Two Chemical Biology Strategies to Interrogate DEAD-box Proteins

by
Megan Moore

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

in

Chemistry and Chemical Biology

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Kevan Shokat
8FDF47C586EF40D...

Kevan Shokat
Chair

DocuSigned by:
Stephen Floor

Stephen Floor

DocuSigned by:
Davide Ruggero
14C3B6AD60C5446...

Davide Ruggero

Committee Members

Dedication

To my parents, who have given me their unconditional love and support over the past six years and over the twenty-two before that.

Acknowledgements

Whoever claimed graduate school is like a marathon had clearly never run a marathon because a PhD is a whole lot harder. This journey could not have happened without incredible levels of support from so many people, in and out of the lab, and I'm so grateful to everyone who has helped me make this possible.

I'm grateful to my advisor, Kevan Shokat, for giving me both the opportunity to work in his lab and the space once I was there to follow my curiosity and work on projects that intrigued me. I appreciate the support and scientific advice from my committee members, Davide Ruggero and Stephen Floor, as well as them both opening their labs to me so I could learn more RNA biology. I was lucky to work alongside some of the smartest, hardest working, funniest, and most generous people I have ever met and I'm grateful for the time spent with my labmates and the lessons I learned from them on how to be a better scientist, co-worker, and friend. In particular, I would like to thank Krister Barkovich and Kevin Lou, who in turns mentored and collaborated with me and without whom my research never would have been imagined. Thank you also to the Chemistry and Chemical Biology (CCB) program, my CCB cohort, and the other incredible trainees across so many fields at UCSF whose journeys intersected and elevated my own.

This journey started long before I had ever heard of UCSF. I'm forever thankful for my academic games coaches, Diana Wieberg and Adrian Prather, who introduced me to some of my closest friends, showed me just how cool learning could be, and were with me from fifth grade through the end of high school. I actually never liked science (I hope I won't get in trouble for saying that in my thesis for a PhD in science!) until I took chemistry in tenth grade with Dr. Cindy Philpot and my whole world changed. It's thanks to her and my college organic chemistry professor, Julianne Yost, that I ever thought of studying chemistry. I'm thankful for my college

advisors, Nicholas Hud and Dr. Seiichi Matsuda, for taking a chance on an undergrad who had no idea what she was doing and allowed me to work in their laboratories. I'm forever grateful though for Jing Jin and David Fialho, the two graduate students in their labs who taught me so much about how to conduct research.

Although I frequently joke about it, I am slightly serious when I say I don't think I would've made it through grad school without my bike. Living in San Francisco offered me more access to incredible nature than I previously knew was possible and the trail runs, bike rides, and backpacking adventures were sometimes the only thing keeping me sane through failed experiments and pandemic lockdowns. I was able to experience so much outside of lab and meet so many inspiring people and I'm grateful to my chosen family in San Francisco who were with me through every step, pedal stroke, and pastry bite.

Finally, thank you to my friends and family who put up with me for these six tumultuous years (and I hope will continue to do so) and stopped asking me how my thesis was going. I appreciate it. But truly, I cannot begin to put into words how much of an impact every one of you has had on me and how much that means to me. I'm grateful for my parents and sister for their love. I wouldn't be here without all the sacrifices and hard work of my family and grandparents and I'm so grateful for them all. To Grammy and Grandad, who passed away during my PhD: I finally did it! I love you both and think of you all the time; what a wonderful world. Thank you all for your unconditional love and support; I love you all.

Contributions

The research in this dissertation were performed under the guidance of Dr. Kevan Shokat. Dr. Stephen Floor and Dr. Davide Ruggero also provided key scientific guidance on all the chapters. Several chapters of this thesis contain material from previously published work. They do not represent the final published form and have been edited.

Part of Chapter 2 of this dissertation is a reprint of a previous publication:

Barkovich, K. J., **Moore, M. K**., Hu, Q., & Shokat, K. M. (2018). Chemical genetic inhibition of DEAD-box proteins using covalent complementarity. *Nucleic acids research*, *46*(17), 8689-8699. K.J.B. and M.K.M. performed experiments. K.J.B. performed chemical synthesis. K.J.B. and Q.H. analyzed crystallography data. K.J.B. and K.M.S. designed and analyzed experiments. K.J.B. and K.M.S. wrote the manuscript. All authors edited the manuscript.

Krister Barkovich provided plasmids for bacterial expression and he also provided many helpful comments and protocols. Holly Vickery performed the Selleck Chem library screen. Screen results were analyzed with a python script provided by Adolfo Cuesta. WuXi AppTech provided the DEL materials and performed sequencing for the screen through their DELOpen program. Larry Zhu provided useful comments on small molecule library screens. Srivats Venkataramanan helped with CRISPR knock-out and knock-in experiments and Stephen Floor's lab provided useful comments and resources. Kevin Lou provided useful comments for NPM labeling and CRISPRi experiments. Luke Gilbert's lab also provided useful comments and resources for CRISPRi experiments. Codon optimality reporters were designed with a python script provided by Jack Stevenson.

Chapter 3 of this dissertation contains data that is a reprint of a previous publication:

Lou, K., Wassarman, D. R., Yang, T., Paung, Y., Zhang, Z., O'Loughlin, T. A., **Moore, M. K.**, Egan, R. K., Greninger, P., Benes, C. H., Seeliger, M., Taunton, J., Gilbert, L., A. & Shokat, K.

M. (2022). IFITM proteins assist cellular uptake of diverse linked chemotypes. *Science*, *378*(6624), 1097-1104.

Part of Chapter 3 is also a preprint of a manuscript. Kevin Lou synthesized molecules, performed CRISPRa experiments, and provided helpful discussion. Chris Ott's lab at Massachusetts General Hospital screened the cancer cell line panel and Doug Wassarman shared his python script for the Cellpanelr tool for data analysis. eFFECTOR provided reagents and useful comments. Luke Gilbert's lab provided helpful comments and resources for CRISPRi experiments. Reagents and helpful protocols were provided by eFFFECTOR. Sara Warrington performed all Roc-RiboTAC experiments and Ying Shi helped with synthesis. Celine Wang performed all (de)methylation experiments in the lab of Ziyang Zhang at UC Berkeley.

Each chapter in this dissertation reflects a substantial collective team effort without which none of this work would be possible.

“I was, I am, I will be”

-Chanel Miller

Two Chemical Biology Strategies to Interrogate DEAD-Box Proteins

Megan Moore

Abstract

As mRNA translation is an energetically costly and crucial operation, cells have developed a plethora of mechanisms to regulate protein production at the post-transcriptional level, controlling initiation and elongation as well as mRNA stability to finely tune translation. In yeast, the DEAD-box protein Dhh1 has been implicated in coordinating translational repression and mRNA decay as a sensor of codon optimality but the exact molecular mechanism and whether this is recapitulated in mammalian cells through the homologue DDX6 still remain unclear. We are developing a chemical genetic inhibition strategy to elucidate the functions of DDX6 as well as other DEAD-box proteins using small molecules that covalently bind an engineered cysteine in the enzyme active site. This approach provides improved temporal and dosing control compared to traditional genetic methods. The latest efforts on this project aim to improve upon the first generation of probes, to discover new small molecule scaffolds, and to develop tools to monitor the probe's efficacy in cells.

Dysregulated protein synthesis is also a hallmark of many diseases, including cancer. The natural product Rocaglamide A (RocA) targets the DEAD-box protein eIF4A to selectively inhibit translation initiation on mRNAs with long, structured 5'-UTRs, including many oncogenes. We hypothesize that dimerizing RocA to add a layer of cell specificity through cellular uptake by IFITM proteins could "tune" the effects of such a potent molecule and broaden its therapeutic index by decreasing on target, off tumor side effects.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction: Perspective on targeting DEAD-Box Proteins

Abstract

DEAD-Box proteins are a large, conserved family of RNA helicases that modify RNA secondary structure and RNA-protein complexes. These proteins perform essential roles throughout RNA metabolism and are frequently implicated in a variety of diseases but have thus far been difficult to study with traditional genetic methods. Chemical genetic tools offer a more promising strategy to uncover these proteins functions and identify potential druggable targets but these methods still require further development. DDX6 is a mammalian DEAD-Box protein that would be an ideal candidate for further study, as its yeast homolog Dhh1 has been implicated as a sensor of codon optimality in mRNA translation but it has been more difficult to study DDX6 in its native context. Additionally, the DEAD-Box protein eukaryotic initiation factor-4A (eIF4A) is another exciting target. eIF4A plays an essential role in the translation initiation of many oncogenes and has several known small molecule inhibitors with intriguing gain-of-function effects that can act as a scaffold off of which to build functionalities that modify translation inhibition selectivity.

Introduction

Cells must maintain strict control of protein synthesis to regulate cell growth and division; dysregulation of these systems can result in the rapid growth and proliferation that is a hallmark of cancerous cells (Averous & Proud, 2006). The complex roles played by RNA in this process are still being fully appreciated, from transcription and splicing in the nucleus to export and

translation to sequestration and eventual decay to all the less well elucidated regulatory roles of diverse, noncoding RNAs.

DEAD-Box Proteins

Control at the RNA level stems in large part from post-transcriptional gene regulation through the coordination of mRNA translation and decay pathways (Jacobson et al, 1996). An important but understudied family of proteins that plays key roles throughout these processes is the ATP-dependent, RNA-binding DEAD-Box proteins (Dbp), as seen in **Figure 1.1**. These proteins are canonically known for their non-translocatory RNA duplex unwinding activity (Liu et al, 2008; Chen et al, 2008) but have been more broadly described as RNA chaperones (Jankowsky & Bowers, 2006), locally unwinding secondary structure and acting as a scaffold to disrupt or promote assembly of other RNA-binding proteins (RBP). Dbps are highly conserved across eukaryotes, with approximately 38 proteins identified in *Homo sapiens*, less than double the number found in *Saccharomyces cerevisiae* (Linder & Jankowsky, 2011), which contrasts highly to enzyme families like kinases, which tend to be greatly expanded in higher eukaryotes. Current genetic methods (Ran et al, 2013), such as knockouts, are not sufficient to interrogate DEAD-box proteins because they are often essential. Moreover, the lack of temporal resolution in knockout models can lead to downstream effects that make it difficult to deconvolute the results and determine the primary function of the protein. Dbps often appear to play multiple or overlapping roles (Jamieson & Beggs, 1991; Linder & Jankowsky, 2011; Galicia-Vázquez, 2012), which can lead to compensation over the timescale necessary to validate genetic methods. Pharmacological inhibition offers tight temporal control over experiments, allowing for precise interrogation of this pathway. However, the few well-studied Dbp inhibitors are specific to the

protein eIF4A and act allosterically, prohibiting their use as generalizable tools to study other Dbps. In addition, the ATP binding pocket is highly conserved across DEAD-box proteins (Schütz, 2010), making it difficult to achieve selectivity for a single Dbp. The second chapter of this thesis focuses on a chemical genetic strategy to develop electrophilic inhibitors specific to "electrophile-sensitive" versions of Dbps, with the goal of creating a platform that overcomes the difficulties of pure genetic strategies to study Dbps of interest (Knight & Shokat, 2007; Floor et al, 2016; Barkovich et al, 2018; **Figure 1.3**).

The Role of DDX6

As translation is one of the oldest (Lee et al, 1999) and most energetically costly cellular processes (Warner, 1999), cells also face strong selective pressure to maintain tight control over this system to ensure fitness (**Figure 1.2**). The 3'-untranslated region (UTR) of mRNAs has been shown extensively to play a role in regulation of mRNA decay through microRNA (miRNA) and RNA-binding protein sites (Bartel, 2009; Shyu et al, 1991; Geisberg et al, 2014) and the 5'-UTR is known to be involved in translation initiation as well as degradation (Wilkie et al, 2003). Until recently, the coding sequence was thought of solely as storing the protein encoding information. However, the coding sequence itself has now been shown to play a role in regulation in a number of organisms (Presnyak et al, 2015; Harigaya et al, 2016; Boël et al, 2016; Bazzini et al, 2016). This process appears to be regulated to varying extents in different organisms by codon optimality and amino acid identity (Forrest et al, 2018). A direct correlation between codon optimality and sequence half-life has been demonstrated (Hanson & Coller, 2018), but the molecular mechanism of this relationship is yet to be fully elucidated.

The DEAD-box protein DDX6 and its yeast homolog, Dhh1, are known to function in both translation and degradation of mRNA (Carroll et al, 2011; Presnyak & Coller, 2013). Dhh1 has been shown to couple these two processes in yeast to lead to increased degradation of mRNAs (Radhakrishnan et al, 2016) with lower codon optimality (**Figure 1.4**) but DDX6 partially rescues Dhh1 deletion (Westmoreland et al, 2003) and has thus far only been implicated in translational repression (Freimer et al, 2018). It remains an open question whether DDX6 recapitulates the role of Dhh1 as a sensor of codon optimality or if this process occurs via another mechanism in mammalian cells. The second chapter of this thesis also discusses development of cell lines and reporters with the goal of combining the electrophile-sensitive inhibition approach with other chemical genetic tools to examine the mechanism of codon optimality sensing in humans and better study the various roles of DDX6.

Targeting eIF4A

Although most Dbps lack well-characterized tool compounds that could be used to study them, the protein eIF4A is an exception as it is the target of several structurally unrelated families of natural products (Bordeleau et al, 2008; Lindqvist et al, 2008; Iwasaki et al, 2018) and is considered a potential therapeutic candidate (Wolfe et al, 2014). The third chapter of this thesis is a preprint of a manuscript to be submitted for publication and describes our work modifying a small series of these inhibitors to better understand how their specificity can be modulated.

Recent years have expanded the concept of how a drug-like molecule looks (Lou et al, 2022). Up until lately, modern drug discovery was primarily characterized by molecules that followed guidelines like Lipinski's rule of five (Lipinski et al, 2001) to achieve bioavailability and permeability, despite the fact that many natural products deviate dramatically from these rules. As

the variety of proteins targeted by drug campaigns has expanded, new mechanisms of action have also come under investigation. Molecular glues and bifunctional molecules offer avenues to stabilize or create novel interfaces between biological macromolecules, as typified by proteolysis-targeting chimeras (PROTACs), which induce targeted protein degradation by bringing an E3 ligase in proximity to a protein of interest and ubiquitinating it (Békés et al, 2022).

Targeting protein synthesis offers another mechanism to deplete levels of proteins of interest and can take advantage of the many nodes of post-transcriptional gene control. Cancer cells in particular become dependent on aberrant mRNA translation for their survival and unchecked proliferation (Ruggero & Pandolfi, 2003; Ruggero et al, 2004; Hsieh et al, 2012) so targeting translation potentially opens an avenue to minimize on-target, off-tumor toxicity. Translation initiation is typically the rate-limiting step in translation and has shown a greater therapeutic index than most translation elongation inhibitors, which makes translation initiation a promising area for drug development. The assembly of several eukaryotic initiation factors (eIFs) into the eIF4F complex on the mRNA 5'-untranslated region (5'-UTR) is a critical step for cap-dependent translation to allow the 43S ribosome pre-initiation complex (43S PIC) to scan through the 5'-UTR uninhibited until it reaches the start codon (Pelletier et al, 2015). eIF4A is an ATP-dependent DEAD-box RNA-binding protein (RBP) that belongs to the eIF4F complex and locally unwinds RNA duplexes (Parsyan et al, 2011), which is essential for the 43S PIC to scan through mRNAs with long, highly structured 5'-UTRs (Wolfe et al, 2014).

Several classes of natural products target eIF4A, including rocaglates, which are typified by the molecule rocaglamide A (RocA). RocA has an intriguing mode of action, stabilizing the interaction between eIF4A and mRNA to form a ternary complex (Iwasaki et al, 2019). This inhibition acts as a gain of function, effectively halting translation initiation and trapping eIF4A

on A- and G-rich mRNA sequences through pi-pi stacking interactions with the purine rings. This sequence specificity makes RocA a semi-selective translation inhibitor that is highly potent against many cancer lines. RocA inhibits the translation of many oncogenes, such as KRAS, CYD1, and CMYC, that have long structured 5'-UTRs with polypurine tracts (Iwasaki et al, 2016). A clinical candidate based on the RocA scaffold with improved PK/PD properties is currently in combined Phase I/II clinical trials for breast and KRAS-mutant cancers (eFFECTOR; Ernst et al, 2020).

Other natural and synthetic members of the rocaglate family act very similarly to RocA, with either a polypurine or no sequence preference (Chu et al, 2020), suggesting that traditional medicinal chemistry approaches modifying functional groups on the core scaffold would have negligible improvements on the molecule's potency and specificity. Recent work has broadened our understanding of the role of eIF4A, demonstrating that it multimerizes on RNA to unwind duplexes *in vitro* (Schmidt et al, 2023) and has been found bound to 5'-UTRs outside of the eIF4F complex along the mRNA binding channel of the 40S ribosomal subunit (Querido et al, 2022). This data suggests that multimers of eIF4A could play a role in translation initiation outside of the eIF4F complex and are potentially the target of RocA.

We hypothesized that dimerizing RocA could take advantage of these recently revealed functions by binding to two adjacent eIF4A proteins multimerized on mRNA. We tested this homodimeric molecule (BisRoc) see how dimerization of a molecule affected chemical genetic interactions, something that has not been directly studied before to our knowledge. In addition, we wanted to learn how targeting eIF4A dimers would affect the cellular specificity and potency of the molecule. Our work on BisRoc presents a proof-of-concept that RocA can be linked to other small molecules and maintain its high potency but gain new behaviors.

Figures

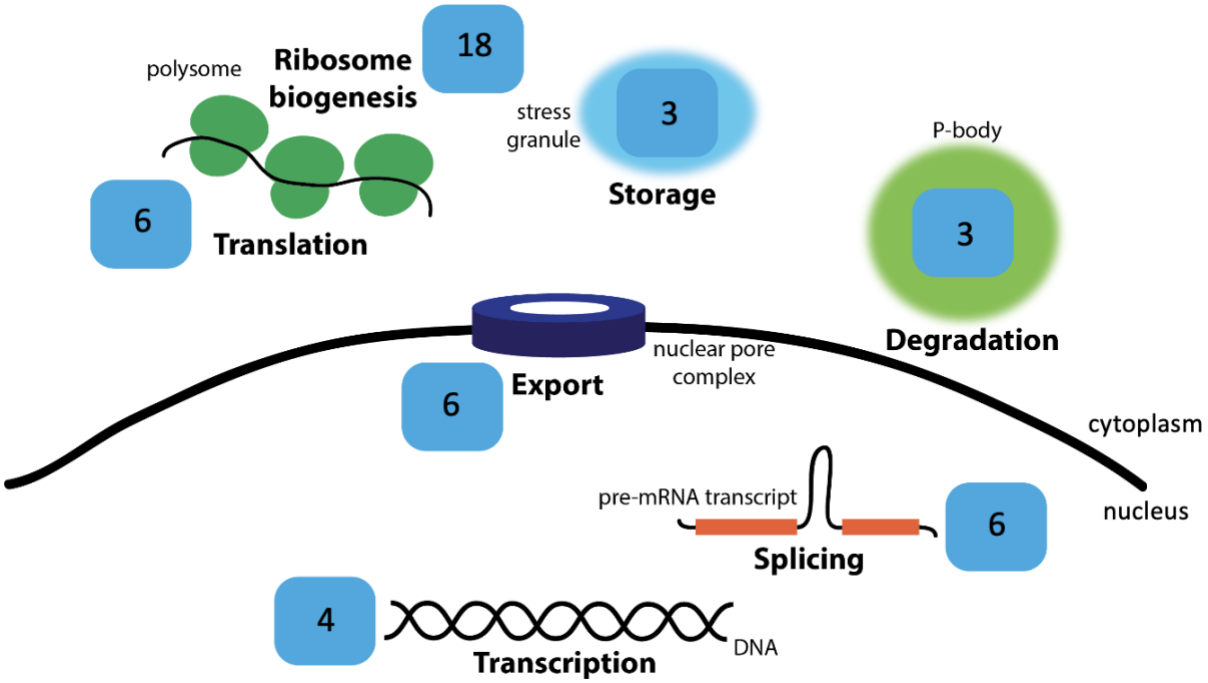


Figure 1.1 DEAD-Box Proteins (DBPs) play roles throughout RNA metabolism in eukaryotic cells. The numbers in each blue box represent estimates of the number of DBPs involved in the associated function in human cells. Adapted from Linder & Jankowsky, 2011.

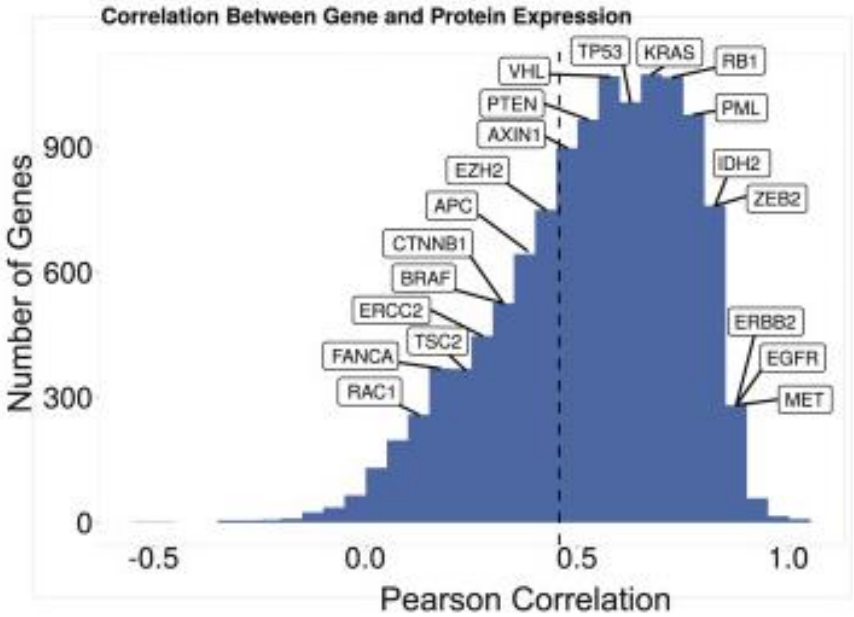


Figure 1.2 Correlation between gene and protein expression. Gene expression is not the sole determinant of protein expression and the correlation between the two varies widely from gene to gene, suggesting that some mRNAs are subject to more post-transcriptional regulation. Figure reprinted from Nusinow et al, 2019.

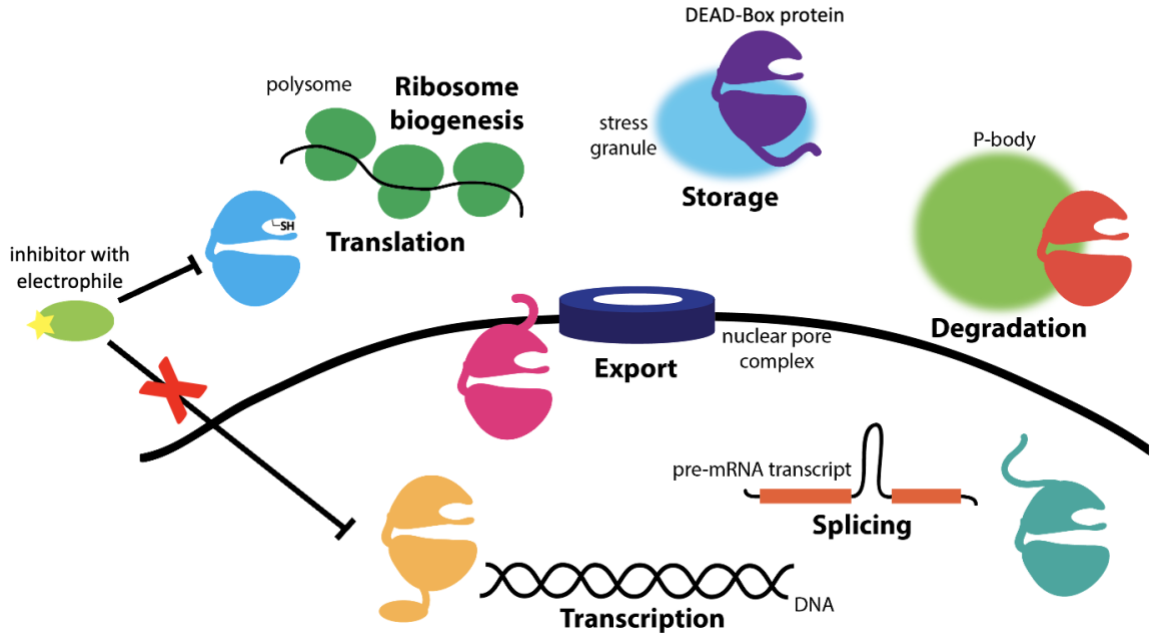


Figure 1.3 Electrophile-sensitive inhibition of DBPs. Chemical genetic methods of inhibition, such as substituting a nucleophilic residue in the protein of interest to selectively react with an electrophilic probe, offer a modular platform to inhibit and study proteins of interest.

