

Mechanistic investigation of protein homeostasis by two key factors, ATAD1 and eIF2A

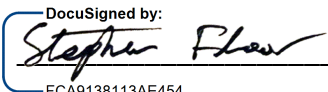
by
Hannah Toutkoushian

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

in
Biochemistry and Molecular Biology

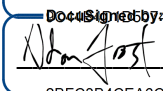
in the
GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

FCA9138113AE454... Stephen Floor
Chair

DocuSigned by:

Peter Walter

DocuSigned by:

2BFC3B4CEA6C478... Adam Frost

Committee Members

ACKNOWLEDGMENTS

It would be impossible to express the importance that the people in my life played up to this point and in pursuit of this degree, however, I am going to try my best.

Starting from the beginning, I would like to thank my advisor from the first lab that I worked in, Dr. Cory Momany and his student Nickolaus Galloway. Without them, I may still be on the misguided course of being an orthopedic surgeon. They demonstrated a joy and obsession with answering questions that was contagious, and instantly hooked me on the scientific process.

I would like to thank my current advisor, Dr. Peter Walter, for the endless support and experiences over my time in graduate school. You have taught me how to be patient with the scientific process and follow the data rather than a desired narrative. I feel very fortunate to have you as a mentor. Thank you for cultivating a positive, collaborative environment that was truly a joy to learn in. Among the great people I worked with in the lab, I would especially like to thank Silvia Ramundo and Lan Wang. I worked with Silvia and Lan at two opposite ends of my PhD, and both have had lasting effects on not only my PhD, but also my thinking as a scientist. Silvia, thank you for being an excellent rotation mentor and friend. Lan, working with you on ATAD1 was the single best thing I did in graduate school. You are going to make an excellent PI and I am going to miss you dearly.

A big thank you to my thesis committee members, Stephen Floor and Adam Frost. Your scientific advice and general advice on navigating graduate school was invaluable to me. While I can't say I ever looked forward to a thesis committee meeting, I always left feeling a renewed sense of purpose and excitement for future experiments.

Thank you to my Tetrad classmates. Irene, thank you for being such a patient, reliable friend that I am so lucky to have. Our various dinners, concerts, and late-night hangouts are a true source of sanity for me.

Thank you, Chris, for all the great conversations and fun times throughout this whole process, I hope we continue to go to Connecticut Yankee for years to come. Paige, thank you for being the insightful, supportive friend that I don't deserve. Your advice on relationships, school, and life have genuinely impacted my mindset, and made the difficult parts of graduate school a lot easier. Adam, I don't know what I would've done without your friendship during this experience. You have been an endless source of laughs, comfort, and support. I am going to miss our weekly lunches but I'm sure we will find an alternative.

To my partner Victor, I am truly so lucky to have lost my debit card and found you. Thank you for being such a bright, thoughtful, generous, funny human. I had no idea how fun life could be until I met you and I can't wait to see what our future looks like. Thank you for bringing me into your family and to your parents, Serge and Olga Charapaev, for accepting me with such generosity. Thank you to our pandemic pup Obie, a small living muppet that has brought an exceptional amount of light and love to our lives.

Finally, I would like to thank my family, in particular my parents, Jenni and Rob Toutkoushian. I am grateful to have had the importance of education instilled in me ever since I could remember. To my Mom, I want to thank you for showing me what unconditional love looks like. As a mother of one dog, I cannot imagine what the task of raising four human beings entails. To my Dad, thank you for teaching me not to take myself too seriously. That life skill has carried me more than I would have ever imagined. Thank you both for the constant support and guidance. I (literally) wouldn't be here if not for you.

CONTRIBUTIONS

The work included in this dissertation was all performed under the guidance of Dr. Peter Walter and Dr. Lan Wang.

Chapter 1 contains work from a manuscript in preparation:

Wang, L., Toutkoushian, H., Belyy, V., Kokontis, C., & Walter, P. 2021. Conserved structural elements specialize ATAD1 as a membrane protein extraction machine. <https://doi.org/10.1101/2021.09.24.461712>

Chapter 2 contains unpublished work in collaboration with Dr. Jin Chen of UT Southwestern (Ribo-seq analysis), Dr. Masaaki Sokabe and Dr. Chris Fraser of UC Davis (*in vitro* binding assays).

Mechanistic investigation of protein homeostasis by two key factors, ATAD1 and eIF2A

Hannah Toutkoushian

ABSTRACT

Maintaining protein homeostasis is essential for a proper functioning cell. From the time a polypeptide chain begins to be synthesized to the time the final product arrives at its destination there is an incredible amount of regulation taking place. The dysregulation of these processes can have disastrous effects for the cell. In this thesis I will be exploring two important areas of regulation: translation initiation and protein targeting. In Chapter 1, I will show structural, biochemical, and cell-based data characterizing the AAA ATPase ATAD1. ATAD1 is tasked with removing mistargeted tail-anchored proteins from the mitochondrial outer membrane. Through cryo-electron microscopy, we solved the structure of this protein with a peptide trapped in the central pore and designed a cell-based assay to validate structure-based predictions. We identified conserved structural elements that are essential for the function of ATAD1, in particular, the unusually aromatic nature of the pore loops that enable ATAD1 to extract its hydrophobic substrates.

In Chapter 2, I will be describing our characterization of an alternative translation initiation factor eukaryotic initiation factor 2A (eIF2A). From the initial discovery of this protein in the 1970's, eIF2A has been implicated in non-canonical translation initiation processes and has been shown to be important for responding to stress. Using available genome-wide genetic interaction datasets for *Saccharomyces cerevisiae*, we identified and characterized the genetic interaction between eIF2A and TIF1 (the DEAD-box helicase component of the eIF4F complex, eIF4A). We found that strains with both proteins deleted exhibited translational reprogramming that resulted in a sensitivity to metabolic stressors. As an exciting entry into detailed *in vitro* characterization, we also overcame a hurdle in the recombinant expression of

the human version of eIF2A and have begun re-examining the previously held assumptions about the mechanism of eIF2A-dependent translation.

TABLE OF CONTENTS

Chapter 1: Conserved structural elements specialize ATAD1 as a membrane protein extraction machine	1
Summary.....	2
Introduction	3
Results.....	6
Hinge motion between the large and the small AAA domains accompanies subunit translocation	7
A unique α -helix at the C-terminus mediates intersubunit interactions	8
Direct visualization and quantification of ATAD1's activity in cells	10
The highly aromatic central pore of ATAD1 is crucial for substrate binding and extraction.....	14
Multiple aromatic amino acids in the central pore are essential for ATAD1's function ..	15
Discussion	16
Figures.....	19
Materials and methods	45
References	51
Chapter 2: Characterization of the alternative translation initiation factor eukaryotic initiation factor 2A (eIF2A)	58
Introduction	59
Results	65
eIF2A and eIF4A (TIF1) interact genetically in <i>Saccharomyces cerevisiae</i>	65
Strains without eIF2A and TIF1 show a strong defect in translation	66

Ribosome profiling shows translational changes in the $\Delta eIF2A\Delta TIF1$ background.....	66
$\Delta eIF2A\Delta TIF1$ yeast exhibit a sensitivity to homeostatic stressors.....	68
Recombinant expression of eIF2A.....	69
Preliminary eIF2A <i>in vitro</i> binding assays	72
Discussion	74
Figures.....	76
Materials and methods	86
References	91

LIST OF FIGURES

Figure 1.1 Architecture of the $\Delta 40$ -ATAD1 ^{E193Q} -substrate complexes.....	20
Figure 1.2 Helix $\alpha 11$ resides at the intersubunit interface of ATAD1.....	21
Figure 1.3 A microscopy assay allows for the direct measurement and quantification of ATAD1's activity.....	22
Figure 1.4 Helix $\alpha 11$ mediates hexamer assembly	24
Figure 1.5 Pore-loop 1 aromatic amino acids are important for ATAD1's function both <i>in vivo</i> and <i>in vitro</i>	25
Figure 1.1 – Figure supplement 1 Sequence alignment of AAA _{MC} proteins	26
Figure 1.1 – Figure supplement 2 Size exclusion trace of $\Delta 40$ -ATAD1 ^{E193Q}	27
Figure 1.1 – Figure supplement 3 Cryo-EM processing of the $\Delta 40$ -ATAD1 ^{E193Q} -substrate complexes.....	28
Figure 1.1 – Figure supplement 4 Hinge motion between the large and the small subdomain accompanies subunit movements.....	30
Figure 1.2 – Figure supplement 1 Helices $\alpha 11$ form an additional spiral staircase beneath pore-loop 3	31
Figure 1.2 – Figure supplement 2 ATAD1 possesses a unique helix at the C-terminus that is structured differently from those in other AAA _{MC} proteins.....	32
Figure 1.3 – Figure supplement 1 Verification of ATAD1 knockout by western blot	33
Figure 1.3 – Figure supplement 2 Live-cell imaging showing the ATAD1 dependent localization of EGFP-Pex26	34
Figure 1.3 – Figure supplement 3 Live-cell imaging showing the ATAD1 dependent localization of EGFP-Gos28 in WT HeLa cells	35

Figure 1.4 – Figure supplement 1 Reproducibility of EGFP-Gos28 mislocalization in live-cell imaging with expression of $\Delta\alpha11$ mutant	36
Figure 1.4 – Figure supplement 2 ATAD1 $\Delta\alpha11$ expression level impacts substrate mislocalization in WT HeLa cells	37
Figure 1.4 – Figure supplement 3 Size exclusion chromatography of $\Delta40$ -ATAD1 and $\Delta40$ -ATAD1 $\Delta\alpha11$	38
Figure 1.4 – Figure supplement 4 Testing the effect of the disease-relevant mutations of ATAD1	39
Figure 1.4 – Figure supplement 5 ATPase assay of $\Delta40$ -ATAD1 and $\Delta40$ -ATAD1 $\Delta\alpha11$	41
Figure 1.5 – Figure supplement 1 Live-cell imaging showing the pore-loop dependent localization of EGFP-Gos28 in ATAD1 ^{-/-} HeLa cells	42
Figure 1.5 – Figure supplement 2 Reproducibility of EGFP-Gos28 mislocalization in live-cell imaging with expression of pore loop mutants	44
Figure 2.1 PTEN does not have isoform specific changes in eIF2A knockout HeLa cells	76
Figure 2.2 eIF2A and TIF1 have a synthetic growth defect.....	77
Figure 2.3 General translation is downregulated in the Δ eIF2A Δ TIF1 background.....	78
Figure 2.4 Ribosome profiling analysis reveals transcripts dependent on the combined presence of eIF2A and TIF1, with a distinct lack of changes in the Δ eIF2A background	79
Figure 2.5 The Δ eIF2A Δ TIF1 strain exhibits metabolic reprogramming and an increased sensitivity to homeostatic stressors.....	80
Figure 2.6 Expression and purification of 6xHis-eIF2A in <i>E.coli</i> results in the purification of the endogenous protein, GlnS.	82
Figure 2.7 eIF2A is readily expressed in Expi293 cells and elutes as a dimer on the size exclusion column.....	83
Figure 2.8 Preliminary binding experiments show 6xHis-ybbR-eIF2A has a moderate affinity for ribosomal subunits and a weak affinity for mRNA/tRNA _i	84

Figure supplement 2.1 WT and $\Delta eIF2A$ cells have similar polysome profiles 85

LIST OF TABLES

Table 1.1: Cryo-EM data analysis.....	19
Table 2.1: Antibodies for validation of ribosome profiling hits.....	89

LIST OF ABBREVIATIONS

AAA	ATPase associated with diverse cellular activities
ADH4	alcohol dehydrogenase 4
ATAD1	ATPase family AAA domain containing 1
ATF4	activation transcription factor 4
BCL	B-cell lymphoma
BiP	binding immunoglobulin protein
Cryo-EM	cryo-electron microscopy
EIF2	eukaryotic initiation factor 2
EIF2A	eukaryotic initiation factor 2A
EIF4A	eukaryotic initiation factor 4A
EGFP	enhanced green fluorescent protein
ER	endoplasmic reticulum
EMAP	epistatic mini array profiles
ESCRT	endosomal sorting complexes required for transport
GCN4	general control nondepressible 4
GET	guided entry of tail-anchored proteins
GLMS	glutamine--fructose-6-phosphate aminotransferase
HCV	hepatitis C virus
HEK	human embryonic kidney
IRES	internal ribosome entry site
ISR	integrated stress response
MOM	mitochondrial outer membrane
MSP1	mitochondrial AAA-ATPase
PARS	parallel analysis of RNA structure

PTEN	phosphatase and tensin homolog
RPKM	reads per kilobase of transcript, per million mapped reads
SCC	squamous cell carcinoma
SNARE	soluble N-ethylmaleimide-sensitive factor attachment protein receptor
TA	tail-anchored
TE	translational efficiency
TORC1	target of rapamycin complex 1
TRC	transmembrane domain recognition complex
tRNA _i	initiator tRNA
uORF	upstream open reading frame

CHAPTER 1

Conserved structural elements specialize ATAD1 as a membrane protein extraction machine

SUMMARY

The mitochondrial AAA protein ATAD1 (in humans; Msp1 in yeast) removes mislocalized membrane proteins, as well as stuck import substrates from the mitochondrial outer membrane, facilitating their re-insertion into their cognate organelles and maintaining mitochondria's protein import capacity. In doing so, it helps to maintain proteostasis in mitochondria. How ATAD1 tackles the energetic challenge to extract hydrophobic membrane proteins from the lipid bilayer and what structural features adapt ATAD1 for its particular function has remained a mystery. Previously, we determined the structure of Msp1 in complex with a peptide substrate (Wang et al., 2020). The structure showed that Msp1's mechanism follows the general principle established for AAA proteins while adopting several structural features that specialize it for its function. Among these features in Msp1 was the utilization of multiple aromatic amino acids to firmly grip the substrate in the central pore. However, it was not clear whether the aromatic nature of these amino acids were required, or if they could be functionally replaced by aliphatic amino acids. In this work, we determined the cryo-EM structures of the human ATAD1 in complex with a peptide substrate at near atomic resolution. The structures show that phylogenetically conserved structural elements adapt ATAD1 for its function while generally adopting a conserved mechanism shared by many AAA proteins. We developed a microscopy-based assay reporting on protein mislocalization, with which we directly assessed ATAD1's activity in live cells and showed that both aromatic amino acids in pore-loop 1 are required for ATAD1's function and cannot be substituted by aliphatic amino acids. In addition, we also discovered a short α -helix at the C-terminus that strongly facilitates ATAD1's oligomerization, a structural feature that distinguishes ATAD1 from its closely related proteins.

INTRODUCTION

Mitochondria serve a multitude of functions, including ATP production, metabolism and proteostasis that all require import of proteins from the cytosol (Pfanner et al., 2019). To ensure proper function, multiple mechanisms facilitate faithful protein targeting and efficient protein import into the organelle. One of these mechanisms is carried out by a protein on the mitochondrial outer membrane (MOM) named ATAD1 (in humans; Msp1 in yeast). ATAD1 has important roles in various biological contexts (Fresenius and Wohlever, 2019; Wang and Walter, 2020), including clearing out mistargeted proteins from the mitochondria (Chen et al., 2014; Nuebel et al., 2021; Okreglak and Walter, 2014), extracting mitochondrial precursor proteins stuck in the protein translocase channel during protein import overload (Weidberg and Amon, 2018), and mediating apoptosis by removing a BCL-family member protein (Winter et al., n.d.). In addition, ATAD1 has also been implicated in the regulation of synaptic plasticity by mediating the endocytosis of neurotransmitter receptors from post-synaptic membranes (Wang and Walter, 2020; Zhang et al., 2011). Among these, the removal of mistargeted tail-anchored (TA) proteins from the MOM is ATAD1's best understood function to date.

TA proteins are integral membrane proteins that are embedded in the membrane by a hydrophobic stretch at the extreme C-terminus. They are targeted to organellar membranes post-translationally. The endoplasmic reticulum (ER)-targeted TA proteins are delivered to the ER membrane by the TRC proteins (Get proteins in yeast) (Favaloro et al., 2008; Schuldiner et al., 2008; Stefanovic and Hegde, 2007). From there, they either stay on the ER membrane or travel to the peroxisome or other membranes along the secretory pathway. By contrast, TA proteins on the MOM either insert spontaneously (Chio et al., 2017) or are actively inserted by the mitochondrial import machinery (Doan et al., 2020). The partitioning of the ER- and the mitochondria- targeted TA proteins relies on different biophysical properties of their targeting signals (consisting of the transmembrane domain and the short segment C-terminal to it) (Chio

et al., 2017). Due to a number of similarities between the ER- and mitochondria-targeted TA proteins, mistargeting happens in both directions. Recently, our lab and the Rutter lab independently discovered that Msp1/ATAD1 could recognize the mislocalized TA proteins and extract them from mitochondria (Chen et al., 2014; Okreglak and Walter, 2014). The extracted proteins are then correctly inserted into the ER membrane, from which they travel to their cognate organelle or become degraded by the proteasome (Dederer et al., 2019; Matsumoto et al., 2019).

ATAD1 is a member of a large family of proteins called the AAA (ATPase Associated with diverse cellular Activities) proteins. It is strictly conserved from yeast to humans. Previously, we determined a series of cryo-EM structures of the cytosolic domain of the *Chaetomium thermophilum* (*C.t.*) Msp1 in complex with a peptide substrate (Wang et al., 2020). The structures revealed that Msp1 follows the general principle established for many AAA proteins (Gates et al., 2017; Monroe et al., 2017; Peña et al., 2018; Puchades et al., 2020, 2017; Zehr et al., 2017): six Msp1 subunits form a helical hexamer resembling a right-handed lock-washer that surrounds the substrate in a hydrophobic central pore. Elements at the intersubunit interface couple ATP hydrolysis with stepwise subunit translocation to unfold the substrate peptide in its central pore. The structures also revealed elements in Msp1 that are adapted for its function of removing membrane proteins, one of which is an unusually hydrophobic central pore. Whereas most AAA proteins extend only one short loop (pore-loop 1) containing a conserved aromatic amino acid to directly contact the substrate, Msp1 utilizes a total of three aromatic amino acids within two short loops to contact the substrate. We previously proposed that these aromatic amino acids enhance the bulkiness and hydrophobicity of the pore-loops, giving Msp1 a firm grip on the substrate to prevent it from backsliding, a feature that may be important for pulling hydrophobic membrane proteins out of the lipid bilayer.

To test the importance of the additional aromatic amino acids, we previously mutated the second (Y167 in *S. cerevisiae* (*S. c.*), corresponding to Y188 in *C. t.*) and the third (H206 in *S. c.*, corresponding to H227 in *C. t.*) aromatic amino acids to aliphatic or polar amino acids and measured their effects on yeast growth. Surprisingly, while the Y167A mutation impacted Msp1's activity, the Y167V or H206A mutation did not have a significant impact on Msp1's function. We hypothesized that this was perhaps because the yeast growth assay provides an indirect readout (cell death) of Msp1's function. In other words, deficiencies in the protein's activity could be masked by compensatory pathways promoting cell survival. Therefore, in this study, we aimed to establish an assay that would allow the direct visualization of protein mislocalization.

In addition to addressing these remaining questions concerning the pore-loops, we also explored unique structural elements in ATAD1 that are not observed in its closely related family members. Within the large AAA protein family, ATAD1 belongs to the "meiotic clade" (AAA_{MC}). Members of this clade include microtubule severing proteins such as katanin and spastin, as well as the protein that disassembles the ESCRT peptides, Vps4 (Frickey and Lupas, 2004). One of the unique structural elements shared by these family members is a C-terminal extension of the classic AAA protein fold, helix α 12. α 12 is important in hexamer assembly and protein function for AAA_{MC} proteins (Sandate et al., 2019; Vajjhala et al., 2008). The secondary structure prediction of ATAD1, however, shows the lack of α 12, and instead predicts a longer α 11 (the alpha helix preceding it, Figure 1.1-figure supplement 1). Of note, there are several disease-relevant mutations clustered around α 11, making it particularly important to understand this region in molecular detail: R9H, D221H and E290K are found in Schizophrenia patients (Umanah et al., 2017); H357Rfs*15, (a frame shift mutation resulting in a 10 amino acid extension at the C-terminus) is found in a family of patients with encephalopathy (Piard et al., 2018); and E267stop results in a truncated protein that is missing a large portion of the C-terminal domain and is found in a family of patients suffering from hypertonia, seizure and death (Ahrens-Nicklas et al., 2017). In the *C.t.* Msp1

structure, we did not observe density for $\alpha 11$, which could be due to degradation or lack of rigid secondary structure, leaving the structure and function of $\alpha 11$ unresolved. To address these questions, we determined the structure of ATAD1 and tested the individual functional contributions of salient structural features in a microscopy-based mislocalization assay.

RESULTS

To obtain a homogenous sample for structural analysis, we expressed human ATAD1 lacking the first 40 amino acids (consisting of the transmembrane helix) with an N-terminal His-tag and bearing a commonly used “Walker B” mutation that inactivates ATP hydrolysis but retains ATP binding ($\Delta 40$ -ATAD1^{E193Q}). $\Delta 40$ -ATAD1^{E193Q} formed oligomers as assessed by size exclusion chromatography (Figure 1.1-figure supplement 2). We took the fraction corresponding to hexamers and incubated it with ATP before preparing samples for cryo-EM imaging.

3D classification of the particles resulted in two distinct hexameric structures. In the first structure, the six ATAD1 subunits (M1-M6) are arranged in a right-handed spiral with an open seam between the top and the bottom subunits (45,003 particles analyzed, Figure 1.1-figure supplement 3B, Table 1.1), an arrangement resembling the “open state” structure of Msp1 (Wang et al., 2020), as well as other related AAA proteins (Cooney et al., 2019; Han et al., 2019; Su et al., 2017; Sun et al., 2017; Twomey et al., 2019; Zehr et al., 2017) (Figure 1.1A). Refinement of this structure resulted in a map with an average resolution of 3.7 Å. In the second structure, the subunits are arranged in a similar fashion as the first, with the exception of M6, which is in transition to convert into M1 as part of the AAA protein functional ATPase cycle and now bridges M1 and M5, closing the hexameric ring (96,577 particles analyzed). This structure closely resembles the “closed state” structure observed in Msp1 (Figure 1.1B). Refinement of this structure yielded a map with an average resolution of 3.2 Å with most of the side chain densities

clearly visible, which allowed us to build an atomic model. Noticeably, the density of the M6 subunit in the “closed state” structure was less well defined than the rest of the complex (local resolution ranging from 4-6 Å, Figure 1.1-figure supplement 3C), indicating a mixture of different states. A similar dynamic nature of this subunit was also observed in the Msp1 structure (Wang et al., 2020) but the density was too poorly resolved to be interpreted. By contrast, for ATAD1, we could identify secondary structure elements clearly, which enabled us to build a polyalanine model for this subunit. In doing so, we were able to analyze the interactions between M6 and neighboring subunits in the closed state, which was not possible for Msp1.

Hinge motion between the large and the small AAA domains accompanies subunit translocation

The overall architecture of ATAD1 closely resembles that of Msp1, pointing to significant structure/function conservation throughout eukaryotic evolution. Six ATAD1 subunits rotate and translocate progressively to form a hexameric spiral assembly. Each ATAD1 subunit consists of two subdomains, a large subdomain followed by a small subdomain (Figure 1.1C), and the nucleotide is bound at the interface between the two. In the “open state”, all six subunits adopt similar overall structures. To initiate the subunit translocation, the M6 subunit loses its bound nucleotide, and translocates upward to bridge the gap with M1 (Figure 1.1-figure supplement 4A). Meanwhile, the small subdomain of M5 rotates towards its large subdomain, allowing it to establish contacts with the small subdomain of M6 in the “closed state” structure (Figure 1.1-figure supplement 4B). In the next step of translocation, M6 continues to move upwards, assuming the M1 position (top position) in the open spiral, detaching fully from M5, which now becomes the new M6, occupying the bottom position. Consequently, the angle between the two subdomains of M5 widens, and the subunit resumes its original conformation as it reaches the M6 position (Figure 1.1-figure supplement 4B and C). Similar hinge motions between the two subdomains were observed previously: the crystal structure of the monomeric *S.c.* Msp1 showed a near 180-degree flip between the small and the large AAA domains and complete disruption of the

nucleotide binding pocket (Wohlever et al., 2017). It remained unclear, however, whether the flip resulted from crystal packing forces or reflected a functionally relevant conformational change. The hinge motion in ATAD1 observed here strongly argues that the two domains are connected by a flexible linker and undergo significant rotations relative to each other in a nucleotide- and subunit position-dependent manner.

A unique α -helix at the C-terminus mediates intersubunit interactions

While examining the interactions of the mobile subunit (M6) in the closed conformation, we noticed that it is held in place by contacts on both sides: on the side of M5, substantial contacts exist between the large and the small subdomains of the two subunits (Figure 1.1 – figure supplement 4D). By contrast, on the side of M1, M6 is held in place by two contact points. One structural element stood out as it was not observed previously in the *C. t.* Msp1 structure (Wang et al., 2020): an α -helix at the extreme C-terminus of M6 points towards the core β sheet of the large subdomain of M1. We next asked whether this α -helix also mediates intersubunit contacts within the other subunits in the spiral. Indeed, slightly extended EM density for this α -helix exists in other subunits as well (Figure 1.1C and Figure 1.2A). We modeled an α -helix (α 11, amino acids 333-346) followed by a short turn (amino acids 347-351) into this density. Density beyond that (amino acids 352-361) was not clearly visible indicating its lack of rigid structure. A similar α -helix existed in the crystal structure of the monomeric *S.c.* Msp1 (Wohlever et al., 2017). However, the last few amino acids in that α -helix in the *S.c.* Msp1 were replaced by amino acids left from protease cleavage, making it difficult to interpret the original structure of this region.

In ATAD1's helical assembly, six α 11s form a staircase beneath pore-loop 3, which does not contact the substrate directly but forms an interconnected network with pore-loops 1 and 2 that directly engage the substrate (Figure 1.2 – figure supplement 1). Zooming in on α 11 reveals that it lies at the interface

between two adjacent subunits (Figure 1.2A). It forms extensive contacts with both the rest of the subunit and the *counterclockwise* (when viewed from the cytosol-facing side, as shown in figure 1.4) adjacent subunit, bridging between the two (Figure 1.2B and D). In $\alpha 11_{M4}$, for example, L336, A339 and I340 pack against the core of M4 (Figure 1.2A); M343 rests on the interface created by M4 and M5, and A349 points to a hydrophobic groove formed by pore-loop 3 of the M5. Finally, both the side chain and the backbone carbonyl of S346 form hydrogen bonds with R254 of M5 (Figure 1.2D). A few charged amino acids (R254, K342, K344, K345, K347, E341 and D348) point into the cytosol (Figure 1.2C and 1.2E).

