SMALL MOLECULE INHIBITORS FOR PROTEIN DEACYLATION TO TREAT CANCER AND COLITIS AND SYNTHESIS OF NUCLEIC ACID AND PROTEIN CROSSLINKERS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Kevin Wielenberg

May 2022

© 2022 Kevin Wielenberg

SMALL MOLECULE INHIBITORS FOR PROTEIN DEACYLATION TO TREAT CANCER AND COLITIS AND SYNTHESIS OF NUCLEIC ACID AND PROTEIN CROSSLINKERS

Kevin Wielenberg, Ph. D.

Cornell University 2022

My graduate research focuses on the development of small molecule probes or inhibitors that can be used in diverse sets of applications.

In chapter 1, I present a new biochemical probe called AP3B that crosslinks and biotinylates nucleic acids. The compound was inspired by a commercially available probe but designed to have better affinity for and reactivity with DNA double helixes. I show that the new crosslinker is more efficient at crosslinking DNA *in vitro*. AP3B also biotinylates nucleic acids in cells better than the commercially available alternative.

In chapter 2, I synthesized a new APT2 inhibitor KW5129. Our lab had previously shown APT2 to be a potential therapeutic target for treating colitis. In this work I attempted to synthesize an APT2 selective covalent inhibitor. Though was met limited success, I synthesized the non-covalent inhibitor KW5129 and show that it is nearly 9 times more potent *in vitro* than the compound it was inspired by, ML349. KW5129 also alleviates dextran sodium sulfate induced inflammation in mice.

The Lin lab has previously synthesized the highly potent SIRT2 selective inhibitor, TM. However, TM's utility is limited by its poor water solubility. I synthesized thioamide and thiourea derivatives of TM that have benzodiazepinedione moieties which improve the water solubility of these compounds. In chapter 3, I present the compound NH-C1-10 as our latest generation SIRT2 inhibitor. I show that NH-C1-10 is an effective SIRT2 selective inhibitor that has broad cytotoxicity in cancer cells and hinders pancreatic

cancer tumor progression in mouse xenografts.

In chapter 4, I present the design and synthesis of several quaternary ammonium containing protein crosslinkers that can be applied for crosslinking mass spectrometry. I identify multiple ways in which using quaternary ammoniums could overcome some of the short comings associated with the well-known crosslinker DSBSO. I show that my crosslinkers work well to both crosslink proteins and identify crosslinked peptides using mass spectrometry. Since this project is still ongoing, future directions are presented including how the crosslinkers are being applied in collaborative projects and how we aim to develop a method for removing dead end crosslinks prior to MS analysis.

BIOGRAPHICAL SKETCH

Kevin Wielenberg was born and raised near the town of Grey Eagle, Minnesota. He attended the University of Minnesota, Duluth from 2013 to 2017 where he earned bachelor's degrees in chemistry and biochemistry and graduated manga cum laude. In the spring of 2017, Kevin was given the Casmir Ilenda award for outstanding undergraduate research for his work studying the synthesis and biological evaluation of benzoxazole conjugates as anticancer agents. Kevin joined the Cornell University Department of Chemistry and Chemical Biology in the summer of 2017 and conducted his doctoral research under the direction of Prof. Hening Lin. Kevin researched and synthesized new APT1/2 and SIRT2 inhibitors as well as DNA and protein crosslinkers. He earned his master's degree from Cornell University in August of 2019. Kevin was named a National Science Foundation Graduate Research Fellowship Program honorable mention for 2018 and was a Cornell University Chemistry Biology Interface Training Program fellow from 2018 to 2020.

ACKNOWLEDGMENTS

I firstly would like to thank my advisor, Prof. Hening Lin, for his mentorship and guidance. Working in his lab has proven to be a challenging yet rewarding experience. I am grateful for his experience, advice, and for allowing me the independence to pursue my own ideas in my graduate research.

I would like to thank my committee members, Prof. Haiyuan Yu and Prof. Frank Schroeder for their expertise and advise which has assisted my research. Prof. Yu has been a wonderful collaborator for my protein crosslinker project. Working with him was a pleasure as he always maintained and upbeat outlook on our projects. I would like to thank Prof. Schroeder for allowing me to rotate in his lab before I had started first semester classes at Cornell. Additionally, the skills I learned from his NMR class proved invaluable to my research.

I would like to thank the Lin lab manager, Xuan Lu. She undoubtedly assisted my research over the past 5 years even more than I realize. I am immensely grateful to her for taking care of ordering supplies, managing group assignments, solving day to day problems, and all the other behind the scenes work she did to keep the lab running. I also would like to thank her for managing and carrying out mouse experiments for me.

My thanks go out to past and current Lin lab members especially Dr. Min Yang, Dr. Jun Young (Nick) Hong, and Dr. Seth Miller. Dr. Yang's extensive synthetic experience helped tremendously whenever I was stuck on a synthetic problem. Dr. Hong and Dr. Miller were great mentors particularly when I first joined the lab. I would also like to thank Dr. Hong and Irma Fernandez for their contributions to chapter 3 of my dissertation as well as all other Lin lab members that contributed to my projects.

I would like to thank all the collaborators I worked with on the 4D Nucleome, RM1, and XL-MS projects, especially Prof. John Lis, Dr. Abdullah Ozer, Dr. Ting-Yi Wang, Dr. Yugandher Kumar, and Dr. Qiuye Zhao.

Lastly, I would like to thank the Howard Hues Medical Institute for their continued funding of the Lin lab and the Cornell Chemistry Biology Interface Training Program for directly funding two years of my graduate research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

An Improved 4'-Aminomethyltrioxsalen-Based DNA Crosslinker for Biotinylation of DNA[a]

Abstract

Nucleic acid crosslinkers that covalently join complementary strands of DNA have applications as both pharmaceuticals and biochemical probes. Psoralen is a popular crosslinking moiety that reacts with double stranded DNA and RNA upon exposure to long wave UV light. The commercially available compound EZ-Link Psoralen-PEG3-Biotin has been used in numerous studies to crosslink DNA and double stranded RNA for genome-wide investigations. Here we present a new probe, AP3B, which uses the psoralen derivative, 4′-aminomethyltrioxsalen, to crosslink and biotinylate nucleic acids. We show that AP3B is 4 to 5 times more effective at labeling DNA in cells and produces a comparable number of crosslinks with over 100 times less compound and less exposure to UV light *in vitro* than EZ-Link Psoralen-PEG3-Biotin.

[a] This is a revised version of our published paper: Wielenberg, K, et. al. An Improved 4'-Aminomethyltrioxsalen-Based DNA Crosslinker for Biotinylation of DNA. ***RSC Adv.***, 2020,**10**, 39870-39874

Introduction

Nucleic acid crosslinkers are important in nucleic acid research and come in many varieties[1, 2]. These range from simple molecules such as formaldehyde to inorganic complexes[3] and have been combined with moieties that can preferentially bind to a target DNA sequence[4]. Crosslinkers can covalently bond to nucleotides by reacting with nucleophilic regions of the bases or the double bonds of the bases. In the latter case, activation of the compound by exposure to light is frequently required to induce pericyclic reactions[5, 6].

Photoactivable DNA and RNA crosslinkers are particularly useful for genomic studies as they allow for spatial and temporal control of the crosslinking reactions by controlling the light source[7]. Psoralen is one such photoactivatable moiety and has natural affinity for DNA due to the compound's planar multicyclic structure which allows it to intercalate nucleotide base pairs. Upon exposure to UV light, psoralen undergoes 2+2 cycloaddition to the 5,6 double bond of pyrimidines; however, it reacts best with thymine bases of complementary 5`-TA-3` dinucleotide stretches[8]. This forms mono and bifunctional adducts,[9] with bifunctional adducts making interstrand crosslinks with the two complementary strands (Figure 1.1).

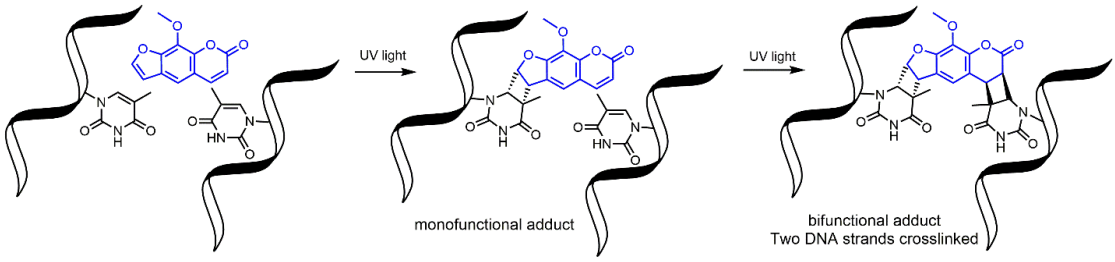


Figure 1.1. Psoralen crosslinking. Psoralen undergoes 2+2 cycloaddition reactions to form mono and bifunctional adducts with DNA.

EZ-Link Psoralen-PEG3-Biotin (PP3B, Figure 1.2 A) is a commercially available probe that utilizes psoralen to biotinylate nucleic acids. Such crosslinkers enable cost effective biotinylation of double-stranded DNA (dsDNA) and dsRNA for experiments including sequencing of psoralen crosslinked, ligated, and selected hybrids (SPLASH), studying RNA interactions, and Chem-seq to study the accessibility of DNA[10-13]. Despite the prevalent use of PP3B for these applications, we noticed that other psoralen derivatives such 4′-aminomethyltrioxsalen (AMT) are more effective nucleic acid crosslinkers than the derivative used in PP3B[14]. In this paper, we present the synthesis of a new psoralen derivative-biotin compound, AMT-PEG3-biotin (AP3B, Figure 1.2 B), and show that it is more efficient at both labeling and crosslinking DNA than the commercially available PP3B.

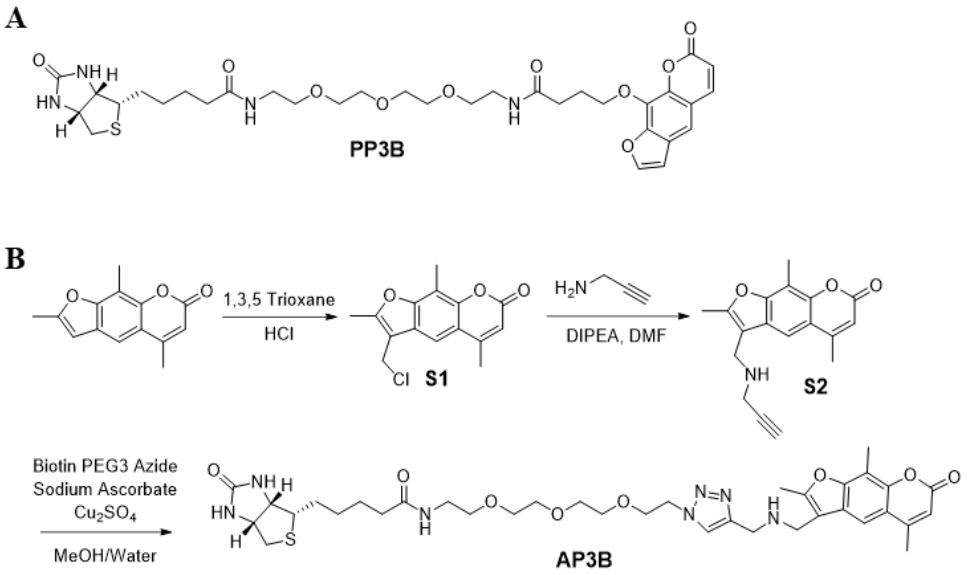


Figure 1.2. Structure of PP3B (**A**) and synthesis of AP3B (**B**).

Results and Discussion

Design and Synthesis of AP3B

AMT was used as the reactive group in our probe as it has several advantages over other psoralen derivatives. Firstly, AMT has increased affinity for DNA due in part to an amino group which becomes protonated *in vivo* and interacts with the negatively charged phosphate backbone of DNA and RNA[15]. The amino group also improves water solubility which we further enhanced in our new compound by including a PEG3 chain. Finally, AMT has a broader range of UV absorbance near 360 nm which allows the compound to absorb more photons near this wavelength compared to other psoralens. This results in increased probability that cycloaddition reactions will occur with AMT[16]. To complete the probe, biotin is used as the affinity tag since biotin streptavidin interaction is the best affinity reagent, with 10^{-14} M affinity and tolerance to harsh conditions such as high temperature (up to 65 ^{o}C) and denaturing conditions (up to 0.5% SDS). The resulting compound is AMT-PEG3-Biotin (AP3B, Figure 1.2B).

AP3B was synthesized from commercially available starting materials (Figure 1.2B). Briefly, 4'-chloromethyl-4,5',8-trimethylpsoralen was reacted with 1,3,5 trioxane in concentrated HCl to functionalize the five-member ring with a chloromethyl group. The alkyl chloride was then substituted with propargyl amine to attach a terminal alkyne which was then reacted with biotin PEG3 azide via click chemistry to produce AP3B.

AP3B has increased crosslinking efficiency in vitro

We examined the efficiency of AP3B to form interstrand crosslinks in purified DNA with that of the commercially available PP3B. For this purpose, we used a denaturing agarose gel mobility shift assay. The use of high pH of the gel prevents base pairing by disrupting hydrogen bonding, there by producing single stranded DNA (ssDNA)[17]. Interstrand crosslinks formed with AP3B or PP3B prevent strands from separating, producing up shifted, higher molecular weight dsDNA bands. We compared the *in vitro* efficiency of AP3B and PP3B and confirmed that the reaction of both compounds with DNA is UV light dependent as expected. The results showed that 800 nM AP3B and 100 μM PP3B were the two most comparable samples, producing similar amounts dsDNA (Figure 1.3).

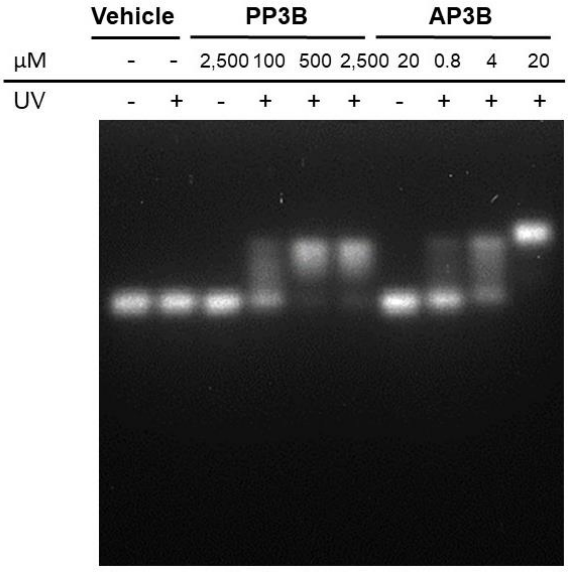


Figure 1.3. Detecting DNA interstrand crosslinks with a denaturing agarose gel. Interstrand crosslinks result in shifting from low mass ssDNA to higher mass dsDNA bands. Both compounds show dependence on exposure to UV light. AP3B shows significant signal depletion of ssDNA bands at concentrations where PP3B has no apparent activity.

The relative rates of interstrand crosslink formation for the two compounds were also compared using the mobility shift assay (Figure 1.4A). In this case, the UV exposure time was varied while the concentrations of PP3B and AP3B were fixed at the highest concentration previously tested (2.5 mM and 20 μM, respectively). These concentrations produced the maximal amount of dsDNA which serves as a reference point for the end of the reaction. As expected, AP3B requires less time to reach completion (5-10 min) compared to PP3B (20-30 min) (Figure 1.4B).

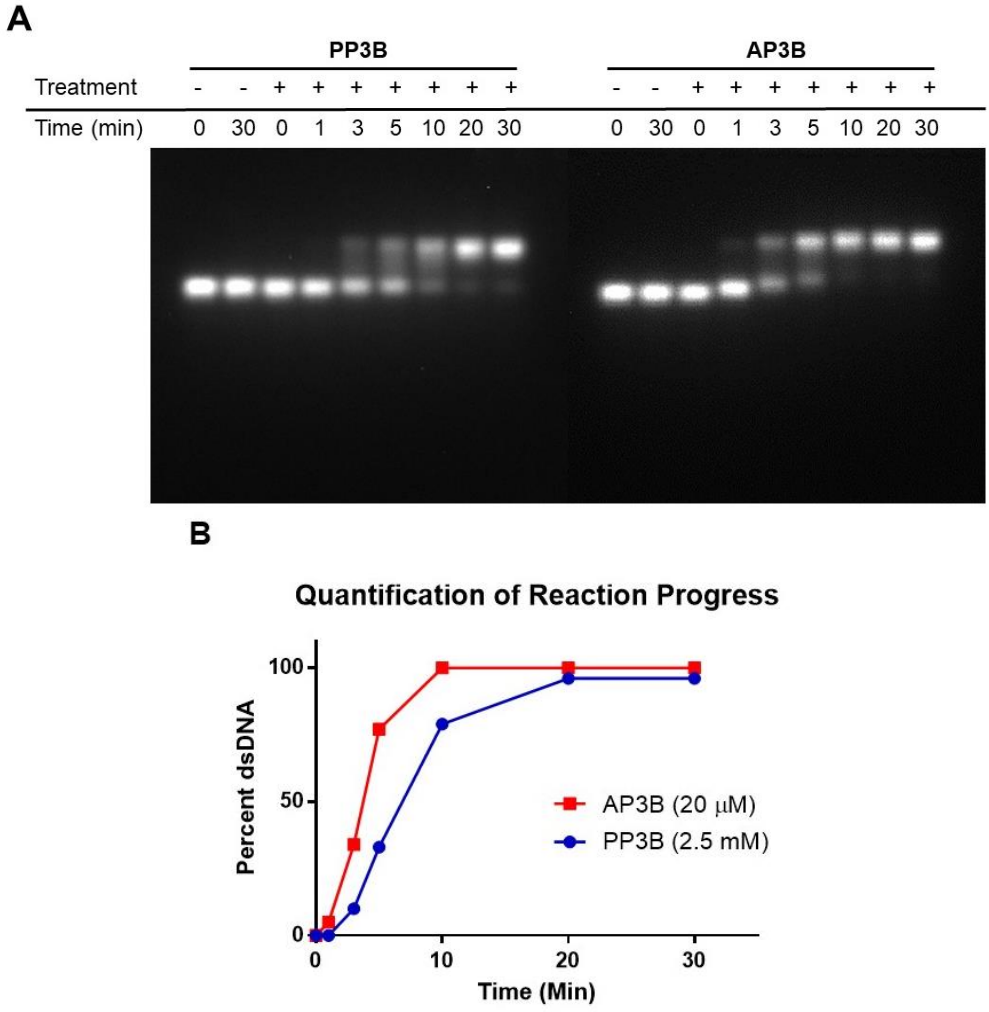


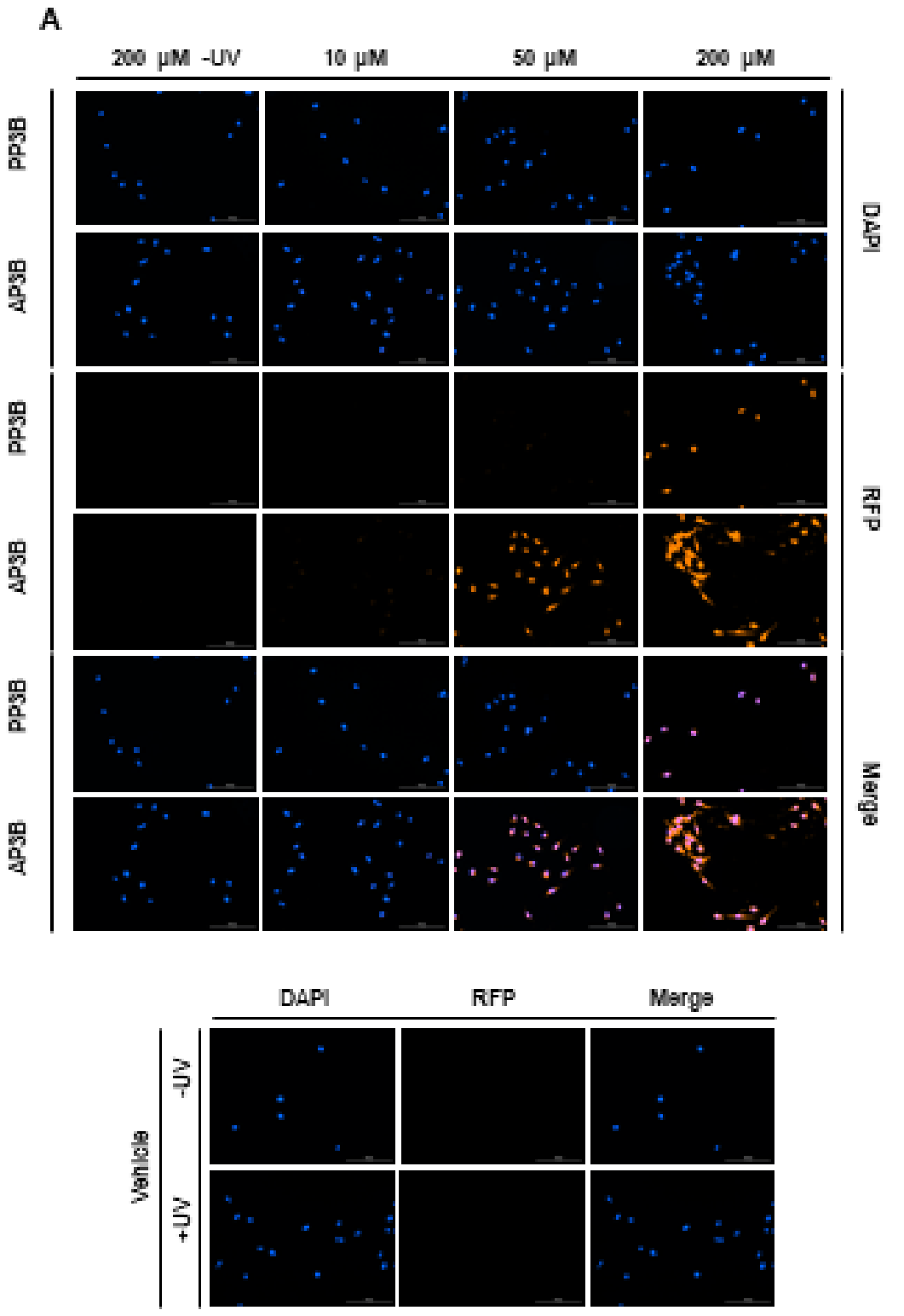
Figure 1.4. Reaction rate comparison of PP3B and AP3B. (**A**) The concentrations of PP3B and AP3B in the treated samples were fixed at 2.5 mM and 20 μM respectively. 6 UV exposure times between 1 and 30 min were compared. (**B**) Quantification of

reaction progress. The integral of the fluorescent intensities of the ssDNA and dsDNA bands were calculated to determine the precent of the total signal for each time point that is in the dsDNA band. 100% dsDNA indicates a complete reaction.

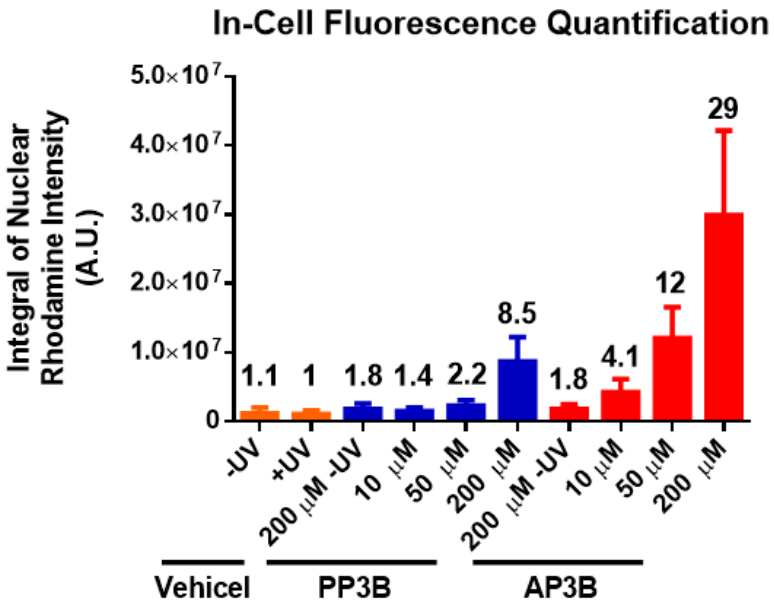
AP3B shows higher labeling efficiency in cells than PP3B

We next determined how the DNA labeling efficiency (measuring both mono-adducts and interstrand crosslinks) of AP3B compares to PP3B in cells. An immunofluorescence assay was developed to compare the two compounds. Hela cells were treated with PP3B, AP3B, or DMSO and exposed to UV light. Unreacted compound was washed away, and the bound probes were recognized by a Streptavidin-rhodamine conjugate which was in turn detected by florescence microscopy. Colocalization of DAPI and rhodamine signals in a cell's nuclei were taken to result primarily from DNA biotinylation, while cytoplasmic staining is due to the probes reacting with dsRNA and possibly proteins. Visual examination of the merged DAPI and rhodamine channels (Figure 1.5A) shows increased biotinylation in the AP3B treated cells compared to the same concentration of PP3B. This observation is confirmed by quantification of the signal intensity of rhodamine in cell's nuclei (Figure 1.5B). In all cases treated cells had higher rhodamine signal intensities than the vehicle treated, or non-UV exposed cells treated with the same concentration of compound. By comparing the signal intensities with the same concentrations of PP3B and AP3B or by comparing the concentrations of PP3B and AP3B that produced similar signal intensities, we concluded that the in-cell labeling efficiency of AP3B is 4 to 5 times that of PP3B.

Figure 1.5. Evaluation of the relative labeling efficiency of PP3B and AP3B in cells. (**A**) In cell fluorescence detection of biotinylation in Hela cells incubated with the indicated concentration of PP3B, AP3B or vehicle (10 μL DMSO). (**B**) Quantification of the intergal of nuclear rhodamine singal intensity with bars normalized the vehicle +UV intensity.


A
200 μM -UV
10 μM
50 μM
200 μM
PP3B
AP3B
PP3B
AP3B
PP3B
AP3B
DAPI
RFP
Merge
DAPI
RFP
Merge
Vehicle
-UV
+UV

B



.

The in-cell results indicate the biotinylation efficiency increase for AP3B is lower than that seen in the gel shift assay. This is due at least in part to the fact that the two assays detect different adducts. While the in-cell experiment reports both mono and bifunctional adducts, the *in vitro* assay detects bifunctional adducts. The discrepancy in the results suggest that AP3B is moderately better than PP3B reacting with nucleic acids in general, however, when focusing on bifunctional adducts, the reaction efficiencies skew heavily in favor of AP3B. This is likely due in part to the increased reaction rate of AP3B with DNA as demonstrated (Figure 1.4B). Since both ends of the crosslinking moiety must react to form interstrand crosslinks, the effect of the increased reaction rate is compounded when only bifunctional adducts are examined. Other factors such as the accessibility of DNA to the crosslinkers in these assays may also account for some of the difference in labeling efficiency reported by the two assays. In cells, DNA is highly compacted in heterochromatin and wrapped around histones in euchromatin. It is

possible that the proteins associated with DNA interfere with the reaction of AP3B more than with the reaction of PP3B, and thus decrease the labeling efficiency of AP3B more. With the *in vitro* assay however, naked DNA was used, and no proteins were present to hinder the reaction.

In summary, we have synthesized a more efficient nucleic acid crosslinking probe containing a photoactivable AMT moiety. Our experiments have shown that AP3B is more reactive with DNA than the currently available product PP3B in terms of both the amount of compound and UV exposure time required to produce similar result. Fluorescent detection of the compounds in cells shows a 5-fold increase in biotinylation for our compound compared to PP3B. The fold-increase is even higher (over 100 times) *in vitro* with naked DNA. These results suggest that AP3B is a highly efficient tool for DNA biotinylation and could potentially be used to improve the efficiency of SPLASH[10], Chem-seq[7], and other experiments.

Materials and Methods

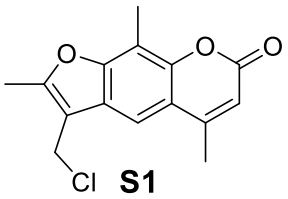
All solvents and reagents were purchased from commercial vendors as analytical or higher-grade purity. Flash chromatography was done using SiliaFlash Irregular Silica Gel, P60, 40 – 63 μm, 60 Å. NMR spectra were collected at the Cornell NMR Facility using a Brucker 500 spectrometer. HRMS data was collected at the Cornell Chemistry Mass Spectrometry Facility using a Thermo Exactive Orbitrap ESI mass spectrometer.

In-cell fluorescence assay

Hela cells were plated on 35 mm imaging petri dishes in DMEM supplemented with 10% FBS and grown for 24 h at 37 ^{o}C with 5% CO_2. The next day the media was removed followed by a PBS wash (1 mL). Cells were then incubated at 37 ^{o}C with 1 mL of blocking solution (1X PBS, 5% BSA, .1% Saponin) for 30 min. PP3B, AP3B, or vehicle (10 μL DMSO) was added and incubated for 10 min at 37 ^{o}C. The solution was removed, and the cells were washed once with PBS to remove excess compound. The dishes that needed to be exposed to UV light were placed on ice and exposed to 368 nm light (5, 1.2 W bulbs) at a distance of 6 cm for 30 min using a UV oven during which time cell death occurred. The unexposed samples were kept under ambient light for 30 min. All dishes were then washed with 40% formamide, PBS, and blocking solution to remove unreacted molecules at which point cell death occurred for the non-UV exposed cells. The cells were then incubated with a streptavidin-rhodamine conjugate (TheremoFisher Scientific, catalog number 21724) solution (1 μg/mL in blocking solution) at 4 ^{o}C for 1 h. Unbound conjugate was removed by 3 washes with blocking solution (1 mL each). Anti-fade reagent containing DAPI (30 μL) was added to each dish and the fluorescent intensities of the DAPI and rhodamine signals were quantified using a Cytation 5 cell imaging multi-mode reader. Integrals of nuclear rhodamine intensity were calculated by taking the integral of the rhodamine signal that overlapped with DAPI signals.

DNA electrophoretic mobility shift assay

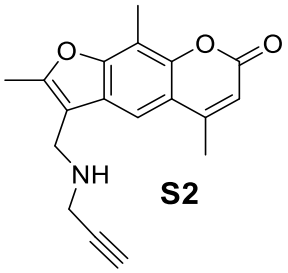
A DNA fragment (~500bp) was purified after PCR amplification for this experiment. A denaturing agarose gel (2% agarose, 30 mM NaOH, 2 mM EDTA, pH 8) was prepared. Reaction solutions were prepared by mixing purified DNA (20 μL, 30 ng/μL), 1 μL of the crosslinker solutions (prepared at different concentrations in DMSO) or DMSO (control without crosslinker). The reaction solutions were placed on ice and exposed to a 368 nm light source (5, 1.2 W bulbs) 6 cm away for 30 min using a UV oven. Then 10x running buffer was added to make the samples a 1x solution. The samples heated to 95 ^{o}C for 3 min prior to loading onto the gel. The gel was run for 180 min at 45 V then stained with a solution of 1X SYBR Gold in 0.5 M Tris at pH 7.5 for 18 h prior to imaging.



S1

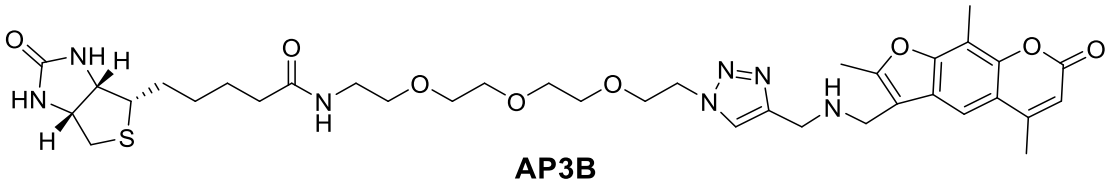
3-(chloromethyl)-2,5,9-trimethyl-7H-furo[3,2-g]chromen-7-one (S1). To a solution of 1,3,5 Trioxane (0.71g, 7.9 mmol) in concentrated HCl (45 mL) was added 2,5,9-trimethyl-7H-furo[3,2-g]chromen-7-one (2.05g 9.0 mmol). The reaction mixture was stirred at room temperature for 16 h then diluted with water (50 mL) and extracted three times with chloroform (100 mL each). The organic layers were combined, washed with brine (100 mL), dried with Na_2SO_4, then concentrated with a rotary evaporator. The crude product was recrystallized from ethyl acetate to yield the final compound (1.19g, 4.3 mmol, 54%).^{1}H NMR (500 MHz, Chloroform-*d*) δ 7.59 (s, 1H), 6.26 (s, 1H), 4.74

(s, *2*H), 2.57 (s, 3H), 2.52 (s, 6H); ^{13}C NMR (126 MHz, Chloroform-*d*) δ 161.3, 155.2, 154.6, 153.12, 149.5, 123.9, 116.4, 113.2, 112.2, 111.2, 109.6, 36.2, 19.4, 12.3, 8.5.



S2

2,5,9-trimethyl-3-((prop-2-yn-1-ylamino)methyl)-7H-furo[3,2-g]chromen-7-one (S2). S1 (0.90g, 0.36 mmol) and propargylamine (0.23 mL, 36 mmol) were combined in toluene (50 mL). The reaction mixture was refluxed for 16 h after which the solution was cooled to room temperature and concentrated with a rotary evaporator. The crude residue was then purified by column chromatography (Hexane: EtOAc 1:1 → 1:2) to give the final product (0.80g 76%). ^{1}H NMR (500 MHz, DMSO): δ 7.79 (s, 1H), 6.31 (s, 1H), 3.87 (s, 2H), 3.27 (s, 2H), 3.14 (s, 1H) 2.47 (s, 6H) 2.44 (s,3H); ^{13}C NMR (126 MHz, DMSO) δ 160.1, 154.1, 153.9, 153.7, 148.3, 125.3, 115.5, 112.8, 112.5, 112.0, 107.5, 82.9, 73.8, 36.2, 18.7, 11.9, 8.2;



AP3B

5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-N-(2-(2-(2-(2-(4-((((2,5,9-trimethyl-7-oxo-7H-furo[3,2-g]chromen-3-yl)methyl)amino)methyl)-1H-1,2,3-triazol-1-yl)ethoxy)ethoxy)ethoxy)ethyl)pentanamide (AP3B). S2 (0.16g, .54

mmol) and Biotin-PEG3-Azide (0.24g, 0.54 mmol) were dissolved in a solution of methanol and water (1:1, 14 mL). $NaHCO_3$ (64 mg, 0.76 mmol), $Cu_2SO_4 \cdot 5H_2O$ (14 mg, .054 mmol) and sodium ascorbate (21 mg, 0.11 mmol) were added to the reaction vessel. This solution was stirred at room temperature for 16 h. The solution was extracted three times with DCM (14 mL). The pooled organic layers were washed with brine (30 mL), dried with Na_2SO_4, and then the solvent was removed with a rotary evaporator. The crude was purified by column chromatography (DCM: MeOH 10:1, 8:1, and then 6:1) to give the final product (0.37g, 0.50 mmol, 93%) ^{1}H NMR (500 MHz, Methanol-d4) δ 7.98 (s, 1H), 7.82 (s, 1H), 6.28 – 6.22 (m, 1H), 4.60 (t, J = 5.0 Hz, 2H), 4.49 (dd, J = 7.9, 4.9 Hz, 1H), 4.29 (dd, J = 7.9, 4.5 Hz, 1H), 3.97 (s, 4H), 3.91 (t, J = 5.1 Hz, 2H), 3.65 – 3.58 (m, 4H), 3.58 – 3.52 (m, 4H), 3.49 (t, J = 5.5 Hz, 2H), 3.22 – 3.14 (m, 1H), 2.92 (dd, J = 12.8, 4.9 Hz, 1H), 2.70 (d, J = 12.8 Hz, 1H), 2.56 (s, 3H), 2.49 (d, J = 13.7 Hz, 6H), 2.18 (t, J = 7.4 Hz, 2H), 1.78 – 1.51 (m, 4H), 1.41 (p, J = 8.7, 8.1 Hz, 2H); ^{13}C NMR (126 MHz, Methanol-d_4) δ 174.6, 164.7, 162.1, 155.3, 155.2, 154.5, 148.7, 145.1, 125.4, 123.8, 115.9, 112.2, 112.0, 111.5, 108.2, 70.13, 70.08, 70.02, 69.8, 69.2, 69.0, 62.0, 60.2, 55.6, 50.0, 42.7, 40.7, 39.6, 38.9, 35.3, 28.3, 28.1, 25.4, 18.1, 10.8, 7.0. HRMS (ESI) m/z: calculated for $C_{36}H_{49}O_8N_7S$ $[M+H]^+$: 740.3436, found 740.3413.