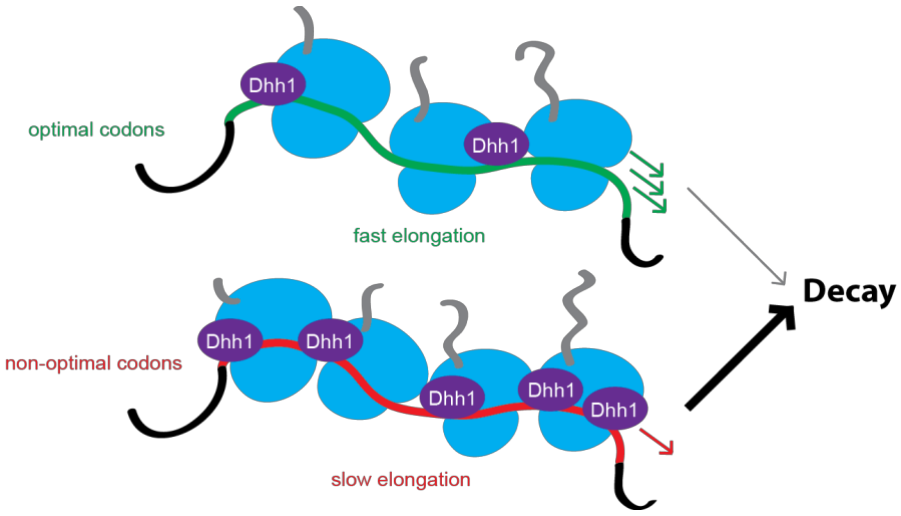


Figure 1.4 Dhh1 is a sensor of codon optimality in yeast. mRNAs containing more non-optimal codons have a slow rate of elongation during translation, which appears to be sensed by Dhh1. More Dhh1 bound to transcripts correlates with lower translation and a higher rate of mRNA decay. Adapted from Presnyak et al, 2015.

References

Averous, J., & Proud, C. G. (2006). When translation meets transformation: the mTOR story. *Oncogene*, *25*(48), 6423-6435.

Barkovich, K. J., Moore, M. K., Hu, Q., & Shokat, K. M. (2018). Chemical genetic inhibition of DEAD-box proteins using covalent complementarity. *Nucleic acids research*, *46*(17), 8689-8699.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *cell*, *136*(2), 215-233.

Bazzini, A. A., del Viso, F., Moreno-Mateos, M. A., Johnstone, T. G., Vejnar, C. E., Qin, Y., ... & Giraldez, A. J. (2016). Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. *The EMBO journal*, *35*(19), 2087-2103.

Békés, M., Langley, D. R., & Crews, C. M. (2022). PROTAC targeted protein degraders: the past is prologue. *Nature Reviews Drug Discovery*, *21*(3), 181-200.

Boël, G., Letso, R., Neely, H., Price, W. N., Wong, K. H., Su, M., ... & Xiao, R. (2016). Codon influence on protein expression in E. coli correlates with mRNA levels. *Nature*, *529*(7586), 358.

Bordeleau, M. E., Robert, F., Gerard, B., Lindqvist, L., Chen, S. M., Wendel, H. G., ... & Pelletier, J. (2008). Therapeutic suppression of translation initiation modulates chemosensitivity in a mouse lymphoma model. *The Journal of clinical investigation*, *118*(7), 2651-2660.

Carroll, J. S., Munchel, S. E., & Weis, K. (2011). The DExD/H box ATPase Dhh1 functions in translational repression, mRNA decay, and processing body dynamics. *The Journal of cell biology*, *194*(4), 527-537.

Chen, Y., Potratz, J. P., Tijerina, P., Del Campo, M., Lambowitz, A. M., & Russell, R. (2008). DEAD-box proteins can completely separate an RNA duplex using a single ATP. *Proceedings of the National Academy of Sciences*, *105*(51), 20203-20208.

Chu, J., Zhang, W., Cencic, R., O'Connor, P. B., Robert, F., Devine, W. G., ... & Pelletier, J. (2020). Rocaglates induce gain-of-function alterations to eIF4A and eIF4F. *Cell reports*, *30*(8), 2481-2488.

Ernst, J. T., Thompson, P. A., Nilewski, C., Sprengeler, P. A., Sperry, S., Packard, G., ... & Reich, S. H. (2020). Design of development candidate eFT226, a first in class inhibitor of eukaryotic initiation factor 4A RNA helicase. *Journal of medicinal chemistry*, *63*(11), 5879-5955.

Floor, S. N., Barkovich, K. J., Condon, K. J., Shokat, K. M., & Doudna, J. A. (2016). Analog sensitive chemical inhibition of the DEAD-box protein DDX 3. *Protein Science*, *25*(3), 638-649.

Forrest, M. E., Narula, A., Sweet, T. J., Arango, D., Hanson, G., Ellis, J., ... & Rissland, O. S. (2018). Codon usage and amino acid identity are major determinants of mRNA stability in humans. *Available at SSRN 3305366*.

Freimer, J. W., Hu, T. J., & Blelloch, R. (2018). Decoupling the impact of microRNAs on translational repression versus RNA degradation in embryonic stem cells. *eLife*, *7*, e38014.

Galicia-Vázquez, G., Cencic, R., Robert, F., Agenor, A. Q., & Pelletier, J. (2012). A cellular response linking eIF4AI activity to eIF4AII transcription. *Rna*, *18*(7), 1373-1384.

Geisberg, J. V., Moqtaderi, Z., Fan, X., Ozsolak, F., & Struhl, K. (2014). Global analysis of mRNA isoform half-lives reveals stabilizing and destabilizing elements in yeast. *Cell*, *156*(4), 812-824.

Hanson, G., & Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. *Nature reviews Molecular cell biology*, *19*(1), 20.

Harigaya, Y., & Parker, R. (2016). Analysis of the association between codon optimality and mRNA stability in Schizosaccharomyces pombe. *BMC genomics*, *17*(1), 895.

Hsieh, A. C., Liu, Y., Edlind, M. P., Ingolia, N. T., Janes, M. R., Sher, A., ... & Ruggero, D. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature*, *485*(7396), 55-61.

Iwasaki, S., Floor, S. N., & Ingolia, N. T. (2016). Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature*, *534*(7608), 558-561.

Iwasaki, S., Iwasaki, W., Takahashi, M., Sakamoto, A., Watanabe, C., Shichino, Y., ... & Ingolia, N. T. (2019). The translation inhibitor rocaglamide targets a bimolecular cavity between eIF4A and polypurine RNA. *Molecular cell*, *73*(4), 738-748.

Jacobson, A., & Peltz, S. W. (1996). Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. *Annual review of biochemistry*, *65*(1), 693-739.

Jamieson, D. J., & Beggs, J. D. (1991). A suppressor of yeast spp81/ded1 mutations encodes a very similar putative ATP-dependent RNA helicase. *Molecular microbiology*, *5*(4), 805-812.

Jankowsky, E., & Bowers, H. (2006). Remodeling of ribonucleoprotein complexes with DExH/D RNA helicases. *Nucleic acids research*, *34*(15), 4181-4188.

Lee, J. H., Choi, S. K., Roll-Mecak, A., Burley, S. K., & Dever, T. E. (1999). Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2. *Proceedings of the National Academy of Sciences*, *96*(8), 4342-4347.

Linder, P., & Jankowsky, E. (2011). From unwinding to clamping—the DEAD box RNA helicase family. *Nature reviews Molecular cell biology*, *12*(8), 505.

Lindqvist, L., Oberer, M., Reibarkh, M., Cencic, R., Bordeleau, M. E., Vogt, E., ... & Pelletier, J. (2008). Selective pharmacological targeting of a DEAD box RNA helicase. *PloS one*, *3*(2), e1583.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. vol. 46, issue 1-3. *Adv Drug Deliv Rev*, 3-26.

Liu, F., Putnam, A., & Jankowsky, E. (2008). ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *Proceedings of the National Academy of Sciences*, *105*(51), 20209-20214.

Knight, Z. A., & Shokat, K. M. (2007). Chemical genetics: where genetics and pharmacology meet. *Cell*, *128*(3), 425-430.

Nusinow, D. P., Szpyt, J., Ghandi, M., Rose, C. M., McDonald III, E. R., Kalocsay, M., ... & Gygi, S. P. (2020). Quantitative proteomics of the cancer cell line encyclopedia. *Cell*, *180*(2), 387-402.

Parsyan, A., Svitkin, Y., Shahbazian, D., Gkogkas, C., Lasko, P., Merrick, W. C., & Sonenberg, N. (2011). mRNA helicases: the tacticians of translational control. *Nature reviews Molecular cell biology*, *12*(4), 235-245.

Pelletier, J., Graff, J., Ruggero, D., & Sonenberg, N. (2015). Targeting the eIF4F translation initiation complex: a critical nexus for cancer development. *Cancer research*, *75*(2), 250-263.

Presnyak, V., & Coller, J. (2013). The DHH1/RCKp54 family of helicases: an ancient family of proteins that promote translational silencing. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1829*(8), 817-823.

Presnyak, V., Alhusaini, N., Chen, Y. H., Martin, S., Morris, N., Kline, N., ... & Coller, J. (2015). Codon optimality is a major determinant of mRNA stability. *Cell*, *160*(6), 1111-1124.

Radhakrishnan, A., Chen, Y. H., Martin, S., Alhusaini, N., Green, R., & Coller, J. (2016). The DEAD-box protein Dhh1p couples mRNA decay and translation by monitoring codon optimality. *Cell*, *167*(1), 122-132.

Querido, J. B., Sokabe, M., Díaz-López, I., Gordiyenko, Y., Fraser, C. S., & Ramakrishnan, V. (2022). The structure of a human translation initiation complex reveals two independent roles for the helicase eIF4A. *bioRxiv*, 2022-12.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature protocols*, *8*(11), 2281-2308.

Ruggero, D., & Pandolfi, P. P. (2003). Does the ribosome translate cancer?. *Nature Reviews Cancer*, *3*(3), 179-192.

Ruggero, D., Montanaro, L., Ma, L., Xu, W., Londei, P., Cordon-Cardo, C., & Pandolfi, P. P. (2004). The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. *Nature medicine*, *10*(5), 484-486.

Schmidt, T., Dabrowska, A., Waldron, J. A., Hodge, K., Koulouras, G., Gabrielsen, M., ... & Bushell, M. (2023). eIF4A1-dependent mRNAs employ purine-rich 5'UTR sequences to activate localised eIF4A1-unwinding through eIF4A1-multimerisation to facilitate translation. *Nucleic acids research*, *51*(4), 1859-1879.

Schütz, P., Karlberg, T., Van Den Berg, S., Collins, R., Lehtiö, L., Högbom, M., ... & Moche, M. (2010). Comparative structural analysis of human DEAD-box RNA helicases. *PloS one*, *5*(9), e12791.

Shyu, A. B., Belasco, J. G., & Greenberg, M. E. (1991). Two distinct destabilizing elements in the c-fos message trigger deadenylation as a first step in rapid mRNA decay. *Genes & development*, *5*(2), 221-231.

Warner, J. R. (1999). The economics of ribosome biosynthesis in yeast. *Trends in biochemical sciences*, *24*(11), 437-440.

Westmoreland, T. J., Olson Jr, J. A., Saito, W. Y., Huper, G., Marks, J. R., & Bennett, C. B. (2003). Dhh1 regulates the G1/S-checkpoint following DNA damage or BRCA1 expression in yeast1. *Journal of Surgical Research*, *113*(1), 62-73.

Wilkie, G. S., Dickson, K. S., & Gray, N. K. (2003). Regulation of mRNA translation by 5′-and 3′-UTR-binding factors. *Trends in biochemical sciences*, *28*(4), 182-188.

Wolfe, A. L., Singh, K., Zhong, Y., Drewe, P., Rajasekhar, V. K., Sanghvi, V. R., ... & Wendel, H. G. (2014). RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature*, *513*(7516), 65-70.

Chapter 2

Chemical genetic inhibition of DEAD-Box Proteins

Abstract

A chemical genetic platform using covalent complementarity between an electrophilic small molecule inhibitor and a nucleophilic mutation in a small molecule binding pocket of the protein was developed as a new approach to better interrogate the activity of DEAD-Box proteins. The initial inhibitor synthesized in this approach was not cell permeable so more work was done to install protecting groups. Additionally, several small molecule screens were conducted to discover new inhibitor scaffolds but none were potent enough to warrant further work. CRISPR was used to develop cell lines with electrophile-sensitive DEAD-Box proteins and several assays were tested to measure inhibitor binding in cells. Finally, it was shown that DDX6 does not cause translational repression of mRNAs with non-optimal codons like its homolog Dhh1 and no link was found with several other known factors of translational repression and degradation.

Introduction

As introduced in Chapter 1, DEAD-Box proteins (Dbps) are the largest family of ATP-dependent RNA chaperones in humans (Jankowsky & Bowers, 2006). These proteins play crucial roles in human health and disease (Linder & Jankowsky, 2011) but most are not well-understood, in part because many of these proteins are intractable with traditional genetic approaches. Additionally, there are no known small molecule inhibitors of DDX6 and inhibitors of other proteins in this family act allosterically (Iwasaki et al, 2016; Ito et al, 2017; Iwatani-Yoshihara et al, 2017), prohibiting their use as generalizable tools to study other Dbps. Furthermore, the ATP binding pocket is highly conserved across DEAD-box proteins (Schütz et al, 2010), making it

difficult to achieve selectivity for a single Dbp. A platform to study these proteins using covalent complementarity was initially described in Barkovich et al, 2018.

Cysteine is the most nucleophilic and second least common of the 20 natural amino acids and cysteine-reactive small molecules have been widely studied, including some which are clinically approved (Singh et al, 1997). These characteristics make it an ideal candidate to be installed as a point mutation in a small molecule-binding pocket of a protein of interest to create an "electrophile-sensitive" (ES) mutant (Garske et al, 2011). We previously identified the third residue of the P-loop in Dbps as an ideal location to install a cysteine in close proximity to the ATP-binding site, the only known, conserved small molecule pocket across the family (**Figure 2.1 A**). The protein DDX3 is a representative member of the family. DDX3 plays a role in translation initiation and has been implicated in the progression of my diseases, including intellectual disabilities and cancer (Lennox et al, 2020; Gadek et al, 2023). An ES version of DDX3, DDX3 S228C, retained (lowered) enzymatic activity (**Figure 2.1 B).** AMP is a known inhibitor of Ded1 (yeast homolog of DDX3) enzymatic activity (Putnam & Jankowsky, 2013) and was thus used as a scaffold to append a series of electrophiles. Out of these, AMP-acrylate was the most potent at labeling DDX3 and inhibiting enzymatic activity (**Figure 2.1 C-D**). This electrophile-sensitive approach with AMP-acrylate has shown promise but this molecule is highly cell impermeable, which greatly limits the range of experiments for which it can be used. Modifications to this scaffold to improve its permeability while maintaining potency and selectivity could greatly increase the utility of this tool compound, particularly for identifying endogenous substrates and functions of Dbps where activity can be disrupted in cell lysates or *in vitro* systems.

One of the initial drivers of this chemical genetic strategy to inhibit DEAD-box proteins was the lack of known small molecule inhibitors of this family of proteins. The natural products hippuristanol, silvestrol, and rocaglamide (RocA) were identified as inhibitors of the helicase eIF4A (Bordeleau et al, 2008; Lindqvist et al, 2008; Wolf et al, 2014; Iwasaki et al, 2018) and RocA has been claimed to inhibit DDX3 to a lesser extent (Chen et al, 2021). However, all of these molecules take advantage of cryptic binding sites on eIF4A that precludes them from being useful starting points for the development of molecules for broader use in the study of DEAD-box proteins. Other inhibitors have been reported in the literature but none of these have been successfully reproduced and did not exhibit any inhibition in biochemical assays in our hands (Floor et al, 2016). Discovery of new classes of inhibitors of Dbps could provide a starting point for a cell permeable probe for our electrophile- sensitive approach as well as molecules of potential therapeutic use (Steimer & Klostermeier, 2012; Shadrick et al, 2013).

A Dbp of particular interest is DDX6, which plays roles in translation repression and whose yeast homolog Dhh1 has been shown to act as a sensor of codon optimality (Presnyak et al, 2015). DDX6 has no known small molecule inhibitors and has proven difficult to study in its context in mammalian cells so new tools are necessary to better understand its functions and substrates relative to its homologs. This work describes the development of DDX6 knock-out and electrophile-sensitive cell lines, as well as work on *in cellulo* assays to identify AMP-acrylate binding without an enzymatic readout. Additionally, the development of codon optimality reporters in human cells provides another tool to test the function of DDX6 and search for other proteins that may play roles in codon optimality sensing in humans.

The long-term aims of this work are two-fold. The first is to better understand the roles and mechanism of DDX6 and how it acts in comparison to Dhh1. The second goal is to further development of an electrophile-sensitive inhibition system so it can be broadly applied to elucidate the functions of other members of the DEAD-box protein family. The objective of this study is to develop a variety of chemical genetic tools to precisely probe potential functions of DDX6 in translational repression and mRNA decay more cleanly than the current standard genetic methods.

Results

X-ray crystallography structure of AMP-acrylamide bound to DDX3^{ES}

To better understand the binding mode of AMP electrophiles to DEAD-Box proteins, a crystal structure of DDX3 S228C bound to AMP-acrylamide was solved to 3 Å resolution (**Table 2.1**). Overlaying this structure (5CZ9) with the structure of DDX3 WT bound to AMP (5E7J; Floor et al, 2016) showed that the overall protein structure remains the same (**Figure 2.2**). AMP-acrylamide was used instead of AMP-acrylate, which undergoes a two-step covalent reaction upon binding (Barkovich et al, 2018) to unambiguously resolve electron density for the compound bound to protein. AMP-acrylamide could be fit in the ATP-binding pocket and showed a covalent bond formed with Cys-228 (Figure 2.2 inset). AMP-acrylamide also maintained hydrogen-bonding interactions with the Q-motif and pi-stacking with Tyr-200. The greatest change upon AMP-acrylamide binding was an observed shifting of the P-loop to accommodate the orientation of the covalent bond, which caused Thr-226 to flip up 9.4Å (**Figure 2.2** inset). Additionally, the AMP-acrylamide phosphoramidate is shifted at least 3Å from the site of the AMP phosphate, losing some of the electrostatic interactions that the phosphate typically makes with the P-loop.

Antiviral protecting group strategy for AMP-acrylates

This current generation of Dbp ES probes employed an AMP scaffold because there were no other known active site inhibitors of DEAD-box proteins (Floor et al, 2016). However, the polar, negatively charged nature of the nucleotide that provides such high affinity for the active site also prevents these molecules from being cell permeable, which severely limits the range of experiments in which they can currently be employed. The antiviral field has faced similar challenges in the development of effective drugs because many of the known small molecule inhibitors are nucleoside and nucleotide analogues. This necessitated the development of prodrug strategies to improve cellular uptake and potency of such compounds (Pradere et al, 2014), which I repurposed to optimize our existing molecules.

One such strategy is employed by the clinically approved antiviral drugs adefovir and tenofovir, which are both similarly based on modified AMP scaffolds. These molecules utilize esters that "cage" the phosphate to improve cell permeability and are then cleaved by esterases upon uptake into the cell, revealing the phosphate for improved contact with the target. Another antiviral, sofosbuvir, makes use of the ProTide (PROdrug + nucleoTIDE) strategy developed by Chris McGuigan (Slusarczyk et al, 2018). This approach modifies the phosphate with an aromatic ester and a simple amino acid, which are cleaved by esterases and amidases, respectively, upon entry into the cell.

A previous graduate student, Krister Barkovich, appended an acrylate to the adefovir scaffold and successfully demonstrated that this new molecule, "adefovir-acrylate," could label DDX3 ES *in vitro* nearly as fast as the best compound, AMP-acrylate (**Figure 2.3 A-B**). Based on this data and the protecting group used on adefovir, I synthesized a version of adefovir with only

one hydroxyl group of the phosphate protected, and then added a methacrylate group to the other hydroxyl (**Figure 2.3 C**). This final product was detected in the reaction mixture by LC-MS but unable to be generated and purified in sufficient yield to be carried forward for biological testing.

Re-purposing kinase inhibitors

Due to the lack of known promiscuous DEAD-box protein small molecule binders and the difficulties encountered by the AMP-based inhibitors, we took a relatively unbiased approach to search for new chemical matter by screening inhibitors of related classes of proteins. Other ATPases, such as kinases, have thus far been more successfully targeted than helicases and contain similar motifs in their ATP-binding pockets that could allow small molecule inhibitors of these enzymes to display some promiscuity against DEAD-box proteins. Screening libraries of such compounds could aid in finding a cell-permeable scaffold upon which we can append an electrophile and modify to improve selectivity for Dbps and weaken affinity for their original kinase targets.

I initially screened the SelleckChem library, which is a curated set of 378 structurally diverse and cell permeable kinase inhibitors, most of which are ATP-competitive. Screening for binding was conducted by competing these noncovalent inhibitors with one of the current covalent inhibitors, AMP-acrylamide, and assessing labeling by whole protein mass spectrometry. This method offers several advantages: it is high throughput as compared to available activity-based assays, competition with AMP-acrylate ensures that hits are most likely ATP competitive, and the assay readout could potentially have a lower false hit rate than certain activity-based assays, such as ATPase activity. Labeling of the protein by AMP-acrylamide in the presence of DMSO serves

as a negative control for an expected maximal rate of labeling, while AMP, a known inhibitor of certain DEAD-box proteins including Ded1, was also competed to provide a positive control.

This initial assay stratified compounds and those that showed percent labeling two standard deviations below the mean were carried forward (**Figure 2.4 A**). To ensure reproducibility and dose sensitivity, this experiment was be repeated with several concentrations of these initial hits (12.5-100 μM). Based on this data, six of the initial screen hits (**Figure 2.5**) were tested to see if they stabilized DDX3 by differential scanning fluorimetry (DSF). Four of these compounds showed 1-3°C stabilization compared to the apo protein, although no hit stabilized as well as the natural inhibitor AMP (**Figure 2.4 B-C**). Inhibition of helicase activity by the best hits was also tested biochemically by a duplex unwinding assay (Jankowsky & Putnam, 2010; Barkovich et al, 2019). This assay uses a short end-labeled RNA oligonucleotide that hybridizes with a longer unlabeled strand. Incubation with a helicase leads to unwinding of the two strands, which can be read out by a native gel. Three of the top hits were tested in this assay and demonstrated unwinding inhibition (**Figure 2.4 D**). However, a major caveat of this screening and hit triaging approach was that all molecules had to be tested at high micromolar concentrations to see activity in the selected assays, which is well above the compound affinities for their intended kinase targets and poses a high risk of aggregation. Dynamic light scattering (DLS) is an important counter assay that should be used to confirm that inhibition of labeling or enzymatic activity is not simply a result of compound aggregation.

DNA-encoded library screen for new chemical matter

To continue the search for new scaffolds for Dbp inhibitors, we turned to a DNA-encoded library (DEL) screen. DEL uses combinatorial chemistry and DNA-barcoding to generate

extremely large libraries of compounds (Mannocci et al, 2011; Yuen & Franzini, 2017). These libraries can be screened in a single well, allowing far larger quantities of compounds to be screened and requiring fewer reagents than many other high-throughput screening methods. The WuXi AppTech DNA-encoded library screened here included approximately 2.8 billion compounds (WuXi DELOpen). To search for ATP-competitive binders of ES Dbps, we performed the screen against DDX3 S228C D1 (132-406) with three rounds of selection. Only the first domain (D1) of the protein was used instead of the longer construct because the second domain (D2) contains many of the RNA-binding motifs that could potentially lead to false positives by interacting with the DNA barcodes. Additionally, a counter screen was performed with the same protein fully labeled by AMP-acrylamide (**Figure 2.6 A**). By only choosing hits that were enriched against the apo protein but not the labeled protein, binders to other parts of the protein (which could still be interesting as potential allosteric inhibitors) and the DNA barcodes could be eliminated. Following sequencing of the enriched libraries for each condition, 270 molecules binding to the apo protein with over 1000-fold enrichment compared to the AMP-acrylamide labeled protein were identified. The greatest enrichment identified for a molecule was 60,000-fold. The top molecules were also compared to those from a DEL screen performed in tandem against the GTPase domain of LRRK2 (unpublished data) and none of the same top molecules were enriched in both screens, nor did selected hits from this screen show activity against LRRK2. Enriched molecules belonged to several different classes of sub-libraries within the DEL and 5 representative molecules were chosen for custom off-DNA synthesis to confirm binding. These 5 molecules were tested in an LC-MS AMP-acrylamide labeling competition assay and an *in vitro* duplex unwinding assay to determine if they inhibited DDX3 helicase activity. One of these molecules, 10115-451, showed activity in both assays (**Figure 2.6 B-C**). However, high

micromolar compound concentrations were necessary to observe this activity, which can often lead to compound aggregation, particularly in assays without detergent, such as the LC-MS assay. Due to their low potency against DDX3 when retested, these compounds were not continued with further. Had they appeared more promising, WuXi AppTech would have sent their structures and we would have attempted to solve co-crystal structures of the molecules bound to DDX3 to collect much structure activity relationship (SAR) data and determine where an electrophile could be appended to the scaffold to use these in our covalent complementary approach.

Generation of electrophile-sensitive DEAD-Box protein cell lines

To take advantage of AMP-acrylate as an inhibitor, biological systems expressing electrophile-sensitive mutants also needed to be developed. We focused on DDX3 because it was a well-behaved model for the family that behaved well in our hands and DDX6 because we felt that current genetic tools most used in the field were not sufficient to study its functions in mRNA decay. Therefore, these methods to study DDX6 would be well complemented by an electrophile-sensitive mutant that could be inhibited to varying degrees on a short timescale and would deconvolute scaffolding and enzymatic functions of the protein that are both lost with other chemical genetic approaches like dTag (Nabet et al, 2018).

