2019

Chemical Biology, Biochemical and Structural Studies of MDN1, an AAA Protein Required for Ribosome Biogenesis

Zhen Chen

CHEMICAL BIOLOGY, BIOCHEMICAL AND STRUCTURAL STUDIES OF MDN1, AN AAA PROTEIN REQUIRED FOR RIBOSOME BIOGENESIS

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Zhen Chen

June 2019

CHEMICAL BIOLOGY, BIOCHEMICAL AND STRUCTURAL STUDIES OF MDN1, AN AAA PROTEIN REQUIRED FOR RIBOSOME BIOGENESIS

Zhen Chen, Ph.D.

The Rockefeller University 2019

Cellular proteins are synthesized by ribosomes, which are ~3 MDa macromolecular complexes comprised of four ribosomal RNAs and ~80 ribosomal proteins in yeast. The biogenesis of such complicated ribonucleoprotein complexes is a highly regulated, multistep process requiring a plethora of more than 200 unique assembly factors. Energy-harnessing enzymes, such as ATPases and GTPases, are needed to remodel the precursors of ribosomes at fast time scales. Mdn1 is an essential dynein-like AAA protein (ATPases Associated with various Activities) that releases specific assembly factors from the precursors of 60S subunit of ribosomes. However, Mdn1's unusually large size (~5000 amino acids in a single polypeptide) and the transient nature of intermediates of ribosome biogenesis have limited our understanding how Mdn1 remodels pre-60S particles. In addition, the limited homology of Mdn1 to other well-studied proteins, including dyneins, has restricted our understanding of its function. Here, I first combined chemical and biochemical approaches to develop and validate ribozinoindoles (Rbins) as the cell-permeable inhibitors of Mdn1, which are the first potent and selective inhibitors of ribosome biogenesis in eukaryotes. These compounds can be further used to dissect the dynamic functions of Mdn1 during the multistep process of ribosome biogenesis. In addition, I solved three cryo-EM structures

of both full-length and truncated Mdn1 (resolution up to 4.0 Å) that provided the first pseudo-atomic models for Mdn1 in two distinct nucleotide states. Remarkably, Mdn1's the C-terminal MIDAS domain (Metal Ion-Dependent Adhesion Site), which interacts with other ribosome assembly factors, docks onto the N-terminal AAA ring in a nucleotide state-specific manner, even though they are separated by more than 2000 aa. These data suggest that conformational changes in the AAA ring can be directly transmitted to the MIDAS domain, thereby driving the selective release of the MIDAS-bound assembly factors from the precursors of 60S subunit of ribosomes. Together, these chemical biology, biochemical and structural studies of Mdn1 reveal how an AAA protein can contribute to the dynamic ribosome biogenesis process in eukaryotes.

To my parents and sister.

Acknowledgements

First and foremost, I want to thank my mentor Dr. Tarun Kapoor for his guidance throughout my graduate study. Dr. Kapoor trained me to analyze scientific data critically and present my work effectively. His critical suggestions have inspired me to reflect on experimental designs and results from a new perspective and these reflections become an invaluable part of my learning experience. I am truly indebted to the discussions that refined my understanding of scientific data and changed my way of thinking. In addition, his generous support and encouragement, especially when various technical difficulties haunted the progress of the projects, are essential for me to move forward. The process of trying to meet his high standards has driven my growth as a scientist. What I learned from him will surely impact my career for many years to come.

I want to thank my collaborator, Dr. Thomas Walz for his sincere welcome to work in his laboratory. Dr. Walz is always willing to share his expertise and experience with me. I am truly indebted to him for the time he dedicated to teaching me. I really appreciate his straightforward approach and meticulous style. I concur that ‘Seeing is believing’ and it would not be possible for my project without his help.

I would also like to thank my thesis committee members, Dr. Jue Chen, Dr. Scott Blanchard, and Dr. Sebastian Klinge, as they have spent substantial efforts on improving the quality of my research, as well as giving me career advice over the years. Dr. Jue Chen and Dr. Sebastian Klinge also offered generous help on computation devices and experimental reagents, respectively,

related to my studies. In addition, I would like to thank Dr. Antonina Roll-Mecak for being my external examiner and sharing her expertise on AAA proteins with me.

I also want to thank my colleagues in the Kapoor laboratory. In particular, I learned a great deal of yeast genetics by working with Dr. Shigehiro Kawashima. Dr. Shih-Chieh Ti has always been enthusiastic to discuss on any molecular biology and biochemistry experiments with me. Dr. Tommasso Cupido, Dr. Jonathan Steinman, Rudolf Pisa and Cristina Santarossa have enriched my experience as a member of the AAA team since the first day. In addition, I enjoyed working closely with Dr. Lisa Hang, Ashley C. Wang and Dr. Subbulakshmi Suresh on the project aiming to develop potential antifungal reagents. Dr. Michal Wieczorek and Megan Kelly have given me numerous thoughtful comments on my writings and I truly appreciate their effort.

I also want to thank my collaborators from the Walz laboratory, who welcome me to their group parties and barbeques on a routine basis. I am in particularly indebted to Dr. Hiroshi Suzuki, who has been a great teacher for me on both the theoretical and practical aspects of EM. I also learned a lot from our scientific discussions on our own data or other literature. I also appreciate the suggestions on data processing from Dr. Hyojin Kim and Dr. Yixiao Zhang, with whom I have shared the same computer room for a long time.

I want to thank my collaborators from Tri-Institutional Therapeutic Discovery Institute, I learned a lot from our discussion and meetings from the drug development perspective. In addition, I have received generous help from

other people in the Rockefeller University. Dr. Gregory Alushin kindly offered his computers when I wanted to utilize cryoSPARC to analyze my data. Mark Ebrahim and Johanna Sotiris at the Evelyn Gruss Lipper Cryo-EM Resource Center made my experience on the state-of-art Cryo-EM instruments as great as I can imagine.

I want to thank Dr. Derek Tan and everyone from the Tri-Institutional Program in Chemical Biology (TPCB) as well as everyone from the David Rockefeller Graduate Program. They always support and care for the students. I immensely appreciate their effort to make our studies free of logistics.

At last, I want to thank my parents and sister, who always support me as much as they can and never ask for anything in returns. I would not be able to do what I like doing today without their understanding. Also, I want to express my special thanks to Dr. Wan-Ying Hsieh, with whom I shared my good and bad days over the past six years. It has been amazing to share the experience of the darkness under the Milkyway, the gorgeous landscapes, as well as numerous unforgettable encounters of wildlives.

Table of Contents

List of Figures

List of Tables

List of Appendices

CHAPTER 1

INTRODUCTION

1.1 Overview of ribosome biogenesis

All cellular proteins are synthesized by ribosomes, which are >3 MDa macromolecular complexes comprised of four ribosomal RNAs (rRNA) and ~80 ribosomal proteins in yeast (Yusupova and Yusupov, 2014). These complex nanomachines consist of two ribonucleoprotein subunits, the large 60S subunit (LSU) and the small 40S subunit (SSU). The 60S subunit contains three rRNAs (25S, 5.8S and 5S) and 46 ribosomal proteins while the 40S subunit is comprised of one rRNA (18S) and 33 different ribosomal proteins. The 60S subunit possesses the peptidyl transfer center and the peptide exit tunnel while the 40S subunit functions as the decoding center to proofread the complementary interaction between messenger RNAs (mRNA) and transfer RNAs (tRNA).

In rapidly dividing cells, a substantial fraction of essential cellular pathways, including transcription, translation and nuclear transport, are dedicated to producing required numbers of these complex macromolecular machines for growth (Thomson et al., 2013; Tschochner and Hurt, 2003; Warner, 1999). For instance, every minute a growing budding yeast cell has to generate ~2000 ribosomes to double the number of ribosomes during the ~90 minutes doubling time (Warner, 1999). To ensure the translation-competency of the final products, ribosome biogenesis is a highly regulated and multi-step

process requiring a plethora of more than 200 unique assembly factors. Consistent with the vital roles of translational fidelity in cells, defects in ribosome assembly are linked to various human diseases referred to as ribosomopathies or could be embryonically lethal (Freed et al., 2010). Currently, understanding the molecular mechanism of these phenotypes is challenging, as the functional roles of different assembly factors remain poorly understood.

Ribosome biogenesis begins from the transcription of ribosomal DNA (rDNA) in the nucleolus to generate the 35S pre-rRNA that includes the mature 18S, 5.8S and 25 S rRNA sequences from 5' to 3', divided by two internal transcribed spacers (ITS1 and ITS2) (Figure 1-1A) (Konikkat and Woolford, 2017). The modification and processing of the pre-rRNA are coupled to ordered association and dissociation of different ribosomal proteins and assembly factors to form various intermediates in the nucleus (Figure 1-1B) (Kressler et al., 2012). The first stable pre-ribosome intermediate, composed of the emerging pre-rRNA and ~70 assembly factors, were directly visualized by electron microscopy as knobs on rDNA chromatin spreads, suggesting the assembling process can be co-transcriptional, likely starting from the 5' end of the pre-rRNA (Miller and Beatty, 1969). These intermediates were found to represent the co-transcriptionally assembled precursors of the 40S subunit (also named as SSU processome or 90S particle) (Grandi et al., 2002). Recent cryo-EM structures of the SSU processome reveal that the rRNA domains at the 5' end of the precursors of 18S rRNA transcript adopt

conformations similar to those in the mature 40S subunit while domains in the 3' end are distinctly different, indicating a hierarchical, 5'-to-3'-oriented assembly process (Chaker-Margot et al., 2017; Kornprobst et al., 2016; Sun et al., 2017).

