UCSF
UC San Francisco Electronic Theses and Dissertations

Title

Beyond cysteine: exploring lysine- and histidine-targeted covalent inhibitors

Permalink

https://escholarship.org/uc/item/554351gw

Author

Weaver, Jacqueline

Publication Date

2024

Supplemental Material

https://escholarship.org/uc/item/554351gw#supplemental

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Beyond cysteine: exploring lysine- and histidine-targeted covalent inhibitors

by
Jacqueline Weaver

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Jack Taunton
0FD9D71201FC43C...

Jack Taunton
Chair

DocuSigned by:
Matthew Jacobson

Matthew Jacobson

DocuSigned by:
Adam Renslo
FBDB6299FCD244B...

Adam Renslo

Committee Members

Copyright 2024

By

Jacqueline Weaver

Acknowledgements

The journey to finish my PhD was a long one that would not have been possible without the support of a great many people. I would first like to thank my PhD advisor, Jack Taunton, for never giving up on me. Jack is an unparalleled scientific mentor, with his wide breadth of knowledge and knack for distilling complex concepts down to their core components. Jack never failed to push me to become a better and better scientist, fostering my confidence to put forth my ideas and connect with other scientists. At the core, Jack always believed in me, even when I didn't believe in myself, and there is nothing greater I could ask for from a mentor.

Next, I'd like to thank all my peers in the many iterations of the Taunton lab. Haoyuan was an outstanding source of advice when I first joined the lab. Nicole was always ready to speak up and take charge when the situation called for it. Keely was an endless source of energy and inspiration. Tangpo and Ying were steady and reliable, always happy to answer questions and provide their expertise. Greg has been a great friend both in and out of lab and is just as good at teaching me how to rock climb as he is at teaching me how to refine a crystal structure. Hao, Nick, and Conor have helped to foster a thriving lab environment to help me cross the finish line of my time in graduate school.

Finally, I'd like to thank my many friends and family who supported me on my journey. My siblings were always ready to get me out of my head and away from science when I needed it. My sister in particular was always ready to make time for me, whether that be to bake a cake, play some videogames, or stay up way past her bedtime to chat about anything and everything. Most of all, I need to thank my parents.

Whenever I found myself floundering, they were always there to support me through my struggles. Whether that be helping me with my taxes or making a 14 h drive to hand deliver some hugs, they were always there when I needed them. Love you, Mom and Dad.

Contributions

Chapter 1 of this thesis is reproduced in its entirety from previously published work cited below:

Weaver, J; Craven, C. B.; Tram, L.; Chen, H.; Taunton, J. Aminomethyl Salicylaldehydes Lock onto a Surface Lysine by Forming an Extended Intramolecular Hydrogen Bond Network. *J. Am. Chem. Soc.* **2024**, *146* (35), 24233–24237.

Beyond cysteine: exploring lysine- and histidine-targeted covalent inhibitors

Jacqueline Weaver

Abstract

Covalent inhibitors are versatile and powerful compounds, useful in their potential as drugs, tools for drug discovery, and probes for chemical biology. Traditionally, covalent inhibitors have targeted cysteine. As the most reactive nucleophilic amino acid at physiological pH, this allows for even weakly reactive electrophiles to exhibit good reaction kinetics towards their target residue. The ability to utilize weakly reactive electrophiles in addition to cysteines low abundance in the proteome reduces the risk of undesired, off-target effects. However, its low abundance comes with the trade-off that many potential protein targets do not contain a cysteine residue proximal to small molecule binding sites, rendering them unviable as targets for cysteine-targeted covalent inhibitors. This leads to the need to develop covalent inhibitors utilizing electrophiles that target non-cysteine amino acids, expanding the breadth of proteins targetable with covalent inhibitors and probes.

We first focused on salicylaldehydes as reversible-covalent inhibitors to target lysines. While salicylaldehydes have previously been reported to form imine adducts with lysine residues on proteins, these adducts have only shown good stability when buried in a hydrophobic pocket where exclusion of water hinders hydrolysis. Whether salicylaldehydes can form highly stable adducts with surface lysines, where solvation makes rapid hydrolysis of the imine more likely, remains an open question. We developed a series of salicylaldehyde probes targeting the surface exposed lysine on the ATPase domain of Hsp90, the best of which engaged the target lysine in a quasi-

irreversible manner. Furthermore, upon reduction of the reversible binding scaffold to a fragment-like size, salicylaldehyde compounds showed significant and sustained target engagement at low micromolar concentrations, demonstrating the potential of salicylaldehydes for use in covalent-fragment based drug discovery.

Second, we explored histidine-targeting covalent probes. Sulfonyl fluorides have previously been shown to react with histidine residues on proteins, but their metabolic stability remains a challenge when developing inhibitors towards in cell applications, necessitating a switch towards similar, but less reactive and more stable electrophiles like fluorosulfates. This balance between compound stability and target modification rate remains a challenge in developing histidine-targeting covalent inhibitors. Utilizing a series of sulfonyl fluoride compounds with different linkers, we showed that conformational constraints are an additional tool for tuning the reaction kinetics of a sulfonyl fluoride with a target histidine outside of altering innate electrophile reactivity. Analysis of histidines proximal to ligands in 65,288 mammalian protein structures from the PDB found that 18% of proteins had at least one ligand-proximal histidine, exemplifying the breadth of proteins containing histidines targetable with covalent inhibitors.

Table of Contents

List of Figures

List of Tables

Chapter 1 : Aminomethyl salicylaldehydes lock onto a surface lysine by forming an extended intramolecular hydrogen bond network

Abstract:

The development of electrophilic ligands that rapidly modify a specific lysine residue remains a major challenge. Salicylaldehyde-based inhibitors have been reported to form stable imine adducts with the catalytic lysine of protein kinases. However, the targeted lysine in these examples is buried in a hydrophobic environment. A key unanswered question is whether this strategy can be applied to a lysine on the surface of a protein, where rapid hydrolysis of the resulting salicylaldimine is more likely. Here, we describe a series of aminomethyl-substituted salicylaldehydes that target a fully solvated lysine on the surface of the ATPase domain of Hsp90. By systematically varying the orientation of the salicylaldehyde, we discovered ligands with long residence times, the best of which engages Hsp90 in a quasi-irreversible manner. Crystallographic analysis revealed a daisy-chain network of intramolecular hydrogen bonds in which the salicylaldimine is locked into position by the adjacent piperidine linker. This study highlights the potential of aminomethyl salicylaldehydes to generate conformationally stabilized, hydrolysis-resistant imines, even when the targeted lysine is far from the ligand binding site and exposed to bulk solvent.

Main

Electrophilic ligands designed to react with protein nucleophiles are an emerging class of therapeutics, exemplified by the cysteine-targeted drugs ibrutinib[1,2], osimertinib[3,4], and sotorasib[5]. The development of targeted covalent inhibitors has mostly focused on cysteine due to its high intrinsic reactivity. However, this strategy cannot be used for most protein targets, which lack a druggable cysteine. Lysine, which

is more prevalent than cysteine in the proteome, is a potential alternative[6,7]. Although lysines are generally less reactive than cysteines, irreversible lysine-targeted inhibitors have been developed for diverse protein targets[8–15]. Many of these inhibitors are sulfonyl fluorides, which can be problematic owing to their metabolic instability and enhanced reactivity toward cysteine and tyrosine residues[16]. Fluorosulfates[17–19], 2-ethynylbenzaldehydes[20], and squarates[15,21] have also shown promise as lysine-targeted electrophiles, with the caveat that irreversible modification of surface lysines often requires hours to reach completion.

Salicylaldehydes can react rapidly and reversibly with biological amines, including lysine residues on proteins, to form stabilized imines[22]. Voxelotor is a salicylaldehyde drug, approved for sickle cell anemia, which covalently modifies the amino terminus of hemoglobin and allosterically prevents its polymerization[23]. Recently, we and others have reported salicylaldehyde-based inhibitors that reversibly target the conserved catalytic lysine of protein kinases[8,24,25]. We found that a subset of kinases bound to a salicylaldehyde probe quasi-irreversibly in cells, whereas the majority of kinases formed covalent complexes that dissociated within 30 minutes[24]. The prolonged residence times observed with certain salicylaldehyde/kinase adducts may require a hydrophobic binding pocket, as well as structured regions of the protein to shield the imine from bulk water. An intramolecular hydrogen bond within the phenol-imine or the keto-enamine tautomer[26,27] could also contribute to the observed long residence times. By contrast, *ortho*-boronic acid benzaldehydes designed to target a surface lysine in the anti-apoptotic protein MCL1, exhibited fast dissociation rates ($t_{1/2}$ < 1 min)[28], implying rapid imine hydrolysis and reformation. Thus, a critical unresolved question is whether

reversible inhibitors can be designed to engage a surface lysine with prolonged residence times. In this study, we describe a series of salicylaldehyde-containing inhibitors that reversibly engage Lys58 on the surface of Hsp90. We find that the precise orientation of the salicylaldehyde relative to the linker and the noncovalent recognition scaffold exerts a profound influence on the kinetic stability of the covalent complex.

We previously reported a series of arylsulfonyl fluoride inhibitors that irreversibly target Lys58 on the surface of the Hsp90 N-terminal ATPase domain[11] (**Fig. 1.1a**). By orienting the sulfonyl fluoride toward Lys58, the chiral piperidine linker attached to the ATPase-binding purine arylthioether enhances the rate of covalent modification. We hypothesized a salicylaldehyde attached to the same linker and scaffold could also trap Lys58. However, it was unclear if the resulting solvent-exposed imine would be sufficiently stable to withstand hydrolysis. Given lysine's inherent conformational flexibility and the distinct geometric requirements of reversible imine formation as compared to the irreversible reaction with sulfonyl fluorides, we systematically evaluated all four possible salicylaldehyde regioisomers 1-4 (**Fig. 1.1b**). Using intact-protein mass spectrometry, we assessed the rate of Hsp90 covalent modification (1 μM, N-terminal ATPase domain) in the presence of 10 μM salicylaldehyde (**Fig. 1.1c**). At each time point, an aliquot was removed and quenched with $NaBH_4$ (5 mM, 10 min) to reduce the Hsp90 imine-adduct and any unreacted salicylaldehyde. Regioisomeric salicylaldehydes 2-4 showed similarly rapid modification rates, reaching an apparent equilibrium between modified (70-100%) and unmodified (0-30%) Hsp90 within 15 min (**Fig. 1.1c**).

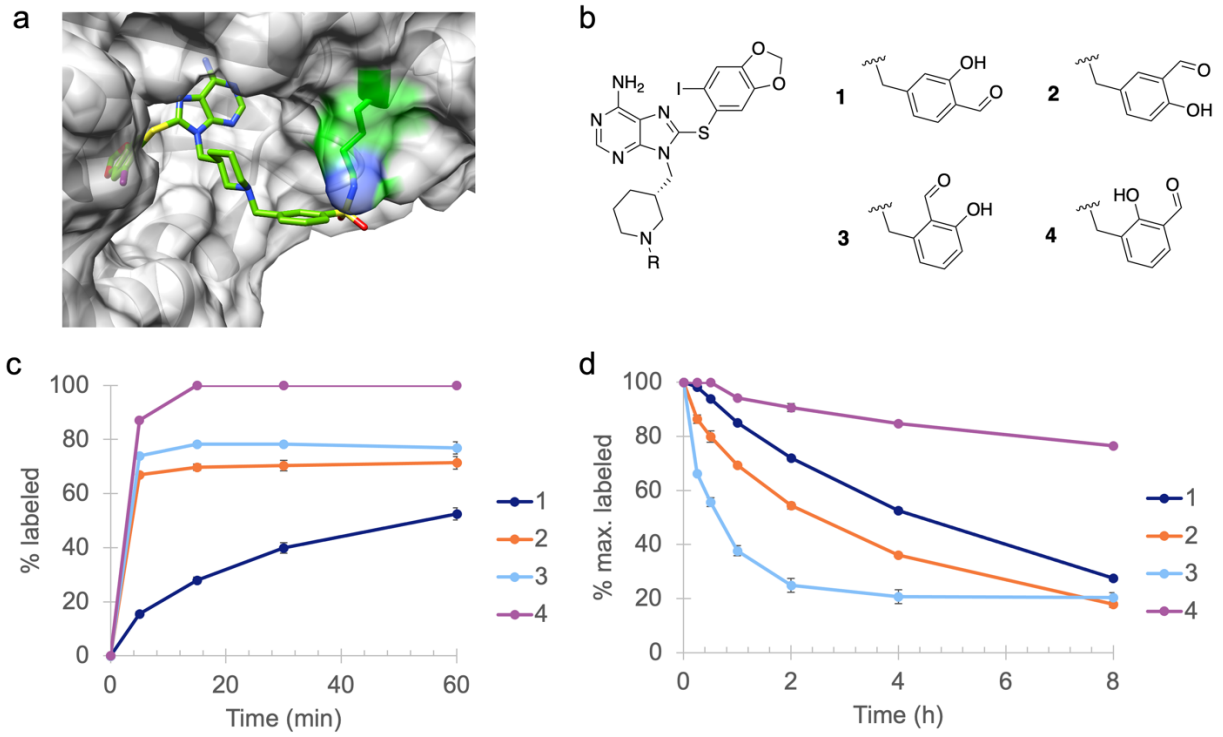


Figure 1.1: Salicylaldehyde regiochemistry determines modification kinetics.

(a) Structure of Hsp90 Lys58 modified by a sulfonyl fluoride probe (PDB: 6U9B). (b) Regioisomeric salicylaldehydes **1**-**4**. (c) Hsp90 (1 μM) was treated with the indicated compounds (10 μM) at room temperature. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of covalently modified Hsp90 quantified by intact-protein LC-MS analysis. (d) Hsp90 (5 μM) was treated with the indicated compounds (10 μM) for 2 h at room temperature, then diluted 10-fold into buffer containing 100 μM PU-H71. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of covalently modified Hsp90 quantified by intact-protein LC-MS. Each data point in (c) and (d) represents the mean value (± SD) from triplicate samples (CV = 1-3%).

Importantly, mutation of Lys58 to Arg prevented detectable modification of Hsp90 by **1-4** (**Fig. 1.2**). Because the concentration of the salicylaldehyde inhibitors in this experiment is >100-fold higher than the binding affinity of related noncovalent inhibitors such as PU-H71[29,30], we assumed that the maximal 70-80% modification observed with **2** and **3** reflects an equilibrium between the imine adducts and the noncovalently bound

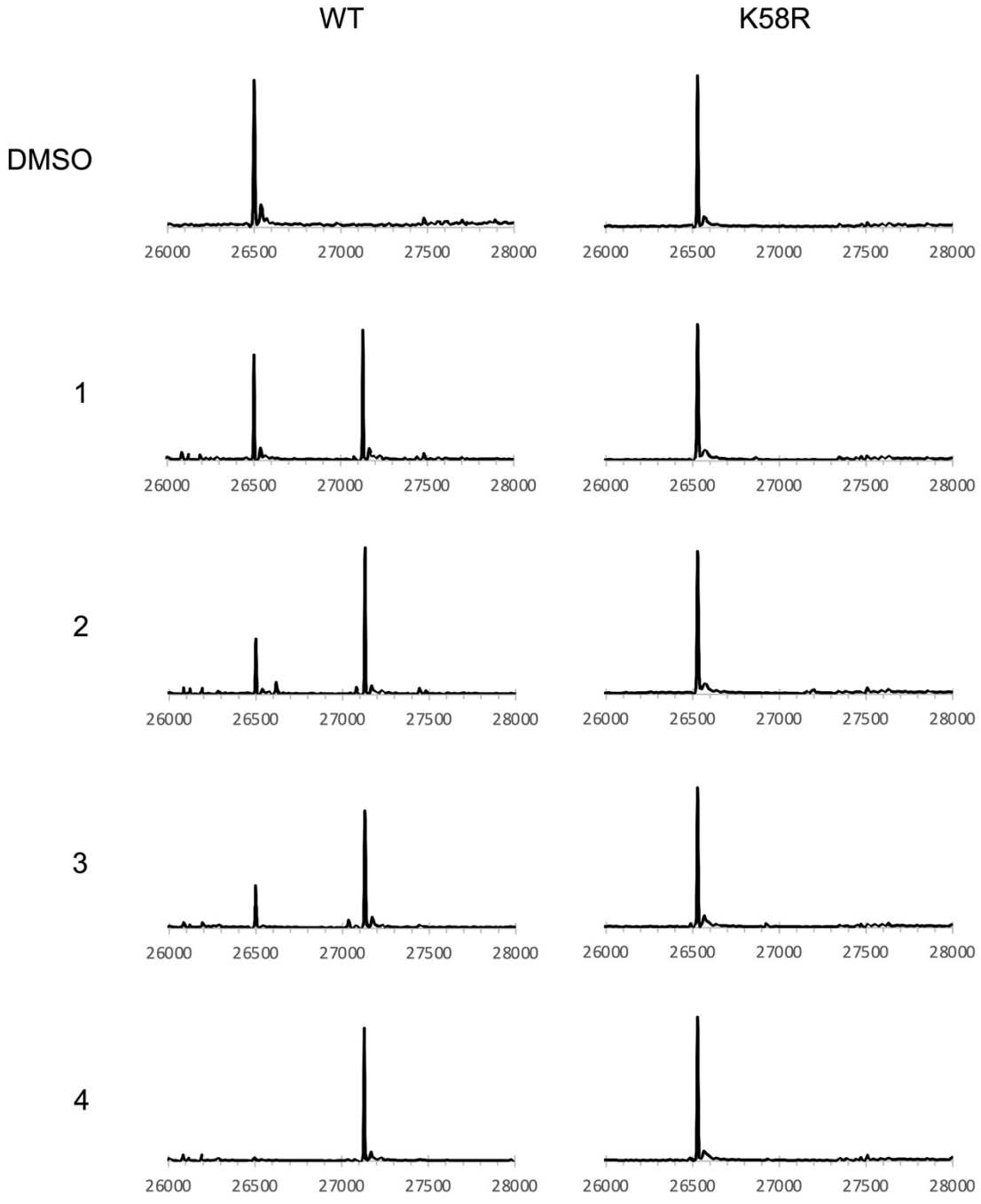


Figure 1.2: Intact-protein MS analysis of Hsp90 treated with compounds 1-4.

Hsp90 (WT or K58R, 1 μM) was treated with DMSO or the indicated compounds (10 μM) for 1 h, reduced with $NaBH_4$ for 10 min, and then analyzed by intact-protein LC-MS.

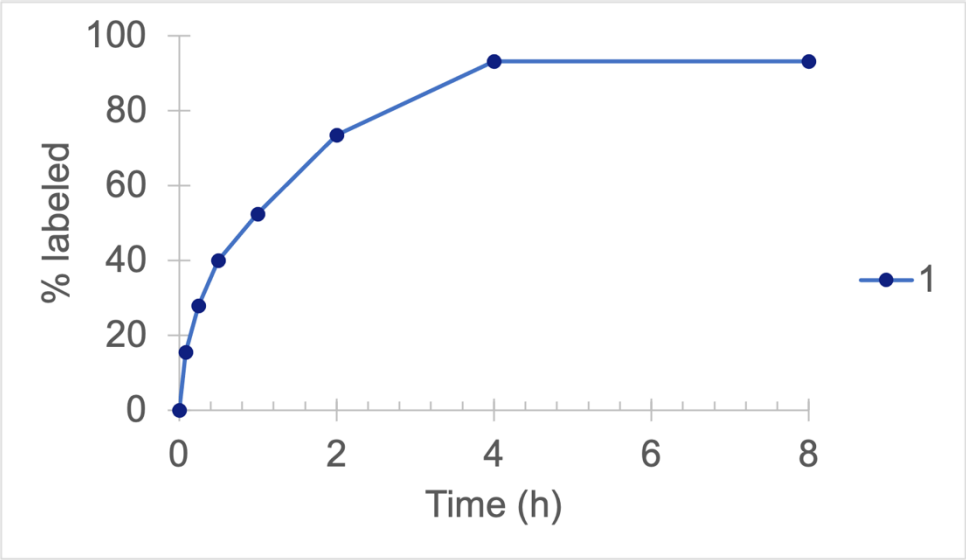


Figure 1.3: Long time course of Hsp90 modification by compound 1.

Hsp90 (1 μM) was treated with compound **1** (10 μM). At the indicated time points, aliquots were removed, reduced with $NaBH_4$ for 10 min, and the percentage of covalently modified protein was quantified by intact-protein LC-MS.

salicylaldehyde ligands. Although salicylaldehyde **1** modified Hsp90 at least 15-fold more slowly ($t_{1/2}$ ~1 h) than the regioisomers **2**-**4**, it ultimately reached 93% modification after 4 h (**Fig 1.3**). Hence, both the rate and the maximum extent of imine formation depend on the salicylaldehyde regiochemistry.

The regioisomeric salicylaldehydes **1**-**4** were further differentiated by comparing their apparent residence times. To estimate the dissociation half-time, we diluted the pre-formed Hsp90/salicylaldehyde complex 10-fold into buffer containing 100 μM PU-H71, a potent competitive inhibitor (K_d = 15 nM)[30], and quantified the disappearance of salicylaldehyde-modified Hsp90 and the formation of unmodified Hsp90 after brief treatment with $NaBH_4$, as described above. Thus, upon imine hydrolysis and dissociation of the inhibitor, rebinding is impeded by the large molar excess of PU-H71. Compound **4** exhibited by far the longest residence time, such that >80% of the covalent

complex remained 8 hours after dilution into 100 μM PU-H71 (**Fig. 1.1d**). By contrast, the other regioisomers dissociated with half-times ranging from 0.5 to 4 h. Collectively, the kinetic data suggest that the rates of imine formation and hydrolysis are strongly dependent on the orientation of both the aldehyde and the phenol (**Fig. 1.1c, 1.1d**). We hypothesized that the preferred 1,2,3-salicyaldehyde regioisomer **4** uniquely facilitates the formation of a conformationally stabilized imine with surface Lys58.

The addition of fluorine at one or more positions of the 1,2,3-salicylaldehyde could potentially modulate multiple chemical properties with varied effects on imine formation, including the electrophilicity of the aldehyde and imine, the pK_a of the phenol and imine, and the preferred conformation of substituents *ortho* to the fluorines. We synthesized three fluorinated variants and assessed the rates of formation and dissociation of the resulting 1,2,3-salicylaldimine adducts with Hsp90 Lys58 (**Fig. 1.4a**). Similar to salicylaldehyde **4**, the fluorinated derivatives **5**-**7** rapidly and completely labeled Hsp90 (**Fig. 1.4b**), with the difluoro compound **7** uniquely attaining 100% modification at the first time point of 5 min. Despite treating Hsp90 with a 10-fold molar excess of compound **7**, intact-protein MS analysis revealed only one modification event, which was not detected with the Lys58Arg mutant (**Fig 1.5**). Compared to the fluorinated variants **5** and **6**, difluoro-salicylaldehyde **7** also exhibited the longest residence time, with less than 5% dissociation during the 8-hour chase with 100-fold excess PU-H71 (**Fig. 1.4c**). The mechanisms underlying the observed fluorine substitution effects are likely complex and require further investigation.

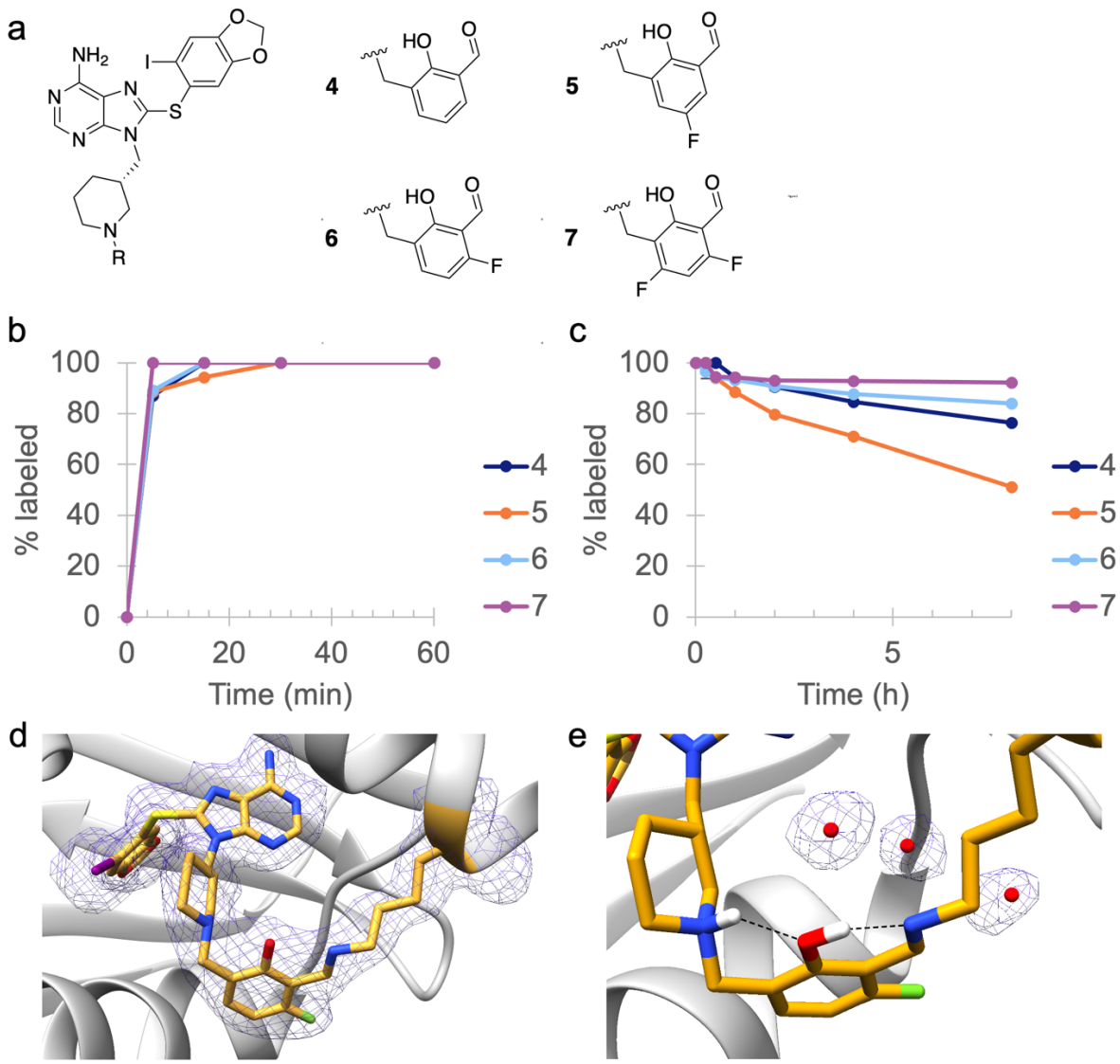


Figure 1.4: Regioselective effect of fluorine on residence time.

(a) Compounds **4**-**7**. (b) Hsp90 (1 μM) was treated with the indicated compounds (10 μM) and the percentage of covalently modified protein was assessed by intact-protein LC-MS. (c) Dissociation of covalently bound compounds was assessed as in Fig. 1. (d) Crystal structure of **6** bound to Hsp90 at 2.0 Å resolution. (e) Ordered water molecules proximal to the imine carbon. Electron density map (2Fo-Fc, 1.0 σ) is shown for **6** and bound waters in (d) and (e), respectively. Each data point in (a) and (b) represents the mean value (± SD) from triplicate samples (CV = 1-3%).

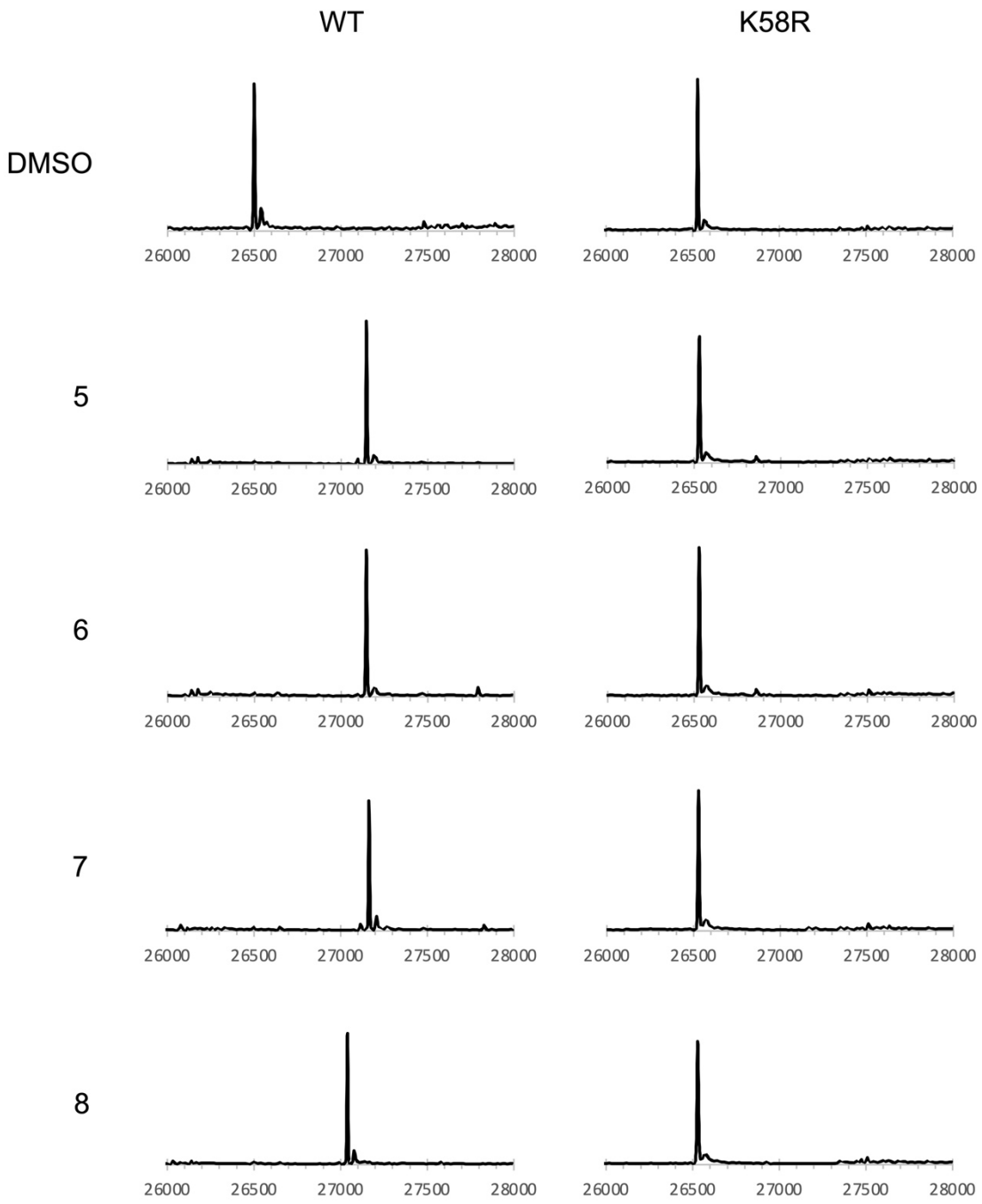


Figure 1.5: Intact-protein MS analysis of Hsp90 modified by compounds 5-8.

Hsp90 (WT or K58R, 1 μM) was treated with DMSO or the indicated compounds (10 μM) for 1 h, reduced with $NaBH_4$ for 10 min, and then analyzed by intact-protein LC-MS.

To gain structural insight into the remarkable kinetic stability conferred by the 1,2,3- salicylaldehyde regioisomer, we determined the cocrystal structure of the monofluoro derivative **6** bound to the Hsp90 ATPase domain without $NaBH_4$ reduction at 2.0 Å resolution. Continuous electron density between Lys58 and the inhibitor confirmed the formation of a *trans*-imine (**Fig. 1.4d**). Notably, one face of the imine is exposed to solvent whereas the other face is proximal to several ordered water molecules, three of which are within 4 Å of the imine carbon (**Fig. 1.4e**). These observations suggest that the hydrolytic stability of the Lys58 imine bond is not due to solvent inaccessibility, as previously observed with a quasi-irreversible, salicylaldehyde inhibitor of the protein kinase, AURKA[24]. Instead, the imine appears to be stabilized by a daisy chain of intramolecular hydrogen bonds: one between the protonated nitrogen of the piperidine and the phenol oxygen, and one between the phenol and the imine nitrogen. We speculate that this extended intramolecular hydrogen bond network, the formation of which is only possible with the 1,2,3-salicylaldehyde, contributes to the prolonged residence times of salicylaldehydes **4**-**7** relative to the regioisomers **1**-**3**. In addition to the covalent bond with Lys58, the crystal structure revealed the expected noncovalent interactions between the purine iodo-arylthioether and the ATPase active site. These interactions are identical to those formed by the related noncovalent inhibitor PU-H71[31].

Given that the difluoro-salicylaldehyde in 7 (MW 690 Da) adds substantial molecular weight relative to PU-H71 (+178 Da), we were motivated to test whether an Hsp90-binding scaffold with reduced molecular weight – and presumably reduced noncovalent binding affinity – could be converted into a potent inhibitor with long residence time, upon derivatization with a 1,2,3-salicylaldehyde**.** Studies of Hsp90

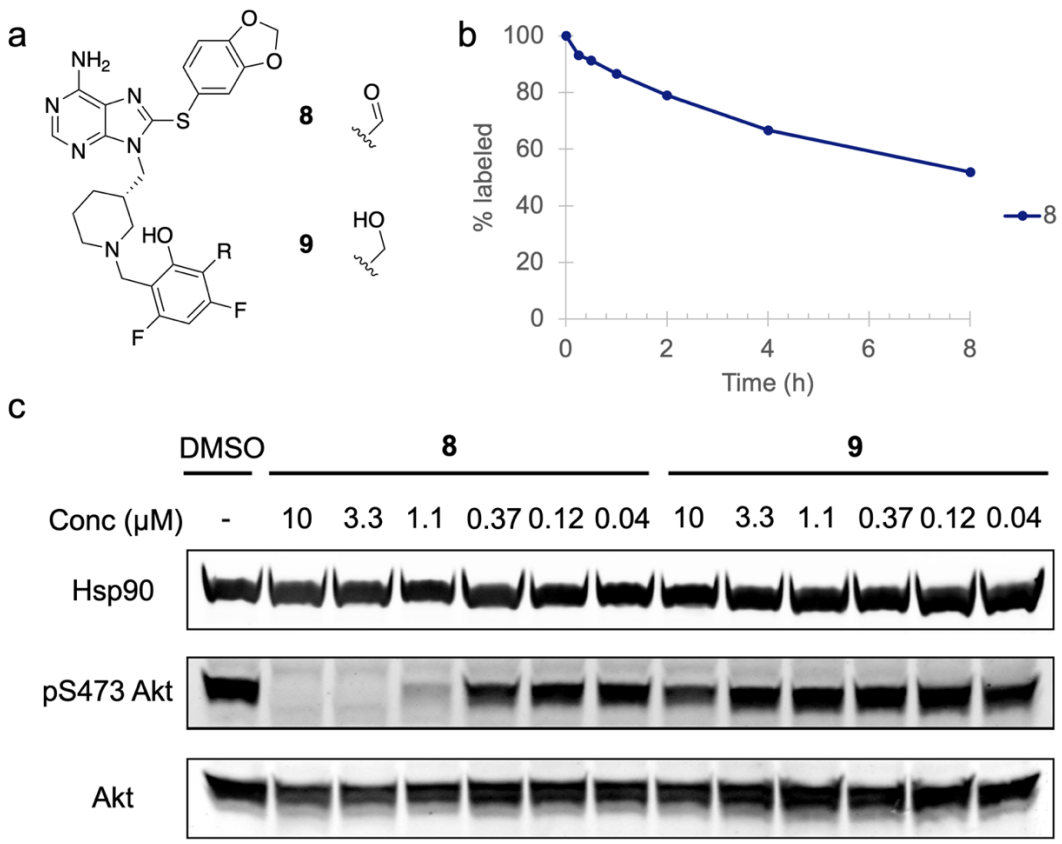


Figure 1.6: Salicylaldehyde electrophile transforms a des-iodo scaffold into a cell-active inhibitor.

(a) Salicylaldehyde **8** and the non-electrophilic alcohol **9**. (b) Compound **8** dissociation was assessed as in Fig. 1. (c) Skbr3 cells were treated with the indicated concentrations of **8** and **9** for 6 h at 37 °C. Cell lysates were analyzed by western blot for total Hsp90, phosphorylated Akt, and total Akt.

inhibitors related to PU-H71 have revealed that removing the iodo substituent results in a ~40-fold loss in potency[29,32]. Thus, we synthesized the des-iodo salicylaldehyde **8**, along with the benzyl alcohol **9** as a non-electrophilic control compound (**Fig. 1.6a**). Compound **8** quantitatively modified wild type Hsp90 but not the Lys58Arg mutant (**Fig 1.5**) and exhibited prolonged residence time ($t_{1/2}$ ~8 h, **Fig. 1.6b**) under both neutral and

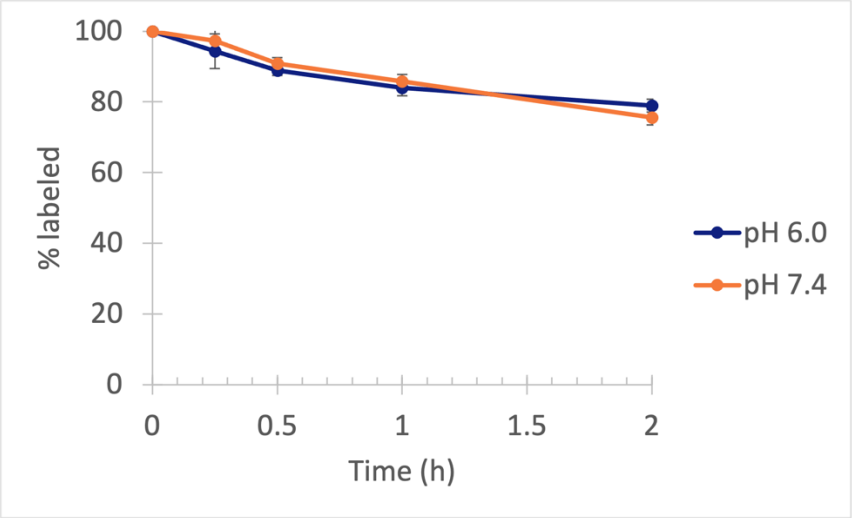


Figure 1.7: Time course of compound 8 dissociation from Hsp90 at low pH.

Hsp90 (5 μM) was treated with compound **8** (10 μM) for 15 min at the indicated pH, then diluted 10-fold into buffer containing 100 μM PU-H71 at the indicated pH. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of labeled protein was quantified by intact-protein LC-MS analysis (mean ± SD, triplicate samples).

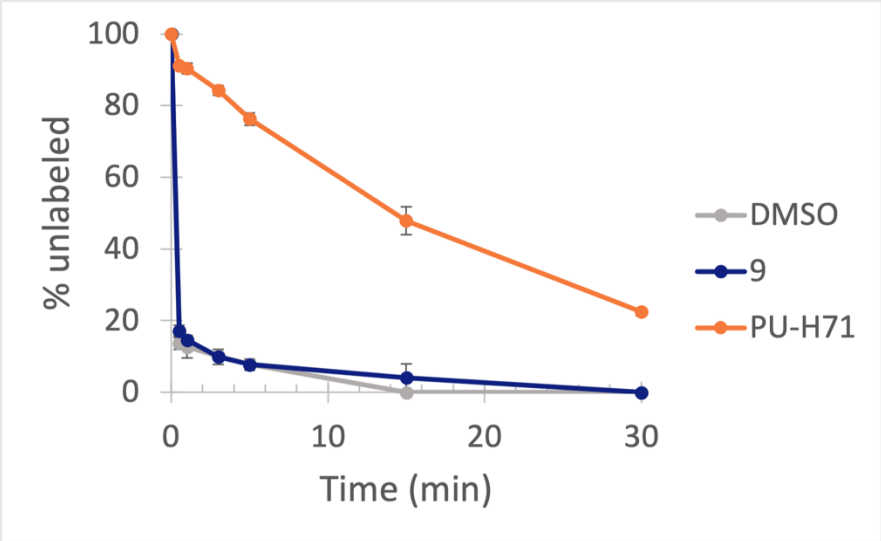


Figure 1.8: Time course of compound 9 dissociation from Hsp90.

Hsp90 (5 μM) was treated with the indicated compounds (10 μM) or DMSO for 15 min, then diluted 10-fold into buffer containing 20 μM compound **7**. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of unlabeled protein was quantified by intact-protein LC-MS analysis (mean ± SD, triplicate samples).

acidic conditions (**Fig 1.7**). By contrast, the benzyl alcohol **9** exhibited a much faster off-rate, dissociating from Hsp90 within seconds ($t_{1/2}$ < 30 s, **Fig 1.8**).

We hypothesized that the increase in residence time due to reversible covalent bond formation by **8** would confer increased potency in cells, even without a washout step to accentuate its kinetic advantage over the noncovalent control **9**. To test this, we treated Skbr3 breast cancer cells with increasing concentrations of compounds **8** or **9** and assessed the levels of phosphorylated Akt. In Skbr3 cells, the Akt kinase is activated downstream of amplified Her2, a receptor tyrosine kinase that requires continuous Hsp90 activity to maintain oncogenic signaling[33]. Strikingly, exposure of Skbr3 cells to salicylaldehyde **8** for 6 h strongly reduced phosphorylated AKT levels at concentrations greater than 370 nM (**Fig. 1.6c**). By contrast, the corresponding benzyl alcohol **9** had negligible effects at concentrations below 10 μM.

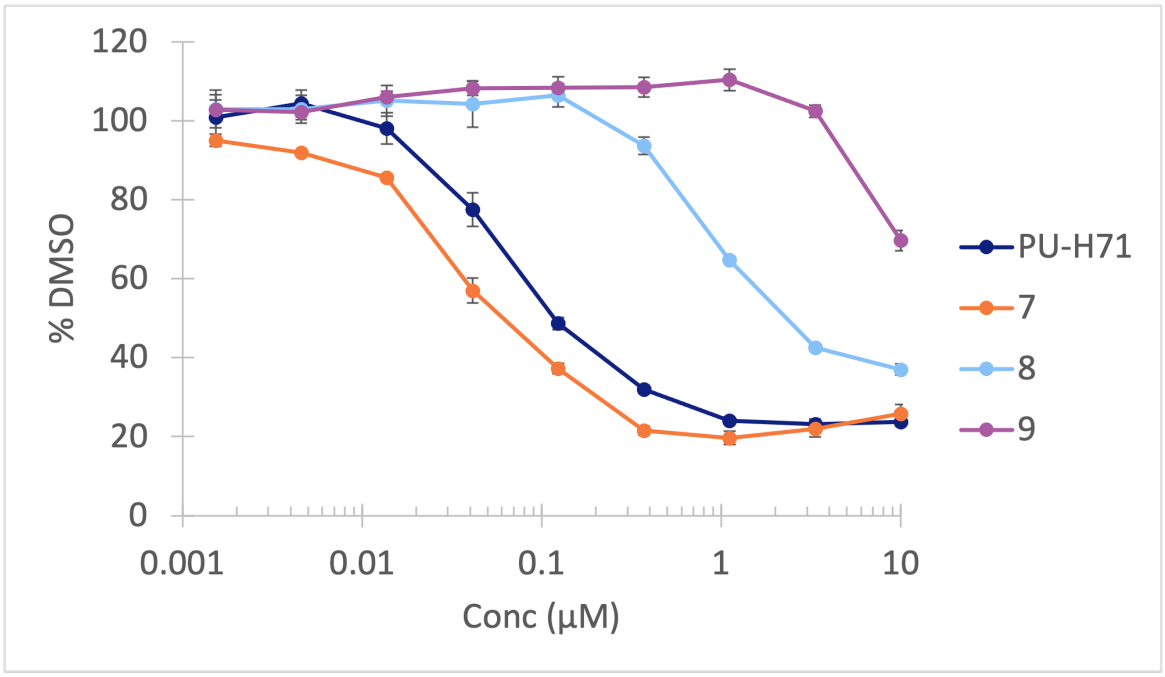


Figure 1.9: Skbr3 cell proliferation assay.

Skbr3 cells were treated with PU-H71, **7**, **8**, or **9** for 72 h at 37 °C. Cell proliferation was quantified using alamarBlue and plotted as a percent of DMSO control.

The increased potency of compound **8** compared to **9** was further demonstrated in a cell proliferation assay, as well as a competitive Hsp90 labeling assay[11] in Skbr3 cells (**Fig 1.9 and 1.10**). Hence, the addition of a 1,2,3-salicylaldehyde can rescue a weak, noncovalent inhibitor by reversibly yet persistently engaging Lys58 on the surface of Hsp90, far away from the primary ligand binding site.

By systematically testing a regioisomeric series of salicylaldehydes linked to a noncovalent recognition element, we found that altering the orientation of the phenol

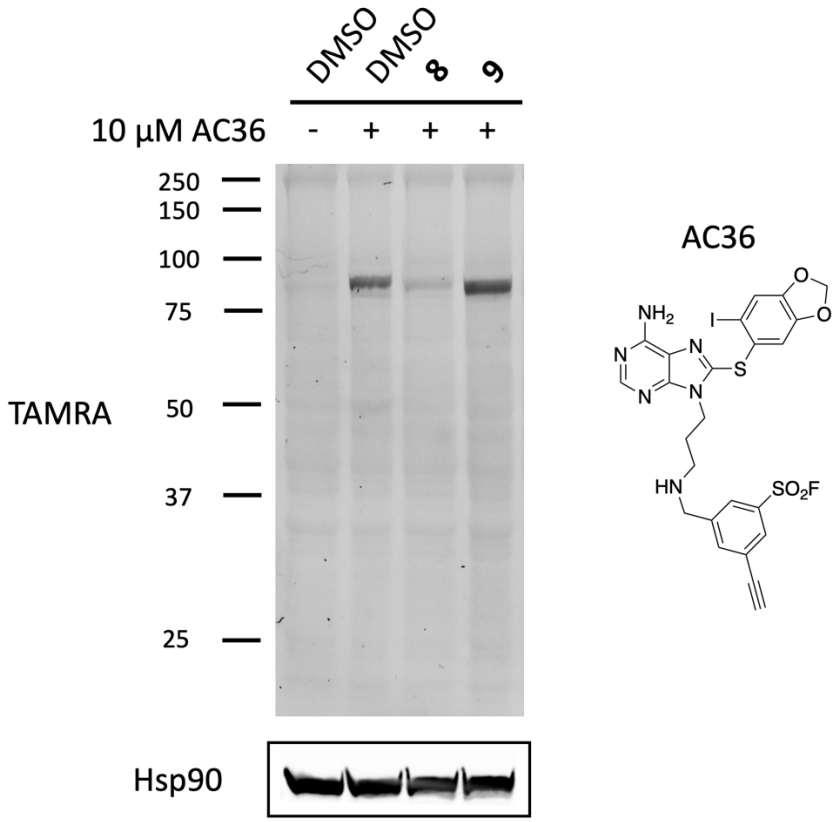


Figure 1.10: Competitive Hsp90 engagement assay in Skbr3 cells.

Skbr3 cells were treated with DMSO, 10 µM compound **8**, or 10 µM compound **9** for 30 min at 37 °C. Clickable probe AC36 (10 µM, referred to as compound **2** in ref. 1) was then added to the cells for 90 min at 37 °C. Cell lysates were prepared, subjected to click conjugation with TAMRA-picolyl azide, and analyzed by in-gel fluorescence (TAMRA) and immunoblotting (Hsp90).

and aldehyde relative to the linker dramatically affects the residence time of the resulting solvent-exposed imine adduct. The 1,2,3-salicylaldehyde regiochemistry, coupled with the addition of two fluorines, led to the discovery of a quasi-irreversible inhibitor that engages a surface lysine. More generally, our study demonstrates the potential of engaging surface lysines with reversible salicylaldehyde-based inhibitors that exhibit prolonged residence times. Such surface-exposed lysines are frequently found within striking distance of small molecule binding sites.

Methods

Protein expression and purification. Hsp90 N-terminal domain (NTD) was expressed and purified as previously described[11].

Time course of Hsp90 NTD modification. Compounds (10 μM) were added to 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$) at RT. At each time point an aliquot was removed and treated with 5 mM $NaBH_4$ for 10 min at RT, then quenched with 1% formic acid. Samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of modified Hsp90.

Time course of salicylaldehyde inhibitor dissociation from Hsp90 NTD. Compounds (10 μM) were added to 5 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$) at RT for 2 h. An aliquot was removed for determination of the initial percentage modification. The mixture was then diluted 10-fold

into labeling buffer containing 100 μM PU-H71 as a competitor. At each time point an aliquot was removed and treated with 5 mM $NaBH_4$ for 10 min at RT, then quenched with 1% formic acid. Samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of modified protein.

Time course of noncovalent inhibitor dissociation from Hsp90 NTD. Benzyl alcohol **9** (10 μM), PU-H71 (10 μM), or DMSO was added to 5 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$) at RT for 15 min. The mixture was then diluted 10-fold into labeling buffer containing excess salicylaldehyde **7** (20 μM). At each time point an aliquot was removed and treated with 5 mM $NaBH_4$ for 10 min at RT, then quenched with 1% formic acid. Samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of unmodified protein.

Co-crystallization and structure determination of compound 6 bound to Hsp90 NTD. Hsp90 NTD was covalently modified by incubating 10 μM **6** with 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$). Complete modification was confirmed by LC-MS. Modified protein was buffer exchanged into 50 mM HEPES pH 8.0, 50 mM KCl, 1 mM DTT with 10% glycerol and concentrated to 20 mg/mL using a 10K MWCO filter (Millipore Amicon Ultra). Crystallization was performed using the hanging drop method with a 1:1 mixture of 20 mg/mL protein-compound adduct and crystallization solution containing 20% PEG-3350, 0.2 M $MgCl_2$, 0.1 M Bis-

Tris pH 6.2. Protein crystals formed within 48 h. X-ray diffraction data was collected at beamline 8.3.1 of the Advanced Light Source (Lawrence Berkley National Laboratory). The resulting X-ray diffraction data was processed using the XDS package[34], followed by molecular replacement with PHASER[35] using PDB code: 6U9B as a starting model. Structure refinement was performed using coot[36] and phenix.refine[35].

Effect of compounds 8 and 9 on phospho-AKT. Skbr3 cells were seeded in 6-well plates and grown in McCoy's 5a media (Gibco) supplemented with 10% fetal bovine serum. Cells were treated with 40 nM to 10 μM of compounds **8** or **9** for 6 h at 37 °C. Cells were washed twice with cold PBS and gently scraped into cold PBS. Cells were pelleted and excess PBS removed. Cells were lysed in 40 μL lysis buffer (0.1% NP-40, 100 mM HEPES, pH 7.4, 150 mM NaCl, 1x PhosSTOP, 1X Roche complete protease inhibitors without EDTA) for 30 min on ice, then cleared by centrifugation at 16,500 *g* for 30 min at 4 °C. Protein concentrations were normalized using the bicinchoninic acid (BCA) assay (Pierce). Lysates were denatured with 6x SDS-PAGE loading buffer and resolved by 10% SDS-PAGE (200 V, 50 min). Gels were transferred to 0.2 μm nitrocellulose membranes (Bio-Rad) using a Bio-Rad Criterion transfer system (70 V, 90 min). Membranes were blocked for 1 h at RT using blocking buffer (2% BSA, 0.1% NaN_3 in TBST). Membranes were incubated with primary antibodies diluted in blocking buffer for 16 h at 4 °C. Membranes were rinsed with TBST (3 x 5 min at RT) and incubated with secondary antibodies diluted in blocking buffer for 1 h at RT. Membranes were rinsed with TBST (1 x 5 min) then TBS (2 x 5 min) and imaged using a Li-Cor Odyssey system. Primary antibodies: Hsp90, Stressmarq SMC-135D, 1:1000; Akt, Cell

Signaling Technology #2920S, 1:5000; pAkt S473, Cell Signaling Technology #4060S, 1:5000; GAPDH, Santa Cruz Biotechnology sc-32233. Secondary antibodies: Goat antimouse (Licor Biosciences #926-32210) or Goat anti-rabbit (Licor Biosciences #926-32211) conjugated to IRDye 800CW, 1:10,000.

Cell proliferation assay. Skbr3 cells were seeded (5000 cells) in 100 µL McCoy's 5a medium (Gibco) supplemented with 10% fetal bovine serum in 96-well flat-bottom plates and left to adhere overnight. The medium was refreshed and the cells were treated with 10 µL of 10x drug stock (0.5% DMSO final) and then incubated for 72 h. AlamarBlue was added and incubated for 3 h. Fluorescence intensity was read out as a measure of cell viability using a Clariostar microplate reader. Data was normalized to the DMSO-treated plate average.

Hsp90 target engagement assay. Skbr3 cells were grown in McCoy's 5a medium (Gibco) supplemented with 10% fetal bovine serum. Cells were treated with DMSO, compound **8** (10 µM), or compound **9** (10 µM) for 30 min at 37 °C. Subsequently, 10 µM AC36 (clickable sulfonyl fluoride probe; compound **2** in ref. 1) was added, and cells were incubated for an additional 90 min at 37 °C. Cells were washed twice with cold PBS and then gently scraped into cold PBS. Cells were pelleted and the excess PBS was removed. Cells were lysed in 40 µL lysis buffer (0.1% NP-40, 100 mM HEPES, pH 7.4, 150 mM NaCl, 1x PhosSTOP, 1X Roche complete protease inhibitors without EDTA) for 30 min on ice and then cleared by centrifugation at 16,500 *g* for 30 min at 4 °C. Protein concentrations were normalized using the bicinchoninic acid (BCA) assay

(Pierce). Click chemistry conjugation with TAMRA-picolyl azide was performed for 1 h at RT (final concentrations: 1% SDS, 5 μM TAMRA-picolyl azide [Click Chemistry Tools, #1254], 1 mM TCEP, 100 μM TBTA, 1 mM CuSO4, 2% DMSO) and then quenched with 6x SDS sample buffer. Proteins were resolved by 10% SDS–PAGE (200 V, 50 min). Gels were washed with deionized water and scanned for TAMRA fluorescence (Typhoon imaging system, Molecular Dynamics) or transferred to 0.2 μm nitrocellulose membranes (Bio-Rad) using a Bio-Rad Criterion transfer system (70 V, 90 min). Membranes were blocked for 1 h at RT using blocking buffer (2% BSA, 0.1% NaN_3 in TBST). Membranes were incubated with primary antibody diluted in blocking buffer for 16 h at 4 °C (Hsp90, Stressmarq SMC-135D, 1:1000). Membranes were rinsed with TBST (3 x 5 min at RT) and incubated with secondary antibody diluted in blocking buffer for 1 h at RT (Licor Biosciences #926-32210, 1:10,000). Membranes were rinsed with TBST (1 x 5 min), then TBS (2 x 5 min), and imaged using a Li-Cor Odyssey system.

