NORTHWESTERN UNIVERSITY

Inhibitors of the Ubiquitin Ligase Nedd4-1 Discovered by Covalent Fragment Screening

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Chemistry

By

Stefan G. Kathman

EVANSTON, ILLINOIS

December 2016

Abstract

The FDA approvals of afatinib and ibrutinib in 2013 led to a heightened interest in cysteine-reactive covalent inhibitors. However, there are few methods to discover new cysteine-reactive inhibitors for enzymes for which reversible binding scaffolds are not known. To this end, we rationally designed a chemical system to attach a cysteine-reactive electrophile to drug-like fragments without significant alterations in the thiol reactivity of the attached electrophile, ensuring that specific binding and not increased reactivity will produce candidate inhibitors. We applied this method, which we call irreversible tethering, to discover inhibitors of the HECT E3 ubiquitin ligase Nedd4-1, an enzyme with no validated inhibitors. Nedd4-1 has a catalytic cysteine and a non-catalytic surface cysteine, and is implicated in viral budding, cancers, and neurodegenerative diseases. We screened our electrophilic fragment library and discovered two fragments which reacted with Nedd4-1 as determined by mass spectrometry. Surprisingly, we found that these inhibitors did not react with the more reactive catalytic cysteine of Nedd4-1, but the other surface cysteine near the non-covalent ubiquitin binding site. This site binds to ubiquitinated substrates in order to extend the length of the ubiquitin chain. The X-ray crystal structure of the most potent fragment in complex with Nedd4-1 was solved, demonstrating that it forms a stable covalent bond with the ubiquitin-binding site cysteine of Nedd4-1 and has specific interactions with residues around this cysteine. This structure has been used to further optimize the fragment into a more potent inhibitor. Due to their proximity to the non-covalent ubiquitin binding site, our inhibitor reduces the binding affinity of Nedd4-1 for ubiquitin. *In vitro* enzymatic assays have shown that these molecules inhibit Nedd4-1 polyubiquitination processivity and switch it to a distributive mechanism. Click chemistry and in-gel fluorescence with an alkyne tagged analog of this inhibitor have shown that the inhibitor reacts with Nedd4-1

in TC71 cells with good selectivity. However, we have not yet been able to show that the optimized inhibitor impairs Nedd4-1 function in cells, perhaps due to insufficient potency or a limitation in its mechanism of action. Further optimization of the inhibitor could result in a compound to study the diseases in which Nedd4-1 polyubiquitination is implicated.

Acknowledgment

Abbreviations:

DUB: deubiquitinase

ESI-MS: electrospray ionization mass spectrometry

FP: fluorescence polarization

HECT: homologous to E6AP carboxyl-terminus

IGF1: insulin-like growth factor 1

MS: mass spectrometry

MW: molecular weight

Nedd4: neural precursor cell-expressed developmentally down-regulated protein 4

PROTAC: proteolysis targeting chimera

PTEN: phosphatase and tensin homolog

RING: really interesting new gene

SAR: structure-activity relationship

Ub: ubiquitin

UPS: ubiquitin proteasome system

Wbp2: WW domain-binding protein 2

Table of Contents

List of Figures

List of Tables

CHAPTER 1: A METHOD FOR THE SCREENING OF COVALENT FRAGMENTS

Portions of this chapter appear in this publication: 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*, 4969-4974.

1.1 INTRODUCTION

Fragment based drug discovery (FBDD) has emerged as a powerful approach to discover drug leads by exploring greater chemical diversity space with smaller libraries.[1,2] The major challenge, however, is to detect weak binding interactions between drug-like fragments and their protein target. Disulfide tethering was developed as one solution to this problem.[3] In this approach, disulfide-containing fragments are covalently trapped on the protein surface via the reversible formation of disulfide bonds. Subsequent MS of the intact protein can identify the covalently bound fragment. The advantages of this method include screening the fragments as mixtures rather than as separate entities. Screening fragments as mixtures increases the throughput capability of the assay and reduces the number of false positives by introducing competition between the fragments. This has proven to be a general and successful approach.[4] Another technique relies on the use of an α-cyanoacrylamide moiety attached to drug-like fragments that react reversibly with non-catalytic cysteines present at the binding site of the protein of interest.[5]

However, when this research was initiated in 2011, it was not known whether it was possible to design a robust system where the protein can select the best binder from a mixture of electrophilic fragments under irreversible conditions. Such an approach would be particularly powerful since the identified fragments can subsequently retain their electrophilic tether while being elaborated into a covalent drug. Irreversible tethering would especially benefit the burgeoning field of covalent drug-discovery.[6] However, one concern with such an approach was the danger of selecting the most reactive fragment rather than the fragment with the most specific binding affinity to the protein target.[7] If the electrophilic fragments are too reactive, cysteines or other nucleophilic residues present on the protein surface could undergo non-specific covalent

modifications by the fragments irrespective of their binding affinity.[8] Alternatively, hyper-reactive cysteines or other nucleophilic residues could non-specifically react with even moderately electrophilic fragments, leading to non-selective covalent inhibitors.[9] In addition, no systematic studies have been done to investigate the kinetic reactivity of cysteine reactive electrophiles attached to a large number of drug-like fragments (~50) in order to outline general principles and design rules for irreversible tethering. While this work was in progress, Nonoo et al. reported an irreversible tethering approach using a small ten-member acrylamide library which included known reversible inhibitor scaffolds for thymidylate synthase.[10] However, a hyper-reactive acrylamide in their library had to be discarded, and no systematic studies were done to further outline design rules for irreversible tethering. Moreover, there are still no reports of irreversible screening of an unbiased library of electrophiles to identify novel and selective binding scaffolds for a protein of interest. Therefore, whether it is possible to rationally design an electrophilic library of drug-like fragments for irreversible tethering is still a concern.

This chapter addresses this concern and shows that the proper selection of a cysteine reactive electrophile yields a chemical system that can select weakly bound electrophilic fragments from a mixture, and covalently trap the best binders at the highly reactive catalytic cysteine of the model cysteine protease papain. The discovered fragments behave as weak and irreversible inhibitors of papain, and have novel non-peptidic structures. The reported method serves as an entry point to discover non-peptidic inhibitors of other cysteine proteases, which are promising drug targets to treat parasitic infections.[11]

1.2 RESULTS AND DISCUSSION

1.2.1 Initial fragment library attempts

At the outset of the project, we chose two cysteine-reactive electrophiles which could be coupled directly to amine fragments: acrylamides and vinyl sulfonamides. We chose these electrophiles because they were the most commonly seen in electrophilic inhibitors that have progressed to clinical trials.[6] Both of these electrophiles were then coupled to the same set of ten diverse amine fragments, which contained a variety of aromatic and aliphatic rings. The inherent cysteine reactivity of these fragments was then determined using NMR kinetics studies. The vinyl sulfonamide and acrylamide fragments (at 10 mM) were reacted with N-acetyl cysteine methyl ester (78 mM) in 2:1 dDMSO:dPBS pD 8, and the rate of disappearance of the vinyl protons was plotted to obtain pseudo first-order rate constants.

There was great variety in the reactivity of the acrylamide fragments (Figure 1.1A), with four being completely unreactive, five being very poorly reactive ($k_1 \sim 10^{-5}$ s^{-1}), and one being highly reactive ($k_1 \sim 10^{-2}$ s^{-1}). Nonoo et al. also noticed that some acrylamides, namely electron deficient amino-aryl fragments, were highly reactive and had to be discarded from the library.[10] This confirms that acrylamides are a poor electrophile for fragment screening since many drug-like heterocycles must be removed from the library due to their hyper-reactivity. By contrast, the 8 vinyl sulfonamides tested all reacted readily with cysteine ($k_1 \sim 10^{-2}$-10^{-3} s^{-1}) (Figure 1.1B). However, the tenfold variation between the most reactive and least reactive fragment was still concerning and unsuitable for screening. To our disappointment, this seemingly validated the concern that irreversible covalent fragment screening would only identify the fragments that are the most reactive rather than the ones which bind the protein the best. Nevertheless, we wondered whether differently-configured electrophiles might be better suited to our approach.
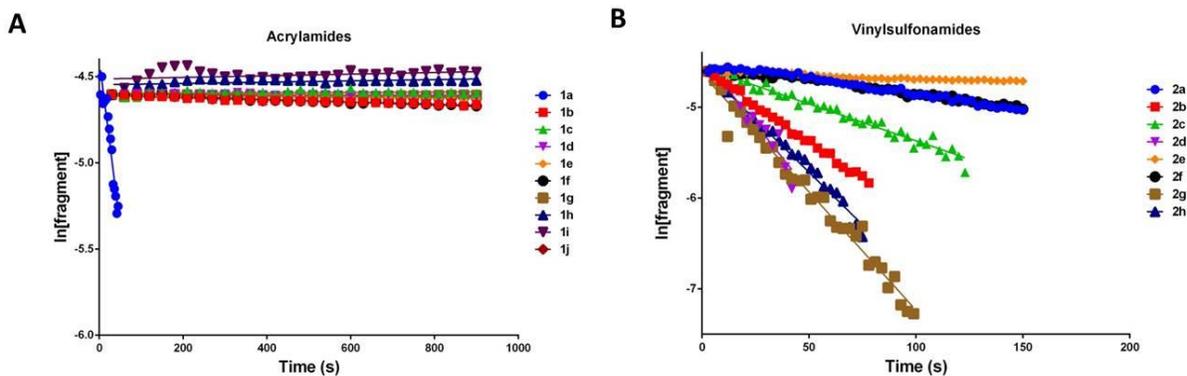
Figure 1.1 Cysteine reactivity of acrylamide and vinylsulfonamide fragments
Pseudo-first order rate plots of the reaction of ten acrylamide (A) and eight vinyl sulfonamide (B) fragments with N-acetyl cysteine methyl ester showing unbalanced reactivity with some hyper-reactive fragments.

1.2.2 Finding an electrophile with balanced reactivity

To find an electrophile which might be suitable for irreversible tethering, we more thoroughly characterized the cysteine reactivity profiles of four Michael acceptors: acrylamides **1** and vinylsulfonamides **2**, and aminomethyl methyl acrylates **3** and methyl vinylsulfones **4** (Figure 1.2). To rigorously test how the cysteine reactivity of these electrophiles would be affected by the structure of attached drug-like fragments, we installed acrylamide and vinyl-sulfonamide electrophiles on aniline, p-MeO-aniline, and p-NO$_2$-aniline to yield electrophiles **1a**-**c** and **2a**-**c**. The methyl acrylate and vinylsulfone electrophiles in **3** and **4** were covalently attached to derivatives of benzoic acid: benzoic acid, p-MeO-benzoic acid, and p-NO$_2$-benzoic acid to yield **3a**-**c** and **4a**-**c**. We envisioned that the different mesomeric and inductive effects of the -OCH$_3$, -H, and –NO$_2$ moieties would cause changes in the reactivity of electrophiles **1**-**4** towards cysteine, and these changes would be representative of fluctuations in the reactivity of drug-like fragments toward cysteines. The electrophile that displayed the least fluctuation in reactivity towards cysteine would be the most optimal electrophile to use for irreversible tethering.
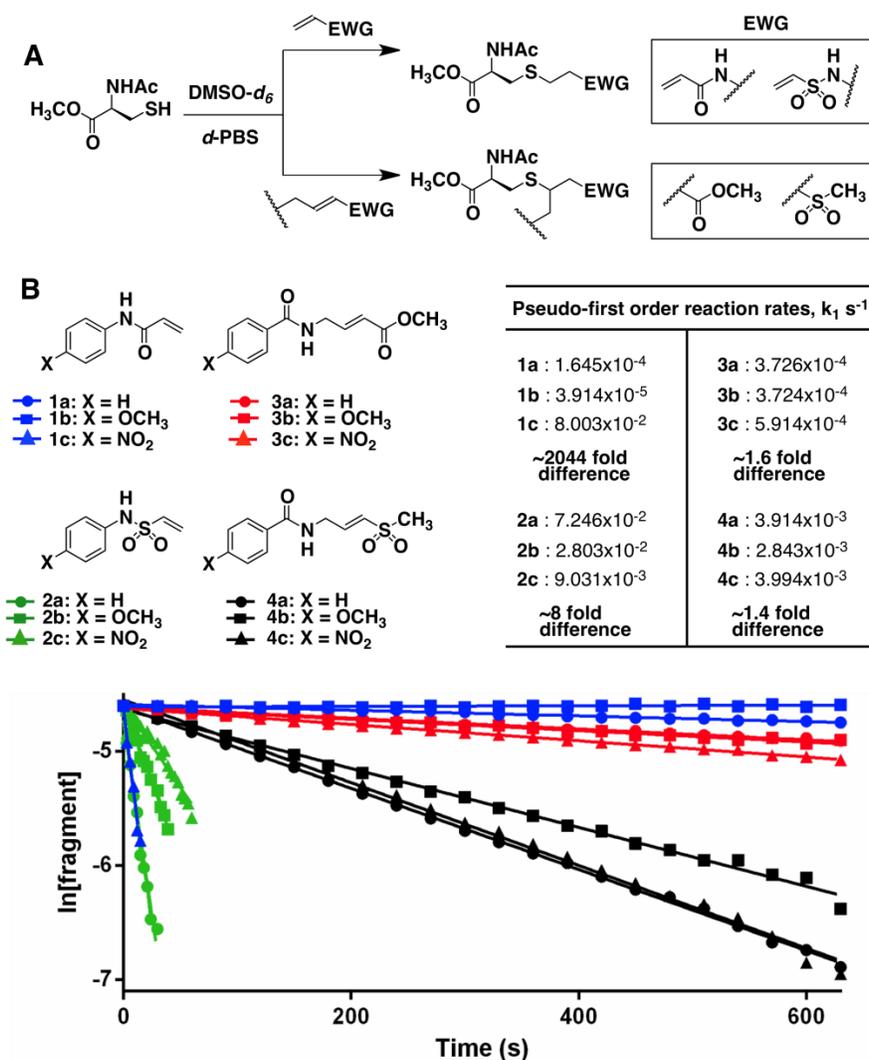
Figure 1.2 NMR rate studies of model fragment electrophiles
A) General scheme of NMR rate studies. B) Chemical structures of the electrophiles **1-4** tested for suitability for irreversible tethering and their pseudo-first order reaction rates with N-acetylcysteine methyl ester at pD 8.0 as measured by NMR spectroscopy.

We therefore measured the pseudo first order reaction rates for each of the compounds **1-4** with N-acetylcysteine methyl ester using NMR spectroscopy (Figure 1.2B).[12] We found that acrylamides **1a-c** displayed a ~2044 fold difference in reactivity, with the –NO$_2$ derivative being the most reactive. Since many drug-like fragments contain an amino group attached directly to electron-deficient aromatic rings, we envisioned that similar to compounds **1a-c** there could be

large fluctuations in the reactivity of such an acrylamide library towards thiols, which would make this library problematic to use, as we observed with our ten diverse acrylamide fragments. Indeed, as we mentioned previously, Nonoo et.al. had to discard one acrylamide fragment from their library due to its hyper-reactivity.[10]

Vinylsulfonamides **2a**-**c** displayed only an ~8 fold difference in reactivity toward N-acetylcysteine methyl ester. This result was also similar to our previous findings with the 8 vinylsulfonamide fragments, so we sought electrophiles with a narrower range of reactivities. Fortunately, both the **3a**-**c** and **4a**-**c** series displayed much more balanced reactivity toward cysteine, with only 1.6 and 1.4 fold differences, respectively, in the reactivity between the least reactive and the most reactive electrophiles. We chose acrylates **3** for further studies because they were tenfold less reactive than vinylsulfones **4**, and therefore less prone to non-specific covalent modifications of nucleophilic amino acid side chains in proteins.[13]

In addition, acrylates are established electrophiles present in irreversible inhibitors of cysteine proteases with activities in *in vitro* biochemical and cell based assays.[14,15] Importantly, *in vitro* k_{inact}/K_I values of acrylate cysteine protease inhibitors vary dramatically (up to 170 fold in the case of falcipain inhibitors) with changes in the structure of the peptide-derived directing group.[14] This indicates that useful levels of kinetic discrimination can be achieved upon structural changes of the directing group, despite the high reactivity of the catalytic cysteine in cysteine proteases. Moreover, the acrylate functionality has been shown to have good pharmacokinetic properties, and is present in an orally bioavailable inhibitor of human rhinovirus 3C protease.[16] These considerations further confirmed to us that acrylate **3** is a good starting point for validating irreversible tethering. Since known acrylate inhibitors are mostly peptidic in nature, we sought to discover novel non-peptidic inhibitors with irreversible tethering.

1.2.3 Building and characterizing a library of vinyl methyl ester fragments

We further validated the utility of electrophile **3** as a thiol-reactive tether by making a library of one hundred structurally diverse drug-like fragments **6**-**105** containing this electrophile. The library was constructed with an HBTU amide coupling with diverse, commercially available carboxylic acid fragments (Figure 1.3A). The acids were selected starting from 94,275 commercially available carboxylic acids. 62,000 were removed because they contained reactive functional groups (e.g. acyl halides) or were unsuitable leads (e.g. nitro compounds). Compounds were then filtered based on a modified "rule of three" criteria[17] to increase the hit rate: molecular weight (MW) ≤ 350 Da; AlogP ≤ 3; hydrogen-bond acceptors ≤ 6; hydrogen-bond donors ≤ 3; rotatable bonds ≤ 4; polar surface area ≤ 90. A principal component analysis and neighborhood algorithm was applied to the 2,958 remaining compounds to produce 755 fragments with a .65 diversity index.[18] 100 of these compounds were then initially selected based on the ease of future analog synthesis and coupled to the acrylate electrophile **5**. We then measured the reaction rates for the first fifty coupled fragments to confirm that this library would have balanced cysteine reactivity and could be used for irreversible tethering (Figure 1.3B). As we expected, these fifty fragments displayed a narrow range of chemical reactivities similar to **3a-c**. Overall we observed only a 2.4 fold difference in the reactivity between the least reactive (k_1 $3.327\times10^{-4}s^{-1}$) and the most reactive (k_1 $7.951\times10^{-4}s^{-1}$) fragment (Figure 1.3B, Table 1.1).
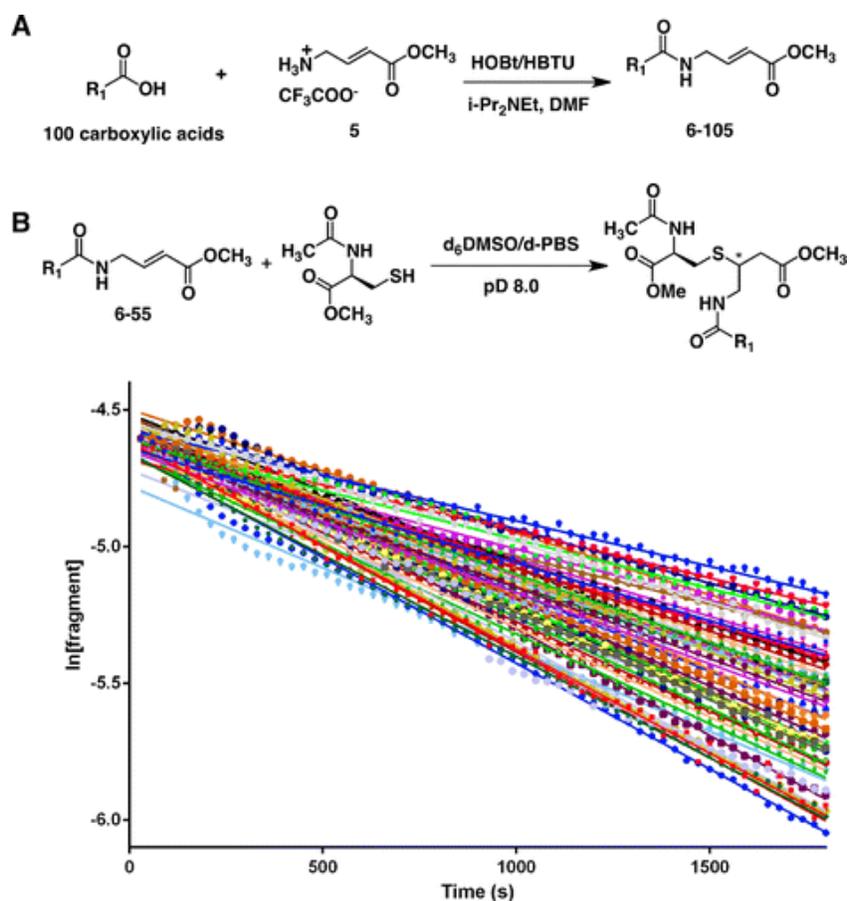
Figure 1.3 Synthesis and cysteine reactivity of 50 fragments
A) Design and synthesis of the fragment library. B). Pseudo-first order NMR rate plots of the reaction of compounds **6**-**55** with N-acetyl cysteine methyl ester. Different colors represent different fragments.

Compound	k pseudo-first order	Compound	k pseudo-first order	Compound	k pseudo-first order
6	0.0007951 s^{-1}	22	0.0007605 s^{-1}	38	0.0007728 s^{-1}
7	0.0006978 s^{-1}	23	0.0004979 s^{-1}	39	0.000651 s^{-1}
8	0.0004232 s^{-1}	24	0.0005202 s^{-1}	40	0.0004793 s^{-1}
9	0.0006824 s^{-1}	25	0.0005202 s^{-1}	41	0.0005635 s^{-1}
10	0.0004656 s^{-1}	26	0.0006107 s^{-1}	42	0.0005281 s^{-1}
11	0.0007414 s^{-1}	27	0.0006665 s^{-1}	43	0.0007616 s^{-1}
12	0.000654 s^{-1}	28	0.0004200 s^{-1}	44	0.0006746 s^{-1}
13	0.0003582 s^{-1}	29	0.0004038 s^{-1}	45	0.0003961 s^{-1}
14	0.0005016 s^{-1}	30	0.0006579 s^{-1}	46	0.0007806 s^{-1}

15	0.0007733 s^{-1}	**31**	0.0005193 s^{-1}	**47**	0.0006539 s^{-1}
16	0.0006414 s^{-1}	**32**	0.0006296 s^{-1}	**48**	0.0004680 s^{-1}
17	0.0006156 s^{-1}	**33**	0.0006348 s^{-1}	**49**	0.0004975 s^{-1}
18	0.0006093 s^{-1}	**34**	0.0007717 s^{-1}	**50**	0.0006514 s^{-1}
19	0.0004396 s^{-1}	**35**	0.0005755 s^{-1}	**51**	0.0005048 s^{-1}
20	0.0005603 s^{-1}	**36**	0.0003400 s^{-1}	**52**	0.0005984 s^{-1}
21	0.0003327 s^{-1}	**37**	0.0004493 s^{-1}	**53**	0.0005521 s^{-1}
Average:	0.000575841 s^{-1}			**54**	0.0005048 s^{-1}
Std Dev:	0.000127064 s^{-1}			**55**	0.0004520 s^{-1}

Table 1.1 Pseudo-first order rate constants of fragments 6-55

1.2.4 Proof-of-concept screen against the cysteine protease papain

Encouraged by these findings, we asked if we could use this library to discover specific covalent enzyme inhibitors with novel structures. As a model protein, we chose the cysteine protease papain. We reasoned that the presence of a highly reactive active site cysteine in papain would serve as a stringent specificity test for the proposed irreversible tethering method. We hypothesized that if the designed chemical system displays specificity in the presence of the highly reactive catalytic cysteine of papain, this system could also be used to discover ligands for less reactive non-catalytic cysteines. In addition, papain is the founding member of a large family of cysteine proteases.[11] Therefore, if the system produced inhibitors of papain, it could serve as an entry point to discover inhibitors of other medically relevant cysteine proteases. For our initial screening, we used a simple MS assay similar to the original disulfide tethering screening conditions (Figure 1.4).
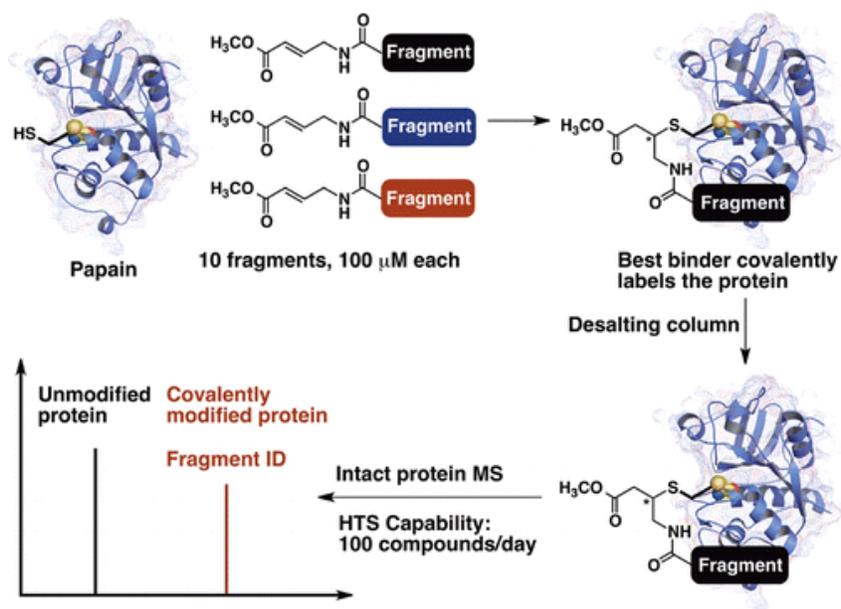
Figure 1. 4 Workflow for screening the covalent fragment library by ESI-MS.

Papain (10 µM) was incubated for one hour with ten reaction mixtures that each contained ten electrophilic fragments (100 µM each) (Table 1.2). Each fragment in the reaction mixture had a unique molecular weight (at least 5 Da difference from the closest fragment) to ensure that whole protein ESI-MS could identify candidate hits unambiguously. Hits were defined as any compounds which labeled papain more than 50%. Remarkably, under these reaction conditions we observed strong mono-labeling of papain by three electrophilic fragments in three separate reaction mixtures: **6**, **7**, and **8** (Figure 1.5). Such selectivity is impressive, given a 9-fold excess of other cysteine reactive electrophiles over compounds **6**, **7**, and **8**. Moreover, we did not detect significant covalent modification of papain with the other seven reaction mixtures. This was despite the fact that these reaction mixtures contained a 100-fold excess of cysteine reactive electrophiles relative to the highly reactive catalytic cysteine of papain. Furthermore, compounds **6**, **7**, and **8** labeled papain even though the corresponding reaction mixtures contained fragments that were equally or even more reactive toward N-acetylcysteine

methyl ester. This observation further suggests that in our system the chemical structure of the

drug-like fragment rather than its reactivity determines the covalent labeling of papain.

Additionally, compounds **6**-**8** demonstrated robust labeling of papain in the presence of 10 mM

glutathione (1000 fold excess relative to papain), confirming that compounds **6**-**8** covalently

label papain due to their specific binding to papain and not simply due to their greater thiol

reactivity.
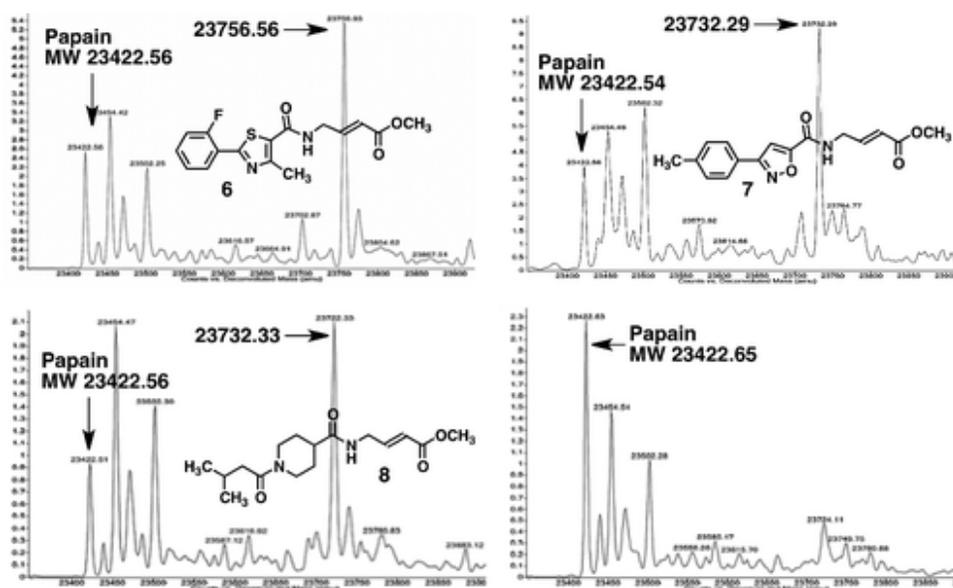
Figure 1.5 Representative MS spectra of 4 fragment mixtures screened against papain.
Papain (10 µM) was incubated with a mixture of 10 electrophilic fragments (100 µM each) for
1h, followed by gel filtration and ESI-MS of the intact protein. Three mixtures with hits and one
with no hits are shown.

Mix 1	MW	Mix 2	MW	Mix 3	MW	Mix 4	MW
25	234.1004	29	234.1004	76	239.1521	103	223.0957
53	254.0458	20	254.163	73	257.0567	28	249.1001
56	272.1161	68	273.1113	21	273.1113	59	259.0768
17	282.1038	60	283.0611	92	285.1113	67	276.1474
16	288.1474	91	288.1474	41	290.1267	50	286.9906
95	298.1317	27	298.1317	89	299.0969	44	290.1267
47	303.1019	86	303.1107	72	304.0882	23	299.1158
8	310.1893	35	312.011	55	314.1267	11	315.0219

102	320.1736	98	323.0561	79	326.1267	88	326.163
51	342.1216	90	342.9878	69	347.1037	54	354.0215
Mix 5	MW	Mix 6	MW	Mix 7	MW	Mix 8	MW
66	249.1113	57	250.0954	37	251.127	39	251.127
36	263.1158	40	266.1267	104	268.0882	64	268.1423
100	277.0773	61	277.1426	62	277.1426	46	279.1583
84	287.1158	12	287.127	45	287.127	48	287.127
101	290.1267	99	291.0831	10	293.1086	19	295.0878
85	299.1158	49	299.127	7	300.111	71	302.0878
26	305.1263	14	306.1016	22	308.0289	63	308.1736
24	315.1219	18	316.1423	43	316.1423	58	316.1423
38	327.0106	42	327.1583	87	330.158	65	334.072
80	357.0688	74	364.119	30	364.1787	83	370.1893
Mix 9	MW	Mix 10	MW				
96	251.127	104	240.1474				
32	270.1004	33	254.0458				
70	280.1059	31	280.1423				
94	287.127	97	288.1222				
77	296.1736	34	297.0001				
13	302.1267	15	302.1267				
78	310.072	52	310.1317				
81	316.1423	75	320.1736				
6	334.0787	93	340.0423				
82	380.0372	9	387.0986				

Table 1.2 Composition of fragment mixes used for screening

We were unable to directly confirm labeling of the catalytic cysteine because the catalytic cysteine peptide was not detectable by ESI-MS or MALDI-TOF upon digestion with trypsin, chymotrypsin, or Glu-C proteases. However, preincubation of papain with compounds **6**-**8**, followed by treatment with **106**, a known papain inhibitor which reacts with its catalytic cysteine,[19] did not cause di-labeling of papain (Figure 1.6A). Additionally, pretreatment of papain with **106** also blocked subsequent labeling by compounds **6**-**8** (Figure 1.6B). These results suggest that compounds **6**-**8** and inhibitor **106** most likely react with the same nucleophilic residue of papain. Compounds **6**-**8** labeled papain in a 1:1 stoichiometry at both 100µM and 1mM concentrations, confirming the specificity of these electrophiles for cysteine.

Moreover, the observed covalent labeling of papain was irreversible, since the covalent adducts were stable to dialysis.
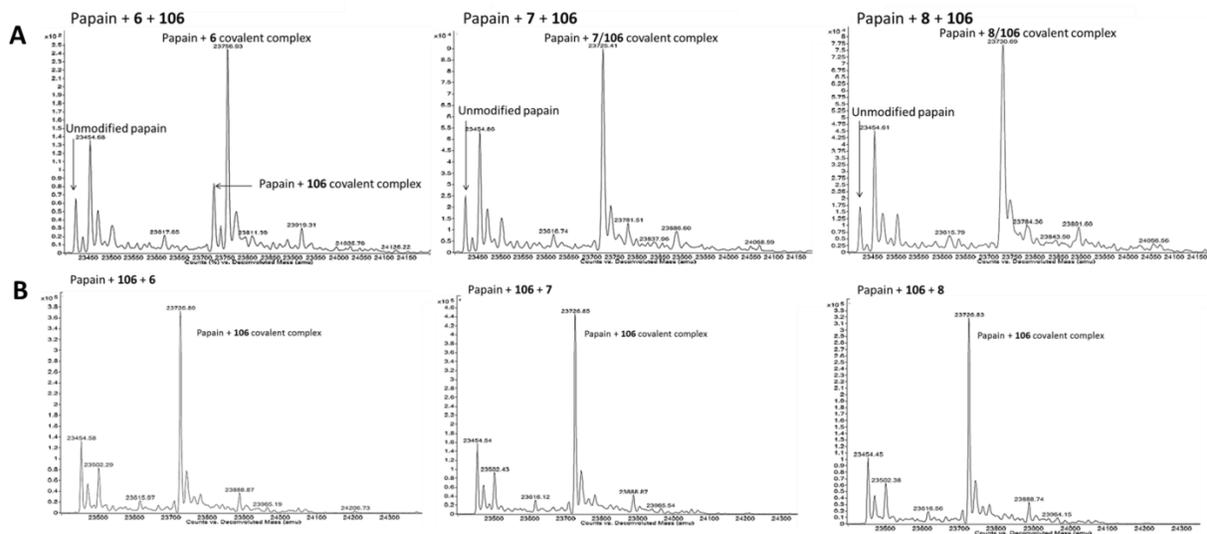
Figure 1.6 Fragment hits and a known papain inhibitor react with the same cysteine
A) ESI-MS of papain treated with 100 μM of **6**, **7**, or **8** for 1h followed by addition of **106** (100μM) and incubation for 1h. The **7**+**106** and **8**+**106** spectra do not show separation between the peaks because inhibitors **7** and **8** are too close in MW to **106**, but the peak is instead a weighted average of the two peaks. However, in no case did treatment with **6-8** followed by **106** result in di-labeling of papain. B) ESI-MS of papain treated with 100 μM of **106** for 1h, followed by addition of 100μM of **6-8** for 1h. In no case did treatment with **106** followed by **6-8** result in dilabeling of papain.

1.2.5 Fragment hits inhibit papain irreversibly

We subsequently tested compounds **6-8** in an enzymatic assay to confirm that they inhibited papain in the concentration and time dependent manner that is characteristic of irreversible inhibitors.[20] Using assay conditions previously described for papain,[19] we determined k_{inact}/K_I values for compounds **6-8** (Figures 1.7, 1.8). Notably, compound **7** was as potent at inhibiting papain as a known moderate peptidic inhibitor **107**,[19] but compounds **6-8** were less potent inhibitors than the known strong peptidic papain inhibitor **106**. This result is expected, since irreversible tethering is designed to detect weak binding interactions between the drug-like fragments and the protein target to identify initial hits. Compounds **6-8** were all more potent

inhibitors than the weak peptidic papain inhibitor **108**.[19] A negative control molecule **19**, which did not label papain in our screen, was ~10 fold less potent at inhibiting papain than the least potent inhibitor **6** and ~33 fold less potent than the most potent inhibitor **7**. Notably, compounds **6-8** do not have a peptidic character in comparison to traditional cysteine protease inhibitors, including known papain inhibitors (Figure 1.8).[21] This result is significant since the proposed method can serve as an entry point to discover other types of non-peptidic inhibitors for medically relevant cysteine proteases, avoiding the undesirable pharmacokinetic liabilities of peptidic inhibitors.[22]
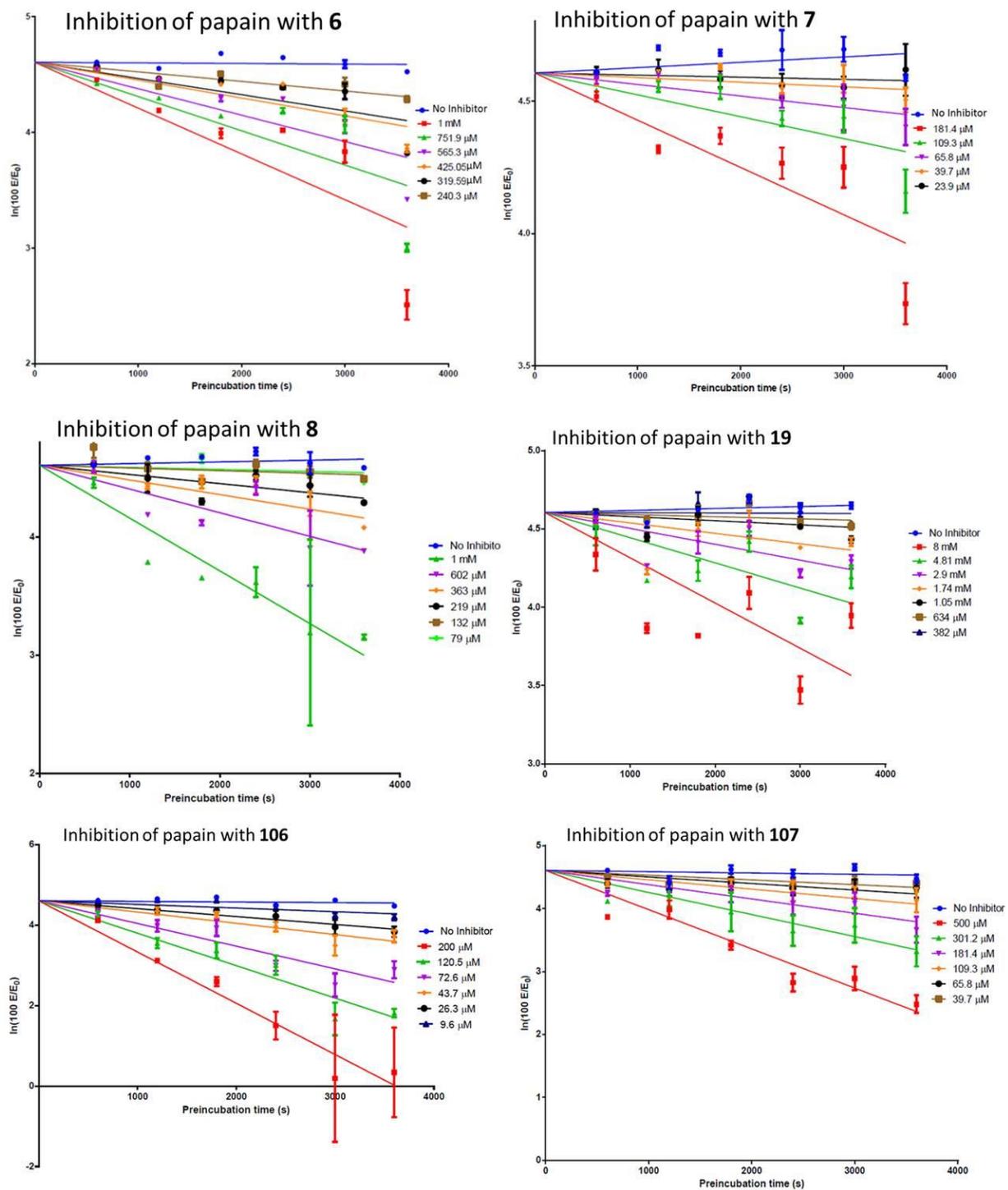
Figure 1.7 Pseudo-first order papain inhibition plots at different concentrations of 6, 7, 8, 19, 106, 107, and 108.