Using CRISPR, I engineered HEK293T cell lines to express an electrophile-sensitive mutant of DDX6, T144C ($DDX6^{ES}$). Given that no suitable Cas9 PAM site exists in the region to be mutated, I first knocked out the endogenous DDX6, which is non-lethal in yeast and certain mammalian cell lines (Presnyak et al, 2015; Lumb et al, 2017; Freimer et al, 2018) but appears to be essential at the organismal level (Karczewski et al, 2019). I selected single cell clones and confirmed the knockout by sequencing and western blotting (**Figure 2.7 A**). The cells displayed

no phenotypic change under normal growth conditions and clones #20 and #33 both showed a full knockout, while #32 had an in-frame deletion that resulted in only partial knockout. I then transfected knockout cells with FLAG-tagged $DDX6^{WT}$ or $DDX6^{ES}$ and used CRISPR to integrate these at a safe harbor locus. After selecting for stable transfections and isolating single cell clones, sequencing indicated that the DDX6 repair template was not integrated at the safe harbor locus and this method was unsuccessful. I instead transduced previously reported TagRFP-T-DDX6 (RFP-DDX6) and SNAP-DDX6 constructs and was able to generate stable cell lines that expressed RFP-DDX6 WT or ES in DDX6 KO lines (**Figure 2.7 B**). Additionally, the FLAG-tagged DDX6 (WT and ES) and previously generated FLAG-tagged DDX3 (WT and ES) constructs could be used to overexpress DDX6 or DDX3 via transient transfection for other experiments.

Assay development for detection of AMP-acrylate labeling in cell lysate

Given that many Dbps do not display canonical helicase activity in *in vitro* assays and lack robust enzymatic activity assays in cells (Linder & Jankowsky, 2011), I wanted to develop assays to measure AMP-acrylamide binding in cell lysate. Such assays would be extremely useful to ensure Dbps were fully bound by compound when using this platform to study novel functions of these proteins and develop new assays for them, particularly if an LC-MS is not available for quantification of labeling.

I first tested selective labeling of Dbps by AMP-acrylate by competing this probe with N-propargyl maleimide (NPM), a semi-selective cysteine probe (Wilson et al, 1989; Russell et al, 2001). NPM can be clicked to TAMRA-N_3 and binding of AMP-acrylate can be read out as a reduction of in-gel fluorescence. NPM non-selectively labels cysteines in many ATP-binding pockets so if AMP-acrylate selectively labels Dbp^{ES}, I expect competition with AMP-acrylate will

cause a reduction in fluorescence of the Dbp band but no change in other bands in a whole cell lysate (**Figure 2.8 A**). NPM successfully competed AMP-acrylate labeling in vitro when tested against purified DDX3 ES and Ste13 ES, an *S. Pombe* ortholog with 83% sequence identity to DDX6 (Figure 2.8 B-C). NPM labeling could not be observed by in-gel fluorescence or electrophoretic mobility shift assay (EMSA) when tested in lysate but was restored when FLAG-tagged DDX6 was immunoprecipitated first (**Figure 2.8 D-E**). However, this assay was not particularly robust.

Cellular thermal shift assay (CETSA) is another option to assay binding of compounds in cellular lysate without enzymatic activity (Molina et al, 2013; Jafari et al, 2014). Ligand binding typically stabilizes (or sometimes destabilizes) the protein, which can be seen as a change in the protein's melting temperature in cell lysate. I first determined the melting (or aggregation) temperature for several Dbps of interest (**Figure 2.9**). Treatment of AMP-acrylate prior to heating lysate showed stabilization of DDX3 ES and destabilization of DDX6 ES but had no effect on DDX3 WT or DDX6 WT (**Figure 2.10 A-B**). However, extremely high concentrations of AMP-acrylate had to be employed to see stabilization and it could not be replicated every time. Using a western blotting readout for CETSA introduces many opportunities for small pipetting errors to occur and, after many optimization attempts, we determined that this assay was not robust enough in our hands to be used further.

Screening potential mammalian codon optimality sensors with fluorescent reporters

To probe the potential role of DDX6 in sensing codon optimality and causing translational repression of transcripts with less optimal codon usage, we designed reporters that expressed mCherry and had an internal ribosomal entry site (IRES) before a green fluorescent protein (GFP)

construct. Using published codon optimality scores for K562 cells (Wu et al, 2019), we designed the mCherry reporters to either use the most optimal codon for each amino acid (OPT) or the least optimal codon for each amino acid (NON) (**Figure 2.11 A**). We used lentivirus to stably integrate these reporters into separate K562 CRISPRi cells and sorted each cell line twice to obtain a pure population. Stable OPT and NON cell lines were analyzed by flow cytometry (**Figure 2.11 B**) and qPCR (**Figure 2.12 C**) and showed clear differences in mCherry expression, as expected.

We then used lentivirus to knockdown DDX6 in both the OPT and NON cell lines to see if mCherry expression of the NON, but not the OPT, reporter changed, which would suggest that DDX6 played a role in translational repression of only the de-optimized reporter. We initially saw that the mCherry/GFP ratio changed in NON cells with a DDX6 knockdown (Figure 2.12 A-B) but not in OPT cells (Figure 2.12 C). However, when we more closely examined mCherry and GFP levels separately, we saw that the DDX6 KD led to a slight reduction in GFP expression, rather than an increase in mCherry expression as we had hypothesized (Figure 2.12 D-E). We also knocked down other genes that have previously been implicated in mRNA decay and translational repression in mammalian cells (CNOT3, CNOT6, DCP2, PATL-1) but none of these knockdowns displayed an increase in mCherry/GFP expression for the NON reporter that we expected to indicate a codon optimality sensor (Figure 2.13).

Discussion

This chapter details the ongoing work to develop a covalent complementary inhibition strategy that could be used to study electrophile-sensitive versions of Dbps in their endogenous context. A new scaffold that bound the ATP-binding site in Dbps could not be identified in a screen

of kinase inhibitors of a DEL screen. The top hits from both screens were not pursued after several validation assays because they only displayed activity when used in high micromolar concentrations. After this LC-MS competition screen and a similar screen against the LRRK2 GTPase domain, we realized that a major difficulty encountered by using this as a screening assay is that it is not compatible with detergent. Between the lack of detergent and the high compound concentrations necessitated by the screen, it is likely that many the molecules screened were above their critical aggregation concentration, which can lead to false positives. Future screens should be accompanied by counter screens using dynamic light scattering (DLS) to check for compound aggregation and should make use of detergent-compatible assays to decrease the likelihood of this problem occurring in the first place. The ATP-binding pocket of Dbps and other ATPases is also notoriously difficult to target with cell-permeable molecules because it is so highly charged and there are no known uncharged pharmacophores that could substitute for the di- or tri-phosphates of ADP and ATP. Using a cleavable protecting group and AMP-based scaffold modeled after antivirals appears to be a more viable strategy for making a cell-permeable version of AMP-acrylate and we believe that more work improving the synthetic route and purification of adefovir-methacrylate could yield sufficient quantity and purity of this molecule to be tested *in vitro* and *in cellulo* and used with electrophile-sensitive Dbp mutants.

We also developed several cell lines that can be used in conjunction with the covalent complementary strategy or separately to probe the roles of several Dbps. We were able to make several DDX6 knockout clones with CRISPR and stably expressed tagged versions of WT or ES DDX6 in these cells, which could also be used with confocal microscopy to examine the sub-cellular localization of DDX6 under different conditions. We worked on two assays to detect AMP-acrylate labeling in cells: NPM competition and CETSA. After several rounds of

optimization for each of these assays, we determined that neither was robust enough be carried forward. Both assays were time-consuming, low throughput, and dependent on qualitative Western blot readouts, which was not ideal.

Finally, we developed stable cell lines expressing codon optimality reporters and CRISPRi machinery and used these to show that DDX6 does not appear to be the mammalian sensor of codon optimality. We believe that this type of reporter could be useful for other experiments. However, using a construct with an IRES GFP means that both the codon optimality/stability reporter (mCherry) and the fluorescent protein for translation normalization (GFP) are expressed on the same transcript, so any changes to the half-life of the transcript will affect the stability of both fluorescent reporters equally. This issue may be partly mitigated by the fact that elongation is expected to occur more slowly on the nonoptimal mCherry, whereas the GFP on both the NON and OPT reporters has the same (optimized) nucleotide sequence and therefore elongation should occur at equal speeds for both GFPs. However, flow cytometry data suggests this only plays a minor, if any role, in protein expression for these reporters as compared to transcript abundance. Further work looking into codon optimality could be improved by creating constructs with a bi-directional promoter so both mCherry and GFP should be transcribed at equal levels but on separate transcripts so transcript stability will not be a confounding factor.

Materials & Methods

Cell culture

HEK293T cells were purchased from ATCC and K562 CRISPRi cells were a gift from the lab of Luke Gilbert at UCSF. Cells were maintained at 37°C in 5% CO_2. 293T cells were cultured

in DMEM (Gibco) and K562 cells were grown in RPMI (Gibco); both were supplemented with 10% heat-inactivated FBS (Axenia Biologix).

Plasmids and cloning

DDX3 plasmids for bacterial expression were constructed in the pHMGWA vector as previously described (Floor et al, 2016; Barkovich et al, 2018). Ste13 plasmids for bacterial expression were constructed in a pET vector and Dhh1 plasmids were constructed in a pMCSG7 vector. DDX3 and DDX6 3x-FLAG-tagged plasmids for mammalian expression were constructed in a pCMV vector with Kan and Neo resistance. SNAP-DDX6, TagRFP-T-DDX6, and pWPI plasmids were obtained from Addgene. The CRISPRi guide vector pLG1 was a gift from the lab of Luke Gilbert at UCSF. Genes and gene blocks were ordered from Twist Biosciences, oligonucleotides were ordered from Eton Biosciences and IDT, modified sgRNAs were ordered from Synthego. All plasmids were Sanger sequenced by Elim Bio and Quintara to ensure the correct sequence.

DDX3 recombinant protein purification

DDX3 S228C D1 (132-406) or D1D2 (132-607) and Ste13 (30-424) were expressed and purified as previously described (Floor et al, 2016; Barkovich et al, 2018). Purified protein was flash frozen in liquid N_2, and stored at -80°C.

X-ray crystallography

Purified DDX3 S228C (132-607) was incubated with 100uM AMP-acrylamide at 4°C until the protein appeared fully labeled by liquid chromatography-mass spectrometry (LC-MS). The

protein was then purified by gel filtration across a Superdex 200 column on an AKTA Pure high performance liquid chromatography (HPLC) system. The column was equilibrated in 20 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, and 0.5 mM TCEP. Collected fractions were analyzed by UV and Coomassie gel stain for protein, combined, concentrated, flash frozen in liquid N_2, and stored at -80°C. Crystals were grown in conditions as previously described (Floor et al, 2016), namely 6% PEG-3000 and 100 mM sodium citrate, pH 5.5. Crystallization was achieved by hanging drop vapor diffusion at 22°C and occurred within 48 hrs. Crystals were looped and stored in liquid N_2 for data collection on Beamline 8.2.2 at the Advanced Light Source (LBNL, Berkeley, CA). Data were indexed and integrated on iMosFlm (Battye et al, 2012), scaled with Scala (Evans, 2006), phased using molecular replacement with PHASER (McCoy et al, 2007) using 5E7J from the Protein Data Bank (PDB) as a search model, and built and refined on PHENIX (Adams et al, 2010) and Coot (Emsley et al, 2010). Structures were visualized using PyMol (Schrodinger) and UCSF Chimera (Pettersen et al, 2021). The atomic coordinates and structure factors for the reported crystal structure were deposited as accession number 6CZ5 with the PDB. Data collection and refinement statistics can be found in **Table 2.1**.

Adefovir-methacrylate synthesis

Adefovir-methacrylate was prepared following the synthetic scheme in **Figure 2.3 C**. Adefovir (AK Scientific) and 1 equivalent 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) were dissolved in tetrahydrofuran (THF) under Ar in a 4 mL scintillation vial with sonication and stirring. In a separate vial, chloromethyl pivalate (Sigma) was dissolved in THF and added slowly to the reaction vial. The reaction was run overnight and monitored by TLC and LC-MS and the

intermediate product (1) was purified by reverse phase HPLC (0-30% MeCN/H_2O) to yield pure compound. The intermediate product (1) was resuspended with 1 equivalent DBU in THF under Ar in 4 mL vial and methacrylic anhydride (prepared as previously described in Barkovich et al, 2018) in THF was slowly added dropwise. The reaction was heated to 80°C and run overnight. The final product (adefovir-methacrylate) was detected by LC-MS but could not be purified by reverse-phase HPLC in sufficient quantity to be tested in esterase cleavage or *in cellulo* assays.

LC-MS competition screen

250 nM DDX3 S228C (132-406) in 10 mM Tris pH 7.5, 100 mM NaCl, and 1 mM $MgCl_2$ was incubated with 50 μM compound (Selleck Chem 378 Kinase Inhibitor Library) for 2 min at 4°C in 96 well plates. 5 μM AMP-acrylamide was added and then the reaction was quenched with 5% formic acid after 6.5 hrs. The extent of modification by AMP-acrylamide (+401 adduct) was determined by whole protein mass spectrometry on a Waters Acquity UPLC/G2-XF QTOF. Plates were maintained at 4°C over the course of the experiment, a wash sample of buffer only was run after every 5 samples, and samples containing 10μM AMP or DMSO instead of an inhibitor were run on each 96-well plate as positive and negative controls, respectively. Reported data points are from single replicates, the fraction of modified protein was determined using a modified python script developed by Adolfo Cuesta in the Taunton lab at UCSF, and data was plotted using Prism software (GraphPad).

Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) was performed as previously described (Niesen et al, 2007; Barkovich et al, 2019). 4 μM DDX3 (132-406) was incubated in 20 mM HEPES pH 7.5,

150 mM NaCl, and 5 mM $MgCl_2$ along with SYPRO Orange. Reaction mixtures were incubated at room temperature with compound or 0.2% DMSO 5 min before the start of the assay. Each condition was run in triplicate on a Bio-Rad CFX96 Real-Time PCR machine and curves were fit using Prism (GraphPad).

RNA duplex unwinding assay

RNA duplex assays were performed as previously described (Jankowsky & Putnam, 2010; Barkovich et al, 2019). 1 μM DDX3 (132-607), 1 nM duplex RNA, 2 mM ATP/$MgCl_2$, and indicated compound concentration were incubated in reaction buffer: 20 mM HEPES pH 7.5, 100 mM NaCl, 500 μM $MgCl_2$, 1 mM TCEP, 0.01% (v/v) NP-40, 5% glycerol, in diethyl pyrocarbonate (DEPC) treated H_2O. The RNA was radiolabeled in advance and the duplex strand sequences are 5′-GCUUUACGGUGC-3′ and 5′-GAACAACAACAACAACCAUGGCACCGUAAAGC-3′. For all the reactions that included chemical inhibitors, the inhibitor, RNA, and protein were incubated 5 min at room temperature before ATP-Mg was added. Reactions were quenched with 1:10 of a 1% SDS 50 mM EDTA stock and run on a 15% PAGE-TBE gel at 5W for 30 min on ice. Gels were dried, transferred, and imaged on a Typhoon imager. Results were quantified on ImageJ software (NIH).

DNA-encoded library screen

The DNA-encoded library (DEL) screen was performed in conjunction with WuXi AppTech using their DELOpen library of approximately 2.8 billion compounds and following the protocol provided with the library. The library was screened against His-MBP-DDX3 S228C D1 and counter-screened against His-MBP-DDX3 S228C D1 fully labeled by AMP-acrylamide and

His-tagged maltose-binding protein (MBP). MBP was used on DDX3 constructs to ensure solubility. In short, proteins were tested with His-Pure Ni-NTA beads to ensure on-bead capture as shown on a Coomassie-stained SDS-PAGE gel. Beads were suspended in wash buffer and incubated for 30 min at 25°C with each of the proteins for the 3 different screening conditions, as well as a non-targeting condition (NTC). Beads were washed and resuspended in the DEL library solution, incubated 1 hr at 25°C, then washed several times with a new buffer before incubating 10 min at 95°C to elute. A fraction of this first round of selection was saved and the rest used for two more rounds of library selection against fresh mixtures of protein immobilized on beads. The supernatant from the second and third rounds of selection was also saved and samples from all three rounds of selection against each protein was stored at -80°C and sent to WuXi for deep sequencing. Based on the sequencing results provided, 5 compounds from different synthetic classes that appeared as hits within the library were ordered for off-bead synthesis and further analysis.

DDX6 KO cell line generation

Nucleofection was done using the Amaxa SF Cell Line Nucleofector Kit (Lonza) following the kit protocol. 2e5 293T cells were suspended in 20 μL nucleofector solution (Lonza SF solution with added 293T supplement). 100 pmol Cas9 was added to 40 pmol of each of the 3 sgRNAs very slowly and incubated 20 min at 22°C before it was mixed with the cell solution in the Lonza cuvette. Nucleofection was performed on a Lonza Nucleofector 4D using the SF cell line program CM-130. Cells were rested in the cuvettes for 10 min post-nucleofection, then plated in 1 mL media in 12-well plates. Cells were grown out and knockout efficiency was assessed by Sanger sequencing. Individual clones were isolated and sequenced by diluting cells to 1 cell/200 μL and

plating 100 µL/well in 96-well plates, growing out 2 wks, identifying wells with single colonies, growing up colonies, and Sanger sequencing. Two separate DDX6 knockout clones were generated, verified by Western blotting, and stored in 5% DMSO at -160°C. Purified Cas9 was ordered from the UC Berkeley QB3 MacroLab and modified sgRNAs were ordered from Synthego; sgRNA sequences can be found in **Table 2.2**.

Immunoprecipitation and NPM labeling competition

293T cells were transfected with DDX6-FLAG (WT or T144C) or DDX3-FLAG (WT or S228C), harvested after 48 hrs, lysed 10 min on ice in 100 mM HEPES pH 7.5, 150 mM NaCl, 0.1% NP-40 lysis buffer with 1x cOmplete EDTA-free protease inhibitor cocktail and 1x PhosSTOP inhibitor cocktail, then clarified and normalized to 2.5 mg/mL protein. Lysate was incubated with magnetic FLAG beads at 4°C for 90 min, then beads were washed 5x in lysis buffer. Beads were washed 2x in 1 mg/mL FLAG peptide in lysis buffer and the supernatant was combined for the immunoprecipitation (IP) fraction. IP protein or cell lysate was incubated with 10 µM AMP-acrylate for 15 min at 22°C, then 1 µM N-propargyl maleimide (NPM) was added for 15 min before the reaction was quenched with 1 µL of 50 µM ß-mercaptoethanol (BME). TAMRA-N_3 or *m*-PEG-N_3 was added via a copper-catalyzed click reaction run for 1 hr at 22°C in the dark. The click reaction was quenched with SDS-loading dye and samples were run by SDS-PAGE gel. In-gel fluorescence was read out on a Typhoon imager and analyzed by ImageJ. Immunoblotting was performed as described in Chapter 3 and images were analyzed by ImageStudioLite software (Li-COR Biosciences).

Cellular thermal shift assay

Cellular thermal shift assay (CETSA) was performed as previously described (Jafari et al, 2014) using 293T or 293T TagRFP-T-DDX6 cells. 293T cells transfected with DDX3-FLAG or DDX6-FLAG were harvested after 72 hrs. Cells were grown in 15 cm dishes, rinsed with ice cold PBS (Gibco), and manually detached. Cell pellets were rinsed 2x with PBS before being lysed with ice cold lysis buffer without detergent (100 mM Hepes (pH 7.5) and 150 mM NaCl with mini-cOmplete Protease Inhibitor Cocktail Tablets and PhosSTOP) by freeze-thawing 3x in liquid N_2 and clarified by spinning at 20,000g for 20 min at 4°C. Lysates were normalized to 2.5 mg/mL protein by a Bradford assay. Standard melt curves for each protein tested were generated by heating lysates in 40-67°C in 3°C intervals. Lysates were incubated 60 min with 50-200 μM AMP-acrylate, then heated 3 min at 52°C, cooled 5 min at 4°C, centrifuged to remove aggregates, and run on an SDS-PAGE gel. Immunoblotting was performed as described in Chapter 3, images were analyzed by ImageStudioLite (Li-Cor Biosciences) and melt curves were fit with Prism (GraphPad).

Transient transfection

Transient transfections in 293T cells were performed following the Lipofectamine 3000 protocol (Thermo Fisher). 293T cells were plated in 10 cm dishes or 6-well plates and grown until 70-90% confluency. Lipofectamine reagent was diluted in Opti-MEM (OMEM) and plasmid plus P-3000 reagent were mixed, diluted in OMEM, then mixed with the diluted lipofectamine reagent and incubated for 5 min at room temperature. This DNA-lipid complex was carefully added dropwise to cells and mixed by gentle shaking. Cells were grown 24-72 hrs then used for downstream analysis. DNA and lipid amounts were determined empirically based off well size and manufacturer's recommendations.

Codon optimality reporters

Codon optimality reporters were designed using published *H. sapiens* codon optimality data (Wu et al, 2019). For each amino acid, the most optimal codon option was used for the "Optimal" mCherry reporter and the least optimal codon option was used for the "Non-optimal" mCherry reporter using a python script written with the help of Jack Stevenson in the Shokat lab. mCherry sequences are listed in **Table 2.3**. Codon optimality reporter constructs were cloned into the vector pWPI (Addgene #12254) that also contained an IRES EGFP. Lentivirus were generated also previously described (Lou et al, 2022) and transduced in K562 CRISPRi cells as described in **Chapter 3**. "OPT" and "NON" cell lines containing either the codon optimized mCherry or non-optimized construct, respectively, were selected through two rounds of cell sorting gated on mCherry and GFP fluorescence. Stable cell lines were routinely monitored by flow cytometry and stored with 5% DMSO at -160°C.

Lentivirus/stable cell line selection

tagRFP-T-DDX6 and SNAP-DDX6 constructs were from Addgene (#119947 and 119948) and electrophile-sensitive point mutations (T144C) were generated by site-directed mutagenesis polymerase chain reaction (PCR) cloning in Stbl3 *E. coli* strains. Lentivirus were generated for tagRFP-T-DDX6 and SNAP-DDX6 constructs and 293T DDX6 KO cells were transduced as described in **Chapter 3**. DDX6-RFP cells were selected by two sequential rounds of fluorescence-activated cell sortings (FACS) on a BD FACSAria II. SNAP-DDX6 cells were also selected by two rounds of FACS using the SNAP-Cell TMR-STAR substrate. K562 CRISPRi OPT and NON cell lines were also transduced with CRISPRi guides as previously described (Lou et al, 2022). Cells were selected with puromycin and verified by BFP expression by flow cytometry. CRISPRi

sgRNA sequences are listed in **Table 2.4**. All stable cell lines were routinely tested and stored with 5% DMSO at -160°C.

Flow cytometry

Flow cytometry was performed on a benchtop Attune II and an Attune Nxt as described in **Chapter 3**.

Figures

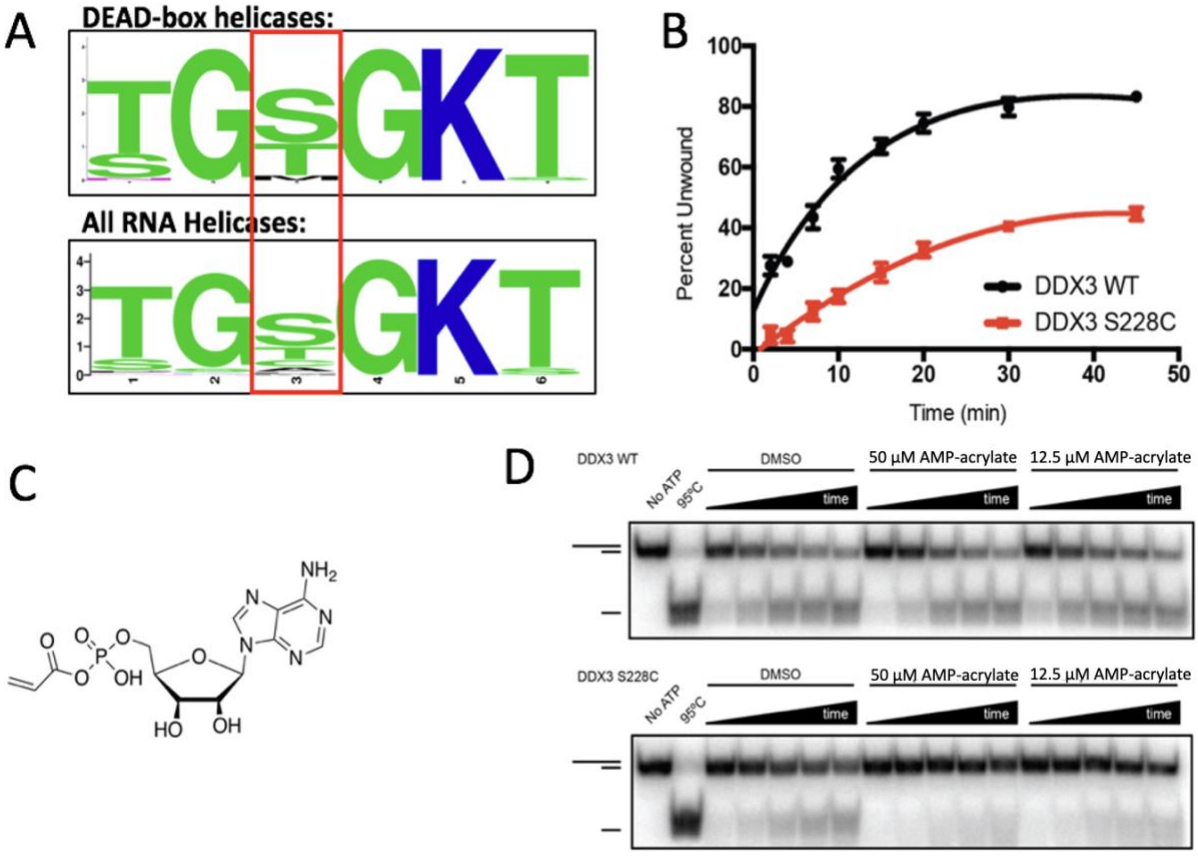


Figure 2.1 Covalent complementary approach to study DEAD-box proteins. A) Alignment of DEAD-box proteins and all RNA helicases along the P loop to identify the site of cysteine mutation (in red box). B) DDX3 S228C displays a 3-fold decrease in unwinding activity compared to DDX3 WT. C) Structure of AMP-acrylate. D) AMP-acrylate does not inhibit WT DDX3 (top). Dose-dependent inhibition of unwinding activity by AMP-acrylate with DDX3 S228C (bottom).