References

1. Sharma, E.; Sterne-Weiler, T.; O'Hanlon, D.; Blencowe, Benjamin J., Global Mapping of Human RNA-RNA Interactions. *Molecular Cell* **2016,** *62* (4), 618-626.
2. Lu, Z.; Zhang, Q. C.; Lee, B.; Flynn, R. A.; Smith, M. A.; Robinson, J. T.; Davidovich, C.; Gooding, A. R.; Goodrich, K. J.; Mattick, J. S.; Mesirov, J. P.; Cech, T. R.; Chang, H. Y., RNA Duplex Map in Living Cells Reveals Higher-Order

Transcriptome Structure. *Cell* **2016,** *165* (5), 1267-1279.
3. Dasari, S.; Tchounwou, P. B., Cisplatin in cancer therapy: molecular mechanisms of action. *Eur J Pharmacol* **2014,** *740*, 364-378.
4. Baraldi, P. G.; Cozzi, P.; Geroni, C.; Mongelli, N.; Romagnoli, R.; Spalluto, G., Novel benzoyl nitrogen mustard derivatives of pyrazole analogues of distamycin A: synthesis and antileukemic activity. *Bioorganic & Medicinal Chemistry* **1999,** *7* (2), 251-262.
5. Sakamoto, T.; Tanaka, Y.; Fujimoto, K., DNA Photo-Cross-Linking Using 3-Cyanovinylcarbazole Modified Oligonucleotide with Threoninol Linker. *Organic Letters* **2015,** *17* (4), 936-939.
6. Song, C.-X.; He, C., Bioorthogonal labeling of 5-hydroxymethylcytosine in genomic DNA and diazirine-based DNA photo-cross-linking probes. *Acc Chem Res* **2011,** *44* (9), 709-717.
7. Dekker, J.; Belmont, A. S.; Guttman, M.; Leshyk, V. O.; Lis, J. T.; Lomvardas, S.; Mirny, L. A.; O'Shea, C. C.; Park, P. J.; Ren, B.; Politz, J. C. R.; Shendure, J.; Zhong, S.; Network, D. N., The 4D nucleome project. *Nature* **2017,** *549* (7671), 219-226.
8. Zhen, W. P.; Buchardt, O.; Nielsen, H.; Nielsen, P. E., Site specificity of psoralen-DNA interstrand cross-linking determined by nuclease Bal31 digestion. *Biochemistry* **1986,** *25* (21), 6598-6603.
9. Couvé-Privat, S.; Macé, G.; Rosselli, F.; Saparbaev, M. K., Psoralen-induced DNA adducts are substrates for the base excision repair pathway in human cells. *Nucleic Acids Res* **2007,** *35* (17), 5672-5682.
10. Aw, Jong Ghut A.; Shen, Y.; Wilm, A.; Sun, M.; Lim, Xin N.; Boon, K.-L.; Tapsin, S.; Chan, Y.-S.; Tan, C.-P.; Sim, Adelene Y. L.; Zhang, T.; Susanto, Teodorus T.; Fu, Z.; Nagarajan, N.; Wan, Y., In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation. *Molecular Cell* **2016,** *62* (4), 603-617.
11. Dadonaite, B.; Gilbertson, B.; Knight, M. L.; Trifkovic, S.; Rockman, S.; Laederach, A.; Brown, L. E.; Fodor, E.; Bauer, D. L. V., The structure of the influenza A virus genome. *Nature Microbiology* **2019,** *4* (11), 1781-1789.
12. Zhong, X.; Feng, L.; Xu, W.-H.; Wu, X.; Ding, Y.-D.; Zhou, Y.; Lei, C.-Q.; Shu, H.-B., The zinc-finger protein ZFYVE1 modulates TLR3-mediated signaling by facilitating TLR3 ligand binding. *Cellular & Molecular Immunology* **2019**.
13. Anders, L.; Guenther, M. G.; Qi, J.; Fan, Z. P.; Marineau, J. J.; Rahl, P. B.; Lovén, J.; Sigova, A. A.; Smith, W. B.; Lee, T. I.; Bradner, J. E.; Young, R. A., Genome-wide localization of small molecules. *Nat Biotechnol* **2014,** *32* (1), 92-96.
14. Isaacs, S. T.; Shen, C.-K. J.; Hearst, J. E.; Rapoport, H., Synthesis and characterization of new psoralen derivatives with superior photoreactivity with DNA and RNA. *Biochemistry* **1977,** *16* (6), 1058-1064.
15. Hearst, J. E., Psoralen Photochemistry and Nucleic Acid Structure. *Journal of Investigative Dermatology* **1981,** *77* (1), 39-44.
16. Oroskar, A.; Olack, G.; Peak, M. J.; Gasparro, F. P., 4'-Aminomethyl-4,5',8-trimethylpsoralen photochemistry: the effect of concentration and UVA fluence on photoadduct formation in poly(dA-dT) and calf thymus DNA. *Photochem Photobiol*

1994, *60* (6), 567-73.
17. Sambrook, J.; Russell, D. W., Alkaline Agarose Gel Electrophoresis. *Cold Spring Harbor Protocols* **2006,** *2006* (1), pdb.prot4027.

CHAPTER 2

Boronic Acid-Containing Acyl Protein Thioesterase Inhibitors Protect Mice in an Inflammatory Bowel Disease Model[a]

Abstract

Acyl Protein Thioesterase 2 (APT2) is a serine hydrolase that removes palmitoyl groups from acylated cysteine residues. Recently, the proinflammatory transcription factor STAT3 was reported to be depalmitoylated by APT2, which promotes the translocation of phosphorylated STAT3 to the nucleus where it can promote the expression of RORC and IL17A genes which induces the differentiation of $CD4^+$ T cells into T helper 17 cells. The activation of STAT3 contributes to the development of inflammatory bowel diseases (IBD) and APT2 is considered a therapeutic target for IBD treatment. To this end, we have synthesized a series of new APT2 inhibitors based on the APT2 selective inhibitor ML349. One of these compounds, KW5129, which contains a boronic acid moiety, is about nine times more potent than ML349 *in vitro* but inhibits both APT1 and APT2. KW5129 increases the level of STAT3 palmitoylation in cells and alleviates the symptoms of colitis in mice. Our study provides important insights for the development of better APT1/2 inhibitors and further supports APT2 as a potential target for inflammatory and autoimmune diseases that involve STAT3 activation.

[a] This is a revised version of our submitted manuscript: Boronic Acid-Containing Acyl Protein Thioesterase Inhibitors Protect Mice in an Inflammatory Bowel Disease Model. Wielenberg, K.; Liu, Y.; Hong, J. Y. ; Yang, M.; Anmangandla, A.; Lu, X,; Lin, H.

Introduction

Cysteine palmitoylation (S-palmitoylation) is a post-translational modification wherein cysteine residues are acylated with palmitic acid. Since the acyl group is attached through a labile thioester bond, the modification is readily enzymatically reversible allowing for dynamic control of the modification[1]. S-palmitoylation is installed by about two dozen zinc finger aspartate-histidine-histidine-cysteine containing proteins (DHHC) and removed by several serine hydrolyzes, including acyl protein thioesterase 1 (APT1 or LYPLA1), APT2 (or LYPLA2), ABHD17A/B/C, and ABHD10[2]. S-palmitoylation can affect protein membrane localization, trafficking, activity, stability, and protein-protein interactions [3-5].

While many reports point to the S-palmitoylation of immune signaling proteins, such as NOD2 and STING, the molecular understanding of the function of S-palmitoylation in immune signaling remains limited[6, 7]. Recently, a palmitoylation-depalmitoylation cycle catalyzed by DHHC7 and APT2 is reported to regulate STAT3, providing a new model for how dynamic palmitoylation could regulate immune signaling[8]. The transcription activator STAT3 aids the differentiation of $CD4^+$ T cells into T Helper 17 (T_H17) cells by promoting the transcription of certain genes, such as RORC and IL17A [9, 10]. STAT3 is activated by the phosphorylation of Y705 by Janus kinase 2 (JAK2)[11] in response to certain cytokine stimulation, such as IL-6. STAT3 C108 palmitoylation catalyzed by DHHC7 promotes the membrane targeting of STAT3, where JAK2 is localized, thus promoting phosphorylation by JAK2. However, subsequent nuclear translocation requires the depalmitoylation by APT2. Thus, the palmitoylation-depalmitoylation cycle together promotes STAT3 activation. Therefore, breaking this

palmitoylation cycle by either disrupting DHHC7 or APT2 could inhibit STAT3 and suppress T_H17 differentiation[8]. Increased T_H17 cell differentiation contributes to inflammatory bowel disease (IBD)[12]. Hence, inhibiting DHHC7 or APT2 could help to treat IBD, as demonstrated in an IBD mouse model[8]. Additionally, APT2 inhibitors have been reported to have potential to treat melanoma[13].

APT1 and 2 share 65% sequence identity and several small molecule inhibitors have been developed for them[14, 15]. The β-lactone Palmostatin B was developed as a covalent APT inhibitor and blocks the catalytic serine by esterifying it. Though very potent, Palmostatin B is not APT1/2 selective and inhibits many serine hydrolases[16 17, 18]. More recently, aided by high-throughput screening studies, certain piperazine amides have been found to be effective inhibitors of APT1 and APT2. Specifically, ML348 and ML349 have been identified as selective, noncovalent inhibitors of APT1 and APT2, respectively. However, neither is as potent *in vitro* as Palmostatin B[18]. Together Palmostatin B, ML348, and ML349 have proven to be highly useful for examining numerous cellular functions of APT1 and APT2, including regulating Scribble localization and protumor signaling, regulating Ras signaling, and regulating MC1R palmitoylation in melanoma[13, 19-21]. The identification of APT2 as a promising target for treating IBD further emphasizes the potential utility of APT2 inhibitors. Similarly, the APT1-selective inhibitor ML348 has been reported to rescue behavior and neuropathology in Huntington disease mice[22]. Given these promising reports, developing more and improved APT1 and APT2 inhibitors will be critical to further test the therapeutic potential of inhibiting APT1 and APT2.

Here we present new covalent and noncovalent APT1 and APT2 inhibitors that are derived from ML349. We show that these compounds are more potent APT1 and APT2 inhibitors *in vitro* and in cells, more effectively prevent STAT3 depalmitoylation, and alleviate the inflammation of Dextran Sodium Sulfate (DSS) induced colitis in mice.

Results and Discussion

New APT2 inhibitor design based on ML349

Currently the most potent APT2-selective inhibitor is ML349. Thus, we used ML349 as the starting point to design and synthesize new APT2 inhibitors. An X-ray crystal structure of APT2 in complex with ML349 showed that the methoxy group of ML349 is located near a relatively hydrophobic channel (Figure 2.1 A)[23]. Thus, we designed MY1 and MY2 (Figure 2.2) to determine whether elongating the methoxy group could enhance APT2 inhibition. MY3 – MY7 (Figure 2.2) were designed to replace the benzene ring or the piperazine ring to see whether they are critical for APT2 inhibition.

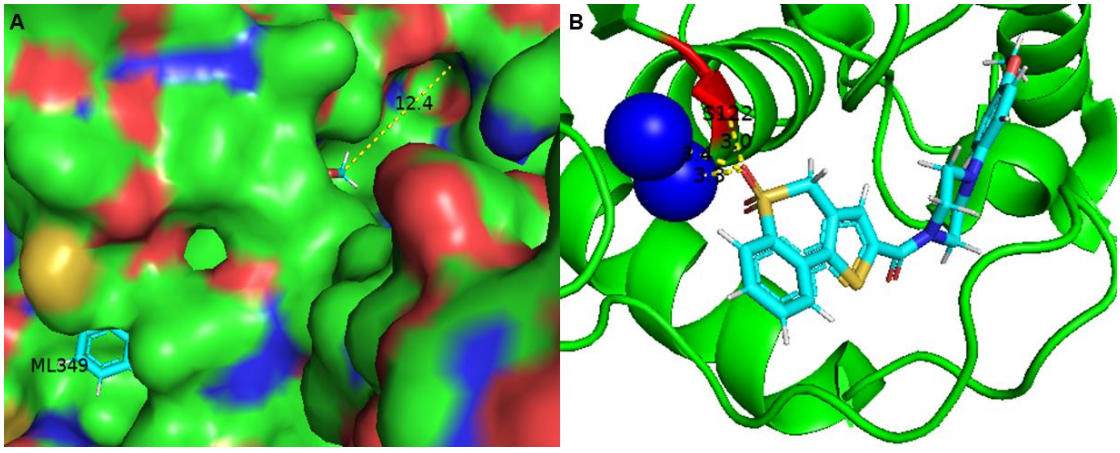


Figure 2.1. Crystal structure for APT2 crystallized with ML349. The para methoxy group of ML349 points toward an unoccupied channel in APT2 (**A**). The sulfone oxygens of ML349 interact with two structural water molecules located near the APT2

active site and are positioned equally near to the catalytic S122 (**B**). The indicated distances are measured in angstroms.

The structure of APT2 in complex with ML349 also showed that the sulfone moiety is positioned near the catalytic serine (S122) and hydrogen bonds to two water molecules near the oxyanion hole (Figure 2.1 B) [23]. We explored whether changing the sulfone moiety could affect APT2 inhibition (Figure 2.2). In MY8, we changed the sulfone in ML349 to a sulfoxide. In KW5108 and KW5116, we used a diethylphosphonate and an ethyl fluorophosphonate, respectively. With the ethyl fluorophosphonate, we hypothesized that it may become a covalent inhibitor by reacting with the catalytic serine residue. In KW5129, we used a boronic acid to mimic the sulfone moiety of ML349. In KW5175 and KW5192, we kept the boronic acid group but modified the ML349 tricyclic ring. In KW5130 and KW5191, we used a hydroxyl group and a sulfonyl fluoride to mimic the sulfone in ML349.

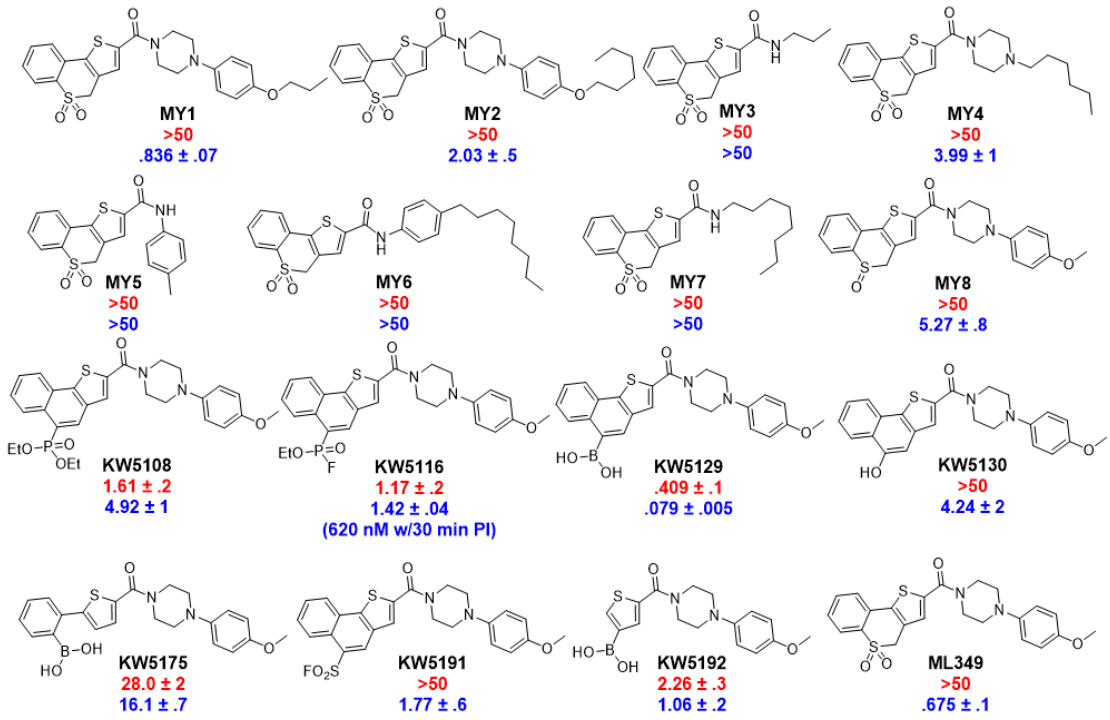


Figure 2.2. The designs of the new inhibitors were derived from ML349. Warheads and long alkyl chains were incorporated better to engage the APT2 binding pocket and the catalytic residue S122. *In vitro* IC_{50} values for APT1 (red) and APT2 (blue) are listed with standard deviations. Inhibition of APT2 with KW5116 increased when the preincubation (PI) time was increased to 30 min.

The MY series of compounds were synthesized based on the established synthesis of ML349[23]. KW5108, KW5116, KW5129, KW5130, and KW5191 followed similar synthetic routes except the initial starting material used was a-tetralone (Figure 2.3). After constructing the thiophene ring, the middle ring was oxidized to make a fully conjugated system by NBS bromination followed by HBr elimination in the same pot. The naphtho[1,2-b]thiophene product readily underwent electrophilic aromatic substitution with bromine or chlorosulfonic acid selectively at the 5 position. The aryl bromide product was cross-coupled with palladium to attach a boronic acid (KW5129),

or diethyl phosphonate which was hydrolyzed and fluorinated to produce KW5116. The remaining compounds in the KW series required their own unique synthetic routes.

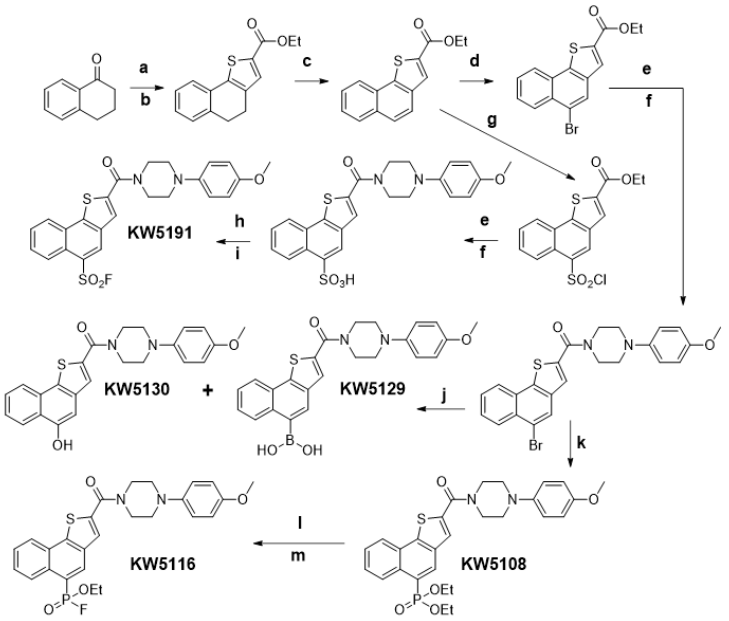


Figure 2.3. Synthesis of KW5108, KW5116, KW5129, KW5130 and KW5191 a) $POCl_3$, DMF, b) Sodium, EtOH, ethyl thioglycolate, c) NBS, benzoyl peroxide, d) Br_2, acetic acid, e) NaOH, f) EDC, DMAP, 1-(4-methoxyphenyl)piperazine, g) chlorosulfonic acid, h) NaOH, $POCl_3$, i) KHF_2, j) bis boronic acid, Xphos, Xphos-Pd-G2, KOAc. k) $Pd(OAc)_2$ PPh_3, DIPEA, Diethyl Phosphonate, l) LiBr, m) Diethylamino Sulfur trifluoride, KW5130 was isolated as a side product of reaction j.

KW5129 and KW5116 are more potent APT2 inhibitors than ML349

The potency of the new compounds against APT1 and APT2 was measured *in vitro* with a fluorogenic assay. Resorufin acetate was used as a fluorogenic substrate to monitor reaction progress[18]. Deacetylation of resorufin acetate by APT1/APT2 produces resorufin which is strongly fluorescent at 584 nm, allowing the reaction to be easily monitored over time to calculate the IC_{50} values of each compound (Figure 2.2).

Extending the para methoxy group with longer alkyl chains (MY1 and MY2) did not improve the potency. Even a short extension as in MY1 was detrimental. Replacing the

benzene or piperazine ring (MY3-MY7) demonstrated that the piperazine amide moiety is critical for APT2 inhibition as its removal abolishes APT2 inhibition while the benzene ring is not as critical.

Modification of the sulfone group in the KW compounds provided better results. The boronic acid-containing KW5129 is ~9 times more potent at inhibiting APT2 compared with ML349 (Figure 2.2). The other two boronic acid compounds (KW5175 and KW5192) without the full tricyclic ring in ML349, were still able to inhibit APT2. KW5192 has an IC_{50} (1.1 μM) that is only slightly worse than that of ML349, while KW5175 is much less potent.

The ethyl fluorophosphonate compound KW5116 gave an IC_{50} value (1.4 μM) slightly worse than that of ML349. However, when the preincubation period was doubled from 15 minutes to 30 minutes, its IC_{50} dropped to 620 nM, which is slightly better than that of ML349, suggesting that KW5116 is a covalent inhibitor. Similarly, the sulfonyl fluoride compound KW5191 inhibits APT2 with an IC_{50} value of 1.8 μM, which is slightly worse than KW5116. Additionally, other compounds that remove the sulfone moiety of ML349 (MY8, KW5108, and KW5130) can inhibit APT2 with IC_{50} of ~5 μM.

We also tested the compounds' inhibition effects on the close homolog, APT1. Although ML349 is selective for APT2, many of the new inhibitors lost APT2 selectivity and start to inhibit APT1 (Figure 2.2). These include the phosphonate compounds KW5108 and KW5116, as well as the boronic acid compound KW5129. KW5129 inhibits APT1 with an IC_{50} value of 0.41 μM, which is about 5.2-fold higher than that (0.079 μM) for APT2 and similar to the reported APT1 selective inhibitor ML348, which we found to have an

APT1 IC_{50} value of 0.82 μM. The fact that these compounds can inhibit both APT1 and APT2 raises an interesting question of how selectivity is achieved, which requires future investigations. Being dual APT1/APT2 inhibitors, some of these compounds could allow us to test whether inhibiting both APT1/APT2 could be problematic or beneficial in later cellular and animal studies.

For further characterization, we decided to focus on two compounds, KW5116 and KW5129, which showed the best APT2 inhibition potency, with KW5116 being potentially a covalent inhibitor. We first determined if KW5116 and KW5129 were covalent or noncovalent inhibitors of APT2. We used a gel filtration assay adapted from a previous paper[18] (Figure 2.4 A). A cell lysate containing APT2 was treated with the APT2 inhibitors and then passed over a size exclusion column multiple times to remove unbound small molecules. The samples were then incubated with tetramethyl rhodamine fluorophosphonate (TAMRA-FP), which can covalently label the active site serine residue of APT2. In the presence of an inhibitor, APT2 will not be fluorescently labeled with TAMRA-FP. If the inhibitor is a non-covalent inhibitor, after passing over the column, the inhibitor will be lost and APT2 will be fluorescently labeled with TAMRA-FP. In contrast, with a covalent inhibitor, APT2 will remain unlabeled by TAMRA-FP even after multiple gel filtrations. KW5129 showed almost complete fluorescent labeling after just one gel filtration while KW5116 showed none after three passes, indicating the fluorophosphonate moiety allowed it to become a covalent inhibitor. In contrast, the boronic acid-containing compound KW5129 is a non-covalent inhibitor of APT2.

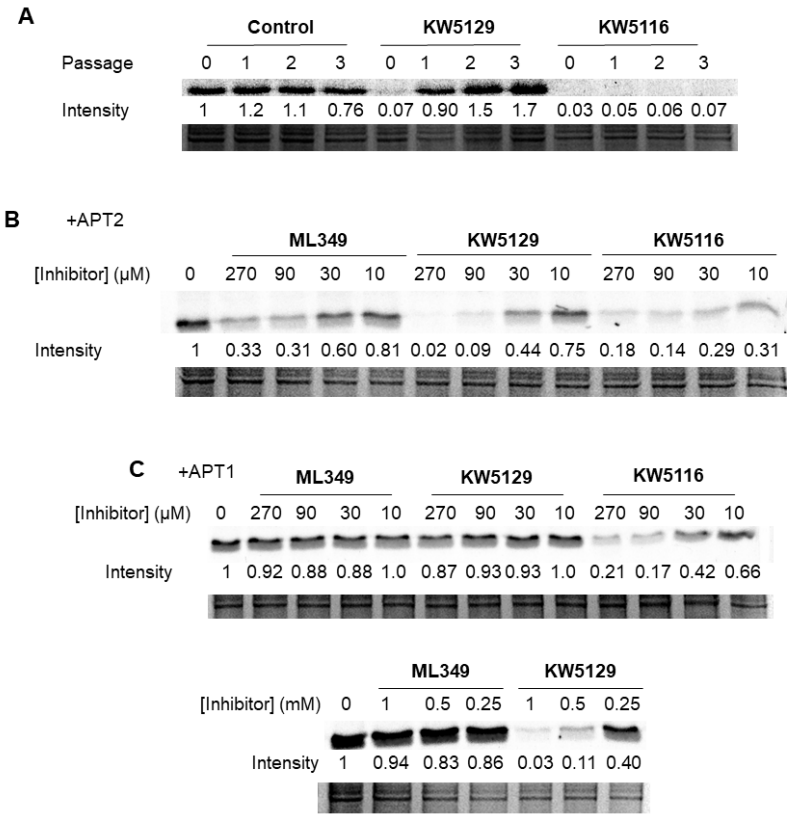


Figure 2.4. TAMRA-FP ABPP shows the mode of action and potency of ML349, KW5116, and KW5129. The fluorescent intensities were normalized using the loading intensity of each lane. After treatment with inhibitor the protein solution was passed over a size exclusion column multiple times then treated with TAMRA-FP (**A**). A cell lysate solution spiked with APT2 (**B**) or APT1 (**C**) was treated with inhibitor for 15 min then TAMRA-FP for 30 min.

We next used *in vitro* competitive activity-based protein profiling (ABPP) to determine the relative potency of KW5116 and KW5129 to that of ML349. A cell lysate containing APT2 was treated with different concentration of ML349, KW5116, or KW5129 and labeled with TAMRA-FP. In this case, compounds with higher binding affinities would better out compete TAMRA-FP for the active site and lower intensity bands would be observed with more potent compounds (Figure 2.4 B). Quantification of the fluorescent intensities of the bands showed that addition of ML349 at 90 μM or 270 μM was only

able to inhibit ~70% of fluorescent labeling intensity. KW5129 almost completely abolished TAMRA-FP fluorescence labeling of APT2 at 270 μM. At 90 μM, KW5129 inhibited APT2 labeling by about 90%. The covalent inhibitor KW5116 produced 70% signal reduction even at a 10 μM and thus was the strongest among the three. In addition, consistent with the IC_{50} values measured with purified proteins (Figure 2.2), KW5116 and KW5129 could also inhibit APT1 in a similar competitive ABBP assay (Figure 2.4 C). All the results indicate that KW5129 and KW5116 are more potent APT2 inhibitor than ML349 but can also inhibit APT1.

KW5116 and KW5129 increase STAT3 palmitoylation

To test whether KW5116 and KW5129 can inhibit APT2 in cells, we detected the S-palmitoylation of STAT3 in HEK 293T cells. Based on previous studies, STAT3 is palmitoylated mainly by DHHC7 on C108 and is depalmitoylated by APT2 especially for phosphorylated STAT3 (pSTAT3). To visualize the palmitoylation level of STAT3, HEK 293T cells were co-transfected with Flag-tagged STAT3 and HA-tagged DHHC7 to increase STAT3 palmitoylation. The cells were treated with different APT2 inhibitors and incubated with alkyne-14 (Alk14), which allowed for later fluorescent detection of STAT3 palmitoylation with click chemistry.

Since only STAT3 and DHHC7 were expressed, this method allowed us to inhibit the endogenous APT2. The overall increase of the palmitoylated STAT3 signal was not dramatic, partially because APT2 prefers to depalmitoylate pSTAT3. With a fixed 16-hour treatment of the inhibitors, we compared KW5116 and KW5129 to ML349. By quantification, 50 μM of ML349 slightly increased the palmitoylation signal of STAT3.

At the same concentration, KW5116 and KW5129 were more potent (Figure 2.5). In addition, we observed a significant increase of palmitoylated STAT3 with only 25 μM of KW5116 and KW5129. Overall, KW5129 and KW5116 showed similar effects, and both appeared to be stronger than ML349 at increasing STAT3 palmitoylation.

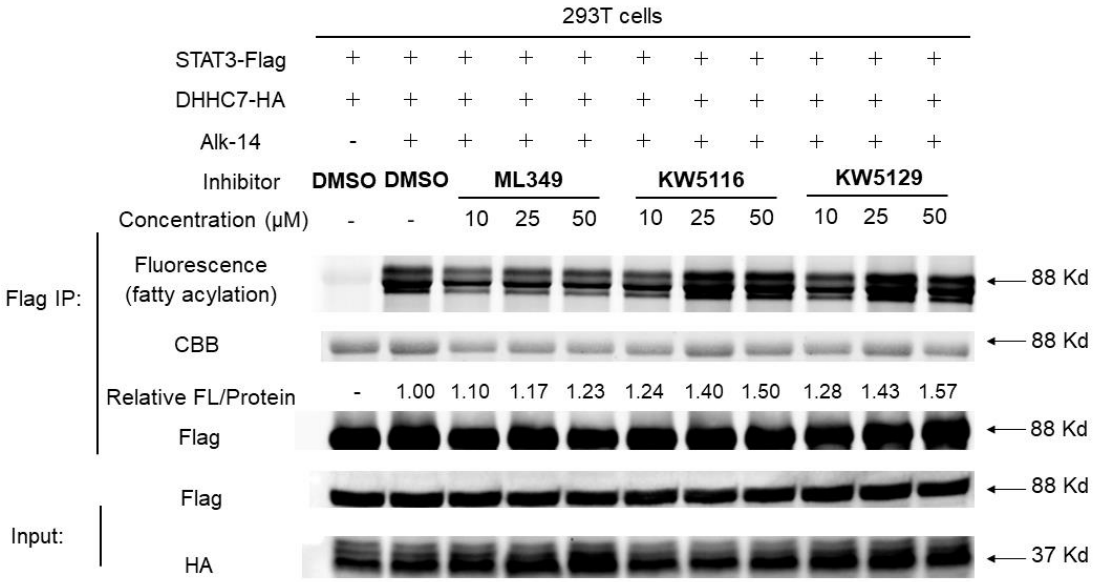


Figure 2.5. KW5116 and KW5129 increased STAT3 palmitoylation in cells. HEK 293T cells expressing flag-tagged STAT3 and HA-tagged DHHC7 were treated with APT2 inhibitors, and metabolically labeled with Alk14 for 6 hours. After obtaining the cell lysate, STAT3 was immunoprecipitated and fluorescently labeled via click chemistry to detect its acylation level. The acylation level was normalized by the protein level as measured by coomassie brilliant blue (CBB) staining.

KW5129 reduces the symptoms of colitis in mice

We then tested the ability of KW5116 and KW5129 to alleviate the symptoms of IBD in mice. The mice were grouped by sex and supplied with dextran sodium sulfate (DSS) in their drinking water to induce colitis. Simultaneously, they were treated with APT2 inhibitors (50 mg/kg) by intraperitoneal injection for 11 days after which the colons were removed and measured. Before testing the compounds in the DSS model, we first carried out an initial toxicity assessment of KW5116 and KW5129. While KW5129 is

well tolerated at 50mg/kg daily injection, KW5116 appears to be toxic. This could be due to KW5116 reacting with off target proteins as was seen with the ABPP experiments (Figure 2.6). Thus, in the DSS model, we only compared KW5129 to ML349.

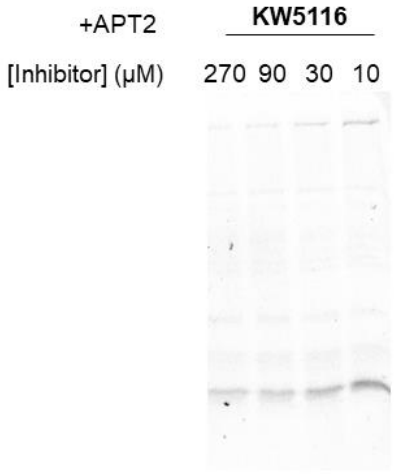


Figure 2.6. TAMRA-FP ABPP with KW5116. Significant fluorescent intensity is lost in bands not corresponding to APT2.

In all cases, treatment with ML349 or KW5129 showed a significant increase in the lengths of the colons compared to the control groups (Figure 2.7 A). However, while KW5129 appeared to be slightly better than ML349, there was no statistically significant difference between the groups treated with ML349 and KW5129. Female groups responded better to treatment than did the male groups. Body weight loss was less severe in females treated with ML349 and KW5129 and the colons were an average of 1.4 cm longer than the control group. In contrast, the colon length difference was only ~ 0.5 cm and there was no statistically significant difference in body weight in the male mice (Figure 2.7 B). The results indicate that KW5129 is at least as effective at

attenuating the symptoms of colitis in mice as ML349 which has been previously shown to treat DSS included colitis in mice [8].

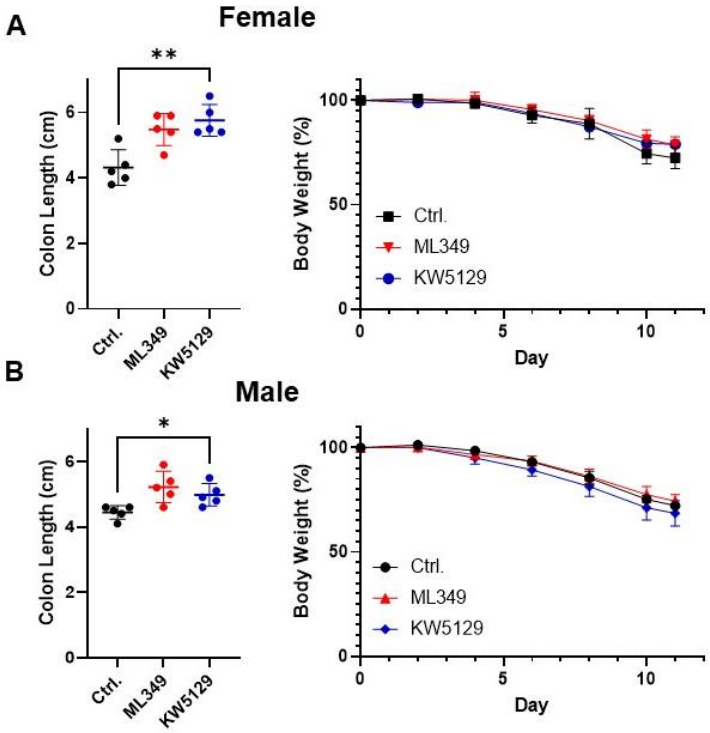


Figure 2.7. DSS mouse colitis model experiment with ML349 and KW5129. Groups of female (**A**) (** P value = .0022), and male (**B**) (* P value = 0.166) mice were treated with 2.5% DSS in their drinking water and injected with ML349 or KW5129 starting on the same day as DSS treatment started. Body weights were tracked throughout the experiment. After sacrificing the mice, their colons were removed and measured.

In summary, our study further confirmed the previous finding that APT2 inhibitors have the potential to function as IBD therapeutics. We synthesized 15 new compounds based on the APT2 inhibitor ML349 and their potency was tested *in vitro*. From these, we selected KW5116 and KW5129 as our lead compounds based on their IC_{50} values. Competitive ABPP showed KW5116 and KW5129 to be a covalent and a noncovalent inhibitor, respectively, and can inhibit APT2 more potently than ML349. Experiments in HEK 293T cells showed a better increase of STAT3 palmitoylation level compared

to ML349. In mouse studies, the covalent inhibitor KW5116 appears to be toxic, likely because it targets proteins other than APT1/APT2, while KW5129 is well tolerated. In the DSS-induced colitis mouse model, treatment with ML349 or KW5129 resulted in a reduction in colon shrinkage indicating a reduction in colitis symptoms. We have therefore demonstrated that our new APT2 inhibitor KW5129 is more potent *in vitro* than ML349 and that KW5129 also inhibits STAT3 depalmitoylation and alleviates symptoms of colitis in mice.

Our results highlight several interesting issues. First, the selectivity between APT1 and APT2 can be affected by very small changes in the inhibitor structure. Even though we started with the APT2-selective inhibitor ML349, many of the changes to the inhibitor structure led to dual inhibitors instead of APT2 selective inhibitors. Thus, future investigations are warranted to understand how the selectivity is determined. Second, our studies identified several warhead structures (boronic acid and phosphonate), which are different from the original sulfone group that interacts with the catalytic serine residue. These findings will be very helpful for future APT1/APT2 inhibitor development. On the therapeutic front, our studies suggest that a dual APT1/APT2 inhibitor KW5129 is well tolerated and did not cause obvious toxicity, which is similar to the APT2-selective inhibitor ML349. Furthermore, our studies here underscore the potential for the use of APT2 inhibitors as IBD therapeutics. The results also hint that sex may affect how well the mice respond to treatment with APT2 inhibitors though more experiments are required to validate this observation.

Materials and Methods

General methods

All solvents and reagents were purchased from commercial vendors as analytical grade or higher purity. Flash chromatography was done using SiliaFlash Irregular Silica Gel, P60, 40 – 63 µm, 60 Å. NMR spectra were collected at the Cornell NMR Facility using a Brucker 500 spectrometer. HRMS data was collected at the Cornell Chemistry Mass Spectrometry Facility using a Thermo Exactive Orbitrap ESI mass spectrometer. For *in vitro* IC_{50} assays, a Cytation 5 cell imaging multi-mode reader was used to monitor the fluorescent intensities with excitation and emission wavelengths set to 535 nm and 590 nm respectively. ActivX TAMRA-FP serine hydrolase probe was purchased from Thermo Fisher Scientific (catalog number 88318). PD-10 Columns prepacked with Sephadex G-25 M for gel filtration assays were purchased from Cytiva.

Cell culture and transfections.

HEK 293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and kept in an incubator set to 37^{o}C with 5% CO_2. To express ZDHHC7 and STAT3, plasmids with the corresponding gene inserts were transfected into 293T cells using polyethylenimine (PEI) (24765, Polysciences) according to the manufacturer's protocol. Flag immunoprecipitation was done with Anti-Flag M2 affinity gel (Millipore Sigma) according to the manufacturer's protocol.

Expression and purification of APT1 and APT2

Human APT1 and APT2 containing a C-terminal 6xHis tag in pQE60 vector were

transformed into BL21(DE3) chemically competent *E. coli*. Typically, 6-8 L of LB broth with 100 µg/mL Ampicillin was inoculated with an overnight starter grown at 30°C. Cultures were grown at 30°C for ~4 hours until the OD_{600} reached 0.6-0.8, then induced with 1 mM IPTG and incubated at 18° C overnight to allow protein expression. Cells were harvested by centrifugation at 6000 *g*. Cell pellets were frozen at -80°C or immediately used for purification. Pellets were resuspended in lysis buffer (50 mM phosphate buffer, pH 8.0, 100 mM NaCl, 10% glycerol, 0.5 mg/mL lysozyme, 100 µM leupeptin, and Pierce universal nuclease). Following a 30 min incubation, cells were sonicated on ice with an ultrasonic liquid processor (Fisherbrand Sonic Dismembrator, FB-505 equipped with a 0.5-inch diameter probe, FB4220) for 4 min total at 60% amplitude. Lysate was clarified by centrifugation at 4°C and 30,000 *g* for 35 min. Clarified lysate was incubated with Ni-NTA resin, washed with 50 mL of wash buffer (50 mM phosphate buffer, pH 8.0, 100 mM NaCl, 10% glycerol, 10 mM imidazole), and eluted in 2 mL of elution buffers containing a 30 – 500 mM imidazole. Fractions containing APT1 or APT2 were pooled, concentrated to 3 mL, and desalted using an Econo-Pac 10DG desalting column equilibrated in the final protein storage buffer (50 mM phosphate buffer pH 8.0, 100 mM NaCl, 10% glycerol). The protein was aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C for future use.

APT1 and APT2 enzymatic assays

This protocol was adapted from a previous paper[18]. Reaction buffer (1X PBS, 0.01% pluronic) was prepared and adjusted to pH 6.5 with HCl. Compounds were dissolved in DMSO to 10 different concentrations, with each being 48.4x of the final concentration

to be used in the enzymatic assay. Purified APT1 or APT2 was diluted in reaction buffer (450 μL) to give 10.8 nM solution. The inhibitor solution (10 μL) was added and incubated at room temperature for 15 min. During the incubation period, a resorufin acetate solution (5 μL, 1 mM in DMSO) was added to wells of a black, flat bottom 96 well plate. Then the enzyme solution (95 μL) was added to each well and the fluorescent intensity of each well was recorded (ex. = 571 nm, em. = 584 nm) at 70 sec. intervals for 30 min. Three biological replicates were performed for each concentration. Fluorescent intensities were plotted against time to calculate v_0 which was then plotted against concentration using GraphPad Prism to calculate IC_{50} values for each compound.