Known inhibitors of papain, k_{inact}/K_I, M^{-1} s^{-1}		Inhibitors of papain in this work, k_{inact}/K_I, M^{-1} s^{-1}	
106	6.575	7	1.228
108	0.386	6	0.409
107	1.102	8	0.461
		19	0.037
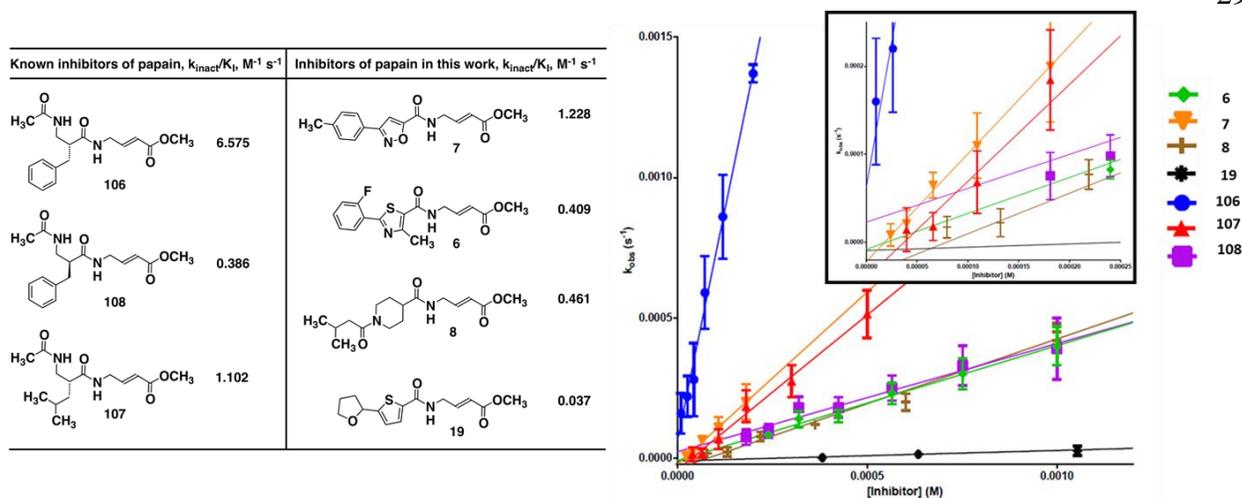
Figure 1.8 Second order inhibition plots and k_{inact}/K_I values
Papain inhibitor compounds **6-8**, known papain inhibitors **106**-**108**, and control compound **19** were tested. Note: compound **7** could not be tested at higher concentrations due to poor solubility.

1.2.6 Counter screens against other enzymes with catalytic cysteines

To further test the specificity of the developed irreversible tethering system, we conducted a counter-screen of the same set of 100 compounds (10 mixtures of 10 compounds each) against three other enzymes with catalytic cysteines: human rhinovirus 3C protease, the E2 ubiquitin-conjugating enzyme UbcH7, and the deubiquitinase USP08, which is a member of the large family of papain-related cysteine proteases.[23] Human rhinovirus 3C protease is a cysteine protease which is an antiviral drug-target with known orally bioavailable acrylate inhibitors.[16] Recent reports have indicated that UbcH7 regulates the entrance into and progression through the S-phase of the cell cycle,[24] while targeting USP08 is a promising approach to overcome gefitinib resistance in lung cancer.[25] We found that HRV3C protease was labeled by compound **22** (~35% labeling) as well as compounds **32** and **98** (~20% labeling) under the same reaction conditions used for papain (Figure 1.9). None of the three papain hits or other electrophilic fragments reacted with HRV3C protease under these conditions, indicating that these hits are selective

binders. Although the three HRV3C hits did not label their target as strongly as the papain hits did, they could eventually be optimized into potent inhibitors of this clinically important cysteine protease. For UbcH7 and USP08, we found that none of compounds **6**-**105** covalently modified these enzymes under the reaction conditions used for papain. When we increased the incubation time with USP08 to 4h, we found two compounds that weakly labeled ~30% of USP08. One was compound **6**, while another was a unique compound (**9**) (Figure 1.10). The other two papain inhibitors **7** and **8** did not label USP08 even after 4h, showing that our system is well behaved and can identify selective binders.
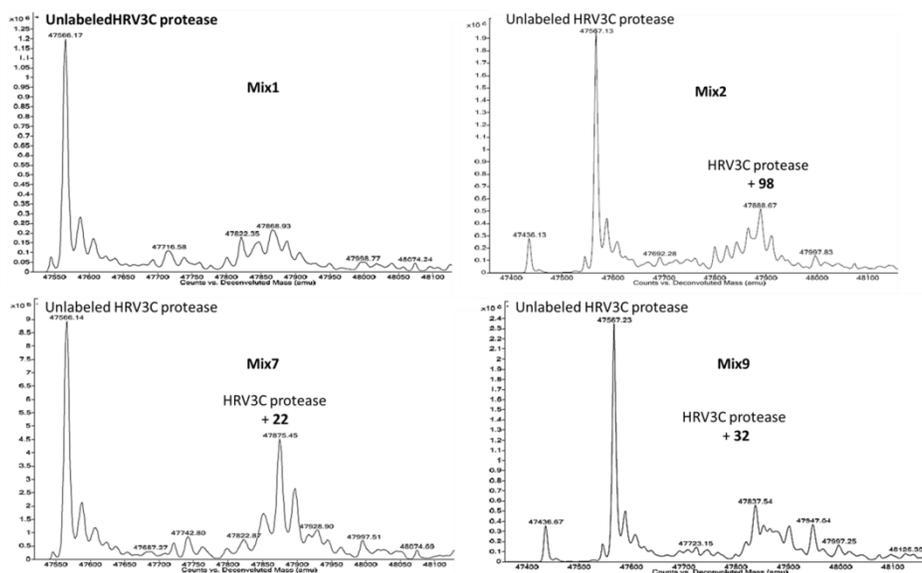
Figure 1.9 Representative MS spectra of four fragment mixtures screened against HRV3C ESI-MS spectra after incubating reaction mixtures containing 10 electrophilic fragments each with GST tagged human rhinovirus 3C protease for 1h. Three mixtures with weak hits and one mixture with no hits are shown.
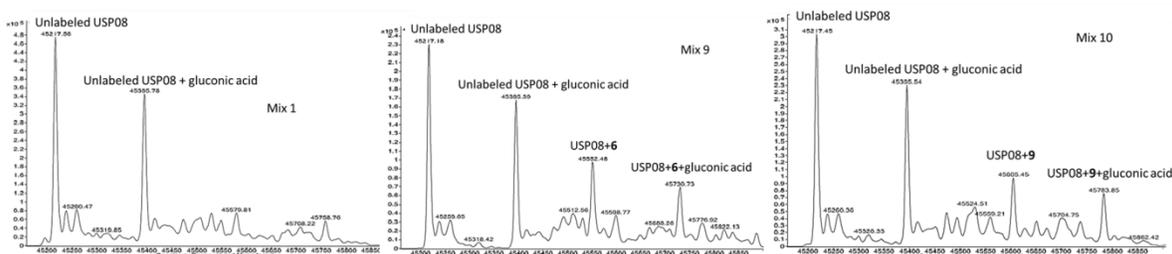
Figure 1.10 Representative MS spectra of four fragment mixtures screened against USP08 ESI-MS spectra of three reaction mixtures containing the catalytic domain of USP08 treated with a mixture of 10 electrophilic fragments for 4h. Two mixtures with weak hits and one mixture with no hits are shown. Notably, USP08 is a member of the papain family of cysteine proteases. USP08 was partially modified by gluconic acid at the N-terminus during bacterial expression.

1.2.7 Collaborations with other research groups

After our work on irreversible tethering was published,[26] several research groups requested our library to screen against their proteins of interest. We synthesized another 100 diverse methyl acrylate fragments to increase the library size to 200 fragments. We sent this library to eleven research groups in the US, the UK, Australia, and Germany. Three groups were able to find and validate unique hits against the parasitic cysteine protease rhodesain, the oncoprotein K-Ras, and the bromodomain protein BRD4.

The screening against rhodesain was conducted by Victor Ogungbe's group at Jackson State University and resulted in our first collaborative paper using the covalent fragment library.[27] Rhodesain is a cathepsin L-like cysteine protease which is a validated drug target in the protozoan *Trypanosoma brucei*, which causes African sleeping sickness.[28] 200 acrylate fragments were screened against rhodesain in an enzymatic assay rather than with ESI-MS. However, since every fragment has the same inherent reactivity with cysteine, one can be confident that any hits will be specific and not simply a hyper-reactive compound. Enzymatic assay screening increases the utility of the fragment library for proteins which are too large to

accurately screen by ESI-MS (>50kDa) or for research groups which do not have access to a mass spectrometer.[29] Seven hits were found for rhodesain, of which compounds **6** and **7** were the most potent (Figure 1.11). Since rhodesain is a member of the papain family of cysteine proteases and has an 81% identical active site and a similar fold to papain, we were not surprised that a fragment molecule would be a hit against both proteases. Further optimization should be able to produce selective inhibitors. Notably, compounds **6** and **7** had anti-trypanosomal IC_{50}s of 30.3 μM and 43.4 μM, respectively, while showing no cytotoxicity towards human Hep G2 cells (IC_{50} >150 μM).
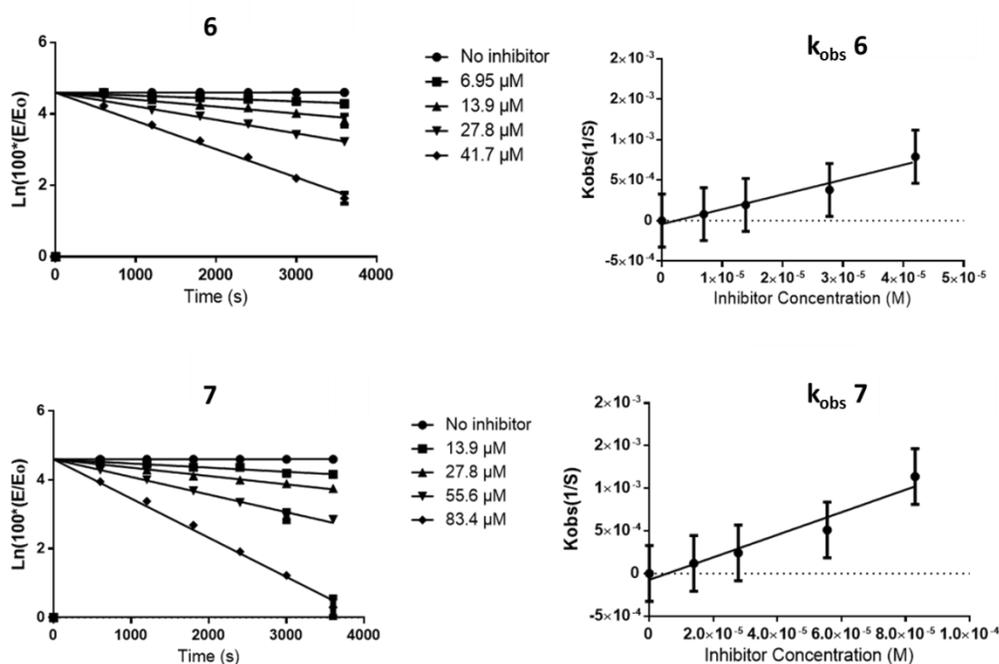
Figure 1.11 Pseudo-first order and second order inhibition plots for compounds 6 and 7 against rhodesain.
These experiments were performed in the lab of Victor Ogungbe at Jackson State University.

1.3 CONCLUSION

In summary, we have rationally designed a chemical system for screening mixtures of electrophilic fragments against the catalytic cysteine of a protein of interest. Our method

eliminates the concern that such an approach would only select the most reactive fragment, or otherwise be nonspecific due to the high reactivity of the catalytic cysteine. Using this method, we identified specific, non-peptidic covalent inhibitors of the cysteine protease papain, which contain novel chemical scaffolds. This was the first example of a successful screen of an unbiased library of electrophilic compounds under irreversible conditions which led to the discovery of specific and novel inhibitor structures for the enzyme of interest.

The key advantage of the reported method is its simplicity. For example, the electrophilic fragments are prepared in one step from commercially available materials using a robust amide bond formation reaction. Moreover, the synthesized electrophilic fragments elicit a predictable and narrow range of chemical reactivities toward thiols and do not react with other nucleophilic residues such as histidine or lysine. The developed screening protocol is simple, and is moderately high-throughput. 200 compounds can be screened in one day without the use of special robotic equipment. Moreover, mixtures of electrophilic fragments can be stored as DMSO stocks, transported, and used to screen fragments against novel protein targets. The developed irreversible tethering method displays a high hit rate (3% for papain and HRV3C protease), and the discovered papain inhibitors have weak potency in enzymatic assays. These are typical characteristics of fragment-based drug discovery methods. Our failure to discover strong inhibitors of USP08 and UbcH7 is most likely not due to the limitations of the method, but rather due to the limited sampling of chemical space since only 100 fragments were screened against these targets. Since USP08 and UbcH7 do not have classical hydrophobic binding pockets like the P2 substrate pocket of papain, it is likely that a larger library will be required to find adequate binders.

While we and others demonstrated that this approach can be used to tether weakly bound fragments to the highly reactive catalytic cysteines of papain and rhodesain, we next asked whether the same approach can be used to tether weakly bound fragments to non-catalytic cysteines on protein surfaces. I will discuss that particular aspect of this technology in the next chapter. Additionally, we are currently working with Rama Mishra at the Center for Molecular Innovation and Drug Discovery to create a virtual library of 1 million electrophilic fragments which can be screened against proteins with catalytic cysteines using recently developed covalent docking methods.[30]

1.4 MATERIALS AND METHODS

Fragment Library Design. Using the Discovery Studio Package with Pipeline Pilot from Accelrys, 94,275 commercially available carboxylic acids were identified from the ChemBridge, ChemDiv, MayBridge, NCI and Sigma-Aldrich libraries using SMARTS query strings. 62,000 of these were removed because they contained reactive functional groups (e.g. acyl halides) or were unsuitable leads (e.g. nitro compounds). Compounds were then filtered based on "rule of three" criteria which were modified to increase the number of resulting compounds: molecular weight (MW) $\leq$ 350 Da; AlogP $\leq$ 3; hydrogen-bond acceptors $\leq$ 3; hydrogen-bond donors $\leq$ 3; rotatable bonds $\leq$ 3; polar surface area $\leq$ 80. A principal component analysis and neighborhood algorithm was applied to the 1,522 remaining compounds to produce 281 fragments with a 0.75 diversity index. 100 of these compounds were then initially selected based on affordability and the ease of future analog synthesis

NMR Rate Studies. N-Acetyl cysteine methyl ester was dissolved in 2:1 deuterated PBS:DMSO-d6 (78mM) with 10mM CH_2Cl_2 as an internal standard. The electrophile (10mM) was then added immediately prior to acquiring spectra. ^1H spectra were taken every 30s for 30

min (or every 4s for 5 min for highly reactive compounds **1c** and **2a-c**). The integrals of the vinyl peaks were used to determine the concentration of the electrophile over time. The natural logarithm of the concentration of the electrophile vs. time was then plotted using GraphPad Prism software. The linear slope of this plot was used to determine the pseudo-first order rate constant. Deuterated PBS recipe: 20mM Na_3PO_4, 50mM NaCl in D_2O was adjusted to pD 8 with DCl solution.

Irreversible Tethering Screening Assay. 10 μM of papain (Sigma P4762), UbcH7 (recombinantly expressed) or USP08 (recombinantly expressed) in 50mM HEPES 150mM NaCl 0.1 mM EDTA pH 7.5 was treated with a mixture of ten fragments (Table 1.2) (10 mM DMSO stock solutions, final concentrations: 100μM of each fragment, and 1% DMSO). The reaction mixture was incubated for 1h or 4h at 23°C before being passed through Zeba gel filtration columns (Thermo, 7K MWCO) to remove unreacted fragments. The protein solution was then immediately analyzed by whole protein LC/ESI-MS.

LC/ESI-MS Protocol. Accurate-mass data were obtained on an Agilent 6210A LC-TOF mass spectrometer in positive ion mode using electrospray ionization. Samples were chromatographed on the LC-TOF instrument using a Poroshell 120 EC-C18 HPLC column (2.1*50 mm, 2.7 micron), an Agilent Series 1200 HPLC binary pump, and an Agilent Series 1200 autoinjector. The HPLC column was held at 45 °C and the autosampler was held at 8 °C. Mobile Phase A was a solution of 0.1% formic acid in water:acetonitrile (19:1). Mobile Phase B was a solution of 0.1 % formic acid in acetonitrile. The flow rate was set to 250 μL/min. The gradient used was 0% B for 2 minutes, ramping linearly to 90% B from 2 minutes to 5 minutes, holding at 90 % B from 5 minutes to 7 minutes, and then returning to 0% B at 7.1 minutes. The column was allowed to equilibrate for 2.7 minutes before the next injection was initiated. The eluent from the column

was diverted to waste for the first 2 minutes. The spectra were acquired from 301 to 3200 daltons using a gas temperature of 340 °C, a gas flow of 7 liters/min, and the nebulizer gas at 35 psi. The following voltages were used: capillary 4200 V, fragmentor 230V, skimmer 64V, and octapole RF peak 250V. Spectra were acquired at a rate of 1 spectra/sec. The data was processed using MassHunter software version B.02.00. Maximum entropy deconvolutions were per-formed with a Mass Step of 1, S/N Threshold of 30, Average Mass at 90% of Peak Height, and 5 Charge States Minimum.

Papain Activity Assays. Papain (4.8 µM) in 50mM Na_3PO_4 2mM EDTA was preactivated with 1mM DTT for 30 min. Activated papain (3.84 µM) in 4:1 mixture of 50mM Na_3PO_4, 2mM EDTA at pH 6.2 and acetronitrile was then preincubated for 1h with varying concentrations of the electrophilic fragment. Every 10 min, 10µL of the reaction mixture was added to a well of a 96-well plate containing 100µL of 4:1 50mM Na_3PO_4 2mM EDTA pH 6.2:acetronitrile with 400µM Cbz-Gly-ONp. p-Nitrophenol product formation was monitored by absorbance at 340nM (ε: 6800 M-1 cm-1) with a Biotek Synergy 4 plate reader. All reactions were performed in duplicate. Product concentration vs. time was plotted with GraphPad Prism software and the initial slope was calculated to determine enzymatic activity (E). The values of k_{inact}/K_I for each inhibitor were then determined according to the method of Kitz and Wilson.[31] Briefly, the slopes of the plots of $ln(100*E_{inhibited}/E_{uninhibited})$ vs. time were used to determine the pseudo-first order inhibition constant k_{obs} for a given concentration of a given inhibitor. The slope of the plot of k_{obs} vs. [Inhibitor] was then used to determine the second order inhibition constant k_{inact}/K_I (since $[I]<<K_I$, the plots were linear at the concentrations tested).

Dialysis experiments (200 µL of papain and compound **6-8** adducts were prepared separately as described above by incubating papain (10 µM) with 100 µM of compounds **6-8** in 1% DMSO

50mM HEPES 150mM NaCl 0.1 mM EDTA pH 7.5. The adducts were loaded into Slide-A-Lyzer dialysis cassettes (Thermo). The cassettes were dialyzed against 500mL of 50mM HEPES 150mM NaCl 0.1mM EDTA pH 7.5 for 16h, then 20 μL were extracted and analyzed by ESI-MS. The buffer was then replaced with fresh buffer and samples continued and switched again after another 24 h, with 20 μL samples extracted for MS analysis prior to each buffer switch.

Recombinant expression of GST-HRV3C protease in E. coli GST-HRV3C protease in a PGEX4T vector plasmid (GST-UbcH7) was transformed into Rosetta (DE3)pLysS cells (Millipore). 1L LB media containing 100μg/ml ampicillin was inoculated with 50 mL overnight cell culture and incubated at 37°C until OD reached ~0.5. Then, IPTG (0.5 mM final concentration) was added to the cell culture media at 28°C, followed by 5 hour incubation at the same temperature. Cells were then harvested and lysed by sonication in phosphate buffered saline (PBS) with 1mM DTT and 1mM PMSF. The supernatant was incubated with glutathione agarose beads (Pierce Biotechnology) for 1 hour at 4°C. The beads were washed three times with PBS + 1mM DTT + 1mM PMSF. The protease was then eluted with with 100 mM Tris pH 8.0, 100 mM NaCl, 10 mM GSH (reduced), 1 mM DTT. The pooled fractions were then dialyzed three times (300, 400, 300 mL) versus 50 mM Tris pH 8.5, 150 mM NaCl, 5 mM DTT, 20% glycerol.

Recombinant expression of UbcH7 in E. coli UbcH7 in a PGEX6P1 vector plasmid (GST-UbcH7) was transformed into BL21 cells (Novagen). 1L LB media containing 100μg/ml ampicillin was inoculated with 50 mL overnight cell culture and incubated at 37°C until OD reached ~1.2. Then, IPTG (1.0 mM final concentration) was added to the cell culture media at 30°C, followed by 4 hour incubation at the same temperature. Cells were then harvested and lysed by sonication in phosphate buffered saline (PBS) with protease inhibitors (Complete Mini

Protease Inhibitor Cocktail, Roche). The supernatant was incubated with glutathione agarose beads (Pierce Biotechnology) for 1 hour at 4°C. The beads were washed three times with PBS and incubated with PreScission Protease overnight at 4°C to elute UbcH7 (50mM HEPES, 150 mM NaCl, 0.1 mM EDTA).

Recombinant expression of USP08 catalytic domain in E. coli USP08 catalytic domain in a PET21a-LIC vector plasmid (6×His-USP08, Addgene) was transformed into BL21 (DE3) cells (Invitrogen). 1L TB media containing 100µM kanamycin and 600µl antifoam 204 (Sigma A-8311) was inoculated with 50ml overnight culture and incubated at 37°C until OD reached ~3. Then, IPTG (100 µM final concentration) was added to the cell culture media at 15°C. The culture was incubated overnight at the same temperature. Cells were then harvested and lysed by sonication in 10 mM Tris-HCl pH 7.0, 0.5 M NaCl 5% glycerol 2 mM imidazole 1 mM β-mercaptoethanol 0.1 µM PMSF. The cleared lysate was then loaded onto TALON metal-affinity beads at 4°C. Beads were washed three times with 10 mM Tris-HCl, pH 7.0 0.5 M NaCl 5% glycerol 10 mM imidazole 1 mM β-mercaptoethanol 0.05% Tween 20. The protein was then eluted with 10 mM Tris-HCl pH 7.0, 0.5 M NaCl 5% glycerol 200 mM imidazole 1 mM β-mercaptoethanol before being exchanged into 50mM HEPES 150mM NaCl 0.1mM EDTA pH 7.5 with PD10 columns (GE Healthcare). MS analysis of USP08 showed that the resulting protein had a cleaved N-terminal methionine residue, and ~50% of the protein had been further modified by gluconic acid at the N-terminus.

1.5 SYNTHESIS

General Information Methanol (ACS grade), ethyl acetate (ACS grade), chloroform (ACS grade), toluene (ACS grade), and diethyl ether (ACS grade), acetonitrile (HPLC grade), and hexanes (ACS grade) were purchased from Fisher Scientific and used without further

purification. Dichloromethane, tetrahydrofuran and dimethylformamide were purified by passing over activated alumina. Commercially available reagents were used without further purification. Unless otherwise specified, reagents were purchased from Sigma-Aldrich and carboxylic acid fragments were purchased from ChemBridge. Reactions were monitored by thin-layer chromatography (TLC) on pre-coated glass backed plates (60 Å silica gel, 0.25mm, Whatman), and components were visualized by UV light (254 and 365 nm) or by treating the plates with anisaldehyde, KMnO4, and ninhydrin stains followed by heating. Flash column chromatography was performed over ultra pure silica gel (230-400 mesh) from Silicycle. ^{1}H and ^{13}C NMR spectra were obtained on a Bruker AVANCE III 500 MHz spectrometer or an Agilent DDR2 400 MHz spectrometer. Chemical shifts were reported in ppm relative to the residual solvent peak (CDCl$_3$ or DMSO-d$_6$). Multiplicity was indicated as follows: s (singlet); d (doublet); t (triplet); q (quartet); m (multiplet); dd (doublet of doublets); ddd (doublet of doublet of doublets); dt (doublets of triplets); td (triplet of doublets); brs (broad singlet). Coupling constants were reported in Hz. Small molecule ESI-MS was performed on an Agilent 1100 MSD quadropole instrument. For compounds tested in enzymatic assays, purity was confirmed by analytical HPLC on a Shimadzu LC-6AD instrument with a Restek Pinnacle C18 column with UV detection at 220nm with a 5→95% acetonitrile/water gradient, 0.1% trifluoracetic acid.

Synthesis of 1a-c Aniline, p-methoxyaniline, or p-nitroaniline (1.07 mmol) was dissolved in THF (0.1 M, 10.7 mL) and cooled to 0°C with stirring. Diisopropylethylamine (1.4 mL, 8.58 mmol) was then added, followed by acryloyl chloride (175 µL, 2.14 mmol). After 5 min., the reaction was warmed to 23°C and stirred for 1 hour. TLC showed a full conversion to product. THF was evaporated under reduced pressure, and the residue was dissolved in 20mL dichloromethane and washed with a saturated aqueous solution of NaHCO$_3$ (2×20 mL). The

organic layer was dried with magnesium sulfate, filtered, and evaporated under reduced pressure. The residue was purified by flash chromatography with an ethyl acetate/hexanes gradient 25% EtOAc→100% EtOAc.
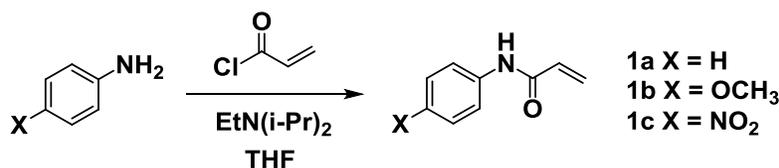
1a (108 mg, 68% yield) ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.52 (d, J = 7.9 Hz, 2H), 7.35 – 7.23 (m, 3H), 7.06 (t, J = 7.4 Hz, 1H), 6.38 (dd, J = 16.9, 1.3 Hz, 1H), 6.19 (dd, J = 16.8, 10.2 Hz, 1H), 5.71 (dd, J = 10.3, 1.2 Hz, 1H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 163.53, 137.71, 131.16, 129.07, 127.90, 124.56, 119.97.

1b (166 mg, 87% yield) ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.52 (d, J = 8.9 Hz, 2H), 7.21 (s, 1H), 6.90 (d, J = 9.0 Hz, 2H), 6.45 (dd, J = 16.8, 1.3 Hz, 1H), 6.26 (dd, J = 16.8, 10.2 Hz, 1H), 5.78 (dd, J = 10.3, 1.3 Hz, 1H), 3.83 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 163.38, 156.55, 131.14, 130.79, 127.50, 121.76, 114.18, 55.49.

1c (44.4 mg, 22% yield) ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.17 (d, J = 9.1 Hz, 2H), 7.72 (d, J = 9.1 Hz, 2H), 7.43 (s, 1H), 6.45 (dd, J = 16.8, 1.0 Hz, 1H), 6.21 (dd, J = 16.8, 10.3 Hz, 1H), 5.83 (dd, J = 10.4, 1.0 Hz, 1H). 13C NMR (126 MHz, CDCl$_3$) δ 164.59, 144.55, 143.20, 130.47, 128.98, 124.92, 119.23.

Synthesis of 2a-c Aniline, p-methoxyaniline, or p-nitroaniline (1.07 mmol) was dissolved in CH2Cl2 (0.1M, 10.7 mL), and cooled to 0°C with stirring. 2-chloroethane sulfonyl chloride (112 μL, 1.07 mmol) was then added, followed by triethylamine (150 μL, 1.07 mmol). After 1h of reaction time, a second equivalent of triethylamine (150 μL, 1.07 mmol) was added and the reaction was warmed to 23°C. After one hour TLC showed full conversion of the starting

material to product, and the reaction was quenched with 20 mL water and extracted with 2×20mL dichloromethane. The combined organic layers were washed with 20 mL 1M HCl and 20 mL saturated aqueous sodium chloride. The organic phase was then dried over magnesium sulfate, filtered, and evaporated under reduced pressure. Purified by flash column chromatography with a CH_3OH/CH_2Cl_2, CH_3OH gradient 0→5%.
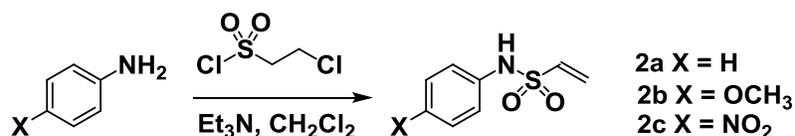
2a (107.6 mg, 55% yield) ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.37 − 7.27 (m, 2H), 7.20 − 7.06 (m, 3H), 6.56 (s,1H), 6.55 (dd, J = 16.5, 9.9 Hz, 1H), 6.27 (d, J = 16.5 Hz, 1H), 5.94 (d, J = 9.9 Hz, 1H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 136.15, 135.02, 129.53, 128.57, 125.42, 120.97.

2b (160 mg, 70% yield) ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.17 (d, J = 8.9 Hz, 2H), 6.88 (d, J = 8.9 Hz, 2H), 6.56 (dd, J = 16.5, 9.9 Hz, 1H), 6.28 (s, 1H), 6.22 (d, J = 16.6 Hz, 1H), 5.96 (d, J = 9.9 Hz, 1H), 3.82 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 158.03, 134.96, 128.44, 128.27, 125.12, 114.61, 55.50.

2c (64.8 mg, 26% yield) ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.24 (d, J = 9.1 Hz, 2H), 7.27 (d, J = 9.1 Hz, 2H), 6.64 (dd, J = 16.4, 9.8 Hz, 1H), 6.49 (d, J = 16.5 Hz, 1H), 6.15 (d, J = 9.8 Hz, 1H). ^{13}C NMR (126 MHz, Chloroform-d) δ 143.55, 143.33, 135.05, 128.84, 125.35, 117.82.

Synthesis of 3a-c: General Scheme
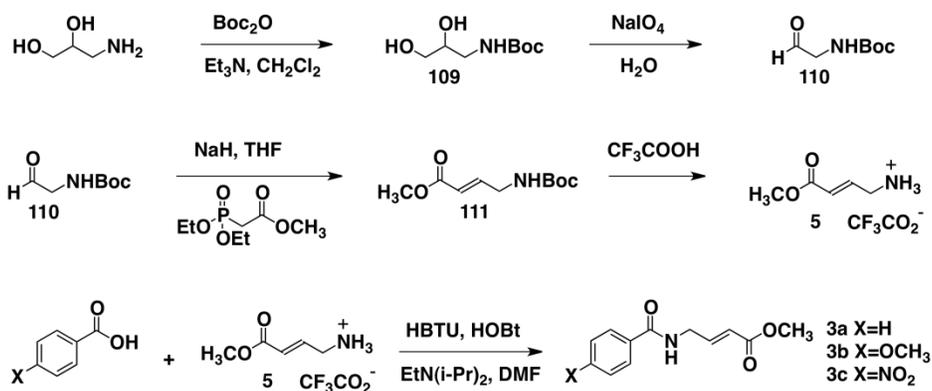
Synthesis of 109 (±)-3-amino-1,2-propanediol (11.29g, 124 mmol) was dissolved in CH_2Cl_2:CH_3OH (1:5) (1M) and triethylamine (2mL, 14.7 mmol) was added. Di-tert-butyl dicarbonate (32.5g, 149 mmol) was dissolved in dichloromethane (0.8M, 186 mL) and added slowly to the reaction mixture. The resulting reaction was stirred at 23°C for 2h, followed by TLC analysis that showed a full consumption of the starting material. The reaction mixture was evaporated under reduced pressure, and the residue was purified by column chromatography with EtOAc:Hexanes 1:4, then dried on high vacuum to yield **109** as a white solid (23.7g, 94% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 5.28 − 4.96 (m, 1H), 3.83 − 3.73 (m, 1H), 3.60 (qd, J = 11.7, 4.9 Hz, 2H), 3.44 (s, 1H), 3.27 (dt, J = 12.9, 6.0 Hz, 2H), 1.46 (s, 9H). ^{13}C NMR (126 MHz, Chloroform-d) δ 157.45, 80.13, 71.37, 63.58, 28.35, 27.42.

Synthesis of 110 109 (10g, 52mmol) was suspended in H_2O (0.6M, 87.2mL) and the flask was covered in foil (to protect $NaIO_4$ from light). $NaIO_4$ (13.4g, 62.8 mmol) was then added and the reaction was stirred for 1h. A white precipitate had formed after 1h, and TLC analysis showed full consuption of the starting material. The precipitate was filtered off, and the aqueous layer was extracted with $CHCl_3$ (8×50 mL). The organic layer was dried with $MgSO_4$, filtered, and evaporated to yield **110** as a yellow oil, which was used immediately without further purification

(7.7g, 93% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 9.68 (s, 1H), 5.23 (s, 1H), 4.10 (d, J = 5.2 Hz, 2H), 1.47 (s, 9H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 197.21, 155.67, 80.19, 51.39, 28.28.

Synthesis of 111 Sodium hydride (60% dispersion in mineral oil) (1.9 g, 46.6 mmol) in tetrahydrofuran (0.17 M, 274 mL) was cooled to 0°C, then triethylphosphonoacetate (8.5 mL, 46.6 mmol) in THF was added dropwise. The reaction was stirred at 0°C for 20 min, then **110** (7.4 g, 46.6 mmol) in THF was added. The reaction was allowed to warm to 23°C and was stirred for 1h. TLC showed a full consuption of the starting materials and conversion to product. THF was removed under reduced pressure, and the residue was then diluted with ethyl acetate (200mL) and water (200 mL). The layers were separated, followed by the extraction of the aqueous layer with EtOAc (2×100 mL). The organic layer was then dried over MgSO$_4$, filtered, and evaporated. The residue was purified by flash column chromatography with an ethyl acetate/hexanes gradient 25% EtOAc → 50% EtOAc to yield **111** (6.6g, 66% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 6.94 (dt, J = 15.7, 4.8 Hz, 1H), 5.97 (dt, J = 15.8, 1.9 Hz, 1H), 4.73 (s, 1H), 3.95 (t, J = 5.6 Hz, 2H), 3.76 (s, 3H), 1.48 (s, 9H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.55, 145.26, 120.71, 79.73, 60.37, 51.58, 41.28, 28.30.

Synthesis of 5 111 (6.6 g, 30.8 mmol) was dissolved in trifluoroacetic acid (47 mL, 617 mmol) and stirred at 23°C for 30 min. TLC at 30 min showed conversion to product. TFA was evaporated and azeotroped with toluene (2×100mL). The residue was then dried on high vacuum for 2 hours, dissolved in 2 mL methanol and dropped into ice cold diethyl ether (200 mL). The ether was then filtered to collect **5** as the TFA salt (6.2 g, 88% yield). ^{1}H NMR (500 MHz, DMSO-d$_6$) δ 8.08 (s, 3H), 6.86 (dt, J = 15.9, 5.6 Hz, 1H), 6.15 (dt, J = 16.0, 1.7 Hz, 1H), 3.70 (s, 3H), 3.70 (d, J = 1.8 Hz, 2H). ^{13}C NMR (126 MHz, DMSO-d$_6$) δ 165.33, 140.61, 123.22, 51.72.

Synthesis of 3a-c Benzoic acid, p-methoxybenzoic acid, or p-nitrobenzoic acid (0.35 mmol) was dissolved in dimethylformamide (0.2M, 1.75 mL), then **5** (42.6mg, 0.35 mmol), HBTU (128mg, 0.34 mmol), and HOBT (51.8 mg, 0.38 mmol) were added, followed by diisopropylethylamine (175 μL, 1.047 mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H_2O (5mL) and extracted with DCM (3×5mL). The combined organic layers were washed with 1M HCl (10mL), saturated aqueous $NaHCO_3$ (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over $MgSO_4$, filtered, and evaporated. Purification with flash column chromatography with CH_3OH/CH_2Cl_2 (CH_3OH gradient 0→5%) yielded **3**.

3a (65.6 mg, 86% yield) ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.93 − 7.71 (m, 2H), 7.62 − 7.52 (m, 1H), 7.53 − 7.37 (m, 2H), 7.03 (dt, J = 15.7, 5.1 Hz, 1H), 6.39 (s, 1H), 6.02 (dt, J = 15.7, 1.9 Hz, 1H), 4.43 − 4.18 (m, 2H), 3.76 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 167.43, 166.42, 144.13, 133.85, 131.87, 128.72, 126.96, 121.61, 51.75, 40.61.

3b (75.7 mg, 87% yield) ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.79 (d, J = 8.8 Hz, 2H), 7.04 (dt, J = 15.7, 5.1 Hz, 1H), 6.97 (d, J = 8.8 Hz, 2H), 6.22 (s, 1H), 6.02 (dt, J = 15.7, 1.9 Hz, 1H), 4.28 (ddd, J = 6.1, 5.1, 1.9 Hz, 2H), 3.89 (s, 3H), 3.76 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.88, 166.46, 162.43, 144.39, 128.79, 126.08, 121.50, 113.87, 55.47, 51.73, 40.57.

3c (59.9 mg, 65% yield) ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.30 (d, J = 8.7 Hz, 2H), 7.94 (d, J = 8.7 Hz, 2H), 6.98 (dt, J = 15.7, 5.3 Hz, 1H), 6.31 (s, 1H), 5.99 (d, J = 15.7 Hz, 1H), 4.28 (td, J = 5.7, 1.9 Hz, 2H), 3.74 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.25, 165.42, 149.78, 143.19, 139.37, 128.24, 123.97, 122.17, 51.86, 40.87.

Synthesis of 4: General Scheme
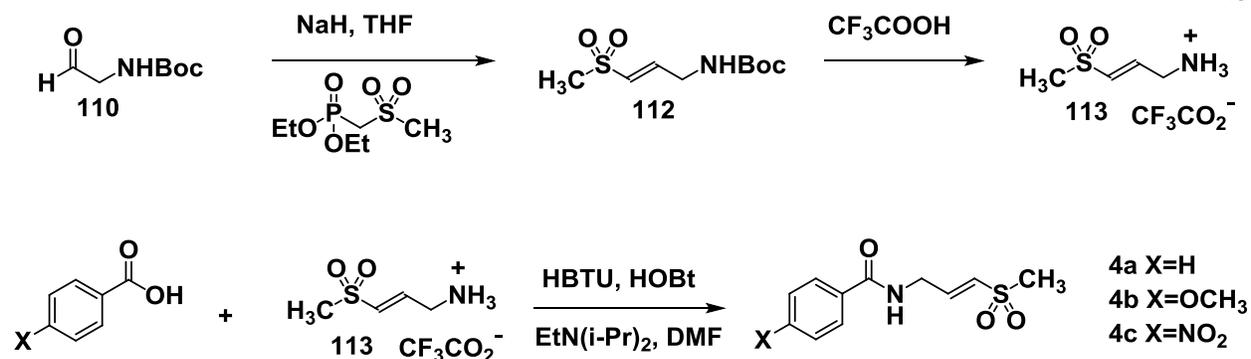
Synthesis of 112 Sodium hydride (60% dispersion in mineral oil) (233.2 mg, 5.83 mmol) in tetrahydrofuran (0.17M, 34.3 mL) was cooled to 0°C with stirring, followed by the dropwise addition of diethyl(methylsulfonylmethyl)phosphonate (Oakwood) (1342.2 mg, 5.83 mmol) in 5 mL THF. The reaction was stirred at 0°C for 20 min, then **110** (928 mg, 5.83 mmol) in 5 mL THF was added. The reaction was allowed to warm to 23°C and was stirred for 1h. TLC showed conversion to the product. THF was removed under reduced pressure, and the residue was then diluted with ethyl acetate (30 mL) and water (30 mL). The layers were separated, and the aqueous layer was extracted with EtOAc (2×30 mL). The organic layer was then dried over MgSO$_4$, filtered, and evaporated. The residue was purified by flash column chromatography with an ethyl acetate/hexanes gradient 25% EtOAc → 50% EtOAc to yield **112** (530.4 mg, 56% yield). ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.38 (dd, J = 11.7, 5.8 Hz, 1H), 6.28 (dt, J = 11.4, 1.7 Hz, 1H), 5.01 (s, 1H), 4.23 (td, J = 6.3, 1.7 Hz, 2H), 3.00 (s, 3H), 1.41 (s, 9H).

