrOF NMR titration of **14** with 5FW BPTF…….. 150

Figure 4-10: Additional **(*S*)-1** analogs 151

Figure 4-11: Ligand deconstruction study of **(*S*)**-1..... 153

Figure 4-12: Molecular dynamics simulation of **(*S*)-1** with 5FW BPTF..... 155

Figure 4-13: Selectivity of **(*S*)-1** toward four different bromodomains 156

Figure 4-14: Sensitivity of cancer cell lines to **(*S*)-1** and **(*R*)-1**..... 160

Figure 5-1: Structures of perfluoro biphenylsulfide (PFS) peptides..... 309

Figure 5-2: Structures of small molecule binders of human serum albumin..... 310

Figure 5-3: ^{19}F NMR titration of **RD-1** with increasing concentrations of HSA 311

Figure 5-4: ^{19}F NMR titration of **RD-2** with increasing concentrations of HSA 312

Figure 5-5: ^{19}F NMR titration of **RD-3** with increasing concentrations of HSA 312

Figure 5-6: ^{19}F NMR titration of **RD-4** with increasing concentrations of HSA 313

Figure 5-7: ^{19}F NMR titration of **RD-5** with increasing concentrations of HSA 313

Figure 5-8: ^{19}F NMR competition experiment with **RD-1**, HSA, and ibuprofen..... 314

Figure 5-9: ^{19}F NMR competition experiment with **RD-1**, HSA, and diclofenac-Na.... 315

Figure 5-10: ^{19}F NMR competition experiment with **RD-1**, HSA, and carbamazepine. 315

Figure 5-11: ^{19}F NMR comparison of **RD-1** with RSA and HSA..... 316

Figure 5-12: ^{19}F NMR comparison of **RD-5** with RSA and HSA..... 317

List of Tables

Table 2-1: Lifetime of model peptide series **2** during tryptic digest..... 41
Table 2-2: MALDI-TOF MS data for ^{19}F-labeled peptides..... 74
Table 2-3: HPLC retention time for peptide series **S1**..... 75
Table 2-4: HPLC retention time for peptide series **S2**..... 75
Table 3-1: Mass spectrometry data for synthesized peptides 138
Table 4-1: Binding affinities of **(*S*)-1** and anlogs against known kinase targets..... 157
Table 4-2: MS of fully fluorinated proteins..... 164
Table 4-3. Unlabeled BPTF Data collection and refinement statistics..... 301
Table 4-4: 5FW BPTF Data collection and refinement statistics 303
Table 4-5: RMSD values between Unlabeled BPTF (moving) and 5FW BPTF (Reference)..... 306
Table 5-1: Components included in ^{19}F NMR assays..... 318

Chapter 1

Introduction

Part I: ^{19}F MRI probes as molecular imaging agents

1.1 ^{19}F MRI in biomedicine

Magnetic resonance imaging (MRI) is a non-invasive method to detect disease or abnormalities in a patient based on the magnetic properties of the nuclei in the body, most often the proton. Despite being developed shortly after the first human scan using ^{1}H detection, ^{19}F MRI is not actively used in the clinic. Initial studies occurred in 1977 involved the imaging of sodium fluoride (NaF) and perfluorotributylamine (PFTBA), showing the capability of detection on the ^{19}F nucleus.[1] Advances in technology, data interpretation, and contrast agents have drastically improved the use of ^{1}H MRI such that it is widely used in the clinic, however there remains an avenue of research to explore regarding development of ^{19}F MRI agents. Traditional ^{1}H MRI scans typically involve the delivery of a contrast agent to invoke changes in the magnetic properties of protons (often from water) that these contrast agents encounter. Thus, detection on the ^{1}H nucleus is based largely on differences in intensity of the same nucleus, which can be complicated by excess noise, or poor resolution. Fluorine is 83% as sensitive compared to proton in the context of magnetic resonance, in addition to being 100% naturally abundant as the MRI-active ^{19}F isotope. Furthermore, there is negligible ^{19}F signal found naturally within the body, and the fluorine that is present in the bones and teeth possesses very short T_2 relaxation times, making them MRI silent. Signal sensitivity is however a challenge in the development of ^{19}F MRI agents, as the limit of detection at clinical field strength (1.5 T) and a 6 min scan time is 30 mM.[2,3]

The state of the art for ^{19}F agents involve the use of perfluorocarbon (PFC) emulsions which yield high ^{19}F-signal, but tend to be difficult to use or have a tendency to accumulate in the body and environment. The work described in this thesis involves my contribution to the development of a new class of ^{19}F MRI agent, based on a water-soluble peptide scaffold that possesses high ^{19}F signal, and can be readily degraded to prevent bioaccumulation and persistence in the environment.

1.2 Molecular imaging

Molecular imaging is the process of using specialized instrumentation and targeted imaging agents to gain insight into tissue characteristics and biochemistry of a given subject. Specifically, an exogenous reporter molecule is designed and delivered to a site within the body for which it has a high affinity. This reporter molecule, upon binding to the site of interest, can then be analyzed by various imaging modalities (e.g., MRI, PET, ultrasound) to provide information about the molecular target studied (Figure 1-1). The goal of molecular imaging is to gain an improved understanding of the biological processes of living organisms. This can be further extended to the imaging of targeted probes for disease diagnosis. This interdisciplinary field of study provides many routes to understanding the nature of disease at the molecular level versus tissue level. There are many widely used methods employed worldwide with various pros and cons, but two that are particularly well suited for molecular imaging techniques are positron emission tomography (PET) and magnetic resonance imaging (MRI).

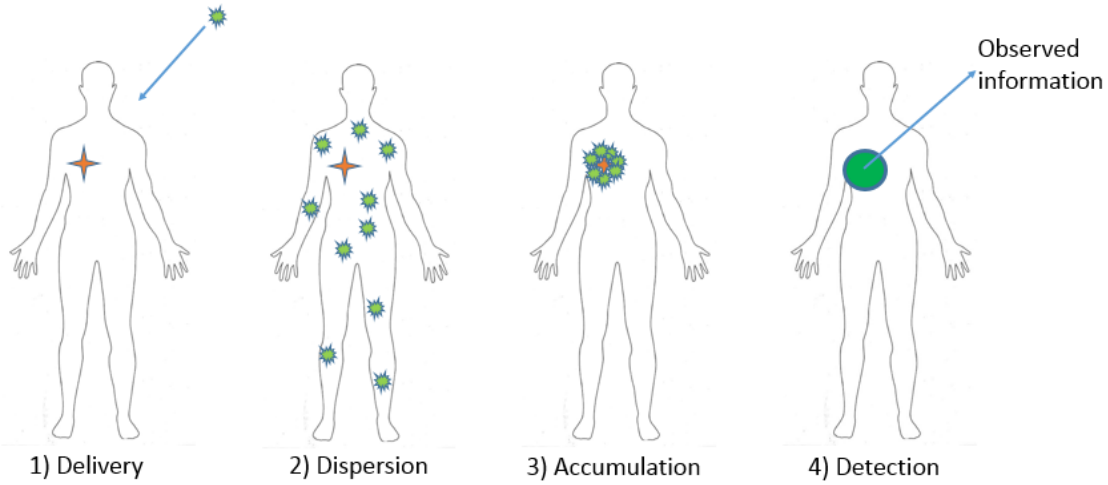


Figure 1-1: General outline for molecular imaging. A reporter molecule is delivered to the body where it disperses throughout, followed by accumulation at the site of interest. The molecule can then be detected using various imaging modalities (MRI, PET, etc.)

Positron emission tomography (PET) scans involve the incorporation of a radioactive nuclide into a biologically active carrier molecule. This radioactive material is generated using a cyclotron to produce nuclides with short half-lives, typically from 2-120 minutes.[4] The unstable radioactive compound will then undergo β-decay, releasing positrons which will annihilate with electrons sending two gamma-ray photons that can be visualized by the PET detectors.[5] Commonly in the clinic, [^{18}F]-2-fluoro-2-deoxy-glucose ([^{18}F]FDG) is the radioactive compound used and its accumulation in cancerous tumors is observed.[6] This compound is a derivative of glucose, and due to the increased metabolism of cancerous cells, [^{18}F]FDG tends to accumulate in tumors more than surrounding healthy cells allowing this to be a powerful imaging domain.[6] PET scans do indeed provide detailed quantitative data as well as high sensitivity, but there are dangers associated with utilizing such powerful ionizing radiation. Due to the short lifetime of the radioactive material, they must be produced shortly before use requiring a cyclotron to be available in the clinic which causes the procedure to be more expensive than other molecular imaging techniques. Although a powerful imaging technique, the need for potentially harmful radiation and on-site development of radioactive material can limit the availability of PET for broad use. MRI, which is non-invasive, non-radioactive, and less expensive to utilize then offers an interesting alternative for development of molecular imaging techniques.

1.3 Magnetic resonance imaging

Magnetic resonance imaging involves the use of powerful magnets and radiofrequency pulses to visualize the tissues within the body. The principles of MRI are identical to those of nuclear magnetic resonance spectroscopy (NMR). These techniques measure the differences in their magnetic dipole in varying environments.[6] MRI is a useful imaging modality as it does not require harmful ionizing radiation, and has a greater ability to collect soft tissue images as compared to computed tomography (CT). Furthermore, it has high depth penetration and good spatial resolution in a clinical setting (1 mm vs 5-7 mm for PET).[6] MRI however has comparatively low signal sensitivity. The concentration required for detection using MRI is 10^{-3}–10^{-5} M, whereas PET can detect concentrations of

10^{-11}–10^{-12} M.[7] Consequently, MRI requires higher concentrations of imaging agent compared to other molecular imaging techniques to adequately see changes in the physiological environment. To increase sensitivity, contrast agents can be used as well as data processing methods such as sparse sampling.[8]

The majority of MRI examinations performed investigate the differences in the environment of protons within the body. Proton MRI studies are quite powerful in showing anatomical information due to the contrast between water and the various tissues of the body; however, there is significant concern about the effectiveness of reporter molecules in ^{1}H MRI due to the large background signal from the high concentration of water universally present in the body.[9] To address the aforementioned deficiency in signal sensitivity, various paramagnetic contrast agents have been developed. These contrast agents affect either the longitudinal (T_1) or transverse (T_2) relaxation times of the local environment, thus providing a difference observable by the MRI instrument between the biomarker being studied and the immediate environment. Gadolinium chelates are often the contrast agent of choice, which primarily affect T_1 times, however superparamagnetic iron oxide (SPIO) compounds are also utilized for the purposes of T_2 weighted contrast.[10] The first clinically used MRI contrast agent is a gadolinium chelate of diethylenetriaminepentaacetate (Gd(DTPA)), called Magnevist® and was distributed by Bayer AG. There are however concerns about the large doses required, and the resultant water pollution from discharged contrast agent.[11] The recommended dose is 0.1 mmol/kg,[11] and considering the average adult body mass of 62 kg,[12] this would equate to nearly 6 grams of Magnevist delivered per patient, and excreted into the water supply. Thus, there exists a need for alternative contrast agents that limit the dosage to patients and downstream accumulation in the environment.

1.3.1 ^{19}F MRI

The use of fluorine in magnetic resonance studies offers an alternative to typical proton centric experiments. Fluorine is an attractive nucleus for MRI studies due to its 100% natural abundance as the ^{19}F isotope, a gyromagnetic ratio of 40.05 MHz T^{-1}, and its nuclear spin of ½.[9] These values are important as they closely match those of hydrogen, which is the most widely used nucleus in MRI studies. An additional benefit to the use of

^{19}F is the very low endogenous signal within the body. While there is a considerable concentration of fluorine in the bones and teeth, due to their immobilization in this state they display very short spin-spin relaxation times (T_2), which renders them invisible to the detector.[3] This makes the introduction of reporter molecules labeled with ^{19}F an intriguing method of undergoing MRI experiments as there will be little ambiguity as to what is producing any signal observed from the instrument. Biomarkers with specialized probes labeled with ^{19}F, could be used to determine the site of diseases. For example, a perfluoro-15-crown-5-ether (Figure 1-2) was formulated in a nanoparticle engineered to bind to $\alpha_v\beta_3$ integrins observed in rabbits on a high cholesterol diet.[13,14] The rabbits with treated with the targeted ^{19}F nanoparticles showed 220% increased fluorine signal than from untargeted perfluoro crown either nanoparticles.[13]

Signal sensitivity is a concern however, as ^{1}H MRI uses contrast based off of the high water concentration (55 M), and hydrogen accounts for nearly 2/3 of the nuclei present in the body.[3] Fluorine has a limit of detection in vivo of 30 mM at 1.5 T and a 6 minute scan time (clinically relevant parameters).[2] This poses the challenge of incorporation of a high concentration of ^{19}F nuclei that does not result in any form of toxicity or difficulty in excretion. There also is concern for the environmental impact of perfluorinated chemicals, such as perfluorooctanoic acid (PFOA) and perfluorooctanylsulfonic acid (PFOS). As such, a key component in the design of ^{19}F MRI agents is a high intensity ^{19}F signal, achieved through the presence of numerous fluorine nuclei on the chemical scaffold, while limiting its ability to accumulate within the body or environment. Coupling the use of an engineered protein as a targeting agent with a pre-established safety profile with multiple iterations of ^{19}F for high signal intensity, it is hypothesized that the previously mentioned limitations can be overcome and will be explored in Chapters 2 and 3.

1.4 Perfluorocarbon compounds for ^{19}F MRI

Perfluorocarbon compounds (PFC) represent the current state of the art in the field of ^{19}F MRI. These compounds typically involve the complete conversion from C-H to C-F bonds within a structure. Carbon-fluorine bonds are considerably stronger than carbon-hydrogen bonds. Favorable orbital overlap makes C-F bonds (485 kJ/mol) the most stable single bond formed in organic chemistry, compared to 425 kJ/mol for C-H.[15] The C-F

bond energy of a trifluoromethyl group has even been reported as high as 531 kJ/mol.[16] Due to this high bond stability, PFCs are generally considered to be chemically inert. Because fluorine possesses seven outer shell electrons, these compounds are effectively shielded by large electron density, which will provide electrostatic repulsion with approaching compounds.[17] Electron density maps show this large cloud of electrons on a sample of varying perfluorinated compounds (Figure 1-2). In addition, the high electronegativity of fluorine combined with its small size contributes to its low polarizability, leads to both hydrophobic and lipophobic character.[15,18] It has also been determined that most PFCs with molecular weights between 460-520 g/mol are not considered toxic or carcinogenic, as such many of these compounds designed for in vivo systems are within that range.[19] Alternatively, highly fluorinated molecules with polar head groups like PFOA and PFOS have been implicated in cancer cases, and have largely been phased out of production in the United States.[20]

The benefits of PFCs however must be weighed against the wealth of limitations that come with their unique properties. As they are very hydrophobic, it is extremely difficult to form solutions with PFCs. Often, PFCs are made into emulsions to aid in delivery within the body. Fluosol®, a PFC emulsion, was approved by the FDA in 1989 for use in angioplasty procedures.[21] This emulsion contained perfluorodecalin (PFDC, Figure 1-2) and perfluorotripropylamine (PFTPA) which had to be frozen and stored separately, and mixed prior to use since the emulsion only remained stable for approximately six hours.[3] Subsequently, due to issues with the storage and thawing of this compound, Fluosol® was removed from the open market by the FDA in 1994.[22] Even more prevalent, was that researchers discovered at least two thirds of the total injected PFC was located in the lung tissue one year after the original injection.[23] This is due to the excretion process of these emulsions or PFCs loaded into nanoparticles. The carrier vehicle of the PFC gets broken down via the reticuloendothelial system (RES) and the freed PFC then binds to blood lipids and is eventually mainly excreted very slowly through the lungs via exhalation.[24] This slow excretion from the body leaves a background signal for future ^{19}F MRI studies.

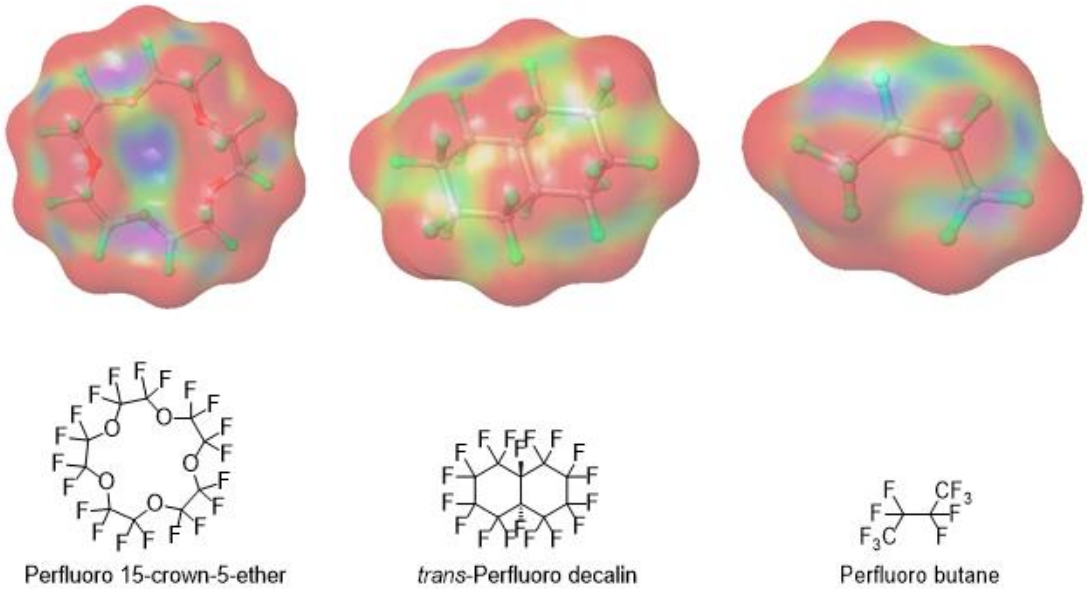


Figure 1-2*:* Poisson-Boltzmann electrostatic map showing electron density of three perfluorinated compounds. The color is a rainbow scale where red = high electron density, violet = low electron density. Image generated using Schrodinger Maestro 9.9.

1.5 ^{19}F MRI agents based on non-PFC scaffolds

In contrast to PFC-centric imaging agents, peptides, polymers, and dendrimers offer an interesting alternative in MRI contrast agents whose physical parameters can be more easily tuned to improve their practical usage (Figure 1-3). PFCs are often highly symmetrical molecules which improves their signal sensitivity by providing an intense ^{19}F signal due to the magnetically equivalent ^{19}F atoms (e.g., PFCE, Figure 1-2). However, if structural changes were introduced in an attempt to improve water solubility, this symmetry would be lost and the outcome would be a potentially dramatic decrease in signal output. A divergent dendrimer termed ^{19}FIT was reported in 2009 that possesses 27 equivalent ^{19}F nuclei and is water soluble, but to date no further reports regarding increased fluorine content or conjugation to drugs has been reported (Figure 1-3).[25] Fluorinated polymers have been studied as a source of ^{19}F signal with improved water solubility by introducing block co-polymers with highly hydrophilic monomers in addition to hydrophobic ^{19}F-

containing monomers.[26] These polymers have been successful in introducing a large amount of ^{19}F signal, however the typically high molecular weight of these polymers will have an impact on the T_2 relaxation of the molecule, leading to broad ^{19}F resonances. Conversely, solid-phase peptide synthesis can be used to develop short sequences of ^{19}F labeled biopolymers with both ^{19}F-rich residues and hydrophilic residues. In addition, these peptides have been engineered to remain unfolded providing a much more narrow resonance and thus higher signal to noise than a similarly fluorinated large polymer.[27] By virtue of being peptides, these molecules are derived from nature and are thus susceptible to enzymatic degradation to prevent their accumulation within the body. For a detailed account of the development of a highly fluorinated, water-soluble peptide, please see Chapter 2 of this thesis.

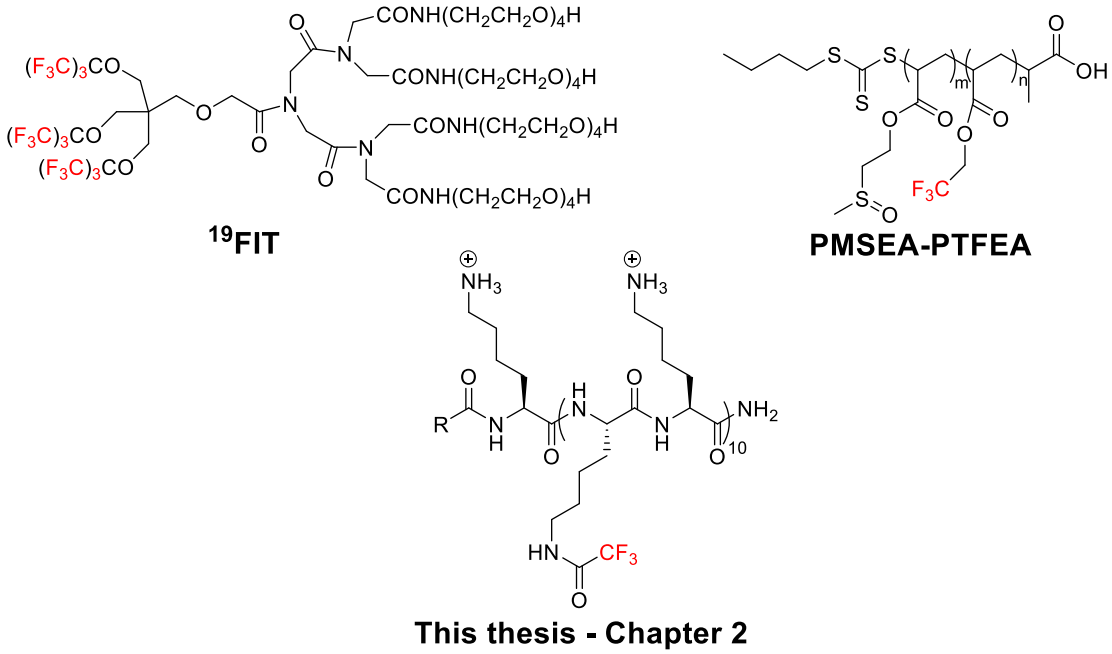


Figure 1-3: Select ^{19}F MRI agents based on non-PFC scaffolds. 19FIT is discussed in Jiang et al[25], PMSEA-PTFEA is described in Fu et al[26].

1.6 Solid-phase peptide synthesis

Pioneered by R.B. Merrifield in his seminal 1963 article, solid-phase peptide synthesis (SPPS) has provided a vast window into the development of peptide scaffolds.[28] The process begins with an insoluble, porous support with a reactive linker. Typically, the C-terminal amino acid residue in the sequence is coupled to this reactive linker. This

residue has both its N-α amine, as well as any reactive side chain protected to avoid cross-reactivity. Following a washing and filtration step, the N-α amine is deprotected orthogonally to the other protecting group on the side chain. All peptides employed in this thesis will utilize base-labile fluorenylmethoxycarbonyl (Fmoc) as the N-α protecting group, and a host of acid-labile protecting groups on the side chain functional groups (i.e., Boc, t-butyl). Following Fmoc-deprotection another residue is then coupled in a manner similar to the first and the process is repeated until the entire sequence has been installed. Often, N-terminal acetylation is performed followed by a global deprotection of the side-chain protecting groups and cleavage from the solid support using a cocktail containing a high concentration (> 90%) of trifluoroacetic acid (Figure 1-4). These peptides can then be purified utilizing reverse-phase HPLC and analyzed via MALDI-TOF MS. SPPS therefore allows for facile development of peptides of high uniformity compared to conventional solution phase or polymerization methodology. Unnatural amino acids as well as D-amino acids can also be incorporated into peptide scaffolds by this manner.

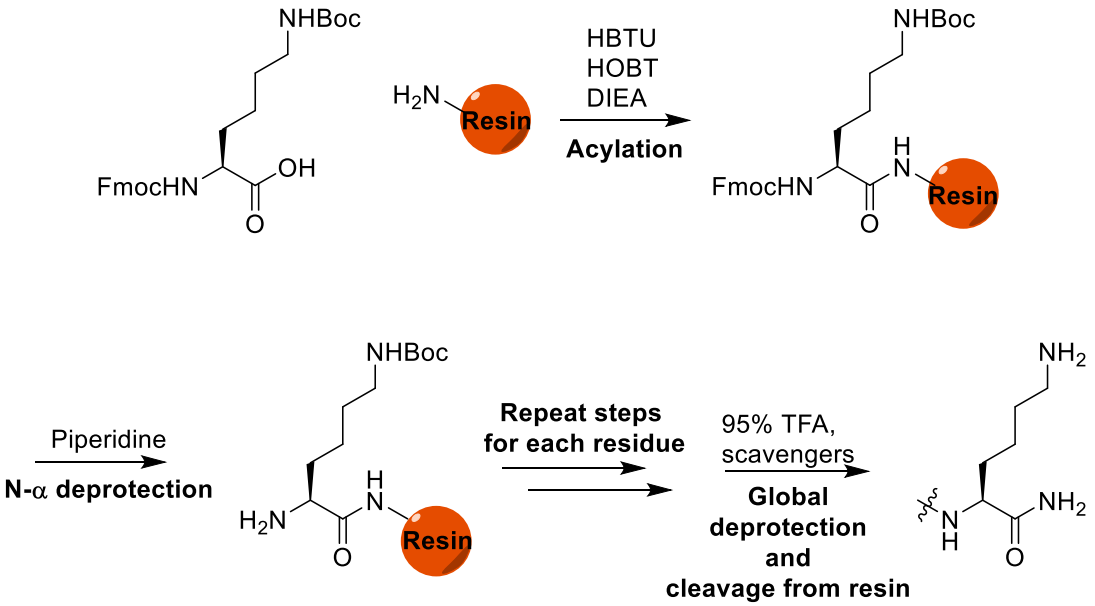


Figure 1-4: General schematic for Fmoc solid phase peptide synthesis. N-α-Fmoc-L-N-ε-Boc lysine is used as a representative amino acid.

1.7 The "Disordered Symmetry" Hypothesis

NMR spectroscopy has been a useful tool in determining the structure of many different proteins. In its folded state, a protein's tertiary structure provides distinct chemical environments for the amino acid side chains.[29] As such, proton resonances within identical residues could have markedly different ^{1}H NMR chemical shifts based on their position in the folded structure. A residue that is buried within the structure is in a different chemical environment than one that is solvent exposed. However, as a protein becomes unfolded, and loses this structural rigidity, the resonances observed from the side chains begin to coalesce as they are in very similar chemical environments.[30] For example, amide protons have a typical chemical shift range in folded proteins of 6.5-10 ppm, but in unfolded proteins this range narrows to 8-8.5 ppm.[31] Therefore, it is plausible that a residue possessing a distinct chemical shift and a high occurrence within the unfolded protein would give rise to an intense singular resonance within a NMR spectrum. Furthermore, denaturation studies of a fluorinated protein have shown a trend toward coalescence in the ^{19}F NMR resonances as shown in Figure 1-5.[32] In this study, Frieden and co-workers site-selectively replaced all eight phenylalanine residues in a murine intestinal fatty acid binding protein (IFABP) with 4-fluorophenylalanine. The ^{19}F NMR spectrum of a well-folded protein showed all eight distinct resonances from each residue. When the protein was denatured chemically using increasing concentration of urea, the ^{19}F resonances began to coalesce as the side chains began sampling similar chemical environment.[32]

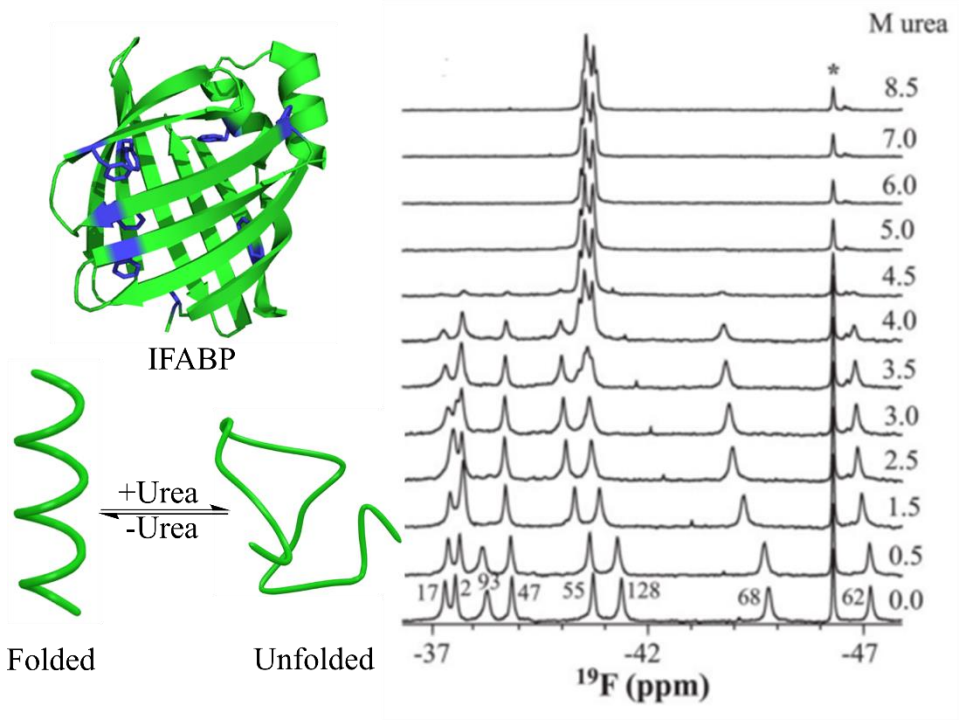


Figure 1-5: ^{19}F NMR spectrum of IFABP, with all phenylalanine residues replaced with 4-fluorophenylalanine (PDB: 1IFB[33]) titrated with increasing concentration of urea. Unfolded protein resonances begin to coalesce. The asterisk marks 6-fluorotryptophan, used as a chemical shift reference, all spectra were normalized to this resonance. Figure Adapted with permission[32] Copyright 2005 American Society for Biochemistry and Molecular Biology.

NMR experiments with proteins often involve two-dimensional techniques such as HSQC NMR to aid in the determination of structure based on correlations between different nuclei, such as ^{1}H and ^{15}N.[30,34–37] These types of experiments are typically done because residues in varying environments will display different chemical shifts. If the environment is identical, in the case of an unfolded domain of the protein, multiple iterations of a distinctive residue could be inserted into the protein to arrive at a singular high signal resonance. This "disordered symmetry" is the central hypothesis in the development of a high signal probe for use in ^{19}F MRI.

1.8 Epidermal Growth Factor Receptor (EGFR) as a Breast Cancer Molecular Imaging Target.

Receptor tyrosine kinases (RTK) are membrane bound proteins whose kinase activity is activated by the binding of a growth factor causing the receptor to dimerize. This activation is due to the proximity of the two receptor molecules upon dimerization allowing them to phosphorylate their respective tyrosine residues in a process called transautophosphorylation (Figure 1-6).[38] Responses to this activation include cell growth and proliferation, stimulating cell survival which in the case of overexpression is likely to lead to tumor formation.[38] A specific RTK, epidermal growth factor receptor (EGFR), is a biomarker commonly overexpressed in breast cancer.[39]

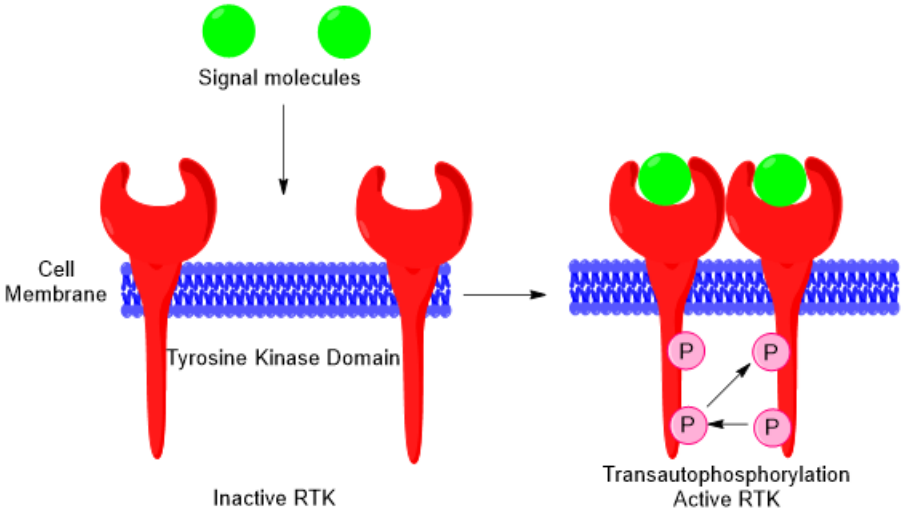


Figure 1-6: Diagram of receptor tyrosine kinase activity. Signal molecules can bind to the extra cellular receptor domain, causing a dimerization. The resulting dimer is phosphorylated due to close proximity of the tyrosine kinase domain of each monomer.

In breast cancer, there are numerous mechanisms by which EGFR activation can stimulate a proliferative response (Figure 1-7). Phosphoinositide-3 kinase (PI3K) is recruited to the activated EGFR, which in turn activates protein kinase B (AKT) via the increasing the amount of phosphatidylinositol-3,4,5-triphosphate (PIP_3) on the cell membrane.[40] AKT is bound to the Pleckstrin Homology (PH) domain of PIP_3, and can then recruit mammalian target of rapamycin (mTOR) which induces cell proliferation leading to cancer.[40]

Another pathway is via the mitogen activated protein (MAP) pathway. Activated EGFR can bind Grb2, a Src homology 2 (SH2) domain protein, through its phosphorylated tyrosine residues. Grb2 in turn recruits a guanine nucleotide exchange factor called SOS, which instigates the removal of guanosine 5'-diphosphate (GDP) from the protein Ras, which can then bind guanosine 5'-triphosphate (GTP) and become active. Activated Ras then binds to Raf, which begins a kinase cascade of MEK, followed by ERK, which then phosphorylates a variety of transcription factors such as MNK and c-Myc leading to tumor formation.[41] Through these and other mechanisms, it has been shown that EGFR is expressed in 50-70% of carcinomas of the lung, colon, and breast.[42] This provides a useful and important biomarker for many cancers, which can be exploited for early detection of the disease, leading to more favorable prognoses.

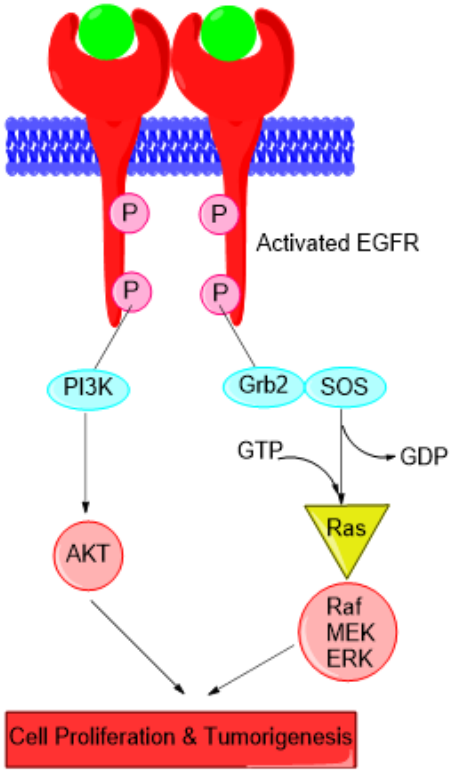


Figure 1-7: EGFR and select biological pathways leading to tumorigenesis.

1.9 Peptides and proteins as EGFR selective targeting agents

Due to the impact of the overexpression of EGFR in various forms of cancer, it is then of importance to develop targeting agents that bind to EGFR-rich cells. Peptides and proteins have emerged as selective targeting agents with the capability of binding the cell surface receptor. In 2005, a phage display screen yielded the sequence YHWYGYTPQNVI (termed GE11) to bind to EGFR with K_d = 22 nM.[43] This was followed by a 2009 study, where a computer-aided design approach yielded a peptide with the sequence LARLLT (termed D4) that was found to bind to EGFR.[44] These two sequences have then been further elaborated by others to target EGFR coupled with an imaging modality. Wagner and co-workers coupled D4 to chemically self-assembled protein nanorings with a DOTA group to chelate ^{64}Cu and detect by PET/CT.[45] Viehweger et al, coupled D4 to a dual-imaging modality of a ^{64}Cu chelate as well as a fluorescent tag (Figure 1-8).[46] The peptide GE11 was used by Kohno et al, in conjugation to an amphipathic peptide comprised of leucines and lysines to selectively lyse EGFR-overexpressing cells (Figure 1-8).[47]

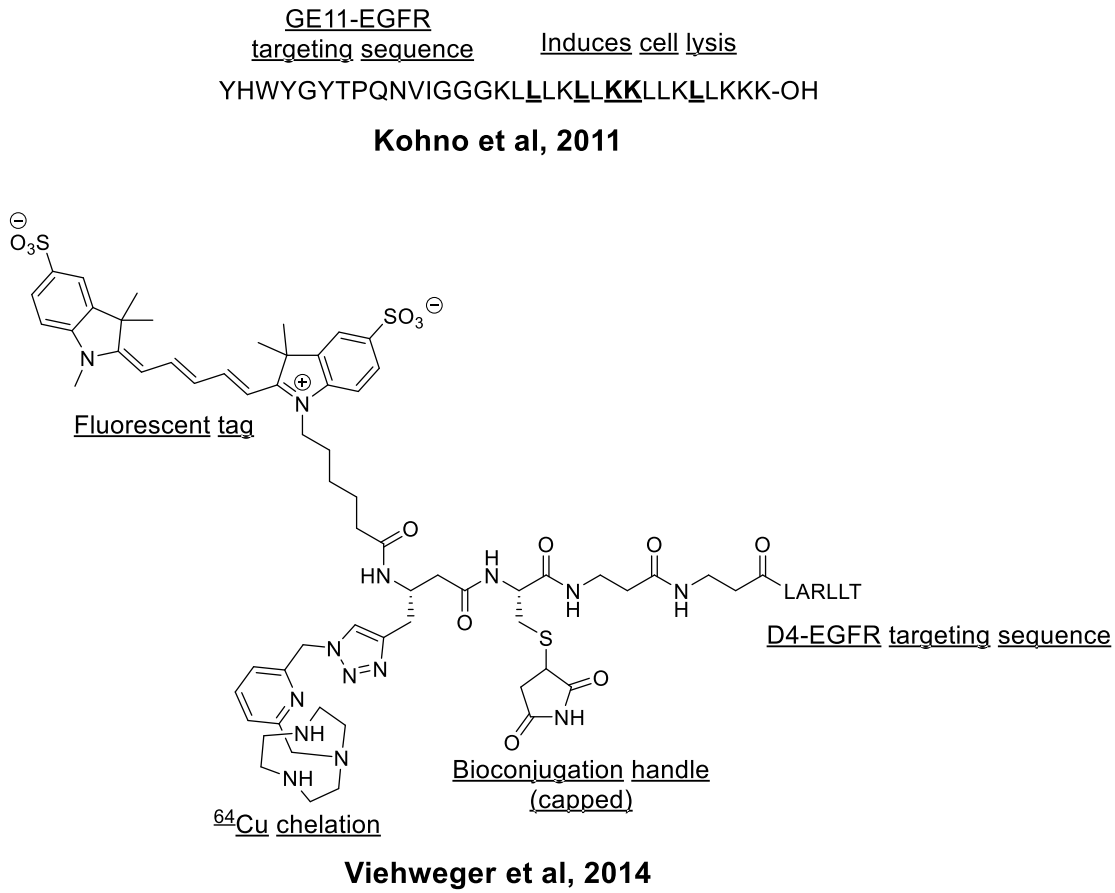


Figure 1-8: Selected EGFR-targeting peptides coupled to a cell lysis-inducing tag (Top, Kohno et al, 2011)[47] and multi-modal imaging domains (Bottom, Viehweger et al, 2014)[46]. Underlined letters indicate D-amino acids.

As an alternative to short peptide sequences, a protein scaffold based on the tenth human fibronectin type III domain was engineered to bind with high specificity to EGFR. This small protein scaffold has a molecular weight near 10 kDa and is highly stable. It is a beta-sandwich protein which possesses three loop regions that can be engineered to bind various biomarkers. In particular, a mutant termed EI3.4.3 was engineered by Hackel et al to bind EGFR with K_d = 250 pM and was coupled to a radioactive ^{64}Cu ligand to serve as a molecular imaging construct for use in PET studies (Figure 1-9).[48,49] The fibronectin domain holds promise for development of selective targeting agents, and in the past several years has been developed to bind with various biomarkers. Adnexus (Bristol Myers Squibb) has developed therapeutics based on the fibronectin domain, termed Adnectins™, which have used the fibronectin scaffold's similarity to antibody domains to bind vascular

epithelial growth factor receptors (VEGFR) and are currently in clinical trials for various cancer studies.[50,51]

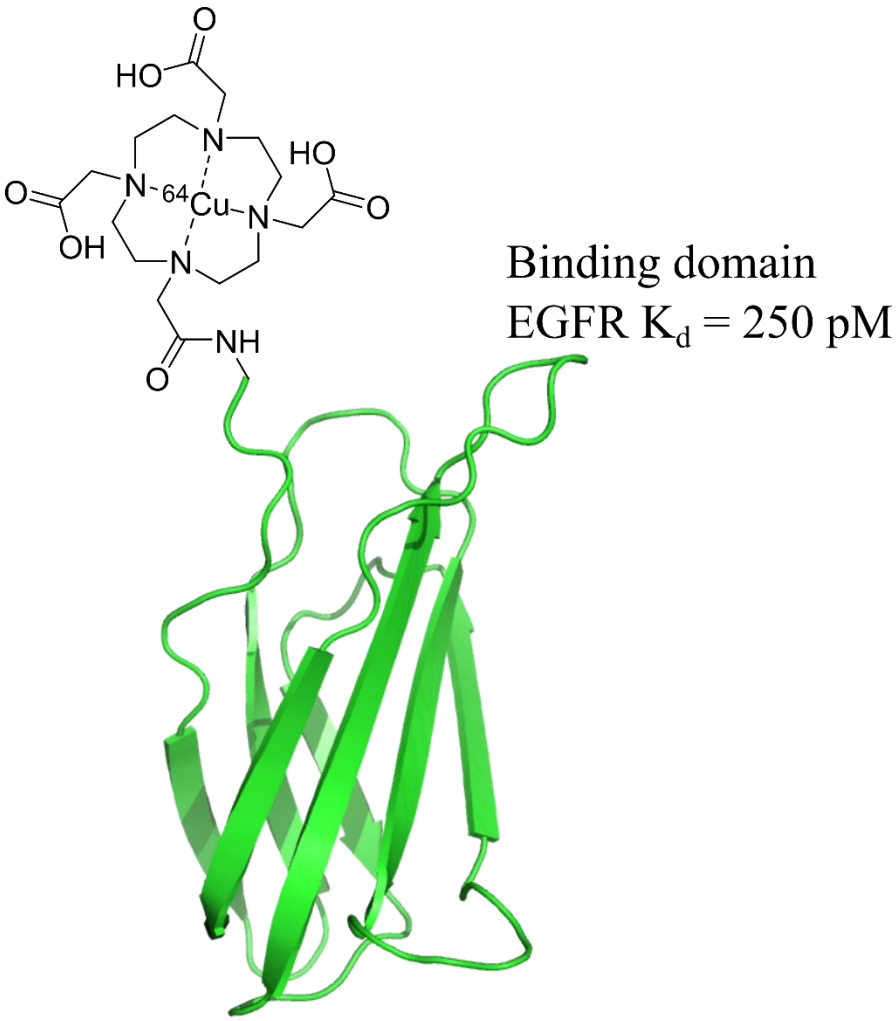


Figure 1-9: NMR structure of Fibronectin scaffold with a ^{64}Cu chelating DOTA tag appended to its N-terminus. (PDB: 1TTB) Figure adapted from Hackel et al.[48]

Transition:

This concludes the introduction to Part I, the development of peptide-based ^{19}F MRI agents for molecular imaging. The remaining portion of this chapter will discuss a second project describing the development of a small molecule inhibitor of an epigenetic regulatory protein termed BPTF.

Part II: Development of a small molecule inhibitor for the bromodomain of BPTF

A portion of this work is described in "BPTF Maintains Chromatin Accessibility and the Self-Renewal Capacity of Mammary Gland Stem Cells" W.D. Frey, A. Chaudhry, P. F. Slepicka, A. M. Ouellette, S. E. Kirberger ,W.C.K. Pomerantz, G.J. Hannon, and C.O. dos Santos, *Stem Cell Reports*, **2017**, *9*, 23-31. Portions of sections 1.11 and 1.12 are scheduled to appear in upcoming manuscripts.

1.10 Epigenetics

Epigenetics is the study of varied and heritable phenotypic responses by changes of gene expression as opposed to alteration of DNA nucleotide sequence.[52,53] While direct modification of DNA (i.e., cytosine methylation) is indeed a factor in epigenetics, alterations of translated proteins are also a major force in determining the phenotype of any given cell population. These post-translational modifications (PTM) impact whether genes will be expressed, and different combinations of modifications can result in drastically different outcomes. Moreover, PTMs are reversible, allowing for an increasingly dynamic and complex system of gene regulation. Most cell types in the human body contain the same genetic code, so these modifications are key in the development and differentiation of cells to sustain life. The mechanism by which this process proceeds largely involves the interactions and accessibility of DNA and histone proteins.

Histone proteins largely serve as a method to package the long strands of DNA in such a way that they can fit inside the nucleus a cell. Wrapped 1.7 times around an octameric core of histone proteins, are 146 base pairs of DNA in a construct called the nucleosome.[54] Four distinct histone proteins (H2a, H2b, H3, and H4) comprise the histone core of the nucleosome.[55] These nucleosomes are all linked together by the polymeric strands of DNA in a higher order assembly called chromatin. Chromatin can exist in two states; euchromatin – where DNA is loosely associated to the histone core and thus more accessible to transcriptional machinery, and heterochromatin – DNA is tightly associated to the histone core and transcription is inhibited (Figure 1-10).[56] Modifications to the termini of these histone proteins can play a key role in the local state that chromatin will take. Amino acid side chains of these proteins can be covalently modified by several enzymes to alter the chemical environment and either encourage or discourage DNA association. This process is dynamic and is largely carried out by three classes of proteins often colloquially referred to as "writers," "readers," and "erasers".[57]

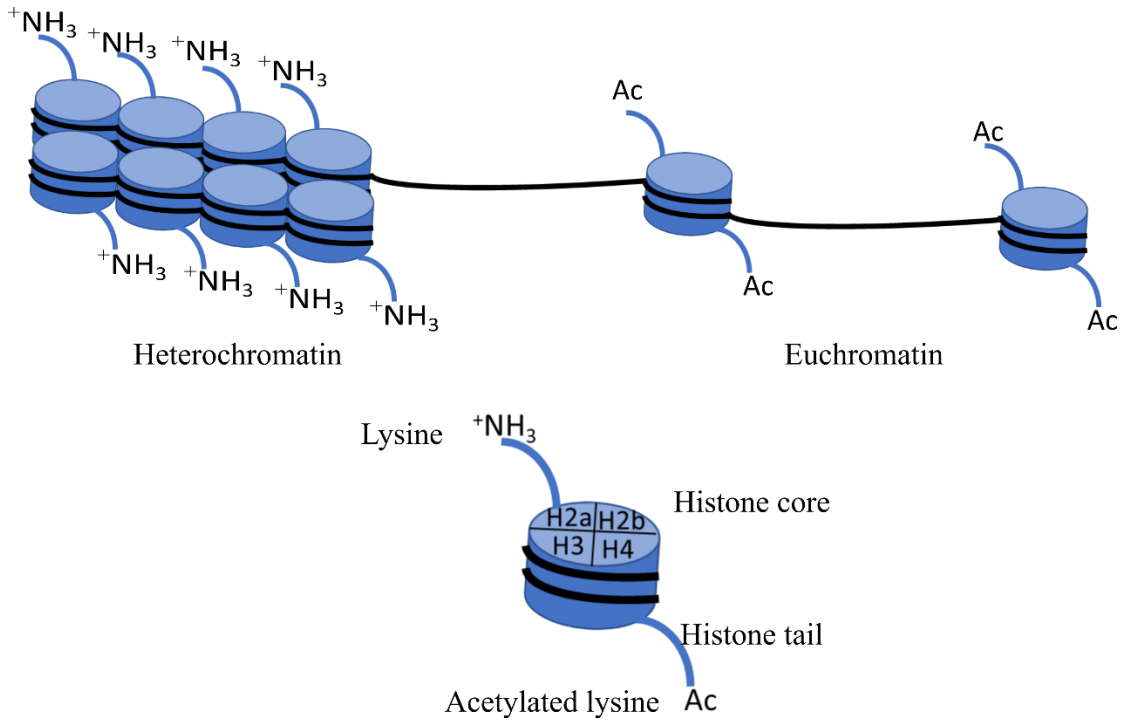


Figure 1-10: General structure of heterochromatin and euchromatin. Acetylated lysine residues decrease interaction between histones and DNA, leading to euchromatin formation and more accessible DNA.

Writer proteins are responsible for installation of PTMs to the side chains of the N-terminal histone protein tails. Examples of these include, but are not limited to, histone acetyl transferases (HAT), histone methyl transferases (HMT), and kinases. These three enzyme classes can directly alter the accessibility of DNA to histone proteins based on charge (Figure 1-11). The DNA double helix contains a negatively charged phosphate backbone that can associate through ionic interactions with the histone octamer, which possesses a high prevalence of cationic amino acid residues. The side chain of lysine, for example, exists in physiologic conditions as a protonated ammonium, displaying a positive charge that attracts DNA and causes heterochromatin formation. HATs can acylate the side chain of lysine, removing its capability to display a charge, and thus loosening the interaction between DNA and the histone, and directing a change to euchromatin formation. The inverse can be true of histone methyltransferases which will alkylate the lysine side chain and enforce cationic character and heterochromatin formation. As such,

acetylated lysines tend to lead to increased gene expression, and methylated lysines tends to suppress expression, but this is not always the case. There exists a degree of sequence-dependence on the impact of PTMs on transcription levels. Acetylation of lysine 27 on histone 3 (stylized – H3K27Ac) for example, will activate transcription, while trimethylation on this same residue (H3K27me3) represses transcription.[58,59] However, variance can exist as H3K4me3 leads to transcriptional activation despite the retention of charge on this lysine.[60] Transcription levels can be altered in a dynamic way as these PTMs are reversible, while "writers" add modifications, they can be removed by enzymes called "erasers."

"Erasers" are an opposing class of enzymes to the writers which remove PTMs, allowing the epigenetic process to be dynamic. The installation and removal of modifications to the histone core can impact which genes are expressed, and when they are expressed. Histone demethylases (HDM) and histone deacetylases (HDAC) are common eraser proteins. Different combinations of PTM modifications can directly impact the cell fate, this has been referred to as the Histone Code Hypothesis.[61] Reader proteins can recognize these modifications, and recruit nucleosome remodeling factors to help drive euchromatin or heterochromatin formation. The work done involving the study of epigenetics in this thesis focuses on a class of reader proteins called bromodomains.

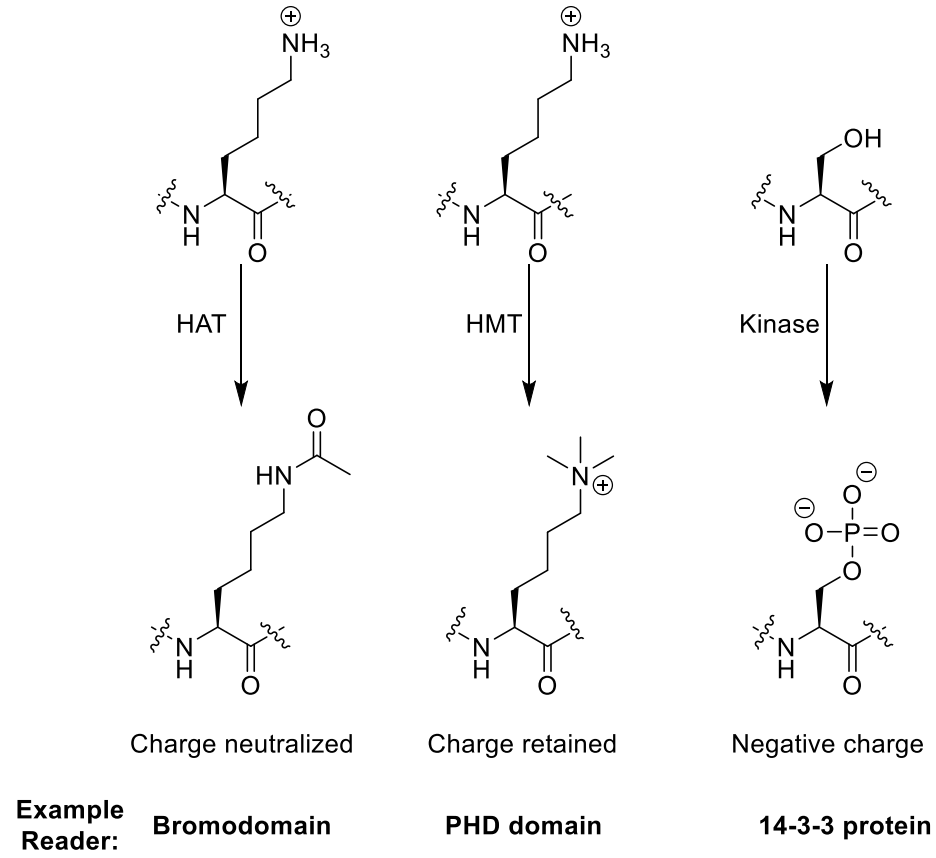


Figure 1-11: Select post translational modifications performed by identified enzyme class, with selected reader domain examples.[62–64]

1.11 Bromodomains

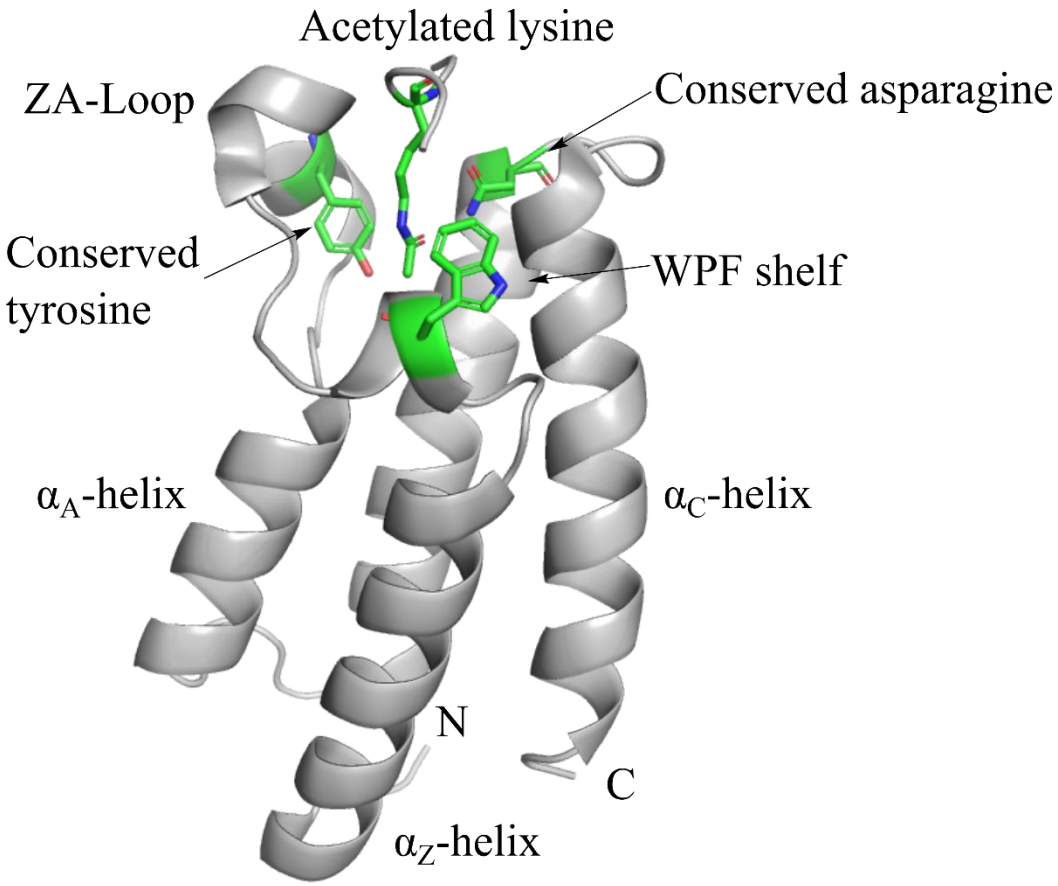


Figure 1-12: Structure of the bromodomain of BPTF with major features labeled. PDB: 3QZT[65].

Bromodomains belong to a class of effector domain, or "reader" that recognizes acylation of the side chain of lysine, primarily acetylation.[62] This class of proteins is responsible for chromatin remodeling and regulating transcription, but has also been implicated in the development of inflammation and cancer.[66] There are 61 known human bromodomains grouped into 7 families, and all of them possess the same basic structure comprised of a left-handed four α-helix bundle (Figure 1-12).[67] The four α-helices, termed the α_z, α_a, α_b, and α_c helices, are connected by two structured loop regions between α_z and α_a (ZA loop), and the other between α_b and α_c (BC loop). These loops can fold near each other making a hydrophobic pocket that instills variability within the acetylated lysine

binding site.[68] This variability is the driving force behind the design of selective inhibitors of bromodomains. As outlined in the phylogenetic tree (Figure 1-14), there exists a subset of families within the overall class of bromodomains.[69] These bromodomains are small domains of much larger protein scaffolds. To study the role of these specific domains, selective inhibitors need to target the domain only; a mass knockdown of the entire protein can remove additional enzymatic and scaffolding functions. Development of domain-selective inhibitors is complicated by the fact that some proteins possess two or more distinct bromodomains.[70,71] Each domain within the seven families tend to have more similarities within the binding site, complicating the task of developing selective inhibitors. One of these families, termed the bromodomain and extra terminal (BET) family, has been widely studied.[72–74] Many inhibitors for BET proteins have been developed (Figure 1-13), including (+)JQ1[74] and I-BET-762,[73] as well as pan-bromodomain inhibitors bromosporine[75] and acetaminophen.[76]

Beyond BET-bromodomain inhibitors, there is a need for the development of more selective probes to determine the downstream effects of individual bromodomains, especially those outside of the BET family. Hydrogen bonds to the conserved asparagine and tyrosine residues play key roles in the binding of ligands to bromodomains, however neighboring residues are variable, allowing for the design of selective inhibitors.[77] There also exists a common WPF shelf comprised of tryptophan, proline, and phenylalanine which are near the binding site in many bromodomains, and its importance will be discussed later. An added challenge in inhibitor design is that bromodomains contain five structurally conserved water molecules, which must be considered during binding events.[78] While BET proteins have been widely studied, another bromodomain containing protein called BPTF is an interesting and underexplored target for inhibitor development.

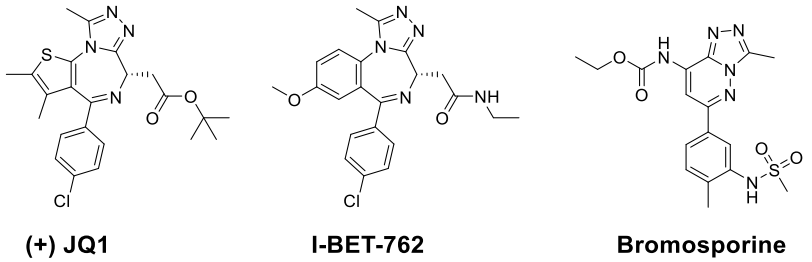


Figure 1-13: Structures of pan-BET bromodomain inhibitors (+) JQ1 and I-BET-762, and pan-bromodomain inhibitor bromosporine.

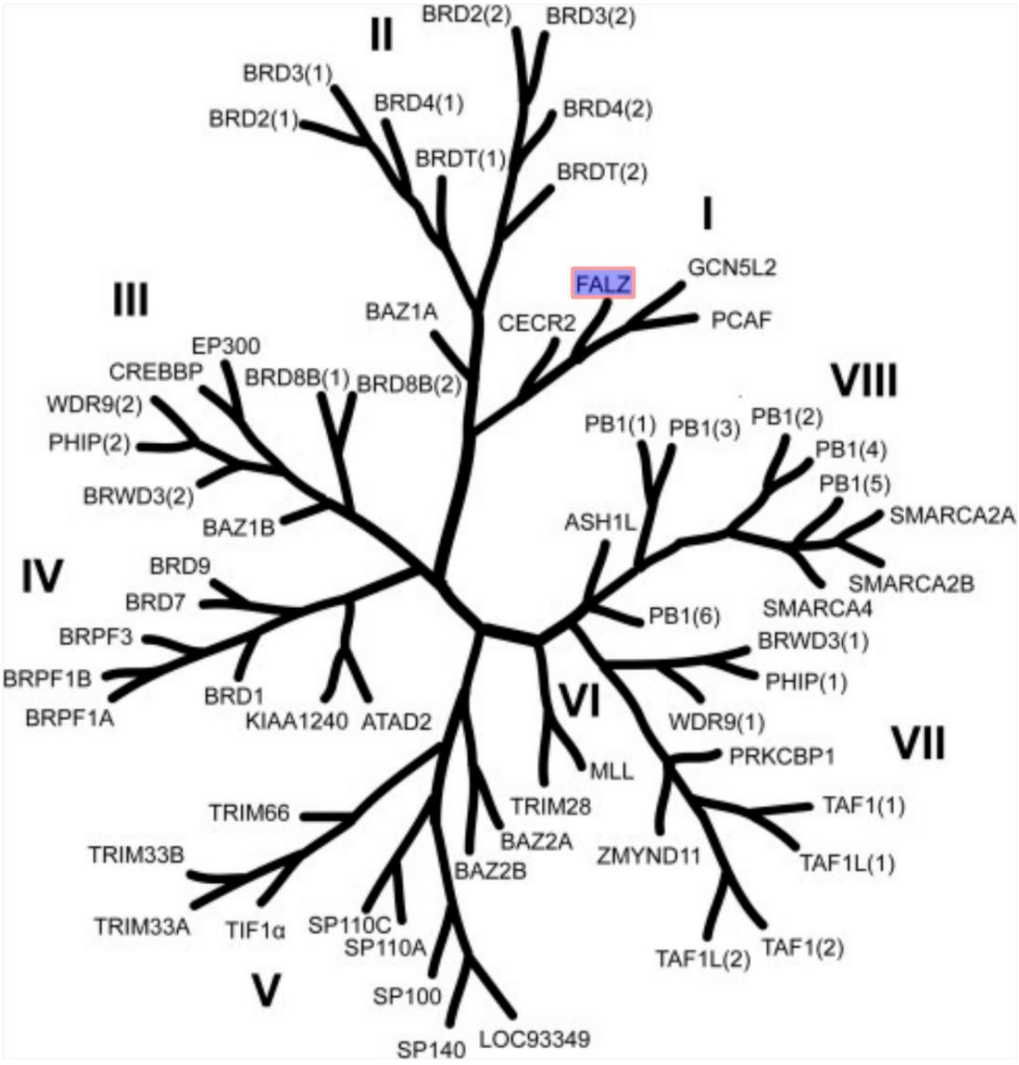


Figure 1-14: Phylogenetic tree of bromodomains organized on structural alignment. BPTF, also called FALZ, is highlighted. Adapted with permission[79] Copyright 2012 American Chemical Society.

1.12 Bromodomain and Planthomeo domain containing Transcription Factor (BPTF)

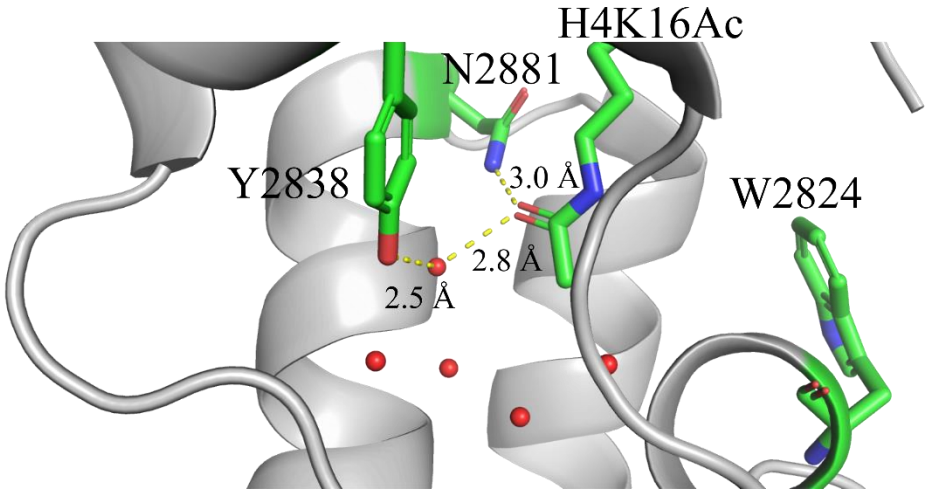


Figure 1-15: Binding site of the bromodomain of BPTF bound to histone 4, acetylated lysine 16 (H4K16Ac) with conserved asparagine and tyrosine labeled. Structurally conserved waters labeled as spheres, plotted according to positions described by Vidler et al.[80] PDB : 3QZT[65]

Bromodomain and planthomeo domain (PHD) containing transcription factor (BPTF, Figures 1-12 and 1-15) is a protein that comprises the largest component of the nucleosome remodeling factor (NURF). NURF is responsible for regulating gene expression by altering the formation of chromatin, which BPTF is responsible for its recruitment and assembly to promoter regions.[81] BPTF contains both a bromodomain as well as two PHD domains, and is implicated in Alzheimer's disease,[82] heart disease, and numerous forms of cancer.[82–89] In lung cancer, Rotter and co-workers demonstrated that knockdown of BPTF prevented cell proliferation, and indicated that BPTF overexpression was linked to a malignant phenotype.[87] The work of dos Santos and colleagues showed BPTF was critical in the development of mammary cells and the depletion of BPTF invoked cell cycle arrest and failure for stem cells to develop into healthy mammary epithelial cells.[94] Landry and co-workers also showed that BPTF inhibits natural killer (NK) cells' antitumor activity.[95]

In a separate study, The lab of Fernando Real in 2013 discovered that knockdown of BPTF in bladder cancer cells showed reduced proliferation, which led them to suspect a role of BPTF in relation to the oncogene c-MYC.[90] This led to a follow up report in 2016

where Real and co-workers describe BPTF and its role with c-MYC. In this report, BPTF is shown to bind in multiple histone sites through both its methyllysine reading PHD domains, as well as through its bromodomain. Silencing the BPTF gene drastically reduces acetylation of histone 3, which is known to be a location for c-MYC activation.[84] Knockdown of the whole BPTF gene showed activity with regards to cancer development, however due to the multiple binding domains with the BPTF protein, it is still unclear what the bromodomain's contribution to disease progress is.[87] To this end, it is necessary to develop inhibitors of BPTF that target its bromodomain, ideally with specificity over other bromodomain containing proteins. BPTF however remains a difficult protein to target, even though inhibitors for structurally similar bromodomains like PCAF,[91,92] GCN5,[91] and CECR2[93] have been reported. A WPF shelf, previously mentioned as a common motif found in bromodomains, is also present in the bromodomain of BPTF. This motif has a key role in the analysis of binding to bromodomains, as proteins can be site-selectively fluorinated to employ ^{19}F NMR as an assay to probe binding interactions with the bromodomain and developed inhibitors.

1.13 Protein-Observed Fluorine NMR (PrOF NMR)

Protein-observed fluorine (PrOF) NMR is a useful tool for detecting ligand interaction to a ^{19}F-labeled protein scaffold.[96] Unlike ligand-observed methods, protein-observed methods for detecting binding have a distinct benefit in that some information of protein stability, as well as binding site identification are possible.[97] In comparison to ^{1}H-^{15}N HSQC experiments, PrOF NMR is not capable of determining the same level of binding location or protein fold, but it has the advantage of utilizing a much simpler spectrum that can be easier to determine and quantify binding interactions. The ^{19}F isotope is 100% isotopically abundant, and there is no detectable amount of naturally occurring ^{19}F in biological systems. This provides a strong signal, without background and negates the need to perform solvent suppression in NMR analyses. Fluorine is also very responsive to changes in its chemical environment, improving sensitivity to binding events, and its large chemical shift range (400 ppm vs 10 ppm for H) provides a likelihood of resonance dispersion in many cases.[98]

The selective incorporation of ^{19}F-labeled amino acids into a protein construct can be conveniently performed in laboratories capable of *E. coli*-based protein expression.[99] The basic premise of expressing fluorinated versions of a protein involves the exchange of a native amino acid or precursor with its fluorinated counterpart. The work performed in this thesis involves the incorporation of 5-fluorotryptophan (5FW) in place of the lone tryptophan found in the bromodomain of BPTF. BL21 cells are used along with 5-fluoroindole to install a single fluorine atom in the WPF shelf, near the acetylated lysine binding site in the bromodomain of BPTF. 5-Fluoroindole can be biosynthetically incorporated as 5FW in *E. coli* cells by taking advantage of the native Shikimate pathway,[100] this is particularly advantageous as 5-fluoroindole ($48.50/g, Sigma-Aldrich) is considerably less expensive than 5-fluorotryptophan ($348/g, racemate, Sigma-Aldrich) from commercial vendors. Aromatic amino acids are often enriched at protein-protein interfaces, as such this method provides a robust pathway for analyzing interactions at these interfaces.[101] The ^{19}F nuclei has a slightly larger van der Waals radius than 1H (1.47 Å vs 1.10 Å), so determination of structure perturbation by a secondary assay such as circular dichroism or x-ray crystallography should also be considered.[102]

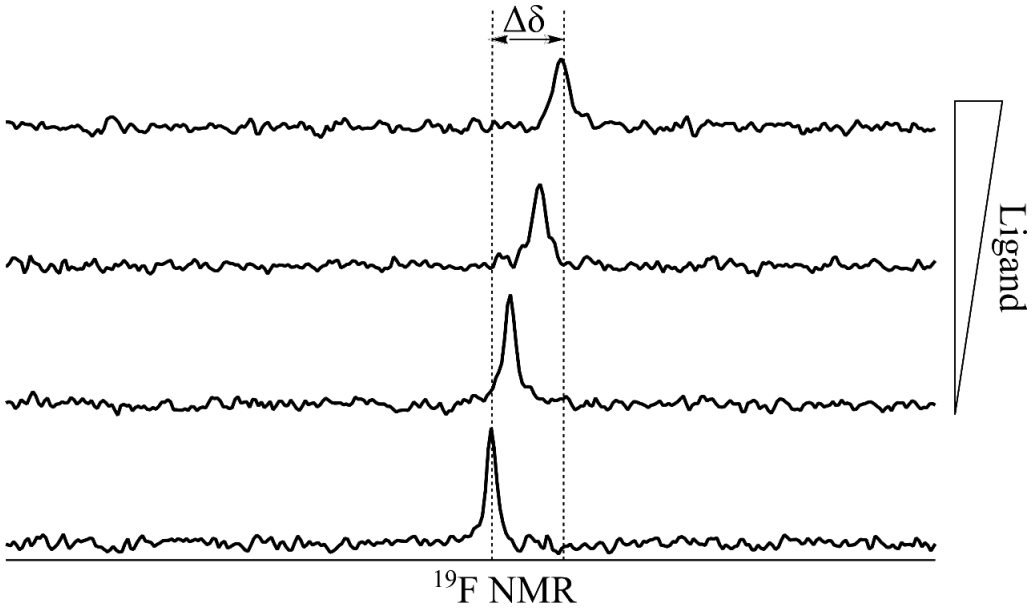


Figure 1-16: Example PrOF NMR titration spectra. Increasing ligand concentration can shift the resonance upon binding, and the change in chemical shift can be plotted to determine K_d. ^{19}F NMR spectra are of the bromodomain of 5FW BPTF, which contains a single tryptophan.

^{19}F NMR spectra can be utilized to observe binding interactions by both quantitative and qualitative means. The chemical shift range of the ^{19}F nucleus is very large, and fluorine is hyperresponsive to its chemical environment, making determinations of binding that influences change in the protein nearby to the fluorinated residue facile. Quantitative measurement of the binding of a ligand to the ^{19}F-labeled protein can be measured using the following equation:

$$Y = \Delta\delta_{max} \frac{(K_d+X+P)-\sqrt{(K_d+X+P)^2-(4PX)}}{2P} \quad (1)$$

In this equation, Y refers to the observed chemical shift, $\Delta\delta_{max}$ is the maximum chemical shift, K_d is the dissociation constant, X is ligand concentration, and P is protein concentration. Following a titration, a curve can be plotted to determine K_d for moderate and weak affinity ligands. In the case of weaker ligands (K_d = mid μM to mM), a change in chemical shift is observed in the ^{19}F NMR spectrum (Figure 1-17). This change is a weighted average of both bound and unbound protein, as the chemical exchange of ligand binding is fast with respect to the NMR time scale. Chemical exchange phenomena can make the quantitative assessment of very strong binding interactions difficult to quantify because the value for k_{off} is sufficiently long that a chemical shift change is not observed, but rather two distinct populations representing a bound and unbound protein.[103] An intermediate-strength interaction, results in neither simply a discrete change in chemical shift nor the observation of two resonances, but rather a broadening of a resonance which can also be difficult to accurately quantify.

Qualitative binding information can also be obtained utilizing PrOF NMR. In the case of a protein construct with more than one fluorinated residue, binding site location can be obtained. By observing which resonance moves and which do not, it can be inferred that a binding event occurs closer to the impacted resonance (Figure 1-16). Allosteric interactions may also be observed when certain resonances move, in particular if this is seen when studying ligands and proteins with known binding locations. Another benefit of protein-observed methods is that an idea of protein behavior is provided, it is possible to observe denaturation, precipitation, and aggregation of the protein in solution (Figure 1-5).

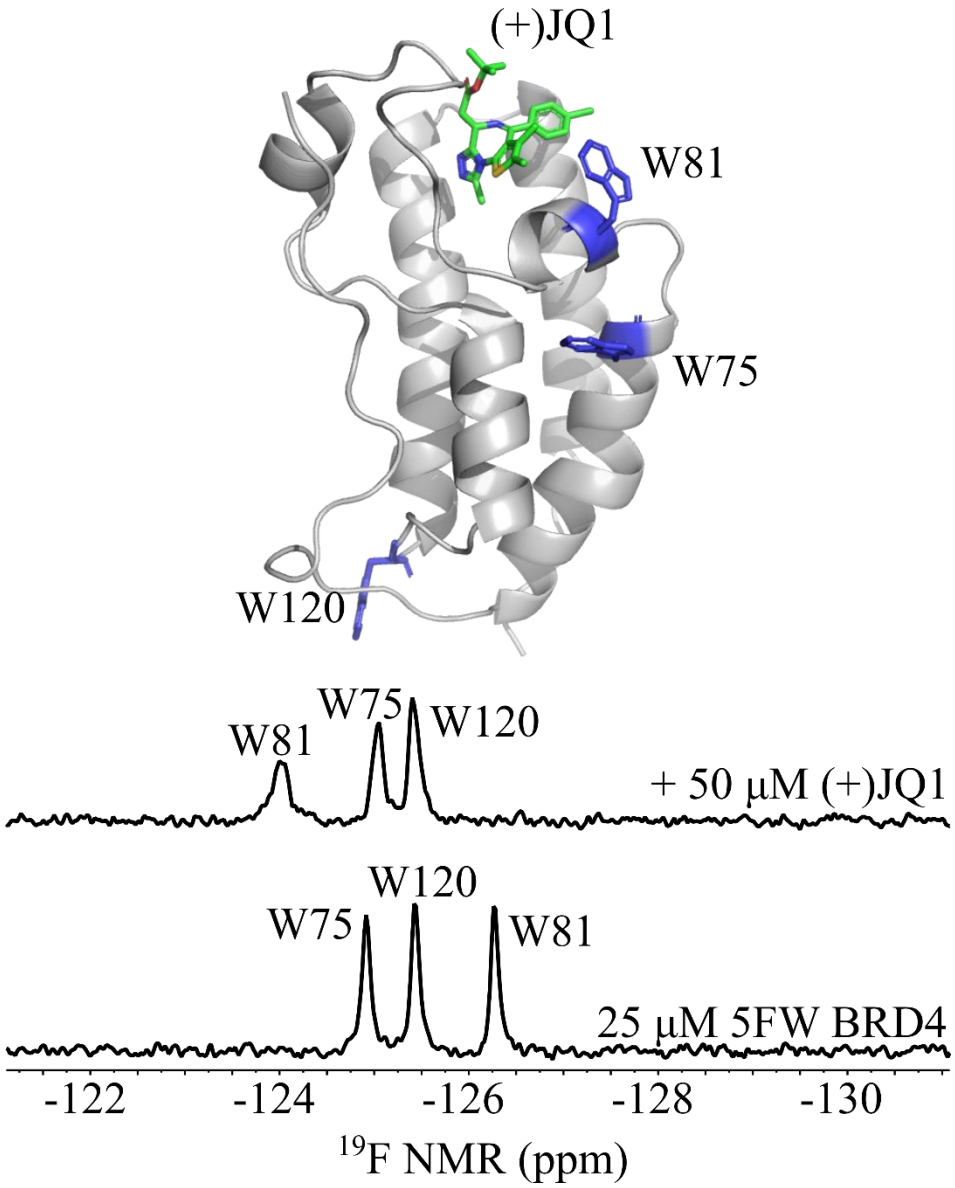


Figure 1-17: Top: Crystal structure of BRD4 bound by (+)-JQ1[74] (PDB: 3MXF) with tryptophan residues labeled. Bottom: ^{19}F NMR of 5FW BRD4 with and without (+)-JQ1. W81, nearest to the binding site, is perturbed significantly, while the distal W120 is unaffected. Figure adapted with permission[96] Copyright 2014 American Chemical Society. https://pubs.acs.org/doi/abs/10.1021/cb5007344. Further permissions related to this material should be directed to the ACS

1.14 Small molecules as protein-protein interaction inhibitors

Protein-protein interactions (PPI) lie at the heart of the study of epigenetics. Enzymes interact with histone proteins to write and erase PTMs , and various effector

domains bind to these proteins influencing gene expression and cell fate. There are an estimated 650,000 PPI, and near 23,000 known proteins encoded in the human genome.[104] The historical description of PPI interfaces is that they are large, shallow, and featureless, making them extremely challenging to target with small molecule drugs.[105] This is likely true of many of these PPI however there are often key residues at these interfaces, termed "hot spots," that play a key role in the binding of two proteins.[106] If these hot spots are mutated into other residues, binding may be lost. Other PPI indeed do have large hydrophobic cavities at their interface which may lend themselves to development of suitable inhibitors. Bromodomains, as previously mentioned, have both key hot spot residues conserved throughout their protein class as well as hydrophobic cavities, indicating that small molecule inhibition can be a suitable choice for drug development. Bromosporine was developed as a pan-bromodomain inhibitor, as well as (+)-JQ1 and I-BET-762, pan BET bromodomain inhibitors, illustrates the promise for small molecule inhibitors of bromodomains as the latter two have been utilized in clinical trials for cancer treatment (Figure 1-13).

1.15 AU1 discovery as first BPTF bromodomain inhibitor

In 2015, the Pomerantz lab disclosed a dual screen of 5FW BPTF, as well as another bromodomain from the BET class, the first bromodomain of BRD4.[97] This screen involved using 229 small molecules from the published kinase inhibitor sets obtained from Glaxo-Smith-Kline. PrOF NMR was utilized to analyze the spectra from the bromodomain of 5FW BPTF (one ^{19}F resonance) and the first bromodomain of 5FW BRD4 (three ^{19}F resonances) concurrently and observe interactions of mixtures of ligands that showed selective binding to either bromodomain. From this screen, the first selective bromodomain inhibitor of BPTF over a much more studied bromodomain of BRD4 was uncovered (Figure 1-18). This inhibitor, termed AU1 (K_d = 2.8 μM, ITC, Figure 1-19) is the focus of Chapter 3 of this dissertation.

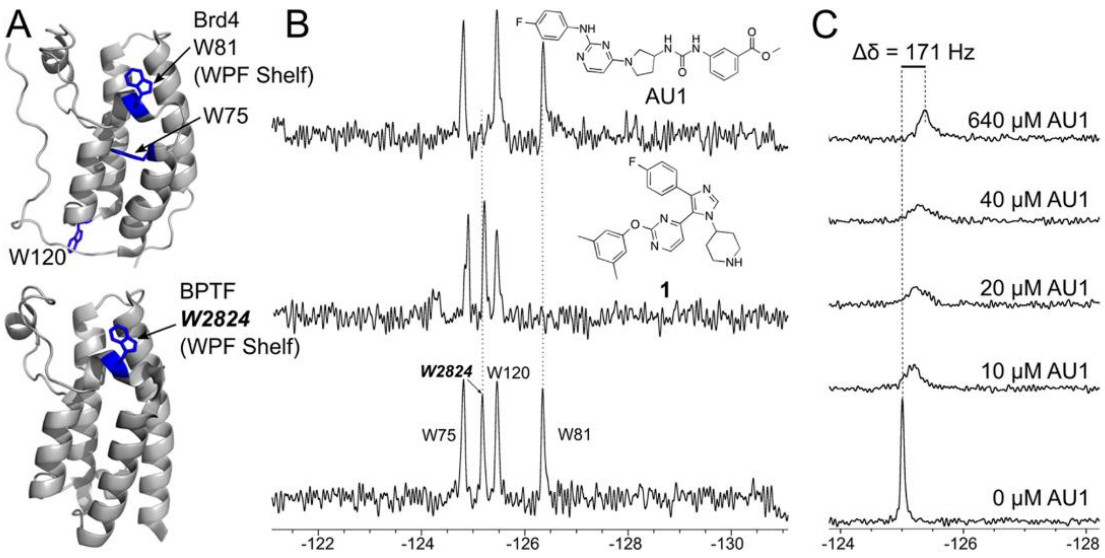


Figure 1-18: (A) Tryptophan at the WPF shelf is in close proximity to the binding site in both Brd4 and BPTF. (B) A sample of NMR spectra from the screen with the second spectrum illustrating a selective Brd4 binder (only the W81 resonance is perturbed), and the third spectrum illustrating a selective BPTF binder (only the W2824 resonance is perturbed). (C) A titration of BPTF binder AU1 into a solution of 50 μM 5FW-BPTF. Consistent with intermediate exchange kinetics, the BPTF resonance broadens into baseline and then begins to sharpen at saturating ligand concentrations that reach the solubility limit of AU1. Figure reprinted with permission[97] Copyright 2015 American Chemical Society.

Activity of AU1 in live mammalian cells has been studied in BPTF-dependent experiments. A firefly luciferase reporter assay was developed in collaboration with Prof. Camila dos Santos (Cold Spring Harbor, NY), where a plasmid containing a firefly luciferase gene coupled to a BPTF-regulatory region called MOBKL3 was transfected into Eph4 cells, along with a firefly luciferase gene in the absence of MOBKL3. The two populations of cells were treated with 5 μM of AU1 for 12 hours, and the cells containing MOBKL3 showed approximately four-fold less luciferase activity indicating the inhibition of BPTF prevents MOBKL3 from regulating luciferase activity (Figure 1-19).[97]

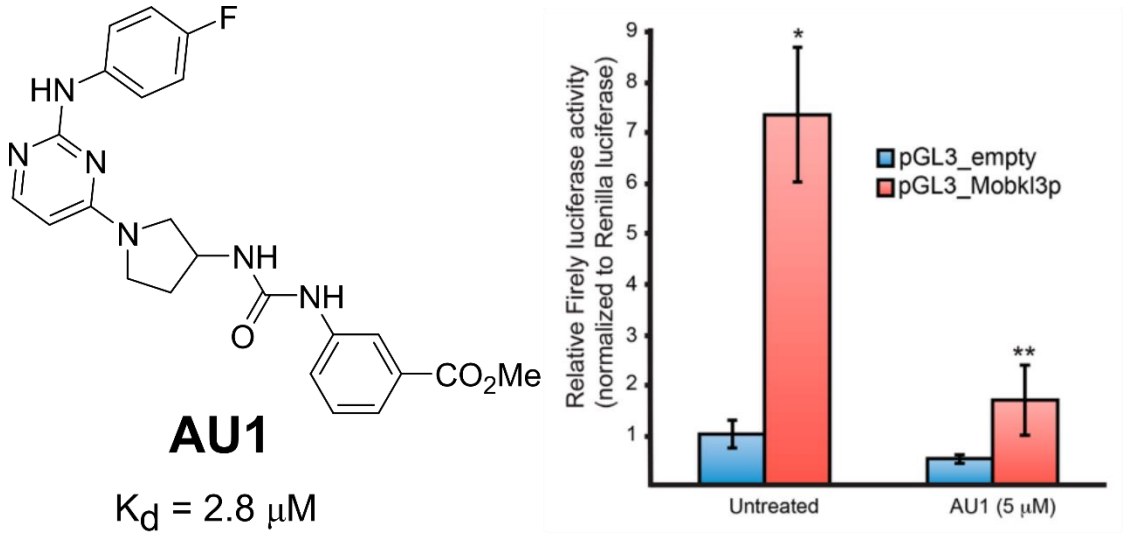


Figure 1-18: Left) Structure of AU1. Dissociation constant determined via isothermal titration calorimetry. Right) Firefly luciferase assay of Eph4 cells with and without treatment of AU1. Figure adapted with permission[97] Copyright 2015 American Chemical Society.

In a separate study in collaboration with the dos Santos lab, it was shown that the (*S*)-enantiomer AU1 reduced the capability for Eph4 cells to proliferate, and led to G1 cell cycle arrest.[94] In addition, after treatment of Eph4 cells with (*S*)-AU1, there was a reduced level of DNA occupancy of c-MYC. The reduction of c-MYC occupancy is an encouraging result because prior studies by Richart et al demonstrated that c-MYC relies on BPTF for association to DNA.[84] Lastly, (*S*)-AU1 treated cells also showed a decrease in chromatin accessibility at promoter sites, which agrees with BPTF deficiency demonstrated in other cell lines.

1.16 Preface to this Dissertation

The following chapters describe my work related to both the development of a ^{19}F MRI agent as well as a small molecule inhibitor for the bromodomain of BPTF. Chapter 2 describes my published work in the design, synthesis, and analyses of a peptide-based ^{19}F MRI agent. Chapter 3 contains a description of progress toward a second generation ^{19}F MRI agent based on our published scaffold. Chapter 4 involves the study of AU1 and development of structural analogs to target the bromodomain of BPTF. Chapter 5 discusses a collaboration with the lab of Prof. Ratmir Derda, using our ^{19}F NMR assays to

characterize fluorinated peptides with binding efficacy to human serum albumin to prolong in vivo half-life of circulating therapeutics.

Chapter 2

Synthesis of Intrinsically Disordered Fluorinated Peptides for Modular Design of High-Signal ^{19}F MRI Agents

Motivation: The purpose of this work was to develop a new class of ^{19}F MRI agents for medical diagnostics. The current state of the art for ^{19}F MRI involves the use of perfluorinated emulsions which suffer from stability issues as well as accumulation both in vivo and in the environment. This study describes our work in developing a high signal ^{19}F MRI contrast agent based on a water-soluble scaffold that is readily degradable by enzymes to prevent bioaccumulation.

2.1 Introduction

Molecular and cellular therapies have the potential to treat numerous diseases, but there is a significant need for noninvasive techniques to monitor the biodistribution of these therapeutics in vivo. The trace amounts of organic fluorine in biological settings enables monitoring of fluorine-tagged molecules in whole organisms by magnetic resonance imaging (MRI).[107] MRI using the ^{19}F nucleus holds great potential for diagnostic molecular imaging with applications for attaching fluorinated molecules to targeting domains,[108] cell tracking,[109] and oxygen sensing.[110] ^{19}F MRI is less developed than ^{1}H MRI, in part due to the need for high signal, biocompatible ^{19}F probe molecules.[111] Synthesis of high-symmetry fluorochemicals, such as perfluorocrownethers,[112] fluoropolymers,[113] and fluorinated dendrimers[114] (e.g., Figure 2-1) leads to an increase the number of magnetically equivalent fluorine nuclei for improving signal intensity. Emulsions of these fluorochemicals represents the current state-of-the-art for ^{19}F MRI probes which are reaching the limit of detection threshold at clinical field strengths.[113,108] However, these emulsions are susceptible to stability problems, poor physicochemical properties, and

retention in the body.[3,115] Environmental persistence and health concerns are also prompting legal regulations of some perfluorochemicals and their degradation products.[116]

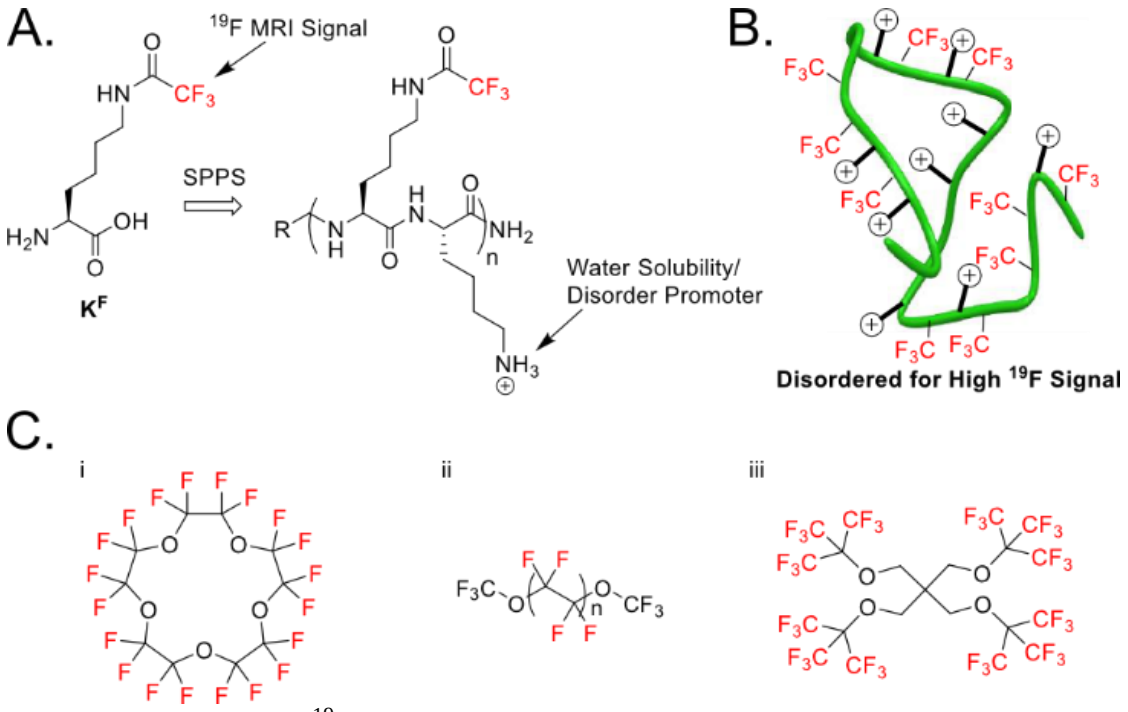


Figure 2-1: Disordered ^{19}F peptide design constructed by solid-phase peptide synthesis (SPPS) and current state-of-the-art molecules. A) Generic structure of disordered ^{19}F peptides. A complete list of peptides used in this study can be found in Figure S1 B) Disordered peptide design showing ^{19}F sources alternating with positive charge. C) Select fluorinated molecules used for ^{19}F MRI: i) Perfluoro-15-crown-5 ether[112] ii) Perfluoro(polyethylene glycol) methyl ether[117] and iii) PERFECTA.[118]

Fluorinated biopolymers are an untapped material for ^{19}F MRI applications. Fluorinated proteins characteristically lead to well-resolved fluorine resonances for which the same fluorine-labeled amino acids can produce resonances that span over 10 ppm.[119,] Such dispersion is due to the high sensitivity of the fluorine chemical shift to distinct protein environments.[120] To overcome this, we were inspired from spectral studies of intrinsically disordered proteins that result in limited resonance dispersion.[121] Therefore, our current approach seeks to override the high environmental sensitivity of the ^{19}F nucleus through synthesis of highly disordered peptide chains lacking secondary or tertiary structure. Here, we describe the design of a series of high-signal ^{19}F MRI agents for future

molecular imaging and cell tracking applications using sequence-defined peptide-based materials that are emulsion-free while possessing tunable biological and environmental stability. As an initial evaluative step towards future molecular imaging applications, we show that the fluorinated polypeptide maintains both a narrow linewidth and high ^{19}F magnetic resonance signal after protein bioconjugation.

Our initial designs focused on the solid-phase synthesis of peptide oligomers containing an alternating pattern of N-ε-trifluoroacetylated lysine (TFA-lysine) and lysine. These amino acids were chosen to favor high solubility, while disfavoring both aggregation as well as secondary structure formation due to electrostatic repulsion (Figure 2-1a, b). TFA-lysine, was chosen based on three magnetically equivalent fluorine nuclei yielding high intensity narrow resonances that are devoid of scalar couplings. Ye et al. also showed that the chemical shift sensitivity of CF_3 groups next to carbonyls are significantly less responsive to environmental changes relative to benzylic CF_3 groups.[122] Regarding faster spectral acquisition time, the T_1 relaxation of CF_3 groups tend to be shorter than those found on aromatic rings.[123] Grage et al. showed by solid state ^{19}F NMR that aliphatic CF_3 groups have T_1 relaxation times between 0.32 and 1.86 s, whereas aryl fluorines had T_1 times ranging from 14.9 to 352 s.[123,124] As a final consideration, the trifluoroacetamide group can also hydrolyze under alkaline conditions for eventual degradation in the environment.

2.2 Results and Discussion

Our first peptide series consisted of a 17 residue peptide **1** containing seven TFA-lysines (Figure 2-2). Encouragingly, when the peptide was dissolved in aqueous solution and studied by ^{19}F NMR, a single resonance was observed with a slight shoulder, indicating that the TFA-lysine residues were experiencing a similar chemical environment. The resonance spanned 16 Hz, with a 7 Hz measurement at half-height. Surprisingly, when the N-terminus was left unacetylated a slight decrease in resonance degeneracy was observed as indicated by the appearance of a set of non-overlapping resonances. We therefore focused on peptides of alternating TFA-lysine and lysine with acylated N-termini.

In order to discern the secondary structure of the peptide sequence and support our disordered peptide designs, we performed circular dichroism (CD) on **1** in both aqueous

and organic media (Figure 2-2). The far-UV CD spectra of **1** display a strong solvent dependence. In methanol, a structure-promoting solvent, an α-helical signature was observed with double minima at 208 and 222 nm. This peptide was determined to be 55% α-helical based on the molar ellipticity at 222 nm. Alternatively, when **1** was dissolved in water, a CD spectrum consistent with a random-coil was observed. Because of the low resolution of CD, we cannot rule out additional population of a polyproline II helix (PPII) based on the small positive absorption at 215 nm, which has been observed for polylysine containing peptides in the PPII conformation[125]. We measured ^{1}H NMR $^{3}J_{HN C\alpha}$ coupling constants from the backbone amides of **1** which range from 6.11-6.80 Hz (Figure 2-3). These values are within the range to slightly below the random coil values calculated by Serrano[126] and higher than 5.16 Hz determined by Shi et al for the PPII helix.[127]

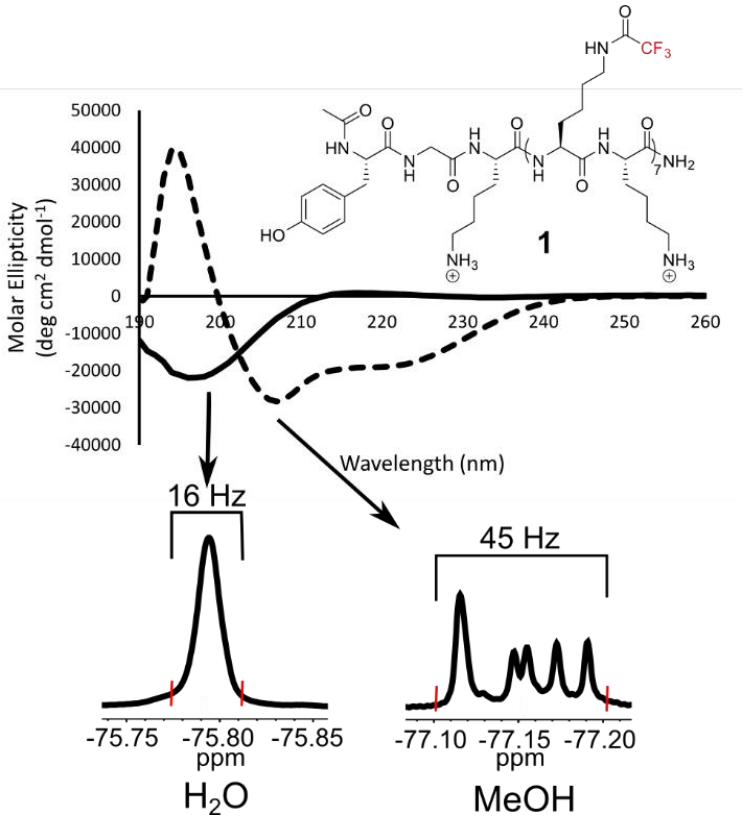


Figure 2-2. Top) Far-UV CD spectra of 100 μM **1** in water (solid lines) and methanol (dashed lines). Bottom) A random-coil is observed in aqueous environment resulting in a degenerate ^{19}F NMR resonance spanning 16 Hz. In methanol, an α-helix is observed resulting in resolved ^{19}F NMR resonances spanning 45 Hz. Linewidths were measured from baseline to baseline at the red hatch marks for comparison.

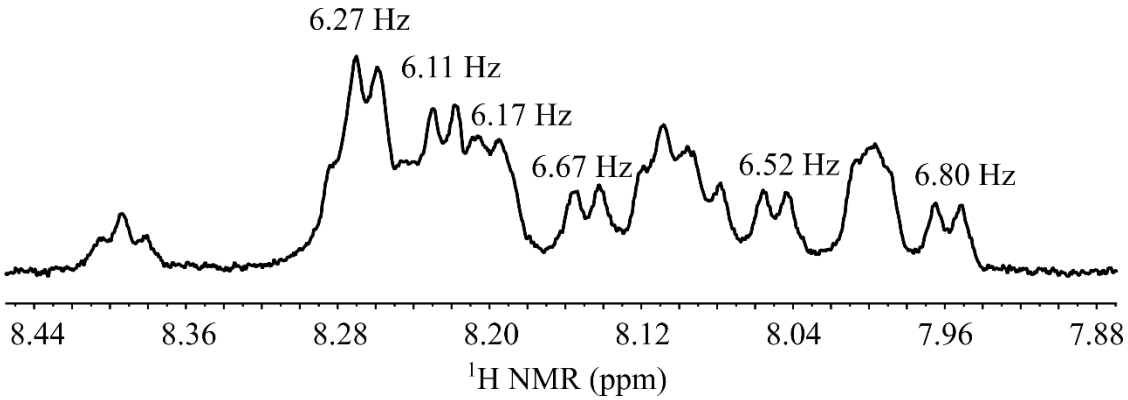


Figure 2-3: ^{1}H NMR (500 MHz, 90% H_2O/10% D_2O, unreferenced spectrum) of **1**, zoomed in on amide protons. Several backbone amide doublets are labeled with their J3 coupling constants, showing a range from 6.11 Hz to 6.80 Hz.

Comparing the ^{19}F NMR spectra from these two solutions also supports the structural design. Whereas all ^{19}F resonances remain degenerate when the peptide is dissolved in water, when methanol is used, the ^{19}F resonances are more resolved, consistent with the hyperresponsive character of the ^{19}F nucleus to distinct chemical environments.[120] This observation is indicative of the ^{19}F-labeled side chains being in distinct environments within the α-helix. (Figure 2-2, Figure 2-4).

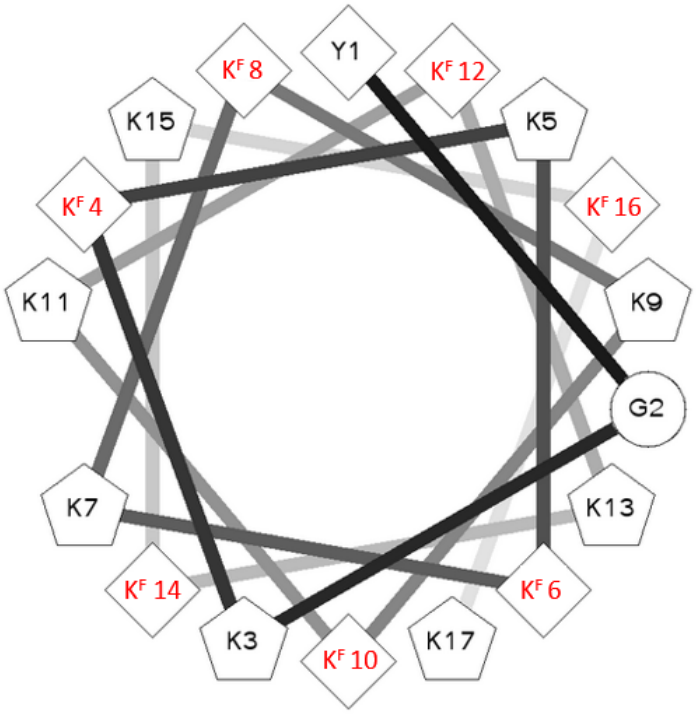
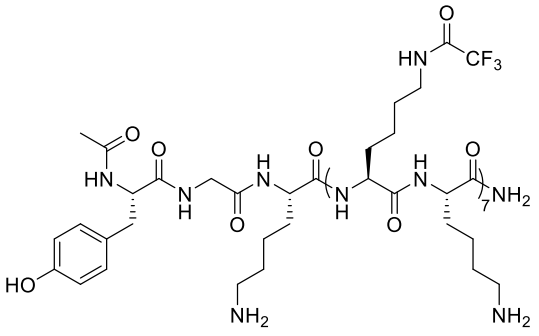
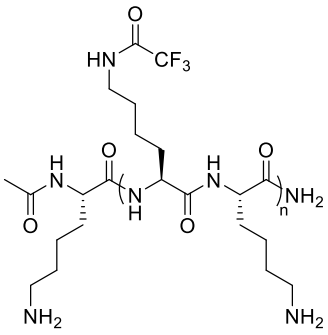


Figure 2-4: α-helical representation of **1** dissolved in methanol. Trifluoroacetyl lysine (K^F) are in distinct environments within the helical structure leading to resolved ^{19}F NMR resonances. Figure was constructed using the helical wheel generator found here: http://rzlab.ucr.edu/scripts/wheel/wheel.cgi

Peptides tend to adopt increased α-helical structure at long chain lengths. To evaluate a length dependence on our designs, peptides **S1a-i** (Figure 2-5) were synthesized containing 5-21 amino acids and up to ten TFA-lysine residues. In many cases a similar degeneracy was observed in the NMR spectra with a slight shoulder in some cases (Figures 2-17 – 2-34). Therefore, even up to 30 equivalent fluorine resonances can easily be obtained using our initial designs. Typically, when symmetrical molecules such as perfluorocrownethers or dendrimers are used to maximize the number of magnetically equivalent ^{19}F resonances, synthesis of these large molecules can become challenging.[128] Additionally, these molecules have a dramatic loss of water solubility.[3] Our peptide constructs however can be readily obtained using solid phase synthesis and readily dissolve in water at millimolar concentrations.

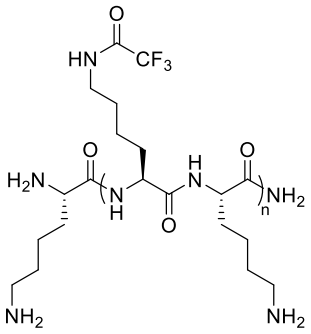


1



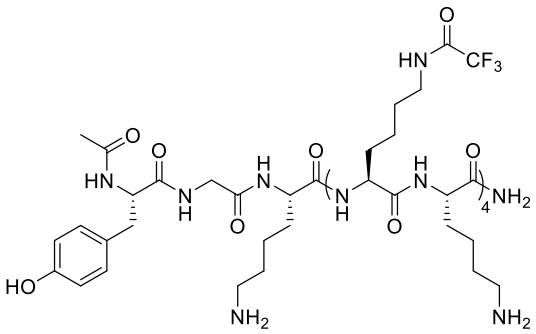
S1a-S1i

S1a: n=2 **S1d:** n=5 **S1g:** n=8
S1b: n=3 **S1e:** n=6 **S1h:** n=9
S1c: n=4 **S1f:** n=7 **S1i:** n=10



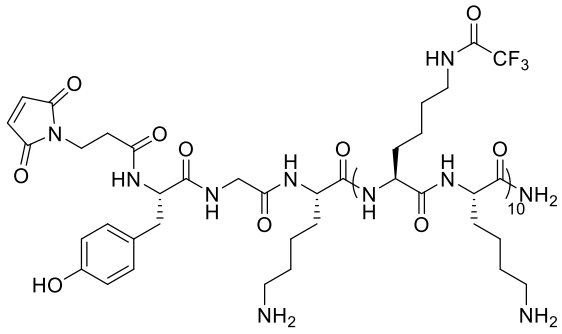
S2a-S2f

Sa: n=2 **Sd:** n=6
Sb: n=4 **Se:** n=7
Sc: n=5 **Sf:** n=8



2a-2c

2a = Ac-YGKK^FKK^FKK^FKK^FK-NH_2
2b = Ac-YGkK^FKK^FkK^FKK^Fk-NH_2
2c = Ac-YGkK^FkK^FkK^FkK^Fk-NH_2
Where K^F = N-ε-trifluoroacetyl lysine
and k = D-Lysine



3

Figure 2-5: Structures of all peptides utilized in this work.

Sequential addition of TFA-lysine to the peptide sequence only minimally affects the hydrophobicity of the peptide chain which we assessed based on reverse-phase HPLC retention time using acetonitrile as the elution solvent, in which case peptides **S1a-i** differ in retention by only 11 min (23-34% acetonitrile). We conclude from these first studies that a disordered peptide sequence that dissolves readily in water, can be maintained at long chain lengths by simple alternation of TFA-lysine and lysine, giving rise to high intensity degenerate ^{19}F NMR spectra.

In addition to a signal sensitivity requirement, a second priority for the design of our fluorinated peptides is the ability to tune their degradation susceptibility in vivo to avoid persistence within the body and the environment. To probe this, we subjected our lysine-rich peptides to trypsin to monitor the rate of degradation. An 11-residue model peptide, **2**, was used for these experiments, which was then modified via incorporation of D-lysine to tune the stability of the molecule. As we expected a 100% D-amino acid containing peptide to be proteolytically stable,[129] we investigated decreased percentages of D-amino acids to allow us to assess a range of stabilities from these diastereomeric peptides, as well as any protection from the fluorinated side-chain.

We monitored the hydrolysis of three peptides over time. In all cases presented, the L-variant of TFA-lysine was used. The all L-lysine containing peptide **2a** was degraded rapidly with complete hydrolysis of the peptide backbone observed approximately one minute after addition of trypsin at 25 °C, and pH 8.1. Alternatively, peptide **2b**, containing 30% D-amino acids, showed a 10-fold increase in protease stability. The final peptide studied contained 50% D-amino acids, **2c**, and showed 72% remaining after 3 hours. After 24 h, 12% of the peptide remained, which matches approximately the behavior of the peptide in buffer alone, indicating the D-lysines provided sufficient protection from trypsin until the hydrolysis of TFA-protecting groups was observed. (Table 2-1, Figures 2-6 – 2-8).

Table 2-1: Lifetime of model peptide series 2 during tryptic digest

Peptide	Sequence[a]	Lifetime
2a	Ac-YGKK^FKK^FKK^FKK^FK-NH_2	1 min
2b	Ac-YGkK^FKK^FkK^FKK^Fk-NH_2	10 min
2c	Ac-YGkK^FkK^FkK^FkK^Fk-NH_2	> 24 h[b]
2b pH = 8.1, buffer only	Ac-YGkK^FKK^FkK^FKK^Fk-NH_2	> 24 h[c]
2b pH 7.1, buffer only	Ac-YGkK^FKK^FkK^FKK^Fk-NH_2	> 48 h[d]

[a] K^F designates N-ε-trifluoroacetyllysine, and lower-case k refers to D-lysine.

[b] 12% of 2c remains after 24 h.

[c] 11% of **2b** remains after 24 h in pH 8.1 buffer without trypsin.

[d] 10% of **2b** remains after 48 h in pH 7.1 buffer without trypsin.

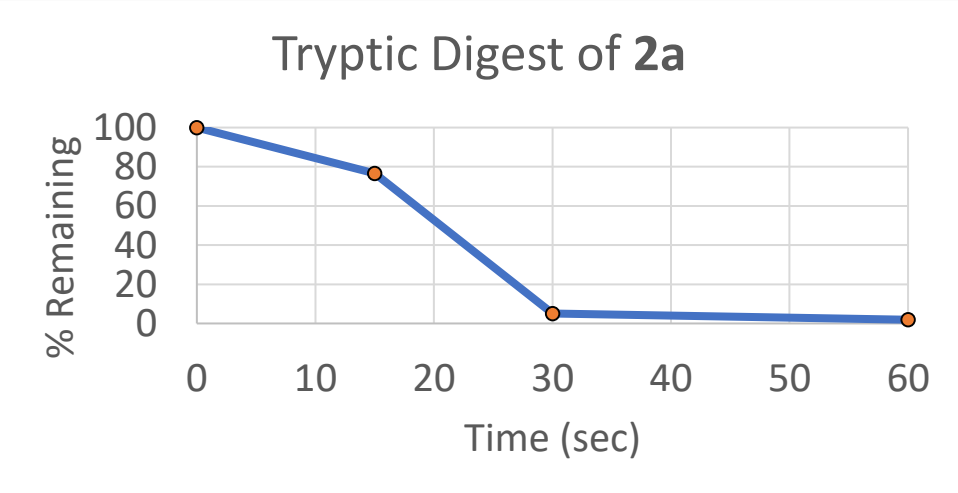


Figure 2-6: Tryptic digest of **2a**, 20:1 substrate to enzyme ratio. The digest was performed in 0.1 M NH_4HCO_3, pH 8.1, ambient temperature.

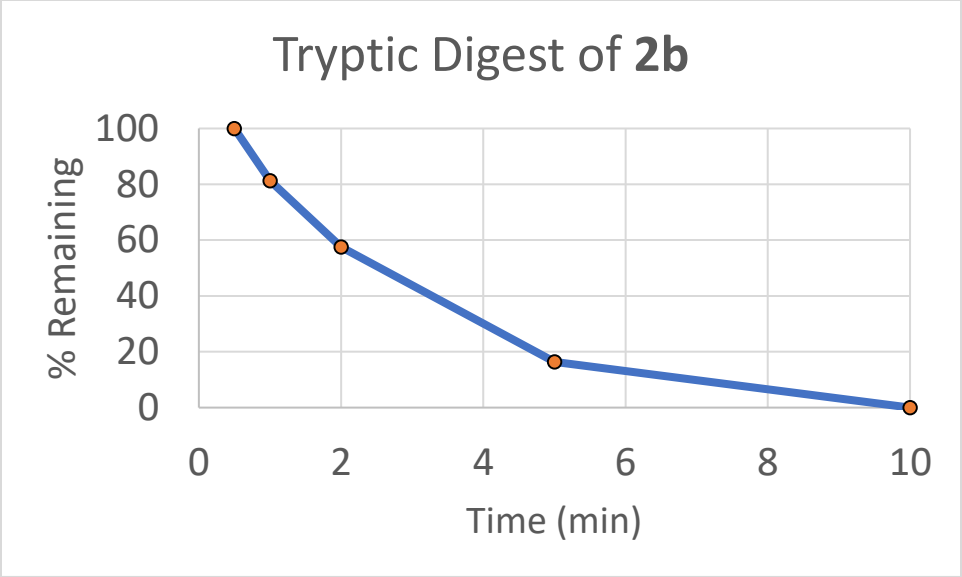


Figure 2-7: Tryptic digest of **2b**, 20:1 substrate to enzyme ratio. The digest was performed in 0.1 M NH_4HCO_3, pH 8.1, ambient temperature.

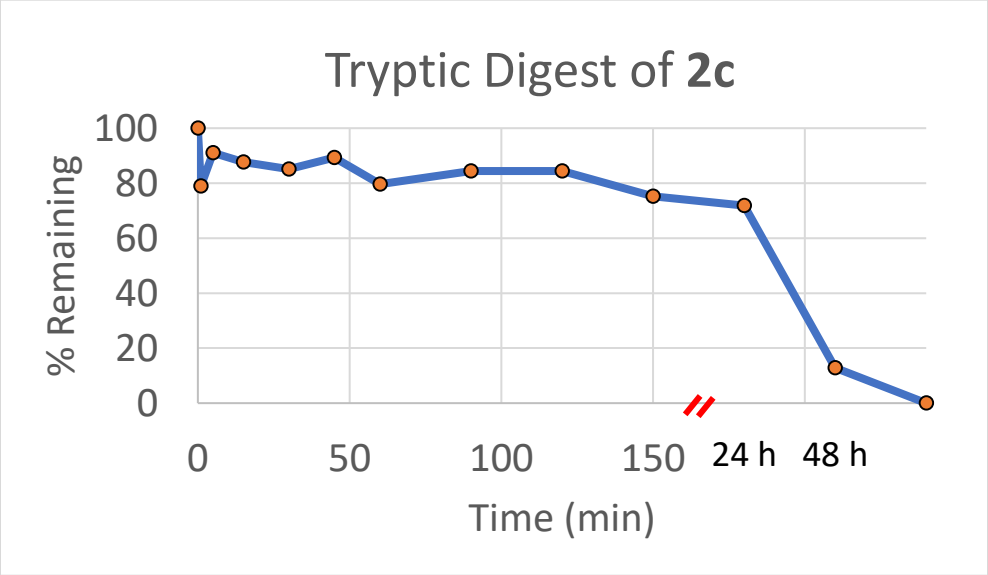


Figure 2-8: Tryptic digest of **2c**, 20:1 substrate to enzyme ratio. The digest was performed in 0.1 M NH_4HCO_3, pH 8.1, ambient temperature. Note change in time units for last two data points.

We tested the buffer conditions of the trypsin assay on the peptide to ensure it would not appreciably interfere with the digest, and found that out to three hours there is an 11% reduction of full length peptide, but near complete hydrolysis of the TFA side chains has occurred by 24 hours. This experiment was also performed at pH 7.1, where the peptide showed no hydrolysis up to five hours, but near complete hydrolysis after 48 hours (Table 2-1, Figures 2-9 and 2-10). Although more stable, ^{19}F NMR spectra in water for peptides **2b** and **2c** showed two resonances, indicating a change in structure (Figures 2-50 and 2-52). Further analysis and design of peptides containing partial incorporation of D amino acids must be performed to reestablish degenerate signals.

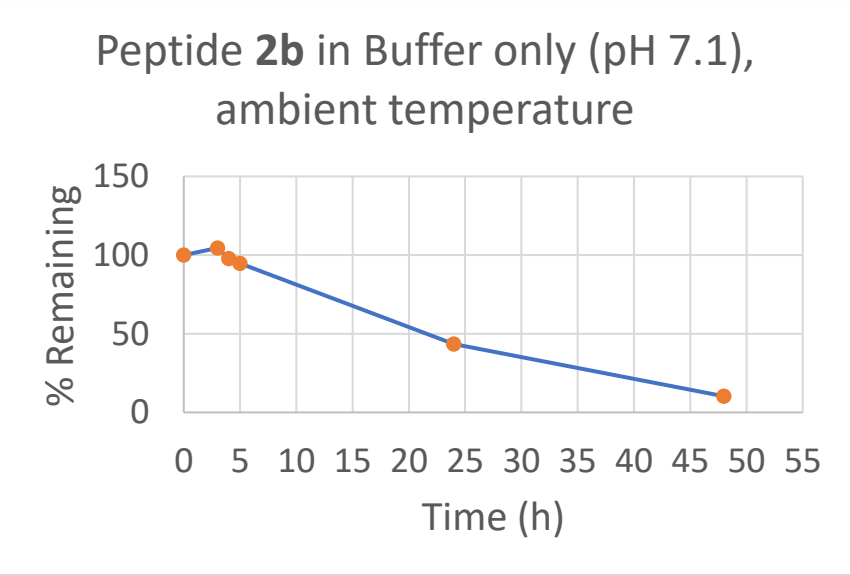


Figure 2-9: Peptide 2b subjected to buffer only (0.1 M NH_4HCO_3, pH 7.1). Percent remaining of parent peptide plotted vs time using HPLC integration

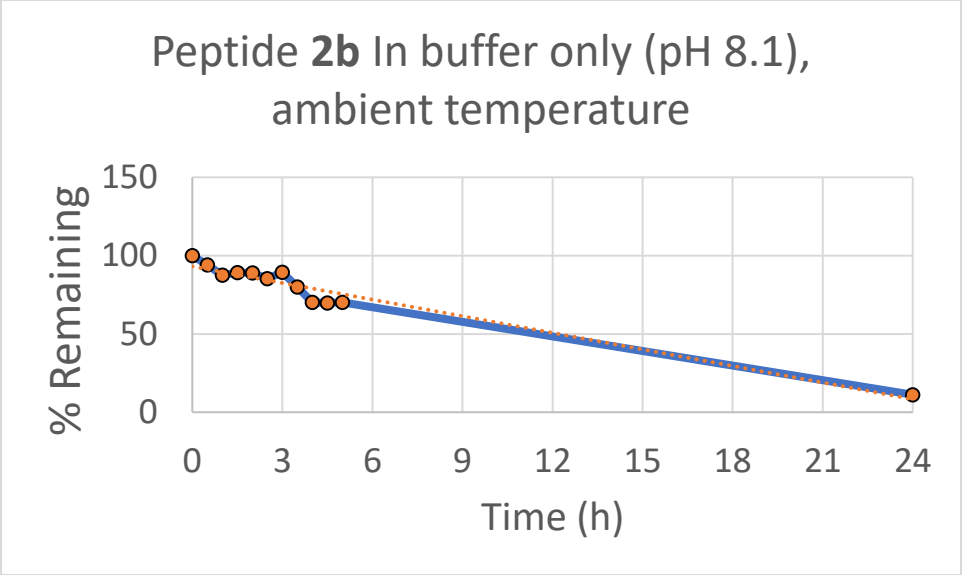


Figure 2-10: Peptide 2b subjected to buffer only (0.1 M NH_4HCO_3, pH 8.1). Percent remaining of parent peptide plotted vs time using HPLC integration

While peptide **2a** maintained overlapping resonances in the ^{19}F NMR spectra in water, in biological systems, local chemical environments can affect chemical shift dispersion. Therefore, despite the instability of **2a** towards proteases, we also measured the ^{19}F NMR spectra of **2a** in an *E. coli* lysate (Figure 2-11). In this case, although the resonance moved upfield by 0.07 ppm, the overlap was maintained. An upfield tail from the ^{19}F resonances was observed from both the **2a** resonances and the reference TFA resonance consistent with imperfect shimming. In the future, to further mitigate image resolution loss, (i.e, blurring) for MRI applications, pulse sequences that encode spatial information such as gradient echo and spin echo MRI can be used.

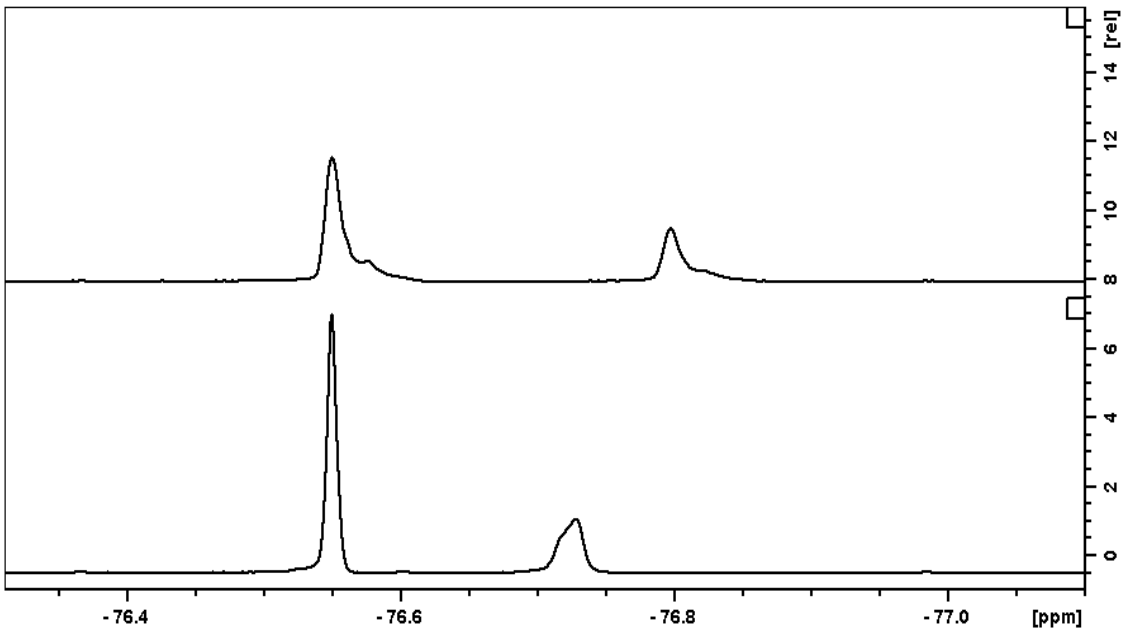
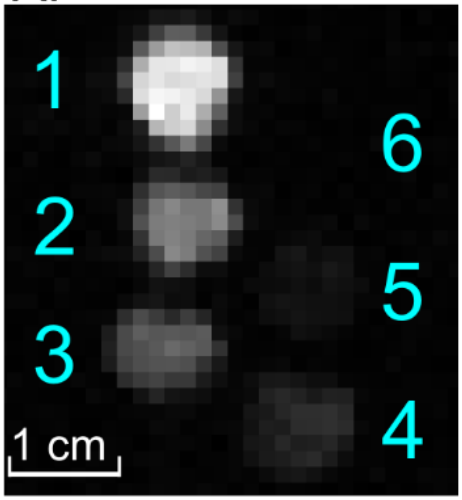


Figure 2-11: ^{19}F NMR spectrum of 2a in buffer alone (50 mM PO_4^-, 300 mM NaCl, pH 7.4, bottom spectrum) and in *E. coli* cellular lysate in the same buffer (top spectrum). The upfield resonance represents **2a**, which shows a movement indicating an interaction with the cellular lysate. Tailing in both resonances is indicative of poor shimming.

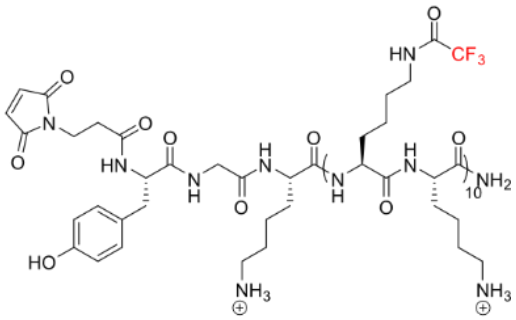
With a tunable peptide system in hand, we turned our attention to peptide **3** containing 30 equivalent fluorine atoms for ^{19}F MRI studies. To assess the limit of detection, a ^{19}F MRI phantom image was acquired of **3** at varying concentrations to evaluate the image quality (Figure 2-12). A 16.4 T Varian MRI instrument was used to scan five samples of **3** with ^{19}F concentrations ranging from 4-64 mM equivalent fluorine (or 0.13 – 2.13 mM peptide). Images were acquired in ten minutes, and a linear dependent intensity gradient can be observed throughout the samples with the lowest concentration still being visible at 4 mM fluorine (0.13 mM peptide) (Figure 2-12D). Highly fluorinated materials tend to aggregate at high concentrations, which is reflected in short T_2 relaxation and results in broadened resonances. Therefore, we also measured the T_1 and T_2* values of the varying peptide concentrations by ^{19}F NMR (470 MHz, 25 °C). We used T_2* as an estimation of T_2 due to the importance of the peak width in the resolution of our spectra. We do not observe a significant concentration effect on T_1 or T_2* consistent with the

narrow line widths (5.9-6.4 Hz) and linear concentration dependence on signal to noise (Figure 2-12C,D) indicating that these peptides do not aggregate in solution.

A.



B.



3

C.

Sample #	[^{19}F]	T_1	T_2*
	(mM)	(ms)	(ms)
1	64	834.0	53.8
2	32	842.5	49.5
3	16	835.1	50.6
4	8	837.8	51.4
5	4	828.0	50.0

D.

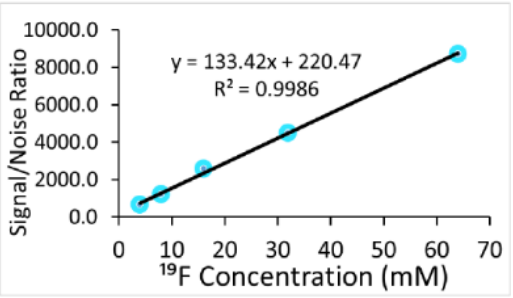


Figure 2-12: A) ^{19}F MRI phantom image of **3** at various concentrations: 1) 2.13 mM (64 mM ^{19}F), 2) 1.07 mM (32 mM ^{19}F), 3) 0.53 mM (16 mM ^{19}F), 4) 0.27 mM (8 mM ^{19}F), 5) 0.13 mM (4 mM ^{19}F), 6) 0 mM. B) Structure of **3** utilized in MRI experiments. C) NMR relaxation times determined for **3** showing no appreciable change as a function of concentration. D) Signal to noise ratio of **3** increases linearly with respect to concentration, indicating no concentration dependent higher order structure.

One potential application of this type of peptide construct is for use as a ^{19}F MRI imaging tag conjugated to a larger targeting protein. To investigate if the newly introduced chemical environment of a bioconjugated protein with the peptide affects its NMR resonances, we used a model protein, bovine serum albumin (BSA), to react via the sulfhydryl moiety of its lone free cysteine residue to a maleimide attached to the N-

terminus of peptide **3**. The bioconjugation reaction was performed by incubating 50 μM BSA with 100 μM **3** for 24 h at 37 °C (0.2 M acetate buffer, pH = 5.0). Protein adduct formation was verified by MALDI-TOF MS (Figure 2-14). After removal of unreacted peptide via dialysis, a ^{19}F NMR spectrum was acquired on the protein-peptide conjugate (Figure 2-13). The resulting spectrum showed a singular resonance with a downfield shoulder for the construct. Although still narrow, the resonance of the conjugate was broadened (11 Hz) in comparison to the free peptide (6.2 Hz), consistent with the attachment of a large protein (66 kDa). The T_2^* of the conjugate was approximately 40% shorter than the peptide alone (29 vs 50 ms). Such a small reduction in T_2 suggests that despite being attached to a large protein, resonance broadening from chemical shift anisotropy is mitigated due to significant rotational freedom of the CF_3 group, and the fluorinated peptide side-chains maintain a significant amount of conformational dynamics, similar to the behavior of intrinsically disordered proteins.[130]

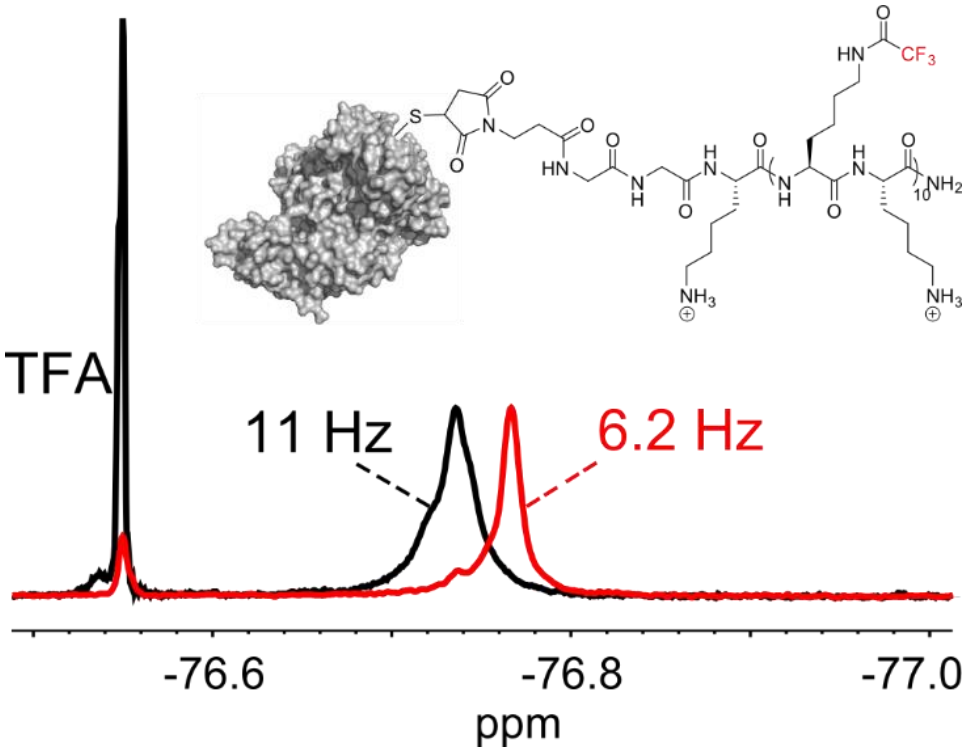


Figure 2-13. ^{19}F NMR of BSA conjugated to **3** (black line). The linewidths are compared to unconjugated **3** (red line), and were measured at the dashed line. Trifluoroacetic acid was added as a reference standard.

These results are encouraging for future molecular imaging applications, in which such peptides could be conjugated to larger antibodies, or smaller targeting agents such as affibodies[131] and fibronectin domains.[48] Although, we note a single peptide possessing 30 equivalent fluorine atoms may be insufficient for imaging applications at clinical field strengths (≤ 3 T), the possibility of attaching multiple fluorinated peptides to a protein construct such as to a tetracysteine tag, or the surface of a nanoparticle opens up the possibility of highly fluorinated imaging constructs.[132]

2.3 Conclusion

In conclusion, these studies lay the groundwork for a molecular design strategy for obtaining high signal ^{19}F MRI agents, based on introducing a high degree of structural disorder. Through careful choice of a peptide backbone and fluorinated side-chain, the stability of the fluorochemical can be rationally tuned for breakdown in vivo or in the environment. Due to the increased interest in new fluorochemicals for imaging applications, and concerns over toxicity and environmental persistence, these new designs will be studied for broader use in ^{19}F MRI-based biomedical applications. We note, that with the increased availability of synthetic methods for more highly fluorinated amino acids,[133,134] even higher signal may be achievable. Paramagnetic T_1 agents,[135] as well as improvement in data processing such as sparse sampling,[136] may help bring designs of fluorinated materials closer to practical applications in the ^{19}F MRI field.

Acknowledgements

This study has been funded via the University of Minnesota McKnight Land Grant Professorship, Research Corporation Cottrell Scholar Award, Heisig-Gleysteen Fellowship, and University of Minnesota Undergraduate Research Scholarship (URS). The MRI work was supported in part by the Minnesota Lions Diabetes Foundation, the Schott Family Foundation, the Carol Olson Memorial Diabetes Research Fund, and the NIH [grants P41 EB015894 and S10 RR025031].

2.4 Experimental

Materials

Solid-phase peptide synthesis was performed using the procedure outlined below. 1-hydroxybenzotriazole monohydrate (HOBT, Chem-Impex Int'l), *O*-benzotriazol-1-yl-*N*, *N*, *N*', *N*'-Tetramethyluronium hexafluorophosphate (HBTU, Novabiochem), and *N*,*N*-Diisopropylethylamine (DIEA, Sigma Aldrich) were used as received without further purification and dissolved into reagent-grade DMF (Fisher Scientific). N-α-Fmoc-N-ε-Boc-L-lysine and N-α-Fmoc-N-ε-trifluoroacetyllysine were received from Chem-Impex. N-α-Fmoc-O-tBu-tyrosine, N-α-Fmoc-glycine, and N-α-Fmoc-N-ε-Boc-D-lysine were received from Novabiochem. Piperidine, trifluoroacetic acid, dichloromethane (Sigma Aldrich), Triisopropylsilane (TCI America), acetic anhydride (Fisher Scientific), 3-maleimidopropionic acid NHS (Chem impex), and triethylamine (Mallinckrodt) were all used as received.

Solid Phase Peptide Synthesis

All peptides were synthesized on NovaSyn TGR® resin using a CEM MARS Xtraction microwave reactor (25 μmol scale). Prior to acylation reactions, N^{α}-Fmoc protected amino acid (100 μmol, 4 eq) was preactivated in DMF (1000 μL), with HOBT (200 μL of 0.5 M solution in DMF, 4 eq), HBTU (200 μL of 0.5 M solution in DMF, 4 eq), and DIEA (200 μL of 1.0 M solution in DMF, 8 eq) followed by addition to the resin. The mixture was heated at 70 °C (400 W power, with a two minute ramp cycle and a four minute hold) to couple the amino acid. The reaction mixture was then drained and the resin washed (3 x DMF, 3 x DCM, 3 x DMF). 9-fluorenylmethyoxycarbonyl (Fmoc) deprotection was achieved by adding piperidine (2 mL, 20% v/v in DMF) to the resin and heating to 80 °C (400 W power, via a two minute ramp and two minute hold cycle). The resin was once again drained and washed. The coupling and deprotection steps were repeated until the full length peptide was synthesized. Following the final deprotection step, the N-terminus was either acetylated, or acylated with a maleimide-conjugated N-hydroxysuccinimide

(NHS) ester. Acetylation reactions were conducted at room temperature with magnetic stirring with 500 µL acetic anhydride, 100 µL triethylamine (TEA), and 1.4 mL DCM for 90 minutes. The maleimide-NHS N-terminal acylation was completed using 3-maleimidopropionic acid NHS (4 eq) dissolved in 1.4 mL DMF and 200 µL DIEA (1.0 M in DMF). The reaction was heated at 70 °C (400 W power, with a two minute ramp and four minute hold cycle). After the resin was drained and washed as previously described, the peptides were cleaved from the resin using 2 mL of a 95:2.5:2.5 trifluoroacetic acid (TFA): triisopropylsilane (TIS): H_2O mixture and stirred for 2 h. The resulting solution was drained from the syringe, precipitated into diethyl ether, and allowed to cool to -20 °C for 15 min. The peptides were then pelleted via centrifugation (3000 x *g*, 3 min at 4 °C). The ether was decanted, and the pellet was dissolved in a 60:40 mixture of 0.1% TFA: acetonitrile. The dissolved peptides were purified using a Dionex Ultimate 3000 RP-HPLC using a C-18 column and a 10-60% water with 0.1% TFA and acetonitrile gradient over 60 minutes. The identity of the purified peptides were determined using an Ab-Sciex 5800 MALDI-TOF/TOF Mass Spectrometer utilizing α-cyano-4-hydroxycinaminic acid matrix.

Circular Dichroism

Circular dichroism (CD) was used to determine the secondary structure of select peptide samples. CD spectra were taken in the far-UV (190-260 nm) range using a Jasco J-815 spectropolarimeter at 25 °C. Peptide samples were taken at 100 µM concentration in either DI H_2O or methanol as a solvent in a cuvette with a 1 mm pathlength. The data was collected at a scan rate of 50 nm/min taken by averaging 10 spectra. The baseline was corrected by comparing the spectra data to the solvent sample alone. The % helicity was calculated according to the equations below. (Farood et al., *PNAS*, **1993**, *90*, 838).

$$\%\ helicity = 100\ x\ ([\theta_{222}/\theta_{222}^{max}).$$

$$\theta_{222}^{max} = -40{,}000\ \left[1 - \left(\frac{2.5}{n}\right)\right],$$

$$where\ n = \#\ of\ backbone\ amides\ in\ peptide\ sequence$$

1D ^{19}F NMR

All ^{19}F NMR spectra were obtained using a 470 MHz Bruker 500 spectrometer with a 5 mm Prodigy TCI cryoprobe. Peptide samples were dissolved in either D_2O or methanol-d4. Aqueous samples were referenced to trifluoroacetate (-76.55 ppm). Spectra were acquired using 128 scans, an acquisition time of 0.5 s, a relaxation delay time of 1 s, and a sweepwidth of 30 ppm. Protein conjugate spectra were acquired using 1500 scans, an acquisition time of 1 s, a relaxation delay time of 1.7 s, and a sweepwidth of 40 ppm.

BSA Bioconjugation

Bovine serum albumin was dissolved at a concentration of 50 μM in buffer (100 mM acetic acid, 300 mM sodium acetate, pH 5.0). Peptide **3** (2 eq) was added to the protein solution and was incubated at 37 °C for 24 h. Reaction progress was monitored using MALDI-TOF MS, samples were desalted using C4 ZipTips and eluted into a sinapinic acid matrix. The protein mixture underwent dialysis (7 KDa cutoff) to remove unreacted peptide and was subjected to the aforementioned ^{19}F NMR procedure. BSA-peptide conjugate was validated utilizing LC-MS (Calculated: 70,468.96 Da, Observed: 70,469.2031 Da) (Figure 2-14).

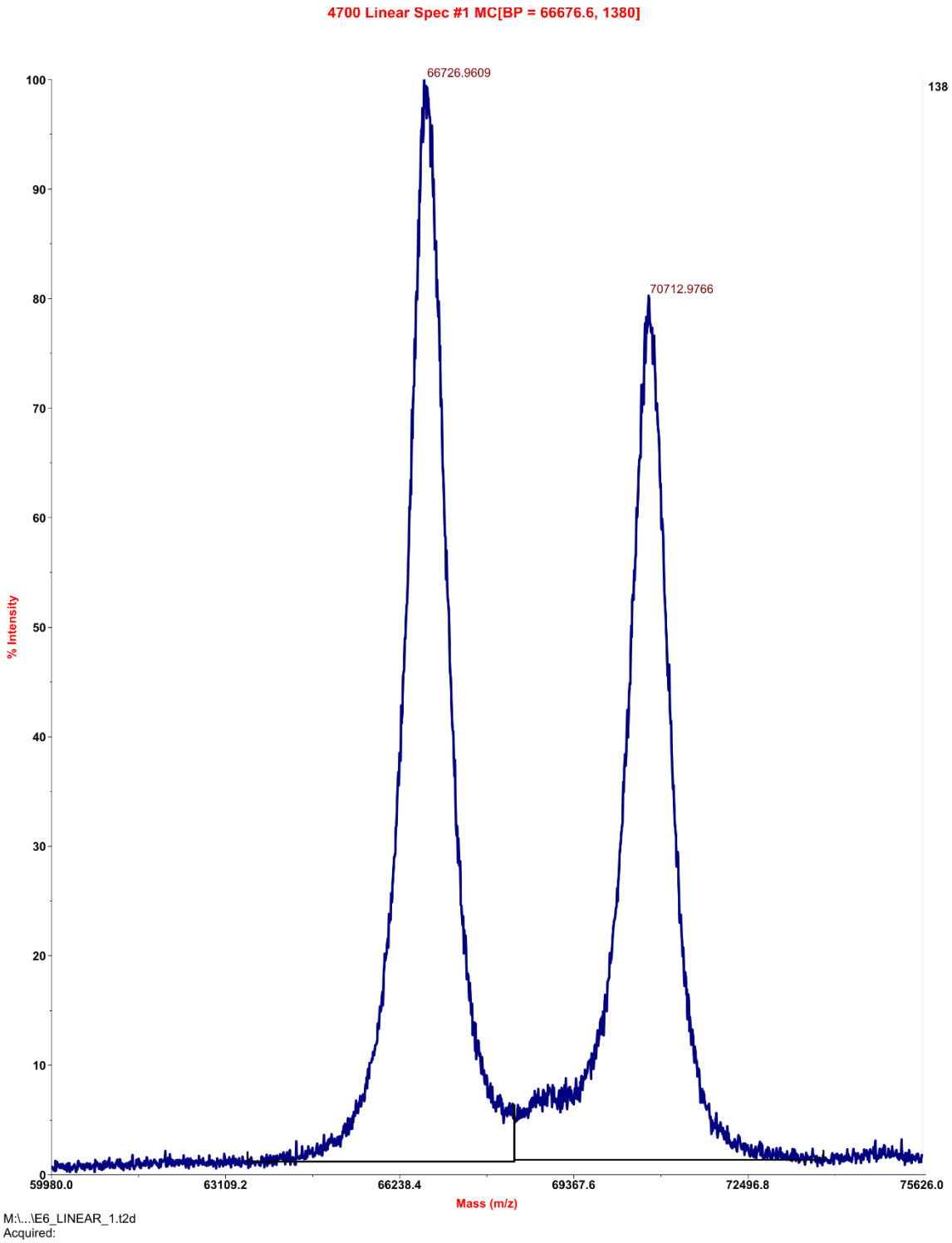


Figure 2-14: MALDI-TOS Mass spectrum of BSA conjugation with **3**. The mass shown here at 70712.9766 indicates product formation. The mass shown at 66726.9009 belongs to unconjugated BSA. The sample was desalted using C4 ZipTips and eluted into sinapinic acid then analyzed via Linear detection mode.

Tryptic Digest

Peptides were dissolved in buffered aqueous solution (100 mM NH_4HCO_3, pH 8.1) at a concentration of 100 μM. Trypsin was then added as a stock of 1 mg/mL in 50 mM acetic acid to a stirring solution of peptide substrate. The substrate to enzyme ratio was varied from 20:1 to 50:1 and time points were quenched by taking 50 μL of reaction solution into 2 μL of neat trifluoroacetic acid. Time points were analyzed via reverse-phase HPLC using the same conditions described above for solid phase peptide synthesis.

^{19}F MRI

To evaluate MRI signal intensity, a phantom was assembled containing a series of six 0.5 mL samples of peptide **4** at various concentrations in 1.5 ml conical tubes: 1) 2.13 mM (64 mM ^{19}F), 2) 1.07 mM (32 mM ^{19}F), 3) 0.53 mM (16 mM ^{19}F) , 4) 0.27 mM (8 mM ^{19}F), 5) 0.13 mM (4 mM ^{19}F), 6) control (0 mM). The tubes were arranged in two offset columns of three (see Figure 2-12) and then centered within a 2.5 cm diameter single-loop surface coil tuned to 656.8 MHz for imaging. ^{19}F MRI images were obtained using a 16.4 T MR system (Agilent Technologies, Santa Clara, CA) using a steady-state free precession sequence (FID-like) with the following parameters, flip angle = 20°, echo time = 1.02 ms, repetition time=2.04 ms, acquisition time = 0.61 ms, slice thickness = 3 mm, FOV = 10x10 cm, matrix = 128x128, signal averages n=2400-10 minute scan.

HPLC and 19F NMR spectra

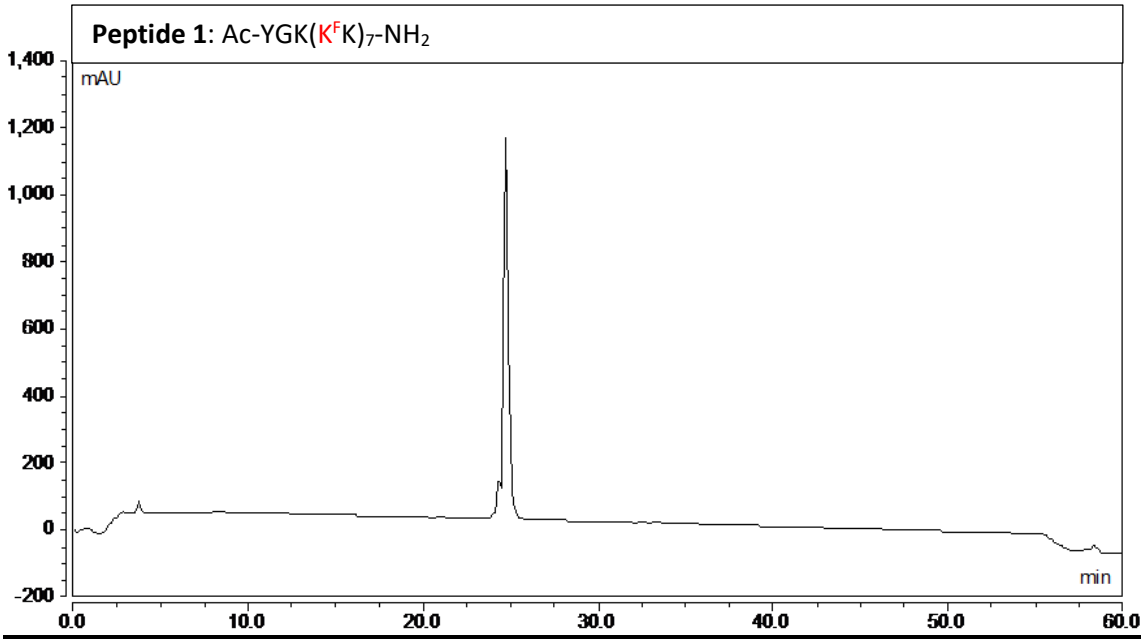


Figure 2-15: HPLC chromatogram of Peptide **1**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

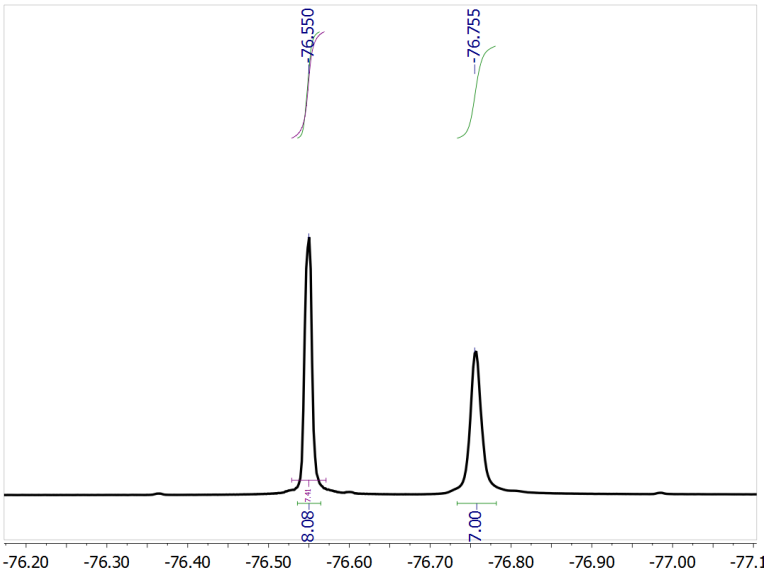


Figure 2-16: ^{19}F NMR of Peptide **1** (470 MHz, D_2O) The downfield resonance at -76.55 is the reference compound TFA

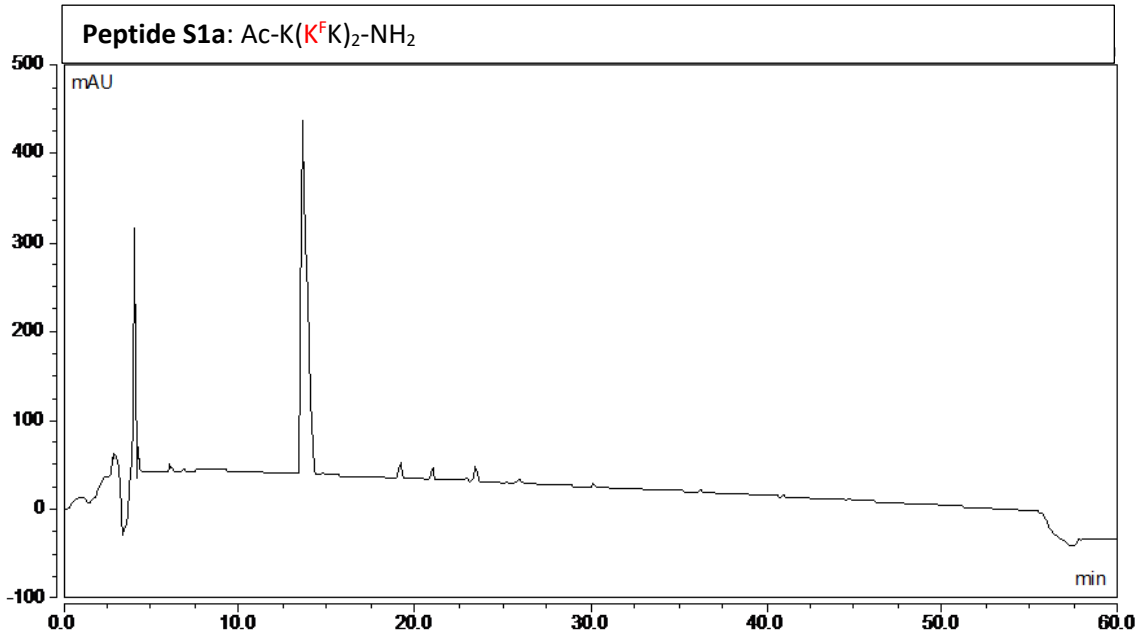


Figure 2-17: HPLC chromatogram of Peptide **S1a**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

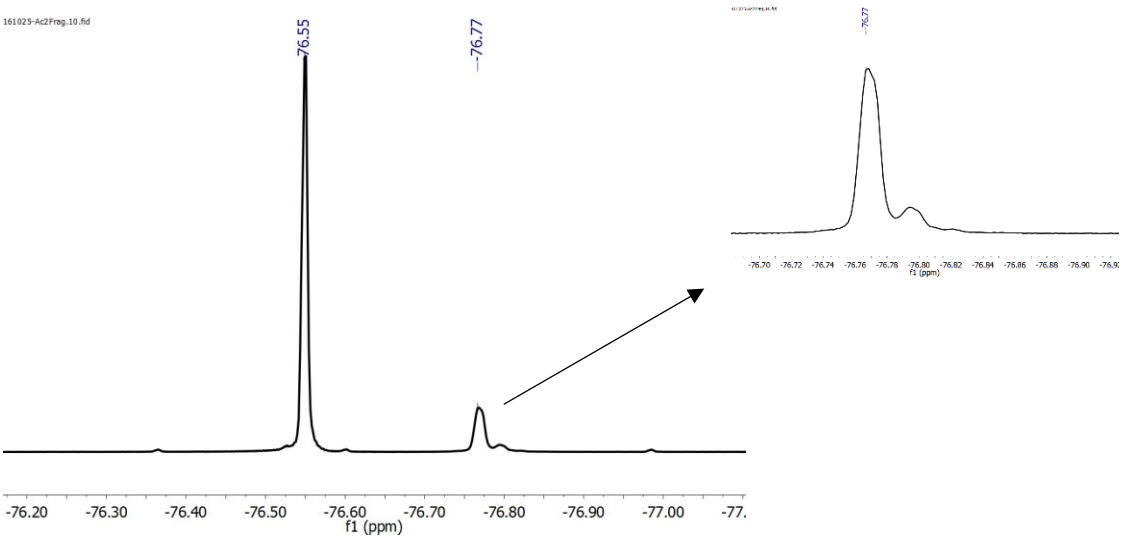


Figure 2-18: ^{19}F NMR of Peptide **S1a** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA. Arrow points to blown-up view of resonances.

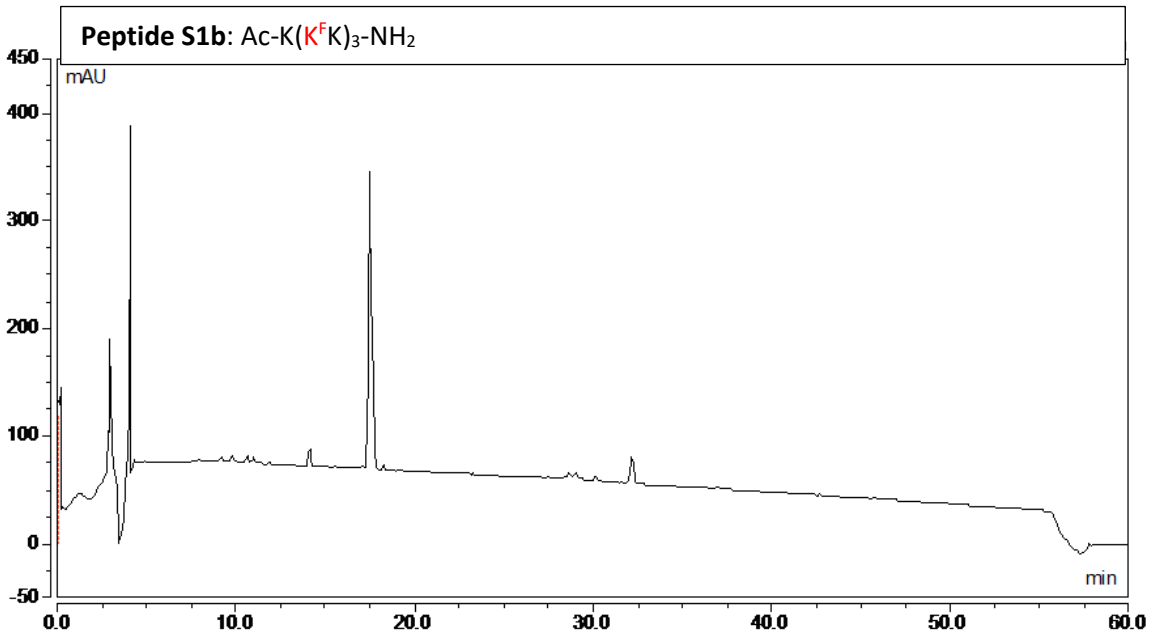


Figure 2-19: HPLC chromatogram of Peptide **S1b**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

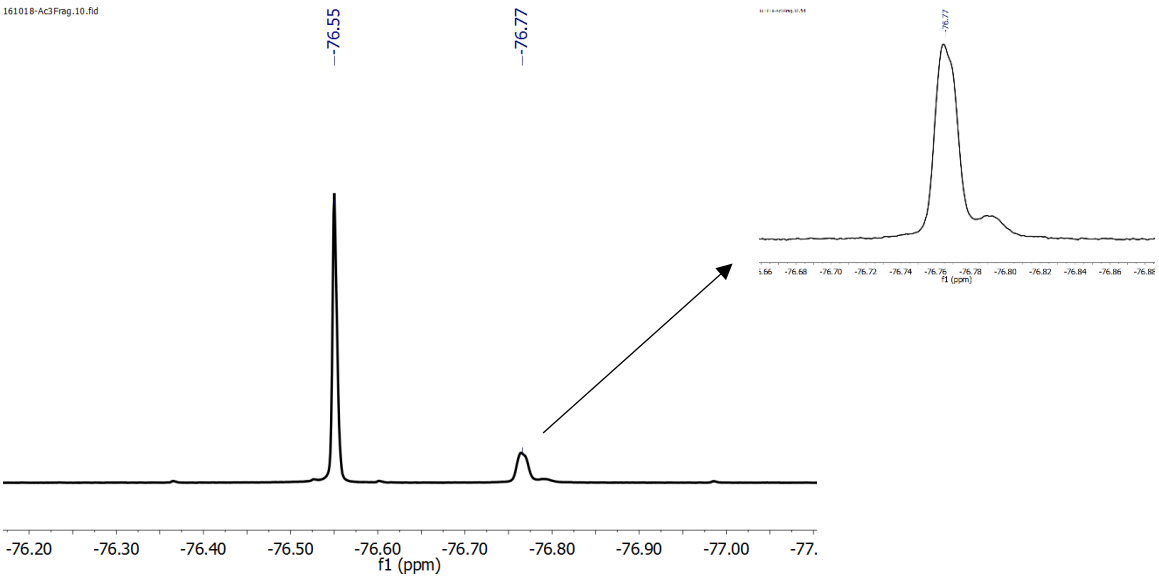


Figure 2-20: ^{19}F NMR of Peptide **S1b** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA. Arrow points to blown up view of resonance.

