

Conserved structural elements specialize ATAD1 as a membrane protein extraction machine

Lan Wang^{1,2,*,#}, Hannah Toutkoushian^{1,2,*}, Vladislav Belyy^{1,2}, Claire Kokontis¹, Peter Walter^{1,2,#}

Affiliations

¹Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, CA, USA.

²Howard Hughes Medical Institute, University of California at San Francisco, San Francisco, CA, USA.

*These authors contributed equally

#To whom correspondence should be addressed; Email: lan@walterlab.ucsf.edu;
peter@walterlab.ucsf.edu;

Subject Areas: Structural biology, cell biology

Abstract

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. We used the yeast growth assay to test the function of key amino acids within these structural elements. 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. ATAD1's distinguishing features include an abundance of aromatic amino acids that line the central pore proposed to give ATAD1 a firm grip on its substrate and a short α -helix at the C-terminus that strongly facilitates ATAD1's oligomerization. We also developed a microscopy-based assay reporting on protein mislocalization, with which we directly assessed ATAD1's activity in live cells.

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, Figure1-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-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-figure supplement 3B, Table 1), an arrangement resembling the “open state” structure of Msp1 (ref), 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 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 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-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 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 (Figure1-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-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-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 – 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 1C and Figure 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 2 – figure supplement 1). Zooming in on α 11 reveals that it lies at the interface between two adjacent subunits (Figure 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 4) adjacent subunit, bridging between the two (Figure 2B and D). In α 11_{M4}, for example, L336, A339 and I340 pack against the core of M4 (Figure 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 2D). A few charged amino acids (R254, K342, K344, K345, K347, E341 and D348) point into the cytosol (Figure 2C and E).

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, α 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, α 12, that reaches across the intersubunit interface to contact the *clockwise* adjacent subunit (Figure 2-figure supplement 2). α 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 3A). To analyze the activity of ATAD1 mutants in the absence of the wild-type enzyme, we deleted both alleles of *ATAD1* in HeLa cells using CRISPR/Cas9 (Figure 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 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 3B, middle row) and observed only the Golgi-localized EGFP-Gos28, 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 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 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 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 3C, Figure 3 – source data 1). In brief, we first identified cells expressing ATAD1-HaloTag.

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 3D). As shown in Figure 3D, expression of an empty vector or ATAD1^{E193Q} in the ATAD1^{-/-} reporter cell line led to a PCC of 0.31 ± 0.035 and 0.35 ± 0.036 , respectively. By contrast, transiently expressing wild-type ATAD1 showed a significantly lower PCC of 0.15 ± 0.024 ($p < 0.0001$), 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 1C), and another in which we kept $\alpha 11$ and removed everything C-terminal to it (ATAD1₁₋₃₅₁, Figure 1C). We first expressed these variants in the ATAD1^{-/-} cells. As shown in Figure 4A, EGFP-Gos28 showed significant mislocalization to mitochondria in cells expressing ATAD1 $\Delta\alpha 11$ but not ATAD1₁₋₃₅₁, suggesting that the $\alpha 11$ 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 4B), indicating its recessive phenotype, i.e., its lack of incorporation into and inactivating the wild-type hexamers present in these cells. While on average, expression of ATAD1 $\Delta\alpha 11$ did not induce substrate mislocalization, it was evident by eye that a small population of wild-type cells expressing ATAD1 $\Delta\alpha 11$ showed EGFP-Gos28 mislocalization (Figure 4- figure supplement 1A). 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 4-figure supplement 1B), suggesting that ATAD1 $\Delta\alpha$ 11 incorporated into wild-type hexamers and inactivated them. 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 4-figure supplement 2), 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 Δ 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 α 11 on ATAD1's substrate binding affinity. As expected and consistent with its inability to form hexamers, Δ 40-ATAD1 $\Delta\alpha$ 11 also showed substantially reduced peptide binding affinity (Figure 4C).

Several disease-related mutations are close to the α 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 4 – figure supplement 3). 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 2 – figure supplement 1). The fact that it lies at the oligomerization interface and also contacts pore-loop 3 prompted us to ask whether it could functionally couple oligomer formation 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 4 -figure supplement 3, $\Delta 40$ -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 4 – figure supplement 4), suggesting that $\alpha 11$ is required to couple subunit oligomerization with substrate binding. 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 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 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 1E). By contrast, in most other known AAA proteins, the X_1 position within pore-loop 1 is an aromatic amino acid, and the X_2 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 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 Y167 to an

aliphatic amino acid in Y167V or Y167L mutants rendered ATAD1 inactive (Figure 5A, figure supplement 1), 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*. As shown in Figure 5B, ATAD1^{W166A}, ATAD1^{Y166A} and ATAD1^{Y167V} showed dramatically reduced (estimated > 100-fold) binding to P13 compared to wild-type ATAD1. ATAD1^{Y167F} bound P13 with a ~10-fold weaker affinity than the wild-type enzyme, but still better than the other pore-loop mutants. We reasoned that this difference resulted because the *in vitro* assay is more sensitive to small changes in activity that cell-based assays fail to capture. 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.

Table 1 Cryo-EM data analysis

Structure	$\Delta 40$ -ATAD1 ^{E193Q} (closed)	$\Delta 40$ -ATAD1 ^{E193Q} (open)
Data collection		
Microscope	Titan Krios	
Voltage (keV)	300	
Nominal magnification	105000x	
Exposure navigation	Image shift	
Electron dose (e ⁻ Å ⁻²)	67	
Dose rate (e ⁻ /pixel/sec)	15	
Detector	K3 summit	
Pixel size (Å)	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 (Å)	3.2	3.5
FSC average resolution, unmasked (Å)	4.1	4.6
Applied B-factor (Å)	121.9	119.2
Reconstruction package	Cryosparc v2.15	
Refinement		
Protein residues	1676	1694
Ligands	17	19
RMSD Bond lengths (Å)	0.003	0.017
RMSD Bond angles (°)	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	