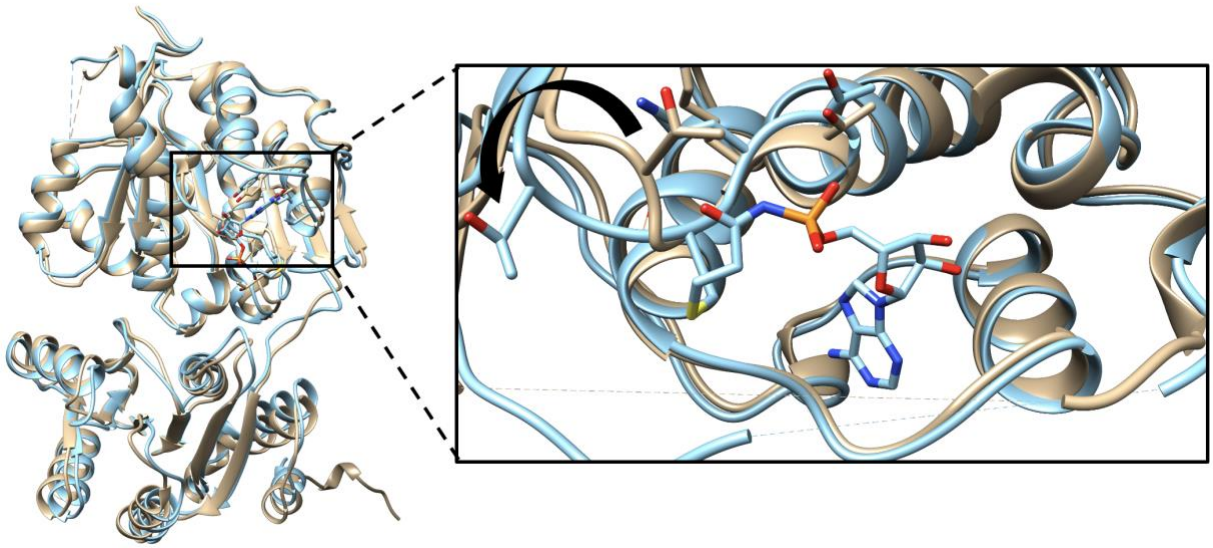


Figure 2.2 Crystal structure of DDX3 (132-607) S228C bound to AMP-acrylamide. DDX3 ES bound to AMP-acrylamide (6ZC5) is shown in blue, overlaid with DDX3 WT (132-607) bound to AMP (5E7J). Inset shows close up of the ATP-binding site with AMP-acrylamide covalently bound at Cys228. Thr226 on the P-loop is shifted 9.4 Å to accommodate the orientation of the covalent bond formation.

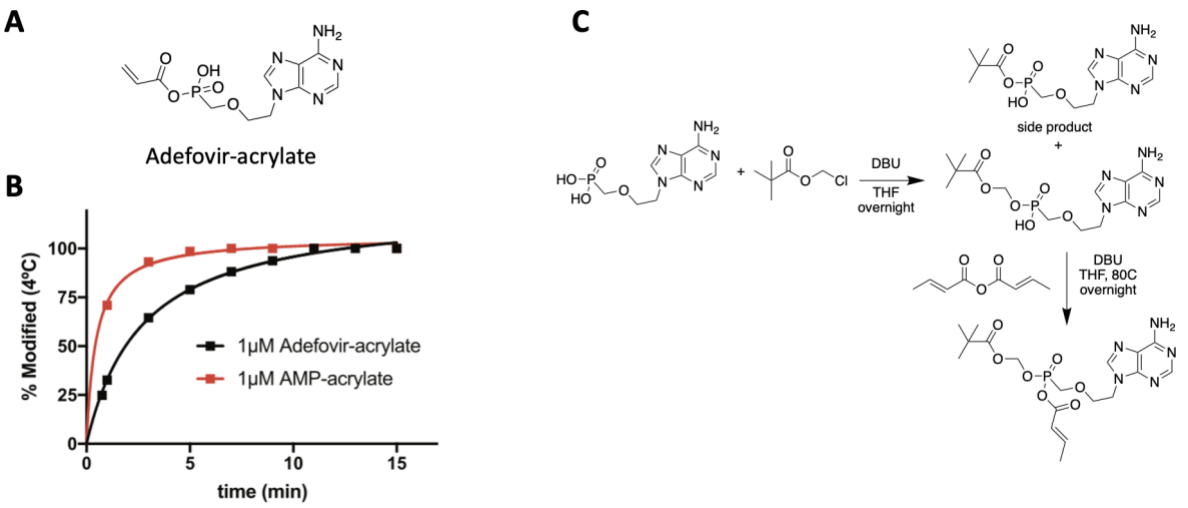


Figure 2.3 Caged phosphate strategy for AMP-based inhibitors. A) Adefovir acrylate was previously synthesized by Krister Barkovich in the Shokat lab. B) Adefovir-acrylate showed full labeling of DDX3 S228C *in vitro*. C) Synthetic scheme for caged Adefovir-methacrylate.

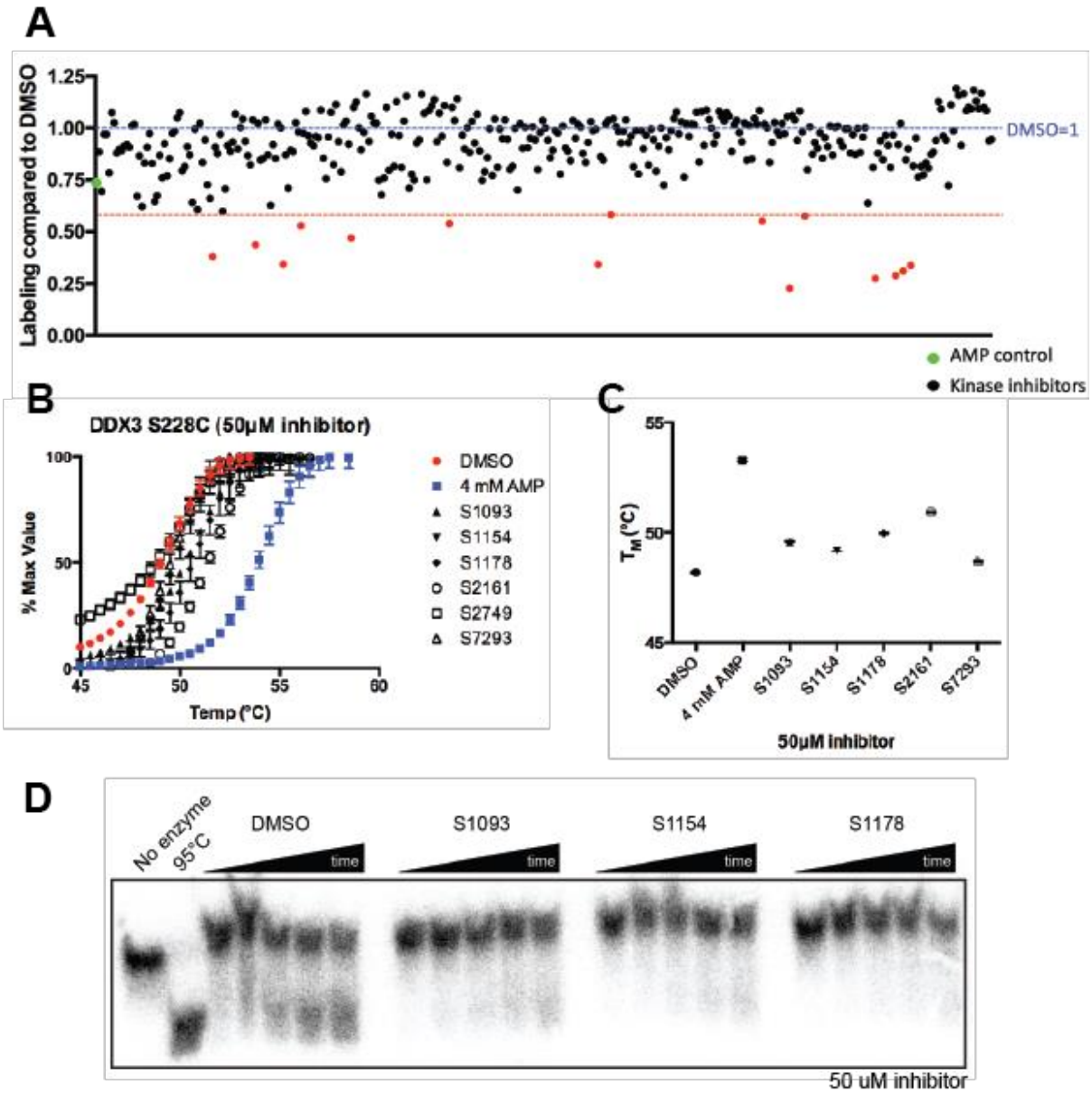


Figure 2.4 Kinase inhibitor library screen. A) Scatter plot of Selleck Chem kinase inhibitor library compounds inhibition of AMP-acrylamide labeling of DDX3 S228C. Blue line is labeling without kinase inhibitor (DMSO control) normalized to 1. Red line represents 2 standard deviations below average of all tested compounds. Compounds marked in red were tested further to check for a dose response in inhibition of labeling assays. B) Top 6 hits were tested for stabilization of DDX3 S228C by DSF. C) DDX3 S228C melting temperatures calculated from Panel B. D) 3 hits were tested for inhibition of DDX3 S228C duplex unwinding activity *in vitro* over 45 min.

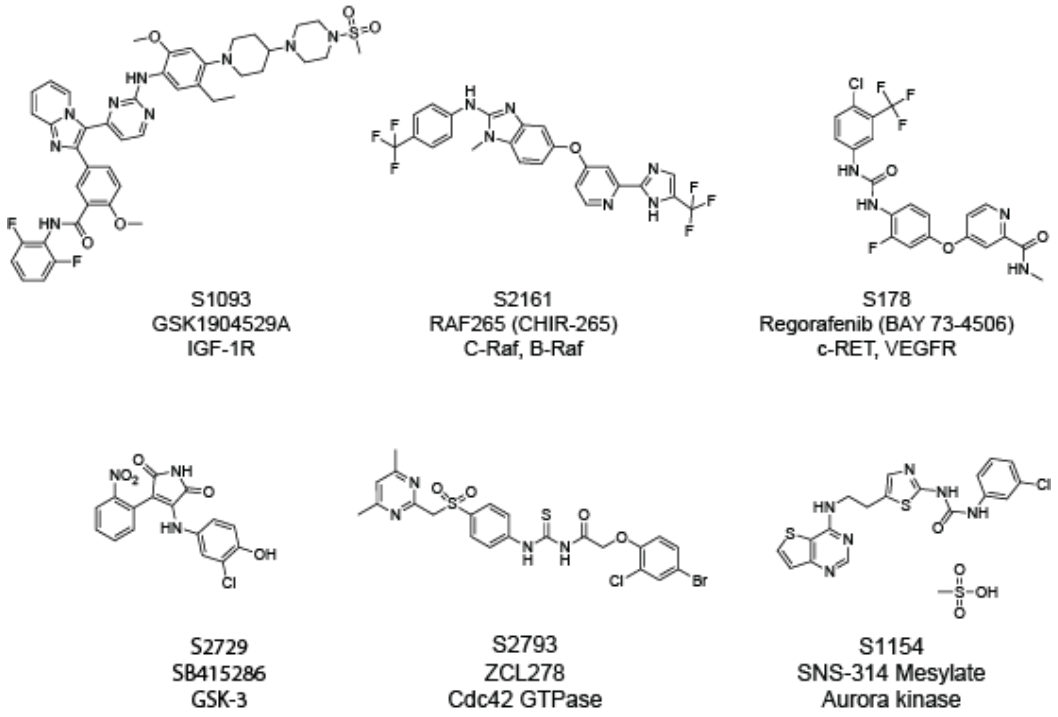


Figure 2.5 Structures of Selleck Chem library hits. Structure of the top six hits identified from the kinase inhibitor library screen against DDX3 S228C. Molecules are shown with their Selleck Chem library ID, commercial or company name, and primary known kinase or GTPase target.

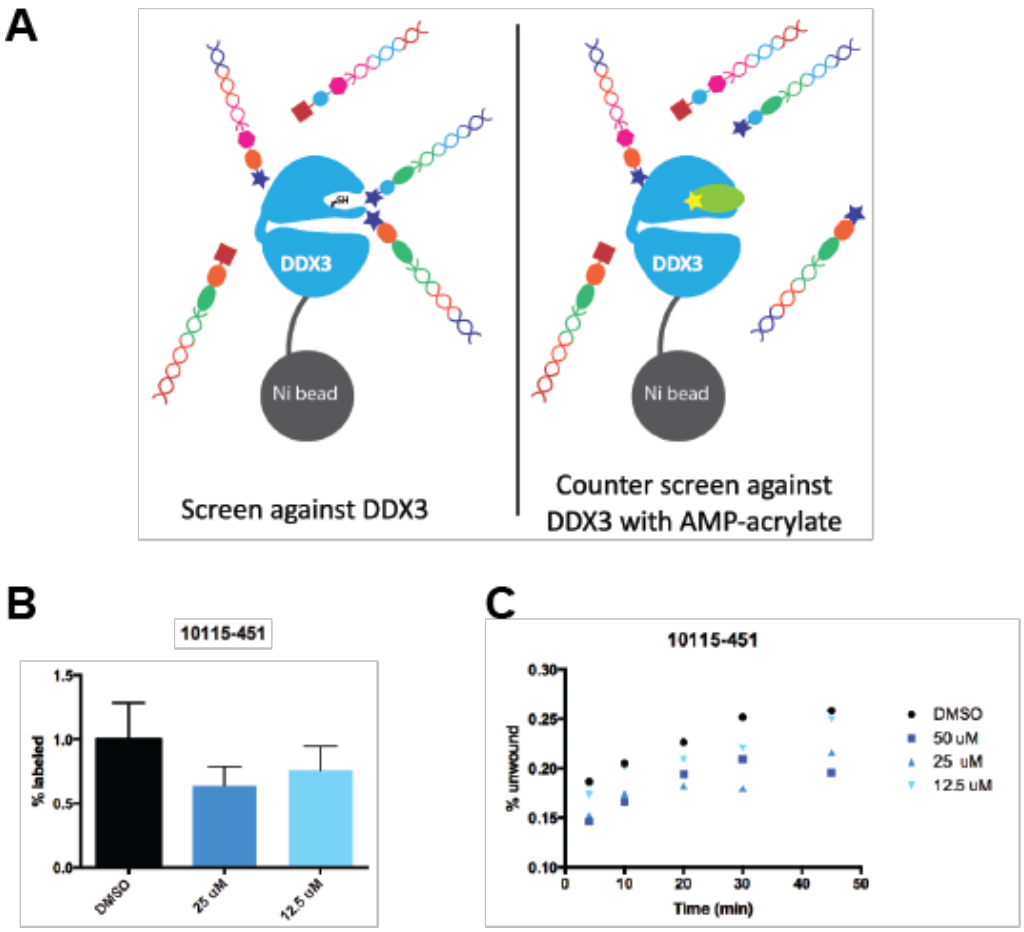


Figure 2.6. DNA-encoded library screen against DDX3 S228C. A) Cartoon interpretation of DEL screen and counter screen against His-MBP-DDX3 (132-406) ES and His-MBP-DDX3 (132-406) ES bound to AMP-acrylamide. B) Molecule 10115-451 decreases AMP-acrylamide labeling of DDX3 ES. C) Molecule 10115-451 inhibits DDX3 unwinding activity in an *in vitro* duplex unwinding assay.

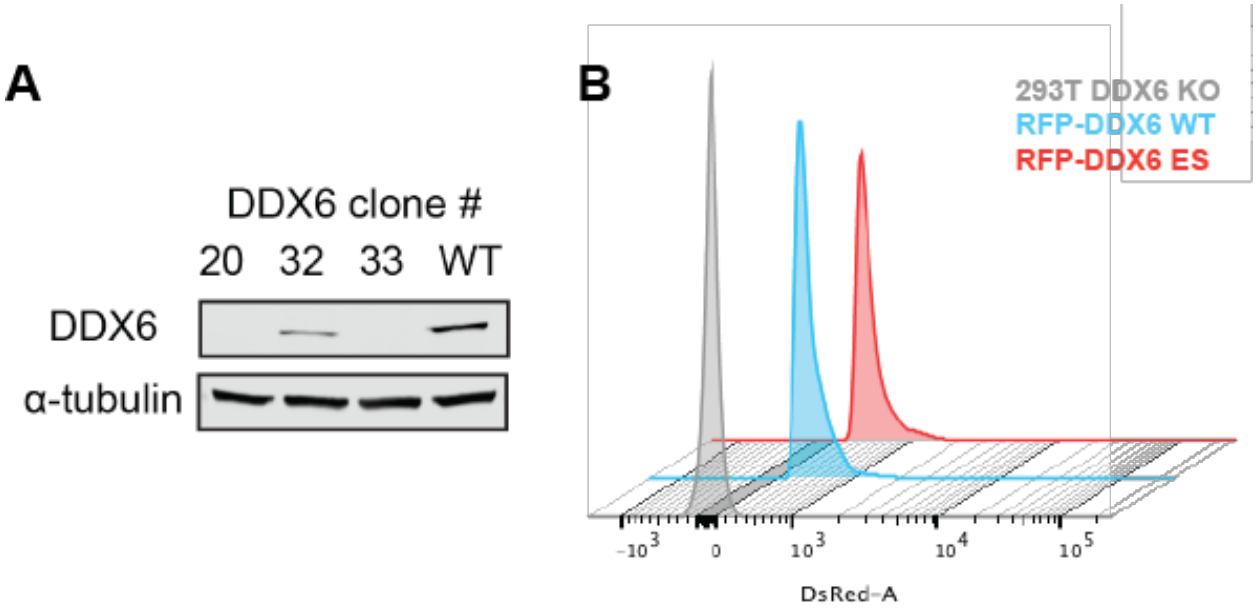


Figure 2.7 DDX6 engineered cell lines. A) 293T clones with DDX6 knockouts were selected and tested by Western blotting. B) DDX6 knockout cell lines were used a background for RFP-DDX6 WT or ES expression via lentivirus. Stable cell lines were identified and sorted by FACS.

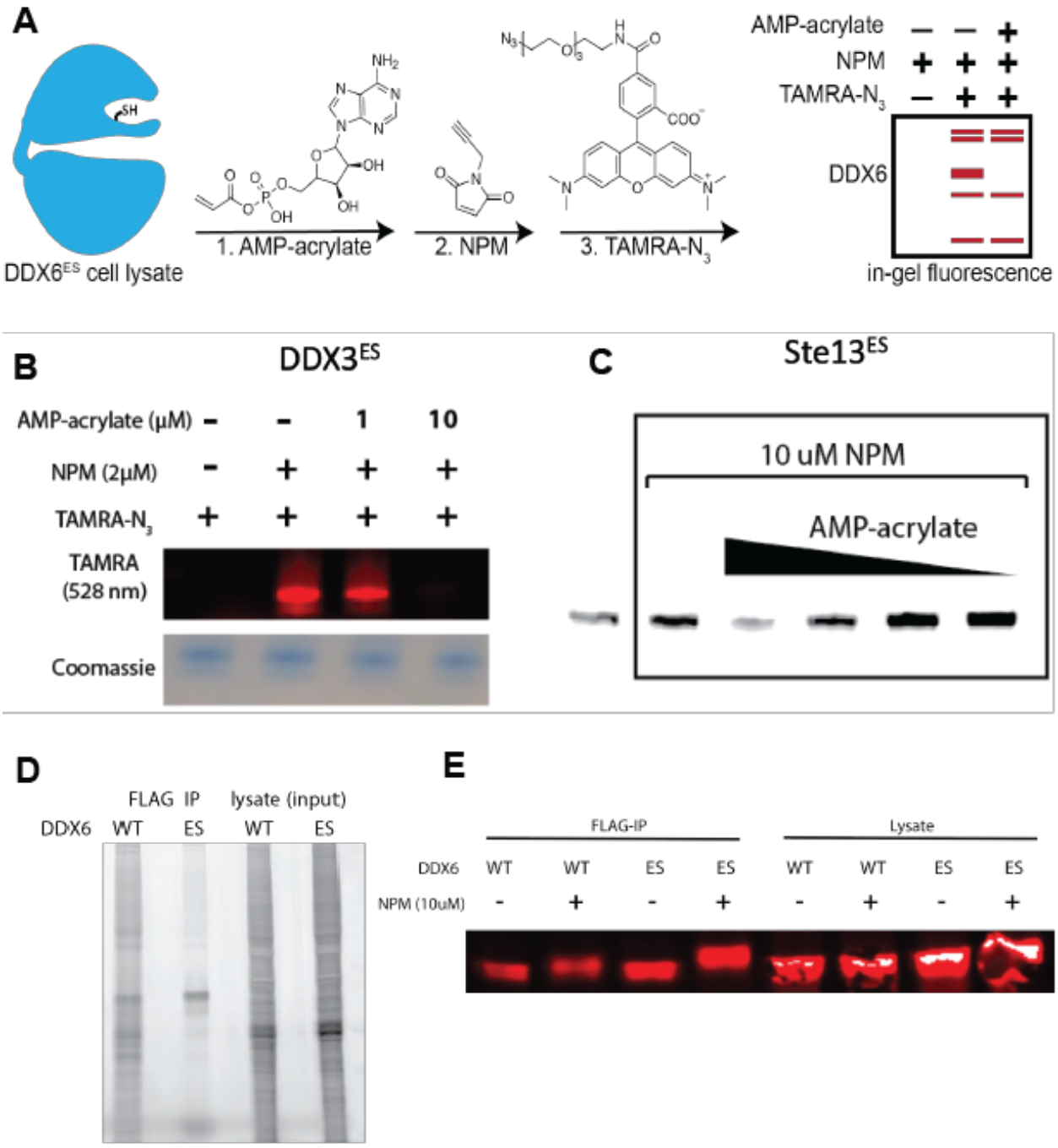


Figure 2.8 NPM competition assay for detection of AMP-acrylate labeling. A) Schematic of assay workflow. B) Competition of NPM labeling of purified DDX3 ES with AMP-acrylate. C) Competition of NPM labeling of purified Ste13 ES with AMP-acrylate. 1 μM protein was used in both B and C. 0.3 μM-20 μM AMP-acrylate was used in C. D) In-gel fluorescence of NPM clicked to TAMRA-N_3 in 293T lysate with DDX6-FLAG WT or ES transiently transfected. E) EMSA assay in Western blot for FLAG tag from same experiment at D.

