UNIVERSITY OF CALIFORNIA
RIVERSIDE

Early Drug Discovery Approaches Against the SUMOylation System and Down Syndrome Related Protein

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Cell, Molecular, and Developmental Biology

by

Jingyi Xu

September 2023

Dissertation Committee:
Dr. Jeff Perry, Chairperson
Dr. Jiayu Liao
Dr. Paul Larsen

The Dissertation of Jingyi Xu is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I am filled with gratitude for everyone that have helped me throughout my Ph.D. program. I would like to express my deepest appreciation to my graduate advisor and major professor, Dr. Jeff Perry for his continuous support, encouragement, patience, and knowledge towards my research. Dr. Perry always encourages me when I questioned myself and shows me the positive side from my experiments, even from the failed experiments. I truly wish him good luck with all the exciting research in the lab.

I would like to express my gratitude to my committee members: Dr. Jiayu Liao and Dr. Paul Larsen. Dr. Liao and Dr. Larsen have served in my committee throughout my Ph.D. program and always encouraged me in difficult times. I thank for their time, insightful comments, and inspiring questions.

I am also grateful to my collaborators who have contributed their time. My collaborator at UCR, Dr. Liao and the former members in the Liao Lab, Dr. Vipul Madahar provided me important samples to support my research of Pias1. My collaborator at UCI, Dr. David Mobley and the members in the Mobley Lab provided me the exciting docking molecules for Uba2 and Dyrk1a projects. My collaborator at City of Hope, the TGen group provided me valuable mass spectrometry data for my research of Ubc9. Dr. Holly Yin and Dr. David Kwon from the HTS group spent a lot of time troubleshooting the HTRF assay and supported my research of Uba2. Dr. Li Du from Dr. Steven Rosen's lab provided insightful suggestions on the HTRF assay as well.

I could not have undertaken this journey without people from the CMDB dept, UCR and people from the Perry Lab. Dr. Morris Maduro has helped all the grad students

throughout the Ph.D. program. Dr. Xuemei Chen and Dr. Ted Karginov have been great instructors throughout the grad courses. I am thankful to Xueyan Xu, Ariana Hardy, Ann Song, Ruthia Soh from the CMDB cohort, for their supports and the friendship we have built. I wish the best in their future endeavors. I am extremely grateful to Dr. Stephanie Dingwall. She has always been a great instructor when I work as teaching assistant for her. I am very thankful for her help and friendship. There are so many great scientists in the Perry Lab, and I would like to extend my sincere thanks to my labmates Dr. Jennifer Jossart and Desiree Aispuro, and my undergrad student Paul Stolarov. It has truly been an honor working with them and I wish them the best in their future endeavors.

Special thanks should also go to my dearest friends: Danni Wu, Wen Xie, Yanming Xue, Danping Wang, and Xikun Wu. My friends from my master program at CSU LA have been extremely supportive as well: Dr. Robert Nissen, Mina Yousefelahiyeh, Estibaliz Alvarado, and Eric Medina. I am very happy to see all of them being very successful in their career.

Last but certainly not the least, I am extremely grateful to my family. My parents and all my family members have been extremely supportive, emotionally, and financially, through my academic career. They have always been my biggest support system. Without my partner James Randolph and our pets (Garrus, Tali, and Ashely) and their support during many of the dark moments, I could not get this far. I wish my partner the best in his future research and career.

DEDICATION

To my loving parents. I love you deeply with all my heart and thank you for your unconditional love.

This work is also dedicated to my dearest friend, Xikun. I hope you are proud of me.

ABSTRACT OF THE DISSERTATION

Early Drug Discovery Approaches Against the SUMOylation System and Down Syndrome Related Protein

by

Jingyi Xu

Doctor of Philosophy, Graduate Program in Cell, Molecular, and Developmental Biology
University of California, Riverside, September 2023
Dr. Jeff Perry, Chairperson

Drug discovery is a costly and time-consuming process. It has four major stages: early drug discovery, pre-clinical phase, clinical trials, and regulatory approval. Early drug discovery is the steps of target validation, hit identification (hit ID), hit-to-lead, and lead optimization for new chemical entities (NCEs) and therapeutic targets. Basic research identifies and validates disease-relevant targets. Hit ID is the process of testing a group of compounds against a target protein. Thermal shift assay (TSA) can detect slight changes in protein melting point and stability, and, therefore, be used to screen for hit ID. X-ray crystallography studies can determine the binding between target protein and hits. *In vitro* enzymatic assays can confirm the inhibition of hits and measure IC_{50} (inhibitory concentration 50%). Hit-to-lead can be done by 'SAR by catalog', in which the features of the hits are used to identify commercially available compounds. During the hit-to-lead stage, hits are developed to improve the binding affinity to move along as a final drug.

Structural-based drug discovery (SBDD) is one of the early-stage drug discovery approaches and has been wildly used in academic and industrial settings since the 1990s.

SBDD uses small molecules that follow Lipinski's Rule of Five that ensures that the small molecules can be administered orally once they become drugs. Fragment-based lead discovery (FBLD) evolved from SBDD. Fragments follow the Rule of Three that helps reduce off-target effects and for a broad search of chemical diversity using smaller libraries in the low thousands of compounds. In silico artificial intelligence-computer-aided drug discovery (AI-CADD) is used to identify hits and develop leads using computer tools and the simulation of human intelligence processed by machines.

Post-translational modifications (PTMs) are the last stage in protein biosynthesis, being either reversible or irreversible chemical events occurring after protein translation. One type of PTM is ubiquitin-like protein is the Small-Ubiquitin-Related Modifier (SUMO) modification. SUMOylation of a target protein occurs via an enzyme cascade that involves three steps: First, the SUMO E1 activation enzyme, the Aos1/Uba2 complex, activates matured SUMO protein and form a high-energy thioester bond to the Uba2 subunit. Next, the activated SUMO is transferred to cysteine 93 of Ubc9 (SUMO E2 conjugating enzymes). Ubc9 conjugates SUMO onto the substrates, or SUMO E3 enzyme, such as Pias1, often aids in this conjugation event to increase substrate specificity. SUMOylation has been studied extensively and found involved in cancer as “non-oncogene addiction”. SUMO enzymes are not mutated or classical oncogenes, but cancer cells essentially depend on them for survival. Notably, non-oncogene addiction may provide novel targets for intractable and undruggable cancers. Another target for cancer is the dual-specificity tyrosine phosphorylation-regulated kinase 1A (Dyrk1a). It plays multiple roles in animal development, with critical effects on neuronal cell cycle regulation and cancer

development. We hypothesis that targeting the SUMOylation modification and Dyrk1a will impact cancer cells as therapeutic treatments. Utilizing SBDD, FBDD, and AI-CADD, we have identified a small molecule, AW2 F5, targets the ATP binding pocket and inhibits Aos1/Uba2 with an IC_{50} of 497.7 nM. We have identified a fragment, LC B3, covalently bind to the backside of Ubc9 as a potential degrader. We also have found a small molecule, DTP C1, targets the ATP binding pocket and inhibit Dyrk1a with IC_{50} of 35.7 nM. In the future, each compound will be optimized and tested in cancer cells for their therapeutic effects.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1
A brief introduction of drug discovery

Chapter 1: A brief introduction of drug discovery

1.1 Abstract

The drug discovery process has four major stages: early drug discovery, pre-clinical phase, clinical trials, and regulatory approval. Early drug discovery is the steps of target validation, hit identification (hit ID), hit-to-lead, and lead optimization. Structural-based drug discovery (SBDD) is one of the early drug discovery approaches. SBDD uses small molecules that follow Lipinski's Rule of Five, which ensures that the small molecules can be administered orally once they become drugs. Fragment-based lead discovery (FBLD) evolved from SBDD, which follows the Rule of Three. The rule provides fragments with high ligand efficiencies (LE), allowing for developing drug-like leads that are typically smaller and have simpler chemistries. Covalent fragments are those with warheads that can form covalent bonds with amino acid such as Cys. Covalent fragments/inhibitors form irreversible bonds with the target protein. They potentially have higher potency and longer duration than non-covalent fragments/inhibitors. Artificial intelligence-computer-aided drug discovery (AI-CADD) is used to identify hits and develop leads using computer tools and the simulation of human intelligence processed by machines. AI-CADD is being used more and more in the drug discovery process since it increases the change to pass tests in development with ideal drug properties.

1.2 Drug discovery strategies and steps

Drug discovery is a costly and time-consuming process. It has four major stages: early drug discovery, pre-clinical phase, clinical trials, and regulatory approval [1, 2]. It requires the collaboration of academic labs, research institutions, biotech companies, and pharmaceutical companies (**Figure 1**).

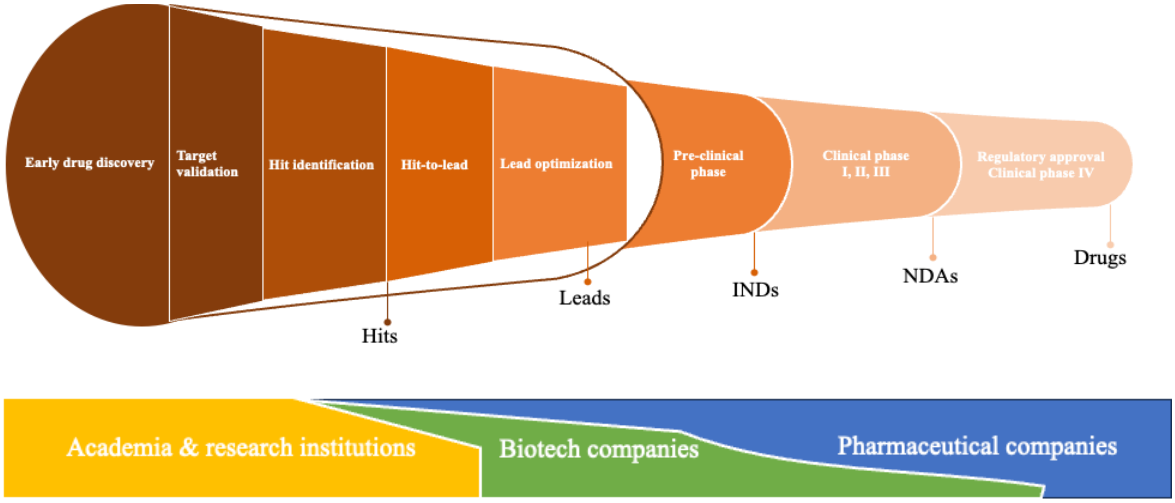


Figure 1. drug discovery pipeline. A summary of drug discovery pipeline. It usually uses more than 10,000 compounds to begin the early drug discovery. About 250 compounds can move to pre-clinic and only 5 pass clinic trials.

Early drug discovery is the steps of target validation, hit identification (hit ID), hit-to-lead, and lead optimization for leads of interest, aiming to create new chemical entities (NCEs) or drugs for therapeutic targets. The leads of interest are then tested in animals or other models with the target diseases in the pre-clinical phase. Once the leads of interest pass the pre-clinical phase, they are identified as investigational new drugs (INDs). Then, the IDNs are tested through clinical trials for drug tolerance, effectiveness, dosage, and interactions with other drugs under current good manufacturing practice (cGMP) by Food

and Drug Administration (FDA) protocol. The clinical trials typically have Phases I, II, and III at this stage, and IV will be done later. The last is regulatory approval when the FDA or similar authorities in other countries approve the new drug application drugs (NDAs) to be sold in the relevant country. Phase IV, the Post-Marketing Surveillance Trial, will test the drugs' long-term effects.

1.2.1 Structural-based drug discovery (SBDD)

Early drug discovery steps are target identification and validation, hit ID, hit-to-lead, and lead optimization. Structural-based drug discovery (SBDD) is one of the early drug discovery approaches. SBDD has been widely used in academic and industrial settings since the 1990s [3], first beginning with the development of HIV-1 inhibitors. SBDD begins by identifying a target protein essential to the progression of a disease pathway [4]. This protein target is then analyzed in detail, either through previously characterized structural data or utilizing bioinformatic predictions, to determine the binding site(s) essential to facilitate the target's function that allows for disease progression [5, 6]. SBDD uses small molecules that follow Lipinski's Rule of Five [7, 8] where: the molecular weight is less than 500 Da, cLogP is less than five, H-bond donors are less than five, and H-bond acceptors are less than 10, polar surface area is less than 140 Å, and rotatable bond are less than 10. Rule of Five ensures that the small molecules can be administered orally once they become drugs.

1.2.2 Fragment-based lead discovery (FBLD)

Fragment-based lead discovery (FBLD) evolved from SBDD. It has several advantages in drug discovery, resulting in 6 FDA-approved drugs (Vemurafenib approved in 2011 [9],

Venetoclax approved in 2016 [10], Erdafitinib approved in 2019 [11], Pexidartinib approved in 2020 [12], Sotorasib approved in 2021 [13], Asciminib approved in 2021 [14]), and more than 50 drugs in clinical trials [15] since the first report of FBLD in 1996 [16]. Fragments follow the Rule of Three [17], where molecule weight is less than 300 Da (100-250 Da original), cLogP is less than three, and H-bond acceptors are less than three. The Rule of Three ensures that the fragments probe interaction with the targets with minimal change of unfavorable interaction [18]. The rule provides fragments with high ligand efficiencies (LE), allowing for developing drug-like leads that are typically smaller and have simpler chemistries. High LE is a good starting point for designing leads. Fragments have less molecular complexity than small molecules; thus, a fragment hit may provide a pathway to drugs with simpler chemistry and high specificity to limit off-site toxicities. Due to the small size of fragments, their binding affinity/impact on the targets in activity assays is low. Fortunately, X-ray crystallography efficiently detects weak interaction (mM – μM). A few thousand highly diverse fragments can provide the equivalent diversity of a few million compounds with high throughput (HTS) screens, with the advantages of far fewer chemicals to screen through and reduced costs in acquiring and storing the libraries. Therefore, FBLD allows for a broad search of chemical diversity using smaller libraries in the low thousands of compounds. FBLD is regarded as a bottom-up approach. Fragment screening can reveal protein hot spots, new binding pockets, allostery, and protein-protein interaction sites. Once a fragment hit is identified, it can be developed into leads and new chemical entities (NCEs) in new areas of IP space [19, 20].

Fragment libraries are designed with different functions. For example, covalent fragments are those with warheads that can form covalent bonds with amino acid residues, for example, Cys, Lys, Ser, His, and Tyr. Covalent fragments are acrylamides, sulfonate esters, thioureas and thiones, Michael acceptors, and other types. Covalent fragments/inhibitors form irreversible bonds with the target protein. They potentially have higher potency and longer duration than non-covalent fragments/inhibitors [21, 22]. When mutations occur, covalent fragments/inhibitors can still bind to the target protein if the mutations do not impact the binding residues. However, covalent fragments/inhibitors might have high toxicity if they bind to off-target proteins [23]. Therefore, the tests to ensure the selectivity of covalent fragments are crucial in the early drug discovery stage. Non-covalent fragments may have very weak electron density, making them difficult to locate in X-ray crystallography studies. On the other hand, brominated fragments can overcome this difficulty since bromine is easy to detect using anomalous scattering with its heavy electron density. Pan-Assay-Interference Compounds (PAINS) are fragments known to react with many proteins and would show false positive results [24]. They are removed from all fragment libraries.

Successful fragment hits will be modified during lead optimization by three main strategies: linking, growing, and merging [25]. From the first linking by structure-activity relationship (SAR) study [16], linking follows five main steps: 1. The first fragment hit ID with the target protein; 2. First hit optimization; 3. The second fragment hit ID with the target protein with the first optimized hit; 4. The second fragment hit optimization; 5. Link the two optimized hits. Growing can be done by SAR by catalog, i.e., Astex [26]. It utilizes

the fragment hit of the target protein and searches for larger molecules that contain the fragment hit / the part binding to the target protein. The larger molecules will be tested against the target protein. Merging uses two fragments as linking. However, the two fragments bind to two different but close enough sites of the target protein in linking; the two fragments bind to overlapping sites of the target protein [27]. If the two fragment hits (optimized) merge, it may yield a bigger and more potent molecule, i.e., Vernalis [28].

1.2.3 Target identification and validation

Basic research identifies and validates disease-relevant targets, which can be proteins, RNA, or genes [1]. Good targets should be safe and "druggable". Therefore, it is vital to determine if the targets are "druggable" in early drug discovery. Since SBDD and FBLD rely on purified and homogeneous target proteins, it could be problematic if post-translational modification and other proteins impact the binding sites of target proteins [25].

In the early drug discovery, the target protein will be purified. With the plasmids with the gene of protein of interest, using expression systems such as bacterial (*E. coli*) or yeast, etc., as host, protein will be expressed and then be purified, i.e., by fast protein liquid chromatography (FPLC). The protein purification strategies are to capture the tag attached to the protein using affinity columns to separate the protein of interest from others and further purify based on the target protein's charge and size using ion exchange and size exclusive columns.

1.2.4 Hit identification (Hit ID)

Hit ID is the process of screening a group of compounds against a target protein. The output of the screening is hits that can inhibit the target protein. The target protein is purified first. The screening methods/strategies include but are not limited to HTS, fragment screening, structural aided drug design, and virtual screening [1]. HTS can detect the target protein's activity with compounds but will need a secondary assay to determine the binding sites [29]. Fragment screening uses fragments with high concentrations. The benefits of using fragments have been talked about in the FBLD section. Structural-aided drug design uses the crystal structure of the target protein and then docks the compounds into it. Regardless of the methods or assays, the Z-factor, a parameter to determine if a screening assay is suitable for hit identification, needs to be determined first. Any value between 0.5 to 1 is acceptable.

Docking can predict and modify the compounds to increase potency. Virtual screening can build compound libraries for the target protein. Structure-based virtual screening (SBVS) and ligand-based virtual screening (LBVS) are the two major strategies of virtual screening [30]. One of the strategies of SBVS is molecular docking, which predicts the pose of compounds with the binding site of the target protein using its crystal structure. LBVS uses chemical databases and similarity principles. Using one of the known binders of the target protein, LBVS can search for other compounds with similar basic scaffolds but better biochemical activity [31, 32], which refers to the scaffold hopping concept. LBVS, such as machine learning, can also be used during lead optimization.

Thermal shift assay (TSA), also known as differential scanning fluorimetry, DSF, can detect slight changes in protein melting point and stability, and, therefore, be used to screen for hit ID. TSA is a rapid fluorescence-based assay that can help identify potential binders by their effect on protein thermal stability in solution [33]; this significantly reduces the hit ID time. TSA indirectly identifies compound-protein interaction by measuring the protein's thermal stability (T_m). T_m is the temperature when 50 % of protein is denatured. When the target protein is fully folded, SYPRO Orange is surrounded by water, and fluorescence is reduced by water quenching. The target protein is heated and unfolded when the real-time PCR system raises the temperature. When the target protein starts unfolding, SYPRO Orange dye binds to any hydrophobic residues of the target protein. Then, fluorescence has less water quenching and can be detected by real-time PCR systems. The more hydrophobic residues expose / the more protein unfolding, the higher signal SYPRO Orange can send. SYPRO Orange will fluoresce, and the real-time PCR system detects the signal and calculates T_m. If a compound binds and stabilizes the protein, causing T_m increases (positively ΔT_m) by at least 0.5 °C, then the compound is defined as a hit. The hit rate of screening helps determine if the protein target of interest is "druggable". For example, in an FBLD screen, a less than 1% hit rate (1% of compounds causing ΔT_m 0.5 °C) suggests that the protein is likely undruggable and 3% or higher infers highly druggable [34]. There may be some false positives (and false negatives) from the screening. Initial TSA screening identifies hits (hits at this stage are binders/shifters of the target protein), and then a dosage test is used to confirm binding, where a higher concentration of the hit should result in higher ΔT_m, or vice versa.

The hits are then tested/confirmed by protein structural analysis to characterize hit-protein interactions. X-ray crystallography studies can determine the binding between target protein and compound. Coupled with SAR and virtual screening, X-ray crystallography will used to involve hits in the hit-to-lead stage. Other assays, such as mass spectrometry and activity assays, can detect covalent fragment or small molecule-protein interactions as well. Therefore, covalent fragments are of strong interest to an early-stage drug discovery project.

In vitro enzymatic assays can confirm that the hits identified by TSA are both binders and inhibitors. It tests hits affinity/inhibition efficiency and the desired therapeutic effect [35], too. Such an assay could be a fluorescence resonance energy transfer (FRET)-based assay, luminescent signal-based assay, etc., and this step may require new assay development. Potencies such as IC_{50} (inhibitory concentration 50%) can be measured during this step and are expected in the μM range. IC_{50} is the concentration of an inhibitor that can reduce the activity of the target in the assay by 50%; therefore, it is assay specific. Low IC_{50} indicates that the inhibitor is potent at low concentrations and will show lower systemic toxicity during medication administration [36].

Binding studies such as isothermal calorimetry (ITC) or surface plasmon resonance (SPR) can measure K_i (inhibition constant) and K_d (dissociation constant). K_i and K_d describe binding affinity, so they are more precise than IC_{50} [37, 38]. Low K_i and K_d indicate strong binding and high binding affinity. However, a very tight binder could be an allosteric activator of the target. So, we need to measure IC50 as it represents the function

of the compound, which is how well the compound inhibits the target. We are looking for the inhibitor with low IC_{50} (high inhibition) and low K_d (strong binding).

1.2.5 Hit-to-lead

During the hit-to-lead stage, hits are developed to improve the binding affinity to move along as a final drug. Hit-to-lead can be done by 'SAR by catalog', in which the features of the hits are used to identify commercially available compounds. Hits can be linked together, growing, or merging. In silico docking can also be conducted based on the hits to search for leads. Potential leads are tested against the protein of interest, and only the best ones move to the next step that is lead optimization, in which the potency should be in the nM range.

Another way to improve the hits in silico is artificial intelligence-computer-aided drug discovery (AI-CADD). AI-CADD is used to identify hits and develop leads using computer tools and the simulation of human intelligence processed by machines [39]. CADD has already been used in drug discovery as it reduces time and expense. Combining AI and deep learning (DL), AI-CADD can screen and identify the most promising compounds at a binding site of a target protein from a library of several million compounds. The multilayered artificial neural network utilizes numerous layers of 'neurons' composed of hundreds or thousands of neurons to screen if the compound interacts with the target protein and predict properties such as ADMET [40], detail in 1.2.7. The docking compounds using AI-CADD are not only against to the target protein but also have ideal drug properties. They have relatively higher change to pass tests in development. Hence, AI-CADD is being used more and more in the drug discovery process.

1.2.6 Lead optimization

Leads are generated and further optimized for better drug properties such as effectiveness and absorption. The same methods or bioassays during hit ID can be used to test leads. The last stage in the early drug discovery / SBDD coupled with FBLD and AI-CADD is testing the leads by *in vivo* assays. The bioavailability will be addressed while maintaining selectivity and potency, i.e., if the lead has oral bioavailability. If any leads fail in the toxic screen, the leads will not be candidates. Back-up leads are also necessary if the preferred lead fails in later pre-clinical phase and clinic trials [1]. From the hit ID to lead optimization, the testing compounds number is from several thousand or even million down to a few candidates for the next step.

1.2.7 Pre-clinical phase

The pre-clinical phase is for drug metabolism and pharmacokinetics (DMPK). DMPK influences the drug candidates from pre-clinical to clinical trials [41], and it describes the drug candidates from ADME(T) (absorption, distribution, metabolism, and excretion) and toxicity and pharmacokinetics (PK). The ADME concept has derived from resorption, distribution, consumption, and elimination [42 & 43] and then became a standard term in drug regulation [44]. It describes the drugs entering the body (absorption), moving in the body (distribution), changing in the body (metabolism), and leaving the body (excretion). Absorption is how drugs enter the body from administration to the desired locations [45]. The administration can be oral, intravenous, intramuscular, etc. The desired locations of absorption are tissues, bloodstream, and stomach (especially vital for oral drugs) anatomically, cellular, and molecular. Bioavailability is one of the major factors of

determining administration based on the properties of the drug in question. For example, suppose an oral medication is processed in large amount by the liver or gut wall and leaves a low amount in circulation, the bioavailability of the medication would be low. Distribution is how drugs spread in the body, and diffusion and convection in the body are the two major factors that influence it, affected by the polarity, size, and binding affinity of drugs. Metabolism is how drugs are inactivated and converted to more hydrophilic subsequent compounds or how pro-drugs are activated and converted to the active parent drugs [46]. According to the data from Phase I and II, the common areas in the body where drugs metabolize are the liver, gastrointestinal tract, skin, kidney, etc. Excretion is the last step in which drugs are removed from the body. When examining excretion and clearance, the ratio of the elimination rate is calculated. The ratio is influenced by the drug and the patients, i.e., the kidney status. Toxicity is a major factor that eliminates drug candidates [47]. On-target toxicity, off-target toxicity, immune hypersensitivity, and bioactivation can contribute to drug toxicity. The development of short-term assays for drug toxicity screening has been the goal. However, pre-clinical toxicity tests in animals and advanced events in human are still the main test for toxicity.

Pharmacodynamics (PD) is the effect of the drug candidates in the body, while PK determines how the drug candidates interact with the body. PK-PD modeling can be used to guide the formula and dosage design during development [48]. The benefit is that the drug candidates themselves and the dosage can be connected to the physiological response in the body. During the pre-clinical (nonclinical) laboratory studies, the FDA required good laboratory practices (GLP) (CFR - Code of Federal Regulations Title 21, n.d.). GLP

provides the requirements for study conduct, personnel, facilities, equipment, written protocols, etc.

1.2.8 Clinic trials and registration

Gathering the information from FDA, Phase I is usually 20 – 100 healthy volunteers with the disease and tests the safety and dosage of the INDs. Low doses and high specificity are needed to meet the safety and efficacy requirements. Phase I can take several months and about 70% of drugs can move to Phase II, which is about several hundred patients with the disease and to further test the efficacy and side effects of the drug from several months up to two years. Phase III is hundreds to thousands of patients with the disease and to further test efficacy. Finally, NDAs are allowed to enter the market, and Phase IV can start. Unlike other phases in the clinical trial, several thousand patients are involved in Phase IV, and more comprehensive data will be collected. Therefore, the long-term benefits, side effects, safety, and even interaction with other drugs will be addressed.

About 90% of INDs fail in clinical trials [2]. They fail mainly because of low efficacy, high toxicity, and poor drug-like properties. The best lead must be selected to increase the chance of any INDs passing the clinic trials. Therefore, high standard must apply to every step in early drug discovery. For example, we ensure that the fragments and AI-CADD small molecules have proper drug-like properties when we build the libraries. During the hit ID, we utilize TSA to identify binders/hits and confirm the hits with dosage test. We utilize *in vitro* enzymatic activity assay to confirm the inhibition of the hits to the target proteins. If possible, we use more than one activity assay. We determine the binding site and pattern of the hits and the target proteins by perform crystallography studies. We select

the ones with the lowest IC_{50} from the hits, and then perform SAR with virtual docking in hit-to-lead. For those target proteins containing high homologues structures with other proteins, we test the hits with the off-target to rule out off-target effect. We ensure that the hits with high efficacy, and therefore low systemic toxicity move to the next step. Hits are future tested by binding assays for potency measurement and *in vivo* tests in lead optimization. We ensure the best leads during the rigorous early drug discovery so that to increase the chance of passing the pre-clinic test and clinic trials.

Chapter 2
SUMOylation and SUMO E1 early drug discovery

Chapter 2: SUMOylation and SUMO E1 early drug discovery

2.1 Abstract

One type of Post-translational modifications, SUMOylation plays a vital role in the regulation of protein stability, activity, and localization. SUMOylation of a target protein occurs via an enzyme cascade that involves three steps: First, the SUMO E1 activation enzyme, the Aos1/Uba2 complex activates matured SUMO protein. Next, the activated SUMO is transferred to cysteine 93 of Ubc9, and Ubc9 conjugates SUMO onto the substrates. SUMO E3 ligase enzyme, such as Pias1, often aids in this conjugation event to increase substrate specificity. Since the first time SUMO1 was found to bind RanGAP1 in the nuclear pore complex protein in 1997, SUMOylation has been studied extensively and found involved in a variety of diseases, such as cancer, human heart failure, atherosclerosis, diabetes, immunology, seizure, and even sudden death. SUMOylation function in human tumorigenesis, as non-oncogene addiction, as it can alter a protein's structure and function, therefore inducing cell proliferation, impacting apoptosis, and metastasis. Uba2 is a synthetic lethal partner of Myc, one of the undruggable oncogenes. Myc-high breast cancer patients show significantly higher metastasis-free survival rate, if they have low level of Uba2, compared with high Uba2-level patients. UBA2 gene is amplified in epithelial cells from a fibrocystic breast and Uba2 protein level is higher in non-small cell lung cancer than normal tissue. There are only a couple of non-clinical inhibitors of Uba2. In this chapter, In this chapter, we utilize SBDD coupled with FBLD and AI-CADD to identify inhibitors of Aos1/Uba2. Our overarching hypothesis is that said inhibitors will inhibit the SUMOylation and impact Myc-driven cancer cells.

2.2 PTM & SUMOylation pathway and function

Post-translational modifications (PTMs) are the last stage in protein biosynthesis, being either reversible or irreversible chemical events occurring after protein translation [49, 50]. PTMs can change protein stability, conformation, folding, localization, and protein-protein interactions. PTMs also regulate protein functions, metabolism, signal transduction, and cellular physiological status [51, 52, 53]—PTMs impact proteins' complex formation and all aspects of protein function.

One type of PTM is the ubiquitin-like modification. The first studied was the Ran-GTPase-activating protein (RanGAP1) regulated by ubiquitin-like modification [54]. Now we know this ubiquitin-like protein is the Small-Ubiquitin-Related Modifier (SUMO), and the modification by SUMO is SUMOylation or SUMO modification. SUMO proteins are highly conserved [55, 56, 57], and there are five proteins in this group (SUMO1-5) that are added to specific target proteins [54, 58, 59, 60, 61]. Although the amino acid sequence of SUMO isoforms is only 18% homology to ubiquitin, another well-studied PTM protein, they are both formed from a ββαββαβ fold [62]. The SUMO precursors contain a flexible N-terminal, a ubiquitin fold, and a short C-terminal tail. SUMO proteases (i.e., SENP1 or SENP2) cleave off four amino acids from the C-terminal tail to reveal a di-Glycine at the C-terminal tail and convert the precursors to their mature forms [64]. All SUMO isoforms modulate target proteins with a ΨK(Q/T)E (or ΨKxD/E for SUMO2) consensus motif, where Ψ is a hydrophobic residue and K reversibly conjugates to SUMO proteins. SUMO2 and SUMO3 are highly homologous and functionally redundant. SUMO2 and SUMO3 are 97% identical to each other, and 46% identical to SUMO1. There is no antibody specific

for SUMO2 and SUMO3. Thus, they are referred to as SUMO2/3. SUMO1 [65] or SUMO3 [66] null mice are viable and have normal development, but SUMO2 null mice die at embryonic day 10 because of reduced cell proliferation. Data suggest that SUMO2 is the most important SUMO isoform due to SUMO2 higher expression level than the other two. SUMO1 and SUMO2/3 are readily expressed in human tissues, but SUMO4 is only expressed in the kidney and has not been studied as well as the others. SUMO4 is stable only under serum-starved condition and modify the stress response proteins [67]. SUMO5 mediates promyelocytic leukemia nuclear bodies and cell growth [68]. Like ubiquitinoylation, SUMOylation can modify substrate proteins in either a mono- or a poly-SUMO manner. SUMO1 can form the mono-SUMO on the target protein, while SUMO2/3 can form mono-SUMO and poly-SUMO chains, as SUMO2/3 contains a conserved Lysine 11 for poly-SUMO, but SUMO1 does not. Data show that the level of SUMOylation alters substrate proteins' function during cellular events [69, 70]. The related ubiquitin moiety was first identified as a marker for protein degradation, while SUMOylation functions have been shown to be involved in the regulation of protein stability, activity, and localization [71, 72, 73].

SUMOylation of a target protein occurs via an enzyme cascade that involves three steps: First, the SUMO E1 activation enzyme, ubiquitin-like 1 activating enzyme (Uba2)/activator of SUMO1(Aos1) complex, serves to bind ATP and activates matured SUMO protein at the C-terminal carboxyl group [74]. This step forms SUMO-adenylate using ATP. The activated SUMO has then transferred to cysteine 113 residue of Uba2 to form a high-energy thioester bond to the Uba2 subunit, with the release of AMP. The

activity site of SUMO E1 is Cys173 of Uba2. Uba2 contains an adenylation domain, a catalytic domain, a ubiquitin-like domain, and a nuclear localization signal at the C-terminal [75, 76]. Inactive Aos1/Uba2 and failed to active SUMO lead to a globally reduced SUMOylation in the cells [77]. Next, the activated SUMO is transferred to cysteine 93 of Ubc9 (ubiquitin conjugating enzyme 9, the only human SUMO E2 conjugating enzymes). Ubc9 interacts with the ubiquitin fold domain of Aos1/Uba2 and then forms a thioester bond with SUMO. Although there is a negatively charged ridge next to the Cys93 of Ubc9, the surface of Cys93 of Ubc9 is positively charged, which interacts with the negatively charged C-terminal of SUMO [78]. A Lys101 residue next to Cys93 generates a positive charge to mask off the negative charge from Cys93. Then, Ubc9 conjugates SUMO onto the substrates [79]. Many of the substrates contain the SUMO consensus motif, and the motif is required for Ubc9 binding. The knockoff Ubc9 is embryonal lethal, but haplo-deficiency is viable [80]. Ubc9 can be degraded by autophagy [81]. Ser71 of Ubc9 can be phosphorylated by Cyclin dependent kinase 1 (CDK1), increasing the thioester bond between Ubc9 and SUMO [82]. SUMO E3 ligase enzyme, such as protein inhibitor of activated STAT 1 (Pias1), often aids in this conjugation event to increase substrate specificity [83, 84]. Coupled with deSUMOylation, which removes SUMO from target proteins catalyzed by SUMO protease family members [85], SUMOylation is reversible and highly dynamic. Six Sentrin/SUMO-specific proteases (SENPs) in mammalian cells can deSUMOylate substates and contain a conserved catalytic domain in the C-terminal. The number of enzymes involved in SUMOylation is less than

ubiquitinylating, but the number of substrates of SUMOylation is large. The specificity of SUMOylation is still unknown.

Functionally, there is a crosstalk between SUMOylation and ubiquitylation, as both can target the same lysine residue [86]. They can compete while ubiquitylation degrades IκBα, but SUMOylation stabilizes it [87]. There are also sequential SUMOylation and then ubiquitylation. SUMO2 modification of cystic fibrosis transmembrane conductance regulator (CFTR) leads to ubiquitylation of CFTR [86]. F508del CFTR (the mutant of cystic fibrosis disease) is modified by SUMO2/3 first, and then SUMO2/3 directs F508del CFTR to the ubiquitin-proteasome pathway via SUMO-targeted ubiquitin ligase (STUbLs). Poly-SUMO chain serve as targeting signals for STUbLs that ubiquitinate the target, typically leading to proteasome degradation [88, 89]. SUMO-interaction motif (SIM) contains hhXSXS/Taaa, in which h represents any hydrophobic amino acid, and a represents any acidic amino acid [90]. STUbLs contain SIMs that bind to the SUMO moiety attached to the protein target and then recruit ubiquitin modification. SUMOylation can also modify the enzyme involved in ubiquitylation and vice versa. For example, SUMOylation can inactivate E2-25k as a ubiquitin-conjugating enzyme [91]. The ubiquitin-proteasome modification degrades SENP2 and SENP3 as SUMO proteases [85].

2.3 Non-oncogene addiction

Since the first time SUMO1 was found to bind RanGAP1 in the nuclear pore complex protein in 1997, SUMOylation has been studied extensively and found involved in a variety of diseases [70], such as cancer [92], heart failure [286], atherosclerosis [93], diabetes [94], immunology [95], seizure [96], and even sudden death [97].

Drugs that can target the activation of oncogenes (e.g., trastuzumab targeting HER-2 the first anti-HER2 therapies), but many of the oncogenes are hard to target (i.e., Myc and Ras as undruggable oncogenes) [98, 99, 100]. Non-oncogene addiction refers to specific genes that are not mutated or classical oncogenes, but cancer cells essentially depend on them for survival. SUMOylation plays a vital role in human tumorigenesis, as non-oncogene addiction, as it can alter a protein's structure and function, therefore inducing cell proliferation, impacting apoptosis, and metastasis [101, 102]. Notably, non-oncogene addiction may provide novel targets for intractable and undruggable cancers [103, 104].

2.4 SUMO E1 in cancer

2.4.1 Myc driven cancer

There are a subset of targets known as the 'undruggable' because they show little impact with or resistance to current treatment regimes, and these undruggable targets include Myc and KRas-driven cancers [105, 106]. The Myc family, including the most common member, proto-oncogene c-Myc, has been found to be mutated in more than 70% of human cancers, including breast, liver, colorectal cancer, and leukemia [107, 108, 109]. Myc has been considered to belong to the "undruggable" group because it lacks a pocket for traditional small molecules to bind [110]. Common mutations of Myc occur at its N-terminal region and T58A. N-terminal mutations cause the induction of cell growth and transformation abnormalities [111]. T58A mutation prevents phosphorylation at this residue and ubiquitinylation of Myc, which subsequently prevents ubiquitin-proteasome degradation of Myc, leading to Myc overexpression [112].

The most common mutation of Myc is an amplification [113]. In acute myeloid leukemia (AML), overexpression or amplification of Myc is caused by AML-associated fusion proteins, such as mutations of Flt3 receptor tyrosine kinase [114]. Overexpressed Myc binds and activates EZH2 to induce cell proliferation. The 5-year survival rate of AML is 25-68% among different ages. In a 2018 study with 132 AML patients, 39% had high Myc protein level, and 61% had low Myc [115], and the low Myc level patients were observed to have a significantly longer overall survival than high Myc patients. In chronic myeloid leukemia (CML), a fusion of *BCR* and *ABL1* genes activates ABL to its oncogenic form, leading to the deregulation of several proteins, including overexpression of Myc [116, 117, 118]. BCR/ABL and overexpression of Myc block differentiation and induce genetic instability. Diagnosed CML patients commonly show high level of Myc mRNA [119]. Among the 66 CML patients, 15 patients (>22%) showed poor responses to imatinib, a front-line tyrosine kinase inhibitor to treat leukemias. Those 15 patients had more than two-fold Myc mRNA expression levels compared to others that showed a response to imatinib. Highly expressed c-Myc with its partner MAX up-regulates BCR/ABL1 expression levels of both mRNA and protein, and depletion of Myc in BCR/ABL1 cell lines leads to cell death [120].

The 5-year survival rate of CML has been improved to 70% with advanced treatments from 2011 to 2017 using chemotherapy in combination with drugs targeting cells with the Philadephia chromosome [121]. However, these long-term treatments significantly impact patients' everyday lives and life spans, as patients need at least three years of treatment and must be closely monitored after stopping treatment. After decades research, two inhibitors

have been found might affecting Myc, a peptide OmoMYC [122] and a small molecule APTO-253 [123]. OmoMYC has complete the phase I clinical trial at the end of 2022. APTO-253 has been terminated in phase I. ATPO-253 had a solubility problem first and then failed to show clinical response in the phase I.

Myc-dependent tumor growth requires SUMO activating enzyme E1 Aos1/Uba2 [124]. Inhibition of Uba2 enzymatic activity represses a transcriptional subprogram of Myc and leads to Myc-high breast cancer cell death. Myc-high breast cancer patients show significantly higher metastasis-free survival rate, if they have low level of Uba2, compared with high Uba2-level patients. In more detail, when Uba2 is knocked down in Myc-high cancer cells, these cells show spindle defects that caused cell apoptosis. Therefore, Uba2 is a synthetic lethal partner of Myc. (Synthetic lethality is a term from genetic studies referring that a mutation in gene A only cause lethality combined with a mutation in gene B, where mutations in gene A and B would not be lethal individually). SUMOylation is required for the stability of MYC [190, 195, 196], which will be discussed in section 3.7.1.

2.4.2 Breast cancer

Breast cancer is the most diagnosed cancer globally (followed by lung and prostate cancers), with about 2.3 million new cases in both sexes [125, 126]. Breast cancers must be diagnosed with physical examination and screening. In many cases the tumor is already spread through blood or lymph when the patient is diagnosed [128, 129]. The standard treatments for breast cancer are surgical resection and chemotherapy [130], but the distant breast cancer 5-year survival rate is still low at only 30% [121]. RT-qPCR data show that the UBA2 gene is amplified in the MCF10 model, an epithelial cell from a fibrocystic

breast [131]. Monensin, a polyether ionophore antibiotic that can impact Na^+ and K^+ concentration gradient in the cell and the physiological function, has been shown to reduce the proliferation of cancers such as pancreatic and prostate cancer [132]. Using the same MCF10 model, Monensin has been found to inhibit MCF10 cell survival rate by apoptosis [129]. Moreover, RT-qPCR and western blot data show that Uba2 mRNA and protein level are significantly decreased in Monensin treated MCF10 cells. Overexpression of Uba2 inhibits the impact of Monensin on MCF10 cells. Thus, inhibition of Uba2 in breast cancer is shown as a potential anti-breast cancer therapeutic target.

2.4.3 Lung cancer

Lung cancer is the leading cause of cancer death and the second most common cancer, with about 2 million new cases in 2020 [125, 133]. Non-small cell lung cancer (NSCLC) is the most common type, for about 81% of lung cancer diagnoses. The 5-year survival rate of localized NSCLC is 65%. However, like breast cancer, about 70% of patients are diagnosed with advanced NSCLC, and the 5-year survival rate drops to 37% for regional NSCLC and 9% for metastatic NSCLC.

Including Uba2, SUMOylation enzymes are highly expressed in NSCLC [134]. Immunostaining of Uba2 in the nucleus is significantly higher in NSCLC than in normal tissues [135]. Overexpression of Uba2 in NSCLC tissue is linked to large size and metastasis. In A549, the cancerous lung cell line, shUBA2 downregulates Uba2 RNA and protein levels. A549 cells show an increase of G1 and G2/M phases and a decreased of the S phase using flow cytometry, with an increase of apoptosis. shUBA2 also inhibits poly (ADP-ribose) polymerase 1 (PARP1), mini-chromosome maintenance (MCM) 2, 3, and 7,

which PARP1 is associated with chemotherapy resistance in NSCLC, and MCM2, 3, 7 is related to the primary tumor (T), regional lymph node involvement (N), and distant metastases (M) (TNM) stage in NSCLC. Knockdown of Uba2 in NSCLC inhibits tumor cell proliferation and shows a potential therapeutic approach.

2.4.4 Acute myeloid leukemia

Acute myeloid leukemia (AML) is the most common leukemia for adults, where the average age of patients being diagnosed is 65 years old and is about 80% of all adult leukemia cases [136]. It is about 20,000 cases diagnosed every year in the United States. It shows as clonal of immature blast cells and causes bone marrow failure. Myelodysplastic syndrome is the most common risk factor of AML, and congenital disorders such as Down syndrome also increase the risk [137]. The 3-year survival rate of AML depends on the age, where it is 66% in patients younger than 60 years old and 33% in patients older than 60 years old. The average 5-year survival rate is 50% [138]. The standard treatments are induction therapy, which is highly toxic to bone marrow and most likely for young patients, consolidation therapy with cell transplantation, and tyrosine kinase inhibitors [136]. summary in Vakiti et al., 2023).

RNA sequencing data has identified the UBA2-WTIP fusion gene in 33.9% of AML patients by RT-PCR and Sanger sequencing [139]. As Uba2 is the catalytic subunit of SUMO E1 enzyme and functions in SUMOylation and has been found overexpressed in various cancers, WTIP functions as a tumor suppressor. Overexpression of the UBA2-WTIP fusion leads to phosphorylation of STAT3 and 5, which promote cell proliferation in AML. Triptolide in AML cells degrades the protein level of the UBA2-WTIP fusion by

western blot, where Triptolide shows induction of leukemia cell apoptosis [140]. Therefore, the UBA2-WTIP fusion has been suggested as a potential therapeutic target for AML.

2.5 Current inhibitors/compounds targeting the SUMO and SUMO E1

A couple of inhibitors are known to target the SUMOylation pathway, the most potent being TAK-981, which forms a covalent adduct with SUMO and decreases the level of protein SUMOylation [141]. TAK-981 is in phase 1 clinical trials since 2018 to check if there are any side effects in people with advanced solid tumors.

For SUMO E1 enzymes, there are a couple of non-clinical inhibitors. A few natural product small molecules, Ginkgolic acid, Anacardic acid, and Kerriamycin B, target SUMO E1 and inhibit SUMOylation [142, 143]. Ginkgolic acid and kerriamycin B block the formation of SUMO and SUMO E1, while anacardic acid is an analog of Ginkgolic acid. Davidiin [144] and Tannic acid [145] are natural products that inhibit SUMOylation. Among the natural products, Dacidiin has the lowest IC_{50} (0.15 μM), but it downregulates EZH2. IC_{50} of others are higher than 2μM. COH-000 is a covalent allosteric inhibitor of Aos1/Uba2 with IC_{50} 0.2μM [146, 147]. COH-000 covalently binds to the Cys30 residue of Uba2 and then recruits ubiquitin-proteasome degradation of Uba2 [147]. COH-000 reduces colorectal carcinoma by promoting apoptosis. Due to the 90% of clinical drug failures, rigorous research on the Uba2 drug/inhibitor is still needed.

In this chapter, we utilized SBDD coupled with FBLD and AI-CADD to identify inhibitors of Aos1/Uba2. Our overarching hypothesis was that said inhibitors can inhibit the SUMOylation and impact Myc-driven cancer cells.

2.6 Methods

2.6.1 Expression

Human Uba2 (1–640) was cloned into vector pD441-CH with a C-terminal 6x His-tag (Uba2-6His) or vector pET21a (+) with a TEV-cleavable C-terminal 6x His-tag (Uba2-TEV-6His). Human Aos1 (1–346) was cloned into vector pD454-GST with a PreScission-cleavable N-terminal GST-tag (GST-Aos1). Human SUMO1 (1–97) and Human SUMO2/3 (1–93) were cloned individually into vector pGEX-6P-1 with a PreScission-cleavable N-terminal GST-tag (GST-SUMO1 and GST-SUMO2/3, respectively). Human RanGap1 (1–587) was cloned into vector pET21a (+) with a TEV-cleavable C-terminal 6x His tag (RanGap1-TEV-6His). All plasmids were purchased from GenScript Biotech. Uba2-6His was co-expressed with GST-Aos1, and Uba2-TEV-6His, GST-Aos1, GST-SUMO1, GST-SUMO2/3, and RanGap1-TEV-6His were individually expressed by transforming plasmids in different *E. coli* cell lines and then small-scale expression test was done for each to determine the optimal cell line and expression condition and saved glycerol stocks. The glycerol stocks were cultured in 50 mL of the desired medium with the desired antibiotic at 37℃ overnight (O/N) for large-scale expression. Each O/N culture was inoculated into 12L medium with antibiotic at 37℃ at 225 rpm (New Brunswick Scientific Innova©43R). When the OD600 reached 50% of the desired concentration, the temperature was lowered to the desired inducement temperature if the inducement temperature was lower than 37℃. Then, the cells were induced with 0.4 mmol/L *isopropyl-β-D-1-thiogalactoside (*IPTG) at 100% of desired OD600 and kept shaking at the desired temperature. Cells were harvested by centrifugation (Sorvall LYNX 4000 Superspeed

Centrifuge with a FiberliteTM F9-6 x 1000 LEX Fixed Angle Rotor) at 4 ℃ and 4,000 rpm for 15 minutes, resuspended at a ratio of 3ml lysis buffer (6His-tagged proteins: 50 mM Tris-HCl pH 8, 300 mM NaCl, 25 mM imidazole, 10 mM BME; GST-tagged proteins: 0 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 mM BME) per gram pellet. Per 40 mL of the resuspended cell, it was sonicated (Q-Sonica Q125, ¼ inch probe) on the ice at 70% amplitude for 4 minutes, with 10 seconds on and 10 seconds off. The supernatant was obtained by centrifugation at 4℃ and 30,000 rpm for 40 minutes. The cell lysate was purified right after or stored at -80℃.

2.6.2 Purification

Supernatant from cell lysate with 6His-tagged in 50 mM Tris-HCl pH 8, 300 mM NaCl, 25 mM imidazole, 10 mM BME was purified by Ni-affinity chromatography (HisTrap FF; Cytiva) using NGC Quest 10 Plus Liquid Chromatography System (Bio-Rad) for fast protein liquid chromatography (FPLC) and eluted by a gradient of 1M imidazole buffer (after elution, TEV was added to cleavage 6His tag if needed), exchanged to a low salt buffer and then loaded onto anion exchange chromatography (HiTrap QFF; Cytiva) and eluted by a gradient of 1M NaCl buffer. Supernatant from cell lysate with GST-tagged in 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 mM BME was purified by GST-affinity chromatography (GSTPrep FF16/10; Cytiva) and eluted by a gradient of 25 mM glutathione buffer (PreScission was added to half of desired fractions to cleavage GST tag if needed), exchanged to a low salt buffer and then loaded onto HiTrap QFF and eluted by a gradient of 1M NaCl buffer. Eluted fractions from Uba2 and Aos1 were collected and combined, purified by gel filtration chromatography (Superdex 200; Cytiva), and finally,

ENrich high-resolution anion exchange chromatography (ENrich Q; Bio-Rad) and eluted by a gradient of 1M NaCl buffer. Fractions from other samples (GST-SUMO1, GST-SUMO2/3, and RanGap1-TEV-6His) were collected and purified by gel filtration chromatography (Superdex 75; Cytiva). 4-12% gradient SDS-PAGE gels (NuPAGE; Invitrogen) confirmed fractions from each step, and then desired fractions were collected for the next step. The final sample was concentrated (Vivaspin 10 kDa; Cytiva) and stored in 50 mM HEPES pH 7.5, 300 mM NaCl at -80℃ after flash frozen in liquid nitrogen.

2.6.3 TSA

The 20 μL of TSA contained:

- 15 μL of optimized reaction buffer,
- 2 μL of 200x SYPRO Orange protein gel stain (S6651; Thermo Scientific),
- 2 μL of protein of interest with optimized concentration, and
- 1 μL of testing compound (fragment final 10 mM or small molecule final 500 μM).

A set of 200 mM covalent fragments was purchased from Life Chemicals Inc. Chemicals such as thioketone, aromatic thiols, or Michael acceptors, acrylamides, and sulfonate esters were used for arraying potential covalent inhibitors. To screen potential covalent binders against Uba2, 1 μL of 200 mM covalent fragments in DMSO or DMSO as negative control were added individually into a white semi-skirted 96-well plate (AB-0900/W; Thermo Scientific) and centrifuged at 1500 rpm for 1 minute in a PlateFuge MicroPlate MicroCentrifuge (Benchmark Scientific). 19 μL of master mix containing 1.36 μM of Uba2/Aos1 and 200 x SYPRO Orange were added to each well, pipetting to mix thoroughly. The plate was sealed with optically clear Microseal® 'B' Seals (MSB1001;

Bio-Rad) and loaded into a CFX Connect™ Real-Time PCR Detection System (Bio-Rad). The reaction plates were heated from 25 °C to 95 °C, at increments of 0.5 °C every 30 s, taking fluorescence scans at excitation/emission ranges of 470/570 nm utilizing the HTRF filter. DMSO negative control determined the apo-T_m as the melting point of the protein. The difference between the covalent fragment and apo-Tm determined the ΔT_m as the shift of protein melting point. Positive ΔT_m indicated the binder of Uba2/Aos1 and would be confirmed.

10 mg of each hit were purchased from Mcule Inc or Enamine Inc. A series dilution of each hit was prepared with DMSO from 200 mM to 50 mM. Following the TSA method, the hit dosage/concentration tests with triplicates were done. Bigger ΔT_m with concentrated binder and smaller ΔT_m with diluted binder indicated the hit of Uba2/Aos1.

2.6.4 SUMOylation reaction

The SUMOylation reaction contained commercial or the purified Uba2-TEV-6His/Aos1 (12.3 μM) and Ubc9 (10 μM), mixed with SUMO1 (50 μM), RanGap (50 μM) and ATP (250 μM) in the reaction buffer (150 mM NaCl, 50 mM Tris pH 7.4, 10 mM MgCl2, 0.005% tween 20, 0.3 mM DTT), 2-hour 37℃ incubation. When testing hits from TSA, the compound was incubated with Uba2-TEV-6His/Aos1 at RT for 30 minutes, and then other components were added for reaction. 5X SDS-PAGE loading buffer (30% glycerol, 250 mM Tris-HCl pH 6.8) was added to each SUMO reaction and then 100°C heated for 5 minutes. 25 μL of each sample was loaded to a 12% SDS-PAGE gel running for 60 minutes at 220 V at room temperature, and then the gel was transferred onto PVDF membrane for 3 hours at 130 V at 4 °C. After the transfer, 10 mL of blocking buffer (5% BSA in TBST

(TBS + 0.1% Tween) was added to the membrane for 1 hour at room temperature. After removing the block buffer, fresh TBST with SUMO1 primary antibody (1:1000) was added to the membrane overnight at 4°C. The following day, the primary antibody was removed, and the membrane was rinsed three times for 15 minutes with TBST and then the secondary antibody (1:2000 of anti-rabbit IgG) for 2 hours at room temperature. Then, the membrane was rinsed three times for 15 minutes with TBST. The membrane was incubated with ECL reagent for 1 minute and then was imaged by ChemiDoc XRS+ (Bio-Rad) and Image Lab Software (Bio-Rad).

After images were taken, images were analyzed by ImageJ, selecting the bands of interest with the rectangular tool, and then plotted for quantitative analysis. The mean value of the expression level of SUMO1 was calculated from three replicates in each reaction. The significance was determined by comparing the quantitation of bands using two-way ANOVA. The analysis collapsed to a simple Student's T-test with Microsoft Excel. In all cases, $P < 0.05$ was used as the threshold for significance.