Gel Filtration Assay to identify covalent inhibitors

This assay was adapted from a previous paper[18]. A cell lysate was made from 293T cells transfected to express APT2. This solution was diluted in 1x PBS to give 2 mL of a 1 mg/mL solution. A 50x solution of the inhibitor in DMSO was made, added to the protein solution, and incubated at 37 ºC for 30 min. An aliquot (60 μL) was taken, and the remaining sample was loaded onto a Sephadex G-25m column and eluted with PBS (2.5 mL). This process was repeated two more times. The protein concentration of each aliquot was calculated by a BCA assay and adjusted to equalize the protein concentration. 25 μL of each aliquot was taken, combined with an ActivX TAMRA-FP solution (0.5 μL, 250 μM in DMSO) and incubated at room temperature for 30 min. Samples were quenched with 6x loading dye then run on a 12% polyacrylamide gel. Gels was imaged with a Typhoon 7000 variable mode imager to determine the

fluorescent intensity of the APT2 bands (~ 25 kDa).

Competitive in vitro ABPP assay

A cell lysate (0.5 mg/ml) was spiked with purified APT1 or APT2 and 25 μL aliquots were taken. The indicated concentrations of the inhibitor (1 μL of a 25x stock) were added to each aliquot and incubated at room temperature for 15 min. ActivX TAMRA-FP (0.5 μL of a 50x stock) was added and the samples were incubated at room temperature for 30 min. The reaction was quenched with 6x loading dye and the resulting samples were run on a 12% polyacrylamide gel. The fluorescent intensities of the APT1 or APT2 bands were then examined (550 nm ex., 570 nm em. filter) with a Typhoon 7000 variable mode imager.

Detection of STAT3 palmitoylation in cells

HEK 293T cells were transfected with STAT3-Flag and ZDHHC7-HA plasmids for 36 h, and were treated with 25 μM ML349, KW5116, KW5129 for 16 h and 50 μM Alk14 (hexadec-15-ynoic acid) for 6 h before collecting. Cells were collected and lysed in the 1% NP-40 lysis buffer (25 mM Tris-HCl pH 8.0, 10% glycerol, 150 mM NaCl, 1% Nonidet P-40) with protease inhibitor cocktail (P8340, Sigma). STAT3-Flag was purified by anti-Flag agarose beads. Click chemistry was done by adding a mixture of 50 μL IP wash buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.2% Nonidet P-40), 3 μL of 2 mM TAMRA-azide (#47130, Lumiprobe), 3.6 μL of 10 mM tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA), 3 μL of 40 mM $CuSO_4$, 3 μL of 40 mM tris(2-carboxyethyl)phosphine HCl (TCEP hydrochloride) to each sample. The mixtures were

mixed by vortex and incubated in the dark at room temperature for 30 min. To each reaction mixture, 20 μL of 6x loading buffer was added and the mixture was boiled at 95 °C for 10 min to denature. The mixture was then separated by SDS-PAGE gel. The gel was scanned by the Typhoon 7000 Variable Mode Imager to visualize the fatty acylation of STAT3 and was stained with Coomassie Brilliant Blue (B7920, Sigma) to confirm protein loading. Relative fluorescent/protein quantification was done by ImageJ.

Mouse studies

Experiments using live mice were conducted with approval of the Cornell Institutional Animal Care and Use Committee (IACUC) and complied with the laws, policies, and animal care and use procedures that govern the use of live vertebrates in research. 10-week-old C57BL/6J mice (5 males or 5 females per group) were treated with 2.5% DSS in their drinking water for 8 days. On the same day as DSS treatment was started, the mice were injected interperitoneally with a 50 mg/kg does of compound. Injections continued every other day for 11 days. On day 12 the mice were sacrificed, and the colons were removed and measured. Body weights and the health of the mice were tracked throughout the experiment.

Synthetic Methods

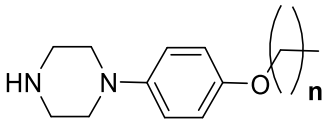
Procedure 1, Amide couplings. The corresponding carboxylic acid (1 eq.) and amine (1 eq.) starting materials were dissolved in DCM. 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC, 3 eq.) was added to the solution followed by 4-

dimethylaminopyridine (DMAP, 0.11 eq.). The solution was stirred at room temperature for 16 h then washed with HCl (0.1 M), water, and brine. The organic layer was then dried over anhydrous sodium sulfate and concentrated under reduced pressure. The resulting crude product was purified by flash chromatography to give the final product.

Procedure 2, Suzuki couplings. An aryl halide (1 eq.), 2-bromophenyl boronic acid (1 eq.), and Pd $(PPh_3)_4$ (0.04 eq) were combined in a in a screw cap vial which was degassed and charged with N_2. 1,4-Dioxane was added to the vial. In a separate vial, sodium carbonate (3 eq.) was dissolved in water (0.25 volume eq. of the dioxane). The water solution was degassed, charged with N_2, and added to the first vial. The mixture was then stirred at 100 ^oC. When the aryl halide starting material was completely consumed as determined by liquid chromatography-mass spectrometry (LC-MS), the reaction mixture was cooled to room temperature. Dioxane was removed under reduced pressure. The resulting crude was diluted with water and extracted three times with an equal volume of DCM. The organic layers were pooled and dried over sodium sulfate. Solvent was removed under reduced pressure to give the second crude material which was purified by flash chromatography to give the final product.

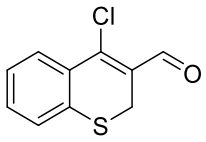
Procedure 3, Borylations. An aryl halide (1 eq.), bis boronic acid (3 eq.), XPhos-Pd-G2 (0.01 eq.), XPhos (0.02 eq.) and KOAc (3 eq.) were combined in a screw cap vial, degassed, and charged with N_2. Ethanol was added to the vial. The mixture was heated to 80 ^oC. LC-MS was used to monitor the reaction progress. When the aryl halide had been completely consumed, the reaction mixture was cooled to room temperature and

the solvent was removed under reduced pressure. The crude product was taken up in water then extracted three times with an equal volume of DCM. The organic layers were collected, pooled, and dried over sodium sulfate. The solvent was removed under reduced pressure to yield the second crude which was purified by flash chromatography to give the final product.



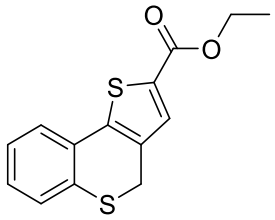
S1 n = 2
S2 n = 5

1-(4-Propoxyphenyl)piperazine (S1) & 1-(4-(hexyloxy)phenyl)piperazine (S2). Tert-butyl 4-(4-hydroxyphenyl) piperazine-1-carboxylate (1 eq.) and the corresponding bromoalkane (3 eq.) were dissolved in acetone. Potassium carbonate (1.5 eq.) was then added, and the mixture was heated to reflux and stirred for 16 h. Solvent was removed under reduced pressure and the residue was taken up into DCM and washed with 1 M aqueous NaOH, water, and brine. The organic layer was dried over sodium sulfate and concentrated under reduced pressure. The resulting solid was dissolved in DCM to which was added trifluoracetic acid (TFA, 0.4 eq. by volume of DCM). This solution was stirred at room temperature for 2 h. Solvent was removed under reduced pressure and the crude product was dissolved in DCM, washed with 1 M sodium carbonate, water, and brine. The organic layer was dried over sodium sulfate and concentrated under reduced pressure. The crude product was purified by flash chromatography to give compounds **S1** and **S2** (58% and 50% yield, respectively).



S3

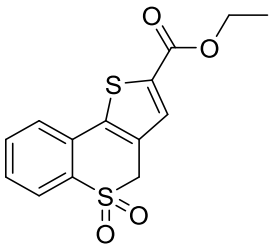
4-Chloro-2H-thiochromene-3-carbaldehyde (S3). $POCl_3$ (0.57 mL, 6.1 mmol) was added dropwise to DMF (6 mL). The solution was chilled to 0 ^oC and thiochroman-4-one (1 g, 6.1 mmol) was added drop wise. The solution was warmed to room temperature and stirred for 16 h. The solution was poured into ice cold water (100 mL) to quench the reaction. This solution was extracted 3 times with DCM (100 mL each). The organic layers were pooled and dried over sodium sulfate. Solvent was removed under reduced pressure and the resulting product was taken to the next step without further purification (1.1 g, 82% yield).



S4

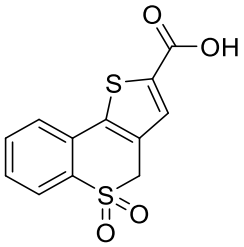
Ethyl 4H-thieno[3,2-c]thiochromene-2-carboxylate (S4). Sodium metal (0.16 g, 6.9 mmol) was slowly added to EtOH (20 mL) and dissolved completely before chilling the solution to 0 ^oC. Ethyl thioglycolate (0.28 mL, 6.6 mmol) was added dropwise followed by compound **S3** (1.1g, 5.0 mmol). The solution was warmed to room temperature while stirring for 16 h. The solution was then heated to 70 ^oC for 2 h. The solvent was removed under reduced pressure and the resulting crude was dissolved in EtOAc (50 mL), washed

with water and brine (50 mL each), and dried over sodium sulfate. The second crude product was concentrated under reduced pressure and purified by flash chromatography (6:1 hexane:EtOAc) to obtain compound **S4** (0.85 g, 60% yield).



S5

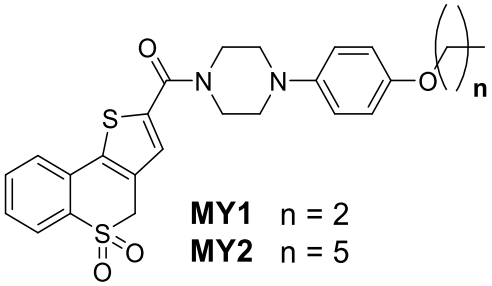
Ethyl 4H-thieno[3,2-c]thiochromene-2-carboxylate 5,5-dioxide (S5). Compound **S4** (2.4 g, 8.8 mmol) was dissolved in acetic acid (24 mL) and heated to 60 ^{o}C. $NaBO_3 \cdot 4H_2O$ (6.8 g, 44 mmol) was added slowly. The solution was stirred at 60 ^{o}C for 5 h. The solution was cooled to room temperature, and solvent was removed under reduced pressure. The residue was washed with water and dried to obtain compound **S5** which was taken to the next step without further purification (1.8 g, 65% yield)



S6

4H-Thieno[3,2-c]thiochromene-2-carboxylic acid 5,5-dioxide (S6). To a solution of compound **S5** (1.2 g, 3.9 mmol) in THF (30 mL) was added water (10 mL) and NaOH (1.2 g, 29 mmol). The resulting solution was heated to 80 ^{o}C and stirred for 3 h. THF

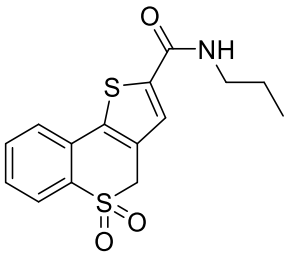
was removed under reduced pressure and the aqueous solution was acidified with 6 M HCl (7 mL). The precipitated solid was collected by vacuum filtration and air dried to obtain compound **S6** (1.1 g, 99% yield).



MY1 n = 2
MY2 n = 5

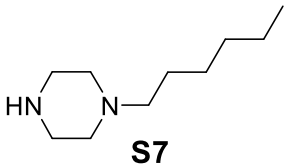
(5,5-Dioxido-4H-thieno[3,2-c]thiochromen-2-yl)(4-(4-propoxyphenyl)piperazin-1-yl)methanone (MY1) & (5,5-dioxido-4H-thieno[3,2-c]thiochromen-2-yl)(4-(4-(hexyloxy)phenyl)piperazin-1-yl)methanone (MY2). **MY1** and **MY2** were synthesized from **S6** and **S1** and **S2** respectively by procedure 1 (52% and 86% yield respectively). **(MY1)** 1H NMR (500 MHz, CDCl3) δ 8.05 (dd, J = 7.8, 1.3 Hz, 1H), 7.71 – 7.60 (m, 2H), 7.54 (td, J = 7.5, 1.3 Hz, 1H), 7.21 (s, 1H), 6.94 – 6.88 (m, 2H), 6.88 – 6.84 (m, 2H), 4.44 (s, 2H), 3.95 – 3.91 (m, 4H), 3.88 (t, J = 6.6 Hz, 2H), 3.18 – 3.04 (m, 4H), 1.87 – 1.71 (m, 2H), 1.03 (t, J = 7.4 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.16, 154.17, 144.89, 137.41, 136.23, 133.98, 133.62, 130.10, 129.82, 129.27, 127.85, 125.87, 124.29, 119.06, 115.30, 69.93, 51.27, 51.22, 22.67, 10.55. HRMS (ESI) calc. for $[M+H]^+$ $C_{25}H_{27}N_2O_4S_2$ 483.14068, obs. 483.41061. **(MY2)** ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.06 (dd, *J* = 7.8, 1.2 Hz, 1H), 7.66 (ddd, *J* = 18.3, 7.7, 1.3 Hz, 2H), 7.55 (td, *J* = 7.5, 1.4 Hz, 1H), 7.24 (s, 1H), 6.96 – 6.90 (m, 2H), 6.90 – 6.85 (m, 2H), 4.46 (s, 2H), 3.94 (td, *J* = 6.7, 5.9, 2.7 Hz, 6H), 3.19 – 3.09 (m, 4H), 1.84 – 1.71 (m, 2H), 1.47 (ddd, *J* = 8.2, 5.3, 2.7 Hz, 2H), 1.35 (ddd, *J* = 7.2, 4.6, 3.3 Hz, 4H), 0.96 –

0.87 (m, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.1, 154.2, 144.9, 137.4, 136.2, 134.0, 133.6, 130.1, 130.0, 129.3, 127.9, 125.9, 124.3, 119.0, 115.3, 68.4, 51.2, 51.2, 31.6, 29.3, 25.8, 22.6, 14.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{28}H_{33}N_2O_4S_2$ 525.18763, obs. 525.18730.



MY3

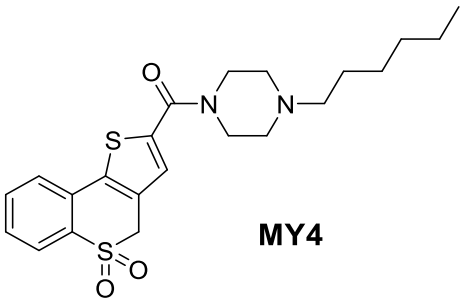
N-Propyl-4H-thieno[3,2-c]thiochromene-2-carboxamide 5,5-dioxide (MY3). **MY3** was synthesized from propylamine and **S6** by procedure 1 (70% yield). ^{1}H NMR (500 MHz, DMSO) δ 8.72 (t, *J* = 5.7 Hz, 1H), 7.96 (dd, *J* = 7.6, 1.1 Hz, 1H), 7.83 – 7.78 (m, 2H), 7.77 (s, 1H), 7.65 (ddd, *J* = 8.5, 6.8, 2.0 Hz, 1H), 4.94 (s, 2H), 3.27 – 3.19 (m, 2H), 1.54 (q, *J* = 7.2 Hz, 2H), 0.90 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (126 MHz, DMSO) δ 160.7, 140.6, 135.9, 134.7, 134.2, 130.8, 130.3, 129.9, 129.7, 126.5, 123.8, 50.9, 41.5, 22.8, 11.9. HRMS (ESI) calc. for $[M+H]^+$ $C_{15}H_{26}NO_3S_2$ 322.05661, obs. 322.05640.



S7

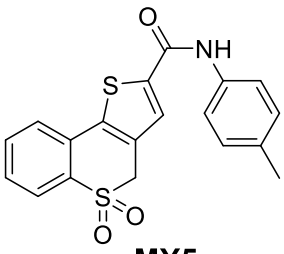
1-hexylpiperazine (S7). A solution of piperazine (1.2 g, 14 mmol) and 1-bromohexane (0.50 mL, 3.6 mmol) in EtOH (30 mL) was heated to reflux for 16 h. Solvent was

removed under reduced pressure. The residue was re-dissolved in DCM (50 mL) and washed with saturated aqueous sodium bicarbonate and brine (50 mL each). The organic layer was dried over sodium sulfate and concentrated under reduced pressure. Crude material was purified by flash chromatography to give compound **S7** (0.51 g, 85% yield).



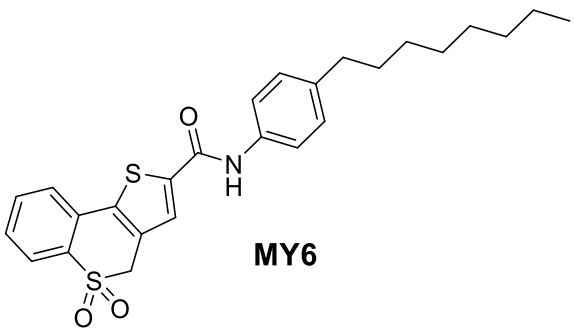
MY4

(5,5-dioxido-4H-thieno[3,2-c]thiochromen-2-yl)(4-hexylpiperazin-1-yl)methanone (MY4). **MY4** was synthesized from **S6** and **S7** by procedure 1 (82% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.04 (dd, *J* = 7.9, 1.3 Hz, 1H), 7.66 (dd, *J* = 7.4, 1.3 Hz, 1H), 7.62 (dd, *J* = 7.9, 1.3 Hz, 1H), 7.54 (td, *J* = 7.5, 1.3 Hz, 1H), 7.18 (s, 1H), 4.44 (s, 2H), 3.79 (t, *J* = 5.0 Hz, 4H), 2.50 (t, *J* = 4.9 Hz, 4H), 2.44 – 2.31 (m, 2H), 1.51 (t, *J* = 7.5 Hz, 2H), 1.40 – 1.21 (m, 6H), 0.96 – 0.83 (m, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.0, 137.6, 136.0, 133.9, 133.6, 130.1, 129.7, 129.2, 127.8, 125.8, 124.2, 58.6, 53.2, 51.2, 31.8, 27.2, 26.7, 22.6, 14.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{22}H_{29}N_2O_3S_2$ 433.16141, obs. 433.16133



MY5

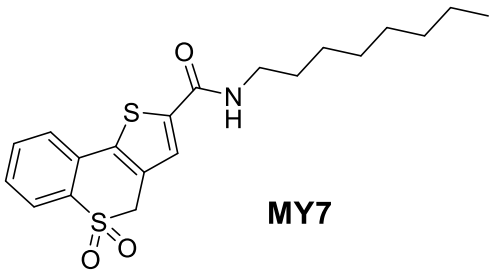
N-(p-tolyl)-4H-thieno[3,2-c]thiochromene-2-carboxamide 5,5-dioxide (5). **MY5** was synthesized from p-toluidine hydrochloride and **S6** by procedure 1 (91% yield). ^{1}H NMR (500 MHz, DMSO) δ 10.41 (s, 1H), 8.03 (s, 1H), 7.98 (dd, *J* = 7.9, 1.2 Hz, 1H), 7.86 (dd, *J* = 7.9, 1.2 Hz, 1H), 7.81 (td, *J* = 7.6, 1.3 Hz, 1H), 7.67 (td, *J* = 7.6, 1.3 Hz, 1H), 7.64 – 7.60 (m, 2H), 7.18 (d, *J* = 8.2 Hz, 2H), 5.00 (s, 2H), 2.29 (s, 3H). ^{13}C NMR (126 MHz, DMSO) δ 159.4, 140.5, 136.9, 136.4, 134.7, 134.3, 133.6, 130.9, 130.7, 130.1, 129.6, 126.6, 123.8, 120.8, 50.9, 21.0. HRMS (ESI) calc. for $[M+H]^+$ $C_{19}H_{15}NO_3S_2$ 370.05661, obs. 370.05646



MY6

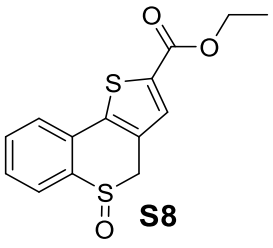
N-(4-octylphenyl)-4H-thieno[3,2-c]thiochromene-2-carboxamide 5,5-dioxide (MY6). **MY6** was synthesized from 4-octylaniline and compound **S6** by procedure 1 (85% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.23 (s, 1H), 7.97 (dd, *J* = 7.7, 1.2 Hz, 1H), 7.53 – 7.45 (m, 2H), 7.45 – 7.37 (m, 4H), 7.08 – 7.02 (m, 2H), 4.34 (s, 2H), 2.54 (dd, *J* = 8.8, 6.7 Hz, 2H), 1.56 (d, *J* = 7.2 Hz, 2H), 1.36 – 1.20 (m, 10H), 0.89 (t, *J* = 6.9 Hz,

3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 158.8, 140.2, 139.5, 137.9, 134.9, 134.1, 133.0, 130.0, 129.3, 128.8, 128.6, 128.2, 126.1, 124.1, 120.1, 51.2, 35.4, 31.9, 31.6, 29.5, 29.4, 29.3, 22.7, 14.2. HRMS (ESI) calc. for $[M+H]^+$ $C_{26}H_{30}NO_3S_2$ 468.16616, obs. 468.16606.



MY7

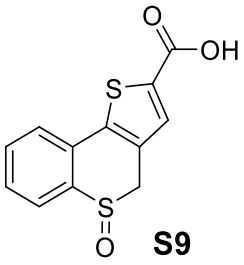
N-octyl-4H-thieno[3,2-c]thiochromene-2-carboxamide 5,5-dioxide (MY7). **MY7** was synthesized from octylamine and compound **S6** by procedure 1 (77% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.04 (dd, *J* = 7.8, 1.3 Hz, 1H), 7.70 – 7.61 (m, 2H), 7.54 (td, *J* = 7.5, 1.4 Hz, 1H), 7.34 (s, 1H), 6.09 (dt, *J* = 11.5, 6.0 Hz, 1H), 4.42 (s, 2H), 3.40 (td, *J* = 7.2, 5.8 Hz, 2H), 1.64 – 1.51 (m, 3H), 1.40 – 1.19 (m, 12H), 0.88 (t, *J* = 6.8 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 160.7, 139.6, 137.1, 134.0, 133.7, 130.3, 129.3, 128.4, 128.3, 125.9, 124.3, 51.3, 40.3, 31.8, 29.6, 29.3, 29.2, 27.0, 22.7, 14.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{20}H_{26}NO_3S_2$ 392.13486, obs. 392.13462.



S8

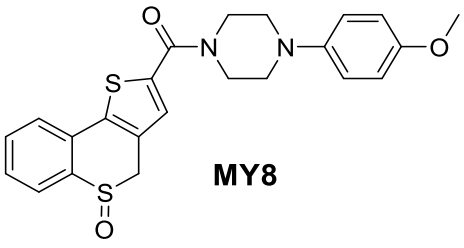
Ethyl 4H-thieno[3,2-c]thiochromene-2-carboxylate 5-oxide (S8). To a solution of

compound **S4** (1.0 g, 3.6 mmol) in DCM (40 mL) was added meta-chloroperoxybenzoic acid (m-CPBA, 0.81 g, 3.6 mmol). The reaction mixture was stirred at room temperature for 1 h then washed twice with saturated aqueous sodium bicarbonate and brine (15 mL each). The organic layer was dried over sodium sulfate and concentrated under reduced pressure. Crude product was purified by flash chromatography (1:1 → 5:6 hexane:EtOAc) to give compound **S8** (78% yield).



S9

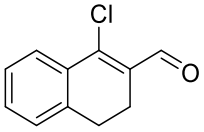
4H-Thieno[3,2-c]thiochromene-2-carboxylic acid 5-oxide (S9). Compound **S8** (0.20 g, 0.68 mmol) was dissolved in MeOH (7 mL) and 1 M aqueous NaOH (3 mL) was added. The mixture was stirred at room temperature for 2 h then heated to 60 ^{o}C for 1 h. MeOH was removed under reduced pressure and the crude material was diluted with water (5 mL). The solution was acidified with 6 M HCl (1 mL) and the resulting precipitate was collected by vacuum filtration to give compound **S9** (0.18 g, 99% yield).



MY8

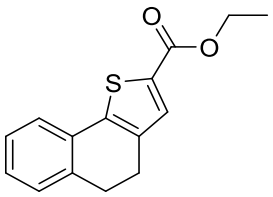
(4-(4-Methoxyphenyl)piperazin-1-yl)(5-oxido-4H-thieno[3,2-c]thiochromen-2-

yl)methanone (MY8). **MY8** was synthesized from 1-(4-methoxypheneyl)piperazine and compound **S9** by procedure 1 (43% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.88 (dd, *J* = 7.6, 1.3 Hz, 1H), 7.67 – 7.59 (m, 2H), 7.52 (td, *J* = 7.4, 1.6 Hz, 1H), 7.28 (s, 1H), 6.97 – 6.92 (m, 2H), 6.91 – 6.86 (m, 2H), 4.49 (d, *J* = 15.3 Hz, 1H), 4.00 – 3.92 (m, 5H), 3.80 (s, 3H), 3.14 (t, *J* = 5.1 Hz, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.5, 154.6, 145.1, 137.9, 136.9, 135.3, 133.1, 130.8, 129.4, 129.3, 128.1, 125.6, 125.5, 119.0, 114.6, 55.6, 51.2, 47.4. HRMS (ESI) calc. for $[M+H]^+$ $C_{23}H_{23}N_2O_3S_2$ 439.11446, obs. 439.11415.



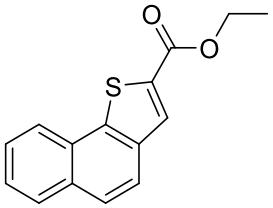
S10

1-chloro-3,4-dihydro-2-naphthaldehyde (S10). $POCl_3$ (0.67 mL, 7.3 mmol) was added dropwise to DMF (6 mL). The solution was chilled to 0 °C and alpha tetralone (0.91 mL, 6.8 mmol) was added drop wise. The solution was warmed to room temperature then heated to 80 °C for 1.5 h. Heating was removed and the solution was cooled to room temperature then poured slowly into 1 M aqueous sodium acetate (50 mL) to quench the reaction. This solution was extracted 3 times with DCM (50 mL each). The organic layers were pooled and dried over sodium sulfate. The solvent was removed under reduced pressure and the resulting product was taken to the next step without further purification.



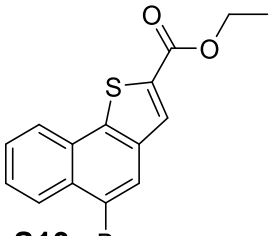
S11

Ethyl 4,5-dihydronaphtho[1,2-b]thiophene-2-carboxylate (S11). Sodium metal (0.20 g, 8.7 mmol) was slowly added to ethanol (20 mL) and dissolved completely before chilling the solution to 0 °C. Ethyl thioglycolate (0.83 mL, 7.5 mmol) was added dropwise followed by compound **S10** from the previous reaction. The solution was warmed to room temperature while stirring for 16 h. The solution was then heated to 70 °C for 2 h. Solvent was removed under reduced pressure after cooling and the resulting crude was dissolved in DCM (15 mL), washed with water and brine (15 mL each), and dried over sodium sulfate. The second crude product was concentrated under reduced pressure and purified by flash chromatography (hexane → 3:2 hexane:DCM) to give compound **S11** (1.0 g, 3.9 mmol, 57% yield over two steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.62 (s, 1H), 7.45 (dt, *J* = 6.1, 1.7 Hz, 1H), 7.29 – 7.22 (m, 3H), 4.39 (q, *J* = 7.1 Hz, 2H), 2.99 (dd, *J* = 8.8, 6.5 Hz, 2H), 2.86 (dd, *J* = 8.6, 6.3 Hz, 2H), 1.42 (t, *J* = 7.2 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.5, 143.1, 137.7, 135.4, 133.6, 130.7, 130.5, 128.3, 128.3, 127.2, 123.6, 61.1, 28.9, 23.8, 14.4.



S12

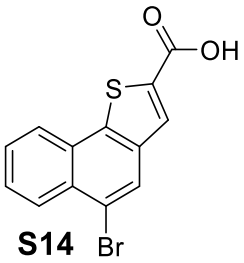
Ethyl naphtho[1,2-b]thiophene-2-carboxylate (S12). Compound **S11** (1.0 g, 3.9 mmol), N-bromosuccinimide (NBS, 0.69 g, 3.9 mmol) and benzoyl peroxide (BPO, 1.0 mg, 0.041 mmol) were dissolved in CCl_4 (30 mL) and heated to reflux for 2 h. Additional NBS and BPO (0.69 g, 3.9 mmol and 1.0 mg, 0.041 mmol respectively) were added and the mixture refluxed again for 2 h. The mixture was cooled, and vacuum filtered to remove suspended solids. The filtrate was concentrated under reduced pressure and the crude material was purified by flash chromatography (hexane → 3:17 EtOAc:hexane) to give compound **S12** (0.88 g, 3.4 mmol, 89% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.21 – 8.10 (m, 2H), 7.92 (dd, *J* = 6.1, 3.7 Hz, 1H), 7.80 (dd, *J* = 8.7, 2.3 Hz, 1H), 7.74 (dd, *J* = 8.8, 2.2 Hz, 1H), 7.60 (qd, *J* = 7.2, 4.9 Hz, 2H), 4.52 – 4.40 (m, 2H), 1.48 (t, *J* = 7.2 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.8, 141.2, 136.6, 132.6, 131.8, 131.3, 128.9, 128.7, 127.0, 127.0, 126.2, 123.9, 122.7, 61.6, 14.4.



S13

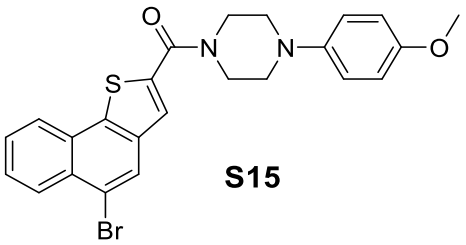
Ethyl 5-bromonaphtho[1,2-b]thiophene-2-carboxylate (S13). Compound **S12** (0.88 g, 3.4 mmol) was dissolved in the acetic acid (17 mL) and bromine was added (0.2 mL, 3.9 mmol). The solution was heated to reflux for 4 h. After cooling to room temperature, the liquid was poured into water (100 mL) and extracted three times with EtOAc (100 mL each). The organic fractions were pooled, dried over anhydrous sodium sulfate, and concentrated under reduced pressure. The crude product was purified by flash

chromatography (hexane → 2:3 DCM:hexane) to give compound **S13** (0.52 g, 1.6 mmol, 47% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.37 (dd, *J* = 8.3, 1.4 Hz, 1H), 8.21 – 8.12 (m, 2H), 8.09 (s, 1H), 7.75 – 7.65 (m, 2H), 4.47 (q, *J* = 7.1 Hz, 2H), 1.47 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.4, 140.8, 136.7, 133.5, 130.2, 130.0, 129.3, 128.6, 128.1, 128.0, 126.4, 124.3, 120.8, 61.8, 14.4.



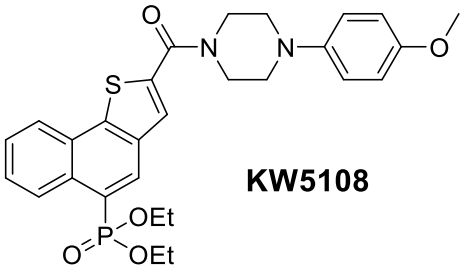
S14

5-Bromonaphtho[1,2-b]thiophene-2-carboxylic acid (S14). Compound **S13** (0.52 g, 1.6 mmol) was dissolved in tetrahydrofuran (THF, 7 mL) to which was added a 2 M aqueous NaOH (7 mL). The mixture was heated to reflux for 16 h. After cooling, THF was removed under reduced pressure. The remaining aqueous solution was acidified with 6 M HCl (5 mL). The precipitated compound **S15** was collected by vacuum filtration and taken to the next step without further purification (0.46 g, 1.5 mmol, 94% yield). ^{1}H NMR (500 MHz, DMSO) δ 13.63 (s, 1H), 8.45 – 8.40 (m, 1H), 8.30 – 8.23 (m, 2H), 8.23 – 8.19 (m, 1H), 7.83 – 7.75 (m, 2H). ^{13}C NMR (126 MHz, DMSO) δ 163.5, 140.3, 137.4, 135.2, 131.0, 129.5, 129.2, 129.1, 129.0, 128.3, 127.3, 125.0, 120.2.



S15

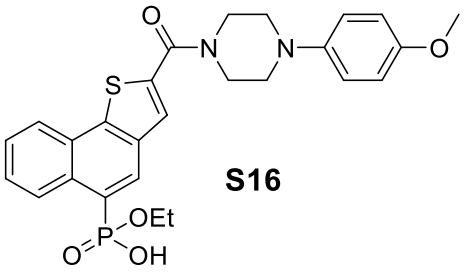
(5-Bromonaphtho[1,2-b]thiophen-2-yl)(4-(4-methoxyphenyl)piperazin-1-yl)methanone (S15). Compound **S15** was synthesized from compound **S14** and 1-(4-methoxypheneyl)piperazine by procedure 1 (47% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.43 – 8.33 (m, 1H), 8.16 – 8.11 (m, 2H), 7.73 – 7.65 (m, 2H), 7.60 (s, 1H), 6.99 – 6.94 (m, 2H), 6.91 – 6.87 (m, 2H), 4.07 – 3.95 (m, 4H), 3.81 (s, 3H), 3.18 (t, *J* = 5.1 Hz, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 163.3, 154.6, 145.1, 138.4, 136.5, 136.5, 129.6, 129.6, 128.6, 127.9, 127.7, 126.0, 125.6, 124.2, 120.7, 119.1, 114.6, 55.6, 53.5, 51.3.



KW5108

Diethyl (2-(4-(4-methoxyphenyl)piperazine-1-carbonyl)naphtho[1,2-b]thiophen-5-yl)phosphonate (KW5108). Compound **S15** (0.51 g, 1.1 mmol), diethyl phosphonate (0.20 mL, 1.5 mmol), N,N-diisopropylethylamine (DIPEA, 0.36 mL, 2 mmol), $Pd(OAc)_2$ (11 mg, 0.049 mmol) and triphenyl phosphine (33 mg, 0.13 mmol) were combined in ethanol (44 mL). The mixture was heated to reflux for 16 h. The reaction was cooled, and the solvent was removed under reduced pressure. The crude material was purified by flash chromatography (1:1 EtOAC:hexane, then 3:97 methanol:DCM)

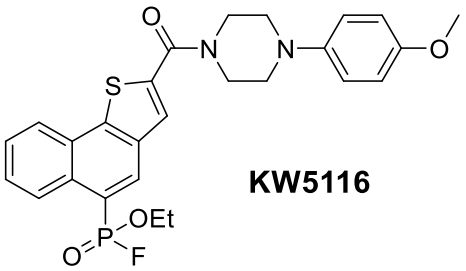
to give **KW5108** (0.534 g, 0.99 mmol, 93% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.66 (d, *J* = 17.1 Hz, 1H), 8.62 – 8.58 (m, 1H), 8.21 (dt, *J* = 7.1, 2.4 Hz, 1H), 7.71 (s, 1H), 7.70 – 7.66 (m, 2H), 6.99 – 6.93 (m, 2H), 6.92 – 6.86 (m, 2H), 4.26 (dp, *J* = 10.2, 7.2 Hz, 2H), 4.12 (ddq, *J* = 10.1, 8.3, 7.1 Hz, 2H), 4.04 – 3.97 (m, 4H), 3.80 (d, *J* = 0.9 Hz, 3H), 3.18 (t, *J* = 5.0 Hz, 4H), 1.34 (t, *J* = 7.0 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 163.2, 154.6, 145.1, 143.5, 143.5, 136.7, 134.9, 134.8, 132.0, 131.9, 129.8, 129.7, 129.0, 128.9, 128.0, 128.0, 127.6, 127.5, 126.8, 126.8, 124.4, 124.4, 123.6, 122.2, 119.1, 114.6, 62.4, 62.3, 55.6, 51.3, 16.4, 16.4. ^{31}P NMR (202 MHz, $CDCl_3$) δ 19.2. HRMS (ESI) calc. for $[M+H]^+$ $C_{28}H_{32}N_2O_5PS$ 539.17641, obs. 539.17735.



S16

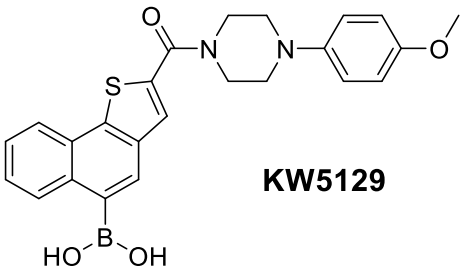
Ethyl hydrogen (2-(4-(4-methoxyphenyl)piperazine-1-carbonyl)naphtho[1,2-b]thiophen-5-yl)phosphonate (S16). **KW5108** (0.50 g, 0.92 mmol) and LiBr (0.951 g, 11 mmol) were combined in acetone (9 mL) and heated to reflux for 3 days. Solvent was removed under reduced pressure. The crude material was purified by flash chromatography (DCM → 1:9 MeOH:DCM) to give compound **S16** (0.168 g, 0.33 mmol, 36% yield). ^{1}H NMR (500 MHz, DMSO) δ 8.64 (dt, *J* = 6.4, 3.5 Hz, 1H), 8.57 (d, *J* = 16.5 Hz, 1H), 8.27 (dt, *J* = 6.0, 2.5 Hz, 1H), 8.12 (s, 1H), 7.73 (dt, *J* = 6.9, 3.5 Hz, 2H), 7.00 – 6.93 (m, 2H), 6.89 – 6.83 (m, 2H), 3.99 – 3.82 (m, 6H), 3.70 (s, 3H),

3.14 (t, *J* = 5.2 Hz, 4H), 1.17 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (126 MHz, DMSO) δ 162.4, 153.9, 145.5, 141.8, 137.3, 135.7, 135.6, 130.8, 130.1, 128.6, 128.5, 128.2, 128.1, 124.7, 118.6, 114.8, 61.3, 55.7, 50.6, 16.8. ^{31}P NMR (202 MHz, DMSO) δ 6.3.