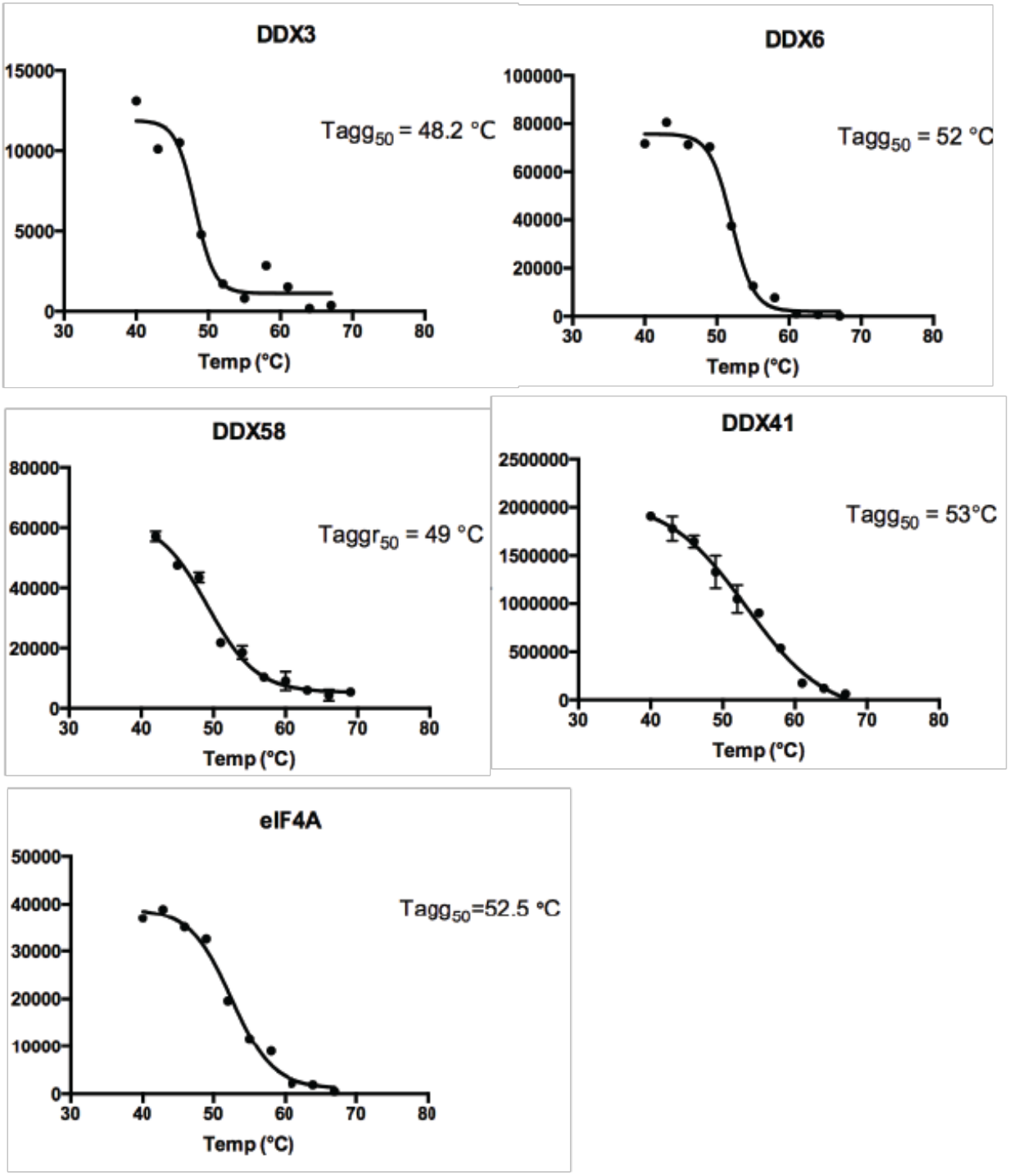


Figure 2.9 CETSA protein melt curves were determined in 293T lysate for several DExD/H-Box proteins of interest.

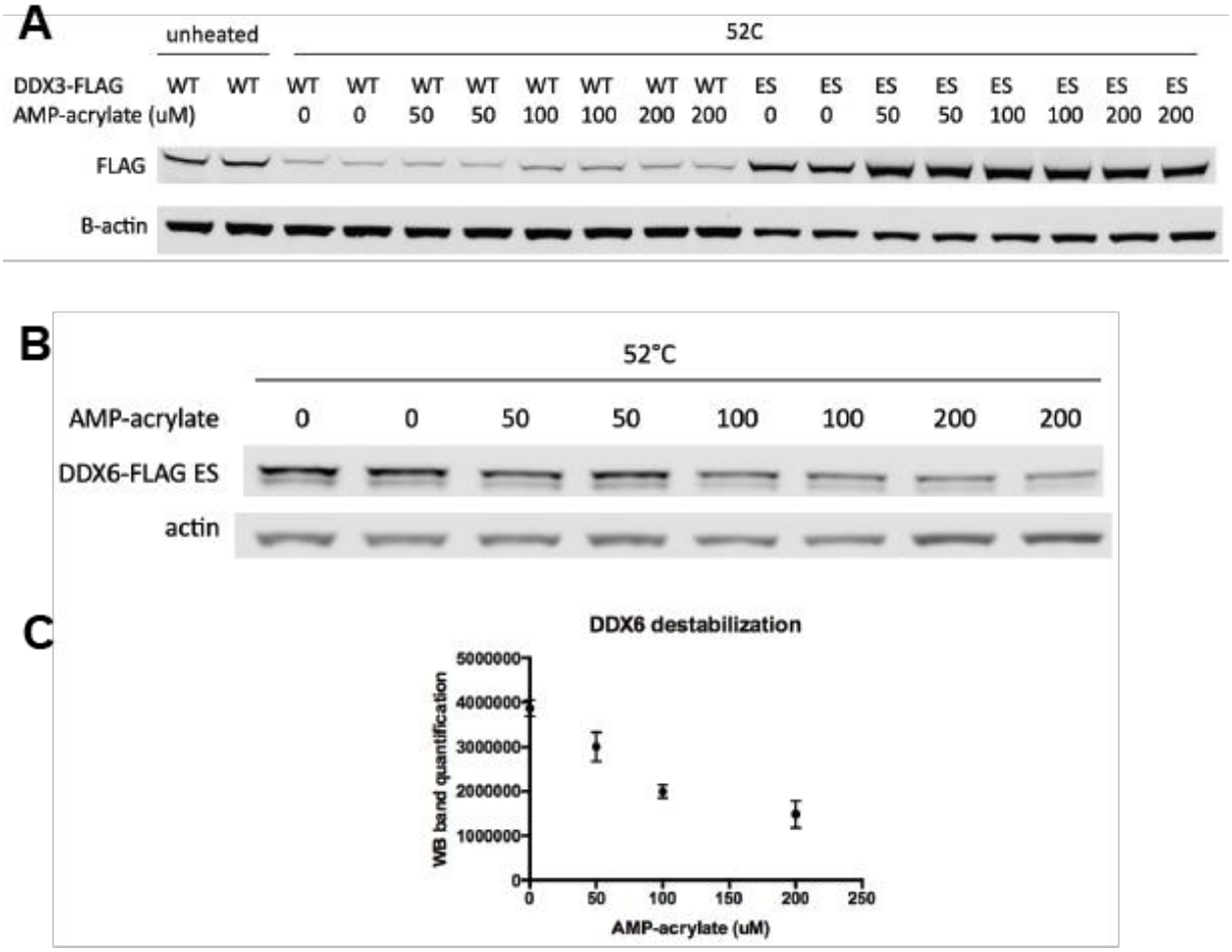


Figure 2.10 AMP-acrylate demonstrates (de)stabilization of ES Dbps by CETSA. A) DDX3-FLAG WT and ES were transiently transfected in 293T cells and lysates were treated with AMP-acrylate before being heated to 52°C. B) DDX6-FLAG WT and ES were transiently transfected in DDX6 KO 293T cells and lysates were treated with AMP-acrylate before being heated to 52°C. C) Quantification of band intensity from 3 Western blot replicates of DDX6 protein.

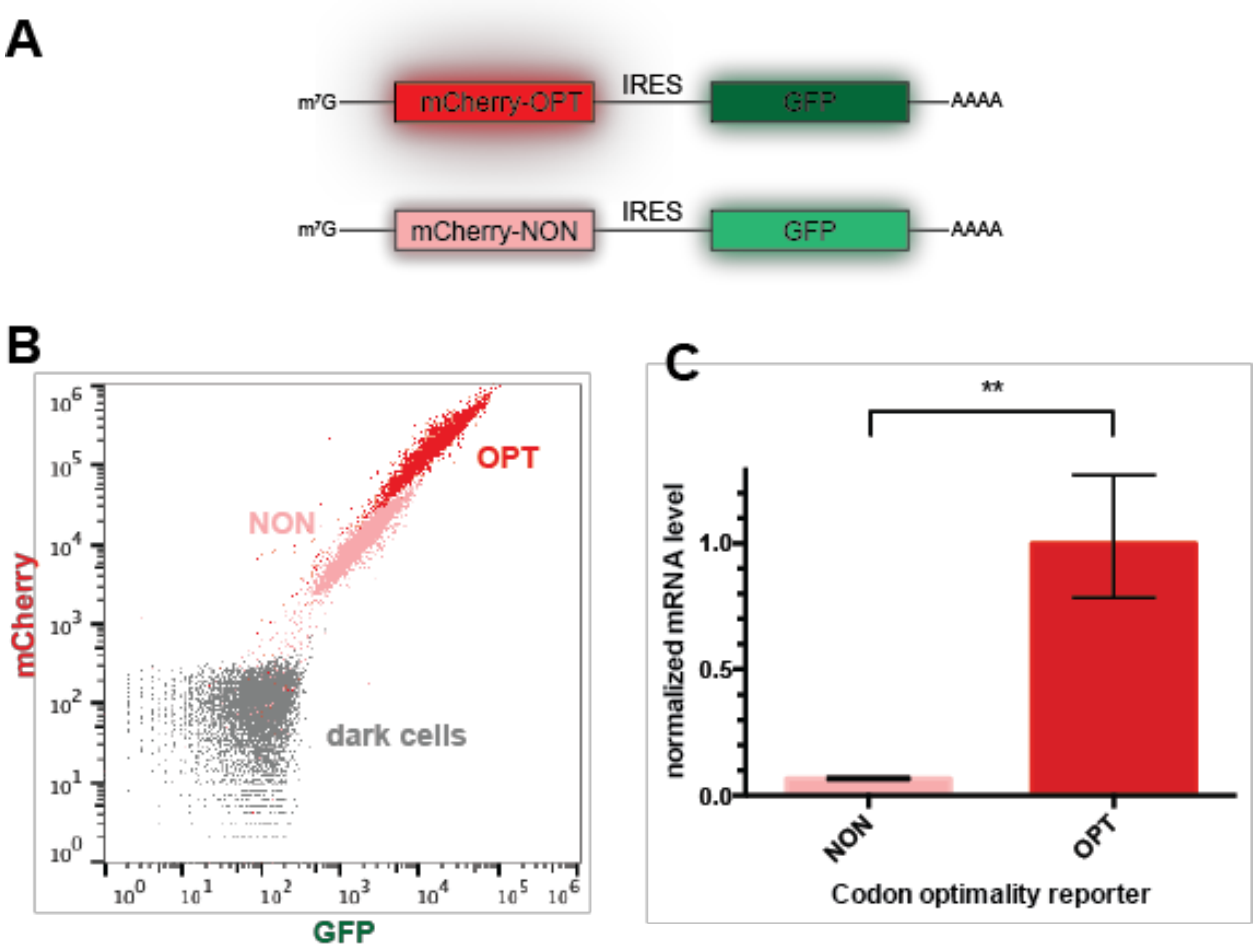


Figure 2.11 Codon optimality reporters. A) Schematic of design of "OPT" (codon optimized mCherry) and "NON" (codon deoptimized mCherry) reporters. B) Cell lines expressing either OPT, NON, or no reporter visualized by flow cytometry. C) Gene expression of OPT and NON reporters obtained by qPCR normalized to GAPDH gene expression.

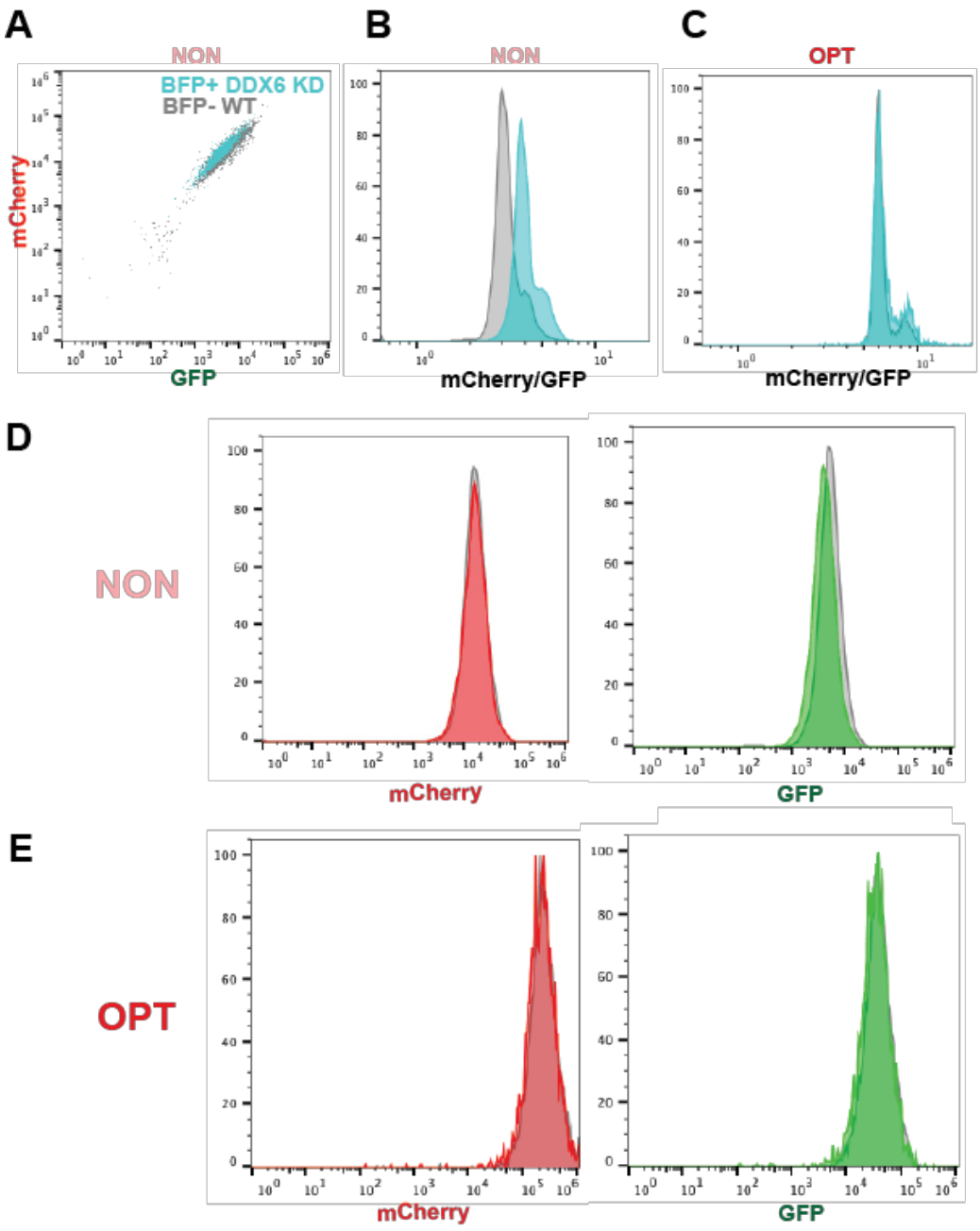


Figure 2.12 DDX6 knockdown lowers GFP expression in NON reporter but not OPT reporter. A) Scatter plot of cells expressing NON reporter with (blue) or without (gray) DDX6 knockdown. B) Same as A, but shown as a ratio of mCherry/GFP expression. C) Histogram of cells expressing OPT reporter with (blue) or without (gray) DDX6 knockdown. D) Histograms of changes in mCherry and GFP expression in cells expressing NON reporter with (red/green) or without (gray) DDX6 knockdown. E) Histograms of changes in mCherry and GFP expression in cells expressing OPT reporter with (red/green) or without (gray) DDX6 knockdown.

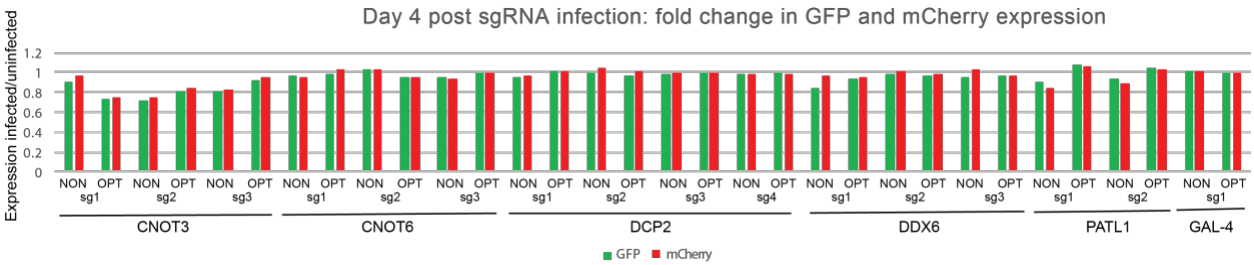


Figure 2.13 Knockdown of known translational repressors in cell lines with NON and OPT codon optimality reporters. Reporter expression is shown as fold-change of GFP and mCherry compared to cells without a knockdown and normalized to GAL-4 negative control knockdowns.

Tables

Table 2.1 X-ray data collection and refinement statistics. Table is reprinted from Barkovich et al, 2019. Atomic factors and structural coordinates for the reported crystal structure can be found in the PDB under accession number 6CZ5.

	DDX3 (132-607) – AMP-acrylamide
Data collection	
Space group	P 21 21 21
Cell dimensions	
a, *b*, *c* (Å)	53.96, 101.09, 105.69
α, *β*, *γ* (°)	90, 90, 90
Resolution (Å)	73.05-3.00 (3.16-3.00)[a]
R_{merge}, R_{meas}, and R_{pim}	0.121 (1.089), 0.135 (1.201), 0.057 (0.500)
$I/\sigma(I)$	8.5 (1.6)
$CC_{1/2}$	0.996 (0.766)
Completeness (%)	100.0 (100.0)
Redundancy	5.4 (5.6)
Refinement	
Resolution (Å)	73.05-3.00
No. reflections	12100
R_{work} / R_{free}	0.2219/0.2667
No. atoms	
Protein	3378
Ligand/ion (specify/describe)	27
Water	0
B factors	
Protein	87.39
Ligand/ion	107.20
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.50

[a] Values in parentheses are for highest-resolution shell.

Table 2.2 DDX6 KO sgRNA sequences. Modified sgRNAs were ordered from Synthego. The 20 nucleotide genome targeting sequences are appended to an 80-mer SpCas9 scaffold to create the guide RNA. Sequences contained 2'-O-Me modifications on the 3 first and last bases and 3' phosphorothioate bonds between the first 3 and last 2 bases.

sgRNA	Sequence
DDX6 sgRNA 1	A*G*U*GCCAUUAUUGAUUGUGU
DDX6 sgRNA 2	C*A*C*UGGUGGCCCUGGAGGAG
DDX6 sgRNA 3	G*G*C*CACCAGUGGGUUUCACA

Table 2.3 mCherry codon optimality reporter sequences.

Reporter	Sequence
"OPT" Codon-optimized mCherry	ATGGTGTCGAAGGGTGAGGAGGACAACATGGCTA TCATCAAGGAGTTCATGCGGTTCAAGGTGCACATG GAGGGTTCGGTGAACGGTCACGAGTTCGAGATCG AGGGTGAGGGTGAGGGTCGGCCCTACGAGGGTAC CCAGACCGCTAAGCTGAAGGTGACCAAGGGTGGT CCCCTGCCCTTCGCTTGGGACATCCTGTCGCCCCA GTTCATGTACGGTTCGAAGGCTTACGTGAAGCACC CCGCTGACATCCCCGACTACCTGAAGCTGTCGTTC CCCGAGGGTTTCAAGTGGGAGCGGGTGATGAACT TCGAGGACGGTGGTGTGGTGACCGTGACCCAGGA CTCGTCGCTGCAGGACGGTGAGTTCATCTACAAGG TGAAGCTGCGGGGTACCAACTTCCCCTCGGACGGT CCCGTGATGCAGAAGAAGACCATGGGTTGGGAGG CTTCGTCGGAGCGGATGTACCCCGAGGACGGTGCT CTGAAGGGTGAGATCAAGCAGCGGCTGAAGCTGA AGGACGGTGGTCACTACGACGCTGAGGTGAAGAC CACCTACAAGGCTAAGAAGCCCGTGCAGCTGCCC GGTGCTTACAACGTGAACATCAAGCTGGACATCAC CTCGCACAACGAGGACTACACCATCGTGGAGCAG TACGAGCGGGCTGAGGGTCGGCACTCGACCGGTG GTATGGACGAGCTGTACAAG
"NON" Codon-nonoptimized mCherry	ATGGTAAGTAAAGGAGAAGAAGATAATATGGCGAT AATAAAAGAATTTATGAGATTTAAAGTACATATGGA AGGAAGTGTAAATGGACATGAATTTGAAATAGAAG GAGAAGGAGAAGGAAGACCTTATGAAGGAACTCA AACTGCGAAATTAAAAGTAACTAAAGGAGGACCT TTACCTTTTGCGTGGGATATATTAAGTCCTCAATTTA TGTATGGAAGTAAAGCGTATGTAAAACATCCTGCG GATATACCTGATTATTTAAAATTAAGTTTTCCTGAA GGATTTAAATGGGAAAGAGTAATGAATTTTGAAGA TGGAGGAGTAGTAACTGTAACTCAAGATAGTAGTT TACAAGATGGAGAATTTATATATAAAGTAAAATTAA GAGGAACTAATTTTCCTAGTGATGGACCTGTAATG CAAAAAAAAACTATGGGATGGGAAGCGAGTAGTG AAAGAATGTATCCTGAAGATGGAGCGTTAAAAGG AGAAATAAAACAAAGATTAAAATTAAAAGATGGA GGACATTATGATGCGGAAGTAAAAACTACTTATAA AGCGAAAAAACCTGTACAATTACCTGGAGCGTATA ATGTAAATATAAAATTAGATATAACTAGTCATAATGA AGATTATACTATAGTAGAACAATATGAAAGAGCGG AAGGAAGACATAGTACTGGAGGAATGGATGAATTA TATAAA

Table 2.4 CRISPRi sgRNA protospacer sequences. Protospacer sequences were chosen from the hCRISPRi v2 Top 5 Library (Horlbeck et al, 2016) and oligos ordered and cloned into expression vector as previously described (Horlbeck et al, 2016; Lou et al, 2022; Chapter 3).

sgRNA name	hCRISPRi v2 library name	Sequence
DDX6 sgRNA 1	DDX6_+_118661671.23-P1P2	GCGGGCGGAGGTTTCCGAGG
DDX6 sgRNA 2	DDX6_+_118661800.23-P1P2	GGCGGGCAGCGGAGGAGGTA
DDX6 sgRNA 3	DDX6_+_118661817.23-P1P2	GTCGGCGGCGCCACGAGAGC

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., ... & Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*, *66*(2), 213-221.

Barkovich, K. J., Moore, M. K., Hu, Q., & Shokat, K. M. (2018). Chemical genetic inhibition of DEAD-box proteins using covalent complementarity. *Nucleic acids research*, *46*(17), 8689-8699.

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., & Leslie, A. G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 271-281.

Bordeleau, M. E., Robert, F., Gerard, B., Lindqvist, L., Chen, S. M., Wendel, H. G., ... & Pelletier, J. (2008). Therapeutic suppression of translation initiation modulates chemosensitivity in a mouse lymphoma model. *The Journal of clinical investigation*, *118*(7), 2651-2660.

Chen, M., Asanuma, M., Takahashi, M., Shichino, Y., Mito, M., Fujiwara, K., ... & Iwasaki, S. (2021). Dual targeting of DDX3 and eIF4A by the translation inhibitor rocaglamide A. *Cell chemical biology*, *28*(4), 475-486.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), 486-501.

Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallographica Section*

D: Biological Crystallography, *62*(1), 72-82.

Floor, S. N., Barkovich, K. J., Condon, K. J., Shokat, K. M., & Doudna, J. A. (2016). Analog sensitive chemical inhibition of the DEAD-box protein DDX 3. *Protein Science*, *25*(3), 638-649.

Floor, S. N., Condon, K. J., Sharma, D., Jankowsky, E., & Doudna, J. A. (2016). Autoinhibitory interdomain interactions and subfamily-specific extensions redefine the catalytic core of the human DEAD-box protein DDX3. *Journal of Biological Chemistry*, *291*(5), 2412-2421.

Freimer, J. W., Hu, T. J., & Blelloch, R. (2018). Decoupling the impact of microRNAs on translational repression versus RNA degradation in embryonic stem cells. *eLife*, *7*, e38014.

Gadek, M., Sherr, E. H., & Floor, S. N. (2023). The variant landscape and function of DDX3X in cancer and neurodevelopmental disorders. *Trends in Molecular Medicine*.

Garske, A. L., Peters, U., Cortesi, A. T., Perez, J. L., & Shokat, K. M. (2011). Chemical genetic strategy for targeting protein kinases based on covalent complementarity. *Proceedings of the National Academy of Sciences*, *108*(37), 15046-15052.

Horlbeck, M. A., Gilbert, L. A., Villalta, J. E., Adamson, B., Pak, R. A., Chen, Y., ... & Weissman, J. S. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *elife*, *5*, e19760.

Ito, M., Tanaka, T., Cary, D. R., Iwatani-Yoshihara, M., Kamada, Y., Kawamoto, T., ... & Imaeda, Y. (2017). Discovery of novel 1, 4-diacylpiperazines as selective and cell-active eIF4A3

inhibitors. *Journal of medicinal chemistry*, *60*(8), 3335-3351.

Iwasaki, S., Floor, S. N., & Ingolia, N. T. (2016). Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature*, *534*(7608), 558-561.

Iwatani-Yoshihara, M., Ito, M., Klein, M. G., Yamamoto, T., Yonemori, K., Tanaka, T., ... & Kawamoto, T. (2017). Discovery of allosteric inhibitors targeting the spliceosomal RNA helicase Brr2. *Journal of Medicinal Chemistry*, *60*(13), 5759-5771.

Jafari, R., Almqvist, H., Axelsson, H., Ignatushchenko, M., Lundbäck, T., Nordlund, P., & Molina, D. M. (2014). The cellular thermal shift assay for evaluating drug target interactions in cells. *Nature protocols*, *9*(9), 2100-2122.