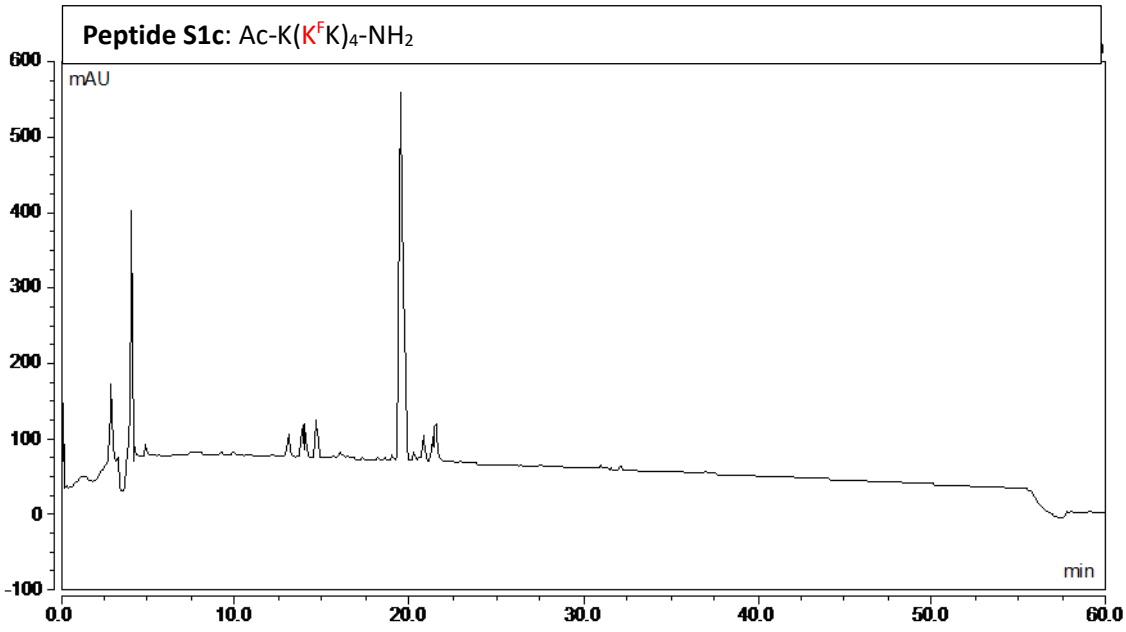


Figure 2-21: HPLC chromatogram of Peptide **S1c**. Gradient of 10:90 $CH_3CN:H_2O$ (0.1% TFA) to 60:40 $CH_3CN:H_2O$ (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

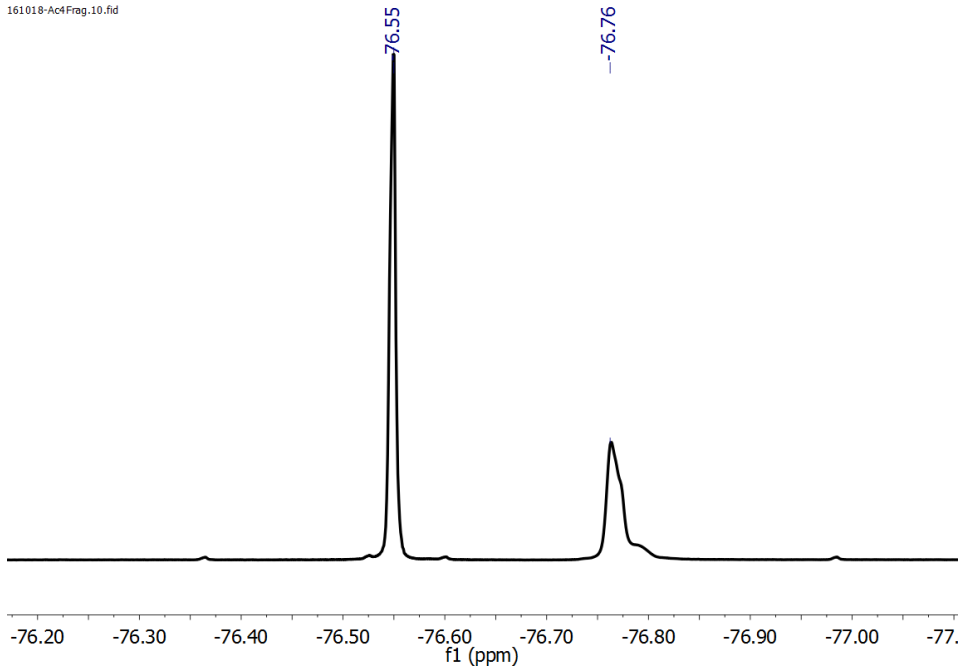


Figure 2-22: ^{19}F NMR of Peptide **S1c** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

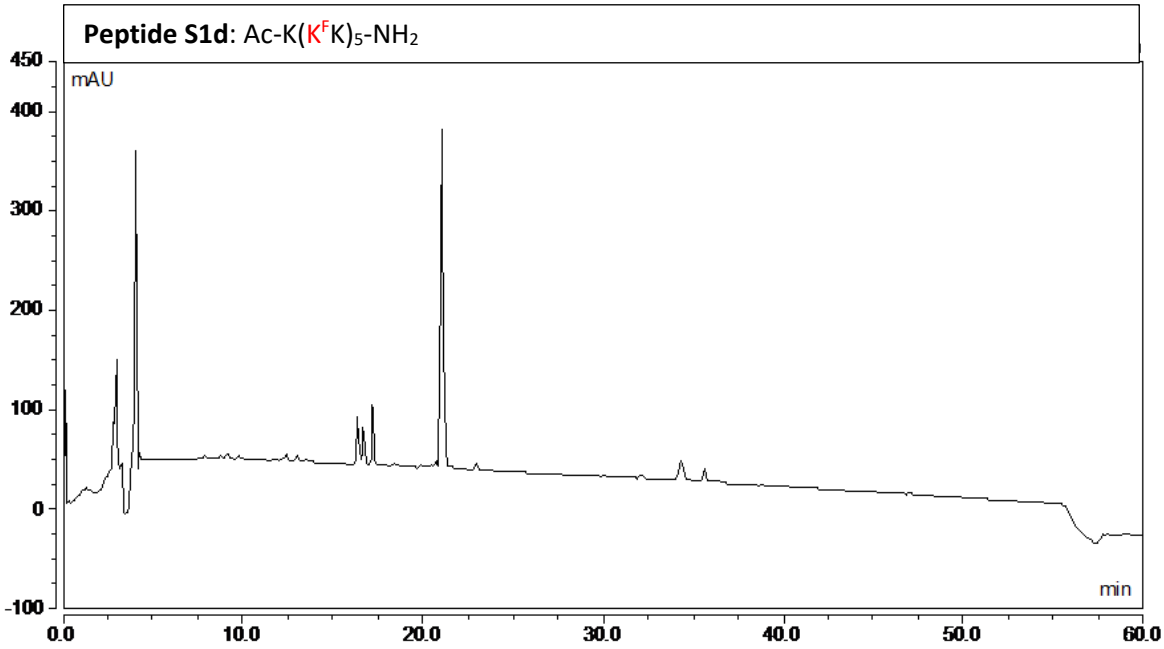


Figure 2-23: HPLC chromatogram of Peptide **S1d**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

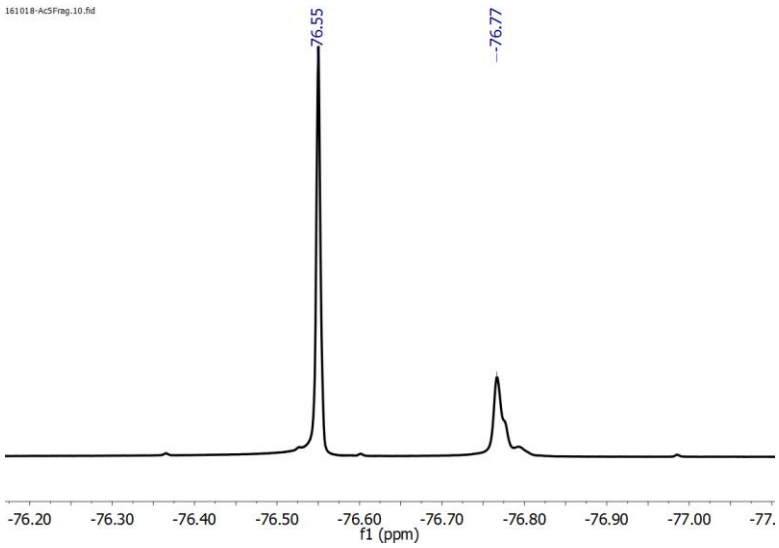


Figure 2-24: ^{19}F NMR of Peptide **S1d** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

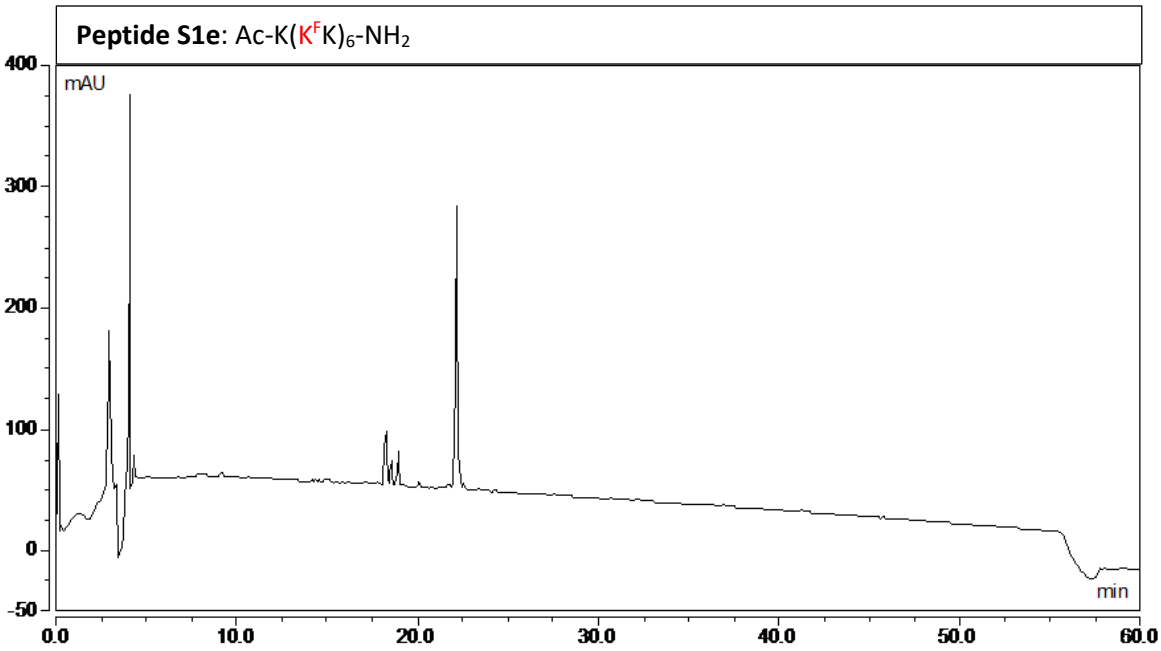


Figure 2-25: HPLC chromatogram of Peptide **S1e**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

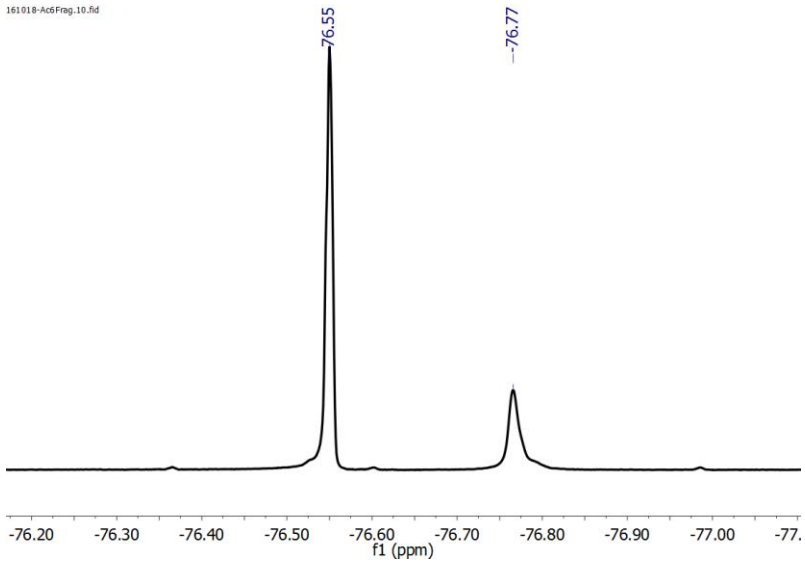


Figure 2-26: ^{19}F NMR of Peptide **S1e** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

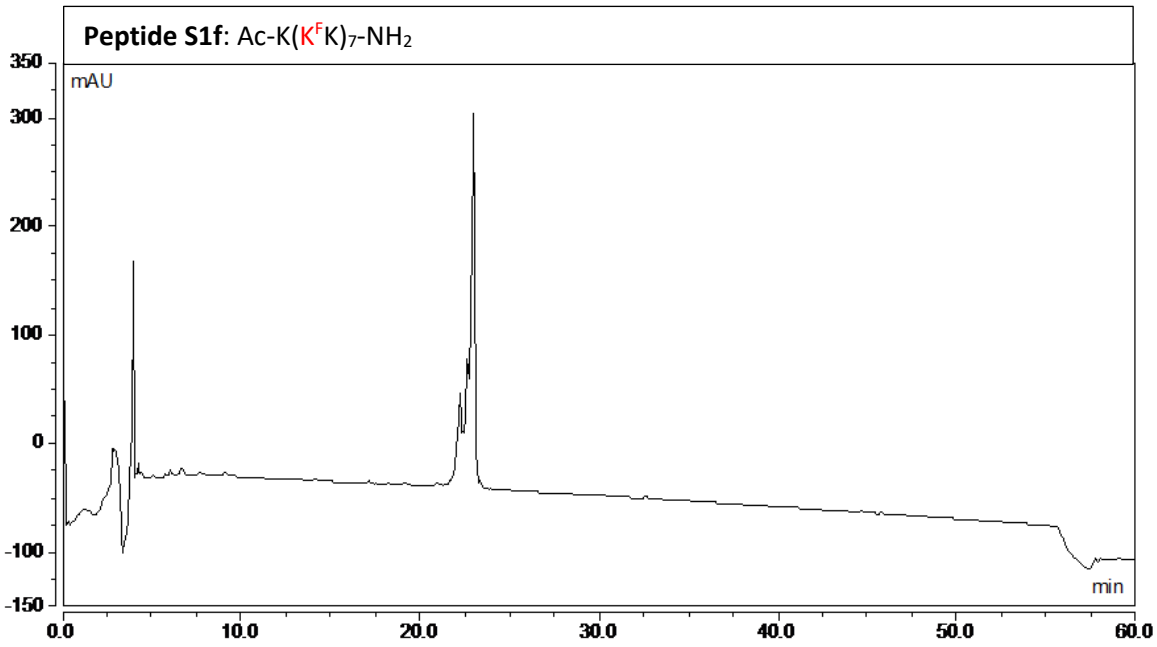


Figure 2-27: HPLC chromatogram of Peptide **S1f**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

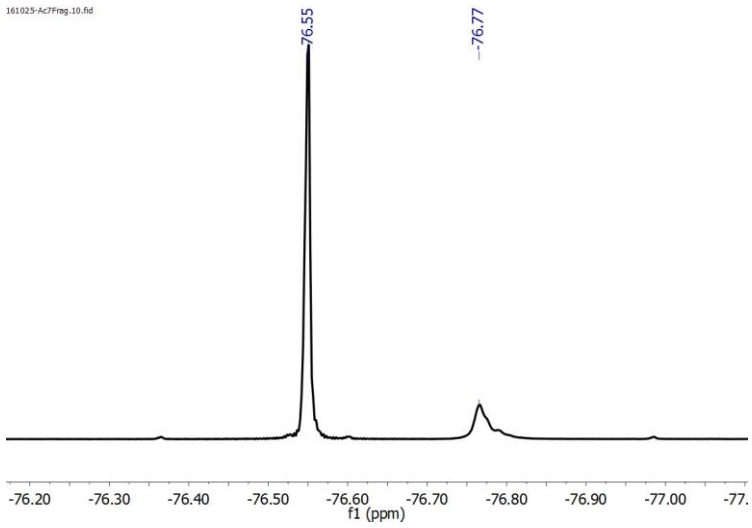


Figure 2-28: ^{19}F NMR of Peptide **S1f** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA. This spectrum includes impurities not separable in our hands via RP-HPLC

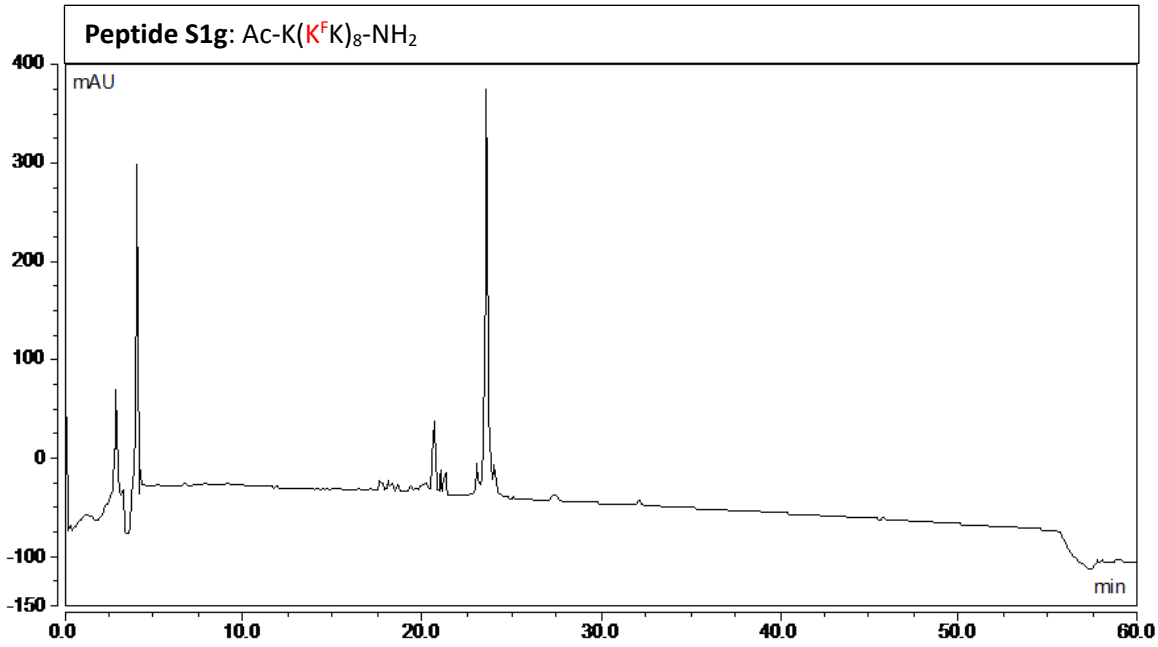


Figure 2-29: HPLC chromatogram of Peptide **S1g**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

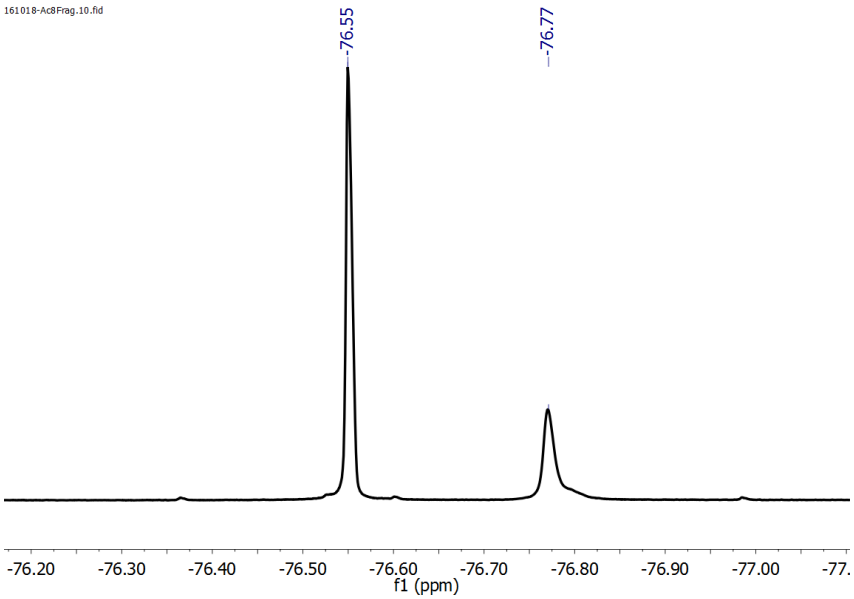


Figure 2-30: ^{19}F NMR of Peptide **S1g** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA. This spectrum includes impurities not separable in our hands via RP-HPLC.

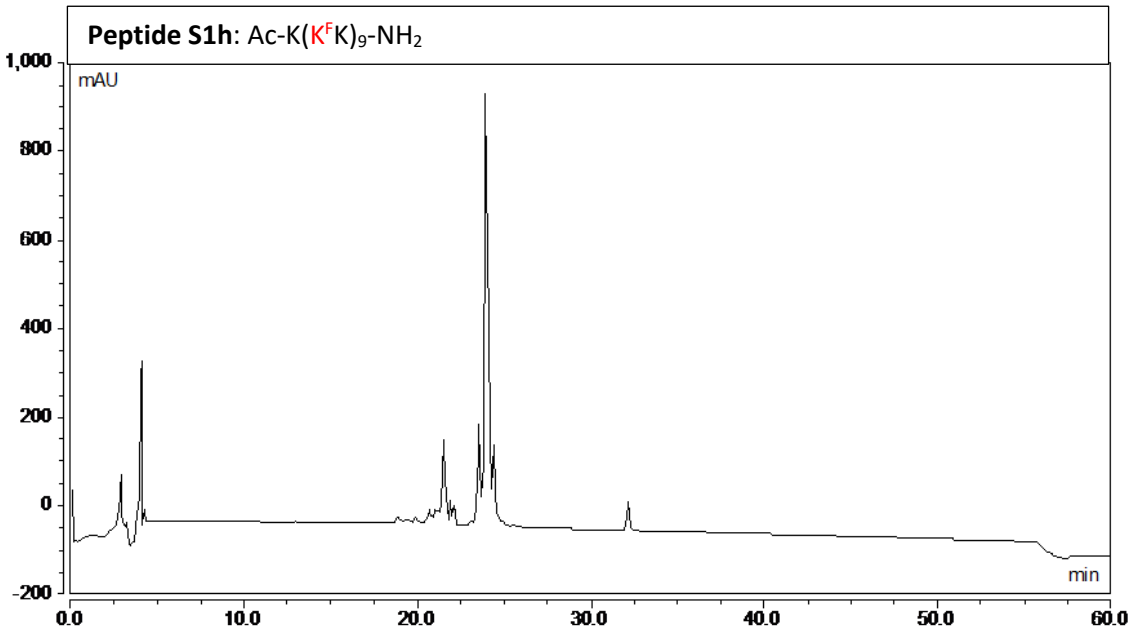


Figure 2-31: HPLC chromatogram of Peptide **S1h**. Gradient of 10:90 $CH_3CN:H_2O$ (0.1% TFA) to 60:40 $CH_3CN:H_2O$ (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

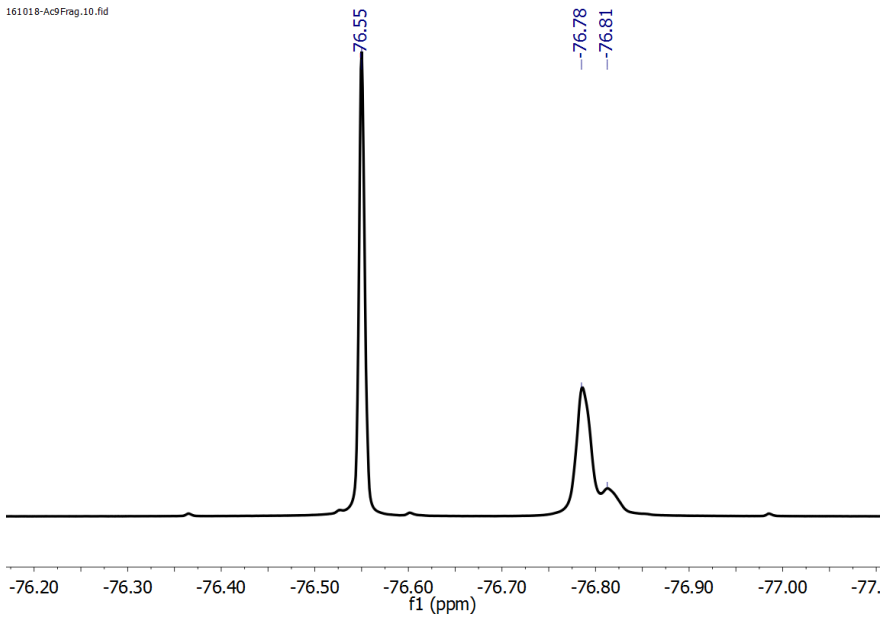


Figure 2-32: ^{19}F NMR of Peptide **S1h** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA. This spectrum includes impurities not separable in our hands via RP-HPLC

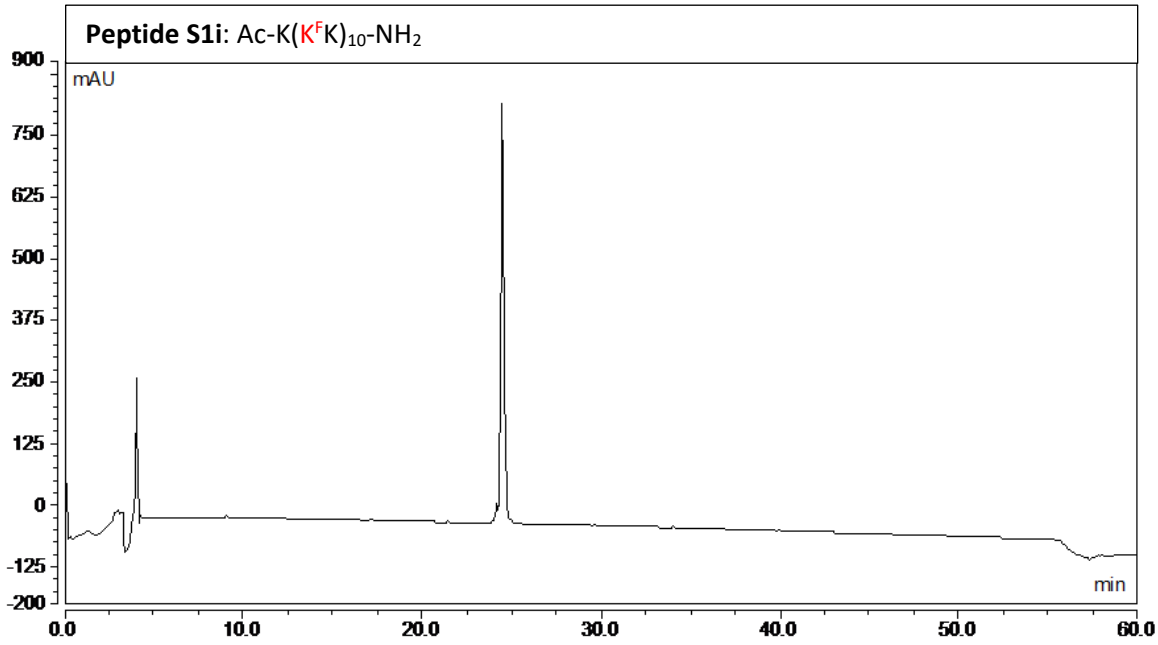


Figure 2-33: HPLC chromatogram of Peptide **S1i**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

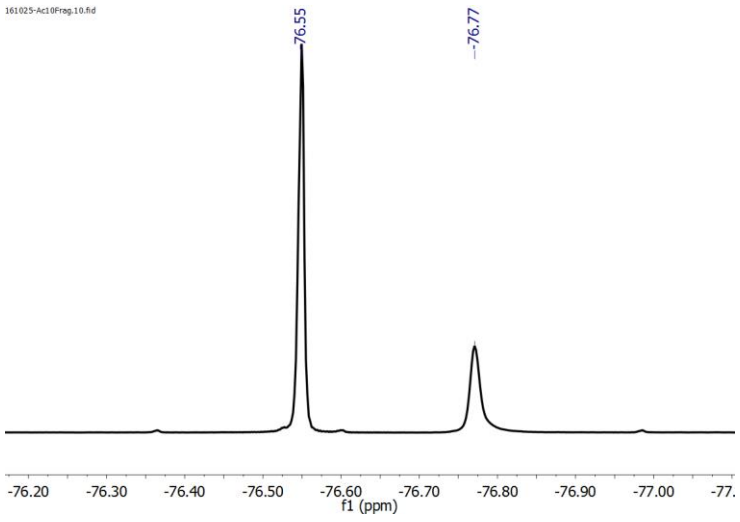


Figure 2-34: ^{19}F NMR of Peptide **S1i** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

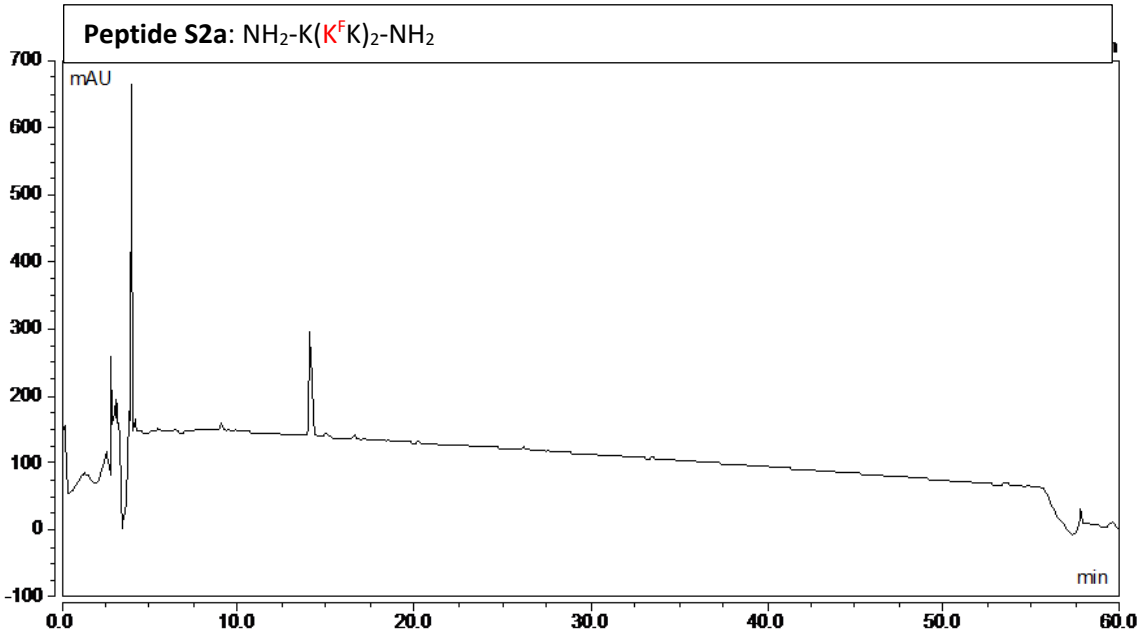


Figure 2-35: HPLC chromatogram of Peptide **S2a**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

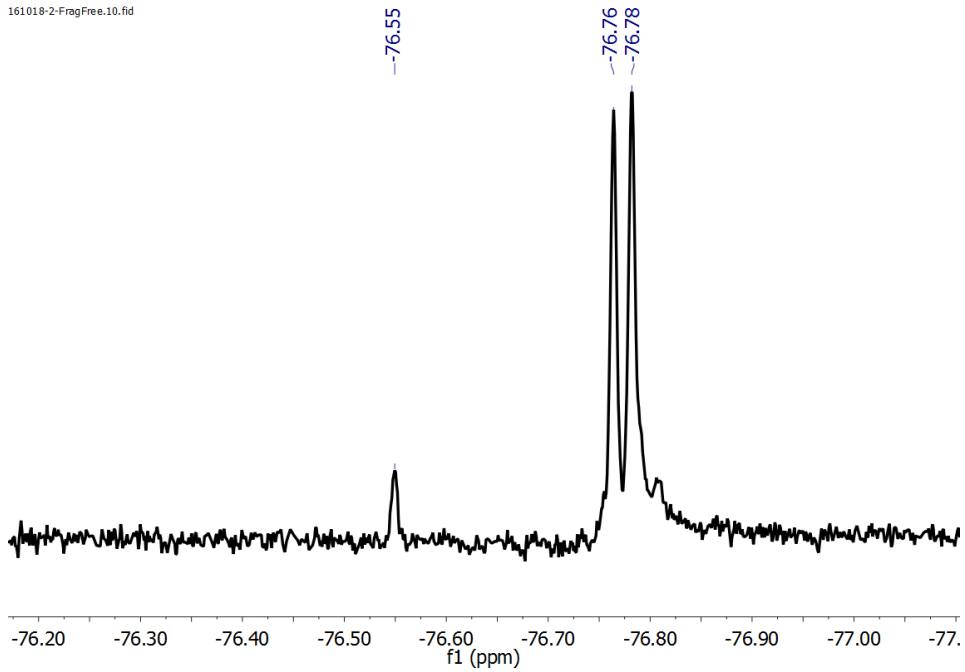


Figure 2-36: ^{19}F NMR of Peptide **S2a** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

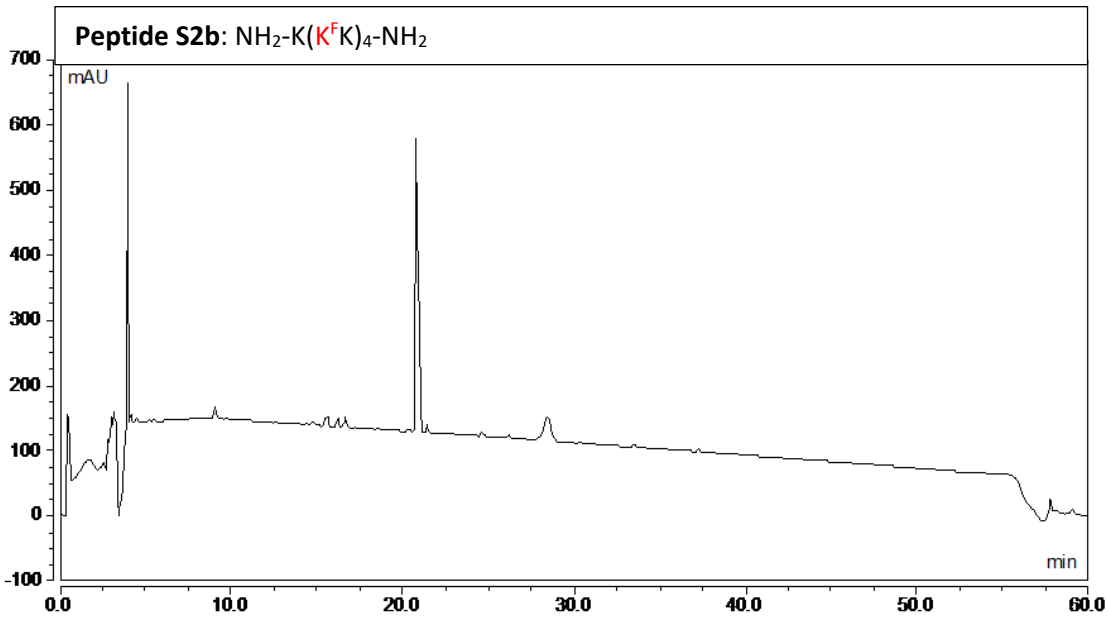


Figure 2-37: HPLC chromatogram of Peptide **S2b**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

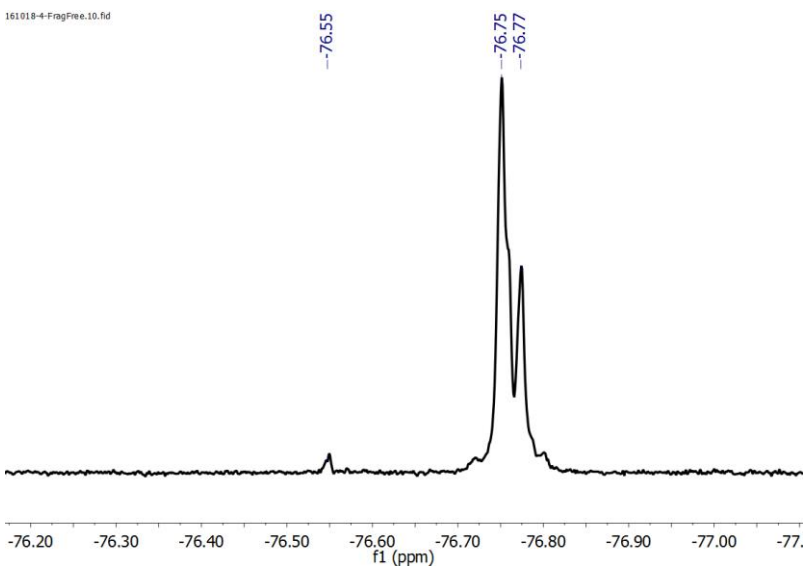


Figure 2-38: ^{19}F NMR of Peptide **S2b** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

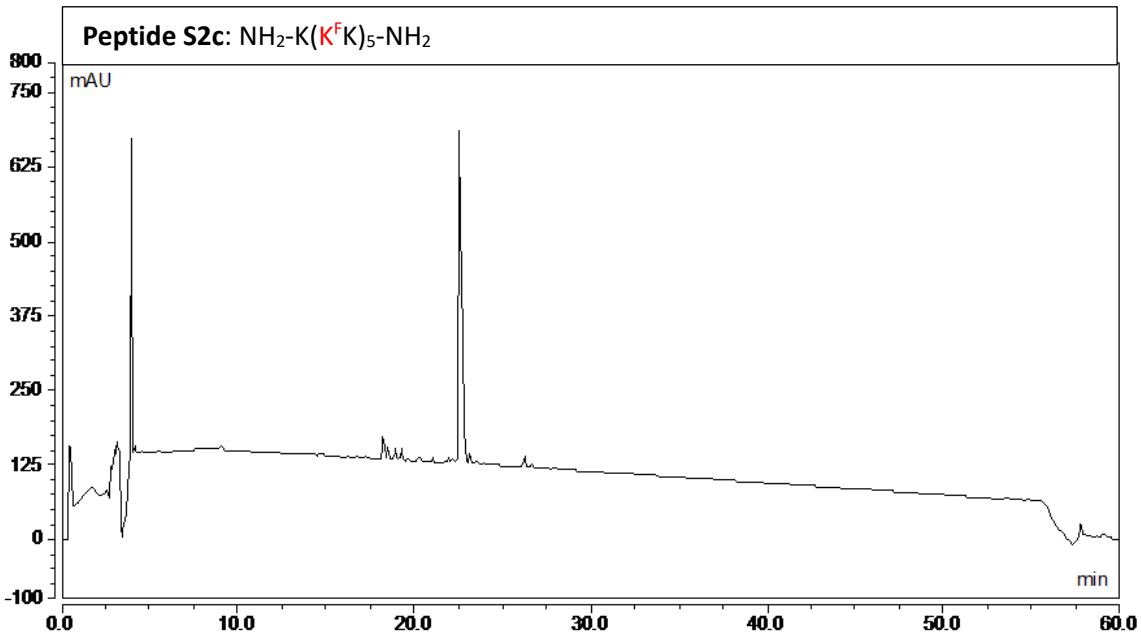


Figure 2-39: HPLC chromatogram of Peptide **S2c**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

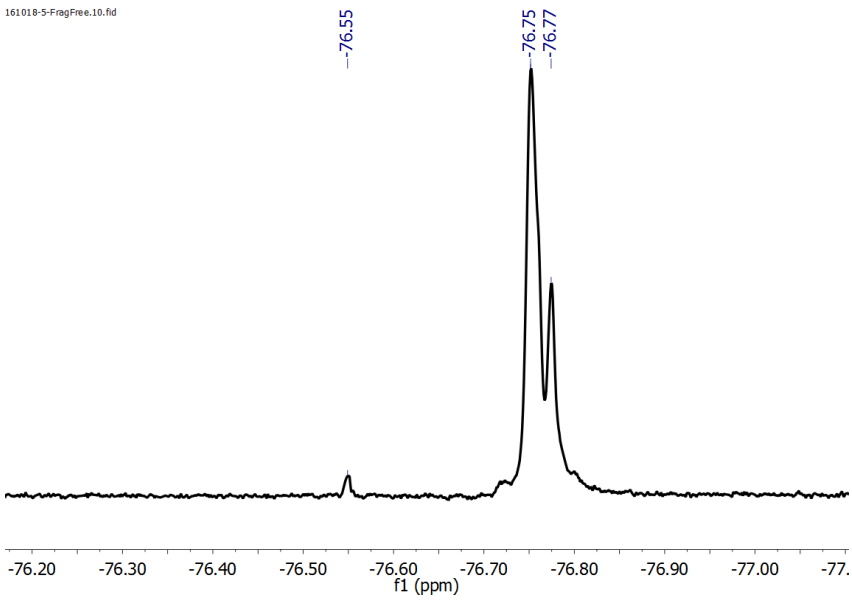


Figure 2-40: ^{19}F NMR of Peptide **S2c** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

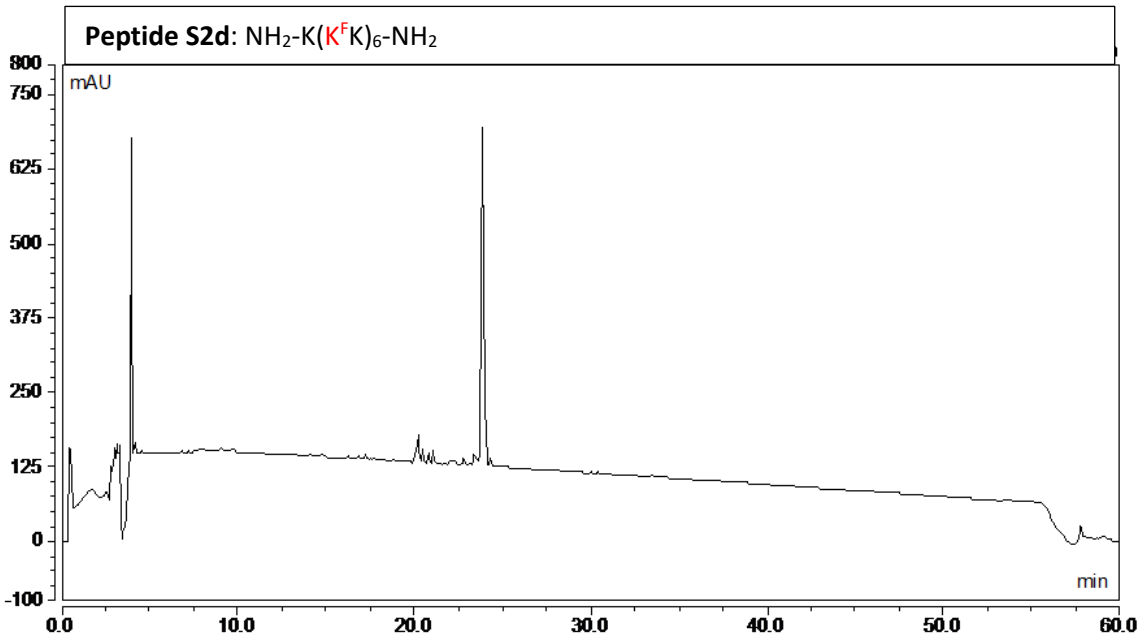


Figure 2-41: HPLC chromatogram of Peptide **S2d**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

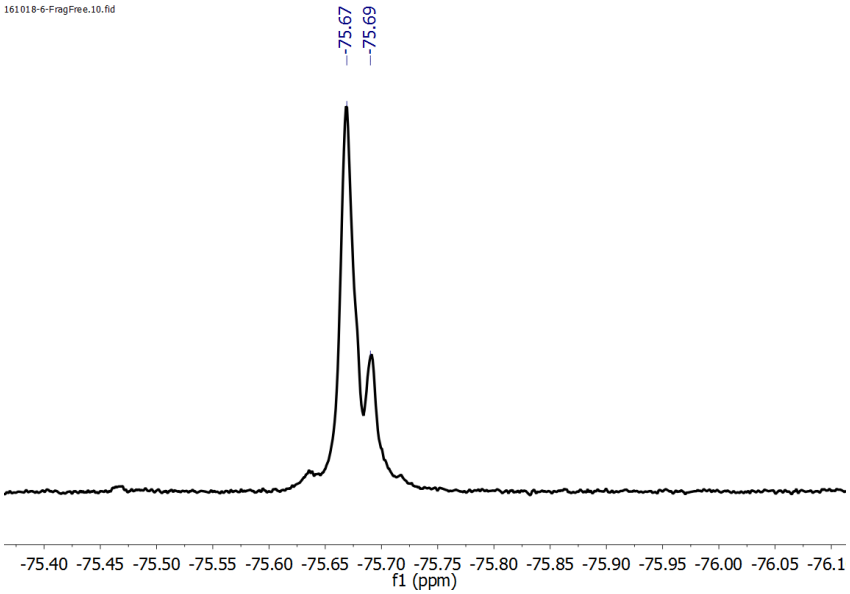


Figure 2-42: ^{19}F NMR of Peptide **S2d** (470 MHz, 10% D_2O, 90% H_2O).

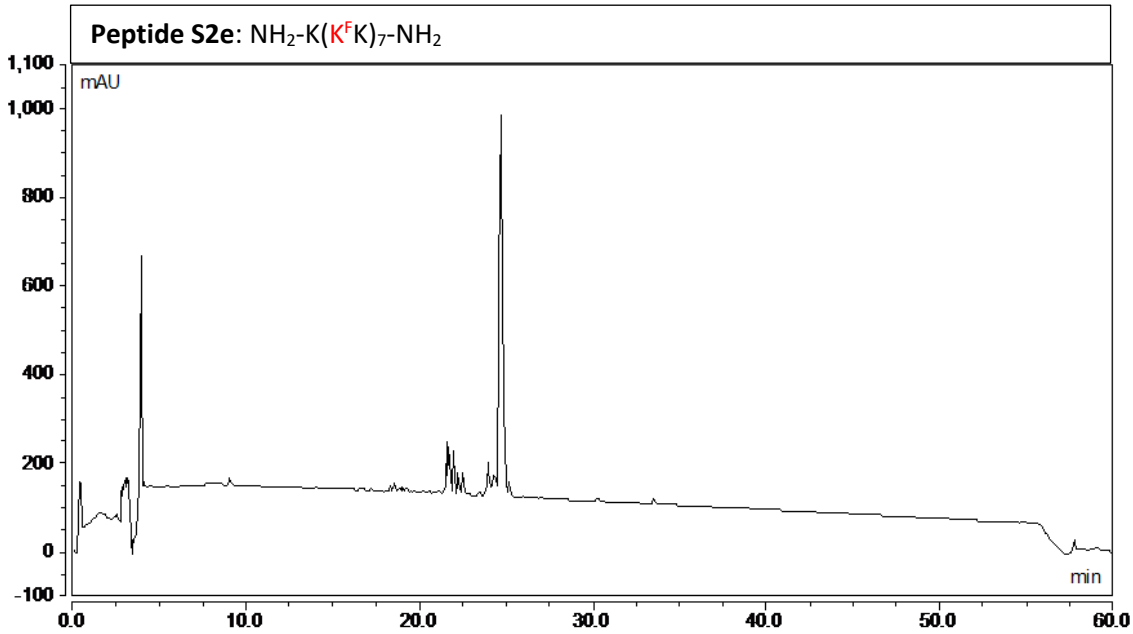


Figure 2-43: HPLC chromatogram of Peptide **S2e**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

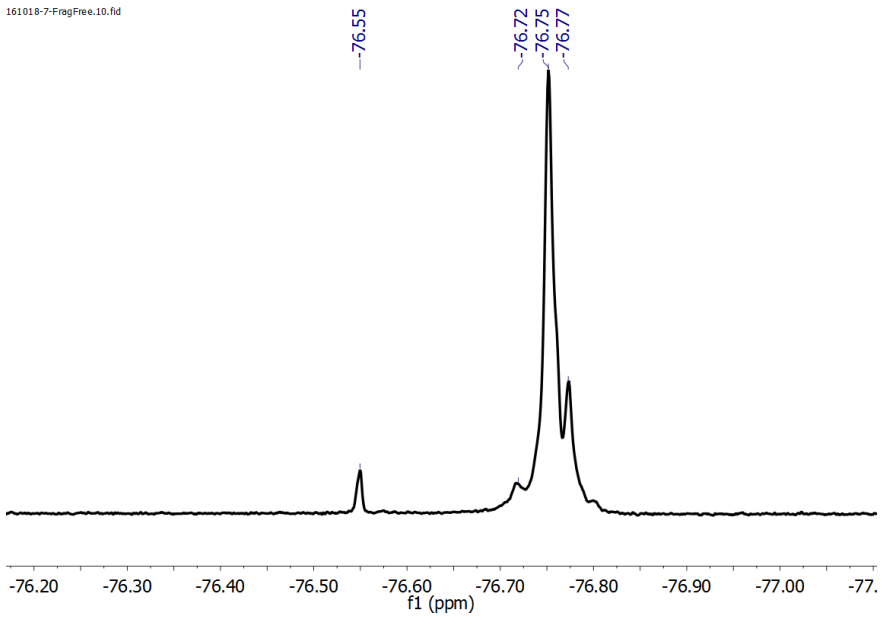


Figure 2-44: ^{19}F NMR of Peptide **S2e** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA. This spectrum includes impurities not separable in our hands via RP-HPLC

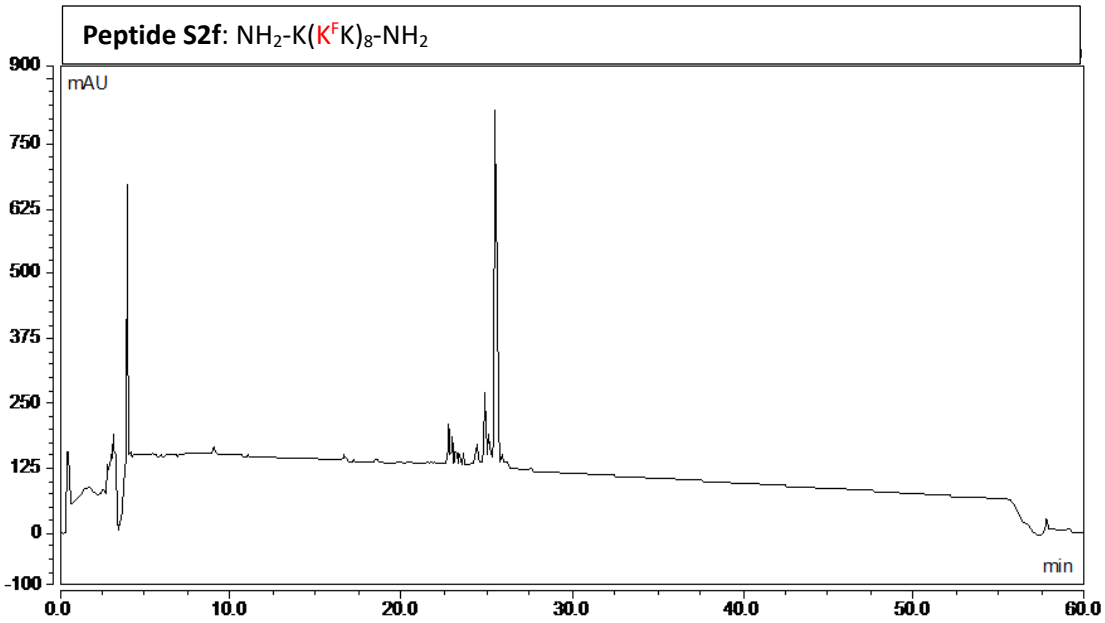


Figure 2-45: HPLC chromatogram of Peptide **S2f**. Gradient of 10:90 $CH_3CN:H_2O$ (0.1% TFA) to 60:40 $CH_3CN:H_2O$ (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

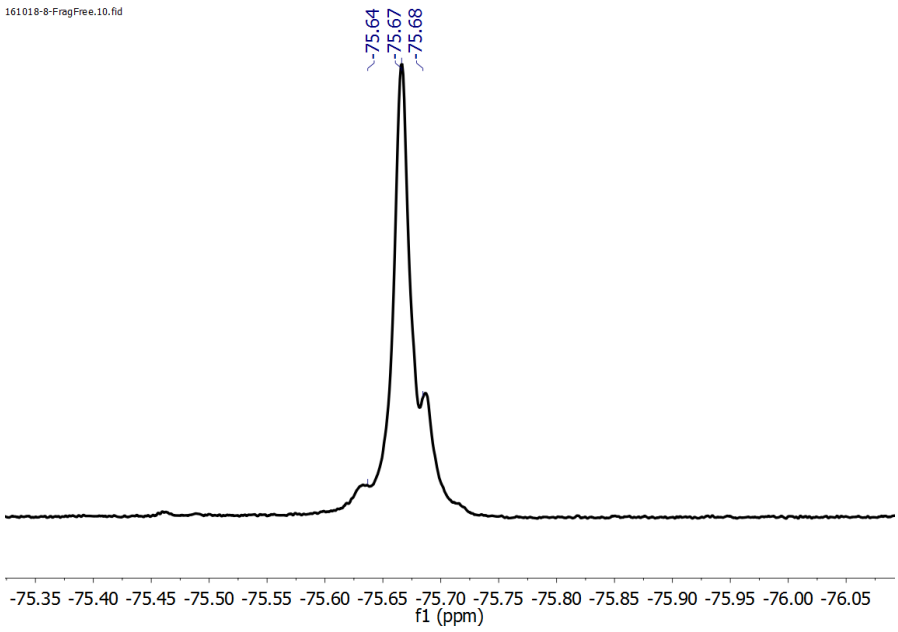


Figure 2-46: ^{19}F NMR of Peptide **S2f** (470 MHz, 10% D_2O, 90% H_2O) The chemical shifts in this spectrum are not referenced. This spectrum includes impurities not separable in our hands via RP-HPLC

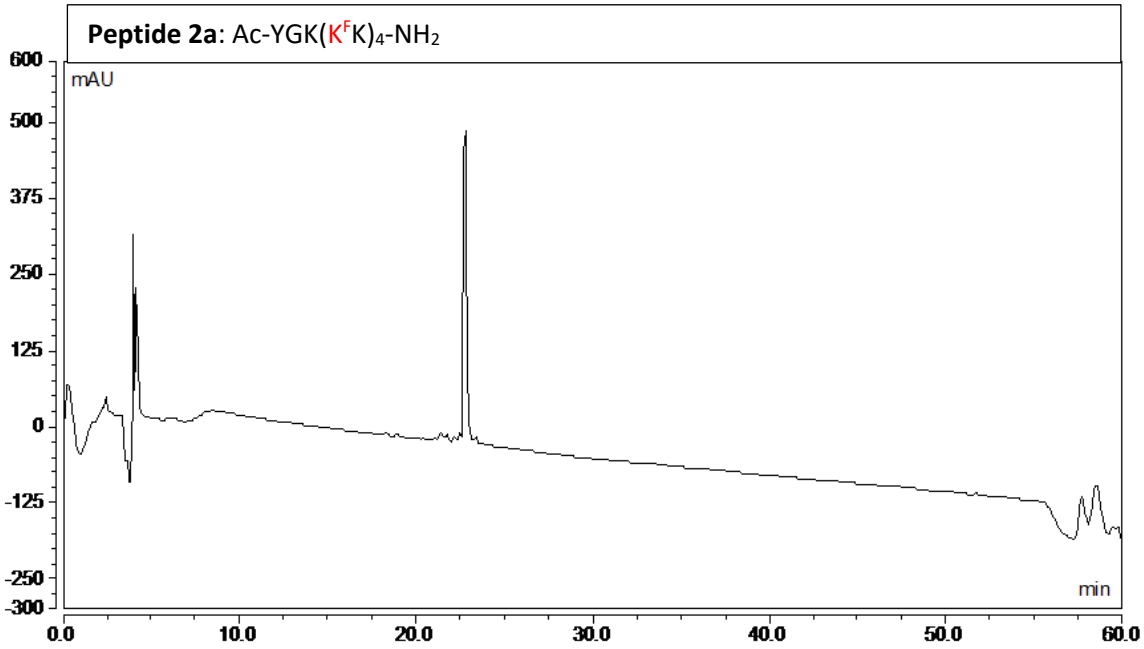


Figure 2-47: HPLC chromatogram of Peptide **2a**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

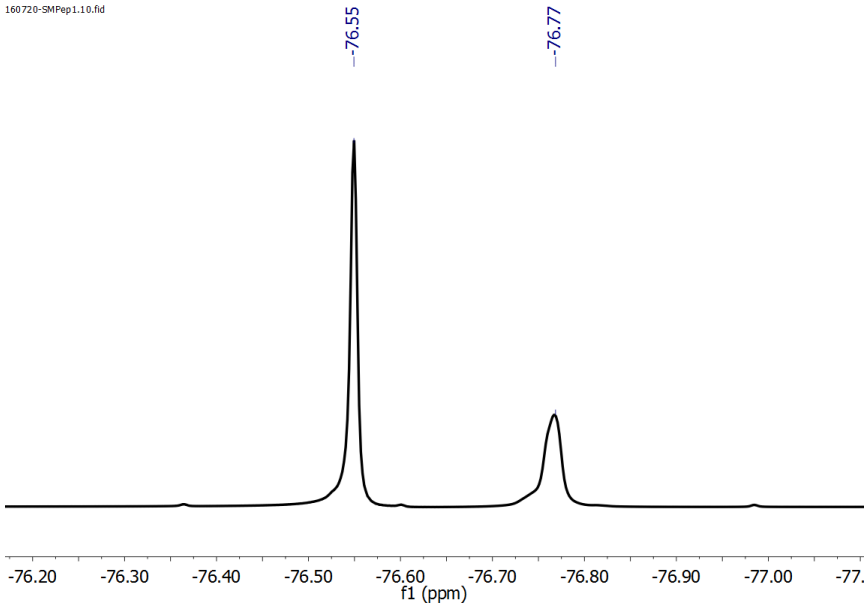


Figure 2-48: ^{19}F NMR of Peptide **2a** (470 MHz, D_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

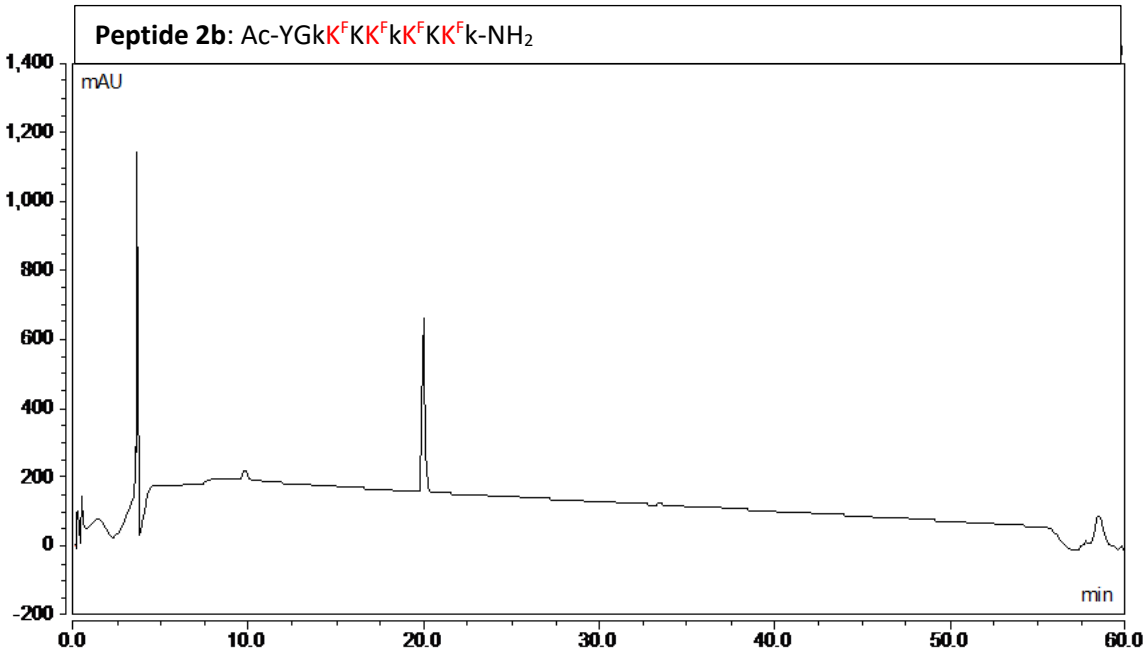


Figure 2-49: HPLC chromatogram of Peptide **2b**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

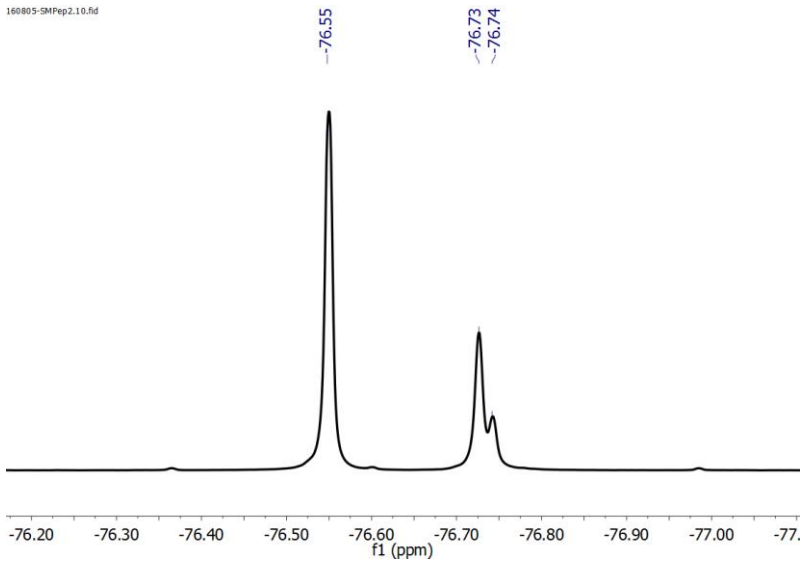


Figure 2-50: ^{19}F NMR of Peptide **2b** (470 MHz, D_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

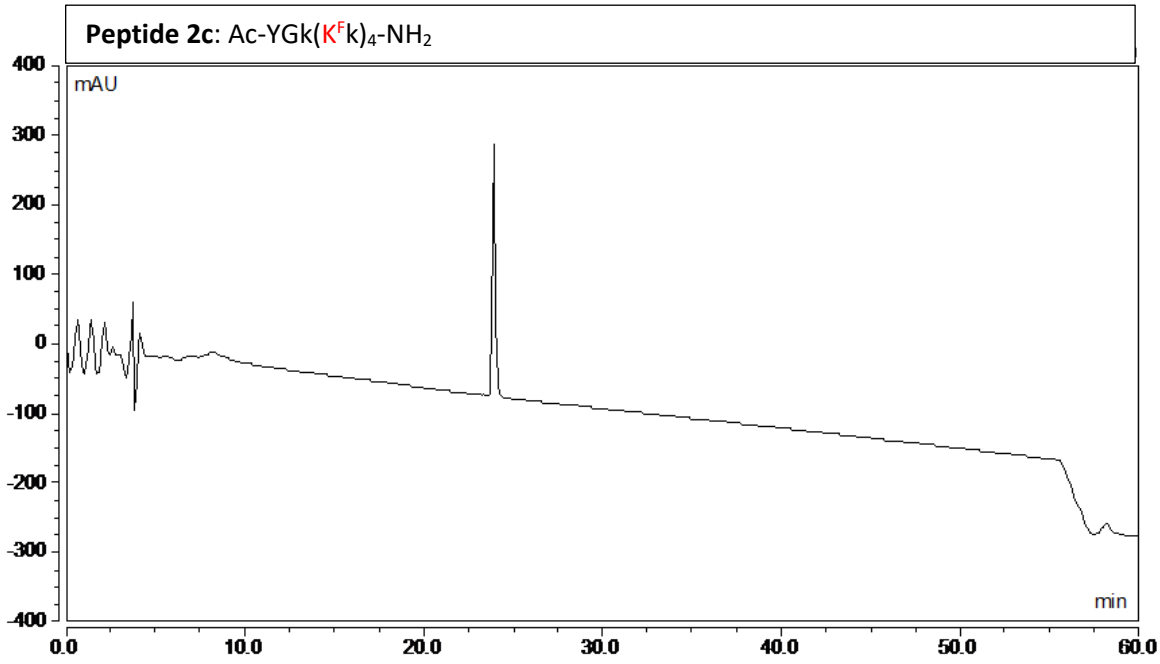


Figure 2-51: HPLC chromatogram of Peptide **2c**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

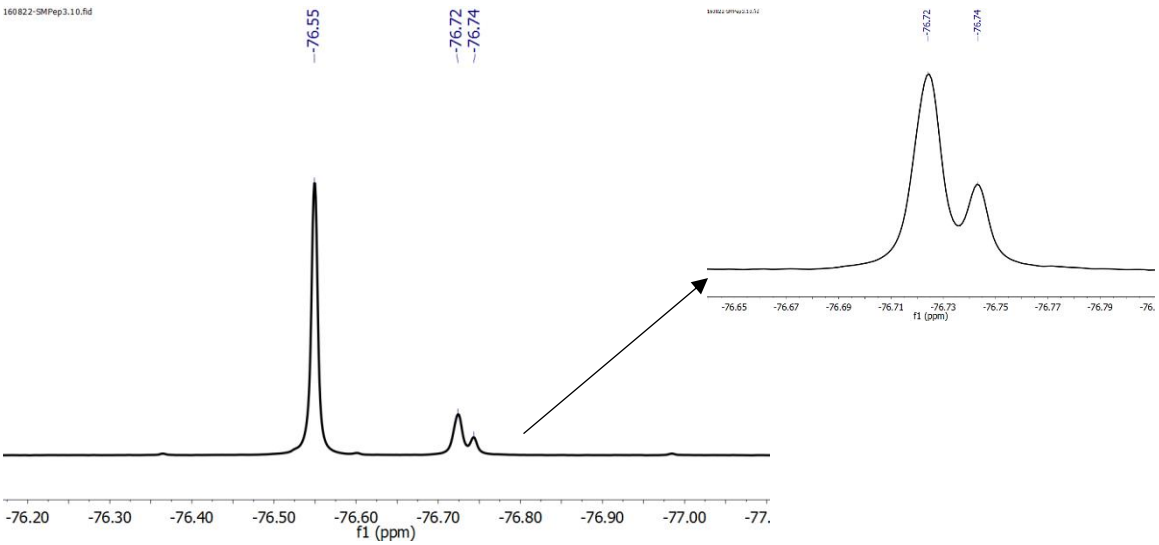


Figure 2-52: ^{19}F NMR of Peptide **2c** (470 MHz, D_2O) The downfield resonance at -76.55 ppm is the reference compound TFA. Arrow points to blown up view of resonances.

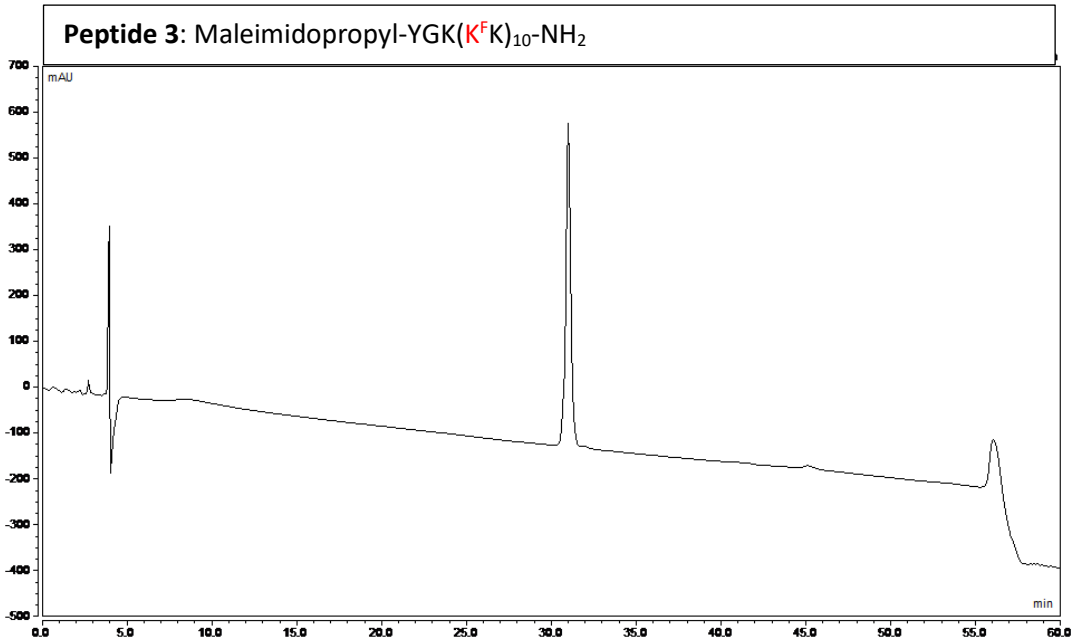


Figure 2-53: HPLC chromatogram of Peptide **3**. Gradient of 10:90 CH_3CN:H_2O (0.1% TFA) to 60:40 CH_3CN:H_2O (0.1% TFA) over 50 minutes with a 5 minute equilibration and 5 minute flush (95% CH_3CN). λ = 220 nm

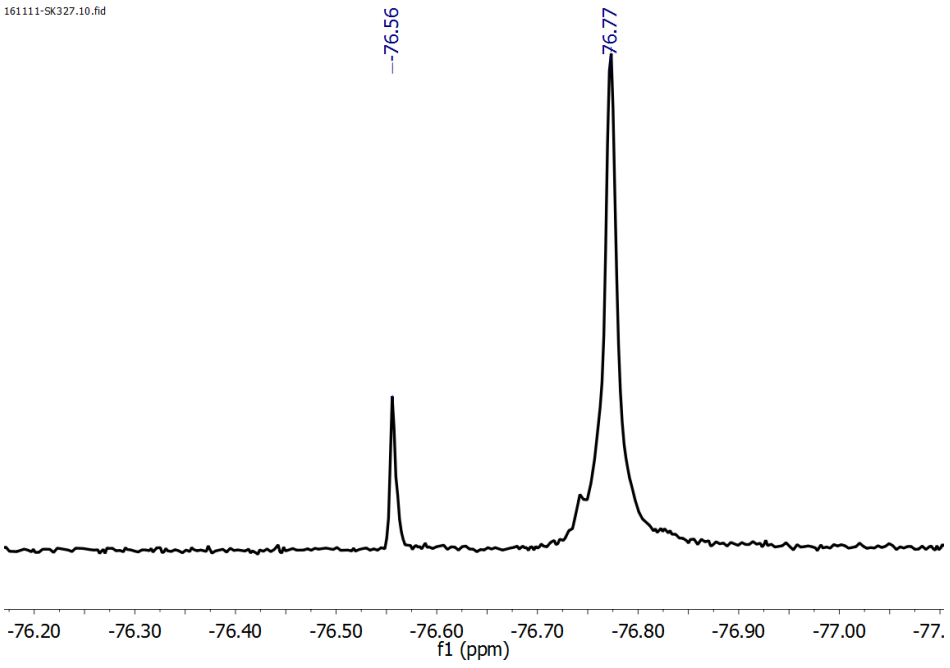


Figure 2-54: ^{19}F NMR of Peptide **3** (470 MHz, 10% D_2O, 90% H_2O) The downfield resonance at -76.55 ppm is the reference compound TFA.

Table 2-2**:** MALDI-TOF MS data for all peptides used (α-cyano-4-hydroxycinnamic acid matrix)

Peptide	**Calculated m/z** *($[M+Na]^+$)*	**Observed m/z** *($[M+Na]^+$)*
1	2895.4	2895.0
S1a	914.5	914.4
S1b	1266.6	1266.5
S1c	1618.8	1618.6
S1d	1971.0	1970.7
S1e	2323.2	2322.8
S1f	2675.3	2674.8
S1g	3027.5	3027.9
S1h	3379.7	3379.0
S1i	3731.8	3731.1
S2a	872.5	872.3
S2b	1224.6	1224.4
S2c	1576.8	1576.6
S2d	1929.0	1928.7
S2e	2281.1	2280.6
S2f	2633.3	2632.9
S2g	2985.5	2985.0
2a	1838.9	1838.6
2b	1838.9	1838.7
2c	1838.9	1838.7
3	4060.9	4061.4

Table 2-3**:** HPLC retention time for peptide series **S1** (10% CH_3CN at t = 0, 1 mL/min gradient)

Peptide	**Retention Time (min)**	**Difference from previous**
S1a	13.560	-
S1b	17.453	+3.893
S1c	19.463	+2.01
S1d	21.043	+1.58
S1e	22.150	+1.107
S1f	22.927	+0.777
S1g	23.547	+0.620
S1h	23.933	+0.386
S1i	24.440	+0.507

Table 2-4: HPLC retention time for peptide series S2 (10% CH_3CN at t = 0, 1 mL/min gradient)

Peptide	**Retention Time (min)**	**Difference from previous**
S2a	14.087	-
S2b	20.800	+2.327
S2c	22.503	+1.703
S2d	23.783	+1.280
S2e	24.647	+0.864
S2f	25.480	+0.833

Chapter 3

Toward a 2nd-generation peptide-based ^{19}F MRI probe

Motivation: The work in Chapter 2 primarily involved the design and synthesis of a peptide-based ^{19}F MRI agent. Following this study, efforts were made to improve upon the scaffold and address three key points that will be important to optimizing this scaffold for future clinical use: exploring alternative ^{19}F-labeled amino acids, appending a scaffold to a biomarker targeting agent, and studying the effects of altering the overall charge of the imaging agent.

3.1 Introduction

In 2017 we published on a peptide-based scaffold for use as a ^{19}F MRI agent. This peptide utilized sequences of N-ε-trifluoroacetylated lysine (TFA-lysine) alternating with unmodified lysine to produce a high intensity ^{19}F NMR signal, which could translate into a sensitive ^{19}F MRI agent.[27] The use of ^{19}F MRI represents an interesting and underexplored avenue for disease diagnosis at an early stage of progression. The rising use of MRI as a medical diagnostic technique provides and avenue for practical use of new technologies. ^{19}F MRI is not currently utilized in the clinic, in part for the need of development and optimization of high signal ^{19}F imaging agents. Perfluorocarbon compounds (PFC) are the current state of the art in the development of these imaging agents. These compounds often involve highly symmetrical molecules that will produce a degenerate ^{19}F signal. However, PFCs are typically employed in unstable emulsions due to their poor water solubility. Moreover, it can be difficult to engineer symmetrical PFCs to target disease biomarkers for use as molecular imaging probes as changes to the chemical structure will disrupt symmetry and potentially reduce the ^{19}F signal by dispersing the resonances.

In our report, we developed a peptide-based scaffold that would produce a high ^{19}F MRI signal while remaining water soluble. In addition, the peptides maintain potential for conjugation to targeting agents and fluorophores to enable practical applications like biomarker targeting or cell tracking. Our prior report discussed the design and synthesis

of these agents, but there still exists a need to elaborate on this scaffold to produce an imaging agent with immediate practical use in a clinic setting. Several avenues of research are in progress, including increasing the ^{19}F signal to higher levels, development of targeting agents for bioconjugation, and reducing non-specific binding to mammalian cells.

3.2 Alternative ^{19}F-labeled amino acids

The first discussion involves the choice of fluorinated amino acid utilized in Chapter 2, N-ε-trifluoroacetylated lysine (TFA-lysine). This source of ^{19}F signal proved to be successful in the development of a highly fluorinated peptide while still maintaining water solubility. Another attractive characteristic of TFA-lysine is its lability of the trifluoroacetamide, providing mechanisms for excretion from the body preventing bioaccumulation.[137] In addition to this, the fluorinated by-product is trifluoroacetate, which has been shown to be environmentally benign.[138] A potential drawback of TFA-lysine containing peptides is the size of peptide that could potentially be needed to meet the limit of detection for clinic-strength MRI equipment (30 mM at 1.5 T)[2,3]. This can be addressed by appending multiple iterations of peptide on a larger scaffold, or by having increasingly long peptides with higher TFA-lysine content. However, an alternative is to explore varied fluorinated amino acids that can provide more magnetically equivalent ^{19}F signal on a per-residue basis. Another potential avenue involves the same number of magnetically equivalent ^{19}F nuclei but invokes a change in functional group which possesses more freedom of rotation around the CF_3 center leading to a potentially more narrow and thus intense fluorine resonance. This section will discuss unpublished work detailing the small molecule and peptide synthesis of three fluorinated amino acids (Figure 3-1).

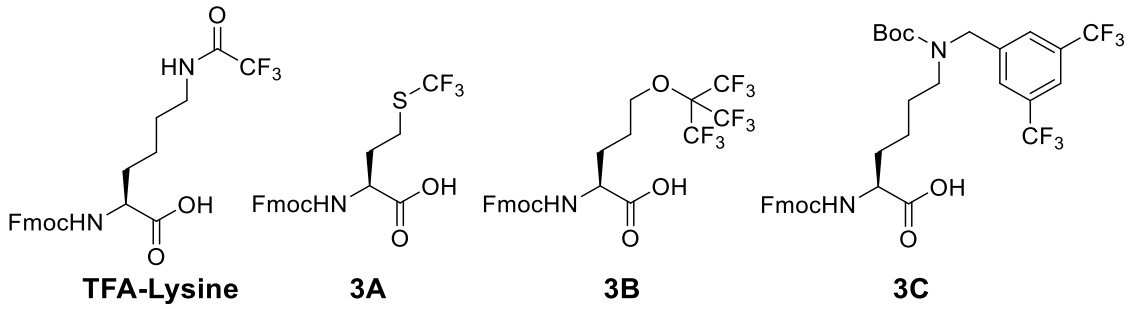


Figure 3-1: Structures of fluorinated amino acids studied here

3.2.1 Trifluoromethyl methionine (4A)

The first fluorinated functional group investigated was a trifluoromethyl sulfide based on three magnetically equivalent fluorine atoms and fast C-S bond rotation, a methionine analog **3A**. However, there are two key differences between **3A** and TFA-lysine. The first difference is that the S-CF_3 bond in **3A** should encounter less of a barrier to rotation compared to the C(O)-CF_3, leading to sharper resonances observed in the NMR. The second key distinction is that **3A** can be incorporated into proteins biosynthetically, similar to the method described in Chapters 1 and 4 involving 5FW BPTF.[139] Due to the lone pair electrons on sulfur in a conventional thioether, there is a significantly lowered steric hindrance about the S-CF_3 center in **3A** relative to a C-CF_3 center (Figure 3-2). This will allow the bond to rotate freely, leading to a sharper ^{19}F NMR resonance. This sharper resonance could potentially lead to a higher MRI signal compared to an equivalently labeled TFA-lysine peptide when conjugated to a larger targeting scaffold due to the likely broadening of the signal. The other intriguing trait of **3A** is the potential of incorporating the fluorinated residue biosynthetically into a protein scaffold with a pre-engineered biomarker targeting capability, negating the need to optimize bioconjugation reactions. This route would allow the design of a peptide to be appended into the DNA encoding a protein of interest, allowing for direct expression and purification of a fluorinated MRI agent with targeting capabilities.

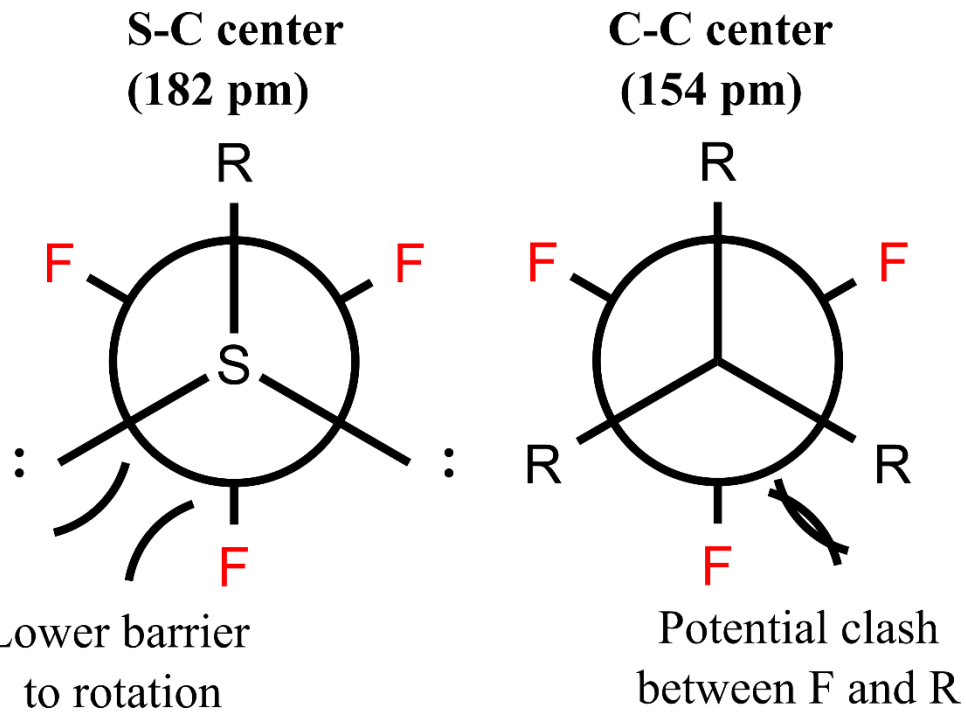


Figure 3-2: Newman projection of a CF_3 moiety attached to a sulfur or carbon center. Less barrier to rotation should be observed in the S-CF_3 case due to a longer S-C bond, and lone pairs providing less steric clash.[140]

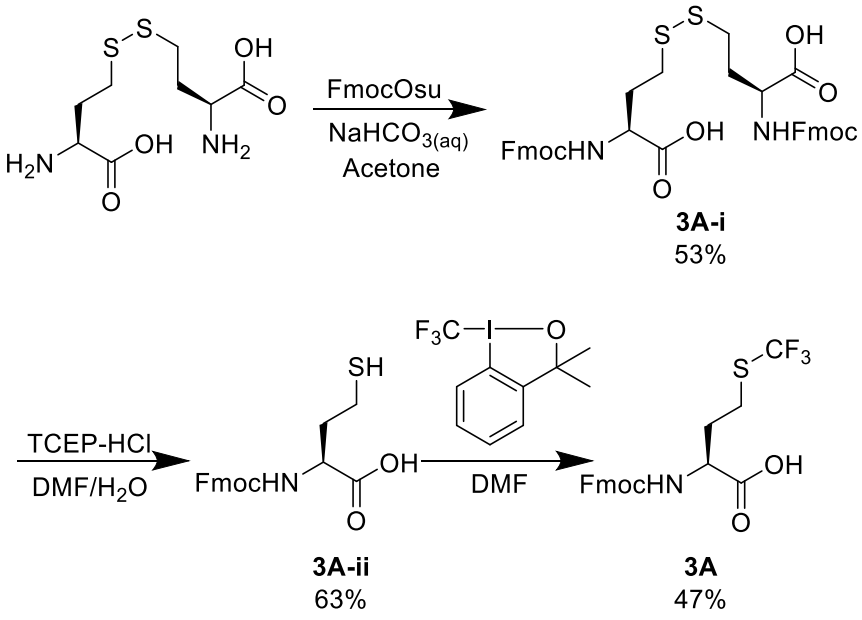


Figure 3-3: Synthetic scheme of **3A**

The synthesis of TFM was based off of a method developed by Suzuki et al (Figure 3-3).[141] The reaction was carried out in three steps beginning with L-homocystine. The N-α amines of L-homocystine were protected with Fmoc, followed by a TCEP-mediated reduction of the disulfide, finally the free thiol was alkylated with a hypervalent iodine-containing fluorinated agent known as the Togni reagent.[142] The scheme for this transformation is outlined below. The purified Fmoc-labeled amino acid was incorporated into a peptide scaffold using the same conditions outlined in Chapter 2 for TFA-lysine peptides. A sequence was synthesized containing 4 iterations of **3A**, leading to a peptide that showed four distinct resonances over a span of 31 Hz. It is likely due to the increased freedom of rotation of the S-CF_3 bond that this apparent resonance dispersion was observed. This data was originally collected prior to our first use of TFA-lysine, and prompted the original investigation of alternative amino acid choices.

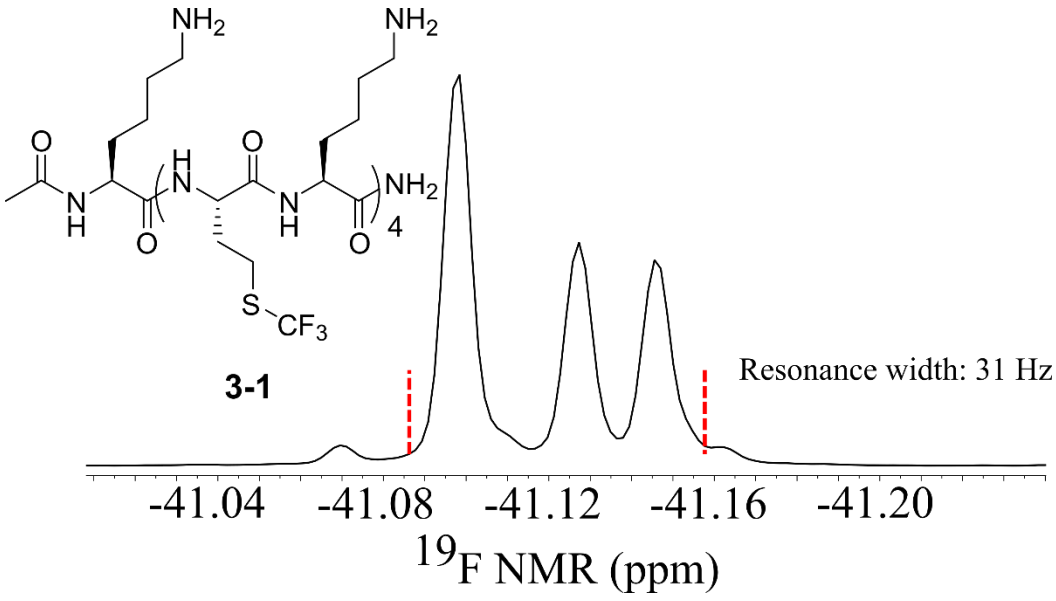


Figure 3-4: Sequence of **3-1** and resulting ^{19}F NMR spectrum. Resonance measurement from dashed lines.

*3.2.2 Perfluorotertbutyl alkoxynorvaline (**3B**)*

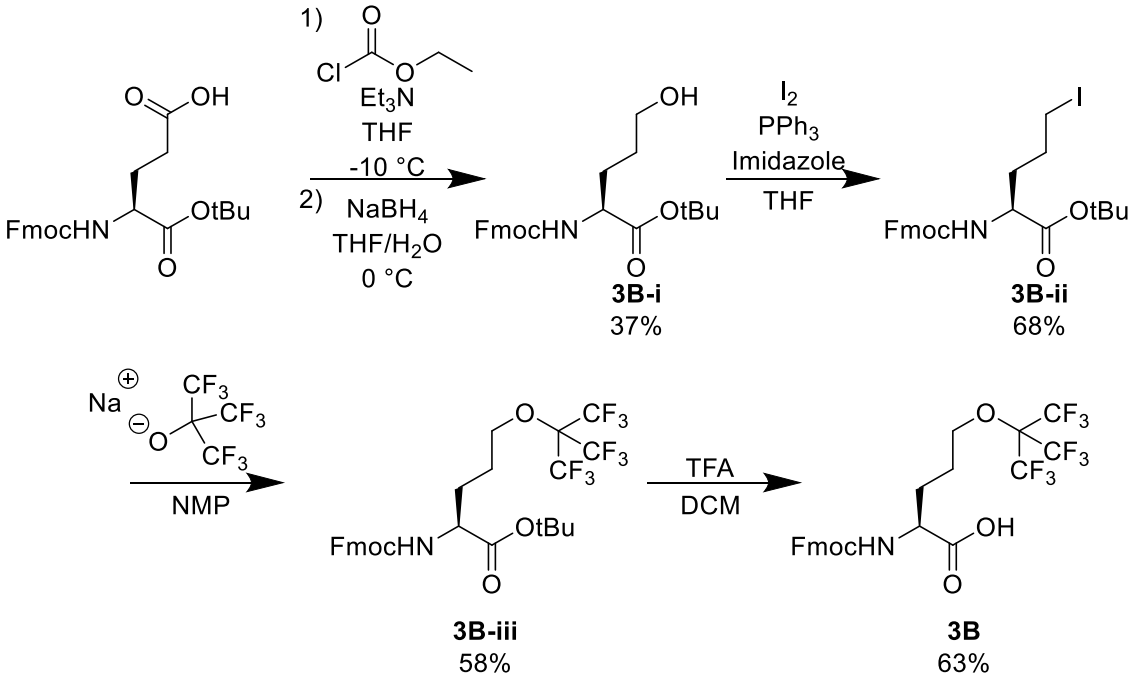


Figure 3-5: Synthesis of **3B**

In an attempt to improve upon the per-residue ^{19}F incorporation of peptides, we sought to utilize perfluorotertbutyl ethers which possess nine magnetically equivalent ^{19}F nuclei. Increasing the ^{19}F incorporation on a per-residue basis will allow the employment of shorter peptides to achieve the same signal. This could potentially be important to help maintain narrower resonances and higher signal to noise, as higher molecular weight peptides will tend to be more susceptible to NMR broadening due to shorter T_2 relaxations.

The synthesis of **3B** was developed by employing and elaborating upon reactions developed by Jobron and Hummel.[143] The reaction began from an Fmoc-protected glutamic acid (Figure 3-5). However as opposed to conventionally used amino acids for solid phase peptide synthesis, this residue was protected with a tert-butyl ester on the α-carbonyl as opposed to the side chain. The side chain was then reduced using ethyl chloroformate to generate an intermediate carbonic anhydride, followed by reduction with sodium borohydride to yield a primary alcohol, **3B-i**. The hydroxyl of **3B-i** was next transformed

into an alkyl iodide (**3B-ii**) through an Appel reaction followed by fluorination using sodium perfluorotertbutoxide to yield **3B-iii**. Finally, the α-tert butyl ester was deprotected under acidic conditions to produce the final amino acid scaffold for solid-phase synthesis.

Following amino acid synthesis, the compatibility of **3B** for solid-phase peptide synthesis was investigated. The coupling of **3B** appeared to be more difficult than TFA-lysine, as evidenced by multiple truncated peptides observed in crude HPLC traces. This could potentially be alleviated using higher equivalents and longer reaction times. Further optimization studies have not currently been developed.

A peptide was produced containing four monomer units of **3B**, yielding 36 fluorine nuclei within the molecule. ^{19}F NMR of the resulting peptide showed again four populations of resonances, however only spanning a total of 17 Hz from the furthest resonances (Figure 3-6). This amino acid possesses potential for further study, when considering that it possesses 36 fluorine residues with only four iterations of fluorinated amino acid as opposed to the required 12 that would be necessary for TFA-lysine.

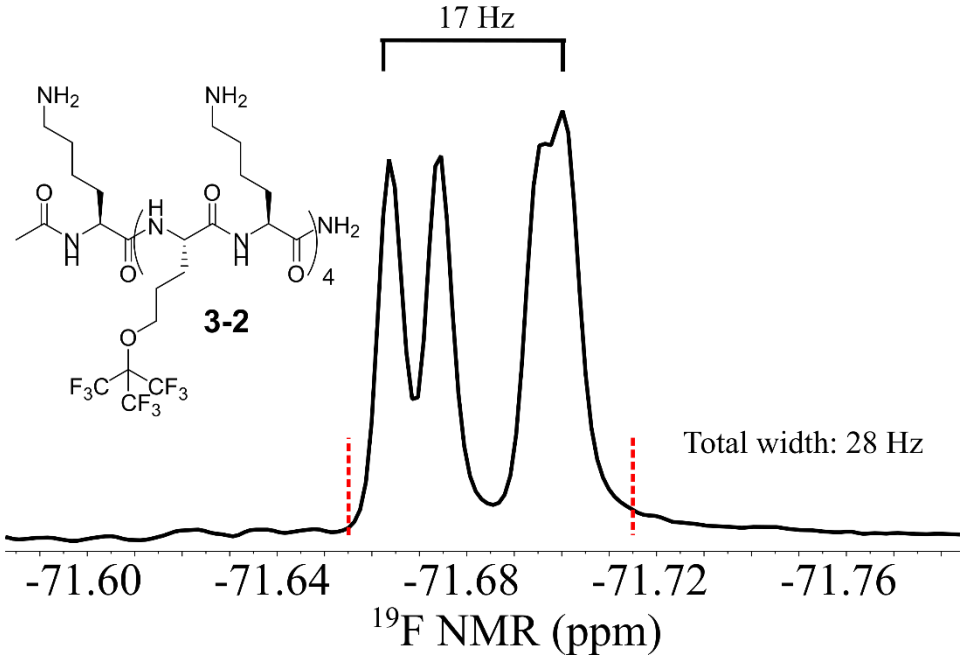


Figure 3-6: Sequence of **3-2** and corresponding ^{19}F NMR spectrum. Total resonance width is measured between the dashed lines.

*3.2.3 Bis(trifluoromethyl) benzyl lysine (**3C**)*

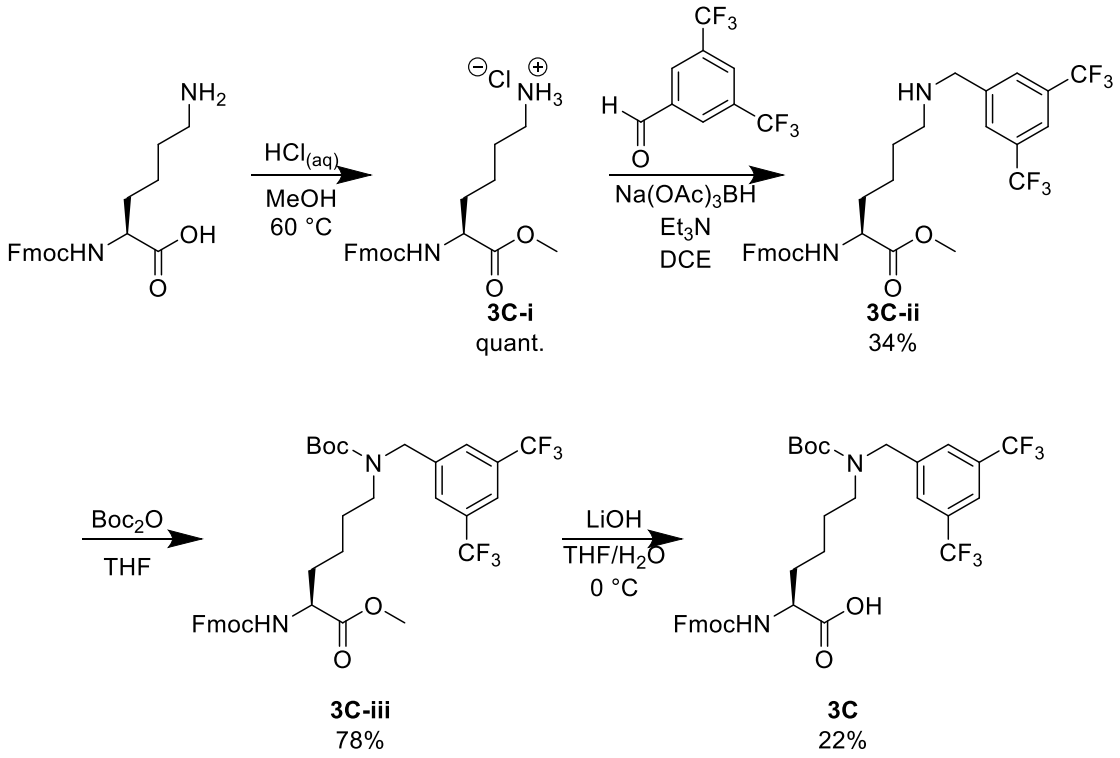


Figure 3-7: Synthesis of **3C**

Another amino acid scaffold investigated was based on magnetically equivalent aryl-CF_3 groups, providing six equivalent ^{19}F nuclei per residue, **3C**. On a per-residue basis, **3C** will provide less ^{19}F signal than **3B**, it has a key distinction of possessing an ionizable secondary amine. The hypothesis with this scaffold is the potential ability for shortening overall peptide sequences by having the fluorinated residue perform both the imaging function that TFA-lysine provided, as well as the water solubility and disorder promoting function of unlabeled lysine. A further inspiration for this is providing an overall reduction of charge in the imaging agent which will be discussed later.

The synthesis of **3C** begins with a methyl ester protection on the carboxylic acid of N-α Fmoc lysine (**3C-i**), followed by a reductive amination of the N-ε amine with 3,5-bis(trifluoromethyl) benzaldehyde to yield intermediate **3C-ii**. The resulting secondary amine is Boc protected (**3C-iii**), followed by a base-mediated hydrolysis of the methyl ester producing **3C** for peptide synthesis (Figure 3-7).

A short peptide sequence with four iterations of **3C** alternating with lysines was synthesized. The ^{19}F NMR spectrum of the peptide showed some dispersion of resonances; however, from peak to peak the total separation was only 15 Hz (Figure 3-8). The synthesis of the peptide was also facile, as the HPLC trace analyzing the crude synthesis showed only a single dominant peak indicating that the coupling efficiency of this fluorinated amino acid is rather high. The ease of synthesis, and narrow dispersion of resonances make peptides based on the bis(trifluoromethyl)benzyl lysine residue an attractive route for further study, and as such will be discussed in section 3.4.

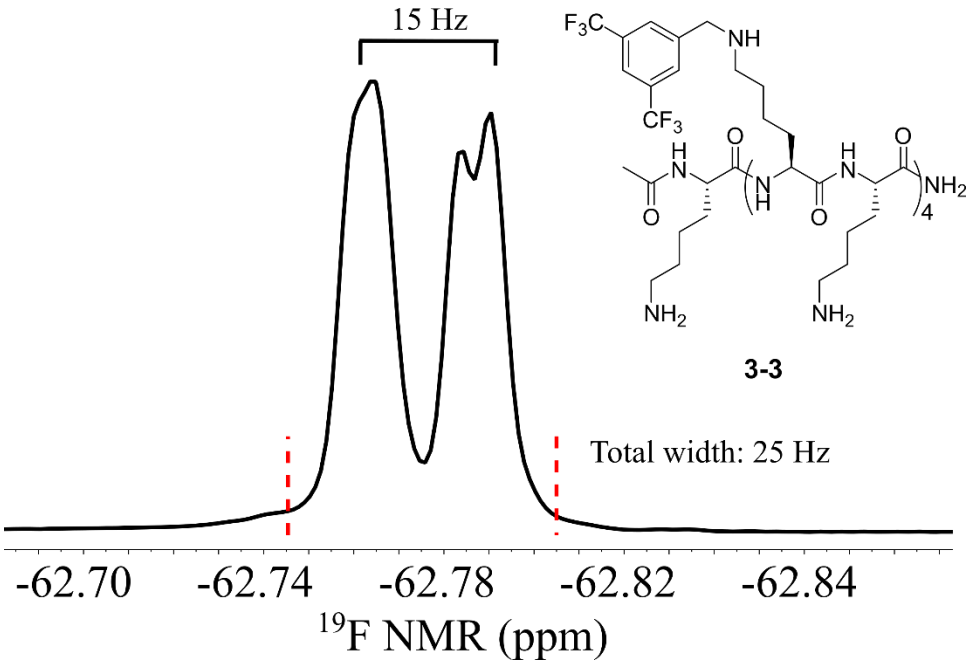


Figure 3-8: Sequence of **3-3** and its ^{19}F NMR spectrum. Total width is measured between the dashed lines.

3.3 Development of biomarker selective targeting agents

The goal for this work is to develop an imaging agent that could be used to target a specified disease biomarker. To this end, we have chosen EGFR as such a biomarker due to its overexpression in many cancers and its cell-surface location which make it an attractive target.[39,41,42,144–147] EGFR is a widely studied biomarker and as such there have already been great efforts to develop targeting agents for this protein. Both engineered proteins as well as short peptide sequences have been used successfully to target EGFR,

often coupled to imaging agents providing a fluorescent readout or radioactive metals for PET signal.[43,46,47,148–151] This section will discuss preliminary work using both an engineered protein domain, fibronectin, as well as EGFR targeting peptide sequences.

3.3.1 Fibronectin

The tenth human fibronectin type III domain is a small protein (MW ~ 10 kDa). This protein is highly stable and has been engineered to bind various molecular targets with high affinity (Figure 3-9). In particular, a mutant termed EI3.4.3 was engineered to bind EGFR with K_d = 250 pM and was coupled to a radioactive ^{64}Cu ligand to serve as a molecular imaging construct for use in PET studies.[49,48] Inspired by this approach, we worked in collaboration with Prof. Benjamin Hackel, the lead researcher who developed this technology with Prof. K. Dane Wittrup,[148] to attempt to translate the scaffold toward our desired application utilizing ^{19}F MRI.

Figure 3-9: Fibronectin scaffold (PDB: 1TTB). Modifiable loops are labeled as binding domain.

Due to the presence of lysines on the ^{19}F peptide scaffolds, we chose to seek out a thiol-mediated bioconjugation approach to append a fibronectin targeting domain to our imaging agent. Researchers in the Hackel lab produced a variant of fibronectin that possessed a mutation encoding for a single cysteine residue distal from the binding epitope. Attempts were made to conjugate the cysteine of this fibronectin scaffold to an N-terminal maleimide **3-5** (Figure 3-10), however no successful reaction was observed. Potentially, the cysteine on this protein scaffold may be positioned in such an orientation preventing the large peptide construct to approach for conjugation. Future attempts with this scaffold could involve using the same mutant that was coupled to a ^{64}Cu ligand, by reacting that protein's lone lysine with 3-maleimidopropyl-NHS ester in advance. The resulting product could be repurified, and another conjugation could be attempted with a ^{19}F-labeled peptide which possesses a cysteine on its N-terminus.

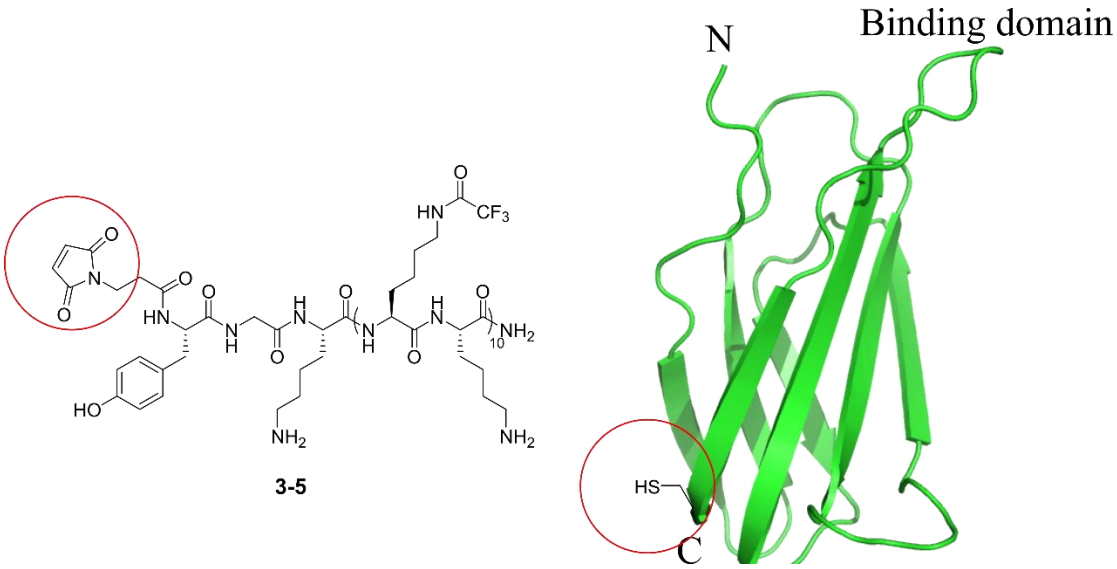


Figure 3-10: Bioconjugation pair between **3-5** (left, ^{19}F imaging agent) and cysteine-containing fibronectin mutant (right, EGFR targeting agent). Hetero-Michael reaction pair motifs are circled.

3.3.2 EGFR-targeting peptides

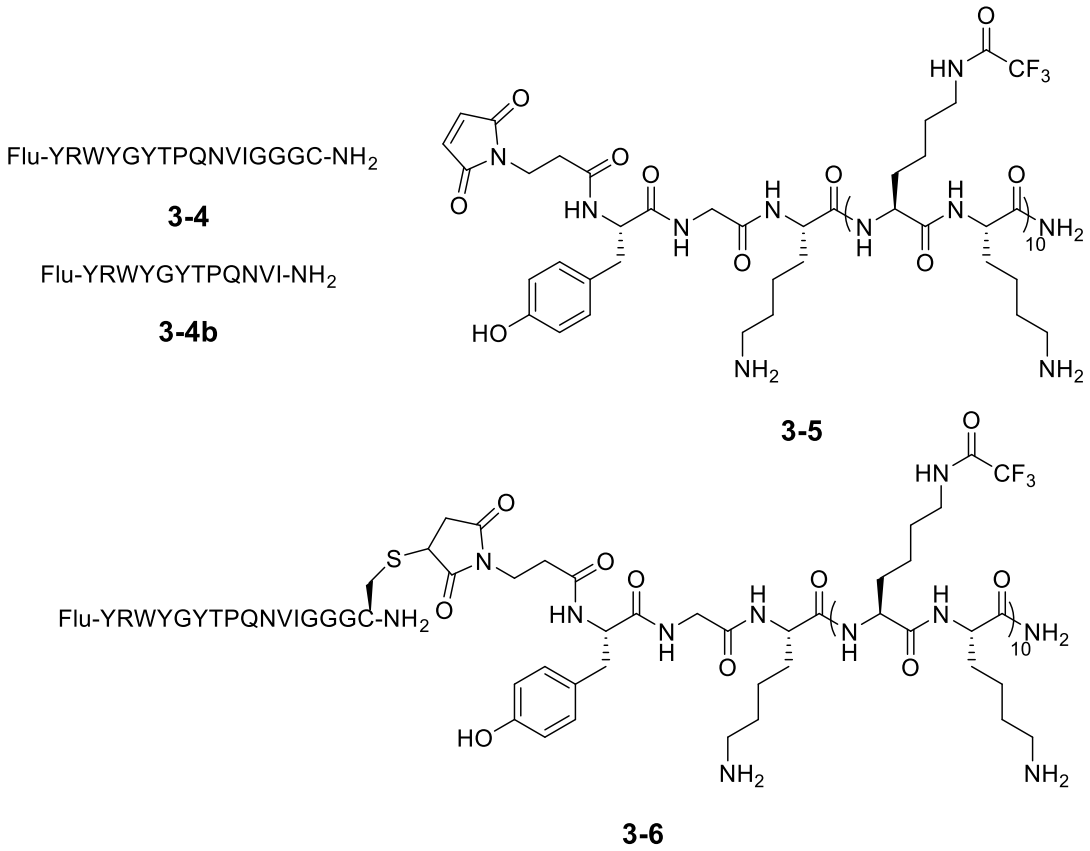


Figure 3-11: Sequences of EGFR targeting (**3-4**, and cysteine-free control **3-4b**) and ^{19}F imaging (**3-5**) peptides, as well as the conjugation product (**3-6**). Fluorescein (Flu) was installed using NHS chemistry on the N-terminus

An alternative strategy that was employed was the use of peptide sequences that have been shown to bind EGFR. We employed a peptide termed GE11 (here reported as a fluorescein-labeled variant, **3-4b**) with an arginine replacing the histidine which was determined to have improved binding to EGFR.[151] On the N-terminus of this peptide a fluorescein tag was installed, while on the C-terminus a cysteine was installed for bioconjugation, **3-4**. Successful alkylation of the C-terminal cysteine was achieved utilizing the maleimide containing ^{19}F peptide (Figure 3-10). The resulting construct was then incubated with two different cell lines: A431 epidermoid carcinoma cells (high EGFR expression, 10^5 copies/cell[152]) and SKOV-3 ovarian carcinoma (low EGFR expression, 10^2 copies/ cell[153]). The fluorescently labeled targeting agent without the inclusion of cysteine

was also incubated separately in both cell lines as a control. We observed via flow cytometry poor binding of the target alone at concentrations below 1 μM, and showed no preference for the high EGFR expressing A431 cells (Figure 3-12). It was hypothesized that the fluorescein tag may have inhibited binding as the C-terminal spacer motif was utilized in other studies successfully, albeit not with a cysteine. The construct of targeting agent and ^{19}F imaging agent however, bound non-specifically in a dose dependent manner across both cell lines (Figure 3-12). This result suggests that the ^{19}F imaging agent possesses cell adherence capability which must be addressed before any logical targeting efforts can be undertaken. Our hypothesis after these experiments was the polycationic nature of the ^{19}F imaging peptide was responsible for non-specific cell adherence. Efforts were thus begun to redesign the scaffold to achieve a reduction in overall charge.

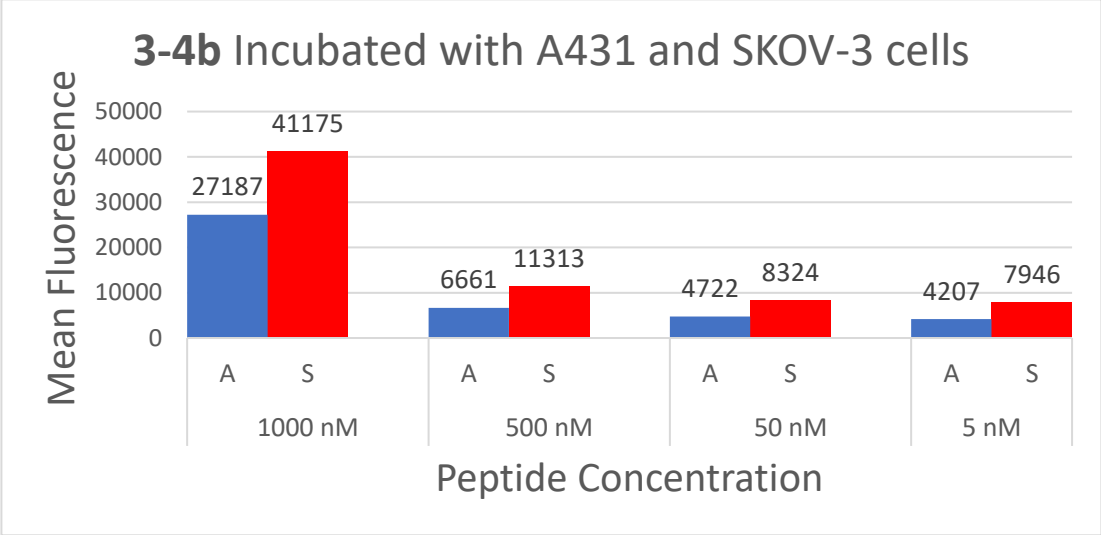


Figure 3-12: Flow cytometry analysis of **3-4b** with A431 (A) and SKOV-3 (S) cells at four concentrations.

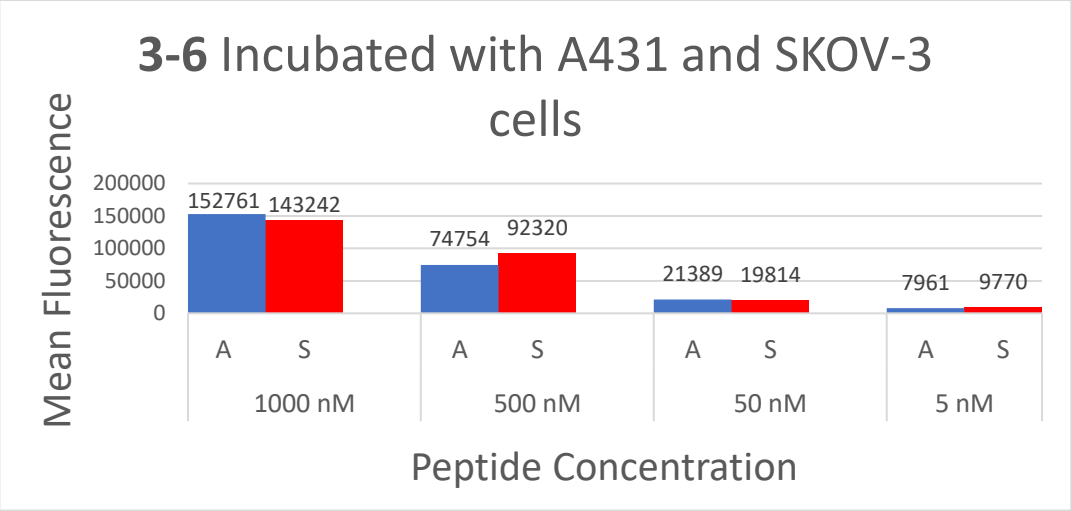


Figure 3-13: Flow cytometry analysis of **3-6** with A431 (A) and SKOV-3 (S) cells at four concentrations.

3.4 Examining the charge of ^{19}F imaging peptides

Following the preliminary EGFR-targeting experiments, it was clear that further optimization of the ^{19}F-labeled peptide scaffold beyond simply its NMR signal and water solubility would be required. The hypothesis for the observed non-specific cell adherence was the polycationic nature of the peptide. The surface of cells is known to be anionic due in part to the phospholipids comprising the cell membrane. A recent study was performed by Zheng Cui and coworkers involving the use of cationic magnetic nanoparticles which were used to investigate cell surface charge.[154] In this work, Cui and co-workers established an array of cell lines, and showed that only cancerous cells adhered to positively charged magnetic nanoparticles, suggesting an ionic interaction and an enrichment of anionic character on the surface of cancerous mammalian cells. To assess this hypothesis, two approaches were taken: determine whether the peptides were adhering to the cell surface or being internalized and making changes to the peptide scaffold to reduce the overall charge of the biopolymer.

3.4.1 Determining cell-surface adhesion

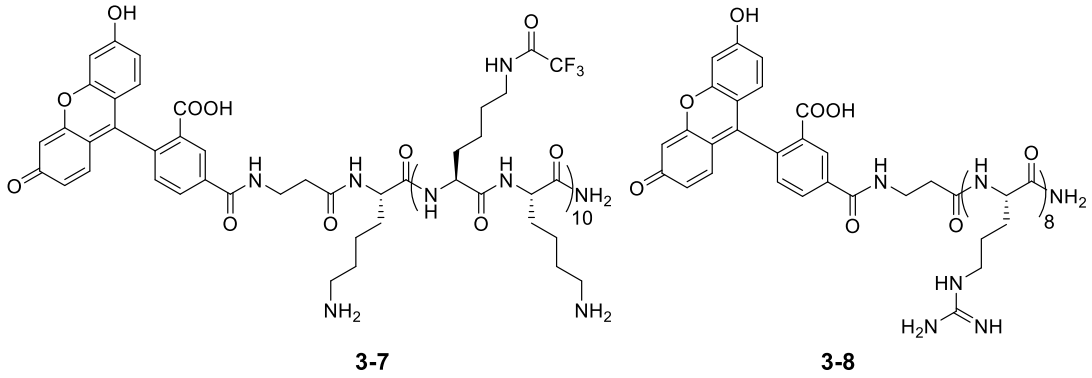


Figure 3-14: Sequences of **3-7** and **3-8**, used for confocal microscopy experiments

Flow cytometry allows for a quantification of peptide interacting with the cell lines studied, however it is difficult to know with certainty the location of the interaction. Cell-surface adhesion would look very similar to internalization of the fluorescently labeled peptide. To more clearly understand the interaction of peptide and cells, confocal microscopy was employed to observe the location of the fluorescent tag. We also used a known cell-internalizing Arg_8 peptide (**3-8**)[155], conjugated to fluorescein for a positive control of internalization (Figure 3-14). HeLa cells were incubated with the ^{19}F labeled peptide, Arg_8 peptide, as well as a vehicle control. Following Hoechst nucleus staining the cells were imaged using an inverted confocal microscope. As seen in Figure 3-15, the ^{19}F labeled peptide is adhering specifically on the cell surface. No punctate fluorescence is observed as opposed to the Arg_8 peptide, which is the likely outcome through endocytosis. This experiment provided strong evidence that a change in the sequence was required to alleviate cell surface adherence.

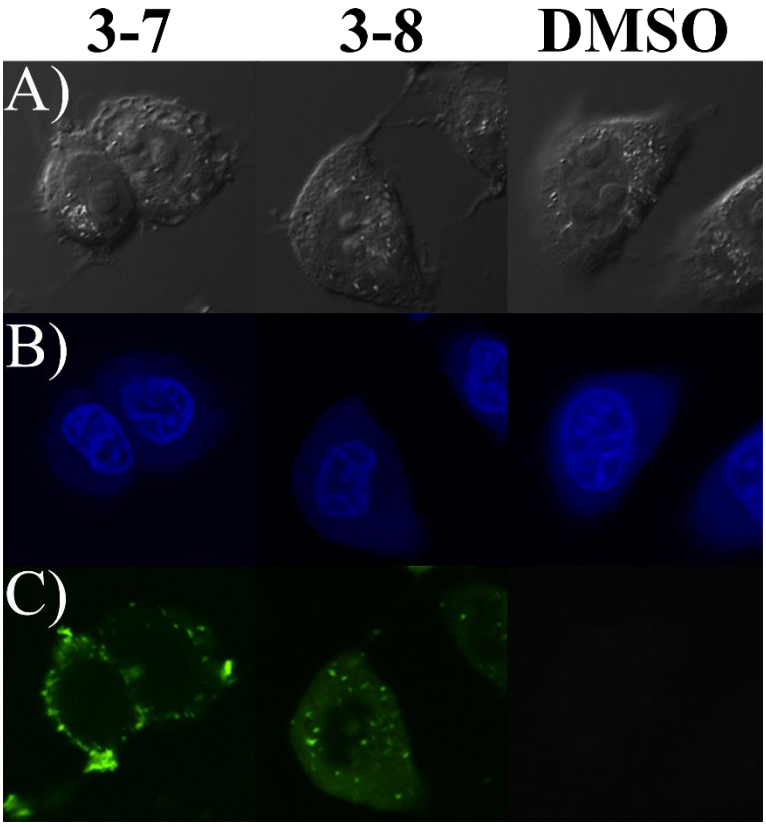
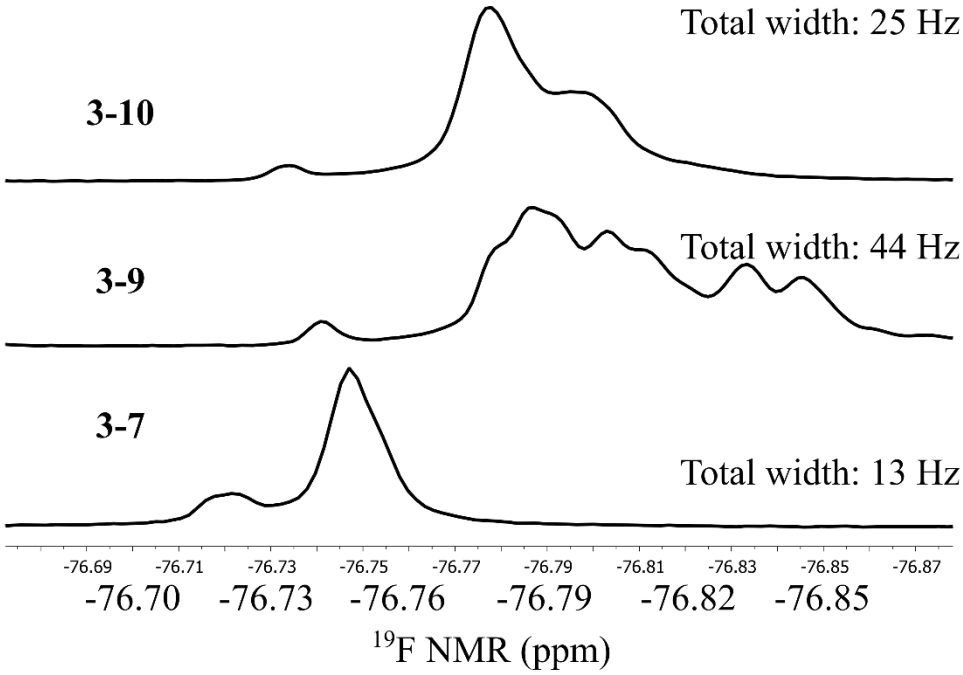


Figure 3-15: Confocal microscopy images of HeLa cells following incubation with fluorescein-labeled ^{19}F peptide (**3-7**), cell internalization positive control (**3-8**), and DMSO vehicle control. A) Bright field image of HeLa cells. B) Hoechst nuclear stain. C) Fluorescein detection

3.4.2 Reduction in peptide charge

Lysine was originally chosen to compliment TFA-lysine due to its ease of ionization at physiological pH would help counteract the hydrophobic nature of multiple fluorinated residues, aiding in water solubility of the scaffold. This ionization would further aid in promoting disorder of the peptide through electrostatic repulsion, leading to a degenerate ^{19}F NMR resonance. Through the original report of this peptide scaffold, little work was done to study how much lysine was necessary to provide the aforementioned benefits. To this end, sequences of TFA-lysine containing peptides were synthesized, incorporating a reduced number of lysine residues to lower the net charge.

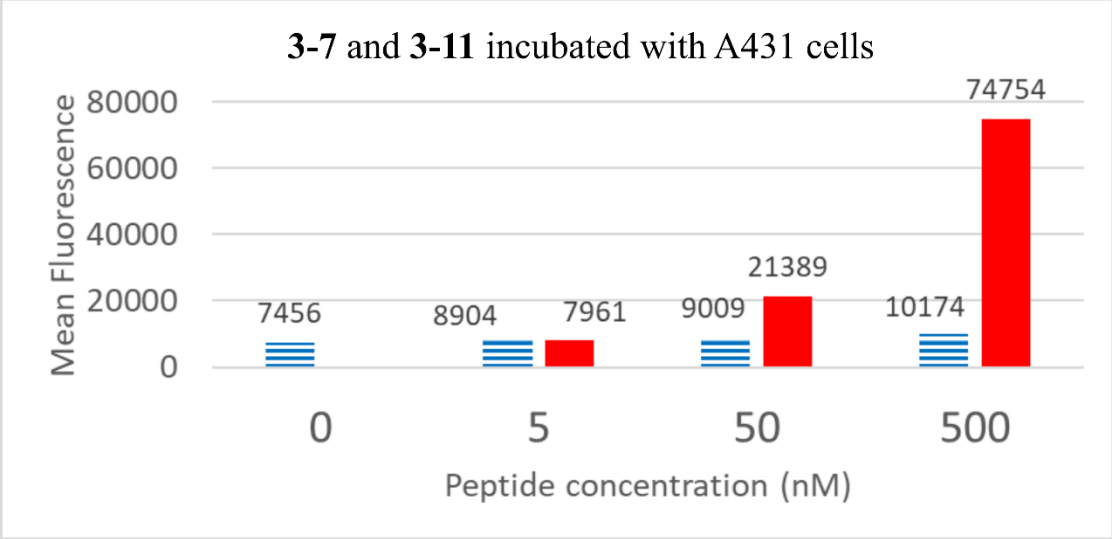
In a first study, three peptides were synthesized, all containing 10 instances of TFA-lysine and 11 instances of accompanying amino acids. In one case, all accompanying residues were lysine. In the other two sequences, three of those lysines were replaced with glutamine, and glutamic acid, respectively. This series was undertaken to explore how a stepwise reduction to approximately half of the overall charge would impact ^{19}F NMR resonances as well as water solubility. In all three cases, the peptides remained water soluble however the ^{19}F NMR resonance in both the Gln_3 and Glu_3 cases was observed to broaden and resolve into multiple distinct resonances (Figure 3-16). Introducing these two residues into the sequence of the peptide may be causing local changes in the chemical environment, causing a shift in the resonance contribution from neighboring TFA-lysines. This is especially evident in the Glu_3 case as the resonances span nearly twice the spectral range as the Gln_3 case, potentially due to local negative charge or a salt bridge interaction within the peptide. Of note, appending fluorescein on the N-terminus of the peptide sequence produces an apparent secondary resonance downfield of the main ^{19}F signal



3-10: Flu-βAla-KK*KK*QK*KK*KK*QK*KK*KK*QK*KK*K-NH_2
3-9: Flu-βAla-KK*KK*EK*KK*KK*EK*KK*KK*EK*KK*K-NH_2
3-7: Flu-βAla-KK*KK*KK*KK*KK*KK*KK*KK*KK*KK*K-NH_2

Figure 3-16: ^{19}F NMR spectra of a series of peptides with reduced lysine content upon replacing residues with E or Q. K* refers to TFA-lysine.

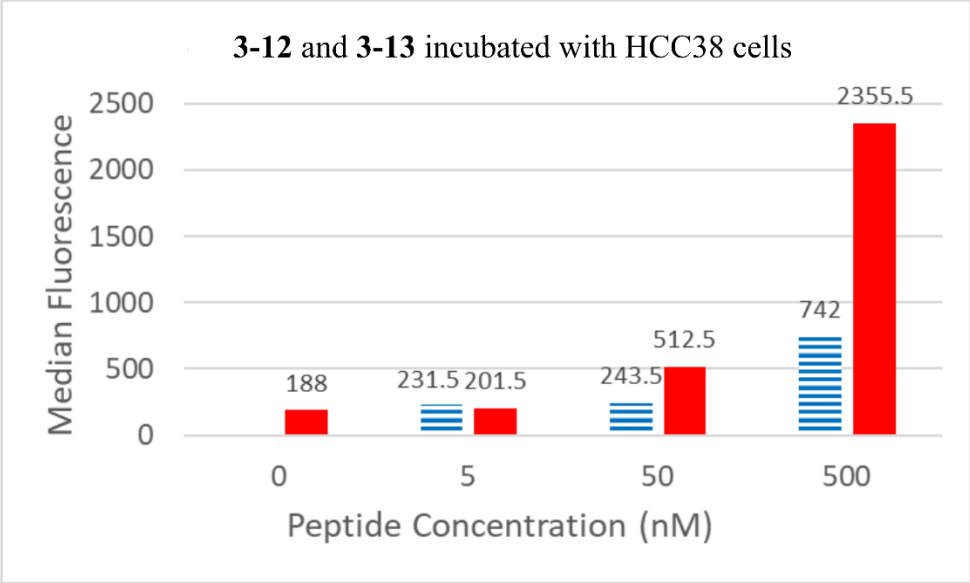
A similar study was then undertaken, where instead of incorporating Glu or Gln to reduce charge a stepwise replacement of lysines for glycines was performed. It was hypothesized that due to the flexibility of the glycine residue that disorder could be promoted without relying solely on electrostatic repulsion. Glycine is still a relatively polar amino acid, but an experiment was undertaken to determine what the impact of removing charge might be on the water solubility of a highly fluorinated peptide. Three sequences were synthesized, one with 11 lysines accompanying 10 TFA-lysines, and two others with five or six of these lysines replaced by glycine respectively. In the case of having six lysines replaced (**3-11b**), solubility was reduced, and the resulting ^{19}F NMR spectrum showed two distinct resonances, spanning over 20 Hz (See section 3.6.6). When one fewer lysine was exchanged (**3-11**), a single narrow (13 Hz) resonance was observed and water solubility was easily attained. A general baseline was then established as the removal of six of lysines showed drastically different physical and spectral characteristics than removing five.



3-7: Flu-βAla-KK*KK*KK*KK*KK*KK*KK*KK*KK*KK*K-NH_2
3-11: Flu-βAla-KK*GK*KK*GK*KK*GK*KK*GK*KK*GK*K-NH_2

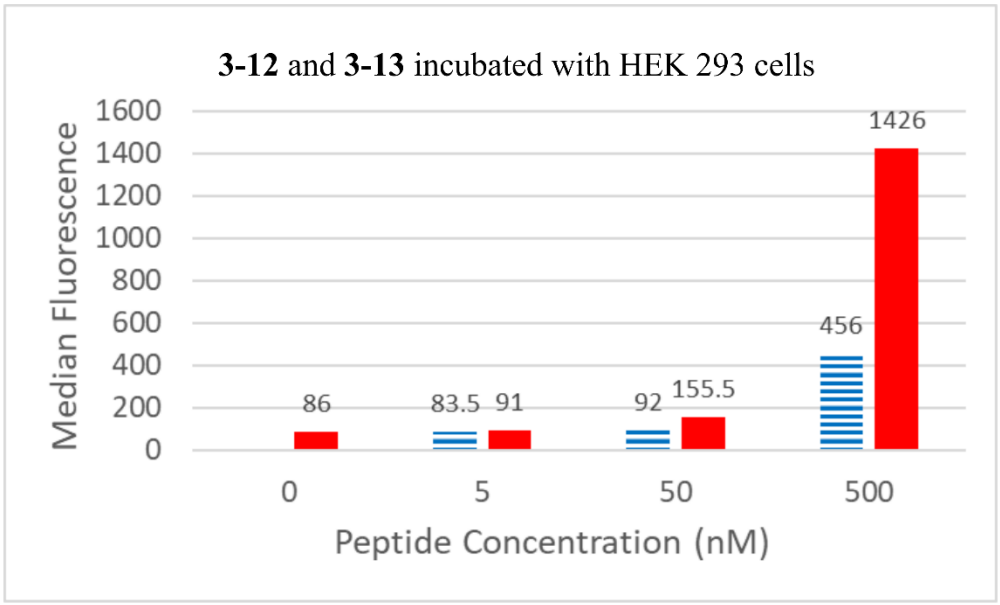
Figure 3-17: Flow cytometry analysis of **3-7** (solid bar) and **3-11** (lines) incubated with A431 cells

After successfully reducing the charge of the peptide from +11 to +6, it needed to be determined if this has an impact on cell surface adherence. Two fluorescently labeled peptides were synthesized, the first resembling the ^{19}F imaging peptide used in the previous cell experiments, with 10 TFA-lysines accompanied by 11 lysines. The second peptide uses the newly discovered replacement of five of these lysines with glycine (**3-11**). Both peptides were incubated again with A431 cells and analyzed using flow cytometry (Figure 3-17). Encouragingly, **3-11** showed a near baseline fluorescence reading as opposed to a dose-dependent increase in fluorescent signal seen in the highly charged peptide **3-7**. This experiment was repeated using two different cell lines, HCC38 breast cancer cells (Figure 3-18), and HEK 293T noncancerous kidney cells (Figure 3-19). In an attempt to lower the impact of the signal from autofluorescence, TAMRA was used as a fluorescent tag as opposed to fluorescein. The goal of this experiment was to both reproduce the data observed with A431 cells, as well as to probe whether cancerous vs non-cancerous cells had an impact on cell surface adherence. Following flow analysis of the two peptides, it was again observed that the lowered charge of the peptide decreased the non-specific adherence, but that the identity of the cell line in this case did not show a convincing difference. At 500 nM in this case, some adherence was observed even in the lowered charge peptide, **311**. A potential explanation for this could be the addition of TAMRA as the fluorophore, a cationic fluorescent tag, contributing to adherence to the cell surface.



3-12: TAMRA-βAla-KK*KK*KK*KK*KK*KK*KK*KK*KK*KK*KK*K-NH_2
3-13: TAMRA-βAla-KK*GK*KK*GK*KK*GK*KK*GK*KK*GK*KK*K-NH_2

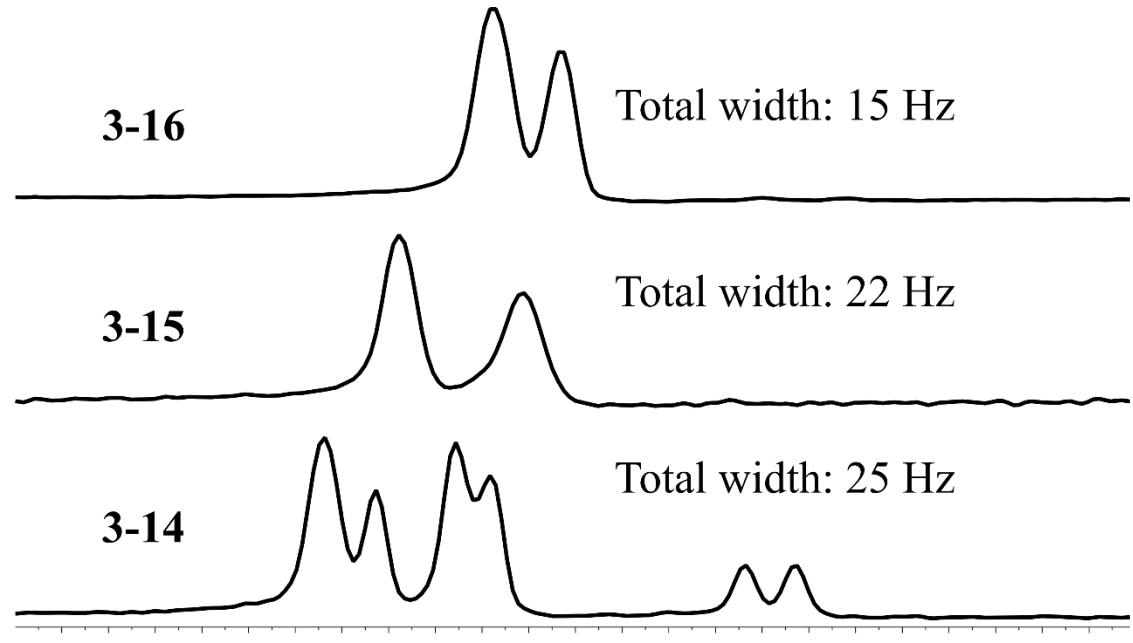
Figure 3-18: Flow cytometry analysis of **3-12** (solid bars) and **3-13** (lines) incubated with HCC-38 cells.



3-12: TAMRA-βAla-KK*KK*KK*KK*KK*KK*KK*KK*KK*KK*KK*K-NH_2
3-13: TAMRA-βAla-KK*GK*KK*GK*KK*GK*KK*GK*KK*GK*KK*K-NH_2

Figure 3-19: Flow cytometry analysis of **3-12** (solid bars) and **3-13** (lines) incubated with HEK-293 cells.

To explore the generalities of our findings that reduced charge lowers the non-specific cell adherence of ^{19}F imaging peptides using TFA-lysine, we tested another fluorinated residue. Three peptide sequences utilizing the previously described bis(trifluoromethyl)benzyl lysine (**3C**) were synthesized to observe how changes in the sequence would impact the observed NMR signal. The sequences all incorporated five iterations of the amino acid to yield a total of 30 fluorine nuclei per peptide. These residues were alternated with lysine in the first peptide, and in the other two a systematic replacement of half of the lysines, and all the lysines, with glycine. The goal is to determine if the reduction in charge from lysine will impact the peptide's ability to unfold and to retain water solubility. The three sequences all showed multiple populations of ^{19}F resonances, and encouragingly the sequence in which all extraneous lysines were replaced with glycine shows the most narrow dispersion of resonances, and retains its water solubility (Figure 3-20). This result will allow us to lower the overall charge of the peptide sequences by having the fluorinated amino acid supply the charge to the peptide. These peptides with high fluorine content and lower extraneous charge from non-fluorinated residues will be interesting candidates for future cell adherence studies.



3-16: Ac-YGGK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}G-NH_2
3-15: Ac-YGKK^{C}GK^{C}KK^{C}GK^{C}KK^{C}GK^{C}KK^{C}GK^{C}KK^{C}GK^{C}K-NH_2
3-14: Ac-YGKK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}K-NH_2

Figure 3-20: ^{19}F NMR overlay of peptides containing **3C** (K^{C}) and increasing iterations of glycine replacing lysine. Total resonance width measurements from baseline to baseline, ignoring additional upfield resonance visible in **3-14**.

3.5 Conclusions and future outlook

Following the original 2017 report, we have designed and synthesized several new ^{19}F-labeled peptide sequences to further ^{19}F MRI probe development. Three additional fluorinated amino acids have been explored which provide interesting routes for future probe design. In particular, peptides based on the bis(trifluoromethyl)benzyllysine (**3C**) hold promise for increasing ^{19}F signal on a per-residue basis, while providing water-solubility and lowering overall probe molecular weights. Targeting agents with specificity to EGFR have been explored in a preliminary manner, but future works should include optimization of reaction conditions or protein scaffolds based on the fibronectin domain. Currently, preliminary work is being explored utilizing chemically self-assembled protein

nanorings, which can displayed the fibronectin domain, for development of bioconjugation reactions with ^{19}F-labeled peptides. Non-specific cell surface adherence has been addressed by reducing the net charge of ^{19}F-labeled peptides to minimize ionic interactions. The second generation of ^{19}F MRI probes should be developed based on these recently explored concepts.

3.6 Experimental

3.6.1 Methods:

All peptides were synthesized and purified as described in Chapter 2.[27] All NMR and MALDI-MS analyses were also performed using the same instruments and parameters as described in Chapter 2. Any reagents used were commercially available and used without further purification. Sodium perfluoro tertbutoxide was prepared according to published protocols.[156]

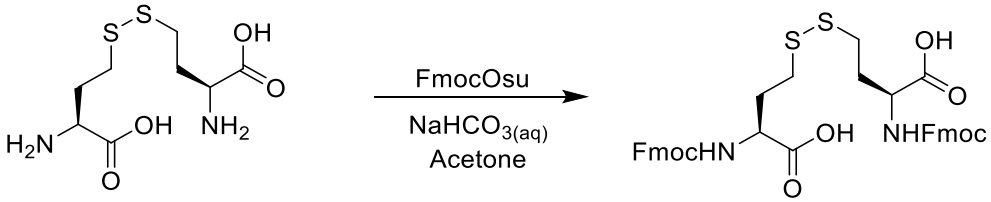
3.6.2 Confocal microscopy:

100,000 HeLa cells were plated into each of six glass-bottomed coated culture dishes in a total volume of 2 mL. The cells were incubated overnight at 37 °C. The following morning, cells were approximately 40% confluent and the media was aspirated. The cells were rinsed 2 X with PBS and were dosed with fluorescent peptide samples (5 μM in PBS). The cells were incubated with fluorescent peptide for 60 min, followed by aspiration and rinsing. Fixing was performed using 1 mL of 4% paraformaldehyde and light shaking for 10 min. The formaldehyde solution was removed and the cells were rinsed 2 X with PBS. Hoechst 34580 (1 mL of 1 μg/mL) was added to the cells to stain the nucleus, followed by shaking lightly for 10 min. The cells were aspirated and rinsed 2 X with PBS followed by addition of 2 mL of PBS. The cells were then imaged using an Olympus FV1000 IX2 Inverted Confocal Microscope with the assistance of the University of Minnesota Imaging Center.

3.6.3 Flow cytometry:

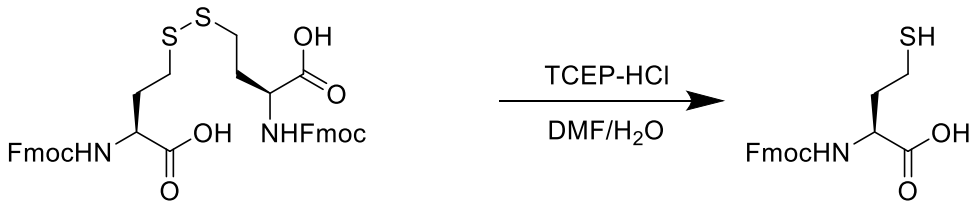
A general procedure for flow cytometry experiments is as follows. Cells (300,000-500,000) were pelleted at 300 x *g* for 5 min. Culture media was decanted, and cells were rinsed with 1 mL of PBS followed by centrifugation and decanting. Fluorescent peptide samples were dosed at varied concentrations (0, 5 nM, 50 nM, 500 nM, 1000 nM) and incubated for 90 min at either ambient temperature, or 4 °C for EGFR-specific experiments to prevent receptor turnover. Peptide solutions were then centrifuged, decanted, and rinsed with PBS followed by centrifugation and decanting. PBS (200 μL) was then added and flow cytometry was performed using either a BD-Accuri C6 (Fluorescein-detection) or BD Fortessa X-20 (TAMRA-detection) with the assistance of the University of Minnesota Flow Cytometry Resource.

3.6.4 Small molecule synthesis



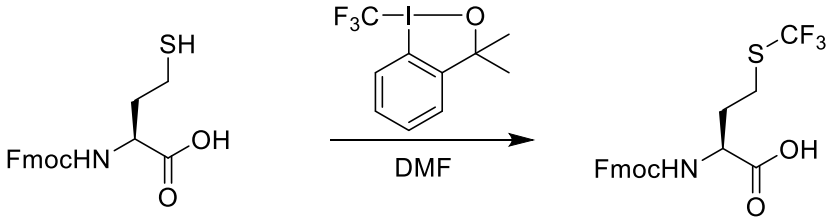
3A-i

L-homocystine (500 mg, 1.86 mmol, 1 eq) was mixed with 10% $NaHCO_3$ to form a turbid suspension. A solution of Fmoc-*N*-hydroxysuccinimide (1.32 g, 3.91 mmol, 2.1 eq) in acetone was added dropwise to the vigorously stirring solution. The white suspension was allowed to stir overnight, followed by removal of acetone by reduced pressure. EtOAc was added and the pH of the aqueous phase was then adjusted to ~3 using 1M HCl. The aqueous phase was extracted 3 X with EtOAc, followed 1 X with water, 1 X with brine, and dried over $MgSO_4$. Following filtration and removal of solvent under reduced pressure yielded **3A-i** as a white solid (0.700 g, 0.982 mmol, 53%). ^{1}H NMR (500 MHz, Acetone-d_6) δ 7.87 – 7.82 (m, 2H), 7.74 – 7.67 (m, 2H), 7.47 – 7.36 (m, 2H), 7.35 – 7.26 (m, 2H), 6.86 (d, *J* = 8.6 Hz, 1H), 4.43 (td, *J* = 8.9, 4.5 Hz, 1H), 4.34 (d, *J* = 7.0 Hz, 2H), 4.23 (t, *J* = 7.2 Hz, 1H), 2.88 (tt, *J* = 14.3, 7.0 Hz, 3H), 2.45 – 2.28 (m, 1H), 2.15 (dtd, *J* = 14.2, 8.7, 5.9 Hz, 1H).



3A-i **3A-ii**

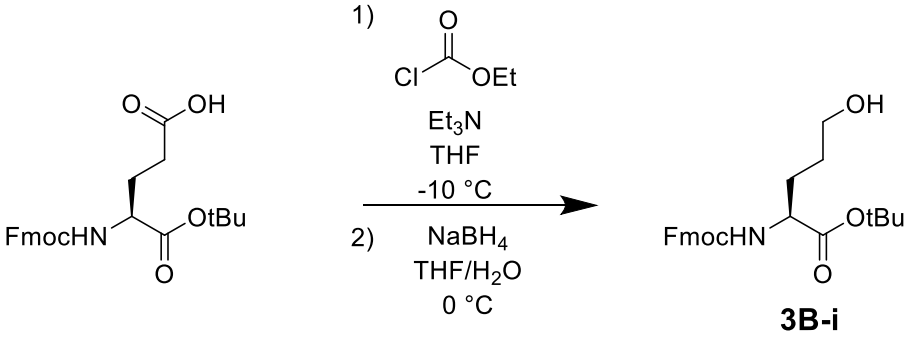
3A-i (0.700 g, 0.98 mmol, 1 eq) was dissolved in a 9:1 solution of DMF and water. Tris-carboxyethylphosphine hydrochloride (0.422 g, 1.47 mmol, 1.5 eq) was added portionwise. The solution was stirred vigorously overnight, the crude material was diluted in EtOAc and washed 3 x with water, 1 x with brine, and dried over $MgSO_4$. The product, **3A-ii**, was isolated after filtration and removal of solvent as a yellow oil (0.438 g, 1.23 mmol, 63%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.79 (d, *J* = 7.5 Hz, 2H), 7.61 (d, *J* = 7.2 Hz, 2H), 7.43 (t, *J* = 7.4 Hz, 2H), 7.34 (t, *J* = 7.5 Hz, 2H), 5.32 (d, *J* = 8.2 Hz, 1H), 4.61 (q, *J* = 7.9 Hz, 1H), 4.48 (d, *J* = 5.2 Hz, 2H), 4.25 (t, *J* = 6.6 Hz, 1H), 2.62 (q, *J* = 7.4 Hz, 2H), 2.22 (dd, *J* = 13.4, 5.7 Hz, 1H), 2.04 (dd, *J* = 14.2, 7.3 Hz, 1H), 1.57 (t, *J* = 8.2 Hz, 1H). LRMS (ESI-TOF) calculated for $C_{19}H_{18}NO_4S^-$ [M-H]$^-$: 356.1, observed 356.0.



3A-ii **3A**

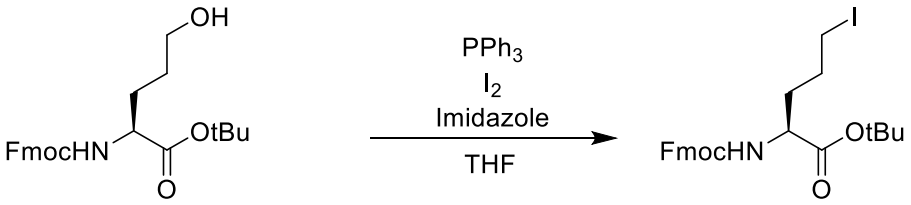
3A-ii (223 mg, 0.624 mmol, 1 eq) was dissolved in 2 mL of DMF, followed by addition of 3,3-dimethyl-1-(trifluoromethyl)-1,2-benziodoxole (226 mg, 0.685 mmol, 1.1 eq) and the reaction mixture was stirred at ambient temperature for 18 h. The solvent was removed, and the mixture was purified via flash column chromatography (1:1 hexanes/EtOAc + 1% AcOH) to afford **3A** as a white solid (125 mg, 47%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.77 (d, *J* = 7.6 Hz, 2H), 7.59 (d, *J* = 7.1 Hz, 2H), 7.41 (t, *J* = 7.5 Hz, 2H), 7.32 (t, *J* = 7.4 Hz, 2H), 5.35 (d, *J* = 7.9 Hz, 1H), 4.50 (dd, *J* = 10.3, 6.8 Hz, 2H), 4.47 – 4.39 (m, 1H), 4.22 (t, *J* = 6.4 Hz, 1H), 2.92 (q, *J* = 6.6, 6.1 Hz, 2H), 2.34 (d, *J* = 7.0 Hz, 1H), 2.08 – 2.01

(m, 1H). ^{19}F NMR as a mixture of rotational isomers (471 MHz, Chloroform-*d*) δ -41.02, -41.10. LRMS (ESI-TOF) calculated for $C_{20}H_{17}F_3NO_4S^-$ [M-H]$^-$: 424.1, observed 424.2.



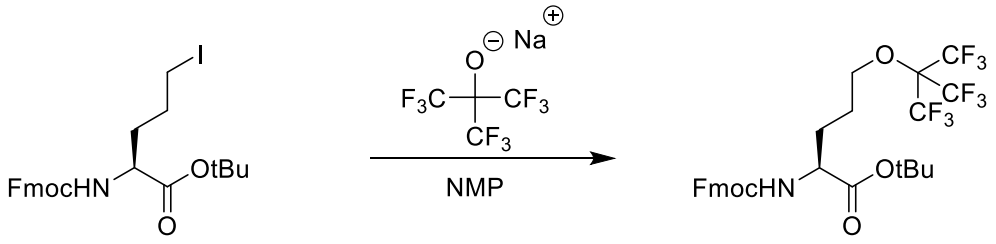
3B-i

Fmoc-L-glutamic acid-α tert butyl ester (1.006 g, 2.364 mmol, 1 eq) was dissolved in 4 mL of anhydrous THF. The solution was cooled to -10 °C in a NaCl-ice water bath, followed by slow addition of triethylamine (363 μL, 2.607 mmol, 1.1 eq). Separately, ethyl chloroformate (250 μL, ρ_{25}= 1.135, 2.615 mmol, 1.1 eq) was dissolved in 2 mL of anhydrous THF and added slowly to the first solution. The mixture became white and heterogenous and was stirred at -10 °C for 90 min, followed by gravity filtration. Separately, sodium borohydride (193.8 mg, 5.123 mmol, 2.2 eq) was suspended in 2 mL of H_2O and cooled to 0 °C. Slowly, the filtrate was added to the borohydride solution was stirred at 0 °C for 30 min. The mixture was quenched with 1M HCl and diluted with EtOAc. The organic layer was washed 3 X with $NaHCO_3$, followed by water and brine then dried over $MgSO_4$. Following filtration and removal of solvent, the mixture was purified on a Combiflash Rf system (hexanes/EtOAc) to afford **3B-i** as a clear oil (357.2 mg, 37%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.76 (d, *J* = 7.6 Hz, 2H), 7.60 (d, *J* = 7.5 Hz, 2H), 7.40 (t, *J* = 7.5 Hz, 2H), 7.31 (t, *J* = 7.5 Hz, 2H), 5.52 (d, *J* = 8.2 Hz, 1H), 4.39 (d, *J* = 7.2 Hz, 2H), 4.30 (q, *J* = 6.8 Hz, 1H), 4.22 (t, *J* = 7.0 Hz, 1H), 3.68 (t, *J* = 6.2 Hz, 2H), 2.07 – 1.88 (m, 2H), 1.76 (tt, *J* = 14.0, 6.7 Hz, 1H), 1.62 (m, 2H), 1.48 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 171.58, 155.99, 143.89, 143.74, 141.27, 127.66, 127.02, 125.08, 119.94, 82.24, 66.88, 62.17, 53.94, 47.17, 31.17, 29.56, 28.08, 27.98. HRMS (ESI-TOF) calculated for $C_{24}H_{29}NO_5H^+$ [M+H]$^+$: 434.1943, observed 434.1908.



3B-i **3B-ii**

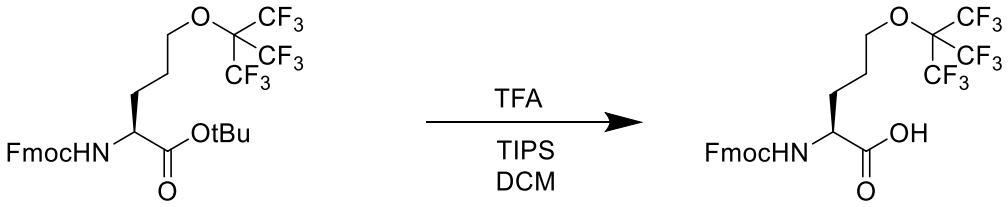
Iodine (267.2 mg, 1.053 mmol, 1.3 eq) was dissolved in 3 mL of THF, followed by addition of triphenylphosphine (277.0 mg, 1.056 mmol, 1.3 eq) and imidazole (72.4 mg, 1.063 mmol, 1.3 eq). The mixture was stirred for 15 min, followed by addition of **3B-i** (333.2 mg, 0.8097 mmol, 1 eq) dissolved in 2 mL of THF. The solution was allowed to stir for 2 h, followed by dilution with EtOAc, and washing 3 X with H_2O. The organic layer was washed with brine, dried over $MgSO_4$, and the solvent was removed after filtration followed by purification on a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **3B-ii** as a brown oil (288.0 mg, 68%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.77 (d, *J* = 7.5 Hz, 2H), 7.60 (d, *J* = 7.5 Hz, 2H), 7.41 (t, *J* = 7.5 Hz, 2H), 7.32 (td, *J* = 7.5, 2.8 Hz, 2H), 5.35 (d, *J* = 8.1 Hz, 1H), 4.41 (qd, *J* = 10.7, 7.0 Hz, 2H), 4.32 – 4.25 (m, 1H), 4.22 (t, *J* = 6.9 Hz, 1H), 3.26 – 3.20 (m, 1H), 3.17 (tt, *J* = 9.8, 4.9 Hz, 1H), 1.93 (m, 2H), 1.79 (m, 2H), 1.49 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 171.12, 155.77, 143.87, 143.71, 141.30, 127.69, 127.05, 125.09, 125.03, 119.98, 119.96, 82.54, 66.88, 53.38, 47.19, 33.69, 28.95, 28.01, 5.73. HRMS (ESI-TOF) calculated for $C_{24}H_{28}INO_4H_+$ $[M+H]^+$: 544.0961, observed 544.0893.