Interestingly, other AAA_{MC} proteins such as katanin, spastin and Vps4 are structured differently in this region (Han et al., 2019, 2017; Sandate et al., 2019; Shin et al., 2019). For those proteins, $\alpha 11$ is shorter (consisting of three helical turns instead of four as seen in ATAD1). It is followed by a loop and another short helix, $\alpha 12$, that reaches across the intersubunit interface to contact the *clockwise* adjacent subunit (Figure 1.2-figure supplement 2). $\alpha 12$ is functionally important as its deletion impacts oligomer assembly and protein function both *in vitro* and *in vivo* (Vajjhala et al., 2008). Since ATAD1 lacks $\alpha 12$, and instead has a longer $\alpha 11$, we hypothesized that $\alpha 11$ could also mediate hexamer assembly. If so, we expected that a mutant version might fail to remove mislocalized membrane proteins in cells lacking wild-type (WT) ATAD1, but be unable to poison WT ATAD1's activity, as it would not be able to incorporate into WT hexamers. In other words, we expected it to display a recessive phenotype, rather than a dominant one. To test this notion, we next sought to establish an assay that distinguishes between dominant and recessive ATAD1 mutations and allows us to measure the activity of ATAD1 variants in a direct and quantifiable way.

Direct visualization and quantification of ATAD1's activity in cells

One of ATAD1's established substrates is Gos28, a Golgi-localized TA SNARE protein. In cells lacking functional ATAD1, Gos28 localizes to mitochondria (Chen et al., 2014). We thus sought to use the localization of Gos28 as a readout for ATAD1's activity (Figure 1.3A). To analyze the activity of ATAD1 mutants in the absence of the wild-type enzyme, we knocked out all alleles of *ATAD1* in HeLa cells using CRISPR/Cas9 (Figure 1.3-figure supplement 1). We next stably expressed Gos28 with an N-terminal EGFP tag in the ATAD1^{-/-} cells and used confocal microscopy to visualize the localization of EGFP-Gos28. In ATAD1^{-/-} cells, we observed two populations of Gos28 molecules: one visualized as an extended perinuclear region characteristic of the Golgi apparatus, and the other in a spread-out network (Figure 1.3B, top row). Overlap of the EGFP signal with the signal from the mitochondria stain, MitoTracker Deep Red, showed that the latter population corresponds to the mitochondrial network, indicating that in ATAD1^{-/-} cells a portion of EGFP-Gos28 molecules are mislocalized to mitochondria, in line with a previous study (Chen et al., 2014). We then expressed wild-type ATAD1 labeled with a C-terminal HaloTag in the ATAD1^{-/-} cells (Figure 1.3B, middle row) and observed a prominent shift of signal towards a Golgi-like distribution, indicating that ATAD1 cleared mislocalized EGFP-Gos28 from the mitochondria. Finally, we expressed the ATAD1 bearing a mutation that inactivates its ATPase activity, ATAD1^{E193Q}. The result mimicked the localization pattern seen in the ATAD1^{-/-} cells (Figure 1.3B, top row), indicating that the ability of clearing mislocalized EGFP-Gos28 was dependent on ATAD1's enzymatic activity. We similarly observed ATAD1-dependent removal of mistargeted Pex26, a peroxisomal TA protein, from mitochondria, suggesting that monitoring the clearance of a mislocalized protein is a reliable method to examine ATAD1-dependent TA protein extraction (Figure 1.3-figure supplement 2).

In addition to expressing the reporter in the ATAD1^{-/-} cell line, we stably expressed Gos28 in a wild-type HeLa cell line. We expected ATAD1 variants that impact its ability to form proper oligomers to induce

Gos28 mislocalization in the ATAD1^{-/-} cell line, but not in the WT cell line. By contrast, inactive ATAD1 variants that retain their ability to assemble into hexamers should act as dominant inhibitors of the enzyme and induce substrate mislocalization.

To test the reliability of the system, we first expressed a known dominant-negative mutant, ATAD1^{E193Q}, within the wild-type reporter cells (Figure 1.3 – figure supplement 3). Expression of ATAD1^{E193Q} induced mislocalization of EGFP-Gos28 to mitochondria, indicating that ATAD1^{E193Q} was incorporated into hexamers, poisoning wild-type activity. As expected, expression of the wild-type ATAD1 or an empty vector did not induce EGFP-Gos28 mislocalization in this background.

Before proceeding to test the function of $\alpha 11$, we wanted to see if we could quantify the degree of EGFP-Gos28 mislocalization in an unbiased fashion and use it as a readout for ATAD1's activity. To this end, we developed a data analysis pipeline using CellProfiler (McQuin et al., 2018) (Figure 1.3C). In brief, we first identified cells expressing ATAD1-HaloTag using the cell-permeable JF549 dye to label and visualize ATAD1 protein. We then measured the colocalization of EGFP-Gos28 and mitochondria (stained with MitoTracker) in ATAD1-expressing cells using the Pearson correlation coefficient (PCC) as the metric, measured between the two channels for each cell (Figure 1.3D). As shown in Figure 1.3D, expression of an empty vector or ATAD1^{E193Q} in the ATAD1^{-/-} reporter cell line led to a PCC of 0.31 ± 0.022 and 0.31 ± 0.017 , respectively. By contrast, transiently expressing wild-type ATAD1 showed a significantly lower PCC of 0.13 ± 0.015 ($p < 0.0001$), confirming the visually evident changes in EGFP-Gos28 localization and validating this approach for the evaluation of functional mutants of ATAD1.

With the EGFP-Gos28 reporter assay established in both an ATAD1^{-/-} background and a WT background, we next used it to test $\alpha 11$'s function. To this end, we made two truncated versions of ATAD1: one in which we deleted 30 amino acids at the C-terminus (removing $\alpha 11$ and everything C-terminal to it,

ATAD1 $\Delta\alpha$ 11, Figure 1.1C), and another in which we kept α 11 and removed everything C-terminal to it (ATAD1₁₋₃₅₁, Figure 1.1C). We first expressed these variants in the ATAD1^{-/-} cells. As shown in Figure 1.4A, EGFP-Gos28 showed significant mislocalization to mitochondria in cells expressing ATAD1 $\Delta\alpha$ 11 but not ATAD1₁₋₃₅₁, suggesting that the α 11 helix is required for its function. By contrast, when expressed in the wild-type cells, ATAD1 $\Delta\alpha$ 11 did not induce significant mislocalization of EGFP-Gos28 (Figure 1.4B), indicating its recessive phenotype, i.e., its lack of incorporation into and inactivating the wild-type hexamers present in these cells (the average values for the biological replicates are shown in Figure 1.4- figure supplement 1A and B). While expression of ATAD1 $\Delta\alpha$ 11 did not induce substrate mislocalization on average, it was evident by eye that a small population of wild-type cells expressing ATAD1 $\Delta\alpha$ 11 showed EGFP-Gos28 mislocalization (Figure 1.4- figure supplement 2A). We assume that this effect resulted from high ATAD1 $\Delta\alpha$ 11 expression in these outlier cells that compensated for ATAD1 $\Delta\alpha$ 11's reduced oligomerization ability and allowed assembly into wild-type hexamers. Indeed, we observed a positive correlation between ATAD1 $\Delta\alpha$ 11 expression level and the degree of Gos28 mislocalization: cells that expressed more ATAD1 $\Delta\alpha$ 11 showed a higher degree of Gos28 mislocalization (Figure 1.4-figure supplement 2B), suggesting that ATAD1 $\Delta\alpha$ 11 incorporated into wild-type hexamers and inactivated them, or that the ATAD1 $\Delta\alpha$ 11 subunits prevented the wild-type enzyme from forming stable hexamers. These data combined showed that α 11 plays an important role in both subunit oligomerization and hexamer function.

To further establish α 11's impact on ATAD1 oligomerization, we examined the elution profile of recombinantly expressed ATAD1 with and without the C-terminal helix. To this end, we purified the cytosolic domain of ATAD1 lacking the α 11 helix (Δ 40-ATAD1 $\Delta\alpha$ 11). On size-exclusion, Δ 40-ATAD1 $\Delta\alpha$ 11 eluted exclusively as a monomer. At similar concentrations, Δ 40-ATAD1 formed oligomers (Figure 1.4-figure supplement 3), indicating that α 11 mediates oligomerization *in vitro*. To evaluate the functional impact of a defect in hexamer formation, we next asked if the ability to bind a

peptide substrate was affected for $\Delta 40$ -ATAD1 $\Delta\alpha 11$. To this end, we sought to identify a peptide that binds ATAD1. First, we designed a peptide array by stepping through ATAD1's known substrates and measured the binding of individual peptides to ATAD1. We picked a few peptides that showed the strongest binding and tested their affinity to ATAD1 in solution using fluorescence polarization. Out of this group, a peptide (P13) that is derived from the sequence of the C-terminus of Pex26 (an established ATAD1 substrate) emerged as the strongest binder, which we then used to test the effect of truncating $\alpha 11$ on ATAD1's substrate binding affinity. As expected and consistent with its inability to form hexamers, $\Delta 40$ -ATAD1 $\Delta\alpha 11$ also showed substantially reduced peptide binding affinity (Figure 1.4C).

Several disease-related mutations are close to the $\alpha 11$ region at the C-terminus (see Introduction). We tested these ATAD1 variants using our imaging assay to assess if their ability to remove mislocalized TA proteins would be impacted. Out of the variants tested, the E267stop variant impacted ATAD1's function significantly, whereas the other variants (H357Rfs*15, D221H, R9H, and E290K) displayed WT-like phenotypes (Figure 1.4 – figure supplement 4). These results indicate that the mechanism that underlies neurological disorders may be separate from the extraction of Gos28.

As previously mentioned, $\alpha 11$ packs against pore-loop 3 from the adjacent subunit and constitutes the additional layer of the interconnected network involving three pore-loops (Figure 1.2 – figure supplement 1). The fact that it lies at the oligomerization interface and also contacts pore-loop 3 prompted us to ask whether $\alpha 11$ could couple oligomerization with substrate binding. Given that oligomerization usually enhances the ATPase activity of AAA proteins, we sought to measure the ATPase activity of $\Delta 40$ -ATAD1 and $\Delta 40$ -ATAD1 $\Delta\alpha 11$ and asked whether the activities can be stimulated by the addition of substrate. As shown in Figure 1.4 -figure supplement 5, $\Delta 40$ -ATAD1 $\Delta\alpha 11$ displays a much lower ATPase activity compared to $\Delta 40$ -ATAD1, indicating the lack of functional hexamers. With the addition of a peptide substrate (P13), the activity of $\Delta 40$ -ATAD1 is stimulated by around 41%, whereas no

significant stimulation was observed for $\Delta 40$ -ATAD1 $\Delta\alpha 11$ (Figure 1.4 – figure supplement 5). These results suggest that $\Delta 40$ -ATAD1 $\Delta\alpha 11$'s ability to oligomerize is significantly impacted and the addition of a substrate is not able to bring the monomeric subunits into a functional oligomeric form. Having examined the oligomer interface in detail, we next examined the substrate interactions within the central pore.

The highly aromatic central pore of ATAD1 is crucial for substrate binding and extraction

Upon building all of ATAD1's amino acids into the EM density, we observed a piece of density in the central pore that resembles a linear peptide in both the closed and the open state structures. This density likely represents a composite of peptides that co-purified with ATAD1 from *E. coli*. The side chain density was clearly visible, and we modeled the peptide as a 10-mer polyalanine. A peptide in the corresponding position was also observed in the Msp1 structures (Wang et al., 2020), suggesting the proteins' high affinity for their substrates.

Like for Msp1, six ATAD1 subunits tightly surround the peptide in the central pore. As summarized in the introduction, amino acids used by ATAD1 to contact the substrate are phylogenetically conserved within the Msp1/ATAD1 family, but differ from other AAA proteins. Each ATAD1 subunit extends two short loops containing three aromatic amino acids to engage the substrate peptide (Figure 1.1D and E). Six pore-loops 1 form a spiral staircase surrounding the peptide substrate. Pore-loop 1 is comprised of a conserved KX_1X_2G motif, where X_1 is a tryptophan (W166) and X_2 a tyrosine (Y167). Both aromatic amino acids intercalate into the side chains of the translocating peptide (Figure 1.1D), holding it in place. Below pore-loop 1, is a second staircase formed by pore-loops 2, which use a histidine (H206) to contact the backbone of the substrate (Figure 1.1E). By contrast, in most other known AAA proteins, the X_1

position within pore-loop 1 is an aromatic amino acid, and the X₂ position is an aliphatic amino acid that does not contact the substrate. Also, their pore-loops 2 usually do not contact the substrate directly.

Multiple aromatic amino acids in the central pore are essential for ATAD1's function

We next revisited our hypothesis that ATAD1 function, like Msp1 function, necessitates engagement of multiple aromatic amino acids in its central pore. To this end, we tested the effect of pore-loop mutations in the ATAD1^{-/-} EGFP-Gos28 reporter cell line (Figure 1.5A) using our cell-based assay. Expression of ATAD1 bearing either the W166A (the first aromatic amino acid in pore-loop 1) or Y167A (the second aromatic amino acid in pore-loop 1) mutation led to significant mislocalization of EGFP-Gos28, suggesting that these mutations inactivate ATAD1. By contrast, expression of the aromatic mutant Y167F cleared Gos28 from mitochondria, indicating functional ATAD1, which is consistent with what we observed for Msp1. Interestingly, changing W166 or Y167 to an aliphatic amino impacted ATAD1's activity: ATAD1^{W166V}, ATAD1^{Y167V} and ATAD1^{Y167L} were inactive, and the activity of ATAD1^{W166L} was also impacted, albeit to a lesser degree (Figure 1.5A, figure supplement 1 and figure supplement 2), supporting the notion that the aromaticity of this position is important. This observation, namely that an aliphatic amino acid did not functionally replace the aromatic amino acid in this position, contrasts with our data from corresponding mutations in Msp1 in the yeast growth assay (Wang et al., 2020), suggesting either a better sensitivity of the microscopy-based assay or an inherent difference between the yeast and the mammalian system. We also tested the effect of mutations on H206 (pore-loop 2) in our assay. While our structural data clearly showed an interaction between H206 and the substrate backbone, mutations in this position did not impact the function of ATAD1 in this context.

Having established the functional importance of the pore-loop 1 side chains in substrate extraction, we further examined their specific impact on ATAD1's ability to bind a substrate *in vitro*. Consistent with the cell-based assay, ATAD1^{W166A}, ATAD1^{W166V}, ATAD1^{W166L}, ATAD1^{Y166A} and ATAD1^{Y167V} showed

dramatically reduced (estimated > 100-fold) binding to P13 compared to wild-type ATAD1. Both ATAD1^{H206A} and ATAD1^{Y167F} showed wild-type like activities in the cell-based assay, but bound P13 with much weaker affinity than the wild-type enzyme in this assay. We reasoned that this difference resulted either because the *in vitro* assay is more sensitive to small changes in activity that cell-based assays fail to capture, or because a different substrate was used in the cell-based assay (Gos28) than the *in vitro* binding assay (P13, derived from Pex26). Taken together, the *in vitro* peptide binding results, implications from the structure, and the cell-based assay all converge on the conclusion that the aromaticity of the second amino acid in pore-loop 1 is important for ATAD1's function.

DISCUSSION

In this work, we present the cryo-EM structures of the soluble domain of the mitochondrial AAA protein ATAD1 in complex with a peptide substrate. Our structures show that the overall architecture of ATAD1 is very similar to that of its yeast homolog Msp1. ATAD1's architecture fully agrees with the conserved mechanism of hand-over-hand spiral propagation established for many AAA proteins in substrate translocation. We also discovered a distinct structural element, helix $\alpha 11$, that was present but remained structurally unresolved in Msp1, which differentiates ATAD1/Msp1 from other AAA_{MC} proteins. Multiple lines of evidence including cell-based mislocalization, substrate binding, and ATPase activity all suggest that $\alpha 11$ is required for the formation of functional oligomers. Although we do not know why it is structured differently than the other AAA_{MC} proteins (which adopt a $\alpha 11$ - $\alpha 12$ organization), one possibility is that $\alpha 11$ is useful in adaptor interaction. In particular, yeast Msp1 docks to the proteasome to channel its substrate directly for degradation (Basch et al., 2020). Another study showed that during mitochondrial import stress, Msp1 is recruited to the TOM complex by the adaptor Cis1 to remove precursor proteins stuck in the TOM translocase (Weidberg and Amon, 2018). We reason that structural

elements facing the cytosolic side of Msp1/ATAD1 (such as $\alpha 11$) may be evolutionarily adapted to interact with these factors. In this view, the $\alpha 11$ - $\alpha 12$ organizational differences in different AAA_{MC} protein subfamilies may have evolved to suit their specialized functions, such as binding to microtubules and ESCRT proteins for katanin and Vps4, respectively. This view is supported by the fact that ATAD1/Msp1 occupy the basal position in the evolutionary tree of this clade, and the other members (katanin, spastin and Vps4) share a common evolutionary ancestor (Frickey and Lupas, 2004).

The cell-based imaging assay allowed us to directly quantify ATAD1-dependent substrate mislocalization and thereby revisit a hypothesis raised in the previous study (Wang et al., 2020) on the functional importance of the pore-loop amino acids. The notion that aromaticity of pore loop/substrate-interacting sidechains is important is supported by single molecule force measurements of other AAA protein translocases. A study on the bacterial AAA proteins ClpXP (Rodriguez-Aliaga et al., 2016), for example, showed that the flat, bulkier pore-loop amino acids give the protein a stronger grip, that is higher coupling efficiency between ATP hydrolysis and substrate unfolding, yet are compromised in the rate of substrate translocation (measured as protein unfolding with ClpXP). Given ATAD1's function is to extract hydrophobic membrane proteins from the lipid membrane, having a tighter grip on the substrate may be important, even if it comes at the cost of a slower translocation speed. Similar to ATAD1, Cdc48, a protein that removes misfolded membrane proteins from the ER membrane, also possesses two aromatic amino acids in its pore-loop 1 (Cooney et al., 2019; Twomey et al., 2019), suggesting that the additional aromatic amino acids in the central pore evolved to aid in the removal of membrane proteins.

To conclude, while ATAD1/Msp1 utilize a conceptually similar mechanism for substrate translocation as many other AAA proteins, the high conservation of several unique features between ATAD1 and Msp1 suggests that evolution fine-tuned these enzymes early in eukaryotic cell evolution for their special role in membrane protein extraction and protein quality control. The phylogenetic comparison of structure and

function, and the complementary experimental opportunities afforded by structural and cell-level analysis, allowed us to extract hints regarding important functional principles and assess their generality over a wide span of evolutionary time. The neurological phenotypes associated with a number of ATAD1 mutations still remain poorly understood, yet they serve to underscore the importance of proteostasis in human physiology and raise hope that understanding these principles may help to develop new treatments for combating devastating pathological dysfunctions arising from proteostasis imbalances.

FIGURES

Table 1.1 Cryo-EM data analysis

Structure	$\Delta 40\text{-ATAD1}^{\text{E193Q}}$ (closed)	$\Delta 40\text{-ATAD1}^{\text{E193Q}}$ (open)
Data collection		
Microscope	Titan Krios	
Voltage (keV)	300	
Nominal magnification	105000x	
Exposure navigation	Image shift	
Electron dose ($\text{e}^{-}\text{\AA}^{-2}$)	67	
Dose rate ($\text{e}^{-}/\text{pixel}/\text{sec}$)	15	
Detector	K3 summit	
Pixel size (\AA)	0.832	
Defocus range (μm)	0.6-2.0	
Micrographs	6937	
Reconstruction		
Total extracted particles (no.)	478463	
Final particles (no.)	96577	45003
Symmetry imposed	C1	C1
FSC average resolution, masked (\AA)	3.2	3.5
FSC average resolution, unmasked (\AA)	4.1	4.6
Applied B-factor (\AA)	121.9	119.2
Reconstruction package	Cryosparc v2.15 and Relion 3.0.8	
Refinement		
Protein residues	1676	1694
Ligands	17	19
RMSD Bond lengths (\AA)	0.003	0.017
RMSD Bond angles ($^{\circ}$)	0.715	1.066
Ramachandran outliers (%)	0.00	0.06
Ramachandran allowed (%)	9.44	7.22
Ramachandran favored (%)	90.56	92.72
Poor rotamers (%)	9.63	11.29
CaBLAM outliers (%)	6.15	5.64
Molprobity score	2.81	2.81
Clash score (all atoms)	10.18	10.87
B-factors (protein)	118.19	171.87
B-factors (ligands)	93.95	161.45
EMRinger Score	1.42	1.05
Refinement package	Phenix 1.17.1-3660-000	

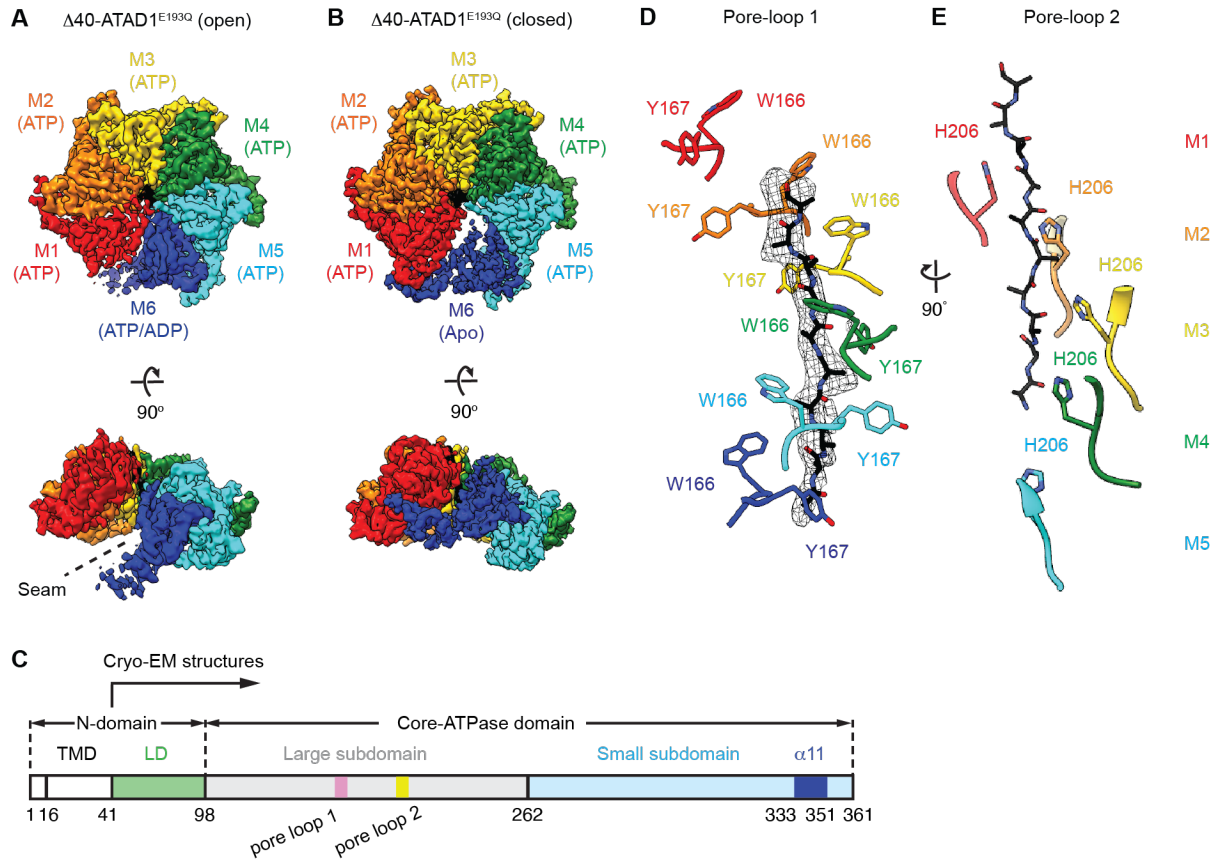


Figure 1.1 Architecture of the $\Delta 40$ -ATAD1^{E193Q}-substrate complexes

(**A** and **B**) Final reconstructions of $\Delta 40$ -ATAD1^{E193Q} (open) and $\Delta 40$ -ATAD1^{E193Q} (closed) complexes shown in top and side views. Each subunit (M1 to M6) is assigned a distinct color, and the substrate is shown in black. The spiral seam of the open conformation (panel A) is denoted with dashed lines. In the top views, the membrane-facing side of ATAD1 is shown. (**C**) Schematic of individual domains and structural elements of human ATAD1. (**D**) Pore-loops 1 (showing those of the closed conformation) form a staircase around the substrate. The peptide density is shown in black mesh. (**E**) Pore-loops 2 form a second staircase below pore-loops 1. H206s directly contact the peptide backbone carbonyls.