2.6.5 AMP assay

The AMP-GloTM Kinase Assay (Promega) measures AMP generated from ATP from any reaction. This assay confirmed that hits could inhibit Uba2/Aos1 activation to SUMO1 using ATP and then determine IC_{50}. Following the AMP-GloTM kinase assay technical manual, a standard curve was generated first to estimate the amount of AMP converted from ATP. A serial dilution from 1 mM of ATP and AMP was added into each well of a 96-well plate with 90 μL of reaction buffer (150 mM NaCl, 50 mM Tris pH 7.4, 10 mM MgCl2, 0.005% tween 20, 0.3 mM DTT) and 25 μL AMP-Glo reagent that can stop the

kinase reaction and remove ATP. After 1 hour of room temperature incubation, a 50 μL detection reagent, which converted AMP to ADP, and then convert ADT to ATP, and introduced luciferase and luciferin to detect ATP, was added to each reaction and incubated another hour. SpectraMac iD5 (Molecular Devices) read and recorded the luminescence and generated a standard curve for relative light units (RLU) and APT-to-AMP conversion.

A SUMO1 concentration test was done to determine the optimal concentration. A serial dilution of SUMO1 (down to 0 as background signal) was mixed with an undiluted Uba2/Aos1 sample and desired ATP in the reaction buffer, one to two hours 37℃ incubation. Then, AMP-Glo reagent was added for each reaction for 1-hour room temperature incubation. Then, a detection reagent was added to each reaction. The plate was measured after one hour of room temperature incubation.

Uba2/Aos1 concentration test was done to determine the optimal concentration. A serial dilution of Uba2/Aos1 (down to 0 as background signal) was mixed with the desired SUMO1 sample and desired ATP in the reaction buffer, 1 to 2 hours 37℃ incubation. Then, AMP-Glo reagent was added for each reaction for 1-hour room temperature incubation. Then, a detection reagent was added to each reaction. The plate was measured after 1-hour room temperature incubation.

Hits were tested using the covalent hits (10 mM final) and AI-CADD hits (500 μM final) determined by TSA. Each test contained hits in DMSO, 6.15 μM of Uba2/Aos1, 25 μM of ATP, 12.5 μM SUMO1, the reaction buffer, and DMSO negative control. A known

Uba2/Aos1 inhibitor, COH000 (MedChemExpress), was used as a positive control. The plate was read and determined hits that inhibited Uba2/Aos1 activation to SUMO1.

IC_{50} was done triplicate with the hits. A serial dilution of inhibitor was tested to generate a curve of Log 10 of inhibitor concentration and RLU (GraphPad). Average and standard deviation were used to plot to visualize the impact of each compound on Dyrk1a activity. CV was less than 10% to express precision and repeatability.

2.6.6 HTRF assay

Following the AID 2006 -uHTS HTRF (homogeneous time-resolved fluorescence) assay protocol [148], Dr. David Kwon and Dr. Holly Yin from HTS Core Facility, City of Hope, tested the DTP compounds. Using the previously purified proteins GST-SUMO2, Uba2/Aos1, Ubc9, RanGap1-TEV-6His, and ATP, SUMO2 was activated by Uba2/Aos1 and then attach to RanGap1 by Ubc9. Anti-GST-Terbium and Anti-6His-d2 measured could recognize GST-SUMO2 and RanGap1-TEV-6His, respectively, and send a fluorescent signal only when SUMO2 SUMOylated RanGap1. Therefore, the signal indicated the reaction activity and enzymes (Uba2/Aos1 and Ubc9) activity.

Reaction buffer (150 mM NaCl, 50 mM Tris pH 7.4, 10 mM MgCl2, 0.005% tween 20, 0.3 mM DTT), Uba2/Aos1, RanGap1-TEV-6His, and testing compound or DMSO were added first, and ATP, Ubc9, and GST-SUMO2 were added for RT incubation 30 minutes. Antibodies were added, and then fluorescence signal was measured.

The concentration of each component was optimized for the desired signal/background ratio, including DMSO toleration. First, the screening test was done by duplicate, and then the IC_{50} of the hit was measured.

2.6.7 Testing compound

- Drug-like fragments from Zenobia. Stock 200 mM in DMSO.
- Covalent or Brominated fragments from Life Chemicals (LC). Stock 200 mM in DMSO.
- DTP docking small molecules targeting the ATP binding pocket from the Mobley Lab, UCI (DTP). Stock 200 mM in DMSO.
- AI-CADD small molecules targeting on Uba2 activity site Cys 173 from Atomwise derived from (AW1) previously hit identified by a former member in the Perry Lab. Stock 10 mM in DMSO.
- AI-CADD small molecules targeting on Uba2 activity site Cys 173 from Atomwise derived from AW33 identified by a former member in the Perry Lab (AW2). Stock 10 mM in DMSO.

A series of assays tested the compounds:

Screening for binders by TSA. Using DMSO as the negative control, positive shift (> / = 0.5 °C) indicated binding and stabilizing protein, and negative shift indicated binding but destabilizing protein. The positive binders were to be focused, as they would benefit future crystallography studies. There was some correlation between ΔT_m and binding affinity/inhibition efficacy. However, a larger ΔT_m did not always indicate a better inhibitor, specifically for fragments, when their binding affinity was not high.

A dosage/concentration of compounds test was conducted to confirm the binder in triplicate by TSA. By diluting compound concentration as a hit, the ΔT_m decreased with

compound dilution, ideally. If ΔT_m did not change or increase as the compound dilution, the binding was most likely non-specific.

Screening test in duplicate for hits by AMP. Using DMSO as the negative control, the luminescent signal of compound/signal of DMOS *100 was % activity. Decreased % activity indicated a decrease of reaction and inhibition from the testing compound.

IC_{50} test of hits in triplicate by AMP. A 2-fold series dilution of compounds was used, and the luminescent signal was plotted against Log10 of concentration by GraphPad, and relative IC_{50} was calculated.

2.6.8 Crystallography study

Crystal trays were set up at RT using 96-well or 24-well crystal trays by sitting drop. 300-400 μL of reservoir was added in the reservoir chamber first. 0.5 or 1 μL of protein (or mixed with small molecule hits) was added to the well, and then mixed with the same amount of reservoir solution. The chamber was sealer by tape and incubate at RT. Crystal trays were examined by a microscope after setup, and every day for 2 weeks.

2.7 Result

2.7.1 Expression

In small-scale expression test, Uba2-6His or Uba2-TEV-6His plasmids were co-transformed with GST-Aos1 or transformed solo in *E. coli* BL21(DE3), ArcticExpress RIL, ArcticExpress RP, ArcticExpress (DE3), or BL21-CondonPlus-RP, and cultured in Luria broth (LB), Terrific Broth (TB), or 2XYT. Cells were induced at OD600 1.2 and set to shake at either 37℃ 4 hours or 14℃ O/N. All cells were collected and then loaded onto nickel spin columns (HisPur Ni-NTA spin column; Thermo Fisher), eluted with high

imidazole buffer, loaded onto GST spin column (Pirece Glutathione spin column; Thermo Fisher), and eluted with high glutathione buffer and run on SDS-PAGE gels for comparison (co-expression samples only [**Supplementary material figure 2.1 A**]). Uba2 is a 74 kDa protein, but the elution samples always showed an around 100 kDa band (**Supplementary material figure 2.2**). To confirm the 100 kDa band, TEV treated the nickel spin column elution, and another nickel spin column test was conducted. As shown in the figure, the 100 kDa band was shown in the flow through but not in elution, indicating the sample did not contain 6His-tagged after TEV treatment (**Supplementary material figure 2.2 B**), ruled out the possibility of contamination or non-specific protein in the sample binding to the nickel spin column in the first round. For Uba2-6His, the best expression was when co-transformed with GST-Aos1 and used the ArcticExpress RP strain, and the condition was induced at 14℃ at OD600 1.2 using 2XYT. Uba2-TEV-6His expression without GST-Aos1 in ArcticExpress RP and induced at 14℃ at OD600 1.2 using 2XYT showed the best result (**Supplementary material figure 2.1 B**). Glycerol stocks were saved. Due to Uba2-TEV-6His having a better expression level, being able to cleave off the 6His tag, and being able to be expressed solo, it was used for most of the experiments after large-scale expression and purification.

GST-SUMO1, GST-SUMO2/3, and RanGap1-TEV-6His were transformed individually in *E. coli* BL21(DE3) and followed by the GST or nickel spin column tests. Inducement at OD600 0.8 and set to shake at 16℃ O/N was the best condition for all three proteins (**Supplementary material figure 2.3**).

2.7.2 Purification

In the large-scale expression for purification, Uba2-TEV-6His would be treated by TEV aiming to gain SUMO E1 without a tag for the HTRF assay. PreScission always treated GST-Aos1 after GSTPrep to cleavage over the GST tag. To remove the GST tag, another GSTPrep would be done to capture Aos1 without GST from the flow-through. As shown in the **Supplementary material figure 2.4 B**, the flow through fractions 8 - 34 all had a 37 kDa band represented Aos1 no tag and a 25 kDa band, matching the PreScission size. QFF could not separate Aos1 and PreScission, and Superdex 75 could not either (**Supplementary material figure 2.4 C**), as their pI and size were similar. Fortunately, Aos1 combined with Uba2 with or without 6His tag, followed by Superdex 100 and ENrich Q, could separate Uba2/Aos1 from PreScission (**Supplementary material figure 2.5 and 2.6; figure 2.1**). The yield was 1.8 mg of Uba2/Aos1 per L of *E. coli* culture and 1.04 mg of Uba2 per L.

Purifications of GST-SUMO1, GST-SUMO2/3, and RanGap1-TEV-6His followed the steps of affinity (6His or GST), HiTrap Q FF, and Superdex 75. Both SUMO1 and SUMO2/3 had the diglycine at their C-terminal, which was the matured form of SUMO and ready to be activated (**Supplementary material figure 2.7; figure 2.2**). The yield was 2.25 mg of GST-SUMO1 or GST-SUMO2/3 per L of *E. coli* culture and 2.5 mg of RanGap1-TEV-6His per L of *E. coli* culture.

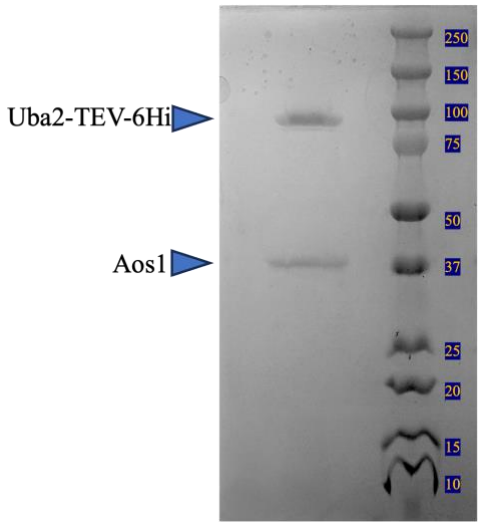


Figure 2.1 Purified Uba2/Aos1 sample. The final post-ENrich S purification sample of Uba2/Aos1 was loaded on an 8-16% SDS-PAGE gel (Novex Tris-Glycine Mini Protein Gel; Thermo Fisher). Lane 1: final sample from ENrich S elution fraction. Lane 2: protein ladder (Precision Plus Protein Dual Color Standards; Bio-Rad).

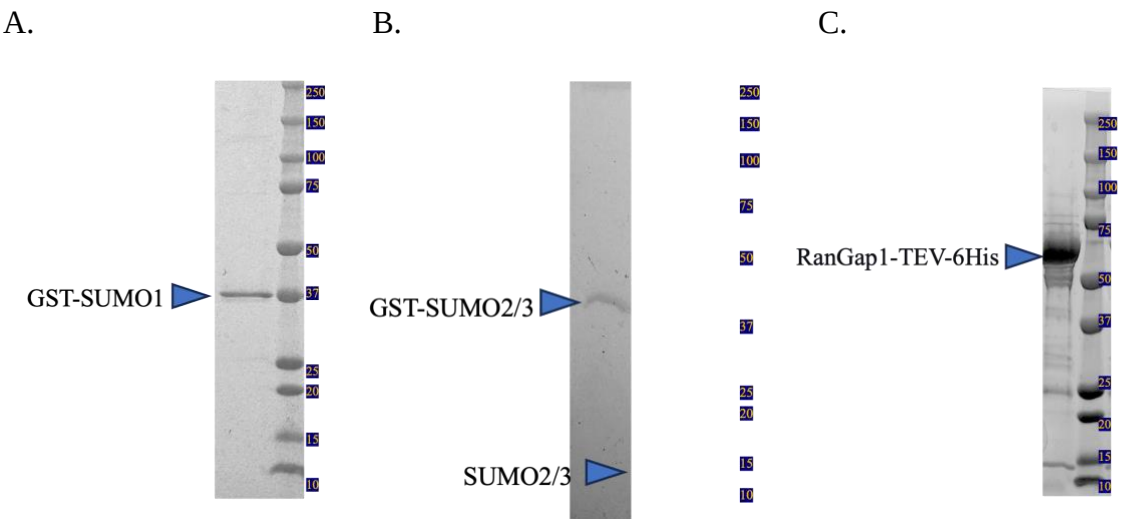
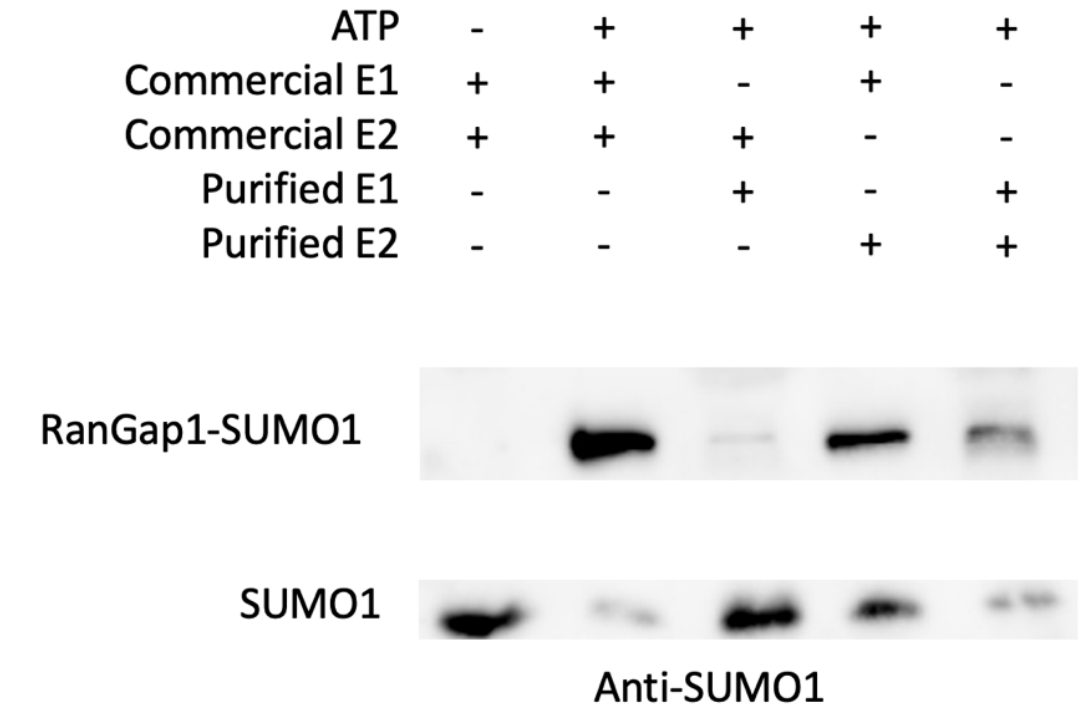


Figure 2.2 Purified GST-SUMO1, SUMO2/3, and RanGap1 sample. A. The final post-Superdex 75 purification sample of GST-SUMO1 was loaded on an 8-16% SDS-PAGE gel. Lane 1: final sample from Superdex 75 elution fraction. Lane 2: protein ladder. B. The final post-Superdex 75 purification sample of GST-SUMO2/3 and SUMO2/3 were loaded on an 8-16% SDS-PAGE gel. Lane 1: final GST-SUMO2/3 sample from Superdex 75 elution fraction. Lane 2: final SUMO2/3 sample from Superdex 75 elution fraction. Lane 2: protein ladder. A. The final post-Superdex 75 purification sample of RanGap1-TEV-6His was loaded on an 8-16% SDS-PAGE gel. Lane 1: final sample from Superdex 75 elution fraction. Lane 2: protein ladder.

2.7.3 SUMOylation activity assay

To confirm the activity of purified proteins, SUMO E1 Uba2-TEV-6His/Aos1, SUMO E2 Ubc9, SUMO E3 Pias1 (in chapter 3), and GST-SUMO1, SUMOylation activity assay was done after purifications. The band of SUMO1 only would represent no SUMOylation reaction of RanGap1, and the band of SUMOylated RanGap1 represented the SUMOylation reaction of RanGap1 and indicated the cascade reaction occurred. As shown in **Figure 2.3 A**, from the left side, lane 1 and lane 2 had all commercial proteins to confirm the experiment setup, as only SUMO1 in lane 1 and SUMOylated RanGap1 in lane 2. When replacing the commercial SUMO E1 with the purified Uba2-TEV-6His/Aos1 or commercial E2 with the purified E2 Ubc9, RanGap1 was still SUMOylated, indicating the function of purified E1 and E2. When the purified GST-SUMO1 was tested, commercial SUMO E1 and E2 were added with or without ATP (**Figure 2.3 B**). Commercial SUMO1 was used as positive control. The band of lane 1 was higher than lane 3, because GST-SUMO1 was 25 kDa bigger than SUMO1.

A.


ATP - + + + +
Commercial E1 + + - + -
Commercial E2 + + + - -
Purified E1 - - + - +
Purified E2 - - - + +
RanGap1-SUMO1
SUMO1
Anti-SUMO1

B.

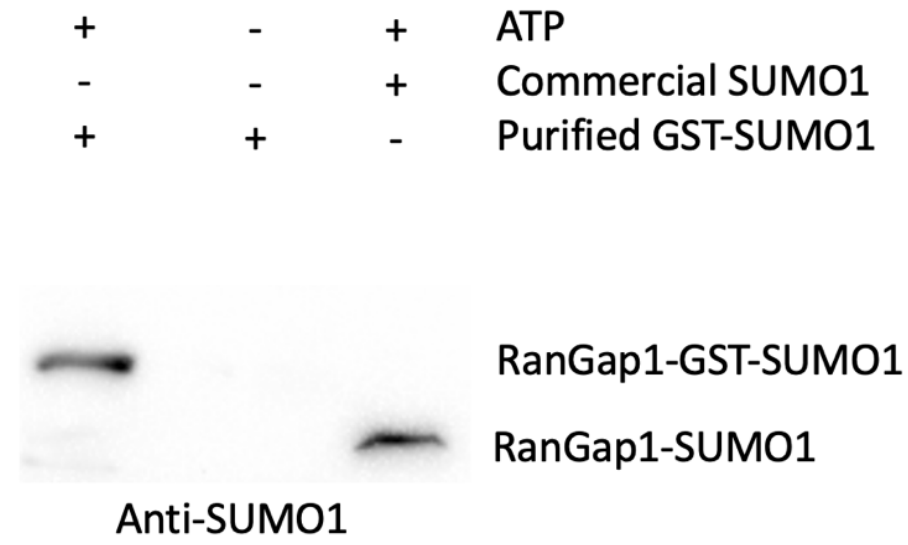

+ - + ATP
- - + Commercial SUMO1
+ + - Purified GST-SUMO1
RanGap1-GST-SUMO1
RanGap1-SUMO1
Anti-SUMO1

Figure 2.3 purified SUMO E1, E2, GST-SUMO1 were tested by SUMOylation reaction kit and western blot. Proteins were added with or without ATP for SUMOylation modification reaction on RanGap1. Reactions were incubated at 37℃ for 2-hour and loaded on an 12% SDS-PAGE gel. Gel was transferred to a PVDF membrane and incubated with anti-SUMO1 antibody at 4℃ overnight. Then membrane was incubated with anti-rabbit secondary antibody at RT for 2-hour and then performed image. A. purified SUMO E1 and E2 were tested with commercial E1 and E2. Lane 1: commercial E1 and E2 without ATP. Lane 2: commercial E1 and E2 with ATP resulting in SUMOylated RanGap1. Lane 3: purified E1 and commercial E2 with ATP resulting in SUMOylated RanGap1. Lane 4: purified E2 and commercial E1 with ATP resulting in SUMOylated RanGap1. Lane 5: purified E1 and E2 with ATP resulting in SUMOylated RanGap1. B. lane 1: purified GST-SUMO1 and commercial E1 and E2 with ATP resulting in SUMOylated RanGap1. Lane 2: purified GST-SUMO1 and commercial E1 and E2 without ATP. Lane 3: Commercial SUMO1, E1, and E2 with ATP resulting in SUMOylated RanGap1.

2.7.4 TSA

A former student in the Perry lab did the optimized protein concentration and reaction buffer for TSA. A final of 9 μM of Uba2/Aos1 in 50 mM HEPES pH 8, 50 mM NaCl, generated an apo-T_m of 36.5℃ $\pm$ 0.5℃.

2.7.5 AMP assay

The AMP standard curve test determined 20 μL of ATP as the desired concentration (**Supplementary material figure 2.8 A**). Higher than 20 μL might cause the reading to be statured. Using undiluted Uba2/Aos1, the SUMO1 concentration test was done by a series dilution (**Supplementary material figure 2.8 B**). 62.5 μL of SUMO1 generated the highest signal difference from the background signal. 5.77 μL of Uba2/Aos1 was the desired concentration (**Supplementary material figure 2.8 C**), generating a high signal as the undiluted sample (11.55 μL) and showing 10.74 of signal/background. One hour of reaction incubation time was determined as two hours might show a statured signal. For the best effect of compound inhibition, compounds were added to Uba2/Aos1 with reaction buffer for a 30-minute RT incubation, and then SUMO1 and ATP were added for the reaction. A Z-factor test was done using COH000 (**Supplementary material figure 2.8 D**). 25 μL of COH000 showed 98.8% inhibition, and the Z-factor was 0.9205, indicating the assay was excellent for compound screening. In the screening test, compounds were tested of inhibition compared with DMSO by duplicate, and then IC_{50} tests were done to confirm. IC_{50} measurements were used to rank compounds.

2.7.6 Drug-like fragments from Zenobia

During the screening test, the first group of compounds to be tested against SUMO E1 was 967 drug-like fragments from Zenobia. The fragments were derived from drug-like cores and followed the Rule of Five. Therefore, the fragments represented safety and good drug properties. 31.66% of the drug-like fragments had positively shifted at least 0.5 °C of the melting point of SUMO E1, and 2.96% had positively shifted at least 2 °C, indicating SUMO E1 is highly druggable (**Figure 2.4**).

2.7.7 Covalent or Brominated fragments from Life Chemicals (LC)

The second group was 1129 covalent fragments and 618 brominated fragments from Life Chemicals. The enzymatic activity site of the catalytic subunit of SUMO E1 is residue Cys 173 of Uba2. Therefore, fragments such as thioketone, aromatic thiols, or Michael acceptors et al. can bind to the Cys residue with a covalent mechanism and irreversibly inhibit it. Brominated fragments can be located during X-ray crystallography via anomalous scattering to overcome the challenge of low electron density of fragments. The percentage of 2 °C ΔTm was 3.37 for covalent and 3.56 for brominated (**Figure 2.4**). A concentration/dosage of fragment test was done with those fragments that showed the most T_m shift, and four fragments had dosage trends (**Supplementary material table 2.1)**: 1C, 1E, 1H, and 2A. The SUMOylation activity assay and the AMP assay tested those (**Figure 2.5 A and B**). Notably, 1H and 2A are both thioketone and very similar structurally.

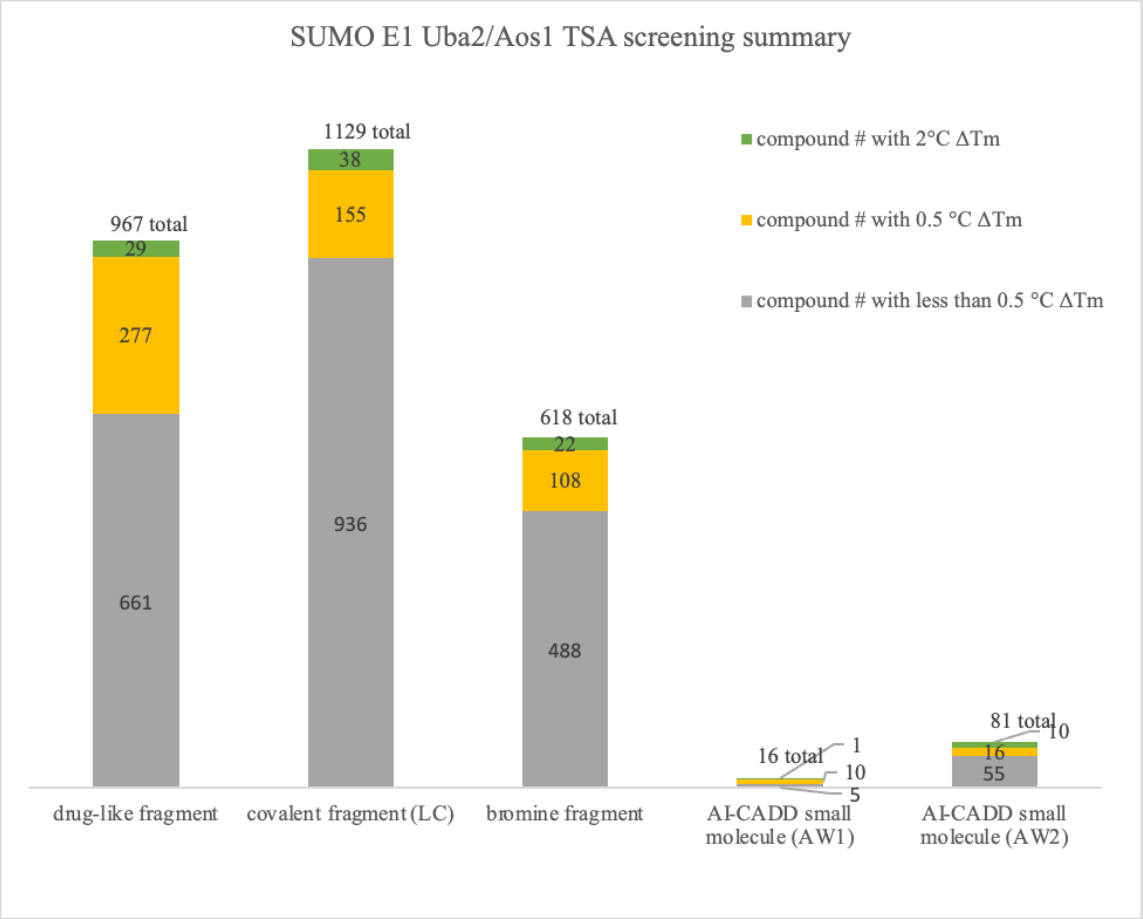
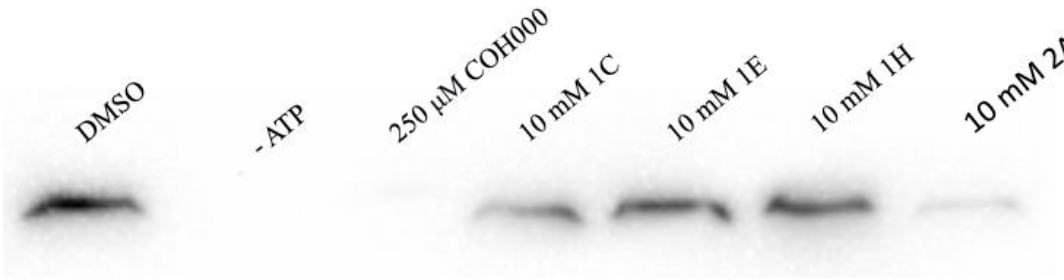


Figure 2.4 SUMO E1 Uba2/Aos1 TSA screening with drug-like, covalent, and brominated fragment, AW1 and 2 AI-CADD small molecules. 2714 fragments (1129 covalent, 618 brominated, and 967 drug-like) were tested against Uba2/Aos1 by TSA. The % of fragments with at least 0.5 °C ΔT_m were 17.09, 21.04, and 31.66, respectively. 16 small molecules from AW1 and 81 small molecules from AW2. The % of small molecules with at least 0.5 °C ΔT_m were 68.75 and 34.7, respectively.

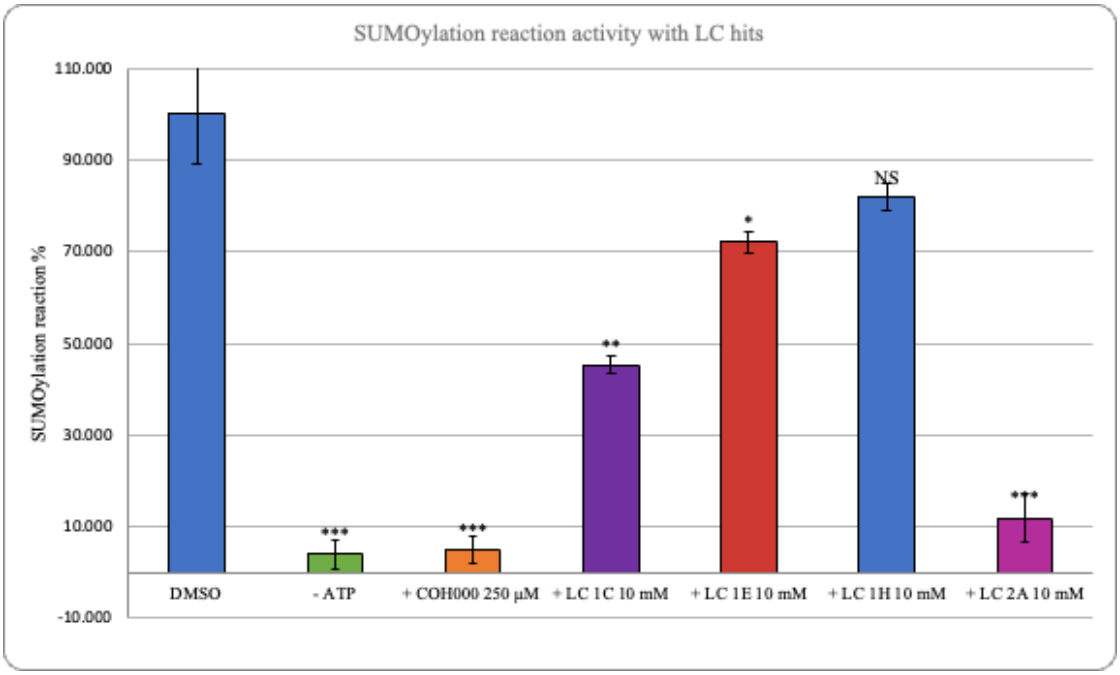
The SUMOylation activity assay was repeated to confirm LC covalent hits from TSA. The band of SUMOylated RanGAP1 represented the reaction (**Figure 2.5 A**). Lane 1 and lane 2 were positive and negative control for the SUMOylation reaction, and lane 1 represented the full activity of the reaction. Lane 3 was the positive control of a good hit of Uba2/Aos1, as COH000 has been identified as a covalent allosteric inhibitor of SUMO E1 (Lv et al., 2018). As shown in the western blot membrane, the reduction of SUMOylated RanGap1 in lane 7 (LC 2A) was significant, confirmed by the statistical result (P=0.0003481). The result determined that LC 2A could inhibit SUMOylation by binding to SUMO E1. However, 1H did not have a P value smaller than 0.05, although 1H and 2A are structurally similar.

The AMP assay tested four fragment hits from TSA and determined relative IC_{50} (hereafter IC_{50}). 2A had 30.21 μM IC_{50}, while 1H had 105.8 μM (**2.5 C**). As a fragment, 2A showed promising results in the first-round drug test.

A.


DMSO
- ATP
250 µM COH000
10 mM 1C
10 mM 1E
10 mM 1H
10 mM 2A

B.


SUMOylation reaction activity with LC hits
SUMOylation reaction %
110.000
90.000
70.000
50.000
30.000
10.000
-10.000
***
***
**
*
NS
***
DMSO
- ATP
+ COH000 250 µM
+ LC 1C 10 mM
+ LC 1E 10 mM
+ LC 1H 10 mM
+ LC 2A 10 mM

C.

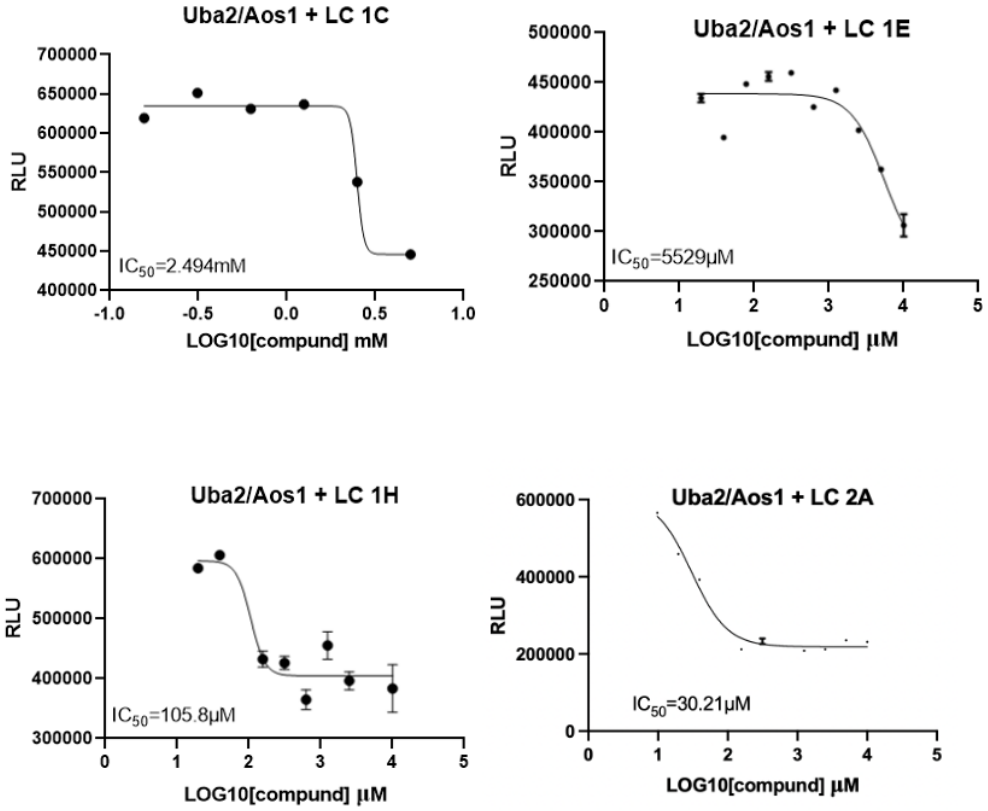


Figure 2.5 LC hits western blot and IC50. A. SUMO1, Uba2/Aos1, Ubc9 and RanGap1 were added with LC covalent fragments (10 mM) or COH000 (250 μM) or DMSO individually for SUMOylation modification reaction on RanGap1. Reactions were incubated at 37℃ for 2-hour and loaded on an 12% SDS-PAGE gel. Gel was transferred to a PVDF membrane and incubated with anti-SUMO1 antibody at 4℃ overnight. Then membrane was incubated with anti-rabbit secondary antibody at RT for 2-hour and then performed image. B. qualification data using ImageJ. P value was calculated using T-TEST. *: $P < 0.9$. **: $P < 0.09$. ***: $P < 0.009$. C. IC_{50} figures of covalent hits using AMP assay and GraphPad (Prism).

2.7.8 AI-CADD small molecules targeting on Uba2 activity site Cys 173 from Atomwise derived (AW1 & 2)

The testing groups were AI-CADD small molecules from Atomwise. Small molecules from AW1 had one out of 16 compounds (6.25%) of 2 °C ΔT_m from the screening test, but none showed a dosage trend in the dosage test. 10 out of 81 compounds from AW2 had at least 2 °C ΔT_m. Most showed a dosage trend in the dosage test against Uba2/Aos1, and six showed a trend in the test against Uba2 solo. They were tested in the AMP assay and obtained IC_{50} (see the AMP section).

For AW1 and 2, a screening AMP test of the final 500 μM of the compound was done first (**figure 2.15 A and B**). 2 out of 14 compounds in AW1 showed less activity in the AMP assay, but none had IC_{50} less than 50 μM. In the AW2 screening, six of 79 compounds showed 40% activity less than DMSO control **(Supplementary material table 2.2)**. IC_{50} was determined for them. 4 compounds were in the 10 μM IC_{50} range, where the best was 0.4977 μM (**Figure 2.6 & Supplementary material figure 2.9-11**).

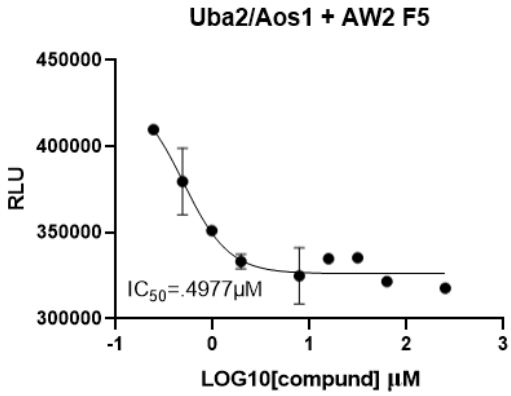


Figure 2. 6 Uba2/Aos1 with AW2 F5. IC_{50} figures of covalent hits using AMP assay and GraphPad.

2.7.9 DTP docking small molecules targeting the ATP binding pocket from the Mobley Lab, UCI (DTP)

The last group was virtual screening small molecules (Uba2 DTP) against the ATP binding pocket of the Uba2 subunit from Dr. Mobley's group at UC Irvine. However, most small molecules are very colorful and show binding to the SYPRO Orange effect. Therefore, the TSA results of Uba2 DTP were not valid.

AMP tests were done with Uba2 DTP compounds (**Figure 2.7 A & Supplementary material table 2.3**). Using DMSO as the negative control, the impact of each compound was indicated by % activity, which was % activity = (signal from compound/signal from DMSO) *100%. Less or equal to 50% was the cutoff. 14 out of 40 were tested to obtain IC_{50}. The best was DTP C7 (18.41 μM IC_{50}). Other good ones were DTP A7 (35.37 μM IC_{50}) and C3 (41.76 μM IC_{50}) (**Figure 2.7 B, C & D)**. Other compounds had IC_{50} larger than 50 μM.

HTRF assay was used to test DTP compounds and the HTRF data was generated by Dr. David Kwon, City of Hope (**Figure 2.7 A and Supplementary material figure 2.12**). DTP C7, A7, and C3 also showed inhibition in the HTRF assay and promising IC_{50} (8.01 μM, 1.01 μM, and 10 μM, respectively). DTP C11 was determined to be IC_{50} of 9.24 μM in the HTRF assay. Another five compounds had IC_{50} in the 10-30 μM range.

A.

compound	AMP % activity (50 uM)	FRET % activity (50 uM)	AMP IC50 μM	FRET IC50 uM
C7	**9.24**	10.24	18.41	8.01
A7	**2.55**	36.35	35.37	1.01
C3	**24.89**	42.83	41.76	10.38
C11	**17.43**	16.72	>50	9.244
C6	**27.38**	17.8	>50	15.99
A3	**30.27**	11.56	>50	16.48
A4	**65.76**	59.11	>50	17.08
A1	**35.25**	65.83	>50	20.72
C9	**39.05**	21.11	>50	25.41
A12	**43.79**	54.61	>50	>50

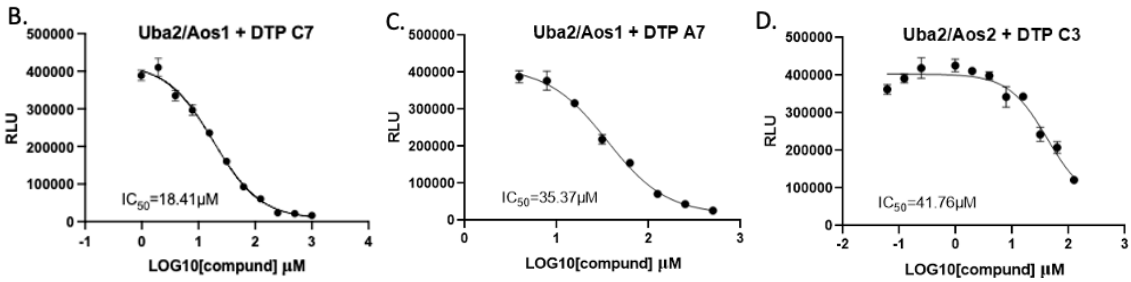


Figure 2.7 Uba2/Aos1 and DTP compound test by AMP assay and HTRF **assay.** A. hits from DTP screening were tested against to Uba2/Aos1 along with DMSO as negative control in triplicates and IC_{50} by AMP and HTRF assay. B, C, D: DTP C7, A7, and C3 IC_{50} figure by AMP assay and GraphPad.

2.8 Discussion

After optimization of the expression condition of the Uba2/Aos1 complex, the protein expression level has been significantly increased, from 0.4 to 1.8 mg of Uba2/Aos1 per L of *E. coli* culture. The expression of Uba2 solo also increased the yield to 1.04 mg of Uba2 per L of *E. coli* culture.

The hit identification studies against Uba2/Aos1 using drug-like fragments, covalent and brominated fragments, and DTP small molecules showed Uba2/Aos1 is highly druggable and provided hit candidates for the following stages. TSA assay indicated that LC 2A could bind and stabilize Uba2/Aos1 because of the positive shift of T_m, while the AMP assay and the SUMOylation activity assay confirmed that LC 2A could inhibit Uba2/Aos1. As LC 2A is a covalent fragment, data suggested that LC 2A covalently bind at the activity site Cys 173. Fragment LC 2A showed IC_{50} of 30.21 μM in the AMP assay and 9.7% of activity (10 mM) in the SUMOylation activity assay compared to the DMSO negative control.

IC_{50}, the biochemical half-maximal inhibitory concertation, has been widely used in early drug discovery and lead optimization [149]. In general, covalent/irreversible inhibitors should be ranked by a single IC_{50}, as the IC_{50} of an irreversible inhibitor decreases over time [150]. A compound with low binding affinity and high chemical reactivity may result in a low value of IC_{50} [151]. Since the experiment setup was the same during the AMP assay, the reaction time and concentration of enzyme (Uba2/Aos1), substrate (SUMO1), and ATP were the same. Therefore, IC_{50} of all covalent fragments

could be used to rank fragment hits from LC and determine which hit(s) could move to the next stage.

Tests on the DTP small molecules also identified promising hit candidates. DTP C7, A7, and C3 were determined using the AMP assay, IC_{50} being in the 50 μM range. The HTRF assay also confirmed their inhibition and promising IC_{50} (8 μM, 1 μM, and 10 μM, respectively). Another six compounds had IC_{50} in the 50 μM range using the HTRF assay. Four of six showed inhibition during the AMP assay screening test, but the IC_{50} values were all larger than 50 μM.

SUMO1 was used in the AMP assay as SUMO1 showed a better signal/background ratio than SUMO2, while SUMO2 had a better ratio than SUMO1 in the HTRF assay. The AMP assay measured the activity of Uba2/Aos1 activating SUMO1, while the HTRF assay measured the activity of SUMO2 SUMOylated RanGap1. However, Uba2/Aos1 activated SUMO1 and SUMO2 via the same chemical reaction, and the compounds targeted Uba2/Aos1, and so results from both assays are valid. The overall IC_{50} values obtained by the HTRF assay were lower than the AMP assay. For DTP C7, A7, and C3, the IC_{50} values indicated that they could move on to the next stage.

IC_{50} measurement is assay specific. The AMP assay measured the change of ATP to AMP as Uba2/Aos1 activated SUMO. The remaining ATP was removed, and AMP was converted to ADP, and then converted to ATP. The ATP then went through luciferin/luciferase reaction for light output. The luciferase signal was converted to IC_{50} values of inhibitors. HTRF assay directly measures the level of SUMOylation of RanGap1

and converted it to IC_{50}. For the ideal signal/background ratio, the setup of two assays were also different. i.e., the amount of ATP added the two assays were different.

AW1 compounds were derived from previously hit. B4, B2, and B1 positively shifted Uba2/Aos1 T_m higher than 1℃. However, they all failed the dosage test and showed no inhibition in the AMP screening test. A7 and A8 inhibited in the AMP screening assay, while they only slightly shifted Uba2/Aos1 T_m. However, their IC_{50} values were larger than 50 μM, which suggests a weak inhibitor that is not of significant interest. As the ΔT_m of A7 and A8 was 0.33 °C and 0.83 °C, respectively, when using the most concentration of the compounds, dosage tests would also be very hard to prove the binding these compounds to the protein was not a false positive.

AW2 compounds were derived from AW33, which negatively shifted Uba2/Aos1 T_m. 4 of the AW2 compounds, F5, C9, E1, and B7, had IC_{50} in the 10 μM range, the best one being 0.5 μM from F5. All four compounds positively shifted Uba2/Aos1 T_m and had dosage trend. As a testing compound in the hit-to-lead, the IC_{50} of F5 and the preliminary data of co-crystals of Uba2/Aos1 and F5 showed a promising result (**figure 2.16**). To optimize these crystal hits, focusing on identifying conditions that produce a better morphology crystal from Uba2/Aos1 and F5 (i.e., 'diffraction quality') will be the next step. Using X-ray crystallography, the interaction between protein and compound can be defined at atomic resolutions, and this information will be combined with further docking and SAR studies, to produce an F5-based analog with improved affinity. The other three compounds of potential interest could also be optimized using the same strategy.

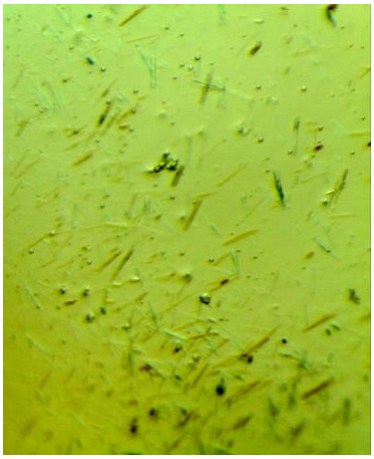
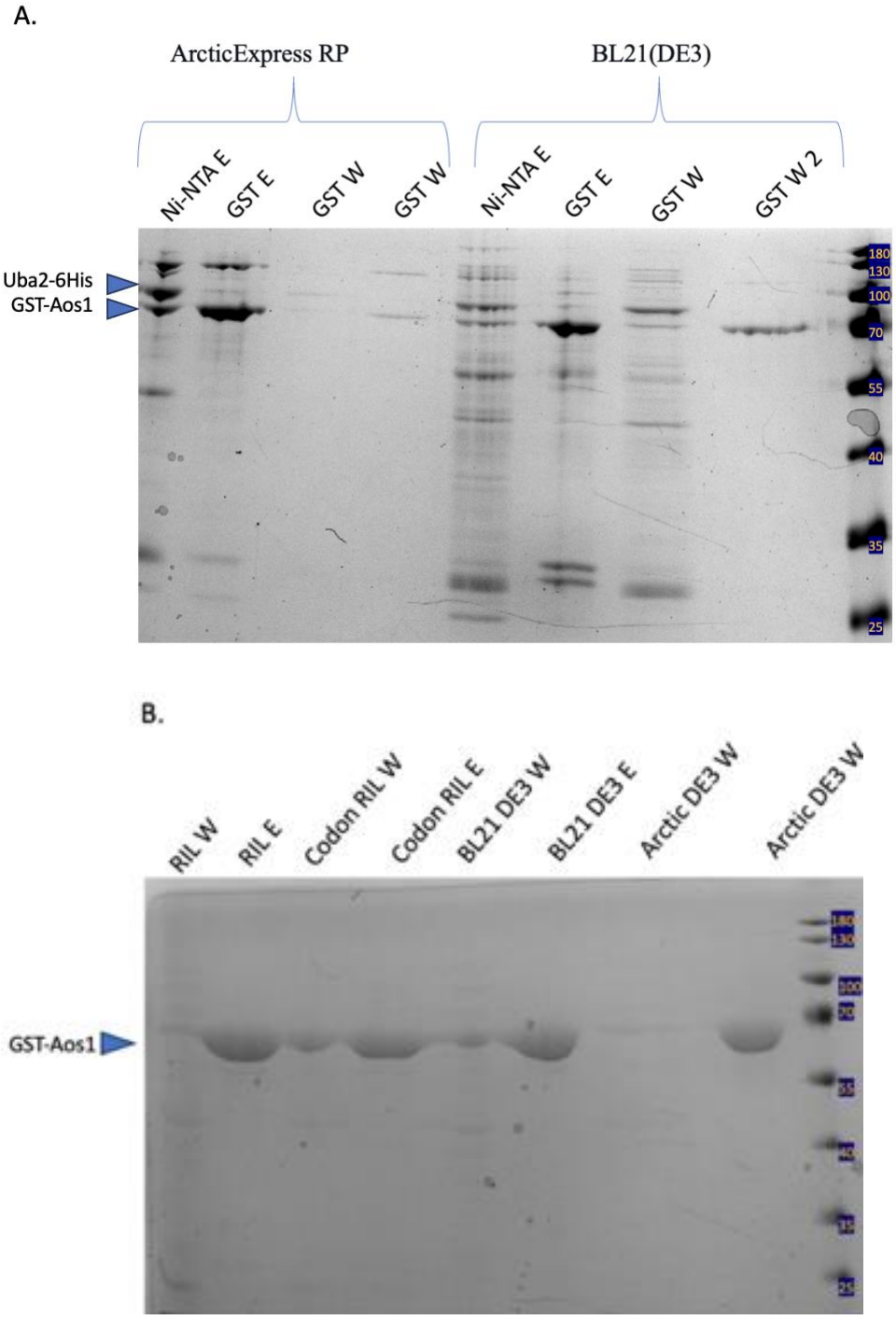


Figure 2. 16. Uba2/Aos1 with AW2 F5 co-crystals. Uba2/Aos1 with 1.2X of AW2 F5 sitting drop in RT.

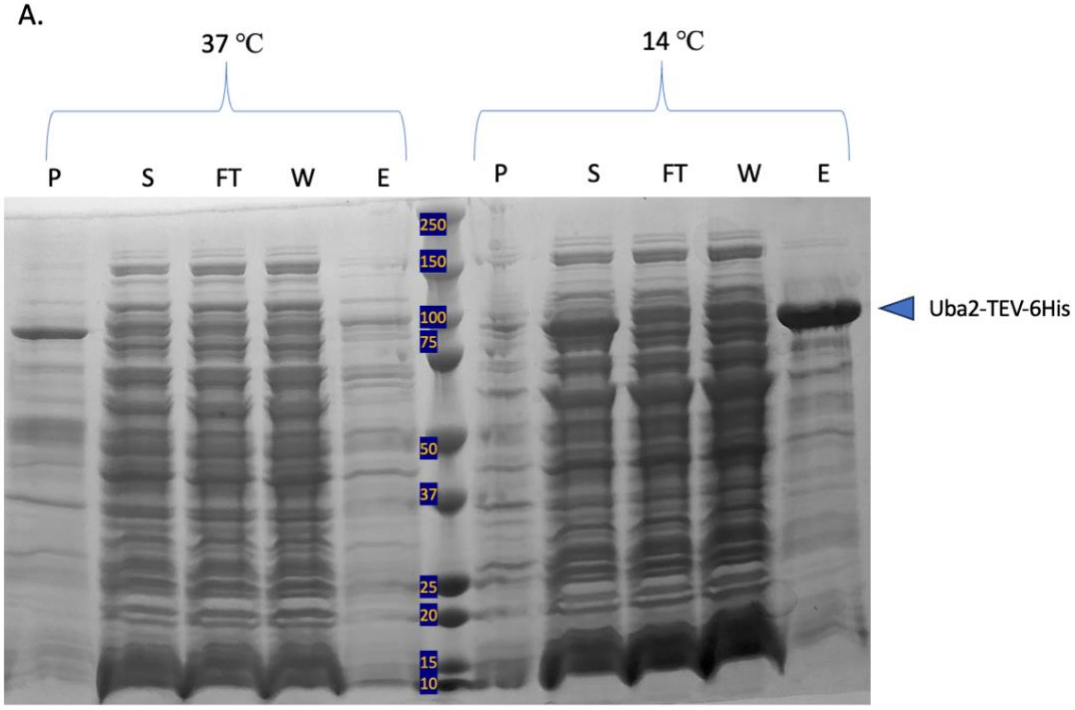
In the future steps, potencies such as K_i and K_d measurement would be necessary for the hit-to-lead and guide lead optimization, as they are the potency that are not assay specific. Low IC_{50} indicates the inhibitor is potent with low concentration, and low K_i and K_d indicate high binding affinity of the inhibitor. SBDD with SAR studies can improve IC_{50} and K_d in the hit-to-lead and lead optimization stage. The lower IC_{50}, K_i, and K_d of the compound, the more likely it passes in the pre-clinic tests, as lower doses can be used to help avoid off-site toxicities.

2.9 Supplementary material

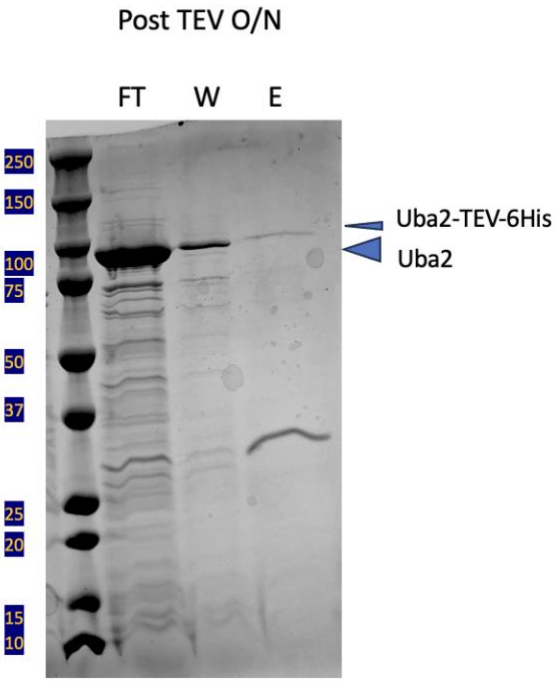


Supplementary material figure 2.1 Uba2/Aos1 co-expression & Aos1 expression small-scale test. A. Uba2-6His co-expressed with GST-Aos1 in *E. coli BL*21(DE3) and ArcticExpress RP with 2XYT and induced by IPTG @ OD 1.2 at 14℃ O/N. Cells were harvested and examined by NI-NTA spin column. Samples were loaded on an 8-16% SDS-PAGE gel. SDS-PAGE gel showed elution samples from Ni-NTA and then GST spin columns. E: elution. W: wash. B. GST-Aos1 expressed in *E. coli BL*21(DE3), ArcticExpress RIL, ArcticExpress RP, ArcticExpress (DE3), or BL21-CondonPlus-RIL with 2XYT and induced by IPTG @ OD 1.2 at 14℃ O/N. SDS-PAGE gels showed elution samples from GST spin column.