3B-ii **3B-iii**

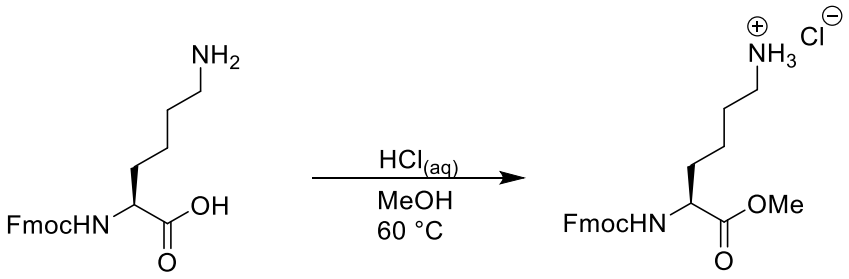
3B-ii (288.0 mg, 0.5524 mmol, 1 eq) was dissolved in 2 mL of NMP, followed by addition of sodium perfluorotertbutoxide (157 mg, 0.608 mmol, 1.1 eq). The mixture was sonicated briefly to aid solubility, and allowed to stir at rt for 18h. Upon completion, the mixture

was diluted with EtOAc, and washed 3 X with H_2O followed by brine, dried over $MgSO_4$, filtered, and concentrated. The material was purified on a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **3B-iii** as a white solid (202.6 mg, 58%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.77 (d, *J* = 7.3 Hz, 2H), 7.60 (dd, *J* = 7.6, 2.4 Hz, 2H), 7.40 (td, *J* = 7.5, 2.7 Hz, 2H), 7.32 (tt, *J* = 7.4, 1.6 Hz, 2H), 5.37 (d, *J* = 8.0 Hz, 1H), 4.44 – 4.38 (m, 2H), 4.30 (q, *J* = 6.8, 6.3 Hz, 1H), 4.22 (t, *J* = 6.9 Hz, 1H), 4.07 – 4.02 (m, 2H), 1.97 (tq, *J* = 15.0, 5.1 Hz, 1H), 1.81 – 1.64 (m, 2H), 1.48 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 171.19, 155.84, 143.88, 143.71, 141.32, 127.70, 127.03, 125.08, 125.03, 119.98, 119.95, 82.50, 68.95, 66.88, 53.62, 47.20, 37.01, 33.70, 28.86, 27.89, 25.39. ^{19}F NMR (471 MHz, Chloroform-*d*) δ -70.37. HRMS (ESI-TOF) calculated for $C_{28}H_{28}F_9NO_5Na^+$ $[M+Na]^+$: 652.1721, observed 652.1695.



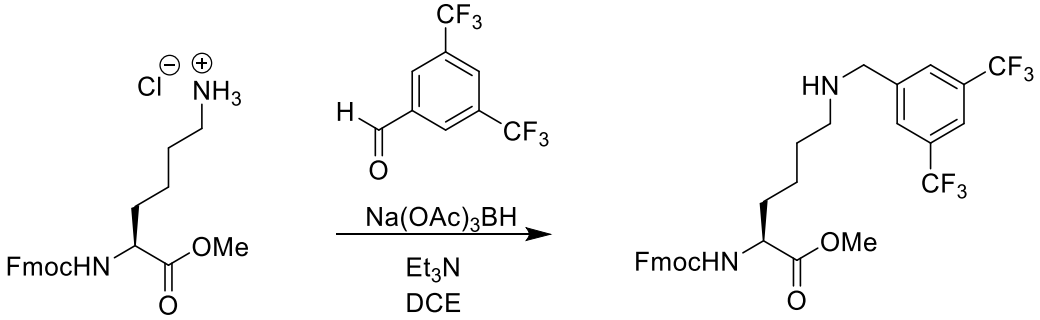
3B-iii **3B**

3B-iii (178.0 mg, 0.2828 mmol, 1 eq) was dissolved in 4 mL of 50/50 DCM and TFA, followed by addition of 25 μL of triisopropylsilane. The reaction was stirred at rt for 3h, followed by removal of solvent. The crude mixture was dissolved in EtOAc and washed 3 X with 1 M HCl, followed by brine and dried over $MgSO_4$. The solvent was removed after filtration to afford **3B** as a white solid (102.4 mg, 63%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.66 (s, 1H), 7.90 (d, *J* = 7.6 Hz, 2H), 7.73 (dd, *J* = 7.5, 3.7 Hz, 2H), 7.70 (d, *J* = 8.2 Hz, 1H), 7.42 (t, *J* = 7.4 Hz, 2H), 7.33 (t, *J* = 7.4 Hz, 2H), 4.34 – 4.29 (m, 2H), 4.23 (t, *J* = 7.1 Hz, 1H), 4.09 (q, *J* = 5.8 Hz, 2H), 3.98 (td, *J* = 8.5, 4.9 Hz, 1H), 1.84 (dq, *J* = 10.0, 5.5 Hz, 1H), 1.77 – 1.61 (m, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 173.46, 156.04, 143.70 (d, *J* = 10.1 Hz), 140.62, 127.52, 126.92, 125.15, 125.14, 121.08, 120.01, 119.99, 118.75, 69.77, 65.47, 53.13, 46.56, 26.54, 25.86. ^{19}F NMR (471 MHz, DMSO-d_6) δ -69.96. HRMS (ESI-TOF) calculated for $C_{24}H_{19}F_9NO_5^-$ $[M-H]^-$: 572.1122, observed 572.1369.



3C-i

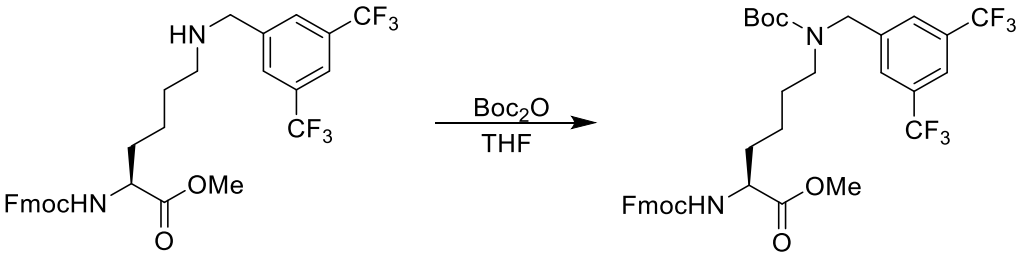
Fmoc-L-Lys-OH (1.1250 g, 3.053 mmol, 1 eq) was suspended in 9 mL of MeOH, followed by approximately 20 drops of 12 M HCl. A condenser was attached and the mixture was heated to 60 °C for 18 h. The solvent was removed and the mixture was precipitated in Et_2O. The precipitate was collected via vacuum filtration and washed with Et_2O to afford **3C-i** as the hydrochloride salt in quantitative yield. ^{1}H NMR (500 MHz, Methanol-d_4) δ 7.83 (d, J = 7.6 Hz, 2H), 7.69 (t, J = 8.5 Hz, 2H), 7.42 (t, J = 7.5 Hz, 2H), 7.34 (td, J = 7.5, 1.1 Hz, 2H), 4.45 (dd, J = 10.6, 6.7 Hz, 1H), 4.37 (dd, J = 10.7, 6.9 Hz, 1H), 4.27 – 4.23 (m, 1H), 4.23 – 4.20 (m, 1H), 3.74 (s, 3H), 2.97 – 2.88 (m, 2H), 1.98 – 1.83 (m, 1H), 1.79 – 1.63 (m, 3H), 1.55 – 1.42 (m, 2H). HRMS (ESI-TOF) calculated for $C_{22}H_{26}N_2O_4H^+$ $[M+H]^+$: 383.1965, observed 383.2026.



3C-i **3C-ii**

3C-i (1.3013 g, 3.11 mmol, 1 eq) was suspended in 25 mL of 1-2 dichloroethane, followed by addition of triethylamine (433 µL, 3.11 mmol, 1 eq). 3,5-Bistrifluoromethylbenzaldehyde (752.1 mg, 3.11 mmol, 1 eq) was added and stirred for 10 min, followed by addition of sodium triacetoxyborohydride (1.0075 g, 4.75 mmol, 1.5 eq) and the reaction was stirred for 18 h. The solution was quenched with saturated $NaHCO_3$ and diluted into DCM and water and extracted into DCM. The organic layers were washed with brine, dried over $MgSO_4$, filtered, concentrated, and purified on a Combiflash Rf

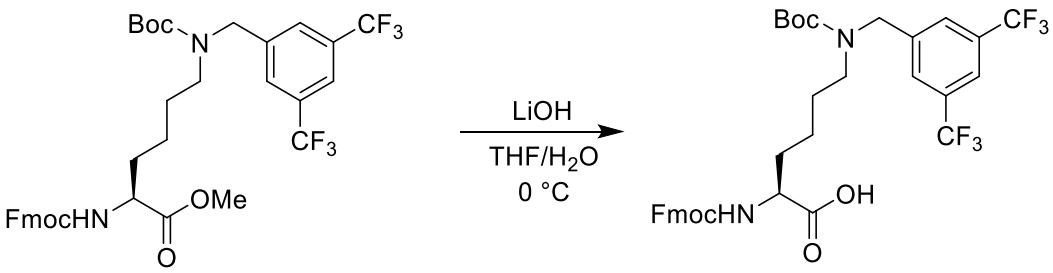
system (DCM/MeOH, 0-10% MeOH) to afford **3C-ii** as a white solid (635.7 mg, 34%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.82 (s, 2H), 7.76 (d, *J* = 8.0 Hz, 3H), 7.59 (dd, *J* = 7.6, 4.4 Hz, 2H), 7.39 (t, *J* = 7.4 Hz, 2H), 7.31 (t, *J* = 7.5 Hz, 2H), 5.39 (d, *J* = 8.4 Hz, 1H), 4.49 – 4.32 (m, 3H), 4.22 (t, *J* = 7.0 Hz, 1H), 3.89 (s, 2H), 3.75 (s, 3H), 2.63 (t, *J* = 7.0 Hz, 2H), 1.96 – 1.76 (m, 2H), 1.76 – 1.63 (m, 1H), 1.63 – 1.50 (m, 2H), 1.41 (m, 2H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 172.93, 155.89, 143.85, 143.70, 142.85, 141.28, 131.54 (q, *J* = 33.1 Hz), 128.13, 128.13, 127.68, 127.02, 125.02, 121.13 – 120.85 (m), 119.96, 119.96, 66.93, 53.71, 52.84, 52.39, 49.00, 47.14, 32.49, 29.35, 22.88. ^{19}F NMR (471 MHz, Chloroform-*d*) δ -62.77 (s). HRMS (ESI-TOF) calculated for $C_{31}H_{30}F_6N_2O_4H^+$ $[M+H]^+$: 609.2143, observed 609.2181.



3C-ii **3C-iii**

3C-ii (608.8 mg, 1.000 mmol, 1 eq) was dissolved in 4 mL of THF. Separately, boc anhydride (333.9 mg, 1.530 mmol, 1.53 eq) was dissolved in 1 mL of THF. The anhydride solution was added to the first, and the reaction proceeded at rt for 3 h. The solvent was removed, and the crude mixture was dissolved in DCM, and washed 3 X with H_2O followed by brine and dried over $MgSO_4$. Following filtration and removal of solvent, the product was purified on a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **3C-iii** as a clear oil (551.9 mg, 78%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.01 (s, 1H), 7.89 (overlapping, 4H), 7.76 (d, *J* = 7.8 Hz, 1H), 7.71 (dd, *J* = 7.6, 4.3 Hz, 2H), 7.41 (t, *J* = 7.5 Hz, 2H), 7.33 (t, *J* = 7.4 Hz, 2H), 4.54 (s, 2H), 4.30 (dd, *J* = 7.1, 4.0 Hz, 2H), 4.22 (t, *J* = 7.0 Hz, 1H), 4.00 (ddd, *J* = 9.5, 7.8, 5.0 Hz, 1H), 3.62 (s, 3H), 3.30 – 3.14 (m, 3H), 1.65 (dq, *J* = 29.6, 11.2, 9.5 Hz, 2H), 1.51 – 1.22 (overlapping m, 14H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 172.86, 155.96, 143.81 (d, *J* = 16.4 Hz), 141.31 (d, *J* = 2.0 Hz), 131.84 (q, *J* = 33.3 Hz), 127.72, 127.06, 125.07, 121.20, 120.00, 80.63, 67.05, 53.70, 52.44, 50.15,

49.55, 47.16, 46.96, 32.45, 32.05, 28.28, 27.84, 27.53, 22.37, 17.71, 12.30. ^{19}F NMR (471 MHz, Chloroform-*d*) δ -62.89. HRMS (ESI-TOF) calculated for $C_{36}H_{38}F_6N_2O_6Na^+$ $[M+Na]^+$: 731.2532, observed 731.2064.

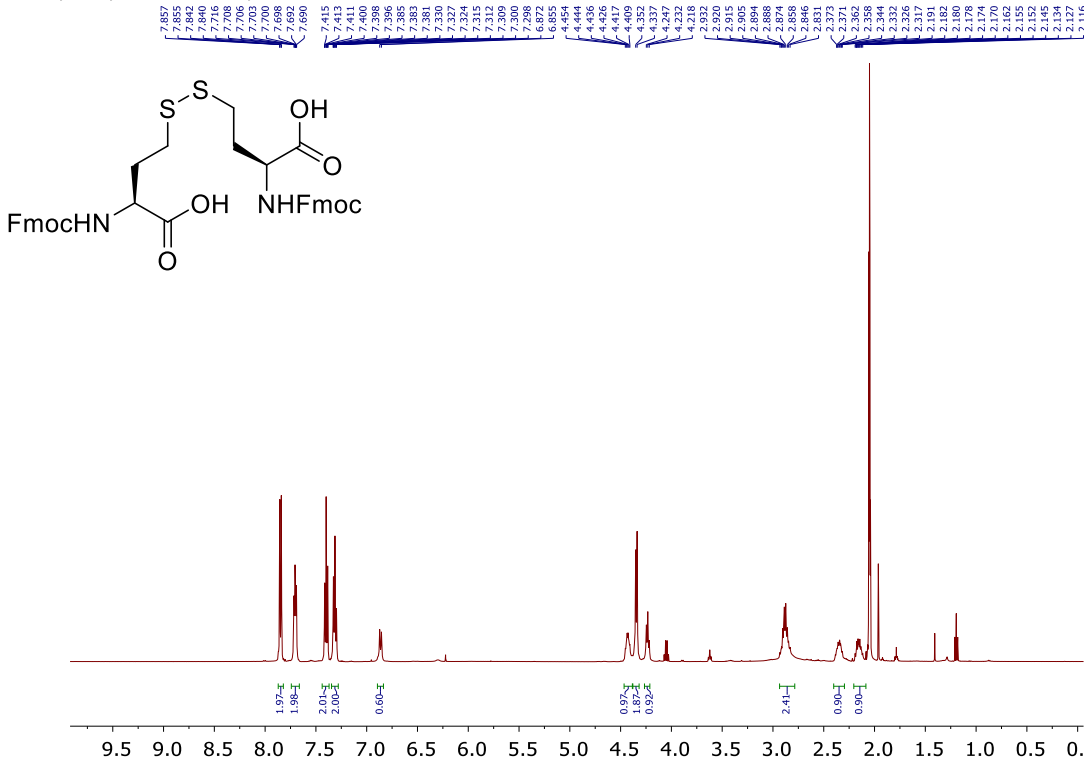


3C-iii **3C**

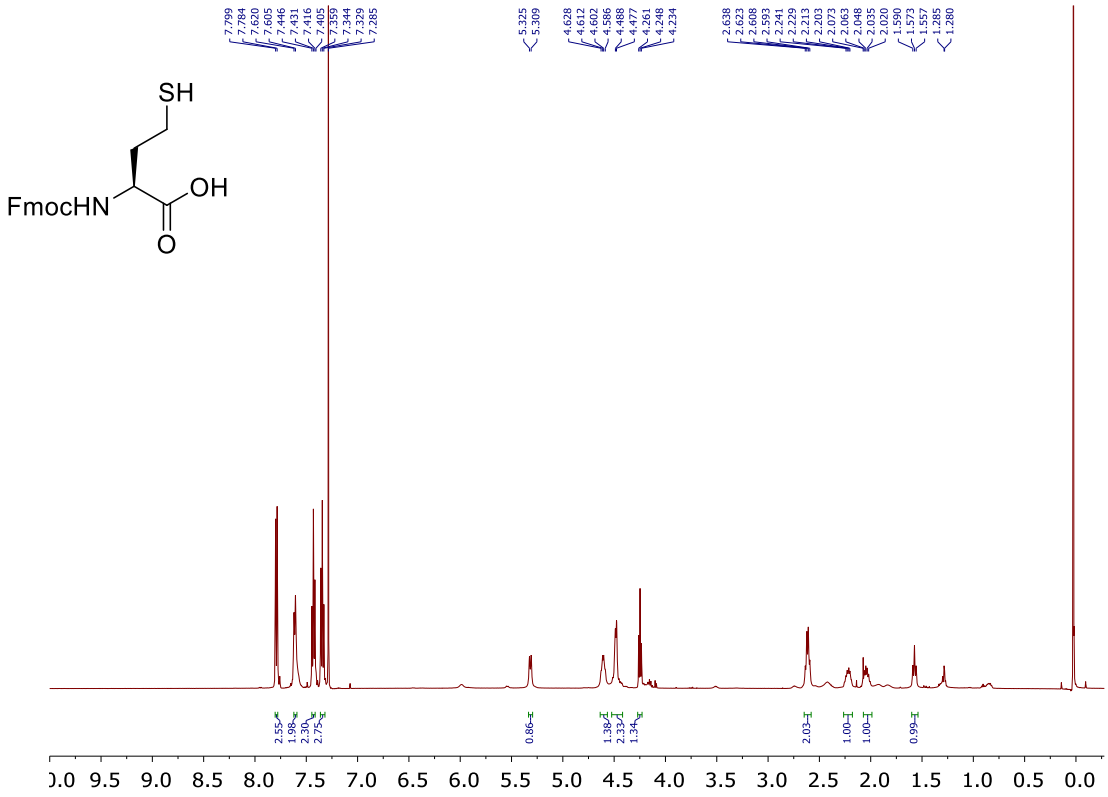
3C-iii (514.7 mg, 0.726 mmol, 1 eq) was dissolved in 5 mL of THF and cooled to 0 °C. Slowly, 2.6 mL of LiOH solution (0.3 M, in H_2O) was added over 30 min. The reaction was quenched with 1 M HCl and extracted into EtOAc. The organic layer was washed with brine, dried over $MgSO_4$, filtered, concentrated, and purified on a Combiflash Rf system (DCM/MeOH, 0-10% MeOH) to afford **3C** as a white solid (111.4 mg, 22%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.78 (s, 1H), 7.75 (d, *J* = 7.7 Hz, 2H), 7.66 (s, 2H), 7.61 – 7.51 (m, 2H), 7.38 (dd, *J* = 8.6, 6.1 Hz, 2H), 7.29 (t, *J* = 7.5 Hz, 2H), 5.65 – 5.29 (m, 2H), 4.60 – 4.30 (m, 5H), 4.21 (t, *J* = 7.0 Hz, 1H), 3.34 – 3.12 (m, 2H), 1.84 (d, *J* = 67.7 Hz, 1H), 1.56 (s, 2H), 1.48 (s, 3H), 1.39 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 175.94, 156.22, 143.78, 143.65, 141.27, 131.70 (q, *J* = 33.41 Hz), 127.71, 127.04, 125.05, 121.19 (d, *J* = 5.5 Hz), 119.97, 67.15, 53.58, 50.13, 47.09, 46.87, 28.25, 22.21. ^{19}F NMR (471 MHz, Chloroform-*d*) δ -62.89. HRMS (ESI-TOF) calculated for $C_{35}H_{35}F_6N_2O_6^-$ $[M-H]^-$: 693.2402, observed 693.2002.

3.6.5 NMR spectra of small molecules

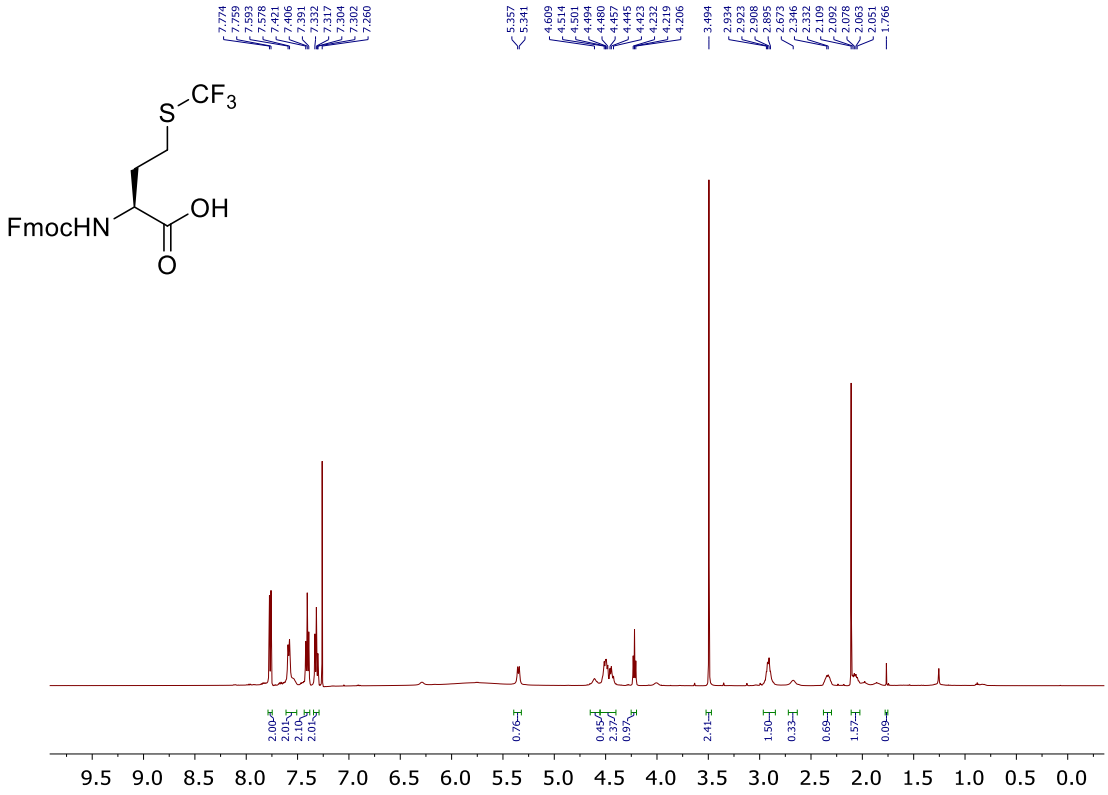
3A-i, ^{1}H, 500 MHz, Acetone-d_6



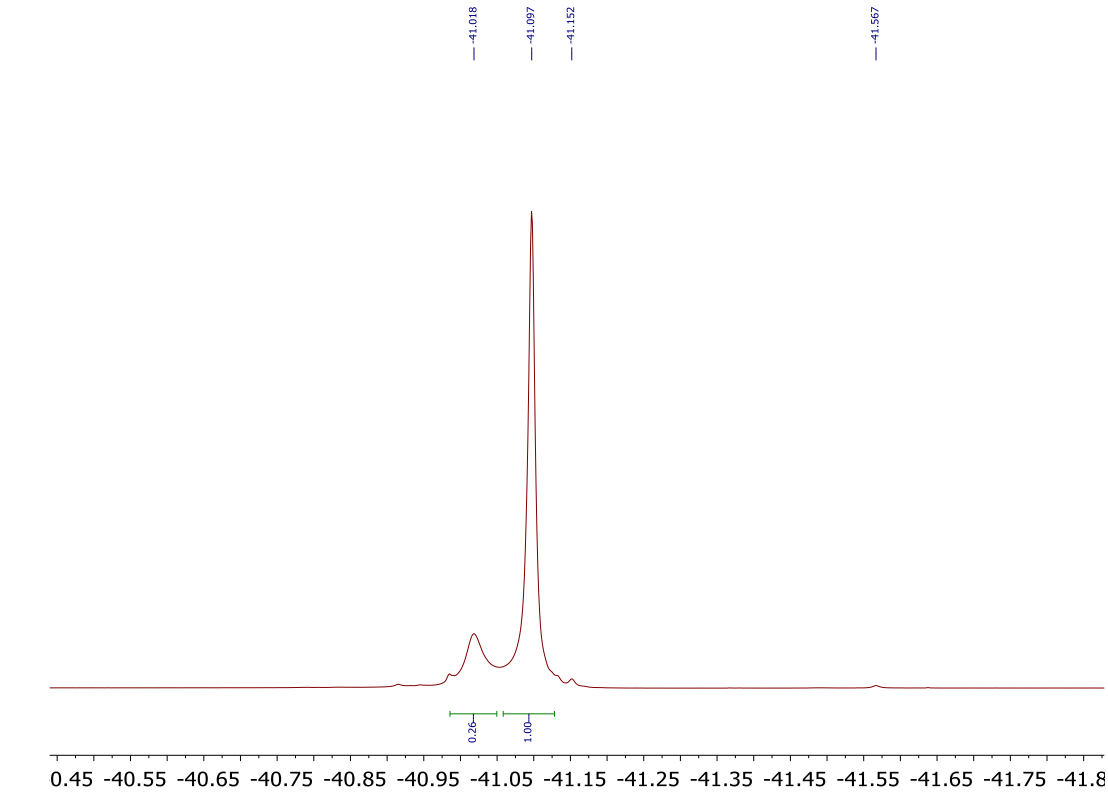
3A-ii, ^{1}H, 500 MHz, $CDCl_3$



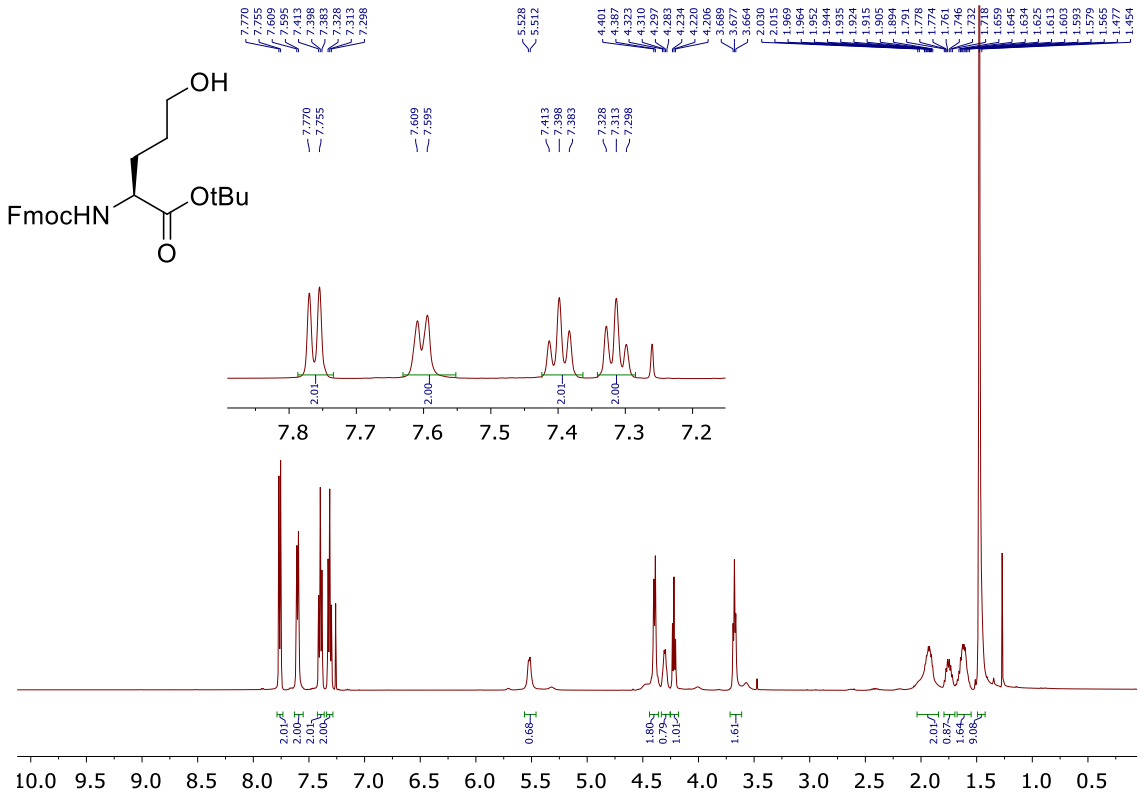
3A, ^{1}H, 500 MHz, $CDCl_3$



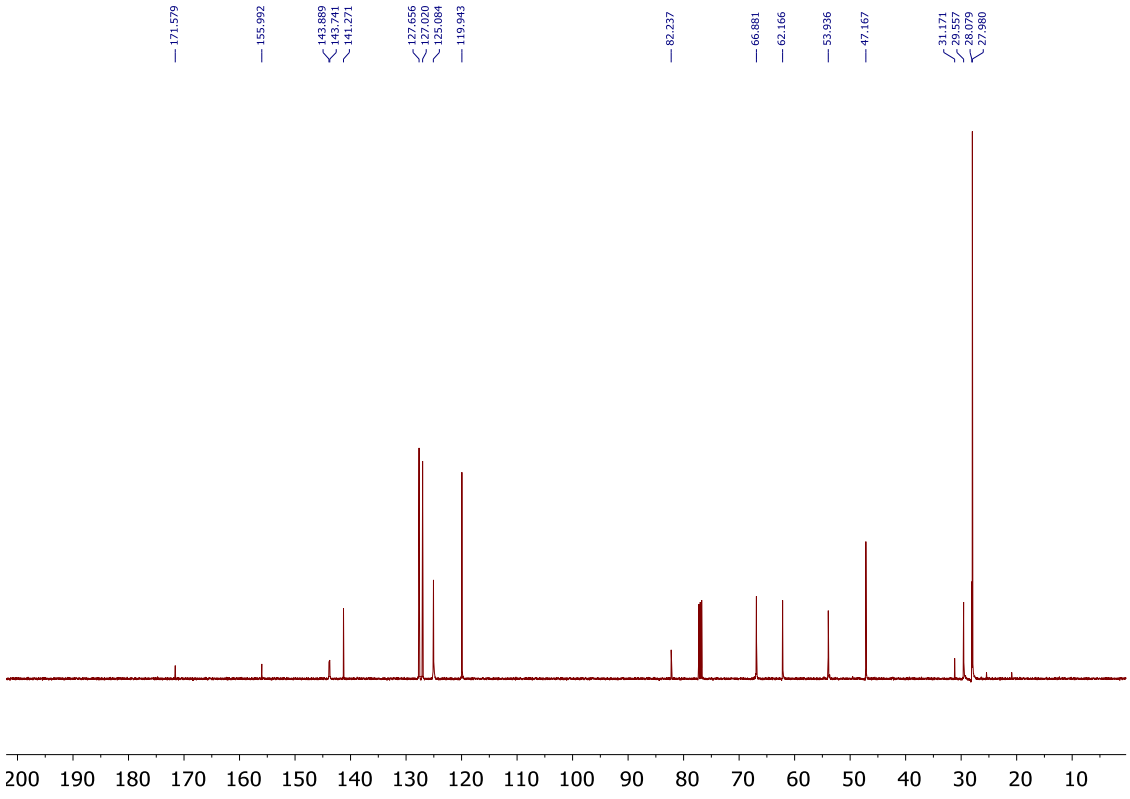
3A, ^{19}F, 471 MHz, $CDCl_3$



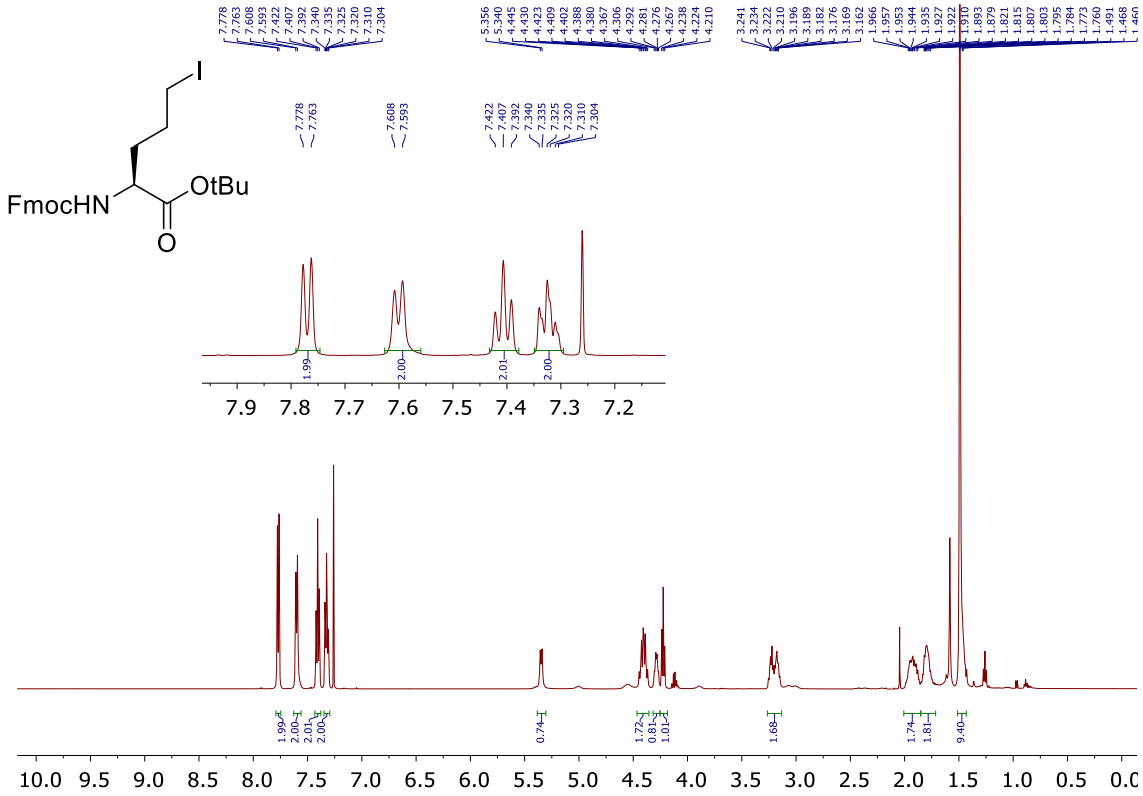
3B-i, ^{1}H, 500 MHz, $CDCl_3$



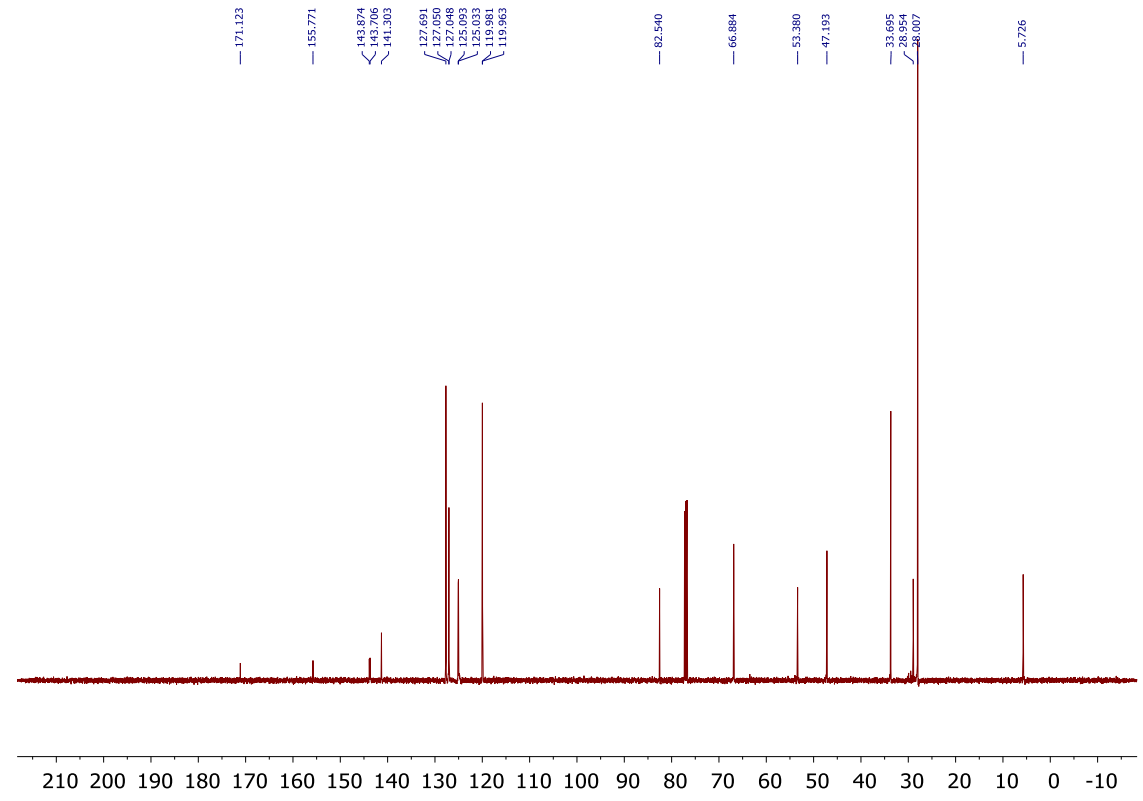
3B-i, ^{13}C, 126 MHz, $CDCl_3$



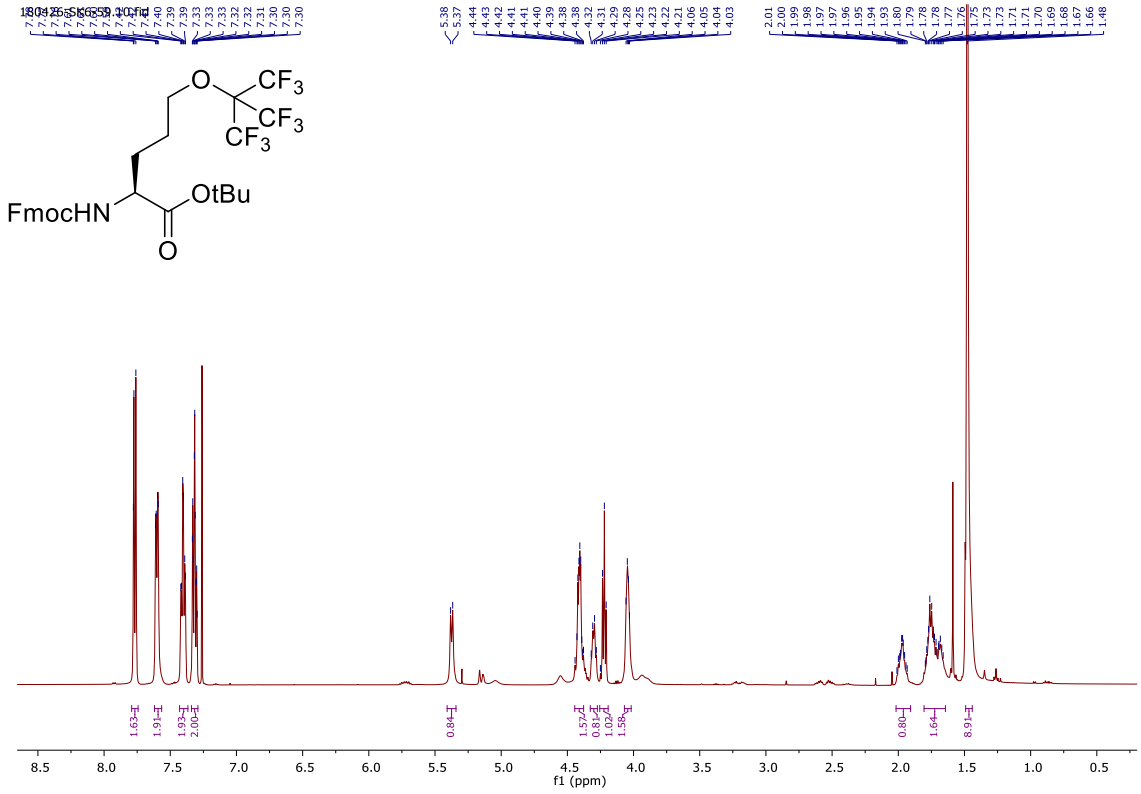
3B-ii, ^{1}H, 500 MHz, $CDCl_3$



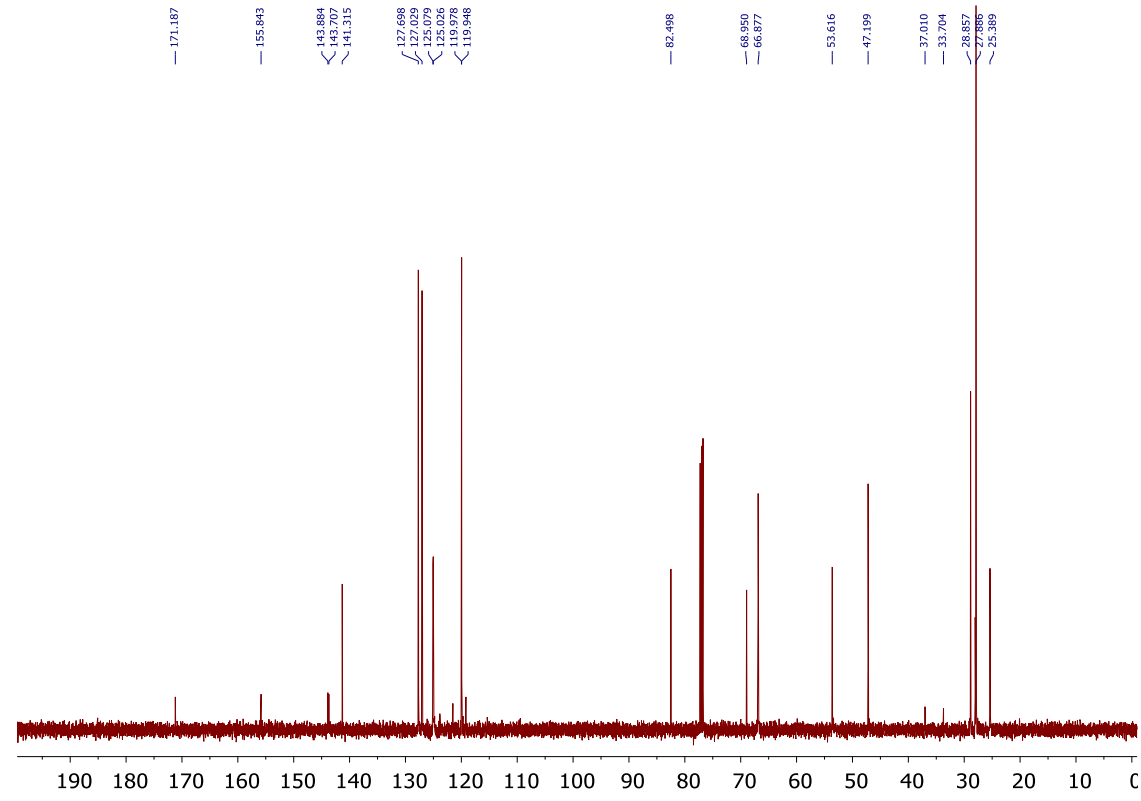
3B-ii, ^{13}C, 126 MHz, $CDCl_3$



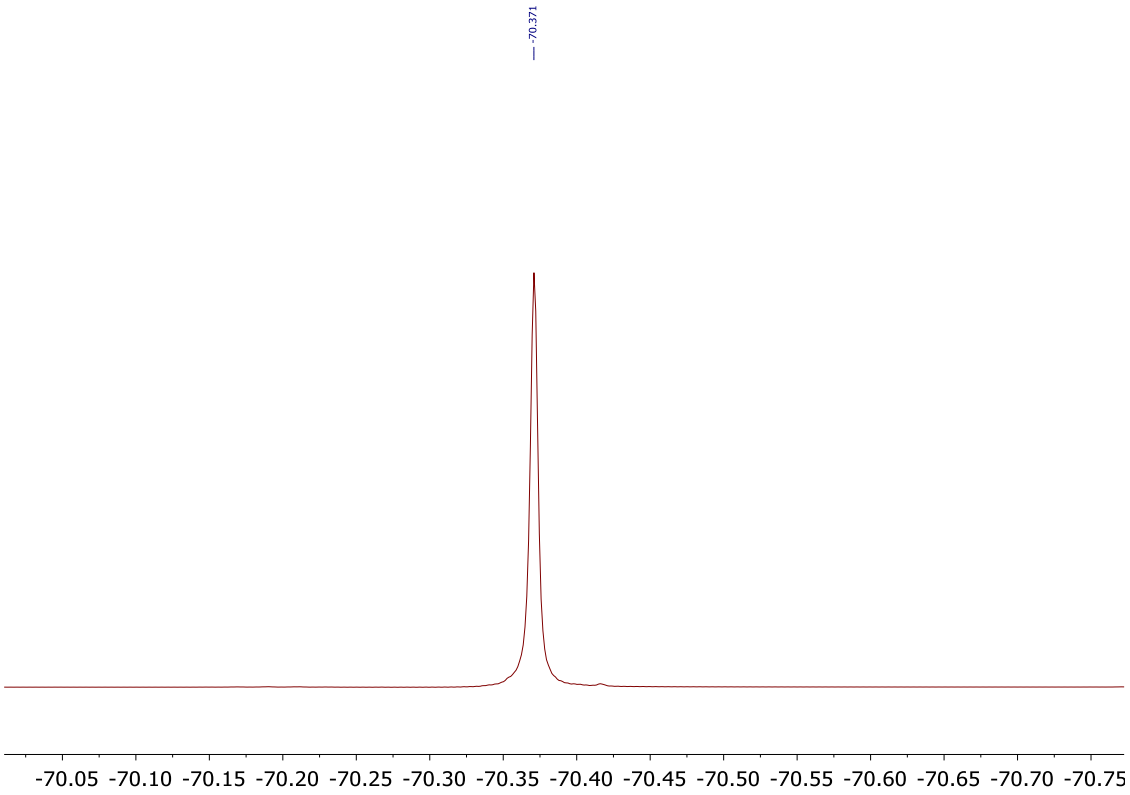
3B-iii, ^{1}H, 500 MHz, $CDCl_3$



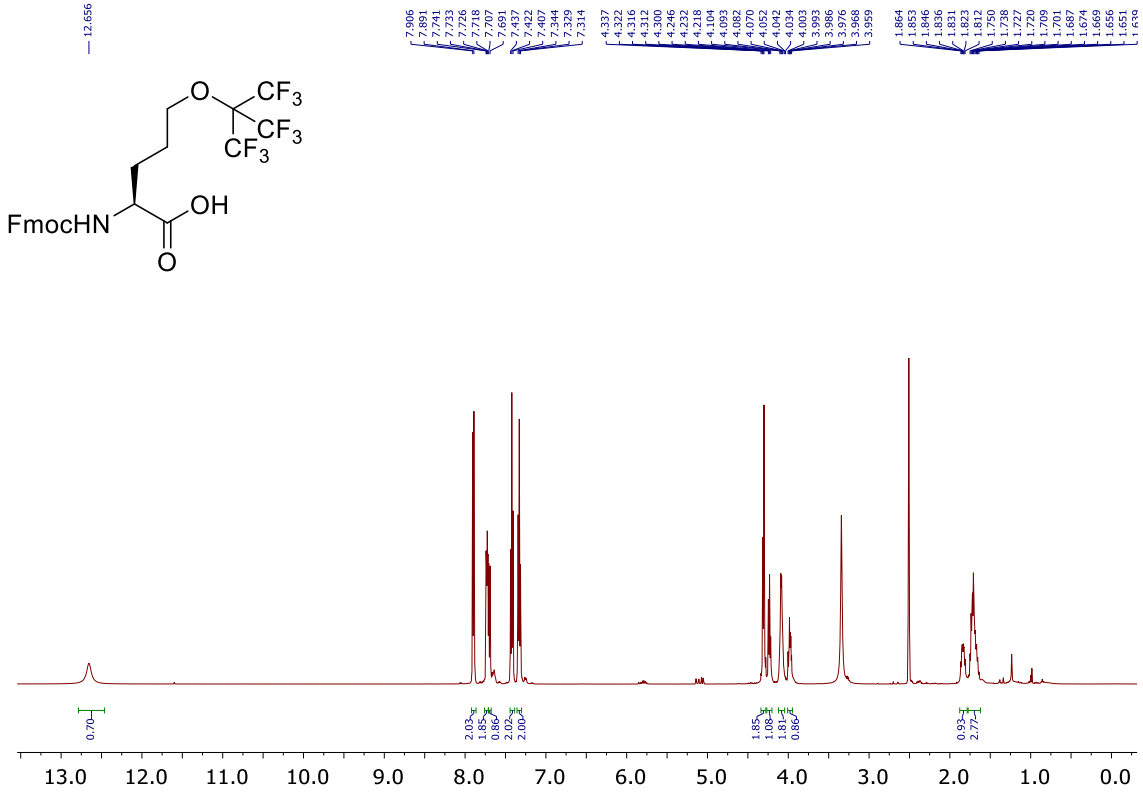
3B-iii, ^{13}C, 126 MHz, $CDCl_3$



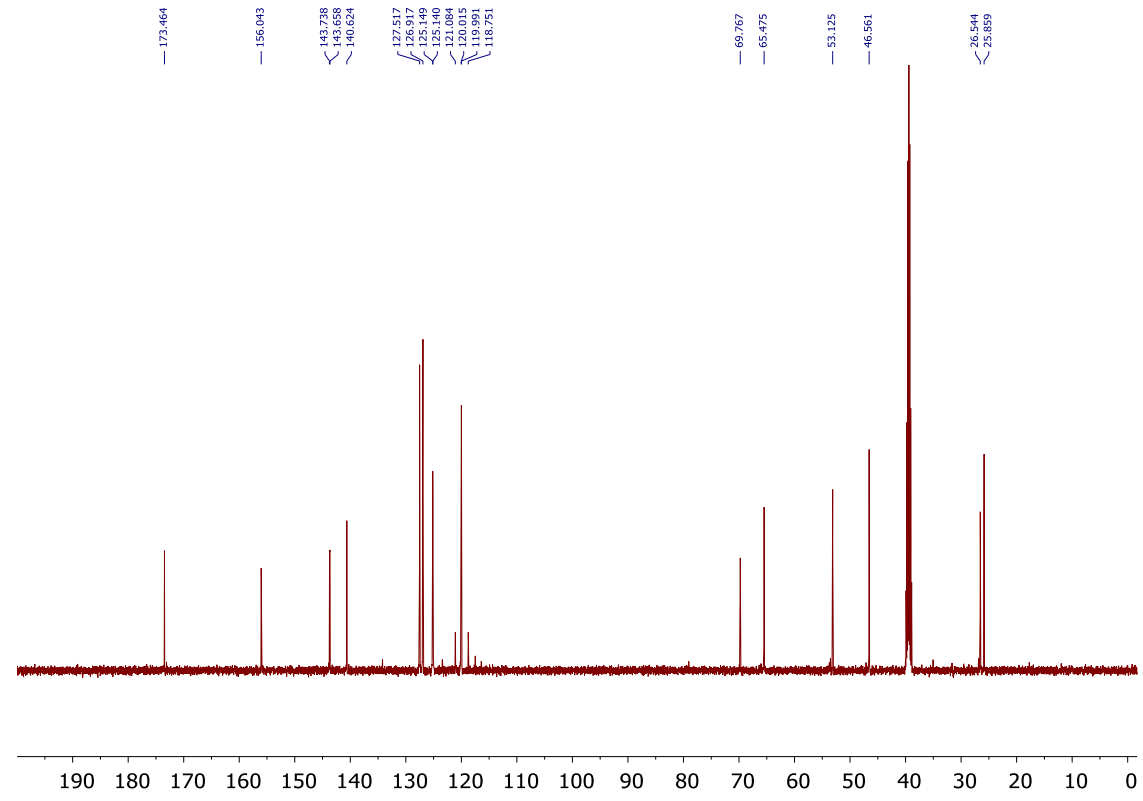
3B-iii, ^{19}F, 471 MHz, $CDCl_3$



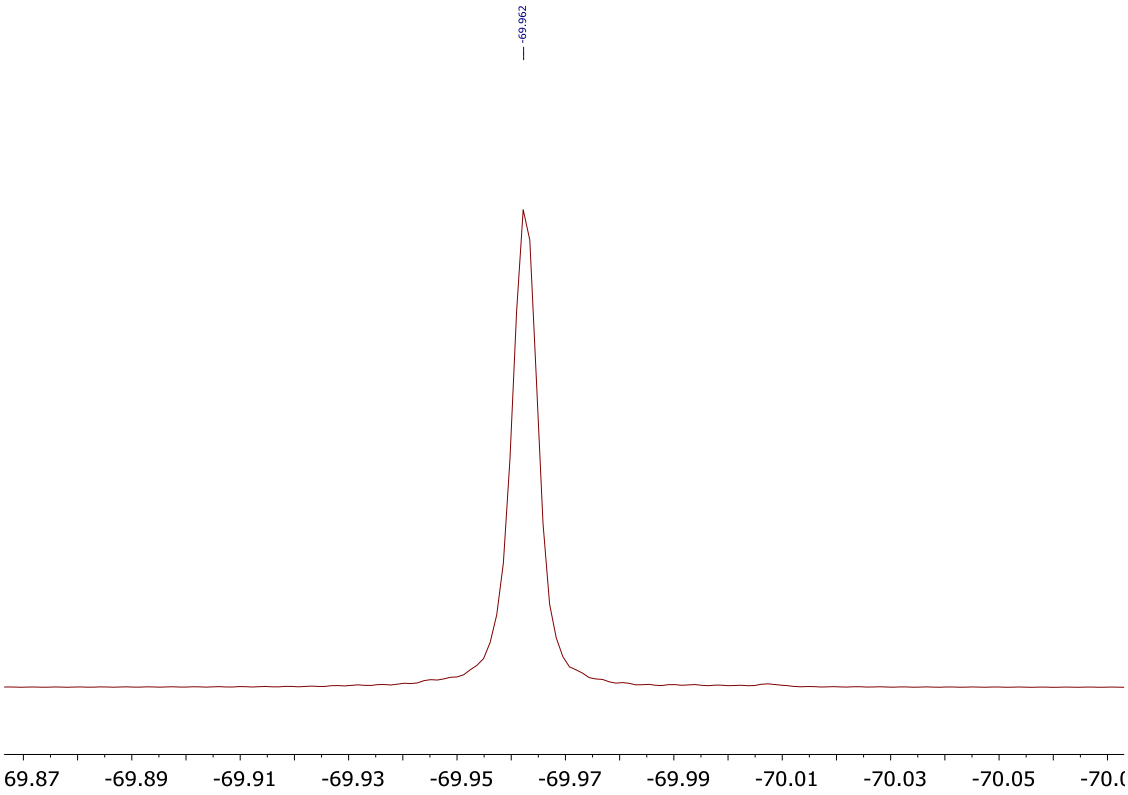
3B, ^{1}H, 500 MHz, DMSO-d_6



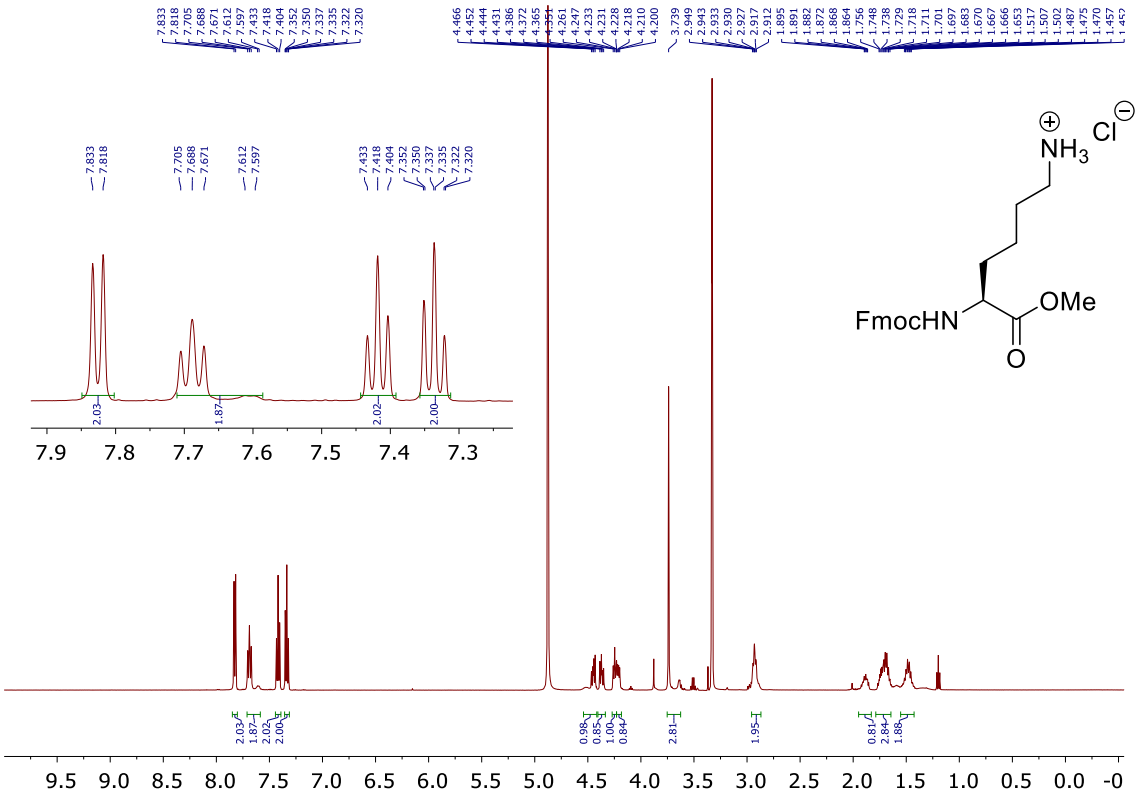
3B, ^{13}C, 126 MHz, DMSO-d_6



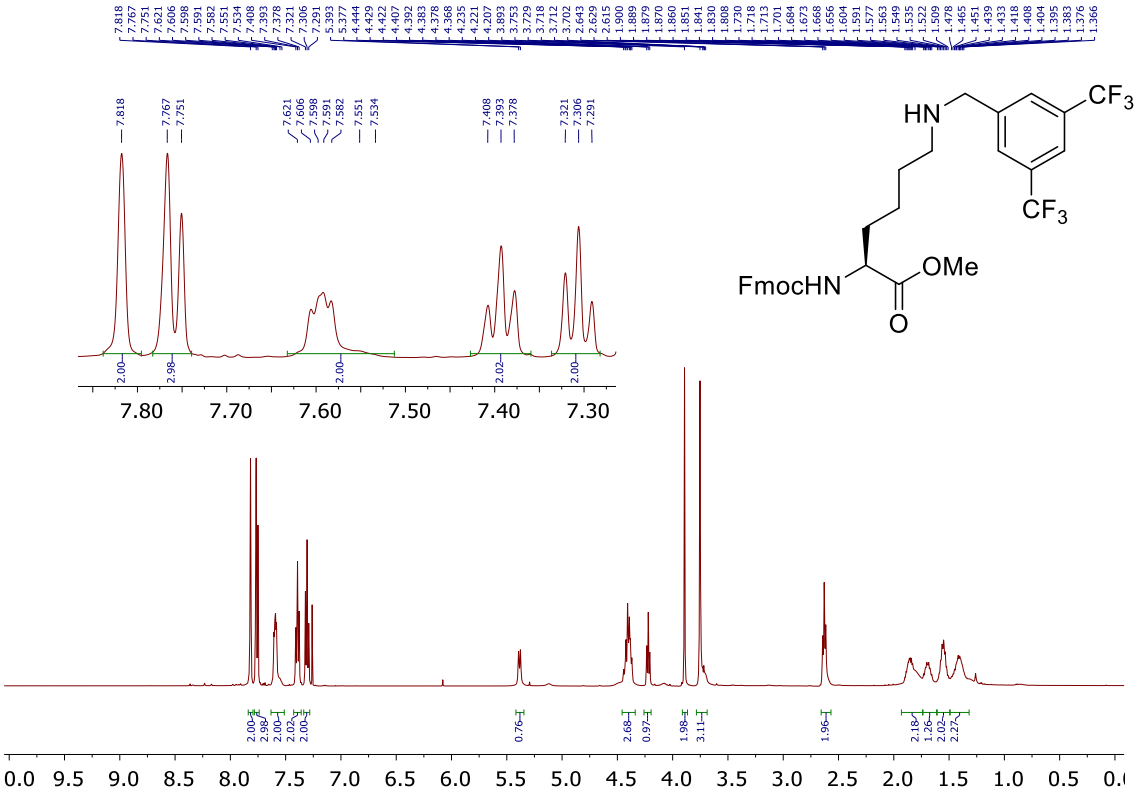
3B, ^{19}F, 471 MHz, DMSO-d_6



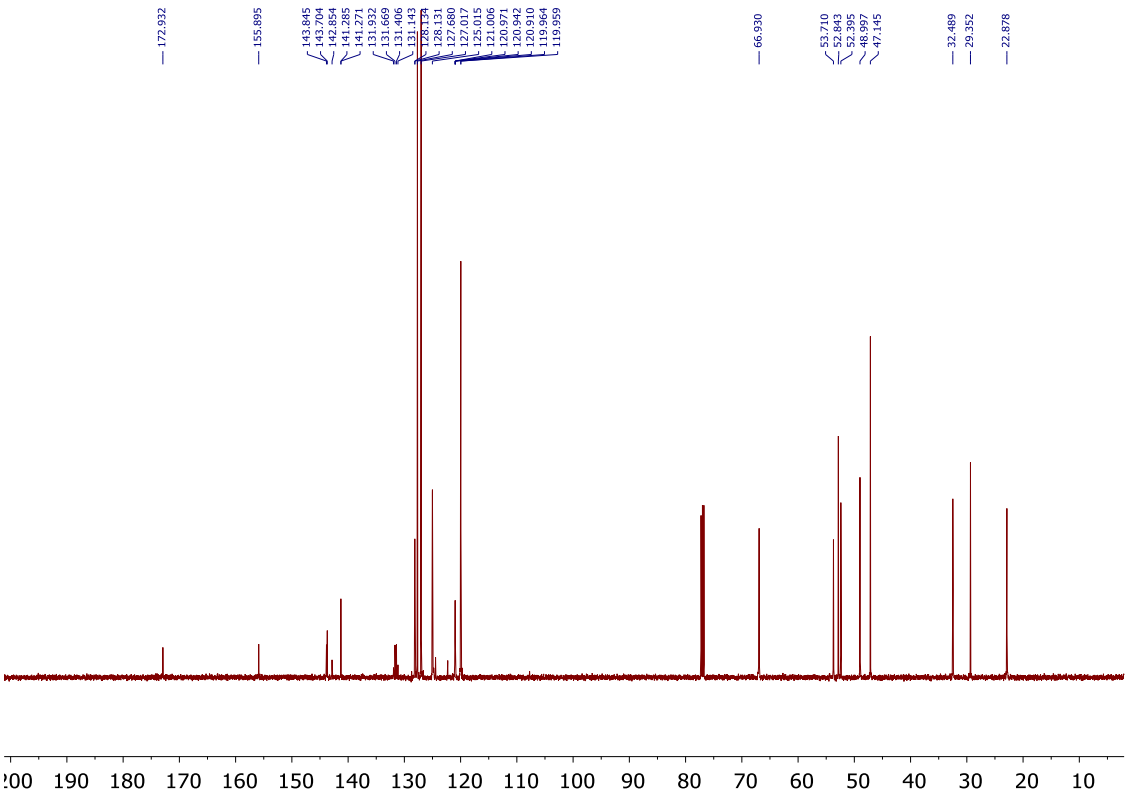
3C-i, ^{1}H, 500 MHz, CD_3OD



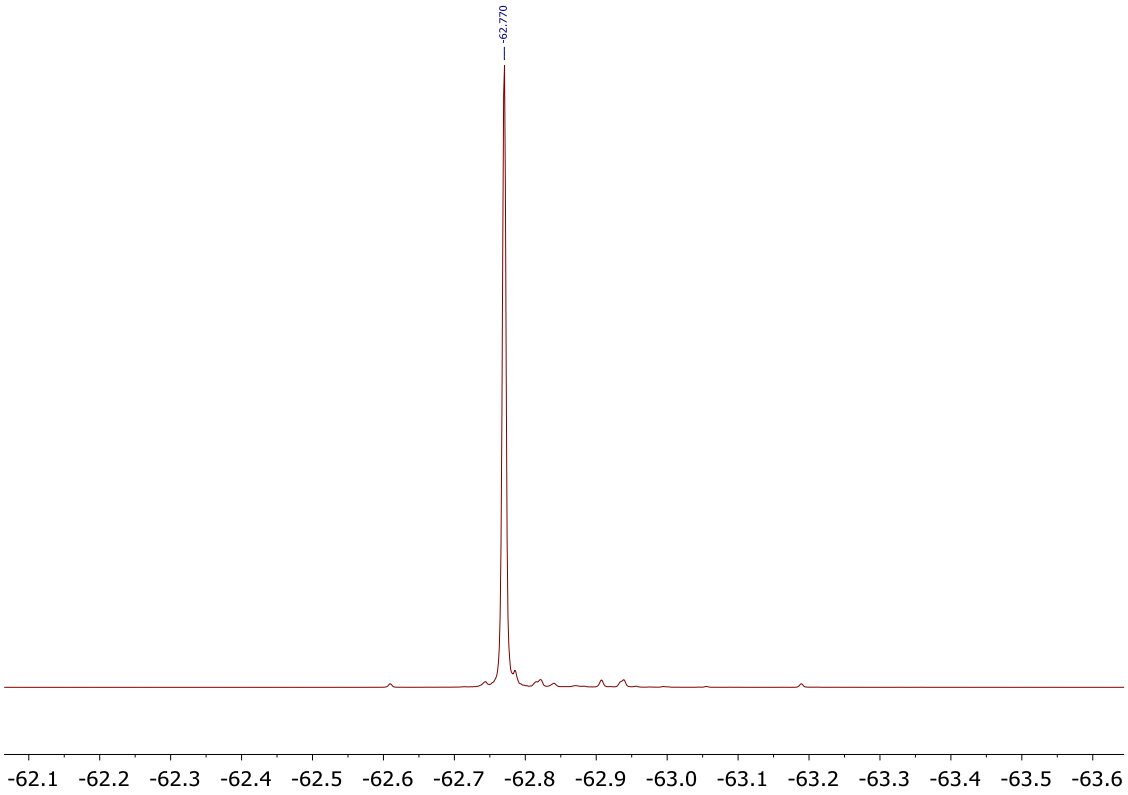
3C-ii, ^{1}H, 500 MHz, $CDCl_3$



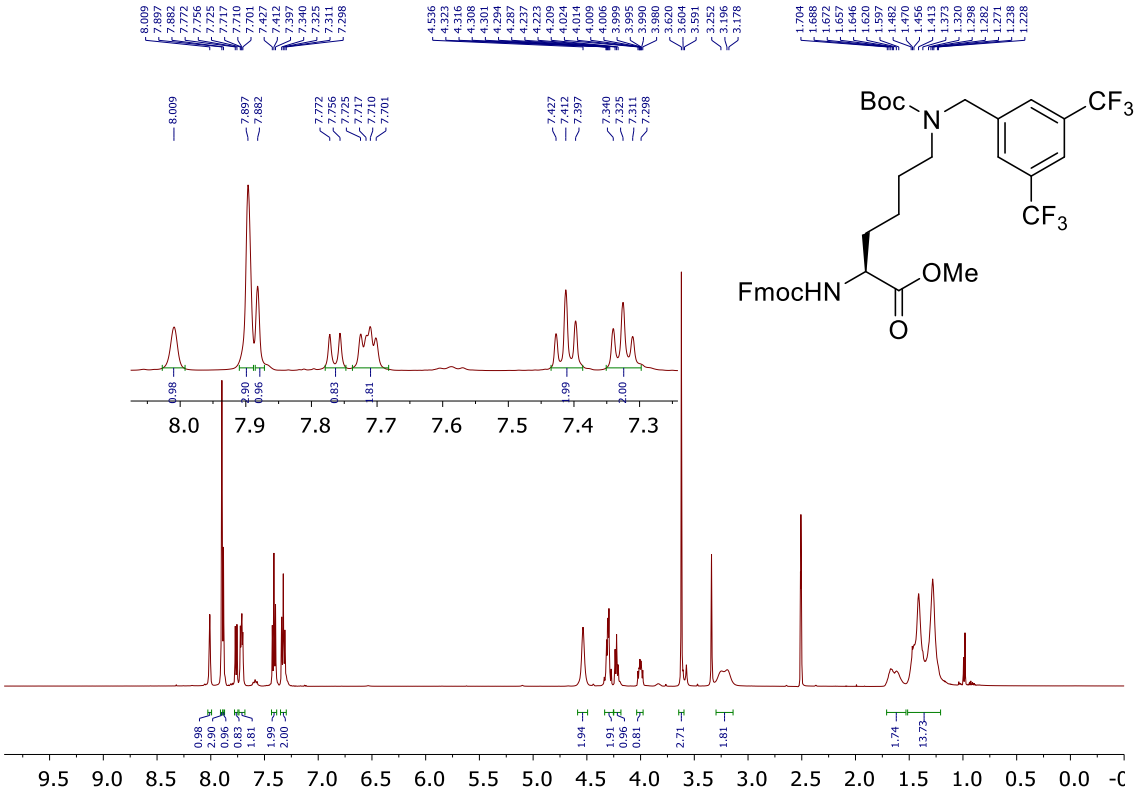
3C-ii, ^{13}C, 126 MHz, $CDCl_3$



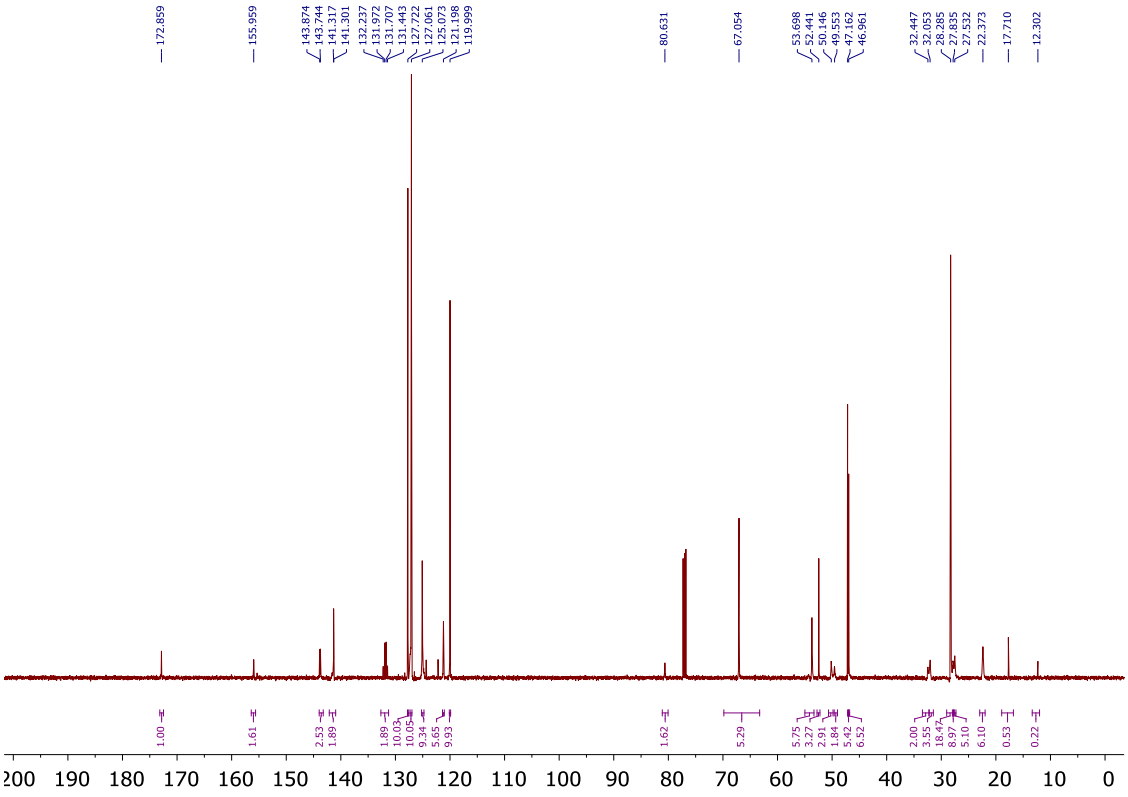
3C-ii, ^{19}F, 471 MHz, $CDCl_3$



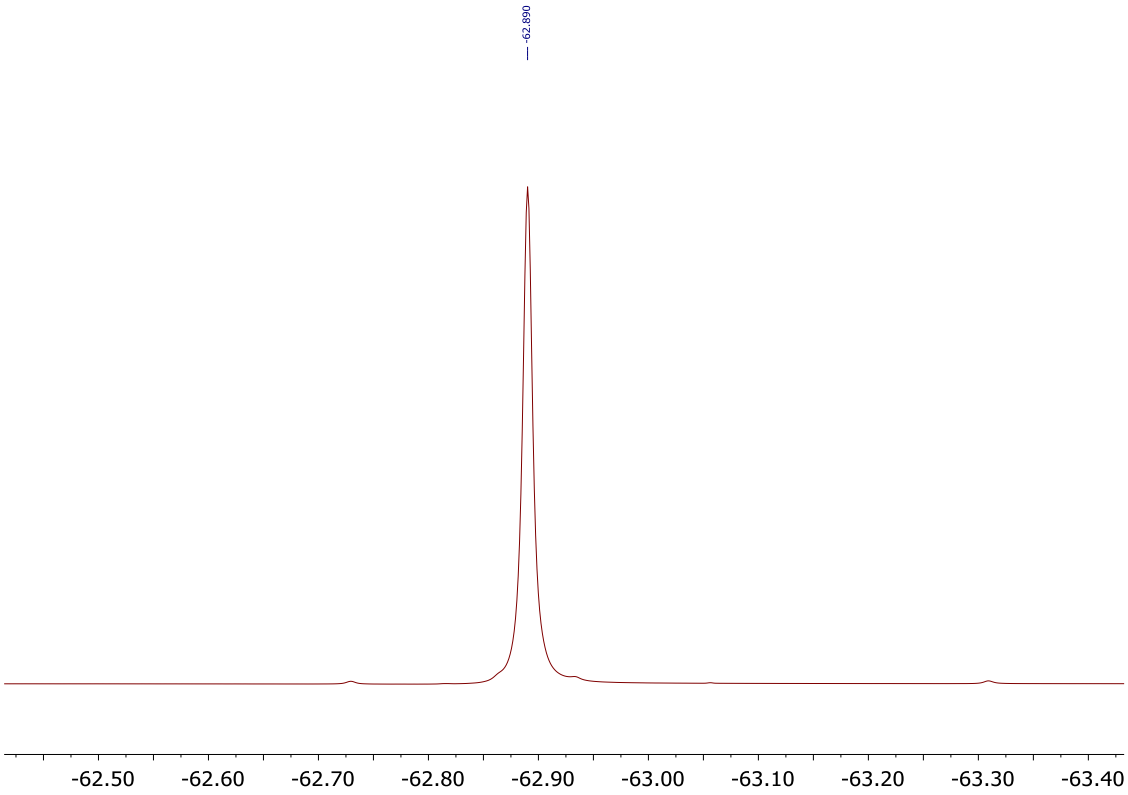
3C-iii, ^{1}H, 500 MHz, DMSO-d_6



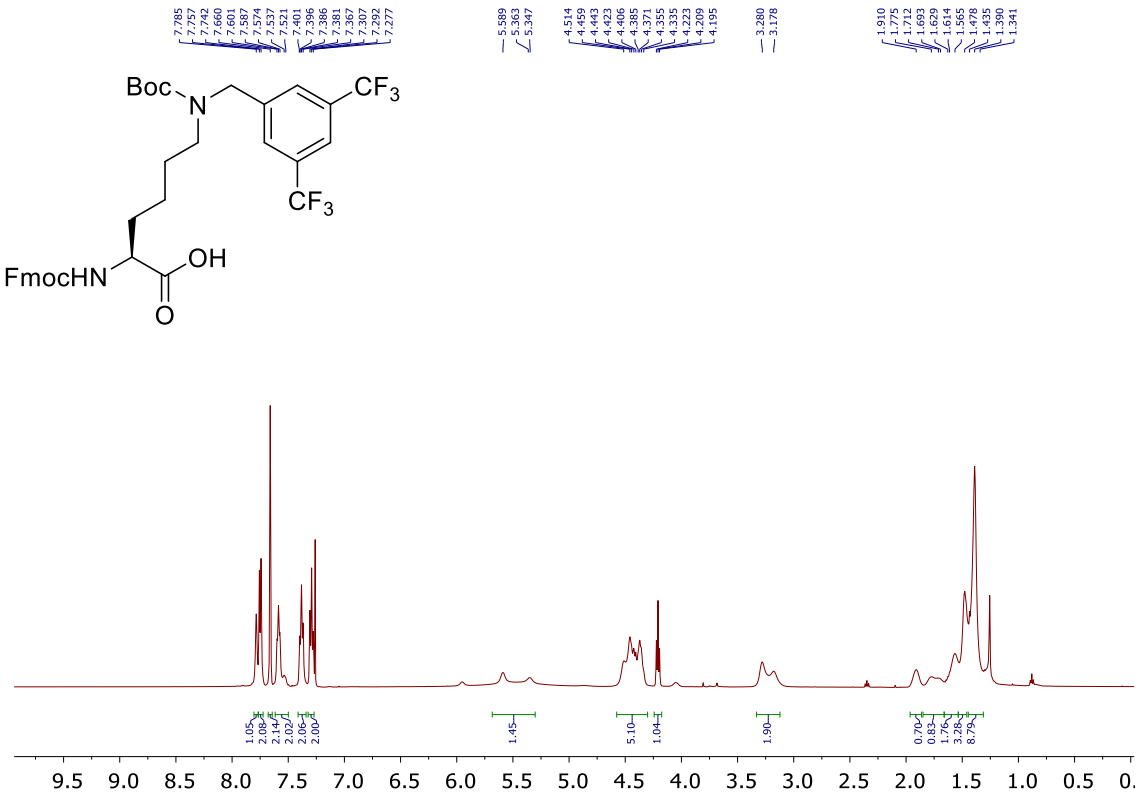
3C-iii, ^{13}C, 126 MHz, $CDCl_3$



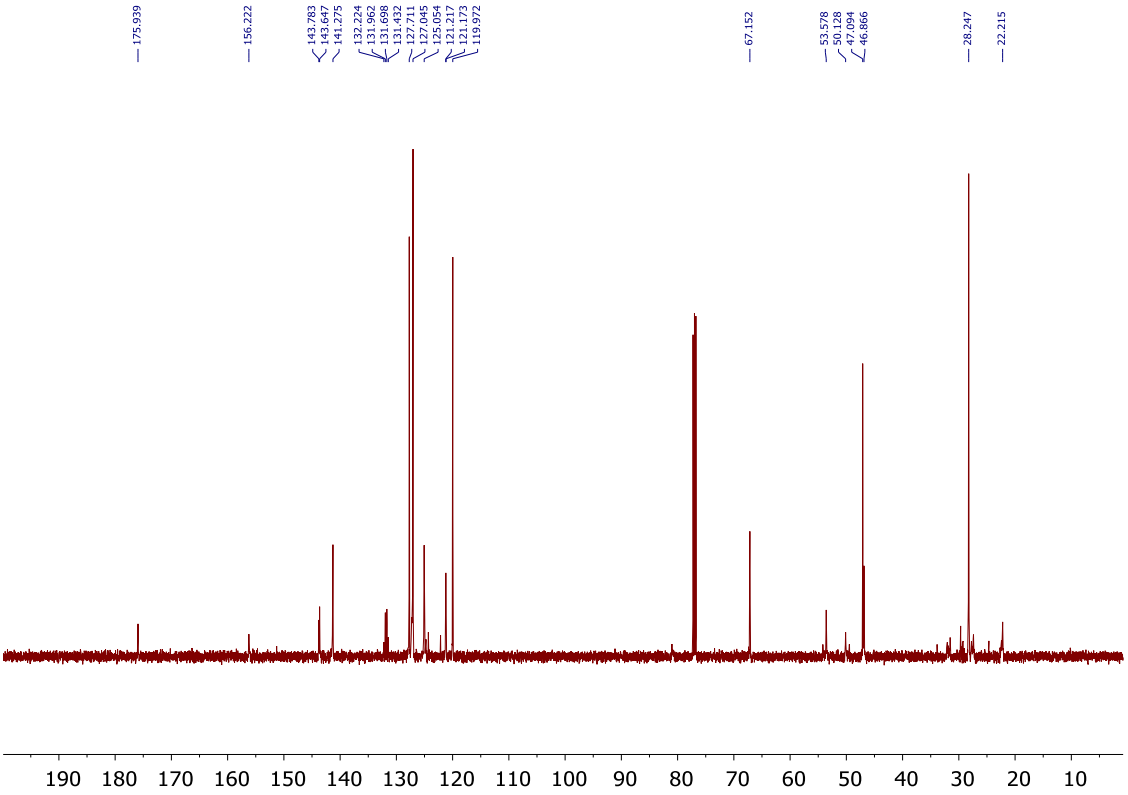
3C-iii, ^{19}F, 471 MHz, $CDCl_3$



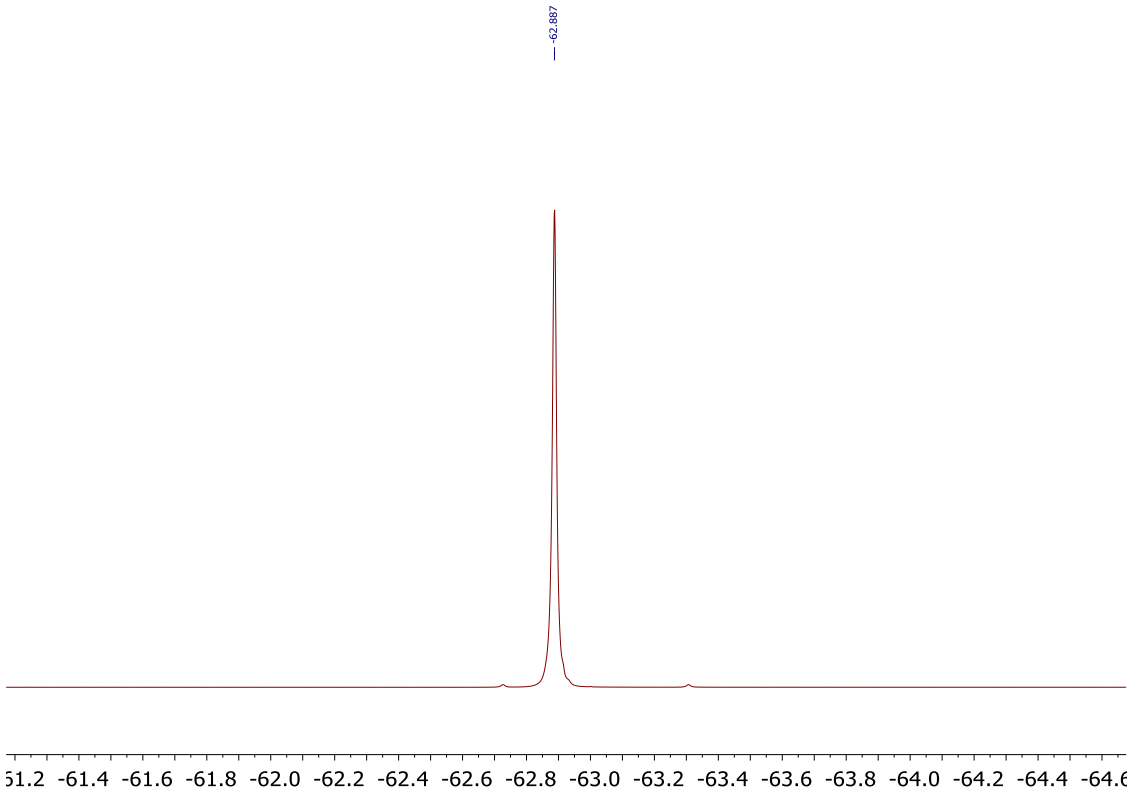
3C, ^{1}H, 500 MHz, $CDCl_3$



3C, ^{13}C, 126 MHz, $CDCl_3$

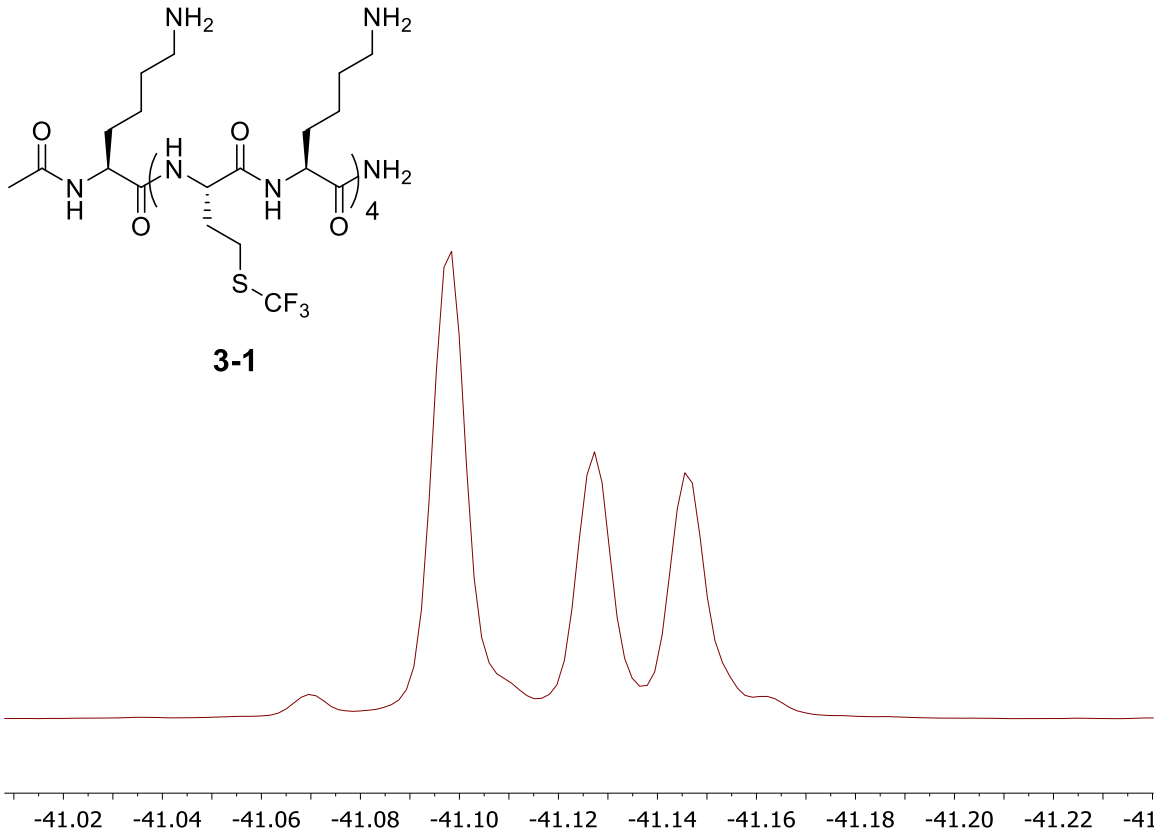


3C, ^{19}F, 471 MHz, $CDCl_3$

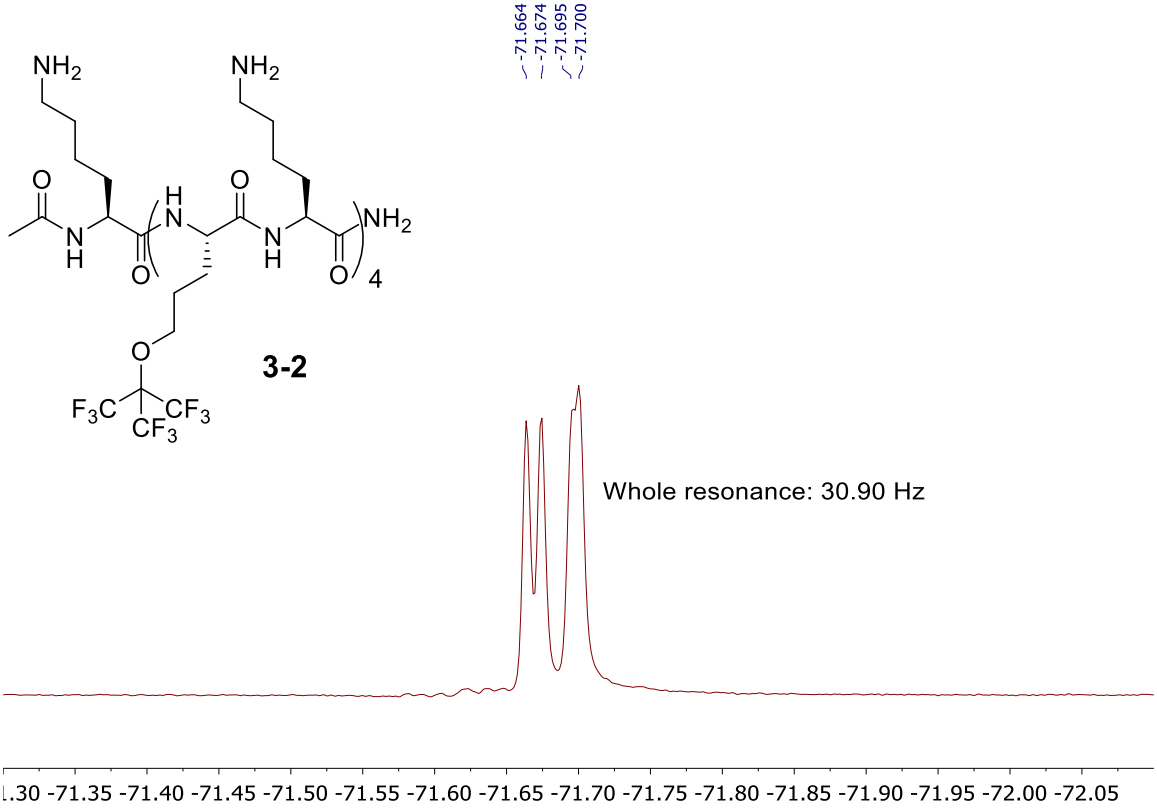


3.6.6 ^{19}F NMR and HPLC traces of synthesized peptides

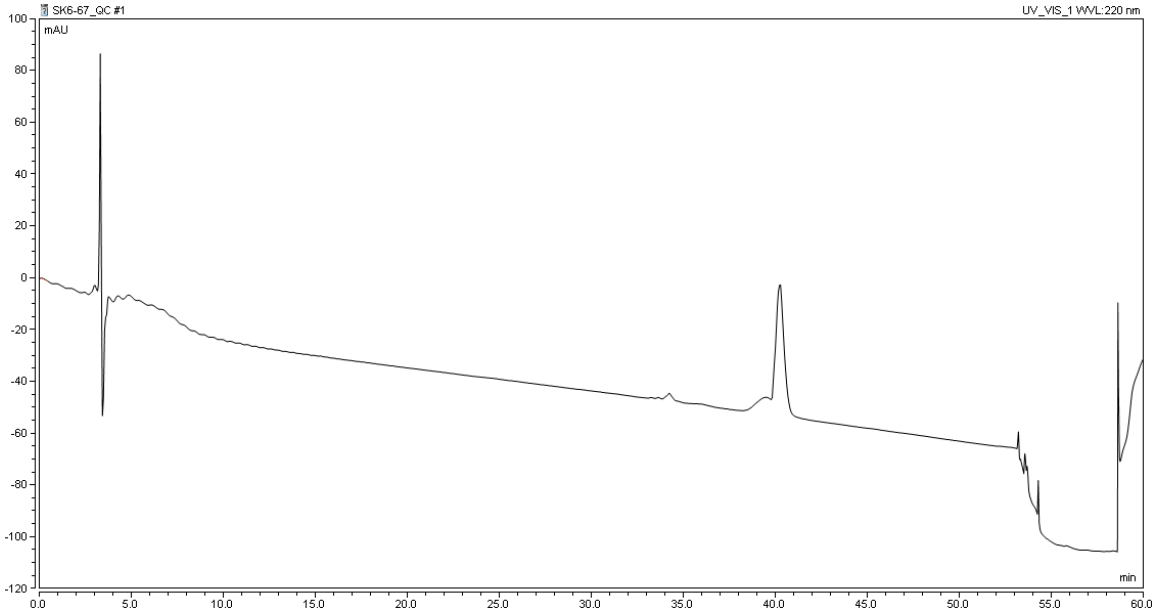
3-1, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O



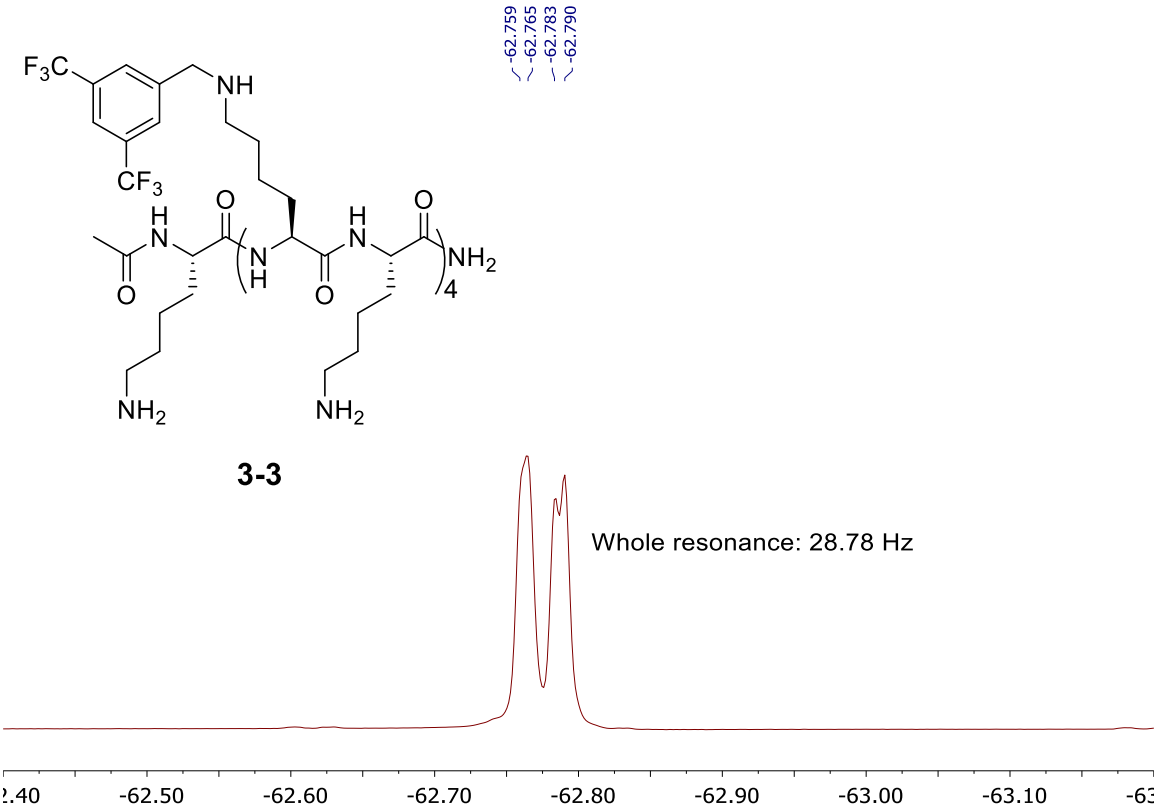
3-2, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O



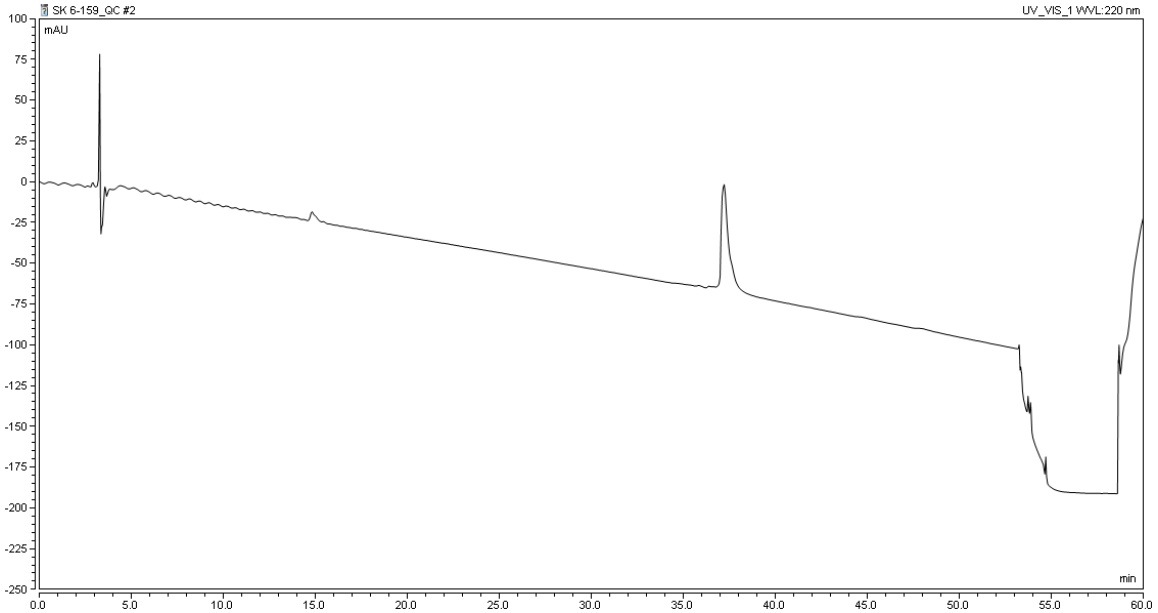
3-2, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min



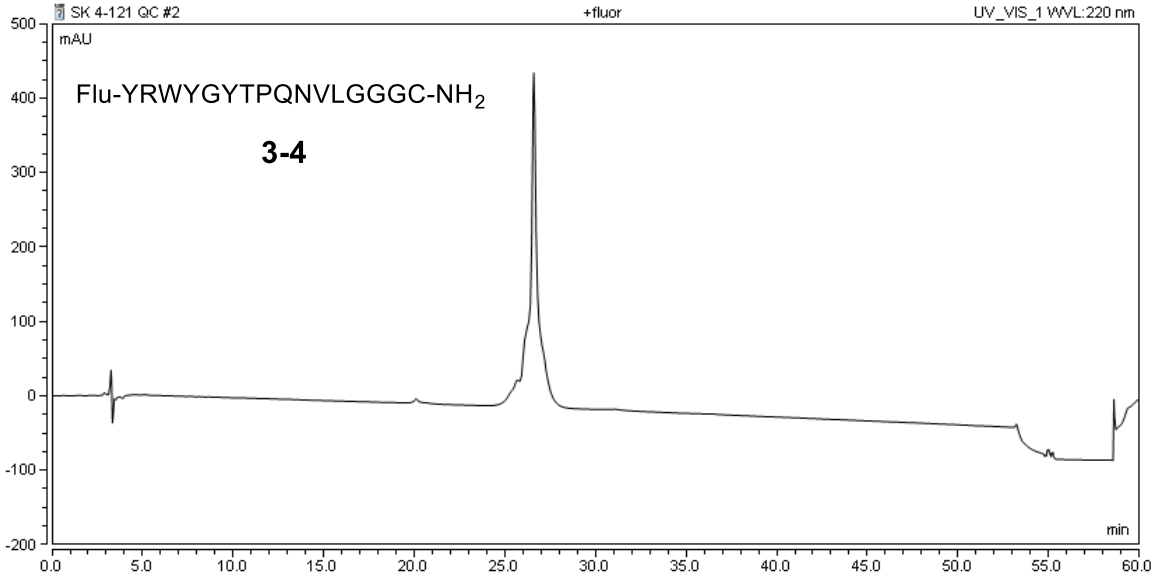
3-3, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O



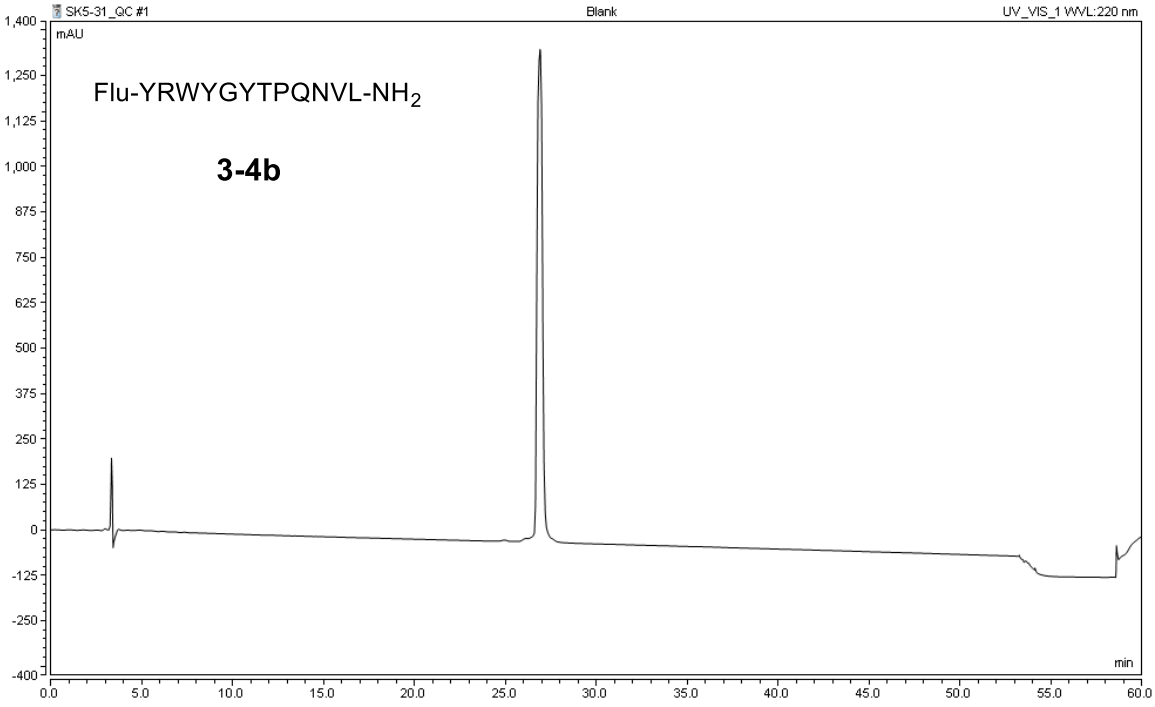
3-3, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min



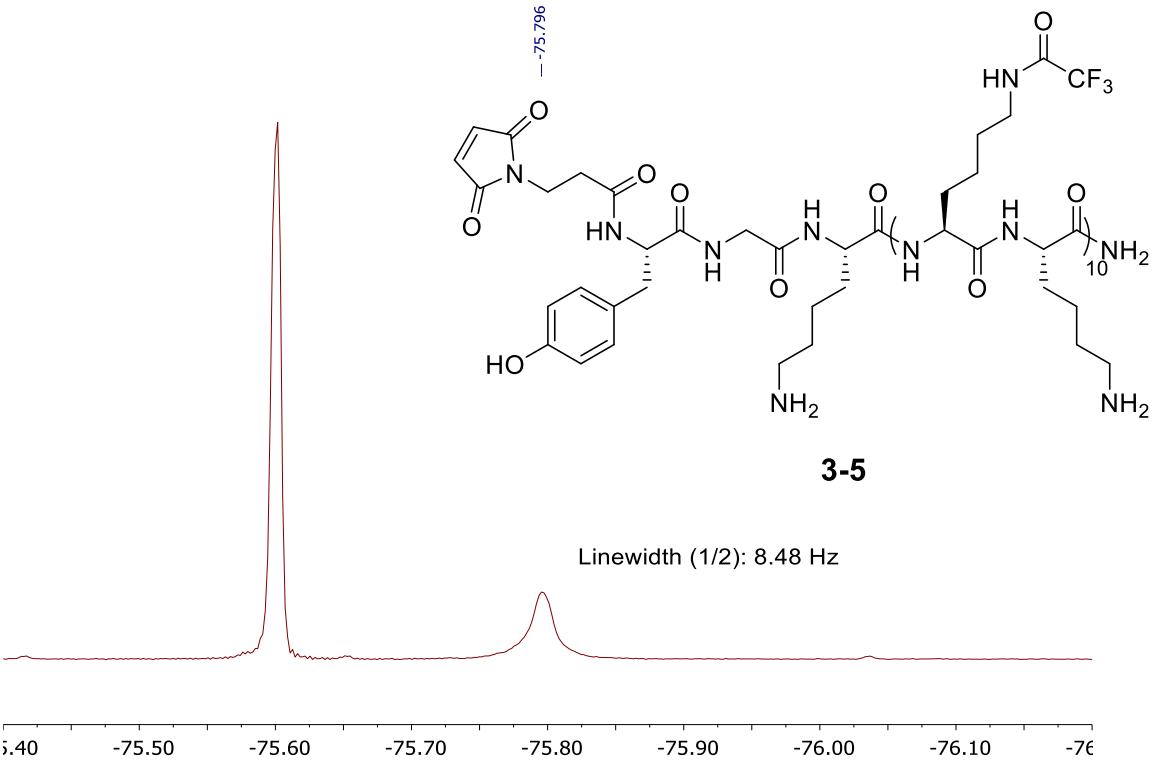
3-4, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min



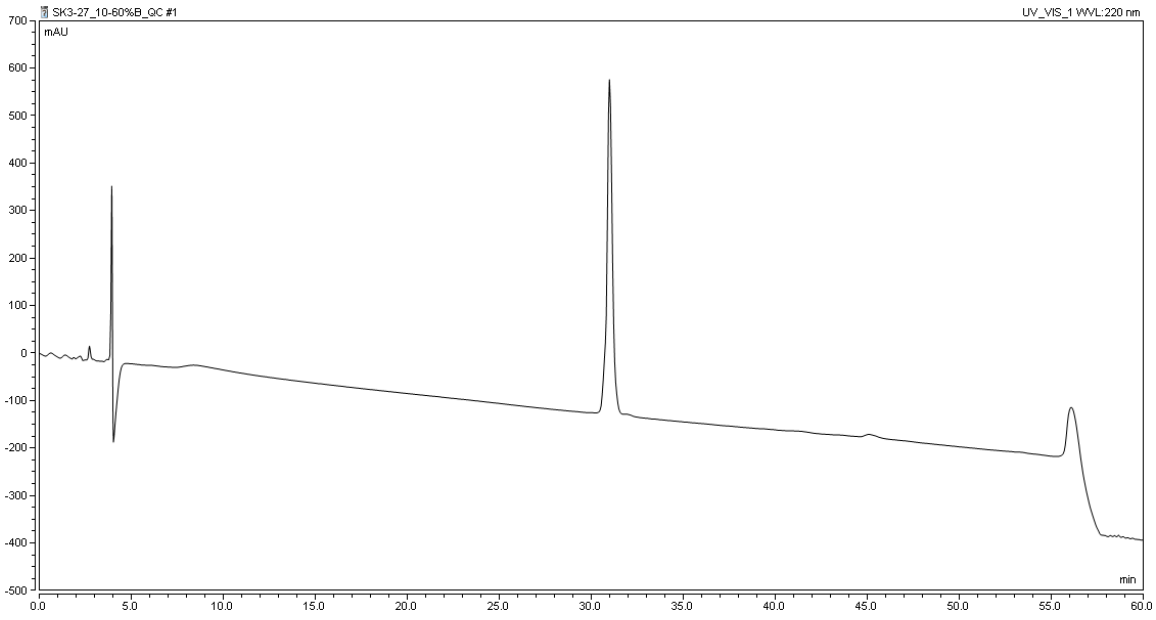
3-4b, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min



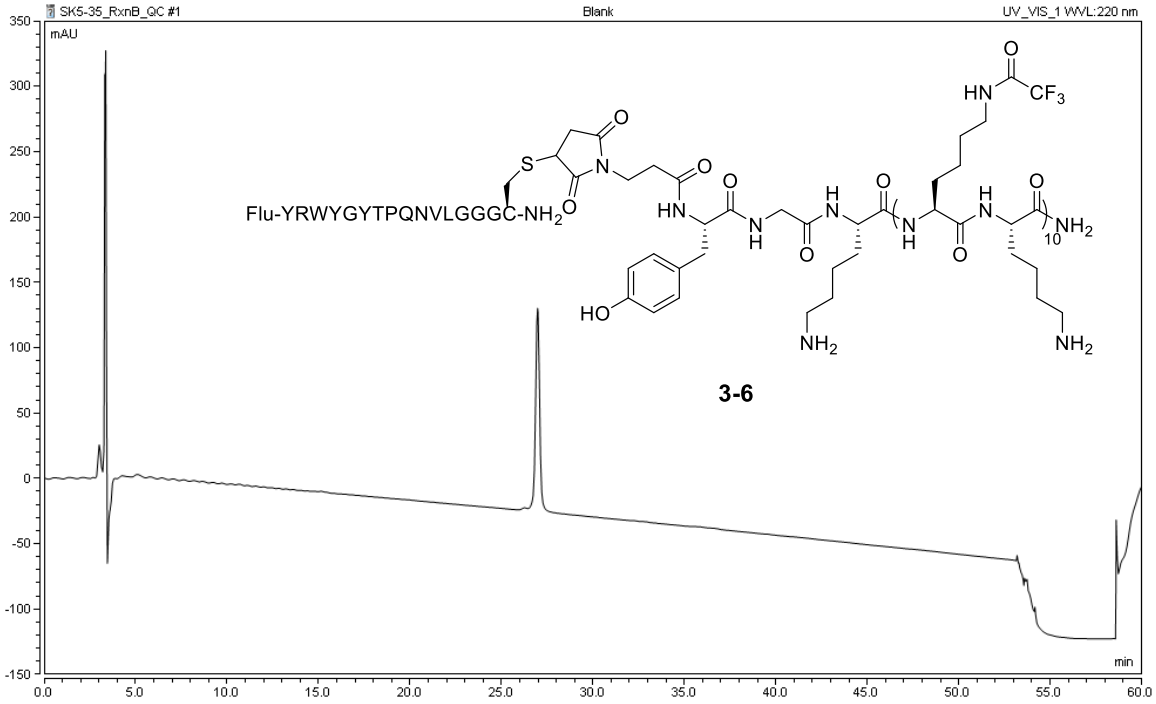
3-5, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O



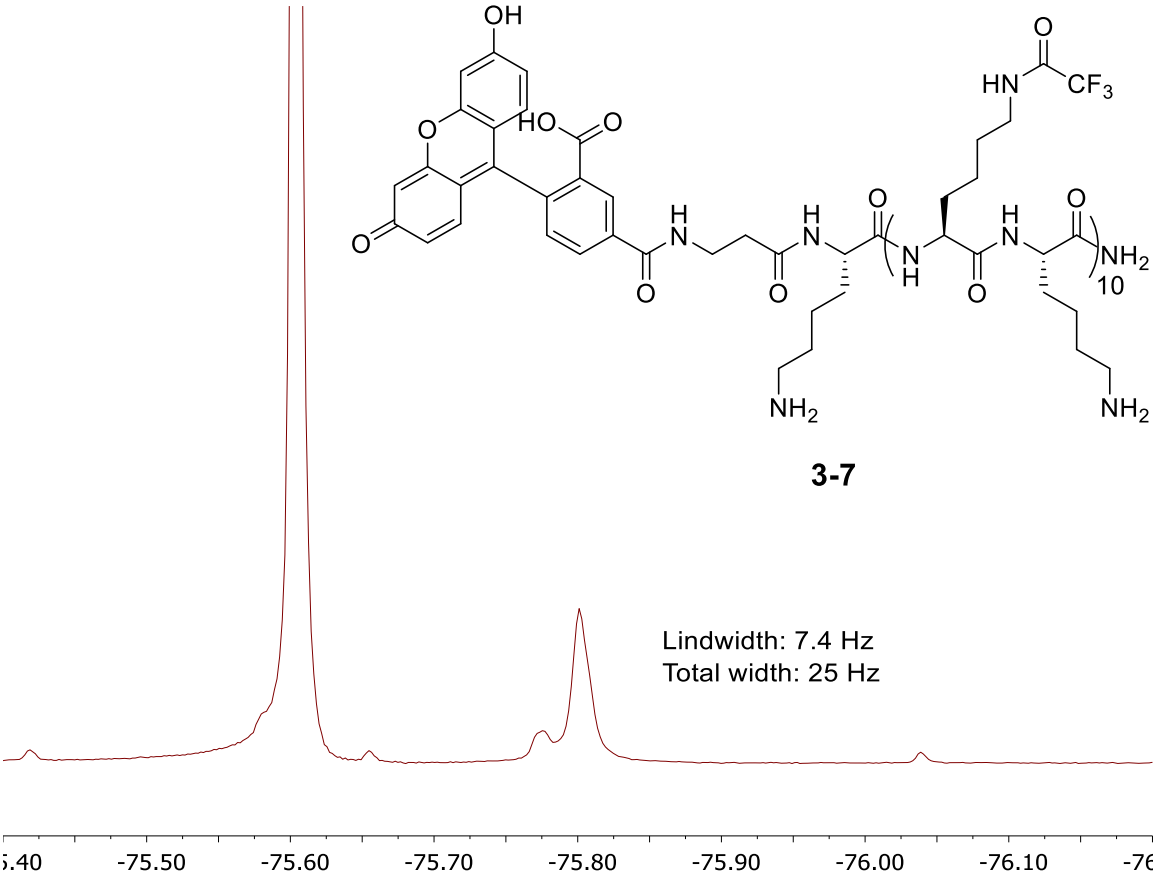
3-5, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min



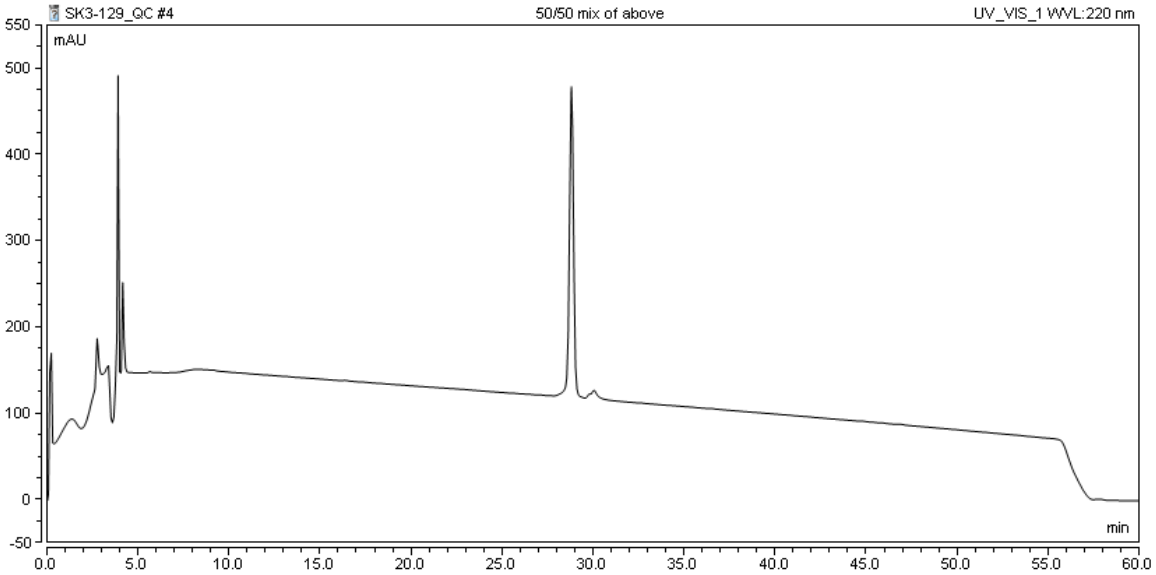
3-6, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min



3-7, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

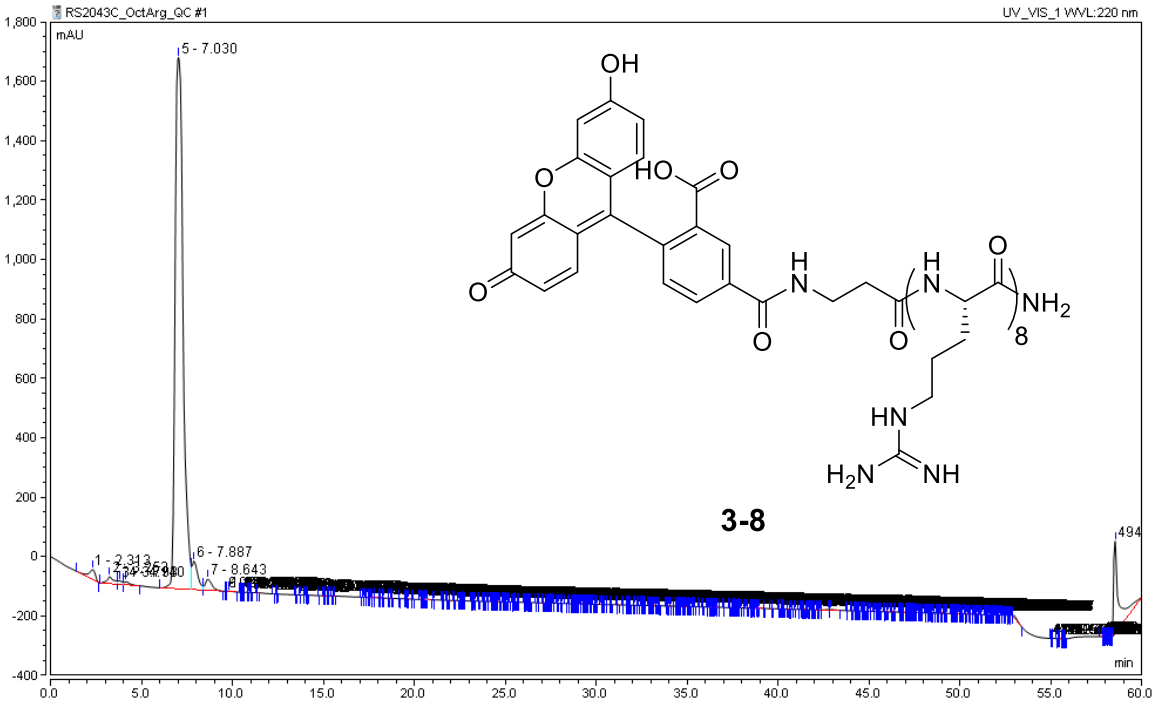


3-7, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min



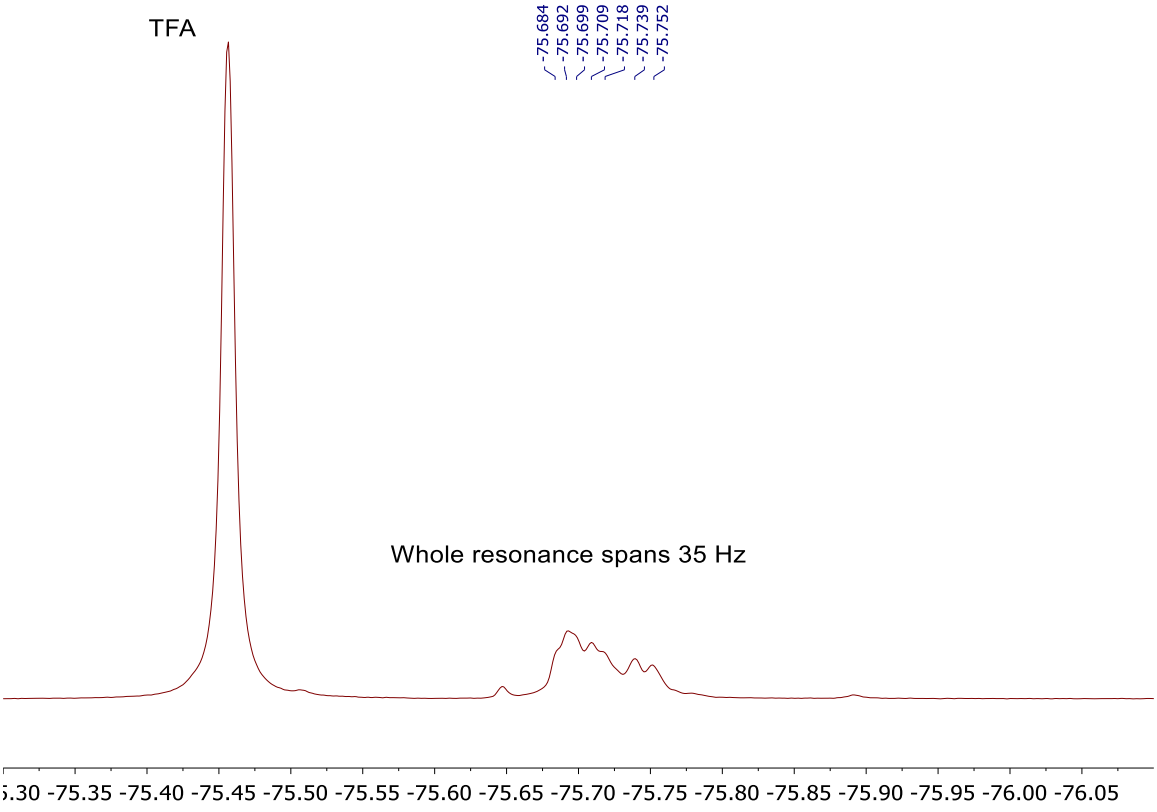
3-8, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

Synthesized by Rachel Staebell, Pomerantz lab

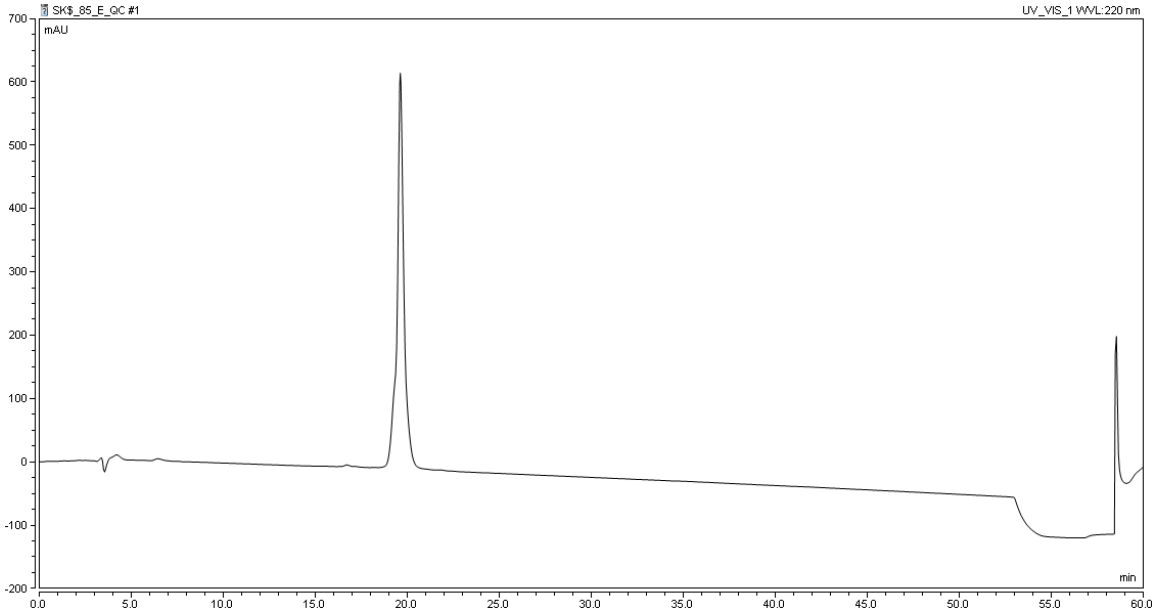


3-9, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-9: Flu-βAla-KK*KK*EK*KK*KK*EK*KK*KK*EK*KK*K-NH_2

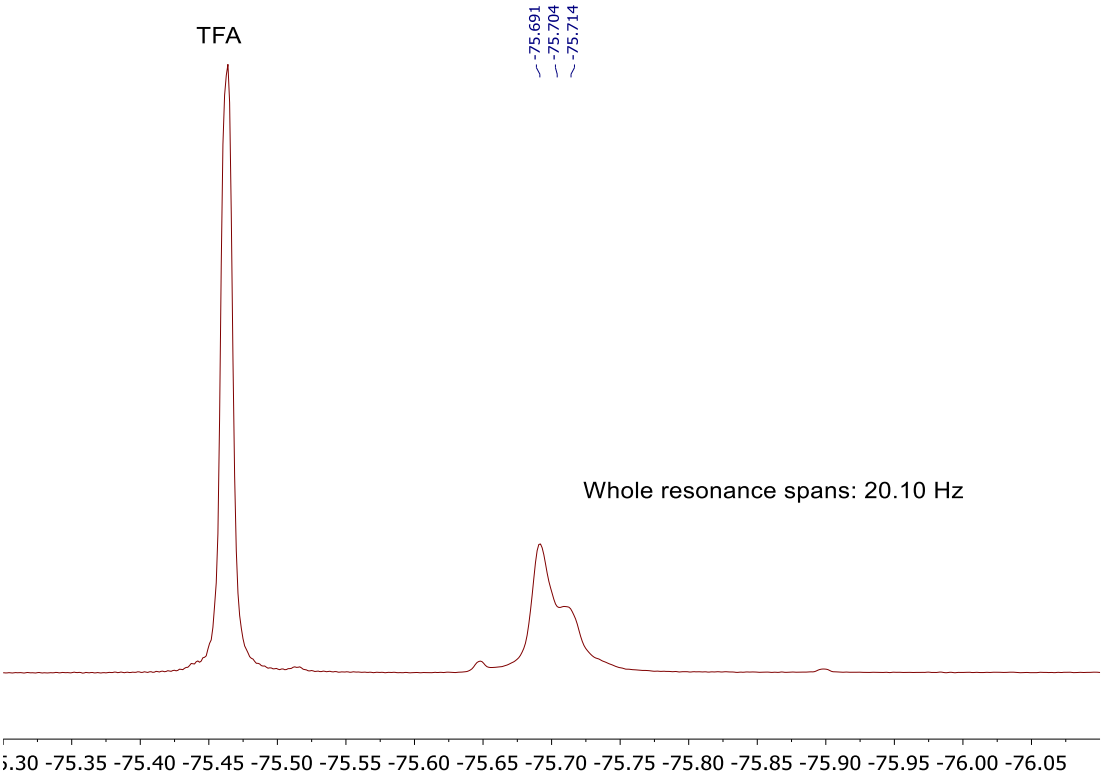


3-9, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

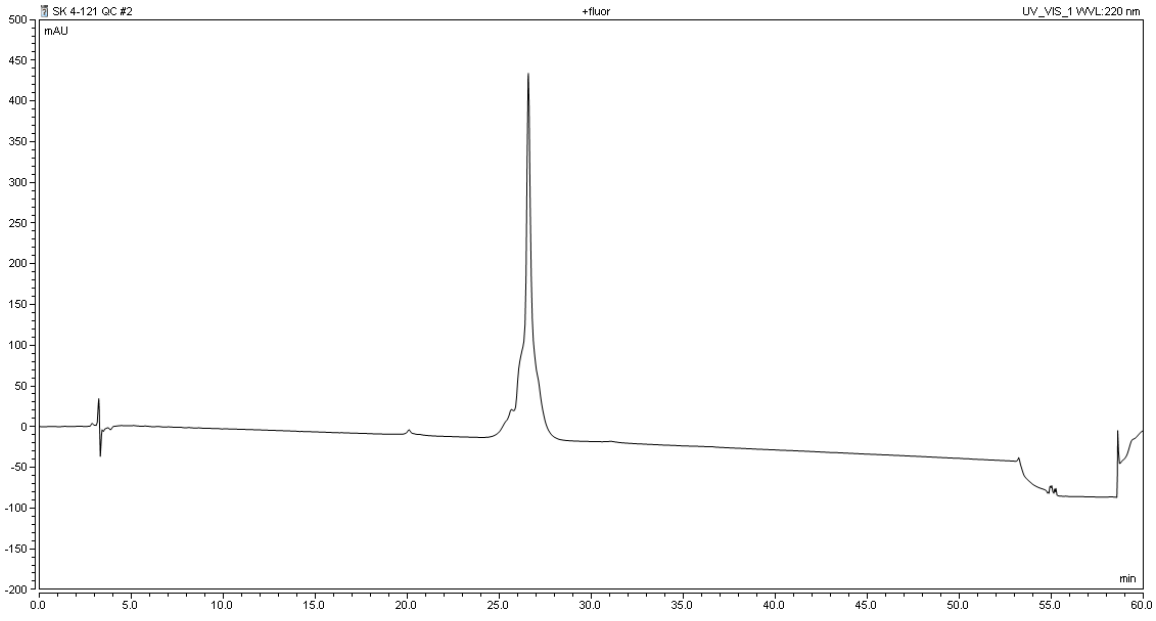


3-10, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-10: Flu-βAla-KK*KK*QK*KK*KK*QK*KK*KK*QK*KK*K-NH_2

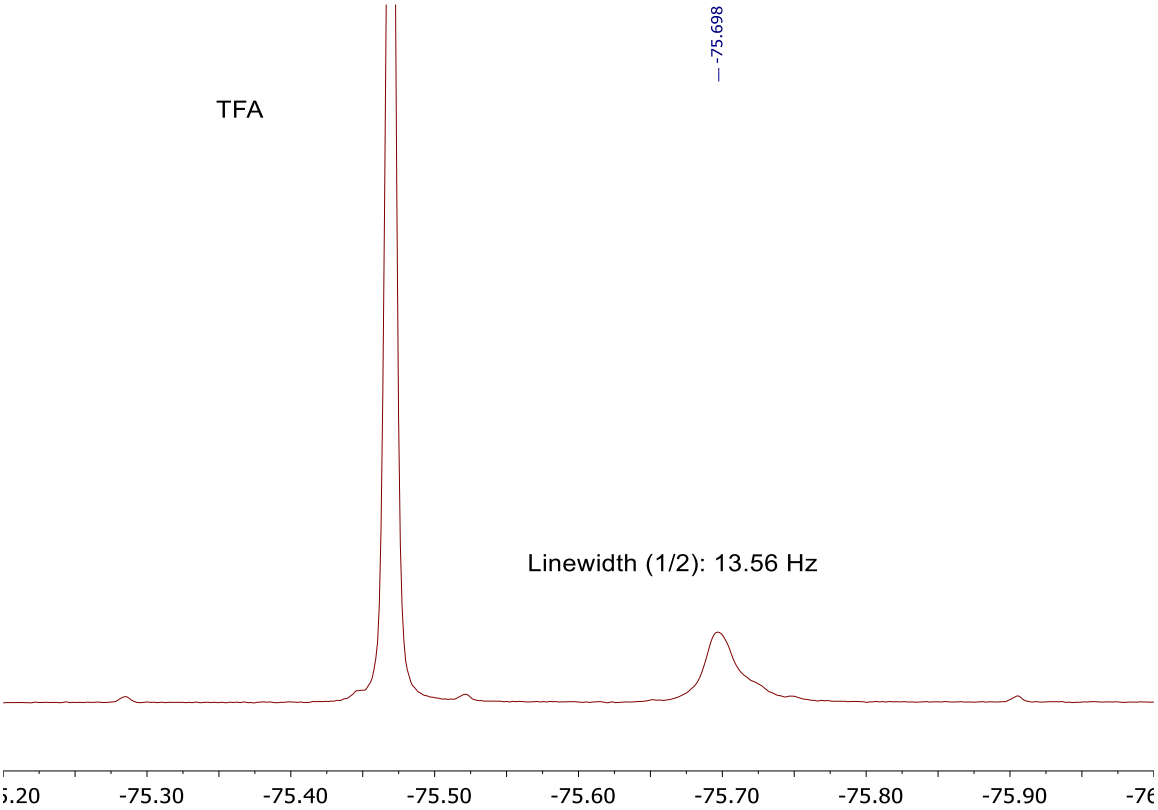


3-10, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

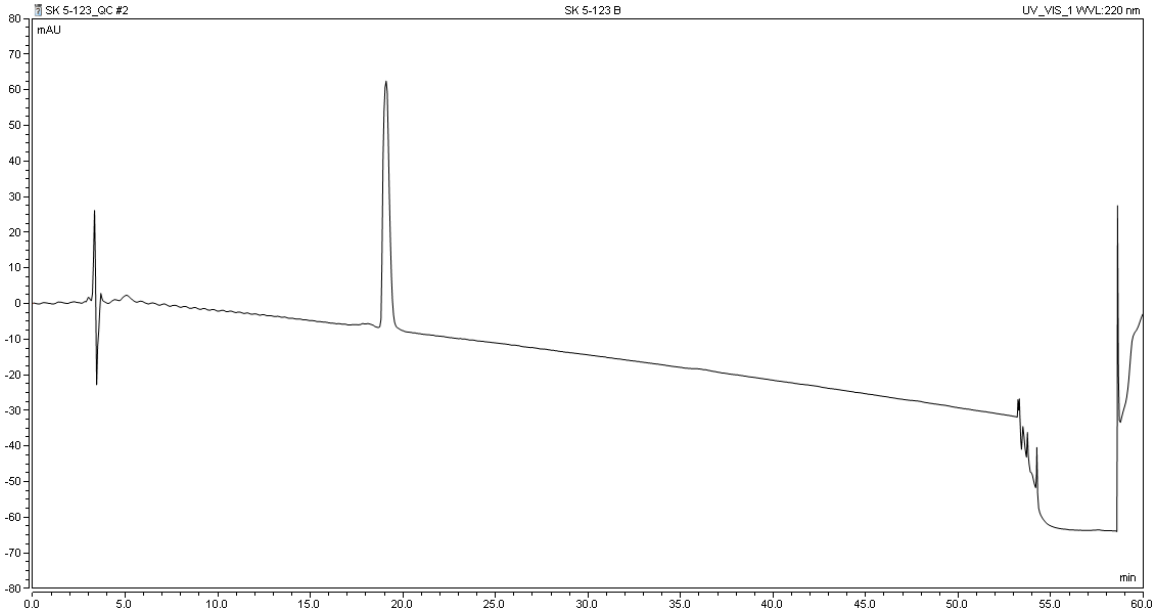


3-11, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-11: Flu-βAla-KK*GK*KK*GK*KK*GK*KK*GK*KK*GK*K-NH_2

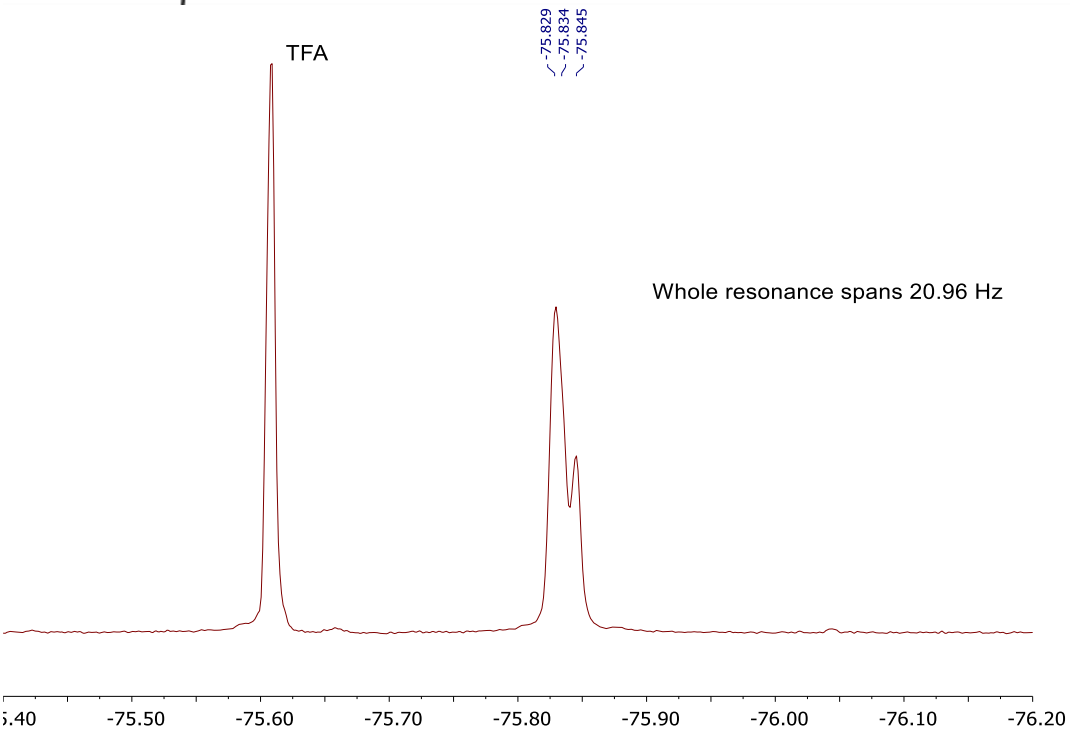


3-11, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

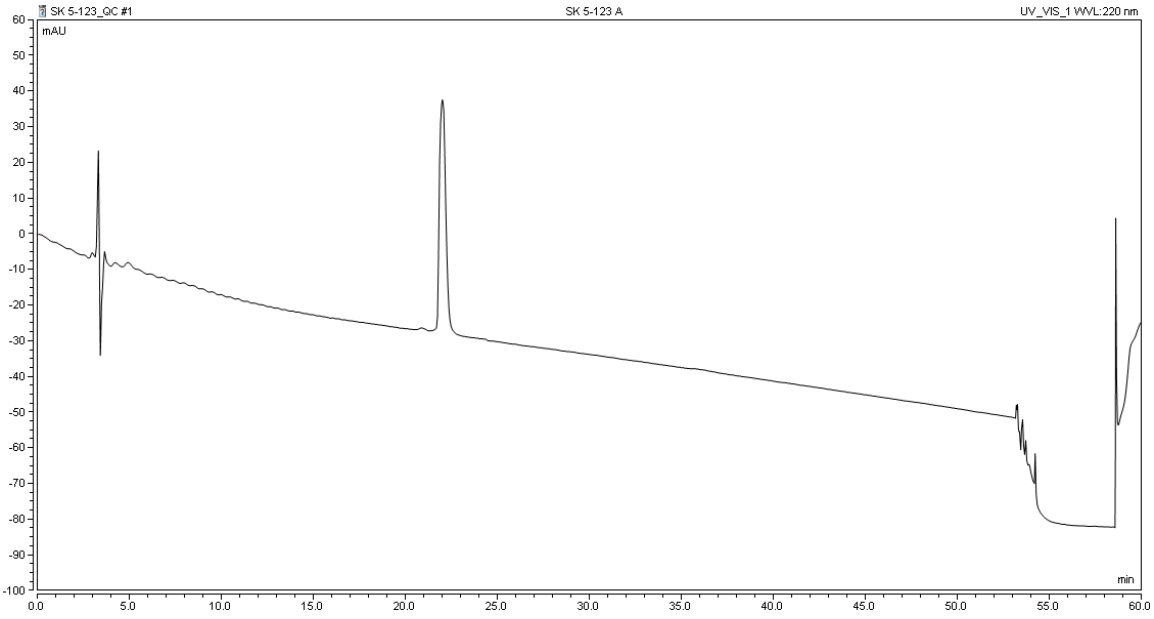


3-11b, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-11b: Flu-βAla-GK*KK*GK*KK*GK*KK*GK*KK*GK*KK*G-NH_2

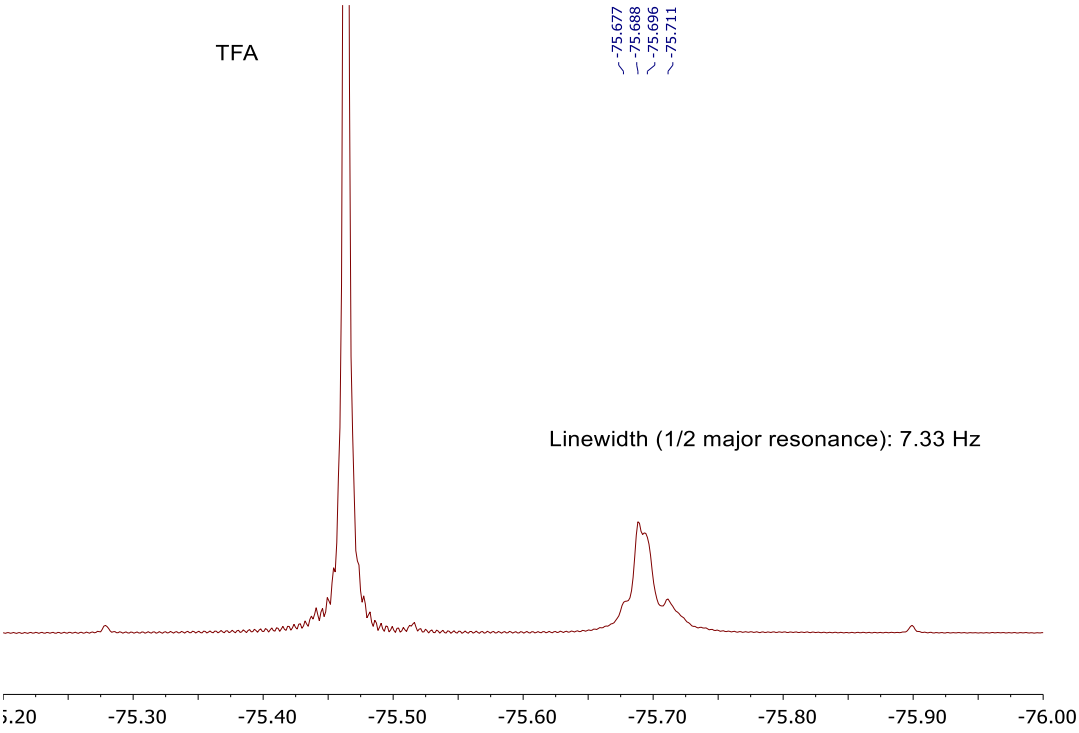


3-11b, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

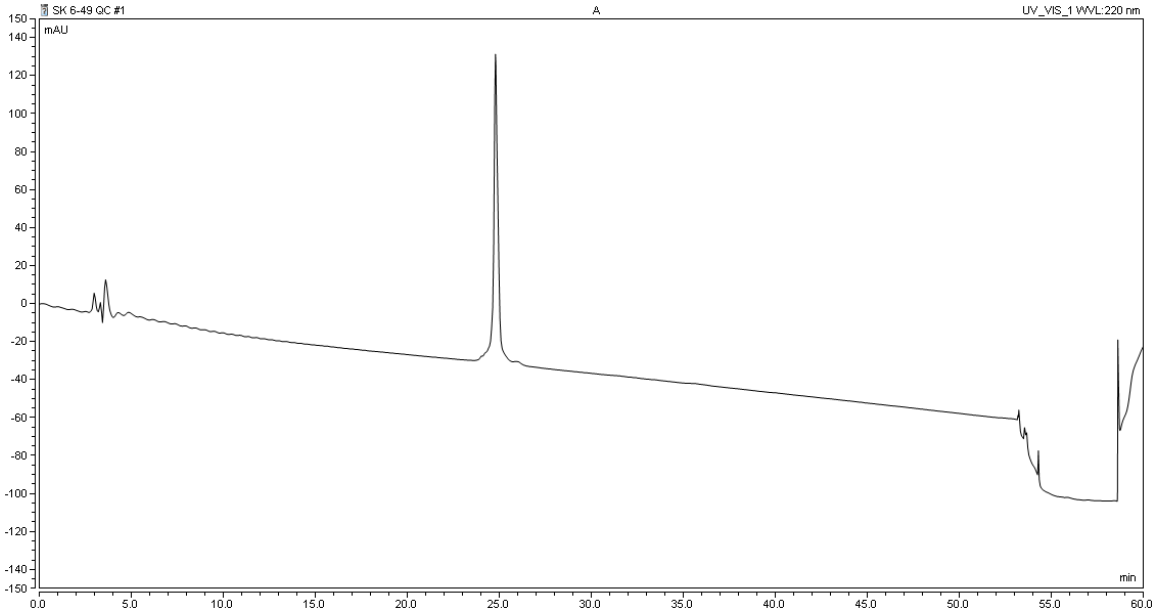


3-12, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-12: TAMRA-βAla-KK*KK*KK*KK*KK*KK*KK*KK*KK*KK*KK*K-NH_2

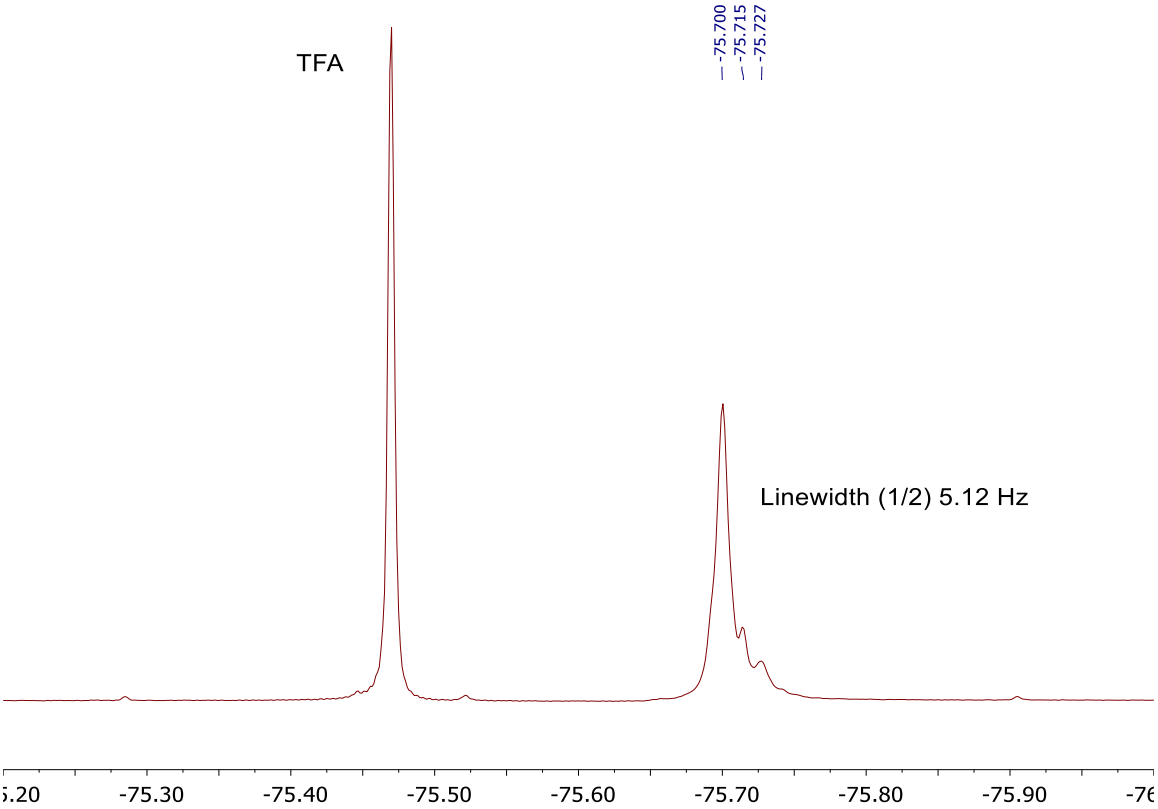


3-12, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

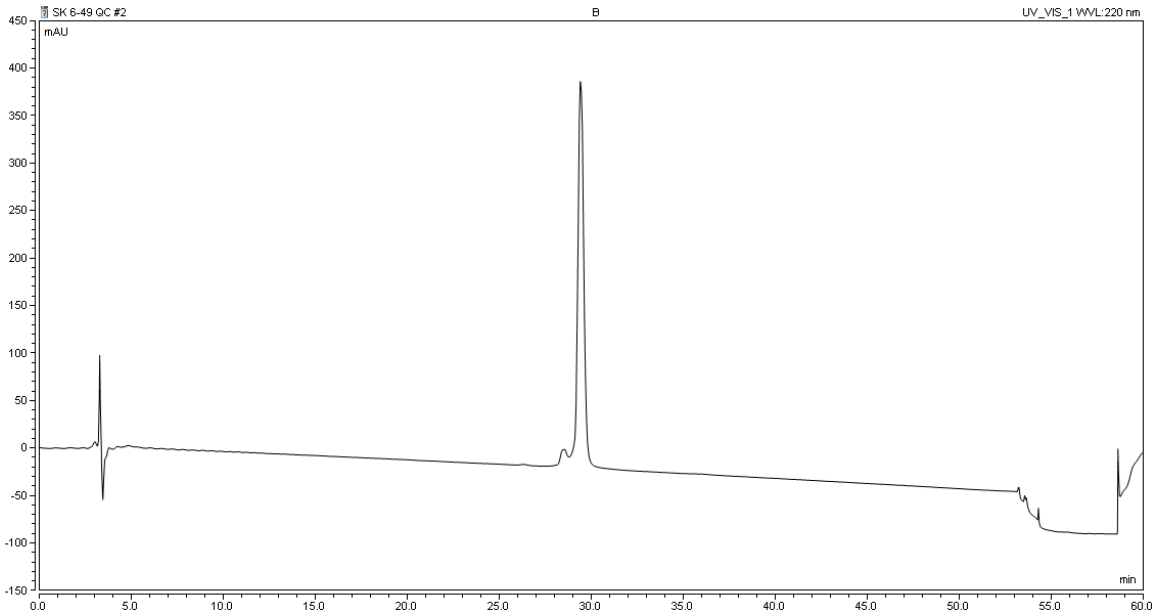


3-13, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-13: TAMRA-βAla-KK*GK*KK*GK*KK*GK*KK*GK*KK*GK*K-NH_2

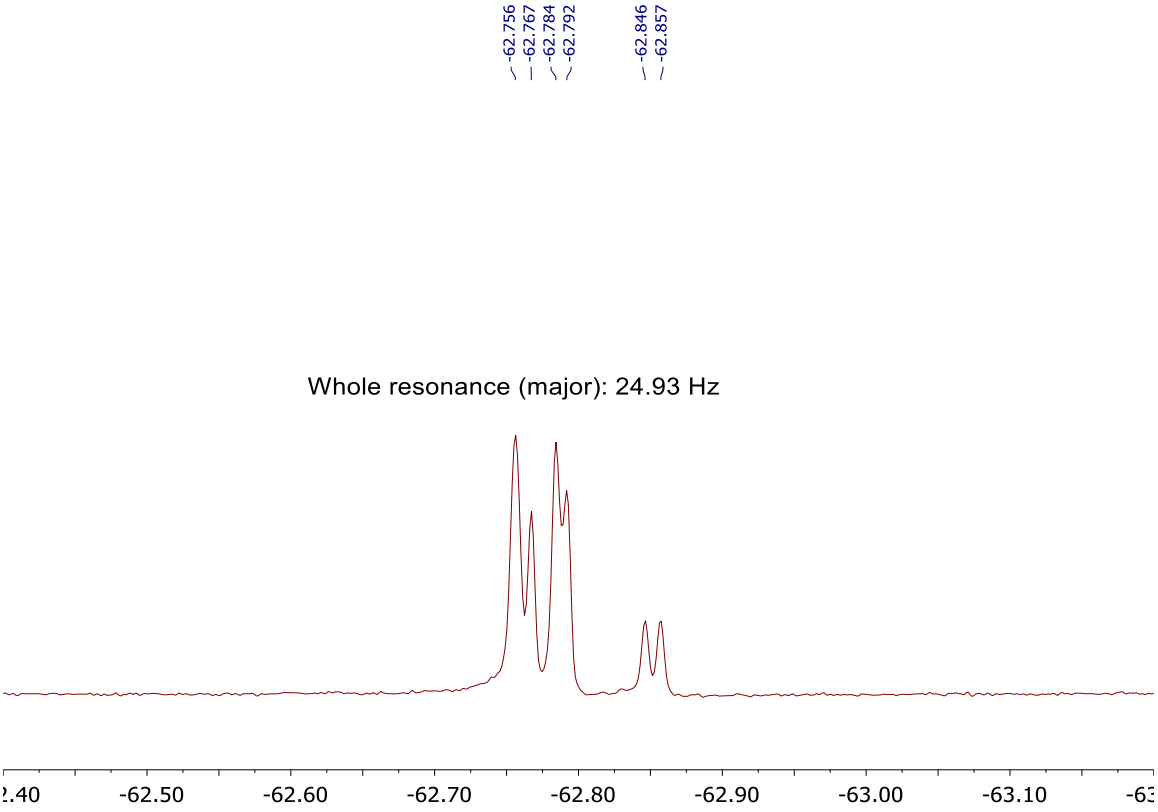


3-13, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

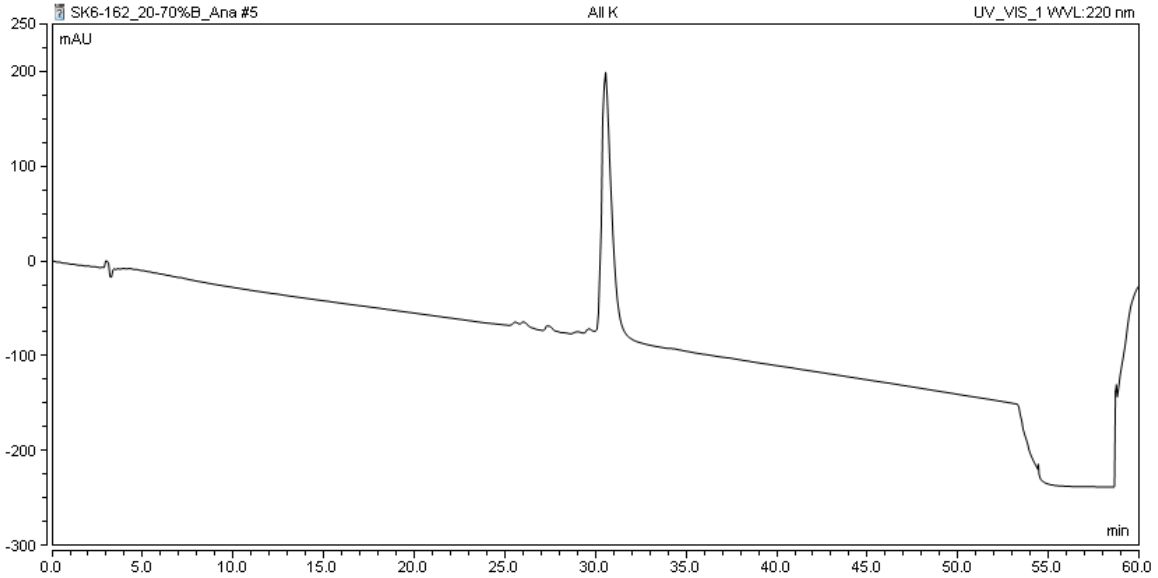


3-14, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-14: Ac-YGKK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}KK^{C}K-NH_2

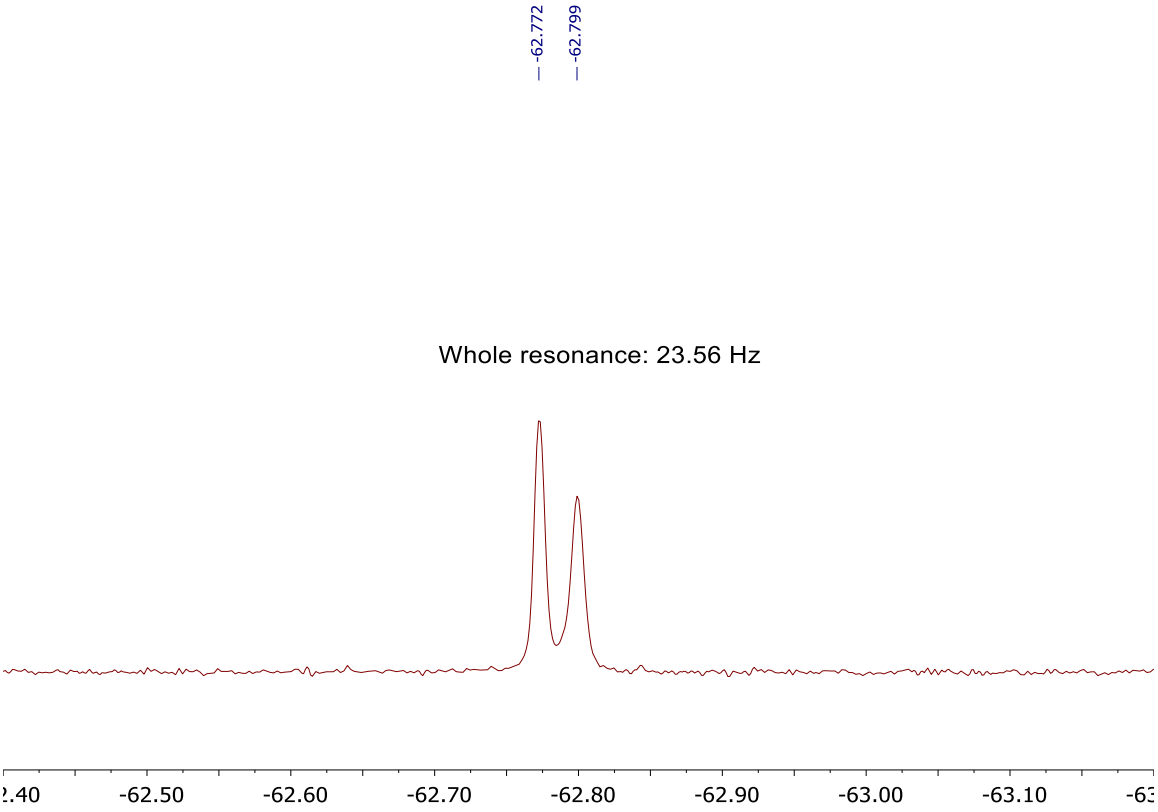


3-14, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

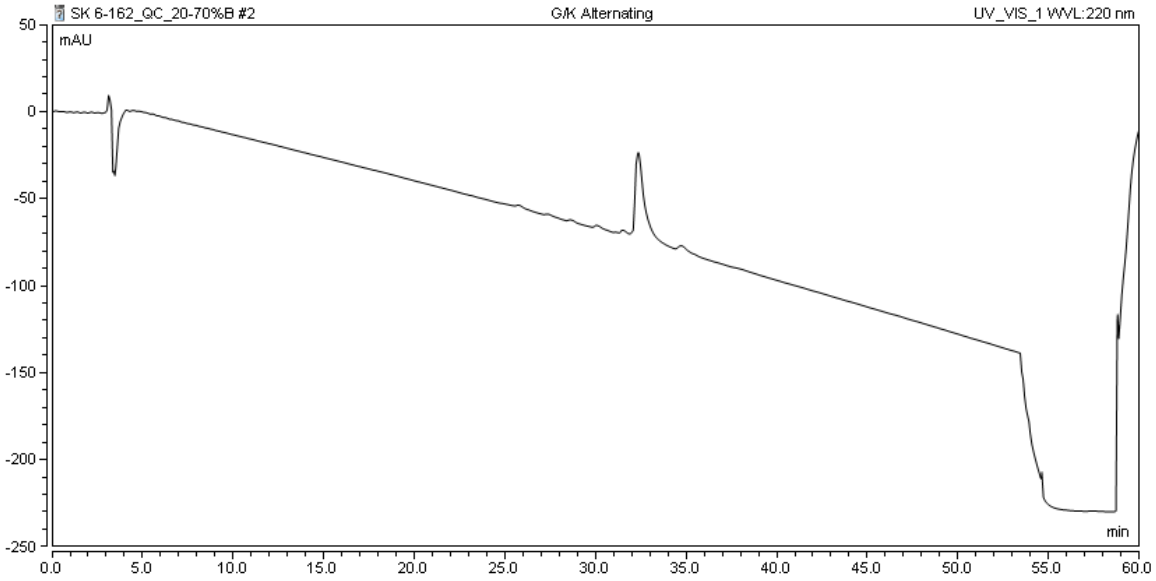


3-15, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-15: Ac-YGKK^{C}GK^{C}KK^{C}GK^{C}KK^{C}GK^{C}KK^{C}GK^{C}KK^{C}GK^{C}K-NH_2

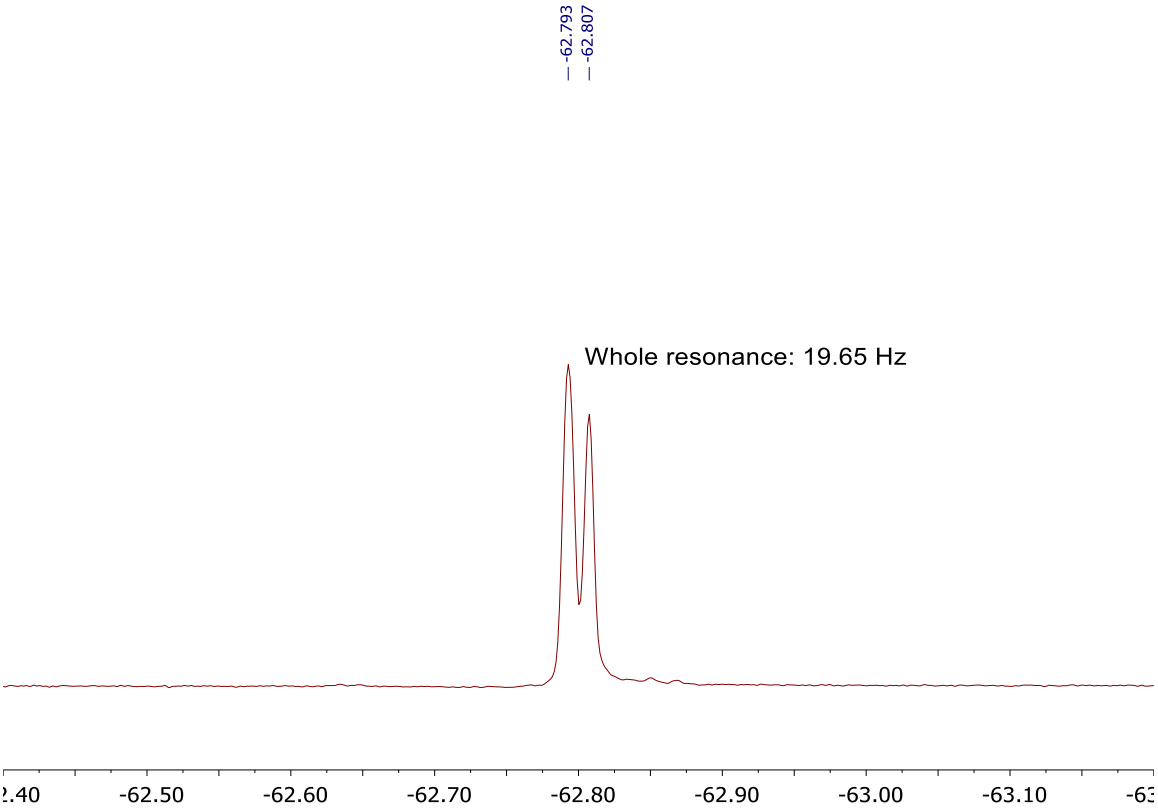


3-15, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min

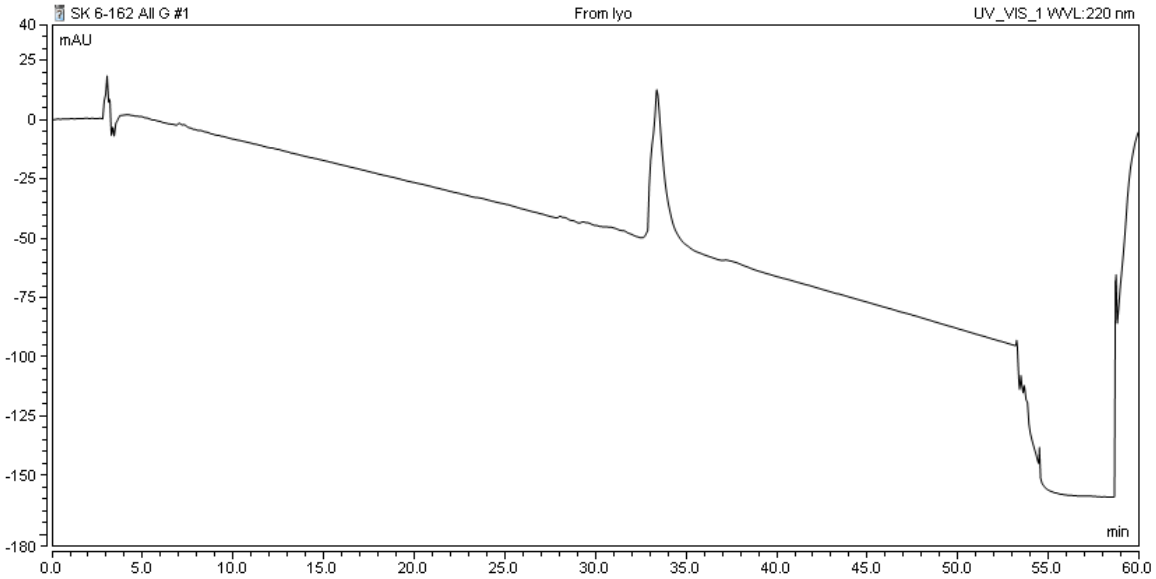


3-16, ^{19}F NMR, 471 MHz, 10% D_2O/H_2O

3-16: Ac-YGGK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}GK^{C}G-NH_2



3-16, RP-HPLC Quality Control. 0.1% TFA/ 10-60% ACN, 1% ACN gradient/min



3.6.7 MALDI-TOF MS data for synthesized peptides

All peptides were analyzed using an AB-Sciex 5800 MALDI/TOF-MS instrument using α-cyano hydroxycinnamic acid as a matrix (saturated solution in 50/50 0.1% TFA/ CH_3CN), with the exception of **3-1**, which was analyzed using a Bruker BioTOF II ESI/TOF-MS directly infused in 0.1% TFA/CH_3CN.