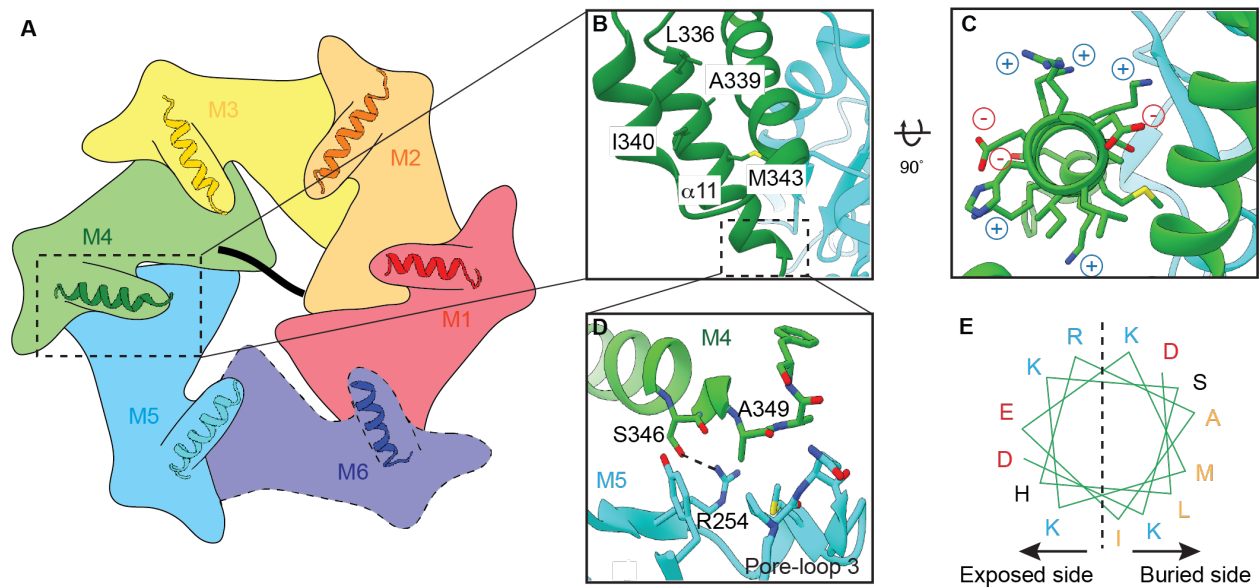


Figure 1.2 Helix $\alpha 11$ resides at the intersubunit interface of ATAD1

(A) Schematic showing the position of helix $\alpha 11$ within each ATAD1 subunit. ATAD1 subunits are colored as in Figure 1. $\alpha 11$ is shown in cartoon representation. Different from Figure 1, here, the cytosol-facing side (instead of the membrane-facing side) of ATAD1 is shown, and the order of subunits from M1, M2 to M6 goes counterclockwise. (B to D) Zoomed-in view of $\alpha 11$ interaction with the neighboring subunits. (E) Helical wheel showing the amphipathic property of $\alpha 11$. Hydrophobic amino acids are colored in orange, positive charged amino acids in blue, negative charged amino acids in red and polar amino acids in purple.

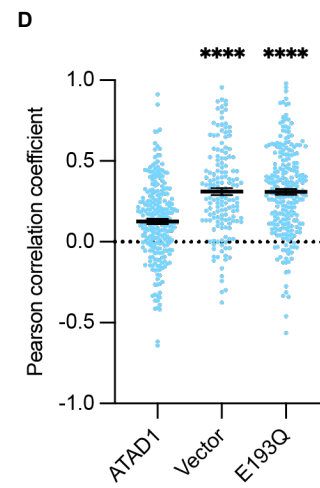
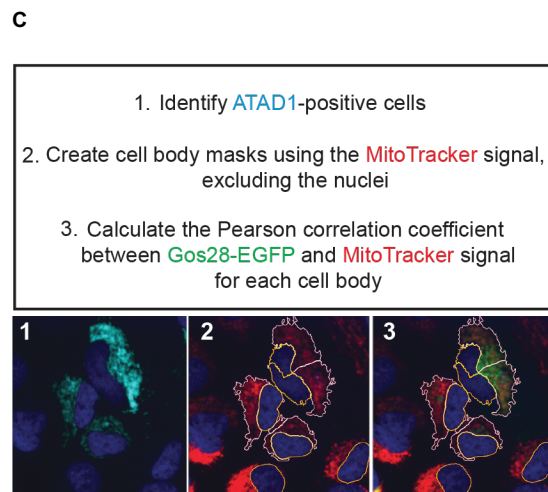
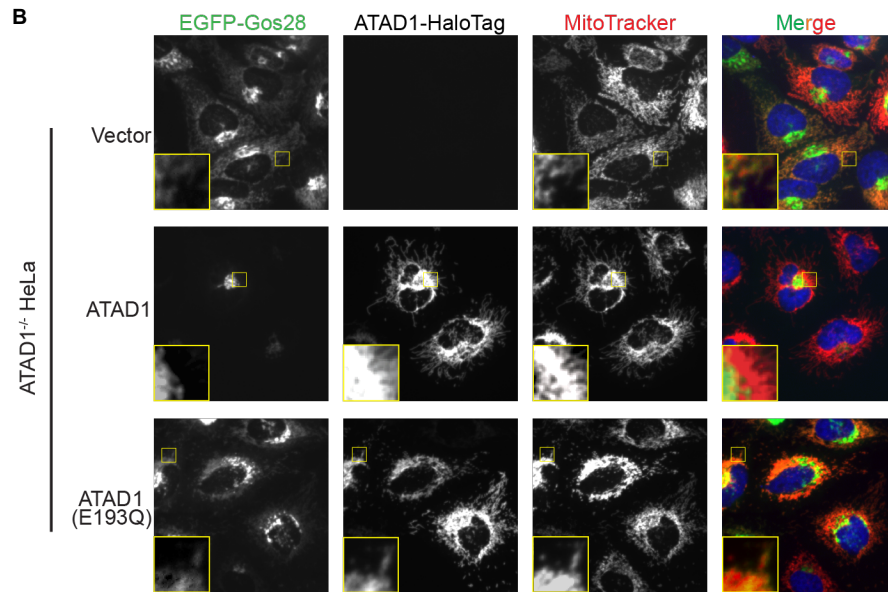
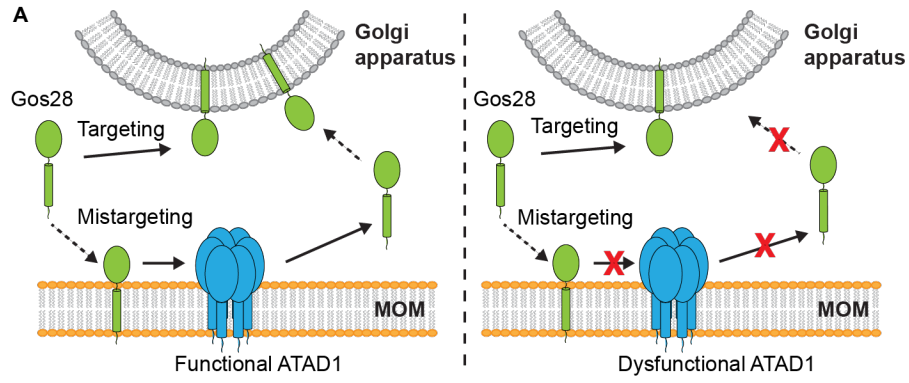


Figure 1.3 A microscopy assay allows for the direct measurement and quantification of ATAD1's activity

(A) Model for ATAD1-dependent extraction of Gos28: in cells expressing a functional ATAD1 (left panel), mislocalized Gos28 on the MOM is extracted by ATAD1 and given a second chance to return to the Golgi apparatus. In cells expressing a dysfunctional ATAD1 (right panel), Gos28 accumulates on the MOM. **(B)** Representative average intensity projection images of live HeLa ATAD1^{-/-} cells stably expressing EGFP-Gos28 (green channel) and transiently expressing empty vector (*top row*), ATAD1-HaloTag (*middle row*), and ATAD1(E193Q)-HaloTag (*bottom row*). Mitochondria are stained with MitoTracker (red channel). The individual channels are shown in black and white and overlay of the EGFP and the MitoTracker channels are shown in the right-most column with Hoechst-stained nuclei in blue. Insets are included to better show the absence or presence of co-localization between EGFP-Gos28 and the mitochondria. **(C)** Workflow of the CellProfiler pipeline for measuring EGFP-Gos28 mislocalization. **(D)** Mean Pearson correlation coefficient (PCC) values and the SEM between EGFP-Gos28 and the mitochondria when expressing the indicated construct. Individual cell PCC values are represented as a single dot. Significance values were calculated using the Mann-Whitney test. **** p<0.0001.

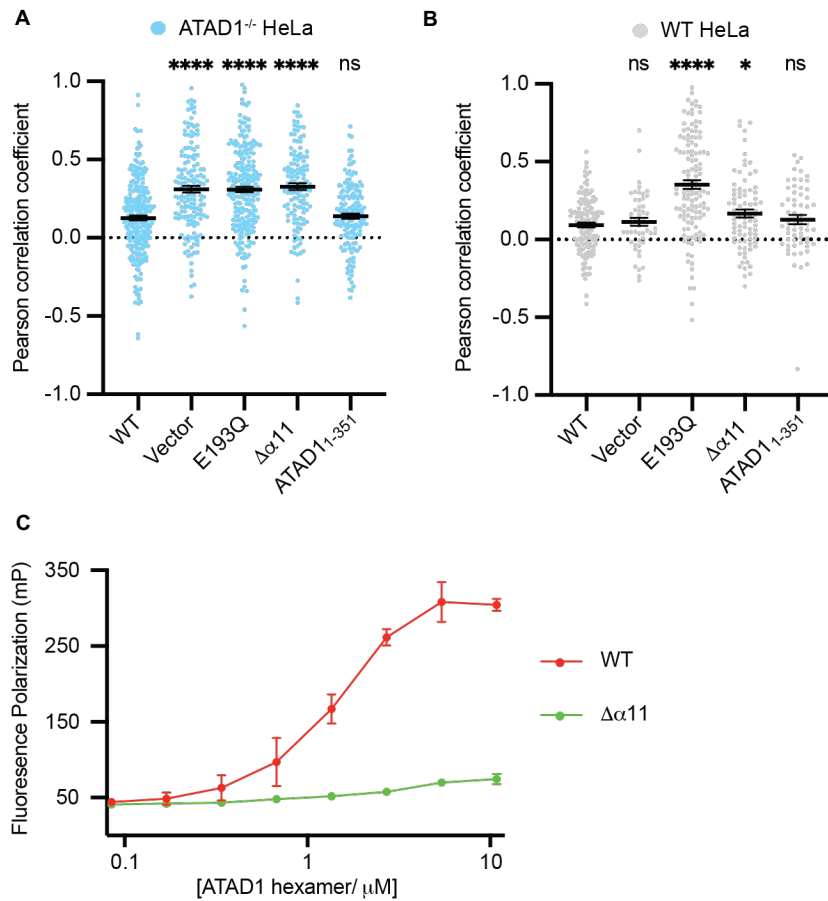


Figure 1.4 Helix α 11 mediates hexamer assembly

(A) Mean Pearson correlation coefficient (PCC) values and the SEM between EGFP-Gos28 and the mitochondria in ATAD1^{-/-} HeLa cells, transiently expressing controls or the ATAD1 Δ α 11 mutant. **(B)** Mean Pearson correlation coefficient (PCC) values and the SEM between EGFP-Gos28 and the mitochondria in WT HeLa cells, transiently expressing controls or the ATAD1 Δ α 11 mutant. Significance values were calculated using the Mann-Whitney test. * $p < 0.05$, **** $p < 0.0001$. **(C)** Fluorescence polarization assay showing the different peptide binding abilities of the Δ 40-ATAD1 (WT) versus Δ 40-ATAD1 Δ α 11 (Δ α 11). ATAD1 concentrations are expressed as hexamer concentrations. Each dot represents an average of two replicates and the error bar represents SEM.

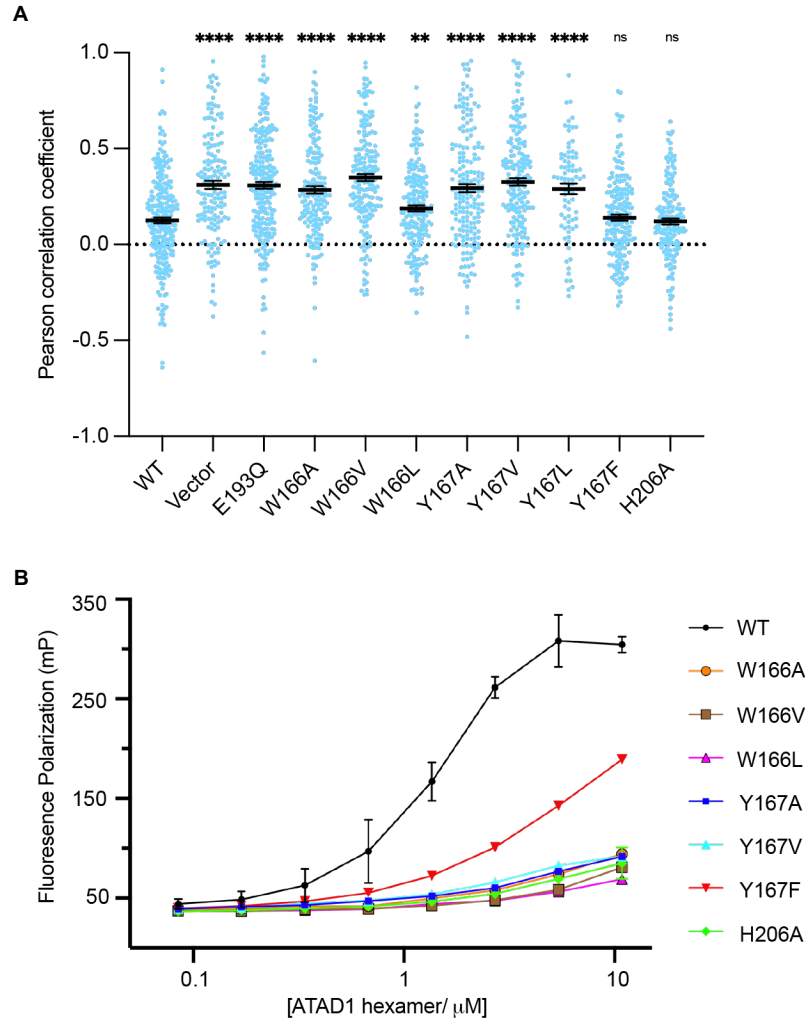


Figure 1.5 Pore-loop 1 aromatic amino acids are important for ATAD1's function both *in vivo* and *in vitro*

(A) Mean Pearson correlation coefficient (PCC) values and the SEM between EGFP-Gos28 and the mitochondria in cells expressing ATAD1 variants. Individual cell PCC values are represented as a single dot. Significance values were calculated using the Mann-Whitney test. ** $p < 0.01$, **** $p < 0.0001$. **(B)** Different ATAD1 variants' peptide binding abilities as measured by fluorescence polarization. ATAD1 concentrations are expressed as hexamer concentrations. Here, WT refers to $\Delta 40$ -ATAD1, and each mutant refers to $\Delta 40$ -ATAD1 bearing that mutation. Each dot represents the average of two replicates, and the error bar represents the SEM.

ATAD1_ <i>H. sapiens</i>	333	QQDLHRAIEKMKKSKDAAFQNVLTHVCLD-----	361
Spastin_ <i>H. sapiens</i>	582	LSDFTESLKKIKRSVSPOTLEAYIRWNKDFGDTTV	616
Katanin_ <i>C. elegans</i>	440	NIDFEAALQAVSPSAGPDTMLKCKEWCDSFGAM--	472
Vps4_ <i>S. cerevisiae</i>	403	IKDFLKAIKSTRPTVNEDDLLKQEQFTRDFGQEGN	437

Figure 1.1 – Figure supplement 1 Sequence alignment of AAA_{MC} proteins

Sequence alignment of AAA_{MC} proteins showing the secondary structure of the C-terminal helices. While ATAD1 has a longer $\alpha 11$, the other members of the AAA_{MC} subfamily have two shorter helices ($\alpha 11$ and $\alpha 12$) separated by a linker. Secondary structure assignment is based on experimentally determined structures: ATAD1 from this study; Spastin from PDB ID 6PEN; Vps4 from PDB ID: 6AP1 and Katanin from PDB ID: 6UGE.

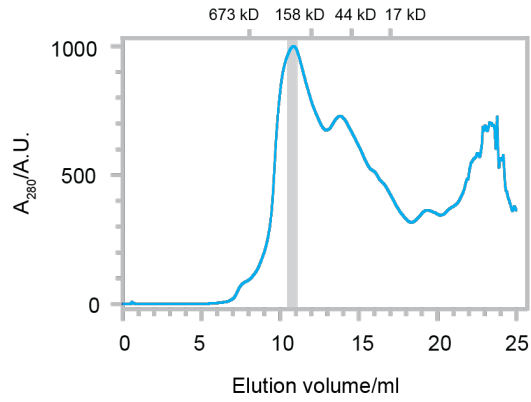


Figure 1.1 – Figure supplement 2 Size exclusion trace of $\Delta 40$ -ATAD1^{E193Q}

The SEC trace of $\Delta 40$ -ATAD1^{E193Q} show that ATAD1 forms oligomeric species on the SEC column. The grey bar indicated the fraction that was imaged by cryo-EM. Size markers are indicated on the top.

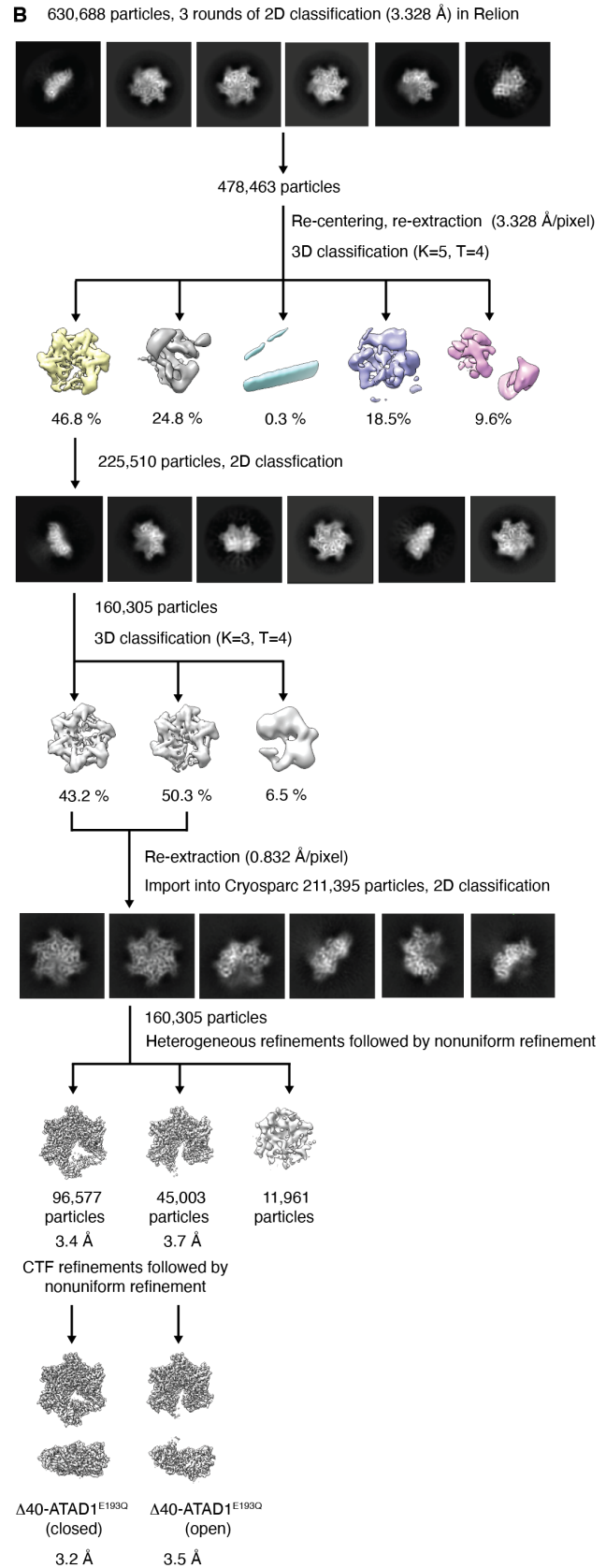
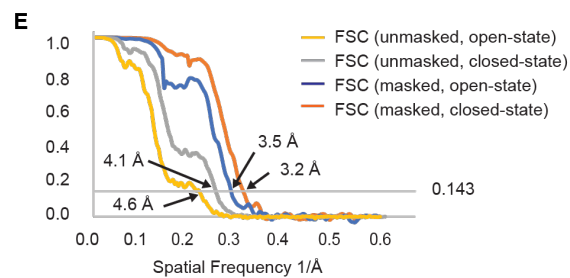
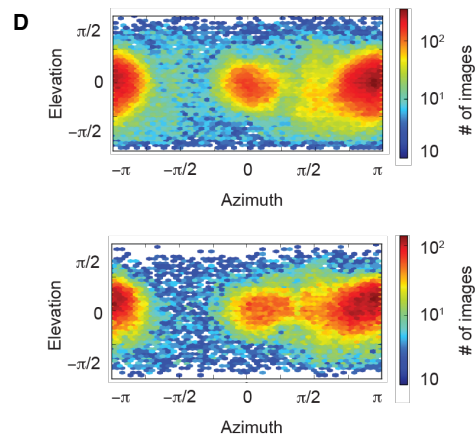
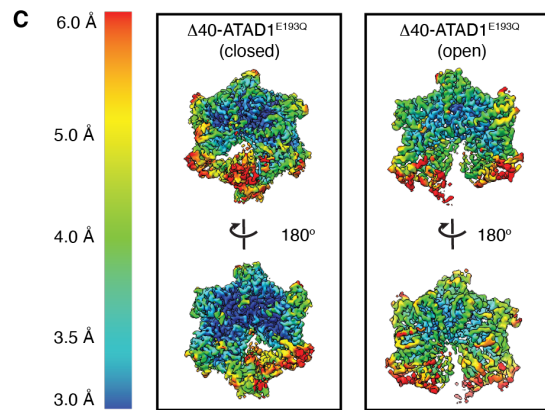
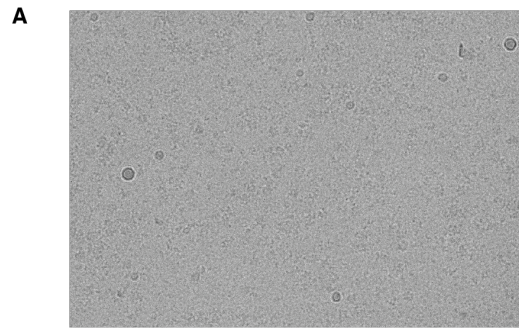


Figure 1.1 – Figure supplement 3 Cryo-EM processing of the $\Delta 40$ -ATAD1^{E193Q}-substrate complexes

(A) Representative micrograph showing the quality of data used for the final reconstruction of the $\Delta 40$ -ATAD1^{E193Q} structures. (B) Data processing scheme showing the 2D and 3D classification done using the RELION software. After two rounds of 3D classification, particles corresponding to ATAD1 are imported into Cryosparc for another round of 2D classification to further purify particles. Particles selected from the 2D classification were then subjected to heterogeneous refinement in which the open and the closed conformations emerged as two distinct classes. CTF refinement followed by nonuniform refinement generated the final reconstruction at 3.2 Å and 3.5 Å for the closed and the open conformations respectively. (C) Local resolution maps of the open and the closed conformations show that the core of the protein complex including the central pore and the nucleotide binding pockets are the best resolved regions. (E) Fourier Shell Correlation (FSC) plots of the 3D reconstructions of $\Delta 40$ -ATAD1^{E193Q}: $\Delta 40$ -ATAD1^{E193Q} (closed) masked (orange), $\Delta 40$ -ATAD1^{E193Q} (closed) unmasked (grey), $\Delta 40$ -ATAD1^{E193Q} (open) masked (blue), $\Delta 40$ -ATAD1^{E193Q} (open) unmasked (yellow).

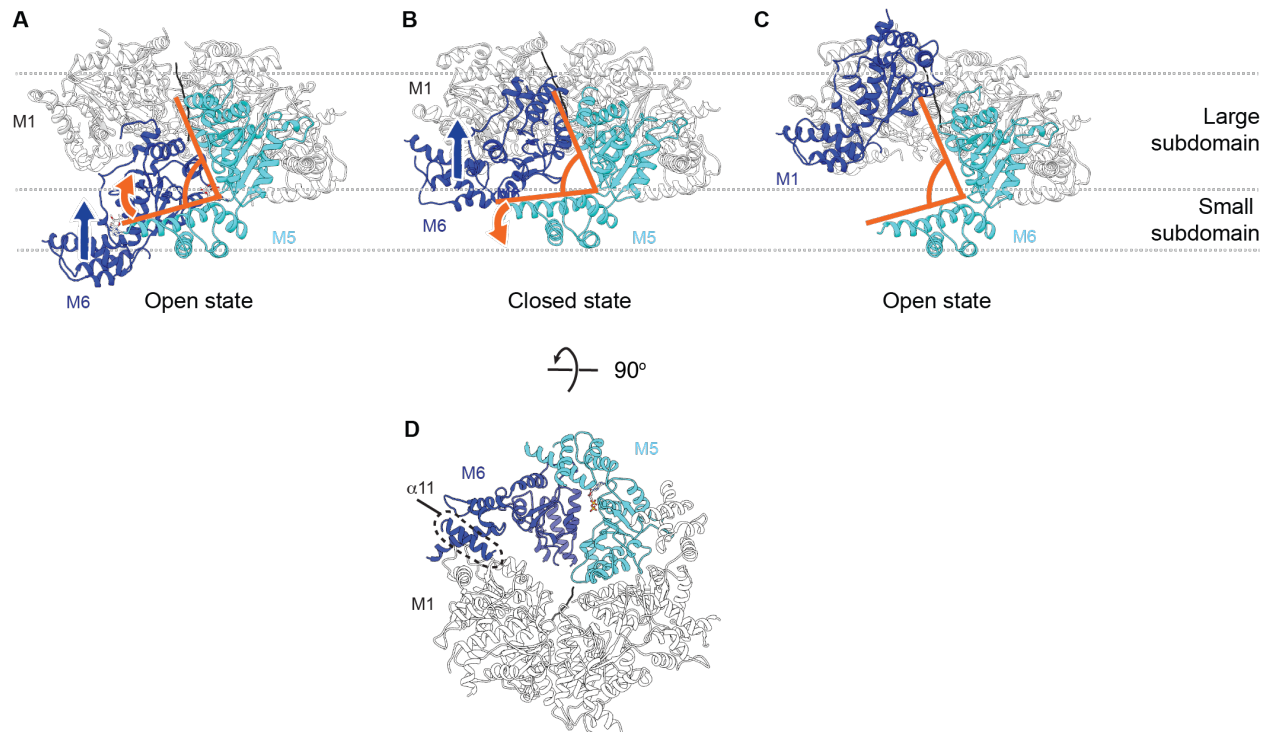


Figure 1.1 – Figure supplement 4 Hinge motion between the large and the small subdomain accompanies subunit movements

(A) In the open state, all six subunits have similar domain organization. M6 loses its bound nucleotide as it translocates from the bottom position to the closed position. The angle between the large and the small subdomain of M5 narrows (indicated by the orange lines). (B) M6 rebinds an ATP molecule as it continues to translocate upward to assume the top position in the open spiral (C). M5 loses contacts with M6 and the angle between its two subdomains widens. M6 is colored in dark blue, M5 in cyan, the rest of the subunits in white and the peptide substrate in black. The blue arrow indicates the direction of M6 translocation and the orange arrow indicates the direction of hinge movement between the two subdomains of M5. (D) 90 degrees flip from panel (B) showing the contact points between M6 and adjacent subunits. Multiple contact points exist between the M6 and the M5 subunits, whereas fewer contacts are made between the M6 and the M1 subunits. Among those, one is made by helix $\alpha 11$ (highlighted in the dashed circle).