Synthesis of 113 **112** (530.4 mg, 2.26 mmol) was dissolved in trifluoroacetic acid (3.45 mL, 45.1 mmol) and stirred at 23°C for 30 min. TLC at 30 min showed conversion to product. Trifluoroacetic acid was evaporated off and azeotroped with toluene (2×30 mL). The residue was then dried on high vacuum for 2 hours, dissolved in 1 mL methanol and dropped into ice cold diethyl ether (100 mL). The resulting mixture was filtered to collect **113** as the TFA salt

(435 mg, 77% yield). ^{1}H NMR (500 MHz, DMSO-d6) δ 8.18 (s, 1H), 6.99 (dt, J = 15.4, 1.7 Hz, 1H), 6.75 (dt, J = 15.4, 5.5 Hz, 1H), 3.76 (dd, J = 5.3, 1.7 Hz, 2H), 3.06 (s, 3H). ^{13}C NMR (126 MHz, DMSO-d6) δ 137.95, 133.13, 48.56, 42.02.

Synthesis of 4a-c Benzoic acid, p-methoxybenzoic acid, or p-nitrobenzoic acid (0.2 mmol) was dissolved in dimethylformamide (0.2M, 1mL), then **113** (50mg, 0.2 mmol), HBTU (73.8 mg, 0.16 mmol), and HOBT (29.8 mg, 0.22 mmol) were added, followed by DIPEA (100.7 µL, 0.6 mmol). The reaction was stirred at 23°C for 16h. TLC at 16 h showed conversion to product. The reaction was quenched with H_2O (5mL) and extracted with CH_2Cl_2 (3×5mL). The combined organic layers were washed with 1M HCl (10mL), saturated aqueous $NaHCO_3$ (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over $MgSO_4$, filtered, and evaporated. Purified by flash column chromatography with a CH_3OH/CH_2Cl_2, CH_3OH gradient 0→5% to yield **4a** (30.5 mg, 64% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.83 (dt, J = 7.1, 1.4 Hz, 2H), 7.74 − 7.56 (m, 1H), 7.56 − 7.44 (m, 2H), 7.04 (dt, J = 15.2, 4.5 Hz, 1H), 6.56 (d, J = 15.1 Hz, 1H), 6.48 (t, J = 5.8 Hz, 1H), 4.37 (ddd, J = 6.2, 4.5, 1.9 Hz, 2H), 2.98 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 167.46, 144.22, 133.31, 132.21, 130.08, 128.84, 126.98, 42.85, 39.88.

Synthesis of 6-108. The carboxylic acid fragment (0.2 mmol) was dissolved in dimethylformamide (0.2M, 1mL), then 5 (46 mg, 0.2 mmol), HBTU (73.8 mg, 0.16 mmol), and HOBt (29.8 mg, 0.22 mmol) were added, followed by EtN(i-Pr)$_2$ (100.7 µL, 0.6 mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H2O (5mL) and extracted three times with CH_2Cl_2 (5 mL). The combined organic layers were washed with 1M HCl (10 mL), saturated aqueous $NaHCO_3$ (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over $MgSO_4$, filtered, and

evaporated. Purified by flash column chromatography with a CH_3OH/CH_2Cl_2, CH_3OH gradient 0-5% to yield compounds **6**-**108**. Yields ranged from 11% to 100%, with an average yield of 60%. For initial library creation, compounds were characterized by ^{1}H NMR and low resolution MS. All compounds tested in enzymatic assays were also characterized by ^{13}C NMR and ≥95% purity was confirmed by HPLC.

Structures of **6-108**:

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

102

103

104

105

106

107

108
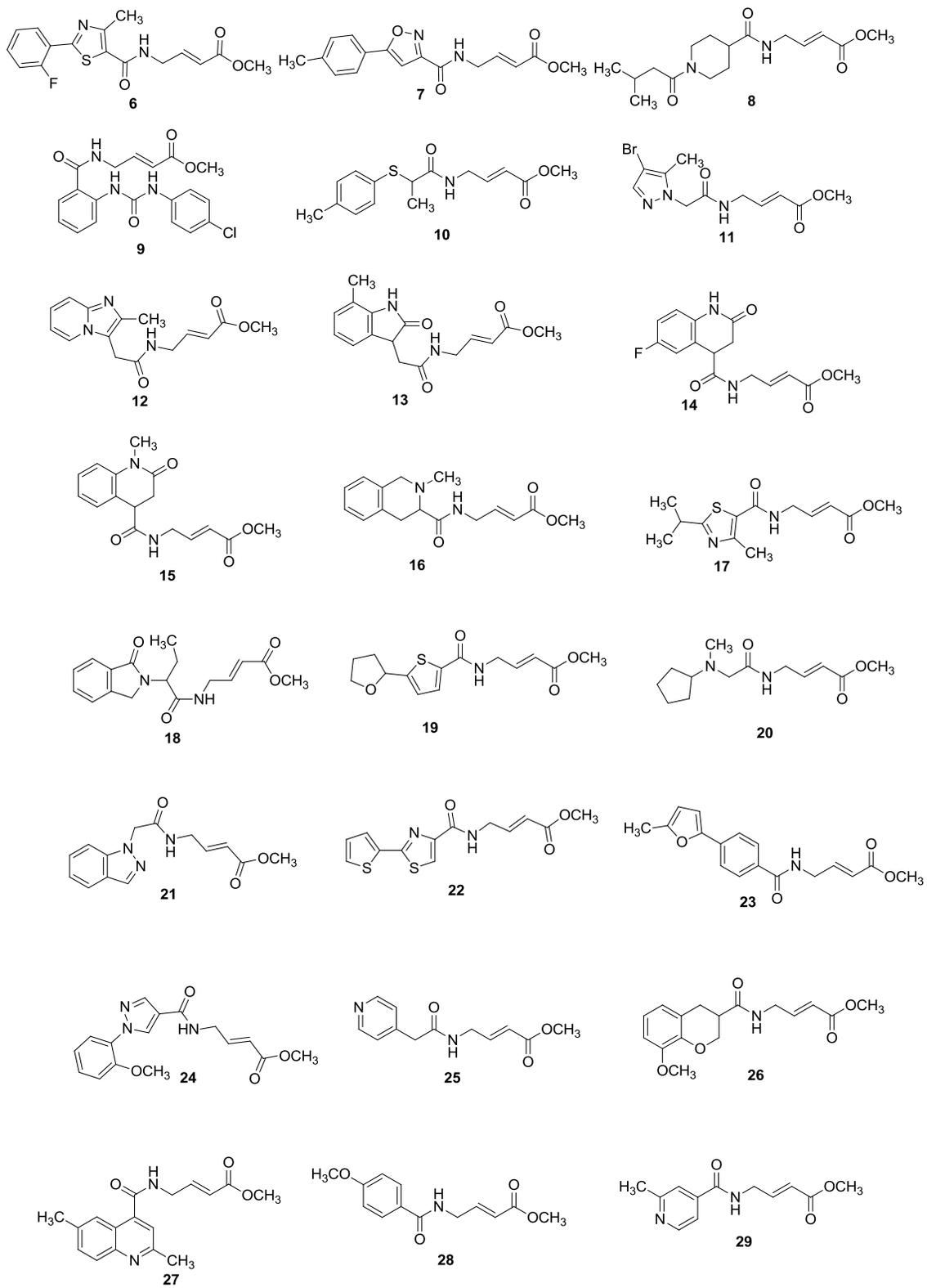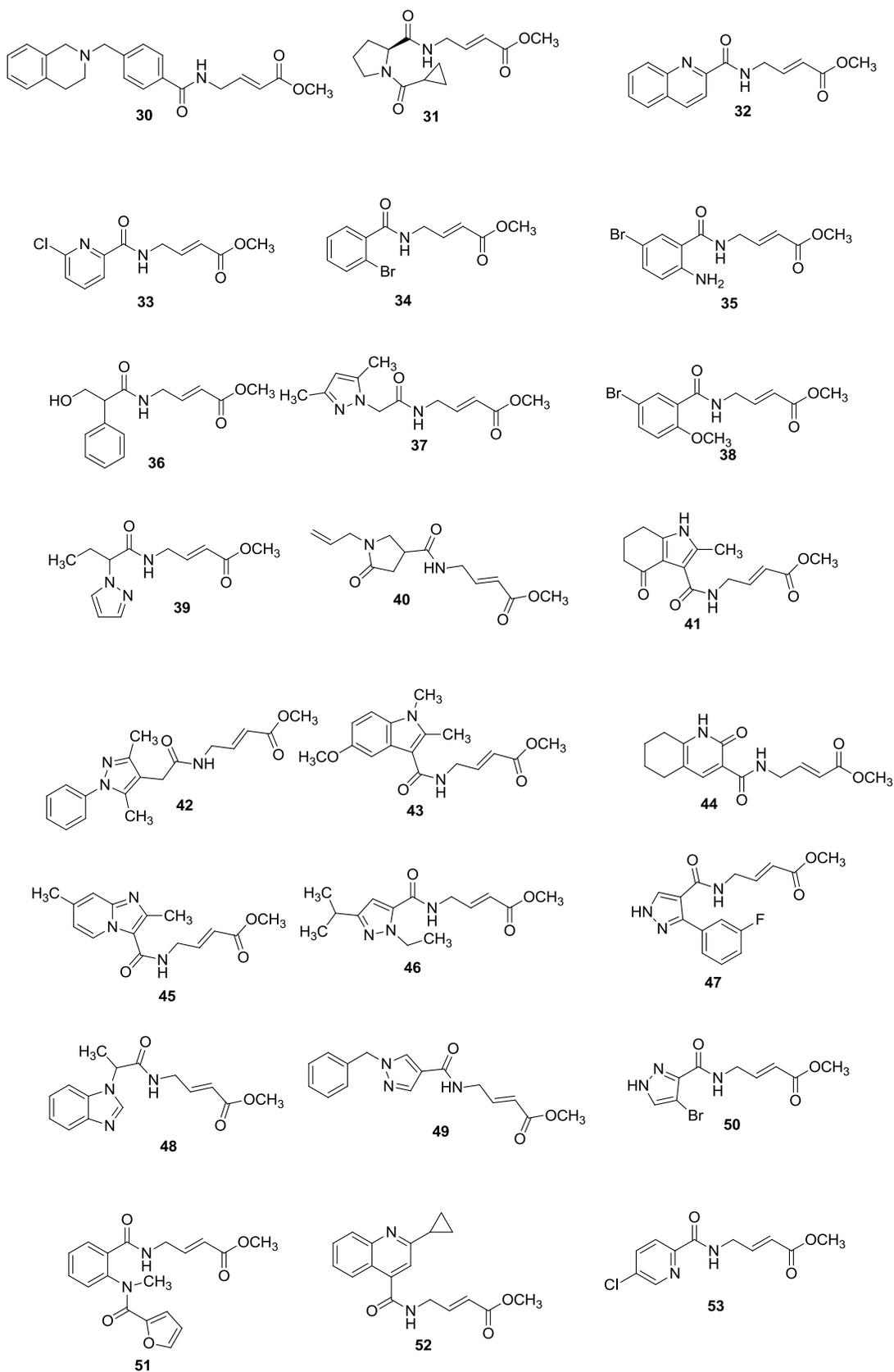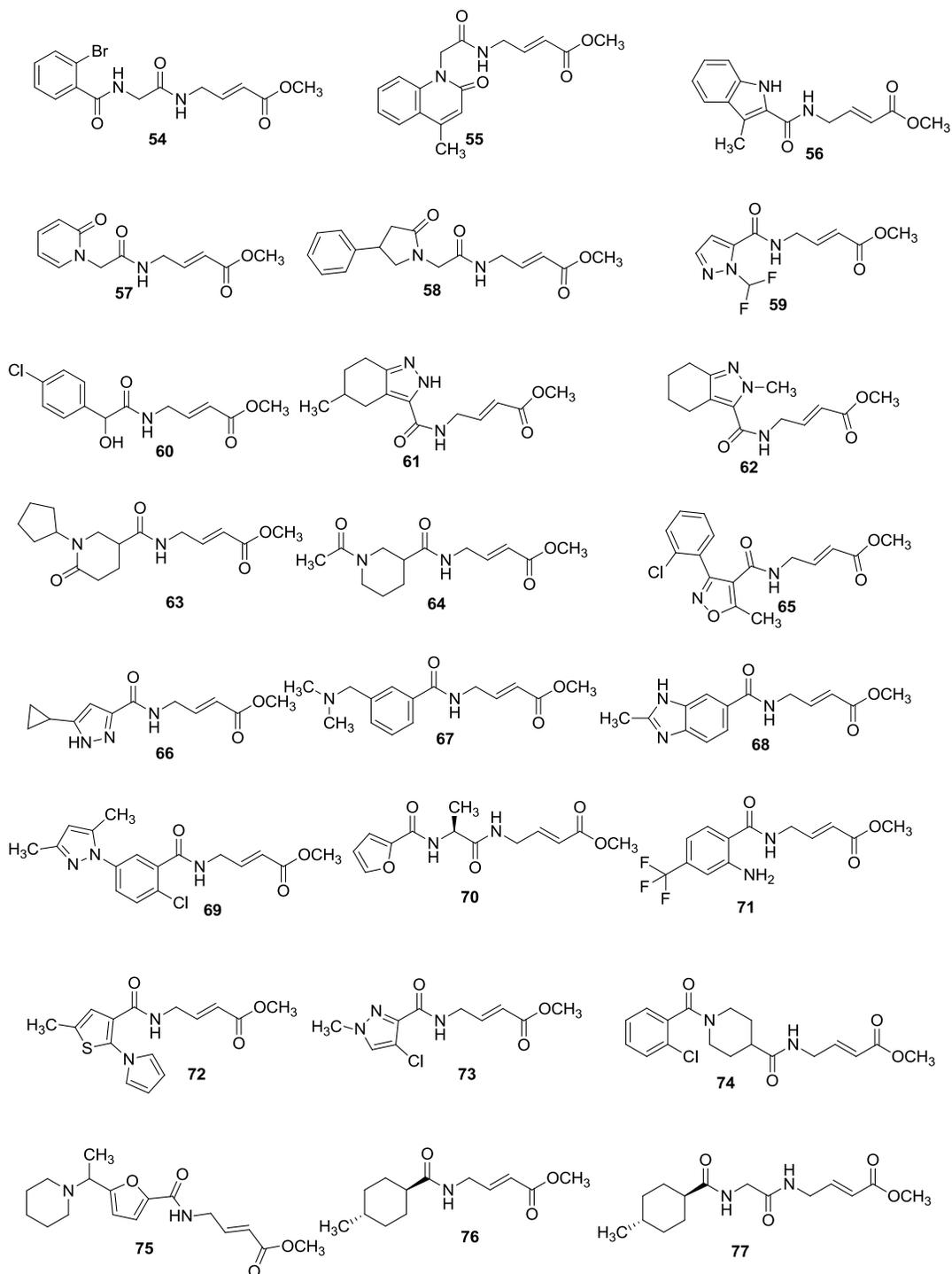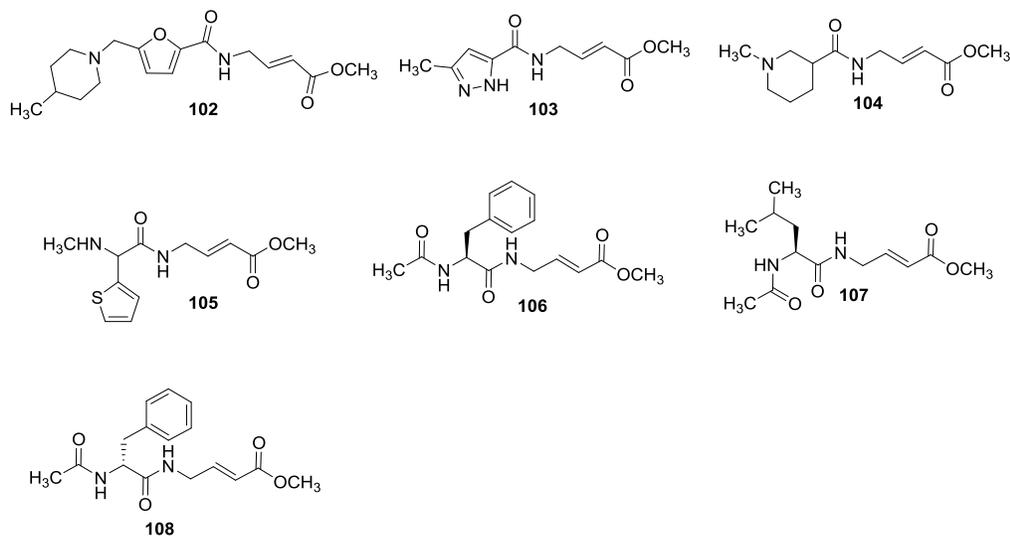
Characterization of compounds tested in enzymatic assays

6
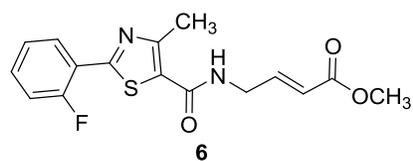
^{1}H NMR (500 MHz, CDCl$_3$) δ 8.29 (td, J = 7.7, 1.8 Hz, 1H), 7.41 (dddd, J = 8.7, 7.3, 5.3, 1.8 Hz, 1H), 7.24 (td, J = 7.9, 1.4 Hz, 1H), 7.17 (ddd, J = 11.4, 8.1, 1.2 Hz, 1H), 6.95 (dt, J = 15.7, 5.2 Hz, 1H), 6.07 − 5.75 (m, 2H), 4.20 (td, J = 5.6, 1.9 Hz, 2H), 3.71 (s, 3H), 2.75 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.29, 161.82, 161.22, 159.39, 156.23, 143.55, 132.07, 128.98, 125.44, 124.78, 122.00, 120.57, 116.22, 51.75, 40.69, 17.38. [M+Na]: 357.1 Da. HPLC purity: 95%.

7
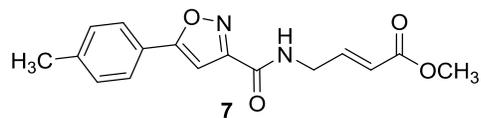
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.63 (d, J = 8.1 Hz, 2H), 7.23 (d, J = 8.0 Hz, 2H), 6.98 (t, J = 6.3 Hz, 1H), 6.93 (dt, J = 15.7, 5.1 Hz, 1H), 6.86 (s, 1H), 5.96 (d, J = 15.7 Hz, 1H), 4.20 (ddd, J = 6.5, 5.1, 1.9 Hz, 2H), 3.68 (s, 3H), 2.35 (s, 3H). ^{13}C

NMR (126 MHz, CDCl$_3$) δ 172.07, 166.25, 159.05, 158.64, 143.10, 141.27, 129.84, 125.90, 123.99, 121.99, 98.50, 51.73, 39.98, 21.55. [M+H]: 301.1 Da. HPLC purity: 97%.

8
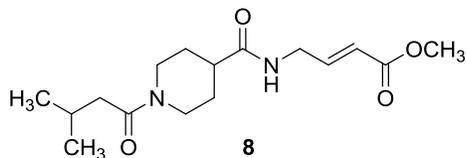
^{1}H NMR (500 MHz, CDCl$_3$) δ 6.87 (d, J = 15.7 Hz, 1H), 6.05 − 5.57 (m, 2H), 4.03 (t, J = 5.5 Hz, 2H), 3.71 (s, 3H), 3.03 (s, 1H), 2.63 (s, 1H), 2.44 − 2.27 (m, 1H), 2.19 (d, J = 6.8 Hz, 2H), 2.07 (dt, J = 13.4, 6.7 Hz, 1H), 1.87 (d, J = 11.3 Hz, 3H), 1.64 (s, 3H), 0.94 (d, J = 6.6 Hz, 6H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 173.93, 171.03, 166.31, 144.04, 121.52, 51.72, 43.11, 42.12, 40.03, 25.84, 22.76. [M+H]: 311.2 Da. HPLC purity: 98%.

19
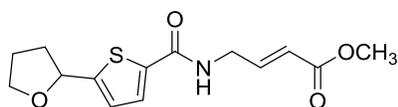
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.24 (d, J = 3.8 Hz, 1H), 6.80 (dt, J = 15.7, 5.1 Hz, 1H), 6.74 (dd, J = 3.8, 0.9 Hz, 1H), 6.08 − 5.96 (m, 1H), 5.81 (d, J = 15.7 Hz, 1H), 4.97 (t, J = 6.6 Hz, 1H), 4.18 − 3.99 (m, 2H), 3.89 (dd, J = 7.6, 5.9 Hz, 1H), 3.74 (dd, J = 7.5, 1.0 Hz, 1H), 3.56 (s, 3H), 2.34 − 2.04 (m, 1H), 1.93 − 1.69 (m, 4H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.38, 161.88, 153.36, 144.00, 136.10, 128.68, 123.75, 121.66, 68.64, 51.70, 40.47, 34.77, 25.84. [M+Na]: 318.1 Da. HPLC purity: 95%.

106
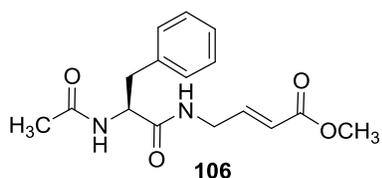
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.55 − 7.00 (m, 5H), 6.72 (dt, J = 15.7, 5.2 Hz, 1H), 6.07 (d, J = 7.8 Hz, 1H), 5.94 (s, 1H), 5.69 (dt, J = 15.9, 1.7 Hz, 1H), 4.60 (td, J = 8.0, 6.2 Hz, 1H), 4.04 − 3.80 (m, 2H), 3.71 (s, 3H), 3.21 − 2.78 (m, 2H), 1.98 (s, 3H). ^{13}C

NMR (126 MHz, CDCl$_3$) δ 170.85, 170.20, 166.27, 143.28, 136.39, 129.18, 128.87, 127.25, 121.61, 54.83, 51.66, 40.03, 38.24, 23.21. [M+Na]: 327.1 Da. HPLC purity: 95%.

107
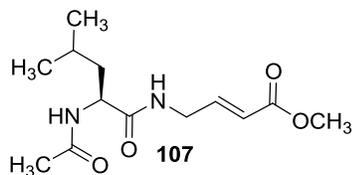
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.33 − 7.19 (m, 1H), 7.11 (dt, J = 15.8, 5.0 Hz, 1H), 6.52 (d, J = 8.4 Hz, 1H), 6.13 (dt, J = 15.9, 1.9 Hz, 1H), 4.74 (td, J = 8.5, 5.9 Hz, 1H), 4.24 (dt, J = 5.8, 3.0 Hz, 2H), 3.95 (s, 3H), 2.23 (s, 3H), 1.89 (tt, J = 13.2, 6.2 Hz, 2H), 1.84 − 1.71 (m, 1H), 1.17 (dd, J = 11.5, 6.1 Hz, 6H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 172.34, 170.54, 166.42, 143.91, 121.38, 51.69 (d, J = 7.5 Hz), 40.86, 40.05, 24.80, 23.12, 22.82, 22.25. [M+Na]: 293.1 Da. HPLC purity: 95%.

108
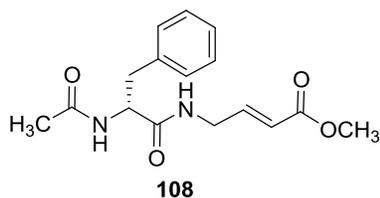
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.51 − 7.12 (m, 5H), 6.80 (dt, J = 15.7, 5.0 Hz, 2H), 6.61 (d, J = 8.0 Hz, 1H), 5.76 (dd, J = 15.8, 2.0 Hz, 1H), 4.78 (q, J = 7.6 Hz, 1H), 4.09 − 3.86 (m, 2H), 3.76 (s, 3H), 3.10 (dd, J = 7.5, 1.9 Hz, 2H), 2.00 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 171.24, 170.32, 166.35, 143.55, 136.46, 129.22, 128.70, 127.11, 121.40, 54.75, 51.63, 39.99, 38.56, 23.10. [M+Na]: 327.1 Da. HPLC purity: 99%.

CHAPTER 2: DISCOVERY AND OPTIMIZATION OF ELECTROPHILIC COMPOUNDS THAT REACT WITH NEDD4-1

2.1 INTRODUCTION

Initially discovered as the means by which cells tag proteins for degradation,[32] the ubiquitin-proteasome system (UPS) has since emerged as a critical mediator of cell signaling,[33] endocytosis,[33] and DNA repair.[34] The UPS covalently attaches ubiquitin (Ub), an 76 amino acid protein, to the ε-amino group of a lysine residue of a substrate protein in a three-step enzymatic cascade.[35] An E1 enzyme activates Ub by forming a thioester bond between the Ub C-terminal glycine and the E1 catalytic cysteine. Ub is then transthiolated onto the catalytic cysteine of an E2 ubiquitin-conjugating enzyme. Finally, an E3 ubiquitin ligase transfers Ub onto to the substrate lysine. It can subsequently build polyubiquitin chains on the substrate (Figure 2.1). E3 enzymes are the largest group in the cascade and the one most responsible for substrate specificity. E3 enzymes can be subdivided into two classes: the RING family (>600 members), which function as scaffolding proteins that bring the Ub-charged E2 and the substrate lysine into proximity,[36] and the much smaller group of HECT E3s (~30 members),[37] which contain a catalytic cysteine of their own that accepts Ub from the E2 to form an intermediate Ub~HECT E3 thioester bond before directly transferring Ub onto the substrate.[38]
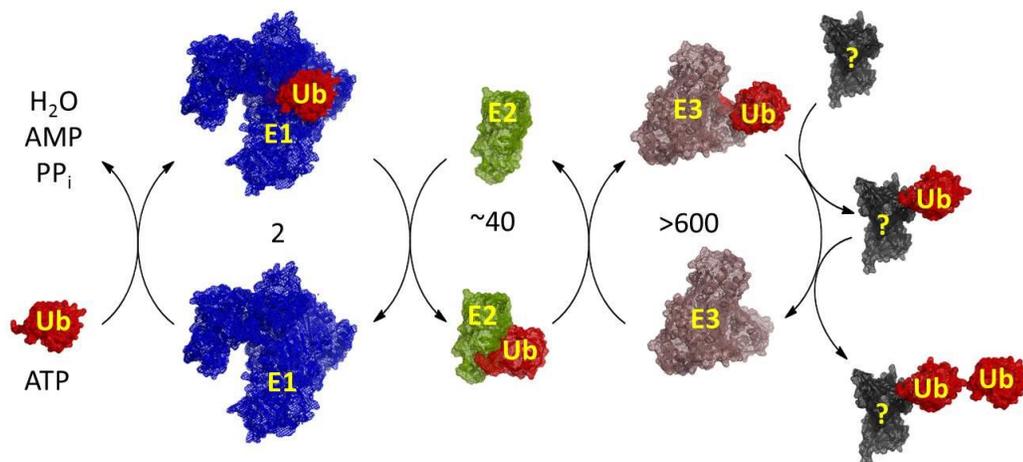
Figure 2.1 Summary of the ubiquitin activation and ligation cascade.

The UPS is rivaled only by protein kinases in its importance to cellular regulation, with ~19,000 known ubiquitination sites on ~5000 substrates.[39] However, due to the three-step activation cascade and the variety of polyUb chain linkages, it greatly exceeds the kinome in complexity, and is therefore more poorly understood. This disparity in understanding can also be attributed to a general lack of pharmacological tools to study the UPS,[40] whereas the study of protein kinases has benefited from ~19,000 potent and selective inhibitors.[41,42] The discovery of E3 inhibitors is hampered by the lack of a small molecule cofactor or substrate which can serve as a pharmacological starting point. E3s lack deep hydrophobic drug-binding pockets, and they mostly achieve their effects through protein-protein interactions, which are notoriously difficult to drug.[43]

We reasoned that HECT E3s might be more pharmacologically tractable since they contain a catalytic cysteine residue. This catalytic cysteine presents a potential handle for the discovery of covalent inhibitors. Therefore, we sought to develop a potent and selective covalent inhibitor of a HECT E3 using our covalent tethering library. We chose Nedd4-1 as a target HECT E3 because it regulates mammalian metabolism, growth, and development,[44] and is a promising drug target to treat cancers,[45,46] obesity,[47] Parkinson's disease,[48] and viral infections.[49,50] However, despite its promise as a therapeutic target, there are no validated small molecule inhibitors of Nedd4-1.[51]

The crystal structure of the Nedd4-1 catalytic HECT domain reveals that it has two important surface cysteines at functional interfaces.[52,53] The HECT domain contains the catalytic Cys^{867} in the C-lobe, which forms the thioester with the C-terminus of Ub, as well as a non-catalytic surface Cys^{627} in the N-lobe non-covalent ubiquitin binding site, which is important for polyubiquitination of protein substrates (Figure 2.1). Therefore, electrophilic fragments that react

with either of these cysteines could, respectively, either inhibit Nedd4-1 completely or disrupt its ability to form polyubiquitin chains.
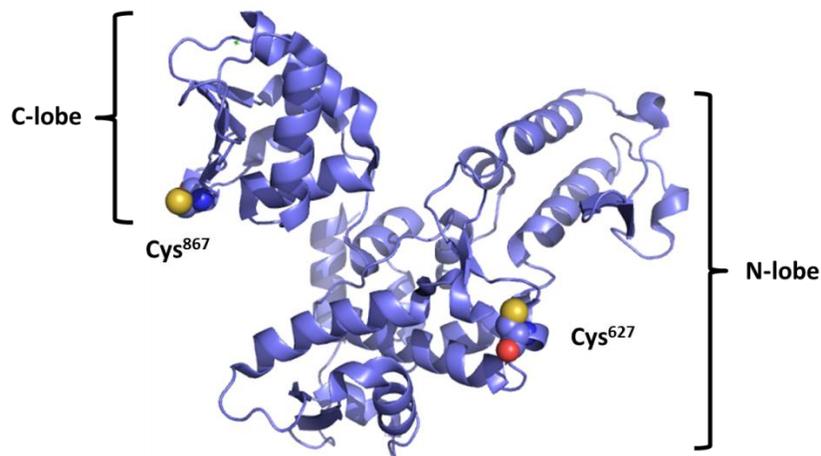
Figure 2.2 Crystal structure of the Nedd4-1 HECT domain.
There are two surface cysteines. The C-lobe contains the catalytic Cys^{867} and the N-lobe contains Cys^{627} in the non-covalent Ub binding site.

2.2 RESULTS AND DISCUSSION

2.2.1 Covalent fragment screen against Nedd4-1 and validation of hits

To search for Nedd4-1 inhibitors, we used the covalent tethering method described in chapter 1. The HECT domain of Nedd4-1 was treated with the same mixtures of electrophilic fragments, and compounds **30** and **43** were identified as weak covalent modifiers of cysteine residues in Nedd4-1 using mass spectrometry (Figure 2.3). These fragments reacted with Nedd4-1 in a time- and concentration-dependent manner, indicating that they bind Nedd4-1 irreversibly. Point mutation studies showed that **30** and **43** reacted with the non-catalytic N-lobe Cys^{627} and not the catalytic Cys^{867} (Figure 2.4). This was a somewhat surprising result, because the catalytic Cys^{867} of Nedd4-1 was approximately five-fold more reactive than Cys^{627} with the non-specific N-acetyl electrophile **114** (Figure 2.5). However, this demonstrates that compounds **30** and **43** are specific hits, since they are able to modify the less reactive non-catalytic cysteine in the

presence of the more reactive catalytic cysteine. Furthermore, compound **43** demonstrated an

initial structure-activity relationship (SAR), an additional indication that covalent labeling of

Nedd4-1 by **43** was specific (Figure 2.6).
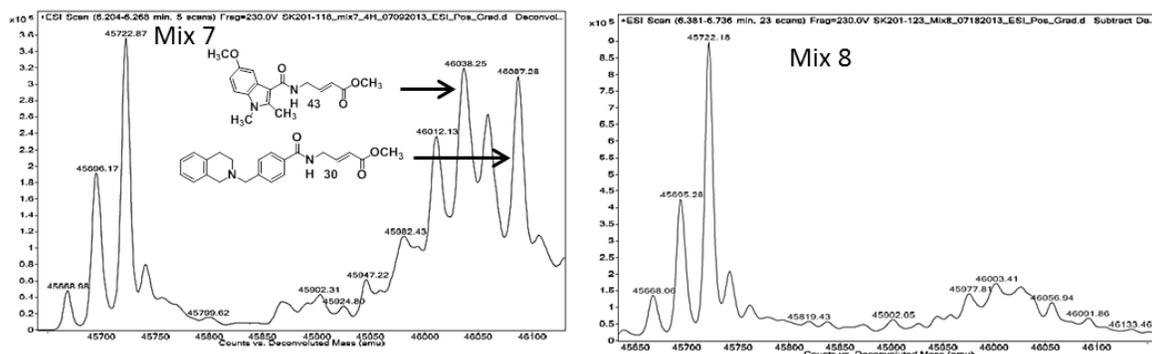
Figure 2.3 Screening of fragment library against Nedd4-1 HECT domain found two hits
Mixtures of ten fragments (100 µM each) were incubated with Nedd4-1 HECT domain for 4 h, followed by gel filtration and whole protein electrospray ionization MS (ESI-MS). Hits **30** and **43** were both in mix 7. A mix with no hits is shown for reference.
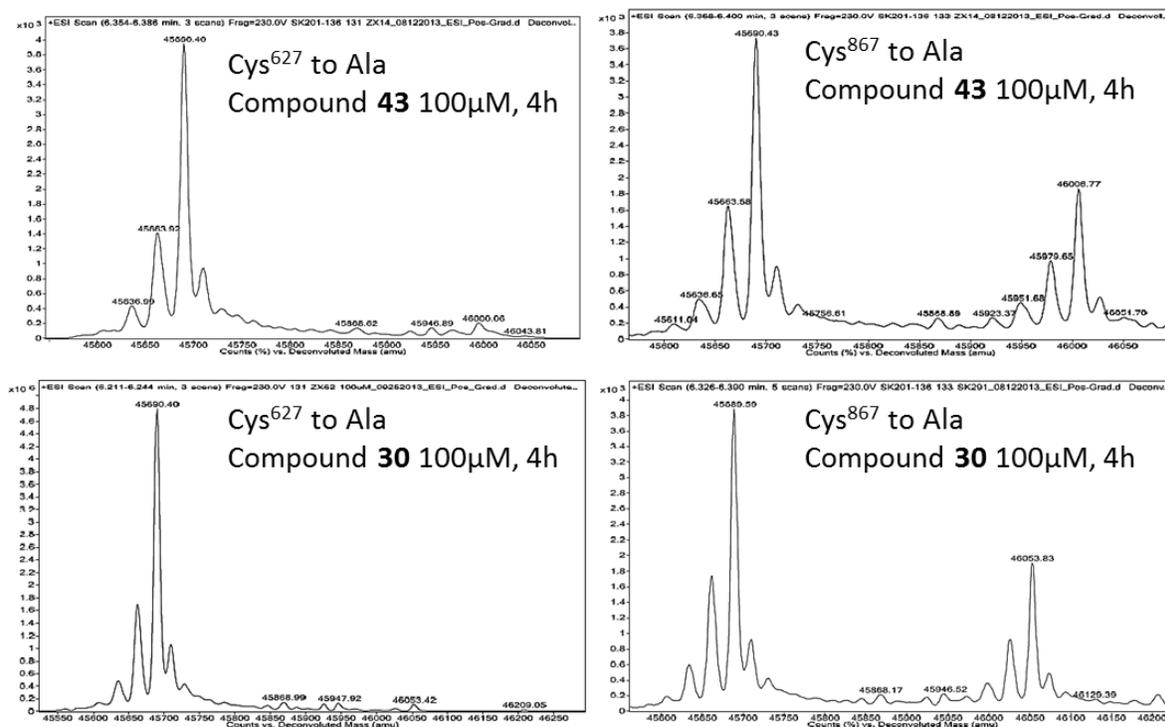
Figure 2.4 Compounds 30 and 43 selectively modify the non-catalytic Cys^{627} of Nedd4-1.
Cys^{627}Ala mutation impairs labeling by compounds **30** and **43**, while the Cys^{867}Ala mutation has no effect.
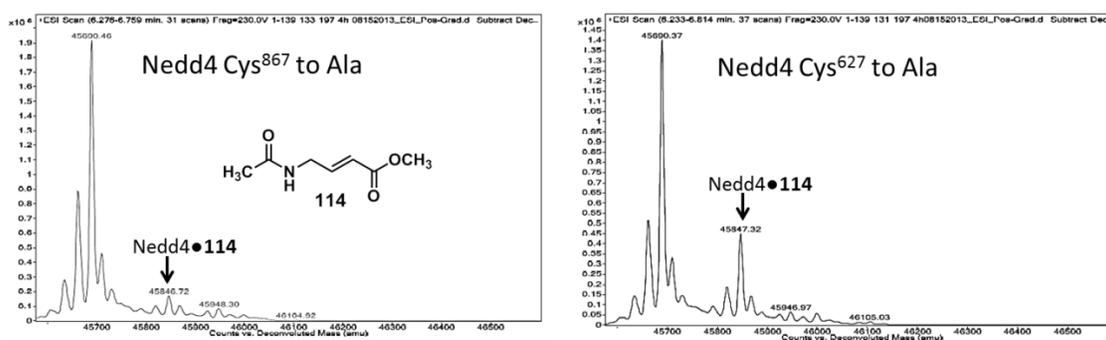
Figure 2.5 The catalytic Cys[867] of Nedd4-1 is more reactive with the non-specific N-acetyl electrophile 114 than Cys[627], as determined by the corresponding Cys to Ala mutations. Compound **114** at 1 mM in 1% DMSO was incubated with the indicated Nedd4-1 HECT domain mutant (10 μM) for 4h. Since compound **114** has no fragment directing group, any difference in reactivity between Cys[867] and Cys[627] with **114** should be due to an inherent difference in the reactivity of the cysteines with the electrophile itself.
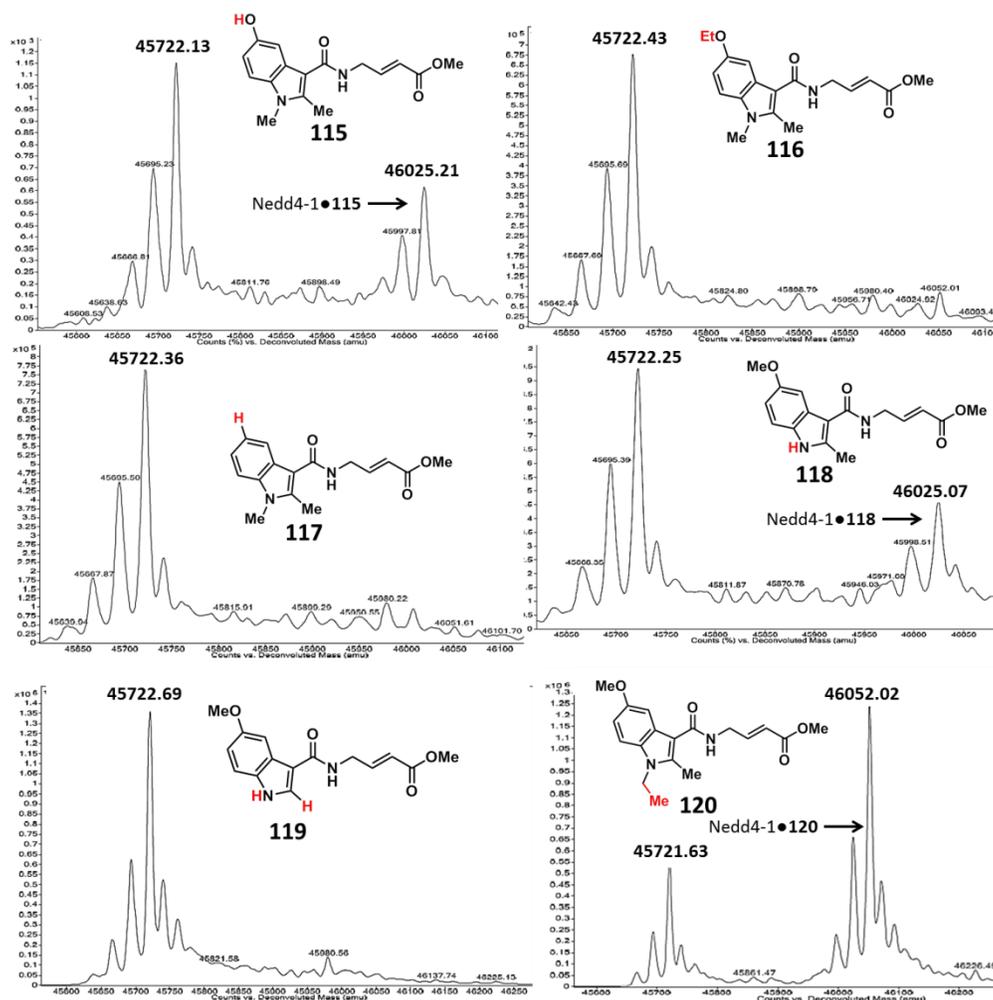
Figure 2.6 Compound 43 demonstrates a structure-activity relationship with Nedd4-1.
Nedd4-1 HECT domain (10 μM) was treated with the indicated compounds at 100 μM for 4 h. Notably, the 5-position of the indole does not tolerate a 5-CH_3-CH_2-O- substitution, while labeling is improved when N-CH_3 is replaced by N-CH_2-CH_3.