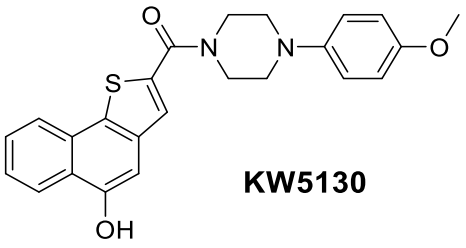
KW5116

Ethyl (2-(4-(4-methoxyphenyl)piperazine-1-carbonyl)naphtho[1,2-b]thiophen-5-yl)phosphonofluoridate (KW5116). Compound **S16** (42 mg, 0.082 mmol) was dissolved in DCM (25 mL) to which (diethylamino)sulfur trifluoride (22 μL, 0.17 mmol) was added drop wise. The solution was stirred at room temperature for 3 h. Solvent was removed under reduced pressure. The resulting material was purified by flash chromatography (1:49 methanol:DCM) to give **KW5116** (16 mg, 0.031 mmol, 38% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.61 (d, *J* = 18.5 Hz, 1H), 8.45 (dq, *J* = 7.9, 2.8 Hz, 1H), 8.22 (dt, *J* = 7.2, 2.4 Hz, 1H), 7.74 – 7.68 (m, 3H), 6.97 – 6.91 (m, 2H), 6.90 – 6.85 (m, 2H), 4.50 – 4.36 (m, 2H), 3.98 (t, *J* = 5.1 Hz, 4H), 3.78 (s, 3H), 3.17 (t, *J* = 5.0 Hz, 4H), 1.45 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.9, 154.6, 145.1, 144.4, 144.4, 137.2, 134.6, 134.5, 131.9, 131.8, 131.8, 129.2, 129.1, 128.9, 128.8, 128.2, 128.0, 127.3, 127.3, 126.6, 124.6, 124.6, 119.1, 114.6, 77.3, 77.0, 76.8, 76.8, 64.4, 64.3, 55.6, 51.3, 29.7, 16.4, 16.4. ^{19}F NMR (470 MHz, $CDCl_3$) δ -60.3, -62.5. ^{31}P NMR (202 MHz, $CDCl_3$) δ 20.2, 15.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{26}H_{27}FN_2O_4PS$ 513.14077, obs. 513.14041.



KW5129

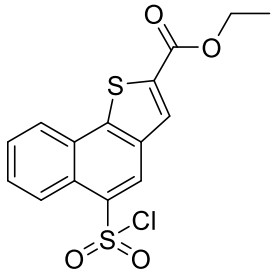
(2-(4-(4-Methoxyphenyl)piperazine-1-carbonyl)naphtho[1,2-b]thiophen-5-yl)boronic acid (KW5129). **KW5129** was synthesized from compound **S15** by procedure 3 (34% yield). **KW5129** was converted to the corresponding aryl trifluoroborate salt to obtain NMR spectra due the compound having poor solubility in CD_3OD and forming a mixture of dehydrated adducts in other solvents. ^{1}H NMR (500 MHz, DMSO) δ 8.53 (d, *J* = 8.0 Hz, 1H), 8.03 – 7.99 (m, 1H), 7.95 (d, *J* = 3.7 Hz, 1H), 7.83 (s, 1H), 7.51 – 7.41 (m, 1H), 6.97 (d, *J* = 6.6 Hz, 2H), 6.86 (d, *J* = 8.6 Hz, 2H), 3.89 (s, 4H), 3.70 (s, 3H), 3.14 (t, *J* = 5.0 Hz, 4H). ^{13}C NMR (126 MHz, DMSO) δ 163.3, 137.2, 136.2, 135.7, 134.8, 131.9, 128.3, 128.0, 125.8, 125.4, 125.4, 125.2, 123.5, 118.6, 114.8, 55.7, 50.7, 46.2. ^{19}F NMR (470 MHz, DMSO) δ -135.6. ^{11}B NMR (160 MHz, DMSO) δ 3.3. HRMS data was obtained from **KW5129** directly. HRMS (ESI) calc. for $[M+H]^+$ $C_{24}H_{24}N_2O_4BS$ 447.15499, obs. 447.15580



KW5130

(5-Hydroxynaphtho[1,2-b]thiophen-2-yl)(4-(4-methoxyphenyl)piperazin-1-

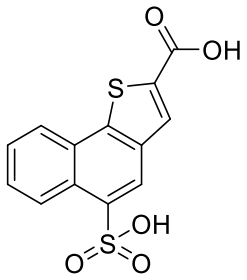
yl)methanone (10A). **KW5130** was isolated as a side product of the synthesis of **KW5129** (9% yield). ^{1}H NMR (500 MHz, DMSO) δ 10.34 (s, 1H), 8.27 (dd, *J* = 8.2, 1.4 Hz, 1H), 8.07 (d, *J* = 8.0 Hz, 1H), 7.81 (s, 1H), 7.66 (ddd, *J* = 8.2, 7.0, 1.4 Hz, 1H), 7.60 (ddd, *J* = 8.3, 6.9, 1.3 Hz, 1H), 7.27 (s, 1H), 6.96 (d, *J* = 8.8 Hz, 2H), 6.91 – 6.80 (m, 2H), 3.87 (t, *J* = 5.0 Hz, 4H), 3.70 (s, 3H), 3.12 (t, *J* = 5.0 Hz, 4H). ^{13}C NMR (126 MHz, DMSO) δ 162.9, 153.9, 152.1, 145.5, 137.5, 136.4, 129.3, 128.9, 128.0, 127.1, 126.3, 125.0, 124.0, 124.0, 118.6, 114.8, 103.6, 55.7, 55.4, 50.7. HRMS (ESI) calc. for $[M+H]^+$ $C_{24}H_{23}N_2O_3S$ 419.14239, obs. 419.14238



S17

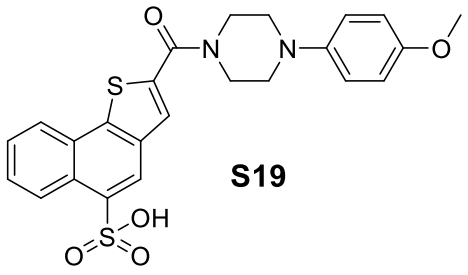
Ethyl 5-(chlorosulfonyl)naphtho[1,2-b]thiophene-2-carboxylate (S17). Chlorosulfonic acid (78 μL, 1.2 mmol) was added to chloroform (6 mL) and chilled to 0 °C. In a sperate vessel, compound **S12** (0.10 g, 0.38 mmol) was dissolved in chloroform (3 mL) and added dropwise to the solution on ice. The solution was warmed gradually to room temperature and stirred for 4 h. The mixture was poured into ice cold water (50 ml) and stirred for 30 min. The water was saturated with NaCl and the precipitate that formed was decanted into a vacuum filter apparatus to collect the solid. This solid was purified by flash chromatography (acetone → 1:9 MeOH:acetone) to give compound **S17** (30 mg, 0.83 mmol, 22% yield). ^{1}H NMR (500 MHz, DMSO) δ

9.01 – 8.96 (m, 1H), 8.45 (s, 1H), 8.39 (s, 1H), 8.24 – 8.19 (m, 1H), 7.68 – 7.65 (m, 2H), 4.39 (q, *J* = 7.1 Hz, 2H), 1.37 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (126 MHz, DMSO) δ 162.3, 143.3, 141.5, 135.6, 133.1, 132.6, 129.7, 128.7, 128.4, 127.6, 127.4, 124.1, 121.8, 62.0, 14.7.



S18

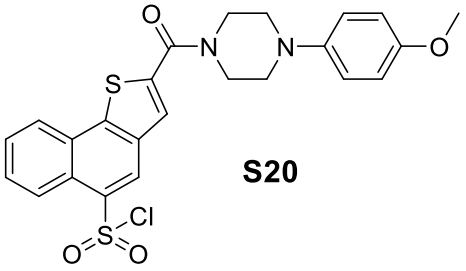
5-sulfonaphtho[1,2-b]thiophene-2-carboxylic acid (S18). Compound **S17** (0.16 g, 0.44 mmol) was added to a 2M aqueous NaOH solution (8 mL) and stirred and room temperature for 16 h. The solution was acidified with 6 M HCl (5 mL) and saturated with NaCl to precipitate compound **S18**. The solid was collected by vacuum filtration and taken to the next step without further purification (0.053 g, 0.17 mmol, 39% yield) ^{1}H NMR (500 MHz, DMSO) δ 9.00 – 8.95 (m, 1H), 8.44 (s, 1H), 8.31 (s, 1H), 8.22 – 8.16 (m, 1H), 7.70 – 7.62 (m, 2H). ^{13}C NMR (126 MHz, DMSO) δ 163.8, 143.0, 141.5, 135.7, 134.4, 132.6, 129.6, 128.7, 128.2, 127.6, 127.2, 124.1, 121.9.



S19

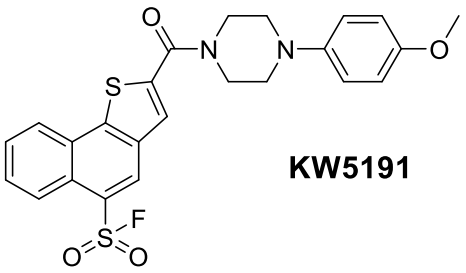
2-(4-(4-methoxyphenyl)piperazine-1-carbonyl)naphtho[1,2-b]thiophene-5-

sulfonic acid (S19). Compound **S19** was synthesized from compound **S18** and 1-(4-methoxypheneyl)piperazine by procedure 1 (25% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.88 (d, *J* = 8.5 Hz, 1H), 8.74 (d, *J* = 2.0 Hz, 1H), 8.29 (dd, *J* = 8.1, 1.5 Hz, 1H), 7.85 (ddd, *J* = 8.6, 7.1, 1.4 Hz, 1H), 7.80 (ddd, *J* = 8.1, 7.0, 1.2 Hz, 1H), 7.76 (d, *J* = 2.1 Hz, 1H), 6.96 – 6.91 (m, 2H), 6.90 – 6.85 (m, 2H), 3.99 (t, *J* = 5.1 Hz, 4H), 3.78 (s, 3H), 3.17 (t, *J* = 5.1 Hz, 4H). 13C NMR (126 MHz, CDCl3) δ 162.4, 154.7, 146.0, 145.0, 138.6, 138.0, 133.2, 129.4, 129.1, 128.8, 126.7, 126.6, 125.7, 124.8, 124.4, 119.1, 114.6, 55.6, 53.4, 50.9.



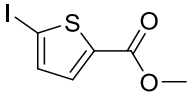
S20

2-(4-(4-Methoxyphenyl)piperazine-1-carbonyl)naphtho[1,2-b]thiophene-5-sulfonyl chloride (S20). Compound **S19** (0.24 g, 0.49 mmol) was dissolved in water (3 mL) and 2M aqueous NaOH (270 μL) was added. The solution was stirred at room temperature for 15 min then evaporated to dryness under reduced pressure. $POCl_3$ (5 mL) was added to the resulting solid and the mixture was stirred at room temperature for 16 h. The solution was concentrated under reduced pressure and the resulting product was taken to the next step without further purification.



KW5191

2-(4-(4-Methoxyphenyl)piperazine-1-carbonyl)naphtho[1,2-b]thiophene-5-sulfonyl fluoride (KW5191). Compound **S20** (30 mg, 0.06 mmol) was dissolved in a 1:1 solution of acetonitrile and water (0.6 mL). In a separated vessel, KHF_2 (52 mg, 0.67 mmol) was dissolved in water (0.6 mL) and added to the first solution drop wise. The resulting solution was stirred at room temperature. The progress of the reaction was monitored by LC-MS. When all the starting material had been consumed, the solution was diluted with water (10 mL) and extracted with EtOAc 3 times (10 mL each). The organic fractions were pooled, dried over sodium sulfate, and concentrated under reduced pressure to give **KW5191** (11 mg, 0.023 mmol, 38% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.74 (d, *J* = 1.6 Hz, 1H), 8.62 (dt, *J* = 8.1, 2.2 Hz, 1H), 8.28 (ddd, *J* = 8.5, 5.3, 1.7 Hz, 1H), 7.88 – 7.77 (m, 2H), 7.76 (d, *J* = 2.3 Hz, 1H), 6.97 – 6.91 (m, 2H), 6.90 – 6.85 (m, 2H), 4.02 – 3.96 (m, 4H), 3.79 (s, 3H), 3.17 (t, *J* = 5.0 Hz, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.4, 154.7, 146.0, 145.0, 138.5, 133.5, 133.2, 129.2, 128.8, 128.3, 126.5, 125.7, 124.8, 119.1, 114.6, 55.6, 51.3. ^{19}F NMR (470 MHz, $CDCl_3$) δ 62.8. HRMS (ESI) calc. for $[M+H]^+$ $C_{24}H_{22}N_2O_4S_2F$ 485.09995, obs. 485.09997.

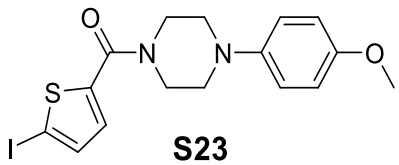


S21

Methyl 5-iodo-2-thiophenecarboxylate (S21). Methyl-2-thiophenecarboxylate (0.18 g, 1.3 mmol), [bis(trifluoroacetoxy)iodo]benzene (0.32 g, 0.73 mmol), and iodine (0.17 g, 0.67 mmol) were dissolved in CCl_4 (1.8 mL) and stirred at room temperature for 2 h. The solvent was removed under reduced pressure. The crude product was purified by flash chromatography (hexane → 1:3 DCM:hexane) to give compound **S21** (0.20 g, 0.76 mmol, 58% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.45 (d, *J* = 3.9 Hz, 1H), 7.28 (d, *J* = 3.9 Hz, 1H), 3.89 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 161.3, 139.2, 137.8, 134.5, 82.8, 52.3.

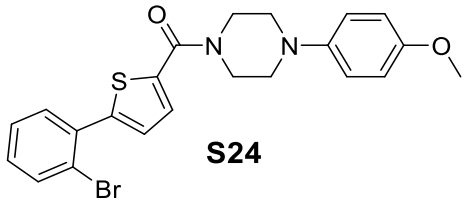
S22

5-Iodo-2-thiophenecarboxylic acid (S22). LiOH (0.152 g, 3.6 mmol) was dissolved in water (5.7 mL) to which was added a solution of compound **S21** (0.20 g, 0.76 mmol) in THF (0.76 mL). The mixture was stirred at room temperature for 4 h. The solution was acidified with 6 M HCl to pH 4 and then extracted with EtOAc three times (10 mL each). The organic fractions were pooled, dried over sodium sulfate, and concentrated under reduced pressure to give compound **S22**, which was taken to the next step without further purification (0.131 g, 0.52 mmol, 68% yield). ^{1}H NMR (500 MHz, DMSO) δ 13.29 (s, 1H), 7.42 (d, *J* = 3.8 Hz, 1H), 7.40 (d, *J* = 3.8 Hz, 1H). ^{13}C NMR (126 MHz, DMSO) δ 162.1, 140.6, 138.5, 135.1, 86.3.



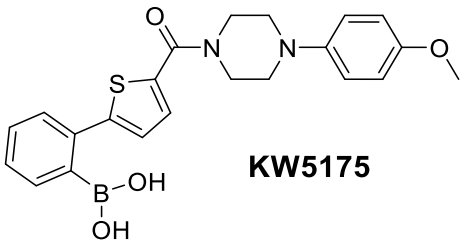
S23

(5-Iodothiophen-2-yl)(4-(4-methoxyphenyl)piperazin-1-yl)methanone (S23). Compound **S23** was synthesized from 1-(4-methoxypheneyl)piperazine and compound **S22** by procedure 1 (60% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.21 (d, *J* = 3.8 Hz, 1H), 6.99 (d, *J* = 3.8 Hz, 1H), 6.93 – 6.88 (m, 2H), 6.87 – 6.83 (m, 2H), 3.91 – 3.84 (m, 4H), 3.77 (s, 3H), 3.13 – 3.05 (m, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.2, 154.5, 145.1, 143.0, 136.7, 130.5, 119.0, 114.6, 55.6, 51.2, 45.5.



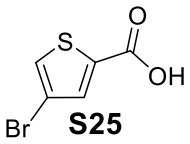
S24

(5-(2-Bromophenyl)thiophen-2-yl)(4-(4-methoxyphenyl)piperazin-1-yl)methanone (S24). Compound **S24** was synthesized from compound **S23** by procedure 2 (92% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.69 (dd, *J* = 8.1, 1.2 Hz, 1H), 7.49 (dd, *J* = 7.7, 1.7 Hz, 1H), 7.35 (td, *J* = 7.5, 1.2 Hz, 1H), 7.33 (s, 1H), 7.24 (d, *J* = 3.8 Hz, 1H), 7.22 (ddd, *J* = 8.0, 7.5, 1.8 Hz, 1H), 6.95 – 6.90 (m, 2H), 6.88 – 6.83 (m, 2H), 3.95 (dd, *J* = 6.0, 4.1 Hz, 4H), 3.78 (s, 3H), 3.17 – 3.08 (m, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 163.3, 154.5, 145.2, 145.1, 137.2, 134.3, 133.9, 131.9, 129.7, 129.0, 127.6, 127.4, 122.7, 119.0, 114.6, 55.6, 51.3, 45.1.



KW5175

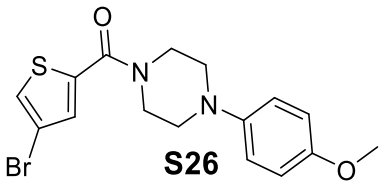
(2-(5-(4-(4-methoxyphenyl)piperazine-1-carbonyl)thiophen-2-yl)phenyl)boronic acid (KW5175). **KW5175** was synthesized from compound **S23** by procedure 3 (45% yield). ^{1}H NMR (500 MHz, MeOD) δ 7.58 (dt, *J* = 7.6, 1.0 Hz, 1H), 7.48 – 7.38 (m, 4H), 7.14 (d, *J* = 3.8 Hz, 1H), 7.01 – 6.96 (m, 2H), 6.89 – 6.84 (m, 2H), 3.98 – 3.89 (m, 4H), 3.75 (s, 3H), 3.17 – 3.08 (m, 4H). ^{13}C NMR (126 MHz, MeOD) δ 163.7, 154.7, 149.5, 145.2, 136.1, 135.9, 132.0, 130.3, 129.1, 127.8, 127.7, 124.2, 118.8, 114.1, 54.5, 51.0, 48.5. ^{11}B NMR (160 MHz, MeOD) δ 31.2. HRMS (ESI) calc. for $[M+H]^+$ $C_{22}H_{24}N_2O_4BS$ 423.1550 obs. 423.1150



S25

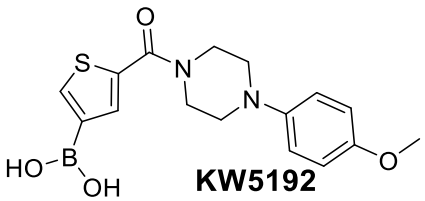
4-bromo-2-thiophenecarboxylic acid (S25). LiOH (0.48 g, 11 mmol) was dissolved in water (18 mL) to which was added a solution of methyl 4-bromo-2-thiophenecarboxylate (0.55 g, 2.5 mmol) in THF (2.5 mL). The mixture was stirred at room temperature for 4 h. The liquid was acidified to pH 4 with 6 M HCl and extracted three times with EtOAc (20 mL each). The organic fractions were pooled, dried over sodium sulfate, and concentrated under reduced pressure to give compound **S25** which was taken to the next step without further purification (0.44 g, 2.1 mmol, 84% yield). ^{1}H NMR (500 MHz, DMSO) δ 13.49 (s, 1H), 8.01 (s, 1H), 7.69 (s, 1H). ^{13}C NMR (126

MHz, DMSO) δ 162.3, 136.6, 135.1, 131.3, 110.0.



S26

(4-Bromothiophen-2-yl)(4-(4-methoxyphenyl)piperazin-1-yl)methanone (S26). Compound **S26** was synthesized from compound **S25** and 1-(4-methoxypheneyl)piperazine by procedure 1 (73% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 (s, 1H), 7.21 (s, 1H), 6.94 – 6.89 (m, 2H), 6.88 – 6.83 (m, 2H), 3.91 – 3.85 (m, 4H), 3.77 (s, 3H), 3.13 – 3.07 (m, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 162.0, 154.5, 145.1, 138.3, 130.9, 126.2, 119.0, 114.6, 109.4, 55.6, 51.2, 46.0.



KW5192

(5-(4-(4-Methoxyphenyl)piperazine-1-carbonyl)thiophen-3-yl)boronic acid (KW5192). **KW5192** was synthesized by procedure 3 from compound **S25** (59% yield). ^{1}H NMR (500 MHz, MeOD) δ 8.05 (s, 1H), 7.64 (s, 1H), 7.01 – 6.91 (m, 2H), 6.90 – 6.81 (m, 2H), 3.89 (t, J = 5.1 Hz, 4H), 3.74 (s, 3H), 3.08 (t, J = 5.1 Hz, 4H). ^{13}C NMR (126 MHz, MeOD) δ 164.2, 154.7, 145.1, 137.7, 136.1, 134.1, 118.8, 114.1, 54.5, 51.0, 48.5. ^{11}B NMR (160 MHz, MeOD) δ 26.3. HRMS (ESI) calc. for $[M+H]^+$ $C_{16}H_{20}N_2O_4BS$ 347.1237, obs. 347.1235.

References

1. Chamberlain, L. H.; Shipston, M. J., The physiology of protein S-acylation. *Physiol Rev* **2015,** *95* (2), 341-376.
2. Lin, H., Protein cysteine palmitoylation in immunity and inflammation. *Febs j* **2021**.
3. Pan, Y.; Xiao, Y.; Pei, Z.; Cummins, T. R., S-Palmitoylation of the sodium channel Nav1.6 regulates its activity and neuronal excitability. *J Biol Chem* **2020,** *295* (18), 6151-6164.
4. Han, J.; Wu, P.; Wang, F.; Chen, J., S-palmitoylation regulates AMPA receptors trafficking and function: a novel insight into synaptic regulation and therapeutics. *Acta Pharm Sin B* **2015,** *5* (1), 1-7.
5. Noland, Cameron L.; Gierke, S.; Schnier, Paul D.; Murray, J.; Sandoval, Wendy N.; Sagolla, M.; Dey, A.; Hannoush, Rami N.; Fairbrother, Wayne J.; Cunningham, Christian N., Palmitoylation of TEAD Transcription Factors Is Required for Their Stability and Function in Hippo Pathway Signaling. *Structure* **2016,** *24* (1), 179-186.
6. Mukai, K.; Konno, H.; Akiba, T.; Uemura, T.; Waguri, S.; Kobayashi, T.; Barber, G. N.; Arai, H.; Taguchi, T., Activation of STING requires palmitoylation at the Golgi. *Nature communications* **2016,** *7*, 11932-11932.
7. Dixon, C. L.; Fairn, G. D., S-palmitoylation of NOD2 controls its localization to the plasma membrane. *J Lipid Res* **2021,** *62*, 100097-100097.
8. Zhang, M.; Zhou, L.; Xu, Y.; Yang, M.; Xu, Y.; Komaniecki, G. P.; Kosciuk, T.; Chen, X.; Lu, X.; Zou, X.; Linder, M. E.; Lin, H., A STAT3 palmitoylation cycle promotes TH17 differentiation and colitis. *Nature* **2020,** *586* (7829), 434-439.
9. Ivanov, I. I.; McKenzie, B. S.; Zhou, L.; Tadokoro, C. E.; Lepelley, A.; Lafaille, J. J.; Cua, D. J.; Littman, D. R., The Orphan Nuclear Receptor RORγt Directs the Differentiation Program of Proinflammatory IL-17+ T Helper Cells. *Cell* **2006,** *126* (6), 1121-1133.
10. Durant, L.; Watford, W. T.; Ramos, H. L.; Laurence, A.; Vahedi, G.; Wei, L.; Takahashi, H.; Sun, H.-W.; Kanno, Y.; Powrie, F.; O'Shea, J. J., Diverse Targets of the Transcription Factor STAT3 Contribute to T Cell Pathogenicity and Homeostasis. *Immunity* **2010,** *32* (5), 605-615.
11. Johnson, D. E.; O'Keefe, R. A.; Grandis, J. R., Targeting the IL-6/JAK/STAT3 signalling axis in cancer. *Nat Rev Clin Oncol* **2018,** *15* (4), 234-248.
12. Gálvez, J., Role of Th17 Cells in the Pathogenesis of Human IBD. *ISRN Inflamm* **2014,** *2014*, 928461-928461.
13. Chen, S.; Han, C.; Miao, X.; Li, X.; Yin, C.; Zou, J.; Liu, M.; Li, S.; Stawski, L.; Zhu, B.; Shi, Q.; Xu, Z. X.; Li, C.; Goding, C. R.; Zhou, J.; Cui, R., Targeting MC1R depalmitoylation to prevent melanomagenesis in redheads. *Nat Commun* **2019,** *10* (1), 877.
14. Toyoda, T.; Sugimoto, H.; Yamashita, S., Sequence, expression in Escherichia coli, and characterization of lysophospholipase II1The nucleotide sequence data reported in this paper have been submitted to the GenBank/EMBL/DDBJ Nucleotide Sequence Databases under the accession number AB009653.1. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1999,** *1437* (2), 182-193.

15. Duncan, J. A.; Gilman, A. G., A Cytoplasmic Acyl-Protein Thioesterase That Removes Palmitate from G Protein α Subunits and p21RAS*. *Journal of Biological Chemistry* **1998,** *273* (25), 15830-15837.
16. Hedberg, C.; Dekker, F. J.; Rusch, M.; Renner, S.; Wetzel, S.; Vartak, N.; Gerding-Reimers, C.; Bon, R. S.; Bastiaens, P. I.; Waldmann, H., Development of highly potent inhibitors of the Ras-targeting human acyl protein thioesterases based on substrate similarity design. *Angew Chem Int Ed Engl* **2011,** *50* (42), 9832-7.
17. Hoover, H. S.; Blankman, J. L.; Niessen, S.; Cravatt, B. F., Selectivity of inhibitors of endocannabinoid biosynthesis evaluated by activity-based protein profiling. *Bioorganic & Medicinal Chemistry Letters* **2008,** *18* (22), 5838-5841.
18. Adibekian, A.; Martin, B. R.; Chang, J. W.; Hsu, K. L.; Tsuboi, K.; Bachovchin, D. A.; Speers, A. E.; Brown, S. J.; Spicer, T.; Fernandez-Vega, V.; Ferguson, J.; Cravatt, B. F.; Hodder, P.; Rosen, H., Characterization of a Selective, Reversible Inhibitor of Lysophospholipase 2 (LYPLA2). In *Probe Reports from the NIH Molecular Libraries Program*, National Center for Biotechnology Information (US): Bethesda (MD), 2010.
19. Hernandez, J. L.; Davda, D.; Cheung See Kit, M.; Majmudar, J. D.; Won, S. J.; Gang, M.; Pasupuleti, S. C.; Choi, A. I.; Bartkowiak, C. M.; Martin, B. R., APT2 Inhibition Restores Scribble Localization and S-Palmitoylation in Snail-Transformed Cells. *Cell Chem Biol* **2017,** *24* (1), 87-97.
20. Turcotte, C.; Dumais, É.; Archambault, A.-S.; Martin, C.; Blanchet, M.-R.; Bissonnette, É.; Boulet, L.-P.; Laviolette, M.; Di Marzo, V.; Flamand, N., Human leukocytes differentially express endocannabinoid-glycerol lipases and hydrolyze 2-arachidonoyl-glycerol and its metabolites from the 15-lipoxygenase and cyclooxygenase pathways. *Journal of Leukocyte Biology* **2019,** *106* (6), 1337-1347.
21. Vujic, I.; Sanlorenzo, M.; Esteve-Puig, R.; Vujic, M.; Kwong, A.; Tsumura, A.; Murphy, R.; Moy, A.; Posch, C.; Monshi, B.; Rappersberger, K.; Ortiz-Urda, S., Acyl protein thioesterase 1 and 2 (APT-1, APT-2) inhibitors palmostatin B, ML348 and ML349 have different effects on NRAS mutant melanoma cells. *Oncotarget* **2016,** *7* (6), 7297-306.
22. Virlogeux, A.; Scaramuzzino, C.; Lenoir, S.; Carpentier, R.; Louessard, M.; Genoux, A.; Lino, P.; Hinckelmann, M. V.; Perrier, A. L.; Humbert, S.; Saudou, F., Increasing brain palmitoylation rescues behavior and neuropathology in Huntington disease mice. *Sci Adv* **2021,** *7* (14).
23. Won, S. J.; Davda, D.; Labby, K. J.; Hwang, S. Y.; Pricer, R.; Majmudar, J. D.; Armacost, K. A.; Rodriguez, L. A.; Rodriguez, C. L.; Chong, F. S.; Torossian, K. A.; Palakurthi, J.; Hur, E. S.; Meagher, J. L.; Brooks, C. L.; Stuckey, J. A.; Martin, B. R., Molecular Mechanism for Isoform-Selective Inhibition of Acyl Protein Thioesterases 1 and 2 (APT1 and APT2). *ACS Chemical Biology* **2016,** *11* (12), 3374-3382.

CHAPTER 3

SIRT2 Inhibitors with Benzodiazepinedione Cores Have Improved Water Solubility, Bioavailability, and Anticancer Activity[a]

Abstract

SIRT2 is a class III histone deacetylase that has been shown to promote tumor progression though multiple mechanisms. Our lab has previously synthesized and tested the mechanism based SIRT2 inhibitor TM and shown it be effective at reducing tumor burden in mice. TM contains a thiomyrisoyl lysine core and its poor water solubility has been a major limiting factor for its therapeutic development. Here we present a new SIRT2 selective inhibitor, NH-C1-10, which replaces the thiomyristoyl lysine with a benzodiazepinedione. We show NH-C1-10 is more water soluble and more potent in cells than TM and deters pancreatic tumor progression in mice.

[a] This project was done in collaboration with a former lab member, Dr. Jun Young (Nick) Hong. Dr. Hong synthesized the NH series compounds, and collected the data presented in figures 3.2 and 3.3 B as well as Figure 3.5 with the help of Irma Fernandez.

Introduction

Sirtuins are the class III histone deacetylases of which there are seven in human (SIRT1-7)[1]. Though classified as deacetylases, sirtuins can remove numerous acyl groups from acylated lysine residues[2-4]. SIRT2 can remove acetyl, myristyl, 4-oxononaoyl, and benzoyl groups from acylated lysine residues[1]. Sirtuins use of nicotinamide adenine dinucleotide (NAD^+) as a co-substrate for the deacylation reaction[5]. SIRT2 has been shown to regulate many cellular functions including mitosis, oxidative stress response, and autophagy among others, by regulating the activity of numerous enzymes through lysine deacetylation[2, 6-10].

Deacetylation of proteins by SIRT2 is also known to promote tumorigenesis through multiple mechanisms. SLUG is stabilized by SIRT2 deacetylation which promotes breast cancer metastasis[11]. c-Myc is a transcription factor which promotes transcription proliferation genes. It is a frequently overexpressed oncoprotein in many cancers, and like SLUG, deacetylation hinders its degradation[12]. SIRT2 also increases the metabolic rate of cancer cells by increasing the activity of lactate dehydrogenase A (LDH-A)[13]. As a result, inhibitors of SIRT2 have been pursued as anti-cancer agents. Our lab has previously developed a mechanism-based SIRT2 selective inhibitor, a thiomyristoyl lysine compound called TM (Figure 3.1). TM treatment promotes c-Myc degradation and shows broad anti-cancer effects in mice[14].

A major factor that has limited TM as a possible cancer therapeutic is its poor water solubility due to the 14-carbon acyl chain. Attempts were made to improve water solubility by replacing the thioamide moiety with thioureas and attaching a quaternary ammonium or glucose to the lysine core. Thioureas did not significantly improve the

water solubility and the inhibitors containing quaternary ammonium or glucose had very poor membrane permeability[15-17]. To improve water solubility while maintaining membrane permeability, here we synthesized thioamides and thioureas with benzodiazepinediones to replace the lysine core structure. We show that one of the new inhibitors, NH-C1-10, has greatly improved water solubility compared to TM. We also show that this compound inhibits SIRT2 *in vitro* and in cells and reduces growth of pancreatic tumors in mice.

Results and Discussion

Design and Synthesis of Benzodiazepinedione SIRT2 Inhibitors

To make more water soluble SIRT2 inhibitors we first changed the cbz protected lysine core of TM to a benzodiazepinedione (BDD) core containing a 5-member aliphatic ring derived from L-proline. We used the program DataWarrior to simulate the octanol water partition coefficients (cLogP) for TM and the new BDD compounds. Lower cLogP values indicate better water solubility. For TM this value is 9.25. Several BDD core structures were synthesized to further improve water solubility by shorting the aliphatic ring or by incorporating a morpholine or piperazine moiety. A BDD made from D-proline was synthesized to examine the effects of reversing the chirality of the core structure. Thiourea and thioamide derivatives with shortened acyl chains were also targeted to decrease hydrophobicity. Lastly, NH-C3-10 was synthesized without a BDD to confirm the BDD structure is required for SIRT2 inhibition (Figure 3.1).

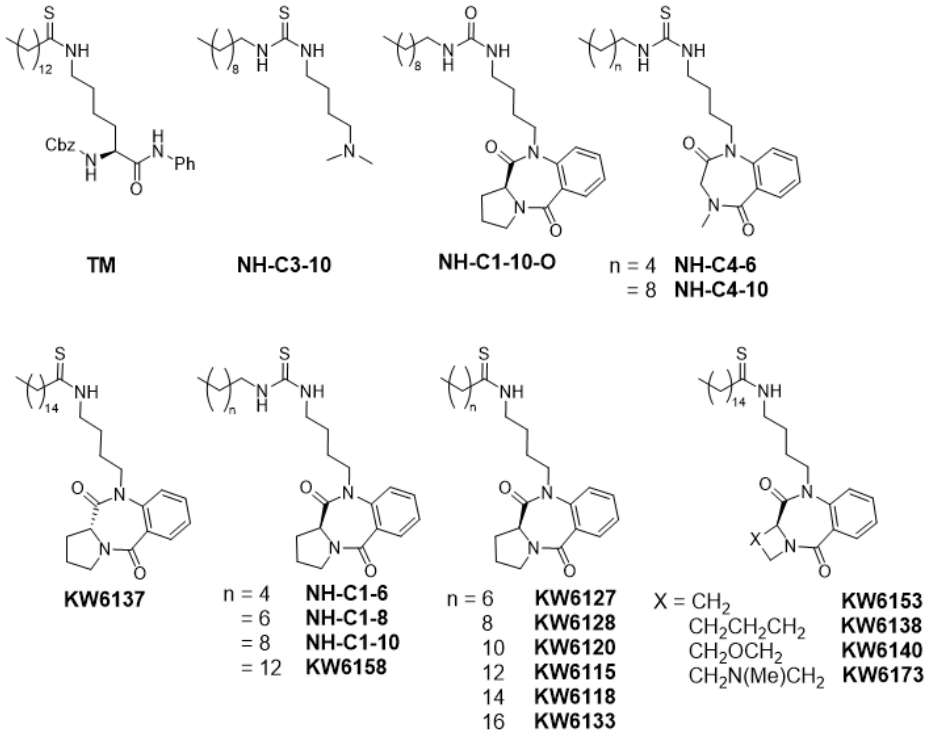


Figure 3.1. SIRT2 inhibitor structures. Derivatives were synthesized to examine how the acyl chain length, sulfur containing functional group, BDD structure, and chirality effected SIRT2 inhibition. NH-C3-10 and NH-C1-10-O were synthesized as controls to confirm the necessity of the BDD core and thiourea\thioamide for SIRT2 inhibition.

Thioureas and Thioamides are Equally Potent Inhibitors

The ability of the new compounds to inhibit SIRT2 was tested with a sirtuin deacetylation enzymatic assay to calculate half maximal inhibition concentrations (Table 3.1). *In vitro* IC_{50} values were determined for SIRT1, 2, 3, and 5 to confirm selectivity for SIRT2. Acyl chain lengths ranging between 8 and 18 carbons were synthesized and tested. Thioamides derived from palmitic acid, such as KW6118, were the most potent with the DBB cores whereas previous inhibitors with lysine-based cores, such as TM, had the greatest inhibition with myristic thioamides. This suggests the BDDs may be positioned slightly differently in the SIRT2 binding pocket. Reversing the chirality of the enantiomeric carbon in KW6137 reduced the potency 4-fold as well

as the selectivity for SIRT2, indicating SIRT2 prefers BDDs derived from L amino acids. There was a small increase in IC_{50} values between the thioamides and thioureas of the same chain length as seen with NH-C1-10 and KW6120 having IC_{50} values of 6.09 and 5.88 μM respectively. However, thioureas produced the same trend of decreased IC_{50} values with increasing chain length.

SIRT2 inhibition increased with increasing acyl chain length up to 16 carbons with KW6118 being the most potent compound *in vitro* having an IC_{50} value of 0.037 μM. Neither NH-C3-10 nor NH-C1-10-O inhibit SIRT2 below 83 μM, indicating the BDD core and the sulfur in the thioamide\thiourea groups are critical components of the new inhibitors. The size of the aliphatic ring, between 4 and 6 members, did not have a major impact on the on the potency of the compounds, nor did removing the ring completely as was done with NH-C4-6 and NH-C4-10. KW6138 with a 6-member aliphatic ring containing 5 carbons had and IC_{50} value of 0.147 μM, 4-fold higher than KW6118. However, when the larger ring BDDs contain a second heteroatom as is the case with KW6140 and KW6173, the loss in potency is only 2-3-fold and cLogP is reduced from 8.31 in KW6118 to ~7.5.

Table 3.1. *In Vitro* IC_{50} Values for SIRT2 Inhibitors (μM)

Compound	SIRT1	SIRT2	SIRT3	SIRT5
NH-C3-10	>83	>83	>83	>83
NH-C1-10-O	>83	>83	>83	>83
NH-C4-6	>83	>83	>83	>83
NH-C4-10	>83	7.86 ± 3	>83	>83
NH-C1-6	>83	>83	>83	>83
NH-C1-8	>83	49.7 ± 20	>83	>83
NH-C1-10	>83	6.09 ± 3	>83	>83
KW6158	>83	0.035 ± 0.009	54.4 ± 20	>83
KW6137	21.6 ± 6	0.159 ± 0.06	23.0 ± 9	>83
KW6127	>83	18.1 ± 4	>83	>83
KW6128	>83	14.2 ± 7	>83	>83
KW6120	>83	5.88 ± 4	>83	>83
KW6115	>83	1.75 ± 0.8	>83	>83
KW6118	>83	0.037 ± 0.01	>83	>83
KW6133	38.2 ± 9	0.180 ± 0.1	35.9 ± 10	>83
KW6153	>83	0.083 ± 0.03	42.6 ± 20	>83
KW6138	33.6 ± 10	0.147 ± 0.06	18.5 ± 8	>83
KW6140	65.9 ± 20	0.060 ± 0.04	60.1 ± 20	>83
KW6173	>83	0.098 ± 0.01	>83	>83

NH-C1-10 is Cytotoxic in Cancer Cells and Forms Stalled Covalent Intermediates

SIRT2 is a known tumor activator and treatment with SIRT2 inhibitors reduces the proliferation of cancer cells. Cell survival assays were done to determine the cytotoxicity of the compounds in MFC7 (estrogen receptor and progesterone receptor positive breast cancer), MDA-MB-231 (triple negative breast cancer), BxPC-3 (pancreatic cancer), and Hela (cervical carcinoma) cells. (Table 3.2). Overall, the compounds ranged from nontoxic to mildly cytotoxic. Only a few compounds were able to produce a cytotoxic effect in all the tested cell lines, and of these NH-C1-10 and KW6158 were the most toxic. Only two thioamides showed toxicity in all cell lines which were the ones containing a unique BDD designed to further enhance water solubility (KW6153 and KW6140). This could be because increased solubility leads to increased bioavailability. Based on the *in vitro* results, NH-C1-10 was selected as the lead compound as a compromise between *in vitro* potency, cytotoxicity, and water solubility. Though 16 carbon acyl chain thioamides were more potent *in vitro*, they also had much higher cLogP values ranging between 8.3 and 7.5 compared to 5.64 for NH-C1-10. Furthermore, decreased *in vitro* IC_{50} values had not resulted in greater cytotoxicity. Thus, we decided to move forward with NH-C1-10 as it was expected to have better bioavailability due to being smaller and having better aqueous solubility.