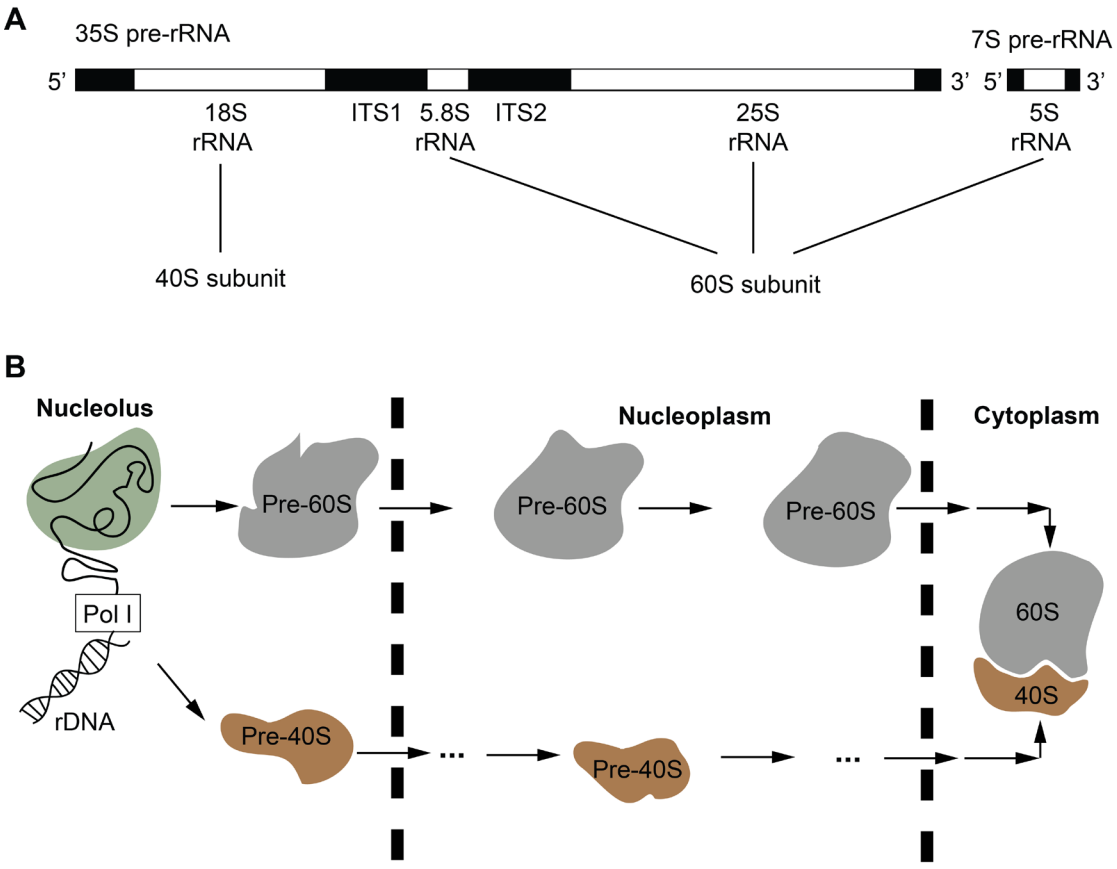


Figure 1-1. Overview of ribosome biogenesis in eukaryotes. (A) Schematics for the pre-RNAs and the composition of rRNA in mature ribosomes. The segments of rRNA in mature ribosomes and the internal transcribed spacers (ITS1 and ITS2) are labeled. (B) Ribosome biogenesis starts with the transcription of rDNA in the nucleolus. Processing and remodeling of precursors of 60S and 40S subunits (pre-60S and pre-40S) are separated in the nucleolus, nucleoplasm and cytoplasm. The matured 60S and 40S subunits combine in the cytoplasm to form the functional ribosome.

Cleavage of ITS1 leads to separation of the rRNA transcripts of the 60S subunit from that of the 40S subunit (Figure 1-1A), also marking the point after which the two subunits are assembled largely independently from each other (Tschochner and Hurt, 2003; Woolford and Baserga, 2013). Properly assembled precursors of 60S and 40S subunits are then exported to the cytoplasm, where the final remodeling steps lead to functional ribosomes (Figure 1-1B).

In contrast to the relatively straightforward 5'-to-3' sequential folding of the 18S rRNA in the 40S subunit, assembly of the 60S subunit appears more complex as the six conserved rRNA domains of 25S and 5.8S rRNA are intertwined with the ribosomal proteins in the mature ribosome (Klinge et al., 2011). The folding and interaction of rRNA are assisted by specific assembly factors under tight spatiotemporal regulation, producing multiple intermediates collectively referred to as pre-60S particles. Identification of different assembly factors involved at different stages of ribosome biogenesis and the implementation of affinity-purification approaches led to the purification of specific pre-60S particles (Nissan et al., 2002). Their molecular compositions were defined through RNA analyses by northern blots and protein analyses by mass spectrometry. The rRNA processing events can thus be correlated with the association and dissociation of specific sets of assembly factors. While the overall reaction pathway is known, the molecular mechanism of the RNA processing and the functions of the assembly factors remain outstanding questions in the field of ribosome biogenesis.

Recent progress in cryo-EM allowed the visualization of key pre-60S particles at atomic or near-atomic resolution, providing not only insights into the processing and folding process of the rRNA but also the structural framework for the functional analysis of assembly factors (Konikkat and Woolford, 2017). These particles include multiple states of Nsa1 associated pre-60S particles in the nucleolus, the early nucleoplasmic Nog2 particles and the late nucleoplasmic Rix1 particles, and the cytoplasmic Nmd3 particle. The insights from these structural studies into the ordered assembling of ribosomes are briefly summarized below.

Comparisons of high-resolution structures of different nucleolar pre-ribosomes isolated by pulldown of tagged Nsa1 and the early nucleoplasmic Nog2 particle suggest that individual rRNA domains in pre-60S particles are assembled in a modular manner (Kater et al., 2017; Sanghai et al., 2018; Wu et al., 2016). In particular, domains I and II of 25S rRNA, which constructs the solvent-exposed side of the 60S subunits, already folds into its mature conformation in the earliest Nsa1-containing pre-60S particle, suggesting they function as a scaffold for further construction. In addition, Nsa1-associated intermediates with either rRNA domain III or domain VI folded are observed, likely representing two different pathways to generate a converging intermediate in which these two domains are later joined by domain IV as observed in the nucleoplasmic Nog2 particle. The stepwise construction of different rRNA domains is accompanied by the formation of the peptide exit tunnel that is surrounded by these rRNA domains.

In the nucleolar pre-60S particles, Nsa1 is associated with a group of known early factors, including Rpf1 that functions as a placeholder for the pre-matured PET (Kater et al., 2017). The removal of Nsa1 and its tightly associated proteins from the nucleolar pre-60S particle, most likely driven by an AAA protein (ATPase Associated with various Activities), Rix7, would result in an empty pre-PET (Kressler et al., 2008). This empty pre-PET can be probed by other nucleoplasmic factors in a later pre-60S particle, such as Nog1 in the nucleoplasmic Nog2 particle, as an additional quality control step to ensure the generation of properly assembled PET (Wu et al., 2016). Another evident difference between the nucleolar particles and the Nog2 particle is the absence of Ytm1-Erb1 subcomplex in the later. In the Nsa1 particles, the Ytm1-Erb1 subcomplex and its associated proteins covers the ITS2 site of the rRNA that separates the 25S and 5.8S rRNA transcripts (See Figure 1-1A) (Kater et al., 2017; Sanghai et al., 2018). The release of Ytm1-Erb1 and its associated assembly factors are required for the recruitment of the nuclease Las1 to cleave the ITS2 site (Gasse et al., 2015). In addition, the binding site of Ytm1-Erb1 overlaps with where Nop53 binds in the Nog2 particle, which is an adaptor protein that recruits the exosome to further process the cleaved ITS2 (Wu et al., 2016). Therefore, proper removal of Ytm1-Erb1 subcomplex is a required for further irreversible cleavage and processing of rRNAs. Previous yeast two-hybrid experiments revealed that Ytm1 could directly interact with the MIDAS domain (Metal Ion Dependent Adhesion Site) in Mdn1, which is the largest AAA protein in eukaryotes (~5000

amino acids in a single polypeptide chain) (Bassler et al., 2010). Disruption of this interaction leads to retention of nucleolar assembly factors in later nucleoplasmic pre-60S particles and accumulation of rRNA intermediates with uncleaved ITS2, suggesting that this interaction is required for the removal of nucleolar factors and processing of ITS2. At last, the Nog2 particle, but none of the Nsa1-associated pre-60S particles, contains the 5S RNP (ribonucleoprotein), assembled separately by the 5S rRNA, Rpl5 and Rpl11. However, this 5S RNP is oriented ~180°-rotated compared to its position in the mature ribosome. These comparisons reveal the timing of assembly of 5S RNP and also suggest that its additional remodeling is required.

The structure of Rix1 particle revealed features of late nucleoplasmic pre-ribosomes before the nuclear export (Barrio-Garcia et al., 2016). In comparison to the Nog2 particle, the 5S RNP is rotated by ~180° in the Rix1 particle, forming the basis of central protuberance close to its mature conformation. The Rpf2-Rrs1 subcomplex that anchors the 5S RNP in its immature conformation in the Nog2 particle is absent in the Rix1 particle. Instead, additional factors, including the Rix1 subcomplex and its interactor Mdn1, are recruited to regions that would clash with the Rpf2-Rrs1 subcomplex if it was retained, suggesting the binding of these later assembly factors are coupled to the repositioning of 5S RNP. Interestingly, Mdn1's MIDAS domain also interacts with another assembly factor Rsa4, and this interaction is required for the ATP-dependent release of Rsa4 from the Rix1 particle and relocation of helix structures of rRNA in the peptidyl transfer

center (Bassler et al., 2014; Ulbrich et al., 2009). This energy-dependent remodeling process by Mdn1 appears to be the key irreversible, proofreading step to ensure that only pre-ribosomes with properly oriented 5S RNP could be further processed for the maturation of peptidyl transfer center. In addition, the Mdn1 remodeling step also triggers hydrolysis of GTPase Nog2 and its release, so that Nmd3, the adaptor protein that recruits the nuclear exportin Crm1, can bind to the site previously occupied by Nog2 in the nucleoplasmic pre-60S particles (Matsuo et al., 2014).

A recent structure of Nmd3 particle provides structural insights into the cytoplasmic maturation process of pre-ribosomes (Ma et al., 2017). In particular, the Nmd3 interacts with the peptidyl transfer center in its near-mature conformation, suggesting the proofreading function of Nmd3. In addition, the Nog1 is removed, most likely promoted by another AAA protein Drg1 (Kappel et al., 2012; Lo et al., 2010), leaving the peptide exit tunnel available for Reh1. The C-terminal helix of Reh1 is inserted into the PET in the Nmd3 particle, functioning as another placeholder to check the structural integrity of this functional site.

In summary, previous biochemical and structural studies of pre-ribosomes revealed three common themes for the function of assembly factors in ribosome biogenesis: (i) the exploitation of steric hindrance and molecular mimicry of assembly factors or ribosomal proteins to block pre-matured binding of later factors (Sanghai et al., 2018); (ii) the reoccurring quality controls coupled to the complementary interaction between the assembly

factors and intermediate functional sites; (iii) the multiple involvements of energy-harnessing enzymes, in particular AAA proteins, for the irreversible removal of assembly factors as key remodeling events. However, how these AAA proteins perform their mechanical functions in cells remains poorly understood due to the transient nature of these remodeling steps in ribosome biogenesis.

1.2 Overview of AAA proteins

The AAA protein family is a large and functionally diverse group of enzymes that can couple ATP-hydrolysis with mechanical work (Erzberger and Berger, 2006; Hanson and Whiteheart, 2005). AAA proteins are required for different essential cellular processes, such as protein homeostasis, molecular transport and DNA replication, as well as ribosome biogenesis. Exactly how different AAA proteins contribute to diverse cellular pathways has remained largely an open question and been actively studied.

