Development of Protein-Observed Fluorine Nuclear Magnetic Resonance Spectroscopy as a Ligand Discovery Technique

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Andrew K. Urick

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ADVISOR:
Dr. William C. K. Pomerantz

APRIL 2017

Acknowledgements

Thank you to all of my coworkers, collaborators, friends, and family who have helped me during my scientific career. In particular:

Will Pomerantz, who helped me develop as a scholarly scientist.
Haitao Hu, who helped me develop as an industrial scientist.
Neeraj Mishra, who helped me develop as a resourceful scientist.
Martin Hulce, who helped me develop as a thoughtful scientist.
Gary Michels, who helped me develop as a passionate scientist.

Abstract

Fragment-based drug design (FBDD) has been rapidly gaining traction in the drug discovery process. A central tenant of fragment-based molecular screening is to use less sophisticated small molecules to sample chemical space more efficiently. With Vemurafenib and Venetoclax as FDA approved therapeutics from FBDD and several others in Phase III clinical trials, FBDD is becoming a validated technique for drug discovery. However, because of their small size these fragments are likely to bind to their target with a low affinity, necessitating more sensitive methods to detect protein-ligand interactions during a screen. Nuclear magnetic resonance spectroscopy has emerged as one of several powerful biophysical techniques for conducting fragment screens. In this thesis, a ^{19}F protein-observed NMR method for detecting bromodomain−ligand interactions using fluorine-labeled aromatic amino acids due to the conservation of aromatic residues in the bromodomain binding site is described. Therein, we test the sensitivity, accuracy, and speed of this method with small molecule ligands. Experiment times on the order of a few minutes and the simplicity of the NMR spectra obtained make this approach well-suited to the investigation of small- to medium-sized proteins, as well as the screening of multiple proteins in the same experiment. Simplified ^{19}F NMR spectra allowed for simultaneous testing of multiple bromodomains to assess selectivity and identification of a new BPTF ligand. Fluorine labeling only modestly affected the Brd4 structure and function assessed by isothermal titration calorimetry, circular dichroism, and X-ray crystallography. To benchmark its potential as a ligand discovery tool, we compare the protein-observed ^{19}F NMR screening method with the well-characterized ligand-observed ^{1}H CPMG NMR screen. We selected the first bromodomain of Brd4 as a model system because of the high ligandability of Brd4 and the need for small molecule inhibitors of related epigenetic regulatory proteins. We conclude that for the protein class understudy here, protein-observed ^{19}F NMR and ^{1}H CPMG have similar sensitivity, with both being effective tools for ligand discovery. The speed, ease of interpretation, and low

concentration of protein needed for binding experiments affords a new method to discover and characterize both native and new ligands.

Table of Contents

List of Tables

List of Figures

Preface

Chapter 1 of this dissertation provides a background on the biological motivations of the project, as well as the various biophysical methods utilized. Bromodomains were chosen due to the availability of the well-characterized bromodomain of Brd4 coupled with the need for chemical probes for the majority of the other bromodomains. In order to study these proteins, various biophysical methods to assess structure and function of the bromodomains were utilized such as circular dichroism, x-ray crystallography, nuclear magnetic resonance spectroscopy (NMR), and others.

In Chapter 2 of this thesis, a ^{19}F NMR method for detecting bromodomain−ligand interactions using fluorine-labeled aromatic amino acids due to the conservation of aromatic residues in the bromodomain binding site is described. Therein, we test the sensitivity, accuracy, and speed of this method with small molecule ligands (+)-JQ1, BI2536, Dinaciclib, TG101348, and acetaminophen using three bromodomains Brd4, BrdT, and BPTF. Simplified ^{19}F NMR spectra allowed for simultaneous testing of multiple bromodomains to assess selectivity and identification of a new BPTF ligand. Fluorine labeling only modestly affected the Brd4 structure and function assessed by isothermal titration calorimetry, circular dichroism, and X-ray crystallography. The speed, ease of interpretation, and low concentration of protein needed for binding experiments affords a new method to discover and characterize both native and new ligands.

The protocol in Chapter 3 describes the sequence-selective labeling of three proteins (the first bromodomains of Brd4 and BrdT, and the KIX domain of the CREB-binding protein) using commercially available fluorinated aromatic amino acids and fluorinated precursors as example applications of the method developed by our research group. NMR spectroscopy can be used to quantify the binding affinity between proteins and low-complexity molecules, termed ‘fragments’; this versatile screening approach allows researchers to assess the druggability of new protein targets. Protein-observed ^{19}F-NMR (PrOF NMR) using ^{19}F-labeled amino acids generates relatively simple spectra that

are able to provide dynamic structural information toward understanding protein folding and function. Changes in these spectra upon the addition of fragment molecules can be observed and quantified. Fragment-screening approaches are discussed, as well as K_d determination, ligand-efficiency calculations and druggability assessment, i.e., the ability to target these proteins using small-molecule ligands. Experiment times on the order of a few minutes and the simplicity of the NMR spectra obtained make this approach well-suited to the investigation of small- to medium-sized proteins, as well as the screening of multiple proteins in the same experiment. To further decrease the time of PrOF NMR experiments, paramagnetic relaxation enhancement through the addition of chelated Ni(II) can be used to shorten longitudinal relaxation time. Enhancing relaxation time leads to shorter experiments without perturbing the binding of low- or high- affinity ligands. This method allows for time-efficient screening of potential ligands for a wide variety of proteins in the growing field of fragment-based ligand discovery.

Selective inhibition of bromodomain function is a newly proposed therapeutic strategy. In Chapter 4, a ^{19}F NMR dual screening method for small molecule discovery using fluorinated tryptophan resonances on two bromodomain-containing proteins is described. The chemical shift dispersion of ^{19}F resonances within fluorine-labeled proteins enables the simultaneous analysis of two fluorinated bromodomains by NMR. A library of 229 small molecules was screened against the first bromodomain of Brd4 and the BPTF bromodomain. We report the first small molecule selective for BPTF over Brd4, termed AU1. The K_d = 2.8 μM for AU1, which is active in a cell-based reporter assay. No binding is detected with Brd4. Three new Brd4 inhibitors with submicromolar affinity were also discovered. Brd4 hits were validated in a thermal stability assay and potency determined via fluorescence anisotropy. The speed, ease of interpretation, and low protein concentration needed for protein-observed ^{19}F NMR experiments in a multiprotein format offers a new method to discover and characterize selective ligands for bromodomain-containing proteins.

In Chapter 5, to evaluate its potential as a ligand discovery tool, we compare a newly developed 1D protein-observed fluorine NMR screening method with the well-

characterized ligand-observed ^{1}H Carr-Purcell-Meiboom-Gill (CPMG) NMR screen. We selected the first bromodomain of Brd4 as a model system to benchmark PrOF NMR because of the high ligandability of Brd4 and the need for small molecule inhibitors of related epigenetic regulatory proteins. We compare the two methods' hit sensitivity, triaging ability, experiment speed, material consumption, and the potential for false positives and negatives. To this end, we screened 930 fragment molecules against Brd4 in mixtures of five and followed up these studies with mixture deconvolution and affinity characterization of the top hits. In selected examples, we also compare the environmental responsiveness of the ^{19}F chemical shift to ^{1}H in 1D-protein observed ^{1}H NMR experiments. To address concerns of perturbations from fluorine incorporation, ligand binding trends and affinities were verified via thermal shift assays and isothermal titration calorimetry. We conclude that for the protein understudy here, PrOF NMR and ^{1}H CPMG have similar sensitivity, with both being effective tools for ligand discovery. In cases where an unlabeled protein can be used, 1D protein- observed ^{1}H NMR may also be effective; however, the ^{19}F chemical shift remains significantly more responsive.

A key necessity of PrOF NMR, is resonance assignment of the protein's ^{19}F NMR spectrum for structural analysis. In Chapter 6, a quantum chemical method has been developed as an initial approach to facilitate the assignment of a fluorinated protein's ^{19}F NMR spectrum. The epigenetic "reader" domain of protein Brd4 was taken as a case study to assess the strengths and limitations of the method. The overall modeling protocol predicts chemical shifts for residues in rigid proteins with good accuracy; proper accounting for explicit solvation of fluorinated residues by water is critical.

Together these initial findings in this thesis describe my contributions to the field of fragment-based ligand discovery through development of a new protein-based ^{19}F NMR biophysical technique.

Citations from previously published work in this dissertation are as follows:

Chapter 2:

"Fluorinated Aromatic Amino Acids Are Sensitive ^{19}F NMR Probes for Bromodomain-Ligand Interactions," N. K. Mishra, A. K. Urick, S. W. Ember, E. Schönbrunn, W. C. Pomerantz, *ACS Chemical Biology* **2014**, 2755-2760.

Chapter 3:

"Protein-observed ^{19}F-NMR for fragment screening, affinity quantification and druggability assessment," C. T. Gee, K. E. Arntson, A. K. Urick, N. K. Mishra, L. M. L. Hawk, A. J. Wisniewski, W. C. K. Pomerantz, *Nature Protocols* **2016**, 1414-1427.

*"*Paramagnetic relaxation enhancement for protein-observed ^{19}F NMR as an enabling approach for efficient fragment screening," L. M. L. Hawk, C. T. Gee, A. K. Urick, H. Hu, W. C. K. Pomerantz, *RSC Advances* **2016**, 95715-95721.

Chapter 4:

"Dual Screening of BPTF and Brd4 Using Protein-Observed Fluorine NMR Uncovers New Bromodomain Probe Molecules," A. K. Urick, L. M. L. Hawk, M. K. Cassel, N. K. Mishra, S. Liu, N. Adhikari, W. Zhang, C. O. dos Santos, J. L. Hall, W. C. K. Pomerantz, *ACS Chemical Biology* **2015**, 2246-2256.

Chapter 5:

"Protein-Observed Fluorine NMR is a Complementary Ligand Discovery Method to ^{1}H CPMG Ligand-Observed NMR," A. K. Urick, L. P. Calle, J. F. Espinosa, H. Hu, W. C. K. Pomerantz, *ACS Chemical Biology* **2016***,* 3154-3164.

Chapter 6:

"Prediction of ^{19}F NMR Chemical Shifts in Labeled Proteins: Computational Protocol and Case Study," W. C. Isley III, A. K. Urick, W. C. K. Pomerantz, C. J. Cramer, *Molecular Pharmaceutics* **2016**, 2376-2386.

Chapter 1. Introduction

1.1 **Epigenetics**

Epigenetics is the study of mitotically and meiotically heritable changes in gene function that are not solely-dependent on DNA sequence.[1] Every cell in a human body contains the same DNA sequence in the nucleus, but each can exhibit different phenotypic properties. These differences are in part due to epigenetic factors, and so the ability to modulate these factors bestows control over biological systems. Dysregulation of these factors is frequently implicated in disease, such as cancer,[2] heart disease,[3] and inflammation[4]. Epigenetic mechanisms include RNA transcription, nucleosome positioning, DNA modification independent of nucleotide sequence, and covalent chromosomal modifications. The biological motivations of this work focus primarily on covalent chromosomal modifications.

A critical factor in determining whether DNA is expressed or not is the accessibility of the chromatin environment.[5] Euchromatic environments, where DNA is not coiled around histone proteins, are generally accessible while heterochromatic environments, where DNA is coiled around histone proteins, are generally inaccessible. To compact DNA sufficiently to fit within the nucleus of a cell, DNA is tightly folded and constrained by histone and nonhistone proteins in a dynamic polymer called chromatin. Chromatin organization is dependent on higher order structure with the basic repeating unit of nucleosomes. Each nucleosome is roughly two superhelical turns of DNA wrapped around an octamer core of histone proteins. While wrapped in the chromatin and nucleosome complex, DNA is transcriptionally inert.[6] In order to regulate whether or not DNA is accessible, and consequently whether portions of DNA are expressed, amino acid side-chains of histones can be covalently modified. Such modifications include acetylation, methylation, phosphorylation, glycosylation, crotonylation, and citrullination, among others (Figure 1.1).[7] Modification of these tails is highly selective for specific amino acid sequences.[8,9] Heritable and reversible changes to histone tails have also been observed to occur.[10]

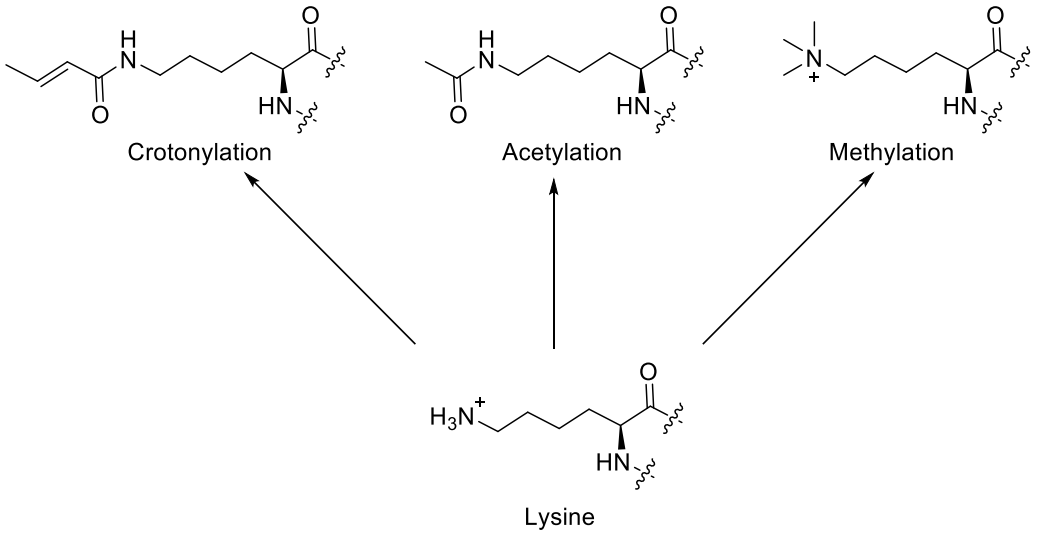


Figure 1.1 Example posttranslational modifications of lysine.

The mechanism of histone modifications regulating DNA expression has been termed the Histone Code Hypothesis.[11] The installation, removal, and recognition of histone modifications is facilitated by the colloquially termed "writers", "readers", and "erasers". A small subset of these regulatory proteins work with acetylation and methylation of lysine sidechains. Writers, such as histone acetyltransferases (HATs), and histone methyltransferases (HMTs), post-translationally modify histones with functional groups such as methyl or acetyl groups. Readers, such as bromodomains and chromodomains, recognize the post-translational modifications made by the writers, and moderate DNA transcription. Erasers, such as histone deacetylase (HDAC) and histone demethylase (HMD) remove the marks made by HATs and HMTs respectively.[11]

1.2 **Bromodomains**

Bromodomains act as readers of acetylated-lysines of histones, playing an integral role in DNA transcription. There are 46 distinct human proteins containing a total of 61 bromodomains (Figure 1.2).[12] All bromodomain modules share a conserved fold comprising a left-handed bundle of four α-helices (α_Z, α_A, α_B, α_C) linked by loop regions (ZA and BC loops) that contribute to substrate specificity. X-ray co-crystal structures with peptides containing acetylated-lysine residues show that the binding site is a central hydrophobic cavity with the acetylated-lysine anchored with a hydrogen bond to a highly

conserved asparagine residue (e.g, N140 of Brd4, Figure 1.3). The loop regions surrounding the acetylated lysine binding site are shown to have high conformational flexibility by computational dynamics studies.[13] Despite the highly conserved fold of the bromodomains, their surface properties around the acetylated-lysine binding site are highly diverse. The electrostatic potential of the surface area around the binding site ranges from highly positively charged (+10 kT/e (electrostatic potential; k ≡ Boltzmann's constant, T ≡ temperature, e ≡ charge of an electron)) to strongly negatively charged (-10 kT/e).[12] Often the bromodomain motif is flanked by other epigenetic regulatory proteins, including plant homeodomains (PHD), other bromodomains, and even histone acetyltransferases.

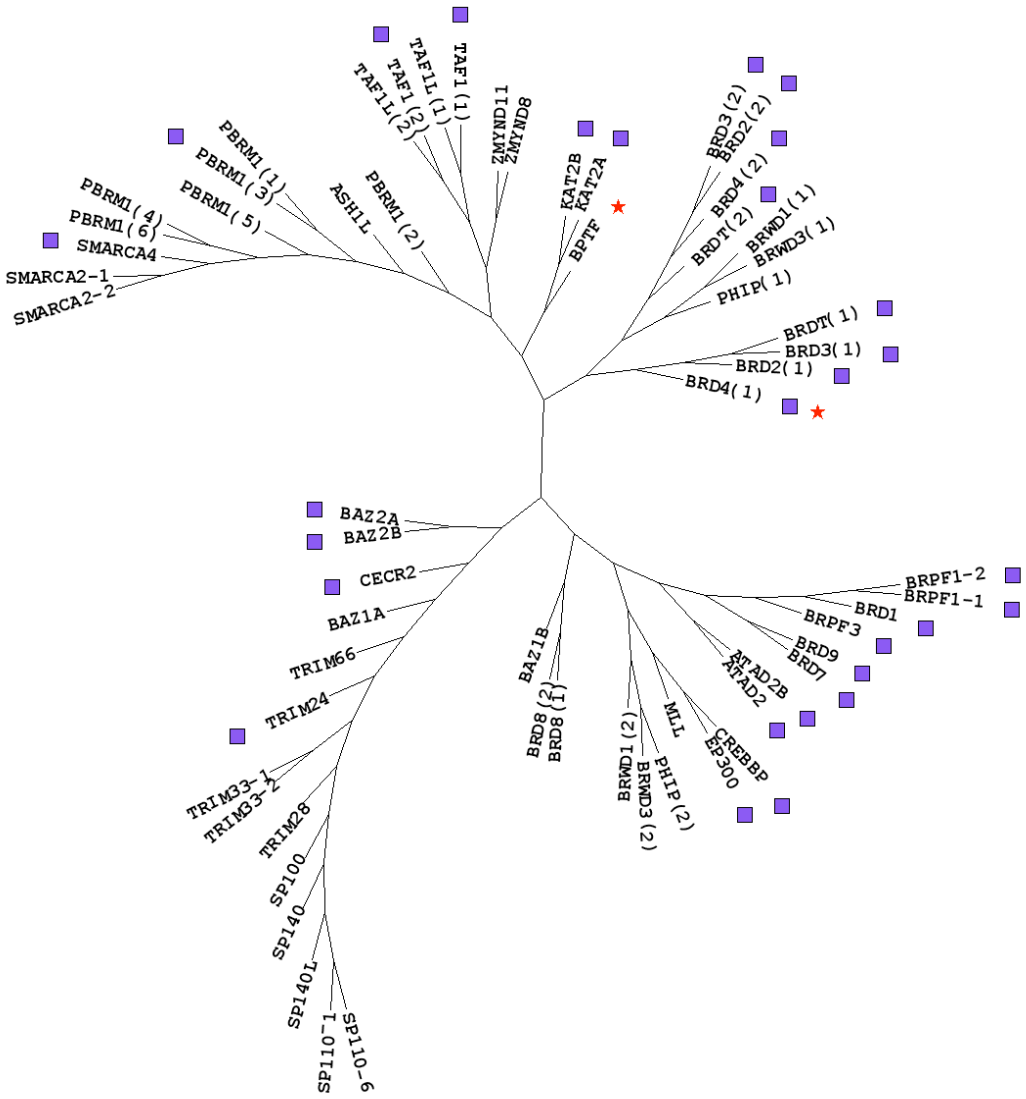


Figure 1.2 A phylogenetic tree of the 61 human bromodomains.
This plot clusters proteins by sequence similarity. The primary proteins under study in this dissertation are highlighted with stars. Bromodomains with a chemical probe (potent, selective, and cell active) with an $IC_{50}/K_d/K_i < 0.5$ μM are highlighted with a square. This figure was generated using ChromoHub.[14]

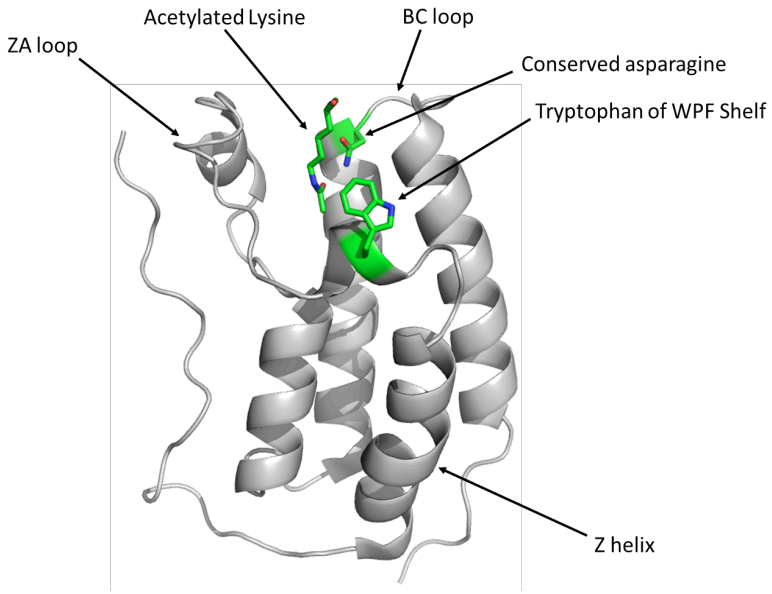


Figure 1.3 The first bromodomain of Brd4.
This is an x-ray structure of the first bromodomain of Brd4 bound with a histone peptide containing an acetylated lysine. Various structural areas are labeled. PDB ID: 3UVW

Bromodomains have been evidenced to play a role in inflammation, obesity, and cancer.[15] They were long viewed as undruggable, as were most protein-protein interactions due to the interface being large, featureless, and lacking an initial natural substrate as a starting place.[16] DMSO also binds in the acetylated-lysine binding site which would convolute high throughput screens in which the molecules are dissolved in DMSO stocks. Despite this view, in 2010 two BET bromodomain inhibitors were developed, JQ1 (K_d= 50 nM)[13] and I-BET (K_d= 55.2 nM)[17]. The molecule JQ1 has been used to study the BET (Bromodomain and extraterminal) bromodomains, but it does not have high specificity toward any one BET bromodomain. It cannot yet be unambiguously determined which bromodomain interaction is leading to specific phenotypes because an inhibitor that only targets a single bromodomain has not been disclosed. Of the 46 bromodomain containing proteins, 14 have been targeted with modest degrees of affinity and selectivity, (CBP, Brd2(1), Brd2(2), Brd3(1), Brd3(2), Brd4(1), Brd4(2), BrdT(1), BrdT(2), BPTF, Brd7, Brd9, BAZ2B, EP300).[18] Chemical probes that target a specific BET bromodomain, as well as probes to target the rest of the 61 bromodomains, are needed. Despite the few small-molecule inhibitors known for bromodomains at the time,

analysis using Schrödinger's SiteMap by Vidler *et al.* suggested that thirteen bromodomains are characterized with high druggability, with five more being suggested as moderately druggable.[19] The BET bromodomains are included in the category of being druggable, which fits with the SiteMap analysis because small-molecule inhibitors have been found for these bromodomains. Other bromodomains, such as the BPTF bromodomain (also known as fetal Alz-50 clone 1 protein (FALZ)), were predicted to be highly druggable. This lack of inhibitors motivated us to choose the bromodomain of BPTF as our protein target as described further in Chapter 4.

Structural biology efforts by both NMR and x-ray crystallography have provided detailed atomic level information for guiding designs of small molecules inhibitors as well as a characterize native histone interactions. As mentioned previously, the bromodomains contain a hydrophobic pocket to recognize acetylated lysines which can vary substantially in composition.[20] Using NMR structural analysis, the seminal study by Dhalluin *et al.* concluded that ligands binding to the bromodomain of PCAF (a bromodomain part of the same family as BPTF) frequently make significant contact with the side chains of three tyrosine residues in the binding site (Y809, Y802, Y760).[20] The acetylated-lysine of the ligand forms a water-mediated hydrogen bond with a highly conserved asparagine residue (N803),[19] results which were further confirmed by x-ray crystallography (Figure 1.4).[21] All of these residues are highly conserved across bromodomains, making fluorinated tyrosine residues an appealing target for binding analysis by PrOF NMR as described in Chapter 2. While the inner regions of the acetylated-lysine binding pocket are largely hydrophobic, the electrostatics of the surface vary among bromodomains. As such, the surface surrounding the pocket of the acetylated-lysine binding site is postulated to be the ideal area to target to affect selectivity between the various bromodomains.[22]

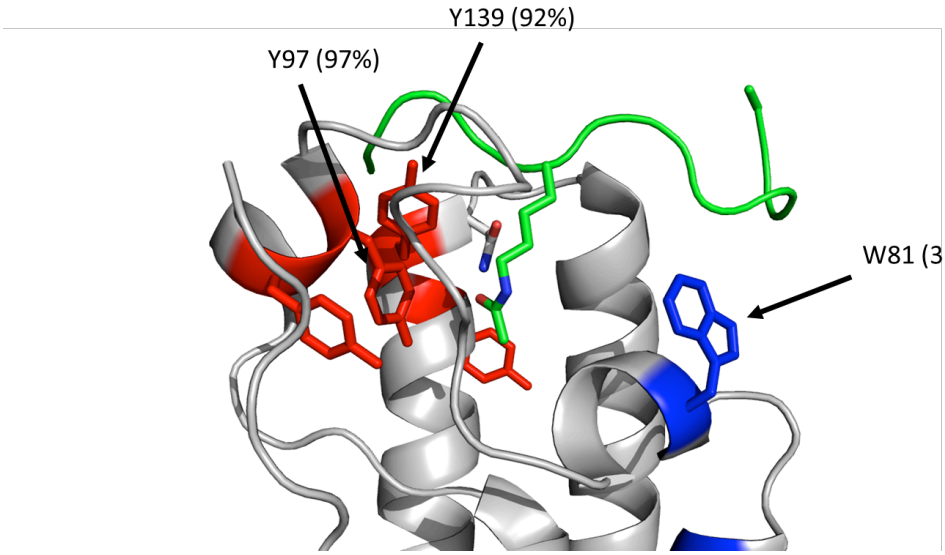


Figure 1.4 Tyrosine and tryptophan residues in the binding site of Brd4.
The percentages represent the conservation of aromatic amino acids at the homologous position in the bromodomain family.

The surface of the acetylated-lysine binding pocket contains a tryptophan-proline-phenylalanine sequence, known as the WPF shelf. This shelf is present in many bromodomains, including Brd4 and BPTF, and is always near the acetyl-lysine binding site, though does vary in that distance between different bromodomains. In Brd4, the WPF shelf is within 7 Å of the acetylated lysine itself, but being near the opening of the binding site, it is frequently under 4 Å from a heavy atom in a binding ligand. Additionally, b-factor analysis of x-ray crystal structures of BPTF indicate greater dynamics and flexibility of the WPF shelf tryptophan, indicating that it is likely closer to the binding site at times than the static structure would indicate. The WPF shelf has been known to also participate in dual binding of acetylated-lysine ligands, wherein one acetylated-lysine binds in the standard binding pocket, while one will bind to the tryptophan residue on the WPF shelf (Figure 1.5).[23]

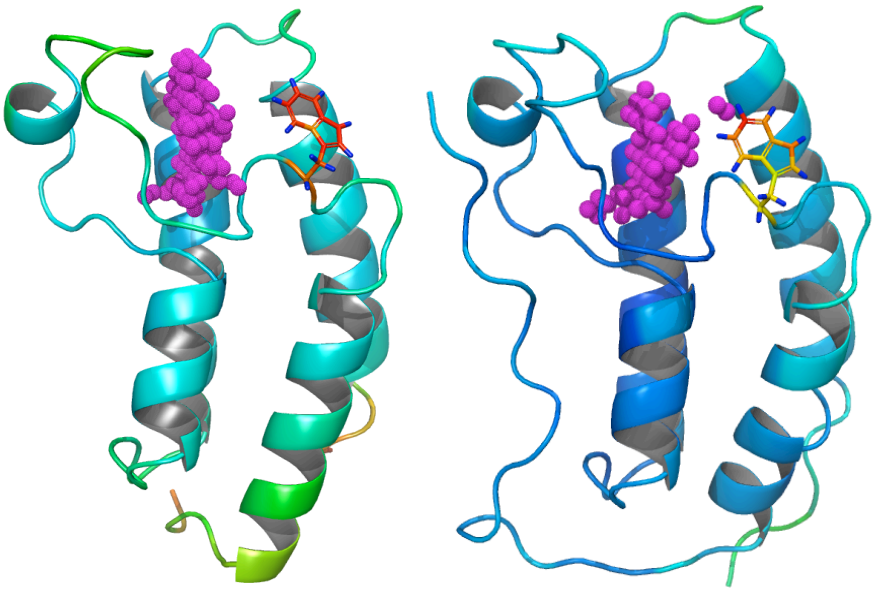


Figure 1.5 The bromodomains of BPTF (Left) and Brd4 (Right) colored by b-factor. More red indicates greater flexibility while more blue indicates greater rigidity. The explicitly shown tryptophan residues have a higher b-factor, indicated greater dynamics of that sidechain than the amide backbone of the protein. The spheres indicate the acetylated-lysine binding site.

1.2.1 The bromodomain of Brd4

Bromodomain containing protein 4 (Brd4) is a BET bromodomain, dysregulation of which has been linked with various diseases. It contains two bromodomains, BD1 and BD2. It was initially identified in purifications of mammalian Mediator complex, which links transcription factors to the Pol II C-terminal domain.[24] Brd4 has been found to be overexpressed in breast cancer biopsies, with Brd4 transcriptional signatures directly correlating to the rate of disease progression.[25] Brd4 dysregulation has also been implicated in nuclear protein in testis (NUT) midline carcinoma.[26] Consistent with these observations, JQ1 treatment significantly inhibited tumor growth in mice in two models of NUT midline carcinoma.[13] Additionally, a small hairpin RNA screen identified Brd4 as playing an essential role in leukemia.[27]

Molecules repurposed from the kinase inhibitor field were found to act as privileged scaffolds for bromodomain inhibition. Dinaciclib was the first disclosed kinase inhibitor with activity toward Brd4,[28] followed up by several more extensive studies that

found many kinase inhibitor and kinase-inhibitor-like molecules with activity toward bromodomains.[29,30] Though there are currently no published small molecules that are selective for Brd4 exhibiting submicromolar potency, there are many molecules that target the general BET-family. Such structures include isoxazoles,[31] triazoles,[13,17] amidines,[32] ureas,[32,33] and oxadiazoles.[32] Additionally, because the natural substrate of Brd4 is acetylated-lysine residues on histone tails, histone peptides containing acetylated-lysine also bind Brd4 with appreciable affinity (Table 1.1). The more acetylated-lysines the histone peptides contain, the tighter the affinity tends to be with the bromodomain of Brd4.

Table 1.1 Peptide Mimetic Binders for Brd4-BD1

Histone Mimic	K_d (μM) for Brd4-BD1	Full Sequence	Method
H4K5/8/12/16	2.8 ± 0.2	YSGRG**K**$_{ac}$GG**K**$_{ac}$GLG**K**$_{ac}$GGA**K**$_{ac}$RHRK	ITC[34]
H4K12/16/20	20.4 ± 0.8	G**K**$_{ac}$GGA**K**$_{ac}$RHR**K**$_{ac}$V	ITC[34]
H4K8/12	27.4 ± 0.9	G**K**$_{ac}$GLG**K**$_{ac}$GGAKR	ITC[34]
H4K5/8	38 ± 3.7	GRG**K**$_{ac}$GG**K**$_{ac}$GLGKG	ITC[35]
H4K12/16	46.1 ± 0.9	G**K**$_{ac}$GGA**K**$_{ac}$RHRKV	ITC[34]
H4K8	84.7 ± 9.2	RGKGG**K**$_{ac}$GLGKGY	ITC[12]
H3K14	118 ± 28	GG**K**$_{ac}$APRKQ	ITC[35]
H3K9	301 ± 40.9	KQTAR**K**$_{ac}$STGGKY	ITC[12]
H4K12	650 ± 11	SGRGKGGKGLG**K**$_{ac}$GGAK	NMR[12]
H4K5	810 ± 57	SGRG**K**$_{ac}$GGKGLGKGGAK	NMR[34]

1.2.2 BPTF

Bromodomain PHD finger transcription factor (BPTF) is a protein that contains both a bromodomain and a planthomeo domain (PHD) finger, for which no small molecule inhibitors were known prior to our study. Originally isolated in Drosophila, BPTF is the largest member of the Nucleosome Remodeling Factor complex (NURF), which was shown through *in vivo* whole genome expression studies to regulate

transcription of several hundred Drosophila genes.[36] Human NURF contains homologues of three of the four NURF components in Drosophila, including BPTF.[37] It has also been found that BPTF is overexpressed in various human tumors, such as lung cancers. Knockdown of BPTF production resulted in reduced embryonic lung cancer cell proliferation.[38] Since then, BPTF has been correlated with various disease states, including heart disease, melanoma,[39] lung cancer,[38] Alzheimer's,[40] gastric and colorectal cancers[41] (Figure 1.7). While the direct role of BPTF in cancer has not been unambiguously elucidated, Real et al. has shown that BPTF is required for c-myc transcriptional activity in tumor cells.[42] Additionally, BPTF has been correlated to roles in spermatogenesis,[43] hematopoiesis,[44] and essential roles in mouse embryonic development[45].

Figure 1.6 Bar representation of the BPTF protein.
The bromodomain and PHD fingers are explicitly shown. The bromodomain is a small fraction of the overall protein. The entire BPTF protein weighs 338 kDa, with the bromodomain portion weighing 14 kDa (4% of total protein weight).

Because BPTF contains a PHD finger, studies involving gene knockdown of BPTF do not offer specific information as to whether the PHD finger or the bromodomain of BPTF is responsible for any phenotypic changes resulting from the knockdown of BPTF. A selective inhibitor of the acetylated-lysine binding site of the BPTF bromodomain would allow for greater elucidation into the specific rolls BPTF plays in various diseases, including lung cancer.

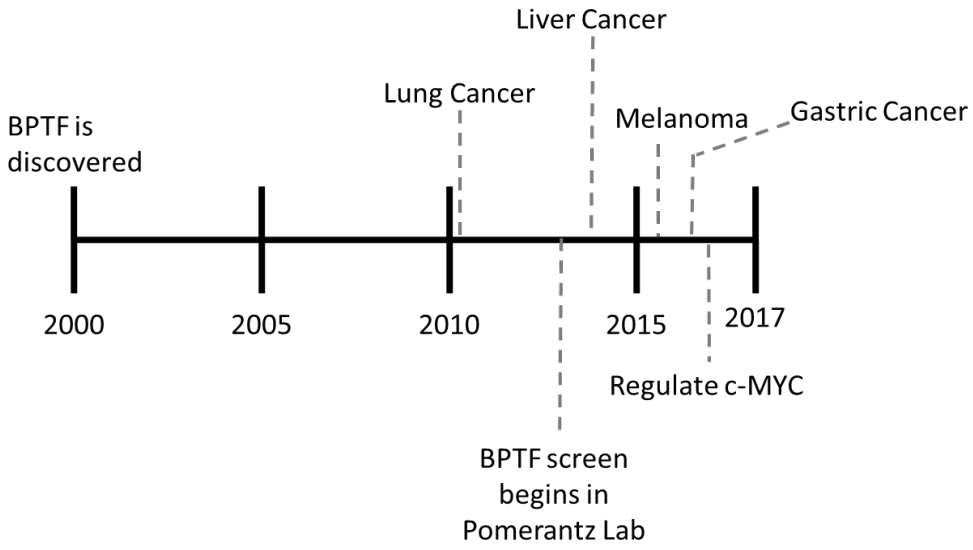


Figure 1.7 A timeline of BPTF research.

1.2.3 Targeting bromodomain protein-protein interactions

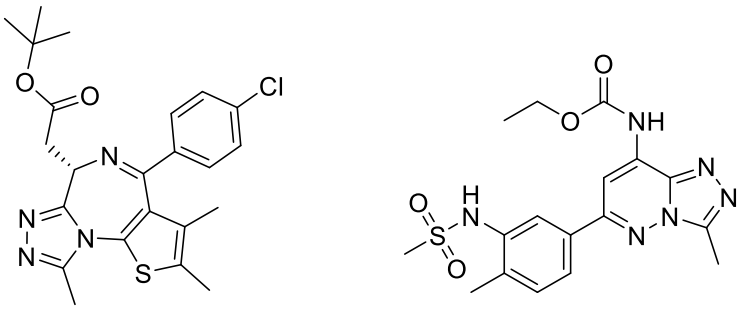
The Histone Code Hypothesis, along with other epigenetic and biological aspects, involves protein-protein interactions (PPIs). There are an estimated 650,000 protein-protein interactions in the human body,[46] with an estimated 23,000 proteins encoded in human DNA. This means each protein is involved with an average of 27 protein-protein interactions. PPIs are challenging drug targets, often characterized as large, featureless, and shallow interfaces. Nevertheless, there are sometimes hotspot amino acids involved in PPIs that contribute large fractions (usually greater than 1.0 kcal/mol) to the Gibb's Free Energy of a binding interaction.[47] Normalizing for the natural abundance of amino acids in proteins, the tryptophan, tyrosine, and arginine are the most likely residues to be at a protein-protein interface.[48] The aromatic amino acids are also the most likely to act as hotspots in α-helical peptide-protein interactions, often contributing ~3.0 kcal/mol to the binding energy.[49]

Despite their challenging nature, small molecules have been successfully used to modulate PPIs. Nutlin-3 is one of the more prominent examples, successfully binding mouse double minute 2 homolog (MDM2) to activate the p53 pathway in cancer cells.[50] Recently, the AbbVie compound Venetoclax was approved by the FDA and is the first small molecule by which the primary mechanism is inhibition of a protein-protein interaction involving Bcl-2.[51] PPIs represent valuable therapeutic targets, but there are

many PPIs that do not yet have known small molecule binders due to the inherent difficulty of finding a small molecule that can mimic a protein. It is challenging to configure assays that effectively identify inhibitors of PPIs, requiring new technologies and screening methodologies to better enable the discovery of new drugs to modulate PPIs, including bromodomains.[52]

Because until 2010 there were no potent (i.e., submicromolar K_d) known small-molecule inhibitors for bromodomains, there were limited biophysical assays available with which to assess bromodomain-ligand interactions. Methods that rely on competitors, such as fluorescence polarization or "spy" based NMR experiments, were not useful for this target class due to not having potent small-molecule inhibitors. Jay Bradner disclosed JQ1, with off-target effects to other bromodomains assessed by thermal shift assays and α-screen. Thermal shift assays measure the melting temperature of a protein in presence and absence of ligand, and assume that the favorable binding energy of a ligand will increase the required temperature to denature the protein. While thermal shift assays tend to be poor at rank-ordering affinities among different ligands, for bromodomains the correlation tends to be modestly high. An α-screen assay attaches protein and a known ligand to separate beads which fluoresce when near each other. If an added molecule displaces the control ligand, the beads move away from each other and the decrease in fluorescence can be measured. The α-screen assays were enabled by using a truncated histone tail containing acetylated-lysine, the natural substrate of bromodomains. While the affinity for the histone tails is low, the high density of histone tail peptides on the beads allows for an avidity effect, heightening the sensitive of the assay. Thus, α-screen was still useful as a basic competition experiment. Direct binding was then verified and quantified by isothermal titration calorimetry.[13]

Due to the increased efforts to inhibit bromodomains in the past several years, there are now many inhibitors available for several different bromodomains. Bromosporine offers a general bromodomain inhibitor, potentially allowing general access to competition experiments for various biophysical assays such as fluorescence polarization as well as competition NMR experiments (Figure 1.8).[53]



(+)-JQ1 Bromosporine

Figure 1.8 Chemical structures of bromodomain binders (+)-JQ1 and bromosporine.

1.3 **Fragment-based drug design**

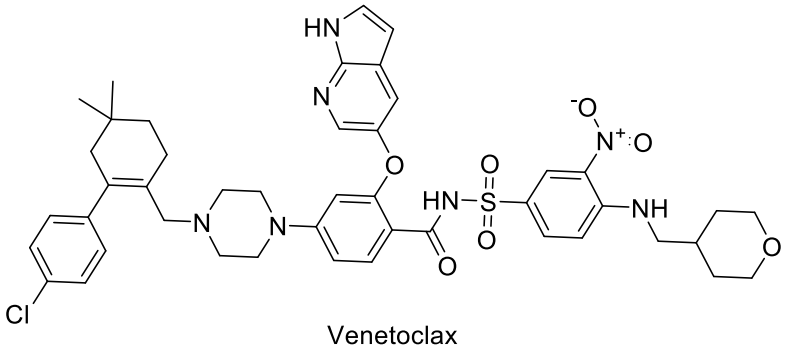
The principle purpose of drug discovery and development is to identify molecules that interact strongly with biomolecules to produce a desired biological response.[54] This initially took the form of phenotypic drug discovery, which requires no prior knowledge of the desired biological target nor its function. The majority of current drug targets are G-protein-coupled receptors, nuclear receptors, ion channels, and enzymes; all of which were initially targets of phenotypic drug discovery approaches. With the advent of genomics and proteomics, biological target identification has been greatly accelerated. The opportunity is now present to efficiently interrogate the role of certain proteins in a given disease state, and better allows for a target-first strategy of drug design. New approaches to target-first drug discovery often take the forms of structure-based drug design, or fragment-based drug discovery.

Over the last 20 years, fragment-based drug discovery has been gaining increasing traction in the pharmaceutical industry. Over 30 drug candidates derived from fragments have entered the clinic, with two approved and several currently in advanced clinical trials.[55–57] Fragment based drug discovery offers a method to sample a broader area of chemical space, compared with high throughput drug screening. A report in Medicinal Research Reviews states the total possible combination of small drug-like molecules (<31 atoms with only C, N, O, S, and H) is 10^{63}; as such, sampling a substantial fraction of that space is extremely challenging.[58–60] High-throughput screens sometimes test up to two

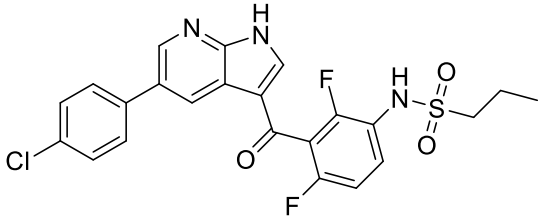
million compounds.[61] Even large screens such as these result in only a small percentage (10^{-52}) of chemical space being sampled. Indeed, despite the ability to screen millions of compounds, high-throughput screens will at times yield very few hits, or actives that are false positives.[62] While the space tested can be made more effective by rationally choosing categories of molecules (i.e. privileged scaffolds), in combination with computational docking studies, serendipity is still a substantial factor in whether or not a useful molecule is screened. A more effective way to increase the amount of chemical space being sampled is to screen with less complex molecules; fragments.[63] A fragment screen will sample a greater amount of potential chemical space with each compound than the more complex, and consequently more specific molecules utilized in a high-throughput screen. Molecular weight is often a proxy for complexity, so fragment libraries tend to contain molecules with a weight < 300 g/mol. This is encompassed in the fragment "Rule of three" which specifies molecular weight < 300, the number of hydrogen bond donors is ≤ 3, the number of hydrogen bond acceptors is ≤ 3, and clogP (computational derived octanol-water partition coefficient) is ≤ 3.[64] Of particular importance for fragments due to their likely low affinity for targets is high water solubility to enable high screening concentrations. Additionally, a strong advantage of fragment screening over high-throughput screening is that due to the simplicity of the molecules being screened, there is lower probability that a specific steric interaction will prevent an otherwise favorable molecule from binding to the desired binding site. Fragments are less sophisticated than the compounds utilized in high-throughput screening, and bind more weakly than compounds used in high-throughput screening. Thus, fragment screening requires more lead optimization, as the hits from such a screen will have low dissociation constants initially (typically between 50 μM and 1 mM)[65], requiring a substantial amount of work to modify the hits into a molecule that binds with high affinity.[66]

The two FDA approved therapeutics derived from fragment based drug design are Vemurafenib[55] and Venetoclax[56] (Figure 1.9). Vemurafenib targets BRAF and is a treatment for melanoma, and is elaborated from an initially identified azaindole fragment. Venetoclax is the result of the seminal work in protein-observed NMR screening

disseminated by Abbott Laboratories, and targets BCL-2 to treat lymphoma.[67] A linking strategy between two fragments was used to develop Venetoclax, and 20 years passed between initial discovery of the fragments and FDA approval. The substantial time required casts doubt on the efficacy of linking fragments to develop high-affinity inhibitors, and consequently compound elaboration continues to be the more common method to develop lead compounds from fragment hits.



Venetoclax



Vemurafenib

Figure 1.9 Structures of the two FDA approved drugs derived from fragments.

1.4 **The role of biophysical assays in FBDD**

One of the major disadvantages of fragment screening compared with high-throughput screening is the increased sensitivity of detectors required for sufficient analysis. Because of this, traditional phenotypic screens as well as biochemical screens frequently lack the required sensitivity to effectively conduct a fragment screen. Biophysical assays satisfy these increased requirements for sensitivity, and have been rapidly adapted to allow for increasing automation for use in screen. The classical fragment screening methods are NMR spectroscopy, surface plasmon resonance (SPR), and x-ray crystallography. Recent developments such as thermal shift assays (TSA) and

microscale thermophoresis (MST), have successfully been used to analyze fragment binding and are also amiable for screening.

1.4.1 NMR spectroscopy

Due to the high-responsiveness of NMR spectroscopy, NMR experiments were the first methods utilized for fragment screening. Since the initial reports from Abbott Laboratories, NMR has become increasingly utilized in the drug discovery process as both a primary screening technique and as a follow-up assay.[67] The wealth of information provided, as well as the ability to perform both protein- and ligand-observed experiments make NMR an incredibly versatile tool in the drug discovery process. The recent advent of cryo-probe technology has allowed NMR screening to be a realized technique, with the capability to conduct protein-observed NMR experiments at modest concentrations (< 70 μM) in under ten minutes.[68] As the low-hanging fruit in the realm of drug discovery grows increasingly barren, techniques with greater response are required to enable screening of different libraries. Ligand- and protein-observed NMR techniques rank among the most responsive biophysical assays currently available.[69]

Ligand-observed NMR experiments have been widely used in both academics and the pharmaceutical industry as a rapid screening technique. These methods take two forms, observing the ligands directly, and observing a bound ligand's displacement as a competition experiment (the "spy" molecule approach).[70] The ligand-observed CPMG method detects ligand-binding interactions as a function of drop in signal intensity (Figure 1.10). While any nucleus present in the molecules in the library can theoretically be observed, this most often takes the form of ^{19}F CPMG or ^{1}H CPMG. Fluorine CPMG is widely used and effective, but a disadvantage is the requirement that fluorine be present in the ligands used for the screen, reducing the span of chemical space that can be sampled. Proton CPMG can be used to overcome this shortfall, with the disadvantage that fragment molecules will usually have many more protons than fluorine nuclei, increasing the difficulty of data interpretation and likelihood of ligand signals overlapping. Due to the hyper-responsiveness of the ^{19}F chemical shift to changes in the atom's local chemical environment, in some cases a simple 1-dimensional ligand-observed NMR

experiment is sufficient to assess target engagement. A recent report from Amgen describes a ^{19}F NMR screen against β-secretase that resulted in significant ligand resonance broadening upon interaction with the BACE protein.[71]

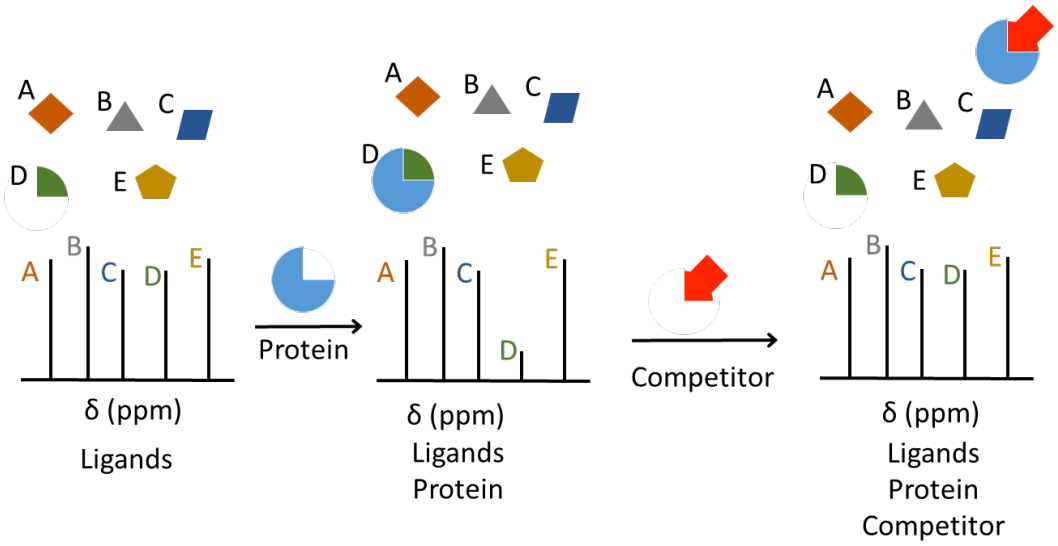


Figure 1.10 A diagram of a CPMG screen.
Because the ligand resonances are observed it is immediately known which ligand is responsible for the binding interaction.

The two other common ligand-observed NMR screening methods are saturation transfer difference (STD), and WaterLOGSY.[72] Both techniques are based on intermolecular nuclear Overhauser effects (NOEs) to transiently bound ^{1}H atoms. WaterLOGSY excites bulk water and transfers the magnetization to bound water, and then to bound ligand. STD selectively excites the protein (usually irradiating upfield of 0 ppm), saturating the bound ligand due to the effective spin diffusion across large molecules (Table 1.2). These methods also allow for epitope mapping, as ^{1}H resonances from atoms that interact with the ligand should be effected more than those farther away.[73] A distinct advantage of STD and WaterLOGSY over CPMG is that the first two methods can be accomplished in a single NMR tube with no follow-up experiments. With CPMG, after the initial spectrum is collected protein must then be added, and a subsequent spectrum acquired. CPMG tends to be more sensitive, but the additional steps slow the time of screening and adds additional labor time as liquid handling capabilities for the protein addition are less common.

Protein-observed NMR methods have become valued for the ability to quantify binding interactions over a wide range of affinities, including particularly weak ligands. While protein-observed ^{1}H-^{15}N HSQC methods are limited by protein size (generally < 40 kDa),[74] pulse sequences such as TROSY have increased this limit to greater than 100 kDa when using high-field magnets,[75] although a deuterated protein is typically employed. With protein-observed NMR techniques a perturbation of protein NMR resonance is observed between the bound and unbound state, thus allowing a dissociation constant to be determined using protein-observed methods if the interaction exhibits a fast chemical exchange rate, commonly found for low affinity ligands. Protein-based methods using labeled amides (e.g., $^{1}H-^{15}N$ HSQC) provide additional structural information for developing small molecules; however, the experiment can be material-intensive and time-consuming. A valuable strength of protein-observed NMR is the ability to obtain accurate dissociation constants of low-affinity ligands by ligand titration while monitoring the dose dependent perturbation of protein resonances. Additionally, the chemical shifts of the protein resonances allow for additional structural assessments. In this context, well-dispersed resonances provide an indication that the protein is in solution and not aggregated. The location of the resonances can offer a "fingerprint" of the protein, and substantial resonance deviation from this fingerprint can indicate large conformational changes. This serves to act as an in situ quality control for the protein NMR assays, reducing false positives and negatives.

Structure-activity relationship (SAR) by NMR is a method for monitoring binding interactions utilizing NMR techniques. Traditional SAR by NMR involves the 2D technique heteronuclear single quantum correlation (HSQC) to monitor the changes in chemical shifts of the amide backbone of proteins, and was initially described in the seminal SAR by NMR paper from Abbott Laboratories.[67] When the chemical environment of the amides in a protein change, different resonances result from the amides, producing different 2D NMR spectra and providing binding information. Using a least squares analysis determines the degree of difference between the two spectra, presumably bound and unbound. Additional information can be gleaned from the spectra by assigning each resonance, and then determining specifically which amides within the

protein are experiencing a difference in chemical environment upon binding. This method will help determine how the protein changes as a function of binding, but does not unambiguously identify the location of the binding site. The amide resonances can shift due to a nearby binding interaction, but can also shift if that binding interaction is inducing a change in the conformation of the protein away from the binding site. A disadvantage of SAR by NMR is that it is slow, reducing its utility in screening techniques.[66]

Table 1.2 A table comparing the different NMR screening methods.

Method	**Parameter observed**	**Physical cause**
Protein-observed NMR	Chemical shift perturbation	Perturbation to local chemical environment
CPMG	Decrease in bound ligand resonance intensity	Ligand exhibiting rotational correlation time corresponding to the protein
STD	Ligand resonance intensity correlates with distance to the protein surface	NOE transferred from protein to bound ligand
WaterLOGSY	Phase inversion of bound ligand resonances	NOE transferred from bulk water to protein-bound water to bound ligand

1.4.2 Surface plasmon resonance

Surface plasmon resonance (SPR) is the other most widely used technique for fragment screening.[76] SPR requires immobilization of the target protein to a sensor chip, often achieved by either bioconjugation reactions or by affinity tags. Because the protein is immobilized, very little protein is consumed during an SPR screen. Changes in mass at the interface of the chip are detected by changes in the angle of incidence from plane-

polarized light. While the mass fraction difference used to be limiting, modern SPR instruments have no restriction on ligand size for analysis.

SPR techniques are unique in that they allow monitoring of protein-ligand interactions in real-time, and thus SPR is one of the few methods by which binding kinetics can be elucidated. While measuring kinetics in principle allows determination of affinity, in practice kinetics on the order of the rate of diffusion (1×10^8 M/s) are challenging to detect. Due to this limitation, a titration experiment will often be used with fragments to determine a dissociation constant.

1.4.3 X-ray crystallography

Though x-ray crystallography is primarily viewed as a technique for structure determination, the high solubility of fragments frequently enables a soaking system to generate co-crystals of protein and ligand. Fragment screening by x-ray crystallography was pioneered by Astex Pharmaceuticals.[77] A key advantage of crystallographic screening is that necessarily a crystal structure is obtained, allowing rational design for further fragment elaborations as well as verification of the binding site. However, x-ray crystallography provides no indication of binding affinity, substantially reducing the ability to triage actives. The active rate of crystallographic screening is also limited to proteins that crystallize readily, and will frequently have lower active rates than other biophysical techniques. However, while crystallographic screening frequently results in false negatives, it is exceptional at avoiding false positives because the presence of electron density corresponding to the ligand unambiguously verifies that the ligand bound in an organized manner.

1.4.4 Thermal shift assays for fragment screening

Thermal shift assays take advantage of the principle that favorable binding interactions tend to require greater energy to disrupt. The temperature at which a protein denatures will be increased due to the favorable binding interaction, allowing rank-ordering of ligands by the degree to which they stabilize the protein. While differential scanning calorimetry (DSC) offers a label-free technique to assess melting temperature as

well as derive thermodynamic parameters, the throughput of DSC precludes it from effective fragment screening. Differential scanning fluorimetry (DSF) dramatically increases the throughput of the assay, allowing thousands of compounds to be screened in a day with minimal protein consumption. DSF involves the incorporation of a fluorescent dye such as Sypro-Orange. An effective DSF dye will undergo fluorescence quenching in the presence of water, and also bind to the accessible hydrophobic patches of a denatured protein. When the protein is thermally denatured, the fluorescent dye will interact with the protein and consequently be less solvated by water, which will reduce the fluorescence quenching and result in an increase in fluorescence.

Despite being a label-free technique, fluorescent dyes have been reported to interact with proteins, ligands, and surfactants.[78] Dye interaction with the protein can stabilize either the folded or thermally-denatured form, potentially shifting the melting temperature in either direction. Dye interaction with the ligand will usually preclude ligand binding, and interaction with surfactant will result in very high background signal. Additionally, a given ligand can interact with the thermally-denatured protein, resulting in a decrease in the melting temperature. All of these phenomenon make DSF very prone to false-positives and false-negatives. Rank-ordering by DSF can also be problematic. The degree of binding is determined at the melting temperature of the protein, usually higher than relevant biological temperatures. The entropic nature of the binding interaction also plays a greater role in melting temperature than the enthalpic portion.[79] Nonetheless, for certain classes of proteins DSF works very well and effectively rank-orders ligands, and has the highest throughput of the modern fragment screening techniques.

1.4.5 Microscale thermophoresis

Microscale thermophoresis (MST) is an emerging fragment screening technique, first being described to assess ligand binding in 2012.[80] The movement of molecules in a temperature gradient is dependent on size, charge, and solvation entropy. Microscale thermophoresis exploits this phenomenon by heating a small portion of the sample with an infrared laser, and measure the rate of diffusion by monitoring fluorescence. The

fluorescent signal can either be from the intrinsic fluorescence of one or more tryptophan residues, or from an attached fluorophore. Because the primary signal is fluorescence, very low concentrations of protein can be used (< 1 μM). Due to the low concentrations allowed, problems arising due to solubility or aggregation of the protein or ligand can generally be avoided with MST. Accurate dissociation constants can be rapidly determined with a serial dilution of ligand, and aggregation of either ligand or protein can be readily observed in the thermophoresis signal.[81]

1.5 **Protein-observed Fluorine NMR**

Both the ligand and drug discovery fields have been increasingly taking advantage of ^{19}F NMR as a biophysical technique for exploring perturbations to protein structure as well as characterizing protein-ligand and protein-protein interactions. The organic fluorine has been eloquently entitled the "odd man out" from the rest of the organic compounds.[82] While the most abundant halogen in the earth's crust, fluorine is rarely found in biological systems. The fluoride ion has a very high redox potential, and many fluorine-containing minerals are insoluble in water, reducing the ability of fluorine to be present in living organisms.[83] It is estimated that there are between six[84] and one-hundred[83] known natural products that contain fluorine. Because of this, fluorine is an ideal atom to use as an NMR probe when studying biomolecules because there are no background signals from the biologics under study, and common buffer additives such as detergents generally do not contain fluorine.

Nuclear magnetic resonance of fluorine nuclei offers a highly sensitive method for observing changes in chemical environment due to a binding interaction. The NMR active ^{19}F atom is the only stable isotope of fluorine, simplifying incorporation because low abundance isotopes are not necessary for this NMR analysis. Additionally, fluorine has the third highest gyromagnetic ratio of the NMR active nuclei, exceeded only by proton and tritium (Table 1.3). Fluorine signals appear as singlets when a proton decoupling sequence is utilized, so if need be, this can be utilized to further increase signal intensity when working with small molecules.[70] These factors, coupled with the newly acquired cryo-prodigy probe allow for rapid (less than 5 minutes) acquisition of

PrOF NMR spectra (Figure 1.11). Because the method of observing binding interactions or conformational changes is generally based on chemical shift, a reduced signal to noise ratio is required compared to integration based analysis. This is explained further in Chapter 5.

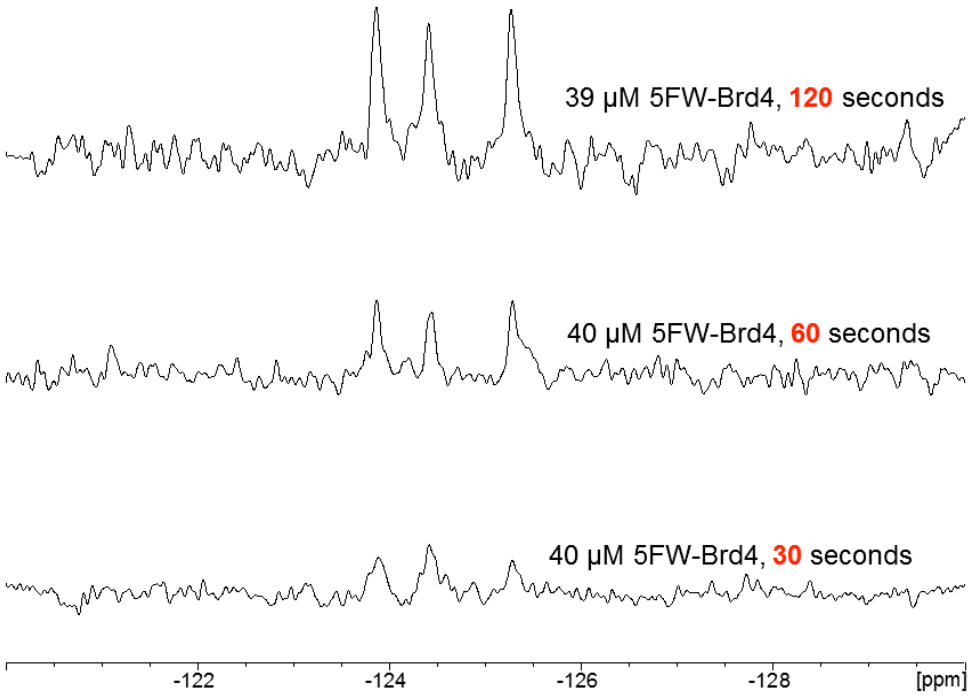


Figure 1.11 ^{19}F NMR spectra of 5FW-Brd4 taken at various experiment times.

Table 1.3 Properties of NMR Nuclei[85]

Nucleus	Gyromagnetic Ratio (MHz/T)	Natural Abundance
^{3}H	45.557	trace
^{1}H	42.577	99.98%
^{19}F	40.053	100%
^{31}P	17.235	100%
^{13}C	10.705	1.11%
^{17}O	5.772	0.0373%
^{15}N	-4.316	0.37%

The resonances of fluorine are highly sensitive to changes in chemical environment, with a range of chemical shifts spanning over 300 ppm.[85] Such a large range results in a low probability of overlapping peaks, simplifying interpretation of NMR spectra. Additionally, the resonances corresponding to fluorinated tryptophan, tyrosine, and phenylalanine, are all predominantly located in different regions of the spectrum, allowing labeling with multiple aromatic amino acid types without concern for overlapping peaks (Figure 1.12).[86]

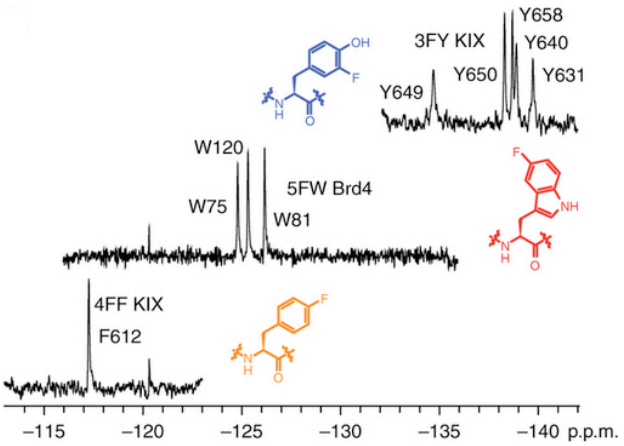


Figure 1.12 PrOF NMR spectra of proteins containing fluorinated amino acids. The ^{19}F spectra of the different aromatic amino acids appear in distinct regions and do not overlap.

PrOF NMR has been used for over 40 years to assess protein conformational perturbations, beginning with the seminal work on alkaline phosphatase,[87] and even being expanded to the 380 kDa ATPase[88]. Because the side-chains are labeled with fluorine, proteins of very large size or restricted mobility are still able to be assessed by PrOF NMR. Labeling of the β_2-adrenergic receptor, a membrane bound protein, with 2,2,2-trifluoroethylcysteine disulfide resulted in resolvable resonances that allowed studying of conformational changes.[89] Despite the responsiveness of fluorine-labeled proteins, until very recently PrOF NMR was not utilized for ligand discovery. Only the recent developments of cryo-probe technology have enabled sufficiently rapid ^{19}F NMR experiments to allow for screening by this method.

The chemical shifts of fluorine nuclei resonances are primarily controlled by lone-pair electrons, resulting in a large paramagnetic term in the shielding equation.[90] Thus, fluorine chemical shifts are particularly sensitive to changes in the local van der Waals environment, as well as local electrostatic fields, with surface exposed fluorine atoms on proteins being able to detect the differences between H_2O and D_2O with chemical shift perturbations up to 0.2 ppm.[86] This sensitivity to local electrostatic environments contributes largely to the fluorine atom's utility for detecting changes in biomolecules.[85] The range of chemical shifts for ^{19}F nuclei is ~100 times larger than ^{1}H nuclei, allowing it to be used to monitor weak binding, folding, enzyme kinetics, and conformational

change, and consequently the related thermodynamic properties.[91–93] Chemical shift agents can further broaden these chemical shift differences, separating surface exposed residues from those on the interior of the protein.

The tools to fluorinate proteins are readily available in most biochemistry labs. At the most basic level, the lab needs to be set up for standard protein *Escherichia coli* (*E. coli*) expression. To express proteins with fluorinated tyrosine or phenylalanine, the aromatic amino acid biosynthesis pathway must be inhibited, or the bacteria must be auxotrophic for the desired amino acids. DL39 cells are auxotrophic for phenylalanine and tyrosine, and so are an excellent choice for fluorophenylalanine and fluorotyrosine incorporation, and are commonly used in labs for *E. coli* work. However, since DL39 cells are not auxotrophic for tryptophan the W3110TrpA33 cell line can be used to instead to incorporate fluorinated tryptophan.[94,95] Other common methods for fluorine incorporation involve inhibiting the aromatic amino acid biosynthesis pathway using glyphosate, a somewhat expensive reagent, but a new method reported by Crowley *et al.* exploits the aromatic amino acid biosynthesis pathway, utilizing fluorinated precursors to produce fluorinated tryptophan. Consequently, in this method the bacterial strain must not be auxotrophic for aromatic amino acids.[96] Instead, Crowley utilizes BL21 cells and adds 5-fluoroindole, a precursor to tryptophan. Addition of 5-fluoroindole to the media used for production of BL21 cells, with removal of natural tryptophan, results in incorporation of the bacterial produced 5-fluorotryptophan into the protein of interest.

PrOF NMR offers a complimentary technique to currently existing techniques for assessing protein binding interactions. While PrOF NMR does not provide the same information regarding the overall structure of a protein that ^{1}H-^{15}N HSQC NMR can, PrOF NMR has the advantage of being simple to interpret because of the low number of resonances, while also placing the chemical probe very close to the binding interaction. ^{1}H-^{15}N HSQC detects the amide backbone of proteins, while PrOF NMR detects fluorine resonances that are of the side chains of the amino acids, much closer to the binding interaction. Because the dominant isotope of fluorine is NMR active, it is much cheaper to incorporate than the rarer ^{13}C or ^{15}N nuclei. ^{13}C labeled tyrosine is ~$850 for 100 mg from Sigma-Aldrich, and ^{13}C labeled tryptophan is ~$2000 per 100 mg. By comparison,

3-flurotyrosine is ~$84 for 500 mg and 5-fluoroindole is $50 for 1 g. Substituting a fluorine atom for a hydrogen atom results in a compound that is sterically similar, with fluorine atoms having a van der Waal's radius of 1.47 Å while hydrogen is 1.20 Å.[97] This offers a substantial advantage over many labeling technologies, such as fluorophore labeling that greatly increase the size of the ligand.

Fluorine nuclear magnetic resonance spectroscopy (^{19}F NMR) offers a hyper-responsive method to analyze protein-ligand interactions and is ideally suited to develop these probes. Chemical shifts of fluorine nuclei resonances are particularly sensitive to changes in chemical environment due to their large number of unpaired electrons, and fluorine has a very high gyromagnetic ratio, allowing for rapid acquisition of spectra.[85] In previous studies, PrOF NMR has been utilized for kinetics experiments, acquiring a ^{19}F NMR spectrum in under two seconds.[98] Protein-observed fluorine (PrOF) NMR provides information regarding binding interactions as a function of chemical shift.[99] Aromatic amino acids are enriched at protein-protein interfaces,[100,48] and therefore fluorinated aromatic amino acids offer an efficient way to monitor binding interactions as a function of chemical shift of the fluorine resonances. PrOF NMR is an analysis technique that can offer great detection of binding interactions because the observed resonances come from atoms in close proximity to binding interactions, but other techniques, including FP, ITC, and HSQC, are necessary in order to fully understand protein-ligand binding interactions.

1.6 **Biophysical techniques to assess structure and function**

Fluorescence anisotropy is affected by the rotational correlation of fluorophores in solution. The fluorescence anisotropy is measured by irradiating the solution with polarized light, and measuring the anisotropy of the returning fluoresced light. The rotational correlation of molecules in solution is proportional to the size of the molecule: the larger the molecule is, the more slowly it tumbles, as indicated by a long rotation correlation time. FP exploits this relationship by attaching a fluorophore to a small molecule that binds to the protein of interest, which results in high anisotropy in the bound state and low anisotropy when free in solution (Figure 1.13). Following the change

in fluorescence anisotropy as a function of concentration of either the protein or the fluorophore allows calculation of the direct dissociation constant (K_d). A disadvantage of this technique is that the binding ligand is modified by incorporation of the fluorophore, and so measurements of the K_d of the fluorescent ligand may be modified as a result of fluorophore incorporation.[101]

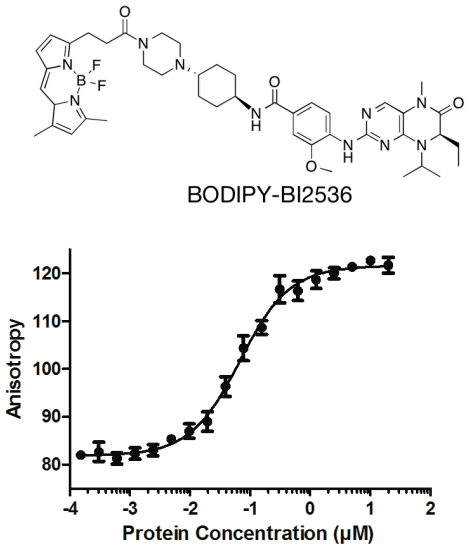


Figure 1.13 Fluorescence anisotropy direct binding with Brd4.
Direct binding between BODIPY-BI2536 and varying concentrations of Brd4 yields a K_i of 50 ± 11 nM.

$$y = c + (b - c)\frac{(Kd + a + x) - \sqrt{(Kd + a + x)^2 - 4ax}}{2a}$$

K_i values were obtained using a variant of the Cheng-Prussof equation from Huang et al.[102]

A common method to avoid only measuring the binding constant of the modified fluorophore-containing ligand is to perform a competition assay. In this assay, the fluorophore containing ligand is already bound to the protein of interest (ideally only at the binding site of interest). The ligand being studied (the one without a fluorophore) is then titrated into the mixture, and competes off the fluorophore containing ligand. The decrease in fluorescence anisotropy is measured as a function of protein, ligand, and fluorophore containing ligand concentration to obtain the IC_{50}, which can be used to calculate the indirect dissociation constant (K_I). In addition to measuring the binding

constant of a ligand unmodified by incorporation of the fluorophore, if the fluorophore containing ligand only binds at the binding-site under study, then it can be determined whether or not the ligand under study binds to the desired binding site or another site on the protein.[101]

Isothermal titration calorimetry (ITC) allows very accurate determination of binding energies in addition to binding stoichiometry and the thermodynamic properties of binding (Figure 1.14). This method works by gradually titrating in a known amount of protein or ligand, and measuring the change in heat as a result of the reaction. The change in heat is measured indirectly, back calculated as a function of the amount of energy required to maintain the solution at a constant temperature.[103] A disadvantage of ITC is that it provides no structural reference to where the binding event is occurring, and not fast enough to be used as an effective screening technique.

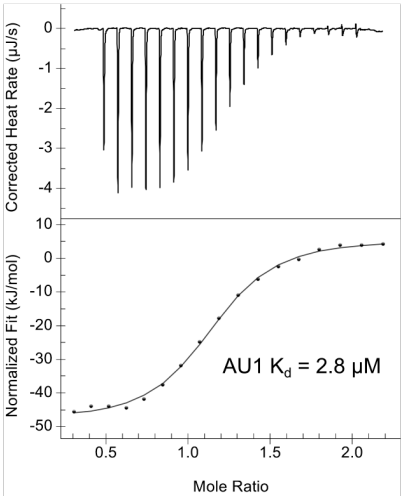


Figure 1.14 An example ITC thermogram of BPTF with AU1 (discussed in chapter 4).

1.7 **Preface to this Dissertation**

The follow chapters describe my role in the development of protein-observed ^{19}F NMR as a ligand discovery and screening technique. Chapter 2 describes the initial application of PrOF NMR to bromodomains, with its sensitivity highlighted by uncovering the previously unknown Brd4-acetaminophen interaction and BPTF-BI2536

interactions. Chapter 3 is a disclosed protocol to disseminate the technique of making fluorinated protein and subsequent NMR analysis, including a brief discussion of using paramagnetic relaxation agents to reduce NMR experiment time by ~50%. Chapter 4 entails the use of a dual PrOF NMR screen to uncover isoform selective binders between BPTF and Brd4, resulting the discovery of the first molecule selective for BPTF over Brd4. Chapter 5 details a collaboration with Lilly Research Laboratories to compare PrOF NMR with the industry standard ^{1}H CPMG ligand observed NMR screening technology. Chapter 6 describes the results of a collaboration with William Isley to develop a computational method to assign PrOF NMR resonances.

Chapter 2. Fluorinated aromatic amino acids are sensitive ^{19}F NMR probes for bromodomain-ligand interactions

Motivation: The purpose of this research was to use the bromodomain of Brd4 as a representative protein of this structural class of epigenetic "reader" proteins to validate the effectiveness of protein-observed fluorine NMR for characterizing protein-ligand interactions. Brd4 represented a good model system as it had structurally characterized small-molecule inhibitors at the time we began this work, so good positive controls were readily available as well as detailed structural biology data.

2.1 **Introduction**

Bromodomains are epigenetic "reader" proteins that were the first structural modules for recognizing the acetylated ε–nitrogen of lysine on histones and play essential roles in diverse diseases, particularly in cancer.[15] Structural biology efforts in NMR[104] and particularly x-ray crystallography have accelerated chemical probe development with more than 150 x-ray structures of 43 different bromodomains reported.[34] The BET (Bromodomain and extraterminal) subfamily of proteins, Brd2, 3, 4, and T are the most intensely studied due to the discovery of tool compounds including pan-BET inhibitors such as (+)-JQ1, first used to validate Brd4 inhibition for treating NUT midline carcinoma (Figure 2.1).[13] Selective targeting of BET bromodomains is one of the most significant challenges for this emerging therapeutic class.[15] Discovery of new binding modes can yield insight into selective inhibitor designs. While BET research is developing rapidly, the majority of the remaining bromodomains are not as well-characterized due to a lack of both small molecule inhibitors and optimal screening methods.[52] Here, we describe a new ligand discovery and characterization method for

bromodomains using ^{19}F NMR, and highlight the value of fluorine labeling aromatic side-chains for selectivity and binding analyses.

Protein-observed NMR has proven to be a powerful ligand discovery method since the original reports from Abbott labs.[67] In these experiments, a protein is labeled with NMR active nuclei, typically on the backbone amides. Small molecule induced changes in protein chemical shifts are used to characterize a binding event. This method can be used to elucidate the binding site, discriminate specific from non-specific interactions, and quantify weak affinities of low molecular weight compounds (MW < 300) termed fragments, commonly used for NMR screening. Fragments can be missed in the current ligand discovery methods for bromodomains.[33] This is particularly challenging in the absence of an initial ligand for competition studies, which is the case for many bromodomains.[52] However, in amide detected, protein-observed NMR, the need for high concentration of protein, detailed resonance assignments, and speed of experiment can be limiting.[105]

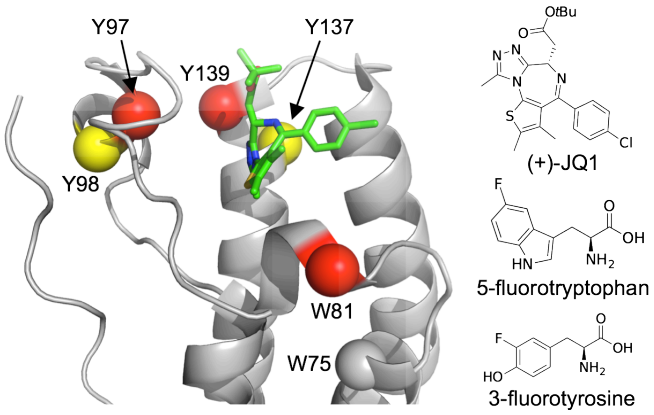


Figure 2.1 Aromatic amino acids are close to the bromodomain binding site.
Ribbon diagram of Brd4(1) bound to (+)-JQ1 (PDB ID 3MXF). The α-carbons of tyrosine and tryptophan are indicated as spheres. Y97, Y139, and W81 are within 5 Å and Y98 and Y137 are within 10 Å. (Right) (+)-JQ1 and fluorinated amino acids used in this study.