Table 3.2 GI_{50} values (μM) of SIRT2 Inhibitors in Four Different Cancer Cell Lines

Compound	MCF7	MDA-MB-231	BxPC-3	Hela
NH-C3-10	17.7 ± 1	30.5 ± 3	-	24.8 ± 4
NH-C1-10-O	>50	>50	-	>50
NH-C4-6	>50	>50	-	>50
NH-C4-10	>50	>50	-	>50
NH-C1-6	>50	>50	-	>50
NH-C1-8	>50	>50	-	>50
NH-C1-10	14.0 ± 3	19.2 ± 3	43.9 ± 10	12.0 ± 1
KW6158	9.56 ± 1	12.1 ± 1	17.7 ± 4	16.8 ± 3
KW6137	13.9 ± 2	15.8 ± 1	>50	>50
KW6127	32.4 ± 10	>50	>50	>50
KW6128	23.1 ± 4	>50	>50	22.5 ± 3
KW6120	13.6 ± 0.5	19.3 ± 2	>50	18.3 ± 2
KW6115	12.6 ± 2	18.6 ± 1	>50	12.6 ± 1
KW6118	14.5 ± 1	17.8 ± 2	>50	24.8 ± 6
KW6133	27.6 ± 7	29.0 ± 6	>50	>50
KW6153	15.8 ± 2	21.2 ± 10	45.6 ± 8	15.1 ± 3
KW6138	17.6 ± 2	17.3 ± 5	>50	34.1 ± 5
KW6140	14.4 ± 1	17.2 ± 1	48.7 ± 10	25.2 ± 3
KW6173	43.5 ± 6	34.9 ± 8	-	>50

Lysine-based thioureas and thioamides have been shown to inhibit SIRT2 by forming a covalent intermediate with NAD^+ [15]. We confirmed that NH-C1-10 also formed a covalent intermediate with NAD^+ and that the new core structure had not changed the mode of inhibition. This was done by LC-MS analysis of the products that formed after SIRT2 was treated with NH-C1-10 and NAD^+ (Figure 3.2). We identified the stalled covalent intermediate, which is similar to that previously observed with TM. This indicates NH-C1-10 is a mechanism-based inhibitor and works by the mechanism previously established for our lysine-based thioamide inhibitors.

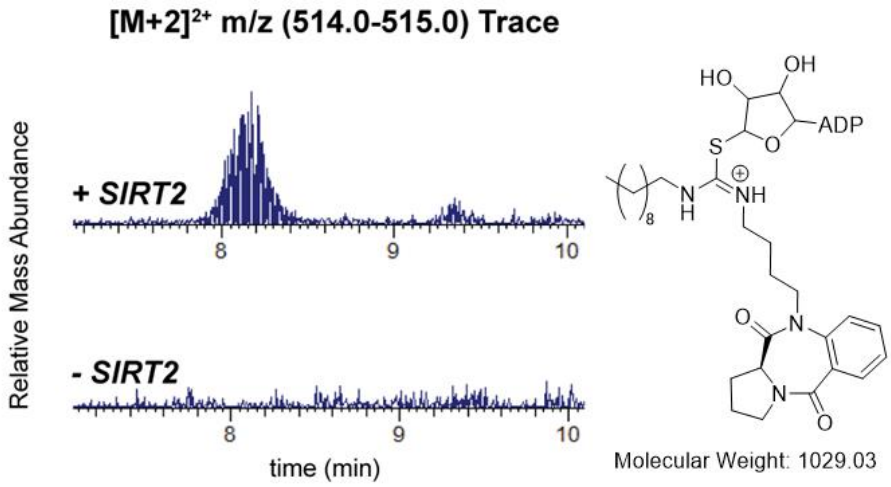


Figure 3.2. Stalled covalent intermediate trapping with NH-C1-10. Products of SIRT2 enzymatic reactions were detected by LC-MS. The NH-C1-10/NAD^+ conjugate is observed only in the presence of SIRT2 indicating the thiourea inhibitor undergoes a reaction with SIRT2 similar to TM, forming a stalled covalent intermediate.

NH-C1-10 Increases Alpha Tubulin Acetylation and Deters Tumor Progression

We confirmed the new BDD compounds could inhibit SIRT2 in cells with an acetyl α tubulin immunofluorescence assay. SIRT2 deacetylates α tubulin so its inhibition results in higher fluorescent intensities when cells are imaged with an acetyl α tubulin antibody.

NH-C1-10, KW6118, and KW6140 were compared using 100 μM concentrations (Figure 3.3 A). Testing these compounds allowed us to compare the effect of changing sulfur containing group, the DBB, and the cLogP values of the inhibitors. The compounds all increased tubulin acetylation to different extents. NH-C1-10 showed the largest increase, doubling the amount of tubulin acetylation. KW6140 produced a 72% increase while KW6118 treatment only resulted in 21% increase in acetylation over the untreated control. These results again show that the most potent inhibitors *in vitro* do not produce the most SIRT2 inhibition in cells and that increasing water solubility has the desired effect of increasing bioavailability.

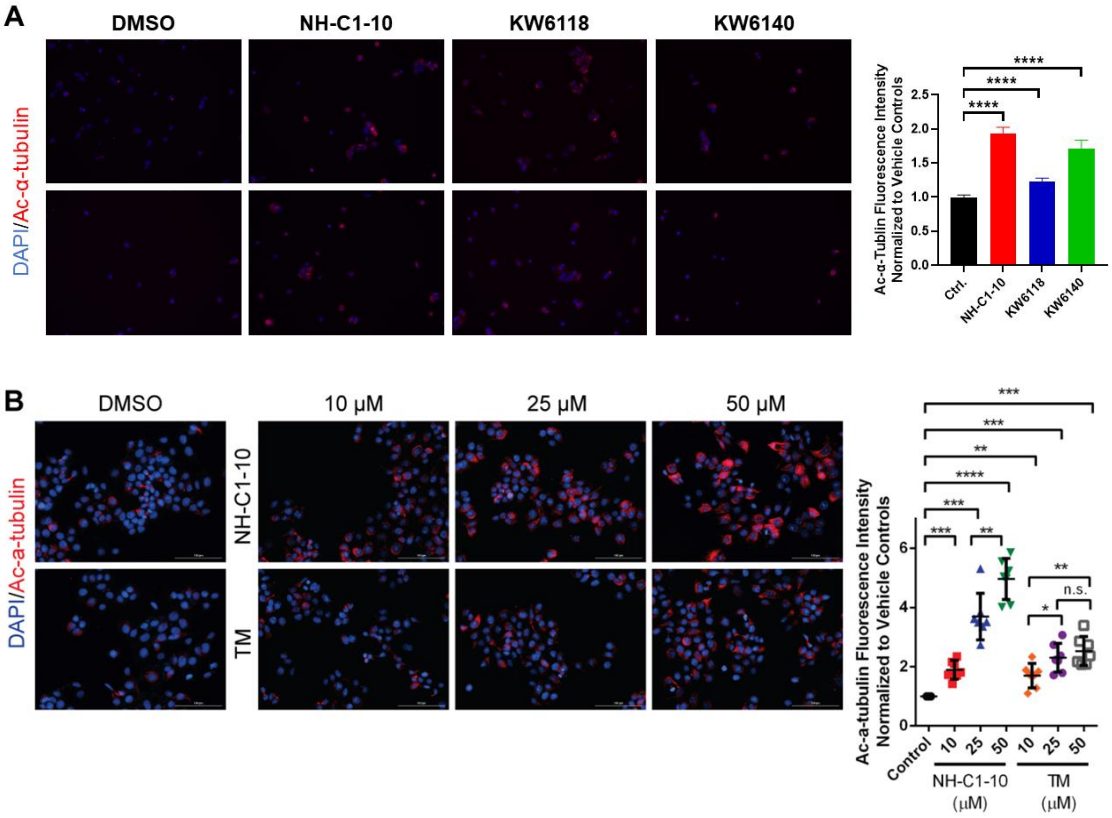


Figure 3.3. Acetyl α tubulin immunofluorescence imaging with SIRT2 inhibitors. MCF7 cells were incubated with SIRT2 inhibitors for 6 h then fixed and incubated with anti-acetyl α tubulin rabbit monoclonal antibody then a TAMRA conjugated anti-rabbit antibody. TAMRA signal intensity was integrated to determine tubulin acetylation. (**A**)

Cells were treated with 100 μM of NH-C1-10, KW6118, and KW6140. (**B**) NH-C1-10 and TM were used to treat cells at the indicated concentrations.

NH-C1-10 was also compared with TM (Figure 3.3 B). Lower concentrations were used due to TM's poor water solubility. Both compounds produced increased fluorescent levels over the vehicle treated samples, though for NH-C1-10 the increase was 5-fold at 50 μM compared to only about 3-fold with TM at the same concentration. TM did not show any significant increase in fluorescence when the inhibitor concentration was increased from 25 μM to 50 μM which likely resulted from its poor water solubility.

NH-C1-10 was then tested in mice to determine how well it prevents solid tumor growth. We grafted the pancreatic cancer cell line BxPC-3 into immunocompromised male mice. BxPC-3 cells used because while SIRT2 inhibitors were tested in several tumor models, they have not been tested in pancreatic cancer mouse models. After tumors were established, the mice were treated with 50 mg/kg of NH-C1-10 by intraperitoneal injection every day for 15 days. The health of the mice, relative tumor volumes, and body weights were tacked over the course of the experiment. After 15 days the mice were sacrificed, and tumors were massed.

NH-C1-10 reduced tumor volume consistently during the experiment (Figure 3.4 A). NH-C1-10 also significantly hindered tumor growth and resulted in an average tumor mass of 1030 mg compared to 1550 mg in vehicle treated mice (Figure 3.4 B). No significant changes were observed in body weight (Figure 3.4 C) or health status during the experiment indicating the compound is well tolerated by mice.

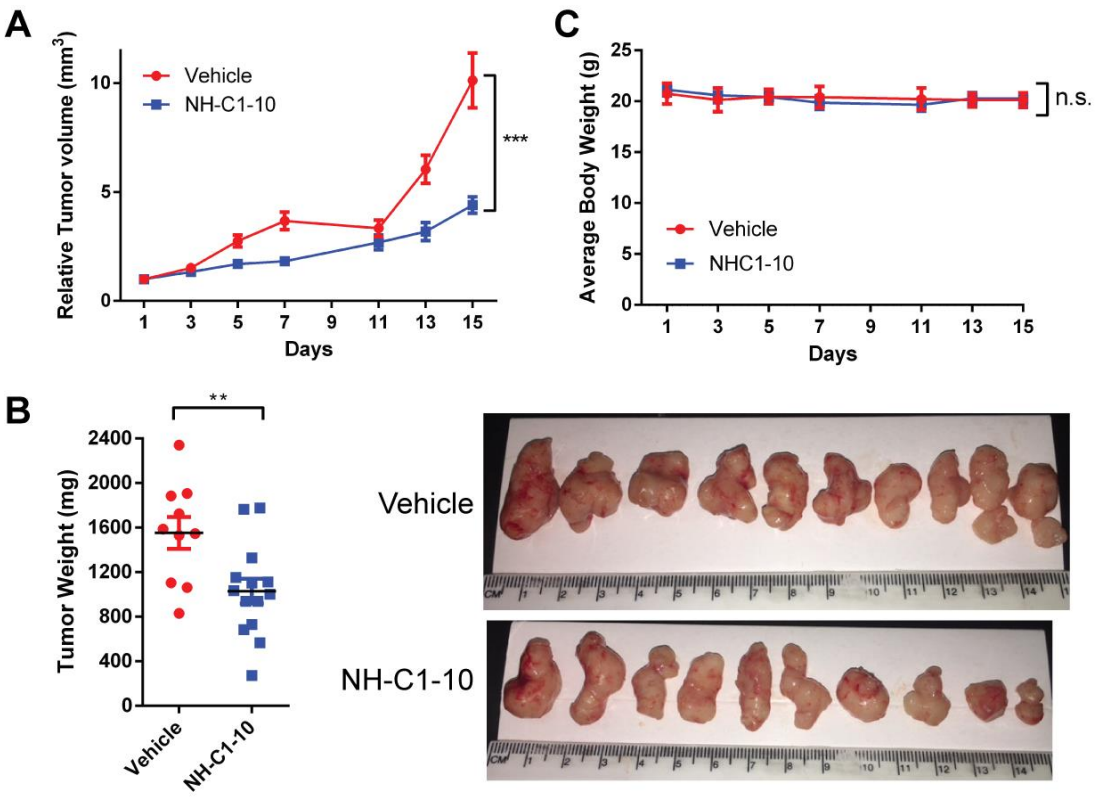


Figure 3.4. NH-C1-10 decreases tumor growth in BxPC-3 tumor xenografts. (**A**) Tumor volume over time for NH-C1-10 (50 mg/kg) or vehicle (10% DMSO, 90% PBS) treated mice. NSG mice injected with BxPC-3 human pancreatic cancer cells were treated every day for 15 days. (*** P value = 0.0005, Two-tailed Student's t-test, n=10 tumors for vehicle, n=14 tumors for NH-C1-10). (**B**) Tumor weights and representative tumor images of BxPC-3 xenograft mice treated with NH-C1-10 (50 mg/kg) or vehicle for 15 days (** P value = 0.0096, Two-tailed Student's t-test, n = 10 tumors for vehicle, n = 14 tumors for NH-C1-10). (**C**) Average body weight of mice treated with NH-C1-10 (50 mg/kg) or vehicle for 15 days. No difference in body weight was observed between the groups. N = 5 mice for vehicle and n = 7 mice for NH-C1-10. Data shown in Figures A, B, and C are the mean values ± standard deviation.

In summary, we synthesized 20 new thioamide and thiourea SIRT2 inhibitors containing BDD cores with the aim of improving the water solubility and bioavailability of the inhibitors. We showed these compounds remain highly potent *in vitro* and cytotoxic in breast, pancreatic, and cervical cancer cells. The thiourea NH-C1-10 selectively inhibits SIRT2 *in vitro* and in cells by forming stalled covalent intermediates

with NAD^+. NH-C1-10 inhibits SIRT2 in cells better than TM indicating the BDD core structure improves inhibitor bioavailability. Tumor xenograft experiments in mice showed that NH-C1-10 greatly impairs the growth of BxPC-3 derived tumors and that treatment with this compound is well tolerated by mice. Taken together, these results show we have synthesized an effective SIRT2 selective inhibitor with greatly improved water solubility and that SIRT2 remains a promising target for developing cancer therapeutics.

Materials and Methods

All solvents and reagents were purchased from commercial vendors as analytical or higher-grade purity. Flash chromatography was done using SiliaFlash Irregular Silica Gel, P60, 40 – 63 μm, 60 Å. NMR spectra were collected at the Cornell NMR Facility using Brucker 400 and 500 spectrometers. HRMS data was collected at the Cornell Chemistry Mass Spectrometry Facility using a Thermo Exactive Orbitrap ESI mass spectrometer. For the enzymatic assays a Shimadzu HPLC LC20-AD connected to Kinetex 5u EVO C18 100 Å column (100 mm x 4.60 mm, 5 μm) and a Shimadzu SPD-20AV UV detector were used. The UV detector was set to monitor 215 nm and 280 nm absorbance to detect acetylated and deacetylated H3K9 peptides. HPLC grade water and acetonitrile with 0.1% HPLC grade TFA were used as mobile phases for the separations with a flow rate of 0.5 mL/min.

All animals used in this study were handled in accordance with federal and institutional guidelines, under a protocol approved by the Cornell University Institutional Animal Care and Use Committee (IACUC). Mice were housed under specific pathogen-free

conditions in an Association for the Assessment and Accreditation of Laboratory Animal Care International accredited facility and cared for in compliance with the Guide for the Care and Use of Laboratory Animals. Mice were on a 12:12 light:dark cycle in individually ventilated cages and received irradiated food and reverse osmosis, hyper-acidified water *ad libitum*.

Cell cultures and cytotoxicity assays

MCF7, MDA-MB-231, and Hela cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS). BxPC-3 cells were cultured in RPMI-1640 media supplemented with 10% FBS. All cells were kept in an incubator at 37 ^{o}C and 5% CO_2. For cytotoxicity assays, 3000 cells in 100 μL of media were added to each well of a 96 well plate. Plates were incubated for 24 h. 100 μL of 2X stock solutions of inhibitor in media was added to each well. Plates were incubated for an additional 3 days. 100 μL of media was removed from the treated wells and replaced with 20 μL of Cell Titer Blue (Promega G8080). Cells were incubated at 37 ^{o}C with the fluorescent intensities recorded after 2, 3, and 4 h to determine cell survival.

Expression and purification of SIRT1, 2, 3, & 5

These enzymes were purified following previously reported procedures.[18,19]

Enzymatic Assays for SIRT1, 2, 3, & 5 IC_{50} deacetylation activity

These assays were performed as reported in pervious papers.[18,19]

Acetyl α tubulin immunofluorescence imaging

This protocol was performed as reported in a previous paper.[16]

Treatment of human pancreatic cancer xenograft mice with NH-C1-10

Xenotransplantation studies were performed as previously described. Briefly, five million BxPC-3 human pancreatic cancer cells suspended in 100 μL of PBS were injected subcutaneously bilaterally on the flanks of male NOD.*Cg-Prkdc*^{scid}*Il2rg*^{tm1Wjl}/SzJ (NSG) mice from the Jackson Laboratory. When tumor size reached 1 cm^3, mice were euthanized, and their primary tumors used to implant into more NSG mice for subsequent cohorts for drug testing. BxPC-3 tumors were minced into 1-2mm pieces under sterile conditions. One tumor piece was then implanted per mouse subcutaneously bilaterally on the flanks of male NSG mice. Once the tumor size grew to 100-200 mm^3, mice were treated daily by intraperitoneal injection for 15 days with either NH-C1-10 (50 mg/kg) or vehicle control (10% DMSO, 10% Kolliphor, in 1X PBS). Overall health, body weight, and tumor development were monitored every other day with measurements of body weight and tumor volume using a caliper. After 2 weeks of treatment or if mice met humane endpoint criteria, mice were sacrificed by CO_2 asphyxiation and necropsied. The tumors were weighed for the data collection and processed/stored for subsequent analyses.

Synthetic Protocols

Procedure 1, benzodiazepinediones. An amino acid (1.3 eq.) and isatoic anhydride (1 eq.) were combined in DMSO (1M) and heated to 140 ^{o}C for 16 h. The solution was cooled to room temperature then poured into ice water. This aqueous solution was allowed to warm to room temperature and extracted 3 times with EtOAc. The organic layers were combined dried over sodium sulfate and concentrated. The resulting crude solid was stirred in Et_2O for 12 h. Vacuum filtration was used to collect the resulting solid which was taken to the next step without further purification.

Procedure 2, benzodiazepinedione nitriles. A benzodiazepinedione (1 eq.) was dissolved in THF (0.15M) and hexamethylphosphoramide (6 eq.) and chilled to 0 ^{o}C. Sodium hydride (2 eq.) was added portion-wise, and the reaction was stirred at 0 ^{o}C for 1 h. 4-Bromobutyronitrile (2 eq.) was added dropwise. The reaction was allowed to warm to room temperature and stirred for 16 h. Solvent was removed under reduced pressure. The resulting crude material was diluted with DCM and washed with water and brine. The organic layer was dried over sodium sulfate and concentrated under reduced pressure. The second crude was purified by flash chromatography (97:3, DCM : MeOH).

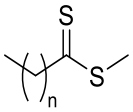
Procedure 3, Boc amino benzodiazepinediones. A benzodiazepinedione nitrile was dissolved in MeOH (0.15M) and chilled to 0 ^{o}C. Nickel chloride hexahydrate (0.2 eq.) and Boc_2O (2 eq.) were added to the solution followed by portion wise addition of $NaBH_4$ (6 eq.). The solution was allowed to warm to room temperature and stirred for 16 h. Water (1.5 volumes) and ammonia (0.2 volumes, 7M in MeOH) were added to the

reaction vessel and stirred for 1 h. The solution was extracted 3 times with EtOAc. Organic layers were combined dried over sodium sulfate and concentrated under reduced pressure. The crude material was purified by flash chromatography (3:2 → 4:1, EtOAc : Hexane).

Procedure 4, thioamide benzodiazepinediones. A Boc protected amino benzodiazepinedione (1 eq.) was dissolved in DCM (0.3M) and TFA (0.2 volumes). The solution was stirred at room temperature for 2 h then concentrated under reduced pressure. The resulting solid was dissolved in EtOH (0.1M) to which was added Et_3N (3 eq.) and a methyl dithioate (1 eq.). This solution was heated to 60 ^oC and stirred for 16 h. Solvent was removed under reduced pressure. The crude material was diluted with water and extracted 3 times with EtOAc. Organic layers were combined, washed with brine, dried over sodium sulfate, and concentrated under reduced pressure. The second crude material was purified by flash chromatography (2:3 → 3:2, EtOAc : Hexane) to yield thioamide products.

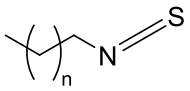
Procedure 5, thiourea benzodiazepinediones. A boc amino benzodiazepinedione (1 eq.) was dissolved in DCM (0.3M) and TFA (0.2 volumes). The solution was stirred at room temperature for 2 h then concentrated under reduced pressure. The resulting material was dissolved in DCM (0.1M) and Et_3N (3 eq.) and chilled to 0 ^oC. The corresponding isothiocyanate (1 eq.) was added slowly with stirring. This solution was allowed to warm to room temperature and stirred for 16 h. DCM was added to dilute the reaction which was then washed with water and brine and dried over sodium sulfate.

Solvent was removed under reduced pressure and the crude material was purified by flash chromatography (3:1, EtOAc : Hexane) to yield thiourea products.



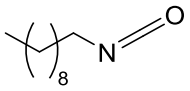
methyl octanedithioate (n = 6), methyl decanedithioate (n = 8), methyl dodecanedithioate (n = 10), methyl tetradecanedithioate (n = 12), methyl hexadecanedithioate (n = 14), and **methyl octadecanedithioate (n = 16).** Octanoic acid, decanoic acid, dodecanoic acid, tetradecanoic acid, hexadecenoic acid, or octadecanoic acid (1 eq.) was dissolved in toluene (0.1M) and Davy Reagent Methyl (0.5 eq.) was added. The solution was heated to reflux for 1 h. The reaction mixture cooled to room temperature and solvent was removed under reduced pressure. Crude material was purified by column chromatography (100% hexane) to give the corresponding methyl dithioate (36%, 44%, 37%, 26%, 38%, and 37% yield respectively). **methyl octanedithioate** ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.00 – 2.93 (m, 2H), 2.54 (s, 3H), 1.83 – 1.69 (m, 2H), 1.36 – 1.12 (m, 8H), 0.86 – 0.75 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 240.1, 52.0, 31.7, 31.4, 29.0, 28.8, 22.7, 20.0, 14.1. **methyl decanedithioate** ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.01 – 2.93 (m, 2H), 2.55 (s, 3H), 1.81 – 1.69 (m, 2H), 1.34 – 1.12 (m, 12H), 0.86 – 0.73 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 240.3, 52.1, 31.9, 31.4, 29.5, 29.3, 29.3, 28.8, 22.7, 20.0, 14.1. **methyl dodecanedithioate** ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.02 – 2.92 (m, 2H), 2.54 (s, 3H), 1.82 – 1.70 (m, 2H), 1.32 – 1.12 (m, 16H), 0.88 – 0.74 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 240.3, 52.1, 31.9, 31.4, 29.6, 29.5, 29.4, 29.3, 28.8, 22.7, 20.0, 14.4. **methyl tetradecanedithioate** ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.01 – 2.92 (m, 2H), 2.55 (s, 3H),

1.83 – 1.72 (m, 2H), 1.32 – 1.10 (m, 20H), 0.81 (td, *J* = 6.9, 1.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 240.3, 52.1, 31.9, 31.6, 31.4, 29.7, 29.7, 29.6, 29.5, 29.4, 29.3, 28.8, 22.7, 22.7, 20.0, 14.1. **methyl hexadecanedithioate** ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.04 – 2.91 (m, 2H), 2.54 (s, 3H), 1.81 – 1.70 (m, 2H), 1.18 (s, 26H), 0.81 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 240.2, 52.1, 32.0, 31.4, 29.7, 29.7, 29.7, 29.7, 29.7, 29.6, 29.5, 29.4, 29.3, 28.8, 22.7, 20.0, 14.1. **methyl octadecanedithioate** ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.00 – 2.93 (m, 2H), 2.55 (s, 3H), 1.81 – 1.69 (m, 2H), 1.18 (s, 27H), 0.81 (t, *J* = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 240.3, 52.1, 32.0, 31.4, 29.7, 29.7, 29.7, 29.6, 29.5, 29.4, 29.3, 28.8, 22.7, 20.0, 14.1.

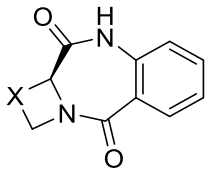


1-isothiocyanatohexane (n = 4) 1-isothiocyanatooctane (n = 6) 1-isothiocyanatodecane (n = 8), and **1-isothiocyanatotetradecane (n = 12).** 1-hexylamine, 1-octylamine, 1-decylamine or 1-tetradecylamine (1 eq.), Et_3N (1.5 eq.) and THF (0.1M) were combined. 1,1'-Thiocarbonyldiimidazole (1.5 eq.) was added, and the solution was stirred at room temperature for 16 h. The solution was diluted with water (2 volumes) and extracted 3 times with EtOAc. Organic layers were combined, washed with brine (1 volume), dried over sodium sulfate, and concentrated under reduced pressure. The resulting material was purified by flash chromatography (82% & 63% yield respectively). **1-isothiocyanatohexane** ^{1}H NMR (500 MHz, MeOD) δ 3.57 (t, *J* = 6.5 Hz, 2H), 1.79 – 1.65 (m, 2H), 1.52 – 1.42 (m, 2H), 1.37 (dtd, *J* = 9.1, 6.1, 4.0 Hz, 4H), 1.01 – 0.89 (m, 3H). ^{13}C NMR (126 MHz, MeOD) δ 44.6, 30.7, 29.7, 26.0,

22.2, 12.9. **1-isothiocyanatodecane** ^{1}H NMR (500 MHz, MeOD) δ 3.57 (t, *J* = 6.5 Hz, 2H), 1.80 – 1.64 (m, 2H), 1.53 – 1.42 (m, 2H), 1.42 – 1.29 (m, 8H), 1.03 – 0.82 (m, 3H). ^{13}C NMR (126 MHz, MeOD) δ 44.5, 31.5, 29.7, 28.9, 28.5, 26.3, 22.3, 13.0. **1-isothiocyanatodecane** ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.50 (t, *J* = 6.7 Hz, 2H), 1.75 – 1.62 (m, 2H), 1.41 (p, *J* = 6.9 Hz, 2H), 1.36 – 1.20 (m, 12H), 0.88 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 45.1, 31.9, 30.0, 29.5, 29.4, 29.3, 28.8, 26.6, 22.7, 14.1. **1-isothiocyanatotetradecane** ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.43 (t, *J* = 6.6 Hz, 2H), 1.62 (p, *J* = 6.8 Hz, 2H), 1.34 (dd, *J* = 12.9, 5.4 Hz, 2H), 1.20 (d, *J* = 9.0 Hz, 20H), 0.81 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 68.0, 45.1, 31.9, 30.0, 29.7, 29.7, 29.6, 29.5, 29.4, 29.4, 28.8, 26.6, 22.7, 14.1.



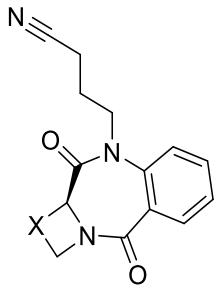
1-isocyanatodecane. Decylamine (1 mL, 5.00 mmol) and TEA (2.1 mL, 15.0 mmol) were combined in THF (40 mL). Triphosgene (1.78 g, 6.00 mmol) was slowly added at room temperature. After stirring for 16 h, the solvents were removed by rotary evaporation. The residue was re-dissolved in ethyl acetate (100 mL) and washed with water (90 mL) and brine (90 mL). The collected organic layer was dried over Na_2SO_4 and concentrated by rotary evaporation. The crude residue was further purified by silica gel column chromatography (Hexane: EtOAc = 4:1) to afford the final product (0.797 g, 87% yield). ^{1}H NMR (500 MHz, MeOD) δ 3.09 (t, *J* = 7.1 Hz, 2H), 1.49 (p, *J* = 6.9 Hz, 2H), 1.32 (d, *J* = 4.1 Hz, 14H), 0.92 (t, *J* = 6.8 Hz, 3H). ^{13}C NMR (126 MHz, MeOD) δ 158.2, 40.5, 31.7, 29.6, 29.3, 29.3, 29.1, 26.4, 22.3, 13.0, 13.0.



X = CH_2 **S1**
= CH_2CH_2 **S2**
= $CH_2CH_2CH_2$ **S3**
= CH_2OCH_2 **S4**
= $CH_2N(Boc)CH_2$ **S5**

(S)-1,10a-dihydroazeto[1,2-a]benzo[e][1,4]diazepine-4,10(2H,9H)-dione (S1), (S)-1,2,3,11a-tetrahydro-5H-benzo[e]pyrrolo[1,2-a][1,4]diazepine-5,11(10H)-dione (S2), (S)-7,8,9,10-tetrahydrobenzo[e]pyrido[1,2-a][1,4]diazepine-6,12(5H,6aH)-dione (S3), (S)-1,3,4,12a-tetrahydro-6H-benzo[e][1,4]oxazino[4,3-a][1,4]diazepine-6,12(11H)-dione (S4), and **tert-butyl (S)-6,12-dioxo-3,4,6,11,12,12a-hexahydrobenzo[e]pyrazino[1,2-a][1,4]diazepine-2(1H)-carboxylate (S5) S1-S5** were synthesized from (S)-azetidine-2-carboxylic acid, L-proline, (S)-piperidine-2-carboxylic acid, (S)-morpholine-3-carboxylic acid, and (S)-4-(tert-butoxycarbonyl)piperazine-2-carboxylic acid respectively by procedure 1 (69%, 54%, 61%, 38%, 42% yield respectively). **(S1)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.63 (s, 1H), 8.03 (dd, *J* = 7.9, 1.7 Hz, 1H), 7.49 (td, *J* = 7.7, 1.7 Hz, 1H), 7.27 (dd, *J* = 15.2, 1.1 Hz, 1H), 7.05 (dd, *J* = 8.0, 1.1 Hz, 1H), 4.72 (dd, *J* = 8.9, 5.8 Hz, 1H), 4.30 – 4.10 (m, 2H), 2.99 – 2.83 (m, 1H), 2.58 (ddt, *J* = 12.0, 8.6, 4.5 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.6, 165.6, 135.4, 132.9, 131.0, 125.2, 124.6, 121.6, 59.3, 46.9, 18.6. **(S2)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.01 (s, 1H), 7.92 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.45 (td, *J* = 7.7, 1.6 Hz, 1H), 7.31 – 7.19 (m, 1H), 7.02 (dd, *J* = 8.1, 1.1 Hz, 1H), 4.53 (dq, *J* = 13.6, 4.2 Hz,

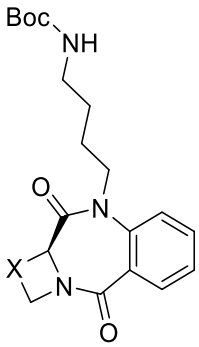
1H), 4.16 (dd, *J* = 6.4, 3.0 Hz, 1H), 3.00 (ddd, *J* = 13.6, 12.2, 4.0 Hz, 1H), 2.34 – 2.18 (m, 1H), 1.96 (dtt, *J* = 12.4, 10.0, 4.8 Hz, 1H), 1.84 (dp, *J* = 12.5, 4.1 Hz, 1H), 1.78 – 1.50 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.1, 168.6, 136.0, 132.1, 131.2, 127.6, 125.1, 120.4, 51.2, 40.4, 23.2, 22.8, 19.2. **(S3)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.50 (s, 1H), 7.91 (dd, *J* = 7.9, 1.6 Hz, 1H), 7.43 (td, *J* = 7.7, 1.6 Hz, 1H), 7.23 (td, *J* = 7.6, 1.2 Hz, 1H), 7.06 (dd, *J* = 8.1, 1.1 Hz, 1H), 4.51 (dt, *J* = 13.7, 4.0 Hz, 1H), 4.16 (dd, *J* = 6.4, 3.1 Hz, 1H), 2.99 (ddd, *J* = 13.6, 12.1, 3.9 Hz, 1H), 2.26 (ddt, *J* = 12.6, 8.0, 4.4 Hz, 1H), 2.03 – 1.89 (m, 1H), 1.82 (dp, *J* = 12.5, 3.9 Hz, 1H), 1.78 – 1.51 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.2, 168.6, 136.1, 132.1, 131.1, 127.5, 125.0, 120.5, 51.2, 40.3, 23.2, 22.8, 19.2. **(S4)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.29 (s, 1H), 7.85 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.42 (td, *J* = 7.7, 1.6 Hz, 1H), 7.22 (td, *J* = 7.6, 1.1 Hz, 1H), 6.93 (dd, *J* = 8.0, 1.2 Hz, 1H), 4.42 (dq, *J* = 12.3, 1.7 Hz, 1H), 4.34 (ddd, *J* = 14.0, 3.4, 1.4 Hz, 1H), 4.09 – 4.01 (m, 1H), 3.93 (dd, *J* = 4.6, 1.4 Hz, 1H), 3.64 (dd, *J* = 12.3, 4.6 Hz, 1H), 3.57 (td, *J* = 11.9, 3.3 Hz, 1H), 3.21 (ddd, *J* = 13.9, 12.1, 4.4 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.4, 168.6, 135.5, 132.6, 131.3, 126.8, 125.6, 120.7, 65.7, 62.9, 50.8, 39.8. **(S5)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.92 (s, 1H), 7.98 (dd, *J* = 7.9, 1.6 Hz, 1H), 7.48 (td, *J* = 7.7, 1.6 Hz, 1H), 7.29 (t, *J* = 7.5 Hz, 1H), 7.01 (d, *J* = 8.0 Hz, 1H), 4.39 – 4.26 (m, 1H), 4.25 – 4.05 (m, 3H), 3.74 – 3.44 (m, 3H), 1.47 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 170.9, 170.6, 167.2, 155.1, 154.9, 135.4, 132.8, 131.8, 126.2, 126.0, 125.4, 120.9, 80.4, 52.7, 43.3, 42.5, 38.7, 37.8, 28.2.



X = CH_2 **S6**
= CH_2CH_2 **S7**
= $CH_2CH_2CH_2$ **S8**
= CH_2OCH_2 **S9**

(S)-4-(4,10-dioxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepin-9(4H)-yl)butanenitrile (S6), (S)-4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butanenitrile (S7), (S)-4-(6,12-dioxo-6,6a,7,8,9,10-hexahydrobenzo[e]pyrido[1,2-a][1,4]diazepin-5(12H)-yl)butanenitrile (S8), and **(S)-4-(6,12-dioxo-3,4,12,12a-tetrahydro-1H-benzo[e][1,4]oxazino[4,3-a][1,4]diazepin-11(6H)-yl)butanenitrile (S9). S6-S9** were synthesized from **S1-S4** by procedure 2 (61%, 85%, 76%, & 53% yield respectively) **(S6)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.57 (ddd, *J* = 8.8, 7.4, 1.7 Hz, 1H), 7.37 (td, *J* = 7.6, 1.1 Hz, 1H), 7.31 (d, *J* = 8.2 Hz, 1H), 4.55 (dd, *J* = 8.7, 4.8 Hz, 1H), 4.44 (ddd, *J* = 14.5, 8.3, 6.6 Hz, 1H), 4.21 (td, *J* = 9.5, 6.4 Hz, 1H), 4.10 (td, *J* = 9.5, 5.9 Hz, 1H), 3.77 (ddd, *J* = 13.9, 8.2, 5.7 Hz, 1H), 3.03 – 2.89 (m, 1H), 2.56 (dtd, *J* = 11.9, 9.1, 6.4 Hz, 1H), 2.41 – 2.20 (m, 2H), 2.05 – 1.92 (m, 1H), 1.92 – 1.80 (m, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 170.4, 165.3, 138.8, 132.5, 130.0, 128.6, 126.7, 123.4, 118.6, 60.3, 47.5, 46.8, 23.9, 18.2, 14.8. **(S7)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.86 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.49 (ddd, *J* = 8.2, 7.4, 1.7 Hz, 1H), 7.28 (td, *J* = 7.5, 1.0 Hz, 1H), 7.21 (dd, *J* = 8.3, 1.1 Hz, 1H), 4.30 (ddd, *J* = 14.0, 8.4, 6.5 Hz, 1H), 3.98

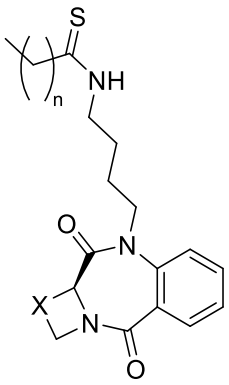
(dd, *J* = 7.8, 2.3 Hz, 1H), 3.79 – 3.65 (m, 2H), 3.55 – 3.43 (m, 1H), 2.64 (dddd, *J* = 9.6, 5.0, 4.0, 1.9 Hz, 1H), 2.31 – 2.14 (m, 2H), 2.07 (tddt, *J* = 11.1, 8.1, 5.6, 1.9 Hz, 1H), 2.01 – 1.86 (m, 3H), 1.86 – 1.74 (m, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.6, 164.8, 138.8, 132.4, 131.0, 130.5, 126.4, 122.3, 118.6, 57.3, 46.8, 46.6, 36.9 26.6, 24.1, 23.8, 14.8. **(S8)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.85 (dd, *J* = 7.7, 1.7 Hz, 1H), 7.55 (ddd, *J* = 8.2, 7.4, 1.7 Hz, 1H), 7.35 (td, *J* = 7.6, 1.1 Hz, 1H), 7.25 (dd, *J* = 8.2, 1.1 Hz, 1H), 4.53 (dddd, *J* = 13.6, 5.0, 2.8, 1.1 Hz, 1H), 4.44 – 4.32 (m, 1H), 4.19 (dd, *J* = 6.7, 2.6 Hz, 1H), 3.77 (ddd, *J* = 13.9, 8.1, 5.7 Hz, 1H), 2.90 (ddd, *J* = 13.7, 12.6, 3.9 Hz, 1H), 2.37 – 2.23 (m, 2H), 2.23 – 2.17 (m, 1H), 2.07 – 1.92 (m, 2H), 1.92 – 1.81 (m, 2H), 1.77 – 1.50 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.9, 168.3, 139.0, 132.1, 130.7, 130.4, 126.4, 121.2, 118.6, 53.5, 51.2, 45.5, 40.2, 24.1, 23.4, 23.2, 19.2, 14.7. **(S9)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.85 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.58 (ddd, *J* = 8.3, 7.4, 1.7 Hz, 1H), 7.38 (td, *J* = 7.6, 1.0 Hz, 1H), 7.30 – 7.25 (m, 1H), 4.50 – 4.29 (m, 3H), 4.18 – 4.07 (m, 1H), 4.00 (dd, *J* = 4.8, 1.4 Hz, 1H), 3.79 (ddd, *J* = 13.9, 8.0, 5.8 Hz, 1H), 3.69 (dd, *J* = 12.3, 4.7 Hz, 1H), 3.60 (td, *J* = 12.0, 3.3 Hz, 1H), 3.22 (ddd, *J* = 13.8, 12.2, 4.4 Hz, 1H), 2.41 – 2.21 (m, 2H), 2.09 – 1.83 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 168.6, 168.5, 139.0, 132.6, 130.5, 129.6, 126.7, 121.5, 118.5, 65.4, 63.4, 50.8, 45.8, 39.8, 24.1, 14.8.