The hallmark of the AAA proteins is a ~250 amino acids AAA domain comprised of a large and a small subdomains, which usually oligomerize to form the functional enzymes (Erzberger and Berger, 2006). The ATPase sites in AAA proteins lie at the interface between two neighboring AAA domains within the functional oligomers. During the ATPase cycle, the nucleotide state of the ATPase site can be recognized by the neighboring AAA domains or subdomains. Alterations of the nucleotide state can drive the change of the relative orientation of surrounding AAA domains or subdomains, thus the

overall conformation of the oligomers and apply mechanical force to their cellular substrates. As AAA proteins are involved in different cellular pathways, the force generated is applied to a large array of client proteins, including other soluble/membrane proteins with specific signaling peptide or biophysical properties, DNA double helix, cytoskeleton filaments or pre-ribosomes. During evolution, different AAA proteins are tailored based on the specific mechanical tasks for the corresponding client proteins, leading to diverse mechanochemical mechanisms among different members of the superfamily.

The “classic clade” of AAA family includes a group of extensively studied AAA proteins, such as ClpX, Yme1 and P97 (Erzberger and Berger, 2006). These AAA proteins can recognize either unstructured or misfolded peptide sequence and unfold their client proteins, therefore they are also called as AAA unfoldases. Biochemical and structural studies reveal that they all function as homohexamers of six identical polypeptide chains that assemble into a ring-like structure (Figure 1-2). Both ClpX and Yme1 have one AAA domain in a single polypeptide and function by forming hexameric AAA rings while P97 has two AAA domains therefore forms a double-layered ring structure (Figure 1-2B). In these AAA unfoldases, conserved ‘Ar-Φ-Gly’ tripeptide (Ar: aromatic residue; Φ: hydrophobic residue, named ‘pore loop’) from each AAA domain interacts with the client proteins inside the central hole of the AAA ring (Bodnar and Rapoport, 2017; Martin et al., 2008; Puchades et al., 2017). During ATP hydrolysis, the conformational changes of the AAA domains can drive the movement of pore loops and thread their substrates

through the central hole processively (Figure 1-2B). Interestingly, Rix7 and Drg1, the two AAA proteins required for the removal of Nsa1 and Rpl24 from the nucleolar and cytoplasmic pre-60S particles, respectively, are highly homologous to P97 (Kressler et al., 2012) and have the conserved 'Ar-Φ-Gly' tripeptide in their orthologs. It is likely that Rix7 and Drg1 recognize exposed sequence of Nsa1 and Rpl24, respectively, and extract them out from the corresponding pre-60S particles during ATPase cycle similar to other AAA unfoldases.

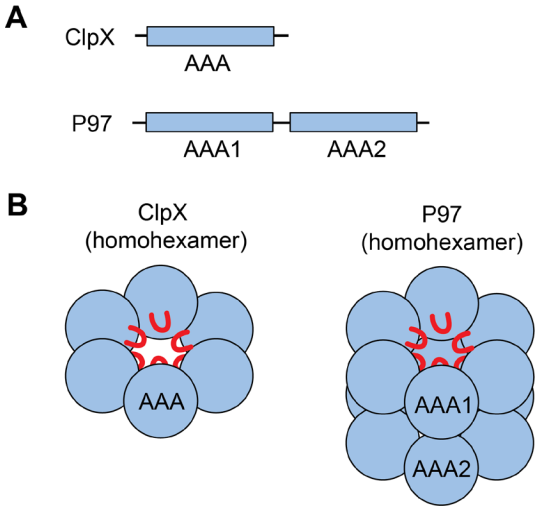


Figure 1-2. Domain organizations and schematics for the functional hexamers of ClpX and P97. (A) Domain organizations of ClpX and P97 and (B) schematics of their homohexamers are shown. The conserved pore loops (red curves) are highlighted in (B).

Another extensively studied AAA protein is dynein, the molecular motor that can transport cargos along microtubules during ATP hydrolysis (Schmidt and Carter, 2016). The force-generation properties of different dynein isoforms have been implicated to play important roles in cell division, intraflagellar

transport, signaling and cilia beatings. Particular cellular functions have been linked to diverse N-terminal sequences from different dyneins, which acts as a cargo-binding domain targeting different isoforms to the corresponding substrates (Figure 1-3A). On the other hand, the C-terminal sequence (~3500 amino acids) forms a relatively conversed AAA motor domain, an engine that transforms the chemical energy stored in ATP into directional movement. An unusual feature of the motor domain is that six non-equivalent AAA domains are linked in a single polypeptide chain, forming a pseudo-hexameric AAA ring (Figure 1-3B). This asymmetry leads to functional specializations of individual AAA domains: (i) The AAA1 site is the “main” site for ATP hydrolysis that drives the conformational changes of dynein and AAA3, also being competent in ATP hydrolysis, acts as a regulatory switch to control the overall activities of AAA1. The AAA2 and AAA4 do not possess conserved key motifs required for hydrolysis. The AAA5 and AAA6 have highly degenerate sequences lacking motifs for both ATP binding and hydrolysis, suggesting these four sites play structural, instead of enzymatic, roles. (ii) The sequence immediately before AAA1 forms multiple helical bundles and connects to the N-terminal domains that appear as a protrusion from the AAA ring. (iii) Structural studies revealed that a long insertion in AAA4 domain forms an elongated “stalk” extending ~20 nm out of the AAA ring. The microtubule-binding domain (MTBD) lies at the tip of the stalk.

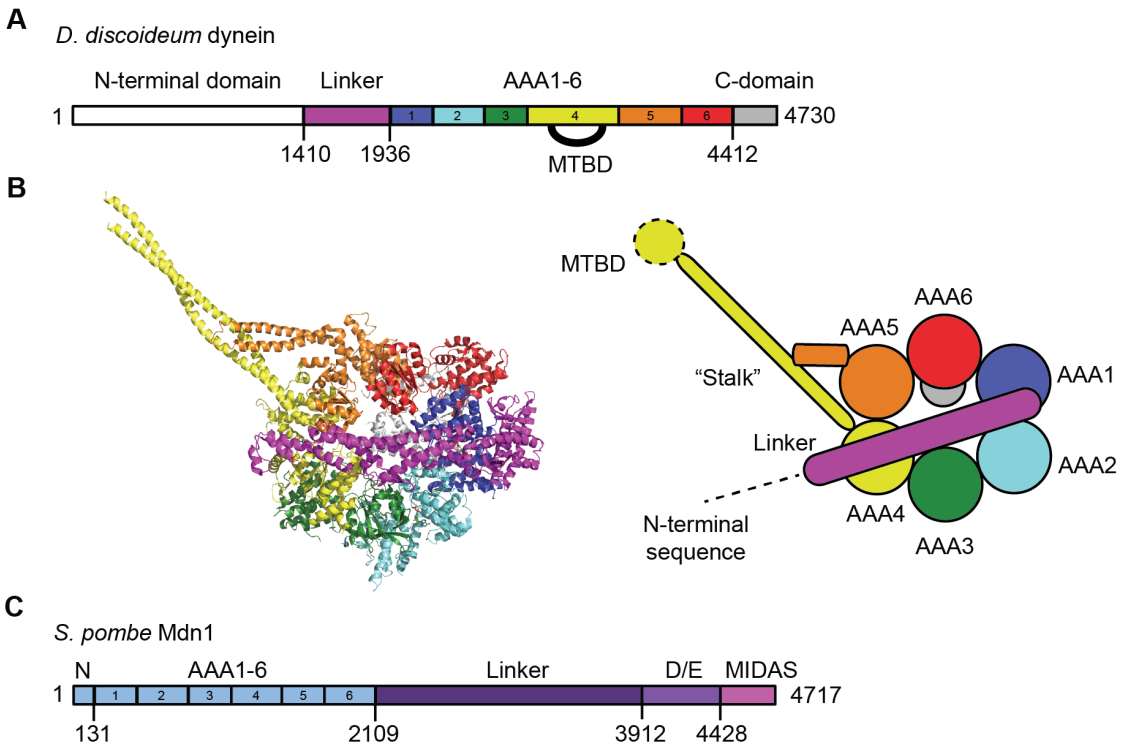


Figure 1-3. Domain organization and structure of *D. discoideum* dynein and *S. pombe* Mdn1. (A) Domain organization (aa 1-4730) and (B) structural model of *D. discoideum* dynein (PDB: 3VKH). Different domains are colored based on (A). A schematic of the model is also shown on the right. The N-terminal sequence and microtubule-binding domain (MTBD) not observed in the crystal structure are marked as dashed lines. (C) Domain organization of *S. pombe* Mdn1 (aa 1-4717). Note the linker domain in Mdn1 and the linker domain in dynein have the same name for historical reasons but otherwise not related to each other.

Our current model for how dynein performs its mechanical functions is largely based on recent single-molecular and high-resolution structural analyses of cytoplasmic dyneins from different species (Bhabha et al., 2014; DeWitt et al., 2015; Kon et al., 2012; Reck-Peterson et al., 2006). During ATP hydrolysis in AAA1, the opening and closing of the AAA1 ATPase site are

coupled to the conformational changes of the microtubule-binding domain through the stalk, leading to binding or unbinding to microtubules and also drive the bending of the linker domain, applying force to its cargoes that are associated with the N-terminal sequence of dynein (See domain organization in Figure 1-3B).