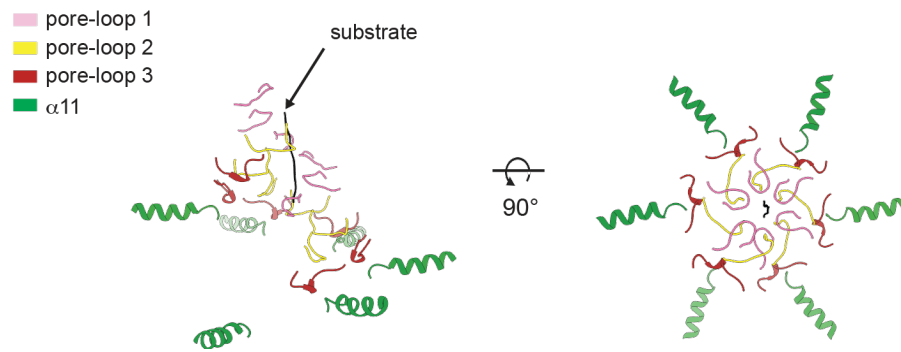


Figure 1.2 – Figure supplement 1 Helices $\alpha 11$ form an additional spiral staircase beneath pore-loop3

Cartoon representation of $\Delta 40$ -ATAD1^{E193Q} (open) showing the multiple staircases formed by three pore-loops and $\alpha 11$ that surround the substrate. Pore-loops 1, 2 and 3 are colored in pink, yellow and dark red respectively. Helix $\alpha 11$ is colored in green and the substrate peptide in black.

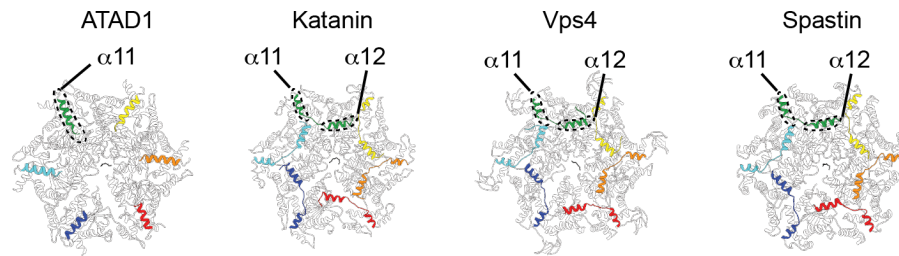


Figure 1.2 – Figure supplement 2 ATAD1 possesses a unique helix at the C-terminus that is structured differently from those in other AAA_{MC} proteins

Structures of different AAA_{MC} proteins shown in cartoon representation. The cytosol-facing side is shown. The C-terminal helices of each protein subunit ($\alpha 11$ for ATAD1, and $\alpha 11$ - $\alpha 12$ for the others) are colored as in Figure 1.1. The rest of the protein is colored in white. Whereas $\alpha 11$ of ATAD1 contacts the counterclockwise subunit, the $\alpha 12$ of Katanin (PDB ID: 6UGE), Vps4 (PDB ID: 6AP1) and Spastin (PDB ID: 6PEN) reach across the intersubunit interface and contacts the clockwise subunit.

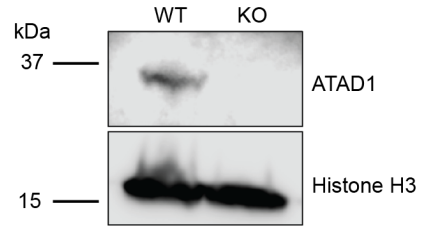


Figure 1.3 – Figure supplement 1 Verification of ATAD1 knockout by Western blot

Representative western on total protein extracts prepared from wild-type HeLa cells (left lane) and ATAD1^{-/-} HeLa cells (right lane). Knockout of ATAD1 was shown by the absence of ATAD1 signal (41 kDa). Histone H3 (17 kDa) served as the loading control.

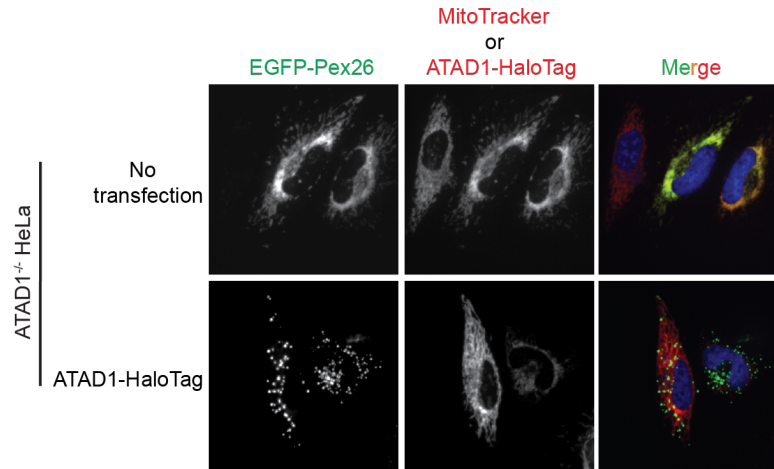


Figure 1.3 – Figure supplement 2 Live-cell imaging showing the ATAD1 dependent localization of EGFP-Pex26

Representative average intensity projection images of live HeLa ATAD1^{-/-} cells stably expressing EGFP-Pex26 (*top row*) and transiently expressing ATAD1-HaloTag (*bottom row*). Mitochondria are stained with MitoTracker red. The individual channels are shown in black and white and overlay of the EGFP and the MitoTracker red channels are shown in the right-most column with Hoechst-stained nuclei in blue.

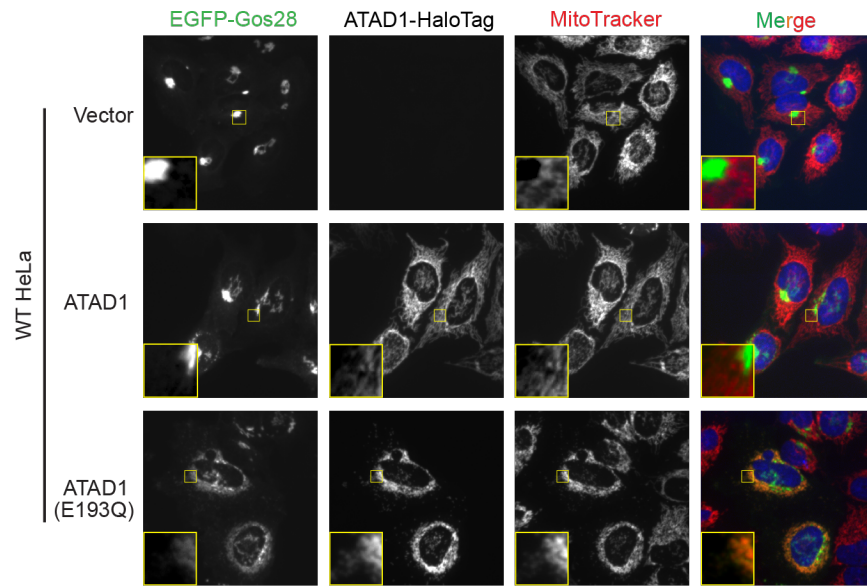


Figure 1.3 – Figure supplement 3 Live-cell imaging showing the ATAD1 dependent localization of EGFP-Gos28 in WT HeLa cells

Representative average intensity projection images of live WT HeLa cells stably expressing EGFP-Gos28 (green channel) and transiently expressing empty vector (*top row*), ATAD1-HaloTag (*middle row*), and ATAD1(E193Q)-HaloTag (*bottom row*). Mitochondria are stained with MitoTracker (red channel). The individual channels are shown in black and white and overlay of the EGFP and the MitoTracker channels are shown in the right-most column with Hoechst-stained nuclei in blue. Insets are included to better show the absence or presence of co-localization between EGFP-Gos28 and the mitochondria.

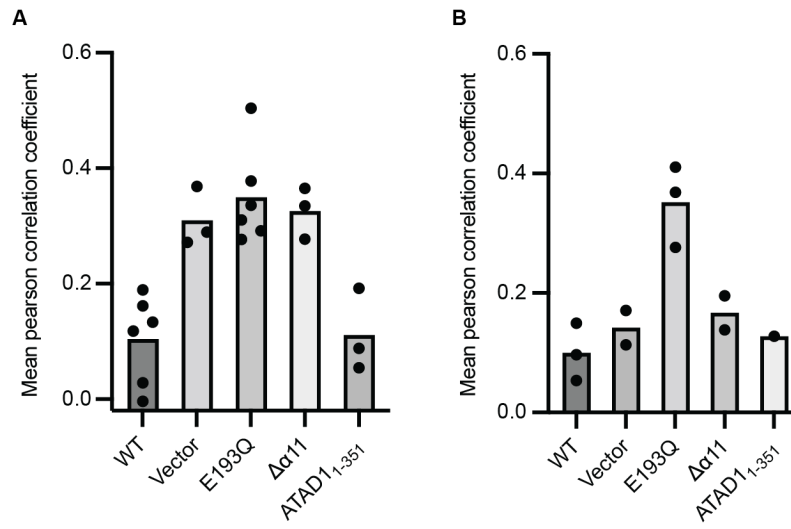


Figure 1.4 – Figure Supplement 1 Reproducibility of EGFP-Gos28 mislocalization in live-cell imaging with expression of $\Delta\alpha11$ mutant

Mean PCC values derived from the localization pipeline from each biological replicate were plotted for each mutant in **(A)** ATAD^{-/-} HeLa cells and **(B)** WT HeLa cells. Each dot represents a biological replicate with an average of 47 cells per replicate.

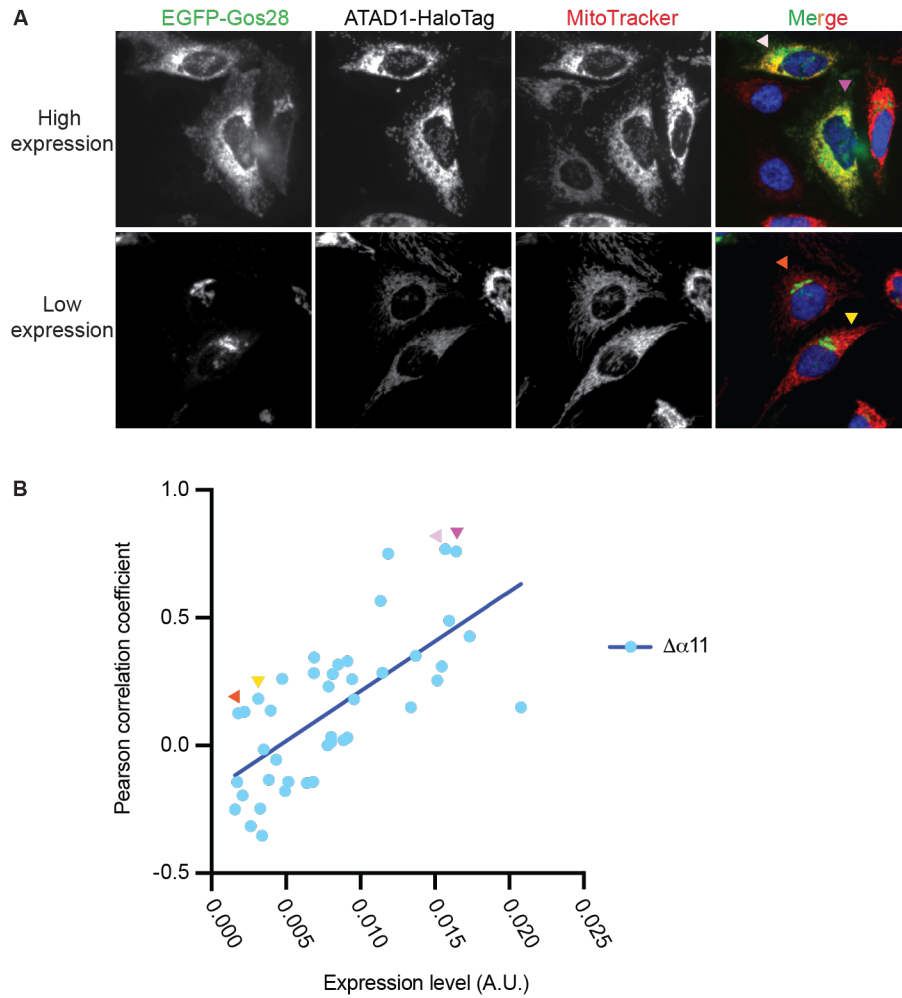


Figure 1.4 – Figure supplement 2 ATAD1 $\Delta\alpha$ 11 expression level impacts substrate mislocalization in WT HeLa cells.

(A) Representative average intensity projection images of live WT HeLa cells stably expressing EGFP-Gos28 and transiently expressing ATAD1 $\Delta\alpha$ 11-HaloTag. EGFP-Gos28 was only seen on the Golgi apparatus (*bottom row*). However, a small population of cells (*top row*) showed mislocalization of EGFP-Gos28 to the mitochondria. These cells also had a higher level of ATAD1 $\Delta\alpha$ 11-HaloTag expression (the PCC of cells shown in the images are indicated in panel B by arrows of matching colors). (B) The expression level of ATAD1 $\Delta\alpha$ 11 was plotted against the PCC for each cell. Linear regression analysis was preformed and there was a significant positive correlation between the two variables ($p < 0.0001$).

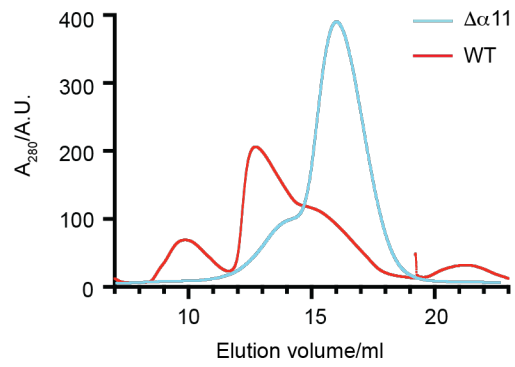


Figure 1.4 – Figure supplement 3 Size exclusion chromatography of $\Delta 40$ -ATAD1 and $\Delta 40$ -ATAD1 $\Delta\alpha 11$

The SEC traces of $\Delta 40$ -ATAD1 (WT) and $\Delta 40$ -ATAD1 $\Delta\alpha 11$ ($\Delta\alpha 11$) show that the former forms oligomeric species on the SEC column while the latter is predominantly a monomer. The SEC trace of $\Delta 40$ -ATAD1 is shown in a red line, and the SEC trace of $\Delta 40$ -ATAD1 $\Delta\alpha 11$ in a blue line.

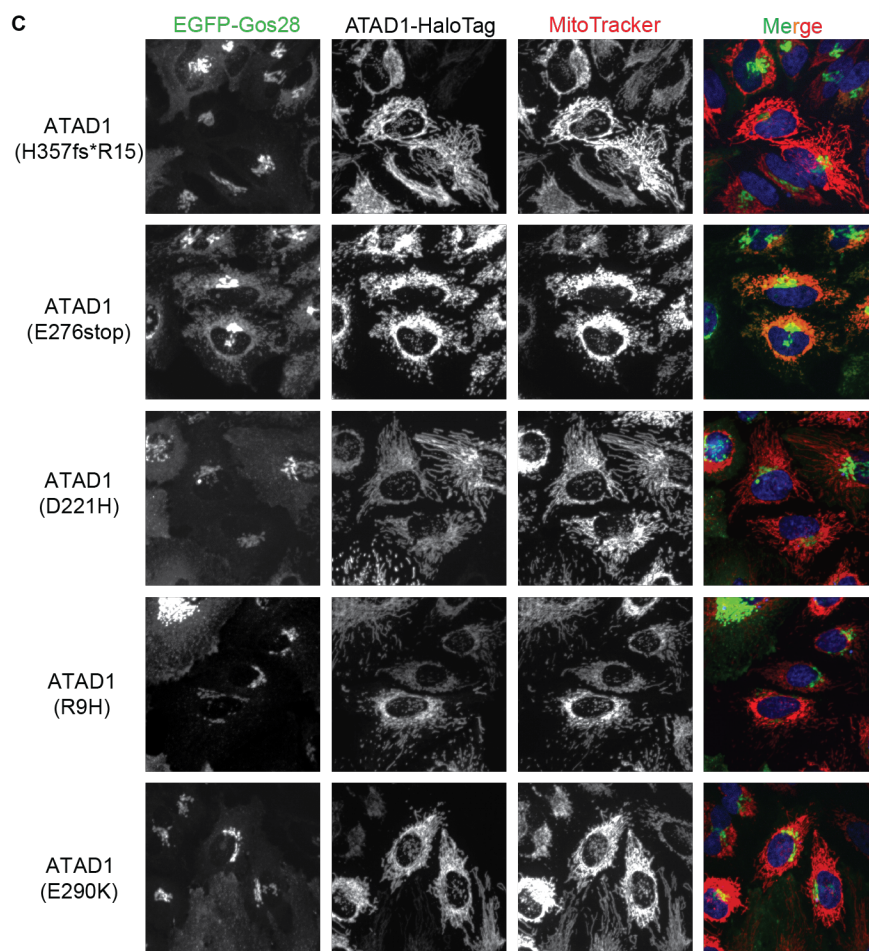
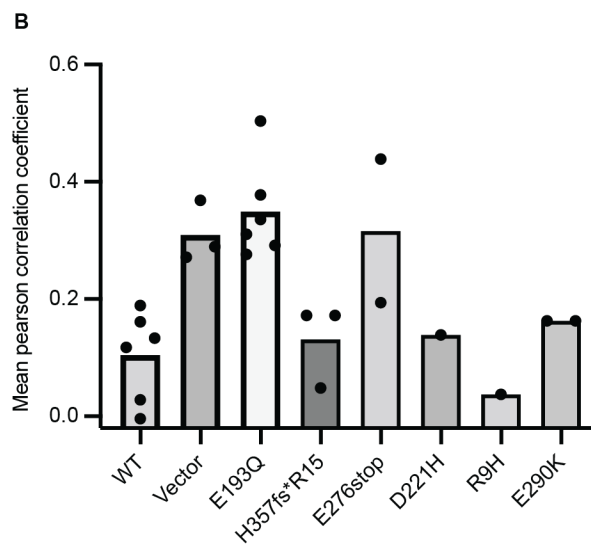
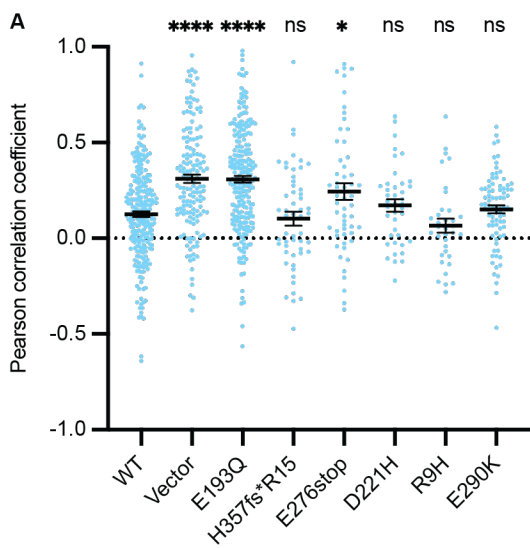


Figure 1.4 – Figure supplement 4 Testing the effect of the disease-relevant mutations of ATAD1

(A) Mean Pearson correlation coefficient (PCC) values and the SEM between EGFP-Gos28 and the mitochondria when expressing the ATAD1 bearing disease-relevant mutations. Individual cell PCC values are represented as a single dot. Significance values were calculated using the Mann-Whitney test. **** $p < 0.0001$, * $p < 0.05$. (B) Mean PCC values derived from the localization pipeline from each biological replicate were plotted for each mutant in ATAD^{-/-} HeLa cells. Each dot represents a biological replicate, with an average of 43 cells per replicate. (C) Representative average intensity projection images of live HeLa ATAD1^{-/-} cells stably expressing EGFP-Gos28 and transiently expressing ATAD1 mutants (as indicated on the far-left panel). The individual channels are shown in black and white and overlay of the EGFP and the MitoTracker channels are shown in the right-most column with Hoechst-stained nuclei in blue.

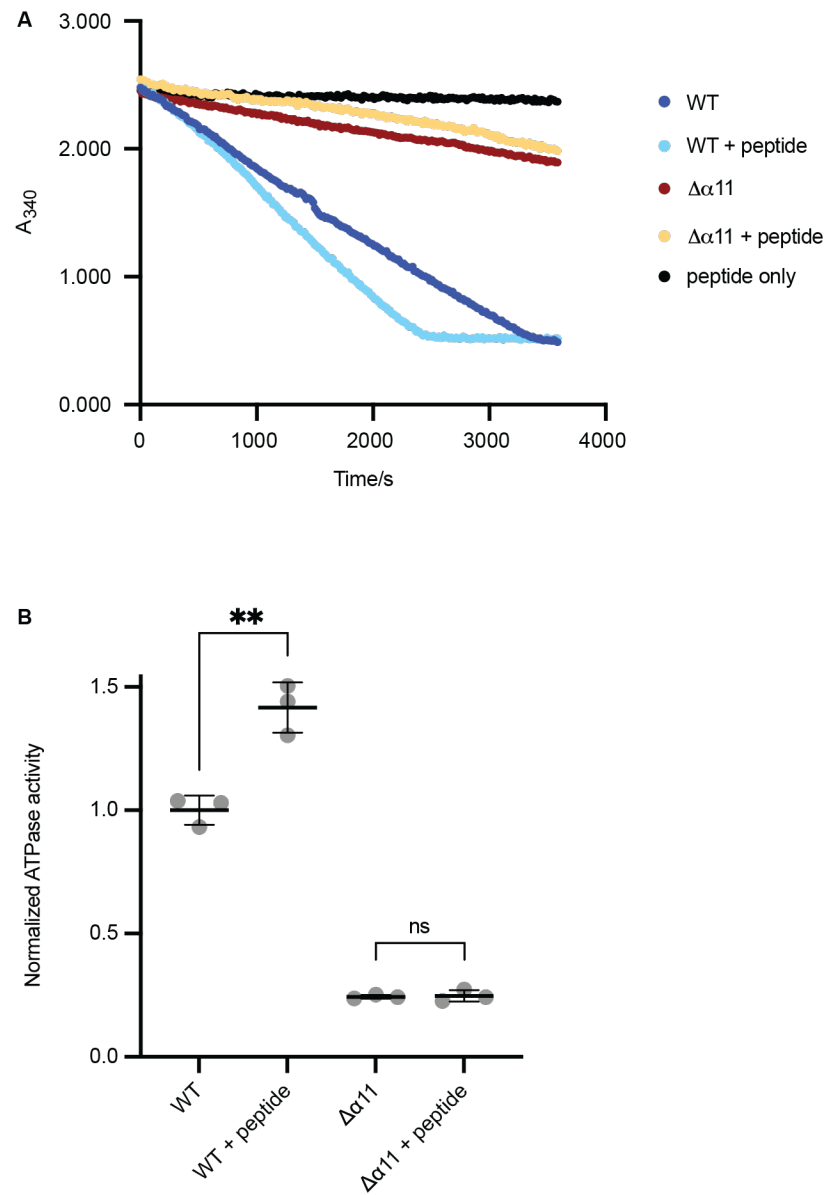


Figure 1.4 – Figure supplement 5 ATPase assay of $\Delta 40$ -ATAD1 and $\Delta 40$ -ATAD1 $\Delta \alpha 11$

(A) Representative ATPase assay showing that the ATPase activity of $\Delta 40$ -ATAD1 (WT) is stimulated by the substrate (P13) whereas $\Delta 40$ -ATAD1 $\Delta \alpha 11$ ($\Delta \alpha 11$) is not. (B) Quantification of panel (A). Each bar represents an average of three replicates and the error bar represents the SD. ** $p < 0.01$.