We described the use of protein-observed ^{19}F NMR (PrOF NMR) for fragment screening using fluorinated aromatic amino acid side chains demonstrated by a 50 compound fragment pilot screen on a non-bromodomain containing protein.[99] Here we first test the PrOF NMR approach on the N-terminal bromodomain of Brd4 (Brd4(1)),

using known ligands and include a small molecule whose binding affinity was too weak to characterize by prior methods. Sequence analyses of 61 bromodomains identified The conservation of aromatic amino acids, many positioned near the histone binding site.[12]

Aromatic amino acid analysis across all 61 Bromodomains Relative to Brd4(1)

Table 2.1 Conservation of aromatic amino acids within bromodomains at the same site as Brd4(1).
Additional aromatic amino acids are all present in bromodomains but found at sites distinct from Brd4(1). (e.g., see the BPTF sequence below). Data was compiled from the bromodomain sequence alignments from Filippakopoulos, P et al. *Cell* 2012, *149,* 214-231. Residues near the histone binding site are highlighted in red.

Brd4(1) Sequence	# Tyr	# Trp	# Phe	Sum	Percent
Y65	6		2	8	13
Y97	59			59	97
Y98	29		2	31	51
Y118	9		3	12	20
Y119	59			59	97
Y137	8	2	1	11	18
Y139	40		16	56	92
W75	10	6	4	20	33
W81		18	1	19	31
W120		3		3	5
F79	6		7	13	21
F83			57	57	93
F129			13	13	21
F133			37	37	61
F157	2		36	38	62

```
BRD2(1)  97 WPFRQPV...LPDYHKIIKQPMDMGT..ENNYYWAASECMQDFNTMFTNCYIYNK
BRD2(2) 370WPFYKPV...LHDYHDIIKHPMDLST..ENRDYRDAQEFAADVRLMFSNCYKYNP
BRD3(1)  57 WPFYQPV...LPDYHKIIKNPMDMCT..ENNYYWSASECMQDFNTMFTNCYIYNK
BRD3(2) 332WPFYKPV...LHDYHDIIKHPMDLST..DGREYPDAQGFAADVRLMFSNCYKYNP
BRD4(1)  81 WPFQQPV...LPDYYKIIKTPMDMGT..ENNYYWNAQECIQDFNTMFTNCYIYNK
BRD4(2) 374WPFYKPV...LHDYCDIIKHPMDMST..EAREYRDAQEFGADVRLMFSNCYKYNP
BRDT(1)  50 WPFQRPV...LPDYYTIIKNPMDLNT..ENKYYAKASECIEDFNTMFSNCYLYNK
BRDT(2) 293WPFYNPV...LHNYYDVVKNPMDLGT..DNQEYKDAYKFAADVRLMFMNCYKYNP
BPTF       WPFLEPV...APDYYGVIKEPMDLAT..QRRYYEKLTEFVADMTKIFDNCRYYNP
```

For new selectivity studies, we apply this method to a second BET bromodomain BrdT, as well as a non-BET BPTF, for which no small molecule probes are known. The hyper-responsiveness of the fluorine-chemical shift to ligand binding and simplicity of

the rapidly acquired NMR spectrum,[99,106] provides the speed and structural information necessary for developing potential inhibitors. These studies lead us to propose PrOF NMR as a structure-based tool for characterizing new and native ligands for the aromatic-rich bromodomains.

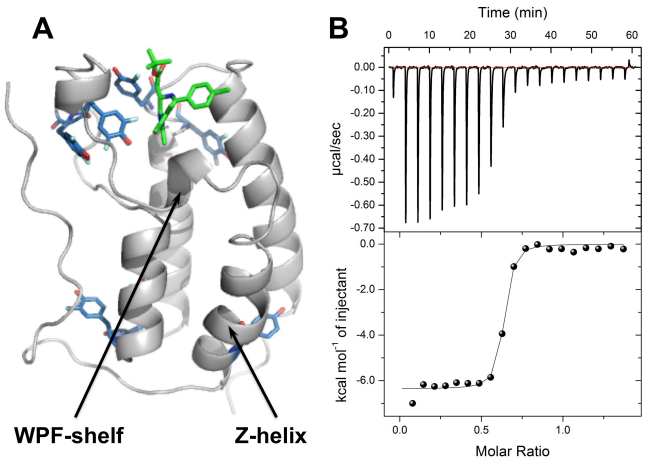


Figure 2.2 Structure and function of 3FY-Brd4(1).
A) Cartoon representation of a 3FY-labeled Brd4(1) crystal structure complexed with (+)-JQ1. B) Isothermal titration calorimetric analysis of binding of (+) JQ1 binding to 3FY-Brd4(1) at 25 ^{o}C.

2.2 **Results and Discussion**

We analyzed the crystal structure of Brd4(1) in complex with inhibitor (+)-JQ1 (Figure 2.1) to determine a fluorine-labeling strategy. The aromatic side chains of Y97, Y139 and W81 are within 3-5 Å of (+)-JQ1. Y98 and Y137 are within 10 Å of the thiophene ring of (+)-JQ1. Although W75 is not in close contact, it is located on the dynamic Z helix of Brd4(1) (Figure 2.2). Fewer phenylalanine side chains were in close proximity. For these reasons, we decided to investigate the ligand detection ability of PrOF NMR using either 3-fluorotyrosine (3FY) or 5-fluorotryptophan (5FW)-labeled Brd4(1). High labeling efficiency with 3FY and 5FW was achieved (78-90%) with good to high protein yields (10-60 mg/L).

We used circular dichroism (CD), isothermal titration calorimetry (ITC), and x-ray crystallography to assess the structural and functional effects on fluorine

incorporation into Brd4(1). Far UV CD spectra of the fluorinated proteins shows similar levels of secondary structure but a slightly lower thermal stability (T_m = 50 °C vs 52 °C) (Figure 2.4, Table 2.2). We used ITC to assess how these effects may alter binding. (+)-JQ1 binds to Brd4(1) with a K_d of 49 nM at 15 °C.[13] We detect a similar binding affinity of 75 ± 4 nM for our Brd4(1) and 88 ± 6 nM for 3FY-Brd4(1) (Figure 2.2, Figure 2.3) at 25 °C. We solved the crystal structure of 3FY-Brd4(1) bound to (+)-JQ1. Aligning this structure onto the unlabeled protein complex with (+)-JQ1 yielded an RMSD of 0.089 Å (Figure 2.2). These results lead us to conclude that fluorine incorporation has a minimal effect on structure and function. Crystallization of 3FY-Brd4(1) is itself significant as less than 20 fluorinated protein crystal structures have been reported.

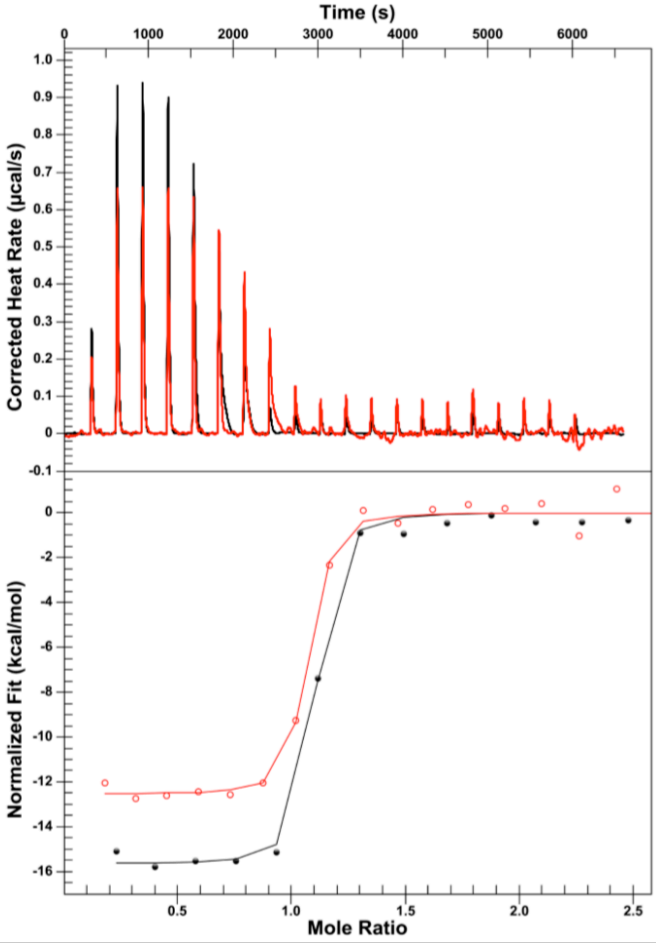


Figure 2.3 ITC Binding isotherm of (+)-JQ1 with unlabeled Brd4(1) and 5FW Brd4. Isothermal titration calorimetric analysis of binding of (+)-JQ1 to unlabeled (filled circle) and 5FW labeled (open circle) Brd4(1) at 25 C.

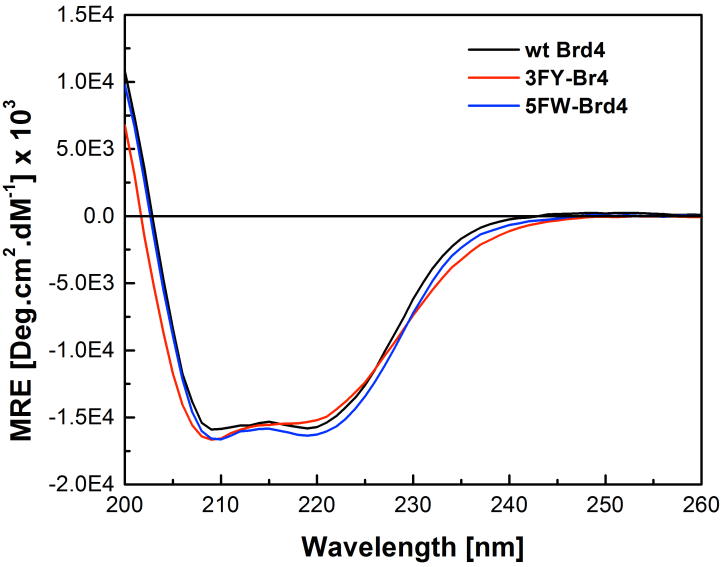


Figure 2.4 Circular dishcorism of Brd4, 3FY-Brd4, and 5FW-Brd4.
Far-UV spectra of unlabelled, 3FY and 5FW labelled Brd4(1). Each spectrum was measured at 25 ^{o}C and protein concentration for each spectrum was 20 µM in 50 mM Tris buffer pH 7.4 containing 100 mM NaCl.

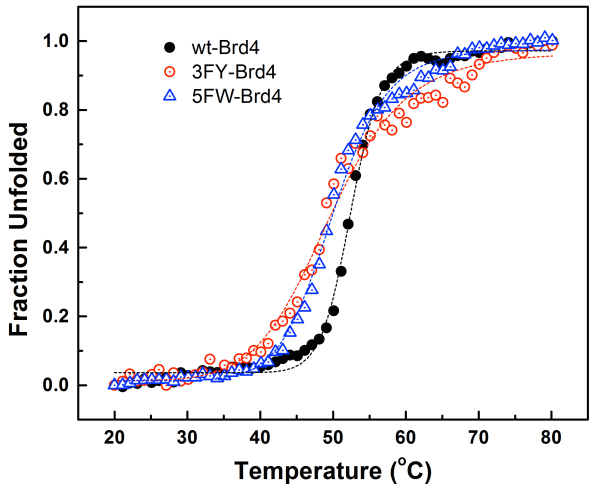


Figure 2.5 Circular dichroism thermal melts of Brd4, 3FY-Brd4, and 5FW-Brd4.
Thermal denaturation curve of unlabeled (filled circle), 3FY labeled (open circle) and 5FW labeled (open triangle). The change in ellipticity at 222 nm was measured with increasing temperature from 20 ^{o}C to 80 ^{o}C. 20 µM protein was used in 50 mM Tris buffer pH 7.4 and 100 mM NaCl.

Table 2.2 CD data of unlabeled and fluorine-labeled Brd4(1)

Protein	T_m (^oC)	MRE at 222 nm
Unlabeled Brd4 (1)	52.40 ± 0.1	-14893
3FY labeled Brd4 (1)	49.57 ± 0.3	-14413
5FW labeled Brd4 (1)	49.90 ± 0.2	-15677

^{19}F NMR spectra of both proteins display well-dispersed resonances indicative of folded proteins. From the ^{19}F NMR spectrum of 5FW-Brd4(1), resonances for W75, W81, and W120 were clearly resolved (Figure 2.9). The seven 3FY resonances in 3FY-Brd4(1) span over 12 ppm suggesting a diverse environment of the aromatic side chains (Figure 2.9). Two of the resonances at -136.6 ppm are partially overlapping. To our knowledge, this is the largest chemical shift range reported for a fluorinated protein but remains to be tested if this range is similar in other bromodomains, which have similar arrangements of aromatic amino acids.

Resonance assignments were made via a combination of site-directed mutagenesis and ligand binding experiments (explained below). For site-directed mutagenesis, a single tryptophan or tyrosine to phenylalanine mutant was expressed and the disappearance of a single resonance was used to assign the side-chain in the parent ^{19}F NMR spectrum. Complete resonance assignments enabled us to test the sensitivity of our fluorine NMR method for characterizing binding footprints and quantifying affinity of weak binding molecules for fragment-based screens (Figure 2.6, Figure 2.7).

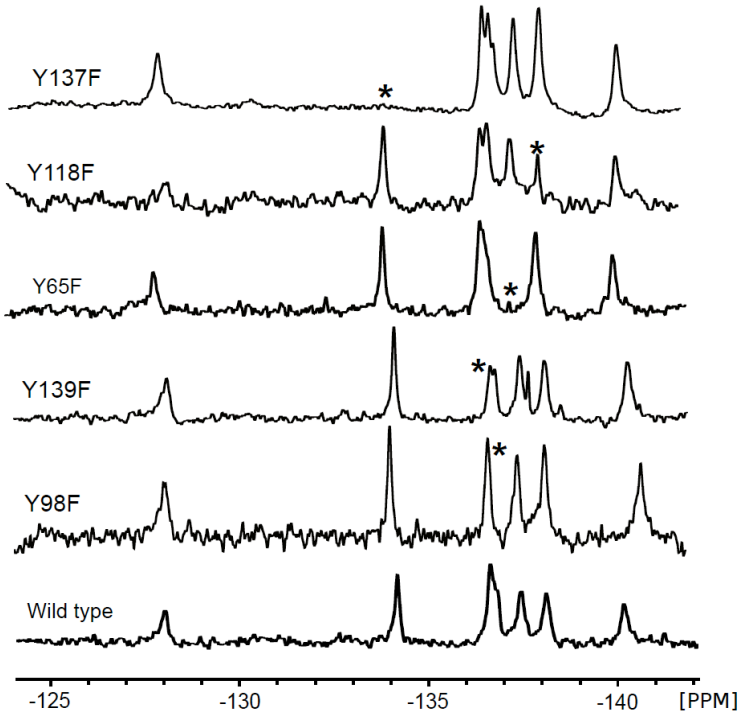


Figure 2.6 Assignment of 3FY tyrosine chemical shifts on 1D ^{19}F NMR spectra.
1D ^{19}F NMR spectra of 3FY labeled wild-type and tyrosine to phenylalanine point mutated Brd4(1). Asterisks represent the position of missing peaks of corresponding mutated tyrosine. The Y118F expressed poorly, and had low stability. In the Y118F spectrum a partial resonance (which integrates to 0.49) grows in obscuring the absence of the resonance at this position. We have identified this resonance as either a degradation product or resonance from aggregation of the protein (Figure 2.8). The two rotamers for Y98, as seen in the Y139F spectrum, aided in the assignment of the shifted Y139 resonance in the acetaminophen binding experiment.

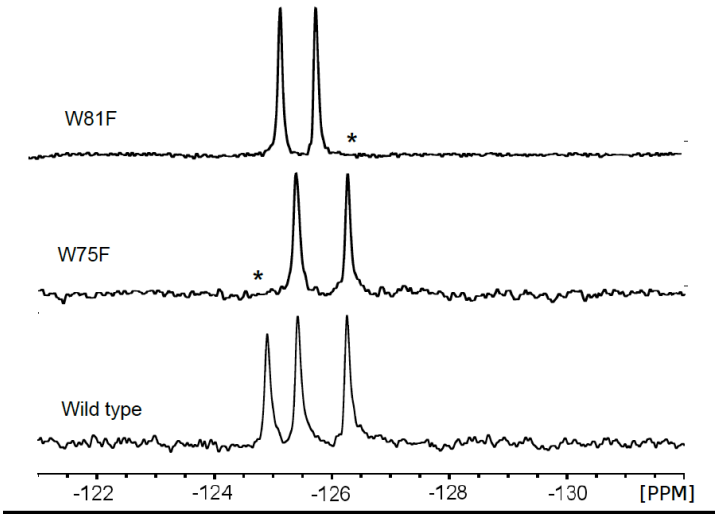


Figure 2.7 Assignment of tryptophan chemical shifts on 1D 19F NMR spectra.
1D ^{19}F NMR spectra of 5FW-labeled wild type and tryptophan to phenylalanine point mutated Brd4(1). Asterisks represent the position of missing peaks corresponding of the mutated tryptophan.

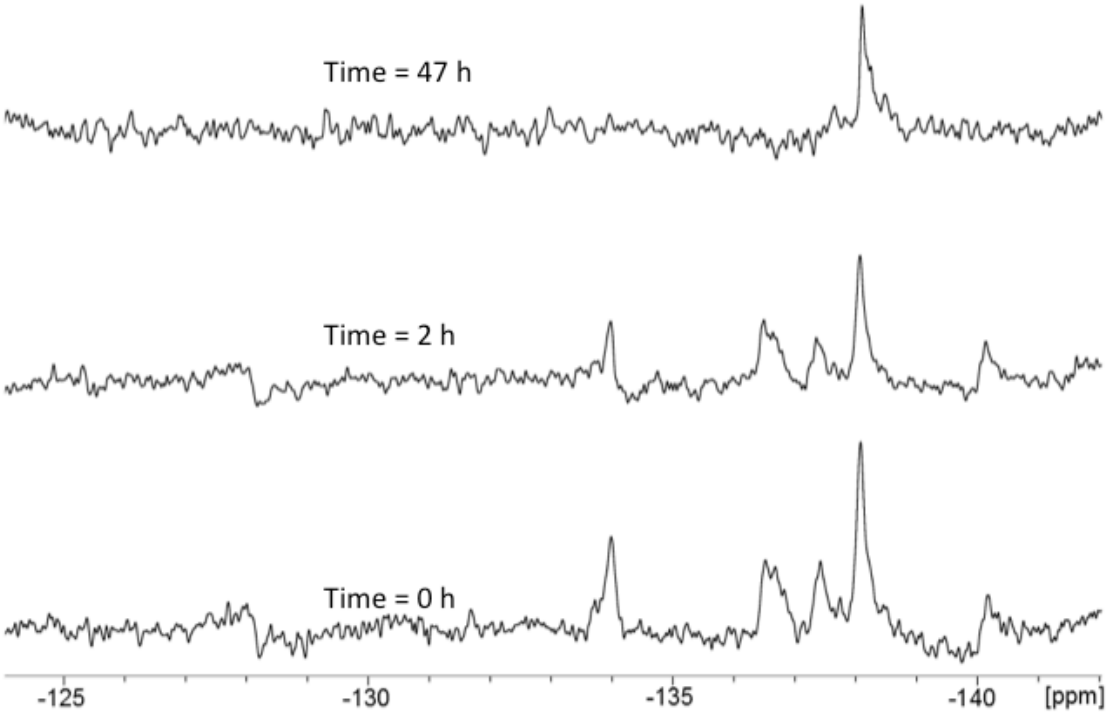


Figure 2.8 Time dependent degradation of a phenylalanine 3FY-Brd4 mutant.
19F NMR spectra collected over time of a 3FY-Brd4 mutant. A new resonance gradually increases in intensity at -138 ppm. The narrowness of the resonance is consistent with degradation of the protein, but may also result from aggregation.

Crystallographic evidence supports the role of W81 for determining the specificity of (+)-JQ1 for binding BET bromodomains in a region termed the WPF shelf.[12,13] W81 was first tested as a diagnostic residue for binding studies. Slow exchange binding was readily detected upon titrating (+)-JQ1, in which case the intensity of the W81 resonance at -126.3 ppm disappeared as a new downfield resonance at -124.0 ppm grew in (Figure 2.9). The observed slow exchange of bound and free protein is consistent with submicromolar binding. Closer inspection of the spectrum revealed a small upfield shift of the W75 resonance located underneath W81 on the Z-helix of Brd4(1), but outside the binding site (Figure 2.9).

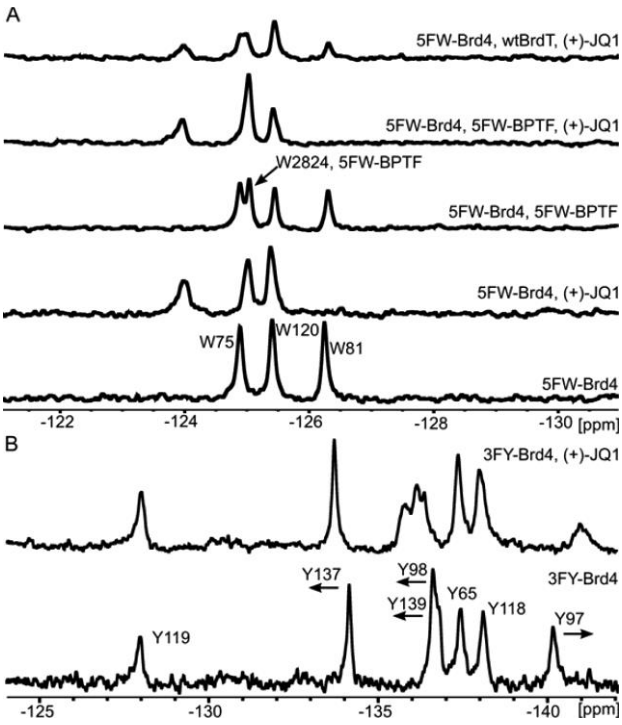


Figure 2.9 PrOF NMR titration of (+)-JQ1 with Brd4(1).
A) Bottom to top: 5FW-Brd4(1) (25 μM); with 1 eq. (+)-JQ1; with 5FW-BPTF (25 μM); with 5FW-BPTF (25 μM) and 1 eq. (+)-JQ1, with unlabeled BrdT(1) (50 μM) and 1 eq. (+)-JQ1. B) 3FY-Brd4(1) (47 μM) titrated with 2 eq. (+)-JQ1.

Titrating (+)-JQ1 with 3FY-Brd4(1) provided further binding information. Y97, Y98, and Y139 are located within the binding site for (+)-JQ1. All three resonances shifted (slow exchange) and broadened upon titration. The most upfield resonance was assigned to Y97, which forms a bridging hydrogen bond to (+)-JQ1 via a structurally conserved water molecule (Figure 2.9). Y137, on the outside of the binding site, moved downfield but remained sharp. The BC loop near Y137 is highly dynamic.[13] Our observations are consistent with a protein conformational change with considerable flexibility at this site.

Fragment screening is a promising method for early stage ligand discovery and is well-suited for NMR analyses due to the ability to characterize protein ligand interactions at the high concentrations needed to detect small molecule binding.[105] Acetaminophen represented an ideal test case as it was shown to bind BET bromodomains via x-ray crystallography, but its affinity for its target was reported as too weak for characterizing

by fluorescence anisotropy or TR-FRET.[33] The small size of this molecule (MW = 160 g/mol) is representative of many compounds in fragment libraries. Using 3FY-Brd4(1), we detect binding as low as (47 μM) with chemical shift changes > 0.6 ppm when fully bound (Figure 2.10, Figure 2.11, Figure 2.12, Table 2.3, Table 2.4). Based on the crystal structure with Brd2 and site directed mutant analysis, we assigned this shifted resonance to Y139. Y97, which is expected to make a water-mediated hydrogen bond, also broadens. Titration of acetaminophen yielded a K_d of 290 ± 40 μM. This results in a high ligand efficiency (0.44 kcal/mol/non-hydrogen atom) for acetaminophen as a bromodomain inhibitor scaffold. Titration with 5FW-Brd4(1) yielded a similar K_d based on the chemical shift perturbation of W81 (230 ± 30 μM). The two K_ds are within error arguing against a specific fluorine perturbing binding. Due to the speed of data collection (chemical shift information can be acquired in < 5 min), low protein concentration (40-50 μM), and conserved aromatic contacts, we anticipate this method will be well-suited for bromodomain fragment screening. This is especially useful where there is a lack of suitable ligands for competition-based experiments or affinities are weak (K_d > 100 μM).[30]

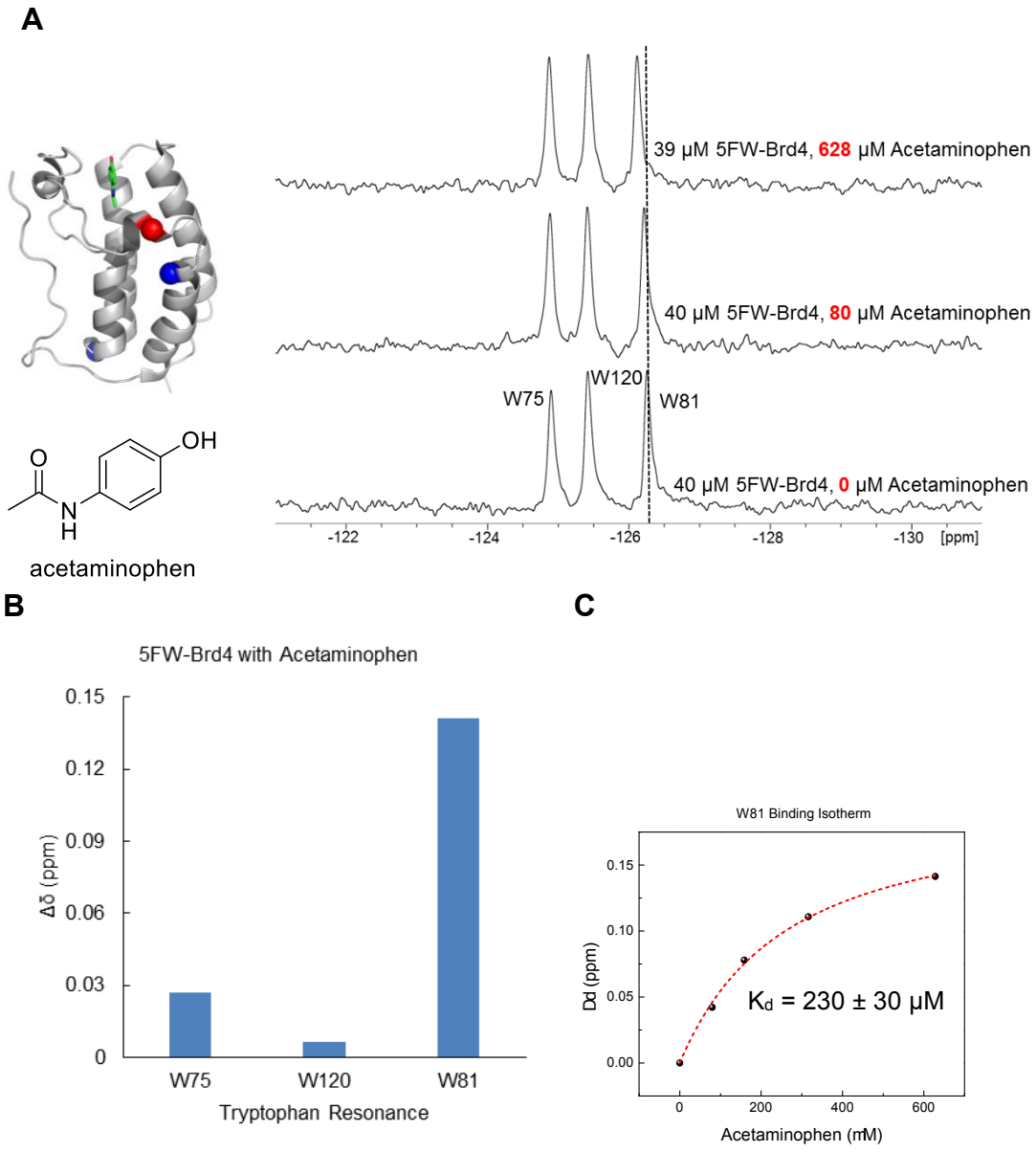


Figure 2.10 ^{19}F NMR spectral analysis of 5FW-Brd4 with acetaminophen.
A) Representative ^{19}F NMR spectra from titration with acetaminophen. Ribbon diagram of acetaminophen bound to Brd2(1) (PDB Code: 4A9J) was aligned to Brd4(1) (PDB Code: 3MXF), Brd2(1) was removed to indicate probable binding mode of acetaminophen to Brd4(1). The α-carbon of tryptophan are indicated as spheres. B) Absolute value of chemical shift perturbation for the 5FW-Brd4 tryptophan resonances at 628 μM acetaminophen. C) A binding isotherm of W81 perturbation for the acetaminophen titration.

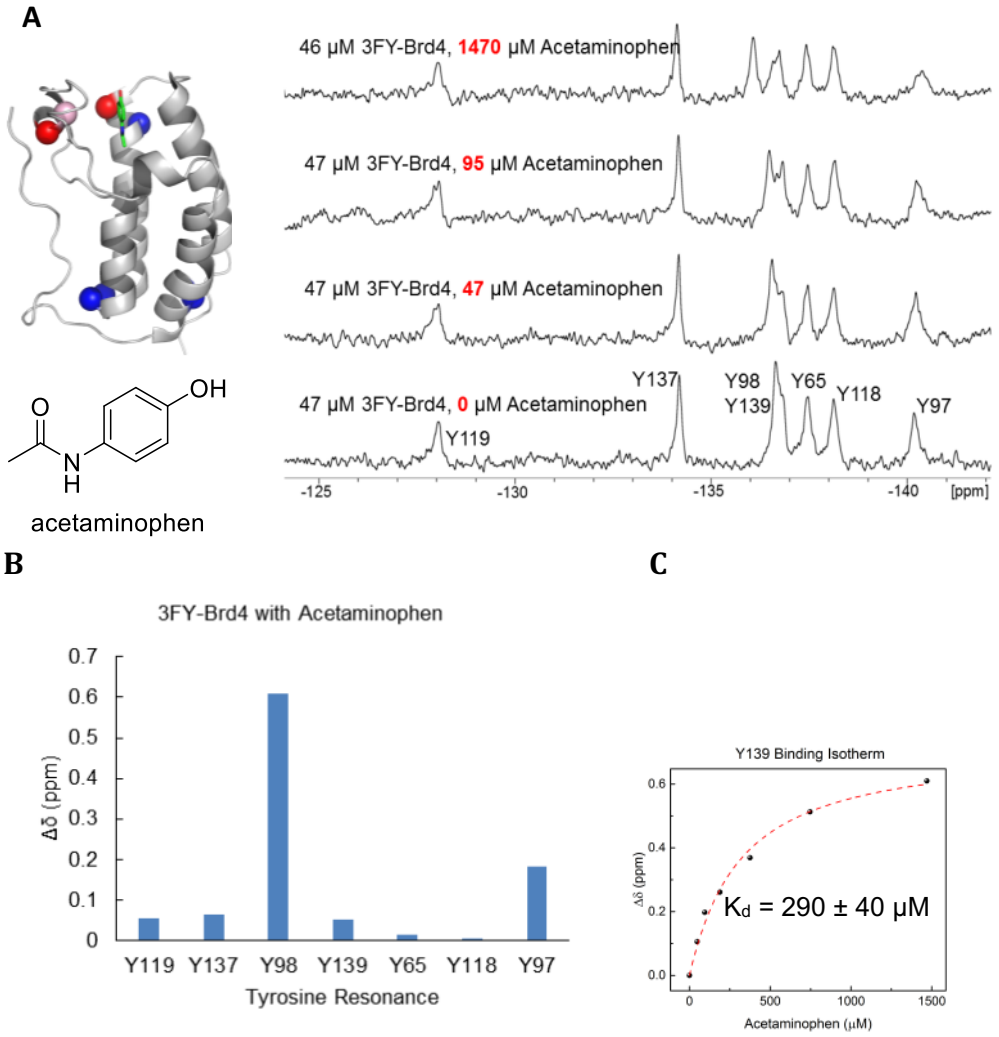


Figure 2.11 ^{19}F NMR spectral analysis of 3FY-Brd4 with acetaminophen.
A) Representative ^{19}F NMR spectra from titration with acetaminophen. Ribbon diagram of acetaminophen bound to Brd2(1) (PDB Code: 4A9J) was aligned to Brd4(1) (PDB Code: 3MXF), Brd2(1) was removed to indicate probable binding mode of acetaminophen to Brd4(1). The α-carbon of tyrosine are indicated as spheres. B) Absolute value of chemical shift perturbation for the 3FY-Brd4 tyrosine resonances at 1470 μM acetaminophen. C) A binding isotherm of Y139 perturbation for the acetaminophen titration.

Table 2.3 5FW-Brd4 19F NMR chemical shift perturbations at different acetaminophen concentrations:

Acetaminophen Concentration	Chemical Shift	W75	W120	W81
0 μM	Chemical shift (ppm)	-124.90	-125.420	-126.259
80 μM	Δδ (ppm)	0.02	0.01	0.04
158 μM	Δδ (ppm)	0.02	0.01	0.07
316 μM	Δδ (ppm)	0.02	0.00	0.11
628 μM	Δδ (ppm)	0.03	-0.01	0.14

Table 2.4 3FY-Brd4 19F NMR chemical shift perturbations at different acetaminophen concentrations:

Ligand Conc.	Chemical Shift	Y119	Y137	Y98/139	Y98/139	Y65	Y118	Y97
0 μM	Chemical shift (ppm)	-128.02	-134.01	-136.60	-136.60	-137.43	-138.05	-140.09
48 μM	Δδ (ppm)	0.04	0.02	0.11	-0.08	-0.01	-0.02	-0.06
95 μM	Δδ (ppm)	0.04	0.04	0.20	-0.05	0.01	-0.02	-0.28
188 μM	Δδ (ppm)	0.09	0.03	0.26	-0.06	-0.01	-0.09	-0.13
375 μM	Δδ (ppm)	0.02	0.04	0.37	-0.02	-0.02	-0.05	-0.15
746 μM	Δδ (ppm)	0.12	0.07	0.51	0.04	0.03	-0.01	-0.22
1470 μM	Δδ (ppm)	0.05	0.06	0.61	0.05	0.01	-0.01	-0.18

BET bromodomains are inhibited by several classes of kinase inhibitors affording molecules with dual modes of actions and a new source of ligand diversity for screening against bromodomains.[30] We chose three inhibitors covering a wide range of IC_{50} values: BI2536 (Figure 2.15, Figure 2.16, Table 2.7, Table 2.8), (0.025 μM) TG101348 (Figure 2.17, Figure 2.18, Table 2.9, Table 2.10), (0.29 μM) and Dinaciclib (Figure 2.13, Figure 2.14, Table 2.5, Table 2.6) (19 μM) to compare their binding footprints to x-ray structures with Brd4 (Figure 2.12).

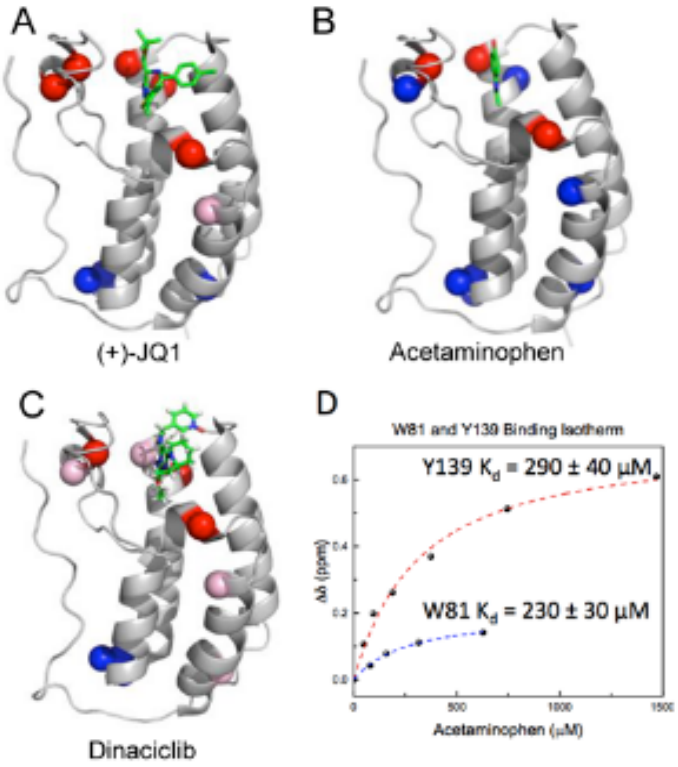


Figure 2.12 Binding footprint and affinity determination with Brd4(1) and various ligands.
A-C) Residues whose fluorinated resonance were, highly perturbed tend to cluster near the binding site. Acetaminophen was overlaid on Brd4(1) based on the crystal structure with Brd2(1) (PDB IDs 3MXF, 4O70, 4A9J) D) Binding isotherms for acetaminophen titration with 3FY-and 5FW-Brd4(1).

During Dinaciclib titration, the resonance for W81 broadened significantly. Perturbation of the fluorine resonance for W75, gave rise to a small chemical shift perturbation in fast exchange. We used this perturbation to estimate a K_d of 60 µM ± 20. BI2536 and TG101348 have nanomolar binding affinity. In both cases W81 was highly perturbed. With BI2536, a new resonance in slow exchange grows in upfield and is consistent with ring current effects based on analysis of the co-crystal structure and the reported nanomolar affinity. Similar to (+)-JQ1 binding, the resonance for W75 showed a small perturbation. W75 shows the most pronounced effect for TG101348 binding. Inspection of crystal structure overlays reveals very minor perturbations to the residues and neighboring side-chains highlighting the sensitivity of ^{19}F NMR for identifying small conformational effects. 3FY-Brd4(1) was used to further characterize the kinase inhibitor

binding modes for Dinaciclib (Figure 2.12) BI2536, and TG101348. The results were consistent with x-ray data and induced conformational effects along the Z-helix.

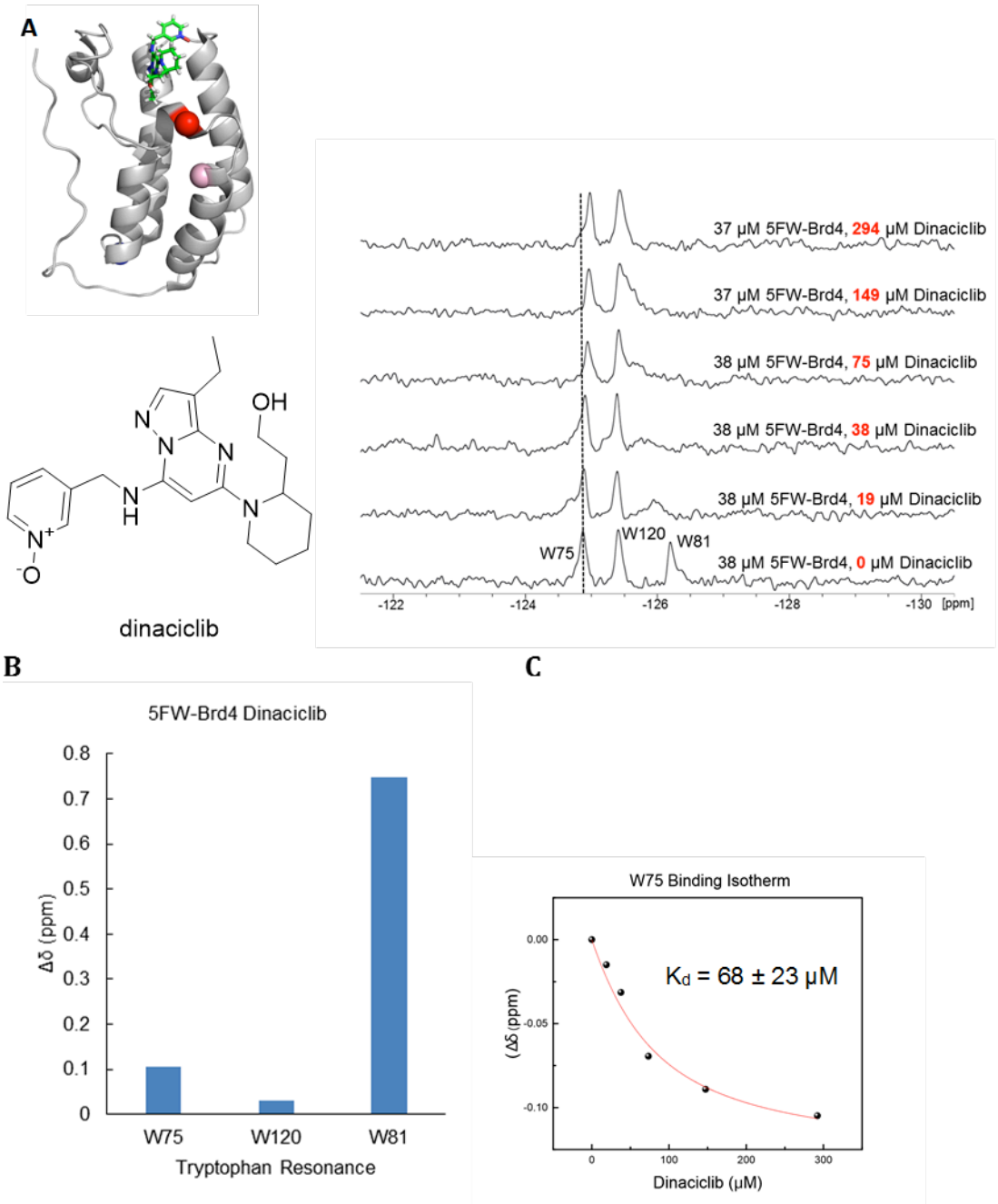


Figure 2.13 ^{19}F NMR spectral analysis of 5FW-Brd4 with dinaciclib.
A) Representative ^{19}F NMR spectra from titration with dinaciclib. Ribbon diagram of dinaciclib bound to Brd4(1) (PDB Code: 4O70). The α-carbon of tryptophan are indicated as spheres. B) Absolute value of chemical shift perturbation for the 5FW-Brd4(1) tryptophan resonances at 294 μM dinaciclib. C) A binding isotherm of W75 perturbation for the dinaciclib titration.

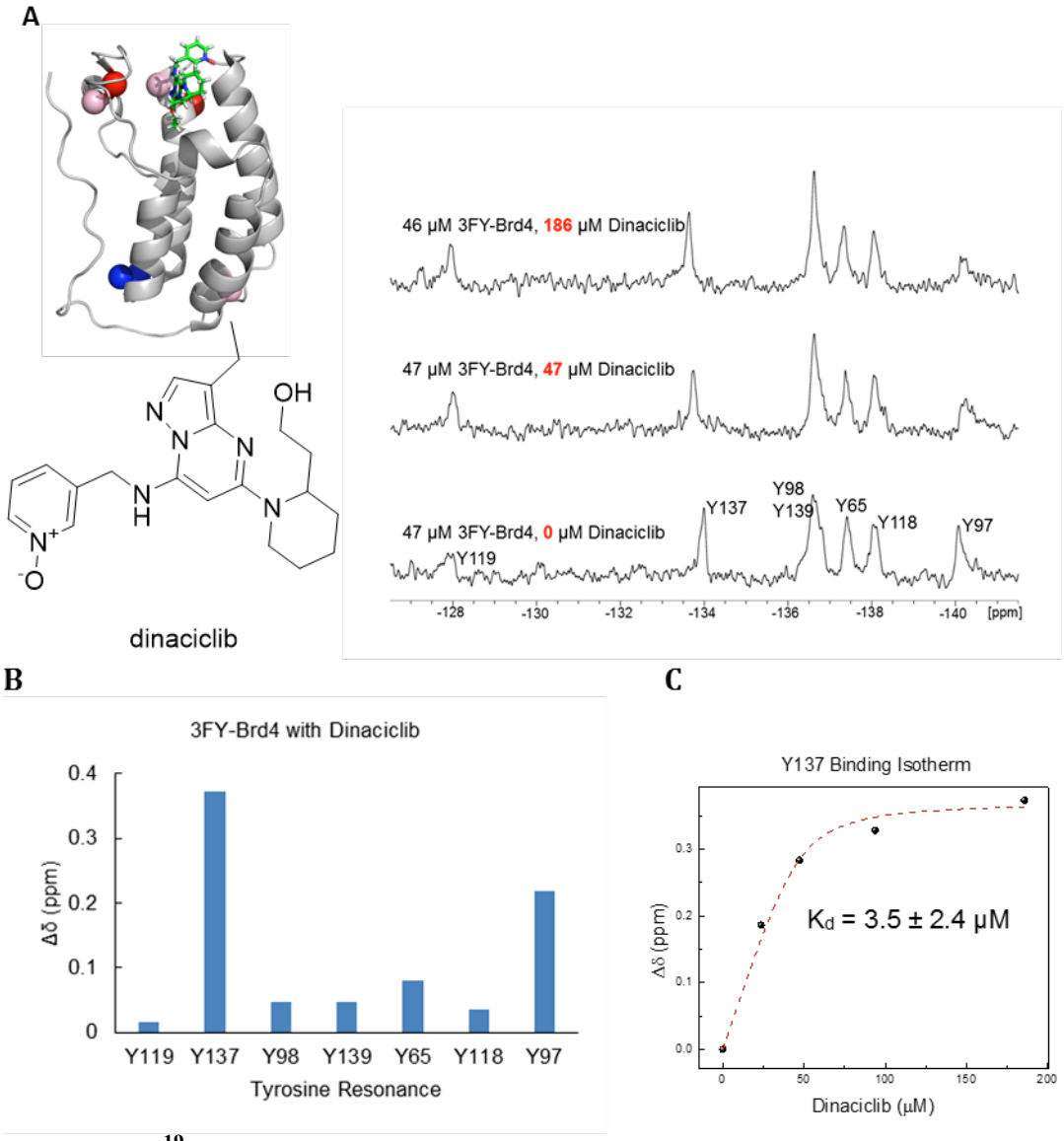


Figure 2.14 ^{19}F NMR spectral analysis of 3FY-Brd4 with dinaciclib.
A) Representative ^{19}F NMR spectra from titration with dinaciclib. Ribbon diagram of dinaciclib bound to Brd4(1) (PDB Code: 4O70). The α-carbon of tyrosine are indicated as spheres. B) Absolute value of chemical shift perturbation for the 3FY-Brd4(1) tyrosine resonances at 186 μM dinaciclib. C) A binding isotherm of Y137 perturbation for the dinaciclib titration. These values are slightly different than obtained from the 5FW-Brd4(1) titration, suggesting potential effects from fluorine incorporation and/or the error in the affinity estimate.

Table 2.5 5FW-Brd4 19F NMR chemical shift perturbations at different dinaciclib concentrations:

Dinaciclib Concentration	Chemical Shift	W75	W120	W81
0 μM	Chemical shift (ppm)	-124.875	-125.400	-126.178
19 μM	Δδ (ppm)	-0.015	0.005	0.233
38 μM	Δδ (ppm)	-0.031	0.009	*
75 μM	Δδ (ppm)	-0.069	-0.016	0.546
149 μM	Δδ (ppm)	-0.089	-0.032	*
294 μM	Δδ (ppm)	-0.105	-0.031	0.748

Table 2.6 3FY-Brd4 19F NMR chemical shift perturbations at different dinaciclib concentrations:

Ligand Conc.	Chemical Shift	Y119	Y137	Y98/139	Y98/139	Y65	Y118	Y97
0 μM	Chemical shift (ppm)	-127.88	-133.99	-136.59	-136.59	-137.41	-138.03	-140.07
24 μM	Δδ (ppm)	-0.10	0.18	-0.04	-0.04	-0.00	-0.06	-0.03
47 μM	Δδ (ppm)	-0.08	0.28	-0.02	-0.02	0.04	-0.01	-0.18
94 μM	Δδ (ppm)	-0.03	0.32	-0.03	-0.03	0.05	-0.01	-0.19
186 μM	Δδ (ppm)	-0.01	0.37	-0.04	-0.04	0.08	-0.03	-0.21

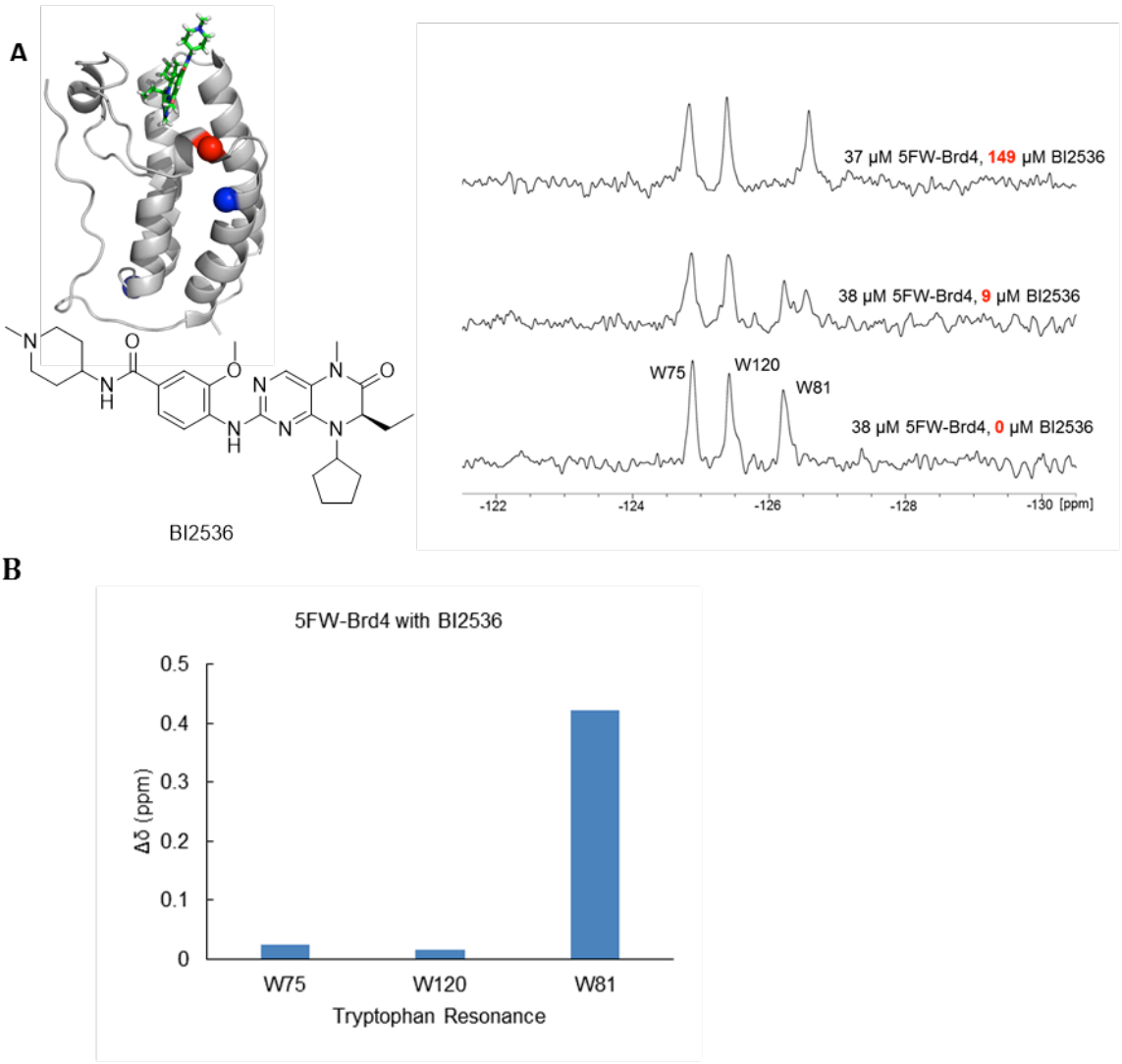


Figure 2.15 ^{19}F NMR spectral analysis of 5FW-Brd4 with BI2536.
A) Representative ^{19}F NMR spectra from titration with BI2536. Ribbon diagram of BI2536 bound to Brd4(1) (PDB Code: 4O74). The α-carbon of tryptophan are indicated as spheres. B) Absolute value of chemical shift perturbation for the 5FW-Brd4(1) tryptophan resonances at 149 μM BI2536.

Table 2.7 5FW-Brd4 19F NMR chemical shift perturbations at different BI2536 concentrations:

BI2536 Concentration	Chemical Shift	W75	W120	W81
0 μM	Chemical shift (ppm)	-124.884	-125.417	-126.207
149 μM	Δδ (ppm)	0.024	0.015	-0.421

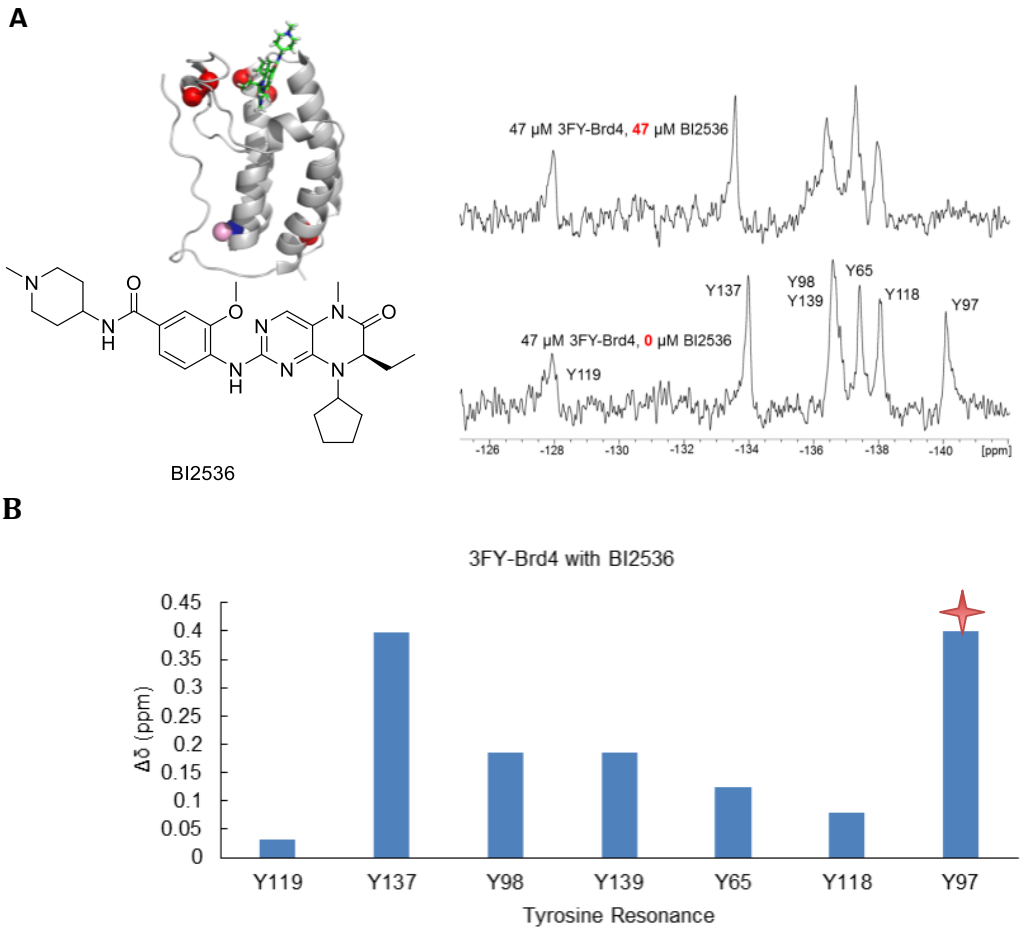


Figure 2.16 ^{19}F NMR spectral analysis of 3FY-Brd4 with BI2536.
A) Representative ^{19}F NMR spectra from titration with BI2536. Ribbon diagram of BI2536 bound to Brd4(1) (PDB Code: 4O74). The α-carbon of tyrosine are indicated as spheres. B) Absolute value of chemical shift perturbation for the 3FY-Brd4(1) tyrosine resonances at 47 μM BI2536. The star indicates broadening into baseline.

Table 2.8 3FY-Brd4 ^{19}F NMR chemical shift perturbations at different BI2536 concentrations:

BI2536 Concentration	Chemical Shift	Y119	Y137	Y98/139	Y98/139	Y65	Y118	Y97
0 μM	Chemical shift (ppm)	-127.9	-134	-136.6	-136.6	-137.4	-138	-140.084
47 μM	Δδ (ppm)	-0.032	0.398	0.1853	0.1853	0.1233	0.0788	*

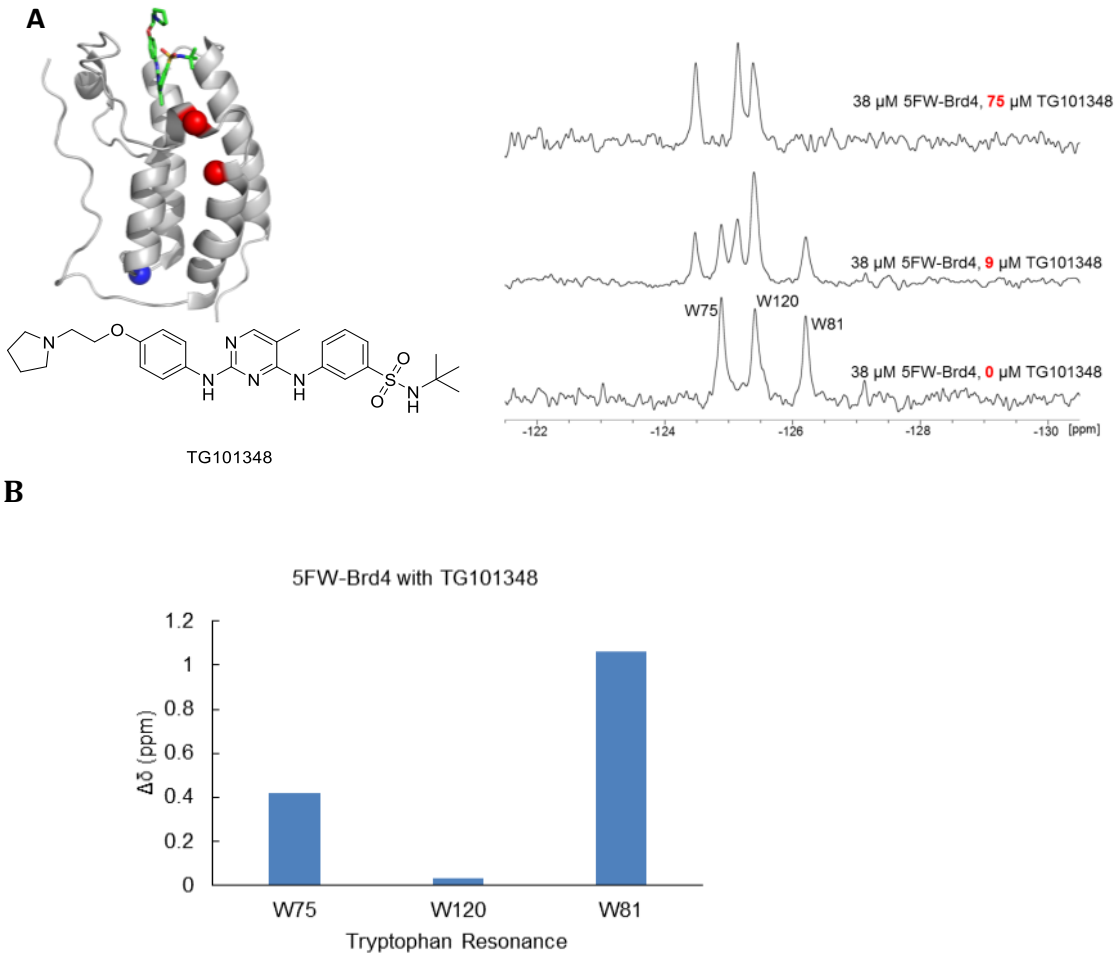


Figure 2.17 ^{19}F NMR spectral analysis of 5FW-Brd4 with TG101348.
A) Representative ^{19}F NMR spectra from titration with TG101348. Ribbon diagram of TG101348 bound to Brd4(1) (PDB Code: 4PS5). The α-carbon of tryptophan are indicated as spheres. B) Absolute value of chemical shift perturbation for the 5FW-Brd4(1) tryptophan resonances at 75 μM TG101348. Because the new resonances cannot be unambiguously assigned, two graphs have been generated representing the two possible tryptophan assignments of the new resonances.

Table 2.9 5FW-Brd4 ^{19}F NMR chemical shift perturbations at different TG101348 concentrations with the first possible resonance assignment:

TG101348 Concentration	Chemical Shift	W75	W120	W81
0 μM	Chemical shift (ppm)	-124.887	-125.415	-126.210
75 μM	Δδ (ppm)	-0.259	0.031	1.742

Table 2.10 5FW-Brd4 ^{19}F NMR chemical shift perturbations at different TG101348 concentrations with the second possible resonance assignment:

TG101348 Concentration	Chemical Shift	W75	W120	W81
0 μM	Chemical shift (ppm)	-124.887	-125.415	-126.210
75 μM	Δδ (ppm)	0.420	0.031	1.064

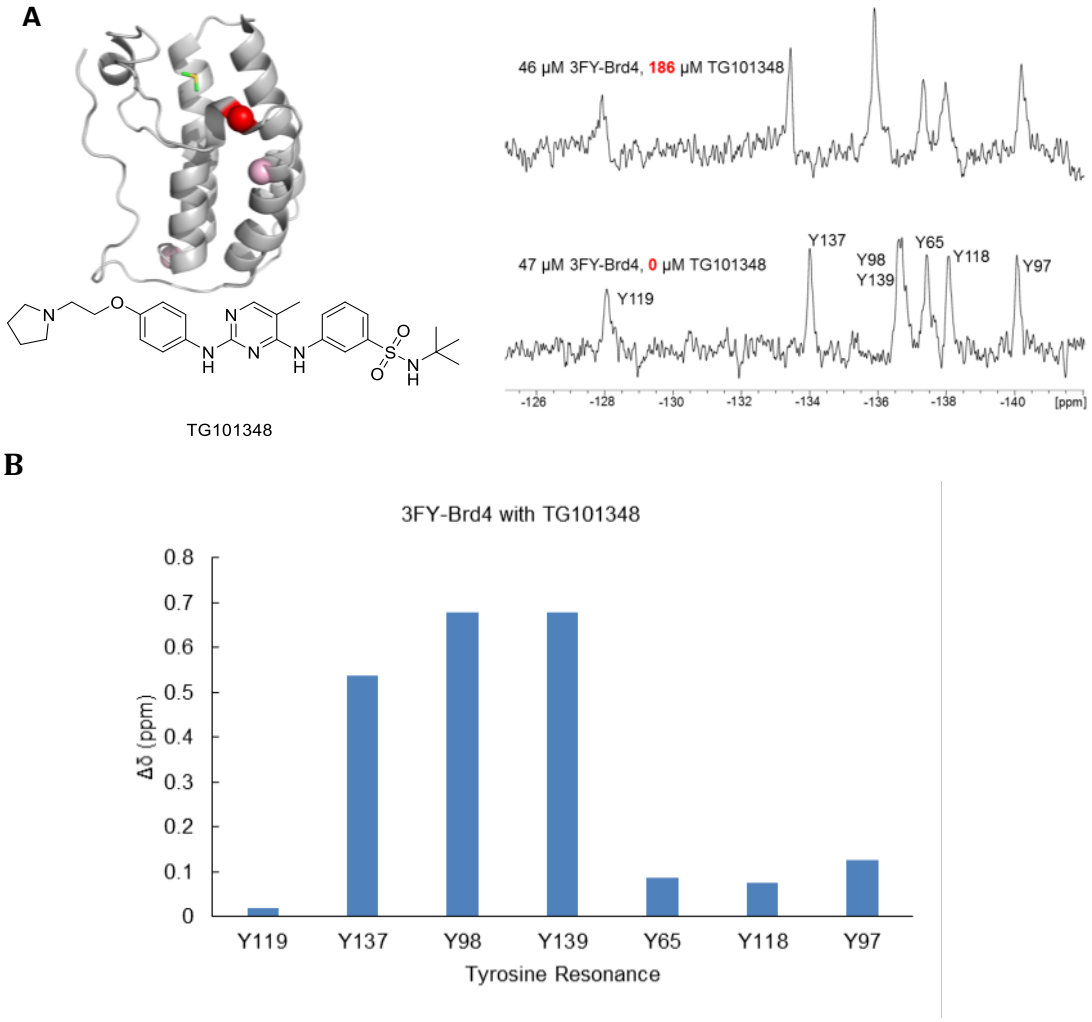


Figure 2.18 ^{19}F NMR spectral analysis of 3FY-Brd4 with TG101348.
A) Representative ^{19}F NMR spectra from titration with TG101348. Ribbon diagram of TG101348 bound to Brd4(1) (PDB Code: 4PS5). The α-carbon of tyrosine are indicated as spheres. B) Absolute value of chemical shift perturbation for the 3FY-Brd4(1) tyrosine resonances at 186 μM TG101348.

Table 2.11 3FY-Brd4 ^{19}F NMR chemical shift perturbations at different TG101348 concentrations:

Ligand Conc.	Chemical Shift	Y119	Y137	Y98/139	Y98/139	Y65	Y118	Y97
0 μM	Chemical shift (ppm)	-127.950	-133.980	-136.577	-136.577	-137.416	-138.055	-140.073
186 μM	Δδ (ppm)	0.020	0.538	0.679	0.679	0.087	0.074	-0.126

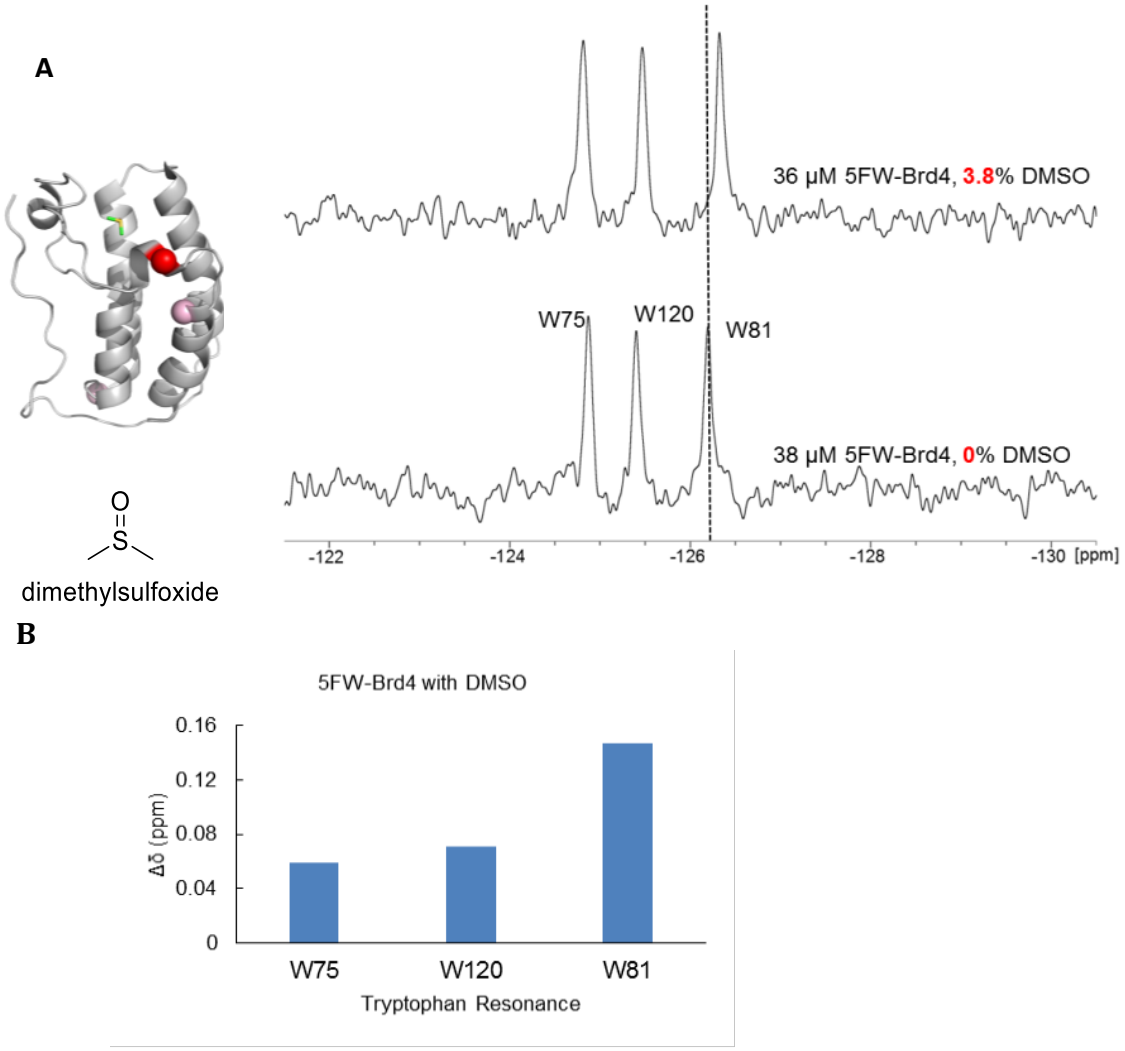


Figure 2.19 ^{19}F NMR spectral analysis of 5FW-Brd4 with DMSO.
A) Representative ^{19}F NMR spectra from titration with DMSO. Ribbon diagram of DMSO bound to Brd4(1) (PDB Code: 4IOR). The α-carbon tryptophan are indicated as spheres. B) Absolute value of chemical shift perturbation for the 5FW-Brd4(1) tryptophan resonances at 3.8% DMSO.

Table 2.12 5FW-Brd4 ^{19}F NMR chemical shift perturbations at different DMSO concentrations:

DMSO Amount	Chemical Shift	W75	W120	W81
0%	Chemical shift (ppm)	-124.8749	-125.3996	-126.1784
3.8%	Δδ (ppm)	0.0588	-0.0709	-0.1473

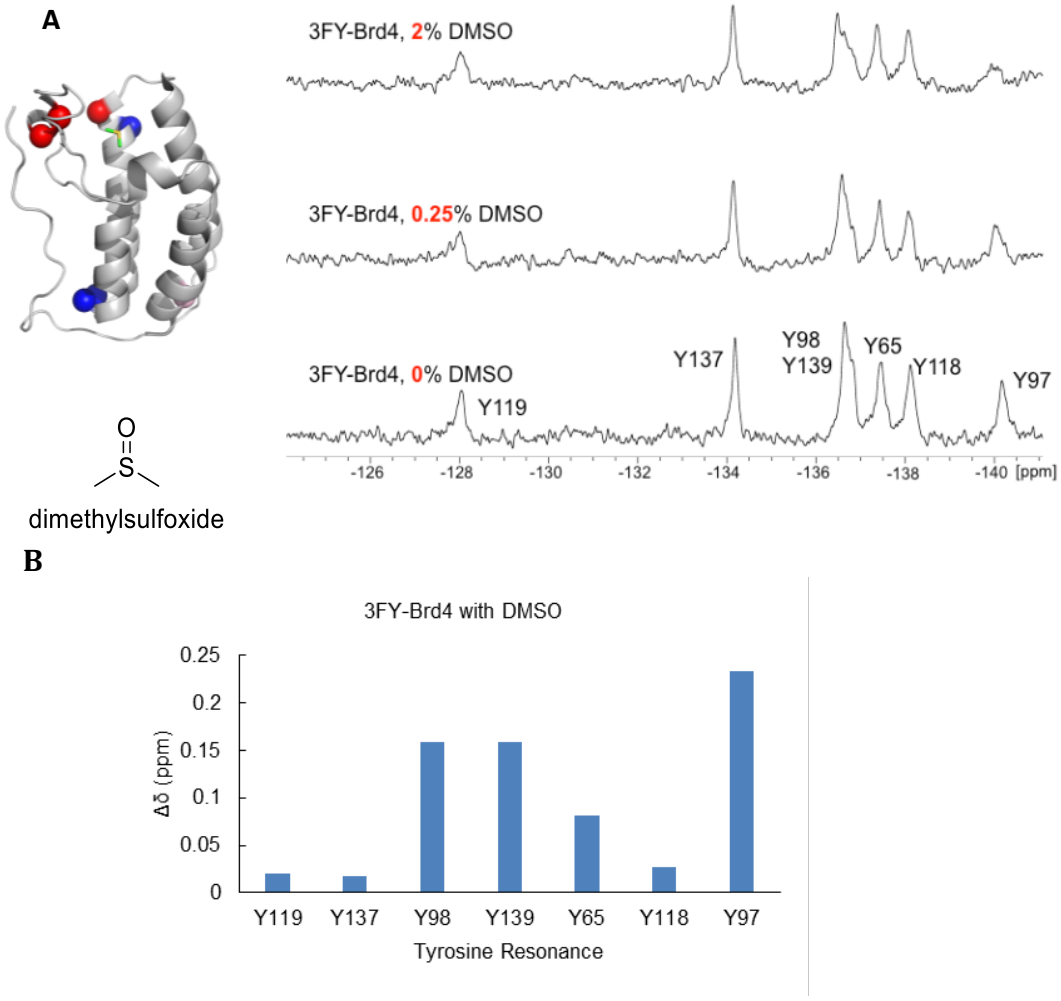


Figure 2.20 ^{19}F NMR spectral analysis of 3FY-Brd4 with DMSO.
A) Representative ^{19}F NMR spectra from titration with DMSO. Ribbon diagram of DMSO bound to Brd4(1) (PDB Code: 4IOR). The α-carbon of tyrosine are indicated as spheres. B) Absolute value of chemical shift perturbation for the 3FY-Brd4(1) tyrosine resonances at 2% DMSO.

Table 2.13 3FY-Brd4 ^{19}F NMR chemical shift perturbations at different DMSO concentrations:

DMSO Amount	Chemical Shift	Y119	Y137	Y98/139	Y98/139	Y65	Y118	Y97
0%	Chemical shift (ppm)	-128.04	-134.15	-136.64	-136.64	-137.45	-138.10	-140.17
0.25%	Δδ (ppm)	0.03	0.01	0.06	0.06	0.02	0.03	0.15
2%	Δδ (ppm)	0.02	0.02	0.16	0.16	0.08	0.03	0.23

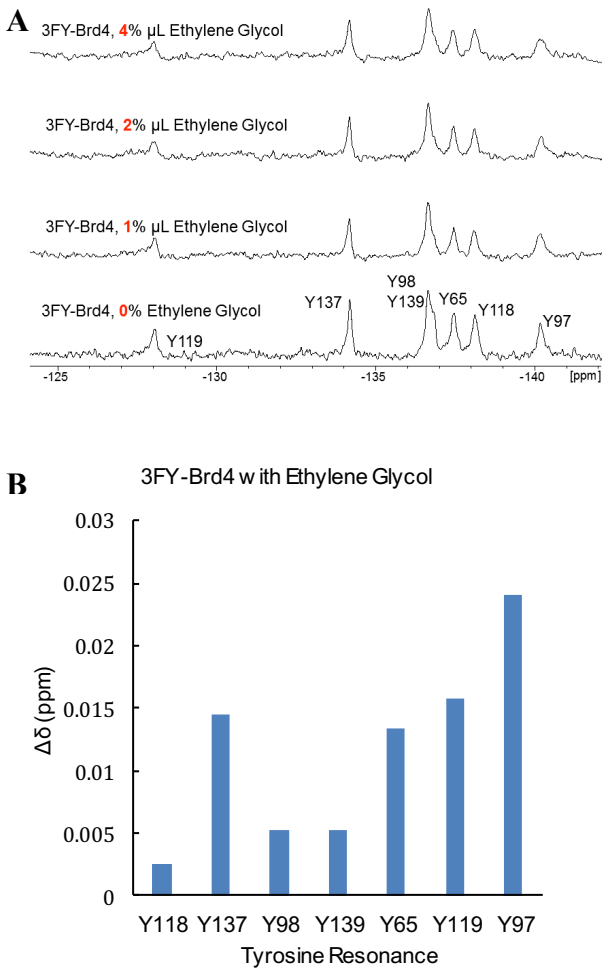


Figure 2.21 ^{19}F NMR spectral analysis of 3FY-Brd4 with ethylene glycol.
A) Representative ^{19}F NMR spectra from titration with ethylene glycol. B) Absolute value of chemical shift perturbation for the 3FY-Brd4(1) tyrosine resonances at 2% ethylene glycol.