2.2.2 Crystal structure of fragment 43 bound to the Nedd4-1 HECT domain

To visualize the binding mode of compound **43**, we crystallized the Nedd4-1:**43** complex

and solved the structure to 2.44 Å resolution (PDB ID: 5C91) (Figure 2.7, Table 2.1). Notably,

this structure is the first of a HECT E3 bound to a small molecule, as well as the first structure of

any E3 ligase covalently bound to a small molecule. The overall conformation is virtually

identical to the previously reported structure of Nedd4-1[52,53] (root-mean-square deviation (rmsd):

0.316 Å) (Figure 2.8). Our structure confirms that **43** forms a stable covalent bond with Cys[627]

and reveals that the hydrophobic indole core of **43** is oriented towards a pocket of the N-lobe

formed by residues Leu[553], Glu[554], Asn[602], Tyr[604], Tyr[605], Leu[607], and Tyr[634] (Figure 2.7). This

ligand orientation explains why compound **116**, which contains an EtO- group on the indole

core, did not label Nedd4-1, since the ligand binding pocket cannot accommodate this sterically

bulkier group (Figure 2.7). The aromatic edge-to-face interactions of Tyr[605] and Tyr[634] with the

indole moiety of **43** provide further stabilization of the ligand conformation, while a hydrogen

bond between the backbone carbonyl oxygen of Tyr[605] and the amide NH of **43** positions the

connecting region between Cys[627] and the indole group. The ester methoxy group of **43** points

inward and towards the cavity formed by Gly[606]O, Asn[621]N, Asn[623]O, and Asn[623]Cβ, which are

within 3.4 – 4.4 Å of the methyl group. However, since the methyl ester group is freely rotatable

around the C-C bond, our crystallographic data cannot exclude a partial conformation in which

the methoxy and the carbonyl groups are exchanged.
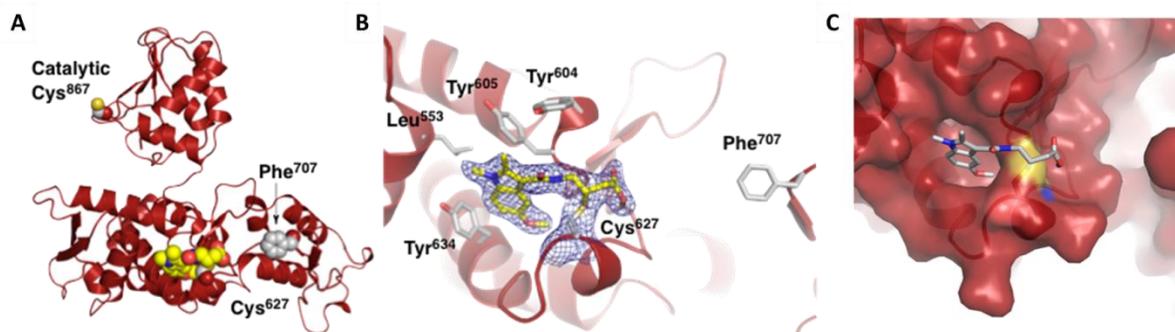
Figure 2.7 Crystal structure of fragment 43 covalently bound to Nedd4-1 HECT domain
A) Cartoon depiction of the crystal structure of Nedd4-1 bound to fragment hit **43** with key side chains and the fragment shown as spheres. B) Close-up view of the small molecule-binding site with the key side chains depicted as sticks and colored by atom type. The $2F_O - F_C$ electron density map (blue mesh, contoured at 1.0 σ) is presented for Cys^{627} and **43**. C) Surface representation of the hydrophobic pocket that **43** binds, which shows areas for potential fragment growth and areas that are sterically constrained.
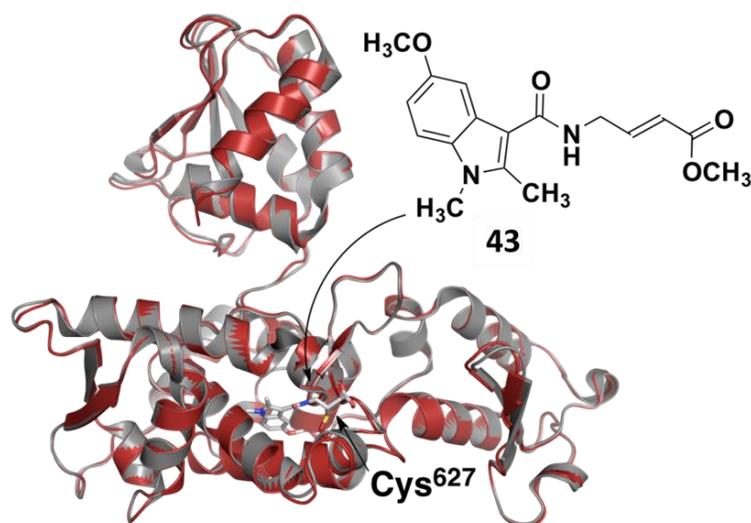
Figure 2.8 Superposition of Nedd4-1 HECT domain and the Nedd4-1:43 complex.
The proteins are depicted as cartoons and the inhibitor as well as the side chain of Cys^{627} are shown as sticks. The unbound Nedd4-1 HECT domain (PDB ID: 2XBF) is shown in red and Nedd4-1 bound to **43** is shown in gray.

	NEDD4-1•compound 43
Data collection	
Space group	*C*2
Cell dimensions	
a, b, c (Å)	175.20, 38.73, 60.36
α, β, γ (°)	90.00, 93.13, 90.00
Resolution (Å)	32.20 - 2.44 (2.50 – 2.44)

R_{merge} (%)	9.1 (60.2)
I / σ (I)	11.2 (2.0)
Completeness (%)	98.2 (98.7)
Multiplicity	3.2 (3.1)
CC(1/2)	0.994 (0.626)
Refinement	
Resolution (Å)	31.86 -2.44
No. reflections	14371
R_{work} / R_{free}	0.248 / 0.298
No. atoms	
Protein	3175
N-lobe (residues 519-780)	2239
C-lobe (residues 780-893)	936
Ligand/ion	23
Water	7
Average B-factors ($Å^2$)	
Protein	50.842
N-lobe (residues 519-780)	42.472
C-lobe (residues 780-893)	70.864
Ligand/ion	63.217
Water	35.437
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	1.036
Ramachandran plot (%)	95.4 / 4.3 / 0.3

Table 2.1 Crystallographic table for Nedd4-1 HECT covalently bound to compound 43

2.2.3 Optimization of compound 43 into a more potent binder

Our initial experiments showed that labeling of Nedd4-1 with **43** could be completely inhibited by 60 μM of Ub (Figure 2.9), which is the approximate concentration of Ub in cells.[54] Therefore further improvements in the potency of **43** were necessary. Since the N-methyl group in **43** was shown to tolerate substitutions, we prepared a series of N-substituted analogues **121**-**124** (Figure 2.10). Of these, the N-cyclopentyl analogue **123** was the most potent as measured by degree of labeling of Nedd4-1 in the ESI-MS assay.
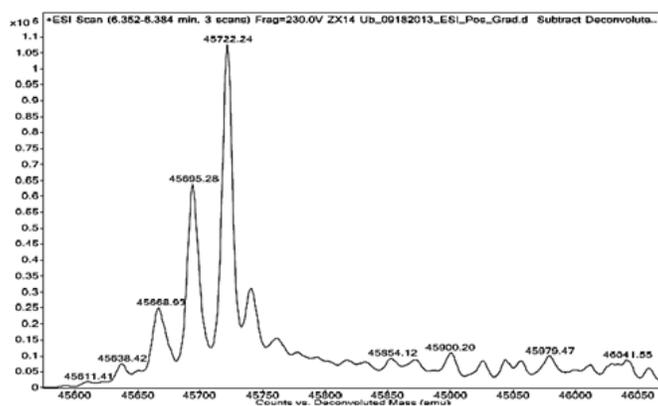
Figure 2.9 Inhibition of Nedd4-1 labeling with compound 43 in the presence of ubiquitin.
Compound **43** (100μM) in 1% DMSO was incubated with Nedd4-1 HECT domain (10 μM) and ubiquitin (60 μM) for 4 h, followed by gel filtration and whole protein ESI-MS.
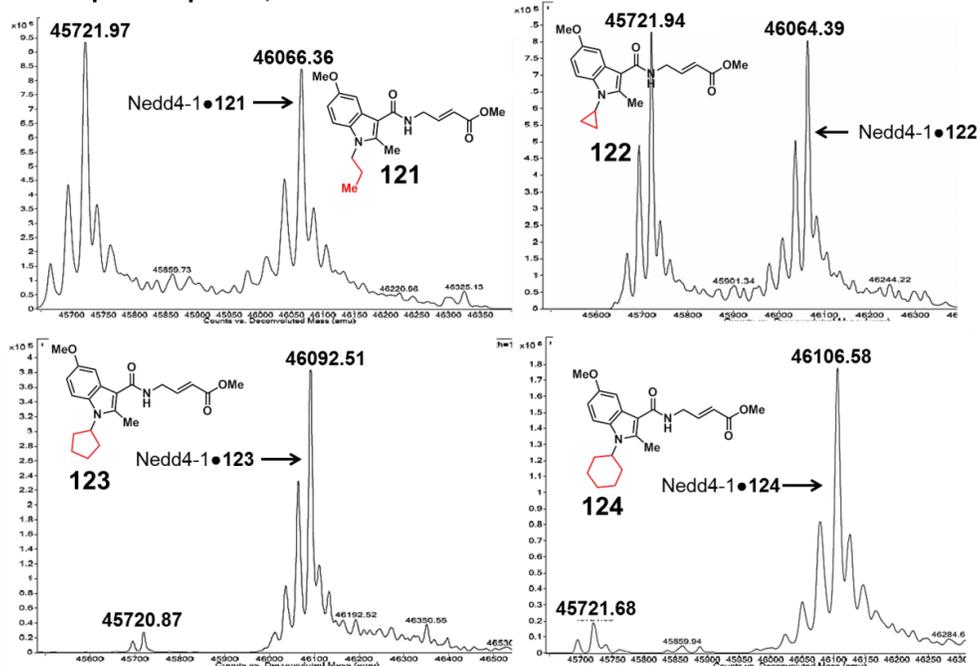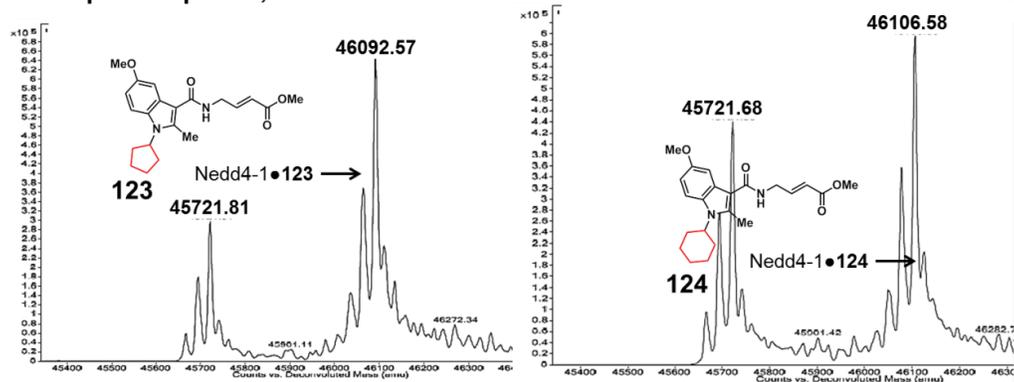
Figure 2.10 SAR studies of N-substituted indole analogs.
A) Analogs **121**-**124** were designed to improve the potency of compound **43**. Nedd4-1 HECT domain (10 μM) was treated with the indicated compounds in 1% DMSO at 100 μM for 1 h, followed by gel filtration and whole protein ESI-MS. B) Studies with 50 μM compound show that **123** is slightly more potent than **124**. Nedd4-1 HECT domain (10 μM) was treated with the indicated compounds in 1% DMSO at 50 μM for 1 h, followed by gel filtration and whole protein ESI-MS.

An extensive SAR using commercially available and synthetically tractable analogs was explored to further improve the potency of **123**, but most of these analogs had limited to no reactivity with Nedd4-1 (Figure 2.11). The 6-bromo indole analog and a benzofuran analog only

minimally reacted with Nedd4-1. Replacing the N-cyclopentyl group with heterocycles or larger

hydrophobic groups also reduced labeling. The electrophile did not tolerate the addition of larger

substituents or replacement of the acrylate with an acrylamide. Indole 2-position alkyl analogs

were equally potent to **123** but no improvement in potency was seen.
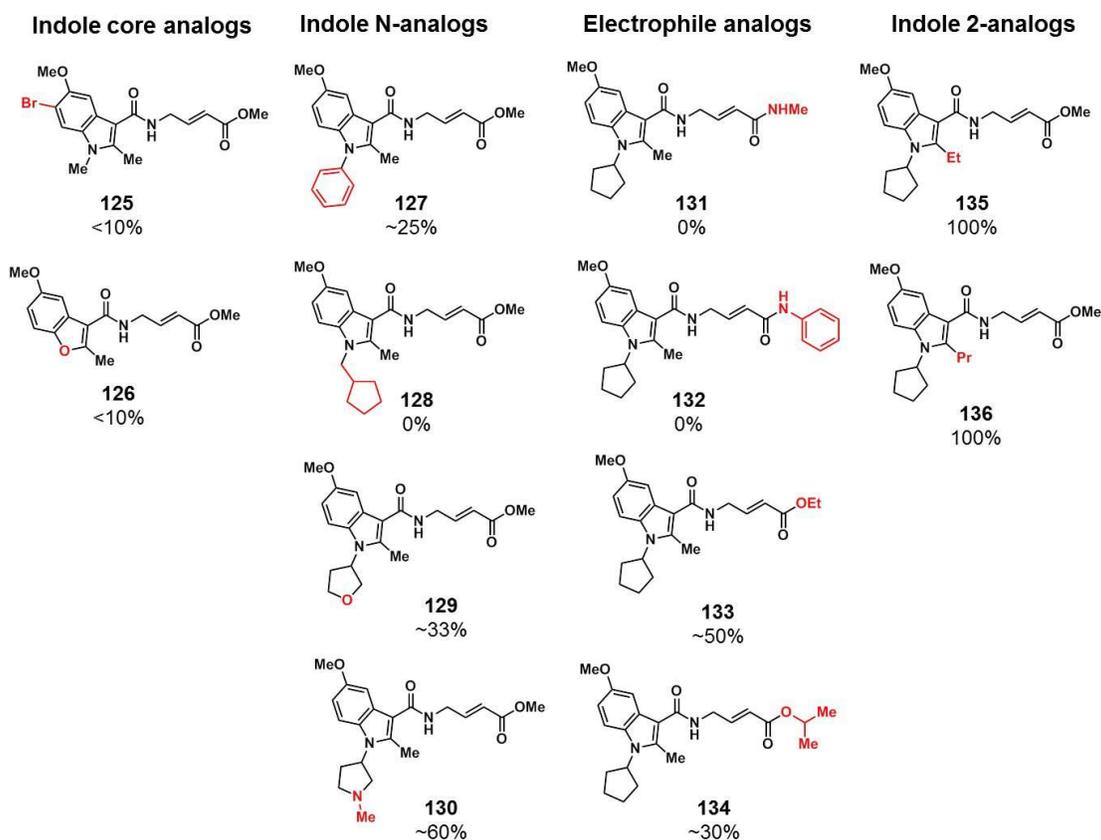
Figure 2.11 Additional analogs of 123 do not improve its potency
Further modifications of the indole core, the electrophile, and indole N-position result in reduced
labeling of Nedd4-1. Alkyl analogs at the indole 2-position were equally potent to **123**.
Percentages are the percent of Nedd4-1 HECT domain that is labeled after 1h with 100 μM of
the indicated compound.

Indicative of its selectivity, compound **123** did not label the HECT domains of the related

ligases WWP1 or E6-AP, which have catalytic cysteines but do not have cysteines at this

position labeled by **123** (Figure 2.12). However, it was effective at labeling the highly

homologous HECT ligase Nedd4-2, which has an almost identical non-covalent Ub-binding site

and a cysteine in the N-lobe ubiquitin binding site (Figure 2.13). Furthermore, compound **123**

did not react with the deubiquitinase USP8, Human Rhinovirus 3C protease, the E1 enzyme

Ube1, or the E2 enzyme UbcH5a, all of which have reactive catalytic cysteines.
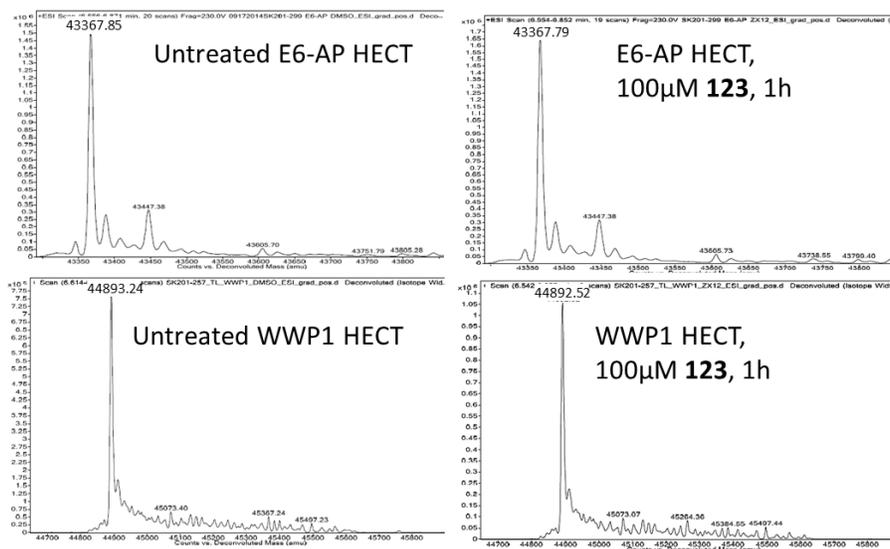
Figure 2.12 Compound 123 does not label the HECT domains of E6-AP and WWP1.
Compound **123** at 100 µM in 1% DMSO was incubated with the catalytic domain of the indicated HECT E3 (10 µM) for 1 h, followed by gel filtration and whole protein ESI-MS.
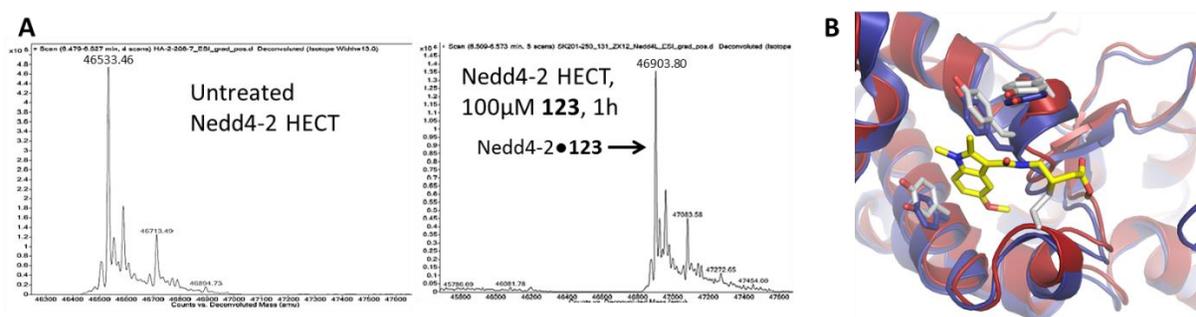
Figure 2.13 Compound 123 reacts with the homologous HECT domain of Nedd4-2.
A) Compound **123** at 100 µM in 1% DMSO was incubated with the catalytic HECT domain of Nedd4-2 (10 µM) for 1 h, followed by gel filtration and whole protein ESI-MS. B) Superposition of Nedd4-2 (PDB ID 2ONI: blue) and the binding site of **43** in the Nedd4-1:**43** complex (red); the protein is depicted as a cartoon, the inhibitor and the side chains of Cys[627], Tyr[604], Tyr[605], Tyr[634], Tyr[659] (2ONI), Tyr[660] (2ONI), Tyr[689] (2ONI) are shown as sticks.

2.3 CONCLUSION

In summary, we successfully screened our covalent fragment library against the catalytic HECT domain of the therapeutically relevant E3 Nedd4-1 and found 2 hits. These hits were unique from the 3 papain hits, indicating the specificity of our covalent tethering method. Surprisingly, we found that our hits did not label the more reactive catalytic cysteine of Nedd4-1, but instead another surface cysteine near the N-lobe non-covalent ubiquitin binding site. We were able to crystalize the covalent complex between the indole hit **43** and the Nedd4-1 HECT domain and solve its structure. This is the first structure of a HECT E3 bound to a small molecule, proving that these challenging drug targets are at least "ligandable".[55] The structure allowed us to optimize the fragment into compound **123**, which labels Nedd4-1 much more potently. However, it was difficult to optimize compound **123** further because most positions on the electrophile or the indole core did not tolerate substitutions. Compound **123** demonstrated good selectivity *in vitro* since it only reacted with the highly homologous Nedd4-2 and not a variety of other HECT E3s, UPS enzymes, and proteases with catalytic cysteines. Although compound **123** cannot completely inhibit Nedd4-1 since it does not react with its catalytic cysteine, we reasoned that it might be able to inhibit non-covalent binding of Ub to the Nedd4-1 HECT N-lobe and disrupt polyubiquitination. This mode of inhibition will be discussed in the next chapter.

2.4 MATERIALS AND METHODS

Recombinant expression of Nedd4-1 (HECT) in *E. coli* Nedd4-1 in a PGEX6P1 vector plasmid (GST-Nedd4-1) was transformed into BL21 cells (Novagen). 1L TB media containing 100μg/ml ampicillin was inoculated with 50 mL overnight cell culture and incubated at 37°C until OD reached ~3. Then, IPTG (1.0 mM final concentration) was added to the cell culture

media at 18°C, followed by 16 hour incubation at the same temperature. Cells were then harvested and lysed by sonication in phosphate buffered saline (PBS) with protease inhibitors (Complete Mini Protease Inhibitor Cocktail, Roche). The supernatant was incubated with glutathione agarose beads (Pierce Biotechnology) for 1 hour at 4°C. The beads were washed three times with PBS and incubated with PreScission Protease (GE Healthcare) for 4h at 23°C to elute Nedd4-1 (elution buffer: 50mM HEPES, 150 mM NaCl, 0.1 mM EDTA 1mM DTT). Mutant plasmids were prepared with Agilent QuickChange kit.

Nedd4-1 HECT sequence:

GPLGSRDYKRKYEFFRRKLKKQNDIPNKFEMKLRRATVLEDSYRRIMGVKRADFLKAR
LWIEFDGEKGLDYGGVAREWFFLISKEMFNPYYGLFEYSATDNYTLQINPNSGLCNEDH
LSYFKFIGRVAGMAVYHGKLLDGFFIRPFYKMMLHKPITLHDMESVDSEYYNSLRWILE
NDPTELDLRFIIDEELFGQTHQHELKNGGSEIVVTNKNKKEYIYLVIQWRFVNRIQKQMA
AFKEGFFELIPQDLIKIFDENELELLMCGLGDVDVNDWREHTKYKNGYSANHQVIQWF
WKAVLMMDSEKRIRLLQFVTGTSRVPMNGFAELYGSNGPQSFTVEQWGTPEKLPRAHT
CFNRLDLPPYESFEELWDKLQMAIENTQGFDGVD

Irreversible Tethering Screening Assay with Nedd4-1 HECT domain 10 μM of Nedd4-1 HECT domain in 50mM HEPES 150mM NaCl 0.1 mM EDTA 1mM DTT pH 7.5 was treated with a mixture of ten fragments (from 10 mM each DMSO stock solution mixtures; final concentrations: 100μM of each fragment, and 1% DMSO). Fragment structures and mixture compositions were the same as reported previously (ref. 14 of the main text). The reaction mixture was incubated for 4h at 23°C before being passed through Zeba gel filtration columns (Thermo, 7K MWCO) to remove unreacted fragments. The protein solution was then immediately analyzed by whole protein LC/ESI-MS.

Preparation of fully labeled Nedd4-1•inhibitor complexes for crystallography 10 μM of

Nedd4-1 HECT domain in 50mM HEPES 150mM NaCl 0.1 mM EDTA 1mM DTT pH 7.5 was

treated with 1mM of inhibitor **1** (from 100 mM DMSO stock solution; final concentrations: 1mM

of inhibitor, 1% DMSO, and 0.2% CHAPS to solubilize the inhibitors at 1mM). The reaction

mixture was incubated for 4h at 23°C before being passed through Zeba gel filtration columns

(Thermo, 7K MWCO) to remove unreacted inhibitor.   The protein solution was then

immediately used for crystallization or enzymatic assay.

Crystallization Crystals of the Nedd4 HECT:inhibitor **1** complex were obtained by the sitting-

drop vapor diffusion method using MiTeGen - XtalQuest Plates with a 1:1 ratio of protein

(6.3mg/mL) and reservoir solution at 20 °C. The precipitant was similar to that used

previously,[52] and consisted of 100 mM 2-(N-morpholino)ethanesulfonic acid (MES), pH 6.5, 35

mM $CaCl_2$, 5 mM tris(2-carboxyethyl)phosphine (TCEP), 6% polyethylene glycol (PEG) 400.

Crystals were soaked in cryoprotectant (100 mM MES, pH 6.0, 6% PEG 400, 20% ethylene

glycol) for 1 min, mounted on loops, and flash frozen in $N_{2(l)}$.

Data Collection and Structure Determination Native data sets were collected using

synchrotron radiation at the LS-CAT 21-ID-D beamline at the Advanced Photon Source,

Argonne National Laboratory, using a Mar 300 CCD detector. The datasets were processed using

Xia2[56] and solved by molecular replacement using Phaser.[57] The coordinates of the Nedd4

HECT structure with Protein Data Bank (PDB) ID 2XBF were used as a starting model.

Descriptions of the inhibitor and link were generated with the program JLigand[58] utilizing the

appropriate library obtained from the Grade Web Server. Model building and refinement were

performed with Coot[59] and REFMAC5,[60,61] respectively. The new chiral center generated upon

covalent binding of the inhibitor to the Cys side chain was initially modeled as both *S* and *R*

enantiomers. Since the refinement with the *S* enantiomer resulted in a lower R_{work}/R_{free}, we modeled this center as the *S* stereoisomer. However, we note that further experimental evidence is necessary to determine the absolute stereochemistry at this site. Translation liberation screw-rotation (TLS)[60,62] parameters and restrained refinement options in REFMAC5 were used for the final refinement cycles. Ramachandran plots were calculated with PROCHECK,[60,63] and validation was performed using both PHENIX[64] and SFCHECK.[65] Data collection and refinement statistics are shown in Table S1. Electron density maps were calculated using FFT,[60,66] and figures were prepared using PyMOL. The atomic coordinates have been deposited in the PDB, Research Collaboratory for Structural Bioinformatics at Rutgers University, ID 5C91.

2.5 SYNTHESIS
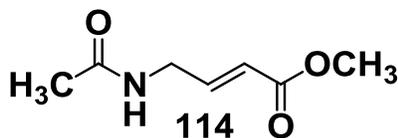
Synthesis of 114 (methyl (E)-4-acetamidobut-2-enoate) 5, TFA (50mg, 0.218 mmol) and triethylamine (77.8µL, 0.558 mmol) were dissolved in anhydrous CH_2Cl_2 (2.4 mL, 0.09M), then acetic anhydride (32.1µL, 0.34 mmol) was added dropwise. The reaction was stirred at 23°C for 24h, at which point TLC showed conversion to product. The reaction was quenched with 10mL saturated NH_4Cl, then extracted with 10% CH_3OH/CH_2Cl_2 (3x10mL). The combined organic layers were dried over $MgSO_4$, filtered, and evaporated. Purification with flash column chromatography with CH_3OH/CH_2Cl_2 (CH_3OH gradient 0→5%) yielded compound **114** (23.3 mg, 68% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 6.93 (dt, J = 15.7, 5.1 Hz, 1H), 5.95 (dt, J = 15.8,

1.8 Hz, 1H), 5.68 (s, 1H), 4.08 (td, J = 5.8, 1.8 Hz, 2H), 3.76 (s, 3H), 2.07 (s, 3H). ^{13}C NMR

(126 MHz, CDCl$_3$) δ 169.98, 166.39, 144.13, 121.48, 51.70, 40.21, 23.14. [M+Na]: 179.828 Da.

Synthesis of 30, 43, 115-124
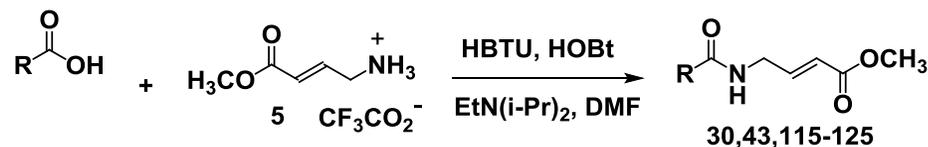
The commercially available carboxylic acid starting material (0.35 mmol) was dissolved in

dimethylformamide (0.2M, 1.75 mL), then **5**, TFA (80.2 mg, 0.35 mmol), HBTU (128mg, 0.34

mmol), and HOBT (51.8 mg, 0.38 mmol) were added, followed by diisopropylethylamine (175

μL, 1.047 mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to

product. The reaction was quenched with H$_2$O (5mL) and extracted with DCM (3×5mL). The

combined organic layers were washed with 1M HCl (10mL), saturated aqueous NaHCO$_3$

(10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over MgSO$_4$, filtered,

and evaporated. Purification with flash column chromatography with CH$_3$OH/CH$_2$Cl$_2$ (CH$_3$OH

gradient 0→5%) yielded compounds **30,43,115-124**.
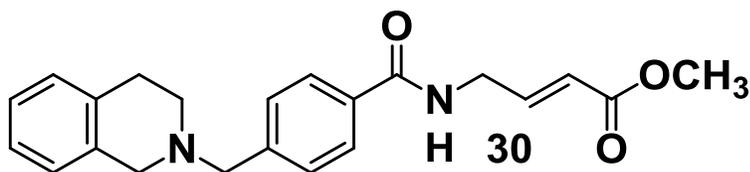
Methyl(E)-4-(4-((3,4-dihydroisoquinolin-2(1H)-yl)methyl)benzamido)but-2-enoate

(compound 30) Carboxylic acid starting material purchased from ChemBridge. 103.22 mg,

80.7% yield. ^{1}H NMR (500 MHz, DMSO-d$_6$) δ 8.82 (t, J = 5.7 Hz, 1H), 7.87 (d, J = 8.2 Hz, 2H),

7.47 (d, J = 8.2 Hz, 2H), 7.18 − 7.04 (m, 3H), 7.04 − 6.86 (m, 2H), 5.92 (dt, J = 15.7, 1.9 Hz,

1H), 4.23 − 3.99 (m, 2H), 3.71 (s, 2H), 3.66 (s, 3H), 3.55 (s, 2H), 2.82 (d, J = 5.8 Hz, 2H), 2.69

(s, 3H). ^{13}C NMR (126 MHz, DMSO-d$_6$) δ 166.08, 165.89, 146.28, 142.08, 134.66, 134.00, 132.69, 128.51, 128.42, 127.26, 126.32, 125.97, 125.45, 119.93, 61.35, 55.40, 51.37, 50.25, 38.21, 28.64. [M+H]: 363.255 Da.
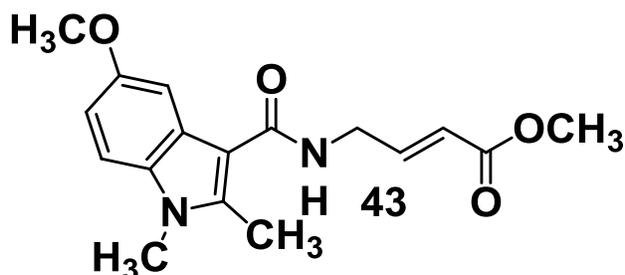
Methyl(E)-4-(5-methoxy-1,2-dimethyl-1H-indole-3-carboxamido)but-2-enoate (compound 43) Carboxylic acid starting material purchased from ChemBridge. 55.6 mg, 50.2% yield. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.28 − 7.21 (m, 2H), 7.11 (dt, J = 15.7, 4.9 Hz, 1H), 6.91 (dd, J = 8.8, 2.4 Hz, 1H), 4.34 (ddd, J = 6.1, 4.9, 2.0 Hz, 2H), 3.91 (s, 3H), 3.76 (s, 3H), 3.70 (s, 3H), 2.74 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.57, 166.28, 155.32, 145.26, 142.70, 131.76, 125.59, 121.15, 110.59, 110.30, 106.82, 101.91, 56.05, 51.67, 40.16, 29.68, 11.84. [M+Na]: 339.244 Da.
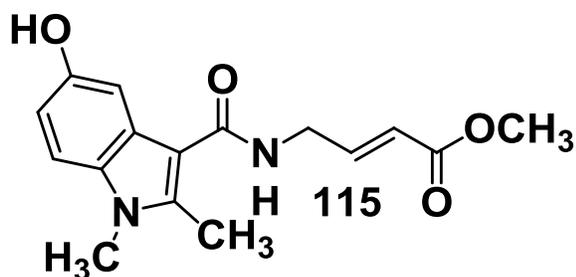
Methyl(E)-4-(5-hydroxy-1,2-dimethyl-1H-indole-3-carboxamido)but-2-enoate (compound 115) Carboxylic acid starting material purchased from Chembridge. 26.2 mg, 24.7% yield. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.19 − 7.10 (m, 2H), 7.04 (dt, J = 15.7, 5.1 Hz, 1H), 6.78 (dd, J = 8.7, 2.4 Hz, 1H), 6.01 (dt, J = 15.6, 1.9 Hz, 1H), 4.27 (td, J = 5.5, 1.9 Hz, 2H), 3.72 (s, 3H), 3.64 (s, 3H), 2.68 (s, 3H).

Methyl(E)-4-(5-ethoxy-1,2-dimethyl-1H-indole-3-carboxamido)but-2-enoate (**compound 116**) Carboxylic acid starting material purchased from Chembridge. 48.3 mg, 41.6% yield. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.23 − 7.16 (m, 2H), 7.06 (dt, J = 15.7, 5.0 Hz, 1H), 6.91 − 6.79 (m, 1H), 6.02 (d, J = 15.7 Hz, 1H), 5.99 − 5.91 (m, 1H), 4.28 (ddd, J = 5.9, 5.0, 1.9 Hz, 2H), 4.08 (q, J = 7.0 Hz, 2H), 3.72 (s, 3H), 3.65 (s, 3H), 2.69 (s, 3H), 1.43 (t, J = 7.0 Hz, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.56, 166.27, 154.61, 145.27, 142.83, 131.80, 125.59, 121.15, 111.08, 110.22, 106.76, 103.03, 64.42, 51.62, 40.16, 29.62, 15.06, 11.79. [M+Na]: 352.952 Da.

Methyl(E)-4-(1,2-dimethyl-1H-indole-3-carboxamido)but-2-enoate (**compound 117**) Carboxylic acid starting material purchased from Chembridge. 71.05 mg, 70.7% yield. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.75 − 7.58 (m, 1H), 7.41 − 7.28 (m, 1H), 7.27 − 7.15 (m, 3H), 7.06 (dt, J = 15.7, 5.0 Hz, 1H), 6.08 (s, 1H), 6.02 (dt, J = 15.7, 1.9 Hz, 1H), 4.30 (ddd, J = 5.9, 5.0, 1.9 Hz, 2H), 3.72 (s, 3H), 3.69 (s, 3H), 2.73 (s, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.58, 166.17, 145.14, 142.87, 136.52, 124.83, 121.31, 118.31, 109.73, 106.97, 51.64, 40.18, 29.53, 11.67. [M+Na]: 308.858 Da.

Methyl(E)-4-(5-methoxy-2-methyl-1H-indole-3-carboxamido)but-2-enoate (compound 118)

Carboxylic acid starting material purchased from Matrix Scientific. 43.3 mg, 40.88%. ^{1}H NMR

(400 MHz, CDCl$_3$) δ 7.23 − 7.17 (m, 2H), 7.05 (dt, J = 15.7, 5.0 Hz, 1H), 6.82 (dd, J = 8.7, 2.4

Hz, 1H), 6.03 (d, J = 15.7 Hz, 0H), 5.96 (t, J = 6.1 Hz, 1H), 4.29 (ddd, J = 6.0, 5.0, 1.9 Hz, 2H),

3.85 (s, 3H), 3.72 (s, 3H), 2.68 (s, 3H).