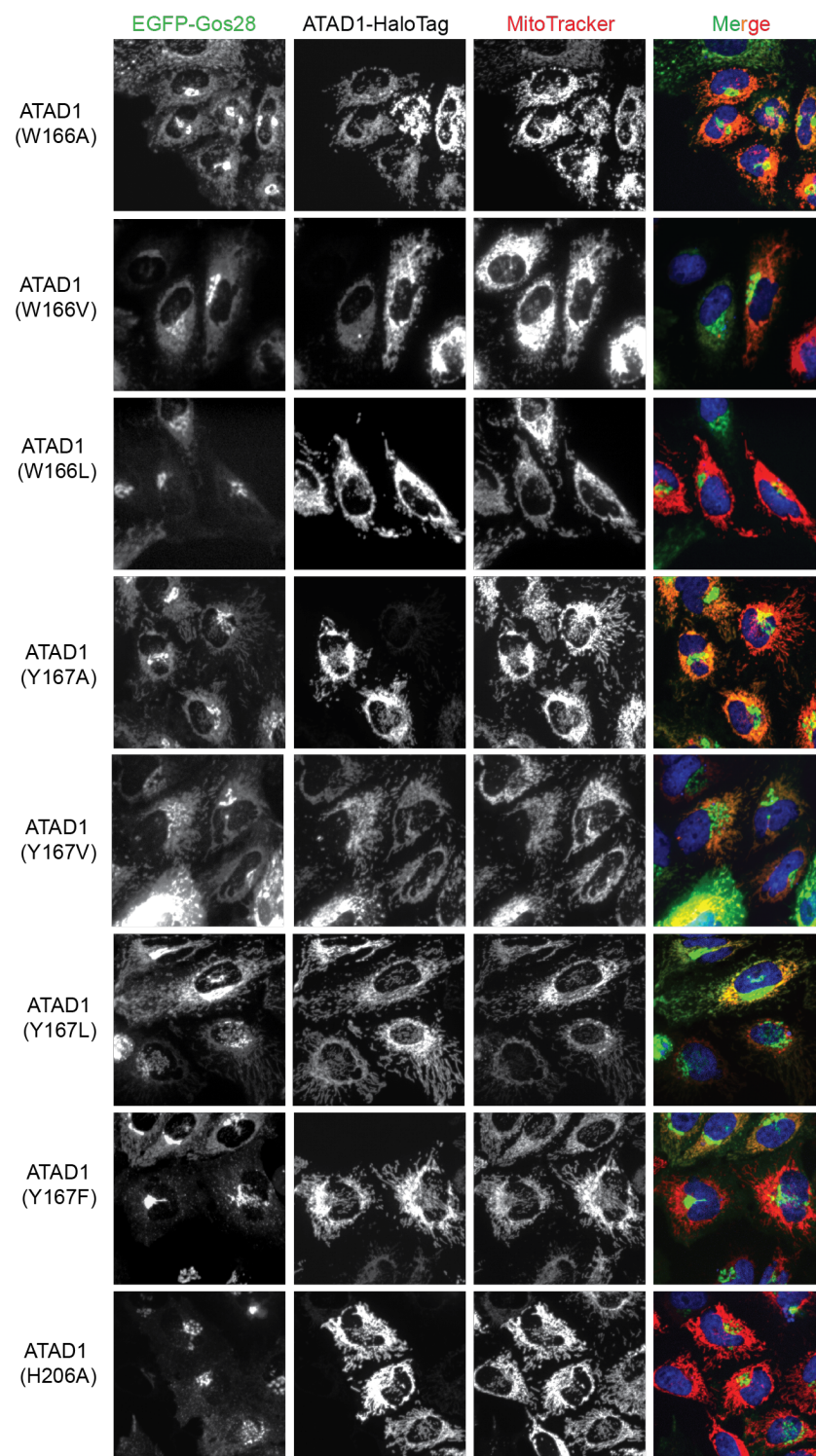


Figure 1.5- Figure supplement 1 Live-cell imaging showing the pore-loop dependent localization of EGFP-Gos28 in ATAD1^{-/-} HeLa cells

Representative average intensity projection images of live HeLa ATAD1^{-/-} cells stably expressing EGFP-Gos28 and transiently expressing ATAD1-HaloTag pore-loop mutants (as indicated on the far-left panel). The individual channels are shown in black and white and overlay of the EGFP and the MitoTracker channels are shown in the right-most column with Hoechst-stained nuclei in blue.

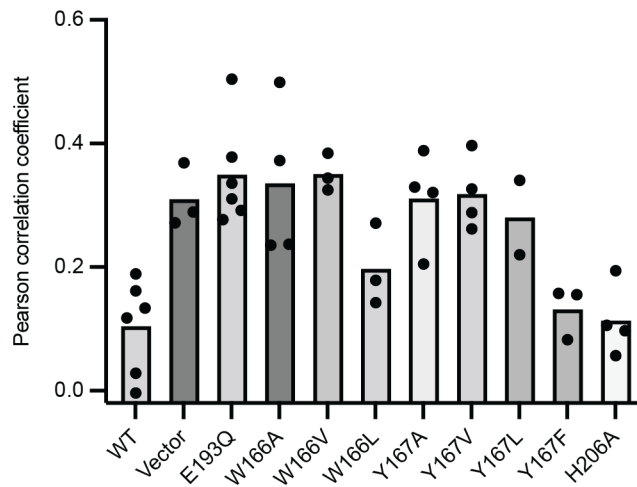


Figure 1.5- Figure supplement 2 Reproducibility of EGFP-Gos28 mislocalization in live-cell imaging with expression of pore loop mutants

Mean PCC values derived from the localization pipeline from each biological replicate were plotted for each pore loop mutant in *ATAD^{-/-}* HeLa cells. Each dot represents a biological replicate, with an average of 50 cells per replicate.

MATERIALS AND METHODS

Molecular cloning

To generate the construct used for cryo-EM studies, the gene encoding the cytosolic domain of human $\Delta 40$ -ATAD1 was PCR amplified and subcloned into a pET28 vector encoding an N-terminal 6XHis tag followed by a thrombin cleavage site. The Walker B mutation (E214Q) was introduced by QuickChange site-directed mutagenesis. To generate the constructs used for imaging, the gene blocks for the human *ATAD1* and *Gos28* were synthesized by Twist Bioscience. *Gos28* was C-terminally fused to *EGFP* and cloned downstream of a CMV promoter within a lentivirus production vector. *ATAD1* was N-terminally fused to *HaloTag* and cloned downstream of a truncated CMV promoter (CMVd3) for transient transfections. Mutations to *ATAD1* were made by QuickChange site-directed mutagenesis. All the constructs are verified by Sanger sequencing.

Protein purification

His- $\Delta 40$ -ATAD1^{E193Q} was expressed and purified as previously described for His- $\Delta 30$ -Msp1^{E214Q} (Wang et al., 2020).

Sample preparation of cryo electron microscopy

His- $\Delta 40$ -ATAD1^{E193Q} was diluted to around 100 μ M in buffer containing 25 mM HEPES pH 7.5, 300 mM NaCl, 1 mM DTT, 2.5% glycerol, 2 mM ATP and 2 mM MgCl₂. The sample was incubated on ice for 1 - 2 h before plunge freezing. A 3 μ l aliquot of the sample were applied onto the Quantifoil R 1.2/1/3 400 mesh Gold grid and incubated for 15 s. A 0.5 μ l aliquot of 0.1-0.2% Nonidet P-40 substitutes was added immediately before blotting. The entire blotting procedure was performed using Vitrobot Mark IV (FEI) at 10° C and 100% humidity.

Electron microscopy data collection

Cryo-EM data was collected on a Titan Krios transmission electron microscope operating at 300 keV and micrographs were acquired using a Gatan K3 summit direct electron detector. The total dose was 60 e⁻/Å², fractioned over 100 frames during a 10 s exposure. Data was collected at 105,000 x nominal magnification (0.832 Å/pixel at the specimen level) and nominal defocus range of -1.0 to -2.0 μm.

Cryo-EM data analysis

Micrograph frames were aligned using MotionCorr2. The contrast transfer function (CTF) parameters were estimated with GCTF (Zhang, 2015). Particles were automatically picked using Gautomatch and extracted in RELION (Scheres, 2012) using a 320-pixel box size. Images were down-sampled to a pixel size of 3.328 Å and classified in 2D in RELION. Classes that showed clear protein features were selected and extracted with re-centering and then subjected to 3D classification. Particles from the best class emerging from 3D classification were then subjected to 2D classification followed by another round of 3D classification to further purify the particles. Particles that showed clear hexameric features were then re-extracted (pixel size = 0.832 Å /pixel) and imported into cryoSPARC (Punjani et al., 2017). Within cryoSPARC, particles were subjected to another round of 2D classification followed by heterogenous refinement, from which two distinct conformations (the open versus the closed conformations) were discovered. CTF refinement followed by nonuniform refinement was performed on both conformations to yield final reconstructions of at 3.2 Å and 3.5 Å, for the closed and the open conformations.

Atomic model building and refinement

Model building and refinement was done in a similar way as previously described (Wang et al., 2020). Briefly, the big and the small AAA domain of the crystal structure of the monomeric *S. cerevisiae* Msp1 (Wohlever et al., 2017) was used to generate the predicted structures of the human ATAD1 in SWISS-MODEL (Waterhouse et al., 2018). The six big AAA domains and the six small AAA domains were individually docked into the map of His-Δ40-ATAD1^{E193Q} in *UCSF Chimera* (Pettersen et al., 2004) using

the *Fit in Map* function. The resulting model was subjected to rigid body refinement in Phenix (Adams et al., 2010), followed by real space refinement in *Coot* (Emsley et al., 2010). After the protein part has been modeled, a piece of continuous density was left in the central pore, into which we modeled a polyalanine sequence. Significant density was visible in the nucleotide-binding pockets within subunits M1 through M5, and an ATP molecule was modeled into that density. For the M6 subunit, in the open conformation the density there was not clear enough to distinguish between ATP and ADP, so although an ADP molecule was modeled, we indicated in figure 1 that it could be either. In the closed conformation, no significant density was observed in the nucleotide binding pocket. The figures displaying structures were prepared with *UCSF Chimera*.

Fluorescence Polarization

All fluorescence polarization experiments were done in the FP assay buffer containing 25 mM HEPES (pH 7.5), 150 mM KCl, 2 mM MgCl₂, 1 mM DTT and 2 mM ATP γ S and measured in 384-well non-stick black plates using ClarioStar PLUS (BMG LabTech) at room temperature. Prior to the reaction set up, ATAD1 was diluted in 2-fold dilution series and incubated with 100 nM fluorescently-labeled peptide (P13: 5-FAM-FSRLYQLRIR, purchased from Genscript) for 20 minutes at room temperature. Then, the mixture was subjected to measurement of parallel and perpendicular intensities (excitation: 482 nm, emission: 530 nm). Data was plotted using GraphPad Prism 8.

ATPase assay

ATPase activity of ATAD1 was measured using the oxidation of NADH as a readout of ATP hydrolysis. An enzyme mixture containing 0.2 mM NADH, 1 mM phosphoenol-pyruvate (PEP), 50 U/mL of pyruvate kinase and lactate dehydrogenase was added to the wells of a 384-well black plate (Corning). 10 μ M ATAD1, assay buffer (25 mM HEPES pH 7.5, 100 mM KCl, 2 mM MgCl₂, 10 mM BSA, 0.05% Tween-20) and either 50 μ M of the unlabeled version of P13 peptide (FSRLYQLRIR) or a blank was

added to the enzyme mixture. Samples were incubated for 20 minutes at 37°C before 1 mM ATP was added to start the reactions. Absorbance at 340 nm was measured every 15 seconds for a total of 60 minutes using the CLARIOstar Plus (BMG LabTech) microplate reader. Data was plotted using GraphPad Prism 8.

Cell culture and transduction

HeLa cells were cultured in DMEM supplemented with 10% FBS, 100 U/ml penicillin/streptomycin, and 6 mM L-glutamine. A pooled ATAD1^{-/-} cell population was generated by Synthego with a guide RNA targeting Exon 5 (CGGUCAGUGUCGAAGGCUGA). Monoclonal populations were obtained using limiting dilution in a 96-well plate and expanding single cell populations into a 6-well plates. Knockouts were confirmed by Sanger sequencing and Western blotting using anti-ATAD1 antibody (N125/10, NeuroMab) and anti-Histone H3 (ab1791, Abcam) as a loading control.

WT HeLa and ATAD1^{-/-} HeLa cell lines expressing the reporter EGFP-Gos28 were generated by lentiviral infection. In brief, Vesicular Stomatitis Virus (VSV)-G pseudotyped lentiviral particles were produced in 293METR packaging cells (kind gift of Brian Ravinovich, formerly at MD Anderson Cancer Center, Camden, NJ) using standard protocols. WT and ATAD1^{-/-} HeLa cells were infected with concentrated virus (supplemented with 8 µg/mL polybrene) by centrifugal inoculation at 2,000 rpm for 2 hrs. Viral supernatant was removed following overnight incubation and cells were expanded for FACS sorting. EGFP positive cells were sorted using SONY SH800 FACS into high and low EGFP-expressing populations. For all imaging experiments, the population that has high EGFP expression was used.

EGFP-Gos28 imaging

Expression of the ATAD1 variants in the cell-based assay was done through transient transfections. The day before transfections, cells were seeded on Ibidi 8-well glass bottom µ-slides in FluoroBrite DMEM

(Life Technologies) media supplemented with glutamine and 10% FBS. Plasmids with the ATAD1 variants under a truncated CMVd3 promoter were transfected into cells using the FuGENE HD transfection reagent (Promega), following the manufacturer's protocol. Cells were incubated for 48 hours before imaging. For visualizing cell structures, nuclei were stained with Hoechst 33342 and mitochondria were stained with MitoTracker Deep Red FM (ThermoFisher). Transfected cells were stained with 25 nM Janelia Fluor 549 dye conjugated with the HaloTag ligand (JF549-HaloTag; kind gift of Dr. Luke Lavis). Cells were incubated with the dyes for 15 minutes at 37°C followed by 3 washes with FluoroBrite DMEM media.

Microscopy

Confocal imaging was carried out on a Nikon Ti-E inverted microscope equipped with a Yokogawa CSU-X high-speed confocal scanner unit and an Andor iXon 512 × 512 EMCCD camera. All images were acquired through a 40× 1.3 NA oil immersion objective. Images were typically acquired with 60× EM gain and 100-ms exposure. The four lasers used were 405 nm (operated at 10 mW), 488 nm (operated at 25 mW), 561 nm (operated at 25 mW), and 640 nm (operated at 15 mW). All components of the microscope were controlled by the μ Manager open-source platform (Edelstein et al., 2010). The microscope stage was enclosed in a custom-built incubator that maintained preset temperature and CO₂ levels for prolonged live-imaging experiments. To avoid unintentional selection bias, fields-of-view were selected by only looking at stained cell nuclei in the 405-nm channel. No cells or fields of view were subsequently excluded from analysis, ensuring that the data faithfully capture the distribution of fluorescence across the entire cell population.

Image quantification with CellProfiler

For each experiment, 15 fields-of-view were imaged per condition with an average number around 50 cells total. Average intensity z-projections were made for each image in each of the four channels and

used as the input for the CellProfiler pipeline (McQuin et al., 2018). For automated image analysis, we developed a pipeline in CellProfiler 4.0.5. First, the images were background-corrected in every channel. Then, nuclei were identified as starting points for the propagation of cell body masks. Cells in which integrated ATAD1 signal (labeled with the JF549-HaloTag dye) passed a manually chosen threshold were identified as ATAD-positive cells. Within those cells, the MitoTracker signal was used to create a pseudo-cell boundary mask, closing all gaps. This serves as a reasonably good proxy for total cell area because the mitochondrial network is broadly distributed throughout HeLa cells. Measurements for colocalization were made inside this mask and outside of the nucleus. Pearson correlation coefficients (PCCs) were calculated for each individual cell by performing a pixel-wise comparison of the EGFP and MitoTracker channels.

REFERENCES

- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biological Crystallogr* 66:213–21. doi:10.1107/s0907444909052925
- Ahrens-Nicklas RC, Umanah GKE, Sondheimer N, Deardorff MA, Wilkens AB, Conlin LK, Santani AB, Nesbitt A, Juulsola J, Ma E, Dawson TM, Dawson VL, Marsh ED. 2017. Precision therapy for a new disorder of AMPA receptor recycling due to mutations in ATAD1. *Neurology Genetics* 3:1–9. doi:10.1212/nxg.0000000000000130
- Basch M, Wagner M, Rolland S, Carbonell A, Zeng R, Khosravi S, Schmidt A, Aftab W, Imhof A, Wagener J, Conradt B, Wagener N. 2020. Msp1 cooperates with the proteasome for extraction of arrested mitochondrial import intermediates. *Mol Biol Cell* 31:753–767. doi:10.1091/mbc.e19-06-0329
- Chen Y, Umanah GKE, Dephoure N, Andrabi SA, Gygi SP, Dawson TM, Dawson VL, Rutter J. 2014. Msp1/ATAD1 maintains mitochondrial function by facilitating the degradation of mislocalized tail-anchored proteins. *Embo J* 33:1548–1564. doi:10.15252/emj.201487943
- Chio US, Cho H, Shan S-O. 2017. Mechanisms of Tail-Anchored Membrane Protein Targeting and Insertion. *Annu Rev Cell Dev Bi* 33:417–438. doi:10.1146/annurev-cellbio-100616-060839
- Cooney I, Han H, Stewart MG, Carson RH, Hansen DT, Iwasa JH, Price JC, Hill CP, Shen PS. 2019. Structure of the Cdc48 segregase in the act of unfolding an authentic substrate. *Science* 365:502–505. doi:10.1126/science.aax0486

- Dederer V, Khmelinskii A, Huhn AG, Okreglak V, Knop M, Lemberg MK. 2019. Cooperation of mitochondrial and ER factors in quality control of tail-anchored proteins. *Elife* 8:e45506. doi:10.7554/elife.45506
- Doan KN, Grevel A, Mårtensson CU, Ellenrieder L, Thornton N, Wenz L-S, Opaliński Ł, Guiard B, Pfanner N, Becker T. 2020. The Mitochondrial Import Complex MIM Functions as Main Translocase for α -Helical Outer Membrane Proteins. *Cell Reports* 31:107567. doi:10.1016/j.celrep.2020.107567
- Edelstein A, Amodaj N, Hoover K, Vale R, Stuurman N. 2010. Computer Control of Microscopes Using μ Manager. *Curr Protoc Mol Biology* 92:14.20.1-14.20.17. doi:10.1002/0471142727.mb1420s92
- Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr Sect D Biological Crystallogr* 66:486–501. doi:10.1107/s0907444910007493
- Favaloro V, Spasic M, Schwappach B, Dobberstein B. 2008. Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *J Cell Sci* 121:1832–1840. doi:10.1242/jcs.020321
- Fresenius HL, Wohlever ML. 2019. Sorting out how Msp1 maintains mitochondrial membrane proteostasis. *Mitochondrion* 49:128–134. doi:10.1016/j.mito.2019.07.011
- Frickey T, Lupas AN. 2004. Phylogenetic analysis of AAA proteins. *J Struct Biol* 146:2–10. doi:10.1016/j.jsb.2003.11.020
- Gates SN, Yokom AL, Lin J, Jackrel ME, Rizo AN, Kendsersky NM, Buell CE, Sweeny EA, Mack KL, Chuang E, Torrente MP, Su M, Shorter J, Southworth DR. 2017. Ratchet-like polypeptide translocation mechanism of the AAA⁺ disaggregase Hsp104. *Science* 357:273–279. doi:10.1126/science.aan1052

- Han H, Monroe N, Sundquist WI, Shen PS, Hill CP. 2017. The AAA ATPase Vps4 binds ESCRT-III substrates through a repeating array of dipeptide-binding pockets. *Elife* 6:e31324. doi:10.7554/elifesciences.31324
- Han H, Schubert HL, McCullough J, Monroe N, Purdy MD, Yeager M, Sundquist WI, Hill CP. 2019. Structure of spastin bound to a glutamate-rich peptide implies a hand-over-hand mechanism of substrate translocation. *J Biol Chem* 295:435–443. doi:10.1074/jbc.119.009890
- Matsumoto S, Nakatsukasa K, Kakuta C, Tamura Y, Esaki M, Endo T. 2019. Msp1 Clears Mistargeted Proteins by Facilitating Their Transfer from Mitochondria to the ER. *Mol Cell* 76:191-205.e10. doi:10.1016/j.molcel.2019.07.006
- McQuin C, Goodman A, Chernyshev V, Kametsky L, Cimini BA, Karhohs KW, Doan M, Ding L, Rafelski SM, Thirstrup D, Wiegand W, Singh S, Becker T, Caicedo JC, Carpenter AE. 2018. CellProfiler 3.0: Next-generation image processing for biology. *Plos Biol* 16:e2005970. doi:10.1371/journal.pbio.2005970
- Monroe N, Han H, Shen PS, Sundquist WI, Hill CP. 2017. Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. *Elife* 6:e24487. doi:10.7554/elifesciences.24487
- Nuebel E, Morgan JT, Fogarty S, Winter JM, Lettlova S, Berg JA, Chen Y, Kidwell CU, Maschek JA, Clowers KJ, Argyriou C, Chen L, Wittig I, Cox JE, Roh-Johnson M, Braverman N, Bonkowsky J, Gygi SP, Rutter J. 2021. The biochemical basis of mitochondrial dysfunction in Zellweger Spectrum Disorder. *Embo Rep* e51991. doi:10.15252/embr.202051991
- Okreglak V, Walter P. 2014. The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins. *Proc National Acad Sci* 111:8019–8024. doi:10.1073/pnas.1405755111

- Peña AH de la, Goodall EA, Gates SN, Lander GC, Martin A. 2018. Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-driven translocation. *Sci New York NY* 362:eaav0725. doi:10.1126/science.aav0725
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera?A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612. doi:10.1002/jcc.20084
- Pfanner N, Warscheid B, Wiedemann N. 2019. Mitochondrial proteins: from biogenesis to functional networks. *Nat Rev Mol Cell Bio* 20:267–284. doi:10.1038/s41580-018-0092-0
- Piard J, Umanah GKE, Harms FL, Abalde-Atristain L, Amram D, Chang M, Chen R, Alawi M, Salpietro V, Rees MI, Chung S-K, Houlden H, Verloes A, Dawson TM, Dawson VL, Maldergem LV, Kutsche K. 2018. A homozygous ATAD1 mutation impairs postsynaptic AMPA receptor trafficking and causes a lethal encephalopathy. *Brain* 141:651–661. doi:10.1093/brain/awx377
- Puchades C, Rampello AJ, Shin M, Giuliano CJ, Wiseman RL, Glynn SE, Lander GC. 2017. Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. *Science* 358:eaao0464. doi:10.1126/science.aao0464
- Puchades C, Sandate CR, Lander GC. 2020. The molecular principles governing the activity and functional diversity of AAA+ proteins. *Nat Rev Mol Cell Bio* 21:43–58. doi:10.1038/s41580-019-0183-6
- Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14:290–296. doi:10.1038/nmeth.4169

- Rodriguez-Aliaga P, Ramirez L, Kim F, Bustamante C, Martin A. 2016. Substrate-translocating loops regulate mechanochemical coupling and power production in AAA+ protease ClpXP. *Nat Struct Mol Biol* 23:974–981. doi:10.1038/nsmb.3298
- Sandate CR, Szyk A, Zehr EA, Lander GC, Roll-Mecak A. 2019. An allosteric network in spastin couples multiple activities required for microtubule severing. *Nat Struct Mol Biol* 26:671–678. doi:10.1038/s41594-019-0257-3
- Scheres SHW. 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180:519–30. doi:10.1016/j.jsb.2012.09.006
- Schuldiner M, Metz J, Schmid V, Denic V, Rakwalska M, Schmitt HD, Schwappach B, Weissman JS. 2008. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134:634–45. doi:10.1016/j.cell.2008.06.025
- Shin SC, Im S-K, Jang E-H, Jin KS, Hur E-M, Kim EE. 2019. Structural and Molecular Basis for Katanin-Mediated Severing of Glutamylated Microtubules. *Cell Reports* 26:1357-1367.e5. doi:10.1016/j.celrep.2019.01.020
- Stefanovic S, Hegde RS. 2007. Identification of a Targeting Factor for Posttranslational Membrane Protein Insertion into the ER. *Cell* 128:1147–1159. doi:10.1016/j.cell.2007.01.036
- Su M, Guo EZ, Ding X, Li Y, Tarrasch JT, Brooks CL, Xu Z, Skiniotis G. 2017. Mechanism of Vps4 hexamer function revealed by cryo-EM. *Sci Adv* 3:e1700325. doi:10.1126/sciadv.1700325
- Sun S, Li L, Yang F, Wang X, Fan F, Yang M, Chen C, Li X, Wang H-W, Sui S-F. 2017. Cryo-EM structures of the ATP-bound Vps4(E233Q) hexamer and its complex with Vta1 at near-atomic resolution. *Nat Commun* 8:16064. doi:10.1038/ncomms16064

- Twomey EC, Ji Z, Wales TE, Bodnar NO, Ficarro SB, Marto JA, Engen JR, Rapoport TA. 2019. Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. *Science* 365:eaax1033. doi:10.1126/science.aax1033
- Umanah GKE, Pignatelli M, Yin X, Chen R, Crawford J, Neifert S, Scarffe L, Behensky AA, Guiberson N, Chang M, Ma E, Kim JW, Castro CC, Mao X, Chen L, Andrabi SA, Pletnikov MV, Pulver AE, Avramopoulos D, Bonci A, Valle D, Dawson TM, Dawson VL. 2017. Thorase variants are associated with defects in glutamatergic neurotransmission that can be rescued by Perampanel. *Sci Transl Med* 9:eaah4985. doi:10.1126/scitranslmed.aah4985
- Vajjhala PR, Nguyen CH, Landsberg MJ, Kistler C, Gan A, King GF, Hankamer B, Munn AL. 2008. The Vps4 C-terminal helix is a critical determinant for assembly and ATPase activity and has elements conserved in other members of the meiotic clade of AAA ATPases. *Febs J* 275:1427–1449. doi:10.1111/j.1742-4658.2008.06300.x
- Wang L, Myasnikov A, Pan X, Walter P. 2020. Structure of the AAA protein Msp1 reveals mechanism of mislocalized membrane protein extraction. *Elife* 9:e54031. doi:10.7554/elife.54031
- Wang L, Walter P. 2020. Msp1/ATAD1 in Protein Quality Control and Regulation of Synaptic Activities. *Annu Rev Cell Dev Bi* 36:1–24. doi:10.1146/annurev-cellbio-031220-015840
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46:gky427-. doi:10.1093/nar/gky427
- Weidberg H, Amon A. 2018. MitoCPR—A surveillance pathway that protects mitochondria in response to protein import stress. *Science* 360:aan4146. doi:10.1126/science.aan4146

- Winter JM, Fresenius HL, Keys HR, Cunningham CN, Ryan J, Sirohi D, Berg JA, Tripp SR, Barta P, Agarwal N, Letai A, Sabatini DM, Wohlever ML, Rutter J. n.d. Co-deletion of ATAD1 with PTEN primes cells for BIM-mediated apoptosis. doi:10.1101/2021.07.01.450781
- Wohlever ML, Mateja A, McGilvray PT, Day KJ, Keenan RJ. 2017. Msp1 Is a Membrane Protein Dislocase for Tail-Anchored Proteins. *Mol Cell* 67:194-202.e6. doi:10.1016/j.molcel.2017.06.019
- Zehr E, Szyk A, Piszczek G, Szczesna E, Zuo X, Roll-Mecak A. 2017. Katanin spiral and ring structures shed light on power stroke for microtubule severing. *Nat Struct Mol Biol* 24:717–725. doi:10.1038/nsmb.3448
- Zhang J, Wang Y, Chi Z, Keuss MJ, Pai Y-ME, Kang HC, Shin J, Bugayenko A, Wang H, Xiong Y, Pletnikov MV, Mattson MP, Dawson TM, Dawson VL. 2011. The AAA+ ATPase Thorase Regulates AMPA Receptor-Dependent Synaptic Plasticity and Behavior. *Cell* 145:284–299. doi:10.1016/j.cell.2011.03.016
- Zhang K. 2015. Gctf: Real-time CTF determination and correction. *J Struct Biol* 193:1–12. doi:10.1016/j.jsb.2015.11.003

CHAPTER 2

Characterization of the alternative translation initiation factor eukaryotic initiation factor

2A (eIF2A)

INTRODUCTION

Before a protein is synthesized and targeted to its final destination, there are a myriad of regulatory steps to ensure the correct protein is being made at the right time. Faithful regulation of protein synthesis is necessary for all living organisms. To adapt to changing conditions, various signaling pathways converge on the control of protein synthesis. It is no surprise that cancerous cells often utilize aberrant signaling pathways to maintain their high demand for protein synthesis while still growing and dividing. With the increasing body of evidence that the translational landscape is changed in these cells, it is imperative to determine on a mechanistic level how aberrant signaling pathways alter protein synthesis. Eukaryotic initiation factor 2A (eIF2A; denoted with a capital “A” — eIF2A is entirely different from trimeric eIF2 and its subunit eIF2-alpha) has been implicated as an important piece of this puzzle: it is (1) an alternative to eIF2-dependent translation, (2) essential for tumorigenesis in an aggressive form of skin cancer, and (3) has been shown to promote sustained translation of certain transcripts during global attenuation. Elucidating the pathway in which eIF2A functions will not only provide possible drug targets but also a deeper understanding of how proteotoxic stress is managed by cells.