Table 2.14 3FY-Brd4 ^{19}F NMR chemical shift perturbations at different ethylene glycol concentrations:

Ethylene Glycol Amount	Chemical Shift	Y119	Y137	Y98/139	Y98/139	Y65	Y118	Y97
0%	Chemical shift (ppm)	-128.044	-134.186	-136.643	-136.643	-137.469	-138.120	-140.172
1%	Δδ (ppm)	0.025	0.015	0.004	0.004	0.015	0.034	-0.018
2%	Δδ (ppm)	-0.002	0.014	-0.005	-0.005	0.013	0.016	-0.024
4%	Δδ (ppm)	0.017	0.008	-0.020	-0.020	0.056	0.007	0.002

Based on the Brd4(1) results, we tested the generality of using PrOF NMR with a second BET bromodomain, BrdT, of interest as a male contraception target,[107] and a non-BET bromodomain, BPTF, for which no small molecule screens nor small molecule x-ray complexes have been reported. Both proteins contain a WPF shelf; therefore, we labeled these proteins with 5FW. BI2536 and (+)-JQ1 bind to BrdT. We detect a perturbation of the fluorinated protein resonances in slow exchange, consistent with their reported nanomolar affinity. (+)-JQ1 is selective for BETs, and thus in the presence of 5FW-BPTF, we do not detect binding (Figure 2.25, Table 2.16). Surprisingly, PrOF NMR reveals BI2536 binds to 5FW-BPTF (Figure 2.22, Figure 2.23, Table 2.15), which went undetected in existing selectivity screens using BROMO*scan*^{SM} and thermal shifts,[29,30] highlighting the sensitivity of ^{19}F NMR. Deconstruction of the BI2536 ligand may lead to a useful starting point for BPTF chemical probe development to study its role in cancer.[38]

Finally, screening in the presence of other BET proteins can increase the stringency and binding information for finding selective BET inhibitors. Binding induced shifts from pan-BET inhibitor (+)-JQ1 are reversed in the presence of excess unlabeled BrdT(1) (Figure 2.9, Figure 2.24), The ^{19}F NMR spectrum with BI2536 is similarly perturbed. The non-BET, BPTF bromodomain has one tryptophan in the WPF shelf. As a new experiment, due to the significant chemical shift dispersion and simplified ^{19}F NMR spectra, we decided to simultaneously test two fluorinated proteins. Consistent with single protein binding studies, addition of (+)-JQ1 to 5FW-Brd4(1) and 5FW-BPTF shows selectivity for Brd4(1) (Figure 2.9), whereas BI2536 exhibits binding to both bromodomains.

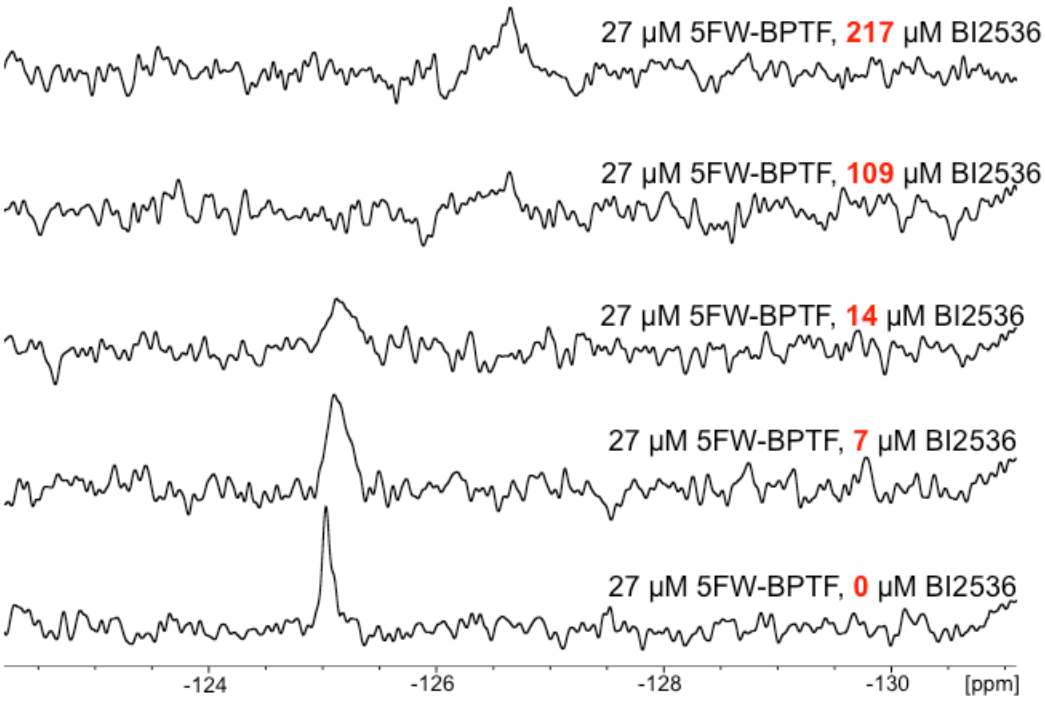


Figure 2.22 ^{19}F NMR spectral analysis of 5FW-BPTF with BI2536.
Above are representative ^{19}F NMR spectra from titration with BI2536.

Table 2.15 5FW-BPTF 19F NMR chemical shift perturbations at different BI2536 concentrations:

BI2536 Concentration	Chemical Shift	W2824
0 μM	Chemical shift (ppm)	-125.031
7 μM	Δδ (ppm)	-0.072
14 μM	Δδ (ppm)	-0.092
109 μM	Δδ (ppm)	*
217 μM	Δδ (ppm)	-1.621

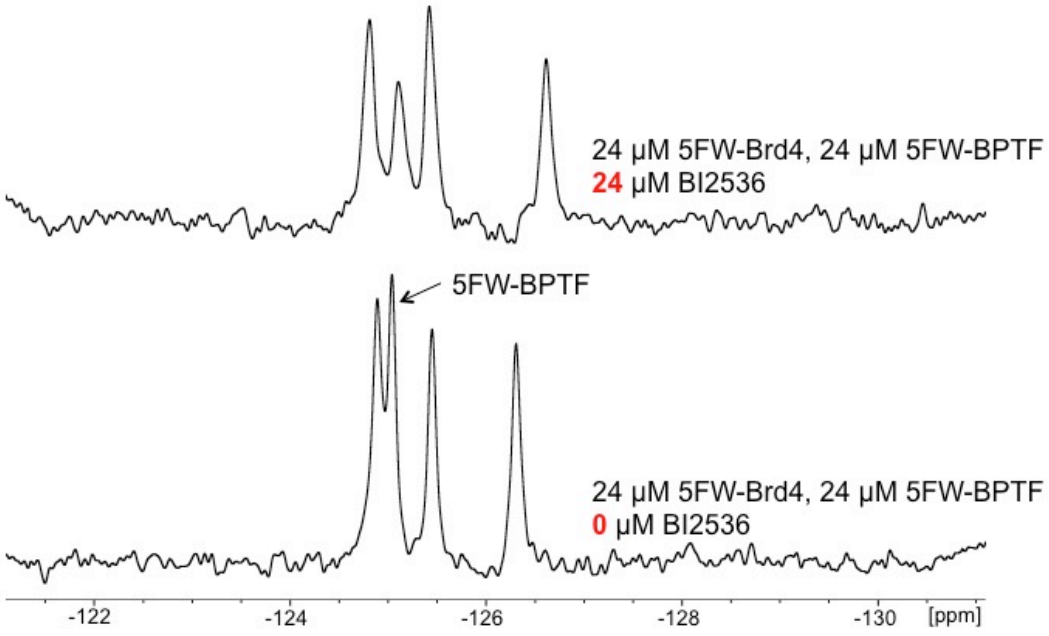


Figure 2.23 ^{19}F NMR spectral analysis of 5FW-Brd4 and 5FW-BPTF with BI2536. The 5FW-Brd4 W81 resonance is perturbed, and the 5FW-BPTF W2824 resonance broadens and shifts upfield. This is indicative of intermediate exchange of BI2536 with 5FW-BPTF in the presence of 5FW-Brd4.

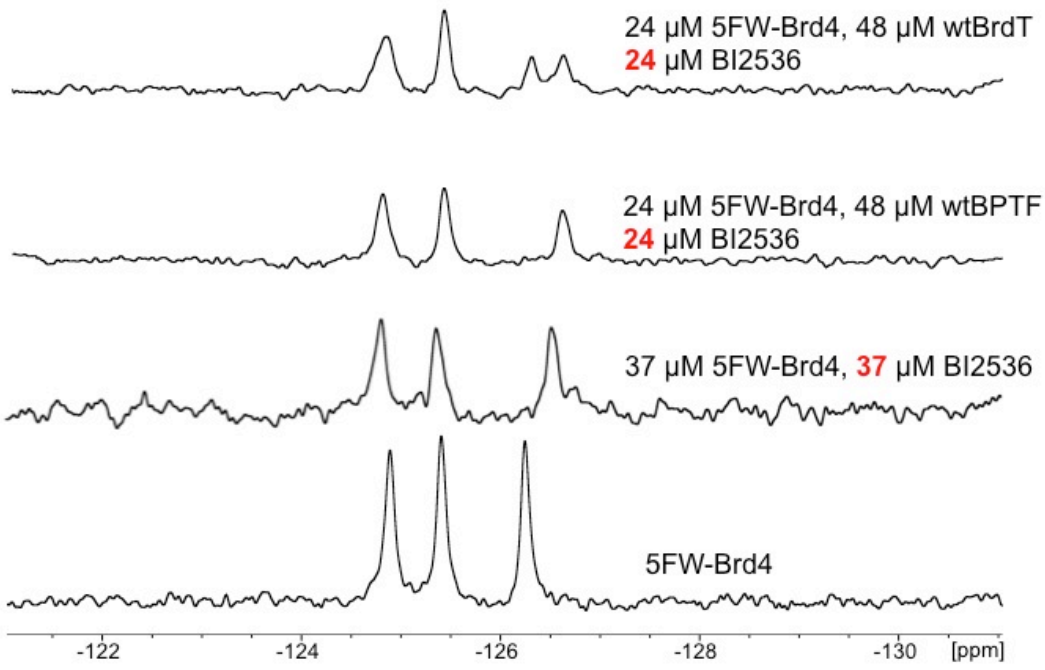


Figure 2.24 ^{19}F NMR spectral analysis of 5FW-Brd4 with BI2536 in the presence of other bromodomains.
Above, wtBrdT is shown to partially compete off BI2536 from Brd4 as there are two populations of 5FW-Brd4 present in the NMR spectrum. Protein wtBPTF is shown to not compete off BI2536 from Brd4, as there is only one population present in the NMR spectrum indicative of the fully bound state obtained from using only 5FW-Brd4 and BI2536.

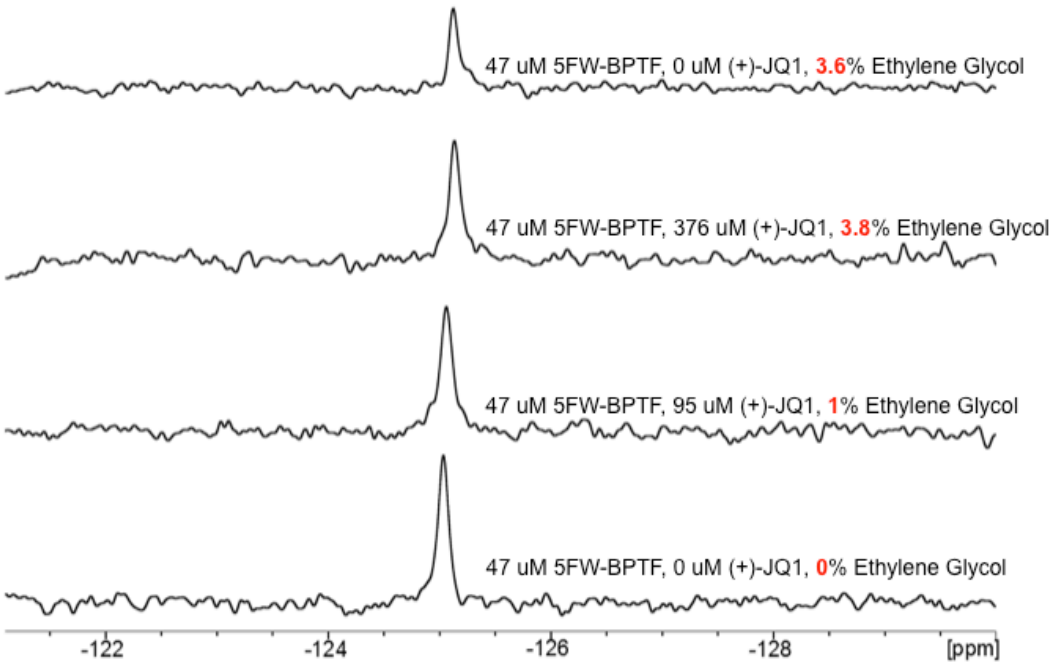


Figure 2.25 ^{19}F NMR spectral analysis of 5FW-BPTF with (+)-JQ1.
Above are ^{19}F NMR spectra from titration with (+)-JQ1. Small chemical shift perturbations in fast exchange at high micromolar concentrations of (+)-JQ1 could not be separated from solvent effects with just ethylene glycol, supporting selectivity of (+)-JQ1 for BET bromodomains.

Table 2.16 5FW-BPTF ^{19}F NMR chemical shift perturbations at different (+)-JQ1 concentrations:

(+)-JQ1 Concentration	Chemical Shift	W2824
0 µM	Chemical shift (ppm)	-125.036
95 µM	Δδ (ppm)	-0.026
376 µM	Δδ (ppm)	-0.099
3.6% Ethylene Glycol	Δδ (ppm)	-0.088

2.3 **Conclusion**

Using a validated therapeutic epigenetic protein, we evaluated fluorinated bromodomains as sensitive tools for detecting ligand binding modes via ^{19}F NMR. By labeling two different aromatic amino acids, we characterized variations in binding modes of (+)-JQ1 and new ligands repurposed from the kinase field. We further showed that our method can detect protein binding with weak fragments, DMSO (Figure 2.19, Figure 2.20, Table 2.12, Table 2.13) and acetaminophen. This is encouraging due to the success of developing fragments into lead molecules in drug discovery campaigns. Based on the high conservation of aromatic amino acids across 61 bromodomains, we anticipate our method should be generalizable for bromodomains, many lacking specific tool compounds to study their biology such as BPTF. We further demonstrate the non-perturbing nature of fluorine for Brd4(1), by solid and solution state methods. Future studies will use PrOF NMR in a full fragment screen. This side-chain labeling technique has been employed for studying small to medium-sized proteins (10 to 60 kDa),[106,108] as well as G-protein coupled receptors.[109] Here we apply it to bromodomains both in isolation and in mixtures. Due to the speed of PrOF NMR, ease of interpretation, and availability of ^{19}F-tuned NMR probes in academic and industrial settings, this approach should be of broad appeal for early stage ligand discovery.

2.4 **Methods**

Brd4(1)(42-168), BrdT(1)(29-134), and BPTF (2793-2911) Protein Expression:

The pNIC28-BSA4 plasmid containing the Brd4(1)and BPTF genes were kind gifts from the laboratory of Stefan Knapp. For protein expression, either the *E. coli* Rosetta (DE3) strain (Novagen) was first transformed with the respective expression plasmid or the BL21(DE3) strain was cotransformed along with the pRARE (Novagen) plasmid and plated onto agar plates containing kanamycin (100 mg/L) and chloramphenicol (35 mg/L). Following overnight incubation at 37 ^{o}C, a single colony was selected from the agar plate and inoculated in 50 mL of LB media containing kanamycin (100 mg/L) and chloramphenicol (35 mg/L). The primary culture was grown

overnight at 25 ^{o}C while shaking at 250 rpm. For secondary culture growth, 1 L of LB media containing kanamycin (100 mg/L) was inoculated with the primary culture and cultured at 37 ^{o}C while shaking at 250 rpm. When the O.D. of culture at 600 nm reached 0.6, the shaker temperature was reduced to 20 ^{o}C. After 30 minutes, the expression was induced with 1 mM IPTG overnight for 12-16 h. Cells were harvested by centrifugation.

Expression of 3FY and 5FW labeled Brd4(1), 5FW-BrdT(1), and 5FW-BPTF:

3FY and 5FW labeled Brd4(1) were expressed based on established methods[96,99] using *E. coli* DL39(DE3) + pRARE and *E. coli* Bl21(DE3) + pRARE strains respectively. To express the labeled protein, the secondary culture in LB media was grown until an O.D. at 600 nm of 0.6 was reached followed by harvesting and washing the cells with PBS. Washed cells were resuspended in defined media of Muchmore et al.[110] containing either 3FY (70 mg/L) in place of tyrosine, or 5-fluoroindole (60 mg/L) in place of tryptophan. The resuspended E. coli were incubated at 37 ^{o}C while shaking for 1 h followed by the cooling to 20 ^{o}C and media temperature equilibration for 30 min. Protein expression was induced with 1 mM IPTG overnight (14-16 h) at 20 ^{o}C. The cells were harvested and stored at -20 ^{o}C. 5FW-BrdT(1) and 5FW-BPTF were expressed by the same protocol. Cell pellets were thawed at room temperature followed by the addition of lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl and 10% v/v glycerol) containing protease inhibitor PMSF (5 mM) and purified according to methods described in the appendix using Ni-Affinity chromatography. Yields following purification are 60 mg/L Brd4(1), 60 mg/L, 5FW-Brd4(1) (78-90% 5FW incorporation) and 10 mg/L 3FY-Brd4(1) (>95% 3FY incorporation) 26 mg/L 5FW-BPTF (94%, incorporation) and 22 mg/L BrdT, 11 mg/L 5FW-BrdT(1) (91%, incorporation). Purity of proteins was assessed by SDS-PAGE. Fluorinated amino acid incorporation efficiency in proteins was measured by mass spectrometry as described in the appendix. Concentration was determined via absorbance at 280 nm according to published values.[99]

1D ^{19}F NMR Parameters:

^{19}F NMR spectra were acquired at 470 MHz on a Bruker 500 spectrometer with a 5 mm Prodigy TCI Cryoprobe without proton decoupling. Samples containing 40-50 μM

bromodomains labeled in 50 mM TRIS, 100 mM NaCl, 1 mM CHAPS, 2 mM DTT, 5% D_2O, pH 7.4 for binding assays unless otherwise stated. Spectra were referenced to trifluoroacetate (−76.55 ppm). Measurement parameters included a relaxation delay time of 0.7 s for 5FW-Brd4(1) and a 90^o flip angle, and a relaxation delay of 0.2 s for 3FY-Brd4(1) containing a 30^o flip angle. An acquisition time of 0.05 s was used for all experiments. A sweepwidth of 10 ppm was used for 5FW-Brd4(1) spectra, and 18 ppm for 3FY-Brd4(1). A 20 Hz line-broadening was applied after 500-3000 transients unless otherwise stated.

Ligand Binding Studies:

Small molecules were titrated into the protein solution from concentrated stock solutions of ethylene glycol (10 mM for all small molecules except acetaminophen, which was 50 mM). Final ethylene glycol concentrations were kept below 4% ethylene glycol. For small molecule titrations 500 scans were acquired with 5FW-Brd4(1) and 3000 with 3FY-Brd4(1) to ensure good S/N resolution for improved fitting of the data. However, good chemical shift estimates can be readily acquired at 200−400 scans in under 5 min for initial screening as described in the appendix. Stock solutions of kinase inhibitors were prepared from preweighed 5 mg samples which were used to estimate final concentrations. K_d values were obtained using a one-site-binding equation accounting for ligand depletion. Reported errors are from the non-linear regression fit of the data.

Structure determination:

Protein crystallization was performed with the mosquito LCP (TTP Labtech) crystallization robot at 18 °C using the sitting drop vapor diffusion method. Crystals of 3FY-Brd4(1) were grown in the presence of 1 mM (+)-JQ1 and 10 % (v/v) DMSO from 0.2 M $NH_4C_2H_3O_2$, 0.1 M HEPES (pH 7.5) and 25 % (w/v) PEG 3350, harvested in cryoprotectant (reservoir containing 25 % (v/v) ethylene glycol and 0.5 mM ligand) and flash frozen in a stream of nitrogen gas. X-ray diffraction data were collected at -180 °C using station 22-ID, SER-CAT, Advanced Photon Source, Argonne National Laboratories.

Protein Expression and Molecular Biology Materials:

For *E. coli* growth, LB agar, LB media, defined media components including unlabeled amino acids, uracil, thiamine-HCl, nicotinic acid, biotin and buffer components were purchased from RPI corp. 3-Fluorotyrosine, thymine, cytosine, guanosine were purchased from Alfa Aesar. Magnesium chloride, manganese sulfate, succinic acid, calcium chloride and 5-fluoroindole were purchased from Sigma-Aldrich. Miniprep plasmid purification kit was purchased from Clontech.

Unlabeled Brd4(1), BrdT(1), and BPTF Protein Expression:

The pNIC28-BSA4 plasmid containing the Brd4(1) and BPTF genes were kind gifts from the laboratory of Stefan Knapp. For protein expression, either the *E. coli* Rosetta (DE3) strain (Novagen) was first transformed with the respective expression plasmid or the BL21(DE3) strain was cotransformed along with the pRARE (Novagen) plasmid and plated onto agar plates containing kanamycin (100 mg/L) and chloramphenicol (35 mg/L). Following overnight incubation at 37 ^{o}C, a single colony was selected from the agar plate and inoculated in 50 mL of LB media containing kanamycin (100 mg/L) and chloramphenicol (35 mg/L). The primary culture was grown overnight at 25 ^{o}C while shaking at 250 rpm. For secondary culture growth, 1 L of LB media containing kanamycin (100 mg/L) was inoculated with the primary culture and cultured at 37 ^{o}C while shaking at 250 rpm. When the O.D. of culture at 600 nm reached 0.6, the shaker temperature was reduced to 20 ^{o}C. After 30 minutes, the expression was induced with 1 mM IPTG overnight for 12-16 h. Cells were harvested by centrifugation and stored at -20 ^{o}C.

Bromodomain Purification:

To purify fluorinated and unlabeled Brd4(1), the cell pellet was thawed at room temperature followed by the addition of lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl and 10% v/v glycerol) containing protease inhibitor PMSF (5 mM). Cells were lysed by sonication and the cell lysate was centrifuged at 6500 g for 30 minutes followed by supernatant filtration over Whatman filter paper. Filtrate containing the histidine-

tagged Brd4(1) was loaded on to a nickel-NTA affinity column and eluted with an imidazole gradient on an AKTA FPLC system monitoring the O.D. at 280 nm. Imidazole was removed from the buffer using a HiPrep column (GE) for buffer exchange into either 50 mM potassium phosphate, pH 7.4, and 100 mM NaCl or 50 mM Tris pH 7.4, 100 mM NaCl. Purified and buffer exchanged protein was treated with TEV protease for either 2 hours at room temperature or alternatively at 4 ^{o}C overnight on a rotating carrousel. The cleaved His-tag, TEV protease and uncleaved Brd4 were removed using nickel-NTA affinity resin.

Site-directed mutagenesis:

To assign the fluorinated resonances in the ^{19}F NMR spectrum, we mutated single tyrosine and tryptophan residues to phenylalanine using a standard PCR amplification method. One common reverse primer and forward primers containing point mutation for (Y65F, Y98F, Y118F, Y137F, Y139F, W75F and W120F) and a primer pairs for (Y97F, Y119F and W81F) were used for PCR amplification. Primer sequences are shown in Table 2.17. Parent plasmid template was digested by Dpn1 nuclease followed by transformation into E. coli BL21(DE3). Three colonies for each mutant were picked and grown in LB media containing kanamycin followed by miniprep plasmid purification (kits from Clontech). Point mutations on plasmids were confirmed by Sanger sequencing. Auxotrophic E. coli DL39(DE3) + pRARE and E. coli BL21(DE3) + pRARE cells were transformed with tyrosine and tryptophan mutation containing plasmids respectively and subjected to protein expression and purification. After several attempts Y97, Y119 and W81 mutants did not express well, and were not purified. These assignments were either inferred (last resonance unassigned for W120) or validated by binding a known ligand (Y119 vs. Y97).

Table 2.17 Primer sequences for point mutation of tyrosine and tryptophan to phenylalanine:

S. No.	Residue Name	Primer	Sequence
1.	Y65F	Forward	CCAACTGCAATTCCTGCTCAGAG
2.	Y98F	Forward	CCTCCCTGATTACTTCAAGATCAT
3.	Y118F	Forward	GGAAAACAACTTCTACTGGAATGC
4.	Y137F	Forward	ACAAATTGTTTCATCTACAACAAGCC
5.	Y139F	Forward	TGTTACATCTTCAACAAGCCTGG
6.	W75F	Forward	CAAGACACTATTTAAACACCAGTT
7.	W120F	Forward	CAACTATTACTTTAATGCTCAGGAAT
8.	Common	Reverse	CAGCAGCCAACTCAGCTTCCT
9.	Y97F	Reverse	GATCTTATAGAAATCAGGGAGG
10.	Y119F	Reverse	CCTGAGCATTCCAGAAATAGTTG
11.	W81F	Reverse	CTGCTGGAAAGGAAATGCAAACT

Protein Mass Spectral Analysis:

Product molecular weight was confirmed by electrospray ionization mass spectrometry (ESI-MS) using a Thermo Scientific Orbitrap Velos LC-MS, Table 2.18. To determine the percent incorporation for fluorinated proteins the integration values of the different deconvoluted mass peaks are entered into the following equation to determine the relative incorporation e..g, FWBrd4:

$$\%_{incorporation} = \frac{(0FWBrd4)*0+(1FWBrd4)*1+(2FWBrd4)*2+(FWBrd4)*3}{(0FWBrd4)*3+(1FWBrd4)*3+(2FWBrd4)*3+(FWBrd4)*3} * 100$$

0FWBrd4 is 5FWBrd4 with no fluorine substitutions, 1FWBrd4 is 5FWBrd4 with one fluorine substitution, 2FWBrd4 has two fluorines substituted and FWBrd4 has 3 fluorines substituted.

Table 2.18 Deconvoluted mass spectral data of wild-type Brd4(1) as well as fluorinated variants and mutants. Only data for fully labeled or unlabeled proteins are shown.

Protein	Calculated	Observed
wtBrd4	15083.33	15083.2
5FW-Brd4	15137.3	15136.8
3FY-Brd4	15209.26	15206.6
Y98F Brd4	15175.27	15173
Y118F Brd4	15175.27	15174.2
Y137F Brd4	15175.27	15172.8
Y139F Brd4	15175.27	15174.3
W75F Brd4	15080.27	15079.4
W120F Brd4	15080.27	15080.1
wtBPTF	14437.40	14434.4
5FW-BPTF	14455.40	14454.3
wtBrdT	14148.44	14153.3
5FW-BrdT	14184.42	14188.5

Circular Dichroism:

To check the secondary structural content, far-UV CD spectra (200-260 nm) of unlabeled and labeled proteins were collected using a peltier equipped temperature controlled Jasco J-815 spectropolarimeter at 25 ^{o}C. For all measurements, 20 ¥ M (50 mM Tris buffer pH 7.4 containing 100 mM NaCl) of protein and a 1 mm cuvette path-length were used. Spectral data were collected at a scan rate of 50 nm/min with averaging of 5 spectra. Processed data were baseline corrected against spectra taken with buffer alone.

Thermal Melting:

Thermal stabilities of labeled and unlabeled proteins were measured by the change in ellipticity at 222 nm with the increase in temperature from 20 oC to 80 oC at the scan rate of 60 degrees/h. The Mid-point of transition was calculated by a sigmoidal fit to determine the Tm.

Isothermal titration calorimetry of unlabeled, 3FY and 5FW labeled Brd4(1):

Auto ITC200 (GE) and Nano-ITC were used to perform isothermal titration calorimetry experiments. A 3FY-Brd4(1) solution was loaded in the Auto ITC200 calorimeter and titrated with (+)-JQ1 diluted in the same buffer as protein (50 mM

potassium phosphate pH 7.4, 100 mM NaCl). Note: to facilite solubility, JQ1 solutions were sonicated immediately prior to use. We used 47 μM of 3FY-Brd4(1). A 50 mM (+)-JQ1 stock solution was prepared in DMSO and diluted in the same buffer as protein at concentration of 500 μM and used for the titration. Titrations were carried out by using 2 μL per injection volume of 500 μM (+)-JQ1. Heat liberated by each injection was integrated (ΔH) and plotted against molar ratio of ligand and protein. Obtained data was fitted by using predefined one and two binding site modes. The fitting of data was not sufficient with one binding site whereas a two binding site model produced an appropriate fit, indicating a potential low affinity binding (K_d = 1.6 mM) site which could be from DMSO. The reported K_d is an average value from of two separate titrations.

For binding of (+)-JQ1 (50 μM) to unlabeled Brd4(1) at 25 ^{o}C, the experiments were performed using Nano-ITC (TA Instruments). In this experiment, 550 μM unlabeled in 50 mM Tris pH 7.4, 100 mM NaCl and 4 mM DTT and 500 μM 5FW labeled proteins in 50 mM HEPES pH 7.5, 150 mM NaCl and 4 mM DTT were used as a ligand and loaded to the microsyringe (50 μL). The first injection of 0.5 μL was followed by 20 and 24 identical injections of 2 and 2.5 μL respectively with a releasing period of 8 seconds per injection with 200-300 second spacing time between injections. Control heat of dilution was measured by independent protein titration into buffer and was subtracted from the protein-ligand binding experimental data. Data was analyzed by NanoAnalyze software to calculate enthalpy of binding (ΔH) and dissociation constants (K_d). In all cases an independent binding model was used for data fitting.

Table 2.19 ITC experimental conditions and obtained values:

Exp.	Protein	Condition	Kd nM	n value
1	Brd4	50 mM Tris pH 7.4, 100mM NaCl, 4 mM DTT	76.0	1.02
2	Brd4	50 mM Tris pH 7.4, 100mM NaCl	75.1 ± 4.5	1.51
3	5FW-Brd4	50mM HEPES pH 7.5 150mM NaCl+ 4 mM DTT	77.6 ± 2.0	0.95
4	3FY-Brd4	50 mM Tris pH 7.4, 100mM NaCl	88.7 ± 4.7	0.96

Table 2.20 Data collection and refinement statistics for the crystal structure determination of 3FY-Brd4(1)

Space group and unit cell dimensions	$P2_12_12_1$; a = 36.8, b = 44.9, c = 78.6; α = β = γ = 90°
Resolution (Å)	1.45 - 20.0 (1.45-1.47)
Unique reflections	22,274 (1,016)
Rsym[a] (%)	7.6 (33.0)
Completeness (%)	93.5 (86.8)
I/σI	21.6 (2.9)
Rcryst[b] (%)	14.0
Rfree[c] (%)	18.5
Average B all[d] (Å²)	12.54
Average B protein[d] (Å²)	11
Average B ligand[d] (Å²)	8.36
Average B solvent[d] (Å²)	22.76
Wilson B (Å)	9.78
rmsd[e] bonds (Å)	0.009
rmsd angles (deg)	1.46
Coordinate Error (Å)	0.13
Ramachadran favored (%)	97.74
Ramachadran allowed (%)	2.26

Values in paranthesis are for the highest resolution shell.

[a] Rsym = 100 xΣhΣi|Ihi-Ih|/ΣhiI where h are unique reflection indices.

[c] Rfree is Rcryst calculated for 1112 randomly chosen unique reflections.

[d] Excluding hydrogen atoms

[e] rmsd = root-mean-square deviation from ideal values.

Chapter 3. Protein-observed 19F-NMR for fragment screening, affinity quantification and druggability assessment

Note: The following was a strongly collaborative work with Clifford Gee, and will also be reproduced in his dissertation.

Motivation: The purpose of this protocol was to disseminate the ability to express fluorinated proteins, and describe the basics of analyzing binding interactions using a titration experiment. This is to facilitate wide-spread use of the screening technology.

NMR spectroscopy, using either labeled proteins or labeled small molecules, is emerging as a preferred method for screening low-complexity molecules (typically <300 Da with a minimal number of functional groups), termed 'fragments' in early-stage ligand discovery campaigns.[111,112] Fragments typically bind to their protein target with low affinity (mid-micromolar to millimolar dissociation constants). These low-affinity interactions are readily detected using NMR methods.[113,114] Fragment molecules can be compared with higher-molecular-weight counterparts found in traditional high-throughput screening libraries, via evaluation of their ligand efficiency (LE), which compares binding affinity or activity relative to the number of atoms in the molecules.[115,116] Highly ligand-efficient compounds can be developed in an atom-economical manner into more potent compounds by fragment linking or growing

approaches.[67,117] Enthusiasm for this approach remains high with the approval of vemurafenib in 2011, discovered through an initial fragment screening campaign, and several more lead molecules that emerged from fragment screens are in late-stage clinical trials.[55,118]

Fluorine NMR is an attractive approach for fragment screening because the spin-1/2 nucleus ^{19}F is stable, has a natural abundance of 100% and is nearly absent in biological systems. Many fluorinated amino acids and building blocks are commercially available, including aromatic amino acids 3-fluorotyrosine (3FY), 4-fluorophenylalanine (4FF) and 5-fluoroindole, described herein. In many cases, minimal structural and functional perturbation has been observed.[106,119,120] ^{19}F chemical shifts are also sensitive to changes in the molecular environment, and therefore ^{19}F is an ideal background-free NMR-active nucleus for studying challenging problems of molecular recognition by biopolymers.[85,106] In the case of fluorine-labeled proteins, the environmental sensitivity of fluorine nuclei typically results in well-resolved 1D ^{19}F NMR spectra of proteins whose fluorine-labeled side chains are observed at low to mid-micromolar concentrations (e.g., 25–100 μM).[106]

Fragment screening using low-molecular-weight, low-complexity molecules has attracted considerable attention because of the reduction of chemical space compared with that of higher-molecular-weight, functional-group-rich small molecules used in high-throughput screening. As a result, fragment libraries are typically smaller than high-throughput screening libraries.[111,112,118,121] An analysis by Scanlon and co-workers[122] of 20 different fragment libraries developed in the context of academic or industrial research yielded an average library size of 4,543 fragment library members and a median size of 1,280. The use of smaller library sizes is further supported by the hit rates from these fragment screens, in which the researchers detected a binding event averaging 8.2%, as reported by 11 different screening centers. The high hit rates suggest that adequate chemical space is being covered. Fragments identified as hits have been used to develop efficient ligands with favorable physicochemical properties.[123]

One of the challenges when screening fragment molecules for binding to a specific protein target is the detection and quantification of low-affinity interactions.

Achieving this important research goal often forces researchers to use high ligand concentrations. NMR is a technique that is well-suited for working at these high concentrations. With NMR, mixtures of fragment molecules can also be tested simultaneously. In experiments using proteins labeled with NMR-active nuclei, the NMR spectrum of a mixture that results in a large change in chemical shift of the NMR-active nucleus is deconvoluted by obtaining NMR spectra of the protein with individual molecules to find the small molecule that actively binds the protein (thus causing the observed change in chemical shift). One advantage of this fragment mixture approach is that it enables researchers to test a large number of compounds in a shorter period of time than would be needed to test individual compounds one at a time. A second advantage when using a labeled protein (i.e., protein-observed NMR methods) is the added structural information that the protein resonances provide. These specific resonance perturbations can be used to guide molecular designs that are aimed at increasing fragment affinity.

Ligand-observed NMR methods, such as saturation transfer difference or transverse relaxation Carr–Purcell–Meiboom–Gill-based experiments,[124,125] provide complementary information that can be used in parallel with protein-observed methods. Ligand-observed and PrOF NMR-based methods are tolerant of a variety of experimental conditions. For example, contrary to many ^{1}H NMR-based experiments, NMR spectra of fluorine-labeled small molecules or fluorine-labeled proteins are not affected by the presence of detergents and buffers that are traditionally used.[112] Advantages associated with the implementation of ligand-observed experiments with respect to protein-observed experiments are the lower protein concentration (0.1–10 μM) needed in many cases—although increasing protein concentration can lead to a better signal[126]—and the lack of an upper limit in protein size. In addition, with respect to protein-observed experiments, the active ligand can be readily identified from the fragment mixture without deconvolution. However, in saturation transfer difference NMR, false-positive hit rates can be as high as 50%,[122] although the occurrence of such false positives can be partly mitigated by repeating the experiment in the presence of a competitor ligand (if one is known).

We have proposed a PrOF NMR screening method in which we rapidly monitor changes in the chemical shifts of fluorine resonances in ^{19}F-labeled protein side chains induced by the presence of small molecules (Figure 3.1). This approach is analogous to that adopted in determining structure–activity relationships by NMR using labeled amides in ^{1}H–^{15}N heteronuclear single-quantum coherence spectroscopy experiments[7]. The increasing availability of improved instrumentation using ^{19}F-tuned cryoprobes (e.g., the QCI-F cryoprobe from Bruker), speed of PrOF NMR experiments and ease of spectral interpretation have increased the accessibility of these experiments in academic and industrial settings. We recently applied PrOF NMR for fragment screening with the transcription-factor-binding domain KIX.[99,127] In this study, we analyzed 85 mixtures (comprising a total of 508 small molecules) in 10 h using a total of 20 mg of protein. This method has been used by others, including follow-up screens against the SPRY-domain-containing SOCS box protein 2[128] and AMA1[129], and it was shown to be more than twice as fast as ^{1}H–^{15}N heteronuclear single-quantum coherence spectroscopy NMR for small proteins.[127,130] We have also used the bromodomain BrdT, which is described below.[131] G-protein-coupled receptor agonists and antagonists can also be identified by this method.[89] 2D PrOF NMR methods have validated binding modes of fluorinated ligands via ^{19}F–^{19}F homonuclear nuclear Overhauser effect experiments with BcL-xL.[132] As the incorporation of fluorine in drug molecules increases, nuclear Overhauser effect experiments provide an additional structural biology tool for the characterization of ligand-binding modes. Our lab recently demonstrated a simultaneous analysis (i.e., multiplexed) of small-molecule binding using two 15-kDa bromodomains, Brd4 and BPTF.[131] In this study, 229 small molecules were screened using this approach.[32] Because a protein and the potential off-target are screened together, this multiplexed experiment is similar to the RAMPED-UP 2D NMR experiments with differently labeled proteins.[133] These approaches are advantageous for studies in which protein selectivity is important.

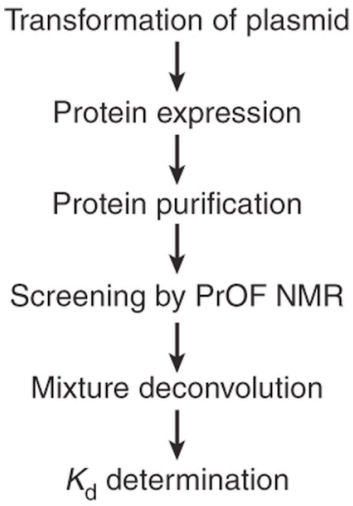


Figure 3.1 General workflow for using PrOF NMR for fragment screening and ligand characterization

In 1974, Sykes et al.[87] first reported the 19F NMR analysis of an 86-kDa protein, alkaline phosphatase, using 3FY to label the protein. Sequence-selective labeling of recombinant proteins with fluorinated amino acids is now a well-documented methodology that facilitates the general use of the present screening method based on PrOF NMR.[95,96] In this protocol, using sequence-selective labeling with fluorinated aromatic amino acids, we describe the application of PrOF NMR to the screening of small-molecule libraries for potential protein ligands and the quantification of the micromolar to millimolar dissociation constants from chemical shift perturbation analysis. We demonstrate our ligand-binding screening method with two proteins, the transcription-factor-binding domain of the CREB-binding protein, KIX, and the bromodomain Brd4 (Figure 3.2). Application of PrOF NMR with a third protein, the first bromodomain of BrdT, will be subsequently described to highlight several additional important aspects of the PrOF NMR protocol. The sequence-selective incorporation of the three fluorinated aromatic amino acids, 3FY, 4FF and 5-fluorotryptophan (5FW), into recombinant proteins in Escherichia coli using either the auxotrophic bacterial cell lines (e.g., DL39(DE3)) or standard bacterial strains (e.g., BL21(DE3)) will first be detailed, as previously described.[95,96,99] This section of the PROCEDURE will then be followed by our ligand-discovery procedures.

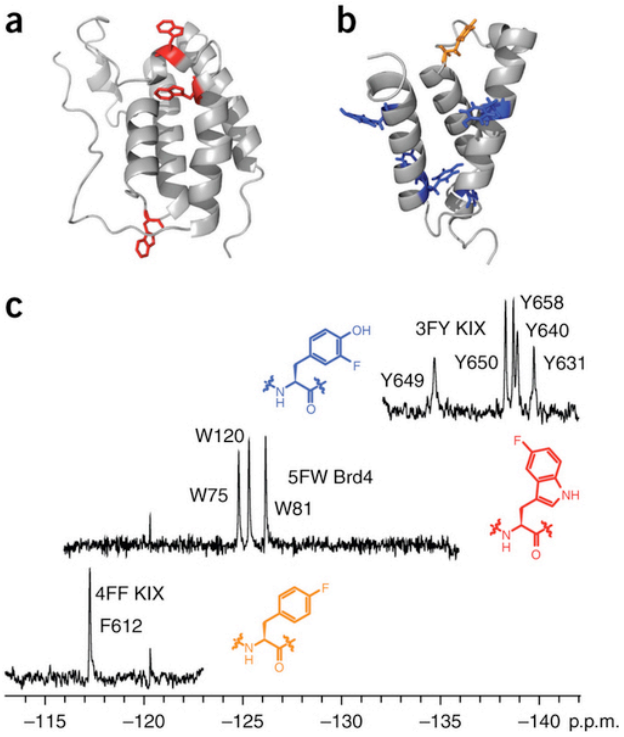


Figure 3.2 PrOF NMR examples.
(a) Crystal structure of Brd4 (PDB 3MXF) with its three tryptophan residues shown in red. (b) Crystal structure of KIX (PDB 1KDX). (c) PrOF NMR spectra for each of these variants, in which the residues highlighted in a and b were replaced by 5FW (Brd4), 3FY (KIX) or 4FF (KIX). In these spectra, the wide chemical shift dispersion of the signals due to the fluorine-labeled, aromatic amino acid analogs are highlighted. The structures of the mentioned fluorinated amino acids are shown next to their corresponding spectra.

3.1 **Experimental design**

Fluorinated protein expression and characterization. All expressed proteins in the PROCEDURE are characterized for purity and fluorine incorporation via SDS–PAGE and protein mass spectrometry via electrospray ionization. Protein yields vary on the basis of the protein system, fluorinated amino acid and the cell line used. We have achieved yields as high as 70 mg/l for 3FY-labeled KIX and 65 mg/l for 4FF-labeled KIX using auxotrophic DL39(DE3) cells, and obtained up to 62 mg/l 5FW-labeled KIX[23] (C.T.G., unpublished data) using non-auxotrophic BL21(DE3) cells with 5-fluoroindole added to the cell culture medium. All expressions led to a high labeling efficiency. We observed even higher yields with our fluorinated bromodomains (84 mg/l for 5FW BPTF

and 88 mg/l for 5FW Brd4).[32] The SDS–PAGE gel in Figure 3.3 shows four KIX protein samples that are unlabeled, 3FY-labeled, 4FF-labeled or 5FW-labeled.

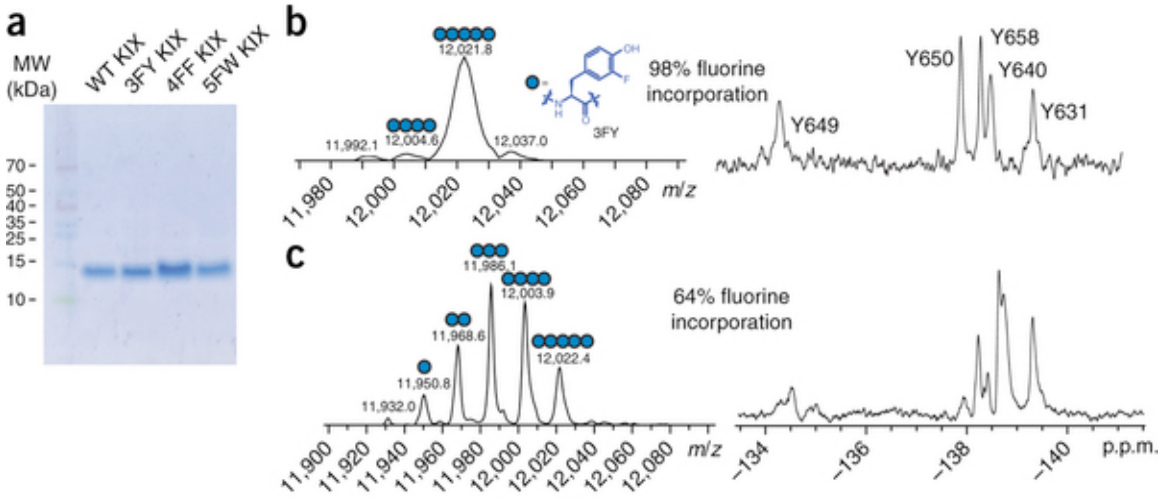


Figure 3.3 Characterization of fluorinated proteins expressed according to present protocol.
(a) SDS–PAGE gel stained with Coomassie blue showing wild-type and three fluorinated variants of KIX. (b) Deconvoluted electrospray ionization mass spectrometry (ESI-MS) spectrum (left) and ^{19}F NMR spectrum (right) of 3FY-labeled KIX with 98% fluorine incorporation. The dominant population present is the fully fluorinated variant. (c) Deconvoluted ESI-MS spectrum (left) and ^{19}FNMR spectrum (right) of 3FY-labeled KIX with 64% fluorine incorporation.[127]

Mass spectrometry is used to assess fluorine incorporation into the protein. For the mass spectrogram of 3FY-labeled KIX, the dominant mass of 12,021.8 Da corresponds to the incorporation of a fully labeled protein (five 3FY residues). The minor mass of 12,004.6 Da corresponds to four of the five tyrosine residues being replaced with 3FY. These major and minor populations lead to a 98% labeled protein (Figure 3.3). Low levels of incorporation can result in a heterogeneous protein sample, which can complicate the analysis of ^{19}F NMR spectra. A ^{19}F NMR spectrum of 3FY-labeled protein that is only 64% labeled is reported for comparison (Figure 3.3). The resulting spectrum is a statistical mixture of multiply labeled proteins, and it can result in additional (e.g., Y650) or broadened resonances. In some instances, if needed, the concentration of fluorinated amino acids can be increased in the culture medium with respect to those recommended in the PROCEDURE, to increase the extent of protein

labeling. Low levels of labeling can also be attributed to residual unlabeled amino acid from the initial expression conditions, an inconvenience that can be reduced with careful washings of bacterial pellets and preincubations with fluorinated amino acids.

Fluorine is not an element that is naturally found in proteins. Therefore, before the implementation of the screening protocol, the structural and functional perturbation caused by the introduction of the fluorine-based label into the protein must be carried out. We have characterized our proteins by a variety of methods, using X-ray crystallography, circular dichroism and thermal stability measurements to assess structure.[99,131] When a ligand was known beforehand, we have also used isothermal titration calorimetry or fluorescence anisotropy ligand-binding experiments to compare the affinity of the known ligand with those of the fluorinated and nonfluorinated proteins. A direct binding experiment with a fluorescently labeled bromodomain ligand, BI-BODIPY, with 5FW-labeled Brd4 and unlabeled Brd4 yielding dissociation constants of 110 and 55 nM, respectively. In our experience, we have considered a two- to threefold change in binding affinity between fluorine-labeled and unlabeled protein to be acceptable for continuing on with the labeled protein in a ligand screen. We recommend trying alternative labeling approaches or using different amino acids (e.g., 6-fluoro versus 5-fluorotryptophan) if larger perturbations are observed.

PrOF NMR. PrOF NMR spectra of small- to medium-sized fluorinated proteins typically reveal well-resolved resonances for each labeled aromatic amino acid, but not in all cases. The different aromatic residues are seen in close but distinct chemical shift regions (Figure 3.2). A well-dispersed NMR spectrum is an additional confirmation of a well-folded protein. Protein resonances tend to be broad. Fluorine resonances from small-molecule impurities or peptides resulting from proteolytic degradation appear as sharp signals in ^{19}F NMR spectra, and their presence can help assess the stability and purity of the protein. A resonance due to a small-molecule impurity can be seen in Figure 3.2 at −120 p.p.m. In some cases, protein precipitation or aggregation leads to loss of signal for all of the protein resonances, which can be useful for the identification of false positives in a ligand screen.

Resonances can be assigned by recording the spectrum of a fluorinated protein for which one residue has been mutated to an amino acid that can no longer be fluorine-labeled in the expression conditions because the corresponding fluorinated analog is not present in the culture medium (e.g., Trp to Phe). As an example, three relevant spectra are reported in Figure 3.4. In this case, the bromodomain Brd4 is labeled with three 5FW residues, which results in three well-resolved resonances. The disappearance of the resonances at −124.9 and −126.2 p.p.m. upon directed mutation of specific tryptophan residues leads to the assignment of these resonances to fluorine-labeled W75 and fluorine-labeled W81 (Figure 3.4, middle and top spectra). However, because of the sensitivity of fluorine to subtle changes in environment, additional chemical shift changes can occur, which, in some cases, preclude assignment. In this case, the addition of a known ligand can help identify side-chain resonances in the ligand-binding site. For example, the first bromodomain of BrdT has two tryptophan resonances, and W50 is known to be located in the 'WPF shelf' at the histone-binding site. Addition of the known BrdT ligand dinaciclib (IC_{50} = 61 μM) was used to help assign the resonance due to the fluorine-labeled W50. The broadening and the 0.25-p.p.m. chemical shift perturbation of the upfield resonance is consistent with X-ray crystallography data on the co-crystal structures, which show a conformational change of W50 upon binding[28] (Figure 3.5). Chemical shift perturbation upon the addition of a known ligand can also serve as a useful confirmation of the functional integrity of a fluorine-labeled protein.[99,131] Fragment screening can be carried out once a well-folded, highly labeled protein has been obtained, although the protein NMR spectrum does not necessarily need to be assigned first.

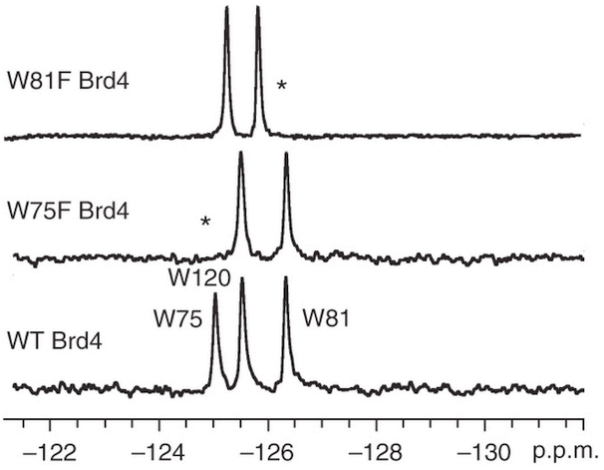


Figure 3.4 Using site-directed mutagenesis to assign PrOF NMR resonances. Asterisks denote the resonance that has disappeared after mutation of the relevant (fluorine-labeled) tryptophan into a phenylalanine, which allows assignment of that resonance to the specific mutated amino acid residue.[131]

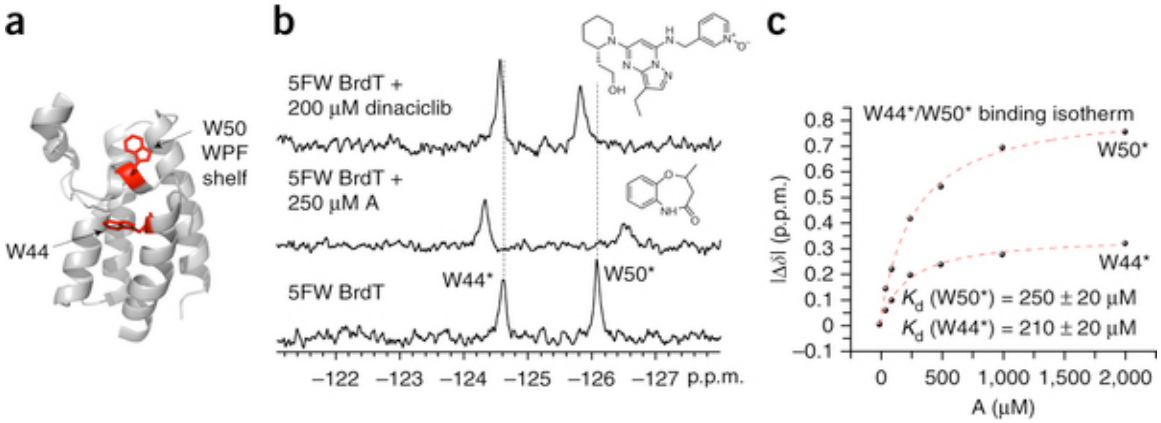


Figure 3.5 Results from experiments involving small-molecule ligands for BrdT. (a) Crystal structure of BrdT (PDB 4FLP) showing the two tryptophans. Please note that W50 is located in the protein's WPF shelf. (b) PrOF NMR spectra of 5FW-BrdT by itself (bottom spectrum) and in the presence of two different small molecules, fragment A, middle spectrum (see also the structure of the compound above the spectrum, to the right-hand side), and dinaciclib, top spectrum (see also the structure of the compound above the spectrum, to the right-hand side). (c) Binding isotherms for fragment A generated by plotting the change in chemical shift for both 5FW resonances as a function of ligand concentration, which yield comparable K_d values. *Resonance assignments that were inferred from small-molecule binding.

Dissociation constant and LE determination. The affinity of fragment hits from the screen can be readily assessed for protein ligands with up to millimolar dissociation constants via ligand titration and subsequent nonlinear regression of the binding isotherm

produced by monitoring changes in chemical shift. This chemical shift perturbation experiment is valid when the molecules are in fast exchange, as indicated by the presence of a single resonance signal whose chemical shift is the weighted average of the bound and unbound states. The speed of the experiment, the low concentration of proteins needed for it and the possible automation of the procedure enable the rapid testing of multiple small molecules.

In the present approach, separate samples are set up with varying concentrations of the small molecule. The dissociation constant (K_d) is obtained by fitting the obtained data to a nonlinear regression curve using the following equation, accounting for receptor depletion.[130] Analysis of the affinity of the small molecule and location of perturbed resonances can then be used to assess valuable structure–activity relationships between molecules and differential binding to various surfaces on a protein.

$$S = D * \frac{(K_d + L + P) - \sqrt{(K_d + L + P)^2 - 4PL}}{2P}$$

In equation 1, S is the observed change in chemical shift, D is the maximum change in chemical shift, K_d is the dissociation constant of the ligand, L is the ligand concentration and P is the protein concentration. The maximal shift, D, can be obtained through nonlinear regression, or experimentally, when further addition of a small molecule no longer perturbs the resonance. In addition, as a means of prioritizing ligands for future development, these affinity data can be used to calculate the LE using the following equation, which enables this efficiency to be calculated on the basis of the binding affinity of the ligand and the contribution of each nonhydrogen atom to the free energy of the interaction.

$$LE = \frac{-RT\ln(K_d)}{\text{Nonhydrogen atoms}}$$

Although this scenario is not commonly encountered in fragment screens, ligands with high affinity may have a sufficiently long residence time on the protein that the bound and unbound states are partially resolved (intermediate exchange) or fully resolved (slow exchange) during the ^{19}F NMR experiment. In these instances, titration will not result in a binding isotherm. Intermediate exchange kinetics results when interchanging

species at equal populations (e.g., bound and unbound states) lead to coalescence into a single resonance.[134,135]

$$\tau_{\text{coalescence}} = (\sqrt{2}\pi\Delta\nu)^{-1} = k_{-1}$$

In the coalescence equation, the frequency difference in Hertz between the two resonances ($\Delta\nu$)—bound versus unbound state—establishes a relationship that enables researchers to estimate the residence time of the ligand on the protein based on the rate of dissociation (k_{-1}). In the case in which the rate of exchange approaches the coalescence point ($\tau \approx \Delta\nu$) during a protein–ligand titration, the resonances will coalesce so that the signal will initially broaden, but it will become sharp again, in a dose-dependent manner, as the chemical shift is perturbed by the titration process. With slow exchange kinetics (likely to be associated with a low dissociation constant), both resonances can be well resolved, and, during a protein–ligand titration, one resonance will gradually disappear while another will grow from the baseline in a dose-dependent manner.

Experimental considerations and assay limitations. The additional rotational freedom in protein side chains versus the amide backbone can lead to smaller linewidths relative to amide resonances. One consideration that should be made before using ^{19}F-labeled proteins for NMR is the sensitivity of the fluorine nucleus to substantial chemical shift anisotropy (CSA) relaxation effects. CSA relaxation leads to resonance broadening, which is enhanced by protein dynamics, including long rotational correlation times of amino acids found in large proteins.[136] CSA relaxation is proportional to the square of the magnetic field strength. From a practical standpoint, we carry out NMR experiments at 471–564 MHz using fluorinated aromatic amino acids. 5FW experiences a lower degree of CSA relative to other fluorinated tryptophan analogs, and in some cases it has been used with proteins as large as 65 kDa.[137] As an alternative to 5FW, fluorine-labeled cysteine or methionine derivatives with substantially reduced CSA effects are recommended for proteins larger than 40 kDa, as has been demonstrated in experiments with G-protein-coupled receptors and amyloid-β aggregates.[89,138–140] Site-selective fluorine labeling with trifluoromethylphenylalanine has also been used for large proteins.[141,142] For small- to medium-sized proteins, PrOF NMR is applicable using a variety of fluorine-labeled amino acids, including singly fluorinated (tyrosine,

phenylalanine and tryptophan, Figure 3.2) and more heavily fluorinated amino acids such as di- and trifluoromethylmethionine,[143,144] and tri- and hexafluoroleucine.[145–147] We propose using singly fluorinated aromatic amino acids, as this approach reduces spectral complexity with respect to using amino acids with multiple nonequivalent fluorines. Furthermore, it allows high levels of enrichment in labeled amino acids at protein–protein interaction interfaces[100] while minimizing the perturbing effect of fluorine substitution. Nevertheless, we reiterate that protein structural and functional perturbation caused by fluorine substitution should always be assessed by conducting preliminary complementary structural and functional biophysical experiments.

3.2 **Materials**

Reagents: Unless noted, reagents of comparable quality from alternative suppliers can be substituted.

Protein Expression Components

- Fragment mixtures. <CRITICAL> These are prepared according to guidelines outlined in Box 1. Sources of fragment libraries include Maybridge and Chembridge
- Plasmid encoding your protein of choice that also encodes for the resistance to an antibiotic of choice

CRITICAL: For nickel affinity purification**,** ensure that the plasmid includes a His tag described in step 9 and in the appendix.

- Plasmid vectors pRARE (Novagen), pRSETB-HIS6KIX (Invitrogen), pNIC28-BSA4 (Addgene, used in Supplementary Method)
- Competent DL39(DE3)* (DL39, CGSC) or BL21(DE3) (Novagen) *Escherichia coli* cell lines
  Note: We thank I. Ropson for providing the DL39(DE3) cell line.

Critical: Choice of cell lines must be made carefully. Auxotrophic cell lines (e.g., DL39(DE3)) are necessary for labeling with 3-fluorotyrosine or 4-fluorophenylalanine. For tryptophan labeling, 5-fluoroindole with non-auxotrophic cell lines (e.g., BL21(DE3)) is recommended for higher protein yields.

- QiAprep Spin Miniprep Kit (250) (Qiagen, Cat. No. 27106)
- Lennox L. Broth (LB) (RPI, powdered or granulated, L24066)Lennox L Agar (RPI, powdered or granulated, L24030)
- Super Optimal Broth with Catabolite Repression (SOC) (SOB with added glucose) (RPI, S25000)

- Antibiotics: Ampicillin (RPI, A40040), Kanamycin (RPI, K22000), Chloramphenicol (RPI, C61000)
- Natural amino acids (Table 3.1) (RPI)
- Isopropyl β–d-1-thiogalactopyranoside (IPTG)(RPI, I56000)
- Imidazole (RPI, I52000)
- Sodium Phosphate Monobasic, monohydrate (RPI, S23120)
- Sodium Phosphate Dibasic, heptahydrate (Sigma, S9390)
- Sodium chloride(Fisher, S2713)**.**
- HEPES (Fisher, BP310)
- Tris (Acros, 167620010)
- 3-fluoro-DL-tyrosine (Alfa Aesar, L01479)
- 4-fluoro-DL-phenylalanine (Sigma Aldrich, F5251)
- 5-fluoro-indole (Sigma Aldrich, F9108)
- Ultrapure water 18.2 MΩ x cm

Protein Purification Components

- Lysozyme (Gold Biotech, L040)
- β-mercaptoethanol (Sigma Aldrich, M6250)
- Phenylmethanesulfonylfluoride (PMSF), (RPI, P20270)

NMR Components

- Dimethylsulfoxide (Fisher, D128)
- Ethylene Glycol (Sigma Aldrich, 102466)
- Trifluoroacetic acid (TFA, Sigma Aldrich, T6508)
- Deuterium oxide (Cambridge Isotope Labs, DLM-6)

Prepared Buffer Solutions (Filtered and stored at room temperature)

- Lysis Buffer: 50 mM Phosphate, 300 mM NaCl, pH = 7.4
- Wash Buffer: 50 mM Phosphate, 100 mM NaCl, 30 mM Imidazole, pH = 7.2
- Elution Buffer: 50 mM Phosphate, 100 mM NaCl, 400 mM Imidazole, pH = 7.2
- NMR Buffer for Brd4: 50 mM Tris, 100 mM NaCl, pH = 7.4
- NMR Buffer for KIX: 50 mM HEPES, 100 mM NaCl, pH = 7.2

Protein Characeterization/SDS-PAGE Components:

- Formic Acid (Sigma, 56302)
- Acetonitrile (J.T. Baker, 9853)
- 40% Acrylamide/Bis solution (Bio-Rad, 1610148)
- Ammonium persulfate (Sigma Aldrich, 248614)
- Bis-Tris (RPI, B75000)
- Tetramethylethylenediamine, (TEMED) (RPI, T18000)
- Coomassie Brilliant Blue R-250 (Bio-Rad, 161-0400)

Equipment

- Refrigerated Incubator Shaker
- UV/Vis spectrophotometer (e.g., Varian Cary 50 Bio UV/Vis spectrophotometer)
- High speed centrifuge
- Sonicator (Fisher FB505) with 1/8" microtip (FB 4418)
- FPLC (GE Äktapurifier)
- HiPrep 26/10 Desalting Column (GE Healthcare 17-5087-01)
- 50 mL Superloop (GE Healthcare 18-1113-82)
- Ni-NTA Agarose Beads (Life Technologies R901-01)
- NMR spectrometer (Recommended ^{19}F S/N ≥ 550:1)

Critical: NMR probe must be able to tune to the ^{19}F nucleus (e.g. Bruker BBO probe), For optimal PrOF NMR experiments, a cryoprobe (e.g. Bruker TCI-Prodigy) should be used for increased S/N.)

- ^{19}F tuned NMR probe, e.g. (TCI-Prodigy cryoprobe, Bruker)
- Electrospray Ionization Mass Spectrometer

Critical: Mass spectrometer must be capable of resolving proteins to 2-3 amu for accurate protein characterization as fluorinated proteins will differ by 18 amu per incorporated fluorine.

- SDS PAGE gel apparatus and power supply (for use in protein characterization
- NMR tubes
- Microcentrifuge tube
- 15-ml Falcon Tubes
- 50-ml Falcon Tubes
- Protein concentrators (Millipore Centricon UFC900324, UFC800324)

Software

- NMR processing software (e.g. TopSpin, MNova, etc…)
- Non-linear regression software (e.g. OriginPro, GraphPad, etc…)

Reagent Setup

LB Media/Agar: Autoclave media prior to using.

Preparation of the defined media

To simplify the preparation if it is going to be done multiple times, a sterile-filtered stock solution consisting of the reagents listed in Step 3 ($CaCl_2$ through biotin) may be

prepared and stored at 4°C for several months; an appropriate volume of stock solution would then be added to the defined media instead of adding the components individually.

1. Combine the amino acids, salts, and nucleotide bases listed below in a total volume of 1 L of deionized water, omitting the amino acid that will be labeled, and autoclave the solution. Natural amino acids were all purchased from RPI.

   Pause Point: If the defined media is not going to be used right away, do not continue media preparation beyond this step. Media with just the amino acids and salts with no carbon source (glucose) or vitamins can be stored covered at room temperature for several months.

Table 3.1

Component	Supplier (Cat. #)	Amount for 1L expression
Alanine	RPI (A20060)	500 mg
Arginine	(A50010)	400 mg
Asparagine	(A50030)	400 mg
Aspartic Acid	(A50060)	400 mg
Cystine HCl	(C81020)	50 mg
Glutamine	(G36040)	400 mg
Glutamic Acid	(G36020)	650 mg
Glycine	(G36050)	550 mg
Histidine	(H75040)	100 mg
Isoleucine	(I54020)	230 mg
Leucine	(L22000)	230 mg
Lysine HCl	(L37040)	420 mg
Methionine	(M22060)	250 mg
Phenylalanine	(P20260)	130 mg
Proline	(P50200)	100 mg
Serine	(S22020)	2.1 g
Threonine	(T21060)	230 mg
Tyrosine	(T68500)	170 mg
Valine	(V42020)	230 mg
Sodium acetate	Macron (7372)	1.5 g
Succinic acid	Sigma Aldrich (S3674)	1.5 g
Ammonium chloride	Fisher (A661)	500 mg
Sodium hydroxide	Alfa Aesar (A16037)	850 mg
Potassium phosphate (dibasic)	RPI (P41300)	10.5 g
Adenine	RPI (A11500)	500 mg
Guanosine	Alfa Aesar (A11328)	650 mg
Thymine	TCI (T0234)	200 mg
Uracil	RPI (U32000)	500 mg
Cytosine	Alfa Aesar (A14731)	200 mg

2. Once the solution is sterilized, add the components sequentially, adding $MgSO_4$ last

Component	Supplier (Cat. #)	Amount for 1-L expression
40% (wt/vol) glucose solution	RPI (G32030)	50 ml
0.01M $FeCl_3$	Sigma Aldrich (157740)	1 ml

Table 3.2 Vitamin solution composition

Vitamin Solution		
Component		Amount for 1-l expression
Tryptophan	RPI (T60080)	50 mg
$CaCl_2$ dihydrate	Sigma Aldrich (C3306)	2 mg
$ZnSO_4$ heptahydrate	Sigma Aldrich (Z4750)	2 mg
$MnSO_4$ monohydrate	Sigma Aldrich (M7634)	2 mg
Thiamine	RPI (T21020)	50 mg
Niacin	Sigma Aldrich (72340)	50 mg
Biotin	RPI (B40040)	1 mg
1M $MgSO_4$	Sigma Aldrich (M7506)	4 ml

3. Adjust the pH to 7.2 using HCl or NaOH. Media can be stored at room temperature for a few weeks, though it is recommended that the glucose and vitamin solutions not be added until immediately prior to use (Table 3.2).

4. Add the necessary fluorinated amino acid(s) or amino acid precursor for appropriate fluorine labeling.

Table 3.3 Fluorinated amino acids or precursor for defined media.

Fluorine Label	Fluorinated Substitution	Amount for 1-L expression	Final concentration in media
3FY	3-fluoro-DL-tyrosine	80 mg	400 μM
4FF	4-fluoro-DL-phenylalanine[a]	29 mg	160 μM
5FW	5-fluoroindole	60 mg	444 μM

[a]Addition of 5 μM of phenylalanine to the defined media when using 4-fluoro-DL-phenylalanine has led to increased protein yield but minimal effects on fluorinated amino acid incorporation.

Equipment Setup

Sonicator Program. Sonication time = 4 minutes at 30% amplitude (this entails eight x 30 second pulses with 60 second rest times between pulses.) Total elapsed time: 12 minutes

Ni Affinity Gradient Elution Method. Set up the FPLC method that will be used in Step 9. In our laboratory we use a GE Äktapurifier with a column packed with Ni-NTA Agarose for affinity purification. The four steps are as follows: Wash the nickel column with 15 column volumes of wash buffer. Perform a gradient elution across 20 column volumes ramping from 0 to 100% elution buffer. Wash with 5 column volumes of 100% elution buffer. Return to 0% elution buffer by applying the reverse gradient in 5 column volumes. UV absorbance is monitored at 280 nm. Total time (at 1 mL/min flow rate): 4 hours

Buffer Exchange Method. Set up the FPLC method that will be used in step 10. In our laboratory, we use a GE Äktapurifier with a GE HiPrep 26/10 Desalting column (17-5087-01). Equilibrate the column with 0.5-1 column volume of buffer prior to loading the sample. Elute the column with 1 column volume of buffer and monitor the UV absorbance at 280 nm. Total time (at 2 mL/min flow rate): 40 minutes.

3.3 **Procedure**

Bacterial Transformation (Timing: 1 day)

1. Prepare bacterial colonies transformed with the plasmid for the protein of interest using standard transformation and inoculation methods. A detailed sample protocol is provided in Supplementary Method.

Critical As noted in the materials section, choice of cell lines must be made carefully. Auxotrophic cell lines (e.g., DL39(DE3)) are necessary for labeling with 3-fluorotyrosine or 4-fluorophenylalanine. For tryptophan labeling, 5-fluoroindole with non-auxotrophic cell lines (e.g., BL21(DE3)) is recommended for higher protein yields.

Pause Point Culture plates with transformed bacteria may be stored at 4 °C for up to a month.

Fluorinated protein expression Timing: 2 days

Critical: Maintain sterile conditions while working with media. Standard protein expression methods can be used; these steps are those that are used routinely in our laboratory and provide information about when in the process, the fluorinated amino acids should be added.

2. Inoculate primary and secondary cultures using standard *E. coli* expression methods. Detailed steps of a sample protocol are provided in the appendix.

3. Using a spectrophotometer, determine the optical dispersion at 600 nm (OD_{600}). Remove the culture from the shaker when the OD_{600} is between 0.6 and 0.8. Please note that this value may vary based on the cuvette distance to the detector in the spectrophotometer.

4. Centrifuge the culture for 20 min at 6000 *g* at 4 °C.

5. Decant the LB media and resuspend the pellet in an equivalent amount of the defined media containing the desired fluorinated amino acid or amino acid precursor [see Reagent setup] and appropriate antibiotic.

6. Shake the solution prepared in the previous step at 37 °C and 250 RPM for 90 min as a recovery time for the bacteria then decrease the temperature to 20 °C to cool down the media. Continue to shake for an additional 30 min to allow the solution to equilibrate. Please note that the most suitable temperature and recovery time will vary for each

protein. Our lab has generally found the above conditions to work well, though there are cases when a shorter recovery time has produced better results.

7. Induce protein expression by adding IPTG to the solution to a final 1-mM concentration. Continue to shake at 20 °C and 250 RPM for 16–20 h. Please note that the most suitable IPTG concentration and induction time will vary for each protein.

8. Centrifuge the cell culture at 6000 *g* at 4 °C for 20 min. For ease of purification, we recommend centrifuging cultures in 500-ml aliquots or smaller. Decant the supernatant medium and store the cell pellet at –20 °C or –80 °C.

Pause Point The cell pellet may be stored at –20 °C or –80 °C for months.

9. Purify and characterize overexpressed protein. Sample purification methods are described in the appendix and mass spectrometry characterization is described in **Box 1**. SDS PAGE characterization can be performed at this step to evaluate the presence of the desired protein. However, due to the low resolution nature of SDS PAGE characterization, it will not be able to quantify fluorine incorporation or distinguish between fluorinated variants. (Figure 3.3)

10. Buffer exchange the protein into the desired buffer. In our lab we use a HiPrep desalting column, but other methods work as well (Nap-5 columns, PD-10 columns, and dialysis are all viable options). For PrOF NMR, it is best to avoid buffers with high concentrations of high mobility salts due to their impact on NMR signal sensitivity. When possible, the lower mobility salts (e.g. Tris, HEPES, etc…) are preferred.

Pause Point: The purified protein can be stored at 4 °C or flash-frozen and stored at –20 °C for several months. For long-term storage, we recommend flash-freezing and storing at –20 °C. Exact storage conditions may vary from one protein to another.

Box 1 – Fluorine Incorporation Characterization Timing 0.5-1 h

1. Concentrate protein to at least a low µM concentration. In our lab we utilize centrifugal protein concentrators with either a 3k or 5k MW cutoff. The exact membrane cutoff needed will depend on the size of the protein of interest

2. On an ESI mass spectrometer equipped with a liquid chromatography system, use 0.1% formic acid in water (v/v) and acetonitrile as solvents and a C18 column for separation. Run a gradient elution ramping from 8% acetonitrile to 80% acetonitrile.

3. Select the protein peak on the chromatogram and deconvolute the corresponding mass spectrum.

4. Integrate the peaks corresponding to the protein of interest and its fluorinated variants.

5. Use the following formula to calculate the percent of fluorine incorporation.

$$\%\ Incorporation = \frac{(0\ F\ protein * 0) + (1F\ protein * 1) + \cdots (nF\ protein * n)}{(0F\ protein * n) + (1F\ protein * n) + \cdots (nF\ protein * n)} * 100$$

Where n corresponds to the number of incorporated fluorinated residues.

Protein-Observed Fluorine (PrOF) NMR Timing 5 min–1 h

11. Concentrate the protein solution to 40–50 µM as done in **Box 1**

12. To a microcentrifuge tube, add 2 µl of 0.1% (v/v) TFA, 25 µl of D_2O, and 473 µl of protein solution just prepared. Mix well and transfer solution to a 5-mm NMR tube.

13. Acquire two fluorine NMR spectra. Focus the first experiment on the TFA reference peak, which can be obtained within several scans, and focus the second experiment on the protein. We find that a spectral width (sw) of 10–20 ppm is sufficient with an offset (O1P/tof) of –76.5 ppm for the TFA, –136 ppm for 3FY-labeled proteins, –125 ppm for 5FW-labeled proteins, or –117 ppm for 4FF-labeled proteins. Experiment time and

number of scans required will be dependent on availability of a cryoprobe versus a room-temperature probe as well as protein On a 500 MHz NMR with a Prodigy inverse cryoprobe (19F S:N 2100:1) 400 scans are sufficient for proteins KIX and bromodomain Brd4, which would lead to a 4–5-min experiment.

14. Process the data setting the TFA reference peak to –76.5 ppm and applying the same correction factor to the second experiment. Resonance assignments can be performed as described in **Box 2**.

(Optional) Assigning PrOF Resonances Timing ~3 weeks

Note: PrOF NMR screening and binding experiments can still be performed without resonances assigned. The assignments can provide additional structural information but are not necessary to evaluate small molecule binding.

1. Perform site-directed mutagenesis to mutate each amino acid of interest to an alternate amino acid chosen such that it will have very little effect on the protein structure (e.g., Y->F)
2. Obtain PrOF NMR spectra of each mutant (see Steps 11-14 of the Procedure). The resonance that has disappeared corresponds to the mutated amino acid residue.