Jankowsky, E., & Bowers, H. (2006). Remodeling of ribonucleoprotein complexes with DExH/D RNA helicases. *Nucleic acids research*, *34*(15), 4181-4188.

Jankowsky, E., & Putnam, A. (2010). Duplex unwinding with DEAD-box proteins. *Helicases: Methods and Protocols*, 245-264.

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., ... & Gauthier, L. D. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *BioRxiv*, 531210.

Lennox, A. L., Hoye, M. L., Jiang, R., Johnson-Kerner, B. L., Suit, L. A., Venkataramanan, S., ... & Sherr, E. H. (2020). Pathogenic DDX3X mutations impair RNA metabolism and neurogenesis during fetal cortical development. *Neuron*, *106*(3), 404-420.

Linder, P., & Jankowsky, E. (2011). From unwinding to clamping—the DEAD box RNA

helicase family. *Nature reviews Molecular cell biology*, *12*(8), 505.

Lindqvist, L., Oberer, M., Reibarkh, M., Cencic, R., Bordeleau, M. E., Vogt, E., ... & Pelletier, J. (2008). Selective pharmacological targeting of a DEAD box RNA helicase. *PloS one*, *3*(2), e1583.

Lou, K., Wassarman, D. R., Yang, T., Paung, Y., Zhang, Z., O'Loughlin, T. A., ... & Shokat, K. M. (2022). IFITM proteins assist cellular uptake of diverse linked chemotypes. *Science*, *378*(6624), 1097-1104.

Lumb, J. H., Li, Q., Popov, L. M., Ding, S., Keith, M. T., Merrill, B. D., ... & Carette, J. E. (2017). DDX6 represses aberrant activation of interferon-stimulated genes. *Cell reports*, *20*(4), 819-831.

Mannocci, L., Leimbacher, M., Wichert, M., Scheuermann, J., & Neri, D. (2011). 20 years of DNA-encoded chemical libraries. *Chemical communications*, *47*(48), 12747-12753.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of applied crystallography*, *40*(4), 658-674.

Molina, D. M., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E. A., Dan, C., ... & Nordlund, P. (2013). Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science*, *341*(6141), 84-87.

Nabet, B., Roberts, J. M., Buckley, D. L., Paulk, J., Dastjerdi, S., Yang, A., ... & Scott, T. G. (2018). The dTAG system for immediate and target-specific protein degradation. *Nature*

chemical biology, *14*(5), 431.

Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature protocols*, *2*(9), 2212-2221.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., ... & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1), 70-82.

Pradere, U., Garnier-Amblard, E. C., Coats, S. J., Amblard, F., & Schinazi, R. F. (2014). Synthesis of nucleoside phosphate and phosphonate prodrugs. *Chemical reviews*, *114*(18), 9154-9218.

Presnyak, V., & Coller, J. (2013). The DHH1/RCKp54 family of helicases: an ancient family of proteins that promote translational silencing. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1829*(8), 817-823.

Presnyak, V., Alhusaini, N., Chen, Y. H., Martin, S., Morris, N., Kline, N., ... & Coller, J. (2015). Codon optimality is a major determinant of mRNA stability. *Cell*, *160*(6), 1111-1124.

Putnam, A. A., & Jankowsky, E. (2013). AMP sensing by DEAD-box RNA helicases. *Journal of molecular biology*, *425*(20), 3839-3845.

Radhakrishnan, A., Chen, Y. H., Martin, S., Alhusaini, N., Green, R., & Coller, J. (2016). The DEAD-box protein Dhh1p couples mRNA decay and translation by monitoring codon optimality. *Cell*, *167*(1), 122-132.

Russell, S. J., Gonzalez, F., Joshua-Tor, L., & Johnston, S. A. (2001). Selective chemical inactivation of AAA proteins reveals distinct functions of proteasomal ATPases. *Chemistry & Biology*, *8*(10), 941-950.

Schrodinger, LLC2015; The PyMOL Molecular Graphics System, Version 1.8.

Schütz, P., Karlberg, T., Van Den Berg, S., Collins, R., Lehtiö, L., Högbom, M., ... & Moche, M. (2010). Comparative structural analysis of human DEAD-box RNA helicases. *PloS one*, *5*(9), e12791.

Shadrick, W. R., Ndjomou, J., Kolli, R., Mukherjee, S., Hanson, A. M., & Frick, D. N. (2013). Discovering new medicines targeting helicases: challenges and recent progress. *Journal of biomolecular screening*, *18*(7), 761-781.

Singh, J., Dobrusin, E. M., Fry, D. W., Haske, T., Whitty, A., & McNamara, D. J. (1997). Structure-based design of a potent, selective, and irreversible inhibitor of the catalytic domain of the erbB receptor subfamily of protein tyrosine kinases. *Journal of medicinal chemistry*, *40*(7), 1130-1135.

Slusarczyk, M., Serpi, M., & Pertusati, F. (2018). Phosphoramidates and phosphonamidates (ProTides) with antiviral activity. *Antiviral Chemistry and Chemotherapy*, *26*, 2040206618775243

Steimer, L., & Klostermeier, D. (2012). RNA helicases in infection and disease. *RNA biology*, *9*(6), 751-771.

Wilson, D. W., Wilcox, C. A., Flynn, G. C., Chen, E., Kuang, W. J., Henzel, W. J., ... & Rothman, J. E. (1989). A fusion protein required for vesicle-mediated transport in both mammalian cells and yeast. *Nature*, *339*(6223), 355-359.

Wolfe, A. L., Singh, K., Zhong, Y., Drewe, P., Rajasekhar, V. K., Sanghvi, V. R., ... & Wendel, H. G. (2014). RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature*, *513*(7516), 65-70.

Wu, Q., Medina, S. G., Kushawah, G., DeVore, M. L., Castellano, L. A., Hand, J. M., ... & Bazzini, A. A. (2019). Translation affects mRNA stability in a codon-dependent manner in human cells. *elife*, *8*, e45396.

https://hits.wuxiapptec.com/delopen

Yuen, L. H., & Franzini, R. M. (2017). Achievements, Challenges, and Opportunities in DNA-Encoded Library Research: An Academic Point of View. *ChemBioChem*, *18*(9), 829-836.

Chapter 3

Dimerizing an RNA-Protein Molecular Glue Refines its Cell Specificity

Abstract

Small molecule therapeutics classically target enzymatic activity or protein-protein interactions. New strategies, such as targeted protein degradation, have widened how we think about using small molecules to target proteins but they still rely on the discovery of small molecule binders for proteins of interest. As an alternative, targeting protein synthesis opens a variety of new opportunities, particularly for "undruggable" proteins with no known small molecule binders. The natural product rocaglamide A (RocA) inhibits the essential cap-dependent translation initiation complex eIF4F through formation of a ternary complex with the RNA helicase eIF4A and its mRNA substrate. Formation of this complex traps eIF4A on long, highly structured 5'-untranslated regions (5'-UTRs) and inhibits protein synthesis. This inhibition is semi-selective and prevents the synthesis of a variety of oncoproteins, which tend to have more structured 5'-UTRs, without affecting the synthesis of many other proteins in the cell. However, there is a need to further improve this selectivity. Zotatifin, a synthetic derivative of Roc (Ernst et al, 2020), is currently in Phase I/II clinical trials for various KRAS- and growth receptor-driven solid tumors where toxicity issues as a result of on-target, off-tumor translation inhibition must be avoided.

We hypothesized that binding multiple copies of eIF4A with a single RocA-based molecule could improve selectivity. Based on a crystal structure of the eIF4A::RNA::RocA complex and unpublished HDX data of eIF4A multimer interactions, we have synthesized several dimers of RocA. Brief studies examining different linkers suggest that these Roc dimers (BisRoc) inhibit multiple copies of eIF4A. BisRoc also demonstrates increased potency against a subset of known Roc targets, including Myc, Cyclin D, and CDK6. Most intriguingly, while Roc and Zotatifin both are highly potent across a large panel of cell lines, BisRoc exhibits a far broader range of potencies, which suggests it could increase the therapeutic index of this inhibitor class. BisRoc appears to

have different cellular accumulation properties as well as a modified 5'-UTR sequence specificity. Characterization to identify the cellular factors that distinguish the behavior of BisRoc from that of RocA and Zotatifin revealed that factors governing entry and export have the highest impact on the potency of BisRoc and eIF4A2 levels affect the potency of the dimer more than either of the monomers. Ongoing work to characterize the 5'-UTR specificity of BisRoc, as well as to test other RocA-linked molecules, is also described here.

Introduction

Dysregulated mRNA translation is a hallmark of cancer development (Ruggero & Pandolfi, 2003; Robichaud et al, 2019; Fabbri et al, 2021). Targeting the distinct steps in mRNA translation can lead to approved therapies such as the rapamycins which block mTOR-mediated translation in cell-specific contexts (Chiu et al, 1994). A major hurdle in identifying inhibitors of mRNA translation with therapeutic potential lies in their ability to selectively disrupt important translational processes in cancer cells while sparing non-transformed cells (Merric & Hunt, 2002; Fan & Sharp, 2021). There are currently no precise inhibitors of a single cancer-associated mRNA with the exception of macromolecular approaches (e.g., siRNA, antisense oligonucleotide, and CRISPR based technologies), which have significant drug-delivery challenges (Shegokar, 2011; Taha et al, 2022).

Natural products have been a rich source of inhibitors of multiple steps of mRNA translation from initiation (rapamycin/mTOR, rocaglamide A/eIF4A, hipuristanol/eIF4A; Burgers & Fürst, 2021) and ribosomal synthesis (antibiotics which bind the bacterial ribosome; Hutchings et al, 2019). Translation initiation is highly regulated at the level of specific mRNAs through structured 5' UTR sequences (Bugaut et al, 2012; Leppek et al, 2018) and affords an opportunity

for *selective mRNA translation inhibition*. The rocaglate natural products, exemplified by rocaglamide A, possess a remarkable mechanism whereby rocaglamide A serves as a molecular glue that binds at the interface of RNA and protein, contacting a stretch of purines of the mRNA substrate as well as eIF4A residues (Iwasaki et al, 2019). This leads to trapping of eIF4A helicase on a subset of mRNA sequences with polypurine rich motifs, resulting in blockade of 43S ribosomal scanning and inhibition of the expression of numerous target proteins, including oncogenes (Iwasaki et al, 2016). Despite a crystal structure of the complex and ribosome profiling of mRNA translation effects of rocaglamide A, a complete understanding of the cellular specificity of these agents on various cancer cells is still being pursued (Chen et al, 2021).

We set out to create a modified rocaglamide A which exhibits differential selectivity for growth inhibition of specific cell types compared to the parent molecule which has high potency (<10 nM EC50) across hundreds of cancer cell lines. To achieve such differential selectivity, we turned to exploit the modulation of cytosolic accumulation of molecules with long linkers by IFITM proteins (Lou et al, 2022). We reasoned that a dimer of rocaglamide A possessing IFITM dependence might result in differential uptake, and thereby inhibition, across a large panel of cancer cell lines compared to monomeric rocaglamide A. The dimer of RocA might also provide additional selectivity for mRNAs based on the engagement of two eIF4A helicases (Schmidt et al, 2023). The dimerization of synthetic and natural products has been traditionally carried out to enhance potency through the chelate effect (Poulin-Kerstein & Dervan, 2003; Cowell & Lee, 2016; Maniaci et al, 2017; Guan et al, 2022). In this case we hoped to "tune" the selectivity through cell access and differential eIF4A dependence of mRNA translation.

Results

Design and synthesis of RocA dimers

We used the crystal structure of eIF4A bound to RocA and a poly-AG RNA as a starting point for designing RocA dimers (Iwasaki et al, 2019). Two symmetry mates in the crystal structure showed the proteins oriented such that the RocA molecules faced one another (**Figure 3.1 A**). This orientation aligns with unpublished hydrogen-deuterium exchange mass spectrometry (HDX-MS) data that suggests eIF4A1 oligomers form contacts at several residues along that face. Based off this data, we modeled that two RocA molecules could be linked together across that space to form a dimer (BisRoc) that bound two copies of eIF4A simultaneously (**Figure 3.1 B**). Looking more closely at the orientation of RocA in its binding pocket, we identified a solvent-exposed amide (**Figure 3.1 C**) as a location that has previously been derivatized (Chen et al, 2021) and ideal for linker attachment. We designed a series of three linker lengths: PEG-11, PEG-4, and PEG-2 (**Figure 3.1 D**). Based off the approximate 14 Å distance between two RocA molecules in the crystal structure, the PEG-11 and PEG-4 linkers should both span the distance to interact with two eIF4A proteins, while the PEG-2 linker was expected to be prohibitively short and was designed as a negative control. All RocA dimers were synthesized as previously described through a one pot amide coupling of the respective PEG linker diamine and rocagloic acid (Lou et al, 2022; **Figure 3.2 A**).

All three BisRoc molecules were tested in a cell viability assay against RocA and retained similar cell activity to RocA in K562 cells, an immortalized chronic myeloid leukemia cell line (**Figure 3.1 E**). Western blotting for previously reported RocA targets showed that the BisRoc molecules retained similar specificity to RocA and BisRoc-1 had increased potency against a subset of targets (**Figure 3.1 F** and **Figure 3.2 B**). BisRoc potency generally showed an inverse

correlation with linker length, so we chose to focus primarily on BisRoc-1 (PEG-11) for further characterization from this point forward.

Cancer cell line panel screen

Although RocA and BisRoc-1 performed similarly in preliminary experiments in K562 cells, we were curious how potency would vary across cell lines and what other cellular factors might play a role (Rees et al, 2016). In a five-day cell viability assay across more than 300 cancer cell lines (Garnett et al, 2012), RocA had a single log-fold range in potency with EC_{50} calculations in the single to double digit nanomolar range in almost all cell lines, which agrees with previously reported studies (Iwasaki et al, 2016); Chu et al, 2020. This high potency across a broad range of cancer cell lines makes sense with the mechanism of action of RocA, which targets the synthesis of many key growth and signaling proteins, as opposed to inhibitors that target a specific driver oncogene or mutation. Intriguingly, BisRoc-1 had a four log-fold spread in potency, where some cell lines responded similarly to both molecules, but a subset of lines were far more resistant to BisRoc-1 (**Figure 3.3 A**). This data suggested that interactions with cellular level factors, rather than affinity for individual mRNAs, accounts for the main difference between RocA and BisRoc-1 potency. (**Figure 3.3 B-C** and **Figure 3.4 B-C**).

To search for effects of individual gene expression on differences between RocA and BisRoc-1 potency, we looked for correlations between mRNA-Seq data and molecule potency across cell lines (**Figure 3.3 D-E** and **Figure 3.4 D**) (DepMap). Individual mutations and tumor type showed no correlation but such analyses were also limited by the size of the dataset. The strongest genetic correlation by far with BisRoc-1 potency was ABCC1, which encodes multidrug resistance-associated protein 1 (MRP1). Increased ABCC1 levels correlated with decreased BisRoc-1 potency, likely through export of the linked molecule, but had no significant correlation

with RocA potency. Interferon induced transmembrane protein 1 (IFITM1) had a strong positive correlation with BisRoc-1 potency, which aligned with our previous findings that IFITM proteins (IFITM 1-3) appear to play a role in the cellular import of large, linked molecules (Lou et al, 2022). Although some strong genetic correlations like these were able to be pulled out of the screen, the amount of noise in the screen from comparing levels of almost 20,000 protein-coding genes across 298 cell lines meant that many signals were too low to identify. Additionally, genes with low expression variability across the cell panel could not be identified as potential interactors.

CRISPRi screen

To better understand which individual genes most contributed to the differences in potency between RocA and BisRoc-1, we turned to a genome-wide CRISPRi screen, which offers the advantage of being able to knock down individual genes in a single cell line, ensuring that each individual knockdown should be against the same genetic background. K562 CRISPRi cells were treated with DMSO, RocA, or BisRoc-1 over the course of a 11-day experiment at concentrations chosen to induce a selective pressure (**Figure 3.5 A**). All molecules were maintained over the course of treatment, except for BisRoc-1, which was removed on day 7 to ensure enough cells remained at the end of the experiment for downstream analysis.

Volcano plots of the growth phenotype for all the genes tested under RocA and BisRoc-1 treatment reflect some of the top hits from the cell panel screen and indicate genes that did not previously show up (**Figure 3.5 B-C**). Knockdown of several genes which we previously identified to play roles in linked molecule uptake and export induced sensitivity to BisRoc-1 but not RocA. ABCB1, which encodes the transmembrane transporter P-glycoprotein (P-gp), is another member of the ABC transporter family that pumps out many xenobiotic compounds. Similarly to ABCC1, this transporter was one of the strongest hits in the screen. We expect that different related

transporters showed up in the two different screens because ABCC1 has very low expression in K562 cells and may not be picked up in a knockdown, while ABCB1 has similar expression across the cell panel and would not show up in an expression correlation. LDLR and OSBP were two other sensitizing hits that we have seen previously in screens of unrelated linked molecules (Lou et al, 2022). For resistance hits, IFITM1 had the greatest effect on growth phenotype and IFITM2 showed up as well. IFITM3 was not a strong hit, which was not surprising because IFITM3 has the lowest expression of these 3 proteins in K562 cells. Our previous work also used the BisRoc series to show that longer linker length increased the dependence of these molecules on IFITM proteins for cellular entry (Lou et al, 2022). The dependence on IFITM levels was tested here with CRISPRa K562 cell lines overexpressing each of the IFITM proteins (**Figure 3.6 A**).

While EIF4A1 was not a significant hit for either RocA or BisRoc-1, several other initiation factors in the eIF4F complex appeared as hits for BisRoc-1 only, including EIF4B knockdown as one of the strongest sensitizing hits and EIF4G1 and EIF4E as resistance hits. Another of the strongest sensitizing hits was knockdown of the gene EIF4A2, the paralog of EIF4A1. The exact role of the eIF4A2 protein in translation initiation is still being determined but it has been demonstrated to act similarly to eIF4A1 when overexpressed and to play a role in translational repression (Meijer et al, 2019; Wilczynska et al, 2019). To confirm this hit, we selected cell lines expressing EIF4A2 knockdowns (**Table 3.1**). We were unable to select for a full knockdown but achieved partial knockdowns confimed by Western blotting and which decreased over time (**Figure 3.5 B**). EIF4A2 partial knockdown induced slight sensitivity to RocA, no sensitivity to Zotatifin, and a stronger sensitivity to all three BisRoc molecules (**Figure 3.6 C**).

BisRoc methylation

To investigate the mechanism of BisRoc-1 dependency on eIF4A2 levels, we examined the structural differences between it and RocA. In addition to the PEG-11 linker, the other key difference is that solvent-exposed amide on RocA is demethylated, while the BisRoc-1 solvent-exposed amide is linked to the PEG linker and unmethylated. Examining the structures of eIF4A1 and eIF4A2 showed that D198, a residue near that amide in eIF4A1, is replaced by E198 in eIF4A2 (**Figure 3.7 A**). We hypothesized that the extra methylene in the E198 side chain might allow it to reach far enough to form a hydrogen bond with the unmethylated nitrogen of the amide on BisRoc-1. To test this hypothesis, we synthesized mono- and un-methylated versions of RocA (**Figure 3.7 B**), as well as a methylated version of BisRoc-1 (**Figure 3.7 C**) and tested them against the EIF4A2 knockdown cell lines. Zotatifin and the mono- and un-methylated versions of RocA showed no sensitivity to EIF4A2 knockdown, while the methylated BisRoc-1 maintained its sensitivity (**Figure 3.7 D**). This data suggests to us that the methylation status of the solvent-exposed amide does not play much of a role in the compound's sensitivity to eIF4A2 and that the presence of the PEG linker or dimerization has a much larger effect.

Roc-RIBOTAC

We also wanted to test the effects of appending another RBP-binding small molecule to RocA and chose a previously reported RNase L ligand (Costales et al, 2020) to see if we could induce targeted-RNA degradation of Roc-bound RNAs. We synthesized our "Roc-RIBOTAC" compound (Tong et al, 2023) with a PEG-2 linker based on previously reported results with the RNase L ligand (Costales et al, 2020; Haniff et al, 2020; **Figure 3.8 A**) and confirmed that it still was functional, although less potent than RocA, in a cell growth assay (**Figure 3.8 B**). The Roc-

RIBOTAC also lowered protein levels of known RocA targets (**Figure 3.8 C**) but did not cause a decrease in their RNA levels (**Figure 3.8 D**), which suggested to us that it was not inducing degradation of bound mRNAs.

Discussion

We designed and synthesized series of RocA dimers with the hypothesis that we could modulate the activity of RocA by targeting multiple copies of eIF4A on mRNA 5'-UTRs. Testing our lead molecule, BisRoc-1, against a large cancer cell line panel revealed a surprising pattern: while RocA was already known to be highly potent against many cancer lines, BisRoc-1 retained potency against some but was several orders of magnitude less potent against others. This effect is unique from what is typically demonstrated by other linked molecules; rather than improving an already potent molecule, dimerization greatly increased its cell line specificity.

Gene expression correlation data from tested cell lines and a genome-wide CRISPRi screen revealed that this change could not be attributed to a single gene but was due to several key chemical genetic interactions. We hypothesize that each of the genes followed up on contributed to differences in potency in part by tuning the intracellular levels of BisRoc-1, whether through its influx, efflux, or cellular retention. CaCo-2 assays show that BisRoc-1 is essentially impermeable to the cellular membrane. We confirmed our previous work, showing that BisRoc molecules depend on IFITM proteins for cellular entry and this dependence correlates strongly with linker length. On the opposite end of the spectrum, the multi-drug efflux pump ABC transporters were the top resistance hits for BisRoc-1. Comparing their effect on the RocA monomer to its dimer is the first direct confirmation that large, linked molecules are better ABC transporter substrates than the unlinked monomers. This effect can be found in other screens against linked molecules like

PROTACs but has not been directly investigated before. Several translation initiation factors were hits in the CRISPRi screen, the strongest of which was eIF4A2.

We hypothesize that BisRoc-1 could have increased cellular retention in cells with higher eIF4A1 and eIF4A2 protein levels because the molecule's two eIF4A binding sites help maintain it in proximity to eIF4A multimers and binding to eIF4A2 as well could increase the protein-bound fraction of molecule (similarly to the mechanism of rapamycin binding to FKBP12). More work will have to be done to confirm if this is indeed the mechanism by which eIF4A2 levels and BisRoc-1 potency is correlated. Testing a methylated BisRoc molecule and unmethylated RocA molecules suggest that methylation of this molecule is not the reason for eIF4A2 dependency and dimerization could be a factor.

Finally, the BisRoc series acts a proof of concept that RocA activity can be altered by linking it to other small molecules. Here we show that simply dimerizing the molecule alters key chemical genetic interactions, but we expect that this series of linked molecules could be greatly expanded to explore other possibilities. Small molecule targeting of RNA has long proved a challenge, but RocA could be used as a starting scaffold to link to small molecules that have been reported to bind other sequences or secondary structure in RNA to alter the sequence specificity of RocA. Additionally, linking RocA to other molecules targeting RBPs could introduce new protein-protein or protein-RNA interactions to apply concepts that have become mainstream in the protein world, such as targeted degradation or introduction of PTMs, to RNAs bound by RocA. We tested a Roc-RIBOTAC molecule as a proof of concept and, although it did not show signs of RNA degradation, we believe that it does not rule out the possibility of this type of molecule working and much more work could be done testing other linkers and RNase ligands.

Overall, this work furthers our knowledge of mechanisms governing cellular activity generally for larger linked small molecules and specifically for RocA derivatives.

Materials & Methods

Cell culture and reagents

K562 CRISPRi and CRISPRa cell lines were a gift from Luke Gilbert at UCSF and generated as previously described (#). OE-19 cell lines were a gift from John Gordan at UCSF. HEK293T (CRL-3216), A673 (CRL-1598), SW403 (CCL-230), HCC1187 (CRL-2322), and NCI-H1155 (CRL-5818), NCI-H196 (CRL-5823) cell lines were purchased from ATCC. HEL, HEL 92.1.7, 769-P, G-401, and KELLY cell lines were purchased from the UCSF Cell and Genome Engineering Core. All cell lines were grown at 37°C with 5% CO_2 and in media and supplements as specified by ATCC. Cells were tested for mycoplasma with the MycoAlert PLUS Mycoplasma Detection Kit (Lonza LT07-701). Cells were counted by a Countess II FL Automated Cell Counter (Thermo Fisher), Attune II flow cytometer, and Attune Nxt flow cytometer (Thermo Fisher). Rocaglamide A was purchased from MedChemExpress (HY-19356) and BisRoc-1, BisRoc-2, and BisRoc-3 were synthesized as described previously. All compounds were stored as 10 μM-10 mM solutions in dimethyl sulfoxide (DMSO) at -20°C.

Cell panel viability screen

High-throughput drug screening was performed as previously described (Garnett et al, 2012; Lou et al, 2022). 300 cancer cell lines were grown in medium supplemented with penicillin/streptomycin and 5% FBS and grown in a humidified atmosphere at 37°C with 5% CO_2. All cell lines were grown in RPMI or DMEM/F12 medium to facilitate high-throughput screening

and to minimize potential effects of different media on drug sensitivity. Cells were seeded in 384 well plates at varying density to ensure optimal growth of each cell line over the course of the experiment. Tumor subtypes with suspension cells were treated with drugs on the same day as seeding and tumor subtypes with adherent cell lines were treated the following day. Tumor subtypes with a mixture of suspension and adherent cells were treated the same day (eg: small cell lung cancer cell lines were all treated on the day after seeding). Compounds were given in duplicate at 9 doses with a 3-fold dilution factor. The highest dose of RocA was 1 μM and the highest dose of BisRoc-1 was 3 μM. A Day 0 readout was done for growth-rate normalization. Cell viability was determined by CellTiter-Glo after five days of compound treatment and compound-treated wells were normalized to DMSO-treated wells.