In eukaryotes, translation initiation begins with the formation of the ternary complex (TC), which consists of eukaryotic initiation factor 2 (eIF2) bound to GTP and an initiator methionyl-tRNA ($\text{Met-tRNA}_i^{\text{Met}}$). The TC binds to the 40S subunit, along with eukaryotic initiation factor 1, 1A, 3, and 5, forming the 43S preinitiation complex. This complex is recruited to mRNA through the cap-binding complex (eIF4F complex bound to the 5' cap of the mRNA). The eIF4F complex consists of the cap-binding protein, eIF4E, the scaffold eIF4G, and the DEAD-box helicase eIF4A. Following cap recognition, the complex scans the mRNA until it finds an AUG codon in favorable context, forming the 48S preinitiation complex. An intricate series of events allows for release of the initiation factors from the 40S subunit and the binding of the 60S subunit, resulting in an 80S ribosome primed for elongation.

There are several safe-guards to ensure fidelity in start codon selection. In most cases, the first AUG encountered by the ribosome scanning 5' to 3' is the start codon. The sequence surrounding the start codon has conserved features that are predictive of its translational efficiency, referred to as the Kozak sequence (Kozak et al. 1986, 1987). Studies in a yeast reconstituted system have supported a model in which full incorporation (P_{in}) of TC into the P-site requires the dissociation of eIF1, which only occurs upon the proper codon-anticodon interaction between the mRNA in a favorable sequence context and $tRNA_i^{Met}$ (Maag et al., 2005). Similarly, the C-terminal tail of eIF1A promotes the scanning state. Upon recognition of an AUG, the rotational freedom of the tail is reduced and no longer occludes the TC from the P_{in} state. Interactions between these factors as well as eIF5, eIF3, and eIF2 ensure that the ribosome only completes scanning when the correct codon is in the P-site.

With the advent of ribosome profiling, however, it has become clear that canonical translation initiation does not capture the whole picture. Ribosome profiling showed that there is often ribosome occupancy in the 5' untranslated regions (UTRs) of genes at not just upstream AUGs, but also near cognate codons such as CUG and UUG (Ingolia et al., 2009). Further studies confirmed that these were not just artifacts of the protocol, but in several cases, upstream near-cognate codons served as the starting point for a functional product. One particularly interesting example of this was seen with the tumor suppressor PTEN. In addition to the 55 kDa protein, alternative initiation at an upstream CUG results in the extended protein isoform PTEN-long (PTEN-L) (Liang et al., 2014). Translation of the long isoform was shown to be dependent on the Kozak sequence as well as the elusive, and unfortunately named, initiation factor eukaryotic initiation factor 2A (eIF2A). Unlike PTEN-short, the product was targeted to mitochondria where it enhanced ATP production. These studies demonstrated that PTEN isoforms can serve different biological functions and that the expression of these isoforms involves different translational machinery.

Characterization of the implicated factor eIF2A began in the early 1970s while researchers were establishing the basic requirements for protein synthesis. Originally termed IF-M1, eIF2A is a member of

the WD-40 repeat protein family that was first isolated by William Merrick's lab from the ribosomal fraction of rabbit reticulocyte lysate (Adams et al., 1975). His group showed that this factor stimulated initiator tRNA binding to the 40S subunit in the presence of a start codon. From these results, Merrick proposed a model in which initiation through eIF2A is an alternative to the GTP-dependent, eIF2 pathway.

eIF2A was later shown to be important for the translation of non-AUG initiated peptides. Conserved from yeast to humans, the gene was originally thought to be nonessential in the wild-type context of all studied systems. Subsequent studies, however, showed that having eIF2A was important for responding to certain environmental cues. Supporting this idea, eIF2A-dependent translation has been observed in a range of contexts, from MHC class I antigen presentation (Starck et al., 2012), to the integrated stress response (ISR) (Starck et al., 2016), and most recently, squamous cell carcinoma (SCC) (Sendoel et al., 2017). In this aggressive form of skin cancer, the authors observed a prominent shift of ribosome occupancy from AUG to non-AUG codons in the 5'-UTRs of mRNAs, including those encoding the tumor suppressor PTEN. The SCC mouse model also exhibited a striking tumorigenesis phenotype, in which knocking out eIF2A not only dramatically reprogrammed the translational landscape of the cells but also completely abrogated tumor formation. There is compelling evidence that eIF2 α phosphorylation resulting from ISR induction is important for tumorigenesis, suggesting alternative translational pathways may be utilized to maintain protein synthesis in tumor growth (Nguyen et al., 2018).

Considering the proposed role of eIF2A in PTEN-L translation, and its essentiality for SCC malignancy, this non-canonical initiation pathway likely represents an important node in cellular transformations, with potential for therapeutic targeting. Despite the initial biochemical characterizations of this eIF2-independent translation initiation pathway and the more recently demonstrated importance of eIF2A, the specifics of eIF2A molecular requirements and the pathways involved have yet to be elucidated. The

current model in the field is that eIF2A is a leucyl-tRNA carrier that is active when ternary complex is limiting, although this has not been shown directly, i.e., biochemically.

Transcripts dependent on eIF2A for translation

While the current model in the field has eIF2A defined as an alternative initiation factor to eIF2, there have been a limited number of studies that have identified transcripts dependent on eIF2A for translation. There are three classes of mRNAs whose translation eIF2A has been implicated in 1) Internal Ribosome Entry Site (IRES) containing 2) non-AUG initiated and 3) upstream Open Reading Frame (uORF) containing mRNAs.

Hepatitis C viral (HCV) mRNA was the first transcript that was shown to have eIF2A dependence (Kim et al., 2011). HCV mRNA contains an IRES, and as such, utilizes different machinery than the canonical cap-containing mRNAs. Translation of HCV was known to be independent of eIF2-GTP-Met-tRNA_i ternary complex from several studies (Robert et al., 2006, Terenin et al., 2009), leading the Jang lab from Pohang University of Science and Technology to probe the role of eIF2A in this process. Kim et al. used luciferase assays in cells to show that with limited eIF2, the translation of HCV was maintained in control cells. However, when eIF2A was knocked down the protein levels were reduced by around 50%. They also used radioactively labeled Met-tRNA_i to conclude that eIF2A was required for the binding of Met-tRNA_i to the 40S in the presence of the HCV mRNA. Finally, infection of Huh-7.5.1 cells with Hepatitis C virus was reduced around 50% in cells with eIF2A knocked down, using the core of HCV as a readout for viral levels.

Another compelling piece of evidence for eIF2A-dependent translation came from the Yin lab of Peking University (Liang et al., 2014). They discovered a long isoform of the tumor suppressor phosphatase and tensin homolog (PTEN) that was the result of translation initiation at an upstream non-AUG start codon.

Similar to an IRES-containing transcript, they hypothesized that secondary structure in the 5'UTR of this gene recruited alternative initiation factors, such as eIF2A, to initiate translation at this non-canonical start codon. Through western blot analysis, they noted a drastic decrease in the protein levels of the long isoform of PTEN, while the canonical isoform was unaffected. From this study, eIF2A was also implicated in non-AUG translation initiation, a phenomenon that was gaining interest in the field due to the aforementioned ribosome profiling results showing the prevalence of non-AUG start codon usage.

Finally, eIF2A was again implicated in eIF2-independent translation in a study from the Walter lab which examined the relationship between uORF translation and the sustained translation of Binding immunoglobulin protein (BiP) and Activating Transcription Factor 4 (ATF4) (Starck et al., 2016). The mRNA for both genes contain uORFs, however, the uORFs for BiP utilize non-AUG start codons whereas ATF4 uORFs initiate at the canonical AUG. The authors knocked down eIF2A and saw a decrease in uORF translation for BiP but not ATF4. They also saw a slight decrease in BiP levels upon knocking down eIF2A but only in the presence of stress. Mutating one of the uORFs and knocking down eIF2A showed a strong decrease in BiP levels in the presence of stress. From this, the authors concluded that eIF2A and the uORFs for BiP were essential for the sustained translation of the downstream ORF during stress.

Between these three transcripts (HCV, PTEN-L, and BiP) there are some common features. For HCV and BiP, limiting eIF2 α results in a dependence on eIF2A. PTEN-L and BiP share the use of non-canonical start codons upstream of the annotated AUG start codon, and all three have reported secondary structure that could contribute to the recruitment of specific translation factors. The mechanistic detail in how eIF2A is participating in the translation of these transcripts, however, remains unclear, and to-date there is no systematic analysis of what constitutes an “eIF2A-dependent” transcript.

Conflicting data in the field on eIF2A-dependent transcripts

Redundancy in initiation factors, as well as cell line specific effects, makes establishing the role of alternative factors especially difficult. Following the HCV study from Kim et al., another study was published simply titled “The Initiation Factors eIF2, eIF2A, eIF2D, eIF4A, and eIF4G Are Not Involved in Translation Driven by Hepatitis C Virus IRES in Human Cells” (González-Almela et al., 2018). In this study, the authors also used a luciferase assay to look at HCV mRNA translation, however, this time the authors saw no difference in eIF2A⁻ HAP1 cells compared to Wild-Type. Besides using a different cell line, the more recent study differs in that the measurements were made in an eIF2A knockout cell line whereas Kim et al. used a knockdown of eIF2A. Perhaps compensatory methods in the knockout cell line could mask the defect in translation. The role of eIF2A in HCV translation remains controversial.

The drastic and specific decrease in PTEN-L expression in eIF2A knockout cells made this an attractive transcript to use as a reporter for eIF2A-dependent translation. Unfortunately, in our hands we were unable to replicate the PTEN-L result in HeLa cells by western blot analysis (Figure 2.1). Using an antibody against PTEN, we saw a slight decrease in the levels of both the AUG-initiated short isoform (PTEN-S) and the CUG-initiated long isoform of PTEN (PTEN-L) in the knockout cell line. We did not see the same dramatic decrease in PTEN-L levels that Liang et al. reported.

While there is clearly a need for systematic identification and analysis of eIF2A-dependent transcripts, this has proven to be difficult. In the squamous cell carcinoma (SCC) study that identified eIF2A as a key factor for tumorigenesis, they performed Ribo-seq and saw changes in the 5'UTR occupancy in eIF2A knockout cells (Sendoel et al., 2017). There was no accompanying RNA-seq, which makes analysis of these data difficult in regards to translational efficiency. Other unpublished studies have observed phenotypes in cells and organisms lacking eIF2A but were unable to detect differences in translation, suggesting eIF2A might also be playing a role outside of translation.

With the conflicting data in the field, we became interested in re-examining the basic model for eIF2A. Instead of operating under the assumption that eIF2A was an alternative tRNA_i carrier to eIF2, we wanted to utilize two simplified systems to characterize eIF2A: 1) an *in vivo* yeast model 2) an *in vitro* system developed at UC Davis in the Fraser lab for studying translation initiation where we could test the effect of eIF2A on the canonical machinery.

RESULTS

eIF2A and eIF4A (TIF1) interact genetically in *Saccharomyces cerevisiae*

We wanted to take advantage of the large volume of genetic interaction datasets in the budding yeast *Saccharomyces cerevisiae* to create a yeast strain that was sensitive to eIF2A deletion, as deletion of eIF2A alone has no apparent phenotype (Komar et al., 2005). From published Epistatic MiniArray Profiles (E-MAPs) (Wilmes et al., 2008), eIF2A was shown to have a strong negative genetic interaction with TIF1 and TIF2 (Figure 2.2A). As a result of the ancestral whole genome duplication event, TIF1 and TIF2 both encode for the DEAD box helicase eIF4A, an important member of the eIF4F complex. In yeast, eIF4A is one of the two main helicases responsible for unwinding secondary structure in the 5'UTR of mRNA during the initiation of translation. While Δ TIF1 Δ TIF2 double deletion cells are not viable, deleting just one of these genes results in a yeast strain that has reduced levels of eIF4A but is still viable with no observable growth defect. To begin to examine the relationship between eIF4A and eIF2A, we first confirmed the genetic interaction using both serial dilution plating assays and plate reader growth assays. As previously reported, when we deleted eIF2A or TIF1 alone there was no observable growth defect compared to WT (Figure 2.2B and C). When TIF1 was deleted in the Δ eIF2A background, however, the strain had a significant growth defect that was evident in both assays.

Strains without eIF2A and TIF1 show a strong defect in translation

We wanted to test whether the growth defect we had observed was concomitant with a defect in translation. To do so, we ran polysome gradients using lysates from the different genetic backgrounds and looked for changes in the monosome and polysome peaks. As expected, the eIF2A deletion strain showed a similar polysome profile when compared to WT (supplemental figure 2.1). The Δ TIF1 profile showed a slight defect in translation, evident by the increase in the monosome peak and reduction in the A_{254} signal for the polysome peaks (Figure 2.3A, blue trace). When compared to WT and Δ TIF1, the Δ eIF2A Δ TIF1 strain showed a significant increase in signal for the monosome peak and collapse of the polysomes (Figure 2.3A, green trace), suggesting a strong defect in translation initiation. As a complementary approach for measuring bulk translation, we also did a puromycin assay to quantify translation products, where we did puromycin labeling followed by western blot analysis with an antibody against puromycin (Figure 2.3B). Lysate from strains without eIF2A or TIF1 showed similar signal to WT whereas the lysate from the double mutant had a strong reduction in signal, also suggesting a defect in translation.

Ribosome profiling shows changes in translational efficiency in the Δ eIF2A Δ TIF1 background

To further investigate the gene-specific translational changes taking place in the four genetic backgrounds (WT, Δ eIF2A, Δ TIF1, and Δ eIF2A Δ TIF1) we used mRNA-sequencing in conjunction with ribosome footprint sequencing analysis (ribosome profiling or Ribo-Seq). In agreement with the growth assays and bulk translation assays, ribosome profiling showed significant changes in the translational efficiency (TE) of genes in the double mutant. 424 genes showed a greater than 0.5-fold \log_2 decrease in translational efficiency in the Δ eIF2A Δ TIF1 background compared to just 74 in the Δ TIF1 background and 87 in the Δ eIF2A background.

To compare the general properties of the mRNAs that were downregulated across the different genetic backgrounds, we calculated the z-score for the TE of each gene and compared the groups of genes with a z-score below -1.5 (Figure 2.4A). We then used the genome-wide parallel analysis of RNA structure (PARS) dataset to quantify the degree of mRNA structure in these different groups (Kertesz et al., 2010). PARS analysis of the 5'UTRs of the translationally downregulated genes in the Δ TIF1 background and double mutant showed a significant increase ($p < 0.0001$) in the total PARS score of the 5'UTR when compared to all mRNA in the dataset (Figure 2.4B). This difference could partially be explained by the increase in 5'UTR length of the genes downregulated in these two backgrounds (Figure 2.4C), although other measurements that are length-independent (First30 and Max30) showed a significant increase in the double mutant as well. Of note, there did not appear to be any significant changes in the mRNA length or structure of the genes downregulated in the Δ eIF2A background when compared to all mRNA in the dataset.

One of the most translationally downregulated genes in the ribosome profiling dataset was alcohol dehydrogenase isoenzyme type IV (ADH4), which showed a strong dependence on the presence of eIF2A and TIF1. ADH4 has an unusually long 5'UTR for *Saccharomyces cerevisiae* with a length of 268 nucleotides (compared to 79 nucleotides for all mRNA) that may contribute to the requirement for these two factors. Furthermore, the average PARS score for the 5'UTR is 0.19 (compared to 0.06 for all mRNA), reflecting a higher propensity to form secondary structures in the 5'UTR. While RNA Reads Per Kilobase of transcript, per Million mapped reads (RPKM) values were consistent among the four backgrounds, there was a 3-fold reduction in ribosome protected fragment RPKM values in the double mutant. We C-terminally FLAG-tagged endogenous ADH4 and were able to see a strong reduction in the protein levels of ADH4 by western blot analysis (Figure 2.4D). Additional experiments mutating this region are needed to show whether the translational dependence is due to properties of the 5'UTR.

Δ eIF2A Δ TIF1 yeast exhibit a sensitivity to homeostatic stressors

GO Term analysis of the translationally downregulated genes in the double mutant showed an enrichment for stress responsive genes as well as metabolic genes (Figure 2.5A). One of the genes included in this list is the well-characterized uORF-containing transcript for General Control Nondepressible 4 (GCN4).

GCN4 is a transcriptional activator that responds to amino acid starvation and other metabolic stressors. It has a long 5'UTR with four uORFs that regulate the translation of the downstream coding sequence.

Using an antibody against endogenous GCN4, we saw by western blot analysis that levels of GCN4 protein were indeed drastically reduced in the Δ eIF2A Δ TIF1 background (Figure 2.5B). Unlike with ADH4, however, there were lower levels of GCN4 in the single mutants as well, suggesting that the decrease in the double mutant is an additive effect from losing both factors. Since expression of GCN4 has been shown to be the result of phosphorylation of eIF2 α , we also blotted for eIF2 α using a phosphoserine-51 specific antibody. In agreement with the downregulation of GCN4 in the double and single mutants, we also saw a decrease in the levels of phospho-eIF2 α . Future experiments are needed to determine whether this result is due to an overall decrease in eIF2 α levels or a deficiency in the phosphorylation of eIF2 α .

Consistent with a decrease in GCN4, Panther GO-slim analysis of the differentially expressed genes from RNA-seq in the double mutant showed a significant enrichment for alpha-amino acid metabolic processes (GO:1901605) among other metabolic processes (Figure 2.5C). We wanted to see if there were consequences for the cell in response to stressors that upregulate GCN4 such as H₂O₂ and rapamycin. H₂O₂ is an oxidative stressor for the cell and rapamycin inhibits TORC1, mimicking amino acid starvation. Both stressors activate the lone eIF2 α kinase, GCN2. Treatment with H₂O₂ and rapamycin did indeed result in phenotypic changes in the different genetic backgrounds, with the most notable growth defect in the double mutant following rapamycin treatment (Figures 2.5D-F). Interestingly, the sensitivity

to stressors did not follow the same trend as GCN4 expression, suggesting the phenotype was not solely due to the impairment in GCN4 expression. The $\Delta eIF2A$ strain showed a similar delay to $\Delta eIF2A\Delta TIF1$ in recovery from H_2O_2 stress, however the rate of growth following that delay was comparable to WT (Figure 2.5D). Following rapamycin treatment, $\Delta eIF2A$ had a reproducible growth defect that was exacerbated by the deletion of TIF1 (Figure 2.5E).

From these experiments, we were able to show that when the canonical translational apparatus is limited (TIF1 deletion, limiting eIF2 α through GCN2, limiting eIF4F through TORC1) eIF2A becomes important in yeast. This contextual requirement could be the result of eIF2A participating in an alternative complex, assisting the canonical machinery, or an indirect effect that is not necessarily due to modulation of translation initiation.

Recombinant expression of human eIF2A

To help get at the true mechanism of eIF2A-dependent translation we were particularly interested in obtaining a reliable, purified protein preparation. Due to the powerful existing *in vitro* reconstituted initiation systems that utilize the human proteins, we decided to attempt to purify the human version of eIF2A. Previous reports have conflicting results regarding recombinant expression and purification of eIF2A. To date, only one group has published results using *Escherichia coli* (*E.coli*) purified protein (Kim et al., 2011). They demonstrated that eIF2A stimulated tRNA binding to the 40S in the presence of HCV mRNA using filter binding assays. Another group, however, showed that the stimulation of tRNA binding to the 40S subunit that was originally shown by William Merrick's group was misattributed to eIF2A due to a

contaminated preparation, and rather, an alternative initiation factor eIF2D was the factor responsible for GTP-independent binding of tRNA to the small subunit (Dmitriev et al., 2010). Due to these discrepancies in the literature, we were interested in preparing our own sample to examine the basics *in vitro* of eIF2A's role in translation initiation.

Recombinant expression of human eIF2A in *Escherichia coli* (*E.coli*)

Recombinant protein expression in *E.coli* has several advantages over mammalian systems, mostly in regards to the cost and ease of growing a large quantity of cells. As previously mentioned, one group had success in using *E.coli* to produce eIF2A for *in vitro* binding assays and we used this protocol as a starting point. In our hands, expression and purification of human 6xHis-eIF2A appeared to be successful at first. We were able to obtain a clean preparation of a protein that eluted around 12 mL on the S200 size exclusion column (Figure 2.6A). and ran around the expected molecular weight on an SDS page gel (Figure 2.6B).

There were a couple problematic aspects of this preparation. The first being that the yield was low, with 6 liters of culture only producing a couple hundred micrograms of protein. The second was that incubation with TEV protease did not result in detectable tag cleavage. Finally, by mass spectrometry the size of the product was around 0.4 kDa too small (Figure 2.6C). To address these concerns, we performed protein identification using mass spectrometry in collaboration with QB3 at UC Berkeley. We were surprised to discover that the primary protein in the sample was the *E.coli* protein Glutamine-fructose-6-phosphate aminotransferase (GlmS) that was 66.9 kDa and co-eluted with eIF2A (Figure 2.6D). There were also a small number of peptide counts

for eIF2A, which explains the positive signal for eIF2A by western blot analysis that we had previously seen. We were unable to separate eIF2A from this contaminating species.

Recombinant expression of human eIF2A in Expi293 cells

Based on these results we decided to change our expression system to a mammalian system. eIF2A is not conserved in bacterial species and we reasoned that the protein folding machinery in a mammalian system may help with its expression and purification. The cell line we used, Expi293, is a commonly used cell line for expression that is suspension-adapted and derived from Human Embryonic Kidney (HEK) cells. We expressed eIF2A with an N-terminal 6xHis-tag and a ybbR tag either following the His-tag or at the C-terminus. These constructs readily expressed in the Expi293 cells (Figure 2.7A) and resulted in a purified product following size exclusion chromatography (SEC) (Figure 2.7B and C). 6xHis-eIF2A-ybbR and 6xHis-ybbR-eIF2A eluted around 13 mL on the S200 column, which was earlier than we would expect for a protein that was 68 kDa. Indeed, after running the standards through the column, we were able to confirm that eIF2A eluted as a dimer (Figure 2.7D). Finally, mass spectrometry analysis of the purified product showed the primary product was 68.9 kDa, the expected size of the 6xHis-eIF2A-ybbr construct (Figure 2.7E). In contrast to the *E.coli* preparation, eIF2A appeared to be readily expressed and isolated from Expi293 cells.

Preliminary *in vitro* binding assays suggest eIF2A possesses weak general RNA binding properties

In collaboration with Dr. Sokabe of Dr. Fraser's lab at UC Davis, we conducted a series of preliminary binding assays. In order to observe anisotropy, we first covalently labeled eIF2A with fluorescein by Sfp phosphopantetheinyl transferase-catalyzed conjugation to the ybbR tag on the recombinantly expressed eIF2A (Yin et al., 2006). We then incubated increasing amounts of 40S or 60S ribosomal subunits with 18 nM of fluorescently labeled eIF2A and measured the anisotropy change (Figure 2.8A). We were surprised to see in this initial trial that while eIF2A showed a moderate affinity for the 40S subunit (K_d of 270 nM), it had a lower K_d for the 60S subunit (K_d of 30 nM). At this stage, it is unclear if these are functional interactions. By chasing labeled eIF2A off 40S or 60S subunits with unlabeled eIF2A, we saw that the inhibitor constant (K_i) was 9-fold higher than the K_d for the 60S (K_i of 270 nM) (Figure 2.8B). This could indicate that eIF2A is binding on multiple sites of the 60S, which would contribute to the surprisingly low K_d . From polysome gradient analysis we see that eIF2A primarily co-sediments with the 40S subunit, so the interaction with the 60S that we observed may be non-specific.

Finally, we tested the binding of eIF2A to fluorescently labeled RNA molecules. eIF2A has been reported to bind both mRNA and tRNA_i. What we saw in these initial assays, however, was a rather weak affinity for the tested RNA species (Figure 2.8C). We tested a CAA-repeat containing mRNA as well as a more structured Globin-UTR mRNA. Binding of the CAA-repeat mRNA to eIF2A showed the tightest binding with a K_d around 700 nM. Both the Globin-UTR mRNA and tRNA_i bound to eIF2A in the low micromolar range. For comparison, the canonical initiator-tRNA carrier, eIF2, binds unacylated tRNA_i with a K_d of 130 nM (Kapp et al., 2004).

These assays are all preliminary and there are countless additional conditions that we can test now that we have a reliable preparation of eIF2A. We can examine how the presence of additional initiation factors affects eIF2A's association with the 40S subunit and if eIF2A physically interacts with these factors or is

competing for their binding sites. There are several potential mRNA molecules that have been associated with eIF2A-dependent translation that could serve as a starting template for building up this alternative initiation complex. The previous experiments in the literature examining eIF2A's interaction with these different components were all qualitative, so being able to deeply explore eIF2A's interaction with the translational machinery in a quantitative manner will be incredibly valuable to the field.