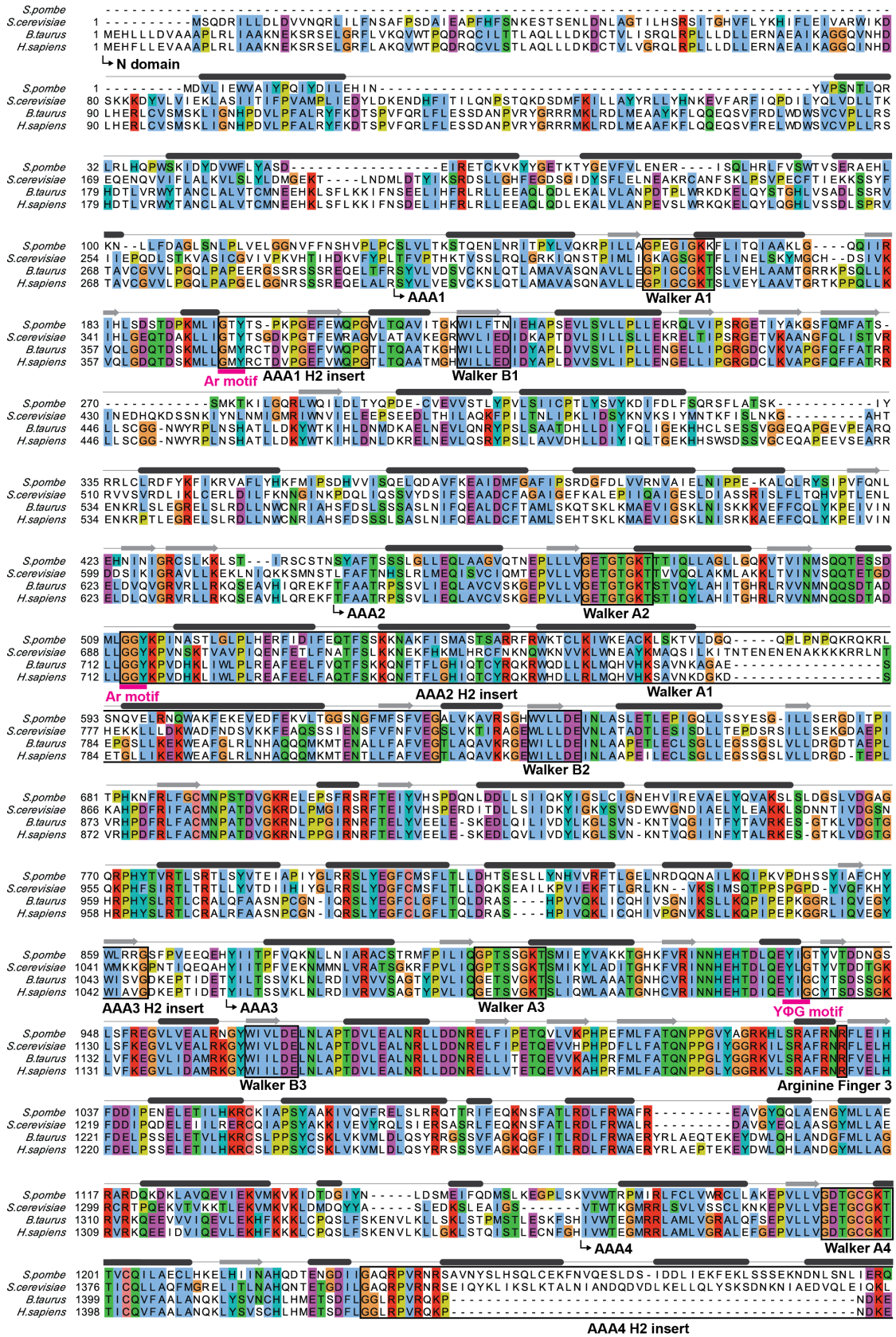
Interestingly, Mdn1, the AAA protein that remodels both nucleolar and nucleoplasmic pre-60S particle, also has six non-equivalent AAA domains linked in a single polypeptide chain (domain organization shown in Figure 1-3C). In addition, Mdn1 and dynein are close relatives in the AAA superfamily based on the alignment of primary sequences (Kressler et al., 2012; Ulbrich et al., 2009). Therefore, Mdn1 has been regarded as a dynein-like AAA protein. However, the AAA domains in *S. pombe* Mdn1 share less than 15% identity with those in *S. pombe* dynein (see details in 1.4 Method). The rest of Mdn1's sequence, which accounts for more than 50% of the total protein, does not share similarities with dynein. Therefore, it has been difficult to develop reliable models for Mdn1 function based on our understanding of dynein or other AAA proteins.

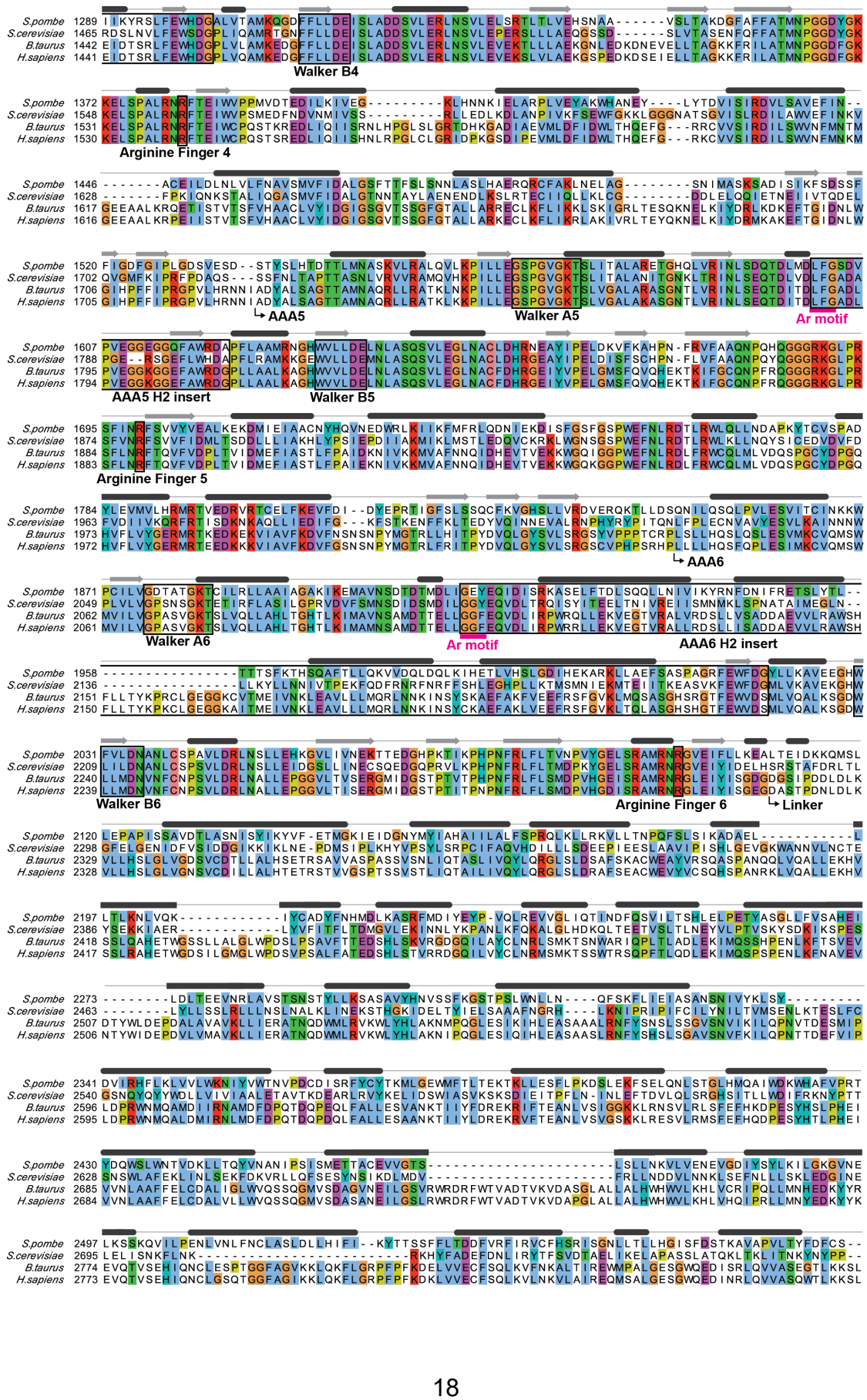
1.3 Previous biochemical and structural characterization of Mdn1

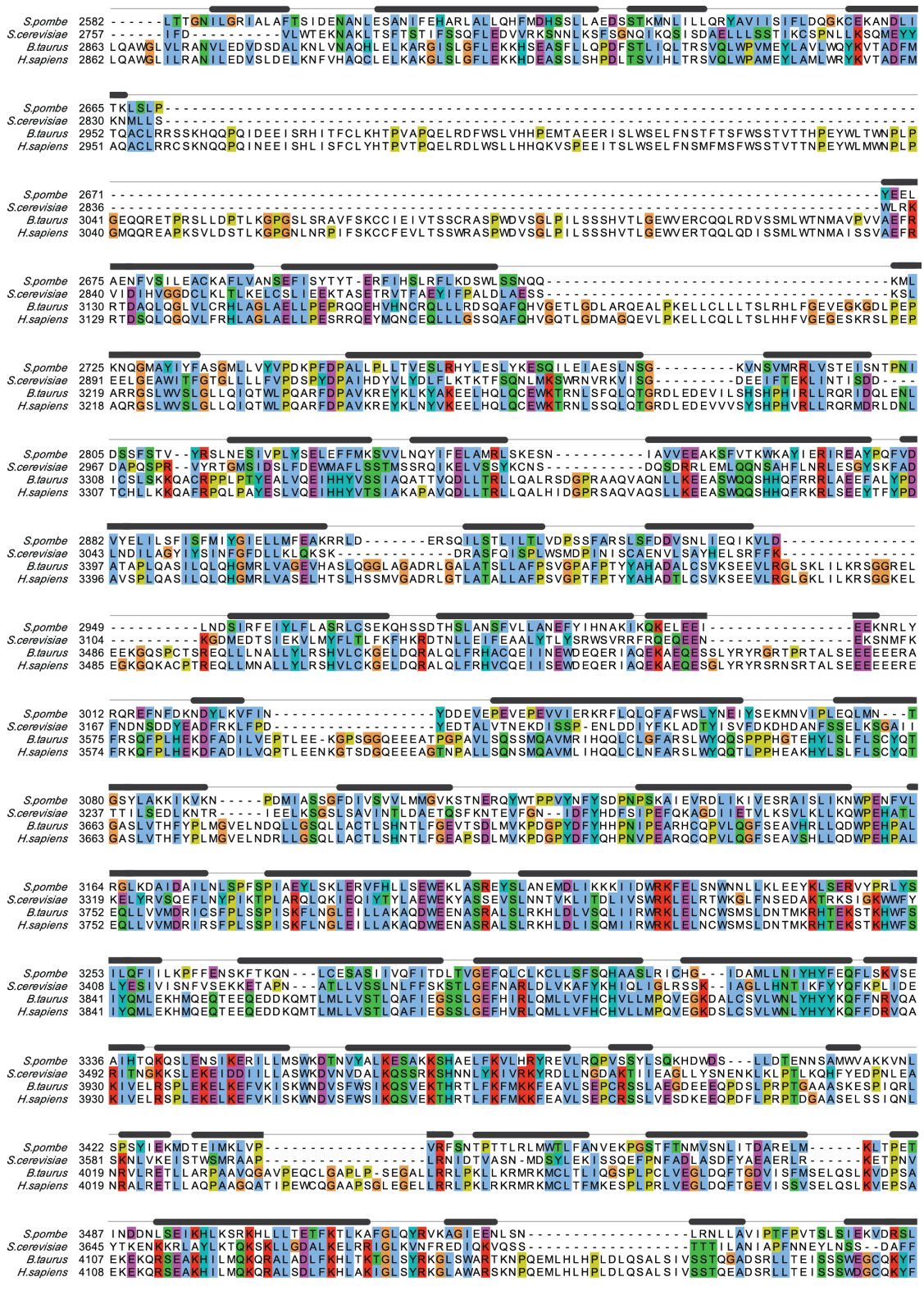
Alignment of primary sequences of Mdn1 orthologs reveals a conserved domain organization from yeasts to human proteins (Figure 1-4). As an example, *S. pombe* Mdn1 has a small domain at the N terminus, named the N domain (~130 aa in *S. pombe* Mdn1), followed by six AAA domains, and a