Fragment Mixture Preparation Timing Variable

1. Obtain or prepare concentrated ligand stock solutions (200 mM) in a given solvent (e.g., DMSO or ethylene glycol) for the fragments of interest. Please note that solvent choice will depend on ligand solubility and solvent effects on the protein of interest.

Critical Step The maximum volume of solvent that will be tolerated without significantly perturbing the chemical shifts or the shapes of the ^{19}F NMR resonances must be determined in order to accurately assess ligand binding. We recommend that the final solution is composed of 1–5% organic solvent.

2. Determine the desired number of fragments for each screening mixture. Commonly screening mixtures include 5–10 fragment compounds. Fragment mixtures of five of six

compounds will reduce the number of compounds needed in the deconvolution step, if a high hit rate is anticipated.

3. Decide the concentration at which each fragment will be screened. Fragment mixture concentrations will depend on the expected K_d's for ligands in the mixture. Due to the low affinity nature of fragment compounds, mid-μM to low-mM K_ds for individual fragments are common. For example, using six fragments per mixture starting from 200 mM DMSO stock solutions yields mixtures containing ligands at 33.3 mM and a final ligand concentration of 833 μM at 2.5% (v/v) DMSO in the NMR sample. Screening ligands at lower concentrations will yield smaller changes in chemical shift, potentially resulting in more false negatives. If the ligandability and/or druggability of the protein is unknown, a pilot screen with ligands at various concentrations can be performed.

4. Compile the fragment mixtures, taking into account the number of acidic, basic, and neutral compounds present to avoid significant pH dependent effects.

Fragment screening via PrOF NMR Timing Variable

15. Prepare a blank sample (See step 10) with the addition of the selected amount of the organic solvent for the fragment mixtures (no ligand or ligand mixtures), maintaining a total sample volume of 500 μL. (e.g. 2 μL 0.1% (v/v) TFA, 25 μL D_2O, 468 μL protein solution, and 5 μL DMSO)

Critical Step This blank sample must be identical to the screening samples in regards to the protein concentration and amount of solvents used. The only difference should be that the blank contains organic solvent (with no ligands) while the screening samples will contain the same amount of organic solvent (with the ligand mixtures).

16. Prepare the fragment mixture NMR samples in a similar manner to Step 15 with the ligand mixture in place of just the organic solvent.

17. Acquire and process PrOF NMR spectra for each mixture (see Steps 13 and 14)

18. Record the chemical shift of each fluorine resonance and calculate the change in chemical shift for each resonance relative to the 'blank' spectrum.

19. Using the results obtained in the library screen, identify promising mixtures by statistically analyzing the changes in chemical shifts for each resonance from each mixture, and select the mixtures that yield a change in chemical shift between one and two standard deviations above the average change in chemical shifts from all experiments obtained from the library screen.

Deconvolution of fragment mixtures Timing Variable

20. Identify the compounds comprising the mixtures that yielded significant changes in chemical shift.

21. Prepare NMR samples with each of these compounds separately maintaining equivalent ligand concentrations.

22. Collect the PrOF NMR spectra for each new sample (see Step 13)

23. Process and analyze the data as before calculating the changes in chemical shift to identify the ligand or ligands that bind to the protein target. (see Steps 14 and 19)

K_d determination of fragment compounds Timing Variable

24. Select a range of ligand concentrations for the binding isotherm being sure to include points below and above the anticipated Kd.

25. Prepare NMR samples with varying concentrations of ligands, making sure to add solvent until the final volume of ligand solution is equal to the total amount of solvent

that was added to the blank prepared in step 15. (1 µL ligand/4 µL DMSO, 2 µL ligand/3 µL DMSO, etc…)

26. Mix the solutions well and transfer them to NMR tubes.

27. Collect PrOF NMR spectra for each sample and analyze data (see Steps 13, 14 and 18).

28. Plot the change in chemical shift (Δδ) as a function of ligand concentration.

29. Fit the data using non-linear regression software (e.g. Originpro or Graph Pad Prism) the following equation to solve for K_d.

$$S = D * \frac{(K_d + L + P) - \sqrt{(K_d + L + P)^2 - 4PL}}{2P}$$

Timing

Steps 1, Bacterial Transformation: 1 d

Steps 2-10, Fluorinated Protein Expression and Purification: 2 d (plus characterization time; see **Box 1**)

Steps 11-14, Protein Observed Fluorine (PrOF) NMR: 5 min-1 h (plus resonance assignments; see **Box 2**)

Steps 15-19, Fragment Screening via PrOF NMR: Variable (plus fragment mixture preparation time; see **Box 3**)

Steps 20-23, Deconvolution of Fragment Mixtures: Variable

Steps 24-27, K_d determination of Fragment Compounds: Variable

Troubleshooting

Troubleshooting advice can be found in Chapter 7

During fragment screening, resonance assignments facilitate the characterization of the binding site of small molecules and of the environmental changes experienced by

the labeled protein side chain. Fragment-mixture hits can result in a large change in chemical shift (e.g., 0.1–1 p.p.m.) due to the responsiveness of fluorine to changes in its chemical environment. The ^{19}F NMR spectra shown in Figure 3.6 show that a peak at −134.5 p.p.m. observed in the spectrum of the fluorinated protein 3FY-KIX undergoes a change in chemical shift of 0.32 p.p.m. in the spectrum of the same protein in the presence of a mixture of different potential ligands.

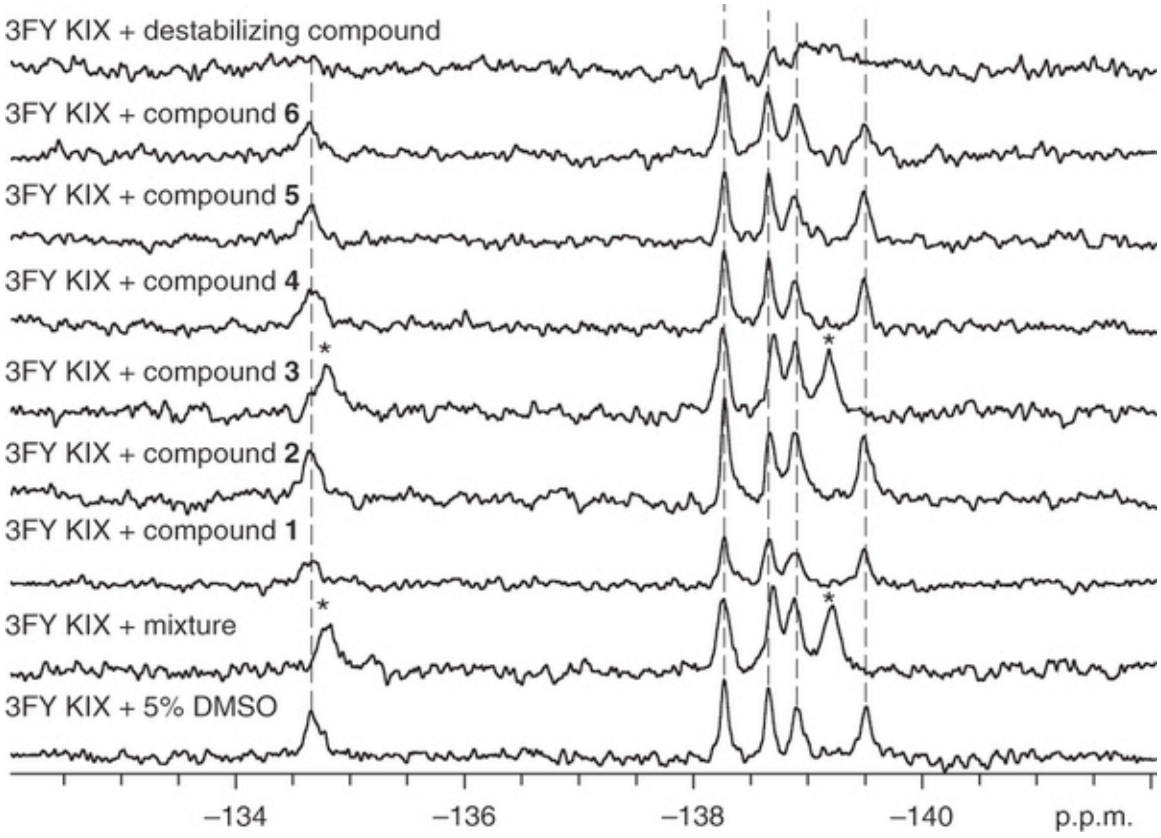


Figure 3.6 Deconvolution of fragment mixture.
The bottom spectrum corresponds to the fluorinated protein with no ligands added. Asterisks denote resonances that have been significantly perturbed. In this particular mixture, compound 3 is responsible for the chemical shift perturbation seen in the mixture. The top spectrum represents a global reduction in signal due to nonspecific effects.

Protein aggregators and denaturants, whose presence is common in many screens, can be detected by global analysis of all of the resonances present in the protein. Global coalescence, broadening or a decrease in intensity are all indications of nonspecific effects induced by the small molecule on the protein, which may lead to false positives in

ligand-observed experiments (e.g., top spectrum in Figure 3.6). Active molecules within a mixture are then identified by testing the molecules from the fragment mixture hit one at a time. Figure 3.6 shows an example of such a 'deconvolution', which helps to identify compound 3 as the active compound, i.e., the compound that caused the change in chemical environment sensed by the fluorine labels in the protein.

Chemical shifts are easily measured and highly reproducible. In particular, 3FY-KIX, across 23 replicate experiments, had a standard deviation in chemical shift over five resonances from 0.011 to 0.037 p.p.m. The latter measurement was from the broadest resonance, the one assigned to fluorine-labeled Y649. Although only one resonance was observed for fluorine-labeled Y649, the potential for rotamers of the ortho-substituted phenol, from restricted motion in this environment, leads to a broadened resonance. In the case of 5FW-labeled Brd4, for which the resonances are sharper, the standard deviation range for the fluorine resonances was narrower, from 0.009 to 0.026 p.p.m. In addition, although aromatic amino acids are commonly found at protein binding sites,[100] labeled amino acids that are far away from the binding site serve as important internal control resonances for nonspecific binding and other factors that may globally affect the chemical shift. In the 229-compound screen with 5FW-labeled Brd4, the average change in chemical shift ($\Delta\delta$) of fluorine-labeled W120, which is located outside the ligand binding site, was 0.001 ± 0.014 p.p.m. However, in the case of W81, which is located at the ligand binding site, the average $\Delta\delta$ of W81's fluorine-labeled analog was 0.033 ± 0.026 p.p.m.

Compound affinity is determined for each fragment-hit molecule. Figure 3.7 shows three binding isotherms. The first binding isotherm was used to identify molecule 9B11, which is a ligand that binds to the protein KIX in the presumed MLL-binding site near Y631 (K_d = 1.6 mM, LE = 0.27). The six experiments can be completed in 30 min. In the following two binding isotherms, the protein Brd4 was separately labeled with 5FW and 3FY. Titration with the Brd4 ligand acetaminophen yielded a similar K_d value for the two proteins (LE = 0.44). Titration of two alternatively labeled proteins is a useful control for assessing any perturbing effects of fluorine on ligand–protein binding. In addition, the magnitude of chemical shift change was substantially larger for the protein

labeled with 3FY than for that labeled with 5FW (0.61 p.p.m. versus 0.14 p.p.m.). These data are consistent with acetaminophen binding farther away from W81 than the affected tyrosine side chains. In some cases, more than one resonance in the NMR spectrum is affected by ligand binding. In these instances, binding isotherms can be obtained based on perturbation of each resonance, thus providing multiple dissociation constant determinations from the same experiment. An example of two affected resonance perturbations for the 5FW-labeled bromodomain, BrdT, in the presence of a new fragment is shown in Figure 3.5, yielding comparable dissociation constants of 250 and 210 μM. The larger magnitude of the change in chemical shift is consistent with ligand binding near the presumed WPF resonance W50.

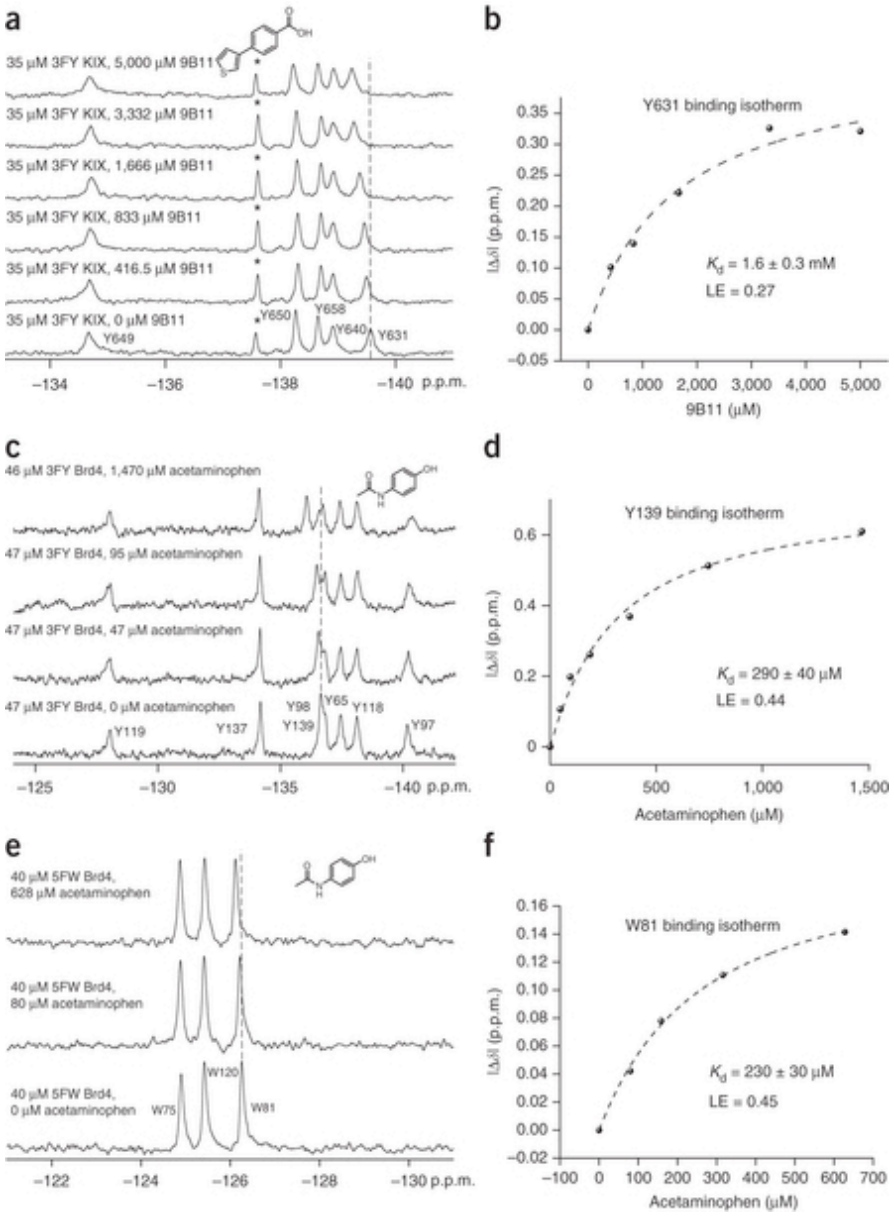


Figure 3.7 Binding isotherms from PrOF NMR titrations.
(a) Stacked spectra for the titration of 3FY KIX with molecule 9B11 (see structure at the top of the panel). A dashed line is added for reference. The asterisk denotes a partial degradation resonance. (b) Binding isotherm generated by monitoring the change in chemical shift of fluorine-labeled Y631 as a function of small-molecule concentration. (c) Stacked spectra for the titration of 3FY Brd4 with acetaminophen (see structure at the top of the panel). A dashed line is added for reference. (d) Binding isotherm generated by monitoring the change in chemical shift of fluorine-labeled Y139 as a function of small-molecule concentration. (e) Stacked spectra for the titration of 5FW Brd4 with acetaminophen (see structure at the top of the panel). A dashed line is added for reference. (f) Binding isotherm generated by monitoring the change in chemical shift of fluorine-labeled W81 as a function of small-molecule concentration.

Hajduk *et al*.[148] have demonstrated the efficacy of NMR-based screening methods for assessing the druggability of a protein target. Proteins with screening hit rates that are lower than 0.1% using fragment libraries are characterized as having low druggability. Upon completion of the fragment screen, PrOF NMR can also provide druggability information based on the screening hit rate, both for different proteins and for different binding sites within a single protein. As an example, of the 508 compounds screened against KIX, the MLL site was preferential for ligand binding (0.8% hit rate) over the CREB site (0% hit rate),[127] consistent with previously reported work[149].

To further increase the speed of PrOF NMR experiments, paramagnetic relaxation agents can be used to enhance the relaxation of atomic nuclei. In NMR experiments, the recycle delay, during which the spectrometer is idle while the magnetization relaxes after a radio frequency pulse, comprises up to 80% or more of the total experiment time (Figure 3.8).[150] Shorter recycle delays allow more scans per unit time, leading to more efficient instrument use. One strategy employed to shorten recycle delays is the addition of a paramagnetic metal to decrease the longitudinal relaxation time (T_1) of ^{1}H, ^{13}C, and ^{15}N nuclei in labelled proteins while minimizing unfavorable line-broadening from decreasing transverse relaxation time (T_2).[150–155]

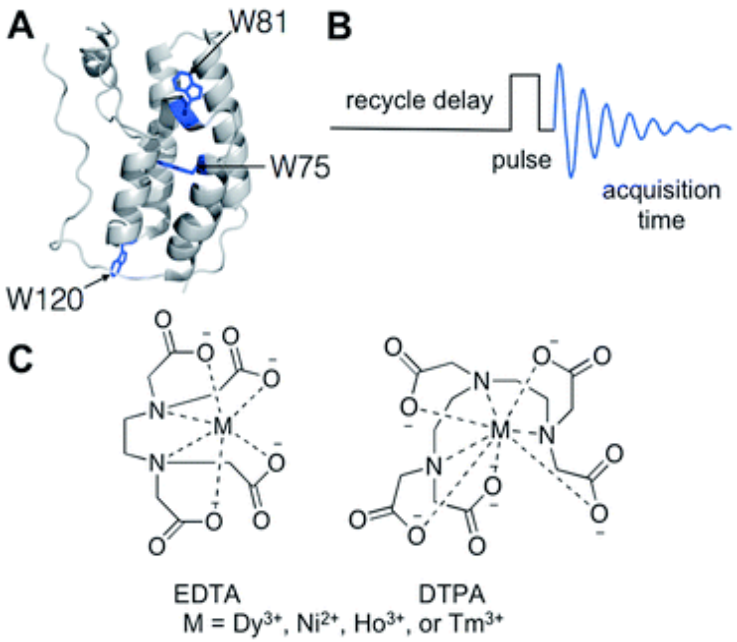


Figure 3.8 Mechanism of paramagnetic relaxation enhancement molecules.
(A) Brd4 with sites of fluorine-labelled tryptophans highlighted in blue. (B) Standard 1D 19F NMR pulse sequence with 90° pulse. (C) Structures of EDTA and DTPA chelates used in these experiments.

We explored the application of paramagnetic relaxation enhancement (PRE) to PrOF NMR, utilizing a solvent-accessible side-chain labelling scheme. This labelling scheme is well-suited to studying protein–protein interaction interfaces, which contain a high proportion of aromatic amino acids[100] that can readily be biosynthetically replaced with their fluorinated counterparts. The appended fluorine atoms on the side chains can be more solvent exposed than backbone amides, and thus more accessible to paramagnetic additives in the bulk solvent. We demonstrate that PRE can reduce experiment times significantly by shortening the recycle delay with minimal line broadening from T_2 effects, ultimately leading to a marked enhancement in sensitivity. In the context of low-complexity molecules commonly used in fragment-based ligand discovery screening, we further show improvements by using this method for rapidly obtaining dissociation constants for weak-binding fragments. In addition to fragment screening, this method may also be useful for proteins available in limited abundance such as GPCRs,[89,156] as the signal to noise increases by the square root of the number of scans. Importantly, this approach is also compatible with proteins containing metal binding hexahistidine affinity tags and with various fluorinated amino acids.

In the absence of Ni-DTPA, experiments were carried out with a recycle delay of 1 second (1.2 times the measured T_1 value of W81, the residue most affected by ligand binding in the binding pocket), with a total experiment time of five minutes (280 scans) at modest protein concentrations of 50 μM. With 20 mM Ni-DTPA present, the recycle delay was shortened to 0.24 seconds for a consistent value of 1.2 times the measured T_1 of W81. In this case, the experiment time was reduced to two minutes (350 scans) at the same protein concentration with similar signal-to-noise for the W81 resonance (11.8 with Ni-DTPA vs. 10.9 without), a 60% reduction in experiment time (Figure 3.9). The signal-to-noise ratio may be further improved if Q-damping effects could be mitigated. The T_1 value of W120 (0.15 s) is more reduced than that of W81 (0.20 s) by the addition of 20 mM Ni-DTPA. Consequently, the relative intensity of W120 increases with the addition of 20 mM Ni-DTPA because the net magnetization has more fully relaxed back to the ground state before application of the next pulse. In contrast, W75 has a longer T_1 value (0.27 s) than W81 and thus its resonance is truncated by the short recycle delay of this experiment. Because W75 is farther from the binding pocket, the experiment was optimized for maximizing signal-to-noise per unit time of W81. Because binding sites must be accessible to their binding partners, amino acid side chains located at protein–protein interaction interfaces should also be accessible to paramagnetic additives and thus subject to substantial PRE effects, allowing for shortened recycle delays and experiment times. Protein stability does not undergo large perturbations upon titration of Ni-DTPA, as measured by differential scanning fluorimetry, a thermal shift assay. Changes in melting temperature were ≤1.2 °C by this method (Figure 3.10).

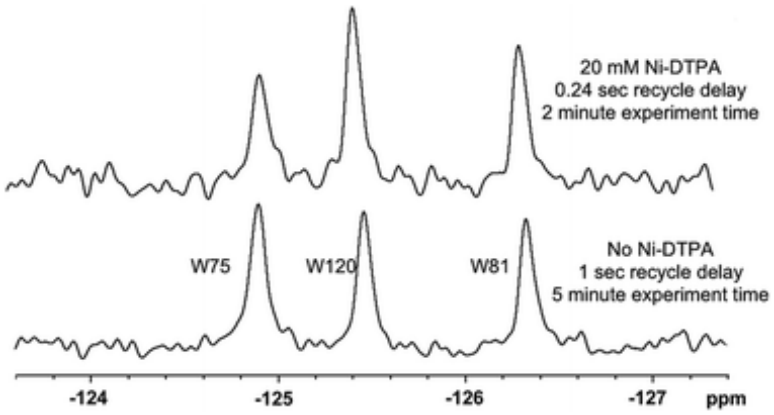


Figure 3.9 Comparison of 19F NMR spectra of 50 µM 5FW-Brd4 in the presence and absence of paramagnetic additive.
The addition of 20 mM Ni-DTPA allows a 60% reduction in experiment time while preserving the signal-to-noise ratio of W81, perturbations of which report on ligand binding. The signal-to-noise ratio of W81 without Ni-DTPA was 11.8, and with Ni-DTPA was 10.9.

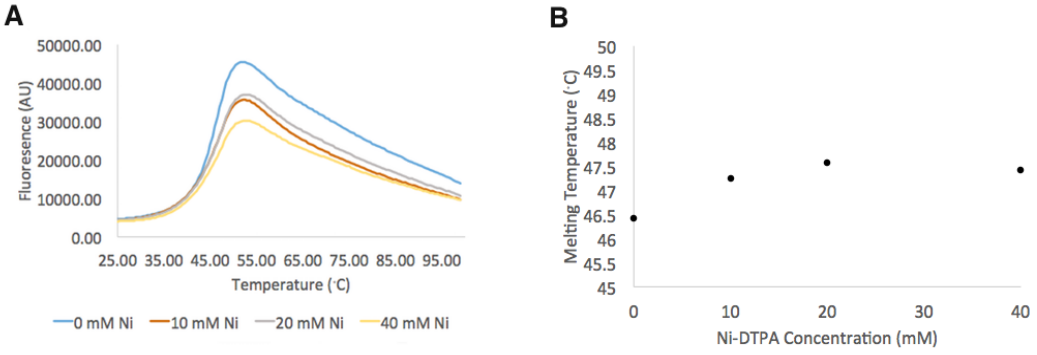


Figure 3.10 Differential scanning fluorimetry performed on Brd4 at increasing Ni-DTPA concentrations.
Only small changes are seen in the melting temperature. (A) Melting curves at three different Ni-DTPA concentrations showing changes in fluorescence upon increasing temperature. (B) Melting temperature of Brd4 plotted against Ni-DTPA concentration. The melting temperature was 46.4 °C in the absence of Ni-DTPA, 47.3 °C in the presence of 10 mM Ni-DTPA, 47.6 °C with 20 mM Ni-DTPA, and 47.4 °C with 40 mM Ni-DTPA.

In conclusion, PrOF NMR offers the medicinal chemist a useful screening platform for small-molecule fragments in early-stage ligand discovery campaigns as either a primary or secondary screen informing future ligand optimization efforts. Since the seminal studies by Sykes *et al.*[87] on alkaline phosphatase, ^{19}F probe developments and

improvements in fluorinated amino acid incorporation have led to ligand identification and affinity quantification with small- to medium-sized proteins in a time-efficient manner.[119] However, in all cases, fluorine perturbation to protein function should be assessed. Although current limitations to PrOF NMR exist—including mixture deconvolution steps and CSA relaxation effects on fluorine nuclei located on protein side chains with limited mobility or in large proteins—these challenges are beginning to be addressed by complementary ligand-observed NMR methods, as well as new fluorinated amino acid labeling strategies[106].

Table 3.4 Troubleshooting

Step	Problem	Possible Reason	Solution
9	Low protein yield **relative to protein expression yields using natural amino Acids**	Cell line may not be robust for protein expression	Change cell lines. Auxotrophic cell lines tend to provide lower yields than non-auxotrophic cell lines
			Use glyphosate to induce auxotrophy in a standard cell line [e.g. BL21(DE3)][53,54]
		Poor compatibility with tRNA synthetase	Cotransform with a plasmid to increase tRNA synthetase levels[47]
		Toxic levels of fluorinated amino acids	Utilize cell-free methods[55]
			Reduce the concentration of fluorinated amino acids
Box 1	Poor label incorporation (< 90%)	Non-optimized recovery time	Experiment with longer recovery times following the media swap
		Not enough fluorinated amino acid or amino acid precursor	Increase the amount of the fluorinated amino

			acid
13	Extremely broad NMR resonances	Magnetic field strength is too high or side chain dynamics are too slow, which can cause significant broadening of resonances due to chemical shift anisotropy Protein is too large	We find ideal field strengths to be 500 MHz – 600 MHz. The signal sensitivity and resolution benefits of higher field strength magnets can be cancelled out by CSA and shortened T_2 relaxation. Use alternate labeling strategies that utilize fluorinated amino acids with more rapidly rotating fluorine groups (e.g., trifluoromethyl)[43,46]
	Unknown fluorine resonances in the NMR spectrum	A fluorinated impurity is present	Fluorinated impurities in your sample may be removed by a buffer exchange Alter spectral window. Polymer background is most pronounced beyond -140

14	Fluorine referencing problems	Fluorine reference is sensitive to solution conditions or the presence of small molecule ligands	Alternative fluorine references (e.g. trifluoroethanol) can be used Alternative NMR pulses (e.g. ERETIC) can be utilized.[56,57]
Box 2, Step 2	Site directed mutagenesis failure	The specific amino acid may be necessary for the protein to adopt its proper secondary structure or affects more than one resonance	Site-selective labeling of 4FF and 3FY can be used to label individual residue positions rather than globally changing all of a given amino acid[58,59] Kitevski-LeBlanc et al. describe a multidimensional NMR method to assign resonances[60] Paramagnetic metals (e.g., Gd^{3+}) can be used to broaden surface exposed residues
19	Protein signal overwhelmed by additional fluorine resonances	A fluorine-containing molecule is present in the mixture	Deconvolute the mixture to ascertain if other ligands are binding Test fluorinated molecules separately and at lower concentrations Use a different fluorinated amino acid labeled protein
23	No deconvoluted compounds exhibit the same change in chemical shift as the mixture	Small molecules may have additive effects on the protein signals	Select the compound that contributes the most towards the change Prioritize other mixtures with more distinct effects from the

			compounds
29	Data do not fit well for K_d determination (Poor R^2 value)	Saturation point not yet reached, making it difficult to accurately determine a K_d using the equation Nonspecific or multi-site binding	Increase the concentration of the ligand to approach saturation Deprioritize this compound Change non-linear regression analysis to include multiple binding sites

Chapter 4. Dual screening of BPTF and Brd4 uncovers new probe molecules

Motivation: Bromodomain-containing protein dysregulation is linked to cancer, diabetes, and inflammation. Selective inhibition of bromodomain function is a proposed therapeutic strategy. The bromodomain of BPTF was implicated in various cancers, so we developed a probe to enable these studies. The chemical shift dispersion of 19F NMR is sufficient to allow simultaneous analysis of two bromodomains labeled with fluorinated tryptophan. This allows the screen and counter-screen simultaneously with a single NMR experiment. Isoform selectivity is known to be challenging between bromodomains, so in order to elucidate the role of the bromodomain of BPTF in disease we knew that selectivity would be challenging but required. Using Brd4 as the counter screen, this was the first communication of a small molecule inhibitor selective for BPTF over Brd4. This aryl-urea containing inhibitor was termed AU1, and is a scaffold for further chemical probe development.

4.1 **Introduction**

Lysine acetylation is an important post-translational modification that is significant in the epigenetic regulation of both health and disease. Histone proteins that are acetylated are bound by bromodomain-containing proteins facilitating assembly of transcription complexes. Small molecules that enable characterizing the role of these epigenetic proteins will improve our understanding of signaling pathways and may ultimately lead to new therapeutics.[157] Clinical trials are underway evaluating inhibition of several members of the bromodomain and extra terminal family (BET) bromodomains

(i.e, Brd2, 3, 4 and T) in cancer, diabetes, and inflammation, supporting bromodomain modulation as a potential therapeutic approach.[158] The bromodomain PHD finger transcription factor, BPTF, is thought to play a significant role in melanoma,[39] leukemia,[38] colorectal,[159] bladder cancer,[109] and is required for c-MYC transcriptional activity[42]. Both the PHD finger and bromodomain are important for chromatin binding.[44] No selective inhibitors for the BPTF bromodomain had been reported to test its role in regulating transcription or cancer, which motivated this research.

The lack of suitable ligands for competition-based experiments provides a challenge for developing reliable screens for bromodomain ligand development, specifically for BPTF.[52] Direct binding experiments using NMR have become a valued method for screening, due to the ability to quantify small molecule protein interactions over a wide range of affinities, particularly weak ligands, and have been used for bromodomain ligand discovery.[68,104,160] Protein-based methods using labelled amides (e.g., ^{1}H-^{15}N HSQC) provide additional structural information for developing small molecules; however, the experiment can be material intensive and time consuming.[70] The fluorine nucleus is highly sensitive to changes in chemical environment. Using this environmental sensitivity, we and others reported on a protein-observed fluorine NMR method (PrOF NMR) for characterizing ligand binding at protein-protein interaction sites using fluorine-labelled side chains which showed a time enhancement of at least 2-fold over HSQC on 12 and 15 kDa proteins.[99,127,130,161,162] ^{19}F is 83% as sensitive as ^{1}H and 100% isotopically abundant (thus inexpensive), facilitating detection of ^{19}F at low concentrations (μM) for small and medium-sized proteins. We previously applied PrOF NMR for fragment-based screening of over 500 small molecules, and previously characterized the bromodomains Brd4, BrdT, and BPTF.[99,127,131]

In this report, we demonstrate how the bromodomain for BPTF and the first bromodomain of Brd4 can be screened simultaneously due to the significant chemical shift dispersion and simplified ^{19}F NMR spectra. This approach is similar to RAMPED-UP NMR developed by Zartler et al. who demonstrated the screening potential of three differentially labelled proteins via 2D-HSQC NMR experiments.[133] These multiplexed protein experiments offer selectivity information up front and can lead to the discovery of

new ligands for two to three biological targets in one screen. Selective targeting of bromodomains remains a significant challenge due to binding site similarity.[131] We aimed to test if screening in the presence of other BET bromodomains could increase assay throughput, stringency, and binding information for finding selective inhibitors. Several structural classes of kinase inhibitors show preferential binding to BET bromodomains, including the PLK-1 kinase inhibitor BI2536.[29,30] We reported its additional binding to a non-BET bromodomain BPTF using PrOF NMR.[131] We have now screened 229 related compounds and disclose our findings for both Brd4 and BPTF selective compounds, including the first reported for BPTF. We validated our ligands using non-fluorinated proteins in protein stability, fluorescence anisotropy, and isothermal titration calorimetry (ITC) experiments (**Figure 1**). These led to submicromolar inhibitors for Brd4 and a selective BPTF ligand (K_d = 2.8 μM) that we term AU1, which is active in cell culture. Finally, we compare our screening hits to a prior screen using x-ray crystallography.[29] We anticipate new Brd4 and BPTF ligands will be valuable probes for studying the role of these bromodomains in various cancers.

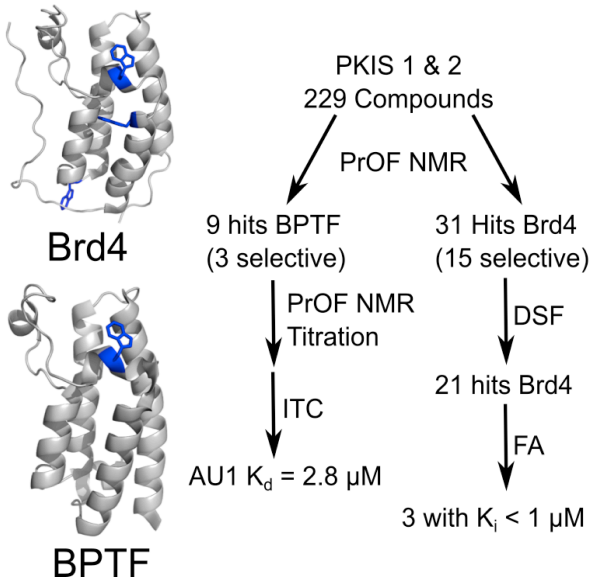


Figure 4.1 The dual screening workflow.
A selection of compounds from the Public Kinase Inhibitor Set from GSK were screened against 5FW-BPTF and 5FW-Brd4. Hits for BPTF were then rank-ordered by titration of ligand, and then one final compound was analyzed by isothermal titration calorimetry (ITC). Hits for Brd4 were rank ordered by differential scanning fluorimetry (DSF), and then K_i information determined by fluorescence anisotropy (FA).

4.2 **Results and Discussion**

4.2.1 Design of a dual Protein-Observed ^{19}F NMR screening assay with bromodomains Brd4 and BPTF

At least four aromatic amino acids for fluorine labeling are found in the histone binding site for more than half of all bromodomains.[131] We chose to incorporate fluorinated tryptophan residues into both proteins (three for Brd4: W120, W81, and W75; one for BPTF: W2824). Tryptophan 81 of the first bromodomain of Brd4 is located in the "WPF shelf", and is important for recognition of small molecule inhibitors including (+)-JQ1[13] and BI2536[29]. Tryptophan 2824 is the homologous residue in BPTF (Figure 2A). We can isolate our fluorinated proteins in high yield (84-88 mg/L with high fluorine labeling, >95%). We found fluorination to be only modestly perturbing to bromodomain structure and function.[131] For Brd4, the binding affinity of (+)-JQ1 is 78 and 75 nM for the 5-fluorotryptophan (5FW)-labeled and unlabeled proteins respectively. Thermal stability is only affected by a two degree change of the thermal melting temperature (Tm) from 52 to 50 °C for 5FW-labeled Brd4. In the case of BPTF, we measure a negligible change in Tm for the single 5FW incorporation relative to the unlabeled protein (T_m = 53 $^{\circ}$C in both cases) using variable temperature circular dichroism (Figure 4.3**,** Figure 4.4).

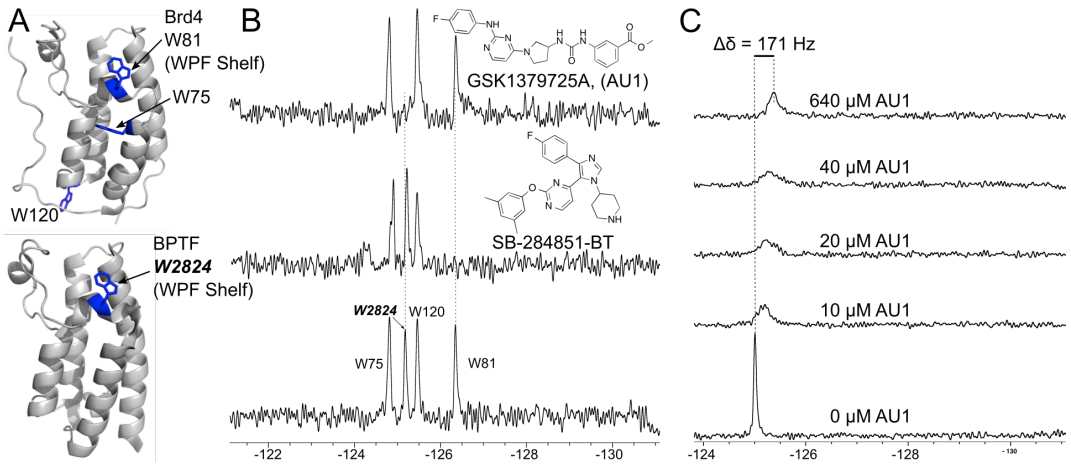


Figure 4.2 PrOF NMR dual screening examples.
A.) The tryptophan at the WPF shelf is in close proximity to the binding site in both Brd4 and BPTF. B.) A sample of NMR spectra from the screen with the second spectrum illustrating a selective Brd4 binder (only the W81 resonance is perturbed), and the third 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.

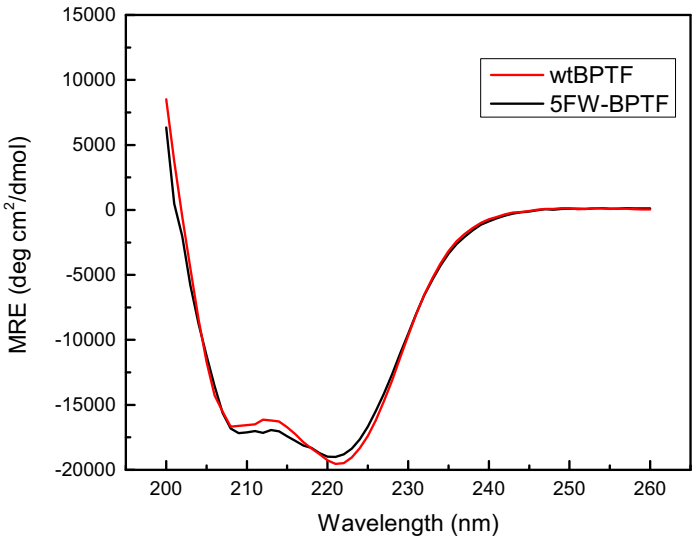


Figure 4.3 Circular dichroism of wtBPTF and 5FW-BPTF.
Secondary structure circular dichroism experiments indicate that fluorinating the tryptophan of BPTF only modestly perturbs the secondary structure.

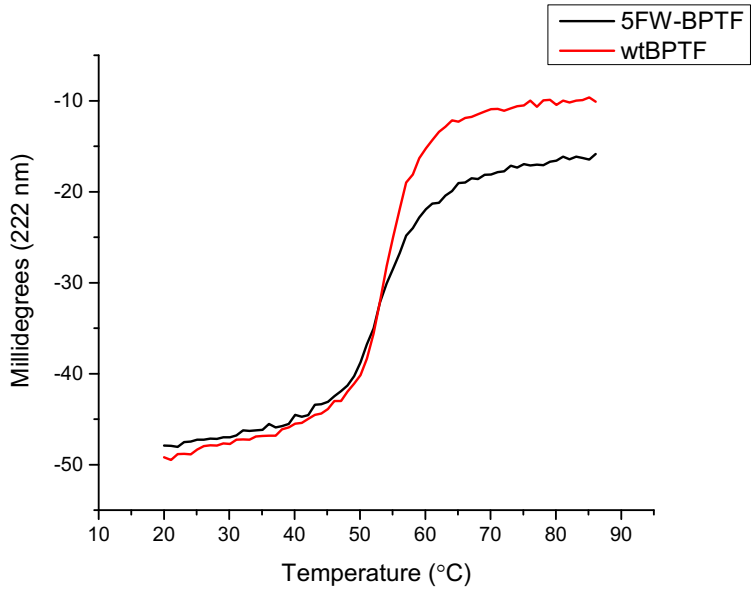


Figure 4.4 Circular dichroism thermal melts of wtBPTF and 5FW-BPTF. Thermal melting experiments at 222 nm result in negligible differences in melting temperatures (Tm = 53 °C in both cases).

Fluorine NMR spectra can be obtained for both proteins at 50 μM in 5 minutes (which includes a short reference experiment) leading to four resolved fluorinated tryptophan resonances corresponding to fluorine signals from both proteins (Figure 4.2). The commonly used solvent dimethylsulfoxide (DMSO) for screening libraries can also inhibit bromodomains.[163] For NMR, addition of this solvent does not dramatically alter the chemical shift dispersion nor the binding ability of the dual Brd4/BPTF ligand BI2536 at 1% (v/v). Therefore all screening experiments were carried out at 1% DMSO in the presence of 50 μM Brd4 and BPTF.

The simultaneous screen of Brd4 and BPTF offers an efficient method to discover small molecule binders for both bromodomains at the same time. Due to the similarity of the acetylated-lysine binding pocket, it is common for molecules to bind with multiple bromodomains; therefore, selectivity information is necessary when designing bromodomain inhibitors. The published kinase inhibitor set (PKIS I and II) contains small molecules designed for kinase inhibitor campaigns and is made widely available by GlaxoSmithKline.[164] A portion of this library has already been screened by x-ray crystallography against Brd4 and allows a basis for comparison of our results.[29] As an

initial test case, we studied two reported kinase inhibitors (TG-101348 and SB-202190) reported to be selective for Brd4 over BPTF.[29,30] Our NMR experiment reproduced this selectivity (Figure 4.5). We then screened 229 compounds individually at 100 μM. Chemical shift perturbations for all compounds were analyzed and the compounds were annotated as dual binders, selective BPTF or Brd4 binders, or non-binders (Table 4.4).

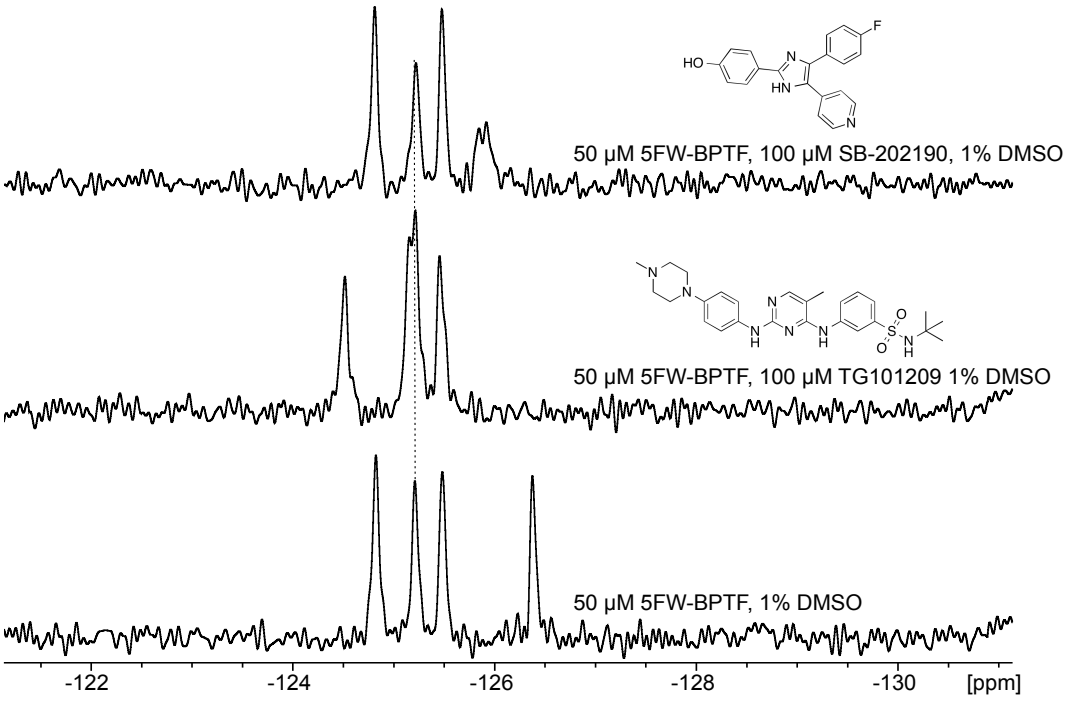


Figure 4.5 PrOF NMR experiments to assess screening conditions with known binders.

The PrOF NMR binding information from the simultaneous screen was readily interpreted for most binders that induced significant resonance broadening and chemical shift changes (Figure 4.6). In some cases, the resonance was completely broadened into baseline (Figure 4.2). NMR resonances exhibiting this broadening behavior from binding are dependent on both the residence time in the bound state and the chemical shift difference between the free and bound state.[134] In our experience, this is consistent with molecules with approximate K_d values from 0.1 to 100 μM. Small molecules can be further titrated into the protein solution until the resonance grows out of baseline, and

relative binding information can be estimated by the magnitude of change in chemical shift (Figure 4.2). We and others have also described PrOF NMR experiments yielding both slow and fast exchange binding kinetics, demonstrating the ability to detect molecules with dissociation constants in the low nanomolar up to millimolar range.[99,127,131,161,162] Finally, binding site position can be estimated by which resonance is perturbed. In most cases, the W81 resonance of Brd4 located in the WPF shelf is the most dramatically perturbed by a binding interaction in the histone recognition pocket, while W120 on the other side of the protein remains unperturbed. The stability of W120 offers confidence that the protein is still well-folded and in solution. On BPTF, W2824, the only tryptophan present, is also located on the WPF shelf.

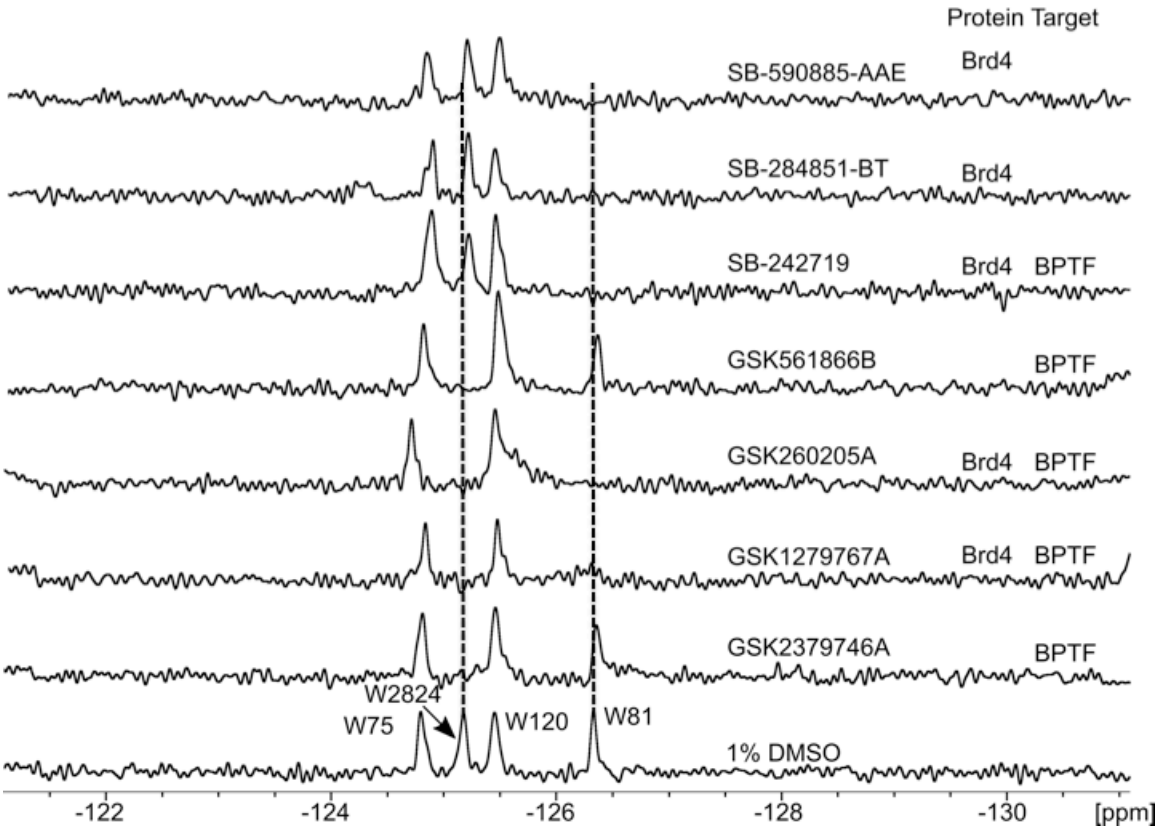


Figure 4.6 A selection of NMR spectra from the dual bromodomain screen.
There are spectra indicative of nonselective binders as well as ligands that selectively target Brd4 or BPTF. The protein target is indicated on the right of each spectrum.

4.2.2 PrOF NMR Screening Results for Brd4 Ligands

From our initial screen we further elucidated the 1,4,5- and 2,4,5-substituted-imidazole class that was previously disclosed as a ligand class for Brd4, and uncovered two new binding classes for both BPTF and Brd4, 1,2,5-oxadiazoles and arylureas containing 2,4-disubstituted pyrimidines. From the 1,2,5-oxadiazoles, 2/29 bound to BPTF (1 selective over Brd4). For the arylureas, 6/24 bound to BPTF (2 selective over Brd4, Figure 4.2**,** Figure 4.7**)**. A hit is described as the resonance broadening into baseline, and selective hits are defined as perturbing the resonance of the other protein by less than one standard deviation based on the average from all NMR experiments in the screen for both chemical shift perturbation and change in resonance width at half height. Both of these classes also contain compounds that resulted in moderate hits, where resonances were broadened but not significantly into baseline.

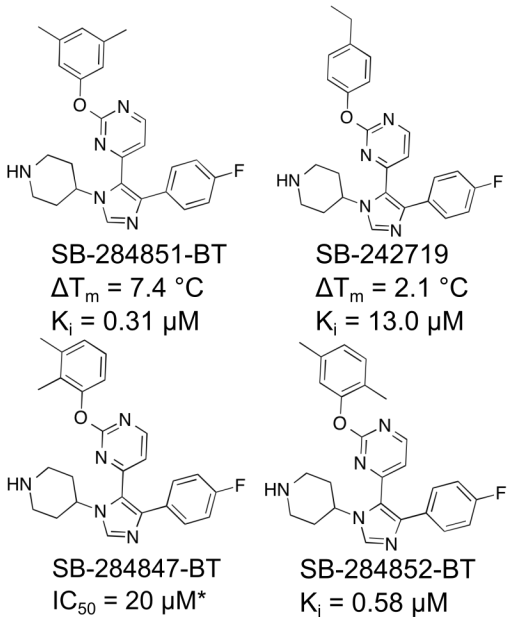


Figure 4.7 Trisubstituted imidazole ligands.
Structures of trisubstituted imidazole-containing compounds that are selective for Brd4 over BPTF. Thermal stabilization and inhibitory potency is reported from differential scanning fluorimetry and fluorescence anisotropy experiments respectively. *SB-284847-BT reported by Ember et al.[19] is shown for comparison.

The ligandability from the PKIS set for the first bromodomain of Brd4 was found to be slightly higher than BPTF (six selective strong binders versus three selective strong

binders), consistent with the druggability analysis of Vidler et al.[19] However, due to the similarity of many compounds this may also represent a limited amount of explored chemical space. Trisubstituted imidazole-containing compounds have been previously found to bind to Brd4, and we found additional compounds of this class that bind to this protein. Of the seven active trisubstituted imidazole-containing compounds (six of which were selective for Brd4), the most dramatic effects observed were for alkyl substitution on the phenylether at the 2-position on the pyrimidine. Ember et al. had previously reported a 2,3-dimethyl substitution pattern,[29] whereas in this study a 3,5-dimethyl, 2,5-dimethyl, and 4-ethyl were discovered (molecules **1**, **2**, and **4**, Figure 4.7). As NMR could only provide an estimate of the affinity, we turned our attention to complementary binding assays, using non-fluorinated bromodomains.

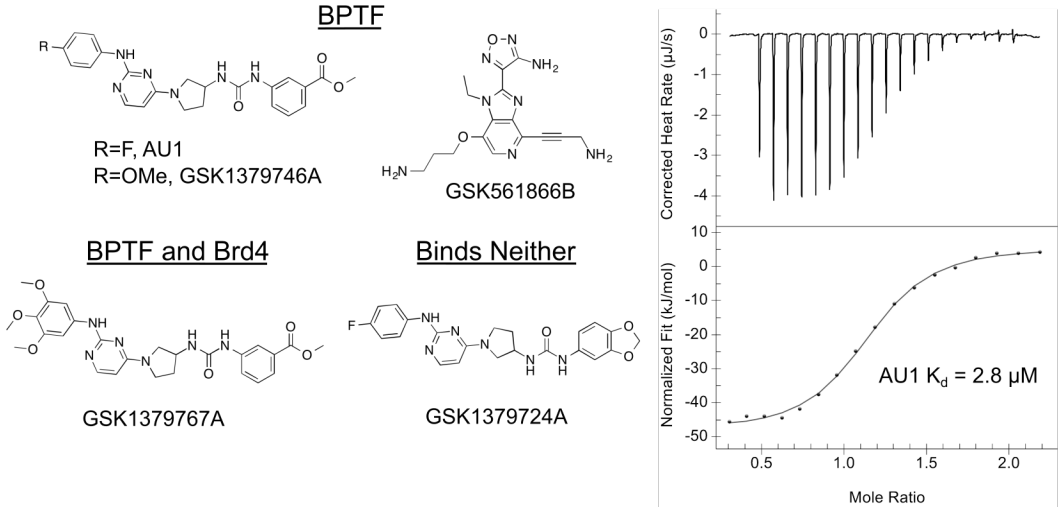
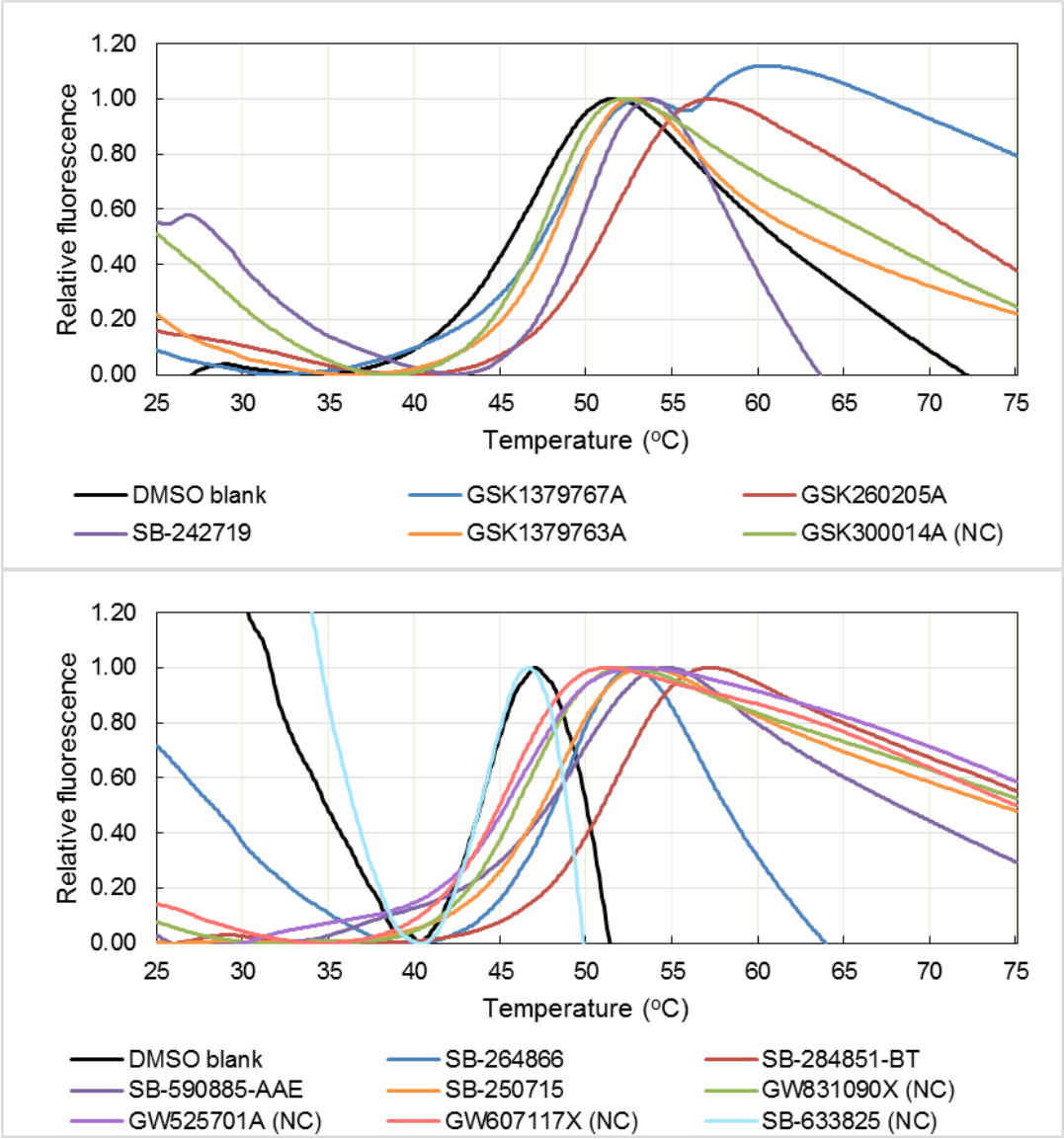


Figure 4.8 Selectivity of different actives from the screen.
Left.) Selected compounds of varying specificity for BPTF over Brd4, and non-binder GSK1379724A. Right.) ITC binding isotherm for AU1 (46 µM) titrated with bromodomain of BPTF (635 µM)

4.2.3 Thermal Stability Evaluation with Brd4 Ligands using Differential Scanning Fluorimetry

Differential scanning fluorimetry (DSF) has proven to be an effective screening tool for Brd4, demonstrating a linear correlation of ligand dissociation constants with thermal stability.[13,29] To rank order our newly discovered Brd4 ligands using a non-fluorinated protein, we generated a correlation curve based on published affinity values using six known Brd4 binders and thermal melting temperatures measured by DSF under

our experimental conditions (Figure 4.9, Figure 4.10). We used this curve to estimate the affinity of newly discovered small molecules from our screen, as has previously been demonstrated by Ember et al.[29] Of the 31 discovered Brd4 ligands, 21 showed a stabilization > 1 $^{\circ}$C. Eight non-binding compounds for Brd4 were also tested, five of which did not result in stabilization larger than 1 $^{\circ}$C; none of the non-binders stabilized greater than 2.5 $^{\circ}$C (Table 4.1). The high confirmation rate supported the sensitivity of PrOF NMR for ligand detection, reducing concerns over potential false positive effects from fluorine incorporation. The trisubstituted imidazole SB-284851-BT possessed one of the largest changes in thermal stability (ΔT_m = 7.4 $^{\circ}$C, Table 4.2, Figure 4.7, Figure 4.10). Due to the only modest linear correlation in our calibration curve, the affinity of this compound and analogs were subsequently quantified in competitive inhibition assays.


1.20
1.00
0.80
0.60
0.40
0.20
0.00
Relative fluorescence
25
30
35
40
45
50
55
60
65
70
75
Temperature (°C)
DMSO blank
GSK1379767A
GSK260205A
SB-242719
GSK1379763A
GSK300014A (NC)
1.20
1.00
0.80
0.60
0.40
0.20
0.00
Relative fluorescence
25
30
35
40
45
50
55
60
65
70
75
Temperature (°C)
DMSO blank
SB-264866
SB-284851-BT
SB-590885-AAE
SB-250715
GW831090X (NC)
GW525701A (NC)
GW607117X (NC)
SB-633825 (NC)

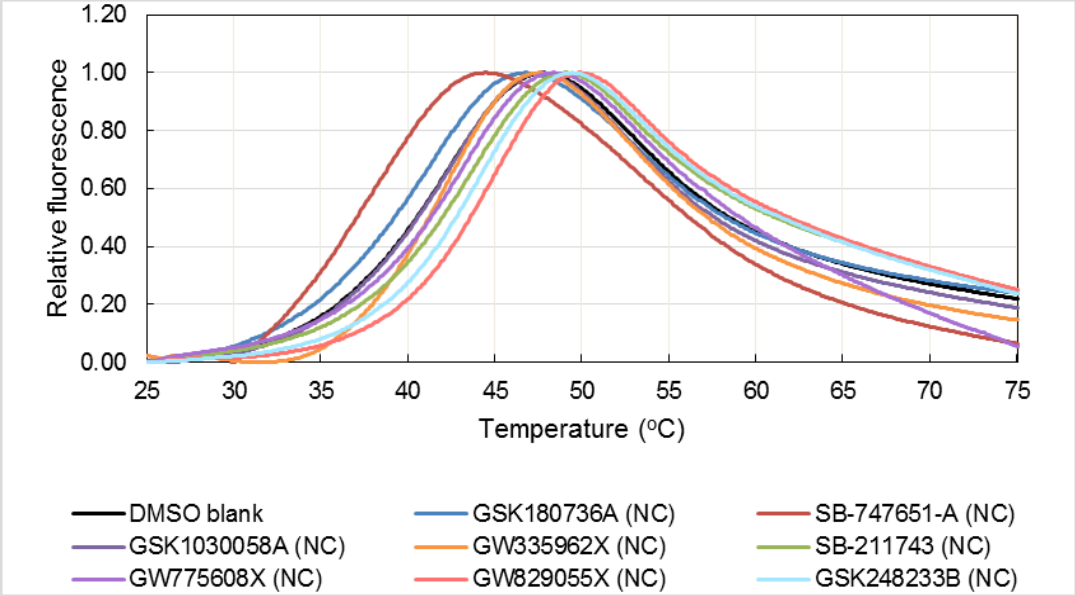


Figure 4.9 Thermal shift data from Brd4 with various ligands identified from the NMR screen.

Compounds that were not detected as hits in the NMR screen are labeled as negative controls (NC).

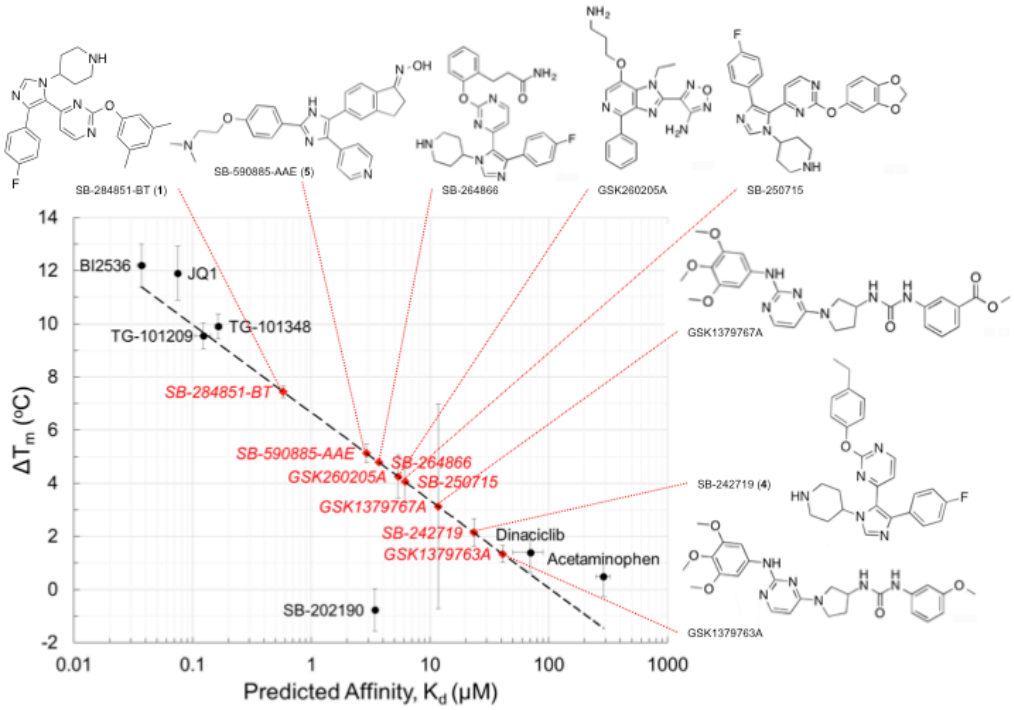


Figure 4.10 Callibration curve for DSF with Brd4.
A DSF calibration curve of previously published Brd4 binders is shown, correlating change in thermal melting with reported binding affinity. Compounds from this screen are illustrated with their respective lines, plotted according to change in melting temperature for estimation of dissociation constants from the linear regression shown in hashed lines.

4.2.4 Competitive Inhibition using Fluorescence Anisotropy for Brd4 Ligand Affinity Determination.

To quantify the ligand affinity for Brd4, we designed a competitive inhibition assay using fluorescence anisotropy (FA) and a fluorescently labeled derivative of BI2536, which was chemically synthesized and reported previously as a fluorescent PLK-1 kinase probe.[165] Chung et al. previously published a similar assay using fluorescently labeled analog I-BET762 as a Brd4 ligand.[33] A competition experiment eliminates the need to fluorescently label all newly discovered small molecules which could alter binding and is higher throughput than ITC. We measured a K_d of 50 ± 11 nM for a BODIPY-conjugated BI2536 probe in direct binding experiments (Figure 4.11), similar to reported values of 48 nM.[30] Importantly, the probe could be self-competed with unlabeled BI2536 (IC_{50} = 130 nM for Brd4). We also measured a K_d of 67 ± 21 μM

binding to BPTF, consistent with the moderate affinity detected in our NMR binding study and within 2-fold of our ITC measurements (K_d = 37 μM) (Figure 4.11). We used this probe to determine dissociation constants from a competitive inhibition experiment (i.e., K_i) for our new Brd4 ligands.

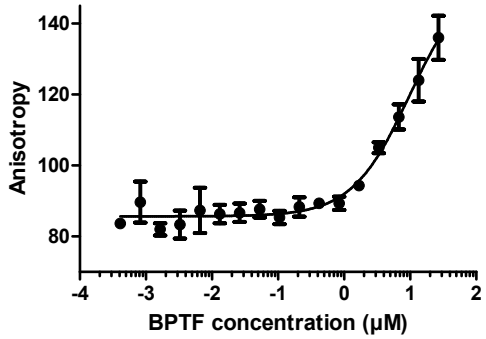


Figure 4.11 Fluoresence anistropoy with BTPF and BODIPY-BI2536.
Direct binding fluorescence anisotropy experiments between BODIPY-BI2536 and varying concentrations of BPTF were used to determine the Kd = 67 μM +/- 21.

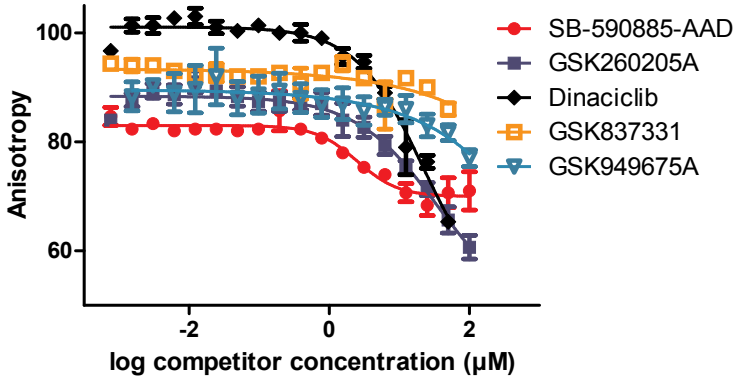


Figure 4.12 Determination of Ki of Brd4 binders by fluorescence anisotropy competition assay.
Fluorescence anisotropy signal is shown for several Brd4 binders, including dinaciclib (diamonds, Ki=41 μM), GSK260205A (squares, Ki=6.7 μM), and SB-590885-AAD (circles, Ki= 0.39 μM), GSK837331 (unfilled squares, 18% inhibition at 100 μM), and GSK949675A (unfilled triangles, 60% inhibition at 100 μM).

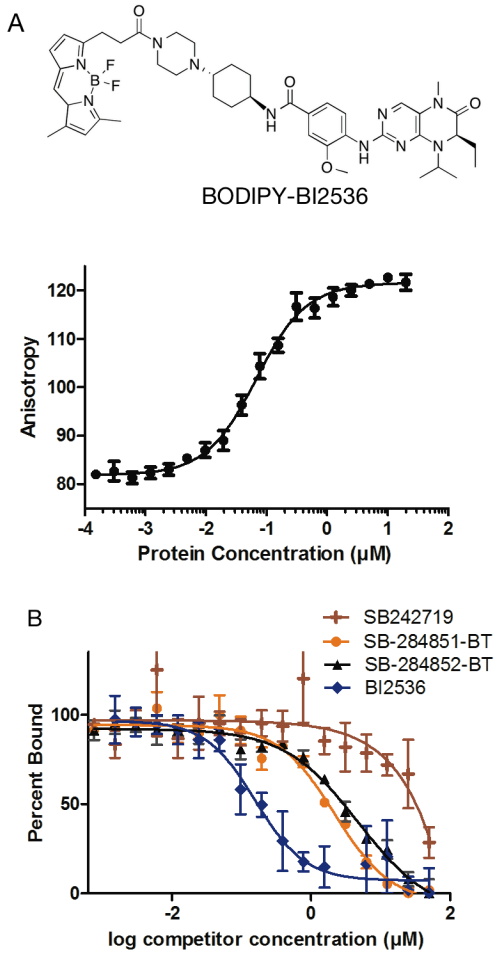
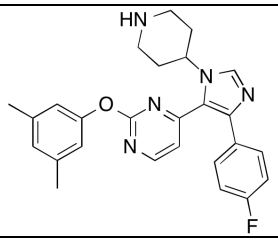
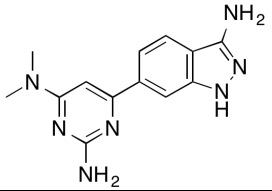
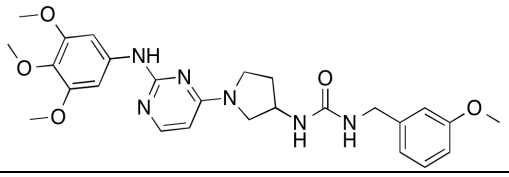
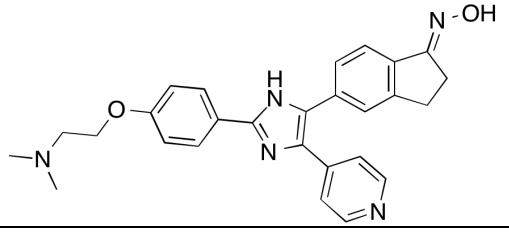
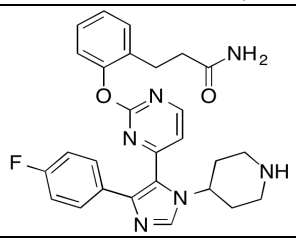
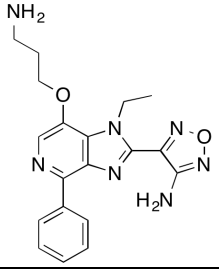


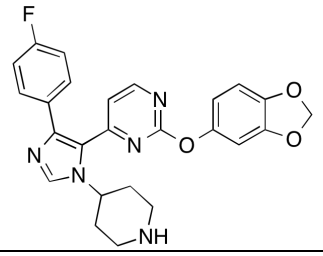
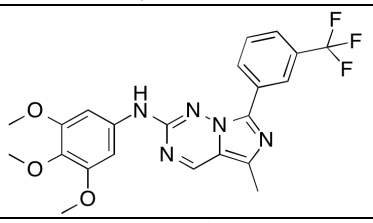
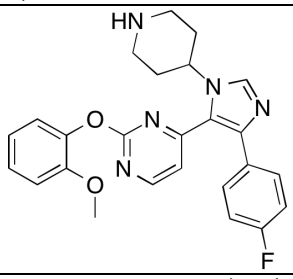
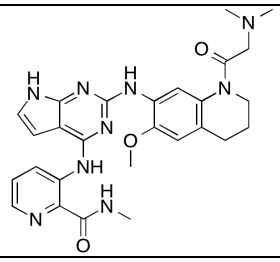
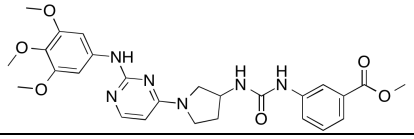
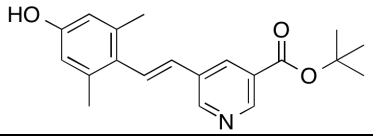
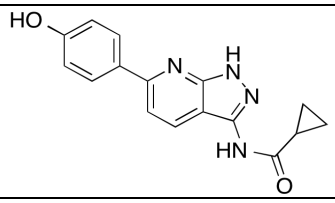
Figure 4.13 Fluorescence anisotropy direct binding and competition experiments with Brd4.
A) Direct binding between BODIPY-BI2536 and varying concentrations of Brd4 yields a K_i of 50 ± 11 nM B) Normalized fluorescence anisotropy signal is shown for Brd4 binders of various potencies, including SB242719 (crosses), SB242851-BT (circles), SB242852-BT (triangles), and unlabeled BI2536 (diamonds). Competition experiments were performed against bound BODIPY-BI2536.

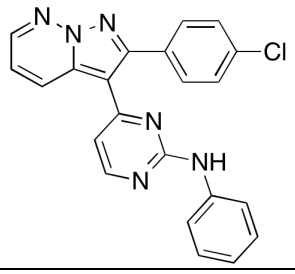
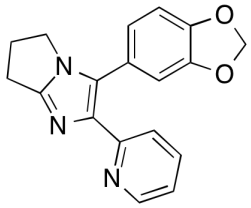
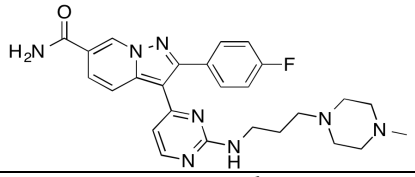
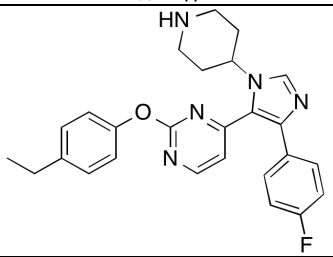
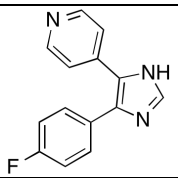
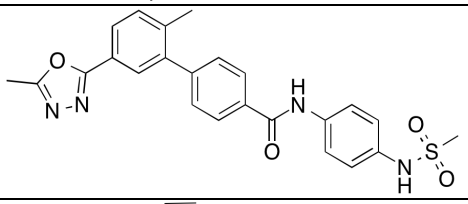
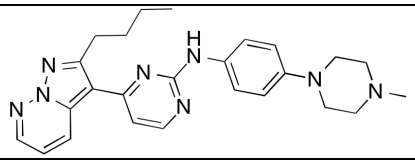
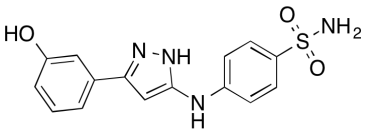
We previously determined a K_d of 70 ± 20 μM of a small molecule Dinaciclib binding to fluorinated Brd4,[131] a known moderate affinity inhibitor.[29] Using our FA assay, when Dinaciclib was dissolved in ethylene glycol, a bromodomain compatible solvent, we obtained a partial inhibition curve with an estimated IC_{50} of 220 μM (Figure 4.12**,** Table 4.1). From this study, we conclude that inhibitors with weaker affinity than

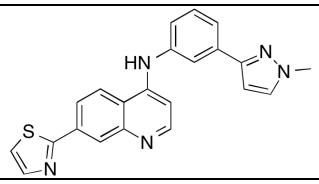
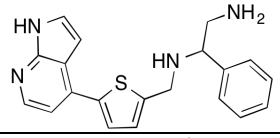
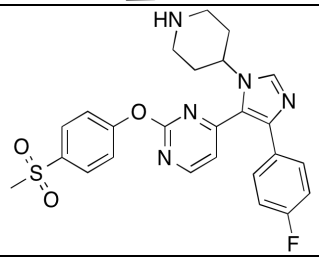
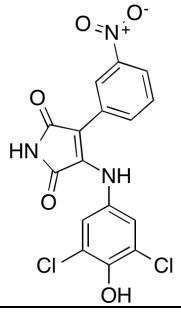
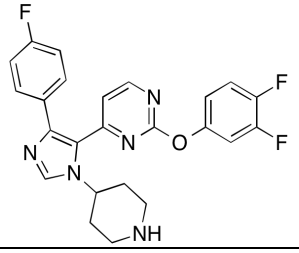
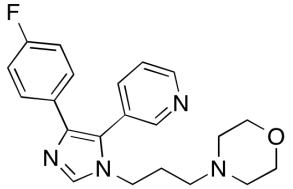
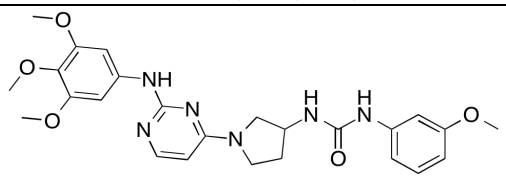
Dinaciclib will be difficult to characterize in this current FA format, and therefore we report only on a limited set.