X = CH_2 **S10**
= CH_2CH_2 **S11**
= $CH_2CH_2CH_2$ **S12**
= CH_2OCH_2 **S13**

tert-butyl (S)-(4-(4,10-dioxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepin-9(4H)-yl)butyl)carbamate (S10) tert-butyl (S)-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)carbamate (S11), tert-butyl (S)-(4-(6,12-dioxo-6,6a,7,8,9,10-hexahydrobenzo[e]pyrido[1,2-a][1,4]diazepin-5(12H)-yl)butyl)carbamate (12), & **tert-butyl (S)-(4-(6,12-dioxo-3,4,12,12a-tetrahydro-1H-benzo[e][1,4]oxazino[4,3-a][1,4]diazepin-11(6H)-yl)butyl)carbamate (S13). S10-S13** were synthesized from **S6-S9** by procedure 3 (37%, 60%, 52% & 48% yield respectively). **(S10)** ^{1}H NMR (400 MHz, DMSO) δ 7.72 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.61 (ddd, *J* = 8.7, 7.2, 1.8 Hz, 1H), 7.52 (dd, *J* = 8.4, 1.1 Hz, 1H), 7.36 (td, *J* = 7.5, 1.1 Hz, 1H), 6.73 (t, *J* = 5.7 Hz, 1H), 4.65 (dd, *J* = 8.6, 4.8 Hz, 1H), 4.30 – 4.18 (m, 1H), 4.10 – 3.89 (m, 2H), 3.64 (ddd, *J* = 13.6, 8.0, 5.0 Hz, 1H), 2.82 (q, *J* = 6.2 Hz, 2H), 2.68 (ddt, *J* = 11.0, 9.1, 6.0 Hz, 1H), 2.55 – 2.40 (m, 1H), 1.35 (s, 14H). ^{13}C NMR (101 MHz, DMSO) δ 170.6, 165.1, 139.4, 132.7, 129.5, 128.9, 126.4, 124.8, 77.8, 60.3 47.7, 46.6, 28.7, 27.0, 25.1, 18.3. **(S11)**^{1}H NMR (400 MHz, $CDCl_3$) δ 7.84 (dd, *J* = 7.9, 1.7 Hz, 1H), 7.48 – 7.41 (m, 1H), 7.25 (dd, *J* =

7.6, 1.0 Hz, 1H), 7.22 (d, *J* = 4.4 Hz, 1H), 4.52 (s, 1H), 4.16 (ddd, *J* = 13.8, 8.9, 6.5 Hz, 1H), 3.96 (dd, *J* = 7.8, 2.3 Hz, 1H), 3.73 (ddd, *J* = 11.6, 8.0, 2.8 Hz, 1H), 3.58 (ddd, *J* = 13.8, 8.5, 5.0 Hz, 1H), 3.49 (ddd, *J* = 11.6, 9.4, 6.6 Hz, 1H), 2.98 (q, *J* = 6.6 Hz, 2H), 2.64 (ddd, *J* = 11.5, 7.8, 2.9 Hz, 1H), 2.06 (dtd, *J* = 20.1, 9.0, 4.5 Hz, 1H), 1.97 – 1.85 (m, 2H), 1.76 – 1.67 (m, 1H), 1.54 (tdd, *J* = 11.5, 7.6, 3.1 Hz, 1H), 1.34 (s, 12H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.3, 165.1, 156.0, 139.3, 132.1, 131.1, 130.3, 126.0, 122.5 57.4, 47.9, 46.5, 28.4, 27.2, 26.6, 25.2, 23.9. **(S12)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.83 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.51 (ddd, *J* = 8.9, 7.4, 1.7 Hz, 1H), 7.31 (td, *J* = 7.6, 1.1 Hz, 1H), 7.24 (dd, *J* = 8.2, 1.0 Hz, 1H), 4.62 (s, 1H), 4.58 – 4.50 (m, 1H), 4.26 (ddd, *J* = 13.8, 8.3, 6.9 Hz, 1H), 4.20 – 4.15 (m, 1H), 3.62 (ddd, *J* = 13.5, 8.0, 5.0 Hz, 1H), 3.05 (q, *J* = 6.5 Hz, 2H), 2.90 (td, *J* = 13.2, 3.8 Hz, 1H), 2.19 (ddq, *J* = 14.0, 4.3, 2.7 Hz, 1H), 2.09 – 1.96 (m, H), 1.91 – 1.77 (m, 2H), 1.76 – 1.30 (m, 19H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.6, 168.6, 156.0, 139.4, 131.8, 130.8, 130.1, 126.0, 121.5, 79.0, 51.2, 46.4, 40.3, 39.8, 28.4, 27.0, 25.2, 23.5, 23.3, 19.3. **(S13)** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.74 (dt, *J* = 7.8, 2.0 Hz, 1H), 7.50 – 7.42 (m, 1H), 7.26 (t, *J* = 7.6 Hz, 1H), 7.18 (d, *J* = 8.2 Hz, 1H), 4.54 (s, 1H), 4.34 (d, *J* = 12.2 Hz, 1H), 4.24 (d, *J* = 14.0 Hz, 2H), 4.04 (ddd, *J* = 9.9, 4.9, 2.5 Hz, 1H), 3.90 (d, *J* = 4.6 Hz, 1H), 3.65 – 3.46 (m, 3H), 3.21 – 3.08 (m, 1H), 2.96 (q, *J* = 6.6 Hz, 2H), 1.53 (dqd, *J* = 13.9, 8.8, 6.2 Hz, 1H), 1.46 – 1.23 (m, 13H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 168.7, 168.3, 156.0, 139.3, 132.2, 130.1, 129.9, 126.3, 121.8, 79.0, 65.4, 63.4, 50.9, 46.5, 39.7, 28.4, 27.0, 25.0.



X = CH_2CH_2		n = 14	
n = 6	**KW6127**	X = CH_2	**KW6153**
8	**KW6128**	= $CH_2CH_2CH_2$	**KW6138**
10	**KW6120**	= CH_2OCH_2	**KW6140**
12	**KW6115**		
14	**KW6118**		
16	**KW6133**		

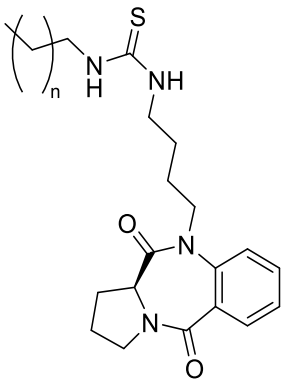
(S)-N-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)octanethioamide (KW6127), (S)-N-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)decanethioamide (KW6128), (S)-N-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)dodecanethioamide (KW6120), (S)-N-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)tetradecanethioamide (KW6115), (S)-N-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)hexadecanethioamide (KW6118), (S)-N-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)octadecanethioamide (KW6133), (S)-N-(4-(4,10-dioxo-1,2,10,10a-tetrahydroazeto[1,2-a]benzo[e][1,4]diazepin-9(4H)-yl)butyl)hexadecanethioamide (KW6153), (S)-N-(4-(6,12-dioxo-6,6a,7,8,9,10-hexahydrobenzo[e]pyrido[1,2-

a][1,4]diazepin-5(12H)-yl)butyl)hexadecanethioamide (KW6138), & (S)-N-(4-(6,12-dioxo-3,4,12,12a-tetrahydro-1H-benzo[e][1,4]oxazino[4,3-a][1,4]diazepin-11(6H)-yl)butyl)hexadecanethioamide (KW6140). The compounds were synthesized from the corresponding Boc amino benzodiazepinediones (**S10-S13**) and dithioates by procedure 4 (59%, 30%, 52%, 41%,11%, 29%, 15%, 53%, & 48% yield respectively). **KW6127** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.95 – 7.85 (m, 1H), 7.80 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.47 (ddd, *J* = 8.9, 7.4, 1.7 Hz, 1H), 7.27 (dd, *J* = 7.6, 1.1 Hz, 1H), 7.23 – 7.19 (m, 1H), 4.24 (dt, *J* = 14.5, 7.7 Hz, 1H), 4.01 (dd, *J* = 7.4, 2.1 Hz, 1H), 3.72 (ddd, *J* = 12.0, 8.1, 2.9 Hz, 1H), 3.60 (dddd, *J* = 18.0, 10.7, 7.5, 5.4 Hz, 2H), 3.53 – 3.41 (m, 2H), 2.70 – 2.59 (m, 1H), 2.52 (tt, *J* = 8.3, 6.3 Hz, 2H), 2.18 – 2.01 (m, 1H), 2.01 – 1.89 (m, 3H), 1.72 – 1.44 (m, 4H), 1.37 (ddt, *J* = 14.6, 11.4, 7.8 Hz, 2H), 1.20 (dt, *J* = 11.3, 4.7 Hz, 10H), 0.86 – 0.74 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.1, 169.5, 165.5, 138.8, 132.3, 131.4, 129.9, 126.3, 123.0, 57.6, 47.2, 47.0, 46.5, 44.9, 31.7, 29.5, 29.0, 29.0, 26.5, 25.2, 24.5, 23.9, 22.61, 14.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{24}H_{36}N_3O_2S$ 430.25282, obs. 430.25257. **KW6128** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.82 (ddd, *J* = 13.5, 6.6, 2.6 Hz, 2H), 7.47 (td, *J* = 7.8, 1.7 Hz, 1H), 7.26 (d, *J* = 7.6, 1H), 7.21 (dd, *J* = 8.2, 1.1 Hz, 1H), 4.25 (ddd, *J* = 14.5, 8.2, 6.5 Hz, 1H), 4.01 (dd, *J* = 7.7, 2.2 Hz, 1H), 3.78 – 3.69 (m, 1H), 3.59 (tdd, *J* = 13.8, 7.3, 3.1 Hz, 2H), 3.54 – 3.41 (m, 2H), 2.65 (tdd, *J* = 9.5, 3.4, 2.0 Hz, 1H), 2.52 (ddd, *J* = 8.8, 6.9, 2.2 Hz, 2H), 2.16 – 2.01 (m, 1H), 2.01 – 1.88 (m, 2H), 1.73 – 1.44 (m, 4H), 1.44 – 1.28 (m, 2H), 1.28 – 1.12 (m, 13H), 0.80 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.2, 169.6, 165.5, 138.8, 132.3, 131.5, 129.9, 126.3, 123.0, 57.6, 47.2, 47.0, 46.5, 44.9, 31.9, 29.5, 29.5, 29.4, 29.3, 29.0, 26.5, 25.1, 24.5, 23.9, 22.7, 14.1. HRMS (ESI) calc. for $[M+H]^+$

$C_{26}H_{40}N_3O_2S$ 458.28412, obs. 458.28370. **KW6120** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.82 (ddd, *J* = 9.5, 5.8, 1.9 Hz, 2H), 7.47 (ddd, *J* = 8.2, 7.4, 1.7 Hz, 1H), 7.29 – 7.19 (m, 2H), 4.25 (ddd, *J* = 14.5, 8.2, 6.6 Hz, 1H), 4.01 (dd, *J* = 7.6, 2.1 Hz, 1H), 3.76 – 3.68 (m, 1H), 3.68 – 3.54 (m, 2H), 3.54 – 3.41 (m, 2H), 2.65 (tdd, *J* = 9.6, 3.4, 2.1 Hz, 1H), 2.52 (tt, *J* = 8.6, 6.4 Hz, 2H), 2.08 (dddd, *J* = 16.3, 11.0, 7.1, 2.3 Hz, 1H), 2.01 – 1.88 (m, 3H), 1.72 – 1.44 (m, 4H), 1.44 – 1.29 (m, 2H), 1.20 (dd, *J* = 9.4, 2.2 Hz, 18H), 0.86 – 0.75 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.17, 169.6, 165.5, 138.8, 132.3, 131.5, 129.9, 126.3, 123.0, 57.6, 47.2, 47.0, 46.5, 44.9, 31.9, 29.6, 29.6, 29.5, 29.4, 29.3, 29.0, 26.5, 25.1, 24.5, 23.8, 22.7, 14.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{28}H_{44}N_3O_2S$ 486.31542, obs. 486.31528. **KW6115** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90 – 7.83 (m, 1H), 7.80 (dd, *J* = 7.7, 1.7 Hz, 1H), 7.47 (td, *J* = 7.8, 1.7 Hz, 1H), 7.26 (td, *J* = 7.5, 1.1 Hz, 1H), 7.21 (dd, *J* = 8.3, 1.1 Hz, 1H), 4.24 (ddd, *J* = 14.5, 8.2, 6.6 Hz, 1H), 4.02 – 3.98 (m, 1H), 3.76 – 3.68 (m, 1H), 3.67 – 3.53 (m, 2H), 3.48 (dtd, *J* = 17.4, 8.7, 6.1 Hz, 2H), 2.65 (dddd, *J* = 11.6, 9.4, 3.3, 2.1 Hz, 1H), 2.52 (ddd, *J* = 8.7, 6.8, 2.2 Hz, 2H), 2.08 (dddd, *J* = 15.7, 7.4, 4.5, 2.2 Hz, 1H), 2.01 – 1.88 (m, 2H), 1.72 – 1.44 (m, 4H), 1.42 – 1.29 (m, 2H), 1.28 – 1.11 (m, 22H), 0.86 – 0.75 (m, 3H). HRMS (ESI) calc. for $[M+H]^+$ $C_{30}H_{48}N_3O_2S$ 514.34672, obs. 514.34662. **KW6118** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.76 (d, *J* = 5.8 Hz, 1H), 7.47 (td, *J* = 7.7, 1.7 Hz, 1H), 7.26 (td, *J* = 7.5, 1.1 Hz, 1H), 7.23 – 7.19 (m, 1H), 4.26 (ddd, *J* = 14.5, 8.3, 6.3 Hz, 1H), 4.01 (dd, *J* = 7.7, 2.2 Hz, 1H), 3.77 – 3.69 (m, 1H), 3.69 – 3.52 (m, 2H), 3.52 – 3.41 (m, 2H), 2.65 (dddd, *J* = 11.7, 9.6, 3.4, 2.0 Hz, 1H), 2.52 (ddd, *J* = 8.9, 6.8, 2.5 Hz, 2H), 2.08 (dddd, *J* = 21.4, 10.9, 6.2, 2.2 Hz, 1H), 2.01 – 1.88 (m, 2H), 1.72 – 1.44 (m, 4H), 1.44 – 1.28 (m, 2H), 1.28 – 1.11 (m, 26H), 0.87 – 0.77 (m, 3H). ^{13}C NMR

(101 MHz, $CDCl_3$) δ 206.2, 169.6, 165.6, 138.8, 132.3, 131.6, 129.9, 126.4, 123.0, 57.6, 47.1, 47.1, 46.5, 44.9, 31.9, 29.7, 29.7, 29.7, 29.6, 29.5, 29.4, 29.4, 29.1, 26.5, 25.1, 24.4, 23.9, 22.7, 14.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{32}H_{52}N_3O_2S$ 542.37802, obs. 542.37778. **KW6133** ^1H NMR (400 MHz, $CDCl_3$) δ 7.81 (dd, *J* = 7.8, 1.7 Hz, 2H), 7.47 (td, *J* = 7.8, 1.7 Hz, 1H), 7.26 (t, *J* = 7.5 Hz, 1H), 7.21 (d, *J* = 8.2 Hz, 1H), 4.26 (ddd, *J* = 14.4, 8.2, 6.4 Hz, 1H), 4.01 (dd, *J* = 7.8, 2.3 Hz, 1H), 3.72 (ddd, *J* = 11.6, 8.1, 2.8 Hz, 1H), 3.68 – 3.53 (m, 2H), 3.53 – 3.41 (m, 2H), 2.71 – 2.60 (m, 1H), 2.51 (ddt, *J* = 13.1, 8.6, 4.4 Hz, 2H), 2.17 – 2.02 (m, 1H), 2.02 – 1.88 (m, 2H), 1.71 – 1.44 (m, 4H), 1.36 (dq, *J* = 14.2, 8.1 Hz, 2H), 1.18 (s, 32H), 0.81 (t, *J* = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.2, 169.6, 165.5, 138.8, 132.3, 131.5, 129.9, 126.3, 123.0, 57.6, 47.1, 47.0, 46.5, 44.9, 31.9, 29.7, 29.7, 29.7, 29.6, 29.6, 29.4, 29.4, 29.1, 26.5, 25.1, 24.5, 23.9, 22.7, 14.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{34}H_{56}N_3O_2S$ 570.40932, obs. 570.40908. **KW6153** ^1H NMR (400 MHz, $CDCl_3$) δ 7.78 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.66 (d, *J* = 5.7 Hz, 1H), 7.48 (td, *J* = 7.8, 1.7 Hz, 1H), 7.28 (td, *J* = 7.6, 1.1 Hz, 1H), 7.23 (d, *J* = 8.2 Hz, 1H), 4.48 (dd, *J* = 8.6, 4.5 Hz, 1H), 4.33 (ddd, *J* = 14.2, 8.1, 6.4 Hz, 1H), 4.16 (td, *J* = 9.4, 6.6 Hz, 1H), 4.02 (td, *J* = 9.6, 5.5 Hz, 1H), 3.69 – 3.43 (m, 3H), 2.89 (ddt, *J* = 13.4, 9.5, 4.7 Hz, 1H), 2.51 (dtd, *J* = 9.1, 6.8, 2.6 Hz, 3H), 1.74 – 1.47 (m, 4H), 1.47 – 1.30 (m, 2H), 1.30 – 1.11 (m, 23H), 0.81 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.2, 170.3, 166.0, 138.7, 132.5, 129.5, 129.2, 126.8, 124.1, 60.7, 48.1, 47.1, 46.9, 45.0, 31.9, 29.7, 29.7, 29.6, 29.5, 29.4, 29.4, 29.0, 25.1, 24.6, 22.7, 18.2, 14.1. HRMS (ESI) calc. for $[M+H]^+$ $C_{31}H_{50}N_3O_2S$ 528.36237, obs. 528.36234. **KW6138** ^1H NMR (400 MHz, $CDCl_3$) δ 7.87 (d, *J* = 5.6 Hz, 1H), 7.71 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.45 (td, *J* = 7.8, 1.7 Hz, 1H), 7.29 – 7.23 (m, 1H), 7.17 (d, *J* = 8.2 Hz, 1H), 4.45 (dt, *J* =

13.6, 4.2 Hz, 1H), 4.31 (ddd, *J* = 14.2, 9.3, 5.3 Hz, 1H), 4.15 (dd, *J* = 6.8, 2.3 Hz, 1H), 3.65 (ddt, *J* = 14.0, 7.9, 5.6 Hz, 1H), 3.52 (ddd, *J* = 14.1, 6.1, 3.9 Hz, 1H), 3.38 (dq, *J* = 13.6, 5.3 Hz, 1H), 2.81 (td, *J* = 13.2, 3.8 Hz, 1H), 2.55 – 2.41 (m, 2H), 2.17 – 2.08 (m, 1H), 1.96 (dtt, *J* = 13.9, 9.5, 4.8 Hz, 2H), 1.80 (dt, *J* = 13.1, 4.0 Hz, 1H), 1.71 – 1.41 (m, 7H), 1.19 (s, 27H), 0.81 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.3, 169.8, 169.3, 138.8, 132.1, 131.1, 129.6, 126.4, 122.0, 51.3, 46.9, 45.3, 44.9, 40.6, 31.9, 29.7, 29.7, 29.7, 29.6, 29.4, 29.4, 29.1, 25.0, 24.2, 23.5, 23.2, 22.7, 19.2, 14.1. (ESI) calc. for $[M+H]^+$ $C_{33}H_{54}N_3O_2S$ 556.39367, obs. 556.39362. **KW6140** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.72 (dd, *J* = 7.7, 1.7 Hz, 2H), 7.49 (td, *J* = 7.8, 1.7 Hz, 1H), 7.28 (td, *J* = 7.5, 1.0 Hz, 1H), 7.20 (d, *J* = 8.2 Hz, 1H), 4.39 – 4.28 (m, 2H), 4.27 – 4.19 (m, 1H), 4.05 (qd, *J* = 6.0, 2.9 Hz, 1H), 3.96 – 3.92 (m, 1H), 3.69 – 3.38 (m, 5H), 3.12 (ddd, *J* = 13.8, 12.3, 4.4 Hz, 1H), 2.56 – 2.42 (m, 2H), 1.97 (s, 1H), 1.71 – 1.43 (m, 4H), 1.19 (d, *J* = 7.0 Hz, 27H), 0.81 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.2, 169.3, 168.5, 138.9, 132.5, 130.1, 129.8, 126.6, 122.2, 65.3, 63.4, 51.0, 47.0, 45.8, 44.8, 40.0, 31.9, 29.7, 29.7, 29.7, 29.6, 29.5, 29.4, 29.4, 29.0, 25.0, 24.4, 22.7, 14.1. (ESI) calc. for $[M+H]^+$ $C_{32}H_{52}N_3O_3S$ 558.37294, obs. 558.37287.

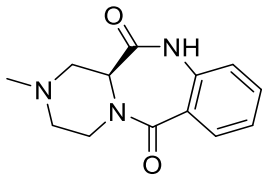


n = 4 **NH-C1-6**
= 6 **NH-C1-8**
= 8 **NH-C1-10**
= 12 **KW6158**

(S)-1-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)-3-hexylthiourea (NH-C1-6), (S)-1-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)-3-octylthiourea (NH-C1-8), (S)-1-decyl-3-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)thiourea (NH-C1-10) & (S)-1-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)-3-tetradecylthiourea (KW6158). These compounds were synthesized from **S11** and the corresponding isothiocyanates by procedure 5 (72% & 59% yield respectively). **NH-C1-6,** ^{1}H NMR (500 MHz, MeOD) δ 7.84 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.65 (ddd, *J* = 8.8, 7.4, 1.7 Hz, 1H), 7.53 (dd, *J* = 8.3, 1.1 Hz, 1H), 7.40 (td, *J* = 7.6, 1.1 Hz, 1H), 4.38 (dt, *J* = 14.2, 7.2 Hz, 1H), 4.20 – 4.14 (m, 1H), 3.77 (dtd, *J* = 11.8, 8.0, 3.5 Hz, 2H), 3.54 (qd, *J* = 8.3, 3.3 Hz, 1H), 3.38 (d, *J* = 14.3 Hz, 2H), 2.70 – 2.61 (m, 1H), 2.19 – 2.01 (m, 3H), 1.57 (tdd, *J* = 14.8, 9.0, 4.8 Hz, 3H), 1.52 – 1.40 (m, 3H), 1.40 – 1.28 (m, 6H), 0.97 – 0.88 (m, 3H). ^{13}C NMR (126 MHz, MeOD) δ 169.8, 166.0, 139.3, 132.4, 130.7, 129.3, 125.8, 123.2, 57.6, 46.3, 31.3, 26.3, 26.1, 24.7, 23.4,

22.3, 13.0. (ESI) calc. for $[M+H]^+$ $C_{23}H_{35}N_4O_2S$ 431.24807, obs. 431.24681 **NH-C1-8,** ^{1}H NMR (500 MHz, MeOD) δ 7.84 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.65 (ddd, *J* = 8.7, 7.3, 1.7 Hz, 1H), 7.53 (dd, *J* = 8.3, 1.0 Hz, 1H), 7.39 (td, *J* = 7.6, 1.1 Hz, 1H), 4.37 (dt, *J* = 14.2, 7.2 Hz, 1H), 4.20 – 4.14 (m, 1H), 3.77 (ddt, *J* = 14.6, 7.8, 4.2 Hz, 2H), 3.54 (qd, *J* = 8.3, 3.3 Hz, 1H), 3.39 (s, 3H), 2.68 – 2.55 (m, 1H), 2.18 – 1.98 (m, 3H), 1.57 (ddd, *J* = 15.4, 7.7, 4.1 Hz, 3H), 1.51 – 1.39 (m, 3H), 1.34 (q, *J* = 5.3 Hz, 10H), 0.98 – 0.85 (m, 3H). ^{13}C NMR (126 MHz, MeOD) δ 169.7, 165.9, 139.3, 132.4, 130.7, 129.3, 125.8, 123.2, 57.6 31.6, 29.1, 29.0, 26.6, 26.1, 24.7, 23.4, 22.3, 13.1. (ESI) calc. for $[M+H]^+$ $C_{25}H_{39}N_4O_2S$ 459.27937, obs. 459.27795. **NH-C1-10** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.84 (d, *J* = 7.7 Hz, 1H), 7.55 (t, *J* = 7.8 Hz, 1H), 7.35 – 7.27 (m, 2H), 6.32 (s, 1H), 6.03 (d, *J* = 7.5 Hz, 1H), 4.35 (dt, *J* = 14.6, 6.8 Hz, 1H), 4.07 (d, *J* = 6.9 Hz, 1H), 3.78 (ddd, *J* = 11.7, 8.4, 3.0 Hz, 1H), 3.71 – 3.21 (m, 6H), 2.71 (dq, *J* = 10.1, 2.8 Hz, 1H), 2.25 – 2.08 (m, 1H), 2.01 (h, *J* = 7.0 Hz, 2H), 1.67 – 1.47 (m, 4H), 1.29 (d, *J* = 18.9 Hz, 16H), 0.88 (t, *J* = 6.6 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 182.2, 169.5, 165.7, 138.8, 132.5, 131.4, 129.7, 126.3, 123.1, 57.6, 47.1, 46.5, 44.4, 43.5, 31.9, 29.6, 29.4, 29.3, 29.1, 27.0, 26.5, 26.1, 25.0, 23.8, 22.7, 14.1. (ESI) calc. for $[M+H]^+$ $C_{27}H_{43}N_4O_2S$ 487.31067, obs. 487.31040. **KW6158** ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76 (dd, *J* = 8.1, 1.7 Hz, 1H), 7.47 (td, *J* = 7.8, 1.7 Hz, 1H), 7.24 (dd, *J* = 8.2, 5.9 Hz, 2H), 6.42 (s, 1H), 6.24 (t, *J* = 5.5 Hz, 1H), 4.22 (dt, *J* = 14.4, 7.2 Hz, 1H), 4.00 – 3.96 (m, 1H), 3.70 (ddd, *J* = 11.5, 7.9, 2.8 Hz, 1H), 3.61 – 3.43 (m, 2H), 3.42 – 3.16 (m, 4H), 2.63 (ddt, *J* = 8.9, 6.5, 2.4 Hz, 1H), 2.07 (dtd, *J* = 15.1, 7.6, 3.9 Hz, 1H), 1.96 (s, 3H), 1.44 (dtd, *J* = 22.4, 12.1, 6.1 Hz, 5H), 1.31 – 1.10 (m, 25H), 0.80 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 182.1, 169.3, 169.3, 165.6, 165.3, 138.9, 138.9, 132.4, 132.2, 131.1, 131.0,

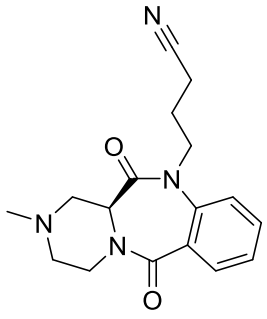
130.0, 129.7, 126.1, 126.1, 123.0, 122.7, 57.5, 57.4, 47.3, 47.1, 46.5, 44.3, 43.5, 38.9, 31.8, 29.6, 29.6, 29.6, 29.6, 29.5, 29.3, 29.3, 29.1, 26.9, 26.5, 26.2, 25.4, 25.0, 24.8, 23.8, 23.8, 22.6, 14.1. (ESI) calc. for $[M+H]^+$ $C_{31}H_{51}N_4O_2S$ 543.37327, obs. 543.37336.



S14

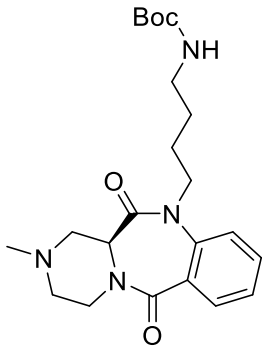
(S)-2-methyl-1,3,4,12a-tetrahydrobenzo[e]pyrazino[1,2-a][1,4]diazepine-6,12(2H,11H)-dione (S14). S5 (0.44 g, 1.3 mmol) was dissolved in DCM (5 mL) and TFA (1 mL). The solution was stirred at room temperature for 2 h then concentrated under reduced pressure. The resulting material was dissolved in acetonitrile (13 mL). K_2CO_3 (0.56 g, 4.1 mmol) and methyl iodide (0.11 mL, 1.8 mmol) were added, and the mixture was stirred at 60 ^oC for 16 h. The mixture was cooled to room temperature and solvent was removed under reduced pressure. Crude solid was diluted with water (20 mL) and extracted with EtOAc (20 mL) three times. Organic layers were combined, washed with brine, dried over sodium sulfate, and concentrated under reduced pressure. The second crude was purified by flash chromatography (DCM → 1:9 DCM : MeOH) to give compound **S14** (0.18 g, 0.73 mmol, 55% yield). ^1H NMR (400 MHz, DMSO) δ 10.34 (s, 1H), 7.70 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.49 (td, *J* = 7.7, 1.6 Hz, 1H), 7.21 (td, *J* = 7.5, 1.1 Hz, 1H), 7.12 – 7.05 (m, 1H), 4.28 (ddd, *J* = 13.3, 3.6, 1.9 Hz, 1H), 4.18 – 4.08 (m, 1H), 3.22 – 3.05 (m, 1H), 2.92 (td, *J* = 12.7, 3.9 Hz, 1H), 2.80 (ddt, *J* = 11.5, 4.1, 2.0 Hz, 1H), 2.23 (s, 3H), 2.06 (dd, *J* = 12.1, 5.5 Hz, 1H), 1.96 (td, *J* = 11.8, 3.6 Hz,

1H). ^{13}C NMR (101 MHz, DMSO) δ 171.0, 168.1, 137.5, 132.4, 130.8, 127.1, 124.4, 120.9, 53.2, 52.1, 51.2, 46.3.



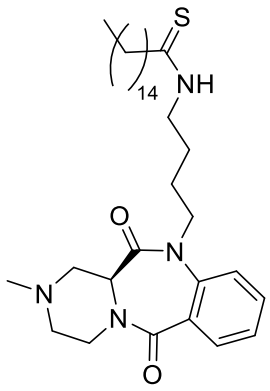
S15

(S)-4-(2-methyl-6,12-dioxo-1,2,3,4,12,12a-hexahydrobenzo[e]pyrazino[1,2-a][1,4]diazepin-11(6H)-yl)butanenitrile (S15). S15 was synthesized from **S14** by procedure 2 (47% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.88 – 7.80 (m, 1H), 7.60 – 7.52 (m, 1H), 7.39 – 7.31 (m, 1H), 7.26 (dd, *J* = 8.2, 3.0 Hz, 1H), 4.52 – 4.32 (m, 2H), 4.13 (t, *J* = 4.3 Hz, 1H), 3.75 (ddd, *J* = 13.8, 8.5, 5.9 Hz, 1H), 3.27 (dd, *J* = 12.2, 3.5 Hz, 1H), 3.16 (td, *J* = 12.9, 4.1 Hz, 1H), 2.99 – 2.90 (m, 1H), 2.41 (d, *J* = 3.7 Hz, 3H), 2.34 – 2.17 (m, 3H), 2.11 (td, *J* = 12.0, 3.7 Hz, 1H), 1.94 (ddt, *J* = 42.2, 14.0, 6.9 Hz, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.2, 168.1, 139.1, 132.3, 130.2, 130.0, 126.4, 121.3, 118.5, 53.1, 52.3, 51.3, 46.4, 45.9, 40.8, 24.1, 14.7.



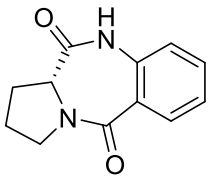
S16

tert-butyl (S)-(4-(2-methyl-6,12-dioxo-1,2,3,4,12,12a-hexahydrobenzo[e]pyrazino[1,2-a][1,4]diazepin-11(6H)-yl)butyl)carbamate (S16). **S16** was synthesized from **S15** by procedure 3 (27% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.81 (dd, *J* = 7.7, 1.7 Hz, 1H), 7.51 (ddd, *J* = 8.7, 7.6, 1.7 Hz, 1H), 7.31 (t, *J* = 7.5 Hz, 1H), 7.23 (d, *J* = 8.2 Hz, 1H), 4.55 (s, 1H), 4.44 (ddd, *J* = 13.6, 3.7, 2.0 Hz, 1H), 4.31 (dt, *J* = 13.6, 7.6 Hz, 1H), 4.12 – 4.07 (m, 1H), 3.61 (ddd, *J* = 13.5, 7.8, 5.0 Hz, 1H), 3.27 (dt, *J* = 12.0, 1.7 Hz, 1H), 3.17 (td, *J* = 12.9, 4.0 Hz, 1H), 3.01 (p, *J* = 7.1 Hz, 2H), 2.93 (ddt, *J* = 11.6, 4.0, 2.0 Hz, 1H), 2.40 (s, 3H), 2.17 (dd, *J* = 12.0, 5.5 Hz, 1H), 2.10 (td, *J* = 11.9, 3.6 Hz, 1H), 1.67 – 1.54 (m, 1H), 1.40 (s, 13H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.0, 168.45, 156.0, 139.5, 132.0, 130.3, 130.0, 126.1, 121.7, 79.0 53.2, 52.5, 51.5, 46.6, 46.4, 40.8, 39.7, 28.4, 27.0, 25.0.



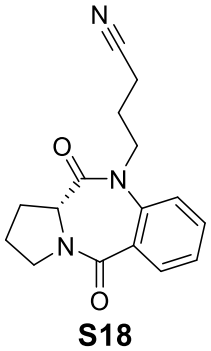
KW6173

(S)-N-(4-(2-methyl-6,12-dioxo-1,2,3,4,12,12a-hexahydrobenzo[e]pyrazino[1,2-a][1,4]diazepin-11(6H)-yl)butyl)hexadecanethioamide (KW6173). KW6173 was synthesized from **S16** by procedure 4 (17% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.78 (dd, *J* = 7.7, 1.6 Hz, 1H), 7.70 (t, *J* = 5.7 Hz, 1H), 7.54 (td, *J* = 7.8, 1.7 Hz, 1H), 7.33 (td, *J* = 7.6, 1.0 Hz, 1H), 7.25 (d, *J* = 8.2 Hz, 1H), 4.50 – 4.31 (m, 2H), 4.14 (dd, *J* = 6.2, 4.6 Hz, 1H), 3.65 (dddd, *J* = 29.8, 14.0, 7.1, 3.6 Hz, 2H), 3.49 (dq, *J* = 13.8, 5.2 Hz, 1H), 3.33 – 3.24 (m, 1H), 3.18 (td, *J* = 12.9, 4.0 Hz, 1H), 2.95 (ddt, *J* = 11.7, 4.1, 2.2 Hz, 1H), 2.61 – 2.47 (m, 2H), 2.41 (d, *J* = 4.1 Hz, 3H), 2.25 – 2.17 (m, 1H), 2.17 – 2.06 (m, 1H), 1.62 (dddd, *J* = 36.3, 19.8, 8.1, 5.8 Hz, 4H), 1.48 – 1.17 (m, 27H), 0.88 (t, *J* = 6.8 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.3, 169.1, 139.0, 132.3, 130.5, 129.6, 126.5, 122.1, 53.0, 52.4, 51.6, 47.0, 46.4, 45.67, 44.7, 41.1, 31.9, 29.7, 29.7, 29.7, 29.6, 29.5, 29.4, 29.4, 29.1, 24.8, 24.4, 22.7, 14.1. (ESI) calc. for $[M+H]^+$ $C_{33}H_{55}N_4O_2S$ 571.40457, obs. 571.40441.



S17

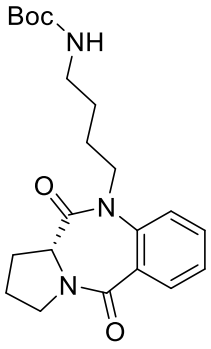
(R)-1,2,3,11a-tetrahydro-5H-benzo[e]pyrrolo[1,2-a][1,4]diazepine-5,11(10H)-dione (S17). S17 was synthesized from D-proline by procedure 1 (62% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.11 (d, *J* = 10.7 Hz, 1H), 8.00 (d, *J* = 7.8 Hz, 1H), 7.47 (td, *J* = 7.8, 1.7 Hz, 1H), 7.26 (t, *J* = 7.7 Hz, 1H), 7.06 (d, *J* = 8.0 Hz, 1H), 4.08 (d, *J* = 6.3 Hz, 1H), 3.82 (ddt, *J* = 12.1, 5.6, 2.9 Hz, 1H), 3.69 – 3.53 (m, 1H), 2.87 – 2.67 (m, 1H), 2.13 – 1.89 (m, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 171.5, 171.5, 165.5, 135.4, 132.4, 131.1, 127.2, 125.0, 121.1, 56.7, 47.4, 26.3, 23.6.