Figures

Figure 1

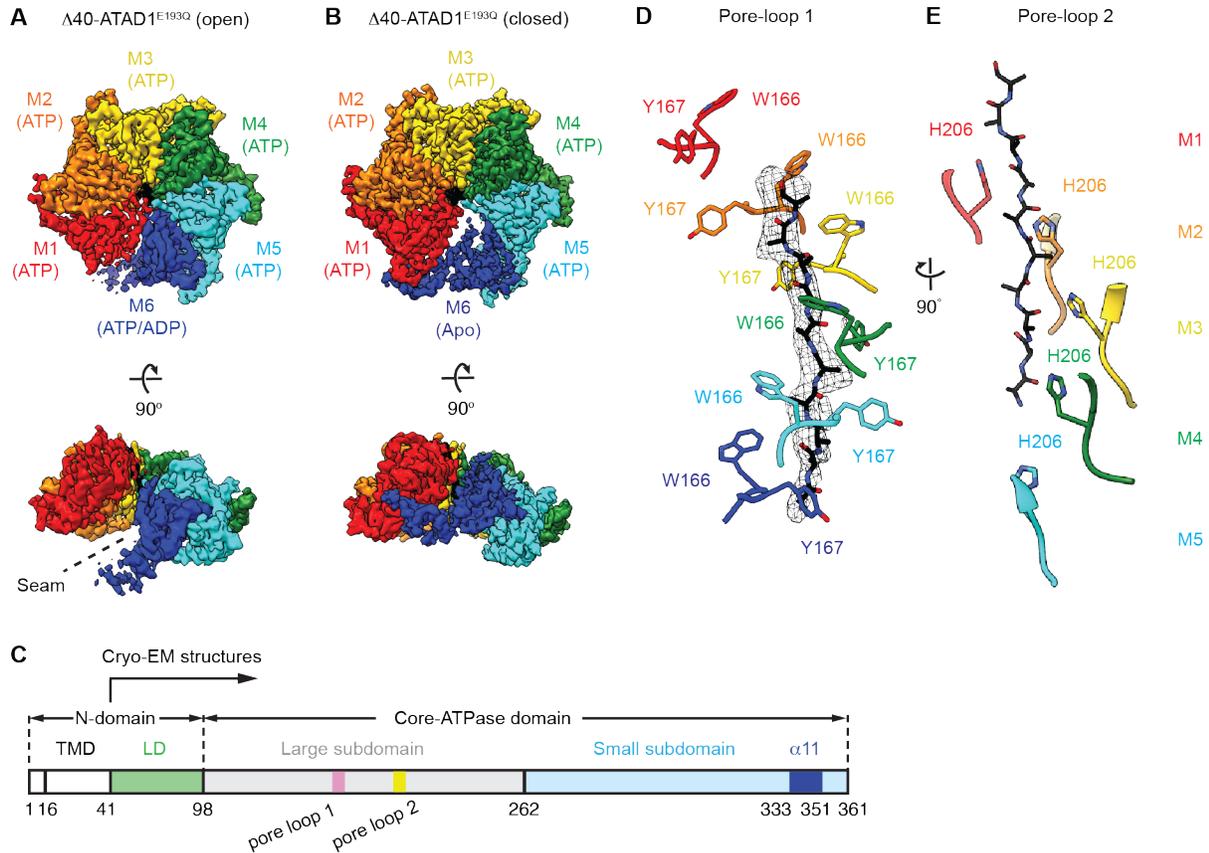


Figure 1 Architecture of the $\Delta 40$ -ATAD1^{E193Q}-substrate complexes

(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. **(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 2

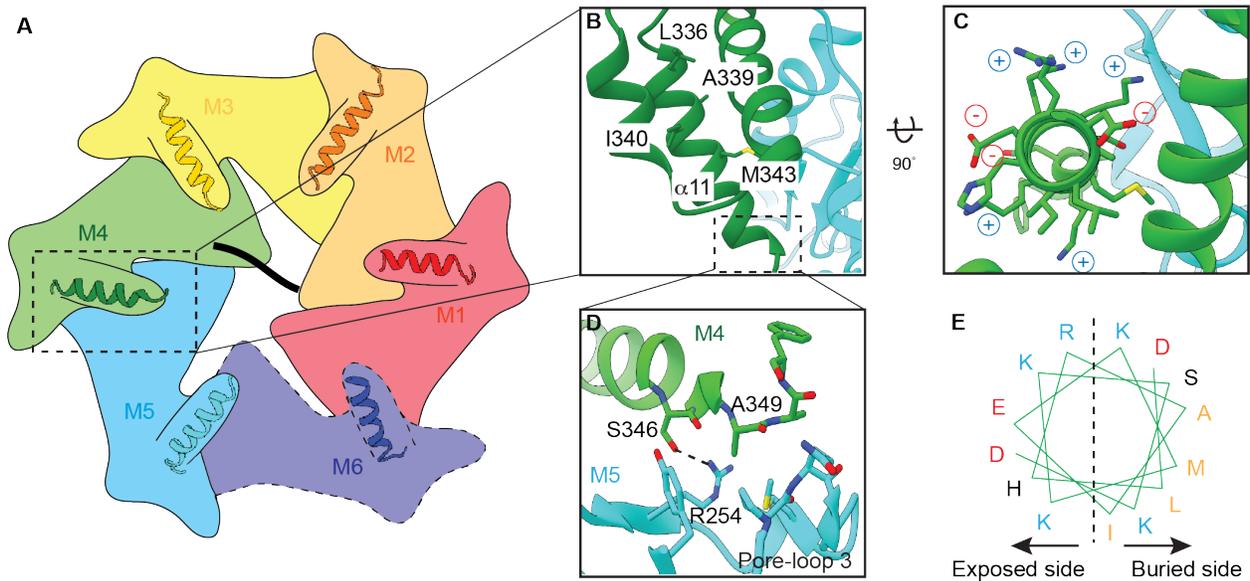


Figure 2 Helix $\alpha 11$ resides at the intersubunit interface of ATAD1

(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. **(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 3