Table 4-1: Mass spectrometry data for synthesized peptides

Entry	**Calculated m/z (g/mol)**	**Observed m/z (g/mol)**	**ion**
3-1	656.2 (z=2) 438.3 (z=3)	656.8 438.2	M+H
3-2	2032.7	2032.0	M+H
3-3	2351.1	2351.4	M+H
3-4	2261.9	2261.9	M+H
3-4b	1987.9	1987.9	M+H
3-5	4061.0	4060.6	M+Na
3-6	6299.9	6300.3	M+H
3-7	4120.9	4120.6	M+Na
3-8	1695.9	1696.3	M+H
3-9	4099.8	4100.4	M+H
3-10	4098.8	4098.3	M+H
3-11	3705.6	3705.8	M+Na
3-11b	3634.5	3634.7	M+Na
3-12	4151.0	4151.3	M+H
3-13	3795.7	3795.1	M+H
3-14	2833.3	2832.9	M+H
3-15	2620.1	2619.6	M+H
3-16	2405.9	2406.4	M+H

Chapter 4

Selectivity, ligand deconstruction, and cellular activity analysis of a BPTF bromodomain inhibitor.

This text is under review for submission as "Selectivity, ligand deconstruction, and cellular activity analysis of a BPTF bromodomain inhibitor." S.E. Kirberger,† P.D. Ycas,† J.A. Johnson, C. Chen, M. Ciccone, R. Woo Wan Lu, A.K. Urick, H. Zahid, K. Shi, H. Aihara, S. D. McAllister, M. Kashani-Sabet, J. Shi, A. Dickson, C.O. dos Santos,* W.C.K. Pomerantz*

†These authors contributed equally to this manuscript

Motivation: The focus of this work is to optimize a small molecule inhibitor for the bromodomain of BPTF, termed AU1. Prior to the 2015 report from Urick et al, there were no small molecule inhibitors for this domain.[97] As BPTF is implicated in numerous diseases and has known binding to c-myc, it has the potential to be an important therapeutic target. My efforts involve synthesizing and testing analogs of AU1 in an attempt to improve its binding affinity and behavior in aqueous solution.

4.1 Introduction

BPTF (bromodomain and PHD finger containing transcription factor) is the largest component of the NURF (nucleosome remodeling factor) complex and is essential for transcriptional regulation of many processes in mammalian chromatin biology including early embryo development,[157] stem cell pluripotency,[94] and T-cell immune function.[157] The 3046 amino acid protein was found to contain a bromodomain and two PHD fingers, domains commonly found on histone and DNA binding proteins (Figure 4-1).[158] To facilitate transcriptional regulation, structural studies have shown that BPTF engages the nucleosome in a bivalent manner through its C-terminal PHD finger and bromodomain via modified histones H3K4me$_3$ and H4K16ac.[65]

Motivating interests for therapeutic development, BPTF associates with the oncoprotein c-MYC and is required for tumorigenesis in high-expressing MYC cells.[84] Dysregulation of the protein has been implicated in a number of diseases including:

pancreatic cancer, melanoma,[159] colorectal cancer,[160] hepatocellular carcinoma,[83] breast cancer,[94] bladder cancer,[161] and lung cancer.[162] Given the relevance of BPTF in disease and its association with NURF, we sought to discover what role the engagement of the bromodomain with chromatin plays in the context of disease states. Using a newly developed protein-observed fluorine NMR (PrOF NMR) assay for bromodomain ligand discovery, we reported the first small molecule inhibitor of the BPTF bromodomain (AU1, ***rac*-1**), selective for BPTF over BRD4.[96,163] ***Rac*-1** was further shown to be cell active, as treatment of breast epithelial cells decreased c-MYC regulated genes, c-MYC chromatin occupancy, and led to G1 cell cycle arrest.[94]

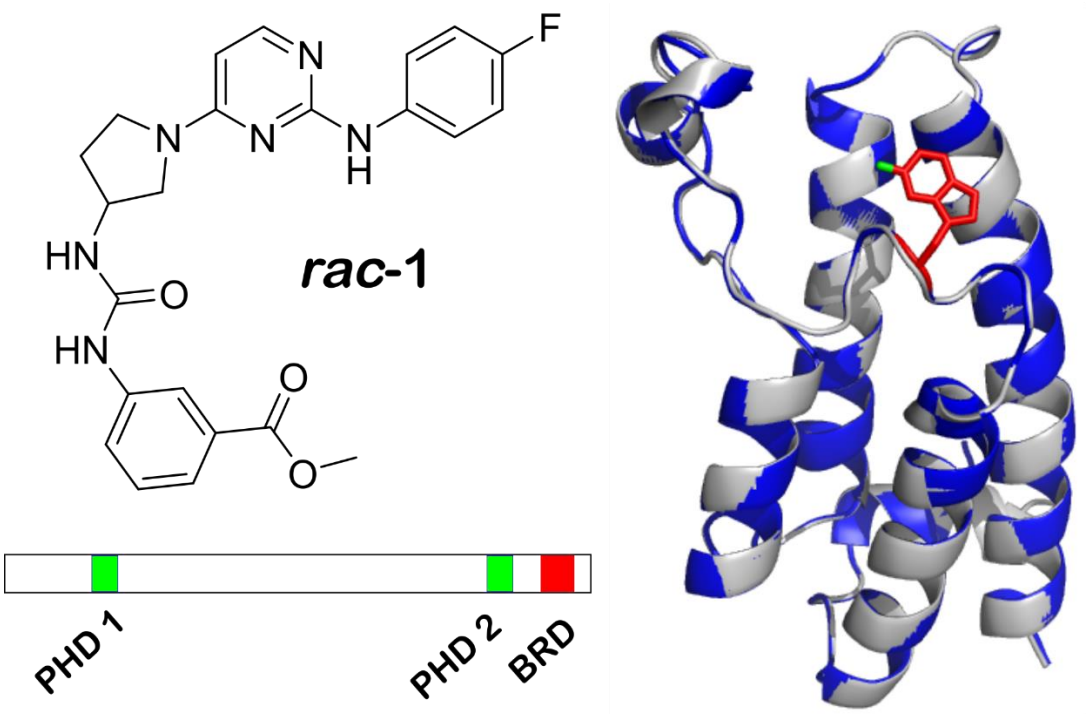


Figure 4-1: Left) structure of racemic AU1 (***rac*-1**). Right) Ribbon diagrams from the crystal structures of the bromodomain of BPTF (gray) overlaid with the bromodomain of 5FW BPTF (blue) with the fluorine of W2824 labeled in green. Bottom) Domain scheme for BPTF with PHD and bromodomain (BRD) labeled.

Despite the biological importance of BPTF, outside of ***rac*-1**, ligands for the BPTF bromodomain are lacking. ***Rac*-1** has only a moderate potency (K_d = 2.8 μM), and its selectivity profile has not been fully characterized. Here we report our efforts to more thoroughly characterize ***rac*-1** and use PrOF NMR in structure-property studies for

improving its metabolic stability and selectivity. We further analyze the mode and efficacy of binding through ligand deconstruction in combination with weighted ensemble molecular dynamics simulations. These studies highlight the versatility of PrOF NMR for bromodomain ligand characterization, and identify areas for improving future BPTF bromodomain inhibitors.

4.2 PrOF NMR identification of the active enantiomer of *rac*-1

For chemical probe development, having access to an inactive stereoisomer with similar physicochemical properties can be a useful control. In the original identification of AU1 (***rac*-1**) as a ligand for the BPTF bromodomain, the racemate was tested. Using our previously reported methods, we synthesized both the (*R*) and (*S*) enantiomers (**(*R*)-1**, and **(*S*)-1**) and tested them using PrOF NMR.[94] PrOF NMR is a structure-based method for characterizing protein binding sites and ligand affinity based on perturbation of fluorine resonances. The BPTF bromodomain contains a single tryptophan at position 2824 in the WPF shelf of the protein near the histone binding site (Figure 4-1). As described previously, for NMR analysis we labeled the BPTF bromodomain with 5-fluorotryptophan (5FW).[99] We monitored the ^{19}F resonance upon addition of both enantiomers. PrOF NMR binding isotherms can be used to quantify the dissociation constant based on measuring dose-dependent changes in chemical shifts of weak protein ligands that exhibit fast chemical exchange rates. In the case of moderate affinity ligands such as ***rac*-1**, significant resonance broadening is observed as the intermediate chemical exchange rate for binding is longer, but not sufficiently long enough to resolve the bound and unbound state. In the PrOF NMR spectrum for the (*S*) enantiomer, more significant broadening of the 5FW resonance into baseline than the racemate was observed. (Figure 3-2B). Conversely, **(*R*)-1** showed no broadening in the NMR assay, indicating that only the (*S*) enantiomer significantly contributed to the binding affinity originally reported, whereas the (*R*) enantiomers of future analogs can serve as useful control molecules.

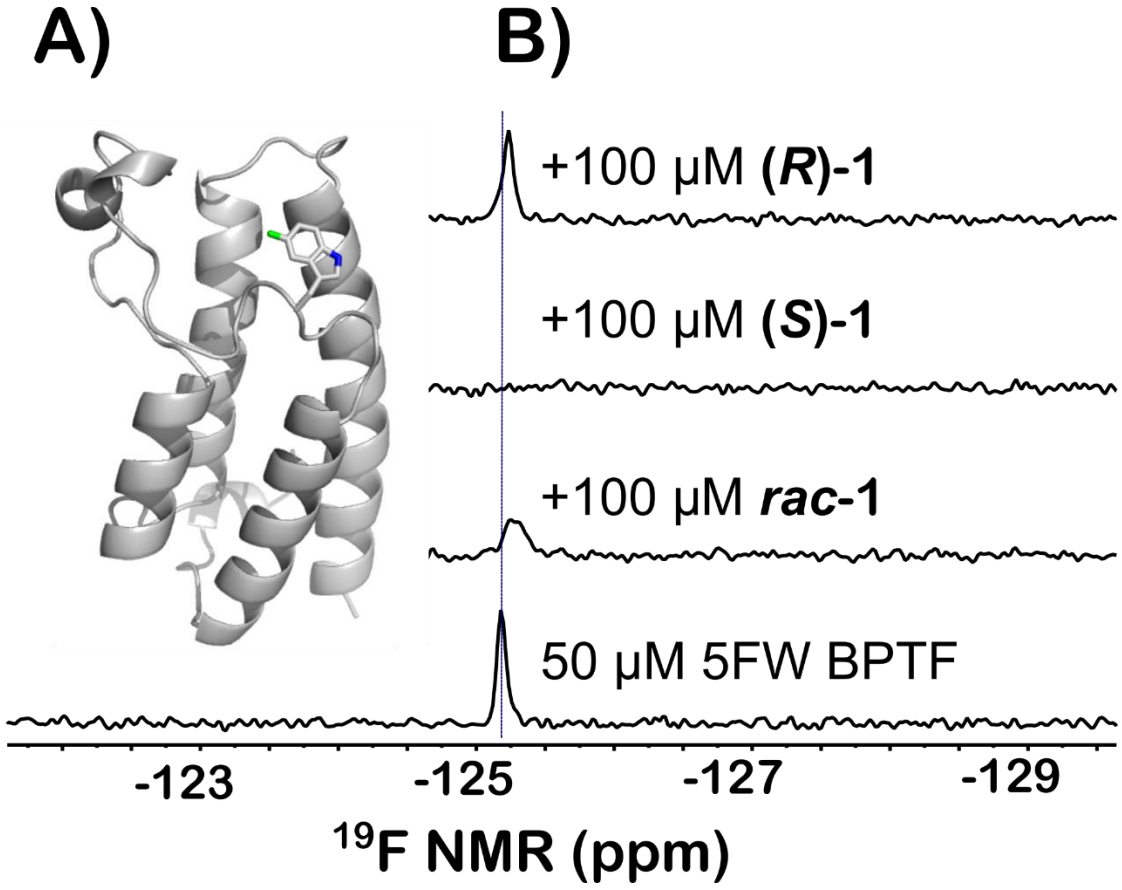


Figure 4-2: A) Ribbon structure of BPTF with 5FW displayed. The fluorine atom is colored in green. B) PrOF NMR using the BPTF bromodomain labeled with 5FW at 2824 of ***rac*-1** and both enantiomers against 5FW BPTF. Each ligand is in a two-fold excess of the protein. Both experiments with ***rac*-1** and **(*S*)-1** resulted in significant broadening of the fluorine resonance.

4.3 X-ray crystal structure of 5FW BPTF supports a lack of structure perturbation from fluorine for PrOF NMR experiments.

To gain further insight into how **(*S*)-1** interacts with the BPTF bromodomain, a co-crystal structure was highly desirable. Although to date we have been unsuccessful, during these studies we investigated the crystallization of the 5FW-BPTF bromodomain. Previously, we characterized the stability of the fluorine-labeled bromodomain by circular dichroism and determined a negligible change in protein thermal stability from fluorination relative to the unlabelled protein.[163] To further verify the native fold is intact, we have now solved the crystal structures of the BPTF bromodomain in both the fluorinated and non-

fluorinated form. Alignment of the structures gave a root-mean-square deviation (RMSD) of 0.1548 Å for the protein backbone and a RMSD of 0.3516 Å for W2824 (See Figure 4-3). Proline 2836, far from the binding site, had the largest RMSD at 0.8093 Å. The similarity in structure of the 5FW and unlabeled BPTF further supports fluorine labeling and PrOF NMR as a viable binding assay. Given the lack of a co-crystal structure of **(*S*)-1),** and further verification of the non-perturbing nature of the fluorine in the bromodomain, we thus turned to indirect methods using PrOF NMR to guide medicinal chemistry efforts using structure-activity relationships (SAR) and ligand deconstruction.

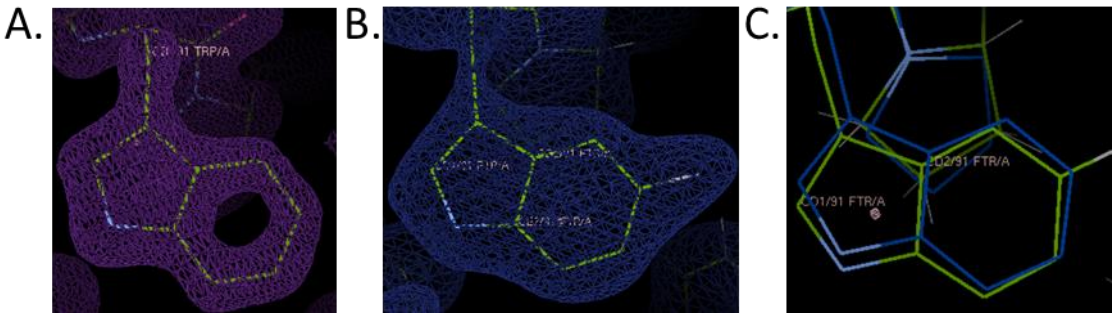


Figure 4-3: A) Electron density and models of W residue for unlabeled BPTF structure. B) Electron density and modeled 5FW residue for the 5FW BPTF structure. C) Overlay of the modeled W (blue) and 5FW (green) residues from the unlabeled and 5FW BPTF structures.

4.4 (*S*)-1 Structure-Activity Relationships.

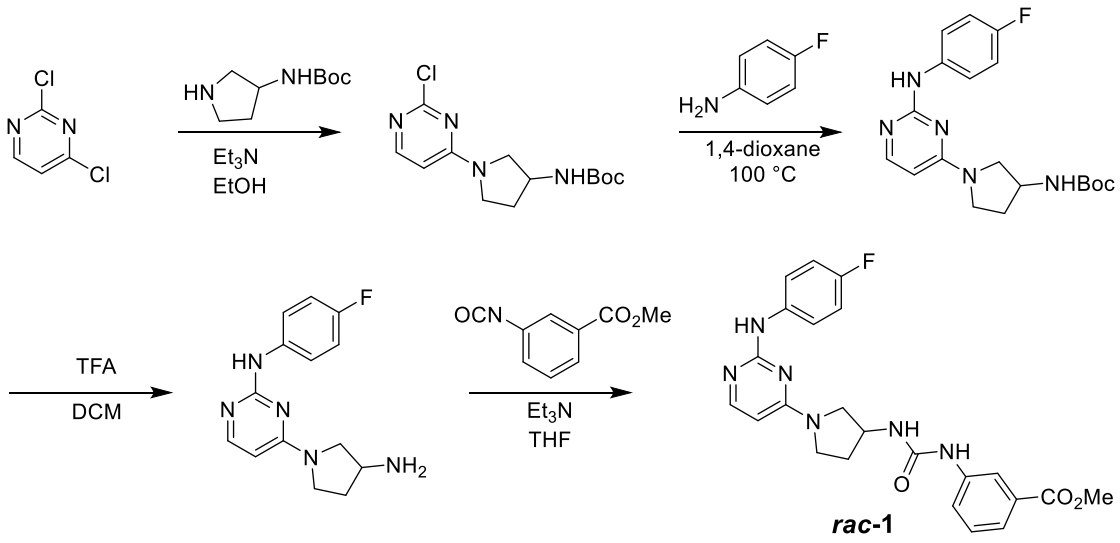


Figure 4-4: General synthetic scheme for ***rac*-1**

In addition to a modest binding affinity of 2.8 μM, ***rac*-1** (Figure 4-4), and by analogy **(*S*)-1,** possess several properties which we wished to improve upon. One significant challenge was the low solubility of ***rac*-1,** as observed by an increase in solution turbidity at concentrations above ~100 μM. Additionally, given its cellular activity, for future use of **(*S*)-1** analogs, we were concerned about the susceptibility to cellular esterase hydrolysis of the methyl ester. Consistent with concern involving metabolic liabilities, pharmacokinetics (PK) studies of **(*S*)-1** in mice, produced a C_{max} of 4,540 ng/mL (10 μM) in the plasma, and a rapid half-life at 0.68 h (Figure 4-5).

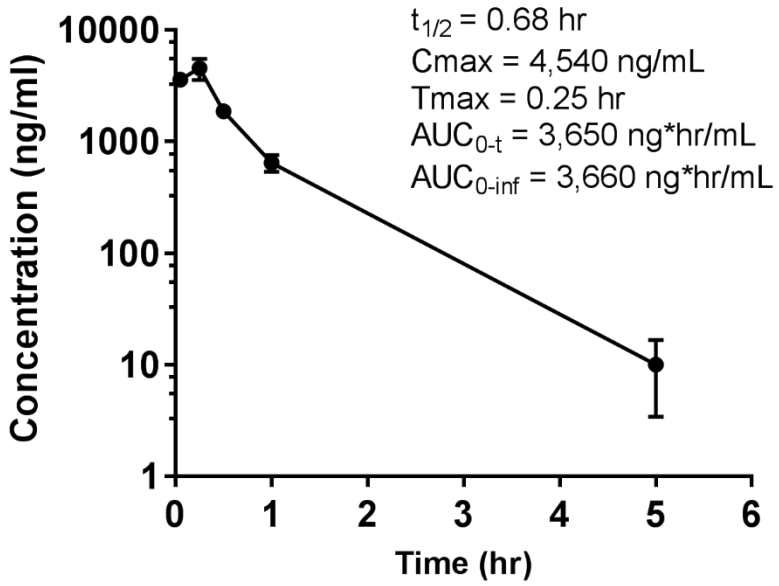


Figure 4-5: Plasma concentration-time curve for **(*S*)-1** (ng/mL) and calculated PK parameters. Due to complications with injections, the 5 and 30-minute data points only represent data from one mouse

Given the concerns above, we took advantage of our four-step synthesis for SAR studies (Figure 4-4). To test for ways to improve the stability of our compound based on ester hydrolysis concerns, we synthesized new analogs of **(*S*)-1** (Figure 4-6). We first tested the would-be product from ester hydrolysis, **12**, against 5FW BPTF by PrOF NMR and determined that binding to the bromodomain was eliminated. To improve the stability of the scaffold in the presence of cellular esterases, we replaced the ester group of **(*S*)-1** with a series of amides, **9-11**, to reduce susceptibility to esterase activity in vivo and cell-based experiments. For these analogs, only small amides were tolerated although binding was weakened. The methyl amide (**9**) showed over an order of magnitude loss in affinity relative to ***rac*-1**. Fast chemical exchange allowed for titration to derive a K_d = 30 μM (Figure 4-7). The ethyl amide (**10**) also showed fast exchange binding behaviour consistent with a weak interaction, but the affinity was unable to be quantified. A larger amide (**11**) resulted in a complete loss of binding. Unfortunately, upon changing the ester to an amide, we also observed an attenuated selectivity with respect to BRD4(1) by PrOF NMR (Figure

4-8). Future analogs containing the more stable methyl amide will need to have improved binding activity and selectivity through modifications on alternative sites on the molecule.

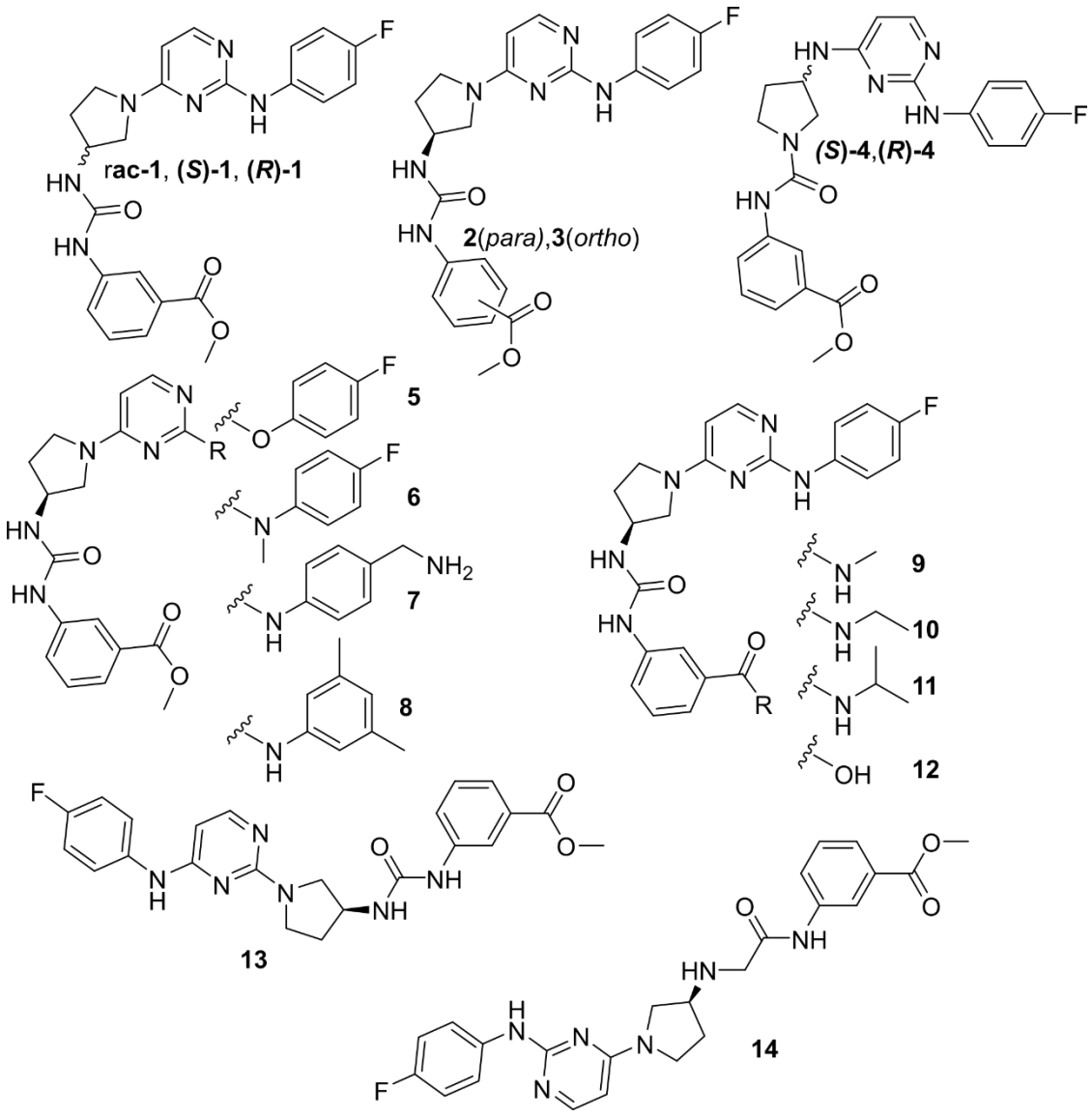


Figure 4-6: Structures of select ***rac*-1** analogs.

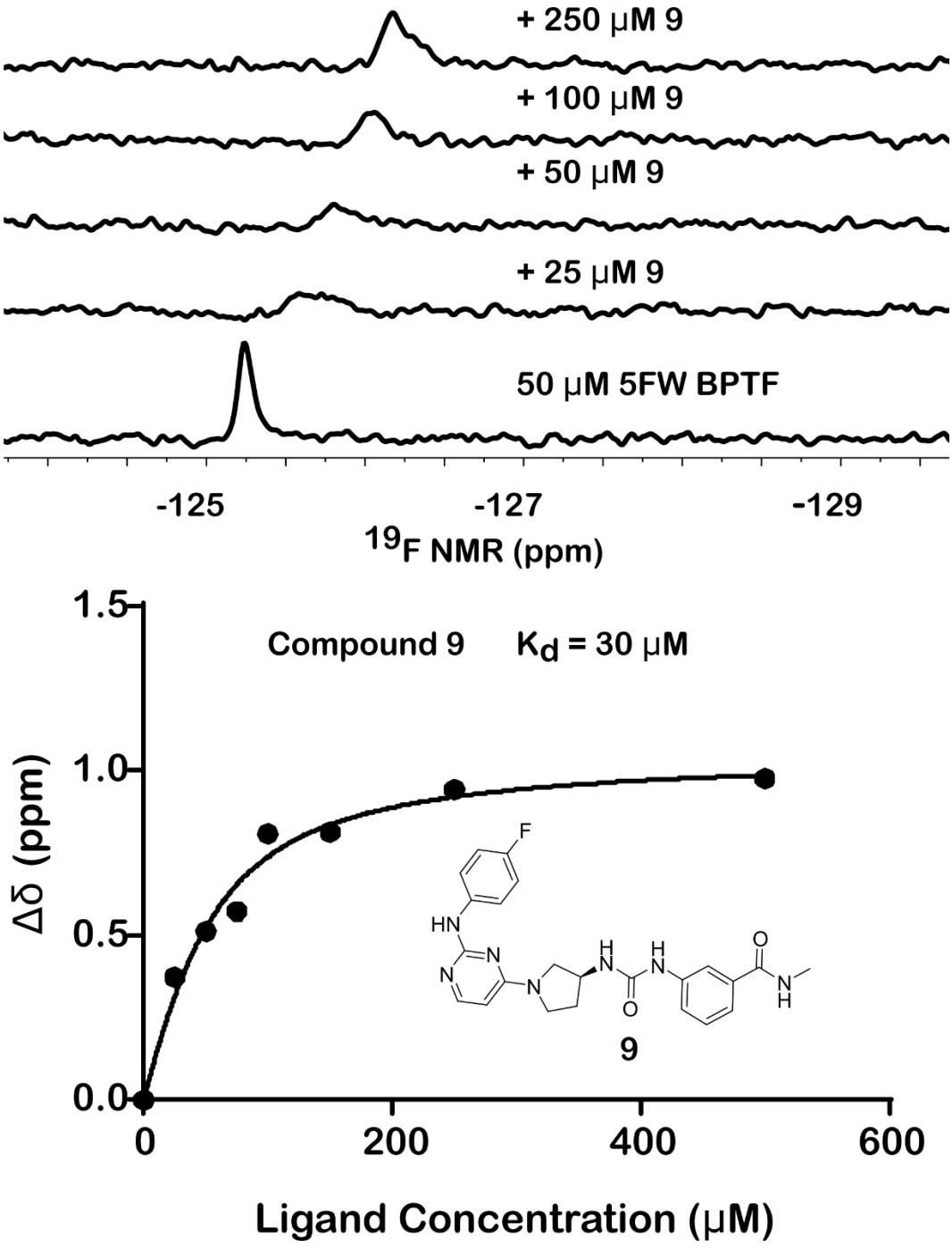


Figure 4-7: Sample titration data of **9**, a broadening and shifting of the 5FW BPTF resonance is observed. Full titration can be found in the Section 4.12.

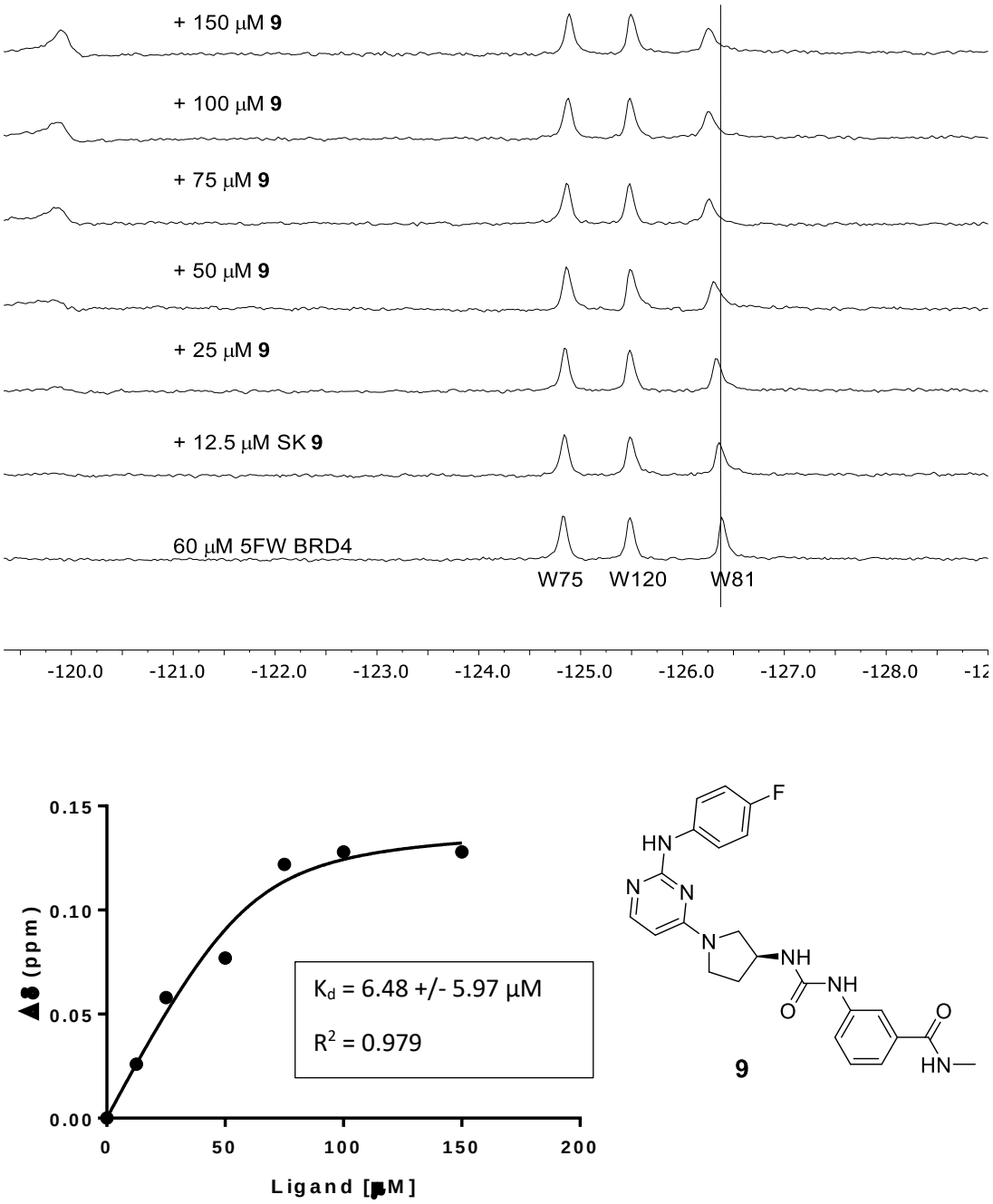


Figure 4-8: PrOF NMR titration of **9** with 5FW BRD4(BD1)

Due to the importance of the ester, we also explored isomers of the compound, substituting the methyl ester around the aromatic ring (**2** and **3**). In these cases, changing the *meta* substitution of the methyl ester on the aryl urea showed a complete loss of binding

in both **2** and **3**, indicating *meta* substituted methyl esters were critical for binding. Conversely, substitutions made on the aniline ring consistently show binding activity through our PrOF NMR assay. Substituting the fluorine for a benzyl amine (**7**) as well as a 3,5-dimethyl substitution pattern (**8**) show similar characteristics as **(*S*)-1** in an NMR titration, indicating a reduced sensitivity to compound modification. The regioisomer of **(*S*)-*1***, **13**, also showed similar binding characteristics to the original scaffold by PrOF NMR. We conclude from these PrOF NMR SAR studies, that the amino-pyrimidine terminus of the molecule may be more solvent exposed when binding to the protein, whereas the methyl ester is bound more deeply in the bromodomain binding site.

(*S*)-1 has limited solubility in aqueous solution and the broadening of its ^{19}F resonance by NMR indicates a highly associated state. We looked at the urea moiety within the scaffold as a potential source of aggregation and sought to alter this functional group to improve the behavior of **(*S*)-1** in aqueous solution. The urea was separated by inserting a methylene as seen in compound **14**, drastically improving the water solubility to > 1 mM at pH 7.4 as assessed by a lack of solution turbidity. However, we observed a reduction in binding affinity using PrOF NMR titrations (K_d = 39 μM, Figure 4-9) when the urea was interrupted. We also attempted to reverse the substitution of the pyrrolidine ring with the hopes of interrupting the aggregation potential of the urea by reducing the hydrogen bond capability of one of the nitrogens (**(*S*)-4** and **(*R*)-4**). Upon this change, we observed by PrOF NMR in **(*S*)-4** a stoichiometric binding at the concentration of protein that we used, indicating that its binding affinity was at least an order of magnitude below our protein concentration, i.e., 5 μM or lower. We then had this compound tested via AlphaScreen (Reaction Biology) as a competition-based assay between the bromodomain and an acylated histone; however, an IC_{50} of 70 μM was observed. This difference in binding affinity observed via PrOF NMR and AlphaScreen could potentially be due to the reduced DMSO concentration necessary for the the AlphaScreen assay 0.25% vs. the 1% used in PrOF NMR, this reduction in co-solvent concentration limits the solubility of this analog for the assay, and points to further improvements that are needed for solubility enhancement of our compounds.

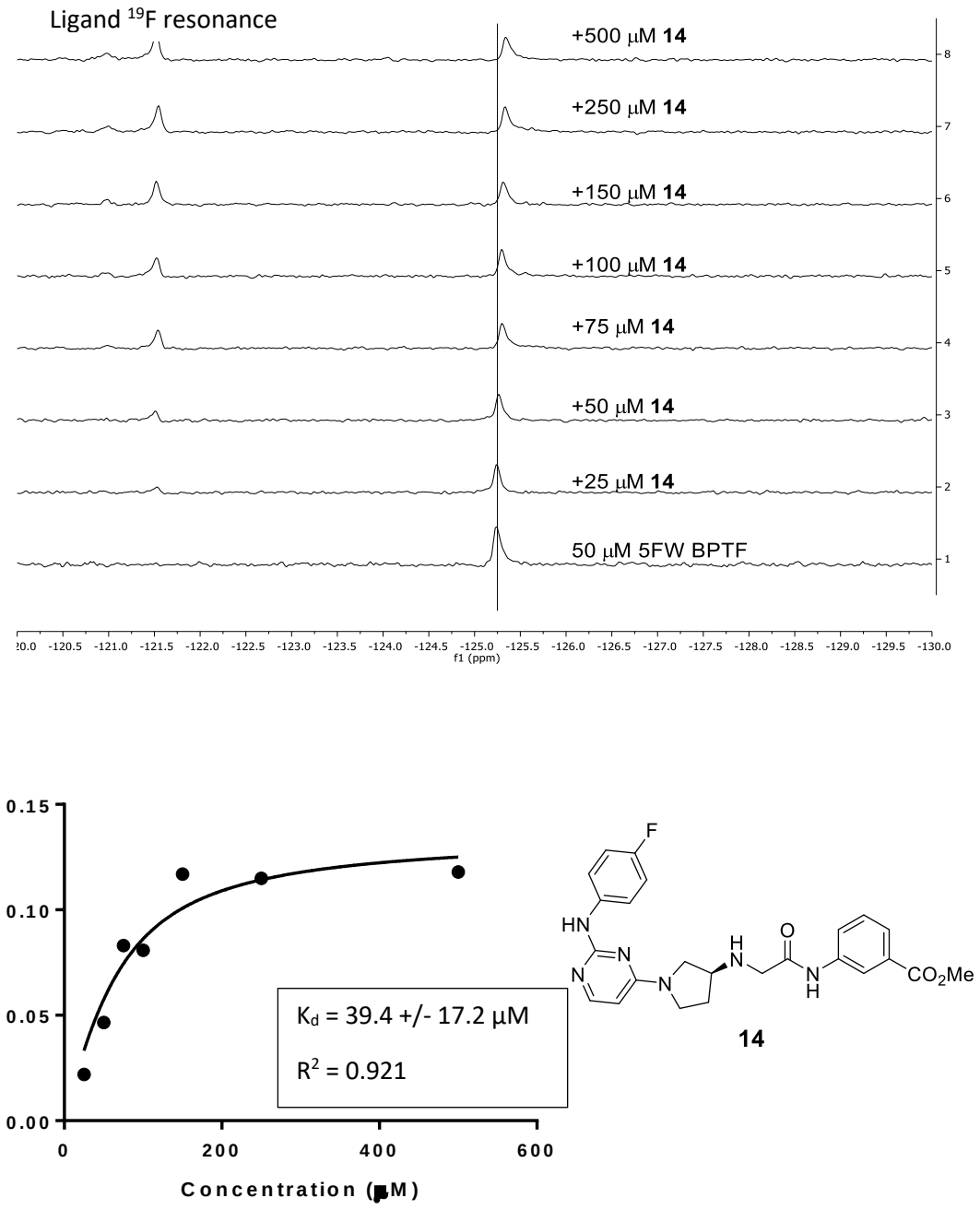


Figure 4-9: PrOF NMR titration of **14** with 5FW BPTF

4.5 Additional SAR

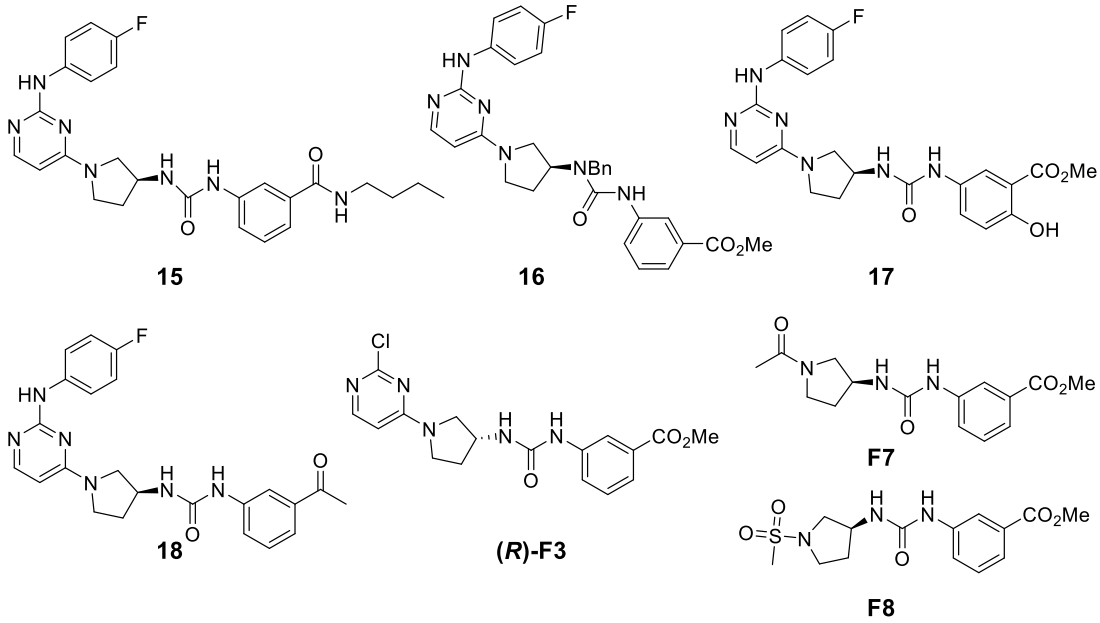
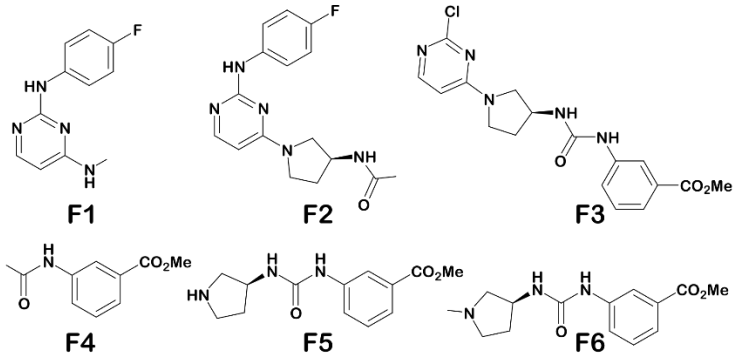


Figure 4-10: Additional **(*S*)-1** analogs

SAR studies were expanded beyond those mentioned in Section 4.4 and were not included in the manuscript describing this work. Seven additional molecules were synthesized (Figure 4-10) and tested via our PrOF NMR assay. N-Butylamide, **15**, continues the line of investigation described by compounds **9-11**. Continuing the trend observed, **15** showed no binding via PrOF NMR, indicating that amides in this position larger than ethyl inhibit binding. The goal of **16** was to attempt to inhibit the hydrogen bonding potential of the urea moiety by alkylating one of the protic nitrogens, this however completely erased binding of the molecule. In an attempt to increase the hydrolytic stability of the methyl ester, **17** was synthesized with an ortho hydroxyl to stabilize the ester. The addition of the ortho hydroxyl also prevented binding to the bromodomain in our PrOF NMR assay. Lastly, we tested **18**, with the exchange of ester for a ketone, and this encouragingly showed binding in a dose dependent manner with a K_d = 1 +/- 3 μM, approximating the binding affinity of ***rac*-1**. Future selectivity studies or analogs could potentially explore this motif over the enzyme-labile ester.

4.6 Ligand deconstruction of (*S*)-1 by PrOF NMR

Although PrOF NMR SAR experiments provided insights for improving solubility and stability of the ***rac*-1** scaffold and the identification of the active enantiomer, **(*S*)-1**, none of the changes led to a significant increase in affinity. These results led us to explore a ligand deconstruction analysis of **(*S*)-1** to evaluate the contribution to binding of specific pharmacophores of the parent scaffold. PrOF NMR has mainly been used for hit identification, and only a preliminary ligand deconstruction study was reported from our group previously.[164] However, NMR is well-suited for deconstruction analysis due to its ability to quantify the affinities of weak binding fragments. We synthesized fragments **F1-F6** of the parent scaffold, which could be used to test the relative contributions of each portion of the molecule by PrOF NMR. In the original study which identified ***rac*-1** as a BPTF bromodomain ligand, apparent important pharmacophores for binders of BPTF were the aryl urea and the meta carbonyl group (ester or ketone). In light of this and the SAR results described above, we predicted a stronger interaction from this portion of the molecule when taken alone. Whereas fragment **F1** which contains the amino-pyrimidine portion of **(*S*)-1** did not bind BPTF, fragment **F4** containing the methyl ester displayed a K_d as determined by PrOF NMR of 1.3 mM with the highest ligand efficiency of 0.28. This can be compared with ***rac*-1** with a ligand efficiency of 0.22. We tested additional fragments, **F3**, **F5**, and **F6** with results summarized in Figure 4-11. In these cases, a similar ligand efficiency was maintained (0.22-0.23) of fragments tested. Extending fragment **F1** to **F2** to include the amino-pyrrolidine and the carbonyl group of the urea also bound with a K_d of 1.3 mM, albeit with a with a reduced ligand efficiency of 0.17. We note, that fragments can bind in multiple orientations if not anchored by a strong non-covalent interaction. Despite this possibility, these data support that the binding of **(*S*)-1** to the bromodomain of BPTF is the result of an additive effect of several weak contacts throughout the **(*S*)-1** scaffold.



Fragment	K_d	Ligand Efficiency
F1	N/A	N/A
F2	1300 μM	0.17 kcal/mol
F3	35 μM	0.23 kcal/mol
F4	1300 μM	0.28 kcal/mol
F5	690 μM	0.23 kcal/mol
F6	580 μM	0.22 kcal/mol

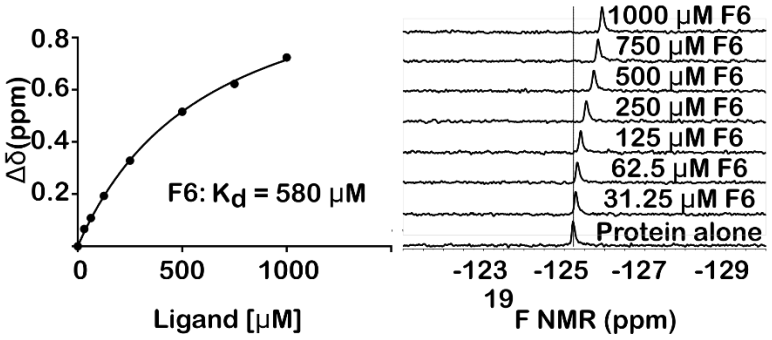


Figure 4-11: Ligand deconstruction study of **(*S*)**-1, with the exception of **F1**, all fragments bind 5FW BPTF with similar ligand efficiencies, indicating binding contacts are dispersed throughout the molecule. Using PrOF NMR, a binding isotherm can be generated to quantify weak binding ligands. **F6** is shown as a representative example. For reference, ***rac*-1** ligand efficiency = 0.22. NB = No observable binding.

4.7 Molecular dynamic simulations of (*S*)-1 with BPTF

We have not solved a co-crystal structure of **(*S*)-1**, to validate a binding mode to support our SAR and ligand deconstruction analysis. necessitating alternative methods to provide higher resolution information on binding. We therefore turned towards molecular

dynamics (MD) simulations of **(*S*)-1** binding to BPTF using a recently reported weighted ensemble approach[165] that has been used to investigate bromodomain-ligand interactions.[166] The inclusion of flexibility of the entire protein over our prior docking studies was anticipated to improve our predicted binding pose. Starting from multiple docked initial structures, a network model of ligand poses was constructed, and the weights of each state were used to predict an ensemble of probable poses (Figure 4-12A). These poses support that the methyl ester serves as an acetylated lysine mimic as it can hydrogen bond to the conserved N2881 of BPTF. The importance of the methyl ester in the acetyl lysine binding site is consistent with the higher ligand efficiency of **F4** and is consistent our experimental findings that alterations of the methyl ester such as substitutions around the aromatic ring obliterate binding, as well as the intolerance for large amide groups at this position. MD simulations show the attached aryl ring also interacts Phe2887 (via a stacking interaction, while the urea is involved in H bonding interactions with D2834. Finally, the *p*-fluoroaniline is exposed to solvent near W2824 in the WPF shelf, consistent with our findings that other groups can be installed at this position (e.g., **7**, **8**) and that replacement of the *p*-fluoroaniline with a chlorine maintains modest affinity (**F3**, K_d = 35 μM). In addition to helping explain our SAR, this docking pose will be used to guide future analog designs in the absence of a co-crystal structure.

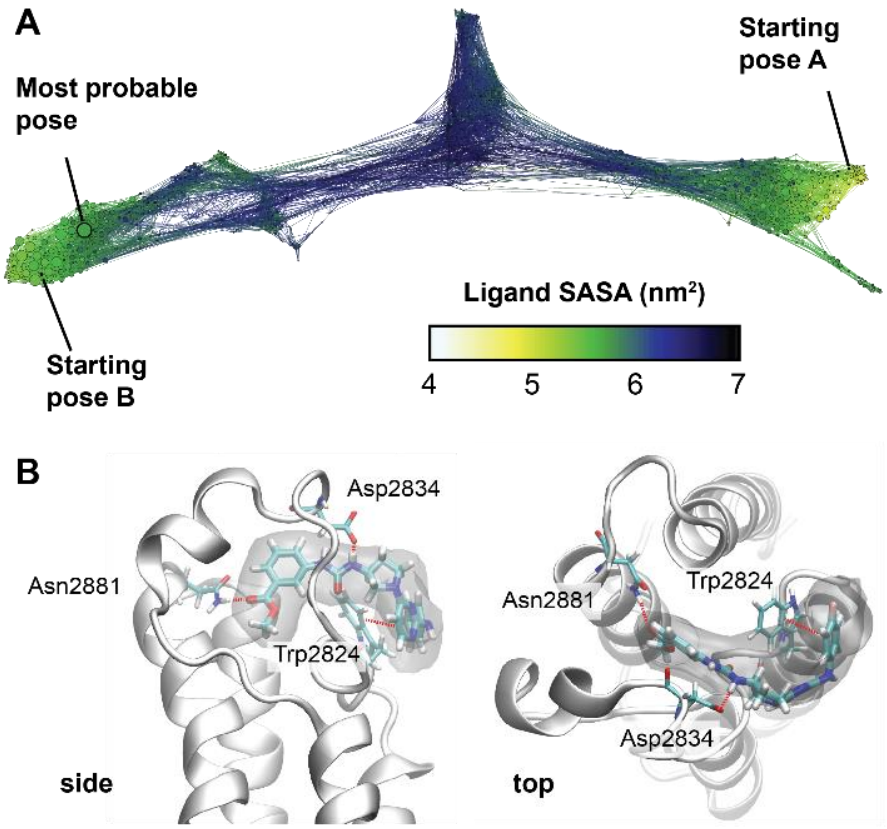


Figure 4-12: A) Conformation space network describing all of the binding poses found and their interconnectivity. Each node represents a ligand pose and nodes are colored according to the solvent accessible surface area of the ligand (blue/black = unbound, green/yellow = fully bound). The two starting poses, as well as the predicted most probable pose are labeled. B) Views of the most probable pose. A density isosurface was computed using 10 random ligand structures taken from most probable state (shown in transparent gray). Interacting residues are labeled.

4.8 Bromodomain selectivity analysis of (*S*)-1

As **(*S*)-1** remained our best compound, we next conducted a preliminary analysis of bromodomain selectivity for **(*S*)-1**. ***Rac*-1** was originally shown to be selective for BPTF over the N-terminal BRD4 bromodomain, BRD4(1), which also contains a WPF shelf and is a member of the bromodomain and extra-terminal domain (BET) family of proteins. In addition to 5FW-labeled BRD4(1), we tested three other WPF shelf-containing proteins: a second BET bromodomain, 5FW BRDT(1), PCAF,[67] which is on the same branch of the phylogenetic tree as BPTF , and a highly homologous bromodomain from *Plasmodium falciparum*, *Pf*GCN5[92] (Figure 4-13). Although *Pf*GCN5 is not a human bromodomain,

acetylated peptides have been shown to bind to both the *Pf*GCN5 bromodomain and BPTF with similar affinity. Additionally, PCAF inhibitors, also inhibit *Pf*GCN5 suggesting high similarity in binding sites of these structurally related bromodomains. In the PrOF NMR spectra, **(*S*)-1** did not perturb resonances for either BRD4(1) or BRDT(1). In the case of the PCAF bromodomain, although PCAF and BPTF are very close on the phylogenetic tree, only a small chemical shift was observed with slight broadening, without a significant dose dependence (see Section 4.12). These results indicate a moderate level of selectivity is maintained for BPTF over the structurally similar PCAF bromodomain. Finally, in the case of *Pf*GCN5, a significant broadening of the most downfield resonance was observed indicating an interaction similar to BPTF. We have tentatively assigned this resonance to the tryptophan in the WPF shelf based on prior binding studies with pan-inhibitor bromosporine.[167] Interestingly ***rac*-1** has shown antimalarial activity[168], consistent with bromodomain targeting, although target identification for ***rac*-1** has not been carried out against the parasite. When potency is increased, a more complete bromodomain profiling will be conducted.

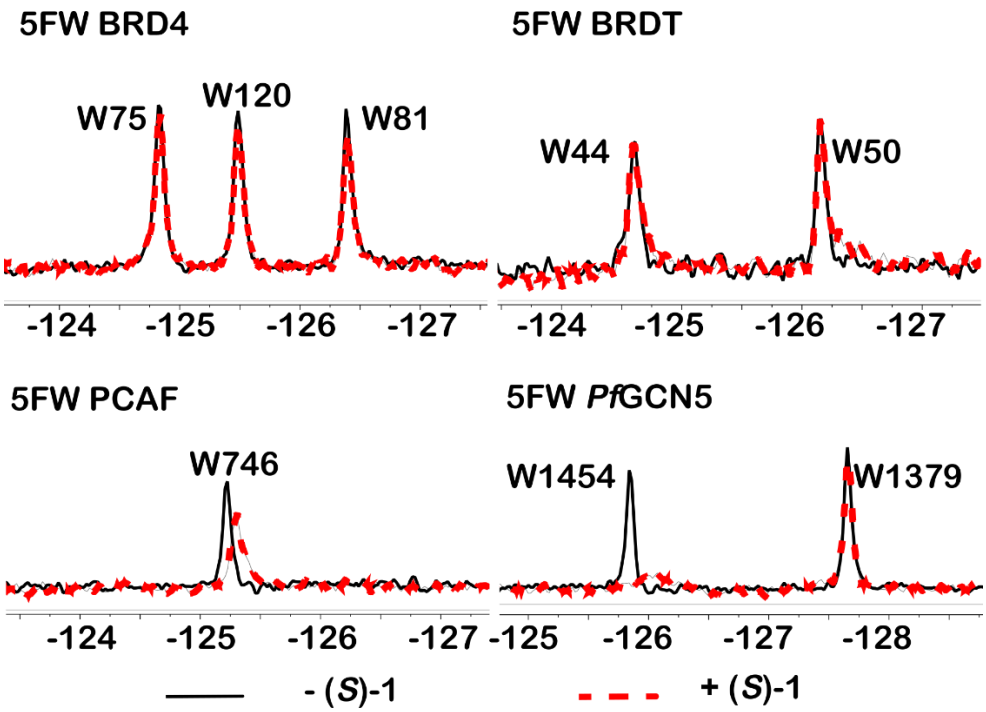


Figure 4-13: Selectivity of **(*S*)-1** toward four different bromodomains, BRD4(1), BRDT(1), PCAF, and *Pf*GCN5. All experiments were carried out with 50 μM protein in the absence (black spectrum) or presence of 100 μM **(*S*)-1** (red spectrum).

4.9 Kinase selectivity analysis of *rac*-1 stereoisomers

***Rac*-1** was originally discovered using the published kinase inhibitor set library. Therefore, we also tested for kinase activity of **(*S*)-1**. In a recent report evaluating inhibition against the kinome, ***rac*-1** showed 82% inhibition of two kinases, CDKL2 and TRKC, when tested at 1 μM.[169] Four additional kinases (CDK11, TRKB, HPK1, p38-δ) were inhibited between 51-68%. Based on this data, we utilized the KINOME*scan* service from DiscoverX to test **(*R*)-1** and **(*S*)-1** against CDKL2 and TRKC. **(*S*)-1** showed dissociation constants of 260 nM and 200 nM against CDKL2 and TRKC respectively, whereas **(*R*)-1** had dissociation constants of 1200 nM and 400 nM against the two kinases (Table 4-1). While there is a nearly five-fold difference in affinity toward CDKL2, there is only a two-fold difference toward TRKC, indicating that the stereocenter is not critical to the binding of these targets, in particular TRKC. These kinases need to be considered in future cell-based studies.

Table 4-1*: Binding affinities of **(S)-1** and anlogs against known kinase targets*

Compound Number	TRKC K_d	CDKL2 K_d
(*S*)-1	200 nM	260 nM
(*R*)-1	400 nM	1,200 nM
5	NB	NB
6	2,600 nM	19,000 nM

We looked to make changes to limit the kinase binding of the molecule, thereby reducing potential off-target effects in vivo. The amino-pyrimidine pharmacophore found in **(*S*)-1** is a known kinase-binding functionality which recognizes the hinge region of the ATP binding pocket in kinases. Two of the simplest changes that we envisioned were replacing the protic nitrogen with an oxygen (**5**), and the alkylation of the protic nitrogen (**6**). N-methylation of the aniline nitrogen yielded stoichiometric binding again in the PrOF NMR assay but a lack of activity was observed in AlphaScreen, potentially for the same solubility reason previously discussed. Upon exchanging the nitrogen for an oxygen,

binding was maintained, albeit weakened relative to **(*S*)-1**, as the ligand approached fast-intermediate exchange in the PrOF NMR assay. Encouragingly, after testing these two compounds against CDKL2 and TRKC we observed a drastic reduction in kinase binding in the N-methyl case (**6**), and a complete loss of binding with compound **5** (Table 4-1). The phenoxypyrimidine motif will be a useful design element for future ligands to remove kinase binding.

4.10 Cancer cell lines sensitivity to genetic altering of *BPTF* gene and *rac*-1 stereoisomers

The studies of ***rac*-1**, lead to the identification of the active enantiomer, **(*S*)-1**, as well as additional kinase off-targets. As a preliminary analysis of the cellular effects of **(*S*)-1** we aimed to test the sensitivity to both **(*S*)-1** and **(*R*)-1** treatment, to identify cell lines that were less susceptible to off target effects shared by **(*S*)-1** and **(*R*)-1,** and were sensitive to BPTF inhibition. We first tested whether cancer cell lines with different BPTF dependencies would grow differently when treated with either **(*S*)-1**, or **(*R*)-1** at 5 μM using an AlamarBlue cell viability assay. We tested lines, HepG2 (Liver hepatocellular carcinoma cell line), K562 (Chronic myelogenous leukemia cell line), MCF-7 cells (Hormone positive breast cancer cell lines) and A549 cells (Lung adenocarcinoma cell line). Our results demonstrated that HepG2 cells were not sensitive to treatment with either **(*S*)-1** or **(*R*)-1**, suggesting that neither BPTF inhibition (**(*S*)-1**) nor additional effects from small molecule composition (**(*R*)-1**) are essential for cellular growth. K562 cells were less viable when treated with **(*S*)-1** whereas **(*R*)-1** treatment did not perturb cell growth. This result suggests that BPTF bromodomain inhibition may play a significant role on controlling K562 cell growth. Alternatively, MCF-7 cells were sensitive to both **(*S*)-1** and **(*R*)-1** treatment, suggesting that for this cell line, off-target effects in addition to those of BPTF inhibition may be conferring sensitivity to small molecule treatment (Figure 4-14C). Conversely, HepG2 was not affected by either **(*S*)-1** or **(*R*)-1** demonstrating that **(*S*)-1** is not broadly toxic under these conditions.

To compare our small molecule inhibition results, we tested the effects of BPTF deletion on HepG2, K562, and MCF-7 cells which exhibit high to moderate BPTF mRNA levels. Deletion of the entire protein, removes additional functionality beyond the

bromodomain, such as scaffolding, but is useful to identify an overall protein-dependence to compare our cell viability data. To delete BPTF, we cloned single stranded guide RNA sequences (sgRNA) targeting two distinct domains of the BPTF gene, in addition to sgRNAs targeting the RNA polymerase RPA3 (depletion control) into an all-in-one CRISPR-Cas9-GFP lentiviral vector. The percentage of GFP+ cells, a mark for lentiviral infection, were monitored for several weeks using flow cytometry, as a proxy for the effects of BPTF deletion of growth of HepG2, K562 and MCF-7 cell lines. Our results demonstrated that HepG2 liver cancer cells were the least sensitive to BPTF depletion, given that the percentage of GFP+ cells carrying BPTF sgRNAs remained largely unchanged through the experiment time course (Figure 4-14B, top panel). These results are also consistent with the inactivity of **(*S*)-1** in this cell line. Conversely, the percentage of GFP+ cells gradually decreased over time in MCF-7 cells and K562 expressing BPTF sgRNAs, similarly to the kinetics of GFP expressing cells carrying RPA3 sgRNA depletion control, suggesting that in these cells deletion of BPTF impairs cell growth (Figure 4-14B, middle and bottom panels). K562 cells were the most sensitive to BPTF depletion, with the percentage of GFP+ cells dropping ~2-fold over time. Together, these results support an overall BPTF dependency, either via gene depletion (CRISPR-Cas9) or small molecule inhibition from **(*S*)-1**) in specific cell types. For future investigations, our preliminary analysis indicate K562 cells may be a useful model cell line to define epigenetic mechanisms associated with BPTF dependency and for testing with improved BPTF bromodomain inhibitors.

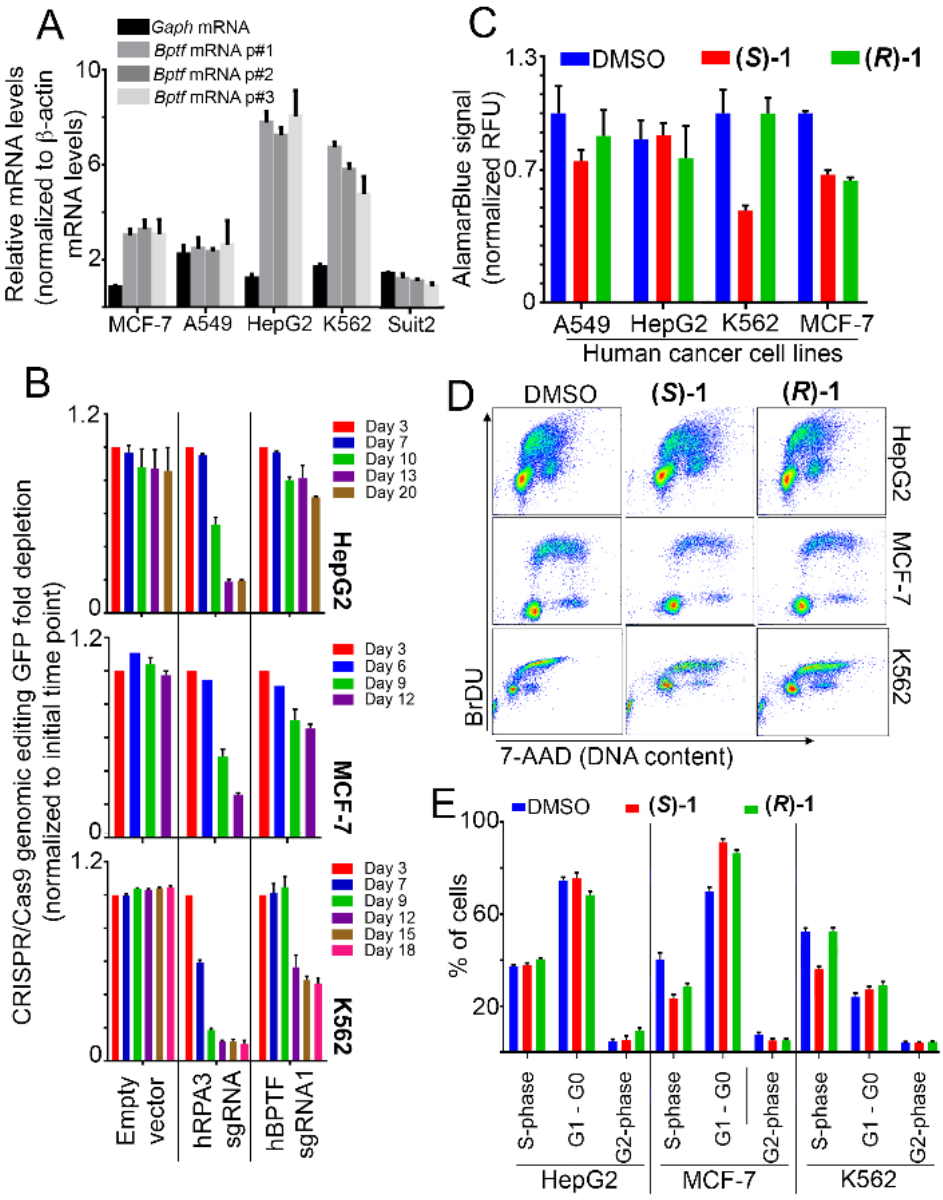


Figure 4-14: Sensitivity of cancer cell lines to **(*S*)-1** and **(*R*)-1**. A) BPTF mRNA quantification across a series of human cancer cell lines. B) CRISPR-Cas9 BPTF-targeting. HepG2, MCF-7 and K562 cell lines were infected with all in one CRISPR-Cas9-GFP lentiviral particles, expressing sgRNAs targeting RPA3 gene (depletion control) and the BPTF gene. Fold change was calculated by comparing final measure of GFP to initial infection GFP. n=2 technical replicates. C) Cell viability analysis in cancer cell lines treated with 5 μM of either **(*S*)-1** or **(*R*)-1** for 72 h, following incubation with AlamarBlue for 4 h. Error bars demonstrate standard deviation across biological replicates. D) Cell cycle analysis of **(*S*)-1** and **(*R*)-1** treated cell lines. HepG2, MCF-7 and K562 cells were cultured with 5 μM of either **(*S*)-1** or **(*R*)-1** for 72 h. E) Cell cycle information from panel D.

4.11 Conclusions

In conclusion, we have studied the SAR of ***rac*-1**, the first reported inhibitor of the bromodomain of BPTF over that of BRD4(1). We have identified that the (***S***)-enantiomer is the dominant stereoisomer for binding to the bromodomain. In addition, substitution of the amino-pyrimidine moiety led to the diminished binding of the substrate to CDKL2 and TRKC kinases. These results also showcase PrOF NMR as a useful protein-ligand characterization tool for quantifying weak binding ligands during ligand deconstruction, SAR analysis, and testing selectivity against related bromodomains. Finally, cell-based experiments including *BPTF* depletion, and chemical inhibition, support a BPTF dependence in K562 cells, a chronic myelogenous leukemia cell line. Richart and co-workers, identified cancer cell lines that could be sensitive to BPTF inhibtion with high c-Myc levels.[84] Chronic myelogenous leukemia cells, were identified to have some of the highest c-Myc levels, supporting our BPTF sensitivity studies. Together these studies provide a framework for the design of inhibitors for the BPTF bromodomain, an important epigenetic effector domain, lacking potent ligands.

4.12 Experimental

4.12.1 Materials:

All reagents used which were commercially available were used as received without further purification. Chloroacetyl chloride and N,N-diisopropylethylamine (Acros). *R*-3-(Boc-amino)pyrrolidine and *S*-3-(Boc-amino)pyrrolidine, iron powder (Alfa Aesar). Acetic acid, ammonium hydroxide (British Drug Houses). 2,4-Dichloropyrimidine (Chem-Impex). (*R*)-1-boc-3-aminopyrrolidine, (*S*)-1-boc-3-aminopyrrolidine, 4-Nitrophenylchloroformate (Combi-blocks). Acetic anhydride, diethyl ether, hexanes, isopropylamine, and sodium chloride (Fisher). Triethylamine (Mallinkrodt). Sodium bicarbonate (EMD Millipore). Ethanol (Pharmco-Aaper). 3-aminobenzoic acid, 4-fluoroaniline, N,N-dimethylaminopyridine, sodium hydroxide, dichloromethane, 1,2-dichloroethane, 1,4-dioxane, ethyl acetate, formaldehyde (37% solution in H_2O/MeOH), imidazole, magnesium sulfate, methanol, methyl 2-isocyanatobenzoate, methyl 3-isocyanatobenzoate, methyl 4-isocyanatobenzoate, methylamine (2.0 M in THF), N-methyl 4-fluoroaniline, sodium triacetoxyborohydride, tetrahydrofuran, thionyl chloride, trifluoroacetic acid, and tris base (Sigma-Aldrich). 3-Nitrobenzoyl chloride (TCI). Ethylamine (70% in H_2O) (VWR).

4.12.2 ^{1}H and ^{13}C NMR:

All NMR spectra were acquired on a Bruker 500 MHz spectrometer equipped with a prodigy TCI cryoprobe. Small molecule NMR spectra were taken in $CDCl_3$ (^{1}H ref. 7.26 ppm, ^{13}C ref. 77.0 ppm), MeOD (^{1}H ref. 3.31 ppm), or DMSO-*d*6 (^{1}H ref. 2.50 ppm, ^{13}C ref. 39.51 ppm).

4.12.3 Protein-Observed Fluorine (PrOF) NMR:

Fluorinated BPTF was produced and purified as described elsewhere.[163] All ^{19}F NMR spectra were acquired at 25 °C and 471 MHz. Protein samples with or without ligand were prepared in a buffer comprised of 50 mM Tris, 100 mM NaCl, pH = 7.4, with 40-50 μM protein, 52 mM trifluoroacetic acid, 5% D_2O, and 1% DMSO. Protein spectra were acquired with 500-1000 scans, acquisition time of 0.05 s, relaxation delay of 0.7 s, and a transmitter offset of -125 ppm with a 20-40 ppm spectral window. A second spectrum was

acquired for each sample to observe trifluoroacetate (^{19}F ref. -76.55 ppm) with 16 scans, acquisition time of 0.5 s, relaxation delay of 1 s, and a transmitter offset of -75 ppm with a 10 ppm spectral window. Each protein sample was referenced by applying the difference between trifluoroacetate and its reference chemical shift.

Binding affinity (K_d) was acquired by observing the change in chemical shift ($\Delta\delta$) compared with a no-ligand control spectrum. The values for $\Delta\delta$ were plotted and a curve was produced using GraphPad Prism software by fitting to the following equation:

$$Y = \Delta\delta_{max} \frac{(K_d+X+P)-\sqrt{(K_d+X+P)^2-(4PX)}}{2P} \quad (1)$$

X = ligand concentration (μM)

Y= experimental $\Delta\delta$ (ppm) as an absolute value

$\Delta\delta_{max}$ = maximum chemical shift (ppm)

K_d = dissociation constant (μM)

P = protein concentration (μM)

4.12.4 Mass Spectrometry:

HRMS Conditions: For high-resolution ESI-MS analysis of small molecules, either a Bruker BioTOF II ESI/TOS-MS direct infusion instrument was used in conjunction with polyethylene glycol standards for mass calibration, or a Thermo-Fisher Orbitrap-Velos equipped with an autosampler.

UPLC/MS Conditions: For UPLC/MS analysis of proteins, a Waters Acquity UPLC coupled to a Waters Synapt G2 HDMS quadrupole orthogonal acceleration time of flight mass spectrometer was used (Waters Corp., Milford, MA USA). A Waters Acquity UPLC Protein BEH C_4 2.1 mm x 100 mm column (1.7 μm diameter particles) at 35°C was used for the following 15 min linear gradient separation at a flow rate of 0.400 mL/min using A: water containing 0.1% formic acid and B: acetonitrile containing 0.1% formic acid: 3% B, 0 min to 3 min; 3% B to 97% B, 3 min to 9 min; 97% B, 9 min to 11 min; 97% B to 3% B, 11 min to 13 min; 3% B 13 min to 15 min. Mass spectra were collected in profile mode over the range *m/z* 300-2500 every 0.1s during the chromatographic separation. MS

parameters in positive electrospray ionization mode were as follows: capillary, 0.3 kV; sampling cone, 35.0 V; extraction cone, 4.0 V; desolvation gas flow, 800 L/h; source temperature, 100°C; desolvation temperature, 350°C; cone gas flow, 20 L/h; trap CE, off. Lockspray (on-the-fly mass calibration) configuration consisted of infusion of a 5 μg/mL solution of leucine-enkephalin and acquisition of one mass spectrum (0.2s scan, *m/z* 50-1200) every 10s. Three lockspray *m/z* measurements of protonated (positive ionization mode) leucine-enkephalin were averaged and used to apply a mass correction to measured *m/z* values during the course of the analysis.

Table 4-2*: MS of fully fluorinated proteins*

Protein	**Calculated m/z (Da)**	**Observed m/z (Da)**	**% Fully-fluorinated**
5FW BPTF (BD)	14455	14455.9	100%
5FW BRD4 (BD1)	15137	15137.0	93%
5FW BRDT (BD1)	14184	14183.2	97%
5FW PCAF (BD)	17182	17181.5	86%
5FW *Pf* GCN5 (BD)	12640	12639.3	96%

4.12.5 Cell lines:

Cell lines were purchased from ATCC and further certified at CSHL tissue culture shared resources. For this specific study we utilized MCF-7 cells (Hormone positive breast cancer cell lines) A549 cells (Lung adenocarcinoma cell line), HepG2 (Liver hepatocellular carcinoma cell line), K562 (Chronic myelogenous leukemia cell line), Suit2 (Metastatic liver tumor of pancreatic carcinoma), and 293ft-NEO (virus production). Cells were grown according to ATCC recommendations.

4.12.6 mRNA quantification:

Cells were resuspended in Trizol (Thermo Fisher) for RNA purification. cDNA was synthesized with SuperScript III (Thermo Fisher) according to manufacturer's instructions. cDNA was utilized on qPCR analysis for the quantification of *BPTF* mRNA levels (3 primer sets: 1F: ACC CAG AGA ATT TGC ATT GG-3',1R: ATT TTA CCC ATG TCG CTT GC- 3'; 2F: AGC AGA AGC CGA CAG TGA TT-3', 2R: GGC CTT GCT TAA CCC ATG TA- 3'; 3F: TTG GCA TCT TGC AAA GTG AG- 3', 3R:TTA TGG GCC TGT AAG GAA CG-3'), *GAPDH* mRNA (F: 5'CCA CAT CGC TCA GAC ACC AT 3', R: 5' CCA GGC GCC CAA TAC G 3'), and *β-ACTIN* mRNA (normalization, 5' AGA GCT ACG AGC TGC CTG AC 3', 5' AGC ACT GTG TTG GCG TAC AG 3'). qPCR reactions were performed on a QuantStudio 6 (Thermo Fisher).

4.12.7 Cell Viability assay:

Cell lines were treated for 72 hours with DMSO control, **(*S*)-1** (5 μM) or **(*R*)-1** (5 μM). Following treatment, cells were incubated for 3 hours with AlamarBlue (Thermo Fisher), a resazurin dye. The AlamarBlue is actively reduced to another compound, resorufin, which is fluorescent in 580 nm excitation and 590 nm emission. Fluorescence was measured using a plate reader. All data was normalized to the control which were the DMSO treated wells.

4.12.8 CRISPR/Cas-9 genomic editing:

Short-guide RNA (sgRNA) sequences were predicted utilizing previously published algorithm.[170] sgRNAs targeting hRPA3 and hBPTF were cloned into the all-in-one CRISPR-CAS9-GFP plasmid. All in one plasmids and virus packaging/envelope coding plasmids (VSVG and Pax2) were transfected into 293 ft-Neo cells using PEI (Sigma). Lentivirus supernatant was collected every 24 hours, for a total of 72 hours, and further utilized to infect HepG2, MCF-7 and K562 cells. GFP measurements (a readout for virus infection) were taken using MACSQuant (Miltenyi Biotec) every 3 days, or when culture plates were ~70 confluent. FlowJo (Tree Star) was used to analyze flow cytometry data.

Genomic DNA from sgRNA infected cells was utilized on qPCR for quantification of editing efficiency, utilizing previously published methods.[171] qPCR reactions were carried out on a QuantStudio 6 (Thermo Fisher) utilizing the following primer sequences for BPTF genomic locus (gBPTF primer #1 - 1F 5' - CGT GAA GAA GAC ACT TCC AAT AC - 3', 1R 5' -CCA CGT CAT CCC ATC TAT GAA- 3'; gBPTF primer #2 - 2F 5' -AGC GTT AAT TCC ACA CTG TAT TTC- 3', 2R 5' -AAG AAC GTG ATG GTA CTC CTT ATC- 3'; gBPTF primer #3 - 3F 5' -CTC GAG AGG AAT TGA TGT CTG AA- 3', 3R 5' -ATG TCT CAC AGC AAA GCA AAT C- 3'; gBPTF primer #4 - 4F 5' -TGT GAG ACA TGT TCA GCA GTA T- 3', 4R 5' -CCT TGT GTG CTA CAC AGA CT- 3')

4.12.9 Cell cycle analysis:

Cell lines were treated for 72 hours with DMSO control, **(*S*)-1** (5 μM) or **(*R*)-1** (5 μM), following BrdU labeling (BD Biosciences) according to manufacturer's instructions. Cells were analyzed using MACSQuant (Miltenyi Biotec), and FlowJo (Tree Star) was used to analysis flow cytometry data.

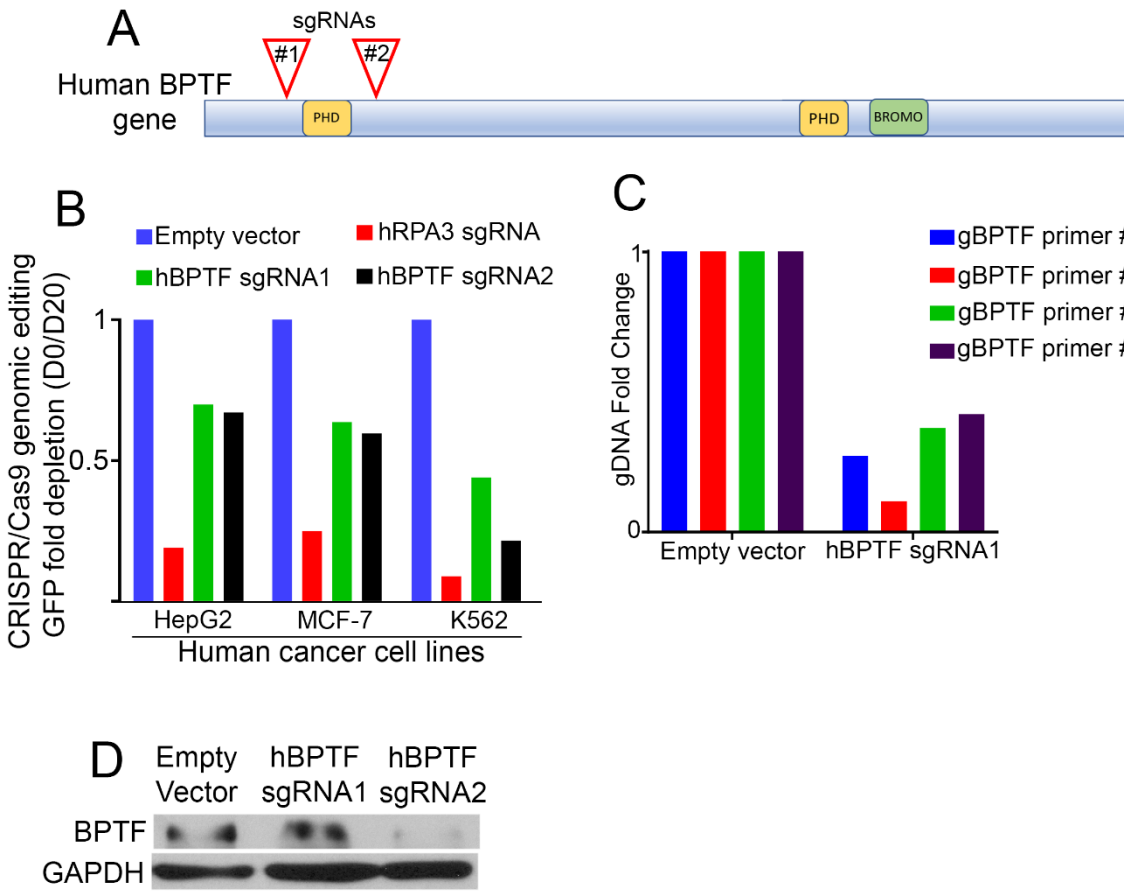


Figure 4-15: **Characterizing the effects of BPTF inhibition and S-AU-1 treatment.** (**A**) BPTF genomic scheme showing location of sgRNAs. (**B**) Summary of CRISPR-Cas9 genomic editing in HepG2, MCF-7 and K562 cells. Cells were infected with all in one CRISPR-Cas9-GFP lentiviral particles, expressing sgRNAs targeting RPA3 gene (depletion control) and two differed domains of BPTF gene. Fold change was calculated by comparing final measure of GFP to initial infection GFP. n=2 technical replicate. (**C**) CRISPR-Cas9 genomic editing efficiency quantification. Genomic DNA from HepG2 cells infected with control lentivirus (Empty vector) or with lentiviral particles expressing BPTF sgRNA1 was purified and utilized for qPCR analysis with primer sets surrounding the BPTF genomic region targeted by sgRNAs. (**D**) Western blot showing BPTF protein levels and GAPDH protein levels (loading control) in HepG2 cells infected with control lentivirus (Empty vector) or lentiviral expressing sgRNAs against BPTF.

4.12.10 Pharmacokinetic (PK) studies.

To assess the PK of AU1, female BALBc mice were injected intraperitoneal with 5 mg/kg of AU1. Blood samples were collected at 5, 15, 30 min and 1, 5, 12, 24 h after i.p. administration. Three mice were used per time point. For each time point, the mice were rendered unconscious by CO_2, followed by surgically thoracotomy. The pooled blood (~1 mL) from the chest cavity was collected into heparinized tubes. The plasma fraction was separated by centrifugation (12,000 rpm, 10 min, 4°C) and stored at -20°C until LC/MS/MS analysis was carried out by the company Ianalytical using standard approaches.

4.12.11 Computational modeling

Initial docked structures

Autodock Vina[172] was used to generate starting poses of **(*S*)-1** bound to BPTF. The structure of BPTF from PDB ID 2RIV[173], where the binding pocket is occupied only by a molecule of isopropyl alcohol, which is removed prior to docking. The Y17E mutation was inconsequential as it is far from the binding pocket, and it is in a segment that was not included in the molecular dynamics simulation. For docking we use a cubic grid with 80 points along each axis, with a 0.375 Å spacing, centered at the position of the C3 atom of the isopropyl alcohol. Crystallographic water molecules were not included during docking. This procedure was chosen to follow prior retrospective predictions of bromodomain-ligand structure as described previously.[174] As the ligand is roughly linear in shape, we chose as initial structures two poses of opposite orientation: one with the *p*-fluoroaniline bound to the recognition pocket (pose A) and one with the methyl ester bound to the recognition pocket (pose B). In both cases we chose the pose with the lowest predicted binding free energy.

Molecular dynamics

Only residues 2797-2907 were used for the dynamics. The two systems (pose A and pose B) were both solvated with a 12 Å cutoff, and neutralized with three sodium ions. Parameters from CGENFF were used for the AU1 ligand.[175,176], The systems were prepared for dynamics using energy minimization with harmonic restraints on the protein

and ligand atoms (500 steps of steepest descent followed by 500 steps of the adopted basis Newton Raphson method) and this procedure was then repeated with the restraints removed. The CHARMM program with an OpenMM interface for GPU dynamics was used to run dynamics, with a 2 fs timestep. The SHAKE algorithm, with a tolerance of 10^{-8} is used to constrain covalent bonds to hydrogen. Nonbonded interactions are computed with the particle mesh Ewald method with a Gaussian width of 0.32 and 96 grid points along the x, y and z dimensions. An 8.5 Å cutoff is used for Lennard-Jones interactions.

The energy-minimized systems were then heated gradually from 50 to 300 K by increments of 25 K, using 5000 dynamics steps at each temperature. Temperature was controlled using a Langevin heatbath with a friction coefficient of 1.0 ps^{-1}. Following heating, equilibration runs were performed at 300 K using constant pressure dynamics, implemented with a Monte Carlo barostat with a reference pressure of 1.0 atm and volume moves attempted every 50 timesteps. The resulting systems were used as starting points for WExplore simulations. All subsequent dynamics were run at 300 K at constant pressure.

WExplore sampling

The WExplore[165] implementation again followed previous work on predicting bromodomain-ligand poses,[174] and more information on the WExplore algorithm and its application to protein-ligand interactions can be found elsewhere.[174,165,177,178,179] Briefly, the WExplore algorithm enhances the sampling of ligand release pathways by running an ensemble of weighted trajectories forward simultaneously and periodically (here, every 20 ps) altering, or "resampling", the trajectory ensemble using cloning and merging operations. The weights (initialized to all be equal) govern the weight with which the trajectory at that time contributes to averages of observables. During cloning, a single trajectory is split in two, and its weight is evenly divided among the clones. When two trajectories are merged into one, their weight is combined, and a single trajectory is chosen for continuation with a probability that is proportional to its weight. Cloning and merging steps are a common feature of all weighted ensemble applications.[165]

WExplore is a variant of weighted ensemble where the total number of walkers is fixed (each cloning event is coupled to a merging event), and cloning and merging decisions are made using the occupancy of a set of hierarchical regions that are defined using Voronoi polyhedra. To determine which region a given trajectory is in, we measure the distances between that trajectory and the set of "images" that define the Voronoi polyhedra. The trajectory is then assigned to the region that has the closest image. Here we use a four-level region hierarchy, with critical distances of 2.5, 3.5, 5.0 and 10 Å (again following previous work[174,178,179]). To improve sampling efficiency we employ a minimum weight (10^{-12}) and a maximum weight (0.1) that a walker can achieve, enforced by disallowing cloning or merging operations that would violate these rules.

For both pose A and pose B we run four WExplore simulations, with 48 trajectories each. Each simulation was run for 830 cycles of dynamics and resampling. In total, 3.19 μs of dynamics was run for each starting pose (6.37 μs combined).

Network analysis and pose determination

To analyze the structures from WExplore sampling and predict an ensemble of final poses, we performed clustering and Markov state model analysis. Clustering was performed using a three-step process. In the first step, each frame from the MD simulation was "featurized" using a large set of protein-ligand distances. This distance set was the set of all possible distances between two subsets of atoms: 1) the ligand atoms N2, C10, C14, C21, N4, H1, N1, C18, O2, C23, O3, N6, H10, and F1, which were chosen to evenly cover the ligand, and 2) all protein CA atoms that were within 8 A of the ligand center of mass. There were 938 distances in the feature set. In the second step, the feature space was projected onto five dimensions using time-lagged independent coordinate analysis (or, tICA).[180] This was performed using a tICA lag time of 0.2 ns. In the third step, a k-means clustering was performed using the tICA projections, dividing the structures into 1200 states.

Markov state models were then constructed for a range of lag times (1, 2, 5, 10, 20, 50 sampling periods), and various subsets of the data (p = 0, 10, 20, 30, 40, 50, 60) where p is the percent of initial data discarded. For each Markov model the weights of each node were computed using the top eigenvector of the transition matrix. The highest such node is shown in Figure 4-16A. There is very little variation as a function of lag time, and three distinct states are predicted in the entire map (P1, P2, and P3). Figure 4-16B shows that these states are in the same community in the network model, which demonstrates that these predictions are relatively robust.

Each node in the network visualizations in Figure 4-12 and Figure 4-13B shows one of the clusters. The weights of the edges between the nodes were computed using the conditional transition probabilities between the clusters as follows:

$$e_{ij} = 100\frac{1}{2}\left(t_{ij} + t_{ji}\right) \quad (2)$$

where t_{ij} is an element of the transition probability matrix that represents the conditional probability of transitioning between states i and j. These weights are used in the Force Atlas algorithm in the Gephi program[181] to construct a minimized network layout.

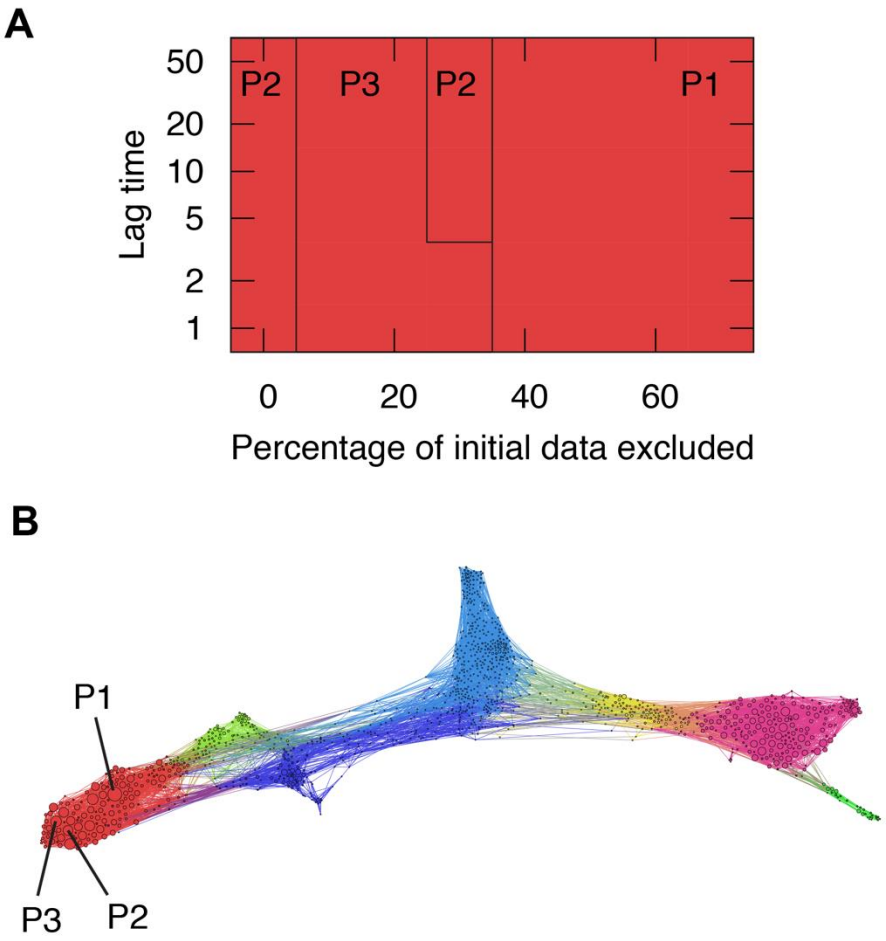
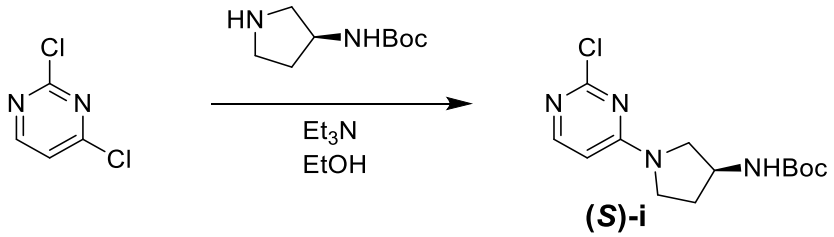


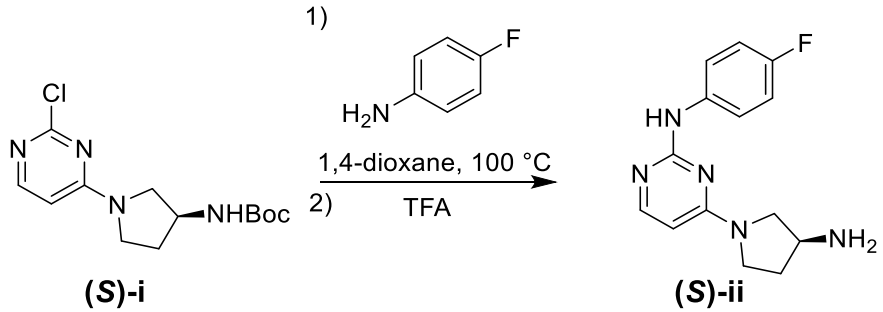
Figure 4-16: A. The highest weighted state is shown as a function of lag time and exclusion of initial data. Each square in the map is colored by the community that the highest weighted state belongs to for that set of parameters, which here is constant across the entire map. B. The network of poses is colored according to communities that were determined by modularity optimization in Gephi. The three highest weighted states are labeled and we observe that they are in the same community.

4.12.12 Small molecule synthesis



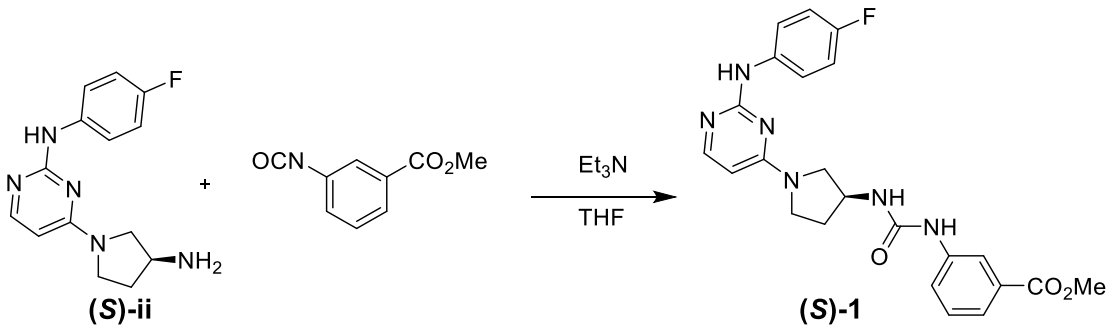
(*S*)-i

(*S*)-i was synthesized as reported previously.[94] ^{1}H NMR, as a mixture of rotational isomers (500 MHz, DMSO-d_6) δ 8.01 (dd, *J* = 11.0, 6.0 Hz, 1H), 7.26 – 7.20 (m, 1H), 6.49 (d, *J* = 6.0 Hz, 1H), 4.11 (dq, *J* = 43.4, 6.0 Hz, 1H), 3.59 (ddd, *J* = 25.8, 11.6, 6.4 Hz, 2H), 3.47 (dq, *J* = 13.8, 7.4, 7.0 Hz, 1H), 3.36 (ddd, *J* = 26.0, 10.1, 5.8 Hz, 1H), 3.17 (dd, *J* = 11.0, 4.5 Hz, 1H), 2.11 (ddq, *J* = 39.9, 13.7, 7.1 Hz, 1H), 1.88 (ddq, *J* = 25.6, 12.8, 6.2 Hz, 1H), 1.39 (s, 9H). ^{13}C NMR, as a mixture of rotational isomers (126 MHz, DMSO-d_6) δ 160.75, 160.63, 159.39, 156.27, 155.20, 103.16, 102.95, 77.96, 51.93, 51.83, 49.80, 49.20, 44.69, 30.43, 29.90, 28.19. HRMS (ESI-TOF) calculated for $C_{13}H_{19}ClN_4O_2Na^+$ $[M+Na]^+$: 321.1094, observed 321.1068.



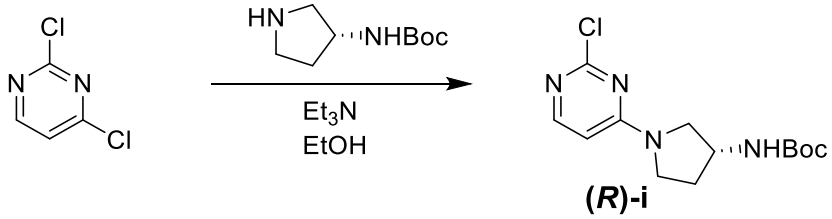
(*S*)-i **(*S*)-ii**

(*S*)-ii was synthesized as reported previously.[94] ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.01 (s, 1H), 7.88 (d, *J* = 5.9 Hz, 1H), 7.80 (dd, *J* = 8.9, 5.1 Hz, 2H), 7.05 (t, *J* = 8.8 Hz, 2H), 5.94 – 5.83 (m, 1H), 3.71 – 2.87 (m, 5H), 2.08 – 1.95 (m, 1H), 1.72 (m, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 160.06, 158.36 (d, *J* = 244.9 Hz), 137.87 (d, *J* = 2.2 Hz), 119.65 (d, *J* = 7.3 Hz), 114.64 (d, *J* = 21.8 Hz), 95.48, 54.46, 50.91, 50.10, 44.59, 33.85. HRMS (ESI-TOF) calculated for $C_{14}H_{16}FN_5H^+$ $[M+H]^+$: 274.1423, observed 274.1504.



(***S***)-ii (***S***)-1

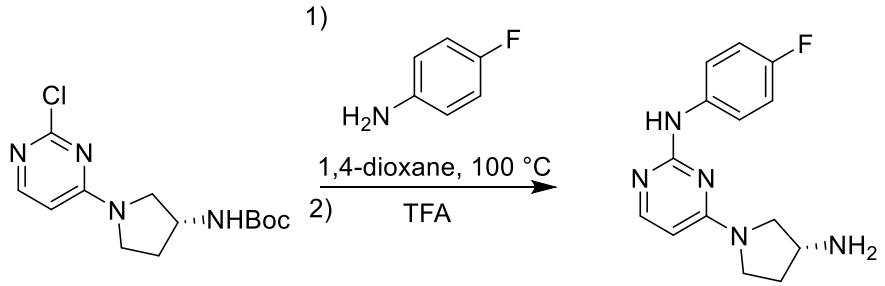
(*S*)-1 was synthesized as reported previously.[94] ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.06 (s, 1H), 8.59 (s, 1H), 8.14 (s, 1H), 7.93 (d, *J* = 5.9 Hz, 1H), 7.79 (dd, *J* = 8.6, 5.0 Hz, 2H), 7.55 (dd, *J* = 8.1, 2.1 Hz, 1H), 7.52 – 7.47 (m, 1H), 7.37 (t, *J* = 7.9 Hz, 1H), 7.06 (t, *J* = 8.7 Hz, 2H), 6.60 (d, *J* = 6.7 Hz, 1H), 5.98 (d, *J* = 5.8 Hz, 1H), 4.33 (m, 1H), 3.84 (s, 3H), 3.75-3.18 (m, 4H), 2.21 (m, 1H), 1.96 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.26, 160.19, 158.40 (d, *J* = 237.5 Hz), 155.64, 155.58, 154.75, 140.64, 137.77, 130.05, 129.06, 122.20, 121.87, 119.72 (d, *J* = 7.3 Hz), 118.09, 114.69 (d, *J* = 21.9 Hz), 95.65, 52.09, 51.92, 44.27. HRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_3H^+$ $[M+H]^+$: 451.1849, observed 451.1860.



(***R***)-i

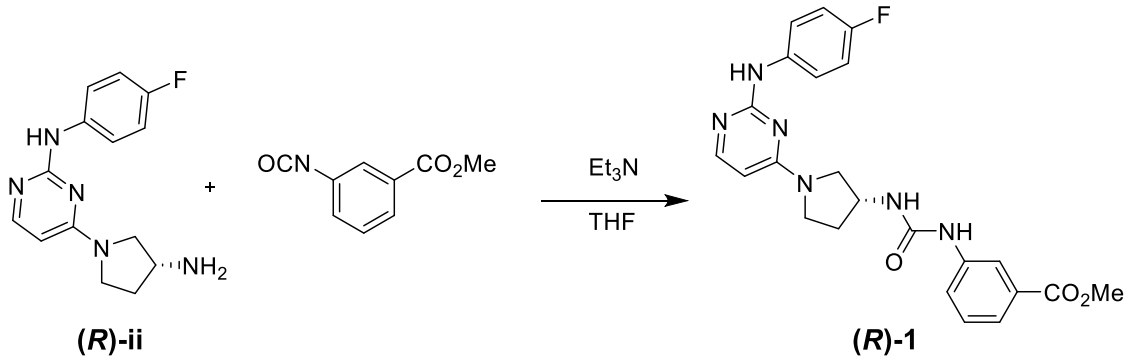
2,4-Dichloropyrimidine (292 mg, 1.96 mmol, 1 eq) was dissolved in 5 mL of EtOH. Triethylamine (300 μL, 1.1 eq) and (*R*)-3-(Boc-amino)pyrrolidine (400 mg, 2.15 mmol, 1.1 eq) were added and the reaction was stirred at ambient temperature for 4 h. The solvent was removed and the crude material was dissolved in EtOAc and washed 3 X with H_2O followed by brine and dried over $MgSO_4$. The crude material was filtered and concentrated followed by purification using a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **(*R*)-i** (105 mg, 18%). ^{1}H NMR as a mixture of rotational isomers (500 MHz, DMSO-d_6) δ 8.02 (d, *J* = 6.1 Hz, 0.5H), 8.00 (d, *J* = 6.2 Hz, 0.5H), 7.26 (d, *J* = 7.4 Hz, 0.5H), 7.23 (d, *J* = 7.3 Hz, 0.5H), 6.49 (d, *J* = 6.0 Hz, 1H), 4.15 (m, 0.5H), 4.05 (m, 0.5H),

3.61 (dd, *J* = 11.9, 6.1 Hz, 0.5H), 3.56 (dd, *J* = 11.1, 6.7 Hz, 0.5H), 3.47 (dq, *J* = 13.7, 7.0 Hz, 1H), 3.42 – 3.32 (m, 1H), 3.17 (d, *J* = 4.7 Hz, 0.5H), 3.15 (d, *J* = 4.7 Hz, 0.5H), 2.15 (dq, *J* = 13.8, 7.1 Hz, 0.5H), 2.06 (dt, *J* = 13.6, 6.7 Hz, 0.5H), 1.94 – 1.88 (m, 0.5H), 1.84 (dt, *J* = 12.8, 6.6 Hz, 0.5H), 1.39 (s, 9H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 160.65, 160.54, 159.29, 156.17, 155.05, 103.07, 102.85, 77.86, 51.83, 51.73, 49.70, 49.10, 44.59, 30.33, 29.80, 28.09. HRMS (ESI-TOF) calculated for $C_{13}H_{19}ClN_4O_2Na^+$ $[M+Na]^+$: 321.1094, observed 321.1141



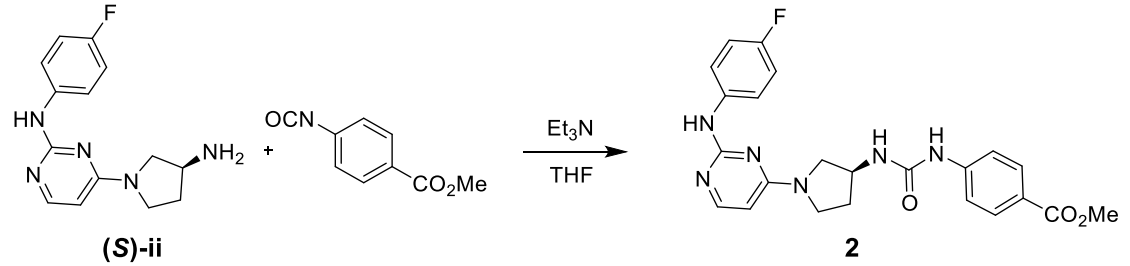
(*R*)-i **(*R*)-ii**

(*R*)-i (105 mg, 0.352 mmol) was dissolved in 3 mL of 1,4-dioxane followed by addition of 4-fluoroaniline (50 μL, 0.528 mmol) and heated to reflux for 18 h. The solvent was removed, and the crude material was redissolved in TFA and stirred at ambient temperature for 2 h. The TFA was removed under a stream of nitrogen, and the crude material was dissolved in H_2O and DCM. The aqueous layer was washed 3 X with DCM and separated, the pH was adjusted to ~12 using 1M NaOH and extracted into EtOAc. The organic layers were combined, washed with brine, and dried over $MgSO_4$. Following filtration, the solvent was removed under reduced pressure to afford **(*R*)-ii** as a white solid (60 mg, 62%, 2 steps). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.01 (s, 1H), 7.88 (d, *J* = 5.8 Hz, 1H), 7.79 (dd, *J* = 8.9, 5.0 Hz, 2H), 7.05 (t, *J* = 8.9 Hz, 2H), 5.89 (d, *J* = 5.8 Hz, 1H), 3.75 – 2.92 (m, 6H), 2.03 (s, 2H), 1.71 (s, 1H), 1.28 – 1.12 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 159.98, 158.26 (d, *J* = 244.5 Hz), 155.41, 155.29, 137.77, 137.75, 119.54 (d, *J* = 7.3 Hz), 114.54 (d, *J* = 21.8 Hz), 95.39, 54.25, 44.47. HRMS (ESI-TOF) calculated for $C_{14}H_{16}FN_5H^+$ $[M+H]^+$: 274.1423, observed 274.1429.



(*R*)-ii **(*R*)-1**

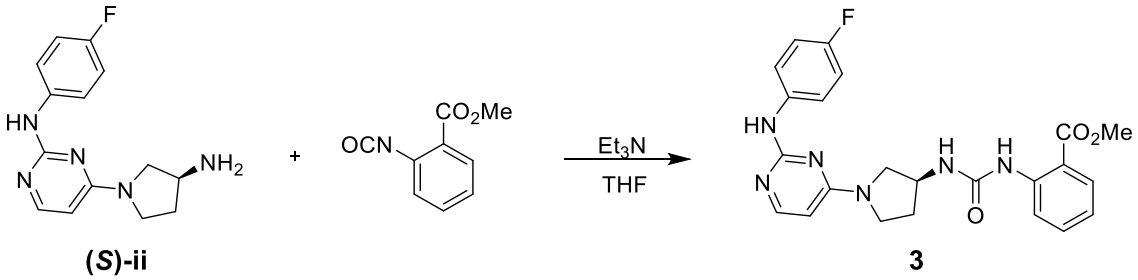
(*R*)-ii (60 mg, 0.22 mmol) was dissolved in 2 mL of THF, followed by addition of triethylamine (46 μL, 0.33 mmol) and methyl 3-isocyanatobenzoate (58 mg, 0.33 mmol). The reaction proceeded at ambient temperature for 2 h. The crude material was concentrated followed by purification using a Combiflash Rf system (Hexanes/EtOAc, 0-100% EtOAc) to isolate **(*R*)-1** (70 mg, 0.16 mmol, 71%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.06 (s, 1H), 8.59 (s, 1H), 8.14 (t, J = 1.9 Hz, 1H), 7.93 (d, J = 6.0 Hz, 1H), 7.85 – 7.76 (m, 1H), 7.54 (dt, J = 8.1, 1.6 Hz, 1H), 7.52 – 7.48 (m, 1H), 7.37 (t, J = 7.9 Hz, 1H), 7.06 (t, J = 8.9 Hz, 2H), 6.60 (d, J = 6.8 Hz, 1H), 5.98 (d, J = 5.9 Hz, 1H), 4.33 (br. m, 1H), 3.84 (s, 3H), 3.71 – 3.39 (m, 4H), 2.21 (br. m, 1H), 1.93 (d, J = 25.9 Hz, 1H), 1.29 – 1.13 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.26, 160.19, 158.40 (d, J = 237.9 Hz), 155.65, 154.75, 140.64, 137.76, 130.05, 129.06, 122.20, 121.87, 119.72 (d, J = 7.3 Hz), 118.09, 114.69 (d, J = 21.9 Hz), 95.65, 52.09, 51.92, 44.27. HRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_3H^+$ $[M+H]^+$: 451.1849, observed 451.1883.



(*S*)-ii **2**

(*S*)-ii (40 mg, 0.15 mmol) was dissolved in 2 mL of THF and 30 μL of triethylamine (0.23 mmol). Methyl-4-isocyanatobenzoate (38.9 mg, 0.220 mmol) was added and the mixture was stirred at room temperature for 3 h. Upon reaction, the solution turned white and

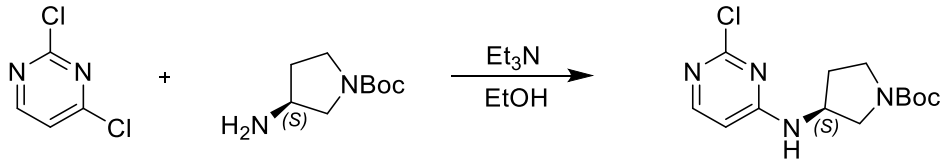
heterogeneous. Aliquots were taken and solubilized into methanol to monitor product formation via TLC (100% EtOAc, R_f = 0.2). Following completion of the reaction, the crude material was concentrated and purified via flash column chromatography (2:1 hexanes:EtOAc, to 100% EtOAc) was used to isolate **2** as a yellow solid (40.0 mg, 0.089 mmol, 57%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.06 (s, 1H), 8.76 (s, 1H), 7.93 (d, *J* = 5.9 Hz, 1H), 7.84 (d, *J* = 8.6 Hz, 2H), 7.79 (dd, *J* = 8.8, 5.0 Hz, 2H), 7.51 (d, *J* = 8.7 Hz, 2H), 7.06 (t, *J* = 8.9 Hz, 2H), 6.72 (d, *J* = 6.8 Hz, 1H), 5.98 (d, *J* = 5.9 Hz, 1H), 4.34 (s, 1H), 3.80 (s, 3H), 3.74 – 3.39 (m, 5H), 3.24 (s, 1H), 2.21 (s, 1H), 1.96 (s, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 165.92, 160.18, 158.40 (d, *J* = 237.3 Hz), 155.66, 155.57, 154.39, 144.86, 137.76 (d, *J* = 2.2 Hz), 130.35, 121.81, 119.72 (d, *J* = 7.3 Hz), 116.75, 114.69 (d, *J* = 21.8 Hz), 95.64, 51.88, 51.69, 44.23. HRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_3H^+$ $[M+H]^+$: 451.1849, observed 451.1884.



(*S*)-ii **3**

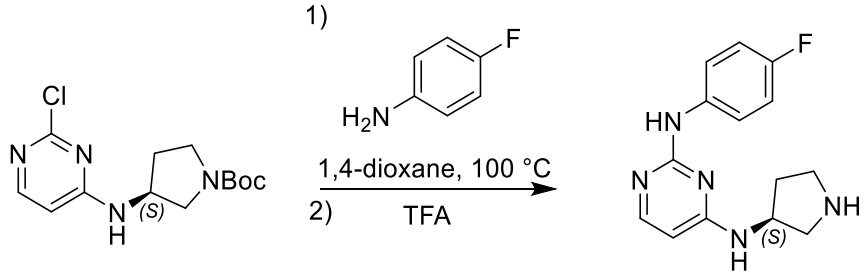
(*S*)-ii (30 mg, 0.11 mmol) was dissolved in 2 mL of THF and 25 μL of triethylamine (0.17 mmol). Methyl-2-isocyanatobenzoate (30 mg, 0.17 mmol) was added and the mixture was stirred at room temperature for 2 h. Following completion of the reaction, the crude material was concentrated and purified via flash column chromatography (9:1 EtOAc/10% NH_4OH in methanol) was used to isolate **3** as a white solid (20.0 mg, 0.044 mmol, 40%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 10.37 (s, 1H), 8.49 (dd, *J* = 8.6, 1.2 Hz, 1H), 7.98 (dd, *J* = 8.1, 1.7 Hz, 1H), 7.84 (d, *J* = 6.1 Hz, 1H), 7.59 – 7.55 (m, 2H), 7.51 (ddd, *J* = 8.7, 7.2, 1.7 Hz, 1H), 6.98 (m, 3H), 5.78 (d, *J* = 6.1 Hz, 1H), 5.61 (s, 1H), 4.57 (q, *J* = 5.4 Hz, 1H), 3.88 (s, 3H), 3.82 – 3.28 (m, 5H), 2.53 (s, 3H), 2.32 (p, *J* = 6.9, 6.5 Hz, 1H), 2.10 (s, 1H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 169.14, 160.46, 158.35, 158.19 (d, *J* = 241.1 Hz), 154.48, 142.88, 135.86, 134.64, 130.78, 120.97, 120.72 (d, *J* = 7.7 Hz), 119.57, 115.26 (d,

J = 22.3 Hz), 113.89, 95.76, 52.24, 44.60. HRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_3H^+$ $[M+H]^+$: 451.1849, observed 451.1884.



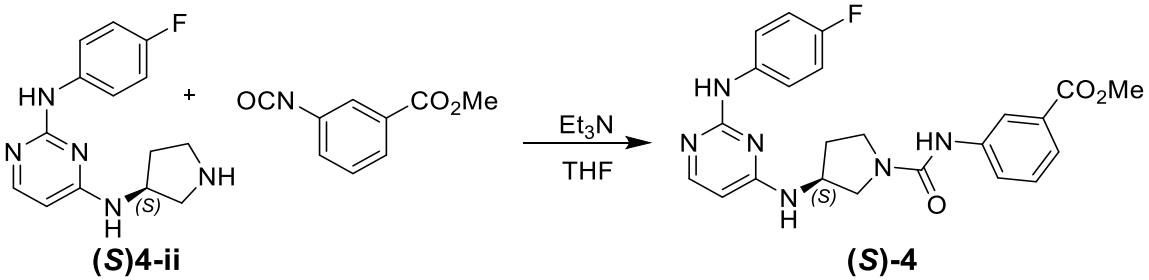
(*S*)4-i

2,4-Dichloropyrimidine (300 mg, 2.0 mmol) was dissolved in 5 mL of EtOH, the mixture was briefly warmed to 50 °C to solubilize. Triethylamine (308 μL, 2.2 mmol) was added, followed by (S)-1-boc-3-aminopyrrolidine (403 μL, 2.2 mmol). The reaction was stirred at ambient temperature for 18 h. The solvent was removed under reduced pressure, and the crude mixture was diluted in EtOAc followed by washing 3 X with H_2O, brine, and dried over $MgSO_4$. The mixture was concentrated and purified on a Combiflash Rf system (Hexanes/EtOAc, 0-100% EtOAc) to afford **(*S*)4-i** as a colorless oil (389 mg, 1.4 mmol, 65%). ^{1}H NMR, mixture of rotational isomers (500 MHz, DMSO-d_6) δ 8.15 (d, *J* = 6.3 Hz, 1H), 7.93 (d, *J* = 5.9 Hz, 1H), 6.46 (d, *J* = 5.9 Hz, 1H), 4.38 (d, *J* = 27.9 Hz, 1H), 3.54 (dt, *J* = 10.7, 5.0 Hz, 1H), 3.48 – 3.29 (m, 4H), 3.11 (dd, *J* = 10.9, 4.1 Hz, 1H), 2.11 (dq, *J* = 13.3, 6.5 Hz, 1H), 1.81 (tt, *J* = 12.2, 6.2 Hz, 1H), 1.40 (s, *J* = 4.8 Hz, 9H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 163.00, 159.71, 155.65, 153.49, 105.23, 78.41, 50.98, 50.68, 49.97, 49.26, 43.89, 43.72, 30.68, 29.92, 28.15. HRMS (ESI-TOF) calculated for $C_{13}H_{19}ClN_4O_2H^+$ $[M+H]^+$: 299.1167, observed 299.1267.



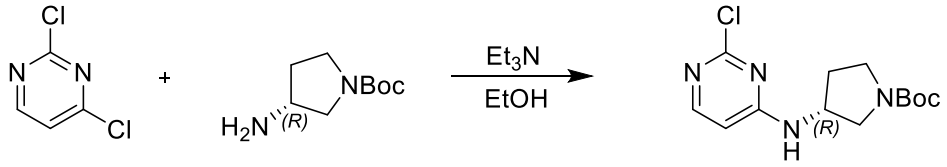
(*S*)4-i **(*S*)4-ii**

(*S*)4-i (268 mg, 0.899 mmol) was dissolved in 5 mL of 1,4-dioxane, followed by addition of 4-fluoroaniline (127 μL, 1.35 mmol). The mixture was heated to reflux for 5 h, followed by removal of solvent. Upon cooling to ambient temperature, diethyl ether was used to precipitate out the intermediate, and was collected via vacuum filtration. The filtered solid was dissolved in 3 mL of trifluoroacetic acid, and stirred for 1 h. The solvent was removed, and the mixture was dissolved in H_2O. The aqueous layer was washed 3 X with DCM and separated. The aqueous layer was adjusted to pH 12 using 1M NaOH, and extracted into EtOAc. The organic layer was washed with brine, and dried over $MgSO_4$ followed by filtration and concentration to isolate **(*S*)4-ii** (215 mg, 0.788 mmol, 87% over 2 steps). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.04 – 8.93 (m, 1H), 8.32 (s, 1H), 7.77 (dd, J = 9.0, 5.0 Hz, 3H), 7.48 – 7.19 (m, 1H), 7.05 (t, J = 8.9 Hz, 2H), 5.99 – 5.89 (m, 1H), 4.31 (s, 1H), 3.63 – 3.16 (m, 1H), 3.09 (dd, J = 11.3, 6.4 Hz, 1H), 2.95 (dt, J = 10.9, 7.3 Hz, 1H), 2.85 (ddd, J = 10.9, 8.1, 5.8 Hz, 1H), 2.70 (dd, J = 11.3, 4.4 Hz, 1H), 2.04 (dq, J = 14.6, 7.5 Hz, 1H), 1.64 (ddt, J = 12.9, 7.8, 5.2 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 161.95, 159.48, 156.40 (d, J = 236.8 Hz), 137.67, 119.73 (d, J = 7.3 Hz), 114.52 (d, J = 21.8 Hz), 79.06, 52.55, 45.01, 32.34. HRMS (ESI-TOF) calculated for $C_{14}H_{16}FN_5H^+$ $[M+H]^+$: 274.1423, observed 274.1461.



(*S*)4-ii **(*S*)-4**

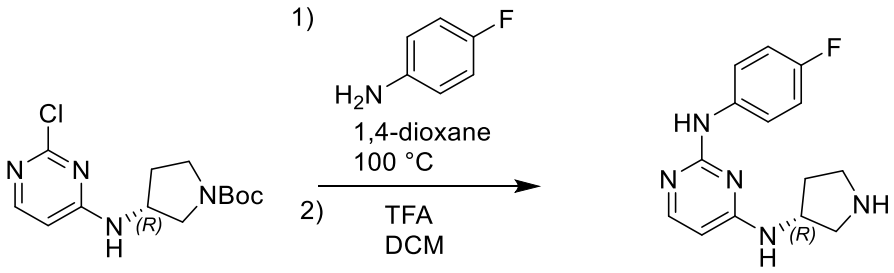
(*S*)4-ii (84 mg, 0.31 mmol) was dissolved in 3 mL of THF, followed by addition of triethylamine (0.47 mmol) and methyl 3-isocyanatobenzoate (84 mg, 0.47 mmol). The reaction was stirred at ambient temperature for 2 h, followed by concentration and purification on a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **(*S*)-4** as a white solid (113 mg, 0.25 mmol, 82%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.03 (s, 1H), 8.46 (s, 1H), 8.19 (t, *J* = 1.9 Hz, 1H), 7.86 – 7.81 (m, 2H), 7.81 – 7.75 (m, 2H), 7.52 (dt, *J* = 7.7, 1.3 Hz, 1H), 7.47 (s, 1H), 7.37 (t, *J* = 7.9 Hz, 1H), 7.05 (t, *J* = 8.9 Hz, 2H), 5.98 (d, *J* = 5.8 Hz, 1H), 4.51 (s, 1H), 3.83 (s, 3H), 3.76 (dd, *J* = 10.7, 6.1 Hz, 1H), 3.58 (dt, *J* = 10.2, 7.3 Hz, 1H), 3.51 (td, *J* = 9.8, 8.9, 5.5 Hz, 1H), 3.36 (dd, *J* = 10.7, 4.3 Hz, 1H), 2.21 (dq, *J* = 13.5, 7.0 Hz, 1H), 1.95 (dq, *J* = 12.7, 6.1 Hz, 1H).^{13}C NMR (126 MHz, DMSO-d_6) δ 166.23, 161.97, 159.49, 156.45 (d, *J* = 236.9 Hz), 154.73, 153.69, 140.91, 137.61 (d, *J* = 2.2 Hz), 129.59, 128.57, 123.62, 122.04, 119.81 (d, *J* = 7.4 Hz), 119.73, 114.56 (d, *J* = 21.8 Hz), 97.74, 51.93, 51.33, 49.66, 43.98, 30.42. HRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_3H^+$ $[M+H]^+$: 451.1849, observed 451.1883.



(*R*)4-i

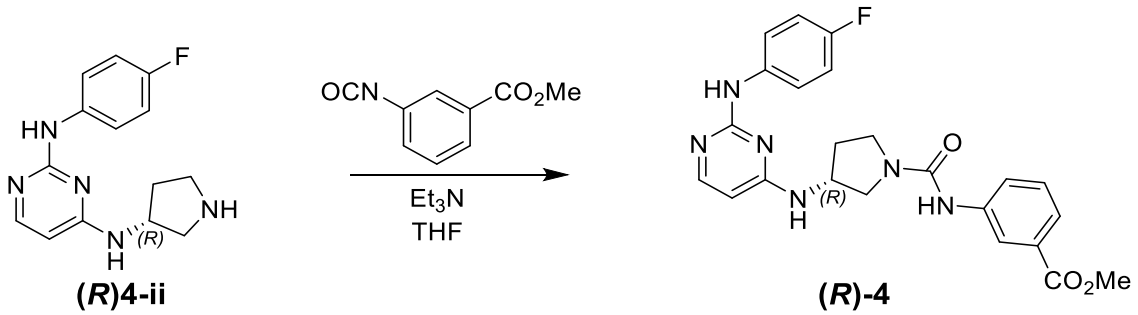
2,4-Dichloropyrimidine (300 mg, 2.0 mmol) was dissolved in 5 mL of EtOH, the mixture was briefly warmed to 50 °C to solubilize. Triethylamine (2.2 mmol) was added, followed by (*R*)-1-boc-3-aminopyrrolidine (403 µL, 2.2 mmol). The reaction was stirred at ambient temperature for 18 h. The solvent was removed under reduced pressure, and the crude mixture was diluted in EtOAc followed by washing 3 X with H_2O, brine, and dried over $MgSO_4$. The mixture was concentrated and purified on a Combiflash Rf system (Hexanes/EtOAc, 0-100% EtOAc) to afford **(*R*)4-i** as a colorless oil (357 mg, 1.2 mmol, 60%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.05 – 7.96 (m, 1H), 6.30 (d, *J* = 5.9 Hz, 1H), 5.83 – 5.58 (m, 1H), 3.67 (dd, *J* = 11.5, 5.9 Hz, 1H), 3.53 – 3.40 (m, 2H), 3.36 – 3.20 (m, 1H), 2.21 (p, *J* = 6.8 Hz, 1H), 2.04 (s, 1H), 1.99 – 1.85 (m, 1H), 1.45 (s, 9H). ^{13}C NMR

(126 MHz, Chloroform-*d*) δ 162.95, 160.72, 154.50, 79.86, 51.61, 51.05, 50.24, 43.89, 43.64, 31.61, 30.86, 28.45. HRMS (ESI-TOF) calculated for $C_{13}H_{19}ClN_4O_2H^+$ $[M+H]^+$: 299.1276, observed 299.1249.



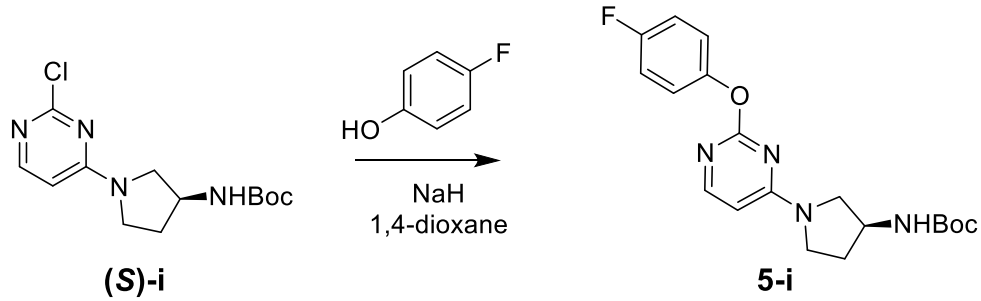
(*R*)4-i **(*R*)4-ii**

(*R*)4-i (330.7 mg, 1.107 mmol) was dissolved in 4 mL of 1,4-dioxane, to which 4-fluoroaniline (158 μL, 1.66 mmol) was added. The solution was heated to reflux for 18 h, followed by removal of solvent. When cooled to rt, diethyl ether was added, and the resulting precipitate was collected by vacuum filtration and washed with additional diethyl ether. The dried precipitate was dissolved in 3 mL of TFA and 1 mL of DCM and stirred at ambient temperature for 1 h. The solvent was removed, and the crude mixture was dissolved in H_2O and DCM. The aqueous layer was washed 3 X with DCM and separated. The pH was adjusted to > 10 using 1M NaOH and extracted 3 X with EtOAc. The combined organic layers were washed with brine and dried over $MgSO_4$. Following filtration, the solvent was removed yielding **(*R*)4-ii** as a brown oil (148.9 mg, 0.5448 mmol, 49%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.91 (d, *J* = 5.8 Hz, 1H), 7.52 (dd, *J* = 8.8, 4.8 Hz, 2H), 7.08 (d, *J* = 26.9 Hz, 1H), 6.98 (t, *J* = 8.6 Hz, 2H), 5.85 (d, *J* = 5.8 Hz, 1H), 5.22 (d, *J* = 6.8 Hz, 1H), 4.34 (s, 1H), 3.18 (ddd, *J* = 19.6, 13.0, 6.7 Hz, 2H), 3.04 – 2.99 (m, 1H), 2.97 (dd, *J* = 12.0, 3.6 Hz, 1H), 2.63 (d, *J* = 14.4 Hz, 3H), 2.21 (ddd, *J* = 15.5, 13.8, 7.4 Hz, 1H), 1.78 (td, *J* = 10.7, 8.0, 4.6 Hz, 1H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 162.28, 159.80, 158.22 (d, *J* = 240.8 Hz), 156.07, 135.98, 121.11 (d, *J* = 7.7 Hz), 115.23 (d, *J* = 22.3 Hz), 53.30, 51.63, 45.36, 33.06. HRMS (ESI-TOF) calculated for $C_{14}H_{16}FN_5H^+$ $[M+H]^+$: 274.1423, observed 274.1438.



(*R*)4-ii **(*R*)-4**

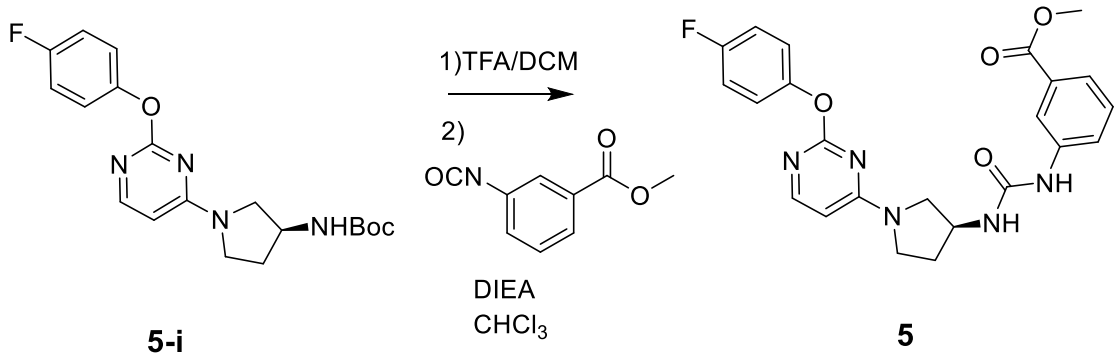
(*R*)4-ii (120 mg, 0.44 mmol) was dissolved in 5 mL of THF, followed by addition of triethylamine (92 μL, 0.66 mmol) and methyl 3-isocyanatobenzoate (120 mg, 0.66 mmol). The reaction was stirred at ambient temperature for 2 h, followed by removal of solvent and purification on a Combiflash Rf System (hexanes/EtOAc, 0-100% EtOAc) to afford **(*R*)-4** as a white solid (110 mg, 55%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.03 (s, 1H), 8.46 (s, 1H), 8.19 (t, *J* = 1.9 Hz, 1H), 7.87 – 7.81 (m, 2H), 7.78 (dd, *J* = 9.0, 5.0 Hz, 2H), 7.52 (dt, *J* = 7.8, 1.2 Hz, 1H), 7.50 – 7.45 (m, 1H), 7.37 (t, *J* = 7.9 Hz, 1H), 7.05 (t, *J* = 8.9 Hz, 2H), 5.98 (d, *J* = 5.8 Hz, 1H), 4.50 (s, 1H), 3.84 (s, 3H), 3.76 (dd, *J* = 10.7, 6.1 Hz, 1H), 3.58 (dt, *J* = 10.2, 7.3 Hz, 1H), 3.51 (td, *J* = 10.2, 9.1, 5.6 Hz, 1H), 3.36 (dd, *J* = 10.7, 4.4 Hz, 1H), 2.21 (dq, *J* = 13.5, 7.0 Hz, 1H), 1.95 (qd, *J* = 11.2, 9.9, 3.9 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.23, 161.97, 159.47, 156.46 (d, *J* = 236.7 Hz), 153.68, 140.90, 137.59 (d, *J* = 2.3 Hz), 129.59, 128.57, 123.62, 122.04, 119.85, 119.76 (d, *J* = 8.6 Hz), 114.56 (d, *J* = 21.9 Hz), 51.93, 51.32, 49.66, 43.97, 30.41, 19.01. HRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_3H^+$ $[M+H]^+$: 451.1849, observed 451.1882.



(*S*)-i **5-i**

In a sealed reaction vessel 16.4 mg (0.68 mmol) of NaH was dissolved in 0.5 mL of 1,4-dioxane followed by addition of 76.7 mg (0.68 mmol) of 4-fluorophenol and stirred for 0.25 h. 51.1 mg (0.17 mmol) of **(*S*)-i** was then added and the mixture was heated to 100°

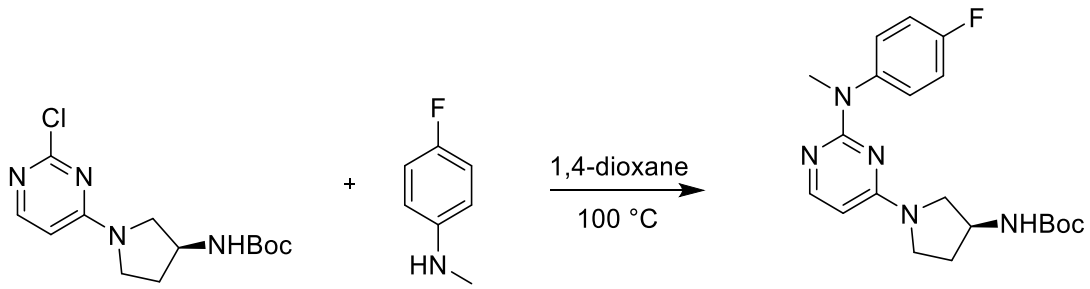
for 48 h. The solution was diluted with 10 mL H_2O and extracted with 10 mL of EtOAc (3 X). The crude material was concentrated and purified via column chromatography (hexanes:EtOAc) to afford **5-i**, used without further purification. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.98 (d, *J* = 6.0 Hz, 1H), 7.18 – 7.11 (m, 2H), 7.09 – 7.03 (m, 2H), 6.03 (d, *J* = 6.0 Hz, 1H), 4.84 (s, 1H), 4.32 (s, 1H), 3.81 – 3.17 (m, 4H), 2.25 (s, 1H), 2.04 (s, 1H), 1.45 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 171.18, 164.90, 161.91, 159.64 (d, *J* = 242.8 Hz), 156.58, 155.30, 148.99 (d, *J* = 2.8 Hz), 123.21 (d, *J* = 8.4 Hz), 115.82 (d, *J* = 23.3 Hz), 99.12, 60.40, 52.51, 50.39, 44.52, 40.90, 31.72, 30.80, 28.35, 21.04, 14.19. HRMS (ESI-TOF) calculated for $C_{19}H_{23}FN_4O_3H^+$ $[M+H]^+$: 375.1827, observed 375.1814.



5-i **5**

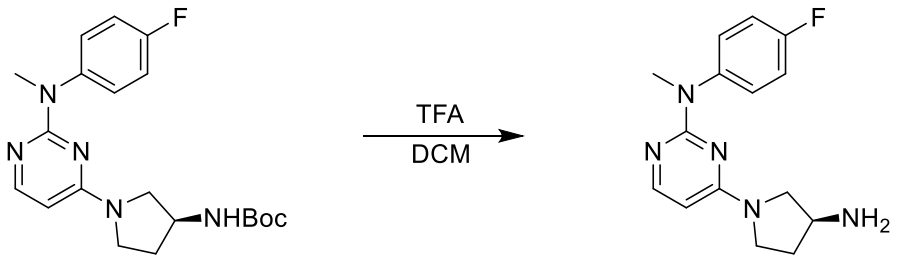
41.6 mg (0.11 mmol) of **5-i** was titrated with 1 mL of trifluoroacetic acid for 12 hr. The solution was neutralized and the resulting amine was extracted with 5 mL of ethyl acetate. The ethyl acetate was removed under vacuum and the resulting oil was resuspended in 1 mL of chloroform. 19.7 mg (0.11 mmol) of methyl 3-isocyanobenzoate was added and the solution was stirred for 12 h. 200 μL of diisopropylethylamine (0.11 mmol) was added and the reaction was stirred for an additional 12 h. The mixture was diluted with 2 mL of DI water and extracted with 3 X 10 mL of DCM. The solvent was removed and **5** was isolated by column chromatography (hexanes:ethyl acetate), 44% yield. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.61 (d, *J* = 11.6 Hz, 1H), 8.14 (s, 1H), 8.01 – 7.90 (m, 1H), 7.59 – 7.48 (m, 2H), 7.37 (t, *J* = 7.9 Hz, 1H), 7.25 – 7.12 (m, 4H), 6.67 – 6.52 (m, 1H), 6.31 (d, *J* = 5.9 Hz, 1H), 4.30 (d, *J* = 44.1 Hz, 1H), 3.84 (s, 3H), 3.61 (s, 1H), 3.46 (s, 2H), 3.38-3.26 (m, 1H), 2.19 (d, *J* = 38.8 Hz, 1H), 1.94 (d, *J* = 42.0 Hz, 1H) ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.73, 164.80, 159.33 (d, *J* = 240.1 Hz), 156.78, 155.23, 149.65, 141.10, 130.53, 129.53,

123.95 (d, J = 8.5 Hz), 122.71, 122.36, 118.60, 116.21 (d, J = 23.3 Hz), 99.99, 52.57, 52.12, 49.71, 49.05, 44.98, 44.80, 31.47, 30.70. HRMS (ESI-TOF) calculated for $C_{23}H_{22}FN_5O_4H^+$ $[M+H]^+$: 452.1729, observed 452.1710.



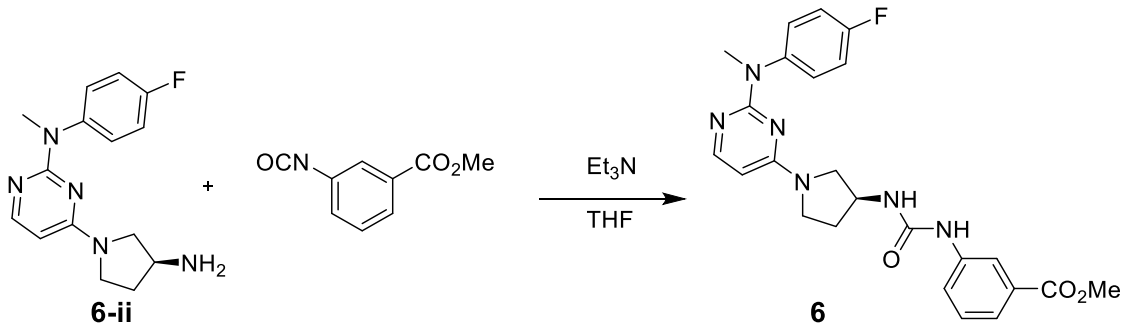
(*S*)-i **6-i**

(*S*)-i (200 mg, 0.67 mmol) was dissolved in 5 mL of 1,4-dioxane, followed by addition of N-methyl-4-fluoroaniline (180 µL, 2.0 mmol) and was heated to reflux for 18 h. The solvent was removed under reduced pressure and the mixture was redissolved in DCM. The organic layer was washed with 10% $NaHCO_3$, followed by H_2O and brine and dried over $MgSO_4$. Following filtration, the crude mixture was concentrated and purified by flash chromatography (1:1 Hexanes/EtOAc) to afford **6-i** as a beige solid (160 mg, 0.41 mmol, 62%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.81 (d, J = 5.8 Hz, 1H), 7.34 (dd, J = 8.8, 5.1 Hz, 2H), 7.14 (t, J = 8.8 Hz, 2H), 5.84 (d, J = 5.8 Hz, 1H), 4.06 (s, 1H), 3.78 – 3.02 (m, 3H), 3.39 (s, 3H), 2.06 (br. m, 1H), 1.82 (br. m, 1H), 1.38 (s, 9H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 160.76, 158.86 (d, J = 240.7 Hz), 155.35, 142.12, 128.02 (d, J = 8.2 Hz), 114.78 (d, J = 22.2 Hz), 94.70, 77.86, 51.18, 44.02, 37.66, 28.20. HRMS (ESI-TOF) calculated for $C_{20}H_{26}FN_5O_2H^+$ $[M+H]^+$: 388.2104, observed 388.2141.



6-i **6-ii**

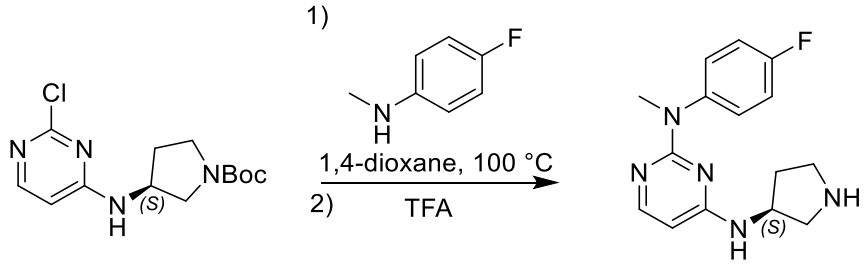
6-i (41.8 mg, 0.108 mmol) was dissolved in 3 mL of TFA and 1 mL of DCM. The reaction was stirred at ambient temperature for 1 h. The solvent was removed, and the crude mixture was dissolved in H_2O and DCM. The aqueous layer was washed 3 X with DCM and separated. The aqueous layer was basified to pH > 10 using 1M NaOH, and then extracted 3 X with EtOAc. The combined organic layers were washed with $MgSO_4$ and brine, followed by filtration and concentration to isolate **6-ii** as a brown oil (20.8 mg, 67%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.89 (d, *J* = 5.9 Hz, 1H), 7.30 – 7.26 (m, 2H), 7.02 (t, *J* = 8.7 Hz, 2H), 5.71 (d, *J* = 5.9 Hz, 1H), 3.69 – 3.49 (m, 3H), 3.46 (s, 3H), 3.41 – 3.00 (m, 1H), 2.18 – 2.09 (m, 1H), 1.88 – 1.70 (m, 3H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 161.36, 160.35, 159.88 (d, *J* = 243.6 Hz), 155.42, 142.08 (d, *J* = 3.0 Hz), 128.13 (d, *J* = 8.2 Hz), 115.28 (d, *J* = 22.3 Hz), 94.32, 54.31, 44.38, 38.27, 34.32. HRMS (ESI-TOF) calculated for $C_{15}H_{18}FN_5H^+$ $[M+H]^+$: 288.1580, observed 288.1624.



6-ii **6**

6-ii (68 mg, 0.24 mmol) was dissolved in 3 mL of THF, followed by addition of triethylamine (49 μL, 0.36 mmol) and methyl 3-isocyanatobenzoate (63 mg, 0.36 mmol).