Synthetic Methods

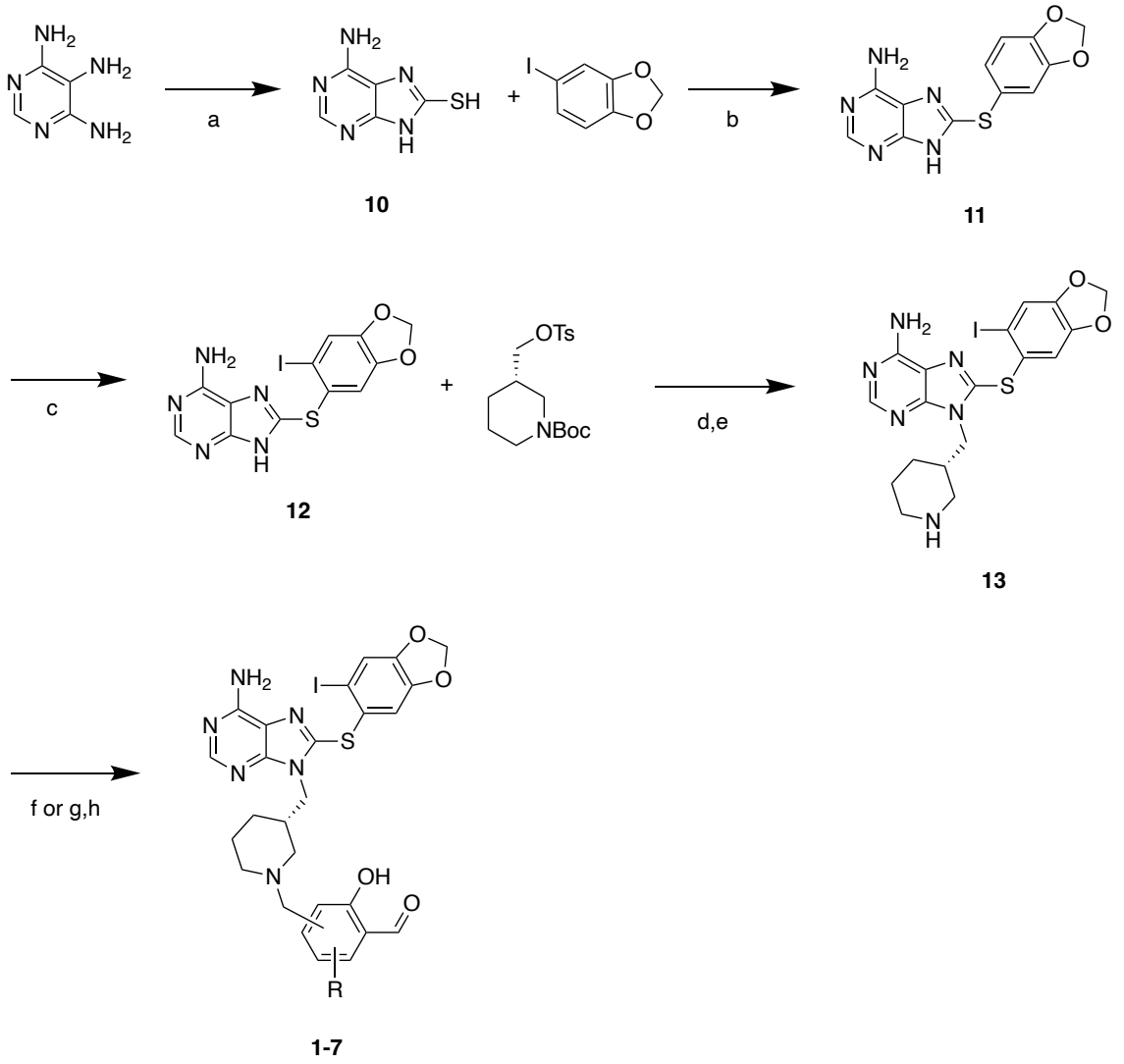
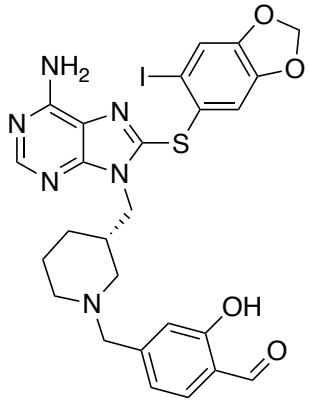


Figure 1.11. Synthetic scheme 1.1.

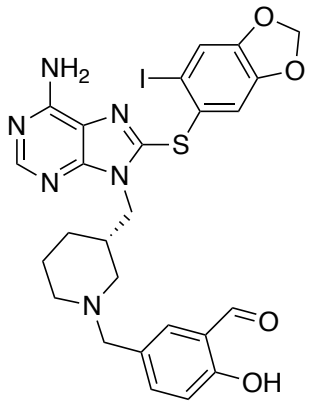
Scheme 1.1: a) $NaHCO_3$, EtOH, water, 60 °C, 3 days. b) CuI, neocuproine, NaO*t*Bu, DMF, 80 °C, for 16 h. c) NIS, TFA, ACN, RT, for 16 h. d) Cs_2CO_3, DMF, 80 °C, for 16 h. e) TFA:CH_2Cl_2 1:1, 0 °C, 1 h. f) DIPEA, DMF, RT, 1 h. g) $NaCNBH_3$, MeOH:$CH_2Cl_2$1:1, RT, for 16 h. h) HCl, THF, RT, 2 h. R = H or F.

(*S*)-4-((3-((6-Amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-2-hydroxybenzaldehyde (1): 13 (13 mg, 0.025



mmol) and DIPEA (18 µL, 0.10 mmol) were dissolved in DMF (0.5 mL), followed by the addition of a solution of 4-(bromomethyl)-2-hydroxybenzaldehyde (5.5 mg, 0.025 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 9.6 mg, 51%. ^{1}H NMR (400 MHz, DMSO) δ 10.18 (s, 1H), 8.15 (s, 1H), 7.57 (d, *J* = 8.0 Hz, 1H), 7.49 (s, 1H), 7.38 (s, 2H), 6.91 (s, 1H), 6.84 (d, *J* = 5.7 Hz, 2H), 6.06 (s, 2H), 4.12–3.97 (m, 2H), 3.43–3.35 (m, 2H), 2.56 (d, *J* = 10.5 Hz, 2H), 2.22–2.11 (m, 1H), 2.03–1.86 (m, 2H), 1.70–1.59 (m, 1H), 1.51–1.43 (m, 1H), 1.40–1.29 (m, 1H), 1.07–0.95 (m, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{26}IN_6O_4S^+$ [M + H]$^+$, 645.0775; found, 645.0793.

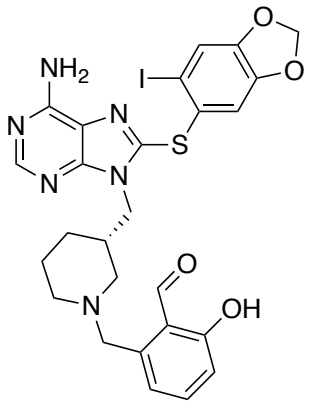
(*S*)-5-((3-((6-Amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-2-hydroxybenzaldehyde (2): 13 (8 mg, 0.016 mmol)



and DIPEA (12 µL, 0.069 mmol) were dissolved in DMF (0.5 mL), followed by the addition of a solution of 5-(chloromethyl)-2-hydroxybenzaldehyde (2.7 mg, 0.016 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 6.5 mg, 63%. ^{1}H NMR (400 MHz, DMSO) δ: 10.31 (s, 1H), 8.36 (s, 1H), 7.74 (d, *J* = 1.9 Hz, 1H), 7.52 (s,

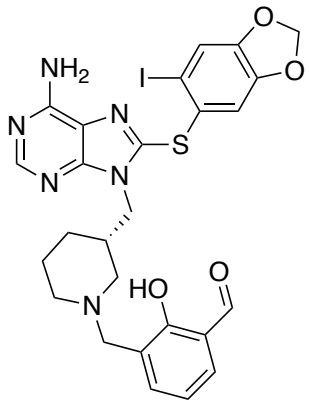
2H), 7.06 (d, *J* = 8.5 Hz, 1H), 7.02 (s, 1H), 6.10 (s, 2H), 4.21 (s, 2H), 4.13 (d, *J* = 6.6 Hz, 2H), 3.31 (t, *J* = 10.2 Hz, 2H), 2.77 (d, *J* = 9.5 Hz, 2H), 1.85 (d, *J* = 13.4 Hz, 1H), 1.70 (d, *J* = 10.9 Hz, 1H), 1.64–1.53 (m, 1H), 1.28–1.24 (m, 2H). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{26}IN_6O_4S^+$ $[M + H]^+$, 645.0775; found, 645.0793.

(*S*)-2-((3-((6-Amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-6-hydroxybenzaldehyde (3): 13 (20 mg, 0.039 mmol) and **3b** (15 mg, 0.063 mmol) were dissolved in a 1:1 MeOH:CH_2Cl_2 mixture (1 mL) and stirred for 15 min. $NaCNBH_3$ (5.0 mg, 0.079 mmol) was added and the reaction was left to stir at RT for 16 h. The reaction was then poured into sat. aqueous $NaHCO_3$ and extracted 3 times with ethyl acetate. The organic fractions were pooled and solvent was removed under vacuum.



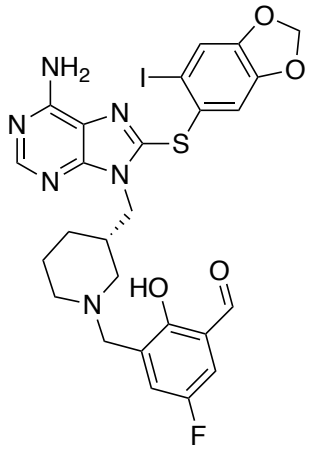
The crude mixture was redissolved in THF (1 mL), 1 M aqueous HCl (40 µL, 0.040 mmol) added, and the reaction left to stir for 16 h. When the reaction was complete by TLC, the solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 10.0 mg, 33%. ^{1}H NMR (400 MHz, DMSO) δ 11.57 (s, 1H), 10.35 (s, 1H), 8.14 (s, 1H), 7.47 (s, 1H), 7.45–7.35 (m, 3H), 6.89–6.81 (m, 3H), 6.06 (s, 2H), 4.07–3.95 (m, 2H), 3.70–3.60 (m, 2H), 2.58 (d, *J* = 9.8 Hz, 1H), 1.99 (dt, *J* = 24.3, 10.2 Hz, 3H), 1.66–1.50 (m, 2H), 1.48–1.40 (m, 1H), 1.34–1.24 (m, 1H), 1.07–0.96 (m, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{26}IN_6O_4S^+$ $[M + H]^+$, 645.0775; found, 645.0793.

(*S*)-3-((3-((6-Amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-2-hydroxybenzaldehyde (4): 13 (15 mg, 0.029 mmol) and **4b** (12 mg, 0.050 mmol) were dissolved in a 1:1 MeOH:CH_2Cl_2 mixture (1 mL) and stirred for 15 min. $NaCNBH_3$ (5.0 mg, 0.072 mmol) was added and the reaction was left to stir at RT for 16 h. The reaction was then poured into sat. aqueous $NaHCO_3$ and extracted 3 times with ethyl acetate. The organic fractions were pooled, and solvent was removed under vacuum.



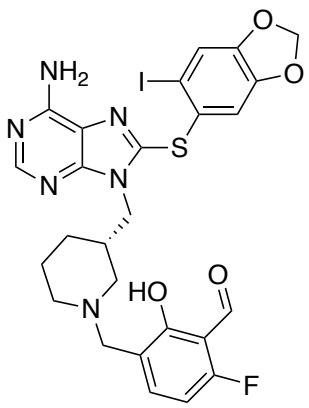
The crude mixture was redissolved in THF (1 mL), 1 M aqueous HCl (40 µL, 0.040 mmol) added, and the reaction left to stir. When the reaction was complete by TLC, the solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 4.5 mg, 20%. ^{1}H NMR (400 MHz, DMSO) δ 10.26 (s, 1H), 8.15 (s, 1H), 7.54 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.49 (s, 1H), 7.43–7.33 (m, 3H), 6.90–6.82 (m, 2H), 6.08–6.04 (m, 2H), 4.08–4.03 (m, 2H), 3.73–3.64 (m, 2H), 2.80–2.70 (m, 2H), 2.18–2.01 (m, 3H), 1.74–1.65 (m, 1H), 1.56–1.48 (m, 1H), 1.43–1.29 (m, 1H), 1.14–1.02 (m, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{26}IN_6O_4S^+$ [M + H]$^+$, 645.0775; found, 645.0793.

(*S*)-3-((3-((6-Amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-5-fluoro-2-hydroxybenzaldehyde (5): 13 (15 mg,



0.029 mmol) and DIPEA (22 µL, 0.13 mmol) were dissolved in DMF (0.5 mL), followed by the addition of a solution of **5c** (6.0 mg, 0.032 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 2 h. Solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 11.0 mg, 55%. ^{1}H NMR (400 MHz, DMSO) δ 10.10 (s, 1H), 8.34 (s, 1H), 7.73–7.66 (m, 2H), 7.52 (s, 1H), 7.02 (s, 1H), 6.10 (s, 2H), 4.29 (s, 2H), 4.14 (d, *J* = 6.1 Hz, 2H), 3.35 (s, 2H), 2.91 (s, 2H), 1.84 (d, *J* = 10.7 Hz, 1H), 1.69 (d, *J* = 10.5 Hz, 2H), 1.36–1.14 (m, 2H). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{25}FIN_6O_4S^+$ $[M + H]^+$, 663.0681; found, 663.0687.

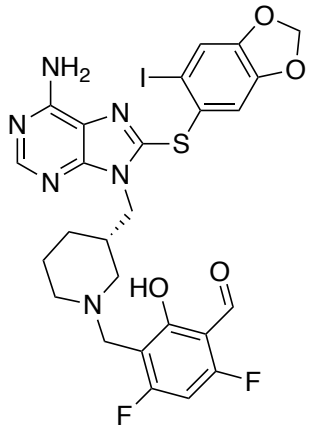
(*S*)-3-((3-((6-Amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-6-fluoro-2-hydroxybenzaldehyde (6): 13 (20 mg,



0.039 mmol) and **6b** (13 mg, 0.051 mmol) were dissolved in a 1:1 MeOH:CH_2Cl_2 mixture (1 mL) and stirred for 15 min. $NaCNBH_3$ (5.0 mg, 0.072 mmol) was added and the reaction was left to stir at RT for 16 h. The reaction was then poured into sat. aqueous $NaHCO_3$ and extracted 3 times with ethyl acetate. The organic fractions were pooled, and solvent was removed under vacuum. The crude mixture was redissolved in THF (1 mL), 1 M aqueous HCl (40 µL, 0.040 mmol) added, and the reaction left to stir. When the reaction was complete by

TLC, the solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 5.1 mg, 19%. ^{1}H NMR (400 MHz, DMSO) δ 10.23 (s, 1H), 8.14 (s, 1H), 7.49 (s, 1H), 7.43–7.33 (m, 3H), 6.83 (s, 1H), 6.61 (dd, *J* = 10.6, 8.5 Hz, 1H), 6.08–6.04 (m, 2H), 4.06–4.03 (m, 2H), 3.59 (s, 2H), 2.72 (d, *J* = 9.1 Hz, 2H), 2.16–2.00 (m, 3H), 1.71–1.63 (m, 1H), 1.54–1.47 (m, 1H), 1.40–1.30 (m, 1H), 1.12–1.01 (m, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{25}FIN_6O_4S^+$ [M + H]$^+$, 663.0681; found, 663.0687.

(*S*)-3-((3-((6-Amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-4,6-difluoro-2-hydroxybenzaldehyde (7): 13 (10



mg, 0.020 mmol) and DIPEA (15 µL, 0.086 mmol) were dissolved in DMF (0.5 mL), followed by the addition of a solution of **7c** (4.2 mg, 0.020 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 8.8 mg, 55%. ^{1}H NMR (400 MHz, DMSO) δ 10.12 (s, 1H), 8.14 (s, 1H), 7.48 (s, 1H), 7.40 (s, 2H), 6.86 (s, 1H), 6.35 (t, *J* = 11.0 Hz, 1H), 6.06 (s, 2H), 4.06 (d, *J* = 7.1 Hz, 2H), 3.80 (s, 2H), 2.94 (d, *J* = 11.2 Hz, 2H), 2.45–2.35 (m, 2H), 2.27–2.19 (m, 1H), 1.76–1.70 (m, 1H), 1.57–1.52 (m, 1H), 1.46–1.40 (m, 1H), 1.15–1.09 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 192.05 (d, *J* = 6.8 Hz), 167.68 (dd, *J* = 120.4, 18.1 Hz), 165.61 (dd, *J* = 123.5, 18.0 Hz), 163.37 – 163.01 (m), 152.79 (s), 151.26 (s), 149.62 (s), 149.37 (s), 149.33 – 149.28 (m), 147.11 (s), 127.84 (s), 119.73 (s), 119.21 (s), 117.93 (s), 115.58 (s), 112.61 (s), 108.44 (d, *J* = 11.4 Hz), 103.40

– 102.60 (m), 96.96 (t, *J* = 26.8 Hz), 92.50 (s), 54.13 (s), 52.02 (s), 47.23 (s), 46.31 (s), 34.97 (s), 25.52 (s), 22.19 (s). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{24}F_2IN_6O_4S^+$ [M + H]$^+$, 681.0587; found, 681.0597.

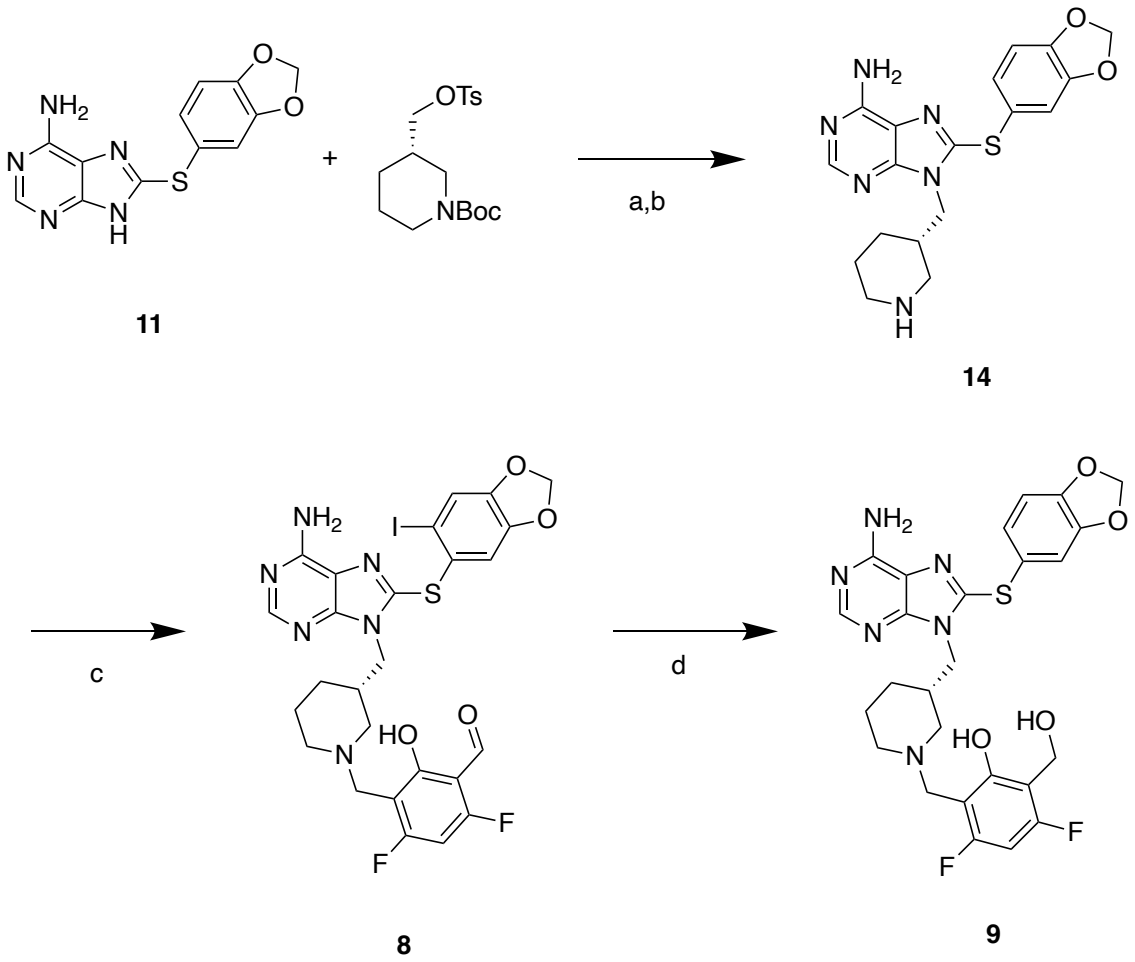
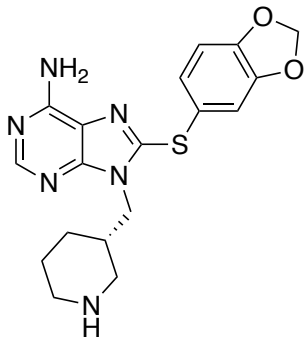


Figure 1.12. Sythetic scheme 1.2.

Scheme 1.2: a) Cs_2CO_3, DMF, 80 °C, 16 h. b) TFA:CH_2Cl_2 1:1, 0 °C, 1 h. c) DIPEA, DMF, RT, 1 h. d) $NaBH_4$, MeOH:CH_2Cl_2 1:1, RT, 1 h.

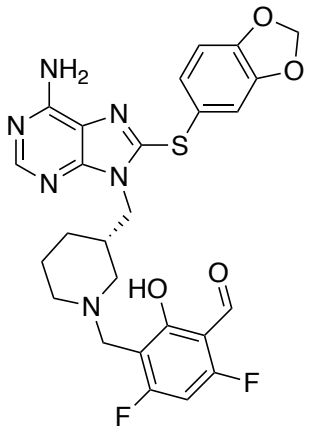
(*S*)-8-(Benzo[*d*][1,3]dioxol-5-ylthio)-9-(piperidin-3-ylmethyl)-9*H*-purin-6-amine (14):



Tert-butyl(3S)-3-(hydroxymethyl)piperidine-1-carboxylate (1.4 g, 6.5 mmol) and *p*-toluenesulfonyl chloride (1.5 g, 7.8 mmol) were dissolved in CH_2Cl_2 (20 mL), then pyridine (440 µL, 7.8 mmol) was added. The reaction was left to stir for 16 h at RT. The reaction was then diluted with saturated ammonium chloride and extracted 3 times with dichloromethane. The combined organic extracts were washed with brine, dried over sodium sulfate and concentrated under vacuum. The resulting tosylate was used without further purification. **11** (205 mg, 0.71 mmol) and Cs_2CO_3 (283 mg, 0.78 mmol) were suspended in DMF (4 mL) and heated to 80 °C under Ar. The tosylate (718 mg, 2.0 mmol) was dissolved in DMF (2 mL) and added to the reaction dropwise. The reaction was left to stir at 80 °C for 16 h. The solvent was removed under vacuum and the residue was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). The resulting mixture of purine regioisomers was concentrated, re-dissolved in dichloromethane (4 mL) and set to stir on ice. Trifluoroacetic acid (4 mL) was then added dropwise. After 1 h, the reaction mixture was concentrated and purified by silica gel chromatography using a 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 75 mg, 28%. ^{1}H NMR (400 MHz, DMSO) δ 8.71–8.59 (m, *J* = 9.4 Hz, 1H), 8.32–8.18 (m, *J* = 10.7 Hz, 1H), 8.15 (s, 1H), 7.41 (s, 2H), 7.09 (d, *J* = 1.8 Hz, 1H), 7.02 (dd, *J* = 8.1, 1.8 Hz, 1H), 6.96 (d, *J* = 8.1 Hz, 1H), 6.07 (s, 2H), 4.13–4.07 (m, 2H), 3.16 (dd, *J* = 22.8, 11.4 Hz, 2H), 2.84–2.72 (m, 2H), 2.36–2.26 (m, 1H), 1.79 (d, *J* = 14.0 Hz, 1H), 1.65–

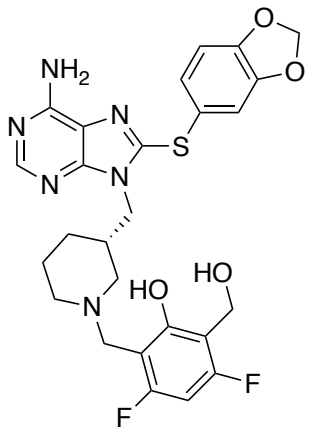
1.46 (m, 2H), 1.31 (ddd, *J* = 24.2, 12.3, 3.4 Hz, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{18}H_{21}N_6O_2S^+$ $[M + H]^+$, 385.1441; found, 385.1367.

(*S*)-3-((3-((6-Amino-8-(benzo[*d*][1,3]dioxol-5-ylthio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-4,6-difluoro-2-hydroxybenzaldehyde (8): 14 (40 mg, 0.10 mmol) and

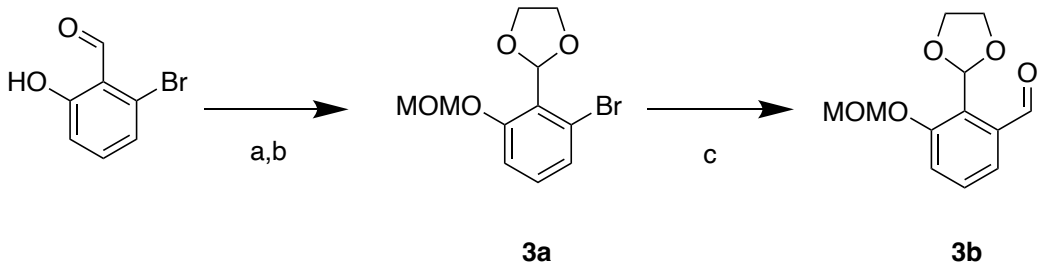


DIPEA (60 µL, 0.35 mmol) were dissolved in DMF (1 mL), followed by the addition of a solution of **7c** (18 mg, 0.087 mmol) in DMF (1 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 13.0 mg, 24%. ^{1}H NMR (400 MHz, DMSO) δ 10.13 (s, 1H), 8.12 (s, 1H), 7.35 (s, 2H), 7.05 (d, *J* = 1.7 Hz, 1H), 6.97 (d, *J* = 1.8 Hz, 1H), 6.94 (s, 1H), 6.38 (t, *J* = 10.9 Hz, 1H), 6.05 (s, 2H), 4.07 (d, *J* = 7.2 Hz, 2H), 3.84 (s, 2H), 2.99–2.92 (m, 2H), 2.41 (d, *J* = 11.4 Hz, 1H), 2.29–2.21 (m, 1H), 1.76–1.69 (m, 1H), 1.53 – 1.43 (m, 2H), 1.27–1.23 (m, 1H), 1.16–1.09 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 192.04 (d, *J* = 6.5 Hz), 167.69 (dd, *J* = 124.9, 17.4 Hz), 165.62 (dd, *J* = 127.7, 17.9 Hz), 163.25 (dd, *J* = 10.4, 6.7 Hz), 151.85 (s), 151.11 (s), 148.94 (s), 148.73 (s), 148.45 (s), 148.32 (s), 127.22 (s), 121.46 (s), 119.36 (s), 117.88 (s), 115.54 (s), 113.17 (s), 109.77 (s), 108.46 (d, *J* = 11.3 Hz), 103.06 (d, *J* = 16.5 Hz), 102.39 (s), 96.96 (t, *J* = 26.7 Hz), 54.08 (s), 51.96 (s), 47.13 (s), 46.22 (s), 34.89 (s), 25.41 (s), 22.09 (s). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{25}F_2N_6O_4S^+$ $[M + H]^+$, 555.1621; found, 555.1615.

(*S*)-2-((3-((6-Amino-8-(benzo[*d*][1,3]dioxol-5-ylthio)-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-3,5-difluoro-6-(hydroxymethyl)phenol (9): 8 (10 mg, 0.018 mmol) was



dissolved in a 1:1 MeOH:CH_2Cl_2 mixture (1 mL) and $NaBH_4$ (1.0 mg, 0.026 mmol) was added. The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by HPLC (5-95% acetonitrile in water with 0.1% TFA). Yield: 7.1 mg, 82%. ^{1}H NMR (400 MHz, DMSO) δ 8.13 (s, 1H), 7.35 (s, 2H), 7.04 (d, *J* = 1.8 Hz, 1H), 6.98 (dd, *J* = 8.1, 1.8 Hz, 1H), 6.93 (d, *J* = 8.0 Hz, 1H), 6.55 (t, *J* = 10.2 Hz, 1H), 6.09–6.04 (m, 2H), 4.39 (s, 2H), 4.10–4.03 (m, 2H), 3.67 (s, 2H), 2.80–2.69 (m, 2H), 2.18–2.04 (m, 3H), 1.73–1.64 (m, 1H), 1.51–1.43 (m, 1H), 1.38–1.28 (m, 1H), 1.12–1.01 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 152.67 (s), 151.51 – 150.73 (m), 149.54 (s), 148.77 (d, *J* = 15.9 Hz), 147.63 (s), 126.97 (d, *J* = 9.0 Hz), 121.75 (d, *J* = 7.6 Hz), 120.26 (s), 119.42 (s), 117.91 (s), 115.57 (s), 113.22 (s), 112.96 (d, *J* = 5.5 Hz), 109.75 (s), 102.37 (s), 93.33 – 91.95 (m), 55.80 (s), 54.45 (s), 54.04 (s), 52.89 (s), 51.84 (d, *J* = 25.4 Hz), 48.38 (s), 46.08 (s), 34.93 (s), 25.41 (s), 22.28 (s). HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{27}F_2N_6O_4S^+$ [M + H]$^+$, 557.1777; found, 557.1747.

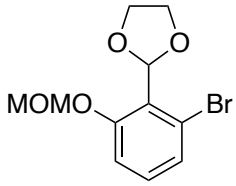


3a **3b**

Figure 1.13. Synthetic scheme 1.3.

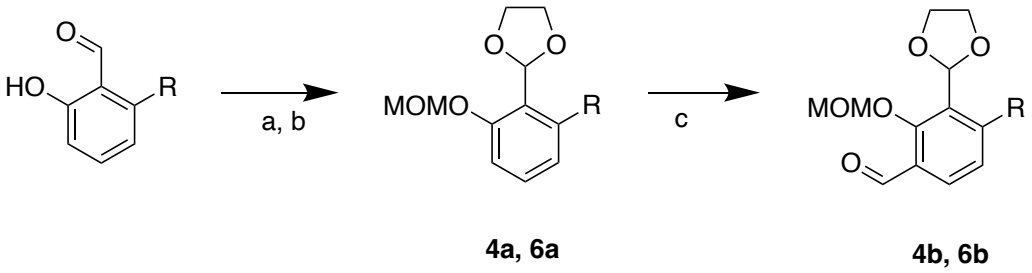
Scheme 1.3: a) triethyl orthoformate, Bu_4NBr_3, ethylene glycol, 45°C, for 16 h. b) MOMBr, DIPEA, CH_2Cl_2, RT, 1h. c) *i*PrMgCl·LiCl, RT, 3 h.

2-(2-Bromo-6-(methoxymethoxy)phenyl)-1,3-dioxolane (3a): 2-bromo-6-



hydroxybenzaldehyde (5.0 g, 25 mmol), triethyl orthoformate (8.3 mL, 50 mmol), and Bu_4NBr_3 (240 mg, 0.05 mmol) were suspended in ethylene glycol (28 mL, 500 mmol) and heated to 45 °C. The reaction was left to stir for 16 h then poured in sat. aqueous $NaHCO_3$ and extracted 3 times with diethyl ether. The organic fractions were pooled, dried over Na_2SO_4, and the solvent removed under vacuum. The crude mixture was redissolved in CH_2Cl_2 (100 mL) and DIPEA (8.7 mL, 37 mmol) was added. MOMBr (8.7 mL, 37 mmol) was added to the reaction dropwise. The reaction was poured into sat aqueous $NaHCO_3$ and extracted 3 times with diethyl ether. The organic fractions were pooled, dried over Na_2SO_4, and the solvent was removed under vacuum. The product was purified by silica gel chromatography in 0-30% ethyl acetate in hexane + 0.1% triethylamine. Yield: 2.0 g, 28% yield. ^{1}H-NMR (400 MHz, $CDCl_3$) δ: 7.25 (dd, *J* = 7.6, 1.2, 1H), 7.15-7.01 (m, 2H), 6.46 (s, 1H), 5.19 (s, 2H), 4.31-4.23 (m, 2H), 4.08-4.00 (m, 2H), 3.48 (s, 3H). Product does not ionize by ESI.

2-(1,3-Dioxolan-2-yl)-3-(methoxymethoxy)benzaldehyde (3b): *i*PrMgCl·LiCl (0.70 mL, 0.91 mmol) was added to **3a** (220 mg, 0.76 mmol) neat and stirred at RT for 3 h. The reaction was cooled to -10°C and DMF (87 µL, 1.1 mmol) was added dropwise. The reaction was allowed to warm to RT over 30 min, then poured into sat. aqueous NH_4Cl and extracted 3 times with ethyl acetate. The organic fractions were pooled, washed with water twice, then brine once, and then dried over Na_2SO_4. The solvent was removed under vacuum. The product was purified by silica gel chromatography in 0-40% ethyl acetate in hexane + 0.1% triethylamine. Yield: 83 mg, 35% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.52 (d, *J* = 0.8 Hz, 1H), 7.54 (dd, *J* = 7.6, 1.4 Hz, 1H), 7.41-7.36 (m, 1H), 7.32 (dd, *J* = 8.3, 1.3 Hz, 1H), 6.33 (s, 1H), 5.18 (s, 2H), 4.24-4.14 (m, 2H), 4.07-3.97 (m, 2H), 3.45 (s, 4H). HRMS (ESI-TOF) m/z calcd. $C_{12}H_{14}O_5Na^+$ [M + Na]$^+$, 261.0733; found, 261.0728.

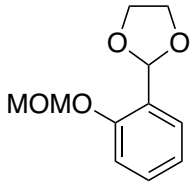


4a, 6a **4b, 6b**

Figure 1.14. Synthetic scheme 1.4.

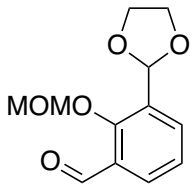
Scheme 1.4: a) triethyl orthoformate, Bu_4NBr_3, ethylene glycol, 45°C, 16 h. b) MOMBr, DIPEA, CH_2Cl_2, RT, 1h. c) *n*BuLi, THF, 0 °C, 2 h. R = H or F

2-(2-(Methoxymethoxy)phenyl)-1,3-dioxolane (4a): Salicylaldehyde (1.0 g, 8.2 mmol),



triethyl orthoformate (2.7 mL, 16 mmol), and Bu_4NBr_3 (39 mg, 0.081 mmol) were suspended in ethylene glycol (4.6 mL, 82 mmol) and heated to 50°C. The reaction was left to stir for 16 h, then poured in sat. aqueous $NaHCO_3$ and extracted 3 times with ethyl acetate. The organic fractions were pooled, dried over Na_2SO_4, and the solvent was removed under vacuum. The crude mixture was redissolved in CH_2Cl_2 (80 mL) and DIPEA (2.9 mL, 17 mmol) was added. MOMBr (1.0 mL, 12 mmol) was added to the reaction dropwise. Additional portions of DIPEA (2.9 mL, 17 mmol) and MOMBr (1.0 mL, 12 mmol) were added every hour until the reaction reached 100% completion by TLC. The reaction was then poured into sat. aqueous $NaHCO_3$ and extracted 3 times with CH_2Cl_2. The organic fractions were pooled, dried over Na_2SO_4, and the solvent was removed under vacuum. The product was purified by silica gel chromatography in 0-30% ethyl acetate in hexane + 0.1% triethylamine. Yield: 560 mg, 33% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.57 (dd, J = 7.6, 1.6 Hz, 1H), 7.37–7.27 (m, 1H), 7.15 (d, J = 8.2 Hz, 1H), 7.06 (t, J = 7.5 Hz, 1H), 6.22 (s, 1H), 5.26 (s, 2H), 4.25–3.97 (m, 4H), 3.52 (s, 3H). Product does not ionize by ESI.

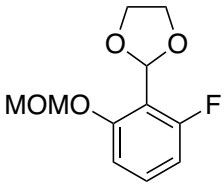
3-(1,3-Dioxolan-2-yl)-2-(methoxymethoxy)benzaldehyde (4b): 4a (450 mg, 2.1



mmol) was dissolved in THF (50 mL) and cooled to 0°C. *n*BuLi (1.3 mL, 2.6 mmol) was added and the reaction was left to stir at 0 °C for 2 h, then the reaction was cooled to -78°C and DMF (250 µL, 3.2 mmol) was added dropwise. The reaction was allowed to warm to RT and stirred for 1 h. The reaction was poured into sat. aqueous NH_4Cl and extracted 3 times with diethyl ether.

The organic fractions were pooled, washed with water twice, then brine once, and then dried over Na_2SO_4. The solvent was removed under vacuum. The product was purified by silica gel chromatography in 0-50% ethyl acetate in hexane + 0.1% triethylamine. Yield: 143 mg, 36% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.32 (s, 1H), 7.85 (dd, J = 7.7, 1.8 Hz, 1H), 7.81 (dd, J = 7.6, 1.8 Hz, 1H), 7.27 (t, J = 7.6 Hz, 1H), 6.09 (s, 1H), 5.13 (s, 2H), 4.30–3.91 (m, 4H), 3.55 (s, 3H). HRMS (ESI-TOF) m/z calcd. $C_{12}H_{14}O_5Na^+$ [M + Na]$^+$, 261.0733; found, 261.0677.

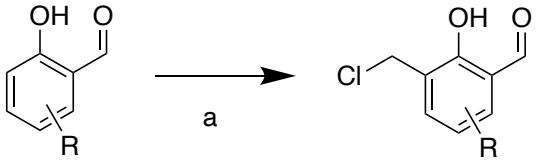
2-(2-Fluoro-6-(methoxymethoxy)phenyl)-1,3-dioxolane (6a): 2-Fluoro-6-



hydroxybenzaldehyde (1.0 g, 7.1 mmol), triethyl orthoformate (2.4 mL, 14 mmol), and Bu_4NBr_3 (34 mg, 0.071 mmol) were suspended in ethylene glycol (4.0 mL, 71 mmol) and heated to 45°C. The reaction was left to stir for 16 h then poured in sat. aqueous $NaHCO_3$ and extracted 3 times with diethyl ether. The organic fractions were pooled, dried over Na_2SO_4, and the solvent was removed under vacuum. The crude mixture was redissolved in CH_2Cl_2 (75 mL) and DIPEA (2.5 mL, 14 mmol) was added. MOMBr (890 µL, 11 mmol) was added to the reaction dropwise. Additional portions of DIPEA (2.5 mL, 14 mmol) and MOMBr (890 µL, 11 mmol) were added every hour until the reaction reached 100% completion by TLC. The reaction was poured into sat. aqueous $NaHCO_3$ and extracted 3 times with diethyl ether. The organic fractions were pooled, dried over Na_2SO_4, and the solvent was removed under vacuum. The product was purified by silica gel chromatography in 0-40% ethyl acetate in hexane + 0.1% triethylamine. Yield: 615 mg, 38% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.23 (td, *J* = 8.4, 6.4 Hz, 1H), 6.90 (d, *J* = 8.5 Hz, 1H), 6.73 (dd, *J* =

10.1, 8.7 Hz, 1H), 6.38 (d, *J* = 0.9 Hz, 1H), 5.19 (s, 2H), 4.25–4.15 (m, 2H), 4.05–3.95 (m, 2H), 3.47 (s, 3H). Product does not ionize by ESI.

3-(1,3-Dioxolan-2-yl)-4-fluoro-2-(methoxymethoxy)benzaldehyde (6b): 6a (190 mg, 0.83 mmol) was dissolved in THF (15 mL) and cooled to 0°C. *n*BuLi (510 µL, 1.0 mmol) was added and the reaction was left to stir at 0 °C for 2 h, then the reaction was cooled to -78°C and DMF (95 µL, 1.3 mmol) was added dropwise. The reaction was allowed to warm to RT and stirred for 1 h. The reaction was poured into sat. aqueous NH_4Cl and extracted 3 times with diethyl ether. The organic fractions were pooled, washed with water twice, then brine once, and then dried over Na_2SO_4. The solvent was removed under vacuum. The product was purified by silica gel chromatography in 0-50% ethyl acetate in hexane + 0.1% triethylamine. Yield: 60 mg, 28% yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.24 (d, *J* = 0.6 Hz, 1H), 7.91 (dd, *J* = 8.8, 6.3 Hz, 1H), 7.00 (t, *J* = 9.3 Hz, 1H), 6.24 (d, *J* = 1.0 Hz, 1H), 5.14 (s, 2H), 4.28–4.18 (m, 2H), 4.09–4.00 (m, 2H), 3.58 (s, 3H). HRMS (ESI-TOF) m/z calcd. $C_{12}H_{13}FO_5Na^+$ [M + Na]$^+$, 279.0639; found, 279.0555.

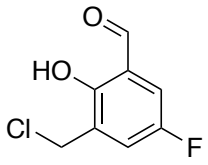


5c, 7c

Figure 1.15. Synthetic scheme 1.5.

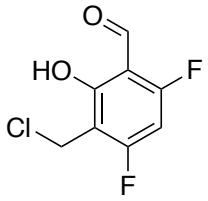
Scheme 1.5: a) paraformaldehyde, H_2SO_4, HCl, 70 °C. R = F

3-(Chloromethyl)-5-fluoro-2-hydroxybenzaldehyde (5c): 5-Fluoro-2-



hydroxybenzaldehyde (500 mg, 3.6 mmol) and paraformaldehyde (220 mg, 7.2 mmol) were suspended in 37% aqueous HCl (1.5 mL) and heated to 70°C, followed by the addition of 2 drops of H_2SO_4. The reaction was allowed to stir at 70°C for 3 days. The reaction was allowed to cool to RT, poured into water, and extracted 3 times with ethyl acetate. The organic layers were pooled, dried over Na_2SO_4, and the solvent was removed under vacuum. The crude product was purified by recrystallization in hexane. Yield: 237 mg, 35%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 11.24 (s, 1H), 9.89 (s, 1H), 7.47 (dd, *J* = 8.5, 3.1 Hz, 1H), 7.31–7.24 (m, 1H), 4.69 (s, 2H). Product does not ionize by ESI.

3-(Chloromethyl)-4,6-difluoro-2-hydroxybenzaldehyde (7c): 2,4-difluoro-6-



hydroxybenzaldehyde (330 mg, 2.1 mmol) and paraformaldehyde (130 mg, 4.3 mmol) were suspended in 37% aqueous HCl (1 mL) and heated to 70°C, followed by the addition of 2 drops of concentrated H_2SO_4. The reaction was allowed to stir at 70°C for 16 h. The reaction was allowed to cool to RT, poured into water, and extracted 3 times with ethyl acetate. The organic layers were pooled, dried over Na_2SO_4, and the solvent was removed under vacuum. Product was purified by recrystallization in hexane. Yield: 104 mg, 24%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 12.31 (d, *J* = 1.7 Hz, 1H), 10.17 (s, 1H), 6.51–6.44 (m, 1H), 4.64 (d, *J* = 1.3 Hz, 2H). Product does not ionize by ESI.

Chapter 2 : Fragment salicylaldehydes target surface-exposed lysines on Hsp90

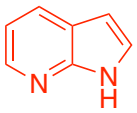
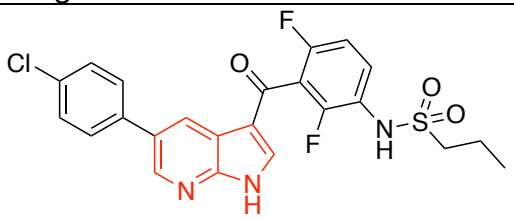
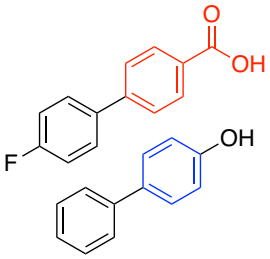
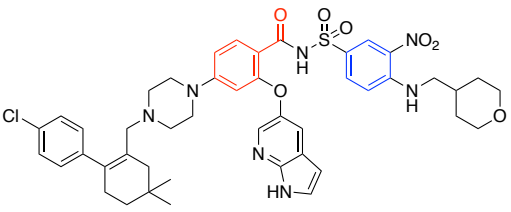
Abstract

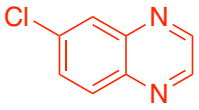
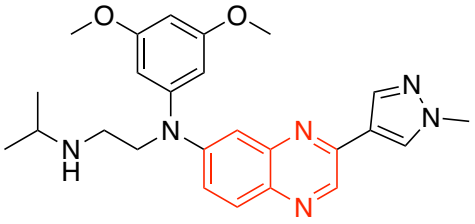
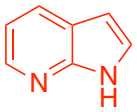
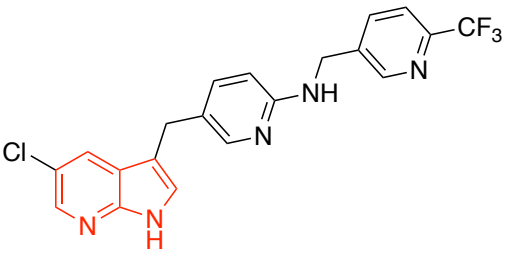
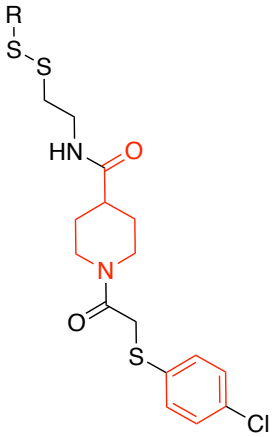
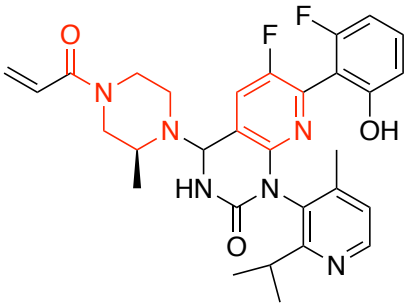
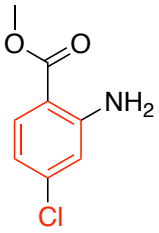
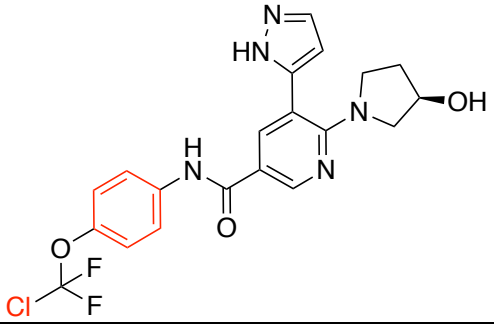
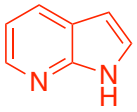
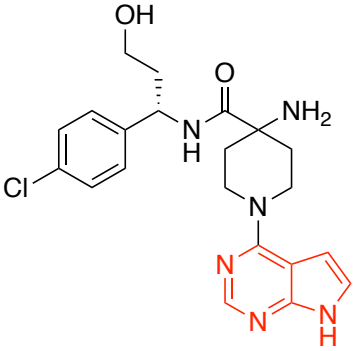
Fragment-based drug discovery allows for small molecule hit identification through a rapid and efficient exploration of chemical space. The use of covalent fragment libraries simplifies identification of hit fragments because unequivocal verification of covalent adduct formation is, in principle, straightforward. While most methods of covalent fragment screening have focused on cysteine-targeting electrophiles, methods for lysine-targeted fragment screening have also been developed on both a target-focused and proteome-wide scale. Salicylaldehydes have the potential to be used as lysine-targeted covalent fragments, and thus we were motivated to test the ability of salicylaldehyde fragments to specifically label a residue on a target protein. Herein, we describe a series of salicylaldehydes appended to an adenosine fragment. I screened these compounds against the ATPase domain of Hsp90, which has a total of 15 lysines. One compound, JW-2-128, completely and specifically modified Hsp90 at low micromolar concentrations in less than 2 h and showed an extended residence time relative to the non-covalent inhibitor PU-H71. JW-2-128 primarily modifies Lys58, as shown by a high-resolution crystal structure. Surprisingly, JW-2-128 efficiently and specifically modified a distinct lysine, likely Lys112, when Lys58 was mutated to arginine. These results suggest that the small size of the adenosine fragment in JW-2-128 can facilitate a conformational change in Hsp90, resulting in modification of a previously inaccessible lysine residue.

Introduction

Small molecule fragment-based screens are a powerful method of hit identification for small molecule drug development. Compared to traditional high-throughput screening, fragment-based screening focuses on molecules with lower complexity, typically 20 or fewer heavy atoms and less than 300 Da in molecular weight, allowing compound libraries to cover a larger proportion of chemical space[37,38]. Following hit identification, fragment molecules are further developed and optimized into high affinity ligands. Fragment-based drug discovery has proven its success with seven clinically approved drugs[39–47] (**Table 2.1**) tracing their origins back to a fragment hit and more than 50 others in various stages of clinical trials[48]. Despite its successes, fragment-based screening comes with notable drawbacks. Fragments typically have very weak binding affinities, often in the low millimolar to high micromolar range, and developing techniques capable of detecting specific, yet weak binding is a major hurdle.

Table 2-1: List of clinically approved drugs developed from an initial fragment hit.

Drug Name	Target	Fragment Hit	Drug Structure
Vemurafenib (Zelboraf)	B-RAF^{V600E}		
Venetoclax (Venclexta/ Venclyxto)	BCL-2		

Drug Name	Target	Fragment Hit	Drug Structure
Erdafitinib (Balversa)	FGFR1-4		
Pexidartinib (Turalio)	CSF1R, KIT		
Sotorasib (Lumakras/ Lumykras)	KRAS^{G12C}		
Asciminib (Scemblix)	BCR-ABL1		
Capivasertib (Truqap)	AKT		

Covalent-fragment screening simplifies the detection of hits, as identifying covalent binders is straightforward in principle, through the use of methods such as intact-protein mass spectrometry. Additionally, the covalent bond can simplify the process of hit optimization, as it anchors the fragment in the binding pocket and reduces the chances of the fragment binding mode changing during scaffold elaboration. However, a major concern with screening irreversible covalent fragments lies in the specificity of the binding, as off-target nucleophilic residues can react with electrophiles if their reactivity is too high, especially if they are used at micromolar or higher concentrations. This also raises the concern of a screen selecting for the most reactive fragments rather than those with the best reversible binding affinity and optimal electrophile geometry and distance relative to the nucleophilic residue. Despite these concerns, irreversible covalent fragment screening has found success, identifying binders for the E3 ligase HOIP[49], PIN1[50], and other targets [51–53].

One method of covalent fragment screening, known as disulfide tethering, leverages a reversible covalent electrophile to rapidly screen fragment libraries. In this approach, a library of disulfide-containing fragments is incubated with the target protein where they can undergo disulfide exchange with a cysteine residue. The reversible nature of this covalent bond means that this method selects for fragments that are stabilized by non-covalent interactions with the protein in addition to the covalent bond. Sotorasib, a clinically approved drug for KRAS G12C positive non-small cell lung cancer, originated from a disulfide tethering hit identified by Ostrem et al[43,46,47](**Fig 2.1**). Disulfide tethering has also been utilized in the identification of a PDK1 inhibitor[54] and 14-3-3/client protein-protein interaction stabilizers[55], among others[56–60]. Disulfide

fragments, however, are not suitable as cellular probes, as the disulfide is not stable to biological thiols like glutathione, which is present at millimolar concentrations within cells[61,62]. This necessitates further development of the fragments to include a more suitable electrophile, such as an acrylamide, which is not always straightforward due to the differing geometric constraints of these electrophiles relative to a disulfide.

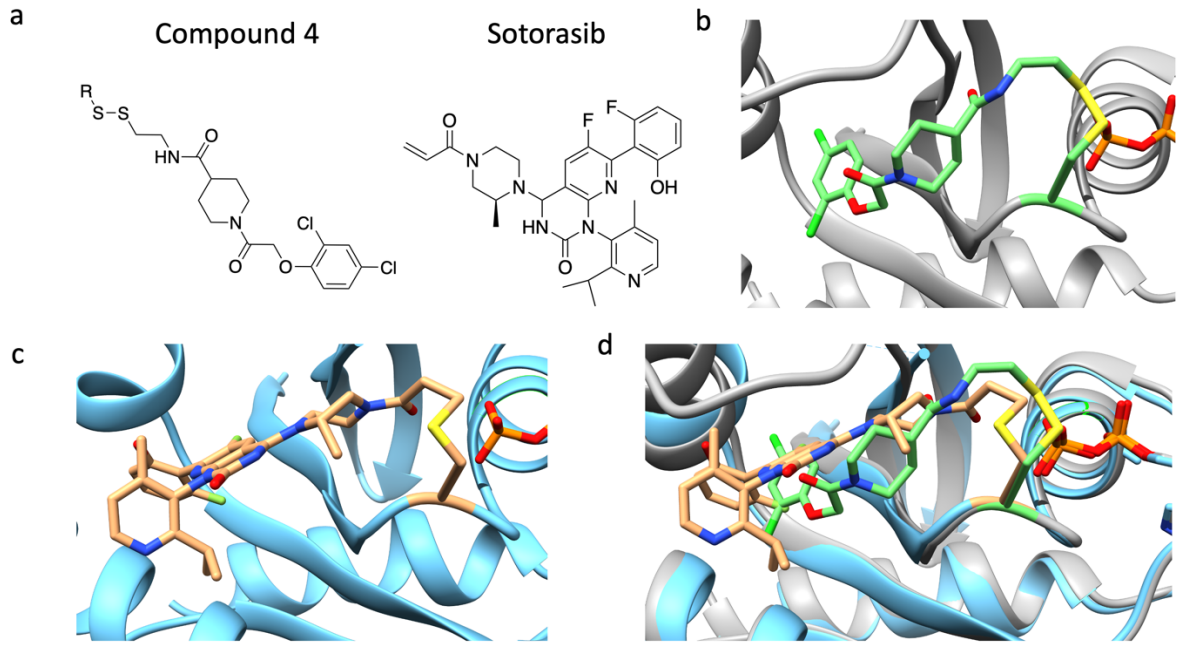


Figure 2.1. Crystal structures of KRAS G12C inhibitors compound 4 and sotorasib.

(a) Structure of disulfide compound **4**, a derivative of the initial fragment hit, and sotorasib, the final clinically approved drug. (b) Crystal structure of disulfide compound **4** covalently bound to (PDB: 4LV6). (c) Crystal structure of sotorasib covalently bound to HRAS G12C (PDB: 8TLR). (d) Overlay of disulfide compound **4** and sotorasib structures.

Other groups have taken an alternative method to covalent fragment-based screening, known as competitive isoTOP-ABPP (isotopic tandem orthogonal proteolysis-activity-based protein profiling), where fragments are screened directly in cells[15,63–68]. In short, cells are first treated with electrophilic fragments or DMSO as a

control, then treated with a broadly reactive electrophile, such as iodoacetamide desthiobiotin for cysteine, containing an affinity handle. Reactive, free cysteine-containing peptides are then enriched, identified, and quantified by mass spectrometry Through chemoproteomic profiling, the residues competed by the electrophilic fragments are identified. This cellular-based approach allows for screening against proteins that may be difficult to express and purify, including membrane proteins and large macromolecular complexes. Additionally, it assesses chemoselectivity in a cellular environment, which is not always observed with hits that derive from screening pure, recombinantly expressed proteins. This comes with the drawback that target engagement is measured indirectly and requires follow-up validation experiments to rule out false positives caused by complex cellular interactions. Additionally, the low throughput of this method is extremely limiting in the number of fragments that can be screened. Finally, the stochastic nature of mass spectrometry data acquisition means missing out on potential residue sites, a major issue for low-abundance proteins.