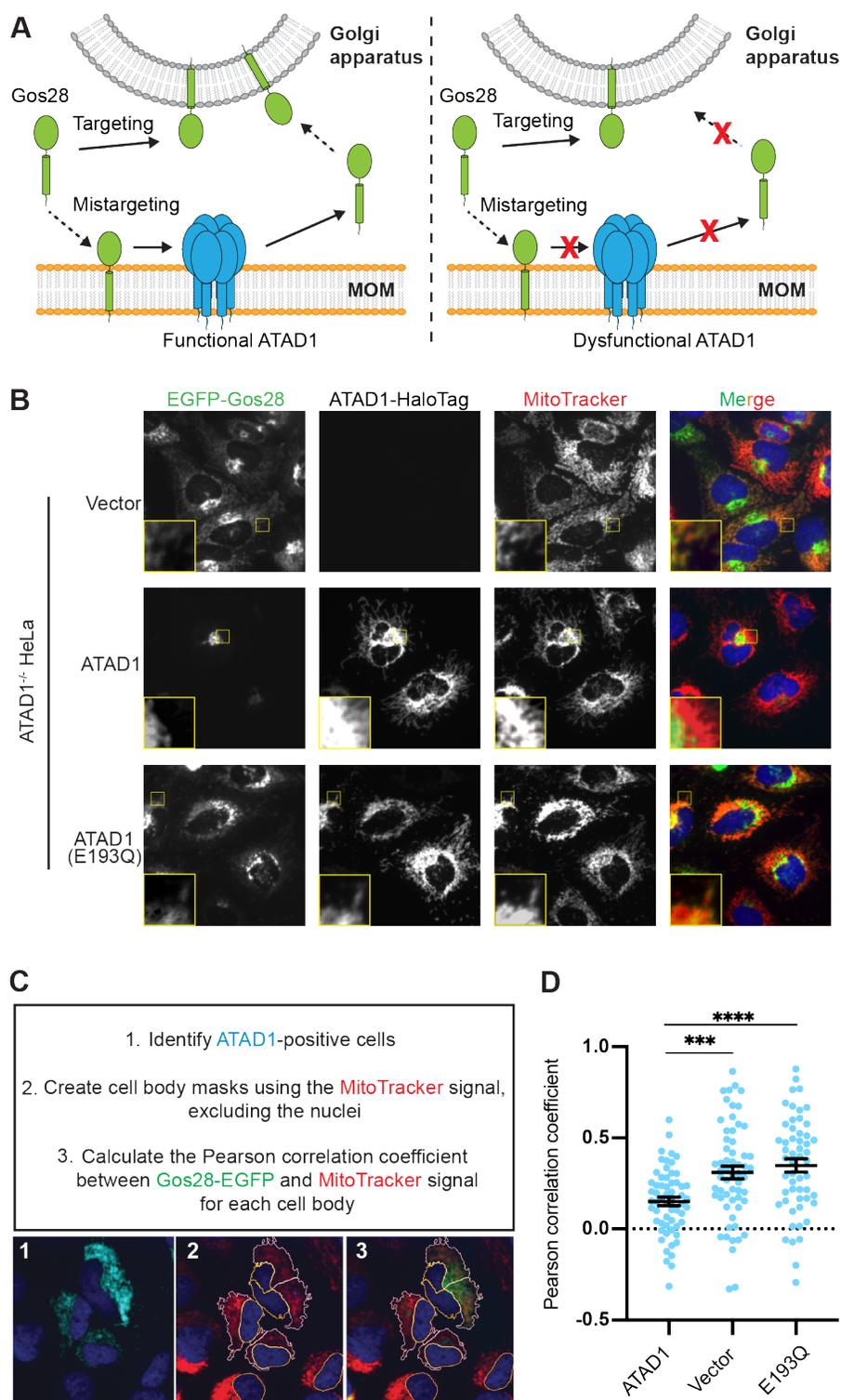


Figure 3 A microscopy assay allows for the direct measurement and quantification of ATAD1's activity

(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 Hoescht-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. *** p< 0.001, **** p<0.0001.

Figure 4

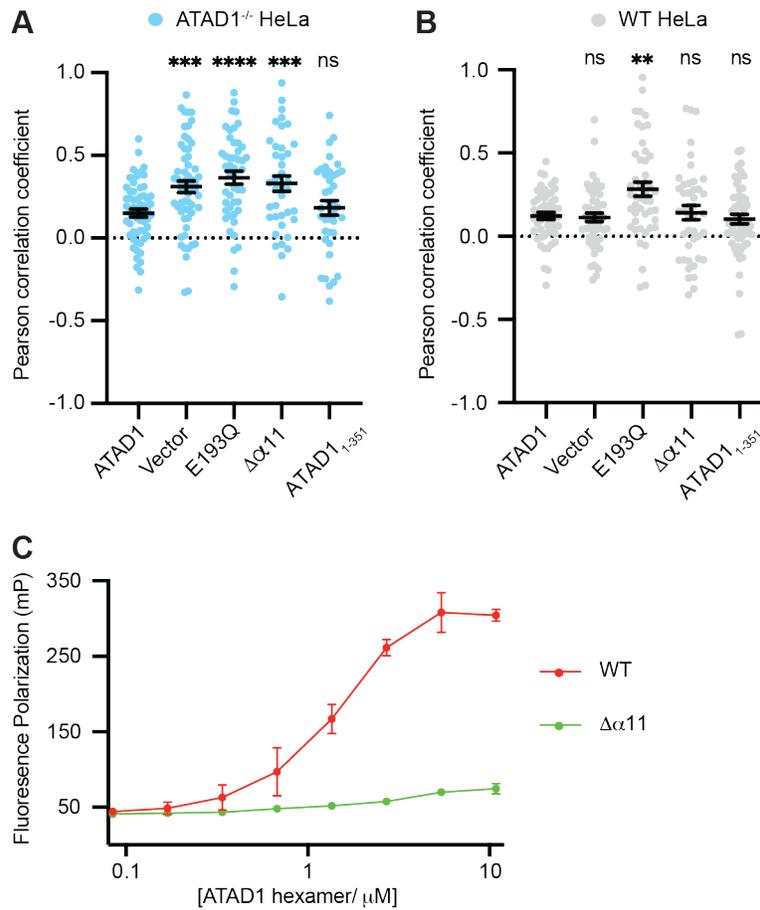


Figure 4 Helix $\alpha11$ mediates hexamer assembly

(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 $\Delta\alpha11$ 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 $\Delta\alpha11$ mutant. ** p < 0.01, *** p < 0.001, **** p < 0.0001. **(C)** Fluorescence polarization assay showing the different peptide binding abilities of the $\Delta40$ -ATAD1 (WT) versus $\Delta40$ -ATAD1 $\Delta\alpha11$ ($\Delta\alpha11$). ATAD1 concentrations are expressed as hexamer concentrations. Each dot represents an average of two replicates and the error bar represents SEM.