Correlation analysis between cell viability and gene expression

Dose-dependent inhibition of 298 cancer cells line by RocA and BisRoc-1 was determined as described above. Correlations of drug sensitivity were determined using the cellpanelr web-app and R package (biorxiv citation) with gene expression and proteomics data obtained from the DepMap (22Q1 release) and plotted in Prism 9 (GraphPad Software).

Genome-wide CRISPRi screen

K562 CRISPRi cells were grown at 37°C and 5% CO_2 in 1 L Nalgene Disposable Erlenmeyer flasks with Vented Closure (Thermo Scientific 4112-1000) in a shaking culture (1300 rpm) in a Multitron Incubator (Infors HT). Cells were cultured in RPMI (Gibco) supplemented with 10% fetal bovine serum (FBS) (Avantor Seradigm), penicillin (100 U/mL, Gibco), streptomycin (100 μg/mL, Gibco), glutamine (29.2 μg/mL, Gibco 10378016), and 0.1% Pluronic

F-68 (Gibco, Thermo 24040032). Cells were transduced with the 5-sgRNA/gene human CRISPRi v2 library (hCRISPRi-v2) with 8 µg/mL polybrene. Library virus was tittered to maximally transduce cells while maintaining a multiplicity of infection (MOI) less than 1. Transduction efficiency and cell growth and viability was monitored daily throughout the experiment by flow cytometry (Attune II, Thermo Fisher). sgRNA+ cells were selected with doses of 1 µg/mL puromycin at 48 and 72 hrs post-transduction until the population was 80-90% sgRNA positive, then cells were allowed to recover in fresh media for 24 hrs before compound treatment. T0 samples were collected before treatment at approximately 750-fold coverage of the library (750 million cells) and the rest of the cells were split into 4 treatment arms with 2 biological replicates per arm: DMSO, 10 nM RocA, 5 nM BisRoc-1. Cells were cultured at a minimum of 750- to 1000-fold library coverage and continued to be monitored daily by flow cytometry. Dilutions were made daily as necessary with media supplemented with treatments at the indicated concentration to maintain selective pressure until the end of the experiment on Day 11 of compound treatment. The only exception was the BisRoc-1 treatment arms, in which compound was removed after Day 5 of treatment in order to decrease selective pressure to maintain adequate library coverage. Samples were harvested on Day 11 at a minimum of 1000-fold coverage (or 750-fold coverage for BisRoc-1 arms). NucleoSpin Blood XL kits (Macherey-Nagel, Takara Bio 740950.10) were used to extract genomic DNA (gDNA) from all samples and sgRNA protospacers were amplified directly from this with Ti Taq polymerase (Takara Bio 639209), analyzed on an Agilent Tapestation on an High Sensitivity DNA D1000 ScreenTape (Agilent 5067-5584 and 5067-5585), and then sequenced on an Illumina HiSeq 4000 through the UCSF Center for Advanced Technology (UCSF CAT, supported by UCSF PBBR, RRP IMIA, and NIH 1S10OD028511-01 grants).

Individual sgRNA cloning

sgRNA protospacers targeting individual genes or a negative control were chosen from the library (Table 3.1) and cloned into the vector pCRISPRia-v2 (Addgene 84832) as previously described. Complementary oligonucleotides (Integrated DNA Technologies) were annealed by mixing in Nuclease-Free Duplex Buffer (Integrated DNA Technologies) and heated at 95°C for 5 min, then cooled at 22°C for 1 hr. The vector pCRISPRia-v2 was digested by BstXI and BlpI (New England Biolabs) and gel purified, then ligated to the annealed duplex with fresh T4 DNA Ligase (New England Biolabs M0202S). Ligated plasmids were transformed into Stellar Competent Cells (Takara Bio 636763) and grown overnight. Individual colonies were grown overnight in 3 mL liquid cultures, plasmids were purified with the Qiagen Spin Miniprep Kit, and sequenced by standard Sanger sequencing (Elim Bio).

Lentivirus production

HEK293T cells were plated (5e5 cells) in 3 mL media in a 6 well plate and grown overnight. Plasmids were transfected with sgRNA vectors and transfection reagents following the TransIT-LT1 transfection protocol (Mirus MIR2300). Additional media (1.5 mL) was added 48 hrs post-transfection and virus was harvested, aliquoted, and immediately used or frozen at -80°C.

Stable cell line generation

K562 CRISPRi cells (2e5) were seeded in 1 mL media in 24-well plates, grown overnight, and treated with lentivirus containing individual sgRNA vectors (this vector also contained blue fluorescent protein (BFP) and puromycin resistance for cell selection) and 8 μg/mL polybrene (Fisher Scientific TR-1003-G). Cells were treated with 2 μg/mL puromycin (Invivogen ANT-PR-

1) beginning 48 hours post-transduction and puromycin was maintained for 10 days until populations were pure by flow cytometry. Knockdowns were checked by immunoblotting and cell line stability was routinely checked by flow cytometry and immunoblotting over the course of subsequent experiments.

Cell viability assays

Cells were seeded (1e4 K562 or 2-5e4 of slow growing cell lines) in 90 µL media in 96 well plates and grown overnight. All wells along the edges of the plate were instead filled with 200 µL PBS to help prevent evaporation over the course of the assay. Cells were treated in triplicate with compounds in 10 µL media with 0.1% DMSO. Compounds were given in 2-fold or 3-fold dilution with a highest final concentration of 1 µM or 500 nM. Plates were incubated at 37°C with 5% CO_2 for 72 hrs following treatment, then cooled at room temperature for 20 min. 100 µL of a solution of 1:5 CellTiter-Glo (Promega G7572) to PBS was added per well and plates were shaken for 20 min to mix before luminescence was read out on a Tecan Spark plate reader.

Immunoblotting

Cells were seeded (0.7e6 HEK 293T or 0.5e6 K562) in 3 mL media in a 6-well plate and grown overnight, then treated with compounds in 0.1% DMSO. HEK293T cells were harvested by scraping after incubating 5 min in ice-cold phosphate-buffered saline (PBS, Gibco 100110049). Cells were pelleted by spinning at 500g for 5 min, then washed twice with ice-cold PBS. Cell pellets were snap frozen and stored at -80°C. Cells were lysed by incubating 10 min on ice in lysis buffer [100 mM Hepes (pH 7.5), 150 mM NaCl, and 0.1% NP-40] with mini-cOmplete Protease Inhibitor Cocktail Tablets (Sigma 11836170001) and PhosSTOP (Roche 4906837001), then

centrifuged at 18,000g for 20 min. Protein concentrations of the clarified lysates were determined by Bradford (Biorad 5000006) on a Nanodrop and lysates were normalized with lysis buffer. Proteins were separated on 4-12% Bis-Tris gels (Thermo Fisher) by polyacrylamide gel electrophoresis (PAGE) and transferred to 0.2 um nitrocellulose membranes (Biorad) in ice-cold Towbin buffer. Membranes were blocked in 5% bovine serum albumin (BSA, Millipore 12659) in Tris-buffered saline, 0.1% Tween 20 (TBS-T) with 0.02% NaN_3. Primary antibodies against the following targets were used to probe the membranes: Cyclin D3 (DCS22, 2936), Cyclin D1 (92G2, 2978), c-MYC (D84C12, 5605), B-actin (8H10D10, 3700), GAPDH (14C10, 2118), TUB (DM1A, 3873), CDK 4 (D9G3E, 12790), CDK 6 (D4S8S, 1331 and DCS83, 3136), IFITM1 (13126), IFITM2 (13530), IFITM3 (59212) from Cell Signaling Technology, pan-Ras (ab108602) and eIF4AI (ab31217) from AbCam, PTGES3 (MAB-100391-100) from R&D Systems, eIF4AII (H-5, SC-137148) from Santa Cruz Biotechnologies. Primary antibodies were diluted 1:1000 in blocking buffer (except for PTGES3, which was diluted 1:250). Membranes were probed overnight with IRDye secondary antibodies (Li-COR) diluted 1:10,000 in blocking buffer, read out on an Odyssey imager (Li-COR), and analyzed on Image Studio Lite 5.2.5 software (Li-COR).

Figures

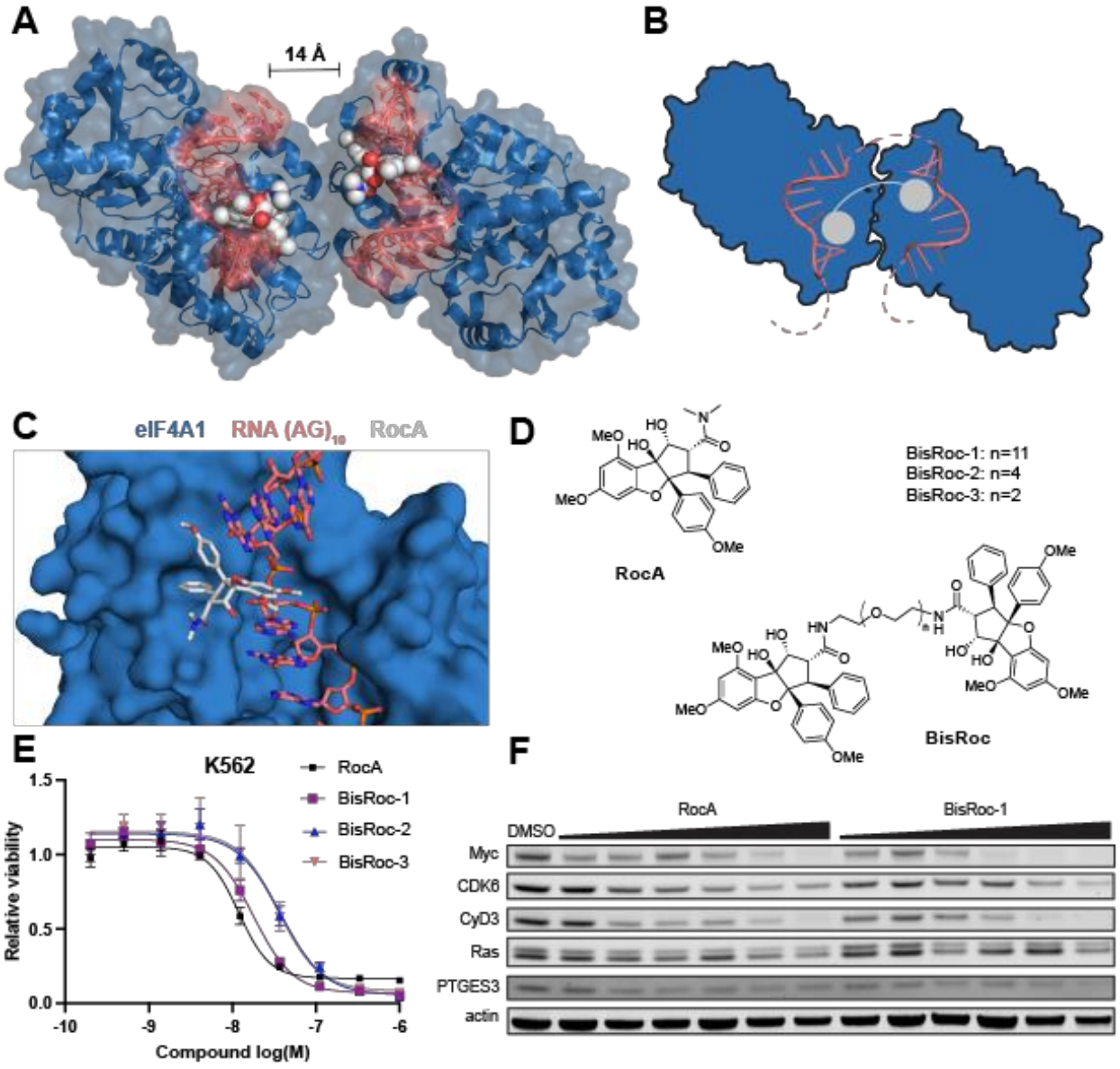


Figure 3.1 Design and testing of RocA dimers (BisRoc). A) Crystal structure of eIF4A1 (blue), RocA (grey), and poly-AG RNA (pink) in a ternary complex (PDB: 5ZC9) reveals 14 Å distance between RocA molecules bound to eIF4A1 in adjacent packing units. B) Proposed model of a BisRoc molecule (gray) bound to two interacting copies of eIF4A. C) Close up of RocA-binding site shows solvent-exposed amide that can be used for RocA derivatization. D) Structures of RocA and BisRoc molecules. E) BisRoc molecules were tested in a 3-day cell growth assay in K562 cells. F) Western blot of established RocA targets shows lowered protein levels when treated with RocA or BisRoc-1. K562 cells were treated 24 hrs with 0.1% DMSO or 1, 3, 10, 30, 100, or 300 nM compound.

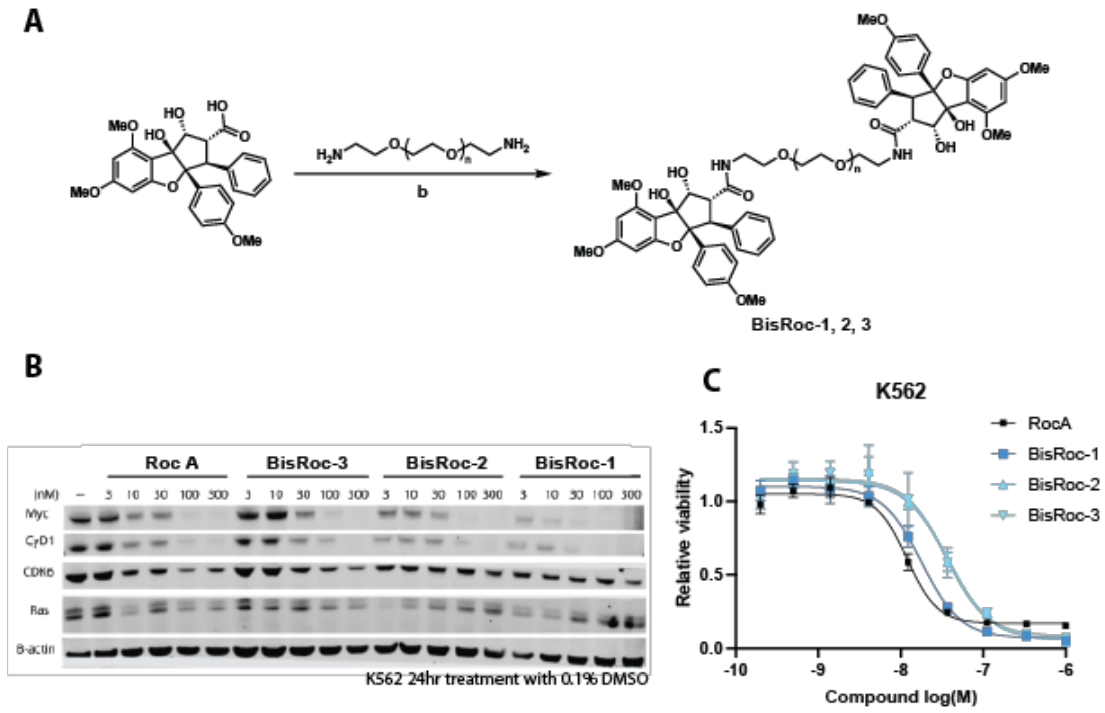
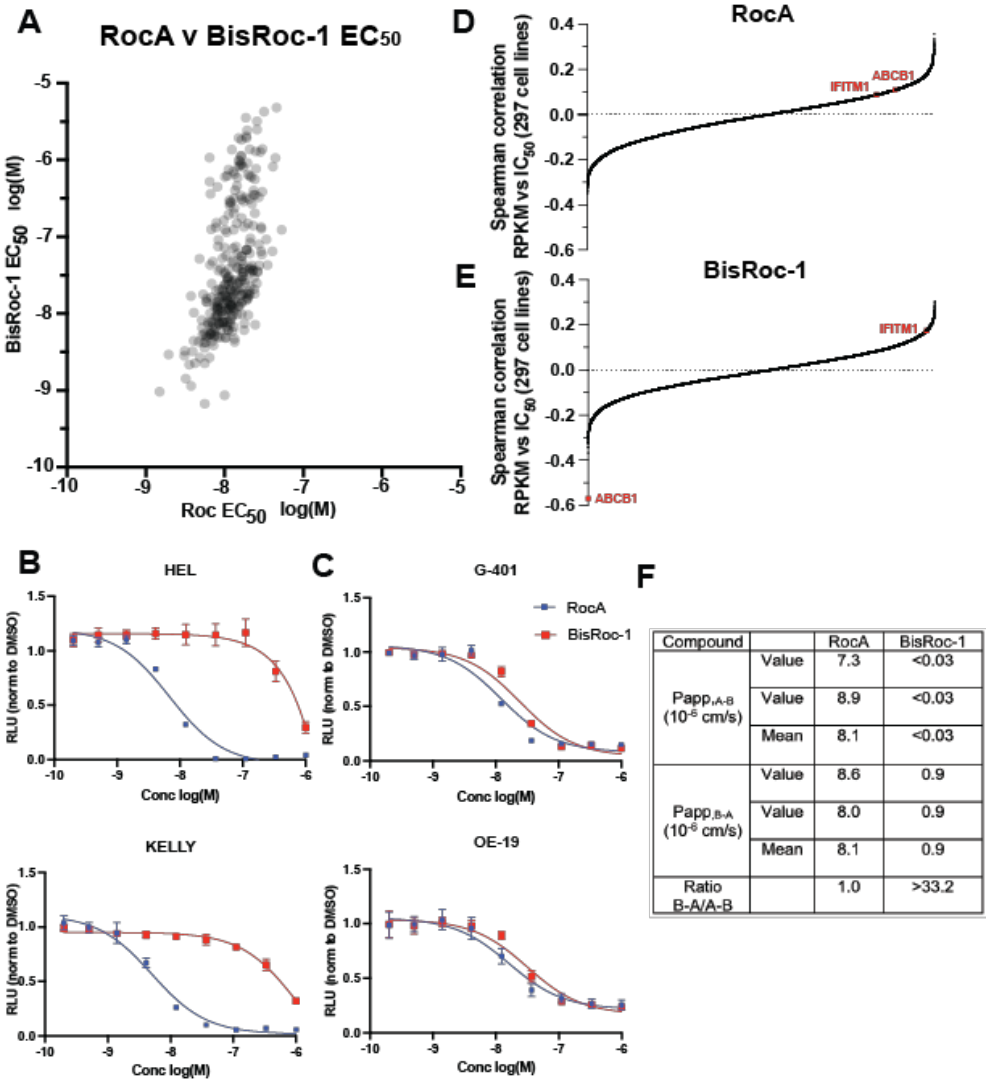


Figure 3.2 Synthesis and testing of RocA dimers (BisRoc). A) General one-step synthetic scheme to synthesize BisRoc molecules from rocagloic acid. B) Western blot of established RocA targets shows lowered protein levels when treated with RocA or BisRoc molecules. K562 cells were treated 24 hrs with 0.1% DMSO or compound. C) BisRoc molecules were tested in a 72 hrs cell growth assay in K562 cells. BisRoc-1 was chosen for further testing based on the results of B and C.



F

Compound		RocA	BisRoc-1
$Papp_{A-B}$ (10^{-6} cm/s)	Value	7.3	<0.03
	Value	8.9	<0.03
	Mean	8.1	<0.03
$Papp_{B-A}$ (10^{-6} cm/s)	Value	8.6	0.9
	Value	8.0	0.9
	Mean	8.1	0.9
Ratio B-A/A-B		1.0	>33.2

Figure 3.3 RocA and BisRoc-1 show different cell specificities in a cancer cell line panel. A) RocA and BisRoc-1 were tested in a 5-day cell growth screen against 300 cancer cell lines to determine their potencies. B) A subset of cell lines was sensitive to RocA but several orders of magnitude less sensitive to BisRoc-1, which recapitulated in an independent 3-day cell growth assay. C) A subset of cell lines was sensitive to RocA and similarly sensitive to BisRoc-1, which recapitulated in an independent 3-day cell growth assay. D) Gene expression data (DepMap) for 297 cell lines was correlated with RocA potency in the screen. E) Gene expression data (DepMap) for 297 cell lines was correlated with BisRoc-1 potency in the screen. Genes of interest that significantly correlated are marked in red. F) RocA and BisRoc-1 were tested in a CaCo-2 assay where RocA was able to pass through the cell membrane but BisRoc-1 was not.

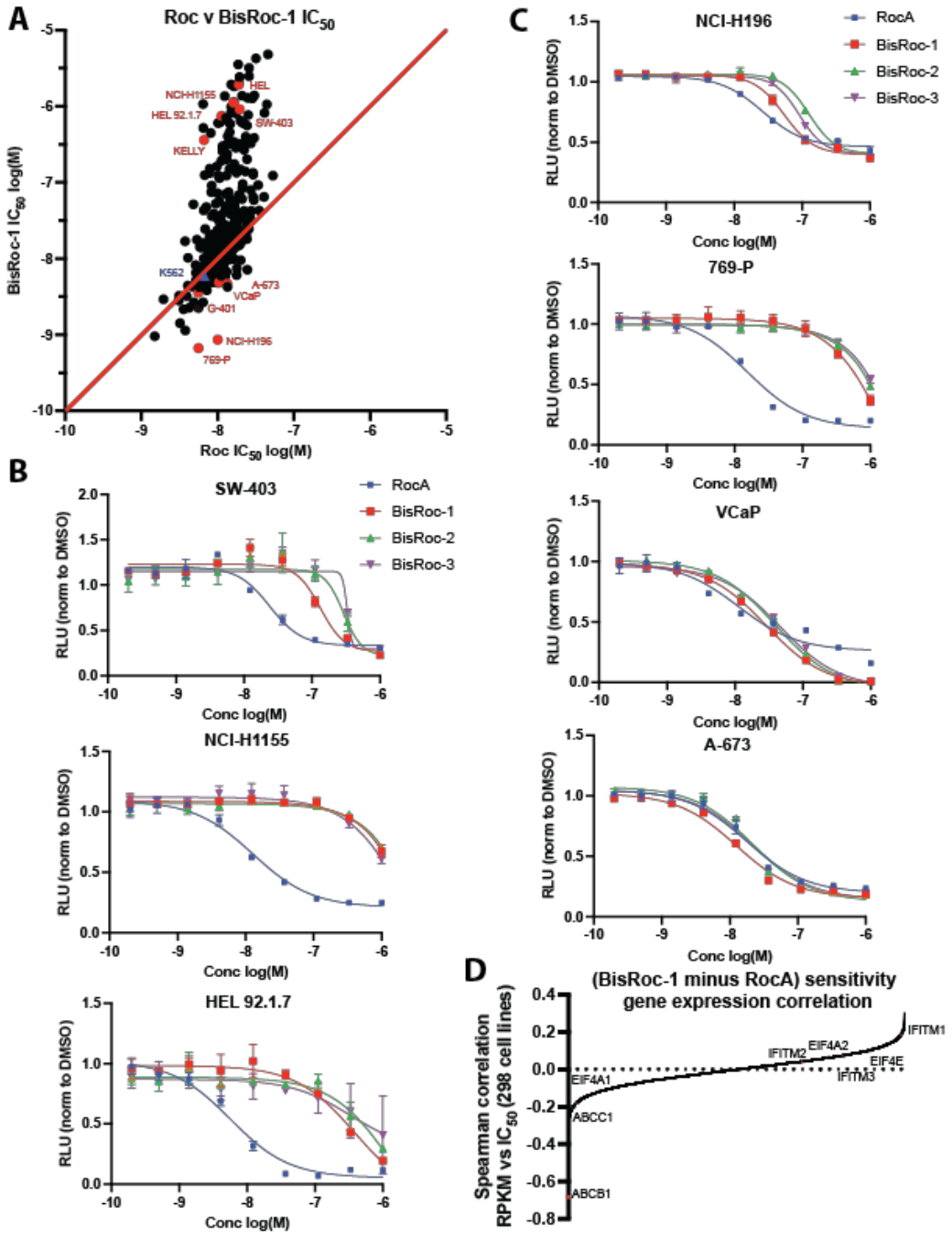


Figure 3.4 RocA and BisRoc-1 show different specificities in a cancer cell line panel. A) RocA potency values in a cancer cell line panel are plotted against BisRoc-1 potencies. K562 is marked in blue. Labeled cell lines marked in red were chosen for independent validation. B) A subset of cell lines was sensitive to RocA but several orders of magnitude less sensitive to BisRoc-1. C) A subset of cell lines was sensitive to RocA and similarly sensitive to BisRoc-1. D) Gene expression data for 298 cell lines (DepMap) was correlated with the difference between BisRoc-1 and RocA sensitivity.