Covalent fragment-based screening has focused primarily on cysteine-targeted electrophiles on account of cysteine's intrinsic reactivity. In contrast, there have been relatively few lysine-targeted covalent-fragment screens. Given that lysine is more prevalent than cysteine in the proteome, developing methods for lysine-targeted covalent-fragment screening increases the generalizability of this approach and expands the breadth of potential protein targets. Broadly lysine-reactive electrophiles, like NHS-ester-alkyne and sulfotetrafluorophenyl-ester-alkyne, have been used for lysine-targeted competitive isoTOP-ABPP, but only in cell lysates[15,64,69]. Aldehyde fragments have also been used to identify lysine-targeted 14-3-3/client protein-protein

interaction stabilizers through reversible-covalent imine-tethering, with one fragment, **28**, showing a greater than 90-fold stabilization of the 14-3-3/Pin1 complex at 100 μM[70,71](**Fig. 2.2**). Given that aldehydes are, in principle, suitable for cellular and *in vivo* experiments (unlike disulfides), this simplifies the further development and use of the initial fragment hits. However, very few examples of imine-tethering with aldehydes have been reported.

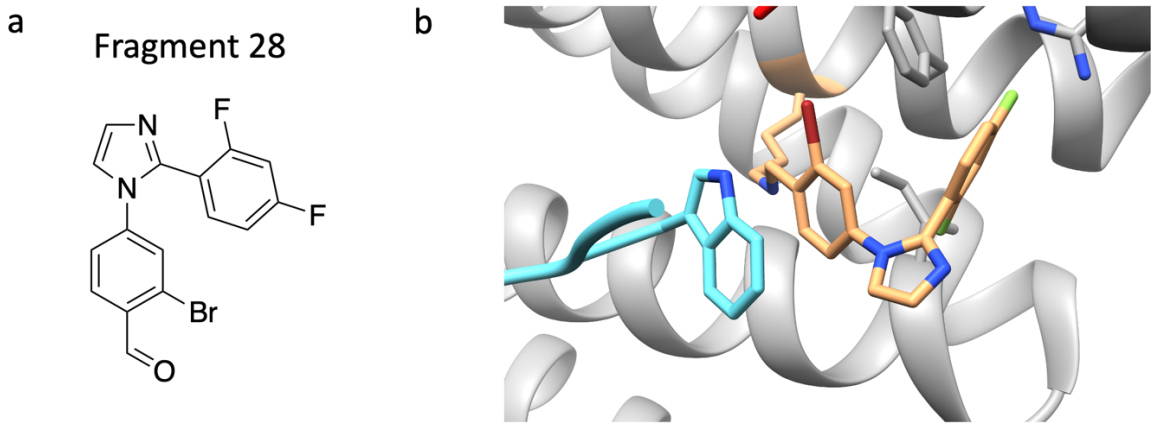


Figure 2.2. Aldehyde fragment stabilizes 14-3-3/Pin1 complex.

(a) Structure of aldehyde fragment 28. (b) Crystal structure of fragment 28 bound to 14-3-3/Pin1 complex.

We reasoned that salicylaldehydes would have even greater potential as imine-tethering fragments than simple aldehydes. Similar to aldehydes, salicylaldehydes form reversible imine adducts with biological amines like lysine, but the *ortho*-hydroxyl group has been previously shown to stabilize the imine through an intramolecular hydrogen bond[26,27]. For example, in the clinically approved salicylaldehyde inhibitor, voxelotor, removal of the hydroxyl results in a 10-fold decrease in the residence time ($T_{1/2}$ 19.1 h vs. 2.1 h)[23]. Given that we previously showed the potential of salicylaldehydes to covalently engage a surface lysine on Hsp90 (Chapter 1), we were inspired to push this

idea further and test smaller, fragment-like salicylaldehydes. Herein, we show that a small, adenosine-like fragment appended to a salicylaldehyde is sufficient to completely label surface lysine K58 on Hsp90 at low μM concentrations within 2 h. We further show that our fragment binds with an extended residence time, approximately 6-fold greater than the potent non-covalent inhibitor PU-H71, despite the greater molecular weight of PU-H71. Lastly, we observe that the small size of our fragment allows for greater protein flexibility within the binding site, allowing the salicylaldehyde to engage an alternative surface-exposed lysine on Hsp90.

Results

To screen lysine-targeted, reversible-covalent fragments against Hsp90, I synthesized a series of compounds with a salicylaldehyde linked to an adenosine fragment by a variable amine-containing linker (**Fig 2.3a**). I then assessed covalent modification of Hsp90 WT and K58H (1 μM) using intact-protein mass spectrometry (**Fig 2.3b, Fig 2.3c**). JW-2-85, with a salicylaldehyde linked to a simple methyl piperidine, acted as a control for promiscuous labeling. Satisfyingly, JW-2-85 at 50 μM did not label Hsp90 at all. The benzaldehyde JW-2-127 showed minimal labeling (< 10%) at 10 μM. Of the four salicylaldehydes tested, JW-2-128 was the only one to achieve 100% labeling within 2 h. JW-2-162, which has the same salicylaldehyde moiety as JW-2-128 but a more flexible linker, also labeled Hsp90 rapidly. However, the extent of modification appeared to plateau around 80%. The increased labeling efficiency of JW-2-128 relative to JW-2-127 demonstrates the utility of salicylaldehydes over benzaldehydes and how much imine stability is provided by the *ortho*-hydroxyl group. Of

note, JW-2-128 has the same chirality and salicylaldehyde regiochemistry as compound **7**, the best salicylaldehyde inhibitor identified in Chapter 1. This suggests that the adenosine fragments likely adopt a similar binding mode as the fully elaborated scaffold. JW-2-128 also showed a dose dependence from 0.3-10 μM (**Fig 2.4**), indicating the adenosine fragment does not immediately saturate the binding pocket. All four salicylaldehydes labeled the K58R mutant as well, suggesting that they can potentially modify an alternative lysine residue on Hsp90. In the presence of excess PU-H71, labeling of both WT and K58R was attenuated (**Fig 2.5a**), suggesting that covalent engagement of the alternative lysine still requires the fragment to non-covalently bind in the ATP pocket. The enantiomer of JW-2-128, JW-2-159, labeled WT and K58R at similar rates, suggesting that it may preferentially target this alternative lysine. JW-2-162 labeled the K58R mutant even faster than JW-2-128, but slower than it labeled the WT protein. This suggests that the flexible propyl linker can more readily adopt the geometry required to modify the alternative lysine, relative to the conformationally rigid piperidine linker. Nevertheless, Lys58 appears to be the preferred site of modification. JW-2-128 shows a similar kinetic preference for Lys58 over the alternative lysine residue.

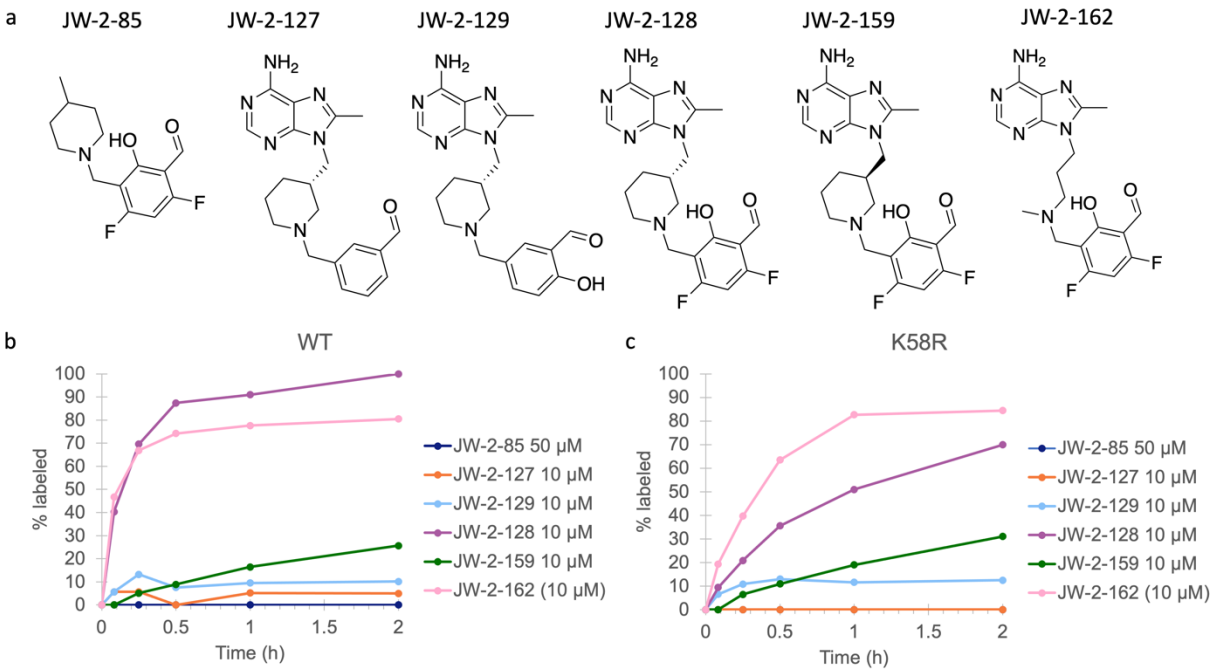


Figure 2.3. Fragment salicylaldehydes modify Hsp90 WT and K58R.

(a) Structures of salicylaldehyde fragments. (b) and (c) Hsp90 WT (b) and K58R (c) (1 µM) was treated with the indicated compounds (10 µM) at room temperature. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of covalently modified Hsp90 quantified by intact-protein LC-MS analysis.

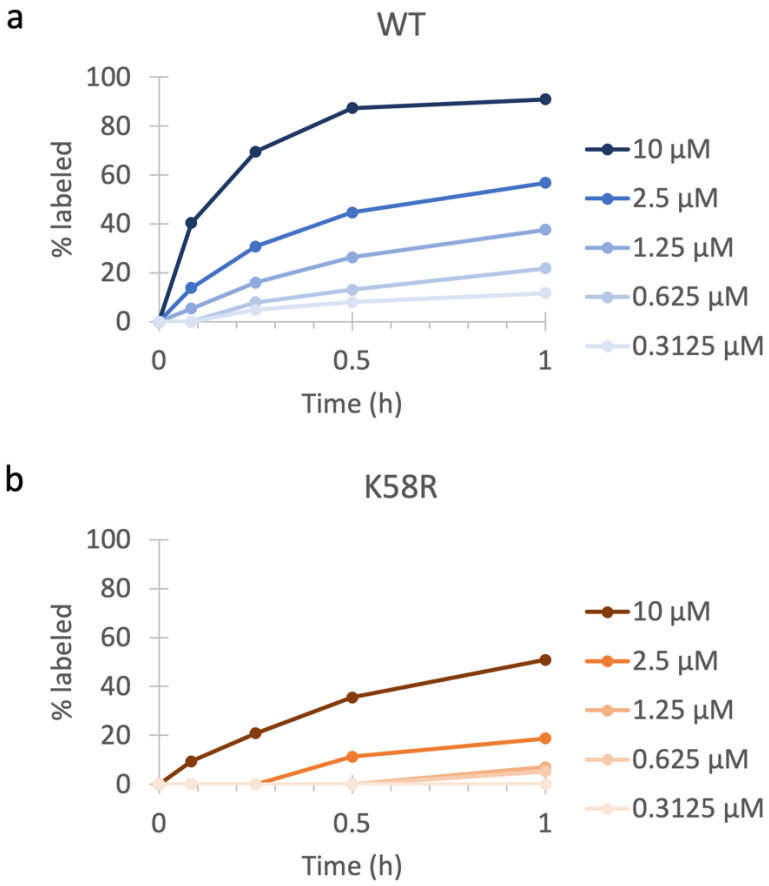


Figure 2.4. JW-2-128 labels Hsp90 in a dose dependent manner.

Hsp90 WT (a) and K58R (b) (1 µM) were treated with JW-2-128 at the indicated concentrations at room temperature. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of covalently modified Hsp90 quantified by intact-protein LC-MS analysis.

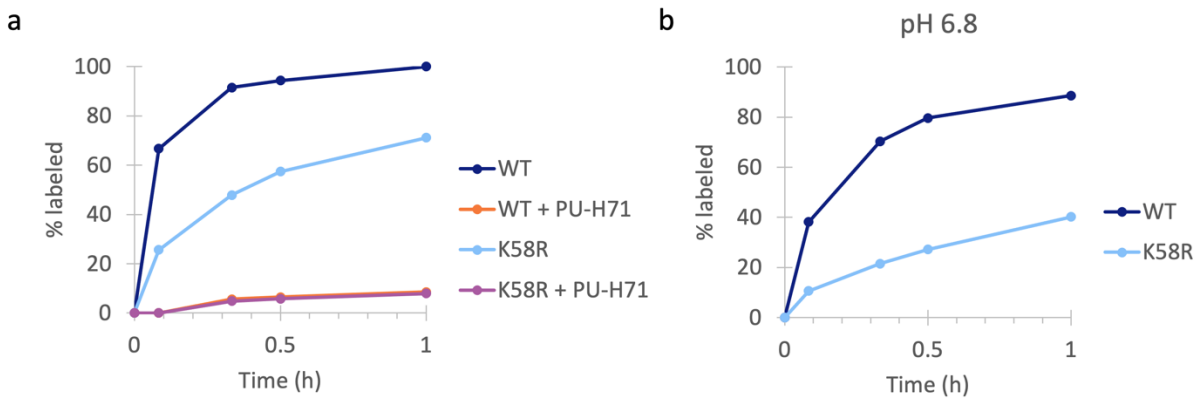


Figure 2.5. JW-2-128 competes with a non-covalent Hsp90 inhibitor.

(a) Hsp90 WT or K58R (1 μM) were first treated with PU-H71 (10 μM) for 15 min at room temperature, then JW-2-128 (10 μM) was added. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of covalently modified Hsp90 quantified by intact-protein LC-MS. (b) Hsp90 WT or K58R (1 μM) were treated with JW-2-128 (10 μM) at room temperature in pH 6.8 buffer. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of covalently modified Hsp90 quantified by intact-protein LC-MS analysis.

We estimated the off-rate of JW-2-128 by first incubating it with WT or K58R Hsp90 (5 μM protein, 10 μM compound), then diluting 10-fold into buffer containing a high concentration of PU-H71 (100 μM) to limit rebinding after compound dissociation (**Fig 2.6**). JW-2-128 showed an extended residence time ($T_{1/2}$ ~1.5 h) relative to the potent non-covalent inhibitor PU-H71 ($T_{1/2}$ ~15 min) (**Fig 1.8**). This result demonstrates the ability of a salicylaldehyde to form a stable imine adduct even when the non-covalent binding affinity is relatively low. JW-2-128 maintained similar rates of dissociation from both WT and K58R Hsp90, indicating that the imine adducts formed with Lys58 and with the alternative lysine have comparable kinetic stabilities. It seems plausible that, despite targeting two distinct lysines, JW-2-128 is able to adopt an unstrained imine with a hydrogen bond to the *ortho*-phenol in both cases.

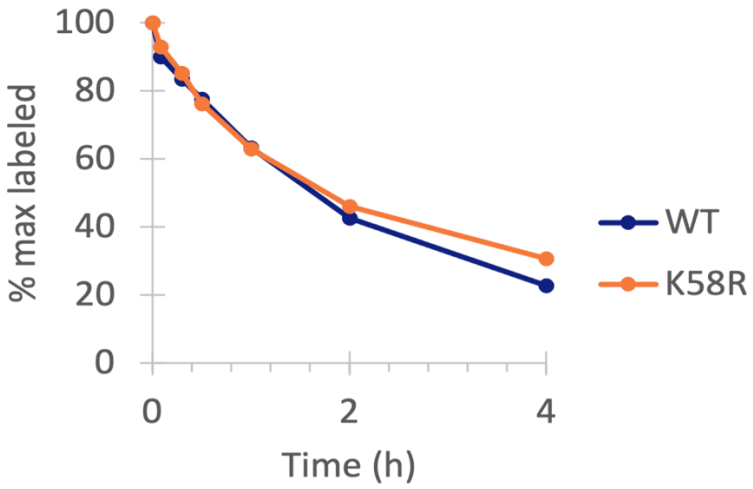


Figure 2.6. JW-2-128 shows extended residence times.

Hsp90 WT or K58R (5 μM) were treated with JW-2-128 (10 μM) for 2 h at room temperature, then diluted 10-fold into buffer containing 100 μM PU-H71. Aliquots were removed at the indicated time points, treated with $NaBH_4$ for 10 min, and the percentage of covalently modified Hsp90 quantified by intact-protein LC-MS.

It is known that pH can have a strong effect on the reaction rates of simple amines with salicylaldehydes, as well as the imine-aldehyde equilibrium[72,73]. We tested the effect of a slightly acidic pH (pH 6.8) on our fragment salicylaldehyde JW-2-128. As expected, the lower pH disfavored imine formation and reduced the rate of labeling Hsp90 by the fragments (**Fig 2.5b**), relative to pH 7.4 (**Fig 2.5a**). This effect was more pronounced in the case of the K58R mutant (labeling reduced by ~50%) relative to the WT (labeling reduced by ~10%). This result suggests that JW-2-128 may form a less stable imine with the K58R mutant. Nevertheless, the similar dissociation kinetics of JW-2-128 from both WT and K58R (**Fig 2.6**) indicate that the imine bonds have comparable kinetic stability at pH 7.4. The differential pH stability may result from the different pK_a's of Lys58 and the alternative lysine, as well as the initial hemiaminal intermediates. A change in the protonation states of the lysine and the intermediate states could affect

the rate of formation of the imine adduct and could explain the observed pH dependencies. The differential effect of pH on different salicylaldehyde-lysine adducts suggests that buffer pH could be a useful method of tuning an imine-tethering screen, with lower pH acting as a more stringent selection criterion and potentially filtering out weaker binders. However, due to the different pK_a's of different lysines, buffer pH would likely need to be optimized on a protein-by-protein basis.

The crystal structure of salicylaldehyde **6** covalently bound to Hsp90 (**Fig 1.4d**) does not reveal an obvious alternative lysine that JW-2-128 could engage on the K58R mutant. This raises the question of what residue on the K58R mutant is engaged by JW-2-128. To determine whether JW-2-128 and compound **6** have the same binding mode, we resolved the co-crystal structure of JW-2-128 bound to WT Hsp90 NTD at 1.8 Å resolution (**Fig 2.7a**). Strong, continuous density confirmed covalent modification of Lys58 by JW-2-128. Comparison to the crystal structure of salicylaldehyde **6** revealed that the methyladenine fragment of JW-2-128 adopted a nearly identical binding pose as the thioadenine moiety of **6**, suggesting this orientation is highly favorable due to the formation of multiple interactions with Hsp90 (**Fig 2.7b**).

Interestingly, the gap in the binding pocket left by the absence of the iodo benzodioxazole allowed for partial unwinding of the helix between Ile104 and Ile110 (**Fig 2.7c**). This helix is known to be flexible, as it is a part of the ATP lid which folds down over the ATP-binding pocket during the Hsp90 catalytic cycle[74]. Moreover, this partly unwound state has been previously observed in liganded crystal structures[75]. This conformational change in the protein causes a series of cascading effects that result in a different conformation of the JW-2-128 salicylaldimine relative to the compound **6**

salicylaldimine (Chapter 1). In the structure, Leu107 moves to occupy the hole left by the benzodioxazole. In doing so, the Gly108 carbonyl moves proximal to the piperidine linker (1.9 Å from the benzyl CH_2 in compound **6**), forcing it to move closer to Lys58 and twisting the position of the adenosine fragment slightly. In order to accommodate this shift, Lys58 adopts a less extended conformation and the torsion angle of the benzene-CH_2 bond rotates 33º. This rotation moves the piperidine nitrogen further from the phenol (2.6 Å to 3.2 Å) and eliminates the possibility of forming a similar extended intramolecular hydrogen bond network as to compound **6**. While conformational changes observed in the JW-2-128 do not immediately suggest which alternative lysine could be modified in the K58R mutant, overlaying this structure with another Hsp90 NTD structure where the helix is partially unwound reveals Lys112 as a strong candidate (**Fig 2.7d**). An alternative conformation of the unwound helix, which would likely be accommodated by JW-2-128, brings Lys112 to within 3.7 Å of the benzyl CH_2. Although this hypothesis requires direct confirmation, the lack of any other lysines proximal to the ATP-binding pocket strongly suggests that Lys112 is the alternative residue labeled in the K58R mutant.

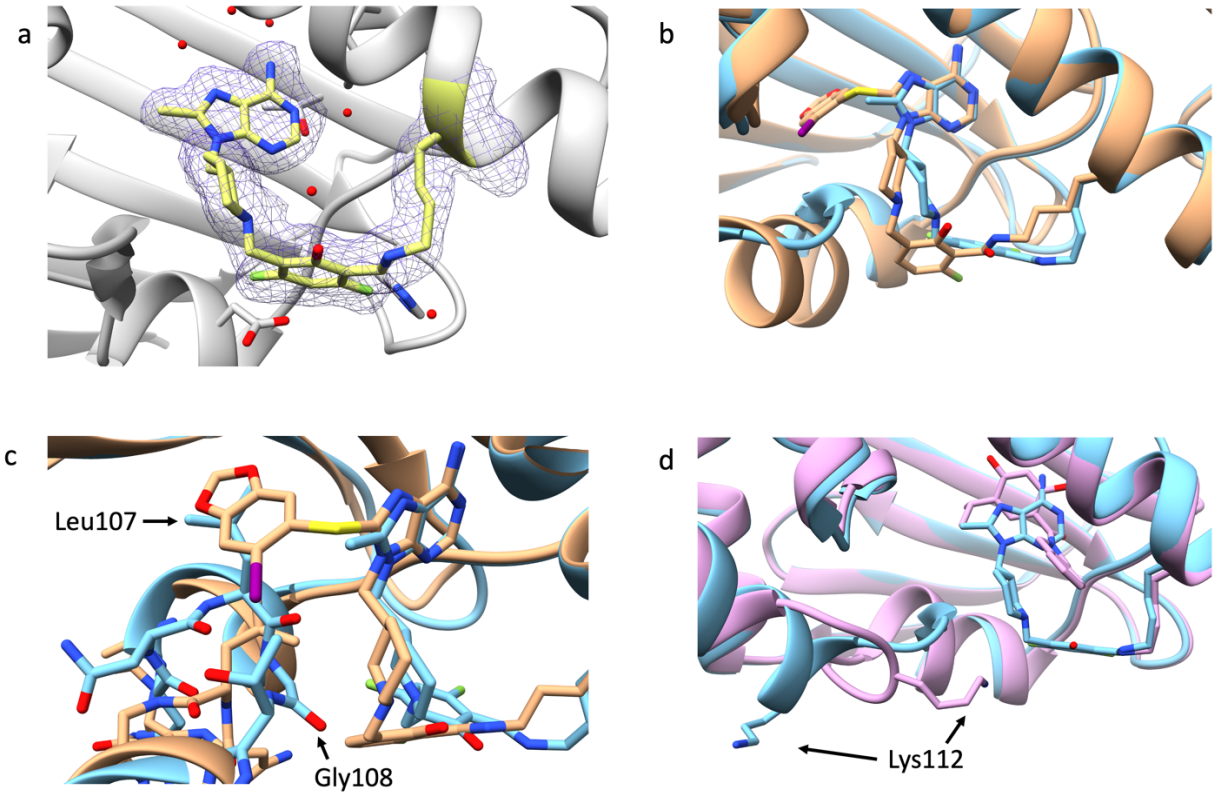


Figure 2.7. Crystal structure of JW-2-128 bound to WT Hsp90 NTD.

(a) Crystal structure of JW-2-128 covalently bound to Hsp90 NTD at 1.8 Å resolution. Electron density map (2Fo-Fc, 1.0 σ) is shown for Lys58 and JW-2-128. (b) Overlay of JW-2-128-Hsp90 structure with compound **6**-Hsp90 structure. (c) Zoomed in overlay of JW-2-128-Hsp90 structure with compound **6**-Hsp90 structure with key residues labeled. (d) Overlay of JW-2-128-Hsp90 structure with PDB 2XAB with Lys112 labeled.

By testing a series of adenosine fragments linked to salicylaldehydes, we demonstrated that fragment salicylaldehydes have the potential to strongly and selectively engage a lysine residue on Hsp90. The high degree of target engagement, even at low μM concentrations, suggests that a salicylaldehyde library could have potential for use in an imine-tethering screen.

Methods

Protein expression and purification. Hsp90 N-terminal domain (NTD) was expressed and purified as previously described[11].

Time course of Hsp90 NTD modification. Compounds (10 μM) were added to 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$) at RT. At each time point an aliquot was removed and treated with 5 mM $NaBH_4$ for 10 min at RT and then quenched with 1% formic acid. Samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of modified Hsp90.

Dose dependent Hsp90 NTD modification. JW-2-128 at the indicated concentrations (0.3 – 10 μM) was added to 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$) at RT. At each time point an aliquot was removed and treated with 5 mM $NaBH_4$ for 10 min at RT and then quenched with 1% formic acid. Samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of modified Hsp90.

Low pH Hsp90 NTD modification. JW-2-128 (10 μM) was added to 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 6.8, 50 mM KCl, 25 mM $MgCl_2$) at RT. At each time point an aliquot was removed and treated with 5 mM $NaBH_4$ for 10 min at RT and then quenched with 1% formic acid. Samples were analyzed on a Waters Xevo G2-XS QToF

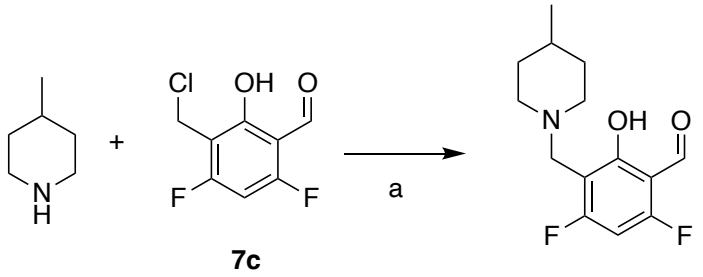
LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of modified Hsp90.

Time course of JW-2-128 dissociation from Hsp90 NTD. JW-2-128 (10 μM) was added to 5 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$) at RT for 1 h. An aliquot was removed for determination of the initial percentage modification. The mixture was then diluted 10-fold into labeling buffer containing 100 μM PU-H71 as a competitor. At each time point an aliquot was removed and treated with 5 mM $NaBH_4$ for 10 min at RT and then quenched with 1% formic acid. Samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of modified protein.

Co-crystallization and structure determination of JW-2-128 bound to Hsp90 NTD. Hsp90 NTD was covalently modified by incubating 10 μM JW-2-128 with 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$). Complete modification was confirmed by LC-MS. Modified protein was buffer exchanged into 50 mM HEPES pH 8.0, 50 mM KCl, 1 mM DTT with 10% glycerol and concentrated to 20 mg/mL using a 10K MWCO filter (Millipore Amicon Ultra). Crystallization was performed using the hanging drop method with a 1:1 mixture of 20 mg/mL protein-compound adduct and crystallization solution containing 25% PEG-3350, 0.2 M $MgCl_2$, 0.1 M Bis-Tris pH 6.5. Protein crystals formed within 48 h. X-ray diffraction data was collected at beamline 8.3.1 of the Advanced Light Source (Lawrence Berkley National Laboratory).

The resulting X-ray diffraction data was processed using the XDS package[34], followed by molecular replacement with PHASER[35] using PDB code: 6U9B as a starting model. Structure refinement was performed using coot[36] and phenix.refine[35].

Synthetic Methods

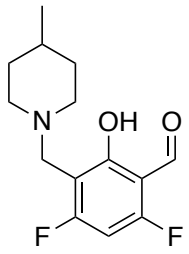


7c **JW-2-85**

Figure 2.8. Synthetic scheme 2.1.

Scheme 2.1: a) DIPEA, DCM, RT, 1h

4,6-difluoro-2-hydroxy-3-((4-methylpiperidin-1-yl)methyl)benzaldehyde (JW-2-85):



4-methylpiperidine (120 mg, 1.2 mmol) and DIPEA (365 µL, 1.6 mmol) were dissolved in DCM (5 mL), followed by the addition of a solution of **7c** (75 mg, 0.4 mmol) in DMF (5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 52 mg, 48%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.27 (s, 1H), 6.29 (dd, *J* = 10.9, 9.9 Hz, 1H), 3.73 (s, 2H), 3.03 – 2.95 (m, 2H), 2.20 (td, *J* = 11.8, 2.6 Hz, 2H), 1.74 – 1.65 (m, 2H), 1.46 – 1.41 (m, 1H), 1.39 – 1.24 (m, 2H), 0.94 (d, *J* = 6.4 Hz, 3H). HRMS (ESI-TOF) *m/z* calcd. $C_{14}H_{18}F_2NO_2^+$ $[M + H]^+$, 270.1300; found, 270.1276.

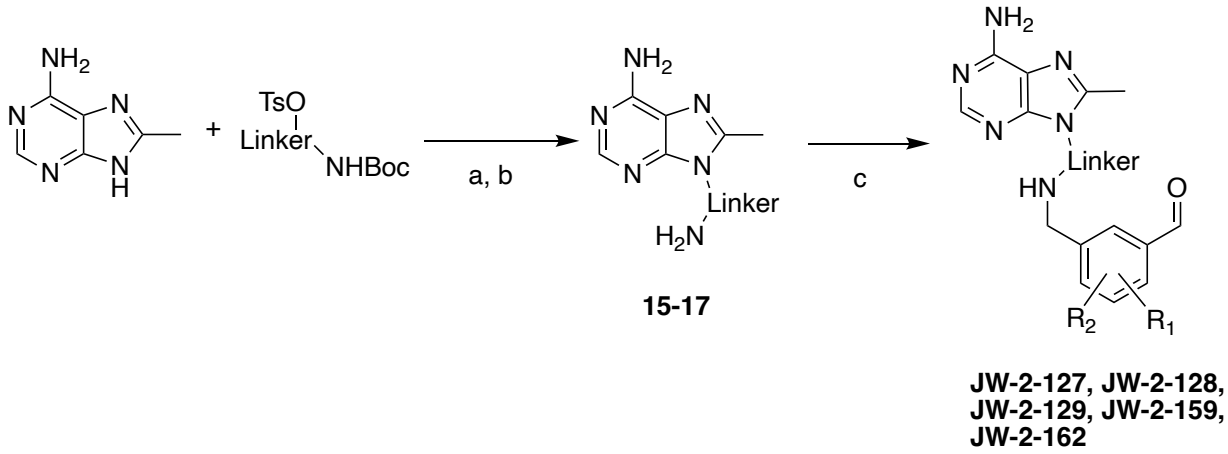
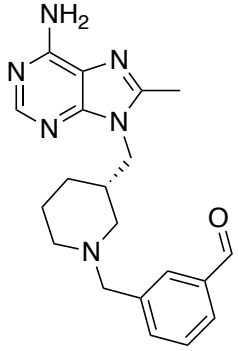


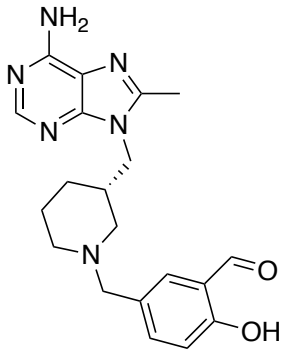
Figure 2.9. Synthetic scheme 2.2.

Scheme 2.2: a) Cs_2CO_3, DMF, 80 °C, 16 h. b) TFA:CH_2Cl_2 1:1, 0 °C, 1 h. c) DIPEA, DMF, RT, 1 h. R_1 = H or OH. R_2 = H or F.

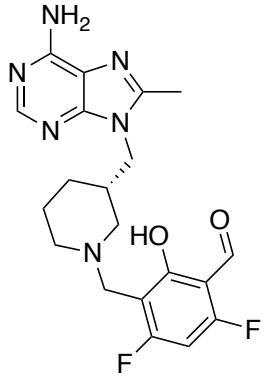
(*S*)-3-((3-((6-amino-8-methyl-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)benzaldehyde (JW-2-127): 15 (30 mg, 0.12 mmol) and DIPEA (52 µL, 0.24 mmol) were dissolved in DMF (1 mL), followed by the addition of a solution of 3-bromobenzaldehyde (30 mg, 0.15 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 11 mg, 25%. ^{1}H NMR (400 MHz, MeOD) δ 9.95 (s, 1H), 8.13 (s, 1H), 7.83 – 7.78 (m, 2H), 7.58 (d, *J* = 7.4 Hz, 1H), 7.49 (t, *J* = 7.4 Hz, 1H), 4.07 (d, *J* = 8.0 Hz, 2H), 3.70 (d, *J* = 2.5 Hz, 1H), 3.61 (d, *J* = 13.0 Hz, 1H), 2.86 (d, *J* = 11.7 Hz, 1H), 2.64 (d, *J* = 10.8 Hz, 1H), 2.28 – 2.17 (m, 2H), 2.05 (t, *J* = 10.7 Hz, 1H), 1.82 – 1.73 (m, 1H), 1.73 – 1.65 (m, 1H), 1.65 – 1.52 (m, 1H), 1.26 – 1.12 (m, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{20}H_{25}N_6O^+$ $[M + H]^+$, 365.2084; found, 365.2014.



(*S*)-5-((3-((6-amino-8-methyl-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-2-hydroxybenzaldehyde (JW-2-129): 15 (30 mg, 0.12 mmol) and DIPEA (52 µL, 0.24 mmol) were dissolved in DMF (1 mL), followed by the addition of a solution of 5-bromo-2-hydroxybenzaldehyde (32 mg, 0.15 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 7.1 mg, 15%. ^{1}H NMR (400 MHz, MeOD) δ 10.02 (s, 1H), 8.15 (s, 1H), 7.69 (d, *J* = 8.1 Hz, 2H), 7.64 (d, *J* = 2.3 Hz, 1H), 7.48 (dd, *J* = 8.5, 2.3 Hz, 1H), 7.22 (d, *J* = 8.3, 2.0 Hz, 2H), 6.91 (d, *J* = 8.5 Hz, 1H), 4.14 – 4.08 (m, 2H), 3.91 (d, *J* = 13.0 Hz, 1H), 3.81 (d, *J* = 13.0 Hz, 1H), 3.14 (d, *J* = 11.0 Hz, 1H), 2.88 (d, *J* = 11.0 Hz, 1H), 1.92 – 1.82 (m, 1H), 1.81 – 1.73 (m, 2H), 1.73 – 1.64 (m, 1H), 1.35 – 1.21 (m, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{20}H_{25}N_6O_2^+$ $[M + H]^+$, 381.2034; found, 381.2059.

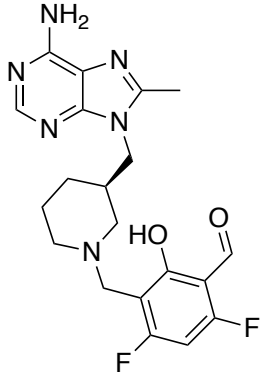


(*S*)-3-((3-((6-amino-8-methyl-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-4,6-difluoro-2-hydroxybenzaldehyde (JW-2-128): 15 (30 mg, 0.12 mmol) and DIPEA (52 µL, 0.24 mmol) were dissolved in DMF (1 mL), followed by the addition of a solution of **7c** (30 mg, 0.15 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane



(solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 7.3 mg, 14%. ^{1}H NMR (400 MHz, DMSO) δ 10.16 (s, 1H), 8.33 (s, 1H), 7.13 (t, *J* = 10.5 Hz, 1H), 4.21 (s, 2H), 4.11 (dd, *J* = 7.2, 3.0 Hz, 2H), 3.37 (d, *J* = 11.6 Hz, 1H), 3.29 (d, *J* = 11.2 Hz, 1H), 2.96 – 2.85 (m, 2H), 2.55 (s, 3H), 1.83 (d, *J* = 14.0 Hz, 1H), 1.63 (d, *J* = 12.4 Hz, 3H). HRMS (ESI-TOF) *m/z* calcd. $C_{20}H_{23}F_2N_6O_2^+$ [M + H]$^+$, 417.1845; found, 417.1890.

(*R*)-3-((3-((6-amino-8-methyl-9*H*-purin-9-yl)methyl)piperidin-1-yl)methyl)-4,6-difluoro-2-hydroxybenzaldehyde (JW-2-159): 16 (30 mg, 0.12 mmol) and DIPEA (52 µL, 0.24 mmol) were dissolved in DMF (1 mL), followed by the addition of a solution of **7c** (30 mg, 0.15 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 9.7 mg, 19%. ^{1}H NMR (400 MHz, DMSO) δ 10.17 (s, 1H), 8.39 (s, 1H), 7.15 (t, *J* = 10.6 Hz, 1H), 4.22 (s, 2H), 4.14 (dd, *J* = 7.2, 3.3 Hz, 2H), 3.41 – 3.37 (m, 1H), 3.34 – 3.27 (m, 1H), 2.98 – 2.86 (m, 2H), 2.57 (s, 3H), 1.88 – 1.80 (m, 1H), 1.64 (d, *J* = 12.4 Hz, 2H). X HRMS (ESI-TOF) *m/z* calcd. $C_{20}H_{23}F_2N_6O_2^+$ [M + H]$^+$, 417.1845; found, 417.1848.

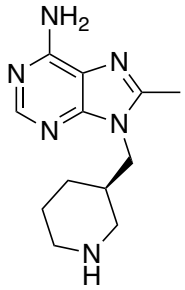


3-(((3-(6-amino-8-methyl-9*H*-purin-9-yl)propyl)(methyl)amino)methyl)-4,6-difluoro-2-hydroxybenzaldehyde (JW-2-162): 17 (35 mg, 0.16 mmol) and DIPEA (70 µL, 0.32 mmol) were dissolved in DMF (1 mL), followed by the addition of a solution of **7c** (40 mg, 0.16 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 9.3 mg, 15%. ^{1}H NMR (400 MHz, DMSO) δ 10.15 (s, 1H), 8.47 (s, 1H), 7.14 (t, *J* = 10.6 Hz, 1H), 4.27 (d, *J* = 11.4 Hz, 4H), 2.61 (s, 3H), 2.55 (t, *J* = 5.6 Hz, 2H), 2.27 – 2.14 (m, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{18}H_{21}F_2N_6O_2^+$ [M + H]$^+$, 391.1689; found, 391.1703.

(*S*)-8-methyl-9-(piperidin-3-ylmethyl)-9*H*-purin-6-amine (15): Tert-butyl(3S)-3-(hydroxymethyl)piperidine-1-carboxylate (2.3 g, 10.7 mmol) and *p*-toluenesulfonyl chloride (2.5 g, 12.8 mmol) were dissolved in CH_2Cl_2 (50 mL) and then pyridine (730 µL, 12.8 mmol) was added. The reaction was left to stir for 16 h at RT. The reaction was then diluted with saturated ammonium chloride and extracted 3 times with dichloromethane. The combined organic extracts were washed with brine, dried over sodium sulfate and concentrated under vacuum. The resulting tosylate was used without further purification. 8-methyl-7H-purin-6-amine (125 mg, 0.84 mmol) and Cs_2CO_3 (273 mg, 0.84 mmol) were suspended in DMF (15 mL) and heated to 80 °C under Ar. The tosylate (775 mg, 2.1 mmol) was dissolved in DMF (5 mL) and added to the reaction dropwise. The reaction was left to

stir at 80 °C for 16 h. The solvent was removed under vacuum and the residue was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). The resulting mixture of purine regioisomers was concentrated, re-dissolved in dichloromethane (4 mL) and set to stir on ice. Trifluoroacetic acid (4 mL) was then added dropwise. After 1 h, the reaction mixture was concentrated and purified by silica gel chromatography using a 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 75 mg, 28%. ^{1}H NMR (400 MHz, MeOD) δ 8.16 (s, 1H), 4.17 (d, *J* = 7.4 Hz, 2H), 3.22 – 3.17 (m, 1H), 2.98 – 2.86 (m, 2H), 2.60 (s, 3H), 2.02 – 1.91 (m, 1H), 1.87 – 1.78 (m, 1H), 1.78 – 1.63 (m, 1H), 1.45 (qd, *J* = 12.4, 3.9 Hz, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{12}H_{19}N_6^+$ [M + H]$^+$, 247.1666; found, 247.1689.

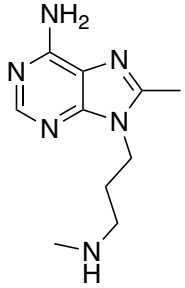
(*R*)-8-methyl-9-(piperidin-3-ylmethyl)-9*H*-purin-6-amine (16): Tert-butyl(3R)-3-



(hydroxymethyl)piperidine-1-carboxylate (2.3 g, 10.7 mmol) and *p*-toluenesulfonyl chloride (2.5 g, 12.8 mmol) were dissolved in CH_2Cl_2 (50 mL) and then pyridine (730 µL, 12.8 mmol) was added. The reaction was left to stir for 16 h at RT. The reaction was then diluted with saturated ammonium chloride and extracted 3 times with dichloromethane. The combined organic extracts were washed with brine, dried over sodium sulfate and concentrated under vacuum. The resulting tosylate was used without further purification. 8-methyl-7H-purin-6-amine (125 mg, 0.84 mmol) and Cs_2CO_3 (273 mg, 0.84 mmol) were suspended in DMF (15 mL) and heated to 80 °C under Ar. The tosylate (775 mg, 2.1 mmol) was

dissolved in DMF (5 mL) and added to the reaction dropwise. The reaction was left to stir at 80 °C for 16 h. The solvent was removed under vacuum and the residue was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). The resulting mixture of purine regioisomers was concentrated, re-dissolved in dichloromethane (4 mL) and set to stir on ice. Trifluoroacetic acid (4 mL) was then added dropwise. After 1 h, the reaction mixture was concentrated and purified by silica gel chromatography using a 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 171 mg, 83%. ^{1}H NMR (400 MHz, MeOD) δ 8.16 (s, 1H), 4.17 (d, *J* = 7.4 Hz, 2H), 3.26 – 3.20 (m, 1H), 3.00 – 2.83 (m, 2H), 2.61 (s, 3H), 2.45 – 2.37 (m, 1H), 2.02 – 1.92 (m, 1H), 1.87 – 1.79 (m, 1H), 1.77 – 1.64 (m, 1H), 1.52 – 1.37 (m, 1H). HRMS (ESI-TOF) *m/z* calcd. $C_{12}H_{19}N_6^+$ $[M + H]^+$, 247.1666; found, 247.1623.

8-methyl-9-(3-(methylamino)propyl)-9*H*-purin-6-amine (17): tert-butyl N-(3-



hydroxypropyl)-N-methylcarbamate (1.6 g, 8.5 mmol) and *p*-toluenesulfonyl chloride (2.0 g, 10.3 mmol) were dissolved in CH_2Cl_2 (40 mL) and then pyridine (590 µL, 10.3 mmol) was added. The reaction was left to stir for 16 h at RT. The reaction was then diluted with saturated ammonium chloride and extracted 3 times with dichloromethane. The combined organic extracts were washed with brine, dried over sodium sulfate and concentrated under vacuum. The resulting tosylate was used without further purification. 8-methyl-7H-purin-6-amine (65 mg, 0.44 mmol) and Cs_2CO_3 (142 mg, 0.44 mmol) were suspended in DMF

(6 mL) and heated to 80 °C under Ar. The tosylate (375 mg, 1.09 mmol) was dissolved in DMF (4 mL) and added to the reaction dropwise. The reaction was left to stir at 80 °C for 16 h. The solvent was removed under vacuum and the residue was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). The resulting mixture of purine regioisomers was concentrated, re-dissolved in dichloromethane (4 mL) and set to stir on ice. Trifluoroacetic acid (4 mL) was then added dropwise. After 1 h, the reaction mixture was concentrated and purified by silica gel chromatography using a 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 77 mg, 81%. ^{1}H NMR (400 MHz, MeOD) δ 8.36 (s, 1H), 4.39 (d, *J* = 7.0 Hz, 2H), 3.07 (t, *J* = 7.7 Hz, 2H), 2.71 (s, 3H), 2.23 (p, *J* = 7.1 Hz, 2H). HRMS (ESI-TOF) *m/z* calcd. $C_{10}H_{17}N_6^+$ $[M + H]^+$, 221.1509; found, 221.1532.

Chapter 3 : Conformationally constrained linkers drive covalent modification of histidine using sulfonyl fluoride exchange.

Abstract

Histidine is an appealing target residue for covalent inhibitors on account of its frequent presence near small-molecule binding sites. Histidine's imidazole side chain has a pK_a of 6-7 and is substantially deprotonated at physiological pH, promoting nucleophilic addition to electrophilic groups. It has previously been shown that sulfonyl fluorides can be used as histidine-targeted electrophiles. Herein, we explore the ability of conformationally constrained linkers to tune the reaction kinetics of a sulfonyl fluoride with a target histidine, irrespective of the innate electrophile reactivity. Using the Hsp90 mutant K58H, we found that two compounds, AC45 and AC50, which contain an azetidine and piperidine linker respectively, showed strong and opposite K58H and WT Hsp90 selectivities. The structural drivers behind these preferences were revealed in a high-resolution crystal structure of AC45 covalently bound to Hsp90 K58H. Additionally, we show that a more stable electrophile, either sulfonyl imidazole or *ortho*-methoxy sulfonyl fluoride, maintains good rates of target modification and improved K58H vs. WT selectivity, despite having attenuated reactivity.

Introduction

With the expanding interest in covalent drug discovery, the rational design of compounds targeting non-cysteine nucleophiles is a key unmet need in the field. Although numerous lysine-targeted[8–15] and tyrosine-targeted[76–83] inhibitors have been reported in recent years, almost none have been tested in animal models and none have entered human clinical trials. The core challenge in targeting these and other non-cysteine amino acids is their attenuated reactivity. This often necessitates the use of

more intrinsically reactive electrophiles to develop compounds with a sufficiently fast k_{inact}. More reactive warheads increase the risk of undesired off-target modifications, a risk compounded by the increased prevalence of these alternative amino acids in the proteome, relative to cysteine. The design of fast and selective covalent inhibitors targeting non-cysteine amino acids remains a major challenge.

Histidine is an attractive and underexplored target for covalent inhibitor design. Histidines are highly prevalent near small-molecule binding sites[84,85] and, as opposed to lysines, are typically not protonated at physiological pH[86], thus facilitating nucleophilic addition. Electrophiles such as epoxides[87,88], α,β-unsaturated ketones[89,90], and others[91–94] have been shown to react with histidine residues (**Fig. 3.1**), but examples are limited, with typically one example of an inhibitor being reported for each electrophile. In the past two years, three examples of sulfonyl fluorides reacting with a histidine residue on a target protein have been reported[18,95,96]. The downside of sulfonyl fluorides is their low metabolic stability, with compounds often showing $T_{1/2}$ values in plasma on the order of minutes[16,95,97,98]. Aromatic substituents have been shown to affect the stability of aryl sulfonyl fluorides, for example, an *ortho*-methoxy group, as utilized in the unnatural amino acid SFY, showed an almost 10-fold decrease in the rate of hydrolysis relative to an unsubstituted aryl sulfonyl fluoride[16,99,100]. However, these substituted aryl sulfonyl fluorides remain generally unexplored in targeted covalent inhibitors. Fluorosulfates pose a more stable alternative to sulfonyl fluorides, but their attenuated reactivity means covalent modification is often slow. While the fastest fluorosulfate probes have been reported to have $T_{1/2}$ values in the range of 30 min[17,78], the majority of fluorosulfates require on the order of hours or days to react with their protein targets[18,19,76,77,95,101–103].

EM12-FS, a fluorosulfate probe targeting His353 on cereblon, showed a 28-fold increase in stability in human plasma relative to the sulfonyl fluoride EM12-SO_2F (**Fig. 3.2**), but had less than 60% target modification at 4 h[95]. Striking a balance between compound stability and modification rate is thus a major challenge in the field of histidine-targeted covalent inhibitors.

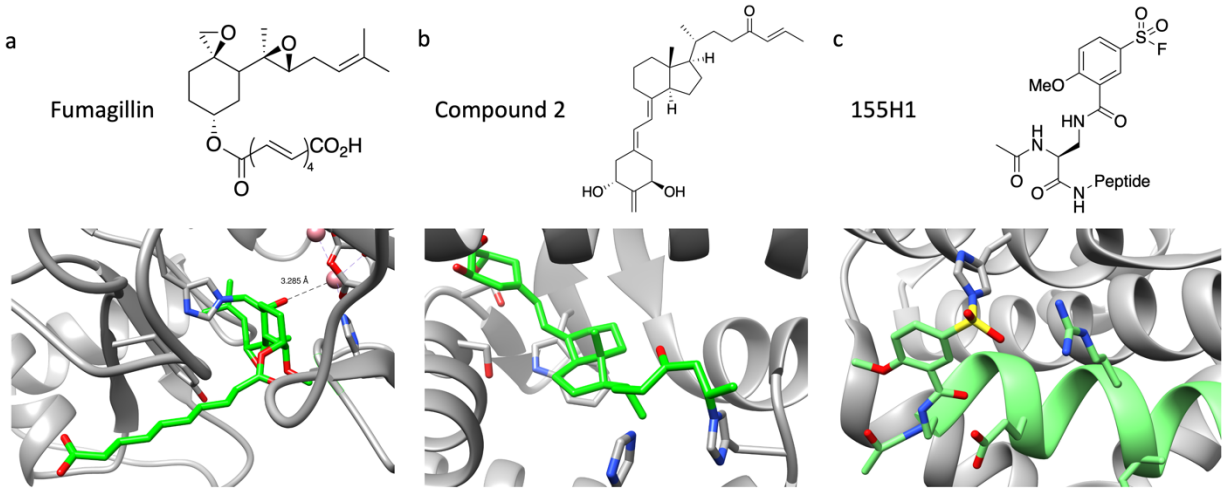
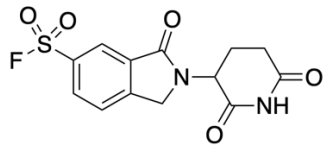


Figure 3.1. Examples of covalent small molecules targeting histidine residues.

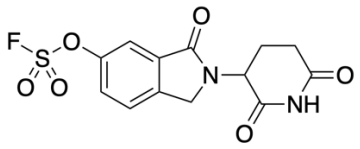
(a) Crystal structure of fumagillin bound to methionine aminopeptidase 2 (PDB: 1BOA). (b) Crystal structure of compound **2** bound to vitamin D receptor (PDB: 5ZWF). (c) Crystal structure of 155H1 bound to hMCL-1 (PDB: 8VJP).

EM12-SO_2F



Human plasma $T_{1/2}$ = 7 min

EM12-FS



Human plasma $T_{1/2}$ = 196 min

Figure 3.2. Fluorosulfates show increased plamsa stability over sulfonyl fluorides.

Structures of EM12-SO_2F and EM12-FS and reported stability in human plasma.

Previously, we have shown that conformationally constrained linkers have a dramatic effect on the modification rate of a lysine-targeted covalent aryl sulfonyl fluoride inhibitor, independent of the electrophilicity of the warhead[11]. Using a sulfonyl fluoride inhibitor that targets Hsp90 Lys58, a piperidine linker with inverted stereochemistry increased the k_{inact} over 30 fold. We reasoned that a similar principle would govern the relative reactivity of sulfonyl fluoride probes for histidine, particularly since it is more conformationally constrained than lysine. Where the side chain of lysine has 4 rotatable bonds, histidine only has two, giving it fewer degrees of freedom to accommodate the geometric constraints of the transition state during covalent bond formation. Building on our previous work, we hypothesized that linker optimization would provide a powerful means of improving the modification rate of histidine-targeting covalent inhibitors. By mutating Lys58 to His, we utilize Hsp90 as a model system for testing the effects of conformational constraints on sulfonyl fluoride reactivity towards a

histidine residue. Herein, we report a series of Hsp90 K58H-reactive sulfonyl fluoride probes and demonstrate that linker rigidity has the ability to confer selectivity between WT and K58H mutant Hsp90. Additionally, we show that an *ortho*-methoxy sulfonyl fluoride and a sulfonyl imidazole exhibit increased stability over the equivalent sulfonyl fluoride while maintaining fast rates of K58H Hsp90 modification.

Results

To explore the effects of conformationally constrained linkers on the ability of sulfonyl fluorides to covalently modfity a target histidine, we tested a series of previously reported *meta*-substituted aryl sulfonyl fluoride Hsp90 probes, AC23, AC44, and AC50 (**Fig 3.3a**), against both WT and K58H Hsp90 (1 μM protein, 10 μM compound). We assessed the rate of covalent modification by intact-protein mass spectrometry. By using compound concentrations >>10-fold higher than the K_i, the observed rate provides an estimate of the first-order rate of covalent bond formation (k_{inact}) within the noncovalent inhibitor/Hsp90 complex. With a flexible propyl linker, AC23 displayed rapid modification of both WT and K58H (**Fig 3.3b, Fig 3.3c**), albeit with a preference for K58H ($T_{1/2}$ WT ~15 min, $T_{1/2}$ K58H ~5 min). The ability of the flexible linker to accommodate the different geometric constraints of histidine and lysine likely accounts for it to efficiently react with both side chains; the more conformationally rigid histidine side chain may facilitate the more rapid reaction with AC23. The more rigid linkers in AC44 and AC50 showed clear and opposite preferences for either K58H or WT, respectively. While AC44 did not label K58H faster than AC23, its preference for K58H over WT protein motivated us to test the *para*-sulfonyl fluoride regioisomer AC45. AC45

showed even greater selectivity for K58H over WT, labeling K58H as fast as AC23 and WT even slower than AC44 ($T_{1/2}$ WT ~2 h). Based on the data shown in **Fig. 3**, AC45 emerged as the sulfonyl fluoride with the greatest selectivity for K58H over WT Hsp90.

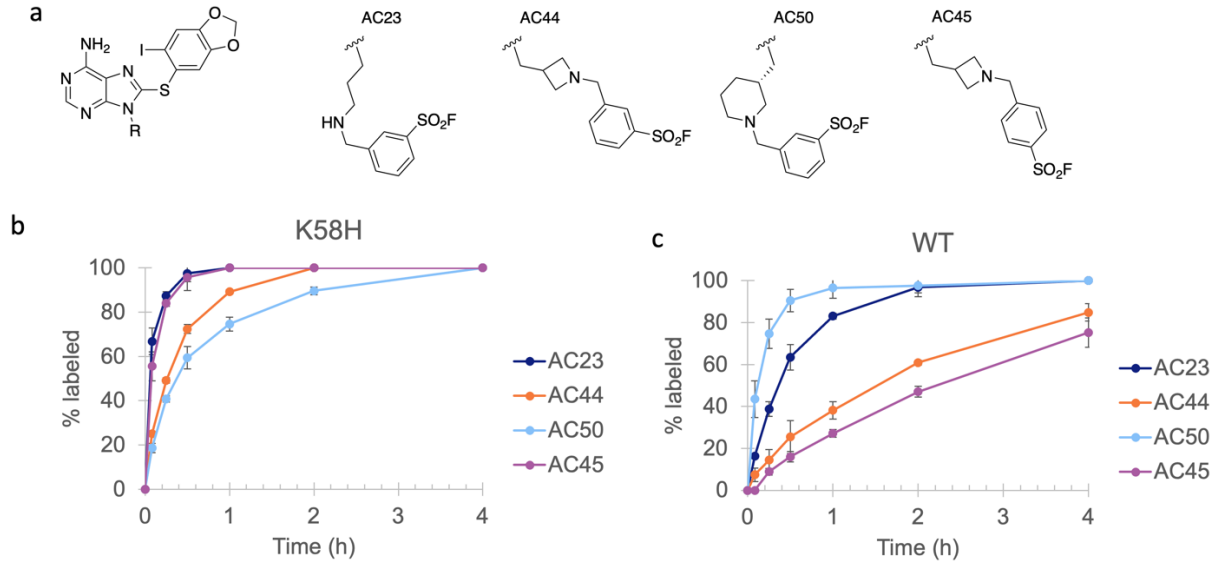


Figure 3.3: Sulfonyl fluorides modify Hsp90 K58H.