Figure 5

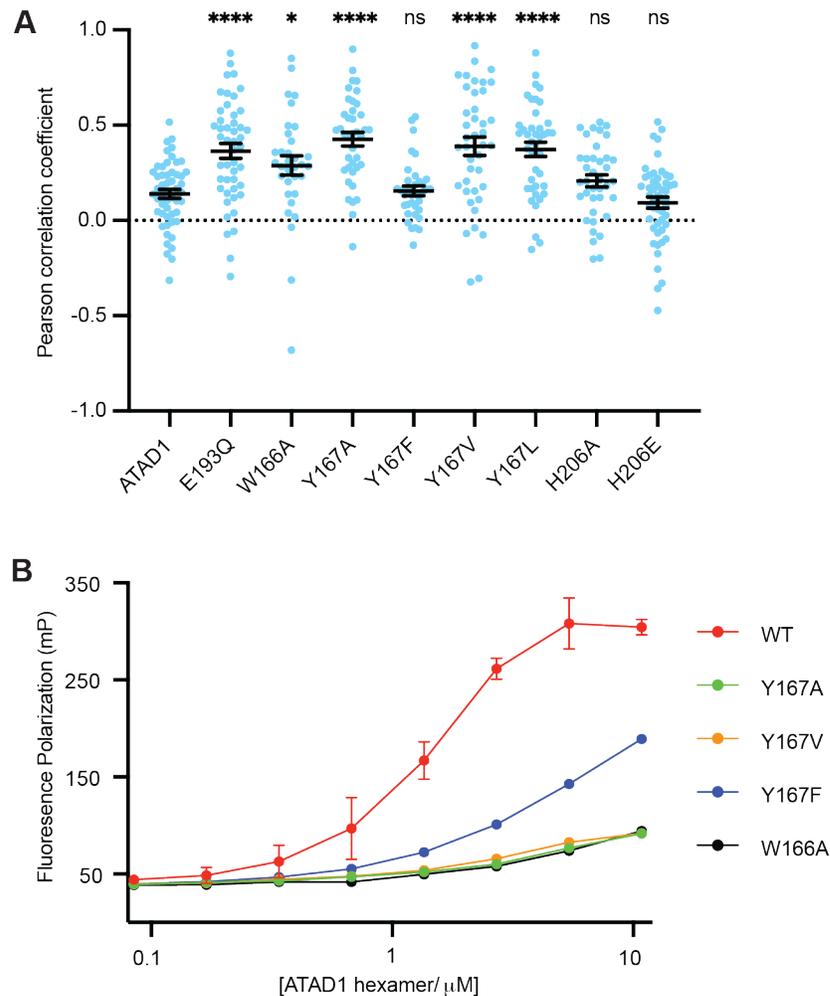


Figure 5 Pore-loop 1 aromatic amino acids are important for ATAD1's function both *in vivo* and *in vitro*

(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. * $p < 0.05$, **** $p < 0.0001$. **(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.

Figure 1-figure supplement 1

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 – 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 α 11, the other members of the AAA_{MC} subfamily have two shorter helices (α 11 and α 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- figure supplement 2

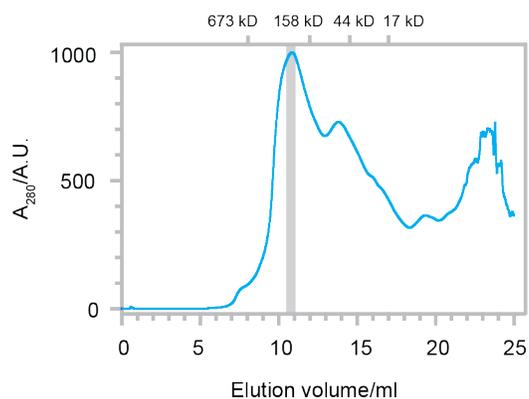


Figure 1 – Figure supplement 2 Size exclusion trace of $\Delta 40$ -ATAD1^{E193Q}

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 - Figure supplement 3

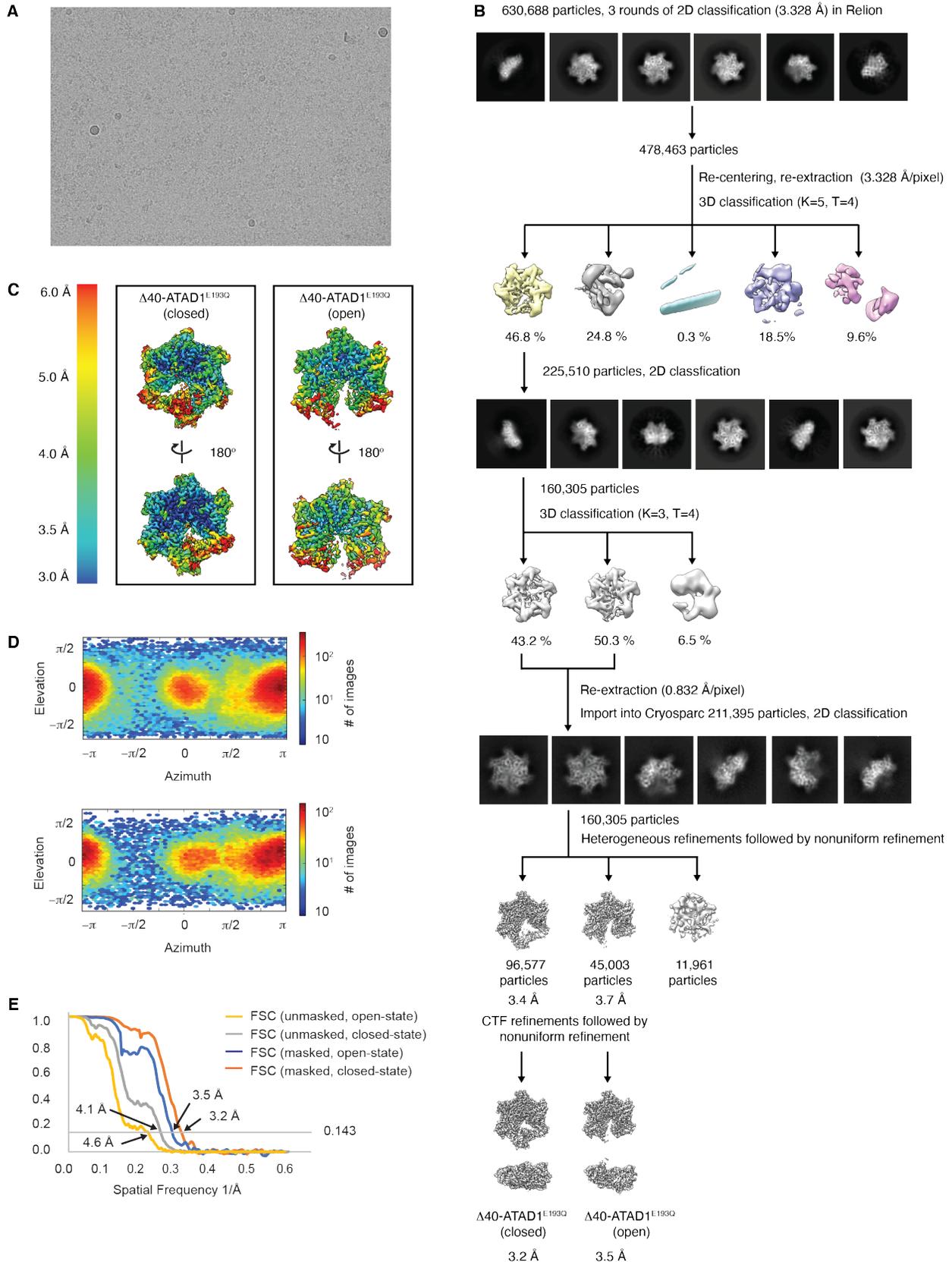


Figure 1 – Figure supplement 3 Cryo-EM processing of the $\Delta 40$ -ATAD1^{E193Q}-substrate complexes