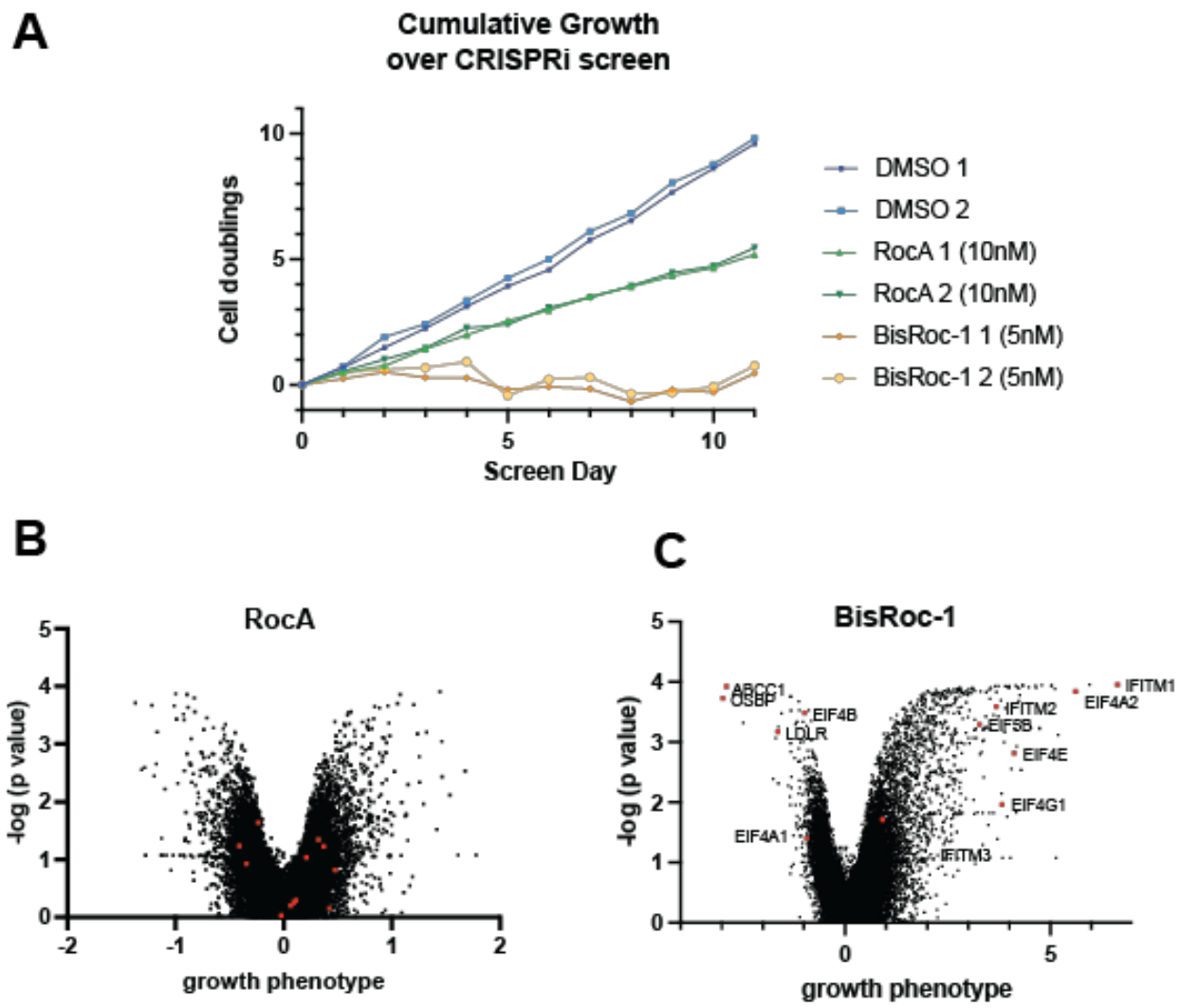


Figure 3.5 CRISPRi screen reveals chemical genetic interactors affecting BisRoc-1 potency. A) Cumulative growth curves for each treatment arm over the course of the 11-day genome-wide CRISPRi screen in K562 cells. BisRoc-1 was removed from both replicates on Day 7, while RocA treatment remained constant throughout. B) Volcano plot of growth phenotype scores for individual genes for the RocA-treated population. Genes marked in red did not have a significant growth phenotype for RocA but did for BisRoc-1 (except for EIF4A1 and IFITM3, which were nonsignificant in both). C) Volcano plot of growth phenotype scores for individual genes for the BisRoc-1 treated population.

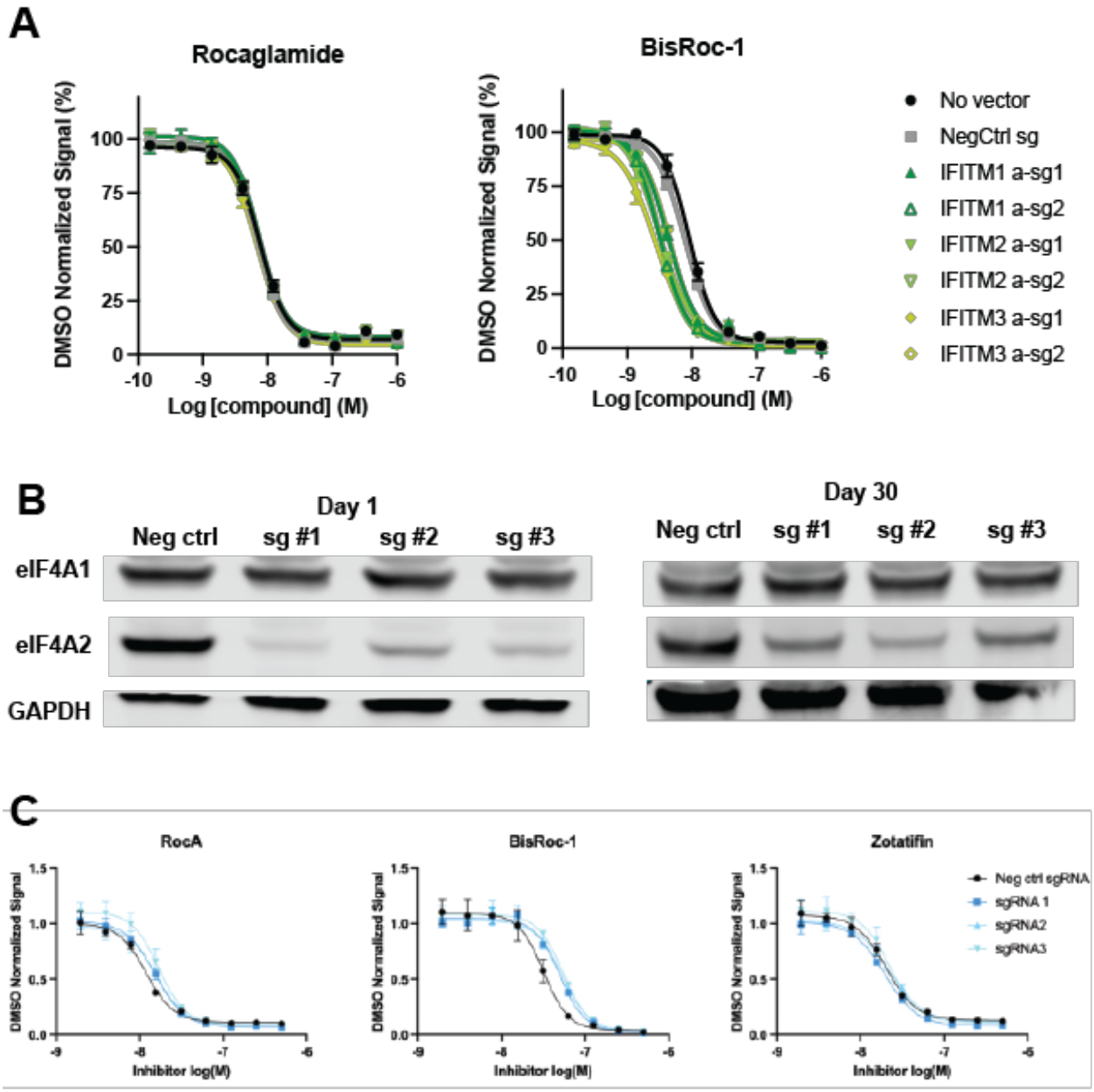


Figure 3.6 CRISPRi screen validation. A) RocA and BisRoc-1 were tested in a 3-day cell growth assay using CRISPRa to overexpress IFITM1-3. B) CRISPRi was used to make cell lines with partial EIF4A2 knockdowns. Knockdowns were validated by Western blot and re-validation over time showed that protein expression levels partly recovered over a 30 day period. C) RocA, BisRoc-1, and Zotatifin were tested in a 3-day cell growth assay in EIF4A2 knockdown cells from Panel B.

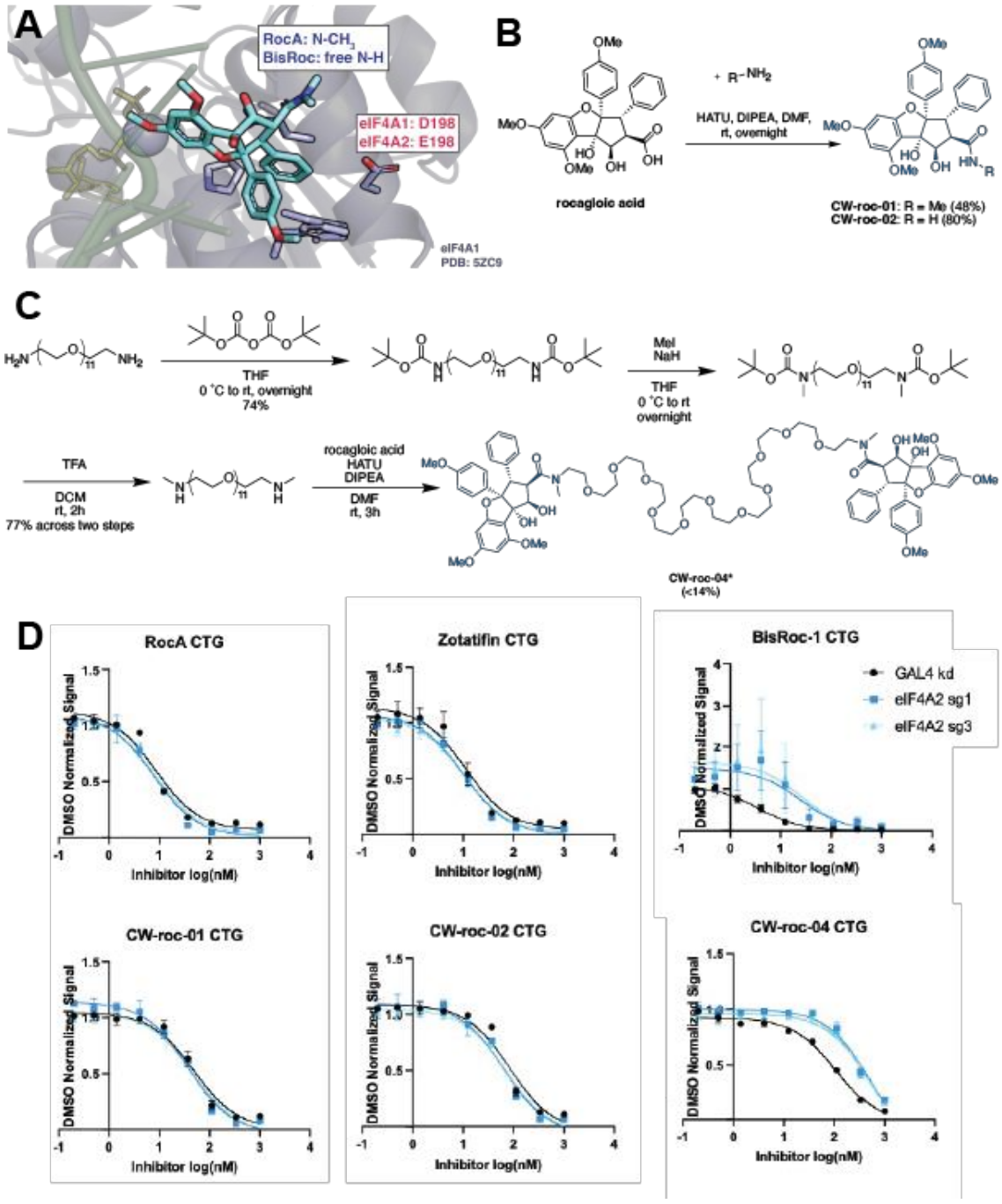


Figure 3.7 Methylation of the RocA or BisRoc-1 amide does not affect dependence of compound sensitivity on eIF4A2 levels. A) Crystal structure of RocA bound to eIF4A1 with methylated amide and D198 shown (PDB: 5ZC9). B) Synthetic route for mono- and unmethylated RocA (CW-Roc-01 and CW-Roc-02). C) Synthetic route for methylated BisRoc-1 (CW-Roc-04). D) The new series of compounds was tested in a 3-day cell growth assay in K562 cells with a negative control guide (GAL4) or EIF4A2 knockdowns. Potency of RocA, Zotatifin, CW-Roc-01, and CW-Roc-02 did not change with EIF4A2 knockdowns. BisRoc-1 and CW-Roc-04 were less potent with EIF4A2 knockdowns.

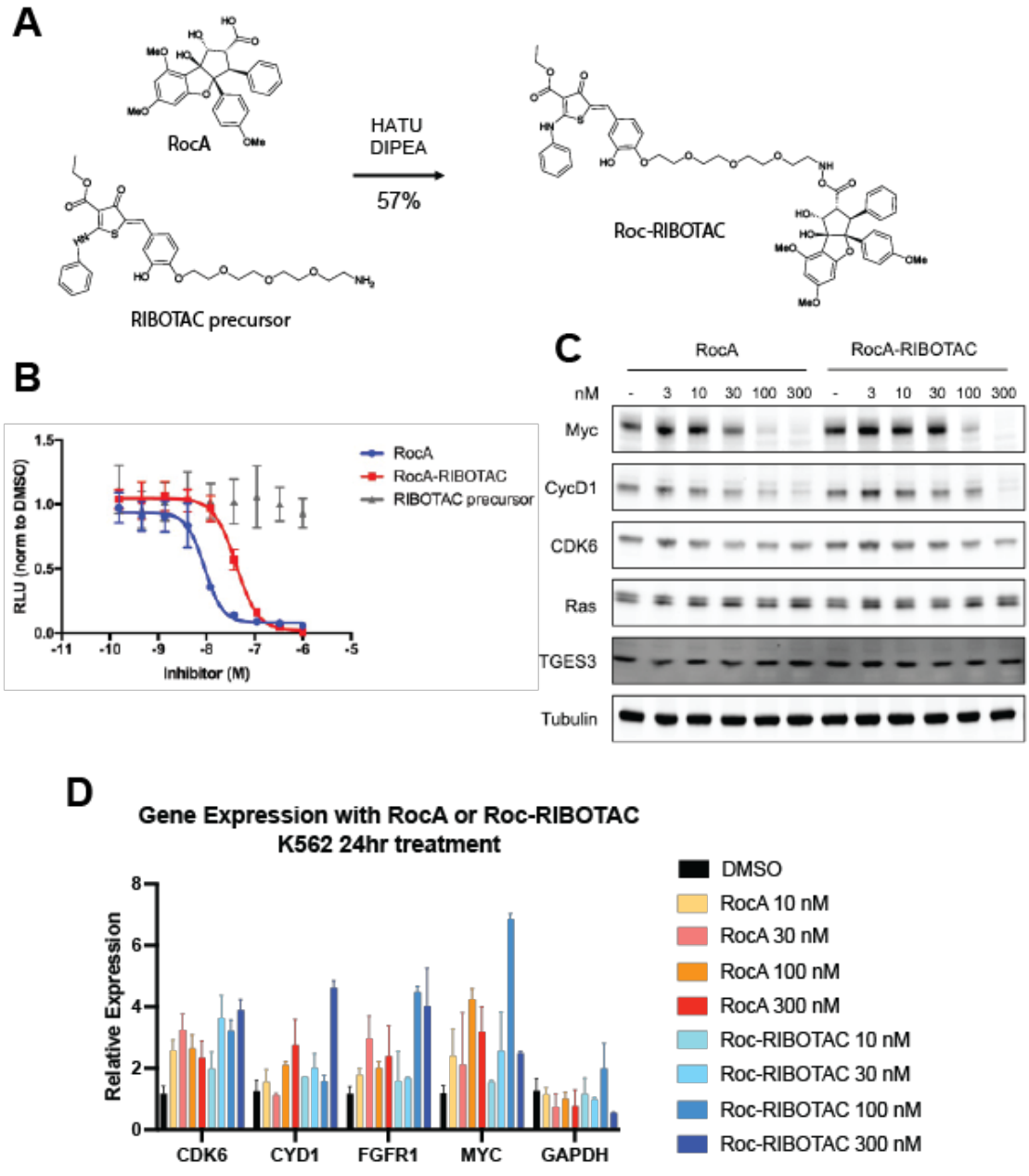


Figure 3.8 Roc-RIBOTAC synthesis and testing. A) Synthetic scheme of one-step reaction of rocagloic acid and RIBOTAC precursor to make Roc-RIBOTAC. B) Roc-RIBOTAC was tested in a 3-day cell growth assay in K562 cells. C) Roc-RIBOTAC was tested against established RocA targets and protein levels were determined by Western blotting. K562 cells with treated 24 hrs with 0.1% DMSO or compound. D) In the same experiment as C, gene expression levels for the tested RocA targets were determined by qPCR, as well as expression levels for the house-keeping protein GAPDH.

Tables

Table 3.1 CRISPRi sgRNA protospacer sequences. Protospacer sequences were chosen from the hCRISPRi v2 Top 5 Library (Horlbeck et al, 2016) and oligos ordered and cloned into expression vector as previously described (Horlbeck et al, 2016; Lou et al, 2022; Chapter 3).

sgRNA name	Sequence
eIF4A2 sgRNA1	GCGCCCGACTGAAAAGACAG
eIF4A2 sgRNA2	GGCATACCTGTTATAATCCG
eIF4A2 sgRNA3	GCGCCGCTGTCTTTTCAGTC

References

Bugaut, A., & Balasubramanian, S. (2012). 5′-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic acids research*, *40*(11), 4727-4741.

Burgers, L. D., & Fürst, R. (2021). Natural products as drugs and tools for influencing core processes of eukaryotic mRNA translation. *Pharmacological Research*, *170*, 105535.

Cancer Dependency Map Portal (RRID:SCR_017655)

Chen, M., Asanuma, M., Takahashi, M., Shichino, Y., Mito, M., Fujiwara, K., ... & Iwasaki, S. (2021). Dual targeting of DDX3 and eIF4A by the translation inhibitor rocaglamide A. *Cell chemical biology*, *28*(4), 475-486.

Chiu, M. I., Katz, H., & Berlin, V. (1994). RAPT1, a mammalian homolog of yeast Tor, interacts with the FKBP12/rapamycin complex. *Proceedings of the National Academy of Sciences*, *91*(26), 12574-12578.

Chu, J., Zhang, W., Cencic, R., O'Connor, P. B., Robert, F., Devine, W. G., ... & Pelletier, J. (2020). Rocaglates induce gain-of-function alterations to eIF4A and eIF4F. *Cell reports*, *30*(8), 2481-2488.

Costales, M. G., Aikawa, H., Li, Y., Childs-Disney, J. L., Abegg, D., Hoch, D. G., ... & Disney, M. D. (2020). Small-molecule targeted recruitment of a nuclease to cleave an oncogenic RNA in a mouse model of metastatic cancer. *Proceedings of the National Academy of Sciences*, *117*(5), 2406-2411.

M Cowell, S., & Sun Lee, Y. (2016). Biphalin: the foundation of bivalent ligands. *Current medicinal chemistry*, *23*(29), 3267-3284.

Ernst, J. T., Thompson, P. A., Nilewski, C., Sprengeler, P. A., Sperry, S., Packard, G., ... & Reich, S. H. (2020). Design of development candidate eFT226, a first in class inhibitor of eukaryotic initiation factor 4A RNA helicase. *Journal of medicinal chemistry*, *63*(11), 5879-5955.

Fabbri, L., Chakraborty, A., Robert, C., & Vagner, S. (2021). The plasticity of mRNA translation during cancer progression and therapy resistance. *Nature Reviews Cancer*, *21*(9), 558-577.

Fan, A., & Sharp, P. P. (2021). Inhibitors of eukaryotic translational machinery as therapeutic agents. *Journal of Medicinal Chemistry*, *64*(5), 2436-2465.

Garnett, M. J., Edelman, E. J., Heidorn, S. J., Greenman, C. D., Dastur, A., Lau, K. W., ... & Benes, C. H. (2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature*, *483*(7391), 570-575.

Guan, X., Cheryala, N., Karim, R. M., Chan, A., Berndt, N., Qi, J., ... & Schonbrunn, E. (2022). Bivalent BET bromodomain inhibitors confer increased potency and selectivity for BRDT via protein conformational plasticity. *Journal of medicinal chemistry*, *65*(15), 10441-10458.

Haniff, H. S., Tong, Y., Liu, X., Chen, J. L., Suresh, B. M., Andrews, R. J., ... & Disney, M. D. (2020). Targeting the SARS-CoV-2 RNA genome with small molecule binders and ribonuclease targeting chimera (RIBOTAC) degraders. *ACS Central Science*, *6*(10), 1713-1721.

Horlbeck, M. A., Gilbert, L. A., Villalta, J. E., Adamson, B., Pak, R. A., Chen, Y., ... & Weissman, J. S. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *elife*, *5*, e19760.

Hutchings, M. I., Truman, A. W., & Wilkinson, B. (2019). Antibiotics: past, present and future. *Current opinion in microbiology*, *51*, 72-80.

Iwasaki, S., Floor, S. N., & Ingolia, N. T. (2016). Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature*, *534*(7608), 558-561.

Iwasaki, S., Iwasaki, W., Takahashi, M., Sakamoto, A., Watanabe, C., Shichino, Y., ... & Ingolia, N. T. (2019). The translation inhibitor rocaglamide targets a bimolecular cavity between eIF4A and polypurine RNA. *Molecular cell*, *73*(4), 738-748.

Leppek, K., Das, R., & Barna, M. (2018). Functional 5′ UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nature reviews Molecular cell biology*, *19*(3), 158-174.

Lou, K., Wassarman, D. R., Yang, T., Paung, Y., Zhang, Z., O'Loughlin, T. A., ... & Shokat, K. M. (2022). IFITM proteins assist cellular uptake of diverse linked chemotypes. *Science*, *378*(6624), 1097-1104.

Maniaci, C., Hughes, S. J., Testa, A., Chen, W., Lamont, D. J., Rocha, S., ... & Ciulli, A. (2017). Homo-PROTACs: bivalent small-molecule dimerizers of the VHL E3 ubiquitin ligase to induce self-degradation. *Nature communications*, *8*(1), 830.

Meijer, H. A., Schmidt, T., Gillen, S. L., Langlais, C., Jukes-Jones, R., de Moor, C. H., ... & Bushell, M. (2019). DEAD-box helicase eIF4A2 inhibits CNOT7 deadenylation activity. *Nucleic Acids Research*, *47*(15), 8224-8238.

Meric, F., & Hunt, K. K. (2002). Translation initiation in cancer: a novel target for

therapy. *Molecular Cancer Therapeutics*, *1*(11), 971-979.

Poulin-Kerstien, A. T., & Dervan, P. B. (2003). DNA-templated dimerization of hairpin polyamides. *Journal of the American Chemical Society*, *125*(51), 15811-15821.

Rees, M. G., Seashore-Ludlow, B., Cheah, J. H., Adams, D. J., Price, E. V., Gill, S., ... & Schreiber, S. L. (2016). Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nature chemical biology*, *12*(2), 109-116.

Robichaud, N., Sonenberg, N., Ruggero, D., & Schneider, R. J. (2019). Translational control in cancer. *Cold Spring Harbor perspectives in biology*, *11*(7), a032896.

Ruggero, D., & Pandolfi, P. P. (2003). Does the ribosome translate cancer?. *Nature Reviews Cancer*, *3*(3), 179-192.

Shegokar, R., Al Shaal, L., & Mishra, P. R. (2011). SiRNA delivery: challenges and role of carrier systems. *Die Pharmazie-An International Journal of Pharmaceutical Sciences*, *66*(5), 313-318.

Schmidt, T., Dabrowska, A., Waldron, J. A., Hodge, K., Koulouras, G., Gabrielsen, M., ... & Bushell, M. (2023). eIF4A1-dependent mRNAs employ purine-rich 5'UTR sequences to activate localised eIF4A1-unwinding through eIF4A1-multimerisation to facilitate translation. *Nucleic acids research*, *51*(4), 1859-1879.

Taha, E. A., Lee, J., & Hotta, A. (2022). Delivery of CRISPR-Cas tools for in vivo genome editing therapy: Trends and challenges. *Journal of Controlled Release*, *342*, 345-361.

Tong, Y., Lee, Y., Liu, X., Childs-Disney, J. L., Suresh, B. M., Benhamou, R. I., ... & Disney, M. D. (2023). Programming inactive RNA-binding small molecules into bioactive degraders. *Nature*, 1-11.

Wilczynska, A., Gillen, S. L., Schmidt, T., Meijer, H. A., Jukes-Jones, R., Langlais, C., ... & Bushell, M. (2019). eIF4A2 drives repression of translation at initiation by Ccr4-Not through purine-rich motifs in the 5′ UTR. *Genome Biology*, *20*, 1-21.

Publishing Agreement

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

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

DocuSigned by:
Megan Moore
2A20F69703A54C9...
Author Signature

8/17/2023
Date