(a) Structure of sulfonyl fluoride probes. (b) Hsp90 K58H (1 μM) was treated with the indicated compounds (10 μM) at 37 °C. At the indicated time points, aliquots were removed and diluted 2-fold into citric acid – K_2HPO_4 buffer pH 4.0 on ice. The percentage of Hsp90 K58H modification was quantified by intact-protein LC-MS analysis. (c) As in (b), but with Hsp90 WT. Each data point in (b) and (c) represents the mean value (± SD) from triplicate samples.

Given that the average pK_a of a histidine residue (6.45) is much lower than that of a lysine residue (10.68)[86], we also tested the effect of pH on the reaction kinetics of AC45 with both Hsp90 WT and K58H. While the rate of reaction with the WT lysine showed a strong pH dependence, reacting more rapidly at higher pH (**Fig. 3.4a**), the rate of reaction with K58H was relatively unaffected by pH (**Fig.3.4b**). As position 58 is fully solvated and not engaged in polar interactions with other side chains, we anticipate the pK_a of Lys58 and His58 to be minimally perturbed and thus close to the average

value for lysines and histidines, respectively. This puts the range of tested pH values above the anticipated pK_a of the histidine, meaning it will be predominantly deprotonated. By contrast, the tested pH values are well below the expected pK_a of the lysine, meaning it will be predominately protonated. As the pH increases, deprotonation of the lysine will be more favorable, thus increasing the reaction rate; by contrast, modification of the histidine is predicted to show a lower pH dependence. This differential effect of pH on reactivity with lysine compared to histidine has previously been observed using the fluorosulfate containing unnatural amino acid FSY[104].

In order to investigate the structural basis of the selectivity for K58H over WT Hsp90, we determined the cocrystal structure of AC45 bound to Hsp90 NTD K58H at 2.1 Å resolution (**Fig 3.5a**). The resolved electron density confirmed the presence of an adduct between the AC45 sulfonyl and the His58 Nε atom. Although the sulfonyl fluoride electrophile can react with either the Nδ or Nε atom, the Nε atom is predicted to react predominantly due to its decreased steric hindrance relative to the Nδ atom. Reaction with Nε was also seen in the crystal structure of the sulfonyl fluoride-containing peptide 155H1 bound to His252 on MCL-1[96]. Comparing the AC45-K58H structure to the previously reported AC50-WT structure (**Fig 3.5b**), the χ1 values of Lys58 and His58 are different, with Lys58 in the *trans* rotamer while His58 is in the *gauche*$^+$ rotamer. This difference in rotamers is likely an effect of the respective compound linkers adopting different preferred conformations in the binding pocket. We speculate that the ability of Lys58 and His58 to accommodate the necessary geometries to react with these different conformations results in the opposite specificities observed for AC45 and AC50.

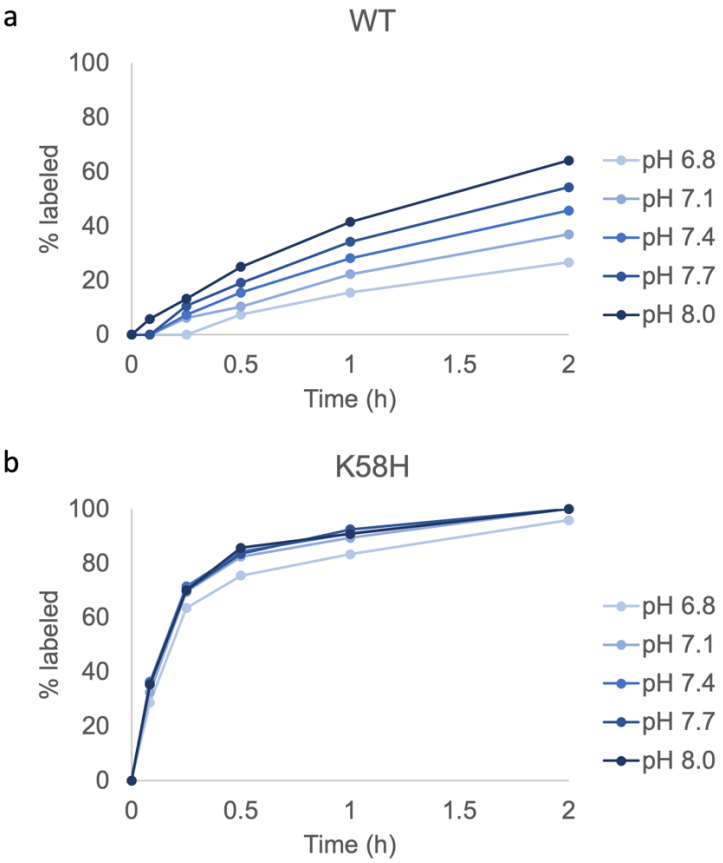


Figure 3.4. Effect of pH on sulfonyl fluoride reaction kinetics.

Hsp90 WT (a) or K58H (b) (1 μM) was treated with the AC45 (10 μM) at 37 °C. At the indicated time points, aliquots were removed and diluted 2-fold into citric acid – K_2HPO_4 buffer pH 4.0 on ice. The percentage of Hsp90 K58H modification was quantified by intact-protein LC-MS analysis.

In order to investigate the structural basis of the selectivity for K58H over WT Hsp90, we determined the cocrystal structure of AC45 bound to Hsp90 NTD K58H at 2.1 Å resolution (**Fig 3.5a**). The resolved electron density confirmed the presence of an adduct between the AC45 sulfonyl and the His58 Nε atom. Although the sulfonyl fluoride electrophile can react with either the Nδ or Nε atom, the Nε atom is predicted to react predominantly due to its decreased steric hindrance relative to the Nδ atom. Reaction

with Nε was also seen in the crystal structure of the sulfonyl fluoride-containing peptide 155H1 bound to His252 on MCL-1[96]. Comparing the AC45-K58H structure to the previously reported AC50-WT structure (**Fig 3.5b**), the χ1 values of Lys58 and His58 are different, with Lys58 in the *trans* rotamer while His58 is in the *gauche*$^+$ rotamer. This difference in rotamers is likely an effect of the respective compound linkers adopting different preferred conformations in the binding pocket. We speculate that the ability of Lys58 and His58 to accommodate the necessary geometries to react with these different conformations results in the opposite specificities observed for AC45 and AC50.

The relative instability of sulfonyl fluorides and their unsuitability for *in vivo* experiments is a limiting factor for their development. Fluorosulfates have previously been utilized as a more stable alternative, but the inclusion of an additional oxygen atom changes the geometry of the electrophile. In addition, fluorosulfates tend to react with proximal nucleophiles much more slowly, on the order of hours to days[76,95]. *Ortho*-methoxy aryl sulfonyl fluorides, on the other hand, have minimal changes to the electrophile geometry. They are less reactive than their unsubstituted counterparts[16] and have been utilized in the unnatural amino acid SFY and incorporated into proteins for use in chemical crosslinking[99,100,104]. Sulfonyl imidazoles are another analogous electrophile to sulfonyl fluorides that retain similar geometry but have been shown to have up to 10-fold greater plasma stability[105].

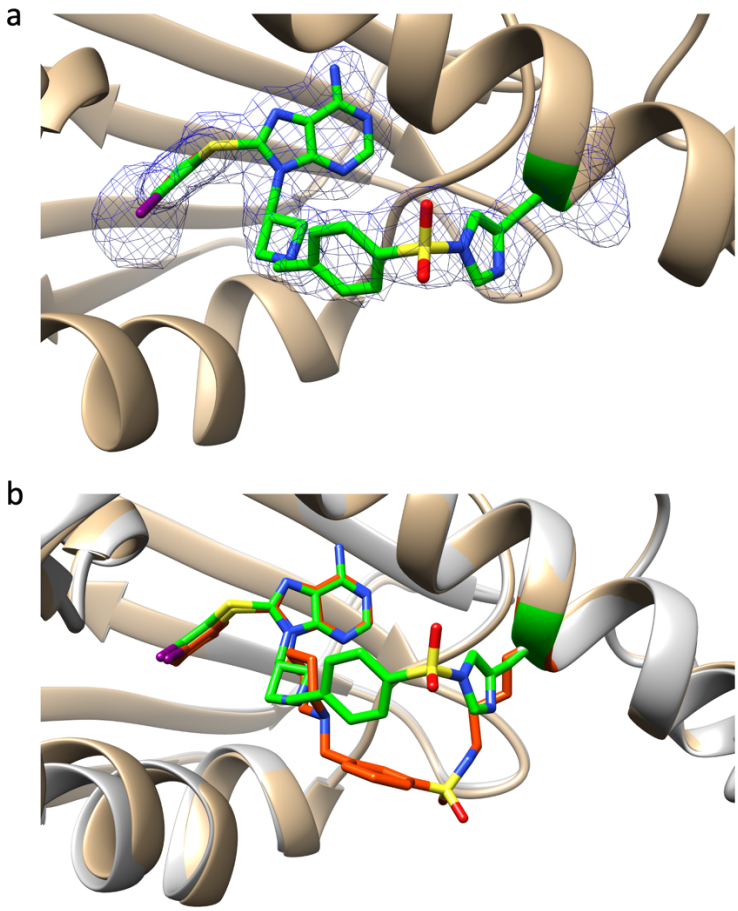


Figure 3.5: Structure of AC45 covalently bound to Hsp90 K58H.

(a) Crystal structure of AC45 covalently bound to Hsp90 K58H at 2.1 Å. Electron density map (2Fo-Fc, 1.0 σ) is shown for AC45 and His58. (b) Overlay of AC45 modified Hsp90 K58H with AC50 modified Hsp90 WT (PDB: 6U9B).

The relative instability of sulfonyl fluorides and their unsuitability for *in vivo* experiments is a limiting factor for their development. Fluorosulfates have previously been utilized as a more stable alternative, but the inclusion of an additional oxygen atom changes the geometry of the electrophile. In addition, fluorosulfates tend to react with proximal nucleophiles much more slowly, on the order of hours to days[76,95]. *Ortho*-methoxy aryl sulfonyl fluorides, on the other hand, have minimal changes to the electrophile geometry. They are less reactive than their unsubstituted counterparts[16]

and have been utilized in the unnatural amino acid SFY and incorporated into proteins for use in chemical crosslinking[99,100,104]. Sulfonyl imidazoles are another analogous electrophile to sulfonyl fluorides that retain similar geometry but have been shown to have up to 10-fold greater plasma stability[105].

To explore sulfonyl fluoride alternatives with increased stability, we synthesized the *ortho*-methoxy substituted sulfonyl fluoride, JW-2-63, and the sulfonyl imidazole, JW-2-183 (**Fig 3.6a**). We compared the stability of these compounds and AC45 towards 20 mM glutathione (GSH) in buffer at pH 7.5, monitoring loss of the parent electrophile by liquid chromatography mass spectrometry. Both JW-2-63 and JW-2-183 showed significantly improved stability in the presence of GSH than AC45, with <20% reacted after 4 h for both compounds compared to $T_{1/2}$ ~ 2 h for AC45 (**Fig 3.6b**). Despite this increased stability, JW-2-63 and JW-2-183 still showed useful rates of modification, with the K58H $T_{1/2}$ increasing to ~30 min for JW-2-63 (~5-fold increase) and ~2 h for JW-2-183 (~20-fold increase) (**Fig 3.6c**). Notably, these weaker electrophiles showed increased selectivity for the K58H mutant over the WT, particularly at later time points. For example, after 4 h, JW-2-63 showed ~5-fold and JW-2-183 approximately ~15-fold selectivity between K58H and WT, whereas AC45 showed less than a 2-fold difference (**Fig 3.6c, Fig 3.6d**). This ability to trade reactivity for selectivity has been previously observed in covalent BTK inhibitors, where a methyl acrylamide showed greater selectivity than the corresponding acrylamide in both concentration-dependent and time-dependent experiments[106]. As we have shown, the ability to tune electrophile reactivity is a powerful approach for designing covalent inhibitors with increased stability and selectivity.

In summary, I have demonstrated how conformationally constrained linkers can influence the reactivity of a sulfonyl fluoride towards a target histidine, irrespective of the innate reactivity. Our results show the potential for exploiting the sulfonyl fluoride's higher reactivity during initial probe development, and then later tuning the electrophile reactivity for increased stability and selectivity. Compared to fluorosulfates, *ortho*-methoxy sulfonyl fluorides and sulfonyl imidazoles allow for a more similar replacement of the electrophile, as they maintain the same distance and geometry as a sulfonyl fluoride relative to the nucleophilic histidine. Further testing of these stabilized electrophiles is needed in order to determine their suitability for cellular experiments.

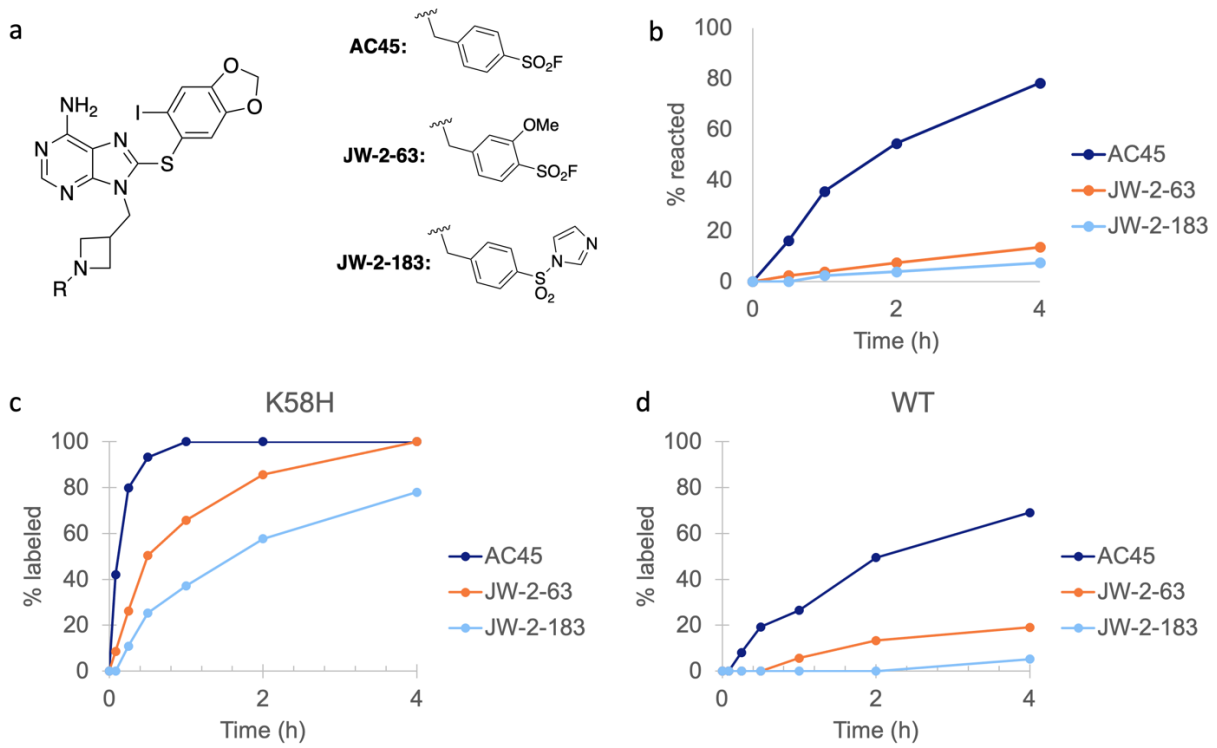


Figure 3.6: JW-2-63 and JW-2-183 show increased stability and retain Hsp90 K58H labeling.

(a) Structure of azetidine linker probe series. (b) Compounds (100 μM) were treated with GSH (20 mM) at pH 7.5 in 50% ACN at 37 °C. At the indicated time points, aliquots were removed, briefly cooled on ice, and the percentage of hydrolyzed product was determined LC-MS analysis and UV absorbance (check wavelength). (c) Hsp90 K58H (1 μM) was treated with the indicated compounds (10 μM) at 37 °C. At the indicated time points, aliquots were removed and diluted 2-fold into citric acid – K_2HPO_4 buffer pH 4.0 on ice. The percentage of Hsp90 K58H modification was quantified by intact-protein LC-MS analysis. (d) As in (c), but with Hsp90 WT. Each data point in (b), (c), and (d) represents the mean value (± SD) from triplicate samples.

Methods

Protein expression and purification. Hsp90 N-terminal domain (NTD) was expressed and purified as previously described[11]. K58H Hsp90AA1 NTD plasmid was prepared by site directed mutagenesis, and then expressed and purified using the same protocol.

Time course of Hsp90 NTD modification. Compounds (10 μM) were added to 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$) at 37 ^oC. At each time point an aliquot was removed and diluted 2-fold into citric acid – K_2HPO_4 buffer pH 4.0 and stored on ice. Samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of modified Hsp90.

pH dependent time course of Hsp90 NTD modification. AC45 (10 μM) was added to 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES, 50 mM KCl, 25 mM $MgCl_2$) at the indicated pH (6.8 – 8.0) at 37 ^oC. At each time point an aliquot was removed and diluted 2-fold into citric acid – K_2HPO_4 buffer pH 4.0 and stored on ice. Samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. LC-MS runs were auto-processed using the Waters OpenLynx software to determine the percentage of modified Hsp90.

Co-crystallization and structure determination of AC45 bound to Hsp90 NTD K58H. Hsp90 NTD K58H was covalently modified by incubating 10 μM AC45 with 1 μM Hsp90 NTD in labeling buffer (20 mM HEPES pH 7.4, 50 mM KCl, 25 mM $MgCl_2$). Complete modification was confirmed by LC-MS. Modified protein was buffer

exchanged into 50 mM HEPES pH 8.0, 50 mM KCl, 1 mM DTT with 10% glycerol and concentrated to 20 mg/mL using a 10K MWCO filter (Millipore Amicon Ultra). Crystallization was performed using the hanging drop method with a 1:1 mixture of 20 mg/mL protein-compound adduct and crystallization solution containing 25% PEG-3350, 0.2 M $MgCl_2$, 0.1 M Bis-Tris pH 5.5. Protein crystals formed within 48 h. X-ray diffraction data was collected at beamline 8.3.1 of the Advanced Light Source (Lawrence Berkley National Laboratory). The resulting X-ray diffraction data was processed using the XDS package[34], followed by molecular replacement with PHASER[35] using PDB code: 6U9B as a starting model. Structure refinement was performed using coot[36] and phenix.refine[35].

Time course of GSH stability. Compounds (100 μM) were treated with GSH (20 mM) at pH 7.5 in PBS with 50% ACN at 37 ^{o}C. At the indicated time points, aliquots were removed, briefly cooled on ice, and samples were analyzed on a Waters Xevo G2-XS QToF LC-MS. The percentage of aryl sulfinic acid product was tracked by integration of the product UV absorbance (254 nm) peak.

Synthetic Methods

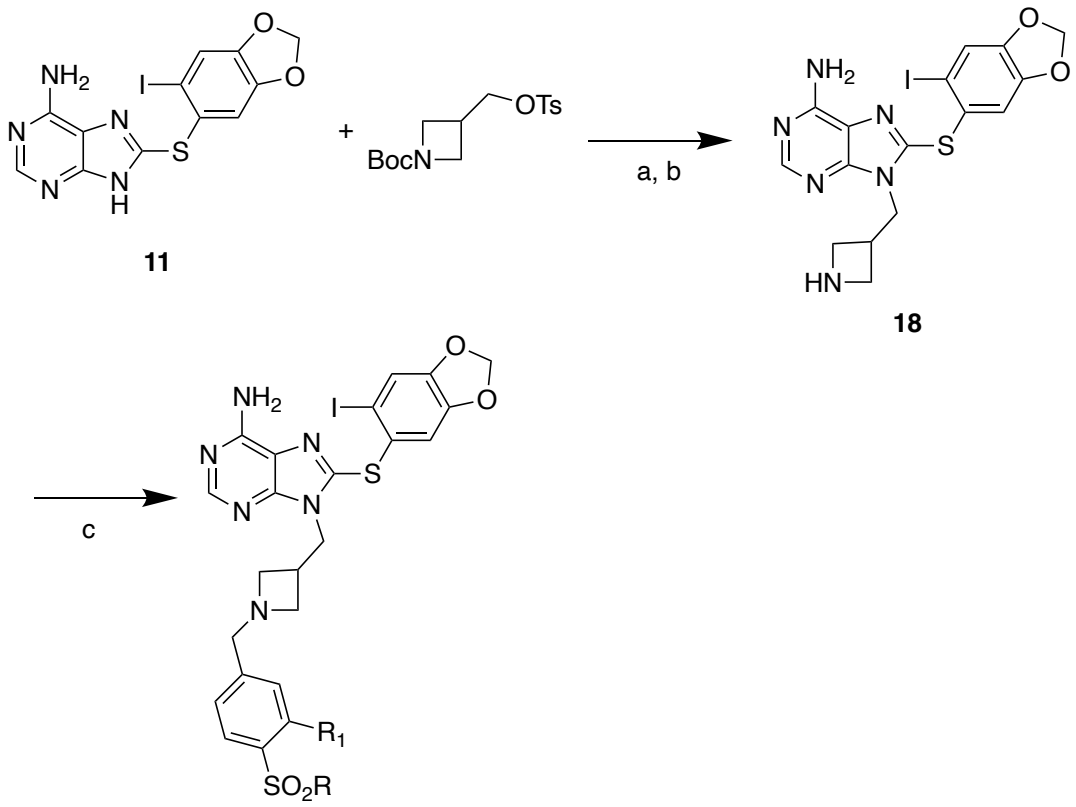
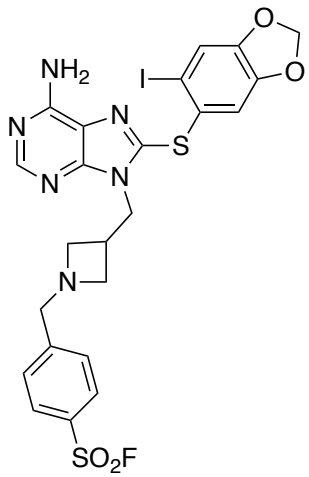


Figure 3.7. Synthetic scheme 3.1.

Scheme 3.1: a) Cs_2CO_3, DMF, 80 °C, 16 h. b) TFA:CH_2Cl_2 1:1, 0 °C, 1 h. c) DIPEA, DMF, RT, 1 h. R = F or imidazole. R_1 = H or OMe.

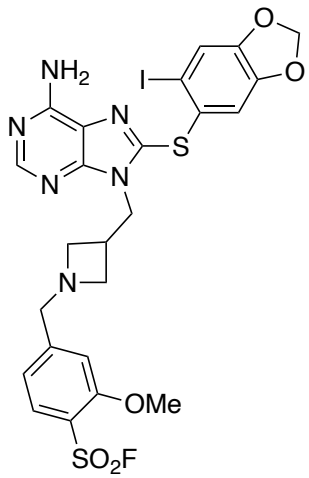
4-((3-((6-amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)azetidin-1-yl)methyl)benzene-1-sulfonyl fluoride (AC45): 18 (30 mg, 0.062 mmol) and DIPEA (26 µL, 0.12 mmol) were dissolved in DMF (0.5 mL), followed by the addition of a solution of 4-(bromomethyl) benzene-1-sulfonyl fluoride (19 mg, 0.074 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was

removed under vacuum and the product was purified by silica gel chromatography in 0-



50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 11 mg, 27 %. ^{1}H NMR (400 MHz, DMSO) δ 8.17 (s, 1H), 8.06 (d, *J* = 8.4 Hz, 2H), 7.65 (d, *J* = 8.4 Hz, 2H), 7.49 (s, 1H), 7.40 (s, 2H), 6.83 (s, 1H), 6.06 (s, 1H), 4.39 (d, *J* = 6.8 Hz, 2H), 3.70 (s, 2H), 3.24 (t, *J* = 7.2 Hz, 2H), 3.03 (t, *J* = 6.0 Hz), 2.92 – 2.84 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 155.26, 153.14, 151.19, 148.94, 148.65, 144.27, 141.50, 136.28, 131.74, 131.56, 130.56, 128.33, 127.31, 126.95, 119.55, 118.80, 111.42, 102.56, 91.34, 60.31, 57.23, 44.40, 43.08. HRMS (ESI-TOF) *m/z* calcd. $C_{23}H_{21}FIN_6O_4S_2^+$ [M + H]$^+$, 655.0050; found, 655.0027.

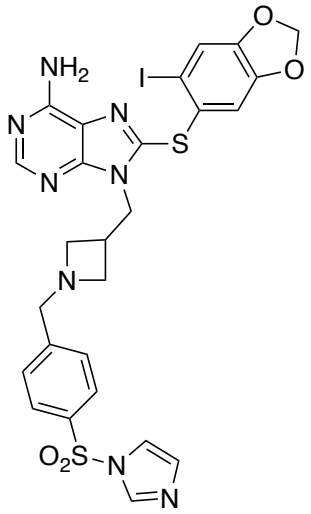
4-((3-((6-amino-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-9-yl)methyl)azetidin-1-yl)methyl)-2-methoxybenzene-1-sulfonyl fluoride (JW-2-63):



18 (14 mg, 0.028 mmol) and DIPEA (24 µL, 0.11 mmol) were dissolved in DMF (0.5 mL), followed by the addition of a solution of **20** (7.4 mg, 0.031 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 3.0 mg, 15 %. ^{1}H NMR (400 MHz, DMSO) δ 8.15 (s, 1H), 7.84 (d, *J* = 8.2 Hz, 1H), 7.47 (s, 1H), 7.40 (s, 2H), 7.27 (s, 1H), 7.10 (d, *J* = 8.2 Hz, 1H), 6.83 (s, 1H), 6.05 (s,

2H), 4.39 (d, *J* = 7.0 Hz, 2H), 3.99 (s, 3H), 3.66 (s, 2H), 3.26 (t, *J* = 7.3 Hz, 2H), 3.02 (t, *J* = 6.5 Hz, 2H), 2.94 – 2.84 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 157.71, 155.28, 153.12, 151.14, 150.65, 148.87, 148.49, 144.00, 130.60, 128.49, 120.13, 119.53, 118.70, 117.78 (d, *J* = 22.0 Hz), 112.81, 111.11, 102.46, 90.95, 61.82, 57.38, 56.90, 46.20. HRMS (ESI-TOF) *m/z* calcd. $C_{24}H_{23}FIN_6O_5S_2^+$ [M + H]$^+$, 685.0155; found, 685.0128.

9-((1-(4-((1*H*-imidazol-1-yl)sulfonyl)benzyl)azetidin-3-yl)methyl)-8-((6-iodobenzo[*d*][1,3]dioxol-5-yl)thio)-9*H*-purin-6-amine (JW-2-183): 18 (25 mg, 0.052 mmol) and DIPEA (22 µL, 0.10 mmol) were dissolved in DMF (0.5 mL), followed by the addition of a solution of **21** (19 mg, 0.062 mmol) in DMF (0.5 mL). The reaction was left to stir at RT for 1 h. Solvent was removed under vacuum and the product was purified by silica gel chromatography in 0-50% gradient of solvent A in dichloromethane (solvent A: 20% methanol, 2% sat. aqueous ammonia in dichloromethane). Yield: 7.5 mg, 21 %. ^{1}H NMR (500 MHz, DMSO) δ 8.37 (s, 1H), 8.15 (s, 1H), 8.02 (d, *J* = 8.4 Hz, 2H), 7.75 (t, *J* = 1.6 Hz, 1H), 7.55 (d, *J* = 8.2 Hz, 2H), 7.47 (s, 1H), 7.41 (s, 2H), 7.12 (d, *J* = 1.8 Hz, 1H), 6.81 (s, 1H), 6.04 (s, 2H), 4.36 (d, *J* = 7.1 Hz, 2H), 3.66 (s, 2H), 3.23 (t, *J* = 7.2 Hz, 2H), 3.03 (t, *J* = 6.4 Hz, 2H), 2.91 – 2.83 (m, 1H). ^{13}C NMR (126 MHz, DMSO) δ 155.27, 153.11, 151.12, 148.85, 148.45, 143.92, 137.25, 135.58, 131.33, 129.44, 128.50, 127.35, 119.52, 118.68, 118.27, 111.02, 102.45, 90.85, 61.38, 53.57, 46.12, 41.82,



30.70, 16.76, 12.53. HRMS (ESI-TOF) *m/z* calcd. $C_{26}H_{24}IN_8O_4S_2^+$ [M + H]$^+$, 703.0362; found, 703.0461.

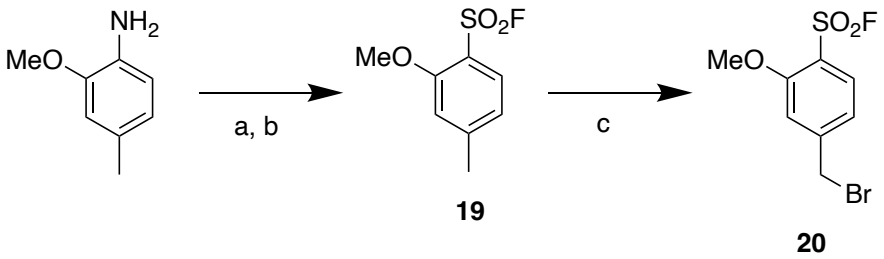
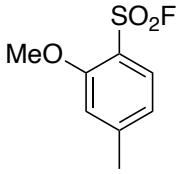


Figure 3.8. Synthetic scheme 3.2.

Scheme 3.2: a) Thionyl chloride, $NaNO_2$, $CuCl_2$, acetic acid, HCl, ACN, 0 ^oC -> RT, 16 h. b) KF•HF, dioxane, RT, 16 h. c) NBS, AIBN, CCl_4, 80 ^oC, 16 h.

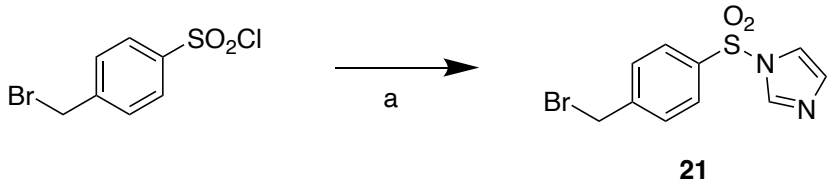
2-methoxy-4-methylbenzene-1-sulfonyl fluoride (19): Thionyl chloride (5 mL, 69



mmol) was added dropwise into water (30 mL) at 0 ^oC over 30 min. In a separate flask, 2-methoxy-4-methylaniline (2.0 g, 14.6 mmol) was dissolved in ACN (117 mL), and acetic acid (11.7 mL) and 37% aqueous HCl (5.7 mL) were added. This solution was cooled to 0 ^oC and a solution of $NaNO_2$ (1.17 g, 17.0 mmol) in water (4 mL) was added dropwise over 30 min. The thionyl chloride solution was then added dropwise over 30 min. The reaction was stirred at 0 ^oC for 20 min and a solution of $CuCl_2$ (3.0 g, 17.6 mmol) in water (4 mL) was added dropwise over 30 min. The reaction was stirred for another 30 min at 0 ^oC and then warmed to RT and stirred for another 16 h. The ACN was removed under vacuum and the remaining solution was poured into water and extracted with ethyl acetate 3 times. The organic fractions were pooled, then dried over Na_2SO_4, and the solvent was removed under vacuum. The crude mixture was dissolved in dioxane (20 mL) and then added to an aqueous solution of 2 M KF•HF (10 mL). The reaction was stirred at RT for

16 h, filtered, poured into water, and extracted 3 times with ethyl acetate. The organic fractions were pooled, dried over Na_2SO_4, and the solvent removed under vacuum to give the pure product. Yield: 1.03 g, 34%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.80 (d, *J* = 8.0 Hz, 1H), 6.90 (d, *J* = 10.1 Hz, 2H), 3.99 (s, 3H), 2.45 (s, 3H). HRMS (ESI-TOF) m/z calcd. $C_8H_{10}FO_3S^+$ [M + H]$^+$, 205.0329; found, 205.0287.

4-(bromomethyl)-2-methoxybenzene-1-sulfonyl fluoride (20): 19 (450 mg, 2.2 mmol) was dissolved in CCl_4 (12 mL). AIBN (36 mg, 0.22 mmol) and then NBS (430 mg, 2.4 mmol) were added. The solution was heated to 80 °C and left to stir for 16 h. The solution was cooled to RT, filtered, and the solvent removed under vacuum. The product was purified by silica gel chromatography in 0-30% gradient of ethyl acetate in hexane. Yield: 250 mg, 40%. ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90 (d, *J* = 8.5 Hz, 1H), 7.15 – 7.07 (m, 2H), 4.47 (s, 2H), 4.03 (s, 3H). HRMS (ESI-TOF) m/z calcd. $C_8H_9BrFO_3S^+$ [M + H]$^+$, 282.9395; found, 282.9412.



21

Figure 3.9. Synthetic scheme 3.3.

Scheme 3.3: a) imidazole, DCM, RT, 1 h.

1-((4-(bromomethyl)phenyl)sulfonyl)-1*H*-imidazole (21): 4-(bromomethyl) benzene-1-sulfonyl chloride 200 mg, 0.75 mmol) was dissolved in DCM (8 mL) and imidazole (100 mg, 2.9 mmol) was added. The reaction

was stirred at RT for 1 h. The reaction was filtered and the solvent was removed under vacuum. Yield: 110 mg, 49%.^{1}H NMR (500 MHz, $CDCl_3$) δ 8.03 (s, 1H), 7.95 – 7.89 (m, 2H), 7.61 – 7.56 (m, 2H), 7.31 (d, *J* = 1.5 Hz, 1H), 7.11 (d, *J* = 1.7 Hz, 1H), 4.47 (s, 2H). HRMS (ESI-TOF) m/z calcd. $C_{10}H_{10}BrN_2O_2S^+$ $[M + H]^+$, 300.9602; found, 300.9598.

Chapter 4 : Mining the Protein DataBank for ligandable histidines

Abstract

While cysteine remains the target of choice for covalent inhibitors owing to its intrinsic reactivity, cysteine has a low prevalence in the proteome and many proteins lack a targetable cysteine proximal to a small molecule binding site. Covalent targeting of alternative nucleophilic acids, like lysine, histidine, and tyrosine, greatly expands the breadth of potential protein targets. While the extent of ligand-proximal cysteines, lysines, and tyrosines has been previously explored, the number of ligand-proximal histidines remains unknown. Herein, we searched 65,288 mammalian protein structures from the Protein DataBank (PDB) to identify ligands with proximal histidine, cysteine, lysine, and tyrosine residues. We found that 18% of the unique proteins included in the search have at least one ligand-proximal histidine. We highlight SphK1 and IDH1 as two potential protein targets for the rational design of histidine-targeted covalent inhibitors.

Introduction

Inclusion of an irreversible or reversible covalent electrophile on a small molecule drug is an effective method to increase compound potency and prolong residence time on its target protein. While most covalent inhibitors target cysteines, only approximately 20% of ligand binding sites contain a cysteine for covalent targeting[85]. Designing covalent inhibitors that target other amino acids, such as lysine, tyrosine, and histidine, is more difficult due to their lower intrinsic nucleophilicity. However, such a strategy has the potential to greatly expand the breadth of proteins that can be targeted with covalent inhibitors.

Work from the Chandra group in 2018 utilized three complementary algorithms (SiteHound, PocketDepth and FPocket) to predict small-molecule binding pockets contained in 113,486 structures from the Protein DataBank (PDB) (from May 2015)[107–110]. In brief, Sitehound is an energy-based method that looks for patches of favorable interactions on the protein surface; PocketDepth is a geometry-based method that calculates pocket depth; and FPocket is a Voronoi tessellation-based algorithm that generates spheres that simultaneously contact 4 amino acids on the protein surface and predicts pockets based on how these spheres cluster. Using the compiled dataset of binding pockets, known as the augmented pocketome, Cuesta analyzed the frequency of specific amino acids in the representative binding sites[85]. Although lysine is more than twice as prevalent as histidine in the overall proteome, the frequency of histidine in small molecule binding pockets is equivalent to that of lysine. Soga et al. developed a method to predict ligand binding sites based on amino acid composition and found that histidine has the second highest propensity for ligand binding among the twenty standard amino acids, falling only behind tryptophan[84]. The frequent occurrence of histidine near small-molecule binding sites motivates the exploration of proteins with histidine residues proximal to small-molecule ligands for covalent probe development.

While Cuesta previously identified small-molecule ligands within 5 Å of a cysteine, lysine, or tyrosine residue by searching structures from the PDB, the extent of ligands proximal to histidines is still unknown. In this chapter, we mined the PDB for histidine residues proximal to small-molecule ligands with the goal of identifying ‘ligandable’ histidines on exciting protein targets for future covalent probe development. From this search, we identified 2,088 unique proteins containing a ligand-proximal

histidine, 194 of which did not contain any other proximal nucleophilic residues. We further examine two of the identified proteins, sphingosine kinase 1 and isocitrate dehydrogenase 1, as potential targets for histidine-targeted covalent inhibitors.

Results

In order to explore the extent of targetable histidine residues in the proteome, we utilized all mammalian protein structures in the PDB with a resolution of 3 Å or better for our analysis (from September 2024). We searched these structures for small-molecule ligands, defined as any atom denoted as 'HETATM' in the structure file. From the identified ligands, we filtered out 'bad ligands' such as ions, buffers, and known detergents and precipitants. A full list of excluded ligands can be found in the script included in appendix A. Of the 65,288 protein structures that fit the search criteria, 36,288 contained at least one ligand, totaling 56% of the structures. However, this number is heavily biased by the nature of the dataset, as many structures in the PDB are of redundant proteins, generally those that crystallize readily, are frequently studied, and/or have lots of known ligands. To reduce the effects of this bias, we used UniProt IDs to count the number of unique proteins. 11,904 unique proteins were included in the search, of which 5,224 had at least one structure that contained a ligand, accounting for 44% of the proteins.

From the structures identified as containing one or more ligands, we identified all histidine residues that were proximal to a ligand. Proximal was defined as at least one ligand atom being within a 5 Å distance of either the histidine Nδ atom or Nε atom. This search was repeated for cysteine, lysine, and tyrosine, with the distance being defined

to the thiol sulfur, amine nitrogen, and phenol oxygen, respectively. A full list of identified residues, with the PDB code, ligand name, residue name, distance between ligand and residue, and UniProt ID can be found in the supplementary files. Of the 5,224 proteins with a known ligand, 3868 of them were found to have at least one of a proximal histidine, cysteine, lysine, or tyrosine, totaling 74% (**Fig 4.1a**). 2,088 proteins were identified as having a proximal histidine, 18% of the proteins included in the search and 40% of those containing a ligand. Of these 2,088 proteins, 194 did not contain a proximal cysteine, lysine, or tyrosine.

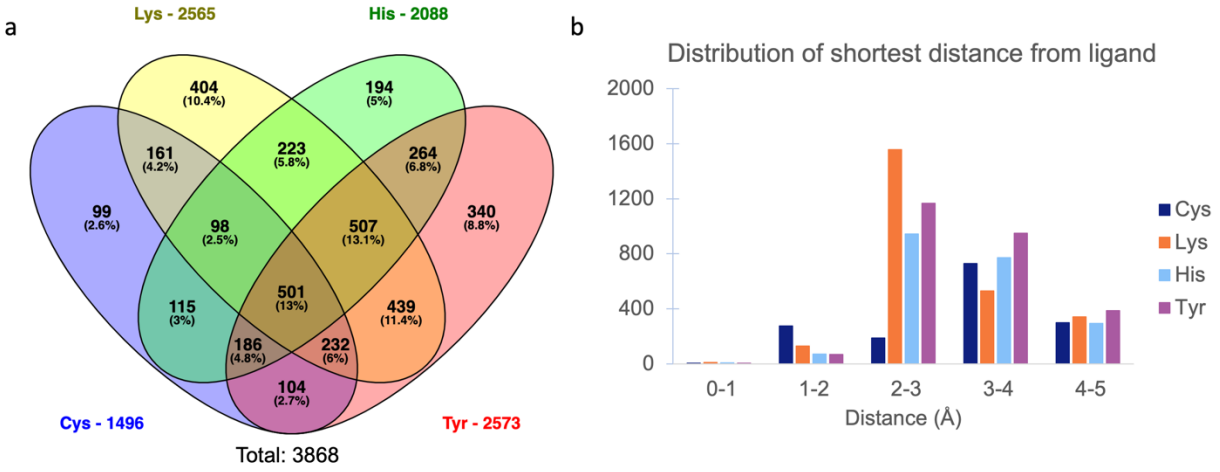


Figure 4.1. Overview of proximal ligand search results.

(a) Venn diagram of unique proteins found with proximal cysteines, lysines, histidines, and tyrosines. Diagram created using Venny[111]. (b) Histogram of the shortest distance found between a ligand and proximal residue within a given protein.

Looking at the shortest distance observed between a ligand and given residue for each individual protein, most of the liganded proteins fall within the 2-3 Å range for each residue type excluding cysteine, with the number of additional liganded proteins decreasing as the distance increases (**Fig 4.1b**). This peak at 2-3 Å can be explained in most cases by hydrogen bonding interactions between the nucleophilic residue and the

ligand. Looking at the full distribution of distances between ligands and proximal histidines, there exist two major peaks at approximately 2.8 Å and 3.4 Å (**Fig 4.2c**). The 2.8 Å peak mirrors those found in lysine and tyrosine (**Fig 4.2b, Fig 4.2d**), likely indicative of hydrogen bonding interactions. The 3.4 Å peak, likewise, mirrors those found in cysteine and tyrosine (**Fig 4.2a, Fig 4.2d**), and likely indicates hydrophobic/van der Waals interactions.

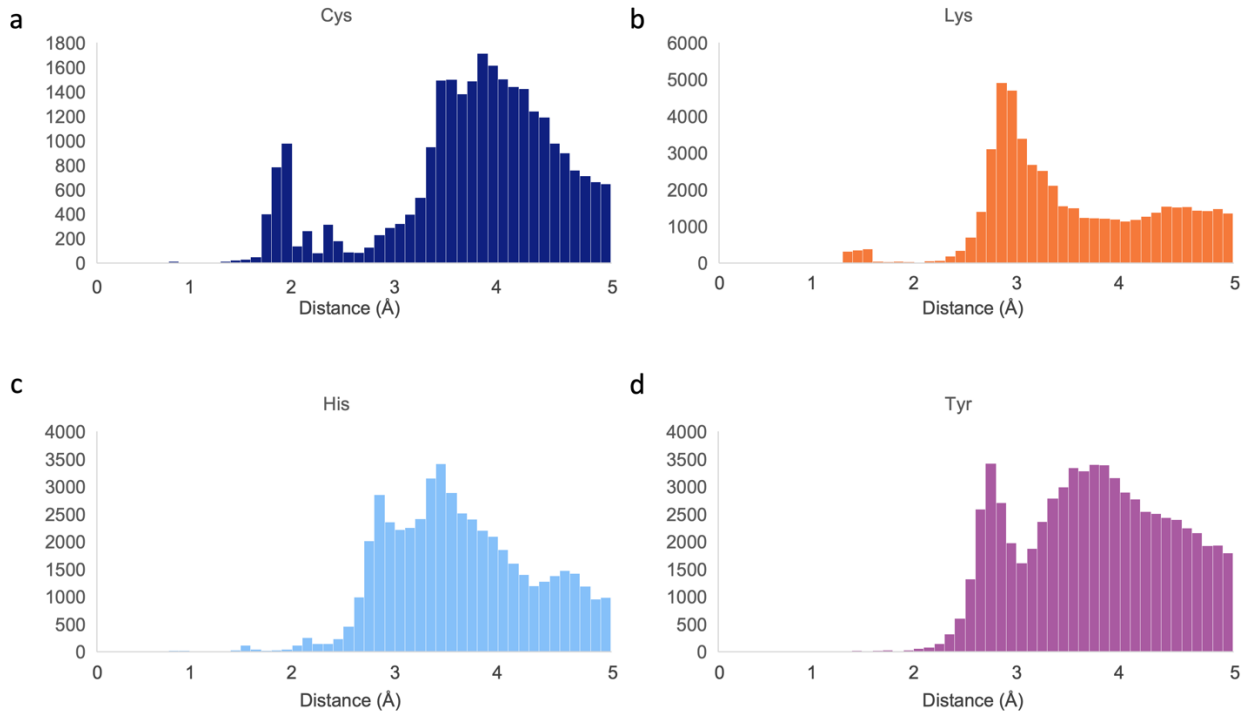


Figure 4.2. Distribution of proximal ligand distances.

Histograms of the measured distance between ligand and the nucleophilic atom of cysteine (a), lysine (b), histidine (c), and tyrosine (d) for all proximal sites found.

Among the 194 proteins identified by our search to uniquely contain a proximal histidine, sphingosine kinase 1 (SphK1) is one example of an interesting target for the development of a covalent inhibitor. SphK1, along with its paralog SphK2, is a lipid kinase that phosphorylates sphingosine to sphingosine-1-phosphate (S1P). S1P, once

produced intracellularly, is excreted from the cell by specific transporters. There it signals through G-protein coupled receptors to help regulate cell-cell and cell-matrix adhesion, among other roles[112,113]. In cancer, S1P is involved in processes that are strongly linked to patient mortality, such as neovascularization and metastasis[114–116]. SphK1 is also often overexpressed in cancers such as lung, gastric, and breast cancers[113,114]. Overexpression of SphK1 has also been associated with drug resistance in cancer[117–121]. Currently, there are no known covalent inhibitors or probes targeting SphK1. PF-543 is a competitive inhibitor of SphK1 that binds in the lipid binding pocket with a K_I of 3.6 nM[122]. It also has a 100-fold selectivity for SphK1 over SphK2. In the structure of PF-543 bound to SphK1, His397 is 3.6 Å away from a benzene ring in PF-543 (**Fig 4.3**). It is possible that installing a sulfonyl fluoride directly onto this benzene ring would yield a covalent inhibitor of SphK1, denoting it as a good starting point for covalent probe design.

Isocitrate dehydrogenase 1 (IDH1) was also identified by our search as a potential target for the development of histidine-targeted covalent inhibitors. Isocitrate dehydrogenases are metabolic enzymes that catalyze the conversion of isocitrate into α-ketoglutarate (αKG). IDH1 is the cytoplasmic isoform and is frequently mutated at position 132 in gliomas, including 88% of secondary glioblastomas and 68% of low-grade gliomas[123,124]. These mutants lose wild-type activity and gain the ability to catalyze the conversion of α-KG into the oncometabolite 2-hydroxyglutarate[125]. The specific IDH1 mutant R132H is by far the most common mutation at this position,

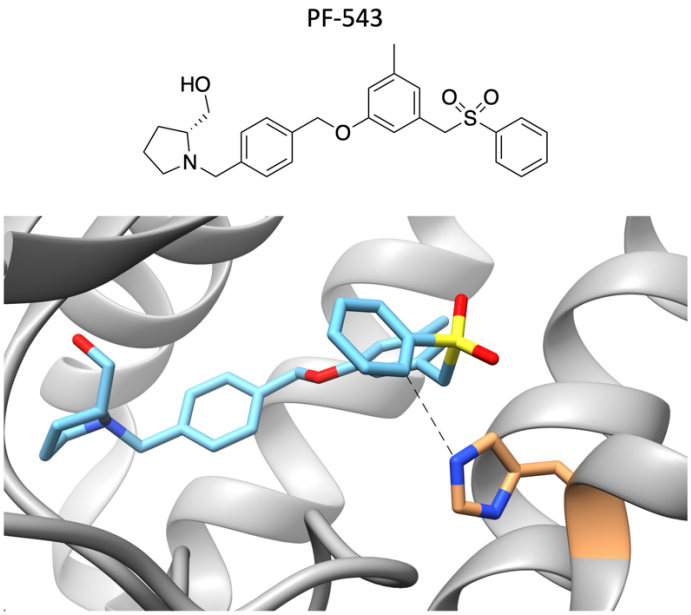


Figure 4.3. SphK1 His397 as a target for covalent inhibitors.

Crystal structure of PF-543 bound to SphK1 (PDB: 4V24). PF-543 (shown in blue) is located 3.6 Å away from the Nε atom of His397 (shown in orange).

comprising 93% of the R132 mutants in brain tumors[124]. To date, three IDH1 mutant inhibitors have been clinically approved, but due to the difficulties in designing drugs capable of crossing the blood-brain barrier, only one has been approved for treatment of glioma[126–128]. A cysteine-targeted covalent IDH1 inhibitor, LY3410738, entered phase I clinical trials but has since been removed from the development pipeline[129–131]. Covalent targeting of His132 has the potential to yield a compound with exquisite mutant selectivity and potency. Analysis of crystal structures of small molecules bound to IDH1 R132H reveals several starting points for the development of a histidine-targeted covalent inhibitor. IHMT-IDH1-053, a covalent vinyl sulfonamide inhibitor targeting Cys269 adjacent to His132, had an IC50 of 28 nM in cellular assays[132] and the crystal structure reveals close proximity between the sulfonamide oxygen and His132 (3.4 Å) (**Fig 4.4a**). Additionally, the structure of compound 1, which is closely related to

BAY1436032, a pan-IDH1 inhibitor with an IC50 of 15 nM for the R132H mutant[133], revealed a 2.9 Å distance between His132 and the carbonyl group (**Fig 4.4b**). Either of these compounds would be a good starting point for rational design of a covalent inhibitor targeting His132.

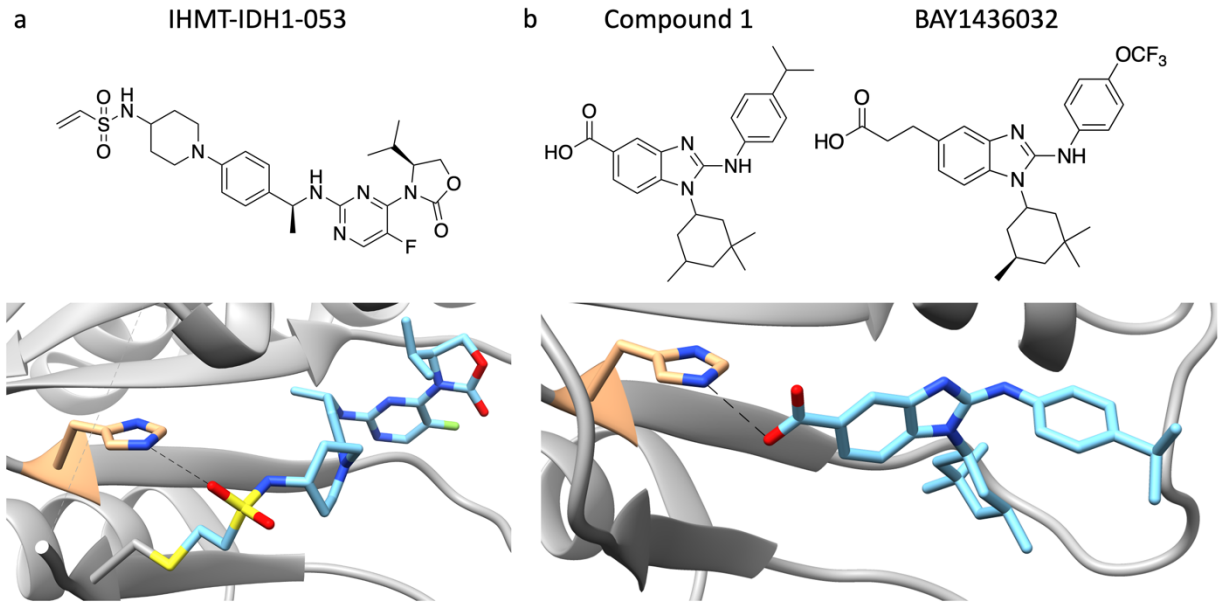


Figure 4.4. IDH1 R132H as a target for covalent inhibitors.

(a) Crystal structure of IHMT-IDH1-053 bound to IDH1 R132H (PDB: 8HB9). IHMT-IDH1-053 (shown in blue) is located 3.4 Å away from the Nε atom of His132 (shown in orange). (b) Crystal structure of compound 1 bound to IDH1 R132H (PDB: 5LGE). Compound 1 (shown in blue) is located 2.9 Å away from the Nε atom of His132 (shown in orange).

Overall, the results of our search provide a basis for identifying the 'low-hanging fruit' for a histidine-targeted covalent inhibitor approach: proteins with known small-molecule ligands that have a proximal histidine residue. 194 proteins identified with proximal histidines had no other proximal nucleophilic residues, exemplifying the need to develop covalent targeting motifs for a breadth of amino acid nucleophiles. Moving

forward, covalent targeting of histidine residues shows great promise, opening up avenues for "best in class" covalent inhibitors of drug targets like SphK1 and IDH1.

References

(1) Honigberg, L. A.; Smith, A. M.; Sirisawad, M.; Verner, E.; Loury, D.; Chang, B.; Li, S.; Pan, Z.; Thamm, D. H.; Miller, R. A.; Buggy, J. J. The Bruton Tyrosine Kinase Inhibitor PCI-32765 Blocks B-Cell Activation and Is Efficacious in Models of Autoimmune Disease and B-Cell Malignancy. *Proceedings of the National Academy of Sciences* **2010**, *107* (29), 13075–13080. https://doi.org/10.1073/pnas.1004594107.

(2) Byrd, J. C.; Furman, R. R.; Coutre, S. E.; Flinn, I. W.; Burger, J. A.; Blum, K. A.; Grant, B.; Sharman, J. P.; Coleman, M.; Wierda, W. G.; Jones, J. A.; Zhao, W.; Heerema, N. A.; Johnson, A. J.; Sukbuntherng, J.; Chang, B. Y.; Clow, F.; Hedrick, E.; Buggy, J. J.; James, D. F.; O'Brien, S. Targeting BTK with Ibrutinib in Relapsed Chronic Lymphocytic Leukemia. *New England Journal of Medicine* **2013**, *369* (1), 32–42. https://doi.org/10.1056/NEJMoa1215637.