(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-supplement 4

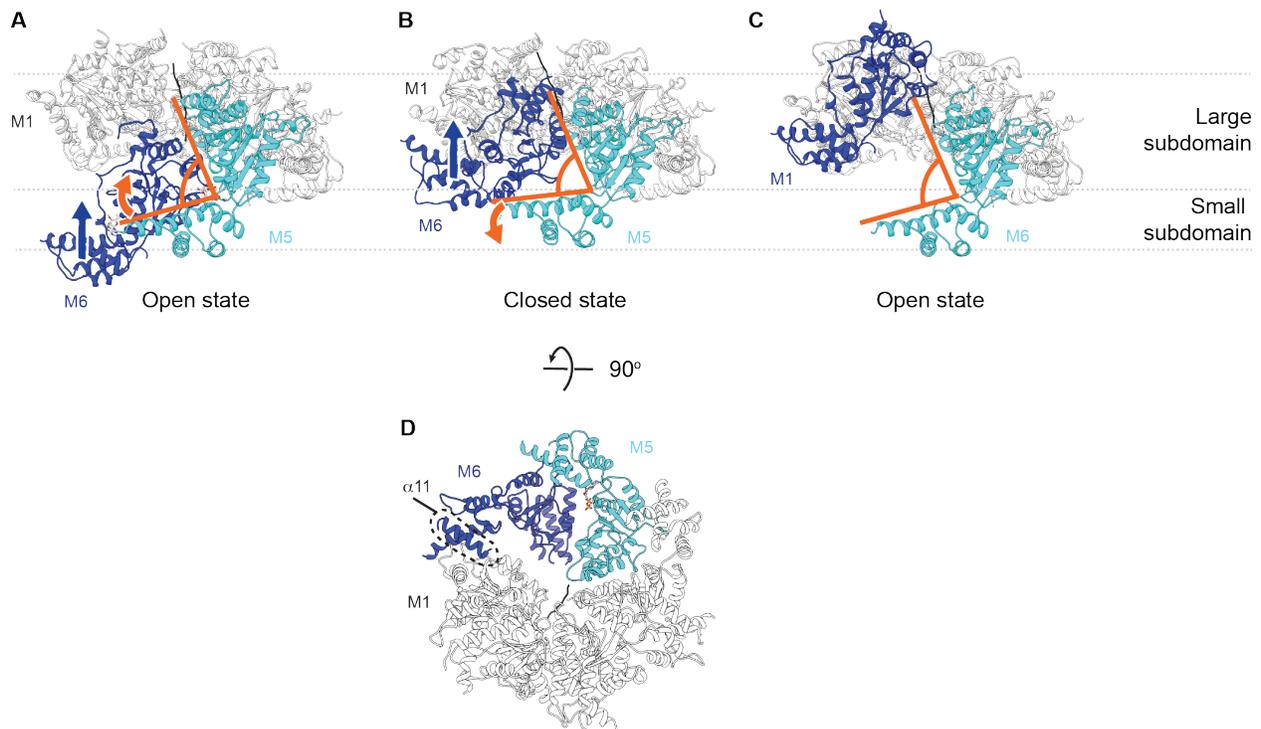


Figure 1 – Figure supplement 4 Hinge motion between the large and the small subdomain accompanies subunit movements

(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 2 -supplement 1

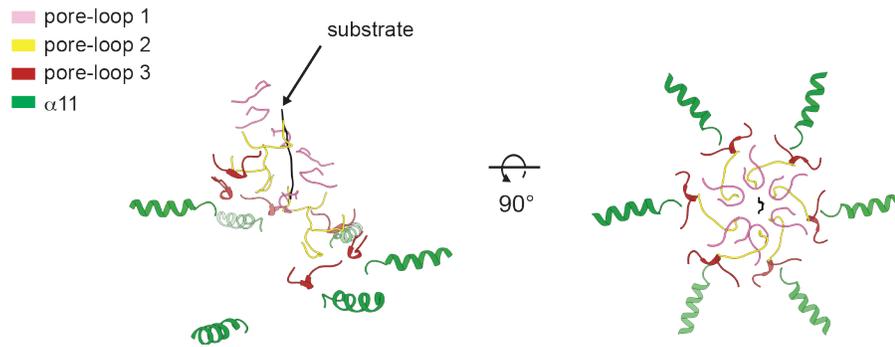


Figure 2 – Figure supplement 1 Helices $\alpha 11$ form an additional spiral staircase beneath pore-loop3

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 2 - figure supplement 2

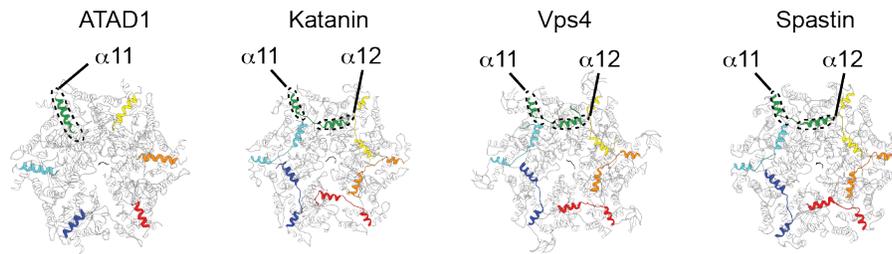


Figure 2 – Figure supplement 2 ATAD1 possesses a unique helix at the C-terminus that is structured differently from those in other AAA_{MC} proteins

Structures of different AAA_{MC} proteins shown in cartoon representation. The 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. 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.