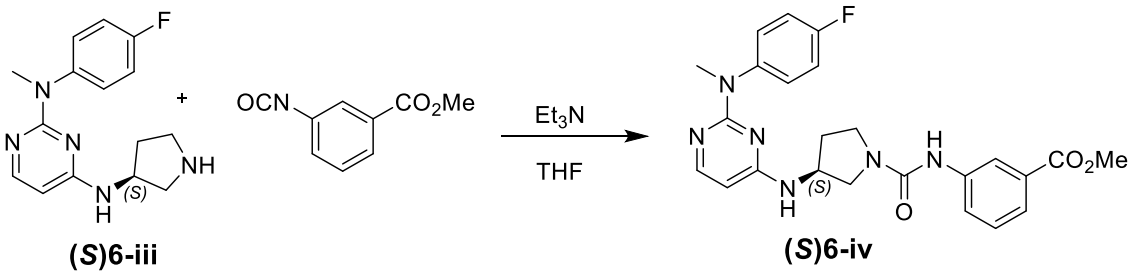
The reaction was stirred at ambient temperature for 2 h. The solvent was removed under reduced pressure and the crude mixture was redissolved in EtOAc and washed 3 X with H_2O followed by brine and $MgSO_4$. Following filtration, the organic layer was concentrated and purified on a Combiflash Rf system (Hexanes/EtOAc, 0-100% EtOAc) to afford **6** as a yellow solid (83 mg, 0.18 mmol, 75%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.67 (s, 1H), 8.13 (d, *J* = 1.9 Hz, 1H), 7.83 (d, *J* = 5.8 Hz, 1H), 7.55 (dd, *J* = 8.1, 2.2 Hz, 1H), 7.50 (d, *J* = 7.6 Hz, 1H), 7.37 (d, *J* = 8.2 Hz, 1H), 7.34 (d, *J* = 7.2 Hz, 2H), 7.15 (t, *J* = 8.7 Hz, 2H), 6.63 (d, *J* = 6.8 Hz, 1H), 5.89 (d, *J* = 5.8 Hz, 1H), 4.27 (s, 1H), 3.83 (s, 3H), 3.57 (dd, *J* = 11.2, 6.0 Hz, 1H), 3.40 (s, 3H), 2.14 (d, *J* = 12.6 Hz, 1H), 1.90 (s, 1H), 1.23 (s, 1H ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.27, 160.80, 159.96, 158.90 (d, *J* = 241.0 Hz), 155.47, 154.76, 142.13 (d, *J* = 2.9 Hz), 140.68, 130.04, 129.04, 128.07 (d, *J* = 8.4 Hz), 122.17, 121.83, 118.06, 114.83 (d, *J* = 22.2 Hz), 94.72, 52.08, 43.97, 37.69. HRMS (ESI-TOF) calculated for $C_{24}H_{25}FN_6O_3H^+$ $[M+H]^+$: 465.2006, observed 465.2041.



(*S*)4-i **(*S*)6-iii**

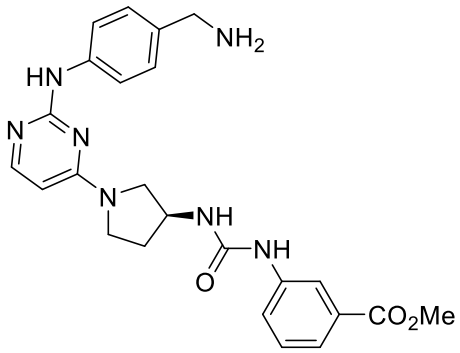
(*S*)4-i (304 mg, 1.02 mmol, 1 eq) was dissolved in 8 mL of 1,4-dioxane, followed by addition of N-methyl-4-fluoroaniline (200 µL, 1.53 mmol). The mixture was heated to reflux for 18 h, followed by removal of solvent. Upon cooling to ambient temperature, diethyl ether was used to precipitate out the intermediate, and was collected via vacuum filtration. The filtered solid was dissolved in 3 mL of trifluoroacetic acid and stirred for 1 h. The solvent was removed, and the mixture was dissolved in H_2O. The aqueous layer was washed 3 X with DCM and separated. The aqueous layer was adjusted to pH 12 using

1M NaOH and extracted into EtOAc. The organic layer was washed with brine and dried over $MgSO_4$. The material was filtered and concentrated to isolate **(*S*)6-iii** as a brown oil (267 mg, used without further purification). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.79 – 7.66 (m, 1H), 7.41 – 7.26 (m, 2H), 7.18 – 7.10 (m, 2H), 5.84 (d, J = 5.8 Hz, 1H), 4.26 – 3.96 (m, 1H), 3.67 – 3.48 (m, 2H), 3.38 (s, 3H), 2.92 (dd, J = 11.3, 6.5 Hz, 1H), 2.86 (dt, J = 11.0, 7.3 Hz, 1H), 2.78 – 2.69 (m, 1H), 2.64 – 2.54 (m, 1H), 1.96 – 1.85 (m, 1H), 1.78 – 1.71 (m, 1H), 1.66 – 1.49 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 160.96, 158.80 (d, J = 240.8 Hz), 142.11 (d, J = 2.9 Hz), 128.07 (d, J = 8.2 Hz), 114.70 (d, J = 22.3 Hz), 66.90, 52.69, 45.11, 37.62, 32.35, 30.09, 25.01. HRMS (ESI-TOF) calculated for $C_{15}H_{18}FN_5H^+$ $[M+H]^+$: 288.1617, observed 288.1580.



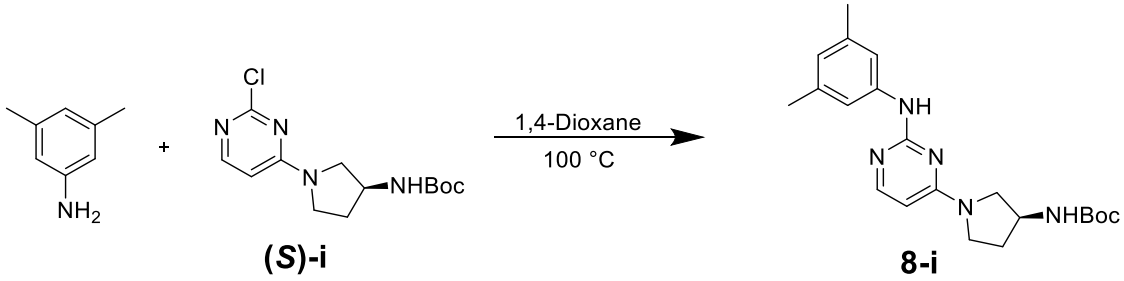
(*S*)6-iii **(*S*)6-iv**

(*S*)6-iii (93 mg, 0.32 mmol, 1 eq) was dissolved in 2 mL of THF, followed by addition of triethylamine (68 μL, 0.49 mmol) and methyl 3-isocyanatobenzoate (86 mg, 0.49 mmol) and was stirred at ambient temperature for 2 h. The crude mixture was concentrated and purified on a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **(*S*)6-iv** as a yellow solid (103 mg, 0.22 mmol, 69%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.41 (s, 1H), 8.19 (d, J = 2.0 Hz, 1H), 7.84 (dd, J = 8.3, 2.1 Hz, 1H), 7.74 (d, J = 5.5 Hz, 1H), 7.52 (dt, J = 7.8, 1.2 Hz, 1H), 7.40 – 7.31 (m, 4H), 7.14 (t, J = 8.8 Hz, 2H), 5.89 (d, J = 5.7 Hz, 1H), 4.26 (s, 1H), 3.84 (s, 3H), 3.63 (dd, J = 10.7, 6.2 Hz, 1H), 3.52 (dt, J = 10.2, 7.1 Hz, 1H), 3.44 (d, J = 18.4 Hz, 1H), 3.28 (dd, J = 10.7, 4.6 Hz, 1H), 2.09 (dq, J = 13.4, 7.0 Hz, 1H), 1.88 (dt, J = 12.5, 7.2 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.23, 161.66, 160.94, 158.84 (d, J = 241.0 Hz), 154.62, 153.64, 142.01 (d, J = 2.9 Hz), 140.92, 129.58, 128.55, 128.09 (d, J = 8.2 Hz), 123.61, 122.01, 119.70, 114.72 (d, J = 22.2 Hz), 96.74, 51.93, 51.21, 49.62, 43.91, 37.64, 30.22. HRMS (ESI-TOF) calculated for $C_{24}H_{25}FN_6O_3H^+$ $[M+H]^+$: 465.2006, observed 465.2070.



7

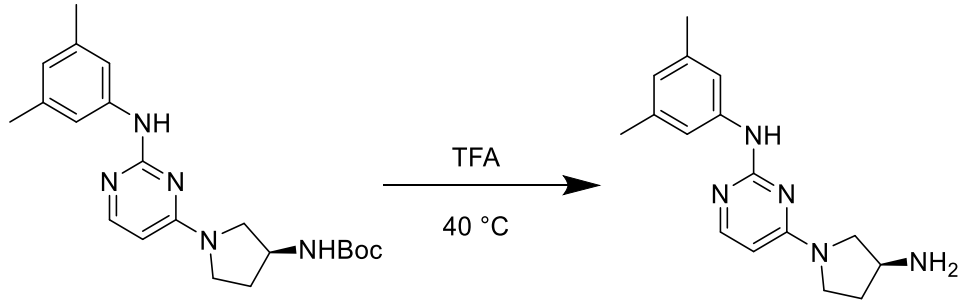
Compound **7** was synthesized as reported previously.[167]



(*S*)-i **8-i**

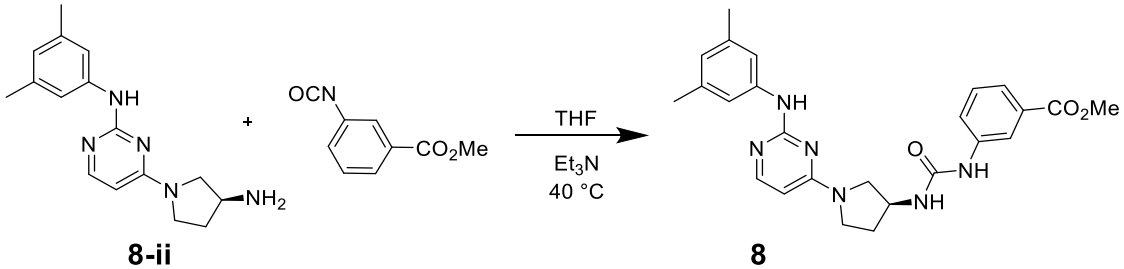
(*S*)-i (302 mg, 1.01 mmol) was dissolved in 8 mL of 1,4-dioxane. 3,5-Dimethylaniline (302.4 µL, 1.18 mmol) was added and the reaction mixture was stirred at 100 ℃ for 6 h. The temperature was reduced to rt and the reaction mixture was stirred for 72 h. Upon completion of the reaction the solvent was removed and the orange/white solid was dissolved in EtOAc. The organic layer was washed 3 X with water, 1 X with brine, and dried over $MgSO_4$. The crude material was filtered and concentrated followed by purification using a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to give the white solid **8-i** (172.7 mg, 46.0%). ^{1}H NMR as a mixture of rotational isomers (500 MHz, Chloroform-*d*) δ 7.59 (s, 1H), 7.33 (s, 1H), 6.77 (s, 2H), 5.90 (d, *J* = 7.1 Hz, 0.5H (rot.)), 5.86 (d, *J* = 7.1 Hz, 0.5H (rot.) 5.14 (s, 0.5H (rot.)), 5.06 (s, 0.5H (rot.)), 4.93 (s, 0.5H (rot.)), 4.38 (m, 1H), 3.95 (m, 0.5H (rot.)), 3.88 – 3.71 (m, 1H), 3.66 – 3.51 (m, 1H), 3.40 (d, *J* = 11.0 Hz, 0.5H (rot.)), 3.17 (s, 0.5H (rot.)), 2.31 (s, 6H), 2.15 (s, 1H), 1.86 (s, 1H), 1.47 (s, 9H). ^{13}C NMR as a mixture of rotational isomers (126 MHz, Chloroform-*d*) δ

159.98, 138.26, 125.96, 125.00, 118.59, 95.28, 95.04, 53.05, 50.84, 45.81, 45.21, 44.82, 32.18, 31.44, 30.47, 29.69, 28.28, 26.38, 23.42, 21.45 HRMS (ESI-TOF) calculated for $C_{21}H_{29}N_5O_2H^+$ $[M+H]^+$: 384.2355, observed 384.2412.



8-i **8-ii**

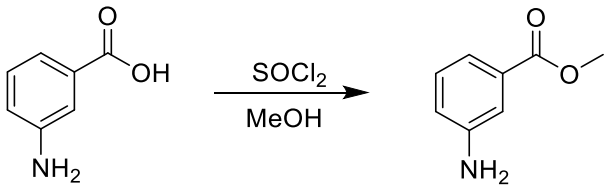
8-i (173 mg, 0.450 mmol) was dissolved in 8 mL neat TFA. The reaction was stirred at 40 °C for 2 h. The solvent was removed. The white solid was dissolved in EtOAc and the organic layer was washed 3 X with 1M NaOH, 1 X brine, and dried over $MgSO_4$. The material was filtered and concentrated to the white solid **8-ii** (109.6 mg, 85.6%) that was taken forward without further purification. ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.95 (d, *J* = 5.8 Hz, 1H), 7.32 (s, 2H), 7.01 (s, 1H), 6.65 (s, 1H), 5.82 (d, *J* = 5.9 Hz, 1H), 3.89-3.02 (m, 4H), 3.75 (t, *J* = 5.6 Hz, 1H), 2.32 (s, 6H), 2.22 (dd, *J* = 12.9, 6.6 Hz, 1H), 1.92 – 1.79 (m, 1H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 160.70, 159.60, 155.59, 140.27, 138.20, 123.33, 116.70, 95.50, 54.78, 44.71, 34.50, 21.58. HRMS (ESI-TOF) calculated for $C_{16}H_{21}N_5H^+[M+H]^+$: 284.1821, observed 284.1877.



8-ii **8**

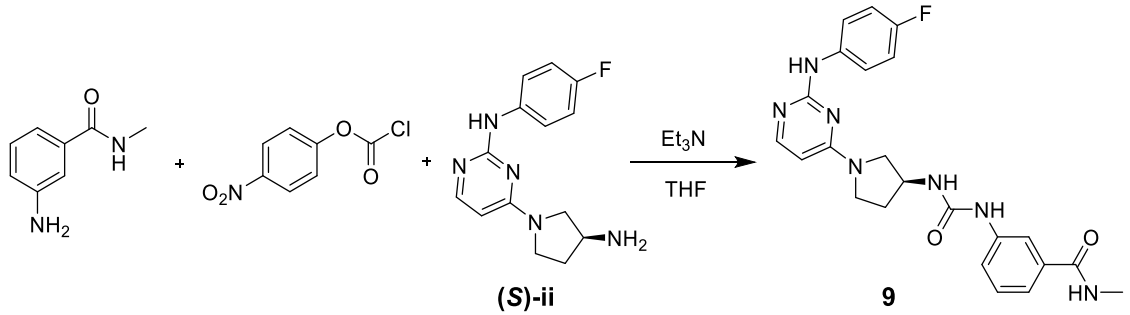
8-ii (100 mg, 0.353 mmol) and methyl-3-isocyanatobenzoate (83.3 mg, 0.470 mmol) were dissolved in 7 mL of THF. Triethylamine (60 µL, 0.462 mmol) was added and the reaction mixture was stirred at 40 ℃ for 90 min. Upon reaction completion the solvent was removed and the white solid that remained was dissolved in EtOAc. The organic layer was washed

3 X with water, 1 X brine, dried over $MgSO_4$. The material was filtered and concentrated followed by purification using a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to give the white solid **8** (70.2 mg, 44.5%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.87 (s, 1H), 8.65 (s, 1H), 8.15 (s, 1H), 7.93 (d, *J* = 5.9 Hz, 1H), 7.55 (dd, 8.0 Hz, 1H), 7.50 (d, *J* = 7.6 Hz, 1H), 7.45 (s, 2H), 7.37 (t, *J* = 7.9 Hz, 1H), 6.66 (d, *J* = 6.6 Hz, 1H), 6.52 (s, 1H), 5.96 (d, *J* = 5.9 Hz, 1H), 4.32 (s, 1H), 3.84 (s, 3H), 3.77-3.38 (m, 4H), 2.22 (s, 6H), 2.00 (s, 1H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 167.04, 159.85, 139.53, 138.59, 137.89, 130.81, 129.03, 123.61, 123.14, 119.49, 117.26, 52.14, 21.40. HRMS(ESI-TOF) calculated for $C_{25}H_{28}N_6O_3H^+$ $[M+H]^+$: 461.2256, observed 461.2327.



9-i

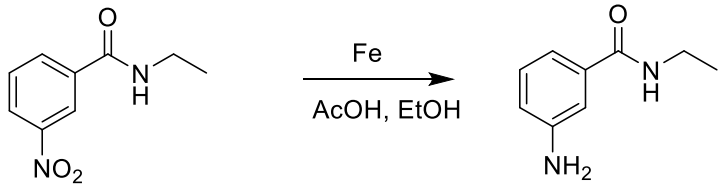
3-Aminobenzoic acid (500 mg, 3.64 mmol) was dissolved in 12 mL of MeOH. Briefly, the solution was warmed to improve solubility. Thionyl chloride (660 μL, 10.9 mmol) was added dropwise, and the solution was stirred at ambient temperature for 3 h. The solvent was removed, and the material was dissolved in EtOAc, washed with 10% $NaHCO_3$, followed by brine and dried over $MgSO_4$. Following filtration, the solvent was removed to afford **9-i** as a brown oil (220 mg, 40%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.21 – 7.18 (m, 1H), 7.13 (t, *J* = 7.6 Hz, 1H), 7.09 (dt, *J* = 7.7, 1.5 Hz, 1H), 6.79 (ddd, *J* = 7.8, 2.4, 1.3 Hz, 1H), 5.37 (s, 2H), 3.79 (s, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.64, 148.88, 130.06, 128.95, 118.20, 116.19, 113.96, 51.70. HRMS (ESI-TOF) calculated for $C_8H_9NO_2H^+$ $[M+H]^+$: 152.0667, observed 152.0705.



4-Nitrophenylchloroformate (161 mg, 0.798 mmol) was dissolved in 2 mL of anhydrous THF and cooled to 0 °C. Separately, 3-aminomethylbenzamide (70 mg, 0.798 mmol) and triethylamine (112 μL, 0.798 mmol) were dissolved in 1 mL of anhydrous THF. The second solution was added dropwise to the first, and the reaction was warmed to ambient temperature for 15 minutes. The mixture was diluted with EtOAc and washed 3 X with H_2O. The organic fraction was washed with brine and dried over $MgSO_4$ followed by filtration. **(*S*)-ii** (110 mg, 0.40 mmol) and triethylamine (112 μL) were added to the EtOAc fraction, and the reaction was stirred at ambient temperature for 1 h. The organic layer was washed 3 X with 1M NaOH, followed by brine and $MgSO_4$. The material was filtered and concentrated followed by purification via column chromatography (9:1 EtOAc/MeOH) to isolate **9** as a white solid (40 mg, 40%, 2 steps) ^{1}H NMR, mixture of rotational isomers (500 MHz, DMSO-d_6) δ 9.09 (s, 1H), 8.48 (s, 1H), 8.32 (dd, J = 4.5 Hz, 1H), 7.93 (d, J = 5.9 Hz, 1H), 7.82 – 7.76 (m, 3H), 7.56 – 7.51 (m, 1H), 7.33 (d, J = 7.7 Hz, 1H), 7.29 (t, J = 7.8 Hz, 1H), 7.07 (t, J = 8.7 Hz, 2H), 6.60 (d, J = 6.8 Hz, 1H), 5.99 (d, J = 5.9 Hz, 1H), 4.33 (s, 1H), 3.80 – 3.19 (m, 4H), 2.76 (s, 1H), 2.75 (s, 2H), 2.21 (s, 1H), 2.01 – 1.89 (m, 1H).^{13}C NMR (126 MHz, DMSO-d_6) δ 166.75, 166.67, 160.17, 159.15, 156.58 (d, J = 236.7 Hz), 155.36, 154.80, 140.28, 140.18, 137.67, 135.28, 128.52, 120.16, 119.83 (d, J = 7.4 Hz), 119.57, 116.79, 116.69, 114.72 (d, J = 21.9 Hz), 95.69, 51.99, 44.31, 26.25. HRMS (ESI-TOF) calculated for $C_{23}H_{24}FN_7O_2H^+$ $[M+H]^+$: 450.2009, observed 450.2044.

10-i

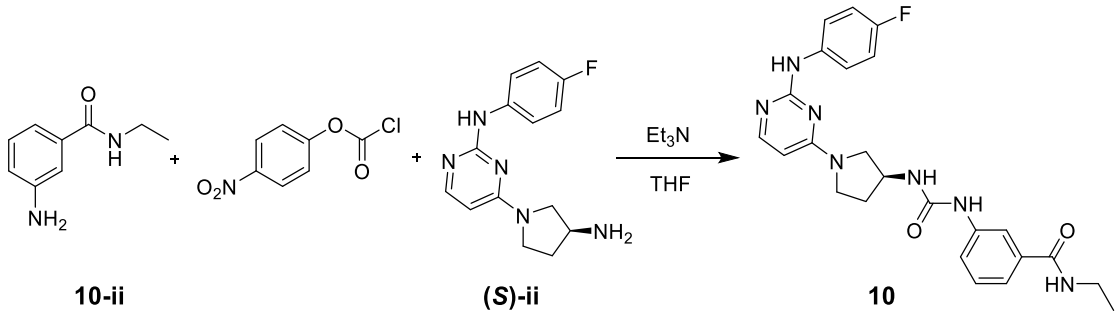
3-Nitrobenzoyl chloride (543.3 mg, 2.93 mmol) was dissolved in 5 mL of DCM. Next, Ethylamine (500 µL, 70% in H_2O, 5.86 mmol) was added, and an exotherm was observed. The reaction was stirred at ambient temperature for 1 h, followed by washing 3 X with H_2O and brine, then dried over $MgSO_4$. The material was filtered and concentrated followed by purification on a Combiflash Rf system (Hexanes/EtOAc, 0-100% EtOAc) to afford **10-i** as a white solid (278.8 mg, 49%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.85 (t, J = 5.5 Hz, 1H), 8.67 (t, J = 2.0 Hz, 1H), 8.42 – 8.34 (m, 1H), 8.28 (dt, J = 7.7, 1.3 Hz, 1H), 7.77 (t, J = 8.0 Hz, 1H), 3.32 (qd, J = 7.2, 5.5 Hz, 2H), 1.15 (t, J = 7.2 Hz, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 163.60, 147.64, 135.89, 133.46, 129.95, 125.57, 121.72, 34.21, 14.47. HRMS (ESI-TOF) calculated for $C_9H_{10}N_2O_3H^+$ $[M+H]^+$: 195.0725, observed 195.0763.



10-i **10-ii**

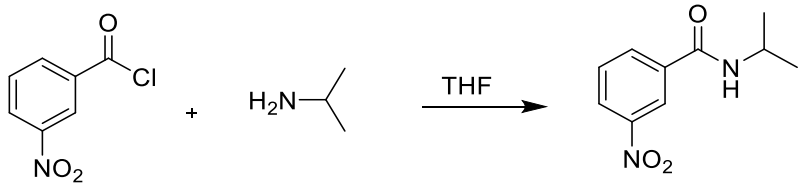
3-Nitroethylbenzamide(**10-i**) (230 mg, 1.19 mmol) was dissolved in 3 mL of EtOH. Iron powder (331 mg, 5.95 mmol) and acetic acid (677 µL, 5.95 mmol) were added, and the solution was stirred at ambient temperature for 3 h. The mixture was filtered, and the filtrate was diluted with EtOAc, followed by washing with 10% $NaHCO_3$. The organic layer was washed with brine and dried over $MgSO_4$. Following filtration and concentration **10-ii** was obtained as a brown oil (150 mg, 77%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.05 (t, J = 7.8 Hz, 1H), 7.00 (t, J = 2.0 Hz, 1H), 6.94 – 6.90 (m, 1H), 6.66 (dd, J = 7.9, 2.3 Hz, 1H), 5.21 (s, 1H), 3.23 (qd, J = 7.2, 5.5 Hz, 2H), 1.08 (t, J = 7.2 Hz, 3H). ^{13}C NMR (126

MHz, DMSO-d_6) δ 166.56, 148.48, 135.58, 128.42, 116.07, 114.05, 112.68, 33.75, 14.74. HRMS (ESI-TOF) calculated for $C_9H_{12}N_2OH^+$ [M+H]$^+$: 165.0983, observed 165.1017.



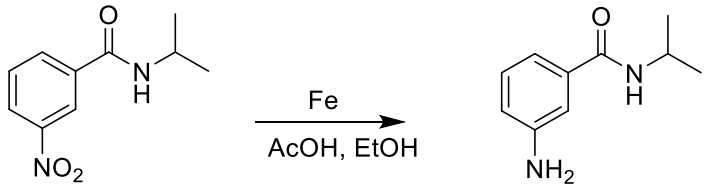
10-ii **(*S*)-ii** **10**

3-Aminoethylbenzamide(**10-ii**) (26.0 mg, 0.159 mmol) and triethylamine (33 μL, 0.236 mmol) were dissolved in 1 mL of THF. Separately, 4-nitrophenylchloroformate (47.6 mg, 0.236 mmol,) was dissolved in 2 mL of THF. The solutions were cooled to 0 °C, and the aniline solution was added dropwise to the chloroformate solution. The mixture was stirred at ambient temperature for 2h. **(*S*)-ii** was then added, and the reaction was stirred for 18 h. The solvent was removed, and it was diluted with EtOAc, followed by washing 3 X with H_2O. The organic layer was washed with brine and dried over $MgSO_4$. The material was filtered and concentrated followed by purification on a Combiflash Rf system (DCM/MeOH, 0-10% MeOH) to afford **10** (16.3 mg, 22%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.06 (s, 1H), 8.46 (s, 1H), 8.36 (t, *J* = 5.6 Hz, 1H), 7.93 (d, *J* = 5.8 Hz, 1H), 7.83 – 7.75 (m, 3H), 7.53 (dt, *J* = 7.9, 1.5 Hz, 1H), 7.36 – 7.32 (m, 1H), 7.28 (t, *J* = 7.8 Hz, 1H), 7.06 (t, *J* = 8.9 Hz, 2H), 6.58 (d, *J* = 6.8 Hz, 1H), 5.98 (d, *J* = 5.9 Hz, 1H), 4.33 (s, 1H), 3.53 (d, *J* = 95.3 Hz, 3H), 3.29 – 3.22 (m, 2H), 2.21 (s, 1H), 1.95 (s, 1H), 1.10 (t, *J* = 7.2 Hz, 3H).^{13}C NMR (126 MHz, DMSO-d_6) δ 165.96, 160.09, 158.30 (d, *J* = 238.2 Hz), 155.55, 155.47, 154.69, 140.13, 137.66, 135.37, 128.36, 120.04, 119.62 (d, *J* = 7.3 Hz), 119.55, 116.76, 114.59 (d, *J* = 21.8 Hz), 95.55, 51.87, 44.17, 33.88, 14.69. HRMS (ESI-TOF) calculated for $C_{24}H_{26}FN_7O_2H^+$ [M+H]$^+$: 464.2166, observed 464.2202.



11-i

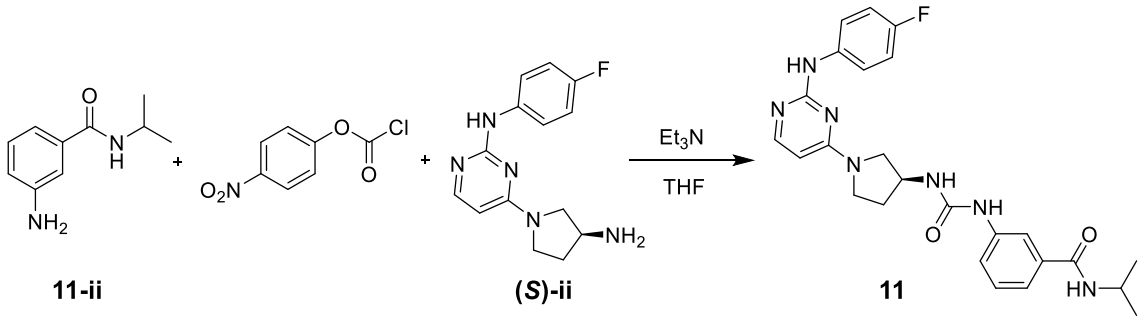
3-Nitrobenzoyl chloride (289.6 mg, 1.561 mmol) was dissolved in 2 mL of THF followed by dropwise addition of isopropylamine (667 µL, 7.797 mmol). The reaction was stirred at ambient temperature for 1.5 h. The mixture was diluted with EtOAc and washed 3 X with H_2O followed by brine and dried over $MgSO_4$. The material was filtered and the solvent was removed under reduced pressure to afford **11-i** as a white solid (65.6 mg, 0.315 mmol, 20%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.55 (t, *J* = 1.9 Hz, 1H), 8.36 – 8.31 (m, 1H), 8.15 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.63 (t, *J* = 8.0 Hz, 1H), 6.14 (s, 1H), 4.31 (dq, *J* = 13.5, 6.7 Hz, 1H), 1.30 (d, *J* = 6.6 Hz, 6H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 164.20, 148.09, 136.52, 133.25, 129.79, 125.86, 121.54, 42.46, 22.71. HRMS (ESI-TOF) calculated for $C_{10}H_{12}N_2O_3Na^+$ $[M+Na]^+$: 231.0746, observed 231.0745.



11-i **11-ii**

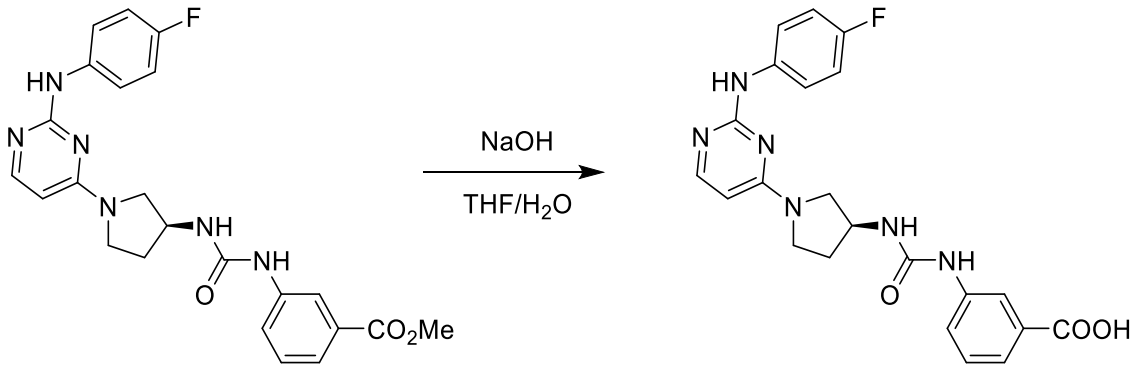
11-i (322 mg, 1.55 mmol) was dissolved in 3 mL of EtOH. Iron powder (743 mg, 7.75 mmol), and acetic acid (1.54 mL, 15.5 mmol) were added and the reaction was stirred at ambient temperature for 2 h. After completion, the solution was filtered, and the filtrate was diluted with EtOAc. The organic layer was washed 3 X with 10% $NaHCO_3$, followed by brine and dried over $MgSO_4$. After filtration, the solvent was removed under reduced pressure to afford **11-ii** as a white solid (80 mg, 29%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.95 (d, *J* = 7.9 Hz, 1H), 7.05 (t, *J* = 7.8 Hz, 1H), 7.00 (t, *J* = 2.0 Hz, 1H), 6.97 – 6.90 (m, 1H), 6.66 (dd, *J* = 7.9, 2.3 Hz, 1H), 4.14 – 3.99 (m, 1H), 1.13 (d, *J* = 6.6 Hz, 6H). ^{13}C NMR

(126 MHz, DMSO-d_6) δ 166.15, 148.52, 135.89, 128.44, 116.11, 114.36, 112.85, 40.70, 22.33. HRMS (ESI-TOF) calculated for $C_{10}H_{14}N_2OH^+$ $[M+H]^+$: 179.1140, observed 179.1177.



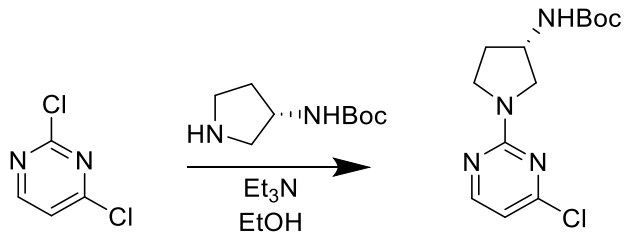
11-ii **(*S*)-ii** **11**

4-Nitrophenylchloroformate (96 mg, 0.48 mmol) was dissolved in 2 mL of anhydrous THF and cooled to 0 °C. Separately, **11-ii** (50 mg, 0.48 mmol) and triethylamine (66 μL, 0.48 mmol) were dissolved in 1 mL of anhydrous THF. The second solution was added dropwise to the first, and the reaction was warmed to ambient temperature for 15 minutes. The mixture was diluted with EtOAc and washed 3 X with H_2O. The organic fraction was washed with brine and dried over $MgSO_4$ followed by filtration. **(*S*)-ii** (76 mg, 0.28 mmol) and triethylamine (100 μL, 0.718 mmol) were added to the EtOAc fraction, and the reaction was stirred at 50 °C temperature for 1 h. The organic layer was washed 3 X with 1M NaOH, followed by brine and $MgSO_4$. Following filtration and concentration, the material was purified via column chromatography (9:1 EtOAc/MeOH) to isolate **11** as a white solid (30 mg, .063 mmol, 23%, 2 steps). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.08 (s, 1H), 8.47 (s, 1H), 8.12 (d, *J* = 7.8 Hz, 1H), 7.93 (d, *J* = 7.5 Hz, 1H), 7.79 (t, *J* = 6.9 Hz, 2H), 7.75 (t, *J* = 1.9 Hz, 1H), 7.54 (dd, *J* = 8.1, 2.1 Hz, 1H), 7.35 (dt, *J* = 7.7, 1.3 Hz, 1H), 7.28 (t, *J* = 7.8 Hz, 1H), 7.06 (t, *J* = 8.9 Hz, 2H), 6.58 (d, *J* = 6.8 Hz, 1H), 5.99 (d, *J* = 5.9 Hz, 1H), 4.33 (s, 1H), 4.13 – 4.01 (m, *J* = 6.7 Hz, 1H), 3.80 – 3.14 (m, 4H), 2.21 (s, 1H), 1.95 (s, 1H), 1.14 (d, *J* = 6.6 Hz, 6H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 165.44, 160.07, 159.08, 156.46 (d, *J* = 237.1 Hz), 155.29, 154.70, 140.05, 137.57, 135.57, 128.27, 119.99, 119.70 (d, *J* = 6.9 Hz), 116.88, 114.61 (d, *J* = 21.9 Hz), 95.57, 51.90, 44.19, 40.78, 22.20. HRMS (ESI-TOF) calculated for $C_{25}H_{28}FN_7O_2H^+$ $[M+H]^+$: 478.2322, observed 478.2357.



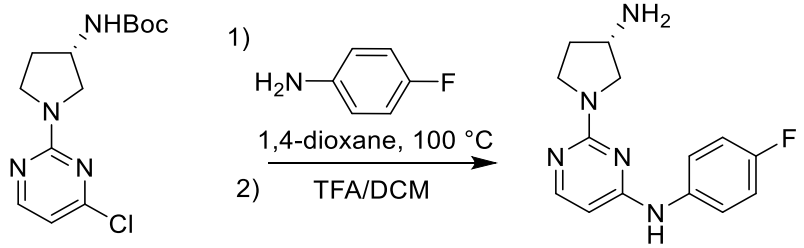
(*S*)-1 12 **(*S*)-1** 12

(*S*)-1 (101.6 mg, 0.372 mmol) was suspended in 5 mL of THF. Next, 1 mL of 1M NaOH was added, and the solution became clear. The reaction proceeded for 18 h, the solvent was removed under a stream of nitrogen and the crude material was dissolved in EtOAc and MeOH. The organic layers were washed 3 X with 1M HCl followed by brine and dried over $MgSO_4$. The material was filtered and concentrated followed by purification on a Combiflash Rf system (EtOAc/MeOH, 0-30% MeOH) to afford **12** as a white solid (12.9 mg, 13%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 12.73 (s, 1H), 10.57 (m, 1H), 9.69 (m, 1H), 8.00 (s, 2H), 7.75 – 7.64 (m, 3H), 7.61 (d, J = 8.2 Hz, 1H), 7.44 (d, J = 7.5 Hz, 1H), 7.30 (t, J = 7.9 Hz, 1H), 7.17 (t, J = 8.7 Hz, 2H), 6.25 (d, J = 6.8 Hz, 1H), 4.40 – 4.25 (m, 1H), 3.81 – 3.30 (m, 5H), 2.29 – 2.14 (m, 1H), 2.00 – 1.84 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 167.36, 159.61, 157.99 (d, J = 239.8 Hz), 155.26, 140.83, 134.89, 131.34, 128.84, 121.80, 121.62 (d, J = 7.8 Hz), 121.39, 118.00, 115.45 (d, J = 22.3 Hz), 96.57, 53.07, 49.07, 48.35, 45.66, 45.46, 31.12, 30.18. HRMS (ESI-TOF) calculated for $C_{22}H_{21}FN_6O_3H^+$ $[M+H]^+$: 437.1693, observed 437.1727



13-i

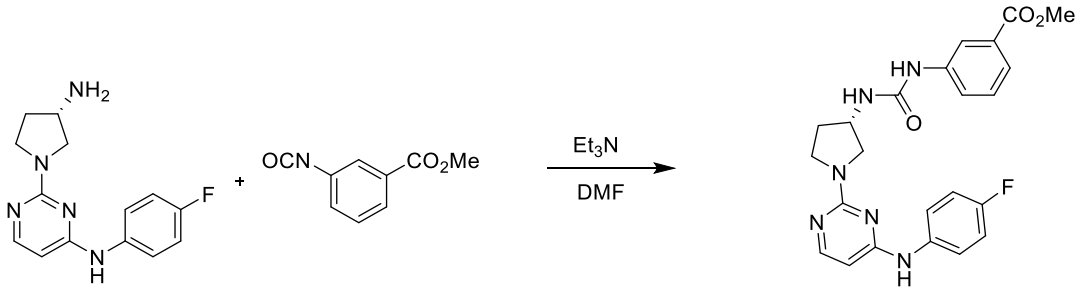
Following the synthetic protocol to form **(*S*)-i**, **13-i** can be isolated as an approximately 20% minor product, observed to be less polar than **(*S*)-i** (R_f of **13-i** = 0.4, 2:1 hexanes/EtOAc). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.28 (d, J = 5.1 Hz, 1H), 7.22 (d, J = 6.5 Hz, 1H), 6.69 (d, J = 5.1 Hz, 1H), 4.07 (p, J = 6.0 Hz, 1H), 3.67 – 3.60 (m, 1H), 3.60 – 3.50 (m, 1H), 3.46 (q, J = 9.0 Hz, 1H), 3.35 – 3.26 (m, 2H), 2.11 (ddd, J = 14.3, 12.8, 6.9 Hz, 1H), 1.86 (dq, J = 12.8, 6.3 Hz, 1H), 1.39 (s, 9H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 159.62, 108.18, 77.77, 51.89, 49.59, 44.72, 44.60, 30.34, 28.10. HRMS (ESI-TOF) calculated for $C_{13}H_{19}ClN_4O_2Na^+$ $[M+Na]^+$: 321.1094, observed 321.1086.



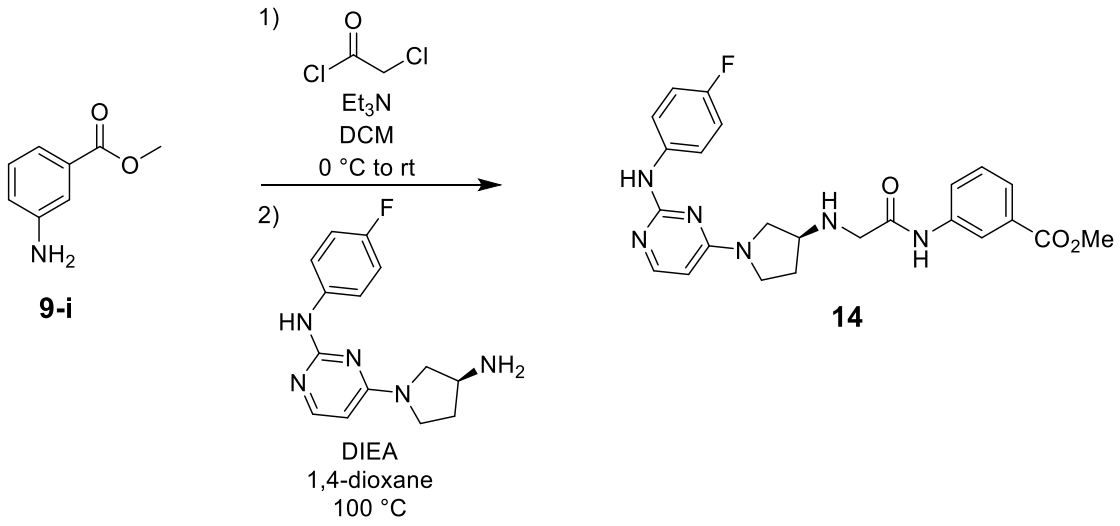
13-i **13-ii**

13-i (100.8 mg, 0.337 mmol) was dissolved in 3 mL of 1,4-dioxane, to which 4-fluoroaniline (48 μL, 0.506 mmol) was added. The solution was heated to reflux for 18 h, followed by removal of solvent. When cooled, the material was subjected to Et_2O and the precipitate was collected via vacuum filtration and washed with subsequent portions of Et_2O. The dried solid was then dissolved in 3 mL of TFA and 1 mL of DCM and stirred for 1h. The solvent was removed, and the crude mixture was dissolved in H_2O and DCM. The aqueous layer was washed 3 X with DCM and separated. The aqueous layer was basified to pH > 10 using 1M NaOH, and then extracted 3 X with EtOAc. The combined organic layers were washed with brine and dried over $MgSO_4$. The material was filtered

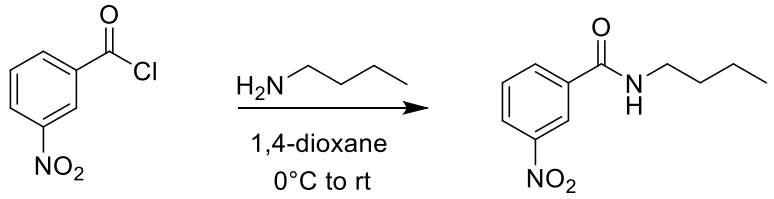
followed by concentration to isolate **13-ii** as a white solid (17.7 mg, 2 steps, 19%) ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.20 (s, 1H), 7.88 (d, J = 5.6 Hz, 1H), 7.76 (dd, J = 9.0, 5.0 Hz, 2H), 7.11 (t, J = 8.7 Hz, 2H), 5.96 (d, J = 5.7 Hz, 1H), 3.67 – 3.55 (m, 2H), 3.55 (s, 1H), 3.50 – 3.41 (m, 1H), 3.19 – 3.12 (m, 1H), 2.02 (dq, J = 13.7, 6.9 Hz, 1H), 1.68 (dq, J = 12.7, 6.3 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 159.97, 158.74 (d, J = 241.0 Hz), 155.86, 137.02 (d, J = 2.3 Hz), 120.41 (d, J = 7.3 Hz), 114.93 (d, J = 21.9 Hz), 95.50, 54.40, 50.54, 44.61, 33.73, 30.31. HRMS (ESI-TOF) calculated for $C_{14}H_{16}FN_5H^+$: $[M+H]^+$: 274.1423, observed 274.1444.



13-ii (135 mg, 0.494 mmol) was dissolved in 3 mL of DMF, followed by triethylamine (0.740 mmol). Methyl 3-isocyanatobenzoate (131 mg, 0.740 mmol) was added and the reaction was stirred for 3 h. The solvent was removed and the reaction was purified via flash chromatography (9:1 EtOAc/MeOH) to afford **13** as a yellow solid (98.9 mg, 44%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.41 (s, 1H), 8.78 (s, 1H), 8.12 (t, J = 1.9 Hz, 1H), 7.90 (d, J = 5.7 Hz, 1H), 7.80 – 7.71 (m, 2H), 7.56 (dd, J = 8.1, 2.1 Hz, 1H), 7.49 (dd, J = 7.7, 1.5 Hz, 1H), 7.36 (t, J = 7.9 Hz, 1H), 7.13 (t, J = 8.7 Hz, 2H), 6.75 (d, J = 6.8 Hz, 1H), 6.04 (d, J = 5.8 Hz, 1H), 4.30 (q, J = 6.0, 5.5 Hz, 1H), 3.83 (s, 3H), 3.70 (dd, J = 11.8, 6.0 Hz, 1H), 3.64 – 3.53 (m, 2H), 3.43 (dd, J = 11.3, 3.9 Hz, 1H), 2.19 (dq, J = 13.6, 7.1 Hz, 1H), 1.92 (dq, J = 12.0, 5.8 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.75, 160.60, 159.64, 157.62 (d, J = 238.3 Hz), 155.28, 141.22, 137.27, 130.52, 129.51, 122.56, 122.23, 121.26 (d, J = 7.5 Hz), 118.44, 115.60 (d, J = 22.1 Hz), 96.65, 52.68, 52.56, 49.58, 44.99, 31.63. HRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_3H^+$ $[M+H]^+$: 451.1849, observed 451.1900.

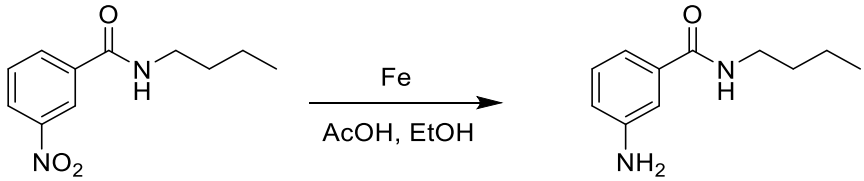


9i (100 mg, 0.66 mmol) was dissolved in 2.5 mL of DCM and cooled to 0 °C followed by addition of triethylamine (101 µL, 0.73 mmol). Separately, chloroacetyl chloride (58 µL, 0.73 mmol) was dissolved in 2 mL of DCM and was added slowly to the first solution. The mixture was warmed to ambient temperature and stirred for 1 h. Upon completion, the reaction was washed 3 X with H_2O, followed by brine and dried over $MgSO_4$. Upon filtration and concentration, and orange solid was isolated and used without further purification. The crude product was dissolved in 3 mL of 1,4-dioxane, followed by addition of N,N-diisopropylethylamine (0.99 mmol) and **(*S*)-ii** (216 mg, 0.79 mmol) and heated to reflux for 18 h. The solvent was removed under reduced pressure, and the product was purified on a Combiflash RF system (9:1 EtOAc/MeOH) to isolate **14** as a brown solid (80 mg, 0.172 mmol, 26%, 2 steps). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.05 (s, 1H), 9.00 (s, 1H), 8.29 (t, J = 2.0 Hz, 1H), 7.88 (d, J = 5.8 Hz, 1H), 7.86 – 7.81 (m, 1H), 7.81 – 7.74 (m, 3H), 7.63 (d, J = 7.9 Hz, 1H), 7.43 (t, J = 7.9 Hz, 1H), 7.04 (t, J = 8.9 Hz, 2H), 6.20 – 5.85 (m, 1H), 3.84 (s, 3H), 3.46-3.19 (m, 4H), 2.09 (s, 1H), 1.98 – 1.84 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 170.49, 165.93, 159.99, 158.26 (d, J = 239.2 Hz), 155.42, 155.35, 138.88, 137.72, 129.97, 129.05, 123.79, 123.54, 119.61, 119.55, 114.54 (d, J = 21.8 Hz), 95.40, 52.06, 51.46, 50.83, 44.43. HRMS (ESI-TOF) calculated for $C_{24}H_{25}FN_6O_3H^+$ $[M+H]^+$: 465.2006, observed 465.2041.



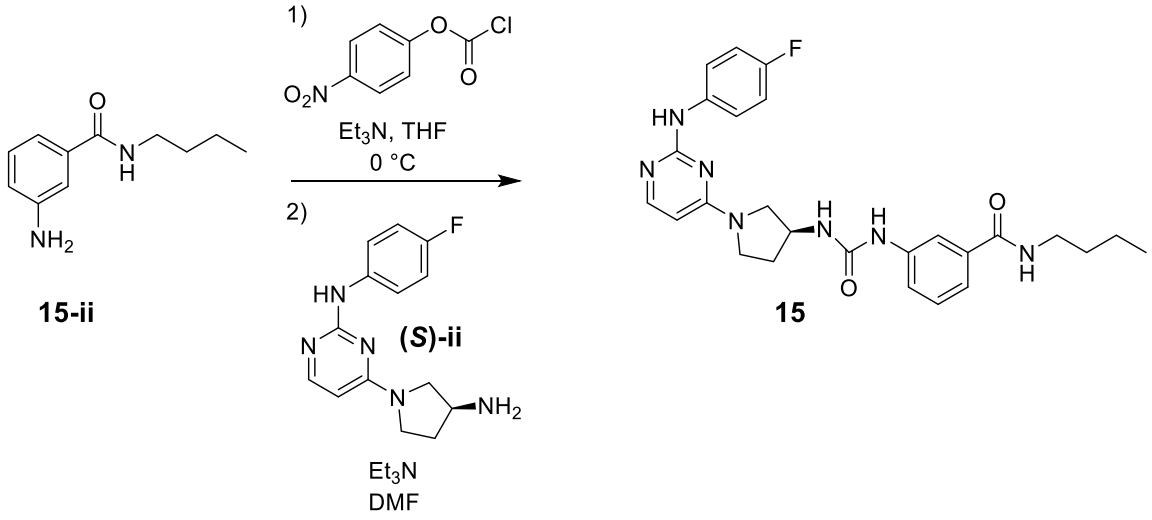
15-i

3-Nitro benzoylchloride (500 mg, 2.7 mmol) was dissolved in 6 mL of 1,4-dioxane and cooled to 0 °C. n-Butylamine (1.35 mL, 13.7 mmol) was added by syringe slowly. The mixture was allowed to room to ambient temperature and stirred for 90 min. The solvent was removed and the crude mixture was dissolved in EtOAc and washed 3 X with H_2O, followed by brine and dried over $MgSO_4$. The solution was filtered and the solvent was removed to isolate **15-i** as a white solid (401 mg, 67%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.56 (t, *J* = 2.0 Hz, 1H), 8.35 (ddt, *J* = 8.2, 1.9, 0.8 Hz, 1H), 8.18 – 8.13 (m, 1H), 7.65 (t, *J* = 7.9 Hz, 1H), 6.22 (s, 1H), 3.50 (td, *J* = 7.2, 5.7 Hz, 2H), 1.64 (p, *J* = 7.5 Hz, 2H), 1.44 (h, *J* = 7.4 Hz, 2H), 0.98 (t, *J* = 7.3 Hz, 3H). LRMS (ESI-TOF) calculated for $C_{11}H_{14}N_2O_3Na^+$ $[M+Na]^+$: 245.1, observed 245.1.

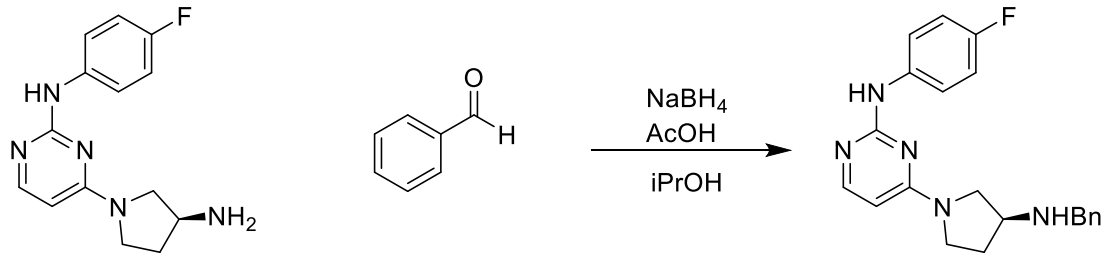


15-i **15-ii**

15-i (562 mg, 2.53 mmol) was dissolved in 5 mL of EtOH, and combined with iron powder (743 mg, 13.3 mmol) and acetic acid (1.54 mL, 26.9 mmol) and stirred for 2 h. After completion, the mixture was filtered and the filtrate was diluted with EtOAc and washed 3 X with 10% $NaHCO_3$ followed by brine and dried over $MgSO_4$. After filtration and removal of solvent **15-ii** was isolated as a white solid (60 mg, 12%). ^{1}H NMR, as a mixture of rotational isomers (400 MHz, Chloroform-*d*) δ 7.21 (s, 0.5 H), 7.19 (s, 0.5 H), 7.17 (d, *J* = 1.4 Hz, 0.5 H), 7.16 (d, *J* = 2.1 Hz, 0.5 H), 7.10 (t, *J* = 1.3 Hz, 0.5 H), 7.09 (t, *J* = 1.3 Hz, 0.5 H), 6.86 (dd, *J* = 2.4, 1.0 Hz, 0.5 H), 6.84 (dd, *J* = 2.4, 1.0 Hz, 0.5 H), 3.39 (t, *J* = 7.2 Hz, 2H), 1.61 – 1.52 (m, 2H), 1.44 – 1.33 (m, 2H), 0.93 (t, *J* = 7.3 Hz, 3H). LRMS (ESI-TOF) calculated for $C_{11}H_{16}N_2OH^+$ $[M+H]^+$: 193.1, observed 193.1.

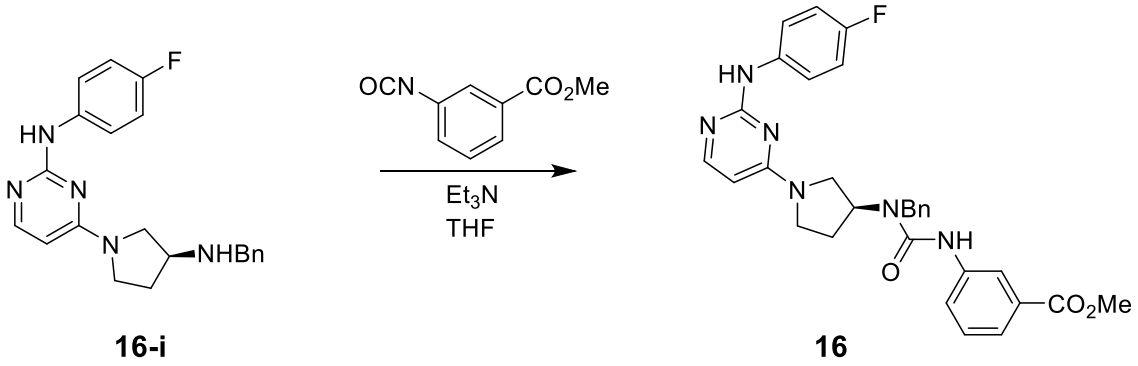


15-ii (40 mg, 0.21 mmol) was dissolved in 1 mL of anhydrous THF with 70 μL Et_3N. Separately, p-nitrophenyl chloroformate (100 mg, 0.50 mmol) was dissolved in 2 mL of anhydrous THF and cooled to 0 °C. The first solution was added dropwise to the second, and stirred for 3 h while warming to rt. The mixture was then diluted into EtOAc and washed with H_2O, followed by brine and $MgSO_4$. After filtration and removal of solvent, the crude mixture was dissolved in 4 mL of DMF, and **(*S*)-ii** (50 mg, 0.18 mmol) and triethylamine (100 μL) were added. The solution was heated for 18 h at 50 °C. The crude mixture was diluted with EtOAc and washed 3 X with 1 M NaOH, followed by brine and dried over $MgSO_4$. The solution was filtered, concentrated, and the material was purified via flash chromatography (9:1 EtOAc/MeOH) to afford **15** as a white solid (16 mg, 16%, 2 steps). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.13 (s, 1H), 8.49 (s, 1H), 8.33 (t, J = 5.7 Hz, 1H), 7.92 (d, J = 5.9 Hz, 1H), 7.81 – 7.73 (m, 3H), 7.54 (ddd, J = 8.0, 2.3, 1.2 Hz, 1H), 7.34 (dt, J = 7.7, 1.4 Hz, 1H), 7.28 (t, J = 7.8 Hz, 1H), 7.12 – 7.03 (m, 2H), 6.61 (d, J = 6.8 Hz, 1H), 6.00 (d, J = 6.0 Hz, 1H), 4.33 (s, 1H), 3.78 – 3.39 (m, 4H), 3.22 (td, J = 7.1, 5.6 Hz, 2H), 2.21 (s, 1H), 1.95 (s, 1H), 1.49 (tt, J = 7.7, 6.3 Hz, 2H), 1.37 – 1.26 (m, 2H), 0.89 (t, J = 7.3 Hz, 3H). LRMS (ESI-TOF) calculated for $C_{26}H_{30}FN_7O_2H^+$ $[M+H]^+$: 492.2, observed 492.2.



(*S*)-ii **16-i**

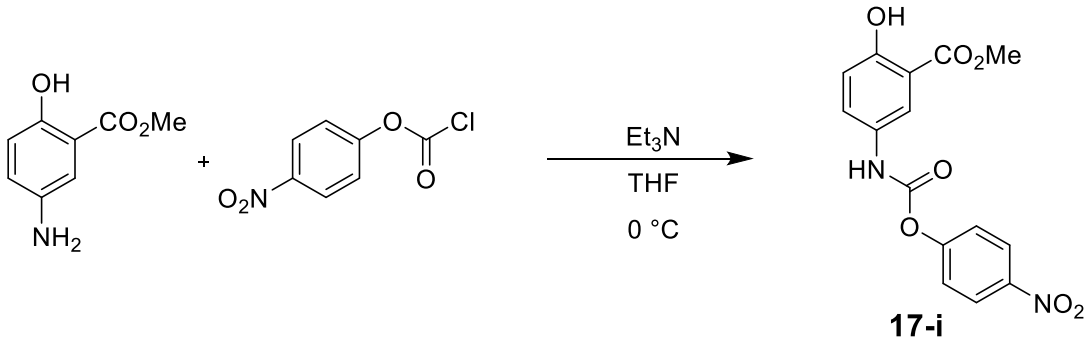
Sodium borohydride (20 mg, 0.529 mmol) was dissolved in 3 mL of isopropanol with 4 Å sieves. The mixture was cooled to 0 °C followed by dropwise addition of acetic acid (91 µL, 1.59 mmol). After bubbling ceases, **(*S*)-ii** (95 mg, 0.348 mmol) and benzaldehyde (35 µL, 0.348 mmol) were added and the mixture was allowed to warm to rt and stirred for 18 h. Upon completion, the mixture was diluted into EtOAc and washed 3 X with 1M NaOH followed by brine and dried over $MgSO_4$. The solution was filtered and the solvent was removed to isolate **16-i** as a clear oil (53 mg, 42%). ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.91 (d, *J* = 6.0 Hz, 1H), 7.60 – 7.53 (m, 2H), 7.34 (d, *J* = 4.4 Hz, 4H), 7.19 (s, 1H), 6.98 (t, *J* = 8.7 Hz, 2H), 5.80 (d, *J* = 5.9 Hz, 1H), 3.85 (s, 2H), 3.66 – 3.11 (m, 2H), 2.19 (dq, *J* = 13.0, 6.4 Hz, 1H), 1.86 (d, *J* = 39.5 Hz, 3H). LRMS (ESI-TOF) calculated for $C_{21}H_{22}FN_5H^+$ $[M+H]^+$: 364.2, observed 364.1.



16-i **16**

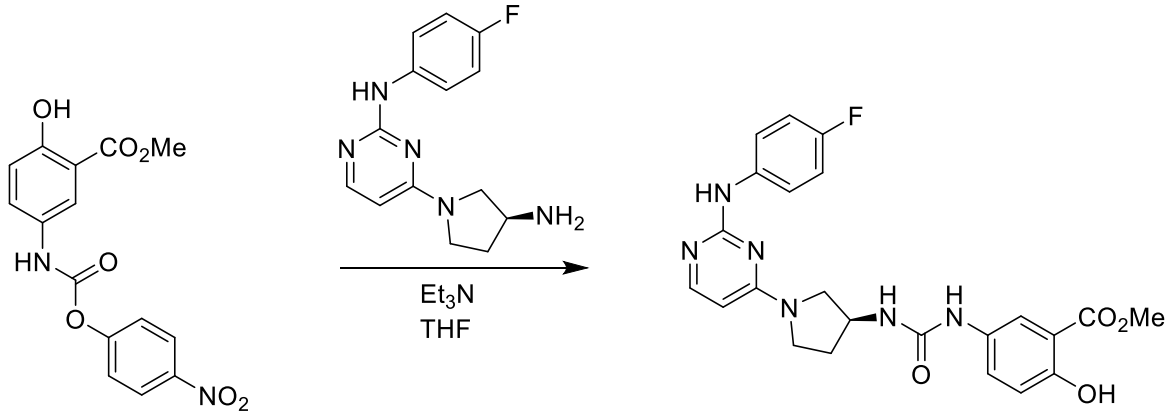
16-i (30 mg, 0.083 mmol) was dissolved in 2 mL of THF followed by addition of methyl-3-isocyanatobenzoate (18 mg, 0.102 mmol) and triethylamine (14 µL, 0.100 mmol). The reaction was stirred for 18 h, followed by concentration and purification on a Combiflash Rf system (DCM/MeOH, 0-10% MeOH) to afford **16** as a white solid (12 mg, 27%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.91 (d, *J* = 5.9 Hz, 1H), 7.71 (t, *J* = 1.9 Hz, 1H), 7.68

(d, *J* = 7.6 Hz, 1H), 7.53 – 7.49 (m, 3H), 7.46 (t, *J* = 7.5 Hz, 2H), 7.39 (t, *J* = 7.4 Hz, 1H), 7.35 – 7.32 (m, 3H), 7.30 (d, *J* = 7.9 Hz, 1H), 7.14 (s, 1H), 6.95 (t, *J* = 8.7 Hz, 2H), 6.40 (s, 1H), 5.80 (d, *J* = 5.9 Hz, 1H), 5.33 – 5.24 (m, 1H), 4.57 (s, 2H), 3.87 (s, 3H), 3.85 (d, *J* = 8.9 Hz, 1H), 3.74 – 3.64 (m, 1H), 3.45 (td, *J* = 10.1, 7.4 Hz, 1H), 3.38 (dd, *J* = 11.1, 8.1 Hz, 1H), 2.34 (dtd, *J* = 14.6, 7.1, 2.7 Hz, 1H), 2.11 (dt, *J* = 12.4, 9.2 Hz, 1H), 1.98 (s, 1H). LRMS (ESI-TOF) calculated for $C_{30}H_{29}FN_6O_3H^+$ $[M+H]^+$: 541.2, observed 541.1



17-i

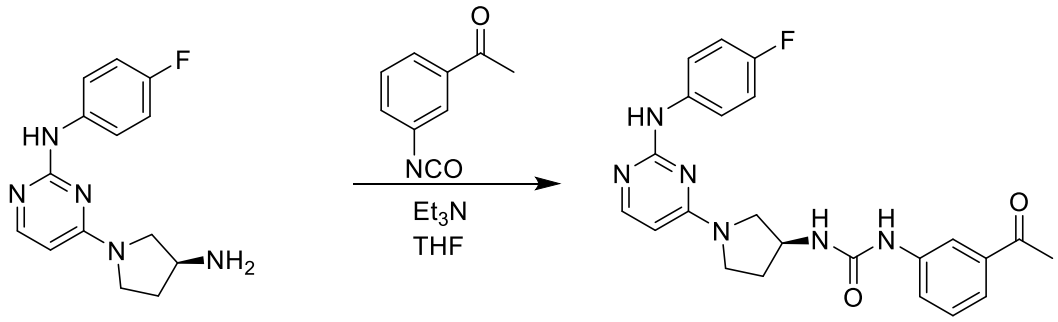
Methyl 5-aminosalicylate (50 mg, 0.299 mmol) was dissolved in 2 mL of THF, followed by addition of triethylamine (60 µL, 0.431 mmol) and cooled to 0 °C. Separately, p-nitrophenylchloroformate (80 mg, 0.397 mmol) was dissolved in THF and cooled to 0 °C. The first solution was added dropwise to the second, followed by warming to rt. The mixture was stirred for 2 h, and taken forward crude to the next step.



17-i **17**

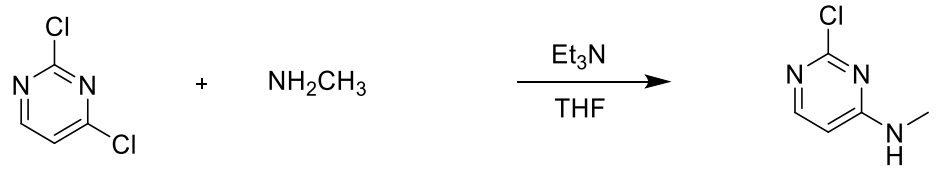
(*S*)-ii (82 mg, 0.300 mmol) was added directly to the crude mixture of **17-i** with an additional 120 µL of triethylamine (3 eq). The solution began to turn yellow in color and the mixture was stirred at rt for 18 h. The solvent was then removed, and the mixture was

dissolved in EtOAc and washed 3 X with 1 M NaOH, followed by brine and drying over $MgSO_4$. Following filtration and concentration, the crude material was purified on a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **17** as a white solid (51 mg, 37%, 2 steps). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.15 (s, 1H), 9.07 (s, 1H), 8.29 (s, 1H), 7.99 (d, *J* = 2.8 Hz, 1H), 7.93 (d, *J* = 5.8 Hz, 1H), 7.80 (dd, *J* = 9.0, 5.0 Hz, 2H), 7.40 (dd, *J* = 8.9, 2.8 Hz, 1H), 7.07 (t, *J* = 8.9 Hz, 2H), 6.89 (d, *J* = 8.9 Hz, 1H), 6.49 (d, *J* = 6.8 Hz, 1H), 5.98 (d, *J* = 5.8 Hz, 1H), 4.31 (s, 1H), 3.89 (s, 3H), 3.77-3.14 (m, 4H), 2.20 (m, 1H), 1.94 (m, 1H). LRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_4H^+$ $[M+H]^+$: 467.2, observed 467.1



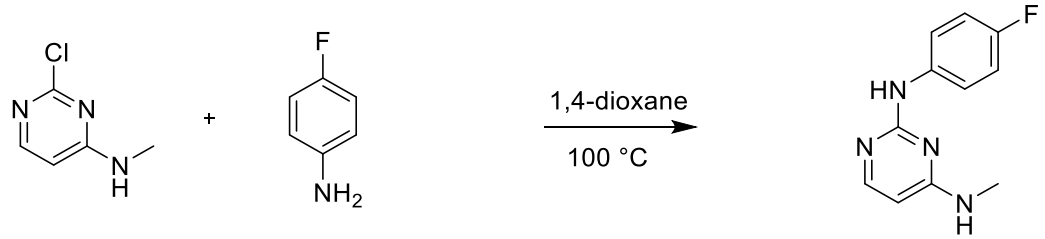
(*S*)-ii **18**

(*S*)-ii (30 mg, 0.110 mmol) was dissolved in 2 mL of THF, followed by addition of 3-acetylphenyl isocyanate (27 mg, 0.165 mmol) and triethylamine (23 μL, 0.165 mmol). The reaction was stirred for 2 h at rt, producing a white precipitate. The mixture was dissolved in DCM/MeOH and purified on a Combiflash Rf system (DCM/MeOH, 0-10% MeOH) to afford **18** as a white solid (26.2 mg, 55%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.08 (s, 1H), 8.57 (s, 1H), 8.01 (t, *J* = 2.0 Hz, 1H), 7.93 (d, *J* = 5.9 Hz, 1H), 7.85 – 7.74 (m, 2H), 7.61 (ddd, *J* = 8.1, 2.3, 1.1 Hz, 1H), 7.52 (dt, *J* = 7.7, 1.4 Hz, 1H), 7.38 (t, *J* = 7.9 Hz, 1H), 7.07 (t, *J* = 8.9 Hz, 2H), 6.62 (d, *J* = 6.8 Hz, 1H), 5.99 (d, *J* = 5.9 Hz, 1H), 4.34 (s, 1H), 3.79-3.18 (m, 4H), 2.54 (s, 3H), 2.22 (m, 1H), 1.97 (m, 1H). LRMS (ESI-TOF) calculated for $C_{23}H_{23}FN_6O_2H^+$ $[M+H]^+$: 435.2, observed 435.4.



F-i

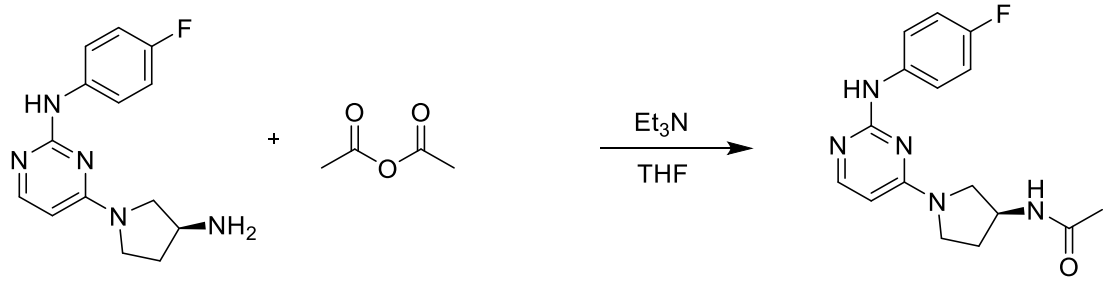
2,4-Dichloropyrimidine (500.2 mg, 3.36 mmol) was dissolved in 10 mL of THF, followed by addition of triethylamine (702 µL, 5.04 mmol). Methylamine (2 mL of a solution 2M in THF, 4.03 mmol) was added dropwise yielding a turbid mixture. The reaction was stirred at ambient temperature for 18 h, followed by removal of solvent. The crude material was redissolved in EtOAc and washed 3 X with H_2O and brine, and dried over $MgSO_4$. The material was filtered and the solvent was removed followed by purification on a Combiflash Rf system (Hexanes/EtOAc, 0-100% EtOAc) to afford **F-i** as a white solid (172.1 mg, 36%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.86 (d, J = 5.9 Hz, 1H), 6.42 (d, J = 5.9 Hz, 1H), 2.78 (d, J = 4.7 Hz, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 163.78, 157.68, 154.97, 104.93, 26.82. HRMS (ESI-TOF) calculated for $C_5H_6ClN_3H^+$ $[M+H]^+$: 144.0329, observed 144.0322.



F-i **F1**

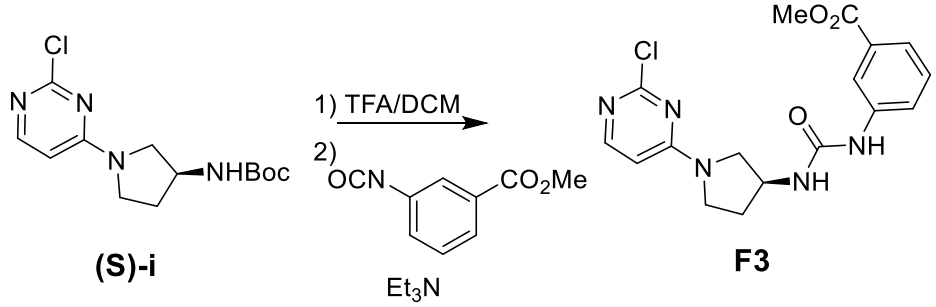
F-i (67.1 mg, 0.467 mmol) was dissolved in 3 mL of 1,4-dioxane followed by addition of 4-fluoroaniline (66.4 µL, 0.701 mmol). The solution was heated to reflux for 12 h, followed by removal of solvent. The crude material was dissolved in EtOAc and washed with 1 M NaOH followed by brine and dried over $MgSO_4$. Following filtration, the solvent was removed and the material was purified on a Combiflash Rf system (Hexanes/EtOAc, 0-100% EtOAc) to afford **F1** as a tan solid (77.6 mg, 76%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.97 (s, 1H), 7.83 – 7.74 (m, 2H), 7.10 (s, 1H), 7.04 (t, J = 8.9 Hz, 2H), 5.91 (d, J = 5.7 Hz, 1H), 2.88 – 2.74 (m, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 162.97, 159.59, 156.37

(d, *J* = 236.6 Hz), 137.77, 137.76, 119.66 (d, *J* = 7.3 Hz), 114.52 (d, *J* = 21.8 Hz), 26.97. HRMS (ESI-TOF) calculated for $C_{11}H_{11}FN_4H^+$ $[M+H]^+$: 219.1001, observed 219.1038.



(*S*)-ii **F2**

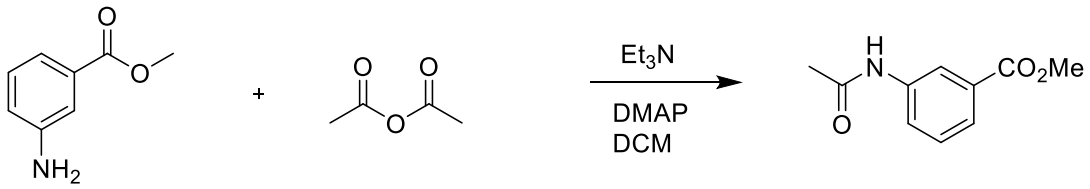
(*S*)-ii (30 mg, 0.11 mmol) was dissolved in 2 mL of THF, followed by addition of triethylamine (17 µL, 0.12 mmol) and acetic anhydride (12 µL, 0.12 mmol) and the solution was stirred at ambient temperature for 18 h. The solvent was removed, and the crude material was redissolved in DCM and washed with 0.1 M HCl, followed by brine and dried over $MgSO_4$. Following filtration, the solvent was removed under reduced pressure to afford **F2** as a white solid (31 mg, 90%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.40 (s, 1H), 8.17 (d, *J* = 6.7 Hz, 1H), 7.91 (d, *J* = 6.2 Hz, 1H), 7.73 (dd, *J* = 9.0, 4.9 Hz, 2H), 7.11 (t, *J* = 8.9 Hz, 2H), 6.05 (d, *J* = 6.4 Hz, 1H), 4.34 (s, 1H), 3.63 (s, 2H), 3.46-3.15 (m, 2H), 2.16 (d, *J* = 17.7 Hz, 1H), 1.94 – 1.86 (m, 1H), 1.81 (s, 3H). ^{13}C NMR, broadening prevented detection of some ^{13}C resonances (126 MHz, DMSO-d_6) δ 169.05, 159.78, 157.05 (d, *J* = 237.5 Hz), 136.48, 120.65 (d, *J* = 7.6 Hz), 114.86 (d, *J* = 21.9 Hz), 95.94, 51.71, 44.60, 22.47. HRMS (ESI-TOF) calculated for $C_{16}H_{18}FN_5OH^+$ $[M+H]^+$: 316.1529, observed 316.1565.



(S)-i **F3**

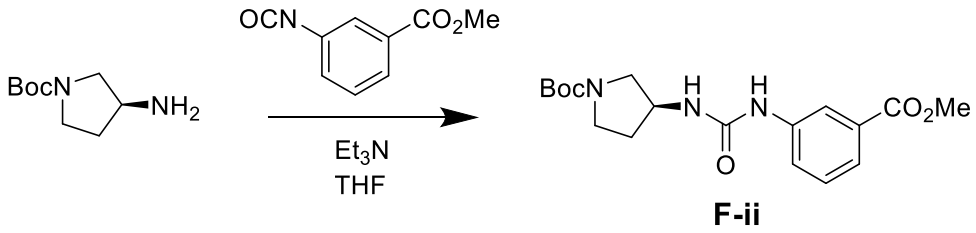
448 mg (1.49 mmol) of **(*S*)-i** was stirred in 3 mL of TFA in 1 mL of DCM for 1h. The TFA was removed and the solution was neutralized with 2 mL of triethylamine in 15 mL of

DCM. 532 mg (3.00 mmol) of methyl 3-isocyanobenzoate was added and stirred for 12 h. The reaction mixture was diluted with 10 mL of water and extracted 3 X with 15 mL of EtOAc. The solvent was removed and **F3** was isolated with silica chromatography (hexanes:ethyl acetate), 77% yield. ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.62 (d, J = 9.9 Hz, 1H), 8.25 (s, 1H), 8.14 (s, 1H), 8.12 – 8.01 (m, 1H), 7.72 (d, J = 8.2 Hz, 1H), 7.68 (d, J = 7.8 Hz, 1H), 7.46 (t, J = 7.9 Hz, 1H), 6.56 (s, 1H), 4.87 (d, J = 62.8 Hz, 1H), 3.85 (s, 3H), 3.81 – 3.72 (m, 1H), 3.66 – 3.46 (m, 3H), 2.39 – 2.22 (m, 1H), 2.19 – 2.02 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 181.13, 166.45, 161.28, 159.90, 156.94, 140.49, 130.23, 129.35, 124.95, 103.71, 103.52, 52.69, 52.09, 45.26, 30.70, 30.21. HRMS (ESI-TOF) calculated for $C_{17}H_{18}ClN_5O_3H^+$ $[M+H]^+$: 376.1171, observed 376.1168.



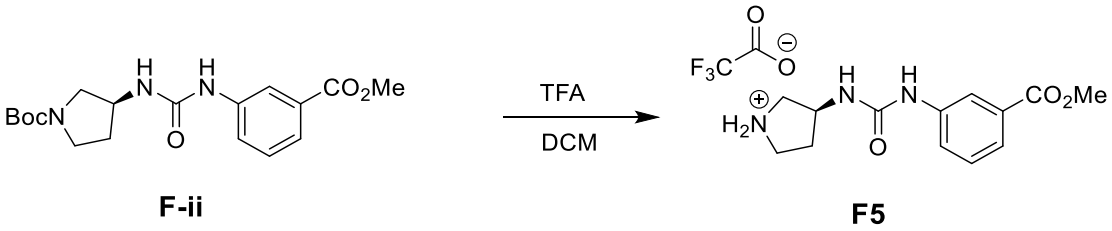
9-i **F4**

9-i (94.8 mg, 0.627 mmol) was dissolved in 3 mL of DCM. Acetic anhydride (88.7 µL, 0.941 mmol, 1), triethylamine (131.2 µL, 0.941 mmol), and N,N-dimethylaminopyridine (5.5 mg, 0.045 mmol) were added. The reaction was stirred at ambient temperature for 18 h. Upon completion, the solution was added to a separatory funnel and washed with 1 M HCl, followed by brine and dried over $MgSO_4$. The material filtered, concentrated, and then was purified on a Combiflash Rf system (Hexanes/EtOAc, 0-100% EtOAc) to afford **F4** as a white solid (39.8 mg, 33%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 10.15 (s, 1H), 8.25 (t, J = 1.9 Hz, 1H), 7.85 – 7.80 (m, 1H), 7.62 (dt, J = 7.8, 1.4 Hz, 1H), 7.44 (t, J = 7.9 Hz, 1H), 3.85 (s, 3H), 2.06 (s, 3H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 168.48, 165.99, 139.55, 129.93, 129.05, 123.47, 123.26, 119.28, 52.06, 23.90. HRMS (ESI-TOF) calculated for $C_{10}H_{11}NO_3H^+$ $[M+H]^+$: 194.0772, observed 194.0810.



F-ii

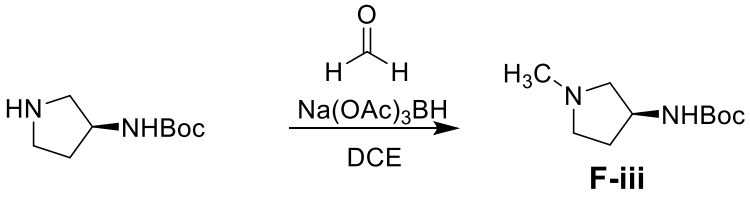
(*S*)-1-boc-3-aminopyrrolidine (100 mg, 0.54 mmol) was dissolved in 3 mL of THF, followed by addition of triethylamine (90 µL, 0.65 mmol) and methyl 3-isocyanatobenzoate (114 mg, 0.65 mmol). The reaction was stirred at ambient temperature for 18 h, followed by removal of solvent and purification on a combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **F-ii** as a clear oil (96 mg, 49%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.59 (d, *J* = 3.3 Hz, 1H), 8.12 (s, 1H), 7.54 (dd, *J* = 8.1, 2.2 Hz, 1H), 7.52 – 7.47 (m, 1H), 7.36 (t, *J* = 7.9 Hz, 1H), 6.49 (t, *J* = 5.3 Hz, 1H), 4.14 (p, *J* = 5.6 Hz, 1H), 3.84 (s, 3H), 3.46 (ddd, *J* = 17.7, 10.8, 5.9 Hz, 1H), 3.30 (s, 1H: solvent obscuring second H), 3.09 (dd, *J* = 10.9, 4.3 Hz, 1H), 2.05 (tq, *J* = 13.3, 7.1, 5.2 Hz, 1H), 1.78 (dh, *J* = 12.7, 6.1 Hz, 1H), 1.40 (s, 9H). ^{13}C NMR, mixture of rotational isomers (126 MHz, DMSO-d_6) δ 166.16, 154.64, 153.45, 140.53, 129.95, 128.95, 122.09, 121.75, 117.97, 78.25, 51.99, 51.60, 51.15, 49.26, 48.48, 43.76, 43.54, 31.18, 30.21, 28.06. HRMS (ESI-TOF) calculated for $C_{18}H_{25}N_3O_5Na^+$ $[M+Na]^+$: 386.1692, observed 386.1683.



F-ii **F5**

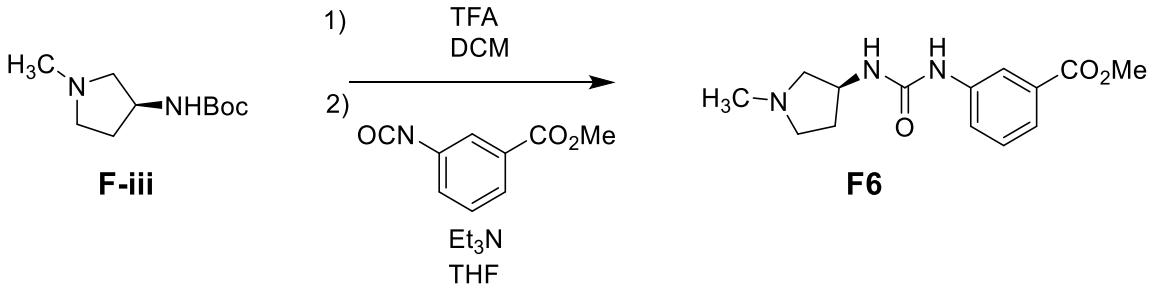
F-ii (569.8 mg, 1.565 mmol) was dissolved in 3 mL of TFA and 1 mL of DCM. The reaction was stirred at ambient temperature for 1 h. After completion, the solvent was removed and the product was precipitated in diethyl ether. The solid was collected by filtration and was dried under vacuum to afford **F5** as a trifluoroacetate salt (503.7 mg, 85%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 9.31 (s, 1H), 9.03 – 8.86 (m, 2H), 8.18 (t, *J* = 2.0 Hz, 1H), 7.58 (dd, *J* = 8.1, 2.1 Hz, 1H), 7.50 (dd, *J* = 7.8, 1.4 Hz, 1H), 7.37 (t, *J* = 7.9 Hz,

1H), 7.19 (d, *J* = 6.0 Hz, 1H), 4.26 (h, *J* = 6.1 Hz, 1H), 3.83 (s, 3H), 3.38 (dq, *J* = 11.9, 6.1 Hz, 1H), 3.31 (h, *J* = 5.9 Hz, 1H), 3.23 (hept, *J* = 6.1 Hz, 1H), 3.09 (dq, *J* = 11.2, 5.4 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.32, 155.08, 140.77, 130.07, 129.03, 122.33, 121.96, 118.22, 52.10, 49.79, 48.91, 43.83, 29.99. HRMS (ESI-TOF) calculated for $C_{13}H_{17}N_3O_3H^+$ $[M+H]^+$: 264.1303, observed 264.1341.



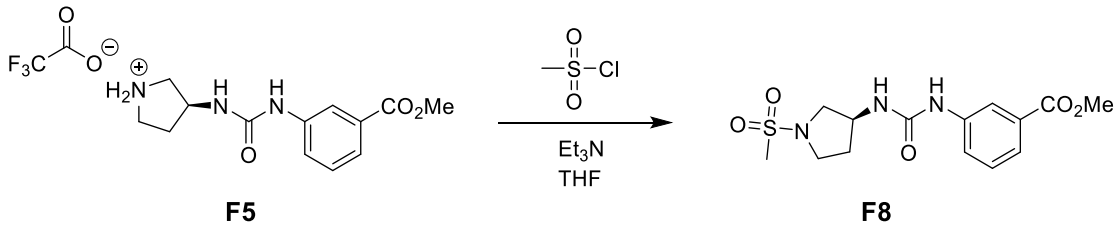
F-iii

(*S*)-3-(Boc-amino)pyrrolidine (312.7 mg, 1.679 mmol) was dissolved in 5 mL of 1,2-dichloroethane. Formaldehyde (185 μL of a 37% solution in H_2O/MeOH, 6.71 mmol) was added and the mixture was stirred at ambient temperature for 20 min. Sodium triacetoxyborohydride (725.0 mg, 3.421 mmol) was then added over 10 min and the mixture was stirred at ambient temperature for 18 h. The reaction was quenched with saturated NH_4Cl and diluted with H_2O at pH 5. The aqueous layer was washed 3 X with DCM and separated. The pH was adjusted to > 10 using 1M NaOH and extracted 3 X into DCM. The combined organic layers were washed with brine and dried over $MgSO_4$. After filtration and removal of solvent **F-iii** was isolated as a white solid (257.0 mg, 1.283 mmol, 77%). ^{1}H NMR (500 MHz, Methanol-d_4) δ 4.09 (q, *J* = 6.6, 5.8 Hz, 1H), 2.82 (dd, *J* = 10.0, 7.2 Hz, 1H), 2.67 (td, *J* = 8.8, 6.3 Hz, 1H), 2.54 (q, *J* = 8.0 Hz, 1H), 2.40 (dd, *J* = 10.1, 5.3 Hz, 1H), 2.36 (s, 3H), 2.24 (dtd, *J* = 14.1, 8.4, 6.2 Hz, 1H), 1.69 – 1.62 (m, 1H), 1.45 (s, 9H). ^{13}C NMR, mixture of rotational isomers (126 MHz, Chloroform-*d*) δ 155.40, 155.35, 79.08, 63.27, 63.25, 54.97, 50.35, 50.23, 41.89, 33.17, 33.13, 28.39. HRMS (ESI-TOF) calculated for $C_{10}H_{20}N_2O_2H^+$ $[M+H]^+$: 201.1558, observed 201.1626



F-iii **F6**

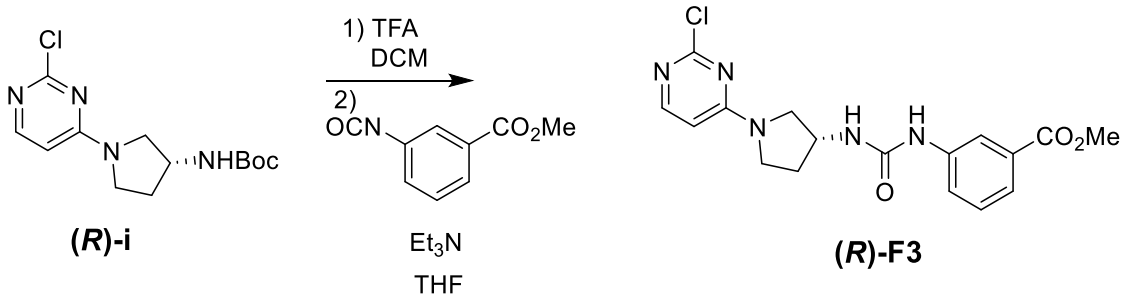
F-iii (257.0 mg, 1.283 mmol) was dissolved in 3 mL of TFA and 1 mL of DCM. The reaction was stirred at ambient temperature for 90 min and the solvent was removed. The crude material was then dissolved in 3 mL of THF, and triethylamine (537 µL, 3.85 mmol) was added slowly and gas was observed. Methyl 3-isocyanatobenzoate (340 mg, 1.92 mmol) was then added and the reaction was stirred for 18 h. The mixture was then diluted with EtOAc and washed 3X with 10% Na_2CO_3 followed by brine and dried over $MgSO_4$. Following filtration, the solvent was removed and the material was purified on a Combiflash Rf system (EtOAc/MeOH, 0-10% MeOH) to afford **F6** as a white solid (177.9 mg, 0.6415 mmol, 50%, 2 steps). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.94 (d, *J* = 2.3 Hz, 1H), 7.65 (dd, *J* = 7.8, 1.7 Hz, 2H), 7.31 (t, *J* = 7.9 Hz, 1H), 6.34 (d, *J* = 7.8 Hz, 1H), 4.24 (s, 1H), 3.85 (s, 3H), 3.37 (s, 1H), 3.01 (t, *J* = 9.2 Hz, 1H), 2.85 (d, *J* = 10.3 Hz, 1H), 2.47 (dd, *J* = 10.2, 6.7 Hz, 1H), 2.39 (s, 3H), 2.32 (ddt, *J* = 12.4, 8.7, 4.4 Hz, 1H), 2.20 (q, *J* = 8.8 Hz, 1H), 1.86 – 1.75 (m, 1H). ^{13}C NMR (126 MHz, Chloroform-*d*) δ 167.04, 155.66, 139.84, 130.76, 129.04, 123.94, 123.67, 120.25, 62.94, 55.14, 52.12, 50.28, 41.59, 32.98. HRMS (ESI-TOF) calculated for $C_{14}H_{19}N_3O_3H^+$ $[M+H]^+$: 278.1460, observed 278.1495.



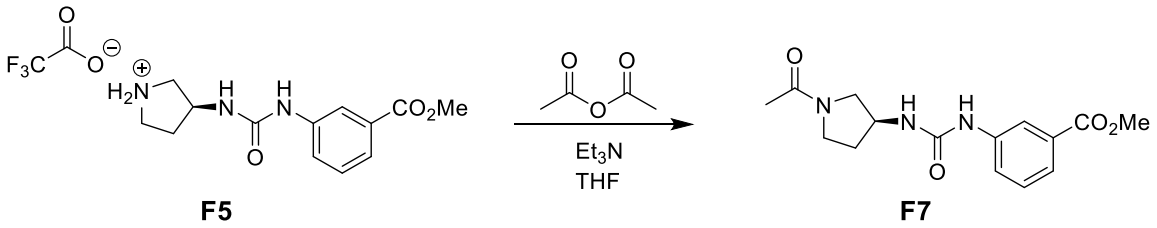
F5 **F8**

F5 (100.8 mg, 0.267 mmol) was dissolved in 3 mL of THF, followed by addition of triethylamine (185 µL, 1.33 mmol). Mesyl chloride (30.8 µL, 0.398 mmol) was then added and the reaction stirred at rt for 3 h. The solvent was removed and the mixture was

dissolved in EtOAc and washed 3 X with H_2O followed by brine and dried over $MgSO_4$. The solution was filtered, concentrated and was purified on a Combiflash Rf system (hexanes/EtOAc, 0-100% EtOAc) to afford **F8** as a white solid (68.1 mg, 75%). ^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.88 (t, *J* = 2.0 Hz, 1H), 7.68 (tt, *J* = 7.9, 7.5, 3.4, 1.7 Hz, 2H), 7.34 (t, *J* = 7.9 Hz, 1H), 7.12 (s, 1H), 5.53 (d, *J* = 7.0 Hz, 1H), 4.51 (qt, *J* = 6.5, 3.7 Hz, 1H), 3.89 (s, 3H), 3.57 (dt, *J* = 9.8, 7.6 Hz, 1H), 3.46 (dd, *J* = 10.5, 5.4 Hz, 1H), 3.42 (dd, *J* = 10.5, 3.0 Hz, 1H), 3.33 (td, *J* = 9.3, 4.9 Hz, 1H), 2.91 (s, 3H), 2.25 (td, *J* = 13.9, 7.9 Hz, 1H), 2.01 (ddt, *J* = 12.4, 8.1, 4.5 Hz, 1H). LRMS (ESI-TOF) calculated for $C_{14}H_{19}N_3O_5SNa^+$ $[M+Na]^+$: 364.1, observed 364.3.



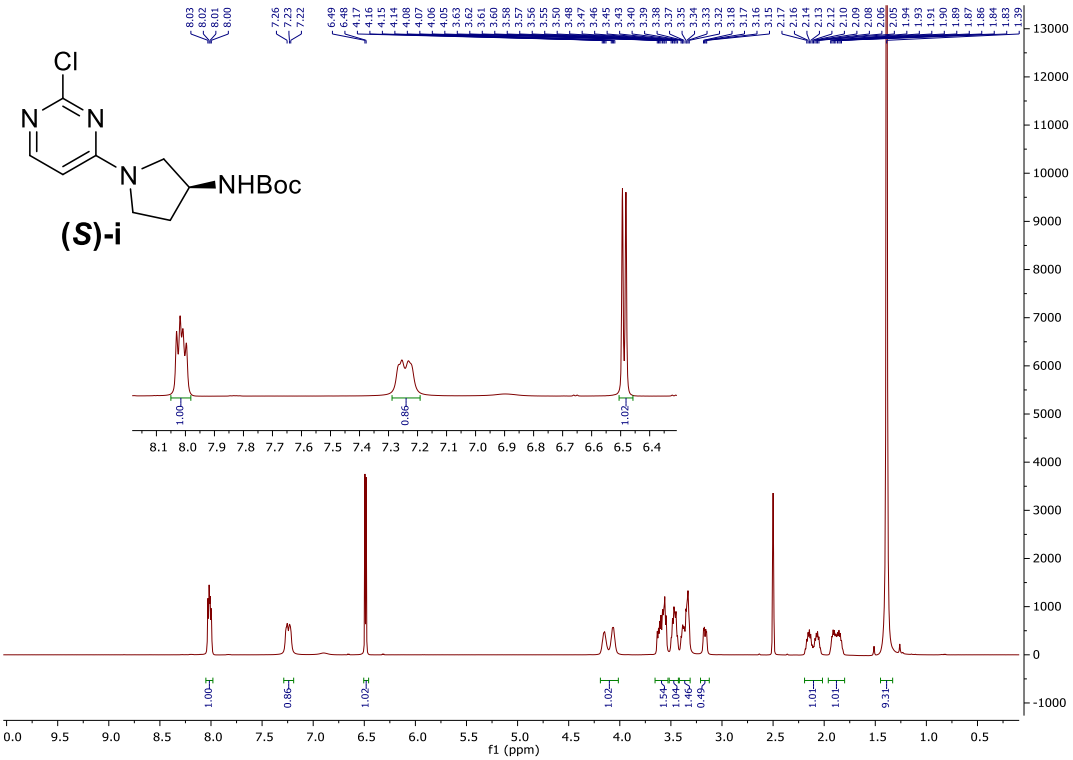
(*R*)-i (125 mg, 0.419 mmol) was dissolved in 1 mL of DCM and 3 mL of TFA. The reaction was stirred at ambient temperature for 1 h, followed by removal of solvent and precipitation in Et_2O. The precipitate was dissolved in 3 mL of THF, and triethylamine (112 μL, 2.10 mmol) was added along with methyl-3-isocyanatobenzoate (28.4 mg, 0.503 mmol) and the reaction was stirred for 16 h. The solvent was removed and the material was purified on a Combiflash Rf system (Hexanes:EtOAc, 0-100% EtOAc) to afford **(*R*)-F3** as a white solid (54.4 mg, 35%, 2 steps). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.61 (d, *J* = 13.8 Hz, 1H), 8.13 (s, 1H), 8.04 (dd, *J* = 14.9, 6.0 Hz, 1H), 7.58 – 7.53 (m, 1H), 7.50 (d, *J* = 7.6 Hz, 1H), 7.37 (t, *J* = 7.9 Hz, 1H), 6.61 (dd, *J* = 14.0, 6.5 Hz, 1H), 6.55 (d, *J* = 6.0 Hz, 1H), 4.41 – 4.21 (m, 1H), 3.84 (s, 3H), 3.67 (ddt, *J* = 21.6, 10.7, 5.9 Hz, 1H), 3.58 (td, *J* = 12.4, 11.7, 5.7 Hz, 1H), 3.46 (q, *J* = 7.6 Hz, 2H), 3.30 – 3.25 (m, 1H), 2.27 – 2.13 (m, 1H), 2.04 – 1.88 (m, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 166.15, 156.35, 140.50, 129.94, 128.95, 122.17, 121.81, 118.05, 103.11, 102.87, 52.20, 52.00, 49.07, 48.54, 44.59, 30.72, 30.11. HRMS (ESI-TOF) calculated for $C_{17}H_{18}ClN_5O_3H^+$ $[M+H]^+$: 376.1176, observed 376.1167.



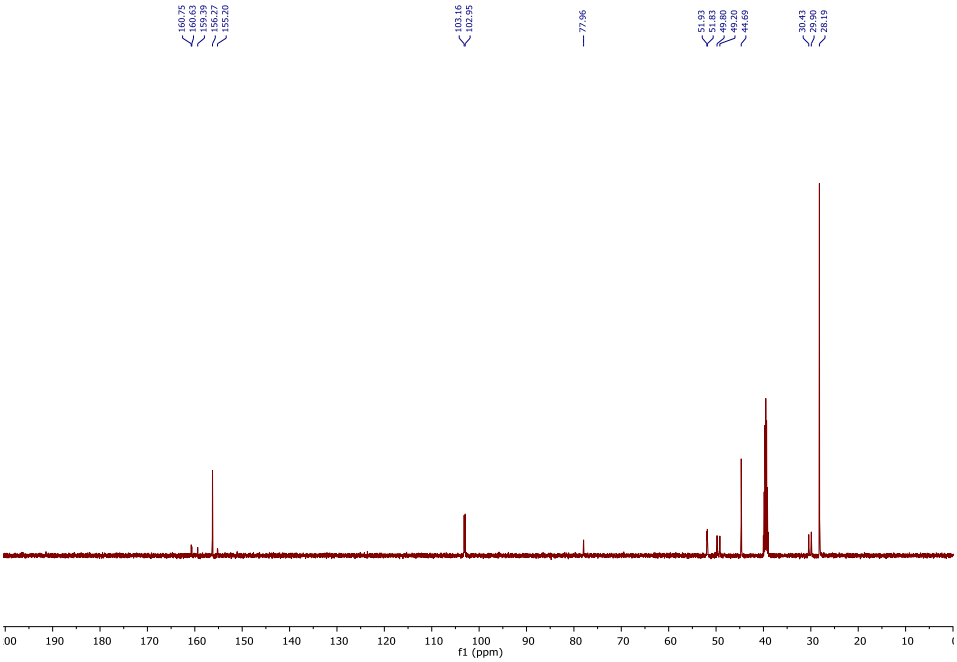
F5 (233.3 mg, 0.618 mmol) was dissolved in 2 mL of THF, followed by addition of triethylamine (617 μL, 3.09 mmol). Acetic anhydride (419 μL, 3.09 mmol) was added, and the mixture was stirred at rt for 3 h. The solvent was removed, and the crude mixture was dissolved in EtOAc and washed 3 X with H_2O, followed by brine and dried over $MgSO_4$. Upon filtration and removal of solvent, the material was purified using a Combiflash Rf system (DCM/MeOH, 0-10% MeOH) to isolate **F7** as a white solid (43.3 mg, 23%).^{1}H NMR as a mixture of rotational isomers (500 MHz, Chloroform-*d*) δ 8.10 (s, 0.5H), 8.09 (s, 0.5H), 7.89 (t, *J* = 1.9 Hz, 0.5H), 7.87 (t, *J* = 1.9 Hz, 0.5H), 7.78 (tdd, *J* = 8.9, 2.4, 1.1 Hz, 1H), 7.65 – 7.63 (m, 0.5H), 7.62 (dd, *J* = 3.1, 1.4 Hz, 0.5H), 7.32 (td, *J* = 7.9, 4.4 Hz, 1H), 6.33 (d, *J* = 7.0 Hz, 0.5H), 6.20 (d, *J* = 6.0 Hz, 0.5H), 4.50 (dtt, *J* = 7.4, 5.0, 2.2 Hz, 0.5H), 4.39 (h, *J* = 5.0 Hz, 0.5H), 3.88 (s, 1.5H), 3.87 (s, 1.5H), 3.67 – 3.35 (m, 4H), 2.25 (dtd, *J* = 13.3, 8.1, 5.2 Hz, 0.5H), 2.20 – 2.11 (m, 0.5H), 2.10 (s, 1.5H), 2.05 (s, 2H), 1.97 – 1.90 (m, 1H). ^{13}C NMR as a mixture of rotational isomers (126 MHz, Chloroform-*d*) δ 170.56, 170.42, 167.13, 155.43, 155.32, 139.76, 130.64, 129.08, 129.05, 123.31, 123.10, 123.06, 119.31, 119.25, 54.45, 52.16 (d, *J* = 2.3 Hz), 51.37, 50.07, 48.92, 45.65, 43.90, 32.72, 30.57, 22.49, 22.19. HRMS (ESI-TOF) calculated for $C_{15}H_{19}N_3O_4Na^+$ $[M+Na]^+$: 328.1274, observed 328.1287.

4.12.13 ^{1}H & ^{13}C NMR spectra:

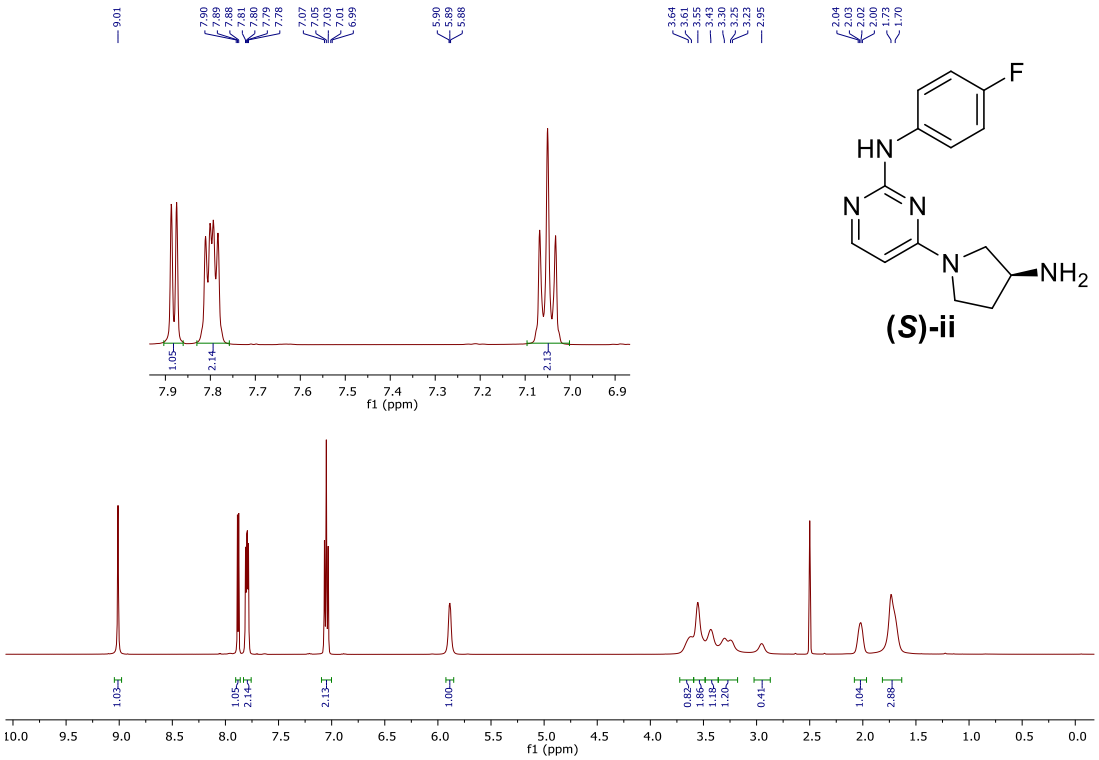
(*S*)-i, ^{1}H, 500 MHz, DMSO-d_6



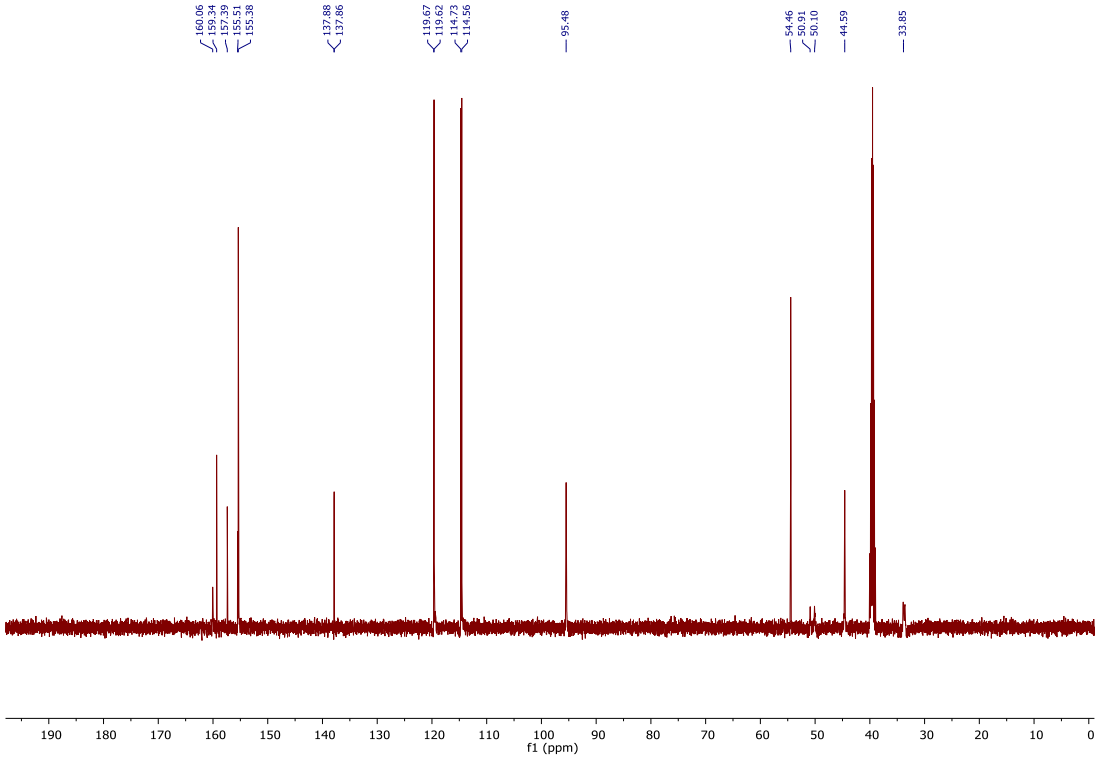
(*S*)-i, ^{13}C, 126 MHz, DMSO-d_6



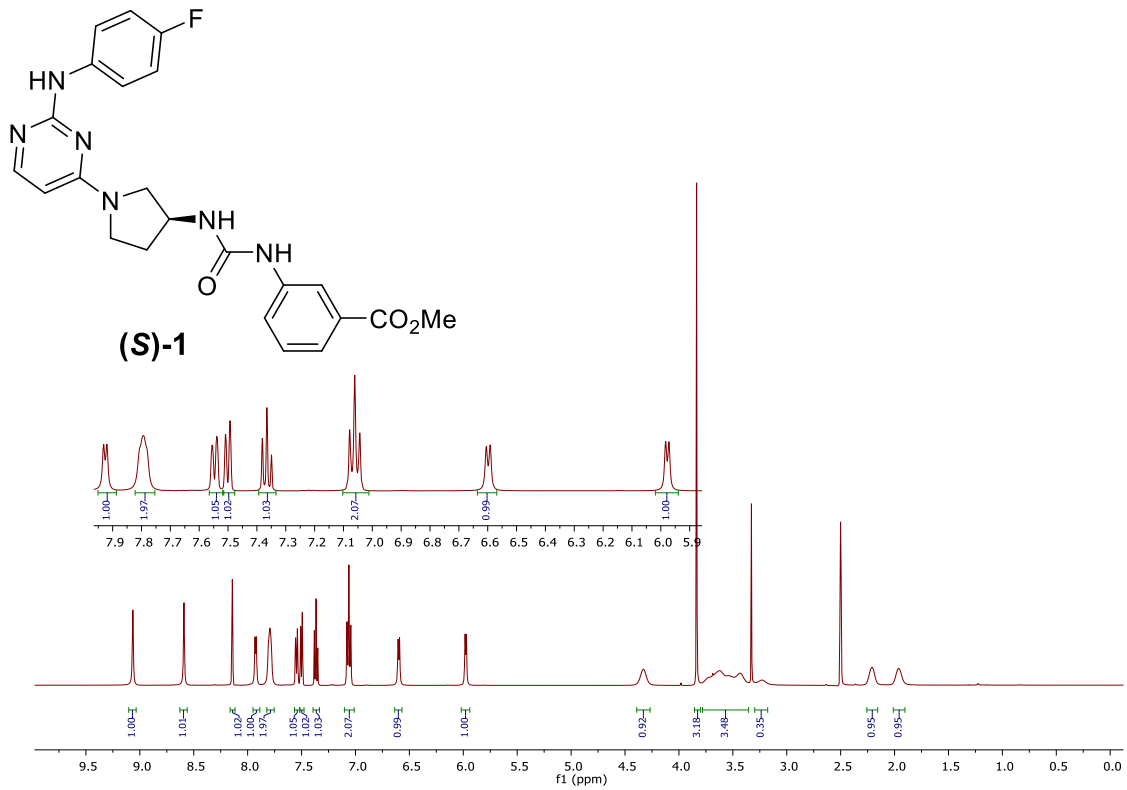
(*S*)-ii, ^{1}H, 500 MHz, DMSO-d_6



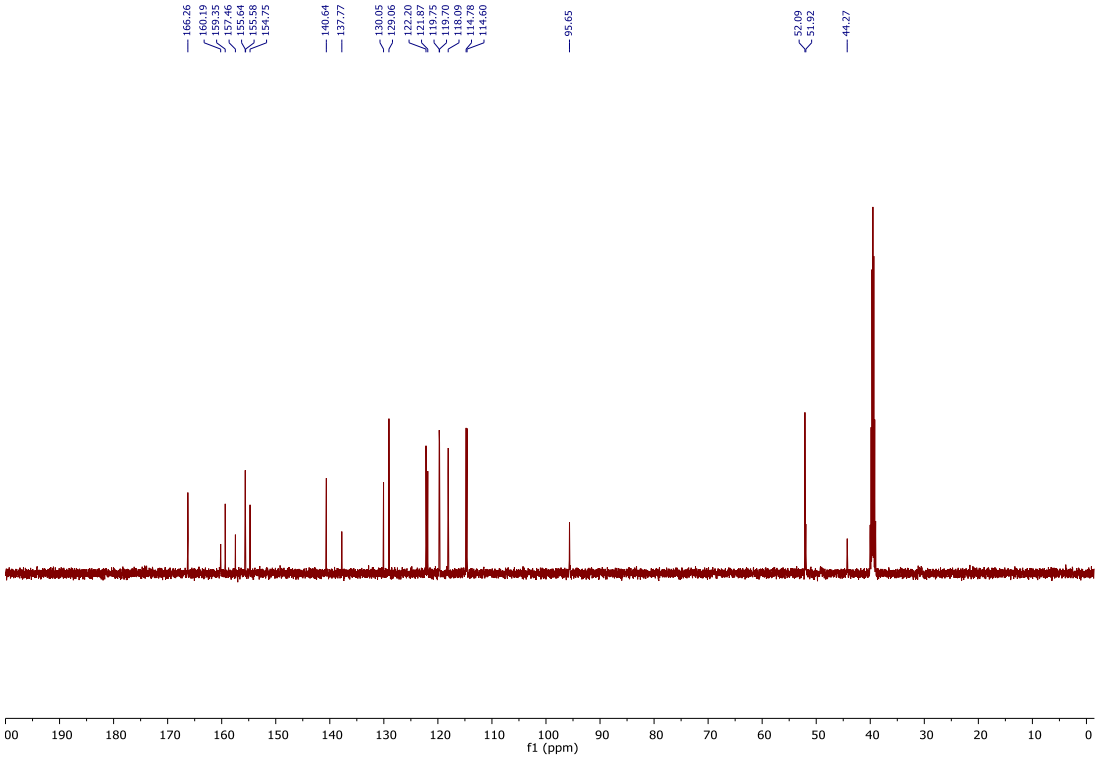
(*S*)-ii, ^{13}C, 126 MHz, DMSO-d_6



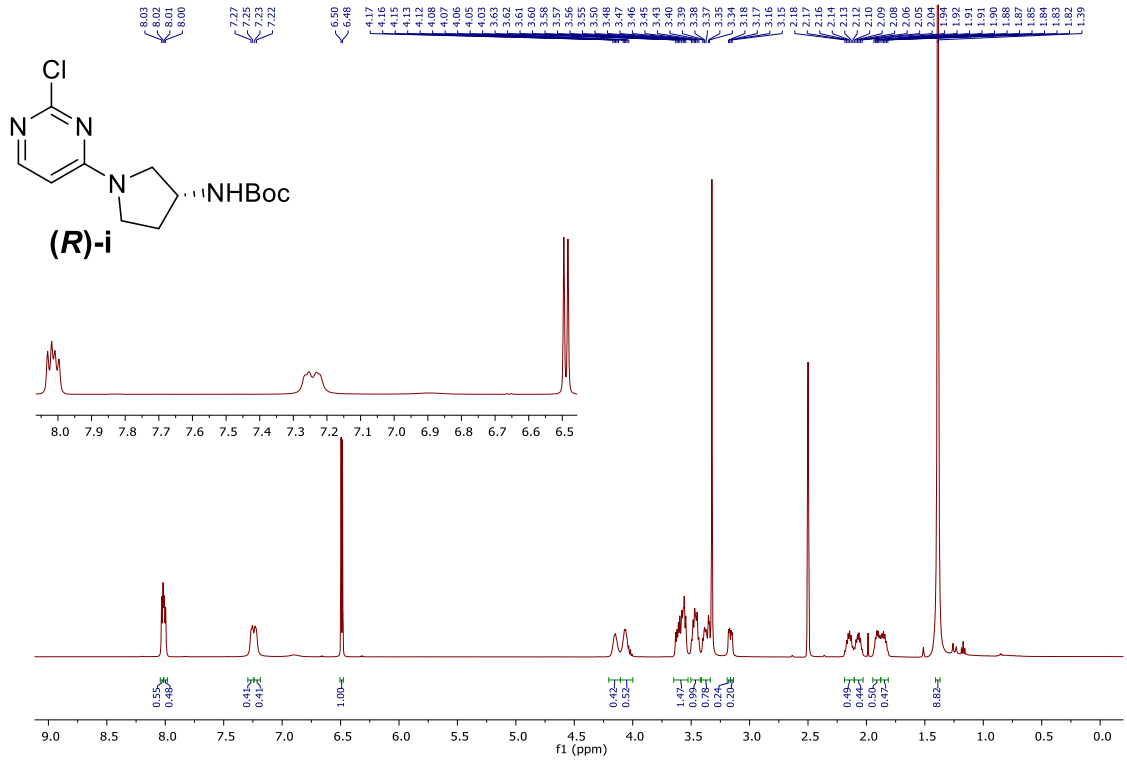
(*S*)-1, ^{1}H, 500 MHz, DMSO-d_6



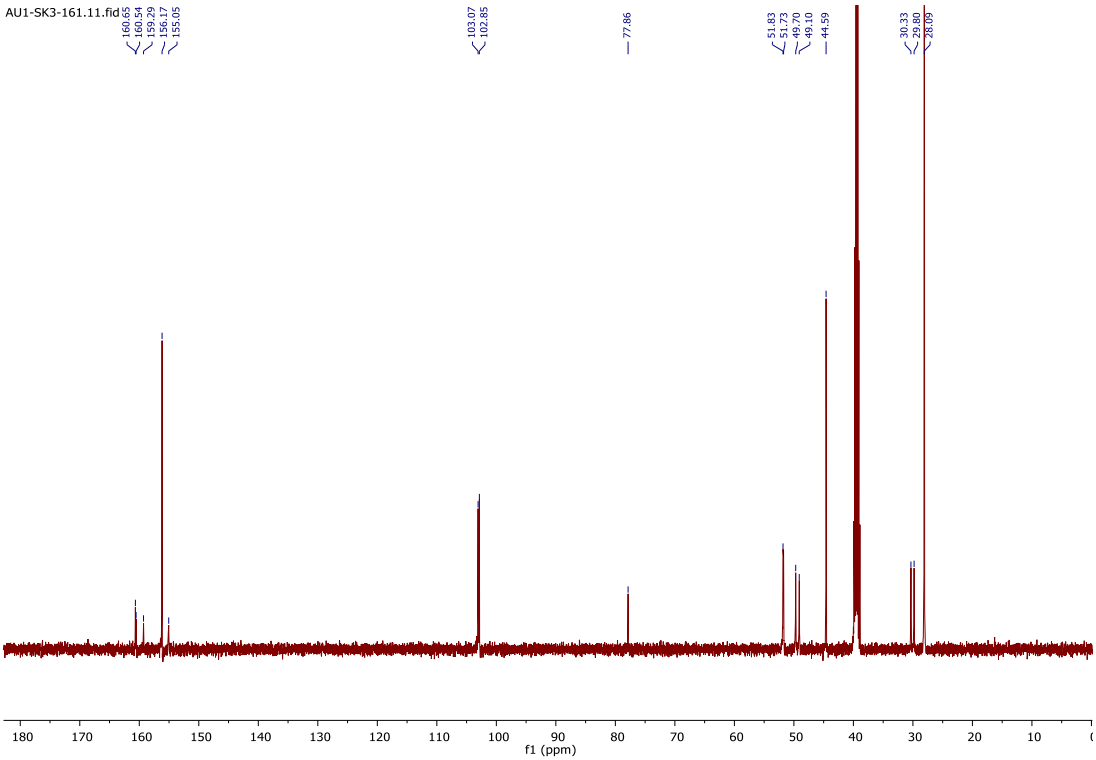
(*S*)-1, ^{13}C, 126 MHz, DMSO-d_6



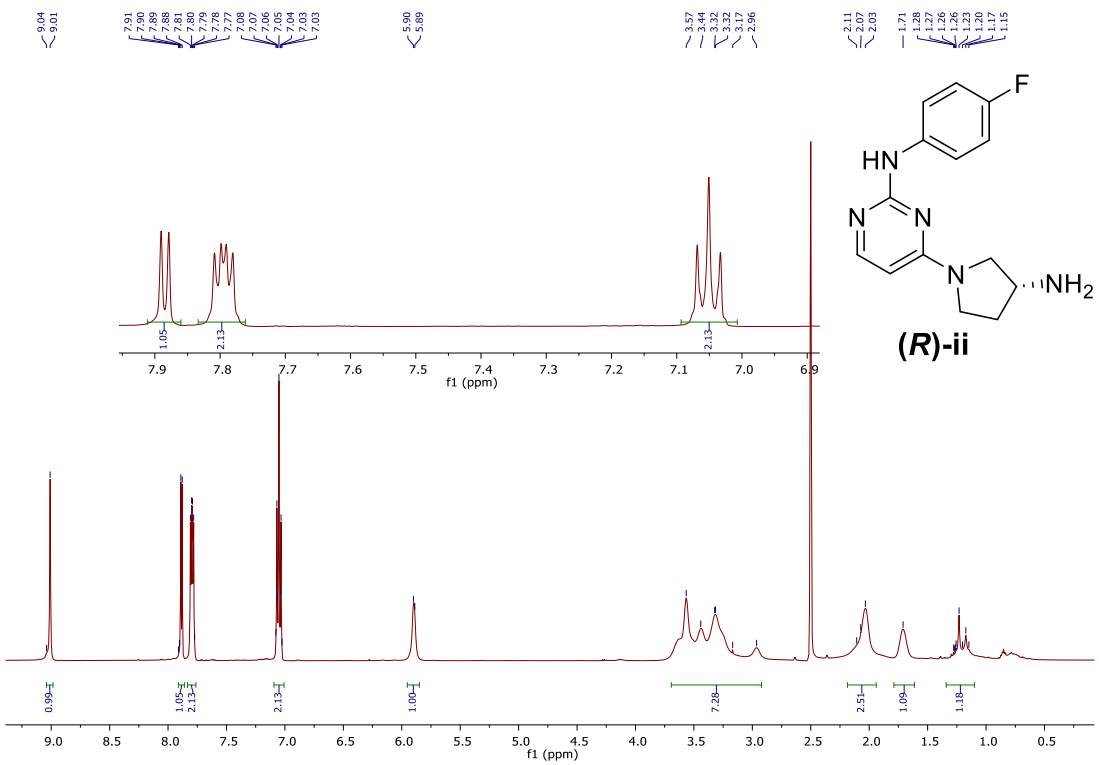
(*R*)-i, ^{1}H, 500 MHz, DMSO-d_6



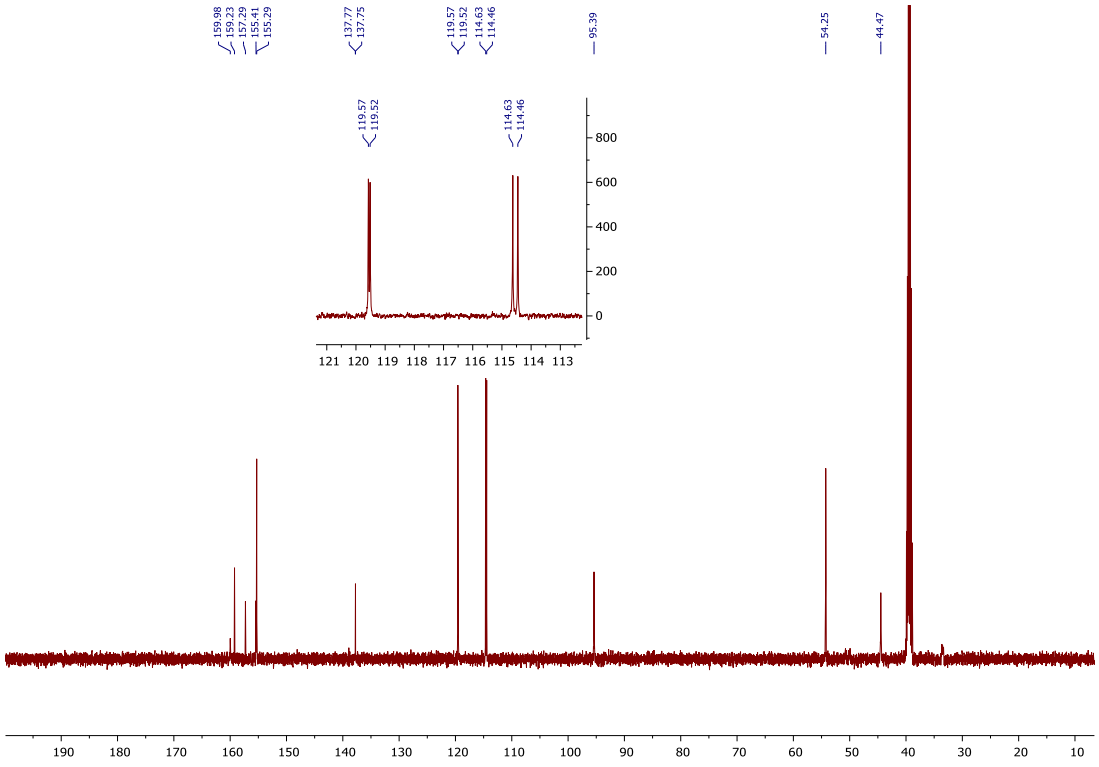
(*R*)-i, ^{13}C, 126 MHz, DMSO-d_6



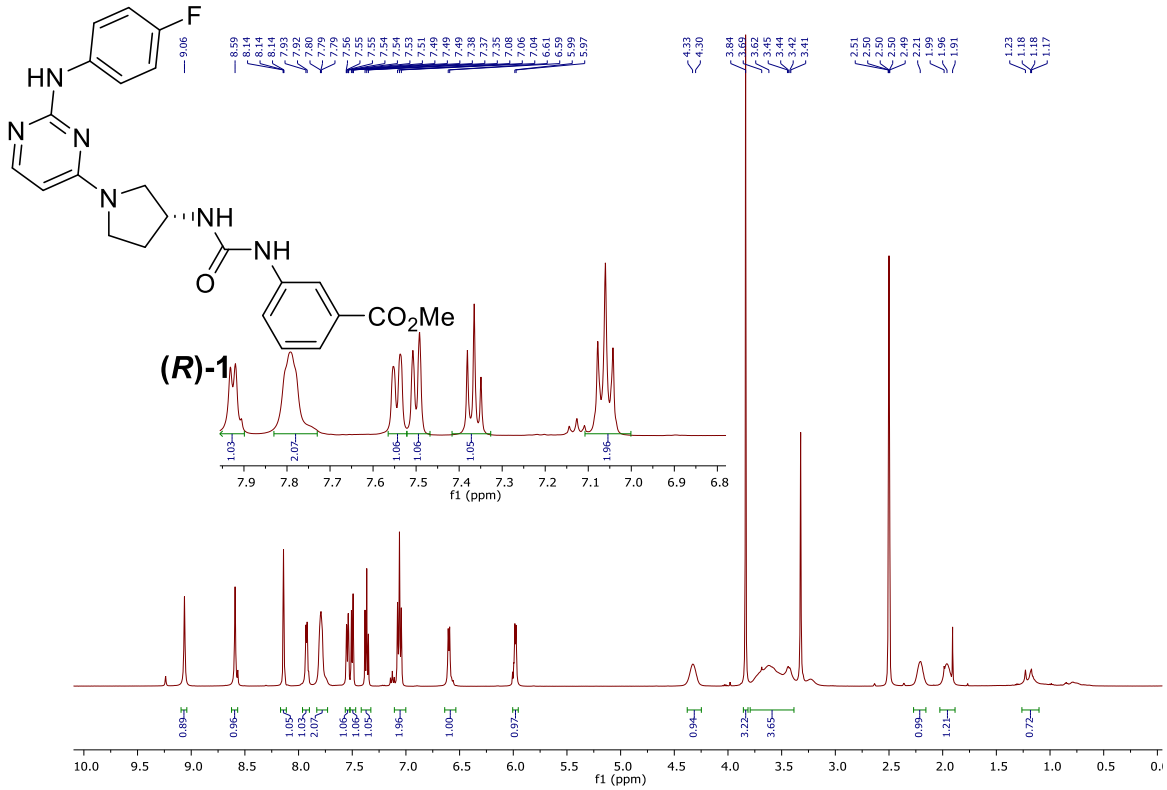
(*R*)-ii, ^{1}H, 500 MHz, DMSO-d_6



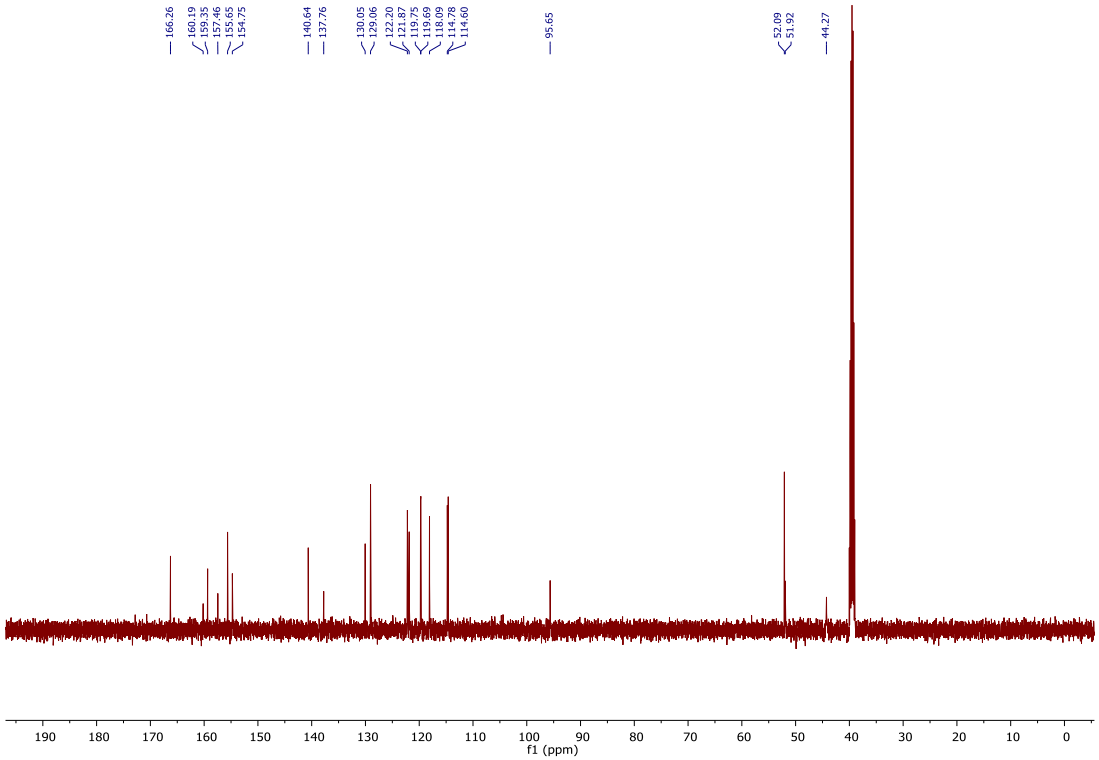
(*R*)-ii, ^{13}C, 126 MHz, DMSO-d_6



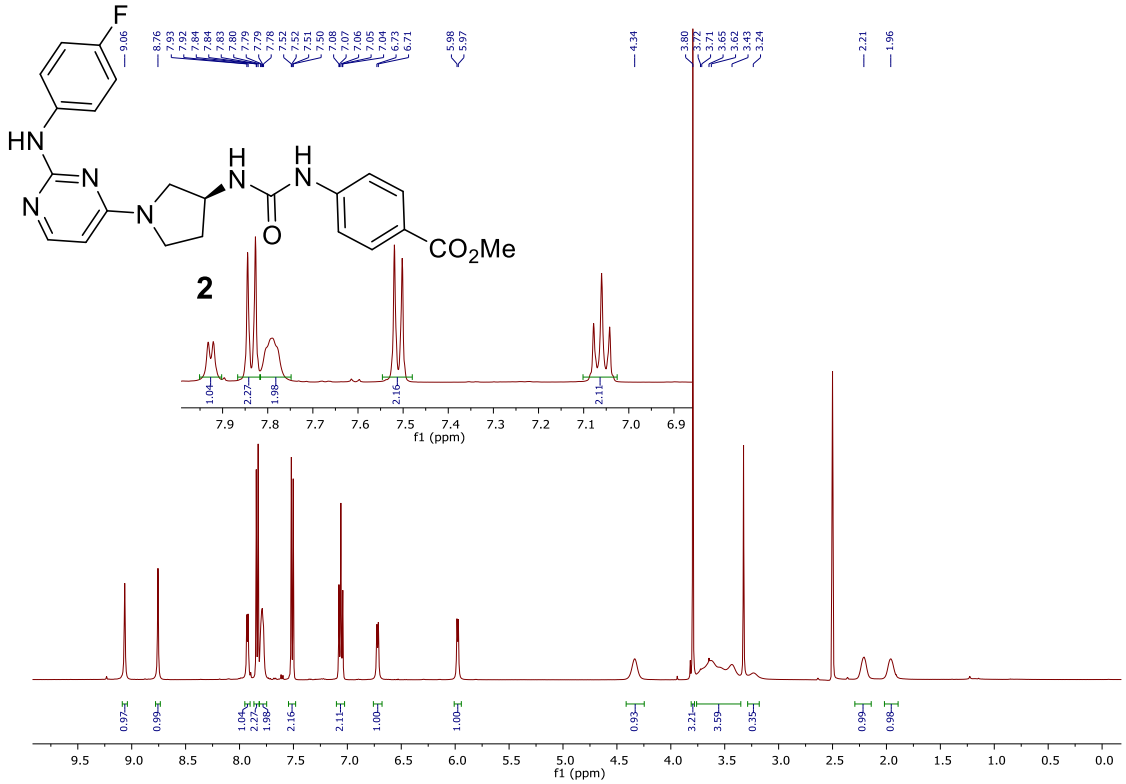
(*R*)-1, ^{1}H, 500 MHz, DMSO-d_6



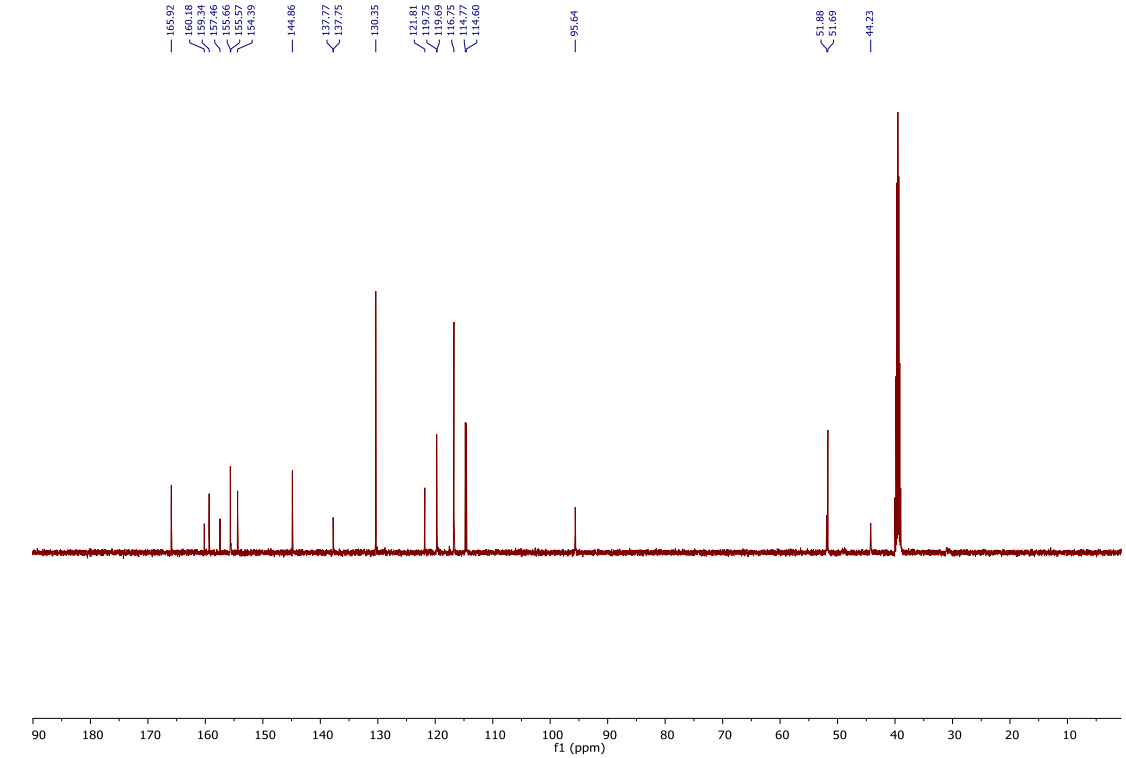
(*R*)-1, ^{13}C, 126 MHz, DMSO-d_6



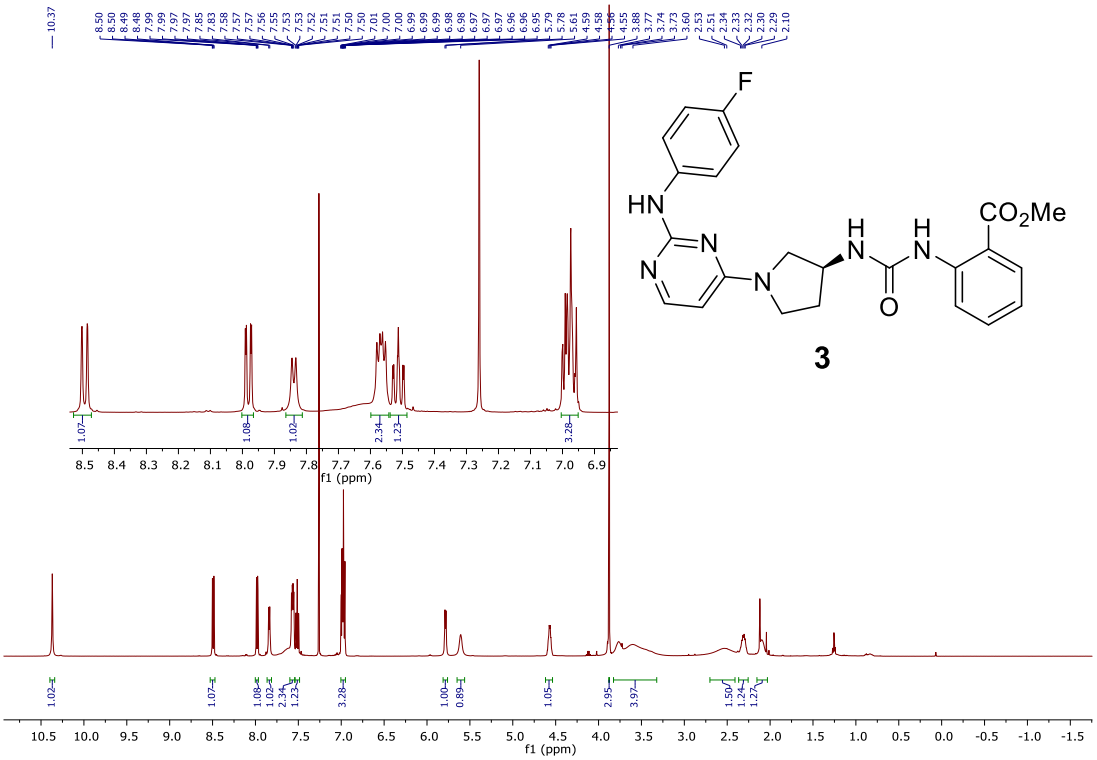
2, ^{1}H, 500 MHz, DMSO-d_6



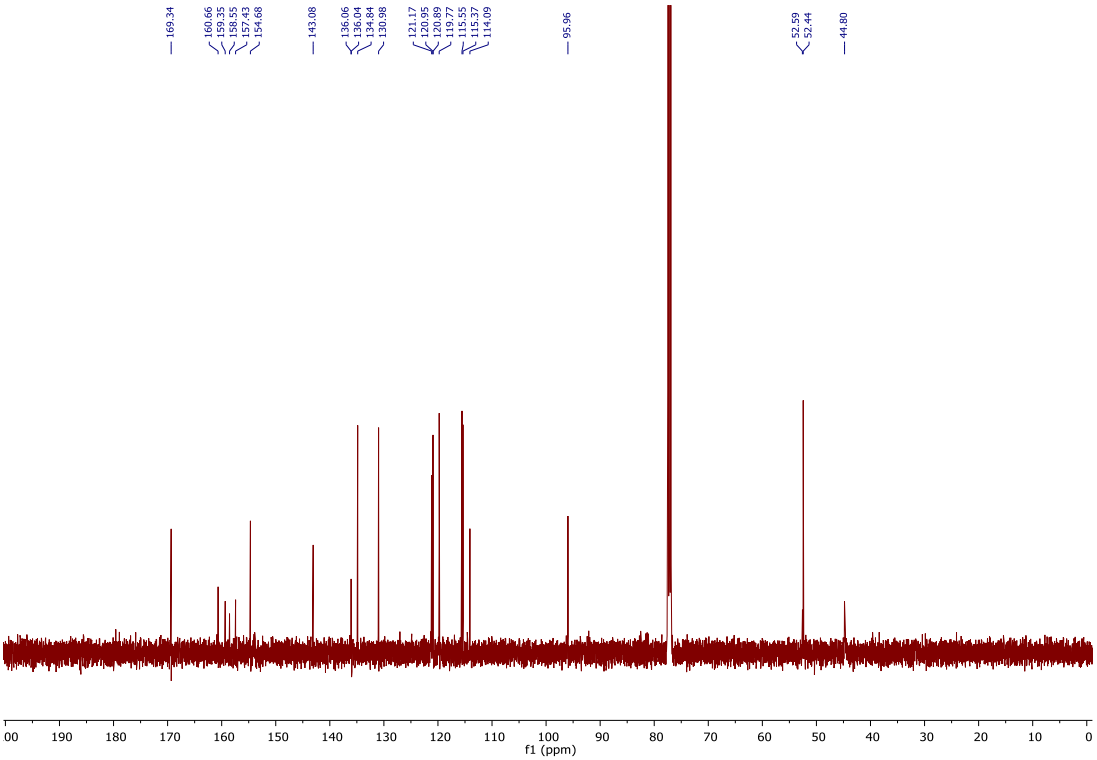
2, ^{13}C, 126 MHz, DMSO-d_6



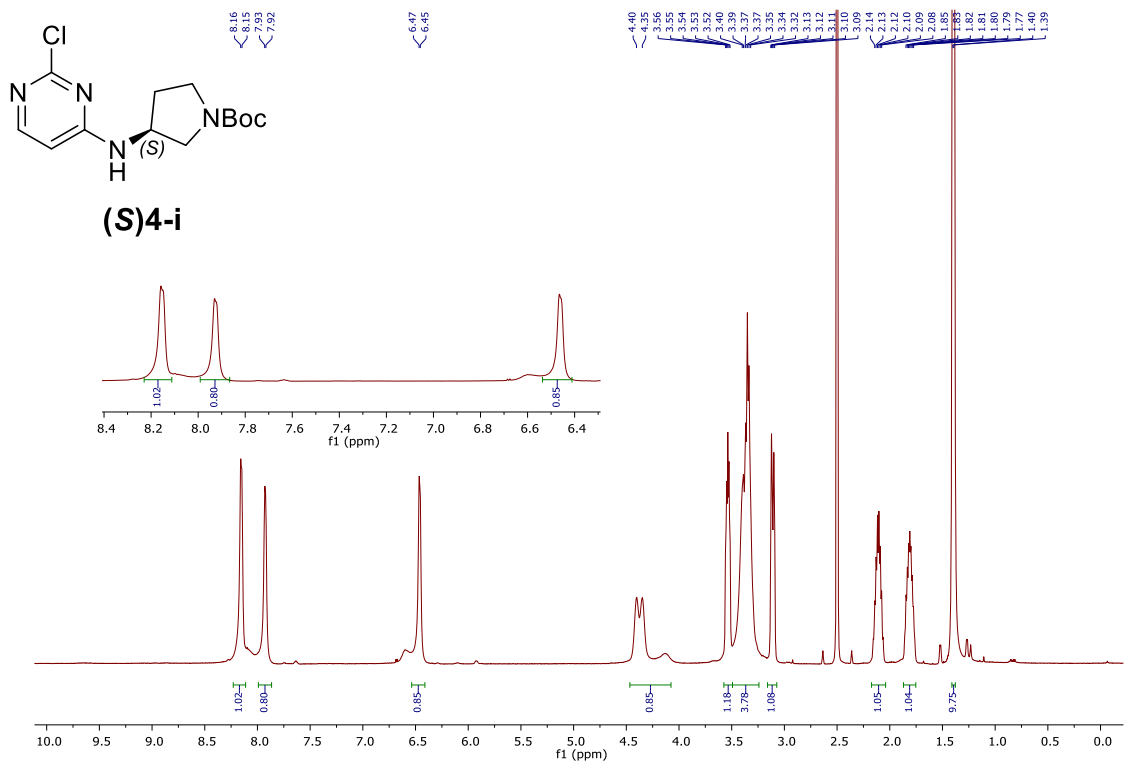
3, ^{1}H, 500 MHz, $CDCl_3$



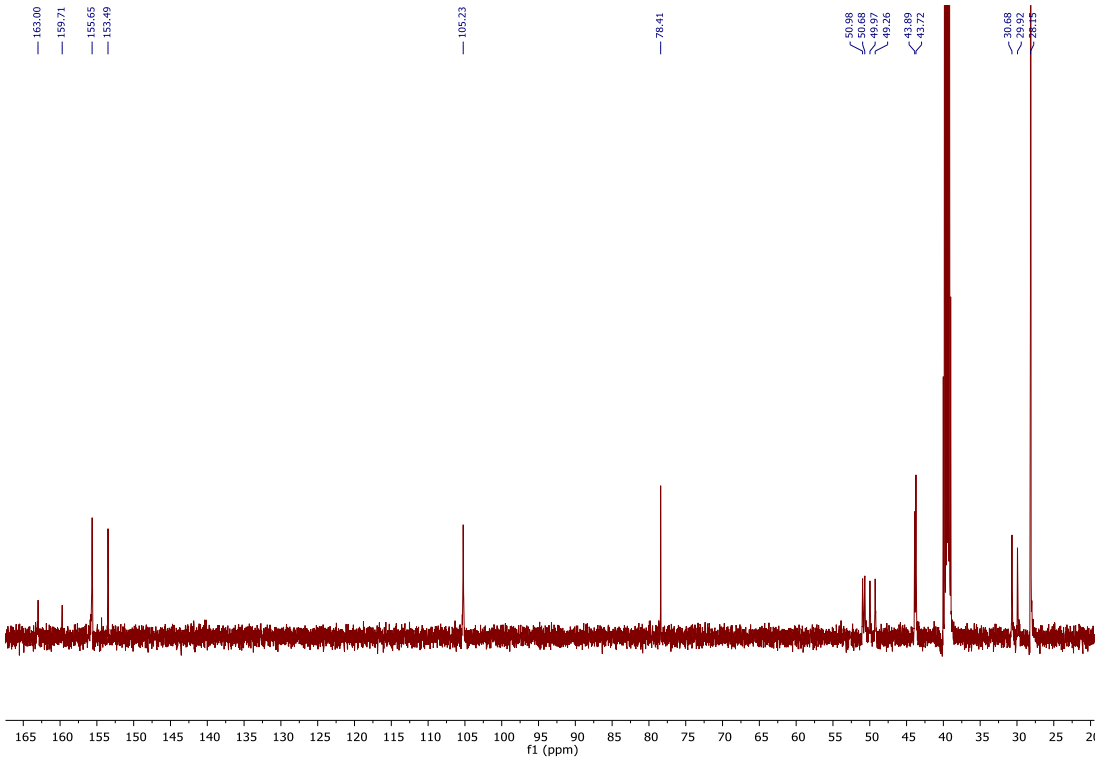
3, ^{13}C, 126 MHz, $CDCl_3$



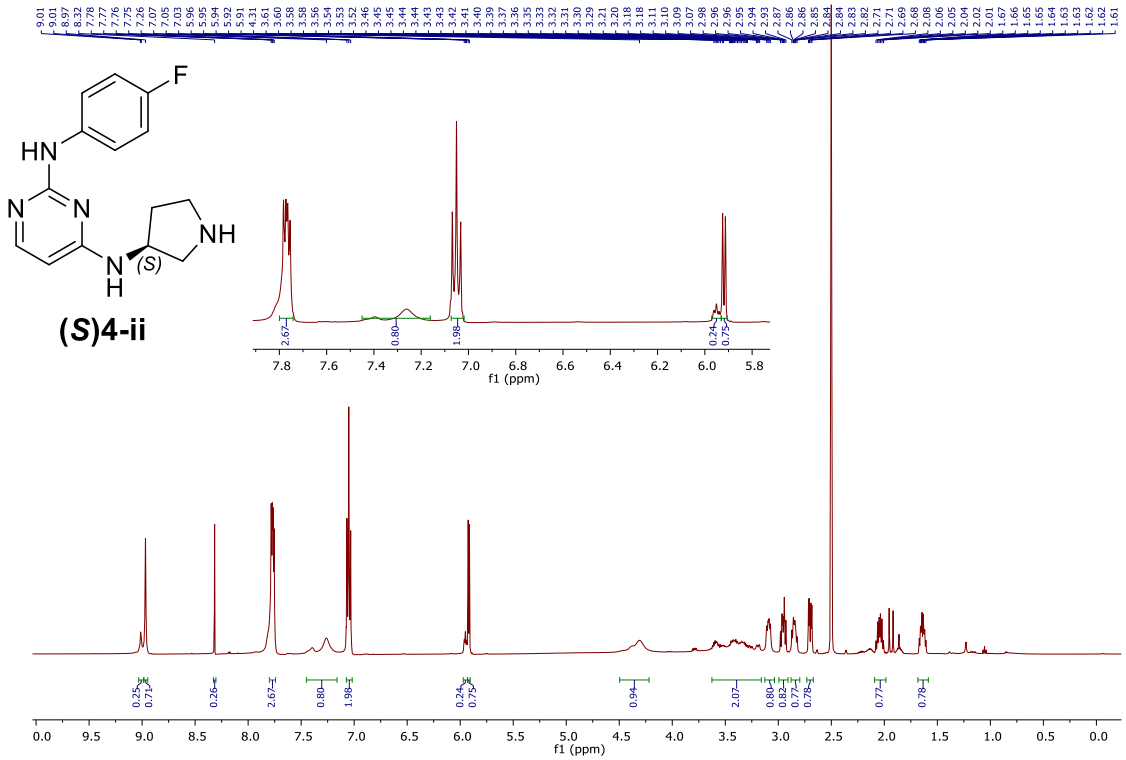
(*S*)4-i, ^{1}H, 500 MHz, DMSO-d_6



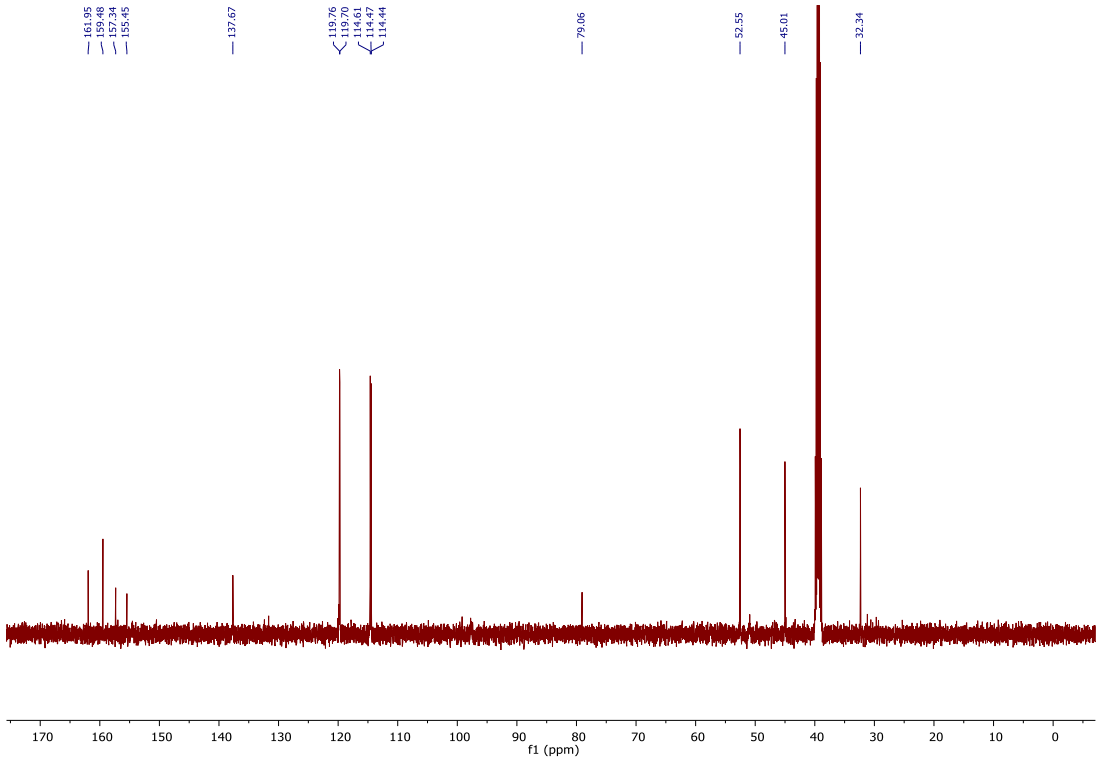
(*S*)4-i, ^{13}C, 126 MHz, DMSO-d_6



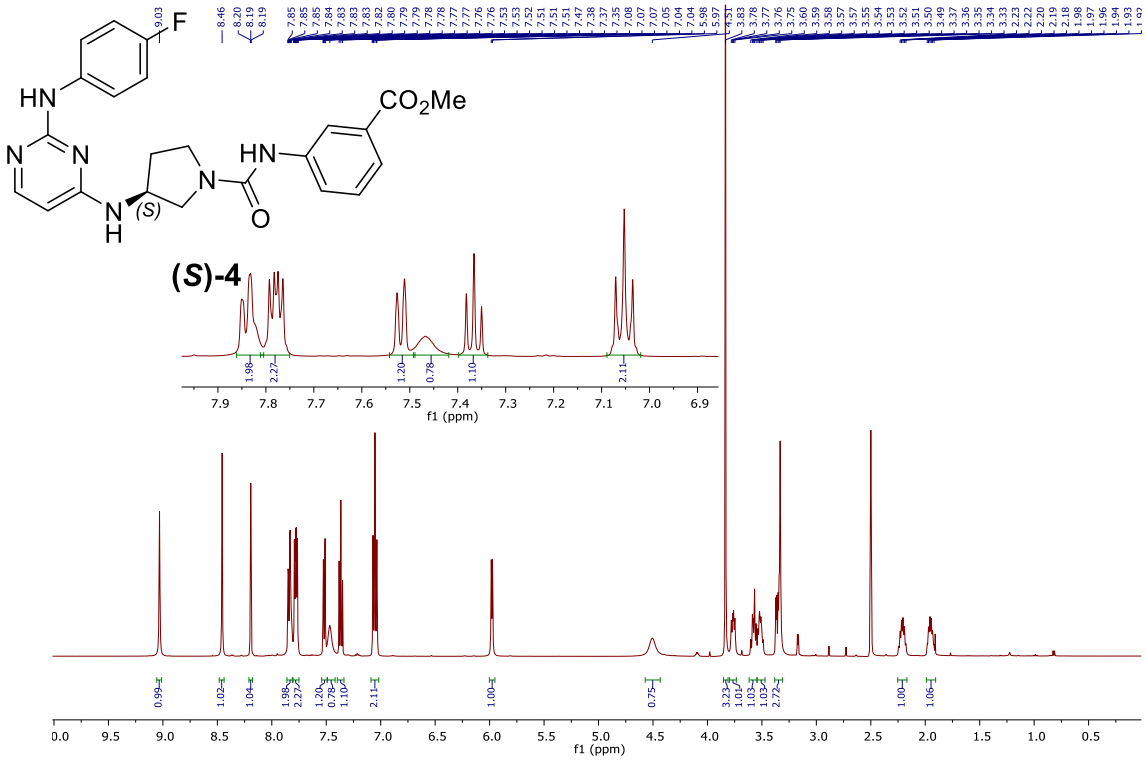
(*S*)4-ii, ^{1}H, 500 MHz, DMSO-d_6



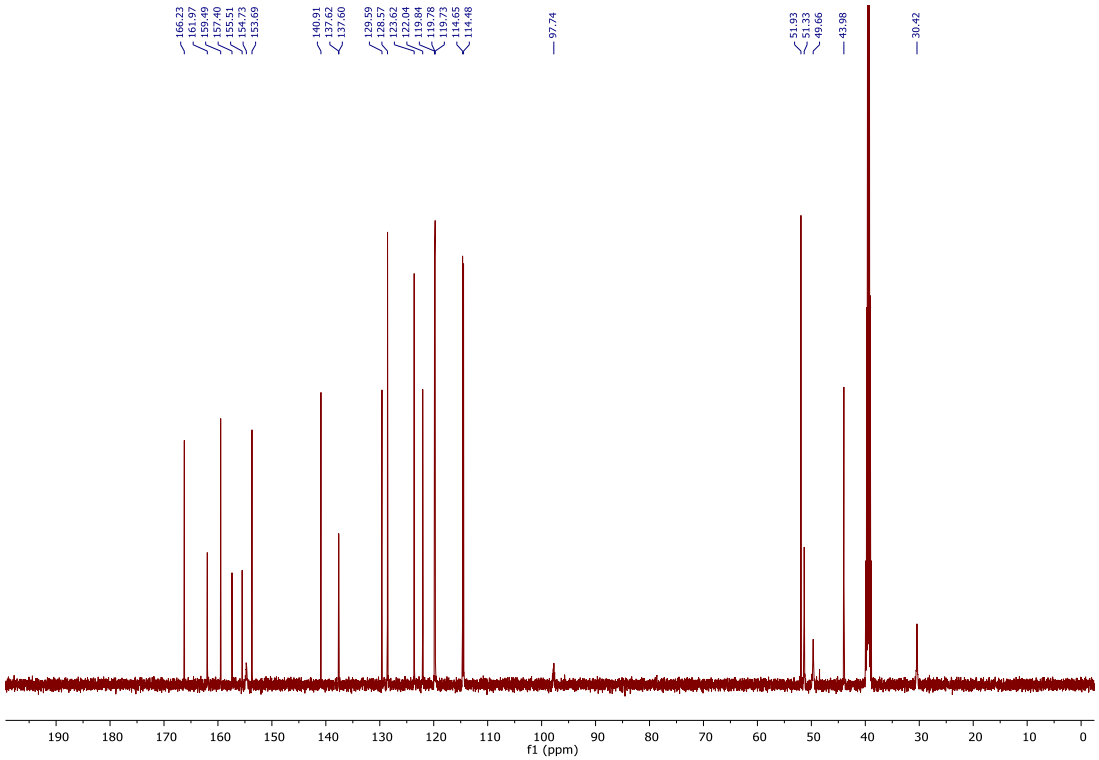
(*S*)4-ii, ^{13}C, 126 MHz, DMSO-d_6



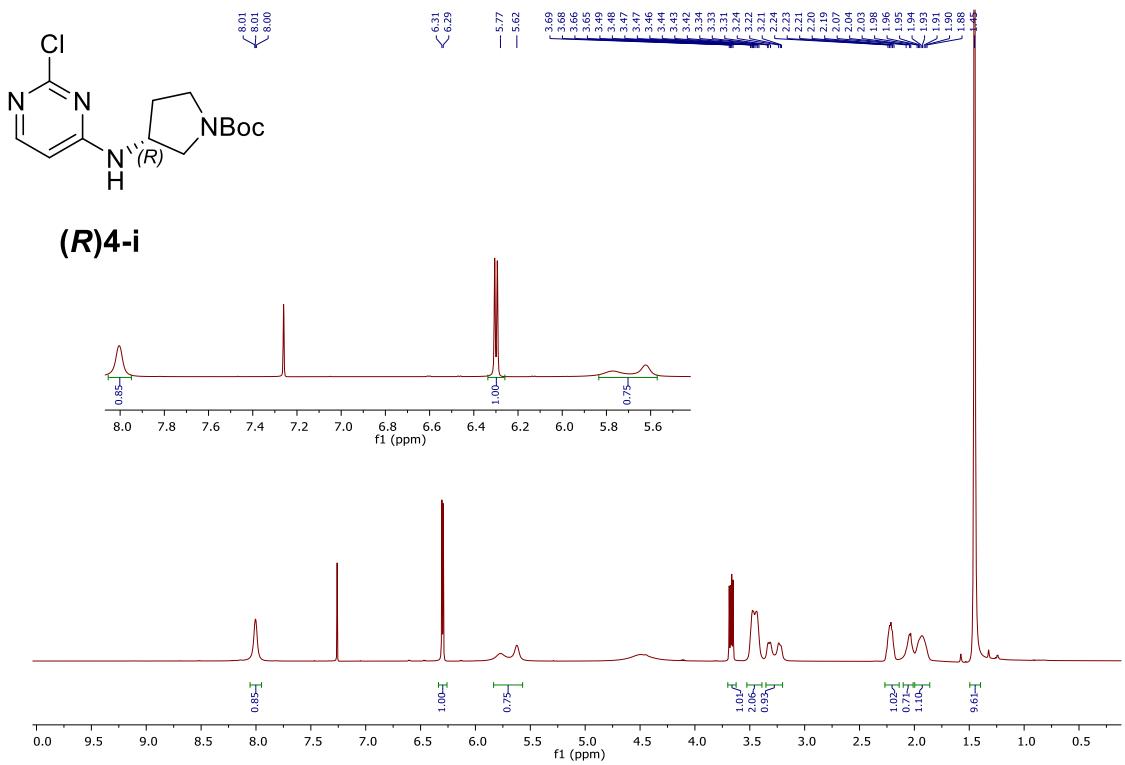
(*S*)-4, ^{1}H, 500 MHz, DMSO-d_6



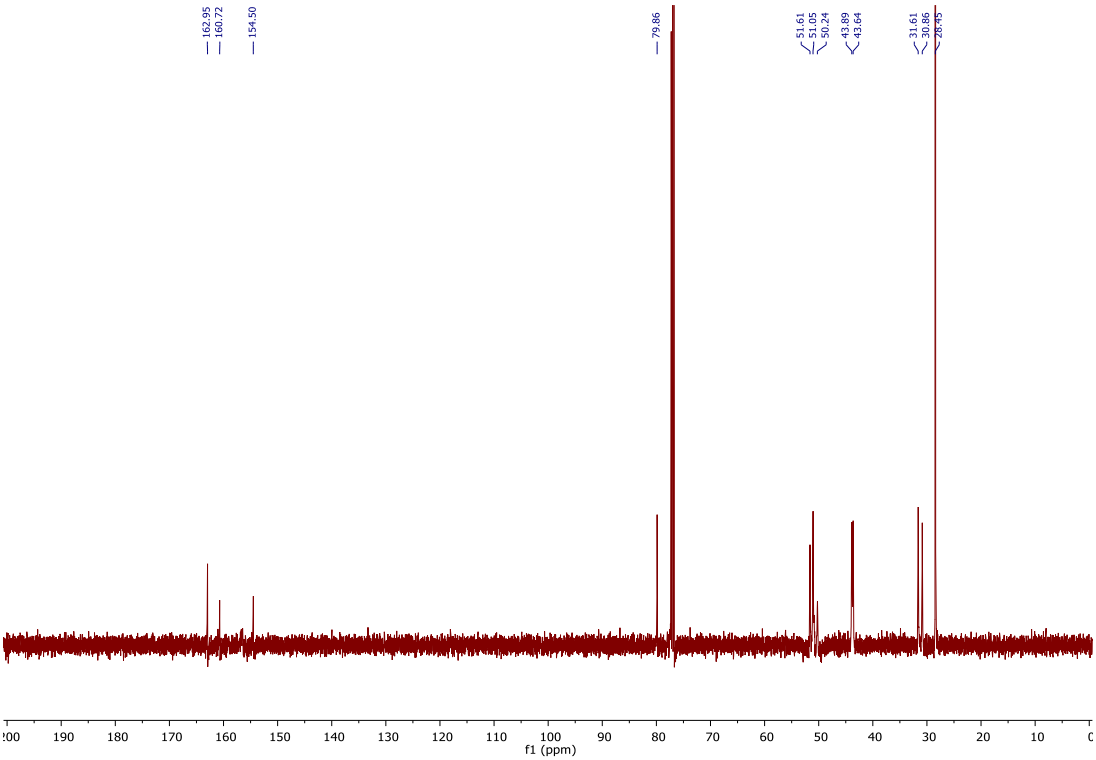
(*S*)-4, ^{13}C, 126 MHz, DMSO-d_6



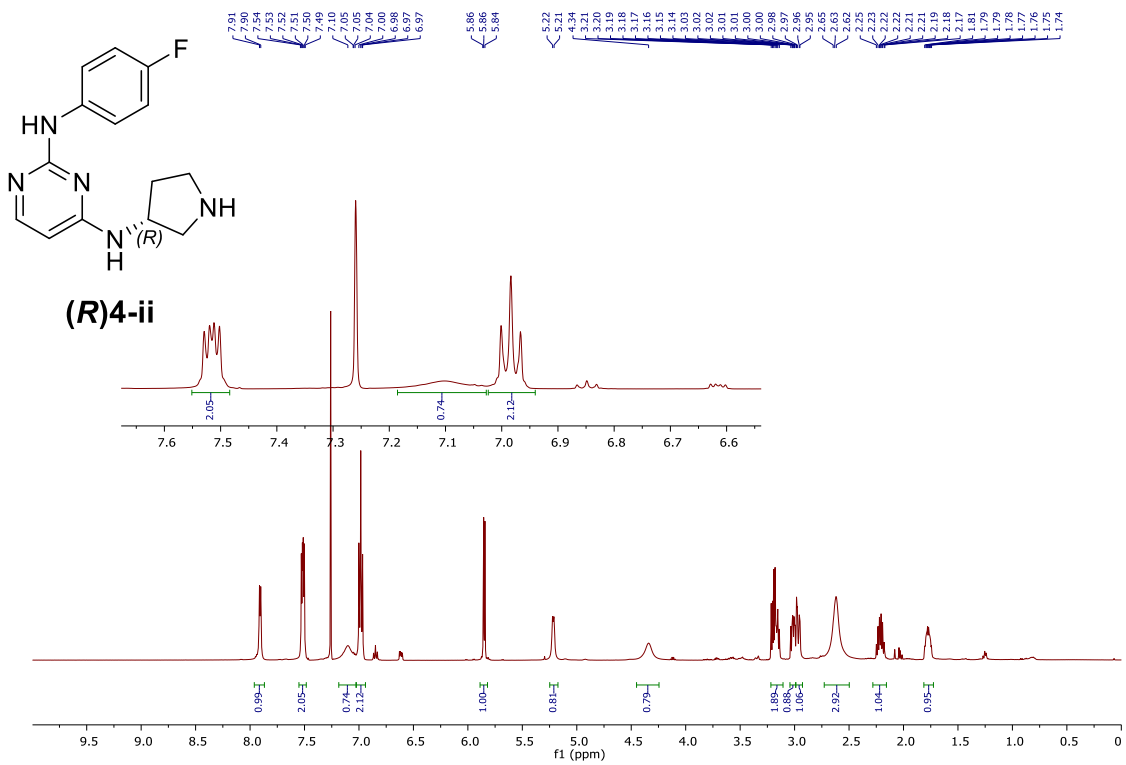
(*R*)4-i, ^{1}H, 500 MHz, $CDCl_3$



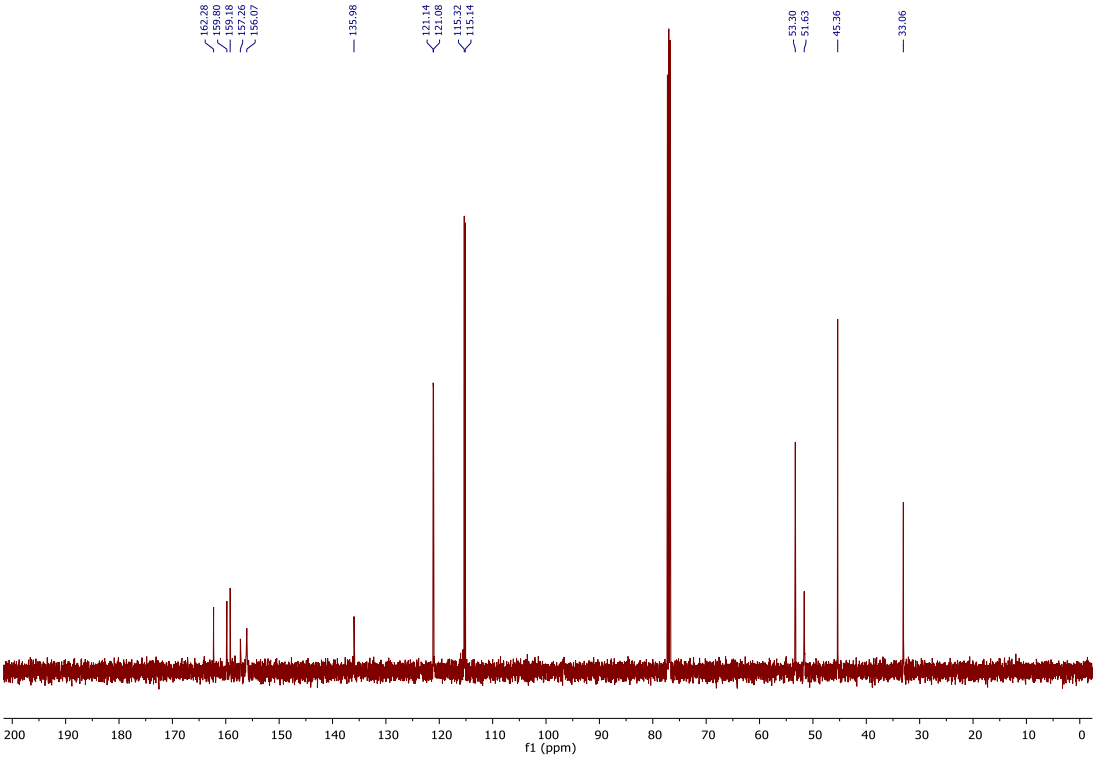
(*R*)4-i, ^{13}C, 126 MHz, $CDCl_3$



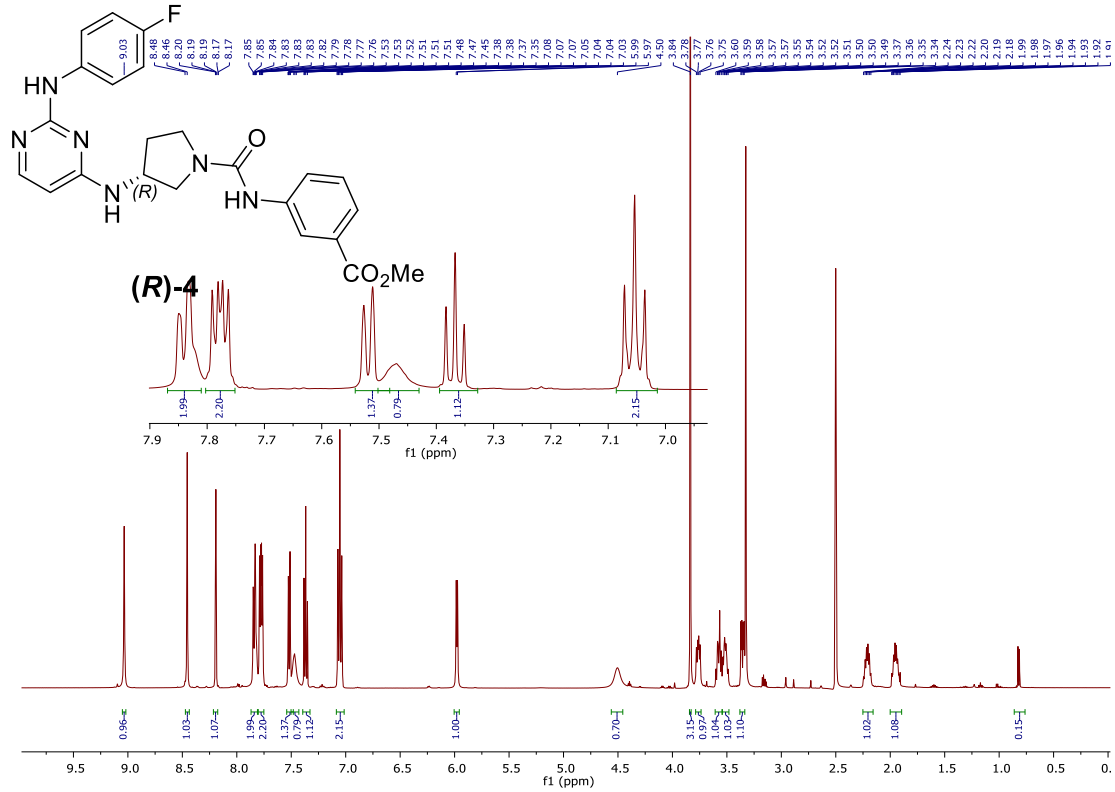
(*R*)4-ii, ^{1}H, 500 MHz, $CDCl_3$



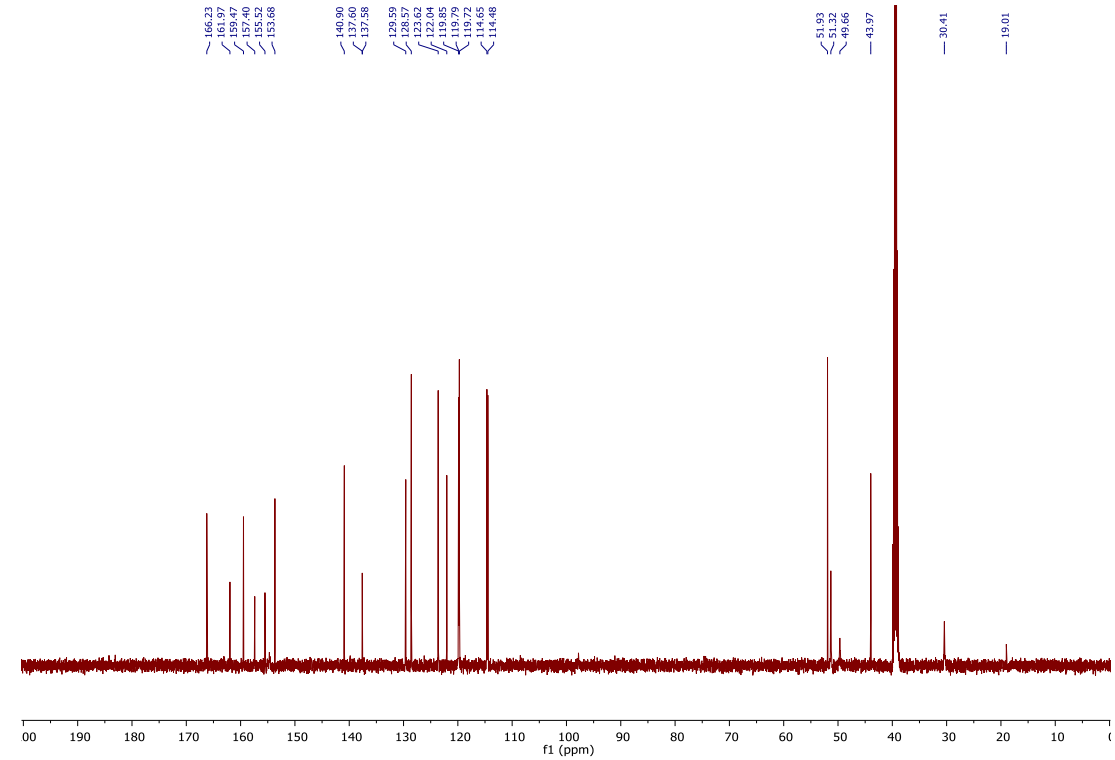
(*R*)4-ii, ^{13}C, 126 MHz, $CDCl_3$



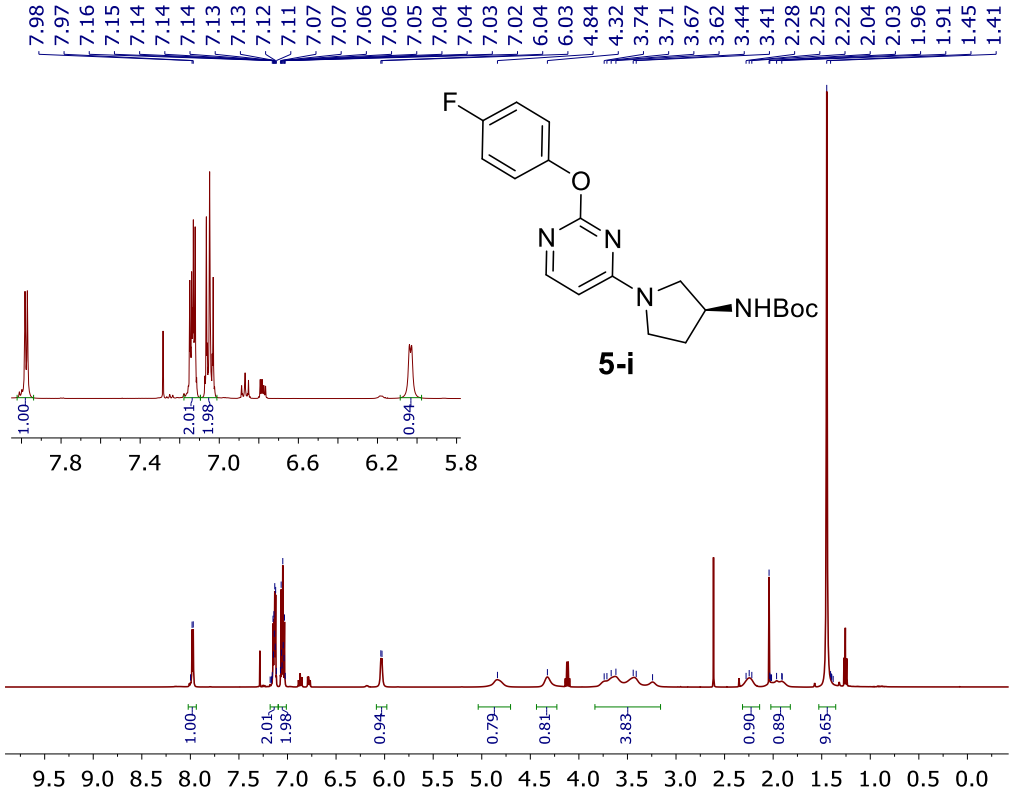
(*R*)-4, ^{1}H, 500 MHz, DMSO-d_6



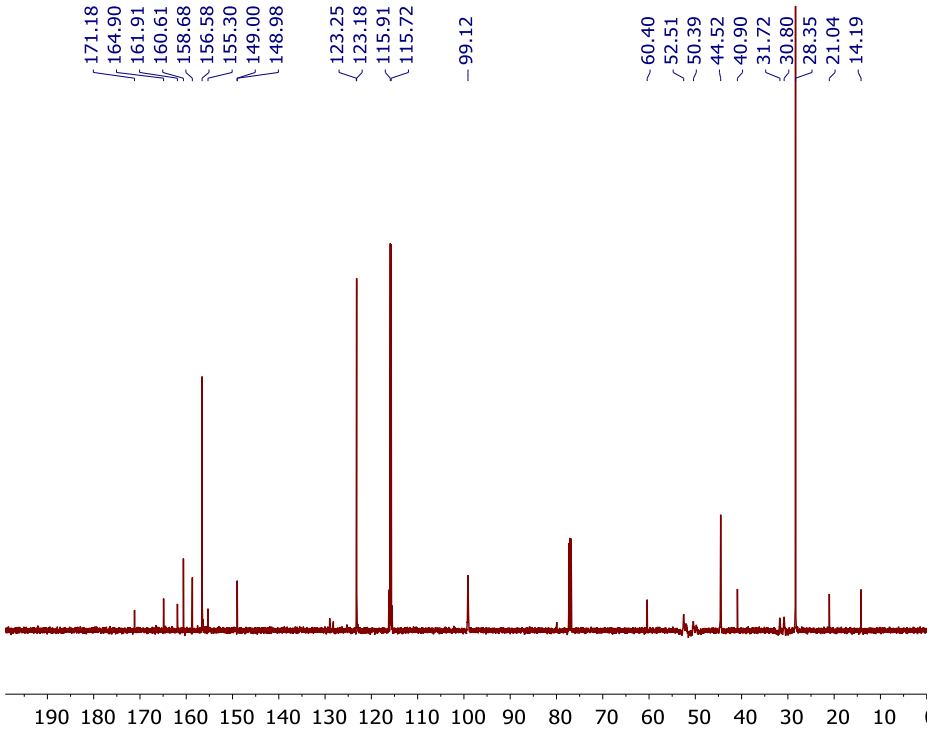
(*R*)-4, ^{13}C, 126 MHz, DMSO-d_6



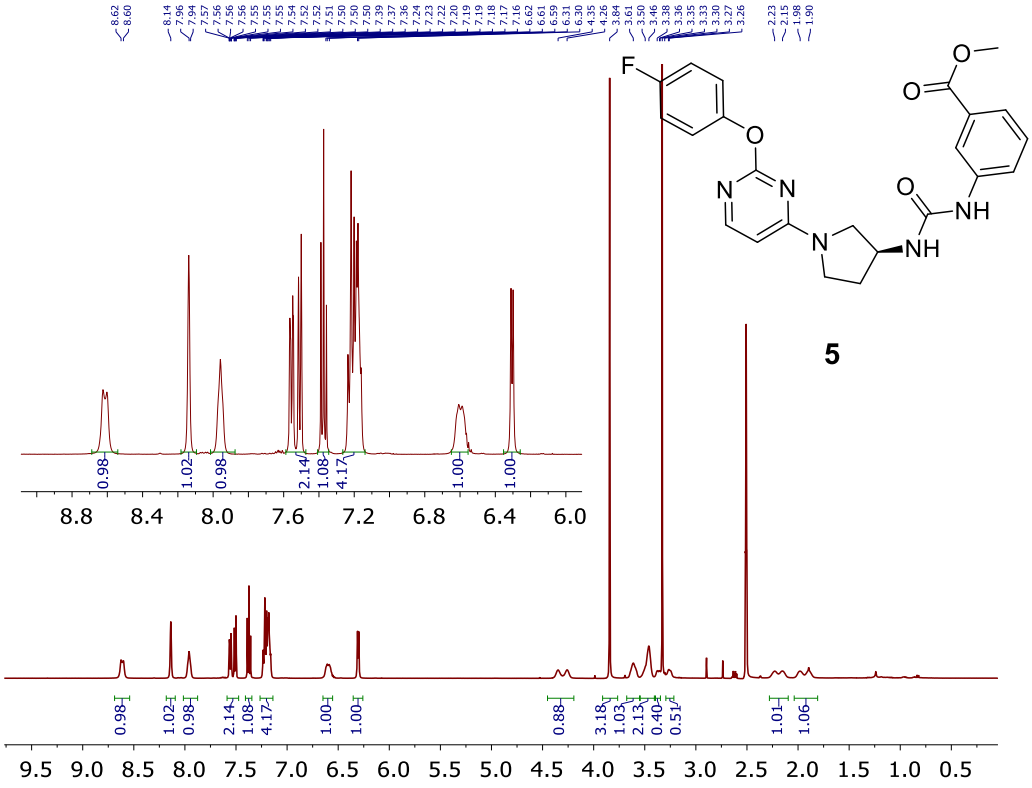
5-i, ^{1}H, 500 MHz, $CDCl_3$



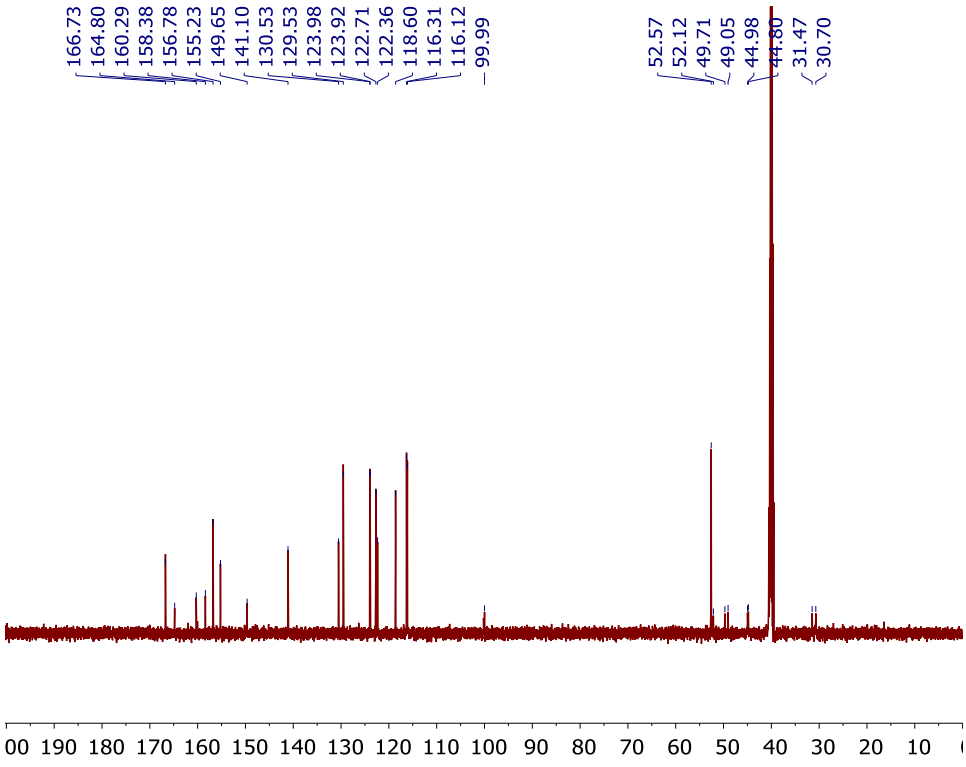
5-i, ^{13}C, 126 MHz, $CDCl_3$



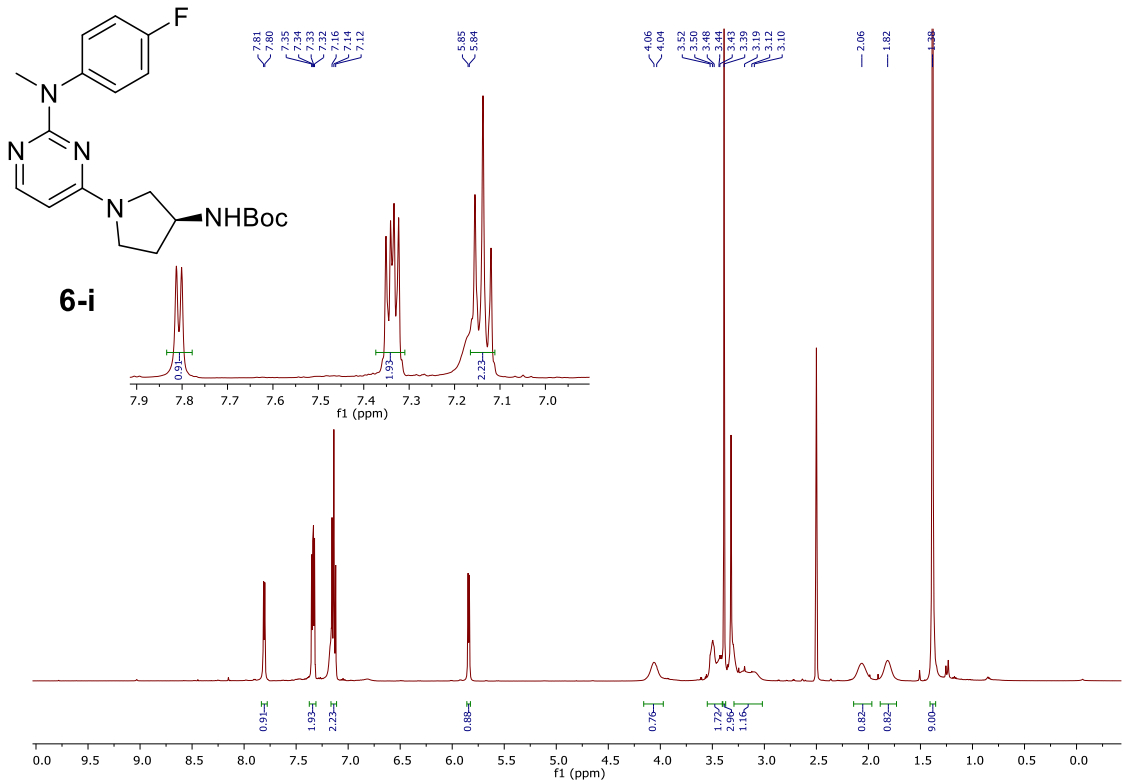
5, ^{1}H, 500 MHz, DMSO-d_6



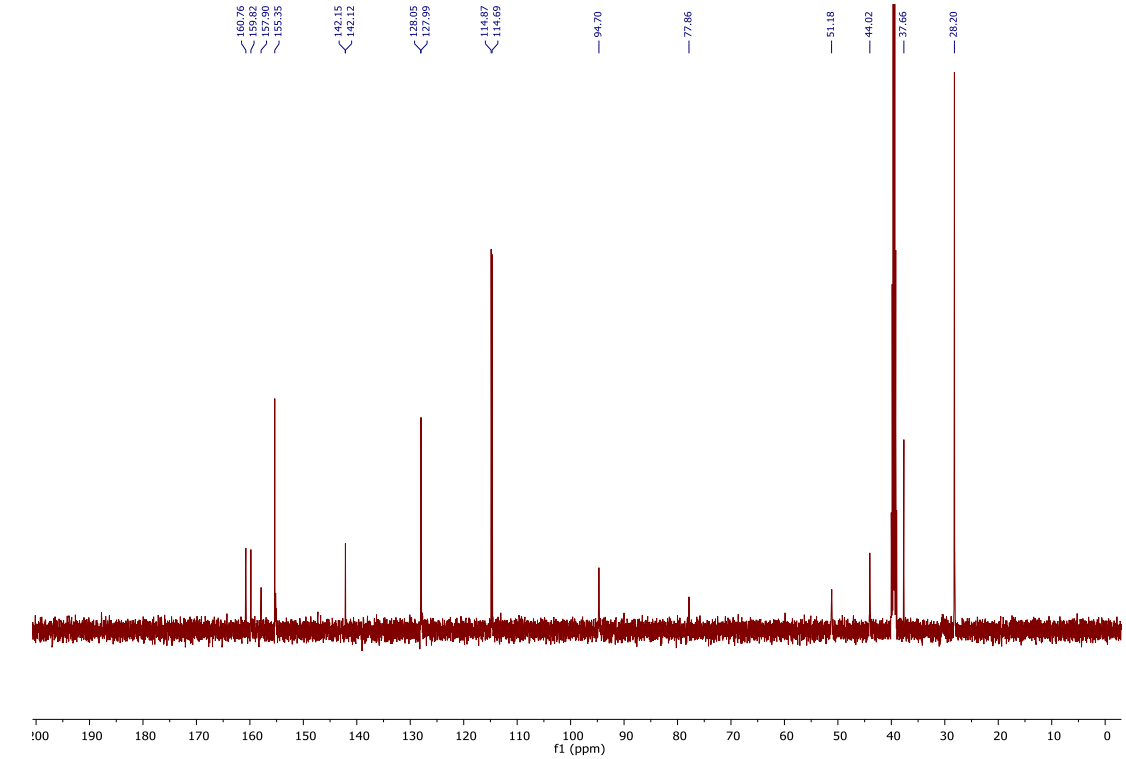
5, ^{13}C, 126 MHz, DMSO-d_6



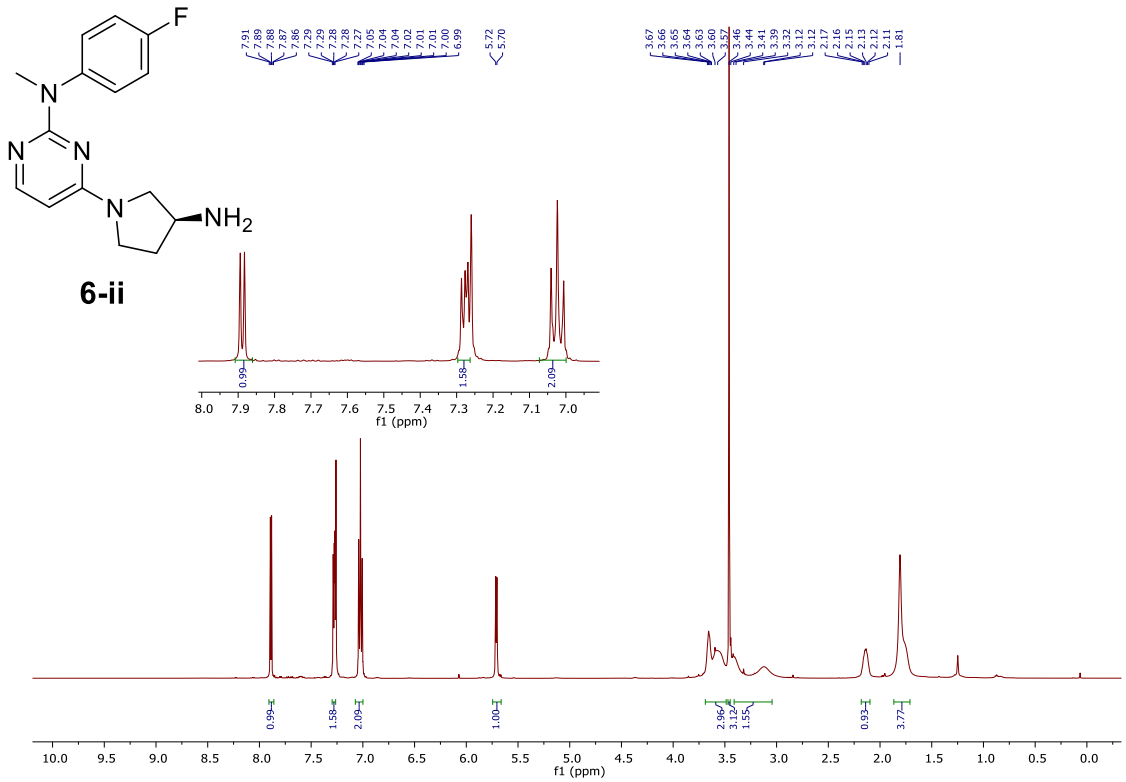
6-i, ^{1}H, 500 MHz, DMSO-d_6



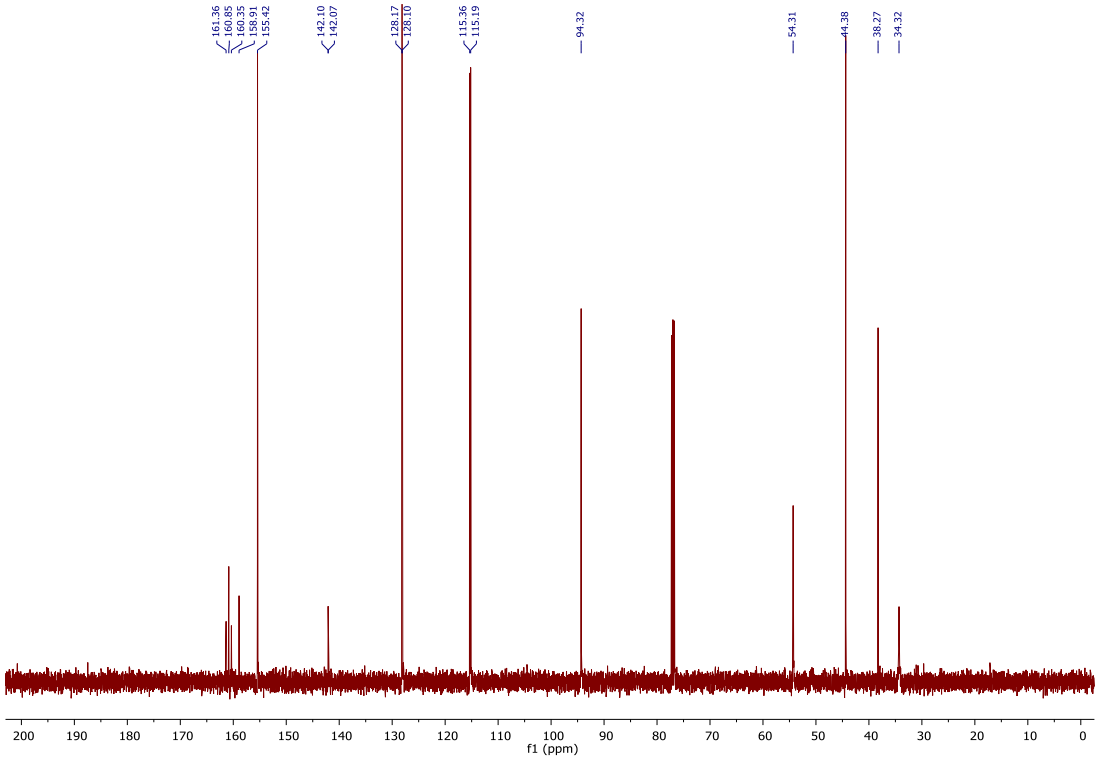
6-i, ^{13}C, 126 MHz, DMSO-d_6



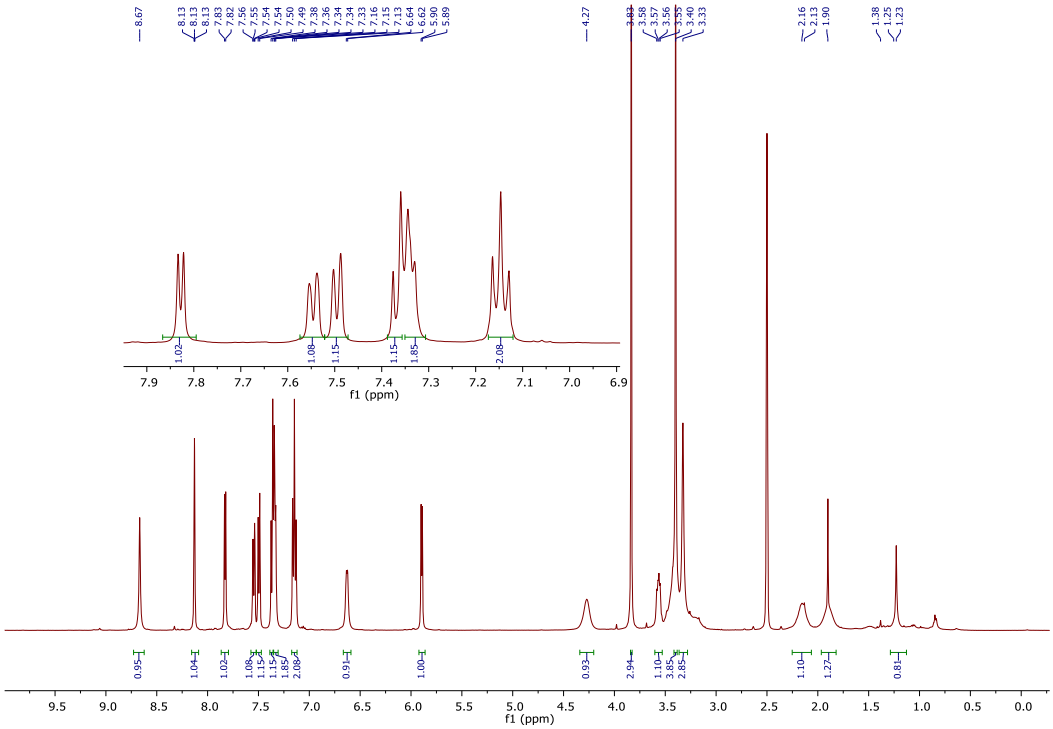
6-ii, ^{1}H, 500 MHz, $CDCl_3$



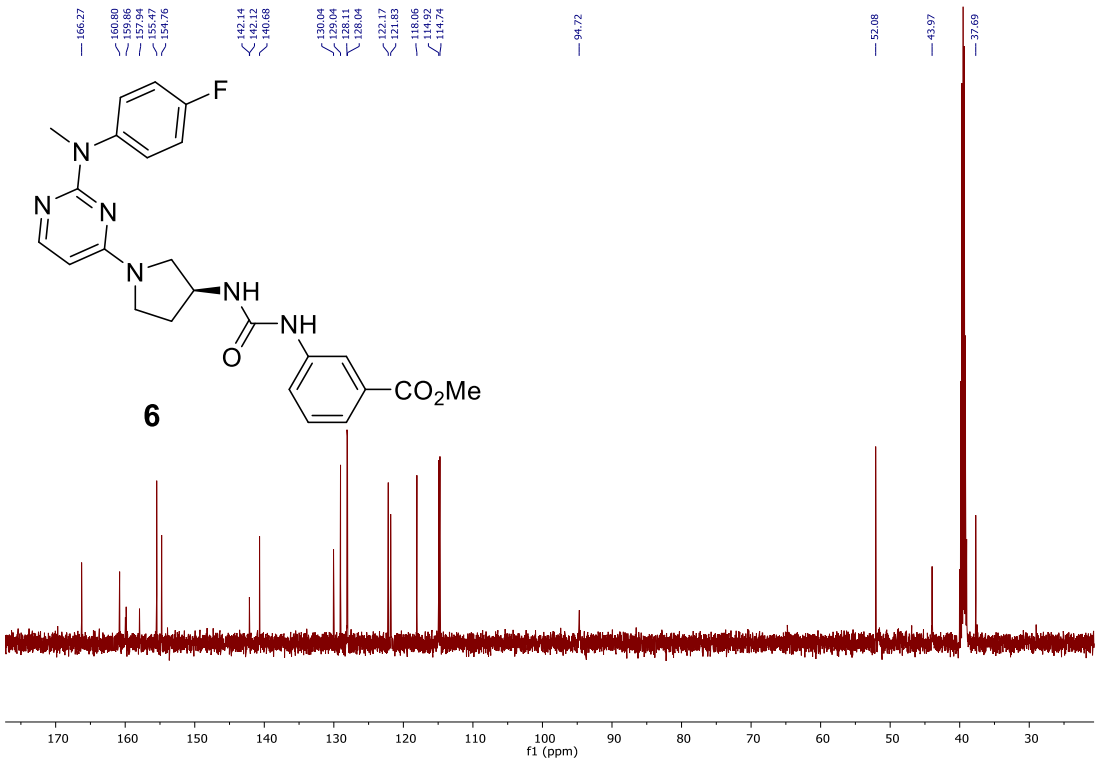
6-ii, ^{13}C, 126 MHz, $CDCl_3$



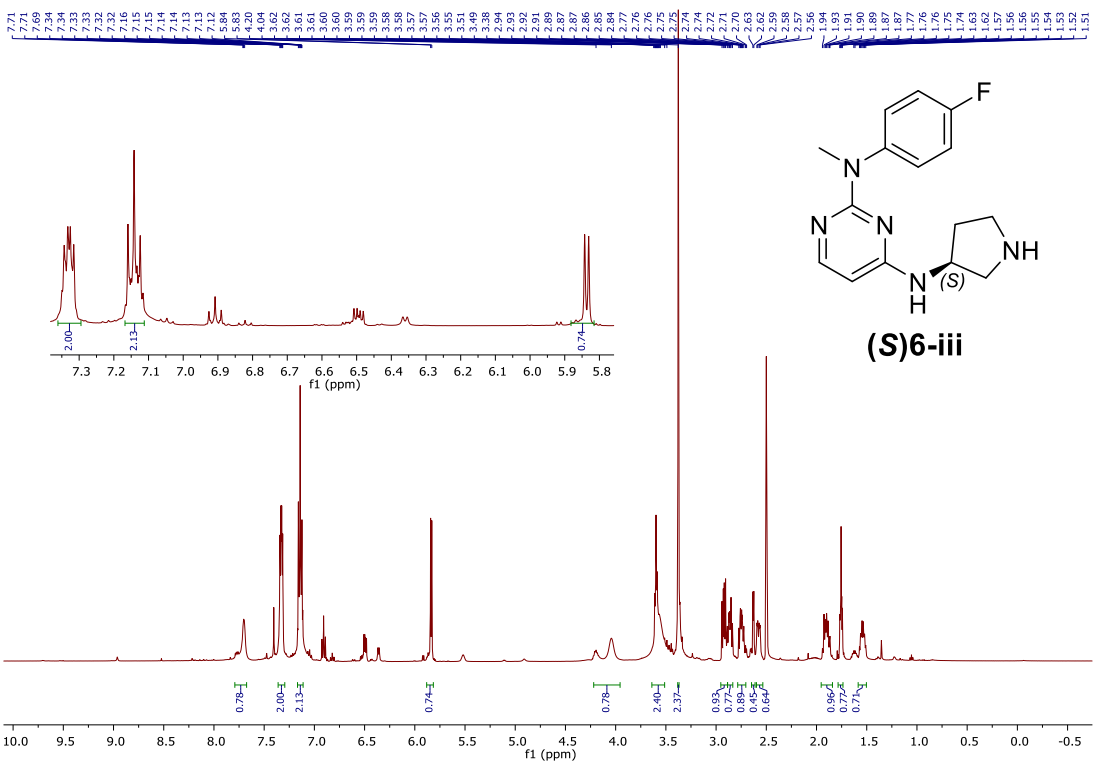
6, ^{1}H, 500 MHz, DMSO-d_6



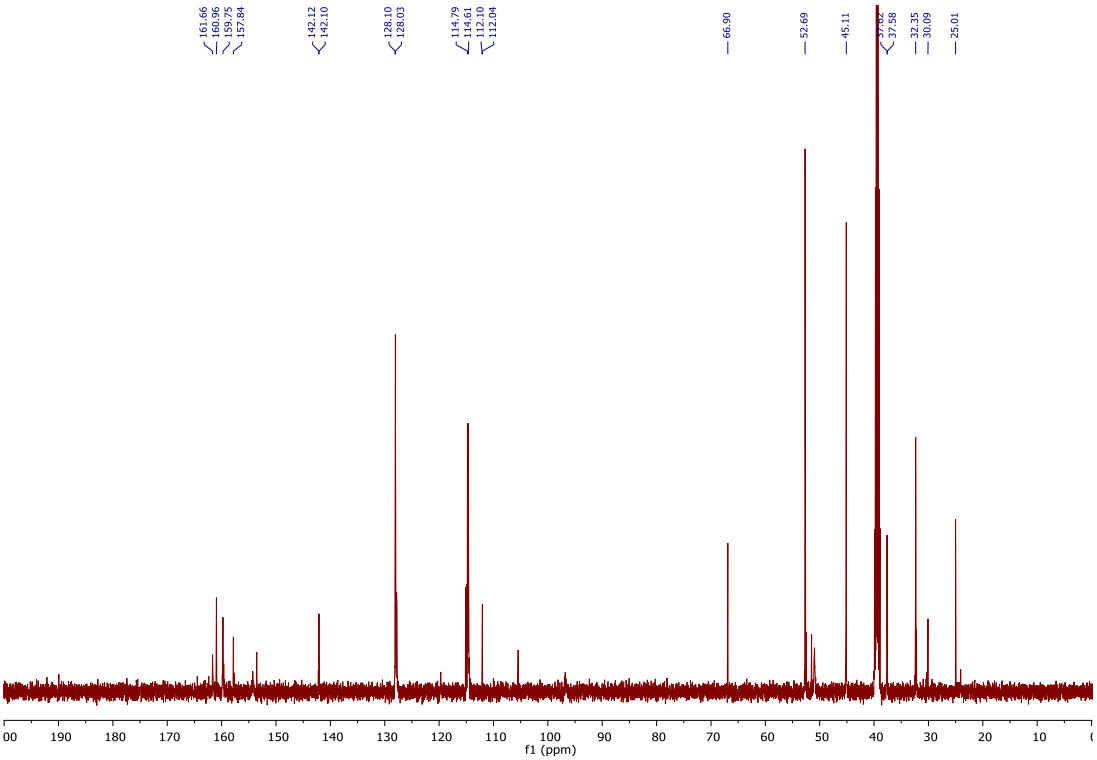
6, ^{13}C, 126 MHz, DMSO-d_6



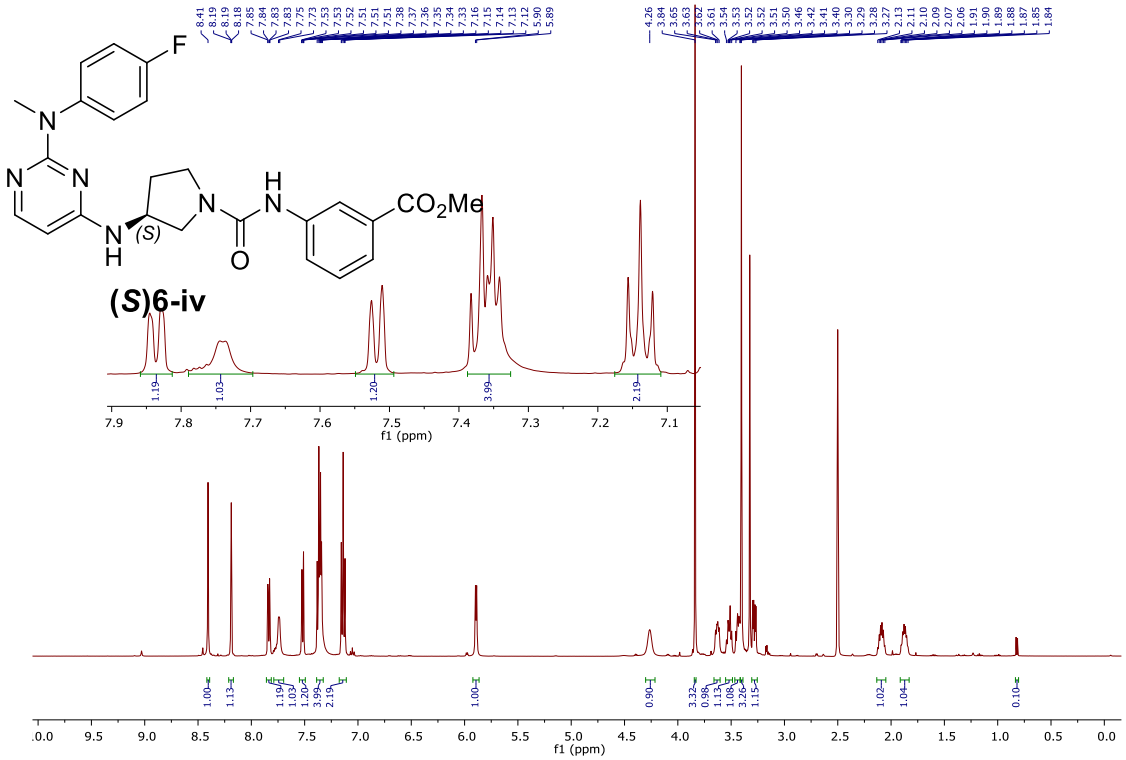
(*S*)6-iii, ^{1}H, 500 MHz, DMSO-d_6



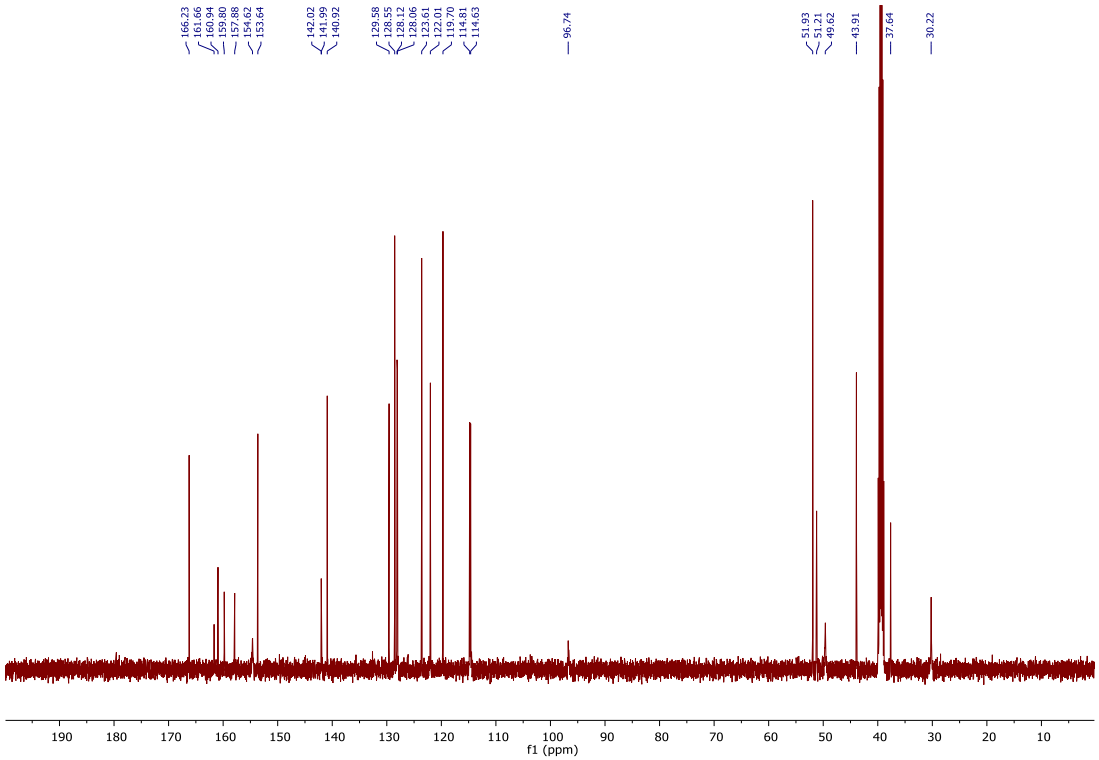
(*S*)6-iii, ^{13}C, 126 MHz, DMSO-d_6



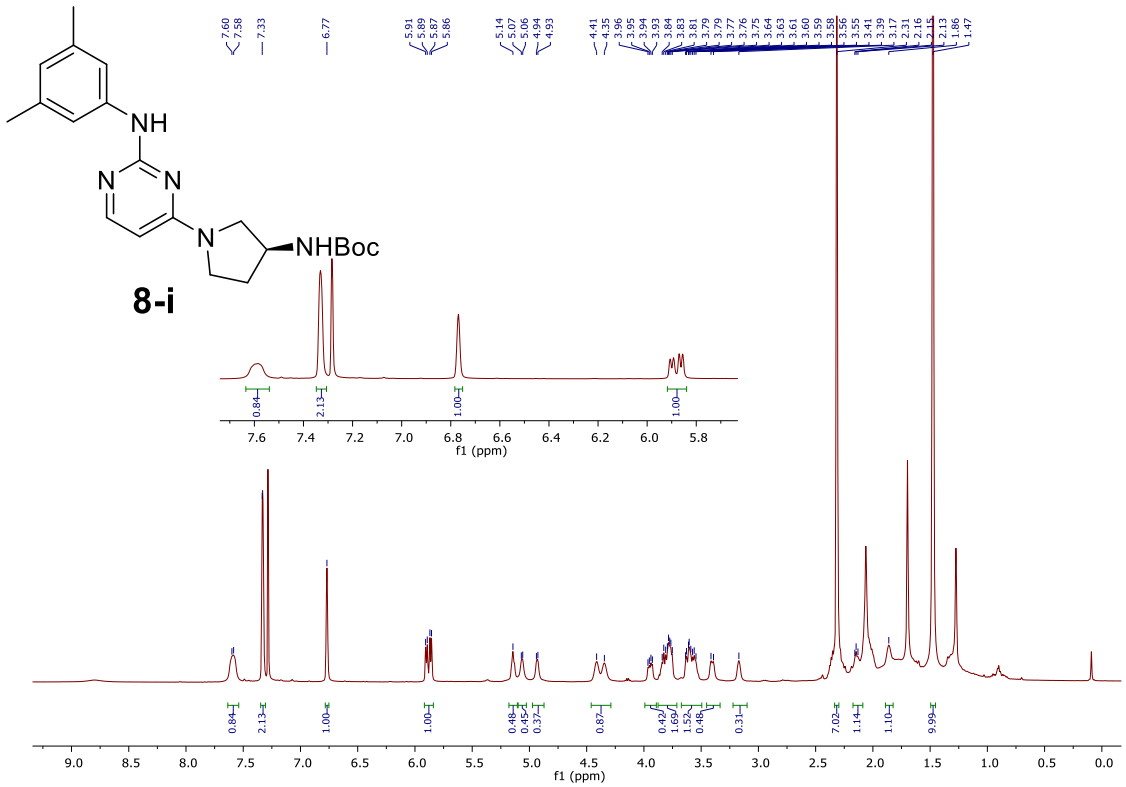
(*S*)6-iv, ^{1}H, 500 MHz, DMSO-d_6



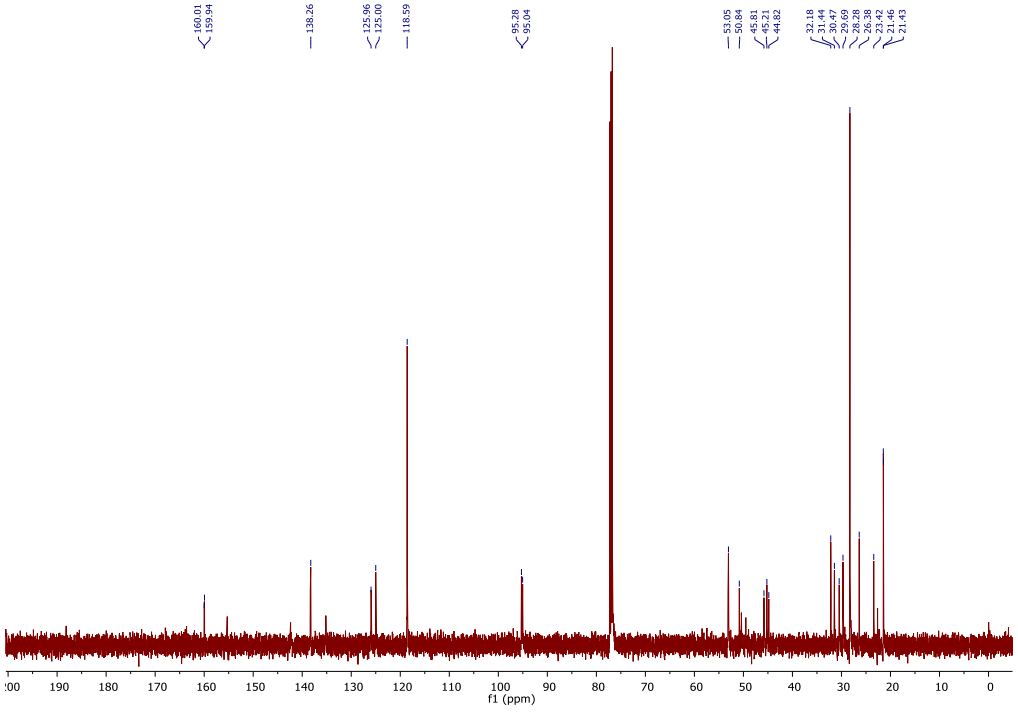
(*S*)6-iv, ^{13}C, 126 MHz, DMSO-d_6



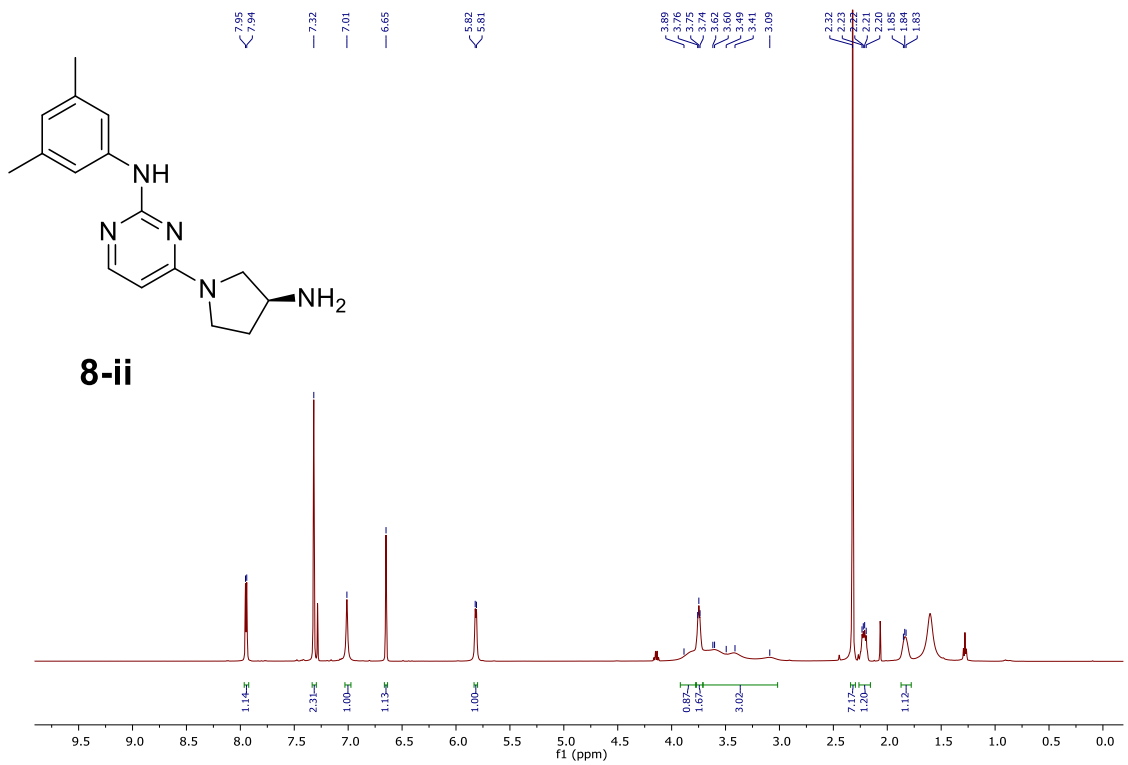
8-i, ^{1}H, 500 MHz, $CDCl_3$



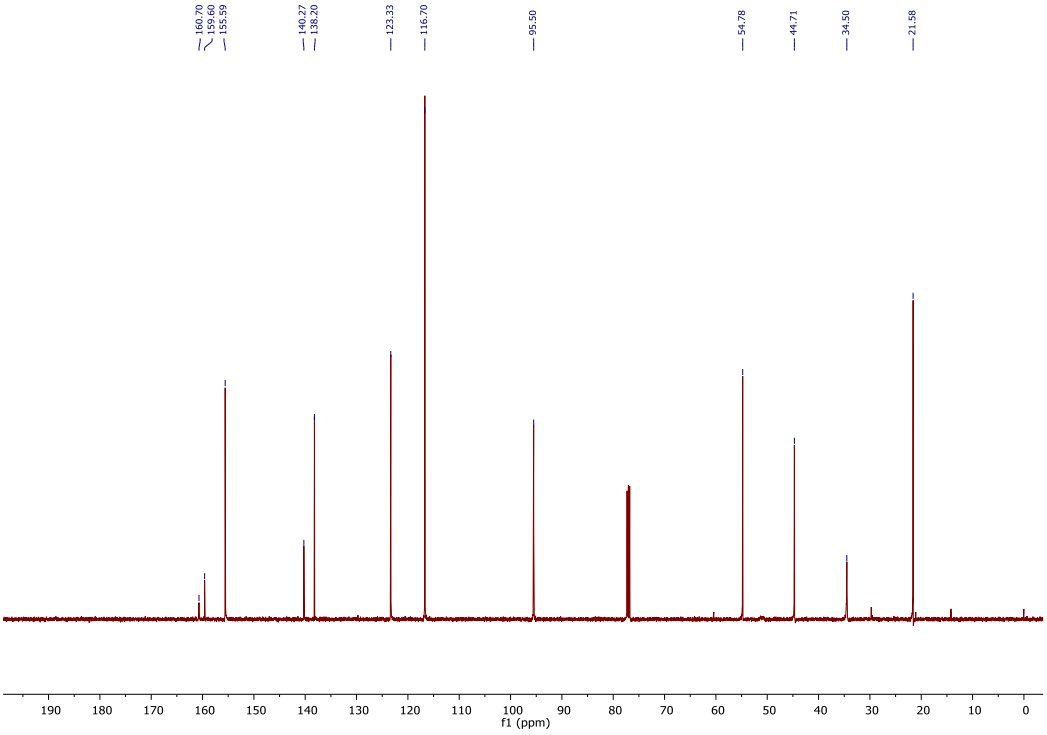
8-i, ^{13}C, 126 MHz, $CDCl_3$



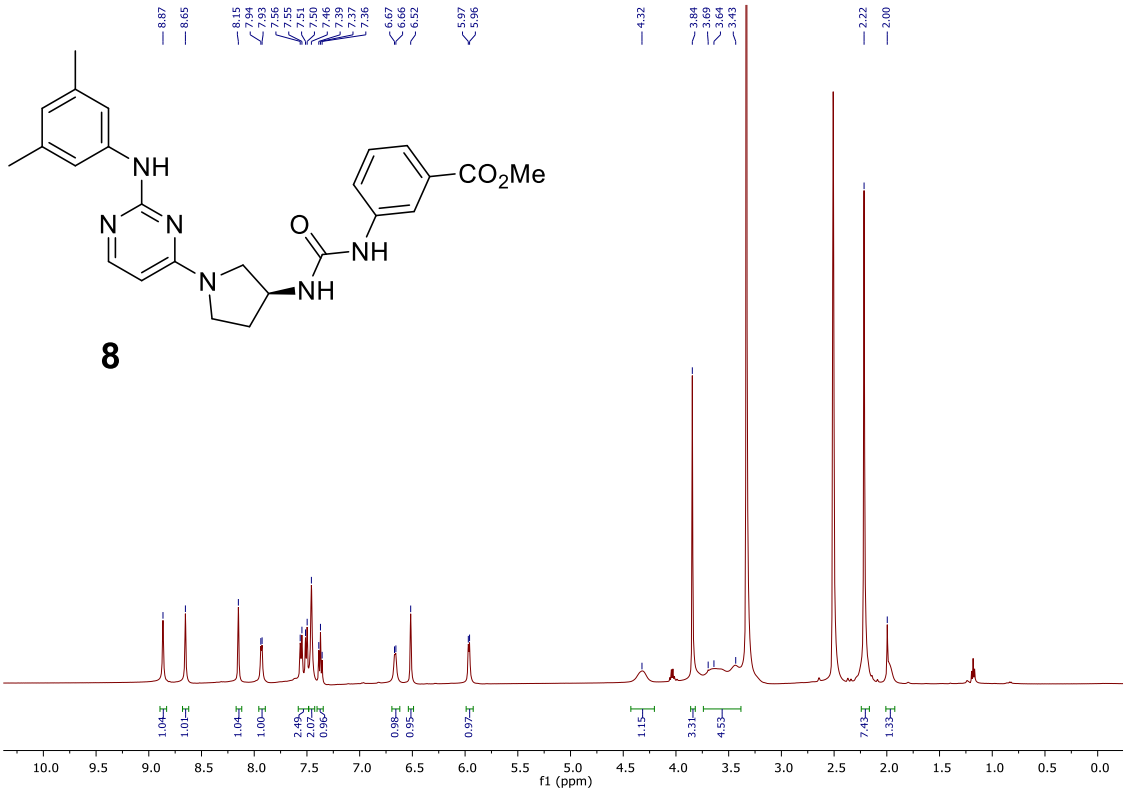
8-ii, ^{1}H, 500 MHz, $CDCl_3$



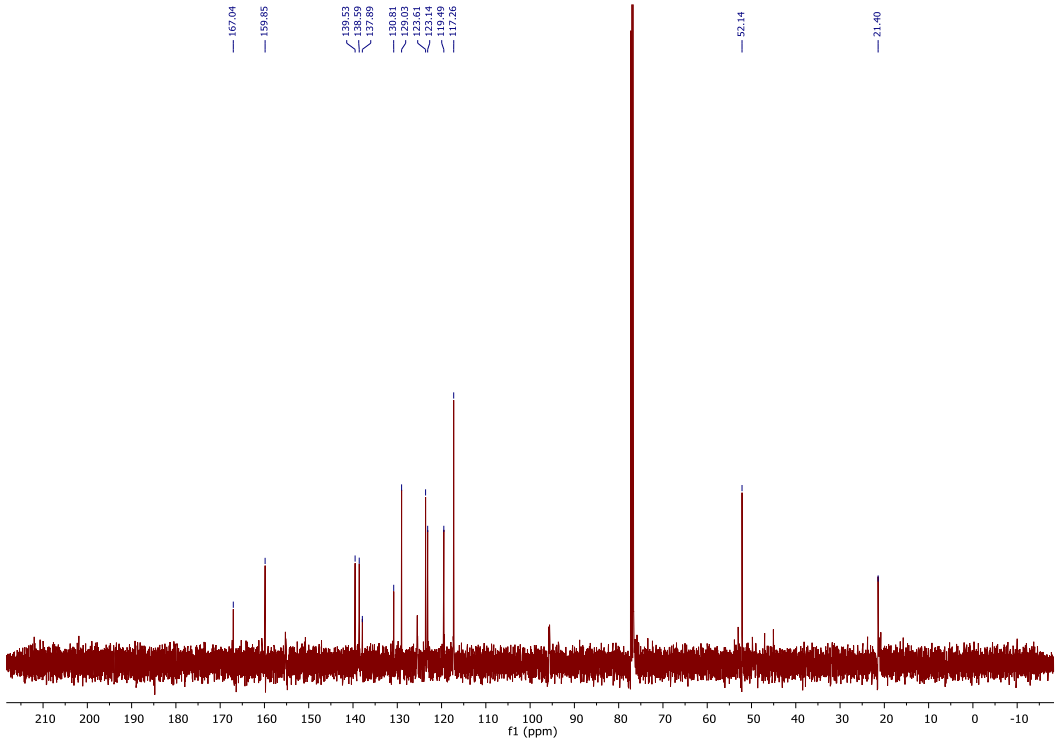
8-ii, ^{13}C, 126 MHz, $CDCl_3$



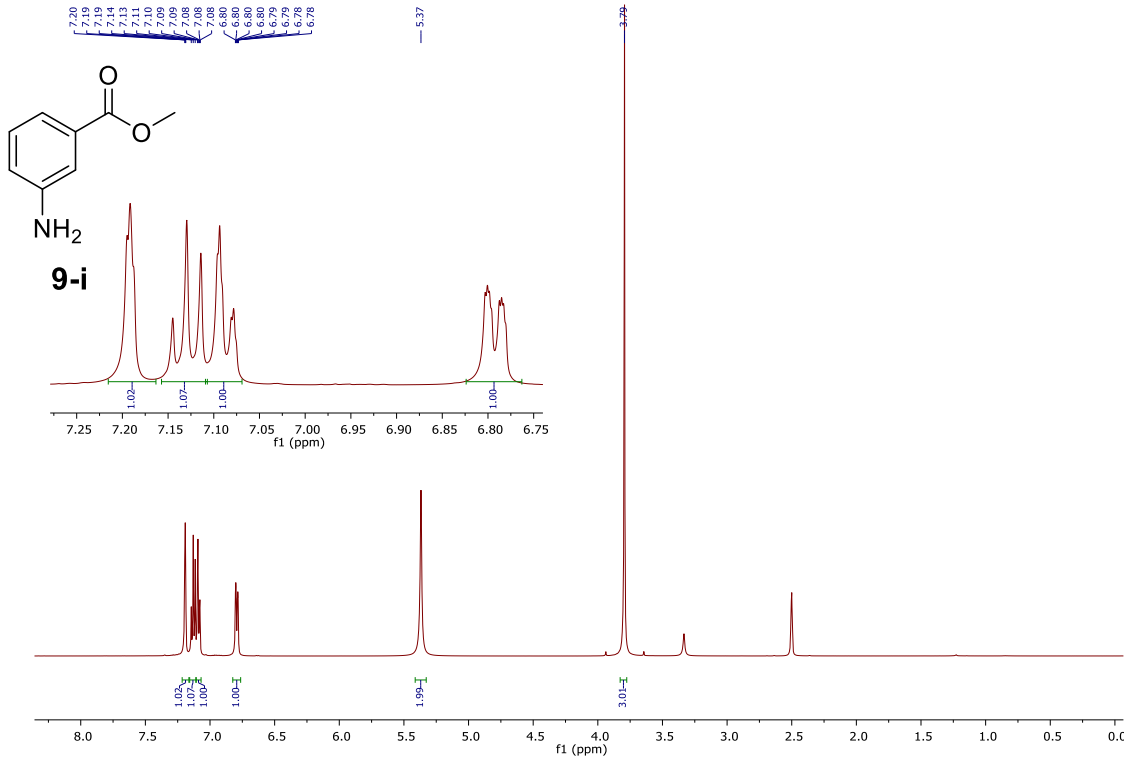
8, ^{1}H, 500 MHz, DMSO-d_6



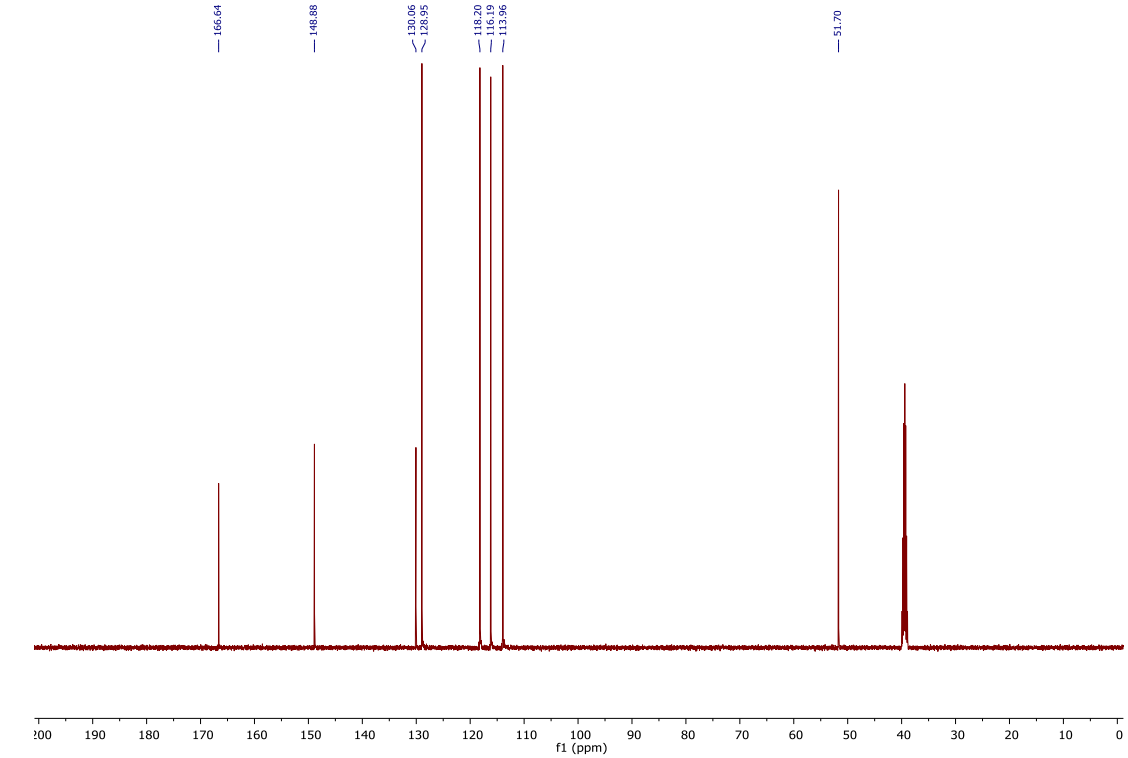
8, ^{13}C, 126 MHz, DMSO-d_6



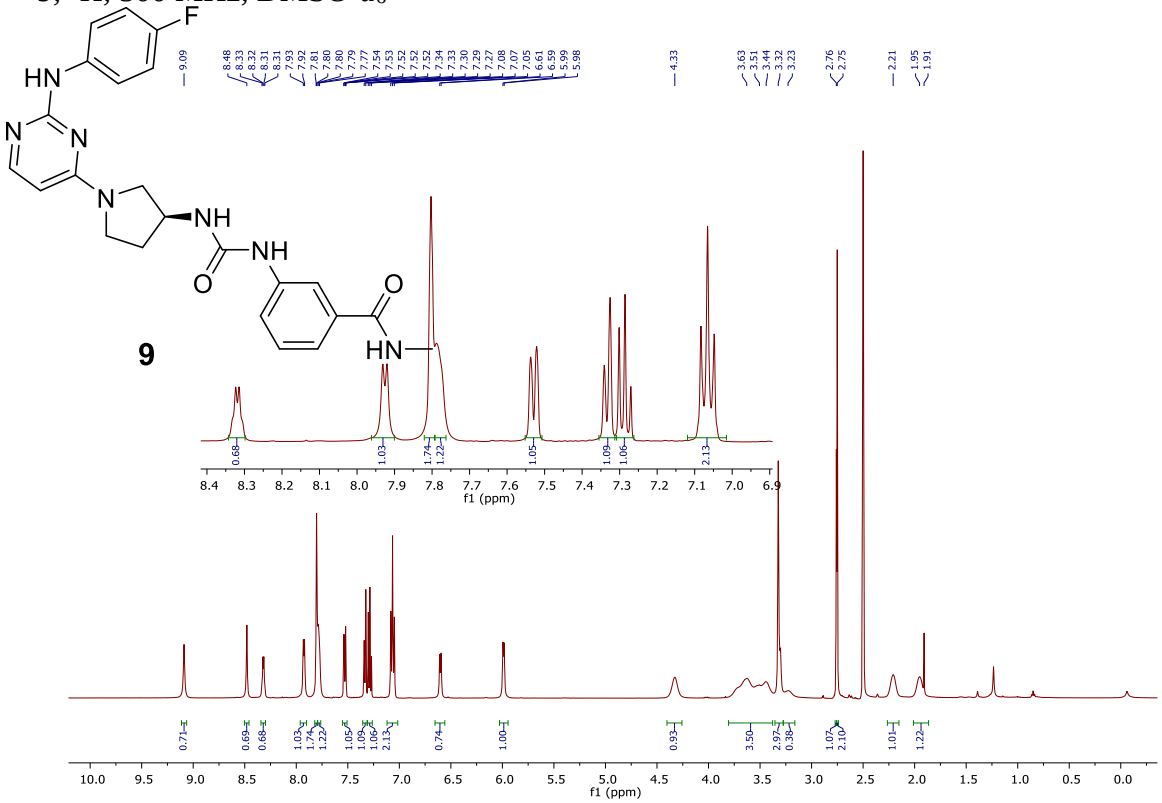
9-i, ^{1}H, 500 MHz, DMSO-d_6



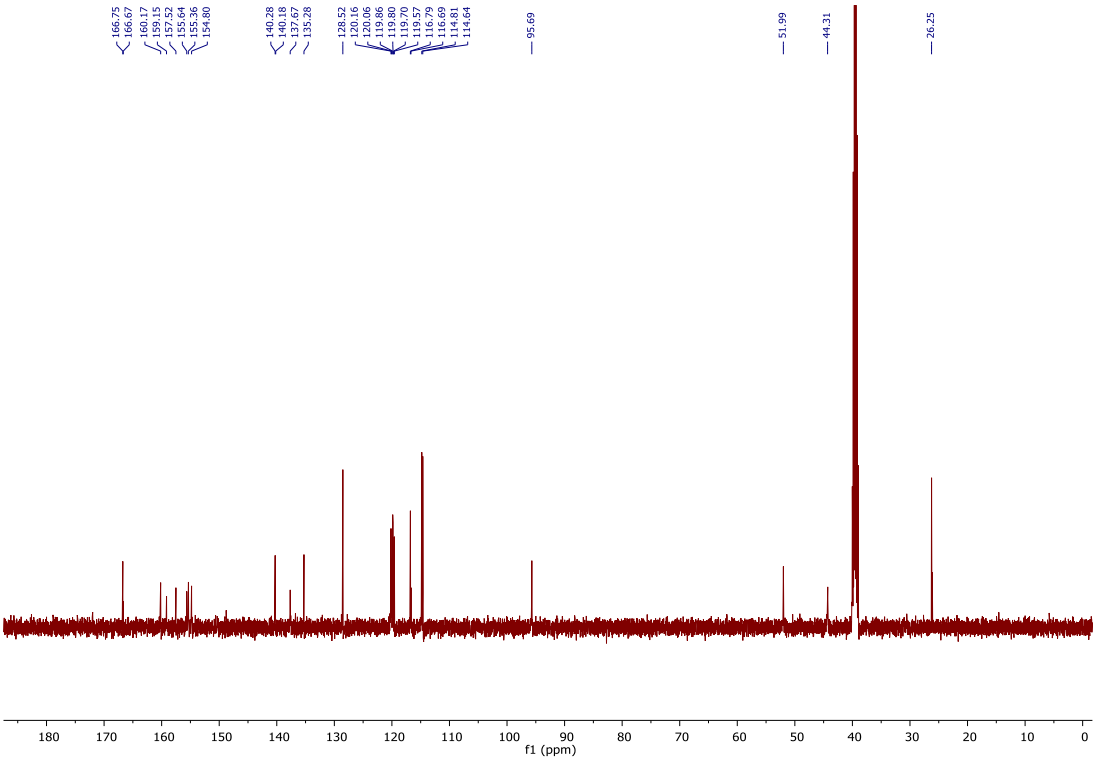
9-i, ^{13}C, 126 MHz, DMSO-d_6



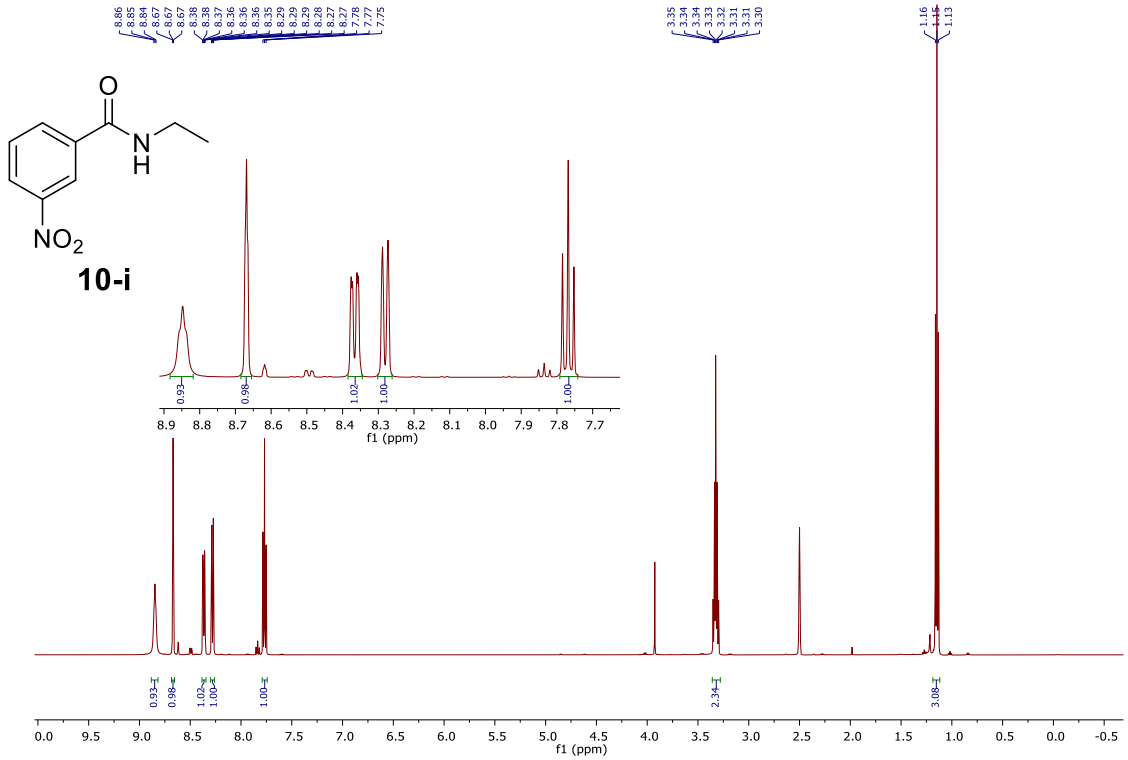
9, ^{1}H, 500 MHz, DMSO-d_6



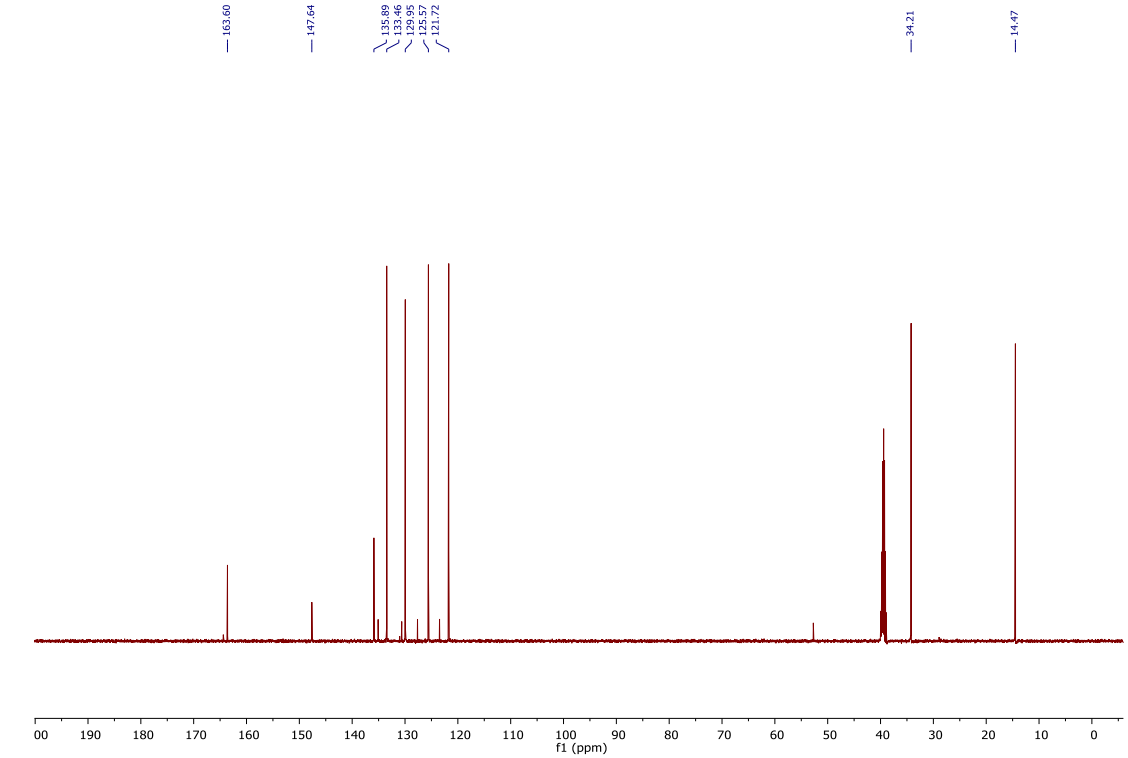
9, ^{13}C, 126 MHz, DMSO-d_6



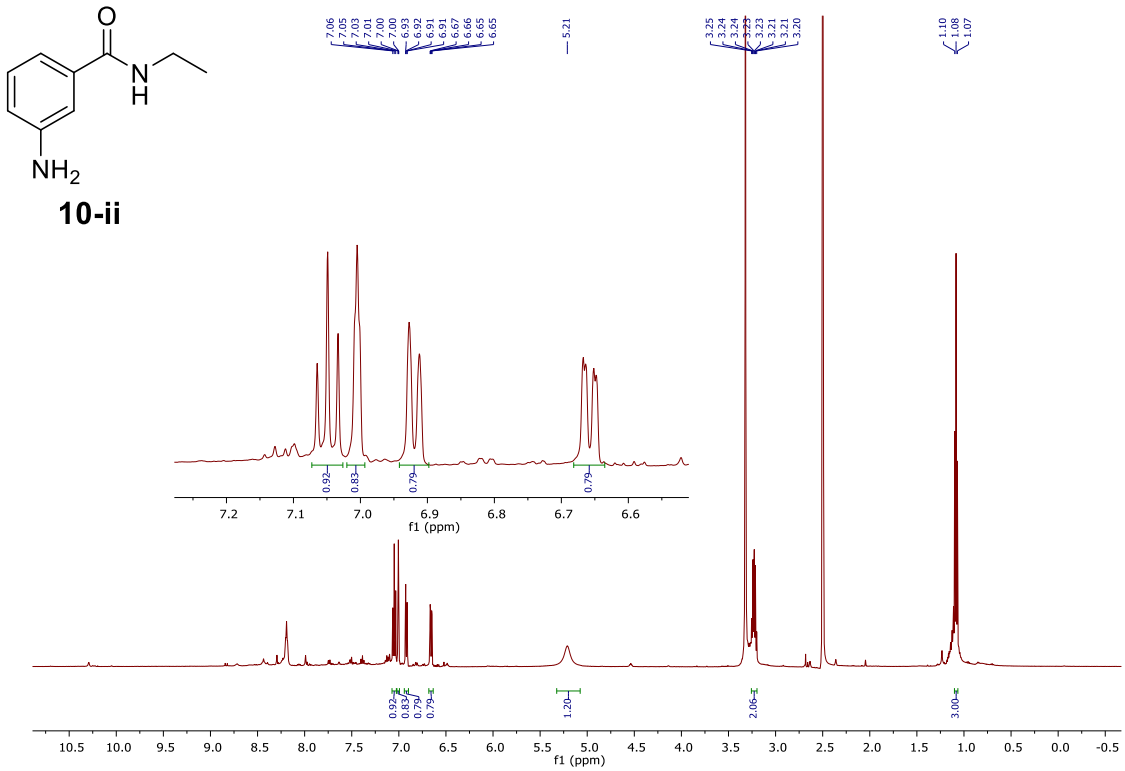
10-i, ^{1}H, 500 MHz, DMSO-d_6



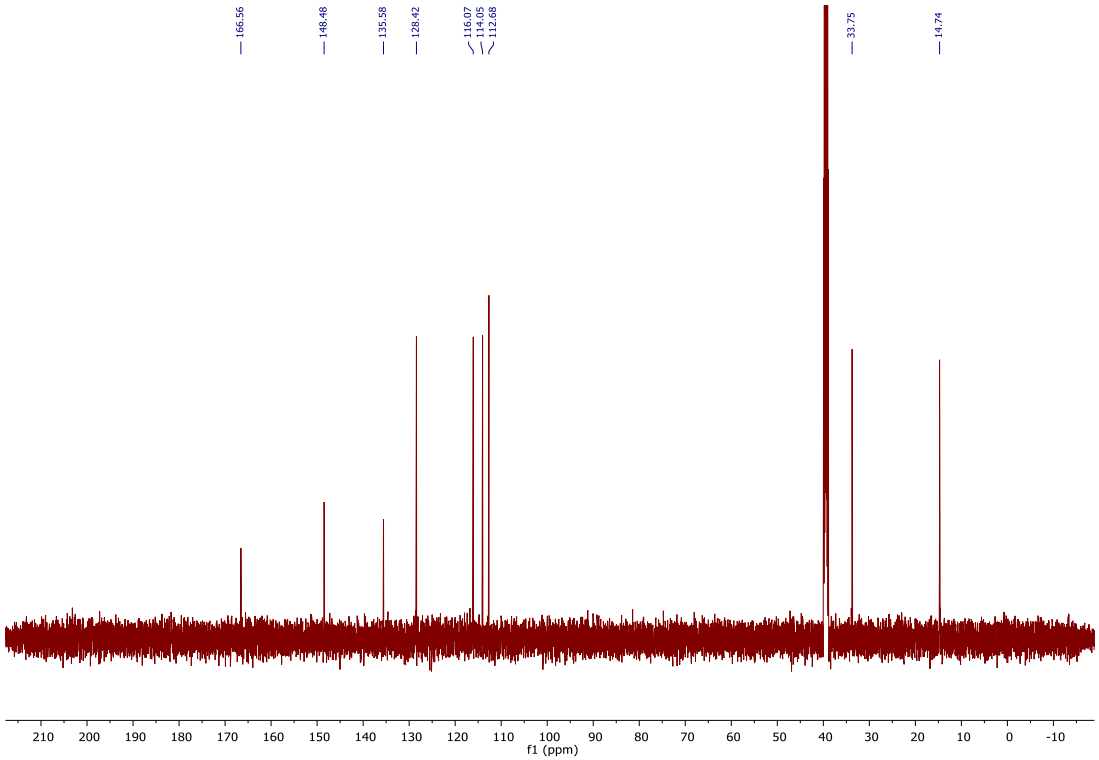
10-i, ^{13}C, 126 MHz, DMSO-d_6



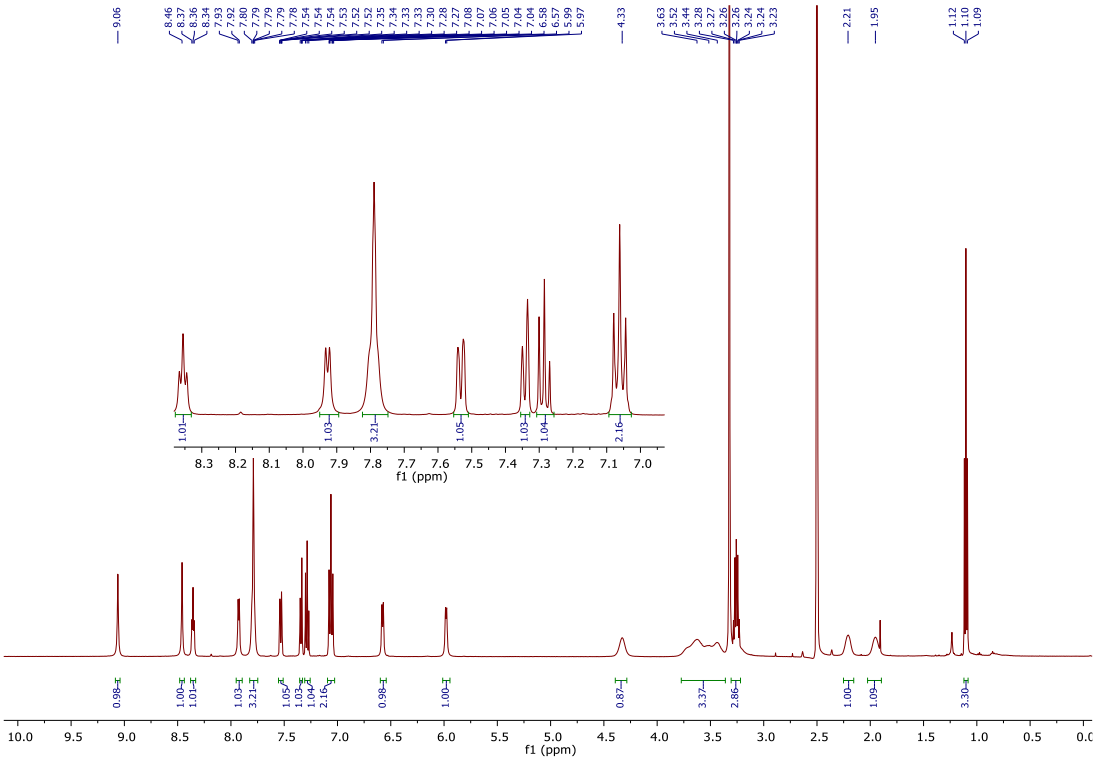
10-ii, ^{1}H, 500 MHz, DMSO-d_6



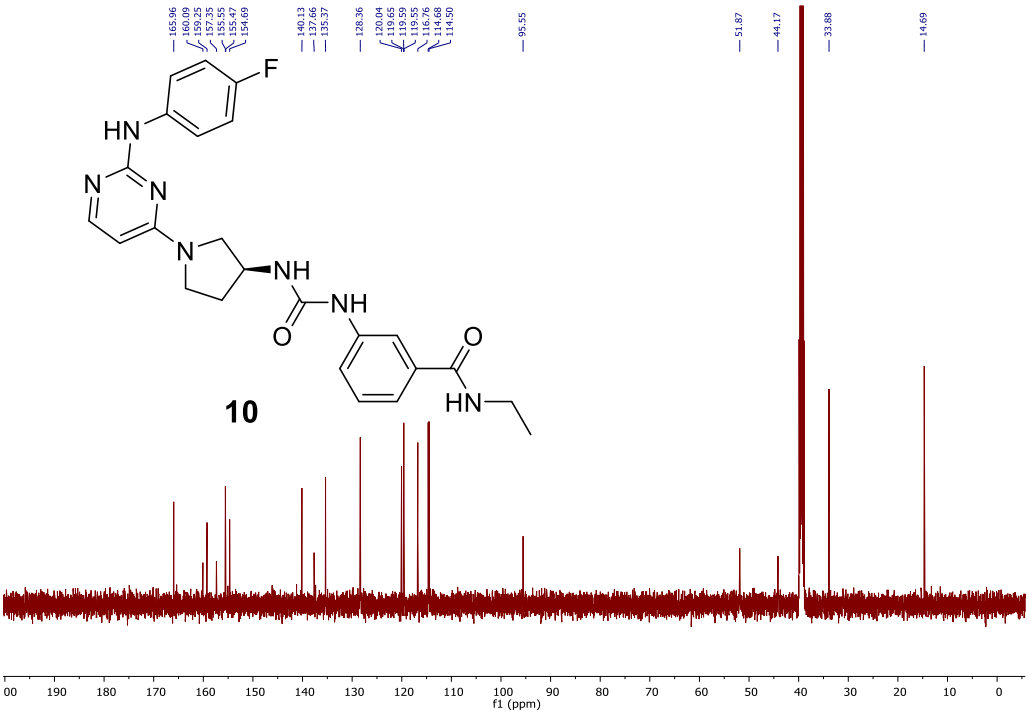
10-ii, ^{13}C, 126 MHz, DMSO-d_6



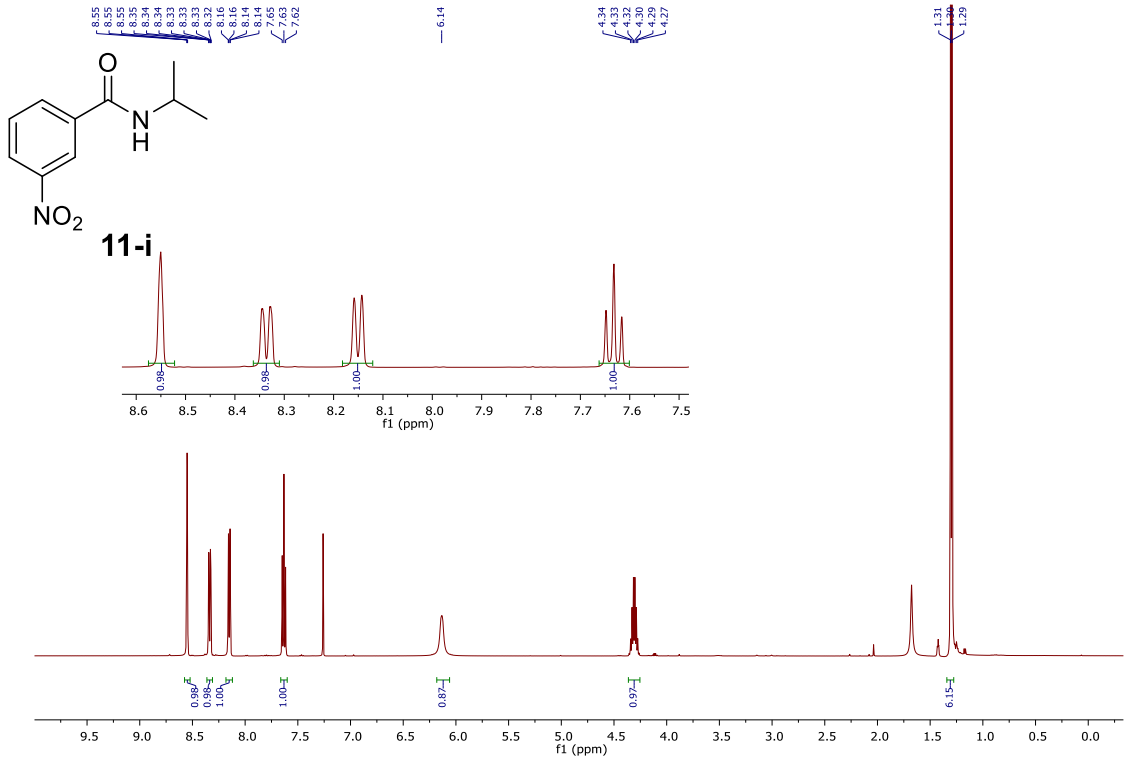
10, ^{1}H, 500 MHz, DMSO-d_6



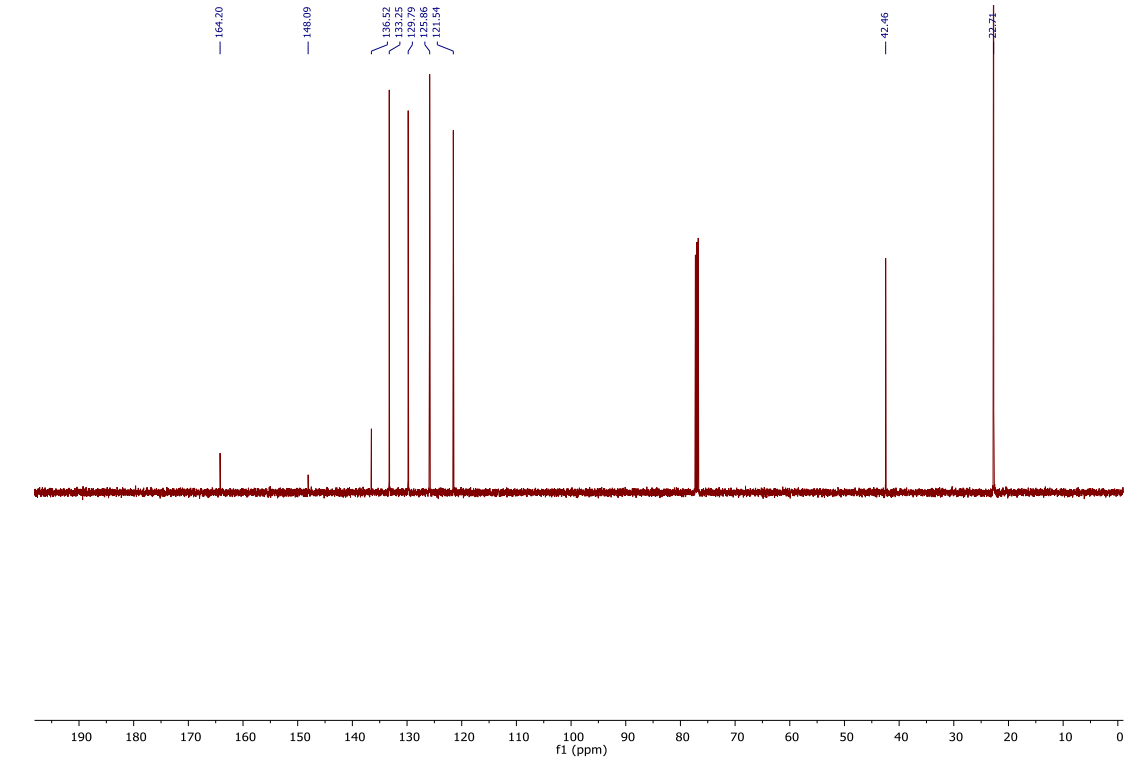
10, ^{13}C, 126 MHz, DMSO-d_6



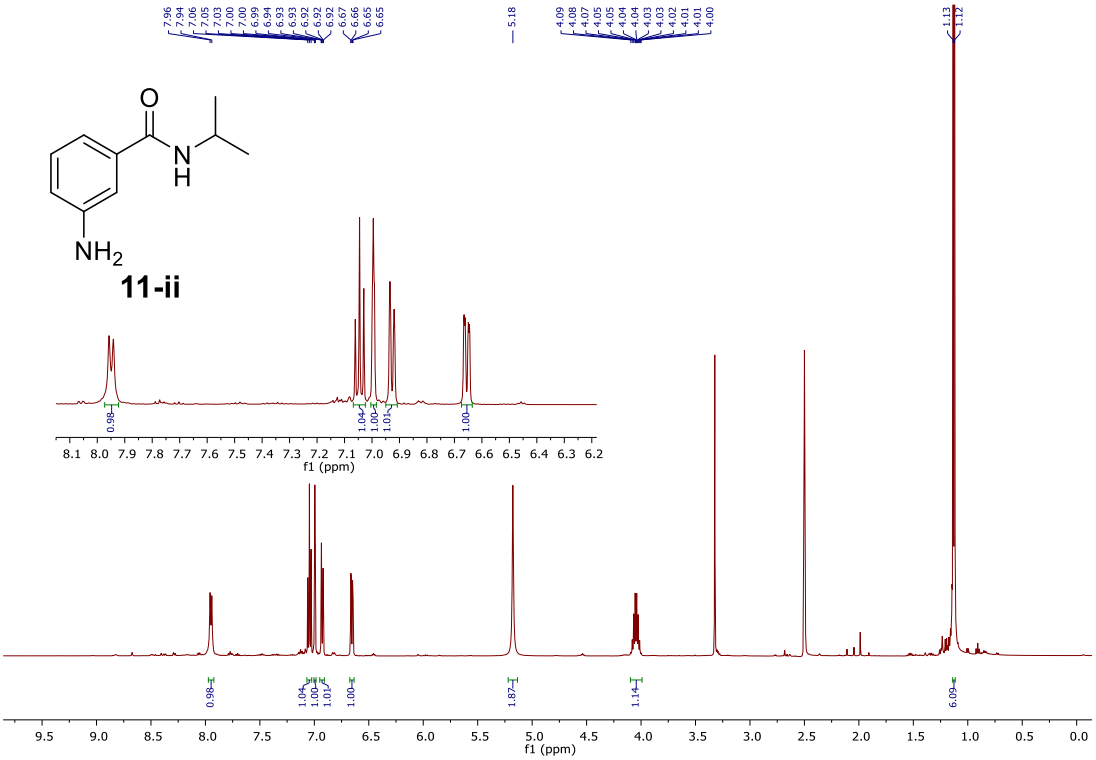
11-i, ^{1}H, 500 MHz, $CDCl_3$



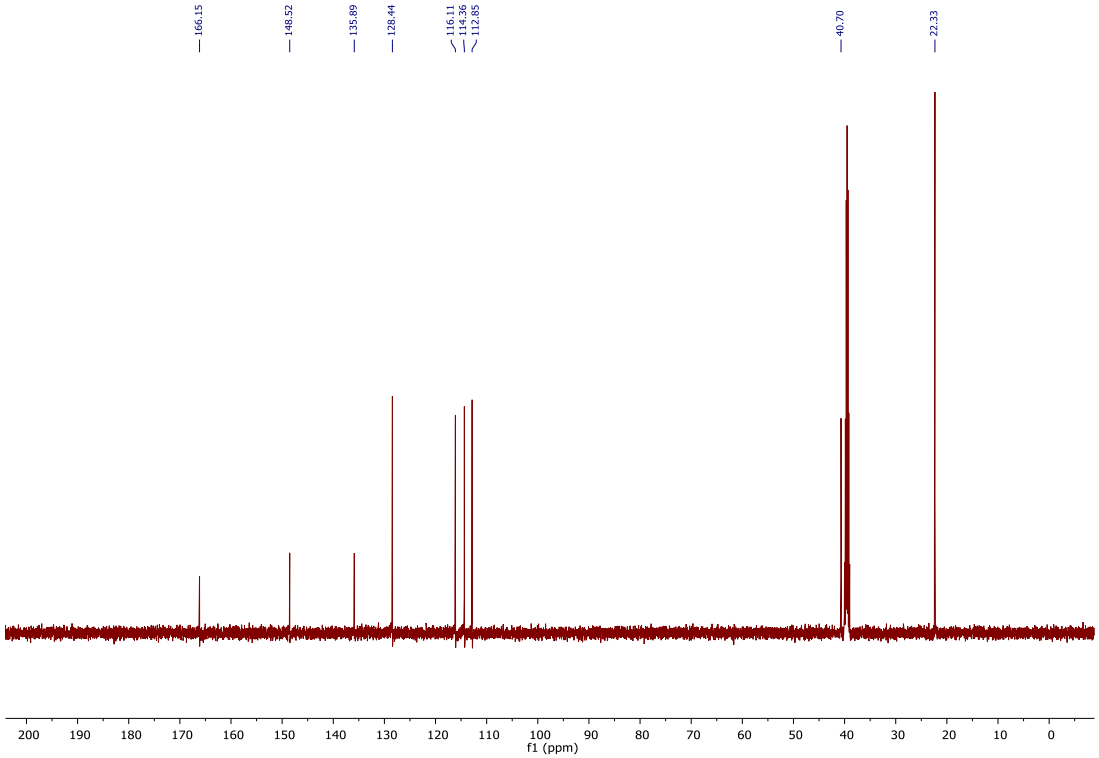
11-i, ^{13}C, 126 MHz, $CDCl_3$



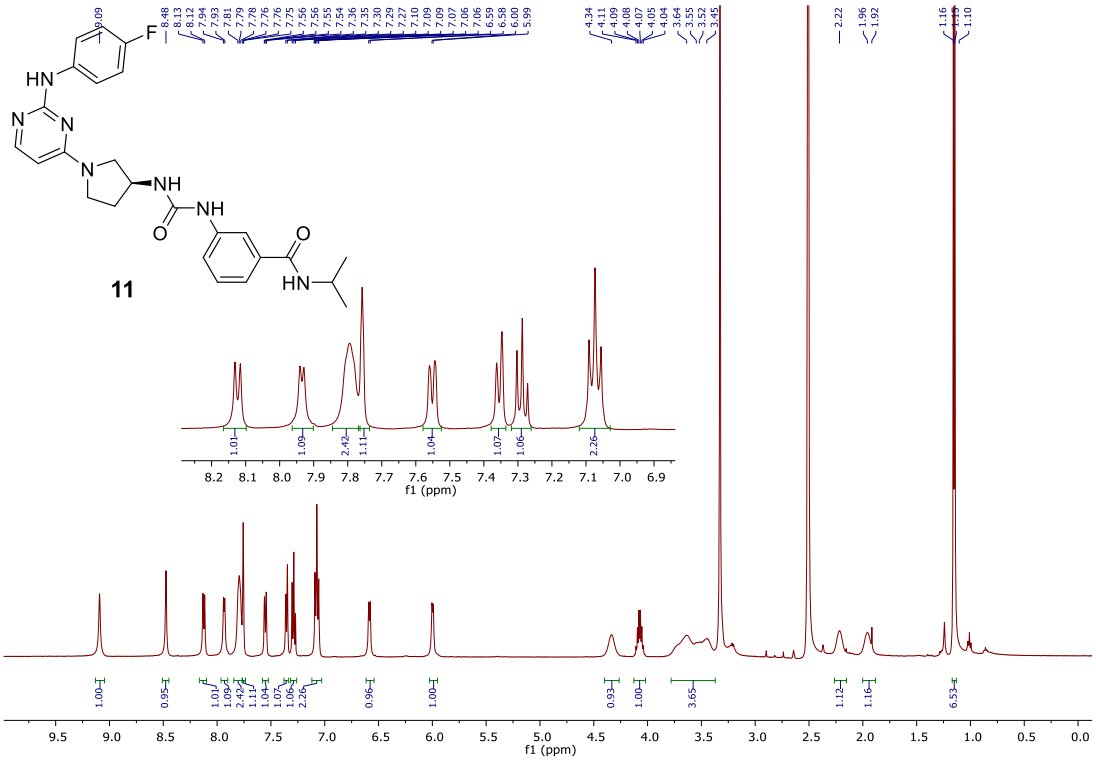
11-ii, ^{1}H, 500 MHz, DMSO-d_6



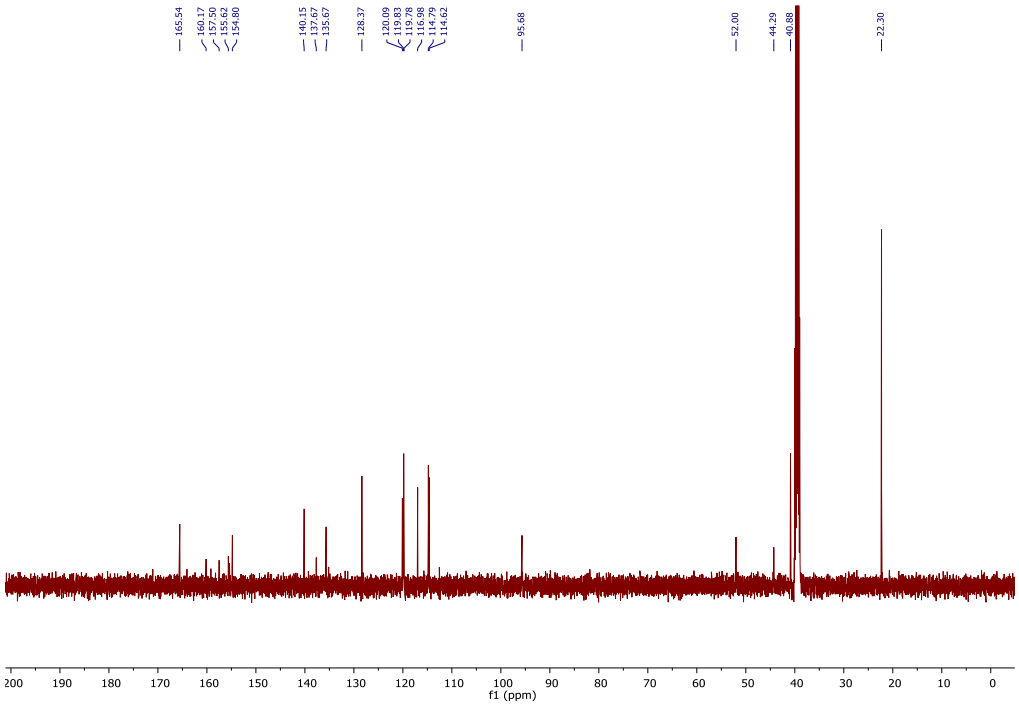
11-ii, ^{13}C, 126 MHz, DMSO-d_6



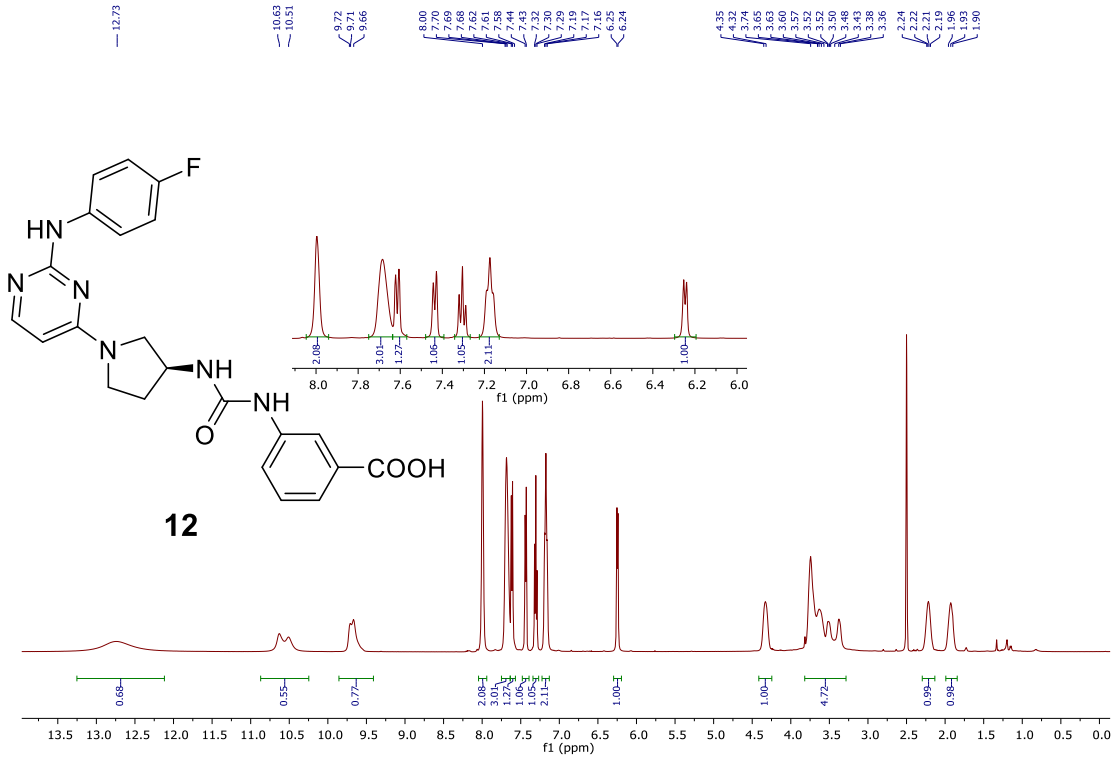
11, ^{1}H, 500 MHz, DMSO-d_6



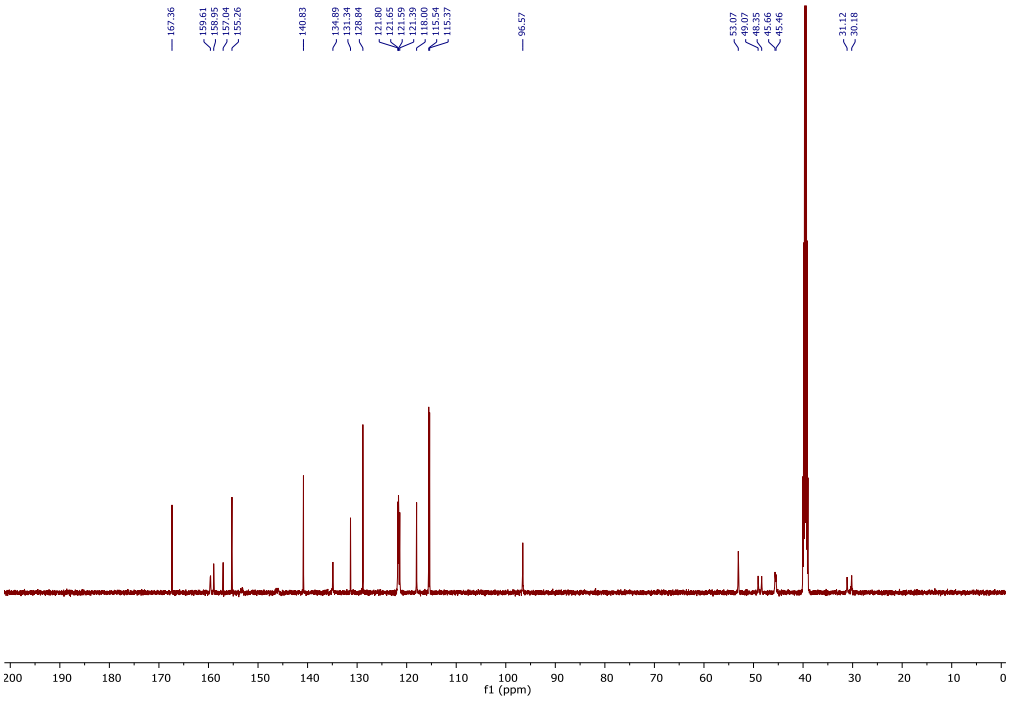
11, ^{13}C, 126 MHz, DMSO-d_6



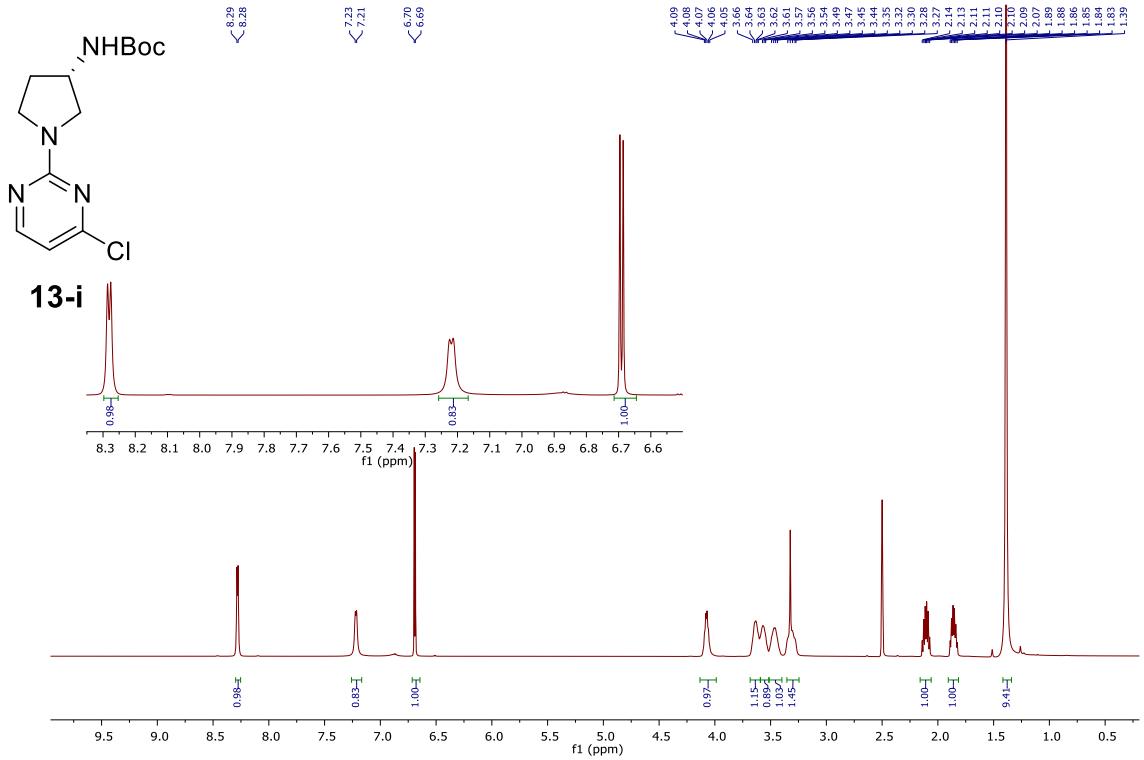
12, ^{1}H, 500 MHz, DMSO-d_6



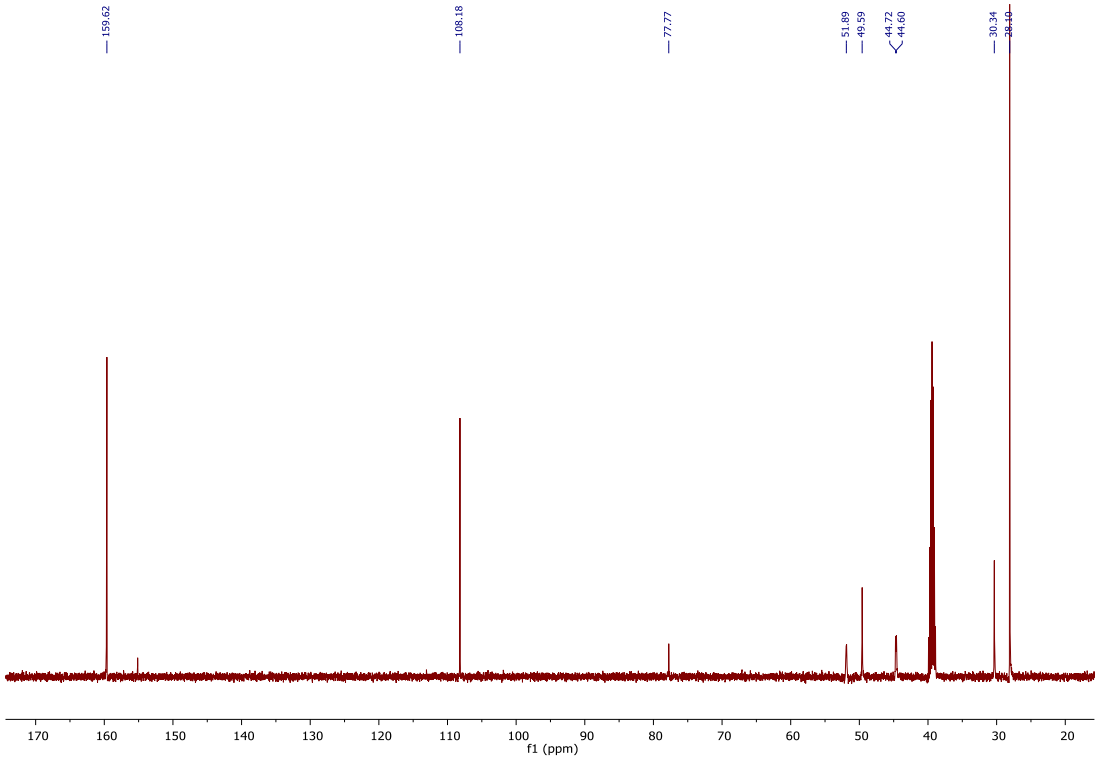
12, ^{13}C, 126 MHz, DMSO-d_6



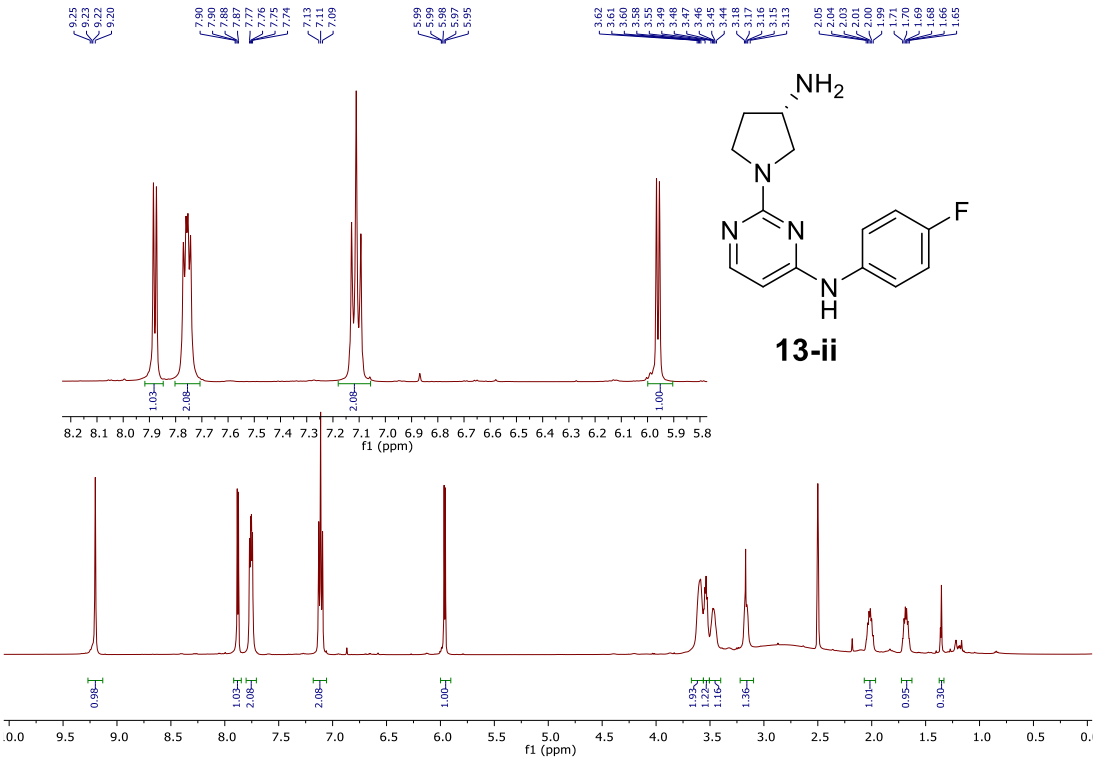
13-i, ^{1}H, 500 MHz, DMSO-d_6



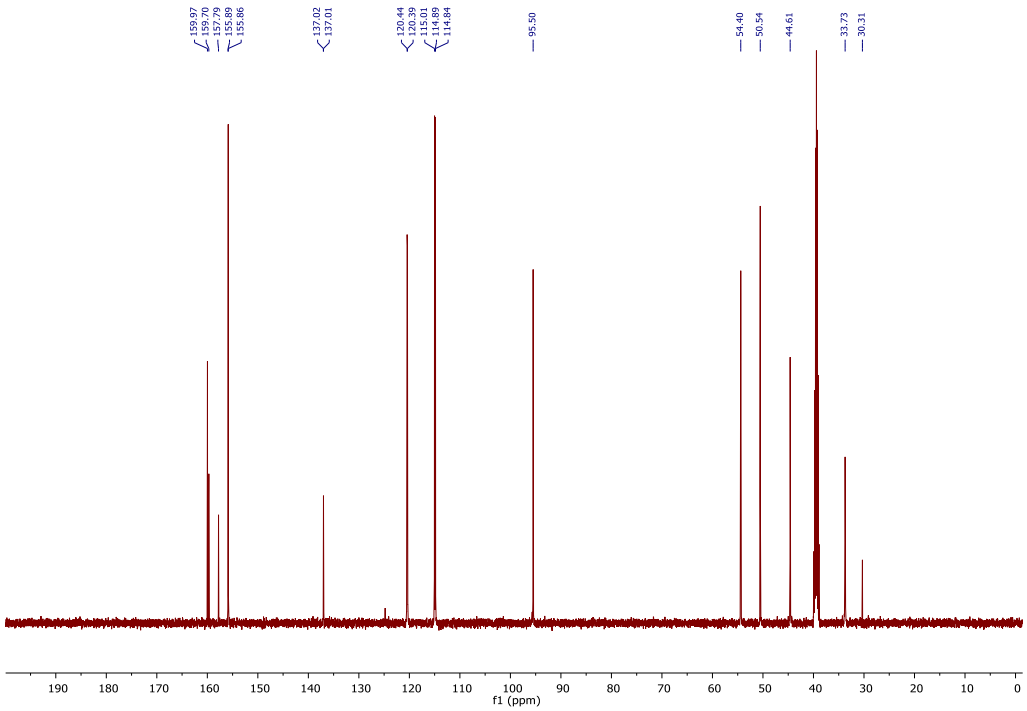
13-i, ^{13}C, 126 MHz, DMSO-d_6



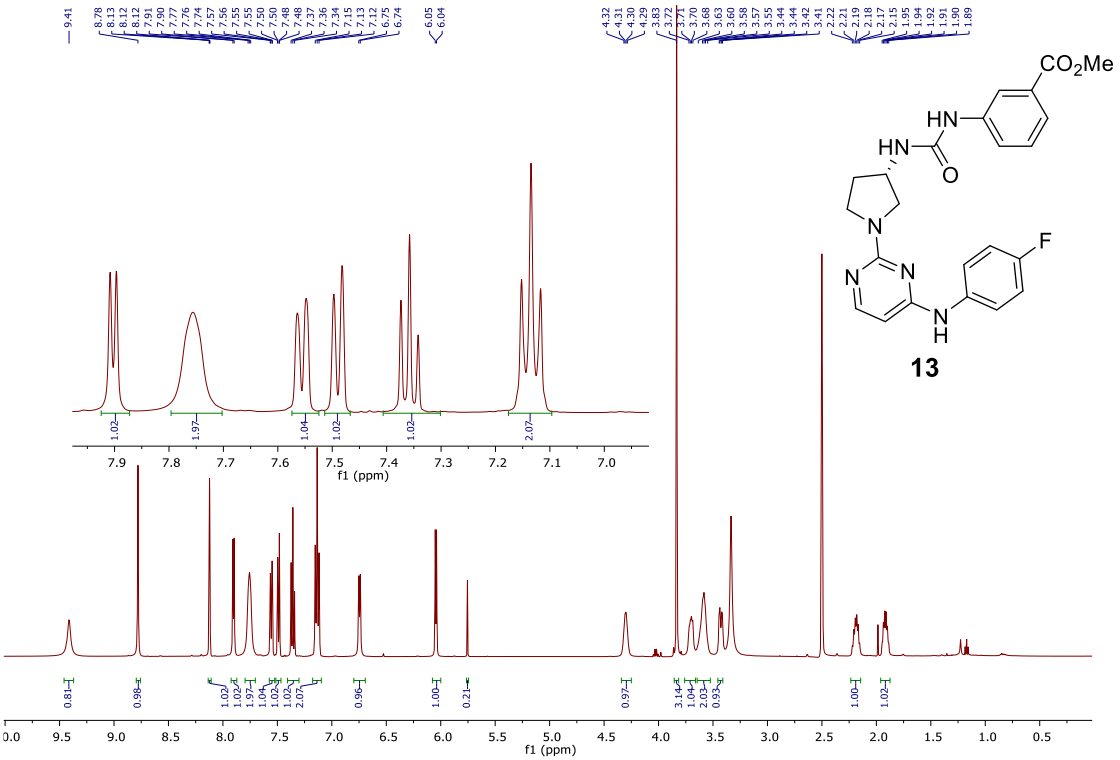
13-ii, ^{1}H, 500 MHz, DMSO-d_6



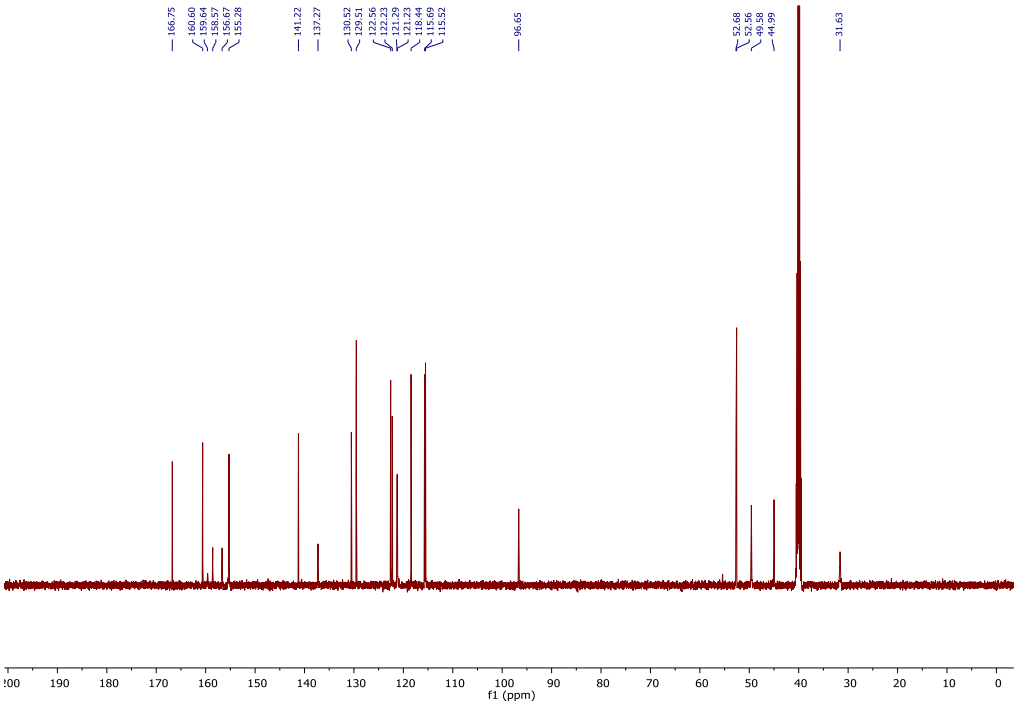
13-ii, ^{13}C, 126 MHz, DMSO-d_6



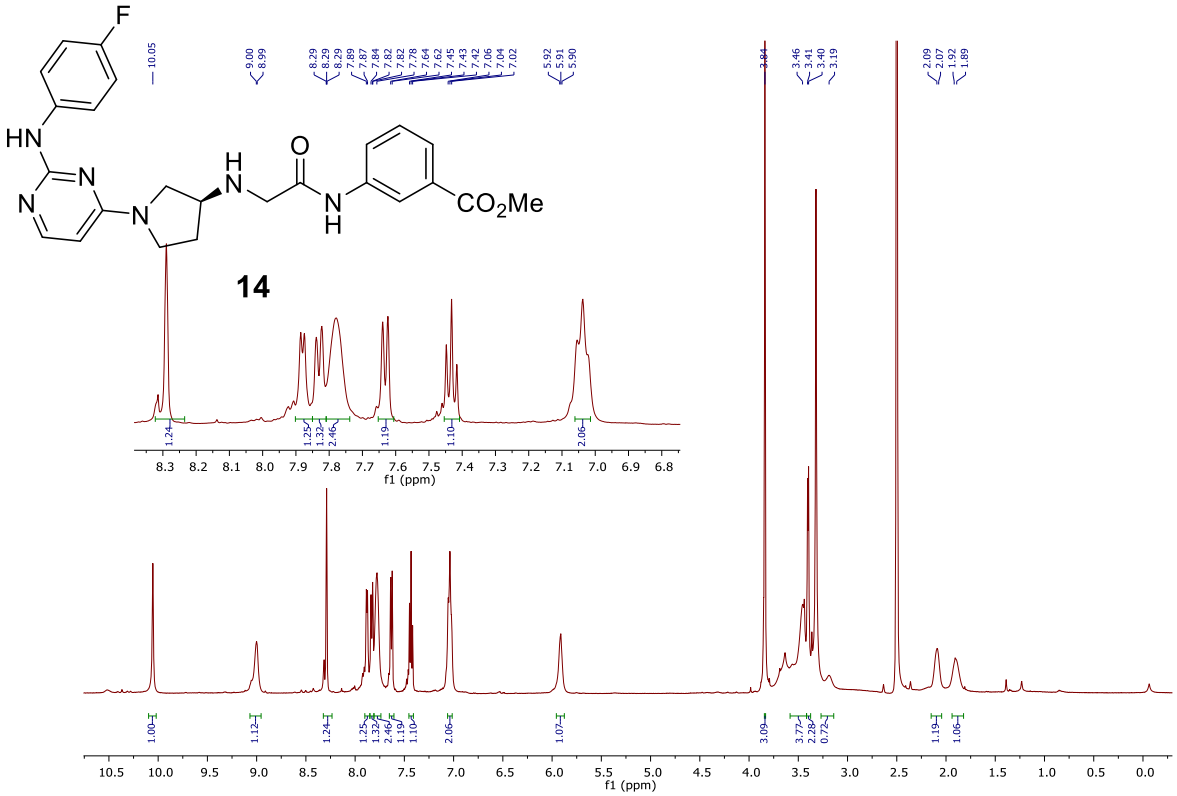
13, ^{1}H, 500 MHz, DMSO-d_6



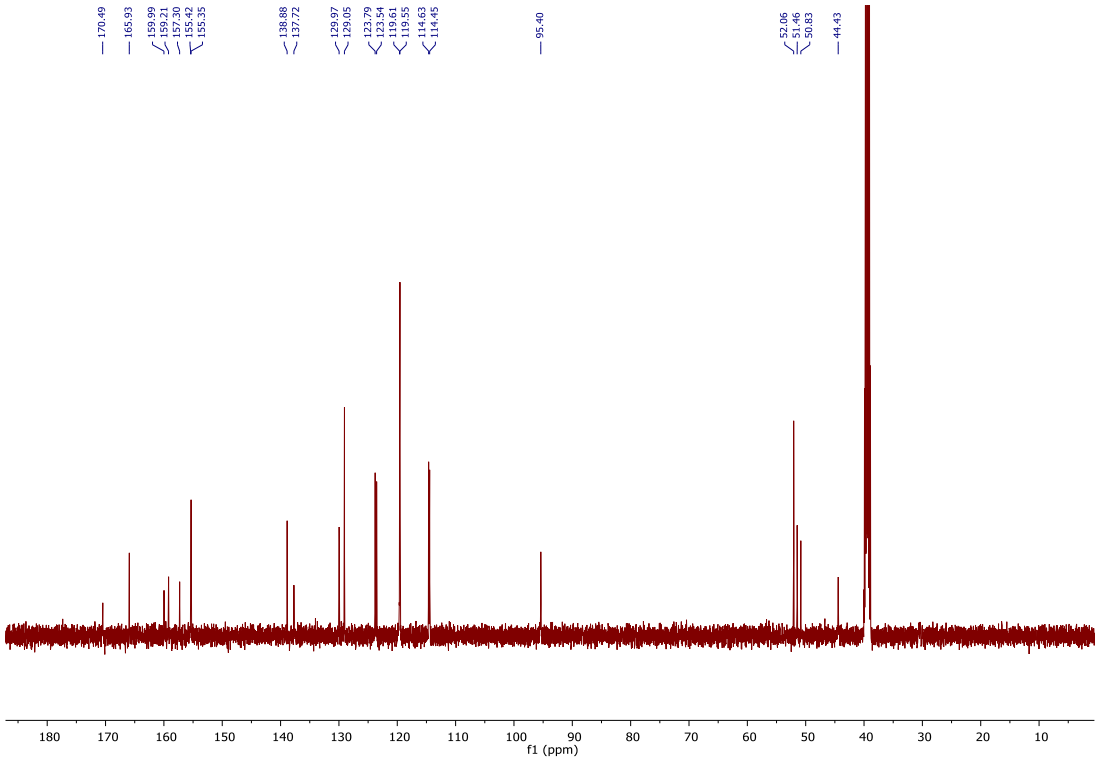
13, ^{13}C, 126 MHz, DMSO-d_6



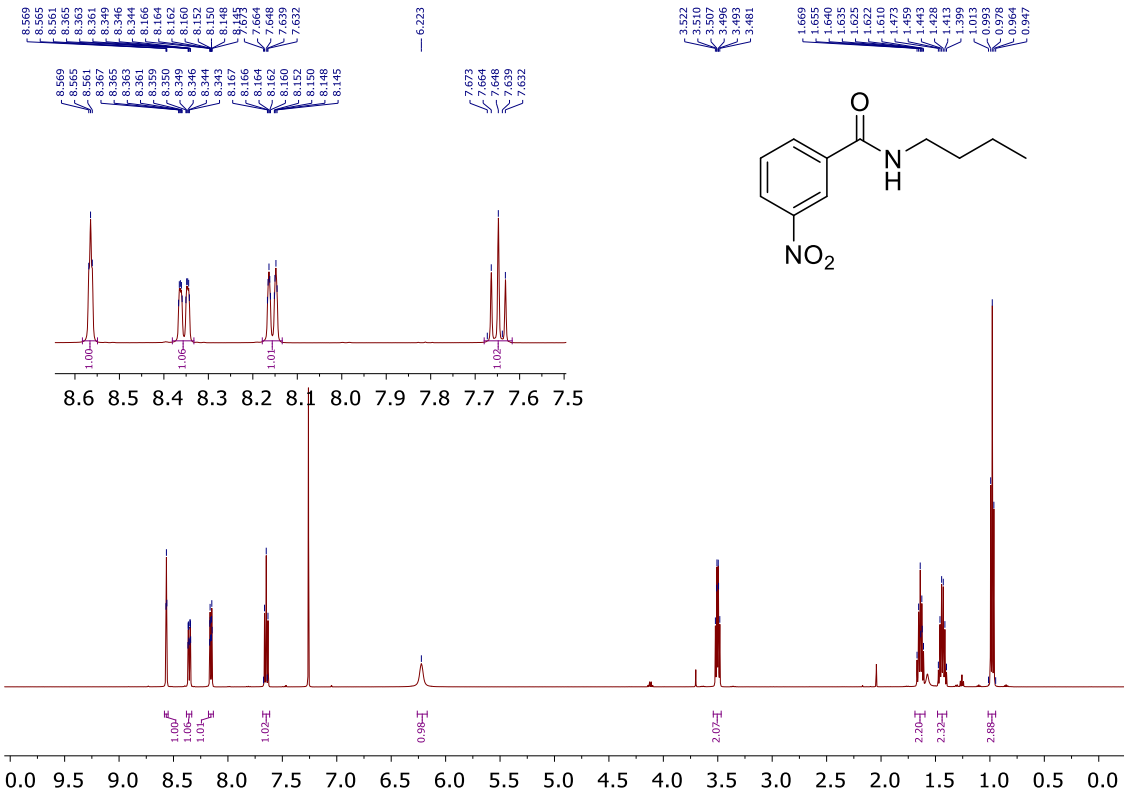
14, ^{1}H, 500 MHz, DMSO-d_6



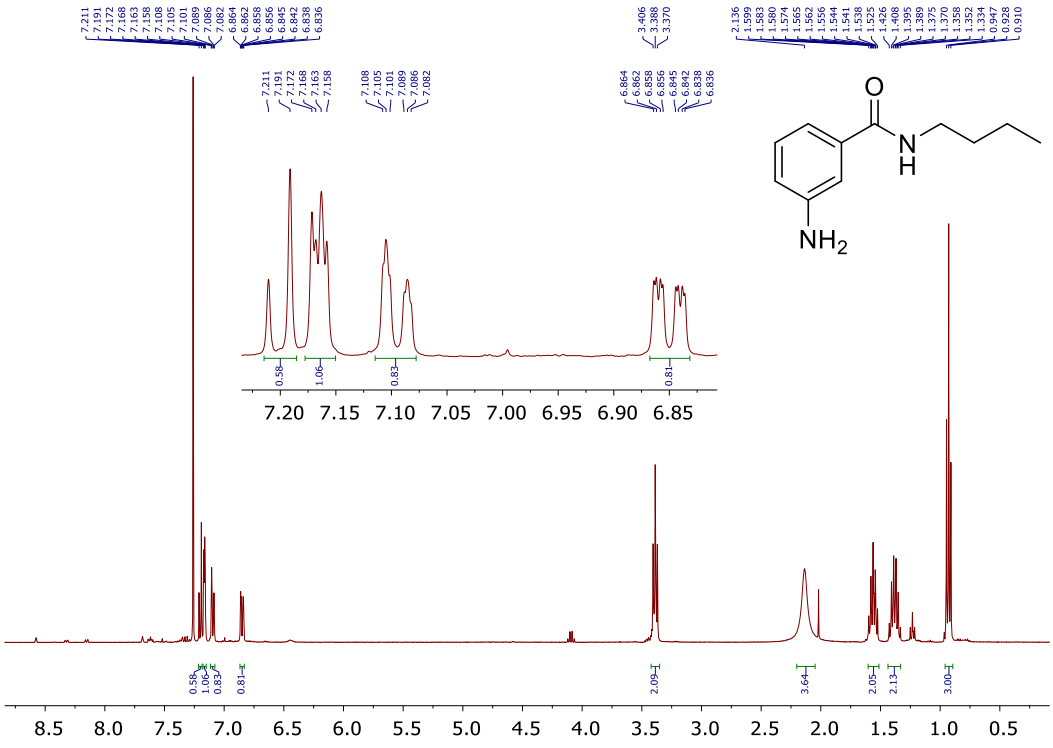
14, ^{13}C, 126 MHz, DMSO-d_6



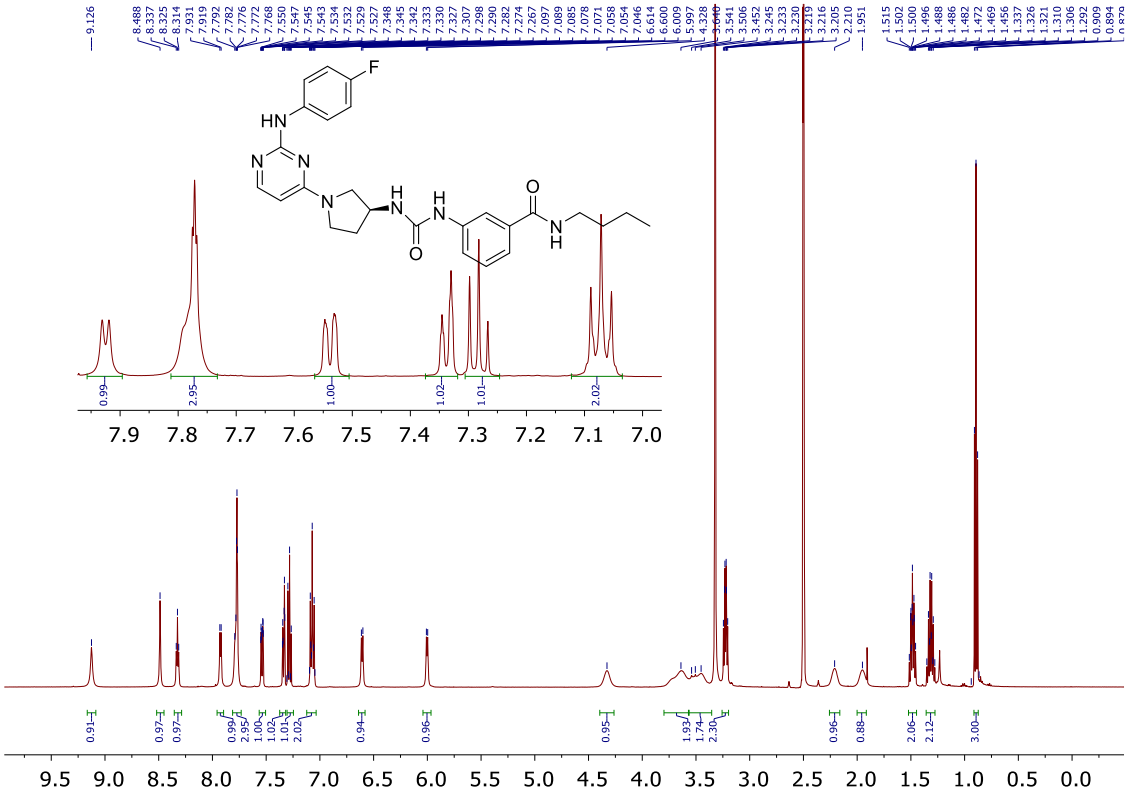
15-i, ^{1}H, 500 MHz, $CDCl_3$



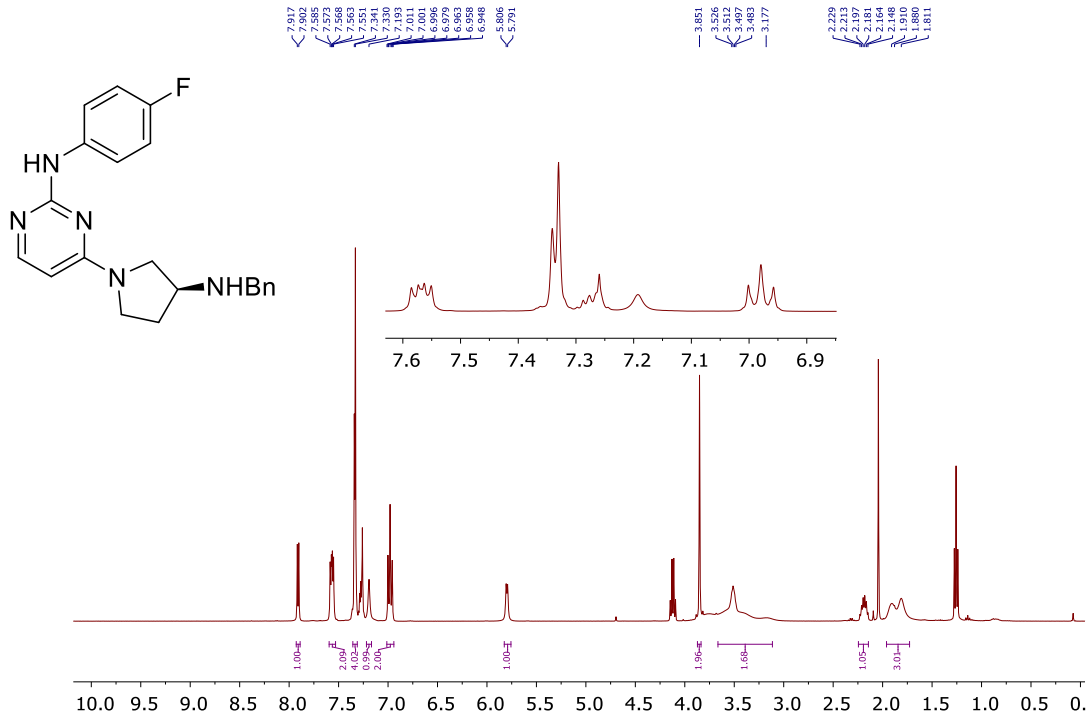
15-ii, ^{1}H, 500 MHz, $CDCl_3$



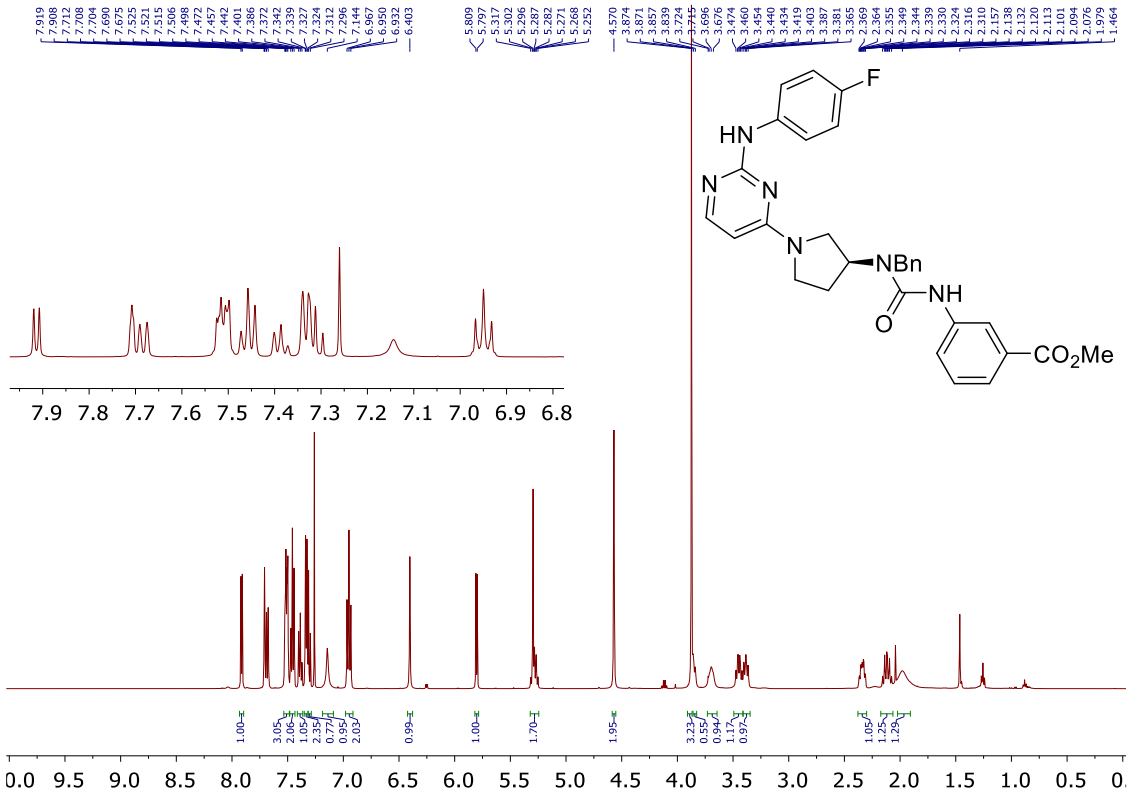
15, ^{1}H, 500 MHz, DMSO-d_6



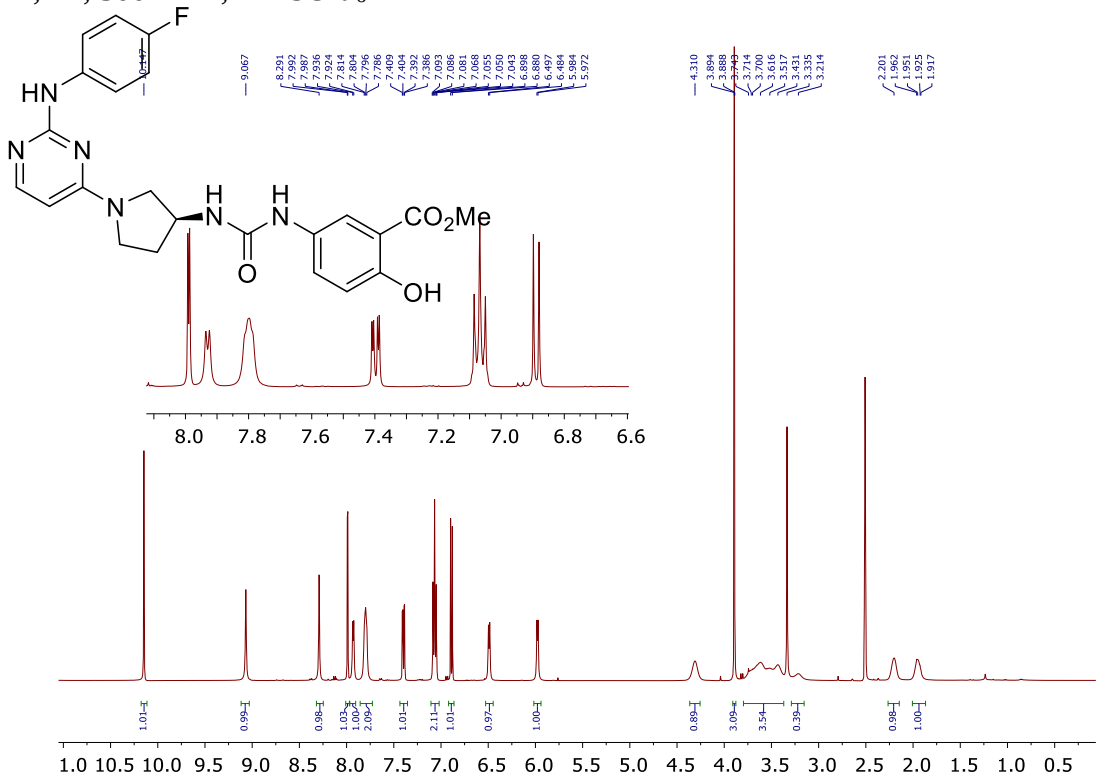
16-i, ^{1}H, 500 MHz, $CDCl_3$



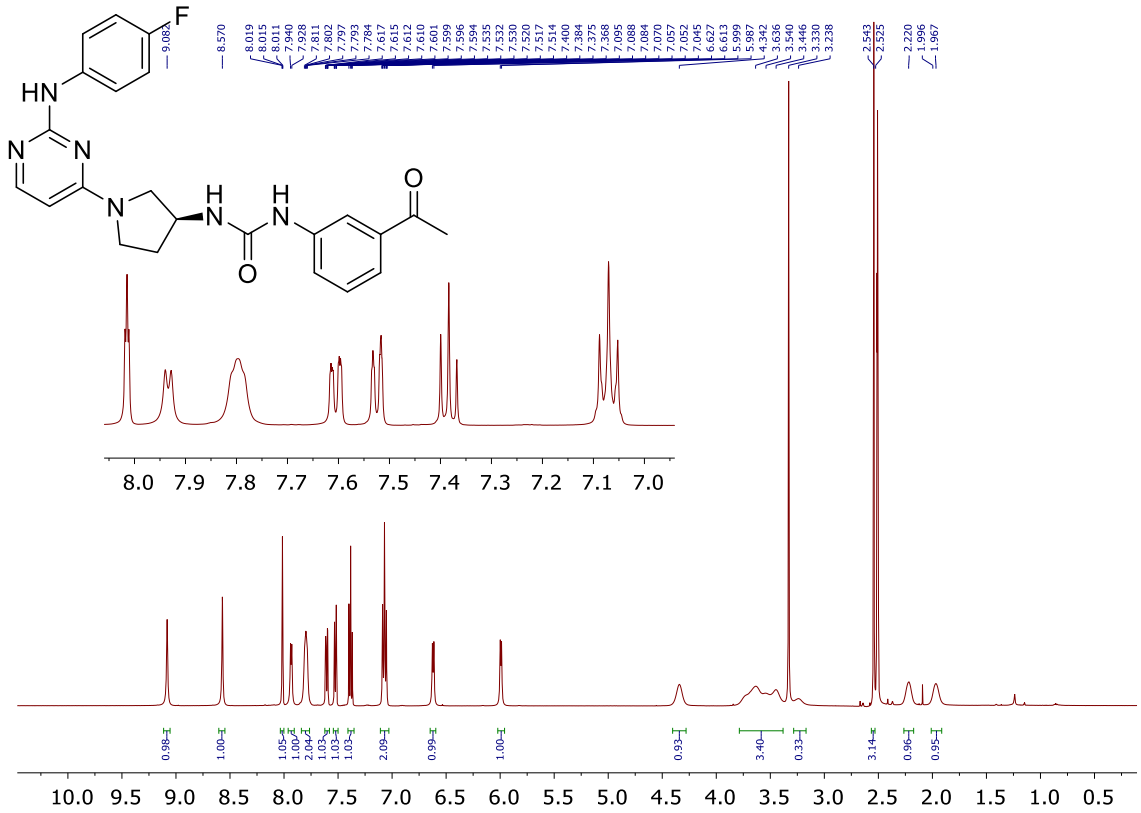
16, ^{1}H, 500 MHz, $CDCl_3$



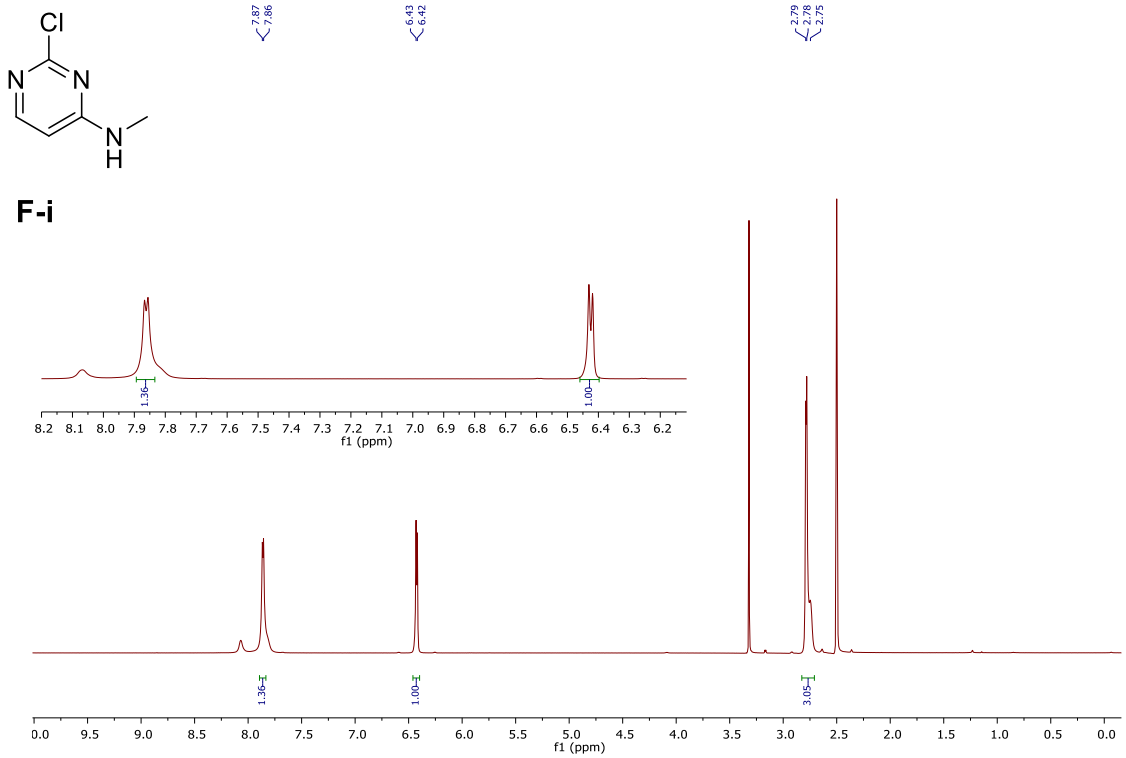
17, ^{1}H, 500 MHz, DMSO-d_6



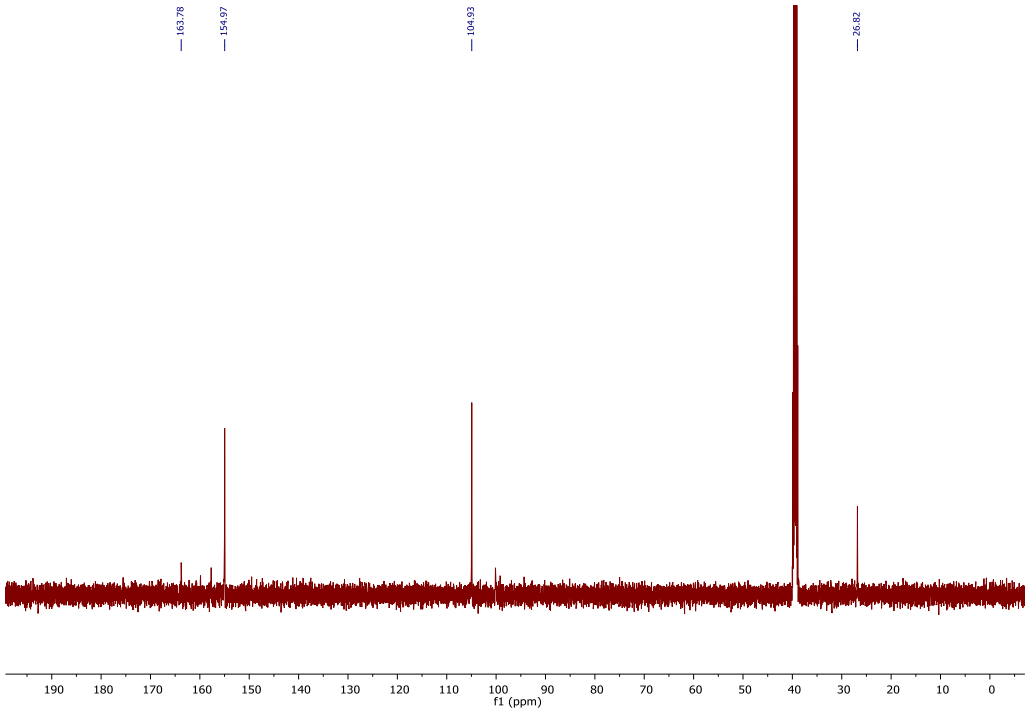