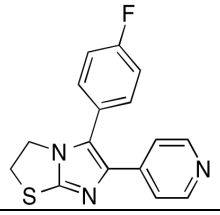
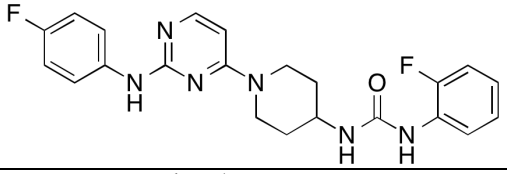
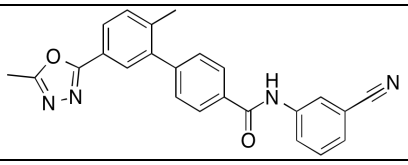
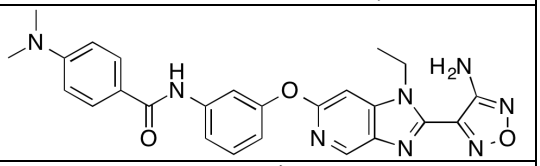
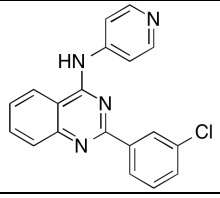
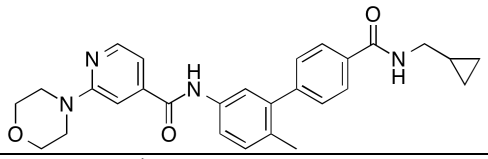
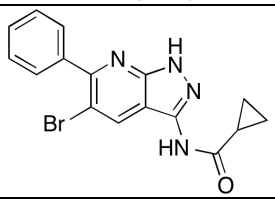
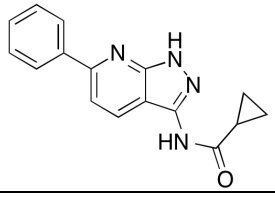
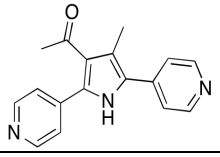
Table 4.1 Comparative assay data from the Dual Bromodomain Screen

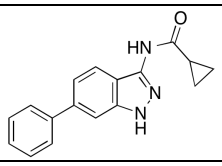
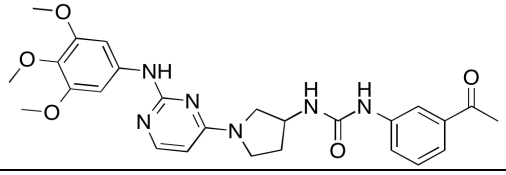
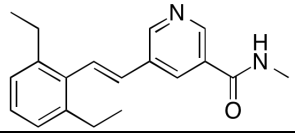
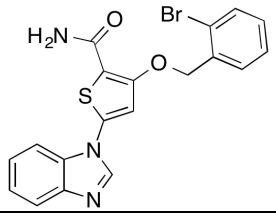
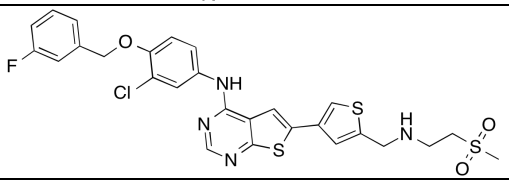
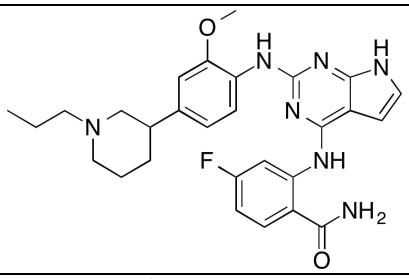
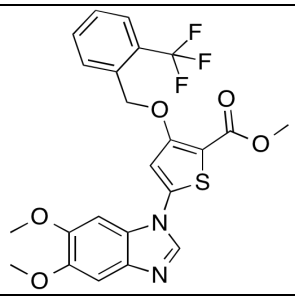
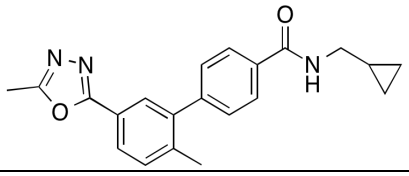
Compound Regno	[a]*PrOF* NMR	ΔT_m (°C)	[b]IC_{50} (μM)	[b]K_i (μM)	Structure
SB-284851-BT	+++	7.4 ± 0.2	2.2 ± 0.4	0.31 ± 0.04	
GSK2219329A	+++	6.3 ± 0.5			
GSK1379765A	+++	6.1 ± 0.3			
SB-590885-AAE	+++	5.1 ± 0.4	2.5±1.4	0.40 ± 0.23	
SB-264866	+++	4.8 ± 0.1	>100		
GSK260205A	+++	4.2 ± 0.8	39 ± 65	7 ± 12	

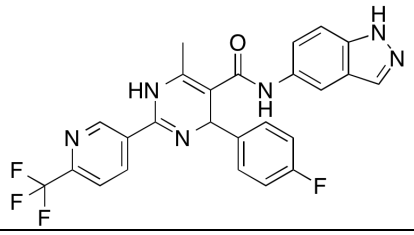
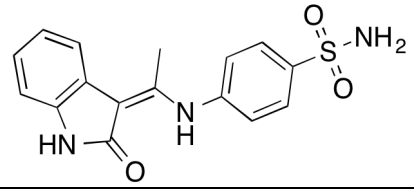
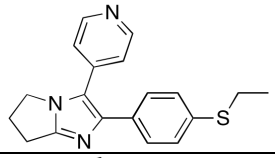
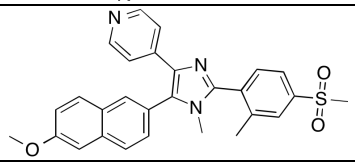
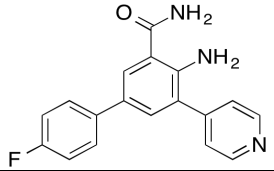
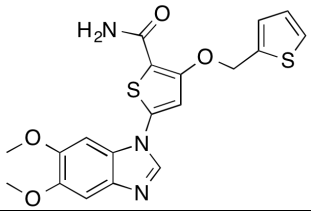
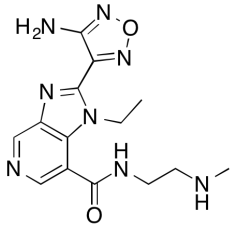
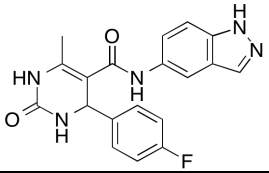
SB-250715	+++	4.1 ± 0.8	>100		
GW837331X	+++	3.8 ± 0.5	18% inhibited at 100		
SB-251527	+++	3.4 ± 0.6			
GSK2220400A	+++	3.1 ± 1.2			
GSK1379767A	+++	3.1 ± 3.9			
GW439255X	+++	3.0 ± 1.0			
SB-710363	++	2.7 ± 0.9			

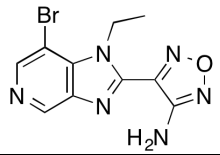
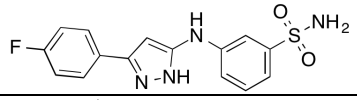
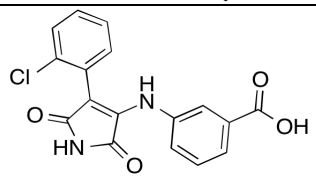
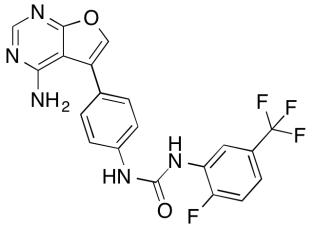
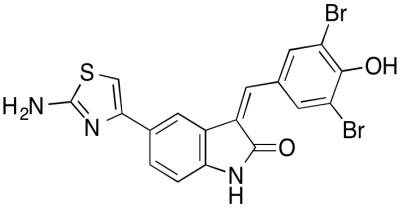
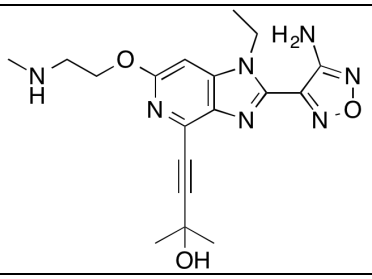
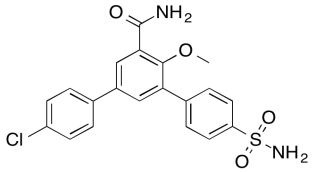
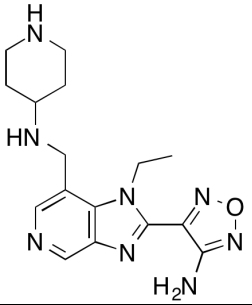
GW829055X	-	2.5 ± 0.7			
SB-400868-A	-	2.3 ± 0.8			
GW622475X	+++	2.2 ± 0.8			
SB-242719	+++	2.1 ± 0.5	70 ± 17	13 ± 3	
SB-217360	++	2.0 ± 0.8			
GW702865X	+++	2.0 ± 0.7			
GW684941X	+++	1.9 ± 0.5			
GW831090X	+	1.8 ± 0.5			

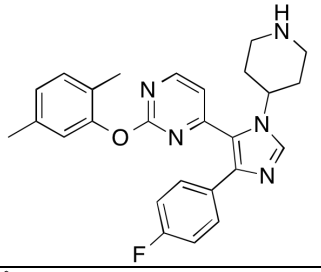
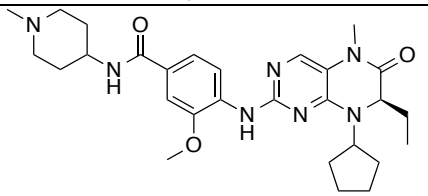
GW525701A	+	1.8 ± 0.6			
GSK1321565A	++	1.5 ± 0.6			
SB-254169	+++	1.5 ± 0.5			
SB-358518	+++	1.4 ± 0.1			
SB-253228	+++	1.3 ± 0.1			
SB-211743	+	1.3 ± 1.1			
GSK1379763A	+++	1.3 ± 0.3			

SKF-86055	+++	1.2 ± 0.3			
GSK1379800A	-	1.1 ± 0.5			
GW607117X	+	1.0 ± 0.4			
GSK248233B	-	0.9 ± 1.5			
GW873004X	++	0.8 ± 0.4			
GW775608X	-	0.7 ± 0.7			
SB-708998	++	0.6 ± 0.7			
SB-708999	++	0.6 ± 0.7			
GW493036X	++	0.6 ± 0.7			

SB-732941	++	0.6 ± 0.7			
GSK1379753A	+++	0.5 ± 0.5			
GW450241X	-	0.5 ± 0.5			
GSK1023156A	+++	0.4 ± 1.1			
GSK300014A	++	0.4 ± 0.5			
GSK1173862A	++	0.4 ± 0.5			
GSK1030058A	-	0.2 ± 1.0			
GW679662X	+++	0.2 ± 1.0			

GSK317354A	+++	0.0 ± 0.7			
GW335962X	+	-0.2 ± 0.8			
SKF-106164-A2	+++	-0.3 ± 1.2			
SB-633825	+	-0.4 ± 0.9	>100		
GSK620503A	+++	-0.7 ± 1.1			
GSK312948A	++	-0.8 ± 0.8			
GW876790X	++	-0.9 ± 0.9			
GSK180736A	+	-1.5 ± 0.8			

SB-738004	+++	-1.9 ± 1.1	>100		
GW824645A	++	-1.9 ± 0.7			
SB-390526	++	-1.9 ± 1.0	>100		
SB-707548-A	++	-2.1 ± 0.5	>100		
GW407323A	+++	-2.2 ± 1.0	>100		
GSK949675A	++	-2.4 ± 0.7	60% inhibition at 100		
GSK711701A	++	-2.7 ± 0.3	>100		
SB-747651-A	-	-4.6 ± 1.4			

SB-284852-BT			4.7 ± 1.4	0.58 ± 0.23	
BI2536			0.16 ± 0.05	Not determined	

[a] +++, a resonance from the protein in question has been completely broadened into baseline

++, a resonance from the protein in question has been perturbed by 2 standard deviations in either chemical shift or linewidth +, a resonance from the protein in question has been perturbed by 1 standard deviation in either chemical shift or linewidth

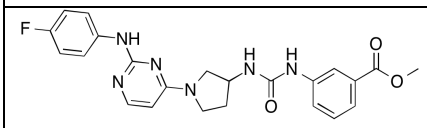
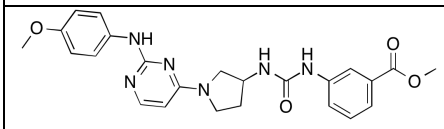
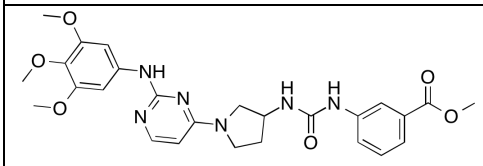
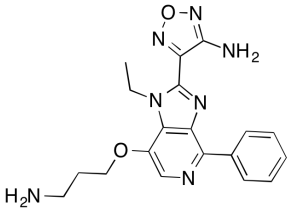
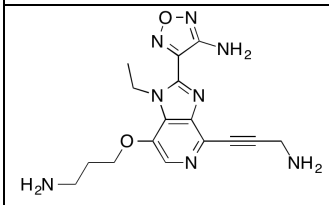
-, resonances are minimally to non perturbed

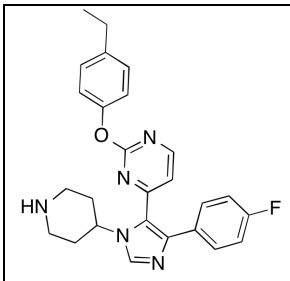
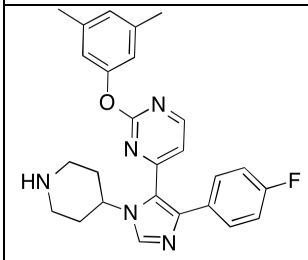
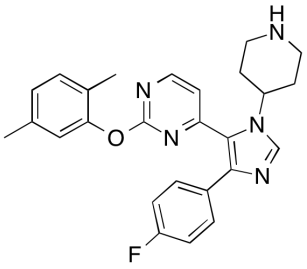
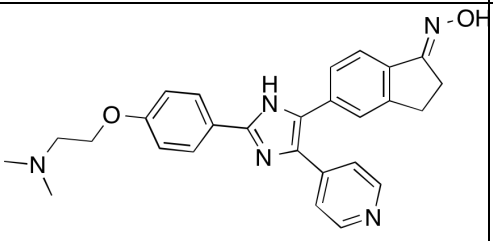
[b] values were obtained for via FA for Brd4 only

The labs of Schönbrunn[29] and Knapp[30] found that 1,4,5- and 2,4,5-trisubstituted imidazoles bound to Brd4. Trisubstituted imidazoles were also found to bind to BAZ2A/B.[166] We have now discovered analogs in our PrOF NMR experiment with submicromolar affinity, the most potent of these scaffolds for Brd4 to date. For 1,4,5-trisubstituted imidazoles in the FA competitive inhibition experiment, we determined a K_i of 310 nM for SB-284851-BT (Figure 4.13**,**Table 4.2), our most promising candidate from DSF and PrOF NMR experiments, a modestly higher affinity than estimated from the calibration curve. The arylalkyl substitution pattern bonded to the pyrimidine is important for binding. Based on the results with SB-284851-BT having a 3,5-dimethyl substitution pattern, we tested SB-284852-BT (previously not screened by PrOF NMR) with a 2,5-dimethyl group substituted on the aromatic ring. SB-284852-BT has a slightly higher K_i of 580 nM (Figure 4.13,Table 4.2). A regioisomer with a 2,3-dimethyl

substituted aromatic ring (SB-284847-BT) was co-crystallized by Ember et al. with Brd4 but inhibited with a much weaker potency (IC_{50} = 20 μM).[29] A 4-ethyl group on SB-242719 also results in 40-fold weaker affinity with a K_i of 13 μM in our assay and a ΔT_m = 2.1 °C by DSF (Figure 4.13**,**Table 4.2).

Table 4.2 Comparative assay data from the Dual Bromodomain Screen

Structure	BPTF ^{19}F NMR[a]	Brd4 ^{19}F NMR[a]	Brd4 ΔT_m (°C)	K_i, (FP, Brd4)	K_d (ITC, BPTF)
 GSK1379725A (AU1)	+++	-	N/A	N/A	K_d = 2.8 μM
 GSK1379746A	+++	-	N/A	N/A	N/A
 GSK1379767A*	+++	+++	3.1 ± 3.9	*b*	N/A
 GSK260205A	+++	+++	4.2 ± 0.8	K_i = 7 ± 12 μM	N/A
 GSK561866B	++	-	N/A	N/A	N/A

 SB-242719	+	+++	2.1 ± 0.5	K_i = 13 ± 3 μM	N/A
 SB-284851-BT	-	+++	7.4 ± 0.2	K_i = 0.31 ± 0.04 μM	N/A
 SB-284852-BT	N/A	N/A	N/A	K_i = 0.58 ± 0.23 μM	N/A
 SB-590885-AAE	-	+++	5.1 ± 0.4	K_i = 0.40 ± 0.23 μM	N/A
[a] +++, a resonance from the protein in question has been completely broadened into baseline ++, a resonance from the protein in question has been perturbed by 2 standard deviations in either chemical shift or linewidth +, a resonance from the protein in					

question has been perturbed by 1 standard deviation in either chemical shift or linewidth -, resonances are minimally to non-perturbed [b] Compound was not soluble in assay buffer	

Our newly discovered Brd4 inhibitors SB-284851-BT and a 2,4,5-trisubsitutedimidazole SB-590885-AAE (K_i = 400 nM, Table 4.2) were previously characterized as selective B-Raf kinase (K_d = 0.3 nM)[167] and p38α kinase (IC_{50} = 41 nM)[168] inhibitors. Although a variety of trisubstituted imidazoles have been reported to bind to Brd4, to date these two compounds are the most potent analogs, and may represent a starting point for ligand design or use in cell culture experiment to test the polypharmacology and potential synergy for p38α or B-Raf and Brd4 inhibition in various cancer models.

4.2.5 PrOF NMR Screening Results and Structure Activity Relationships for BPTF Ligands

In addition to new Brd4 ligands, several selective compounds were also found to bind to BPTF. These results are encouraging due to an absence of reported ligands for the bromodomain of BPTF. The arylurea containing compounds were the tightest binders that we found for BPTF exhibiting greater than two standard deviation changes in chemical shift and resonance broadening into baseline. Two arylurea containing compounds completely selective for BPTF over Brd4 were obtained, GSK1379746A and GSK1379725A. GSK1379725A was selected for follow-up studies, and so we now refer to it as AU1 (Figure 4.2**,** Figure 4.8**,** Figure 4.14). Negligible chemical shift changes were observed for Brd4 in the presence of AU1 at up to 160 μM (Figure 4.15). In addition, a 1,2,5-oxadiazole (GSK561866B) was highly selective for BPTF (Figure 4.8). Non-selective arylurea containing compounds from the screen that have a trimethoxy-substituted aryl ring and bound in the intermediate exchange regime by PrOF NMR, are anticipated to bind with the trimethoxyphenyl ring in the acetyl-lysine binding pocket based on the binding mode of the trimethoxyphenyl ring of GW61228X bound to Brd4

reported by Ember et al.[29]. Arylurea-containing compounds that lacked the trimethoxy-substituted aryl ring in some cases bound more weakly to Brd4.

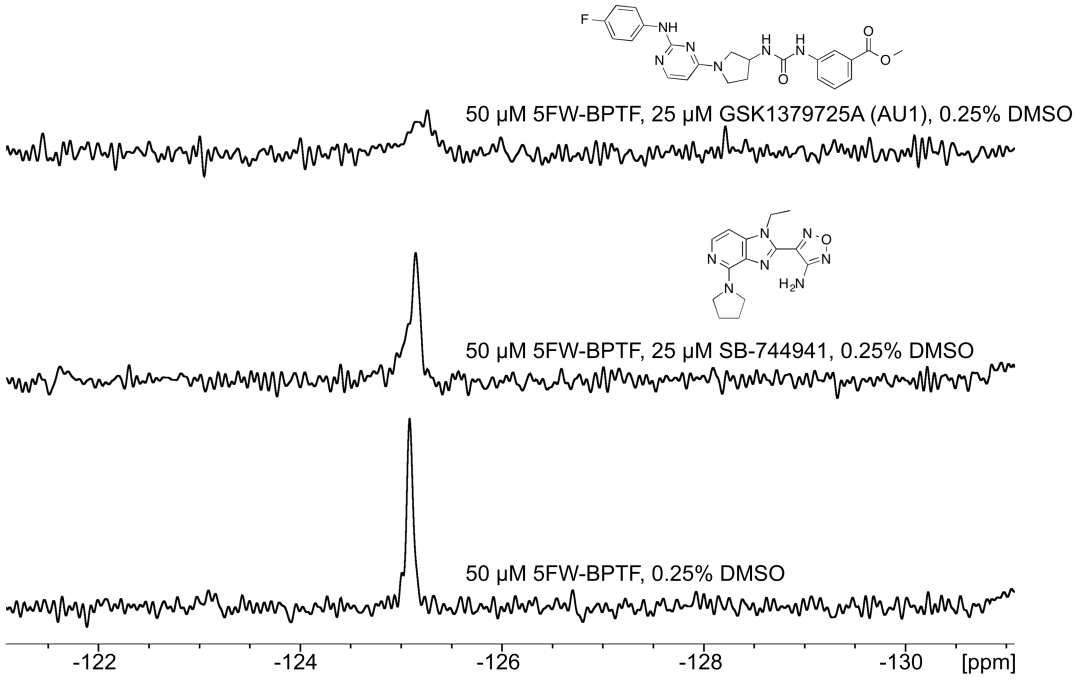


Figure 4.14 BPTF binders at lower concentrations.
The strongest selective binders were added into 5FW-BPTF at ¼ the concentration of the screen to rank-order the compounds. The rapid broadening of AU1 suggests a lower dissociation constant than the slight broadening of SB-744941.

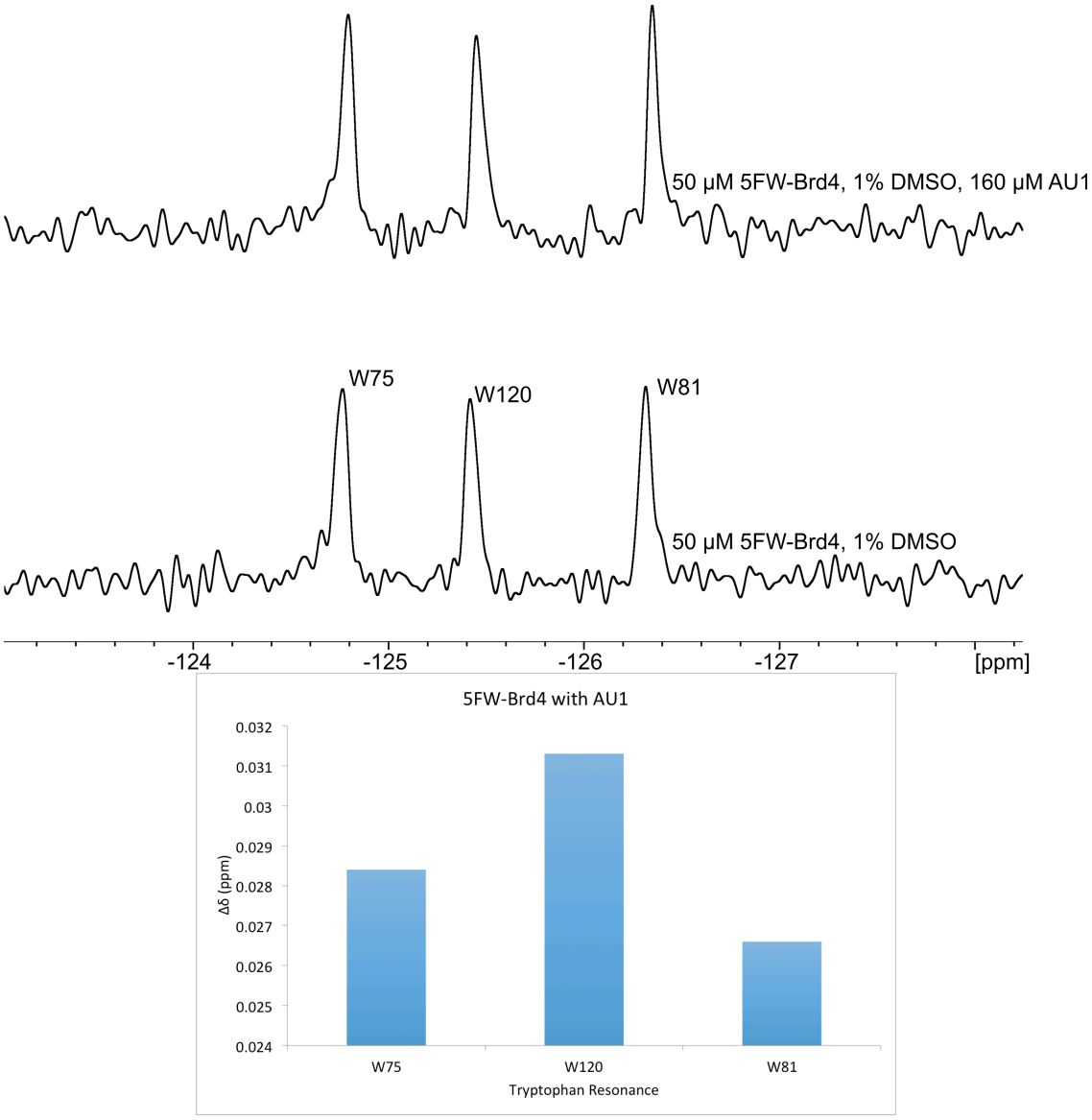


Figure 4.15 Fluorine NMR spectra analyzing 5FW-Brd4 with AU1.
A graph accompanying graph showing absolute chemical shift perturbation of each resonance. The low chemical shift perturbations as well as no significant broadening indicate minimal interaction of AU1 with 5FW-Brd4.

Of the six high affinity BPTF binding compounds of the arylurea class, binding appears to be correlated to a benzene ring with a meta-substituted carbonyl present. Computational binding studies using the Phase[169–171] module of Schrödinger indicate the placement of both the aromatic ring and hydrogen bond acceptor strongly correlate with binding (Table 4.3**,** Figure 4.16**,** Figure 4.17). Docking studies indicate probable π–π

interactions with the nearby phenylalanine 2887 and hydrogen bonds to the asparagine 2881 side chain in the binding site.[172,173] Compound GSK1379724A is similar to AU1 (Figure 4.8), but contains a cyclic acetal in place of a carbonyl group on the phenyl ring. This compound is unable to bind to either BPTF or Brd4, and represents a useful negative control compound for future studies. Related molecules without a carbonyl group showed poor BPTF binding. Additional arylurea-containing compounds will be developed to determine a more extensive structure activity relationship prior to obtaining higher resolution structural information.

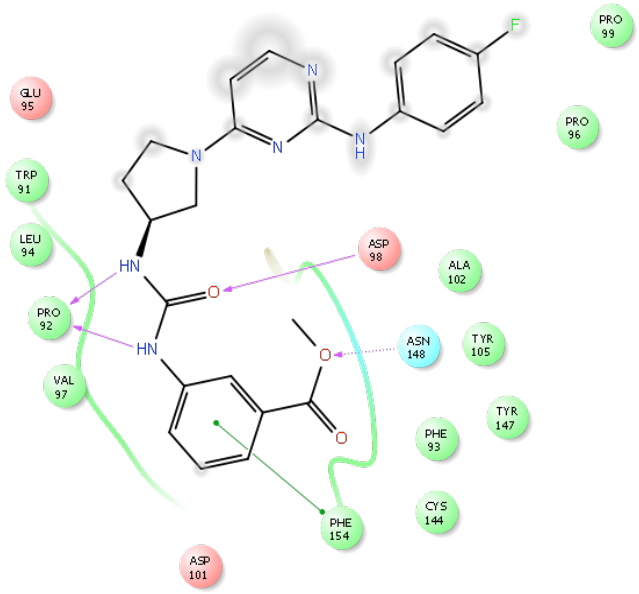


Figure 4.16 A per-residue interaction map of AU1 with BPTF developed from QM/MM docking studies in Glide.
Please note that the residue numbers do not correspond to standard numbering for BPTF. The phenylalanine engaging in the pi-pi interaction is F2887 and the asparagine hydrogen bonding to the oxygen of AU1 is N2881. While (S)-AU1 is shown here, the SAR of the different stereoisomers has yet to be elucidated.

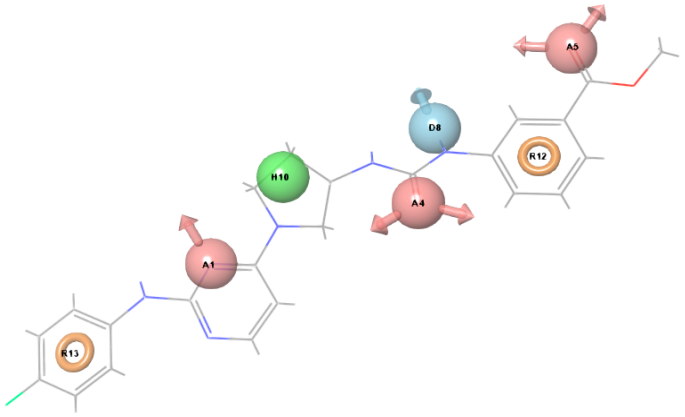


Figure 4.17 A pharmacophore model of the aryl-urea compounds generated using Phase.
Of particular interest are R12 and A5, both of which seem to correlate heavily with binding.

Table 4.3 A list of arylurea compounds used for the phase analysis, and their required activity by the generated pharmacophore model.

Name	Pharm Set
GSK1379706A	inactive
GSK1379710A	inactive
GSK1379712A	inactive
GSK1379714A	inactive
GSK1379717A	inactive
GSK1379720A	inactive
GSK1379721A	inactive
GSK1379722A	inactive
GSK1379723A	inactive
GSK1379724A	inactive
GSK1379725A	inactive
GSK1379727A	inactive
GSK1379731A	inactive
GSK1379741A	inactive
GSK1379742A	inactive
GSK1379745A	inactive
GSK1379746A	inactive
GSK1379748A	inactive
GSK1379753A	active
GSK1379762A	inactive
GSK1379763A	active
GSK1379765A	active
GSK1379766A	active
GSK1379767A	active

4.2.6 Dissociation Constant Determination for a Selective BPTF Ligand AU1

Due to the absence of ligands for the BPTF bromodomain, a correlation between DSF thermal stability values and K_d have yet to be reported. In fact, for BI2536 we measured a slight decrease in the T_m for BPTF of -1.2 ^oC consistent with other reports[30], although we have determined a K_d of 37 μM and 67 μM via ITC and FA respectively (Figure 4.11). The origins of the slight destabilization are yet unclear. From the NMR titration of AU1, the bound and unbound resonance were separated by 171 Hz (Figure 4.2), providing an upper bound for the chemical exchange rate.[134] Assuming an

association rate of 1 x 10^8 M^{-1} s^{-1} as a high end for a range of protein-small molecule interactions, (e.g., chymotrypsin:proflavin k_1 = 1.2 x 10^8 M^{-1} s^{-1}),[174] we estimate an upper K_d of 8 μM from this experiment. For a more accurate determination with a non-fluorinated protein, we used ITC as a complementary direct binding assay using unlabeled BPTF. We obtained a K_d of 2.8 μM (Figure 4.8), consistent with our intermediate exchange resonance broadening by PrOF NMR. We are encouraged by this initial potency as a starting point for future cell-based studies, although new analogs are anticipated to be necessary, including the identification of the active stereoisomer and removal of the methyl ester hydrolysis susceptibility. A summary of our recorded data from the biochemical experiments described above is found in Table 4.1.

Although AU1 has been demonstrated to be selective over Brd4, a full selectivity panel against other bromodomains will be needed. A database search using ChEMBL only showed AU1 to be active in five whole cell assays with an EC_{50} of 500 nM carried out by Gamo et al.[175] However, the targets are unknown, which we are following up. Additionally, no kinase activity has been reported for AU1 despite the growing screening use of the PKIS library. These results are encouraging regarding off target effects for this BPTF ligand.

4.2.7 AU1 transcriptional inhibition experiments in live cells.

As an initial test for the effect of BPTF bromodomain inhibition on transcription, we first measured the cellular toxicity against healthy human cells. The prior study by Gamo et al. did not detect cellular toxicity in hepatocytes at 10 μM compound,[175] which is more than 3-fold above the K_d of AU1. We also observe no cellular toxicity against HEK 293T cells after 24 hours at 3 and 10 μM. Only after 48 hours of incubation at 10 μM do cells have a reduced viability (76% viable) (Figure 4.18).

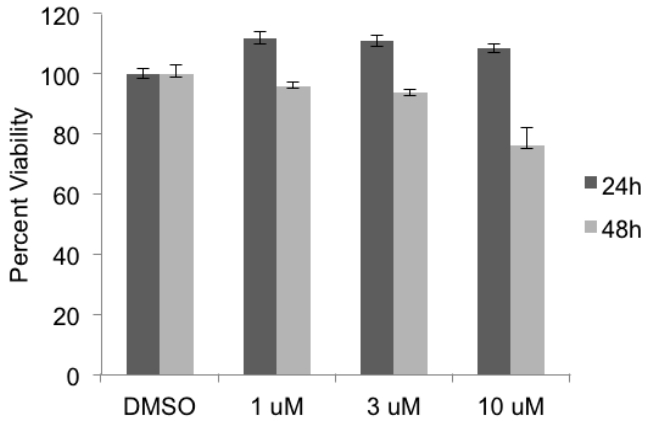


Figure 4.18 Toxicity assay of AU1 with mamallian cells.
HEK 293T cells were plated in 96-well plates. 24h post plating, cells were treated with DMSO or AU1 (1 μM, 3 μM and 10 μM) for 24 (n=6, per treatment) and 48h (n=5, per treatment). Cell viability was assessed by using a CellTiter Blue Viability kit from Promega.

Ruthenberg et al. demonstrated that both the PHD finger and bromodomain of BPTF were important for chromatin binding.[44] Native chromatin precipitation experiments showed colocalization of BPTF at the MOBKL3 locus with histone H4, acetylated at lysine 16.[44] To investigate the effect of AU1 treatment on BPTF-mediated gene regulation, we cloned a BPTF-regulatory region, localized upstream to the Mobkl3 promoter region,[44] into a firefly luciferase reporter plasmid (pGL3_Mobkl3p). Transfection of the pGL3_Mobkl3p into Eph4, a normal mouse mammary gland cell line[176] displayed robust luciferase activity (Figure 4.19), ~7-fold greater than those observed with a control firefly luciferase reporter plasmid (pGL3_empty) lacking a BPTF regulatory region (Figure 4.19). This data supports the association of factors to the Mobkl3 regulatory region, such as BPTF contributed to the increased luciferase activity.

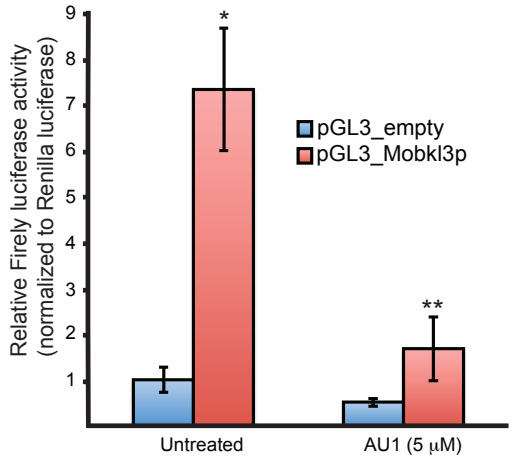


Figure 4.19 AU1 treatment interferes with BPTF activity in cells.
Eph4 normal mouse mammary gland cells were transfected with firefly pGL3_empty (background control), firefly pGL3_Mobkl3p and pRL Renilla control vectors and luciferase activity was measured. Firefly luciferase activity was normalized to Renilla luciferase activity. N = 4 biological replicates. * pGL3_empty pGL3_Mobkl3p, p=0.0002. ** pGL3_Mobkl3p untreated x pGL3_Mobkl3p plus AU-1, p=0.0006.

To test if the pGL3_Mobkl3p luciferase levels were dependent on the bromodomain of BPTF, the Eph4 were again transfected with either the pGL3_Mobkl3p or the pGL3_empty vector and then treated with 5 μM AU1 for 12 hours. A significant loss of luciferase activity (~4-fold) was observed in cells transfected with pGL3_Mobkl3p vector, suggesting that the Mobkl3 regulatory region requires BPTF activity to support the regulation of luciferase activity (Figure 4.19). Together our data suggests that binding of AU1 to the BPTF bromodomain interferes with its regulatory activity in cultured cells. This may represent a suitable system to study the net effects of the BPTF-bromodomain on gene regulation.

4.2.8 Comparison of PrOF NMR to a prior X-ray Crystallography Screen

A combination of PrOF NMR, DSF, and FA have led to several new small molecules for both Brd4 and BPTF. In 2014, two parallel screens using kinase inhibitor libraries against Brd4 were reported providing a useful data set for assessing the application of PrOF NMR in a discovery mode.[29,30] From the x-ray crystallography screen[29], 61 compounds overlapped with the molecules studied here. Of the 28

compounds where a high-resolution crystal structure was obtained (either apo or a co-crystal), PrOF NMR uncovered seven hits while x-ray crystallography uncovered two, with one compound identified by both screening techniques. Although our PrOF NMR results reveal more binders than this screen with the PKIS set; both methods discovered ligands with similar structural classes. In combination, PrOF NMR offers a complementary technique to x-ray crystallography, potentially guiding efforts for subsequent crystallization studies and improving on the structural biology efforts targeting bromodomain ligand discovery.

In conclusion, this study presents the first simultaneous screen of two proteins by PrOF NMR, enhancing the throughput and stringency for finding selective ligands for structurally related proteins. This study is reminiscent of RAMPED-UP NMR by Zartler et al. highlighting the screening potential of three differentially labeled proteins via 2D-HSQC NMR experiments.[133] NMR screening has typically been employed in the area of fragment-based discovery. The hinge-binding kinase inhibitor classes have proven to be a privileged structural class for bromodomains allowing efficient exploration of chemical space and may be useful against other bromodomain family members outside the BET class. Several p38α B-Raf inhibitors with submicromolar affinity for Brd4 will allow for the testing of synergistic effects of dual protein inhibition in cancer and inflammation. Our newest ligand for BPTF, AU1, represents a useful starting point for testing the unexplored role of its bromodomain in various cancer models. Biological studies are currently underway. Our method has been applied to BPTF and BET bromodomain Brd4. Discovery of highly isoform selective ligands for bromodomains within the BET bromodomain family has yet to be successful. In the future, PrOF NMR screening in the presence of multiple BET bromodomains may provide a method to achieve this goal.

4.3 **Future Directions**

4.3.1 The role of the bromodomain of BPTF and optimization of AU1

Since initial publication of the preceding work, the structure of AU1 has been further optimized and preliminary data obtained to assess its effect in diseased cells. The

S enantiomer was confirmed to be the active enantiomer against BPTF, with an affinity confirmed by microscale thermophoresis (Figure 4.20). The *R* enantiomer showed no binding by PrOF NMR. The methylester as well as the urea were found to be important for BPTF binding, while the aniline portion of the molecule can be substituted with little effect on binding (Figure 4.21). The urea seems to be partially responsible for the poor solubility of this compound, as remove of the on the urea nitrogen results in a more soluble compound with a reduced binding affinity. In order to better enable successful SAR studies with AU1, we have attempted co-crystallization studies with (S)-AU1 and the bromodomain of BPTF. While we have successfully crystalized the apo-structure of BPTF several times, attempts to generate a co-crystal structure are still ongoing (Figure 4.22).

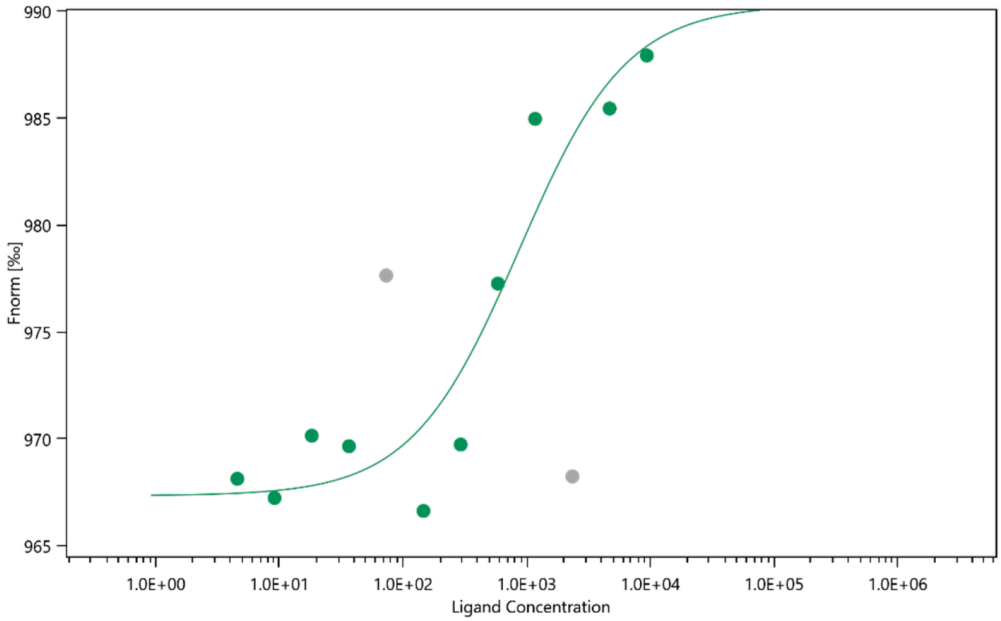


Figure 4.20 Microscale thermophoresis trace of (S)-AU1 with BPTF, resulting in a Kd of 807 nM +/- 440 nM.

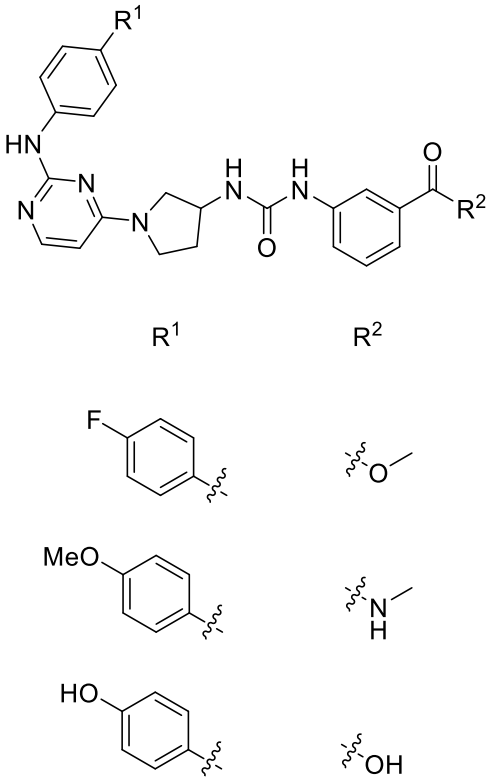


Figure 4.21 Limited set of derivatives of AU1 that have been tested for binding against BPTF.

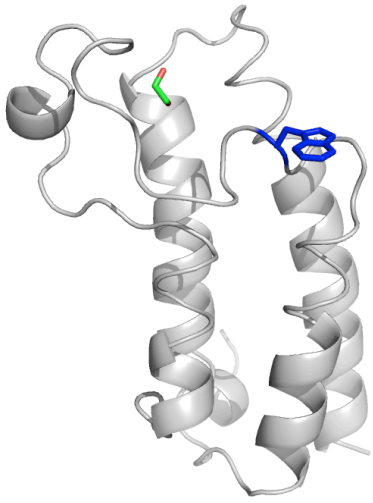
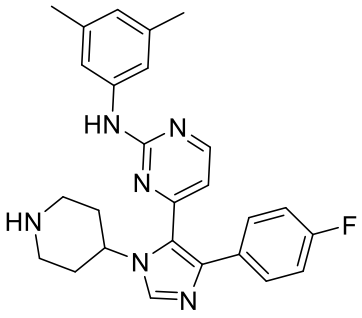


Figure 4.22 X-ray crystal structure of BPTF with ethylene glycol in the binding site. Encouraged by how prone this protein is to crystallization, co-crystallization attempts with BPTF and various aryl-urea analogs are ongoing.

Exposure of melanoma cells to (S)-AU1 in a synergistic experiment with vemurafenib resulted in almost complete reduction of cell proliferation. Further experiments will be assessed when a more potent version of (S)-AU1 can be developed.

4.3.2 Development of HU-10

The trisubstituted imidazole compound SB-284851-BT was further elaborated into HU-10 to increase binding affinity for Brd4 (Figure 4.23). It is being used to elucidate the polypharmacology of inhibiting the bromodomain Brd4 and kinases simultaneously.



HU-10

Figure 4.23 The structure of HU-10

4.4 **Methods**

Expression of 5FW-Brd4 (42−168), and 5FW-BPTF (2793−2911):

3FY and 5FW labeled Brd4(1) were expressed based on established methods[99] using E. coli Bl21(DE3) + pRARE strains. To express the labeled protein, the secondary culture in LB media was grown until an O.D. at 600 nm of 0.6 was reached followed by harvesting. Cells were resuspended in defined media of Muchmore et al.[110] containing 5-fluoroindole (60 mg/L) in place of tryptophan[96]. The resuspended E. coli were incubated at 37 °C while shaking for 1.5 h followed by the cooling to 20 °C and media temperature equilibration for 30 min. Protein expression was induced with 1 mM IPTG overnight (14−16 h) at 20 °C. The cells were harvested and stored at −20 °C. 5FW-BPTF was expressed by the same protocol. Cell pellets were thawed at RT followed by the addition of lysis buffer (50 mM Phosphate pH 7.4, 300 mM NaCl) containing protease inhibitor PMSF (5 mM) as well as the Halt protease inhibitor and purified according to methods

described in the appendix using Ni-affinity chromatography. Yields following purification are 88 mg/L 5FW-Brd4 (>95% fluorine incorporation), 84 mg/L 5FW-BPTF (>95% fluorine incorporation). Purity of proteins was assessed by SDS-PAGE. Fluorinated amino acid incorporation efficiency in proteins was measured by mass spectrometry as described in the appendix. Concentration was determined via absorbance at 280 nm.[177]

Protein Observed Fluorine NMR:

1D ^{19}F NMR Parameters. ^{19}F NMR spectra were acquired at 470 MHz on a Bruker 500 spectrometer with a 5 mm Prodigy TCI Cryoprobe without proton decoupling. Samples containing 40−50 μM bromodomains were labeled in 50 mM TRIS, 100 mM NaCl, and 5% D_2O, pH 7.4 with 100 μM ligand for binding assays unless otherwise stated. Spectra were referenced to trifluoroacetate (−76.55 ppm). Measurement parameters included a relaxation delay time of 0.7 s for 5FW-Brd4(1) and a 90° flip angle. An acquisition time of 0.05 s was used for all experiments. A sweepwidth of 10 ppm was used for 5FW-Brd4 spectra. All screening experiments used 400 transients. Small molecules were titrated into the protein solution from concentrated stock solutions of DMSO (10 mM for all small molecules except acetaminophen). Final DMSO concentrations were kept at or below 1% DMSO.

Differential Scanning Fluorimetry (DSF):

The inhibitory activities of compounds against Brd4 were assessed by DSF using a CFX384 Real-Time PCR Detection System (BioRad). Purified Brd4 (4 μM final concentration; 10 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM DTT) were assayed, in replicates of ten, in a 384-well plate. Inhibitors were added to a final concentration of 100 μM and 1 % DMSO or ethylene glycol. SYPRO Orange Protein Gel Stain (1:5000; Sigma Aldrich) was used as the fluorescent probe. Protein stability was using 0.1 °C increments and 10 s incubations per increment from 25 to 75 ^{o}C. The inflection point of the transition curve/melting temperature (T_m) was calculated using the Boltzmann equation within the Protein Thermal Shift Software (v.1.1) (Applied Biosystems). (+)-

JQ1 and dinaciclib were used as controls for strong and weak binders of Brd4, respectively. The ΔT_m was calculated by using DMSO and ethylene glycol control wells as a reference.

Fluorescence Anisotropy Direct Binding and Competition Experiments:

Fluorescence anisotropy was measured from an absorption wavelength of 485 nm and an emission wavelength of 535 nm on a Tecan Infinite 500 plate reader using 384-well plates (Corning 3676). All experiments were carried out in 50 mM Tris, 100 mM NaCl, 5 mM CHAPS at pH 7.4. The fluorescently labeled tracer, BODIPY-BI2536, was synthesized according to published methods.[165] For both direct binding and competition experiments, 25 μM stocks of a BODIPY-BI2536 stock solution in DMSO were diluted to a final concentration of 25 nM. For direct binding experiments, Brd4 and BPTF were serially diluted from micromolar to subnanomolar concentrations. For competition experiments, Brd4 was kept at a constant concentration of 156 nM, equivalent to 80% bound tracer as determined in direct binding experiments, and the concentration of competing ligand was serially diluted from micromolar to subnanomolar concentrations. Because of the capacity of DMSO to perturb bromodomain binding, competitors were diluted from 10 mM ethylene glycol stocks. Data were collected within 30 minutes after plating to minimize Brd4 binding to the plate surface. Measurements taken after 30 minutes led to a significant decrease in the apparent dissociation constant at longer time points. All experiments were carried out in triplicate. Acquired data were fit using GraphPad Prism. K_d values were determined by fitting to the following equation, which accounts for ligand depletion. In this equation, b and c are the maximal and minimal anisotropy values, respectively, a is the concentration of fluorescently labeled tracer, x is the protein concentration, and y is the observed anisotropy value.

$$y = c + (b - c)\frac{(Kd + a + x) - \sqrt{(Kd + a + x)^2 - 4ax}}{2a}$$

IC_{50} values were determined using GraphPad Prism's log(inhibitor) vs response function. K_i values were obtained using a variant of the Cheng-Prussof equation from Huang et al.[102]

Isothermal Titration Calorimetry:

A Nano-ITC was used to perform isothermal titration calorimetry experiments. AU1 (46 μM) was diluted into the same buffer as BPTF (50 mM sodium phosphate pH 7.0) from a 10 mM DMSO stock solution. The resulting solution was sonicated immediately prior to use. Titrations were carried out by using a first injection of 0.5 μM followed by 19 injections of 2.5 μL volume of 635 μM BPTF with a releasing period of 8 seconds per injection with 300 second spacing time between injections. Heat liberated by each injection was integrated (ΔH) and plotted against molar ratio of ligand and protein. The obtained data was fit by using predefined two binding site model. The fitting of data was not sufficient with one binding site whereas a two binding site model produced an appropriate fit, indicating a potential low affinity binding (K_d = 1.0 mM) site. Data was analyzed by NanoAnalyze software to calculate enthalpy of binding (ΔH) and dissociation constants (K_d). A two binding site binding model was used for data fitting. ITC data was obtained in a similar fashion for BI2536, using a concentration of 0.5 mM of BI2536 and 7 mM BPTF.

Viability Assays:

Cell viability assay: Cell viability assay of HEK 293 cells were performed using resazurine dye based CellTiter-Blue® (Promega). HEK 293T cells were plated in 96 well plates and experiments conducted when cells were 80% confluent. Cells were treated with the 0, 1.0, 3.0 and 10.0 μM AU1 for 24 and 48 hours in 5% CO_2 (n=5-6, per condition). 20 μL of CellTiter-Blue® was added to the each well and incubated for 2.5 h at 37^{o}C. As resazurin dye is reduced by viable cells to resorufin. Resorufin is fluorescent at 580 excitation and 590 emission. The data is normalized with the control (DMSO treated), (mean +/- SE).

Luciferase reporter array:

The BPTF-regulatory region localized upstream to the Mobkl3 promoter region (~380bp) was amplified by PCR, cloned into pCR2.1-TOPO vector (Invitrogen) and

transferred to the firefly luciferase pGL3-Basic vector (Promega) using XhoI and HindIII restriction enzymes. pGL3_Mobkl3p and pGL3_empty vectors were propagated in TOP10 chemically competent bacteria (Invitrogen) and plasmid purification was performed with Qiagen kit (Qiagen). Purified plasmids were co-transfected into Eph4 cells, a normal mouse mammary gland cell line (gift from Dr. Celeste M. Nelson) with pRL Renilla luciferase control vector (Promega) using PEI (Sigma). Transfection was carried out for 8 hours, followed by media change and treatment with AU1 (12 hours). Cells were then washed once with 1xPBS and lysed with Passive Lysis Buffer (Promega) for 20 minutes with constant agitation. Luciferase activity was measured with the Dual-Glow luciferase reporter assay kit (Promega) on a Glomax 20/20 luminometer (Promega). Firefly luciferase levels were normalized to Renilla luciferase levels.

Protein Expression and Molecular Biology Materials:

For *E. coli* growth, LB agar, LB media, defined media components including unlabeled amino acids, uracil, thiamine-HCl, nicotinic acid, biotin and buffer components were purchased from RPI corp. Thymine, cytosine, guanosine were purchased from Alfa Aesar. Magnesium chloride, manganese sulfate, succinic acid, calcium chloride and 5-fluoroindole were purchased from Sigma-Aldrich. Miniprep plasmid purification kit was purchased from Clontech.

Unlabeled Brd4, and BPTF Protein Expression:

The pNIC28-BSA4 plasmid containing the first bromodomain of Brd4 and BPTF genes were kind gifts from the laboratory of Stefan Knapp. For protein expression, either the *E. coli* Rosetta (DE3) strain (Novagen) was first transformed with the respective expression plasmid or the BL21(DE3) strain was cotransformed along with the pRARE (Novagen) plasmid and plated onto agar plates containing kanamycin (100 mg/L) and chloramphenicol (35 mg/L). Following overnight incubation at 37 °C, a single colony was selected from the agar plate and inoculated in 50 mL of LB media containing kanamycin (100 mg/L) and chloramphenicol (35 mg/L). The primary culture was grown overnight at 25 °C while shaking at 250 rpm. For secondary culture growth, 1 L of LB

media containing kanamycin (100 mg/L) was inoculated with the primary culture and cultured at 37 °C while shaking at 250 rpm. When the O.D. of culture at 600 nm reached 0.6, the shaker temperature was reduced to 20 °C. After 30 minutes, the expression was induced with 1 mM IPTG overnight for 12-16 h. Cells were harvested by centrifugation and stored at -20 °C.

Bromodomain Purification:

To purify fluorinated and unlabeled Brd4, the cell pellet was thawed at room temperature followed by the addition of lysis buffer (50 mM Phosphate pH 7.4, 300 mM NaCl) containing protease inhibitor PMSF (5 mM) as well as the Halt protease inhibitor. Cells were lysed by sonication and the cell lysate was centrifuged at 7500 g for 30 minutes followed by supernatant filtration over Whatman filter paper. Filtrate containing the histidine-tagged Brd4 was loaded on to a nickel-NTA affinity column and eluted with an imidazole gradient on an AKTA FPLC system monitoring the O.D. at 280 nm. Imidazole was removed from the buffer using a HiPrep column (GE) for buffer exchange into 50 mM Tris pH 7.4, 100 mM NaCl. Purified and buffer exchanged protein was treated with TEV protease for either 2 hours at room temperature or alternatively at 4 °C overnight on a rotating carrousel. The cleaved His-tag, TEV protease and uncleaved Brd4 were removed using nickel-NTA affinity resin.

Protein Mass Spectral Analysis

Product molecular weight was confirmed by electrospray ionization mass spectrometry (ESI-MS) using a Thermo Scientific Orbitrap Velos LC-MS, reported previously. To determine the percent incorporation for fluorinated proteins the integration values of the different deconvoluted mass peaks are entered into the following equation to determine the relative incorporation e.g, FWBrd4:

$$\%_{incorporation} = \frac{(0FWBrd4)*0+(1FWBrd4)*1+(2FWBrd4)*2+(FWBrd4)*3}{(0FWBrd4)*3+(1FWBrd4)*3+(2FWBrd4)*3+(FWBrd4)*3} * 100$$

0FWBrd4 is 5FWBrd4 with no fluorine substitutions, 1FWBrd4 is 5FWBrd4 with one fluorine substitution, 2FWBrd4 has two fluorines substituted and FWBrd4 has 3 fluorines substituted.

Circular Dichroism:

To check the secondary structural content, far-UV CD spectra (200-260 nm) of unlabeled and labeled proteins were collected using a peltier equipped temperature controlled Jasco J-815 spectropolarimeter at 25 °C. For all measurements, 20 μM (50 mM Tris buffer pH 7.4 containing 100 mM NaCl) of protein and a 1 mm cuvette path-length were used. Spectral data were collected at a scan rate of 50 nm/min with averaging of 10 spectra. Processed data were baseline corrected against spectra taken with buffer alone.

Thermal Melting:

Thermal stabilities of labeled and unlabeled proteins were measured by the change in ellipticity at 222 nm with the increase in temperature from 20 °C to 80 °C at the scan rate of 60 degrees/h.

K_d estimation of AU1 by NMR:

The upper limit of k_{off} can be determined by the following equation:

$$\tau_{coalescense} = (\sqrt{2}\pi\Delta\upsilon)^{-1} = k^{-1}$$

Where $\Delta\upsilon$ is the energy difference between the two resonances in Hz and $k^1 = k_{off}$. The k_{on} rate can be estimated as diffusion controlled (1.0×10^8 M^{-1}/s). The following equation can then be used to estimate an upper limit to the K_d:

$$K_d = \frac{k_{off}}{k_{on}}$$

The energy difference between the bound and unbound resonance of 5FW-BPTF when binding with AU1 is a difference of 171 Hz, which results in an upper limit K_d of 8 μM

Docking Studies of AU1 with Brd4 and BPTF:

Brd4 (PDB ID: 3UVW) and BPTF (PDB ID: 3QZT) were prepared using the Protein Preparation Wizard workflow as follows: adding hydrogen, assigning partial charges using the OPLS-2005 force field, and assigning protonation states. AU1 was prepared using the Ligprep module, generating protonation states from pH range 6.2-8.2. The docking grid was generated by removing the bound peptides and analyzing the protein for binding sites using SiteMap with more restricted notion of hydrophobicity and fine grid. The known acetylated-lysine binding site was selected as the binding site to generate the grid. All tyrosines were allowed to rotate during the docking process. AU1 was then docked using QM/MM ligand polarized docking with initial docking being Glide XP with ligand-van der waals scaling of 0.8. The partial charges of the ligand atoms were calculated using B3LYP with 6-31G*/LACVP* basis set with Ultra-fine SCF accuracy level. AU1 was then redocked using Glide XP and sorted by gscore. Gscores were favorable for AU1 binding to both BPTF and Brd4, so predictions based upon the Gscore are unlikely to be helpful. The docking study was primarily used to suggest binding poses for further analysis. One such pose showed the methyl ester of AU1 hydrogen bonding with asparagine 2881 and a $\pi-\pi$ interaction of the aryl ring with phenylalanine 2887.

Pharmacophore modeling studies:

Pharmacophore modeling was carried out using PHASE (version 3.9, 2014) module of the Schrödinger suite implemented in Maestro (Maestro 2014-2) molecular modeling package. PHASE identifies the spatial arrangement of functional groups that are common and essential for the biological activity of the compounds. The structures of compounds were prepared using LigPrep (version 3.0, 2014). Conformers were generated using ConfGen by applying OPLS-2005 force field. Four features/sites were considered in generating pharmacophore variants: hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H) and aromatic ring (R). The maximum number of sites was set to 7 and minimum to 3. Common pharmacophores were required in all 5 of the active compounds with a final box size of 1 Å and minimum intersite distance of 2 Å.

Resulting pharmacophore hypotheses were then scored using default weights of scoring parameters for both active (survival score) and inactive ones (survival-inactive scores). Many of the pharmacophore sites generated in the hypotheses were consistent across many of the developed models due to the similarity of the various aryl-urea compounds analyzed, but location of the aryl ring with a hydrogen-bond acceptor present at the carbonyl seemed to correlate heavily with binding.

Cell Viability Experiments in HEK 293T cells:

A cell viability assay of HEK 293T cells was done by using a resazurine dye based CellTiter-Blue® from Promega assay. HEK 293T cells were plated in a 96 well plate and grew until 80% confluency. Desired confluent cells were treated with the 0, 1, 3 and 10 μM AU1 for 24 and 48 hours in 5% CO_2 (n=6 for 24 hours and n=5 for 48 hours treatment, per condition). After the desired time point of treatment 20 μL of CellTiter-Blue® was added to the each well and incubated for 2.5 h at 37 ^{o}C. The fluorescence of the reduced dye, resofurin was measured at 580 excitation and 590 emission wavelengths by using an omega plate reader. The presented data is normalized with the control (DMSO treated), error bars are depicting the standard error of the mean.

Table 4.4 Summary of NMR screening data including change of Chemical Shift and line broadening upon ligand addition for Brd4 resonances and BPTF

	W75		W2824		W120		W81	
Compound Regno	Δδ	ΔLW	Δδ	ΔLW	Δδ	ΔLW	Δδ	ΔLW
GSK260205A	0.0928	-0.0194	N/A	N/A	-0.0033	0.0483	N/A	N/A
GSK1379767A	-0.0104	-0.0273	N/A	N/A	-0.0007	-0.0076	N/A	N/A
SKF-106164-A2	-0.0177	0.0052	N/A	N/A	0.0275	0.0406	N/A	N/A
GSK1379753A	-0.0101	-0.0005	N/A	N/A	0.0067	-0.0107	N/A	N/A
GSK1379710A	-0.0005	-0.0078	N/A	N/A	-0.007	-0.0034	-0.0511	0.0255
GSK1379731A	-0.0062	-0.0025	N/A	N/A	-0.0002	0.0151	-0.058	0.0292
SB-744941	-0.0004	0.0006	N/A	N/A	0.0085	0.0337	0.0014	0.0037
GSK1379746A	-0.015	0.0061	N/A	N/A	-0.0028	-0.0026	-0.0244	0.0086
GSK1379725A	0.0006	0.003	N/A	N/A	0.0013	0.0064	-0.0092	0.0096
GSK1379765A	-0.0585	0.0187	-0.1275	0.0104	-0.0061	-0.0126	N/A	N/A
SB-738004	-0.0385	0.0399	0.0098	-0.0212	0.0199	0.0256	N/A	N/A
SKF-86055	-0.0043	0.0069	-0.1329	0.0635	0.0072	-0.0047	N/A	N/A

GW876790X	0.0007	0.0325	-0.0468	0.0393	0.0019	0.0377	0.0049	0.0417
GSK949675A	0.0024	-0.0046	-0.1232	0.0113	-0.0049	0.0111	0.0186	0.0313
GSK711701A	-0.0716	0.0211	0.1324	0.0146	0.0171	-0.0073	0.0656	-0.009
SB-732881	-0.038	-0.0039	-0.1668	0.0252	-0.0081	0.0061	-0.0589	0.037
GSK1007102A	-0.015	-0.0055	-0.1787	0.014	0.0062	0.0151	0.1039	-0.0016
GSK619487A	0.0005	0.0073	-0.1237	0.0125	-0.0042	-0.0142	0.0189	-0.0087
GW301789X	0.0009	0.0148	-0.0361	0.0304	-0.0098	-0.0137	0.0049	-0.0018
GSK507274A	0.0033	-0.0148	-0.1481	0.0313	0	0.0132	0.0177	-0.0008
GSK902056A	-0.0022	-0.0197	-0.162	0.0279	-0.0015	0.0071	0.011	0.0125
GSK561866B	0.0047	0.002	-0.3035	0.0405	-0.0073	0.0312	0.0066	0.0008
GW445017X	-0.0105	0.0047	-0.0577	0.0395	0.0007	-0.0129	0.0095	0.0099
SB-736290	-0.0072	-0.0049	-0.0313	-0.0224	-0.007	0.0047	0.0059	0.0067
GSK943949A	-0.0084	-0.0079	-0.1467	0.0217	-0.0033	0.0103	0.0569	0.0131
GW679662X	0.0377	-0.0159	0.0481	0.0038	0.0223	-0.0099	N/A	N/A
SB-242719	-0.0727	0.0343	-0.0176	0.0175	0.0145	0.0148	N/A	N/A
GSK2219329A	-0.0552	0.018	-0.057	0.0234	-0.0029	0.0009	N/A	N/A
SB-253228	-0.0378	0.0131	-0.0156	-0.0107	0.0074	-0.025	N/A	N/A
GSK1379763A	-0.0438	0.0042	-0.0832	0.0036	-0.0075	-0.0173	N/A	N/A
GSK620503A	-0.0456	0.0176	0.0653	0.0313	0.0069	-0.0023	N/A	N/A
GW684941X	-0.0434	0.0006	-0.0908	0.0267	0.0017	0.0115	N/A	N/A
GSK1379766A	-0.0169	0.0059	-0.0739	-0.0032	-0.0073	-0.0051	N/A	N/A
SB-251527	-0.0177	-0.0064	-0.0045	-0.0082	0.0054	0.0067	N/A	N/A
GSK312948A	-0.0676	0.0269	0.0453	-0.0055	0.0029	-0.001	0.1733	0.0003
GW873004X	0.0009	-0.0008	-0.0216	0.0167	0.0007	0.0244	0.0174	0.0324
SB-708998	0.007	-0.0307	-0.0066	0.0154	0.0002	-0.0052	-0.0312	0.0047
SB-708999	0.0544	-0.0093	0.0018	-0.0093	0.0051	-0.0056	0.1249	0.0052
GSK2297542A	-0.0338	0.0012	0.0524	-0.0012	0.0166	-0.0035	0.1581	0.0136
GSK1379748A	-0.0446	0.0093	-0.0904	0.0226	-0.0058	-0.0089	0.1348	0.0568
GW782907X	-0.0141	0.0151	0.0592	0.0127	0.0141	-0.0248	0.0212	-0.0156
SB-711805	-0.0435	-0.0162	-0.0926	0.0061	0.0019	0.0003	0.1369	0.004
SB-390527	-0.0097	0.0125	-0.0301	0.0189	0.007	0.0103	0.0294	-0.0101
GW305178X	-0.0018	-0.0029	-0.0795	0.0065	-0.0219	-0.008	0.1015	-0.0064
SB-751148	-0.0135	0.0102	-0.0631	0.0169	0.0063	-0.0255	0.1431	-0.007
GW620972X	0.0057	-0.0046	-0.0268	0.0173	-0.0062	0.013	0.0024	-0.0089
SB-698596-AC	-0.0199	-0.0023	-0.0769	0.0027	-0.0089	0.0134	0.0197	-0.0086

GW784307A	0.0046	0.0037	-0.0227	0.0165	0.0009	-0.02	0.0128	-0.0147
GW786460X	-0.0174	-0.0009	-0.0097	-0.0116	-0.0018	0.0009	0.026	-0.0102
SB-360741	-0.0153	-0.0033	-0.001	0.0201	0.0016	0.0099	0.0192	-0.0085
GW276655X	-0.0086	-0.0013	-0.0116	-0.0118	0.0019	-0.0046	0.0205	-0.0087
GW694077X	-0.0002	-0.002	-0.0738	0.0003	0.0063	0.0086	0.0926	-0.0093
SB-707548-A	-0.1039	0.0316	-0.0095	-0.0093	0.0091	-0.0065	0.0826	0.0005
GSK1379723A	-0.0086	-0.0188	-0.0186	-0.0093	-0.0013	-0.0177	0.0351	0.0013
GSK2297099A	-0.0416	0.0053	0.0542	-0.0046	0.0105	-0.0073	0.0464	0.0054
GW781483X	0.0154	0.0065	-0.0872	0.0154	-0.0028	0.0127	0.0743	-0.0027
GSK192082A	0.0013	0.0118	-0.0129	0.0154	0.0009	-0.023	0.0094	-0.0049
SB-747651-A	0.0018	0.0165	-0.0652	0.0003	-0.0016	0.0035	0.0128	0.0038
GW796410X	-0.0108	-0.0088	-0.006	-0.0084	0.0082	0.0272	-0.0056	0.0022
GSK1379714A	-0.0101	-0.0045	-0.0311	-0.0093	-0.0043	-0.0179	0.0147	-0.0041
GSK1379720A	-0.0219	-0.0014	0.0329	0.0018	-0.0021	-0.0147	0.0139	0.0019
GSK1535721A	-0.0244	0.0055	-0.0952	0.0045	0.0106	-0.0029	0.0407	-0.0079
GW778894X	-0.0187	0.0111	0.036	0.0051	0.0107	-0.0002	0.0352	0.0028
GSK614526A	0.0031	0.0037	-0.1008	-0.006	-0.005	0.0221	0.0186	0.0024
GW300660X	-0.0141	-0.0016	-0.0757	0.0097	-0.0036	-0.0082	0.0457	0.0122
GW811761X	0.0011	-0.0031	-0.0033	-0.0109	0.0008	-0.0067	0.0069	0.0084
GW827105X	-0.0035	-0.0031	0.0052	0.0209	0.0011	0.0069	0.032	0.0142
GW810372X	-0.0055	0	-0.0068	-0.0099	-0.0043	0.006	0.004	-0.007
SKF-104365	-0.0018	-0.0073	-0.0145	0.0246	-0.0016	-0.0029	0.0009	0.008
GSK1030058A	0.01	0.0073	-0.0239	-0.0087	0.0005	0.0007	-0.0114	-0.0049
GW440137A	-0.0239	-0.0095	-0.0117	0.0161	-0.0041	-0.0073	0.0575	0.0034
GW709213X	-0.0059	0.0013	0.0019	-0.0111	0.0018	-0.0044	0.0761	0.0121
GW693481X	-0.0056	0.0032	-0.0165	0.016	0.0036	-0.007	0.0676	0.0056
SB-400868-A	-0.0047	0.0031	-0.0056	0.0201	0.0071	0.0017	-0.001	0.0085
SB-772077-B	-0.0168	0.0075	0.029	0.0024	0.0075	-0.0119	0.0376	0.0005
SB-750140	-0.0011	-0.0088	-0.0072	-0.0139	-0.0014	0.0021	0.0313	0.0072
SB-751399-B	0.003	-0.0023	0.0076	-0.0171	0.0015	-0.0071	0.0139	-0.0025
GW439255X	N/A	N/A	-0.0137	0.0029	0.0007	0.0119	N/A	N/A
SB-264866	-0.0717	0.0281	-0.0124	-0.0002	0.0075	0.0251	N/A	N/A
SB-284851-BT	-0.0779	-0.0229	-0.0195	0.0042	0.02	0.0107	N/A	N/A
GW837331X	-0.0686	0.017	-0.0591	0.0056	0.0018	-0.0055	N/A	N/A

SB-358518	-0.072	0.0104	-0.0176	0.0117	0.0026	-0.005	N/A	N/A
SB-590885-AAE	-0.0401	0.0098	-0.0037	0.0122	-0.0146	0.0252	N/A	N/A
SB-250715	-0.041	0.0096	-0.0396	0.0037	0.0028	0.0249	N/A	N/A
GW622475X	-0.0389	0.0037	-0.033	0.0065	0.0142	0.0086	N/A	N/A
GSK317354A	-0.0444	0.0006	0	0.0079	-0.0067	-0.0106	N/A	N/A
GW407323A	-0.0415	0.0234	0.0197	0.0003	0.0037	-0.0135	N/A	N/A
GSK2220400A	0.0311	0.0017	-0.0098	-0.002	0.0002	0.0017	N/A	N/A
GSK1023156A	-0.0324	0.0115	0.011	-0.0066	0.0078	-0.0116	N/A	N/A
SB-254169	-0.0186	0.0068	-0.0225	0.0061	0.013	0.0034	N/A	N/A
SB-223133	-0.0024	0.0017	-0.0048	0.0058	-0.004	-0.0079	N/A	N/A
GW702865X	-0.0093	0.0081	0.0062	-0.0053	-0.0003	-0.0075	N/A	N/A
GSK1321565A	-0.0537	0.0024	-0.0186	-0.0061	0.0053	-0.0041	-0.2799	0.0237
SB-217360	0.0177	-0.008	-0.0019	-0.0067	0.0245	-0.0046	0.3748	0.0054
SB-732941	-0.0144	0.0144	-0.0159	0.0102	-0.0058	-0.008	0.1942	-0.0076
SB-710363	0.0262	0.0126	-0.0553	-0.0021	-0.0039	-0.0011	0.377	-0.0001
GW824645A	0.0005	-0.0005	0.0162	-0.0028	-0.0009	-0.0128	0.1708	0
GW493036X	-0.0299	-0.0014	0.0155	-0.0055	0.0096	-0.0047	0.2674	0.0277
GSK1173862A	0.0034	0.006	0.0068	-0.0072	-0.0038	-0.0105	0.0661	0.0328
GW57482A	0.0009	0.016	0.0069	-0.0025	0.0073	-0.0259	0.009	-0.0104
SB-226605	-0.0041	-0.017	0.0013	-0.0032	0.0115	-0.002	0.1069	0.0029
GW335962X	0.0034	0.0157	-0.0552	-0.0035	-0.0077	0.0088	0.0105	-0.0088
GW831090X	-0.0336	-0.012	-0.0239	-0.0064	0.0014	-0.0014	-0.0476	0.042
GSK1653539A	0.0105	-0.013	-0.0051	-0.0013	-0.0035	-0.0013	0.0213	0.0215
SB-814597	0.0093	0.0078	-0.0322	0.0032	-0.0196	-0.014	-0.0026	0.022
GW583373A	0.0015	-0.0077	-0.0048	-0.0057	-0.0088	-0.0038	0.0145	0.017
GSK1379727A	-0.0196	-0.0052	-0.0417	0.0079	-0.0113	0.0021	-0.0481	0.0075
GI261520A	-0.0152	0.0113	-0.0044	-0.005	0.0055	0.0154	0.0281	-0.0169
GW771127A	-0.01	0.004	-0.0192	-0.0061	0.0027	0.0145	0.012	-0.0132
GW632580X	-0.0111	-0.007	-0.0109	0.0014	-0.004	0.013	-0.0164	-0.0114
GW525701A	-0.0239	-0.004	0.0145	0.0036	0.0105	-0.0066	0.1124	0.0225
GW697465A	0.0102	-0.002	0.0165	-0.004	0.0136	-0.0062	0.0104	-0.011
SB-317658	0.0065	-0.006	0.0117	0.0006	0.0092	-0.0009	0.025	0.0165
GW856804X	-0.011	0.0073	-0.0412	0.0039	-0.0069	-0.0062	0.0151	0.0186
SB-737198	-0.0017	0.0021	-0.0458	0.0136	0.0099	-0.0114	0.0575	-0.0115
GW711782X	0.0006	-0.0048	-0.009	-0.0017	0.0014	-0.0108	0.1141	0.0051
GSK1379762A	-0.0069	0.001	-0.0046	0.0133	0.0076	0.0058	-0.0076	0.0261