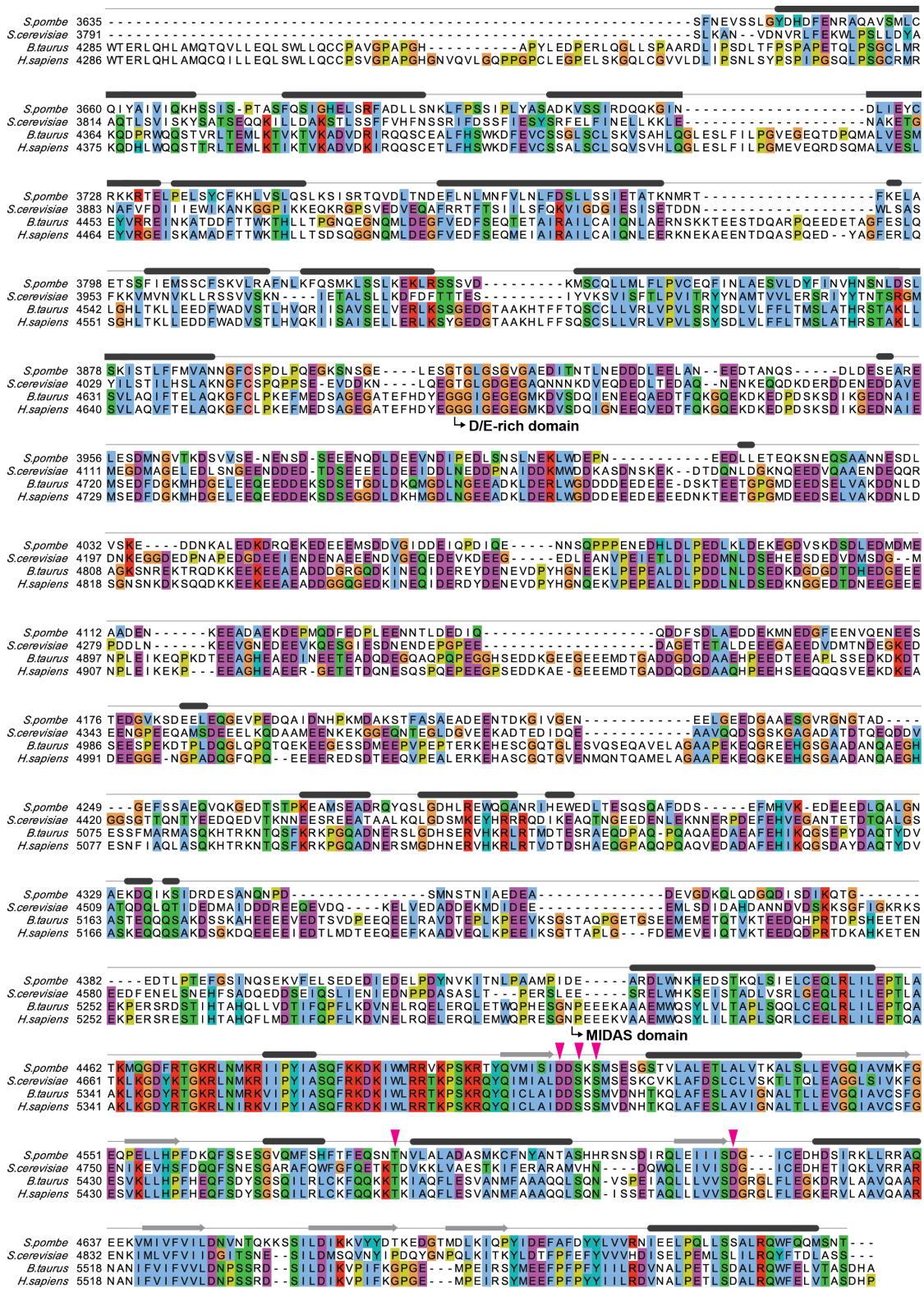
'tail', which is comprised of the linker, Asp/Glu-rich (D/E-rich), and MIDAS domains (Figure 1-3C). Currently, the biochemical studies of Mdn1 are limited, partially due to its large size (~5000 amino acids in a single polypeptide) and the difficulties in obtaining recombinant full-length protein for *in vitro* biochemical and structural studies.

Figure 1-4. Sequence alignment of Mdn1 orthologs. The sequences were aligned and displayed in Jalview (Waterhouse et al., 2009). Domain boundaries and selected structural and catalytic motifs are marked. The numbers following the Walker A, Walker B and arginine finger motifs indicate the AAA domain in which they are located (e.g. Walker A2 denotes the Walker A motif in the AAA2 domain). Secondary-structure elements in *S. pombe* Mdn1 were predicted using SPIDER2 (Yang et al., 2017) and the potential α-helices (black cylinders) and β-strands (grey arrows) are shown above the *S. pombe* sequence. Residues that are important for the interaction between the MIDAS domain and Rsa4 in *S. cerevisiae* are indicated (magenta triangles for Asp-4511, Ser-4513, Ser-4515, Thr-4582 and Asp-4619 in *S. pombe* Mdn1).


S.pombe
S.cerevisiae
B.taurus
H.sapiens
N domain
AAA1
Walker A1
Ar motif
AAA1 H2 insert
Walker B1
AAA2
Walker A2
AAA2 H2 insert
Walker B2
AAA3 H2 insert
AAA3
Walker A3
YΦG motif
Walker B3
Arginine Finger 3
AAA4
Walker A4
AAA4 H2 insert