Methyl(E)-4-(5-methoxy-1H-indole-3-carboxamido)but-2-enoate **(compound** **119)**

Carboxylic acid starting material purchased from Sigma-Aldrich. 17.1 mg, 16.9% yield. [1]H

NMR (400 MHz, CDCl$_3$) δ 8.36 (s, 1H), 7.66 (d, J = 3.0 Hz, 1H), 7.55 (d, J = 2.5 Hz, 1H), 7.30

(dd, J = 8.9, 0.6 Hz, 1H), 7.04 (dt, J = 15.7, 5.0 Hz, 1H), 6.92 (dd, J = 8.9, 2.4 Hz, 1H), 6.03 (d, J

= 15.7 Hz, 1H), 5.97 (s, 1H), 4.28 (ddd, J = 6.0, 5.0, 1.9 Hz, 2H), 3.88 (s, 3H), 3.72 (s, 3H). [13]C

NMR (126 MHz, CDCl$_3$) δ 166.93, 166.00, 155.39, 145.31, 131.49, 131.34, 128.13, 127.96,

125.89, 125.86, 113.09, 112.62, 112.57, 110.68, 110.63, 102.23, 55.85, 51.68, 39.96. [M+Na]:

310.842 Da.

Methyl(E)-4-(1-ethyl-5-methoxy-2-methyl-1H-indole-3-carboxamido)but-2-enoate

(compound 120) Carboxylic acid starting material purchased from Enamine. 68.95 mg, 59.4% yield. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.22 (d, J = 10.8 Hz, 1H), 7.06 (dt, J = 15.7, 5.0 Hz, 1H), 6.86 (dd, J = 8.8, 2.4 Hz, 1H), 6.03 (d, J = 15.7 Hz, 1H), 5.96 (t, J = 5.9 Hz, 1H), 4.29 (ddd, J = 5.9, 5.0, 1.9 Hz, 2H), 4.11 (t, J = 7.2 Hz, 2H), 3.86 (s, 3H), 3.71 (s, 3H), 2.69 (s, 3H), 1.33 (t, J = 7.2 Hz, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.50, 155.36, 144.97, 142.00, 130.63, 125.84, 121.32, 110.71, 110.37, 102.00, 56.06, 51.65, 40.27, 37.98, 29.73, 14.94, 11.59. [M+Na]: 352.918 Da.
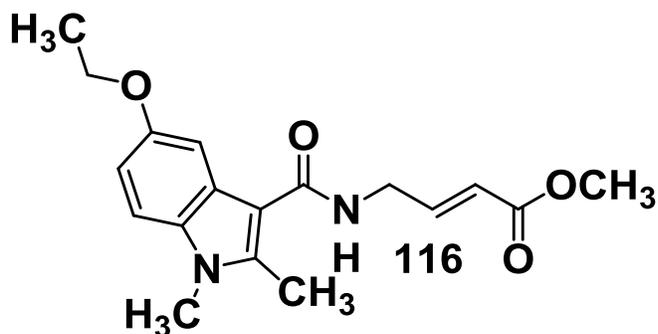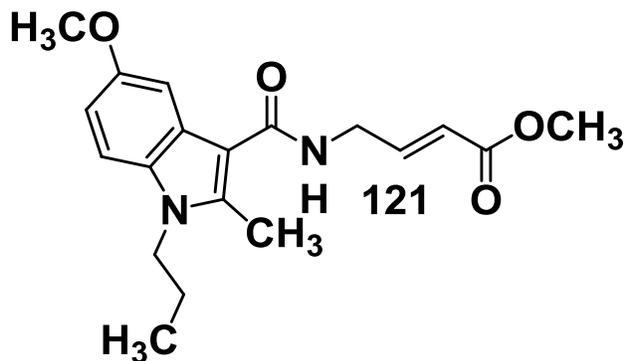

Methyl(E)-4-(5-methoxy-2-methyl-1-propyl-1H-indole-3-carboxamido)but-2-enoate

(Compound 121) Carboxylic acid starting material purchased from Enamine. 57.4 mg, 47.6% yield. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.24 (d, J = 1.5 Hz, 1H), 7.10 (dt, J = 15.7, 4.9 Hz, 1H), 6.89 (dd, J = 8.8, 2.4 Hz, 1H), 6.07 (d, J = 15.7 Hz, 1H), 6.03 (dd, J = 12.2, 5.7 Hz, 1H), 4.33

(ddd, J = 6.0, 4.9, 1.9 Hz, 2H), 4.06 (t, J = 7.4 Hz, 2H), 3.89 (s, 3H), 3.75 (s, 3H), 2.72 (s, 3H), 1.81 (q, J = 7.4 Hz, 2H), 0.98 (t, J = 7.4 Hz, 3H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.54, 166.34, 155.26, 145.23, 142.21, 131.15, 125.77, 121.19, 110.62, 110.57, 106.90, 101.93, 56.05, 51.63, 44.83, 40.17, 23.16, 11.80, 11.44. [M+Na]: 366.934 Da.
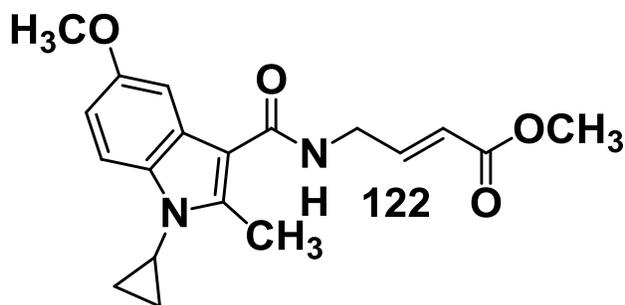
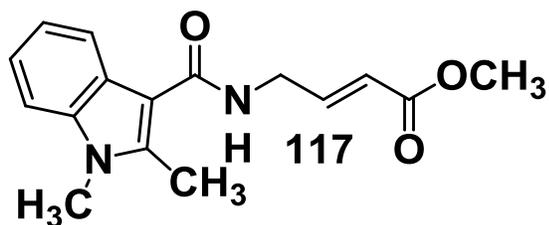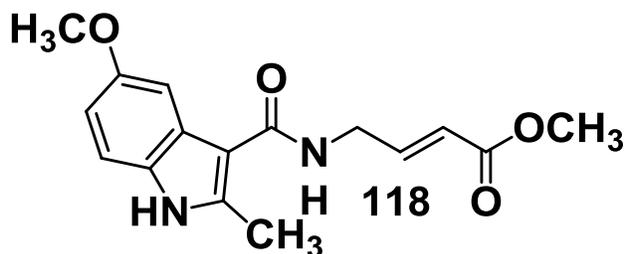**Methyl(E)-4-(1-cyclopropyl-5-methoxy-2-methyl-1H-indole-3-carboxamido)but-2-enoate (compound 122)** Carboxylic acid starting material purchased from Enamine. 71.2 mg, 59.4% yield. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.51 (d, J = 8.9 Hz, 1H), 7.24 (d, J = 2.2 Hz, 1H), 7.09 (ddd, J = 15.7, 5.8, 4.7 Hz, 1H), 6.89 (ddd, J = 9.2, 2.3, 1.1 Hz, 1H), 6.06 (dd, J = 16.0, 1.9 Hz, 1H), 6.02 (d, J = 7.0 Hz, 1H), 4.32 (td, J = 5.4, 4.7, 1.6 Hz, 2H), 3.89 (d, J = 1.1 Hz, 3H), 3.76 (d, J = 1.2 Hz, 3H), 3.15 (dt, J = 7.0, 3.1 Hz, 1H), 2.79 (d, J = 1.2 Hz, 3H), 1.25 (dt, J = 6.9, 1.6 Hz, 2H), 1.15 − 0.96 (m, 2H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.51, 155.35, 145.09, 144.42, 132.36, 125.68, 121.23, 111.81, 110.62, 56.03, 51.64, 40.18, 24.95, 12.96, 7.54. [M+Na]: 364.920 Da.
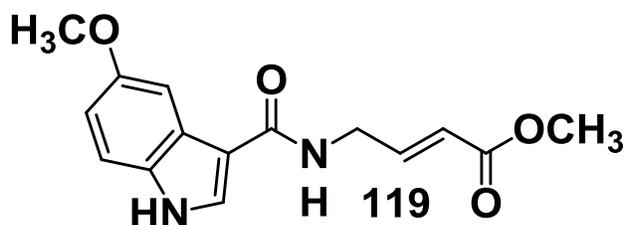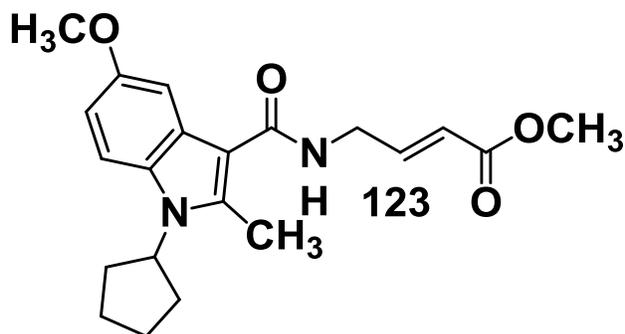
Methyl(E)-4-(1-cyclopentyl-5-methoxy-2-methyl-1H-indole-3-carboxamido)but-2-enoate

(compound 123) Carboxylic acid starting material purchased from Enamine. 78.8 mg, 60.8%.

^{1}H NMR (500 MHz, CDCl$_3$) δ 7.33 (d, J = 8.9 Hz, 1H), 7.24 (d, J = 2.4 Hz, 1H), 7.08 (dt, J = 15.7, 4.9 Hz, 1H), 6.83 (dd, J = 8.9, 2.4 Hz, 1H), 6.05 (dt, J = 15.6, 1.7 Hz, 1H), 6.01 − 5.90 (m, 1H), 4.82 (t, J = 9.0 Hz, 1H), 4.31 (td, J = 5.6, 1.8 Hz, 2H), 3.87 (s, 3H), 3.74 (s, 3H), 2.72 (s, 3H), 2.24 (td, J = 8.6, 5.3 Hz, 2H), 2.06 (dt, J = 10.7, 7.2 Hz, 5H), 1.81 (q, J = 6.5 Hz, 2H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.53, 154.88, 145.17, 142.29, 129.10, 126.85, 121.19, 112.42, 110.16, 107.19, 102.08, 56.00, 55.96, 51.63, 40.21, 29.99, 25.38, 12.28. [M+Na]: 392.971 Da.
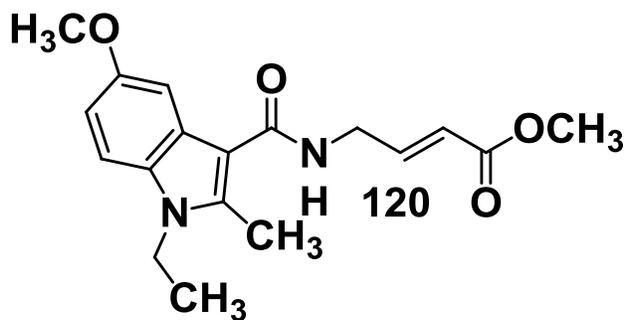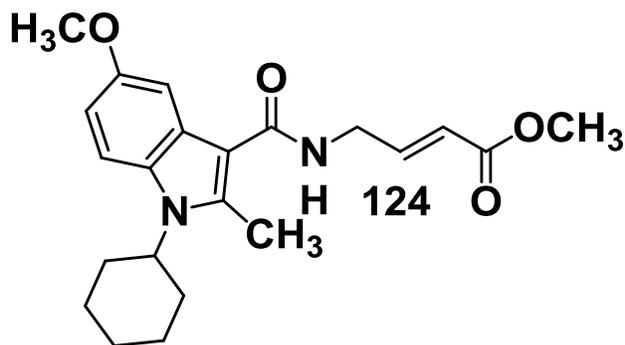

Methyl(E)-4-(1-cyclohexyl-5-methoxy-2-methyl-1H-indole-3-carboxamido)but-2-enoate

(compound 124) Carboxylic acid starting material purchased from Enamine. 63.5 mg, 47.2% yield ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.50 (d, J = 9.1 Hz, 1H), 7.24 (d, J = 2.5 Hz, 1H), 7.14 − 7.03 (m, 1H), 6.85 (dd, J = 9.0, 2.5 Hz, 1H), 6.07 (d, J = 15.7 Hz, 1H), 6.02 (d, J = 7.0 Hz, 1H), 4.33 (ddd, J = 5.9, 4.9, 2.0 Hz, 2H), 4.24 (tt, J = 12.3, 4.1 Hz, 1H), 3.89 (s, 3H), 3.76 (s, 3H), 2.74 (s, 3H), 2.43 − 2.19 (m, 2H), 2.00 (dt, J = 14.0, 3.3 Hz, 2H), 1.87 (ddd, J = 24.5, 12.7, 4.0 Hz, 4H), 1.57 − 1.43 (m, 2H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.53, 154.73, 145.16, 141.62, 121.19, 110.19, 101.74, 55.95, 51.63, 40.22, 31.25, 26.37, 25.47. [M+Na]: 406.993 Da.

CHAPTER 3: CHARACTERIZATION OF THE MECHANISM OF ACTION OF THE NEDD4-1 INHIBITORS

Portions of this chapter appear in this publication: 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*, 12442-12445.

3.1 INTRODUCTION

Current biochemical studies of HECT E3s suggest a mode of chain elongation which may occur by either a processive or a distributive mechanism.[52,53,67-69] A processive enzyme remains continuously bound to the substrate as it grows the polyUb chain, while a distributive enzyme dissociates from the substrate after each addition of a Ub to the chain (Figure 3.1). In either model, the last ubiquitin (Ub) of the growing polyubiquitin chain binds the N-lobe of the catalytic HECT domain proximal to the C-lobe, which positions this polyUb chain for the addition of another Ub molecule (Figure 3.2A). However, whether HECT E3 ligases are processive or distributive enzymes, and whether this process might be targeted for inhibition,[70] had not been completely investigated.
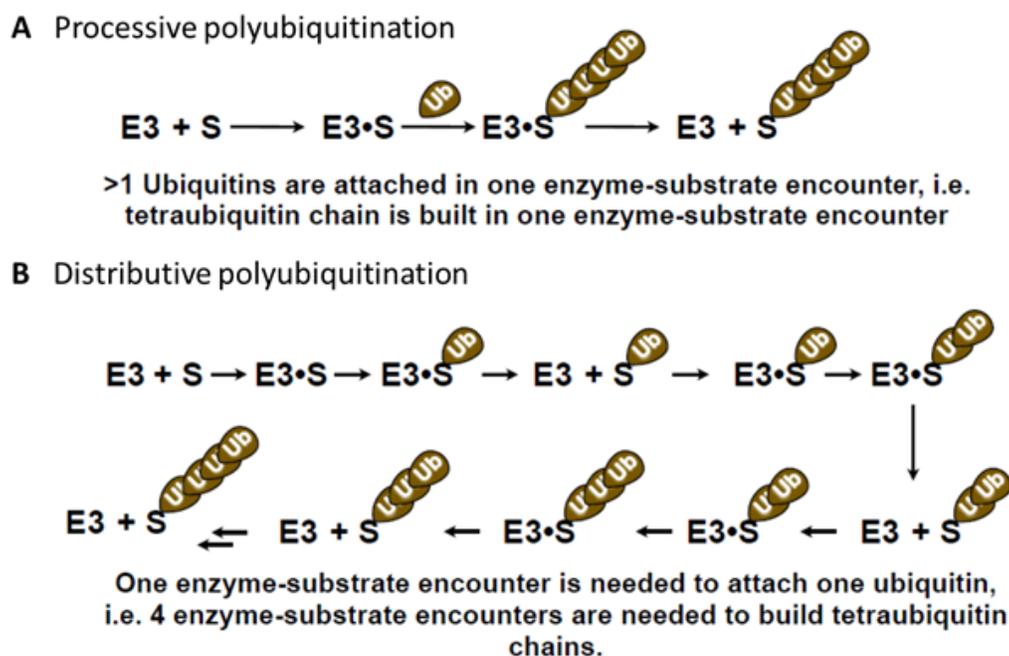
Figure 3.1 Processive and distributive mechanisms of polyubiquitination.
E3 ligases can build polyubiquitin chains with either a processive (A) or distributive (B) enzymatic mechanism. E3: E3 ubiquitin ligase; S: substrate, Ub: ubiquitin.

Cys^{627} is located in the N-lobe of Nedd4-1 near Phe^{707} of Nedd4-1 and Ile^{44} of Ub, two

residues that form critical hydrophobic contacts (Figure 3.2B). Mutation of Phe^{707} to Ala in

Nedd4-1 disrupts Nedd4-1:Ub binding and affects the kinetics of polyubiquitin chain growth.[52]

The equivalent F618A mutation in Rsp5, the *S. cerevisiae* homolog of Nedd4-1, also disrupts its

binding to Ub and results in temperature-sensitive growth defects, suggesting an essential

function of this site *in vivo*.[68] Therefore small molecules that target Cys^{627} should inhibit the

ability of Nedd4-1 to elongate polyubiquitin chains. This chapter will examine this hypothesis

and further explore the consequences of this mode of inhibition.
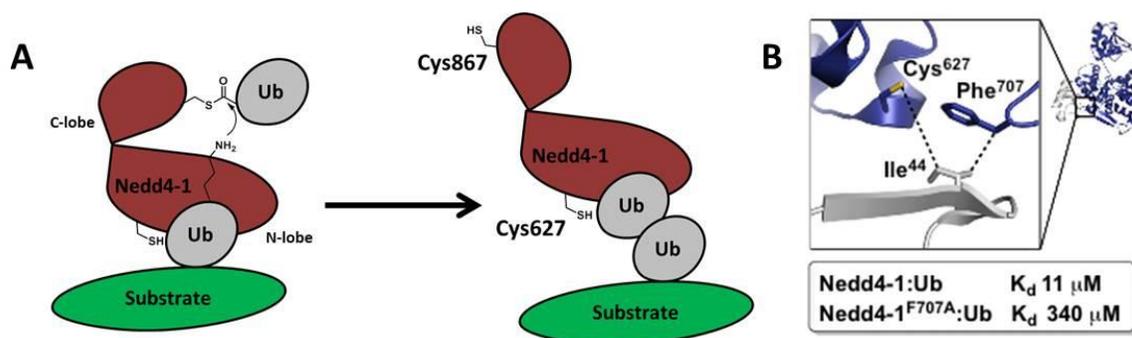
Figure 3.2 The Nedd4-1 HECT mechanism for building polyubiquitin chains.
A) The HECT N-lobe non-covalently binds the end of the growing polyubiquitin chain. Cys^{627}, which is modified by compound **123**, is part of this binding site B) The hotspot interface of Nedd4-1 (blue) and ubiquitin (gray) in the Nedd4-1:Ub non-covalent complex (PDB ID 2XBB) responsible for polyubiquitin chain elongation, with the key side chains depicted as sticks.

3.2 RESULTS AND DISCUSSION

3.2.1 Fragments and optimized inhibitors disrupt Ub binding to Nedd4-1

We first wanted to test our hypothesis that compounds **43** and **123** could disrupt non-

covalent Ub binding to Nedd4-1. Interestingly, it was shown that Tyr^{605} of Nedd4-1 is also

important for non-covalent Ub binding and polyubiquitin chain elongation by Nedd4-1.[52] Since

43 forms an edge-to-face binding interaction with Tyr^{605}, ligands at this position should block the

interaction between Leu[73] of Ub and Tyr[605] of Nedd4-1 observed in the Nedd4-1 HECT:Ub complex structure.

Fluorescence polarization (FP) experiments using fluorescein-Ub were used to confirm that **123** irreversibly inhibits the Nedd4-1:Ub interaction and to rigorously quantify the potency of this inhibition. Compound **123** disrupted Nedd4-1:Ub binding with second-order inhibition kinetics (K_I 29.3 μM, k_{inact} 5.8×10^{-5} s^{-1}; k_{inact}/K_I = 1.98 M^{-1} s^{-1}) and was 22.2-fold more potent than compound **43**. (Figure 3.3), which roughly correlates with the qualitative increased potency of labeling observed by ESI-MS. Notably, the FP assay requires a high concentration of Nedd4-1 (8 μM); so with a K_I of 29.3 μM we achieve half-maximal covalent inhibition at only a 3.7-fold excess of inhibitor relative to Nedd4-1. Furthermore, in contrast to fragment **43**, compound **123** was able to label Nedd4-1 even in the presence of 60 μM Ub, indicative of its increased potency (Figure 3.4).

A)

B)

	k_{inact}	K_I	k_{inact}/K_I
43	$2.2 \pm .232 \times 10^{-5}$ s^{-1}	245.8 ± 54.8 μM	$.089 \pm .029$ M^{-1}s^{-1}
123	$5.8 \pm .358 \times 10^{-5}$ s^{-1}	29.3 ± 5.91 μM	$1.98 \pm .52$ M^{-1}s^{-1}
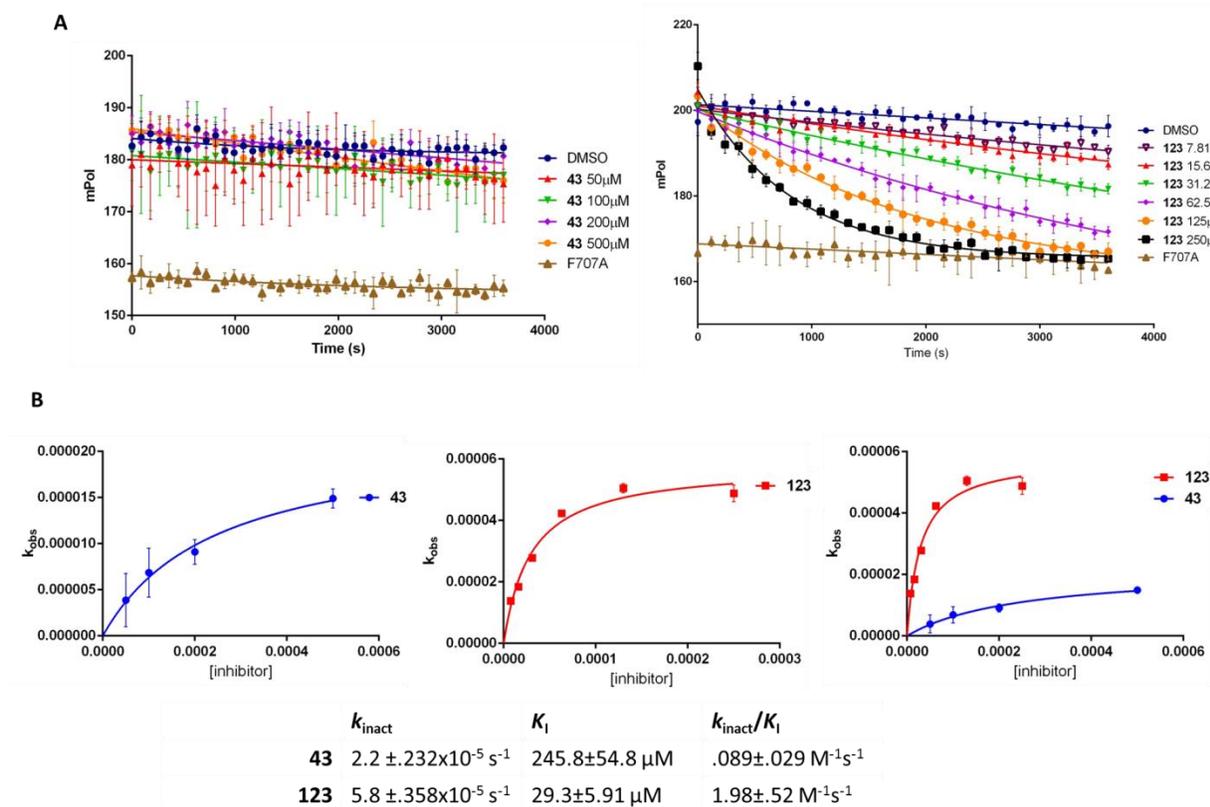
Figure 3.3 FP studies show compounds 43 and 123 inhibit Nedd4-1:Ub binding
A) Potency of inhibitors **43** and **123** at disrupting Nedd4-1:Ub interactions as assessed by fluorescence polarization. Nedd4-1 HECT and ubiquitin-fluorescein were treated with the indicated concentration of inhibitor in 1% DMSO. Changes in fluorescence polarization were monitored over 1 h. All reactions were performed in triplicate and plotted as mean ± s.e.m. B) k_{obs} vs [inhibitor] plots showing a two-step mechanism for the covalent modification of Nedd4-1, in which the initial non-covalent Nedd4-1:inhibitor complex is formed, followed by the covalent bond formation step. k_{obs} values were determined from the slopes of the log plots of (A). Compound **123** is 22.2-fold more potent than compound **43**.
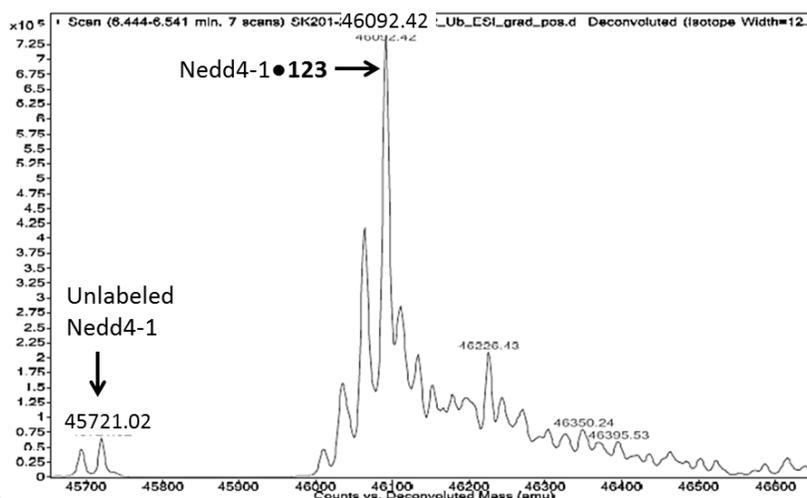
Figure 3.4 Compound 123 completely labels Nedd4-1 in the presence of 60 μM ubiquitin. This concentration of Ub is significantly above the Nedd4-1:Ub K_d value of 11 μM. Compound **123** at 100 μM in 1% DMSO was incubated with Nedd4-1 HECT domain (10 μM) and ubiquitin (60 μM) for 4 h, followed by gel filtration and whole protein ESI-MS.

3.2.2 Compound 123 disrupts the formation of polyubiquitin chains by Nedd4-1

To investigate if compound **123** affects the ability of Nedd4-1 to elongate polyUb chains we used Wbp2-C-K222 as a substrate.[69,71] Wbp2-C-K222 has only one acceptor lysine residue (Lys[222]) and a single cysteine residue (Cys[73]), which we modified with 5-iodo-acetamidofluorescein for quantification purposes (abbreviated Flu-Wbp2; Figure 3.5). Remarkably, we observed that the catalytic HECT domain of Nedd4-1 covalently modified with compound **123** was not able to build long polyUb chains on Flu-Wbp2, as compared to Nedd4-1 treated with DMSO or the negative control electrophile **47** (Figure 3.6). Importantly, this inhibition does not occur because **123** inhibits the formation of the Nedd4-1~Ub thioester (Figure 3.7). Nedd4-1 Phe[896]Ala, which can form a Nedd4-1~Ub thioester but cannot discharge the Ub,[72] showed no difference in the rate of Nedd4-1~Ub thioester formation when untreated or completely labeled by **123**. However, since the Nedd4-1:**123** complex could still build short

polyUb chains on Flu-Wbp2, we asked if these chains are built with a processive or a distributive

mechanism.
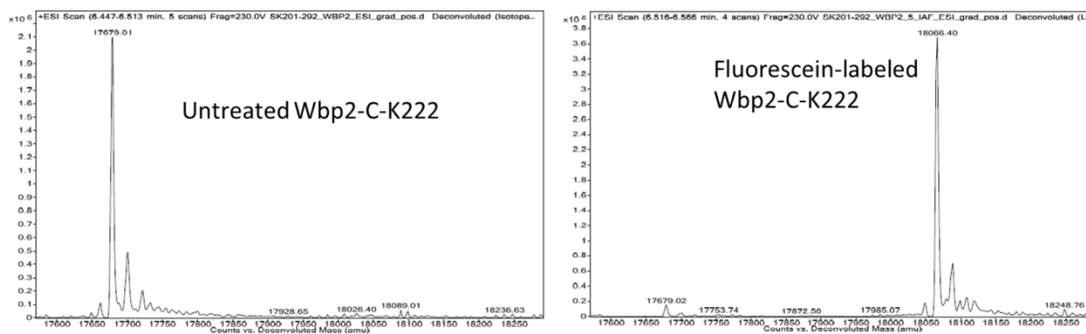
Figure 3.5 Covalent labeling of Wbp2-C-K222 with 5–iodoacetamidofluorescein.
Wbp2-C-K222 was treated with 3 mM 5-iodoacetamidofluorescein for 90 min at 4°C.
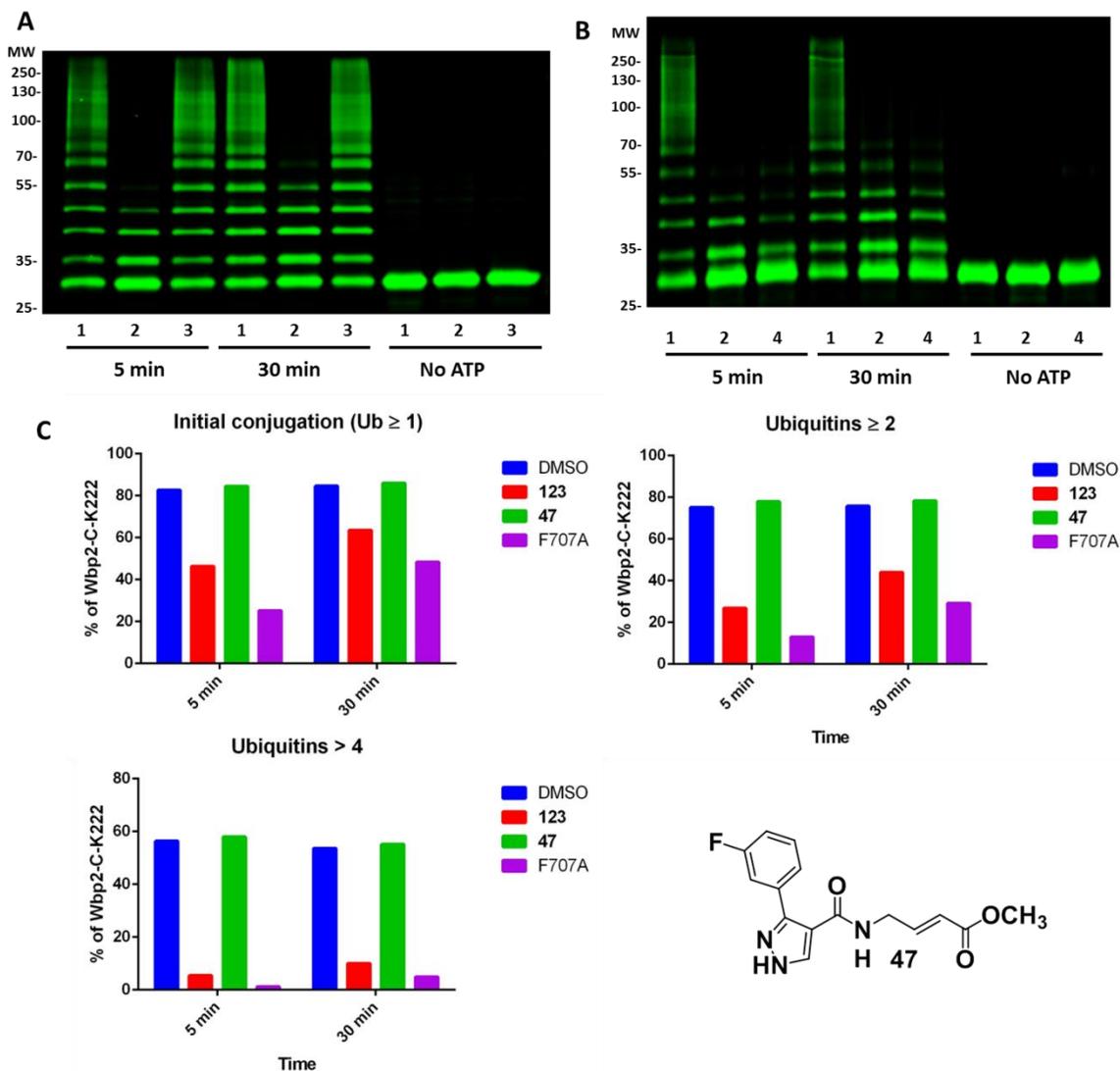
Figure 3.6 Nedd4-1 HECT domain labeled by compound 123 has an impaired ability to form polyubiquitin chains

A) Nedd4-1 HECT domain treated with 1% DMSO (lane 1), compound **123** (lane 2), or non-hit electrophile **47** (lane 3) was incubated with E1 and E2 enzymes, ubiquitin, Flu-Wbp2 and ATP. Reaction mixtures were quenched at the indicated times and the amount of ubiquitinated Flu-Wbp2 was determined using in-gel fluorescence. B) Same as in (A) but with the F707A mutant (lane 4) C) Quantification of fluorescent bands from (A) and (B) showing that both initial monoubiquitination and polyubiquitination are disrupted by inhibitor **123**, but polyubiquitination is more greatly affected. This effect is comparable to the Nedd4-1 F707A mutation. Fluorescent Wbp2-C-K222 bands with the indicated number of ubiquitins are plotted as percent of total Wbp2-C-K222 bands (non-ubiquitinated + polyubiquitinated). Compound **47** is an electrophilic compound from the original screening library that did not react with Nedd4-1 HECT domain.
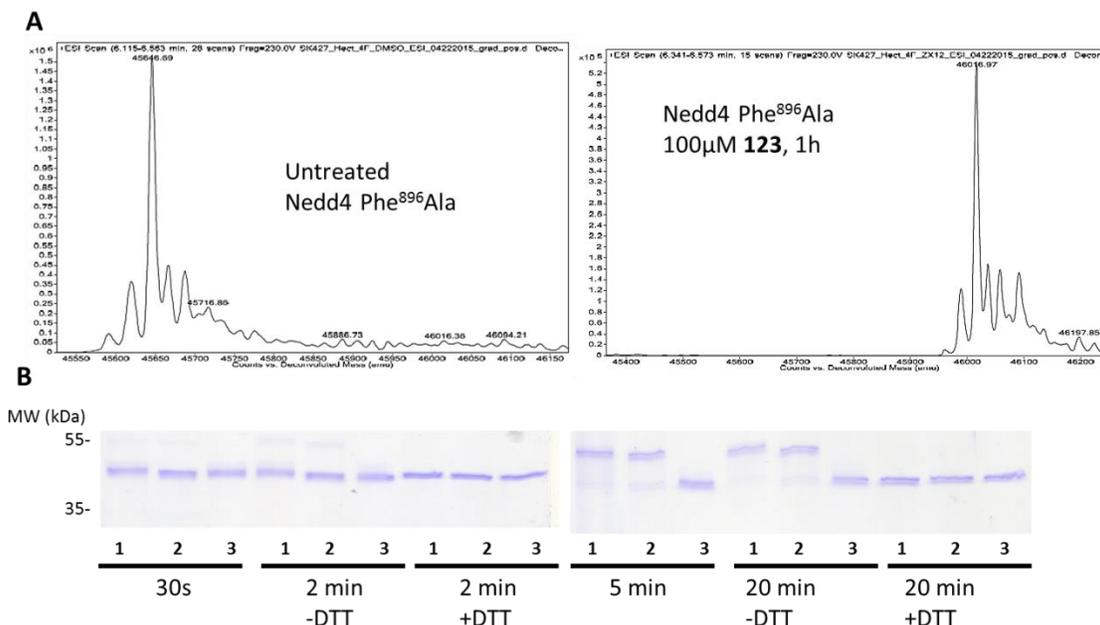
Figure 3.7 Nedd4-1~Ub thioester formation is unaffected by inhibitor 123.
A) Nedd4-1 HECT Phe[896]Ala, which can form a Nedd4-1~Ub thioester but cannot discharge Ub, was fully labeled by compound **123** (100μM, 1h). B) Nedd4-1 HECT Phe[896]Ala was treated with 1% DMSO (lane 1) or compound **123** (lane 2) without subsequent Zeba column gel filtration, then incubated with E1 and E2 enzymes, ubiquitin, and ATP. Reaction mixtures were quenched with Laemmli buffer (-DTT) at the indicated times and the amount of Nedd4-1~Ub thioester was visualized by Coomassie staining. Lane 3 is a no ATP control reaction. This figure shows that compound **123** does not inhibit E1, E2, or the E2~Ub to E3~Ub transthiolation.

3.2.3 Compound 123 switches Nedd4-1 from a processive to a distributive enzyme

We hypothesized that since compound **123** disrupts Ub binding to the non-covalent Ub binding site of Nedd4-1, the ubiquitinated substrate would be more likely to dissociate from the Nedd4-1:inhibitor complex in between rounds of addition of Ub to the growing chain. Therefore, inhibitor-bound Nedd4-1 should assemble polyUb chains via a distributive mechanism. To distinguish between processive and distributive mechanisms, we used an assay wherein full length Nedd4-1 with or without inhibitor pretreatment is allowed to ubiquitinate Flu-Wbp2 for 1 min, followed by addition of a 200-fold excess of non-fluorescent Wbp2.[73-75] If Nedd4-1 is processive, it should remain bound to Flu-Wbp2-Ub$_x$ and continue to elongate the polyUb chain on Flu-Wbp2 even in the presence of the large excess of non-fluorescent Wbp2. If Nedd4-1 is

distributive, it should dissociate from Flu-Wbp2-Ub_x between rounds of ubiquitination. In this case, Flu-Wbp2-Ub_x will be outcompeted by non-fluorescent Wbp2 and polyUb chain growth on Flu-Wbp2 will be inhibited.