GW576609B	-0.005	-0.0062	-0.0051	-0.0053	0.0078	0.0057	0.0183	-0.009
GW607117X	-0.0039	-0.0079	0.0034	0.0027	0.0023	-0.0053	-0.0158	0.0223
GW651576X	-0.0043	-0.0106	-0.0055	0.0015	0.0014	0.002	0.005	-0.0157
GW673715X	-0.0097	0.0057	-0.0048	0.0056	0.0007	0.0021	0.0147	-0.0125
GSK466314A	0.0024	0.0067	-0.0165	0.0006	-0.0028	-0.0059	0.0306	-0.0166
GW572399X	-0.0087	-0.0003	-0.0144	-0.0057	0.0022	-0.0023	0.0377	-0.0163
GW829877X	-0.0043	-0.0047	-0.0033	-0.0006	-0.0027	0.0031	0.0089	-0.0155
GSK257997A	0.0035	-0.0055	-0.0229	0.0029	0.0016	-0.0015	0.0314	-0.0108
SB-211743	-0.0026	0.0017	-0.0357	0.0011	0.0005	0.0093	0.0291	-0.0088
SB-282975-A	0.0043	0.0105	-0.0276	0.0016	-0.0008	0.0079	0.1219	0.0029
SB-633825	-0.0036	-0.0009	-0.0324	-0.0064	-0.0007	0.0084	-0.0766	-0.0035
GSK180736A	-0.0239	0.0107	-0.004	0.0104	0.0035	-0.0068	0.1289	0.0041
GSK300014A	-0.0054	0.0263	-0.0123	0.0067	-0.0016	-0.0174	0.0067	-0.0005
SB-390526	0.0016	-0.0286	0.0107	0.0067	-0.0002	0.0064	0.0095	0.0024
GSK1379706A	-0.008	-0.0121	-0.0186	-0.0032	-0.0002	-0.019	0.0056	0.0002
GSK1379742A	-0.0138	-0.0113	0.0072	-0.0008	-0.0063	-0.0185	-0.0004	0.0029
GSK1379717A	-0.0007	-0.0188	-0.0068	0.0023	0.0074	-0.0146	0.0441	0.0124
SB-341528	-0.0077	-0.0119	-0.0222	-0.0051	0.0029	0.0129	0.0019	0.0002
SKF-97623	0.0045	0.0136	0.0189	0.0011	0.0105	0.0136	0.0034	-0.0014
GW743024X	-0.0033	0.0221	-0.0088	0.0007	0.0006	-0.0193	-0.0077	-0.005
GW843682X	-0.0356	-0.002	0.006	-0.0019	-0.0027	-0.0058	-0.0079	0.0104
GSK1379745A	-0.0264	-0.0126	-0.0079	-0.0009	-0.0033	-0.0106	-0.0004	0.0028
GW693917X	-0.0113	-0.0108	-0.0023	-0.0058	0.0044	0.0054	0.0089	-0.0052
GW829874X	-0.0122	-0.011	-0.0094	0.0001	0.0044	0.0105	0.0105	-0.0064
GW631581B	-0.0156	-0.0114	-0.0142	-0.0024	-0.0003	-0.0009	0.0274	-0.0033
GW830365A	-0.0179	-0.0122	-0.0294	-0.0052	-0.0031	0.0018	0.0142	0.0112
GW852849X	-0.0438	0.0011	0.0012	0.0079	-0.005	0.0043	0.0414	0.0139
GSK292658A	0.0008	-0.0185	-0.0012	-0.0063	0.0036	-0.0116	0.0077	0.0086
GSK346294A	-0.0024	-0.0141	-0.006	-0.0042	0.0039	0.0115	0.0636	0.0063
GW709199X	0.0033	-0.0197	-0.0071	-0.0013	0.0001	0.0071	0.0012	0.0028
GW828206X	0.0015	-0.017	-0.0009	1E-04	-0.0049	-0.0041	0.0036	-0.002
GW867588X	0.0023	0.0126	-0.0047	0.004	0.0016	0.0023	0.0168	-0.0006
SB-361058	-0.0364	0.0081	-0.0083	0.0113	0.0003	-0.0048	0.0475	0.0009

SB-711237	0.0002	0.0148	-0.0157	0.0046	-0.0044	-0.009	0.0075	0.0082
SB-759335-B	-0.0016	-0.0119	-0.0002	-0.0021	-0.0014	0.0096	0.0151	0.0053
GSK1379741A	-0.0012	0.0055	-0.0131	-0.002	-0.0043	-0.0153	-0.0065	-0.005
GSK1379724A	0.0009	-0.0018	0.0195	1E-04	0.0117	0.0177	0.0304	0.0022
GSK2213727A	-0.0065	-0.0087	0.0089	-0.0061	0.0035	0.0202	0.0604	0.0057
GSK2186269A	0.0034	-0.0026	0.0011	0.0101	0.0096	0.0133	0.0116	-0.0046
GW642125X	-0.0069	0.006	-0.0175	0.0102	-0.0055	0.0052	0.0504	-0.0012
GW695874X	-0.0021	0.0035	-0.0271	0.0056	-0.0052	-0.0056	0.0749	0
GW679410X	-0.0234	-0.0001	-0.0114	-0.0067	-0.0002	0.0141	0.0867	-0.0062
SB-735465	-0.0277	-0.0067	-0.0158	-0.0041	0.0107	0.0127	0.0671	0.0003
SB-738482	-0.023	-0.0023	-0.0251	-0.0041	-0.0054	0.0006	0.0886	-0.001
GI261656A	0.0023	-0.006	-0.0187	0.0026	0.0049	-0.0125	0.0184	-0.0003
GSK1379800A	-0.003	-0.006	0.0243	0.0081	0.0123	-0.0079	0.0904	0.0119
GSK1398463A	0.0093	-0.0011	0.015	-0.004	-0.0055	0.0148	0.0169	-0.0006
GSK429286A	-0.009	-0.0002	0.0193	0.0042	0.0117	0.0083	0.0159	0.0031
GW869979X	0.007	0.0033	-0.0418	-0.0022	0.0106	0.0199	0.0168	-0.0041
SB-249175	-0.0006	-0.0084	0.0154	0.0024	0.0156	-0.0067	0.0601	-0.0073
GSK361065A	-0.0083	0.0093	-0.0024	0.0045	-0.0103	-0.007	-0.0137	0.0069
GSK938890A	0.0061	0.0006	-0.0573	0.0087	0.0122	-0.001	0.0269	-0.0021
GW580496A	0.0064	0.0052	-0.007	0.0111	0.0048	-0.0227	0.0044	0.0026
GW622055X	-0.0054	0.006	-0.0075	0.0062	0.0021	-0.0154	0.056	0.0035
GW641155B	0.0017	0.0055	0.0044	-0.0006	0.0056	-0.0241	0.0084	0.0081
GW795486X	0.0063	0.0009	0.0017	0.0091	0.0084	-0.0137	0.0124	0.0058
GW809897X	-0.0018	0.0116	-0.0077	0.0006	0.0008	-0.0206	0.0012	-0.0069
GW785404X	0.0029	0.0006	-0.0064	0.0089	-0.0043	0.0175	0.0148	0.0079
SB-693162	-0.0104	-0.0093	0.0089	0.0089	0.0005	0.0152	0.0263	-0.0037
GSK718429A	0.0047	0.0004	-0.0029	0.0013	0.0031	-0.0085	0.0202	0.0072
GW569530A	-0.0005	-0.0064	0.0181	0.0003	0.003	-0.0029	0.0134	0.0045
GW775608X	-0.0005	-0.0067	0.0206	-0.0054	0.0003	-0.0061	0.0163	0.0095
SB-210313	0.0032	0.0027	0.01	-0.0035	-0.0044	-0.0047	0.0449	0.0017
SB-220455	-0.0019	-0.0002	-0.0153	0.0023	-0.0038	-0.0014	0.0578	0.0141
GSK1379712A	-0.0044	-0.0002	-0.0142	0.014	-0.0046	-0.0075	0.0049	0.0049
GSK1379721A	-0.0118	-0.0057	-0.0282	-0.0046	-0.0016	-0.0052	0.0072	0.0132
GSK1379722A	-0.0062	-0.0005	-0.0305	-0.0003	-0.0019	-0.0052	0.027	0.0005
GW589961A	-0.0079	-0.0009	-0.0007	-0.0061	0.0075	0.006	0.0142	-

								0.0064
SB-742864	-0.0124	-0.0096	-0.0146	-0.0043	0.0004	0.0041	0.0512	0.0045
GW459057A	-0.0003	0.0031	-0.021	0.0009	0.0003	0.0098	0.0004	-0.0069
GW819077X	0.004	0.0032	-0.0203	0.0093	-0.0019	0.004	-0.004	-0.0034
SB-347804	0.0111	0.0011	-0.0248	0.0027	0.0027	0.0081	-0.0112	0.0043
SB-409513	0.002	0.0003	-0.0267	-0.0002	-0.0041	0.0026	0.0025	-0.0032
SB-734117	-0.0087	0.0009	-0.0546	0.0041	-0.0004	0.0043	0.0655	-0.0002
GW809885X	-0.0063	-0.0032	-0.0094	-0.0038	-0.001	0.0061	0.0234	0.0026
GW440146A	0.0059	-0.0037	-0.003	-0.0022	0.0038	0.0028	0.005	0.006
GW514786X	0.0077	-0.007	-0.0202	0.0032	0.0035	-0.0035	0.0117	0.0144
SB-376715	-0.0113	-0.0053	-0.0168	0.0038	0.0048	-0.0015	0.0018	0.0029
AH5015X	0.0068	-0.0104	-0.0132	0.0033	0.0034	0.0051	0.0611	-0.0046
GSK2372690A	0.0005	-0.0073	0.0156	0.0009	0.0058	0.0009	-0.0266	0.0026
GW681251X	0.0031	0.0084	0.0093	0.0029	0.0087	-0.0027	0.0371	0.004
GW810578X	0.0024	-0.0076	0.015	-0.002	0.0084	-0.0021	0.0302	0.0043
SB-390766	-0.017	-0.0029	0.0007	-0.0029	0.0057	0.0038	0.0879	0.0054
SB-404290	-0.0001	-0.0049	0.011	-0.0037	0.0049	-1E-04	-0.0266	0
GSK466317A	-0.0172	0.0117	-0.0268	-0.0006	-0.0017	0.0027	0.0559	-0.0068
GSK483724A	0.0091	0.0073	-0.0255	-0.0041	-0.0003	-0.0017	0.0313	0.0084
GW572738X	0.0109	0.0085	-0.0265	-0.0052	-0.0026	0.0024	-0.0033	-0.0068
GW583340C	0.0118	0.0079	-0.0136	0.0101	-0.0039	-0.0031	-0.001	-0.0067
GW812171X	-0.0117	-0.0078	0.0185	0.0009	0.0019	-0.0009	0.0297	-0.0036
SB-431542-A	-0.0106	0.0093	-0.0399	0.0129	0.0041	0.0065	0.0412	0.0049
SB-733416	0.0021	-0.0008	0.0071	0.0041	0.0053	-0.0119	0.001	0.0005
SB-736302	0.0066	0.0082	-0.0259	-0.0065	0.0014	-0.0003	0.001	-0.0027
GW701427A	-0.0107	-0.0023	-0.0127	0.0004	0.0048	-0.0058	-0.0251	0.0134
GW829055X	-0.0025	0.006	0.0063	0.0006	0.0029	-0.0068	0.0389	0.0047
GSK248233B	0.0055	-0.0091	0.0038	0.0011	-0.0005	0.0105	0.0054	0.0096
SB-738561	0.0006	-0.0035	-0.0152	0.0116	0.0041	0.0108	0.0391	0.0041
GW276655X	0.0026	-0.0032	0.0065	-0.0009	0.0002	-0.0053	0.0009	0.0012
GW278681X	-0.0041	-0.0008	-0.0219	-0.0094	-0.0057	0.0096	0.0178	0.0101
GW279320X	-0.0037	-0.0038	-0.0135	0.0038	-0.0032	0.0107	-0.0008	0.007
GW282974X	-0.0051	-0.0037	0.0047	-0.0053	0.0017	0.0034	0.0242	0.0113
GW300657X	-0.0038	-0.0022	0.0059	-0.0036	0.0043	-0.0004	0.0233	0.0154
GW432441X	-0.0168	0.0052	-0.0175	0.0075	0.0006	0.009	-0.1791	0.031

GW513184X	-0.0115	-0.0026	-0.0237	0.0054	-0.0096	-0.0007	-0.0468	0.0077
GW549034X	-0.009	0.0042	0.0051	-0.001	-0.0044	-0.0018	0.0161	0.0045
GW618013X	-0.0157	-0.0017	-0.0569	0.0003	-0.0087	0.0084	0.082	0.02
GW644007X	-0.0859	-0.0011	-0.0583	0.0028	-0.0151	-0.0181	0.0822	-0.0015
GW679410X	-0.013	0.0061	-0.0119	-0.0036	-0.0013	0.0051	0.069	0.0054
GW708336X	-0.008	0.005	-0.0286	-0.002	-0.0039	-0.0113	0.0056	0.0103
GW784684X	-0.0113	0.0012	0.0101	-0.0059	-0.0026	-0.001	0.0133	0.0058
GW804482X	-0.0158	0.0036	0.0064	0.0087	-0.0086	-0.0111	0.1298	0.0238
GW810372X	-0.001	0.0072	0.0177	0.0017	-0.0034	0.0177	0.0082	0.0135
GW827396X	-0.0068	-0.0046	-0.0105	0.008	-0.0032	-0.0032	0.0256	0.0324
SB-264865	-0.0291	0.0076	-0.0003	-0.0027	0.0032	0.0156	125.244	-0.056
SB-747651-A	-0.0017	0.0112	-0.039	0.0044	0.0009	0.0042	0.022	0.012

Chapter 5. Protein-Observed Fluorine NMR is a Complementary Ligand Discovery Method to ^{1}H CPMG Ligand-Observed NMR

Motivation: We compared the newly developed 1D protein-observed fluorine NMR (PrOF NMR) screening method with the well-characterized ligand-observed 1H CPMG NMR screen to evaluate its potential as a ligand discovery tool. We selected the first bromodomain of Brd4 as a model system to benchmark PrOF NMR because of the high ligandability of Brd4 and the need for small molecule inhibitors of related epigenetic regulatory proteins. We compare the two methods' hit sensitivity, triaging ability, experiment speed, material consumption, and the potential for false positives and negatives.

5.1 **Introduction**

Fragment based ligand discovery (FBLD) and fragment based drug design (FBDD) have been rapidly gaining traction in the chemical probe development and drug discovery process for difficult targets such as RNA,[178] transcription factors,[179] and epigenetic regulatory proteins.[118,180] Using molecules with lower molecular weight and lower functional group density than molecules found in traditional high throughput screening libraries, fragment screens seek to sample chemical space more efficiently. With vemurafenib[55] and venetoclax[56] as approved drugs that originated from FBDD and several other drug candidates in Phase III clinical trials that also started as fragments, FBDD is becoming a validated technique for drug discovery. These success stories highlight the impact that new methods can have in the drug discovery process.

Fragment screening facilitates drug/chemical probe development at the earliest stages of the discovery process. However, because of their small size (molecular weight typically less than 300 g/mol), fragments are likely to bind with low affinity,

necessitating sensitive methods to detect protein-ligand interactions during a screen. Among these techniques are thermal shift analysis,[79] surface plasmon resonance,[181] X-ray crystallography,[77] and NMR techniques such as saturation transfer difference (STD),[182] Carr-Purcel-Meiboom-Gill (CPMG),[70] WaterLOGSY,[183] and HSQC NMR[67]. Remarkably, some studies show little overlap in hits detected by different screening methods.[184–186] With the potential for low overlap, it is prudent and common practice to use an orthogonal biophysical method to verify an active.[187] Using protein-observed and ligand-observed NMR experiments allows sensitive follow-up with different assays that utilize similar conditions. The goal of this study is to compare the effectiveness of a new protein-observed ^{19}F NMR (PrOF NMR) screening method with an established ligand-observed ^{1}H CPMG screen to provide insight on the appropriate way to implement PrOF NMR in a ligand discovery setting as a complimentary assay.

Nuclear magnetic resonance spectroscopy has been used, among other techniques, as a way to conduct fragment screens. NMR screens can be conducted by ligand-observed techniques or protein-observed techniques, both of which have inherent advantages over the other. Ligand-observed NMR techniques take advantage of protein-specific properties (relaxation rates, diffusion coefficients, nuclear Overhauser effects) being transferred to bound small molecules. The ligand-observed ^{1}H CPMG method used here detects ligand-binding interactions as a function of a drop in signal intensity due to a transfer of the protein's short transverse relaxation time (T_2) properties to the ligand. ^{19}F CPMG or ^{1}H CPMG are most common due to the high gyromagnetic ratios of ^{19}F and ^{1}H nuclei and thus lead to readily acquired spectra. ^{19}F CPMG is widely used and effective due to the large chemical shift range exhibited by fluorine, the reduced spectral overlap due to its much simplified spectra, and the absence of background signals. However, because we wanted to assess a wide array of compounds without being limited to ligands containing fluorine, we used ^{1}H CPMG as the ligand-observed experiment for our benchmark study.

Protein-observed NMR experiments monitor perturbations to protein resonances. These protein resonances provide an in situ quality control since protein aggregation or precipitation events are readily detected by reduced intensity of the NMR resonances.

Because different regions of the protein give rise to unique NMR resonances, protein-observed experiments also provide information on changes in the chemical environment on the protein surface which can be used to characterize ligand binding sites and binding-induced conformational changes. Typically, these experiments monitor the amide backbone in the case of ^{1}H-^{15}N HSQC NMR or through selective side-chain labeling for ^{1}H-^{13}C HSQC NMR, but this approach can be material-intensive, time-consuming, and expensive due to the need for isotopic labeling.[67,188] The fluorine nucleus, as an alternative, is highly responsive to changes in the chemical environment. ^{19}F is 83% as sensitive as ^{1}H and 100% isotopically abundant (thus inexpensive,[119] e.g., at $52 per gram, 5-fluoroindole is over 10-20 fold cheaper for either ^{13}C or ^{15}N-labeled indole) facilitating detection of ^{19}F at low concentrations (40-50 μM) for small and medium-sized proteins.[85]

Although, alternative protein-observed NMR methods are available, PrOF NMR benefits from both a simplified NMR spectrum as well as being a sensitive reporter of weak binding interactions. As such, ^{19}F NMR has been used in ligand discovery for both ligand-observed and protein-observed experiments.[32,99,127,131,162,189–193] In some cases, protein-observed ^{19}F NMR has proven advantageous when used in conjunction with other protein-observed NMR methods. In studying Equinatoxin II, Anderluh et al. found that ^{1}H-^{13}C HSQC NMR resulted in longer experiment times and lower response than ^{19}F NMR.[194] Harner et al. reported a 7-minute experiment using SOFAST ^{1}H-^{15}N HSQC on a comparable bromodomain at 70 μM protein, compared with the 2-minute PrOF NMR experiments reported here at 50 μM protein.[68] Additionally, Richards et al. reported in studies with Δ-somatostatin that ^{19}F NMR offered improved K_d precision due to higher spectrum resolution and greater chemical environment sensitivity.[195]

As an alternative to 2D-NMR methods, chemical shift perturbation experiments using 1D protein-observed NMR can be valuable using an unlabeled protein.[196] For the protein under study here, we found that chemical shift response from the ^{1}H of the N-H of tryptophan[196] to be 6-20x less responsive than the equivalent PrOF NMR shift with acetaminophen and two compounds uncovered in this screen which will be described below (**Figure 5.1**,

Figure 5.2). We have previously applied PrOF NMR for fragment-based screening of over 500 small molecules and characterized the bromodomains Brd4, BrdT, and BPTF using fluorine-labeled aromatic amino acids 3-fluorotyrosine and 5-fluorotryptophan (5FW).[131] Using this method, we reported the first molecule selective for BPTF over Brd4 using a simultaneous dual protein screening approach.[32] However, due to a reduced surface coverage from labeling only a few side chains, the potential for false negatives was a concern which we sought to evaluate in this study.

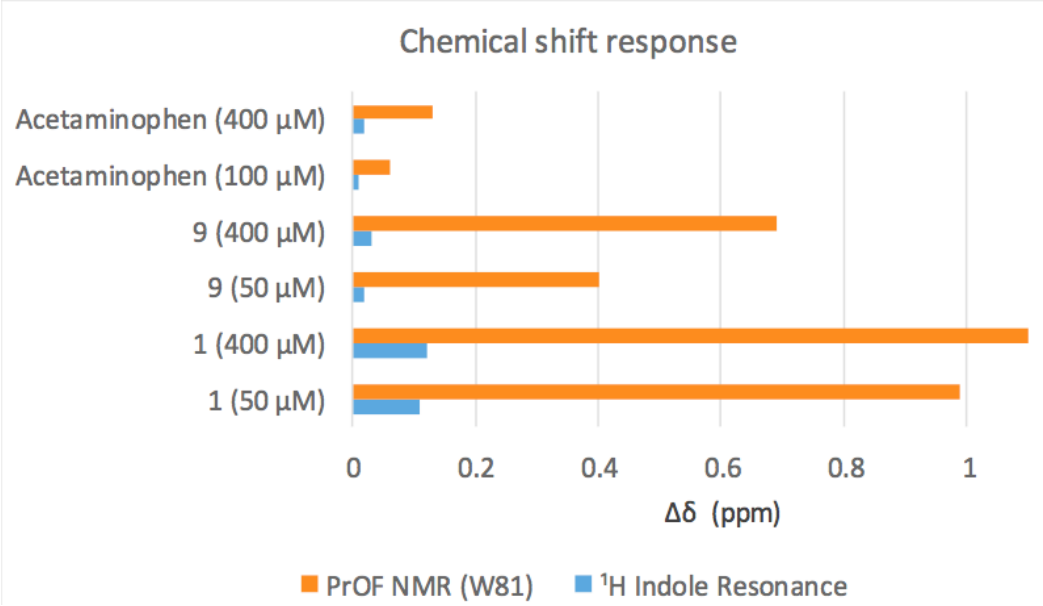


Figure 5.1 Comparison of the perturbations induced upon the W81 PrOF NMR resonance and the assumed W81 indole resonance with various binders.

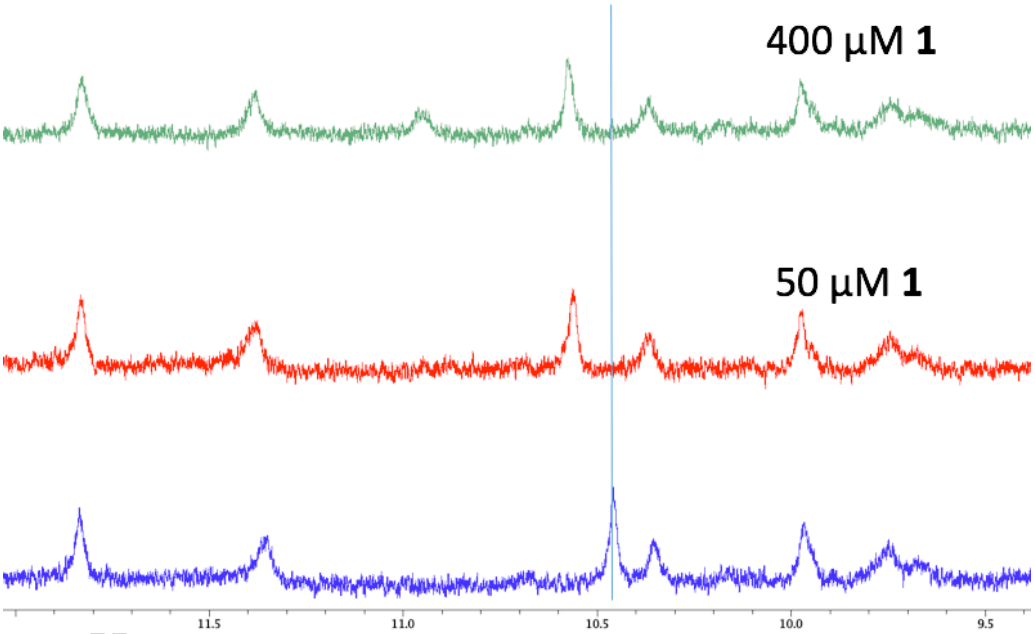


Figure 5.2 The area of the ^{1}H NMR spectrum corresponding to the N-H tryptophan resonances.

The first bromodomain of Brd4 represents an important model protein because of its high ligandability and the need for small molecule inhibitors of other bromodomains due to a growing understanding of their importance in regulation of disease. Bromodomains are epigenetic regulatory proteins, recognizing acetylated lysine residues present on histone tails. Of the 61 human bromodomains, Brd4 in particular has been shown to be involved in cancer,[13] inflammation,[17] and heart disease.[197] However, many other bromodomains lack specific chemical probes for evaluating the pharmacological inhibition effects on their biology. In this study we sought to compare PrOF NMR with the well-established ^{1}H CPMG NMR method to evaluate the effectiveness of this technique as a screening method, as well as the complementarity between ligand-observed and protein-observed NMR techniques. ^{1}H CPMG was chosen because it is frequently used as a screening technique,[198] and we found it to be robust for this protein system. We sought to compare assay speed, hit overlap, and the potential for false negatives and false positives with PrOF NMR. To this end, we screened 930 fragments against Brd4 using each method and discuss our findings regarding the potential advantages and limitations of the two methods.

5.2 **Results**

5.2.1 PrOF NMR Results

In the PrOF NMR screen, 930 fragments were tested in mixtures of five; induced chemical shift perturbations and resonance broadening were used to rank-order hits. The PrOF NMR screen was performed in the presence of 2% DMSO, and because DMSO is a known bromodomain ligand that affects fluorine chemical shifts,[131] all PrOF NMR screening experiments are compared with a 2% DMSO control sample. For assessing the speed and reproducibility of the experiment the screen was conducted twice with different numbers of scans leading to a thirteen- and two-minute NMR experiment time. The comparison of resonance perturbations between the long and short PrOF NMR experiments indicates that two minutes is a sufficient experiment time for acquiring screening data under these conditions (Figure 5.3, Figure 5.4). Chemical shift

perturbations from fragment binding between the thirteen- and two-minute experiments were very similar, with modest concordance for the linewidth broadening between the two screening times (Figure 5.3). In shorter experiments resulting in lower signal-to-noise ratios (S/N = 7:1), the ability to accurately measure linewidth and integration values of resonances is diminished, while the induced chemical shift is still readily interpreted. When chemical shift is used as the primary indicator of a binding interaction, as is the case for this study, S/N $\geq$ 7:1 are sufficient.

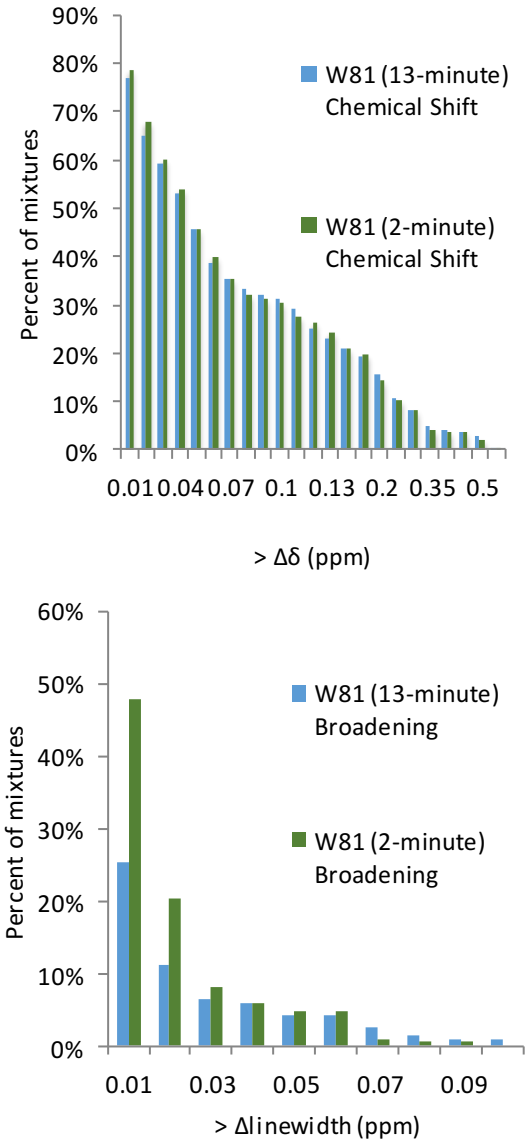


Figure 5.3 A comparison of the data in the 13-minute PrOF NMR experiment and the 2-minute PrOF NMR experiment.
Comparing the percentage of mixtures perturbing W81 greater than a specific chemical shift (top) or linewidth (bottom). The average difference in perturbation of chemical shift of W81 between the 13-minute experiment and the 2-minute experiment was 0.0049 ppm, with a standard deviation in chemical shift perturbation average of 0.0055 ppm. Chemical shift perturbation is very robust even with low signal to noise, while linewidth is comparable after the widths are greater than 0.03 ppm. The average difference in linewidth between the 13-minute and 2-minute experiment is 0.0116 ppm, with a standard deviation of 0.0111 ppm.

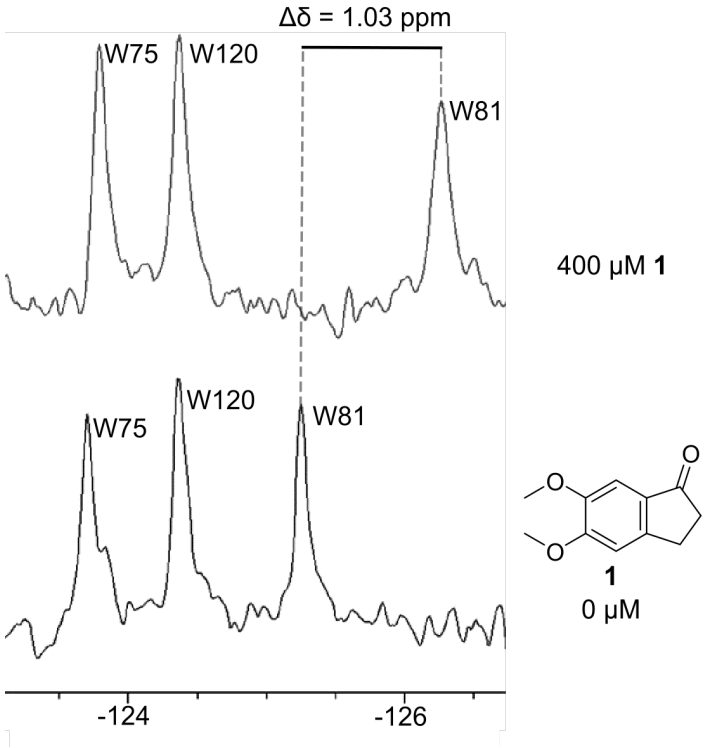


Figure 5.4 A 2-minute PrOF NMR experiment (S/N: 7:1).
The perturbation due to the binding interaction of 1 is clearly visible in this spectra as it was with the 13-minute PrOF NMR experiments.

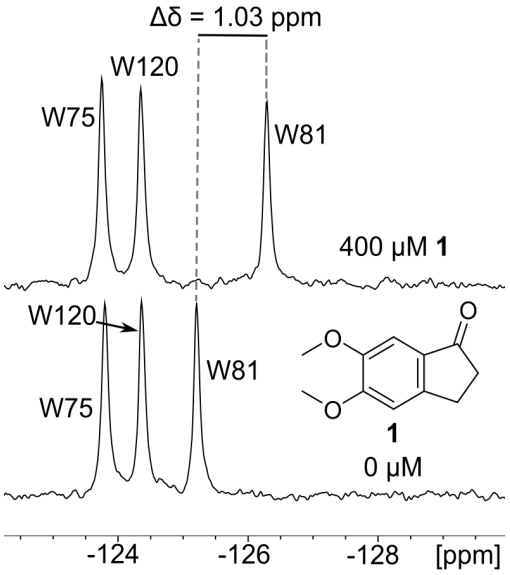


Figure 5.5 PrOF NMR spectrum of 5FW-Brd4 and 1.
In the presence of 1, the resonance for W81 is significantly shifted upfield consistent with binding near the acetylated lysine interaction site of Brd4.

Perturbation of chemical shift is used as the initial indicator of a binding interaction because we expect most of the active fragments to be relatively weak ligands. Weak ligand interactions (K_d > 100 μM) typically fall into the fast exchange regime on the NMR time scale, with highly reproducible chemical shifts. Broadening of a resonance is also an indication of a binding event, potentially in the intermediate exchange regime, and was used as a secondary measure. The first bromodomain of Brd4 has three tryptophan residues: W81 (11 Å from N140, a conserved residue in bromodomains) is at the binding site, W75 is 18 Å from N140, and W120 is on the opposite end of the protein, 36 Å from N140 (Figure 5.6).[199–201] Because of the distance of W120 from the binding site, it is assumed that W120 is not perturbed by ligands during the screen. Thus, when the ^{19}F resonances corresponding to W81 or W75 are perturbed more than the greatest observed perturbation of W120, we conclude that a binding interaction is taking place. To capture as many potential binding interactions as possible for the analysis, we take a perturbation of 0.03 ppm for either W75 or W81 to be the cut-off value indicating the lowest limit for a binding interaction because this was the highest perturbation observed for W120 (Figure 5.7). In this screen 59.1% of mixtures yielded a perturbation of W81

greater than 0.03 ppm. As an example of seeking a more stringent hit-rate, the percentage of detected hits drops to 21% with a cut-off perturbation of 0.14 ppm as illustrated in Figure 5.7.

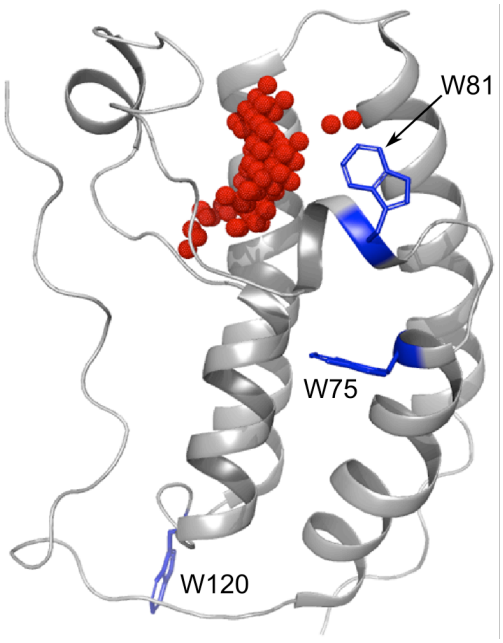


Figure 5.6 First bromodomain of Brd4 with all three tryptophan residues displayed and labeled by residue number.
Spheres indicate the acetylated lysine binding site (generated by SiteMap). PDB ID: 3UVW.

Because PrOF NMR is a protein-observed technique, it cannot be determined which compound is binding to the protein in the mixture of five without further deconvolution in which each compound is tested individually. Because of the high hit-rate with Brd4, 20% of the mixtures were deconvoluted in order to reduce protein consumption while assessing a large representative sample of the hits in the screen. 10% of the mixtures in the screen were selected to be deconvoluted due to broadening of the W81 resonance in their corresponding PrOF NMR spectra, and an additional 10% were chosen to span a range of PrOF NMR chemical shift perturbations (0.03 ppm – 0.31 ppm) so as to capture both low- and high-affinity interactions, as well as assess five mixtures causing global loss of protein signals, potentially indicating the presence of aggregators.

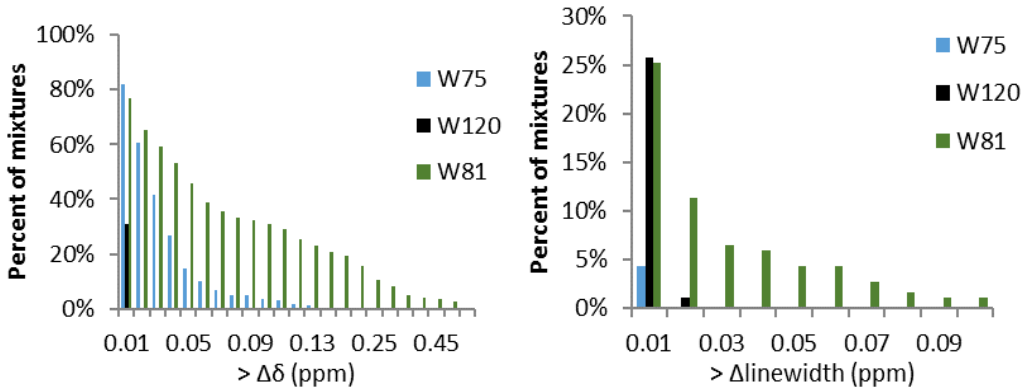


Figure 5.7 Analysis of the chemical shifts of the three resonances of 5FW-Brd4, comparing the percentage of mixtures perturbing the chemical shift of a particular resonance.
Left) With a chemical shift change of 0.01 ppm or higher W120 is perturbed in over 30% of mixtures, but with a 0.03 ppm cutoff W120 is not perturbed. Thus, 0.03 ppm perturbations of either W75 or W81, which are closer to the binding site, are used as the hit cutoff for PrOF NMR. Right) A 0.03 ppm linewidth change for the resonances for W75 and W81 were also considered a hit cutoff.

A valuable strength of protein-observed NMR methods is their ability to quantify ligand affinities by titration. While chemical shift perturbation is a standard method for detecting ligand interactions, the magnitude of the perturbation also varies due to local chemical environment so a larger shift does not necessarily correspond to a stronger interaction. To quickly rank-order hits from the PrOF NMR screen, we took 22 ligands (11 of which are publically available) and titrated each ligand into a single NMR tube containing 5-fluorotryptophan-labeled Brd4 (5FW-Brd4) (Figure 5.8, Figure 5.9). Using a single tube reduces protein consumption, but results in error propagation from pipetting, does not control for the effect of DMSO (a known bromodomain ligand) or protein dilution. Additionally, the concentration of the ligands was not verified by quantitative NMR. We term these values $\Delta\delta_{50}$, whereas we ascribe dissociation constants to titrations that use distinct NMR samples for each titration point to keep the DMSO concentration consistent, from fresh ligand stocks with a concentration verified by quantitative NMR. Using the more rigorous titration method, we obtained dissociation constants for five actives found in the screen.

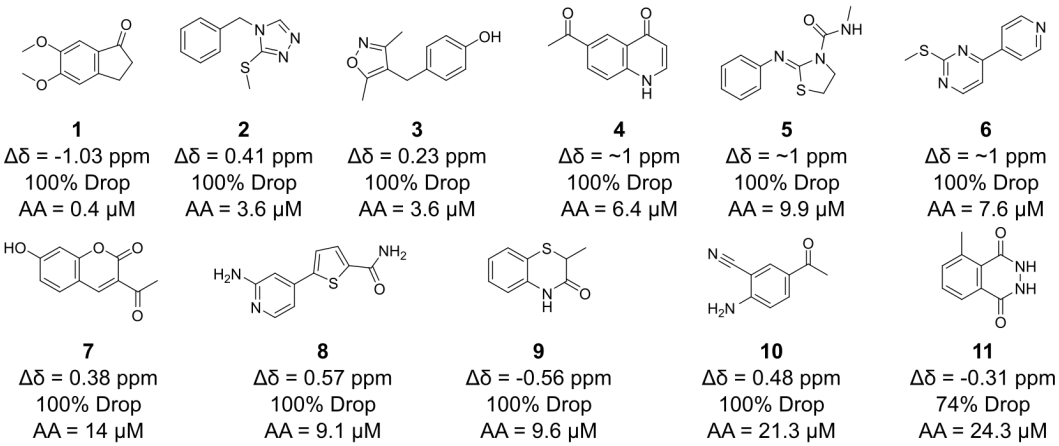


Figure 5.8 Selected molecules that were hits in the PrOF NMR screen.
The Δδ50 was measured by PrOF NMR titration. Percent reduction in signal in the ^{1}H CPMG experiment is indicated by percent drop, and the change in chemical shift upon the addition of ligand at 400 μM is indicated with Δδ.

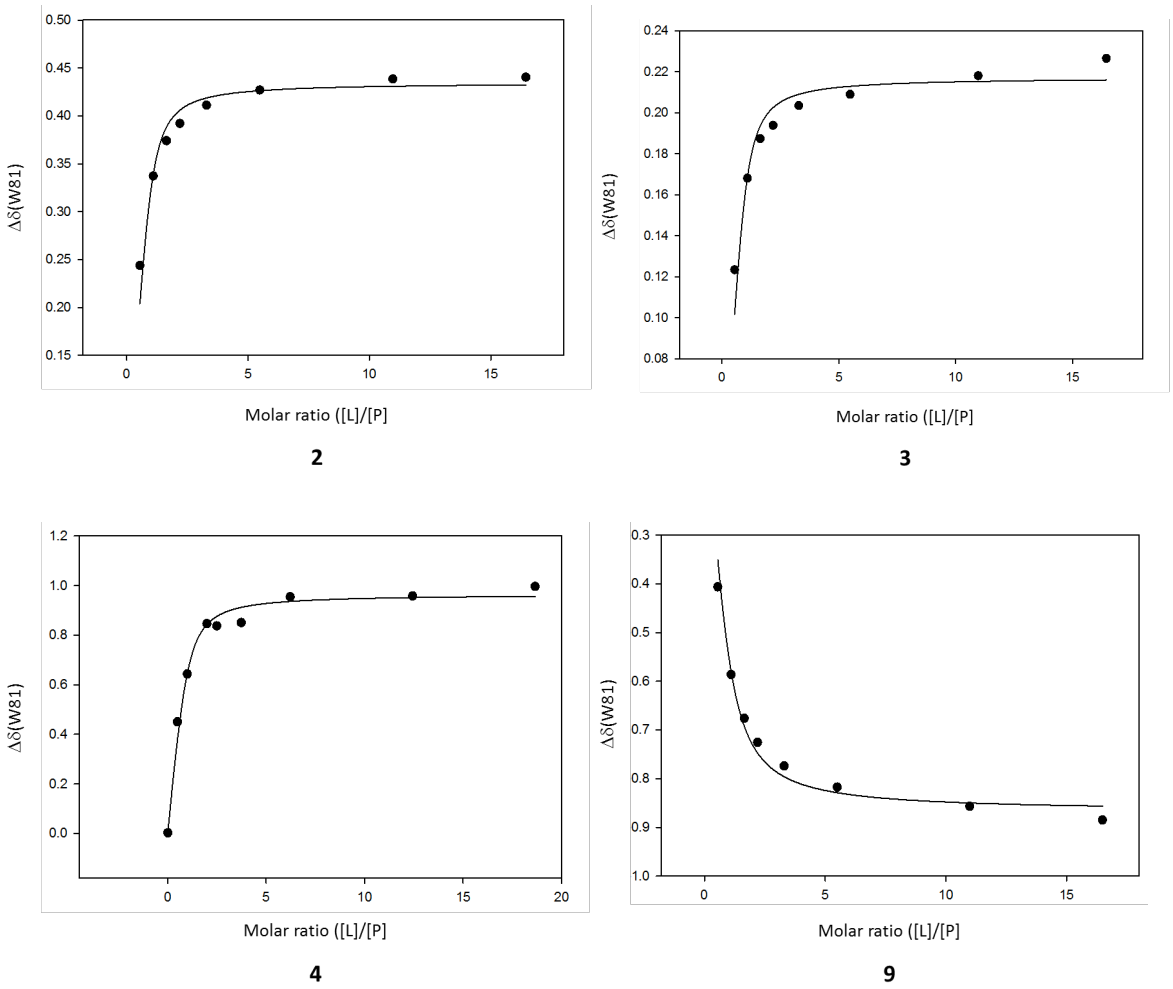


Figure 5.9 Fitting graphs for selected compounds for Δδ50 determination using PrOF NMR titrations.

The data was fitted to the following equation using SigmaPlot (Systat Software Inc., San Jose, California, USA):

$$\Delta\delta_{obs} = \Delta\delta_{max} - \frac{1}{2}\Delta\delta_{max}\left\{\sqrt{\left(\frac{[L]_{tot}}{[E]_0} + \frac{k_D}{[E]_0} - 1\right)^2 + 4\frac{k_D}{[E]_0}} - \left(\frac{[L]_{tot}}{[E]_0} + \frac{k_D}{[E]_0} - 1\right)\right\}$$

5.2.2 ^1H CPMG Results

In the ^{1}H CPMG screen, the same 930 compounds were screened in the same mixtures of five. Some of the more potent compounds showed a 100% signal intensity

reduction (Figure 5.10). Due to the signal-to-noise from using 100 μM ligand, for a reliable measurement we assume the limit of detection for change in resonance signal intensity is 10%. A hit was thus defined by a 20% reduction in ligand signal intensity upon addition of the protein to allow for a competition experiment. For the ^{1}H CPMG screen, a competition step with a known potent ligand can significantly reduce false positive rates and at the same time provide valuable information on the inferred binding site for the hit. For this reason, we performed the competition step using (+)-JQ1 as the competitor, and a competitive hit required a 50% recovery of the signal intensity upon the addition of (+)-JQ1 (i.e., a 10% reduction of the original signal intensity in the limiting case).[13] Of the 930 compounds screened, 230 of the compounds were identified as hits (24.7%) with Brd4 and 190 (20.4%) compounds were competitive hits against (+)-JQ1.

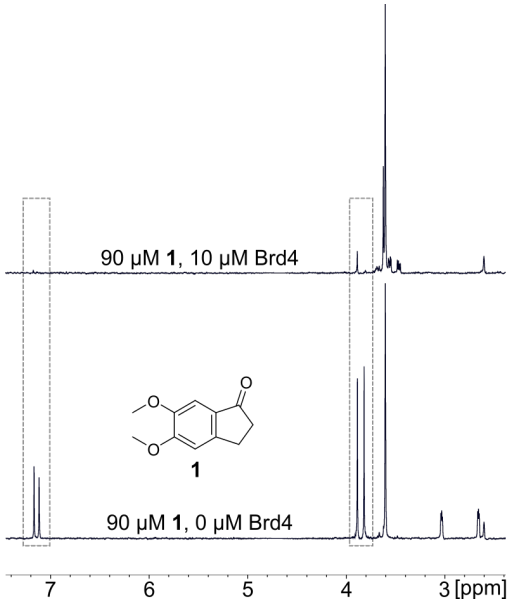


Figure 5.10 A ^{1}H CPMG experiment of 1 and Brd4.
Upon the addition of Brd4, the ^{1}H resonances of fragment 1 signal intensity are reduced up to 100% in the top spectrum. Examples of reduced resonances are indicated by the dashed lines.

5.2.3 Comparison of Results

Due to the high hit-rate in the PrOF NMR screen that is in part ascribed to the conservative chemical shift cut-off, not every mixture which contained a hit was deconvoluted. Because the same mixtures were used for both the PrOF NMR and ^{1}H

CPMG screens, we offer a comparison of the data as mixtures as well as the 190 compounds individually tested by PrOF NMR. When comparing the fragment mixture results between ^{1}H CPMG and PrOF NMR, the two assays are 77% similar when using a competition ^{1}H CPMG experiment (74% when results from the competition step are omitted). From the data for the 190 individual compounds deconvoluted by PrOF NMR (38 mixtures), there was 85% similarity between the two assays when using a competition ^{1}H CPMG experiment (80% when omitting the competition experiments) (Figure 5.11). After deconvolution, nearly every mixture that was a hit by PrOF NMR contained at least one individual compound that bound, with 17 of the 33 mixtures containing more than one hit. The only mixture that was a hit by PrOF NMR that did not identify a small molecule binder after deconvolution of the mixture had an initial W81 chemical shift perturbation of 0.033 ppm, barely above the cut-off of what is considered a hit in this screen.

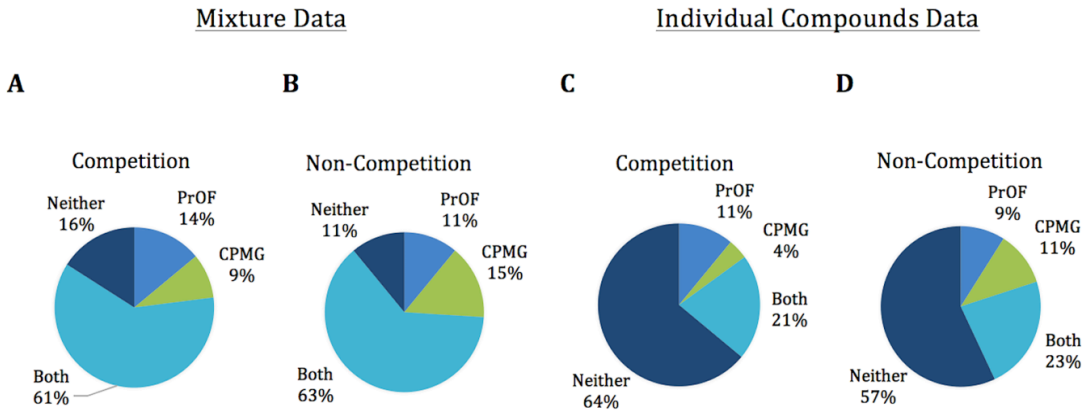


Figure 5.11 Comparison of hit data between PrOF NMR and 1H CPMG from the fragment mixture data (A,B) as well as the individual deconvoluted compounds (C,D).

While both assays have a large overlap in detected hits, we analyzed disagreements between the two screens to identify sources of potential false positives and negatives. For weak binding ligands, there were 10 discrepancies exhibiting only minor perturbation in one assay (i.e. chemical shift changes of 0.03 ppm - 0.05 ppm in PrOF NMR; or signal intensity drop of 20 - 30% in ^{1}H CPMG) and no perturbation in the other. In this case, the discrepancies can be explained by the weakest binders being missed by

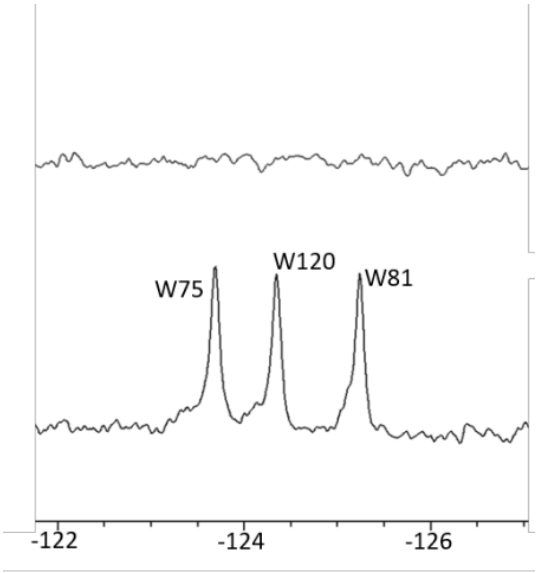
one of the two assays as the sensitivity limits of these assays are approached. In the case of stronger binding compounds, there were 13 cases where there was a moderate to strong perturbation in one assay that was a non-binder by the other assay (Table 5.1, 0.05 - 0.2 ppm, PrOF NMR; 30 – 100% signal drop, ^{1}H CPMG). In the case of ^{1}H CPMG, seven ligands were detected as competitive hits but not detected by PrOF NMR, though only two of those exhibit a signal drop of greater than 50%. These two compounds may bind deeply in the binding site, or at an allosterically-coupled location, such that PrOF NMR may not have a fluorine probe close enough to detect the binding event, and could thus be false negatives in the context of PrOF NMR. Additionally, there were two PrOF NMR hits that were non-competitive hits by ^{1}H CPMG, indicating that the ligands bound but were not competitive at the acetylated lysine binding site. These molecules may be consistent with either non-specific binders or binding to a second site on the protein. By PrOF NMR both of these hits perturbed W81 as well as W75, but because there are more than 10 fragments that perturbed both W81 and W75 in the PrOF NMR screen that were competitive hits by ^{1}H CPMG it is unlikely that W75 could be used to ascertain differential binding. From this analysis we conclude that although a small fraction of false positives and negatives were identified between the two methods, the majority of compounds that were missed by one of the two screening methods correspond to low- to moderate-affinity fragments. Importantly, only two fragments (1% of competitive ^{1}H CPMG hits, 0.2% of total molecules in the screen) that scored as strong and competitive hits by ^{1}H CPMG are potential false negatives by PrOF NMR (Table 5.1).

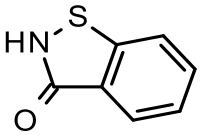
Table 5.1 A list of discrepancies between hits identified by PrOF NMR and 1H CPMG

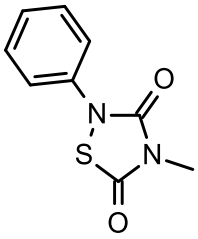
Compound No.	*Δδ W81 (ppm)*	*ΔlinewidthW81 (ppm)*	*Δδ W75 (ppm)*	*% Drop (CPMG)*	*% recovery (CPMG)*
13	0.264	0.0738	-0.0058	<10%	N/A
14	-0.2107	0.0123	-0.0021	14.4%	50.6%
15	0.0878	0.0123	-0.0772	<10%	N/A
16	-0.0763	0.0246	-0.0168	<10%	N/A
17	-0.062	0.0123	-0.0176	<10%	N/A
18	-0.0443	0	-0.0007	<10%	N/A
19	0.0389	0.0123	-0.0039	<10%	N/A
20	0.0313	0	0.0001	10.4%	258.5%
21	0.0271	0	-0.0187	100.0%	29.2%
22	0.0262	0.0123	-0.0291	90.2%	119.4%
23	0.02	0	0.0014	28.5%	-34.0%
24	0.0198	0.0123	-0.0191	41.6%	82.1%
25	0.0171	0	-0.0181	22.3%	35.5%
26	0.0168	0.0123	-0.0236	61.7%	26.6%
27	0.0159	0.0123	-0.0042	23.4%	102.9%
28	-0.0113	0	-0.0193	43.6%	160.2%
29	-0.008	0	-0.0167	20.2%	33.3%
30	0.0074	0	-0.0199	36.8%	130.0%
31	0.007	0	-0.0242	100.0%	40.0%
32	-0.0034	0	-0.0165	65.8%	119.7%
33	0.0028	0	-0.0088	41.2%	14.8%
34	-0.0014	0	-0.0111	69.0%	5.3%
35	0.0013	0	-0.0068	56.1%	-4.5%
36	-0.0011	0.0123	-0.0054	20.4%	-20.6%
37	-0.0009	0	-0.0144	43.4%	10.9%
38	0.0005	0.0123	-0.0275	28.6%	32.6%
39	-0.0005	0	0.0015	30.9%	-77.5%
40	0.0001	0	-0.0247	28.7%	134.8%

Small molecule-induced protein aggregation is a common artifact in many screens[202] and can be a significant cause of false positives. Therefore, we also evaluated this effect in our study. In our screen, there were five mixtures where the resonances in the PrOF NMR spectrum were no longer visible, and these five mixtures were

deconvoluted to reveal one compound from each mixture responsible for the signal suppression, potentially acting as protein aggregators (Figure 5.12). These molecules were not detected as ^{1}H CPMG hits because of insufficient signal-to-noise ratios to analyze the data. However, if a competitive ligand-observed experiment were conducted detecting the displacement of a "spy molecule",[70] these aggregators could result in false positives. Adding detergents to the screening buffer can be one way to reduce such effects. Due to the absence of fluorine in many commercial detergents, they are compatible with ^{19}F-based ligand-observed and protein-observed experiments, although they were not used in our current study.







Aggregator 1 **Aggregator 2**

Figure 5.12 Potential protein aggregators found by PrOF NMR.
In the bottom spectrum is the protein in the absence of the small molecule mixture. In the top spectrum all resonances for the protein disappear indicating that the small molecule mixture induces the formation of a large protein aggregate, or the protein has crashed out of solution. Two such compounds that resulted in this spectrum are shown on the right, and are similar to known PAINs compound categories (Baell, J. & Walters, M., Nature, 2014, 481-483.)

Prioritizing compounds based on affinity is important for carrying compounds forward in the lead discovery process. This is particularly true in cases of high hit-rates in a fragment screen, such as described here. From our screening data and follow-up affinity

determination, we found that both methods are fairly effective at triaging hits from the screening data. We used PrOF NMR titrations to determine the $\Delta\delta_{50}$ for 22 compounds. We then compared several parameters with $\Delta\delta_{50}$, and found that chemical shift perturbation in PrOF NMR and signal intensity reduction in ^{1}H CPMG correlate but only modestly with $\Delta\delta_{50}$ measured by PrOF NMR (Figure 5.13). Despite broadening being related to chemical exchange kinetics in PrOF NMR, there is a poor correlation between broadening and $\Delta\delta_{50}$ with Brd4. Interpretation of broadening in terms of relative affinity is complicated by the differences in chemical shifts between bound and unbound state which varies between ligands, as well as additional protein dynamics such as restricted mobility of side-chains due to binding interactions. Additionally, with higher affinity fragments there are cases where the W81 resonance is sharpening out of the baseline, approaching a fully bound state. This can result in situations where a broader resonance does not necessarily indicate a tighter binder. Thus, resonance broadening was not an effective method for ranking compounds in this PrOF NMR screen. However, when screening proteins with a lower ligandability, broadening of resonances in PrOF NMR could potentially be a useful parameter for rank-ordering compounds as well.

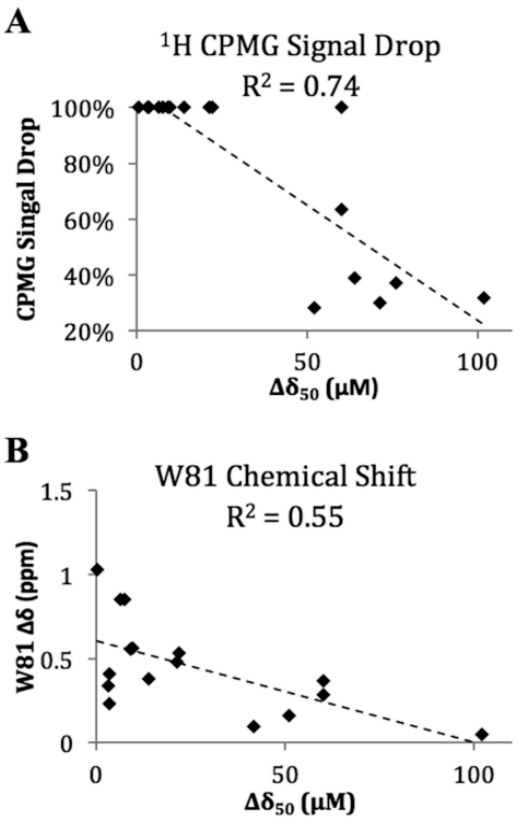


Figure 5.13 A comparison of 20 different hits ranked by Δδ50 by PrOF NMR and different assay data.
A). Signal reduction in 1H CPMG in panel A or PrOF NMR chemical shift perturbation in panel B). A correlation is noticeable between Δδ50 and 1H CPMG signal drop, as well as Δδ50 and W81 chemical shift perturbation. While one could expect broadening of W81 to also correlate well, with Brd4 this is not the case. Using a protein with a lower ligandability could result in better rank-ordering with broadening, but with Brd4 many of the tighter binders are already coalescing at 400 μM ligand.

As further support for the successful identification of true bromodomain binders, several of the compounds identified in this screen have similar pharmacophores to known bromodomain inhibitors: isoxazoles,[31] triazoles,[13,17] amidines,[32] and ureas[32,33] have all been previously reported.[180] Additionally, we identified several ligands containing ketones, potentially acting as acetylated lysine mimetics. We thus conclude, that the conditions for screening fragments against Brd4 by both PrOF NMR and ^{1}H CPMG were effective for uncovering new ligands.

5.2.4 Dissociation Constant Determination

Protein-observed NMR techniques, including PrOF NMR, have been utilized to quantify dissociation constants via ligand titration at a fixed protein concentration. These derived dissociation constants rely on chemical shift perturbations from a titration experiment. While titrations can be performed with ^{1}H CPMG ligand-observed experiments, these experiments require rigorous controls to present off-setting effects from convoluting the measured dissociation constants, and can also be time consuming.[190,203] These titrations are easily accomplished with PrOF NMR and can be used to prioritize compounds based on affinity and/or ligand efficiency. The dissociation constants obtained by this method were used to establish ligand affinities described above for comparing with resonance perturbations. Original affinity estimates were obtained by titrating a small molecule from the library stock directly into the protein solution, which can introduce errors as described above. To verify the rank ordering and more accurately measure the dissociation constants, the affinities of the top five compounds (molecules **1**, **2**, **3**, **4**, and **12**) were determined by PrOF NMR (Figure 5.14, Figure 5.15, Figure 5.16, Figure 5.17, Figure 5.18). The affinities shown in Table 5.2 ranged from 6.2 to 256 μM and were consistent with the rank ordering by $\Delta\delta_{50}$. Using the same samples a 1D protein-observed experiment was conducted monitoring the protein indole N-H resonances, however due to the lack of significant chemical shift upon ligand binding only two of five K_d values could be obtained but were in agreement with PrOF NMR values (32 vs 39 μM for **1**, and 148 vs 142 μM for **2**).

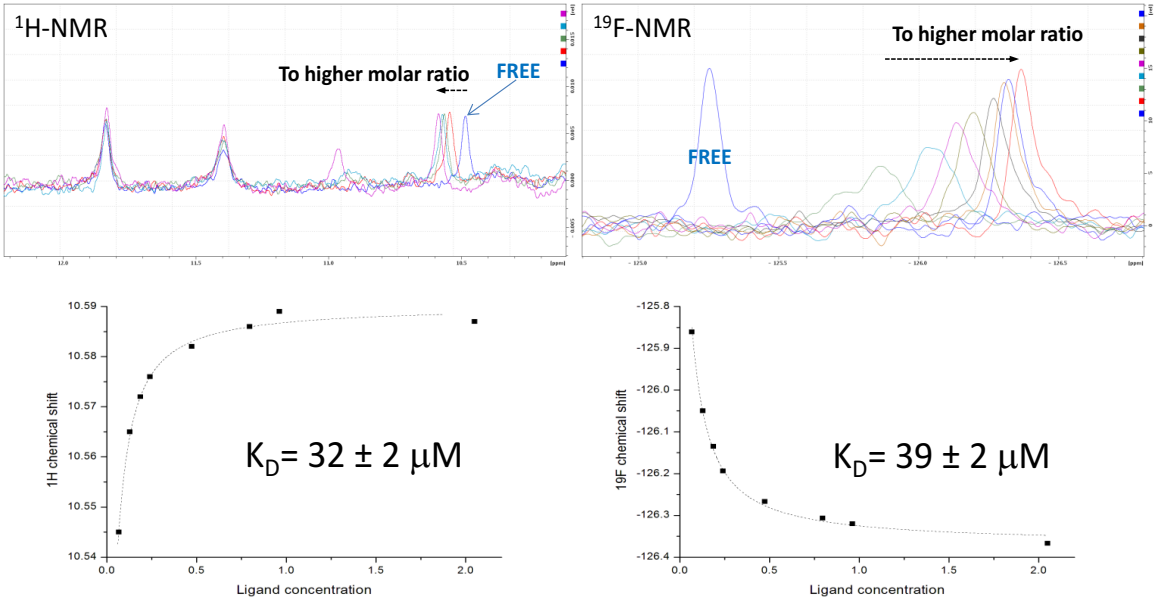


Figure 5.14 A titration of 5FW-Brd4 (40 µM active protein) with 1 monitored by both PrOF NMR and 1H NMR of the N-H tryptophan resonance (2% DMSO).

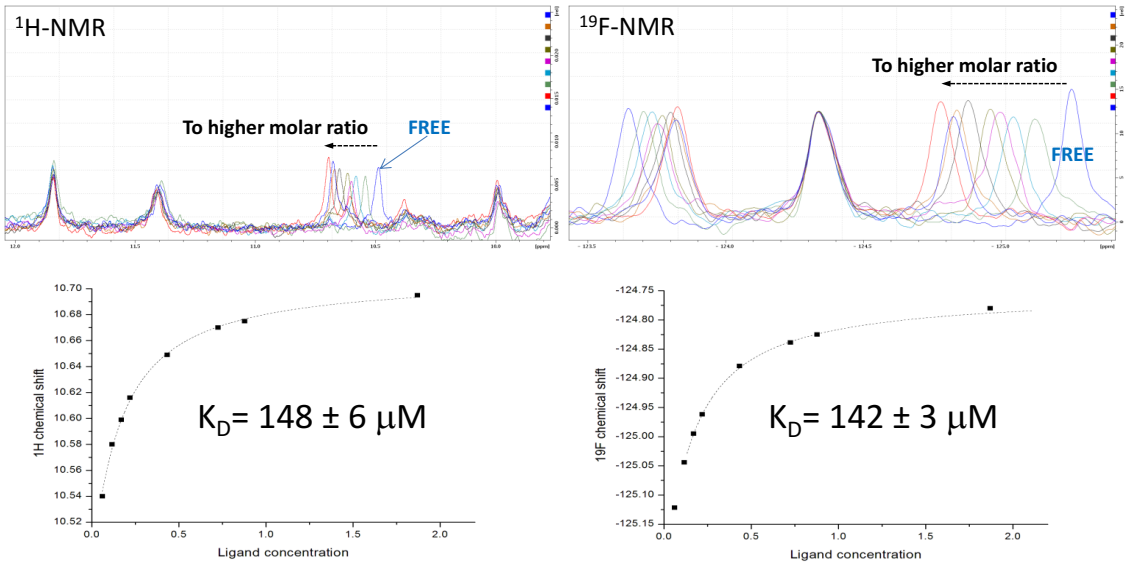
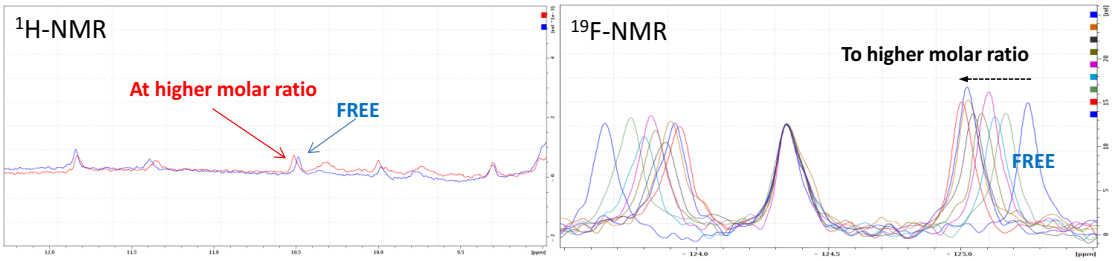


Figure 5.15 A titration of 5FW-Brd4 (40 µM active protein) with 2 monitored by both PrOF NMR and 1H NMR of the N-H tryptophan resonance (2% DMSO).



Monitoring ^{1}H-NMR of the N-H tryptophan resonances no significant chemical shift changes were recorded even at the highest molar ratio.

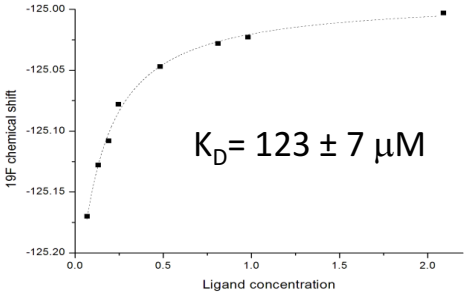
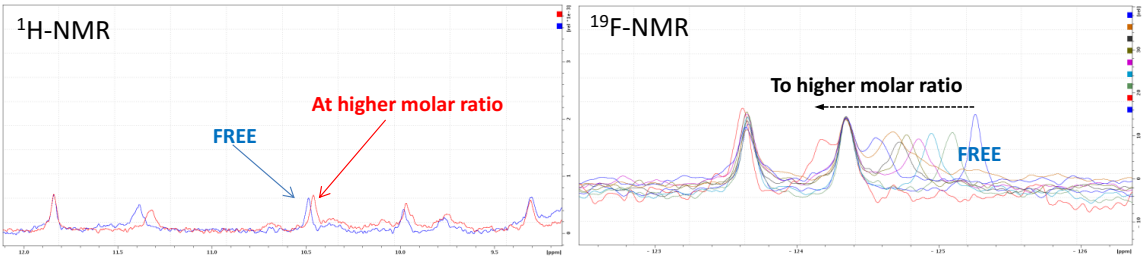


Figure 5.16 A titration of 5FW-Brd4 (40 µM active protein) with 3 monitored by both PrOF NMR and ^{1}H NMR of the N-H tryptophan resonance (2% DMSO).



Monitoring ^{1}H-NMR of the N-H tryptophan resonances no significant chemical shift changes were recorded even at the highest molar ratio.

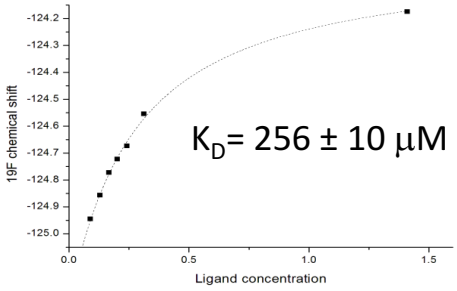


Figure 5.17 A titration of 5FW-Brd4 (40 µM active protein) with 4 monitored by both PrOF NMR and ^{1}H NMR of the N-H tryptophan resonance (2% DMSO).

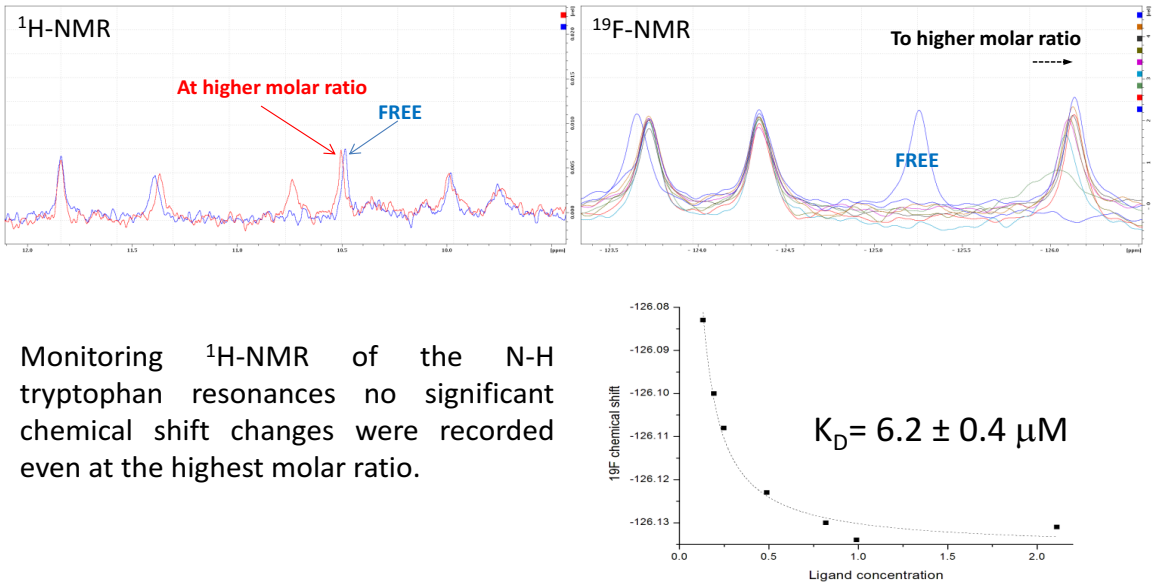


Figure 5.18 A titration of 5FW-Brd4 (40 µM active protein) with 12 **monitored by both PrOF NMR and ^{1}H NMR of the N-H tryptophan resonance (2% DMSO).**

To verify the rank-ordering of NMR hits by additional methods, we measured the top eight fragments by $\Delta\delta_{50}$ with thermal shift analysis (TSA) and isothermal titration calorimetry (ITC) with unlabeled Brd4. All eight showed an increased melting temperature of Brd4 by 0.5 °C or greater by TSA (Figure 5.19), and five led to quantifiable dissociation constants by ITC (Figure 5.20, Figure 5.21, Figure 5.22). Two of the molecules for which we were unable to quantify the affinities showed an incomplete binding isotherm trend, consistent with weak binding. One of the fragments that showed incomplete binding was **9**. Seeking to verify binding measurements by another method, we attempted to measure the K_i of **9** and **1** by fluorescence anisotropy. While **1** resulted in a K_i of 10.8 µM (comparable to ITC measurements), the fluorescence anisotropy experiments with **9** still led to an incomplete binding curve supporting a weak binding affinity (Figure 5.23). In summary of these results, the rank-ordering of the hits by ITC and TSA matches the rank-ordering of hits by PrOF NMR $\Delta\delta_{50}$ determined by the initial titrations.

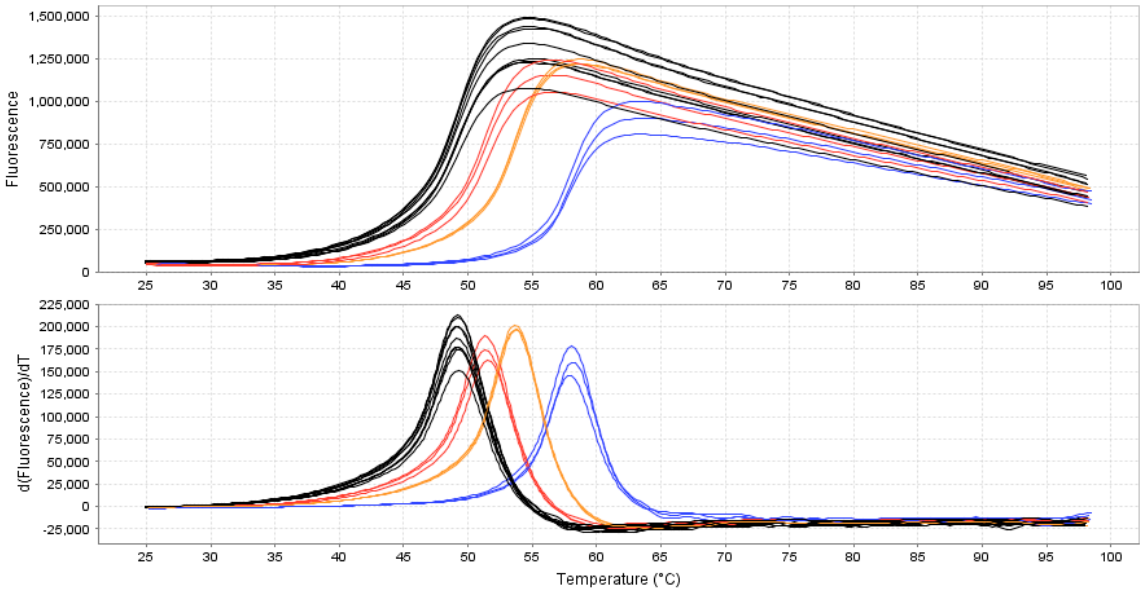


Figure 5.19 Differential scanning fluorimetry of apo-form of Brd4 and three different binders.