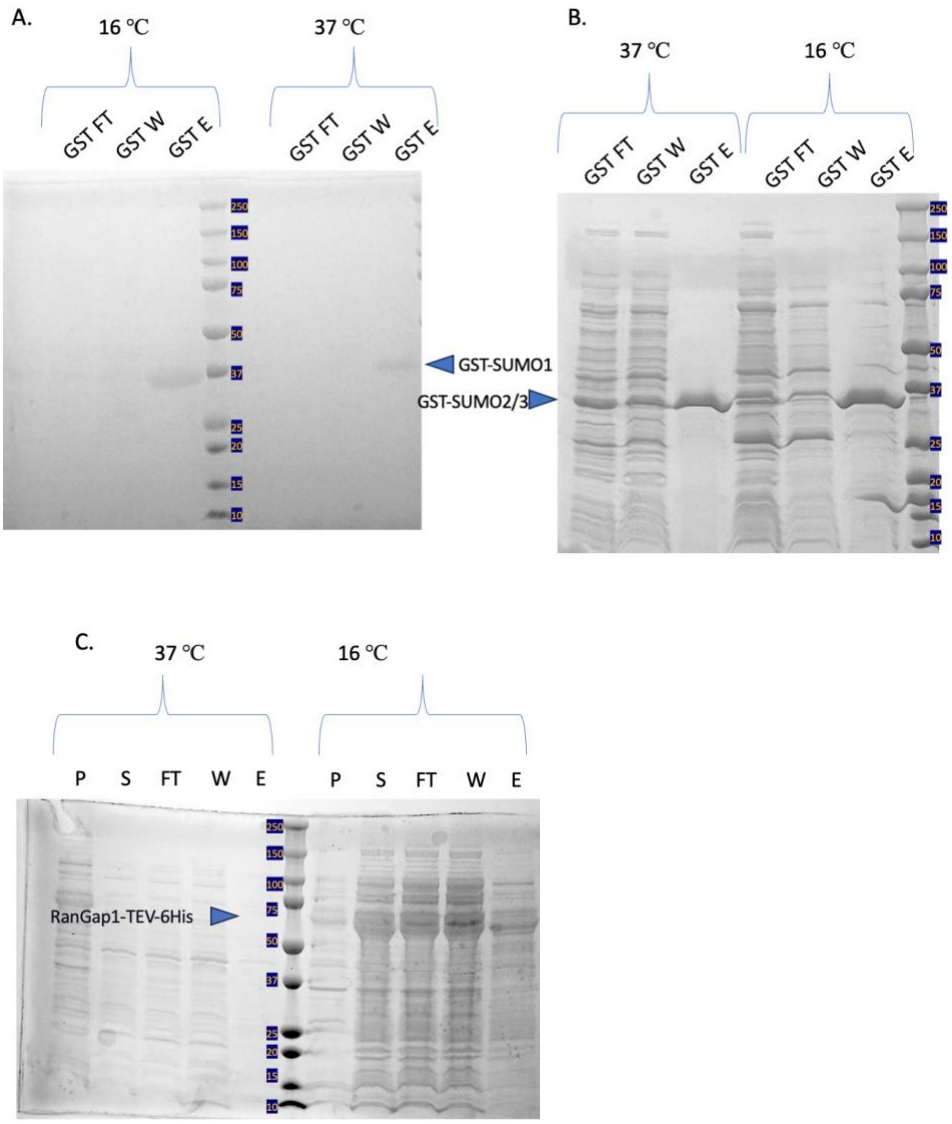
A.


37 ℃
14 ℃
P
S
FT
W
E
P
S
FT
W
E
250
150
100
75
50
37
25
20
15
10
Uba2-TEV-6His

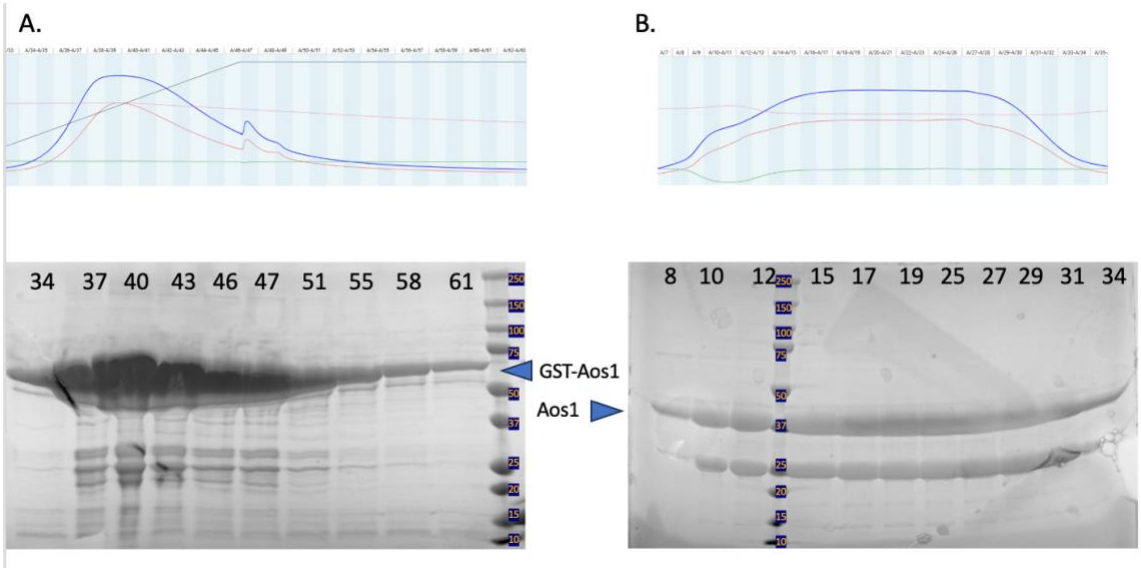
B.

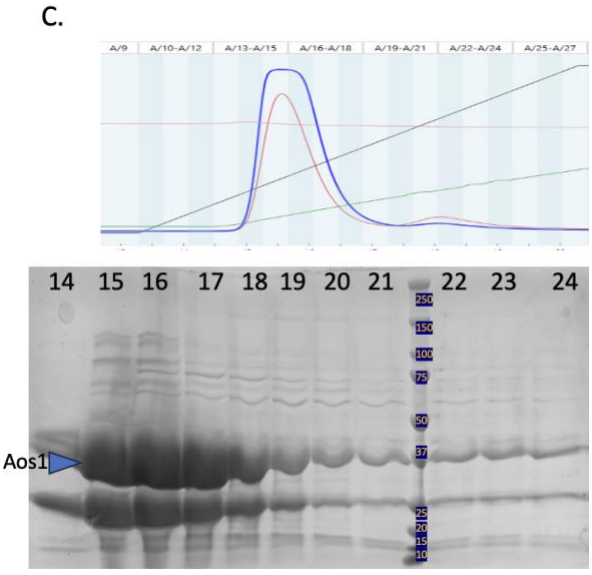

Post TEV O/N
FT
W
E
250
150
100
75
50
37
25
20
15
10
Uba2-TEV-6His
Uba2

Supplementary material figure 2.2 Uba2 expression small-scale test. Uba2-TEV-6His expressed in *E. coli ArcticExpress* RP with 2XYT and induced by IPTG @ OD 1.2 at 14℃ O/N or 37℃ 3-hour. Cells were harvested and examined by NI-NTA spin column. Samples were loaded on an 8-16% SDS-PAGE gel. SDS-PAGE gels showed elution samples from Ni-NTA (A) and after TEV O/N treatment (B). P: precipitant. S: supernatant. FT: flow through.



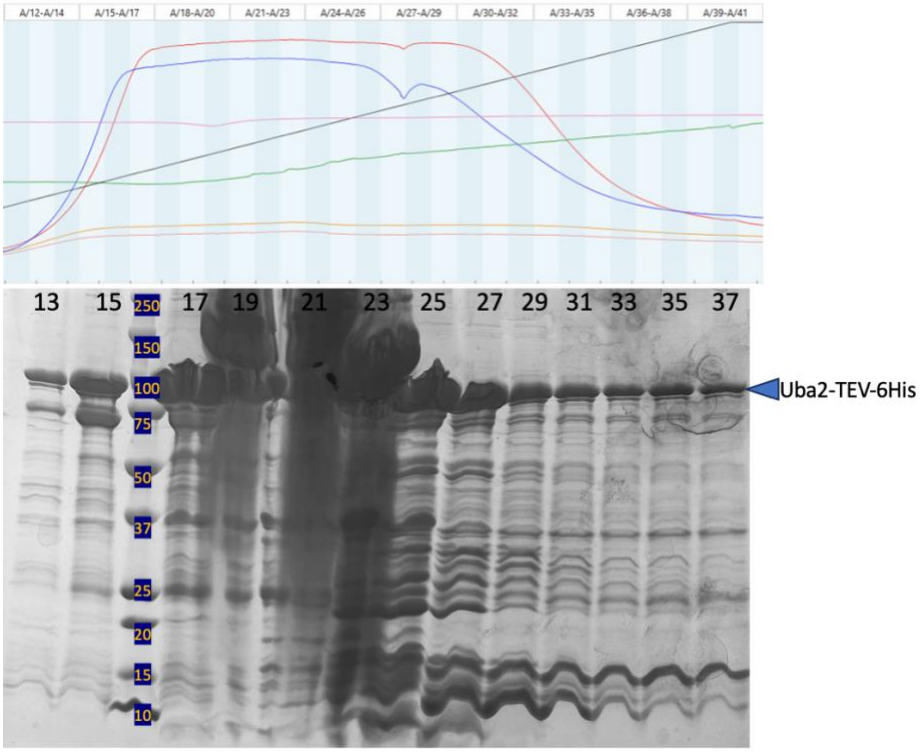
Supplementary material figure 2.3 SUMO1, SUMO2/3, & RanGap1 expression small-scale test. A. GST-SUMO1, B. GST-SUMO2/3, and C. RanGap1-TEV-6His were transformed individually in *E. coli* BL21(DE3) and followed by the GST or Ni-NTA spin column tests. Inducement at OD600 0.8 and set to shake at 16℃ O/N or 37℃ 3 hours. Cells were harvested and examined by NI-NTA spin column. Samples were loaded on an 8-16% SDS-PAGE gel.


A.
B.
34 37 40 43 46 47 51 55 58 61
GST-Aos1
Aos1
8 10 12 15 17 19 25 27 29 31 34

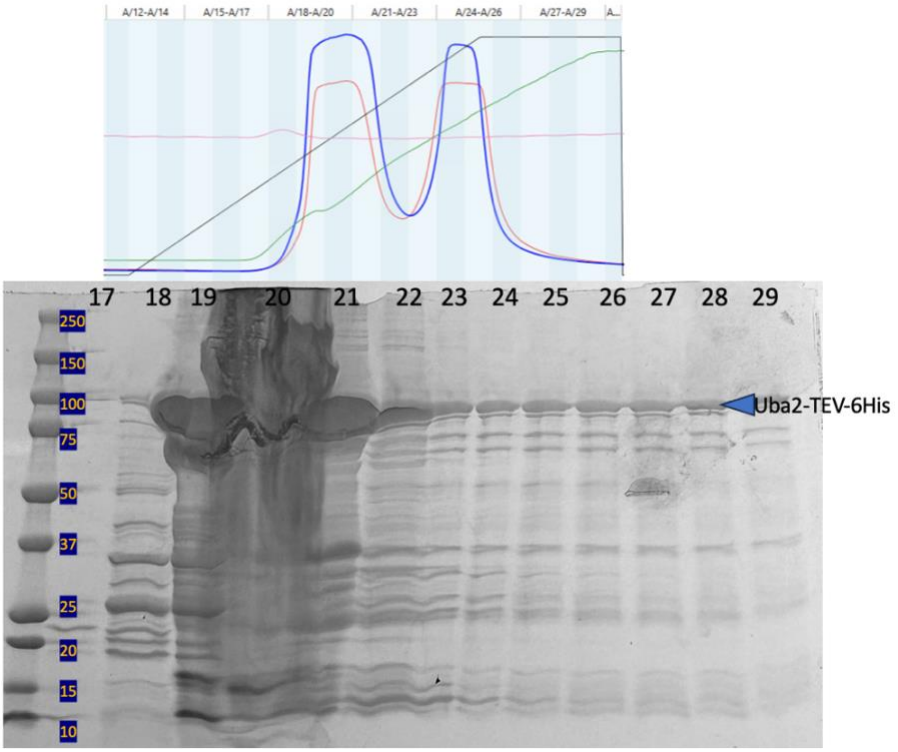

C.
A/9
A/10-A/12
A/13-A/15
A/16-A/18
A/19-A/21
A/22-A/24
A/25-A/27
14 15 16 17 18 19 20 21 22 23 24
Aos1

Supplementary material figure 2.4 Aos1 purification by fast protein liquid chromatography (FPLC). GST-Aos1 in *E. coli BL*21(DE3) with 2XYT induced by IPTG @ OD 1.2 at 14℃ O/N. A. GSTPrep captured GST-Aos1 in elution fractions 34-63. B. PreScission O/N treatment removed GST tag from Aos1, and another GSTPrep captured Aos1 in flow through fractions 8-34. C. After buffer exchange, sample was purified by QFF in fraction 14-19. Buffers were used: GST A 50mM HEPES ph8.5, 300mM NaCl, 1mM BME & GST B 50mM HEPES ph8.5, 300mM NaCl, 25mM glutathione, 1mM BME; QFF A 50mM HEPES ph8.5, 50mM NaCl, 1mM BME & QFF B 50mM HEPES ph8.5, 1M NaCl, 1mM BME.

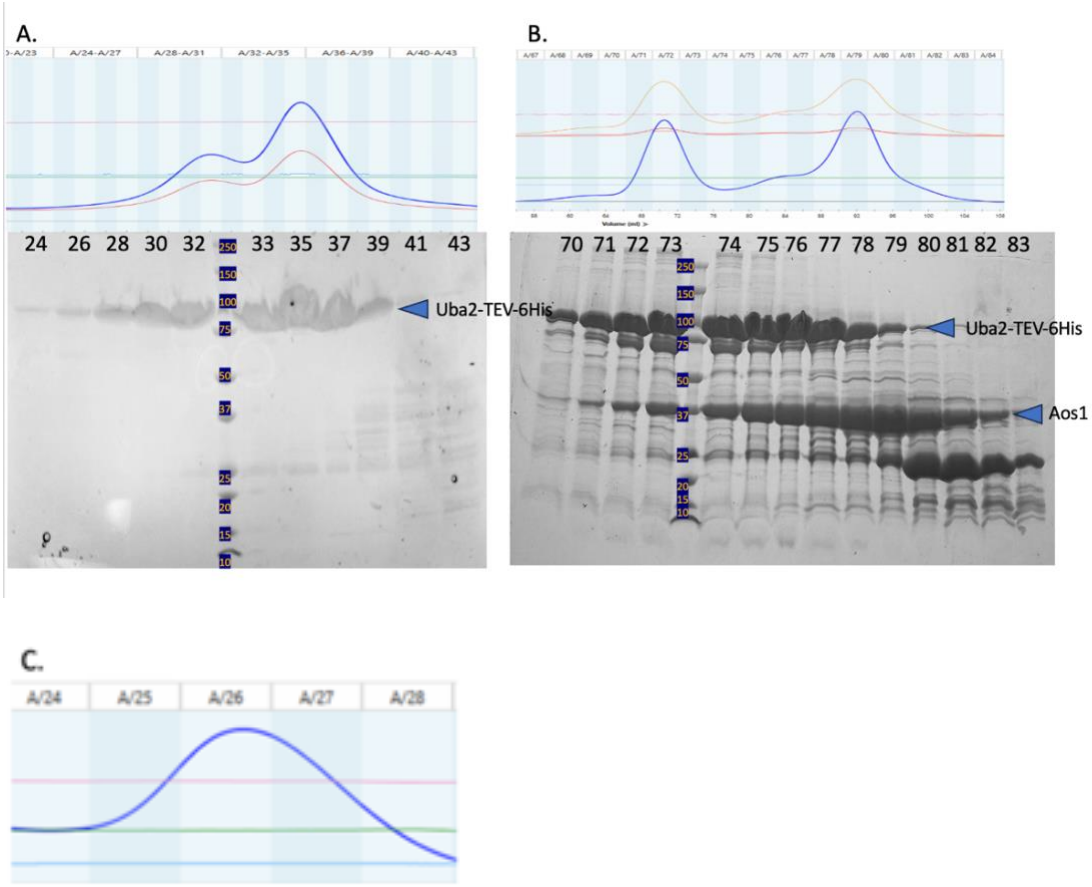
A.


A/12-A/14
A/15-A/17
A/18-A/20
A/21-A/23
A/24-A/26
A/27-A/29
A/30-A/32
A/33-A/35
A/36-A/38
A/39-A/41
13 15 17 19 21 23 25 27 29 31 33 35 37
250
150
100
75
50
37
25
20
15
10
Uba2-TEV-6His

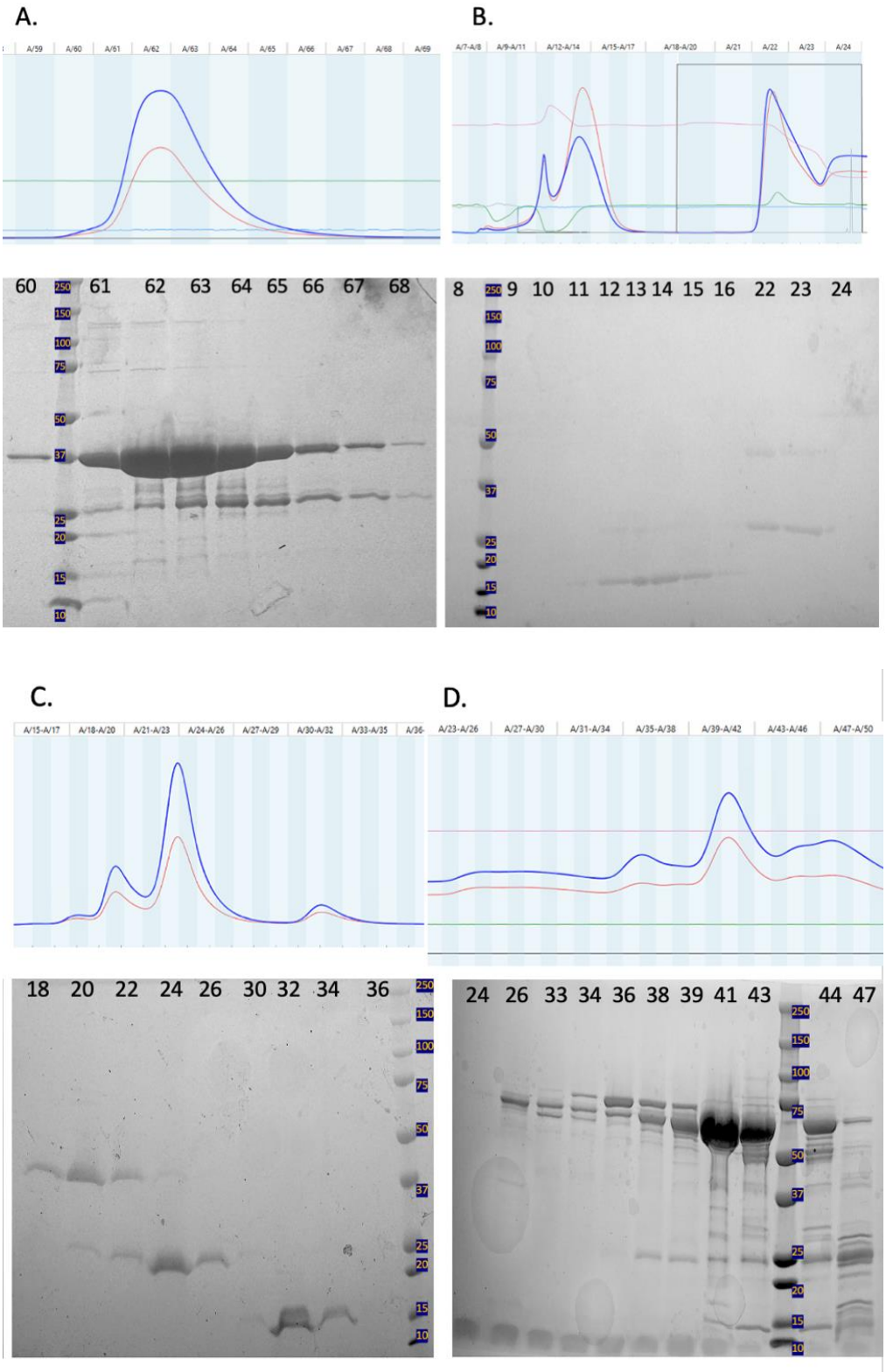
B.


A/12-A/14
A/15-A/17
A/18-A/20
A/21-A/23
A/24-A/26
A/27-A/29
17 18 19 20 21 22 23 24 25 26 27 28 29
250
150
100
75
50
37
25
20
15
10
Uba2-TEV-6His

Supplementary material figure 2.5 Uba2 purification by FPLC. Uba2-TEV-6His in *E. coli ArcticExpress* RP with 2XYT and induced by IPTG @ OD 1.2 at 14℃ O/N. A. HisTrap FF captured Uba2 in elution fractions 13-27. B. Buffer exchanged and then purified by QFF in fractions 19-22. Buffers were used: His A 50mM HEPES ph8.5, 300mM NaCl, 25mM imidazole, 1mM BME & His B 50mM HEPES ph8.5, 300mM NaCl, 1M imidazole, 1mM BME; QFF A 50mM HEPES ph8.5, 50mM NaCl, 1mM BME & QFF B 50mM HEPES ph8.5, 1M NaCl, 1mM BME.

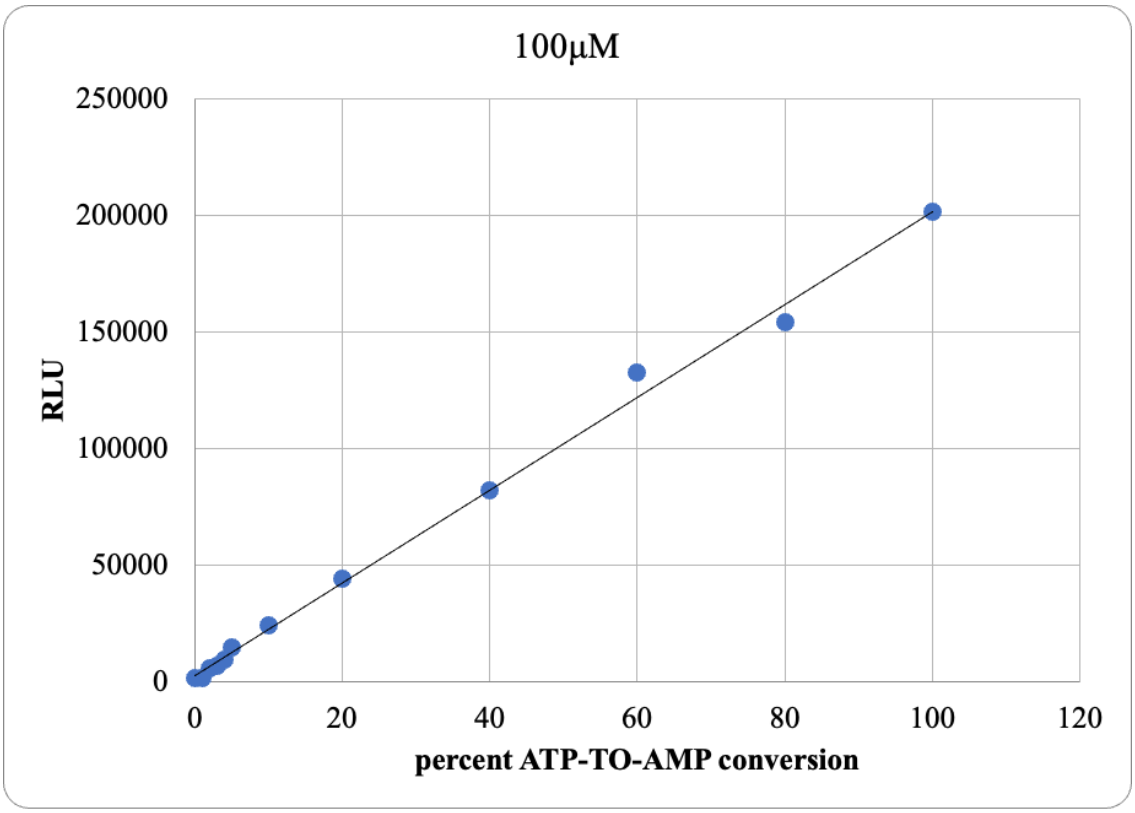


Supplementary material figure 2.6 Uba2 FPLC. A. Uba2-TEV-6His after QFF was purified by Superdex 200 without in fraction 28-31 or B. with Aos1 (right) in fraction and C. followed by ENrich Q in fractions 26-27. and buffer exchange. Buffers were used: 50mM HEPES ph7.5, 300mM NaCl; ENrich Q A 50mM HEPES ph7.5, 50mM NaCl, 1mM BME & ENrich Q B 50mM HEPES ph7.5, 1M NaCl, 1mM BME. Yield: 1.04 mg of Uba2 / L of *E. coli* culture & 1.8 mg of Uba2/Aos1 / L.


A.
B.
C.
D.
60 61 62 63 64 65 66 67 68
8 9 10 11 12 13 14 15 16 22 23 24
18 20 22 24 26 30 32 34 36
24 26 33 34 36 38 39 41 43 44 47

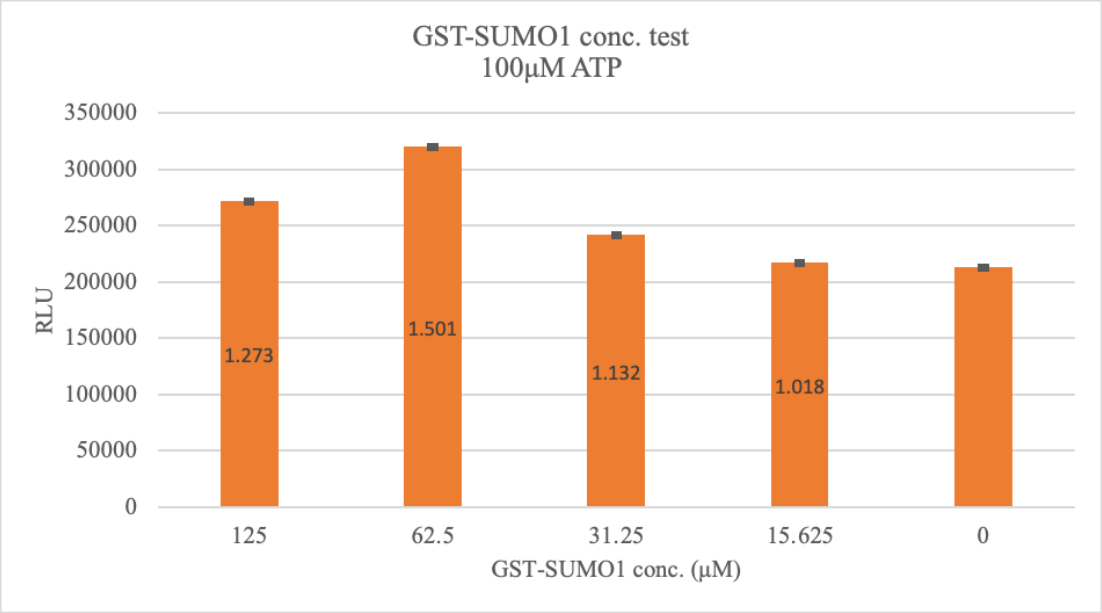
Supplementary material figure 2.7 SUMO1, SUMO2/3, RanGap FPLC. GST-SUMO1, GST-SUMO2/3, and RanGap1-TEV-6His followed the steps of affinity (6His or GST), HiTrap Q FF, and Superdex 75. A. GST-SUMO1 after Superdex 75 fractions 60-64. B. after PreScission treatment, SUMO1 was in the flow through after GSTPrep in fraction 12-16. C. GST-SUMO2 was in fractions 18-20 & SUMO2 was in fractions 32-34 after Superdex 75. D. RanGap1-TEV-6His was in fraction 40-43 after Superdex 75. The yield was 4.5 mg of GST-SUMO1 or GST-SUMO2/3 per 2 L of *E. coli* culture and 10 mg of RanGap1-TEV-6His per 4 L of *E. coli* culture.

A.

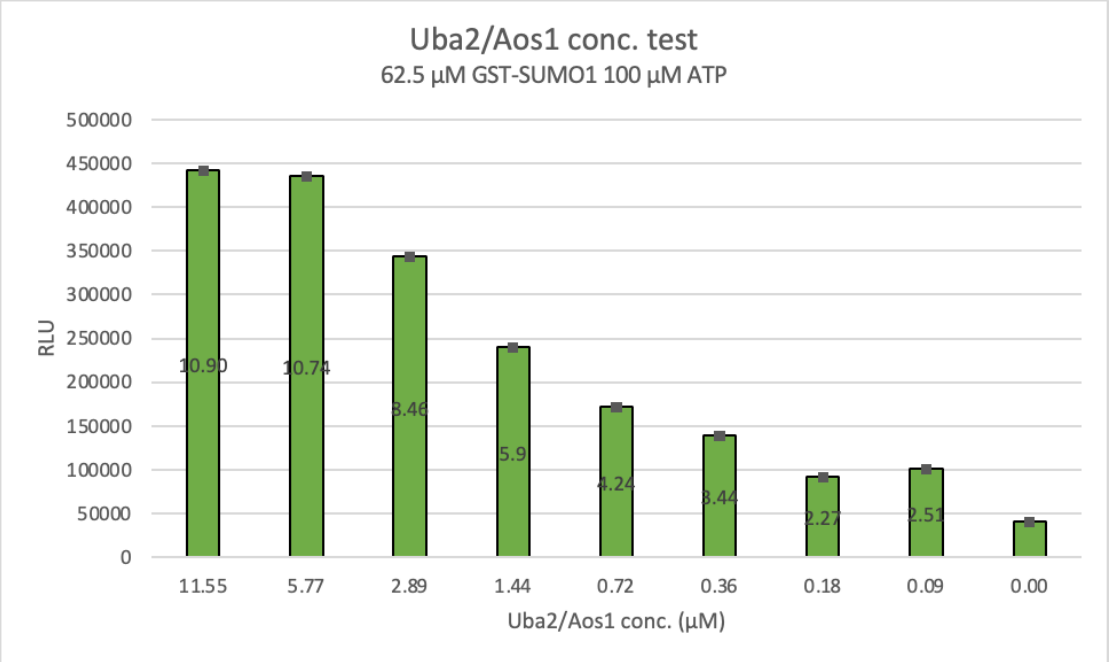


	2	3	4	5	6	7	8	9	10	11	12	
100	80	60	40	20	10	5	4	3	2	1	0	ADP % in ATP+ADP mixture
0	20	40	60	80	90	95	96	97	98	99	100	ATP % in ATP+ADP mixture

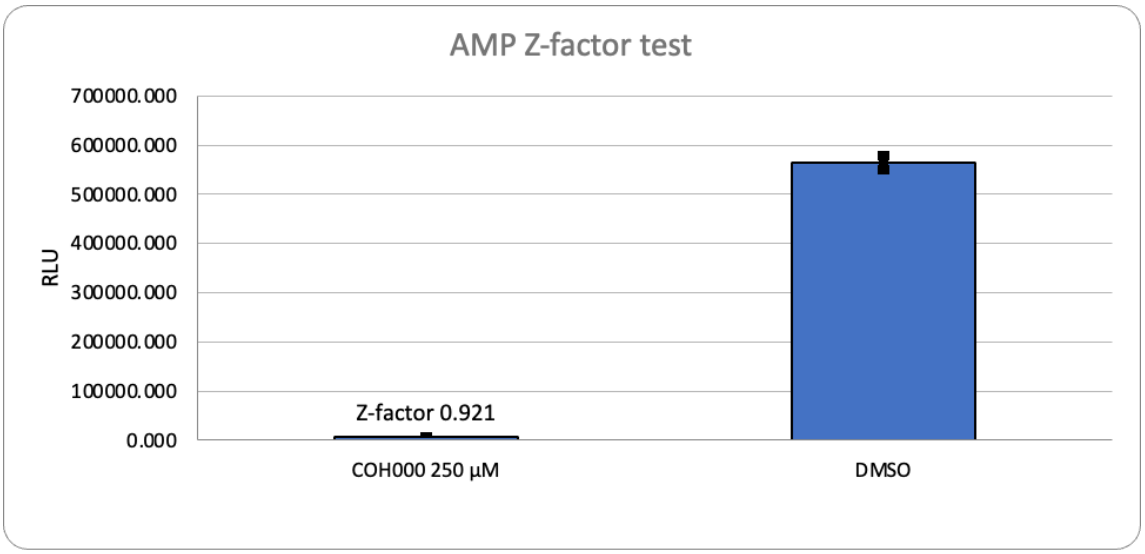
B.


GST-SUMO1 conc. test
100μM ATP
RLU
350000
300000
250000
200000
150000
100000
50000
0
1.273
1.501
1.132
1.018
125
62.5
31.25
15.625
0
GST-SUMO1 conc. (μM)

C.


Uba2/Aos1 conc. test
62.5 μM GST-SUMO1 100 μM ATP
RLU
500000
450000
400000
350000
300000
250000
200000
150000
100000
50000
0
10.90
10.74
3.46
5.9
4.24
3.44
2.27
2.51
11.55
5.77
2.89
1.44
0.72
0.36
0.18
0.09
0.00
Uba2/Aos1 conc. (μM)

D.



Supplementary material figure 2.8 preparing for the AMP assay. A. ATP-to-AMP conversion standard curve was generated by mixing the appropriate volumes of ATP and AMP. The standard curve could be used to calculate the ATP concentration used in later kinase reactions. B. GST-SUMO1 concentration test was done with a serial dilution of GST-SUMO, from 125 µM to 15.625 µM, and no Dyrk1a as background signal. Uba2/Aos1 was undiluted 11.55 µM. The S/B ratio was labelled in each column. C. Uba2/Aos1 concentration test was done with a serial dilution of Uba2/Aos1, from 11.55 µM to 0.09 µM, and no Uba2/Aos1 as background signal. GST-SUMO1 was 62.5 µM. The S/B ratio was labelled in each column. D. AMP assay Z-factor test using the known Uba2/Aos1 inhibitor COH000.

LC covalent fragment	TSA average ΔT_m ℃ (500 μM)	TSA dose trend	AMP IC_{50} μM
2A	2.50	Yes	39.8
1H	2.17	Yes	105.8
1C	2.33	Yes	2494
1E	3.50	Yes	>50
1F	3.00	Yes	>50
2E	1.17	No	
1B	-1.50		
1D	0.67		
1G	-1.67		
2B	-2.83		
2C	-2.00		
2D	-2.17		
2F	-1.17		

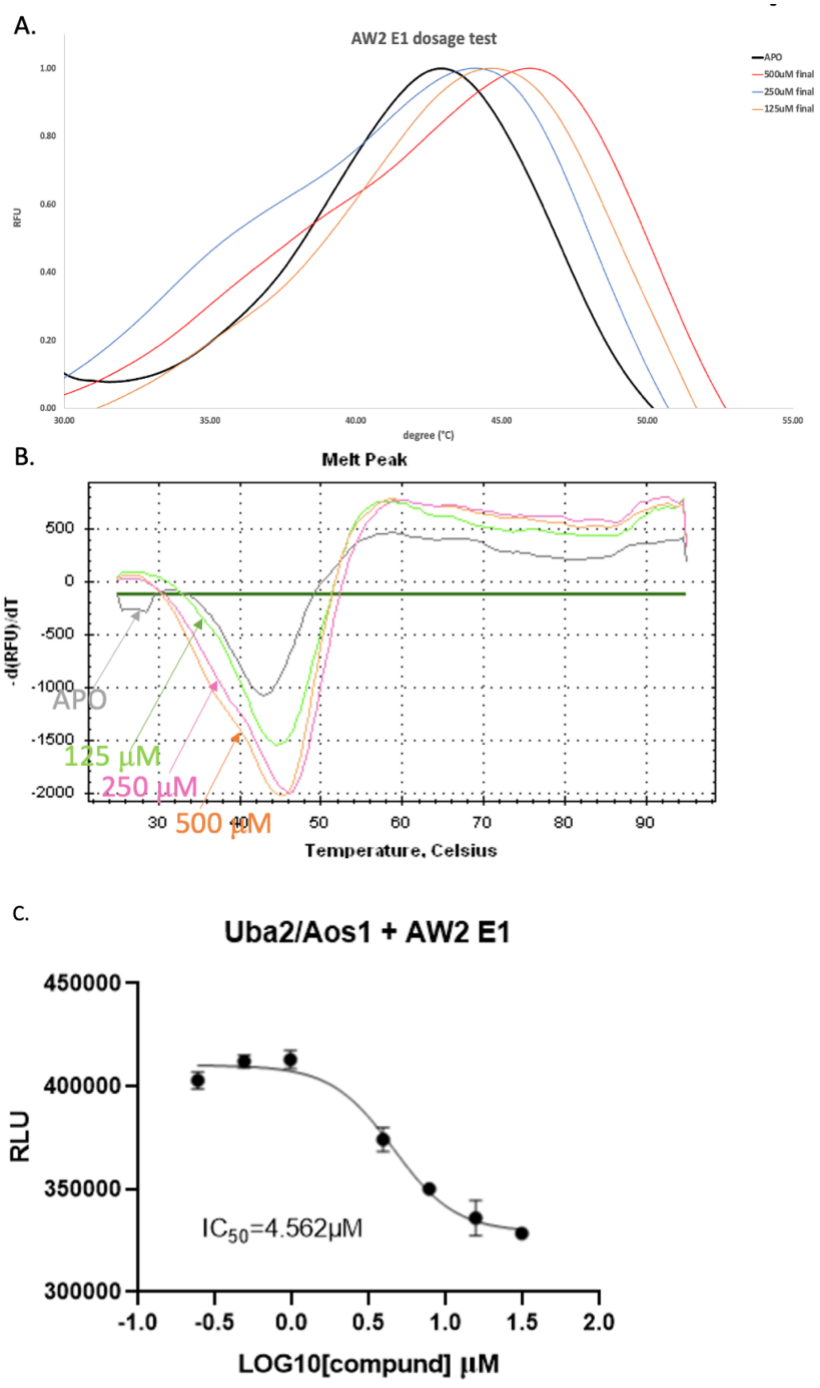
Supplementary material table 2.1 summary of Uba2/Aos1 with LC covalent fragment binders using TSA and AMP assay.

13 LC covalent fragments positively shifted T_m of Uba2/Aos1 in the first round of screening (10mM final concentration). And then they were tested with diluted concentration (500 μM) and dosage/concentration test by TSA. Their IC_{50} values were measured by ADP assay and calculated by GraphPad.

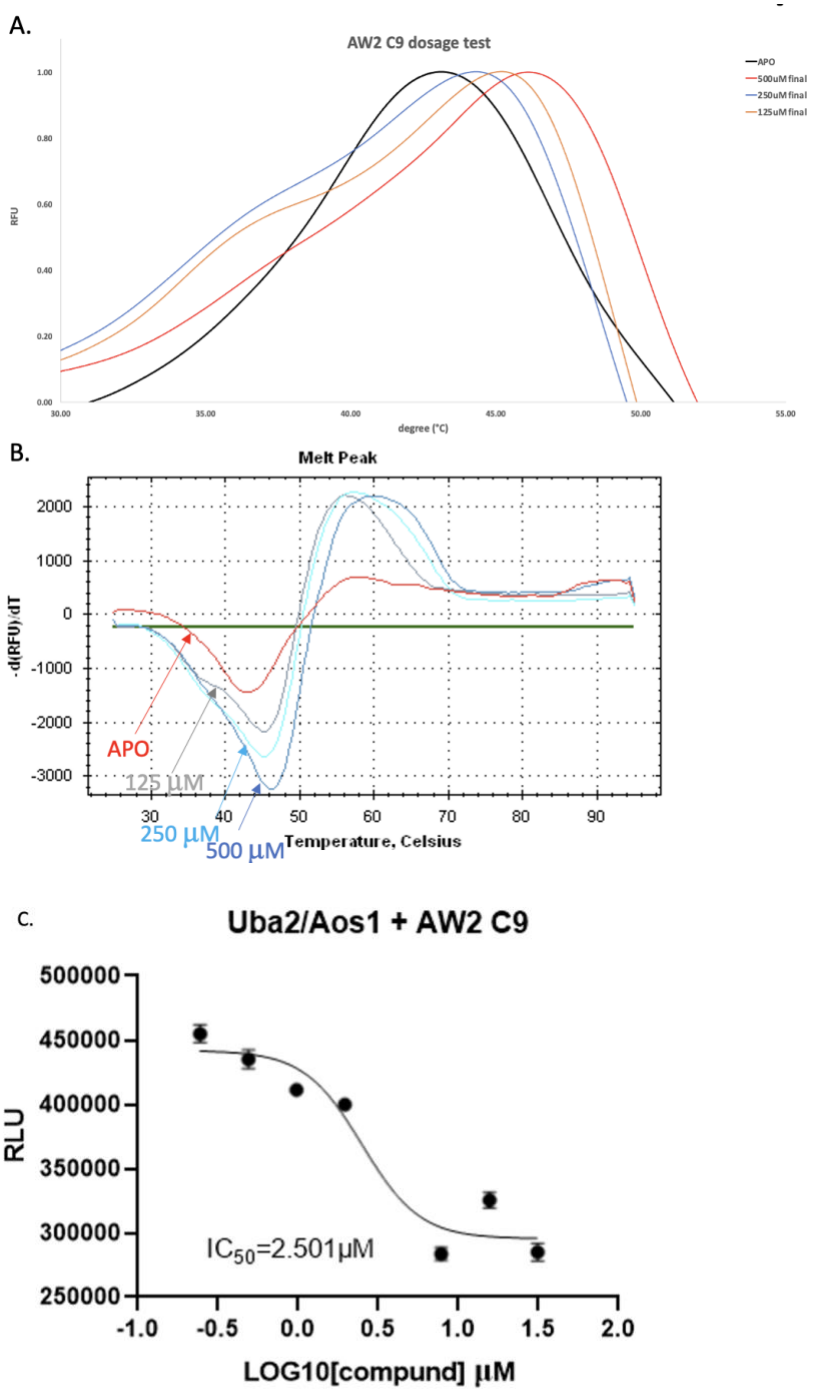
AW1 compound	TSA average ΔT_m (500 µM)	amp % activity (500 µM)	TSA dose trend	TSA dose trend uba2 only	AMP IC_{50} µM
A07	0.33	10.67			>50
A08	0.83	1.10			>50
A09	0.17	101.54			>50
B01	1.17	115.19	✓	✗	
B02	1.33	101.26	✓	✗	
A01	-0.67	107.39			
A02	0.83	106.68			
A03	0.50	92.43			
A04	0.50	119.73			
A05	N/R	110.99			
A10	0.67	121.53			
A12	0.83	118.69			
B03	-0.83	98.67			
B04	2.83	101.06	✗		
AW2 compound	TSA ave del tm (500uM)	amp % activity (500uM)	TSA dose trend	TSA dose trend uba2 only	AMP IC50 uM *
F05	2.50	58.95	✓	✓	0.498
C09	3.67	45.92	✓	✓	2.501
E01	1.33	65.94	✓	✓	4.562
B07	3.50	69.90	✓	✓	10.880
C10	4.33	105.07	✓	✓	>50
F03	5.50	116.53	✓	✓	>50
A07	1.50	122.41	✓	✗	>50
C02	1.17	121.65	✓	✗	>50
D06	1.00	119.02	✓	✗	>50
E08	1.67	111.99	✓	✗	>50
A08	2.33	117.50	✗		
A12	2.17	87.79	✗		
B05	1.00	80.20	✗		
B08	3.50	114.75	✗		
D03	1.00	117.94	✗		
D04	2.33	111.26	✗		
E04	1.50	116.32	✗		
G07	5.33	109.49	✗		

Supplementary material table 2.2 summary of Uba2/Aos1 with AW1 & 2 AI-CADD small molecules using TSA and AMP assay.

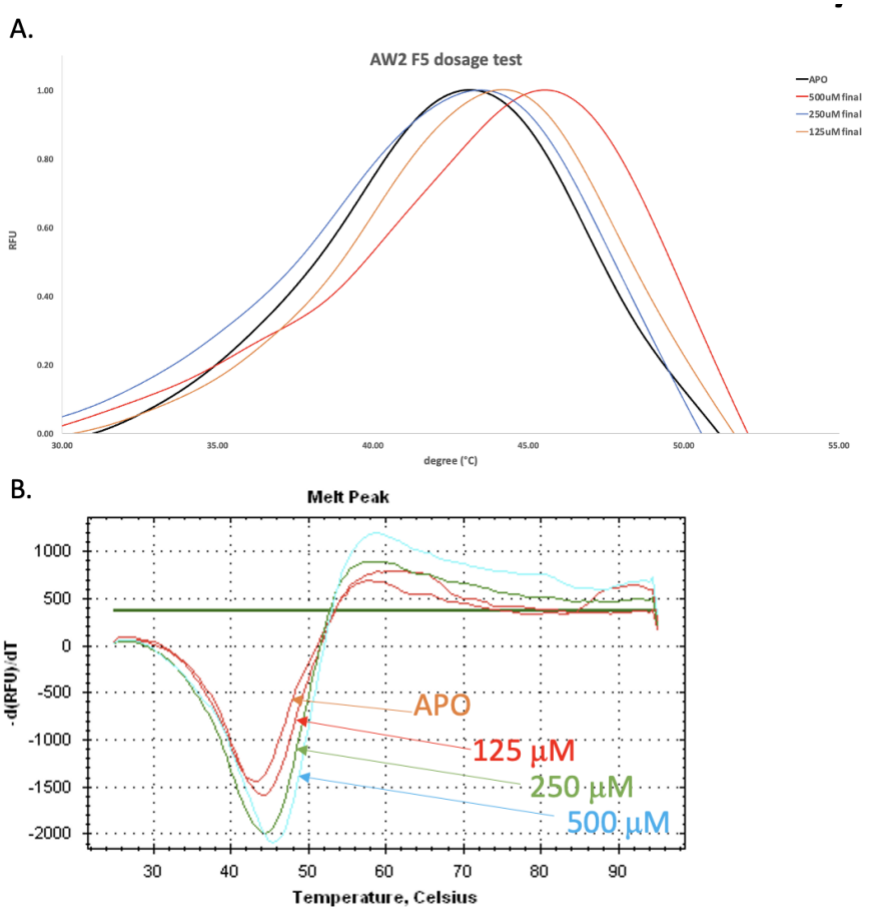
AW1 & 2 AI-CADD small molecules were tested against to Uba2/Aos1 along with DMSO as negative control in triplicates. The average of ΔT_m and dosage trend in TSA determined whether the small molecule was a binder of Uba2/Aos1. The average of % activity and dosage trend in the AMP assay determined whether the small molecule was an inhibitor of Uba2/Aos1.IC_{50} values were measured for inhibitors.



Supplementary material figure 2.9 Uba2/Aos1 with AW2 E1. AW2 E1 dosage/concentration test by TSA against to Uba2/Aos1 A. RFU vs. degree °C. B. -d(RFU)/dT vs. degree °C. C, IC_{50} figure by AMP assay and GraphPad.



Supplementary material figure 2.10 Uba2/Aos1 with AW2 C9. AW2 C9 dosage/concentration test by TSA against to Uba2/Aos1 A. RFU vs. degree °C. B. -d(RFU)/dT vs. degree °C. C, IC_{50} figure by AMP assay and GraphPad.

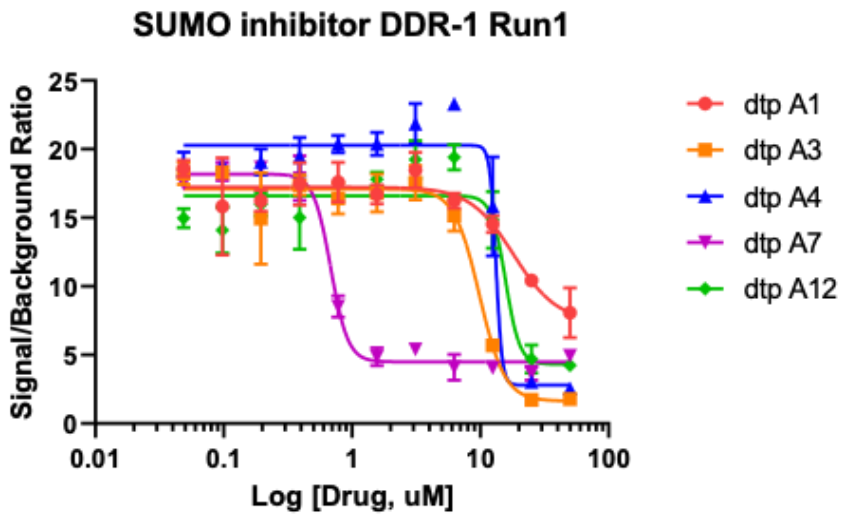


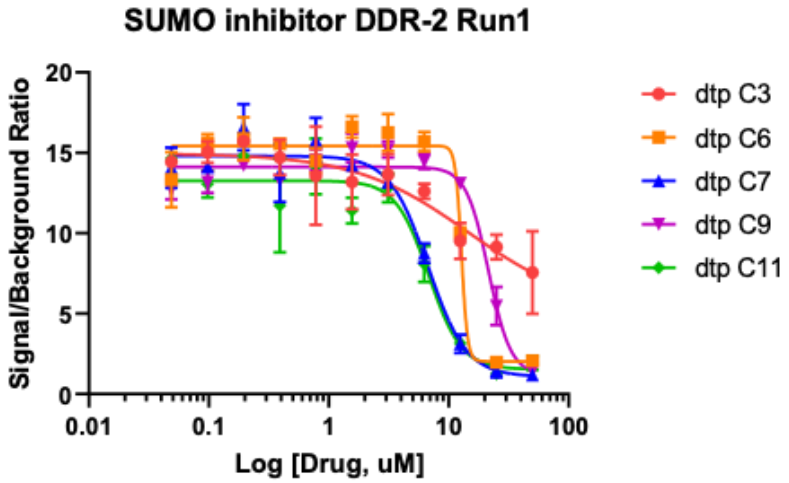
Supplementary material figure 2.11 Uba2/Aos1 with AW2 F5. AW2 F5 dosage/concentration test by TSA against to Uba2/Aos1 A. RFU vs. degree °C. B. -d(RFU)/dT vs. degree °C.

DTP	AMP % activity (50 µM)	HTRF % activity (50 µM)	AMP best fit IC_{50} µM *	HTRF IC_{50} µM	AMP 95% IC_{50} µM	AMP R^2
C7	9.24	10.24	18.41	8.01	14.75 to 22.46	0.9952
A7	2.55	36.35	35.37	1.01	29.11 to 42.45	0.9928
C3	24.89	42.83	41.76	10.38	22.33 to 980.3	0.938
C11	17.43	16.72	>50	9.24		
C6	27.38	17.80	>50	15.99		
A3	30.27	11.56	>50	16.48		
A4	65.76	59.11	>50	17.08		
A1	35.25	65.83	>50	20.72		
C9	39.05	21.11	>50	25.41		
A12	43.79	54.61	>50	>50		
A11	6.89	49.97	>50			
C8	10.16	123.53	>50			
B10	4.34	96.72	>50			
B3	12.28	93.78	>50			
C10	14.76	129.33	>50			
B12	24.5	139.71	>50			
B4	27.44	95.24	>50			
B2	53.75	102.87				
C5	64.32	107.3				
B11	69.45	119.13				
A2	82.9	93.32				
A5	83.04	95.12				
D2	84.88	79.08				
D4	85.83	86.85				
D1	87.59	73.09				
C4	93.01	112.9				
A10	98.82	88.99				
D3	100.3	90.49				
B1	106.92	89.12				
B7	107.95	102.89				
B6	110.54	107.26				
A6	112.71	87.2				
B8	117.14	60.92				
A8	120.22	119.37				
B5	125.73	94.85				

Supplementary material table 2.3 summary of Uba2/Aos1 with DTP small molecules using TSA and AMP assay.

DTP small molecules were tested against to Uba2/Aos1 along with DMSO as negative control in triplicates. The average of % activity and dosage trend in the AMP assay determined whether the small molecule was an inhibitor of Uba2/Aos1.IC_{50} values were measured for inhibitors.





Supplementary material figure 2.12 IC50 figure of Uba2/Aos1 with DTP small molecules using HTRF assay. Collaborating with the HTS group, City of Hope, DTP small molecules were tested against to Uba2/Aos1 along with DMSO as negative control in triplicates. The IC_{50} values were measured with the inhibitors. Figure was generated by Dr. David Kwon.

Chapter 3
SUMO E2 Ubc9 and SUMO E3 Pias1 early drug discovery

Chapter 3: SUMO E2 Ubc9 and SUMO E3 Pias1 early drug discovery

3.1 Abstract

The best characterized and sole human SUMO conjugating enzyme E2, is Ubc9, also referred to as Ube2I, which is a 148 amino acid protein with an active cysteine residue, Cys73. Ubc9 interacts with the activated SUMO via a thioester bond and can conjugate SUMO to its substrate. Ras, a family of small GTPase proteins, falls within the "undruggable group". KRas, one member of the RAS family, has mutations that account for approximately 30% of all human cancers, ~90% of pancreatic cancers, and 30-50% of colorectal carcinomas. As a synthetic lethality partner, Ubc9 is essentially required for KRas cancer cell viability. SUMOylation modification at the conserved Lys42 of Ras and SUMO-resistant Ras has significantly reduced activity compared with the wild type. SUMO E1 and E2 are sufficient for SUMOylation, but the SUMO E3 ligases enhance SUMOylation substrate recognition and specificity. Pias1 belong to the largest class of SUMO E3, is required for the SUMOylation of PCNA, which PCNA is involved in DNA replication. Pias1 SUMOylates at Cys42 of MYC, which promotes phosphorylation at Ser62 by ERK to stabilizes MYC, and dephosphorylation at Thr58 by GSK3b, which recruit ubiquitin-proteasome degradation of MYC. Pias1 is required for MYC in the primary B-cell lymphomas and the cell's viability. Pias1 also contributes to immunological functions by negatively regulating NF-κB/STAT signaling. We hypothesized that the Ubc9 inhibitors/chemical tools can inhibit SUMOylation and impact KRas-driven cancer cells, and the Pias1 inhibitors can inhibit SUMOylation and impact MYC-driven cancer cells.