For these experiments, we used full-length Nedd4-1 with the activating E554A mutation, which disrupts the autoinhibitory conformation of wild type full length Nedd4-1.[76] We found that E554A Nedd4-1 was processive and efficiently converted Flu-Wbp2 into $\geq Ub_4$ modified Flu-Wbp2 even after addition of a 200-fold excess of non-fluorescent Wbp2 (Figure 3.8). However, in the case of the Nedd4-1:**123** complex (Figure 3.9), we found that ubiquitination of Flu-Wbp2 was significantly inhibited upon addition of a 200-fold excess of non-fluorescent Wbp2 (Figure 3.10). Furthermore, consumption of monoubiquitinated Flu-Wbp2 and the formation of Ub_2/Ub_3 and $\geq Ub_4$ modified Flu-Wbp2 were also inhibited. This observation indicates that inhibitor-bound Nedd4-1 dissociates from Flu-Wbp2-Ub_x before adding Ub_{x+1}, and is therefore distributive. Similar results were observed for the Nedd4-1 E554A F707A mutant (Figure 3.11). These experiments prove for the first time that Nedd4-1 is processive, and when the non-covalent interaction between the N-lobe and Ub is disrupted by compound **123** or the F707A mutation the enzyme becomes distributive. Previously, it was assumed, but not rigorously proven, that HECT E3s are processive and not distributive enzymes.
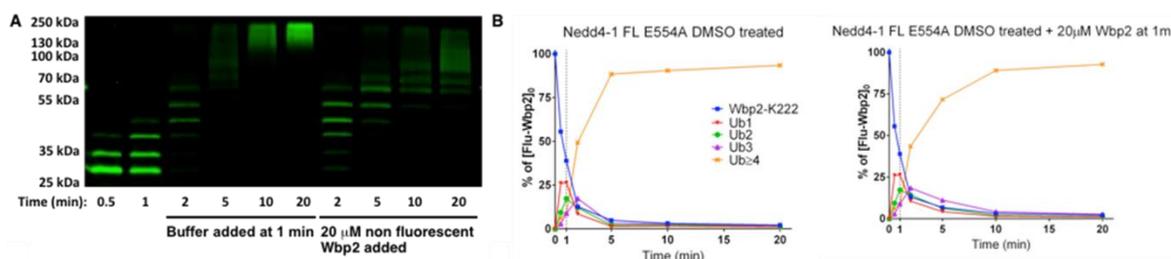
Figure 3.8 Unlabeled Nedd4-1 is a processive enzyme.
A) Full length Nedd4-1 with the activating E554A mutation (150 nM) was incubated with fluorescent Flu-Wbp2 substrate (100 nM) in the presence of ATP, Ub, E1 and E2 enzymes. After

1 min, a 200-fold excess of non-fluorescent Wbp2 substrate or empty buffer was added to the reaction mixture, and further ubiquitination of Flu-Wbp2 was monitored. B) The amount of ubiquitinated Flu-Wbp2 was plotted as a function of time.
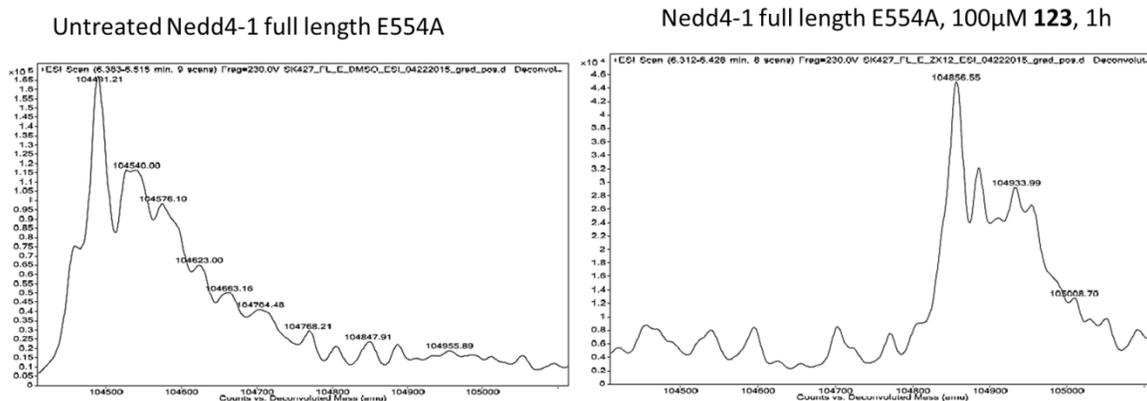
Figure 3.9 Compound 123 reacts with Nedd4-1 full-length E554A.
Compound 123 at 100 μM in 1% DMSO was incubated with Nedd4-1 full-length E554A (10 μM) for 1 h, followed by gel filtration and whole protein ESI-MS.
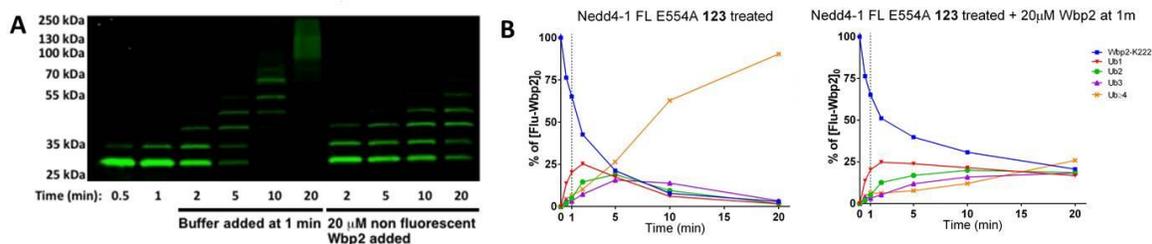
Figure 3.10 Inhibitor 123 switches Nedd4-1 from a processive to a distributive enzyme.
A) Full length Nedd4-1 E554A covalently modified with compound **123** (150 nM) was incubated with fluorescent Flu-Wbp2 substrate (100 nM) in the presence of ATP, Ub, E1 and E2 enzymes. After 1 min, a 200-fold excess of non-fluorescent Wbp2 substrate or empty buffer was added to the reaction mixture, and further ubiquitination of Flu-Wbp2 was monitored. B) The amount of ubiquitinated Flu-Wbp2 in A) was plotted as a function of time.
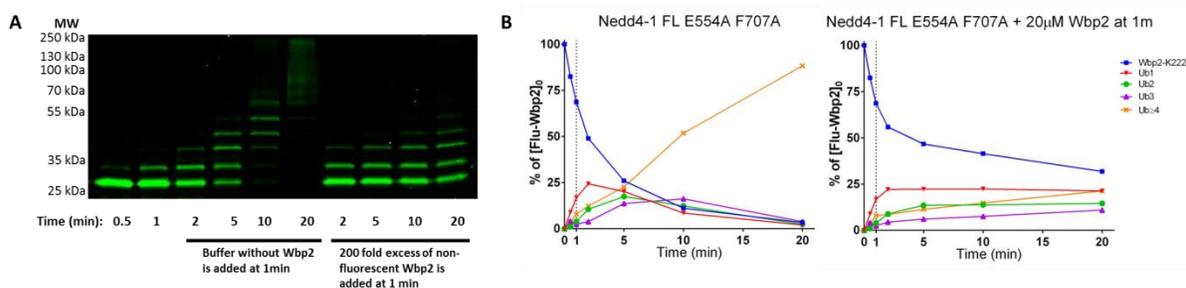
Figure 3.11 Nedd4-1 full-length E554A F707A is also distributive.

Full-length Nedd4-1 E554A F707A (150 nM) was incubated with fluorescent Flu-Wbp2 substrate (100 nM) in the presence of ATP, Ub, E1 and E2 enzymes. After 1 min, a 200-fold excess of non-fluorescent Wbp2 substrate or empty buffer was added to the reaction mixture, and further ubiquitination of Flu-Wbp2 was monitored. B) The amount of ubiquitinated Wbp2 in (A) was plotted as a function of time.

3.2.4 Distributive, but not processive, Nedd4-1 is completely inhibited by a DUB

Since endogenous intracellular deubiquitinating enzymes (DUBs) reverse protein ubiquitination, we hypothesized that distributive Nedd4-1 would be more susceptible to antagonism by DUBs than processive Nedd4-1. To test this hypothesis, full length Nedd4-1 E554A with or without compound **123** bound and the Nedd4-1 E554A F707A mutant were allowed to ubiquitinate Flu-Wbp2 in the presence of the DUB USP8. *In vitro*, USP8 can disassemble K48- and K63-linked polyUb chains,[77] while Nedd4-1 predominantly makes K63-linked chains.[52] Remarkably, we observed that untreated Nedd4-1 robustly polyubiquitinated Flu-Wbp2 in the presence of USP8 after 30 min. However, neither compound **123**-treated Nedd4-1 nor the F707A mutant were able to consume Flu-Wbp2 or build $\geq Ub_4$ chains on Flu-Wbp2 in the presence of USP8, even though they consumed Flu-Wbp2 in the absence of USP8 (Figures 3.12, 3.13). Since distributive Nedd4-1 dissociates from the substrate in between rounds of Ub addition to the growing polyUb chain, this provides an opportunity for the DUB to hydrolyze the Ub chain before another Ub can be added to it by Nedd4-1. Processive catalysis appears to be a necessary condition for the formation of $\geq Ub_4$ chains on the substrate in the presence of DUB.
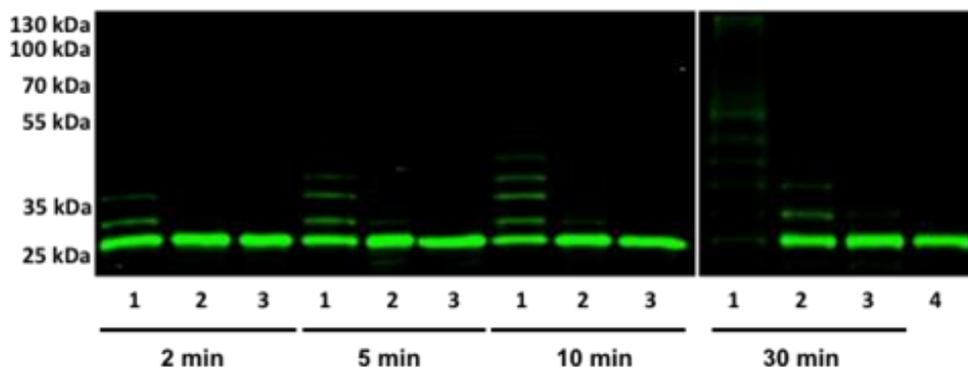
Figure 3.12 Distributive Nedd4-1, but not processive Nedd4-1, is inhibited in the presence of a deubiquitinase enzyme.
Ubiquitination of Flu-Wbp2 (100 nM) by full length Nedd4-1 (150 nM) at different time points in the presence of the deubiquitinase USP8 (200 nM) shows that distributive Nedd4-1 is inhibited by USP8, but processive Nedd4-1 is not. Lane 1: DMSO treated Nedd4-1 full length E554A, lane 2: compound **123** treated Nedd4-1 full length E554A, lane 3: Nedd4-1 full length E554A F707A, and lane 4: no ATP control.
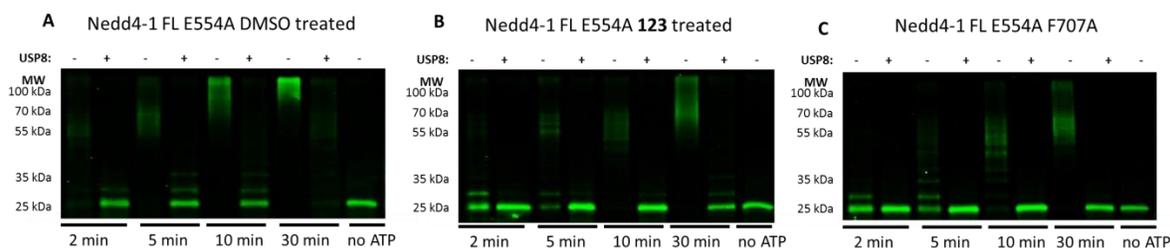
Figure 3.13 Side-by-side comparison of Nedd4-1 activity with or without USP8 added.
Ubiquitination of Flu-Wbp2 at different time points in the presence or absence of the deubiquitinating enzyme USP8 (200 nM catalytic domain) shows that distributive Nedd4-1 is inhibited by USP8 but processive Nedd4-1 is slowed down but not impaired. A) DMSO-treated Nedd4-1 full-length E554A; B) compound **123** treated Nedd4-1 full-length E554A; C) Nedd4-1 full-length E554A F707A.

3.3 CONCLUSION

In summary, we have rigorously proven that the HECT E3 Nedd4-1 is a processive enzyme, and that disrupting non-covalent binding of Ub to the N-lobe of Nedd4-1 switches Nedd4-1 to a distributive enzyme. Furthermore, we discovered and characterized a covalent Nedd4-1 inhibitor that targets this processive site and switches Nedd4-1 from a processive to a distributive mechanism. Additionally, we showed that introducing a DUB antagonist into the

assay augments the inhibitory effect of compound **123** on Nedd4-1 while the untreated enzyme is still able to build long polyUb chains. E3 ligase inhibitors with this mechanism of action were not previously known and this work outlines a conceptually new design strategy for E3 ligase inhibitors. Taken together, these studies, which were published in 2015,[78] provide fundamental insights into the HECT E3 enzymatic mechanism and represent an important case study in the emerging area of E3 ligase inhibitor discovery.[79] In the next chapter we will explore whether compound **123** can bind to and inhibit Nedd4-1 in cells.

3.4 MATERIALS AND METHODS

Recombinant expression of Nedd4-1 (full length) in *E. coli* Nedd4-1 in a PGEX6P1 vector plasmid (GST-Nedd4-1) was transformed into BL21 cells (Novagen). 1L TB media containing 100μg/ml ampicillin was inoculated with 50 mL overnight cell culture and incubated at 37°C until OD reached ~3. Then, IPTG (1.0 mM final concentration) was added to the cell culture media at 18°C, followed by 16 hour incubation at the same temperature. Cells were then harvested and lysed by sonication in phosphate buffered saline (PBS) with protease inhibitors (Complete Mini Protease Inhibitor Cocktail, Roche). The supernatant was incubated with glutathione agarose beads (Pierce Biotechnology) for 1 hour at 4°C. The beads were washed three times with PBS and incubated with PreScission Protease (GE Healthcare) for 4h at 23°C to elute Nedd4-1 (elution buffer: 50mM HEPES, 150 mM NaCl, 0.1 mM EDTA 1mM DTT). Mutant plasmids were prepared with Agilent QuickChange kit.

Nedd4-1 full length E554A sequence:

GPLGSMATCAVEVFGLLEDEENSRIVRVRVIAGIGLAKKDILGASDPYVRVTLYDPMNG

VLTSVQTKTIKKSLNPKWNEEILFRVHPQQHRLLFEVFDENRLTRDDFLGQVDVPLYPLP

TENPRLERPYTFKDFVLHPRSHKSRVKGYLRLKMTYLPKTSGSEDDNAEQAEELEPGW

VALDQPDAACHLQQQQEPSPLPPGWEERQDILGRTYYVNHESRRTQWKRPTPQDNLTD

AENGNIQLQAQRAFTTRRQISEETESVDNQESSENWEIIREDEATMYSSQAFPSPPPSSNL

DVPTHLAEELNARLTIFGNSAVSQPASSSNHSSRRGSLQAYTFEEQPTLPVLLPTSSGLPP

GWEEKQDERGRSYYVDHNSRTTTWTKPTVQATVETSQLTSSQSSAGPQSQASTSDSGQ

QVTQPSEIEQGFLPKGWEVRHAPNGRPFFIDHNTKTTTWEDPRLKIPAHLRGKTSLDTSN

DLGPLPPGWEERTHTDGRIFYINHNIKRTQWEDPRLENVAITGPAVPYSRDYKRKYEFFR

RKLKKQNDIPNKFEMKLRRATVLADSYRRIMGVKRADFLKARLWIEFDGEKGLDYGGV

AREWFFLISKEMFNPYYGLFEYSATDNYTLQINPNSGLCNEDHLSYFKFIGRVAGMAVY

HGKLLDGFFIRPFYKMMLHKPITLHDMESVDSEYYNSLRWILENDPTELDLRFIIDEELFG

QTHQHELKNGGSEIVVTNKNKKEYIYLVIQWRFVNRIQKQMAAFKEGFFELIPQDLIKIF

DENELELLMCGLGDVDVNDWREHTKYKNGYSANHQVIQWFWKAVLMMDSEKRIRLL

QFVTGTSRVPMNGFAELYGSNGPQSFTVEQWGTPEKLPRAHTCFNRLDLPPYESFEELW

DKLQMAIENTQGFDGVD

Fluorescence Polarization Assay Nedd4-1 HECT (8 μM) and ubiquitin-fluorescein (50 nM, Lifesensors) in 50mM HEPES pH 7.5 150mM NaCl 0.1 mM EDTA were treated with DMSO or varying concentrations of inhibitor in 1% DMSO in black 96 well plates. Changes in fluorescence polarization were monitored over 1h with a Biotek Synergy 4 plate reader. Slopes of ln(polarization) vs. time were plotted with GraphPad Prism and used to determine the pseudo-first order rate contstant k_{obs} for a given concentration of inhibitor. The values of k_{inact}/K_I for each inhibitor were then determined by fitting the k_{obs} vs. [inhibitor] plot to the equation $k_{obs} = k_{inact}*[inhibitor]/([inhibitor]+K_I)$. All reactions were performed in triplicate and plotted as mean ±s.e.m.

Preparation of fluorescein-Wbp2-C-K222 Wbp2-C-K222 in a PGEX6P1 vector plasmid (GST-Wbp2-C-K222) was transformed into BL21 cells (Novagen). 1L LB media containing 100μg/ml ampicillin was inoculated with 50 mL overnight cell culture and incubated at 37°C until OD reached ~0.6. Then, IPTG (0.1 mM final concentration) was added to the cell culture media at 18°C, followed by 16 hour incubation at the same temperature. Cells were then harvested and lysed by sonication in phosphate buffered saline (PBS) with protease inhibitors (Complete Mini Protease Inhibitor Cocktail, Roche). The supernatant was incubated with glutathione agarose beads (Pierce Biotechnology) for 1 hour at 4°C. The beads were washed three times with PBS and incubated with PreScission Protease (GE Healthcare) for 16h at 4°C to elute Wbp2-C-K222 (50mM HEPES, 150 mM NaCl, 0.1 mM EDTA 1mM DTT). Wbp2-C-K222 was then treated with 1mM TCEP for 15 min, then 5-iodoacetamidofluorescein (Santa Cruz Biotechnology) in DMSO was added (final concentration: 3mM, 5% DMSO). The reaction was rocked at 4°C for 90 min in the dark, then passed through a Zeba gel filtration column (Thermo, 7K MWCO). The tagged protein was further purified by size exclusion with an S75 column (GE Healthcare). Elution buffer: 20mM Tris-HCl pH 8.0 200mM NaCl 1mM EDTA 5% glycerol 1mM DTT.

Wbp2-C-K222 sequence:

GPLGSSRRASVGSPEFTMLTFTAGGAIEFGQRMLQVASQASRGEVPSGAYGYSYMPSGA
YVYPPPVANGMYPCPPGYPYPPPPPEFYPGPPMMDGAMGYVQPPPPPYPGPMEPPVSGP
DVPSTPAAEAKAAEAAASAYYNPGNPHNVYMPTSQPPPPPYYPPEDRRTQ

In vitro **ubiquitination assays – Nedd4-1 HECT domain** Reaction mixtures were composed of 80nM Ube1 E1 enzyme (Boston Biochem), 1.5 μM UbcH5a E2 enzyme (Boston Biochem), 1.5 μM Nedd4-1 HECT domain, 1 μM Flu-Wbp2 substrate, 6 μM ubiquitin (Sigma-Aldrich), and

ATP in 25mM HEPES pH 7.6 100mM NaCl 4mM $MgCl_2$. 30 µL reactions were quenched with 6X Laemmli buffer and 3 µL beta-mercaptoethanol and analyzed by SDS-PAGE. Fluorescent gels were imaged with a Typhoon 9400 (GE Healthcare) and fluorescent bands were quantified with ImageQuant TL. For inhibitor treated assays, Nedd4-1 HECT domain (10µM) was pretreated with inhibitor (100µM in 1% DMSO) for 1h, then used immediately in the enzymatic assay.

In vitro **ubiquitination assays – Nedd4-1 full length E554A** Reaction mixtures were composed of 60nM Ube1 E1 enzyme (Boston Biochem), 150 nM UbcH5a E2 enzyme (Boston Biochem), 150 nM Nedd4-1 full length E554A, 100 nM Flu-Wbp2 substrate, 80 µM ubiquitin (Sigma-Aldrich), and ATP in 25mM HEPES pH 7.6 100mM NaCl 4mM $MgCl_2$. 30 µL reactions were quenched with 6X Laemmli buffer and 3 µL beta-mercaptoethanol and analyzed by SDS-PAGE. Fluorescent gels were imaged with a Typhoon 9400 (GE Healthcare, at Northwestern's Keck Biophysics Facility) and fluorescent bands were quantified with ImageQuant TL. For inhibitor treated assays, Nedd4-1 full length E554A (10µM) was pretreated with inhibitor (100µM in 1% DMSO) for 1h, then diluted to 1µM and used immediately in the enzymatic assay.

CHAPTER 4: CELL BASED STUDIES OF NEDD4-1 INHIBITOR SPECIFICITY AND EFFICACY

4.1 INTRODUCTION

Having established that inhibitor **123** binds to Nedd4-1 with reasonable selectivity *in vitro*, our next goal was to establish its selectivity amongst the entire proteome in cells. Due to the cysteine reactive electrophile on **123**, there is the possibility of non-specific labeling of reactive cysteines in the proteome.[9,80] Fortunately, a main advantage of the covalent handle is that it allows for straightforward and high quality selectivity profiling via competition studies and click chemistry with an alkyne tagged analog of the inhibitor.[81,82] The alkyne-tagged analog can be reacted with an azido-fluorophore or azido-biotin, and the proteome-wide reactivity of the probe can be visualized with in-gel fluorescence or analyzed by biotin capture and MS proteomics, respectively.[83] Pretreatment with the electrophilic inhibitor before adding the probe will more specifically determine the targets and off-targets of the inhibitor.

We also desired to see what biological effects compound **123** might have in cells. Although compound **123** had only modest potency *in vitro*, and proved to be difficult to optimize, it could still be useful as a tool compound to determine the substrates and functions of Nedd4-1 if it has good selectivity. Such a tool compound would have advantages compared to current methods of discovering E3 substrates. One such method is genetic perturbation by RNA interference or genetic knockout,[84,85] which has a slow response time compared to small molecules[86] and is not suitable to live animal studies due to the essentiality of Nedd4-1 in embryonic development. Another method is co-immunoprecipitation of the E3 and its bound substrates,[87] which is challenging because E3-substrate interactions are weak and transient so many potential substrates will not be picked up. An additional method is finding substrates from an *in vitro* proteome array that is treated with the E3 of interest,[88] which provides information that is replicable in cells but may also miss substrates due to the lack of a live cellular milieu.

Understandably, these suboptimal methods have produced often contradictory findings about the substrates of Nedd4-1, and thus it remains difficult to know which substrates to focus on when testing for the effects of compound **123** in cells. It was initially reported that Nedd4-1 can mono- or polyubiquitinate the tumor suppressor PTEN;[89] respectively resulting in nuclear import of PTEN,[90] which is tumor-suppressive, or proteasomal degradation of PTEN,[91] which is oncogenic. It has also been shown that ubiquitination of PTEN itself by Nedd4-1 is sufficient to impair its tumor-suppressive phosphatase activity.[92] However, in later studies, reduction in Nedd4-1 expression has had no effect on PTEN stability or function,[93] and other E3 ligases were shown to ubiquitinate PTEN instead.[94] Like the PTEN controversy, one study showed that Nedd4-1 negatively regulates insulin-like growth factor (IGF1) signaling,[95] while others have concluded that Nedd4-1 positively regulates IGF1 signaling.[44,46] Similarly, one paper showed that Nedd4-1 has no effect on epidermal growth factor (EGF) signaling,[46] while others demonstrated that Nedd4-1 positively regulates epidermal growth factor receptor (EGFR) expression.[96,97] Other papers have shown that Nedd4-1 negatively regulates fibroblast growth factor (FGF) signaling by promoting endocytosis and degradation of the activated FGF1 receptor.[88,98,99] Our hope was that compound **123** could resolve some of these controversies by providing an orthogonal approach to genetic knockdown to studying the effects of impaired Nedd4-1 activity. However, it was unclear if disrupting Nedd4-1 processivity in cells is sufficient to completely inhibit Nedd4-1 function as it was *in vitro* when USP8 was present.

4.2 RESULTS AND DISCUSSION

4.2.1 Alkyne probes of Nedd4-1 inhibitors for click chemistry

An electrophilic probe analog **137** of inhibitor **123** was made by replacing the methoxy group of the acrylate with a propargyl group, which contains an alkyne for the click reaction.

Probe **137** labels Nedd4-1 *in vitro* with equal potency to inhibitor **123**. Treatment of TC71

Ewing's sarcoma cells, which overexpress Nedd4-1,[46] with **137,** followed by lysis, click

chemistry with rhodamine-azide and in-gel fluorescence showed that **137** reacts with reasonable

selectivity with a protein of a molecular weight corresponding to Nedd4-1 (Figure 4.2).

Pretreatment with inhibitor **123**, but not the inactive analog **117**, competed the Nedd4-1

molecular weight band as well as one other band of MW >250 kDa. All other bands were shown

to be off targets unique to the probe. Encouragingly, Nedd4-1 and Nedd4-2 were also pulled

down by probe **137** after a click reaction with biotin azide (Figure 4.3). The pulldown of Nedd4-

1 and Nedd4-2 was again competed by **123** but not **117**. However, a more rigorous

characterization of all of the targets of probe **137** by trypsin digestion of the biotin pulldown and

identification of the protein targets by tandem MS of the tryptic peptides must still be done. Even

though probe **137** has many off targets, which can be explained by its 4.2-fold greater reactivity

with cysteine than inhibitor **123** by NMR (Figure 4.4), the specific targets of inhibitor **123** can be

identified by comparative quantitative SILAC proteomics of samples pretreated or not pretreated

with **123** grown in heavy or light isotope media, respectively.[81]
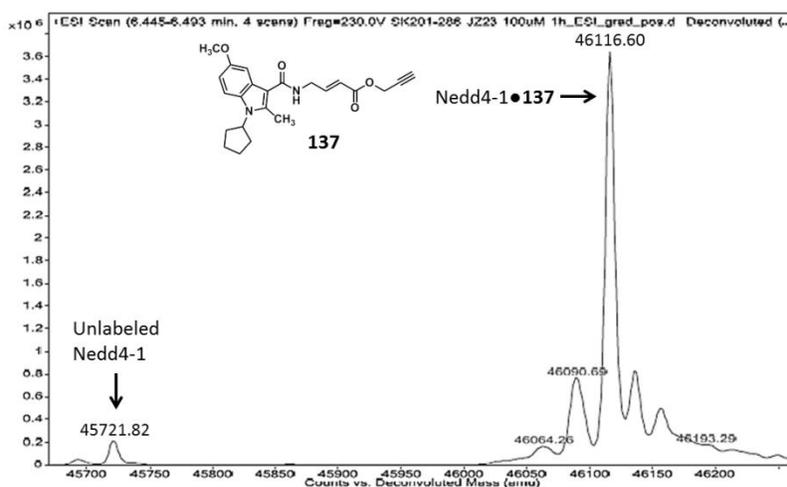
Figure 4.1 Probe 137 completely labels Nedd4-1 HECT *in vitro*.

Probe **137** at 100 µM in 1% DMSO was incubated with Nedd4-1 HECT domain (10 µM) for 1 h, followed by gel filtration and whole protein ESI-MS.
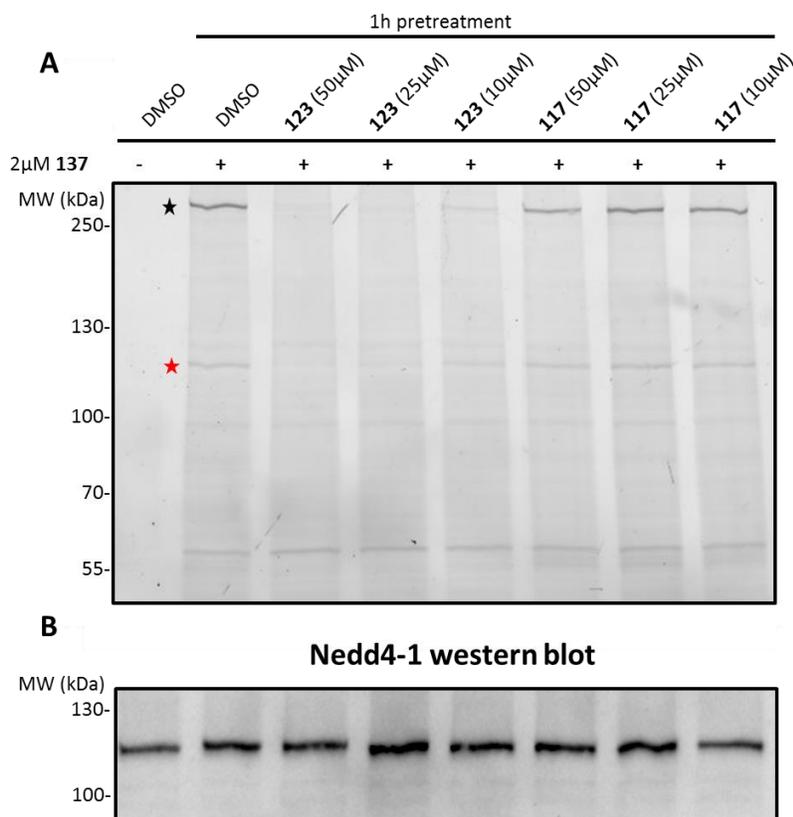
Figure 4.2 An alkyne-tagged probe 137 demonstrates that compound 123 is cell membrane-permeable and has good apparent selectivity for Nedd4-1 in TC-71 cells.
A) Pretreatment with compound **123** for 1h, but not the inactive indole **117**, abolished labeling of a band at ~120 kDa (red star), which matches Nedd4-1 by Western blot (B). There was also one notable specific off target of **123** at >250 kDa (black star). TC-71 cells were treated with the indicated compounds, lysed, click reacted with rhodamine azide and targets were visualized with in-gel fluorescence.
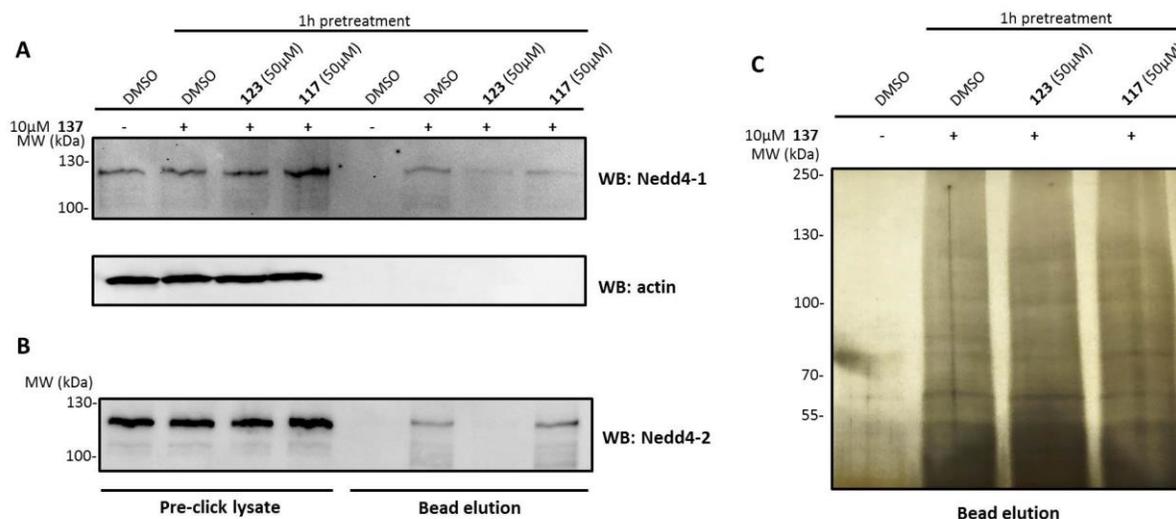
Figure 4.3 Biotin pulldown confirms probe 137 and inhibitor 123 react with Nedd4-1 and Nedd4-2 in TC71 cells
A) TC71 cells were treated with 10 μM **137**, followed by lysis, click reaction with biotin azide, streptavidin bead capture, and elution. Nedd4-1 is pulled-down by **137**, and this pulldown is somewhat inhibited by compound **123** but less so by the inactive analog **117**. B) Same as (A), but for Nedd4-2. C) Silver stain of the streptavidin bead elutions showing all the proteins pulled down by probe **137**.
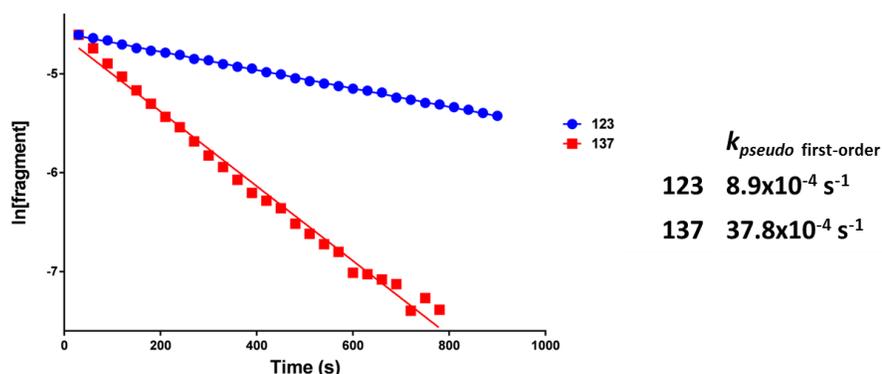
Figure 4.4 Probe 137 is more reactive than inhibitor 123
NMR rate studies with N-acetylcysteine methyl ester show that probe **137** is 4.2-fold more reactive than inhibitor **123**. This likely explains the amount of unique off-targets seen for the probe.

4.2.2 TC71 cell growth inhibition by compound 123 and a non-hydrolyzable analog

A previous study reported that knocking down Nedd4-1 inhibits the growth of TC71 cells.[46] In this paper, Nedd4-1 was shown to be essential for IGF1 signaling, which TC71 cells

are dependent upon for survival. Therefore we used the CellTiter-Glo assay to test if compound

123 can inhibit TC71 cell growth (Figure 4.5). **123** had an IC_{50} of 29.9 µM, which was better

than that of the inactive electrophilic compound **117** (105.2 µM) by a factor of 3.5. However, it

was far less potent than the IGF1 receptor kinase inhibitor OSI-906, which is currently in clinical

trials for Ewing's sarcoma.[100] Notably, compound **123** did demonstrate synergy with OSI-906

while compound **117** did not (Figure 4.6), indicating that compound **123** and OSI-906 might be

inhibiting the same pathway or multiple connected pathways.[101] The combination index (CI) of

123 and OSI-906 is 0.5, while that of **117** and OSI-906 is 0.904. A CI <1 is a hallmark of

synergistic drug interaction, while a CI ≈ 1 indicates an additive interaction.[102]
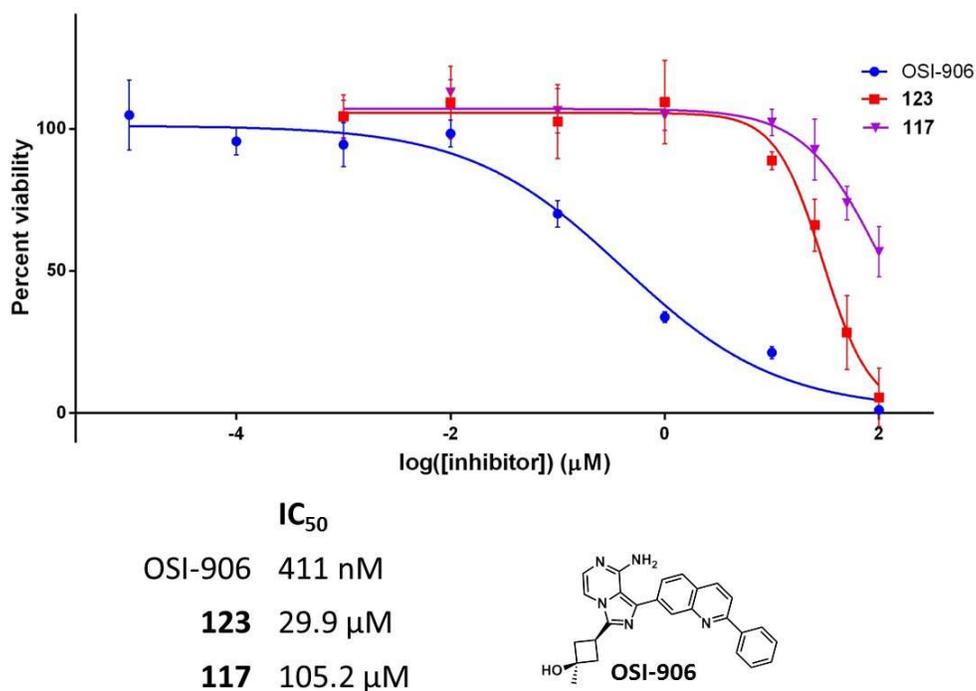
Figure 4.5 TC71 cell growth inhibition by compound 123
Compound **123** inhibited TC71 cell growth with an IC_{50} of 29.9 µM, which was 3.5-fold more potent than the inactive control **117** but 72.7-fold less potent than the IGF1R inhibitor OSI-906. TC71 cells were treated with the indicated concentration of compound for 2 days before growth was quantified with the CellTiter-Glo assay.
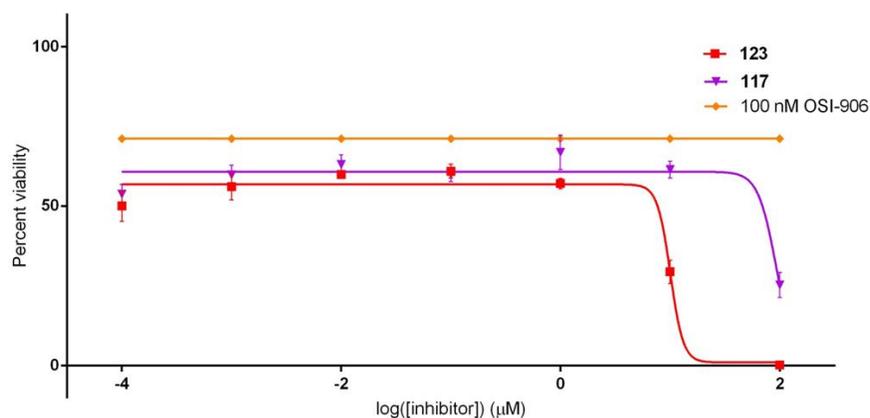
$$Combination\ Index(CI) = \frac{C(A)}{IC_{50}(A)} + \frac{C(B)}{IC_{50}(B)}$$

	C-IC$_{50}$	CI
123	7.29 µM	0.500
116	73.4 µM	0.904

Figure 4.6 Compound 123 demonstrates synergy with OSI-906 in inhibiting cell growth
Compound **123** and sub-IC$_{50}$ concentration of OSI-906 combine to reduce the cell growth IC$_{50}$ of **123**, while **117** and OSI-906 do not. A combination index <1 indicates synergy, while a CI of 1 indicates additivity.

One possible reason for the low cell growth inhibition potency of compound **123** is that the acrylate methyl ester could be hydrolyzed by esterases in the cell.[103] The hydrolyzed methyl ester compound **138** does not label Nedd4-1 (Figure 4.7). Amides such as **131** and **132** also do not label Nedd4-1, but compound **139**, in which the methyl acrylate is replaced with a vinyl ketone, robustly labels Nedd4-1 (Figure 4.8A). Compound **139** completely labels Nedd4-1 after 10 min at 100 µM, but there was also a small amount of di-labeling of Nedd4-1. It labels 50% of Nedd4-1 at 25 µM, and compound **140**, the vinyl ketone analog of the inactive compound **117**, is much less potent but it does slightly react with Nedd4-1 (Figure 4.8B). Compound **139** also disrupts non-covalent binding of Ub to Nedd4-1 in the FP assay (Figure 4.9), with 3.4-fold increased potency compared to compound **123** (k_{inact}/K_I 6.83 M^{-1}s^{-1} vs. 1.98 M^{-1}s^{-1}), as was also

seen in the MS assay. However, this added potency is entirely due to an increase in k_{inact} ($2.1\times10^{-4}s^{-1}$ vs. $5.8\times10^{-5}s^{-1}$), so this increased reactivity may cause selectivity problems, as demonstrated by the di-labeling of Nedd4-1 by compound **139** in the MS assay. The increased potency of compound **139** was also reflected in the TC71 cell growth inhibition assay (Figure 4.10). Compound **139** completely inhibited TC71 cell growth at 10 μM, although the inhibition plot was too steep to accurately determine an IC_{50}. However, compound **140** also inhibited TC71 cell growth at 10 μM, suggesting that the increased potency may be due to the greater cytotoxicity of the more reactive vinyl ketone electrophile rather than specific inhibition of Nedd4-1. Therefore, studies with compound **139** should be treated with caution due to the higher likelihood of off targets. Nevertheless, it provides a useful complement to compound **123** since it is not possible for the electrophile to be hydrolyzed and rendered inactive.
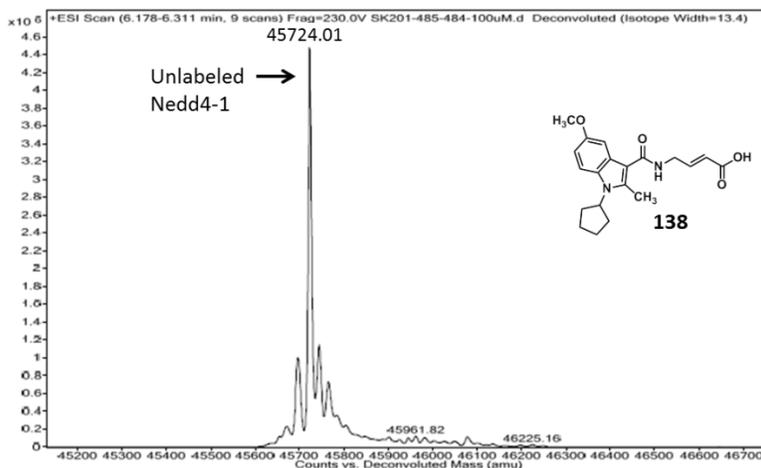
Figure 4.7 Compound 138, the hydrolyzed methyl ester analog of 123, does not label Nedd4-1.
Compound **138** at 100 μM in 1% DMSO was incubated with Nedd4-1 HECT domain (10 μM) for 1 h, followed by gel filtration and whole protein ESI-MS.
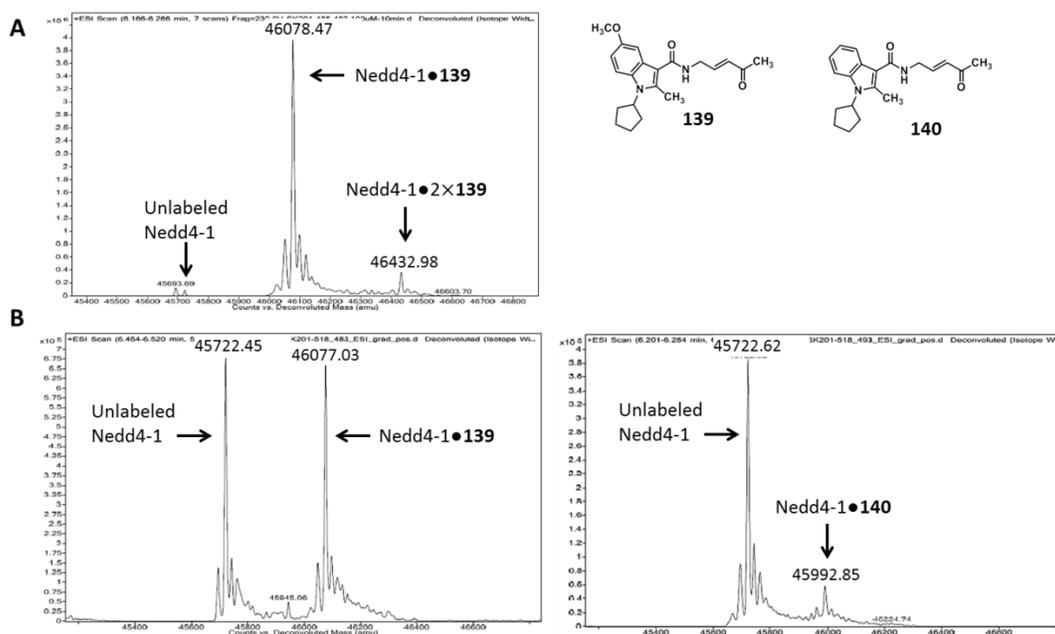
Figure 4.8 Compound 139, the non-hydrolyzable vinyl ketone analog of 123, robustly labels Nedd4-1.