(3) Finlay, M. R. V.; Anderton, M.; Ashton, S.; Ballard, P.; Bethel, P. A.; Box, M. R.; Bradbury, R. H.; Brown, S. J.; Butterworth, S.; Campbell, A.; Chorley, C.; Colclough, N.; Cross, D. A. E.; Currie, G. S.; Grist, M.; Hassall, L.; Hill, G. B.; James, D.; James, M.; Kemmitt, P.; Klinowska, T.; Lamont, G.; Lamont, S. G.; Martin, N.; McFarland, H. L.; Mellor, M. J.; Orme, J. P.; Perkins, D.; Perkins, P.; Richmond, G.; Smith, P.; Ward, R. A.; Waring, M. J.; Whittaker, D.; Wells, S.; Wrigley, G. L. Discovery of a Potent and Selective EGFR Inhibitor (AZD9291) of Both Sensitizing and T790M Resistance Mutations That Spares the Wild Type Form of the Receptor. *J Med Chem* **2014**, *57* (20), 8249–8267. https://doi.org/10.1021/jm500973a.

(4) Jänne, P. A.; Yang, J. C.-H.; Kim, D.-W.; Planchard, D.; Ohe, Y.; Ramalingam, S. S.; Ahn, M.-J.; Kim, S.-W.; Su, W.-C.; Horn, L.; Haggstrom, D.; Felip, E.; Kim, J.-H.; Frewer, P.; Cantarini, M.; Brown, K. H.; Dickinson, P. A.; Ghiorghiu, S.; Ranson, M. AZD9291 in EGFR Inhibitor–Resistant Non–Small-Cell Lung Cancer. *New England Journal of Medicine* **2015**, *372* (18), 1689–1699. https://doi.org/10.1056/NEJMoa1411817.

(5) Canon, J.; Rex, K.; Saiki, A. Y.; Mohr, C.; Cooke, K.; Bagal, D.; Gaida, K.; Holt, T.; Knutson, C. G.; Koppada, N.; Lanman, B. A.; Werner, J.; Rapaport, A. S.; San Miguel, T.; Ortiz, R.; Osgood, T.; Sun, J.-R.; Zhu, X.; McCarter, J. D.; Volak, L. P.; Houk, B. E.; Fakih, M. G.; O'Neil, B. H.; Price, T. J.; Falchook, G. S.; Desai, J.; Kuo, J.; Govindan, R.; Hong, D. S.; Ouyang, W.; Henary, H.; Arvedson, T.; Cee, V. J.; Lipford, J. R. The Clinical KRAS(G12C) Inhibitor AMG 510 Drives Anti-Tumour Immunity. *Nature* **2019**, *575* (7781), 217–223. https://doi.org/10.1038/s41586-019-1694-1.

(6) Cuesta, A.; Taunton, J. Lysine-Targeted Inhibitors and Chemoproteomic Probes. *Annu Rev Biochem* **2019**, *88* (1), 365–381. https://doi.org/10.1146/annurev-biochem-061516-044805.

(7) Pettinger, J.; Jones, K.; Cheeseman, M. D. Lysine-Targeting Covalent Inhibitors. *Angewandte Chemie International Edition* **2017**, *56* (48), 15200–15209. https://doi.org/10.1002/anie.201707630.

(8) Chen, P.; Sun, J.; Zhu, C.; Tang, G.; Wang, W.; Xu, M.; Xiang, M.; Zhang, C.; Zhang, Z.; Gao, L.; Yao, S. Q. Cell-Active, Reversible, and Irreversible Covalent

Inhibitors That Selectively Target the Catalytic Lysine of BCR-ABL Kinase. *Angewandte Chemie International Edition* **2022**, *61* (26). https://doi.org/10.1002/anie.202203878.
(9) Zhao, Q.; Ouyang, X.; Wan, X.; Gajiwala, K. S.; Kath, J. C.; Jones, L. H.; Burlingame, A. L.; Taunton, J. Broad-Spectrum Kinase Profiling in Live Cells with Lysine-Targeted Sulfonyl Fluoride Probes. *J Am Chem Soc* **2017**, *139* (2), 680–685. https://doi.org/10.1021/jacs.6b08536.
(10) Quach, D.; Tang, G.; Anantharajan, J.; Baburajendran, N.; Poulsen, A.; Wee, J. L. K.; Retna, P.; Li, R.; Liu, B.; Tee, D. H. Y.; Kwek, P. Z.; Joy, J. K.; Yang, W.; Zhang, C.; Foo, K.; Keller, T. H.; Yao, S. Q. Strategic Design of Catalytic Lysine-Targeting Reversible Covalent BCR-ABL Inhibitors**. *Angewandte Chemie International Edition* **2021**, *60* (31), 17131–17137. https://doi.org/10.1002/anie.202105383.
(11) Cuesta, A.; Wan, X.; Burlingame, A. L.; Taunton, J. Ligand Conformational Bias Drives Enantioselective Modification of a Surface-Exposed Lysine on Hsp90. *J Am Chem Soc* **2020**, *142* (7), 3392–3400. https://doi.org/10.1021/jacs.9b09684.
(12) Tamura, T.; Ueda, T.; Goto, T.; Tsukidate, T.; Shapira, Y.; Nishikawa, Y.; Fujisawa, A.; Hamachi, I. Rapid Labelling and Covalent Inhibition of Intracellular Native Proteins Using Ligand-Directed N-Acyl-N-Alkyl Sulfonamide. *Nat Commun* **2018**, *9* (1), 1870. https://doi.org/10.1038/s41467-018-04343-0.
(13) Wan, X.; Yang, T.; Cuesta, A.; Pang, X.; Balius, T. E.; Irwin, J. J.; Shoichet, B. K.; Taunton, J. Discovery of Lysine-Targeted EIF4E Inhibitors through Covalent Docking. *J Am Chem Soc* **2020**, *142* (11), 4960–4964. https://doi.org/10.1021/jacs.9b10377.
(14) Dalton, S. E.; Dittus, L.; Thomas, D. A.; Convery, M. A.; Nunes, J.; Bush, J. T.; Evans, J. P.; Werner, T.; Bantscheff, M.; Murphy, J. A.; Campos, S. Selectively Targeting the Kinome-Conserved Lysine of PI3Kδ as a General Approach to Covalent Kinase Inhibition. *J Am Chem Soc* **2018**, *140* (3), 932–939. https://doi.org/10.1021/jacs.7b08979.
(15) Abbasov, M. E.; Kavanagh, M. E.; Ichu, T.-A.; Lazear, M. R.; Tao, Y.; Crowley, V. M.; am Ende, C. W.; Hacker, S. M.; Ho, J.; Dix, M. M.; Suciu, R.; Hayward, M. M.; Kiessling, L. L.; Cravatt, B. F. A Proteome-Wide Atlas of Lysine-Reactive Chemistry. *Nat Chem* **2021**, *13* (11), 1081–1092. https://doi.org/10.1038/s41557-021-00765-4.
(16) Mukherjee, H.; Debreczeni, J.; Breed, J.; Tentarelli, S.; Aquila, B.; Dowling, J. E.; Whitty, A.; Grimster, N. P. A Study of the Reactivity of S $^{(VI)}$ –F Containing Warheads with Nucleophilic Amino-Acid Side Chains under Physiological Conditions. *Org Biomol Chem* **2017**, *15* (45), 9685–9695. https://doi.org/10.1039/C7OB02028G.
(17) Baggio, C.; Udompholkul, P.; Gambini, L.; Salem, A. F.; Jossart, J.; Perry, J. J. P.; Pellecchia, M. Aryl-Fluorosulfate-Based Lysine Covalent Pan-Inhibitors of Apoptosis Protein (IAP) Antagonists with Cellular Efficacy. *J Med Chem* **2019**, *62* (20), 9188–9200. https://doi.org/10.1021/acs.jmedchem.9b01108.
(18) Udompholkul, P.; Garza-Granados, A.; Alboreggia, G.; Baggio, C.; McGuire, J.; Pegan, S. D.; Pellecchia, M. Characterization of a Potent and Orally Bioavailable Lys-Covalent Inhibitor of Apoptosis Protein (IAP) Antagonist. *J Med Chem* **2023**, *66* (12), 8159–8169. https://doi.org/10.1021/acs.jmedchem.3c00467.

(19) Wang, N.; Yang, B.; Fu, C.; Zhu, H.; Zheng, F.; Kobayashi, T.; Liu, J.; Li, S.; Ma, C.; Wang, P. G.; Wang, Q.; Wang, L. Genetically Encoding Fluorosulfate-<scp>l</Scp> -Tyrosine To React with Lysine, Histidine, and Tyrosine via SuFEx in Proteins *in Vivo*. *J Am Chem Soc* **2018**, *140* (15), 4995–4999. https://doi.org/10.1021/jacs.8b01087.
(20) Chen, P.; Tang, G.; Zhu, C.; Sun, J.; Wang, X.; Xiang, M.; Huang, H.; Wang, W.; Li, L.; Zhang, Z.-M.; Gao, L.; Yao, S. Q. 2-Ethynylbenzaldehyde-Based, Lysine-Targeting Irreversible Covalent Inhibitors for Protein Kinases and Nonkinases. *J Am Chem Soc* **2023**, *145* (7), 3844–3849. https://doi.org/10.1021/jacs.2c11595.
(21) Taylor, K. I.; Ho, J. S.; Trial, H. O.; Carter, A. W.; Kiessling, L. L. Assessing Squarates as Amine-Reactive Probes. *J Am Chem Soc* **2023**, *145* (46), 25056–25060. https://doi.org/10.1021/jacs.2c05691.
(22) Mason, M.; Belvisi, L.; Pignataro, L.; Dal Corso, A. A Tight Contact: The Expanding Application of Salicylaldehydes in Lysine-Targeting Covalent Drugs. *ChemBioChem* **2024**, *25* (7). https://doi.org/10.1002/cbic.202300743.
(23) Metcalf, B.; Chuang, C.; Dufu, K.; Patel, M. P.; Silva-Garcia, A.; Johnson, C.; Lu, Q.; Partridge, J. R.; Patskovska, L.; Patskovsky, Y.; Almo, S. C.; Jacobson, M. P.; Hua, L.; Xu, Q.; Gwaltney, S. L.; Yee, C.; Harris, J.; Morgan, B. P.; James, J.; Xu, D.; Hutchaleelaha, A.; Paulvannan, K.; Oksenberg, D.; Li, Z. Discovery of GBT440, an Orally Bioavailable R-State Stabilizer of Sickle Cell Hemoglobin. *ACS Med Chem Lett* **2017**, *8* (3), 321–326. https://doi.org/10.1021/acsmedchemlett.6b00491.
(24) Yang, T.; Cuesta, A.; Wan, X.; Craven, G. B.; Hirakawa, B.; Khamphavong, P.; May, J. R.; Kath, J. C.; Lapek, J. D.; Niessen, S.; Burlingame, A. L.; Carelli, J. D.; Taunton, J. Reversible Lysine-Targeted Probes Reveal Residence Time-Based Kinase Selectivity. *Nat Chem Biol* **2022**, *18* (9), 934–941. https://doi.org/10.1038/s41589-022-01019-1.
(25) Wang, W.; Wang, X.; Tang, G.; Zhu, C.; Xiang, M.; Xiao, Q.; Zhang, Z.-M.; Gao, L.; Yao, S. Q. Multitarget Inhibitors/Probes That Target LRRK2 and AURORA A Kinases Noncovalently and Covalently. *Chemical Communications* **2023**, *59* (72), 10789–10792. https://doi.org/10.1039/D3CC03530A.
(26) Bruyneel, W.; Charette, J. J.; De Hoffmann, E. Kinetics of Hydrolysis of Hydroxy and Methoxy Derivatives of N-Benzylidene-2-Aminopropane. *J Am Chem Soc* **1966**, *88* (16), 3808–3813. https://doi.org/10.1021/ja00968a024.
(27) Crugeiras, J.; Rios, A.; Riveiros, E.; Richard, J. P. Substituent Effects on the Thermodynamic Stability of Imines Formed from Glycine and Aromatic Aldehydes: Implications for the Catalytic Activity of Pyridoxal-5′-Phosphate. *J Am Chem Soc* **2009**, *131* (43), 15815–15824. https://doi.org/10.1021/ja906230n.
(28) Akçay, G.; Belmonte, M. A.; Aquila, B.; Chuaqui, C.; Hird, A. W.; Lamb, M. L.; Rawlins, P. B.; Su, N.; Tentarelli, S.; Grimster, N. P.; Su, Q. Inhibition of Mcl-1 through Covalent Modification of a Noncatalytic Lysine Side Chain. *Nat Chem Biol* **2016**, *12* (11), 931–936. https://doi.org/10.1038/nchembio.2174.
(29) He, H.; Zatorska, D.; Kim, J.; Aguirre, J.; Llauger, L.; She, Y.; Wu, N.; Immormino, R. M.; Gewirth, D. T.; Chiosis, G. Identification of Potent Water Soluble Purine-Scaffold Inhibitors of the Heat Shock Protein 90. *J Med Chem* **2006**, *49* (1), 381–390. https://doi.org/10.1021/jm0508078.

(30) Taldone, T.; Patel, P. D.; Patel, M.; Patel, H. J.; Evans, C. E.; Rodina, A.; Ochiana, S.; Shah, S. K.; Uddin, M.; Gewirth, D.; Chiosis, G. Experimental and Structural Testing Module to Analyze Paralogue-Specificity and Affinity in the Hsp90 Inhibitors Series. *J Med Chem* **2013**, *56* (17), 6803–6818. https://doi.org/10.1021/jm400619b.
(31) Immormino, R. M.; Kang, Y.; Chiosis, G.; Gewirth, D. T. Structural and Quantum Chemical Studies of 8-Aryl-Sulfanyl Adenine Class Hsp90 Inhibitors. *J Med Chem* **2006**, *49* (16), 4953–4960. https://doi.org/10.1021/jm060297x.
(32) Alasia, M.; Bertin, L.; Cerval, V.; Halley, F.; Mailliet, P.; Mendez Perez, M.; Minoux, H.; Ruxer, J.-M. Novel Hsp90 Inhibitory Carbazole Derivatives, Compositions Containing Same, and Use Thereof. WO2009122034A2, October 8, 2009.
(33) Fujita, N.; Sato, S.; Ishida, A.; Tsuruo, T. Involvement of Hsp90 in Signaling and Stability of 3-Phosphoinositide-Dependent Kinase-1. *Journal of Biological Chemistry* **2002**, *277* (12), 10346–10353. https://doi.org/10.1074/jbc.M106736200.
(34) Kabsch, W. *XDS*. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (2), 125–132. https://doi.org/10.1107/S0907444909047337.
(35) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *PHENIX* : A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (2), 213–221. https://doi.org/10.1107/S0907444909052925.
(36) Emsley, P.; Cowtan, K. *Coot* : Model-Building Tools for Molecular Graphics. *Acta Crystallogr D Biol Crystallogr* **2004**, *60* (12), 2126–2132. https://doi.org/10.1107/S0907444904019158.
(37) Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H. Ligand Efficiency and Fragment-Based Drug Discovery. *Drug Discovery Today*. 2009. https://doi.org/10.1016/j.drudis.2008.11.007.
(38) Hann, M. M.; Oprea, T. I. Pursuing the Leadlikeness Concept in Pharmaceutical Research. *Current Opinion in Chemical Biology*. 2004. https://doi.org/10.1016/j.cbpa.2004.04.003.
(39) Perera, T. P. S.; Jovcheva, E.; Mevellec, L.; Vialard, J.; De Lange, D.; Verhulst, T.; Paulussen, C.; Van De Ven, K.; King, P.; Freyne, E.; Rees, D. C.; Squires, M.; Saxty, G.; Page, M.; Murray, C. W.; Gilissen, R.; Ward, G.; Thompson, N. T.; Newell, D. R.; Cheng, N.; Xie, L.; Yang, J.; Platero, S. J.; Karkera, J. D.; Moy, C.; Angibaud, P.; Laquerre, S.; Lorenzi, M. V. Discovery & Pharmacological Characterization of JNJ-42756493 (Erdafitinib), a Functionally Selective Small-Molecule FGFR Family Inhibitor. *Mol Cancer Ther* **2017**, *16* (6). https://doi.org/10.1158/1535-7163.MCT-16-0589.
(40) 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* (2). https://doi.org/10.1038/nm.3048.

(41) Bollag, G.; Tsai, J.; Zhang, J.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P. Vemurafenib: The First Drug Approved for BRAF-Mutant Cancer. *Nature Reviews Drug Discovery*. 2012. https://doi.org/10.1038/nrd3847.

(42) Tap, W. D.; Wainberg, Z. A.; Anthony, S. P.; Ibrahim, P. N.; Zhang, C.; Healey, J. H.; Chmielowski, B.; Staddon, A. P.; Cohn, A. L.; Shapiro, G. I.; Keedy, V. L.; Singh, A. S.; Puzanov, I.; Kwak, E. L.; Wagner, A. J.; Von Hoff, D. D.; Weiss, G. J.; Ramanathan, R. K.; Zhang, J.; Habets, G.; Zhang, Y.; Burton, E. A.; Visor, G.; Sanftner, L.; Severson, P.; Nguyen, H.; Kim, M. J.; Marimuthu, A.; Tsang, G.; Shellooe, R.; Gee, C.; West, B. L.; Hirth, P.; Nolop, K.; van de Rijn, M.; Hsu, H. H.; Peterfy, C.; Lin, P. S.; Tong-Starksen, S.; Bollag, G. Structure-Guided Blockade of CSF1R Kinase in Tenosynovial Giant-Cell Tumor. *New England Journal of Medicine* **2015**, *373* (5). https://doi.org/10.1056/nejmoa1411366.

(43) Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M. K-Ras(G12C) Inhibitors Allosterically Control GTP Affinity and Effector Interactions. *Nature* **2013**, *503* (7477). https://doi.org/10.1038/nature12796.

(44) Schoepfer, J.; Jahnke, W.; Berellini, G.; Buonamici, S.; Cotesta, S.; Cowan-Jacob, S. W.; Dodd, S.; Drueckes, P.; Fabbro, D.; Gabriel, T.; Groell, J. M.; Grotzfeld, R. M.; Hassan, A. Q.; Henry, C.; Iyer, V.; Jones, D.; Lombardo, F.; Loo, A.; Manley, P. W.; Pellé, X.; Rummel, G.; Salem, B.; Warmuth, M.; Wylie, A. A.; Zoller, T.; Marzinzik, A. L.; Furet, P. Discovery of Asciminib (ABL001), an Allosteric Inhibitor of the Tyrosine Kinase Activity of BCR-ABL1. *J Med Chem* **2018**, *61* (18). https://doi.org/10.1021/acs.jmedchem.8b01040.

(45) Donald, A.; McHardy, T.; Rowlands, M. G.; Hunter, L. J. K.; Davies, T. G.; Berdini, V.; Boyle, R. G.; Aherne, G. W.; Garrett, M. D.; Collins, I. Rapid Evolution of 6-Phenylpurine Inhibitors of Protein Kinase B through Structure-Based Design. *J Med Chem* **2007**, *50* (10). https://doi.org/10.1021/jm0700924.

(46) Janes, M. R.; Zhang, J.; Li, L. S.; Hansen, R.; Peters, U.; Guo, X.; Chen, Y.; Babbar, A.; Firdaus, S. J.; Darjania, L.; Feng, J.; Chen, J. H.; Li, S.; Li, S.; Long, Y. O.; Thach, C.; Liu, Y.; Zarieh, A.; Ely, T.; Kucharski, J. M.; Kessler, L. V.; Wu, T.; Yu, K.; Wang, Y.; Yao, Y.; Deng, X.; Zarrinkar, P. P.; Brehmer, D.; Dhanak, D.; Lorenzi, M. V.; Hu-Lowe, D.; Patricelli, M. P.; Ren, P.; Liu, Y. Targeting KRAS Mutant Cancers with a Covalent G12C-Specific Inhibitor. *Cell* **2018**, *172* (3). https://doi.org/10.1016/j.cell.2018.01.006.

(47) Lanman, B. A.; Allen, J. R.; Allen, J. G.; Amegadzie, A. K.; Ashton, K. S.; Booker, S. K.; Chen, J. J.; Chen, N.; Frohn, M. J.; Goodman, G.; Kopecky, D. J.; Liu, L.; Lopez, P.; Low, J. D.; Ma, V.; Minatti, A. E.; Nguyen, T. T.; Nishimura, N.; Pickrell, A. J.; Reed, A. B.; Shin, Y.; Siegmund, A. C.; Tamayo, N. A.; Tegley, C. M.; Walton, M. C.; Wang, H. L.; Wurz, R. P.; Xue, M.; Yang, K. C.; Achanta, P.; Bartberger, M. D.; Canon, J.; Hollis, L. S.; McCarter, J. D.; Mohr, C.; Rex, K.; Saiki, A. Y.; San Miguel, T.; Volak, L. P.; Wang, K. H.; Whittington, D. A.; Zech, S. G.; Lipford, J. R.; Cee, V. J. Discovery of a Covalent Inhibitor of KRASG12C (AMG 510) for the Treatment of Solid Tumors. *J Med Chem* **2020**, *63* (1). https://doi.org/10.1021/acs.jmedchem.9b01180.

(48) Dan Erlanson. *Fragments in the clinic: 2024 edition*. https://practicalfragments.blogspot.com/2024/02/fragments-in-clinic-2024-edition.html (accessed 2024-11-18).
(49) Johansson, H.; Tsai, Y. C. I.; Fantom, K.; Chung, C. W.; Kümper, S.; Martino, L.; Thomas, D. A.; Eberl, H. C.; Muelbaier, M.; House, D.; Rittinger, K. Fragment-Based Covalent Ligand Screening Enables Rapid Discovery of Inhibitors for the RBR E3 Ubiquitin Ligase HOIP. *J Am Chem Soc* **2019**, *141* (6). https://doi.org/10.1021/jacs.8b13193.
(50) Dubiella, C.; Pinch, B. J.; Koikawa, K.; Zaidman, D.; Poon, E.; Manz, T. D.; Nabet, B.; He, S.; Resnick, E.; Rogel, A.; Langer, E. M.; Daniel, C. J.; Seo, H. S.; Chen, Y.; Adelmant, G.; Sharifzadeh, S.; Ficarro, S. B.; Jamin, Y.; Martins da Costa, B.; Zimmerman, M. W.; Lian, X.; Kibe, S.; Kozono, S.; Doctor, Z. M.; Browne, C. M.; Yang, A.; Stoler-Barak, L.; Shah, R. B.; Vangos, N. E.; Geffken, E. A.; Oren, R.; Koide, E.; Sidi, S.; Shulman, Z.; Wang, C.; Marto, J. A.; Dhe-Paganon, S.; Look, T.; Zhou, X. Z.; Lu, K. P.; Sears, R. C.; Chesler, L.; Gray, N. S.; London, N. Sulfopin Is a Covalent Inhibitor of Pin1 That Blocks Myc-Driven Tumors in Vivo. *Nat Chem Biol* **2021**, *17* (9). https://doi.org/10.1038/s41589-021-00786-7.
(51) Kathman, S. G.; Xu, Z.; Statsyuk, A. V. A Fragment-Based Method to Discover Irreversible Covalent Inhibitors of Cysteine Proteases. *J Med Chem* **2014**, *57* (11). https://doi.org/10.1021/jm500345q.
(52) Kathman, S. G.; Span, I.; Smith, A. T.; Xu, Z.; Zhan, J.; Rosenzweig, A. C.; Statsyuk, A. V. A Small Molecule That Switches a Ubiquitin Ligase from a Processive to a Distributive Enzymatic Mechanism. *J Am Chem Soc* **2015**, *137* (39). https://doi.org/10.1021/jacs.5b06839.
(53) Resnick, E.; Bradley, A.; Gan, J.; Douangamath, A.; Krojer, T.; Sethi, R.; Geurink, P. P.; Aimon, A.; Amitai, G.; Bellini, D.; Bennett, J.; Fairhead, M.; Fedorov, O.; Gabizon, R.; Gan, J.; Guo, J.; Plotnikov, A.; Reznik, N.; Ruda, G. F.; Díaz-Sáez, L.; Straub, V. M.; Szommer, T.; Velupillai, S.; Zaidman, D.; Zhang, Y.; Coker, A. R.; Dowson, C. G.; Barr, H. M.; Wang, C.; Huber, K. V. M.; Brennan, P. E.; Ovaa, H.; Von Delft, F.; London, N. Rapid Covalent-Probe Discovery by Electrophile-Fragment Screening. *J Am Chem Soc* **2019**, *141* (22). https://doi.org/10.1021/jacs.9b02822.
(54) Erlanson, D. A.; Arndt, J. W.; Cancilla, M. T.; Cao, K.; Elling, R. A.; English, N.; Friedman, J.; Hansen, S. K.; Hession, C.; Joseph, I.; Kumaravel, G.; Lee, W. C.; Lind, K. E.; McDowell, R. S.; Miatkowski, K.; Nguyen, C.; Nguyen, T. B.; Park, S.; Pathan, N.; Penny, D. M.; Romanowski, M. J.; Scott, D.; Silvian, L.; Simmons, R. L.; Tangonan, B. T.; Yang, W.; Sun, L. Discovery of a Potent and Highly Selective PDK1 Inhibitor via Fragment-Based Drug Discovery. *Bioorg Med Chem Lett* **2011**, *21* (10). https://doi.org/10.1016/j.bmcl.2011.03.032.
(55) Kenanova, D. N.; Visser, E. J.; Virta, J. M.; Sijbesma, E.; Centorrino, F.; Vickery, H. R.; Zhong, M.; Neitz, R. J.; Brunsveld, L.; Ottmann, C.; Arkin, M. R. A Systematic Approach to the Discovery of Protein-Protein Interaction Stabilizers. *ACS Cent Sci* **2023**, *9* (5). https://doi.org/10.1021/acscentsci.2c01449.
(56) Hardy, J. A.; Lam, J.; Nguyen, J. T.; O'Brien, T.; Wells, J. A. Discovery of an Allosteric Site in the Caspases. *Proc Natl Acad Sci U S A* **2004**, *101* (34). https://doi.org/10.1073/pnas.0404781101.

(57) Buck, E.; Wells, J. A. Disulfide Trapping to Localize Small-Molecule Agonists and Antagonists for a G Protein-Coupled Receptor. *Proc Natl Acad Sci U S A* **2005**, *102* (8). https://doi.org/10.1073/pnas.0500016102.
(58) Yang, W.; Fucini, R. V.; Fahr, B. T.; Randal, M.; Lind, K. E.; Lam, M. B.; Lu, W.; Lu, Y.; Cary, D. R.; Romanowski, M. J.; Colussi, D.; Pietrak, B.; Allison, T. J.; Munshi, S. K.; Penny, D. M.; Pham, P.; Sun, J.; Thomas, A. E.; Wilkinson, J. M.; Jacobs, J. W.; McDowell, R. S.; Ballinger, M. D. Fragment-Based Discovery of Nonpeptidic BACE-1 Inhibitors Using Tethering. *Biochemistry* **2009**, *48* (21). https://doi.org/10.1021/bi900017q.
(59) Erlanson, D. A.; Wells, J. A.; Braisted, A. C. Tethering: Fragment-Based Drug Discovery. *Annual Review of Biophysics and Biomolecular Structure*. 2004. https://doi.org/10.1146/annurev.biophys.33.110502.140409.
(60) Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. Site-Directed Ligand Discovery. *Proc Natl Acad Sci U S A* **2000**, *97* (17). https://doi.org/10.1073/pnas.97.17.9367.
(61) Hwang, C.; Sinskey, A. J.; Lodish, H. F. Oxidized Redox State of Glutathione in the Endoplasmic Reticulum. *Science (1979)* **1992**, *257* (5076). https://doi.org/10.1126/science.1523409.
(62) Lee, M. H.; Yang, Z.; Lim, C. W.; Lee, Y. H.; Dongbang, S.; Kang, C.; Kim, J. S. Disulfide-Cleavage-Triggered Chemosensors and Their Biological Applications. *Chemical Reviews*. 2013. https://doi.org/10.1021/cr300358b.
(63) Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.; Horning, B. D.; González-Páez, G. E.; Chatterjee, S.; Lanning, B. R.; Teijaro, J. R.; Olson, A. J.; Wolan, D. W.; Cravatt, B. F. Proteome-Wide Covalent Ligand Discovery in Native Biological Systems. *Nature* **2016**, *534* (7608). https://doi.org/10.1038/nature18002.
(64) Hacker, S. M.; Backus, K. M.; Lazear, M. R.; Forli, S.; Correia, B. E.; Cravatt, B. F. Global Profiling of Lysine Reactivity and Ligandability in the Human Proteome. *Nat Chem* **2017**, *9* (12). https://doi.org/10.1038/nchem.2826.
(65) Kuljanin, M.; Mitchell, D. C.; Schweppe, D. K.; Gikandi, A. S.; Nusinow, D. P.; Bulloch, N. J.; Vinogradova, E. V.; Wilson, D. L.; Kool, E. T.; Mancias, J. D.; Cravatt, B. F.; Gygi, S. P. Reimagining High-Throughput Profiling of Reactive Cysteines for Cell-Based Screening of Large Electrophile Libraries. *Nat Biotechnol* **2021**, *39* (5). https://doi.org/10.1038/s41587-020-00778-3.
(66) Brulet, J. W.; Borne, A. L.; Yuan, K.; Libby, A. H.; Hsu, K. L. Liganding Functional Tyrosine Sites on Proteins Using Sulfur-Triazole Exchange Chemistry. *J Am Chem Soc* **2020**, *142* (18). https://doi.org/10.1021/jacs.0c00648.
(67) Gonzalez-Valero, A.; Reeves, A. G.; Page, A. C. S.; Moon, P. J.; Miller, E.; Coulonval, K.; Crossley, S. W. M.; Xie, X.; He, D.; Musacchio, P. Z.; Christian, A. H.; McKenna, J. M.; Lewis, R. A.; Fang, E.; Dovala, D.; Lu, Y.; McGregor, L. M.; Schirle, M.; Tallarico, J. A.; Roger, P. P.; Toste, F. D.; Chang, C. J. An Activity-Based Oxaziridine Platform for Identifying and Developing Covalent Ligands for Functional Allosteric Methionine Sites: Redox-Dependent Inhibition of Cyclin-Dependent Kinase 4. *J Am Chem Soc* **2022**, *144* (50). https://doi.org/10.1021/jacs.2c04039.

(68) Fu, L.; Jung, Y.; Tian, C.; Ferreira, R. B.; Cheng, R.; He, F.; Yang, J.; Carroll, K. S. Nucleophilic Covalent Ligand Discovery for the Cysteine Redoxome. *Nat Chem Biol* **2023**, *19* (11). https://doi.org/10.1038/s41589-023-01330-5.
(69) Ward, C. C.; Kleinman, J. I.; Nomura, D. K. NHS-Esters As Versatile Reactivity-Based Probes for Mapping Proteome-Wide Ligandable Hotspots. *ACS Chem Biol* **2017**, *12* (6). https://doi.org/10.1021/acschembio.7b00125.
(70) Cossar, P. J.; Wolter, M.; Van Dijck, L.; Valenti, D.; Levy, L. M.; Ottmann, C.; Brunsveld, L. Reversible Covalent Imine-Tethering for Selective Stabilization of 14-3-3 Hub Protein Interactions. *J Am Chem Soc* **2021**, *143* (22). https://doi.org/10.1021/jacs.1c03035.
(71) Wolter, M.; Valenti, D.; Cossar, P. J.; Levy, L. M.; Hristeva, S.; Genski, T.; Hoffmann, T.; Brunsveld, L.; Tzalis, D.; Ottmann, C. Fragment-Based Stabilizers of Protein–Protein Interactions through Imine-Based Tethering. *Angewandte Chemie - International Edition* **2020**, *59* (48). https://doi.org/10.1002/anie.202008585.
(72) Vázquez, M. A.; Echevarría, G.; Muñoz, F.; Donoso, J.; Blanco, F. G. Kinetic Study of the Schiff-Base Formation between Glycine and Pyridoxal 5′-Phosphate (PLP), Pyridoxal (PL), and 5′-Deoxypyridoxal (DPL). *Journal of the Chemical Society, Perkin Transactions 2* **1989**, No. 11. https://doi.org/10.1039/p29890001617.
(73) Vazquez, M. A.; Muñoz, F.; Donoso, J.; Blanco, F. G. Influence of the Side Chain on the Stability of Schiff-bases Formed between Pyridoxal 5′-phosphate and Amino Acids. *Int J Chem Kinet* **1990**, *22* (9). https://doi.org/10.1002/kin.550220903.
(74) Krukenberg, K. A.; Street, T. O.; Lavery, L. A.; Agard, D. A. Conformational Dynamics of the Molecular Chaperone Hsp90. *Q Rev Biophys* **2011**, *44* (2). https://doi.org/10.1017/S0033583510000314.
(75) Woodhead, A. J.; Angove, H.; Carr, M. G.; Chessari, G.; Congreve, M.; Coyle, J. E.; Cosme, J.; Graham, B.; Day, P. J.; Downham, R.; Fazal, L.; Feltell, R.; Figueroa, E.; Frederickson, M.; Lewis, J.; McMenamin, R.; Murray, C. W.; O'Brien, M. A.; Parra, L.; Patel, S.; Phillips, T.; Rees, D. C.; Rich, S.; Smith, D. M.; Trewartha, G.; Vinkovic, M.; Williams, B.; Woolford, A. J. A. Discovery of (2,4-Dihydroxy-5-Isopropylphenyl)-[5-(4-Methylpiperazin-1- Ylmethyl)-1,3-Dihydroisoindol-2-Yl]Methanone (AT13387), a Novel Inhibitor of the Molecular Chaperone Hsp90 by Fragment Based Drug Design. *J Med Chem* **2010**, *53* (16). https://doi.org/10.1021/jm100060b.
(76) Bum-Erdene, K.; Liu, D.; Gonzalez-Gutierrez, G.; Ghozayel, M. K.; Xu, D.; Meroueh, S. O. Small-Molecule Covalent Bond Formation at Tyrosine Creates a Binding Site and Inhibits Activation of Ral GTPases. *Proc Natl Acad Sci U S A* **2020**, *117* (13). https://doi.org/10.1073/pnas.1913654117.
(77) Ippolito, J. A.; Niu, H.; Bertoletti, N.; Carter, Z. J.; Jin, S.; Spasov, K. A.; Cisneros, J. A.; Valhondo, M.; Cutrona, K. J.; Anderson, K. S.; Jorgensen, W. L. Covalent Inhibition of Wild-Type HIV-1 Reverse Transcriptase Using a Fluorosulfate Warhead. *ACS Med Chem Lett* **2021**, *12* (2). https://doi.org/10.1021/acsmedchemlett.0c00612.

(78) Bolding, J. E.; Martín-Gago, P.; Rajabi, N.; Gamon, L. F.; Hansen, T. N.; Bartling, C. R. O.; Strømgaard, K.; Davies, M. J.; Olsen, C. A. Aryl Fluorosulfate Based Inhibitors That Covalently Target the SIRT5 Lysine Deacylase**. *Angewandte Chemie - International Edition* **2022**, *61* (47). https://doi.org/10.1002/anie.202204565.
(79) Chen, Y.; Craven, G. B.; Kamber, R. A.; Cuesta, A.; Zhersh, S.; Moroz, Y. S.; Bassik, M. C.; Taunton, J. Direct Mapping of Ligandable Tyrosines and Lysines in Cells with Chiral Sulfonyl Fluoride Probes. *Nat Chem* **2023**, *15* (11). https://doi.org/10.1038/s41557-023-01281-3.
(80) Ueda, T.; Tamura, T.; Kawano, M.; Shiono, K.; Hobor, F.; Wilson, A. J.; Hamachi, I. Enhanced Suppression of a Protein-Protein Interaction in Cells Using Small-Molecule Covalent Inhibitors Based on an N-Acyl- N-Alkyl Sulfonamide Warhead. *J Am Chem Soc* **2021**, *143* (12). https://doi.org/10.1021/jacs.1c00703.
(81) Shishido, Y.; Tomoike, F.; Kuwata, K.; Fujikawa, H.; Sekido, Y.; Murakami-Tonami, Y.; Kameda, T.; Abe, N.; Kimura, Y.; Shuto, S.; Abe, H. A Covalent Inhibitor for Glutathione S-Transferase Pi (GSTP 1-1) in Human Cells. *ChemBioChem* **2019**, *20* (7). https://doi.org/10.1002/cbic.201800671.
(82) Teng, M.; Teng, M.; Ficarro, S. B.; Ficarro, S. B.; Ficarro, S. B.; Yoon, H.; Che, J.; Che, J.; Zhou, J.; Fischer, E. S.; Fischer, E. S.; Marto, J. A.; Marto, J. A.; Marto, J. A.; Zhang, T.; Zhang, T.; Gray, N. S.; Gray, N. S. Rationally Designed Covalent BCL6 Inhibitor That Targets a Tyrosine Residue in the Homodimer Interface. *ACS Med Chem Lett* **2020**, *11* (6). https://doi.org/10.1021/acsmedchemlett.0c00111.
(83) Yang, X.; Van Veldhoven, J. P. D.; Offringa, J.; Kuiper, B. J.; Lenselink, E. B.; Heitman, L. H.; Van Der Es, D.; Ijzerman, A. P. Development of Covalent Ligands for G Protein-Coupled Receptors: A Case for the Human Adenosine A3 Receptor. *J Med Chem* **2019**, *62* (7). https://doi.org/10.1021/acs.jmedchem.8b02026.
(84) Soga, S.; Shirai, H.; Koborv, M.; Hirayama, N. Use of Amino Acid Composition to Predict Ligand-Binding Sites. *J Chem Inf Model* **2007**, *47* (2). https://doi.org/10.1021/ci6002202.
(85) Cuesta, A. Development of Lysine-Reactive Covalent Inhibitors and Chemoproteomic Probes. Ph.D. Dissertation, University of California, San Francisco, San Francisco, 2019.
(86) Pahari, S.; Sun, L.; Alexov, E. PKAD: A Database of Experimentally Measured PKa Values of Ionizable Groups in Proteins. *Database* **2019**, *2019* (1). https://doi.org/10.1093/database/baz024.
(87) Morgen, M.; Jöst, C.; Malz, M.; Janowski, R.; Niessing, D.; Klein, C. D.; Gunkel, N.; Miller, A. K. Spiroepoxytriazoles Are Fumagillin-like Irreversible Inhibitors of MetAP2 with Potent Cellular Activity. *ACS Chem Biol* **2016**, *11* (4). https://doi.org/10.1021/acschembio.5b00755.
(88) Lowther, W. T.; Mcmillen, D. A.; Orville, A. M.; Matthews, B. W. The Anti-Angiogenic Agent Fumagillin Covalently Modifies a Conserved Active-Site Histidine in the Escherichia Coli Methionine Aminopeptidase. *Proc Natl Acad Sci U S A* **1998**, *95* (21). https://doi.org/10.1073/pnas.95.21.12153.
(89) Yoshizawa, M.; Itoh, T.; Hori, T.; Kato, A.; Anami, Y.; Yoshimoto, N.; Yamamoto, K. Identification of the Histidine Residue in Vitamin D Receptor That Covalently

Binds to Electrophilic Ligands. *J Med Chem* **2018**, *61* (14). https://doi.org/10.1021/acs.jmedchem.8b00774.
(90) Jakob, C. G.; Upadhyay, A. K.; Donner, P. L.; Nicholl, E.; Addo, S. N.; Qiu, W.; Ling, C.; Gopalakrishnan, S. M.; Torrent, M.; Cepa, S. P.; Shanley, J.; Shoemaker, A. R.; Sun, C. C.; Vasudevan, A.; Woller, K. R.; Brad Shotwell, J.; Shaw, B.; Bian, Z.; Hutti, J. E. Novel Modes of Inhibition of Wild-Type Isocitrate Dehydrogenase 1 (IDH1): Direct Covalent Modification of His315. *J Med Chem* **2018**, *61* (15). https://doi.org/10.1021/acs.jmedchem.8b00305.
(91) Tsilikounas, E.; Kettner, C.; Bachovchin, W. W. Identification of Serine and Histidine Adducts in Complexes of Trypsin and Trypsinogen with Peptide and Nonpeptide Boronic Acid Inhibitors by Proton NMR Spectroscopy. *Biochemistry* **1992**, *31* (51). https://doi.org/10.1021/bi00166a019.
(92) Roberts, M. F.; Deems, R. A.; Mincey, T. C.; Dennis, E. A. Chemical Modification of the Histidine Residue in Phospholipase A2 (Naja Naja Naja). A Case of Half Site Reactivity. *Journal of Biological Chemistry* **1977**, *252* (7). https://doi.org/10.1016/s0021-9258(17)40568-0.
(93) Li, T.; Maltais, R.; Poirier, D.; Lin, S. X. Combined Biophysical Chemistry Reveals a New Covalent Inhibitor with a Low-Reactivity Alkyl Halide. *Journal of Physical Chemistry Letters* **2018**, *9* (18). https://doi.org/10.1021/acs.jpclett.8b02225.
(94) Jia, S.; He, D.; Chang, C. J. Bioinspired Thiophosphorodichloridate Reagents for Chemoselective Histidine Bioconjugation. *J Am Chem Soc* **2019**, *141* (18). https://doi.org/10.1021/jacs.8b11912.
(95) Cruite, J. T.; Dann, G. P.; Che, J.; Donovan, K. A.; Ferrao, S.; Ficarro, S. B.; Fischer, E. S.; Gray, N. S.; Huerta, F.; Kong, N. R.; Liu, H.; Marto, J. A.; Metivier, R. J.; Nowak, R. P.; Zerfas, B. L.; Jones, L. H. Cereblon Covalent Modulation through Structure-Based Design of Histidine Targeting Chemical Probes. *RSC Chem Biol* **2022**, *3* (9), 1105–1110. https://doi.org/10.1039/D2CB00078D.
(96) Alboreggia, G.; Udompholkul, P.; Baggio, C.; Muzzarelli, K.; Assar, Z.; Pellecchia, M. Histidine-Covalent Stapled Alpha-Helical Peptides Targeting HMcl-1. *J Med Chem* **2024**, *67* (10), 8172–8185. https://doi.org/10.1021/acs.jmedchem.4c00277.
(97) Jiao, G. S.; Kim, S.; Moayeri, M.; Thai, A.; Cregar-Hernandez, L.; McKasson, L.; O'Malley, S.; Leppla, S. H.; Johnson, A. T. Small Molecule Inhibitors of Anthrax Edema Factor. *Bioorg Med Chem Lett* **2018**, *28* (2). https://doi.org/10.1016/j.bmcl.2017.11.040.
(98) Gambini, L.; Baggio, C.; Udompholkul, P.; Jossart, J.; Salem, A. F.; Perry, J. J. P.; Pellecchia, M. Covalent Inhibitors of Protein-Protein Interactions Targeting Lysine, Tyrosine, or Histidine Residues. *J Med Chem* **2019**, *62* (11). https://doi.org/10.1021/acs.jmedchem.9b00561.
(99) Li, S.; Wang, N.; Yu, B.; Sun, W.; Wang, L. Genetically Encoded Chemical Crosslinking of Carbohydrate. *Nat Chem* **2023**, *15* (1). https://doi.org/10.1038/s41557-022-01059-z.
(100) Sun, W.; Wang, N.; Liu, H.; Yu, B.; Jin, L.; Ren, X.; Shen, Y.; Wang, L. Genetically Encoded Chemical Crosslinking of RNA in Vivo. *Nat Chem* **2023**, *15* (1). https://doi.org/10.1038/s41557-022-01038-4.
(101) Nowak, R. P.; Ragosta, L.; Huerta, F.; Liu, H.; Ficarro, S. B.; Cruite, J. T.; Metivier, R. J.; Donovan, K. A.; Marto, J. A.; Fischer, E. S.; Zerfas, B. L.; Jones, L. H.

Development of a Covalent Cereblon-Based PROTAC Employing a Fluorosulfate Warhead. *RSC Chem Biol* **2023**, *4* (11), 906–912. https://doi.org/10.1039/D3CB00103B.
(102) Fadeyi, O. O.; Hoth, L. R.; Choi, C.; Feng, X.; Gopalsamy, A.; Hett, E. C.; Kyne, R. E.; Robinson, R. P.; Jones, L. H. Covalent Enzyme Inhibition through Fluorosulfate Modification of a Noncatalytic Serine Residue. *ACS Chem Biol* **2017**, *12* (8). https://doi.org/10.1021/acschembio.7b00403.
(103) Alboreggia, G.; Udompholkul, P.; Atienza, E. L.; Muzzarelli, K.; Assar, Z.; Pellecchia, M. Covalent Targeting of Histidine Residues with Aryl Fluorosulfates: Application to Mcl-1 BH3 Mimetics. *J Med Chem* **2024**. https://doi.org/10.1021/acs.jmedchem.4c01541.
(104) Yu, B.; Cao, L.; Li, S.; Klauser, P. C.; Wang, L. The Proximity-Enabled Sulfur Fluoride Exchange Reaction in the Protein Context. *Chem Sci* **2023**, *14* (29). https://doi.org/10.1039/d3sc01921g.
(105) Cruite, J. T.; Nowak, R. P.; Donovan, K. A.; Ficarro, S. B.; Huang, H.; Liu, H.; Liu, Y.; Marto, J. A.; Metivier, R. J.; Fischer, E. S.; Jones, L. H. Covalent Stapling of the Cereblon Sensor Loop Histidine Using Sulfur-Heterocycle Exchange. *ACS Med Chem Lett* **2023**, *14* (11). https://doi.org/10.1021/acsmedchemlett.3c00371.
(106) Lanning, B. R.; Whitby, L. R.; Dix, M. M.; Douhan, J.; Gilbert, A. M.; Hett, E. C.; Johnson, T. O.; Joslyn, C.; Kath, J. C.; Niessen, S.; Roberts, L. R.; Schnute, M. E.; Wang, C.; Hulce, J. J.; Wei, B.; Whiteley, L. O.; Hayward, M. M.; Cravatt, B. F. A Road Map to Evaluate the Proteome-Wide Selectivity of Covalent Kinase Inhibitors. *Nat Chem Biol* **2014**, *10* (9). https://doi.org/10.1038/nchembio.1582.
(107) Bhagavat, R.; Sankar, S.; Srinivasan, N.; Chandra, N. An Augmented Pocketome: Detection and Analysis of Small-Molecule Binding Pockets in Proteins of Known 3D Structure. *Structure* **2018**, *26* (3). https://doi.org/10.1016/j.str.2018.02.001.
(108) Hernandez, M.; Ghersi, D.; Sanchez, R. SITEHOUND-Web: A Server for Ligand Binding Site Identification in Protein Structures. *Nucleic Acids Res* **2009**, *37* (SUPPL. 2). https://doi.org/10.1093/nar/gkp281.
(109) Kalidas, Y.; Chandra, N. PocketDepth: A New Depth Based Algorithm for Identification of Ligand Binding Sites in Proteins. *J Struct Biol* **2008**, *161* (1). https://doi.org/10.1016/j.jsb.2007.09.005.
(110) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinformatics* **2009**, *10*. https://doi.org/10.1186/1471-2105-10-168.
(111) Oliveros, J. C. *Venny. An Interactive Tool for Comparing Lists with Venn's Diagrams. - References - Scientific Research Publishing*. Csic.
(112) Mendelson, K.; Evans, T.; Hla, T. Sphingosine 1-Phosphate Signalling. *Development (Cambridge)* **2014**, *141* (1). https://doi.org/10.1242/dev.094805.
(113) Zheng, X.; Li, W.; Ren, L.; Liu, J.; Pang, X.; Chen, X.; Kang, D.; Wang, J.; Du, G. The Sphingosine Kinase-1/Sphingosine-1-Phosphate Axis in Cancer: Potential Target for Anticancer Therapy. *Pharmacology and Therapeutics*. 2019. https://doi.org/10.1016/j.pharmthera.2018.10.011.
(114) Pyne, N. J.; Pyne, S. Sphingosine 1-Phosphate and Cancer. *Nature Reviews Cancer*. 2010. https://doi.org/10.1038/nrc2875.

(115) Pyne, S.; Adams, D. R.; Pyne, N. J. Sphingosine 1-Phosphate and Sphingosine Kinases in Health and Disease: Recent Advances. *Progress in Lipid Research*. 2016. https://doi.org/10.1016/j.plipres.2016.03.001.
(116) Pyne, N. J.; Pyne, S. Recent Advances in the Role of Sphingosine 1-Phosphate in Cancer. *FEBS Letters*. 2020. https://doi.org/10.1002/1873-3468.13933.
(117) Pchejetski, D.; Doumerc, N.; Golzio, M.; Naymark, M.; Teissié, J.; Kohama, T.; Waxman, J.; Malavaud, B.; Cuvillier, O. Chemosensitizing Effects of Sphingosine Kinase-1 Inhibition in Prostate Cancer Cell and Animal Models. *Mol Cancer Ther* **2008**, *7* (7). https://doi.org/10.1158/1535-7163.MCT-07-2322.
(118) Sukocheva, O.; Wang, L.; Verrier, E.; Vadas, M. A.; Xia, P. Restoring Endocrine Response in Breast Cancer Cells by Inhibition of the Sphingosine Kinase-1 Signaling Pathway. *Endocrinology* **2009**, *150* (10). https://doi.org/10.1210/en.2009-0391.
(119) Nemoto, S.; Nakamura, M.; Osawa, Y.; Kono, S.; Itoh, Y.; Okano, Y.; Murate, T.; Hara, A.; Ueda, H.; Nozawa, Y.; Banno, Y. Sphingosine Kinase Isoforms Regulate Oxaliplatin Sensitivity of Human Colon Cancer Cells through Ceramide Accumulation and Akt Activation. *Journal of Biological Chemistry* **2009**, *284* (16). https://doi.org/10.1074/jbc.M900735200.
(120) Rosa, R.; Marciano, R.; Malapelle, U.; Formisano, L.; Nappi, L.; D'Amato, C.; D'Amato, V.; Damiano, V.; Marfè, G.; Del Vecchio, S.; Zannetti, A.; Greco, A.; De Stefano, A.; Carlomagno, C.; Veneziani, B. M.; Troncone, G.; De Placido, S.; Bianco, R. Sphingosine Kinase 1 Overexpression Contributes to Cetuximab Resistance in Human Colorectal Cancer Models. *Clinical Cancer Research* **2013**, *19* (1). https://doi.org/10.1158/1078-0432.CCR-12-1050.
(121) Guillermet-Guibert, J.; Davenne, L.; Pchejetski, D.; Saint-Laurent, N.; Brizuela, L.; Guilbeau-Frugier, C.; Delisle, M. B.; Cuvillier, O.; Susini, C.; Bousquet, C. Targeting the Sphingolipid Metabolism to Defeat Pancreatic Cancer Cell Resistance to the Chemotherapeutic Gemcitabine Drug. *Mol Cancer Ther* **2009**, *8* (4). https://doi.org/10.1158/1535-7163.MCT-08-1096.
(122) Schnute, M. E.; McReynolds, M. D.; Kasten, T.; Yates, M.; Jerome, G.; Rains, J. W.; Hall, T.; Chrencik, J.; Kraus, M.; Cronin, C. N.; Saabye, M.; Highkin, M. K.; Broadus, R.; Ogawa, S.; Cukyne, K.; Zawadzke, L. E.; Peterkin, V.; Iyanar, K.; Scholten, J. A.; Wendling, J.; Fujiwara, H.; Nemirovskiy, O.; Wittwer, A. J.; Nagiec, M. M. Modulation of Cellular S1P Levels with a Novel, Potent and Specific Inhibitor of Sphingosine Kinase-1. *Biochemical Journal* **2012**, *444* (1). https://doi.org/10.1042/BJ20111929.
(123) Yan, H.; Parsons, D. W.; Jin, G.; McLendon, R.; Rasheed, B. A.; Yuan, W.; Kos, I.; Batinic-Haberle, I.; Jones, S.; Riggins, G. J.; Friedman, H.; Friedman, A.; Reardon, D.; Herndon, J.; Kinzler, K. W.; Velculescu, V. E.; Vogelstein, B.; Bigner, D. D. IDH1 and IDH2 Mutations in Gliomas . *New England Journal of Medicine* **2009**, *360* (8). https://doi.org/10.1056/nejmoa0808710.
(124) Balss, J.; Meyer, J.; Mueller, W.; Korshunov, A.; Hartmann, C.; von Deimling, A. Analysis of the IDH1 Codon 132 Mutation in Brain Tumors. *Acta Neuropathol* **2008**, *116* (6). https://doi.org/10.1007/s00401-008-0455-2.
(125) Dang, L.; White, D. W.; Gross, S.; Bennett, B. D.; Bittinger, M. A.; Driggers, E. M.; Fantin, V. R.; Jang, H. G.; Jin, S.; Keenan, M. C.; Marks, K. M.; Prins, R. M.;

Ward, P. S.; Yen, K. E.; Liau, L. M.; Rabinowitz, J. D.; Cantley, L. C.; Thompson, C. B.; Vander Heiden, M. G.; Su, S. M. Cancer-Associated IDH1 Mutations Produce 2-Hydroxyglutarate. *Nature* **2009**, *462* (7274). https://doi.org/10.1038/nature08617.
(126) Lamb, Y. N. Vorasidenib: First Approval. *Drugs* **2024**, *84* (10), 1325–1331. https://doi.org/10.1007/s40265-024-02097-2.
(127) Dhillon, S. Ivosidenib: First Global Approval. *Drugs* **2018**, *78* (14). https://doi.org/10.1007/s40265-018-0978-3.
(128) Kang, C. Olutasidenib: First Approval. *Drugs* **2023**, *83* (4). https://doi.org/10.1007/s40265-023-01844-1.
(129) Salama, V.; Brooks, N.; Skwarska, A.; Kays, L.; Milligan, P.; Newell, K.; Roth, K.; Geeganage, S.; Gilmour, R.; Chan, S. M.; Sarry, J.-E.; Sabatier, M.; DiNardo, C.; Konopleva, M. Abstract 6417: LY3410738, a Novel Inhibitor of Mutant IDH1 Is More Effective than Ivosidenib and Potentiates Antileukemic Activity of Standard Chemotherapy in Preclinical Models of Acute Myeloid Leukemia (AML). *Cancer Res* **2020**, *80* (16_Supplement). https://doi.org/10.1158/1538-7445.am2020-6417.
(130) Pauff, J. M.; Papadopoulos, K. P.; Janku, F.; Turk, A. A.; Goyal, L.; Shroff, R. T.; Shimizu, T.; Ikeda, M.; Azad, N. S.; Cleary, J. M.; Peters, M. L. B.; Borad, M. J.; Jaeckle, K. A.; Kizilbash, S. H.; Tupper, R.; Furin, C. E.; Hanley, M. P.; Hill, E. G.; Xu, X.; Harding, J. J. A Phase I Study of LY3410738, a First-in-Class Covalent Inhibitor of Mutant IDH1 in Cholangiocarcinoma and Other Advanced Solid Tumors. *Journal of Clinical Oncology* **2021**, *39* (3_suppl). https://doi.org/10.1200/jco.2021.39.3_suppl.tps350.
(131) Plieth, J. *Lilly calls time on IDH1/2 inhibition*.
(132) Liang, Q.; Wang, B.; Zou, F.; Guo, G.; Wang, W.; Wang, W.; Liu, Q.; Shen, L.; Hu, C.; Wang, W.; Wang, A.; Huang, T.; He, Y.; Xia, R.; Ge, J.; Liu, J.; Liu, Q. Structure-Based Discovery of IHMT-IDH1-053 as a Potent Irreversible IDH1 Mutant Selective Inhibitor. *Eur J Med Chem* **2023**, *256*. https://doi.org/10.1016/j.ejmech.2023.115411.
(133) Pusch, S.; Krausert, S.; Fischer, V.; Balss, J.; Ott, M.; Schrimpf, D.; Capper, D.; Sahm, F.; Eisel, J.; Beck, A. C.; Jugold, M.; Eichwald, V.; Kaulfuss, S.; Panknin, O.; Rehwinkel, H.; Zimmermann, K.; Hillig, R. C.; Guenther, J.; Toschi, L.; Neuhaus, R.; Haegebart, A.; Hess-Stumpp, H.; Bauser, M.; Wick, W.; Unterberg, A.; Herold-Mende, C.; Platten, M.; von Deimling, A. Pan-Mutant IDH1 Inhibitor BAY 1436032 for Effective Treatment of IDH1 Mutant Astrocytoma in Vivo. *Acta Neuropathol* **2017**, *133* (4). https://doi.org/10.1007/s00401-017-1677-y.