3.2 SUMO E2 in cancer

The best characterized and sole human SUMO conjugating enzyme E2, is Ubc9, also referred to as Ube2I, which is a 148 amino acid protein with an active cysteine residue, Cys73. Other than Ubc9, there are about 40 E2s involved in ubiquitin or other ubiquitin-like, such as NEDD [152]. Ubc9 contains one N-terminal α helix and five β sheets, a loop with catalytic activity site Cys93, and three α helices. Ubc9 interacts with the activated SUMO via a thioester bond. It conjugates SUMO to its substrate by recognizing the ψ-K-X-D/E consensus motif if it is present in an unstructured region, such as a loop or coil [153]. Ubc9 C93S failed to recruit SUMO from the SUMO-Aos1/Uba2 complex. The N-terminal helix and the loop following the helix of Ubc9 can form a noncovalent interaction with SUMO, and this noncovalent complex promotes SUMO2 chain formation [154].

3.2.1 Ras-driven cancer

Ras, a family of small GTPase proteins, falls within the "undruggable group" [155, 106]. The exception is the G12C Ras mutation that occurs in 7.9% of KRas lung cancer cases, and this is discussed further below (Lindsay & Blackhall, 2019). KRas, one member of the RAS family, has mutations that account for approximately 30% of all human cancers [156], ~90% of pancreatic cancers [157], and 30-50% of colorectal carcinomas [158]. KRas mutations are always point mutations, as compared to Myc, which amplifies and/or overexpresses. KRas plays essential roles in signal transduction: the receptor tyrosine kinase activates Grb2 that activates RAS-GEF (SOS). SOS then activates KRas and KRas stimulates the PI3K/AKT pathway, which could contribute to a variety of cancers, for example, renal neoplasm [159]. Mutation of KRas can additionally cause hyperactivation

of the MAPK pathway, which is also observed in various tumors [160]. The most common mutation of KRas is G12D in pancreatic ductal adenocarcinoma, leading to a poor treatment outcome and an overall 5-year survival rate of only 10% [161]. Another mutation, G12C, has been found in 13% of non-small cell lung cancers and 3-5% of colorectal cancers [162]. The different mutations of KRas are tissue specific [163]. KRas utilizes GTP to phosphorylate downstream proteins. The KRas-GDP/GTP cycle determines the activity level of KRas. Mutations of KRas, such as G12D and G12V, maintain binding with GTP and the activation form of KRas. The other mutation, KRas G12C, still maintains the cycle, and KRas has activation and inactivation forms [164]. However, the affinity of KRas G12C for GTP is in the picomolar range. Therefore, KRas G12C is still very active. AMG 510 is the first drug directly targeting KRas G12C mutation, after decades of development [13, 165]; FDA approved AMG 510 as a covalent inhibitor that targets the Cys12 residue in 2021. AMG 510 covalently binds to Cys12 of KRas G12C and converts KRas to the inactivation form. For KRas G12D, the most common KRas mutation, peptide KRprp-2d, was identified as a KRas G12D inhibitor [166].

Moreover, Ras, the so-call ‘undruggable’ target (except in the case of G12C), also activates another undruggable target, Myc. Ras activates MAPK and CDK2 that phosphate and activate Myc at its S62 site [167]. With Pin1, which targets Myc on residue P63 and converts Myc from a *trans* to a *cis* isomer, Myc can promote the expression of the target gene that drives tumorigenesis. Ras also inhibits GSK3, which functions to phosphorylate Myc on T58 for expected degradation via ubiquitination, and therefore, Ras also prevents Myc degradation.

SUMOylation modification at the conserved Lys42 of Ras and SUMO-resistant Ras has significantly reduced activity compared with the wild type [168]. As a synthetic lethality partner, Ubc9 is essentially required for KRas cancer cell viability [69]. The proliferation of KRas-driven cells is significantly reduced with Ubc9 shRNA treatment by increasing apoptosis.

3.2.2 Liver cancer

4.7% of new cases are liver cancer, yet 8.3% of new cancer deaths are caused by liver cancer. Among primary liver cancer, the most common type is hepatocellular carcinoma (HCC) (around 80% of all cases). Hepatitis B virus (HBV), hepatitis C virus (HCV), type 2 diabetes, and diet are risk factors for HCC. About 40% of patients are diagnosed at an early stage, and the 5-year survival rate is 36 %; this survival rate drops to 13% if cancer spreads to surrounding tissue and it is only 3% if it spreads to other body parts [126, 133].

Using one-way ANOVA, *Ubc9* mRNA is higher expressed in HCC than in normal tissue [169]. Moreover, Ubc9 DNA is overexpressed in pan-digestive system tumors (DSTs), such as cholangiocarcinoma, colon adenocarcinoma, rectal adenocarcinoma, stomach adenocarcinoma, pancreatic adenocarcinoma.

Data shows that the Ubc9 protein level is significantly higher in HCC tissues than in normal tissue as shown by immunohistochemistry and western blot [170]. The level of Ubc9 is associated with tumor size and aggressiveness. When knockdown Ubc9 using shRNA, the expression of cleaved-Caspase3 was increased, indicating Ubc9 inhibition leads to HCC cell apoptosis. Treating HCC cells with shUBC9 and doxorubicin, a treatment for HCC, leads to higher apoptosis than doxorubicin treatment. Data shows Ubc9

plays a role in HCC drug resistance, and inhibition of Ubc9 increases HCC sensitivity to doxorubicin. Together, data has shown that an inhibitor of Ubc9 could be a potential treatment for HCC.

3.2.3 Colon cancer and NF-κB2/p100

Colon cancer incidence of was high among adults older than 50, but it has dropped with increased screening [126, 133]. However, the incidence of colon cancer in young people has increased. The 5-year survival rate at the localized stage is 91%, and 72% at the regional stage. However, it drops to only 13% if it is at a distant stage. Using colon tumor line, the knockdown of Ubc9 by siRNA significantly decreases p100 and p52 as a NF-κB2 precursor, and intimately impact NF-κB2 [171]. Although no data shows that the inhibition of Ubc9 would kill colon cancer, data shows that Ubc9 plays a vital role in the NF-κB pathway.

3.3 E2 Ubc9 inhibitors

A couple of synthetic small molecules target Ubc9, and the most inhibitory is Spectomycin B with IC_{50} 4.4 μM [172]. Spectomycin B is a natural product small molecule that binds to Ubc9 and selectively inhibits the Ubc9-SUMO intermediate. 2-D08 inhibits SUMO1 transferring from Ubc9 to SUMO substrate with IC_{50} 6 μM [173].

This Ubc9 study aimed to identify chemical tools to control the activity of the SUMO E2 Ubc9, which could be a novel solution/finding future drug discovery-related studies. We hypothesized that the Ubc9 inhibitors/chemical tools can inhibit SUMOylation and impact KRas-driven cancer cells. We used SBDD coupled with FBLD approaches.

3.4 Methods

3.4.1 Expression

Human Ubc9 (1–158) was cloned into vector pET21a (+) with an N-terminal 6x His tag (6His-TEV-Ubc9) (GenScript Biotech). 6His-TEV-Ubc9 was expressed by transforming plasmids in different *E. coli* cell lines, and then a small-scale expression test was done to determine the optimal cell line and expression condition, and then saved glycerol stock. The glycerol stock was cultured in 50 mL of LB with ampicillin at 37℃ O/N. The O/N culture was inoculated into 4L of LB with ampicillin at 37℃ and 225 rpm. When the OD600 was 0.3, the temperature was lowered to 30℃, and then the cells were induced with 0.4 mmol/L IPTG at OD600 0.8 and kept shaking in the incubator for five hours at 30℃. Cells were harvested by centrifugation at 4℃ and 4,000 rpm for 15 minutes, resuspended at a ratio of 3 ml lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 25 mM imidazole, 10 mM BME) per gram pellet. Per 40 mL of the resuspended cell, it was sonicated on the ice at 70% amplitude for 4 minutes, with 10 seconds on and 10 seconds off. The supernatant was obtained by centrifugation at 4℃ and 30,000 rpm for 40 minutes. The cell lysate was purified right after or stored at -80℃.

3.4.2 Purification

The supernatant from cell lysate in 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 25 mM imidazole, and 10 mM BME was purified by HisTrap FF and eluted by a gradient of 1M imidazole buffer (TEV was added to desired fractions to cleavage 6His tag), then exchanged to a low salt buffer, loaded onto HiTrap Q FF, and eluted by a gradient of 1M NaCl buffer, followed by Superdex 200 and eluted by a low salt buffer, and finally ENrich

Q and eluted by a gradient of 1 M NaCl buffer. SDS-PAGE gels confirmed fractions from each step, and then desired fractions were collected for the next step. The final sample was stored in 50 mM HEPES pH 8.5, 300 mM NaCl at -80℃ after being flash frozen in liquid nitrogen.

3.4.3 TSA

200 mM of covalent fragments were purchased from Life Chemicals Inc. Following the TSA method from 2.5.2.1, hits were determined using a final of 8 μM of Ubc9 with the reaction buffer (50 mM HEPES pH 8.0, 50 mM NaCl).

3.4.4 Testing compound

Drug-like fragments from Zenobia. Stock 200 mM in DMSO.

Covalent or Brominated fragments from Life Chemicals (LC). Stock 200 mM in DMSO.

3.4.5 Mass spectrometry

Using the purified 6His-Ubc9 mixed with covalent fragment hits or DMSO individually, the TGen group, City of Hope, designed and performed mass spectrometry (LC-MS/MS non-GMP) analysis. Proteins were digested with proteinase K for five minutes to get overlapping fragments and then reduced with iodoacetamide. The digested peptides were analyzed by mass spectrometer and analyzed data. QC passed threshold (>0.9) with a mean correlation coefficient of 0.98.

3.5 Results

3.5.1 Expression

In the small-scale expression test, 6His-TEV-Ubc9 plasmid was transformed within *E. coli* ArcticExpress RP or BL21-CondonPlus-RP and cultured in LB, TB, or 2XYT. Cells

were induced at OD600 0.8 and set to shake at either 30℃ 5 hours or 14℃ O/N. All cells were collected and then loaded onto nickel spin columns, eluted with high imidazole buffer, and run on 12% SDS-PAGE gels. Ubc9 is an 18 kDa protein, and a band of 18 kDa size were shown in elution samples from both cell lines. The expression level was higher using BL21-CondonPlus-RP and LB (**Supplementary material figure 3.1)**.

3.5.2 Purification

In the large-scale expression for purification, 6His-TEV-Ubc9 would be treated by TEV after the first HisTrap FF, aiming to gain SUMO E2 without the tag for the HTRF assay. To remove the His tag, another HisTrap FF would be done to separate Ubc9 from the 6His tag in the flow-through. QFF and Superdex 75 (SEC) were used to purify Ubc9 further. The yield was 2.5 mg of Ubc9/ L of *E. coli* culture (**Figure 3.**1 & **Supplementary material figure 3.2**).

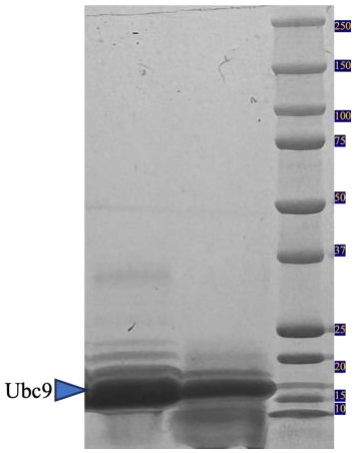


Figure 3.1 Purified Ubc9 sample. The final post-SEC sample of Ubc9 was loaded on an 8-16% SDS-PAGE gel (Novex Tris-Glycine Mini Protein Gel; Thermo Fisher). Lane 1 & 2: final sample from SEC elution fractions. Lane 2: protein ladder (Precision Plus Protein Dual Color Standards; Bio-Rad).

3.5.3 TSA

The optimized protein concentration and reaction buffer for TSA were conducted first. 8 μM of Ubc9 in the reaction buffer 50 mM HEPES pH 8.0, 50 mM NaCl generated the desired RFU and relatively low T_m for TSA of Ubc9, which was 54.5°C (**Supplementary material figure 3.3**). During the screening test, the drug-like fragments were the first group of compounds to be tested against SUMO E2. 57% of the drug-like fragments had positively shifted at least 0.5°C of the melting point of SUMO E2, and 5.27% had positively shifted at least 2°C, indicating SUMO E2 is highly druggable (**Figure 3.2**).

The second group was the covalent fragments and the brominated fragment. The enzymatic activity site of Ubc9 is residue Cys 93. Therefore, covalent fragments can bind and inhibit it. The percentage of 2°C ΔT_m was 0.97 using covalent (18 fragments) and brominated fragments (**Figure 3.2**). A concentration/dosage of fragment test was done with those fragments that showed the most T_m shift, and 17 had a dosage trend. Mass spectrometry determined the binding regions (**Figure 3.3)**.

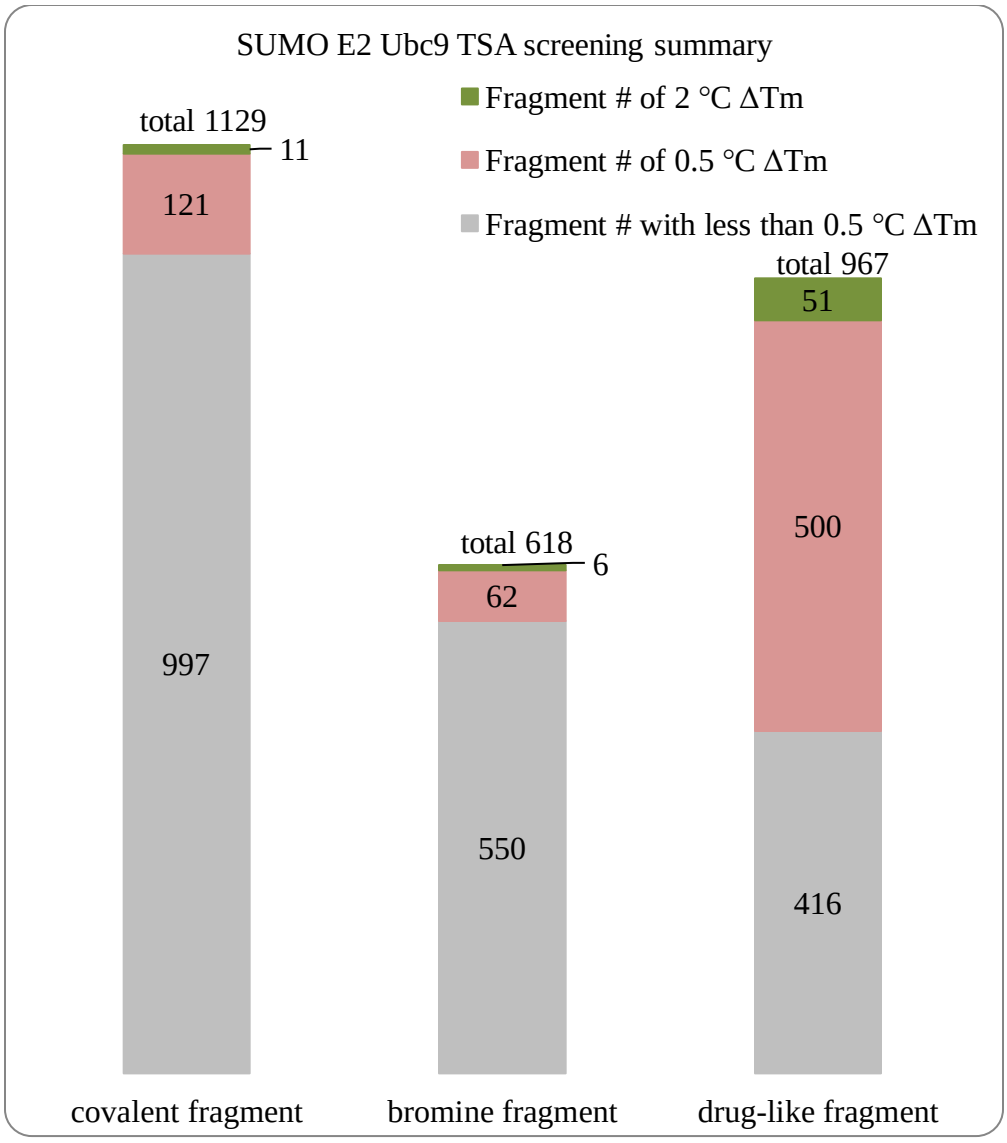


Figure 3.2 Summary of SUMO E2 Ubc9 TSA screening with covalent, brominated, and drug-like fragment. 2714 fragments (1129 covalent, 618 brominated, and 967 drug-like) were tested against Ubc9 by TSA. The % of fragments with at least 0.5 °C ΔT_m were 11.69, 11, and 57, respectively.

hits	10 mM °C ΔTm	5 mM °C ΔTm	2.5 mM °C ΔTm	mass spec binding region
b3	**1**	0.5	0	16 – 44 (2*K, 1*C), 144 – 158 (3*K),
c1	**0.5**	0	0	14 – 45
b8	**0.5**	0.5	0	15 – 45
b5	**2.5**	0.5	0	16 – 45
a8	**0.5**	0	0	31 - 44
a1	**1.5**	1	0.5	NA
a2	**0.5**	0	-0.5	NA
a3	**0.5**	0.5	0	NA
a4	**1**	0	0	NA
a5	**1**	0	0	NA
a6	**1**	0	0	NA
a7	**1**	0.5	0.5	NA
b1	**1**	0.5	0	NA
b2	**1**	0.5	0.5	NA
b4	**1**	0.5	-0.5	NA
b6	**1**	0.5	-0.5	NA
b7	**0.5**	0	-0.5	NA
c2	**1**	0	1	NA

Table 3.1 Ubc9 covalent fragments dosage/concentration test by TSA and binding region by mass spectrometry. Covalent fragment binders from the screening were tested by the dosage/concentration test to confirm the binding event. The final concentration of fragments was 10, 5, and 2.5 mM in the dosage/concentration test. They were tested by mass spectrometry for covalent binding and indicated binding region.

3.5.4 Mass spectrometry

Collaborating with the TGen group at City of Hope, all 18 covalent fragments were tested by liquid chromatography with tandem mass spectrometry (LC-MS-MS non-GMP) to see where in the polypeptide chain the covalent fragments bound. When there were regions with decreased abundance, it corresponded to the presence of covalent fragment inhibitors. Five fragments showed binding to regions of Ubc9, and one of them (LC B3) showed binding to Cys or Lysine residues of Ubc9 (**Table 3.1 & Figure 3.3**). However, the activity site of Ubc9 Cys 93 is on the surface but not in a hydrophobic pocket. LC B3 did not bind to Cys 93 but to other Cys at Ubc9's backside.

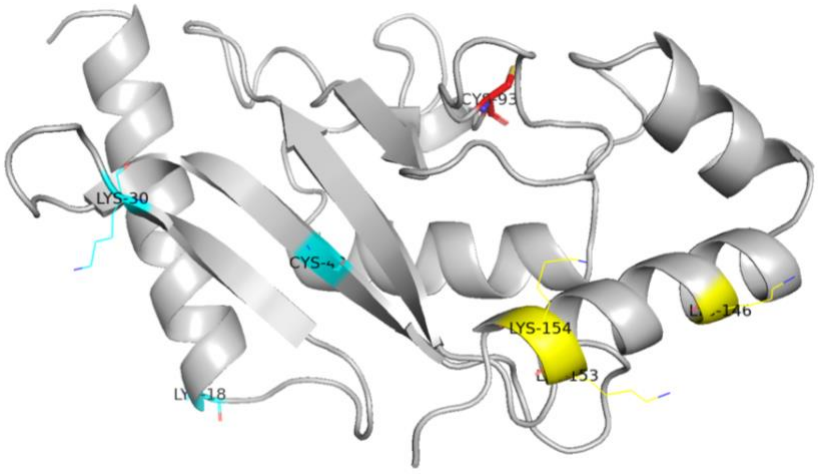
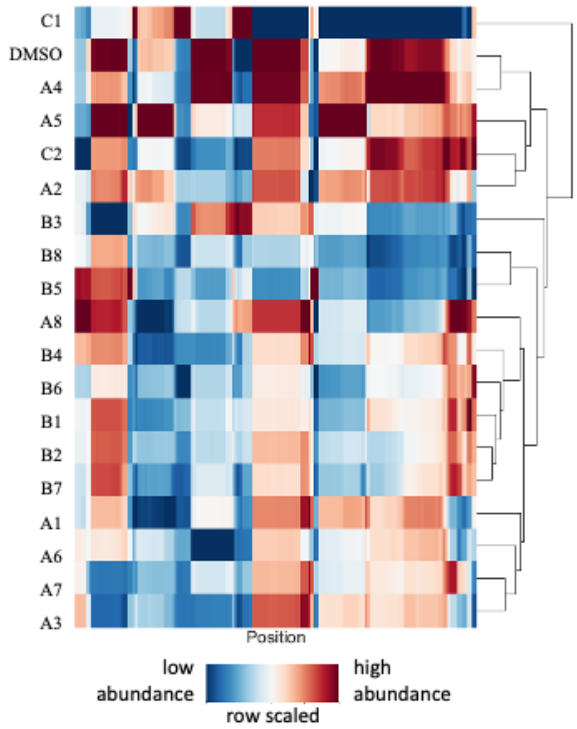


Figure 3.4 Potential binding sites of Ubc9 by covalent fragment B3. Combined the data by mass spectrometry and Ubc9 crystal data (PDB 1U9A), potential binding sites are labelled. K18, K30, and C43 at the N terminal (highlighted by blue) and K146, K153, and K154 at the C-terminal region (highlighted by yellow) are all on the backside of Ubc9, where the activity site C93 (highlighted by red) is located on the surface of another side of Ubc9.



Covalent fragment	Putative Binding Regions ($log_{2ratio} \leq -1$)
B3 (Michael acceptor)	16 – 44 (2*K, 1*C), 144 – 158 (3*K)
B8	15 – 45
B5	16 – 45
C1	14 – 45
A8	31 - 44

Figure 3.3 Mass spectrometry data summary. Top panel: the abundance by Ubc9 amino acid position with covalent fragments or DMSO (negative control). The abundance represented the number of peptides digested with proteinase K. Compared with Ubc9 mixed with DMSO, the low abundance indicated no digestion by covalent fragment binding. Bottom panel: covalent fragments that showed significant low abundance peptides (log_{2ratio} <= -1) with the correlated region.

3.6 Conclusion

The aim for Ubc9 was to identify a covalent fragment that can bind to the active site Cys 93, and thereby inhibit Ubc9. From TSA, data suggested that Ubc9 was highly druggable, as numerous fragments were observed to bind. Furthermore, five LC covalent fragments could form covalent bonds with Ubc9, as shown from mass spec. LC B3 could covalently bind to Lys or Cys, more specifically. However, mass spec did not observe any covalent binding to Cys 93. Cys 93 is located on a flat surface of the 3D structure of Ubc9, which makes it hard to bind to small molecules, which prefer pockets within a surface. LC B3 binding regions were the N or C terminal regions (functioning in DNA repair). At the N-terminal region, the potential binding residue could be K18 by the end of an α-helix, K30 and C43 on a beta sheet. At the C-terminus, the three potential binding sites K146, K153, and K154 are on the same α-helix. They are all located at the backside of Ubc9, where activate site C93 is located on the opposing surface of Ubc9 (**Figure 3.4)**. Besides SUMOylation as PTM of protein, SUMO also functions in DNA repair through Ubc9 [174, 175]. Ubc9 interacts with DNA repair protein Rad60 (NIP45 in humans), and the interaction is required for Rad60's function of replication fork-associated recombination. Therefore, although no covalent fragment could bind to the enzymatic activity site of Ubc9, LC B3 is still a promising hit that can interrupt Ubc9 functions in DNA repair.

FBLD uses a fragment that can bind to the target and inhibit it. Therefore, the target loss function in a pathway. Another way to block the target in the pathway is to degrade the target through a ubiquitin-proteasome system, endosome-lysosome system, or autophagy-lysosome system [176]. The overarching mechanism is to build a

heterobifunctional degradant with ligand binding to the target, a ligase that can trigger polyubiquitination and degradation, and a linker that links the ligand and the ligase. No data has shown that Ubc9 would go through the ubiquitin-proteasome system, but Ubc9 can be targeted by autophagy in human [81]. Therefore, future research on Ubc9 hypothesizes that Autophagy-targeting Chimera (AUTAC) could degrade Ubc9 by triggering K63-linked polyubiquitination and lysosome-mediated degradation (**Figure 3.5**). In the AUTAC system, the ligand is LC F5, which can covalently bind to Ubc9. The tag is an S-guanylation-inspired degradation tag for autophagy. The AUTAC will trigger K63-linked polyubiquitination and lysosome-mediated degradation [177]. One of the challenges is that the molecular weight of the AUTAC might be relatively big. Therefore, tests for drug-like properties during hit-to-lead and lead optimization are needed.

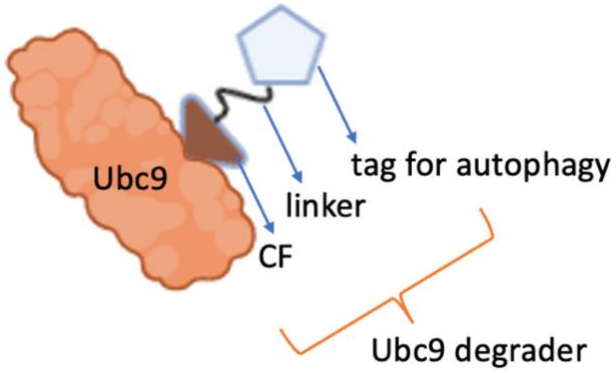


Figure 3.5 Design of AUTAC for Ubc9. Using the concept of AUTAC, a degrader of Ubc9 contains LC B3 (CF) that can covalently bind to Ubc9, a tag for autophagy to trigger K63-linked polyubiquitination and lysosome-mediated degradation, and a linker that links CF and the tag.

3.7 E3 PIAS1 in cancer and immune

SUMO E1 and E2 are sufficient for SUMOylation, but the SUMO E3 ligases enhance SUMOylation substrate recognition and specificity [178, 179]. The largest class of SUMO E3 includes the protein inhibitor of activated signal transducer and activator of transcription (STAT) (PIAS1-4) family and MMS21/NSE2 [180]. The PIAS family regulates DNA damage, cell migration, and checkpoints. For example, Pias1 is required for the SUMOylation of PCNA, which PCNA is involved in DNA replication [181]. Mms21 is required for DNA repair by stimulating SUMOylation of DNA repair protein TRAX [182]. Ran-binding protein 2 (RanBP20) also shows SUMO E3 activity [183]. Poly-comb protein PC2 can recruit Ubc9 and CtBP as SUMO substrates and is identified as SUMO E3 [184]. Lastly, the tripartite motif (TRIM) has eight family members and can recruit SUMO and substates binding via its RING (interesting new gene) and zinc-binding domains [185, 186].

Pias proteins inhibit STAT, as the name implies. There are PIAS1, PIAS2 (Piasx), PIAS3, and PIAS4 (Piasy) genes in the human genome [83]. Structurally, Pias family proteins contain N-terminal scaffold attachment factor-A/B acinus and PIAS (SAP) domain, PINIT motif, SP-RING-type zinc binding structure, SUMO interacting motif (SIM), and Siz-PIAS C-terminus domain (SP-CTD). The Pias family is also called the SP-RING family due to its structure. The SP-RING domain interacts with Ubc9. The SP-CTD domain interacts with SUMO. The PINIT motif is required for PIAS-depended substrate (i.e., PCNA), then transfers SUMO to its substrates [187]. There is no 3D structure information on full-length Pias1, while other members of the Pias family have been studied

more and have structural information. Using RanBP2, another SUMO E3 ligase, and crystallography, the complex of SUMO1, Ubc9, RanBP2, and RanGap1 show that SUMO E3 potentially has three functions: SUMO E3 assists Ubc9 to conjugate SUMO, SUMO E3 binds to SUMO2-Ubc9 covalent complex for the substrate binds to SUMO, and SUMO E3 assists a stable form of SUMO and the substrates [84].

3.7.1 Pias1 in cancer

Although Pias1 does not have as much information as other Pias proteins, data has shown that Pias1 is involved in cancer such as prostate cancer [188, 189] and myeloma [190, 191].

In prostate cancer, androgens are required for tumor growth. Drugs that can target androgen receptor (AR) are the treatments for prostate cancer, for example, Enzalutamide (MDV3100), Abiraterone, and Docetaxel [188, 190, 191, 192]. However, prostate cancer patients can show resistance to all the above drugs. qRT-PCR and western blot data shows that androgens upregulate Pias1 mRNA and protein levels in prostate cancer patients [188]. The overexpressed Pias1 then binds and stabilizes AR and enhances AR activity. Knockdown Pias1 by siRNA in prostate cells shows the better output of Enzalutamide treatment with decreased cell proliferation. The above data shows that Pias1 is a potential prostate cancer target for therapy.

Myeloma or multiple myeloma (MM) is the cancer in which plasma cells grow too fast and take the room from other red blood cells or platelets in the bone marrow [193]. Age is the top risk factor, and exposure to X-rays or being overweight are linked to MM. The overall 5-year survival rate is 58% with treatment such as chemotherapy. Pias1 is upregulated in some MM cases, as shown from recent immunoblot data [194]. Patients

with high levels of Pias1 and Ubc9 have lower survival rates than those with lower levels of Pias1. Together, Pias1 and Ubc9 are identified as therapeutic targets for MM.

Pias1 SUMOylates at Cys42 of MYC. SUMOylation of MYC promotes phosphorylation at Ser62 by ERK, which stabilizes MYC, and dephosphorylation at Thr58 by GSK3b, which recruit ubiquitin-proteasome degradation of MYC [190, 195, 196]. Pias1 is required for MYC in the primary B-cell lymphomas and the cell's viability. Uba2 is the synthetic lethality partner of MYC. Inhibition of SUMOylation promotes MYC lymphoma apoptosis [197]. Therefore, SUMOylation modification is highly required for MYC-driven cancer.

Pias1 also plays a role in cancer metastasis. Metastasis requires loss of cell polarity and cell adhesion, which focal adhesion kinase (FAK) promotes metastasis and is overexpressed in advanced-stage cancers [198]. Pias1 SUMOylates at Lys152 of FAK and then promotes FAK activation via autophosphorylation at Thr397 [199]. In NSCLC, co-IP determines that Pias1 and FAK physically interact with each other and promote cancer progression [200].

3.7.2 Immune functions

Pias1 is linked to cancer initiation and being a p53 repressor [201] and contributes to immunological functions by negatively regulating NF-κB/STAT signaling [202]. Notably, Pias1 knockout mice resist viral and bacterial infection, because depletion of Pias1 enhances interferon (IFN)-mediated responses [203]. Pias1 interacts and inhibits STAT1 by blocking the DNA-binding activity of STAT1. STAT1 then activates the IFN sequence and responsive elements with infection stimulation. Without Pias1, IFN-induced gene

activation is increased, enhancing immune response to viral or microbial challenges. The SUMOylation pathway modulates innate (innate immunity: the natural immunity) and intrinsic immunity (intrinsic immunity: a form of innate immunity directly restricting viral replication) by altering IFN and JAK/STAT via SUMO2/3 to prevent IFN production. Loss of SUMOylation triggers a potent and spontaneous IFN with or without stimulation. Hence, SUMOylation is also a novel target for anti-virus therapy, including potentially coronavirus [204].

Pias1 has yet to be characterized structurally, and in this chapter, the aim is to study the structure of the SUMO E3 Pias1 protein and to identify chemical tools targeting the protein aiming for targeting SUMOylation of MYC-driven cancers, using SBDD coupled with FBLD.

3.8 Methods

3.8.1 Expression

Human Pias1 (1–663) was cloned into vector pET21a (+) with an N-terminal 6x His tag (GenScript Biotech). Pias1 was expressed by transforming plasmids in different *E. coli* cell lines. Then, a small-scale expression test was done to determine the optimal cell line and expression condition and saved glycerol stock. The glycerol stock was cultured in 50 mL of TB with ampicillin at 37℃ overnight (O/N). The O/N culture was inoculated into 12L of TB with ampicillin at 37℃ at 225 rpm (New Brunswick Scientific Innova©43R). When the OD600 was 0.6, the temperature was lowered to 16℃, and then the cells were induced with 0.4 mmol/L IPTG at OD600 1.2 and kept shaking O/N 16℃. Cells were harvested by centrifugation (Sorvall LYNX 4000 Superspeed Centrifuge with a FiberliteTM F9-6 x

1000 LEX Fixed Angle Rotor) at 4℃ and 4,000 rpm for 15 minutes, resuspended at a ratio of 3ml lysis buffer (50 mM Tris-HCl pH 8.5, 300 mM NaCl, 25 mM imidazole, 10 mM BME) per gram pellet. Per 40 mL of the resuspended cell, it was sonicated (Q-Sonica Q125, ¼ inch probe) on the ice at 70% amplitude for 4 minutes, with 10 seconds on and 10 seconds off. The supernatant was obtained by centrifugation at 4℃ and 30,000 rpm for 40 minutes. The cell lysate was purified right after or stored at -80℃.

3.8.2 Purification

Supernatant from cell lysate in 50 mM Tris-HCl pH 8.5, 300 mM NaCl, 25 mM imidazole, 10 mM BME with benzonase, and ETDA-free protease inhibitor cocktail was purified by Ni-affinity chromatography (HisTrap FF; Cytiva) using NGC Quest 10 Plus Liquid Chromatography System (Bio-Rad) and eluted by a gradient of 1M imidazole buffer, exchanged to a low salt buffer and then loaded onto anion exchange chromatography (HiTrap Q FF; Cytiva) and eluted by a gradient of 1 M NaCl buffer, gel filtration chromatography (Superdex 200; Cytiva) and eluted by a low salt buffer, and finally Enrich high-resolution anion exchange chromatography (ENrich Q; Bio-Rad) and eluted by a gradient of 1 M NaCl buffer. 12% SDS-PAGE gels confirmed fractions from each step, and then desired fractions were collected for the next step. The final sample was stored in 50 mM HEPES pH 8.5, 300 mM NaCl at -80℃ after flash froze in liquid nitrogen.

3.8.3 SUMOylation activity assay

Following the methods in 2.5.2.5, purified Pias1 and purified Pias1-specific SUMO substrate PCNA from other members in the Perry Lab were added with SUMO1, SUMO E1, SUMO E2, and ATP to verify the ligase activity of Pias1. Purified SUMO1, E1, E2,

E3, PCNA, and ATP were added into one 1.5 mL microcentrifuge tube. Including every component but not Pias1 was the negative control in another tube. Tubes were incubated at 37℃ for 2 hours. The reaction was stopped by adding SDS-PAGE loading dye and boiling for five minutes. Then, samples were loaded onto a 12% SDS-PAGE gel and a western blot using an anti-SUMO1 antibody.

3.8.4 TSA

Following the TSA method from 2.5.2.1, hits were determined using a final of 1.6 μM of Pias1 with the reaction buffer (50 mM HEPES pH 8.5, 50 mM NaCl). The first round of screening was done by singlet, and concentration/dosage tests were done in triplicate. DMSO served as the negative control.

3.9 Results

3.9.1 Expression

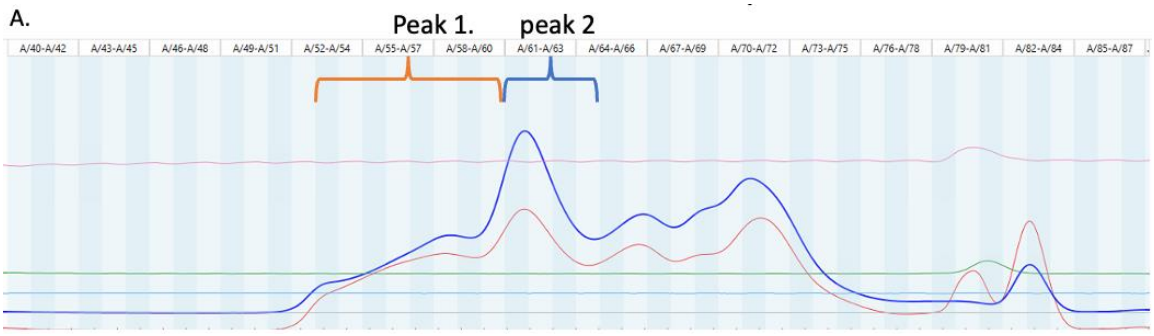
In the small-scale expression test, Pias1 plasmids were transformed in *E. coli* BL21(DE3), ArcticExpress RIL, ArcticExpress RP, ArcticExpress (DE3), or BL21-CondonPlus-RP, and cultured in TB or 2XYT. Cells were induced at OD600 1.2 and set to shake at either 37℃ 4 hours or 16 ℃ O/N. A total of 30 tests were collected and then loaded onto nickel spin columns, eluted with high imidazole buffer, and run on 8-16% SDS-PAGE gels for comparison. Protein was expressed only using the BL21-CondonPlus-RP strain, and the best expression condition was induced at 16℃ at OD600 1.2 and incubate at 16℃ O/N using TB (**Supplementary material figure 3.4**). Glycerol stock was saved for large-scale expression and purification.

3.9.2 Purification

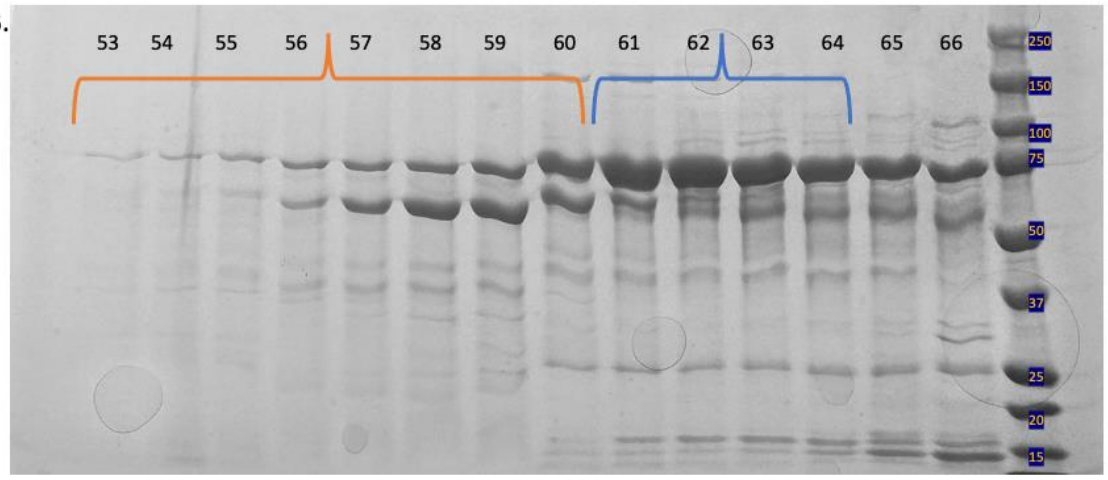
Using FPLC, there were mainly two peaks in post-size exclusion chromatography (SEC). As shown in the chromatography figure, peak #1 was eluted first, and the shape was not an ideal symmetric peak (**Figure 3.6 A**). On the 8-16% SDS-PAGE gel, the correlated fractions (53-59) showed two major bands, a smaller upper band, and a bigger lower band (**Figure 3.6 B**). As the theoretical size of Pias1 is 74 kDa, the upper band was most likely Pias1, and the lower band was thinner than Pias1. Using the semi-log SEC calibration line (**Figure 3.6 C**), the size of the peak 1 sample was 238.98 kDa (**Figure 3.6 D**). Peak #2 had a better shape on the chromatograph figure. The fractions (60-64) also showed two major bands, but the upper bands were more prominent. The peak #2 sample size was 71.91 kDa, calculated by the SEC calibration line. Peak #2 fractions were collected and followed by an Enrich Q column. There were mainly three peaks post-Enrich Q (**Figure 3.7** A). The first peak was the biggest and showed higher A280 (blue) than A260 (red) wavelength—the correlated fractions 36 and 37 on SDS-PAGE gel with a band of 74 kDa (**Figure 3.7** B). The minimal peak #2 showed a very faint band on the gel. The last peak had a faint band on the gel and a higher A260 than A280. The first peak fractions were collected and used as Pias1 in future tests. The yield was 2mg of material per 12 L of *E. coli* culture.

Proliferating cell nuclear antigen (PCNA) is an E3 Pias1 specific SUMOylation substrate. As shown in **Figure 3.8**, from the left side, lane #1 was the negative control without Pias1, and line #2 was the testing group with Pias1. A band of 15 kDa showed in lane #1, indicating SUMO1, and a band of 50 kDa in lane #2, indicating SUMOylated PCNA.

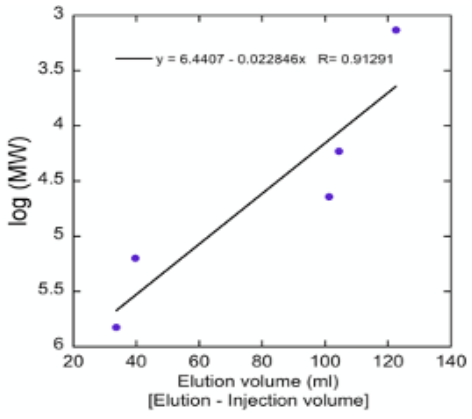
A.



B.



C.



D.

Elution volume (Superdex-200)	elution volume (mL)	mw (kDa)
peak 1	46.5	238.98
peak 2	69.33	71.91

Figure 3.6 Pias1 FPLC SEC. Pias1 after HisTrap FF and HiTrap QFF, was purified by Superdex 200. A. Chromatography of Pias1 SEC contains peak 1 (fraction 53-60) and peak 2 (fraction 61-64). B. Fractions from peak1 and peak2 were loaded on an 8-16% SDS-PAGE gel. C. Semi log figure of SEC elution volume vs log of molecular weight. D. The molecular weight of samples from A was calculated using semi log figure in C. In fraction 53-60, sample size was 239 kDa, and 71.9 kDa in fraction 61-64. Buffers were used: Superdex 50 mM HEPES pH 8, 300 mM NaCl, 1 mM BME.

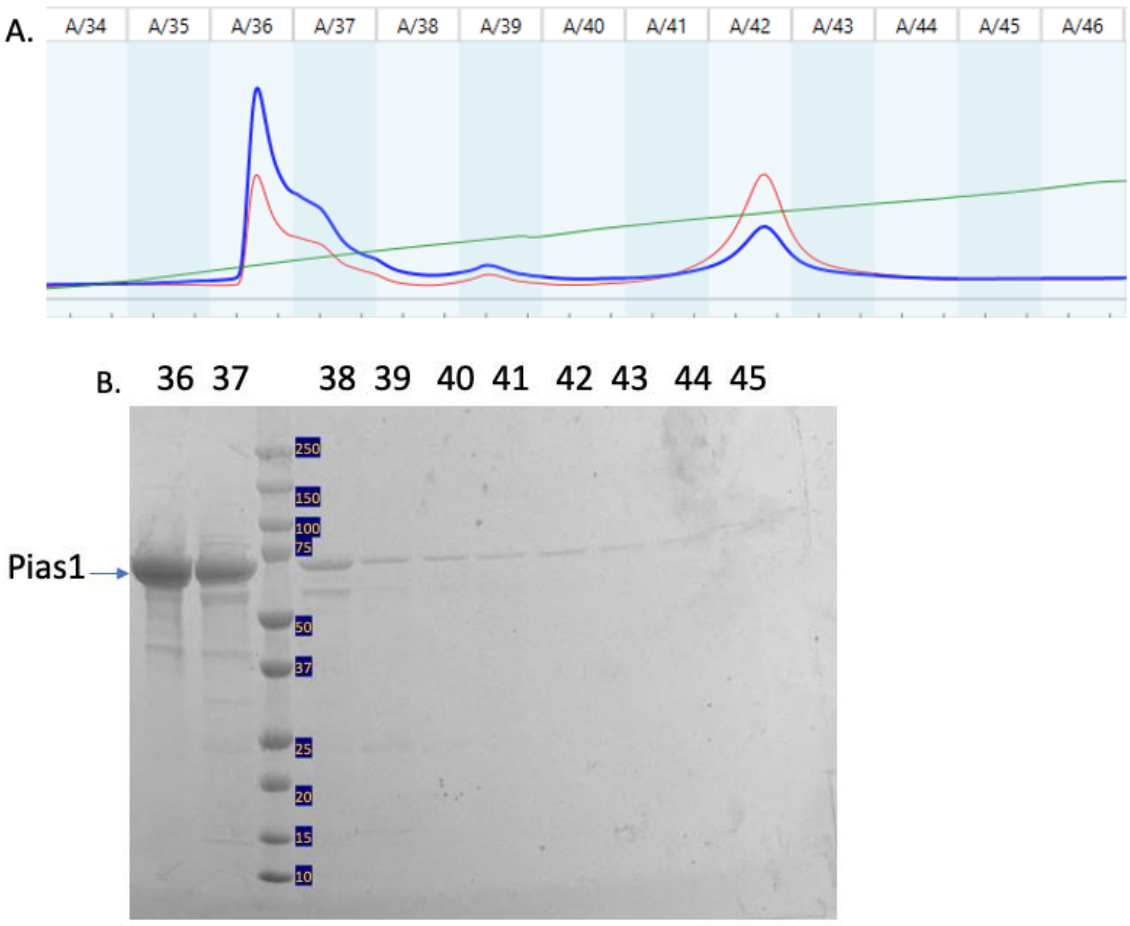


Figure 3.7 Pias1 FPLC IEX. Pias1 after Superdex 200 was purified by Enrich Q and eluted in fraction 36-37. A. The chromatograph figure with threes peaks. Blue line represented the A280 wavelength, indicated the protein signal. Red line represented A260 wavelength by DNA and RNA. Peak at fraction 42 contained more DNA and RNA than protein. B. Fractions were loaded on an 8-16% SDS-PAGE gel. 36 and 37 contained the most Pias1 protein. The other fractions also contained protein but very diluted. Buffers were used: Enrich Q A 50 mM HEPES pH 7.5, 50 mM NaCl, 1 mM BME & Enrich Q B 50 mM HEPES pH 7.5, 1 M NaCl, 1 mM BME. Yield: 0.4 mg of Pias1/ L of *E. coli* culture.

3.9.3 SUMOylation activity assay

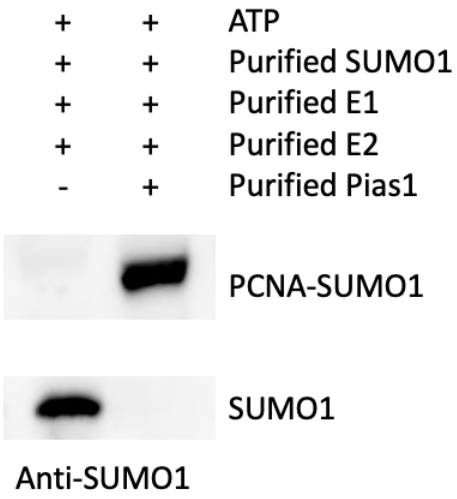


Figure 3.8 Purified SUMO E3 Pias1 was tested by SUMOylation reaction kit and run by western blot. Purified SUMO E1, E2, E3 Pias1, and Pias1 specific substrate PCNA were added with or without ATP for SUMOylation modification reaction on PCNA. Reactions were incubated at 37℃ for 2-hour and loaded on an 12% SDS-PAGE gel. Gel was transferred to a PVDF membrane and incubated with anti-SUMO1 antibody at 4℃ overnight. Then membrane was incubated with anti-rabbit secondary antibody at RT for 2-hour and then performed image. Lane #1: purified E1, E2, SUMO1, PCNA with ATP. Lane #1: purified E1, E2, E3 Pias1, SUMO1, PCNA with ATP resulting in SUMOylated PCNA.

3.9.4 TSA

As shown in **Supplementary material figure 3.5**, Pias1 always showed two peaks. One is approximately 42.5℃, and the second is 56.5℃. Pias1 was druggable shown in the screening test (**Figure 3.9**). When testing with different dosage of LC 1A, both T_m shifted positively and showed a dosage trend (**Figure 3.10**).

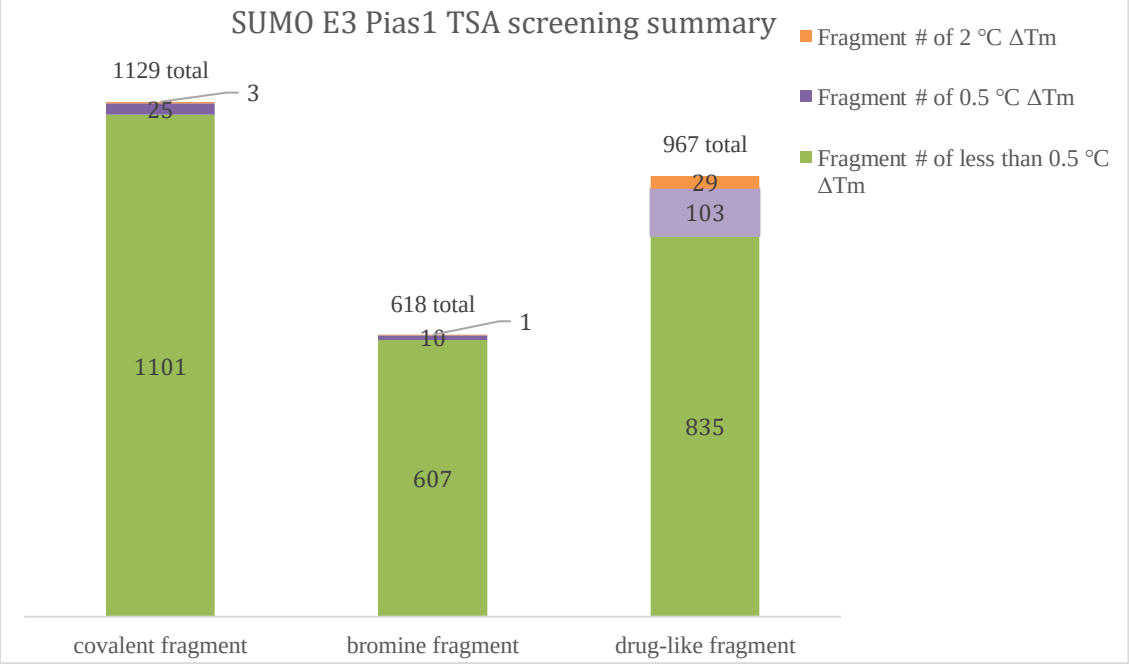


Figure 3.9 SUMO E3 Pias1 TSA screening with drug-like, covalent, and brominated fragment. 2714 fragments (1129 covalent, 618 brominated, and 967 drug-like) were tested against Ubc9 by TSA. The % of fragments with at least 0.5°C ΔT_m were 2.48, 1.78, and 13.62, respectively.

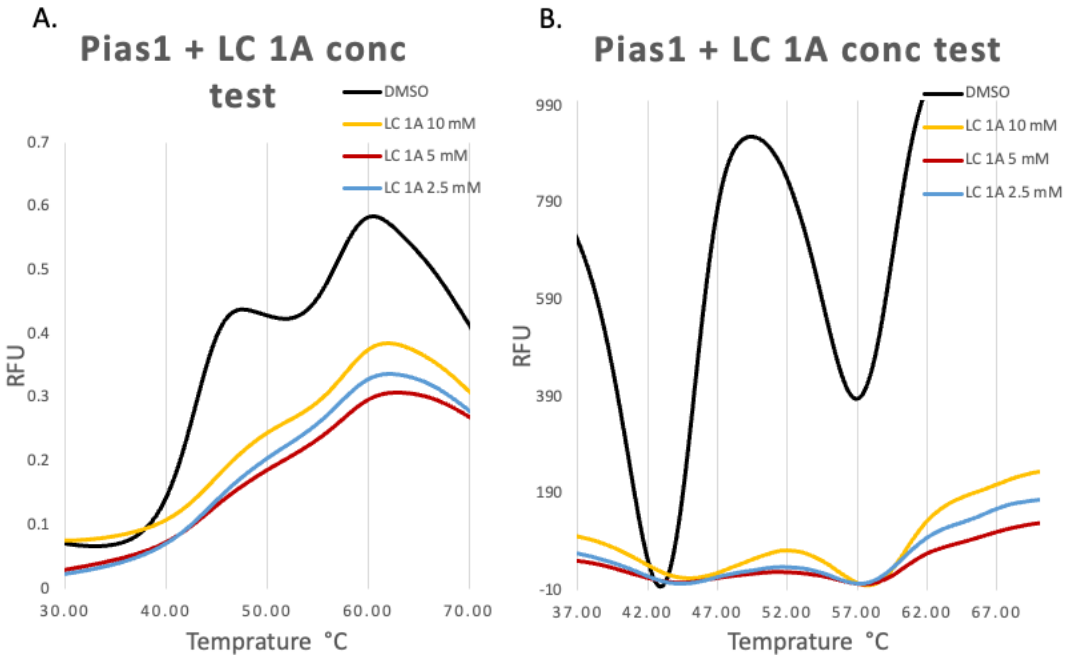


Figure 3.10 Pias1 with LC 1A dosage/concentration test. LC 1A was tested with different concentration against to Pias1. ΔT_m of peak 1 were 1.33°C, 1.33 °C, and 1.00°C from 10 mM, 5 mM, and 2.5 mM of LC 1A, respectively. ΔT_m of peak 2 were 0.5°C, 0.5°C, and 0°C from 10 mM, 5 mM, and 2.5 mM of LC 1A, respectively. A. Normalized figure of RFU vs. degree °C. B. Normalized figure of -d(RFU)/dT vs. degree °C.

3.10 Conclusion

The first challenge of these Pias1 research studies was to purify the protein, as it has not been expressed and purified from *E. coli* before, as far as we were aware. Using the plasmid from the Liao Lab and a construct from GenScript, Pias1 was purified for the first time. The molecular weight of Pias1 is 74 kDa. Expression optimizations were done to increase the expression level and final product yield. However, the yield still needed improvement, which remained an issue.