18, ^{1}H, 500 MHz, DMSO-d_6



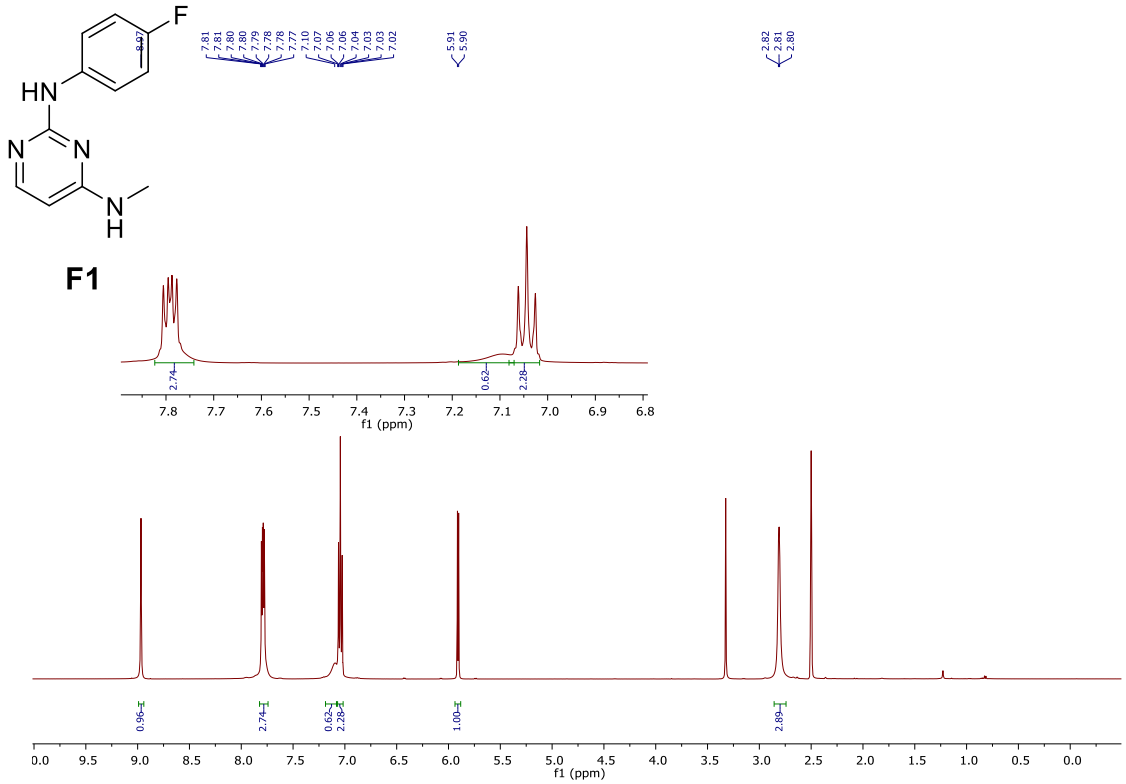
F-i, ^{1}H, 500 MHz, DMSO-d_6



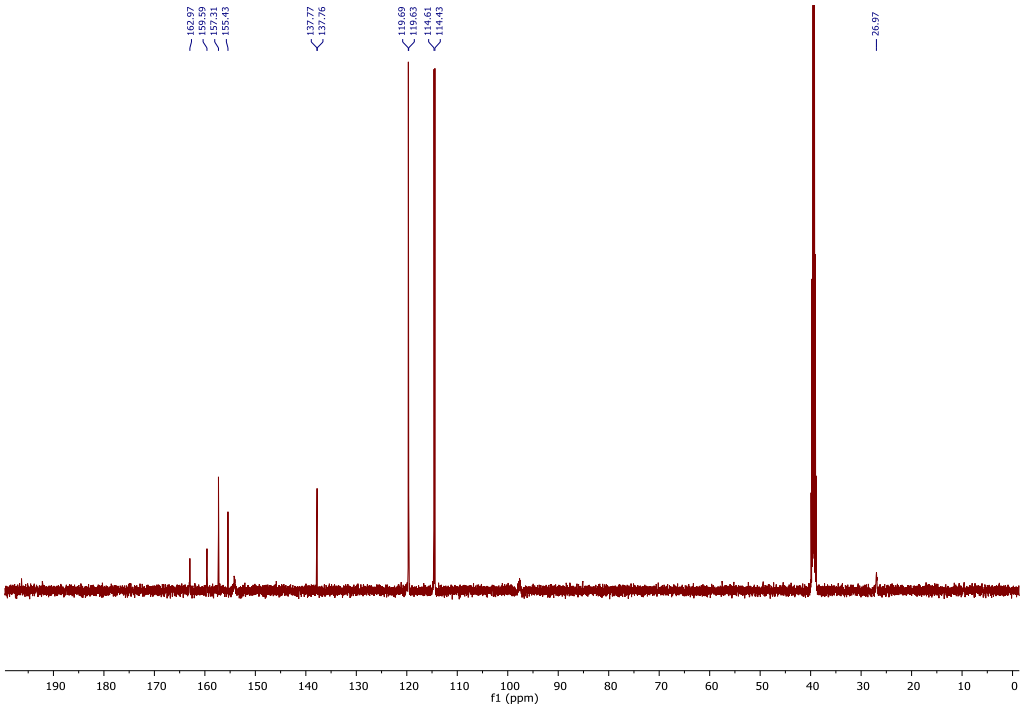
F-i, ^{13}C, 126 MHz, DMSO-d_6



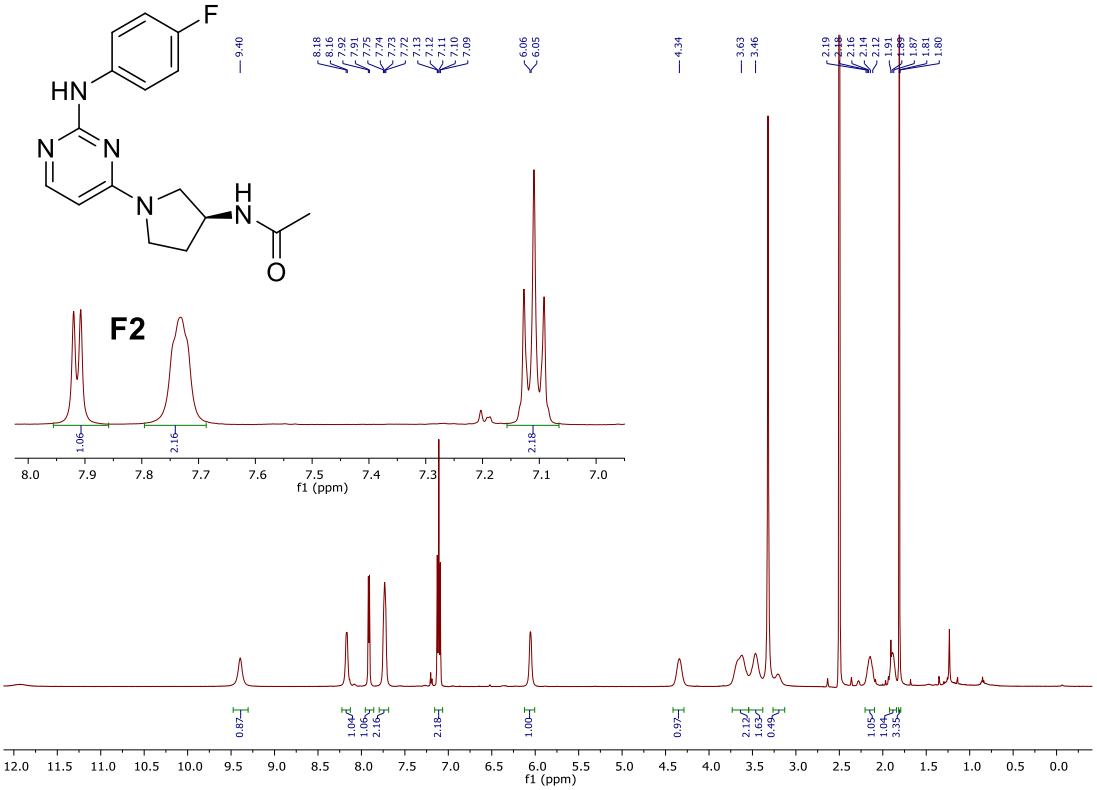
F1, ^{1}H, 500 MHz, DMSO-d_6



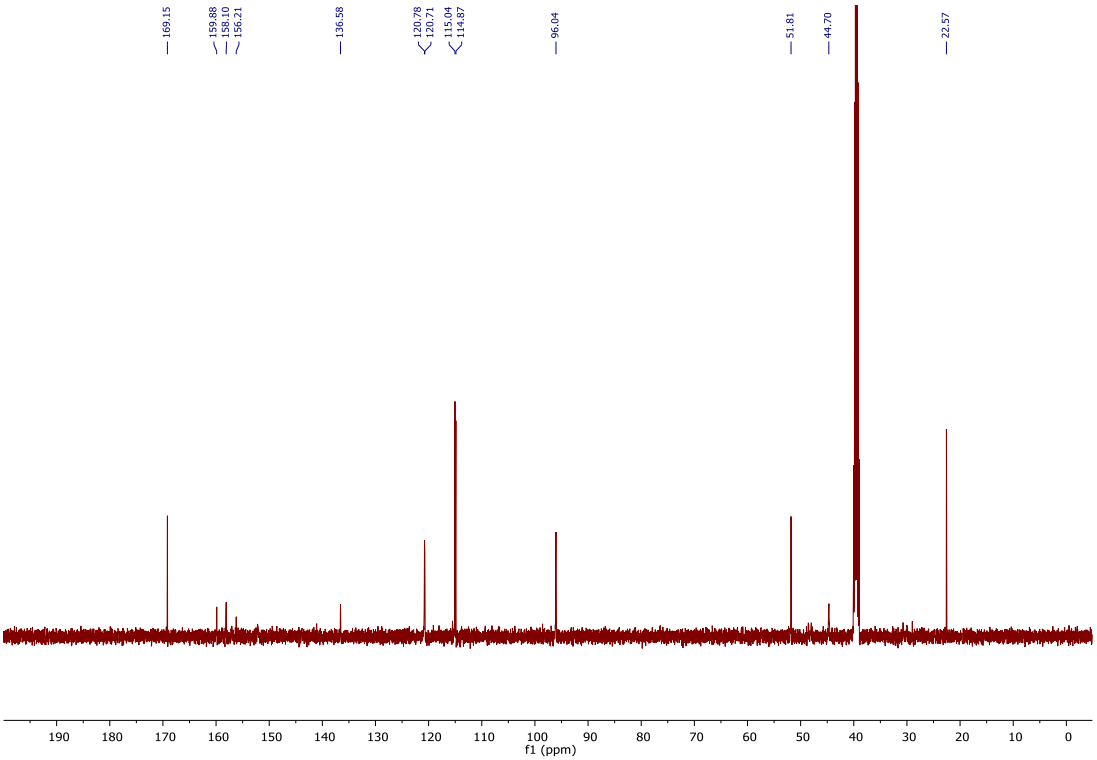
F1, ^{13}C, 126 MHz, DMSO-d_6



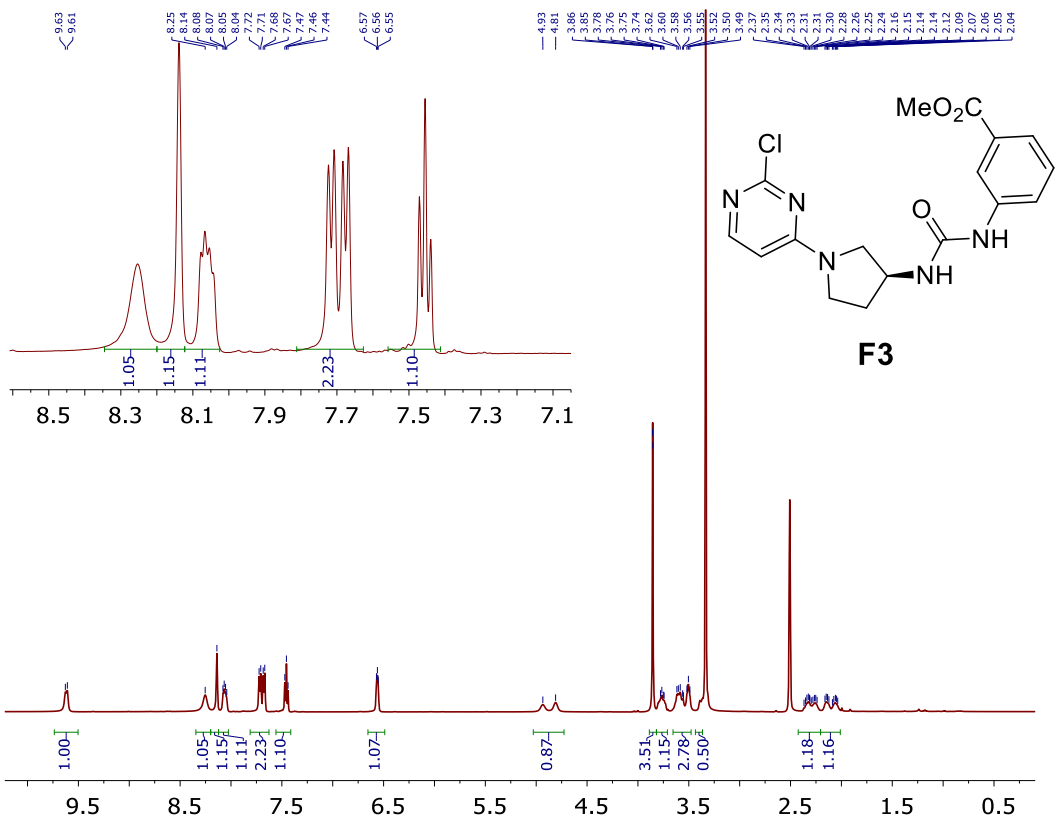
F2, ^{1}H, 500 MHz, DMSO-d_6



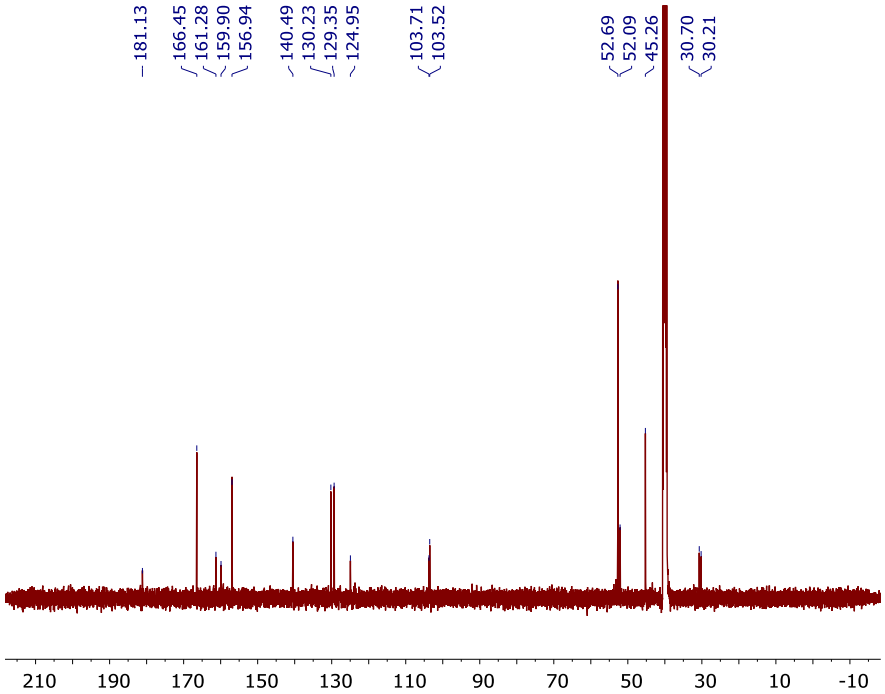
F2, ^{13}C, 126 MHz, DMSO-d_6



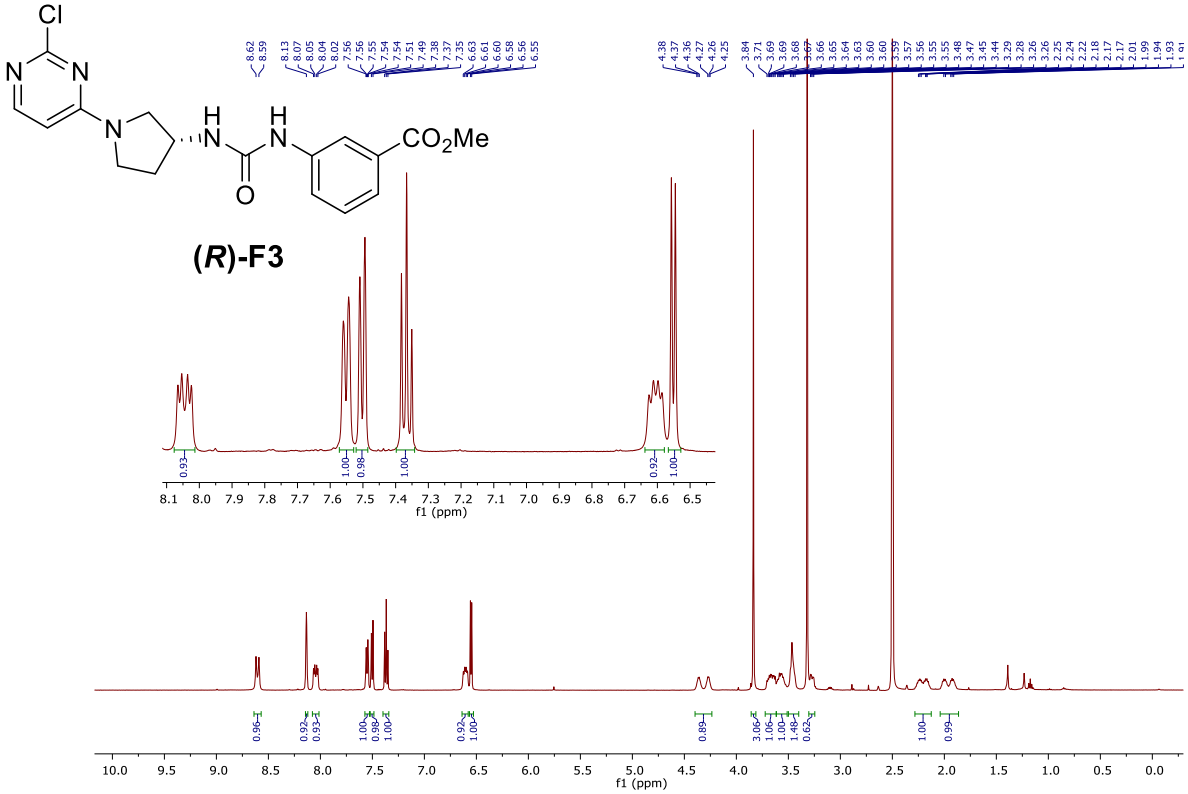
F3, ^{1}H, 500 MHz, DMSO-d_6



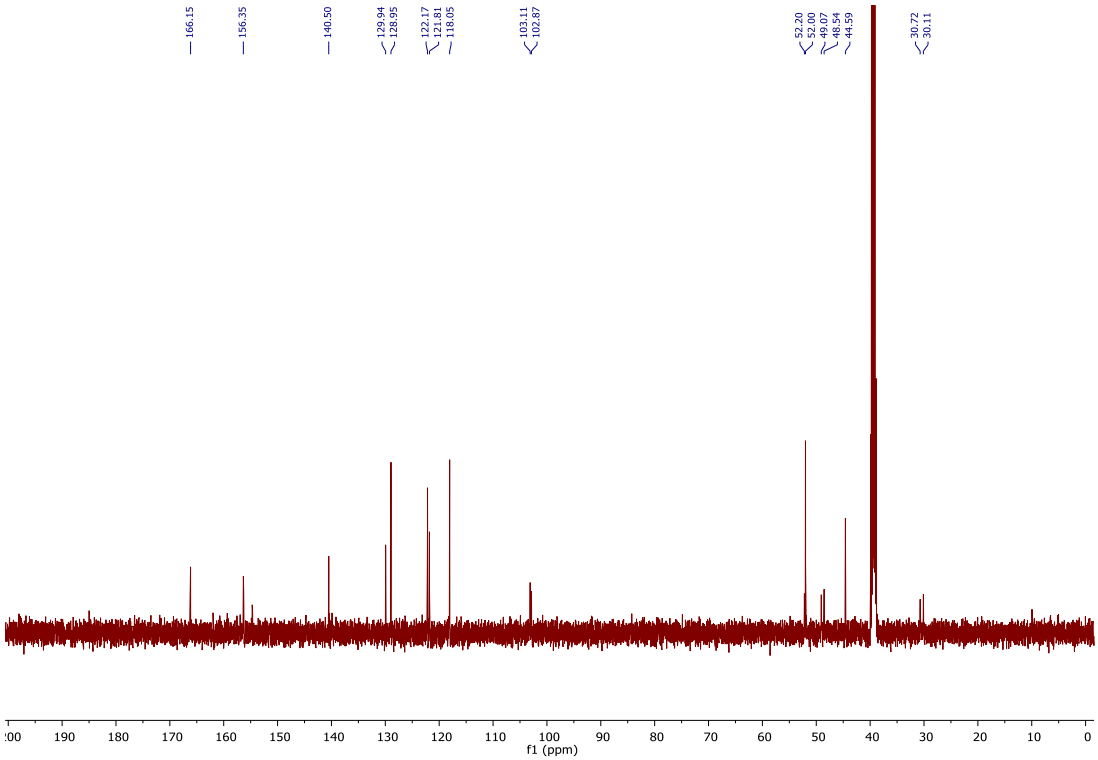
F3, ^{13}C, 126 MHz, DMSO-d_6



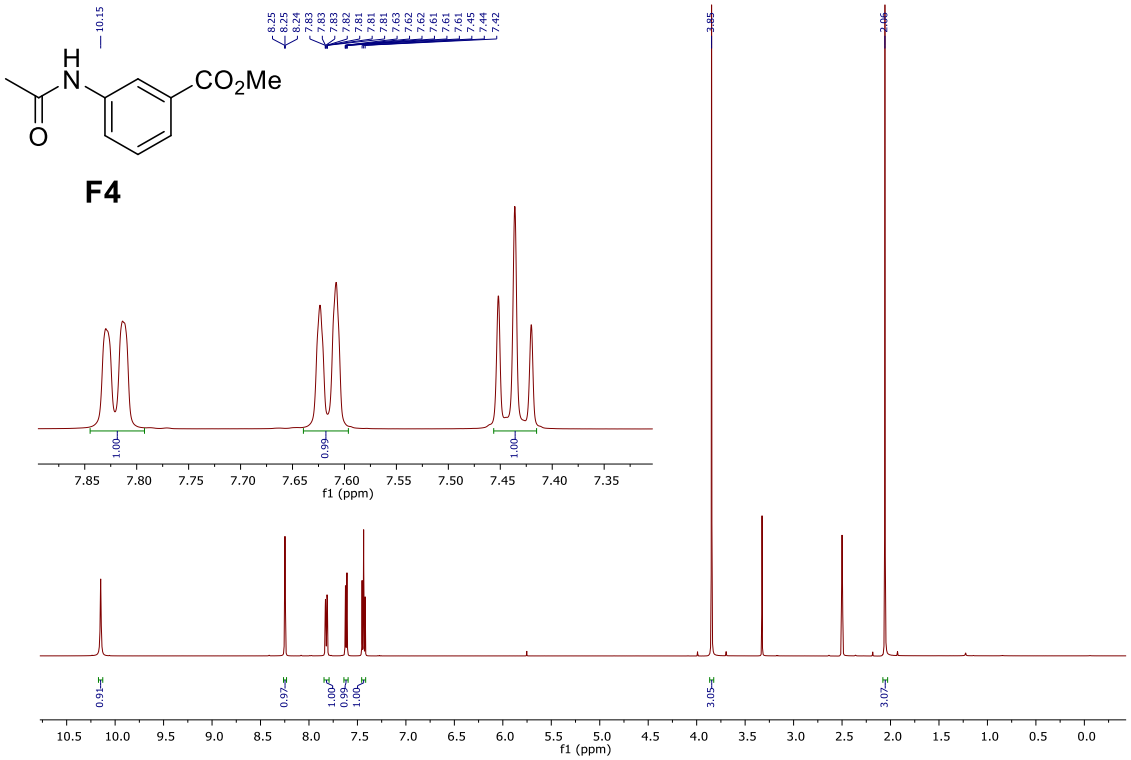
(*R*)-F3, ^{1}H, 500 MHz, DMSO-d_6



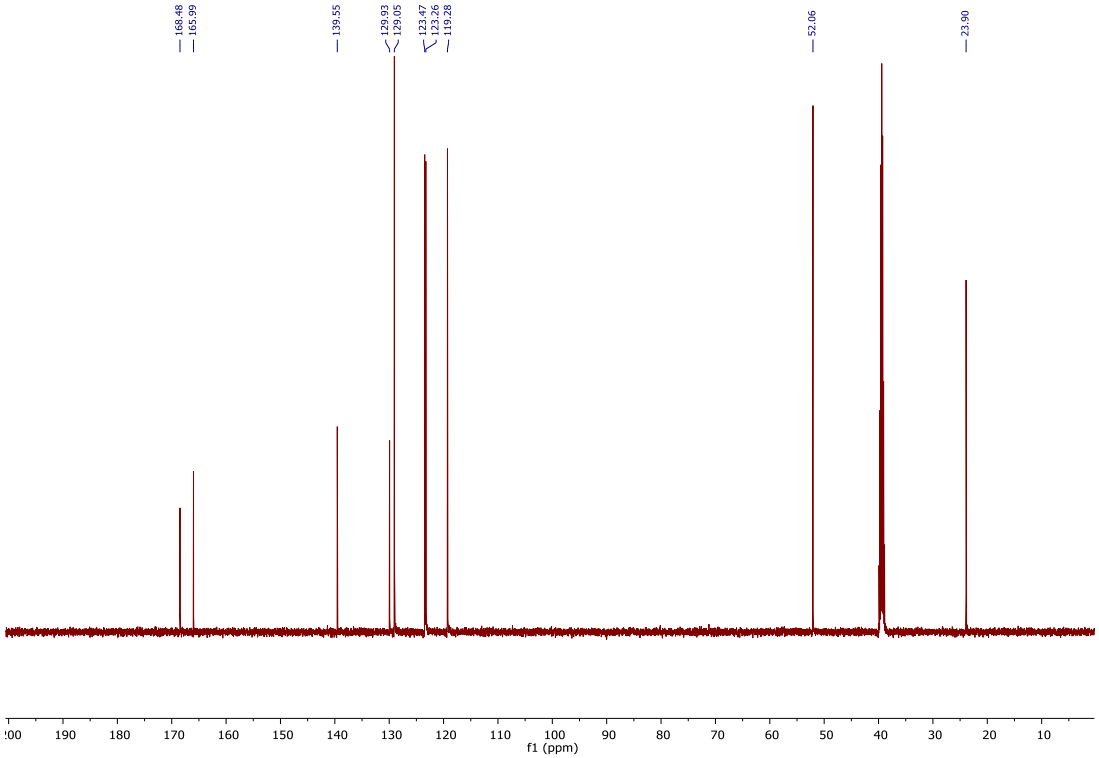
(*R*)-F3, ^{13}C, 126 MHz, DMSO-d_6



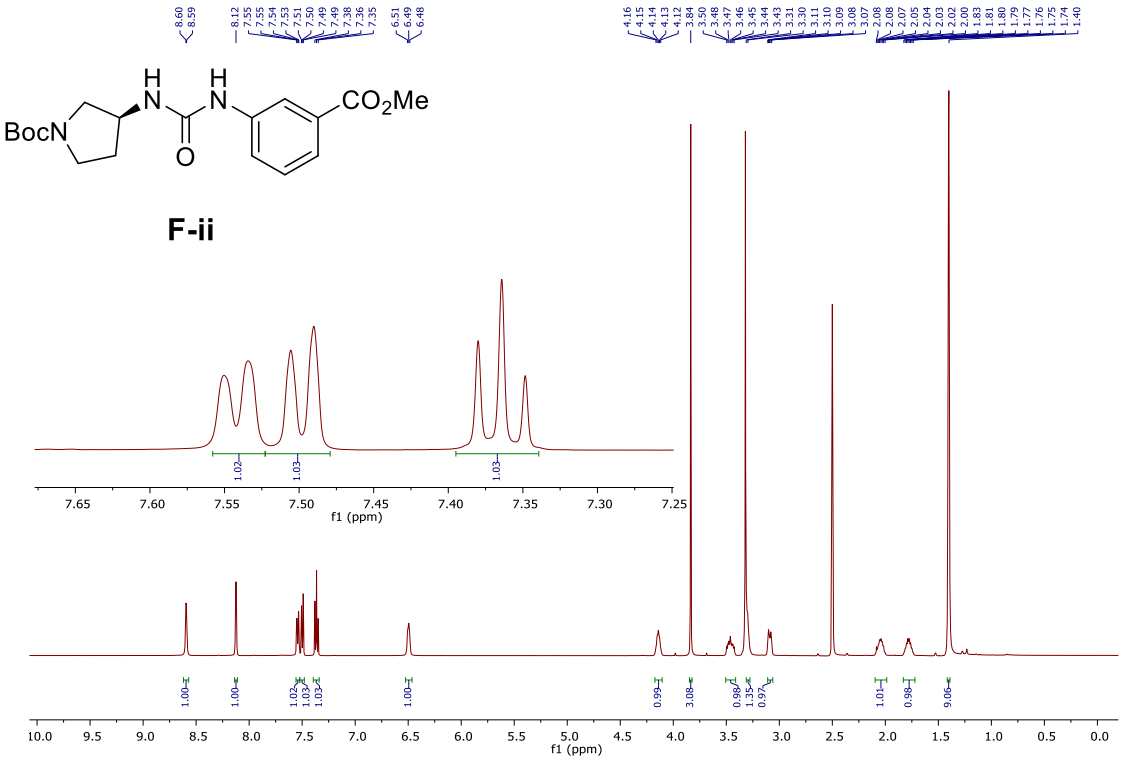
F4, ^{1}H, 500 MHz, DMSO-d_6



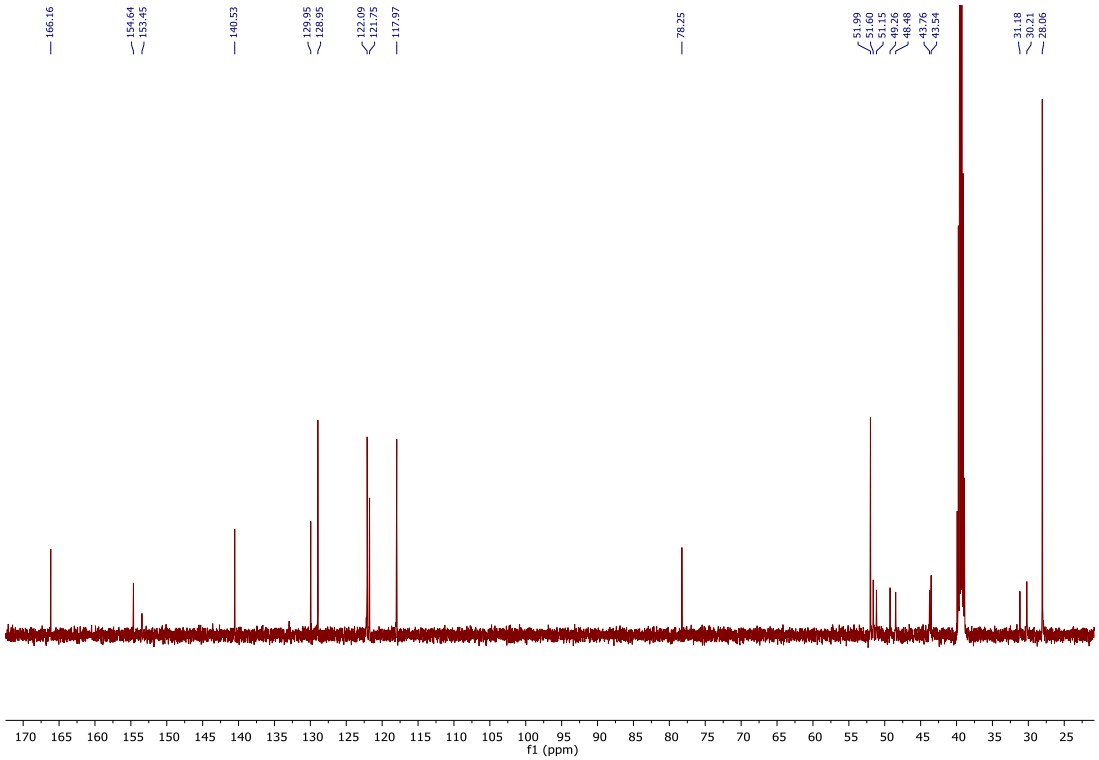
F4, ^{13}C, 126 MHz, DMSO-d_6



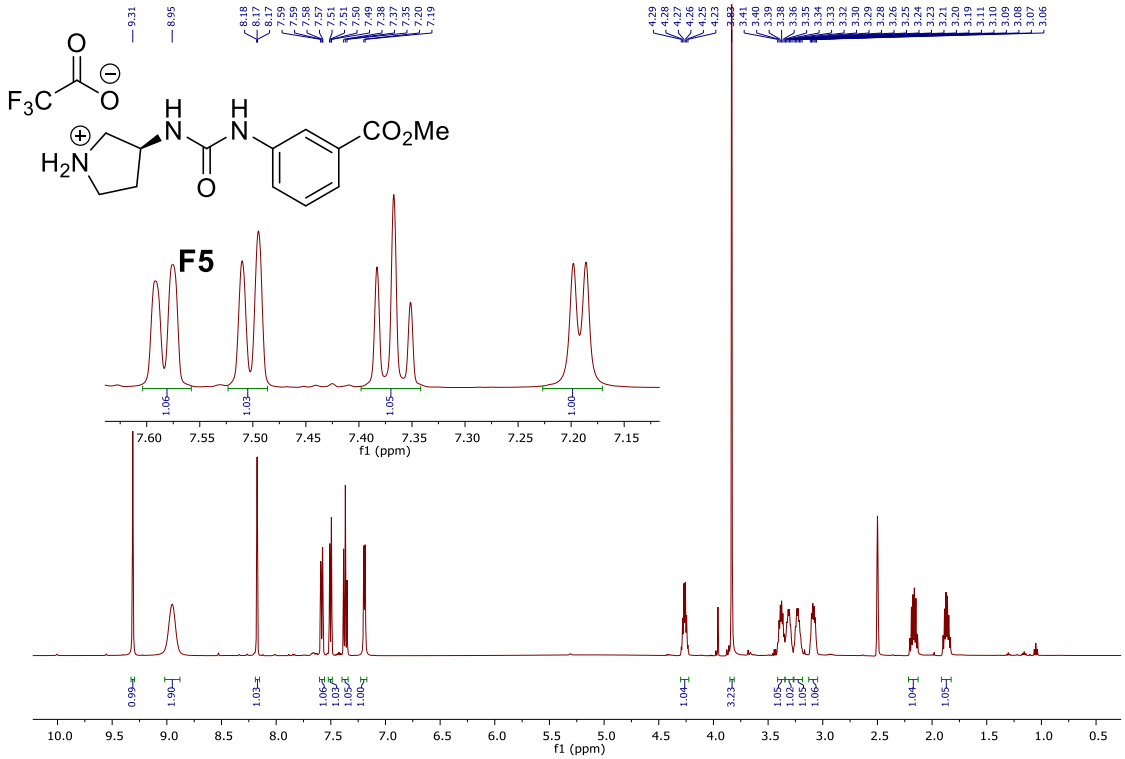
F-ii, ^{1}H, 500 MHz, DMSO-d_6



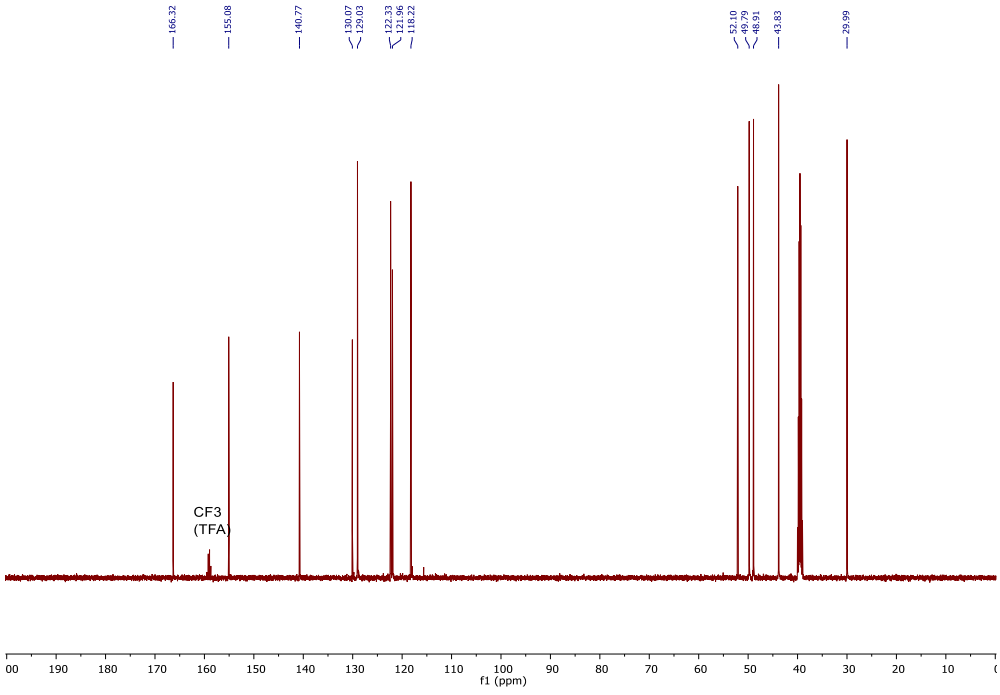
F-ii, ^{13}C, 126 MHz, DMSO-d_6



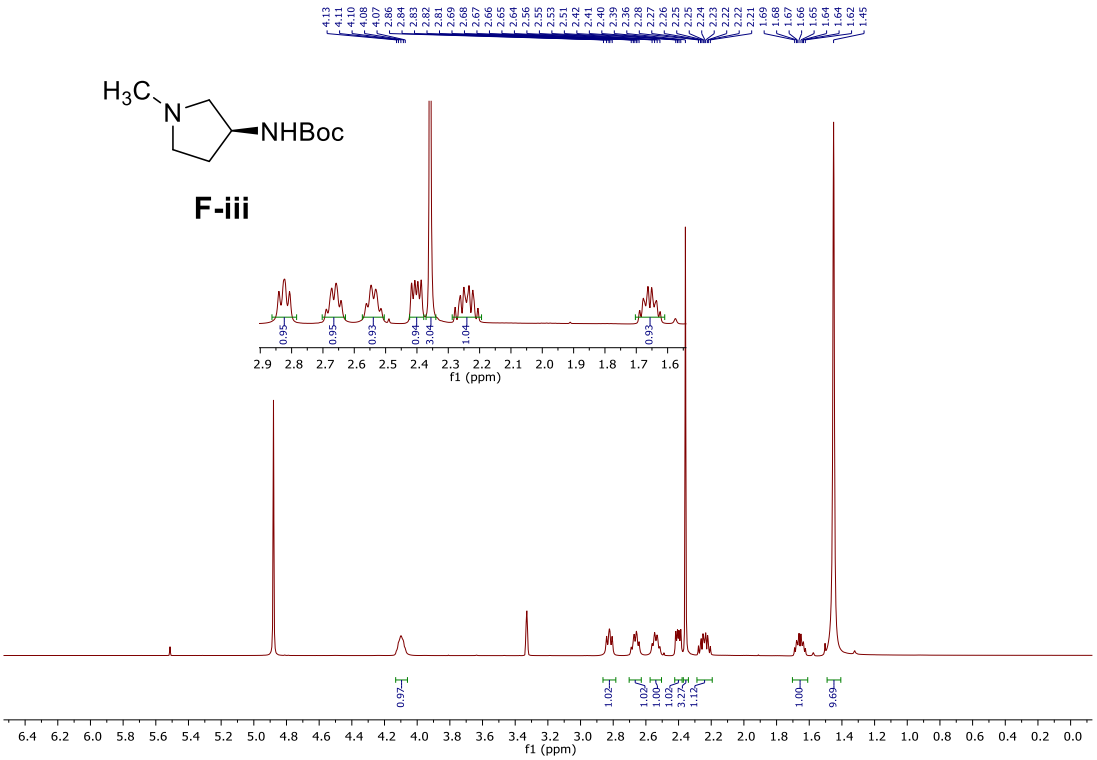
F5, ^{1}H, 500 MHz, DMSO-d_6



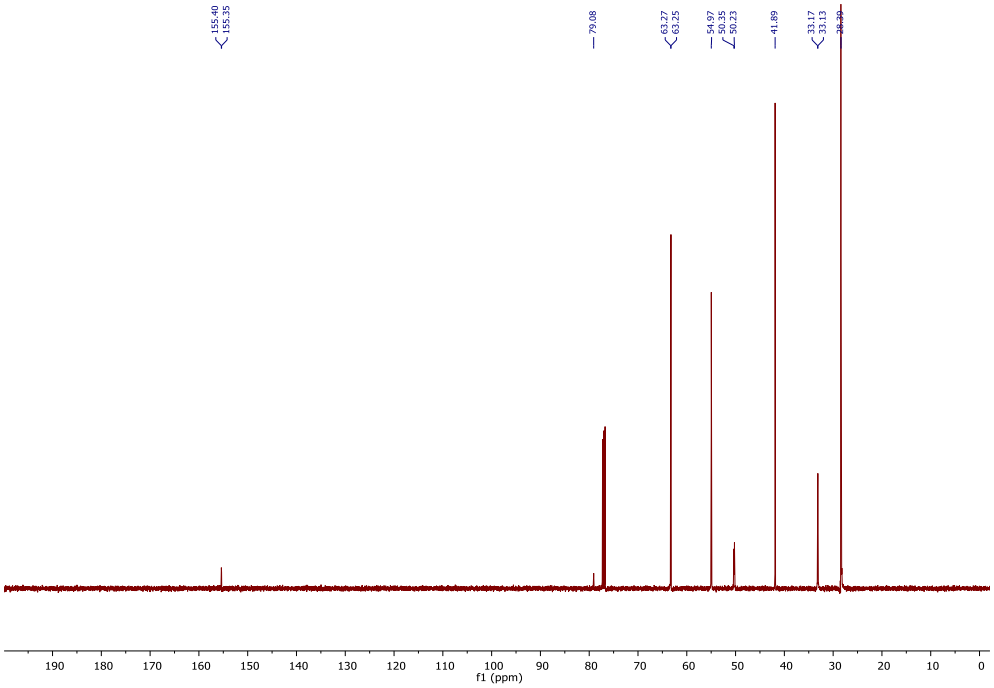
F5, ^{13}C, 126 MHz, DMSO-d_6



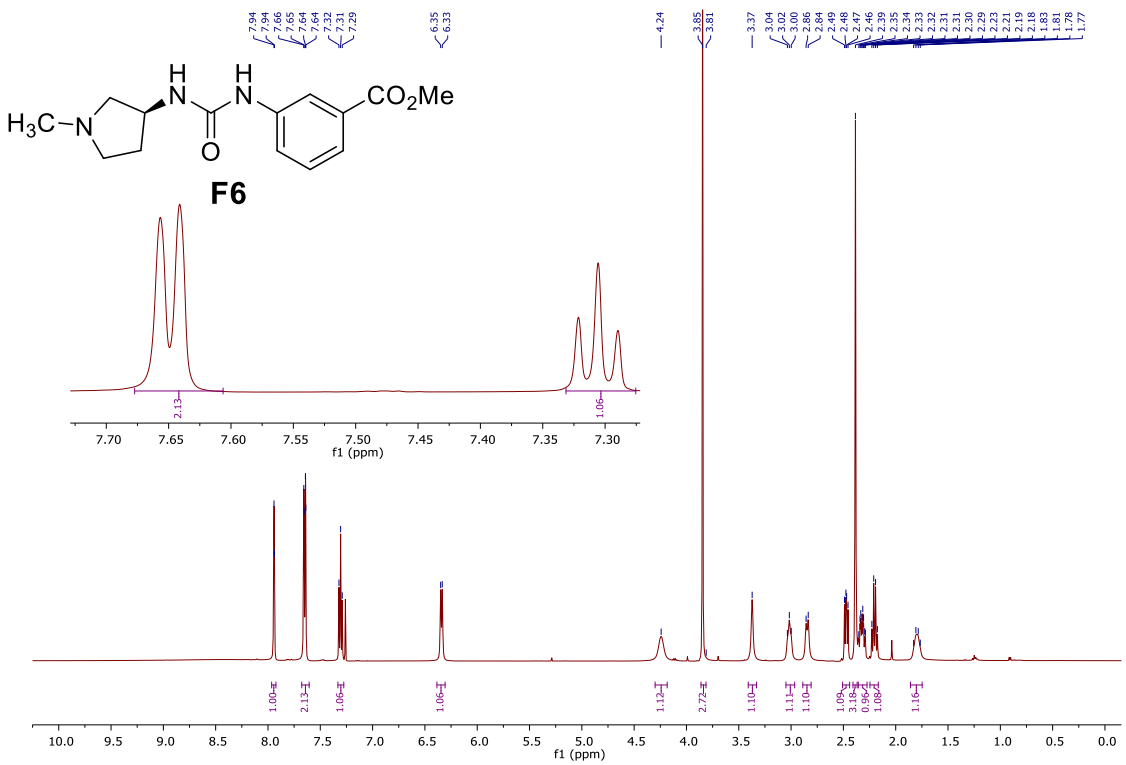
F-iii, ^{1}H, 500 MHz, MeOD



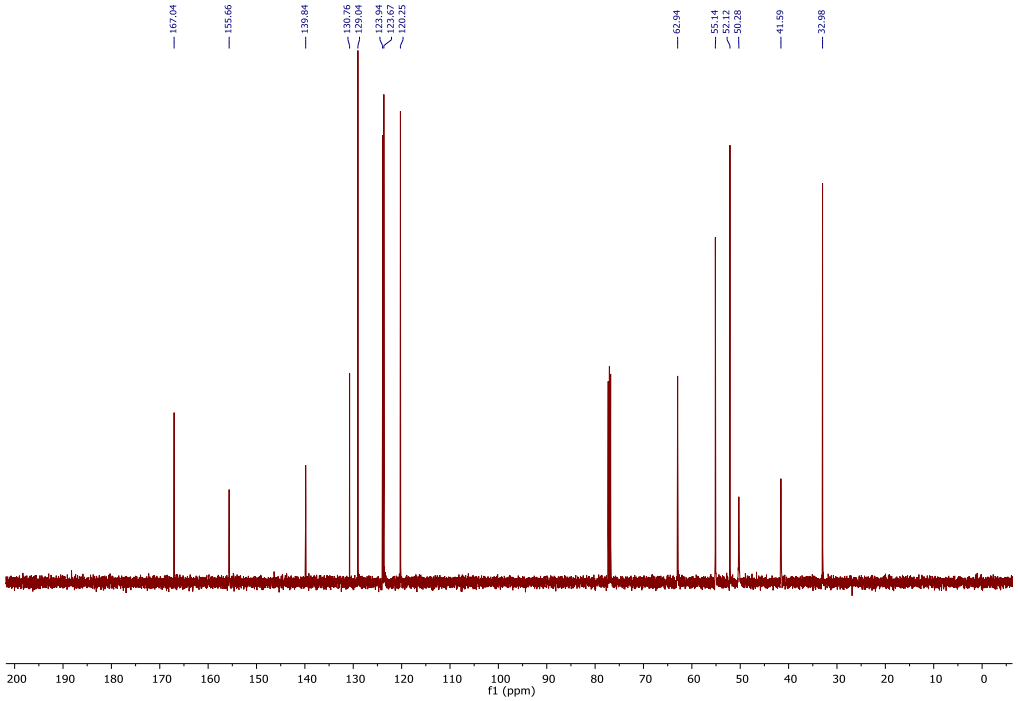
F-iii, ^{13}C, 126 MHz, $CDCl_3$



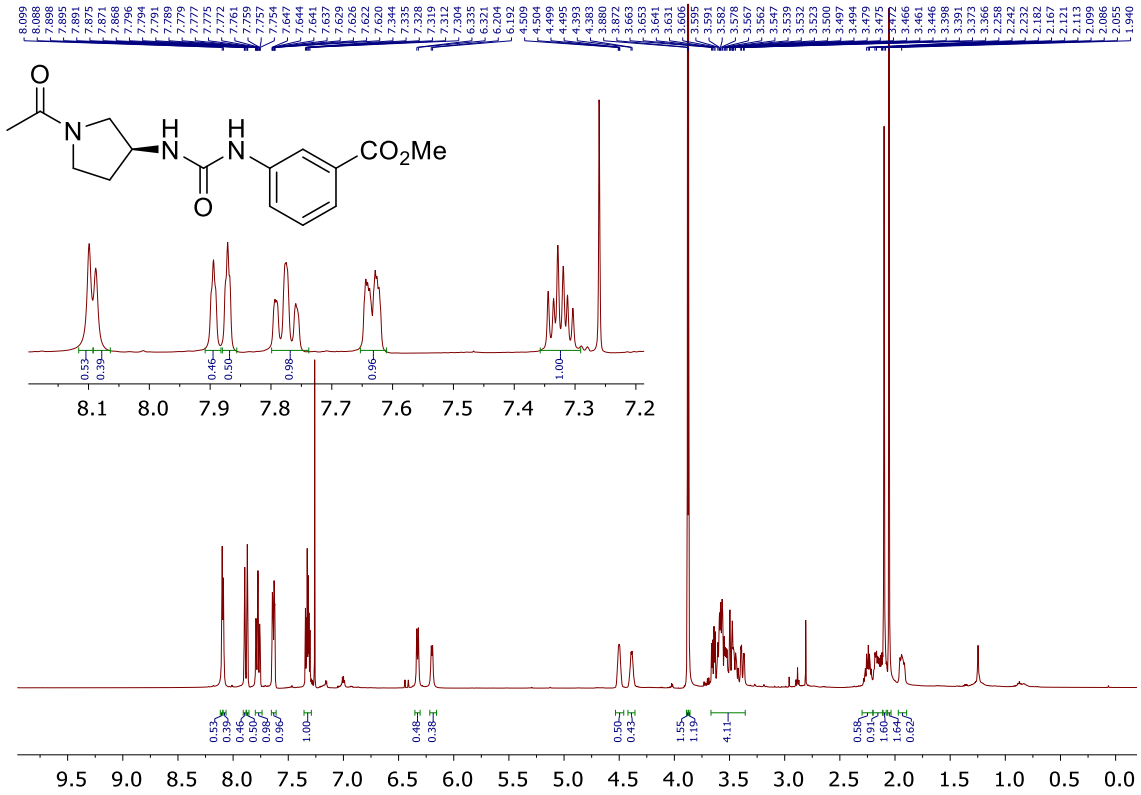
F6, ^{1}H, 500 MHz, $CDCl_3$



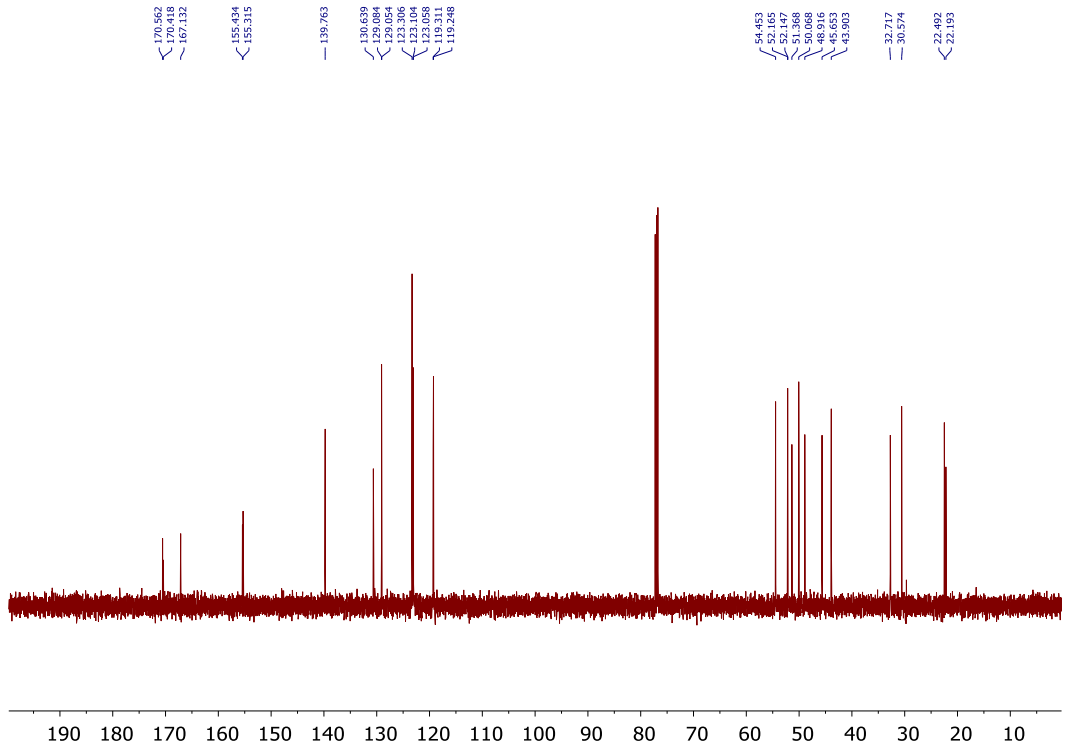
F6, ^{13}C, 126 MHz, $CDCl_3$



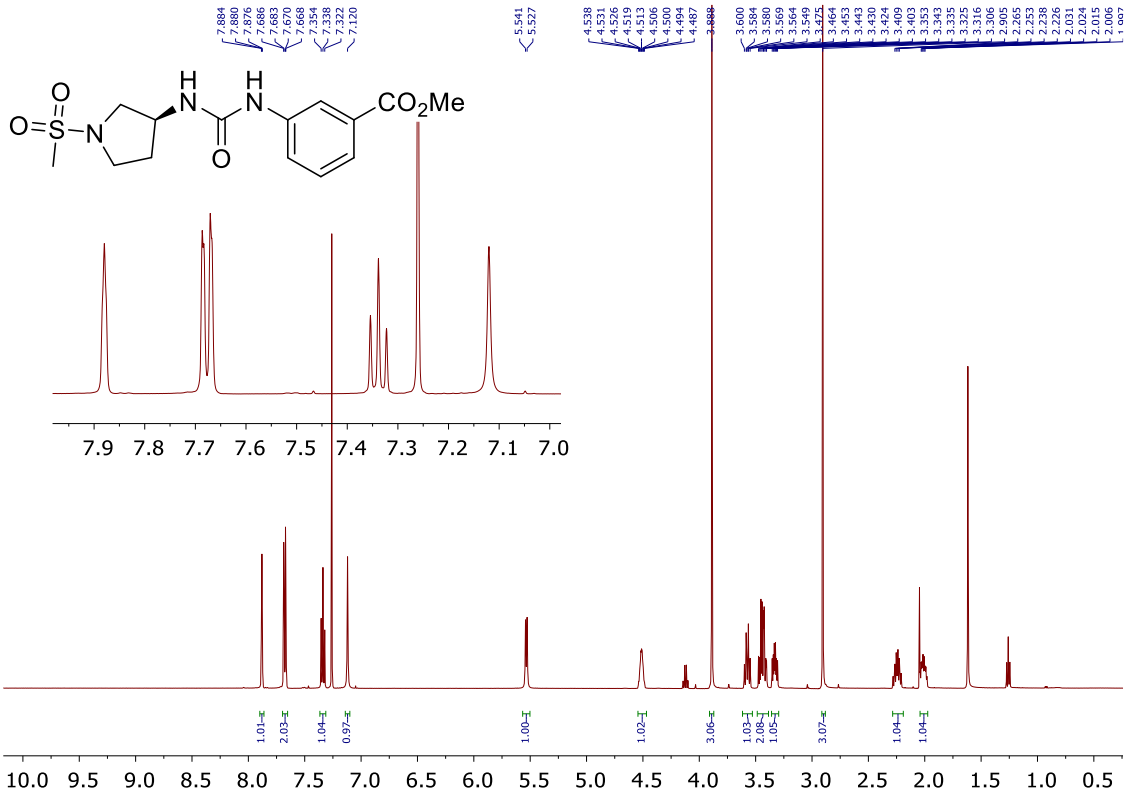
F7, ^{1}H, 500 MHz, $CDCl_3$



F7, ^{13}C, 126 MHz, $CDCl_3$

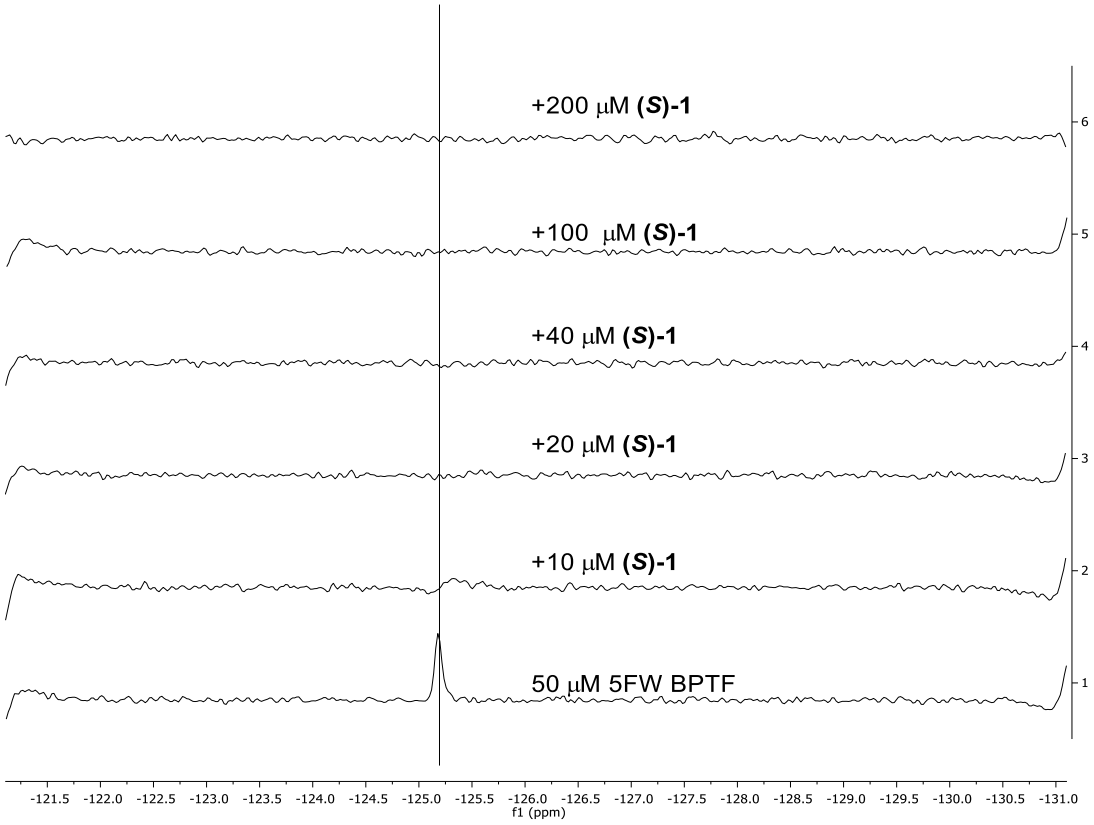


F8, ^{1}H, 500 MHz, $CDCl_3$

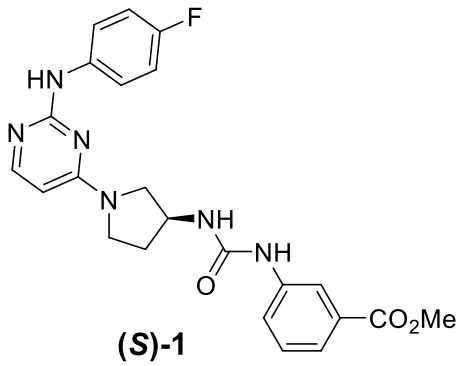


4.12.13 PrOF NMR Spectra

(*S*)-1 PrOF NMR titration

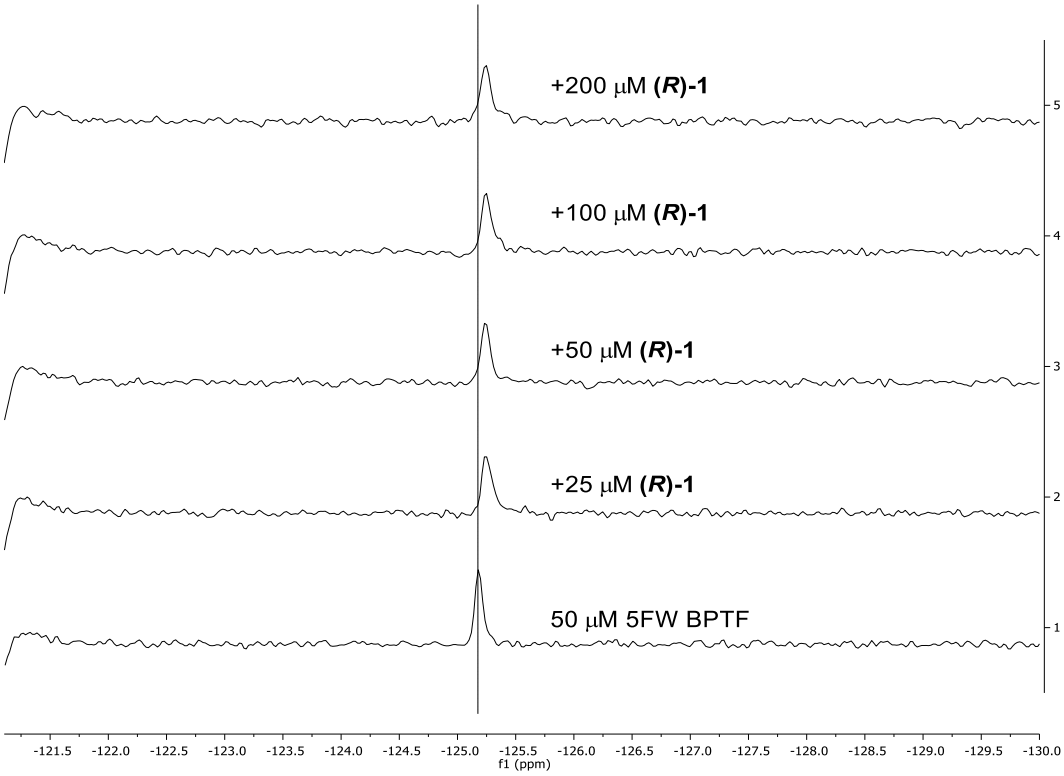


Broadens into baseline, no resonance is recovered.

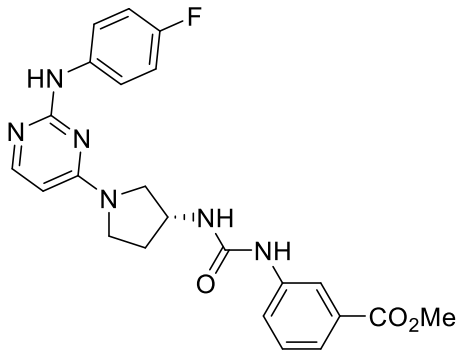


(*S*)-1

(*R*)-1 PrOF NMR titration

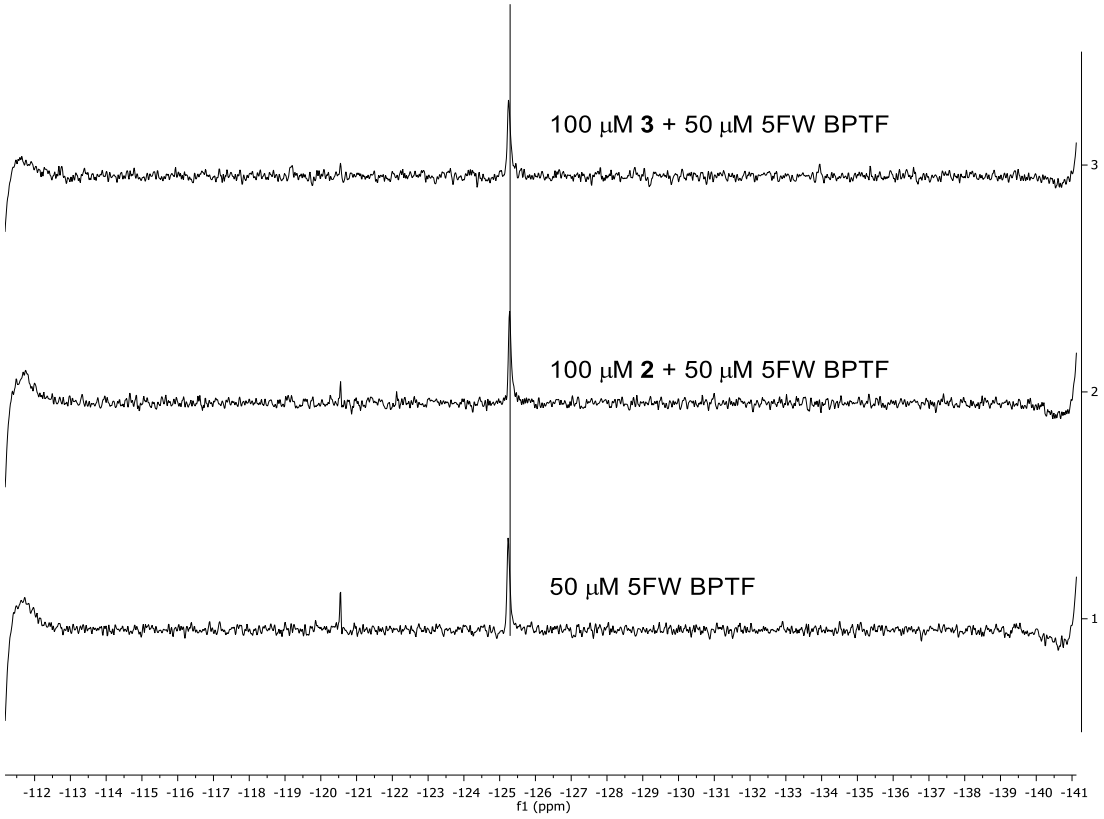


No dose-response observed, slight movement (0.06 ppm change) seen at all concentrations

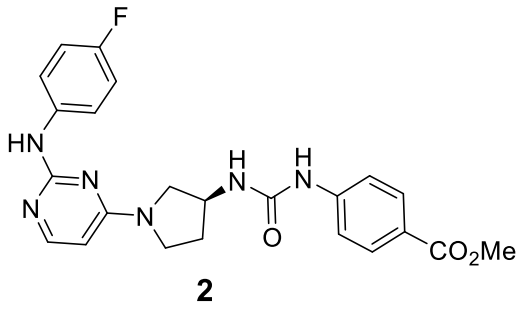


(*R*)-1

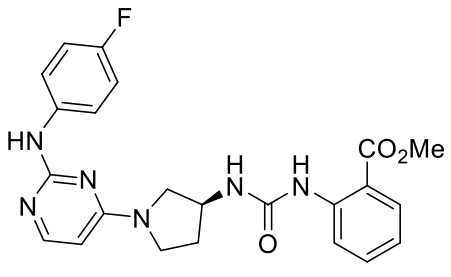
2 and **3**, single point PrOF NMR measurements



No resonance movement or broadening observed, compounds do not bind

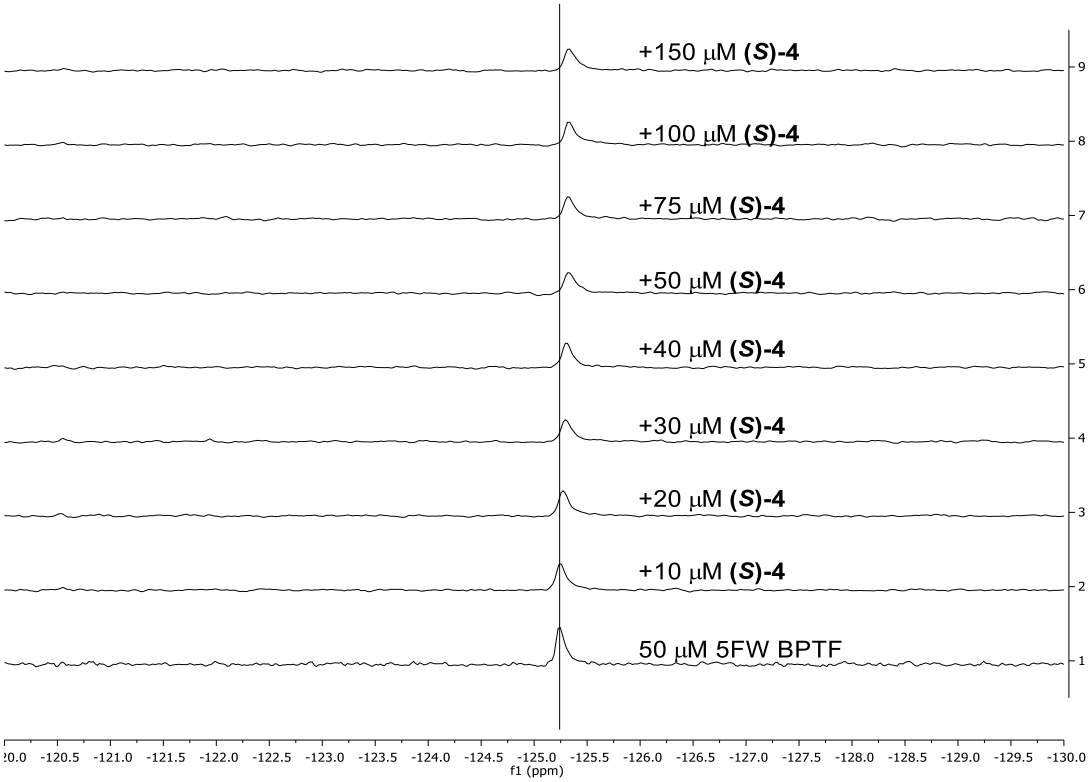


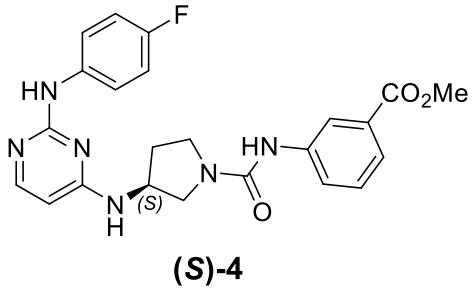
2



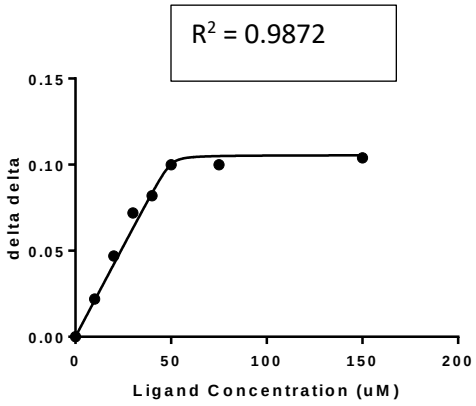
3

(*S*)-4 PrOF NMR titration with 5FW BPTF



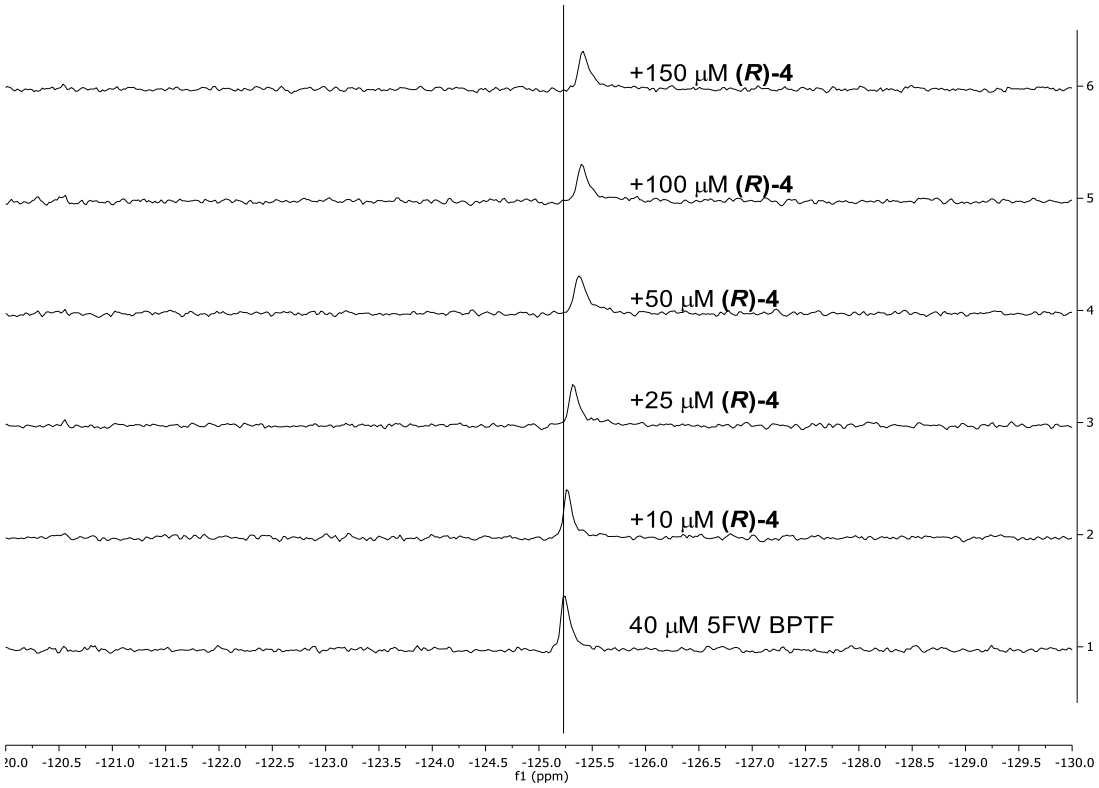


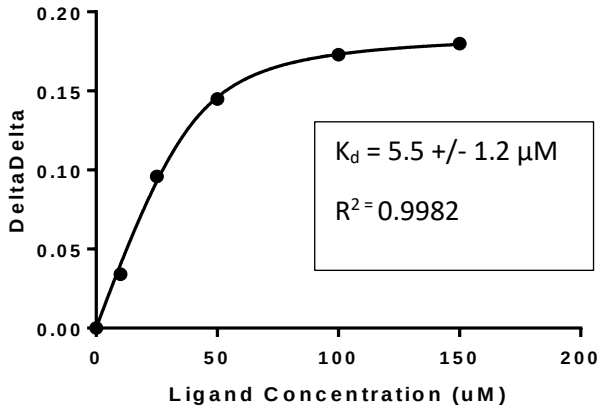
(*S*)-4

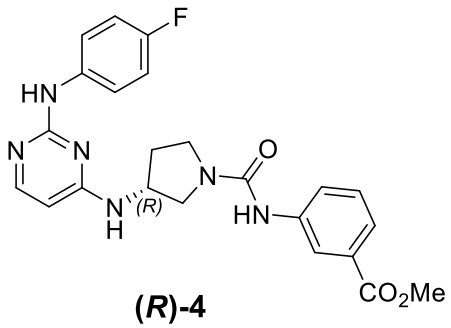


Apparent stoichiometric binding prevented K_d determination, followed up with alpha screen

(*R*)-4 PrOF NMR titration with 5FW BPTF

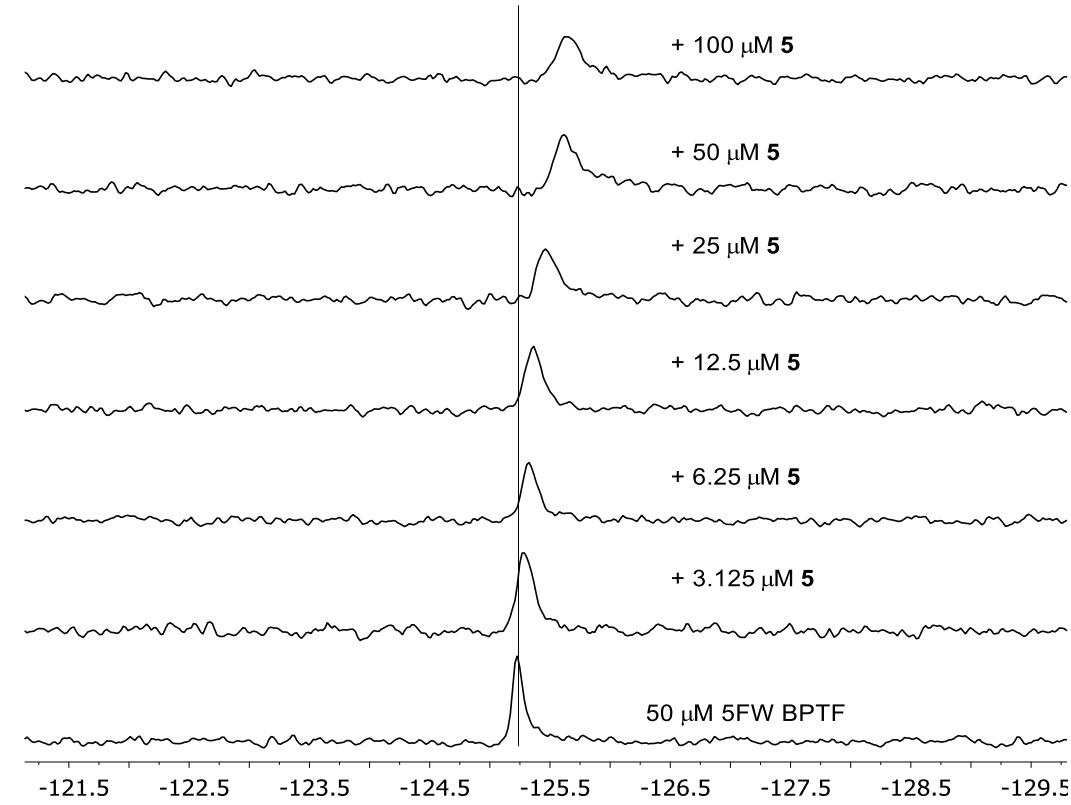


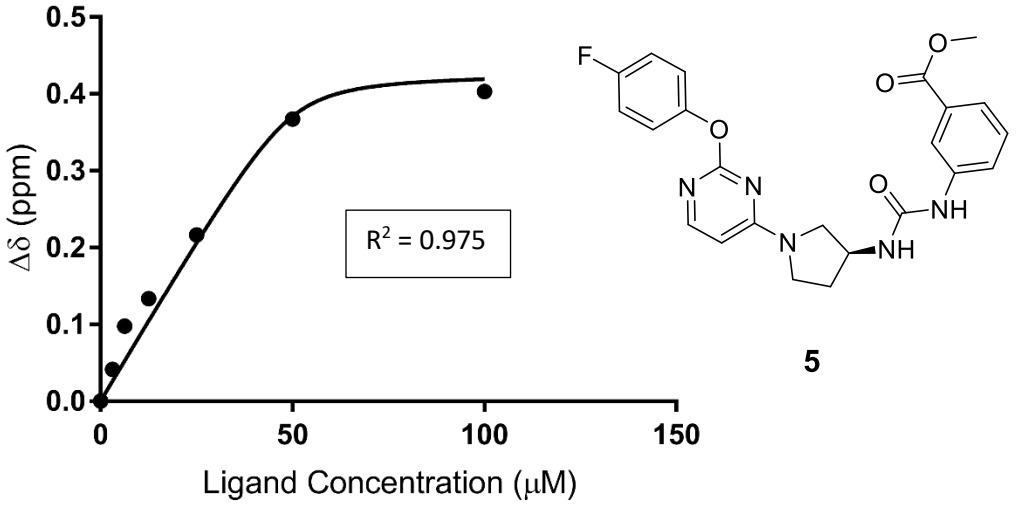




(*R*)-4

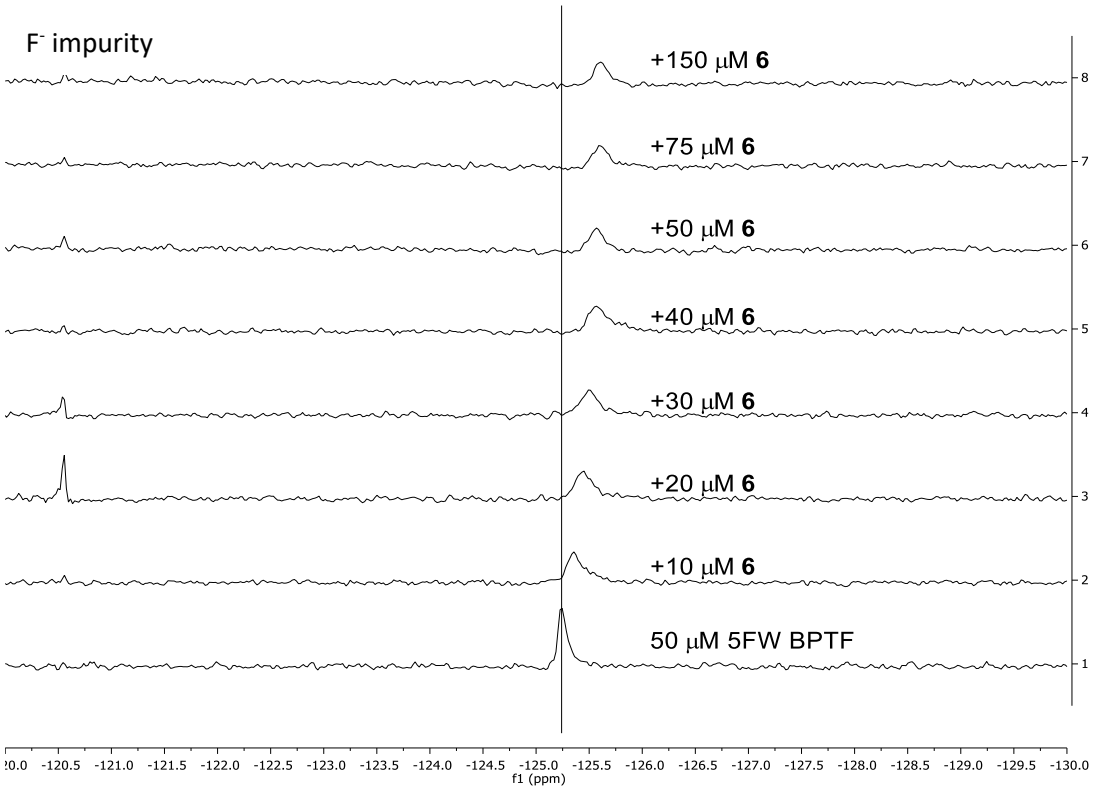
5 PrOF NMR titration with 5FW BPTF

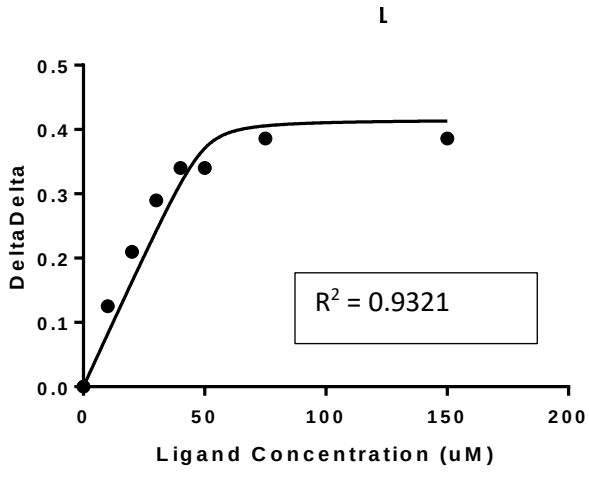


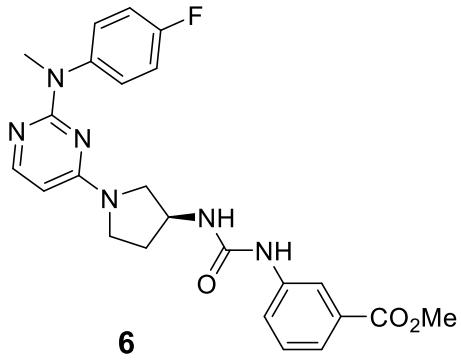


Apparent stoichiometric binding prevents accurate K_d determination

6 PrOF NMR titration with 5FW BPTF





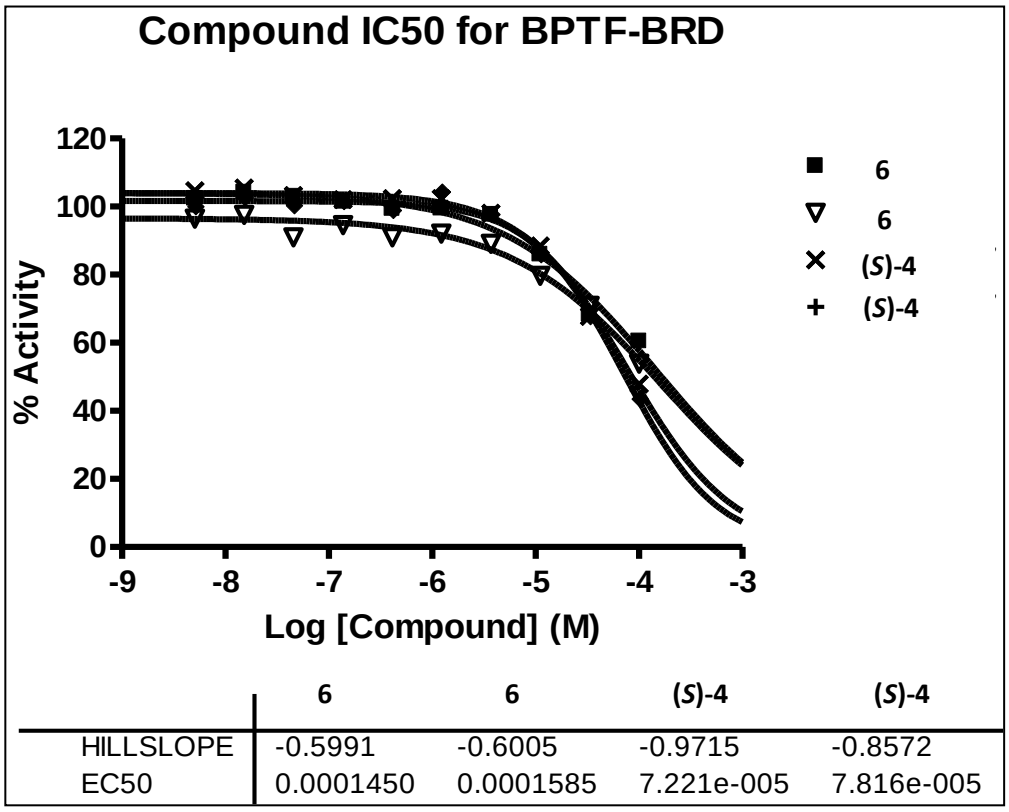


6

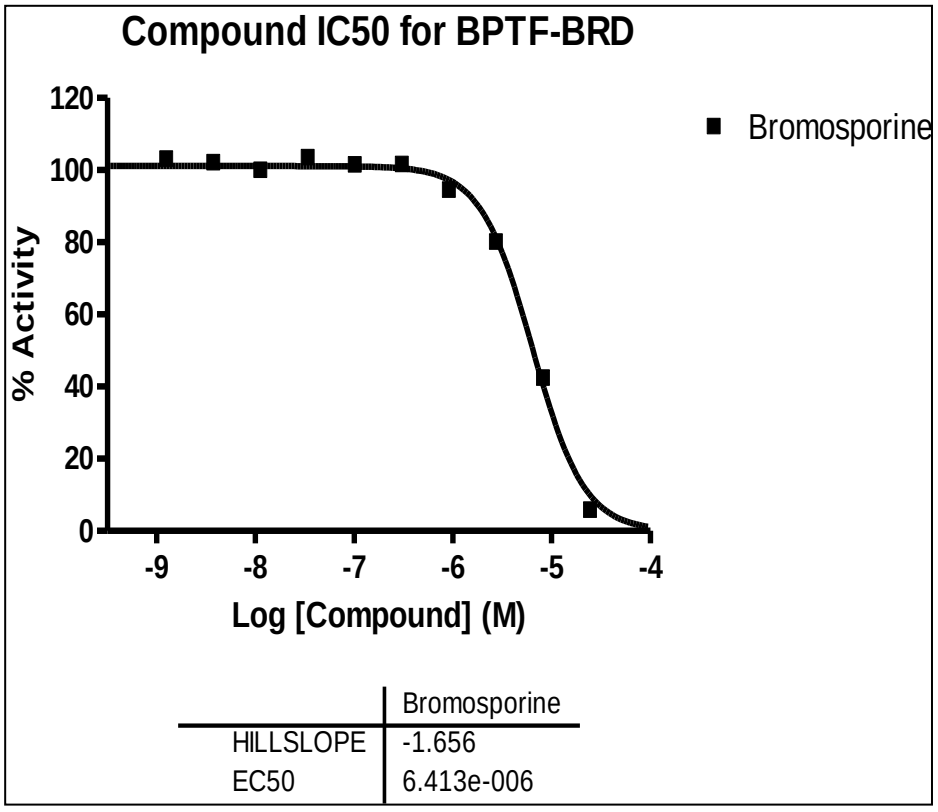
Apparent stoichiometric binding prevented determination of a $K_{d,}$

Followed up with alphascreen assay (next page)

Alpha Screen data for **(*S*)-4**, **6,** and known pan-bromodomain inhibitor, bromosporine

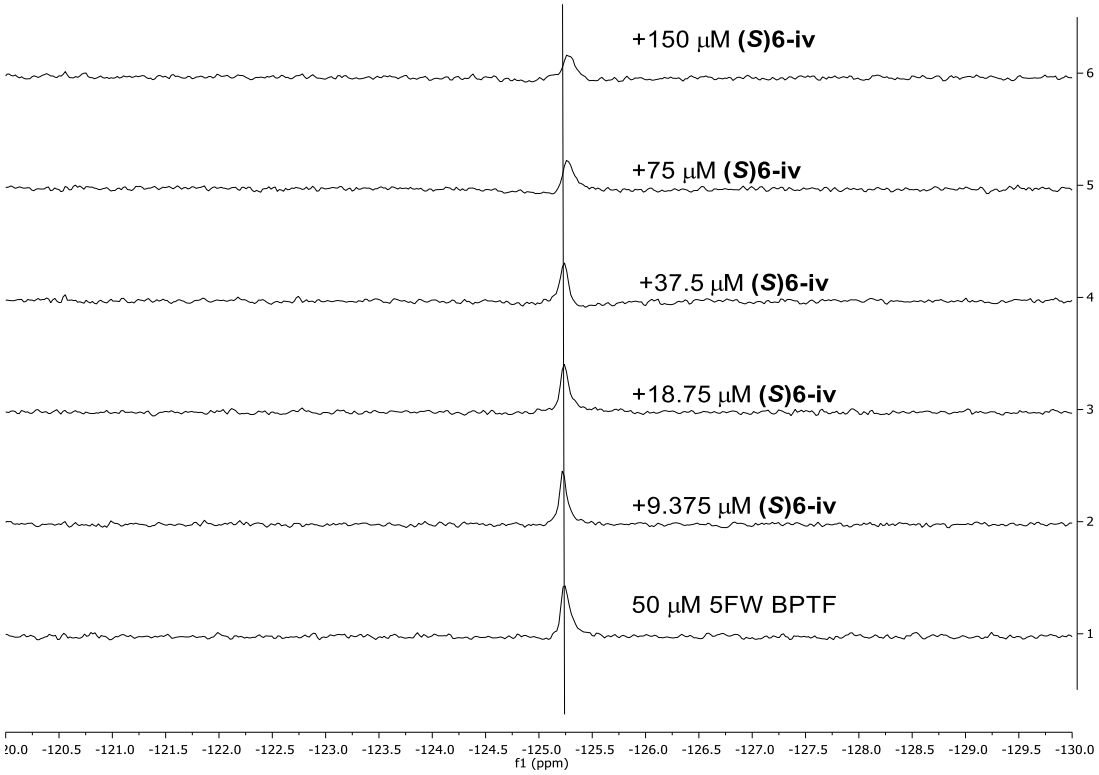


	6	6	(*S*)-4	(*S*)-4
HILLSLOPE	-0.5991	-0.6005	-0.9715	-0.8572
EC50	0.0001450	0.0001585	7.221e-005	7.816e-005

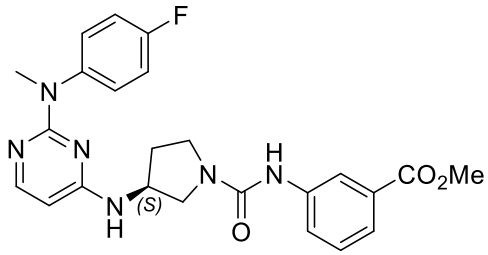


	Bromosporine
HILLSLOPE	-1.656
EC50	6.413e-006

(*S*)6-iv PrOF NMR titration with 5FW BPTF

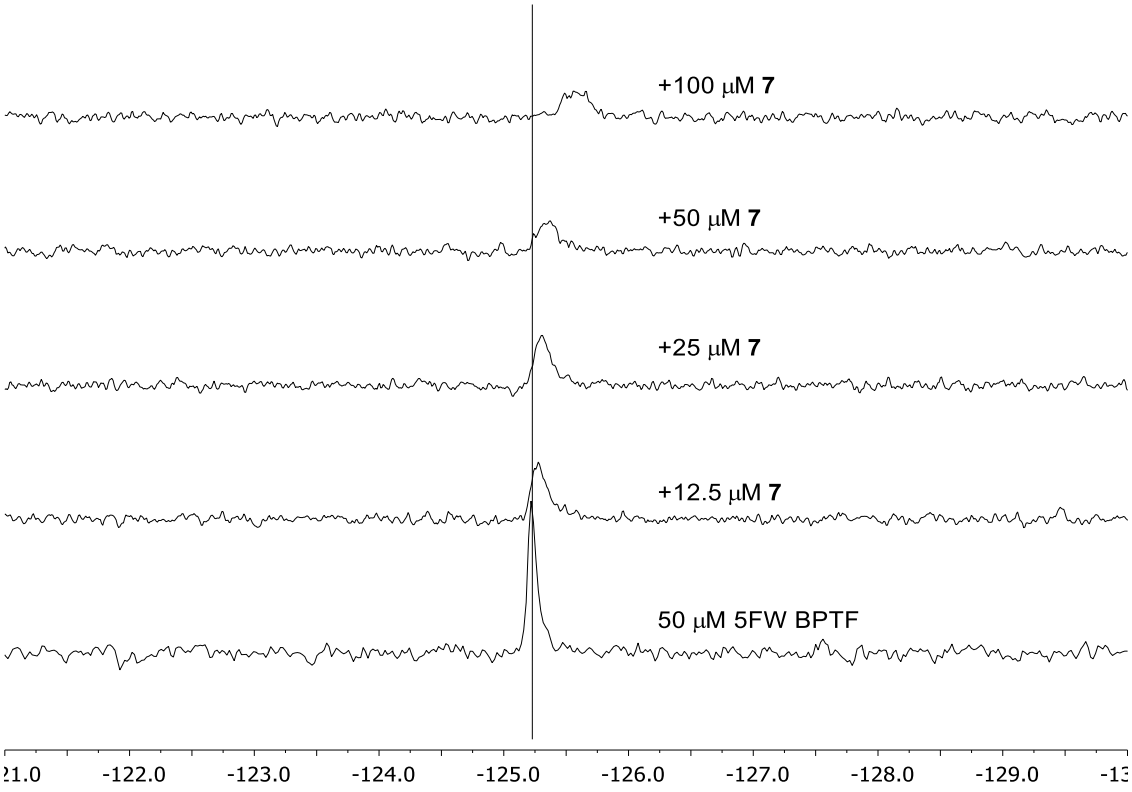


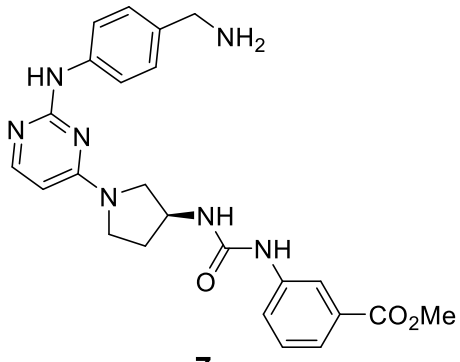
Slight broadening, can not fit to curve



(*S*)6-iv

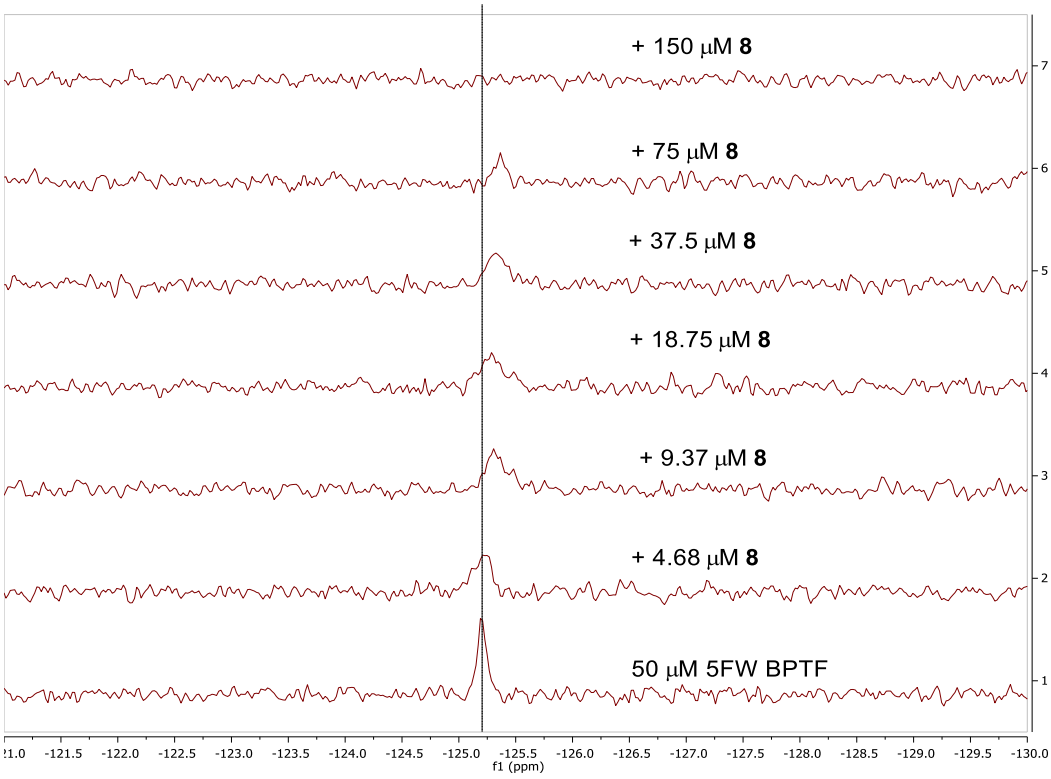
7 PrOF NMR titration with 5FW BPTF



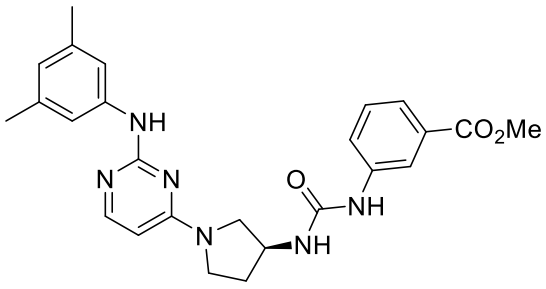


7

8 PrOF NMR titration with 5FW BPTF

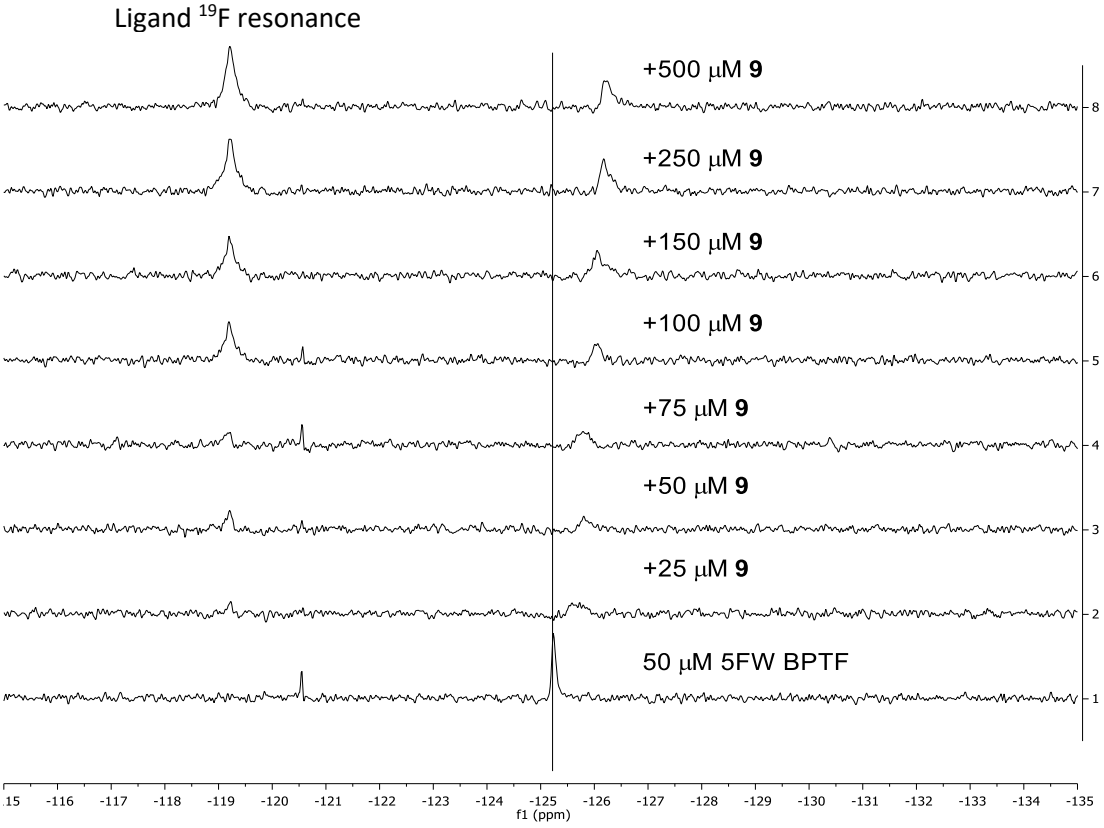


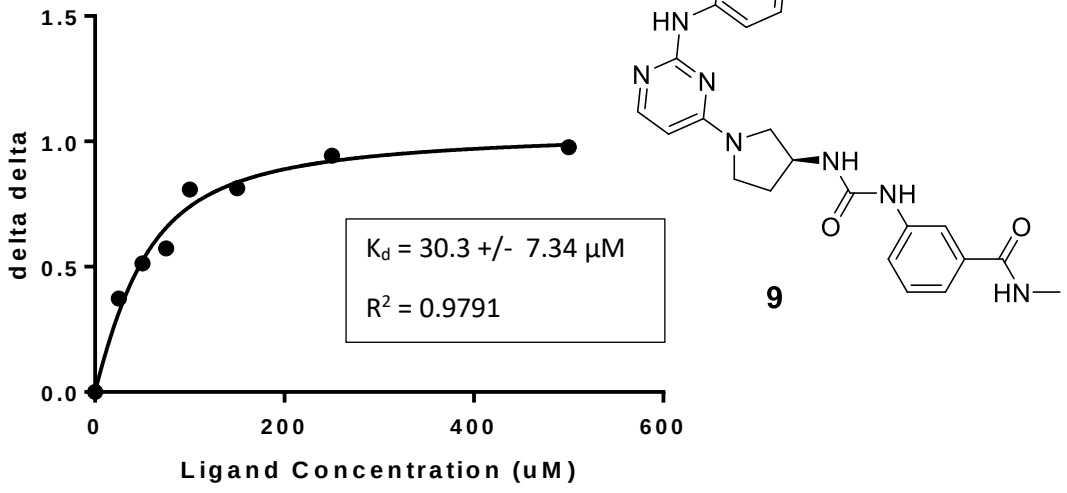
Broadening and eventual disappearance into baseline, can not fit curve



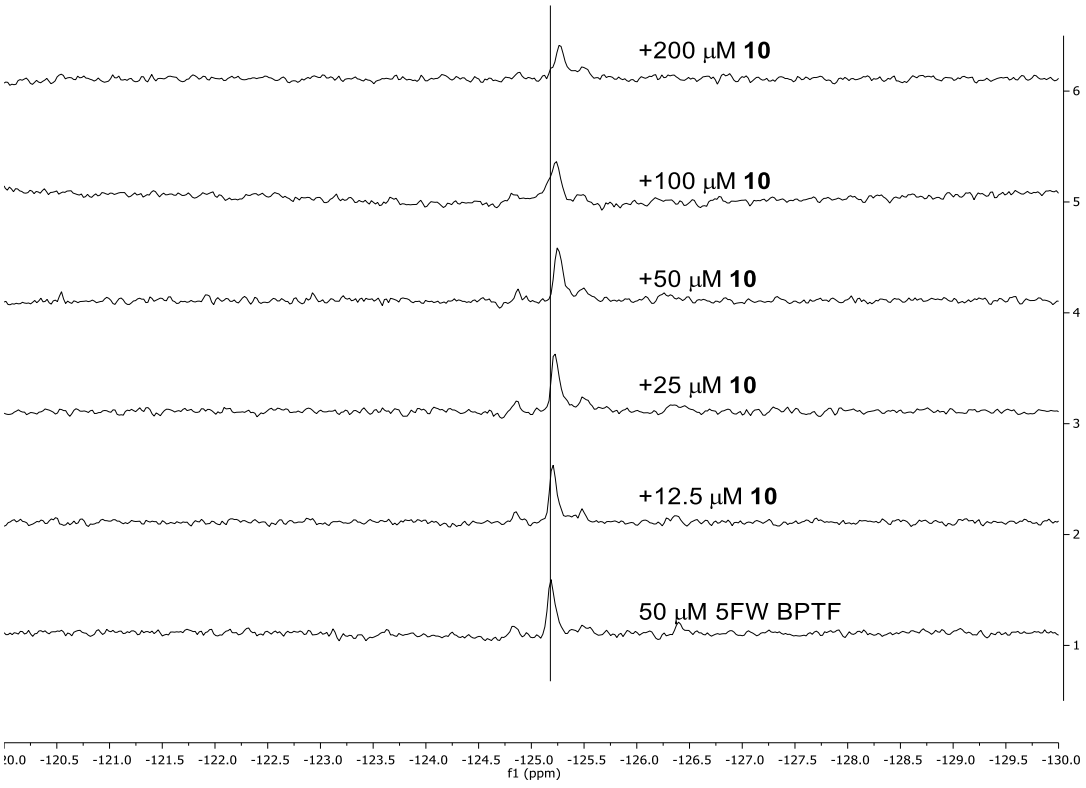
8

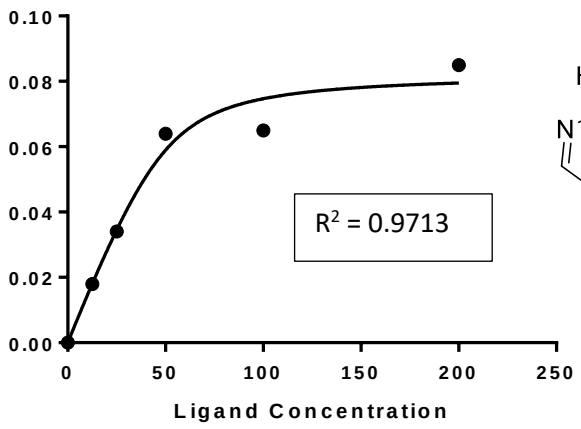
9 PrOF NMR titration

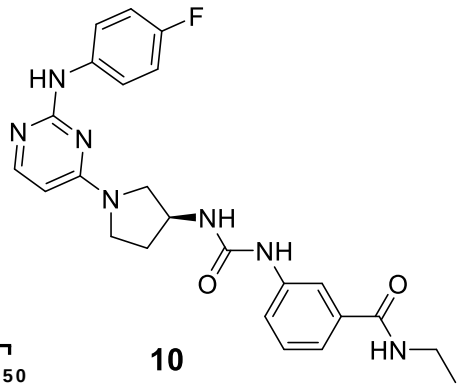




10 PrOF NMR titration with 5FW BPTF



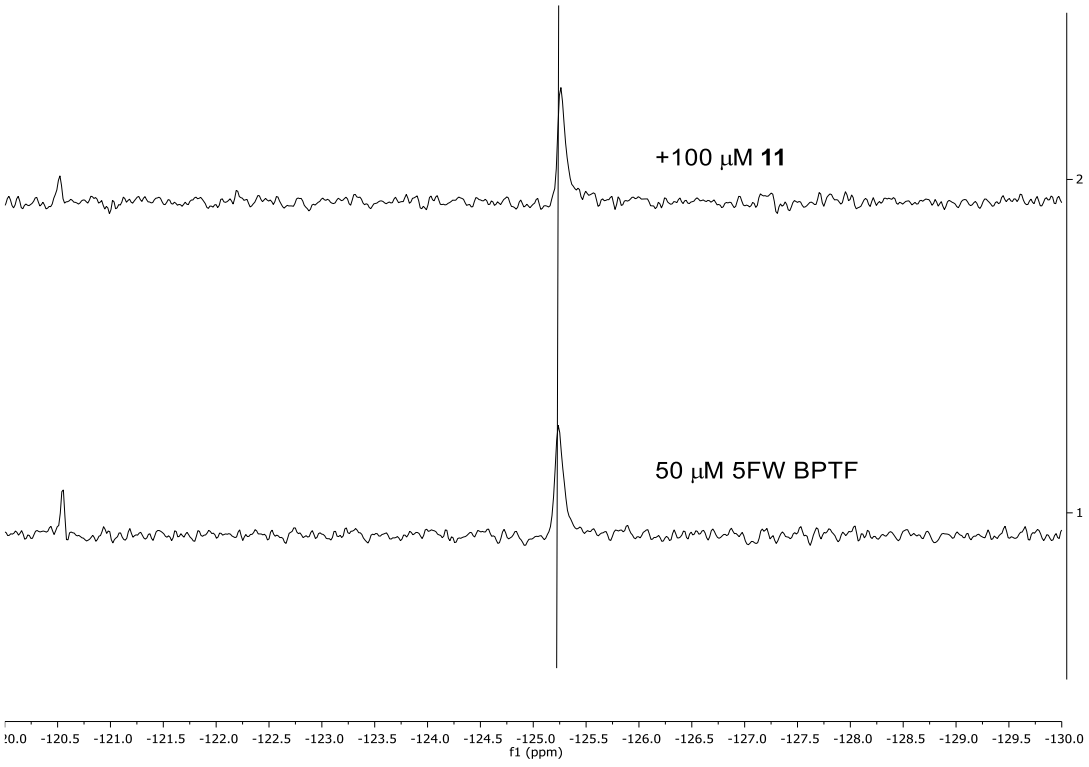




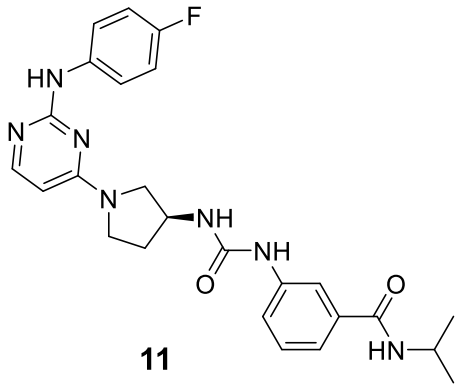
10

Small dynamic range prevents reliable K_d determination

11 PrOF NMR with 5FW BPTF, single point measurement

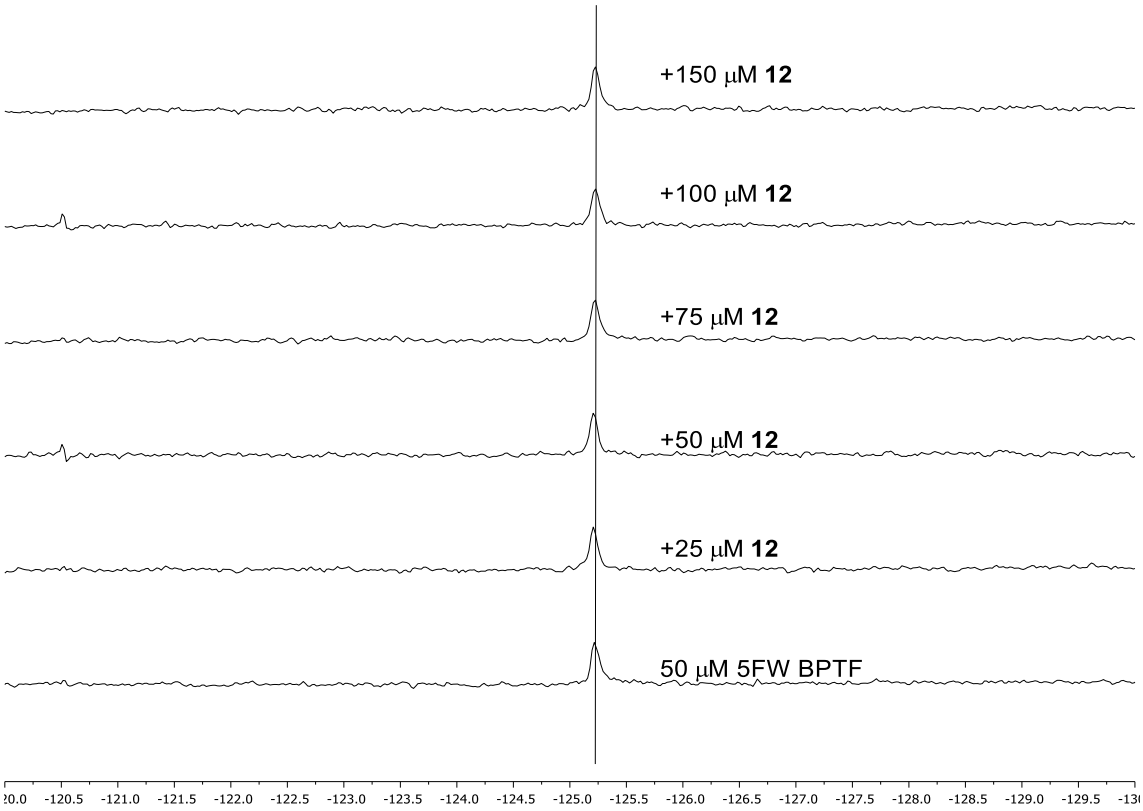


No binding observed

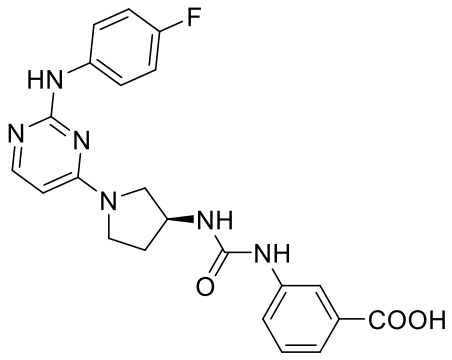


11

12 PrOF NMR titration with 5FW BPTF

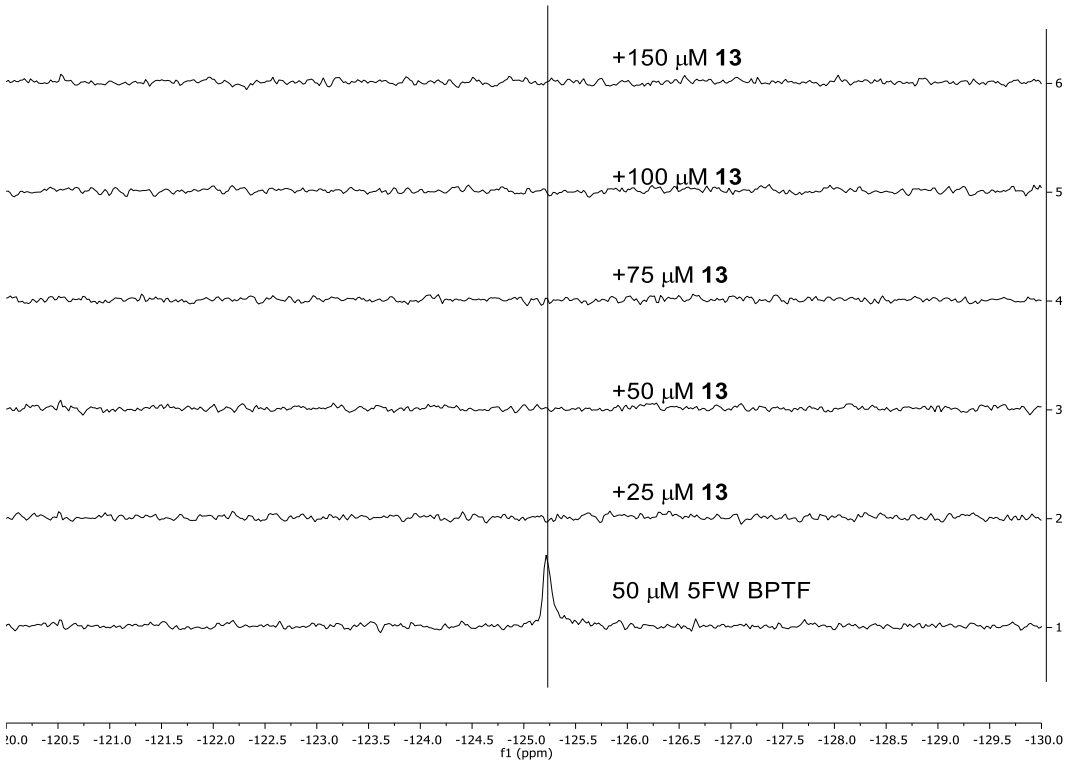


No binding observed

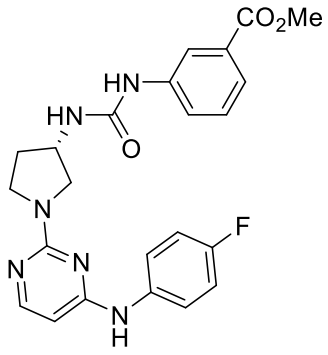


12

13 PrOF NMR titration with 5FW BPTF

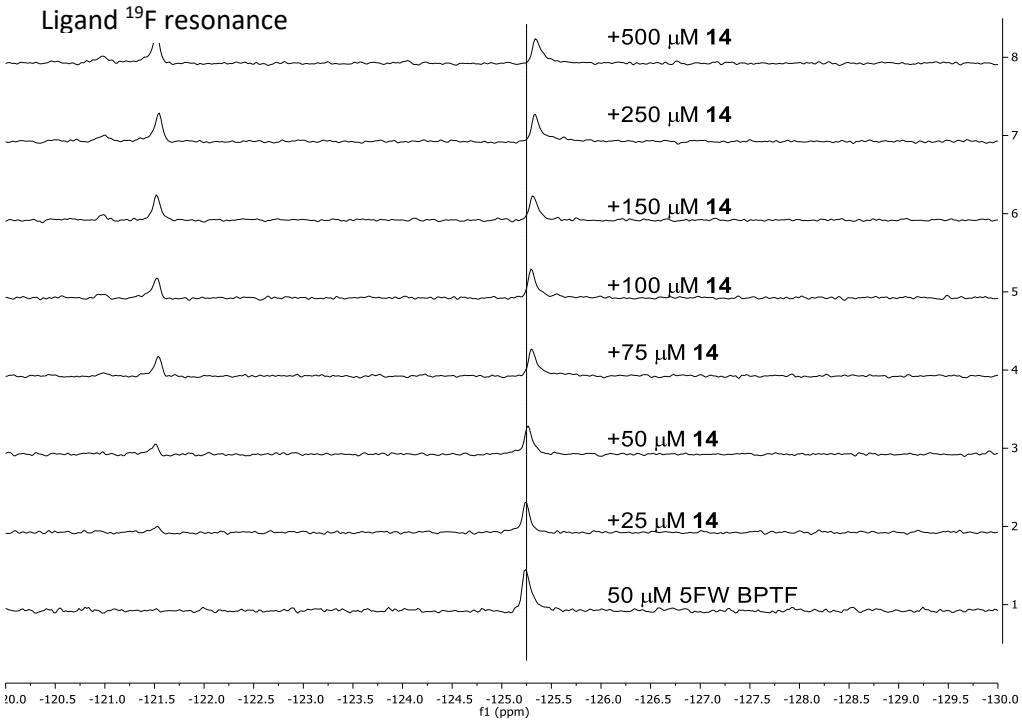


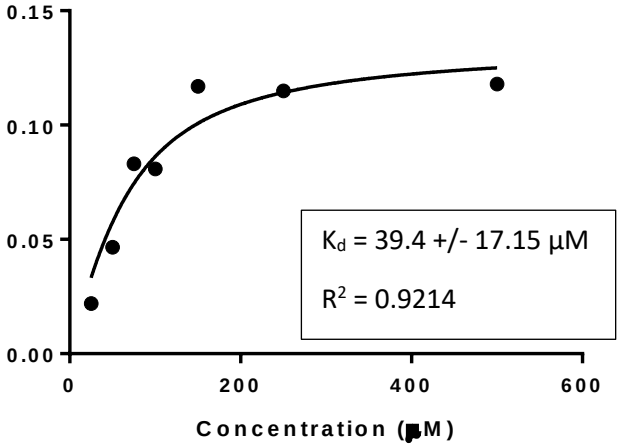
Broadens into baseline and does not return, no quantitative data possible

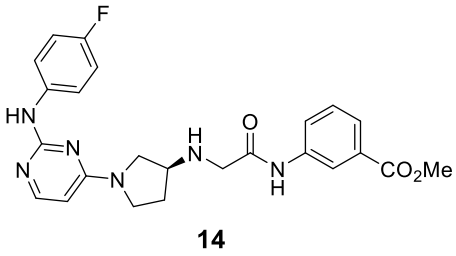


13

14 PrOF NMR titration with 5FW BPTF

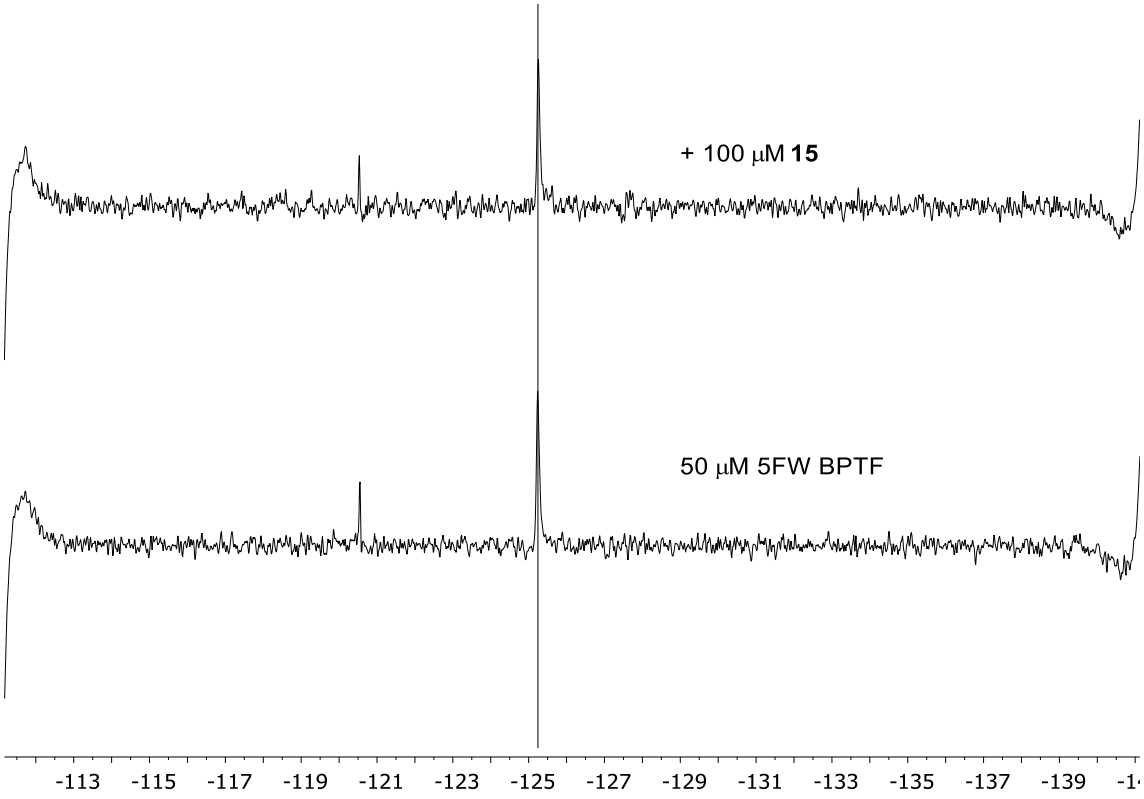




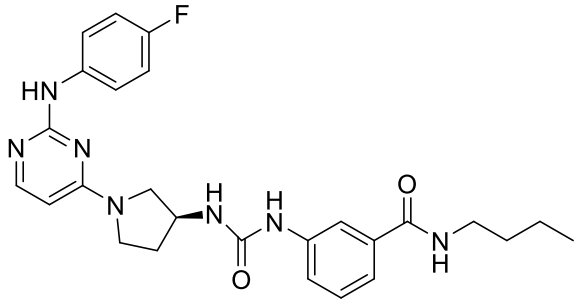


14

15 PrOF NMR single-point measurement with 5FW BPTF

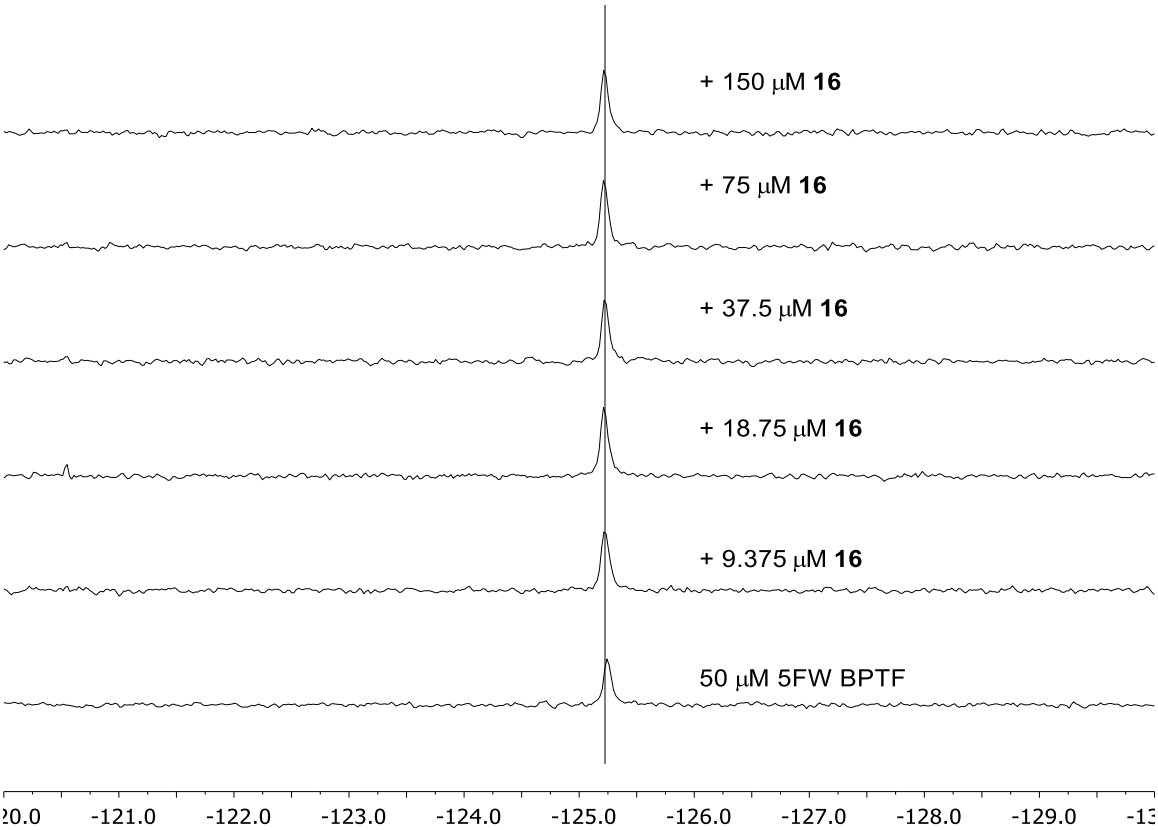


No discernible shift in resonance, no binding observed

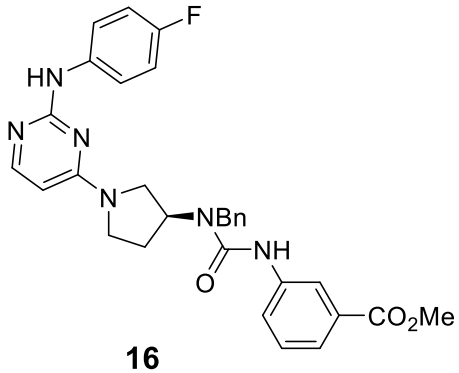


15

16 PrOF NMR titration with 5FW BPTF

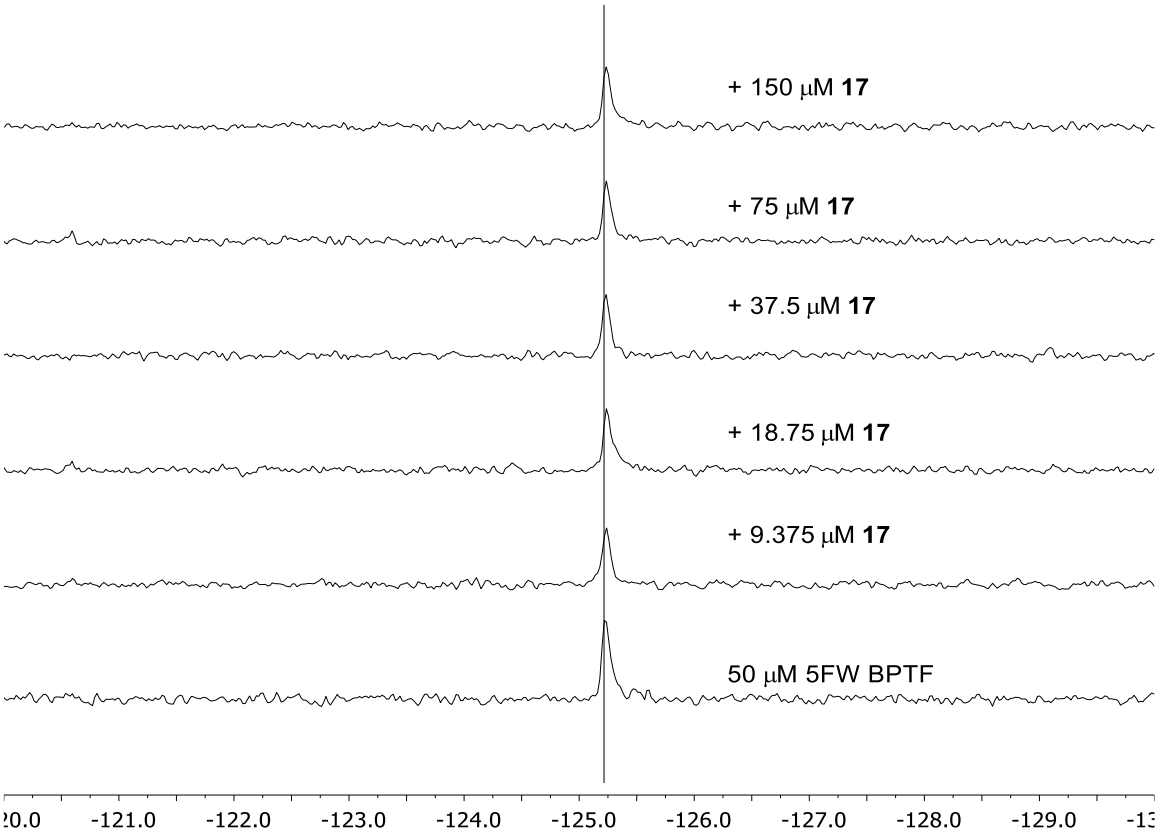


No binding observed

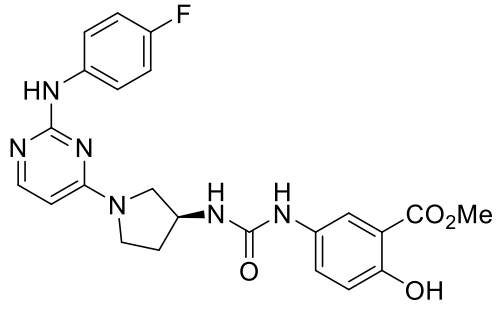


16

17 PrOF NMR titration with 5FW BPTF

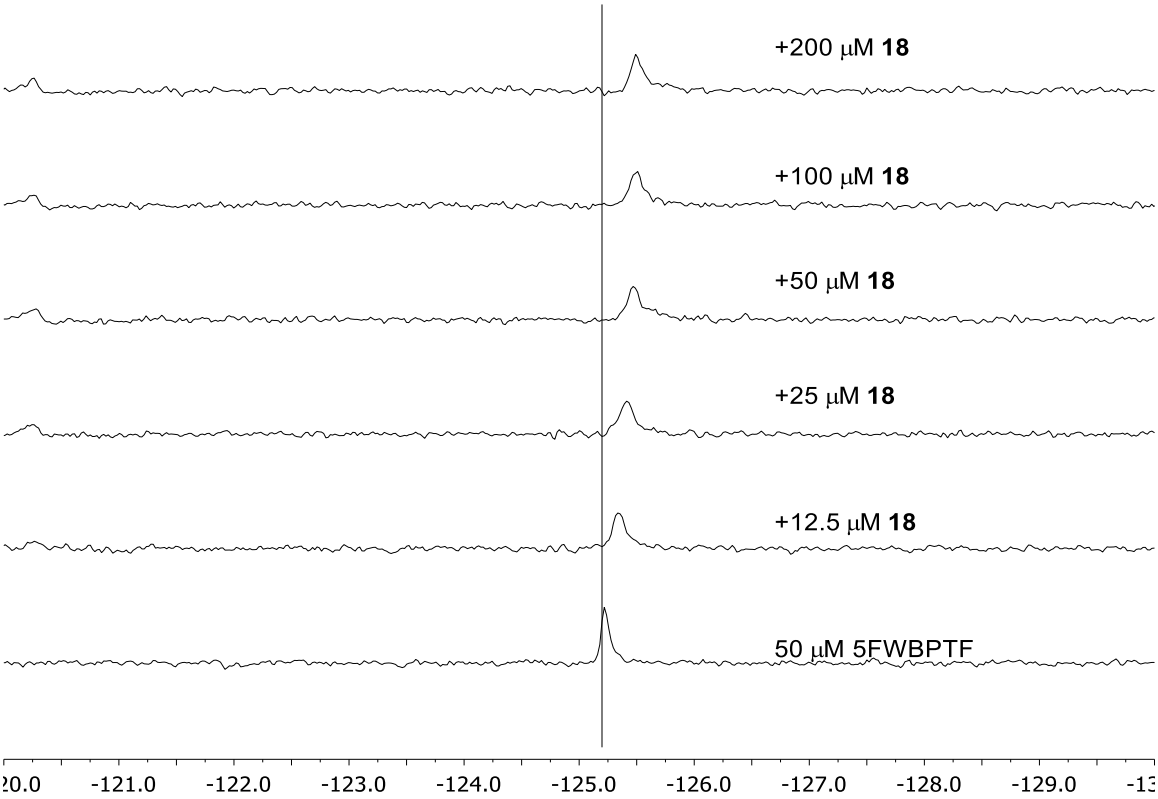


No binding observed



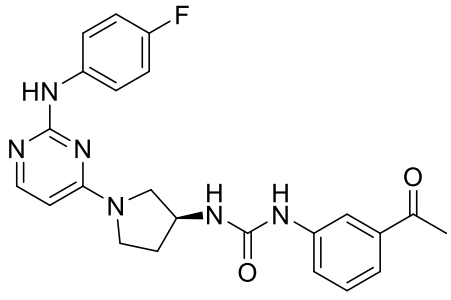
17

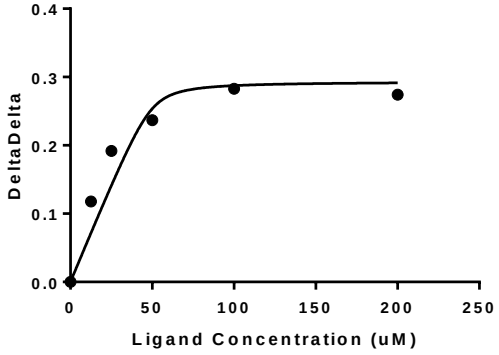
18 PrOF NMR titration with 5FW BPTF



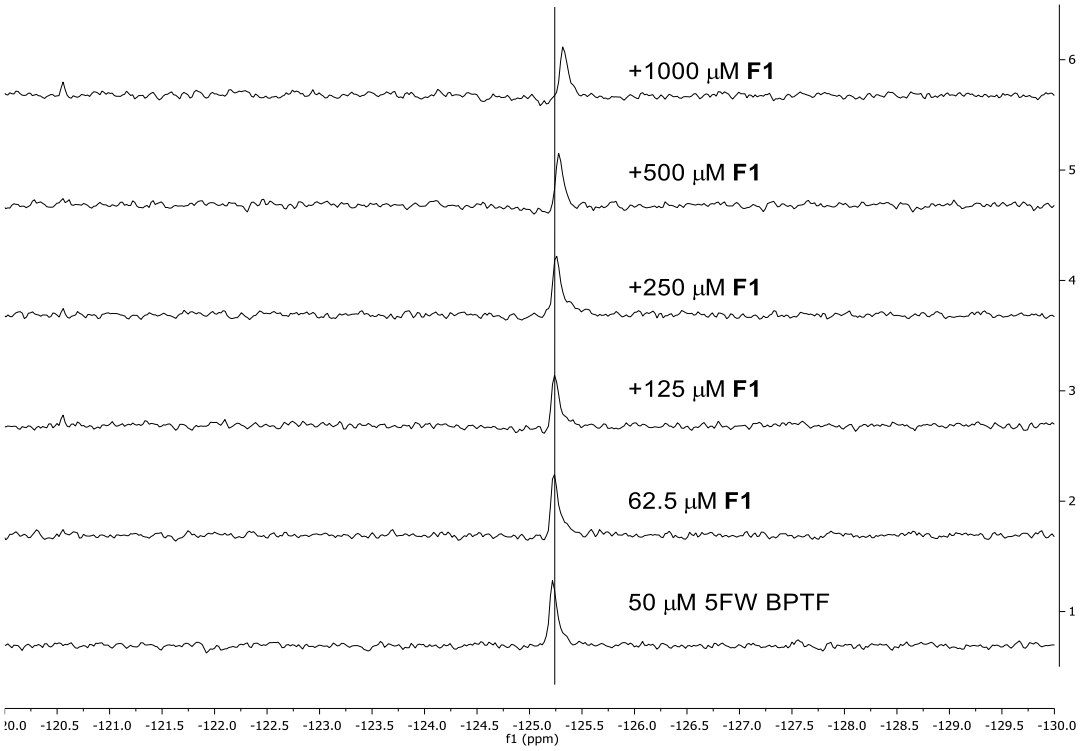
K_d = 1.08 +/- 3.0 μM

R^2 = 0.909

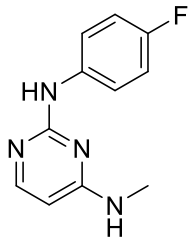




F1 PrOF NMR titration with 5FW BPTF

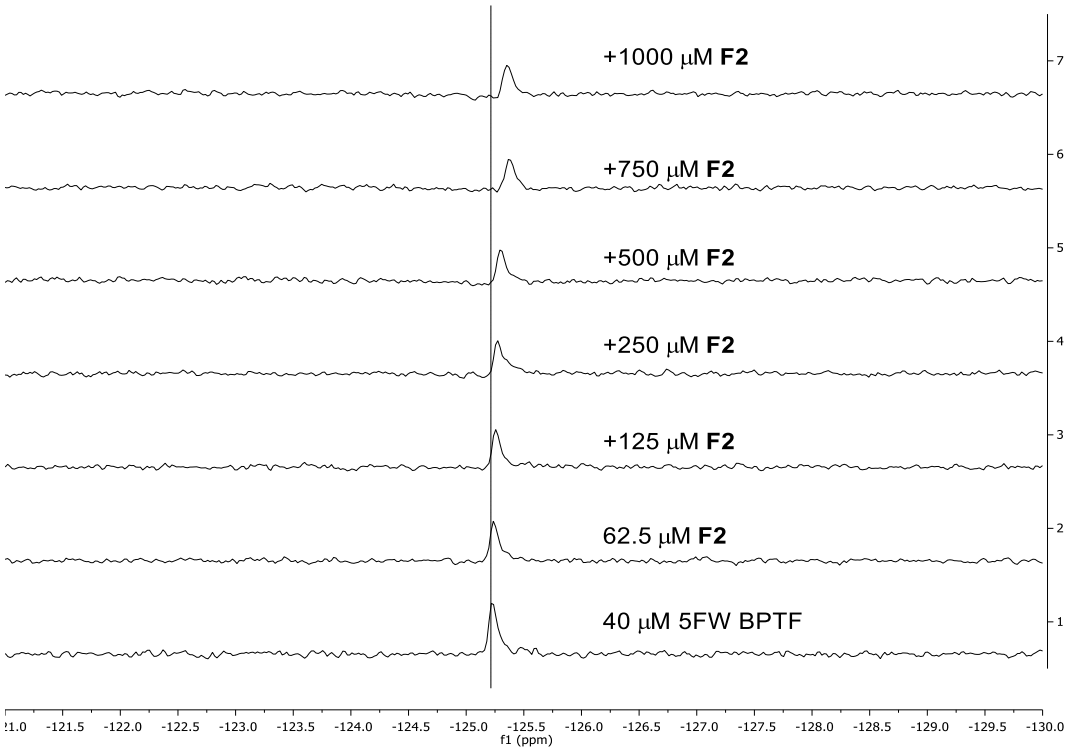


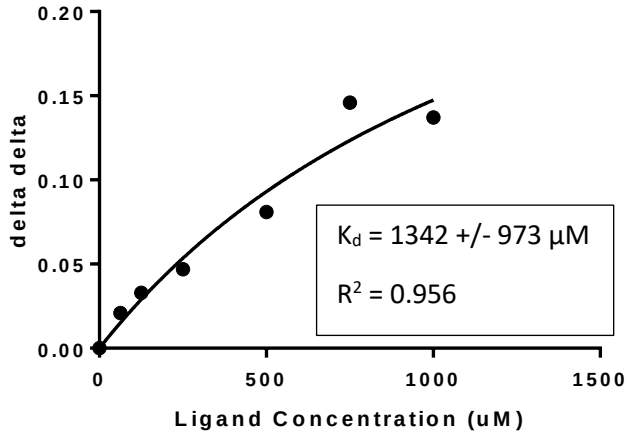
Chemical shift movement only observed at high concentrations, potentially due to non-specific binding.