Appendix A: PDB search script

```
import os, string, sys, math, subprocess, csv
import cPickle as pic
from html import HTML
import time

print("Starting Script\n")
#print(strftime("%Y-%m-%d %H:%M:%S"))

"""
Script requires 3 arguments
1. The location of the PDB files
2. A list of names of the PDB files as a txt file
3. Path/filename for the output
"""

pdbpath = sys.argv[1]
pdblist= sys.argv[2]
outputname = sys.argv[3]

"""
Accumulated list of "bad" ligands to be excluded
"""
badligands = ['NBN', 'ENC', 'MNC', 'NPN', 'GSH', '0HH', 'GTX', 'BEN', 'OAR', 'PYX', 'CFC', 'ETF', 'TDG', 'ADC', 'RIP', 'CFO', 'OFO',
'9DI', 'HPA', 'NOS', 'MET', 'LEU', 'DME', 'TRA', '1DA', '  O', 'MBG', 'SCR', 'AYA', 'ETA', 'HAE', 'MIC', 'DAS', 'BE2', 'DCE', 'ASP', 'HEA',
'LDA', 'SNN', 'IAS', 'HIC', 'BHD', 'EDR', 'SGN', 'NTP', 'PYJ', 'SAC', 'TSU', 'CYH', 'AOP', 'CRA', 'MBR', 'CTO', 'UMG', 'GUM', 'BLV',
'FMS', 'A2G', 'DAR', 'PMB', 'MMC', '2FU', 'PBP', 'INN', 'BRC', 'SSB', 'SNC', 'HGB', 'AJ3', 'ADN', 'BAM', 'NVA', 'H2S', 'ADE', 'ACA',
'MTA', 'DEC', 'NON', 'COH', 'HMC', 'TRI', 'PHG', 'CAD', 'PG6', 'SAH', 'NH3', 'DDQ', 'ABH', 'SO2', 'SET', 'NLE', 'DSS', 'BUQ', 'CYS',
'AR7', '0QE', 'PYZ', 'ITU', ' LA', 'V7O', 'PHQ', 'HRG', 'CMT', 'GER', 'CSS', 'LPB', 'BOM', 'ASG', 'GLU', '4MZ', '1MZ', 'HAR', 'DHM',
'UNA', 'GUP', 'DDZ', 'BO3', 'AAE', 'DTD', 'PBR', 'HLT', 'ALA', 'PRO', 'GCS', 'IPU', 'IUM', 'CDU', 'HEV', 'CRS', 'RCO', 'IOL', 'HDS',
'ASJ', 'CNN', 'MGS', 'MGU', 'DCY', 'ESA', ' CE', 'GCU', 'HXA', 'LAX', 'AZE', 'AIB', 'SE4', 'ATH', 'MOS', 'FRU', 'FGA', 'DAM', 'PUR',
'HNI', 'ARG', ' TL', 'GL2', 'GL5', 'GL7', 'GL9', 'CR1', 'CR6', 'SGA', 'MAG', 'ACN', '3OL', 'AEM', 'EMM', '3OM', 'IUR', 'IDK', 'URA', 'GLY',
'NTA', 'CDG', 'GM2', 'D12', 'TWT', 'AOA', 'OGA', 'BEZ', 'G2F', 'DO3', 'IL0', 'CXL', 'AOE', 'ANL', 'NA1', 'NAA', 'TBR', 'PFB', 'BRB',
'FEO', 'ELA', 'NNH', 'PP9', 'UNK', 'PEF', 'P6L', 'P3A', 'PCW', 'NHE', '15P', 'CPQ', ' SM', 'MPO', 'OPE', 'EPH', 'PEF', 'PLD', 'CGU',
'TMA', 'TBU', '0V5', 'BU3', 'DXE', 'ASO', 'M2P', 'BU1', 'GLF', 'P4G', 'PG5', 'I7P', 'I8P', '3PP', 'ETX', 'MXE', '2PG', 'B3P', 'ETE', 'ARS',
'EYG', 'I3P', 'IP5', 'XPE', 'DVT', 'FII', 'FUM', 'B7G', 'P33', ' HO', '6BP', 'HEX', '6NA', ' PI', 'INS', 'IHP', 'C15', 'OAA', 'IPH', 'PGH', 'PPF',
'7PE', 'PE3', 'PPI', 'PUT', 'AGL', 'RH3', 'SRT', 'SIN', 'SAT', 'TMO', 'DTU', 'LYK', 'L2O', 'PPK', 'PIE', '52N', 'BU2', '13P', 'DIO', 'MA1',
'2FP', 'NBG', 'OLC', 'COM', '1PG', 'NDG', 'ALS', 'FGP', 'FGL', '5MM', 'AKG', 'PGA', 'MGM', 'MAH', 'PC1', 'FUL', 'G6D', 'EIC', 'ACM',
'GLA', 'GLG', 'TRE', 'GLC', 'VG1', 'G1P', 'G16', 'G6P', 'MAK', 'MAN', 'RAM', 'FUC', 'AF3', 'MLR', 'GLS', 'GAL', 'BGC', 'M6D', 'BG6',
'LAT', ' CS', 'CH6', 'TAR', 'DKA', 'HG2', 'CPT', 'MLA', 'DMA', 'CDL', 'LMT', 'SDS', 'PLC', 'PEE', 'EMC', 'FAR', 'FPP', 'GRG', 'F6P', 'F6R',
' GD', 'HEZ', 'GOA', '4IP', 'ISD', 'ICT', 'IDR', 'LMR', 'DAO', 'PEQ', ' PB',' LU', 'MF4', 'MLT', 'MLI', 'MAL', 'MOH', 'SGM', 'MYR', 'OCT',
'2PE', 'IDU', 'SIA', 'IDS', '37X', 'HSJ', 'DR6', 'OLA', ' OS', 'PLM', '7PH', 'PO3', 'PEP', '12P', 'P4C', ' PR', 'POP', 'PYR', ' RB', 'PGO',
'TTN', 'STE', 'SPV', 'SO3', 'SUC', 'ALF', 'TGL', 'TFA', 'MGF', 'BET', 'TAM', '3PO', 'WO4', '  W', 'UNL', 'URE', 'OXL', 'YT3', 'VO4', 'MES',
'LMU', 'CRO', 'NRQ', 'UNX', 'OIM', 'OFM', ' PD', 'SMC', 'PEO', 'PCA', 'NO2', 'NME', 'NAG', 'MN3', 'MME', 'MLZ', ' LI', 'HDD', 'GAI',
'FLC', 'DOD', 'DMF', 'BMA', 'CRQ', 'BOG', 'CIT', 'CPS', 'CSX', '  F', '202', 'SO4', 'GOL', ' ZN', ' MG', ' CA', ' CL', ' NA', 'EDO', 'PO4',
'HEM', 'ACT', ' MN', '  K', 'PEG', ' FE', 'ACE', ' CU', 'MPD', ' NI', 'NH2', 'DMS', 'BME', 'TRS', ' CD', 'PG4', 'FMT', 'ACY', 'SEP', 'EPE',
'SF4', 'PGE', ' CO', 'TPO', 'IOD', 'FE2', 'FES', 'IMD', 'NO3', ' HG', 'CSO', 'PTR', 'IPA', '1PE', 'KCX', 'MRD', 'HEC', 'CMO', ' BR', 'CO3',
'CME', 'TLA', 'SCN', 'OXY', 'P6G', 'NCO', 'CAC', 'TYS', 'CSD', 'MLY', 'DAL', 'F3S', 'EOH', 'CU1', 'AZI', 'HED', 'M3L', 'DTT', 'NH4',
'HYP', 'OCS', 'CYN', 'FME', 'CAS', 'ALY', 'BTB', 'ABA', ' XE', 'CSW', 'BCT', ' SR', ' BA', ' NO', 'C8E', 'BCL', 'BEF', 'PE4', ' AU', ' CP', '
PT', ' YB', 'MSE', 'ORN', '4BF', 'CCS', 'HOH', ' OH']

"""
list of amino acids
"""
aa = ['ALA', 'CYS', 'ASP', 'GLU', 'PHE', 'GLY', 'HIS', 'ILE', 'LYS', 'LEU', 'MET', 'PRO', 'GLN', 'ARG', 'SER', 'THR', 'VAL', 'TRP', 'TYR']


"""
Given two sets of [x,y,z] coordinates, finds the distance between them
"""
def dist (a1,a2):
    return math.sqrt((a1[0]-a2[0])*(a1[0]-a2[0])+(a1[1]-a2[1])*(a1[1]-a2[1])+(a1[2]-a2[2])*(a1[2]-a2[2]))


def unique(List):
    checked = []
    for e in List:
```

```
            if e not in checked:
                checked.append(e)
        return checked



class pdbentry:
        'Base class for pdb files'

        count = 0
        noLigand = 0
        noLigandList = []
        liganded = 0

        uniprotList = {}
        dummyList = {}

        def __init__(self,name):
                self.name = name
                self.path = pdbpath + '/' + name + '.pdb'
                pdbentry.count += 1

        def filelines(self):
                file = open(self.path)
                filelines = file.readlines()
                file.close
                return filelines

        """
        function to get all names and coordinates of amino acids of given type. format: [['HIS A   1',[x,y,z]], ['HIS A   2',[x,y,z]]]
        """
        def getAAResidues(self):
                coordinates = []  #list to store residue names and coordinates. format is [ ['HIS A   1' , [x,y,z] ], ['HIS A   2' , [x,y,z] ]

                for line in self.filelines():
                        if line[0:4] == 'ATOM':
                                if ((line[12:16]==" NE2") and (line[17:20]=="HIS")):
                                        coordinates.append((line[17:26] + line[12:16], [float(line[30:38]),float(line[38:46]),float(line[46:54])]))
                                elif ((line[12:16]==" ND1") and (line[17:20]=="HIS")):
                                        coordinates.append((line[17:26] + line[12:16], [float(line[30:38]),float(line[38:46]),float(line[46:54])]))

                return coordinates


        'function returns [{ligand 3-letter code: ligand full name}, [[3-letter code, [[x,y,z],[x,y,z],...],[3-letter code, [[x,y,z],[x,y,z],...]]'
        def getLigands(self):
                ligandsDict = {}
                ligandsList = []
                ligandCoords = []
                ligandsCoordinates = []
                ligandLast = ''
                flag = 0

                for line in self.filelines():

                        'adds the names of all ligands in structure to a list'
                        if line[0:6] == "HETNAM" and line[11:14] in badligands:
                                pass
                        if line[0:6] == "HETNAM" and line[11:14] not in badligands:
                                flag = 1
                                ID = line[11:14]
                                name = line[15:].strip()
                                if ID not in ligandsDict.keys():
                                        ligandsDict[ID] = (name)
                                        ligandsList.append(line[11:14])

                                'checks if the current line has atom coordinates for the same ligand as last loop, then adds coordinates to
placeholder list'
                        elif line[0:6] == 'HETATM' and line[17:20] == ligandLast:
                                ligandCoords.append([float(line[30:38]),float(line[38:46]),float(line[46:54])])
```

```
                'checks if line contains new ligand. if yes, appends coordinates of last ligand to master list and clears
placeholder list'
            elif line[0:6] == 'HETATM' and line[17:20] in ligandsList:
                if flag == 0:
                    ligandLast = line[17:20]
                    ligandCoords.append([float(line[30:38]),float(line[38:46]),float(line[46:54])])
                    flag = 1
                else:
                    ligandsCoordinates.append([ligandLast,ligandCoords])
                    ligandCoords = []
                    ligandLast = line[17:20]
                    ligandCoords.append([float(line[30:38]),float(line[38:46]),float(line[46:54])])
        ligandsCoordinates.append([ligandLast,ligandCoords])

        'keeps count of number of PDBs with or without a "good" ligand'
        'inaccurate if getLigands function is called for a single PDB code more than once'
        if flag == 0:
            pdbentry.noLigand += 1
            pdbentry.noLigandList.append(self.name)
        else:
            pdbentry.liganded += 1

        return ligandsDict, ligandsCoordinates

    'function returns list of AAs within 10A of ligand. format: [[[3-letter code, "HIS A   1", distance], ...],{ligandDictionary}]'
    def aaNearLigand(self):
        ligandArray = self.getLigands()
        aaArray = self.getAAResidues()
        proxAAs = []
        'nested loops go through for each ligand, for each AA, for all ligand atom coordinates, is AA atom and ligand atom within
10A'
        'if yes, replace 10A cutoff with their distance and continue to scan through ligand atom coordinates'
        'this ensures the final reported distance is the minimum distance between ligand and AA'
        for ligand in ligandArray[1]:
            for aa in aaArray:
                cutoff = 10.0
                for xyz in ligand[1]:
                    if dist(aa[1],xyz) < cutoff:
                        cutoff = dist(aa[1],xyz)
                if cutoff < 10.0:
                    proxAAs.append([ligand[0],aa[0],cutoff])
        return proxAAs, ligandArray[0]


    'function to get additional info for output file. format: [[[Chain,DB abbreviation,DB ID,Entry description],...],[[molecule
name,chain],...]]'
    def getInfo(self):
        geneInfo = []
        names = []
        for line in self.filelines():                                   #scans through all lines in file
            if line[25:29] == ' UNP':                                    #pulls out chain ID, uniprot ID, and species
                geneInfo.append([line[12:13], line[26:32], line[33:41], line[42:54]])
                if line[33:41] not in pdbentry.uniprotList:
                    pdbentry.uniprotList[line[33:41]] = [self.name]
                elif line[33:41] in pdbentry.uniprotList and self.name not in pdbentry.uniprotList[line[33:41]]:
                    pdbentry.uniprotList[line[33:41]].append(self.name)
            elif line[0:6] == 'COMPND' and line[11:19] == 'MOLECULE':          #stores the protein name
                nameproxy = line[21:]
            elif line[0:6] == 'COMPND' and line[11:16] == 'CHAIN':             #associates stored protein name with correct
chains and adds it to list for output
                names.append([nameproxy,line[18:]])
        pdbentry.dummyList = pdbentry.uniprotList.copy()                    #make a dummy list for editing to count number
of uniprots with a ligand and with a proximal ligand

        return geneInfo, names


'formats information for csv creation. format: [[PDB code, protein name, ligand name, ligand abbreviation, residue name, distance,
Uniprot ID, Species],...]'
```

```
'output includes all AAs found in a given PDB'
def FormatOutput(PDBcode):
	Output = []
	PDB = pdbentry(PDBcode)
	print(PDB.name)
	proxAAs,ligandDict = PDB.aaNearLigand()
	info = PDB.getInfo()

	for AA in proxAAs:
		chain = AA[1][4]			#gets the chain ID for a given AA entry
		for proteinName in info[1]:		#checks which protein name is relevant to the given chain ID
			if chain in proteinName[1]:
				chainName = proteinName[0]
				break
		if info [0] != []:
			for geneName in info[0]:		#checks what gene info set is relevant to the given chain ID
				if geneName[0] == chain:
					relInfo = geneName
					break
			try:
				x = relInfo[2]
			except:
				relInfo = [' ',' ',' ',' ']
		else:
			relInfo = [' ',' ',' ',' ']

		placeholder = [PDB.name,chainName,ligandDict[AA[0]],AA[0],AA[1],AA[2],relInfo[2],relInfo[3]]
		Output.append(placeholder)
	return Output

"""
Takes a results array and outputs a CSV file
"""
def writeCSV (outputTable, name):
	h = HTML()
	ta = h.script(' ', src='../scripts/sorttable.js')
	ta.text("</script>", escape=False)
	ta.table(border='1', klass='sortable')
	r = ta.tr
	csv_output = []

	'Adds hyperlinks to PDB code and Uniprot ID'
	for AA in outputTable:
		print(AA)
		csvpdblink = "=HYPERLINK(\"http://www.rcsb.org/structure/%s\"" % AA[0] + ",\"%s\")" % AA[0]
		if AA[6] != '':
			csvuniprotlink = "=HYPERLINK(\"http://uniprot.org/uniprot/%s\"" % AA[6].strip() + ",\"%s\")" % AA[6].strip()
		else:
			csvuniprotlink = AA[6]
		try:
			csv_output.append([csvpdblink] + AA[1:6] + [csvuniprotlink] + [AA[7]])
		except:
			print ("Error1.")
			csv_output.append([csvpdblink] + AA[1:])
	f = open(name, 'w')
	f.write(str(ta))
	f.close()

	with open("%s.csv" % name, 'wb') as csvfile:
		csvwriter = csv.writer(csvfile)
		csvwriter.writerow(['PDB Code', 'Protein Description', 'Ligand', 'Ligand Abbrv.', 'Amino Acid', 'Distance from Ligand (A)',
'Uniprot/Genebank ID', 'Species'])
		for line in csv_output:
			csvwriter.writerow(line)


"""
MAIN
"""
start_time = time.time()
```

```
'extracts PDB codes from text file list'
with open(pdblist) as f:
	pdbs1 = f.read().splitlines()
try:
	if pdbs1[0][4] == ',' :
		pdbs = pdbs1[0].split(',')
except:
	pdbs = pdbs1

'compiles a results list to feed to the writeCSV function'
outputLog = []
holder = ['','','','','']
check = 1;
uniprotsProx = []

for entry in pdbs:
	results = FormatOutput(entry)			#gets results for current PDB code
	for AA in results:
		if AA[6] not in uniprotsProx:		#running list of unique uniprots with a proximal ligand
			uniprotsProx.append(AA[6])
		if AA[4][0:9] == holder[4][0:9]:		#checks if both ND and NE are within the range cutoff
			check = 1;			#if yes, selects the one that is closer
			if AA[5] < holder[5]:
				outputLog.append(AA)
			else:
				outputLog.append(holder)
		elif check == 1:			#checks if last residue was a duplicate entry, thus already appended to the results list
			holder = AA;			#if yes, saves the new residue to compare to the next line to check for duplicate entries
			check = 0;
		else:					#scenario where previous residue was a singlet entry
			outputLog.append(holder)		#adds previous residue to results list and saves current residue for comparison
to the next
			holder = AA

writeCSV(outputLog, outputname)

print('Total PDB codes: ' + str(pdbentry.count))
print('PDB codes with no ligand: ' + str(pdbentry.noLigand))
print('PDB codes with a ligand: ' + str(pdbentry.liganded))

'checks if the getLigands function was run more than once, and warns if it was'
if pdbentry.noLigand + pdbentry.liganded != pdbentry.count:
	print('These numbers are inaccurate!')


print('Number of unique uniprot IDs: ' + str(len(pdbentry.uniprotList)))

'counting how many uniprots have at least one structure with a ligand'
removeList = []
for PDB in pdbentry.noLigandList:			#trimming the dummy list down to just those uniprots that have at least one
associated structure containing a ligand
	for item in pdbentry.dummyList:
		if PDB in pdbentry.dummyList[item]:		#if the PDB code has no ligands in the structure, remove the PDB code fronm
that uniprot
			pdbentry.dummyList[item].remove(PDB)
		if pdbentry.dummyList[item] == []:		#if uniprot has no PDB codes left, there are no structures where it has a ligand
so remove it from the dictionary
			removeList.append(item)

removeList = unique(removeList)

for item in removeList:
	pdbentry.dummyList.pop(item)

print('Number of unique uniprot IDs that have at least one structure with a ligand: ' + str(len(pdbentry.dummyList)))
print('Number of unique uniprot IDs that have at least one proximal ligand: ' + str(len(uniprotsProx)))
```

```
end_time = time.time()

run_time = end_time - start_time

print('Time: ' + str(run_time))
```

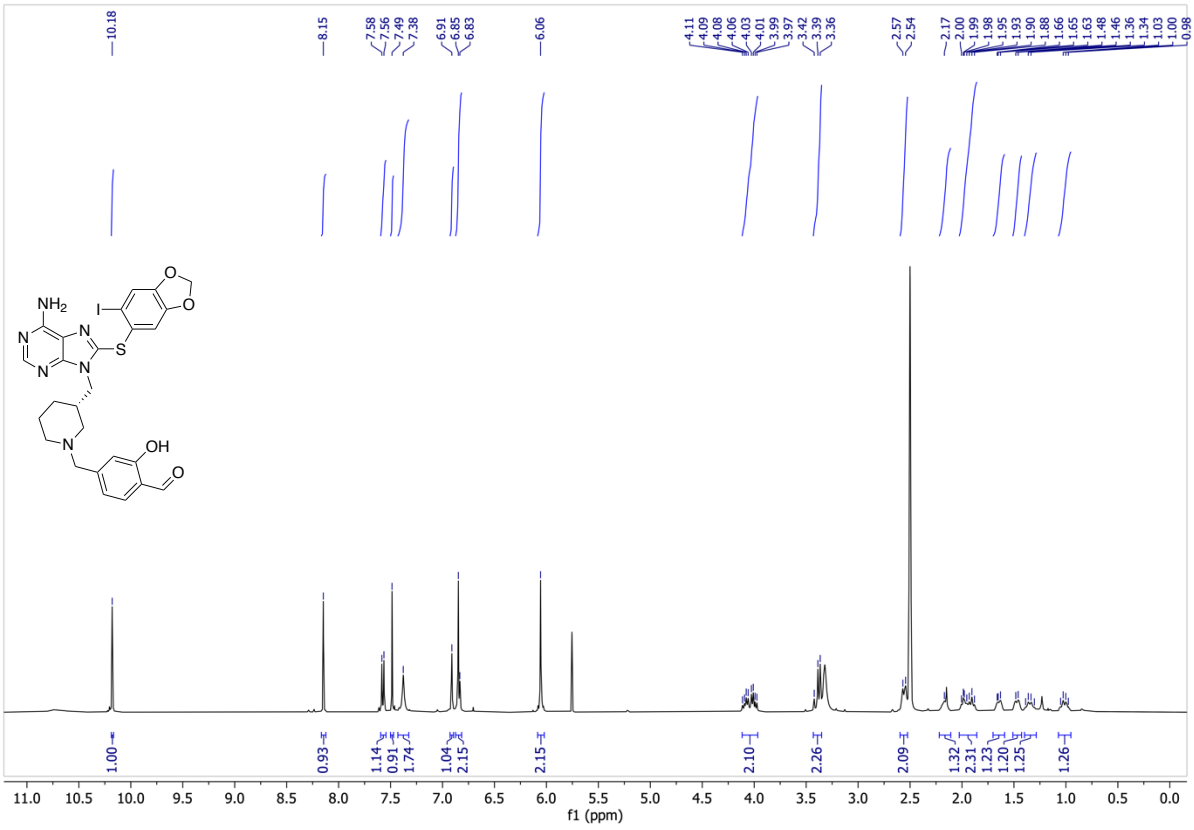
Appendix B: X-ray data collection and refinement statistics

Table S1. X-ray data collection and refinement statistics

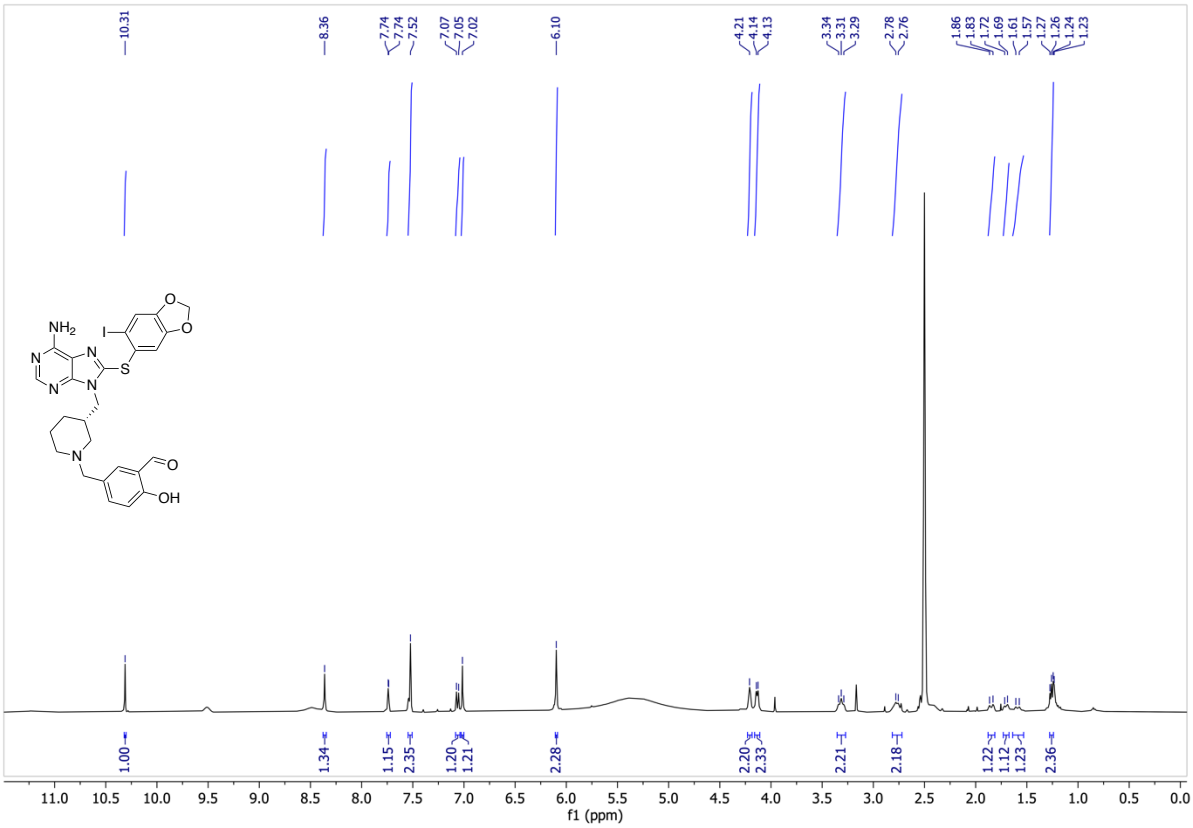
	Hsp90-compound **6** complex	Hsp90 K58H-AC45 complex	Hsp90-JW-2-128 complex
PDB Code	9AUU		
Wavelength	1.11583	1.11583	1.11583
Resolution range	67.16 - 2.0 (2.072 - 2.0)	52.54 - 2.1 (2.175 - 2.1)	32.53 - 1.8 (1.864 - 1.8)
Space group	I 2 2 2	I 2 2 2	I 2 2 2
Unit cell			
a, b, c (Å)	66.89 91.82 98.48	64.01 92.00 97.75	65.06 89.73 98.89
α, β, γ (°)	90 90 90	90 90 90	90 90 90
Total reflections	37882 (3654)	31900 (3187)	54189 (5366)
Unique reflections	20396 (1336)	16451 (1647)	27160 (2687)
Multiplicity	1.9 (1.9)	1.9 (1.9)	2.0 (2.0)
Completeness (%)	91.67 (65.43)	95.12 (94.83)	99.91 (99.93)
Mean I/sigma(I)	13.34 (2.37)	9.73 (2.25)	20.53 (1.27)
R-merge	0.03193 (0.2855)	0.05031 (0.3493)	0.0156 (0.5918)
R-work	0.2244 (0.2950)	0.1880 (0.2519)	0.1897 (0.3520)
R-free	0.2221 (0.2847)	0.2534 (0.2942)	0.2209 (0.3719)
RMS (bonds)	0.007	0.011	0.007
RMS (angles)	0.85	1.18	0.95
Ramachandran favored (%)	96.12	95.15	97.57
Ramachandran outliers (%)	0	0.49	0
Average B-factor	36.21	36.70	40.56
macromolecules	36.19	36.66	40.09
ligands	25.94	38.86	42.63
solvent	37.65	36.44	45.40

Appendix C: NMR Spectra

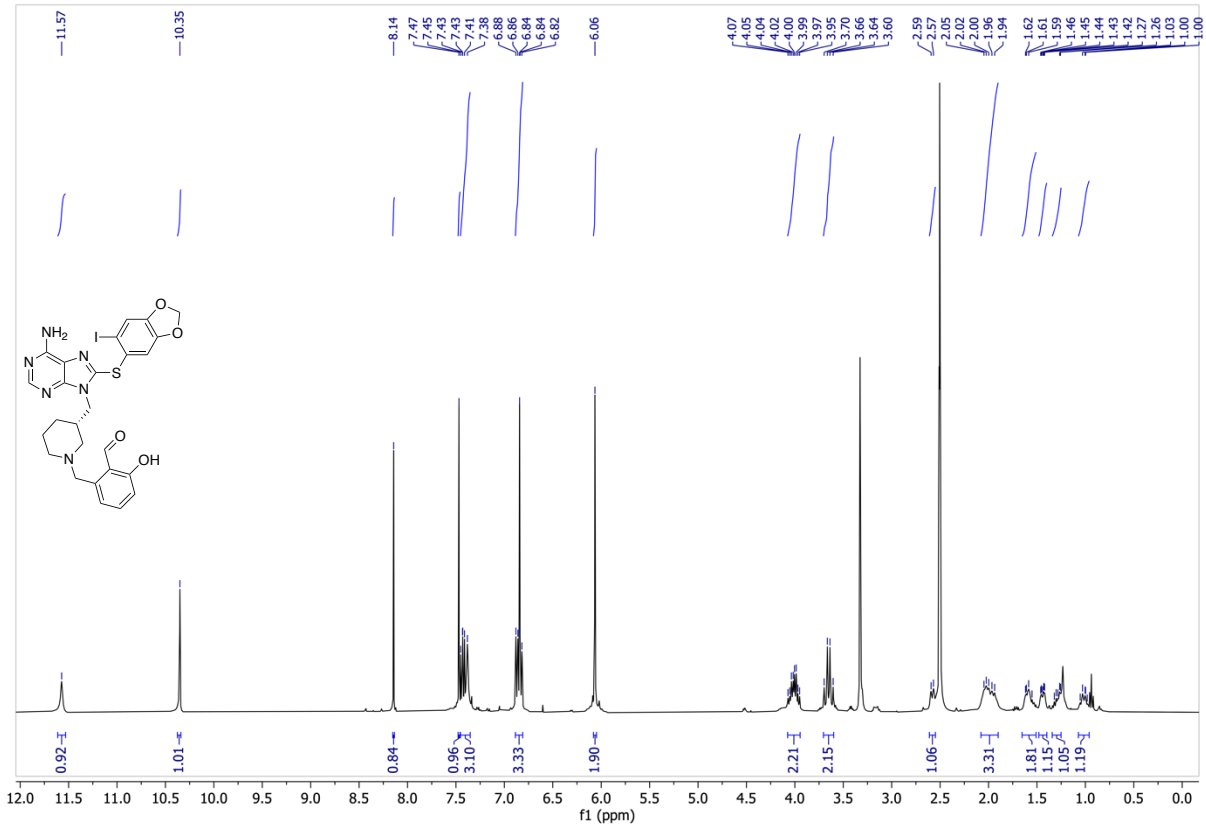
1



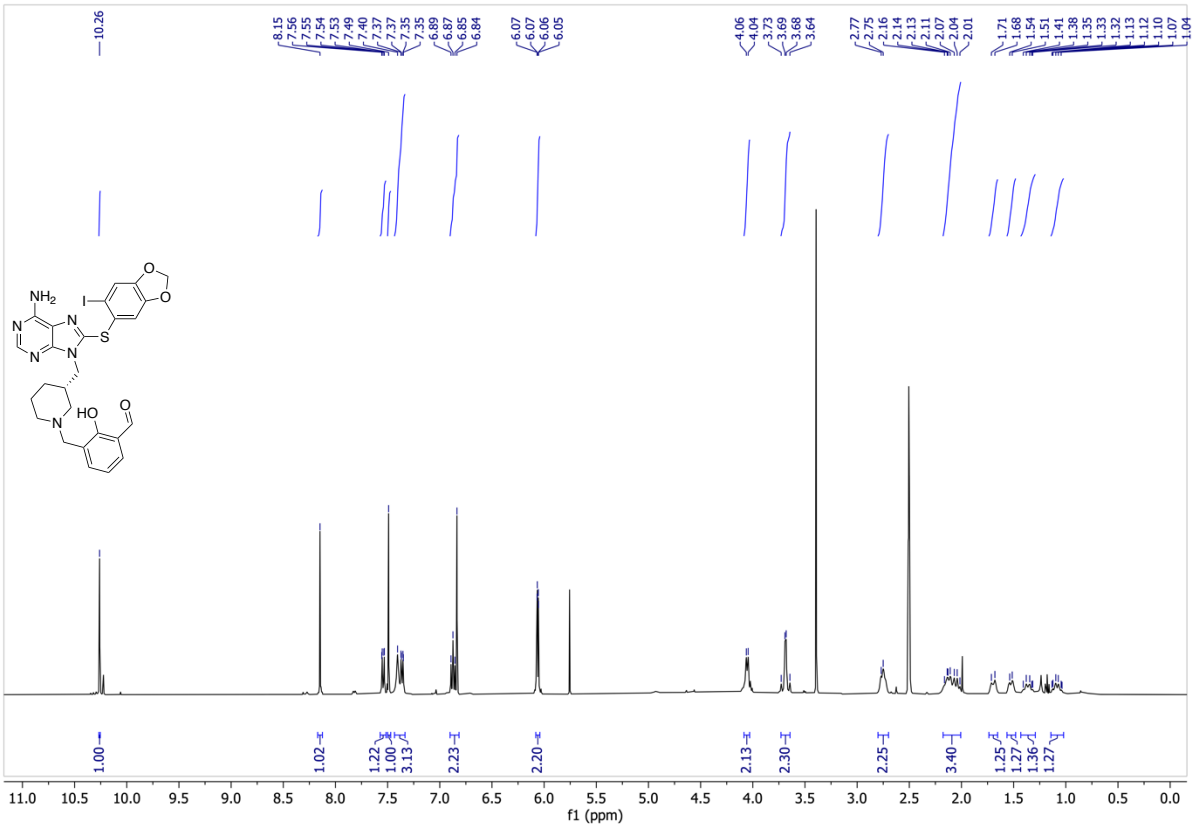
2


10.31
8.36
7.74
7.74
7.52
7.07
7.05
7.02
6.10
4.21
4.14
4.13
3.34
3.31
3.29
2.78
2.76
1.86
1.83
1.72
1.69
1.61
1.57
1.27
1.26
1.24
1.23
NH2
N
N
N
N
S
I
O
O
N
O
OH
1.00
1.34
1.15
2.35
1.20
1.21
2.28
2.20
2.33
2.21
2.18
1.22
1.12
1.23
2.36
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

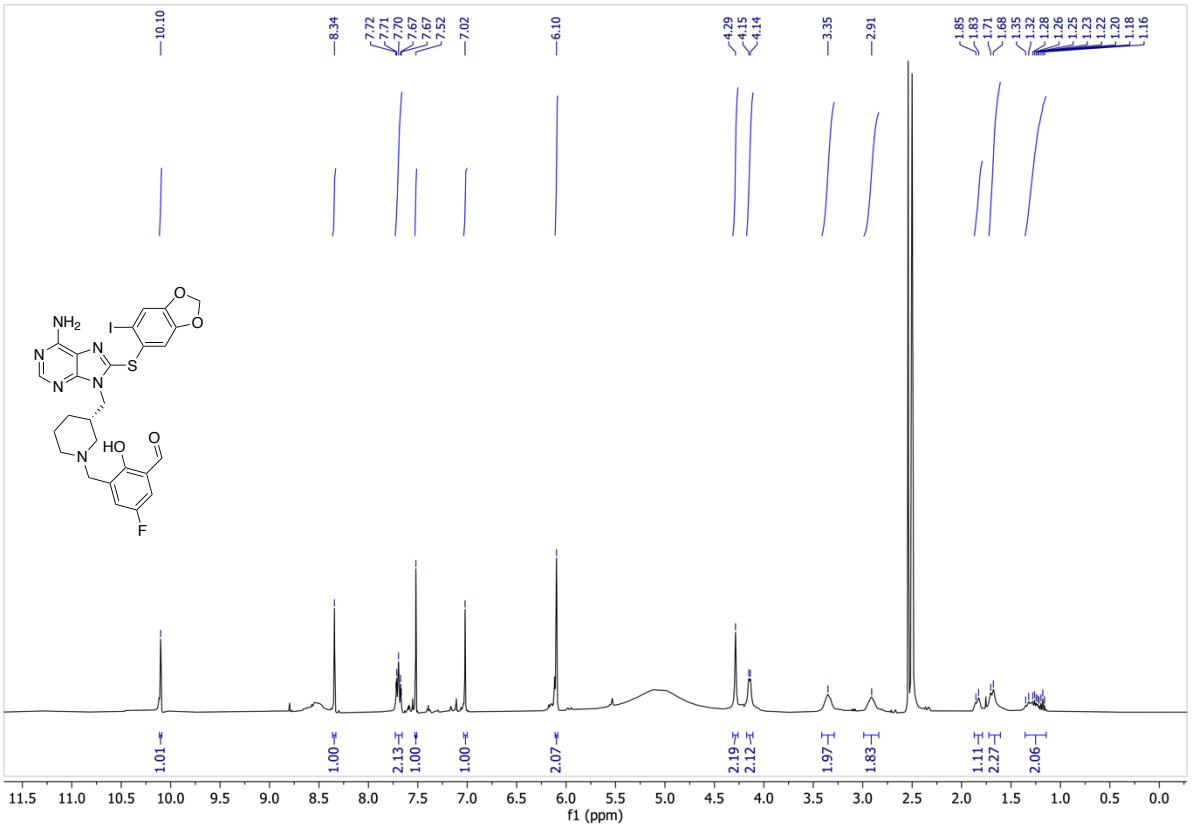
3


11.57
10.35
8.14
7.47
7.45
7.43
7.43
7.41
7.38
6.88
6.86
6.84
6.84
6.82
6.06
4.07
4.05
4.04
4.02
4.00
3.99
3.97
3.95
3.70
3.66
3.64
3.60
2.59
2.57
2.05
2.02
2.00
1.96
1.94
1.62
1.61
1.59
1.46
1.45
1.44
1.43
1.42
1.27
1.26
1.03
1.00
1.00
NH2
N
N
N
N
S
I
O
O
O
N
OH
0.92
1.01
0.84
0.96
3.10
3.33
1.90
2.21
2.15
1.06
3.31
1.81
1.15
1.05
1.19
12.0
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

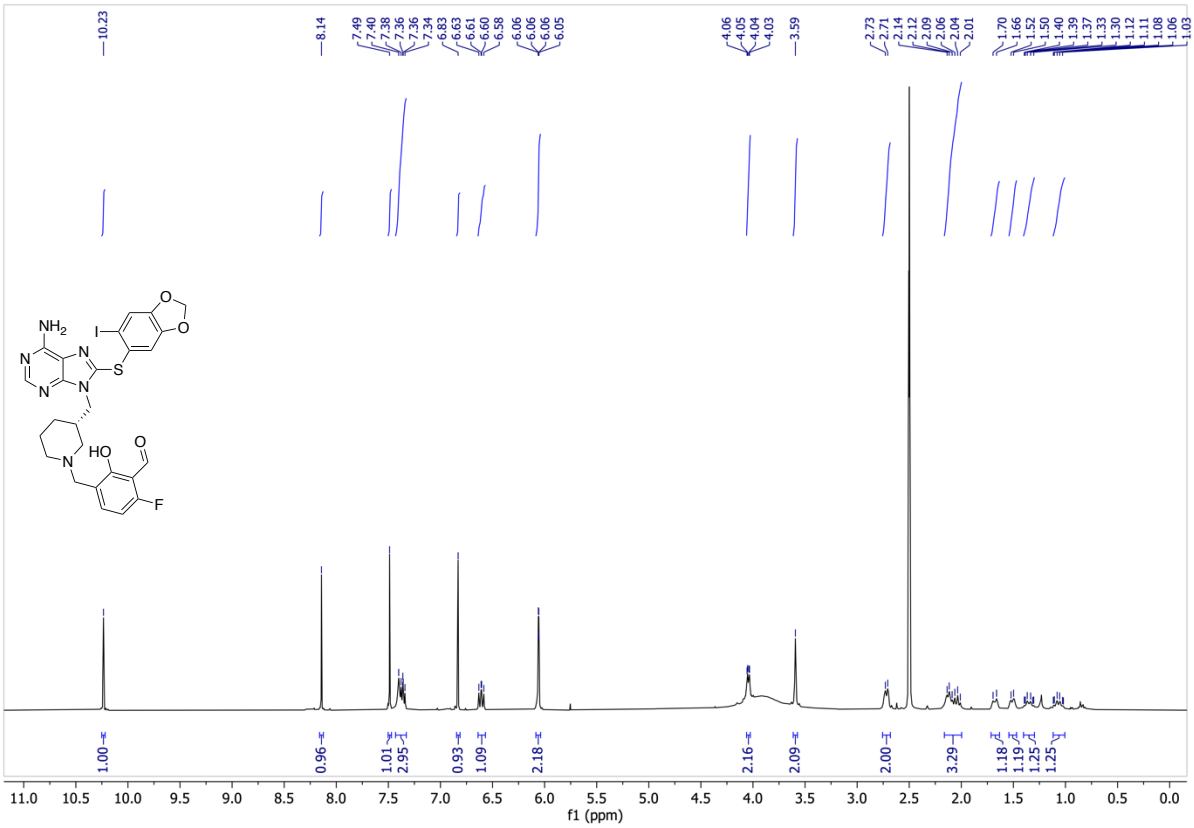
4


10.26
8.15
7.56
7.55
7.54
7.53
7.49
7.40
7.37
7.37
7.35
7.35
6.89
6.87
6.85
6.84
6.07
6.07
6.06
6.05
4.06
4.04
3.73
3.69
3.68
3.64
2.77
2.75
2.16
2.14
2.13
2.11
2.07
2.04
2.01
1.71
1.68
1.54
1.51
1.41
1.38
1.35
1.33
1.32
1.13
1.12
1.10
1.07
1.04
NH2
I
O
O
N
N
N
N
S
HO
O
1.00
1.02
1.22
1.00
3.13
2.23
2.20
2.13
2.30
2.25
3.40
1.25
1.27
1.36
1.27
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

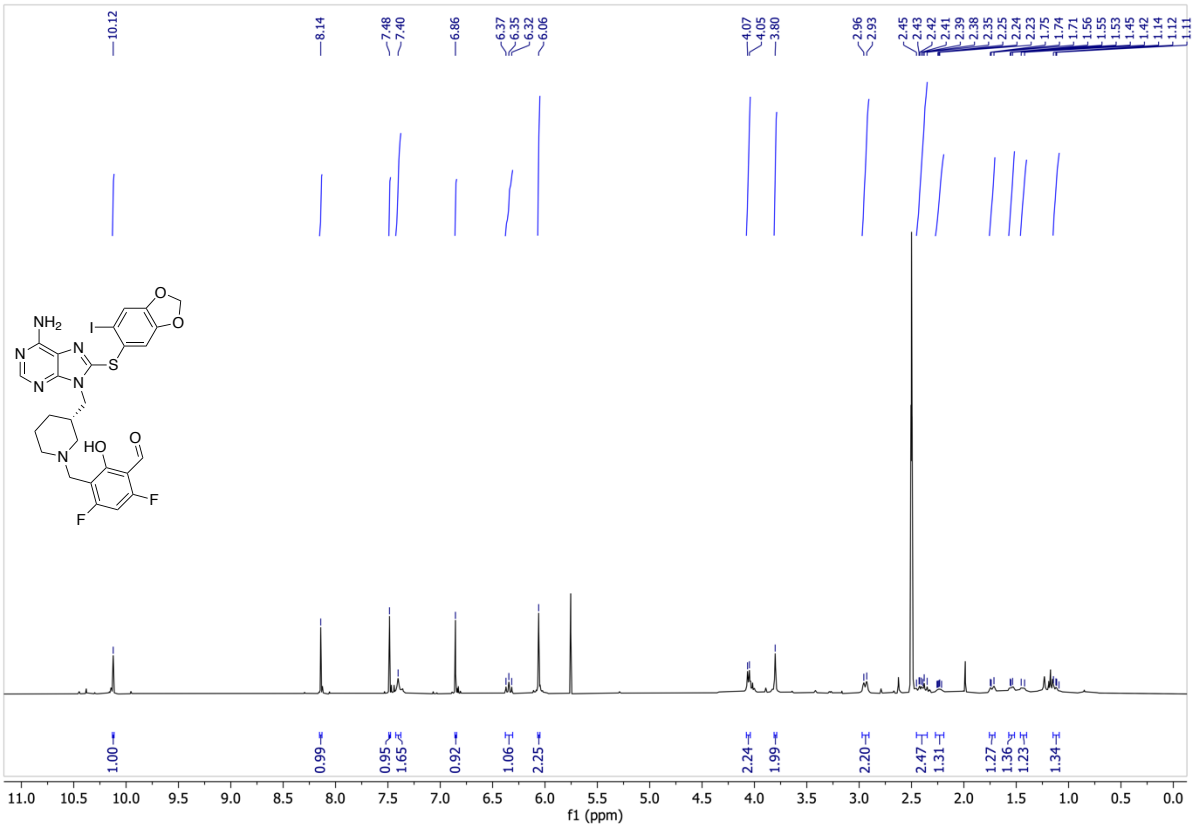
5


10.10
8.34
7.72
7.71
7.70
7.67
7.67
7.52
7.02
6.10
4.29
4.15
4.14
3.35
2.91
1.85
1.83
1.71
1.68
1.35
1.32
1.28
1.26
1.25
1.23
1.22
1.20
1.18
1.16
NH2
N
N
N
N
I
O
O
S
HO
O
N
F
1.01
1.00
2.13
1.00
1.00
2.07
2.19
2.12
1.97
1.83
1.11
2.27
2.06
11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

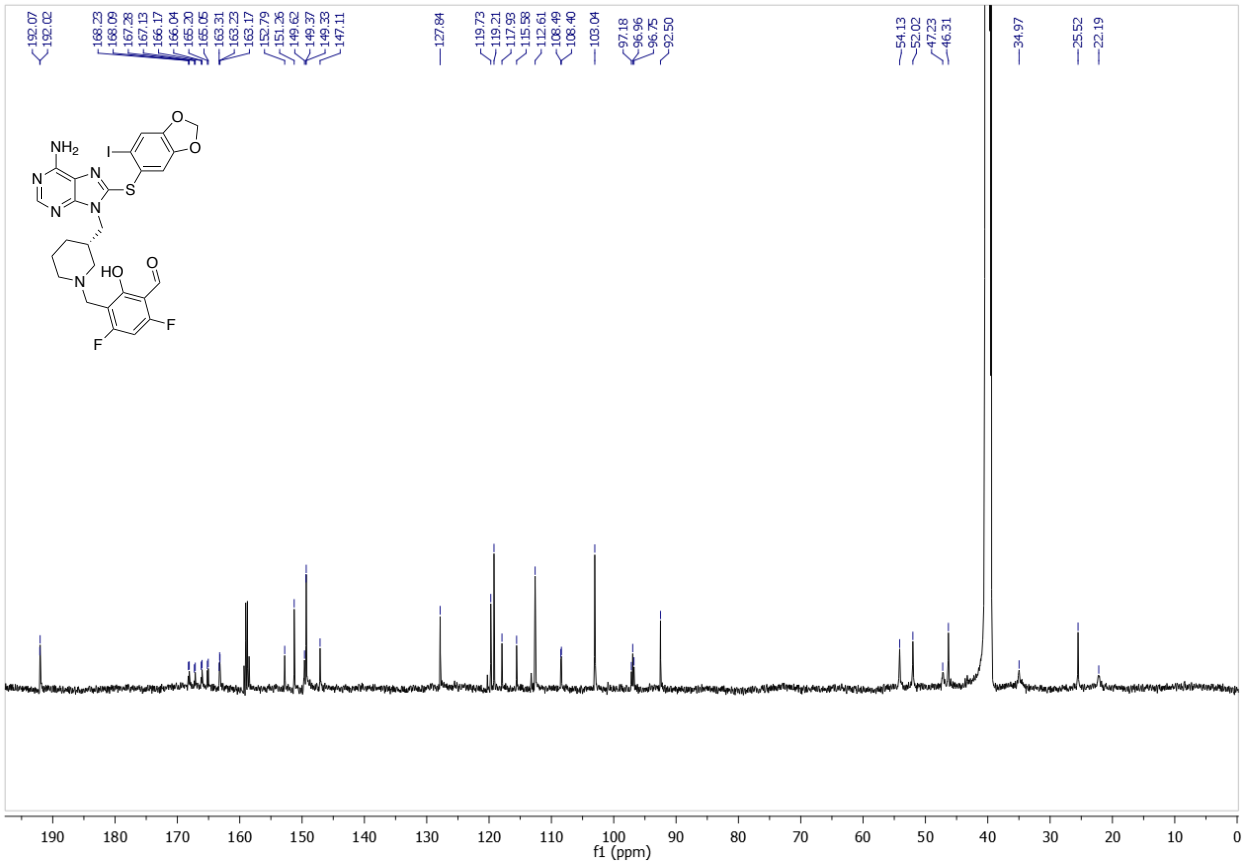
6


10.23
8.14
7.49
7.40
7.38
7.36
7.36
7.34
6.83
6.63
6.61
6.60
6.58
6.06
6.06
6.06
6.05
4.06
4.05
4.04
4.03
3.59
2.73
2.71
2.14
2.12
2.09
2.06
2.04
2.01
1.70
1.66
1.52
1.50
1.40
1.39
1.37
1.33
1.30
1.12
1.11
1.08
1.06
1.03
NH2
I
O
O
N
N
N
N
S
HO
O
F
1.00
0.96
1.01
2.95
0.93
1.09
2.18
2.16
2.09
2.00
3.29
1.18
1.19
1.25
1.25
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

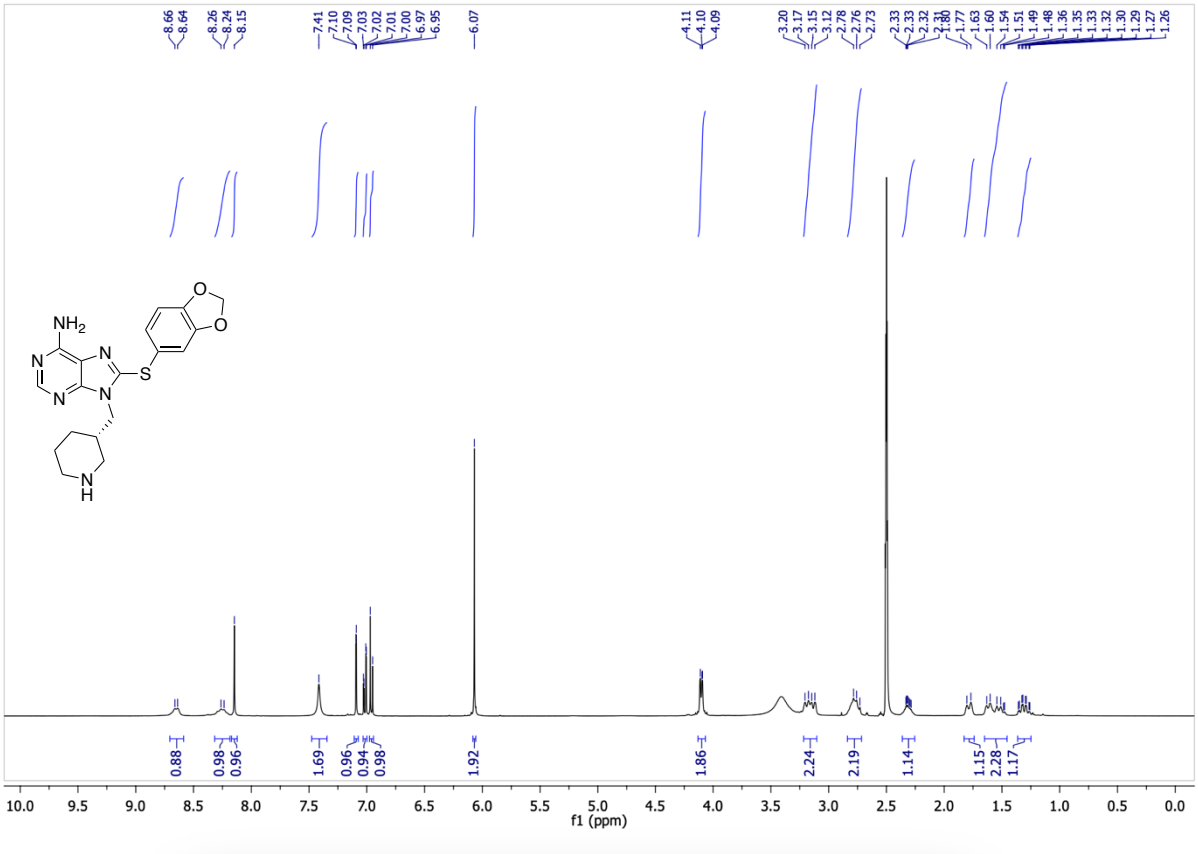
7


10.12
8.14
7.48
7.40
6.86
6.37
6.35
6.32
6.06
4.07
4.05
3.80
2.96
2.93
2.45
2.43
2.42
2.41
2.39
2.38
2.35
2.25
2.24
2.23
1.75
1.74
1.71
1.56
1.55
1.53
1.45
1.42
1.14
1.12
1.11
NH2
I
O
O
N
N
N
N
S
HO
O
N
F
F
1.00
0.99
0.95
1.65
0.92
1.06
2.25
2.24
1.99
2.20
2.47
1.31
1.27
1.36
1.23
1.34
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