The post-SEC Pias1 sample always had one prominent peak (peak #2) and few peaks. One theory was that peak #1 was trimer and peak #2 was monomer, based on the semi-log calibration calculation. The fractions from peak #1 and 2 had one band of 75 kDa and another slightly smaller band on the SDS-PAGE gel. However, fractions from the later peak (i.e., the peak after peak #2) had a similar band on the gel. Data suggested Pias1 could have multimeric form and monomer form, and it might have a truncated form, possibly caused by degradation. However, a protease inhibitor cocktail was added to the lysate buffer. As a future step, this could be confirmed by mass spec for degradation or more details studies, such as Cyro-EM for studying the multimeric state.

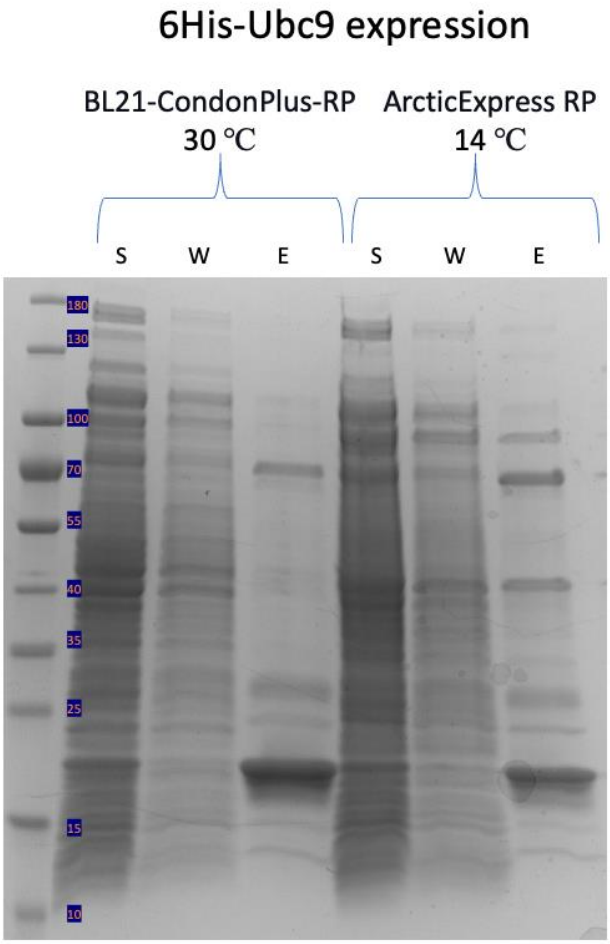
Using the drug-like, covalent, and brominated fragments, data showed that Pias1 was druggable. Intriguing, Pias1 showed two peaks / T_ms during the TSA assay, although the monomer sample of peak #2 from SEC was used. Using reaction buffers with different pH could reduce or deplete the second T_m. pH 4.5, 5.5, and 6.5 had the first peak / T_m, pH 6 had the first peak / T_m and vastly reduced second peak / T_m, and the rest of the pH had two peaks / T_m. The reaction buffer had pH 8.5, as Pias1 in this buffer had relatively low T_m

(both T_ms). The change of T_m suggested that protonation of His residue(s) (pK_a 6.0) mediated multimerization and was pH dependent. The structural study of Pias1 in the future will confirm this.

Focused on the covalent fragments, LC 1A had shown binding during the screening test and binding trend in the dosage/concentration test. Therefore, the next step is to study LC 1A as a potential hit of Pias1.

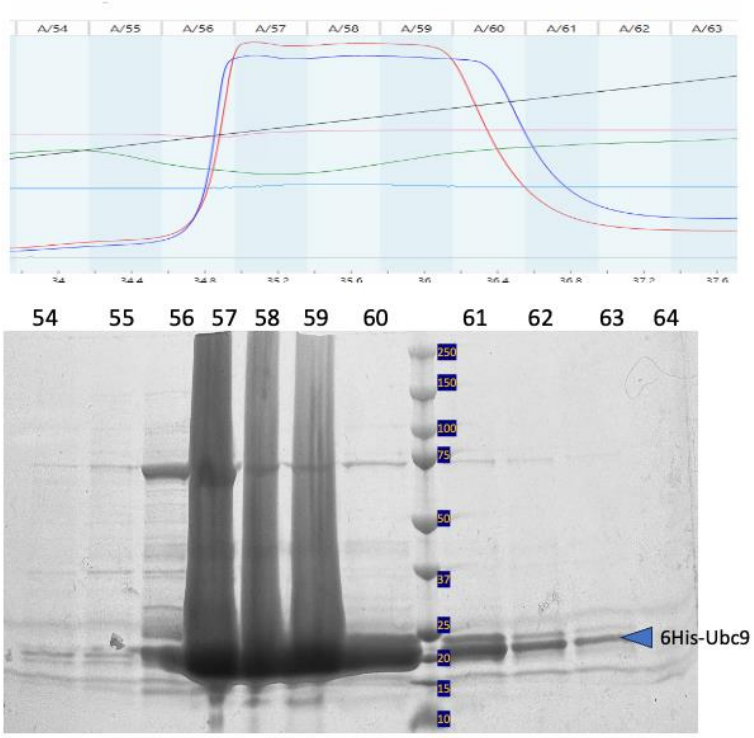
Due to the low yield of Pias1 and the possibility of multimeric and monomer co-exist, crystallography work of Pias1 might be a challenge. However, due to the relatively large molecular weight of Pias1, Cyro-EM might be a better strategy for structural research of Pias1. For future work, a 3D structure of Pias1 would be very useful. Data of Pias1 and LC 1A interaction would conform LC 1A as a hit for Pias1, to allow it to move to the next stages of early-stage drug discovery. At the same time, the IC_{50} test could be done by the HTRF assay, with some modifications. The HTRF assay tested the SUMOylation level of RanGap1 via Uba2/Aos1 and Ubc9, which RanGap1 is an E3 independent SUMOylation substrate. To test Pias1, the HTRF assay setup should be GST-SUMO2, Uba2/Aos1, Ubc9, Pias1, and His-PCNA, as PCNA is a Pias1 specific SUMOylation substrate.

3.11 Supplementary material

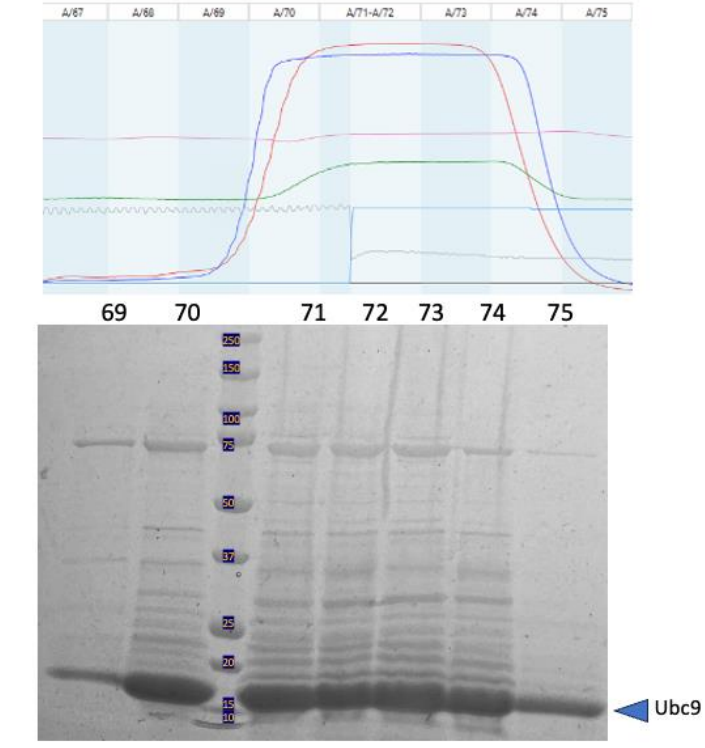


Supplementary material figure 3.1 Ubc9 expression small-scale test. 6His-TEV-Ubc9 expressed in *E. coli* ArcticExpress RP with LB and induced by IPTG @ OD 0.8 at 14℃ O/N or BL21-CondonPlus-RP at 30℃ 3-hour. Cells were harvested and examined by NI-NTA spin column. Samples were loaded on an 8-16% SDS-PAGE gel. Lane 1: protein ladder. Lane 2 & 5: cell supernatant. Lane 2 & 6: wash from NI-NTA spin column. Lane 3 & 7: elution from NI-NTA spin column. Lane 2 – 5 were sample of BL21-CondonPlus-RP, and lane 6 – 8 were sample of ArcticExpress RP.

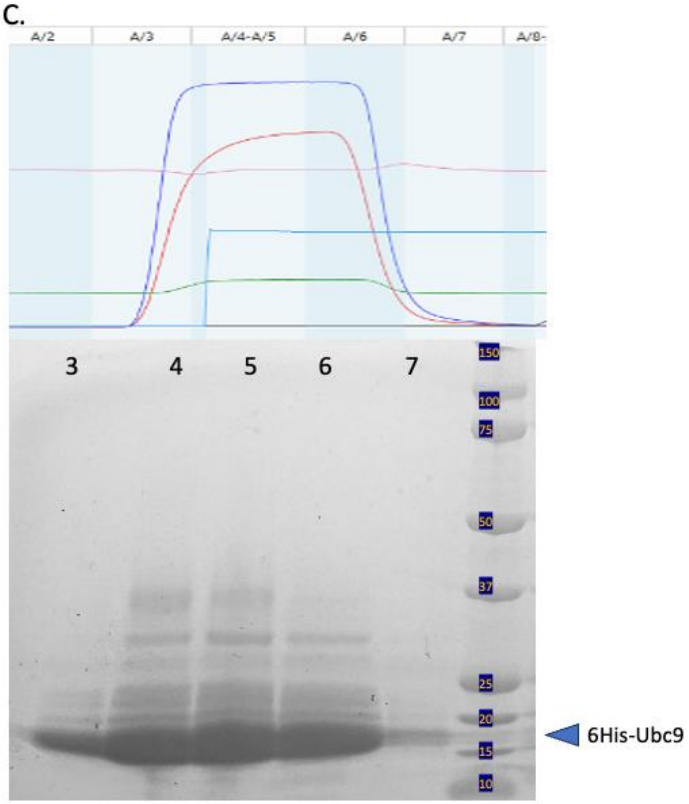
A.


A/54
A/55
A/56
A/57
A/58
A/59
A/60
A/61
A/62
A/63
54
55
56
57
58
59
60
61
62
63
64
250
150
100
75
50
37
25
20
15
10
6His-Ubc9

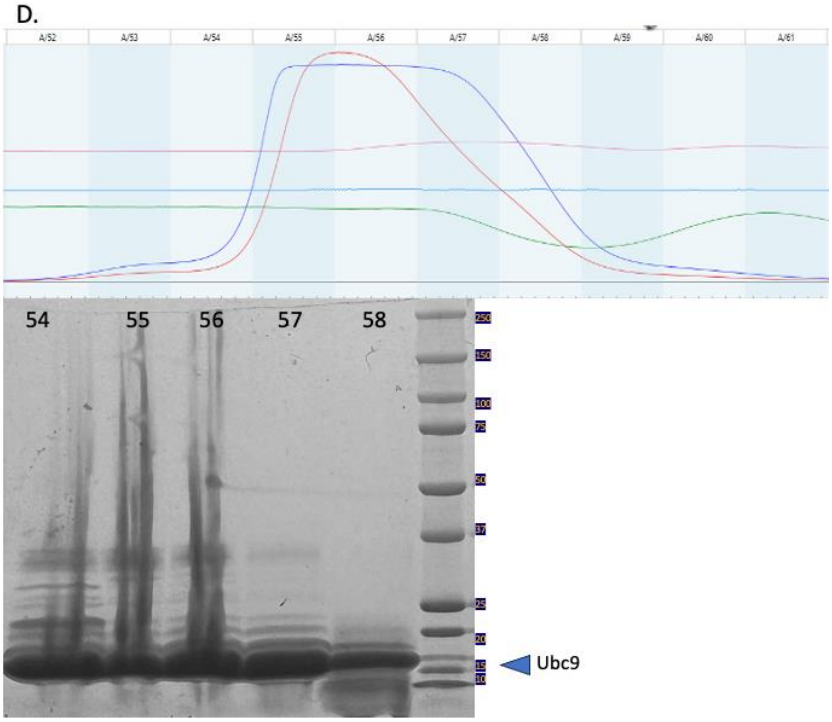
B.


A/67
A/68
A/69
A/70
A/71-A/72
A/73
A/74
A/75
69
70
71
72
73
74
75
250
150
100
75
50
37
25
20
15
10
Ubc9

C.

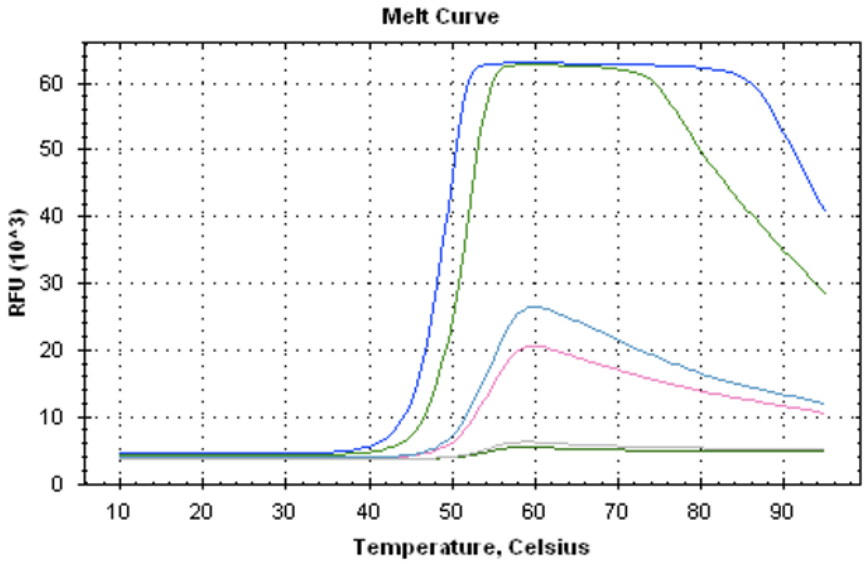

A/2
A/3
A/4-A/5
A/6
A/7
3
4
5
6
7
150
100
75
50
37
25
20
15
10
6His-Ubc9

D.

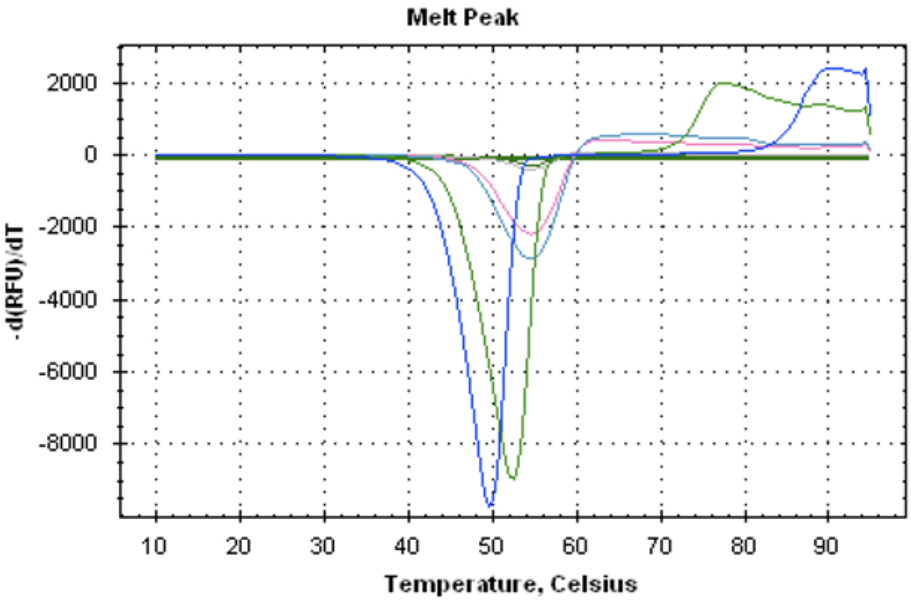

A/52
A/53
A/54
A/55
A/56
A/57
A/58
A/59
A/60
A/61
54
55
56
57
58
Ubc9

Supplementary material figure 3.2 Ubc9 purification by FPLC. 6His-TEV-Ubc9 in *E. coli BL*21-CondonPlus-RP with LB and induced by IPTG @ OD 0.8 at 30 ℃ 3-hour. A. HisTrap FF captured Ubc9 in elution fractions 56-63. B. TEV treatment O/N and then purified by HisTrap FF in fractions 69-75. C. HiTrap QFF in fraction 3-7 and D. Superdex 75 in fraction 57-58. Buffers were used: His A 50 mM HEPES ph8, 300 mM NaCl, 25 mM imidazole, 1 mM BME & His B 50 mM HEPES ph8, 300 mM NaCl, 1 M imidazole, 1 mM BME. HiTrap Q A 50 mM HEPES ph8, 50 mM NaCl, 1 mM BME & HiTrap Q B 50 mM HEPES ph8, 1 M NaCl, 1 mM BME. Superdex 50 mM HEPES ph8, 300 mM NaCl, 1 mM BME. Yield: 2.5 mg of Ubc9/ L of *E. coli* culture.

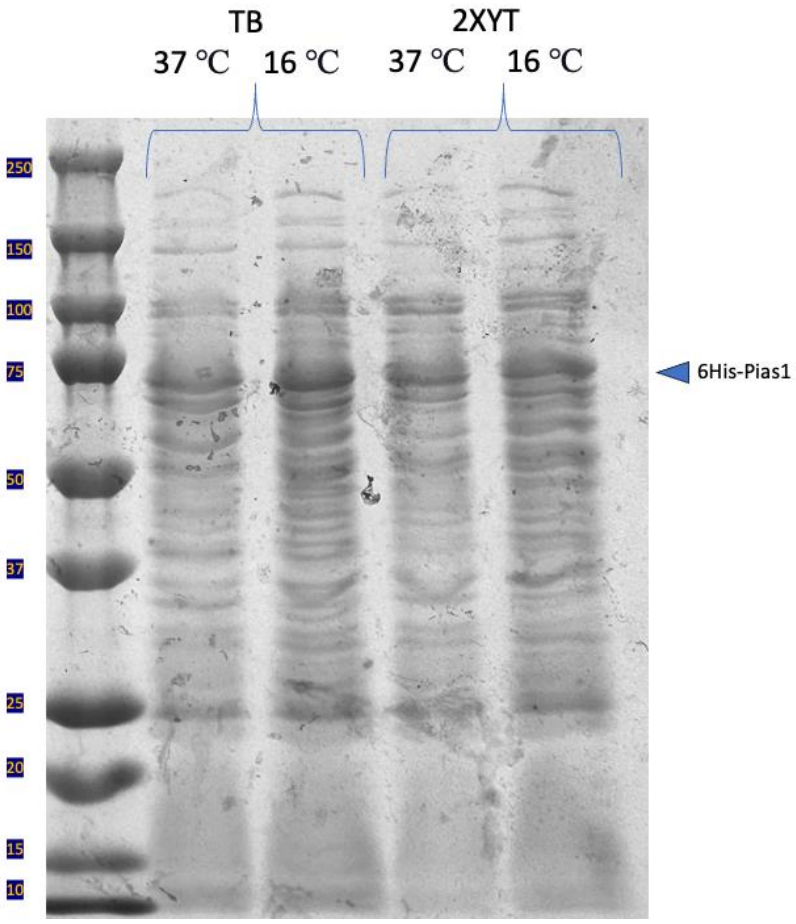
A.



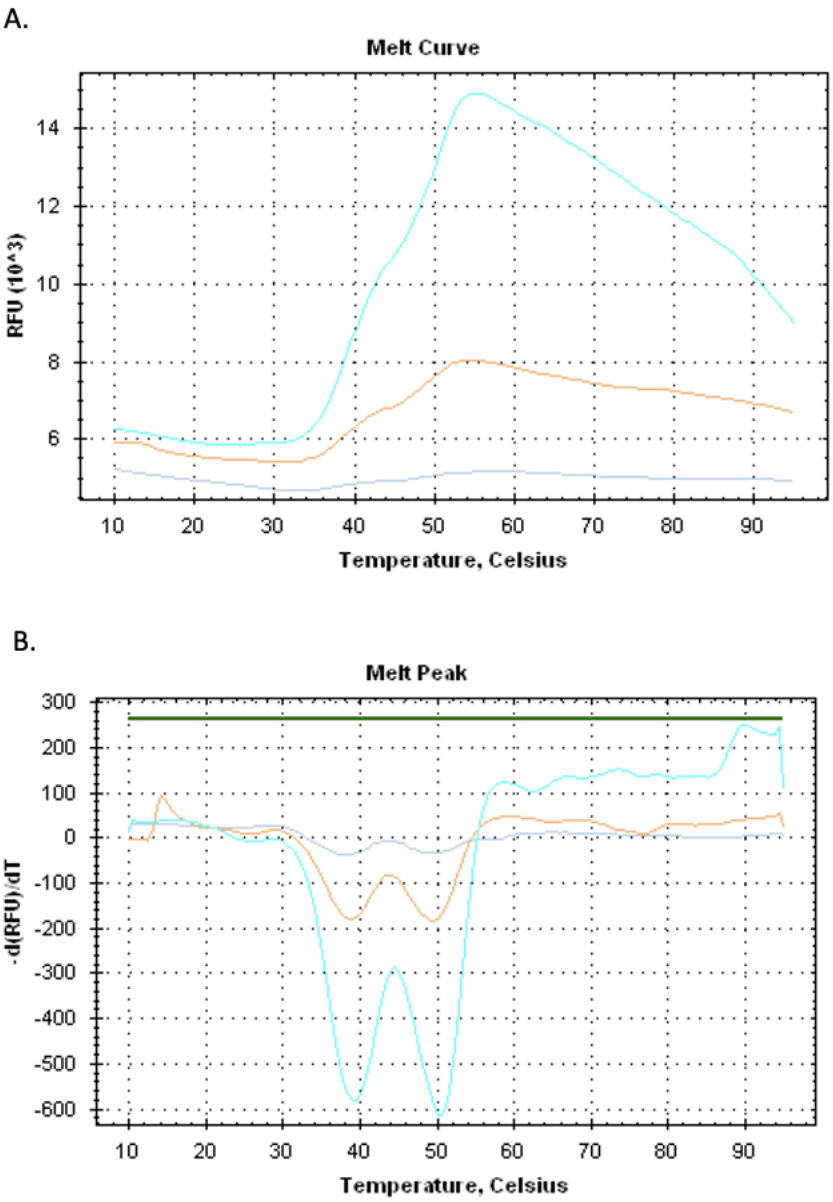
B.



Supplementary material figure 3.3 Ubc9 concentration test for TSA. Ubc9 serial dilution from 32 μM to 1 μM in PBS were tested for the concentration test. A. The figure of RFU vs. degree °C. B. The figure of -d(RFU)/dT vs. degree °C. 32 and 16 μM of Ubc9 (represented by line in blue and green, respectively) saturated the threshold of the rt-PCR.



Supplementary material figure 3.4 Pias1 expression small-scale test. 6His-Pias1 expressed in *E. coli* BL21-CondonPlus-RP with TB or 2XYT and induced by IPTG @ OD 1.2 at 16 ℃ O/N or at 37 ℃ 3-hour. Cells were harvested and examined by NI-NTA spin column. Elution Samples were loaded on an 8-16% SDS-PAGE gel following the order of TB 37 ℃, TB 16 ℃, 2XYT 37 ℃, and 2XYT 16 ℃.



Supplementary material figure 3.5 Pias1 concentration test for TSA. Pias1 serial dilution from 1.6 μM to 0.4 μM in PBS were tested for the concentration test. A. Figure of RFU vs. degree °C. Only 1.6 μM showed the RFU above 10*10^3. B. -d(RFU)/dT vs. degree °C. Two peaks showed in all concentration samples.

Chapter 4
The Down syndrome related protein Dyrk1a in early drug discovery

Chapter 4: The Down syndrome related protein Dyrk1a in early drug discovery

4.1 Abstract

The dual-specificity tyrosine phosphorylation-regulated kinase 1A (Dyrk1a) plays multiple roles in animal development, with critical effects on neuronal cell cycle regulation and cancer development. Dyrk1a has dual specificity because it has autophosphorylation activity on a tyrosine residue (cis and trans) and conducts phosphorylation of a downstream substrate on a serine or threonine residue. Transgenic mice overexpressing Dyrk1a show neurodevelopmental delays like those observed in humans, indicating the critical role of Dyrk1a in Down syndrome (DS) neurodevelopment. Alzheimer's disease (AD) is associated with alterations in Dyrk1a expression as well. Dyrk1a has been linked to type 2 diabetes (T2D), and inhibition of Dyrk1a by small molecule inhibitors shows promised results of treatment. Acute lymphoblastic leukemia (ALL) is the most common cancer for children, especially children with DS. Dyrk1a is significantly overexpressed in leukemia cells from pediatric ALL patients by gene expression profiling and qPCR. Therefore, Dyrk1a is a novel therapeutic target for DS, AD, T2D, and leukemia. There are some inhibitors against Dyrk1a. However, most of them lack selectivity, since Dyrk1a shares at least 85% of the ATP binding site with another member from the DYRK family, Dykr1b. The research efforts described in this chapter have been focused on hit identification and initial hits-to-leads studies that target Dyrk1a. Our future efforts will focus on improving hit small molecules selectivity to Dyrk1a.

4.2 Dyrk1a functions in cellular and diseases

4.2.1 Dyrk1a gene and protein function

The dual-specificity tyrosine phosphorylation-regulated kinase 1A (Dyrk1a) plays multiple roles in animal development, with critical effects on neuronal cell cycle regulation and cancer development [205]. Dyrk1a belongs to the DYRK subfamily, with five isoforms (Dyrk1a, Dyrk1b, Dyrk2, Dyrk3, Dyrk4), in the CMGC kinase family. The CMGC kinase family contains cyclin-dependent kinase (CDK), mitogen-activated protein kinase (MAPK), glycogen synthase kinase (GSK), and CDC-like kinase (CLK), and the CMGC kinases are highly conserved among organisms [206]. Dyrk1a is highly conserved among vertebrates and highly expressed in the brain of vertebrate embryos [207, 208], especially during early development [209, 210]

Dyrk1a contains a DCAF7 binding domain and a nuclear localization signal (NLS) at the N-terminus, following a DYRK homology (DH) and a catalytic domain (**Figure 4.6)**. Unlike other isoforms in the DYRK family, Dyrk1a uniquely contains a second NLS in the catalytic domain. Dyrk1a also contains a PEST (rich in proline, glutamic acid, serine, threonine) region, His (rich in histidine) region, and an S/T (rich in serine and threonine) region in the C-terminus. Like the second NLS, the His and S/T regions are exclusively for Dyrk1a. In the N-terminus, the DH forms five antiparallel β sheets and functions to maintain the structure of the N-terminus and stabilize the folding [211]. The DH is required for the catalytic activity of Dyrk1a. The activity site Tyr 321 and Phe 238, as the gatekeeper in the ATP binding pocket, are in the catalytic domain. Autophosphorylation of Tyr 321 active Dyrk1a and stabilizes the catalytic domain loop by forming a salt bridge with Arg

325 and Arg 328 in the catalytic domain for substrate binding [212]. The Phe 238 in the hinge region is the residue for inhibition [213].

Dyrk1a has dual specificity because it has autophosphorylation activity on a tyrosine residue (cis and trans) and conducts phosphorylation of a downstream substrate on a serine or threonine residue [127, 215]. A cis-autophosphorylation at the Y321 residue (the second tyrosine in the conserved YXY motif) in the catalytic domain activates Dyrk1a during its translation while bound to the ribosome [2289]. This autophosphorylation is intrinsic in the catalytic domain. As Dyrk1a is translated, the N-terminal is available for folding and causes a folded intermediate. This intermediate form contains a different kinase structure than the mature form. Once Dyrk1a is activated, it can undergo trans-autophosphorylation at residues Y140, Y159, Y177, S310, Y319, and Y499 [211]. Dyrk1a autophosphorylates at residue Ser 520 binding to the protein 13-3-3β [212]. The complex with 13-3-3β stimulates the activity of Dyrk1a. Dyrk1a phosphorylates on their serine or threonine side chains. For example, Dyrk1a phosphorylates tau, the hallmark of Alzheimer's disease (AD), on residue threonine 212 *in vitro* [213] and serine 202 and 404 *in vitro* [214]. In DYRK1A transgenic mice, tau's phosphorylation level of T212 significantly increased compared with the negative control. data suggests that overexpression of Dyrk1a contributes to AD.

As a kinase, Dyrk1a has been shown to impact various downstream proteins. For example, Dyrk1a phosphorylates the antiproliferative cyclin-dependent kinase (CDK) inhibitor ($p27^{kip1}$) and upregulates its expression [288]. Overexpression of $p27^{kip1}$ has been shown that can induce apoptosis in HeLa cell line and human lung fibroblasts IMR90 [216],

but also rescue human lung adenocarcinoma derived A549 cells from apoptosis [217]. Dyrk1a also phosphorylates Lin52 to promote the assembly of the DREAM complex (DP, RB-like, E2F4, and Muvb) that represses gene expression of cell cycle-promoting targets during the G0 stage [218]. It is a promising candidate target for therapeutic approaches, where overexpression of Dyrk1a causes various diseases including Down syndrome, AD, diabetes, and cancer, and reduction of the abnormal expression level has been suggested as a treatment option. As Dyrk1a regulate multiple downstream proteins and they function distinctly in cancer, Dyrk1a also play different roles in different type of cancer.

4.2.2 Dyrk1a roles in Down syndrome (DS)

DS is the most common genetic defect in humans, affecting 1 in 691 births [219]. DS patients have intellectual and cognitive disabilities, learning and memory deficits, and maturation delay [220, 221]. DS is caused by an extra copy of all or part of chromosome 21 [222]. In DS patients, Dyrk1a is overexpressed in the brain [223] due to the gene being in the Down syndrome critical region (DSCR) of chromosome 21 [224]. Transgenic mice overexpressing Dyrk1a show neurodevelopmental delays like those observed in humans, indicating the critical role of Dyrk1a in DS neurodevelopment. DS mice show delayed walking activity, retarded psychomotor development, and reduced learning skills [225]. The Dryk1a gene locates at the “critical region” within 21q22 and Dyrk1a protein functions in neuronal cells. Dyrk1a is expressed in neural progenitor cells in early vertebrate embryos [226]. Long-term overexpression of Dyrk1a is thought to be responsible for the neurodevelopmental defects observed in Down syndrome (DS) patients. For example, overexpressed Dyrk1a phosphorylates at Ser 10 of $p27^{Kip1}$, stabilizing $p27^{Kip1}$.

Overexpression of $p27^{Kip1}$ leads to mammalian cell apoptosis [216]. Dykr1a also phosphorylates at Thr 286 of Cyclin D1 and induces the degradation of Cyclin D1 [288]. $P27^{Kip1}$ inhibits Cyclin D1, which promotes the cell cycle from G1 to S [227]. Upregulated $p27^{Kip1}$ and downregulated Cyclin D1 decreases neural cell proliferation and increase differentiation [288]. Overexpression of Dyrk1a also causes hyper-phosphorylating tau at residue Thr212 [228]. Taken together, the overexpressed Dyrk1a in neuronal cells contributes to cell degradation in DS. The "gene-dosage" hypothesis is a very accepted theory for DS [229], and some Dyrk1a inhibitors have been developed as a potential therapeutic strategy.

4.2.3 Dyrk1a functions in Alzheimer's disease

DS patients also show a dramatically increased chance of developing Alzheimer's disease (AD), hereafter referred to as combined DS/AD [230]. AD is associated with alterations in Dyrk1a expression and phosphorylated tau [231]. AD patients and mouse models of AD [232] show higher than normal levels of Dyrk1a mRNA in the hippocampus, and the overexpressed Dyrk1a in the brain leads to hyper-phosphorylated tau protein and neurofibrillary degeneration [228, 232, 233]. Hyper-phosphorylated tau causes neurofibrillary degradation by forming paired helical filaments (PHF) [234]. PHF aggregates and then forms intracellular neurofibrillary tangle (NFT) that leads to neuronal death [235, 236]. Inhibition of Dyrk1a restores tau from hyper-phosphorylation in mice models [237]. Comparing DS patients to healthy individuals, the three copies of chromosome 21 cause 1.4-fold DYRK1A transcription in lymphoblastoid [238], leading to hyperphosphorylated tau and AD symptoms. This mechanism is also thought to underlie

the high prevalence of early-onset AD symptoms in DS patients, where in early-onset AD 55% of those under 60 years old show AD symptoms, and 77% of those over 60 years old show AD symptoms [239]. Overall, overexpression of Dyrk1a is linked to AD and combined DS/AD.

4.2.4 Dyrk1a functions in type 2 and type 3 diabetes

Type 2 diabetes (T2D) and AD were considered two independent diseases. However, with more and more experimental pieces of evidence, the similarity of the pathophysiology of T2D and AD has emerged, such as insulin resistance [240]. Insulin resistance is caused by prolonged metabolic stress and decreased cellular response to insulin in obesity and diabetes [241]. Insulin resistance is a significant factor leading to T2D [242, 243], as it maintains energy homeostasis. β cells promote glucose-stimulated insulin secretion. β cell mass increases first once T2D develops, but then the cells are damaged, and the mass is reduced over time due to high blood sugar and insulin dysfunction [244]. Dyrk1a phosphorylate at Ser28 of LIN52 of the MuvB core of the DREAM complex (dimerization partner, RP-like, E2F, and multi-vulval class B), and then recruit the p130/E2F4/DP1 complex. The DREAM complex then represses genes such as MYBL2 that are required for the cell cycle [245]. Without Dyrk1a phosphorylation, the DREAM complex can no longer repress the cell cycle genes. Therefore, the proliferation of β cells is induced. The N-terminus of Dyrk1a interacts with (insulin receptor substrate 2) IRS2 and phosphorylate at tyrosine residues, which promotes polyubiquitination and degradation of IRS2 [246]. However, IRS2 is required for β cell proliferation [247]. β cell proliferation is restored by

inhibition of Dyrk1a by Harmine [248], which is a Dyrk1a inhibitor that will be discussed in 4.1.6.

At the same time, insulin functions as a neurotrophic and neuroprotective factor. Insulin injection increases rats' brain-derived neurotrophic factor/tropomyosin receptor kinase B, so rats perform better during the Morris water maze test [249]. T2D patients have higher risk of AD development. On another hand, AD patients show significant brain insulin resistance [250], although the function of insulin in the central nervous system is still unclear. In addition to the high risk of AD among T2D patients, AD has been considered a metabolic disease. The concept of type 3 diabetes (T3D) was proposed, in which insulin resistance may play a role in development of AD [251]. Overexpression of Dyrk1a has been linked to AD since overexpressed Dyrk1a causes hyperphosphorylation of tau [252]. Dyrk1a has been linked to T2D and contributes to T3D, and inhibition of Dyrk1a by small molecule inhibitors shows promised results of treatment [253]. For example, one of the studies of inhibitor of Dyrk1a, Harmine, has shown that it can induce human pancreatic B-cell proliferation *in vitro* and *in vivo* by reducing the level of Dyrk1a [254].

4.2.5 Dyrk1a functions in cancer

Acute lymphoblastic leukemia (ALL) is the most common cancer for children, especially children with DS [205]. During treating ALL, genetic markers are tested to determine if an upgrade risk stratification should be used due to the markers linked to relapse and treatment resistance [255]. Intrachromosomal amplification of chromosome 21 (iAMP21) is one of the intermediate genetic markers in B-cell precursor (BCP) ALL that is associated with the

patient's age, low white blood cell, relapse, and prognosis [256, 257, 258]. BCP ALL patients with the iAMP21 marker show that Dyrk1a is overexpressed, determined by DNA arrays and RNA sequencing [259]. Dyrk1a is significantly overexpressed in leukemia cells from pediatric ALL patients by gene expression profiling and qPCR [261, 290]. The overexpression of Dyrk1a is associated with poor prognosis in pediatric ALL patients with iAMP21 [259], in which patients do not show favorable treatment response outcomes and overall survival. With the treatment of two Dyrk1a inhibitors, Harmine and INDY, pediatric ALL cell lines show reduced proliferation and viability. Inhibition of Dyrk1a leads to the degradation of FOXO1, STAT3, and Cyclin D3 and promotes ALL cell chemosensitivity [215]. Treating with Harmine also impacts the mouse model of ALL, in which the number of leukemic cells in the bone marrow and the incidence of relapse are reduced.

Acute myeloid leukemia (AML) is one of the most common cancers in adults. Overexpression of Dyrk1a has also been found in AML. Compared with normal hematopoietic cells, AML cells from patients have overexpressed Dyrk1a by doing microarray [260]. Cells from the bone marrow of AML patients also show high level of the Dyrk1a gene by real-time PCR [291]. Moreover, after the treatment of the chemotherapy drug Cytarabine, the Dyrk1a gene is even more upregulated in AML patient cells [262], suggesting Dykr1a might be involved in chemotherapy resistance. The Dyrk1a inhibitor, Harmine, was used in AML cells and showed reduced cell growth in vitro [263]. Another Dyrk1a inhibitor, AZ191, induces apoptosis in AML cells in the mouse model [264]. When

treating AML cells with Cytarabine and a Dyrk1a inhibitor L41-7, cells were sensitive to Cytarabine and showed reduced growth [245].

Overexpressed Dyrk1a has been found in breast cancer cell lines and breast tumors, and Harmine inhibits the growth of breast cancer cells by inducing apoptosis *in vitro* and *in vivo* [265, 266]. Data have shown that Dyrk1a is expressed highest in basal-like breast cancer, a more aggressive subtype [267]. One of the reasons could be that Dyrk1a promoted cell proliferation by activating the Wnt/β-catenin signaling pathway [268]. Wnt ligand stabilizes β-catenin and then lead to accumulate in the cytoplasm. β-catenin then translocates to the nucleus and interacts with transcription factors such as T-cell factor/lymphoid enhancer factor to activate genes involved in cell proliferation. Dyrk1a can stabilize β-catenin by phosphorylating it at multiple sites such as Ser 45 and Ser33. Using Harmine and siRNA to inhibit Dyrk1a, basal-like breast cancer cells show reduced growth and metastasis.

Overexpression of Dyrk1a has been found in pancreatic cancer tissues and is associated with poor prognosis and increased cell proliferation [269]. Overexpressed Dyrk1a activates the NF-κB pathway and the PI3K/Akt/mTOR pathway to promote pancreatic cancer cell growth [270, 271]. Knockdown of Dyrk1a in pancreatic cancer cells inhibits cell proliferation and induces apoptosis [272].

4.2.6 Targeting Dyrk1a for therapy

There are some inhibitors against Dyrk1a. The neutral product epigallocatechin-gallate (EGCG) from green tea shows inhibition of Dyrk1a but also complicated pharmacokinetic properties [273]. For example, EGCG shows both conjugated and unconjugated forms of

Dyrka in different tissue in mice after administration of EGCG and targeted different pathways. One of the well-known Dryk1a inhibitors, Harmine, is an alkaloid isolated from the South American vine as well as a potent inhibitor of monoamine oxidase [274]. Harmine inhibits Dyrk1a (IC_{50} = 33 nM) by binding to the ATP binding site of Dyrk1a and involving two hydrogen bonds in the ATP binding pocket [205, 275, 276). Harmine shows minimal toxicity in cells at the concentration required to inhibit Dyrk1a completely but did not affect cell viability. CX-4945 has been considered a CK2 (regulates cell growth, proliferation, and apoptosis [277]) inhibitor and is currently in clinical trials for cholangiocarcinoma and basal cell carcinoma as a CK2 inhibitor. However, it also functions as a Dyrk1a (IC_{50} = 6.8 nM) and Dyrk1b (IC_{50} = 6.4 nM) inhibitor [237].

One of the challenges of research on inhibitors of Dyrk1a is selectivity, as Dyrk1a shares at least 85% of the ATP binding site with Dykr1b [206], and most current inhibitors bind to the ATP binding site, and so they lack high selectivity for Dyrk1a; this may cause safety and efficacy issue and potential failure in clinical trials as off-target can lead to unexpected side effects. For example, Harmine binds and inhibits Dyrk1a at 33 nM IC_{50} while it also inhibits Dyrk1b (IC_{50} = 165 nM), Dyrk2 (IC_{50} = 900 nM), and Dyrk3 (IC_{50} = 800 nM) [275]. Inhibition of both Dyrk1a and Dyrk1b can increase human B cell proliferation [278]. However, although Dyrk2 is upregulated in cancer cells, such as glioma and lung adenocarcinoma, it is downregulated in breast cancer, chronic myeloid leukemia [279], and colorectal cancer, and so Harmine would impact Dyrk2's tumor-suppressive function. Lastly, Dyrk3 is a tumor-suppresser down-regulated in hepatocellular carcinoma [280]. As a potential candidate for HCC therapy, Dyrk3 phosphorylates NCOA3 and

negatively regulates ATF4 in the Dyrk3/NCOA3/ATF4 axis, suggesting that If Dyrk3 were inhibited, ATF4 would promote cancer cell proliferation and invasion. Because Harmine inhibit both Dyk1a (33 nM IC_{50}) and Dyrk1b (165 nM IC_{50}), its selectivity is somewhat limited [281] but it entered phase I clinical trials for diabetes mellitus in 2022.

Thus, Dyrk1a is a novel therapeutic target for various disease states, and the research efforts described in this chapter have been focused on hit identification and initial hits-to-leads studies that target Dyrk1a. Our future efforts will focus on improving hit small molecules selectivity to Dyrk1a.

4.3 Methods

4.3.1 Expression

Human Dyrk1a (127–485) was cloned into vector pD441-NH with an N-terminal 6x His tag (6His-Dyrk1a) (ATUM). Dyrk1a was expressed by transforming plasmids in different *E. coli* cell lines, and then a small-scale expression test was done to determine the optimal cell line and expression condition and saved glycerol stock. The glycerol stock was cultured in 50 mL of LB with kanamycin at 37℃ O/N. The O/N culture was inoculated into 12L of LB with kanamycin at 37℃ at 225 rpm. the cells were induced with 0.4 mmol/L IPTG at OD600 0.8 and kept shaking for four hours 37℃. Cells were harvested by centrifugation at 4℃ and 4,000 rpm for 15 minutes, resuspended at a ratio of 3ml lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 25 mM imidazole, 10 mM BME) per gram pellet. Per 40 mL of the resuspended cell, it was sonicated on the ice at 70% amplitude for four minutes, with 10 seconds on and 10 seconds off. The supernatant was obtained by

centrifugation at 4℃ and 30,000 rpm for 40 minutes. The cell lysate was purified right after or stored at -80℃.

4.3.2 Purification

Supernatant from cell lysate in 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 25 mM imidazole, 10 mM BME was purified by HisTrap FF, and eluted by a gradient of 1M imidazole buffer, exchanged to a low salt buffer and then loaded onto cation exchange chromatography (HiTrap SP FF; Cytiva) and eluted by a gradient of 1M NaCl buffer, Superdex 75 and eluted by a low salt buffer, and finally Enrich high-resolution cation exchange chromatography (ENrich S; Bio-Rad) and eluted by a gradient of 1M NaCl buffer. SDS-PAGE gels confirmed fractions from each step, and then desired fractions were collected for the next step. The final sample was stored in 50 mM HEPES pH 7.5, 300 mM NaCl at -80℃ after flash froze in liquid nitrogen.

4.3.3 TSA

10 mM of ATP site AI CADD small molecules were sent from Atomwise Inc, and 200 mM of ATP site DTP small molecules were sent from the Mobley Lab, UCI. Following the TSA method from 2.5.2.1, hits were determined using a final of 4.2 μM of Dyrk1a with the reaction buffer (50 mM HEPES pH 8.0, 50 mM NaCl).

4.3.4 ADP assay

The ADP-Glo^{TM} Kinase Assay (Promega) measures ADP generated from ATP from a kinase reaction. This assay confirmed that hits can inhibit Dyrk1a activity and then determine IC_{50}. Following the ADP-Glo^{TM} kinase assay technical manual, a standard curve was generated first to estimate the amount of ADP generated. A serial dilution from 1 mM

of ATP and ADP was added into each well of a 96-well plate with 90 μL of kinase reaction buffer (40 mM Tris pH 7.5, 20 mM $MgCl_2$, 0.1 mg/ml BSA) and 25 μL ADP-Glo reagent that can stop the kinase reaction and remove ATP. After 40-min room temperature incubation, a 50 μL kinase detection reagent, which converted ADP to ATP and introduced luciferase and luciferin to detect ATP, was added to each reaction and incubated for another 1 hour. The luminescence was read and recorded by SpectraMac iD5 and generated a standard curve for relative light units (RLU) and ATP-to-ADP conversion. A dyrk1a concentration test was done to determine the optimal concentration. A serial dilution of Dyrk1a (down to 0 as background signal) was mixed with desired ATP in the kinase reaction buffer, 1-hour room temperature incubation. Then, ADP-Glo reagent was added for each reaction for 40 minutes of room temperature incubation. Then, a kinase detection reagent was added to each reaction. The plate was measured after 1-hour room temperature incubation. A 0.5 μM of Dyrk1a final concentration was determined as the desired kinase amount.

The small molecules screening was done using 10 mM of ATP site AI-CADD small molecules. Each test contained 0.5 mM of final small molecules concentration in DMSO, 0.5 μM of Dyrk1a final concentration, 100 μM of ATP final concentration, the kinase reaction buffer, and DMSO as the negative control. A known Dyrk1a inhibitor, Verzenio (abemaciclib) was used as a positive control and determined the Z-factor of this ADP assay for Dyrk1a screening. The plate was read and determined hits that inhibited Dyrk1a auto-phosphorylation.

IC_{50} was done with the hits. A serial dilution of inhibitor was tested to generate a curve of Log 10 of inhibitor concentration and RLU (GraphPad).

All tests were done three times. Average and standard deviation were used to plot to visualize the impact of each small molecules on Dyrk1a activity. CV was less than 10% to express precision and repeatability.

4.3.5 Crystallography study

Crystal trays were set up at RT using 96-well or 24-well crystal trays by sitting drop. 300-400 μL of reservoir was added in the reservoir chamber first. 0.5 or 1 μL of protein (or mixed with small molecule hits) was added to the well, and then mixed with the same amount of reservoir solution. The chamber was sealer by tape and incubate at RT. Crystal trays were examined by a microscope after setup, and everyone for 2 weeks.

4.4 Results

4.4.1 Expression

In small-scale expression test, 6His-Dyrk1a plasmid was transformed with in *E. coli* BL21(DE3) or BL21-CondonPlus-RP, and cultured in LB, TB, or 2XYT. Cells were induced at OD600 0.8 and set to shake at either 37℃ 4 hours, 30℃ 5 hours, or 22℃ O/N. All cells were collected and then loaded onto nickel spin columns, eluted with high imidazole buffer, loaded onto nickel spin column, eluted with high glutathione buffer, and run on SDS-PAGE gels for comparison (co-expression samples only). Dyrk1a is a 42 kDa protein, and a band of 40 kDa size showed in elution samples from all samples. The highest expression level was using BL21(DE3) and LB inducement of 37℃ 4 hours (**Supplementary material figure 4.1**).

4.4.2 Purification

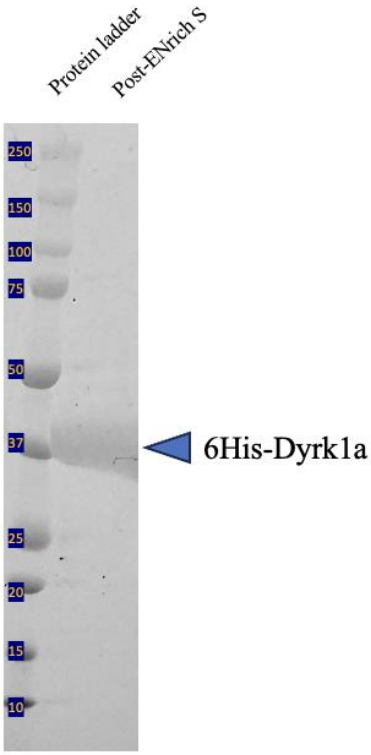


Figure 4.1 Purified Dyrk1a sample. The final post-ENrich S purification sample of Dyrk1a was loaded on an 8-16% SDS-PAGE gel (Novex Tris-Glycine Mini Protein Gel; Thermo Fisher). Lane 1: protein ladder. Lane 2: final sample from ENrich S elution fraction.

In the large-scale expression for purification, 6His-Dyrk1a went through HisTrap FF, HiTrap SP FF, Superdec 75, and ENrich S. Yield was 0.875 mg of Dyrk1a / L of *E. coli* culture. (**figure 4.1** & **Supplementary material figure 4.2**).

4.4.3 TSA

The optimized protein concentration and reaction buffer for TSA were determined first. 4.2 μM of Dyrk1a in the reaction buffer 50 mM HEPES pH 8.0, 50 mM NaCl generated the desired RFU and relatively low T_m for TSA of Dyrk1a, which was 42°C (**Supplementary material figure 4.3**). During the screening test of AI-CADD small molecules from Atomwise, 29.17% of the small molecules had positively shifted at least 0.5°C of the melting point of Dyrk1a, and 7.29% had positively shifted at least 2°C (**figure 4.2** & **supplementary material table 4.1**). The small molecules with 2°C shift were considered as binders and were further tested to confirm as hits.

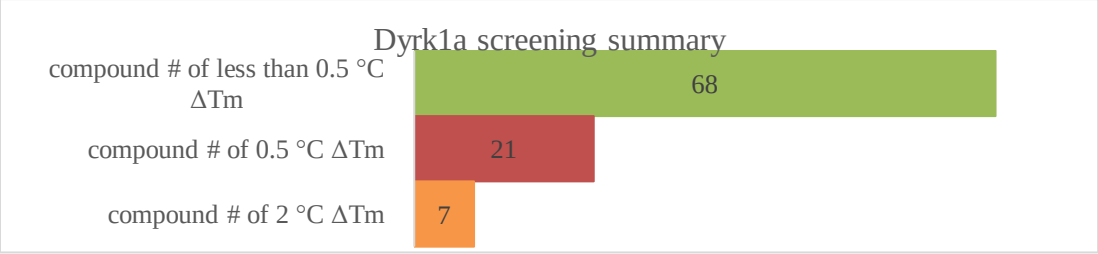


Figure 4.2 Summary of Dyrk1a screening with AI-CADD small molecules using TSA. 96 AW small molecules were derived by virtual screening in the library of Dyrk1a AI-CADD small molecules. 68 AW small molecules shifted Dyrk1a T_m less than 0.5°C (including no shift and negatively shift). 21 AW small molecules (21.9%) shifted Dyrk1a T_m 0.5 - 2°C. 7 AW small molecules (9.3%) shifted Dyrk1a T_m more than 2°C.

4.4.4. ADP assay

To use the ADP assay kit, a 100μM of ATP final concentration was determined as the desired ATP amount (**Supplementary material figure 4.4 A**). A 0.5 μM of Dyrk1a final concentration was determined as the desired kinase amount (**Supplementary material figure 4.4 B**), and the signal/background ratio was 11.06. Using Abemaciclib, a Z-factor of 0.9415 was determined and indicated that the ADP assay was valid to test Dyrk1a (**Supplementary material figure 4.4 D**).

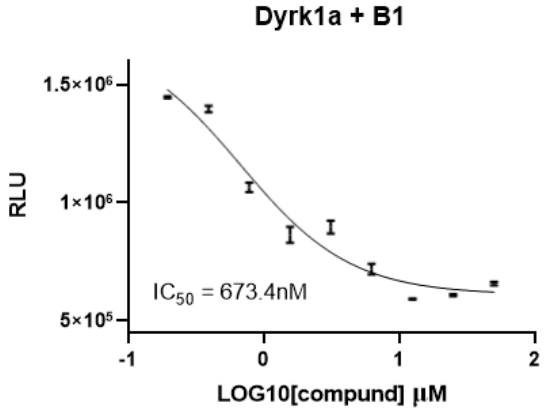
The screenings were done in three rounds with different concentrations of AW small molecules, 100 μM, 50 μM, and 10 μM, aiming to confirm binders and identify hits. During the 100 μM screening, 21 of AW small molecules showed reduction of activity of Dyrk1a. Activity % = (RLU of Dyrk1a with small molecules / RLU of Dyrk1a with DMSO) *100. After three rounds, 13 of AW small molecules showed inhibition trends with different dosages (**table 4.1** & **Supplementary material table 4.2**). IC_{50} of seven small molecules were determined, and the best one was 0.669 μM using AW H4 (figure 4.3). Notably, ΔT_m of AW H4 was 12.17 °C from TSA screening.

Dyrk1a AI-CADD small molecule	average °C ΔT_m (500 uM)	average ADP activity % (100 uM)	ADP dosage trend	ADP best fit IC_{50} µM	ADP 95% IC_{50} µM	ADP IC_{50} R^2
H04	12.17	16.32	yes	0.66	0.5436 to 0.7601	0.99
B01	-1.67	55.9	yes	0.67	0.1039 to 1.074	0.96
D09	4.33	53.3	yes	1.12	1.081 to 2.194	0.93
E09	2.33	47.59	yes	1.73	1.394 to 2.175	0.95
E10	NA	77.9	yes	1.89	1.701 to 2.281	0.96
D08	5	53.18	yes	1.92	1.447 to 2.469	0.94
G06	-1	63.07	yes	3.42	2.484 to 4.884	0.92

Table 4.1 Summary of Dyrk1a with AI-CADD small molecule hits using ADP assay.

Seven AW small molecules reduced activity of Dyrk1a in the first round of screening (100 µM final concentration) and showed inhibition trend in the sconed and third round of screenings (50 µM and 10 µM) using the ADP assay. Their IC_{50} values were measured by ADP assay and calculated by GraphPad. R^2 quantifies goodness of fit. The ΔT_m results of the hits from TSA were added into the table to compare with ADP activity % results.

A.



B.