S.pombe 1289 IIKYRSLFEWHDGALVTAMKQGDFFLLDEISLADDSVLERLNSVLELSRTLTLVEHSNAA------VSLTAKDGFAFFATMNPGGDYGK
S.cerevisiae 1465 RDSLNVLFEWSDGPLIQAMRTGNFFLLDEISLADDSVLERLNSVLEPERSLLLAEQGSSD------SLVTASENFQFFATMNPGGDYGK
B.taurus 1442 EIDTSRLFEWHDGPLVLAMKEDGFFLLDEISLADDSVLERLNSVLEVEKTLLLAEKGNLEDKDNEVELLTAGKKFRILATMNPGGDFGK
H.sapiens 1441 EIDTSRLFEWHDGPLVQAMKEDGFFLLDEISLADDSVLERLNSVLEVEKSLVLAEKGSPEDKDSEIELLTAGKKFRILATMNPGGDFGK
Walker B4
S.pombe 1372 KELSPALRNRFTEIWVPPMVDTEDILKIVEG---------KLHNNKIELARPLVEYAKWHANEY----LYTDVISIRDVLSAVEFIN--
S.cerevisiae 1548 KELSPALRNRFTEIWVPSMEDFNDVNMIVSS---------RLLEDLKDLANPIVKFSEWFGKKLGGGNATSGVISLRDILAWVEFINKV
B.taurus 1531 KELSPALRNRFTEIWCPQSTKREDLIQIISRNLHPGLSLGRTDHKGADIAEVMLDFIDWLTHQEFG---RRCVVSIRDILSWVNFMNTM
H.sapiens 1530 KELSPALRNRFTEIWCPQSTSREDLIQIISHNLRPGLCLGRIDPKGSDIPEVMLDFIDWLTHQEFG---RKCVVSIRDILSWVNFMNKM
Arginine Finger 4
S.pombe 1446 -------ACEILDLNLVLFNAVSMVFIDALGSFTTFSLSNNLASLHAERQRCFAKLNELAG--------SNIMASKSADISIKFSDSSF
S.cerevisiae 1628 -------FPKIQNKSTALIQGASMVFIDALGTNNTAYLAENENDLKSLRTECIIQLLKLCG--------DDLELQQIETNEIIVTQDEL
B.taurus 1617 GEEAALKRQETISTVTSFVHAACLVYIDGIGSGVTSSGFGTALLARRECLKFLIKKLSKIGRLTESQKNELKIYDRLKDKEFTGIDNLW
H.sapiens 1616 GEEAALKRPEIISTVTSFVHAACLVYIDGIGSGVTSSGFGTALLARKECLKFLIKRLAKIVRLTEYQKNELKIYDRMKAKEFTGIDNLW
AAA5
Walker A5
Ar motif
S.pombe 1520 FIGDFGIPLGDSVESD--STYSLHTDTTLMNASKVLRALQVLKPILLEGSPGVGKTSLITALARETGHQLVRINLSDQTDLMDLFGSDV
S.cerevisiae 1702 QVGMFKIPRFPDAQS---SSFNLTAPTTASNLVRVVRAMQVHKPILLEGSPGVGKTSLITALANITGNKLTRINLSEQTDLVDLFGADA
B.taurus 1706 GIHPFFIPRGPVLHRNNIADYALSAGTTAMNAQRLLRATKLNKPILLEGSPGVGKTSLVGALARASGNTLVRINLSEQTDITDLFGADL
H.sapiens 1705 GIHPFFIPRGPVLHRNNIADYALSAGTTAMNAQRLLRATKLKKPILLEGSPGVGKTSLVGALAKASGNTLVRINLSEQTDITDLFGADL
AAA5 H2 insert
Walker B5
S.pombe 1607 PVEGGEGGQFAWRDAPFLAAMRNGHWVLLDELNLASQSVLEGLNACLDHRNEAYIPELDKVFKAHPN-FRVFAAQNPQHQGGGRKGLPR
S.cerevisiae 1788 PGE--RSGEFLWHDAPFLRAMKKGEWVLLDEMNLASQSVLEGLNACLDHRGEAYIPELDISFSCHPN-FLVFAAQNPQYQGGGRKGLPK
B.taurus 1795 PVEGGKGGEFAWRDGPLLAALKAGHWVVLDELNLASQSVLEGLNACFDHRGEIYVPELGMSFQVQHEKTKIFGCQNPFRQGGGRKGLPR
H.sapiens 1794 PVEGGKGGEFAWRDGPLLAALKAGHWVVLDELNLASQSVLEGLNACFDHRGEIYVPELGMSFQVQHEKTKIFGCQNPFRQGGGRKGLPR
Arginine Finger 5
S.pombe 1695 SFINRFSVVYVEALKEKDMIEIAACNYHQVNEDWRLKIIKFMFRLQDNIEKDISFGSFGSPWEFNLRDTLRWLQLLNDAPKYTCVSPAD
S.cerevisiae 1874 SFVNRFSVVFIDMLTSDDLLLIAKHLYPSIEPDIIAKMIKLMSTLEDQVCKRKLWGNSGSPWEFNLRDTLRWLKLLNQYSICEDVDVFD
B.taurus 1884 SFLNRFTQVFVDPLTVIDMEFIASTLFPAIDKNIVKKMVAFNNQIDHEVTVEKKWGQIGGPWEFNLRDLFRWCQLMLVDQSPGCYDPGQ
H.sapiens 1883 SFLNRFTQVFVDPLTVIDMEFIASTLFPAIEKNIVKKMVAFNNQIDHEVTVEKKWGQKGGPWEFNLRDLFRWCQLMLVDQSPGCYDPGQ
AAA6
S.pombe 1784 YLEVMVLHRMRTVEDRVRTCELFKEVFDI--DYEPRTIGFSLSSQCFKVGHSLLVRDVERQKTLLDSQNILQSQLPVLESVITCINKKW
S.cerevisiae 1963 FVDIIVKQRFRTISDKNKAQLLIEDIFG---KFSTKENFFKLTEDYVQINNEVALRNPHYRYPITQNLFPLECNVAVYESVLKAINNNW
B.taurus 1973 HVFLVYGERMRTKEDKEKVIAVFKDVFGSNSNPYMGTRLLHITPYDVQLGYSVLSRGSYVPPPTCRPLSLLHQSLQSLESVMKCVQMSW
H.sapiens 1972 HVFLVYGERMRTEEDKKKVIAVFKDVFGSNSNPYMGTRLFRITPYDVQLGYSVLSRGSCVPHPSRHPLLLLHQSFQPLESIMKCVQMSW
S.pombe 1871 PCILVGDTATGKTCILRLLAAIAGAKIKEMAVNSDTDTMDLIGEYEQIDISRKASELFTDLSQQLLNIVIKYRNFDNIFRETSLYTL--
S.cerevisiae 2049 PLVLVGPSNSGKTETIRFLASILGPRVDVFSMNSDIDSMDILGGYEQVDLTRQISYITEELTNIVREIISMNMKLSPNATAIMEGLN--
B.taurus 2062 MVILVGPASVGKTSLVQLLAHLTGHTLKIMAVNSAMDTTELLGGFEQVDLIRPWRQLLEKVEGTVRALVRDSLLVSADDAEVVLRAWSH
H.sapiens 2061 MVILVGPASVGKTSLVQLLAHLTGHTLKIMAMNSAMDTTELLGGFEQVDLIRPWRRLLEKVEGTVRALLRDSLLISADDAEVVLRAWSH
Walker A6
Ar motif
AAA6 H2 insert
S.pombe 1958 ----------------TTTSFKTHSQAFTLLQKVVDQLDQLKIHETLVHSLGDIHEKARKLLAEFSASPAGRFEWFDGYLLKAVEEGHW
S.cerevisiae 2136 ----------------LLKYLLNNIVTPEKFQDFRNRFNRFFSHLEGHPLLKTMSMNIEKMTEIITKEASVKFEWFDGMLVKAVEKGHW
B.taurus 2151 FLLTYKPRCLGEGGKCVTMEIVNKLEAVLLLMQRLNNKINSYSKAEFAKFVEEFRSFGVKLMQSASGHSRGTFEWVDSMLVQALKSGDW
H.sapiens 2150 FLLTYKPKCLGEGGKAITMEIVNKLEAVLLLMQRLNNKINSYCKAEFAKLVEEFRSFGVKLTQLASGHSHGTFEWVDSMLVQALKSGDW
S.pombe 2031 FVLDNANLCSPAVLDRLNSLLEHKGVLIVNEKTTEDGHPKTIKPHPNFRLFLTVNPVYGELSRAMRNRGVEIFLLKEALTEIDKKQMSL
S.cerevisiae 2209 LILDNANLCSPSVLDRLNSLLEIDGSLLINECSQEDGQPRVLKPHPNFRLFLTMDPKYGELSRAMRNRGVEIYIDELHSRSTAFDRLTL
B.taurus 2240 LLMDNVNFCNPSVLDRLNALLEPGGVLTVSERGMIDGSTPTVTPHPNFRLFLSMDPVHGEISRAMRNRGLEIYISGDGDGSIPDDLDLK
H.sapiens 2239 LLMDNVNFCNPSVLDRLNALLEPGGVLTISERGMIDGSTPTITPNPNFRLFLSMDPVHGDISRAMRNRGLEIYISGEGDASTPDNLDLK
Walker B6
Arginine Finger 6
Linker
S.pombe 2120 LEPAPISSAVDTLASNISYIKYVF-ETMGKIEIDGNYMYIAHAIILALFSPRQLKLLRKVLLTNPQFSLSIKADAEL----------L
S.cerevisiae 2298 GFELGENIDFVSIDDGIKKIKLNE-PDMSIPLKHYVPSYLSRPCIFAQVHDILLLSDEEPIEESLAAVIPISHLGEVGKWANNVLNCTE
B.taurus 2329 VLLHSLGLVGDSVCDTLLALHSETRSAVVASPASSVSNLIQTASLIVQYLQRGLSLDSAFSKACWEAYVRSQASPANQQLVQALLEKHV
H.sapiens 2328 VLLHSLGLVGNSVCDILLALHTETRSTVVGSPTSSVSTLIQTAILIVQYLQRGLSLDRAFSEACWEVYVCSQHSPANRKLVQALLEKHV
S.pombe 2197 LTLKNLVQK----------IYCADYFNHMDLKASRFMDIYEYP-VQLREVVGLIQTINDFQSVILTSHLELPETYASGLLFVSAHEI
S.cerevisiae 2386 YSEKKIAER----------LYVFITFLTDMGVLEKINNLYKPANLKFQKALGLHDKQLTEETVSLTLNEYVLPTVSKYSDKIKSPES
B.taurus 2418 SSLQAHETWGSSLLALGLWPDSLPSAVFTTEDSHLSKVRGDGQILAYCLNRLSMKTSNWARIQPLTLADLEKIMQSSHPENLKFTSVEV
H.sapiens 2417 SSLRAHETWGDSILGMGLWPDSVPSALFATEDSHLSTVRRDGQILVYCLNRMSMKTSSWTRSQPFTLQDLEKIMQSPSPENLKFNAVEV
S.pombe 2273 --------LDLTEEVNRLAVSTSNSTYLLKSASAVYHNVSSFKGSTPSLWNLLN----QFSKFLIEIASANSNIVYKLSY---------
S.cerevisiae 2463 --------LYLLSSLRLLLNSLNALKLINEKSTHGKIDELTYIELSAAAFNGRH----LKNIPRIPIFCILYNILTVMSENLKTESLFC
B.taurus 2507 DTYWLDEPDALAVAVKLLIERATNQDWMLRVKWLYHLAKNMPQGLESIKIHLEASAAALRNFYSNSLSSGVSNVIKILQPNVTDESMIP
H.sapiens 2506 NTYWIDEPDVLVMAVKLLIERATNQDWMLRVKWLYHLAKNIPQGLESIQIHLEASAASLRNFYSHSLSGAVSNVFKILQPNTTDEFVIP
S.pombe 2341 DVIRHFLKLVVLWKNIYVWTNVPDCDISRFYCYTKMLGEWMFTLTEKTKLLESFLPKDSLEKFSELQNLSTGLHMQAIWDKWHAFVPRT
S.cerevisiae 2540 GSNQYQYYWDLLVIVIAALETAVTKDEARLRVYKELIDSWIASVKSKSDIEITPFLN-INLEFTDVLQLSRGHSITLLWDIFRKNYPTT
B.taurus 2596 LDPRWNMQAMDIIIRNAMDFDPQTDQPEQLFALLESVANKTIIYFDREKRIFTEANLVSIGGKKLRNSVLRLSFEFHKDPESYHSLPHEI
H.sapiens 2595 LDPRWNMQALDMIRNLMDFDPQTDQPDQLFALLESAANKTIIYLDREKRVFTEANLVSVGSKKLRESVLRMSFEFHQDPESYHTLPHEI
S.pombe 2430 YDQWSLWNTVDKLLTQYVNANIPSISMETTACEVVGTS----------LSLLNKVLVENEVGDIYSYLKILGKGVNE
S.cerevisiae 2628 SNSWLAFEKLINLSEKFDKVRLLQFSESYNSIKDLMDV----------FRLLNDDVLNNKLSEFNLLLSKLEDGINE
B.taurus 2685 VVNLAAFFELCDALIGLWVQSSQGMVSDAGVNEILGSVRWRDRFWTVADTVKVDASGLALLALHWHWVLKHLVCRIPQLLMNHEDKYYR
H.sapiens 2684 VVNLAAFFELCDALVLLWVQSSQGMVSDASANEILGSLRWRDRFWTVADTVKVDAPGLALLALHWHWVLKHLVHQIPRLLMNYEDKYYK
S.pombe 2497 LKSSKQVILPENLVNLFNCLASLDLLHIFI--KYTTSSFFLTDDFVRFIRVCFHSRISGNLLTLLHGISFDSTKAVAPVLTYFDFCS--
S.cerevisiae 2695 LELISNKFLNK----------RKHYFADEFDNLIRYTFSVDTAELIKELAPASSLATQKLTKLITNKYNYPP--
B.taurus 2774 EVQTVSEHIQNCLESPTGGFAGVKKLQKFLGRPFPFKDELVVECFSQLKVFNKALTIREWMPALGESGWQEDISRLQVVASEGTLKKSL
H.sapiens 2773 EVQTVSEHIQNCLGSQTGGFAGIKKLQKFLGRPFPFKDKLVVECFSQLKVLNKVLAIREQMSALGESGWQEDINRLQVVASQWTLKKSL