DISCUSSION

To maintain protein homeostasis, the cell relies heavily on the regulation of translation initiation. Regulation is achieved by controlling the availability of canonical initiation factors as well as the utilization of alternative factors, such as eIF2A. As translational regulation is critical for the cell, redundancies built into the system make studying the role of a particular factor especially difficult. Through these studies, we have encountered a “slight change” from loss of eIF2A at almost every turn, and a quick search through the literature will corroborate this finding. By sensitizing the yeast system with reduced levels of eIF4A, we not only were able to characterize the relationship of eIF2A to a canonical initiation factor, but we were able to actually observe changes in the cell due to loss of eIF2A. Through ribosome profiling we saw that loss of eIF2A alone has minimal effects on the translational efficiency of genes. We also saw that losing eIF2A in the eIF4A-depleted background did not change the average length or PARS score of downregulated mRNA, however, the translational efficiency of genes were drastically reduced. This result suggests eIF2A is not influencing mRNA selection but rather the efficiency by which mRNA are translated. Further mechanistic studies will be necessary to establish where in the initiation pathway eIF2A is functioning.

Now with a purified sample of eIF2A and the proper assay system, we can both test long-held assumptions about the function of eIF2A as well as start to build-up a more comprehensive model. We have already seen that this positively charged protein interacts with every type of RNA we tested, highlighting the importance of carefully controlled *in vitro* experiments. We know from polysome gradient analysis that eIF2A primarily co-sediments with the 40S subunit and we also saw a moderate affinity for the 40S subunit *in vitro*. Compared to other canonical factors, however, the measured affinity for the 40S subunit was rather low. This result could be indicative that additional proteins are needed to secure this interaction or that eIF2A requires post-translational modification(s). It could also reflect eIF2A's role as an alternative factor in that the interaction is relevant only when other factors are limited.

Combining these *in vitro* assays with additional *in vivo* characterization could prove invaluable for understanding how this protein functions.

These studies are a considerable effort for a gene that was previously considered to be non-essential. Recent reports looking at eIF2A at both the cellular and organismal level, however, have shown that context matters when evaluating the importance of eIF2A. The first characterization of an eIF2A knockout mouse showed that offspring were viable, with no apparent phenotype when examined at 3-5 months of age (Golovko et al., 2016). The same group followed up with a more in-depth analysis that uncovered a decrease in lifespan as well as a strong metabolic phenotype in the knockout mouse (Anderson et al., 2021). After 1 year of age, the animals without eIF2A exhibited severe obesity and enlarged livers when compared to WT. Furthermore, mice on a high fat diet gained weight at a faster rate than WT mice fed the same diet, a phenomenon that was especially pronounced in the female mice. More recently, in *Drosophila melanogaster* it was shown that disruption of the gene for eIF2A was lethal and a hypomorphic allele rendered males sterile (Lowe et al., 2022). Understanding how eIF2A functions is not only an interesting mechanistic question, but also clearly informs on basic developmental and metabolic processes.

Beyond basic science, there are obvious advantages when considering a context-specific factor as a therapeutic target. As discussed previously, eIF2A was implicated as a key component of the co-opted translational machinery in Squamous Cell Carcinoma. Unlike with a factor such as eIF2, whose inhibition would drastically impair the translation of healthy cells as well as diseased cells, inhibition of eIF2A could specifically target that aberrant growth. A better structural and biochemical understanding of what the eIF2A-containing initiation complex could inform on chemical inhibition

FIGURES

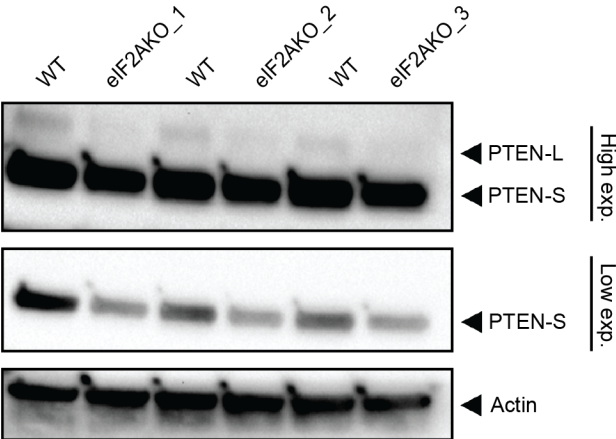


Figure 2.1: PTEN does not have isoform specific changes in eIF2A knockout HeLa cells

Western blot analysis of protein lysate from monoclonal HeLa eIF2A knockout cells shows a slight decrease in both the long and the short isoform of PTEN when compared to WT cells.

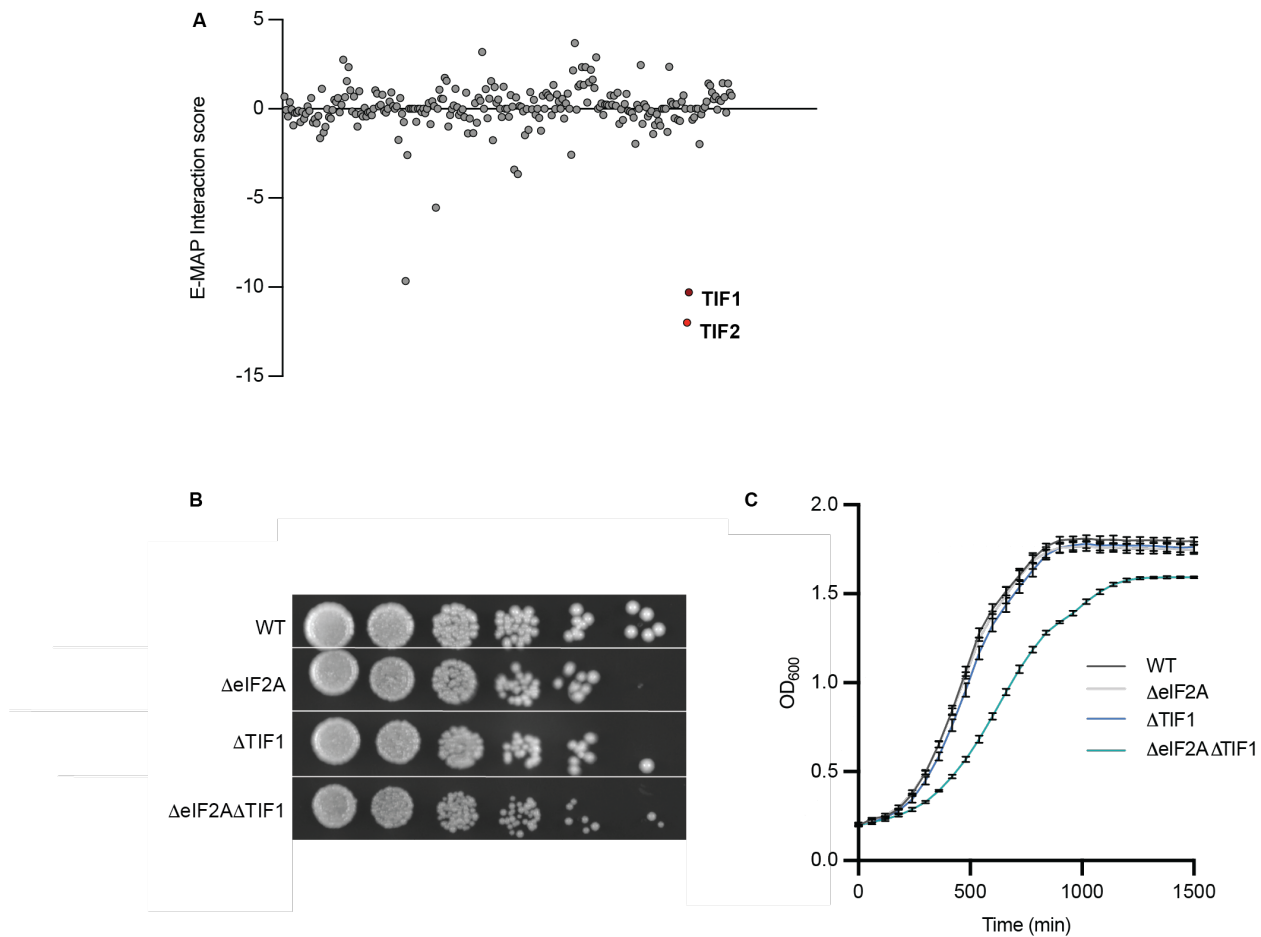


Figure 2.2: eIF2A and TIF1 have a synthetic growth defect

(A) Scatter plot of the E-MAP interaction scores for eIF2A from Wilmes et al. (2008). Each point represents the genetic interaction score between a gene and eIF2A. TIF1 and TIF2 are represented by red points. (B) 5-fold serial dilutions of WT W303 yeast, Δ eIF2A, Δ TIF1 and Δ eIF2A Δ TIF1 on YPD agar after 24 hrs. (C) Growth assay of the above strains in a plate reader format, with OD₆₀₀ as the readout for cell density. Average values of the three replicates are plotted with standard deviation.

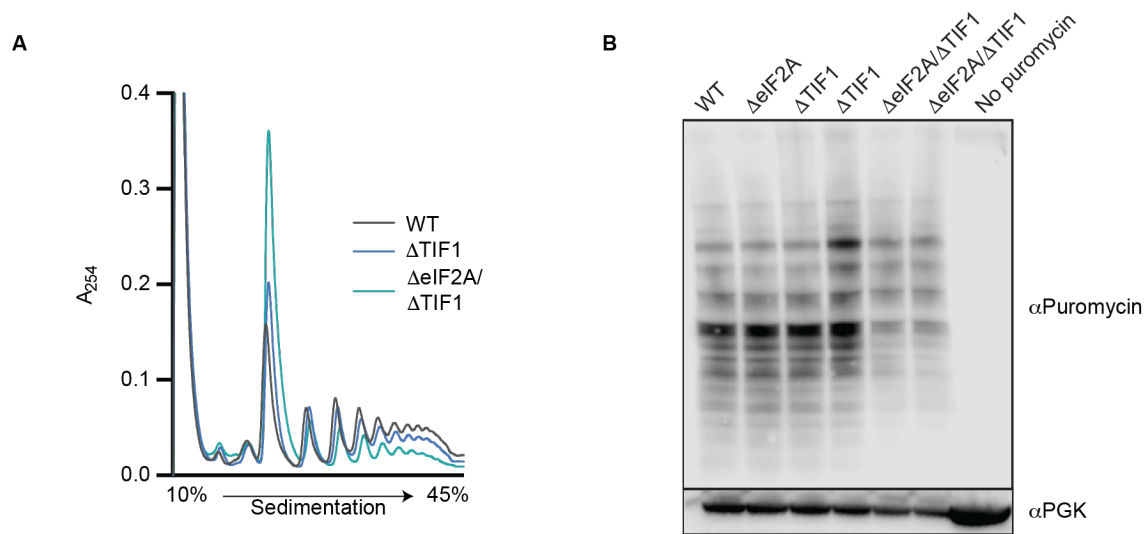


Figure 2.3: General translation is downregulated in the Δ eIF2A Δ TIF1 background

(A) Polysome profiles for WT, Δ TIF1, and Δ eIF2A Δ TIF1 show a strong defect in translation in the double mutant. (B) Similarly, the Δ eIF2A Δ TIF1 mutant shows a decrease in puromycin labeled peptides by western blot analysis when compared to WT and the single mutants. PGK serves as a loading control.

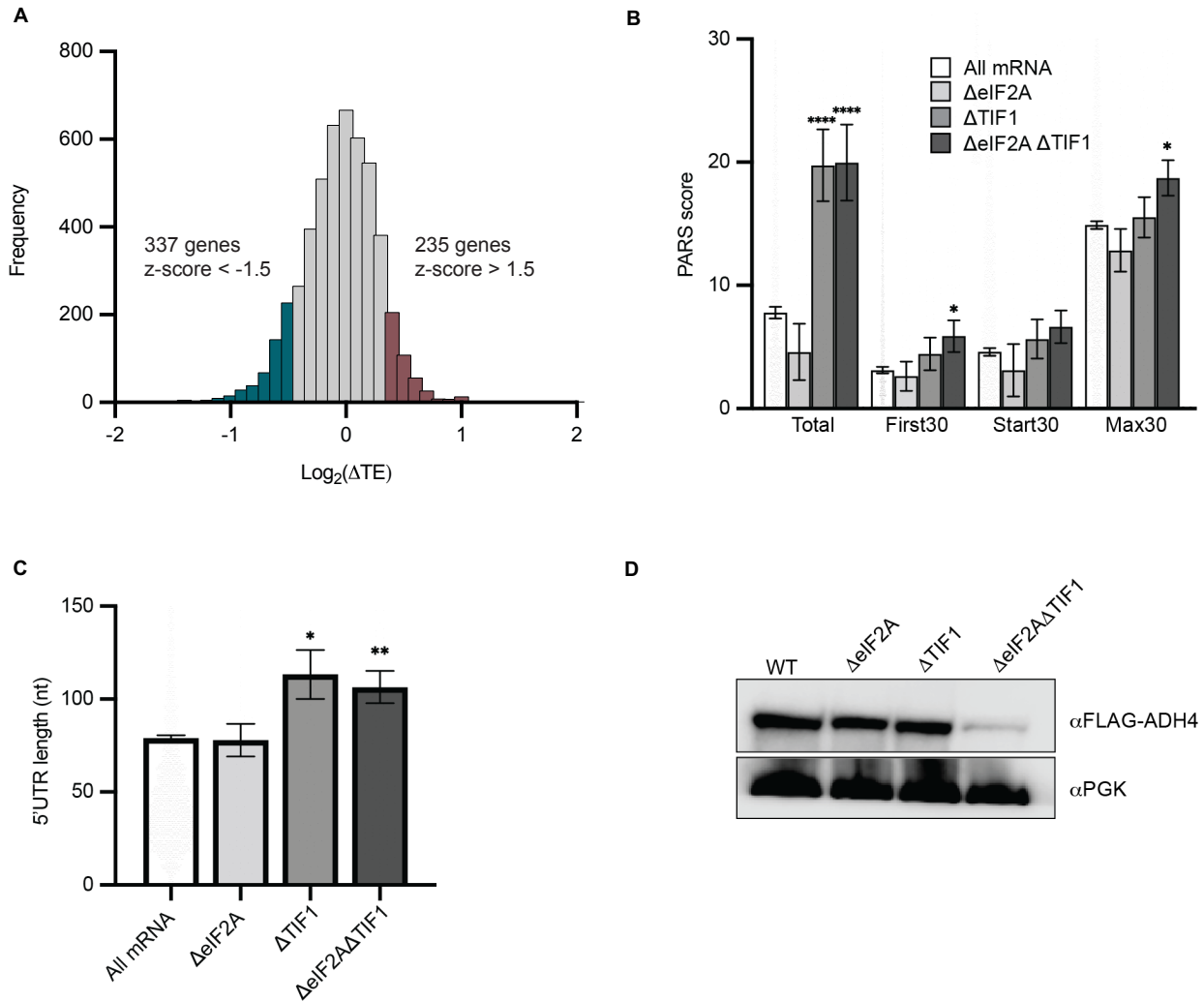
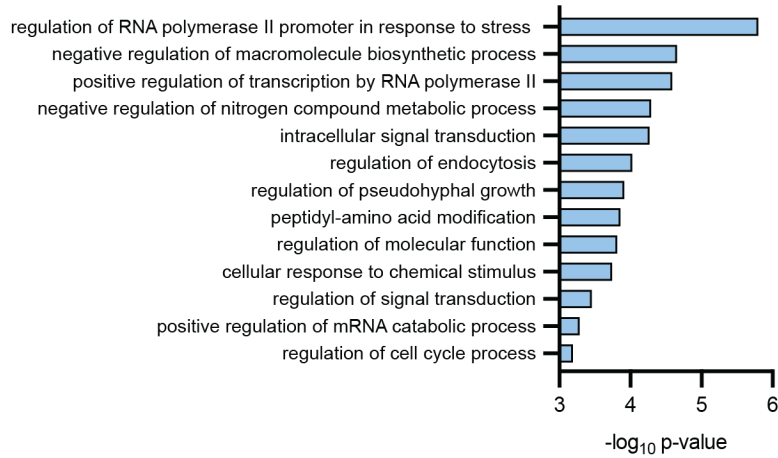
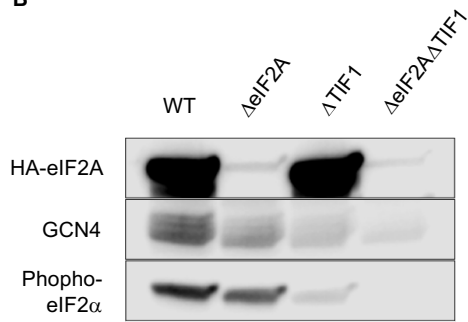
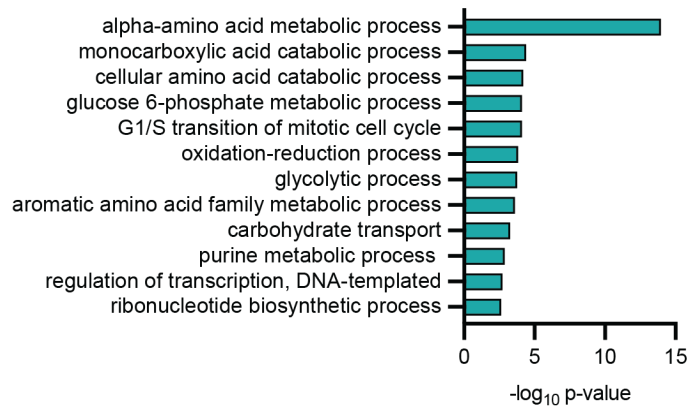
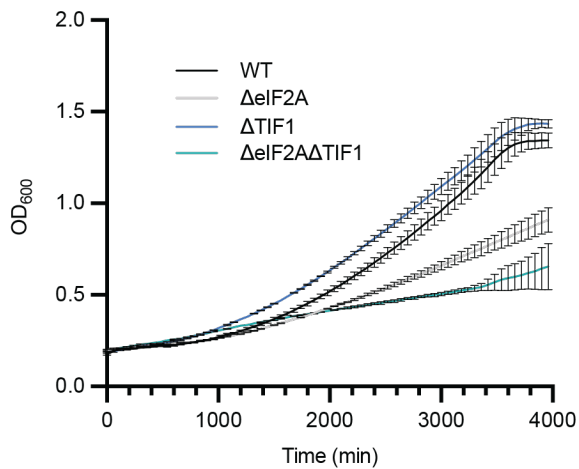
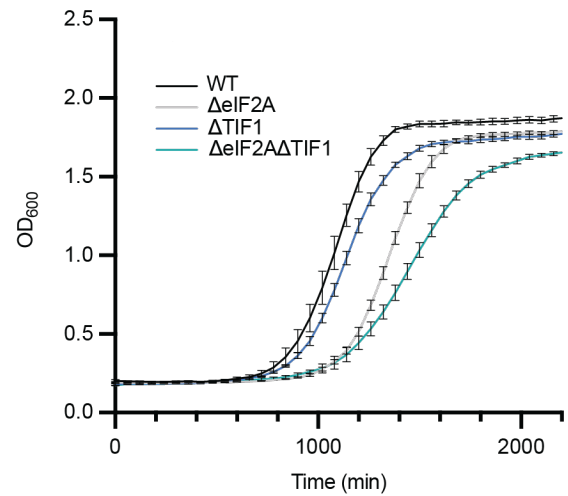


Figure 2.4: Ribosome profiling analysis reveals differentially expressed transcripts under the combined presence of eIF2A and TIF1, with a distinct lack of changes in the ΔeIF2A background

(A) Frequency distribution of translation efficiency changes in the $\Delta\text{eIF2A}\Delta\text{TIF1}$ background used to calculate z-scores for mRNA analysis. (B) PARS scores for the different 5'UTR measurements for each of the genetic backgrounds compared to WT. (C) 5'UTR length distributions for each of the genetic backgrounds compared to WT. (D) Western blot analysis of C-terminally FLAG-tagged ADH4 shows reduced protein levels in the double mutant. PGK serves as a loading control.

A**B****C****D****E**

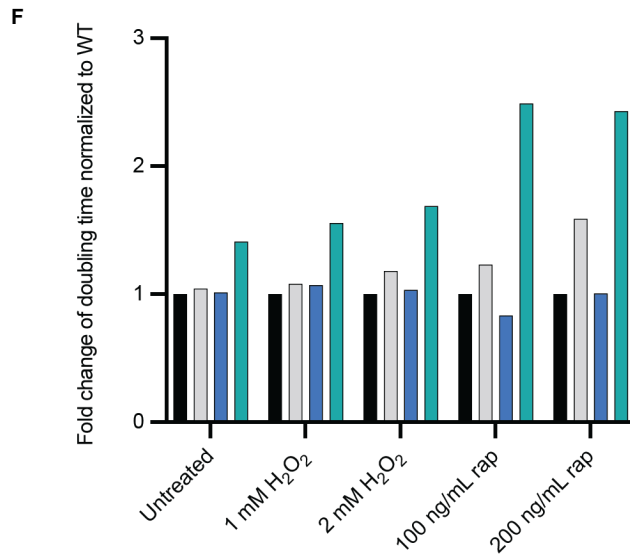


Figure 2.5: The $\Delta eIF2A\Delta TIF1$ strain exhibits metabolic reprogramming and an increased sensitivity to homeostatic stressors.

(A) GO Term analysis of the translationally downregulated genes in the double mutant shows a significant enrichment of stress responsive and metabolic genes. (B) Western blot of lysate from WT, $\Delta eIF2A$, $\Delta TIF1$, and $\Delta eIF2A\Delta TIF1$ with an antibody against GCN4 and PGK as a loading control. (C) GO Slim analysis of the genes downregulated in the double mutant at the RNA level shows a significant enrichment of genes involved in amino acid metabolism. (D) OD_{600} of WT, $\Delta eIF2A$, $\Delta TIF1$, and $\Delta eIF2A\Delta TIF1$ read every hour for 40 hours following treatment with 1 mM H_2O_2 . The mean is plotted with the standard deviation between three replicates. (E) Same as (D) except treated with 200 ng/mL rapamycin for 67 hours. (F) Fold-change of the doubling time relative to WT following treatment with each of the stressors.

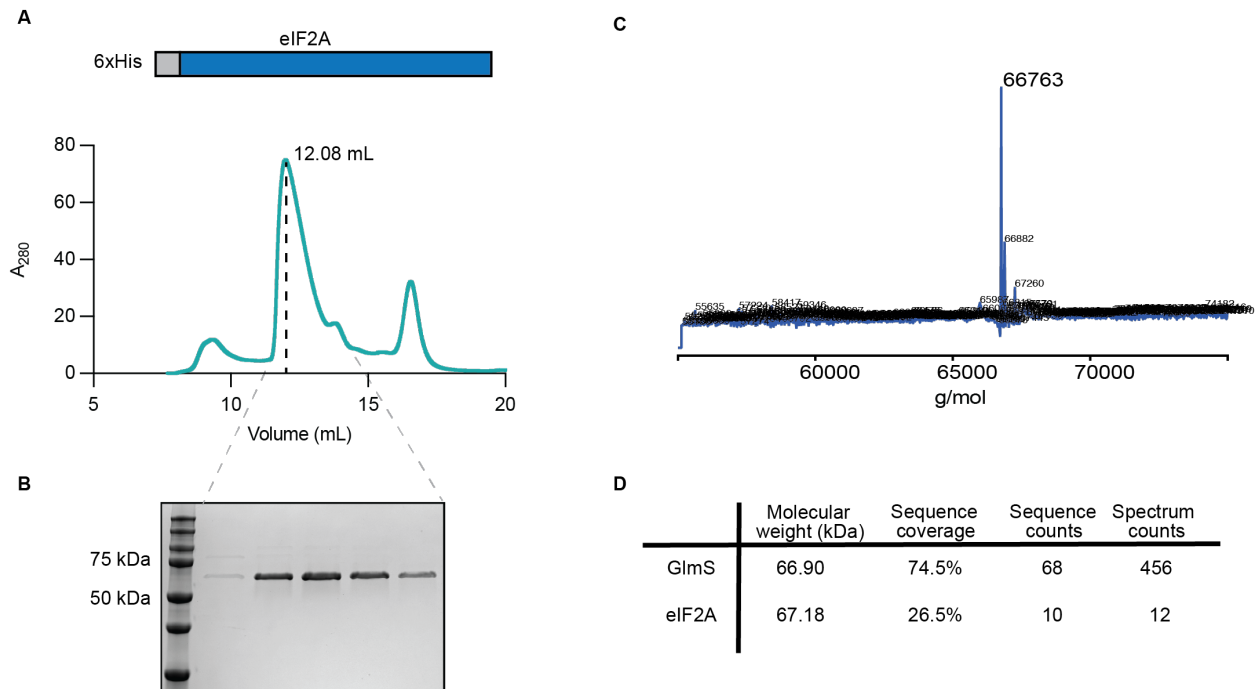


Figure 2.6: Expression and purification of 6xHis-eIF2A in *E. coli* results in the purification of the endogenous protein, GlmS.

(A) Schematic for the expression construct and size exclusion chromatography A_{280} trace for 6xHis-eIF2A purification. (B) SDS-PAGE gel of the fractions from the major peak. (C) Intact protein mass spectrometry on the main peak reveals the major product is around 66.8 kDa. (D) Summary of the results from LC-MS for protein identification. The *E. coli* protein GlmS was the most abundant protein in the sample.