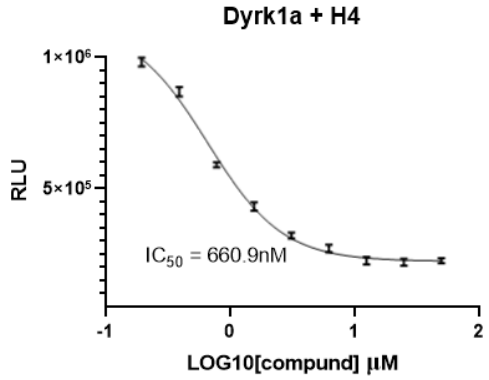


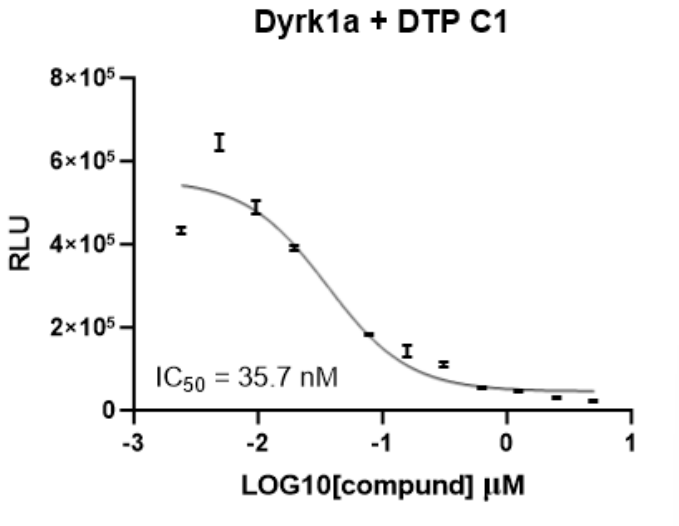
Figure 4.3 IC50 figures of AW B1 and H4. IC_{50} figure and values of AW B1 (panel A) and AW H4 (panel B) were generated by GraphPad. The best fit IC_{50} values were used.

Using AW B1 and H4 as Dyrk1a hits, virtual docking was done by the Mobley Lab, UCI, aiming to generate small molecules with lower IC_{50}, as the first step in the hit-to-lead. 40 small molecules (Dyrk1a DTP small molecule) were generated and tested by ADP assay in duplicate first (100 µM final). 16 small molecules reduced the activity of Dyrk1a and further tested for IC_{50} values (**table 4.2** & **Supplementary material table 4.3**). Notably, DTP C1 and B6 generated IC_{50} 36 nM and 92 nM, respectively (**figure 4.4**).

Dyrk1a DTP small molecule	% activity	AMP best fit IC_{50} µM	AMP 95% IC_{50} µM	AMP IC_{50} R squared
C01	4.80	0.036	0.02189 to 0.05799	0.93
B06	1.36	0.092	0.07240 to 0.1187	0.97
C02	16.55	0.239	0.09719 to ???	0.94
C07	20.12	1.275	0.6747 to 2.549	0.98
D01	21.78	2.058	1.489 to 12.60	0.95
C06	4.75	2.258	1.230 to ???	0.91
A07	25.88	2.335	1.216 to 3.945	0.93
C03	16.51	2.370	???	0.69
C08	23.70	2.406	???	0.59
C10	21.33	2.582	2.145 to ???	0.80
A05	27.30	3.800	0.9729 to 5.843	0.92
C04	12.38	4.657	1.045 to ???	0.71
B03	4.62	>50		
C11	19.92	>50		
A03	22.40	>50		
A08	24.87	>50		

Table 4.2 Summary of Dyrk1a with DTP small molecule hits using ADP assay. 16 DTP small molecules reduced activity of Dyrk1a in the screening (100 µM final concentration) using the ADP assay. Their IC_{50} values were measured by ADP assay and calculated by GraphPad (Prism). R^2 quantifies goodness of fit.

A.



B.

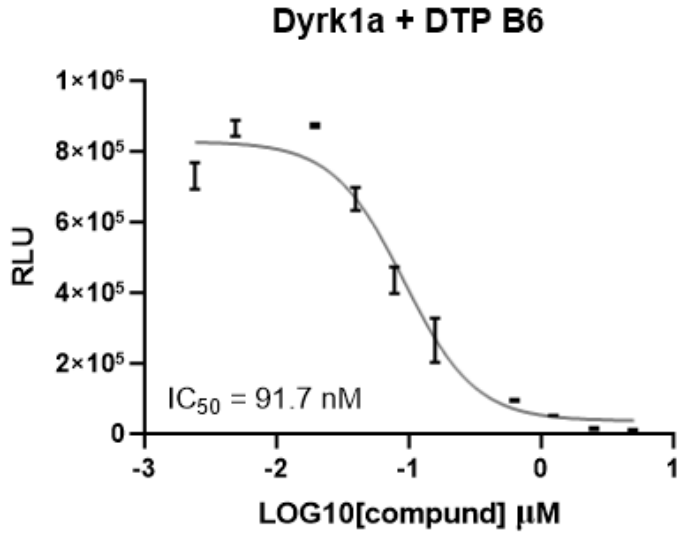


Figure 4.4 IC50 figures of DTP C1 and B6. IC_{50} figure and values of DTP C1 (panel A) and DTP B6 (panel B) were generated by GraphPad. The best fit IC_{50} values were used.

4.4.5 Crystallography study

A. B. C.

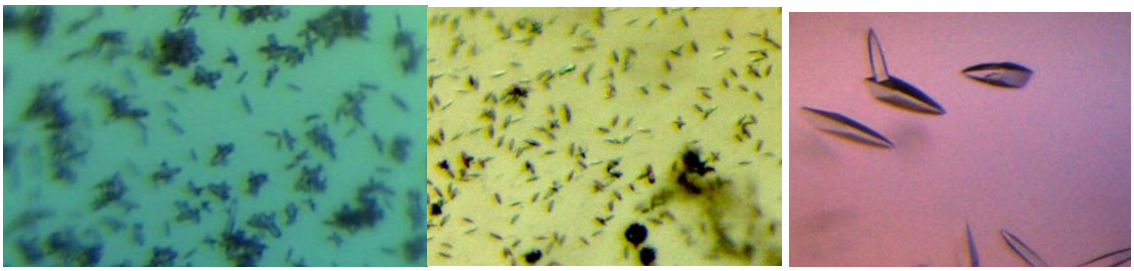


Figure 4.5. Dyrk1a co-crystallization with AW H4, D9, and abemacicle. Crystal condition screening was conducted with 5 mg/mL of Dyrk1a with 1.2X of AW2 H4, D9, or 5X of a known inhibitor abemacicle by sitting drop at RT. Needle crystals of Dyrk1a co-crystal with AW H4 (A.), AW D9 (B.), and abemacicle (C) were formed after 4 days incubation.

To visualize the binding between Dyrk1a and hits, crystallography study was conducted with apo Dyrk1a at first, by screening crystal conditions from Anatrace top 96 crystallization screen (Anatrace), MCSG crystallization suite (Anatrace), PEG/pH 1 & 2 (Hampton research), PEG/Ion 1 & 2 (Hampton research), Grid Screen Ammonium Sulfate (Hampton research), et al. Crystal trays were set up by sitting drop with 10 or 5 mg/mL of Dyrk1a and incubated at RT. The drops were examined under a microscope, but no crystal formed after one month. All drops were clear. Co-crystallization with AW H4 or D9 were performed with the same condition screening procedure. The protein concentration was 5 mg/mL and ratio of protein: small molecule was 1: 1.2. Needle crystals formed after four days incubation at RT (**figure 4.5**). Microseeding with seed beads technique was conducted

aiming to increase the size of the crystals, but the size was still not ideal for X-ray (data not shown).

4.5 Conclusion

As a kinase, Dyrk1a appears as a highly druggable target as 7.29 % of the AW small molecules had positively shifted T_m of Dyrk1a at least 2 °C. There are already some Dyrk1a inhibitors that inhibit Dyrk1a but lack selectivity (i.e., can also inhibit other members in the DYRK family). Screening against AI-CADD small molecules, seven hits were identified with the IC_{50} in the 10 μM range. In the hit identification stage, AW B1 and H4 (IC_{50} 673.4 nM and 550.9 nM, respectively) showed the most promising results (also obtained co-crystal in figure 4.7) for further optimization. To obtain a higher than 10:1 signal/background ratio, the autophosphorylation of Dyrk1a was detected using the ADP assay, but not Dyrk1a phosphorylate DYRKtide. In the future, other assays might be tested if the activity of Dyrk1a phosphorylation DYRKtide. For example, in vitro protein kinase assay uses labeled [^{33}P]ATP[γP]. The labeled [^{33}P] into DYRKtide can be measured and determine the activity of Dyrkr1a [275]. Although microseeding with the seed beads could not increase the size of current co-crystal, further crystallography techniques can be conducted, such as microseed matrix screening and macroseeding, as well as wider sparse matrix screens could be conducted.

As discussed in Chapter 2, IC_{50} measurements are assay specific. The results here were mainly for ranking the hits, so the next stage could focus on the best ones. Therefore, affinity measurements that include K_d or K_i are essential before we conduct lead optimization steps. For Dyrk1a, a selectivity test is also critical by testing the hits with

Dyrk1b. As all the AI-CADD small molecules targeting the ATP binding pocket of Dyrk1a, and Dyrk1a and Dyrk1b have ATP binding pockets 85 % identical, there is a possibility of AW B1 and H4 inhibiting Dyrk1b. Like Dyrk1a, Dyrk1b can undergo autophosphorylation in Ser 520 residue via an intramolecular mechanism *in vitro* [282], and Dyrk1b also undergoes autophosphorylation, on residue Tyr 273 [283]. Therefore, the ADP assay will likely work with Dyrk1b. For example, if the result of AW B1 and H4 showed inhibition of Dyrk1b but with significantly higher IC_{50} than Dyrk1a, the hits might still be valid to pursue and optimize further. Docking studies based on AW B1 and H4 has begun already aiming to optimize the hits. The ADP assay was conducted to determine if any improvement of IC_{50}. From 40 small molecules docking based on AW H4 by the Mobley Lab, UCI, 2 of them showed 35.7 nM and 91.7 nM IC_{50} (figure 4.8). This improvement of IC_{50} is promising for the hit-to-lead stage. Our initial co-crystallization studies on Dyrk1a with AW H4 or D9 was successful in producing crystals, crystallization studies will also be completed on small molecules of interest with promising IC_{50} values. Conducting X-ray crystallography on Dyrk1a and/or Dyrk1b with the hits will allow us to visualize the binding mode(s) at atomic resolutions, which greatly aids hits-to-leads optimization.

Neither ubiquitin-proteasome inhibition by MG132 nor the autophagy inhibitor Chloroquine restores Dyrk1a expression [284], indicating that degradation of Dyrk1a is not the main form of cellular regulation of this kinase. Therefore, proteolysis targeting chimera (PROTAC) and AUTAC are not possible to degrade Dyrk1a.

Although not the focus of this study, another way to increase the selectivity of hits is to identify an allosteric interaction site, since the ATP binding pocket and the activation

loop of autophosphorylation are very similar in Dyrk1a and Dyrk1b [285]. Interestingly, Dyrk1a has a nuclear localization site (NLS) in its catalytic domain, a histidine repeat, and serine/threonine repeats near its C-terminal (**figure 4.6**), which other members in the DYRK family do not have; these could potentially be targeted by an allosteric inhibitor.

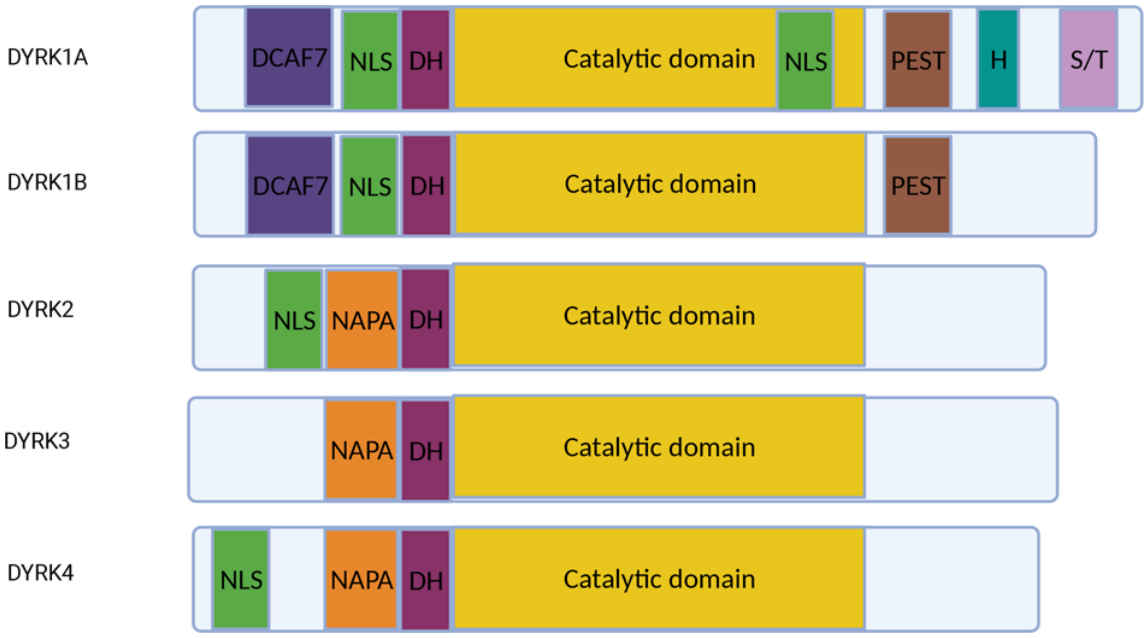
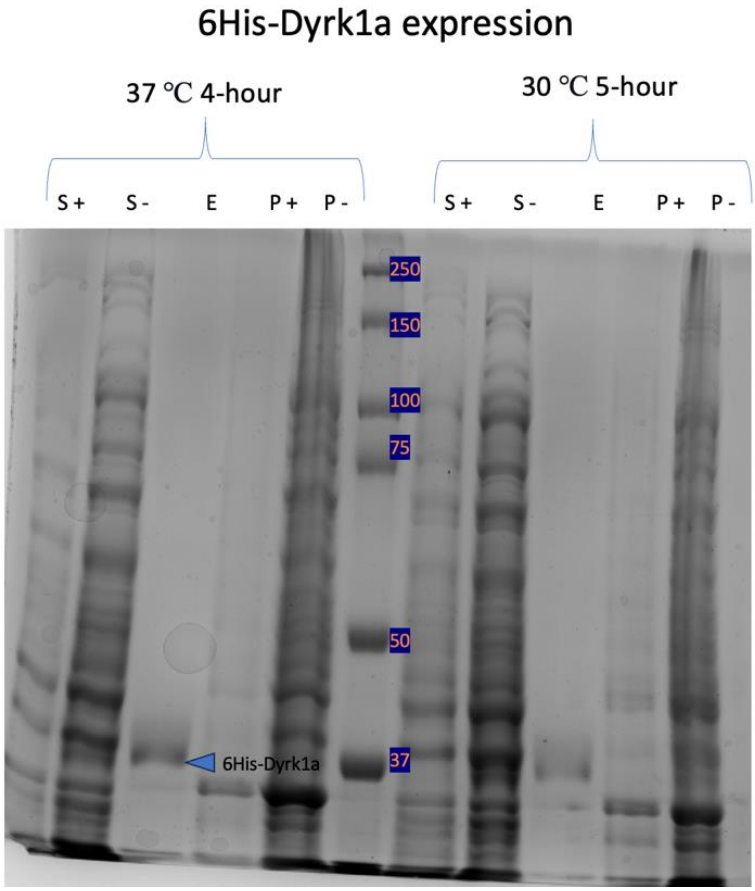
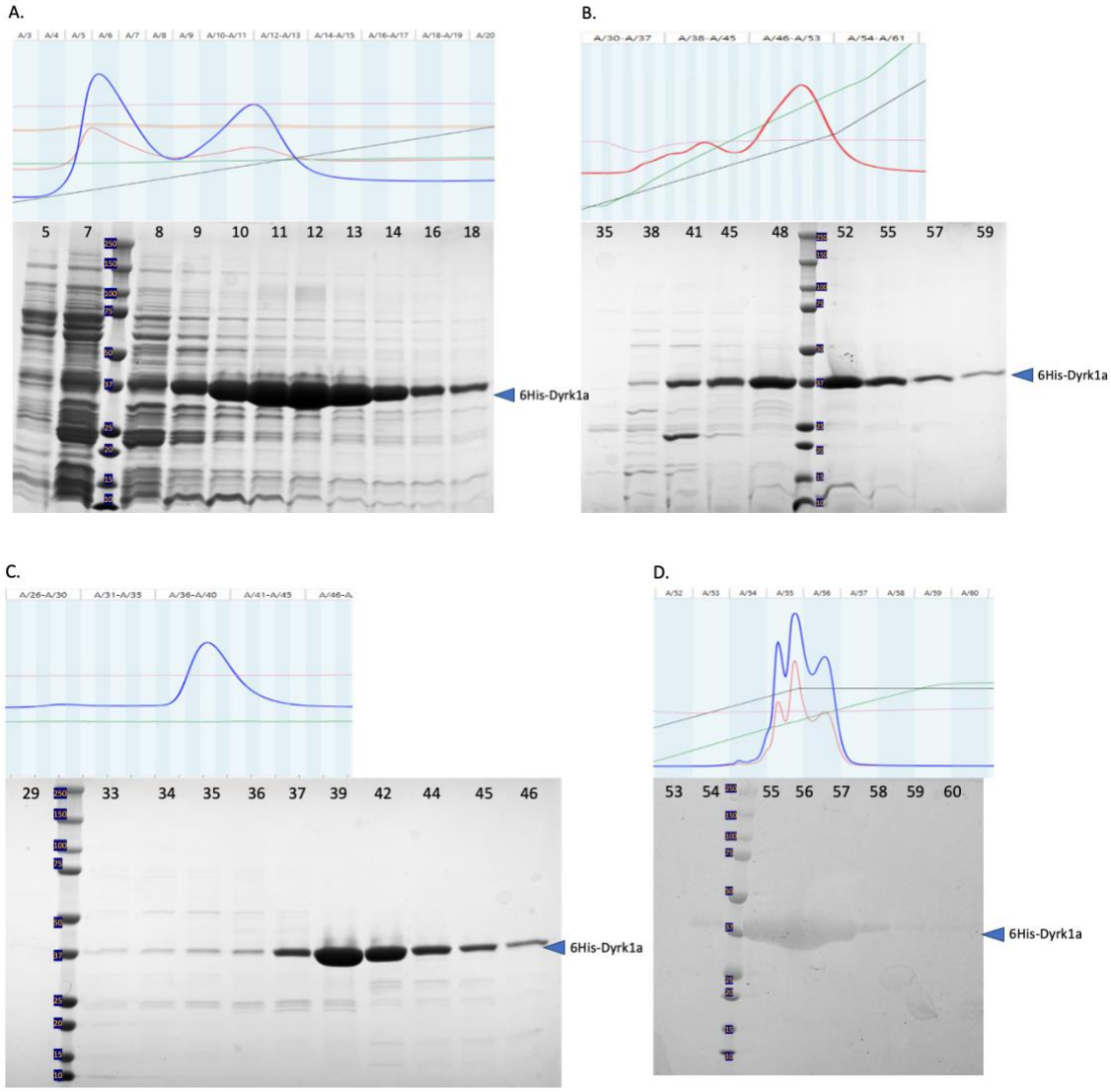


Figure 4.6 schematic diagram of members in the DYRK family The members of the DYRK family contain the DYRK homology (DH) box and the catalytic domain [287]. Dyrk1a and Dyrk1b contain the DCAF7 binding domain, Nuclear localization site (NLS), and the protein, glutamate, serine, threonine (PEST) domain. Only Dyrk1a contains a second NLS in its catalytic domain, a Histidine repeat (H), and a serine threonine (S/T) repeats at its C-terminal.

4.6 Supplementary material

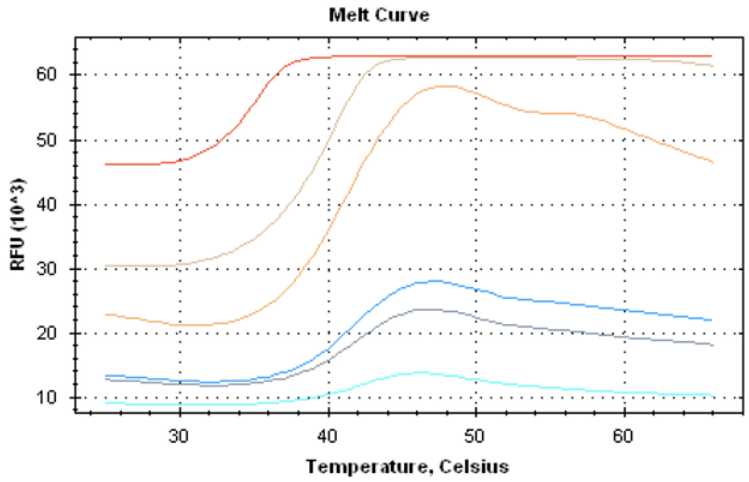


Supplementary material figure 4.1 Dyrk1a expression small-scale test. 6His-Dyrk1a expressed in *E. coli* BL21 (DE3) with LB and induced by IPTG @ OD 0.8 at 37 ℃ 4-hour or at 30 ℃ 5-hour. Cells were harvested and examined by NI-NTA spin column. Samples were loaded on an 8-16% SDS-PAGE gel. Lane 1 & 7: cell supernatant with IPTG inducement (S+). Lane 2 & 8: cell supernatant without IPTG inducement (S-). Lane 3 & 9: elution (E) from NI-NTA spin column. Lane 4 &10: cell precipitate with IPTG inducement (P+). Lane 5 & 11: cell precipitate without IPTG inducement (P-). Lane 6: protein ladder. Lane 1 – 5 were sample from 37 ℃ 4-hour, and lane 7 – 11 were sample from 30 ℃ 5-hour.

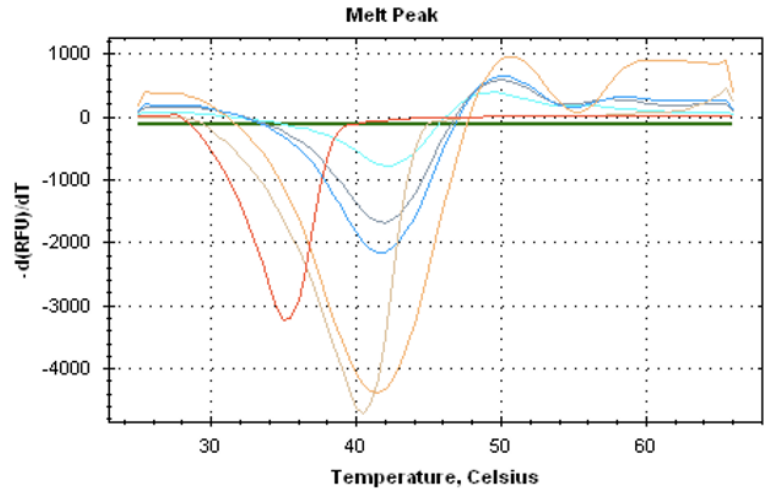

A.
5 7 8 9 10 11 12 13 14 16 18
6His-Dyrk1a
B.
35 38 41 45 48 52 55 57 59
6His-Dyrk1a
C.
29 33 34 35 36 37 39 42 44 45 46
6His-Dyrk1a
D.
53 54 55 56 57 58 59 60
6His-Dyrk1a

Supplementary material figure 4.2 Dyrk1a purification by FPLC. 6His-Dyrk1a in *E. coli* BL21 (DE3) with LB and induced by IPTG @ OD 0.8 at 37 ℃ 4-hour. A. HisTrap FF captured Dyrk1a in elution fractions 9-18. B. HiTrap SPFF in fractions 45-62. C. Superdex 75 in fraction 37-46, and D. ENrich S in fraction 55-57. Buffers were used: His A 50 mM HEPES pH 8, 300 mM NaCl, 25 mM imidazole, 1 mM BME & His B 50 mM HEPES pH 8, 300 mM NaCl, 1 M imidazole, 1 mM BME & HiTrap S A 50 mM HEPES pH 6, 50 mM NaCl, 1 mM BME & HiTrap S B 50 mM HEPES pH 6, 1 M NaCl, 1 mM BME. Superdex 50 mM HEPES pH 8, 300 mM NaCl, 1 mM BME. ENrich S A 50 mM HEPES pH 6, 50 mM NaCl, 1 mM BME & ENrich S B 50 mM HEPES pH 6, 1 M NaCl, 1 mM BME. Yield: 0.875 mg of Dyrk1a / L of *E. coli* culture.

A.



B.



Supplementary material figure 4.3 Dyrk1a concentration test for TSA. Dyrk1a serial dilution from 33.6 µM to 1.05 µM in PBS were tested for the concentration test. A. the figure of RFU vs. degree °C. B. the figure of -d(RFU)/dT vs. degree °C. 33.6, 16.8, and 8.4 µM of Dyrk1a (represented by line in red, light brown, and yellow) saturated the threshold of the rt-PCR.

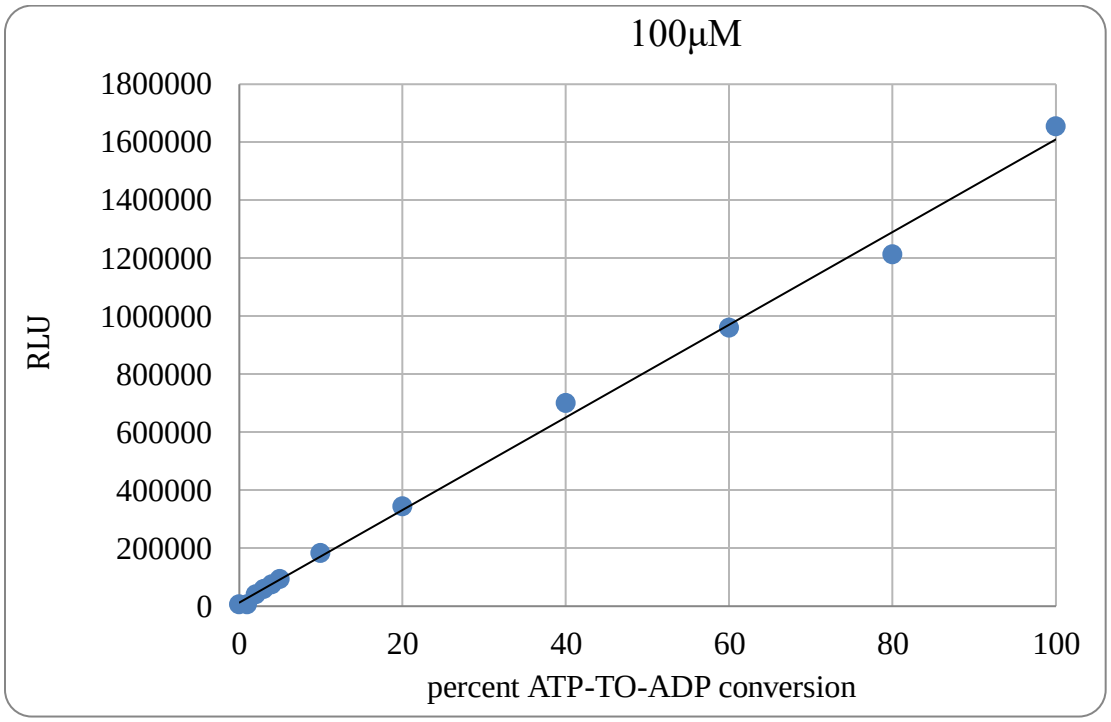
Dyrk1a AI-CADD small molecule	test 1 ΔT_m	test 2 ΔT_m	test 3 ΔT_m	average ΔT_m
A01	0.5	0.5	0.5	0.50
B01	-1.5	-2.0	-1.5	-1.67
C01	4.0	4.0	4.0	4.00
D01	0.0	-0.5	-0.5	-0.33
E01	-2.5	-3.5	-2.5	-2.83
F01	-14.5	-15.0	-15.0	-14.83
G01	0.0	-0.5	-0.5	-0.33
H01	0.0	0.0	0.0	0.00
A02	-0.5	-0.5	-1.0	-0.67
B02	0.0	0.0	0.0	0.00
C02	0.5	0.5	0.5	0.50
D02	1.0	1.5	1.5	1.33
E02	0.0	0.0	0.0	0.00
F02	0.0	-0.5	0.0	-0.17
G02	4.5	4.0	4.0	4.17
H02	-1.0	-1.5	-1.0	-1.17
A03	0.5	0.0	0.5	0.33
B03	0.5	0.0	0.0	0.17
C03	-0.5	0.0	0.0	-0.17
D03	0.0	0.0	0.0	0.00
E03	-0.5	-0.5	-0.5	-0.50
F03	0.0	0.5	0.0	0.17
G03	0.0	0.0	0.0	0.00
H03	0.0	0.5	0.0	0.17
A04	0.5	0.5	0.5	0.50
B04	0.0	0.0	0.0	0.00
C04	0.0	0.0	-0.5	-0.17
D04	-0.5	0.0	0.0	-0.17
E04	0.0	0.0	-0.5	-0.17
F04	-2.0	-3.5	-3.5	-3.00
G04	2.5	3.0	2.5	2.67
H04	12.5	12.0	12.0	12.17
A05	0.5	0.5	0.5	0.50

B05	0.5	0.5	0.5	0.50
C05	-1.5	-1.5	-1.5	-1.50
D05	-1.5	-1.5	-1.5	-1.50
E05	-0.5	-0.5	-0.5	-0.50
F05	-0.5	-0.5	-0.5	-0.50
G05	0.0	0.0	0.5	0.17
H05	-7.5	-8.0	-8.0	-7.83
A06	0.0	0.5	0.0	0.17
B06	0.5	0.5	0.5	0.50
C06	0.5	0.5	0.5	0.50
D06	1.0	1.5	1.5	1.33
E06	-0.5	0.0	0.0	-0.17
F06	1.0	1.0	0.5	0.83
G06	-1.0	-1.0	-1.0	-1.00
H06	0.0	0.0	0.0	0.00
A07	0.5	0.5	0.5	0.50
B07	-1.0	-1.0	-0.5	-0.83
C07	-1.5	-1.5	-1.5	-1.50
D07	1.5	1.5	1.0	1.33
E07	0.5	0.5	0.0	0.33
F07	0.5	0.5	1.0	0.67
G07	0.0	0.0	0.0	0.00
H07	-3.0	-3.5	-3.5	-3.33
A08	0.5	0.5	0.5	0.50
B08	-0.5	-0.5	-0.5	-0.50
C08	-2.0	-2.5	-2.5	-2.33
D08	5.0	5.0	5.0	5.00
E08	-0.5	-1.0	-1.0	-0.83
F08	0.0	0.5	0.0	0.17
G08	1.0	0.5	0.5	0.67
H08	0.5	0.5	0.0	0.33
A09	0.5	0.5	1.0	0.67
B09	-1.0	-0.5	-0.5	-0.67
C09	-1.0	-1.0	-1.5	-1.17

D09	4.0	4.5	4.5	4.33
E09	2.5	2.0	2.5	2.33
F09	-0.5	-0.5	-0.5	-0.50
G09	-1.0	-1.0	-1.0	-1.00
H09	-0.5	-0.5	-0.5	-0.50
A10	0.0	0.5	0.0	0.17
B10	2.5	2.5	2.5	2.50
C10	-2.0	-1.5	-1.5	-1.67
D10	-0.5	-0.5	-0.5	-0.50
E10	-14.5	-15.0	-15.0	-14.83
F10	0.0	0.0	0.0	0.00
G10	-1.0	-0.5	-0.5	-0.67
H10	1.5	2.0	2.0	1.83
A11	-11.0	-11.5	-11.5	-11.33
B11	1.0	1.0	1.0	1.00
C11	0.0	0.0	0.0	0.00
D11	-0.5	-0.5	-0.5	-0.50
E11	-0.5	-0.5	-0.5	-0.50
F11	0.0	0.0	0.0	0.00
G11	-1.0	-1.0	-1.0	-1.00
H11	0.0	0.0	0.0	0.00
A12	-1.0	-0.5	-0.5	-0.67
B12	0.5	0.5	0.5	0.50
C12	-0.5	-0.5	-0.5	-0.50
D12	0.5	0.5	0.5	0.50
E12	0.0	0.0	0.0	0.00
F12	0.5	0.0	0.5	0.33
G12	0.0	0.0	0.0	0.00
H12	0.0	0.0	0.0	0.00

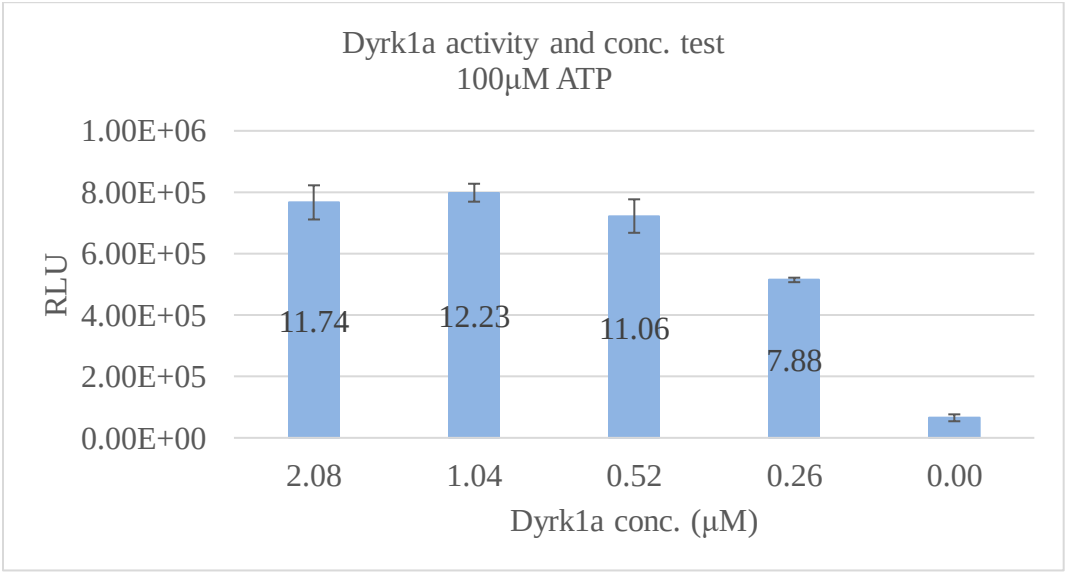
Supplementary material table 4.1 results of screening test of AW AI-CADD small molecules against to Dyrk1a using TSA. AI-CADD small molecules were tested against to Dyrk1a along with DMSO as negative control in triplicates and the average of ΔT_m determined whether the small molecule was a binder of Dyrk1a.

A.

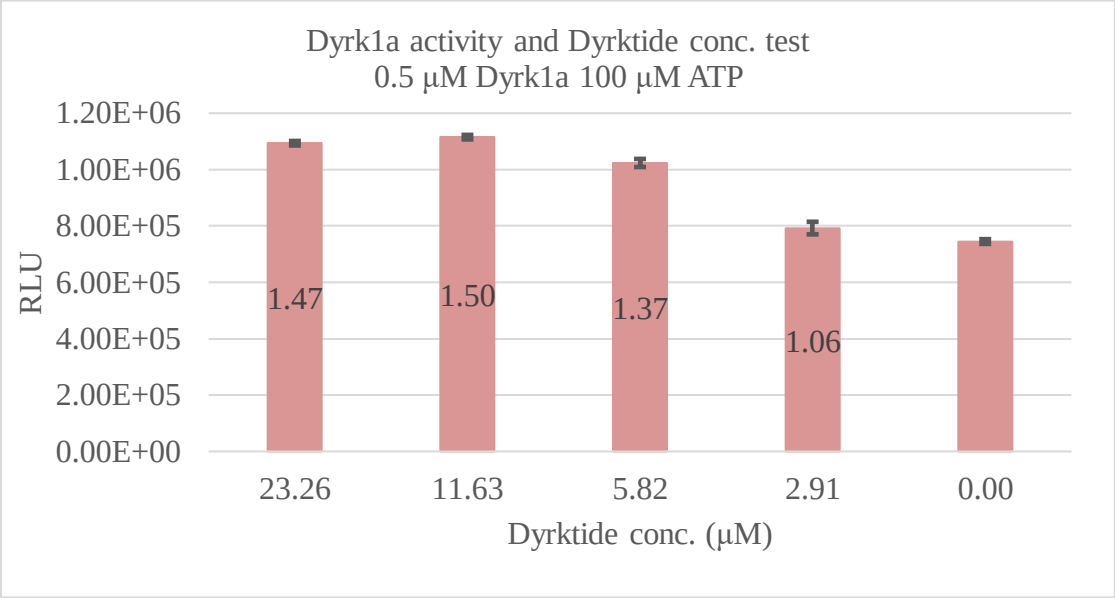


1	2	3	4	5	6	7	8	9	10	11	12	
100	80	60	40	20	10	5	4	3	2	1	0	ADP % in ATP+ADP mixture
0	20	40	60	80	90	95	96	97	98	99	100	ATP % in ATP+ADP mixture

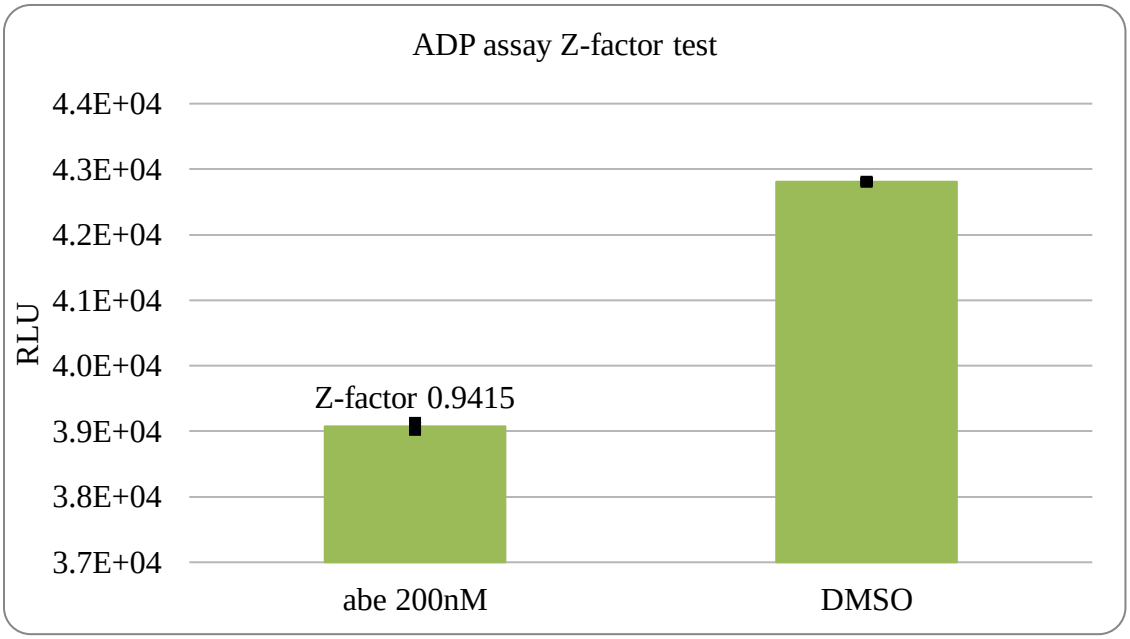
B.



C.


Dyrk1a activity and Dyrktide conc. test
0.5 µM Dyrk1a 100 µM ATP
RLU
1.20E+06
1.00E+06
8.00E+05
6.00E+05
4.00E+05
2.00E+05
0.00E+00
1.47
1.50
1.37
1.06
23.26
11.63
5.82
2.91
0.00
Dyrktide conc. (µM)

D.


ADP assay Z-factor test
RLU
4.4E+04
4.3E+04
4.2E+04
4.1E+04
4.0E+04
3.9E+04
3.8E+04
3.7E+04
Z-factor 0.9415
abe 200nM
DMSO

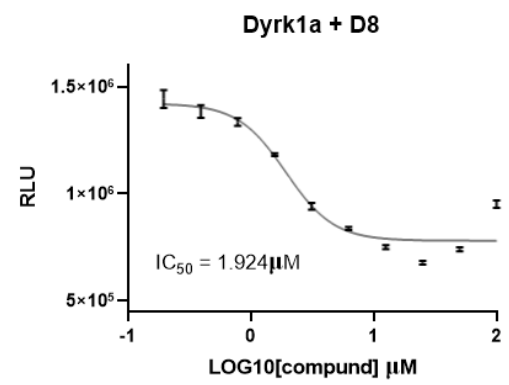
Supplementary material figure 4.4 preparing for the ADP assay. A. ATP-to-ADP conversion standard curve was generated by mixing the appropriate volumes of ATP and ADP. The standard curve could be used to calculate the ATP concentration used in later kinase reactions. B. Dyrk1a concentration test was done with a serial dilution of Dyrk1a, from 2.08 μM to 0.26 μM, and no Dyrk1a as background signal. The S/B ratio was labelled in each column. C. Dyrktide concentration test was done with a serial dilution of Dyrktide, from 23.26 μM to 2.91 μM, and no Dyrktide as background signal. The S/B ratio was labelled in each column. D. ADP assay Z-factor test using the known Dyrk1a inhibitor Abemaciclib.

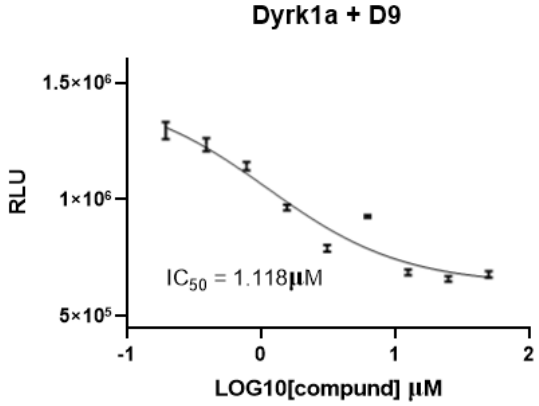
Dyrk1a AI-CADD small molecule	average °C ΔT_m (500 μM)	average AMP activity % (100 μM)	AMP dosage trend	AMP best fit IC_{50} μM	AMP 95% IC_{50} μM	AMP IC_{50} R squared
H04	12.17	16.32	yes	0.66	0.5436 to 0.7601	0.99
B01	-1.67	55.9	yes	0.67	0.1039 to 1.074	0.96
D09	4.33	53.3	yes	1.12	1.081 to 2.194	0.93
E09	2.33	47.59	yes	1.73	1.394 to 2.175	0.95
E10	NA	77.9	yes	1.89	1.701 to 2.281	0.96
D08	5	53.18	yes	1.92	1.447 to 2.469	0.94
G06	-1	63.07	yes	3.42	2.484 to 4.884	0.92
G04	2.67	70.57	yes	>50		
E01	-2.83	84.19	yes	>50		
F06	0.83	86.96	yes	>50		
D06	1.33	87.54	yes	>50		
F01	NA	88.29	yes	>50		
H01	0	89.52	yes	>50		
G02	4.17	75.6	No			
H02	-1.17	84.84	No			
E03	-0.5	68.59	No			
E06	-0.17	79.86	No			
H06	0	88.96	No			
G09	-1	86.27	No			
E11	-0.5	73.2	No			
E12	0	61.22	No			
A01	0.5	126.53				
C01	4	113.3				
D01	-0.33	125.64				
G01	-0.33	94.77				
A02	-0.67	125.47				
B02	0	131.83				
C02	0.5	123.85				
D02	1.33	114.92				

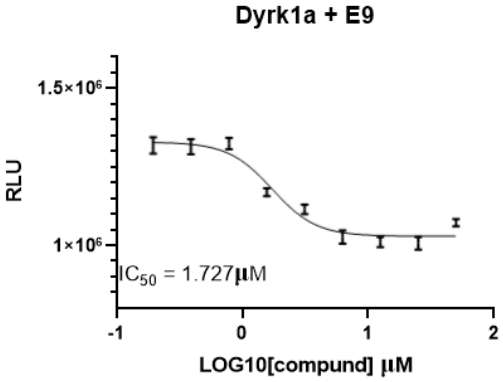
E02	0	97.97				
F02	-0.17	91.84				
A03	0.33	141.96				
B03	0.17	127.72				
C03	-0.17	133.15				
D03	0	131.08				
F03	0.17	97.04				
G03	0	91.72				
H03	0.17	95.8				
A04	0.5	145.66				
B04	0	139.36				
C04	-0.17	129.79				
D04	-0.17	110.52				
E04	-0.17	105.06				
F04	-3	101.92				
A05	0.5	117.6				
B05	0.5	139.31				
C05	-1.5	133.86				
D05	-1.5	129.89				
E05	-0.5	103.15				
F05	-0.5	100.53				
G05	0.17	101.05				
H05	-7.83	91.55				
A06	0.17	117.98				
B06	0.5	112.45				
C06	0.5	99.43				
A07	0.5	112.4				
B07	-0.83	106.95				
C07	-1.5	105.12				
D07	1.33	97.66				
E07	0.33	96.58				
F07	0.67	91.45				
G07	0	93.16				
H07	-3.33	92.24				
A08	0.5	107.25				
B08	-0.5	106.33				
C08	-2.33	100.03				

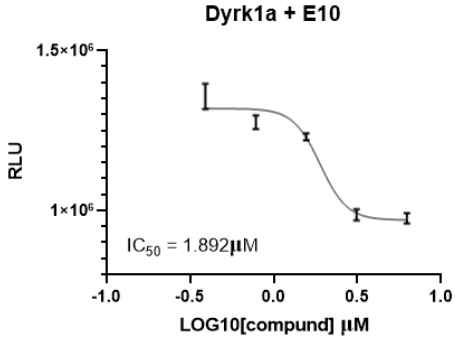
E08	-0.83	91.09				
F08	0.17	90.39				
G08	0.67	90.86				
H08	0.33	90.9				
A09	0.67	102.83				
B09	-0.67	116.44				
C09	-1.17	104.88				
F09	-0.5	94.23				
H09	-0.5	93.06				
A10	0.17	113.81				
B10	2.5	90.71				
C10	-1.67	103.02				
D10	-0.5	97.17				
F10	0	90.26				
G10	-0.67	94.48				
H10	1.83	91.05				
A11	-11.33	94				
B11	1	117.92				
C11	0	110.76				
D11	-0.5	99.32				
F11	0	103.64				
G11	-1	93.19				
H11	0	92.39				
A12	-0.67	93.16				
B12	0.5	117.21				
C12	-0.5	115				
D12	0.5	104.39				
F12	0.33	93.16				

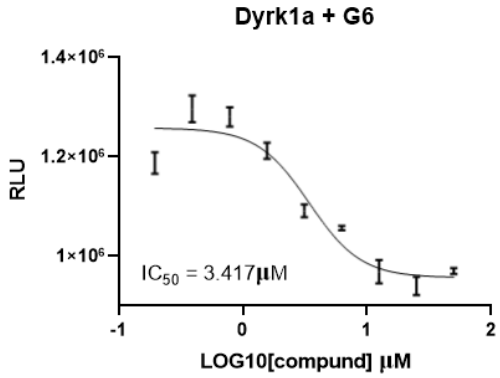
B.


Dyrk1a + D8
RLU
1.5×10⁶
1×10⁶
5×10⁵
IC50 = 1.924μM
-1
0
1
2
LOG10[compund] μM


Dyrk1a + D9
RLU
1.5×10⁶
1×10⁶
5×10⁵
IC50 = 1.118μM
-1
0
1
2
LOG10[compund] μM


Dyrk1a + E9
RLU
1.5×10⁶
1×10⁶
IC50 = 1.727μM
-1
0
1
2
LOG10[compund] μM


Dyrk1a + E10
RLU
1.5×10⁶
1×10⁶
IC50 = 1.892μM
-1.0
-0.5
0.0
0.5
1.0
LOG10[compund] μM


Dyrk1a + G6
RLU
1.4×10⁶
1.2×10⁶
1×10⁶
IC50 = 3.417μM
-1
0
1
2
LOG10[compund] μM

Supplementary material figure 4.2 results of screening test of AW AI-CADD small molecules against to Dyrk1a using TSA and ADP assay. AI-CADD small molecules were tested against to Dyrk1a along with DMSO as negative control in triplicates. A. The average of % activity and dosage trend in the ADP assay determined whether the small molecule was an inhibitor of Dyrk1a. B. IC_{50} values were measured for the inhibitors AW D8, D9, E9, E10, and G6.

Dyrk1a DTP small molecule	% activity	AMP best fit IC_{50} μM	AMP 95% IC_{50} μM	AMP IC_{50} R squared
C01	4.80	0.036	0.02189 to 0.05799	0.93
B06	1.36	0.092	0.07240 to 0.1187	0.97
C02	16.55	0.239	0.09719 to ???	0.94
C07	20.12	1.275	0.6747 to 2.549	0.98
D01	21.78	2.058	1.489 to 12.60	0.95
C06	4.75	2.258	1.230 to ???	0.91
A07	25.88	2.335	1.216 to 3.945	0.93
C03	16.51	2.370	???	0.69
C08	23.70	2.406	???	0.59
C10	21.33	2.582	2.145 to ???	0.80
A05	27.30	3.800	0.9729 to 5.843	0.92
C04	12.38	4.657	1.045 to ???	0.71
B03	4.62	>50		
C11	19.92	>50		
A03	22.40	>50		
A08	24.87	>50		
B11	34.74			
C09	37.87			
B01	50.91			
B05	51.56			
D04	54.94			
C05	59.61			
C12	59.66			
D03	63.60			
B02	66.08			
A12	69.83			
B12	79.09			
B04	81.19			
A10	85.01			
B09	85.94			
A11	86.36			
A06	89.06			

B10	89.08			
D02	91.03			
B08	94.04			
A09	94.31			
B07	96.87			
A02	103.32			
A04	105.50			
A01	111.53			
DMSO	100.00			

Supplementary material table 4.3 results of screening test of DTP small molecules against to Dyrk1a using ADP assay. DTP small molecules were tested against to Dyrk1a along with DMSO as negative control in triplicates. The average of % activity and dosage trend in the ADP assay determined whether the small molecule was an inhibitor of Dyrk1a. IC_{50} values were measured for the inhibitors.

Chapter 5
Discussion

Chapter 5: Discussion

5.1 summary of early drug discovery

The early drug discovery starts from screening compounds, identifying the target's best leads, and then moving to the pre-clinic phase. Drug discovery requires a huge amount of time, investment, and collaboration. Dedicated research and the best candidate for each step are required to improve the success of a project.

Target identification and validation is the first step and is accomplished by basic research. Therefore, selecting the targets that are druggable and there is a market need for a new therapy is the first task in starting an early drug discovery research program. SBDD and FBLD rely on purified protein (to form crystals and test with compounds), so obtaining a large amount of purified proteins by FPLC were the first experimental steps. Then, the screening libraries were developed for each project. We used drug-like, covalent, and brominate fragments to ensure structural diversity and drug-like properties. We also collaborated with Atomwise and utilized their AI-CADD virtual screening. We used TSA as the primary method in hit ID because it is rapid, low-cost, and requires minimal purified protein. Compounds with positive ΔT_m were determined as hits, as ΔT_m indicated binding and positive ΔT_m indicated binding and stabilization of the protein, and compounds that can bind and stabilize protein may led to greater chance of success in crystallization studies. TSA measuring the T_m of the target protein and can identify the change of Tm of the target protein when it binds to a compound. We always performed a dosage/concentration test to confirm the hits (binders at this stage) not a false positive.

SBDD and FBLD also heavily rely on X-ray crystallography or other structural techniques. Once we identified the hits (binders), we used X-ray crystallography to detect and visualize the binding site and pattern, as this step can be critical in optimizing hits to leads. We collaborated with Atomwise and the Molbey Lab at UCI, to utilize the AI-CADD and molecule docking, respectively, for hits-to-leads. To ensure that the hits (binders) affected the target protein's function, we also used *in vitro* biochemical assays, confirming that the hits not only bind to target proteins but also reduce their biochemical activity. To summarize the early drug discovery steps in the Perry Lab, we selected target proteins and purified them by FPLC. We selected test libraries and screened for hits by TSA. We detect and visualize the binding between hits and proteins by X-ray crystallography. We confirm hits inhibiting proteins by *in vitro* assays, potentially binding assays, and improve hits by virtual docking. We did not rule out the compound with negative ΔT_m. Negative ΔT_m binders and protein co-crystallization might be hard to achieve. However, binding can be detected by other methods, such as NMR, ITC, SPR or cryo-EM if necessary.

5.2 SUMO E1 Uba2/Aos1

The first challenge for the SUMO E1 Uba2/Aos1 project was to increase the expression level, which was poor at 1 mg per 12 L, to greatly enable future screens and *in vitro* studies. We tested all *E. coli* cell strains in the Perry Lab, to optimize the expression and we also looked at changing the methods from co-expressing Uba2 and Aos1 to individual expression in case that improved overall yields. The expression level of Aos1 was much higher than Uba2, when co-expressed. Since Uba2 is a relatively large protein of 74 kDa molecular weight (and runs as 100 kDa on SDS-PAGE gels) and Aos1 with the GST tag is

only 60 kDa, expression of Aos1 may be easier for the *E. coli* system than Uba2. Fortunately, our data showed that Uba2 could be expressed without Aos1, with a significant improvement in the expression level of 1.04 mg per L of *E. coli* culture. Uba2 and Aos1 were purified individually by affinity and ion exchange. Then Uba2 and Aos1 were mixed and purified by size exclusion and high-resolution ion exchange to reach higher than 90% purity Some sample is lost during each column step, but purity is important, especially for the structural studies. The yield was 1.8 mg of Uba2/Aos1 per L of *E. coli* culture. For other SUMOylation pathway proteins involved in the Uba2/Aos1 project, i.e.: SUMO1, SUMO2, and RanGap1, the expression and purification steps were more straightforward than Uba2/Aos1.

TSA studies on Uba2/Aos2 was successful. Uba2/Aos2 is a heterodimer, so the reaction buffer screening was completed with extra attention paid to obtaining one T_m but not two as observed by a former member in the Perry lab. The reaction buffer that could yield one peak/T_m was selected, and this T_m of Uba2/Aos1 complex was the same as T_m of Uba2. Uba2/Aos1 complex was used for all tests. During the TSA studies, Uba2/Aos1 was tested with drug-like, covalent (aiming for the activity site Cys 173), brominated fragments, and virtual screening DTP small molecules (focused on the ATP binding pocket). Notably, the covalent fragment hit LC 2A showed significantly reduced SUMOylation on RanGap1 by western blot and 30.21 μM IC_{50} by the AMP assay. LC 2A showed the results as a promising hit, and it is worth pursuing this further in our hits-to-leads study. LC 2A will need to be developed into a larger small molecule, such as through the FBDD linking, growing, or merging approaches.