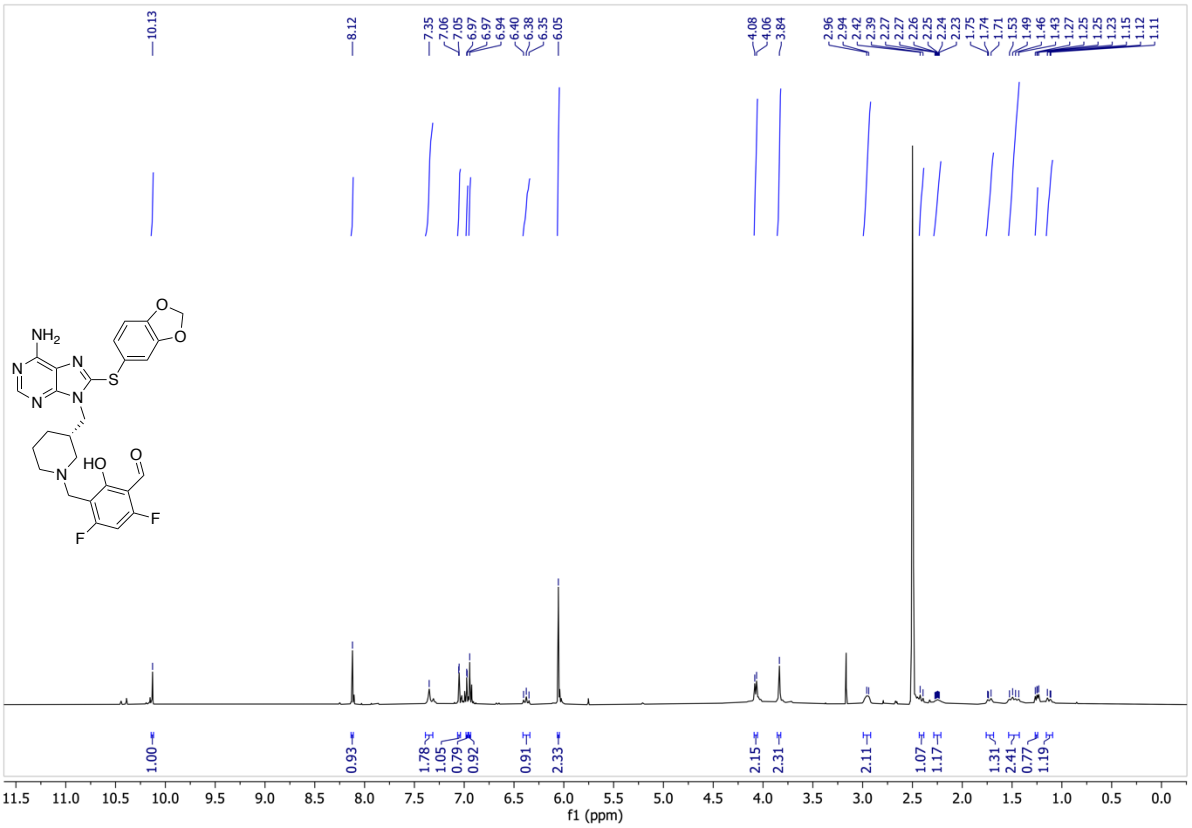
7


192.07
192.02
168.23
168.09
167.28
167.13
166.17
166.04
165.20
165.05
163.31
163.23
163.17
152.79
151.26
149.62
149.37
149.33
147.11
127.84
119.73
119.21
117.93
115.58
112.61
108.49
108.40
103.04
97.18
96.96
96.75
92.50
54.13
52.02
47.23
46.31
34.97
25.52
22.19
NH2
I
O
O
N
N
N
N
S
HO
O
N
F
F
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

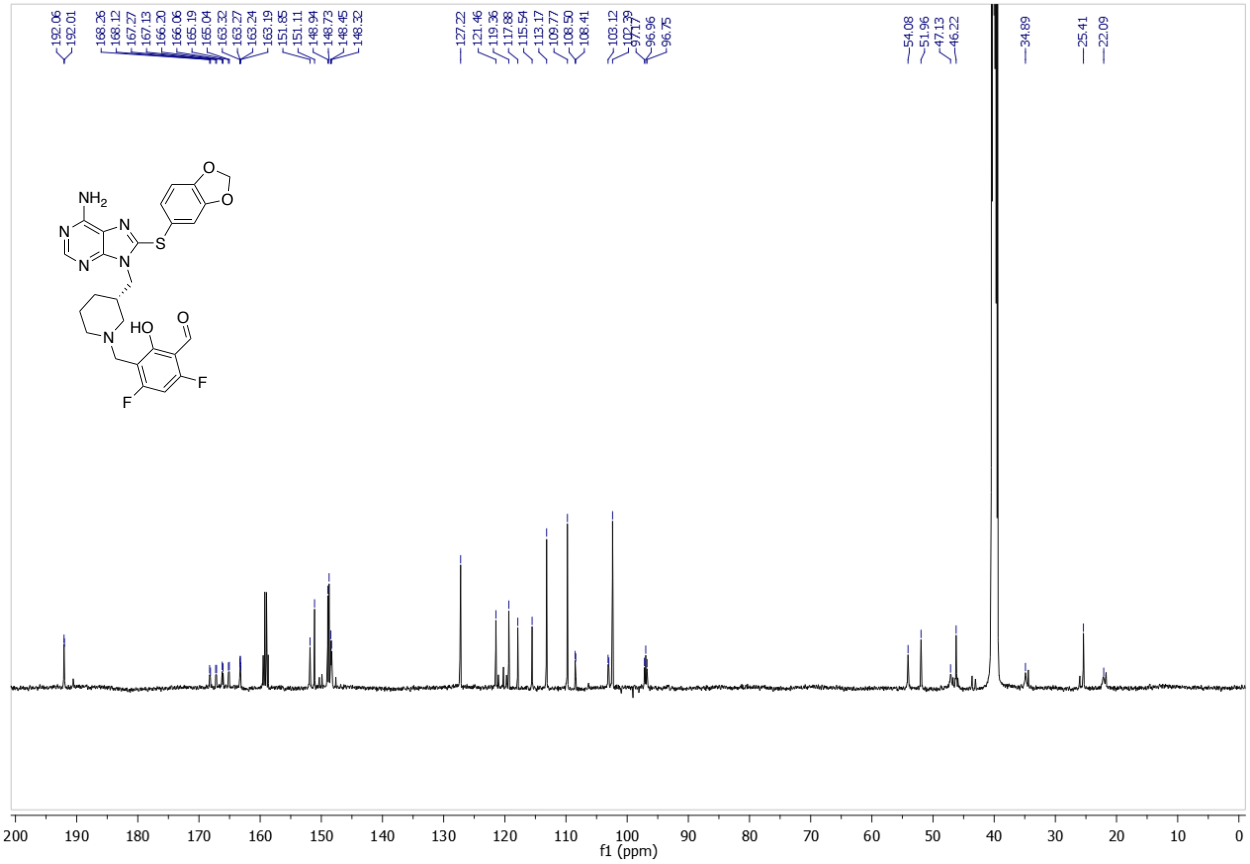
14


8.66
8.64
8.26
8.24
8.15
7.41
7.10
7.09
7.03
7.02
7.01
7.00
6.97
6.95
6.07
4.11
4.10
4.09
3.20
3.17
3.15
3.12
2.78
2.76
2.73
2.33
2.32
1.77
1.63
1.60
1.54
1.51
1.49
1.48
1.36
1.35
1.33
1.32
1.30
1.29
1.27
1.26
NH2
N
N
N
N
S
O
O
N
H
0.88
0.98
0.96
1.69
0.96
0.94
0.98
1.92
1.86
2.24
2.19
1.14
1.15
2.28
1.17
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

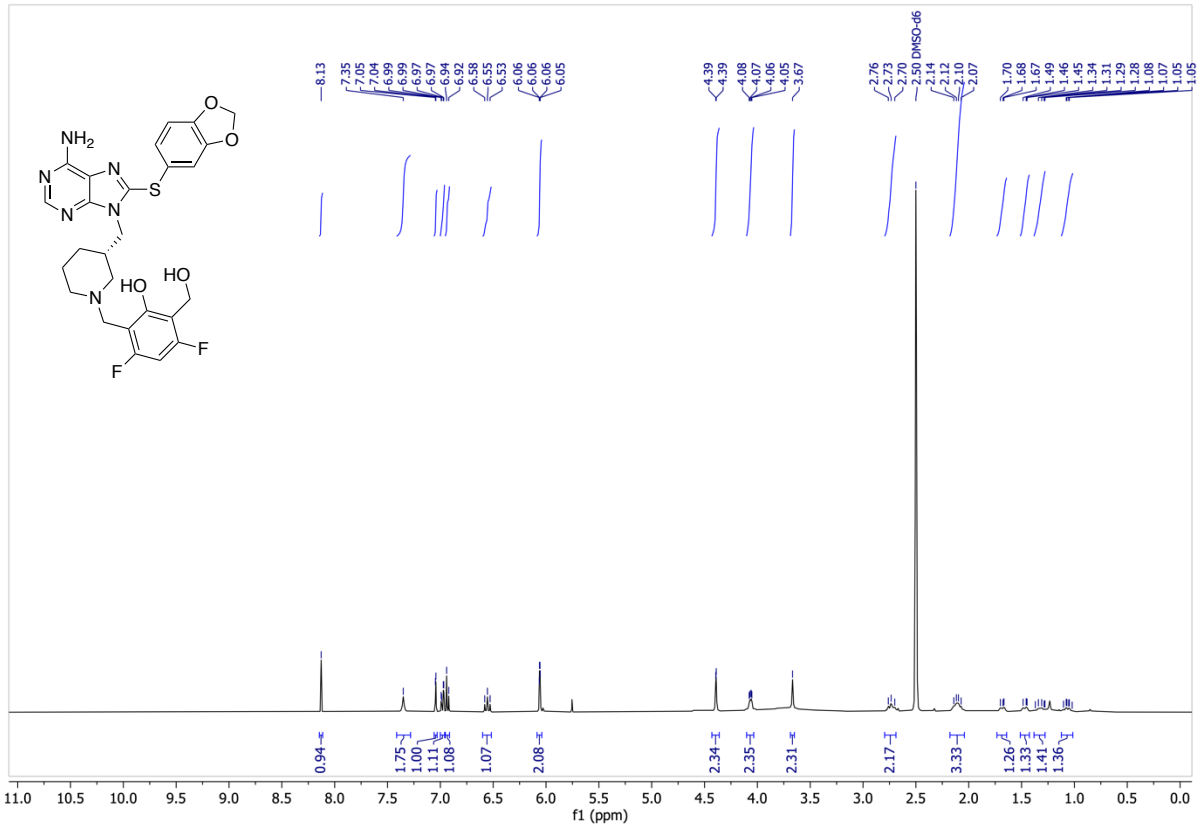
8


10.13
8.12
7.35
7.06
7.05
6.97
6.97
6.94
6.40
6.38
6.35
6.05
4.08
4.06
3.84
2.96
2.94
2.42
2.39
2.27
2.27
2.26
2.25
2.24
2.23
1.75
1.74
1.71
1.53
1.49
1.46
1.43
1.27
1.25
1.25
1.23
1.15
1.12
1.11
NH2
N
N
N
N
S
O
O
HO
N
O
F
F
1.00
0.93
1.78
1.05
0.79
0.92
0.91
2.33
2.15
2.31
2.11
1.07
1.17
1.31
2.41
0.77
1.19
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

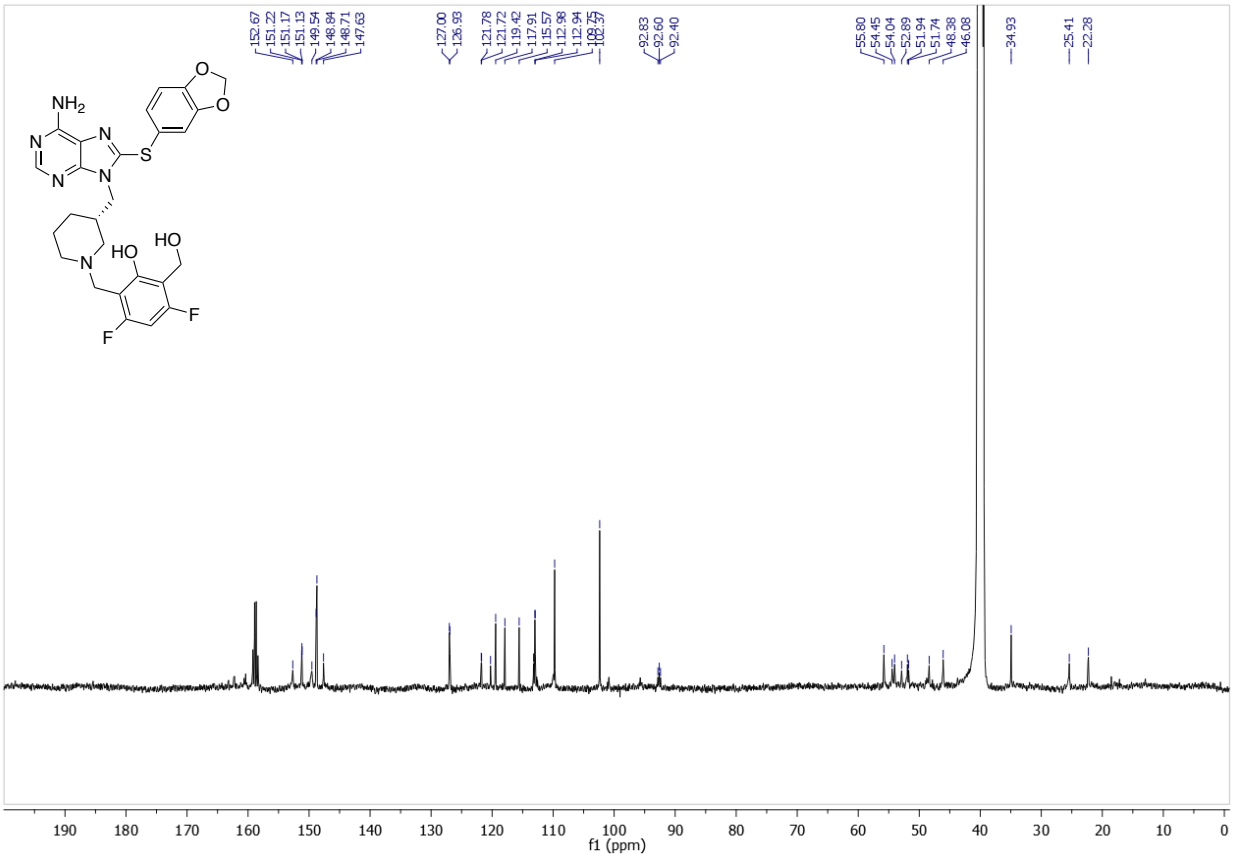
8


192.06
192.01
168.26
168.12
167.77
167.13
166.20
166.06
165.19
165.04
163.32
163.27
163.24
163.19
151.85
151.11
148.94
148.73
148.45
148.32
127.22
121.46
119.36
117.88
115.54
113.17
109.77
108.50
108.41
103.12
97.39
96.96
96.75
54.08
51.96
47.13
46.22
34.89
25.41
22.09
NH2
N
N
N
N
S
O
O
HO
O
N
F
F
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

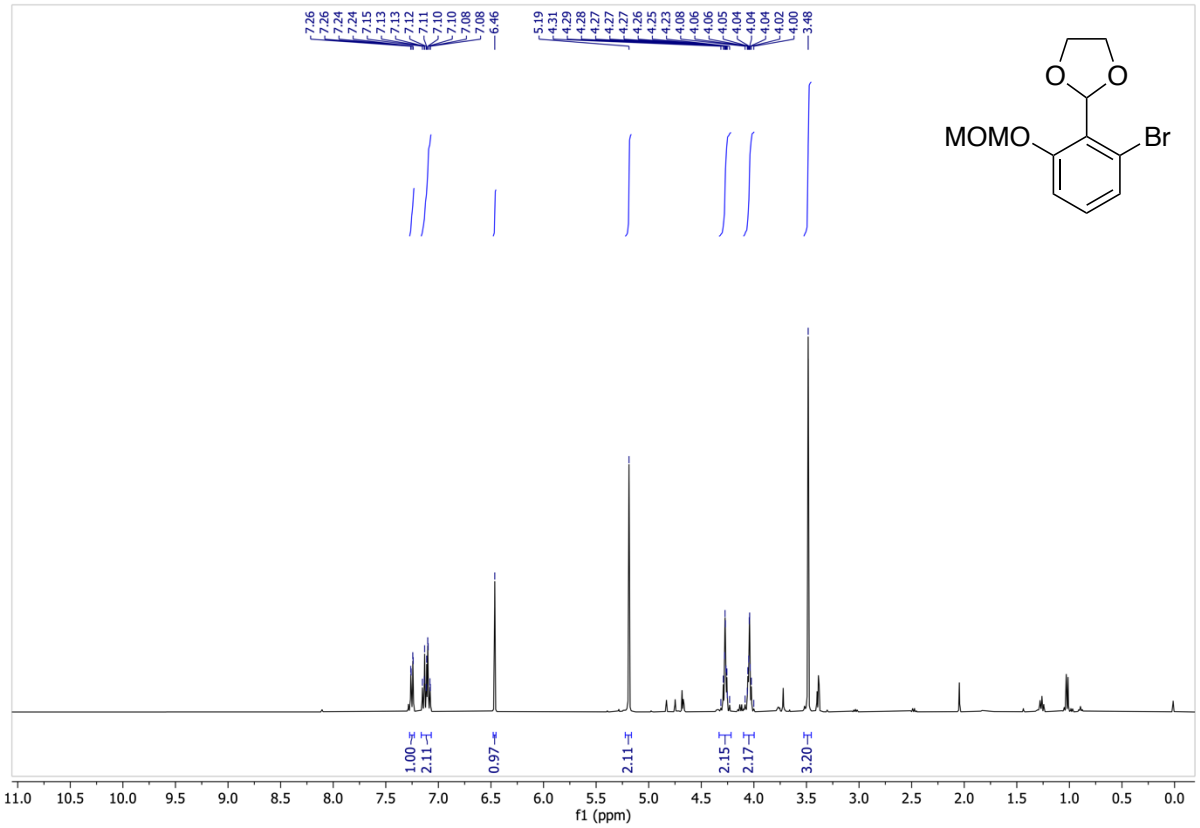
9


8.13
7.35
7.05
7.04
6.99
6.99
6.97
6.97
6.94
6.92
6.58
6.55
6.53
6.06
6.06
6.06
6.05
4.39
4.39
4.08
4.07
4.06
4.05
3.67
2.76
2.73
2.70
2.50 DMSO-d6
2.14
2.12
2.10
2.07
1.70
1.68
1.67
1.49
1.46
1.45
1.34
1.31
1.29
1.28
1.08
1.07
1.05
1.05
0.94
1.75
1.00
1.11
1.08
1.07
2.08
2.34
2.35
2.31
2.17
3.33
1.26
1.33
1.41
1.36
f1 (ppm)

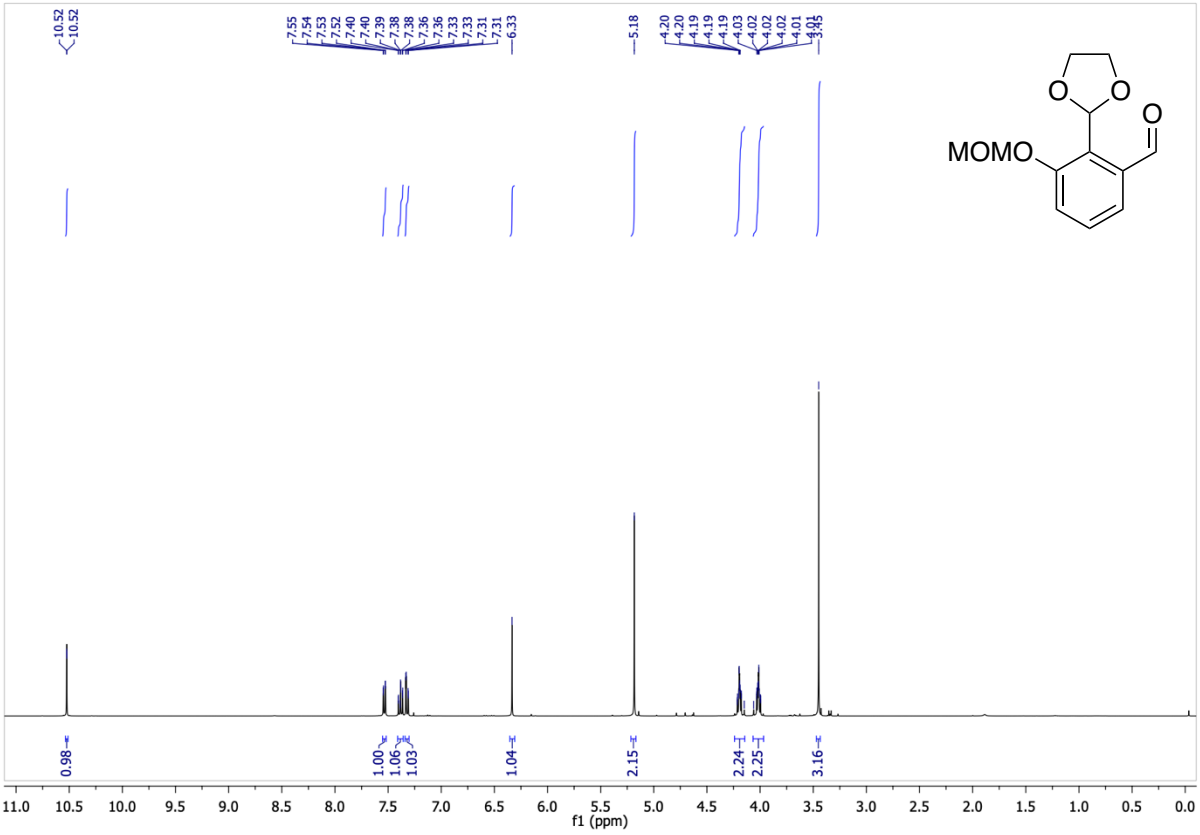
9


152.67
151.22
151.17
151.13
149.54
148.84
148.71
147.63
127.00
126.93
121.78
121.72
119.42
117.91
115.57
112.98
112.94
92.83
92.60
92.40
55.80
54.45
54.04
52.89
51.94
51.74
48.38
46.08
34.93
25.41
22.28
NH2
N
N
N
N
S
O
O
HO
HO
N
F
F
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

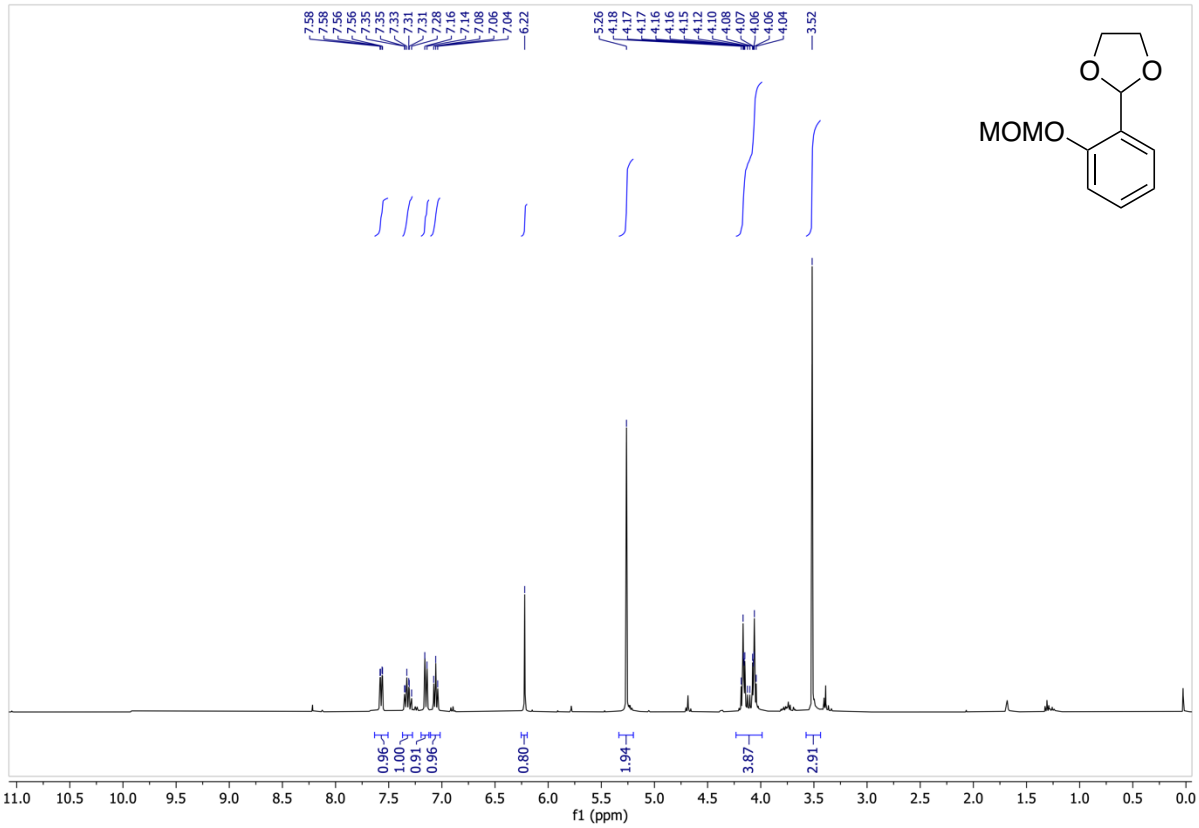
3a


7.26
7.26
7.24
7.24
7.15
7.13
7.12
7.11
7.10
7.10
7.08
7.08
6.46
5.19
4.31
4.29
4.28
4.27
4.27
4.27
4.26
4.25
4.23
4.08
4.06
4.06
4.05
4.04
4.04
4.04
4.02
4.00
3.48
O
O
MOMO
Br
1.00
2.11
0.97
2.11
2.15
2.17
3.20
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

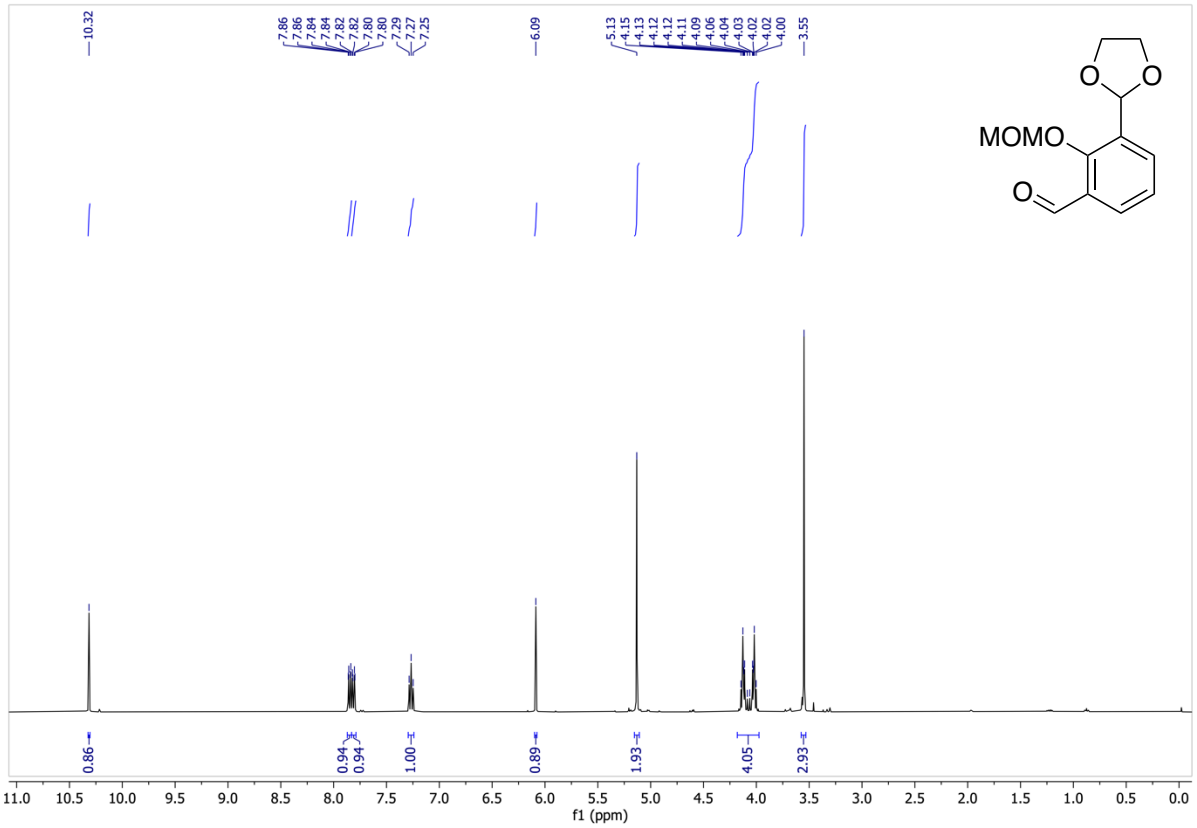
3b


10.52
10.52
7.55
7.54
7.53
7.52
7.40
7.40
7.39
7.38
7.38
7.36
7.36
7.33
7.33
7.31
7.31
6.33
5.18
4.20
4.20
4.19
4.19
4.19
4.03
4.02
4.02
4.02
4.01
4.01
3.45
O
O
O
MOMO
0.98
1.00
1.06
1.03
1.04
2.15
2.24
2.25
3.16
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

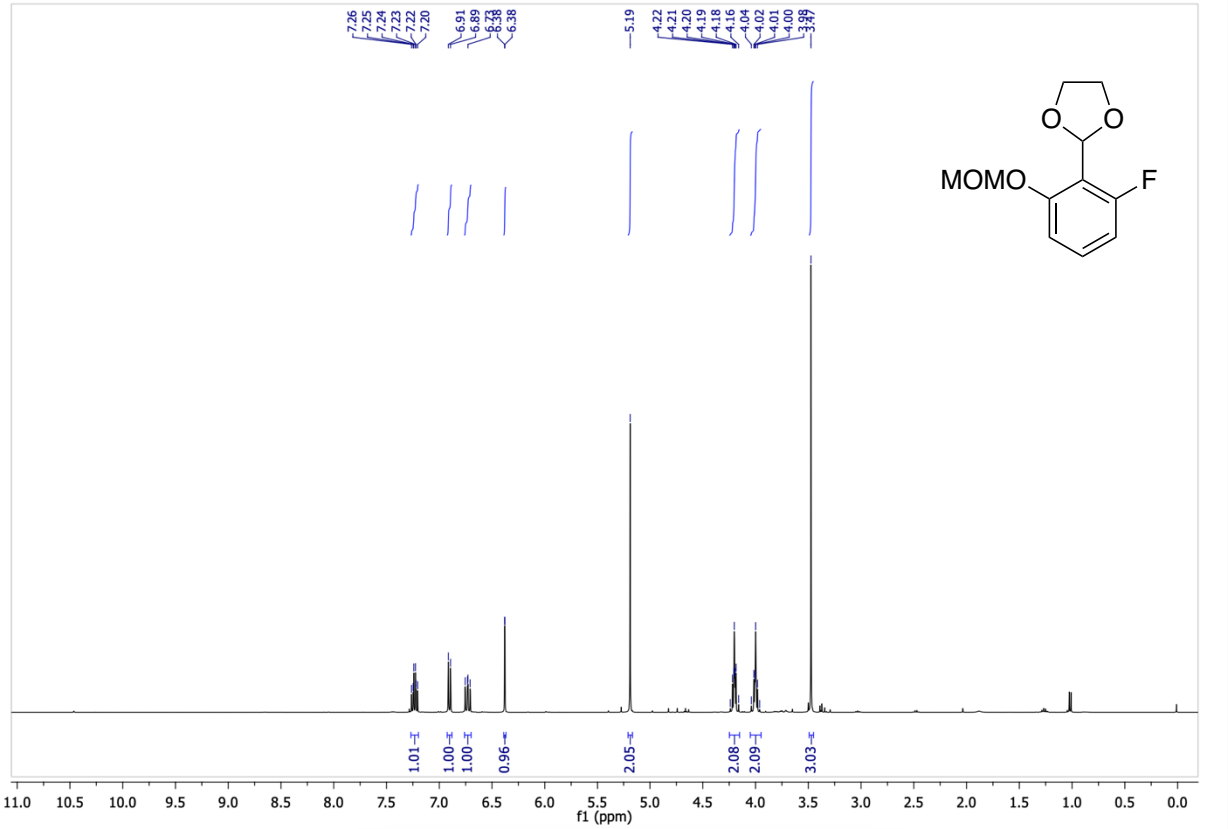
4a


7.58
7.58
7.56
7.56
7.35
7.33
7.31
7.31
7.28
7.16
7.14
7.08
7.06
7.04
6.22
5.26
4.18
4.17
4.17
4.16
4.16
4.15
4.12
4.10
4.08
4.07
4.06
4.06
4.04
3.52
MOMO
0.96
1.00
0.91
0.96
0.80
1.94
3.87
2.91
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

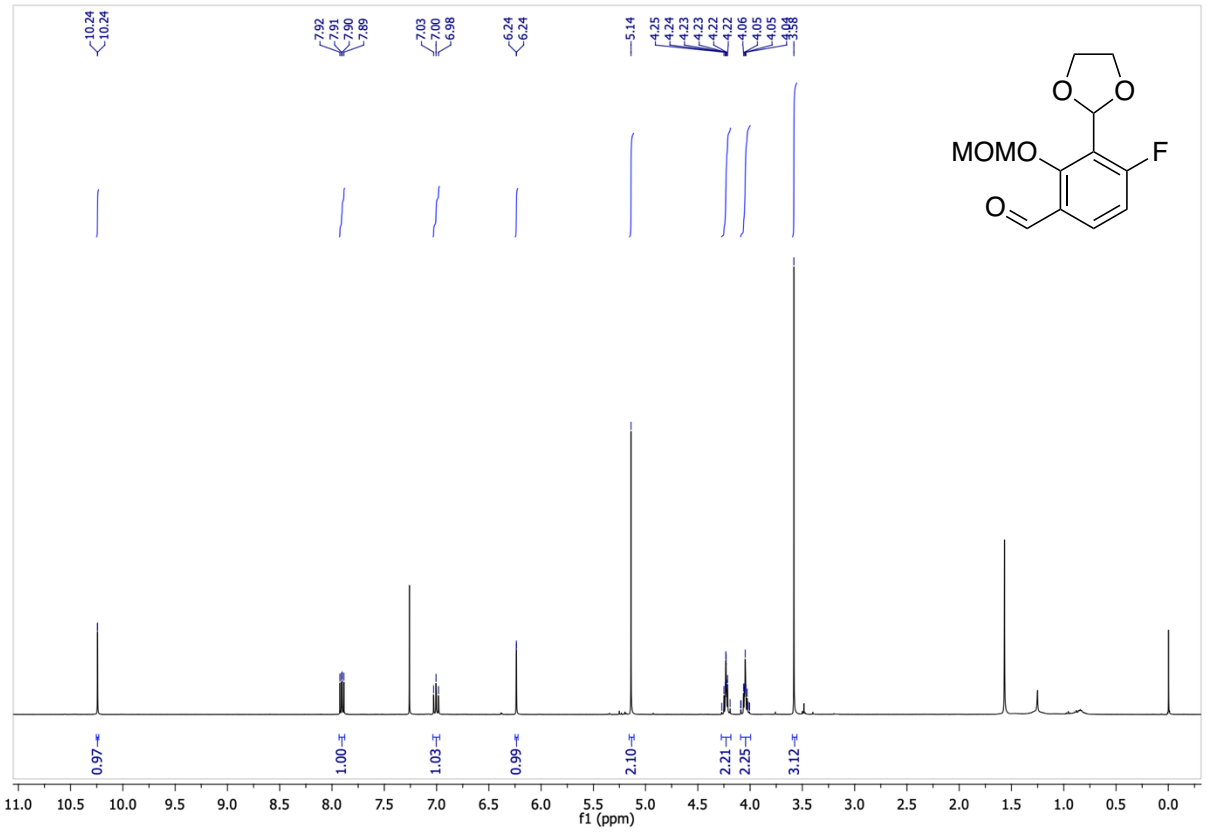
4b


10.32
7.86
7.86
7.84
7.84
7.82
7.82
7.80
7.80
7.29
7.27
7.25
6.09
5.13
4.15
4.13
4.12
4.12
4.11
4.09
4.06
4.04
4.03
4.02
4.02
4.00
3.55
O
O
MOMO
O
0.86
0.94
0.94
1.00
0.89
1.93
4.05
2.93
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

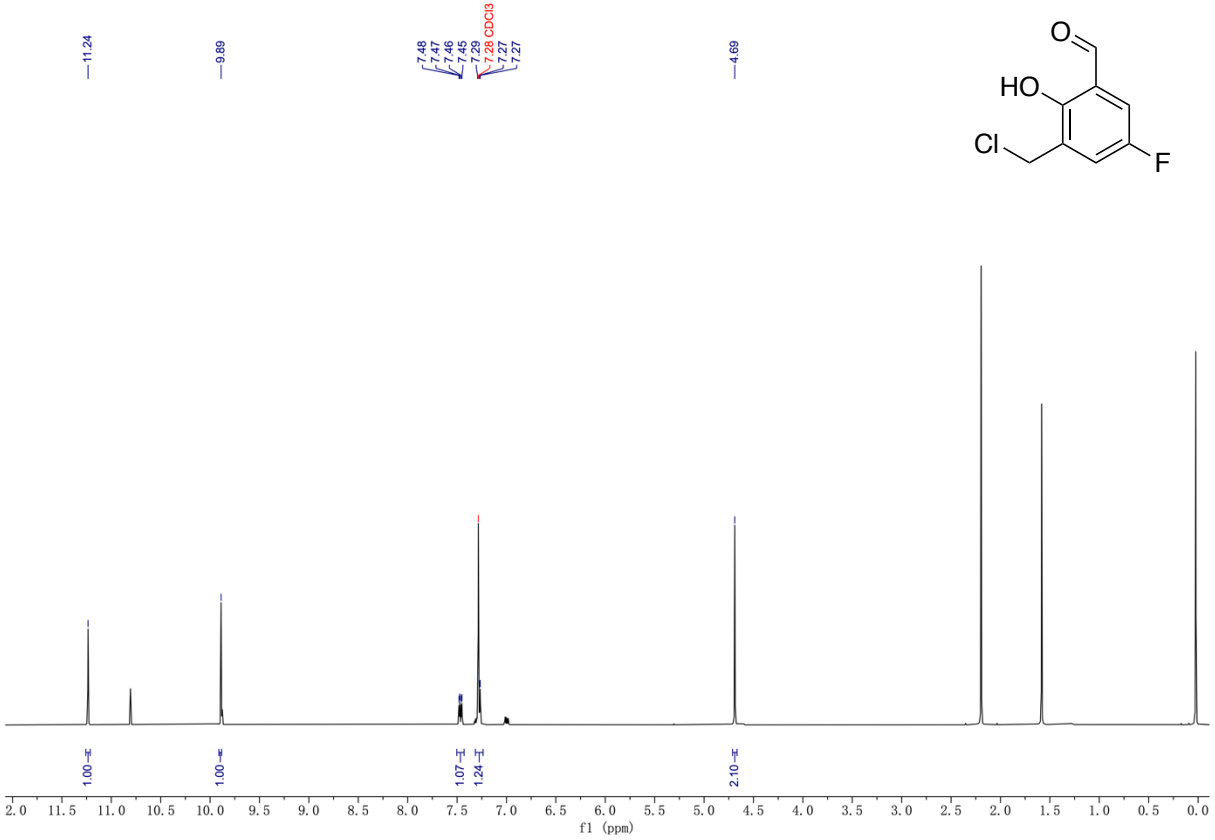
6a


7.26
7.25
7.24
7.23
7.22
7.20
6.91
6.89
6.73
6.38
6.38
5.19
4.22
4.21
4.20
4.19
4.18
4.16
4.04
4.02
4.01
4.00
3.99
3.97
O
O
MOMO
F
1.01
1.00
1.00
0.96
2.05
2.08
2.09
3.03
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

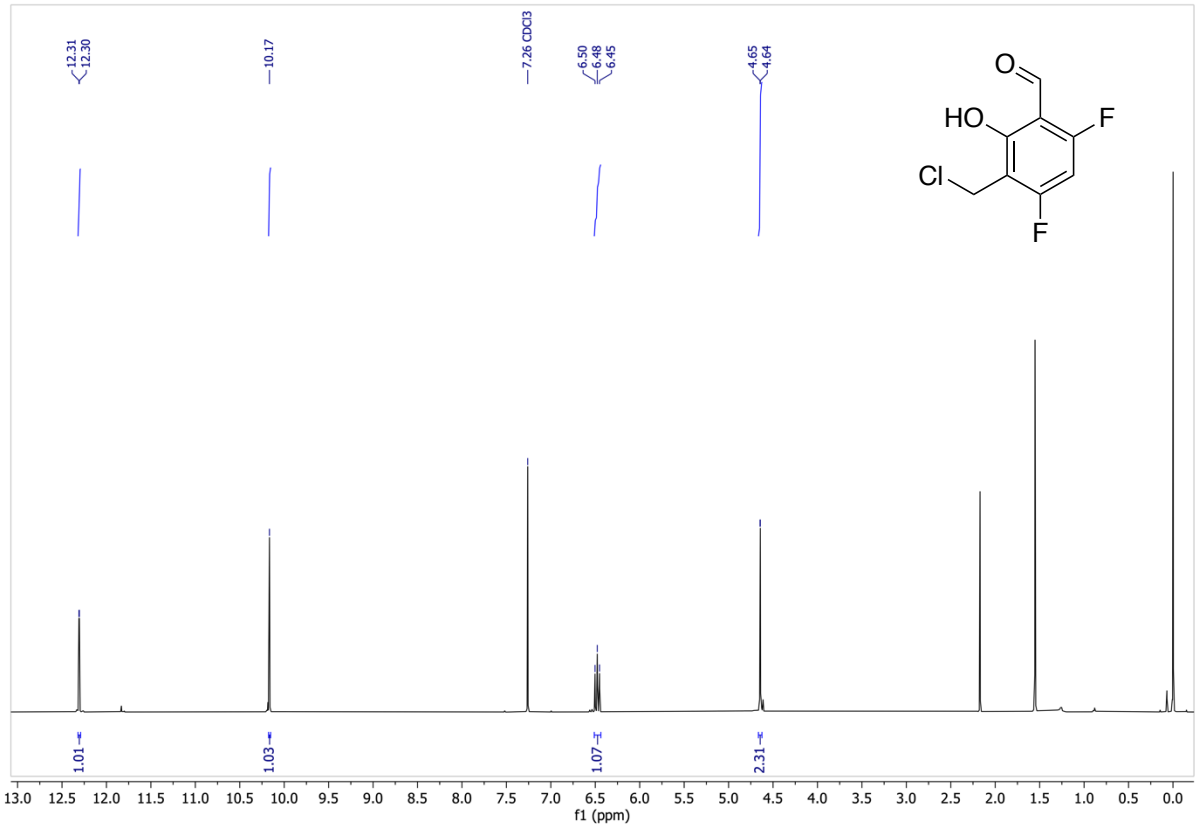
6b


10.24
10.24
7.92
7.91
7.90
7.89
7.03
7.00
6.98
6.24
6.24
5.14
4.25
4.24
4.23
4.23
4.22
4.22
4.06
4.05
4.05
4.04
3.58
MOMO
F
O
O
O
0.97
1.00
1.03
0.99
2.10
2.21
2.25
3.12
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

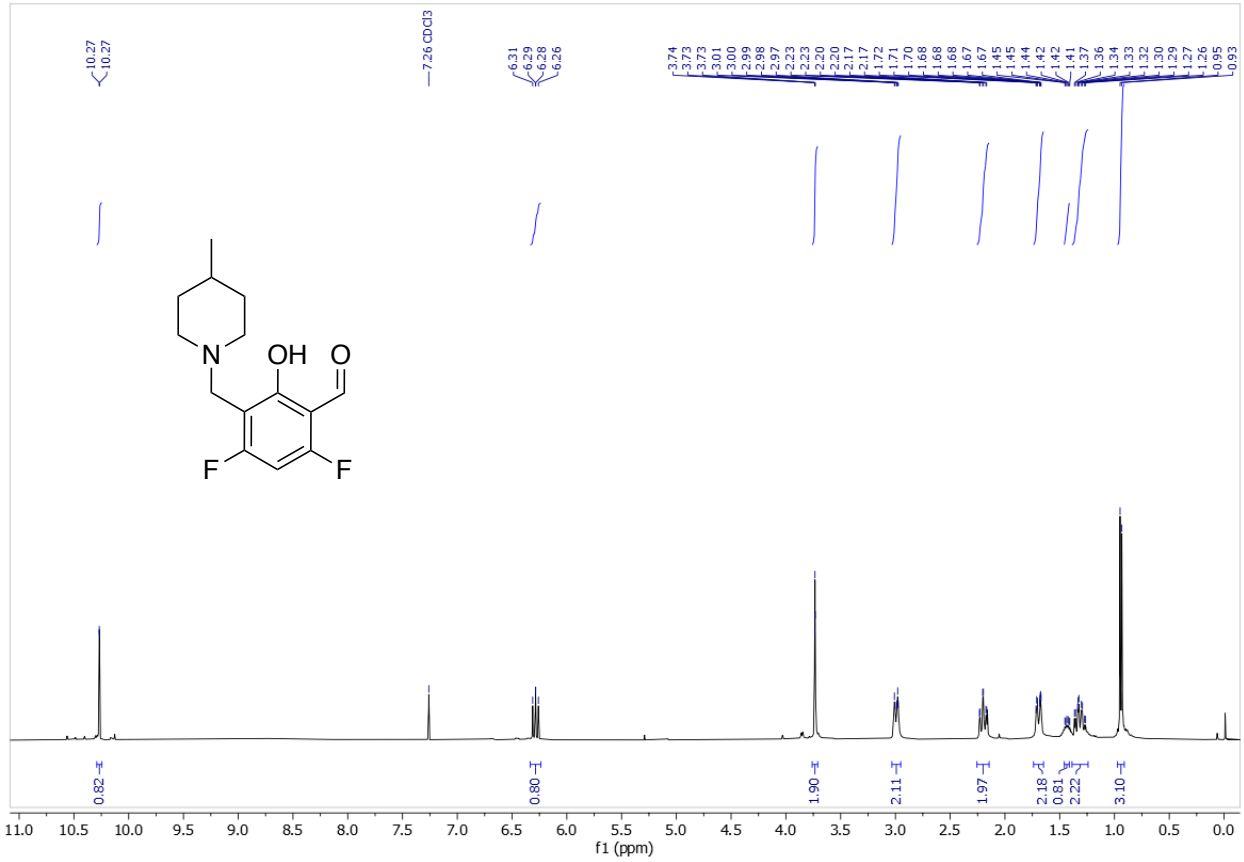
5c


11.24
9.89
7.48
7.47
7.46
7.45
7.29
7.26 CDCl3
7.27
7.27
4.69
O
HO
Cl
F
1.00
1.00
1.07
1.24
2.10
f1 (ppm)

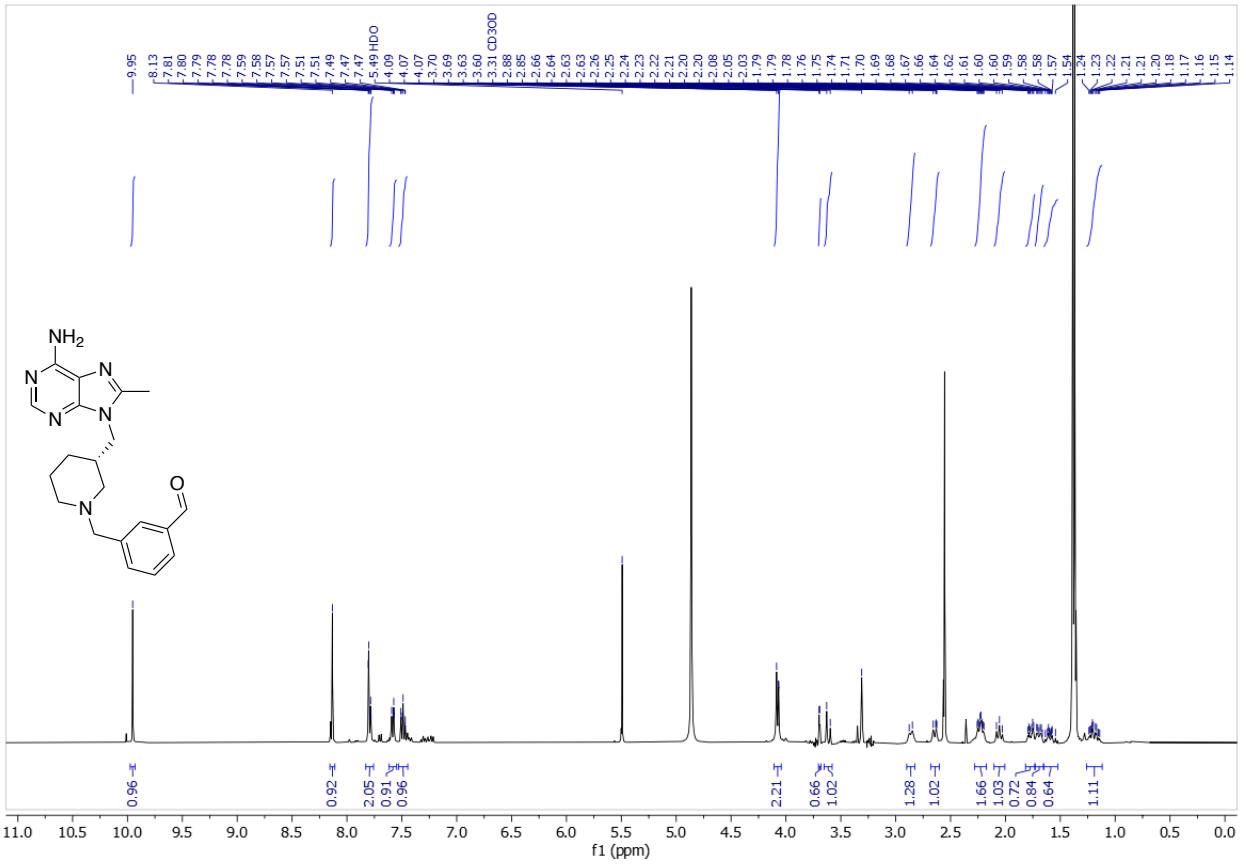
7c


12.31
12.30
10.17
7.26 CDCl3
6.50
6.48
6.45
4.65
4.64
O
HO
F
Cl
F
1.01
1.03
1.07
2.31
13.0 12.5 12.0 11.5 11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

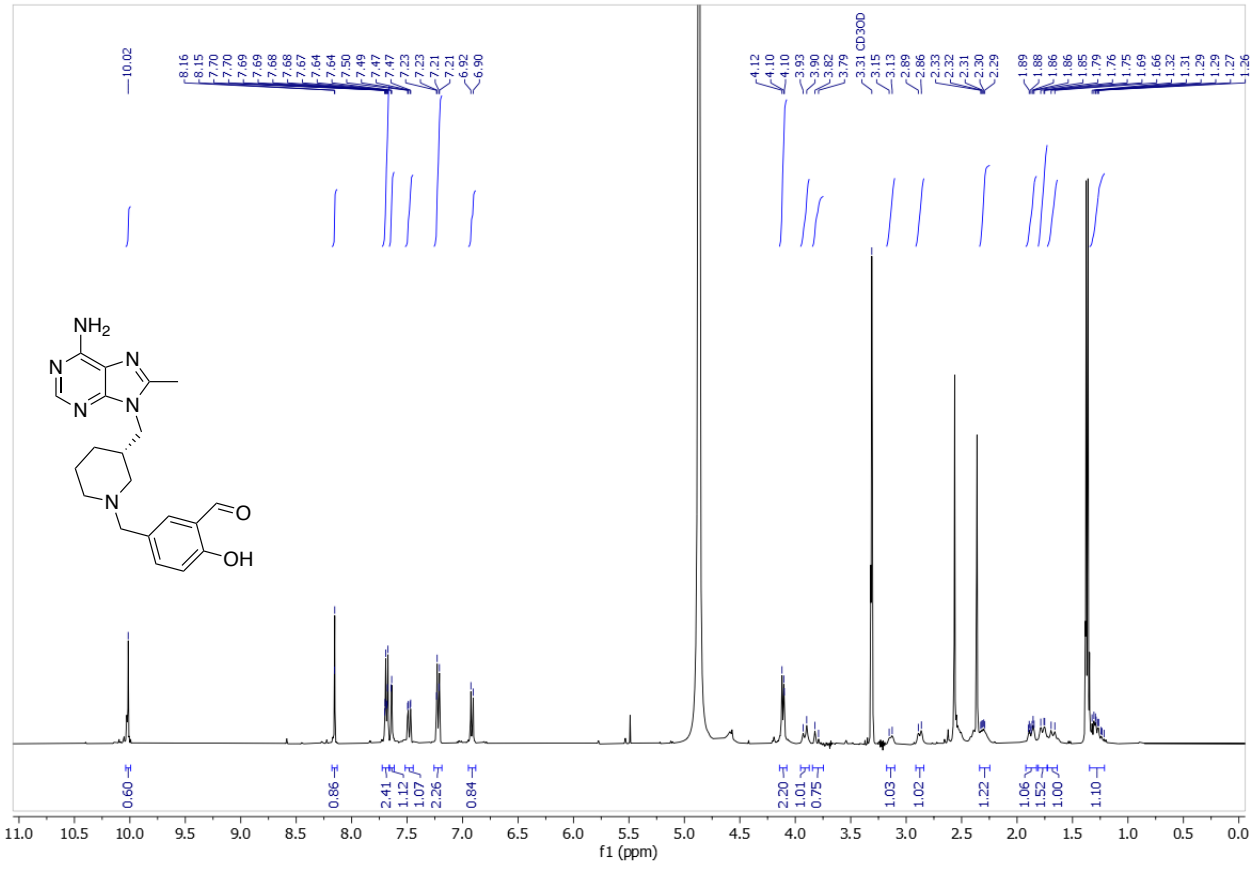
JW-2-85


10.27
10.27
7.26 CDCl3
6.31
6.29
6.28
6.26
3.74
3.73
3.73
3.01
3.00
2.99
2.98
2.97
2.23
2.20
2.20
2.17
1.72
1.71
1.70
1.68
1.68
1.68
1.67
1.67
1.45
1.45
1.44
1.42
1.42
1.41
1.37
1.36
1.34
1.33
1.32
1.30
1.29
1.27
1.26
0.95
0.93
N
OH
O
F
F
0.82
0.80
1.90
2.11
1.97
2.18
0.81
2.22
3.10
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