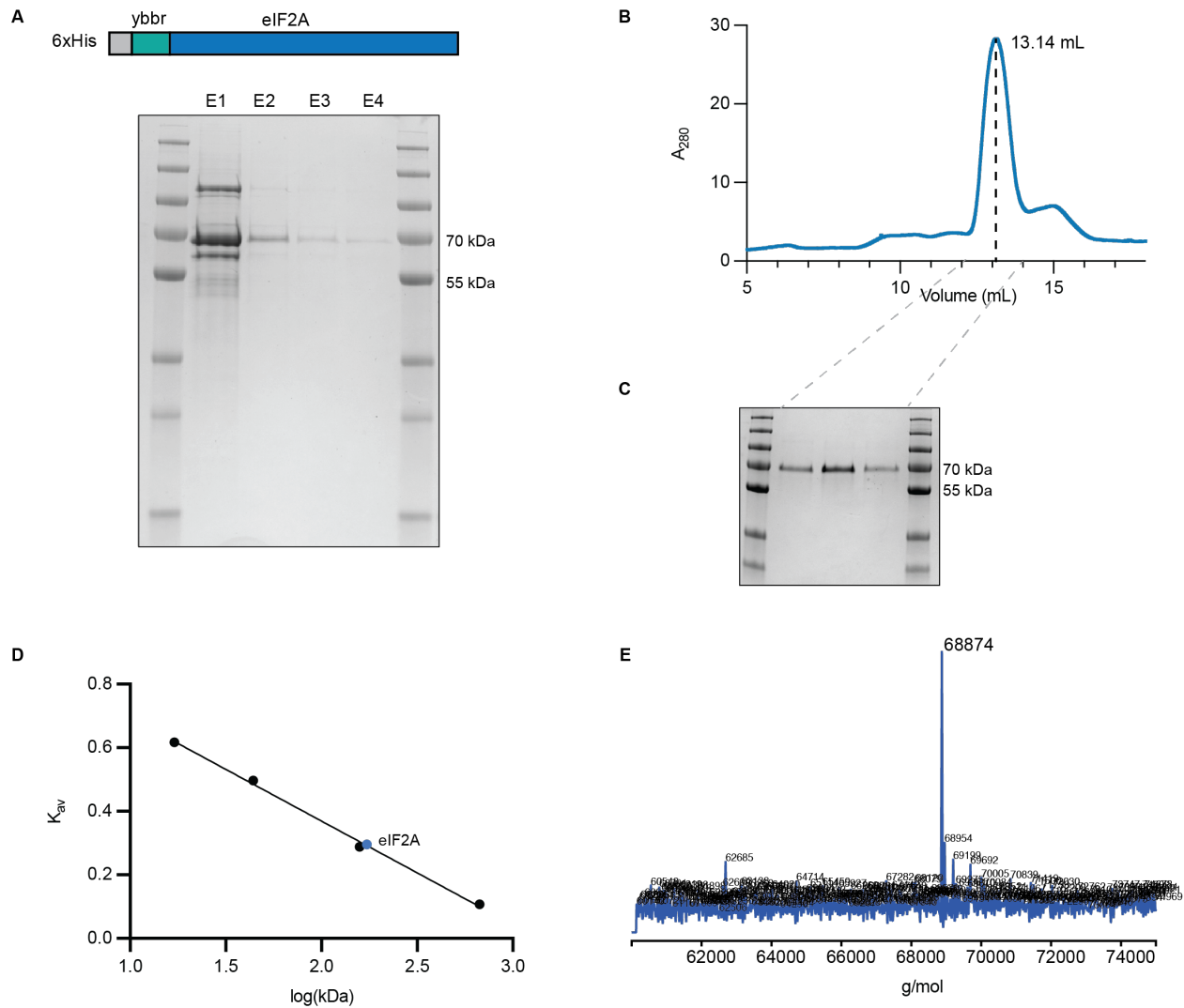


Figure 2.7: eIF2A is readily expressed in Expi293 cells and elutes as a dimer on size exclusion column

(A) Schematic of the construct used for expression and a representative SDS-PAGE gel of the high imidazole elutions from TALON resin. (B) The SEC trace shows that 6xHis-ybbr-eIF2A elutes at 13.14 mL. (C) An SDS-PAGE gel of the fractions from the major peak shows that main product is around 70 kDa. (D) Standard curve for SEC standards shows that eIF2A eluted as a dimer on the S200 SEC column. (E) Mass spectrometry analysis of the final product from the 6xHis-ybbr-eIF2A purification shows a final product of 68.9 kDa.

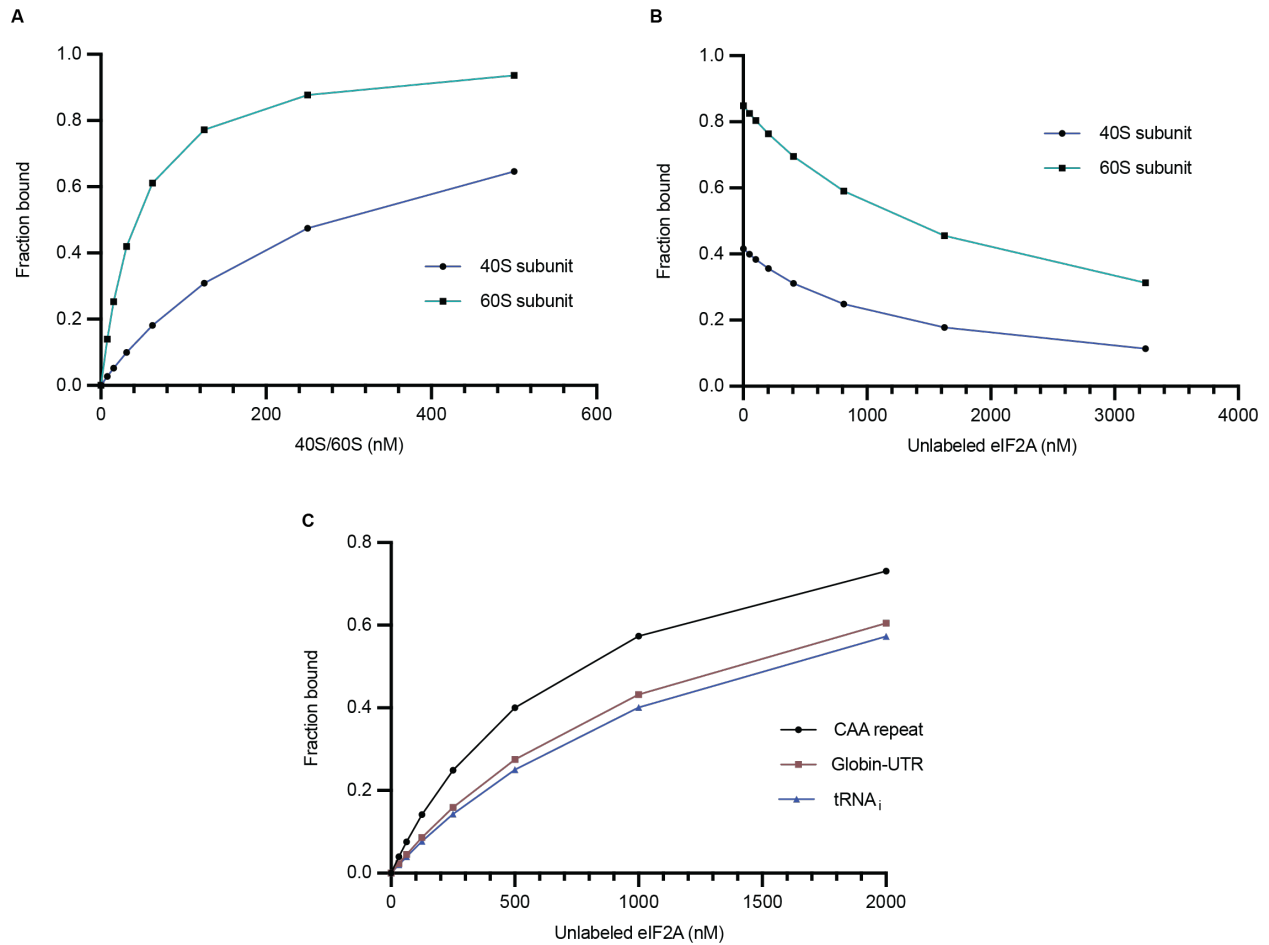
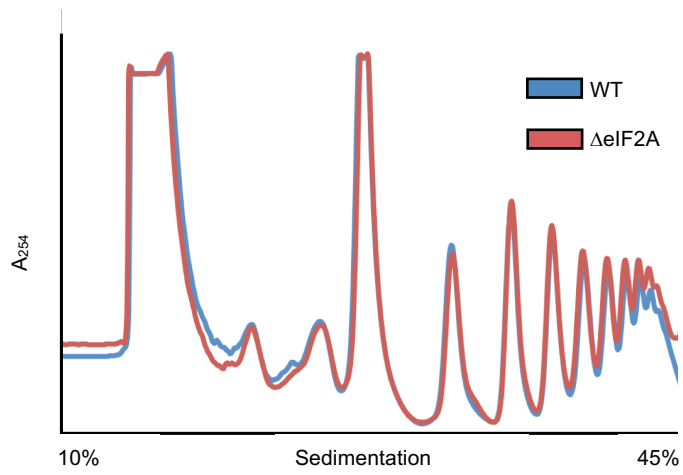


Figure 2.8: Preliminary binding experiments show 6xHis-ybbR-eIF2A has a moderate affinity for ribosomal subunits and a weak affinity for mRNA/tRNA_i

(A) Anisotropy assays with labeled eIF2A and increasing amounts of ribosomal subunits. The fraction bound was calculated and plotted against the concentration of the indicated ribosomal subunit. **(B)** The labeled eIF2A in the complexes formed in (A) were competed off with unlabeled eIF2A to calculate the inhibition constant (K_i). **(C)** Anisotropy assays with labeled mRNA species or labeled tRNA_i and increasing concentrations of unlabeled eIF2A. The fraction bound was calculated and plotted against the concentration of eIF2A.



Supplemental figure 2.1: WT and $\Delta eIF2A$ cells have similar polysome profiles

Polysome gradient analysis of WT (blue trace) and $\Delta eIF2A$ (red trace) lysate run over a 10-45% sucrose gradient.

MATERIALS AND METHODS

Yeast growth assays

For agar plate assays, cultures were grown overnight to saturation then diluted to OD 0.1 and allowed to grow until log phase. From there, cultures were again diluted to OD 0.1 and diluted 5-fold five times consecutively. 4 mL from each dilution were plated on YPD plates and left to grow overnight. The plate reader assay was performed with a similar set-up, instead of diluting the cells 5-fold, however, they were plated at OD 0.1 in triplicate in a 96-well plate and read for 24-72 hrs in the ClarioStar PLUS (BMG LabTech) plate reader at 30°C.

Construction of yeast strains and plasmids

Deletions of *eIF2A* and *TIF1* were performed by PCR-mediated gene-replacement (Longtine et al. 1998). For $\Delta eIF2A$, primers with 40 bases to the 5' and 3' of *eIF2A* were used to amplify the hygromycin B cassette and transformed into WT W303. $\Delta eIF2A$ transformants were selected by growth on hygromycin B plates and confirmed by gDNA PCR amplification of the cassette with a forward primer in the promoter region of *eIF2A*. Similarly, $\Delta TIF1$ was made by amplifying *HIS3* using primers with 40 bases to the 5' and 3' of *TIF1*. Transformants were selected by growth on synthetic complete (SC) -His plates. Deletions were confirmed by gDNA PCR amplification of the *HIS3* gene with a forward primer in the promoter region of *TIF1*. The $\Delta eIF2A\Delta TIF1$ double mutant was constructed by transforming the *HIS3* gene with complementary regions to the 5' and 3' of *TIF1* in the $\Delta eIF2A$ background.

Polysome profiles and polysome gradient fractionation

Lysate was prepared by growing the indicated strain in YPD overnight and diluting to an OD of 0.1 in 500 mL of YPD. Cultures were grown until mid-log phase, and then harvested by vacuum filtration. The cells were rapidly scraped off the filter and frozen in liquid nitrogen. Cells were lysed with the Retsch

CryoMill in lysis buffer (20 mM Tris pH 8.0, 140 mM KCl, 5 mM MgCl₂, 100 µg/mL cycloheximide, 1 mM DTT, 1% Triton-X 100, 20 U/mL RnaseOUT™, and Halt™ Protease Inhibitor Cocktail). The resulting lysate powder was stored at -80°C while the gradients were prepared. 10-45% continuous sucrose gradients were prepared in 1X polysome buffer (20 mM Tris pH 8.0, 140 mM KCl, 5 mM MgCl₂, 100 µg/mL cycloheximide, 1 mM DTT, and 20 U/mL RnaseOUT™). The gradients were left at 4°C for at least 1 hour.

Lysate was thawed at room temperature and centrifuged at 3,000 x g for 5 min at 4°C to pellet the debris. The supernatant was then centrifuged again at max. speed for 10 minutes at 4°C. The A₂₆₀ of the resulting supernatant was determined and each sample was diluted to 1.0 µg/µL. 200 µL of the diluted lysate was loaded on a sucrose gradient and centrifuged in a Beckman SW40Ti rotor at 36,000 RPM for 2 hours.

Ribosome footprinting and RNA-seq

RNA for RNA-seq and Ribo-seq were isolated from yeast cells grown to mid-log phase and harvested by vacuum filtration. Cells were scraped off the filter and flash frozen in liquid nitrogen. Cells were lysed in 1x Lysis buffer (20 mM Tris pH 8.0, 140 mM KCl, 5 mM MgCl₂, 100 µg/mL cycloheximide, 1% Triton X-100, 1 mM DTT, and 20 U/mL RnaseOUT™) using the Retsch CryoMill. Half of the lysate was used for isolating total RNA and the other half was used for isolating ribosome-protected fragments. Total RNA for RNA-sequencing was isolated using the Direct-zol RNA Purification kit (Zymo). Ribosomal RNA (rRNA) was depleted using the RiboMinus™ Transcriptome Isolation Kit (Invitrogen™). rRNA depletion was confirmed by Bioanalyzer analysis before continuing to the library preparation. The RNA-seq library was generated for each mutant using the NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB) according to the manufacturer's protocol.

For Ribo-seq library preparation RNase I (Ambion) treatment was performed on yeast lysate for 1 hr at RT. The digestion was loaded onto a 10-45% sucrose gradient and centrifuged at 36k RPM for 2 hours to

isolate ribosome protected fragments. Library preparation of the isolated fragments were as previously described (Ingolia et al., 2012) with multiplexing of each sample with a unique barcode:

/5App/NNNNNCGTAAAGATCGGAAGAGCACACGTCTGAA/3ddC/

/5App/NNNNNCTAGAAGATCGGAAGAGCACACGTCTGAA/3ddC/

/5App/NNNNNGATCAAGATCGGAAGAGCACACGTCTGAA/3ddC/

/5App/NNNNNGCATAAGATCGGAAGAGCACACGTCTGAA/3ddC/

50 bp single end read sequencing was performed by the Center for Advanced Technology at UCSF on the HiSeq4000.

Western blot analysis

Yeast lysate were prepared as previously described (Zhang et al., 2011). Lysate was run on an SDS-page gel and transferred to a PVDF membrane using the iBlot2 Western blot transfer system. Membranes were blocked in 5% milk PBST and incubated with α FLAG (Sigma F1804), α HA (Biolegend, 901502), α GCN4 (Novus, C11L34), α Phospho-eIF2 α (Cell Signaling, 9721), or α PGK (Invitrogen, 459250) overnight at 4°C (dilutions and species in Table 2.1). Membranes were washed and incubated with either anti-Rabbit or anti-Mouse IgG (H+L) HRP Conjugate (Promega) for 1 hr at RT. Membranes were then developed with SuperSignal West Dura Extended Duration Substrate (Thermo Fisher) and imaged with the Odyssey Imaging System (LiCor).

Table 2.1: Antibodies for validation of ribosome profiling hits

Target	Cat. #	Dilution	Host
FLAG-M2	Sigma F1804	1:3000	Mouse
3xHA	Biolegend, 901502	1:2000	Mouse
GCN4	Novus, C11L34	1:1000	Rabbit
Phospho- eIF2 α	Cell Signaling, 9721	1:1000	Rabbit
PGK	Invitrogen, 459250	1:2000	Mouse

Purification of 6xHis-eIF2A in *E.coli*

6 liters of BL21 *E.coli* transformed with 6xHis-eIF2A expression plasmid were grown to mid-log phase and induced with 0.5 mM IPTG for 18 hours. Cells were harvested and lysed with the Emulsiflex in lysis buffer (20 mM HEPES pH 7.4, 300 mM KCl, 20 mM imidazole, 10% glycerol, 0.5 mM TCEP, and Halt™ Protease Inhibitor Cocktail). The lysate was centrifuged at 16,000 RPM for 30 min at 4°C and the resulting supernatant was used for the purification steps. Lysate was run over a 5 mL HisTrap column and eluted with a high imidazole buffer. The elution was further purified over a MonoQ column and then a Superdex 200 10/300 GL Cytiva column in 20 mM HEPES pH 7.4, 200 mM KCl, 5% glycerol, 0.5 mM TCEP.

Purification of 6xHis-ybbR-eIF2A and 6xHis-eIF2A-ybbR in Expi293 cells

Expi293 cells were cultured in the Expi293™ Expression Medium and transfected as per the manufacturer's instructions. The 6xHis-ybbR-eIF2A and 6xHis-eIF2A-ybbR plasmids were synthesized by TwistBioscience. 1L of transfected Expi293 cells were harvested and lysed for 30 minutes at 4°C in lysis buffer (20 mM HEPES pH 7.4, 300 mM KCl, 0.1% Triton X-100, 5 mM imidazole, 10% glycerol, and Halt™ Protease Inhibitor Cocktail). The lysed cells were centrifuged at 16,000 RPM for 30 min at

4°C to pellet debris, and the supernatant was incubated with TALON resin for another 20 min at 4°C. The TALON resin was washed three times with wash buffer (20 mM HEPES pH 7.4, 300 mM KCl, 5 mM imidazole, 10% glycerol, and Halt[™] Protease Inhibitor Cocktail). 6xHis-eIF2A was eluted from the resin with 1 mL of elution buffer (20 mM HEPES pH 7.4, 300 mM KCl, 200 mM imidazole, and 10% glycerol). The eluted sample was then injected onto the Superdex 200 10/300 GL Cytiva column and buffer exchanged into 20 mM HEPES pH 7.4, 200 mM KCl, 5% glycerol, and 0.5 mM TCEP.

REFERENCES

- Adams, S. L., Safer, B., Anderson, W. F., & Merrick, W. C. 1975. Eukaryotic initiation complex formation. Evidence for two distinct pathways. *The Journal of Biological Chemistry*, 250(23), 9083–9089.
- Anderson, R., Agarwal, A., Ghosh, A., Guan, B., Casteel, J., Dvorina, N., Baldwin, W. M., Mazumder, B., Nazarko, T. Y., Merrick, W. C., Buchner, D. A., Hatzoglou, M., Kondratov, R. v., & Komar, A. A. 2021. eIF2A-knockout mice reveal decreased life span and metabolic syndrome. *The FASEB Journal*, 35(11). <https://doi.org/10.1096/fj.202101105R>
- Dmitriev, S. E., Terenin, I. M., Andreev, D. E., Ivanov, P. A., Dunaevsky, J. E., Merrick, W. C., & Shatsky, I. N. 2010. GTP-independent tRNA Delivery to the Ribosomal P-site by a Novel Eukaryotic Translation Factor. *Journal of Biological Chemistry*, 285(35), 26779–26787. <https://doi.org/10.1074/jbc.M110.119693>
- Golovko, A., Kojukhov, A., Guan, B.-J., Morpurgo, B., Merrick, W. C., Mazumder, B., Hatzoglou, M., & Komar, A. A. 2016. The eIF2A knockout mouse. *Cell Cycle*, 15(22), 3115–3120. <https://doi.org/10.1080/15384101.2016.1237324>
- González-Almela, E., Williams, H., Sanz, M. A., & Carrasco, L. 2018. The Initiation Factors eIF2, eIF2A, eIF2D, eIF4A, and eIF4G Are Not Involved in Translation Driven by Hepatitis C Virus IRES in Human Cells. *Frontiers in Microbiology*, 9. <https://doi.org/10.3389/fmicb.2018.00207>

- Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M., & Weissman, J. S. 2012. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature Protocols*, 7(8), 1534–1550. <https://doi.org/10.1038/nprot.2012.086>
- Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S., & Weissman, J. S. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science (New York, N.Y.)*, 324(5924), 218–223. <https://doi.org/10.1126/science.1168978>
- Kapp, L. D., & Lorsch, J. R. 2004. GTP-dependent Recognition of the Methionine Moiety on Initiator tRNA by Translation Factor eIF2. *Journal of Molecular Biology*, 335(4), 923–936. <https://doi.org/10.1016/j.jmb.2003.11.025>
- Kertesz, M., Wan, Y., Mazor, E., Rinn, J. L., Nutter, R. C., Chang, H. Y., & Segal, E. 2010. Genome-wide measurement of RNA secondary structure in yeast. *Nature*, 467(7311), 103–107. <https://doi.org/10.1038/nature09322>
- Kim, J. H., Park, S. M., Park, J. H., Keum, S. J., & Jang, S. K. 2011. eIF2A mediates translation of hepatitis C viral mRNA under stress conditions. *The EMBO Journal*, 30(12), 2454–2464. <https://doi.org/10.1038/emboj.2011.146>
- Komar, A. A., Gross, S. R., Barth-Baus, D., Strachan, R., Hensold, J. O., Goss Kinzy, T., & Merrick, W. C. 2005. Novel Characteristics of the Biological Properties of the Yeast *Saccharomyces cerevisiae* Eukaryotic Initiation Factor 2A. *Journal of Biological Chemistry*, 280(16), 15601–15611. <https://doi.org/10.1074/jbc.M413728200>

- Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 44(2), 283–292. [https://doi.org/10.1016/0092-8674\(86\)90762-2](https://doi.org/10.1016/0092-8674(86)90762-2)
- Kozak, M. 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *Journal of Molecular Biology*, 196(4), 947–950. [https://doi.org/10.1016/0022-2836\(87\)90418-9](https://doi.org/10.1016/0022-2836(87)90418-9)
- Liang, H., He, S., Yang, J., Jia, X., Wang, P., Chen, X., Zhang, Z., Zou, X., McNutt, M. A., Shen, W. H., & Yin, Y. 2014. PTEN α , a PTEN isoform translated through alternative initiation, regulates mitochondrial function and energy metabolism. *Cell Metabolism*, 19(5), 836–848. <https://doi.org/10.1016/j.cmet.2014.03.023>
- Longtine, M. S., McKenzie, A., Demarini, D. J., Shah, N. G., Wach, A., Brachat, A., Philippsen, P., & Pringle, J. R. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast (Chichester, England)*, 14(10), 953–961. [https://doi.org/10.1002/\(SICI\)1097-0061\(199807\)14:10<953::AID-YEA293>3.0.CO;2-U](https://doi.org/10.1002/(SICI)1097-0061(199807)14:10<953::AID-YEA293>3.0.CO;2-U)
- Lowe, D. D., & Montell, D. J. 2022. Unconventional translation initiation factor EIF2A is required for *Drosophila* spermatogenesis. *Developmental Dynamics*, 251(2), 377–389. <https://doi.org/10.1002/dvdy.403>
- Maag, D., Fekete, C. A., Gryczynski, Z., & Lorsch, J. R. 2005. A Conformational Change in the Eukaryotic Translation Preinitiation Complex and Release of eIF1 Signal Recognition of the Start Codon. *Molecular Cell*, 17(2), 265–275. <https://doi.org/10.1016/j.molcel.2004.11.051>

Nguyen, H. G., Conn, C. S., Kye, Y., Xue, L., Forester, C. M., Cowan, J. E., Hsieh, A. C., Cunningham, J. T., Truillet, C., Tameire, F., Evans, M. J., Evans, C. P., Yang, J. C., Hann, B., Koumenis, C., Walter, P., Carroll, P. R., & Ruggero, D. 2018. Development of a stress response therapy targeting aggressive prostate cancer. *Science Translational Medicine*, *10*(439).
<https://doi.org/10.1126/scitranslmed.aar2036>

Robert, F., Kapp, L. D., Khan, S. N., Acker, M. G., Kolitz, S., Kazemi, S., Kaufman, R. J., Merrick, W. C., Koromilas, A. E., Lorsch, J. R., & Pelletier, J. 2006. Initiation of Protein Synthesis by Hepatitis C Virus Is Refractory to Reduced eIF2 · GTP · Met-tRNA_i^{Met} Ternary Complex Availability. *Molecular Biology of the Cell*, *17*(11), 4632–4644. <https://doi.org/10.1091/mbc.e06-06-0478>

Sendoel, A., Dunn, J. G., Rodriguez, E. H., Naik, S., Gomez, N. C., Hurwitz, B., Levorse, J., Dill, B. D., Schramek, D., Molina, H., Weissman, J. S., & Fuchs, E. 2017. Translation from unconventional 5' start sites drives tumour initiation. *Nature*, *541*(7638), 494–499. <https://doi.org/10.1038/nature21036>

Starck, S. R., Jiang, V., Pavon-Eternod, M., Prasad, S., McCarthy, B., Pan, T., & Shastri, N. 2012. Leucine-tRNA Initiates at CUG Start Codons for Protein Synthesis and Presentation by MHC Class I. *Science*, *336*(6089), 1719–1723. <https://doi.org/10.1126/science.1220270>

Starck, S. R., Tsai, J. C., Chen, K., Shodiya, M., Wang, L., Yahiro, K., Martins-Green, M., Shastri, N., & Walter, P. 2016. Translation from the 5' untranslated region shapes the integrated stress response. *Science*, *351*(6272). <https://doi.org/10.1126/science.aad3867>

Terenin, I. M., Dmitriev, S. E., Andreev, D. E., & Shatsky, I. N. 2008. Eukaryotic translation initiation machinery can operate in a bacterial-like mode without eIF2. *Nature Structural & Molecular Biology*, *15*(8), 836–841. <https://doi.org/10.1038/nsmb.1445>

Wilmes, G. M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S. R., Whitworth, G. B., Kress, T. L., Weissman, J. S., Ideker, T., Guthrie, C., & Krogan, N. J. 2008. A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. *Molecular Cell*, 32(5), 735–746.
<https://doi.org/10.1016/j.molcel.2008.11.012>

Yin, J., Lin, A. J., Golan, D. E., & Walsh, C. T. 2006. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nature Protocols*, 1(1), 280–285.
<https://doi.org/10.1038/nprot.2006.43>

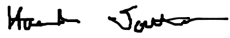
Zhang, T., Lei, J., Yang, H., Xu, K., Wang, R., & Zhang, Z. 2011. An improved method for whole protein extraction from yeast *Saccharomyces cerevisiae*. *Yeast*, 28(11), 795–798.
<https://doi.org/10.1002/yea.1905>

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:



7C567822E2D84F1...

Author Signature

3/10/2022

Date