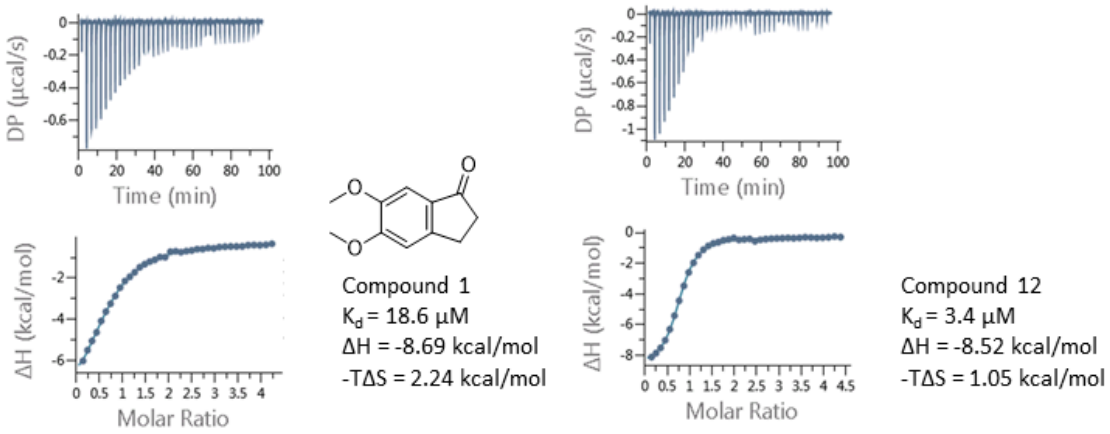


Figure 5.20 ITC results of the highest affinity fragments which were identified by PrOF NMR experiments.
These ITC experiments were performed with Brd4 wild-type in the presence of DMSO (0.5% DMSO).

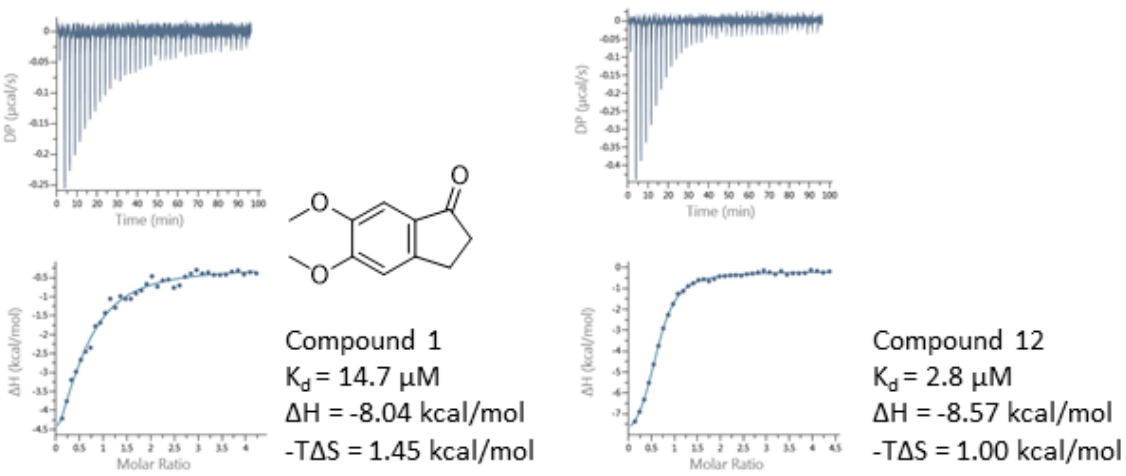
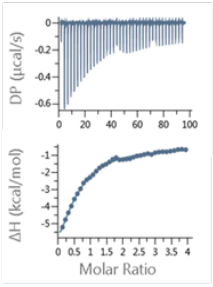
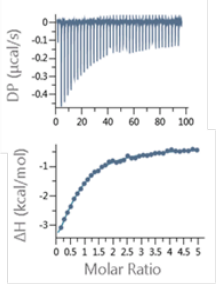


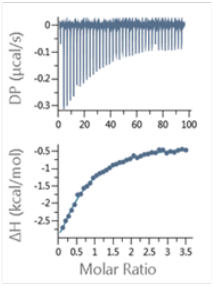
Figure 5.21 ITC results of the highest affinity fragments which were identified by PrOF NMR experiments.
These ITC experiments were performed with 5FW-Brd4 in the presence of DMSO (0.5% DMSO).



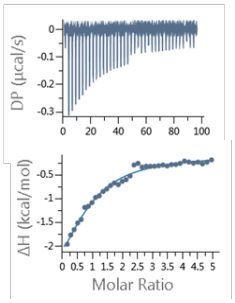
2
Kd = 37.4 µM



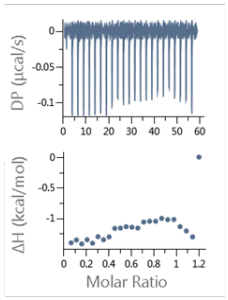
3
Kd = 42.1 µM



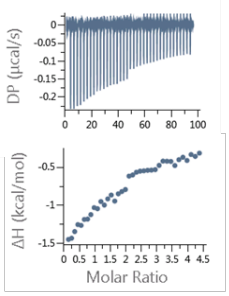
4
Kd = 54.6 µM



6
Unclear by ITC



8
Negative by ITC



9
Very low heat evolved
Unclear by ITC

Figure 5.22 ITC results of the other fragments which were identified by PrOF NMR experiments.
These ITC experiments were performed with Brd4 in the presence of DMSO (0.5% DMSO).

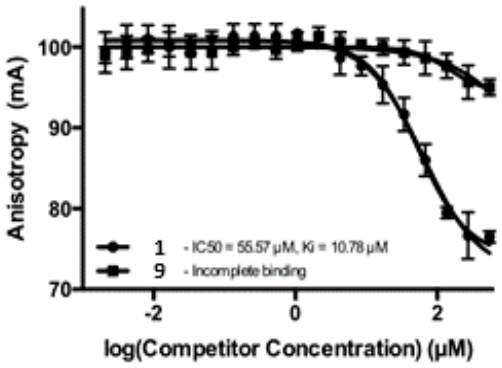


Figure 5.23 Fluorescence anisotropy competitive inhibition measurement of 1 and 9, displacing a fluorescently labeled ligand, BI-BODIPY.
The isotherm for 9 shows incomplete binding, consistent with a low affinity ligand.

Table 5.2 Comparison of biophysical affinities obtained by different methods

Compound	ITC - K_d (µM) (Ligand Efficiency)	5FW-Brd4 ITC (µM)	PrOF NMR K_d (µM)	^{1}H NMR K_d (µM)	$\Delta\delta_{50}$ PrOF NMR (µM)	Thermal Shift (Δ °C)
12	3.4 (LE: 0.53)	2.8	6.2	[b]	<1	4.53
1	18.6 (LE: 0.46)	14.7	39	32	<1	2.23
2	37.4 (LE: 0.42)	[a]	142	148	3.6	0.77
3	42.1 (LE: 0.40)	[a]	132	[b]	3.6	1.1
4	54.6 (LE: 0.41)	[a]	256	[b]	6.4	0.72
5	[a]	[a]	[a]	[a]	9.9	[a]
6	[b]	[a]	[a]	[a]	7.6	1.4
7	[a]	[a]	[a]	[a]	14	[a]
8	[b]	[a]	[a]	[a]	9.1	1.3
9	[b]	[a]	[a]	[a]	9.6	0.57
10	[a]	[a]	[a]	[a]	21.3	[a]
11	[a]	[a]	[a]	[a]	24.3	[a]

[a] Values not determined

[b] Results inconclusive

In comparison of K_d values from PrOF NMR titration and ITC variations of 1.8 to 4.7-fold were observed. To determine the effect of fluorine incorporation on binding to Brd4, the dissociation constants of the two strongest binders, **1** and **12**, were measured by

ITC with 5FW-Brd4 (Figure 5.20, Figure 5.21). In this case, the K_d values by ITC were similar (**12**: 3.4 vs. 2.8 μM, **1**: 18.6 vs. 14.7 μM). Besides error in fitting the data for the weak binding fragments, additional origins of these effects may be the 4-fold higher concentration of DMSO used in the PrOF NMR experiment (2% versus 0.5%) which can attenuate ligand binding. Comparison of the affinities obtained for the fluorinated proteins by PrOF NMR and ITC, suggest this could exert a 2.2 to 2.6-fold effect. The 4.7-fold difference supports an added perturbation in binding beyond differences in experimental conditions. Therefore, we conclude in the small sample of molecules studied that binding interactions seem to be minimally to modestly perturbed by fluorine incorporation (). These results are consistent with our prior studies.[131]

5.3 **Discussion**

Table 5.3 A general comparison of PrOF NMR with 1H CPMG as screening methods

^{1}H CPMG	**PrOF NMR**
Difficult to optimize without known ligand	Does not require reference compound for optimization of experimental conditions
More prone to false positives without competition experiments	Less prone to false positives
Low concentration of unlabeled protein (2-10 μM) needed	Requires moderate concentration of fluorinated protein (≤ 40 μM)
Low concentration of ligand needed (low to mid μM), solubility of ligands can be observed by resonance height	High concentration of ligand needed (high μM to mM), no information on solubility of ligands is provided
No experimental deconvolution needed, but data analysis is time consuming	Data analysis can be readily automated, but experimental deconvolution required
Faster with larger proteins, no theoretical upper bound size limitation	Faster with smaller to medium proteins, approaches a size limitation (<65 kDa for aromatic amino acid labeling)[137]
NMR time for this ^{1}H CPMG screen (15 kDa protein, three 10 minute NMR experiments): 93 hours	NMR time for this PrOF NMR screen (15 kDa protein, 2 minute NMR experiment): 29 hours (23 hours of which would be used to deconvolute 140 mixtures. Note: only 38 mixtures were deconvoluted to conserve protein)
	Facile K_d determination with titration experiment

Because ^{1}H CPMG is a ligand-observed technique, the method can report on ligand solubility as well as ligand binding interactions. However, ligand-observed methods can suffer substantially when there are no known ligands for the desired target for optimizing the experiment. The known ligand allows the tuning of screening parameters, such as ligand and protein concentration, as well as the CPMG filter length. Acetaminophen was used as a test compound for this screen, which we previously determined to bind to Brd4 with a K_d between 230 and 290 μM. With unoptimized parameters the signal change upon binding was negligible, but optimized ligand

concentration (100 μM) and CPMG filter length (800 ms) resulted in a nearly 30% drop in signal intensity (Figure 5.24). Without a reference compound, one could screen an entire library while unknowingly using poorly optimized conditions. With protein-observed techniques, once the labeled protein has been obtained the NMR parameters can all be optimized using only the protein. The effects of different ligand concentrations are more straightforward with protein-observed techniques (greater ligand concentration results in greater response) and can be initially estimated based on protein ligandability prediction. Importantly, a high-affinity reference ligand allows for competition CPMG experiments, which could help to avoid false positives resulting from non-specific binding or protein denaturants. This highlights the utility of broad spectrum inhibitors for a class of proteins, such as bromosporine for bromodomains.[53] As an alternative to a competition experiment to eliminate non-specific binding with ligand-observed NMR experiments, an STD NMR method exists that tests for specific binding by epitope mapping.[73]

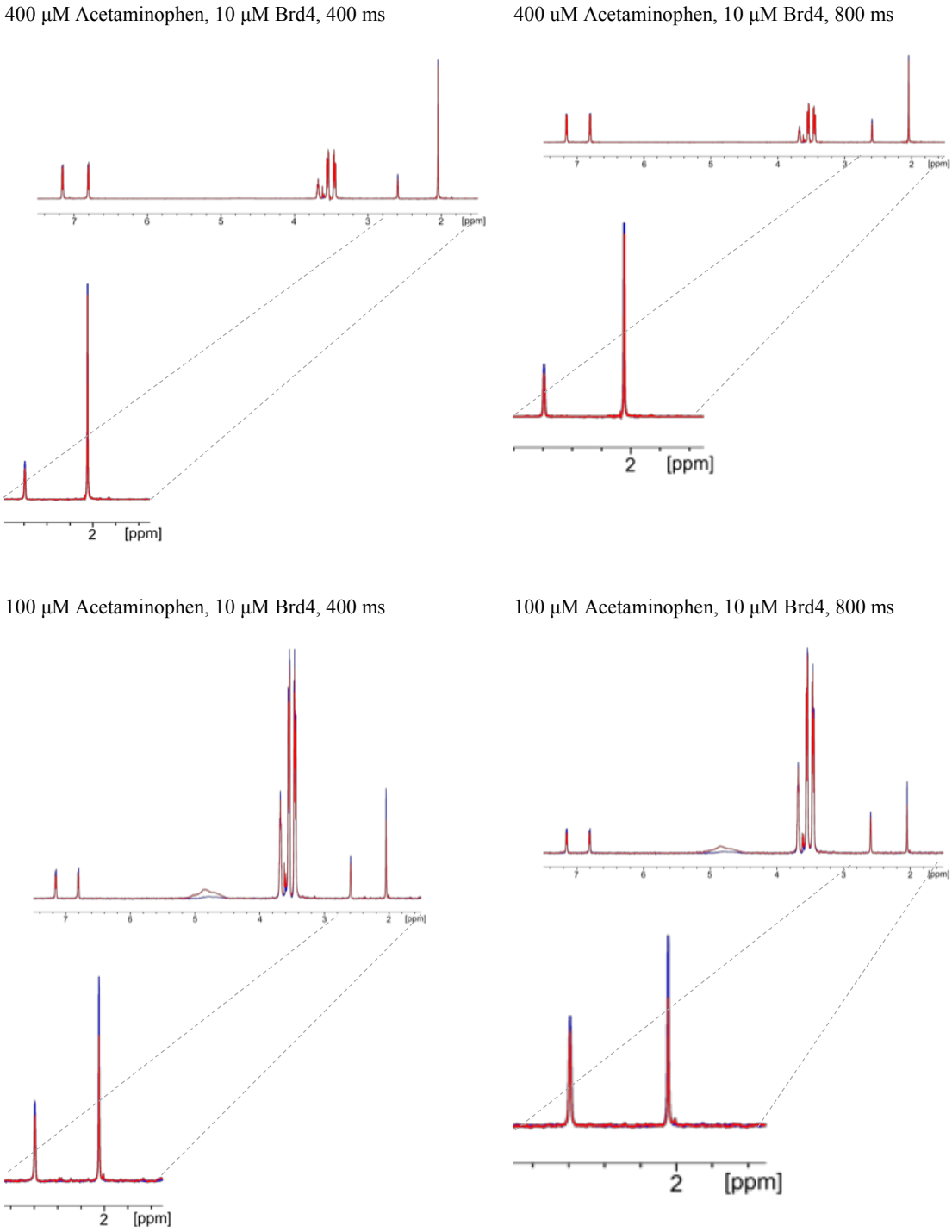


Figure 5.24 A comparison of the different CPMG parameters.
Rear NMR spectra are just ligand, while front NMR spectra are ligand + protein. Both concentration of ligand and CPMG filter length are varied, with 100 µM acetaminophen and an 800 ms filter length resulting in a greater signal drop upon protein addition than 400 µM acetaminophen with a 400 ms filter length.

A potential difficulty with using PrOF NMR is the ability to obtain fluorinated protein in sufficient yields, as well as the potential for the fluorinated protein to behave differently than the unlabeled protein, although neither of these have been problematic for bromodomains. Fluorine substitutions on aromatic rings tend to increase π-π stacking interactions by 0.5 kcal/mol per fluorine atom.[204] Work by Dougherty et al. found that incorporation of 5-fluorotryptophan into proteins can reduce cation-π interactions by 0.6-0.8 kcal/mol.[205,206] These energetic differences result in a 4-fold change in binding affinity or less, and would only occur if the fluorinated amino acid is directly involved in the binding interaction. ITC studies with **1** and **12** indicate no significant change in binding affinities between Brd4 and 5FW-Brd4, consistent with prior studies using different ligands.[131] However, the 4.7-fold difference in binding between the unlabeled protein and fluorine labeled protein for molecule **4** may reflect such a case.

Because only certain amino acids are fluorinated for PrOF NMR, some structural information is useful to determine which amino acids are close to the binding site of interest. Acetaminophen binds weakly with Brd4, and has been crystallized with the highly homologous bromodomain of Brd2. Assuming a similar binding mode between Brd4 and Brd2, the fluorine of W81 is 6.1 Å from the nearest heavy atom of acetaminophen (Figure 5.25). However, with Brd4 bound to (+)-JQ1, W75 is also slightly affected (0.13 ppm) with the fluorine atom 11.3 Å away from the nearest heavy atom of (+)-JQ1. This latter effect may occur from a subtle conformational change not observed in the x-ray structure. Several other ligands in the PrOF NMR screen perturb both W81 and W75, so the distance dependence of the fluorine atom can vary. Aromatic amino acids tend to be enriched at protein-protein interfaces,[100] and with bromodomains we found high conservation of aromatic amino acids near the binding site.[131] Therefore, for these type of interactions it is highly likely at least one aromatic amino acid will be close by to monitor.

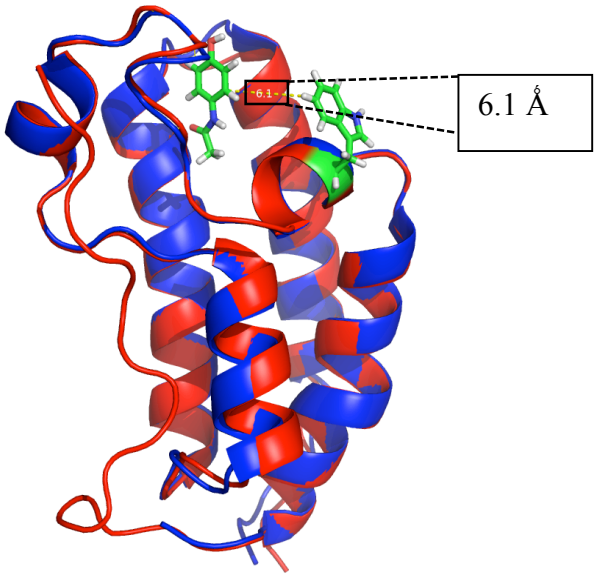


Figure 5.25 An overlay of Brd4 and Brd2 where Brd2 is crystallized with acetaminophen.
Assuming a similar binding mode, this places the nearest heavy atom of acetaminophen 6.1 Å away from the fluorine of W81 in 5FW-Brd4.

Protein size will play a large role in the speed with which spectra can be acquired. A smaller protein results in rapid PrOF NMR experiments with lengthy ^{1}H CPMG experiments, while for larger proteins the ligand-observed becomes the faster of the two methods. At this protein size, a PrOF NMR screen is substantially faster than a ^{1}H CPMG screen (2 minutes vs. 20-30 minutes for a single experiment, though experimental deconvolution is necessary by PrOF NMR), but as protein-size increases PrOF NMR will require longer experiment times due to a longer rotational correlation time leading to resonance broadening. Rule et al. has shown a 65 kDa protein labeled with 5-fluorotryptophan can lead to broad but resolved resonances.[137] In this case, experiment times would be significantly longer unless a more dynamic side chain is labeled.[89,156] Because CPMG exploits the differences in rotational correlation between small ligand and large protein, it exhibits the opposite trend, i.e., the sensitivity of CPMG increases with larger proteins. With larger proteins higher concentrations of ligand can be used, dramatically reducing experiment time. One difficulty with ^{1}H CPMG that could be

resolved with software is the ability to work up screening data. Because of the potential for overlapping resonances, and the need to correlate resonances from the screen with baseline spectra of the individual ligands, automated data processing can greatly expedite the process.[207]

Finally, protein-observed techniques are inherently less prone to false positives as long as more than one resonance is present. Because there are NMR resonances in a specific pattern due to the local chemical environment of each observed nucleus, the spectrum will reveal whether or not a protein is in solution and well-folded. If the protein is denatured or degraded, the resonances coalesce, disappear, or sharpen. This is essentially an in situ quality control for every protein-observed experiment, reducing potential false positives. Additionally, because there are multiple resonances corresponding to amino acids at different positions on the protein, the perturbed resonances frequently correlate with the rough identification of the ligand binding site, although as with all protein-observed NMR experiments that rely on chemical shift perturbations, conformational effects cannot be ruled out. However, we have found that the chemical shift perturbations for bromodomains tend to localize to residues near the binding site.[131] This type of analysis is useful if there are multiple binding sites on a protein construct where only one is the desired target. Gee et al. previously showed this in the context of the protein KIX which possesses two binding sites.[127] Additionally, because multiple labels are present which may be outside known binding sites, there is the potential for serendipitously uncovering a cryptic binding site.

In conclusion, in this study we have benchmarked PrOF NMR as having comparable sensitivity to ^{1}H CPMG with 85% assay overlap when an additional competition experiment is employed in the ^{1}H CPMG NMR experiment in the context of bromodomain screening. Without a competitor the agreement between the two assays dropped to 80%. The similar hit-rate of PrOF NMR with the well-utilized ^{1}H CPMG further validates PrOF NMR as a screening method for ligand discovery for similar proteins. Because the detection abilities of both methods were so similar, the decision of which biophysical screen to use as a primary screen is protein dependent, and the complementary nature of the data supports conducting both ligand and protein-based

experiments. Factors that can influence this decision are the availability of reference compounds, the ability to express protein in sufficient yields, protein size, and presence of multiple binding sites. One advantage of ^{1}H CPMG over PrOF NMR is the ability to avoid a time consuming deconvolution step for the fragment mixtures. We found an 85% similarity between PrOF NMR of a represented set of deconvoluted mixtures and ^{1}H CPMG, similar to the mixture data which was 77% similar between methods, as such no further deconvolution pursued. Due to the similarities in assay conditions, the possibilities of doing a sequential ligand-observed and PrOF NMR experiment during the same NMR screen may prove beneficial eliminating the need for deconvolution.

5.4 **Future directions**

Three dimensional fragments are low molecular weight compounds that have some degree of shape. The utility of 3D fragments is underexplored, despite drug-like compounds having more sp3 or 3D character.[208] Statistical analysis of drug-like compounds has shown that during each phase of the drug development process, successful compounds tend to have a significant amount of Fsp3 character, as well as at least one stereocenter.[208] Additionally, compounds that are less aromatic and have more 3D character tend to be more water soluble,[209] an important quality for use in vivo and in biophysical assays. Therefore, 3D-enriched fragment screens could produce higher quality drug-like hits than traditional 2D-enriched screens.

To evaluate the effectiveness of 3D-enriched fragments, five hundred 3D-enriched fragments were screened against Brd4, a model protein, using ^{1}H CPMG. The active-rate will be compared with the 2D library presented in the preceding text. These compounds will be further elaborated to ascertain the quality of 3D-fragment hits.

5.5 **Materials and Methods**

Preparation of fragment library:

The fragments in the screening library were initially analyzed by ^{1}H NMR, and any compounds with low resonance intensity from low solubility or aggregation propensity were removed from the set. The fragments were then combined into mixtures

of five in such a way to maximize the diversity of compounds in each mixture, while avoiding mixing reactive functional groups. Although the fragment library consists of a majority of commercial fragments, several fragments are proprietary compounds of Eli Lilly & Company. **12** was one such compound, whose structure has been omitted.

Expression of 5FW-Brd4 (42−168):

Unlabeled and 5FW-labeled Brd4 were expressed based on established methods[32,131] using E. coli Bl21(DE3) + pRARE strains. To express the labeled protein, the secondary culture in LB media was grown until an OD_{600} of 0.6 was reached followed by harvesting. Cells were resuspended in defined media of Muchmore et al.[110] containing 5-fluoroindole (60 mg/L) in place of tryptophan.[96] The resuspended E. coli were incubated at 37 °C while shaking for 1.5 h followed by cooling to 20 °C and media temperature equilibration for 30 min. Protein expression was induced with 1 mM IPTG overnight (14−16 h) at 20 °C. The cells were harvested and stored at −20 °C. Cell pellets were thawed at room temperature followed by the addition of lysis buffer (50 mM Phosphate pH 7.4, 300 mM NaCl) containing protease inhibitor PMSF (5 mM) as well as the Halt protease inhibitor and purified according to methods described in the appendix using Ni-affinity chromatography. Yields following purification are 120 mg/L 5FW-Brd4 (>94% fluorine incorporation assessed by mass spectrometry). Purity of proteins was assessed by SDS-PAGE. Fluorinated amino acid incorporation efficiency in proteins was measured by mass spectrometry as described in the appendix. Concentration was determined via absorbance at 280 nm.[177]

Protein-observed fluorine (PrOF) NMR:

1D ^{19}F NMR Parameters: ^{19}F NMR spectra were acquired at 565 MHz on a Bruker Avance III spectrometer equipped with a quadruple resonance HFCN CryoProbe without proton decoupling, unless otherwise specified. Samples for binding assays contained 50 μM 5FW-Brd4 and 5 fragments each at 400 μM in 50 mM Tris, 100 mM NaCl, and 5% D_2O, pH 7.4. Spectra were referenced to trifluoroacetate (−76.55 ppm).

Measurement parameters included a relaxation delay time of 0.7 s and a 58° pulse flip angle (based on the Ernst angle from T1 determination). An acquisition time of 0.05 s and a spectral width of 10 ppm were used for all experiments. Thirteen-minute screening experiments used 1000 scans while two-minute screening experiments used 160 scans. Small molecules were titrated into the protein solution from concentrated stock solutions of DMSO (100 mM). Final DMSO concentrations were kept at or below 2%. Proton decoupling was not used because it results in a reduction in signal intensity due to the negative nuclear Overhauser effect with large molecules. Additionally, the increased NMR linewidth from large biomolecules obscures couplings.

^{1}H CPMG:

^{1}H NMR spectra were first collected on all compounds in the fragment library to be used as reference spectra for deconvoluting screening data. Samples were prepared with 100 μM each of fragment, and a ^{1}H CPMG spectrum acquired of each mixture (with a CPMG filter length of 1.2 s and an interpulse delay of 2.5 ms). In a second step, a concentrated protein stock solution was then added to each sample to a final concentration of 10 μM Brd4, and a ^{1}H CPMG spectrum was recorded. Finally, known competitor (+)-JQ1[13] was added to a concentration of 20 μM, and competition was monitored by recovery of signal intensity.

Fluorescence anisotropy:

Fluorescence anisotropy was measured from an excitation wavelength of 485 nm and an emission wavelength of 535 nm on a Tecan Infinite 500 plate reader using low volume 384-well plates (Corning 4511). All experiments were carried out in 50 mM Tris, 150 mM NaCl, and 4 mM CHAPS at pH 7.4. The fluorescently labeled tracer, BI-BODIPY, was synthesized according to published methods.[165] For both direct binding and competition experiments, 25 μM stocks of a BI-BODIPY stock solution in DMSO were diluted to a final concentration of 25 nM. For direct binding experiments, Brd4 was serially diluted from micromolar to subnanomolar concentrations. For competition

experiments, Brd4 was kept at a constant concentration of 156 nM, equivalent to 80% bound tracer as determined in direct binding experiments, and the concentration of competing ligand was serially diluted from micromolar to subnanomolar concentrations. Data were collected within 30 min after plating to minimize Brd4 binding to the plate surface. All experiments were carried out in triplicate. Acquired data were fit using GraphPad Prism. K_d values were determined by fitting to the following equation, which accounts for ligand depletion. In this equation, *b* and *c* are the maximal and minimal anisotropy values, respectively, *a* is the concentration of fluorescently labeled tracer, *x* is the protein concentration, and *y* is the observed anisotropy value.

$$y = c + (b - c)\frac{(Kd + a + x) - \sqrt{(Kd + a + x)^2 - 4ax}}{2a}$$

IC_{50} values were determined using GraphPad Prism's log(inhibitor) vs response function. K_i values were obtained using a variant of the Cheng-Prussof equation from Huang et al.[102]

Isothermal titration calorimetry:

Experiments were carried out on a MicroCal Auto-iITC200 titration from Malvern with a cell volume of 200 µL and a 40 µL microsyringe. Experiments were carried out at 25 °C while stirring at 750 rpm, in ITC buffer (50 mM phosphate buffer at pH 7.4 and 150 mM NaCl). The microsyringe was loaded with a solution of ligands whose concentrations were accurately measured by quantitative NMR (ITC buffer with 0.5% DMSO-d_6), and was automatically inserted into the calorimetric cell which was filled with an amount of the protein, Brd4 and 5FW-Brd4 (200 µL, 30 µM in ITC buffer with 0.5% DMSO-d_6). The system was first allowed to equilibrate until the cell temperature reached 25 °C and an additional delay of 60 s was applied. These first titrations were conducted using an initial control injection of 0.5 µL followed by 18 identical injections of 2 µL with a duration of 4 s per injection and a spacing of 150 s between injections. Afterwards, a second titration was performed over this first titration

without cleaning the cell. This second titration was executed with Continue Injections as Automation Method in MicroCal Auto-iTC200. Second titrations were conducted using 19 identical injections of 2 µL with duration of 4 s per injection and a spacing of 150 s between injections. The titration experiments were designed in such a fashion, as to ensure complete saturation of the protein before the final injection. The heat of dilution for the ligands was independent of ligand concentration and corresponded to the heat observed from the last injection, following saturation of protein binding, thus facilitating the estimation of the baseline of each titration from the last injection. The collected data were corrected for ligand heats of dilution, and deconvoluted using the MicroCal PEAQ-ITC Analysis Software to yield enthalpy of binding (ΔH) and binding constant (K_d). Thermodynamic parameters were calculated using the basic equation of thermodynamics ($\Delta G = \Delta H - T\Delta S = -RT\ln K_d$, where ΔG, ΔH and ΔS are the changes in free energy, enthalpy and entropy of binding respectively). A single binding site model was employed by MicroCal PEAQ-ITC Analysis Software.

Thermal shift assays:

Thermal shift assay (TSA) measures the thermal stability of the Brd4 wild-type, and the comparison between the melting temperature of apo-protein and melting temperatures in the presence of different compounds. Previous exploratory phase was performed to determine optimal experimental conditions for this assay in the same buffer as ITC (50 mM phosphate buffer at pH 7.4 and 150 mM NaCl). In 96-well PCR plate with a sample of 20 µL in each well (2.5% DMSO) Brd4 was at 26 µM, fragments were assayed at 250 µM, and Protein Thermal Shift™ Dye was utilized at 2X (Protein Thermal Shift™ Dye Kit from ThermoFisher Scientific). Fluorescence data was collected on an Applied Biosystems 7500 FAST RealTime PCR System with an excitation range of 580 ± 10 nm. The fluorescence emission signal at 623 ± 14 nm was used for data analysis. Samples were preheated for 2 minutes at 25 °C, then the temperature was continuously increased 2 °C/min from 24 to 99 °C, and finally samples were maintained for 2 minutes at 99 °C.

Acknowledgements:

This project was funded in part by the NSF-CAREER Award CHE-1352091 (W. C. K. Pomerantz) and the NIH Biotechnology training grant 5T32GM008347-23 (A. K. Urick). Some NMR spectra were recorded on an instrument purchased with support from the NIH Shared Instrumentation Grant program (S10OD011952). We kindly thank A. Ayoub for assistance with the fluorescence anisotropy experiments.

Protein Expression and Molecular Biology Materials:

For *E. coli* growth, LB agar, LB media, defined media components including unlabeled amino acids, uracil, thiamine-HCl, nicotinic acid, biotin and buffer components were purchased from RPI corp. Thymine, cytosine, guanosine were purchased from Alfa Aesar. Magnesium chloride, manganese sulfate, succinic acid, calcium chloride and 5-fluoroindole were purchased from Sigma-Aldrich. Miniprep plasmid purification kit was purchased from Clontech.

Unlabeled Brd4 Protein Expression:

The pNIC28-BSA4 plasmid containing the first bromodomain of Brd4 genes were kind gifts from the laboratory of Stefan Knapp. For protein expression, either the *E. coli* Rosetta (DE3) strain (Novagen) was first transformed with the respective expression plasmid or the BL21(DE3) strain was cotransformed along with the pRARE (Novagen) plasmid and plated onto agar plates containing kanamycin (100 mg/L) and chloramphenicol (35 mg/L). Following overnight incubation at 37 °C, a single colony was selected from the agar plate and inoculated in 50 mL of LB media containing kanamycin (100 mg/L) and chloramphenicol (35 mg/L). The primary culture was grown overnight at 25 °C while shaking at 250 rpm. For secondary culture growth, 1 L of LB media containing kanamycin (100 mg/L) was inoculated with the primary culture and cultured at 37 °C while shaking at 250 rpm. When the O.D. of culture at 600 nm reached 0.6, the shaker temperature was reduced to 20 °C. After 30 minutes, the expression was

induced with 1 mM IPTG overnight for 12-16 h. Cells were harvested by centrifugation and stored at -20 °C.

Bromodomain Purification:

To purify fluorinated and unlabeled Brd4, the cell pellet was thawed at room temperature followed by the addition of lysis buffer (50 mM Phosphate pH 7.4, 300 mM NaCl) containing protease inhibitor PMSF (5 mM) as well as the Halt protease inhibitor. Cells were lysed by sonication and the cell lysate was centrifuged at 7500 g for 30 minutes followed by supernatant filtration over Whatman filter paper. Filtrate containing the histidine-tagged Brd4 was loaded on to a nickel-NTA affinity column and eluted with an imidazole gradient on an AKTA FPLC system monitoring the O.D. at 280 nm. Imidazole was removed from the buffer using a HiPrep column (GE) for buffer exchange into 50 mM Tris pH 7.4, 100 mM NaCl. Purified and buffer exchanged protein was treated with TEV protease for either 2 hours at room temperature or alternatively at 4 °C overnight on a rotating carrousel. The cleaved His-tag, TEV protease and uncleaved Brd4 were removed using nickel-NTA affinity resin.

Protein Mass Spectral Analysis:

Product molecular weight was confirmed by electrospray ionization mass spectrometry (ESI-MS) using a Waters Acquity UPLC/Synapt G2 QTOF LC-MS. To determine the percent incorporation for fluorinated proteins the integration values of the different deconvoluted mass peaks are entered into the following equation to determine the relative incorporation e.g, FWBrd4:

$$\%_{incorporation} = \frac{(0FWBrd4)*0+(1FWBrd4)*1+(2FWBrd4)*2+(FWBrd4)*3}{(0FWBrd4)*3+(1FWBrd4)*3+(2FWBrd4)*3+(FWBrd4)*3} * 100$$

0FWBrd4 is 5FWBrd4 with no fluorine substitutions, 1FWBrd4 is 5FW-Brd4 with one fluorine substitution, 2FWBrd4 has two fluorine atoms substituted and FWBrd4 has 3 fluorine atoms substituted. Below is an example mass spectrum of 5FW-Brd4 (expected mass of triply fluorinated 5FW-Brd4, 15137 m/z).

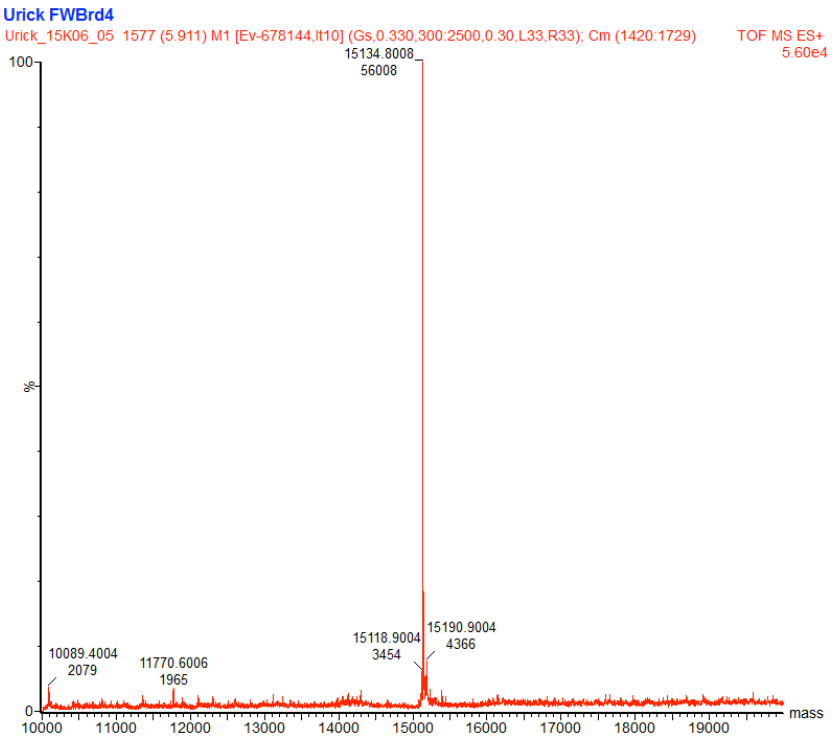


Figure 5.26 Protein mass spectrum of 5FW-Brd4.

Chapter 6. Prediction of ^{19}F NMR chemical shifts in labeled proteins

Note: The following was largely the work of Dr. William C. Isley III, and is also present in Chapter 2 of his dissertation "Factors Affecting Recognition and Chemical Reactivity at the Macromolecular Scale". To place my contributions to this work in context, the entire study has been reproduced below.

Motivation: Resonance assignment of the protein's 19F NMR spectrum is necessary for structural analysis, but can be challenging. Here, a quantum chemical method has been developed as an initial approach to facilitate the assignment of a fluorinated protein's 19F NMR spectrum. The epigenetic "reader" domain of protein Brd4 was taken as a case study to assess the strengths and limitations of the method.

6.1 **Introduction**

Epigenetic proteins regulate the expression of genetic information through addition, removal, or molecular recognition of posttranslational modifications of DNA or DNA associated proteins. Bromodomains are epigenetic protein modules that bind to *N*-ε-acetyl groups on lysine side-chains including those of acetylated histone proteins. Small molecule chemical probes for these proteins are in high demand for their potential therapeutic regulation of disease.[210] Since the first reports[13,17] in 2010 of two nanomolar inhibitors for BET bromodomains Brd2, 3, 4 and T, eighteen clinical trials have been initiated to test the efficacy of BET bromodomain inhibition in the areas of cancer and inflammation.[158,210] There are many other bromodomains, however, which lack specific chemical probes to evaluate their role in both health and disease.[180]

We recently reported a protein-based NMR method for bromodomain ligand discovery, using fluorine-labeled aromatic amino acids.[127,131] Since inception, this method has been used for screening libraries of low-complexity, small molecules termed

fragments,[99,127,128,193] as well as higher complexity molecules based on kinase inhibitor scaffolds.[32] In these experiments, proteins are expressed in the presence of fluorine-labeled amino acids (e.g., 3-fluorotyrosine, or 3FY) resulting in global replacement of the non-fluorine labeled aromatic amino acid. A feature of this method is the sensitivity of the ^{19}F nucleus to different chemical environments, typically leading to rapidly obtained and well-resolved 1D ^{19}F NMR spectra. In the case of the first bromodomain of Brd4, we replaced all seven tyrosine residues with 3FY resulting in dispersed resonances spanning over 12 ppm (Figure 6.1).[131]

A notable challenge when conducting protein-observed ^{19}F NMR experiments (or PrOF NMR), is the initial assignment of the NMR resonances. This is most often facilitated by site directed mutagenesis experiments.[106] In these experiments a particular amino acid that is labeled with fluorine is mutated to a different amino acid, and the ^{19}F NMR spectrum of the mutant protein compared to the wild type protein. The absence of a single resonance is then used to assign the NMR spectrum. However, this method is not generally applicable as not all mutant proteins may express well, or multiple resonances may be perturbed in the NMR spectrum. In the case of 3FY-labeled Brd4, several of the mutant proteins expressed at low levels and some were more susceptible to degradation, thus complicating the NMR analysis particularly for the 3FY resonances for Y118 and Y119.[131] To help overcome challenges of mutagenic protein expression, we reasoned that a computational approach to NMR chemical shift prediction would function as a useful tool in facilitating the ^{19}F NMR assignment motivating this study. Protein based NMR ensemble (NMRE) simulations for ^{1}H, ^{13}C, and ^{14}N have been able to utilize the wealth of available NMR data in the literature to create a predictive protocol from machine learning algorithms; however, the database on ^{19}F NMR in proteins is significantly smaller.[211] Sampling of the protein's conformational ensemble has been shown to significantly improve prediction accuracy over simply using the X-ray crystal structure.[212,213]

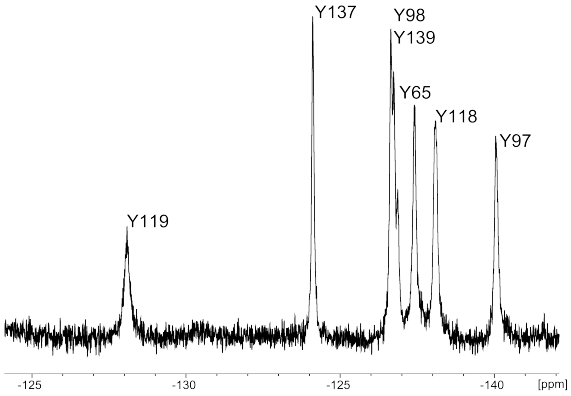


Figure 6.1 Measured 19F NMR spectrum of 3FY-labeled Brd4.
So that the resonance width can be observed unmodified, no line broadening has been applied to this spectrum. However, modest application of line broadening will increase both S/N and reproducibility of chemical shift.

Chemical shift predictions for fluorinated proteins have highlighted several challenges for theory. Using the 5-fluorotryptophan-labeled galactose binding protein, Pearson et al. predicted the ^{19}F NMR spectrum with reasonable agreement with experimental data.[214] Additionally, Sternberg *et al*.[215] developed a semi-empirical protocol for prediction of fluorotryptophan ^{19}F NMR for a solid state membrane bound protein gramicidin A that gave reasonable agreement with more rigorous levels of theory. Conversely, despite systematic analysis of local electrostatic effects and short range contacts, Lau and Gerig were unable to reliably predict the ^{19}F NMR spectrum of the more dynamic 6-fluorotryptophan-labeled dihydrofolate reductase.[216] These mixed results for fluorotryptophan illustrate the challenges associated with making accurate predictions. The challenge is enhanced by the even more narrow spectral range of ≈ 2 ppm for 5-fluortryptophan-labeled Brd4. Recent work by Kasireddy *et al*.[217] demonstrates the challenges and potential utility of ^{19}F NMR predictions on fluorohistidine molecules. By contrast, the prediction of the chemical shifts within 3FY-substituted proteins has not been previously attempted.

In this report, we propose that the wider spectral range of 3FY provides a novel and significantly more accessible platform to predictively assign the full ^{19}F NMR spectrum of a 3FY-labeled bromodomain, Brd4. The enhanced utility of 3FY as a

predictive platform is more challenging than other fluorinated aromatics (4-phenylalanine or fluorotryptophan) given the increased conformational flexibility and additional hydrogen bonding interactions provided by the phenol and asymmetric substitution. We also demonstrate that automation of spectral predictions for ^{19}F NMR in bromodomain-containing proteins is feasible, and that a cluster based method for prediction of ^{19}F NMR chemical shifts shows great promise for further development. While traditional predictions of protein-based ^{1}H NMR involve dynamic simulations to sample the large ensemble of configurations, the relative rigidity and availability of X-ray crystal structures for bromodomains, and the limited number of side-chain resonances requiring assignment, render an exploratory cluster-based model feasible. Due to the high conservation of aromatic amino acids in the majority of the 61 bromodomains,[131] this method may find general utility for these proteins.

6.2 **Theoretical Methods**

In this section, we summarize our overall protocol to arrive at ^{19}F chemical shift predictions starting from a Protein Data Bank (PDB) file for an unlabeled protein. The solution of the crystal structure of 3FY-labeled Brd4 (PDB 4QZS) was shown to have minimal structural perturbation (RMSD 0.089 Å) vs the unlabeled protein (PDB 3MXF).[13,131] We include full details of file manipulation and software employed; scripts developed for process automation are also publically available. This work utilizes the overall scheme below, with details described in the following sections.

1. The protein structure is taken from the PDB file.
2. To obtain an initial solvation environment around the target residues, a molecular dynamic optimization of water molecules on a frozen protein is performed.
3. From the solvated protein, clusters are excised around each 3FY. The NMR ab initio calculations of the full protein are prohibitively expensive, so protein fragments are used.

4. Geometries sampling the fluorine and phenol conformers are generated, and the target 3FY is optimized within a frozen protein fragment using density functional theory.
5. NMR chemical shifts are predicted using a Boltzmann average over the cluster conformers.

6.2.1 Hydration of Protein and Optimized H-atom Positions

We used the online Molecular Dynamics on Web (MDWeb) toolset,[218] initially going through the following steps for a given protein structure:

Select action "Prepare Structure Topology for AMBER ParmFF99SB* (Hornak & Simmerling, including Best & Hummer psi modification)"[219,220]

Select action for structural optimization of 50 or more exterior water molecules using the Classical Molecular Interaction Potentials (CMIP)[221]

Select action to energetically minimize hydrogen atom positions using NAMD[222]

Select action to export PDB structure

File I/O: The resulting PDB exported from MDWeb lacks the column at the end of a regularly formatted PDB file that specifies the actual atom designation for each ATOM type. Opening the PDB file in OpenBabel[223], and choosing to convert PDB -> PDB fixes this issue.

6.2.2 Cluster Generation

The final, properly formatted PDB file serves as input to the cluster generation script. Additional input includes specification of a target residue number and a cutoff radius to be used for cluster generation. Waters solvating the exterior of the protein are optimized with the AMBER force field during step 3, which samples the solvation configurational space rapidly and serves as a platform for cluster exploration.

The script generates a local cluster by identifying other protein residues within the cutoff radius of the target residues' non-hydrogen atoms. Once all such surrounding residues have been found, the script keeps them in position, caps all open backbone termini with acetyl- (N terminus) or *N*-methylamino- (C terminus) groups, and eliminates all remaining atoms. Note that any crystallographically conserved water molecules, counter ions, or small molecules in the PDB file are removed during the prepare topology

step—water molecules are added back to the structure, however, in the next step. The interior water molecules are added during cluster generation, and are manually inserted so as to match the phenol-water distance determined from X-ray diffraction, along with the oxygen atom oriented at same angle from the aromatic plane as in the XRD. The orientations of the hydrogen atoms on the water relative to the phenol (*not* available from XRD) are chosen so that the water can participate in hydrogen bonding (either as a donor or acceptor depending on orientation and nearby side chain functionality). 3FY residues exposed to the protein surface are solvated with one to three water moieties. The coordinates of these waters are allowed to relax, as discussed further below. Once a cluster (with no fluorine atoms) has been generated, four conformers are created manually, corresponding to the four relative orientations of fluorine and the phenolic proton (Figure 6.3).

6.2.3 Optimization

The geometry of the target residue is then optimized with Density Functional Theory, within a frozen cluster, i.e. only the fluorinated residue is optimized within its otherwise fixed cluster framework. This step employs the M06-L[224] density functional and the def2-SVP basis set[225] on all atoms. The aqueous SMD[226] continuum solvation model is also employed, as early surveys of gas-phase results were found to give poor structures and also to suffer from convergence difficulties in clusters having local charge separations.

6.2.4 19F NMR Chemical Shift Prediction

Following restrained optimization of the cluster, the ^{19}F NMR chemical shift is predicted. Trifluoroacetate (δ_{exp} = -76.55 ppm) is employed to compute a reference chemical shielding (σ_{ref}), with computations of 3FY chemical shifts δ_{pred} then being determined as

$$\delta_{pred} = (\sigma_{ref} + \delta_{exp}) - \sigma_{pred}$$

where σ_{pred} is the shielding predicted for 3FY in the cluster. All chemical shifts are reported relative to $CFCl_3$ (set to 0.0 ppm). To predict chemical shifts most accurately, a linear regression of predicted values on experimental measurements is common practice. We benchmarked several computational protocols on a training set of 14 molecules

containing aryl-fluorine bonds (Table 6.1) evaluating such model parameters as 1) gas phase vs SMD implicit solvation, 2) density functional choice (B3LYP vs PBE0), and 3) basis set size (double-zeta vs triple-zeta quality). Most protocols performed well over the benchmark set, PBE0 with implicit solvation provided the highest accuracy for prediction of the chemical shift of 3-fluorotyrosine. We adopted PBE0/SMD with the EPR-II basis set as our recommended protocol, as it offers an optimal combination of accuracy and computational efficiency. The corresponding linear regression to be used with chemical shifts predicted from this level of theory is

$$\delta_{fit} = 0.860 * \delta_{pred} - 24.6$$

Additional analysis is provided below in Results and Discussion.

Table 6.1 Performance of 19F δ Prediction Protocols in Comparison to Experiment.

		^{19}F δ (ppm) [a]					
		Exp	SMD Implicit Solvation				
			PBE0		B3LYP		HF
Molecule			EPR-II	EPR-III	EPR-II	EPR-III	EPR-III
2-fluorophenol [1 b]	1	-141.9	-144.4	-136.6	-147.0	-146.8	-152.0
2,6-difluorophenol [2 b]	2	-139.1	-140.2	-133.5	-143.2	-144.0	-146.2
2-fluoro-*p*-hydroquinone [1 b]	3	-138.7	-140.3	-132.6	-143.1	-143.1	-145.7
3-fluorotyrosine [b]	4	-137.6	-142.6	-135.8	-145.2	-146.3	-151.8
5-fluoroindole [3]	5	-126.55	-135.2	-129.1	-138.7	-140.1	-144.2
6-fluoroindole [3]	6	-121.83	-131.5	-125.3	-135.2	-136.4	-137.9
5-fluorotryptophan [3]	7	-126.15	-134.6	-128.7	-138.1	-139.9	-143.7
6-fluorotryptophan [3]	8	-122.85	-131.1	-125.5	-134.8	-136.8	-138.4
fluorobenzene	9	-113.15	-121.3	-113.8	-125.2	-124.7	-136.3
hexafluorobenzene	10	-164.9	-166.7	-163.9	-168.1	-173.5	-130.8
4-fluorophenylalanine	11	-116.7	-125.8	-119.2	-129.7	-130.4	-148.5
Pentafluoro-benzene (1,5F)	12	-139.9	-143.0	-138.5	-145.3	-148.5	-172.5
Pentafluoro-benzene (2,4F)	13	-162.1	-167.0	-163.9	-168.6	-173.7	-161.8
Pentafluoro-benzene (3F)	14	-153.5	-158.5	-154.6	-160.4	-164.6	-169.3
Statistics							
Slope			0.860	0.930	0.808	0.897	0.204
Intercept			-24.55	-9.22	-34.40	-24.25	-109.69
R**2			0.9737	0.9541	0.9695	0.9477	0.0683

[a] Chemical shifts reported relative to $CFCl_3$ at 0.00 ppm and trifluoroacetate at -76.55 ppm [b] These molecules have one explicit water molecule to accuarely model effect of hydrogen bonding to phenol on ^{19}F chemical shifts.

In many instances, a 3FY residue can adopt multiple poses within its associated cluster, and the phenol group leads further to multiple possible rotamers. To account for an equilibrium distribution of structures, Boltzmann weighted chemical shifts were computed as

$$\delta_{tot} = \sum_{c_i} \delta_{fit}^{c_i} \left(\frac{e^{-\Delta G^{c_i}/RT}}{\sum_{c_i} e^{-\Delta G^{c_i}/RT}} \right)$$

where G is a relative solvated electronic energy for conformer c_i.

Software. All optimization and chemical shift computations were accomplished using the Gaussian09 Rev D.01 suite of electronic structure programs.

6.3 **Results and Discussion**

No benchmarking study for the accuracy of ^{19}F NMR chemical shift predictions for biomolecular moieties at the density functional (DFT) or Hartree Fock (HF) level of theory was available. To validate our own protocol, we explored various options over a 14-molecule training set as outlined in the theoretical methods. This protocol was employed to predict composite ^{19}F chemical shifts using local clusters and accounting for conformational flexibility. We next address the modeling challenges associated with our protocol and the physical insights into the effects of chemical environment on ^{19}F NMR shifts in proteins that it provides.

6.3.1 Optimizing 19F NMR Prediction Protocol

Prior to cluster generation, a method to predict the ^{19}F NMR of aryl fluorine resonances in a training set of different fluorinated aromatic rings spanning about the same spectral frequency range as 3FY (-125 to -145 ppm) was optimized.[131] The training set and experimental NMR data are reported in Table 6.1; data for a subset are shown in Figure 6.3. The phenols in the training set all include one explicit water molecule acting as a hydrogen bond acceptor to represent the solvation shell. Addition of water to 3FY shifts δ_{pred} downfield by 8 ppm, from -150 ppm to -142 ppm. The protocols tested include molecular optimization with M06-L/def2-SVP in either the gas phase or with an aqueous SMD solvation model. After these structures were obtained, NMR predictions were performed using either the PBE0[227] or B3LYP[228] density functionals in combination with either the EPR-II or EPR-III basis sets. Additionally, NMR predictions were made at the Hartree Fock level of theory using the 6-311++G(2d,2p)[229,230] basis set and the aqueous SMD solvation model.

Results for selected regressions of computed training set ^{19}F NMR data on experimental measurements are shown in Table 6.2. Plotted in Figure 6.2, the correlation between experiment and theory selected for further use was found for the protocol

combining the aqueous SMD solvation model,[226] the PBE0 density functional,[227] and the EPR-II basis set. A similar statistical performance was obtained with the larger basis PBE0/EPR-III/SMD model, but we chose to continue with PBE0/EPR-II/SMD based on its lesser computational expense. B3LYP methods offered similar levels of accuracy across the set. PBE0 was selected given its better performance for 3FY (Table 6.1). Although some prior research suggested that ^{19}F chemical shifts for fluorobenzenes were accurately predicted at the Hartree Fock level of theory,[231] we found density functional methods to be much more accurate over our biologically motivated training set. ^{19}F NMR has been found to be more sensitive to changes in the electrostatic potential[232] and it has been shown that including exact Hartree-Fock exchange in hybrid density functionals increasingly degrades the prediction of nuclear shieldings as nuclei become heavier.[233] Compared to ^{1}H and ^{13}C nuclei, modeling of ^{19}F NMR chemical shifts in other systems has found greater errors for HF predictions than for those at the DFT or MP2 levels.[232–234]

Table 6.2: Performance of ^{19}F δ Prediction Protocols Compared to Experiment After Linear Regression.[a]

Protocol	Intercept	Slope	Adj. R^2
PBE0 /EPR-II	-24.6 ± 5.4	0.860 ± 0.039	0.974
PBE0 /EPR-III	-9.2 ± 7.7	0.930 ± 0.057	0.954
B3LYP/EPR-II	-34.4 ± 5.4	0.809 ± 0.040	0.969
B3LYP/EPR-III	-24.3 ± 8.0	0.897 ± 0.058	0.948
HF /6-311++G(2d,2p)	-109.7 ± 27.9	0.285 ± 0.204	0.068

[a] All models employed aqueous SMD continuum solvation.

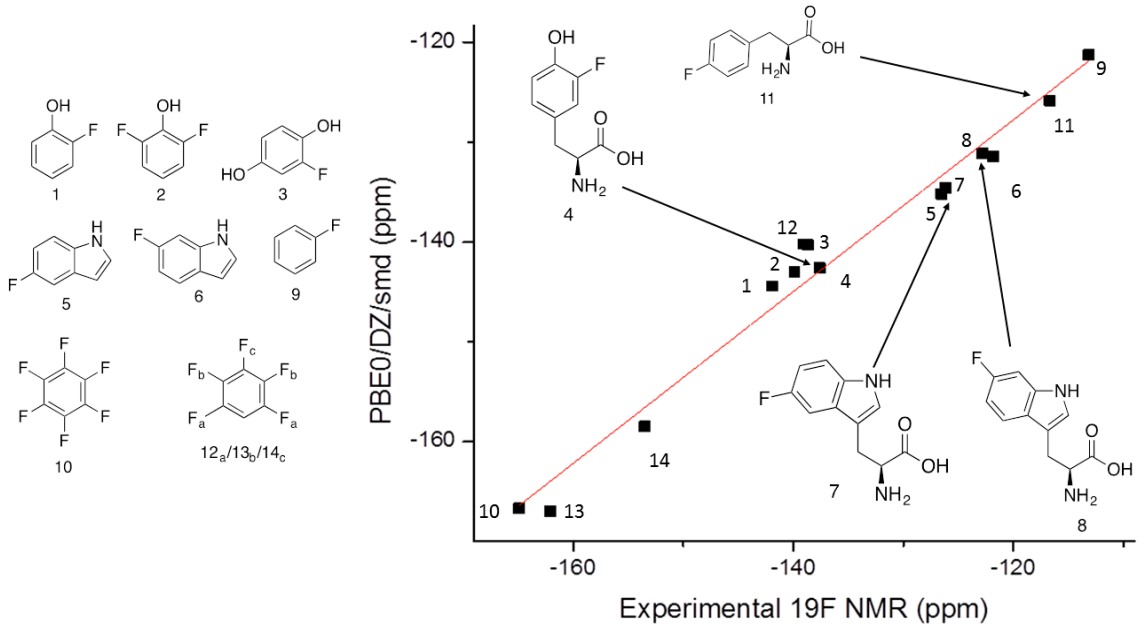


Figure 6.2 All molecules from the training set are shown, as well as predicted 19F δ from PBE0/EPR-II/SMD vs. experiment for the full training set.

6.3.2 Cluster Radial Convergence

Accurate modeling of the fluorine-19 isotope's magnetic behavior requires that cluster models reproduce the local environment derived from the full protein. The fluorinated resonance of Y65 ($3FY_{65}$) of the apo form of the Brd4 bromodomain (PDB ID: 4IOR) was selected to evaluate the convergence of predicted chemical shift with respect to cluster size. Residue $3FY_{65}$ serves as a sensitive test case since it is a solvent-exposed amino acid which might be expected to sample quite different environments in different rotamer states (e.g., protein interior vs exterior directed fluorine). The relative energies and chemical shifts for each conformer are included in Table 6.3. During cluster generation, a radial cutoff of at least 2.75 Å is required in order to include the residue having a carbonyl group that can serve as a hydrogen bond acceptor for the phenol. A cutoff of 3.25 Å is required to encompass surrounding water molecules since hydrogen atoms are not used in cluster generation. After these previously absent possible phenolic hydrogen bond acceptors (and/or donors, although 3FY is generally a better acid than a base in this regard) have been included, the convergence of the predicted fluorine NMR

chemical shift improves. However, the relative stability of each conformer as a function of cluster size is not as well converged. Taking into consideration the computational expense of the chemical shift prediction step, clusters generated with a radial cutoff of greater than 4.00 Å were found to be too large to be conveniently employed in the chemical shift calculation step. In general, to balance computational efficiency and accuracy in modeling the local environment, we recommend employing a cutoff distance between 3.25 Å and 4.00 Å. Unless otherwise noted, results reported below are for a cutoff distance of 4.00 Å.

Table 6.3: Convergence of $3FY_{65}$ ^{19}F NMR δ (ppm) as a Function of Cutoff Distance.[a]

			$\Delta E(_{s\text{-}cis\text{-}in})$	$\Delta E(_{s\text{-}trans\text{-}ex})$	$\Delta E(_{s\text{-}cis\text{-}ex})$	^{19}F NMR δ_{fit} (ppm)[d]				
R (Å)	# AA[b]	# PP[c]	kcal/mol	kcal/mol	kcal/mol	s-*trans*-in.	s-*cis*-in	*s*-*trans*-ex	s-*cis*-ex.	δ_{pred} (ppm)
2.00	2	1	-2.85	0.06	-2.65	-133.8	-153.3	-133.5	-152.2	-152.5
2.25	2	1	--	--	--	--	--	--	--	--
2.50	5	1	-2.63	-2.26	-3.92	-137.0	-156.3	-136.1	-153.8	-155.0
			0.00	-0.46	3.97		-130.9	-146.7		
2.75	6	2				N/A	[e]	[e]	-149.6	-141.7
			0.01	-1.57	6.32		-132.1	-136.5		
3.00	8	3				N/A	[e]	[e]	-145.8	-136.2
			-2.96	-4.33	-2.26		-122.5	-129.4	-133.9	
3.25	11	2				-121.0[f]	[e]	[e]	[f]	-129.6
3.50	11	2	--	--	--	--	--	--	--	--
			-3.56	-4.97	-2.59	-121.0[f]	-127.9	-132.4	-136.6	-132.1
3.75	14	3					[e]	[e]	[f]	
4.00	14	3	--	--	--	--	--	--	--	--
EXP	--	--	--	--	--	--	--	--	--	**-137.4**

a The isomeric environments are separated into interior (in.) vs exterior (ex.) and s-cis vs s-trans. Energy differences are reported relative to the s-trans-in conformer. b AA is the number of residues kept with side chains intact (does not include caps). c PP is the number of disconnected peptide chains. d δfit is the Boltzmann weighted 19F NMR chemical shift e Includes a hydrogen bond to the carbonyl oxygen of an adjacent peptide. f Includes a hydrogen bond to external water.

6.3.3 3-Fluorotyrosine Conformer Weights

One challenge from a chemical modeling standpoint is the accurate sampling of all thermodynamically relevant conformers. For the 3FY systems considered here, there are nearly always four relevant conformers (Figure 6.3**Error! Reference source not found.**). These conformers account for the internal orientation of the phenol hydrogen relative to the fluorine, and the relative orientation of the fluorine to the protein tertiary structure. The number of accessible conformers can increase if additional phenol hydrogen-bond acceptors are present (non-hydrogen bonded conformers are generally much higher in energy). These different conformers expose the sensitive fluorine probes to different magnetic environments.

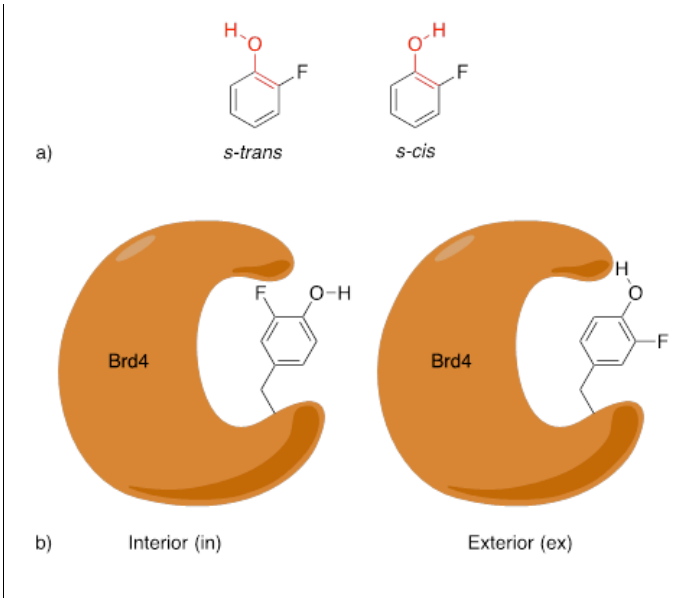


Figure 6.3 Nomenclature of tyrosine conformations.
a) s-cis vs s-trans conformers. b) s-trans 3-fluorotyrosine shown in interior (in) vs exterior (ex) locations for the fluorine atom. For residues at the surface, the exterior orientation effectively places the fluorine into the solvent, while for more buried residues it simply denotes an "outward" vs and "inward" rotation. *In the case of an entirely interior residue, "exterior" implies the environment closest to the surface.

Considering the phenol's conformational effects on the ^{19}F chemical shift, there are two key observations. In the absence of other external groups, s-*cis* (H,F) conformers are slightly lower in energy than s-*trans* conformers, which reflects the expected favorable electrostatic interaction expected in the absence of alternative hydrogen-

bonding opportunities, e.g., with coordinating water molecules. There is a large difference in the chemical shifts for the two fluorine-hydrogen orientations: s-*trans* conformers have $\delta_{fit}^{s-trans} \approx$ -130 ppm, whereas the lower-energy s-*cis* conformers have $\delta_{fit}^{s-cis} \approx$ -150 ppm. A strong upfield shift is consistent with significantly higher nuclear shielding from a dipolar interaction with the hydroxyl group. This result is also consistent with computations and experimental results from Dalvit *et al*. who identify highly shielded fluorine nuclei in close proximity to hydrogen-bond donors.[235] If external hydrogen-bond acceptors for the hydroxyl group are present, however, this disparity in chemical shifts is substantially reduced. These various effects are manifest in Figure 6.4 for residue $3FY_{65}$, for which there are indeed four accessible conformers and for which two phenolic hydrogen-bond acceptors are observed, (1) external water, and (2) an interior peptide backbone carbonyl. In the case of $3FY_{65}$, as predictions converge with increasing cluster size, it is apparent that the most favorable conformer involves externally oriented fluorine, with the s-*trans* phenolic proton hydrogen-bonded to the interior peptide carbonyl. In effort to experimentally assess predicted conformer weighting PrOF NMR spectra of 3FY-Brd4 were acquired at 15 °C, 25 °C, and 35 °C to see if our model could accurately replicate the spectra at different temperatures. However, because the small changes in chemical shift from different temperatures (Avg. change = 0.13 ppm) are much lower than the current error in our method we were unable to draw conclusions from these experiments (Table 6.4).

Table 6.4 Temperature Effect on 3FY-Brd4 19F Chemical Shifts

Residue	δ at 25 °C (ppm)	Δδ at 15 °C (ppm)
Y97	-140.1	-0.26
Y118	-138.0	-0.11
Y65	-137.4	-0.21
Y98	-136.6	-0.06
Y139	-136.6	-0.06
Y137	-134.0	+0.05

Y119	-128.0	-0.16

6.3.4 Accurately Modeling the Phenolic Environment

Accurate modeling of the phenolic proton environment is a critical challenge for the accurate prediction of the 3FY 19F NMR, as a phenolic hydrogen-bond acceptor has a significant impact on the 19F chemical shift. Furthermore, inclusion of only the hydrogen-bond acceptor can be problematic if the acceptor itself is involved in additional strong interactions, e.g., a charged residue interacting with adjacent ionic residues. This happens in clusters where the target 3FY has a phenolic hydrogen bond to a negatively charged aspartate or glutamate. The anionic hydrogen-bond acceptor, in the absence of a positive counter-ion, over-delocalizes electronic density onto fluorine leading to an erroneous upfield shift. For example, a cluster of 3FY118 with a cutoff radius of 3.00 Å involves a hydrogen bond to a glutamate carboxylate and leads to a chemical shift prediction of -148.5 ppm. An increased cluster size, however, ultimately includes ion-pairing of the glutamate's carboxylate with the guanidinium group of arginine 113 (Figure 6.4), and the predicted chemical shift becomes -140.8 ppm (Table 6.7), i.e., substantially reduced in magnitude.

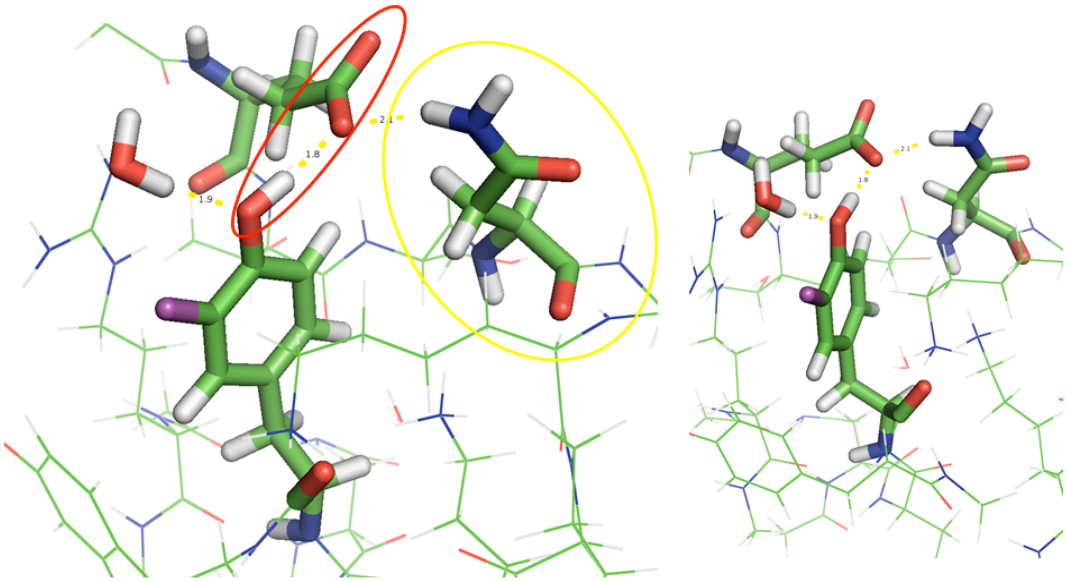


Figure 6.4 Brd4 (4IOR) Residue 118 with hydrogen bond to glutamate 49.

Note that within the right circle is glutamate 49's counter ion, arginine 113, which is only included in clusters with a cut off radius larger than 3.00 Å. The right pane shows a slightly rotated orientation to facilitate visualization.

6.3.5 Assessing Physical Contributions to Chemical Shifts

While it would be difficult to partition contributions from chemically intuitive sources such as van der Waals forces, electrostatic charges, and hydrogen bonds from the ^{19}F chemical shift in the full protein environment, we have performed a series of calculations designed better to assess them in appropriate model systems. In particular, we have predicted the ^{19}F chemical shift and fluorine Mulliken population in fluorobenzene as an argon atom, a sodium cation, a fluoride anion, and a water molecule are adjusted along the C-F axis over a range of lengths (with continuum aqueous solvation; Figure 6.5). These four probes interact predominantly through dispersion, positive charge, negative charge, and hydrogen bonding, respectively. Ar and F^- have a very similar effect on the chemical shift; a significant deshielding effect is predicted at smaller distances. The effect is as large as 21 ppm at a distance of 2.5 Å, but reduces to 2 ppm by 3.5 Å. The behavior with respect to Ar is consistent with ^1H deshielding observed in sterically compressed organic complexes.[236] We note that Ar does not affect the population density on fluorine. The effect of fluoride is further discussed below. Explicit hydrogen bonding from a water molecule to the fluorine atom results in a smaller increase in deshielding, ranging from 11 ppm at 2.5 Å (O–F distance, somewhat shorter than expected for a typical hydrogen bond) to 0 ppm at 3.5 Å. In evaluating hydrogen bonding effects on the fluorine chemical shift, Dalvit and Vulpetti found that fluorines participating in hydrogen bonds exhibit a range of shieldings but are typically more shielded than those in hydrophobic environments.[235] One difference in our model is that we have found that our models require explicit water to obtain the best match with experimental measurements. Although we do not observe a strong shielding effect nor a large accumulation charge on fluorine via water interactions, the net shielding and increased charge accumulation relative to fluoride or argon are consistent with the findings of Dalvit and Vulpetti. The sodium cation has the opposite effect on the ^{19}F chemical shift, significantly shielding the fluorine atom, with the effect ranging from 18 ppm at 2.5 Å to 2.5 ppm at 3.5 Å. These results do not show the same behavior exhibited

by alkali metal fluoride materials.[237] Population analysis of the fluorine atom's effective charge shows that the amount of electronic density localized onto the fluorine tracks with the trends in both fluoride and sodium. Table 6.5 shows that as the sodium cation gets closer, the amount of electrons on fluorine increases; the opposite trend is seen for fluoride. The increased (or decreased) shielding can be explained by an induced dipole moment on the fluorobenzene ring in response to the electrostatic charge getting closer.

Table 6.5 Mulliken population analysis.
Mulliken populations are taken from computations at the PBE0/SMD level of theory with EPR-II basis sets on C, H, F, and O atoms and def2-TZVPP on Ar and Na atoms.

Mulliken Population on Fluorine

Distance	Na^+	Ar	F^-	$O(H_2)$
2.50	-0.155	-0.166	-0.105	-0.194
2.75	-0.133	-0.172	-0.117	-0.194
3.00	-0.125	-0.172	-0.129	-0.186
3.25	-0.122	-0.171	-0.139	-0.178
3.50	-0.123	-0.1710	-0.146	-0.178

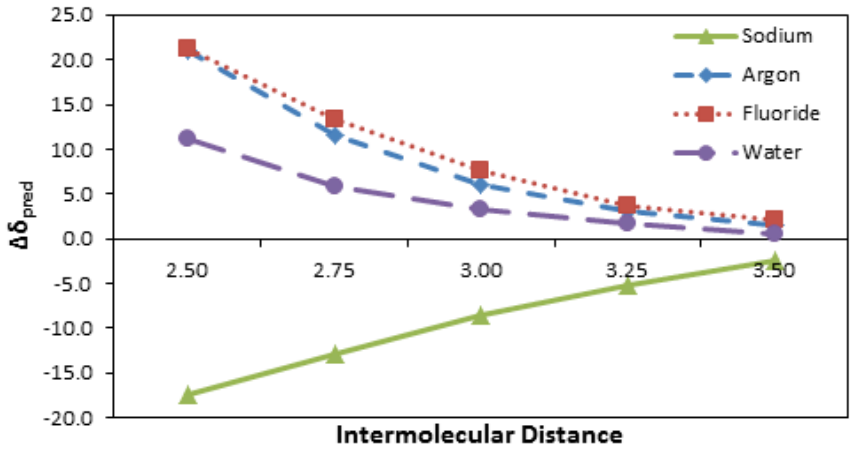


Figure 6.5 Changes in fluorobenzene 19F NMR chemical shift (ppm) vs distances between the fluorine atom and various probe atoms (O for water).

6.3.6 19F NMR Predictions for 3-fluorotyrosine mutant BRD4

Using a cutoff distance of 4.0 Å, we predicted the ^{19}F NMR spectrum for the entire fluorinated mutant protein (Table 6.6). First we considered optimizing only each target residue within a rigid surrounding framework. We found this protocol to be sufficient for a subset of residues, namely $3FY_{98}$ and $3FY_{118}$, but not to be sufficient for the entire protein. Closer examination of clusters with large discrepancies revealed that explicit water molecules directly interact with the 3FY phenol group in other instances. Even though the water positions are optimized during the classical molecular dynamics portion of the cluster generation protocol, the sensitivity of the ^{19}F chemical shifts in the training set to local solvation led us to hypothesize that further optimization of adjacent water molecule positions might improve our chemical shift predictions. Data when nearby water molecules are included in the partial geometry optimization step are shown in Table 6.6.

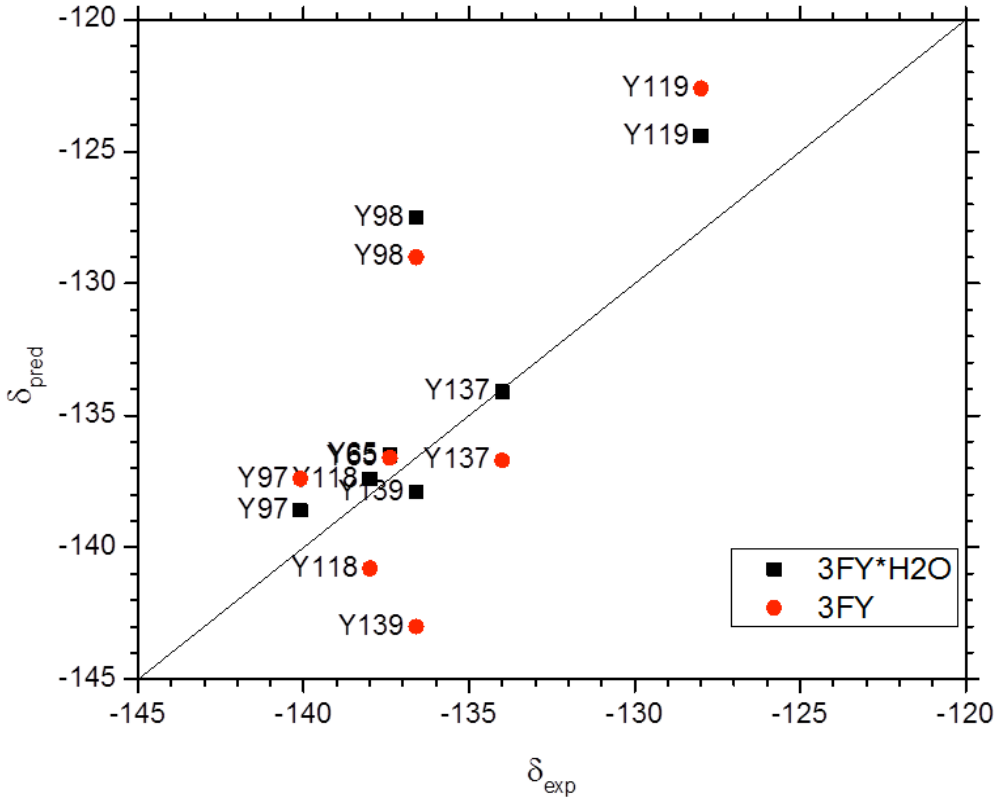


Figure 6.6 Performance of 19F δ for 3FY in Comparison to Experiment.
Points selected are the chemicals shifts of the best fitting shift to experiment.