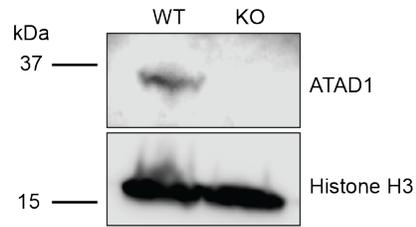


Figure 3 – Figure supplement 1 Verification of ATAD1 knockout by Western blot

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 3 - figure supplement 2

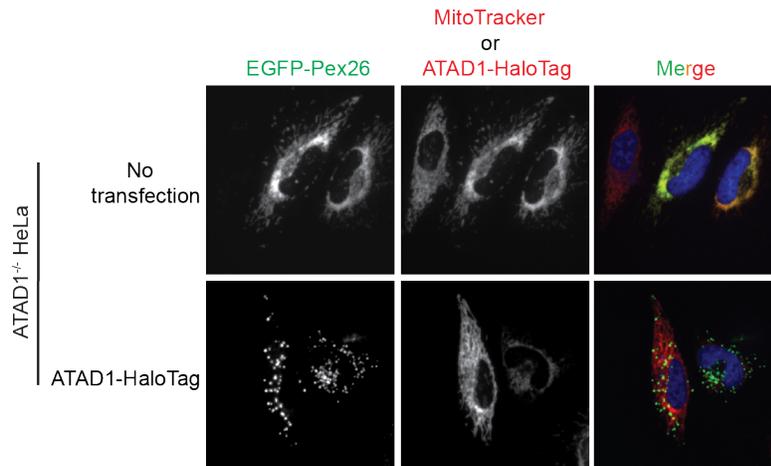


Figure 3 – Figure supplement 2 Live-cell imaging showing the ATAD1 dependent localization of EGFP-Pex26

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 Hoescht-stained nuclei in blue.

Figure 3 - figure supplement 3

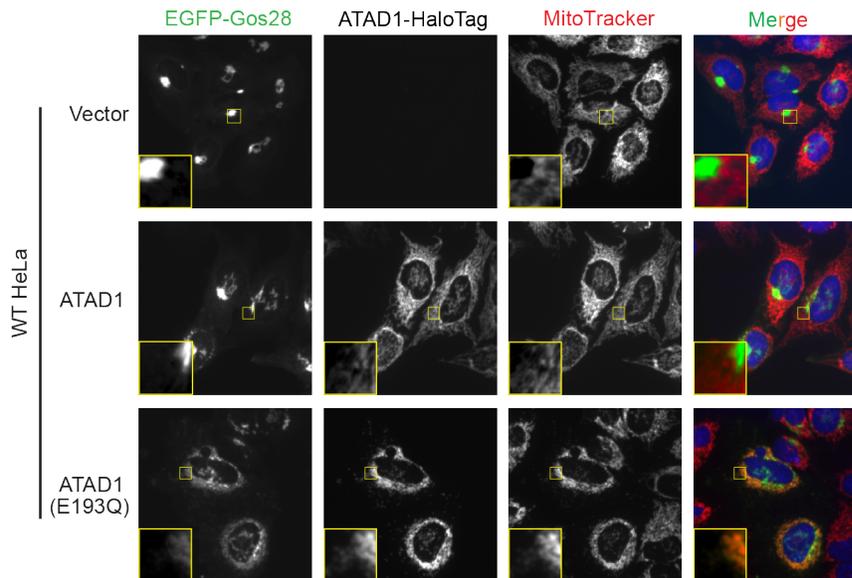


Figure3 – Figure supplement 3 Live-cell imaging showing the ATAD1 dependent localization of EGFP-Gos28 in WT HeLa cells

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 Hoescht-stained nuclei in blue. Insets are included to better show the absence or presence of co-localization between EGFP-Gos28 and the mitochondria.

Figure 4 - figure supplement 1

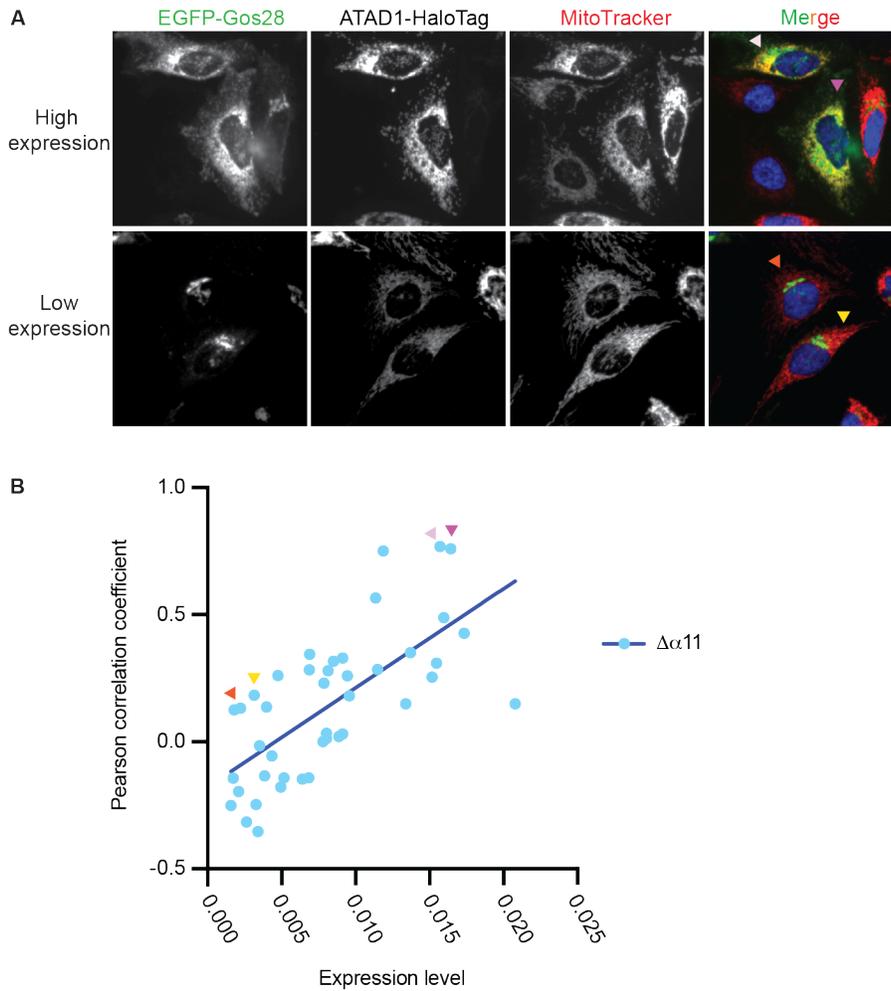


Figure 4 – Figure supplement 1 ATAD1 $\Delta\alpha 11$ expression level impacts substrate mislocalization in WT HeLa cells.

(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 performed and there was a significant positive correlation between the two variables ($p < 0.0001$).