S18

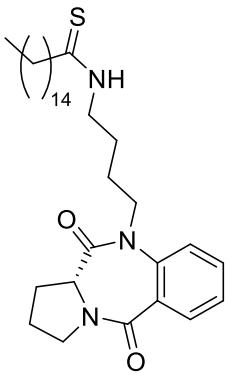
(R)-4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butanenitrile (S18). S18 was synthesized from **S17** by procedure 2 (99% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.94 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.62 – 7.52 (m, 1H), 7.35 (td, *J* = 7.6, 1.1 Hz, 1H), 7.28 (dd, *J* = 8.8, 1.6 Hz, 1H), 4.37 (ddd, *J* = 14.5, 8.4, 6.5 Hz, 1H), 4.12 – 3.99 (m, 1H), 3.87 – 3.70 (m, 2H), 3.64 – 3.43 (m, 1H), 2.65 (m, 1H), 2.41 – 2.20 (m, 2H), 2.20 – 2.08 (m, 1H), 2.08 – 1.94 (m, 3H), 1.94 – 1.80 (m,

1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.6, 164.9, 138.8, 132.4, 131.0, 130.5, 126.4, 122.3, 118.5, 57.3, 46.8, 46.6, 26.6, 24.1, 23.8, 14.8.



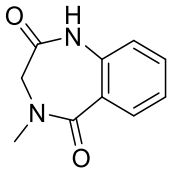
S19

tert-butyl (R)-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)carbamate (S19). S19 was synthesized from **S18** by procedure 3 (45% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.88 – 7.75 (m, 1H), 7.52 – 7.36 (m, 1H), 7.22 (t, *J* = 7.8 Hz, 2H), 4.64 (s, 1H), 4.23 – 4.07 (m, 1H), 4.02 – 3.91 (m, 1H), 3.73 (ddd, *J* = 11.7, 8.1, 2.8 Hz, 1H), 3.58 (ddd, *J* = 13.8, 8.4, 4.9 Hz, 1H), 3.53 – 3.42 (m, 1H), 2.97 (q, *J* = 6.6 Hz, 2H), 2.70 – 2.57 (m, 1H), 2.12 – 1.99 (m, 1H), 1.92 (h, *J* = 7.4 Hz, 2H), 1.54 (dtt, *J* = 12.5, 8.5, 4.5 Hz, 1H), 1.33 (s, 13H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 169.3, 165.1, 156.0, 139.2, 132.1, 131.1, 130.2, 125.9, 122.6, 79.0, 57.3, 47.8, 46.5, 39.8, 39.8, 28.4, 27.1, 26.6, 25.1, 23.8.



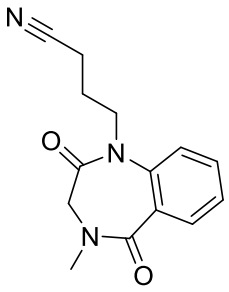
KW6137

(R)-N-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)hexadecanethioamide (KW6137). KW6137 was synthesized from **S19** by procedure 4 (34% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.80 (dd, *J* = 7.8, 1.7 Hz, 2H), 7.47 (ddd, *J* = 8.8, 7.5, 1.7 Hz, 1H), 7.26 (td, *J* = 7.6, 1.1 Hz, 1H), 7.21 (d, *J* = 8.2 Hz, 1H), 4.26 (ddd, *J* = 14.5, 8.2, 6.4 Hz, 1H), 4.01 (dd, *J* = 7.7, 2.2 Hz, 1H), 3.72 (ddd, *J* = 11.7, 8.2, 2.8 Hz, 1H), 3.67 – 3.42 (m, 4H), 2.70 – 2.60 (m, 1H), 2.51 (ddt, *J* = 13.2, 8.6, 4.5 Hz, 2H), 2.16 – 2.02 (m, 1H), 1.96 (tdd, *J* = 11.9, 7.6, 4.9 Hz, 3H), 1.71 – 1.44 (m, 4H), 1.43 – 1.28 (m, 2H), 1.27 – 1.13 (m, 25H), 0.81 (t, *J* = 6.7 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 206.2, 169.6, 165.5, 138.8, 132.3, 131.5, 129.9, 126.4, 123.0, 57.6, 47.1, 47.0, 46.5, 44.9, 31.9, 29.7, 29.7, 29.7, 29.6, 29.4, 29.4, 29.0, 26.5, 25.1, 24.5, 23.9, 22.7, 14.1. (ESI) calc. for $[M+H]^+$ $C_{32}H_{52}N_3O_2S$ 542.37802, obs. 542.37802.



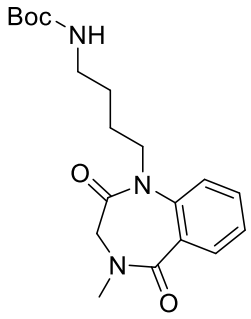
S20

4-methyl-3,4-dihydro-1H-benzo[e][1,4]diazepine-2,5-dione (S20). S20 was synthesized from methyl glycine by procedure 1 (32% yield) ^{1}H NMR (500 MHz, DMSO) δ 10.46 (s, 1H), 7.75 (dd, *J* = 7.8, 1.6 Hz, 1H), 7.50 (ddd, *J* = 8.1, 7.2, 1.6 Hz, 1H), 7.22 (ddd, *J* = 8.1, 7.3, 1.2 Hz, 1H), 7.10 (dd, *J* = 8.2, 1.2 Hz, 1H), 3.85 (s, 2H), 3.12 (s, 3H). ^{13}C NMR (126 MHz, DMSO) δ 170.3, 167.0, 137.4, 132.5, 131.3, 126.7, 124.4, 121.2, 52.6, 36.3.



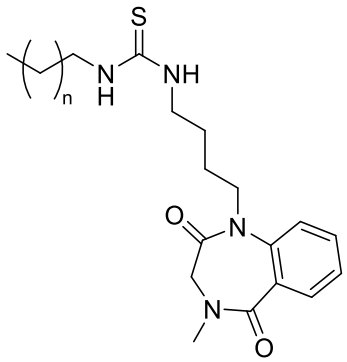
S21

4-(4-methyl-2,5-dioxo-2,3,4,5-tetrahydro-1H-benzo[e][1,4]diazepin-1-yl)butanenitrile (S21). S21 was synthesized from **S20** by procedure 2 (54% yield). ^{1}H NMR (500 MHz, MeOD) δ 7.82 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.66 (ddd, *J* = 8.8, 7.3, 1.7 Hz, 1H), 7.51 (dd, *J* = 8.2, 1.1 Hz, 1H), 7.42 (td, *J* = 7.5, 1.1 Hz, 1H), 4.44 – 4.35 (m, 1H), 4.14 (d, *J* = 14.8 Hz, 1H), 3.87 (ddd, *J* = 13.9, 8.0, 5.4 Hz, 1H), 3.74 (d, *J* = 14.8 Hz, 1H), 3.26 (s, 3H), 2.42 – 2.33 (m, 2H), 1.93 (ddt, *J* = 15.0, 12.9, 6.2 Hz, 1H), 1.88 – 1.78 (m, 1H). ^{13}C NMR (126 MHz, MeOD) δ 168.8, 167.9, 139.3, 132.4, 129.9, 129.6, 126.1, 122.0, 118.8, 52.6, 45.0, 34.6, 23.5, 13.4.



S22

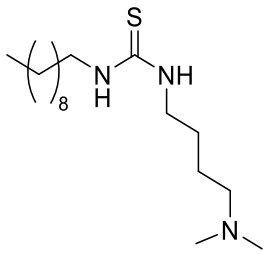
tert-butyl (4-(4-methyl-2,5-dioxo-2,3,4,5-tetrahydro-1H-benzo[e][1,4]diazepin-1-yl)butyl)carbamate (S22). S22 was synthesized from **S21** by procedure 3 (61% yield). ^{1}H NMR (500 MHz, MeOD) δ 7.80 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.64 (ddd, *J* = 8.7, 7.4, 1.7 Hz, 1H), 7.51 (d, *J* = 8.3 Hz, 1H), 7.48 – 7.33 (m, 1H), 4.34 (dt, *J* = 14.6, 7.7 Hz, 1H), 4.12 (d, *J* = 14.6 Hz, 1H), 3.80 – 3.68 (m, 2H), 3.25 (s, 3H), 2.97 (t, *J* = 6.8 Hz, 2H), 1.56 (dt, *J* = 8.2, 5.5 Hz, 1H), 1.42 (s, 10H), 1.38 – 1.27 (m, 3H). ^{13}C NMR (126 MHz, MeOD) δ 170.1, 169.5, 141.0, 133.7, 131.2, 131.1, 127.3, 123.7, 79.8, 54.1, 47.2, 40.6, 36.0, 28.8, 27.9, 26.0.



n = 4 **NH-C4-6**
= 8 **NH-C4-10**

1-hexyl-3-(4-(4-methyl-2,5-dioxo-2,3,4,5-tetrahydro-1H-benzo[e][1,4]diazepin-1-yl)butyl)thiourea (NH-C4-6) & 1-decyl-3-(4-(4-methyl-2,5-dioxo-2,3,4,5-

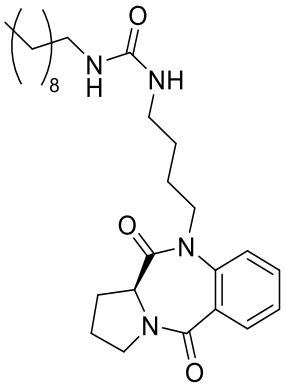
tetrahydro-1H-benzo[e][1,4]diazepin-1-yl)butyl)thiourea (NH-C4-10). NH-C4-6 & NH-C4-10 were synthesized from **S22** by procedure 5 (42% & 50% yield respectively). **NH-C4-6** ^{1}H NMR (500 MHz, MeOD) δ 7.80 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.65 (ddd, *J* = 8.8, 7.3, 1.7 Hz, 1H), 7.52 (dd, *J* = 8.3, 1.1 Hz, 1H), 7.40 (td, *J* = 7.6, 1.1 Hz, 1H), 4.36 (dt, *J* = 14.3, 7.2 Hz, 1H), 4.13 (d, *J* = 14.7 Hz, 1H), 3.81 – 3.68 (m, 2H), 3.40 (s, 3H), 3.26 (s, 3H), 1.57 (tt, *J* = 14.4, 7.4 Hz, 3H), 1.52 – 1.39 (m, 3H), 1.39 – 1.29 (m, 6H), 1.01 – 0.87 (m, 3H). ^{13}C NMR (126 MHz, MeOD) δ 168.7, 168.1, 139.5, 132.3, 129.7, 129.7, 125.9, 123.0, 52.7, 45.8, 34.6, 31.3, 26.2, 24.6, 22.3, 13.0. (ESI) calc. for $[M+H]^+$ $C_{21}H_{33}N_4O_2S$ 405.23242, obs. 405.2313 **NH-C4-10** ^{1}H NMR (500 MHz, MeOD) δ 7.80 (dd, *J* = 7.9, 1.7 Hz, 1H), 7.65 (ddd, *J* = 8.7, 7.3, 1.7 Hz, 1H), 7.52 (dd, *J* = 8.4, 1.1 Hz, 1H), 7.40 (td, *J* = 7.6, 1.1 Hz, 1H), 4.36 (dt, *J* = 14.4, 7.3 Hz, 1H), 4.13 (d, *J* = 14.7 Hz, 1H), 3.84 – 3.69 (m, 2H), 3.40 (s, 3H), 3.26 (s, 3H), 1.69 – 1.24 (m, 20H), 0.92 (t, *J* = 6.9 Hz, 3H). ^{13}C NMR (126 MHz, MeOD) δ 168.7, 168.1, 139.5, 132.3, 129.7, 129.7, 125.9, 122.3, 52.7, 45.8, 34.6, 31.7, 29.3, 29.3, 29.1, 26.6, 24.6, 22.3, 13.0. (ESI) calc. for $[M+H]^+$ $C_{25}H_{41}N_4O_2S$ 461.29502, obs. 461.29374



NH-C3-10

1-decyl-3-(4-(dimethylamino)butyl)thiourea (NH-C3-10). To a stirred solution of 4-dimethylaminobutylamine (0.150 mL, 1.08 mmol) in THF (10 mL), was added 1-

isothiocyanatodecane (0.259 g, 1.30 mmol) and DIPEA (0.377 mL, 2.17 mmol) at room temperature. After stirring for 16 h, the solvents were removed by rotary evaporation. The crude residue was re-dissolved in ethyl acetate (30 mL) and washed with water (25 mL) and brine (25 mL). The collected organic layer was dried over Na_2SO_4 and concentrated by rotary evaporation. The crude residue was further purified by silica gel column chromatography (DCM: MeOH = 8:2 → 75:25) to afford the final product **NH-C3-10** (0.176 g, 77%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.14 (s, 1H), 5.94 (s, 1H), 3.42 (d, *J* = 41.5 Hz, 4H), 2.36 (td, *J* = 6.6, 2.3 Hz, 2H), 2.27 (d, *J* = 2.5 Hz, 6H), 1.71 (h, *J* = 5.1 Hz, 2H), 1.66 – 1.56 (m, 4H), 1.41 – 1.22 (m, 10H), 0.90 (dt, *J* = 7.6, 4.4 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 181.3, 58.8, 45.3, 31.8, 29.3, 29.2, 27.0, 26.7, 24.9, 22.6, 14.1. (ESI) calc. for $[M+H]^+$ $C_{17}H_{38}N_3S$ 316.27864, obs. 316.27747.



NH-C1-10-O

(S)-1-decyl-3-(4-(5,11-dioxo-2,3,11,11a-tetrahydro-1H-benzo[e]pyrrolo[1,2-a][1,4]diazepin-10(5H)-yl)butyl)urea (NH-C1-10-O). NH-C1-10-O was synthesized from **S11** by procedure 5 using 1-isocyanatodecane instead of an isothiocyanate (28% yield) ^{1}H NMR (500 MHz, MeOD) δ 7.84 (dd, *J* = 7.8, 1.7 Hz, 1H), 7.65 (ddd, *J* = 8.8,

7.3, 1.7 Hz, 1H), 7.52 (dd, *J* = 8.3, 1.1 Hz, 1H), 7.39 (td, *J* = 7.5, 1.0 Hz, 1H), 4.37 (dt, *J* = 13.9, 7.6 Hz, 1H), 4.20 – 4.14 (m, 1H), 3.75 (dtt, *J* = 11.7, 7.9, 4.0 Hz, 2H), 3.54 (dddd, *J* = 12.0, 6.8, 5.1, 2.1 Hz, 1H), 3.05 (dt, *J* = 18.3, 6.8 Hz, 4H), 2.69 – 2.60 (m, 1H), 2.16 – 2.00 (m, 3H), 1.56 (dddt, *J* = 11.9, 9.0, 7.2, 5.0 Hz, 1H), 1.48 – 1.23 (m, 20H). ^{13}C NMR (126 MHz, MeOD) δ 169.7, 166.6, 159.8, 139.3, 132.4, 130.7, 129.3, 125.8, 123.2, 57.5, 47.0, 46.3, 39.6, 38.9, 31.7, 29.9, 29.4, 29.3, 29.1, 29.1, 27.0, 26.6, 26.1, 24.7, 23.4, 22.3, 13.0. (ESI) calc. for $[M+H]^+$ $C_{27}H_{43}N_4O_3$ 471.33352, obs. 471.3323.

References

1. Hong, J. Y.; Lin, H., Sirtuin Modulators in Cellular and Animal Models of Human Diseases. *Front Pharmacol* **2021,** *12*, 735044-735044.
2. Bheda, P.; Jing, H.; Wolberger, C.; Lin, H., The Substrate Specificity of Sirtuins. *Annu Rev Biochem* **2016,** *85*, 405-29.
3. Chen, Y.; Sprung, R.; Tang, Y.; Ball, H.; Sangras, B.; Kim, S. C.; Falck, J. R.; Peng, J.; Gu, W.; Zhao, Y., Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteomics* **2007,** *6* (5), 812-9.
4. Tan, M.; Luo, H.; Lee, S.; Jin, F.; Yang, J. S.; Montellier, E.; Buchou, T.; Cheng, Z.; Rousseaux, S.; Rajagopal, N.; Lu, Z.; Ye, Z.; Zhu, Q.; Wysocka, J.; Ye, Y.; Khochbin, S.; Ren, B.; Zhao, Y., Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **2011,** *146* (6), 1016-28.
5. Frye, R. A., Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochem Biophys Res Commun* **2000,** *273* (2), 793-8.
6. Wang, F.; Tong, Q., SIRT2 suppresses adipocyte differentiation by deacetylating FOXO1 and enhancing FOXO1's repressive interaction with PPARgamma. *Mol Biol Cell* **2009,** *20* (3), 801-8.
7. Wang, Y.; Yang, J.; Hong, T.; Chen, X.; Cui, L., SIRT2: Controversy and multiple roles in disease and physiology. *Ageing Research Reviews* **2019,** *55*, 100961.
8. Xu, Y.; Li, F.; Lv, L.; Li, T.; Zhou, X.; Deng, C.-X.; Guan, K.-L.; Lei, Q.-Y.; Xiong, Y., Oxidative Stress Activates SIRT2 to Deacetylate and Stimulate Phosphoglycerate Mutase. *Cancer Research* **2014,** *74* (13), 3630.
9. Serrano, L.; Martínez-Redondo, P.; Marazuela-Duque, A.; Vazquez, B. N.; Dooley, S. J.; Voigt, P.; Beck, D. B.; Kane-Goldsmith, N.; Tong, Q.; Rabanal, R. M.; Fondevila, D.; Muñoz, P.; Krüger, M.; Tischfield, J. A.; Vaquero, A., The tumor suppressor SirT2 regulates cell cycle progression and genome stability by modulating

the mitotic deposition of H4K20 methylation. *Genes Dev* **2013,** *27* (6), 639-53.
10. Zhao, S.; Xu, W.; Jiang, W.; Yu, W.; Lin, Y.; Zhang, T.; Yao, J.; Zhou, L.; Zeng, Y.; Li, H.; Li, Y.; Shi, J.; An, W.; Hancock, S. M.; He, F.; Qin, L.; Chin, J.; Yang, P.; Chen, X.; Lei, Q.; Xiong, Y.; Guan, K.-L., Regulation of Cellular Metabolism by Protein Lysine Acetylation. *Science* **2010,** *327* (5968), 1000-1004.
11. Zhou, W.; Ni, T. K.; Wronski, A.; Glass, B.; Skibinski, A.; Beck, A.; Kuperwasser, C., The SIRT2 Deacetylase Stabilizes Slug to Control Malignancy of Basal-like Breast Cancer. *Cell Rep* **2016,** *17* (5), 1302-1317.
12. Liu, P. Y.; Xu, N.; Malyukova, A.; Scarlett, C. J.; Sun, Y. T.; Zhang, X. D.; Ling, D.; Su, S. P.; Nelson, C.; Chang, D. K.; Koach, J.; Tee, A. E.; Haber, M.; Norris, M. D.; Toon, C.; Rooman, I.; Xue, C.; Cheung, B. B.; Kumar, S.; Marshall, G. M.; Biankin, A. V.; Liu, T., The histone deacetylase SIRT2 stabilizes Myc oncoproteins. *Cell Death Differ* **2013,** *20* (3), 503-14.
13. Zhao, D.; Zou, S. W.; Liu, Y.; Zhou, X.; Mo, Y.; Wang, P.; Xu, Y. H.; Dong, B.; Xiong, Y.; Lei, Q. Y.; Guan, K. L., Lysine-5 acetylation negatively regulates lactate dehydrogenase A and is decreased in pancreatic cancer. *Cancer Cell* **2013,** *23* (4), 464-76.
14. Jing, H.; Hu, J.; He, B.; Negrón Abril, Y. L.; Stupinski, J.; Weiser, K.; Carbonaro, M.; Chiang, Y.-L.; Southard, T.; Giannakakou, P.; Weiss, R. S.; Lin, H., A SIRT2-Selective Inhibitor Promotes c-Myc Oncoprotein Degradation and Exhibits Broad Anticancer Activity. *Cancer Cell* **2016,** *29* (3), 297-310.
15. Farooqi, A. S.; Hong, J. Y.; Cao, J.; Lu, X.; Price, I. R.; Zhao, Q.; Kosciuk, T.; Yang, M.; Bai, J. J.; Lin, H., Novel Lysine-Based Thioureas as Mechanism-Based Inhibitors of Sirtuin 2 (SIRT2) with Anticancer Activity in a Colorectal Cancer Murine Model. *Journal of Medicinal Chemistry* **2019,** *62* (8), 4131-4141.
16. Hong, J. Y.; Price, I. R.; Bai, J. J.; Lin, H., A Glycoconjugated SIRT2 Inhibitor with Aqueous Solubility Allows Structure-Based Design of SIRT2 Inhibitors. *ACS Chemical Biology* **2019,** *14* (8), 1802-1810.
17. Hong, J. Y.; Fernandez, I.; Anmangandla, A.; Lu, X.; Bai, J. J.; Lin, H., Pharmacological Advantage of SIRT2-Selective versus pan-SIRT1–3 Inhibitors. *ACS Chemical Biology* **2021,** *16* (7), 1266-1275.
18. Hong, J. Y.; Zhang, X.; Lin, H., HPLC-Based Enzyme Assays for Sirtuins. *Methods Mol Biol* **2018,** *1813*, 225-234.
19. Du, J.; Jiang, H.; Lin, H., Investigating the ADP-ribosyltransferase Activity of Sirtuins with NAD Analogues and 32P-NAD. *Biochemistry* **2009,** *48* (13), 2878-2890.

CHAPTER 4

Synthesis of Quaternary Ammonium Containing Crosslinkers for Crosslinking Mass Spectrometry

Abstract

Protein crosslinkers used in for crosslinking mass spectrometry must be able to react easily with amino acids, fragment predicably during MS ionization, and in some cases be purified by affinity chromatography while remaining as small as possible. Balancing all these requirements can be difficult. One prominent crosslinker is DSBSO, which sacrifices size and fragmentation simplicity to become affinity enrichable which can increase the rate of both false negatives and false positive hits. Here we present several quaternary ammonium containing protein crosslinkers that we are developing to solve some of the problems associated with DSBSO. We show that the new compounds crosslink proteins as well as DSSO, a simplified version of DSBSO, and report a similar number of GDH1 crosslinks when crosslinked peptides are identified with mass spectrometry.

Introduction

Protein-protein interactions (PPIs) are a crucial for cellular processes. As a natural consequence of their ubiquity, dysregulated PPIs have been implicated in numerous diseases including Huntington's disease, cancer, and diseases which stem from pathogenic infection including SARS-CoV-2[1-6]. Though progress has been made in

developing small molecule drugs to inhibit specific interactions, PPIs are generally much less druggable targets than enzymes with defined binding pockets[7-9]. By identifying new PPIs and mapping interactomes, a better understanding of normal cellular functions and diseases can be achieved.

Numerous studies have mapped myriad interactomes resulting in the construction of multiple databases to compile and organize PPI networks[10-12]. Multiple approaches have been taken to identify PPIs on a proteome wide scale. Of these methods, two that have found a great deal of success are yeast 2 hybrid (Y2H) and affinity purification mass spectrometry (AP-MS). However, both these methods have limitations that prevent them from becoming a truly proteome wide scale approach for PPI detection[6, 12, 13]. Both require choosing specific proteins of interest for which interactors are identified. For Y2H, both proteins must be modified thus requiring the tedious and laborious process of generating fusion proteins for all interactor candidates. With AP-MS, only a single protein of interest needs to be modified with a purification tag and the use of mass spectrometry allows for extremely high-through put detection of PPIs. However, since this method requires the binding affinity to be high enough for copurification of the interacting partners, transiently associated proteins and weakly interacting PPIs can be lost during the purification process[14].

Crosslinking mass spectrometry (XL-MS) does not have the same drawbacks of Y2H or AP-MS and has the potential to be a truly proteome wide approach for identifying PPIs[15, 16]. XL-MS can be applied to any cell or protein solution without the need for engineered proteins. A specially designed bifunctional chemical crosslinker with a MS labile bond is used to crosslink interacting proteins thereby stabilizing even weaker

interactions[17]. The crosslinker fragments in a predicable way during MS ionization producing peptide/linker fragments with a signature mass difference that is used to identify peptides that originated from interacting partners (Figure 4.1)[18]. This means that by simply crosslinking a cell lysate and analyzing the sample with tandem MS, potentially an entirely complete interactome can be mapped with a single experiment. In practice however, there are several limitations holding back XL-MS from attaining its full potential.

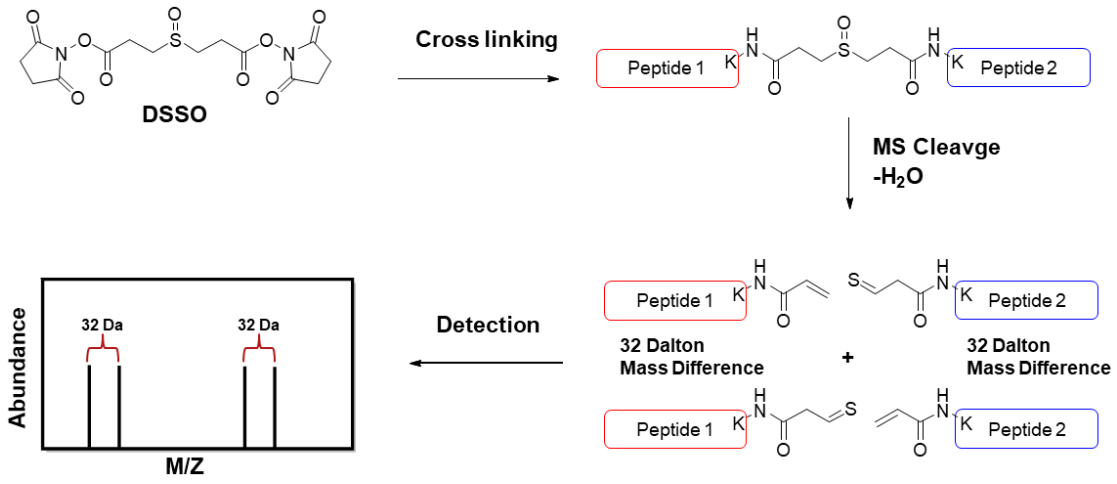


Figure 4.1. Crosslink detection with Disuccinimidyl Sulfoxide. Disuccinimidyl Sulfoxide (DSSO) is used to demonstrate the how crosslinks are identified with XL-MS. Succinimide esters form amide bonds with lysine side chains producing crosslinks. Both C-S bonds in the compound are equally likely to fragment during MS ionization producing pairs of peptide linker fragments that differ in mass by 32 Da. Crosslinks can therefore be identified by searching for doublets split by a difference of 32 Da.

The most common modification of a crosslinking reaction results in only one end of the linker reacting with a protein, which is referred to as a dead end (DE) modification[14]. Intraprotein crosslinks are very common but can only give structural information about a single protein[19]. Interprotein crosslinks that result from PPIs are the least common. As a result, the signal from interprotein crosslinks can be easily drowned out by the noise

of intraprotein links, DEs, and non-crosslinked peptides[20, 21]. Moreover, the crosslinker itself can be a source of problems since the larger the mass difference of the crosslinker, the higher the rate of false positive hits detected[22].

A XL-MS crosslinker that has become very widely used is disuccinimidyl sulfoxide (DSSO) (Figure 4.1)[23-26]. This compound uses succinimide esters to crosslink lysine residues and can produce linker peptide fragments with a 32 Da mass difference[26]. A major limitation of DSSO is that it cannot be enriched by affinity purification and therefore cannot be used in complex samples for proteome-wide PPI identification. To address this issue, a DSSO derivative, (DSBSO) (Figure 4.2 A) was synthesized to include azide or alkyne click chemistry handles which are attached to the linker through acid cleavable bonds. This allows for elution of crosslinked peptides after affinity enrichment by treatment with an acidic buffer. To accommodate this design, a second sulfoxide group was added, which in turn introduced an over fragmentation problem wherein the linker can fragment twice, eliminating any mass difference and rendering crosslinked peptides undetectable[27, 28].

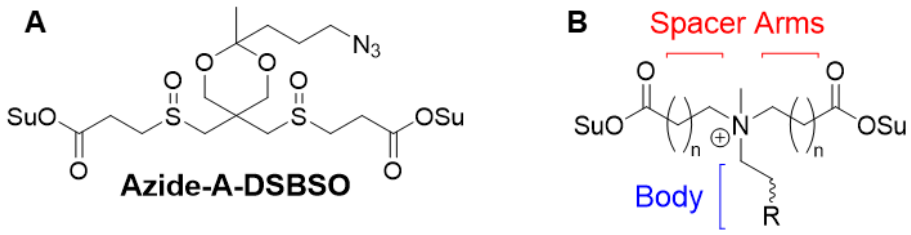


Figure 4.2 Structure of Azide-A-DSBSO and quaternary ammonium crosslinkers. Azide-A-DSBSO (**A**) is the azide containing version of DSSO which allow for click conjugation to biotin alkyne. DSBSO has 2 fragmentable bonds which can produce fragments with no mass difference. The design for our quaternary ammonium crosslinkers (**B**) contains 3 alkyl chains. The chains that terminate in succinimide esters are referred to as the spacer arms. The remaining chain we refer to as the linker body and is where we will incorporate an alkyne or azide group to allow for click chemistry attachment of biotin.

In this work we aim to create an affinity purifiable crosslinker that does not have the over fragmentation problem of DSBSO. We use quaternary ammoniums as the source of the MS liable bonds in our crosslinkers (Figure 4.2 B). During ionization quaternary ammoniums will undergo Hofmann elimination to produce alkene and tertiary amine fragments, neither of which can fragment further[29]. Quaternary ammoniums also allow click chemistry handles to be attached to the linker body. This simplifies the synthesis of the crosslinker and should allow us to minimize the mass difference of our probes. We will also utilize multiple amino acid modification reactions to achieve a broad coverage of PPIs[30-39].

Results and Discussion

Quaternary Ammonium Crosslinker Design and Synthesis

Despite the apparent simplicity, the use of quaternary ammoniums salts in our crosslinkers presented some synthetic challenges. To maximize fragmentation within the spacer arm, probes were designed to not have hydrogens on carbons β to the nitrogen in the linker body. We also aimed to minimize the mass of the linker body to reduce the mass difference of the linker fragments, thereby reducing the rate of false positive hit detection. These constrains greatly limited the amines that could be used to form the linker body. Succinimide esters proved to be unstable with 2 carbons in the spacer arms unless electron deficient amines were used. These amines required harsh electrophiles to alkylate, which required linkers to be designed to avoid incorporating other nucleophilic moieties.

Non-enrichable quaternary ammonium crosslinkers were initially synthesized, tested,

and iteratively improved. This resulted in KW133, 3024A, 3023B, and 3024B (Figure 4.3). Crosslinkers with clickable handles, such as KW5003 and 5101, were made after fragmentation issues with the simpler probes had been addressed (Figure 4.3).

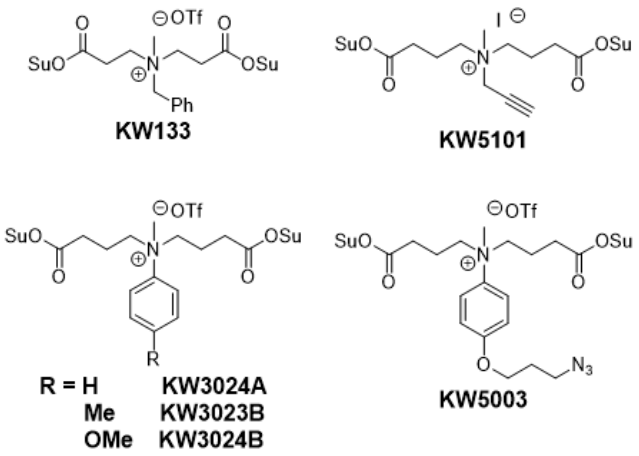


Figure 4.3. Structure of non-enrichable and enrichable quaternary ammonium crosslinkers.

KW133 was the first successful attempt at synthesizing a non-enrichable crosslinker. This compound did not crosslink as well as DSSO though it was able to crosslink proteins to a small degree (Figure 4.4). Fragmentation problems arose when attempting to detect PPIs with MS using KW133. Firstly, the crosslinker fragmented in MS1 instead of only fragmenting when the ionization energy is increased in MS2. This significantly reduced the rate of crosslink detection. KW133 also fragmented at the benzylic position rather than in the spacer arm. To address these problems, we made a derivative from aniline instead of benzyl amine. However, this compound also fragmented at MS1, though only in the spacer arm. We reasoned that since the hydrogens in the spacer arms that were β with respect to the nitrogen were also α to the carbonyl carbons, these protons were likely very acidic making the quaternary

ammonium too unstable and causing it to fragment early. To address this, the spacer arm was extended by an additional methylene group resulting in the crosslinkers KW3023B, KW3024A, and KW3024B. These compounds were more stable fragmented only in MS2 as desired.

Extended Spacer Arms Improved Crosslinking Reactions

Before the linkers could be used in a XL-MS experiment we first confirmed they were able to crosslink proteins. To do this, a gel mobility shift assay was used to visualize the crosslinking of glutamate dehydrogenase 1 (GDH1). GDH1 is a hexameric protein. Lysine crosslinking can stabilize the hexameric structure and allows the detection of the complex on denaturing gel electrophoresis. As a result, in SDS-PAGE gel, crosslinked GDH1 shifts to higher molecular weight bands in intervals of roughly 50 kDa (molecular weight of GDH1 monomer) as the amount of crosslinking increases. This allows for a rough comparison of the crosslinking efficiencies of our new molecules (Figure 4.4). KW133 only has a small amount of crosslinking. When the spacer arms were extended by adding an addition methylene group, the amount of crosslinking significantly increased. KW323B, 324A, 5003, and 5101 all crosslinked GDH1 about as well as DSSO.

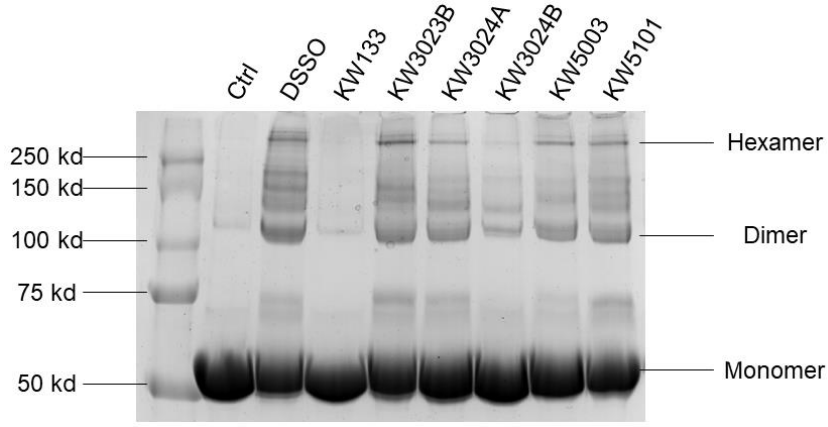


Figure 4.4. GDH1 crosslinking with new protein crosslinkers. KW133 showed only minor crosslinking, however, extending the spacer arm greatly improved the crosslinking efficiency of the subsequent compounds.

Nonenrichable Crosslinkers Report a Similar Number of Crosslinks as DSSO

We examined how well our new probes could detect PPIs compared to DSSO by crosslinking GDH1 and identifying the linked peptides with MS. We found that the quaternary ammoniums reported ~80% the number of unique crosslinks as DSSO, nearly 40% of which were not detected by DSSO (Figure 4.5).

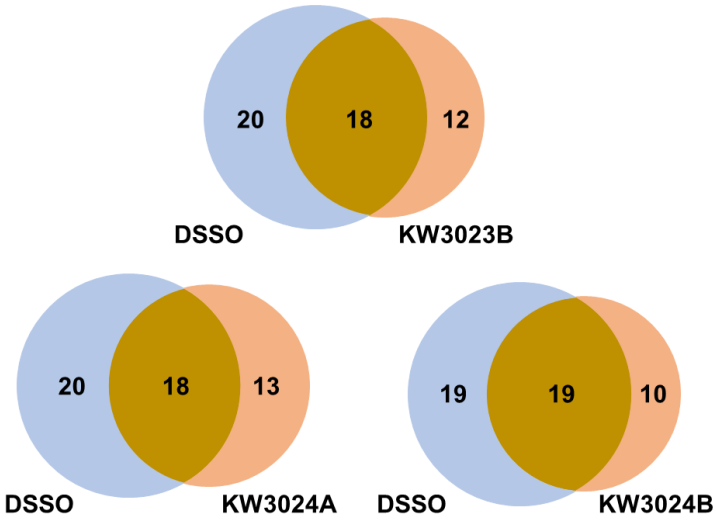


Figure 4.5. Crosslinker comparison by identifying peptide crosslinks. The sequences of crosslinked peptides were determined from MS3 data to identify unique crosslinks for each crosslinker. The identity of crosslinks in overlap and non-overlap regions of each Ven diagram do not necessarily correspond to crosslinks in the same sections of the other diagrams.

From here, KW5003 and KW5101 (Figure 4.3) were made with handles to allows for click chemistry attachment of an affinity purification tag. The protocols for using these crosslinkers to enrich and identity crosslinked peptides in a whole cell lysate are currently being optimized. Attempts were made to synthesize crosslinkers with chemically cleavable bonds in the linker body between the quaternary ammonium and the click chemistry moiety, however none of synthetic routes attempted proved viable due mainly to issues that arose during the alkylation step.

Future Directions

The designs of the crosslinkers can be further refined to minimize their mass differences. One way of achieving this is to place a chemically cleavable moiety between the clickable handle and quaternary nitrogen. The attempts that had been made to this end usually failed due to the need to use harsh electrophiles during the methylation step which often destroyed the cleavable group. To avoid this problem, we could instead remove the requirement for no hydrogens on the β carbon of the linker body which would allow for more electron rich amines to be used. These amines could be methylated with methyl iodide instead of methyl triflate as was used for all but one of the new crosslinkers.

Collaborative projects with other labs will apply our crosslinkers to identify new PPIs. One such project is in progress and will use XL-MS and KW5101 to identify proteins that regulate RNA polymerase II transcription. As part of another collaboration with the Yu lab, we are attempting to use the crosslinkers to build a HRas interactome.

Finally, a method for purifying crosslinked peptides from DEs is being developed. A

lysine/KW5101/lysine conjugate that can be attached to a solid support is being synthesized. The lysine/KW5101 conjugate will serve as a mimic for crosslinked peptides. Immobilizing this molecule on agarose beads should allow for high-throughput selection of DNA aptamers with high affinity for crosslinked peptides. These aptamers will then be used to positively select crosslinks from DEs, thereby reducing unwanted signals during MS analysis and improving crosslink detection.