S.pombe 2582 ------LTTGNILGRIALAFTSIDENANLESANIFEHARLALLQHFMDHSSLLAEDSSTKMNLILLQRYAVIISIFLDQGKCEKANDLI
S.cerevisiae 2757 ------IFD----------VLWTEKNAKLTSFTSTIFSSQFLEDVVRKSNNLKSFSGNQIKQSISDAELLLSSTIKCSPNLLKSQMEYY
B.taurus 2863 LQAWGLVLRANVLEDVDSDALKNLVNAQHLELKARGISLGFLEKKHSEASFLLQPDFSTLIQLTRSVQLWPVMEYLAVLWQYKVTADFM
H.sapiens 2862 LQAWGLILRANILEDVSLDELKNFVHAQCLELKAKGLSLGFLEKKHDEASSLSHPDLTSVIHLTRSVQLWPAMEYLAMLWRYKVTADFM
S.pombe 2665 TKLSLP----------
S.cerevisiae 2830 KNMLLS----------
B.taurus 2952 TQACLRRSSKHQQPQIDEEISRHITFCLKHTPVAPQELRDFWSLVHHPEMTAEERISLWSELFNSTFTSFWSSTVTTHPEYWLTWNPLP
H.sapiens 2951 AQACLRRCSKNQQPQINEEISHLISFCLYHTPVTPQELRDLWSLLHHQKVSPEEITSLWSELFNSMFMSFWSSTVTTNPEYWLMWNPLP
S.pombe 2671 ----------YEEL
S.cerevisiae 2836 ----------WLRK
B.taurus 3041 GEQQRETPRSLLDPTLKGPGSLSRAVFSKCCIEIVTSSCRASPWDVSGLPILSSSHVTLGEWVERCQQLRDVSSMLWTNMAVPVVAEFR
H.sapiens 3040 GMQQREAPKSVLDSTLKGPGNLNRPIFSKCCFEVLTSSWRASPWDVSGLPILSSSHVTLGEWVERTQQLQDISSMLWTNMAISSVAEFR
S.pombe 2675 AENFVSILEACKAFLVANSEFISYTYT-ERFIHSLRFLKDSWLSSNQQ----------KML
S.cerevisiae 2840 VIDIHVGGDCLKLTLKELCSLIEEKTASETRVTFAEYIFPALDLAESS----------KSL
B.taurus 3130 RTDAQLQGLVLCRHLAGLAELLPEPRQQEHVHNCRQLLLRDSQAFQHVGETLGDLARQEALPKELLCLLLTSLRHLFGEVEGKGDLPEP
H.sapiens 3129 RTDSQLQGQVLFRHLAGLAELLPESRRQEYMQNCEQLLLGSSQAFQHVGQTLGDMAGQEVLPKELLCQLLTSLHHFVGEGESKRSLPEP
S.pombe 2725 KNQGMAYIYFASGMLLVYVPDKPFDPALLPLLTVESLRHYLESLYKESQILEIAESLNSG---------KVNSVMRRLVSTEISNTPNI
S.cerevisiae 2891 EELGEAWITFGTGLLLLFVPDSPYDPAIHDYVLYDLFLKTKTFSQNLMKSWRNVRKVISG---------DEEIFTEKLINTISDD----
B.taurus 3219 ARRGSLWVSLGLLQIQTWLPQARFDPAVKREYKLKYAKEELHQLQCEWKTRNLSFQLQTGRDLEDEVILSHSHPHIRLLRQRIDQLENL
H.sapiens 3218 AQRGSLWVSLGLLQIQTWLPQARFDPAVKREYKLNYVKEELHQLQCEWKTRNLSSQLQTGRDLEDEVVVSYSHPHVRLLRQRMDRLDNL
S.pombe 2805 DSSFSTV--YRSLNESIVPLYSELEFFMKSVVLNQYIFELAMRLSKESN----------IAVVEEAKSFVTKWKAYIERIREAYPQFVD
S.cerevisiae 2967 DAPQSPR--VYRTGMSIDSLFDEWMAFLSSTMSSRQIKELVSSYKCNS----------DQSDRRLEMLQQNSAHFLNRLESGYSKFAD
B.taurus 3308 ICSLSKKQACRPPLPTYEALVQEIHHYVSSIAQATTVQDLLTRLLQALRSDGPRAAQVAQNLLKEEASWQQSHHQFRRRLAEEFALYPD
H.sapiens 3307 TCHLLKKQAFRPQLPAYESLVQEIHHYVTSIAKAPAVQDLLTRLLQALHIDGPRSAQVAQSLLKEEASWQQSHHQFRKRLSEEYTFYPD
S.pombe 2882 VYELILSFISFMIYGIELLMFEAKRRLD-------ERSQILSTLILTLVDPSSFARSLSFDDVSNLIEQIKVLD----------
S.cerevisiae 3043 LNDILAGYIYSINFGFDLLKLQKSK----------DRASFQISPLWSMDPINISCAENVLSAYHELSRFFK----------
B.taurus 3397 ATAPLQASILQLQHGMRLVAGEVHASLQGGLAGADRLGALATSLLAFPSVGPAFPTYYAHADALCSVKSEEVLRGLSKLILKRSGGREL
H.sapiens 3396 AVSPLQASILQLQHGMRLVASELHTSLHSSMVGADRLGTLATALLAFPSVGPTFPTYYAHADTLCSVKSEEVLRGLGKLILKRSGGKEL
S.pombe 2949 ----------LNDSIRFEIYLFLASRLCSEKQHSSDTHSLANSFVLLANEFYIHNAKIKQKELEEI----------EEKNRLY
S.cerevisiae 3104 ----------KGDMEDTSIEKVLMYFLTLFKFHKRDTNLLEIFEAALYTLYSRWSVRRFRQEQEEN----------EKSNMFK
B.taurus 3486 EEKGQSPCTSREQLLNALLYLRSHVLCKGELDQRALQLFRHACQEIINEWDEQERIAQEKAEQESSLYRYRGRTPRTALSEEEEEERA
H.sapiens 3485 EGKGQKACPTREQLLMNALLYLRSHVLCKGELDQRALQLFRHVCQEIISEWDEQERIAQEKAEQESGLYRYRSRNSRTALSEEEEEERE
S.pombe 3012 RQREFNFDKNDYLKVFIN----------YDDEVEPEVEPEVVIERKRFLQLQFAFWSLYNEIYSEKMNVIPLEQLMN---T
S.cerevisiae 3167 FNDNSDDYEADFRKLFPD----------YEDTALVTNEKDISSP-ENLDDIYFKLADTYISVFDKDHDANFSSELKSGAII
B.taurus 3575 FRSQFPLHEKDFADILVEPTLEE-KGPSGGQEEEATPGPAVLSQSSMQAVMRIHQQLCLGFARSLWYQQSPPPHGTEHYLSLFLSCYQT
H.sapiens 3574 FRKQFPLHEKDFADILVQPTLEENKGTSDGQEEEAGTNPALLSQNSMQAVMLIHQQLCLNFARSLWYQQTLPPHEAKHYLSLFLSCYQT
S.pombe 3080 GSYLAKKIKVKN-----PDMIASSGFDIVSVVLMMGVKSTNERQYWTPPVYNFYSDPNPSKAIEVRDLIKIVESRAISLIKNWPENFVL
S.cerevisiae 3237 TTILSEDLKNTR-----IEELKSGSLSAVINTLDAETQSFKNTEVFGN--IDFYHDFSIPEFQKAGDIIETVLKSVLKLLKQWPEHATL
B.taurus 3663 GASLVTHFYPLMGVELNDQLLGSQLLACTLSHNTLFGEVTSDLMVKPDGPYDFYHHPNIPEARHCQPVLQGFSEAVHRLLQDWPEHPAL
H.sapiens 3663 GASLVTHFYPLMGVELNDRLLGSQLLACTLSHNTLFGEAPSDLMVKPDGPYDFYQHPNVPEARQCQPVLQGFSEAVSHLLQDWPEHPAL
S.pombe 3164 RGLKDAIDAILNLSPFSPIAEYLSKLERVFHLLSEWEKLASREYSLANEMDLIKKKIIDWRKFELSNWNNLLKLEEYKLSERVYPRLYS
S.cerevisiae 3319 KELYRVSQEFLNYPIKTPLARQLQKIEQIYTYLAEWEKYASSEVSLNNTVKLITDLIVSWRKLELRTWKGLFNSEDAKTRKSIGKWWFY
B.taurus 3752 EQLLVVMDRICSFPLSSPISKFLNGLEILLAKAQDWEENASRALSLRKHLDLVSQLIIRWRKLELNCWSMSLDNTMKRHTEKSTKHWFS
H.sapiens 3752 EQLLVVMDRIRSFPLSSPISKFLNGLEILLAKAQDWEENASRALSLRKHLDLISQMIIRWRKLELNCWSMSLDNTMKRHTEKSTKHWFS
S.pombe 3253 ILQFIILKPFFENSKFTKQN---LCESASIIVQFITDLTVGEFQLCLKCLLSFSQHAASLRICHG---IDAMLLNIYHYFEQFLSKVSE
S.cerevisiae 3408 LYESIVISNFVSEKKETAPN---ATLLVSSLNLFFSKSTLGEFNARLDLVKAFYKHIQLIGLRSSK--IAGLLHNTIKFYYQFKPLIDE
B.taurus 3841 IYQMLEKHMQEQTEEQEDDKQMTLMLLVSTLQAFIEGSSLGEFHIRLQMLLVFHCHVLLMPQVEGKDALCSVLWNLYHYYKQFFNRVQA
H.sapiens 3841 IYQMLEKHMQEQTEEQEDDKQMTLMLLVSTLQAFIEGSSLGEFHVRLQMLLVFHCHVLLMPQVEGKDSLCSVLWNLYHYYKQFFDRVQA
S.pombe 3336 AIHTQKQSLENSIKERILLMSWKDTNVYALKESAKKSHAELFKVLHRYREVLRQPVSSYLSQKHDWDS---LLDTENNSAMWVAKKVNL
S.cerevisiae 3492 RITNGKKSLEKEIDDIILLASWKDVNVDALKQSSRKSHNNLYKIVRKYRDLLNGDAKTIIEAGLLYSNENKLKLPTLKQHFYEDPNLEA
B.taurus 3930 KIVELRSPLEKELKEFVKISKWNDVSFWSIKQSVEKTHRTLFKFMKKFEAVLSEPCRSSLAEGDEEEQPDSLPRPTGAAASKESPIQRL
H.sapiens 3930 KIVELRSPLEKELKEFVKISKWNDVSFWSIKQSVEKTHRTLFKFMKKFEAVLSEPCRSSLVESDKEEQPDFLPRPTDGAASELSSIQNL
S.pombe 3422 SPSYIEKMDTEIMKLVP----------VRFSNTPTTLRLMWTLFANVEKPGSTFTNMVSNLITDARELM------KLTPET
S.cerevisiae 3581 SKNLVKEISTWSMRAAP----------LRNIDTVASN-MDSYLEKISSQEFPNFADLASDFYAEAERLR------KETPNV
B.taurus 4019 NRVLRETLLARPAAVQGAVPEQCLGAPLP-SEGALLRRLPKLLKRMRKMCLTLIQGSPLPCLVEGLDQFTGDVISFMSELQSLKVDPSA
H.sapiens 4019 NRALRETLLAQPAAGQATIPEWCQGAAPSGLEGELLRRLPKLRKRMRKMCLTFMKESPLPRLVEGLDQFTGEVISSVSELQSLKVEPSA
S.pombe 3487 INDDNLSEIKHLKSRKHLLLTETFKTLKAFGLQYRVKAGIEENLSN----------LRNLLAVIPTFPVTSLSIEKVDRSL
S.cerevisiae 3645 YTKENKKRLAYLKTQKSKLLGDALKELRRIGLKVNFREDIQKVQSS----------TTTILANIAPFNNEYLNSS--DAFF
B.taurus 4107 EKEKQRSEAKHILMQKQRALADLFKHLTKTGLSYRKGLSWARTKNPQEMLHLHPLDLQSALSIVSSTQGADSRLLTEISSSWEGCQKYF
H.sapiens 4108 EKEKQRSEAKHILMQKQRALSDLFKHLAKIGLSYRKGLAWARSKNPQEMLHLHPLDLQSALSIVSSTQEADSRLLTEISSSWDGCQKYF
S.pombe 3558 MKSLDFIPKFQTLAGHQHNDLSVPEVQKGVGLFNSMLSLQLGERAQLVEFTNELLALKNVYSEVGVNGSPLESFNNS----------
S.cerevisiae 3714 FKILDLLPKLRSAASNPSDDIPVAAIERGMALAQSLMFSLITVRHPLSEFTNDYCKINGMMLDLEHFTCLKGDIVHS----------
B.taurus 4196 YRSLARHVRLIAVLAAPAKEVGAGHVERCKGFSAHLMKVLIRQRRSLTVLTEQWILLRNLLSCVQEMDGRLTGPPVYPVAFPPQDGVQQ
H.sapiens 4197 YRSLARHARLNAALATPAKEMGMGNVERCRGFSAHLMKMLVRQRRSLTTLSEQWILLRNLLSCVQEIHSRLMGPQAYPVAFPPQDGVQQ