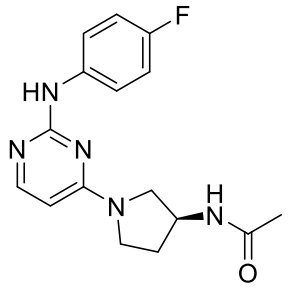


F1

F2 PrOF NMR titration with 5FW BPTF

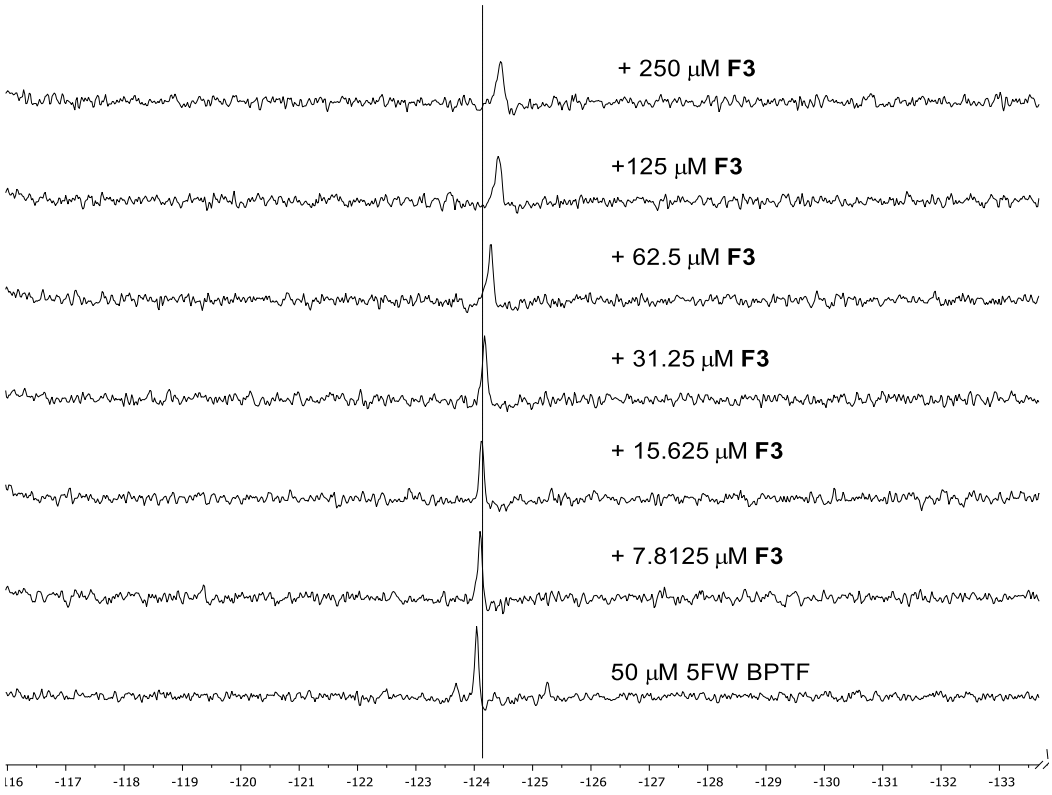


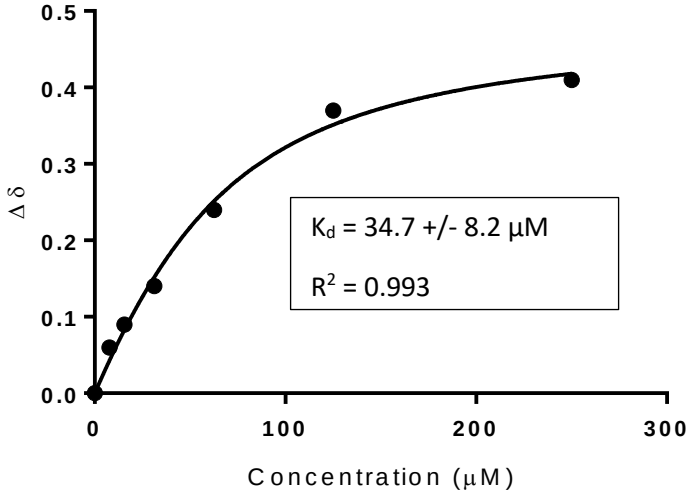


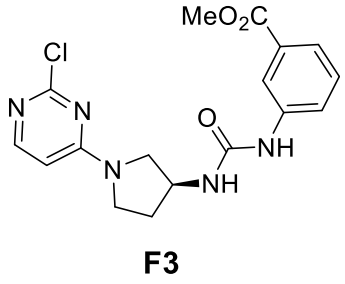


F2

F3 PrOF NMR titration with 5FW BPTF

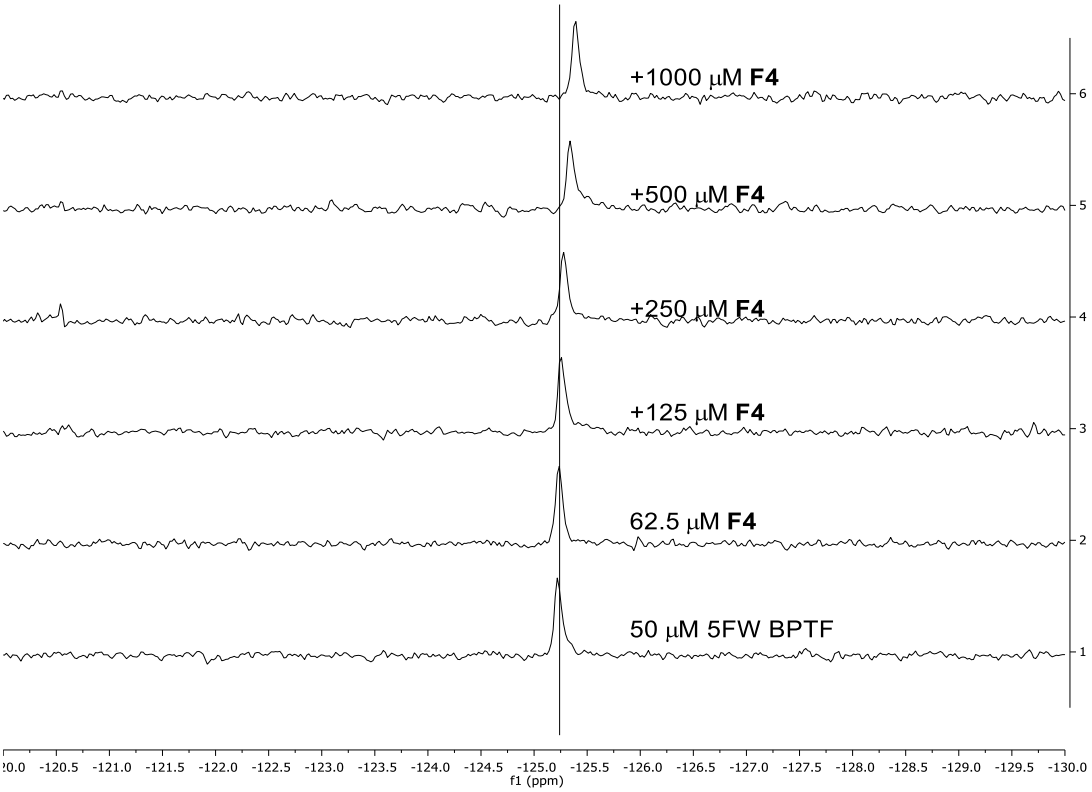


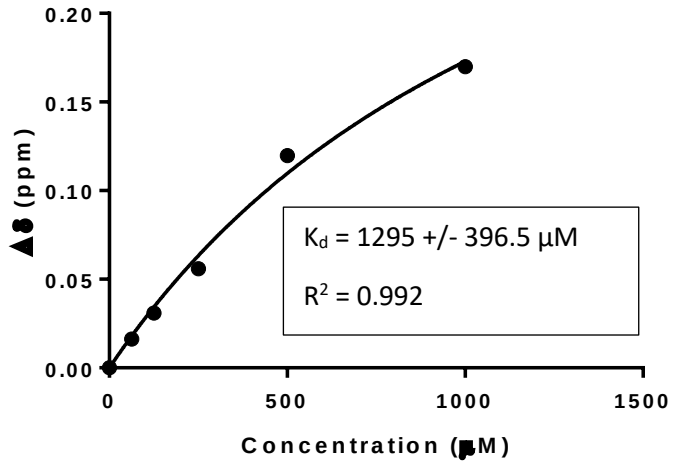


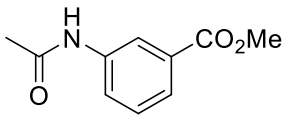


F3

F4 PrOF NMR titration with 5FW BPTF

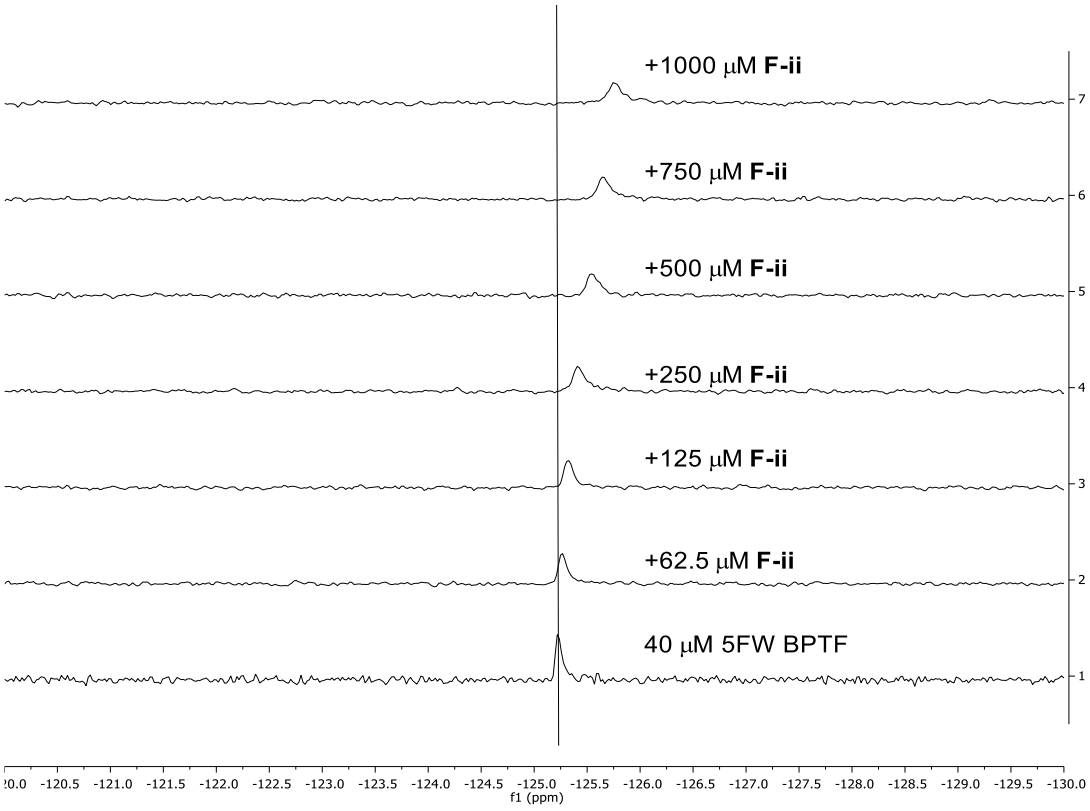


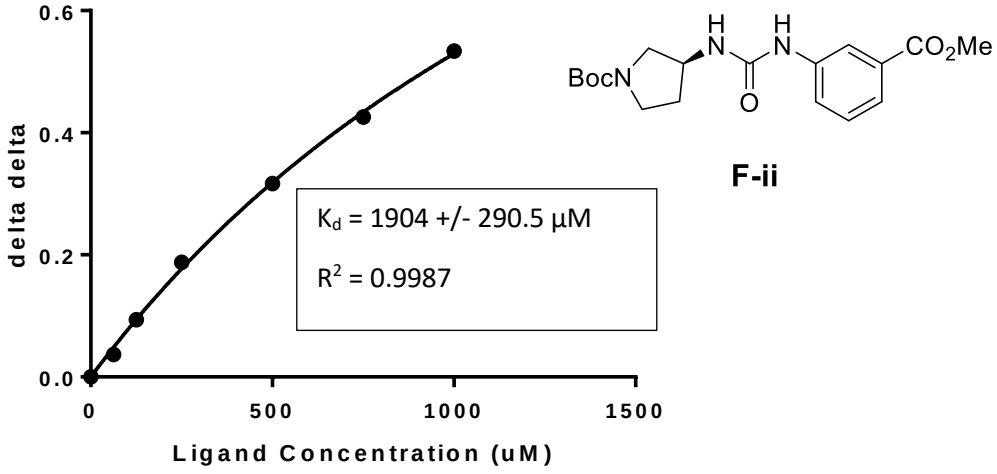




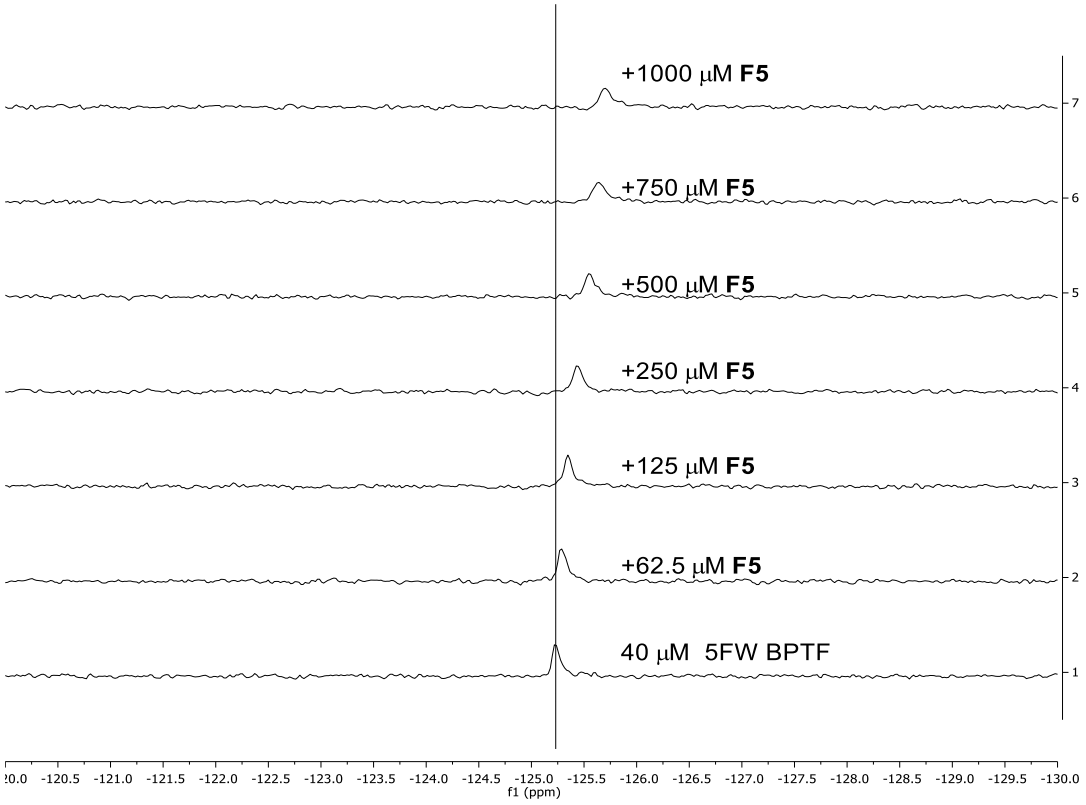
F4

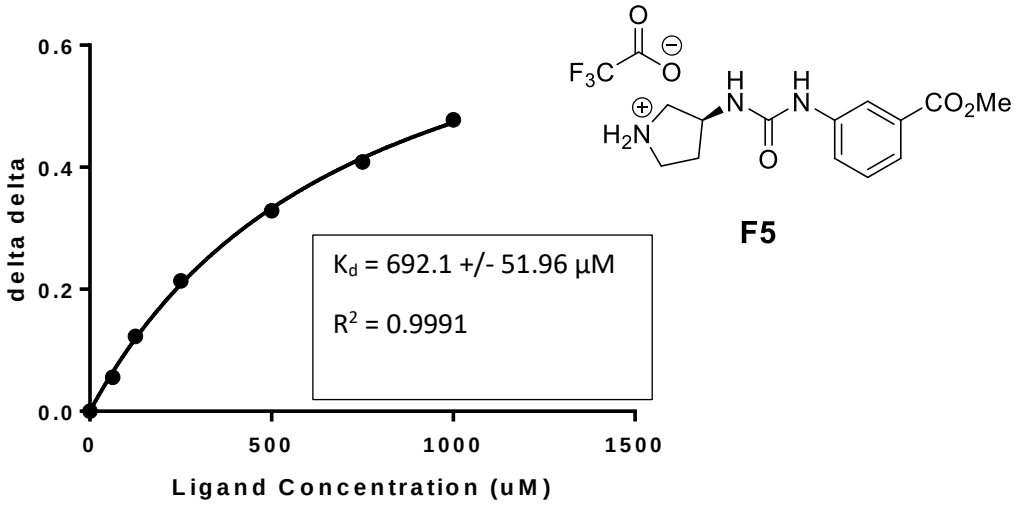
F-ii PrOF NMR titration with 5FW BPTF





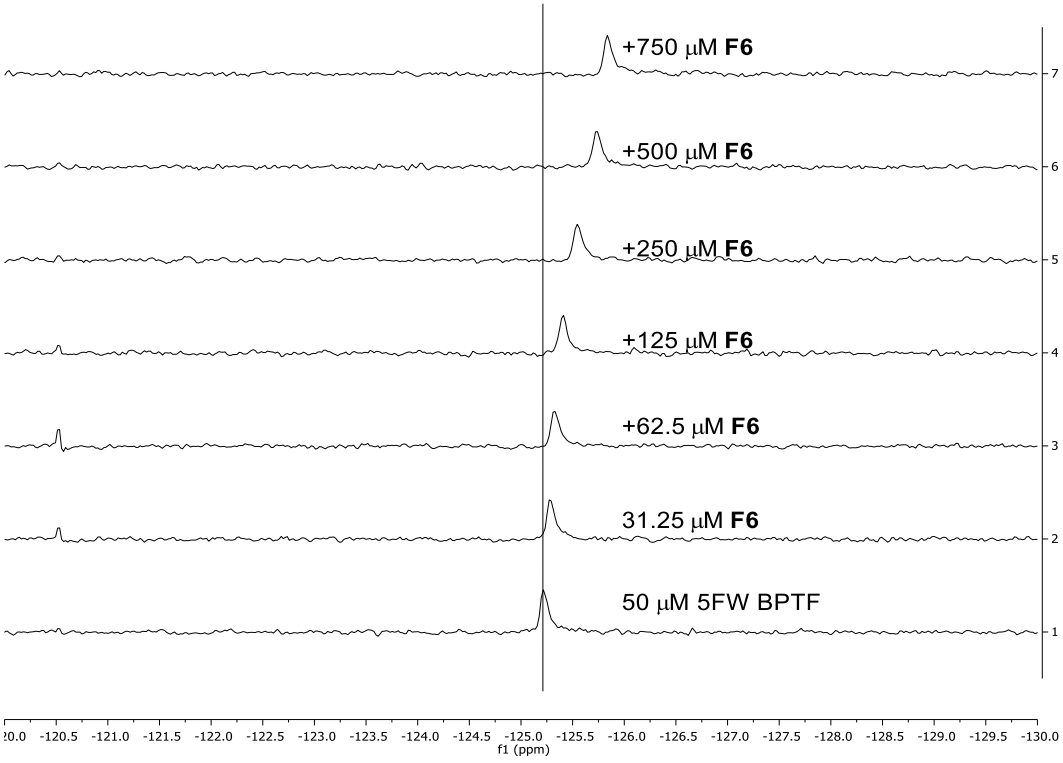
F5 PrOF NMR titration with 5FW BPTF

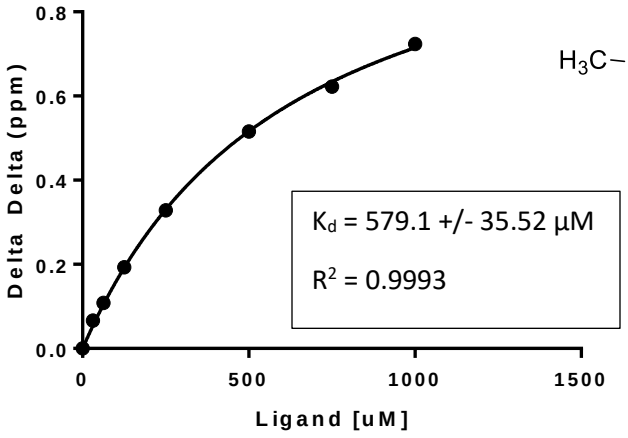


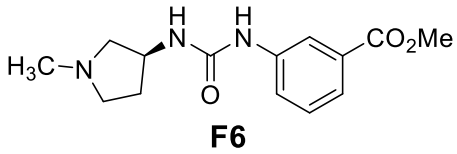


F5

F6 PrOF NMR titration with 5FW BPTF

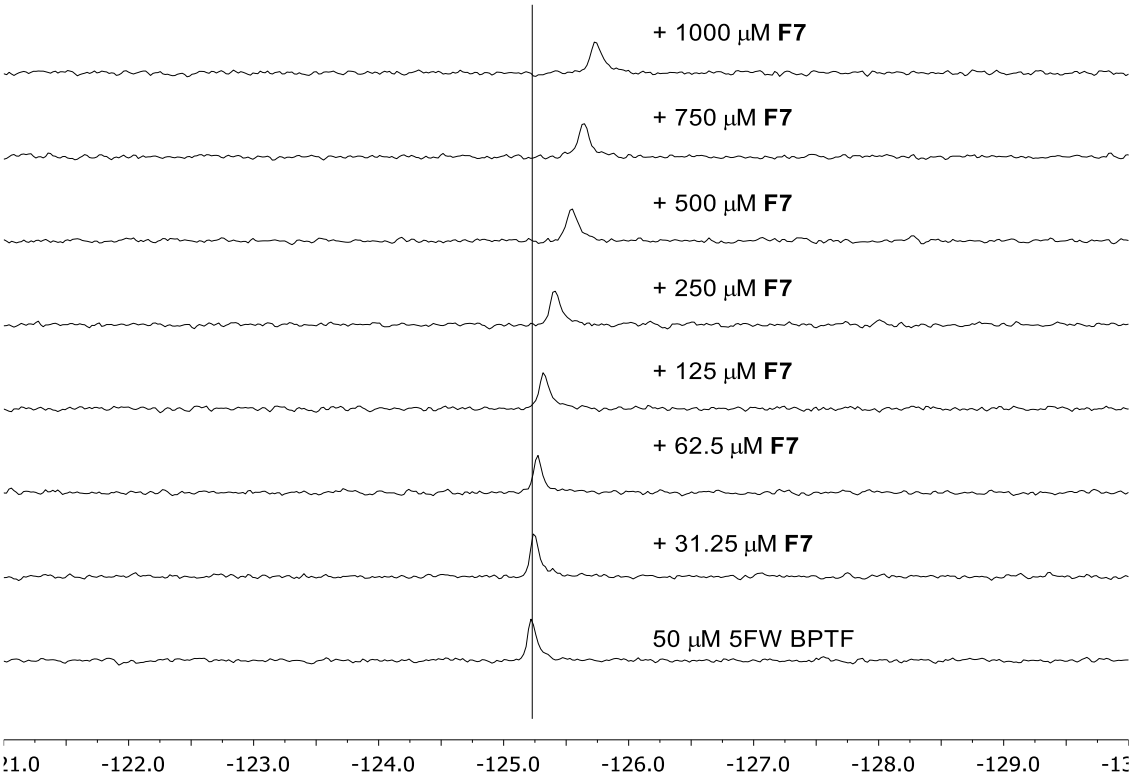


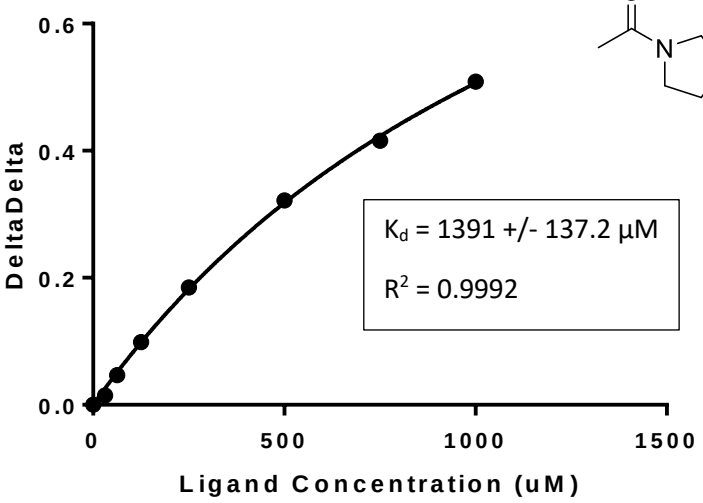


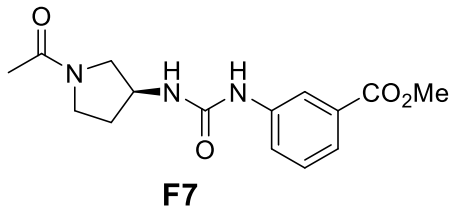


F6

F7 PrOF NMR titration with 5FW BPTF

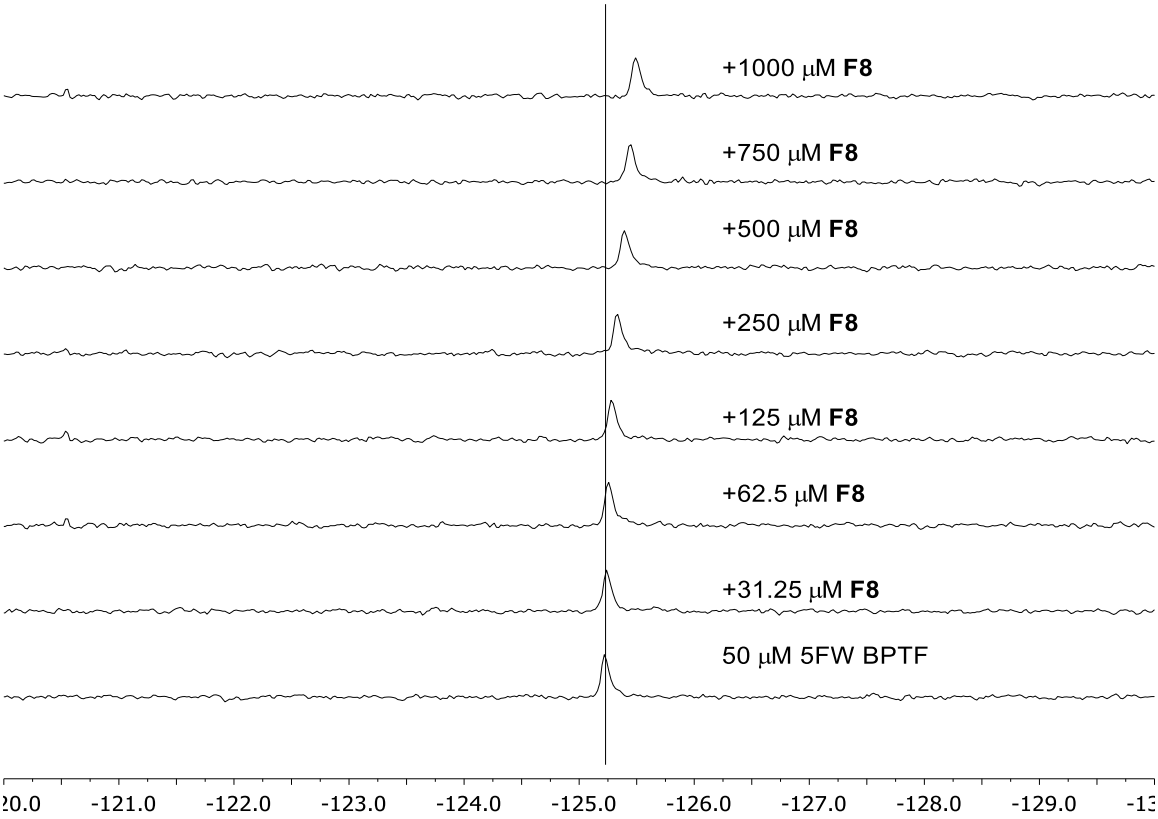






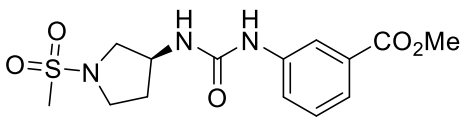
F7

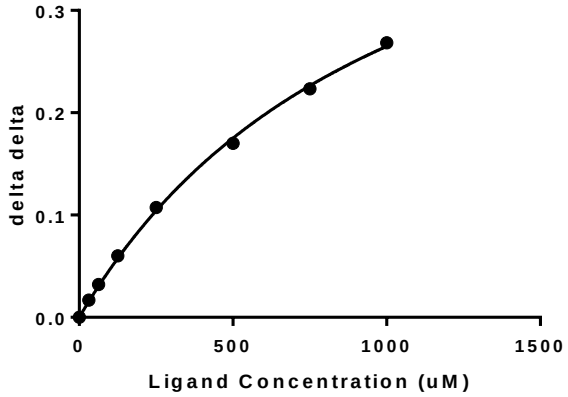
F8 PrOF NMR titration with 5FW BPTF



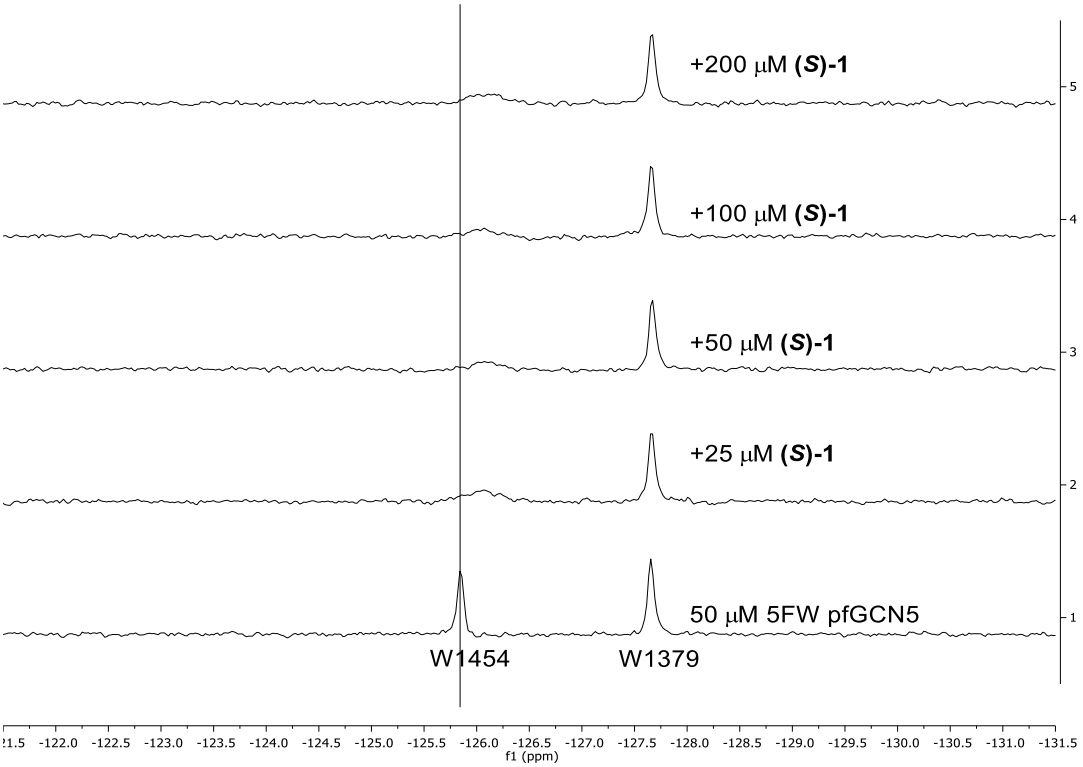
K_d = 991 +/- 86.5 μM

R^2 = 0.999

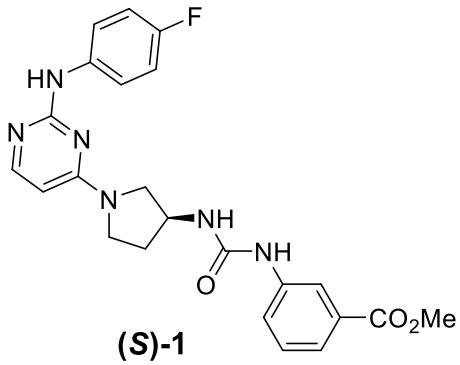




(*S*)-1 PrOF NMR titration with *Pf* GCN5

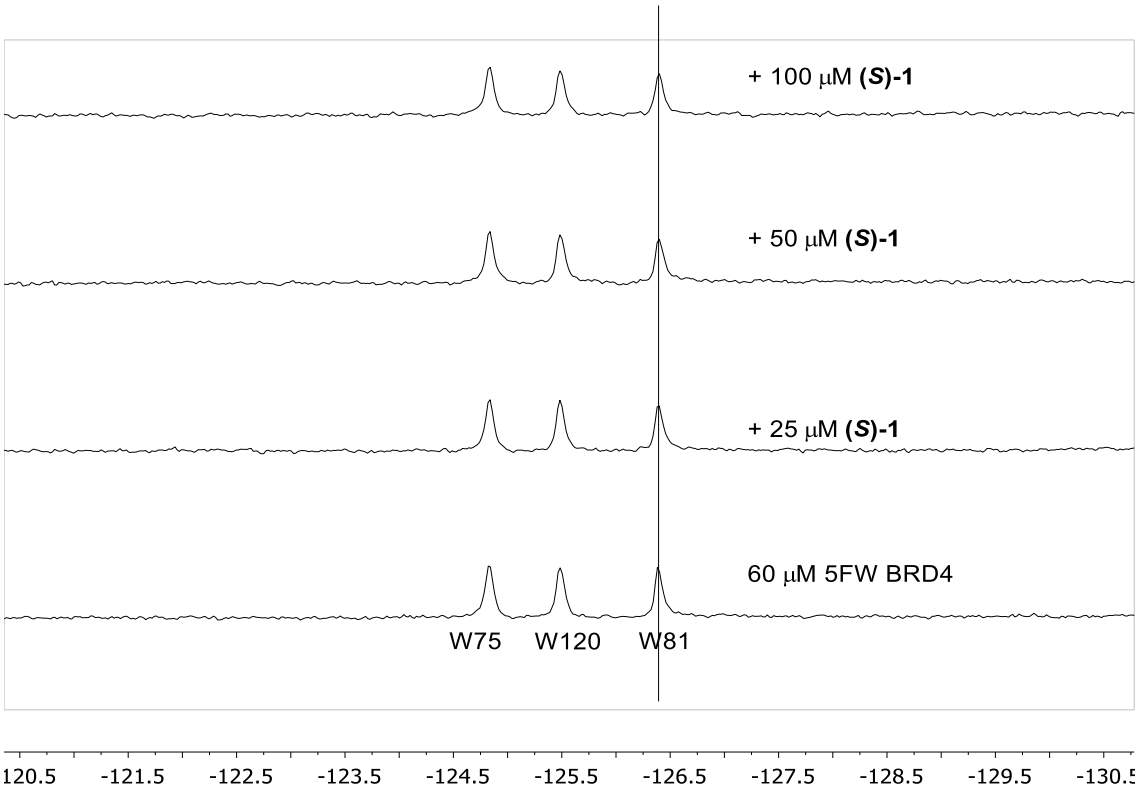


Pf GCN5 possesses two tryptophan residues, **(*S*)-1** impacts only W1454.

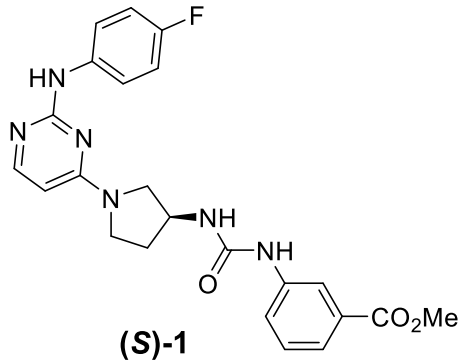


(*S*)-1

(*S*)-1 PrOF NMR titration with 5FW BRD4 (BD1)

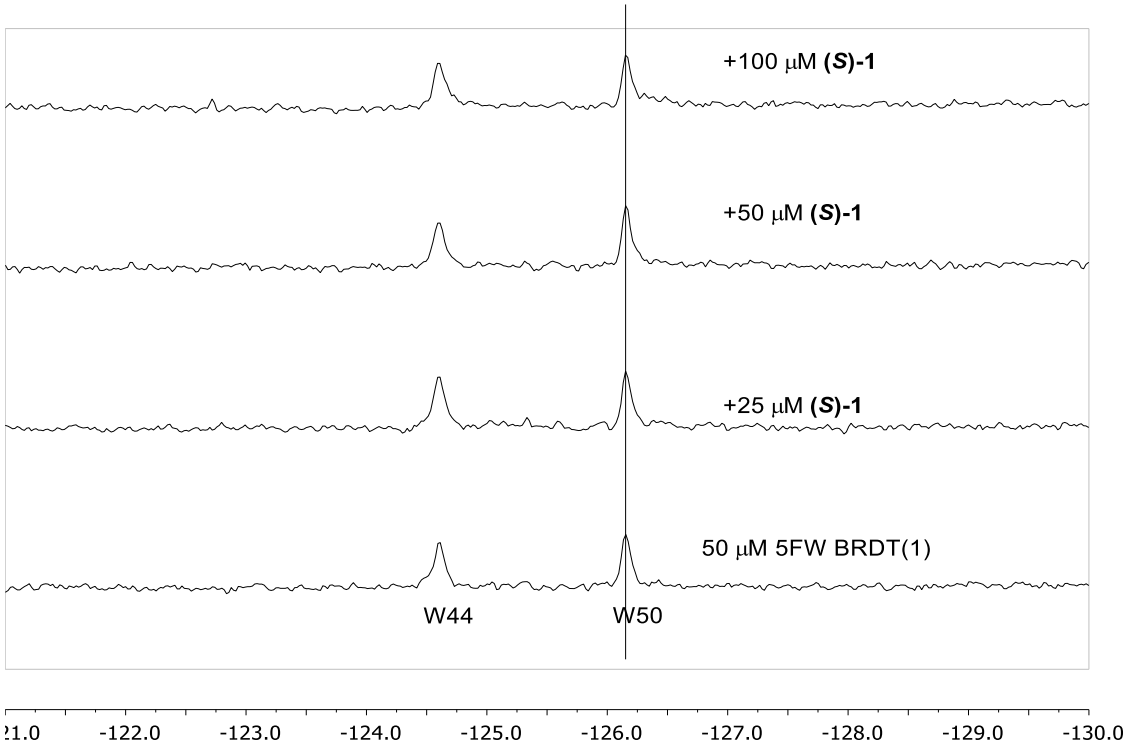


No binding is observed

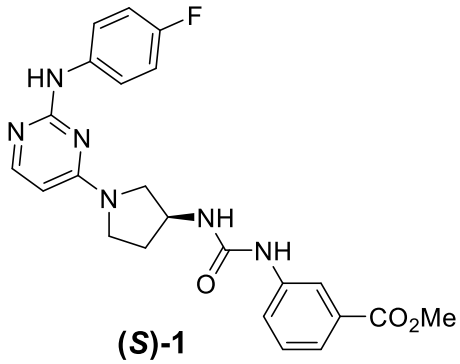


(*S*)-1

(*S*)-1 PrOF NMR titration with 5FW BRDT(BD1)

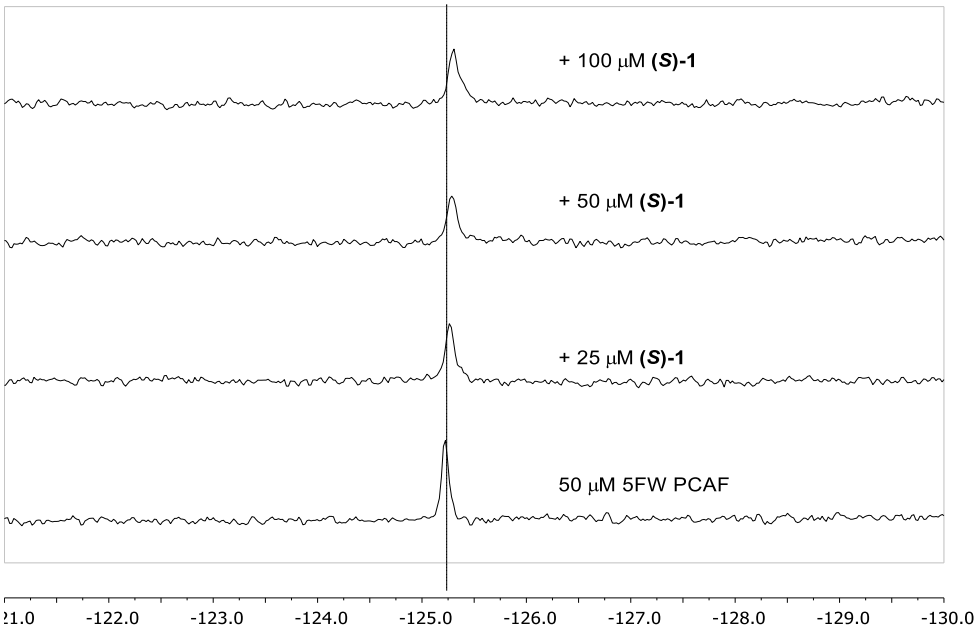


No binding is observed

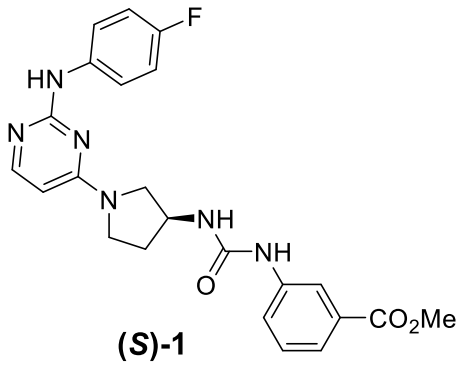


(*S*)-1

(*S*)-1 PrOF NMR titration with 5FW PCAF



Broadening and minor shift is observed, binding is ambiguous



(*S*)-1

4.12.14 Crystallization conditions and X-ray data collection methods

Unlabeled BPTF (445 µM, in 50 mM Tris, 100 mM NaCl, 10% (v/v) glycerol) was crystalized using the sitting drop method with 25% (w/v) polyethylene glycol 1500 and 10% (w/v) proprionate-cacodylate-bis tris propane. Crystals were harvested, cryoprotected with ethylene glycol and flash frozen. A RigakuMSC Micromax 007 X-ray generator with Saturn 944+ CCD Camera was used for data collection. The structure was solved using molecular replacement with Phaser-MR[17] and the PDB structure 3QZS. Phenix[18] refine and Coot were used for model structure refinement.

Unlabeled BPTF sequence:
MSTEDAMTVLTPLTEKDYEGLKRVLRSLQAHKMAWPFLEPVDPNDAPDYYGVI
KEPMDLATMEERVQRRY

5FW BPTF (247 µM, in 50 mM Tris, 100 mM NaCl, 10% (v/v) ethylene glycol) was crystalized at 20 ℃ using the hanging drop method with 100 mM magnesium acetate and 10% (v/v) polyethylene glycol 3500. Crystals were harvested, cryoprotected with ethylene glycol and flash frozen. A RigakuMSC Micromax 007 X-ray generator with Saturn 944+ CCD Camera was used for data collection. The structure was solved using molecular replacement with Phaser-MR[17] and PDB from the unlabeled BPTF. Phenix[18] refine and Coot were used for model structure refinement. Upon refinement there was insufficient electron density to model the N-terminal residues MST.

5FW BPTF Sequences:
EDAMTVLTPLTEKDYEGLKRVLRSLQAHKMAXPFLEPVDPNDAPDYYGVIKEP
MDLATMEERVQRRY

Table 4-3. Unlabeled BPTF Data collection and refinement statistics.	
Wavelength	
Resolution range	37.94 - 1.59 (1.647 - 1.59)
Space group	P 1 21 1
Unit cell	27.323 66.802 39.274 90 104.949 90
Total reflections	66630 (6749)
Unique reflections	18193 (1782)
Multiplicity	3.7 (3.7)
Completeness (%)	97.23 (96.69)
Mean I/sigma(I)	12.06 (2.78)
Wilson B-factor	15.33
R-merge	0.0975 (0.6611)
R-meas	0.114 (0.776)
R-pim	0.05836 (0.4006)
CC1/2	0.996 (0.528)
CC*	0.999 (0.832)
Reflections used in refinement	17872 (1782)
Reflections used for R-free	1788 (178)
R-work	0.1629 (0.2174)
R-free	0.1980 (0.2391)
CC(work)	0.968 (0.880)
CC(free)	0.949 (0.855)

Number of non-hydrogen atoms	1187
macromolecules	1017
ligands	16
solvent	154
Protein residues	122
RMS(bonds)	0.007
RMS(angles)	0.87
Ramachandran favored (%)	100
Ramachandran allowed (%)	0
Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	2.93
Average B-factor	21.05
macromolecules	19.27
ligands	40.86
solvent	30.73

Statistics for the highest-resolution shell are shown in parentheses.

Table 4-4: 5FW BPTF Data collection and refinement statistics.	
Wavelength	
Resolution range	26.4 - 2.065 (2.139 - 2.065)
Space group	P 1 21 1
Unit cell	27.335 66.763 39.26 90 105.016 90
Total reflections	14605 (761)
Unique reflections	7452 (395)
Multiplicity	2.0 (1.9)
Completeness (%)	88.59 (47.88)
Mean I/sigma(I)	19.77 (9.55)
Wilson B-factor	11.57
R-merge	0.03296 (0.09002)
R-meas	0.04661 (0.1273)
R-pim	0.03296 (0.09002)
CC1/2	0.998 (0.98)
CC*	0.999 (0.995)
Reflections used in refinement	7450 (395)
Reflections used for R-free	389 (17)
R-work	0.1535 (0.1522)
R-free	0.2148 (0.3176)
CC(work)	0.965 (0.955)
CC(free)	0.938 (0.673)

Number of non-hydrogen atoms	1168
macromolecules	987
ligands	5
solvent	176
Protein residues	119
RMS(bonds)	0.014
RMS(angles)	1.47
Ramachandran favored (%)	98.25
Ramachandran allowed (%)	1.75
Ramachandran outliers (%)	0
Rotamer outliers (%)	0
Clashscore	6.7
Average B-factor	17.89
macromolecules	16.61
ligands	16.58
solvent	25.08
Statistics for the highest-resolution shell are shown in parentheses.	

RMSD alignment of 5FW and Unlabeled BPTF

RMSD alignment was done using secondary structure matching (SSM).

Moving (Unlabeled BPTF):
MSTEDAMTVLTPLTEKDYEGLKRVLRSLQAHKMAWPFLEPVDPNDAPDYYGVI
KEPMDLATMEERVQRRY

Target (5FW BPTF):
EDAMTVLTPLTEKDYEGLKRVLRSLQAHKMAXPFLEPVDPNDAPDYYGVIKEP
MDLATMEERVQRRY

Rotation - euler (alpha,beta,gamma) 7.2267 179.7618 7.3379

Translation - Angstroms 6.6968 -4.2895 75.8852

INFO: core rmsd achieved: 0.1548 Angstroms

number of residues in reference structure: 119

number of residues in moving structure: 122

number of residues in aligned sections (reference): 119

number of residues in aligned sections (moving): 122

number of aligned residues: 116

number of gaps: 0

number of misdirections: 0

number of SSE combinations: 5.0000

sequence identity: 99.1379%

Table 4-5: RMSD values between Unlabeled BPTF (moving) and 5FW BPTF (Reference)

Moving Reference Distance(Å)	Moving Reference Distance(Å)
A 63 <---> A 63 : 0.1642	A 92 <---> A 92 : 0.1880
A 64 <---> A 64 : 0.1265	A 93 <---> A 93 : 0.1720
A 65 <---> A 65 : 0.1667	A 94 <---> A 94 : 0.0843
A 66 <---> A 66 : 0.1048	A 95 <---> A 95 : 0.1040
A 67 <---> A 67 : 0.0880	A 96 <---> A 96 : 0.1194
A 68 <---> A 68 : 0.1451	A 97 <---> A 97 : 0.1019
A 69 <---> A 69 : 0.1458	A 98 <---> A 98 : 0.1082
A 70 <---> A 70 : 0.0596	A 99 <---> A 99 : 0.1454
A 71 <---> A 71 : 0.1132	A 100 <---> A 100 : 0.1224
A 72 <---> A 72 : 0.2768	A 101 <---> A 101 : 0.2340
A 73 <---> A 73 : 0.0703	A 102 <---> A 102 : 0.0264
A 74 <---> A 74 : 0.0846	A 103 <---> A 103 : 0.8093
A 75 <---> A 75 : 0.1051	A 104 <---> A 104 : 0.6393
A 76 <---> A 76 : 0.1141	A 105 <---> A 105 : 0.0656
A 77 <---> A 77 : 0.0625	A 106 <---> A 106 : 0.0509
A 78 <---> A 78 : 0.0609	A 107 <---> A 107 : 0.1190
A 79 <---> A 79 : 0.0577	A 108 <---> A 108 : 0.1670
A 80 <---> A 80 : 0.0525	A 109 <---> A 109 : 0.1060
A 81 <---> A 81 : 0.1388	A 110 <---> A 110 : 0.1957
A 82 <---> A 82 : 0.0905	A 111 <---> A 111 : 0.1312
A 83 <---> A 83 : 0.1282	A 112 <---> A 112 : 0.1246
A 84 <---> A 84 : 0.0782	A 113 <---> A 113 : 0.0562
A 85 <---> A 85 : 0.0528	A 114 <---> A 114 : 0.0640
A 86 <---> A 86 : 0.0905	A 115 <---> A 115 : 0.0450
A 87 <---> A 87 : 0.1051	A 116 <---> A 116 : 0.0267
A 88 <---> A 88 : 0.1744	A 117 <---> A 117 : 0.0703
A 89 <---> A 89 : 0.1191	A 118 <---> A 118 : 0.0569
A 90 <---> A 90 : 0.1005	A 119 <---> A 119 : 0.0368
A 91 <---> A 91 : 0.3516	A 120 <---> A 120 : 0.015

Moving	Reference	Distance(Å)
A 121	A 121	0.0701
A 122	A 122	0.0763
A 123	A 123	0.0910
A 124	A 124	0.1337
A 125	A 125	0.1118
A 126	A 126	0.0812
A 127	A 127	0.0404
A 128	A 128	0.0887
A 129	A 129	0.0566
A 130	A 130	0.0505
A 131	A 131	0.0686
A 132	A 132	0.1039
A 133	A 133	0.0669
A 134	A 134	0.0680
A 135	A 135	0.0527
A 136	A 136	0.0622
A 137	A 137	0.0533
A 138	A 138	0.0389
A 139	A 139	0.0821
A 140	A 140	0.0450
A 141	A 141	0.0785
A 142	A 142	0.0668
A 143	A 143	0.0401
A 144	A 144	0.0711
A 145	A 145	0.0979
A 146	A 146	0.0130
A 147	A 147	0.0300
A 148	A 148	0.0198
A 149	A 149	0.1103
A 150	A 150	0.3136
A 151	A 151	0.1244
A 152	A 152	0.1499
A 153	A 153	0.1176
A 154	A 154	0.0437
A 155	A 155	0.0465
A 156	A 156	0.0582
A 157	A 157	0.0507
A 158	A 158	0.0629
A 159	A 159	0.0737
A 160	A 160	0.0671
A 161	A 161	0.0948
A 162	A 162	0.0849
A 163	A 163	0.0788
A 164	A 164	0.0741
A 165	A 165	0.0412
A 166	A 166	0.0706
A 167	A 167	0.0464
A 168	A 168	0.0335
A 169	A 169	0.0832
A 170	A 170	0.0986
A 171	A 171	0.1001
A 172	A 172	0.1088

Chapter 5

^{19}F NMR analysis of Human Serum Albumin-binding peptides

Motivation: The following work is my contribution in a collaborative effort with the lab of Prof. Ratmir Derda at the University of Alberta. His lab has developed fluorinated crosslinked peptides which bind to human serum albumin. The goal of this work is to develop these fluorinated peptides as auxiliaries which could potentially be used to elongate the half-life of circulating therapeutics in vivo by limiting their ability to be cleared from the body. My contribution involves the testing of peptides synthesized in the Derda lab by ^{19}F NMR to determine relative binding affinities and explore potential binding sites.

5.1 Introduction

Both covalent and non-covalent association of therapeutic molecules to blood plasma proteins are widely accepted methods of elongating the half-life of drugs in the body.[182] Clinical trials have been undertaken, and many approved, of covalent conjugation or non-covalent association to human serum albumin (HSA), a protein present in blood plasma. Attaching drugs to plasma proteins is a method for prolonging the half-life of drugs in vivo from minutes when unbound, to several hours to days when associated to HSA.[183,184] Recently, the lab of Ratmir Derda has sought to utilize this concept to associate peptides to HSA for development auxiliaries for peptide-based drugs. Derda and co-workers have developed a means to crosslink two cysteine residues via a nucleophilic aromatic substitution reaction within a peptide scaffold surrounding a short motif with the capability to bind HSA with a perfluoro biphenylsulfide (PFS) crosslink (Figure 5-1).[185] Due to stability issues, the sulfone moiety described in their 2016 report has been replaced with sulfide linkers.[185] The Derda lab has developed several peptides of this nature, but their binding affinity to HSA and binding site remain ambiguous. The sequences utilized to bind with HSA were developed in a phage display assay. In a collaboration with the Derda lab, we used ^{19}F NMR to analyze the binding characteristics of five peptide sequences in both direct-binding and competitive experiments.

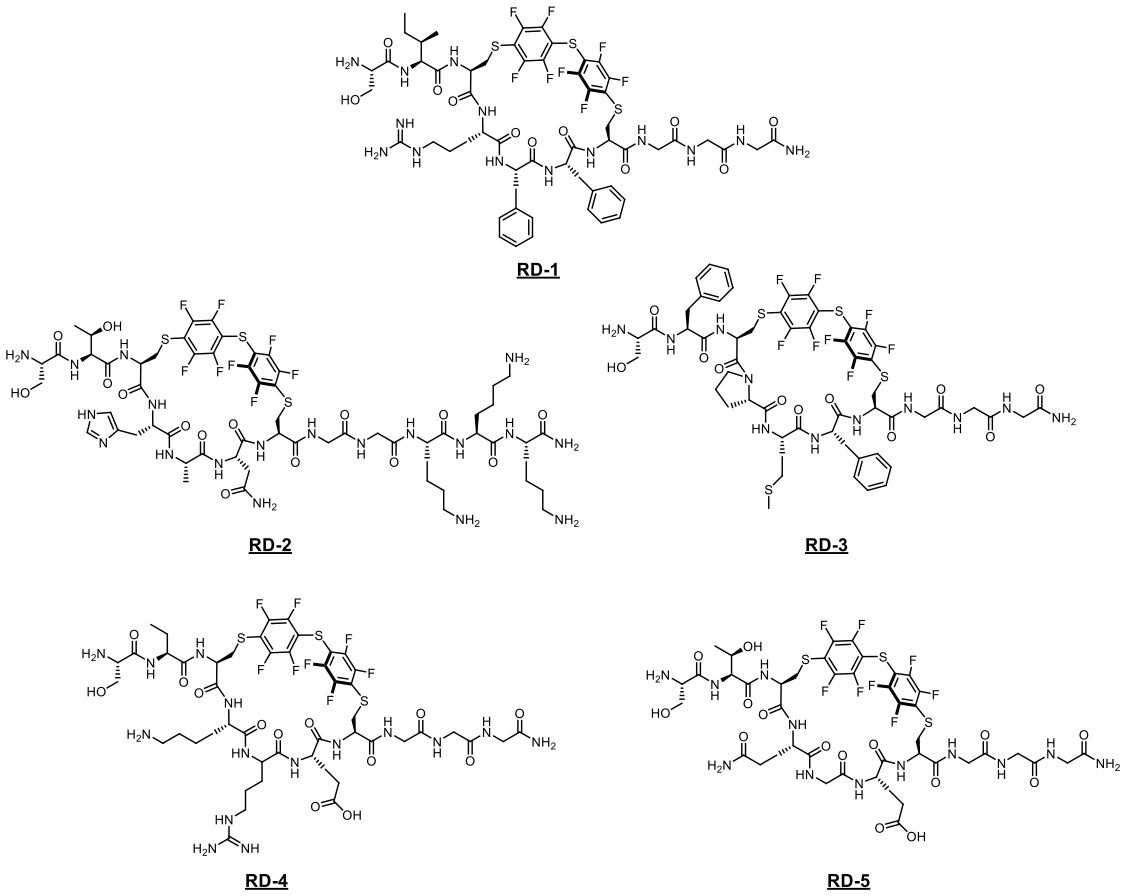


Figure 5-1: Structures of perfluoro biphenylsulfide (PFS) peptides

In contrast to protein observed fluorine (PrOF) NMR that has been discussed in chapters 1 and 3, we utilized the ^{19}F resonances directly from the peptide scaffold and titrated them against varied concentrations of HSA. The objective was to analyze the fluorine resonances of peptides **RD-1** – **RD-5** (Figure 5-1) and determine dissociation constants for each peptide to establish a ranking of binding strength. A second goal of our collaboration was to gain insight as to the binding site of these peptide motifs on HSA. Binding site identification would be performed using three known HSA binders: ibuprofen,[186] diclofenac-Na,[187] and carbamazepine[188] (Figure 5-2). Sudlow et al described two major binding sites for drugs to HSA, termed Site I and Site II. Site I is primarily bound by bulky heterocyclic compounds, while Site II is typically occupied by aryl carboxylates.[189,190] As such, the three drugs chosen all have a range of affinities for these two sites. Ibuprofen binds tightest with a K_d = 370 nM to Site II[190], diclofenac-Na binds

both sites (Site I K_d = 17 μM, Site II K_d = 2 μM)[187], while carbamazepine binds Site I with K_d = 190 μM[188]. These three molecules were each mixed in solution with HSA and the tightest binding PFS-peptide in an attempt to displace the small molecule drug to elucidate binding site.

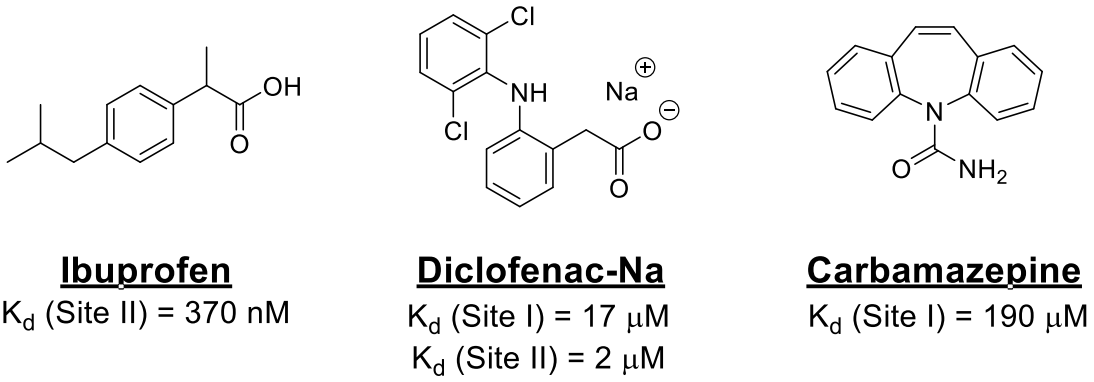


Figure 5-2: Structures of small molecule binders of human serum albumin.

5.2 Binding affinity ranking of PFS peptides

To determine a relative order of binding of the five peptides on hand, we employed ^{19}F NMR to observe the perturbation of the four resonances displayed from the fluorinated peptides. A working PFS peptide concentration was first determined, with resonances having a clear signal to noise ratio (signal:noise = 20 was used here). Based on these parameters a 20 μM solution of peptide was determined to yield a sufficient ^{19}F NMR spectrum for binding experiments. For each peptide tested, a baseline ^{19}F NMR spectrum was taken, followed by a titration of HSA from 160 μM, to 10 μM using two-fold serial dilutions. (Figures 5-3 through 5-7). In all cases, the chemical shift perturbation did not exhibit fast chemical exchange kinetics, prohibiting accurate K_d determination in our assay. In these cases considerable broadening was observed preventing accurate determination of chemical shift, and was consistent with slow to intermediate chemical exchange. As an alternative semi-quantitative method, we rank-ordered the strength of binding based on the HSA concentrations at which the ^{19}F resonances broadened into baseline (see chapter 1 for a more detailed description of chemical exchange analysis). In the case of **RD-1**, ^{19}F resonances were broadened completely by stoichiometric

concentrations of peptide and HSA. Conversely, **RD-5** showed considerably less broadening, as distinct resonances were observed at the highest concentration tested, indicating faster chemical exchange and thus a higher K_d. The relative binding of each peptide is ordered such that **RD-1** is considered the strongest binder, while **RD-5** is considered the weakest.

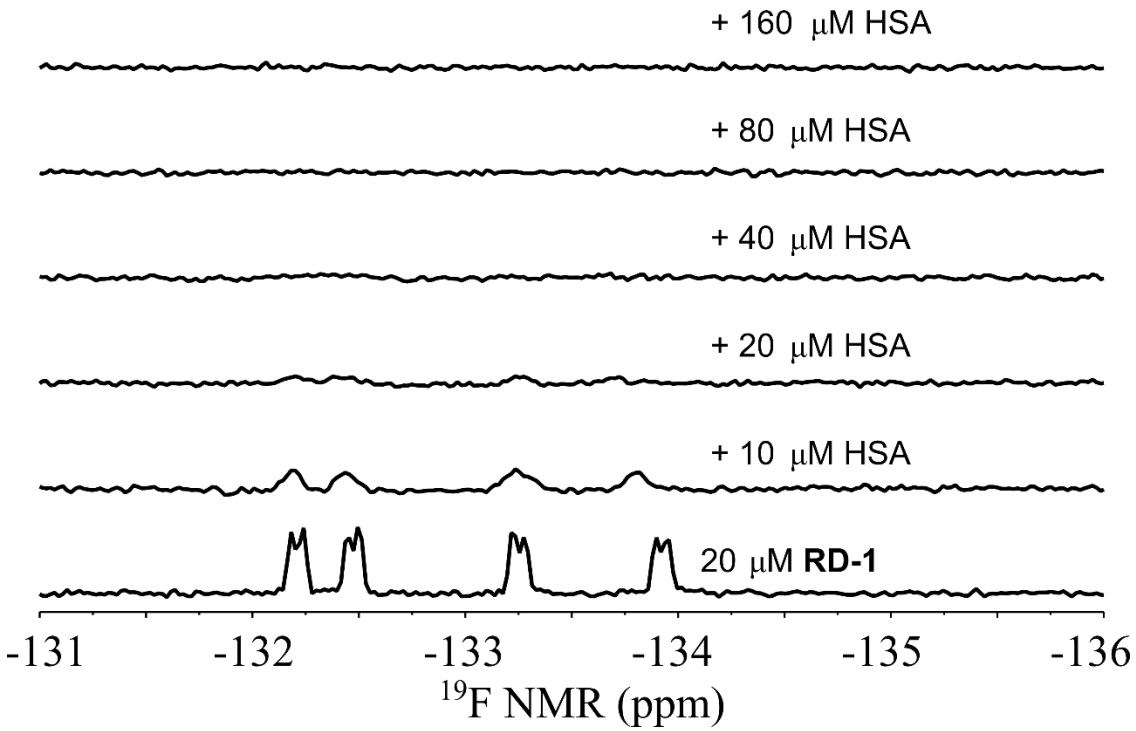


Figure 5-3: ^{19}F NMR titration of **RD-1** with increasing concentrations of HSA.

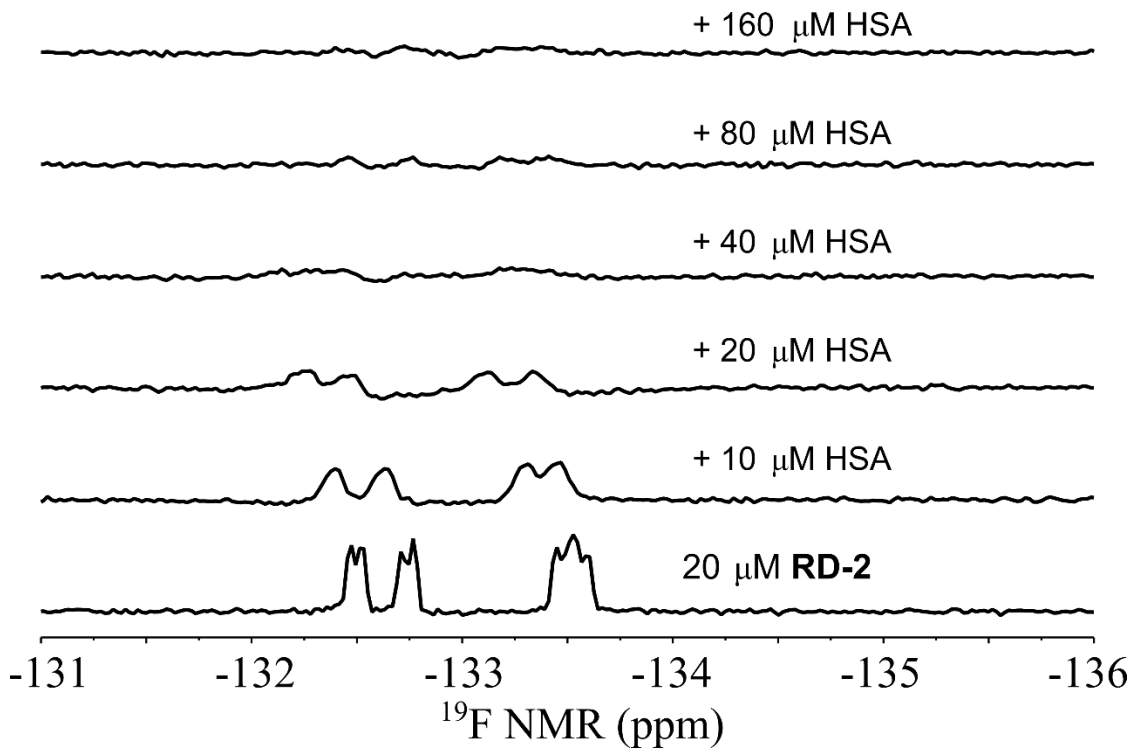


Figure 5-4: ^{19}F NMR titration of **RD-2** with increasing concentrations of HSA.

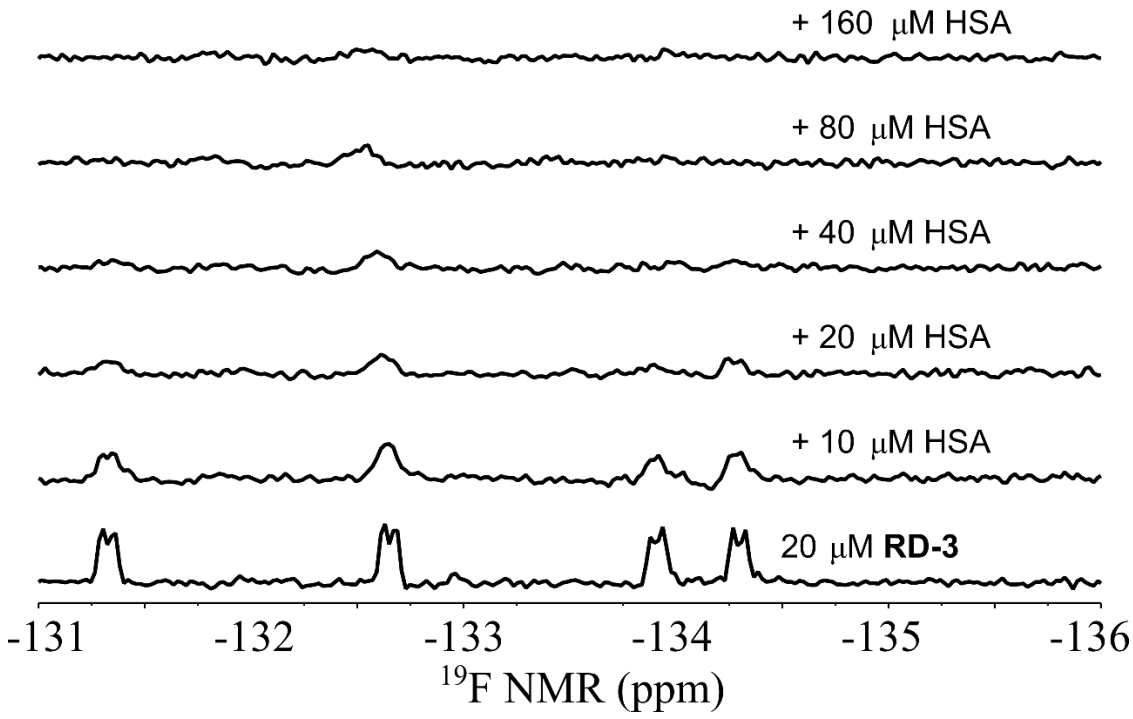


Figure 5-5: ^{19}F NMR titration of **RD-3** with increasing concentrations of HSA.

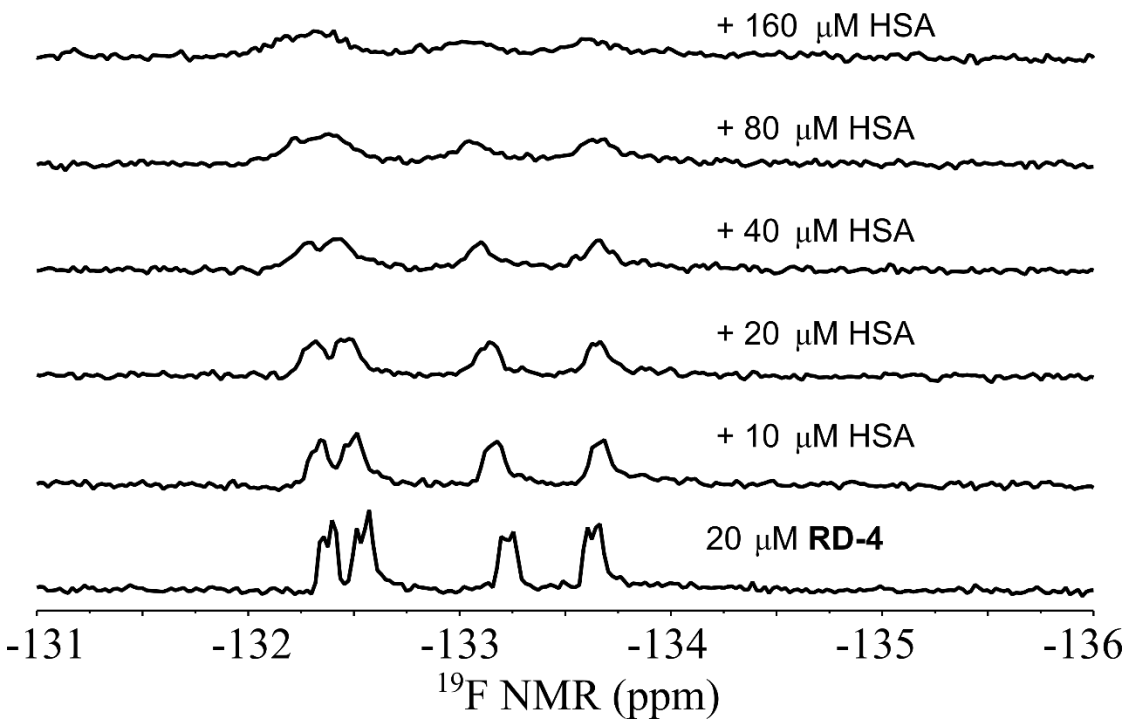


Figure 5-6: ^{19}F NMR titration of **RD-4** with increasing concentrations of HSA.

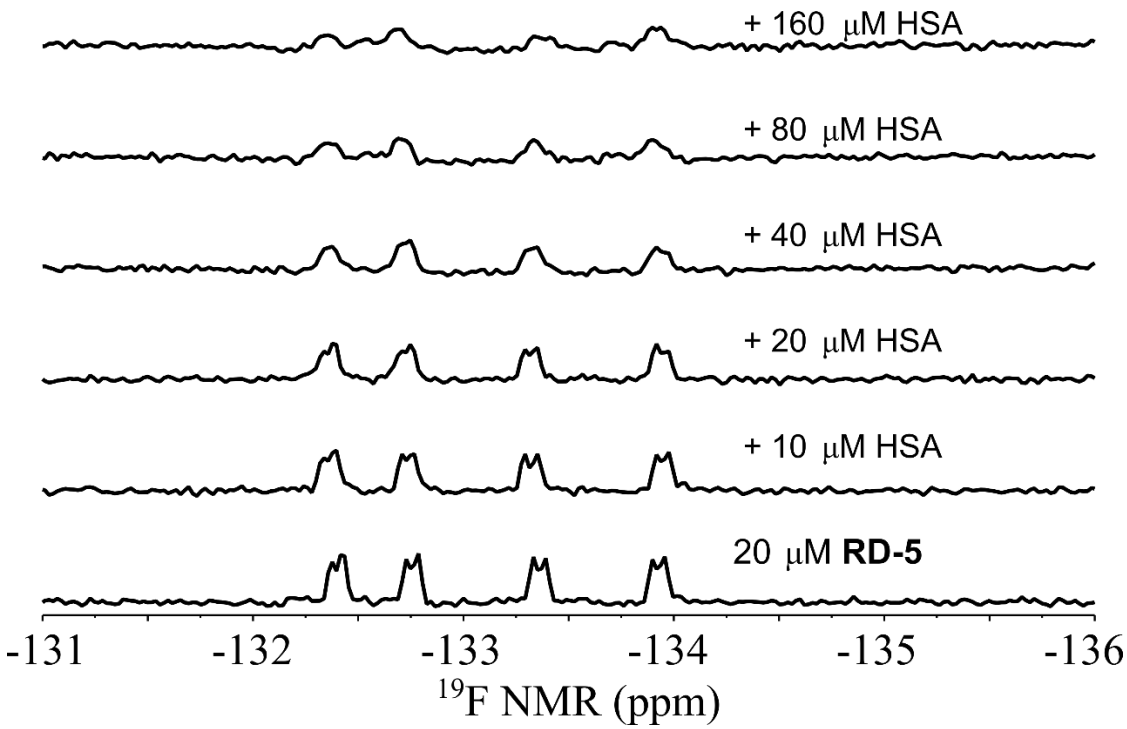


Figure 5-7: ^{19}F NMR titration of **RD-5** with increasing concentrations of HSA.

5.3 Competitive binding analysis with known HSA-binding drugs

With a relative ranking of affinity in hand, we next sought to determine if these fluorinated peptides shared a binding site with other established HSA-binding compounds. **RD-1** showed evidence of stoichiometric binding when the ^{19}F NMR spectrum was taken at equimolar concentrations (20 μM each). This was chosen as the starting point for competition experiments. The objective of the competition experiments was to observe a return in ^{19}F resonances from the peptide when it was outcompeted for its binding site by a small molecule drug. Two concentrations (100 μM and 1000 μM) were chosen to observe competitive binding. Three small molecules: ibuprofen, diclofenac-Na, and carbamazepine were chosen as competitors (Figure 5-2). However, in all three cases no return of ^{19}F resonance was observed for **RD-1** (Figures 5-8 – 5-10). The lack of a returned resonance can indicate that **RD-1** doesn't share a binding site with any of these drugs, it is unlikely that **RD-1** is such a strong binder as to see no return of a resonance as the drugs were used in such a dramatic excess of **RD-1**.

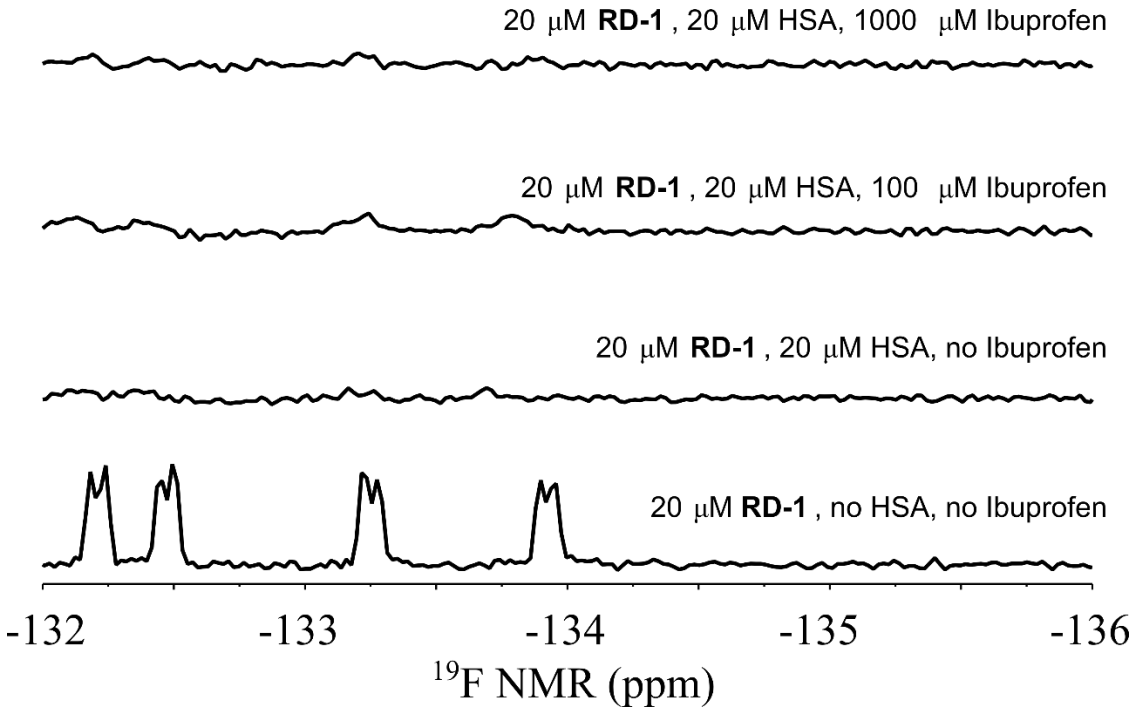


Figure 5-8: ^{19}F NMR competition experiment with **RD-1**, HSA, and ibuprofen.

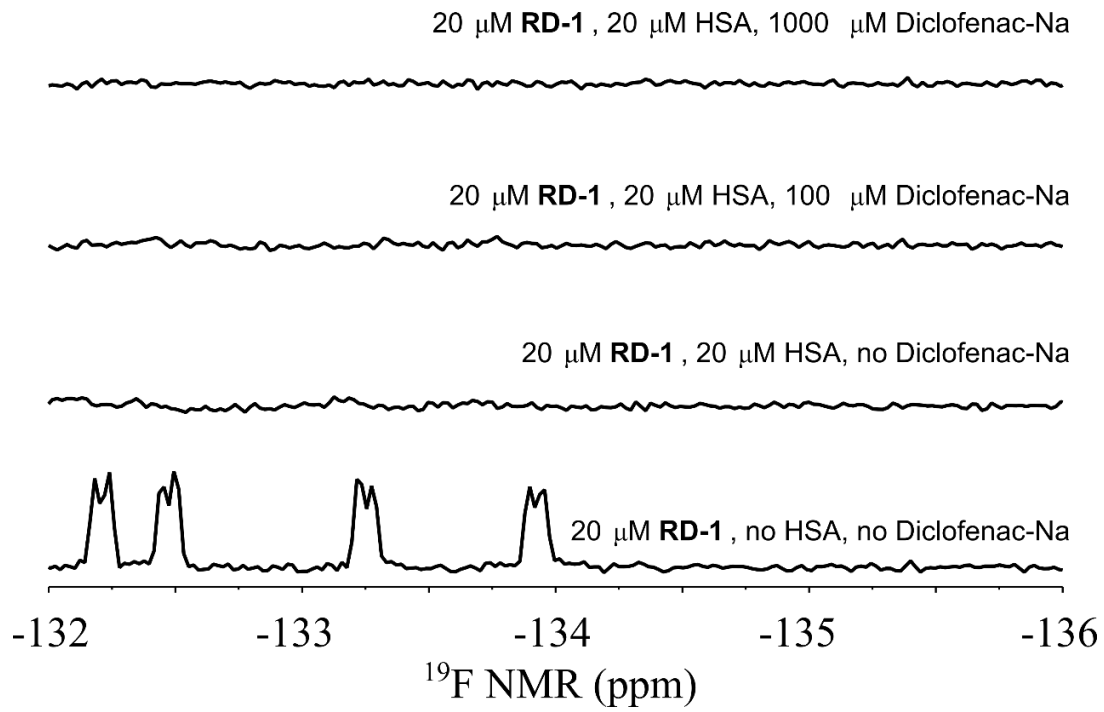


Figure 5-9: ^{19}F NMR competition experiment with **RD-1**, HSA, and diclofenac-Na.

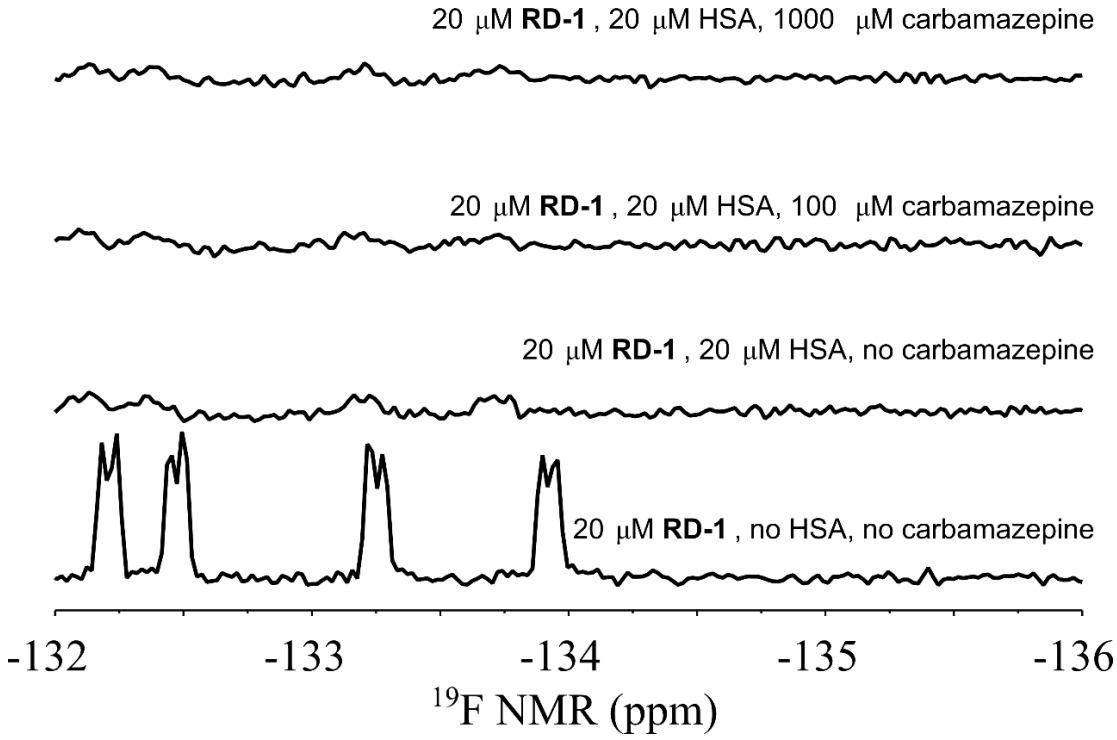


Figure 5-10: ^{19}F NMR competition experiment with **RD-1**, HSA, and carbamazepine.

5.4 Comparison between HSA and rat serum albumin (RSA)

In our NMR binding assays, human serum albumin was used to determine the efficacy of the PFS-peptides. A future experiment planned by Derda and co-workers will be in vivo studies in rats to observe increased serum half-life. As such, it is important to know if rat serum albumin (RSA) will be a reasonable analog to HSA. Two peptides, **RD-1** and **RD-5**, were tested at equimolar concentrations to both RSA and HSA to observe their ^{19}F NMR spectra (Figures 5-11 and 5-12). These two peptides were chosen as they were determined to be the strongest (**RD-1**) and weakest (**RD-5**) binders, respectively. Both peptides appeared to give similar data between RSA and HSA, indicating that in vivo data in rats should be obtained.

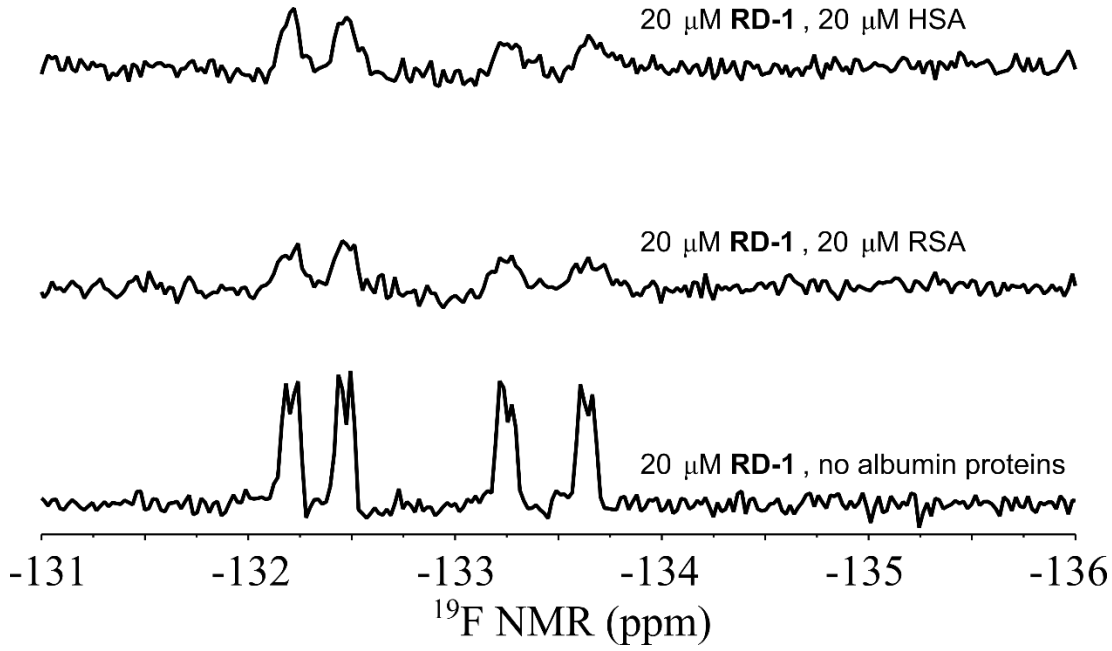


Figure 5-11: ^{19}F NMR comparison of **RD-1** with RSA and HSA.

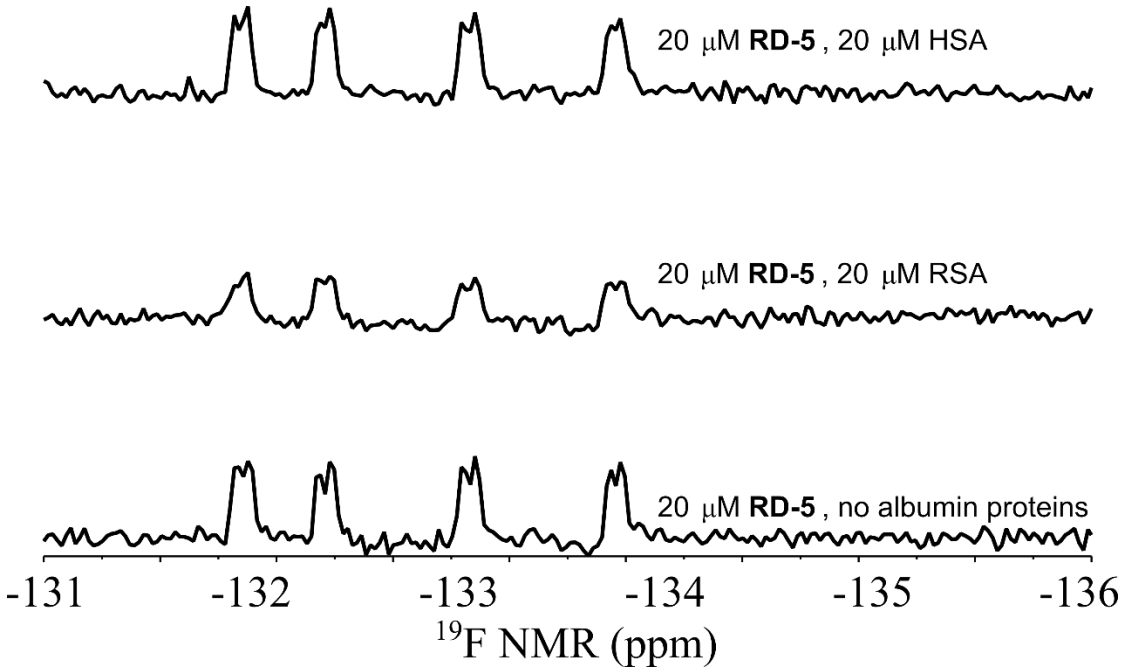


Figure 5-12: ^{19}F NMR comparison of **RD-5** with RSA and HSA.

5.5 Conclusions and future outlook

Although a definitive dissociation constant was not determined in our ^{19}F NMR assays, we were able to qualitatively rank-order the five PFS peptides in terms of their ^{19}F resonance broadening in the presence of HSA, indicating a binding event. **RD-1** bound in approximately a stoichiometric fashion, indicating that its K_d likely lies at least an order of magnitude below the protein concentration used (i.e., 2 μM or less). A competition experiment was then performed using **RD-1** and three known small molecule inhibitors of HSA, however **RD-1** was not competed off even at high concentrations. The lack of an observed change in the ^{19}F NMR after dosing with competitor indicates that **RD-1** does not share a binding site with any of the drugs used. Preliminary experiments with a secondary assay, fluorescence polarization, have been utilized to determine K_d values for **RD-1** (K_d = 6 +/- 0.5 μM) and **RD-5** (K_d = 900 +/- 700 μM) against HSA. In vivo rat experiments as well as attempts at X-ray co-crystal structures are ongoing.

5.6 Experimental

All NMR experiments were performed on a 470 MHz Bruker Avance III HD with a Prodigy TCI cryoprobe (2100:1 S/N for ^{19}F). Experiments were performed with a fluorinated peptide concentration of 20 μM, with varying concentration of human serum albumin (HSA). Parameters used for each experiment are as follows: 750 scans, acquisition time of 0.05 s, relaxation delay of 0.7 s, spectrum centered at -135 ppm with a spectral window of 20 ppm. A reference spectrum observing 2,2,2-trifluoroethanol (TFE) was acquired for each sample with the following parameters: 16 scans, acquisition time of 0.5 s, relaxation delay of 1 s, spectrum centered at -75 with a spectral window of 10 ppm. The observed chemical shift of the TFE reference was deducted from -77.75 ppm and this difference was applied to the peptide spectrum.

A typical experiment involving a titration of HSA was setup as displayed in the table below:

Table 5-1: Components included in ^{19}F NMR assays

Component	**HSA**	**PFS-peptide**	**TFE**	**D_2O**	**PBS**
Stock (μM)	**2000**	**200**	**26.5**	**x**	**x**
	0	50	2	25	423
	2.5	50	2	25	420.5
	5	50	2	25	418
	10	50	2	25	413
	20	50	2	25	403
	40	50	2	25	383

This series would produce a titration with 20 μM PFS-peptide held constant, and HSA varying in a two-fold fashion from 0, 10, 20, 40, 80, 160 μM. TFE refers to 2,2,2-trifluoroethanol as a 1/1000 dilution (approx. 26.5 μM), and PBS components were 50 mM phosphate, 100 mM NaCl, pH 7.4. Nonbolded values indicate volumes of stock used in μL.

In competition experiments, these values were supplemented with 100 μM or 1000 μM of small molecule drug, with the volume needed being deducted from PBS volume.

References

(1) Holland, G. .; Bottomley, P. .; Hinshaw, W. . 19F Magnetic Resonance Imaging. *J. Magn. Reson.* **1977**, *28* (1), 133–136.

(2) Schlemmer, H.; Becker, M.; Bachert, P.; Dietz, A.; Rudat, V.; Vanselow, B.; Wollensack, P.; Zuna, I.; Knopp, M. V; Weidauer, H.; Wannenmacher, M.; Kaick, G. Van. Alterations of Intratumoral Pharmacokinetics of 5-Fluorouracil in Head and Neck Carcinoma during Simultaneous Radiochemotherapy Alterations of Intratumoral Pharmacokinetics of 5-Fluorouracil in Head and Neck Carcinoma during Simultaneous Radiochemotherapy. *Cancer Res* **1999**, *59*, 2363–2369.

(3) Ruiz-Cabello, J.; Barnett, B. P.; Bottomley, P. A.; Bulte, J. W. M. Fluorine (19F) MRS and MRI in Biomedicine. *NMR Biomed.* **2011**, *24* (2), 114–129.

(4) Schubiger, P. A.; Lehmann, L.; Friebe, M. *PET Chemistry: The Driving Force in Molecular Imaging*; Springer: New York, 2007.

(5) Phelps, M. E. Positron Emission Tomography Provides Molecular Imaging of Biological Processes. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (16), 9226–9233.

(6) James, M. L.; Gambhir, S. S. A Molecular Imaging Primer: Modalities, Imaging Agents, and Applications. *Physiol. Rev.* **2012**, *92* (2), 897–965.

(7) Medarova, Z.; Moore, A. Imaging of Pancreatic Cancer : A Promise for Early Diagnosis through Targeted Strategies. *Cancer Ther.* **2004**, *2*, 329–344.

(8) Lustig, M.; Donoho, D.; Pauly, J. M. Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging. *Magn. Reson. Med.* **2007**, *58* (6), 1182–1195.

(9) Yu, J.-X.; Hallac, R. R.; Chiguru, S.; Mason, R. P. New Frontiers and Developing Applications in 19 F NMR. *Prog Nucl Magn Reson Spectrosc* **2013**, *70*, 25–49.

(10) Caravan, P.; Ellison, J. J.; McMurry, T. J.; Lauffer, R. B. Gadolinium(III) Chelates as MRI Contrast Agents: Structure, Dynamics, and Applications. *Chem. Rev.* **1999**, *99* (9), 2293–2352.

(11) Rogowska, J.; Olkowska, E.; Ratajczyk, W.; Wolska, L. Gadolinium as a New

Emerging Contaminant of Aquatic Environments. *Environ. Toxicol. Chem.* **2018**, *37* (6), 1523–1534.

(12) Walpole, S. C.; Prieto-Merino, D.; Edwards, P.; Cleland, J.; Stevens, G.; Roberts, I. The Weight of Nations: An Estimation of Adult Human Biomass. *BMC Public Health* **2012**, *12* (1), 1.

(13) Waters, E. A.; Chen, J.; Allen, J. S.; Zhang, H.; Lanza, G. M.; Wickline, S. A. Detection and Quantification of Angiogenesis in Experimental Valve Disease with Integrin-Targeted Nanoparticles and 19-Fluorine MRI/MRS. *J. Cardiovasc. Magn. Reson.* **2008**, *10* (1), 1–9.

(14) Tirotta, I.; Dichiarante, V.; Pigliacelli, C.; Cavallo, G.; Terraneo, G.; Bombelli, F. B.; Metrangolo, P.; Resnati, G. 19 F Magnetic Resonance Imaging (MRI): From Design of Materials to Clinical Applications. *Chem. Rev.* **2015**, *115* (2), 1106–1129.

(15) Krafft, M. P. Fluorocarbons and Fluorinated Amphiphiles in Drug Delivery and Biomedical Research. *Adv. Drug Deliv. Rev.* **2001**, *47* (2–3), 209–228.

(16) Smart, B. E. *The Chemistry of Functional Groups*; Patai, S., Rappoport, Z., Eds.; Wiley: Chichester, New York, 1983.

(17) Riess, J. G. Perfluorochemical Emulsions for Intravascular Use: Principles, Materials and Methods. In *Fluorine in Medicine in the 21st Century*; University of Manchester, 1994; pp 1–9.

(18) Dunitz, J. D. Organic Fluorine: Odd Man Out. *ChemBioChem* **2004**, *5* (5), 614–621.

(19) Riess, J. Reaassessment of Criteria for the Selection of Perfluorochemicals for Second-Generation Blood Substitutes: Analysis of Structure/Property Relationships. *Artif Organs* **1984**, *8*, 44–56.

(20) *US EPA (U.S. Environmental Protection Agency). Drinking Water Health Advisory for Perfluorooctanoic Acid (PFOA). Available Online:*; 2016.

(21) Kent, K. M.; Cleman, M. W.; Cowley, M. J.; Forman, M. B.; Jaffe, C. C.; Kaplan, M.; King III, S. B.; Krucoff, M. W.; Lassar, T.; McAuley, B. Reduction of Myocardial Ischemia during Percutaneous Transluminal Coronary Angioplasty with Oxygenated Fluosol. *Am J Cardiol* **1990**, *66* (3), 279–284.

(22) Jahr, J. S.; Walker, V.; Manoochehri, K. Blood Substitutes as Pharmacotherapies in Clinical Practice. *Curr. Opin. Anaesthesiol.* **2007**, *20* (4), 325–330.

(23) Nosé, Y. Is There a Role for Blood Substitutes in Civilian Medicine: A Drug for Emergency Shock Cases? *Artif. Organs* **2004**, *28* (9), 807–812.

(24) Wolters, M.; Mohades, S. G.; Hackeng, T. M.; Post, M. J.; Kooi, M. E.; Backes, W. H. Clinical Perspectives of Hybrid Proton-Fluorine Magnetic Resonance Imaging and Spectroscopy. *Invest. Radiol.* **2013**, *48* (5), 341–350.

(25) Jiang, Z.-X.; Liu, X.; Jeong, E.-K.; Yu, Y. B. Symmetry-Guided Design and Fluorous Synthesis of a Stable and Rapidly Excreted Imaging Tracer for (19)F MRI. *Angew. Chem. Int. Ed. Engl.* **2009**, *48* (26), 4755–4758.

(26) Fu, C.; Zhang, C.; Peng, H.; Han, F.; Baker, C.; Wu, Y.; Ta, H.; Whittaker, A. K. Enhanced Performance of Polymeric19F MRI Contrast Agents through Incorporation of Highly Water-Soluble Monomer MSEA. *Macromolecules* **2018**, *51* (15), 5875–5882.

(27) Kirberger, S. E.; Maltseva, S. D.; Manulik, J. C.; Einstein, S. A.; Weegman, B. P.; Garwood, M.; Pomerantz, W. C. K. Synthesis of Intrinsically Disordered Fluorinated Peptides for Modular Design of High-Signal 19F MRI Agents. *Angew. Chemie - Int. Ed.* **2017**, *56* (23), 6440–6444.

(28) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149–2154.

(29) Frieden, C.; Hoeltzli, S. D.; Ropson, I. R. A. J. NMR and Protein Folding : Equilibrium and Stopped-Flow Studies. *Protein Sci.* **1993**, No. 2, 2007–2014.

(30) Dyson, H. J.; Wright, P. E. Unfolded Proteins and Protein Folding Studied by NMR. *Chem. Rev.* **2004**, *104* (8), 3607–3622.

(31) Tompa, P.; Fersht, A. *Structure and Function of Intrinsically Disordered Proteins*; 2010.

(32) Li, H.; Frieden, C. Phenylalanine Side Chain Behavior of the Intestinal Fatty Acid-Binding Protein: The Effect of Urea on Backbone and Side Chain Stability. *J. Biol. Chem.* **2005**, *280* (46), 38556–38561.

(33) Sacchettini, J. C.; Gordon, J. I.; Banaszak, L. J. Refined Apoprotein Structure of Rat Intestinal Fatty Acid Binding Protein Produced in Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (20), 7736–7740.

(34) Mayzel, M.; Saline, M.; Pedersen, A.; Rosenlo, J.; Orekhov, V. Y. Highly Efficient NMR Assignment of Intrinsically Disordered Proteins : Application to B- and T Cell Receptor Domains. *PLoS One* **2013**, *8* (5), 1–8.

(35) Kosol, S.; Contreras-martos, S.; Cedeño, C.; Tompa, P. Structural Characterization of Intrinsically Disordered Proteins by NMR Spectroscopy. *Molecules* **2013**, *18*, 10802–10828.

(36) Jensen, M. R.; Ruigrok, R. W. H.; Blackledge, M. Describing Intrinsically Disordered Proteins at Atomic Resolution by NMR. *Curr. Opin. Struct. Biol.* **2013**, *23* (3), 426–435.

(37) Kwan, A. H.; Mobli, M.; Gooley, P. R.; King, G. F.; Mackay, J. P. Macromolecular NMR Spectroscopy for the Non-Spectroscopist. *FEBS J.* **2011**, *278*, 687–703.

(38) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 5th ed.; Anderson, M., Granum, S., Eds.; Garland Science Group: New York, 2008.

(39) Nicholson, R. I.; Gee, J. M. W.; Harper, M. E. EGFR and Cancer Prognosis. *Eur. J. Cancer* **2001**, *37*, 9–15.

(40) Chalhoub, N.; Baker, S. J. PTEN and the PI3-Kinase Pathway in Cancer. *Annu. Rev. Pathol.* **2009**, *4*, 127–150.

(41) Roberts, P. J.; Der, C. J. Targeting the Raf-MEK-ERK Mitogen-Activated Protein Kinase Cascade for the Treatment of Cancer. *Oncogene* **2007**, *26* (22), 3291–3310.

(42) Normanno, N.; De Luca, A.; Bianco, C.; Strizzi, L.; Mancino, M.; Maiello, M. R.; Carotenuto, A.; De Feo, G.; Caponigro, F.; Salomon, D. S. Epidermal Growth Factor Receptor (EGFR) Signaling in Cancer. *Gene* **2006**, *366* (1), 2–16.

(43) Li, Z. Identification and Characterization of a Novel Peptide Ligand of Epidermal Growth Factor Receptor for Targeted Delivery of Therapeutics. *FASEB J.* **2005**, *19* (14), 1978–1985.

(44) Song, S.; Liu, D.; Peng, J.; Deng, H.; Guo, Y.; Xu, L. X.; Miller, A. D.; Xu, Y. Novel Peptide Ligand Directs Liposomes toward EGF-R High-Expressing Cancer Cells in Vitro and in Vivo. *FASEB J.* **2009**, *23* (5), 1396–1404.

(45) Shah, R.; Petersburg, J.; Gangar, A. C.; Fegan, A.; Wagner, C. R.; Kumarapperuma, S. C. In Vivo Evaluation of Site-Specifically PEGylated Chemically Self-Assembled Protein Nanostructures. *Mol. Pharm.* **2016**, *13* (7), 2193–2203.

(46) Viehweger, K.; Barbaro, L.; García, K. P.; Joshi, T.; Geipel, G.; Steinbach, J.; Stephan, H.; Spiccia, L.; Graham, B. EGF Receptor-Targeting Peptide Conjugate Incorporating a near-IR Fluorescent Dye and a Novel 1,4,7-Triazacyclononane-Based ^{64}Cu(II) Chelator Assembled via Click Chemistry. *Bioconjug. Chem.* **2014**, *25* (5), 1011–1022.

(47) Kohno, M.; Horibe, T.; Haramoto, M.; Yano, Y.; Ohara, K.; Nakajima, O.; Matsuzaki, K.; Kawakami, K. A Novel Hybrid Peptide Targeting EGFR-Expressing Cancers. *Eur. J. Cancer* **2011**, *47* (5), 773–783.

(48) Hackel, B. J.; Kimura, R. H.; Gambhir, S. S. Use of 64Cu-Labeled Fibronectin Domain with EGFR-Overexpressing Tumor Xenograft. *Radiology* **2012**, *263* (1), 179–188.

(49) Hackel, B. J.; Ackerman, M. E.; Howland, S. W.; Wittrup, K. D. Stability and CDR Composition Biases Enrich Binder Functionality Landscapes. *J. Mol. Biol.*

2010, *401* (1), 84–96.

(50) Ramamurthy, V.; Krystek, S. R.; Bush, A.; Wei, A.; Emanuel, S. L.; Das Gupta, R.; Janjua, A.; Cheng, L.; Murdock, M.; Abramczyk, B.; Cohen, D.; Lin, Z.; Morin, P.; Davis, J. H.; Dabritz, M.; McLaughlin, D. C.; Russo, K. A.; Chao, G.; Wright, M. C.; et al. Structures of Adnectin/Protein Complexes Reveal an Expanded Binding Footprint. *Structure* **2012**, *20* (2), 259–269.

(51) Lipovsek, D. Adnectins: Engineered Target-Binding Protein Therapeutics. *Protein Eng. Des. Sel.* **2011**, *24* (1–2), 3–9.

(52) Waddington, C. H. The Epigenotype. *Int. J. Epidemiol.* **2012**, *41* (1), 10–13.

(53) Berger, S. L.; Kouzarides, T.; Shiekhattar, R.; Shilatifard, A. An Operational Definition of Epigenetics. *Genes Dev.* **2009**, *23* (7), 781–783.

(54) Luger, Karolin; Mäder, Armin W.; Richmond, Robin K.; Sargent, David F.; Richmond, T. J. Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution. *Nature* **1997**, *389*, 251–260.

(55) Bhasin, M.; Reinherz, E. L.; Reche, P. A. Recognition and Classification of Histones Using Support Vector Machine. *J. Comput. Biol.* **2006**, *13* (1), 102–112.

(56) Passarge, E. Emil Heitz and the Concept of Heterochromatin: Longitudinal Chromosome Differentiation Was Recognized Fifty Years Ago. *Am J Hum Genet* **1979**, *31* (2), 106–115.

(57) Nicholson, T. B.; Veland, N.; Chen, T. *Writers, Readers, and Erasers of Epigenetic Marks*; Elsevier Inc., 2015.

(58) Cao, R.; Wang, L.; Wang, H.; Xia, L.; Erdjument-Bromage, H.; Tempst, P.; Jones, R.S.; Zhang, Y. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science (80-.).* **2003**, *300* (5616), 131–135.

(59) Zhang, B.; Day, D. S.; Ho, J. W.; Song, L.; Cao, J.; Christodoulou, D.; Seidman, J. G.; Crawford, G. E.; Park, P. J.; Pu, W. T. A Dynamic H3K27ac Signature Identifies VEGFA-Stimulated Endothelial Enhancers and Requires EP300

Activity. *Genome Res.* **2013**, *23* (6), 917–927.

(60) Sedkov, Y.; Cho, E.; Petruk, S.; Cherbas, L.; Smith, S. T.; Jones, R. S.; Cherbas, P.; Canaani, E.; Jaynes, J. B.; Mazo, A. Methylation at Lysine 4 of Histone H3 in Ecdysone-Dependent Development of Drosophila. *Nature* **2003**, *426* (6962), 78–83.

(61) Jenuwein, T.; Allis, C. D. Translating the Histone Code. *Science (80-.).* **2001**, *293* (5532), 1074–1081.

(62) Smith, S. G.; Zhou, M.-M. The Bromodomain as an Acetyl-Lysine Reader Domain. In *Chromatin Signaling and Diseases*; Elsevier, 2016; pp 97–111.

(63) Kamps, J. J. A. G.; Huang, J.; Poater, J.; Xu, C.; Pieters, B. J. G. E.; Dong, A.; Min, J.; Sherman, W.; Beuming, T.; Matthias Bickelhaupt, F.; Li, H.; Mecinović, J. Chemical Basis for the Recognition of Trimethyllysine by Epigenetic Reader Proteins. *Nat. Commun.* **2015**, *6* (1), 8911.

(64) Muslin, A. J.; Tanner, J. W.; Allen, P. M.; Shaw, A. S. Interaction of 14-3-3 with Signaling Proteins Is Mediated by the Recognition of Phosphoserine. *Cell* **1996**, *84* (6), 889–897.

(65) Ruthenburg, A. J.; Li, H.; Milne, T. A.; Dewell, S.; McGinty, R. K.; Yuen, M.; Ueberheide, B.; Dou, Y.; Muir, T. W.; Patel, D. J.; Allis, C. D. Recognition of a Mononucleosomal Histone Modification Pattern by BPTF via Multivalent Interactions. *Cell* **2011**, *145* (5), 692–706.

(66) Belkina, A. C.; Denis, G. V. BET Domain Co-Regulators in Obesity, Inflammation and Cancer. *Nat. Rev. Cancer* **2012**, *12* (7), 465–477.

(67) Filippakopoulos, P.; Picaud, S.; Mangos, M.; Keates, T.; Lambert, J. P.; Barsyte-Lovejoy, D.; Felletar, I.; Volkmer, R.; Müller, S.; Pawson, T.; Gingras, A. C.; Arrowsmith, C. H.; Knapp, S. Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell* **2012**, *149* (1), 214–231.

(68) Dhalluin, C.; Carlson, J. E.; Zeng, L.; He, C.; Aggarwal, A. K.; Zhou, M. Structure and Ligand of a Histone Acetyltransferase Bromodomain. *Nature* **1999**, *399*, 491–

496.

(69) Hewings, D. S.; Rooney, T. P. C.; Conway, S. J. Emerging Epigenetic Therapies—Bromodomain Ligands. In *Epigenetic Cancer Therapy*; Elsevier, 2015; pp 495–524.

(70) Ntranos, A.; Casaccia, P. Bromodomains: Translating the Words of Lysine Acetylation into Myelin Injury and Repair. *Neurosci. Lett.* **2016**, *625*, 4–10.

(71) Sanchez, R.; Zhou, M. M. The Role of Human Bromodomains in Chromatin Biology and Gene Transcription. *Curr. Opin. Drug Discov. Devel.* **2009**, *12* (5), 659–665.

(72) Padmanabhan, B.; Mathur, S.; Manjula, R.; Tripathi, S. Bromodomain and Extra-Terminal (BET) Family Proteins: New Therapeutic Targets in Major Diseases. *J. Biosci.* **2016**, *41* (2), 295–311.

(73) Nicodeme, E.; Jeffrey, K. L.; Schaefer, U.; Beinke, S.; Dewell, S.; Chung, C. W.; Chandwani, R.; Marazzi, I.; Wilson, P.; Coste, H.; White, J.; Kirilovsky, J.; Rice, C. M.; Lora, J. M.; Prinjha, R. K.; Lee, K.; Tarakhovsky, A. Suppression of Inflammation by a Synthetic Histone Mimic. *Nature* **2010**, *468* (7327), 1119–1123.

(74) Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; et al. Selective Inhibition of BET Bromodomains. *Nature* **2010**, *468* (7327), 1067–1073.

(75) Picaud, S.; Leonards, K.; Lambert, J. P.; Dovey, O.; Wells, C.; Fedorov, O.; Monteiro, O.; Fujisawa, T.; Wang, C. Y.; Lingard, H.; Tallant, C.; Nikbin, N.; Guetzoyan, L.; Ingham, R.; Ley, S. V.; Brennan, P.; Muller, S.; Samsonova, A.; Gingras, A. C.; et al. Promiscuous Targeting of Bromodomains by Bromosporine Identifies BET Proteins as Master Regulators of Primary Transcription Response in Leukemia. *Sci. Adv.* **2016**, *2* (10), 15–20.

(76) Chung, C. W.; Dean, A. W.; Woolven, J. M.; Bamborough, P. Fragment-Based Discovery of Bromodomain Inhibitors Part 1: Inhibitor Binding Modes and Implications for Lead Discovery. *J. Med. Chem.* **2012**, *55* (2), 576–586.

(77) Meslamani, Jamel; Smith, Steven G.; Sanchez, Roberto; Zhou, M.-M. Structural Features and Inhibitors of Bromodomains. *Drug Discov Today Technol.* **2016**, *19*, 3–15.

(78) Muller, S.; Filippakopoulos, P.; Knapp, S. Bromodomains as Therapeutic Targets. *Expert Rev. Mol. Med.* **2011**, *13* (September), e29.

(79) Hewings, D. S.; Rooney, T. P. C.; Jennings, L. E.; Hay, D. A.; Schofield, C. J.; Brennan, P. E.; Knapp, S.; Conway, S. J. Progress in the Development and Application of Small Molecule Inhibitors of Bromodomain-Acetyl-Lysine Interactions. *J. Med. Chem.* **2012**, *55* (22), 9393–9413.

(80) Vidler, L. R.; Brown, N.; Knapp, S.; Hoelder, S. Druggability Analysis and Structural Classification of Bromodomain Acetyl-Lysine Binding Sites. *J. Med. Chem.* **2012**, *55* (17), 7346–7359.

(81) Wysocka, J.; Swigut, T.; Xiao, H.; Milne, T. A.; Kwon, S. Y.; Landry, J.; Kauer, M.; Tackett, A. J.; Chait, B. T.; Badenhorst, P.; Wu, C.; Allis, C. D. A PHD Finger of NURF Couples Histone H3 Lysine 4 Trimethylation with Chromatin Remodelling. *Nature* **2006**, *442* (7098), 86–90.

(82) Goller, T.; Vauti, F.; Ramasamy, S.; Arnold, H.-H. Transcriptional Regulator BPTF/FAC1 Is Essential for Trophoblast Differentiation during Early Mouse Development. *Mol. Cell. Biol.* **2008**, *28* (22), 6819–6827.

(83) Xiao, S.; Liu, L.; Fang, M.; Zhou, X.; Peng, X.; Long, J.; Lu, X. BPTF Associated with EMT Indicates Negative Prognosis in Patients with Hepatocellular Carcinoma. *Dig Dis Sci* **2015**, *60* (4), 910–918.

(84) Richart, L.; Carrillo-de Santa Pau, E.; Rio-Machin, A.; de Andres, M. P.; Cigudosa, J. C.; Lobo, V. J.; Real, F. X. BPTF Is Required for C-MYC Transcriptional Activity and in Vivo Tumorigenesis. *Nat Commun* **2016**, *7*, 10153.

(85) Gong, Y.; Liu, D. BPTF Biomarker Correlates with Poor Survival in Human NSCLC. **2017**, 102–107.

(86) Dai, M.; Lu, J.-J.; Guo, W.; Yu, W.; Wang, Q.; Tang, R.; Tang, Z.; Xiao, Y.; Li, Z.; Sun, W.; Sun, X.; Qin, Y.; Huang, W.; Deng, W.; Wu, T. BPTF Promotes Tumor Growth and Predicts Poor Prognosis in Lung Adenocarcinomas. *Oncotarget* **2015**, *6* (32), 33878–33892.

(87) Buganim, Y.; Goldstein, I.; Lipson, D.; Milyavsky, M.; Polak-Charcon, S.; Mardoukh, C.; Solomon, H.; Kalo, E.; Madar, S.; Brosh, R.; Perelman, M.; Navon, R.; Goldfinger, N.; Barshack, I.; Yakhini, Z.; Rotter, V. A Novel Translocation Breakpoint within the BPTF Gene Is Associated with a Pre-Malignant Phenotype. *PLoS One* **2010**, *5* (3).

(88) Stankiewicz, P.; Khan, T. N.; Szafranski, P.; Slattery, L.; Streff, H.; Vetrini, F.; Bernstein, J. A.; Brown, C. W.; Rosenfeld, J. A.; Rednam, S.; Scollon, S.; Bergstrom, K. L.; Parsons, D. W.; Plon, S. E.; Vieira, M. W.; Quaio, C. R. D. C.; Baratela, W. A. R.; Acosta Guio, J. C.; Armstrong, R.; et al. Haploinsufficiency of the Chromatin Remodeler BPTF Causes Syndromic Developmental and Speech Delay, Postnatal Microcephaly, and Dysmorphic Features. *Am. J. Hum. Genet.* **2017**, *101* (4), 503–515.

(89) Lee, J. H.; Kim, M. S.; Yoo, N. J.; Lee, S. H. BPTF, a Chromatin Remodeling-Related Gene, Exhibits Frameshift Mutations in Gastric and Colorectal Cancers. *Apmis* **2016**, *124* (5), 425–427.

(90) Balbás-Martínez, C.; Sagrera, A.; Carrillo-de-Santa-Pau, E.; Earl, J.; Márquez, M.; Vazquez, M.; Lapi, E.; Castro-Giner, F.; Beltran, S.; Bayés, M.; Carrato, A.; Cigudosa, J. C.; Domínguez, O.; Gut, M.; Herranz, J.; Juanpere, N.; Kogevinas, M.; Langa, X.; López-Knowles, E.; et al. Recurrent Inactivation of STAG2 in Bladder Cancer Is Not Associated with Aneuploidy. *Nat. Genet.* **2013**, *45* (12), 1464–1469.

(91) Humphreys, P. G.; Bamborough, P.; Chung, C. W.; Craggs, P. D.; Gordon, L.; Grandi, P.; Hayhow, T. G.; Hussain, J.; Jones, K. L.; Lindon, M.; Michon, A. M.;

Renaux, J. F.; Suckling, C. J.; Tough, D. F.; Prinjha, R. K. Discovery of a Potent, Cell Penetrant, and Selective P300/CBP-Associated Factor (PCAF)/General Control Nonderepressible 5 (GCN5) Bromodomain Chemical Probe. *J. Med. Chem.* **2017**, *60* (2), 695–709.

(92) Moustakim, M.; Clark, P. G. K.; Trulli, L.; Fuentes de Arriba, A. L.; Ehebauer, M. T.; Chaikuad, A.; Murphy, E. J.; Mendez-Johnson, J.; Daniels, D.; Hou, C. F. D.; Lin, Y. H.; Walker, J. R.; Hui, R.; Yang, H.; Dorrell, L.; Rogers, C. M.; Monteiro, O. P.; Fedorov, O.; Huber, K. V. M.; et al. Discovery of a PCAF Bromodomain Chemical Probe. *Angew. Chemie - Int. Ed.* **2017**, *56* (3), 827–831.

(93) Crawford, T. D.; Audia, J. E.; Bellon, S.; Burdick, D. J.; Bommi-Reddy, A.; Côté, A.; Cummings, R. T.; Duplessis, M.; Flynn, E. M.; Hewitt, M.; Huang, H. R.; Jayaram, H.; Jiang, Y.; Joshi, S.; Kiefer, J. R.; Murray, J.; Nasveschuk, C. G.; Neiss, A.; Pardo, E.; et al. GNE-886: A Potent and Selective Inhibitor of the Cat Eye Syndrome Chromosome Region Candidate 2 Bromodomain (CECR2). *ACS Med. Chem. Lett.* **2017**, *8* (7), 737–741.

(94) Frey, W. D.; Chaudhry, A.; Slepicka, P. F.; Ouellette, A. M.; Kirberger, S. E.; Pomerantz, W. C. K.; Hannon, G. J.; dos Santos, C. O. BPTF Maintains Chromatin Accessibility and the Self-Renewal Capacity of Mammary Gland Stem Cells. *Stem Cell Reports* **2017**, *9* (1), 23–31.

(95) Mayes, K.; Elsayed, Z.; Alhazmi, A.; Waters, M.; Alkhatib, S. G.; Roberts, M.; Song, C.; Peterson, K.; Chan, V.; Ailaney, N.; Malapati, P.; Blevins, T.; Lisnić, B.; Dumur, C. I.; Landry, J. W. BPTF Inhibits NK Cell Activity and the Abundance of Natural Cytotoxicity Receptor Co-Ligands. *Oncotarget* **2017**, *8* (38), 64344–64357.

(96) Mishra, N. K.; Urick, A. K.; Ember, S. W. J.; Scho, E.; Pomerantz, W. C. Fluorinated Aromatic Amino Acids Are Sensitive 19F NMR Probes for Bromodomain-Ligand Interactions. *ACS Chem. Biol.* **2014**, *9* (12), 2755–2760.

(97) Urick, A. K.; Hawk, L. M. L.; Cassel, M. K.; Mishra, N. K.; Liu, S.; Adhikari, N.; Zhang, W.; dos Santos, C. O.; Hall, J. L.; Pomerantz, W. C. K. Dual Screening of

BPTF and Brd4 Using Protein-Observed Fluorine NMR Uncovers New Bromodomain Probe Molecules. *ACS Chem. Biol.* **2015**, *10* (10), 2246–2256.

(98) Danielson, M. A.; Falke, J. J. Use of 19F NMR to Probe Protein Structure and Conformational Changes. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25* (1), 163–195.

(99) Gee, C. T.; Arntson, K. E.; Urick, A. K.; Mishra, N. K.; Hawk, L. M. L.; Wisniewski, A. J.; Pomerantz, W. C. K. Protein-Observed 19F-NMR for Fragment Screening, Affinity Quantification and Druggability Assessment. *Nat. Protoc.* **2016**, *11* (8), 1414–1427.

(100) Crowley, P. B.; Kyne, C.; Monteith, W. B. Simple and Inexpensive Incorporation of 19F-Tryptophan for Protein NMR Spectroscopy. *Chem. Commun.* **2012**, *48* (86), 10681–10683.

(101) Bogan, A. A.; Thorn, K. S. Anatomy of Hot Spots in Protein Interfaces. *J. Mol. Biol.* **1998**, *280* (1), 1–9.

(102) Mantina, M.; Chamberlin, A. C.; Valero, R.; Cramer, C. J.; Truhlar, D. G. Consistent van Der Waals Radii for the Whole Main Group. *J. Phys. Chem. A* **2009**, *113* (19), 5806–5812.

(103) Bryant, R. G. The NMR Time Scale. *J. Chem. Educ.* **1983**, *60* (11), 933.

(104) Stumpf, M. P. H.; Thorne, T.; de Silva, E.; Stewart, R.; An, H. J.; Lappe, M.; Wiuf, C. Estimating the Size of the Human Interactome. *Proc. Natl. Acad. Sci.* **2008**, *105* (19), 6959–6964.

(105) Kilburg, D.; Gallicchio, E. *Recent Advances in Computational Models for the Study of Protein–Peptide Interactions*, 1st ed.; Elsevier Inc., 2016; Vol. 105.

(106) Clackson, T.; Wells, J. A. A Hot Spot of Binding Energy in a Hormone-Receptor Interface. *Science (80-.).* **1995**, *267* (5196), 383–386.

(107) Ahrens, E. T.; Zhong, J. In Vivo MRI Cell Tracking Using Perfluorocarbon Probes and Fluorine-19 Detection. *NMR Biomed.* **2013**, *26* (7), 860–871.

(108) Kok, M. B.; De Vries, A.; Abdurrachim, D.; Prompers, J. J.; Grüll, H.; Nicolay, K.; Strijkers, G. J. Quantitative 1H MRI, 19F MRI, and 19F MRS of Cell-Internalized Perfluorocarbon Paramagnetic Nanoparticles. *Contrast Media Mol. Imaging* **2011**, *6* (1), 19–27.

(109) Ahrens, E. T.; Flores, R.; Xu, H.; Morel, P. a. In Vivo Imaging Platform for Tracking Immunotherapeutic Cells. *Nat. Biotechnol.* **2005**, *23* (8), 983–987.

(110) Mignion, L.; Magat, J.; Schakman, O.; Marbaix, E.; Gallez, B.; Jordan, B. F. Hexafluorobenzene in Comparison with Perfluoro-15-Crown-5-Ether for Repeated Monitoring of Oxygenation Using 19F MRI in a Mouse Model. *Magn. Reson. Med.* **2013**, *69* (1), 248–254.

(111) Schlemmer, H. P.; Becker, M.; Bachert, P.; Dietz, A.; Rudat, V.; Vanselow, B.; Wollensack, P.; Zuna, I.; Knopp, M. V.; Weidauer, H.; Wannenmacher, M.; Van Kaick, G. Alterations of Intratumoral Pharmacokinetics of 5-Fluorouracil in Head and Neck Carcinoma during Simultaneous Radiochemotherapy. *Cancer Res.* **1999**, *59* (10), 2363–2369.

(112) Nakamura, T.; Sugihara, F.; Matsushita, H.; Yoshioka, Y.; Mizukami, S.; Kikuchi, K. Mesoporous Silica Nanoparticles for 19 F Magnetic Resonance Imaging, Fluorescence Imaging, and Drug Delivery. *Chem. Sci.* **2015**, *6* (3), 1986–1990.

(113) Janjic, J. M.; Srinivas, M.; Kadayakkara, D. K. K.; Ahrens, E. T. Self-Delivering Nanoemulsions for Dual Fluorine-19 MRI and Fluorescence Detection. *J. Am. Chem. Soc.* **2008**, *130* (9), 2832–2841.

(114) Yu, W.; Yang, Y.; Bo, S.; Li, Y.; Chen, S.; Yang, Z.; Zheng, X.; Jiang, Z.-X.; Zhou, X. Design and Synthesis of Fluorinated Dendrimers for Sensitive 19F MRI. *J. Org. Chem.* **2015**, *80* (9), 4443–4449.

(115) Pfannkuch, F.; Schnoy, N. [Perfluorochemical-Emulsion (Fluorocar- Bon-43) as Blood Gas Carrier – Fate of the Substance after Intravenous Injection in Rat (Author's Transl)]. *Anaesthesist* **1979**, *28*, 511–516.

(116) Nicole, W. PFOA and Cancer in a Highly Exposed Community: New Findings

from the C8 Science Panel. *Environ. Health Perspect.* **2013**, *121* (11–12), A340–A340.

(117) Mountain, G. A.; Jelier, B. J.; Bagia, C.; Friesen, C. M.; Janjic, J. M. Design and Formulation of Nanoemulsions Using 2-(Poly(Hexafluoropropylene Oxide)) Perfluoropropyl Benzene in Combination with Linear Perfluoro(Polyethylene Glycol Dimethyl Ether). *J. Fluor. Chem.* **2014**, *162* (2), 38–44.

(118) Tirotta, I.; Mastropietro, A.; Cordiglieri, C.; Gazzera, L.; Baggi, F.; Baselli, G.; Grazia Bruzzone, M.; Zucca, I.; Cavallo, G.; Terraneo, G.; Baldelli Bombelli, F.; Metrangolo, P.; Resnati, G. A Superfluorinated Molecular Probe for Highly Sensitive in Vivo ^{19}F-MRI. *J. Am. Chem. Soc.* **2014**, *136* (24), 8524–8527.

(119) Mishra, N. K.; Urick, A. K.; Ember, S. W. J.; Schönbrunn, E.; Pomerantz, W. C. Fluorinated Aromatic Amino Acids Are Sensitive 19 F NMR Probes for Bromodomain-Ligand Interactions. *ACS Chem. Biol.* **2014**, *9* (12), 2755–2760.

(120) Kitevski-LeBlanc, J. L.; Prosser, R. S. Current Applications of 19F NMR to Studies of Protein Structure and Dynamics. *Prog. Nucl. Magn. Reson. Spectrosc.* **2012**, *62*, 1–33.

(121) Wang, G. F.; Li, C.; Pielak, G. J. 19F NMR Studies of ??-Synuclein-Membrane Interactions. *Protein Sci.* **2010**, *19* (9), 1686–1691.

(122) Ye, L.; Larda, S. T.; Frank Li, Y. F.; Manglik, A.; Prosser, R. S. A Comparison of Chemical Shift Sensitivity of Trifluoromethyl Tags: Optimizing Resolution in 19F NMR Studies of Proteins. *J. Biomol. NMR* **2015**, *62* (1), 97–103.

(123) Grage, S. L.; Dürr, U. H. N.; Afonin, S.; Mikhailiuk, P. K.; Komarov, I. V.; Ulrich, A. S. Solid State 19F NMR Parameters of Fluorine-Labeled Amino Acids. Part II: Aliphatic Substituents. *J. Magn. Reson.* **2008**, *191* (1), 16–23.

(124) Dürr, U. H. N.; Grage, S. L.; Witter, R.; Ulrich, A. S. Solid State 19F NMR Parameters of Fluorine-Labeled Amino Acids. Part I: Aromatic Substituents. *J. Magn. Reson.* **2008**, *191* (1), 7–15.

(125) Rucker, A. L.; Creamer, T. P. Polyproline II Helical Structure in Protein Unfolded

States: Lysine Peptides Revisited. *Protein Sci.* **2002**, *11* (4), 994–994.

(126) Serrano, L. Comparison between the Phi Distribution of the Amino Acids in the Protein Database and NMR Data Indicates That Amino Acids Have Various Phi Propensities in the Random Coil Conformation. *J. Mol. Biol.* **1995**, *254*, 322–333.

(127) Shi, Z.; Chen, K.; Liu, Z.; Ng, A.; Bracken, W. C.; Kallenbach, N. R. Polyproline II Propensities from GGXGG Peptides Reveal an Anticorrelation with Beta-Sheet Scales. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (50), 17964–17968.

(128) Yu, Y. B. Fluorinated Dendrimers as Imaging Agents for 19F MRI. *Wiley Interdiscip. Rev. Nanomedicine Nanobiotechnology* **2013**, *5* (6), 646–661.

(129) Tsuyuki, E.; Tsuyuki, H.; Stahmann, M. A. The Synthesis and Enzymatic Hyrolysis of Poly-d-Lysine. *J. Biol. Chem* **1956**, *222*, 479–485.

(130) Li, C.; Lutz, E. A.; Slade, K. M.; Ruf, R. A. S.; Wang, G. F.; Pielak, G. J. 19F NMR Studies of ??-Synuclein Conformation and Fibrillation. *Biochemistry* **2009**, *48* (36), 8578–8584.

(131) Antaris, A. L.; Chen, H.; Cheng, K.; Sun, Y.; Hong, G.; Qu, C.; Diao, S.; Deng, Z.; Hu, X.; Zhang, B.; Zhang, X.; Yaghi, O. K.; Alamparambil, Z. R.; Hong, X.; Cheng, Z.; Dai, H. A Small-Molecule Dye for NIR-II Imaging. *Nat. Mater.* **2015**, *15* (2), 235–242.

(132) Hu, G.; Tang, J.; Bai, X.; Xu, S.; Wang, L. Superfluorinated Copper Sulfide Nanoprobes for Simultaneous 19F Magnetic Resonance Imaging and Photothermal Ablation. *Nano Res.* **2016**, *9* (6), 1–9.

(133) Tressler, C. M.; Zondlo, N. J. Synthesis of Perfluoro- Tert -Butyl Tyrosine, for Application in 19 F NMR, via a Diazonium-Coupling Reaction. *Org. Lett.* **2016**, *18* (24), 6240–6243.

(134) Buer, B. C.; Levin, B. J.; Marsh, E. N. G. Perfluoro-Tert-Butyl-Homoserine as a Sensitive 19F NMR Reporter for Peptide-Membrane Interactions in Solution. *J. Pept. Sci.* **2013**, *19* (5), 308–314.

(135) Harvey, P.; Kuprov, I.; Parker, D. Lanthanide Complexes as Paramagnetic Probes for 19F Magnetic Resonance. *Eur. J. Inorg. Chem.* **2012**, No. 12, 2015–2022.

(136) Zhong, J.; Mills, P. H.; Hitchens, T. K.; Ahrens, E. T. Accelerated Fluorine-19 MRI Cell Tracking Using Compressed Sensing. *Magn. Reson. Med.* **2013**, *69* (6), 1683–1690.

(137) European Chemical Agency. 2,2,2-trifluoroacetamide https://echa.europa.eu/registration-dossier/-/registered-dossier/12379/5/3/2.

(138) Boutonnet, J. C.; Bingham, P.; Calamari, D.; Rooij, C. de; Franklin, J.; Kawano, T.; Libre, J.-M.; McCul-loch, A.; Malinverno, G.; Odom, J. M.; Rusch, G. M.; Smythe, K.; Sobolev, I.; Thompson, R.; Tiedje, J. M. Environmental Risk Assessment of Trifluoroacetic Acid. *Hum. Ecol. Risk Assess. An Int. J.* **1999**, *5* (1), 59–124.

(139) Duewel, H.; Daub, E.; Robinson, V.; Honek, J. F. Incorporation of Trifluoromethionine into a Phage Lysozyme: Implications and a New Marker for Use in Protein 19 F NMR †. *Biochemistry* **1997**, *36* (11), 3404–3416.

(140) Cottrell, T. L. *"The Strengths of Chemical Bonds," 2nd Ed.*; 1958.

(141) Suzuki, Y.; Brender, J. R.; Soper, M. T.; Krishnamoorthy, J.; Zhou, Y.; Ruotolo, B. T.; Kotov, N. A.; Ramamoorthy, A.; Marsh, E. N. G. Resolution of Oligomeric Species during the Aggregation of Aβ1-40 Using 19 F NMR. *Biochemistry* **2013**, *52* (11), 1903–1912.

(142) Eisenberger, P.; Gischig, S.; Togni, A. Novel 10-I-3 Hypervalent Iodine-Based Compounds for Electrophilic Trifluoromethylation. *Chem. - A Eur. J.* **2006**, *12* (9), 2579–2586.

(143) Jobron, L.; Hummel, G. Solid-Phase Synthesis of New S -Glycoamino Acid Building Blocks. *Org. Lett.* **2000**, *2* (15), 2265–2267.

(144) Mamot, C.; Drummond, D. C.; Greiser, U.; Cells, T.; Hong, K.; Kirpotin, D. B.; Marks, J. D.; Park, J. W. Epidermal Growth Factor Receptor (EGFR) -Targeted Immunoliposomes Mediate Specific and Efficient Drug Delivery to EGFR- and

EGFRvIII-Overexpressing Tumor Cells Epidermal Growth Factor Receptor (EGFR) -Targeted Immunoliposomes Mediate Specific and Effi. *Cancer Res.* **2003**, *63*, 3154–3161.

(145) Masuda, H.; Zhang, D.; Bartholomeusz, C.; Doihara, H.; Hortobagyi, G. N.; Ueno, N. T. Role of Epidermal Growth Factor Receptor in Breast Cancer. *Breast Cancer Res. Treat.* **2012**, *136* (2), 331–345.

(146) Nieto, Y.; Nawaz, F.; Jones, R. B.; Shpall, E. J.; Nawaz, S. Prognostic Significance of Overexpression and Phosphorylation of Epidermal Growth Factor Receptor (EGFR) and the Presence of Truncated EGFRvIII in Locoregionally Advanced Breast Cancer. *J. Clin. Oncol.* **2007**, *25* (28), 4405–4413.

(147) Bublil, E. M.; Yarden, Y. The EGF Receptor Family: Spearheading a Merger of Signaling and Therapeutics. *Curr. Opin. Cell Biol.* **2007**, *19* (2), 124–134.

(148) Hackel, B. J.; Kapila, A.; Wittrup, K. D. Picomolar Affinity Fibronectin Domains Engineered Utilizing Loop Length Diversity, Recursive Mutagenesis, and Loop Shuffling. *J. Mol. Biol.* **2008**, *381* (5), 1238–1252.

(149) Kikuchi, O.; Ohashi, S.; Horibe, T.; Kohno, M.; Nakai, Y.; Miyamoto, S.; Chiba, T.; Muto, M.; Kawakami, K. Novel EGFR-Targeted Strategy with Hybrid Peptide against Oesophageal Squamous Cell Carcinoma. *Sci. Rep.* **2016**, *6* (February), 1–12.

(150) Ongarora, B. G.; Fontenot, K. R.; Hu, X.; Sehgal, I.; Satyanarayana-Jois, S. D.; Vicente, D. G. H. Phthalocyanine-Peptide Conjugates for Epidermal Growth Factor Receptor Targeting. *J. Med. Chem.* **2012**, *55* (8), 3725–3738.

(151) Tada, N.; Horibe, T.; Haramoto, M.; Ohara, K.; Kohno, M.; Kawakami, K. A Single Replacement of Histidine to Arginine in EGFR-Lytic Hybrid Peptide Demonstrates the Improved Anticancer Activity. *Biochem. Biophys. Res. Commun.* **2011**, *407* (2), 383–388.

(152) Zhang, F.; Wang, S.; Yin, L.; Yang, Y.; Guan, Y.; Wang, W.; Xu, H.; Tao, N. Quantification of Epidermal Growth Factor Receptor Expression Level and

Binding Kinetics on Cell Surfaces by Surface Plasmon Resonance Imaging. *Anal. Chem.* **2015**, *87* (19), 9960–9965.

(153) Gottschalk, N.; Kimmig, R.; Lang, S.; Singh, M.; Brandau, S. Anti-Epidermal Growth Factor Receptor (EGFR) Antibodies Overcome Resistance of Ovarian Cancer Cells to Targeted Therapy and Natural Cytotoxicity. *Int. J. Mol. Sci.* **2012**, *13* (12), 12000–12016.

(154) Chen, B.; Le, W.; Wang, Y.; Li, Z.; Wang, D.; Lin, L.; Cui, S.; Hu, J. J.; Hu, Y.; Yang, P.; Ewing, R. C.; Shi, D.; Cui, Z. Targeting Negative Surface Charges of Cancer Cells by Multifunctional Nanoprobes. *Theranostics* **2016**, *6* (11), 1887–1898.

(155) Suzuki, T.; Futaki, S.; Niwa, M.; Tanaka, S.; Ueda, K.; Sugiura, Y. Possible Existence of Common Internalization Mechanisms among Arginine-Rich Peptides. *J. Biol. Chem.* **2002**, *277* (4), 2437–2443.

(156) Szabó, D.; Mohl, J.; Bálint, A. M.; Bodor, A.; Rábai, J. Novel Generation Ponytails in Fluorous Chemistry: Syntheses of Primary, Secondary, and Tertiary (Nonafluoro-Tert-Butyloxy)Ethyl Amines. *J. Fluor. Chem.* **2006**, *127* (11), 1496–1504.

(157) Landry, J.; Sharov, A. A.; Piao, Y.; Sharova, L. V; Xiao, H.; Southon, E.; Matta, J.; Tessarollo, L.; Zhang, Y. E.; Ko, M. S.; Kuehn, M. R.; Yamaguchi, T. P.; Wu, C. Essential Role of Chromatin Remodeling Protein Bptf in Early Mouse Embryos and Embryonic Stem Cells. *PLoS Genet* **2008**, *4* (10), e1000241.

(158) Jones, M. H.; Hamana, N.; Shimane, M. Identification and Characterization of BPTF, a Novel Bromodomain Transcription Factor. *Genomics* **2000**, *63* (1), 35–39.

(159) Dar, A. A.; Nosrati, M.; Bezrookove, V.; de Semir, D.; Majid, S.; Thummala, S.; Sun, V.; Tong, S.; Leong, S. P.; Minor, D.; Billings, P. R.; Soroceanu, L.; Debs, R.; Miller 3rd, J. R.; Sagebiel, R. W.; Kashani-Sabet, M. The Role of BPTF in Melanoma Progression and in Response to BRAF-Targeted Therapy. *J.Natl.*

Cancer. Inst. **2015**, *107* (5).

(160) Xiao, S.; Liu, L.; Lu, X.; Long, J.; Zhou, X.; Fang, M. The Prognostic Significance of Bromodomain PHD-Finger Transcription Factor in Colorectal Carcinoma and Association with Vimentin and E-Cadherin. *J Cancer Res Clin Oncol* **2015**, *141* (8), 1465–1474.

(161) Kim, K.; Punj, V.; Choi, J.; Heo, K.; Kim, J. M.; Laird, P. W.; An, W. Gene Dysregulation by Histone Variant H2A.Z in Bladder Cancer. *Epigenetics and Chromatin* **2013**, *6* (1), 1–13.

(162) Dai, M.; Lu, J.-J.; Guo, W.; Yu, W.; Wang, Q.; Tang, R.; Tang, Z.; Xiao, Y.; Li, Z.; Sun, W.; Sun, X.; Qin, Y.; Huang, W.; Deng, W.; Wu, T. BPTF Promotes Tumor Growth and Predicts Poor Prognosis in Lung Adenocarcinomas. *Oncotarget* **2015**, *6* (32), 33878–33892.

(163) Urick, A. K.; Hawk, L. M. L.; Cassel, M. K.; Mishra, N. K.; Liu, S.; Adhikari, N.; Zhang, W.; Dos Santos, C. O.; Hall, J. L.; Pomerantz, W. C. K. Dual Screening of BPTF and Brd4 Using Protein-Observed Fluorine NMR Uncovers New Bromodomain Probe Molecules. *ACS Chem. Biol.* **2015**, *10* (10), 2246–2256.

(164) Ayoub, A. M.; Hawk, L. M. L.; Herzig, R. J.; Jiang, J.; Wisniewski, A. J.; Gee, C. T.; Zhao, P.; Zhu, J.-Y.; Berndt, N.; Offei-Addo, N. K.; Scott, T. G.; Qi, J.; Bradner, J. E.; Ward, T. R.; Schönbrunn, E.; Georg, G. I.; Pomerantz, W. C. K. BET Bromodomain Inhibitors with One-Step Synthesis Discovered from Virtual Screen. *J. Med. Chem.* **2017**, *60* (12), 4805–4817.

(165) Dickson, A.; Brooks, C. L. WExplore: Hierarchical Exploration of High-Dimensional Spaces Using the Weighted Ensemble Algorithm. *J. Phys. Chem. B* **2014**, *118* (13), 3532–3542.

(166) Dickson, A. Mapping the Ligand Binding Landscape. *Biophys. J.* **2018**, 1–13.

(167) Perell, G. T.; Mishra, N. K.; Sudhamalla, B.; Ycas, P. D.; Islam, K.; Pomerantz, W. C. K. Speci Fi c Acetylation Patterns of H2A.Z Form Transient Interactions with the BPTF Bromodomain. *Biochemistry* **2017**, *56*, 4607–4615.

(168) Almela, M. J.; Lozano, S.; Lelièvre, J.; Colmenarejo, G.; Coterón, J. M.; Rodrigues, J.; Gonzalez, C.; Herreros, E. A New Set of Chemical Starting Points with Plasmodium Falciparum Transmission-Blocking Potential for Antimalarial Drug Discovery. *PLoS One* **2015**, *10* (8), 1–18.

(169) Elkins, J. M.; Fedele, V.; Szklarz, M.; Abdul Azeez, K. R.; Salah, E.; Mikolajczyk, J.; Romanov, S.; Sepetov, N.; Huang, X. P.; Roth, B. L.; Al Haj Zen, A.; Fourches, D.; Muratov, E.; Tropsha, A.; Morris, J.; Teicher, B. A.; Kunkel, M.; Polley, E.; Lackey, K. E.; et al. Comprehensive Characterization of the Published Kinase Inhibitor Set. *Nat. Biotechnol.* **2016**, *34* (1), 95–103.

(170) Shi, J.; Wang, E.; Milazzo, J. P.; Wang, Z.; Kinney, J. B.; Vakoc, C. R. Discovery of Cancer Drug Targets by CRISPR-Cas9 Screening of Protein Domains. *Nat. Biotechnol.* **2015**, *33* (6), 661–667.

(171) Roe, J. S.; Hwang, C. Il; Somerville, T. D. D.; Milazzo, J. P.; Lee, E. J.; Da Silva, B.; Maiorino, L.; Tiriac, H.; Young, C. M.; Miyabayashi, K.; Filippini, D.; Creighton, B.; Burkhart, R. A.; Buscaglia, J. M.; Kim, E. J.; Grem, J. L.; Lazenby, A. J.; Grunkemeyer, J. A.; Hollingsworth, M. A.; et al. Enhancer Reprogramming Promotes Pancreatic Cancer Metastasis. *Cell* **2017**, *170* (5), 875–888.e20.

(172) Trott, O. Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2009**, *31* (2), 455–461.

(173) Li, H.; Fischle, W.; Wang, W.; Duncan, E. M.; Liang, L.; Murakami-Ishibe, S.; Allis, C. D.; Patel, D. J. Structural Basis for Lower Lysine Methylation State-Specific Readout by MBT Repeats of L3MBTL1 and an Engineered PHD Finger. *Mol. Cell* **2007**, *28* (4), 677–691.

(174) Dickson, A. Mapping the Ligand Binding Landscape. *bioRxiv* **2018**, 346817.

(175) Vanommeslaeghe, K.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* **2012**, *52* (12), 3144–3154.

(176) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* **2012**, *52* (12), 3155–3168.

(177) Dickson, A.; Lotz, S. D. Ligand Release Pathways Obtained with WExplore: Residence Times and Mechanisms. *J. Phys. Chem. B* **2016**, *120* (24), 5377–5385.

(178) Dickson, A.; Lotz, S. D. Multiple Ligand Unbinding Pathways and Ligand-Induced Destabilization Revealed by WExplore. *Biophys. J.* **2017**, *112* (4), 620–629.

(179) Lotz, S. D.; Dickson, A. Unbiased Molecular Dynamics of 11 Min Timescale Drug Unbinding Reveals Transition State Stabilizing Interactions. *J. Am. Chem. Soc.* **2018**, *140* (2), 618–628.

(180) Schwantes, C. R.; Pande, V. S. Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. *J. Chem. Theory Comput.* **2013**, *9* (4), 2000–2009.

(181) Bastian, M.; Heymann, S.; Jacomy, M. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Third Int. AAAI Conf. Weblogs Soc. Media* **2009**, 361–362.

(182) Lee, P.; Wu, X. Review: Modifications of Human Serum Albumin and Their Binding Effect. *Curr. Pharm. Des.* **2015**, *21* (14), 1862–1865.

(183) Elsadek, B.; Kratz, F. Impact of Albumin on Drug Delivery — New Applications on the Horizon. *J. Control. Release* **2012**, *157* (1), 4–28.

(184) Sleep, D.; Cameron, J.; Evans, L. R. Albumin as a Versatile Platform for Drug Half-Life Extension. *Biochim. Biophys. Acta - Gen. Subj.* **2013**, *1830* (12), 5526–5534.

(185) Kalhor-Monfared, S.; Jafari, M. R.; Patterson, J. T.; Kitov, P. I.; Dwyer, J. J.; Nuss, J. M.; Derda, R. Rapid Biocompatible Macrocyclization of Peptides with Decafluoro-Diphenylsulfone. *Chem. Sci.* **2016**, *7* (6), 3785–3790.

(186) Galantini, L.; Leggio, C.; Konarev, P. V.; Pavel, N. V. Human Serum Albumin Binding Ibuprofen: A 3D Description of the Unfolding Pathway in Urea. *Biophys. Chem.* **2010**, *147* (3), 111–122.

(187) Chamouard, J.-M.; Barre, J.; Urien, S.; Houin, G.; Tillement, J.-P. Diclofenac Binding to Albumin and Lipoproteins in Human Serum. *Biochem. Pharmacol.* **1985**, *34* (10), 1695–1700.

(188) Kim, H. S.; Hage, D. S. Chromatographic Analysis of Carbamazepine Binding to Human Serum Albumin. *J. Chromatogr. B* **2005**, *816* (1–2), 57–66.

(189) Sudlow, G.; Birkett, D.J.; Wade, D. N. The Characterization of Two Specific Drug Binding Sites on Human Serum. *Mol. Pharmacol.* **1975**, *11*, 824–832.

(190) Baroni, S.; Mattu, M.; Vannini, A.; Cipollone, R.; Aime, S.; Ascenzi, P.; Fasano, M. Effect of Ibuprofen and Warfarin on the Allosteric Properties of Haem-Human Serum Albumin. *Eur. J. Biochem.* **2001**, *268* (23), 6214–6220.