Materials and Methods

All solvents and reagents were purchased from commercial vendors as analytical or higher-grade purity. Flash chromatography was done using SiliaFlash Irregular Silica Gel, P60, 40 – 63 μm, 60 Å. NMR spectra were collected at the Cornell NMR Facility using Brucker 400 and 500 spectrometers. Disuccinimidyl sulfoxide (DSSO) was purchased from Thermo Fisher Scientific (catalog number A33545).

GDH1 expression and purification

10G E. coli cells containing pET28a plasmids with yeast glutamate dehydrogenase 1 (GDH1) gene inserts were grown in LB (10 mL, 50 μg/mL kanamycin) at 37 ^{o}C for 16 h. Cells were harvested by centrifugation (4500 xG for 5 min). Plasmids were collected and purified by miniprep. The plasmids were then used to transform BL21 E. coli. Transformed bacteria were selected by growing the transformed LB on an agar plate containing kanamycin (50 μg/mL) for 16 h. A single colony was selected and used to inoculate a starter culture (10 mL LB, 50 μg/mL kanamycin) which was grown for 16 h. This culture was then used to inoculate 2L of LB (with 50 μg/mL kanamycin). Cells

were grown at 37 ºC to OD600 of 0.6 then cooled to 16 ºC. Protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPGA, 0.15mM) and grown at 16 ºC for 16 h. Cells were harvested by centrifugation and lysed with an EmulsiFlex-C3 cell disruptor. The lysate was centrifugated to pellet cellular debris. Supernatant was added to a column which was loaded with Ni-NTA resin (1.2 mL) that had been washed with lysis buffer (500 mM NaCl, 50 mM Tris-HCl at pH 8.0, 5% glycerol, 1 mM TCEP, 5 mL). After the supernatant had flowed through, the column was washed with lysis buffer containing 30 mM imidazole (50 mL). Protein was eluted by passing solutions of increasing concentration of imidazole in lysis buffer (50, 75, 100, 150, 200, and 250, mM, 10mL each) over the column. 12 fractions (5 mL each) were collected during elution. SDS-PAGE was performed on the fractions to determine which contained GDH1. GDH1 containing fractions were pooled and desalted by dialysis. The protein concentration of the resulting solution was determined by BCA assay.

GDH1 gel shift assay

Purified yeast glutamate dehydrogenase 1 (GDH1) was diluted to 0.5 mg/mL in PBS (20 µL, pH 7.4). Crosslinker (1 µL, 2 mM in DMSO) was added and the solution was kept at room temperature for 1 h. The reaction was quenched with 6X gel loading dye and the samples were heat denatured and run on a SDS PAGE gel. The gel was stained with Coomassie.

GDH1 crosslinking for XL-MS

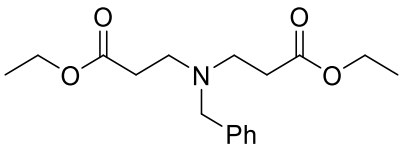
A solution of GDH1 (100 μL, 1mg/mL in HEPES 20 mM, pH 7.5) was combined with a crosslinker solution (2 μL, 50 mM in DMSO), and incubated at room temperature for 30 min. The reaction was quenched by addition of Tris-Cl (2 μL, pH 7.5) followed by at 15 min incubation at room temperature. A 5 μL aliquot was taken for SDS-PAGE analysis to confirm crosslinking had occurred. SDS was added to the remaining sample to 1% followed by a 15 min incubation at 65 ^{o}C. DTT was added (1 μL, 1M) and the sample was incubated at room temperature for 15 min. Alkylating solution (20X stock solution, 0.5M iodoacetamide, 1M Tris, pH 8.0) was added, followed by another 15 min. incubation. A 3X volume of PPT solution (50% acetone, 49.9% ethanol, 0.1% acetic acid) was added and the samples were incubated on ice for 30 min then centrifuged at 16,100 *g* (10 min., 4 ^{o}C). Supernatant was discarded, the pellet was washed with PPT solution (3X volume) then sonicated in a water bath for 10 seconds. The sample was centrifuged again, supernatant was discarded, and the pellet was to air dried for 5 min. The pellet was redissolved in a Urea/Tris solution (50 μL, 8M urea, 50 mM Tris) and NaCl/Tris solution (150 μL, 50 mM Tris, 150 mM NaCl, pH 8.0). Trypsin Gold (1 μL, 1 μg/μL) was added and the sample was incubated at 37 ^{o}C for 16 h. TFA and formic acid (5 μL and 10% by volume each) were added to the digested protein solution which was centrifuged at 16,100 *g* for 1 min. A C18 Sep-Pac column (50 mg) was washed (1 volume 80% acetonitrile, 29.9 % water, 0.1% acetic acid) and equilibrated (1 volume 0.1% TFA). The protein solution was loaded on the column and allowed to enter by gravity. The column was washed twice (0.1% acetic acid, 1 volume each time). Peptide was eluted with 80% acetonitrile, 29.9 % water, 0.1% acetic acid

(200 μL) and solvent was removed by Speed Vac (45 ^{o}C for 40 min). .1P angio solution (.1P angiotensin in 0.1% TFA, .5 μL/μg) was added and the sample was taken for XL-MS analysis.

Synthetic Methods

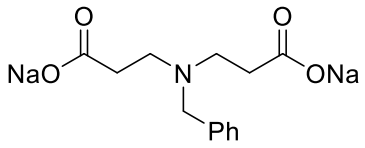
Procedure 1, sodium dicarboxylates. A diester (1 eq.) and NaOH (2 eq.) were combined in EtOH (0.2M) and stirred at room temperature for 16 h. Solid product was collected by vacuum filtration and taken to the next step without further purification.

Procedure 2, succinimide esters. A sodium dicarboxylate (1 eq.), EDC (2.5 eq.), and N-Hydroxyl succinimide (NHS, 2.2 eq.) were combined in DCM (0.1M) and stirred at room temperature for 16 h. The mixture was diluted with DCM (1 volume), washed with water and brine, and dried over sodium sulfate. Solvent was removed under reduced pressure. The resulting crude solid was purified by flash chromatography.



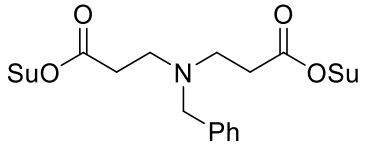
S1

diethyl 3,3'-(benzylazanediyl)dipropionate (S1). Benzylamine (1mL, 9.1 mmol) and ethyl acrylate (2.1 mL, 23 mmol) were combined in EtOH (30 mL) and heated to reflux for 16 h. After cooling, solvent was removed under reduced pressure. Crude liquid was purified by flash chromatography (1:3, EtOAc : hexane) to give product **S1** (1.0g 4.2 mmol, 47% yield).



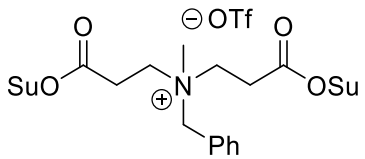
S2

sodium 3,3'-(benzylazanediyl)dipropionate (S2). S2 was synthesized from **S1** by procedure 1 (53% yield).



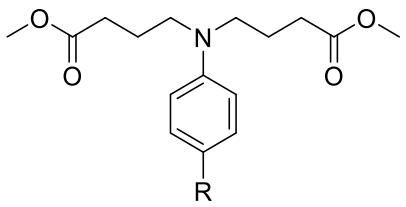
S3

bis(2,5-dioxopyrrolidin-1-yl) 3,3'-(benzylazanediyl)dipropionate (S3). S3 was synthesized from **S2** by procedure 2 (34% yield).



KW133

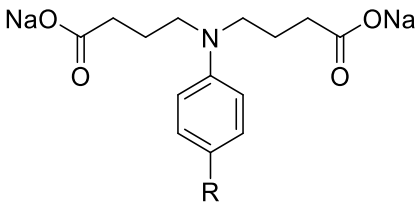
N-benzyl-3-((2,5-dioxopyrrolidin-1-yl)oxy)-N-(3-((2,5-dioxopyrrolidin-1-yl)oxy)-3-oxopropyl)-N-methyl-3-oxopropan-1-aminium trifluoromethanesulfonate (KW133). S3 (0.60g, 1.4 mmol) was dissolved in DCM (5 mL). Methyl triflate (0.22 mL 2 mmol) was added dropwise and the reaction was stirred at room temperature for 16 h. Solvent was removed under reduced pressure. Crude material was purified by flash chromatography (EtOAc → 1:3 MeOH : EtOAc) to give the final product.



R = H S4
Me S5
OMe S6

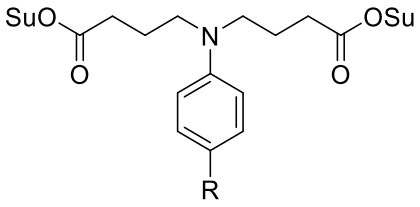
N-benzyl-3-((2,5-dioxopyrrolidin-1-yl)oxy)-N-(3-((2,5-dioxopyrrolidin-1-yl)oxy)-3-oxopropyl)-N-methyl-3-oxopropan-1-aminium trifluoromethanesulfonate

dimethyl 4,4'-(phenylazanediyl)dibutyrate (S4), dimethyl 4,4'-(p-tolylazanediyl)dibutyrate (S5), and dimethyl 4,4'-((4-methoxyphenyl)azanediyl)dibutyrate (S6). Aniline, toluidine, or p-anisidine (1 eq.) were combined with methyl 4-bromobutanoate (4 eq.), NaI (1.5 eq.), and N, N-diisopropylethylamine (8 eq.) in acetonitrile (0.5 M). The reaction mixture was heated to reflux for 16 h. After cooling to room temperature, the mixture was poured into water and extracted 3 times with DCM. Organic layers were combined, dried over sodium sulfate, and concentrated under reduced pressure. The crude material was purified by flash chromatography (1:9 → 2:3, EtOAc : hexane) to give **S4-S6** (59%, 52%, and 55% yield respectively)



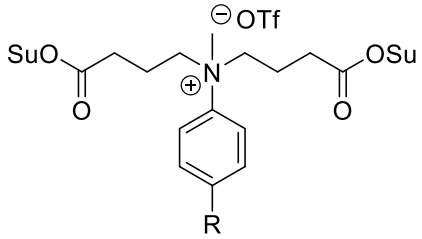
R = H S7
Me S8
OMe S9

sodium 4,4'-(phenylazanediyl)dibutyrate (S7), sodium 4,4'-(p-tolylazanediyl)dibutyrate (S8), and sodium 4,4'-((4-methoxyphenyl)azanediyl)dibutyrate (S9). S7-S9 were synthesized from **S4-S6** by procedure 1 (74%, 39%, and 55% yield respectively).



R = H S10
Me S11
OMe S12

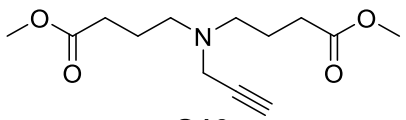
bis(2,5-dioxopyrrolidin-1-yl) 4,4'-(phenylazanediyl)dibutyrate (S10), bis(2,5-dioxopyrrolidin-1-yl) 4,4'-(p-tolylazanediyl)dibutyrate (S11), and bis(2,5-dioxopyrrolidin-1-yl) 4,4'-((4-methoxyphenyl)azanediyl)dibutyrate (S12). S10-S12 were synthesized from **S7-S9** by procedure 2 (36%, 39%, and 49% yield respectively).



R = H KW3024A
Me KW3023B
OMe KW3024B

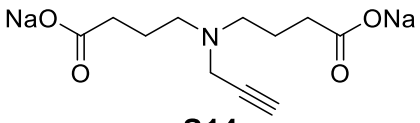
N,N-bis(4-((2,5-dioxopyrrolidin-1-yl)oxy)-4-oxobutyl)-N-methylbenzenaminium trifluoromethanesulfonate (KW3024A), N,N-bis(4-((2,5-dioxopyrrolidin-1-yl)oxy)-4-oxobutyl)-N,4-dimethylbenzenaminium trifluoromethanesulfonate (KW3023B), and N,N-bis(4-((2,5-dioxopyrrolidin-1-yl)oxy)-4-oxobutyl)-4-methoxy-N-methylbenzenaminium trifluoromethanesulfonate (KW3024B). S10, S11, or S12 (1 eq.) was dissolved in DCM and methyl triflate (2 eq.) was added dropwise. The reaction was stirred at room temperature for 16 h. Precipitate was collected by vacuum filtration and washed with DCM to give highly pure products (68%, 37% and 90% yield respectively). **KW3024A** ^{1}H NMR (400 MHz, MeOD) δ 7.83 (dd, *J* = 6.4, 4.0 Hz, 2H), 7.76 – 7.62 (m, 4H), 4.16 (tt, *J* = 13.2, 6.5 Hz, 2H), 3.96 (td, *J* = 12.4, 4.7 Hz, 2H), 3.69 (s, 3H), 2.87 (s, 8H), 2.77 (t, *J* = 6.9 Hz, 4H), 2.15 (tq, *J* = 12.2, 6.8 Hz, 2H), 1.70 (tq, *J* = 11.7, 6.6 Hz, 2H). ^{13}C NMR (126 MHz, MeOD) δ 173.5, 172.5, 141.5, 130.5, 130.4, 121.4, 121.4, 67.7, 67.2, 29.1, 24.9, 18.0. ^{19}F NMR (376 MHz, MeOD) δ -80.1. **KW3023B** ^{1}H NMR (400 MHz, MeOD) δ 7.73 – 7.65 (m, 2H), 7.51 (d, *J* = 8.5 Hz, 2H), 4.13 (td, *J* = 12.5, 4.6 Hz, 2H), 3.93 (td, *J* = 12.4, 4.7 Hz, 2H), 3.65 (s, 3H), 2.86 (s, 8H), 2.76 (t, *J* = 6.9 Hz, 4H), 2.46 (s, 3H), 2.14 (tq, *J* = 12.1, 6.5 Hz, 2H), 1.79 – 1.61 (m, 2H). ^{13}C NMR (101 MHz, MeOD) δ 170.4, 167.8, 141.4, 131.0, 121.1, 67.0, 26.7, 25.1, 19.4, 18.0. ^{19}F NMR (376 MHz, MeOD) δ -80.1. **KW3024B** ^{1}H

NMR (400 MHz, DMSO) δ 7.79 (d, *J* = 8.9 Hz, 2H), 7.20 (d, *J* = 8.9 Hz, 2H), 4.50 (ddd, *J* = 14.5, 10.2, 5.2 Hz, 2H), 4.31 (td, *J* = 11.7, 5.4 Hz, 2H), 3.86 (s, 3H), 3.66 (s, 3H), 3.51 – 3.26 (m, 6H), 2.96 – 2.63 (m, 10H). ^{13}C NMR (101 MHz, DMSO) δ 170.4, 166.2, 160.7, 133.1, 123.8, 116.1, 63.1, 56.2, 48.0, 25.9, 25.8. ^{19}F NMR (376 MHz, DMSO) δ -77.8.



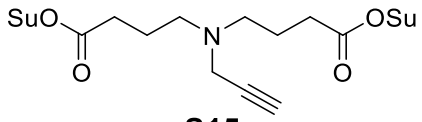
S13

dimethyl 4,4'-(prop-2-yn-1-ylazanediyl)dibutyrate (S13). Propargyl amine (0.52g, 9.4 mmol), $KHCO_3$ (2.9g, 29 mmol), KI (0.15g, 0.9 mmol), and methyl 4-bromobutanoate (2.4 mL, 19 mmol) were combined in acetonitrile (30 mL) and heated to reflux for 16 h. The mixture was cooled to room temperature and solvent was removed under reduced pressure. The resulting crude was taken up into water (30 mL) and extracted 3 times with EtOAc (30 mL each time). Organic layers were pooled, dried over sodium sulfate, and concentrated under reduced pressure. The second crude was purified by flash chromatography (1:4 → 2:3, EtOAc : hexane) to give the final product (1.5g, 5.7 mmol 60% yield).



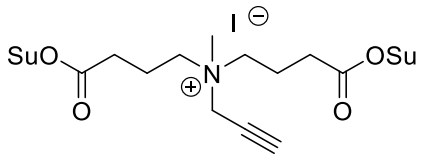
S14

sodium 4,4'-(prop-2-yn-1-ylazanediyl)dibutyrate (S14). S14 was synthesized from **S13** by procedure 1 (51 % yield).



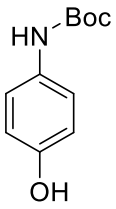
S15

bis(2,5-dioxopyrrolidin-1-yl) 4,4'-(prop-2-yn-1-ylazanediyl)dibutyrate (S15). S15 was synthesized from **S14** by procedure 2 (47% yield).



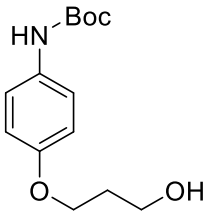
KW5101

4-((2,5-dioxopyrrolidin-1-yl)oxy)-N-(4-((2,5-dioxopyrrolidin-1-yl)oxy)-4-oxobutyl)-N-methyl-4-oxo-N-(prop-2-yn-1-yl)butan-1-aminium iodide (KW5101). **S15** (0.60g 1.4 mmol) and methyl iodide (0.44 mL 7.1 mmol) were combined in acetonitrile and stirred at room temperature overnight. Solvent was removed under reduced pressure. The resulting material was purified by flash chromatography (DCM → 1:9, MeOH : DCM) to give **KW5101** (48% yield). ^{1}H NMR (400 MHz, DMSO) δ 4.45 (d, *J* = 2.5 Hz, 2H), 4.07 (d, *J* = 2.4 Hz, 1H), 3.51 – 3.39 (m, 4H), 3.11 (s, 3H), 2.86 (d, *J* = 17.3 Hz, 12H), 2.07 (p, *J* = 7.4 Hz, 4H). ^{13}C NMR (101 MHz, DMSO) δ 170.6, 168.7, 83.8, 72.4, 60.0, 52.5, 48.5, 27.4, 25.9, 17.6.



S16

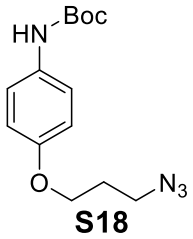
tert-butyl (4-hydroxyphenyl)carbamate (S16). 4-aminophenol (0.99g, 9.1 mmol) was dissolved in THF (30 mL). Boc_2O (2.5 mL, 11 mmol) was added drop wise. The solution was heated to reflux and stirred for 16 h. After cooling, solvent was removed under reduced pressure. The crude material was taken up in EtOAc (30mL), washed with water, saturated aqueous $NaHCO_3$, and brine (30 mL each), and dried over sodium sulfate. Solvent was removed under reduced pressure. The resulting solid was stirred in hexane for 4 h then solid product was collected by vacuum filtration and taken to the next step without further purification (1.4g, 6.7 mmol, 74% yield).



S17

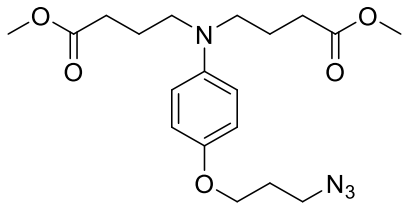
tert-butyl (4-(3-hydroxypropoxy)phenyl)carbamate (S17). S16 (2.5g, 12 mmol), K_2CO_3 (0.50g, 3.6 mmol) and 3-bromopropan-1-ol (2 mL, 23 mmol) were combined in acetonitrile (40 mL) and heated to reflux for 16 h. After cooling solvent was removed under reduced pressure and the resulting material was taken up in water (50 mL). This mixture was extracted 3 times with DCM. Organic layers were pooled, dried over sodium sulfate, and concentrated under educed pressure. The crude then purified by

flash chromatography (2:3 → 1:1, EtOAc : hexane) to give **S17** (2.4g, 9.1 mmol, 76% yield).



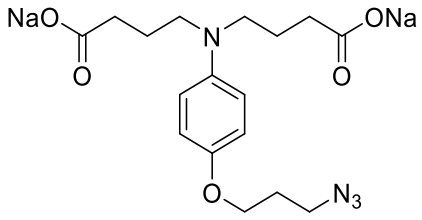
S18

tert-butyl (4-(3-azidopropoxy)phenyl)carbamate (S18). S17 (1.9g, 7.1 mmol) and Et_3N (0.22 mL, 14 mmol) were dissolved in DCM (50 mL) and chilled to 0 ^oC. MsCl (0.83 mL, 11 mmol) was added dropwise. The solution was stirred at 0 ^oC for 30 min then warmed to room temperature and stirred for an additional 5 h. Water (50 mL) was added followed by DCM (50 mL). The organic layer was washed with water, and brine (50 mL each), dried over sodium sulfate and concentrated under educed pressure. The resulting material was redissolved in DMF (50 mL) and sodium azide (0.92g, 14 mmol) was added. This mixture was heated to 70 ^oC for and stirred for 16 h. After cooling solvent was removed under reduced pressure. The crude material was added to water (25 mL) and extracted 3 times with EtOAc (25 mL each). Organic layers were pooled, dried over sodium sulfate, and concentrated under reduced pressure. The final crude was purified by flash chromatography (hexane → 1:4 EtOAc : hexane) to give **S18** (2.0g, 6.8 mmol, 96% yield)



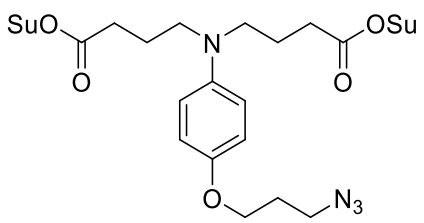
S19

dimethyl 4,4'-((4-(3-azidopropoxy)phenyl)azanediyl)dibutyrate (S19). S18 (2.3g, 7.7 mmol) was dissolved in DCM (25 mL) and TFA (5 mL) and stirred at room temperature for 4 h. Solvent was removed under reduced pressure. This material was combined in acetonitrile (16 mL) with methyl 4-bromobutanoate (3.9 mL, 31 mmol), NaI (1.7g, 12 mmol), and N, N-diisopropylethylamine (12 mL, 70 mmol). The reaction mixture was heated to reflux and stirred for 16 h. After cooling to room temperature, the mixture was poured into water and extracted 3 times with DCM. Organic layers were combined, dried over sodium sulfate, and concentrated under reduced pressure. The crude material was purified by flash chromatography (1:20 → 3:7, EtOAc : hexane) to give **S19** (1.6g, 54% yield).



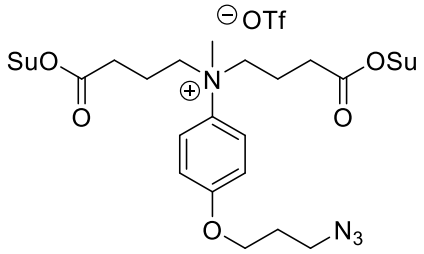
S20

sodium 4,4'-((4-(3-azidopropoxy)phenyl)azanediyl)dibutyrate (S20). S20 was synthesized from **S19** by procedure 1 (80% yield).



S21

bis(2,5-dioxopyrrolidin-1-yl) 4,4'-((4-(3-azidopropoxy)phenyl)azanediyl)dibutyrate (S21). S21 was synthesized from **S20** by procedure 2 (29% yield).



KW5003

4-(3-azidopropoxy)-N,N-bis(4-((2,5-dioxopyrrolidin-1-yl)oxy)-4-oxobutyl)-N-methylbenzenaminium trifluoromethanesulfonate (KW5003). S21 (0.10g, 0.17 mmol) was dissolved in DCM (2 mL) and chilled to 0 ^{o}C. Methyl triflate (22 μL, 0.19 mmol) was added dropwise. The solution was slowly warmed to room temperature and stirred for 16 h. Solvent was removed under reduced pressure. The crude material was purified by flash chromatography (DCM → 1:4 MeOH : DCM) to give **KW5003** (0.053g, 0.073 mmol, 43% yield). ^{1}H NMR (400 MHz, DMSO) δ 7.84 – 7.70 (m, 2H), 7.19 (d, *J* = 9.2 Hz, 2H), 4.14 (t, *J* = 6.1 Hz, 2H), 4.00 (qd, *J* = 12.2, 5.8 Hz, 2H), 3.82 (qd, *J* = 12.0, 4.5 Hz, 2H), 3.56 – 3.49 (m, 4H), 2.77 (m, 15H), 2.00 (m, 4H), 1.50 (m, 2H).

References

1. Ryan, D. P.; Matthews, J. M., Protein–protein interactions in human disease. *Current Opinion in Structural Biology* **2005,** *15* (4), 441-446.
2. Bowler, E. H.; Wang, Z.; Ewing, R. M., How do oncoprotein mutations rewire protein-protein interaction networks? *Expert Rev Proteomics* **2015,** *12* (5), 449-55.
3. Tourette, C.; Li, B.; Bell, R.; O'Hare, S.; Kaltenbach, L. S.; Mooney, S. D.; Hughes, R. E., A large scale Huntingtin protein interaction network implicates Rho GTPase signaling pathways in Huntington disease. *J Biol Chem* **2014,** *289* (10), 6709-6726.
4. Gartel, A. L., FOXM1 in Cancer: Interactions and Vulnerabilities. *Cancer Res* **2017,** *77* (12), 3135-3139.
5. Hunter, T., Oncoprotein networks. *Cell* **1997,** *88* (3), 333-46.
6. Gordon, D. E.; Jang, G. M.; Bouhaddou, M.; Xu, J.; Obernier, K.; O'Meara, M. J.; Guo, J. Z.; Swaney, D. L.; Tummino, T. A.; Hüttenhain, R.; Kaake, R. M.; Richards, A. L.; Tutuncuoglu, B.; Foussard, H.; Batra, J.; Haas, K.; Modak, M.; Kim, M.; Haas, P.; Polacco, B. J.; Braberg, H.; Fabius, J. M.; Eckhardt, M.; Soucheray, M.; Bennett, M. J.; Cakir, M.; McGregor, M. J.; Li, Q.; Naing, Z. Z. C.; Zhou, Y.; Peng, S.; Kirby, I. T.; Melnyk, J. E.; Chorba, J. S.; Lou, K.; Dai, S. A.; Shen, W.; Shi, Y.; Zhang, Z.; Barrio-Hernandez, I.; Memon, D.; Hernandez-Armenta, C.; Mathy, C. J. P.; Perica, T.; Pilla, K. B.; Ganesan, S. J.; Saltzberg, D. J.; Ramachandran, R.; Liu, X.; Rosenthal, S. B.; Calviello, L.; Venkataramanan, S.; Lin, Y.; Wankowicz, S. A.; Bohn, M.; Trenker, R.; Young, J. M.; Cavero, D.; Hiatt, J.; Roth, T.; Rathore, U.; Subramanian, A.; Noack, J.; Hubert, M.; Roesch, F.; Vallet, T.; Meyer, B.; White, K. M.; Miorin, L.; Agard, D.; Emerman, M.; Ruggero, D.; García-Sastre, A.; Jura, N.; von Zastrow, M.; Taunton, J.; Schwartz, O.; Vignuzzi, M.; d'Enfert, C.; Mukherjee, S.; Jacobson, M.; Malik, H. S.; Fujimori, D. G.; Ideker, T.; Craik, C. S.; Floor, S.; Fraser, J. S.; Gross, J.; Sali, A.; Kortemme, T.; Beltrao, P.; Shokat, K.; Shoichet, B. K.; Krogan, N. J., A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing. *bioRxiv* **2020**, 2020.03.22.002386.
7. White, P. W.; Titolo, S.; Brault, K.; Thauvette, L.; Pelletier, A.; Welchner, E.; Bourgon, L.; Doyon, L.; Ogilvie, W. W.; Yoakim, C.; Cordingley, M. G.; Archambault, J., Inhibition of Human Papillomavirus DNA Replication by Small Molecule Antagonists of the E1-E2 Protein Interaction*. *Journal of Biological Chemistry* **2003,** *278* (29), 26765-26772.
8. Carry, J. C.; Garcia-Echeverria, C., Inhibitors of the p53/hdm2 protein-protein interaction-path to the clinic. *Bioorg Med Chem Lett* **2013,** *23* (9), 2480-5.
9. Sang, P.; Zhang, M.; Shi, Y.; Li, C.; Abdulkadir, S.; Li, Q.; Ji, H.; Cai, J., Inhibition of β-catenin/B cell lymphoma 9 protein-protein interaction using α-helix-mimicking sulfono-γ-AApeptide inhibitors. *Proc Natl Acad Sci U S A* **2019,** *116* (22), 10757-10762.
10. Szklarczyk, D.; Gable, A. L.; Nastou, K. C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N. T.; Legeay, M.; Fang, T.; Bork, P.; Jensen, L. J.; von Mering, C., The STRING database in 2021: customizable protein-protein networks, and functional

characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* **2021,** *49* (D1), D605-d612.
11. Huttlin, E. L.; Ting, L.; Bruckner, R. J.; Gebreab, F.; Gygi, M. P.; Szpyt, J.; Tam, S.; Zarraga, G.; Colby, G.; Baltier, K.; Dong, R.; Guarani, V.; Vaites, L. P.; Ordureau, A.; Rad, R.; Erickson, B. K.; Wühr, M.; Chick, J.; Zhai, B.; Kolippakkam, D.; Mintseris, J.; Obar, R. A.; Harris, T.; Artavanis-Tsakonas, S.; Sowa, M. E.; De Camilli, P.; Paulo, J. A.; Harper, J. W.; Gygi, S. P., The BioPlex Network: A Systematic Exploration of the Human Interactome. *Cell* **2015,** *162* (2), 425-440.
12. Li, Z.; Ivanov, A. A.; Su, R.; Gonzalez-Pecchi, V.; Qi, Q.; Liu, S.; Webber, P.; McMillan, E.; Rusnak, L.; Pham, C.; Chen, X.; Mo, X.; Revennaugh, B.; Zhou, W.; Marcus, A.; Harati, S.; Chen, X.; Johns, M. A.; White, M. A.; Moreno, C.; Cooper, L. A.; Du, Y.; Khuri, F. R.; Fu, H., The OncoPPi network of cancer-focused protein-protein interactions to inform biological insights and therapeutic strategies. *Nat Commun* **2017,** *8*, 14356.
13. Rual, J.-F.; Venkatesan, K.; Hao, T.; Hirozane-Kishikawa, T.; Dricot, A.; Li, N.; Berriz, G. F.; Gibbons, F. D.; Dreze, M.; Ayivi-Guedehoussou, N.; Klitgord, N.; Simon, C.; Boxem, M.; Milstein, S.; Rosenberg, J.; Goldberg, D. S.; Zhang, L. V.; Wong, S. L.; Franklin, G.; Li, S.; Albala, J. S.; Lim, J.; Fraughton, C.; Llamosas, E.; Cevik, S.; Bex, C.; Lamesch, P.; Sikorski, R. S.; Vandenhaute, J.; Zoghbi, H. Y.; Smolyar, A.; Bosak, S.; Sequerra, R.; Doucette-Stamm, L.; Cusick, M. E.; Hill, D. E.; Roth, F. P.; Vidal, M., Towards a proteome-scale map of the human protein–protein interaction network. *Nature* **2005,** *437* (7062), 1173-1178.
14. Tang, X.; Munske, G. R.; Siems, W. F.; Bruce, J. E., Mass Spectrometry Identifiable Cross-Linking Strategy for Studying Protein−Protein Interactions. *Analytical Chemistry* **2005,** *77* (1), 311-318.
15. Chavez, J. D.; Bruce, J. E., Chemical cross-linking with mass spectrometry: a tool for systems structural biology. *Curr Opin Chem Biol* **2019,** *48*, 8-18.
16. Leitner, A.; Faini, M.; Stengel, F.; Aebersold, R., Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends in Biochemical Sciences* **2016,** *41* (1), 20-32.
17. Holding, A. N., XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods* **2015,** *89*, 54-63.
18. Tan, D.; Li, Q.; Zhang, M.-J.; Liu, C.; Ma, C.; Zhang, P.; Ding, Y.-H.; Fan, S.-B.; Tao, L.; Yang, B.; Li, X.; Ma, S.; Liu, J.; Feng, B.; Liu, X.; Wang, H.-W.; He, S.-M.; Gao, N.; Ye, K.; Dong, M.-Q.; Lei, X., Trifunctional cross-linker for mapping protein-protein interaction networks and comparing protein conformational states. *Elife* **2016,** *5*, e12509.
19. Sinz, A., Cross-Linking/Mass Spectrometry for Studying Protein Structures and Protein-Protein Interactions: Where Are We Now and Where Should We Go from Here? *Angew Chem Int Ed Engl* **2018,** *57* (22), 6390-6396.
20. Petrotchenko, E. V.; Serpa, J. J.; Borchers, C. H., An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol Cell Proteomics* **2011,** *10* (2), M110.001420.
21. Chakrabarty, J. K.; Naik, A. G.; Fessler, M. B.; Munske, G. R.; Chowdhury, S. M., Differential Tandem Mass Spectrometry-Based Cross-Linker: A New Approach

for High Confidence in Identifying Protein Cross-Linking. *Analytical Chemistry* **2016,** *88* (20), 10215-10222.
22. Yugandhar, K.; Wang, T.-Y.; Wierbowski, S. D.; Shayhidin, E. E.; Yu, H., Structure-based validation can drastically under-estimate error rate in proteome-wide cross-linking mass spectrometry studies. *bioRxiv* **2019**, 617654.
23. Ser, Z.; Cifani, P.; Kentsis, A., Optimized Cross-Linking Mass Spectrometry for in Situ Interaction Proteomics. *J Proteome Res* **2019,** *18* (6), 2545-2558.
24. Wang, X.; Cimermancic, P.; Yu, C.; Schweitzer, A.; Chopra, N.; Engel, J. L.; Greenberg, C.; Huszagh, A. S.; Beck, F.; Sakata, E.; Yang, Y.; Novitsky, E. J.; Leitner, A.; Nanni, P.; Kahraman, A.; Guo, X.; Dixon, J. E.; Rychnovsky, S. D.; Aebersold, R.; Baumeister, W.; Sali, A.; Huang, L., Molecular Details Underlying Dynamic Structures and Regulation of the Human 26S Proteasome. *Mol Cell Proteomics* **2017,** *16* (5), 840-854.
25. Singh, J.; Ponnaiyan, S.; Gieselmann, V.; Winter, D., Generation of Antibodies Targeting Cleavable Cross-Linkers. *Anal Chem* **2021,** *93* (8), 3762-3769.
26. Kao, A.; Chiu, C.-l.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L., Development of a Novel Cross-linking Strategy for Fast and Accurate Identification of Cross-linked Peptides of Protein Complexes. *Molecular & Cellular Proteomics : MCP* **2011,** *10* (1), M110.002212.
27. Burke, A. M.; Kandur, W.; Novitsky, E. J.; Kaake, R. M.; Yu, C.; Kao, A.; Vellucci, D.; Huang, L.; Rychnovsky, S. D., Synthesis of two new enrichable and MS-cleavable cross-linkers to define protein-protein interactions by mass spectrometry. *Org Biomol Chem* **2015,** *13* (17), 5030-7.
28. Matzinger, M.; Kandioller, W.; Doppler, P.; Heiss, E. H.; Mechtler, K., Fast and Highly Efficient Affinity Enrichment of Azide-A-DSBSO Cross-Linked Peptides. *Journal of Proteome Research* **2020,** *19* (5), 2071-2079.
29. Clifford-Nunn, B.; Showalter, H. D. H.; Andrews, P. C., Quaternary Diamines as Mass Spectrometry Cleavable Crosslinkers for Protein Interactions. *Journal of The American Society for Mass Spectrometry* **2012,** *23* (2), 201-212.
30. Iacobucci, C.; Götze, M.; Piotrowski, C.; Arlt, C.; Rehkamp, A.; Ihling, C.; Hage, C.; Sinz, A., Carboxyl-Photo-Reactive MS-Cleavable Cross-Linkers: Unveiling a Hidden Aspect of Diazirine-Based Reagents. *Analytical Chemistry* **2018,** *90* (4), 2805-2809.
31. Gutierrez, C. B.; Block, S. A.; Yu, C.; Soohoo, S. M.; Huszagh, A. S.; Rychnovsky, S. D.; Huang, L., Development of a Novel Sulfoxide-Containing MS-Cleavable Homobifunctional Cysteine-Reactive Cross-Linker for Studying Protein–Protein Interactions. *Analytical Chemistry* **2018,** *90* (12), 7600-7607.
32. Belsom, A.; Mudd, G.; Giese, S.; Auer, M.; Rappsilber, J., Complementary Benzophenone Cross-Linking/Mass Spectrometry Photochemistry. *Analytical Chemistry* **2017,** *89* (10), 5319-5324.
33. Baslé, E.; Joubert, N.; Pucheault, M., Protein Chemical Modification on Endogenous Amino Acids. *Chemistry & Biology* **2010,** *17* (3), 213-227.
34. Leriche, G.; Chisholm, L.; Wagner, A., Cleavable linkers in chemical biology. *Bioorganic & Medicinal Chemistry* **2012,** *20* (2), 571-582.
35. Jones, A. X.; Cao, Y.; Tang, Y.-L.; Wang, J.-H.; Ding, Y.-H.; Tan, H.; Chen,

Z.-L.; Fang, R.-Q.; Yin, J.; Chen, R.-C.; Zhu, X.; She, Y.; Huang, N.; Shao, F.; Ye, K.; Sun, R.-X.; He, S.-M.; Lei, X.; Dong, M.-Q., Improving mass spectrometry analysis of protein structures with arginine-selective chemical cross-linkers. *Nature Communications* **2019,** *10* (1), 3911.
36. Fioramonte, M.; de Jesus, H. C. R.; Ferrari, A. J. R.; Lima, D. B.; Drekener, R. L.; Correia, C. R. D.; Oliveira, L. G.; Neves-Ferreira, A. G. d. C.; Carvalho, P. C.; Gozzo, F. C., XPlex: An Effective, Multiplex Cross-Linking Chemistry for Acidic Residues. *Analytical Chemistry* **2018,** *90* (10), 6043-6050.
37. Bicker, K. L.; Subramanian, V.; Chumanevich, A. A.; Hofseth, L. J.; Thompson, P. R., Seeing Citrulline: Development of a Phenylglyoxal-Based Probe To Visualize Protein Citrullination. *Journal of the American Chemical Society* **2012,** *134* (41), 17015-17018.
38. Yin, X.-G.; Gao, X.-F.; Du, J.-J.; Zhang, X.-K.; Chen, X.-Z.; Wang, J.; Xin, L.-M.; Lei, Z.; Liu, Z.; Guo, J., Preparation of Protein Conjugates via Homobifunctional Diselenoester Cross-Linker. *Organic Letters* **2016,** *18* (22), 5796-5799.
39. Chumsae, C.; Gifford, K.; Lian, W.; Liu, H.; Radziejewski, C. H.; Zhou, Z. S., Arginine modifications by methylglyoxal: discovery in a recombinant monoclonal antibody and contribution to acidic species. *Analytical chemistry* **2013,** *85* (23), 11401-11409.