A) Compound **139** at 100 μM in 1% DMSO was incubated with Nedd4-1 HECT domain (10 μM) for 10 min, followed by gel filtration and whole protein ESI-MS. B) Compound **139** and compound **140**, the vinyl ketone analog of inactive compound **117**, at 25 μM in 1% DMSO were incubated with Nedd4-1 HECT domain (10 μM) for 1 h, followed by gel filtration and whole protein ESI-MS.

	k_{inact}	K_I	k_{inact}/K_I
123	$5.8 \pm .358 \times 10^{-5}$ s^{-1}	29.3 ± 5.91 μM	$1.98 \pm .52$ M^{-1}s^{-1}
139	$2.1 \pm .185 \times 10^{-4}$ s^{-1}	31.2 ± 7.99 μM	$6.83 \pm .34$ M^{-1}s^{-1}

Figure 4.9 Compound 139 disrupts Nedd4-1:Ub binding more potently than compound 123

FP shows that compound **139** is 3.4-fold more potent at disrupting Nedd4-1:Ub non-covalent binding than compound **123**. However, the increased k_{inact}/K_I comes entirely from the k_{inact} term,

presumably due to the greater reactivity of the vinyl ketone electrophile compared to the methyl acrylate electrophile.
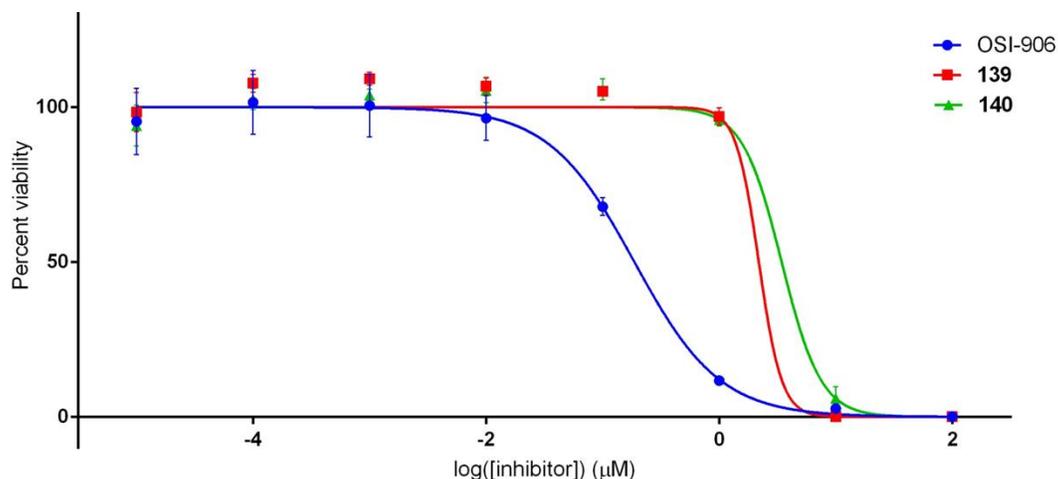
Figure 4.10 Compounds 139 and 140 both potently inhibit growth of TC71 cells
Compound **139** is more potent than compound **123**, completely inhibiting TC71 cell growth at 10 μM. However, compound **140**, which does not robustly label Nedd4-1 *in vitro*, also inhibits TC71 cell growth at 10 μM. This suggests that the greater potency of compound **139** compared to compound **123** is probably due to increased cytotoxicity of the vinyl ketone electrophile.

4.2.3 Nedd4-1 processivity inhibitors do not inhibit IGF1 signaling

Since Nedd4-1 was shown to be essential for IGF1 signaling in TC71 cells,[46] we next sought to determine if our inhibitors disrupted this signaling pathway. Nedd4-1 is thought to ubiquitinate and inactivate the phosphatase PTEN, which itself dephosphorylates a tyrosine on the signaling protein IRS1. When this IRS1 tyrosine is not phosphorylated, the IGF1 signal is no longer carried down the cascade and the downstream kinase Akt is not phosphorylated and activated. Therefore, if our compounds do fully inhibit Nedd4-1 in TC71 cells, we should see a reduction in Akt phosphorylation after IGF1 stimulation.

Unfortunately, a significant reduction in Akt phosphorylation was not observed upon treatment with either compound **123** or compound **139** (Figure 4.11). Akt phosphorylation was diminished by the IGF1R inhibitor OSI-906, but Akt phosphorylation levels were the same upon treatment with compound **123** and **139** as after treatment with their negative controls **117** and

140. There are several possible explanations for this finding, given that compound **123** was shown to react with Nedd4-1 in TC71 cells at 50 µM. One is that compound **123** does indeed disrupt Nedd4-1 processivity in cells, but this mechanism is insufficient to fully inhibit Nedd4-1. A complementary E3 ligase could extend the polyubiquitin chain if Nedd4-1 bound to **123** is able to monoubiquitinate its substrate. It may also be the case that only monoubiquitination of PTEN is required to impair its ability to dephosphorylate IRS1. It is also possible that Nedd4-1 is not in fact necessary for IGF1 signaling, given that other studies have found it to be a negative regulator of IGF1 signaling.[95]
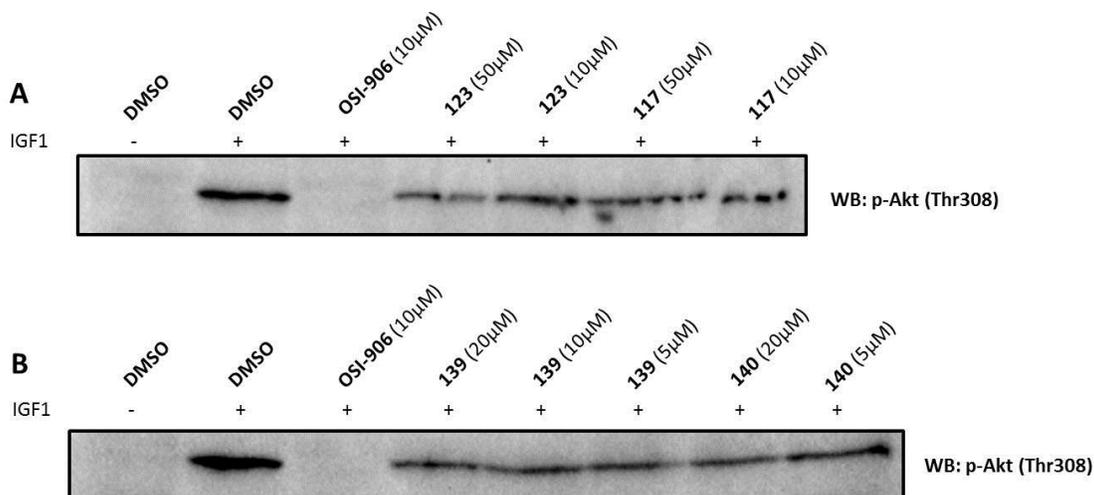
Figure 4.11 Compounds 123 and 139 have no effect on IGF1 signaling in TC71 cells.
TC71 cells were treated with methyl acrylate inhibitors (A) or vinyl ketone inhibitors (B) for 6h before stimulation with IGF1 (100ng/mL) for 15min. A significant reduction in downstream Akt phosphorylation was not seen compared to OSI-906, and p-Akt levels were the same between **123** and **139** and their negative controls **117** and **140**.

4.2.4 Bifunctional PROTAC molecules from Nedd4-1 inhibitors and thalidomide

Although we were not able to show that inhibitor **123** disrupts Nedd4-1 function in cells, we were encouraged by the finding that it does react with Nedd4-1 in cells with apparent selectivity. Therefore inhibitor **123** could be incorporated into a bifunctional molecule known as a PROTAC. PROTACs are chimeric molecules with one half that binds to a protein substrate of

interest and another half that binds to an E3 ligase. The PROTAC can then bring the substrate and the E3 into proximity, resulting in ubiquitination and degradation of the substrate.[104,105] A PROTAC derived from **123** and another E3 binding element could induce degradation of Nedd4-1 rather than merely inhibiting its processivity.

With that in mind, we coupled **123** to thalidomide, which binds the E3 ligase cereblon,[106] to make PROTAC **141**. The linker between thalidomide and **123** was placed at the 2-position of the indole core of **123**, which was previously shown to tolerate longer chain substitutions. Unfortunately PROTAC **141** did not label Nedd4-1 *in vitro* with the same potency as inhibitor **123** (Figure 4.12). The PROTAC only labeled about 33% of Nedd4-1 at 200 μM after 1h. Nevertheless, it is possible that only weak binding to Nedd4-1 in cells would be sufficient to induce its degradation by cereblon, so we tested **141** for this effect. However, treatment of TC71 cells with a range of concentrations of PROTAC **141** (10 μM-1nM) was not shown to induce degradation of Nedd4-1 or Nedd4-2 at a range of concentrations tested (Figure 4.13). Another possible effect of PROTAC **141** is ubiquitination and degradation of cereblon by Nedd4-1, but this was not observed either. It is possible that if a wider variety of linkers between the indole and thalidomide portions of the PROTAC was explored that a PROTAC that efficiently labels Nedd4-1 and achieves its desired effect could be discovered. Indeed, all PROTACs are very sensitive to linker length and typically many linkers must be synthesized before finding the optimal one.[107] However, we decided to stop pursuing PROTACs derived from **123** due to concerns about the micromolar potency of the parent indole compound.
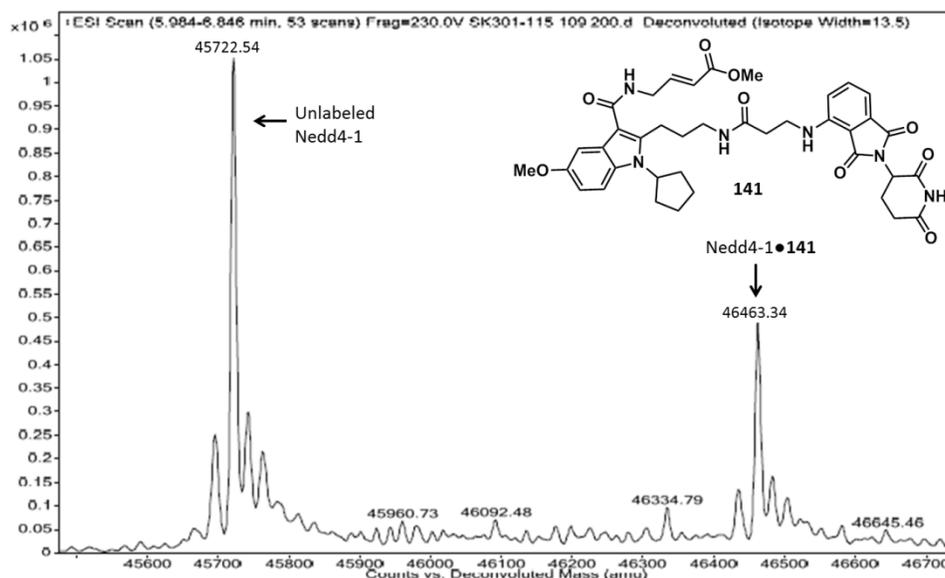
Figure 4.12 PROTAC 141 does not label Nedd4-1 as well as the indole 123 it is derived from.
Compound **141** at 200 μM in 1% DMSO was incubated with Nedd4-1 HECT domain (10 μM) for 1h, followed by gel filtration and whole protein ESI-MS.
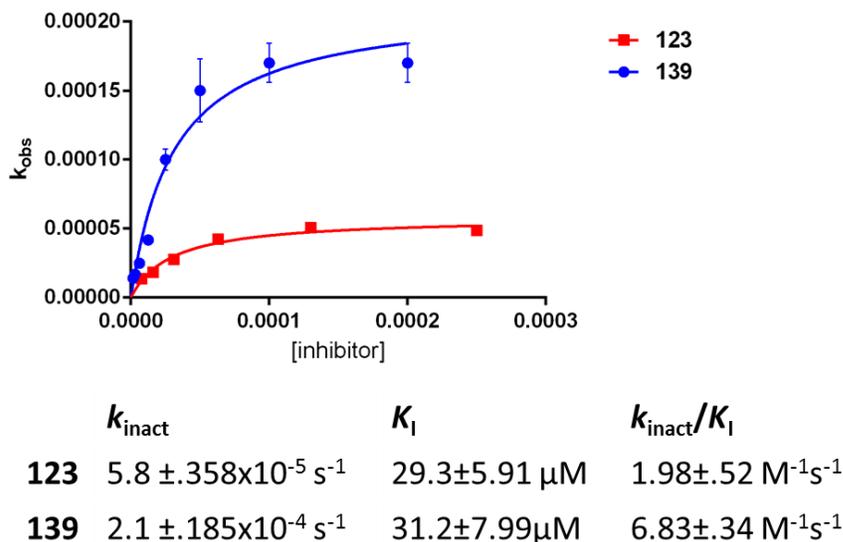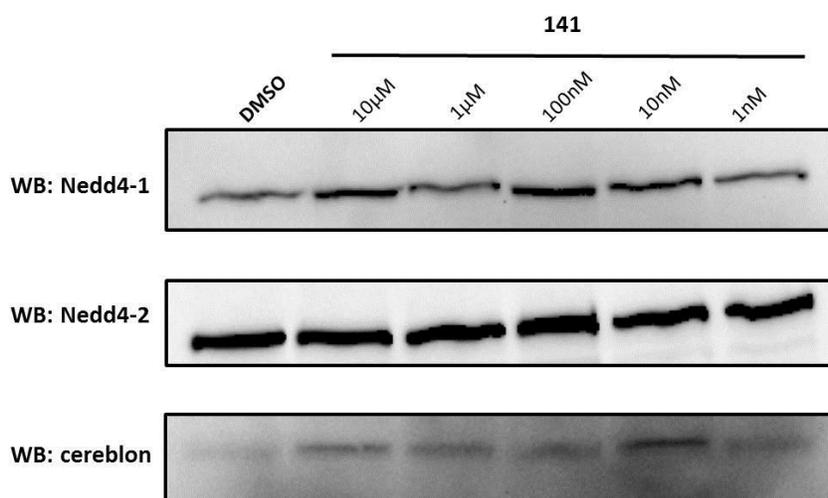

Figure 4.13 PROTAC 141 does not induce degradation of Nedd4-1, Nedd4-2, or cereblon
TC71 cells were treated with PROTAC **141** at the indicated concentrations for 16h, before lysis and Western blot to determine if any of the possible target proteins had reduced levels. Nedd4-1 and Nedd4-2 did not appear to be degraded by cereblon, nor did cereblon appear to be degraded by Nedd4-1.

4.3 CONCLUSION

Inhibitor **123** has both benefits and disadvantages due to its covalent mechanism of action. There are always concerns about non-specific labeling of off-target proteins due to the reactive electrophile, but the covalent interaction allows for rigorous characterization of all these interactions. Alkyne probe **137** was used to determine the selectivity of **123**, and it was found to react with Nedd4-1 and Nedd4-2 in TC71 cells with only one major apparent off-target. However, a quantitative proteomics study must still be performed with probe **137** before the specificity of inhibitor **123** can be fully described.

Despite the fact that compound **123** does bind Nedd4-1 in cells, it was not shown to inhibit Nedd4-1 function in the IGF1 signaling assay we performed. This may be due to the low binding affinity (~30 μM K_I) of **123** and its non-hydrolyzable vinyl ketone analog **139**. It may also be necessary to test compound **123** in many other biological assays since so little is still known about the true substrates of Nedd4-1 and so many contradictory results have been reported. Nevertheless, a compound that selectivity binds Nedd4-1 in cells is still an important breakthrough in the challenging field of HECT E3 inhibitor discovery. Such a molecule could be incorporated into a bifunctional molecule such as a PROTAC, which can induce degradation of the HECT E3 or an unnatural substrate of the HECT E3. Our initial attempt to make a PROTAC from compound **123** did not succeed, but it is possible that a PROTAC with a different linker or based on a more potent indole compound could achieve its desired effects.

4.4 MATERIALS AND METHODS

Antibodies Antibodies were ordered from the following sources: Nedd4-1 (Cell Signaling 3607), actin (Millipore MAB1501), Nedd4-2 (Cell Signaling 4013), p-Akt Thr308 (Cell Signaling 2965), cereblon (Abcam ab98992).

Reagents Cell culture and assay reagents were ordered from the following sources: ITS (Aldrich I3146), OSI-906 (Active Biochem A-1058), IGF1 (Aldrich GF138).

Cell-based selectivity studies with alkyne probe, click reaction, and in-gel fluorescence TC71 cells were cultured in Iscove's Modified Dulbecco's Medium (IMDM), 20% (v/v) FBS, 4mM glutamine, and 1xITS (5µg/mL insulin, 5µg/mL transferrin, 5ng/mL selenous acid). Cells were grown to confluence in 6-well plates (9.6 cm^2 per well). The growth medium was aspirated off and replaced with 2mL serum-free IMDM containing 0.1% DMSO or the indicated concentration of inhibitor in 0.1% DMSO. After 1h incubation time at 37 ºC, cells were treated with 0.1% DMSO or 2 µM of probe **137** in 0.1% DMSO for an additional hour. The medium was then aspirated off and cells were washed with Dulbecco's Phosphate Buffered Saline 3×2 mL. 300 µL of lysis buffer containing Tris HCl (25 mM, pH 7.6), NaCl (150 mM), 1% NP40, 1% sodium deoxycholate, 0.1 % SDS, and protease inhibitors (Sigma-Aldrich, P8340, 1:100 v/v) was added to each well, and cells were lysed by rocking at 4 ºC for 10 min. Cellular debris was cleared by centrifugation (21,000×g) for 45 min at 4 ºC. The total protein concentration of each cell lysate was normalized to 1.0 mg/mL using the BCA assay. 20 µL of each cell lysate was then treated for 30 min in the dark with the click chemistry reagents: $CuSO_4$ (final conc. 1 mM), TBTA (final conc. 100 µM), sodium ascorbate (final conc. 1 mM), TCEP (final conc. 1 mM), and Azide-Fluor-585 (final conc. 100 µM). Proteins were then resolved by SDS-PAGE 7.5% or 15% acrylamide gels, and subjected to in-gel scanning fluorescence imaging on the Typhoon 9400 at 610 nm excitation. The gel was then stained with Coomassie Brilliant Blue to visualize all proteins.

Biotin pulldown studies with alkyne probe TC71 cells were grown to confluence in T-75 flasks. The growth medium was aspirated off and replaced with 15mL serum-free IMDM

containing 0.1% DMSO or the indicated concentration of inhibitor in 0.1% DMSO. After 1h

incubation time at 37 °C, cells were treated with 0.1% DMSO or 10 μM of probe **137** in 0.1%

DMSO for an additional hour. The medium was then aspirated off and cells were washed with

Dulbecco's Phosphate Buffered Saline 3×10 mL. 2 mL of lysis buffer containing Tris HCl (25

mM, pH 7.6), NaCl (150 mM), 1% NP40, 1% sodium deoxycholate, 0.1 % SDS, and protease

inhibitors (Sigma-Aldrich, P8340, 1:100 v/v) was added to each well, and cells were lysed by

rocking at 4 °C for 10 min. Cellular debris was cleared by centrifugation (21,000×g) for 45 min

at 4 °C. The total protein concentration of each cell lysate was normalized to 1.0 mg/mL using

the BCA assay. The lysates were then click reacted with carboxamide-6-azidohexanyl biotin

(200 μM) in the presence of $CuSO_4$ (1 mM), TBTA (100 μM), sodium ascorbate (1 mM), TCEP

(1 mM), and 0.4% SDS for 1h. Proteins were then treated with acetone (8mL), incubated at -

20°C for 30 min, and centrifuged (4500×g, 30 min) to remove the excess amount of

carboxamide-6-azidohexanyl biotin. The proteins were resuspended with 5mL of HEPES

(100mM, pH 7.4, 0.5% SDS) and incubated with 400 μL of streptavidin agarose beads

(Invitrogen) for 1 hour at room temperature. The beads were then washed with 2×4 ml of HEPES

(100 mM, pH 7.4, 0.5% SDS) and 2×4 ml of deionized water. Proteins were eluted from the

streptavidin beads by incubation in 200 μL 0.5M Tris-HCl pH 6.8, 4% SDS, 6 mM biotin for

15min at room temperature, followed by boiling for 15min at 95°C. The eluates were separated

by SDS-PAGE (7.5% gel), transferred to 0.2 μm nitrocellulose membrane (70V, 60 min) and

blocked with TBST buffer containing 5% non-fat milk for 1 hour. Anti-Nedd4-1 or anti-Nedd4-2

antibodies (1:1000 in 5% non-fat milk) were incubated overnight at 4 °C. The membranes were

washed with TBST 3×5min, and incubated with goat anti-rabbit IgG HRP conjugate (Bio-Rad

1:3000 dilution in non-fat milk, 5% in TBST) for one hour, followed by wash with TBST

3×5min. The chemiluminescent reagent (Bio-Rad Clarity) was then added and the immunoblot was imaged on a Bio-Rad ChemiDoc XRS+ imager.

Cell growth IC$_{50}$ TC71 cells (90μL, ~10,000 per well) were seeded into black with clear bottom 96 well plates (Greiner). At least four duplicate wells were tested for every given concentration of a given compound. After 16h, compounds in 1.1% DMSO were added to the wells (final well volume: 110 μL, 0.2% DMSO). Cells were incubated with the compounds for 48h, and the CellTiter-Glo reagent (Promega) was then added. Luminescence was then quantified on a Biotek Synergy 4 to determine cell growth. The luminescence of each well was then expressed as a percentage of that of the DMSO-only control wells and was plotted in GraphPad Prism. Inhibition curves were fitted with the log[inhibitor] vs. response (variable slope, four parameters) function to determine IC$_{50}$.

IGF1 signaling studies 6 well plates were pretreated with 1mL 10μg/mL fibronectin (Aldrich) in DPBS at 37°C for 1h and then washed with 1 mL DPBS. TC71 cells were then grown to 90% confluence in 2 mL IMDM with 10% FBS. Media was then aspirated off and wells were washed once with 1mL DPBS before being serum starved for 16h in 2 mL IMDM. The cells were then treated with compounds for 6h, before stimulation with 100ng/mL IGF1 at 37°C for 15min. Media was aspirated off and wells were washed twice with 1 mL DPBS. Cells were then lysed and Western blots with anti-phospho Akt (Thr308) antibody were performed as described above.

4.5 SYNTHESIS

Synthesis of Probe 137: General Scheme
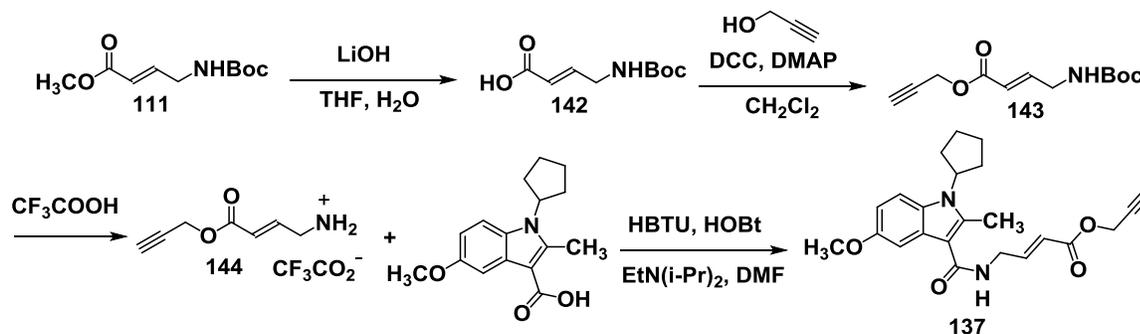
Synthesis of 142 110 (1.5g, 7 mmol) was dissolved in THF (37mL). Lithium hydroxide monohydrate (971 mg, 23 mmol) dissolved in water (22 mL) was then added. TLC at 3h showed full conversion to product. THF was evaporated under reduced pressure and the aqueous residue was adjusted to pH 3 with 1M HCl. The product was extracted with DCM (3x30mL). The combined organic layers were dried over $MgSO_4$, filtered, and evaporated to yield **142** (1.3g, 91%). ^{1}H NMR (400 MHz, DMSO-d$_6$) δ 12.23 (s, 1H), 7.13 (t, J = 5.8 Hz, 1H), 6.70 (dt, J = 15.7, 4.7 Hz, 1H), 5.72 (d, J = 15.8 Hz, 1H), 3.67 (t, J = 4.7 Hz, 2H), 1.35 (s, 9H).

Synthesis of 143 142 (200 mg, 1 mmol) was dissolved in CH_2Cl_2 (5 mL, 0.2M). DCC (199 mg, 1 mmol) and DMAP (21 mg, 0.17 mmol) were then added, followed by propargyl alcohol (179 µL, 3 mmol). The reaction was stirred for 1h at room temperature, at which point TLC showed conversion to product. The white precipitate was then filtered off and washed with $CHCl_3$ (2x5mL). The filtrate was then evaporated, and the residue was purified by flash column chromatography with an ethyl acetate/hexanes gradient 25% EtOAc → 50% EtOAc to yield **143** (120 mg, 50% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.22 (dt, J = 15.9, 4.7 Hz, 1H), 6.21 (dt, J = 15.7, 2.0 Hz, 1H), 5.01 (s, 1H), 4.98 (d, J = 1.7 Hz, 2H), 4.17 (d, J = 5.8 Hz, 2H), 2.72 (d, J = 2.4 Hz, 1H), 1.69 (s, 9H).

Synthesis of 144 143 (120 mg, 0.5 mmol) was dissolved in trifluoroacetic acid (766 µL, 10 mmol) and stirred at 23°C for 30 min. TLC at 30 min showed conversion to product. TFA was

evaporated and azeotroped with toluene (2×100mL). The residue was then dried on high vacuum for 2 hours to yield **144** as the TFA salt (109 mg, 86% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.36 (s, 3H) 7.00 (s, 1H), 6.22 (s, 1H), 5.01 (s, 1H), 4.78 (s, 2H), 3.88 (s, 2H), 2.65 (s, 1H).

Synthesis of 137 1-cyclopentyl-5-methoxy-2-methyl-1H-indole-3-carboxylic acid (117 mg, 0.43 mmol) was dissolved in dimethylformamide (0.2M, 2.14 mL), then **144**, TFA (109 mg, 0.43 mmol), HBTU (158 mg, 0.41 mmol), and HOBT (63.6 mg, 0.47 mmol) were added, followed by diisopropylethylamine (215 µL, 1.28 mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H$_2$O (5mL) and extracted with DCM (3×5mL). The combined organic layers were washed with 1M HCl (10mL), saturated aqueous NaHCO$_3$ (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over MgSO$_4$, filtered, and evaporated. Purification with flash column chromatography with CH$_3$OH/CH$_2$Cl$_2$ (CH$_3$OH gradient 0→5%) yielded compound **137** as a yellow powder (55.4 mg, 33.8% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.34 (d, J = 9.0 Hz, 1H), 7.24 (d, J = 2.5 Hz, 1H), 7.15 (dt, J = 15.7, 4.8 Hz, 1H), 6.84 (dd, J = 9.0, 2.4 Hz, 1H), 6.08 (dt, J = 15.8, 1.9 Hz, 1H), 6.02 (t, J = 6.0 Hz, 1H), 4.82 (t, J = 9.0 Hz, 1H), 4.74 (d, J = 2.5 Hz, 2H), 4.32 (ddd, J = 6.5, 4.9, 2.0 Hz, 2H), 3.88 (s, 3H), 2.73 (s, 3H), 2.48 (t, J = 2.5 Hz, 1H), 2.31 − 2.17 (m, 2H), 2.06 (dd, J = 9.2, 3.1 Hz, 4H), 1.81 (qd, J = 5.1, 4.5, 2.2 Hz, 2H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 166.38, 165.11, 154.90, 146.57, 142.37, 129.10, 126.83, 120.42, 112.45, 110.19, 107.12, 102.05, 74.88, 56.01, 55.99, 51.98, 40.23, 29.99, 25.39, 12.29. [M+H]: 395.360 Da.

Synthesis of 138
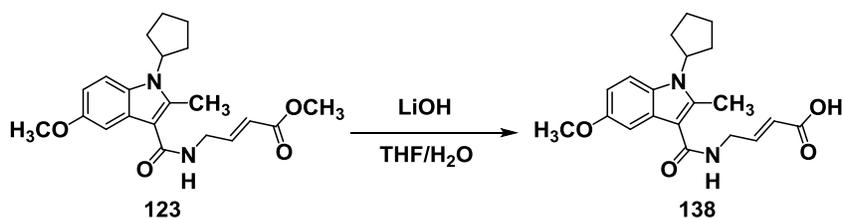
123 (40mg, 0.1 mmol) was dissolved in THF (0.6 mL). Lithium hydroxide monohydrate (14 mg, 0.6 mmol) dissolved in water (0.3 mL) was then added. TLC at 3h showed full conversion to product. THF was evaporated under reduced pressure and the aqueous residue was adjusted to pH 3 with 1M HCl. The product was extracted with DCM (3x10mL). The combined organic layers were dried over MgSO$_4$, filtered, and evaporated to yield **138** (24.7 mg, 69% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.31 (d, J = 8.9 Hz, 1H), 7.21 (d, J = 2.5 Hz, 1H), 7.14 (dt, J = 15.7, 4.8 Hz, 1H), 6.81 (dd, J = 9.0, 2.5 Hz, 1H), 6.02 (dt, J = 15.6, 1.9 Hz, 1H), 5.97 (t, J = 5.9 Hz, 1H), 4.80 (p, J = 8.9 Hz, 1H), 4.32 (ddd, J = 6.6, 4.9, 2.0 Hz, 2H), 3.85 (s, 3H), 2.70 (s, 3H), 2.20 (dt, J = 16.7, 8.2 Hz, 2H), 2.13 – 1.94 (m, 4H), 1.89 – 1.63 (m, 2H).

Synthesis of 139: General Scheme

Synthesis of 145 1-cyclopentyl-5-methoxy-2-methyl-1H-indole-3-carboxylic acid (100 mg, 0.37 mmol) was dissolved in dimethylformamide (0.2M, 1.83 mL), then allylamine (27.41 μL, 0.37 mmol), HBTU (134.6 mg, 0.354 mmol), and HOBT (54.35 mg, 0.40 mmol) were added, followed by diisopropylethylamine (122.5 μL, 0.70 mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H$_2$O (5mL) and extracted with DCM (3×5mL). The combined organic layers were washed with 1M HCl

(10mL), saturated aqueous NaHCO$_3$ (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over MgSO$_4$, filtered, and evaporated. Purification with flash column chromatography with CH$_3$OH/CH$_2$Cl$_2$ (CH$_3$OH gradient 0→5%) yielded compound **145** (55.5 mg, 48.6% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.31 (d, J = 8.9 Hz, 1H), 7.21 (d, J = 2.5 Hz, 1H), 7.05 (dt, J = 15.6, 4.9 Hz, 1H), 6.81 (dd, J = 9.0, 2.5 Hz, 1H), 6.02 (dt, J = 15.7, 1.9 Hz, 1H), 5.96 – 5.89 (m, 1H), 4.79 (p, J = 9.0 Hz, 1H), 4.29 (ddd, J = 5.9, 4.9, 2.0 Hz, 2H), 3.85 (s, 3H), 3.71 (s, 3H), 2.21 (ddd, J = 12.8, 6.9, 3.9 Hz, 2H), 2.13 – 1.95 (m, 4H), 1.90 – 1.74 (m, 2H).

Synthesis of 139 145 (30mg, 0.1mmol) and Hoveyda-Grubbs catalyst II (7mg, 0.01 mmol) were added to a round bottom flask, which was evacuated and filled with nitrogen. CH$_2$Cl$_2$ (0.05M, 2mL) and 3-butene-2-one (40.5 μL, 0.5mmol) were then added, and the flask was fitted with a reflux condenser and heated to 40°C for 16h. At 16h, TLC showed conversion to product. Filtered through celite and removed CH$_2$Cl$_2$ under reduced pressure. Purified by flash column chromatography with CH$_3$OH/CH$_2$Cl$_2$ (CH$_3$OH gradient 0→5%) yielded compound **139** (26.2 mg, 74% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.31 (d, J = 9.0 Hz, 1H), 7.24 (s, 1H), 7.21 (d, J = 2.5 Hz, 1H), 6.87 (dt, J = 16.1, 5.0 Hz, 1H), 6.81 (dd, J = 9.0, 2.5 Hz, 1H), 4.80 (p, J = 9.0 Hz, 1H), 4.30 (ddd, J = 5.9, 5.0, 1.9 Hz, 2H), 3.85 (s, 3H), 2.70 (s, 3H), 2.25 (s, 3H), 2.24 – 2.14 (m, 2H), 2.11 – 1.99 (m, 4H), 1.89 – 1.73 (m, 2H). [M+H]: 355.192 Da.
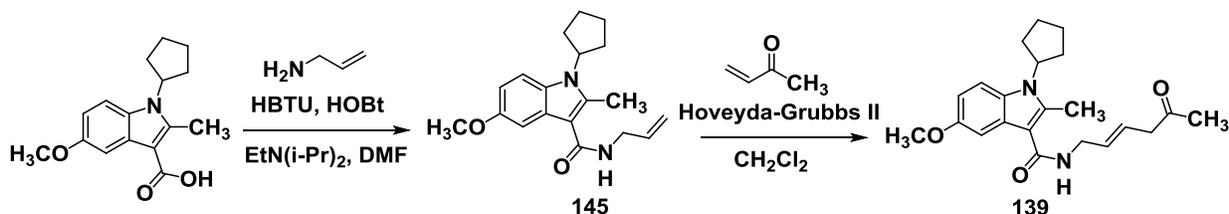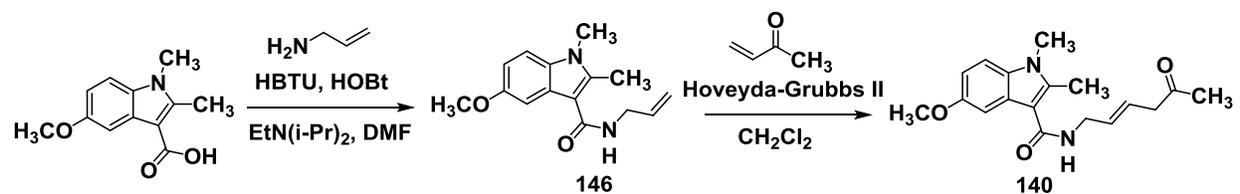
Synthesis of 140: General Scheme

Synthesis of 146 1-methyl-2-methyl-1H-indole-3-carboxylic acid (69.3 mg, 0.37 mmol) was dissolved in dimethylformamide (0.2M, 1.83 mL), then allylamine (27.41 μL, 0.37 mmol),

HBTU (134.6 mg, 0.354 mmol), and HOBT (54.35 mg, 0.40 mmol) were added, followed by diisopropylethylamine (122.5 μL, 0.70 mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H_2O (5mL) and extracted with DCM (3×5mL). The combined organic layers were washed with 1M HCl (10mL), saturated aqueous $NaHCO_3$ (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over $MgSO_4$, filtered, and evaporated. Purification with flash column chromatography with CH_3OH/CH_2Cl_2 (CH_3OH gradient 0→5%) yielded compound **146** (51.3 mg, 60.8% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.31 7.70 (dt, J = 7.4, 0.9 Hz, 1H), 7.35 – 7.29 (m, 1H), 7.20 (td, J = 7.1, 1.5 Hz, 2H), 6.05 – 5.99 (m, 1H), 5.99 – 5.92 (m, 1H), 5.29 (dq, J = 17.1, 1.6 Hz, 1H), 5.18 (dq, J = 10.2, 1.4 Hz, 1H), 4.14 (tt, J = 5.7, 1.6 Hz, 2H), 3.69 (s, 3H), 2.73 (s, 3H).

Synthesis of 140 146 (51.3mg, 0.22mmol) and Hoveyda-Grubbs catalyst II (15.7mg, 0.022 mmol) were added to a round bottom flask, which was evacuated and filled with nitrogen. CH_2Cl_2 (0.05M, 4.5mL) and 3-butene-2-one (90.8 μL, 1.1mmol) were then added, and the flask was fitted with a reflux condenser and heated to 40°C for 16h. At 16h, TLC showed conversion to product. Filtered through celite and removed CH_2Cl_2 under reduced pressure. Purified by flash column chromatography with CH_3OH/CH_2Cl_2 (CH_3OH gradient 0→5%) yielded compound **140** (27.2 mg, 45% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 – 7.59 (m, 1H), 7.34 – 7.31 (m, 1H), 7.26 – 7.22 (m, 1H), 7.21 (dd, J = 7.1, 1.7 Hz, 1H), 6.87 (dt, J = 16.1, 5.0 Hz, 1H), 6.23 (dt, J = 16.0, 1.8 Hz, 1H), 6.13 (d, J = 6.4 Hz, 1H), 4.31 (ddd, J = 6.1, 5.0, 1.9 Hz, 2H), 3.68 (d, J = 0.8 Hz, 3H), 2.73 (s, 3H), 2.25 (s, 3H). [M+H]: 271.063 Da.