The mechanism of covalent inhibitor has been described in the following equation (293):

$E + I \rightleftharpoons E \cdot I \rightleftarrows E - I$

E is the enzyme, and I is the inhibitor. E・I is the initial non-covalent complex. This step places the enzyme and the covalent inhibitor close enough, more specifically, the nucleophile of the protein and the electrophile of the covalent inhibitor. Therefore, the covalent bond forms and generates the E – I covalent complex. The parameters govern the covalent inhibitor is k_2, which is the affinity of the covalent complex. Experimentally, the bond affinity of covalent inhibitors can be measured by surface plasmon resonance (SPR). A chip interacts with the affinity tag of the target protein. Then the covalent inhibitor will bind to the protein and then can be washed off by buffers. SPR can measure the mass change when the covalent inhibitor binds to the protein and when the inhibitor is washed off in a real-time manner. Therefore, the binding affinity (K_{inact}/K_i) can be measured and to evaluate the covalent inhibitor.

Most of the DTP compounds bind to SYPRO Orange, so they were tested by the AMP assay and the HTRF assay. DTP C7 showed inhibition of Uba2/Aos1 by the AMP assay (18.41 μM IC_{50}) and the HTRF assay (8.01 μM IC_{50}) and promising IC_{50} from both assays. Another promising compound was DTP A7 (35.37 μM IC_{50} in AMP assay and 1.01 μM IC_{50} in HTRF assay). X-ray crystallography work will be the next step for these compounds that showed promising results in the hit ID stage, to define the binding site interactions. Therefore, further docking can be done to identify SAR by catalog analogs that can potentially improve binding affinity.

AW2 (docked against the ATP binding pocket by Atomwise) compounds were derived from a hit identified by a former member in the Perry Lab. The hit was a negative shifter in TSA and showed inhibition of Uba2/Aos1 in HTRF. All compounds in AW2 were tested by TSA and AMP assay. AW2 F5 with an IC_{50} of only 497.7 nM in the AMP assay had been the best inhibitor during the research on Uba2/Aos1. AW2 F5 was a positive shifter in TSA, indicating it bind and stabilized Uba2/Aos1. Notably, the co-crystal of Uba2/Aos1 and AW2 F5 was formed. Although the crystal was still tiny, optimization of the crystal condition would be done in the future SAR study.

As Uba2/Aos1 now has promising fragment hits, small molecule hits, the compound in the hit-to-lead process, and finally forms preliminary crystals, it is worth continuing to work on this project.

5.3 SUMO2 Ubc9

For the SUMO2 Ubc9 project, the biggest challenge is that the activity site Cys 93 is located on the surface but and not in a pocket, making it harder to target Cys 93. Nevertheless, we tried to test Ubc9 with covalent fragments, aiming to screen for any fragment that could covalently bond to Cys 93. However, mass spectrometry showed that fragment hit LC B3 bind to Cys or Lys at the backside of Ubc9, far from Cys 93. We also always know that Ubc9 also functions in DNA repair by the backside, forming a noncovalent complex with SLD2, so LC B3 has the potential to interrupt the DNA repair function. The AUTAC strategy is a promising method to convert LC B3 to a degrader. To determine the efficiency of a degrader, a DC_{50} (the half of maximal degrader concentration

that causes the target protein 50% degradation) will be measured by i.e., *in vitro* p62 oligomerization.

5.4 SUMO E3 Pias1

For the SUMO E3 Pias1 project, expression was very challenging. After optimization, the expression level was 2 mg of Pias1 per 12 L of *E. coli* culture, but the improvement was not as good as Uba2/Aos1. Therefore, we had to spend a large amount of time on protein work. During the size exclusion column, we noticed that Pias1 might have a multimeric form and a monomer form. Pias1 yield one T_m with pH 6 buffer and two T_ms with other pH but not 6 buffers by TSA. This suggested that protonation of His residue(s) (pK_a 6.0) mediated multimerization and it was pH dependent. We tried to screen the LC fragments against it. Pias1 was druggable; only the hit rate of Pias1 was the lowest among all SUMO projects at 2.48%. If we want to keep the Pias1 project, a future optimization of protein expression should be done first. A higher yield is required to support further tests. For example, for structure studies, X-ray crystallography requires a lot of purified Pias1. Considering the molecular weight of Pias1 (75 kDa), NMR will not be suitable, but cryo-EM can also be tried.

5.5 Dyrk1a

Dyrk1a is the project that has progressed the furthest. The hit ID screening started using AI-CADD small molecules aiming for the ATP binding pocket), leading to a relatively high hit rate (21.9%) by TSA. We used the ADP assay to test the hits. We first used DYRKtide as the substrate. We showed that purified Dyrk1a had enzymatic activity, but more troubleshooting had to be completed to achieve a better signal/background ratio.

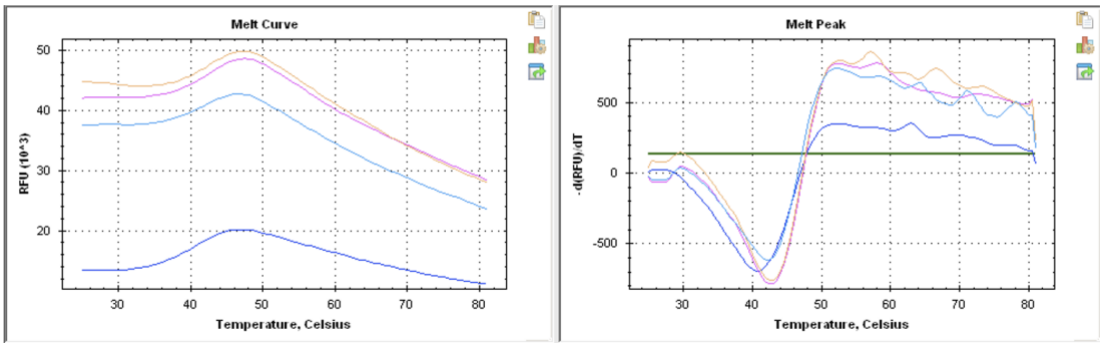
Luckily, Dyrk1a can autophosphorylate itself, and the signal/background ratio (11.06) was ideal after the first test. The best compound was AW H4 with 0.669 μM IC_{50} by the AMP assay. It also formed co-crystals with Dyrk1a. As a hit, AW H4 can move to the next step.

Recently, molecular docking was finished (DTP compounds from the Molbey Lab, UCI), and the best IC_{50} was 35.7 nM from DTP C1, which was a very promising result at the hits stage of a hit-to-lead study. The next step for Dyrk1a will be X-ray crystallography to visualize the binding between Dyrk1a and AW H4 or DTP C1, and further expand for lead and lead optimization. The initial co-crystals of Dyrk1a, AW H4, and D9 will be optimized for bigger-size crystals. Previously, we have had co-crystals of Dyrk1a and abemacicle, and the crystals had good size and morphology. Abemacicle can be potentially removed by wash steps and then the crystals can be soaked with our compounds in the future. Alternatively, much higher concentrations of our compound can be added to the crystallization drop after abemacicle Dyrk1A crystals have been grown to compete out abemacicle with the compound of interest. A binding study will be completed in the near-term to obtain affinity *in vitro*, and we will then obtain compound potency in cells. Meanwhile, due to the DYKR members sharing ATP binding pocket, testing the compounds with Dyrk1b or other DYRK family members for selectivity is essential. Overall, with Atomwise and the Mobley Lab, Dyrk1a hits are encouraging and will be taken through our hits-to-lead pipeline for lead generation for future DMPK-focused studies.

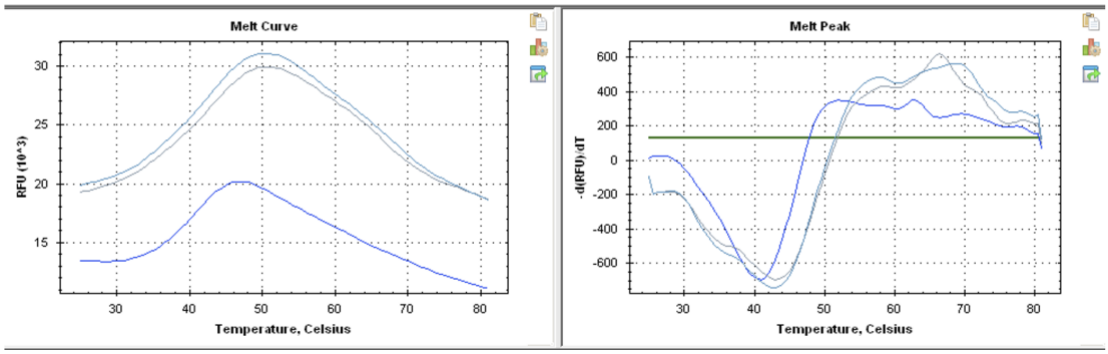
APPENDICES

Appendix A: Melt Curve and Derivative plots for Supplementary material table 2.1.

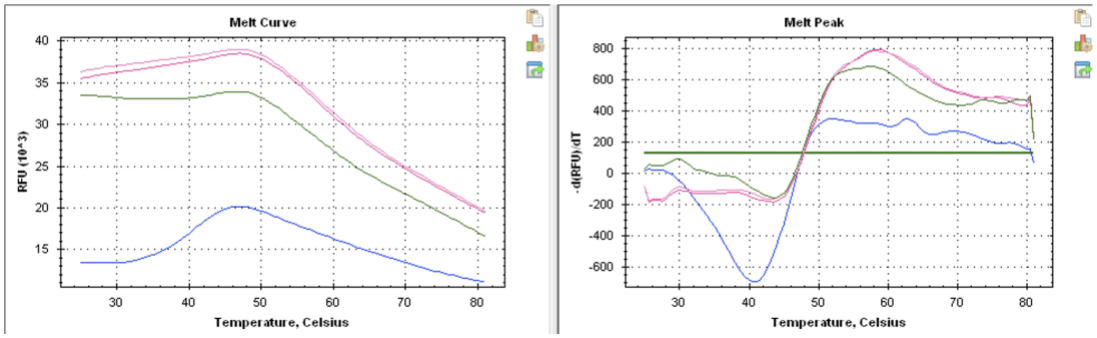
Uba2 + LC 2E
ΔT_m = 1°C



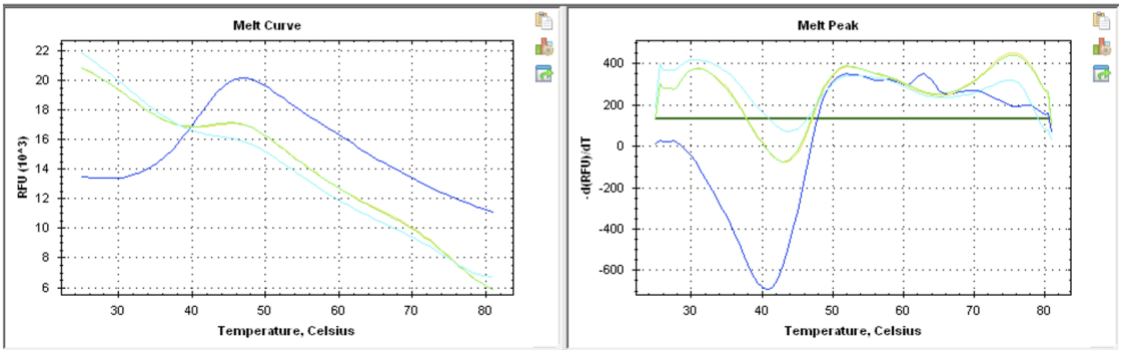
Uba2 + LC 2A
ΔT_m = 2°C



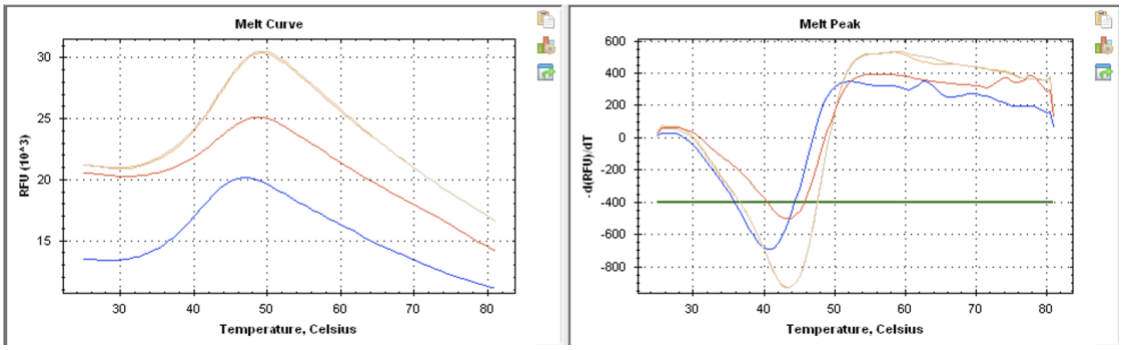
Uba2 + LC 1H
ΔT_m = 2°C



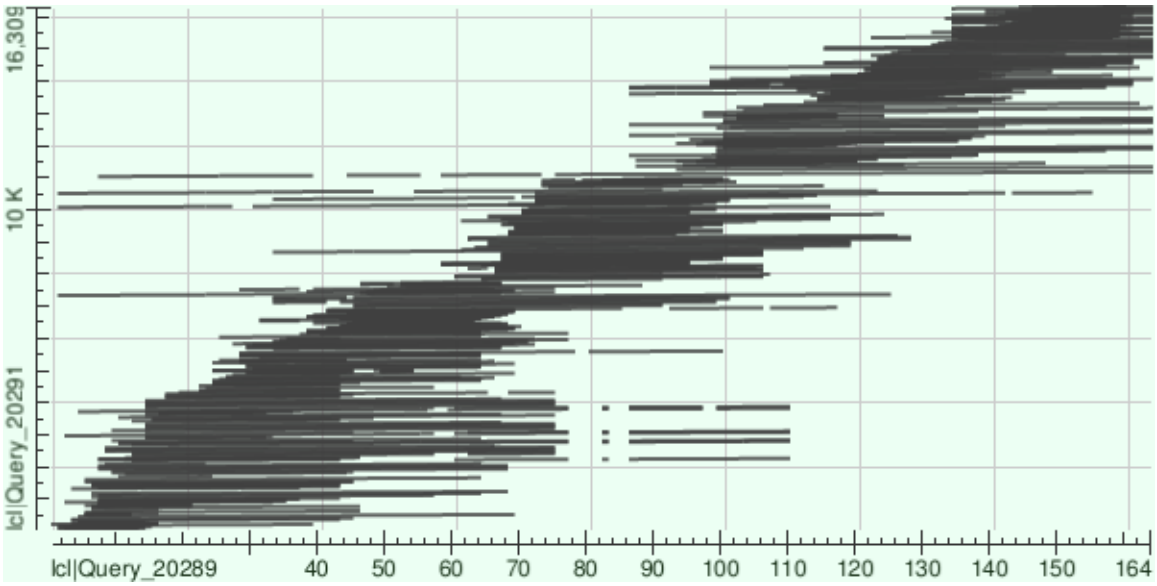
Uba2 + LC 1E
$\Delta T_m = 1.5°C$

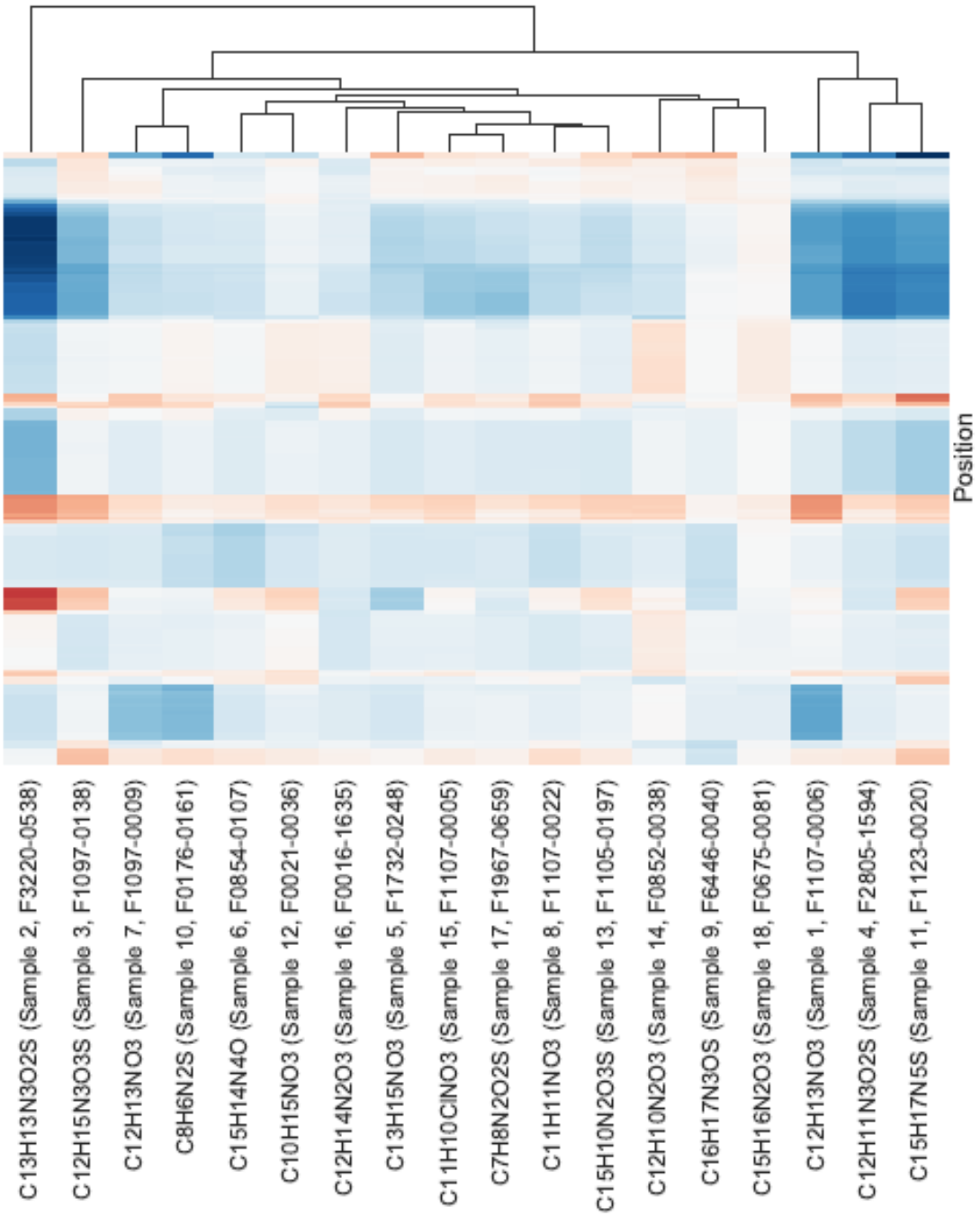


Uba2 + LC 1C
$\Delta T_m = 1.5°C$



Appendix B: Ubc9 peptides for Figure 3.4




C13H13N3O2S (Sample 2, F3220-0538)
C12H15N3O3S (Sample 3, F1097-0138)
C12H13NO3 (Sample 7, F1097-0009)
C8H6N2S (Sample 10, F0176-0161)
C15H14N4O (Sample 6, F0854-0107)
C10H15NO3 (Sample 12, F0021-0036)
C12H14N2O3 (Sample 16, F0016-1635)
C13H15NO3 (Sample 5, F1732-0248)
C11H10ClNO3 (Sample 15, F1107-0005)
C7H8N2O2S (Sample 17, F1967-0659)
C11H11NO3 (Sample 8, F1107-0022)
C15H10N2O3S (Sample 13, F1105-0197)
C12H10N2O3 (Sample 14, F0852-0038)
C16H17N3OS (Sample 9, F6446-0040)
C15H16N2O3 (Sample 18, F0675-0081)
C12H13NO3 (Sample 1, F1107-0006)
C12H11N3O2S (Sample 4, F2805-1594)
C15H17N5S (Sample 11, F1123-0020)
Position

REFERENCES

[1] Hughes, JP, et al. "Principles of Early Drug Discovery." *British Journal of Pharmacology*, vol. 162, no. 6, Feb. 2011, pp. 1239–49, https://doi.org/10.1111/j.1476-5381.2010.01127.x.

[2] Sun, Duxin, et al. "Why 90% of Clinical Drug Development Fails and How to Improve It?" *Acta Pharmaceutica Sinica B*, vol. 12, no. 7, July 2022, pp. 3049–62, https://doi.org/10.1016/j.apsb.2022.02.002. Accessed 31 July 2023.

[3] Takebe, Tohru, et al. "The Current Status of Drug Discovery and Development as Originated in United States Academia: The Influence of Industrial and Academic Collaboration on Drug Discovery and Development." *Clinical and Translational Science*, vol. 11, no. 6, July 2018, pp. 597–606, https://doi.org/10.1111/cts.12577. Accessed 18 Nov. 2019.

[4] Bissaro, Maicol, et al. "The Rise of Molecular Simulations in Fragment-Based Drug Design (FBDD): An Overview." *Drug Discovery Today*, June 2020, https://doi.org/10.1016/j.drudis.2020.06.023.

[5] Bleicher, Konrad H., et al. "Hit and Lead Generation: Beyond High-Throughput Screening." *Nature Reviews Drug Discovery*, vol. 2, no. 5, May 2003, pp. 369–78, https://doi.org/10.1038/nrd1086. Accessed 28 June 2019.

[6] Ciulli, Alessio. "Biophysical Screening for the Discovery of Small-Molecule Ligands." *Protein-Ligand Interactions*, 2013, pp. 357–88, https://doi.org/10.1007/978-1-62703-398-5_13.

[7] Lipinski, Christopher A., et al. "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings." *Advanced Drug Delivery Reviews*, vol. 23, no. 1-3, Jan. 1997, pp. 3–25, https://doi.org/10.1016/s0169-409x(96)00423-1.

[8] Tijjani, Habibu, et al. "Chapter 14 - in Silico Insight into the Interaction of 4-Aminoquinolines with Selected SARS-CoV-2 Structural and Nonstructural Proteins." *ScienceDirect*, edited by Chukwuebuka Egbuna, Elsevier, 1 Jan. 2022, pp. 313–33, www.sciencedirect.com/science/article/abs/pii/B9780323955782000017.

[9] Flaherty, Keith. T., et al. "Vemurafenib." *Nature Reviews Drug Discovery*, vol. 10, no. 11, Oct. 2011, pp. 811–12, https://doi.org/10.1038/nrd3579. Accessed 14 May 2020.

[10] Oltersdorf, Tilman, et al. "An Inhibitor of Bcl-2 Family Proteins Induces Regression of Solid Tumours." *Nature*, vol. 435, no. 7042, June 2005, pp. 677–81, https://doi.org/10.1038/nature03579.

[11] Perera, Timothy P. S., et al. “Discovery and Pharmacological Characterization of JNJ-42756493 (Erdafitinib), a Functionally Selective Small-Molecule FGFR Family Inhibitor.” *Molecular Cancer Therapeutics*, vol. 16, no. 6, Mar. 2017, pp. 1010–20, https://doi.org/10.1158/1535-7163.mct-16-0589. Accessed 4 Feb. 2020.

[12] Benner, Brooke, et al. “Pexidartinib, a Novel Small Molecule CSF-1R Inhibitor in Use for Tenosynovial Giant Cell Tumor: A Systematic Review of Pre-Clinical and Clinical Development.” *Drug Design, Development and Therapy*, vol. Volume 14, May 2020, pp. 1693–704, https://doi.org/10.2147/dddt.s253232. Accessed 17 Dec. 2020.

[13] Ostrem, Jonathan M., et al. “K-Ras(G12C) Inhibitors Allosterically Control GTP Affinity and Effector Interactions.” *Nature*, vol. 503, no. 7477, Nov. 2013, pp. 548–51, https://doi.org/10.1038/nature12796.

[14] Schoepfer, Joseph, et al. “Discovery of Asciminib (ABL001), an Allosteric Inhibitor of the Tyrosine Kinase Activity of BCR-ABL1.” *Journal of Medicinal Chemistry*, vol. 61, no. 18, Aug. 2018, pp. 8120–35, https://doi.org/10.1021/acs.jmedchem.8b01040.

[15] Mureddu, Luca G., and Geerten W. Vuister. “Fragment-Based Drug Discovery by NMR. Where Are the Successes and Where Can It Be Improved?” *Frontiers in Molecular Biosciences*, vol. 9, Feb. 2022, https://doi.org/10.3389/fmolb.2022.834453. Accessed 10 May 2022.

[16] Shuker, S. B., et al. “Discovering High-Affinity Ligands for Proteins: SAR by NMR.” *Science*, vol. 274, no. 5292, Nov. 1996, pp. 1531–34, https://doi.org/10.1126/science.274.5292.1531.

[17] Congreve, Miles, et al. “A ‘Rule of Three’ for Fragment-Based Lead Discovery?” *Drug Discovery Today*, vol. 8, no. 19, Oct. 2003, pp. 876–77, https://doi.org/10.1016/s1359-6446(03)02831-9.

[18] Hann, Michael M., et al. “Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery.” *Journal of Chemical Information and Computer Sciences*, vol. 41, no. 3, May 2001, pp. 856–64, https://doi.org/10.1021/ci000403i. Accessed 3 Dec. 2019.

[19] Erlanson, Daniel A., et al. “Twenty Years On: The Impact of Fragments on Drug Discovery.” *Nature Reviews Drug Discovery*, vol. 15, no. 9, July 2016, pp. 605–19, https://doi.org/10.1038/nrd.2016.109. Accessed 2 Dec. 2021.

[20] Troelsen, Nikolaj S., and Mads H. Clausen. “Library Design Strategies to Accelerate Fragment-Based Drug Discovery.” *Chemistry – a European Journal*, vol. 26, no. 50, July 2020, pp. 11391–403, https://doi.org/10.1002/chem.202000584. Accessed 14 Mar. 2023.

[21] Barf, Tjeerd, and Allard Kaptein. “Irreversible Protein Kinase Inhibitors: Balancing the Benefits and Risks.” *Journal of Medicinal Chemistry*, vol. 55, no. 14, June 2012, pp. 6243–62, https://doi.org/10.1021/jm3003203. Accessed 29 July 2020.

[22] Lonsdale, Richard, and Richard A. Ward. “Structure-Based Design of Targeted Covalent Inhibitors.” *Chemical Society Reviews*, vol. 47, no. 11, 2018, pp. 3816–30, https://doi.org/10.1039/c7cs00220c.

[23] Nakayama, Shintaro, et al. “A Zone Classification System for Risk Assessment of Idiosyncratic Drug Toxicity Using Daily Dose and Covalent Binding.” *Drug Metabolism and Disposition: The Biological Fate of Chemicals*, vol. 37, no. 9, Sept. 2009, pp. 1970–77, https://doi.org/10.1124/dmd.109.027797. Accessed 11 May 2020.

[24] Baell, Jonathan B., and J. Willem M. Nissink. “Seven Year Itch: Pan-Assay Interference Compounds (PAINS) in 2017—Utility and Limitations.” *ACS Chemical Biology*, vol. 13, no. 1, Dec. 2017, pp. 36–44, https://doi.org/10.1021/acschembio.7b00903. Accessed 21 May 2020.

[25] Lamoree, Bas, and Roderick E. Hubbard. “Current Perspectives in Fragment-Based Lead Discovery (FBLD).” *Essays in Biochemistry*, edited by Rob L.M. van Montfort and Paul Workman, vol. 61, no. 5, Nov. 2017, pp. 453–64, https://doi.org/10.1042/ebc20170028. Accessed 18 May 2021.

[26] Howard, Steven, et al. “Fragment-Based Discovery of the Pyrazol-4-Yl Urea (AT9283), a Multitargeted Kinase Inhibitor with Potent Aurora Kinase Activity†.” *Journal of Medicinal Chemistry*, vol. 52, no. 2, Jan. 2009, pp. 379–88, https://doi.org/10.1021/jm800984v. Accessed 20 Apr. 2020.

[27] Singh, Meenakshi, et al. “NMR-Fragment Based Virtual Screening: A Brief Overview.” *Molecules*, vol. 23, no. 2, Jan. 2018, p. 233, https://doi.org/10.3390/molecules23020233. Accessed 4 Aug. 2023.

[28] Hubbard, R. E. “Fragment Approaches in Structure-Based Drug Discovery.” *Journal of Synchrotron Radiation*, vol. 15, no. 3, May 2008, pp. 227–30, https://doi.org/10.1107/S090904950705666X. Accessed 31 July 2022.

[29] Fox, Sandra, et al. “High-Throughput Screening: Update on Practices and Success.” *Journal of Biomolecular Screening*, vol. 11, no. 7, Oct. 2006, pp. 864–69, https://doi.org/10.1177/1087057106292473.

[30] Vázquez, Javier, et al. “Merging Ligand-Based and Structure-Based Methods in Drug Discovery: An Overview of Combined Virtual Screening Approaches.” *Molecules*, vol. 25, no. 20, Jan. 2020, p. 4723, https://doi.org/10.3390/molecules25204723.

[31] Dilip, Athira, et al. “Ligand-Based Virtual Screening Interface between PyMOL and LiSiCA.” *Journal of Cheminformatics*, vol. 8, no. 1, Sept. 2016, https://doi.org/10.1186/s13321-016-0157-z. Accessed 8 Nov. 2020.

[32] Zheng, Minghao, et al. “LBVS: An Online Platform for Ligand-Based Virtual Screening Using Publicly Accessible Databases.” *Molecular Diversity*, vol. 18, no. 4, Sept. 2014, pp. 829–40, https://doi.org/10.1007/s11030-014-9545-3. Accessed 22 Aug. 2019.

[33] Lo, Mei-Chu, et al. “Evaluation of Fluorescence-Based Thermal Shift Assays for Hit Identification in Drug Discovery.” *Analytical Biochemistry*, vol. 332, no. 1, Sept. 2004, pp. 153–59, https://doi.org/10.1016/j.ab.2004.04.031. Accessed 31 Mar. 2021.

[34] Chilton, Molly, et al. *Hot-Spotting with Thermal Scanning: A Ligand- and Structure-Independent Assessment of Target Ligandability*. no. 12, June 2017, pp. 4923–31, https://doi.org/10.1021/acs.jmedchem.7b00208. Accessed 21 July 2023.

[35] Wiseman, Thomas, et al. “Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter.” *Analytical Biochemistry*, vol. 179, no. 1, May 1989, pp. 131–37, https://doi.org/10.1016/0003-2697(89)90213-3. Accessed 25 Mar. 2021.

[36] Berrouet, Catherine, et al. “Comparison of Drug Inhibitory Effects (IC50) in Monolayer and Spheroid Cultures.” *Bulletin of Mathematical Biology*, vol. 82, no. 6, June 2020, https://doi.org/10.1007/s11538-020-00746-7.

[37] Yung-Chi, Cheng, and William H. Prusoff. “Relationship between the Inhibition Constant (KI) and the Concentration of Inhibitor Which Causes 50 per Cent Inhibition (I50) of an Enzymatic Reaction.” *Biochemical Pharmacology*, vol. 22, no. 23, Dec. 1973, pp. 3099–108, https://doi.org/10.1016/0006-2952(73)90196-2.

[38] Alizadeh, Seyedeh Roya, et al. “Drug Design Strategies That Aim to Improve the Low Solubility and Poor Bioavailability Conundrum in Quercetin Derivatives.” *Expert Opinion on Drug Discovery*, July 2023, pp. 1–16, https://doi.org/10.1080/17460441.2023.2241366. Accessed 31 July 2023.

[39] Vemula, Divya, et al. “CADD, AI and ML in Drug Discovery: A Comprehensive Review.” *European Journal of Pharmaceutical Sciences*, vol. 181, Feb. 2023, p. 106324, https://doi.org/10.1016/j.ejps.2022.106324. Accessed 5 Aug. 2023.

[40] Puri, Munish, et al. “Artificial Neural Network for Drug Design, Delivery and Disposition - 1st Edition.” *Shop.elsevier.com*, shop.elsevier.com/books/artificial-neural-network-for-drug-design-delivery-and-disposition/puri/978-0-12-801559-9.

[41] Mak, Kit-Kay, et al. “The Role of DMPK Science in Improving Pharmaceutical Research and Development Efficiency.” *Drug Discovery Today*, Nov. 2021, https://doi.org/10.1016/j.drudis.2021.11.005.

[42] Teorell, Torsten. “Kinetics of Distribution of Substances Administered to the Body, I: The Extravascular Modes of Administration.” *DIVA*, 1937, www.diva-portal.org/smash/record.jsf?pid=diva2%3A513484&dswid=1397. Accessed 21 Aug. 2023.

[43] Nelson, Eino. “Kinetics of Drug Absorption, Distribution, Metabolism, and Excretion.” *Journal of Pharmaceutical Sciences*, vol. 50, no. 3, Mar. 1961, pp. 181–92, https://doi.org/10.1002/jps.2600500302.

[44] Doogue, Matthew P., and Thomas M. Polasek. “The ABCD of Clinical Pharmacokinetics.” *Therapeutic Advances in Drug Safety*, vol. 4, no. 1, Jan. 2013, pp. 5–7, https://doi.org/10.1177/2042098612469335.

[45] Grogan, Sean, and Charles V. Preuss. “Pharmacokinetics.” *PubMed*, StatPearls Publishing, 2022, www.ncbi.nlm.nih.gov/books/NBK557744/#:~:text=This%20is%20closely%20related%20to.

[46] Markovic, Milica, et al. “Prodrugs for Improved Drug Delivery: Lessons Learned from Recently Developed and Marketed Products.” *Pharmaceutics*, vol. 12, no. 11, Oct. 2020, p. 1031, https://doi.org/10.3390/pharmaceutics12111031.

[47] Guengerich, F. Peter. “Mechanisms of Drug Toxicity and Relevance to Pharmaceutical Development.” *Drug Metabolism and Pharmacokinetics*, vol. 26, no. 1, 2011, pp. 3–14, www.ncbi.nlm.nih.gov/pmc/articles/PMC4707670/.

[48] Zou, Huixi, et al. “Application of Pharmacokinetic-Pharmacodynamic Modeling in Drug Delivery: Development and Challenges.” *Frontiers in Pharmacology*, vol. 11, July 2020, https://doi.org/10.3389/fphar.2020.00997. Accessed 31 July 2023.

[49] Kwon, S. J. “Proteomics Studies of Post-Translational Modifications in Plants.” *Journal of Experimental Botany*, vol. 57, no. 7, Mar. 2006, pp. 1547–51, https://doi.org/10.1093/jxb/erj137.

[50] Hashiguchi, Akiko, and Setsuko Komatsu. “Proteomics of Soybean Plants.” *Semantic Scholar*, Elsevier, 2017, pp. 89–105, https://doi.org/10.1016/B978-0-12-804007-2.00006-0.

[51] Beltrao, Pedro, et al. “Evolution and Functional Cross-Talk of Protein Post-Translational Modifications.” *Molecular Systems Biology*, vol. 9, no. 1, Jan. 2013, p. 714, https://doi.org/10.1002/msb.201304521.

[52] Tripodi, Farida, et al. “Post-Translational Modifications on Yeast Carbon Metabolism: Regulatory Mechanisms beyond Transcriptional Control.” *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1850, no. 4, Apr. 2015, pp. 620–27, https://doi.org/10.1016/j.bbagen.2014.12.010.

[53] Lau, Priscilla Nga Ieng, and Peter Cheung. “Elucidating Combinatorial Histone Modifications and Crosstalks by Coupling Histone-Modifying Enzyme with Biotin Ligase Activity.” *Nucleic Acids Research*, vol. 41, no. 3, Dec. 2012, pp. e49–49, https://doi.org/10.1093/nar/gks1247.

[54] Matunis, M. J., et al. “A Novel Ubiquitin-like Modification Modulates the Partitioning of the Ran-GTPase-Activating Protein RanGAP1 between the Cytosol and the Nuclear Pore Complex.” *Journal of Cell Biology*, vol. 135, no. 6, Dec. 1996, pp. 1457–70, https://doi.org/10.1083/jcb.135.6.1457.

[55] Chen, Ling, et al. “Regulating Tumor Suppressor Genes: Post-Translational Modifications.” *Signal Transduction and Targeted Therapy*, vol. 5, no. 1, June 2020, https://doi.org/10.1038/s41392-020-0196-9.

[56] Xue, Xiangfei, et al. “Emerging Role of Protein Post-Translational Modification in the Potential Clinical Application of Cancer.” *Nano LIFE*, vol. 10, no. 01n02, Mar. 2020, pp. 2040008–8, https://doi.org/10.1142/s1793984420400085.

[57] Molfetta, Rosa, et al. “SUMOylation and Related Post-Translational Modifications in Natural Killer Cell Anti-Cancer Responses.” *Cell Dev Biol*, vol. 11, no. 11:1213114., May 2023, https://doi.org/10.3389/fcell.2023.1213114.

[58] Mahajan, Rohit, et al. “A Small Ubiquitin-Related Polypeptide Involved in Targeting RanGAP1 to Nuclear Pore Complex Protein RanBP2.” *Cell*, vol. 88, no. 1, Jan. 1997, pp. 97–107, https://doi.org/10.1016/s0092-8674(00)81862-0.

[59] Saitoh, Hisato, and Joseph Hinchey. “Functional Heterogeneity of Small Ubiquitin-Related Protein Modifiers SUMO-1 versus SUMO-2/3.” *Journal of Biological Chemistry*, vol. 275, no. 9, Mar. 2000, pp. 6252–58, https://doi.org/10.1074/jbc.275.9.6252.

[60] Owerbach, David, et al. “A Proline-90 Residue Unique to SUMO-4 Prevents Maturation and Sumoylation.” *Biochemical and Biophysical Research Communications*, vol. 337, no. 2, Nov. 2005, pp. 517–20, https://doi.org/10.1016/j.bbrc.2005.09.090.

[61] Wang, Cong-Yi, et al. “Characterization of a Negative Feedback Network between SUMO4 Expression and NFκB Transcriptional Activity.” *Biochemical and Biophysical Research Communications*, vol. 381, no. 4, Apr. 2009, pp. 477–81, https://doi.org/10.1016/j.bbrc.2009.02.060.

[62] Bayer, Peter, et al. “Structure Determination of the Small Ubiquitin-Related Modifier SUMO-1.” *Journal of Molecular Biology*, vol. 280, no. 2, July 1998, pp. 275–86, https://doi.org/10.1006/jmbi.1998.1839.

[63] Kamitani, Tetsu, et al. “Preferential Modification of Nuclear Proteins by a Novel Ubiquitin-like Molecule.” *Journal of Biological Chemistry*, vol. 272, no. 22, May 1997, pp. 14001–4, https://doi.org/10.1074/jbc.272.22.14001. Accessed 16 Aug. 2023.

[64] Wang, Jun, et al. *Defective Sumoylation Pathway Directs Congenital Heart Disease*. no. 6, May 2011, pp. 468–76, https://doi.org/10.1002/bdra.20816. Accessed 18 June 2023.

[65] Wang, L., et al. “SUMO2 Is Essential While SUMO3 Is Dispensable for Mouse Embryonic Development.” *EMBO Reports*, vol. 15, no. 8, June 2014, pp. 878–85, https://doi.org/10.15252/embr.201438534.

[66] Wei, Wenzhong, et al. “A Stress-Dependent SUMO4 Sumoylation of Its Substrate Proteins.” *Biochemical and Biophysical Research Communications*, vol. 375, no. 3, Oct. 2008, pp. 454–59, https://doi.org/10.1016/j.bbrc.2008.08.028. Accessed 16 Aug. 2023.

[67] Liang, Ya-Chen, et al. “SUMO5, a Novel Poly-SUMO Isoform, Regulates PML Nuclear Bodies.” *Scientific Reports*, vol. 6, May 2016, https://doi.org/10.1038/srep26509.

[68] Yu, Bing, et al. “Oncogenesis Driven by the Ras/Raf Pathway Requires the SUMO E2 Ligase Ubc9.” *Proceedings of the National Academy of Sciences*, vol. 112, no. 14, Mar. 2015, https://doi.org/10.1073/pnas.1415569112. Accessed 11 Dec. 2022.

[69] Chang, Hui-Ming, and Edward T. H. Yeh. “SUMO: From Bench to Bedside.” *Physiological Reviews*, vol. 100, no. 4, Oct. 2020, pp. 1599–619, https://doi.org/10.1152/physrev.00025.2019.

[70] Geiss-Friedlander, Ruth, and Frauke Melchior. “Concepts in Sumoylation: A Decade On.” *Nature Reviews Molecular Cell Biology*, vol. 8, no. 12, Dec. 2007, pp. 947–56, https://doi.org/10.1038/nrm2293.

[71] Gareau, Jaclyn R., and Christopher D. Lima. “The SUMO Pathway: Emerging Mechanisms That Shape Specificity, Conjugation and Recognition.” *Nature Reviews Molecular Cell Biology*, vol. 11, no. 12, Nov. 2010, pp. 861–71, https://doi.org/10.1038/nrm3011.

[72] Hannoun, Zara, et al. “Post-Translational Modification by SUMO.” *Toxicology*, vol. 278, no. 3, Dec. 2010, pp. 288–93, https://doi.org/10.1016/j.tox.2010.07.013.

[73] Desterro, Joana M. P., et al. “Identification of the Enzyme Required for Activation of the Small Ubiquitin-like Protein SUMO-1.” *Journal of Biological Chemistry*, vol. 274, no. 15, Apr. 1999, pp. 10618–24, https://doi.org/10.1074/jbc.274.15.10618.

[74] Dohmen, R. Jürgen, et al. “An Essential Yeast Gene Encoding a Homolog of Ubiquitin-Activating Enzyme.” *Journal of Biological Chemistry*, vol. 270, no. 30, July 1995, pp. 18099–109, https://doi.org/10.1074/jbc.270.30.18099. Accessed 21 Aug. 2023.

[75] Lois, Luisa Maria, and Christopher D. Lima. “Structures of the SUMO E1 Provide Mechanistic Insights into SUMO Activation and E2 Recruitment to E1.” *The EMBO Journal*, vol. 24, no. 3, Jan. 2005, pp. 439–51, https://doi.org/10.1038/sj.emboj.7600552.

[76] Bossis, Guillaume, and Frauke Melchior. “Regulation of SUMOylation by Reversible Oxidation of SUMO Conjugating Enzymes.” *Molecular Cell*, vol. 21, no. 3, Feb. 2006, pp. 349–57, https://doi.org/10.1016/j.molcel.2005.12.019. Accessed 16 Aug. 2023.

[77] Giraud, M. F., et al. “Structure of Ubiquitin-Conjugating Enzyme 9 Displays Significant Differences with Other Ubiquitin-Conjugating Enzymes Which May Reflect Its Specificity for Sumo rather than Ubiquitin.” *Acta Crystallographica Section D Biological Crystallography*, vol. 54, no. 5, Sept. 1998, pp. 891–98, https://doi.org/10.1107/s0907444998002480. Accessed 24 June 2021.

[78] Tatham, Michael H., et al. “Unique Binding Interactions among Ubc9, SUMO and RanBP2 Reveal a Mechanism for SUMO Paralog Selection.” *Nature Structural & Molecular Biology*, vol. 12, no. 1, Dec. 2004, pp. 67–74, https://doi.org/10.1038/nsmb878.

[79] Nacerddine, Karim, et al. “The SUMO Pathway Is Essential for Nuclear Integrity and Chromosome Segregation in Mice.” *Developmental Cell*, vol. 9, no. 6, Dec. 2005, pp. 769–79, https://doi.org/10.1016/j.devcel.2005.10.007. Accessed 16 Aug. 2023.

[80] Mattoscio, Domenico, et al. “Autophagy Regulates UBC9 Levels during Viral-Mediated Tumorigenesis.” *PLOS Pathogens*, edited by Nancy Raab-Traub, vol. 13, no. 3, Mar. 2017, p. e1006262, https://doi.org/10.1371/journal.ppat.1006262. Accessed 9 Apr. 2023.

[81] Su, Yee-Fun, et al. “Phosphorylation of Ubc9 by Cdk1 Enhances SUMOylation Activity.” *PLoS ONE*, edited by Wenqing Xu, vol. 7, no. 4, Apr. 2012, p. e34250, https://doi.org/10.1371/journal.pone.0034250. Accessed 3 Nov. 2022.

[82] Rytinki, Miia M., et al. “PIAS Proteins: Pleiotropic Interactors Associated with SUMO.” *Cellular and Molecular Life Sciences*, vol. 66, no. 18, June 2009, pp. 3029–41, https://doi.org/10.1007/s00018-009-0061-z. Accessed 12 Mar. 2021.

[83] Reverter, David, and Christopher D. Lima. “Insights into E3 Ligase Activity Revealed by a SUMO–RanGAP1–Ubc9–Nup358 Complex.” *Nature*, vol. 435, no. 7042, June 2005, pp. 687–92, https://doi.org/10.1038/nature03588. Accessed 8 Mar. 2021.

[84] Hickey, Christopher M., et al. “Function and Regulation of SUMO Proteases.” *Nature Reviews Molecular Cell Biology*, vol. 13, no. 12, Nov. 2012, pp. 755–66, https://doi.org/10.1038/nrm3478.

[85] Ahner, Annette, et al. “Small Heat Shock Proteins Target Mutant Cystic Fibrosis Transmembrane Conductance Regulator for Degradation via a Small Ubiquitin-like Modifier–Dependent Pathway.” *Molecular Biology of the Cell*, vol. 24, no. 2, Jan. 2013, pp. 74–84, https://doi.org/10.1091/mbc.e12-09-0678.

[86] Da Silva-Ferrada, Elisa, et al. “Role of Monoubiquitylation on the Control of IκBα Degradation and NF-KB Activity.” *PLoSONE*, vol. 6, no. 10, Oct. 2011, pp. e25397–97, https://doi.org/10.1371/journal.pone.0025397.

[87] Prudden, John, et al. “SUMO-Targeted Ubiquitin Ligases in Genome Stability.” *The EMBO Journal*, vol. 26, no. 18, Nature Portfolio, Sept. 2007, pp. 4089–101, https://doi.org/10.1038/sj.emboj.7601838. Accessed 23 Apr. 2023.

[88] Perry, J. Jefferson P., et al. “A SIM-Ultaneous Role for SUMO and Ubiquitin.” *Trends in Biochemical Sciences*, vol. 33, no. 5, May 2008, pp. 201–8, https://doi.org/10.1016/j.tibs.2008.02.001. Accessed 27 Dec. 2021.

[89] Minty, Adrian, et al. *Covalent Modification of P73␣ by SUMO-1 TWO-HYBRID SCREENING with P73 IDENTIFIES NOVEL SUMO-1-INTERACTING PROTEINS and a SUMO-1 INTERACTION MOTIF**. JBC Papers in Press, 2000, https://doi.org/10.1074/jbc.M004293200. Accessed 16 Aug. 2023.

[90] Praefcke, Gerrit J. K., et al. “SUMO Playing Tag with Ubiquitin.” *Trends in Biochemical Sciences*, vol. 37, no. 1, Jan. 2012, pp. 23–31, https://doi.org/10.1016/j.tibs.2011.09.002.

[91] Bawa-Khalfe, T., and E. T. H. Yeh. “SUMO Losing Balance: SUMO Proteases Disrupt SUMO Homeostasis to Facilitate Cancer Development and Progression.” *Genes & Cancer*, vol. 1, no. 7, July 2010, pp. 748–52, https://doi.org/10.1177/1947601910382555.

[92] Heo, Kyung-Sun, et al. “Disturbed Flow-Activated P90RSK Kinase Accelerates Atherosclerosis by Inhibiting SENP2 Function.” *The Journal of Clinical Investigation*, vol. 125, no. 3, Mar. 2015, pp. 1299–310, https://doi.org/10.1172/JCI76453.

[93] Ferdaoussi, Mourad, et al. “Isocitrate-To-SENP1 Signaling Amplifies Insulin Secretion and Rescues Dysfunctional β Cells.” *Journal of Clinical Investigation*, vol. 125, no. 10, Sept. 2015, pp. 3847–60, https://doi.org/10.1172/jci82498.

[94] Hannoun, Zara, et al. “The Implication of SUMO in Intrinsic and Innate Immunity.” *Cytokine & Growth Factor Reviews*, vol. 29, no. 29:3-16, June 2016, pp. 3–16, https://doi.org/10.1016/j.cytogfr.2016.04.003.

[95] Henley, Jeremy M., et al. “Neuronal SUMOylation: Mechanisms, Physiology, and Roles in Neuronal Dysfunction.” *Physiological Reviews*, vol. 94, no. 4, Oct. 2014, pp. 1249–85, https://doi.org/10.1152/physrev.00008.2014.

[96] Qi, Yitao, et al. “Hyper-SUMOylation of the Kv7 Potassium Channel Diminishes the M-Current Leading to Seizures and Sudden Death.” *Neuron*, vol. 83, no. 5, Sept. 2014, pp. 1159–71, https://doi.org/10.1016/j.neuron.2014.07.042.

[97] Nahta, Rita, et al. “The HER-2-Targeting Antibodies Trastuzumab and Pertuzumab Synergistically Inhibit the Survival of Breast Cancer Cells.” *Cancer Research*, vol. 64, no. 7, Apr. 2004, pp. 2343–46, https://doi.org/10.1158/0008-5472.can-03-3856.

[98] Prochownik, E. V., and P. K. Vogt. “Therapeutic Targeting of Myc.” *Genes & Cancer*, vol. 1, no. 6, June 2010, pp. 650–59, https://doi.org/10.1177/1947601910377494.

[99] Stephen, Andrew G., et al. “Dragging Ras Back in the Ring.” *Cancer Cell*, vol. 25, no. 3, Mar. 2014, pp. 272–81, https://doi.org/10.1016/j.ccr.2014.02.017.

[100] Jackson, Stephen P., and Daniel Durocher. “Regulation of DNA Damage Responses by Ubiquitin and SUMO.” *Molecular Cell*, vol. 49, no. 5, Mar. 2013, pp. 795–807, https://doi.org/10.1016/j.molcel.2013.01.017.

[101] Sarangi, Prabha, and Xiaolan Zhao. “SUMO-Mediated Regulation of DNA Damage Repair and Responses.” *Trends in Biochemical Sciences*, vol. 40, no. 4, Apr. 2015, pp. 233–42, https://doi.org/10.1016/j.tibs.2015.02.006.

[102] Licciardello, Marco P., and Stefan Kubicek. *Pharmacological Treats for SUMO Addicts*. May 2016, pp. 390–97, https://doi.org/10.1016/j.phrs.2016.01.004. Accessed 26 May 2023.

[103] Nagel, Remco, et al. “Drugging the Addict: Non-Oncogene Addiction as a Target for Cancer Therapy.” *EMBO Reports*, vol. 17, no. 11, Oct. 2016, pp. 1516–31, https://doi.org/10.15252/embr.201643030.

[104] Horiuchi, Dai, et al. “Taking on Challenging Targets: Making MYC Druggable.” *American Society of Clinical Oncology Educational Book*, vol. 8, no. 34, May 2014, pp. e497–502, https://doi.org/10.14694/edbook_am.2014.34.e497.

[105] Uprety, Dipesh, and Alex A. Adjei. “KRAS: From Undruggable to a Druggable Cancer Target.” *Cancer Treatment Reviews*, vol. 89, no. 102070, Sept. 2020, p. 102070, https://doi.org/10.1016/j.ctrv.2020.102070.

[106] Green, Andrew R., et al. “MYC Functions Are Specific in Biological Subtypes of Breast Cancer and Confers Resistance to Endocrine Therapy in Luminal Tumours.” *British Journal of Cancer*, vol. 114, no. 8, Mar. 2016, pp. 917–28, https://doi.org/10.1038/bjc.2016.46.

[107] Dang, Hien, et al. “NELFE-Dependent MYC Signature Identifies a Unique Cancer Subtype in Hepatocellular Carcinoma.” *Scientific Reports*, vol. 9, no. 1, Mar. 2019, p. 3369, https://doi.org/10.1038/s41598-019-39727-9.

[108] Chen, Hui, et al. “Targeting Oncogenic Myc as a Strategy for Cancer Treatment.” *Signal Transduction and Targeted Therapy*, vol. 3, no. 1, Feb. 2018, https://doi.org/10.1038/s41392-018-0008-7.

[109] Wang, Chen, et al. “Alternative Approaches to Target Myc for Cancer Treatment.” *Signal Transduction and Targeted Therapy*, vol. 6, no. 1, Mar. 2021, https://doi.org/10.1038/s41392-021-00500-y.

[110] Oster, Sara K., et al. “Functional Analysis of the N-Terminal Domain of the Myc Oncoprotein.” *Oncogene*, vol. 22, no. 13, Apr. 2003, pp. 1998–2010, https://doi.org/10.1038/sj.onc.1206228.

[111] Wang, Xiaoyan, et al. “Phosphorylation Regulates C-Myc’s Oncogenic Activity in the Mammary Gland.” *Cancer Research*, vol. 71, no. 3, Feb. 2011, pp. 925–36, https://doi.org/10.1158/0008-5472.can-10-1032.

[112] Dang, Chi V., et al. “The C-Myc Target Gene Network.” *Seminars in Cancer Biology*, vol. 16, no. 4, Aug. 2006, pp. 253–64, https://doi.org/10.1016/j.semcancer.2006.07.014.

[113] Salvatori, B., et al. “Critical Role of C-Myc in Acute Myeloid Leukemia Involving Direct Regulation of MiR-26a and Histone Methyltransferase EZH2.” *Genes & Cancer*, vol. 2, no. 5, May 2011, pp. 585–92, https://doi.org/10.1177/1947601911416357.

[114] Yun, Seongseok, et al. “Prognostic Significance of MYC Oncoprotein Expression on Survival Outcome in Patients with Acute Myeloid Leukemia with Myelodysplasia Related Changes (AML-MRC).” *Leukemia Research*, vol. 84, no. 2756, Sept. 2019, p. 106194, https://doi.org/10.1016/j.leukres.2019.106194.

[115] Heisterkamp, Nora, et al. “Acute Leukaemia in Bcr/Abl Transgenic Mice.” *Nature*, vol. 344, no. 6263, Mar. 1990, pp. 251–53, https://doi.org/10.1038/344251a0.

[116] Jennings, Barbara A., and Ken I. Mills. “C-Myc Locus Amplification and the Acquisition of Trisomy 8 in the Evolution of Chronic Myeloid Leukaemia.” *Leukemia Research*, vol. 22, no. 10, Oct. 1998, pp. 899–903, https://doi.org/10.1016/s0145-2126(98)00097-6.