Table 6.6: NMR predictions for 3FY residues in BRD4.[a]

Residues		ΔE(s-*cis*-in)	ΔE(s-*trans*-ex)	ΔE(s-*cis*-ex)	^{19}F NMR δ_{fit} (ppm)[d]					EXP
AA	#AA[b]	kcal/mol	kcal/mol	kcal/mol	s-*trans*-in	s-*cis*-in	s-*trans*-ex	s-*cis*-ex	δ_{tot}	
Y97	13	5.10	6.69	2.32	-137.4	-145.4	-126.5	-129.2	-137.3	**-140.1**
Y118	12	--	0.00[e]	4.39	--	--	-140.8	-140.6	-140.8	**-138.0**
Y65	14	-3.56	-4.97	-2.59	-121.0	-127.9	-132.4	-136.6	-132.1	**-137.4**
Y98	14	-0.13	9.52	10.15	-121.6	-120.1	-121.1	-129.0	-120.8	**-136.6**
Y139	10	-9.41	-13.28	-4.43	-119.0	-117.3	-143.0	-148.9	-142.9	**-136.6**
Y137	11	10.00	6.79	-3.49	-116.3	-119.4	-136.7	-140.8	-140.7	**-134.0**
Y119	14	11.18	--	-0.31	-119.4	-121.0	--	-122.6	-122.0	**-128.0**

[a] Only the fluorinated tyrosine residue is optimized for these models. [b] The number of residues (# AA) describes the number of residues kept with side chains intact (does not include caps). [c] The number of disconnected peptide chains (#PP) is denoted by the number of chains. [d] δ_{fit} is the Boltzmann weighted ^{19}F NMR chemical shift for 3-fluorotyrosine. [e] The isomeric environments are separated into interior (in.) vs exterior (ex.) and s-*cis* vs s-*trans*,energy differences are reported relative to the s-*trans*-in conformer except for $3FY_{118}$, for which no s-*trans*-in configuration exists, so the s-*trans*-ex configuration is taken as the zero of energy.

Table 6.7: Hydration effects on NMR predictions for 3FY residues in BRD4.[a]

Residues		$\Delta E(_{s\text{-}cis\text{-}in})$	$\Delta E(_{s\text{-}trans\text{-}ex})$	$\Delta E(_{s\text{-}cis\text{-}ex})$	^{19}F NMR δ_{fit} (ppm)[e]					EXP
AA	#H_2O[d]	kcal/mol	kcal/mol	kcal/mol	s-*trans*-in.	s-*cis*-in	*s*-*trans*-ex	s-*cis*-ex.	δ_{tot}	
Y97	3	-2.18	-0.29	6.65	-137.5	***138.6***	-133.7	-131.9	-138.3	**-140.1**
Y118	1	--	0.00[e]	4.97	--	--	***137.4***	-112.7	-137.4	**-138.0**
Y65	2	-6.32	-8.51	-2.85	-121.5	-129.7	*135.5*	**136.5**	-135.4	**-137.4**
Y98	1	-0.02	7.90	9.97	-120.4	*120.2*	-120.8	-127.5	-120.4	**-136.6**
Y139	2	-9.31	-16.65	-5.28	-116.6	-115.3	*125.2*	**137.9**	-125.2	**-136.6**
Y137	2	-23.08	-28.25	-3.49	-114.1	-115.1	*129.6*	**134.1**	-129.6	**-134.0**
Y119	1*	11.26	--	-1.81	-119.4	-121.2	--	***124.4***	-124.4	**-128.0**

[a] Only the fluorinated tyrosine residue and water molecules are optimized for these models. [b-e] See footnotes to Table 6.6.

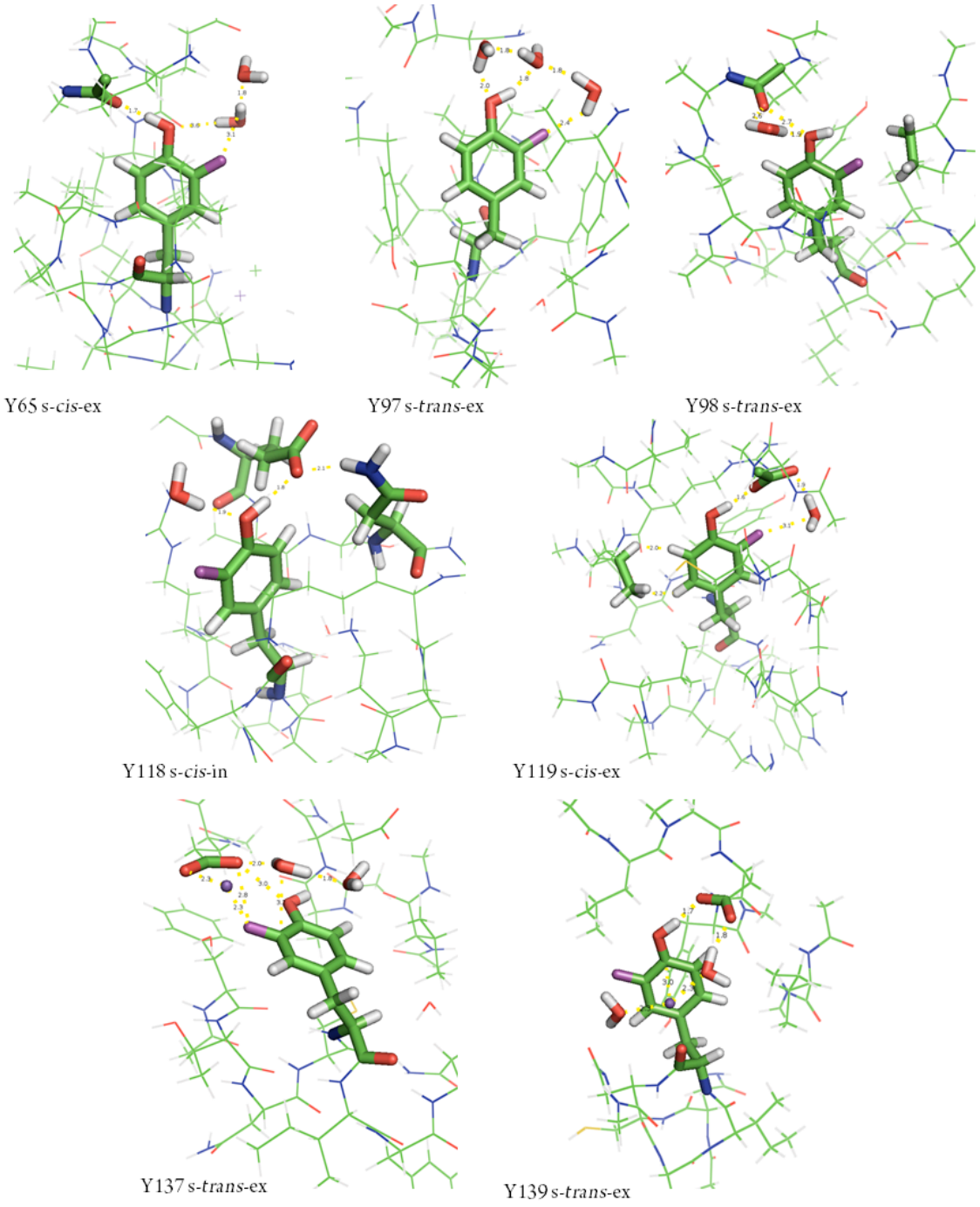


Figure 6.7 Lowest energy configurations of 3-fluorotyrosine clusters.
3FY65 shows the s-cis-ex, 3FY97 shows the s-trans-ex, 3FY98 shows the s-trans-ex, 3FY118 shows the s-cis-in, 3FY119 shows the s-cis-ex, 3FY137 shows the s-trans-ex, 3FY139 shows the s-trans-ex.

$3FY_{65}$ has two water molecules directly participating in hydrogen bonding interactions. We see that optimization of water has a 3 ppm shift on the low-energy s-*trans*-ex conformation (-132.5 vs -135.5 ppm), and dramatically improves agreement of the Boltzmann weighted δ_{tot} with experiment. We also note that for the solvent-optimized protocol, both exterior fluorine orientations predict quite chemical shifts quite similar to one another ($\delta_{fit}^{5\text{-}trans\text{-}ex}$ = -135.5 and $\delta_{fit}^{5\text{-}cis\text{-}ex}$ = -136.6 ppm) and to experiment (δ_{exp} = -137.4 ppm), while interior fluorine conformers are likely not to contribute ($\delta_{fit}^{5\text{-}trans\text{-}in}$ = -121.5 and $\delta_{fit}^{5\text{-}cis\text{-}in}$ = -129.7 ppm).

$3FY_{97}$ has the strongest upfield shift measured at δ_{exp} = -140.1 ppm. We note that this moiety is on the exterior of the protein and has three water molecules that directly participate in hydrogen bonding with the phenol and fluorine. The predicted chemical shift for the s-*cis*-in conformer (Figure 6.7) has the most upfield shift we predict at $\delta_{fit}^{5\text{-}cis\text{-}in}$ = -138.6 ppm and matches quite closely with experimental measurements.

$3FY_{98}$ has one water molecule participating in a hydrogen bond with the phenol. Models for $3FY_{98}$ exhibit different behavior than experimental measurements. For a cutoff radius of 4.0 Å, only one water molecule is found near the residue, compared to the protein x-ray crystal structure where a cluster of six waters is adjacent (only one being < 3.5 Å away, see Figure 6.8). Given the possibility of an extended hydrogen bonding network, further exploration was performed on $3FY_{98}$ with additional water moieties. Results showed negligible change from $H_2O{\cdot}3FY_{98}$. The phenoxide form of 3FY would provide an insufficient shift in neutral conditions with a pK_a over 8.4, which is not enough to account for the discrepancy in magnitude or population.[238] Given that an extended solvation sphere does not account for the difference, we postulate that the discrepancy could be due to a dynamic change in configuration that is not taken into account by our modeling protocol. Consistent with this hypothesis, we note that the carbonyl backbone hydrogen acceptor K102 is very flexible, as measured by its b-factor. Any structural perturbations on the carbonyl backbone would not be taken into account by this protocol.

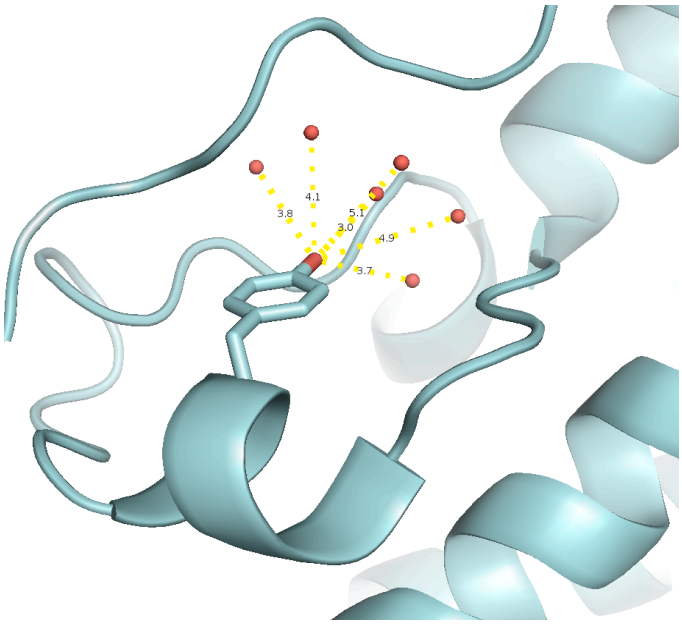


Figure 6.8 Hydration environment of Y98 in 4IOR x-ray structure of bromodomain Brd4.

$3FY_{118}$ has one water molecule directly participating in a hydrogen bond with the phenol. This hydrogen bond slightly reduces the upfield shift by accepting additional negative charge density. As the second further upfield shift measured at δ_{exp} = -138.0 ppm, we see qualitative and quantitative agreement for the chemical shift predicted at $\delta_{fit}^{s\text{-}trans\text{-}ex}$ = -137.4 ppm. We note that $3FY_{118}$ does not have an "interior" fluorine configuration as the interior fluorine clashes directly with carboxylate side chain resulting in an extremely high energy configuration.

$3FY_{119}$ is measured to have the furthest downfield chemical shift, and is largely shielded from the protein surface by two adjacent alpha helix coils. We find $3FY_{119}$ participates in a hydrogen bond with an adjacent carboxylate of aspartate 127. This carboxylate has an ion pair with arginine 122 and has one hydrogen bound water molecule (which does not participate in hydrogen bonding with the $3FY_{119}$). The s-*cis*-ex configuration is predicted to have a chemical shift of $\delta_{fit}^{s\text{-}cis\text{-}ex}$= -124.4 ppm, and with the exception of previously mentioned $3FY_{98}$, is the most downfield residue. We note that the

s-*trans*-ex configuration has no hydrogen acceptor for the phenolic proton, and with a low barrier for rotation of the phenol all initial guesses converge to s-*cis*-ex (Figure 6.7).

$3FY_{137}$ and $3FY_{139}$ are both solvent accessible residues on the exterior of the protein. We find that during molecular dynamic optimization, $3FY_{137}$ has one sodium ion and two waters directly interacting with the phenol. Given the close proximity of sodium to the residue, and the high mobility of sodium cations, sodium was also relaxed during solvent optimization. The s-*cis*-ex configuration is the only configuration that results in a chemical shift in the typical range for an external residue at -134.1 ppm as compared to -134.0 ppm. $3FY_{139}$ has two waters and a sodium atom near the residue with water participating in hydrogen bonding to the phenol. Similar behavior is noted here where the s-cis-ex configuration predicts a chemical shift at -137.9 as compared to a measured chemical shift at -136.6 ppm. We note for both $3FY_{137}$ and $3FY_{139}$, DFT optimization of solvent waters were critical to improve the accuracy of the NMR predictions with changes between 6 to 10 ppm.

The predicted most stable configurations for each residue are shown in Figure 6.7. If we examine this set of residues, we see that the s-*trans*-ex conformer is predicted to be the most stable conformer for all residues except $3FY_{97}$ and $3FY_{98}$, where the s-*cis*-in conformer is more stable. However, if we compare the predicted chemical shifts to their respective measured values, we notice that the s-*cis* isomers show a much smaller deviation than s-*trans* isomers for five out of seven residues. The residues where water does not act as the phenolic hydrogen bond acceptor ($3FY_{97}$, $3FY_{118}$, $3FY_{119}$) have the most reliable predictions compared to measured values (MSE 2 ppm). In applying this protocol to future systems, we note that the assignment of new resonances can be eased by eliminating conformers that have predicted chemical shifts well downfield of observed 3FY chemical shifts. Secondly, the resonances where the phenol directly interacts with a well conserved carboxylate residue are most reliable.

We then examined the effect that optimizing water has on the chemical shift of the fluorinated tyrosine. If $3FY_{98}$ is excluded, which shows not fully understood divergent behavior, the protocol for the best match exclusive target residue relaxation has a mean unsigned deviation (MUD) from experimental data of 3.5 ppm (Table 6.8). Including

optimized water networks with the target residue leads to improvements in the protocol when the residue is directly solvated by water. When explicit water solvent is optimized, the MUD for the same protocol improves to 1.3 ppm. This demonstrates that explicit solvent optimization dramatically improves predictions of the fluorine environment in the bromodomain containing-proteins. Future work should likely focus further on the accurate prediction of relative conformer energetics as they influence averaged chemical shifts.

Table 6.8 Performance of 19F δ for 3FY in Comparison to Experiment

	AA	**97**	**118**	**65**	**98**	**139**	**137**	**119**	**MSE**	**MUE**
	δ_{exp}	-140.1	-138	-137.4	-136.6	-136.6	-134	-128	All - 98	All - 98
Hydrated 3FY	δ_{fit} (Best)	-138.6	-137.4	-136.5	-127.5	-137.9	-134.1	-124.4		
	$\Delta\delta_{fit}$	1.5	0.6	0.9	9.1	-1.3	-0.1	3.6	0.9	1.3
	δ_{fit} (GS)	-138.6	-137.4	-135.5	-120.2	-125.2	-129.6	-124.4		
	$\Delta\delta_{fit}$	1.5	0.6	1.9	16.4	11.4	4.4	3.6	3.9	3.9
	δ_{tot}	-138.3	-137.4	-135.4	-120.4	-125.2	-129.6	-124.4		
	$\Delta\delta_{tot}$	1.8	0.6	2	16.2	11.4	4.4	3.6	4.0	4.0
Only 3FY	δ_{fit} (Best)	-137.4	-140.8	-136.6	-129	-143	-136.7	-122.6		
	$\Delta\delta_{fit}$	2.7	-2.8	0.8	7.6	-6.4	-2.7	5.4	-0.5	3.5
	δ_{fit} (GS)	-137.4	-140.8	-127.9	-120.1	-143	-140.8	-122.6		
	$\Delta\delta_{fit}$	2.7	-2.8	9.5	16.5	-6.4	-6.8	5.4	0.3	5.6
	δ_{tot}	-137.3	-140.8	-132.1	-120.8	-142.9	-140.7	-122		
	$\Delta\delta_{tot}$	2.8	-2.8	5.3	15.8	-6.3	-6.7	6	-0.3	5.0

6.4 **Conclusions**

This work has taken a first step toward automated prediction of ^{19}F NMR spectra for bromodomains. The machinery has been built and tested on Brd4 to take the protein crystal structure, extract clusters based on target residues and desired cluster sizes, and predict chemical shifts. We have shown that water plays a significant role in the prediction of the ^{19}F chemical shifts of 3FY residues, and that models must take account of water's influence on the tyrosine phenol group in order to accurately predict chemical

shifts. This makes 3FY much more challenging to model than 4-fluorophenylalanine labeled proteins, but can provide significant insight into the local structure of hydrogen bonding networks. Further work on accurately sampling the thermodynamically accessible ensemble of configurations, e.g., through sampling of molecular dynamics (MD) trajectories as evaluated by Lehtivarjo *et al*.,[213] could improve this first-generation protocol, albeit at a considerably higher computational cost associated with performing MD simulations. As a reasonable first approximation to this challenge of thermodynamic sampling, our first generation model uses simply the four conformers associated with the orientation of the phenolic proton and the fluorine atom. Further consideration of more sophisticated techniques for ensemble averaging is warranted in future studies. Based on results from BRD4, this method should be able to predict accurate chemical shifts for 3-fluorotyrosine residues where water does not directly participate as a hydrogen bonding partner; for microsolvated residues, further attention to alternative possibilities may be necessary. This protocol shows promise as a tool to facilitate assignment of challenging ^{19}F NMR spectra in labeled proteins.

6.5 **Future directions**

Continuation of this work will include increasing the accuracy of the predictions, as well as expansion to other amino acids. Initial work has been performed to predict the chemical shifts of 5-fluorotryptophan incorporated into Brd4 in a halgonase. Though convergence of the initial minimization step has proven challenging for Brd4, the lack of available rotomers could simplify the prediction process.

Labeled Protein Expression:

3FY-Brd4 was expressed based on established methods using E. coli DL39(DE3) + pRARE strains. To express the labeled protein, the secondary culture in LB media was grown until an O.D. at 600 nm of 0.6 was reached followed by harvesting. Cells were resuspended in defined media of Muchmore et al.[4] containing 3-fluorotyrosine (80 mg/L) in place of tyrosine. The resuspended E. coli were incubated at 37 °C while shaking for

1.5 h followed by cooling to 20 °C and media temperature equilibration for 30 min. Protein expression was induced with 1 mM IPTG overnight (14−16 h) at 20 °C. The cells were harvested and stored at −20 °C. The cell pellet was thawed at room temperature followed by the addition of lysis buffer (50 mM Phosphate pH 7.4, 300 mM NaCl) containing protease inhibitor PMSF (5 mM) as well as the Halt protease inhibitor. Cells were lysed by sonication and the cell lysate was centrifuged at 7500 g for 30 minutes followed by supernatant filtration through Whatman filter paper. Filtrate containing the histidine-tagged Brd4 was loaded on to a nickel-NTA affinity column and eluted with an imidazole gradient on an AKTA FPLC system monitoring the absorbance at 280 nm. Imidazole was removed from the buffer using a HiPrep column (GE) for buffer exchange into 50 mM Tris pH 7.5, 100 mM NaCl. Purified and buffer exchanged protein was treated with TEV protease at 4 °C overnight on a rotating carrousel. The cleaved His-tag, TEV protease and uncleaved Brd4 were removed using nickel-NTA affinity resin.

1D ^{19}F NMR Parameters:

^{19}F NMR spectra were acquired at 470 MHz on a Bruker 500 spectrometer with a 5 mm Prodigy TCI Cryoprobe without proton decoupling. Samples containing 60 μM 3FY-Brd4 in 50 mM TRIS, 100 mM NaCl, 10% D2O, pH 7.5 were used for NMR experiments. Measurement parameters include a relaxation delay of 0.3 s for 3FY-Brd4 containing a 30° flip angle, an acquisition time of 0.1 s, and a spectral width of 18 ppm. A 20 Hz line-broadening was applied after 3000 transients.

Discussion of effective dispersion and relative line width of ^{19}F NMR spectra:

It is worth discussing the effective dispersion associated with the experimental chemical shift measurements based on the line width of broad protein resonances. The linewidths of 3FY-Brd4 are indeed broad and vary between 0.09 ppm and 0.14 ppm at half-height. The median error in the chemical shift prediction is 1.3 ppm, so differences due to linewidth should be a minimal source of error at this point. Despite the larger linewidth the chemical shift of ^{19}F resonances is very reproducible with a standard deviation of 0.03 ppm or less being common.[5] Based on the reproducibility in chemical

shifts, the average resonance line-width, and the average error in the theoretical measurements, as a conservative estimate we anticipate the uncertainties would be minor and only contribute ~11% error in the event that accurate chemical shifts were difficult to measure.

Variable Temperature Chemical Shifts:

To assess the effect of temperature on chemical shift, additional ^{19}F NMR spectra of 3FY-Brd4 were acquired at 15 °C and 35 °C. At 35 °C we observed substantial signal reduction, indicating reduced stability of the protein. However, we were able to acquire an NMR spectrum at 15 °C. In most instances the difference in chemical shift between 15 °C and 25 °C is less than the linewidth at half-height of the resonances. To eliminate variable chemical shifts from trifluoroacetate at different temperatures, the spectra provided for the variable temperature experiments are unreferenced.

Chapter 7. References

(1) Feinberg, A. P. Phenotypic plasticity and the epigenetics of human disease. *Nature* **2007**, *447*, 433–440.

(2) Einav Nili, G.-Y.; Saito, Y.; Egger, G.; Jones, P. a Cancer Epigenetics: Modifications, Screening, and Therapy. *Annu. Rev. Med.* **2008**, *59*, 267–280.

(3) Baccarelli, A.; Rienstra, M.; Benjamin, E. J. Cardiovascular Epigenetics: Basic Concepts and Results From Animal and Human Studies. *Circ. Cardiovasc. Genet.* **2010**, *3*, 567–573.

(4) Bayarsaihan, D. Epigenetic Mechanisms in Inflammation. *J. Dent. Res.* **2011**, *90*, 9–17.

(5) Muller, H. J. Further Studies on the Nature and Causes of Gene Mutations. *Proc. Int. Congr. Genet.* **1932**, *1*, 213–255.

(6) Passarge, E. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am. J. Hum. Genet.* **1979**, *31*, 106–15.

(7) Sadakierska-Chudy, A.; Filip, M. A Comprehensive View of the Epigenetic Landscape. Part II: Histone Post-translational Modification, Nucleosome Level, and Chromatin Regulation by ncRNAs. *Neurotox. Res.* **2015**, *27*, 172–197.

(8) Strahl, B. D.; Allis, C. D. The language of covalent histone modifications. *Nature* **2000**, *403*, 41–45.

(9) Turner, B. M. Histone acetylation and an epigenetic code. *BioEssays* **2000**, *22*, 836–845.

(10) Rosen, C.; Dorsett, D.; Jack, J. A proline-rich region in the Zeste protein essential for transvection and white repression by Zeste. *Genetics* **1998**, *148*, 1865–74.

(11) Jenuwein, T. Translating the Histone Code. *Science* **2001**, *293*, 1074–1080.

(12) 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*, 214–231.

(13) Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E. M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S.; McKeown, M. R.; Wang, Y.; Christie, A. L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. a; Wiest, O.; Kung, A. L.; Knapp, S.; Bradner, J. E. Selective inhibition of BET bromodomains. *Nature* **2010**, *468*, 1067–1073.

(14) Liu, L.; Zhen, X. T.; Denton, E.; Marsden, B. D.; Schapira, M. ChromoHub: A data hub for navigators of chromatin-mediated signalling. *Bioinformatics* **2012**, *28*, 2205–2206.

(15) Belkina, A. C.; Denis, G. V BET domain co-regulators in obesity, inflammation and cancer. *Nat. Rev. Cancer* **2012**, *12*, 465–477.

(16) Darnell, J. E. Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2002**, *2*, 740–749.

(17) Nicodeme, E.; Jeffrey, K. L.; Schaefer, U.; Beinke, S.; Dewell, S.; Chung, C.-W.; Chandwani, R.; Marazzi, I.; Wilson, P.; Coste, H.; White, J.; Kirilovsky, J.; Rice, C. M.; Lora, J. M.; Prinjha, R. K.; Lee, K.; Tarakhovsky, A. Suppression of inflammation by a synthetic histone mimic. *Nature* **2010**, *468*, 1119–1123.

(18) Zhang, G.; Smith, S. G.; Zhou, M. M. Discovery of Chemical Inhibitors of Human Bromodomains. *Chem. Rev.* **2015**, *115*, 11625–11668.

(19) 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*, 7346–7359.

(20) 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.

(21) Owen, D. J. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *EMBO J.* **2000**, *19*, 6141–6149.

(22) Zeng, L.; Li, J.; Muller, M.; Yan, S.; Mujtaba, S.; Pan, C.; Wang, Z.; Zhou, M.-M.

Selective small molecules blocking HIV-1 Tat and coactivator PCAF association. *J. Am. Chem. Soc.* **2005**, *127*, 2376–7.

(23) Morinière, J.; Rousseaux, S.; Steuerwald, U.; Soler-López, M.; Curtet, S.; Vitte, A.-L.; Govin, J.; Gaucher, J.; Sadoul, K.; Hart, D. J.; Krijgsveld, J.; Khochbin, S.; Müller, C. W.; Petosa, C. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **2009**, *461*, 664–668.

(24) Jiang, Y. W.; Veschambre, P.; Erdjument-Bromage, H.; Tempst, P.; Conaway, J. W.; Conaway, R. C.; Kornberg, R. D. Mammalian mediator of transcriptional regulation and its possible role as an end-point of signal transduction pathways. *Pnas* **1998**, *95*, 8538–8543.

(25) Crawford, N. P. S.; Alsarraj, J.; Lukes, L.; Walker, R. C.; Officewala, J. S.; Yang, H. H.; Lee, M. P.; Ozato, K.; Hunter, K. W. Bromodomain 4 activation predicts breast cancer survival. *Proc. Natl. Acad. Sci.* **2008**, *105*, 6380–6385.

(26) French, C. A. NUT midline carcinoma. *Cancer Genet. Cytogenet.* **2010**, *203*, 16–20.

(27) Dawson, M. a; Prinjha, R. K.; Dittmann, A.; Giotopoulos, G.; Bantscheff, M.; Chan, W.-I.; Robson, S. C.; Chung, C.; Hopf, C.; Savitski, M. M.; Huthmacher, C.; Gudgin, E.; Lugo, D.; Beinke, S.; Chapman, T. D.; Roberts, E. J.; Soden, P. E.; Auger, K. R.; Mirguet, O.; Doehner, K.; Delwel, R.; Burnett, A. K.; Jeffrey, P.; Drewes, G.; Lee, K.; Huntly, B. J. P.; Kouzarides, T. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **2011**, *478*, 529–33.

(28) Martin, M. P.; Olesen, S. H.; Georg, G. I.; Schönbrunn, E. Cyclin-dependent kinase inhibitor dinaciclib interacts with the acetyl-lysine recognition site of bromodomains. *ACS Chem. Biol.* **2013**, *8*, 2360–2365.

(29) Ember, S. W. J.; Zhu, J. Y.; Olesen, S. H.; Martin, M. P.; Becker, A.; Berndt, N.; Georg, G. I.; Schonbrunn, E. Acetyl-lysine binding site of bromodomain-containing protein 4 (BRD4) interacts with diverse kinase inhibitors. *ACS Chem. Biol.* **2014**, *9*, 1160–1171.

(30) Ciceri, P.; Müller, S.; O'Mahony, A.; Fedorov, O.; Filippakopoulos, P.; Hunt, J.

P.; Lasater, E. a; Pallares, G.; Picaud, S.; Wells, C.; Martin, S.; Wodicka, L. M.; Shah, N. P.; Treiber, D. K.; Knapp, S. Dual kinase-bromodomain inhibitors for rationally designed polypharmacology. *Nat. Chem. Biol.* **2014**, *10*, 305–312.

(31) Bamborough, P.; Diallo, H.; Goodacre, J. D.; Gordon, L.; Lewis, A.; Seal, J. T.; Wilson, D. M.; Woodrow, M. D.; Chung, C. W. Fragment-based discovery of bromodomain inhibitors part 2: Optimization of phenylisoxazole sulfonamides. *J. Med. Chem.* **2012**, *55*, 587–596.

(32) 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*, 2246–2256.

(33) Chung, C.-W.; Dean, A. W.; Woolven, J. M.; Bamborough, P. Fragment-Based Discovery of Bromodomain Inhibitors Part 1: Inhibitor Binding Modes and Implications for Lead Discovery. *J. Med. Chem.* **2012**, *55*, 576–586.

(34) Filippakopoulos, P.; Knapp, S. The bromodomain interaction module. *FEBS Lett.* **2012**, *586*, 2692–2704.

(35) Vollmuth, F.; Geyer, M. Interaction of Propionylated and Butyrylated Histone H3 Lysine Marks with Brd4 Bromodomains. *Angew. Chemie Int. Ed.* **2010**, *49*, 6768–6772.

(36) Badenhorst, P.; Xiao, H.; Cherbas, L.; Kwon, S. Y.; Voas, M.; Rebay, I.; Cherbas, P.; Wu, C. The Drosophila nucleosome remodeling factor NURF is required for Ecdysteroid signaling and metamorphosis. *Genes Dev.* **2005**, 2540–2545.

(37) Barak, O. Isolation of human NURF: a regulator of Engrailed gene expression. *EMBO J.* **2003**, *22*, 6089–6100.

(38) 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*, e9657.

(39) Dar, A. A.; Nosrati, M.; Bezrookove, V.; de Semir, D.; Majid, S.; Thummala, S.;

Sun, V.; Tong, S.; Leong, S. P. L.; Minor, D.; Billings, P. R.; Soroceanu, L.; Debs, R.; Miller, J. R.; Sagebiel, R. W.; Kashani-Sabet, M. The Role of BPTF in Melanoma Progression and in Response to BRAF-Targeted Therapy. *JNCI J. Natl. Cancer Inst.* **2015**, *107*, djv034-djv034.

(40) Jordan-Sciutto, K. L.; Dragich, J. M.; Caltagarone, J.; Hall, D. J.; Bowser, R. Fetal Alz-50 Clone 1 (FAC1) Protein Interacts with the Myc-Associated Zinc Finger Protein (ZF87/MAZ) and Alters Its Transcriptional Activity †. *Biochemistry* **2000**, *39*, 3206–3215.

(41) 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*, 425–427.

(42) Richart, L.; Carrillo-de Santa Pau, E.; Río-Machín, A.; de Andrés, M. P.; Cigudosa, J. C.; Lobo, V. J. S.-A.; Real, F. X. BPTF is required for c-MYC transcriptional activity and in vivo tumorigenesis. *Nat. Commun.* **2016**, *7*, 10153.

(43) Kwon, S. Y.; Xiao, H.; Wu, C.; Badenhorst, P. Alternative Splicing of NURF301 Generates Distinct NURF Chromatin Remodeling Complexes with Altered Modified Histone Binding Specificities. *PLoS Genet.* **2009**, *5*, e1000574.

(44) 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*, 692–706.

(45) Landry, J.; Sharov, A. a; Piao, Y.; Sharova, L. V; Xiao, H.; Southon, E.; Matta, J.; Tessarollo, L.; Zhang, Y. E.; Ko, M. S. H.; 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*, e1000241.

(46) 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*, 6959–6964.

(47) Clackson, T.; Wells, J. A.; Wellst, J. A. A Hot Spot of Binding Energy in a Hormone-Receptor Interface. **2012**, *267*, 383–386.

(48) London, N.; Movshovitz-Attias, D.; Schueler-Furman, O. The Structural Basis of Peptide-Protein Binding Strategies. *Structure* **2010**, *18*, 188–199.

(49) Bullock, B. N.; Jochim, A. L.; Arora, P. S. Assessing Helical Protein Interfaces for Inhibitor Design. *J. Am. Chem. Soc.* **2011**, *133*, 14220–14223.

(50) Vassilev, L. T. In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science* **2004**, *303*, 844–848.

(51) Mullard, A. Pioneering apoptosis-targeted cancer drug poised for FDA approval. *Nat. Rev. Drug Discov.* **2016**, *15*, 147–149.

(52) Chung, C. -w.; Witherington, J. Progress in the Discovery of Small-Molecule Inhibitors of Bromodomain-Histone Interactions. *J. Biomol. Screen.* **2011**, *16*, 1170–1185.

(53) Structural Genomics Consortium (SGC): Bromosporine. http://www.thesgc.org/?chemical-probes/?bromosporine.

(54) Wells, J. a; McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* **2007**, *450*, 1001–1009.

(55) Bollag, G.; Tsai, J.; Zhang, J.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P. Vemurafenib: the first drug approved for BRAF-mutant cancer. *Nat. Rev. Drug Discov.* **2012**, *11*, 873–886.

(56) Souers, A. J.; Leverson, J. D.; Boghaert, E. R.; Ackler, S. L.; Catron, N. D.; Chen, J.; Dayton, B. D.; Ding, H.; Enschede, S. H.; Fairbrother, W. J.; Huang, D. C. S.; Hymowitz, S. G.; Jin, S.; Khaw, S. L.; Kovar, P. J.; Lam, L. T.; Lee, J.; Maecker, H. L.; Marsh, K. C.; Mason, K. D.; Mitten, M. J.; Nimmer, P. M.; Oleksijew, A.; Park, C. H.; Park, C.-M.; Phillips, D. C.; Roberts, A. W.; Sampath, D.; Seymour, J. F.; Smith, M. L.; Sullivan, G. M.; Tahir, S. K.; Tse, C.; Wendt, M. D.; Xiao, Y.; Xue, J. C.; Zhang, H.; Humerickhouse, R. a; Rosenberg, S. H.; Elmore, S. W. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat. Med.* **2013**, *19*, 202–8.

(57) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **2016**.

(58) Reymond, J.-L.; van Deursen, R.; Blum, L. C.; Ruddigkeit, L. Chemical space as a

source for new drugs. *Med. Chem. Commun.* **2010**, *1*, 30.

(59) Fink, T.; Bruggesser, H.; Reymond, J.-L. Virtual Exploration of the Small-Molecule Chemical Universe below 160 Daltons. *Angew. Chemie Int. Ed.* **2005**, *44*, 1504–1508.

(60) Bohacek, R. S.; McMartin, C.; Guida, W. C. The art and practice of structure-based drug design: A molecular modeling perspective. *Med. Res. Rev.* **1996**, *16*, 3–50.

(61) Dunn, D. Ultra-High Throughput Screen of Two-Million-Member Combinatorial Compound Collection in a Miniaturized, 1536-Well Assay Format. *J. Biomol. Screen.* **2000**, *5*, 177–187.

(62) Barker, A.; Kettle, J. G.; Nowak, T.; Pease, J. E. Expanding medicinal chemistry space. *Drug Discov. Today* **2013**, *18*, 298–304.

(63) Hann, M. M.; Leach, A. R.; Harper, G. Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864.

(64) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A "Rule of Three" for fragment-based lead discovery? *Drug Discov. Today* **2003**, *8*, 876–877.

(65) Carr, R.; Jhoti, H. Structure-based screening of low-affinity compounds. *Drug Discov. Today* **2002**, *7*, 522–7.

(66) Leach, A. R.; Hann, M. M.; Burrows, J. N.; Griffen, E. J. Fragment screening: an introduction. *Mol. Biosyst.* **2006**, *2*, 429.

(67) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996**, *274*, 1531–1534.

(68) Harner, M. J.; Chauder, B. A.; Phan, J.; Fesik, S. W. Fragment-Based Screening of the Bromodomain of ATAD2. *J. Med. Chem.* **2014**, *57*, 9687–9692.

(69) Dalvit, C. NMR methods in fragment screening: theory and a comparison with other biophysical techniques. *Drug Discov. Today* **2009**, *14*, 1051–1057.

(70) Dalvit, C.; Fagerness, P. E.; Hadden, D. T. a; Sarver, R. W.; Stockman, B. J. Fluorine-NMR Experiments for High-Throughput Screening: Theoretical Aspects, Practical Considerations, and Range of Applicability. *J. Am. Chem. Soc.* **2003**,

125, 7696–7703.

(71) Jordan, J. B.; Whittington, D. A.; Bartberger, M. D.; Sickmier, E. A.; Chen, K.; Cheng, Y.; Judd, T. A Fragment Linking Approach Using 19F NMR Spectroscopy to Obtain Highly Potent and Selective Inhibitors of β-secretase. *J. Med. Chem.* **2016**, acs.jmedchem.5b01917.

(72) Antanasijevic, A.; Ramirez, B.; Caffrey, M. Comparison of the sensitivities of WaterLOGSY and saturation transfer difference NMR experiments. *J. Biomol. NMR* **2014**, *60*, 37–44.

(73) Cala, O.; Krimm, I. Ligand-Orientation Based Fragment Selection in STD NMR Screening. *J. Med. Chem.* **2015**, *58*, 8739–8742.

(74) Renaud, J.-P.; Chung, C.; Danielson, U. H.; Egner, U.; Hennig, M.; Hubbard, R. E.; Nar, H. Biophysics in drug discovery: impact, challenges and opportunities. *Nat. Rev. Drug Discov.* **2016**, *15*, 679–698.

(75) Scapin, G.; Yang, X.; Prosise, W. W.; McCoy, M.; Reichert, P.; Johnston, J. M.; Kashi, R. S.; Strickland, C. Structure of full-length human anti-PD1 therapeutic IgG4 antibody pembrolizumab. *Nat. Struct. Mol. Biol.* **2015**, *22*, 953–958.

(76) Giannetti, A. M. From Experimental Design to Validated Hits. In *Methods in Enzymology*; Elsevier Inc., 2011; Vol. 493, pp. 169–218.

(77) Davies, Thomas G.; Hyvönen, M. *Fragment-Based Drug Discovery and X-Ray Crystallography*; Davies, T. G.; Hyvönen, M., Eds.; Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 2012; Vol. 317.

(78) Shi, S.; Semple, A.; Cheung, J.; Shameem, M. DSF method optimization and its application in predicting protein thermal aggregation kinetics. *J. Pharm. Sci.* **2013**, *102*, 2471–2483.

(79) Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2007**, *2*, 2212–2221.

(80) Seidel, S. a I.; Wienken, C. J.; Geissler, S.; Jerabek-Willemsen, M.; Duhr, S.; Reiter, A.; Trauner, D.; Braun, D.; Baaske, P. Label-free microscale thermophoresis discriminates sites and affinity of protein-ligand binding. *Angew.*

Chemie - Int. Ed. **2012**, *51*, 10656–10659.

(81) Seidel, S. A. I.; Dijkman, P. M.; Lea, W. a.; van den Bogaart, G.; Jerabek-Willemsen, M.; Lazic, A.; Joseph, J. S.; Srinivasan, P.; Baaske, P.; Simeonov, A.; Katritch, I.; Melo, F. a.; Ladbury, J. E.; Schreiber, G.; Watts, A.; Braun, D.; Duhr, S. Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods* **2013**, *59*, 301–315.

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

(83) Gribble, G. W. The diversity of naturally produced organohalogens. *Chemosphere* **2003**, *52*, 289–297.

(84) O'Hagan, D.; B. Harper, D. Fluorine-containing natural products. *J. Fluor. Chem.* **1999**, *100*, 127–133.

(85) Gerig, J. T. Fluorine NMR (http://www.biophysics.org/Portals/1/PDFs/Education/gerig.pdf). *Previously Publ. Biophys. Textb. Online* **2001**, 1–35.

(86) 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.

(87) Sykes, B. D.; Weingarten, H. I.; Schlesinger, M. J. Fluorotyrosine Alkaline Phosphatase from Escherichia coli: Preparation, Properties, and Fluorine-19 Nuclear Magnetic Resonance Spectrum. *Proc. Natl. Acad. Sci.* **1974**, *71*, 469–473.

(88) Kim, H. W.; Perez, J. a; Ferguson, S. J.; Campbell, I. D. The specific incorporation of labelled aromatic amino acids into proteins through growth of bacteria in the presence of glyphosate. Application to fluorotryptophan labelling to the H(+)-ATPase of Escherichia coli and NMR studies. *FEBS Lett.* **1990**, *272*, 34–6.

(89) Liu, J. J.; Horst, R.; Katritch, V.; Stevens, R. C.; Wuthrich, K. Biased Signaling Pathways in 2-Adrenergic Receptor Characterized by 19F-NMR. *Science* **2012**, *335*, 1106–1110.

(90) Danielson, M. A.; Falke, J. J. Use of 19F NMR to probe protein structure and conformational changes. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 163–195.

(91) Peng, J. W. Cross-correlated (19)F relaxation measurements for the study of

fluorinated ligand-receptor interactions. *J. Magn. Reson.* **2001**, *153*, 32–47.

(92) Ropson, I. J.; Frieden, C. Dynamic NMR spectral analysis and protein folding: identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by 19F NMR. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 7222–6.

(93) Horng, J.-C.; Raleigh, D. P. Φ-Values beyond the Ribosomally Encoded Amino Acids: Kinetic and Thermodynamic Consequences of Incorporating Trifluoromethyl Amino Acids in a Globular Protein. *J. Am. Chem. Soc.* **2003**, *125*, 9286–9287.

(94) B, R. N. M. R. S. I.; Enzymol, D. E. M.; Berg, V. Den; Lemaster, D. M.; Richards, F. M. NMR Sequential Assignment of Escherichia coli Thioredoxin Utilizing Random Fractional Deuteration. *Biochemistry* **1988**, *27*, 142–150.

(95) Frieden, C.; Hoeltzli, S. D.; Bann, J. G. The Preparation of 19F-Labeled Proteins for NMR Studies. In *Methods in enzymology*; 2004; Vol. 380, pp. 400–415.

(96) Crowley, P. B.; Kyne, C.; Monteith, W. B. Simple and inexpensive incorporation of 19F-Tryptophan for protein NMR spectroscopy. *Chem. Commun.* **2012**, *48*, 10681.

(97) Cobb, S. L.; Murphy, C. D. 19F NMR applications in chemical biology. *J. Fluor. Chem.* **2009**, *130*, 132–143.

(98) Hoeltzli, S. D.; Frieden, C. Stopped-flow NMR spectroscopy: real-time unfolding studies of 6-19F-tryptophan-labeled Escherichia coli dihydrofolate reductase. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 9318–22.

(99) Pomerantz, W. C.; Wang, N.; Lipinski, A. K.; Wang, R.; Cierpicki, T.; Mapp, A. K. Profiling the Dynamic Interfaces of Fluorinated Transcription Complexes for Ligand Discovery and Characterization. *ACS Chem. Biol.* **2012**, *7*, 1345–1350.

(100) Bogan, A. A.; Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**, *280*, 1–9.

(101) Rossi, A. M.; Taylor, C. W. Analysis of protein-ligand interactions by fluorescence polarization. *Nat. Protoc.* **2011**, *6*, 365–387.

(102) Xinyi Huang Fluorescence Polarization Competition Assay: The Range of Resolvable Inhibitor Potency Is Limited by the Affinity of the Fluorescent Ligand.

J. Biomol. Screen. **2003**, *8*, 34–38.

(103) Freyer, M. W.; Lewis, E. a Isothermal Titration Calorimetry: Experimental Design, Data Analysis, and Probing Macromolecule/Ligand Binding and Kinetic Interactions. In *Methods in cell biology*; 2008; Vol. 84, pp. 79–113.

(104) Borah, J. C.; Mujtaba, S.; Karakikes, I.; Zeng, L.; Muller, M.; Patel, J.; Moshkina, N.; Morohashi, K.; Zhang, W.; Gerona-Navarro, G.; Hajjar, R. J.; Zhou, M. M. A small molecule binding to the coactivator CREB-binding protein blocks apoptosis in cardiomyocytes. *Chem. Biol.* **2011**, *18*, 531–541.

(105) Stockman, B. J.; Dalvit, C. NMR screening techniques in drug discovery and drug design. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *41*, 187–231.

(106) 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.

(107) Matzuk, M. M.; McKeown, M. R.; Filippakopoulos, P.; Li, Q.; Ma, L.; Agno, J. E.; Lemieux, M. E.; Picaud, S.; Yu, R. N.; Qi, J.; Knapp, S.; Bradner, J. E. Small-molecule inhibition of BRDT for male contraception. *Cell* **2012**, *150*, 673–684.

(108) Rule, G. S.; Pratt, E. A.; Simplaceanu, V.; Ho, C. Nuclear magnetic resonance and molecular genetic studies of the membrane-bound D-lactate dehydrogenase of Escherichia coli. *Biochemistry* **1987**, *26*, 549–56.

(109) 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 Chromatin* **2013**, *6*, 34.

(110) Muchmore, D. C.; McIntosh, L. P.; Russell, C. B.; Anderson, D. E.; Dahlquist, F. W. *Nuclear Magnetic Resonance Part B Structure and Mechanism*; Methods in Enzymology; Elsevier, 1989; Vol. 177.

(111) Lepre, C. A. Practical Aspects of NMR-Based Fragment Screening. In *Methods in Enzymology*; Elsevier Inc., 2011; Vol. 493, pp. 219–239.

(112) Dalvit, C.; Fagerness, P. E.; Hadden, D. T. A.; Sarver, R. W.; Stockman, B. J. Fluorine-NMR experiments for high-throughput screening: Theoretical aspects, practical considerations, and range of applicability. *J. Am. Chem. Soc.* **2003**, *125*,

7696–7703.

(113) Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Curr. Top. Med. Chem.* **2003**, *3*, 39–53.

(114) Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

(115) Tanaka, D.; Tsuda, Y.; Shiyama, T.; Nishimura, T.; Chiyo, N.; Tominaga, Y.; Sawada, N.; Mimoto, T.; Kusunose, N. A Practical Use of Ligand Efficiency Indices Out of the Fragment-Based Approach: Ligand Efficiency-Guided Lead Identification of Soluble Epoxide Hydrolase Inhibitors. *J. Med. Chem.* **2011**, *54*, 851–857.

(116) Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discov.* **2014**, *13*, 105–121.

(117) Erlanson, D. a; Wells, J. a; Braisted, A. C. Tethering: Fragment-Based Drug Discovery. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 199–223.

(118) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* **2007**, *6*, 211–219.

(119) Arntson, K. E.; Pomerantz, W. C. K. Protein-Observed Fluorine NMR: A Bioorthogonal Approach for Small Molecule Discovery. *J. Med. Chem.* **2016**, *59*, 5158–5171.

(120) Sharaf, N. G.; Gronenborn, A. M. 19F-Modified Proteins and 19F-Containing Ligands as Tools in Solution NMR Studies of Protein Interactions. In *Methods in Enzymology*; Elsevier Inc., 2015; Vol. 565, pp. 67–95.

(121) Wu, B.; Zhang, Z.; Noberini, R.; Barile, E.; Giulianotti, M.; Pinilla, C.; Houghten, R. A.; Pasquale, E. B.; Pellecchia, M. HTS by NMR of Combinatorial Libraries: A Fragment-Based Approach to Ligand Discovery. *Chem. Biol.* **2013**, *20*, 19–33.

(122) Doak, B. C.; Morton, C. J.; Simpson, J. S.; Scanlon, M. J. Design and evaluation of the performance of an NMR screening fragment library. *Aust. J. Chem.* **2013**, *66*, 1465–1472.

(123) Siegal, G.; AB, E.; Schultz, J. Integration of fragment screening and library design.

Drug Discov. Today **2007**, *12*, 1032–1039.

(124) Lim, S. S.; Debono, C. O.; MacRaild, C. A.; Chandrashekaran, I. R.; Dolezal, O.; Anders, R. F.; Simpson, J. S.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Development of inhibitors of Plasmodium falciparum apical membrane antigen 1 based on fragment screening. *Aust. J. Chem.* **2013**, *66*, 1530–1536.

(125) Vom, A.; Headey, S.; Wang, G.; Capuano, B.; Yuriev, E.; Scanlon, M. J.; Simpson, J. S. Detection and prevention of aggregation-based false positives in STD-NMR-based fragment screening. *Aust. J. Chem.* **2013**, *66*, 1518–1524.

(126) Dias, D. M.; Van Molle, I.; Baud, M. G. J.; Galdeano, C.; Geraldes, C. F. G. C.; Ciulli, A. Is NMR Fragment Screening Fine-Tuned to Assess Druggability of Protein–Protein Interactions? *ACS Med. Chem. Lett.* **2014**, *5*, 23–28.

(127) Gee, C. T.; Koleski, E. J.; Pomerantz, W. C. K. Fragment Screening and Druggability Assessment for the CBP/p300 KIX Domain through Protein-Observed 19F NMR Spectroscopy. *Angew. Chemie Int. Ed.* **2015**, *54*, 3735–3739.

(128) Leung, E. W. W.; Yagi, H.; Harjani, J. R.; Mulcair, M. D.; Scanlon, M. J.; Baell, J. B.; Norton, R. S. 19 F NMR as a Probe of Ligand Interactions with the iNOS Binding site of SPRY Domain-Containing SOCS Box Protein 2. *Chem. Biol. Drug Des.* **2014**, *84*, 616–625.

(129) Ge, X.; MacRaild, C. A.; Devine, S. M.; Debono, C. O.; Wang, G.; Scammells, P. J.; Scanlon, M. J.; Anders, R. F.; Foley, M.; Norton, R. S. Ligand-Induced Conformational Change of Plasmodium falciparum AMA1 Detected Using 19 F NMR. *J. Med. Chem.* **2014**, *57*, 6419–6427.

(130) Curtis-Marof, R.; Doko, D.; Rowe, M. L.; Richards, K. L.; Williamson, R. A.; Howard, M. J. 19F NMR spectroscopy monitors ligand binding to recombinantly fluorine-labelled b′x from human protein disulphide isomerase (hPDI). *Org. Biomol. Chem.* **2014**, *12*, 3808.

(131) 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*, 2755–2760.

(132) Yu, L.; Hajduk, P. J.; Mack, J.; Olejniczak, E. T. Structural Studies of Bcl-xL/ligand Complexes using 19F NMR. *J. Biomol. NMR* **2006**, *34*, 221–227.

(133) Zartler, E. R.; Hanson, J.; Jones, B. E.; Kline, A. D.; Martin, G.; Mo, H.; Shapiro, M. J.; Wang, R.; Wu, H.; Yan, J. RAMPED-UP NMR: Multiplexed NMR-based screening for drug discovery. *J. Am. Chem. Soc.* **2003**, *125*, 10941–10946.

(134) Bryant, R. G. The NMR time scale. *J. Chem. Educ.* **1983**, *60*, 933.

(135) Rogers, M. T.; Woodbrey, J. C. A Proton Magnetic Resonance Study of Hindered Internal Rotation in Some Substituted N , N-Dimethylamides. *J. Phys. Chem.* **1962**, *66*, 540–546.

(136) Hull, W. E.; Sykes, B. D. Fluorotyrosine alkaline phosphatase: internal mobility of individual tyrosines and the role of chemical shift anisotropy as a 19F nuclear spin relaxation mechanism in proteins. *J. Mol. Biol.* **1975**, *98*, 121–53.

(137) Ho, C.; Pratt, E. A.; Rule, G. S. Membrane-bound d-lactate dehydrogenase of Escherichia coli: a model for protein interactions in membranes. *Biochim. Biophys. Acta - Rev. Biomembr.* **1989**, *988*, 173–184.

(138) Klein-Seetharaman, J.; Getmanova, E. V; Loewen, M. C.; Reeves, P. J.; Khorana, H. G. NMR spectroscopy in studies of light-induced structural changes in mammalian rhodopsin: Applicability of solution 19F NMR. *Proc. Natl. Acad. Sci.* **1999**, *96*, 13744–13749.

(139) Chung, K. Y.; Kim, T. H.; Manglik, A.; Alvares, R.; Kobilka, B. K.; Prosser, R. S. Role of Detergents in Conformational Exchange of a G Protein-coupled Receptor. *J. Biol. Chem.* **2012**, *287*, 36305–36311.

(140) 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 Aβ 1–40 Using 19 F NMR. *Biochemistry* **2013**, *52*, 1903–1912.

(141) Li, C.; Wang, G.-F.; Wang, Y.; Creager-Allen, R.; Lutz, E. A.; Scronce, H.; Slade, K. M.; Ruf, R. A. S.; Mehl, R. A.; Pielak, G. J. Protein 19 F NMR in Escherichia coli. *J. Am. Chem. Soc.* **2010**, *132*, 321–327.

(142) Hammill, J. T.; Miyake-Stoner, S.; Hazen, J. L.; Jackson, J. C.; Mehl, R. a

Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nat. Protoc.* **2007**, *2*, 2601–2607.

(143) Salopek-Sondi, B.; Vaughan, M. D.; Skeels, M. C.; Honek, J. F.; Luck, L. a 19 F NMR Studies of the Leucine-Isoleucine-Valine Binding Protein: Evidence That a Closed Conformation Exists in Solution. *J. Biomol. Struct. Dyn.* **2003**, *21*, 235–246.

(144) 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*, 3404–3416.

(145) Tang, Y.; Tirrell, D. A. Biosynthesis of a Highly Stable Coiled-Coil Protein Containing Hexafluoroleucine in an Engineered Bacterial Host. *J. Am. Chem. Soc.* **2001**, *123*, 11089–11090.

(146) Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A. Stabilization of Coiled-Coil Peptide Domains by Introduction of Trifluoroleucine †. *Biochemistry* **2001**, *40*, 2790–2796.

(147) Lee, H.-Y.; Lee, K.-H.; Al-Hashimi, H. M.; Marsh, E. N. G. Modulating Protein Structure with Fluorous Amino Acids: Increased Stability and Native-like Structure Conferred on a 4-Helix Bundle Protein by Hexafluoroleucine. *J. Am. Chem. Soc.* **2006**, *128*, 337–343.

(148) Hajduk, P. J.; Huth, J. R.; Fesik, S. W. Druggability Indices for Protein Targets Derived from NMR-Based Screening Data Druggability Indices for Protein Targets Derived from NMR-Based Screening Data. *J. Med. Chem.* **2005**, 2518–2525.

(149) Lodge, J. M.; Justin Rettenmaier, T.; Wells, J. A.; Pomerantz, W. C.; Mapp, A. K. FP tethering: a screening technique to rapidly identify compounds that disrupt protein–protein interactions. *Medchemcomm* **2014**, *5*, 370.

(150) Cai, S.; Seu, C.; Kovacs, Z.; Sherry, A. D.; Chen, Y. Sensitivity Enhancement of Multidimensional NMR Experiments by Paramagnetic Relaxation Effects. *J. Am. Chem. Soc.* **2006**, *128*, 13474–13478.

(151) Theillet, F.-X.; Binolfi, A.; Liokatis, S.; Verzini, S.; Selenko, P. Paramagnetic

relaxation enhancement to improve sensitivity of fast NMR methods: application to intrinsically disordered proteins. *J. Biomol. NMR* **2011**, *51*, 487–495.

(152) Yamamoto, K.; Vivekanandan, S.; Ramamoorthy, A. Fast NMR Data Acquisition From Bicelles Containing a Membrane-Associated Peptide at Natural-Abundance. *J. Phys. Chem. B* **2011**, *115*, 12448–12455.

(153) Eletsky, a; Moreira, O.; Kovacs, H.; Pervushin, K. A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using C-13 spectroscopy. *J. Biomol. NMR* **2003**, *26*, 167–179.

(154) Chan, S. H. S.; Waudby, C. A.; Cassaignau, A. M. E.; Cabrita, L. D.; Christodoulou, J. Increasing the sensitivity of NMR diffusion measurements by paramagnetic longitudinal relaxation enhancement, with application to ribosome–nascent chain complexes. *J. Biomol. NMR* **2015**, *63*, 151–163.

(155) Takeuchi, K.; Frueh, D. P.; Hyberts, S. G.; Sun, Z.-Y. J.; Wagner, G. High-Resolution 3D CANCA NMR Experiments for Complete Mainchain Assignments Using C α Direct Detection. *J. Am. Chem. Soc.* **2010**, *132*, 2945–2951.

(156) Manglik, A.; Kim, T. H.; Masureel, M.; Altenbach, C.; Yang, Z.; Hilger, D.; Lerch, M. T.; Kobilka, T. S.; Thian, F. S.; Hubbell, W. L.; Prosser, R. S.; Kobilka, B. K. Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell* **2015**, *161*, 1101–1111.

(157) Arrowsmith, C. H.; Bountra, C.; Fish, P. V; Lee, K.; Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nat. Rev. Drug Discov.* **2012**, *11*, 384–400.

(158) Filippakopoulos, P.; Knapp, S. Targeting bromodomains: epigenetic readers of lysine acetylation. *Nat. Rev. Drug Discov.* **2014**, *13*, 337–356.

(159) 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**.

(160) Ferguson, F. M.; Dias, D. M.; Rodrigues, J.; Wienk, H.; Boelens, R.; Bonvin, A.; Abell, C.; Ciulli, A. Binding Hotspots of BAZ2B Bromodomain: Histone Interaction Revealed by Solution NMR Driven Docking. *Biochemistry* **2014**, *53*,

6706–6716.

(161) Marsh, E. N. G.; Suzuki, Y. Using 19F NMR to probe biological interactions of proteins and peptides. *ACS Chem. Biol.* **2014**, *9*, 1242–1250.

(162) Ge, X.; MacRaild, C. A.; Devine, S. M.; Debono, C. O.; Wang, G.; Scammells, P. J.; Scanlon, M. J.; Anders, R. F.; Foley, M.; Norton, R. S. Ligand-Induced Conformational Change of Plasmodium falciparum AMA1 Detected Using F-19 NMR. *J. Med. Chem.* **2014**, *57*, 6419–6427.

(163) Philpott, M.; Yang, J.; Tumber, T.; Fedorov, O.; Uttarkar, S.; Filippakopoulos, P.; Picaud, S.; Keates, T.; Felletar, I.; Ciulli, A.; Knapp, S.; Heightman, T. D. Bromodomain-peptide displacement assays for interactome mapping and inhibitor discovery. *Mol. Biosyst.* **2011**, *7*, 2899.

(164) Drewry, D. H.; Willson, T. M.; Zuercher, W. J. Seeding Collaborations to Advance Kinase Science with the GSK Published Kinase Inhibitor Set (PKIS). *Curr. Top. Med. Chem.* **2014**, *14*, 340–342.

(165) Zhang, Z. J.; Kwiatkowski, N.; Zeng, H.; Lim, S. M.; Gray, N. S.; Zhang, W.; Yang, P. L. Leveraging kinase inhibitors to develop small molecule tools for imaging kinases by fluorescence microscopy. *Mol. Biosyst.* **2012**, *8*, 2523–2526.

(166) Drouin, L.; McGrath, S.; Vidler, L. R.; Chaikuad, A.; Monteiro, O.; Tallant, C.; Philpott, M.; Rogers, C.; Fedorov, O.; Liu, M. J.; Akhtar, W.; Hayes, A.; Raynaud, F.; Muller, S.; Knapp, S.; Hoelder, S. Structure Enabled Design of BAZ2-ICR, A Chemical Probe Targeting the Bromodomains of BAZ2A and BAZ2B. *J. Med. Chem.* **2015**, *58*, 2553–2559.

(167) Takle, A. K.; Brown, M. J. B.; Davies, S.; Dean, D. K.; Francis, G.; Gaiba, A.; Hird, A. W.; King, F. D.; Lovell, P. J.; Naylor, A.; Reith, A. D.; Steadman, J. G.; Wilson, D. M. The identification of potent and selective imidazole-based inhibitors of B-Raf kinase. *Bioorganic Med. Chem. Lett.* **2006**, *16*, 378–381.

(168) Boehm, J. C.; Bower, M. J.; Gallagher, T. F.; Kassis, S.; Johnson, S. R.; Adams, J. L. Phenoxypyrimidine inhibitors of p38α kinase. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1123–1126.

(169) Dixon, S. L.; Smondyrev, A. M.; Rao, S. N. PHASE: A novel approach to

pharmacophore modeling and 3D database searching. *Chem. Biol. Drug Des.* **2006**, *67*, 370–372.

(170) Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. a. PHASE: A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results. *J. Comput. Aided. Mol. Des.* **2006**, *20*, 647–671.

(171) Small-Molecule Drug Discovery Suite 2014−2: Phase, version 3.9; Schrödinger, LLC: New York, 2014.

(172) Cho, A. E.; Guallar, V.; Berne, B. J.; Friesner, R. Importance of accurate charges in molecular docking: Quantum Mechanical/Molecular Mechanical (QM/MM) approach. *J. Comput. Chem.* **2005**, *26*, 915–931.

(173) Small-Molecule Drug Discovery Suite 2014−2: Schrödinger Suite 2014−2 QM-Polarized Ligand Docking protocol; Glide, version 6.3; Schrödinger,LLC,New York, NY, 2014. Jaguar, version 8.4; Schrödinger, LLC: New York, 2014. QSite, version 6.3; Schrödinger, LLC: New York, 2014.

(174) Fersht, A. *Structure and Mechanism in Protein Science*; Freeman: New York, 1999.

(175) Gamo, F.-J.; Sanz, L. M.; Vidal, J.; de Cozar, C.; Alvarez, E.; Lavandera, J.-L.; Vanderwall, D. E.; Green, D. V. S.; Kumar, V.; Hasan, S.; Brown, J. R.; Peishoff, C. E.; Cardon, L. R.; Garcia-Bustos, J. F. Thousands of chemical starting points for antimalarial lead identification. *Nature* **2010**, *465*, 305–310.

(176) Montesano, R.; Soriano, J. V; Fialka, I.; Orci, L. Isolation of EpH4 mammary epithelial cell subpopulations which differ in their morphogenetic properties. *Vitr. Cell. Dev. Biol. - Anim.* **1998**, *34*, 468–477.

(177) Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **1995**, *4*, 2411–2423.

(178) Warner, K. D.; Homan, P.; Weeks, K. M.; Smith, A. G.; Abell, C.; Ferré-D'Amaré, A. R. Validating Fragment-Based Drug Discovery for Biological RNAs: Lead Fragments Bind and Remodel the TPP Riboswitch Specifically. *Chem. Biol.*

2014, *21*, 591–595.

(179) Fontaine, F.; Overman, J.; François, M. Pharmacological manipulation of transcription factor protein-protein interactions: opportunities and obstacles. *Cell Regen. (London, England)* **2015**, *4*, 2.

(180) Bamborough, P.; Chung, C. Fragments in bromodomain drug discovery. *Med. Chem. Commun.* **2015**, *6*, 1587–1604.

(181) Shepherd, C. A.; Hopkins, A. L.; Navratilova, I. Fragment screening by SPR and advanced application to GPCRs. *Prog. Biophys. Mol. Biol.* **2014**, *116*, 113–123.

(182) Begley, D. W.; Moen, S. O.; Pierce, P. G.; Zartler, E. R. Saturation Transfer Difference NMR for Fragment Screening. In *Current Protocols in Chemical Biology*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2013; pp. 251–268.

(183) Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J. Biomol. NMR* **2000**, *18*, 65–8.

(184) Schiebel, J.; Radeva, N.; Krimmer, S. G.; Wang, X.; Stieler, M.; Ehrmann, F. R.; Fu, K.; Metz, A.; Huschmann, F. U.; Weiss, M. S.; Mueller, U.; Heine, A.; Klebe, G. Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study. *ACS Chem. Biol.* **2016**, *11*, 1693–1701.

(185) Schiebel, J.; Radeva, N.; Köster, H.; Metz, A.; Krotzky, T.; Kuhnert, M.; Diederich, W. E.; Heine, A.; Neumann, L.; Atmanene, C.; Roecklin, D.; Vivat-Hannah, V.; Renaud, J.-P.; Meinecke, R.; Schlinck, N.; Sitte, A.; Popp, F.; Zeeb, M.; Klebe, G. One Question, Multiple Answers: Biochemical and Biophysical Screening Methods Retrieve Deviating Fragment Hit Lists. *ChemMedChem* **2015**, *10*, 1511–1521.

(186) Kutchukian, P. S.; Wassermann, A. M.; Lindvall, M. K.; Wright, S. K.; Ottl, J.; Jacob, J.; Scheufler, C.; Marzinzik, A.; Brooijmans, N.; Glick, M. Large Scale Meta-Analysis of Fragment-Based Screening Campaigns: Privileged Fragments and Complementary Technologies. *J. Biomol. Screen.* **2015**, *20*, 588–596.

(187) Keserű, G. M.; Erlanson, D. A.; Ferenczy, G. G.; Hann, M. M.; Murray, C. W.;

Pickett, S. D. Design Principles for Fragment Libraries: Maximizing the Value of Learnings from Pharma Fragment-Based Drug Discovery (FBDD) Programs for Use in Academia. *J. Med. Chem.* **2016**, acs.jmedchem.6b00197.

(188) Rodriguez-Mias, R. a; Pellecchia, M. Use of Selective Trp Side Chain Labeling To Characterize Protein−Protein and Protein−Ligand Interactions by NMR Spectroscopy. *J. Am. Chem. Soc.* **2003**, *125*, 2892–2893.

(189) Jordan, J. B.; Whittington, D. A.; Bartberger, M. D.; Sickmier, E. A.; Chen, K.; Cheng, Y.; Judd, T. Fragment-Linking Approach Using 19 F NMR Spectroscopy To Obtain Highly Potent and Selective Inhibitors of β-Secretase. *J. Med. Chem.* **2016**, *59*, 3732–3749.

(190) Wamhoff, E.-C.; Hanske, J.; Schnirch, L.; Aretz, J.; Grube, M.; Varón Silva, D.; Rademacher, C. 19 F NMR-Guided Design of Glycomimetic Langerin Ligands. *ACS Chem. Biol.* **2016**, *11*, 2407–2413.

(191) Devine, S. M.; Mulcair, M. D.; Debono, C. O.; Leung, E. W. W.; Nissink, J. W. M.; Lim, S. S.; Chandrashekaran, I. R.; Vazirani, M.; Mohanty, B.; Simpson, J. S.; Baell, J. B.; Scammells, P. J.; Norton, R. S.; Scanlon, M. J. Promiscuous 2-Aminothiazoles (PrATs): A Frequent Hitting Scaffold. *J. Med. Chem.* **2015**.

(192) Norton, R.; Leung, E.; Chandrashekaran, I.; MacRaild, C. Applications of 19F-NMR in Fragment-Based Drug Discovery. *Molecules* **2016**, *21*, 860.

(193) 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*, 1414–1427.

(194) Anderluh, G.; Razpotnik, A.; Podlesek, Z.; Maček, P.; Separovic, F.; Norton, R. S. Interaction of the Eukaryotic Pore-forming Cytolysin Equinatoxin II with Model Membranes: 19F NMR Studies. *J. Mol. Biol.* **2005**, *347*, 27–39.

(195) Richards, K. L.; Rowe, M. L.; Hudson, P. B.; Williamson, R. A.; Howard, M. J. Combined ligand-observe 19F and protein-observe 15N,1H-HSQC NMR suggests phenylalanine as the key Δ-somatostatin residue recognized by human protein disulfide isomerase. *Sci. Rep.* **2016**, *6*, 19518.

(196) Barile, E.; Pellecchia, M. NMR-Based Approaches for the Identification and Optimization of Inhibitors of Protein–Protein Interactions. *Chem. Rev.* **2014**, *114*, 4749–4763.

(197) Prinjha, R. K.; Witherington, J.; Lee, K. Place your BETs: the therapeutic potential of bromodomains. *Trends Pharmacol. Sci.* **2012**, *33*, 146–153.

(198) Practical Fragments - NMR Poll Results http://practicalfragments.blogspot.com/search/label/CPMG.

(199) Halgren, T. New method for fast and accurate binding-site identification and analysis. *Chem. Biol. Drug Des.* **2007**, *69*, 146–148.

(200) Halgren, T. a Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–389.

(201) Schrödinger Release 2015-3: SiteMap, version 3.6, Schrödinger, LLC, New York, NY, 2016.

(202) Ferreira, R. S.; Bryant, C.; Ang, K. K. H.; McKerrow, J. H.; Shoichet, B. K.; Renslo, A. R. Divergent Modes of Enzyme Inhibition in a Homologous Structure−Activity Series. *J. Med. Chem.* **2009**, *52*, 5005–5008.

(203) Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability. *J. Biomol. NMR* **2001**, *21*, 349–359.

(204) Cozzi, F.; Ponzini, F.; Annunziata, R.; Cinquini, M.; Siegel, J. S. Polar Interactions between Stacked π Systems in Fluorinated 1,8-Diarylnaphthalenes: Importance of Quadrupole Moments in Molecular Recognition. *Angew. Chemie Int. Ed. English* **1995**, *34*, 1019–1020.

(205) Torrice, M. M.; Bower, K. S.; Lester, H. A.; Dougherty, D. A. Probing the role of the cation–π interaction in the binding sites of GPCRs using unnatural amino acids. *Proc. Natl. Acad. Sci.* **2009**, *106*, 11919–11924.

(206) Zhong, W.; Gallivan, J. P.; Zhang, Y.; Li, L.; Lester, H. a; Dougherty, D. a From ab initio quantum mechanics to molecular neurobiology: A cation-π binding site in the nicotinic receptor. *Proc. Natl. Acad. Sci.* **1998**, *95*, 12088–12093.

(207) Peng, C.; Frommlet, A.; Perez, M.; Cobas, C.; Blechschmidt, A.; Dominguez, S.;

Lingel, A. Fast and Efficient Fragment-Based Lead Generation by Fully Automated Processing and Analysis of Ligand-Observed NMR Binding Data. *J. Med. Chem.* **2016**, *59*, 3303–3310.

(208) Lovering, F.; Bikker, J.; Humblet, C. Escape from flatland: Increasing saturation as an approach to improving clinical success. *J. Med. Chem.* **2009**, *52*, 6752–6756.

(209) Ritchie, T. J.; Macdonald, S. J. F. The impact of aromatic ring count on compound developability - are too many aromatic rings a liability in drug design? *Drug Discov. Today* **2009**, *14*, 1011–1020.

(210) Huston, A.; Arrowsmith, C. H.; Knapp, S.; Schapira, M. Probing the epigenome. *Nat. Chem. Biol.* **2015**, *11*, 542–545.

(211) Han, B.; Liu, Y.; Ginzinger, S. W.; Wishart, D. S. SHIFTX2: significantly improved protein chemical shift prediction. *J. Biomol. NMR* **2011**, *50*, 43–57.

(212) Karp, J. M.; Erylimaz, E.; Cowburn, D. Correlation of chemical shifts predicted by molecular dynamics simulations for partially disordered proteins. *J. Biomol. NMR* **2015**, *61*, 35–45.

(213) Lehtivarjo, J.; Tuppurainen, K.; Hassinen, T.; Laatikainen, R.; Peräkylä, M. Combining NMR ensembles and molecular dynamics simulations provides more realistic models of protein structures in solution and leads to better chemical shift prediction. *J. Biomol. NMR* **2012**, *52*, 257–267.

(214) Pearson, J. G.; Oldfield, E.; Lee, F. S.; Warshel, A. Chemical shifts in proteins: a shielding trajectory analysis of the fluorine nuclear magnetic resonance spectrum of the Escherichia coli galactose binding protein using a multipole shielding polarizability-local reaction field-molecular dynamics approach. *J. Am. Chem. Soc.* **1993**, *115*, 6851–6862.

(215) Sternberg, U.; Klipfel, M.; Grage, S. L.; Witter, R.; Ulrich, A. S. Calculation of fluorine chemical shift tensors for the interpretation of oriented 19F-NMR spectra of gramicidin A in membranes. *Phys. Chem. Chem. Phys.* **2009**, *11*, 7048.

(216) Lau, E. Y.; Gerig, J. T. Origins of Fluorine NMR Chemical Shifts in Fluorine-Containing Proteins †. *J. Am. Chem. Soc.* **2000**, *122*, 4408–4417.

(217) Kasireddy, C.; Bann, J. G.; Mitchell-Koch, K. R. Demystifying fluorine chemical

shifts: electronic structure calculations address origins of seemingly anomalous 19 F-NMR spectra of fluorohistidine isomers and analogues. *Phys. Chem. Chem. Phys.* **2015**, *17*, 30606–30612.

(218) Hospital, A.; Andrio, P.; Fenollosa, C.; Cicin-Sain, D.; Orozco, M.; Gelpi, J. L. MDWeb and MDMoby: an integrated web-based platform for molecular dynamics simulations. *Bioinformatics* **2012**, *28*, 1278–1279.

(219) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins Struct. Funct. Bioinforma.* **2006**, *65*, 712–725.

(220) Best, R. B.; Hummer, G. Optimized Molecular Dynamics Force Fields Applied to the Helix−Coil Transition of Polypeptides. *J. Phys. Chem. B* **2009**, *113*, 9004–9015.

(221) Gelpí, J. L.; Kalko, S. G.; Barril, X.; Cirera, J.; de la Cruz, X.; Luque, F. J.; Orozco, M. Classical molecular interaction potentials: Improved setup procedure in molecular dynamics simulations of proteins. *Proteins Struct. Funct. Bioinforma.* **2001**, *45*, 428–437.

(222) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(223) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, 33.

(224) Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006**, *125*, 194101.

(225) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

(226) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by

the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

(227) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158.

(228) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648.

(229) McLean, a. D.; Chandler, G. S. Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z=11–18. *J. Chem. Phys.* **1980**, *72*, 5639.

(230) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650.

(231) Sanders, L. K.; Oldfield, E. Theoretical Investigation of 19 F NMR Chemical Shielding Tensors in Fluorobenzenes. *J. Phys. Chem. A* **2001**, *105*, 8098–8104.

(232) Oldfield, E. CHEMICAL SHIFTS IN AMINO ACIDS,PEPTIDES, AND PROTEINS: From Quantum Chemistry to Drug Design. *Annu. Rev. Phys. Chem.* **2002**, *53*, 349–378.

(233) Allen, M. J.; Keal, T. W.; Tozer, D. J. Improved NMR chemical shifts in density functional theory. *Chem. Phys. Lett.* **2003**, *380*, 70–77.

(234) Liu, P.; Goddard, J. D.; Arsenault, G.; Gu, J.; McAlees, A.; McCrindle, R.; Robertson, V. Theoretical studies of the conformations and 19F NMR spectra of linear and a branched perfluorooctanesulfonamide (PFOSAmide). *Chemosphere* **2007**, *69*, 1213–1220.

(235) Dalvit, C.; Invernizzi, C.; Vulpetti, A. Fluorine as a Hydrogen-Bond Acceptor: Experimental Evidence and Computational Calculations. *Chem. - A Eur. J.* **2014**, *20*, 11058–11068.

(236) Marchand, A. P.; Rose, J. E. On the question of bridge-proton absorptions in the nuclear magnetic resonance spectra of norbornene and related systems. *J. Am. Chem. Soc.* **1968**, *90*, 3724.

(237) Cai, S.-H.; Chen, Z.; Wan, H.-L. Theoretical Investigation of 19 F NMR Chemical Shielding of Alkaline-Earth-Metal and Alkali-Metal Fluorides. *J. Phys. Chem. A*

2002, *106*, 1060–1066.

(238) Seyedsayamdost, M. R.; Reece, S. Y.; Nocera, D. G.; Stubbe, J. Mono-, Di-, Tri-, and Tetra-Substituted Fluorotyrosines: New Probes for Enzymes That Use Tyrosyl Radicals in Catalysis †. *J. Am. Chem. Soc.* **2006**, *128*, 1569–1579.