Synthesis of PROTAC 141: General Scheme

147

148

149

150
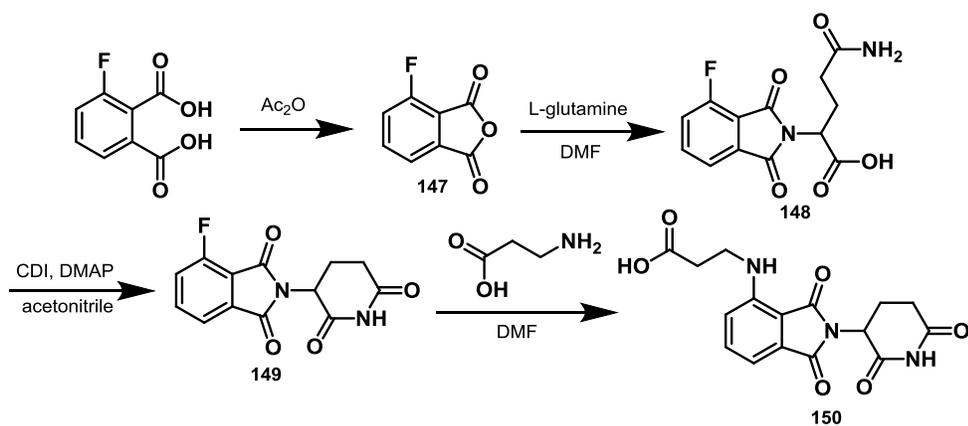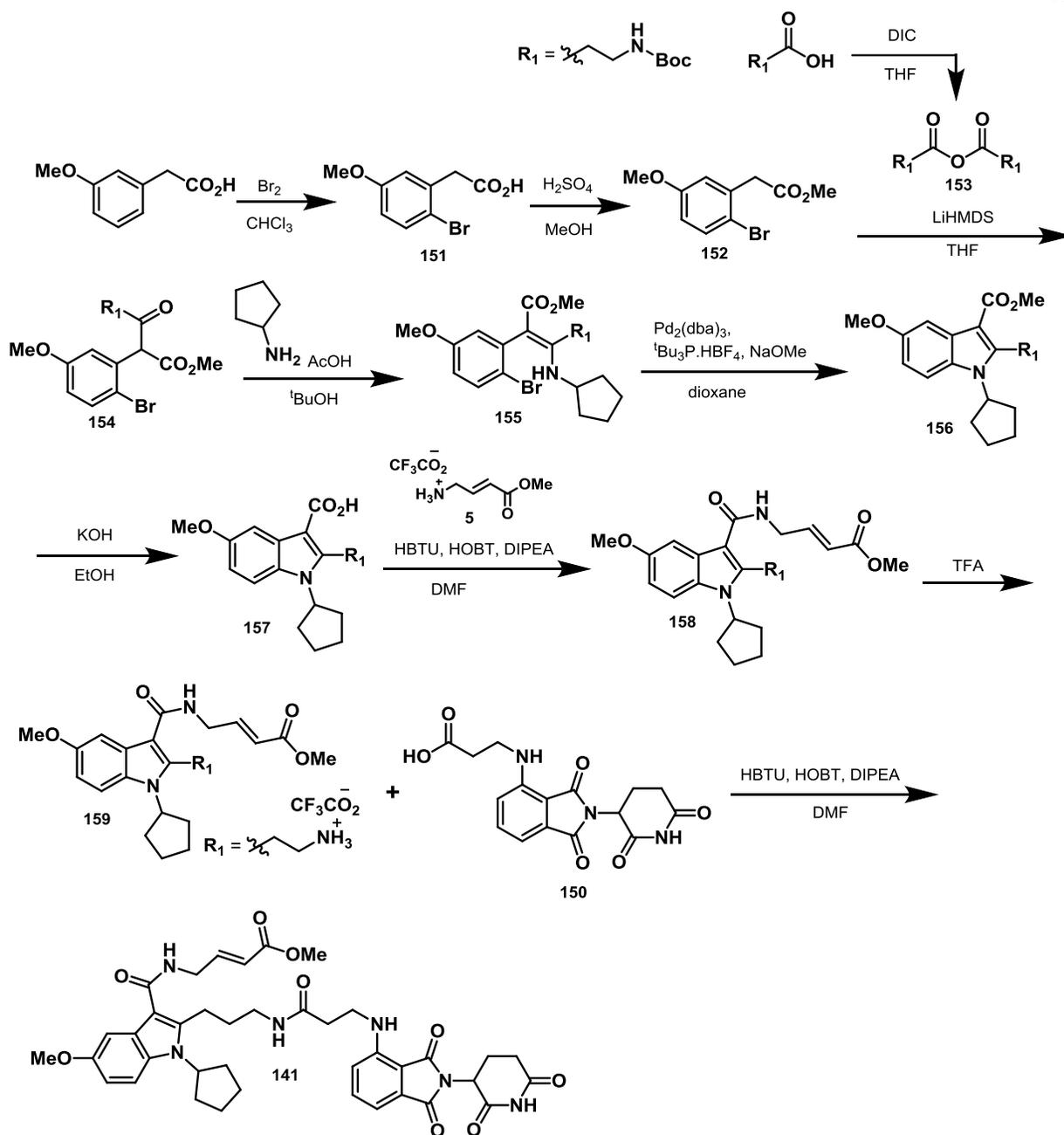
Synthesis of 147 A mixture of 3-fluorophthalic acid (3g, 16.3mmol) in acetic anhydride (0.68M, 24mL) was refluxed for 2 h. The volatiles were removed by vacuum, and the residues were crystallized in acetic anhydride to afford **147** (2.3g, 84% yield). ^{1}H NMR (500 MHz, DMSO-d$_6$) δ 8.06 (ddd, J = 8.3, 7.5, 4.5 Hz, 1H), 7.94 (d, J = 7.5 Hz, 1H), 7.86 (dd, J = 9.3, 8.4 Hz, 1H).

Synthesis of 148 147 (2.3g, 13.8mmol) and L-glutamine (2.023g, 13.8mmol) in dry dimethylformamide (1.2M, 11.5mL) was stirred at 90°C for 8h. The solvent was then removed under reduced pressure. The residue was re-dissolved in 4N HCl (4 mL) and stirred for an additional 8h. The resulting precipitation was collected by filtration, washed with water, and dried to **148** (1.7g, 41% yield) as an off-white solid. ^{1}H NMR (500 MHz, DMSO-d$_6$) δ 8.01 – 7.89 (m, 1H), 7.77 (d, J = 7.4 Hz, 1H), 7.72 (t, J = 8.8 Hz, 1H), 7.18 (s, 1H), 6.71 (s, 1H), 4.74 (dd, J = 10.8, 4.6 Hz, 1H), 2.41 – 2.29 (m, 1H), 2.25 (ddt, J = 14.0, 10.8, 6.7 Hz, 1H), 2.11 (t, J = 7.3 Hz, 2H).

Synthesis of 149 148 (1.3g, 4.3mmol), 1,1′-Carbonyldiimidazole (833mg, 5.14mmol), and 4-(dimethylamino)pyridine (47mg, 0.38mmol) in acetonitrile (1.58M, 2.75mL) was refluxed for 5 h. The reaction mixture was then cooled to room temperature. The resulting solid was collected by filtration, and washed with acetonitrile to afford the crude product. The crude product was purified by flash column chromatography CH$_3$OH/CH$_2$Cl$_2$ (CH$_3$OH gradient 0→5%) to yield **149** (452 mg, 37.9% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.04 (s, 1H), 7.76 (ddd, J = 8.3, 7.4, 4.3 Hz, 1H), 7.73 – 7.67 (m, 1H), 7.41 (td, J = 8.5, 0.9 Hz, 1H), 4.97 (dd, J = 12.5, 5.3 Hz, 1H), 2.96 – 2.84 (m, 1H), 2.85 – 2.78 (m, 1H), 2.78 – 2.67 (m, 1H), 2.24 – 2.08 (m, 1H).

Synthesis of 150 149 (25mg, 0.09mmol) was dissolved in dimethylformamide (0.2M, 0.45mL) and then β-alanine (16mg, 0.18mmol) was added. The reaction was stirred at 90°C for 12 h. The mixture was then cooled to room temperature, poured into water (5 mL) and extracted with ethyl acetate (2x5mL). The combined organic phases were washed with water (5 mL) and brine (5 mL), dried over anhydrous sodium sulfate, and concentrated under reduced pressure. The crude residue was purified by flash column chromatography CH$_3$OH/CH$_2$Cl$_2$ (CH$_3$OH gradient 0→10%) to yield **150** (12 mg, 38.4% yield). ^{1}H NMR (500 MHz, CDCl$_3$) 8.19 (s, 1H), 7.76

(ddd, J = 8.3, 7.3, 4.3 Hz, 1H), 7.70 (dd, J = 7.4, 0.8 Hz, 1H), 7.41 (td, J = 8.5, 0.9 Hz, 1H), 4.97 (dd, J = 12.3, 5.3 Hz, 1H), 2.94 – 2.85 (m, 1H), 2.85 – 2.76 (m, 1H), 2.76 – 2.70 (m, 1H), 2.67 (ddd, J = 7.0, 5.7, 1.6 Hz, 2H), 2.62 (ddd, J = 7.8, 5.7, 1.6 Hz, 2H), 2.19 – 2.09 (m, 1H).

Synthesis of 151 3-methoxyphenylacetic acid (5.1g, 30.9mmol) was dissolved in chloroform (1.54M, 20mL). Bromine (1.67mL, 32.5mmol) was then added dropwise to the soluion. The resulting mixture was allowed to stir overnight at room temperature and then poured into 20% aqueous sodium thiosulfate solution (35.16 mL). The layers were separated and the aqueous layer was extracted with chloroform (200 mL). The combined organics were washed with 20% aqueous sodium thiosulfate solution, dried over $MgSO_4$ and concentrated *in vacuo* to give **151** as an off-white solid (7.6g, 100% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.44 (d, J = 8.8 Hz, 1H), 6.83 (d, J = 3.0 Hz, 1H), 6.71 (dd, J = 8.8, 3.1 Hz, 1H), 3.78 (s, 2H), 3.76 (s, 3H).

Synthesis of 152 **151** (7.6g, 30.9mmol) was dissolved in methanol (1.033M, 30mL). Concentrated sulfuric acid (6 drops) was them added to the solution and the resulting mixture was stirred at room temperature overnight. The reaction was concentrated *in vacuo* to approximately 5mL and the mixture was taken up in EtOAc and washed with 1M NaOH(aq) and then with brine. The organics were dried over $MgSO_4$ and concentrated *in vacuo* to give **152** as a yellow oil (7.1g, 88.7%). ^{1}H NMR (500 MHz, $CDCl_3$) δ δ 7.25 (d, J = 8.8 Hz, 1H), 6.65 (d, J = 3.1 Hz, 1H), 6.52 (dd, J = 8.8, 3.0 Hz, 1H), 3.59 (s, 3H), 3.56 (s, 2H), 3.53 (s, 3H).

Synthesis of 153 4-(*tert*-Butoxycarbonylamino)butyric acid (2g, 9.84mmol) was dissolved in dichloromethane (2M, 4.9mL) and then N,N′-diisopropylcarbodiimide (0.761mL, 4.9mmol) was added. The mixture was stirred for 30min at room temperature, and then filtered to remove the precipitated urea byproduct. The filtrate was evaporated to give **153** (1.47g, 76.9%). ^{1}H NMR

(500 MHz, CDCl$_3$) δ 4.64 (s, 2H), 3.17 (q, J = 6.6 Hz, 2H), 2.49 (t, J = 7.2 Hz, 2H), 1.83 (p, J = 7.0 Hz, 2H), 1.42 (s, 9H).

Synthesis of 154 152 (810mg, 3.12mmol) was dissolved in dry THF (2.18M, 1.43mL) and cooled to -78°C. LiHMDS (1M in THF, 6.88mL) was then added slowly. After stirring for 1h at -78°C, **153** was added. Thesolution was stirred for 1h and subsequently warmed to room temperature. After stirring for another 1h at room temperature, the reaction mixture was partitioned between saturated aqueous NH$_4$Cl (~10mL) and ethyl acetate (10mL). The layers were separated and the aqueous phase was extracted with ethyl acetate (2x10mL). The combined organic extracts were washed with brine, dried over MgSO$_4$, filtered, and concentrated *in vacuo*. The resultant oil was purified via flash column chromatography an ethyl acetate/hexanes gradient 25% EtOAc → 50% EtOAc to produce **154** (1.4g, 100% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 13.02 (s, 1H), 7.53 – 7.42 (m, 1H), 6.94 (d, J = 3.0 Hz, 1H), 6.74 (d, J = 3.1 Hz, 1H), 6.72 (d, J = 0.9 Hz, 1H), 3.77 (d, J = 2.7 Hz, 3H), 3.67 (s, 3H), 3.17 – 2.96 (m, 2H), 2.53 (dt, J = 13.1, 7.1 Hz, 1H), 2.18 – 2.02 (m, 2H), 1.80 – 1.65 (m, 2H), 1.38 (s, 9H).

Synthesis of 155 154 (700mg, 1.58mmol) was dissolved in *tert*-butanol (0.4M, 3.94mL), and then cyclopentylamine (186μL, 1.89mmol) and acetic acid (108 μL, 1.89mmol). The reaction vessel was fitted with a reflux condenser and subsequently heated to reflux for 18 h. After 18 h, the reaction mixture was cooled to room temperature and the solvent was removed *in vacuo*. The resultant oil was partitioned between saturated aqueous NaHCO$_3$ and ethyl acetate. The layers were separated and the aqueous phase was extracted twice with ethyl acetate. The combined organic phases were washed with water, dried with MgSO$_4$, and concentrated *in vacuo* to afford an oil. This resultant oil was purified via flash column chromatography an ethyl acetate/hexanes gradient 25% EtOAc → 50% EtOAc to produce **155** (189mg, 23.5% yield). ^{1}H NMR (500 MHz,

CDCl$_3$) δ 9.55 (d, J = 8.8 Hz, 1H), 7.46 (d, J = 8.7 Hz, 1H), 6.76 (d, J = 3.1 Hz, 1H), 6.68 (dd, J = 8.8, 3.1 Hz, 1H), 4.13 (s, 1H), 3.76 (s, 3H), 3.54 (s, 3H), 2.90 (q, J = 7.6, 6.4 Hz, 2H), 2.17 (ddd, J = 13.4, 11.2, 5.1 Hz, 1H), 1.98 (td, J = 13.1, 12.2, 5.9 Hz, 4H), 1.88 – 1.74 (m, 2H), 1.66 – 1.55 (m, 6H), 1.38 (s, 9H).

Synthesis of 156 155 (189mg, 0.37mmol) was added to a dry round bottom flask equipped with a stir bar and reflux condenser. The system was purged with argon and anhydrous dimethylformamide (0.288M, 1.28mL) was added followed by sodium methoxide (29.9mg, 0.55mmol) and RuPhos (25.9mg, 0.055mmol). The white suspension was thoroughly degassed by bubbling argon gas through the mixture for 10min. RuPhos-Pd-G3 (30.9mg, 0.037mmol) was added and the mixture was stirred at 100°C for 4h. After cooling to room temperature, the reaction mixture was filtered through celite and the bed was washed with ethyl acetate. The filtrate was evaporated under reduced pressure and the residue purified by flash column chromatography an ethyl acetate/hexanes gradient 10% EtOAc → 25% EtOAc to produce **156** (64mg, 40.3% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.63 (d, J = 2.6 Hz, 1H), 7.28 (d, J = 9.0 Hz, 1H), 6.81 (dd, J = 9.0, 2.6 Hz, 1H), 5.18 (s, 1H), 4.78 (p, J = 9.0 Hz, 1H), 3.91 (s, 3H), 3.86 (s, 3H), 3.26 (t, J = 7.5 Hz, 2H), 3.16 (q, J = 6.3 Hz, 2H), 2.36 – 2.19 (m, 2H), 2.12 – 1.98 (m, 4H), 1.79 (dtd, J = 12.9, 6.6, 6.0, 2.2 Hz, 4H), 1.43 (s, 9H).

Synthesis of 157 156 (64 mg, 0.15mmol) was added to a round bottom flask, followed by ethanol (0.391M, 0.38mL) and potassium hydroxide (100mg, 1.78mmol) in water (0.2M, 0.76mL). The mixture was heated at 80ºC for 8h. After cooling and acidification with aq. HCl (2M), the product was extracted twice with ethyl acetate. The organic layer was then dried over MgSO$_4$, filtered, and removed *in vacuo* to afford **157** (35mg, 56.5% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.69 (d, J = 2.6 Hz, 1H), 7.28 (d, J = 9.0 Hz, 1H), 6.84 – 6.76 (m, 2H), 5.06 (s, 1H),

4.79 (p, J = 9.2 Hz, 1H), 3.85 (s, 3H), 3.26 (dd, J = 9.0, 6.5 Hz, 2H), 3.23 − 3.10 (m, 2H), 2.27 (q, J = 9.5, 6.1 Hz, 2H), 2.07 − 2.03 (m, 4H), 1.86 − 1.71 (m, 4H), 1.42 (s, 9H).

Synthesis of 158 157 (35mg, 0.084mmol) was dissolved in dimethylformamide (0.2M, 0.42mL), then **5**, TFA (38.5 mg, 0.17mmol), HBTU (35mg, 0.092mmol), and HOBT (12.5mg, 0.092mmol) were added, followed by diisopropylethylamine (56.3μL, 0.34mmol). The reaction was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H_2O (5mL) and extracted with DCM (3x5mL). The combined organic layers were washed with 1M HCl (10mL), saturated aqueous NaHCO3 (10mL), and saturated aqueous NaCl (10mL). The organic layer was dried over $MgSO_4$, filtered, and evaporated. Purification with flash column chromatography with CH_3OH/CH_2Cl_2 (CH3OH gradient 0→5%) yielded compound **158** (28mg, 65% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33 (dd, J = 9.0, 1.2 Hz, 1H), 7.19 − 7.15 (m, 1H), 7.05 (dtd, J = 16.0, 5.0, 1.3 Hz, 1H), 6.82 (dt, J = 9.0, 1.7 Hz, 1H), 6.07 − 5.98 (m, 2H), 5.95 − 5.83 (m, 1H), 4.76 (p, J = 9.2 Hz, 1H), 4.29 (td, J = 5.4, 4.9, 1.8 Hz, 2H), 3.84 (d, J = 1.2 Hz, 3H), 3.71 (d, J = 1.2 Hz, 3H), 3.18 (t, J = 7.1 Hz, 2H), 3.12 (d, J = 6.2 Hz, 2H), 2.25 (td, J = 8.8, 5.5 Hz, 2H), 2.10 − 2.01 (m, 4H), 1.81 (p, J = 7.0, 6.6 Hz, 4H), 1.38 (s, 9H).

Synthesis of 159 158 (28mg, 0.05mmol) was dissolved in trifluoroacetic acid (83.5μL, 1.09mmol) and stirred at 23°C for 30 min. TLC at 30 min showed conversion to product. TFA was evaporated and azeotroped with toluene (2×5mL). The residue was then dried on high vacuum for 2 hours to yield **158** as the TFA salt (29mg, 100% yield).

Synthesis of 141 150 (5.5mg, 0.015mmol) was dissolved in dimethylformamide (0.075M, 0.2mL), then **158**, TFA (8 mg, 0.015mmol), HBTU (5.7mg, 0.015mmol), and HOBT (2.2mg, 0.016mmol) were added, followed by diisopropylethylamine (7.6μL, 0.45mmol). The reaction

was stirred at 23°C for 16h. TLC at 16h showed conversion to product. The reaction was quenched with H_2O (1mL) and extracted with DCM (3x1mL). The combined organic layers were washed with 1M HCl (1mL), saturated aqueous NaHCO3 (1mL), and saturated aqueous NaCl (1mL). The organic layer was dried over $MgSO_4$, filtered, and evaporated. Purification with flash column chromatography with CH_3OH/CH_2Cl_2 (CH3OH gradient 0→5%) yielded compound **141** (6mg, 53.5% yield). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.37 (t, J = 5.5 Hz, 1H), 7.94 (s, 1H), 7.47 (dd, J = 8.5, 7.1 Hz, 1H), 7.33 (d, J = 9.0 Hz, 1H), 7.16 (d, J = 2.5 Hz, 1H), 7.06 (d, J = 7.1 Hz, 1H), 7.05 – 6.99 (m, 1H), 6.99 – 6.92 (m, 1H), 6.84 (dd, J = 9.0, 2.5 Hz, 1H), 6.54 (t, J = 6.1 Hz, 1H), 6.14 (t, J = 5.9 Hz, 1H), 6.07 – 5.94 (m, 1H), 4.84 (dd, J = 12.2, 5.3 Hz, 1H), 4.75 (t, J = 8.9 Hz, 1H), 4.23 (td, J = 6.1, 5.6, 1.9 Hz, 2H), 3.85 (d, J = 2.8 Hz, 4H), 3.70 (s, 3H), 3.59 (q, J = 6.5 Hz, 2H), 3.16 (q, J = 7.0, 6.6 Hz, 4H), 2.89 – 2.59 (m, 4H), 2.54 (t, J = 6.7 Hz, 2H), 2.24 (q, J = 14.9, 11.5 Hz, 4H), 2.09 – 1.93 (m, 6H), 1.80 (dt, J = 12.3, 6.9 Hz, 6H). ^{13}C NMR (126 MHz, CDCl$_3$) δ 146.68, 144.82, 136.06, 126.52, 121.29, 113.15, 111.47, 102.32, 77.26, 77.21, 77.01, 76.76, 55.99, 51.72, 39.17, 31.36, 30.05, 25.39, 22.74. [M+H]: 741.311 Da.

REFERENCES

(1)   Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. *Nat Rev Drug Discov* **2004**, *3*, 660.

(2)   Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. *Biochemistry* **2012**, *51*, 4990.

(3)   Erlanson, D. A.; Braisted, A. C.; Raphael, D. R.; Randal, M.; Stroud, R. M.; Gordon, E. M.; Wells, J. A. *PNAS* **2000**, *97*, 9367.

(4)   Erlanson, D. A.; Wells, J. A.; Braisted, A. C. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 199.

(5)   Miller, R. M.; Paavilainen, V. O.; Krishnan, S.; Serafimova, I. M.; Taunton, J. *J. Am. Chem. Soc.* **2013**, *135*, 5298.

(6)   Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. *Nat. Rev. Drug. Discov.* **2011**, *10*, 307.

(7)   Hansen, S. K.; Erlanson, D. A. In *Fragment-Based Drug Discovery: A Practical Approach*; John Wiley & Sons, Ltd: 2008, p 245.

(8)   Cardoso, R.; Love, R.; Nilsson, C. L.; Bergqvist, S.; Nowlin, D.; Yan, J.; Liu, K. K.; Zhu, J.; Chen, P.; Deng, Y. L.; Dyson, H. J.; Greig, M. J.; Brooun, A. *Protein Sci.* **2012**, *21*, 1885.

(9)   Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. *Nature* **2010**, *468*, 790.

(10)   Nonoo, R. H.; Armstrong, A.; Mann, D. J. *ChemMedChem* **2012**, *7*, 2082.

(11)   Rosenthal, P. J. *Adv. Exp. Med. Biol.* **2011**, *712*, 30.

(12)   Reddick, J. J.; Cheng, J.; Roush, W. R. *Org. Lett.* **2003**, *5*, 1967.

(13)   Chen, G.; Heim, A.; Riether, D.; Yee, D.; Milgrom, Y.; Gawinowicz, M. A.; D., S. *J. Am. Chem. Soc.* **2003**, *125*, 8130.

(14)   Hanzlik, R. P.; Thompson, S. A. *J. Med. Chem.* **1984**, *27*, 711.

(15)   Ettari, R.; Micale, N.; Schirmeister, T.; Gelhaus, C.; Leippe, M.; Nizi, E.; Di Francesco, M. E.; Grasso, S.; Zappalà, M. *J. Med. Chem.* **2009**, *52*, 2157.

(16)   Patick, A. K.; Brothers, M. A.; Maldonado, F.; Binford, S.; Maldonado, O.; Fuhrman, S.; Petersen, A.; Smith III, G. J.; Zalman, L. S.; Burns-Naas, L. A.; Tran, J. Q. *Antimicrob. Agents Chemother.* **2005**, *49*, 2267.

(17)   Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A. *Drug Discovery Today* **2003**, *8*, 876.

(18)   Golbraikh, A. *Journal of Chemical Information and Computer Sciences* **2000**, *40*, 414.

(19)   Liu, S.; Hanzlik, R. P. *J. Med. Chem.* **1992**, *35*, 1067.

(20)   Krippendorff, B. F.; Neuhaus, R.; Lienau, P.; Reichel, A.; Huisinga, W. *J. Biomol. Screen.* **2009**, *14*, 913.

(21)   Powers, J. C.; Asgian, J. L.; Ekici, O. D.; James, K. E. *Chem. Rev.* **2002**, *102*, 4639.

(22)   Renukuntla, J.; Vadlapudi, A. D.; Patel, A.; Boddu, S. H.; Mitra, A. K. *Int. J. Pharm.* **2013**, *447*, 75.

(23)   Reyes-Turcu, F. E.; Ventii, K. H.; Wilkinson, K. D. *Annual Review of Biochemistry* **2009**, *78*, 363.

(24)   Whitcomb, E. A.; Dudek, E. J.; Liu, Q.; Taylor, A. *Mol. Biol. Cell.* **2009**, *20*, 1.

(25)   Byun, S.; Lee, S. Y.; Lee, J.; Jeong, C. H.; Farrand, L.; Lim, S.; Reddy, K.; Kim, J. Y.; Lee, M. H.; Lee, H. J.; Bode, A. M.; Lee, K. W.; Dong, Z. *Clin. Cancer Res.* **2013**, *19*, 3894.

(26)   Kathman, S. G.; Xu, Z.; Statsyuk, A. V. *J. Med. Chem.* **2014**, *57*, 4969.

(27)   McShan, D.; Kathman, S.; Lowe, B.; Xu, Z.; Zhan, J.; Statsyuk, A.; Ogungbe, I. V. *Bioorg. Med. Chem. Lett.* **2015**, *25*, 4509.

(28)   Steverding, D.; Sexton, D. W.; Wang, X.; Gehrke, S. S.; Wagner, G. K.; Caffrey, C. R. *Int. J. Parasitol.* **2012**, *42*, 481.

(29)   Kathman, S. G.; Statsyuk, A. V. *Med. Chem. Commun.* **2016**, *7*, 576.

(30)   London, N.; Miller, R. M.; Krishnan, S.; Uchida, K.; Irwin, J. J.; Eidam, O.; Gibold, L.; Cimermančič, P.; Bonnet, R.; Shoichet, B. K.; Taunton, J. *Nat Chem Biol* **2014**, *10*, 1066.

(31)   Kitz, R.; Wilson, I. B. *J. Biol. Chem.* **1962**, *237*, 3245.

(32)   Hershko, A.; Ciechanover, A. *Annual Review of Biochemistry* **1998**, *67*, 425.

(33)   Mukhopadhyay, D.; Riezman, H. *Science* **2007**, *315*, 201.

(34)   Jackson, S. P.; Durocher, D. *Molecular Cell* **2013**, *49*, 795.

(35)   Pickart, C. M. *Annu. Rev. Biochem.* **2001**, *70*, 503.

(36)   Deshaies, R. J.; Joazeiro, C. A. *Annu. Rev. Biochem.* **2009**, *78*, 399.

(37)   Huibregtse, J. M.; Scheffner, M.; Beaudenon, S.; Howley, P. M. *Proc. Nat. Acad. Sci.* **1995**, *92*.

(38)   Rotin, D.; Kumar, S. *Nat. Rev. Mol. Cell. Biol.* **2009**, *10*, 398.

(39)   Kim, W.; Bennett, Eric J.; Huttlin, Edward L.; Guo, A.; Li, J.; Possemato, A.; Sowa, Mathew E.; Rad, R.; Rush, J.; Comb, Michael J.; Harper, J. W.; Gygi, Steven P. *Molecular Cell* **2011**, *44*, 325.

(40)   Buckley, D. L.; Crews, C. M. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 2312.

(41)   Hu, Y.; Furtmann, N.; Bajorath, J. *Journal of Medicinal Chemistry* **2015**, *58*, 30.

(42)   Cohen, P.; Tcherpakov, M. *Cell* **2010**, *143*, 686.

(43)   Arkin, Michelle R.; Tang, Y.; Wells, James A. *Chemistry & Biology* **2014**, *21*, 1102.

(44)   Cao, X. R.; Lill, N. L.; Boase, N.; Shi, P. P.; Croucher, D. R.; Shan, H.; Qu, J.; Sweezer, E. M.; Place, T.; Kirby, P. A.; Daly, R. J.; Kumar, S.; Yang, B. *Sci. Signal.* **2008**, *1*, ra5.

(45)   Bernassola, F.; Karin, M.; Ciechanover, A.; Melino, G. *Cancer Cell* **2008**, *14*, 10.

(46)   Shi, Y. J.; Wang, J. R.; Chandarlapaty, S.; Cross, J.; Thompson, C.; Rosen, N.; Jiang, X. J. *Nat. Struct. Mol. Biol.* **2014**, *21*, 522.

(47)   Li, J. J.; Ferry, R. J.; Diao, S. Y.; Xue, B. Z.; Bahouth, S. W.; Liao, F. F. *Endocrinology* **2015**, *156*, 1283.

(48)   Tardiff, D. F.; Jui, N. T.; Khurana, V.; Tambe, M. A.; Thompson, M. L.; Chung, C. Y.; Kamadurai, H. B.; Kim, H. T.; Lancaster, A. K.; Caldwell, K. A.; Caldwell, G. A.; Rochet, J. C.; Buchwald, S. L.; Lindquist, S. *Science* **2013**, *342*, 979.

(49)   Yasuda, J.; Nakao, M.; Kawaoka, Y.; Shida, H. *J. Virol.* **2003**, *77*, 9987.

(50)   Han, Z.; Lu, J.; Liu, Y.; Davis, B.; Lee, M. S.; Olson, M. A.; Ruthel, G.; Freedman, B. D.; Schnell, M. J.; Wrobel, J. E.; Reitz, A. B.; Harty, R. N. *J. Virol.* **2014**, *88*, 7294.

(51)   Boase, N. A.; Kumar, S. *Gene* **2015**, *557*, 112.

(52)   Maspero, E.; Mari, S.; Valentini, E.; Musacchio, A.; Fish, A.; Pasqualato, S.; Polo, S. *EMBO Rep.* **2011**, *12*, 342.

(53)   Maspero, E.; Valentini, E.; Mari, S.; Cecatiello, V.; Soffientini, P.; S.;, P.; Polo, S. *Nat. Struct. Mol. Biol.* **2013**, *20*, 696.

(54)   Kaiser, S. E.; Riley, B. E.; Shaler, T. A.; Trevino, R. S.; Becker, C. H.; Schulman, H.; Kopito, R. R. *Nature Methods* **2011**, *8*, 691.

(55)   Surade, S.; Blundell, Tom L. *Chemistry & Biology* **2012**, *19*, 42.

(56)   Winter, G. *J. Appl. Cryst.* **2010**, *43*, 186.

(57)   McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Cryst.* **2007**, *40*, 658.

(58)   Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. *Acta Cryst. D* **2012**, *68*, 431.

(59)   Emsley, P.; Cowtan, K. *Acta Cryst. D* **2004**, *60*, 2126.

(60)   Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Cryst. D* **2011**, *67*, 235.

(61)   Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Cryst. D* **1997**, *53*, 240.

(62)   Winn, M. D.; Isupov, M. N.; Murshudov, G. N. *Acta Cryst. D* **2001**, *57*, 122.

(63)   Laskowski, R. A.; Rullmannn, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J. Biomol. NMR* **1996**, *8*, 477.

(64)   Adams, P. D.; Afonine, P. V.; Bunkoczi, 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. *Acta Cryst. D* **2010**, *66*, 213.

(65)   Vaguine, A. A.; Richelle, J.; Wodak, S. J. *Acta Cryst. D* **1999**, *55*, 191.

(66)   Read, R. J.; Schierbeek, A. J. J. *J. Appl. Cryst.* **1988**, *21*, 490.

(67)   Kamadurai, H. B.; Qiu, Y.; Deng, A.; Harrison, J. S.; Macdonald, C.; Actis, M.; Rodrigues, P.; Miller, D. J.; Souphron, J.; Lewis, S. M.; Kurinov, I.; Fujii, N.; Hammel, M.; Piper, R.; Kuhlman, B.; Schulman, B. A. *Elife* **2013**, *2*, e00828.

(68)   French, M. E.; Kretzmann, B. R.; Hicke, L. *J. Biol. Chem.* **2009**, *284*, 12071.

(69)   Kim, H. C.; Steffen, A. M.; Oldham, M. L.; Chen, J.; Huibregtse, J. M. *EMBO Rep.* **2011**, *12*, 334.

(70)   Copeland, R. A. *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*; 2nd ed.; John Wiley & Sons, Ltd, 2013.

(71)   Kim, H. C.; Huibregtse, J. M. *Molecular and Cellular Biology* **2009**, *29*, 3307.

(72)   Salvat, C.; Wang, G.; Dastur, A.; Lyon, N.; Huibregtse, J. M. *Journal of Biological Chemistry* **2004**, *279*, 18935.

(73)   Levengood, M. R.; Splain, R. A.; Kiessling, L. L. *Journal of the American Chemical Society* **2011**, *133*, 12758.

(74)   Petroski, M. D.; Deshaies, R. J. *Cell* **2005**, *123*, 1107.

(75)   Pierce, N. W.; Kleiger, G.; Shan, S.-o.; Deshaies, R. J. *Nature* **2009**, *462*, 615.

(76)   Mari, S.; Ruetalo, N.; Maspero, E.; Stoffregen, M. C.; Pasqualato, S.; Polo, S.; Wiesner, S. *Structure* **2014**, *22*, 1639.

(77)   Reyes-Turcu, F. E.; Wilkinson, K. D. *Chem. Rev.* **2009**, *109*, 1495.

(78)   Kathman, S. G.; Span, I.; Smith, A. T.; Xu, Z.; Zhan, J.; Rosenzweig, A. C.; Statsyuk, A. V. *J. Am. Chem. Soc.* **2015**, *137*, 12442.

(79)   Mund, T.; Lewis, M. J.; Maslen, S.; Pelham, H. R. *Proc. Nat. Acad. Sci.* **2014**, *111*, 16736.

(80)   Weerapana, E.; Simon, G. M.; Cravatt, B. F. *Nat Chem Biol* **2008**, *4*, 405.

(81)   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. *Nat Chem Biol* **2014**, *10*, 760.

(82) 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. *Nature* **2016**, *534*, 570.

(83) Su, Y.; Ge, J.; Zhu, B.; Zheng, Y.-G.; Zhu, Q.; Yao, S. Q. *Current Opinion in Chemical Biology* **2013**, *17*, 768.

(84) Yen, H.-C. S.; Elledge, S. J. *Science* **2008**, *322*, 923.

(85) Liu, C.; Choe, V.; Rao, H. *Trends in Biotechnology* **2010**, *28*, 461.

(86) Weiss, W. A.; Taylor, S. S.; Shokat, K. M. *Nat Chem Biol* **2007**, *3*, 739.

(87) Shimoji, T.; Murakami, K.; Sugiyama, Y.; Matsuda, M.; Inubushi, S.; Nasu, J.; Shirakura, M.; Suzuki, T.; Wakita, T.; Kishino, T.; Hotta, H.; Miyamura, T.; Shoji, I. *Journal of Cellular Biochemistry* **2009**, *106*, 1123.

(88) Persaud, A.; Alberts, P.; Amsen, E. M.; Xiong, X.; Wasmuth, J.; Saadon, Z.; Fladd, C.; Parkinson, J.; Rotin, D. *Molecular Systems Biology* **2009**, *5*, 333.

(89) Wang, X.; Trotman, L. C.; Koppie, T.; Alimonti, A.; Chen, Z.; Gao, Z.; Wang, J.; Erdjument-Bromage, H.; Tempst, P.; Cordon-Cardo, C.; Pandolfi, P. P.; Jiang, X. *Cell*, *128*, 129.

(90) Trotman, L. C.; Wang, X.; Alimonti, A.; Chen, Z.; Teruya-Feldstein, J.; Yang, H.; Pavletich, N. P.; Carver, B. S.; Cordon-Cardo, C.; Erdjument-Bromage, H.; Tempst, P.; Chi, S.-G.; Kim, H.-J.; Misteli, T.; Jiang, X.; Pandolfi, P. P. *Cell*, *128*, 141.

(91) Amodio, N.; Scrima, M.; Palaia, L.; Salman, A. N.; Quintiero, A.; Franco, R.; Botti, G.; Pirozzi, P.; Rocco, G.; De Rosa, N.; Viglietto, G. *The American Journal of Pathology*, *177*, 2622.

(92) Maccario, H.; Perera, N. M.; Gray, A.; Downes, C. P.; Leslie, N. R. *Journal of Biological Chemistry* **2010**, *285*, 12620.

(93) Fouladkou, F.; Landry, T.; Kawabe, H.; Neeb, A.; Lu, C.; Brose, N.; Stambolic, V.; Rotin, D. *Proceedings of the National Academy of Sciences* **2008**, *105*, 8585.

(94) Maddika, S.; Kavela, S.; Rani, N.; Palicharla, V. R.; Pokorny, J. L.; Sarkaria, J. N.; Chen, J. *Nat Cell Biol* **2011**, *13*, 728.

(95) Vecchione, A.; Marchese, A.; Henry, P.; Rotin, D.; Morrione, A. *Mol. Cell Biol.* **2003**, *23*, 3363.

(96) Katz, M.; Shtiegman, K.; Tal-Or, P.; Yakir, L.; Mosesson, Y.; Harari, D.; Machluf, Y.; Asao, H.; Jovin, T.; Sugamura, K.; Yarden, Y. *Traffic* **2002**, *3*, 740.

(97) Sun, A.; Yu, G.; Dou, X.; Yan, X.; Yang, W.; Lin, Q. *Molecular Cancer* **2014**, *13*, 248.

(98) Persaud, A.; Alberts, P.; Hayes, M.; Guettler, S.; Clarke, I.; Sicheri, F.; Dirks, P.; Ciruna, B.; Rotin, D. *The EMBO Journal* **2011**, *30*, 3259.

(99) Persaud, A.; Alberts, P.; Mari, S.; Tong, J.; Murchie, R.; Maspero, E.; Safi, F.; Moran, M. F.; Polo, S.; Rotin, D. *Science Signaling* **2014**, *7*, ra95.

(100) Mulvihill, M. J.; Cooke, A.; Rosenfeld-Franklin, M.; Buck, E.; Foreman, K.; Landfair, D.; O'Connor, M.; Pirritt, C.; Sun, Y.; Yao, Y.; Arnold, L. D.; Gibson, N. W.; Ji, Q.-S. *Future Medicinal Chemistry* **2009**, *1*, 1153.

(101) Chen, D.; Zhang, H.; Lu, P.; Liu, X.; Cao, H. *Molecular BioSystems* **2016**, *12*, 614.

(102) Chou, T.-C. *Pharmacological Reviews* **2006**, *58*, 621.

(103) Testa, B.; Mayer, J. M. In *Hydrolysis in Drug and Prodrug Metabolism*; Verlag Helvetica Chimica Acta: 2006, p 365.

(104) Winter, G. E.; Buckley, D. L.; Paulk, J.; Roberts, J. M.; Souza, A.; Dhe-Paganon, S.; Bradner, J. E. *Science* **2015**, *348*, 1376.

(105) Lu, J.; Qian, Y.; Altieri, M.; Dong, H.; Wang, J.; Raina, K.; Hines, J.; Winkler, James D.; Crew, Andrew P.; Coleman, K.; Crews, Craig M. *Chemistry & Biology* **2015**, *22*, 755.

(106) Ito, T.; Ando, H.; Suzuki, T.; Ogura, T.; Hotta, K.; Imamura, Y.; Yamaguchi, Y.; Handa, H. *Science* **2010**, *327*, 1345.

(107) Bondeson, D. P.; Mares, A.; Smith, I. E. D.; Ko, E.; Campos, S.; Miah, A. H.; Mulholland, K. E.; Routly, N.; Buckley, D. L.; Gustafson, J. L.; Zinn, N.; Grandi, P.; Shimamura, S.; Bergamini, G.; Faelth-Savitski, M.; Bantscheff, M.; Cox, C.; Gordon, D. A.; Willard, R. R.; Flanagan, J. J.; Casillas, L. N.; Votta, B. J.; den Besten, W.; Famm, K.; Kruidenier, L.; Carter, P. S.; Harling, J. D.; Churcher, I.; Crews, C. M. *Nat Chem Biol* **2015**, *11*, 611.