Figure 4-figure supplement 2

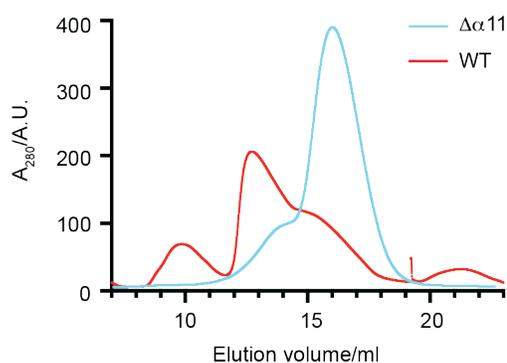


Figure 4 – Figure supplement 2 Size exclusion chromatography of $\Delta40$ -ATAD1 and $\Delta40$ -ATAD1 $\Delta\alpha11$

The SEC traces of $\Delta40$ -ATAD1 (WT) and $\Delta40$ -ATAD1 $\Delta\alpha11$ ($\Delta\alpha11$) show that the former forms oligomeric species on the SEC column while the latter is predominantly a monomer. The SEC trace of $\Delta40$ -ATAD1 is shown in a red line, and the SEC trace of $\Delta40$ -ATAD1 $\Delta\alpha11$ in a blue curve.

Figure 4 - figure supplement 3

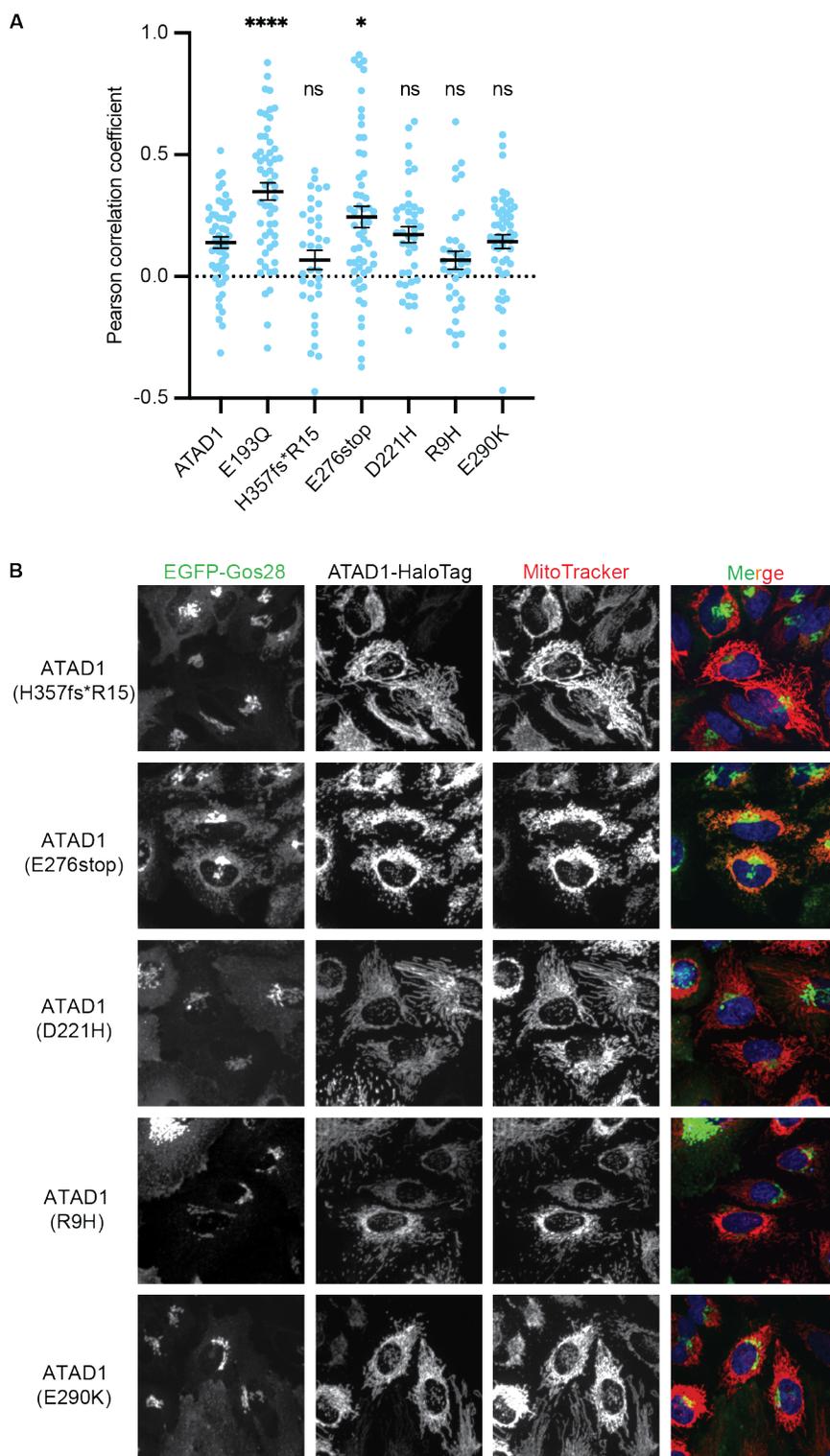


Figure 4 – Figure supplement 3 Testing the effect of the disease-relevant mutations of ATAD1

(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. **** $p < 0.0001$, * $p < 0.05$. **(B)** 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 Hoescht-stained nuclei in blue.