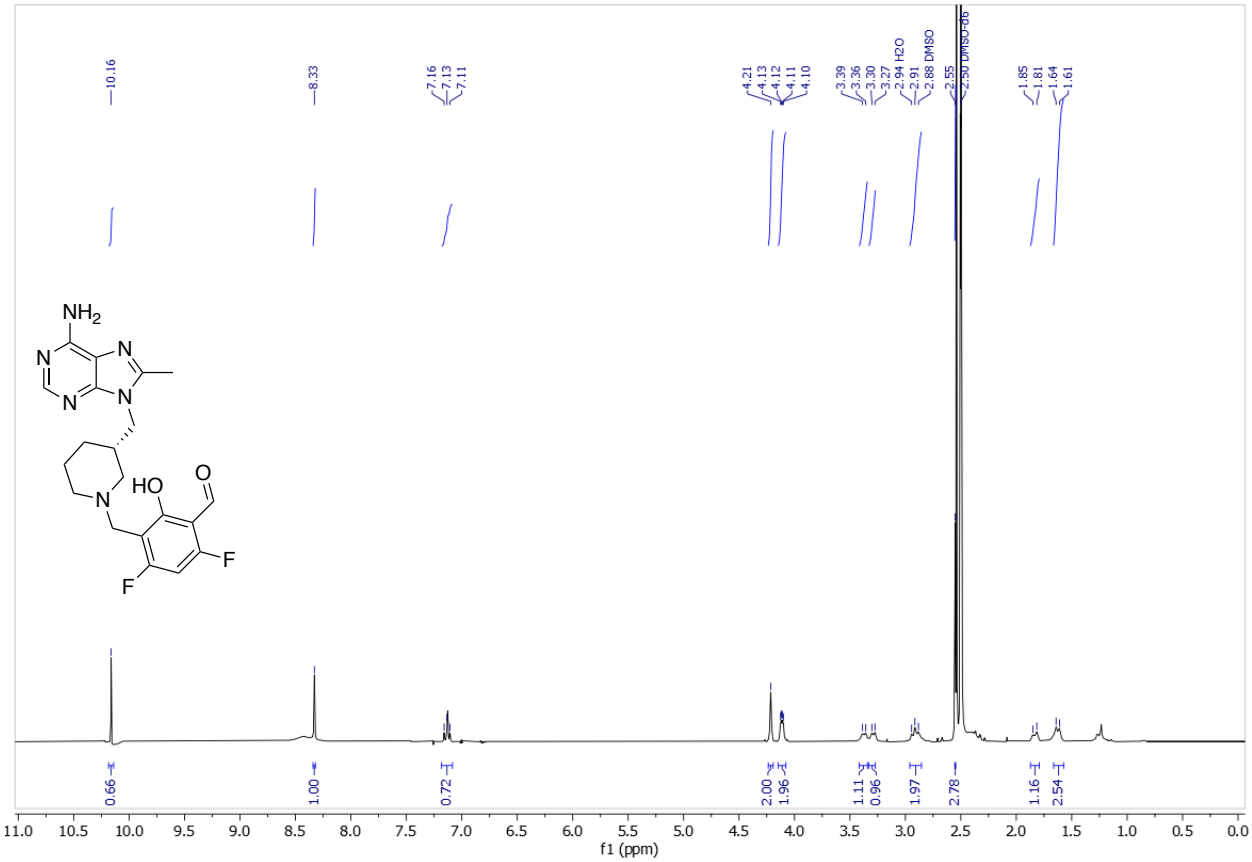
JW-2-127


9.95
8.13
7.81
7.80
7.79
7.78
7.78
7.59
7.58
7.57
7.57
7.51
7.51
7.49
7.47
7.47
5.49 HDO
4.09
4.07
4.07
3.70
3.69
3.63
3.60
3.31 CD3OD
2.88
2.85
2.66
2.64
2.63
2.63
2.26
2.25
2.24
2.23
2.22
2.21
2.20
2.20
2.08
2.05
2.03
1.79
1.79
1.78
1.76
1.75
1.74
1.71
1.70
1.69
1.68
1.67
1.66
1.64
1.62
1.61
1.60
1.60
1.59
1.58
1.58
1.57
1.54
1.24
1.23
1.22
1.21
1.21
1.20
1.18
1.17
1.16
1.15
1.14
NH2
N
N
N
N
O
0.96
0.92
2.05
0.91
0.96
2.21
0.66
1.02
1.28
1.02
1.66
1.03
0.72
0.84
0.64
1.11
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

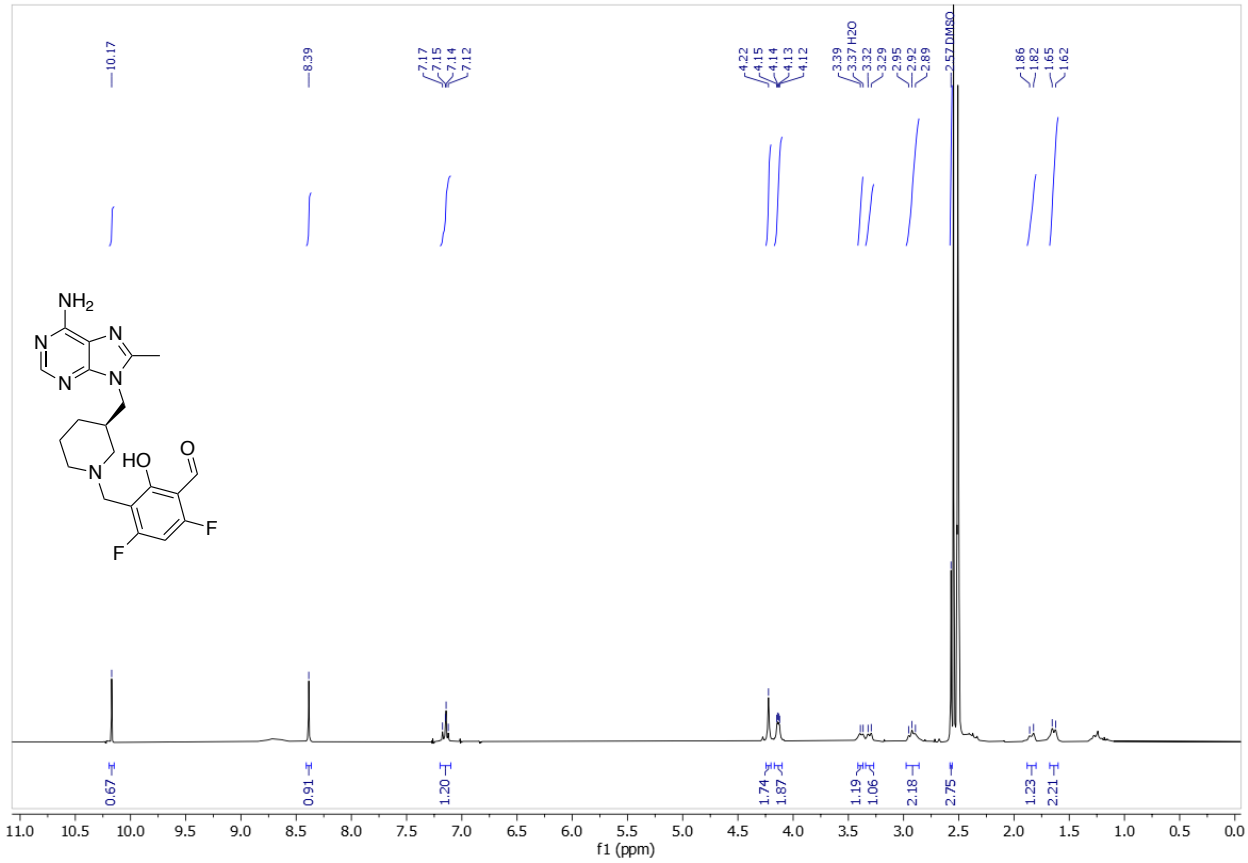
JW-2-129


10.02
8.16
8.15
7.70
7.70
7.69
7.69
7.68
7.68
7.67
7.64
7.64
7.50
7.49
7.47
7.47
7.23
7.23
7.21
7.21
6.92
6.90
4.12
4.10
4.10
3.93
3.90
3.82
3.79
3.31 CD3OD
3.15
3.13
2.89
2.86
2.33
2.32
2.31
2.30
2.29
1.89
1.88
1.86
1.86
1.85
1.79
1.76
1.75
1.69
1.66
1.32
1.31
1.29
1.29
1.27
1.26
NH2
N
N
N
N
N
O
OH
0.60
0.86
2.41
1.12
1.07
2.26
0.84
2.20
1.01
0.75
1.03
1.02
1.22
1.06
1.52
1.00
1.10
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

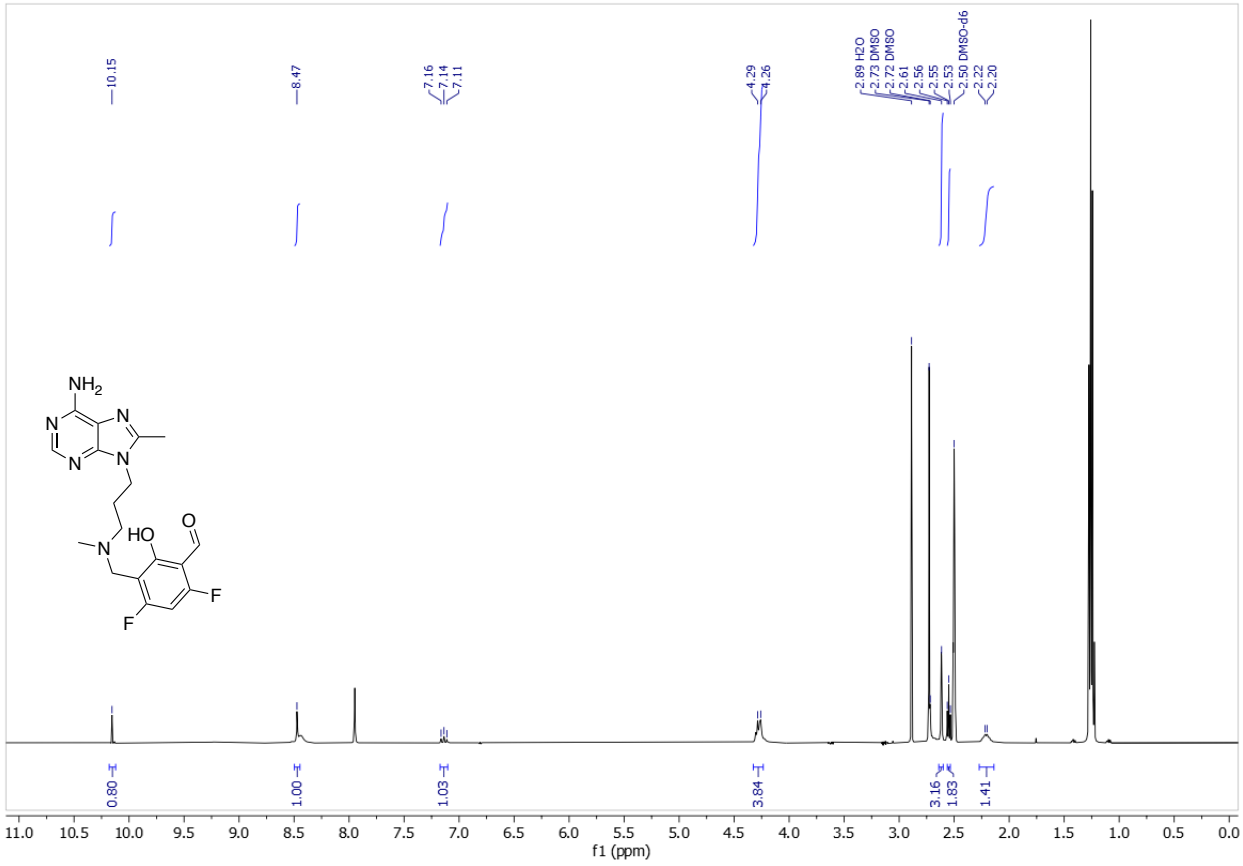
JW-2-128


10.16
8.33
7.16
7.13
7.11
4.21
4.13
4.12
4.11
4.10
3.39
3.36
3.30
3.27
2.94 H2O
2.91
2.88 DMSO
2.55
2.50 DMSO-d6
1.85
1.81
1.64
1.61
NH2
N
N
N
N
HO
O
N
F
F
0.66
1.00
0.72
2.00
1.96
1.11
0.96
1.97
2.78
1.16
2.54
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

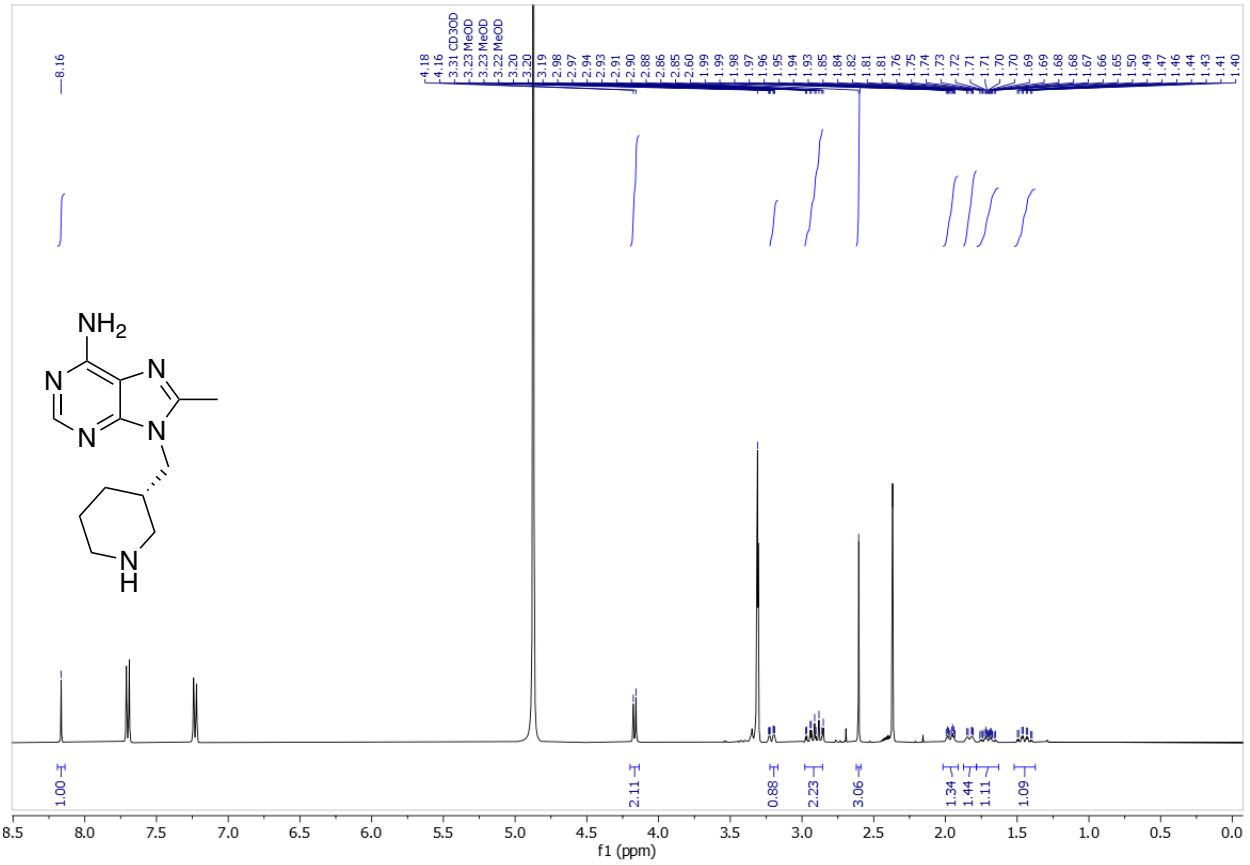
JW-2-159


10.17
8.39
7.17
7.15
7.14
7.12
4.22
4.15
4.14
4.13
4.12
3.39
3.37 H2O
3.32
3.29
2.95
2.92
2.89
2.57 DMSO
1.86
1.82
1.65
1.62
NH2
N
N
N
N
HO
O
F
F
0.67
0.91
1.20
1.74
1.87
1.19
1.06
2.18
2.75
1.23
2.21
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

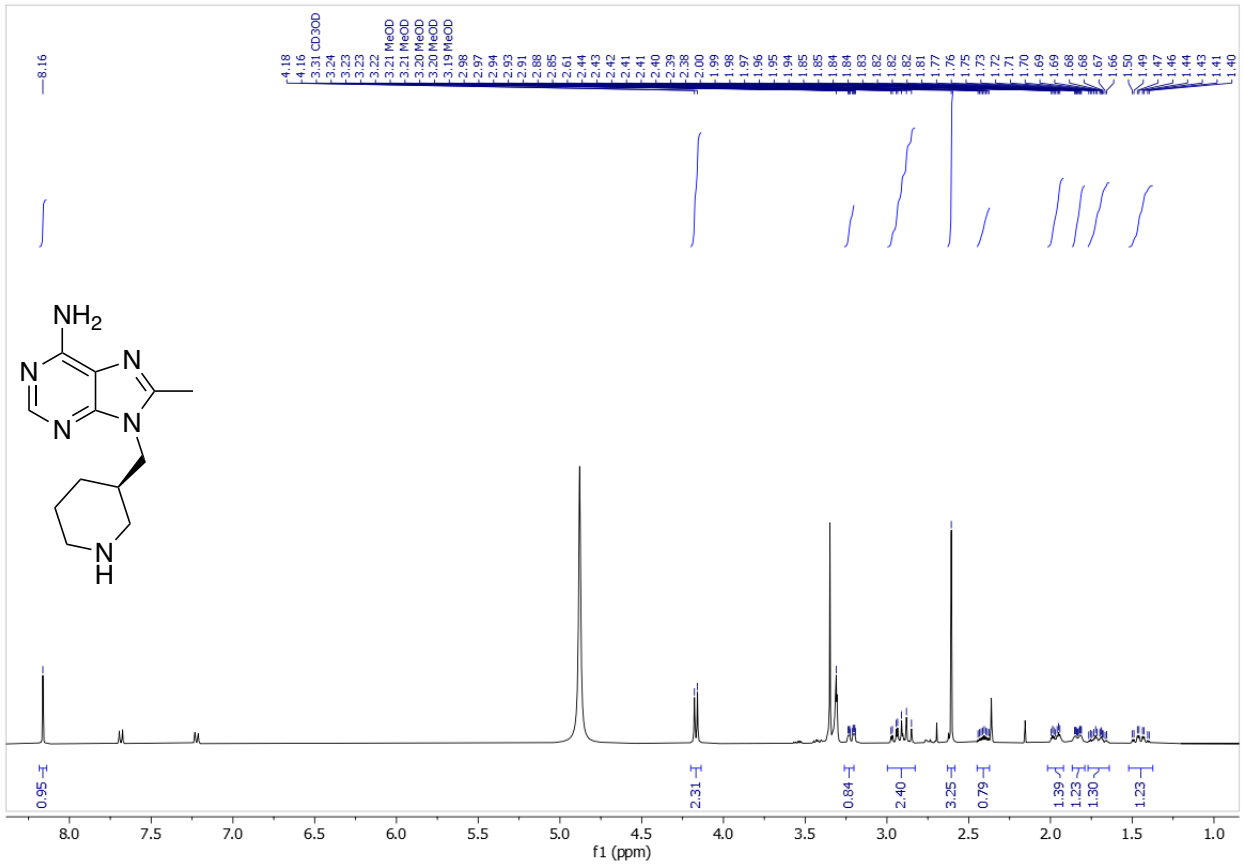
JW-2-162


10.15
8.47
7.16
7.14
7.11
4.29
4.26
2.89 H2O
2.73 DMSO
2.72 DMSO
2.61
2.56
2.55
2.53
2.50 DMSO-d6
2.22
2.20
NH2
N
N
N
N
HO
O
N
F
F
0.80
1.00
1.03
3.84
3.16
1.83
1.41
11.0 10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

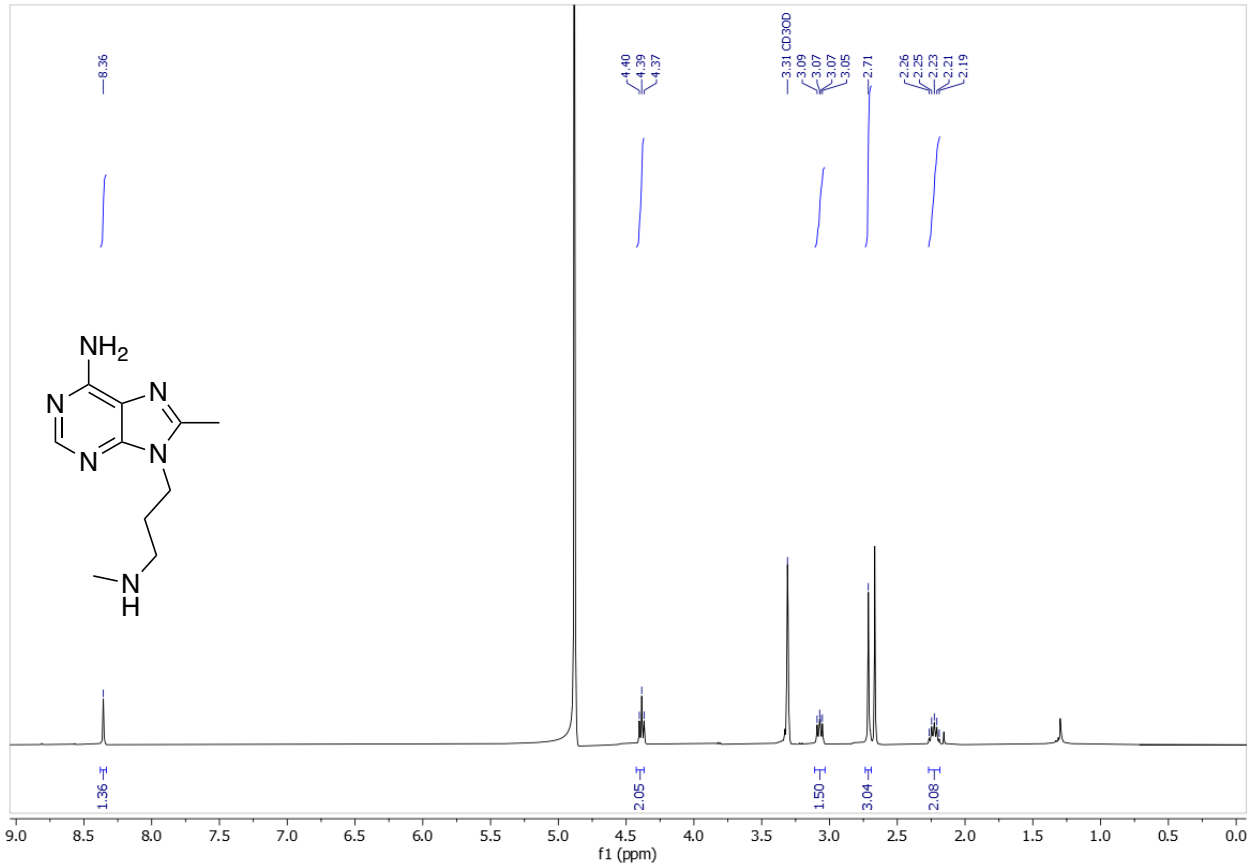
15


8.16
4.18
4.16
3.31 CD3OD
3.23 MeOD
3.23 MeOD
3.22 MeOD
3.20
3.20
3.19
2.98
2.97
2.94
2.93
2.91
2.90
2.88
2.86
2.85
2.60
1.99
1.99
1.98
1.97
1.96
1.95
1.94
1.93
1.85
1.84
1.82
1.81
1.81
1.76
1.75
1.74
1.73
1.72
1.71
1.71
1.70
1.70
1.69
1.69
1.68
1.68
1.67
1.66
1.65
1.50
1.49
1.47
1.46
1.44
1.43
1.41
1.40
NH2
N
N
N
N
N
H
1.00
2.11
0.88
2.23
3.06
1.34
1.44
1.11
1.09
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

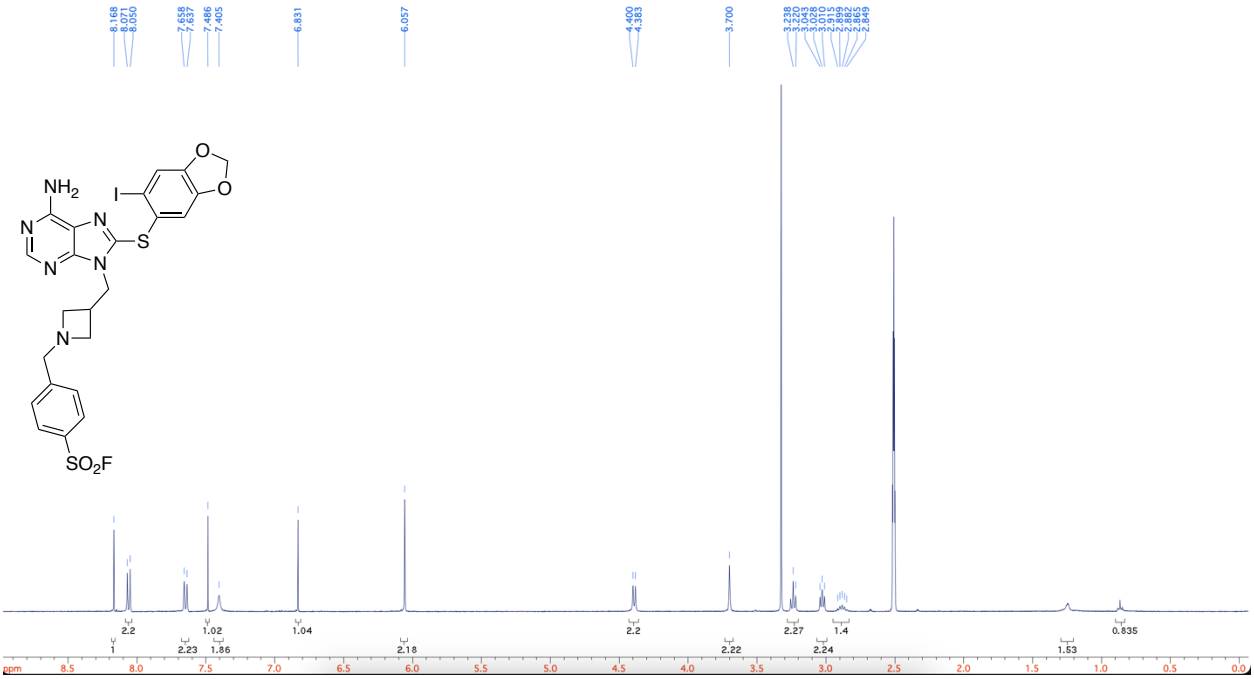
16



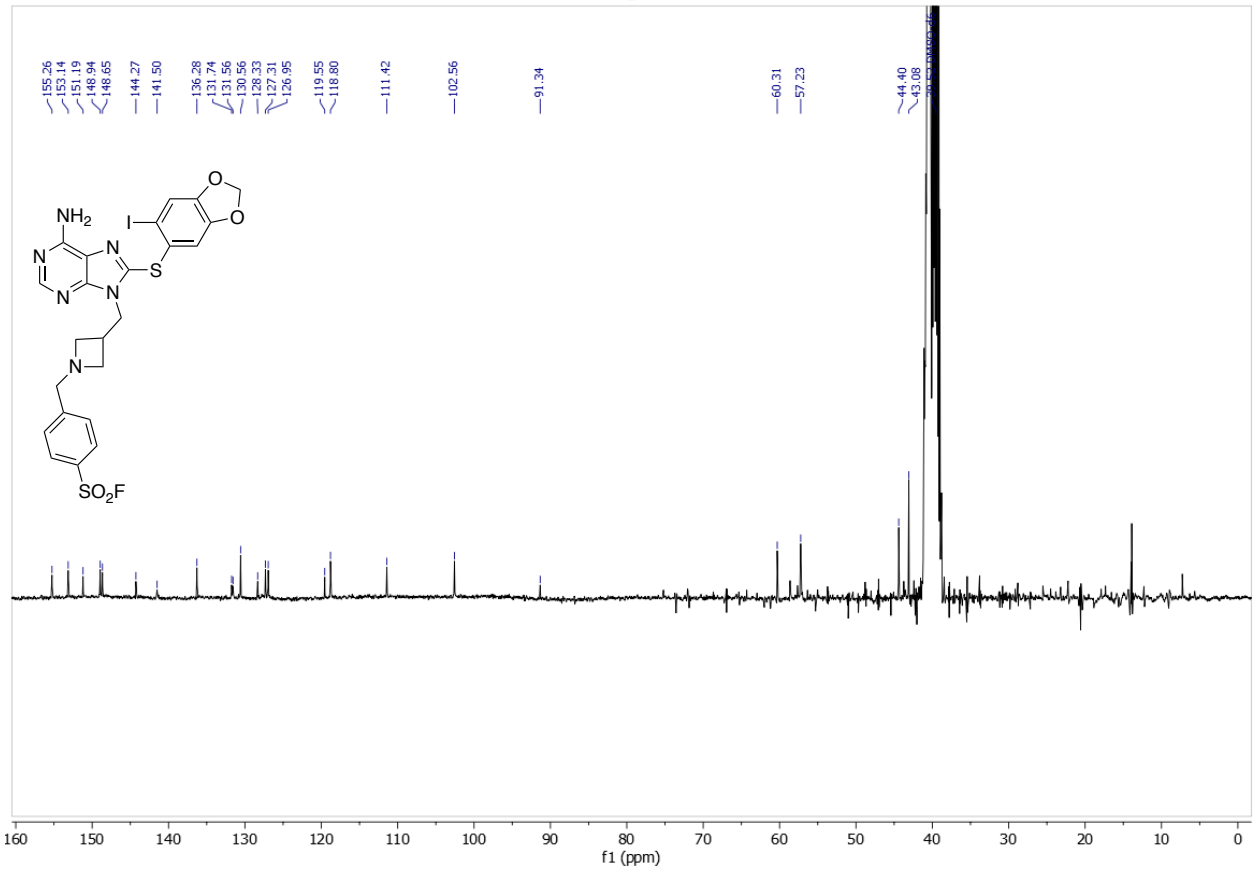
17


8.36
4.40
4.39
4.37
3.31 CD3OD
3.09
3.07
3.07
3.05
2.71
2.26
2.25
2.23
2.21
2.19
NH2
N
N
N
N
N
H
1.36
2.05
1.50
3.04
2.08
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

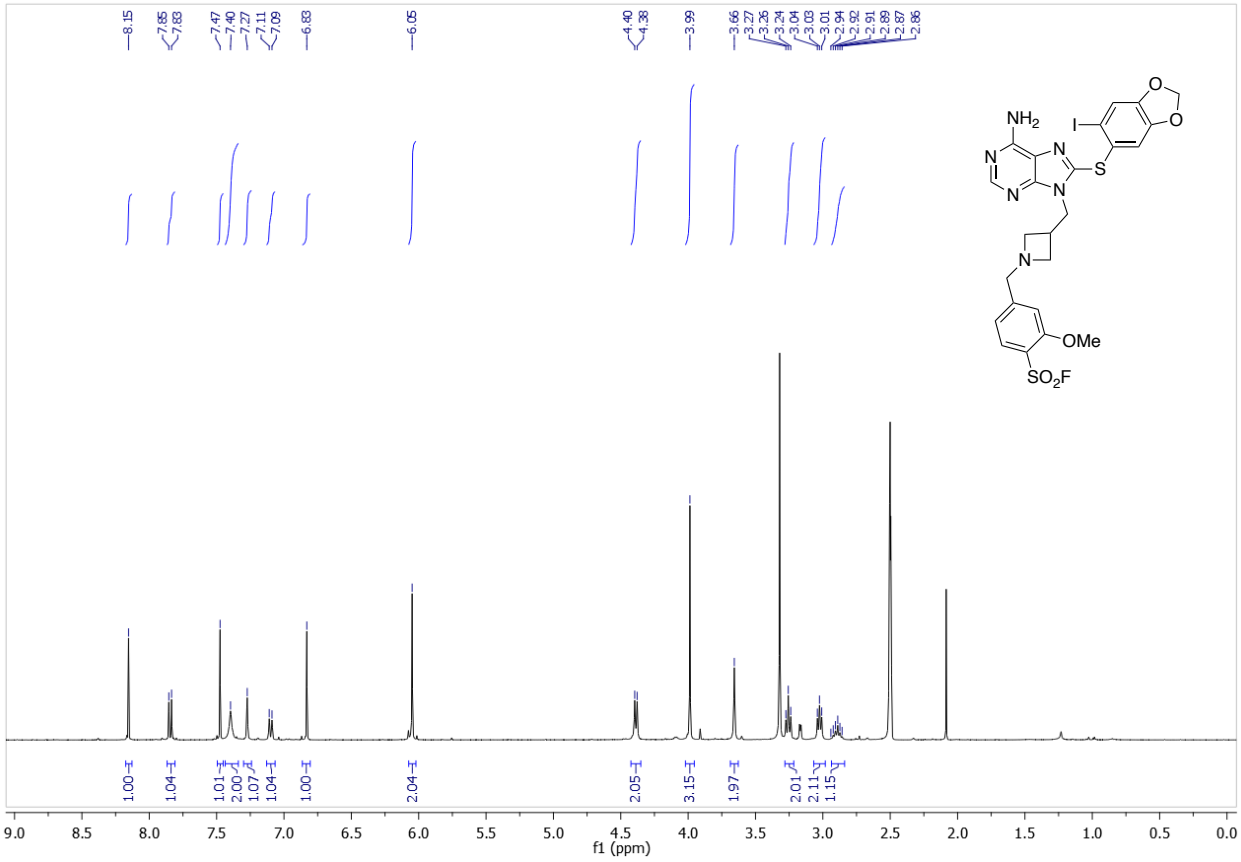
AC45


8.168
8.071
8.050
7.658
7.637
7.486
7.405
6.831
6.057
4.400
4.383
3.700
3.238
3.220
3.043
3.028
3.010
2.911
2.889
2.882
2.865
2.849
NH2
I
O
O
N
N
N
N
S
N
SO2F
2.2
1.02
1.04
2.2
2.27
1.4
0.835
1
2.23
1.86
2.18
2.22
2.24
1.53
ppm
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0

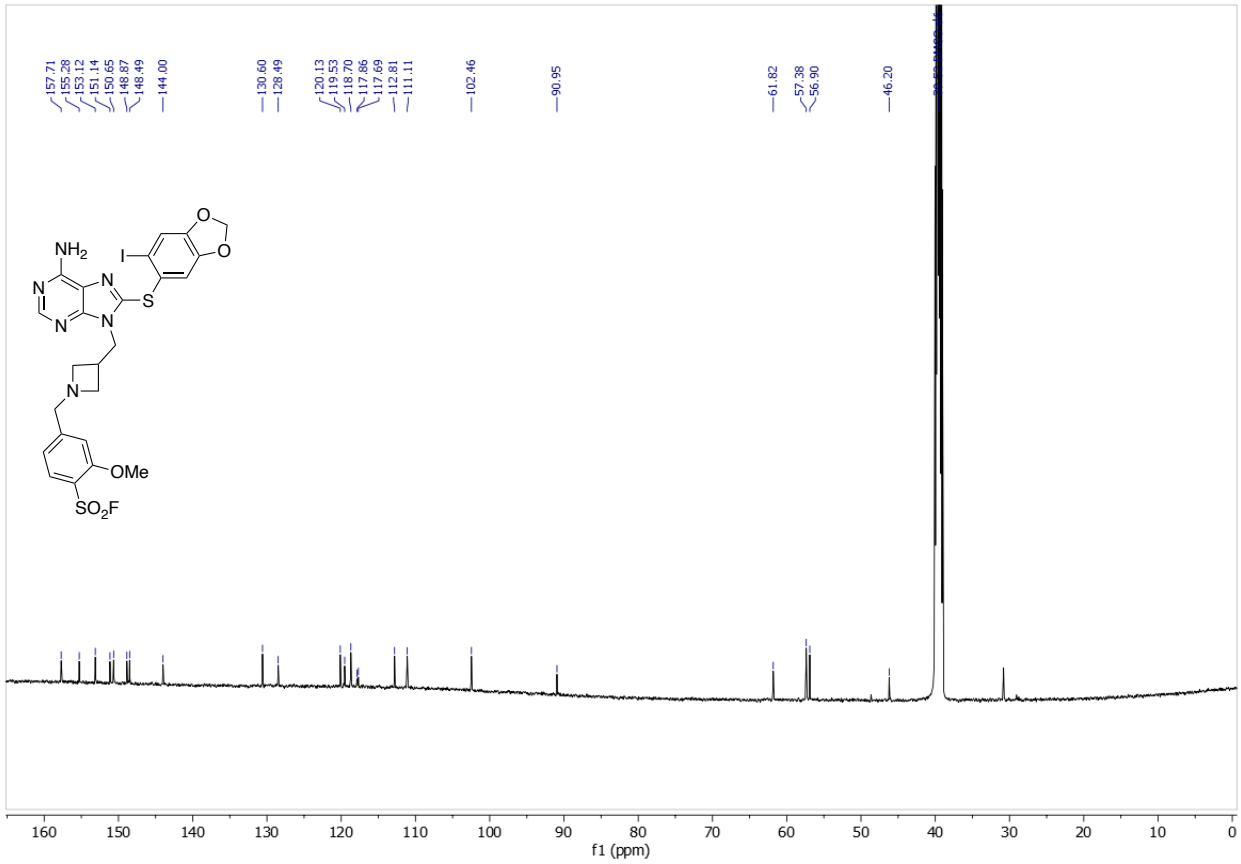
AC45


155.26
153.14
151.19
148.94
148.65
144.27
141.50
136.28
131.74
131.56
130.56
128.33
127.31
126.95
119.55
118.80
111.42
102.56
91.34
60.31
57.23
44.40
43.08
NH2
N
N
N
N
S
I
O
O
N
SO2F
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

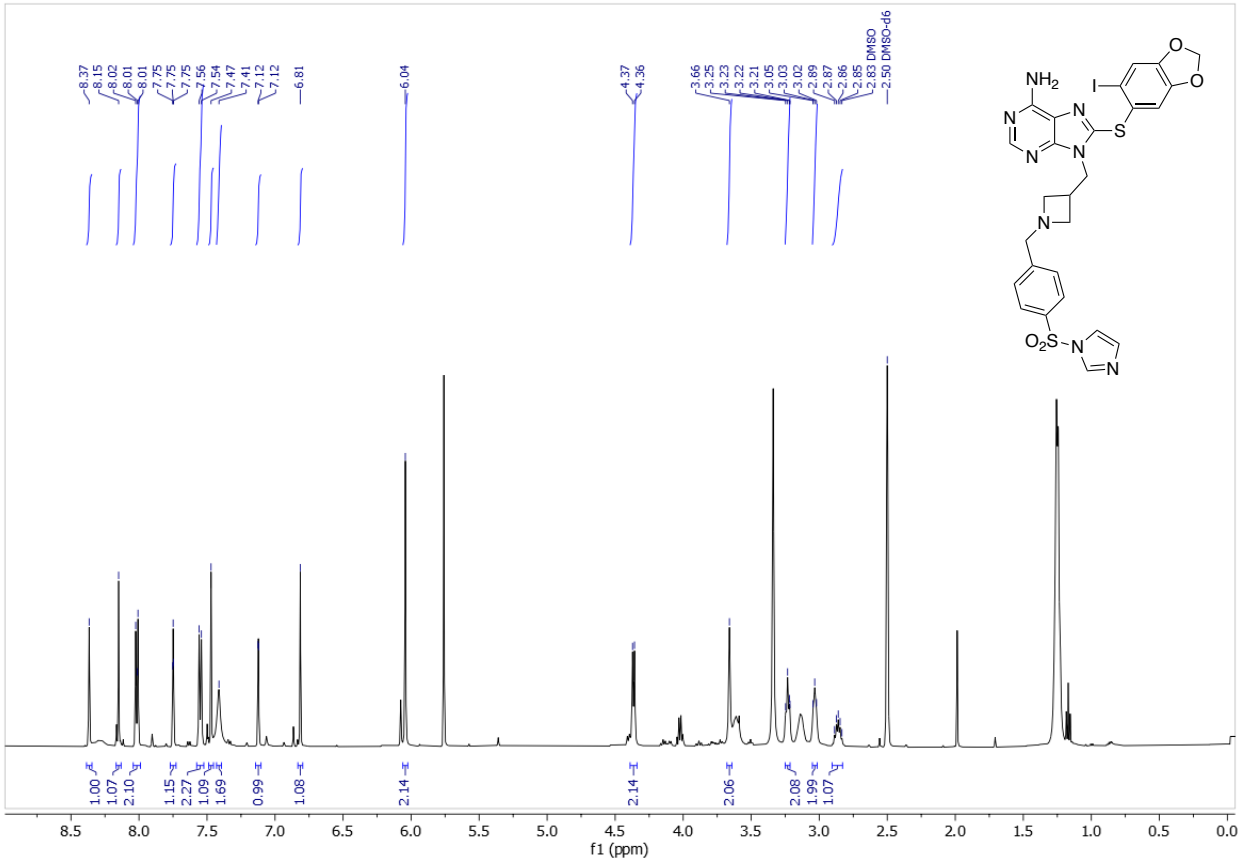
JW-2-63


8.15
7.85
7.83
7.47
7.40
7.27
7.11
7.09
6.83
6.05
4.40
4.38
3.99
3.66
3.27
3.26
3.24
3.04
3.03
3.01
2.94
2.92
2.91
2.89
2.87
2.86
NH2
I
O
O
N
N
N
N
S
N
OMe
SO2F
1.00
1.04
1.01
2.00
1.07
1.04
1.00
2.04
2.05
3.15
1.97
2.01
2.11
1.15
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

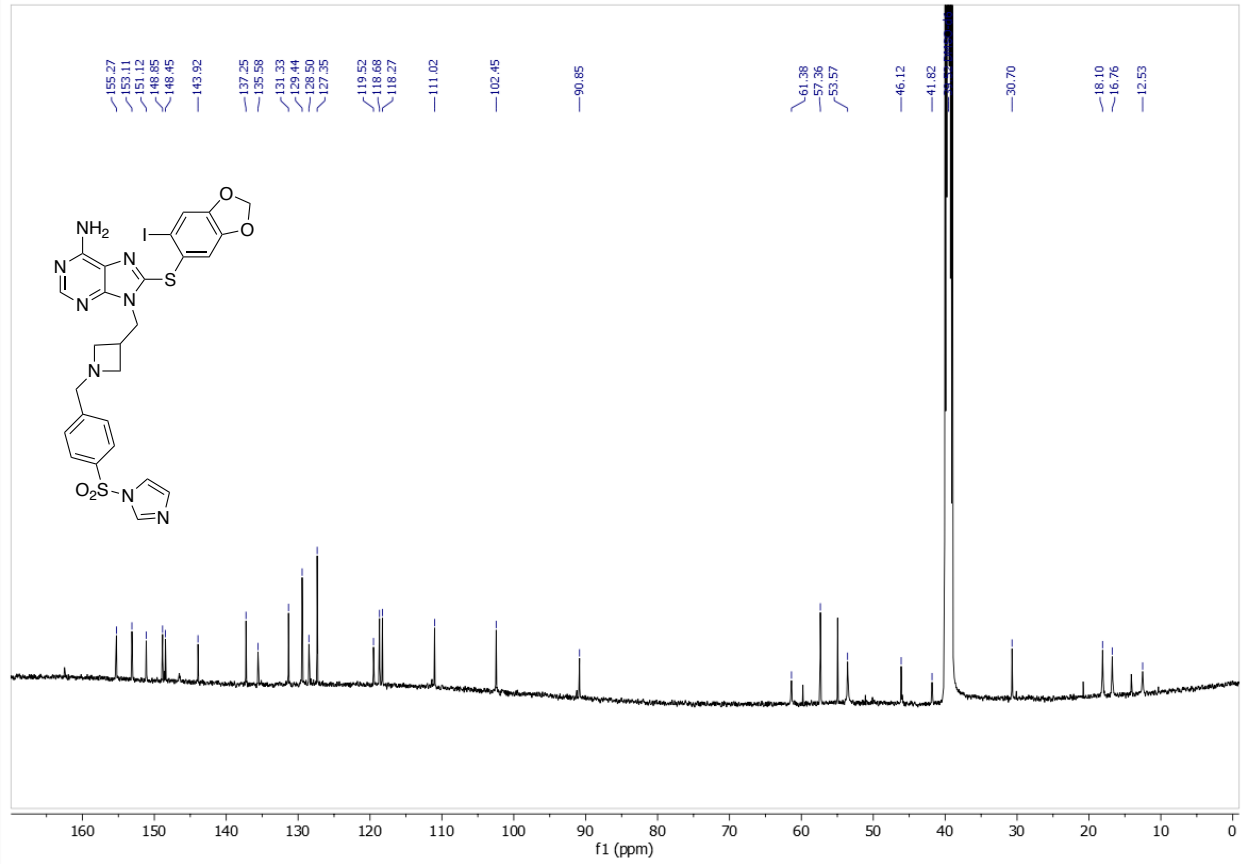
JW-2-63


157.71
155.28
153.12
151.14
150.65
148.87
148.49
144.00
130.60
128.49
120.13
119.53
118.70
117.86
117.69
112.81
111.11
102.46
90.95
61.82
57.38
56.90
46.20
NH2
I
O
O
N
N
N
N
S
N
OMe
SO2F
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

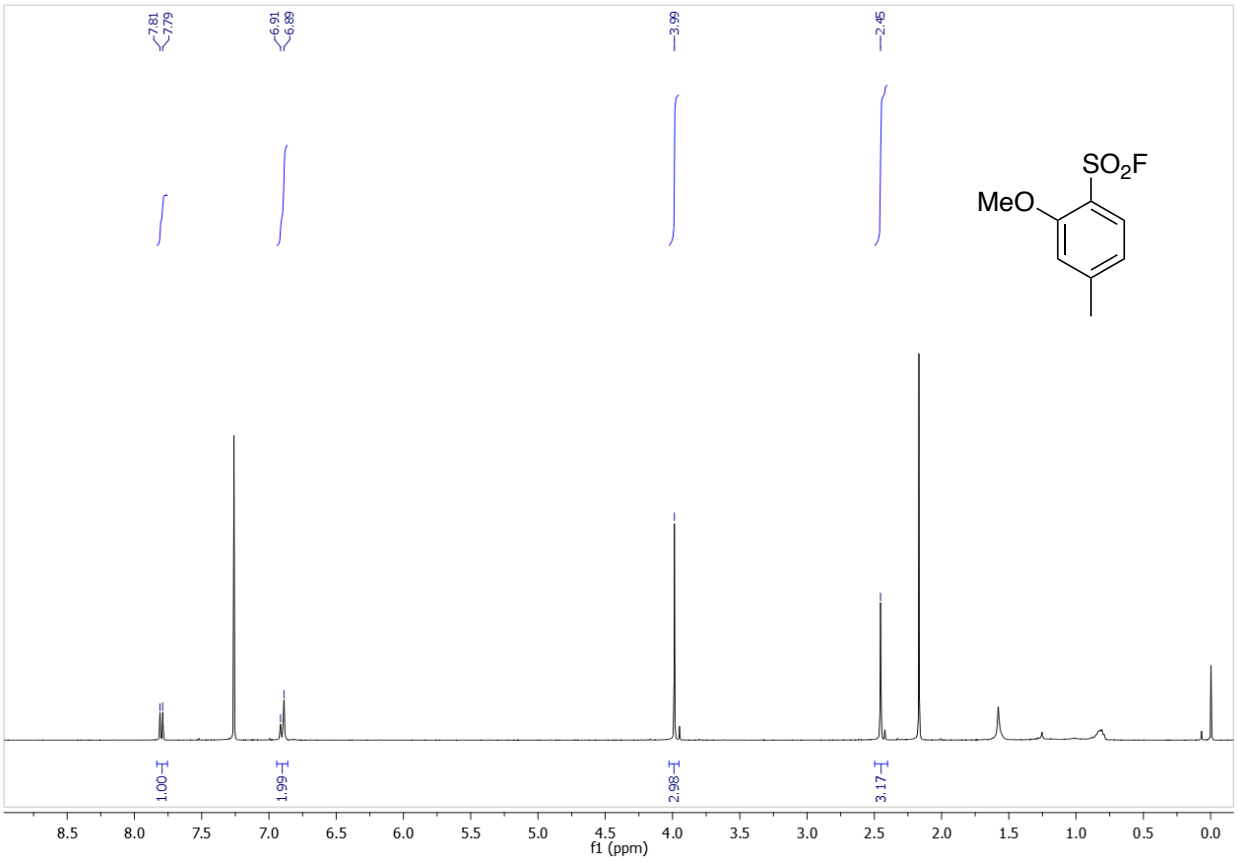
JW-2-183


8.37
8.15
8.02
8.01
8.01
7.75
7.75
7.75
7.56
7.54
7.47
7.41
7.12
7.12
6.81
6.04
4.37
4.36
3.66
3.25
3.23
3.22
3.21
3.05
3.03
3.02
2.89
2.87
2.86
2.85
2.83 DMSO
2.50 DMSO-d6
1.00
1.07
2.10
1.15
2.27
1.09
1.69
0.99
1.08
2.14
2.14
2.06
2.08
1.99
1.07
f1 (ppm)

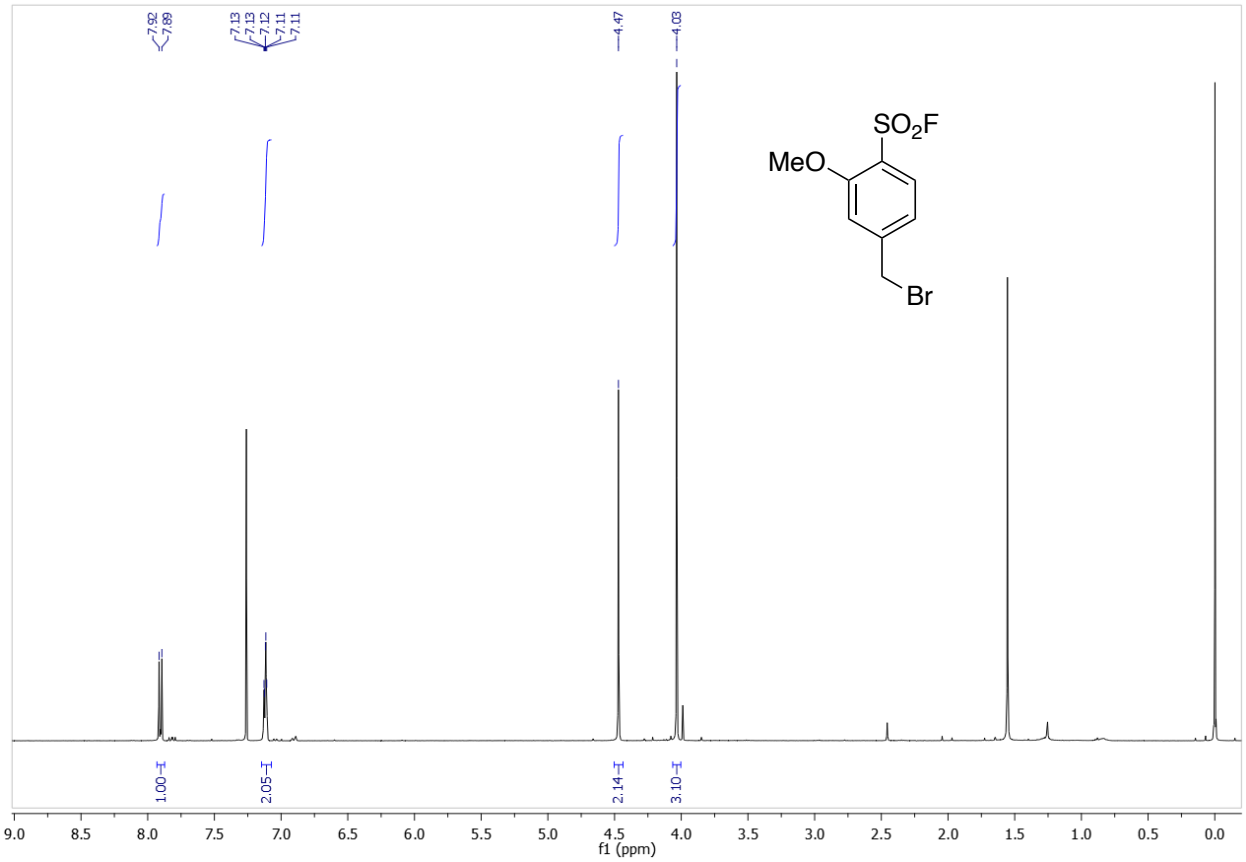
JW-2-183


155.27
153.11
151.12
148.85
148.45
143.92
137.25
135.58
131.33
129.44
128.50
127.35
119.52
118.68
118.27
111.02
102.45
90.85
61.38
57.36
53.57
46.12
41.82
30.70
18.10
16.76
12.53
NH2
I
O
O
N
N
N
N
S
O2S
N
N
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

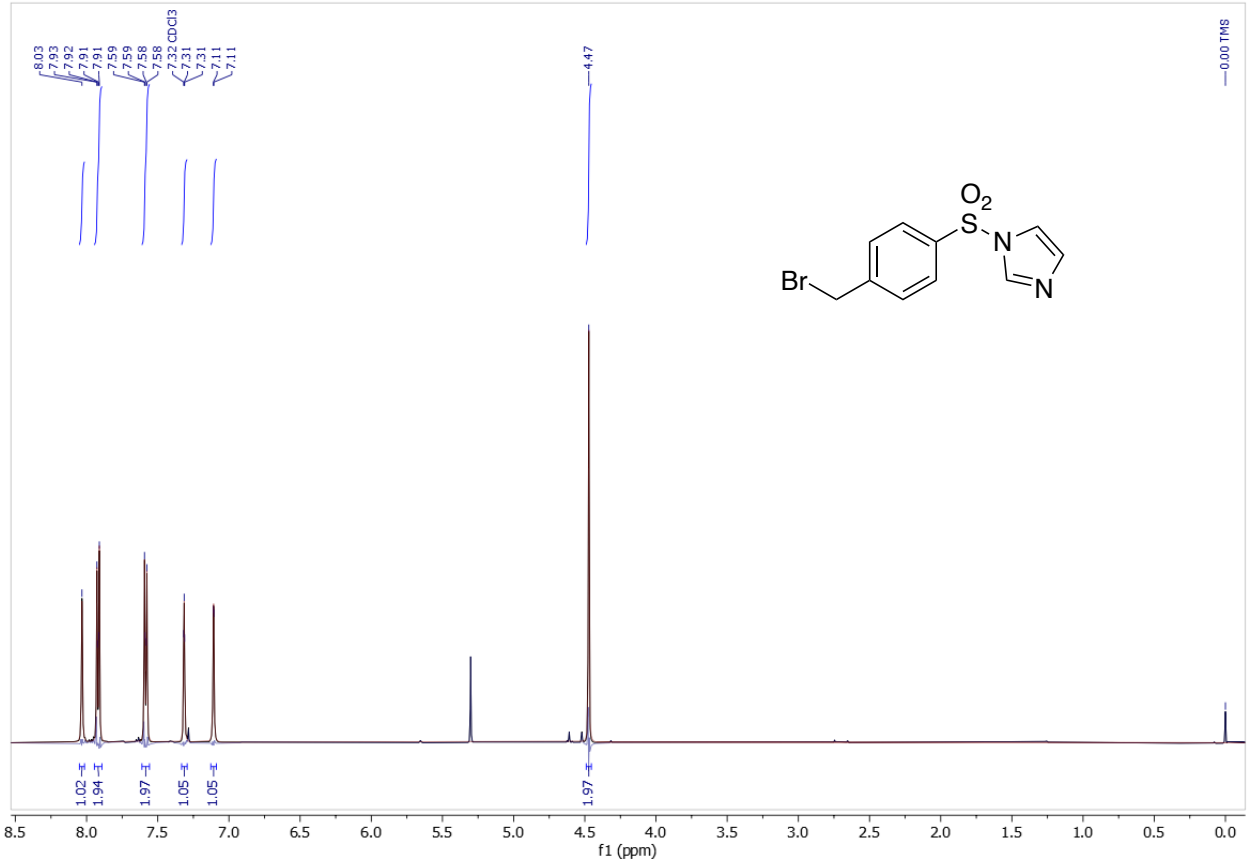
19


7.81
7.79
6.91
6.89
3.99
2.45
SO2F
MeO
1.00
1.99
2.98
3.17
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

20


7.92
7.89
7.13
7.13
7.12
7.11
7.11
4.47
4.03
SO2F
MeO
Br
1.00
2.05
2.14
3.10
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

20


8.03
7.93
7.92
7.91
7.91
7.59
7.59
7.58
7.58
7.32 CDCl3
7.31
7.31
7.11
7.11
4.47
0.00 TMS
O2
S
N
N
Br
1.02
1.94
1.97
1.05
1.05
1.97
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Jacqueline Weaver
AF307A27F4FF45C...
Author Signature

12/12/2024
Date