[117] Quintás-Cardama, Alfonso, and Jorge E. Cortes. “Chronic Myeloid Leukemia: Diagnosis and Treatment.” *Mayo Clinic Proceedings*, vol. 81, no. 7, 2006, pp. 973–88, https://doi.org/10.4065/81.7.973.

[118] Albajar, Marta, et al. “MYC in Chronic Myeloid Leukemia: Induction of Aberrant DNA Synthesis and Association with Poor Response to Imatinib.” *Molecular Cancer Research*, vol. 9, no. 5, May 2011, pp. 564–76, https://doi.org/10.1158/1541-7786.mcr-10-0356.

[119] Sharma, Nitesh, et al. “BCR/ABL1 and BCR Are under the Transcriptional Control of the MYC Oncogene.” *Molecular Cancer*, vol. 14, no. 1, July 2015, https://doi.org/10.1186/s12943-015-0407-0.

[120] American Cancer Society. “Cancer Facts & Figures 2020 | American Cancer Society.” *Www.cancer.org*, 2021, www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2021.html.

[121] Massó-Vallés, Daniel, and Laura Soucek. “Blocking Myc to Treat Cancer: Reflecting on Two Decades of Omomyc.” *Cells*, vol. 9, no. 4, Apr. 2020, https://doi.org/10.3390/cells9040883.

[122] Local, Andrea, et al. “APTO-253 Stabilizes G-Quadruplex DNA, Inhibits MYC Expression, and Induces DNA Damage in Acute Myeloid Leukemia Cells.” *Molecular Cancer Therapeutics*, vol. 17, no. 6, June 2018, pp. 1177–86, https://doi.org/10.1158/1535-7163.MCT-17-1209.

[123] Kessler, Jessica D., et al. “A SUMOylation-Dependent Transcriptional Subprogram Is Required for Myc-Driven Tumorigenesis.” *Science*, vol. 335, no. 6066, Jan. 2012, pp. 348–53, https://doi.org/10.1126/science.1212728. Accessed 11 Nov. 2021.

[124] Sung, Hyuna, et al. “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.” *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, Feb. 2021, pp. 209–49, acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21660.

[125] Siegel, Rebecca L., et al. “Cancer Statistics, 2023.” *CA: A Cancer Journal for Clinicians*, vol. 73, no. 1, Jan. 2023, pp. 17–48.

[126] Soppa, Ulf, and Walter Becker. “DYRK Protein Kinases.” *Current Biology*, vol. 25, no. 12, June 2015, pp. R488–89, https://doi.org/10.1016/j.cub.2015.02.067. Accessed 27 July 2023.

[127] Hill, Deirdre A., et al. “Breast Cancer Survival, Survival Disparities, and Guideline-Based Treatment.” *Breast Cancer Research and Treatment*, vol. 170, no. 2, Mar. 2018, pp. 405–14, https://doi.org/10.1007/s10549-018-4761-7.

[128] Gu, Jiangtao, et al. “Monensin Inhibits Proliferation, Migration, and Promotes Apoptosis of Breast Cancer Cells via Downregulating UBA2.” *Drug Development Research*, vol. 81(6), no. 745-753., May 2020, https://doi.org/10.1002/ddr.21683.

[129] Peart, Olive. “Breast Intervention and Breast Cancer Treatment Options.” *Radiologic Technology*, vol. 86, no. 5, May 2015, pp. 535M558M; quiz 559-562, pubmed.ncbi.nlm.nih.gov/25995413/.

[130] Maguire, S. L., et al. “Three-Dimensional Modelling Identifies Novel Genetic Dependencies Associated with Breast Cancer Progression in the Isogenic MCF10 Model.” *Minerva-Access.unimelb.edu.au*, vol. 240 (3), no. pp. 315 - 328, Nov. 2016, minerva-access.unimelb.edu.au/items/b569bef1-f25a-553d-ab6a-0b7a5a656984.

[131] Wang, Xin, et al. “Monensin Inhibits Cell Proliferation and Tumor Growth of Chemo-Resistant Pancreatic Cancer Cells by Targeting the EGFR Signaling Pathway.” *Scientific Reports*, vol. 8, no. 1, Dec. 2018, p. 17914, https://doi.org/10.1038/s41598-018-36214-5. Accessed 17 June 2022.

[132] Siegel, Rebecca L., et al. “Cancer Statistics, 2022.” *CA: A Cancer Journal for Clinicians*, vol. 72, no. 1, Jan. 2022, pp. 7–33, https://doi.org/10.3322/caac.21708.

[133] Sheng, Hongxu, et al. “Construction and Validation of a Two-Gene Signature Based on SUMOylation Regulatory Genes in Non-Small Cell Lung Cancer Patients.” *BMC Cancer*, vol. 22, no. 1, May 2022, p. 572, https://doi.org/10.1186/s12885-022-09575-4.

[134] Jiang, Biying, et al. “Identifying UBA2 as a Proliferation and Cell Cycle Regulator in Lung Cancer A549 Cells.” *Journal of Cellular Biochemistry*, vol. 120, no. 8, Aug. 2019, pp. 12752–61, https://doi.org/10.1002/jcb.28543.

[135] Anusha Vakiti, and Prerna Mewawalla. “Cancer, Acute Myeloid Leukemia (AML, Erythroid Leukemia, Myelodysplasia-Related Leukemia, BCR-ABL Chronic Leukemia).” *Nih.gov*, StatPearls Publishing, 18 Dec. 2019, www.ncbi.nlm.nih.gov/books/NBK507875/.

[136] Hartmann, Luise, and Klaus H. Metzeler. “Clonal Hematopoiesis and Preleukemia-Genetics, Biology, and Clinical Implications.” *Genes, Chromosomes & Cancer*, vol. 58, no. 12, Dec. 2019, pp. 828–38, https://doi.org/10.1002/gcc.22756.

[137] Shlush, Liran I., et al. “Tracing the Origins of Relapse in Acute Myeloid Leukaemia to Stem Cells.” *Nature*, vol. 547, no. 7661, June 2017, pp. 104–8, https://doi.org/10.1038/nature22993.

[138] Lu, Xiaoya, et al. “Identification of the UBA2-WTIP Fusion Gene in Acute Myeloid Leukemia.” *Experimental Cell Research*, vol. 371, no. 2, Oct. 2018, pp. 409–16, https://doi.org/10.1016/j.yexcr.2018.08.035.

[139] Carter, Bing Z., et al. “Triptolide Induces Caspase-Dependent Cell Death Mediated via the Mitochondrial Pathway in Leukemic Cells.” *Blood*, vol. 108, no. 2, July 2006, pp. 630–37, https://doi.org/10.1182/blood-2005-09-3898.

[140] Berger, Allison, et al. “Abstract 3079: Pharmacodynamic Evaluation of the Novel SUMOylation Inhibitor TAK-981 in a Mouse Tumor Model.” *Cancer Res*, vol. 79, no. 13_Supplement, July 2019, pp. 3079–79, https://doi.org/10.1158/1538-7445.am2019-3079.

[141] Fukuda, Isao, et al. “Ginkgolic Acid Inhibits Protein SUMOylation by Blocking Formation of the E1-SUMO Intermediate.” *Chemistry & Biology*, vol. 16, no. 2, Feb. 2009, pp. 133–40, https://doi.org/10.1016/j.chembiol.2009.01.009.

[142] Fukuda, Isao, et al. “Kerriamycin B Inhibits Protein SUMOylation.” *The Journal of Antibiotics*, vol. 62, no. 4, Mar. 2009, pp. 221–24, https://doi.org/10.1038/ja.2009.10.

[143] Takemoto, Misao, et al. “Inhibition of Protein SUMOylation by Davidiin, an Ellagitannin from Davidia Involucrata.” *The Journal of Antibiotics*, vol. 67, no. 4, Apr. 2014, pp. 335–38, https://doi.org/10.1038/ja.2013.142.

[144] Suzawa, Miyuki, et al. “A Gene-Expression Screen Identifies a Non-Toxic Sumoylation Inhibitor That Mimics SUMO-Less Human LRH-1 in Liver.” *Elife*, vol. 4, no. 4, Dec. 2015, https://doi.org/10.7554/elife.09003.

[145] Lv, Zongyang, et al. “Molecular Mechanism of a Covalent Allosteric Inhibitor of SUMO E1 Activating Enzyme.” *Nature Communications*, vol. 9, no. 1, Dec. 2018, https://doi.org/10.1038/s41467-018-07015-1.

[146] Li, Yijia, et al. “Allosteric Inhibition of Ubiquitin-like Modifications by a Class of Inhibitor of SUMO-Activating Enzyme.” *Cell Chem Biol*, vol. 26, no. 2, Feb. 2019, pp. 278-288.e6, https://doi.org/10.1016/j.chembiol.2018.10.026.

[147] “AID 2006 - UHTS HTRF Assay for Identification of Inhibitors of SUMOylation - PubChem.” *Pubchem.ncbi.nlm.nih.gov*, pubchem.ncbi.nlm.nih.gov/bioassay/2006. Accessed 18 July 2023.

[148] Kalliokoski, Tuomo, et al. “Comparability of Mixed IC50 Data – a Statistical Analysis.” *PLoS ONE*, edited by Andrea Cavalli, vol. 8, no. 4, Apr. 2013, p. e61007, https://doi.org/10.1371/journal.pone.0061007.

[149] Copeland, Robert A. *Evaluation of Enzyme Inhibitors in Drug Discovery*. Mar. 2013, https://doi.org/10.1002/9781118540398. Accessed 29 July 2020.

[150] Kuzmič, Petr. *A Two-Point IC50 Method for Evaluating the Biochemical Potency of Irreversible Enzyme Inhibitors*. June 2020, https://doi.org/10.1101/2020.06.25.171207. Accessed 21 Aug. 2023.

[151] Stewart, Mikaela D., et al. "E2 Enzymes: More than Just Middle Men." *Cell Research*, vol. 26, no. 4, Mar. 2016, pp. 423–40, https://doi.org/10.1038/cr.2016.35.

[152] Bernier-Villamor, Victor, et al. "Structural Basis for E2-Mediated SUMO Conjugation Revealed by a Complex between Ubiquitin-Conjugating Enzyme Ubc9 and RanGAP1." *Cell*, vol. 108, no. 3, Feb. 2002, pp. 345–56, https://doi.org/10.1016/s0092-8674(02)00630-x. Accessed 25 Nov. 2020.

[153] Knipscheer, Puck, et al. "Noncovalent Interaction between Ubc9 and SUMO Promotes SUMO Chain Formation." *The EMBO Journal*, vol. 26, no. 11, May 2007, pp. 2797–807, https://doi.org/10.1038/sj.emboj.7601711.

[154] Barbacid, M. "Ras Genes." *Annual Review of Biochemistry*, vol. 56, no. 1, June 1987, pp. 779–827, https://doi.org/10.1146/annurev.bi.56.070187.004023. Accessed 16 Oct. 2019.

[155] Fernandez-Medarde, A., and E. Santos. "Ras in Cancer and Developmental Diseases." *Genes & Cancer*, vol. 2, no. 3, Mar. 2011, pp. 344–58, https://doi.org/10.1177/1947601911411084.

[156] Hruban, R. H., et al. "K-Ras Oncogene Activation in Adenocarcinoma of the Human Pancreas. A Study of 82 Carcinomas Using a Combination of Mutant-Enriched Polymerase Chain Reaction Analysis and Allele-Specific Oligonucleotide Hybridization." *The American Journal of Pathology*, vol. 143, no. 2, Aug. 1993, pp. 545–54, pubmed.ncbi.nlm.nih.gov/8342602/. Accessed 21 June 2023.

[157] Kafatos, George, et al. "*RAS* Mutation Prevalence among Patients with Metastatic Colorectal Cancer: A Meta-Analysis of Real-World Data." *Biomark Med*, vol. 11, no. 9, Feb. 2016, pp. 751–60, https://doi.org/10.2217/bmm-2016-0358. Accessed 14 June 2023.

[158] Polosukhina, Dina, et al. "Functional KRAS Mutations and a Potential Role for PI3K/AKT Activation in Wilms Tumors." *Molecular Oncology*, vol. 11, no. 4, Apr. 2017, pp. 405–21, https://doi.org/10.1002/1878-0261.12044.

[159] Drosten, Matthias, and Mariano Barbacid. "Targeting the MAPK Pathway in KRAS-Driven Tumors." *Cancer Cell*, vol. 37, no. 4, Apr. 2020, pp. 543–50, https://doi.org/10.1016/j.ccell.2020.03.013.

[160] Hershkovitz, Dov, et al. "Adenoma and Carcinoma Components in Colonic Tumors Show Discordance for KRAS Mutation." *Human Pathology*, vol. 45, no. 9, Sept.

2014, pp. 1866–71, https://doi.org/10.1016/j.humpath.2014.05.005. Accessed 21 June 2023.

[161] Hayama, Tamuro, et al. "G12V and G12C Mutations in the Gene KRAS Are Associated with a Poorer Prognosis in Primary Colorectal Cancer." *International Journal of Colorectal Disease*, vol. 34, no. 8, July 2019, pp. 1491–96, https://doi.org/10.1007/s00384-019-03344-9. Accessed 23 Jan. 2020.

[162] Cook, Joshua H., et al. "The Origins and Genetic Interactions of KRAS Mutations Are Allele- and Tissue-Specific." *Nature Communications*, vol. 12, no. 1, Mar. 2021, p. 1808, https://doi.org/10.1038/s41467-021-22125-z.

[163] Hunter, J. C., et al. "Biochemical and Structural Analysis of Common Cancer-Associated KRAS Mutations." *Molecular Cancer Research*, vol. 13, no. 9, June 2015, pp. 1325–35, https://doi.org/10.1158/1541-7786.mcr-15-0203.

[164] Canon, Jude, et al. "The Clinical KRAS(G12C) Inhibitor AMG 510 Drives Anti-Tumour Immunity." *Nature*, vol. 575, no. 217–223, Oct. 2019, https://doi.org/10.1038/s41586-019-1694-1.

[165] Sakamoto, Kotaro, et al. "K-Ras(G12D)-Selective Inhibitory Peptides Generated by Random Peptide T7 Phage Display Technology." *Biochemical and Biophysical Research Communications*, vol. 484, no. 3, Mar. 2017, pp. 605–11, https://doi.org/10.1016/j.bbrc.2017.01.147. Accessed 3 Oct. 2021.

[166] Allen-Petersen, Brittany L., and Rosalie C. Sears. "Mission Possible: Advances in MYC Therapeutic Targeting in Cancer." *BioDrugs*, vol. 33, no. 5, Aug. 2019, pp. 539–53, https://doi.org/10.1007/s40259-019-00370-5. Accessed 23 Jan. 2021.

[167] Choi, Byeong Hyeok, et al. "K-Ras Lys-42 Is Crucial for Its Signaling, Cell Migration, and Invasion." *The Journal of Biological Chemistry*, vol. 293, no. 45, Nov. 2018, pp. 17574–81, https://doi.org/10.1074/jbc.RA118.003723. Accessed 21 Aug. 2023.

[168] Huang, Shuai, et al. "Combined Evaluation of MRNA and Protein Expression, Promoter Methylation, and Immune Infiltration of UBE2I in Pan-Digestive System Tumors." *Oxidative Medicine and Cellular Longevity*, vol. 2022, 2022, p. 1129062, https://doi.org/10.1155/2022/1129062. Accessed 21 Aug. 2023.

[169] Fang, Sufen, et al. "Down-Regulation of UBC9 Increases the Sensitivity of Hepatocellular Carcinoma to Doxorubicin." *Oncotarget*, vol. 8, no. 30, May 2017, pp. 49783–95, https://doi.org/10.18632/oncotarget.17939. Accessed 14 Dec. 2022.

[170] Vatsyayan, Jaya, et al. "SUMO1 Modification of NF-KB2/P100 Is Essential for Stimuli-Induced P100 Phosphorylation and Processing." *Empo*, vol. 9, no. 9, July 2008, pp. 885–90, https://doi.org/10.1038/embor.2008.122. Accessed 21 June 2023.

[171] Hirohama, Mikako, et al. “Spectomycin B1 as a Novel SUMOylation Inhibitor That Directly Binds to SUMO E2.” *ACS Chemical Biology*, vol. 8, no. 12, Dec. 2013, pp. 2635–42, https://doi.org/10.1021/cb400630z. Accessed 21 June 2023.

[172] Kim, Yeong Shik, et al. “An Electrophoretic Mobility Shift Assay Identifies a Mechanistically Unique Inhibitor of Protein Sumoylation.” *Cellpress*, vol. 20, no. 4, Apr. 2013, pp. 604–13, https://doi.org/10.1016/j.chembiol.2013.04.001. Accessed 18 May 2023.

[173] Prudden, John, et al. “Molecular Mimicry of SUMO Promotes DNA Repair.” *Nature Structural & Molecular Biology*, vol. 16, no. 5, May 2009, pp. 509–16, https://doi.org/10.1038/nsmb.1582. Accessed 20 July 2023.

[174] Prudden, John, et al. “DNA Repair and Global Sumoylation Are Regulated by Distinct Ubc9 Noncovalent Complexes.” *Molecular and Cellular Biology*, vol. 31, no. 11, June 2011, pp. 2299–310, https://doi.org/10.1128/mcb.05188-11. Accessed 11 May 2023.

[175] “12 Types of Targeted Protein Degradation Technologies | Biopharma PEG.” *Www.biochempeg.com*, www.biochempeg.com/article/278.html#:~:text=AUTAC%20molecules%20trigger%20K63%2Dlinked. Accessed 20 July 2023.

[176] Takahashi, Daiki, et al. “AUTACs: Cargo-Specific Degraders Using Selective Autophagy.” *Molecular Cell*, vol. 76, no. 5, Elsevier BV, Dec. 2019, pp. 797-810.e10, https://doi.org/10.1016/j.molcel.2019.09.009. Accessed 16 Aug. 2023.

[177] Yamashita, Daisuke, et al. “Transcription Factor HDREF Is a Novel SUMO E3 Ligase of Mi2α.” *Journal of Biological Chemistry*, vol. 291, no. 22, American Society for Biochemistry and Molecular Biology, May 2016, pp. 11619–34, https://doi.org/10.1074/jbc.m115.713370. Accessed 14 Aug. 2023.

[178] Han, Zhi-Jian, et al. “The Post-Translational Modification, SUMOylation, and Cancer (Review).” *International Journal of Oncology*, vol. 52, no. 4, Feb. 2018, https://doi.org/10.3892/ijo.2018.4280.

[179] Salas-Lloret, Daniel, and Román González-Prieto. “Insights in Post-Translational Modifications: Ubiquitin and SUMO.” *International Journal of Molecular Sciences*, vol. 23, no. 6, Mar. 2022, p. 3281, https://doi.org/10.3390/ijms23063281. Accessed 19 Oct. 2022.

[180] Yunus, Ali A., and Christopher D. Lima. “Structure of the Siz/PIAS SUMO E3 Ligase Siz1 and Determinants Required for SUMO Modification of PCNA.” *Molecular Cell*, vol. 35, no. 5, Sept. 2009, pp. 669–82, https://doi.org/10.1016/j.molcel.2009.07.013. Accessed 7 Apr. 2020.

[181] Potts, Patrick Ryan, and Hongtao Yu. “Human MMS21/NSE2 Is a SUMO Ligase Required for DNA Repair.” *Molecular and Cellular Biology*, vol. 25, no. 16, Aug. 2005, pp. 7021–32, https://doi.org/10.1128/mcb.25.16.7021-7032.2005. Accessed 31 Jan. 2021.

[182] Pichler, Andrea, et al. “The Nucleoporin RanBP2 Has SUMO1 E3 Ligase Activity.” *Cell*, vol. 108, no. 1, Jan. 2002, pp. 109–20, https://doi.org/10.1016/s0092-8674(01)00633-x. Accessed 10 Mar. 2021.

[183] Kagey, Michael H., et al. “The Polycomb Protein Pc2 Is a SUMO E3.” *Cell*, vol. 113, no. 1, Apr. 2003, pp. 127–37, https://doi.org/10.1016/s0092-8674(03)00159-4. Accessed 27 Jan. 2020.

[184] Hatakeyama, Shigetsugu. “TRIM Family Proteins: Roles in Autophagy, Immunity, and Carcinogenesis.” *Trends in Biochemical Sciences*, vol. 42, no. 4, Apr. 2017, pp. 297–311, https://doi.org/10.1016/j.tibs.2017.01.002.

[185] Koliopoulos, Marios G., et al. “Functional Role of TRIM E3 Ligase Oligomerization and Regulation of Catalytic Activity.” *EMBO J*, vol. 35, no. 11, June 2016, pp. 1204–18, https://doi.org/10.15252/embj.201593741. Accessed 14 May 2023.

[186] Kahyo, Tomoaki, et al. “Involvement of PIAS1 in the Sumoylation of Tumor Suppressor P53.” *Molecular Cell*, vol. 8, no. 3, Sept. 2001, pp. 713–18, https://doi.org/10.1016/s1097-2765(01)00349-5. Accessed 17 Mar. 2023.

[187] Puhr, M., et al. “PIAS1 Is a Determinant of Poor Survival and Acts as a Positive Feedback Regulator of AR Signaling through Enhanced AR Stabilization in Prostate Cancer.” *Oncogene*, vol. 35, no. 18, May 2016, pp. 2322–32, https://doi.org/10.1038/onc.2015.292. Accessed 22 June 2023.

[188] Yang, Nanyang, et al. “SUMO3 Modification by PIAS1 Modulates Androgen Receptor Cellular Distribution and Stability.” *Cell Communication and Signaling: CCS*, vol. 17, no. 1, Nov. 2019, p. 153, https://doi.org/10.1186/s12964-019-0457-9. Accessed 22 June 2023.

[189] Scher, Howard I., et al. “Increased Survival with Enzalutamide in Prostate Cancer after Chemotherapy.” *New England Journal of Medicine*, edited by Richard C. Cabot et al., vol. 367, no. 13, Sept. 2012, pp. 1187–97, https://doi.org/10.1056/nejmoa1207506.

[190] Ryan, Charles J., et al. “Abiraterone in Metastatic Prostate Cancer without Previous Chemotherapy.” *New England Journal of Medicine*, vol. 368, no. 2, Jan. 2013, pp. 138–48, https://doi.org/10.1056/nejmoa1209096.

[191] de Bono, Johann Sebastian, et al. “Prednisone plus Cabazitaxel or Mitoxantrone for Metastatic Castration-Resistant Prostate Cancer Progressing after Docetaxel Treatment:

A Randomised Open-Label Trial." *The Lancet*, vol. 376, no. 9747, Oct. 2010, pp. 1147–54, https://doi.org/10.1016/s0140-6736(10)61389-x. Accessed 16 Dec. 2019.

[192] Shi, Xiaoxia, et al. "The Role of SUMO E3 Ligases in Signaling Pathway of Cancer Cells." *International Journal of Molecular Sciences*, vol. 23, no. 7, Mar. 2022, p. 3639, https://doi.org/10.3390/ijms23073639. Accessed 16 Mar. 2023.

[193] "Myeloma | CDC." *Www.cdc.gov*, 8 June 2021, www.cdc.gov/cancer/myeloma/index.htm#:~:text=Myeloma%2C%20also%20called%20multiple%20myeloma.

[194] Driscoll, James F., et al. "The Sumoylation Pathway Is Dysregulated in Multiple Myeloma and Is Associated with Adverse Patient Outcome." *Blood*, vol. 115, no. 14, American Society of Hematology, Apr. 2010, pp. 2827–34, https://doi.org/10.1182/blood-2009-03-211045. Accessed 20 Apr. 2023.

[195] Sears, Rosalie C. "The Life Cycle of C-Myc: From Synthesis to Degradation." *Cell Cycle (Georgetown, Tex.)*, vol. 3, no. 9, Sept. 2004, pp. 1133–37, pubmed.ncbi.nlm.nih.gov/15467447/. Accessed 17 Aug. 2023.

[196] Welcker, Markus, et al. "The Fbw7 Tumor Suppressor Regulates Glycogen Synthase Kinase 3 Phosphorylation-Dependent C-Myc Protein Degradation." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 24, June 2004, pp. 9085–90, https://doi.org/10.1073/pnas.0402770101. Accessed 26 Nov. 2022.

[197] Hoellein, Alexander, et al. "Myc-Induced SUMOylation Is a Therapeutic Vulnerability for B-Cell Lymphoma." *Blood*, vol. 124, no. 13, Sept. 2014, pp. 2081–90, https://doi.org/10.1182/blood-2014-06-584524.

[198] Sulzmaier, Florian J., et al. "FAK in Cancer: Mechanistic Findings and Clinical Applications." *Nature Reviews Cancer*, vol. 14, no. 9, Aug. 2014, pp. 598–610, https://doi.org/10.1038/nrc3792.

[199] Gress Kadaré, et al. "PIAS1-Mediated Sumoylation of Focal Adhesion Kinase Activates Its Autophosphorylationn." *Journal of Biological Chemistry*, vol. 278, no. 48, American Society for Biochemistry and Molecular Biology, Nov. 2003, pp. 47434–40, https://doi.org/10.1074/jbc.m308562200. Accessed 30 Apr. 2023.

[200] Constanzo, Jerfiz D., et al. "PIAS1-FAK Interaction Promotes the Survival and Progression of Non-Small Cell Lung Cancer." *Neoplasia*, vol. 18, no. 5, Elsevier BV, May 2016, pp. 282–93, https://doi.org/10.1016/j.neo.2016.03.003. Accessed 17 Aug. 2023.

[201] Schmidt, Darja, and Stefan Müller. "Members of the PIAS Family Act as SUMO Ligases for C-Jun and P53 and Repress P53 Activity." *Proceedings of the National*

Academy of Sciences, vol. 99, no. 5, Feb. 2002, pp. 2872–77, https://doi.org/10.1073/pnas.052559499. Accessed 16 Mar. 2023.

[202] Tahk, Samuel, et al. "Control of Specificity and Magnitude of NF-KB and STAT1-Mediated Gene Activation through PIASy and PIAS1 Cooperation." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 28, National Academy of Sciences, July 2007, pp. 11643–48, https://doi.org/10.1073/pnas.0701877104. Accessed 16 Aug. 2023.

[203] Liu, Bin, et al. "PIAS1 Selectively Inhibits Interferon-Inducible Genes and Is Important in Innate Immunity." *Nature Immunology*, vol. 5, no. 9, Sept. 2004, pp. 891–98, https://doi.org/10.1038/ni1104. Accessed 22 June 2023.

[204] Ryu, Hong-Yeoul. "SUMO: A Novel Target for Anti-Coronavirus Therapy." *Pathogens and Global Health*, vol. 115, no. 5, Mar. 2021, pp. 292–99, https://doi.org/10.1080/20477724.2021.1906562. Accessed 1 July 2022.

[205] Yang, Yuping, et al. "Function and Inhibition of DYRK1A: Emerging Roles of Treating Multiple Human Diseases." *Biochemical Pharmacology*, vol. 212, no. 115521, June 2023, p. 115521, https://doi.org/10.1016/j.bcp.2023.115521. Accessed 6 June 2023.

[206] Kannan, Natarajan, and Andrew F. Neuwald. "Evolutionary Constraints Associated with Functional Specificity of the CMGC Protein Kinases MAPK, CDK, GSK, SRPK, DYRK, and CK2α." *Protein Science*, vol. 13, no. 8, Aug. 2004, pp. 2059–77, https://doi.org/10.1110/ps.04637904.

[207] Tejedor, Francisco J., and Barbara Hämmerle. "MNB/DYRK1A as a Multiple Regulator of Neuronal Development." *FEBS Journal*, vol. 278, no. 2, Dec. 2010, pp. 223–35, https://doi.org/10.1111/j.1742-4658.2010.07954.x.

[208] Yang, Hsiao-Pei, et al. "Molecular Nature of 11 Spontaneous *de Novo* Mutations in *Drosophila Melanogaster*." *Genetics*, vol. 157, no. 3, Mar. 2001, pp. 1285–92, https://doi.org/10.1093/genetics/157.3.1285. Accessed 27 Apr. 2022.

[209] Okui, M., et al. "High-Level Expression of the Mnb/Dyrk1A Gene in Brain and Heart during Rat Early Development." *Genomics*, vol. 62, no. 2, Dec. 1999, pp. 165–71, https://doi.org/10.1006/geno.1999.5998.

[210] Hämmerle, B., et al. "Expression Patterns and Subcellular Localization of the down Syndrome Candidate Protein MNB/DYRK1A Suggest a Role in Late Neuronal Differentiation." *European Journal of Neuroscience*, vol. 17, no. 11, June 2003, pp. 2277–86, https://doi.org/10.1046/j.1460-9568.2003.02665.x.

[211] Soundararajan, Meera, et al. "Structures of down Syndrome Kinases, DYRKs, Reveal Mechanisms of Kinase Activation and Substrate Recognition." *Structure*, vol. 21,

no. 6, June 2013, pp. 986–96, https://doi.org/10.1016/j.str.2013.03.012. Accessed 27 July 2023.

[212] Nolen, Brad, et al. “Regulation of Protein Kinases.” *Molecular Cell*, vol. 15, no. 5, Sept. 2004, pp. 661–75, https://doi.org/10.1016/j.molcel.2004.08.024. Accessed 8 Feb. 2021.

[213] Arbones, Maria L., et al. “DYRK1A and Cognition: A Lifelong Relationship.” *Pharmacology & Therapeutics*, vol. 194, Feb. 2019, pp. 199–221, https://doi.org/10.1016/j.pharmthera.2018.09.010. Accessed 25 July 2022.

[214] Ryoo, Soo-Ryoon, et al. “DYRK1A-Mediated Hyperphosphorylation of Tau.” *Journal of Biological Chemistry*, vol. 282, no. 48, Nov. 2007, pp. 34850–57, https://doi.org/10.1074/jbc.m707358200.

[215] Kim, Jung-Hyun, et al. “Doubling up on Function: Dual-Specificity Tyrosine-Regulated Kinase 1A (DYRK1A) in B Cell Acute Lymphoblastic Leukemia.” *The Journal of Clinical Investigation*, vol. 131, no. 1, Jan. 2021, https://doi.org/10.1172/JCI142627. Accessed 28 July 2023.

[216] Wang, Xiantao, et al. “P27Kip1 Overexpression Causes Apoptotic Death of Mammalian Cells.” *Oncogene*, vol. 15, no. 24, Dec. 1997, pp. 2991–97, https://doi.org/10.1038/sj.onc.1201450. Accessed 30 May 2021.

[217] Ishii, T., et al. “Effects of P27Kip1 on Cell Cycle Status and Viability in A549 Lung Adenocarcinoma Cells.” *European Respiratory Journal*, vol. 23, no. 5, May 2004, pp. 665–70, https://doi.org/10.1183/09031936.04.00096204. Accessed 12 Dec. 2021.

[218] Litovchick, Larisa, et al. “DYRK1A Protein Kinase Promotes Quiescence and Senescence through DREAM Complex Assembly.” *Genes & Development*, vol. 25, no. 8, Apr. 2011, pp. 801–13, https://doi.org/10.1101/gad.2034211.

[219] Parker, Samantha E., et al. “Updated National Birth Prevalence Estimates for Selected Birth Defects in the United States, 2004-2006.” *Birth Defects Research Part A: Clinical and Molecular Teratology*, vol. 88, no. 12, Sept. 2010, pp. 1008–16, https://doi.org/10.1002/bdra.20735.

[220] Carr, Janet. “Mental and Motor Development in Young Mongol Children.” *APA PsycNet*, Journal of Mental Deficiency Research, 1970, 14(3), 205–220.

[221] Zemel, B. S., et al. “Growth Charts for Children with down Syndrome in the United States.” *PEDIATRICS*, vol. 136, no. 5, Oct. 2015, pp. e1204–11, https://doi.org/10.1542/peds.2015-1652.

[222] Lejeune, J., and Marie Gauthier. *Les Chromosomes Humains En Culture de Tissus.* 1959, 248:602–603.

[223] Guimerá, J., et al. "A Human Homologue of Drosophila Minibrain (MNB) Is Expressed in the Neuronal Regions Affected in down Syndrome and Maps to the Critical Region." *Human Molecular Genetics*, vol. 5, no. 9, Sept. 1996, pp. 1305–10, https://doi.org/10.1093/hmg/5.9.1305.

[224] Marcia Tavares Smith, and Donald L. Dodson. "Facial Expression in Adults with Down's Syndrome." *Journal of Abnormal Psychology*, American Psychological Association, Nov. 1996, https://doi.org/10.1037//0021-843x.105.4.602. Accessed 26 Apr. 2023.

[225] Altafaj, X. "Neurodevelopmental Delay, Motor Abnormalities and Cognitive Deficits in Transgenic Mice Overexpressing Dyrk1A (Minibrain), a Murine Model of Down's Syndrome." *Human Molecular Genetics*, vol. 10, no. 18, Sept. 2001, pp. 1915–23, https://doi.org/10.1093/hmg/10.18.1915. Accessed 7 Apr. 2021.

[226] Hämmerle, B., et al. "Expression Patterns and Subcellular Localization of the down Syndrome Candidate Protein MNB/DYRK1A Suggest a Role in Late Neuronal Differentiation." *European Journal of Neuroscience*, vol. 17, no. 11, June 2003, pp. 2277–86, https://doi.org/10.1046/j.1460-9568.2003.02665.x.

[227] Fu, Maofu, et al. "Minireview: Cyclin D1: Normal and Abnormal Functions." *Endocrinology*, vol. 145, no. 12, Dec. 2004, pp. 5439–47, https://doi.org/10.1210/en.2004-0959.

[228] Liu, Fei, et al. "Overexpression of Dyrk1A Contributes to Neurofibrillary Degeneration in down Syndrome." *The FASEB Journal*, vol. 22, no. 9, May 2008, pp. 3224–33, https://doi.org/10.1096/fj.07-104539. Accessed 1 Apr. 2022.

[229] Kazemi, Mohammad, et al. "Down Syndrome: Current Status, Challenges and Future Perspectives." *International Journal of Molecular and Cellular Medicine*, vol. 5, no. 3, Babol University of Medical Sciences, Aug. 2016, pp. 125–33, www.ncbi.nlm.nih.gov/pmc/articles/PMC5125364/.

[230] Moorhead, Paul S., et al. "Chromosome Studies of Patients with Alzheimer Disease." *American Journal of Medical Genetics*, vol. 14, no. 3, Mar. 1983, pp. 545–56, https://doi.org/10.1002/ajmg.1320140319. Accessed 9 Apr. 2022.

[231] Ferrer, Isidro, et al. "Constitutive Dyrk1A Is Abnormally Expressed in Alzheimer Disease, down Syndrome, Pick Disease, and Related Transgenic Models." *Neurobiology of Disease*, vol. 20, no. 2, Nov. 2005, pp. 392–400, https://doi.org/10.1016/j.nbd.2005.03.020. Accessed 20 Dec. 2022.

[232] Kimura, Ryo, et al. “The DYRK1A Gene, Encoded in Chromosome 21 down Syndrome Critical Region, Bridges between β-Amyloid Production and Tau Phosphorylation in Alzheimer Disease.” *Human Molecular Genetics*, vol. 16, no. 1, Nov. 2006, pp. 15–23, https://doi.org/10.1093/hmg/ddl437. Accessed 20 Dec. 2022.

[233] Wegiel, Jerzy, et al. “The Role of DYRK1A in Neurodegenerative Diseases.” *FEBS Journal*, vol. 278, no. 2, Dec. 2010, pp. 236–45, https://doi.org/10.1111/j.1742-4658.2010.07955.x. Accessed 6 Dec. 2019.

[234] Alonso, Alejandra del C., et al. “Alzheimer’s Disease Hyperphosphorylated Tau Sequesters Normal Tau into Tangles of Filaments and Disassembles Microtubules.” *Nature Medicine*, vol. 2, no. 7, July 1996, pp. 783–87, https://doi.org/10.1038/nm0796-783.

[235] Stotani, Silvia, et al. “DYRK1A Inhibition as Potential Treatment for Alzheimer’s Disease.” *Future Medicinal Chemistry*, vol. 8, no. 6, Apr. 2016, pp. 681–96, https://doi.org/10.4155/fmc-2016-0013. Accessed 23 Feb. 2022.

[236] Chung, Sul-Hee. “Aberrant Phosphorylation in the Pathogenesis of Alzheimer’s Disease.” *BMB Reports*, vol. 42, no. 8, Aug. 2009, pp. 467–74, https://doi.org/10.5483/bmbrep.2009.42.8.467. Accessed 17 Mar. 2021.

[237] Kim, Hyeongki, et al. “A Chemical with Proven Clinical Safety Rescues Down-Syndrome-Related Phenotypes in through DYRK1A Inhibition.” *Disease Models & Mechanisms*, vol. 9, no. 8, Aug. 2016, pp. 839–48, https://doi.org/10.1242/dmm.025668.

[238] Aït Yahya-Graison, E., et al. “Classification of Human Chromosome 21 Gene-Expression Variations in down Syndrome: Impact on Disease Phenotypes.” *The American Journal of Human Genetics*, vol. 81, no. 3, Sept. 2007, pp. 475–91, https://doi.org/10.1086/520000.

[239] Head, Elizabeth, et al. “Alzheimer’s Disease in down Syndrome.” *European Journal of Neurodegenerative Disease*, vol. 1, no. 3, Dec. 2012, pp. 353–64, www.ncbi.nlm.nih.gov/pmc/articles/PMC4184282/.

[240] Chatterjee, Shreyasi, and Amritpal Mudher. “Alzheimer’s Disease and Type 2 Diabetes: A Critical Assessment of the Shared Pathological Traits.” *Frontiers in Neuroscience*, vol. 12, Frontiers Media S.A., 2018, p. 383, https://doi.org/10.3389/fnins.2018.00383.

[241] Gregor, Margaret F., and Gökhan S. Hotamisligil. “Inflammatory Mechanisms in Obesity.” *Annual Review of Immunology*, vol. 29, no. 1, Apr. 2011, pp. 415–45, https://doi.org/10.1146/annurev-immunol-031210-101322.

[242] Cavaghan, Melissa K., et al. “Interactions between Insulin Resistance and Insulin Secretion in the Development of Glucose Intolerance.” *Journal of Clinical Investigation*, vol. 106, no. 3, Aug. 2000, pp. 329–33, https://doi.org/10.1172/jci10761.

[243] Muoio, Deborah M., and Christopher B. Newgard. “Molecular and Metabolic Mechanisms of Insulin Resistance and β-Cell Failure in Type 2 Diabetes.” *Nature Reviews Molecular Cell Biology*, vol. 9, no. 3, Mar. 2008, pp. 193–205, https://doi.org/10.1038/nrm2327.

[244] Eizirik, Décio L., et al. “Pancreatic β-Cells in Type 1 and Type 2 Diabetes Mellitus: Different Pathways to Failure.” *Nature Reviews Endocrinology*, vol. 16, no. 7, May 2020, pp. 349–62, https://doi.org/10.1038/s41574-020-0355-7. Accessed 14 Jan. 2021.

[245] Wang, Peng, et al. “Disrupting the DREAM Complex Enables Proliferation of Adult Human Pancreatic β Cells.” *The Journal of Clinical Investigation*, vol. 132, no. 15, Aug. 2022, https://doi.org/10.1172/JCI157086. Accessed 18 Aug. 2023.

[246] Lu, Mei, et al. “DYRK1A Aggravates β Cell Dysfunction and Apoptosis by Promoting the Phosphorylation and Degradation of IRS2.” *Experimental Gerontology*, vol. 125, Oct. 2019, p. 110659, https://doi.org/10.1016/j.exger.2019.110659. Accessed 18 Aug. 2023.

[247] Stamateris, Rachel E., et al. “Glucose Induces Mouse β-Cell Proliferation via IRS2, MTOR, and Cyclin D2 but Not the Insulin Receptor.” *Diabetes*, vol. 65, no. 4, American Diabetes Association, Jan. 2016, pp. 981–95, https://doi.org/10.2337/db15-0529. Accessed 18 Aug. 2023.

[248] Wang, Peng, et al. “A High-Throughput Chemical Screen Reveals That Harmine-Mediated Inhibition of DYRK1A Increases Human Pancreatic Beta Cell Replication.” *Nature Medicine*, vol. 21, no. 4, Mar. 2015, pp. 383–88, https://doi.org/10.1038/nm.3820. Accessed 14 May 2021.

[249] Haas, Mary E., et al. “The Regulation of ApoB Metabolism by Insulin.” *Trends in Endocrinology and Metabolism: TEM*, vol. 24, no. 8, Aug. 2013, https://doi.org/10.1016/j.tem.2013.04.001.

[250] De Felice, Fernanda G., et al. “How Does Brain Insulin Resistance Develop in Alzheimer’s Disease?” *Alzheimer’s & Dementia*, vol. 10, no. 1, Feb. 2014, pp. S26–32, https://doi.org/10.1016/j.jalz.2013.12.004.

[251] Steen, Eric, et al. “Impaired Insulin and Insulin-like Growth Factor Expression and Signaling Mechanisms in Alzheimer’s Disease – Is This Type 3 Diabetes?” *Journal of Alzheimer’s Disease*, vol. 7, no. 1, Mar. 2005, pp. 63–80, https://doi.org/10.3233/jad-2005-7107.

[252] Kumar, Kushal, et al. “Recent Advances in the Neurobiology and Neuropharmacology of Alzheimer’s Disease.” *Biomedicine & Pharmacotherapy*, vol. 98, no. 2018, Feb. 2018, pp. 297–307, https://doi.org/10.1016/j.biopha.2017.12.053. Accessed 25 Nov. 2020.

[253] Hamzé, Rim, et al. “Type 2 Diabetes Mellitus and Alzheimer’s Disease: Shared Molecular Mechanisms and Potential Common Therapeutic Targets.” *International Journal of Molecular Sciences*, vol. 23, no. 23, Dec. 2022, p. 15287, https://doi.org/10.3390/ijms232315287.

[254] Ogawa, Yasushi, et al. “Development of a Novel Selective Inhibitor of the down Syndrome-Related Kinase Dyrk1A.” *Nature Communications*, vol. 1, no. 1, Oct. 2010, https://doi.org/10.1038/ncomms1090.

[255] Schmiegelow, K., et al. “Long-Term Results of NOPHO ALL-92 and ALL-2000 Studies of Childhood Acute Lymphoblastic Leukemia.” *Leukemia*, vol. 24, no. 2, Dec. 2009, pp. 345–54, https://doi.org/10.1038/leu.2009.251.

[256] Harewood, L., et al. “Amplification of AML1 on a Duplicated Chromosome 21 in Acute Lymphoblastic Leukemia: A Study of 20 Cases.” *Leukemia*, vol. 17, no. 3, Mar. 2003, pp. 547–53, https://doi.org/10.1038/sj.leu.2402849.

[257] Robinson, C. Michael, et al. “Adult Distal Humeral Metaphyseal Fractures: Epidemiology and Results of Treatment.” *Journal of Orthopaedic Trauma*, vol. 17, no. 1, Jan. 2003, pp. 38–47, https://doi.org/10.1097/00005131-200301000-00006. Accessed 4 Jan. 2020.

[258] Toft, N., et al. “Results of NOPHO ALL2008 Treatment for Patients Aged 1–45 Years with Acute Lymphoblastic Leukemia.” *Leukemia*, vol. 32, no. 3, Aug. 2017, pp. 606–15, https://doi.org/10.1038/leu.2017.265.

[259] Ivanov Öfverholm, Ingegerd, et al. “Overexpression of Chromatin Remodeling and Tyrosine Kinase Genes in IAMP21-Positive Acute Lymphoblastic Leukemia.” *Leukemia & Lymphoma*, vol. 61, no. 3, Oct. 2019, pp. 604–13, https://doi.org/10.1080/10428194.2019.1678153.

[260] Cheung, Edna, et al. “The Leukemia Strikes Back: A Review of Pathogenesis and Treatment of Secondary AML.” *Annals of Hematology*, vol. 98, no. 3, Mar. 2019, pp. 541–59, https://doi.org/10.1007/s00277-019-03606-0.

[261] Thijs Booiman, et al. “DYRK1A Controls HIV-1 Replication at a Transcriptional Level in an NFAT Dependent Manner.” *PLoS One*, vol. 10, no. 12, Dec. 2015, pp. e0144229–29, https://doi.org/10.1371/journal.pone.0144229.

[262] Bhansali, Rahul S., et al. “DYRK1A Regulates B Cell Acute Lymphoblastic Leukemia through Phosphorylation of FOXO1 and STAT3.” *Journal of Clinical Investigation*, vol. 131, no. 1, Jan. 2021, https://doi.org/10.1172/jci135937. Accessed 16 Mar. 2022.

[263] Chen, Wenlian, et al. “Enhanced Fructose Utilization Mediated by SLC2A5 Is a Unique Metabolic Feature of Acute Myeloid Leukemia with Therapeutic Potential.” *Epub*, vol. 30, no. 5, Nov. 2016, pp. 779–91, https://doi.org/10.1016/j.ccell.2016.09.006.

[264] Li, Zhuoyan, et al. “Alterations of T-Cell-Mediated Immunity in Acute Myeloid Leukemia.” *Oncogene*, vol. 39, no. 18, Apr. 2020, pp. 3611–19, https://doi.org/10.1038/s41388-020-1239-y.

[265] Zhang, Liang, et al. “Pharmacological Effects of Harmine and Its Derivatives: A Review.” *Archives of Pharmacal Research*, vol. 43, no. 12, Nov. 2020, pp. 1259–75, https://doi.org/10.1007/s12272-020-01283-6. Accessed 22 May 2022.

[266] Nafie, Ebtesam, et al. “Harmine Inhibits Breast Cancer Cell Migration and Invasion by Inducing the Degradation of Twist1.” *PLoS ONE* , vol. 16, no. 2, Feb. 2021, pp. e0247652–52, https://doi.org/10.1371/journal.pone.0247652.

[267] Sato, Katsutoshi, et al. “Simultaneous CK2/TNIK/DYRK1 Inhibition by 108600 Suppresses Triple Negative Breast Cancer Stem Cells and Chemotherapy-Resistant Disease.” *Nature Communications*, vol. 12, no. 1, Aug. 2021, https://doi.org/10.1038/s41467-021-24878-z.

[268] Hong, Ji Yeon, et al. “Down’s-Syndrome-Related Kinase Dyrk1A Modulates the P120-Catenin–Kaiso Trajectory of the Wnt Signaling Pathway.” *Journal of Cell Science*, vol. 125, no. 3, Feb. 2012, pp. 561–69, https://doi.org/10.1242/jcs.086173.

[269] Nishikawa, Yoshihiro, et al. “648021,.” *BioMed Research International*, vol. 2014, no. 648021, 2014, pp. 1–6, https://doi.org/10.1155/2014/648021.

[270] Liu, Pixu, et al. “Targeting the Phosphoinositide 3-Kinase Pathway in Cancer.” *Nature Reviews Drug Discovery*, vol. 8, no. 8, Aug. 2009, pp. 627–44, https://doi.org/10.1038/nrd2926.

[271] Zhu, Yu, et al. “Tissue-Resident Macrophages in Pancreatic Ductal Adenocarcinoma Originate from Embryonic Hematopoiesis and Promote Tumor Progression.” *Immunity*, vol. 47, no. 2, Aug. 2017, pp. 323-338.e6, https://doi.org/10.1016/j.immuni.2017.07.014. Accessed 8 Apr. 2019.

[272] Massey, Andrew, et al. “Targeting DYRK1A/B Kinases to Modulate P21-Cyclin D1-P27 Signalling and Induce Anti-Tumour Activity in a Model of Human Glioblastoma.”

J Cell Mol Med, vol. 25, no. 22, Oct. 2021, pp. 10650–62, https://doi.org/10.1111/jcmm.17002.

[273] Lambert, Joshua D., et al. "Epigallocatechin-3-Gallate Is Absorbed but Extensively Glucuronidated Following Oral Administration to Mice." *The Journal of Nutrition*, vol. 133, no. 12, Dec. 2003, pp. 4172–77, https://doi.org/10.1093/jn/133.12.4172. Accessed 10 May 2020.

[274] Kim, Karl H. S., et al. "Distinct Cortical Areas Associated with Native and Second Languages." *Nature*, vol. 388, no. 6638, July 1997, pp. 171–74, https://doi.org/10.1038/40623.

[275] Göckler, Nora, et al. "Harmine Specifically Inhibits Protein Kinase DYRK1A and Interferes with Neurite Formation." *FEBS Journal*, vol. 276, no. 21, Oct. 2009, pp. 6324–37, https://doi.org/10.1111/j.1742-4658.2009.07346.x.

[276] Becker, Walter, and Wolfgang Sippl. "Activation, Regulation, and Inhibition of DYRK1A." *FEBS Journal*, vol. 278, no. 2, Dec. 2010, pp. 246–56, https://doi.org/10.1111/j.1742-4658.2010.07956.x.

[277] Meggio, Flavio, and Lorenzo A. Pinna. "One-Thousand-And-One Substrates of Protein Kinase CK2?" *The FASEB Journal*, vol. 17, no. 3, Mar. 2003, pp. 349–68, https://doi.org/10.1096/fj.02-0473rev. Accessed 4 Jan. 2022.

[278] Ackeifi, Courtney, et al. "GLP-1 Receptor Agonists Synergize with DYRK1A Inhibitors to Potentiate Functional Human β Cell Regeneration." *Science Translational Medicine*, vol. 12, no. 530, Feb. 2020, https://doi.org/10.1126/scitranslmed.aaw9996.

[279] Park, Chun Shik, and H. Daniel Lacorazza. "DYRK2 Controls a Key Regulatory Network in Chronic Myeloid Leukemia Stem Cells." *Experimental & Molecular Medicine*, vol. 52, no. 10, Oct. 2020, pp. 1663–72, https://doi.org/10.1038/s12276-020-00515-5.

[280] Ma, Fei, et al. "Dual-Specificity Tyrosine Phosphorylation-Regulated Kinase 3 Loss Activates Purine Metabolism and Promotes Hepatocellular Carcinoma Progression." *Hepatology (Baltimore, Md.)*, vol. 70, no. 5, Nov. 2019, pp. 1785–803, https://doi.org/10.1002/hep.30703.

[281] Kumar, Kunal, et al. "Synthesis and Biological Validation of a Harmine-Based, Central Nervous System (CNS)-Avoidant, Selective, Human β-Cell Regenerative Dual-Specificity Tyrosine Phosphorylation-Regulated Kinase a (DYRK1A) Inhibitor." *Journal of Medicinal Chemistry*, vol. 63, no. 6, Jan. 2020, pp. 2986–3003, https://doi.org/10.1021/acs.jmedchem.9b01379.

[282] Nica Alvarez, Mó, et al. “DYRK1A Autophosphorylation on Serine Residue 520 Modulates Its Kinase Activity via 14-3-3 Binding □ D.” *Molecular Biology of the Cell*, vol. 18, 2007, pp. 1167–78, https://doi.org/10.1091/mbc.E06. Accessed 27 July 2023.

[283] Ashford, Anne L., et al. “A Novel DYRK1B Inhibitor AZ191 Demonstrates That DYRK1B Acts Independently of GSK3β to Phosphorylate Cyclin D1 at Thr286, Not Thr288.” *Biochemical Journal*, vol. 457, no. 1, Dec. 2013, pp. 43–56, https://doi.org/10.1042/bj20130461. Accessed 12 Sept. 2022.

[284] Yousefelahiyeh, Mina, et al. “DCAF7/WDR68 Is Required for Normal Levels of DYRK1A and DYRK1B.” *PLOS ONE*, edited by Arun Rishi, vol. 13, no. 11, Nov. 2018, p. e0207779, https://doi.org/10.1371/journal.pone.0207779. Accessed 13 May 2023.

[285] Papenfuss, Marco, et al. “Differential Maturation and Chaperone Dependence of the Paralogous Protein Kinases DYRK1A and DYRK1B.” *Scientific Reports*, vol. 12, no. 1, Feb. 2022, p. 2393, https://doi.org/10.1038/s41598-022-06423-0. Accessed 20 July 2023.

[286] Kho, Changwon, et al. “SUMO1-Dependent Modulation of SERCA2a in Heart Failure.” *Nature*, vol. 477, no. 7366, Sept. 2011, pp. 601–5, https://doi.org/10.1038/nature10407.

[287] Rammohan, Malini, et al. “The Chromosome 21 Kinase DYRK1A: Emerging Roles in Cancer Biology and Potential as a Therapeutic Target.” *Oncogene*, vol. 41, no. 14, Feb. 2022, pp. 2003–11, https://doi.org/10.1038/s41388-022-02245-6. Accessed 28 July 2023.

[288] Soppa, Ulf, et al. “The down Syndrome-Related Protein Kinase DYRK1A Phosphorylates $P27^{Kip1}$and Cyclin D1 and Induces Cell Cycle Exit and Neuronal Differentiation.” *Cell Cycle*, vol. 13, no. 13, May 2014, pp. 2084–100, https://doi.org/10.4161/cc.29104.

[289] Lochhead, Pamela, et al. “Nurse.” *Cell*, vol. 121, 2005, pp. 925–36, https://doi.org/10.1016/j.cell.2005.03.034. Accessed 27 July 2023.

[290] Taussig, David C., et al. “Anti-CD38 Antibody–Mediated Clearance of Human Repopulating Cells Masks the Heterogeneity of Leukemia-Initiating Cells.” *Blood*, vol. 112, no. 3, Aug. 2008, pp. 568–75, https://doi.org/10.1182/blood-2007-10-118331.

[291] Yan, Xiao-Jing, et al. “Exome Sequencing Identifies Somatic Mutations of DNA Methyltransferase Gene DNMT3A in Acute Monocytic Leukemia.” *Nature Genetics*, vol. 43, no. 4, Mar. 2011, pp. 309–15, https://doi.org/10.1038/ng.788. Accessed 22 Apr. 2020.

[292] Zhang, J. -H . “A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays.” *Journal of Biomolecular Screening*, vol. 4, no. 2, Apr. 1999, pp. 67–73, https://doi.org/10.1177/108705719900400206.

[293] Singh, Juswinder, et al. “The Resurgence of Covalent Drugs.” *Nature Reviews Drug Discovery*, vol. 10, no. 4, Apr. 2011, pp. 307–17, https://doi.org/10.1038/nrd3410

[294] Sparks, Robert P., et al. “Use of Surface Plasmon Resonance (SPR) to Determine Binding Affinities and Kinetic Parameters between Components Important in Fusion Machinery.” *Methods in Molecular Biology (Clifton, N.J.)*, vol. 1860, 2019, pp. 199–210, https://doi.org/10.1007/978-1-4939-8760-3_12.