Figure 4- Figure supplement 4

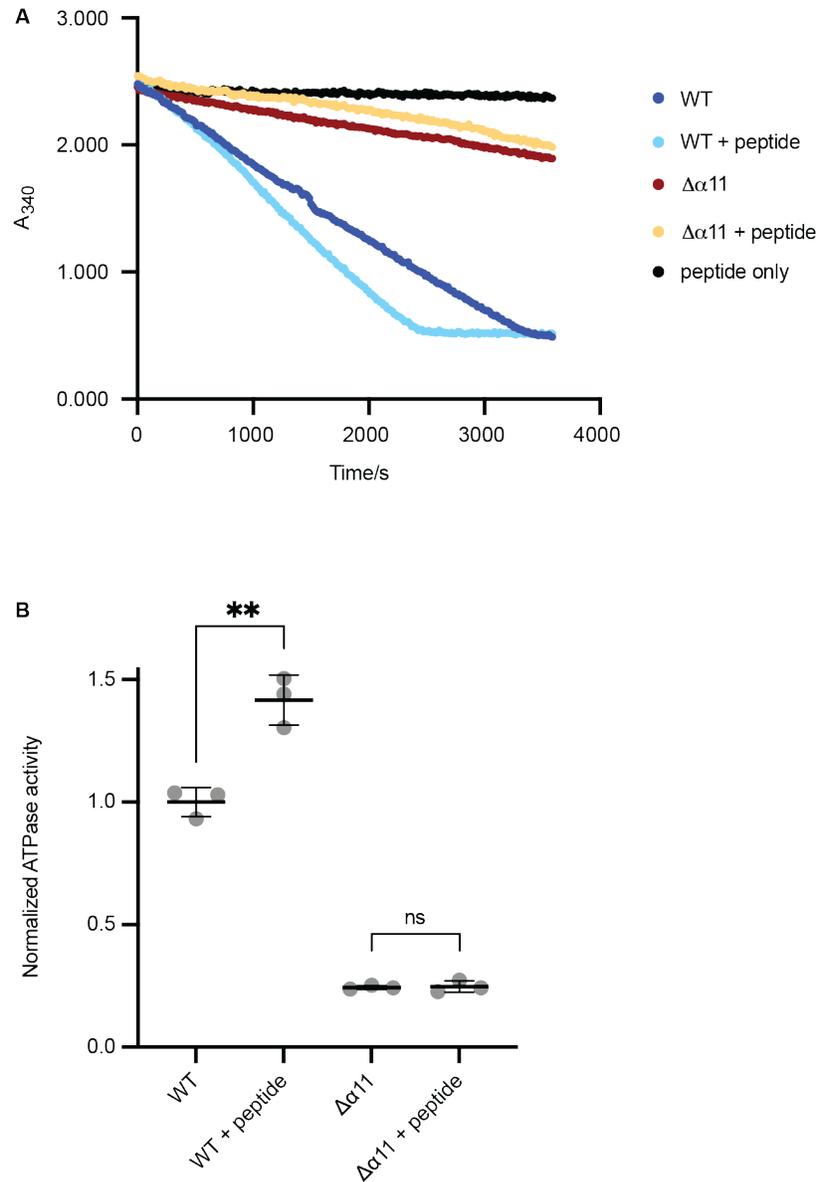


Figure 4 – Figure supplement 4 ATPase assay of $\Delta 40$ -ATAD1 and $\Delta 40$ -ATAD1 $\Delta\alpha 11$

(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 5 - supplement 1

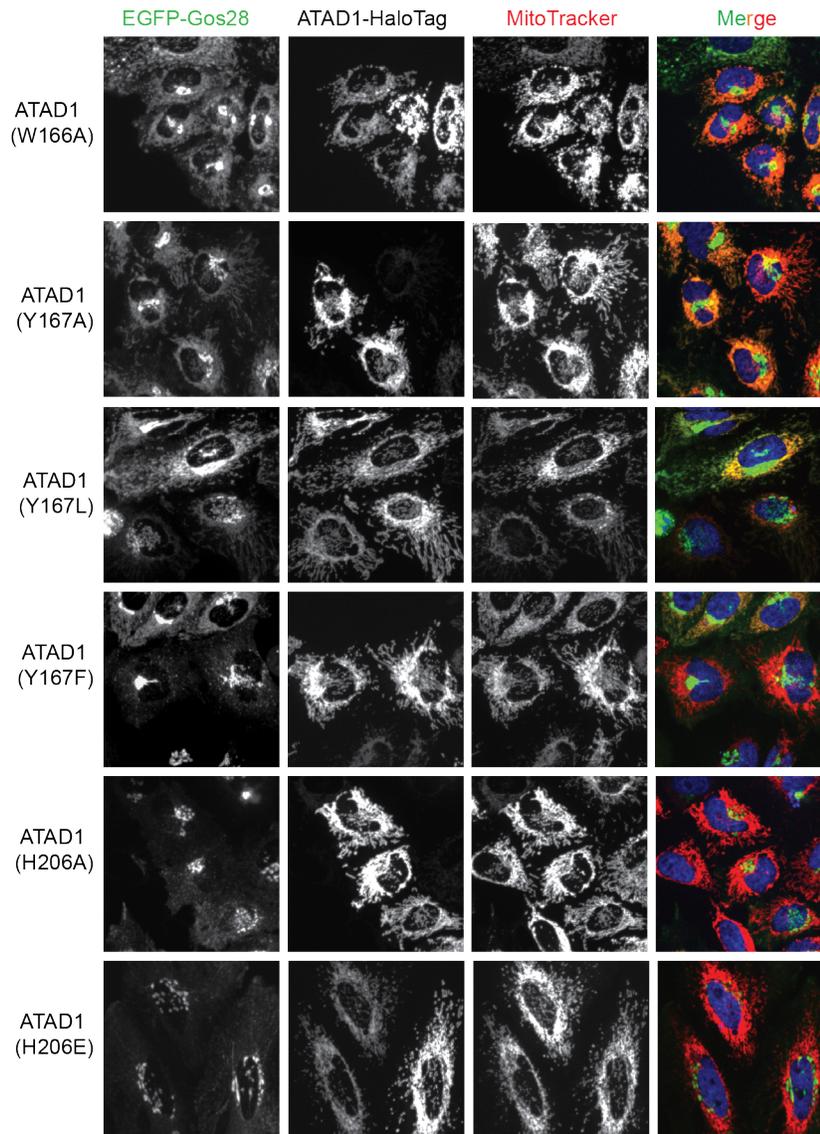


Figure 5- Figure supplement 1 Live-cell imaging showing the pore-loop dependent localization of EGFP-Gos28 in ATAD1^{-/-} HeLa cells

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.

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 $e^-/\text{\AA}^2$, fractioned over 100 frames during a 10 s exposure. Data was collected at 105,000 x nominal magnification (0.832 $\text{\AA}/\text{pixel}$ 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 \AA 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 $\text{\AA}/\text{pixel}$) 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_γS 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 μ M 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. The data analysis pipeline is described in the Figure 3 - source data 1.

Acknowledgments

We thank the Walter lab for helpful discussions throughout the course of this project; H Zhou and Z Yu of the Cryo-EM facility at the HHMI Janelia Research Campus. We thank the QB3 shared cluster for computational support. We also thank Hehua and Obie for their constant support throughout this project.

Funding

This work was supported by the National Institutes of Health (R01GM032384). Lan Wang is a Damon Runyon Cancer Research Foundation fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2312-17), Vladislav Belyy V.B. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2284-17), and PW is an Investigator of the Howard Hughes Medical Institute.

Author Contributions

Conception and design: L Wang, H Toutkoushian and P Walter. Analysis and interpretation of data: L Wang, H Toutkoushian, V Belyy, C Kokontis and P Walter. Acquisition of data: L Wang,

H Toutkoushian, C Kokontis. Writing (original draft): L Wang, H Toutkoushian and P Walter.

Writing (review and editing): L Wang, H Toutkoushian, V Belyy, C Kokontis and P Walter.

Competing Interests

The authors declare no competing interests.

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/embj.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/elife.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.ac119.009890
- 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/elife.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 N Y* 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:eaa0464. 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:eaa4985. 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:eaan4146. 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