S.pombe
S.cerevisiae
B.taurus
H.sapiens
D/E-rich domain
MIDAS domain

Mdn1 is required for the removal of assembly factors, including Ytm1 and Rsa4 from Nsa1 and Rix1 particles, respectively (Bassler et al., 2010; Ulbrich et al., 2009). The N-terminal AAA domains in Mdn1 is likely capable of harnessing the energy from ATP hydrolysis and then provide the driving force for the remodeling of pre-60S particles. Consistent with this notion, mutants defective in ATP hydrolysis in Mdn1 was found to be non-functional in cells (Barrio-Garcia et al., 2016; Matsuo et al., 2014). On the other hand, the C-terminal MIDAS domain interacts with assembly factors Rsa4 and Ytm1, and disruption of either interaction leads to retention of early factors on ribosome intermediates (Bassler et al., 2010; Ulbrich et al., 2009). The key question is how are conformational changes or force generated by the N-terminal AAA domains transmitted to the C-terminal MIDAS domain, thus to its associated assembly factors in the pre-60S particles. This question is challenging to address as the AAA and MIDAS domains are separated by ~2000 amino acids that share little readily detectable homology with other well-studied proteins, leaving their biochemical functions unknown in this unusually large protein.

Previous negative-stain electron microscopy (EM) studies suggest a model for how Mdn1's conformational changes during its ATPase cycle are coupled to its function (Ulbrich et al., 2009). Two-dimensional (2D) averages of *S. cerevisiae* Rea1 reveal that it has a hexagonal ring-shaped structure (diameter ~15 nm) that is similar to other AAA proteins like ClpX or dynein. In addition, an elongated 'tail' (~20 nm) extending from the ring structure at different relative orientations was also observed. It was assumed that the C-

terminal MIDAS domain of Mdn1 lies at the tip of the elongated 'tail' so that the changes of relative orientations between the ring and tail will lead to movement of the MIDAS domain. In this model, the entire ~20-nm tail can transmit force, similar to the lever-arm motion described for dynein or myosin (Schmidt et al., 2015; Spudich, 2001), from the AAA domains to the MIDAS domain (Bassler et al., 2010). An alternative hypothesis was also proposed that long-range structural communication may exist between the AAA and MIDAS domains and conformational changes could propagate across the tail formed by ~2000 aa, similar to the conformational communication between the AAA ring and MTBD domain at the tip of the stalk in dyneins (Ulbrich et al., 2009). However, the exact position of the MIDAS domain cannot be determined due to the limited resolution in these studies. High-resolution data are needed to determine if these proposed models explain Mdn1 function.

The cryo-EM structure of the nucleoplasmic Rix1 particle at ~7 Å provides insights into how Mdn1 and its client protein Rsa4 bind the core of the pre-60S particle. In particular, the AAA ring of Mdn1 interacts with the core of the pre-60S particle through a limited portion of its outer edge (Figure 1-5). The C-terminal tail of Mdn1 could not be resolved in the map, leaving localization of MIDAS domain unknown. On the other hand, while Rsa4 locates in close proximity to the AAA ring of Mdn1, robust density with unknown identities can be observed between the AAA ring and the N-terminal domain of Rsa4. How Mdn1 interacts with Rsa4 in the pre-60S particle and then drive its release remains to be an open question.

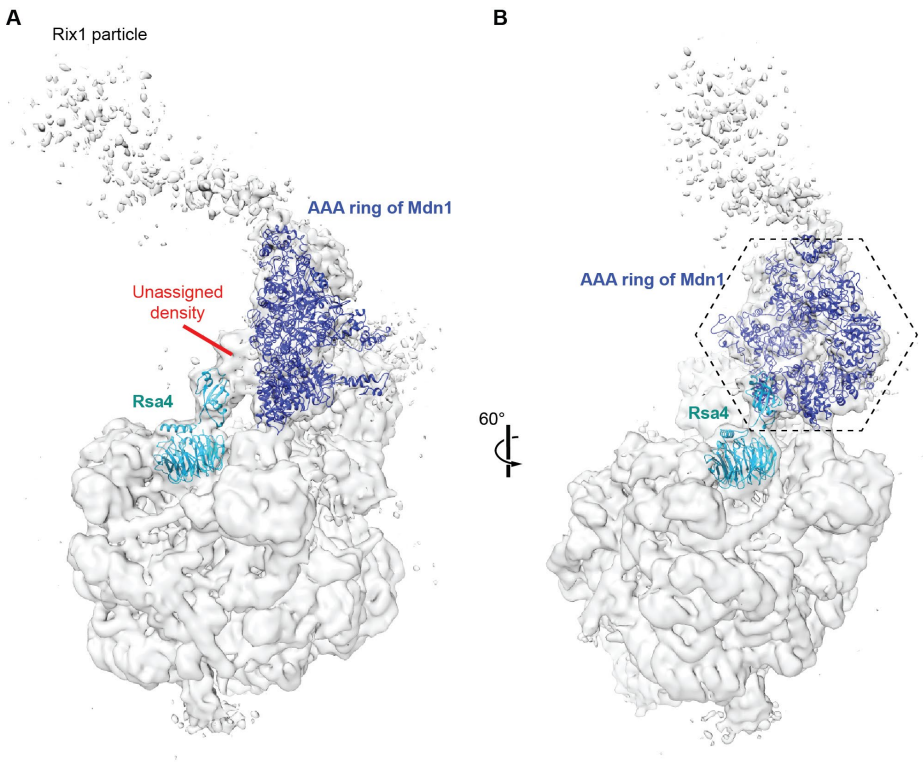


Figure 1-5. Cryo-EM density map of the Rix1 particle and models for Mdn1 and Rsa4. (A and B) Two views of the density map of the *S. cerevisiae* Rix1 particle with models for Mdn1 and Rsa4. In particular, the model for Mdn1 is based on the published model (PDB: 5FL8) for the Rix1 particle map (EMDB: 3199) and the model for Rsa4 crystal structure (cyan, *C. thermophilum*, PDB: 4WJS) is placed into the same map based on the published model (PDB: 5FL8). In (A), unassigned density with unknown identity is indicated. The hexagonal AAA ring of Mdn1 is outlined in (B). For clarity, models for other components of the Rix1 particle are omitted.

1.4 Outline

Conventional genetic analyses, using temperature sensitive strains or overexpression of dominant-negative mutants, suppress protein function over hours. These approaches may not be suitable to study the dynamic steps of

ribosome biogenesis that are completed within minutes, especially when the protein is involved in multiple stages of the same process, such as Mdn1. In order to dissect the functions of protein like Mdn1 in living cells, we need acute inhibition so that we can distinguish between direct effects of Mdn1 inhibition from cumulative defects resulting from blocking earlier stages of ribosome biogenesis. However, we currently lack selective perturbation tools to trap transient intermediates of ribosome biogenesis, in particular for the steps require the dynamic function AAA proteins. In addition, how Mdn1 performs its key remodeling function remains poorly understood due to the lack of biochemical and structural characterization of this unusually large protein.

In order to address these limitations, I sought out to (i) develop cell-permeable chemical inhibitors for Mdn1; (ii) establish purification strategies and biochemical assays for the full-length recombinant Mdn1; (iii) perform cryo-EM analysis of recombinant Mdn1 at different nucleotide states.

In the second chapter, I present the discovery of Ribozinoindoles (Rbins) as the first potent and selective chemical inhibitors of Mdn1 in S. pombe. I then focus on biochemical reconstitution and characterization of recombinant full-length S. pombe Mdn1. Using the purified protein and established biochemical assays, I then validate Mdn1 as the physiological target of Rbins in cells.

The subsequent chapters focus on structural characterization of Mdn1 in the presence of different nucleotides, nucleotide analogs and Rbin-1 using negative-stain EM (Chapter 3) and cryo-EM (Chapters 4 and 5). In particular, I

present three cryo-EM structures of full-length or truncated S. pombe Mdn1, up to the resolution of 4.0 Å, at two different nucleotide states. These structures reveal unexpected intramolecular dynamics of Mdn1, which is very different compared to dyneins, and also suggest a mechanism of how Mdn1 contribute to ribosome biogenesis.

Based on my work, I present the final chapter as a summary and an outlook in which I discuss the potential future application of Rbins to dissect the functions of Mdn1 and other biophysical experiments to study the mechanochemical cycle of Mdn1.

1.5 Methods

Protein sequence alignment and comparison. The sequences of Mdn1 orthologs (*S. pombe* Mdn1, Uniprot: O94248; *S. cerevisiae* Mdn1, Uniprot: Q12019; *B. taurus* Mdn1, Uniprot: E1BC24; *H. sapiens* Mdn1, Uniprot: Q9NU22) were aligned in ClustalW, which was also used to align and analyze the sequences of the AAA domains from *S. pombe* Mdn1 (Uniprot: O94248, aa 1-2196) and *S. pombe* dynein (Uniprot: O13290, aa 1852-4196).

CHAPTER 2

INHIBITORS OF RIBOSOME BIOGENESIS THAT TARGET MDN1

Note to readers: the results discussed below arose from a collaborative effort between myself and several colleagues in the Kapoor Lab at The Rockefeller University and the Kawashima lab at The University of Tokyo. A closely related version was published in Cell (DOI: 10.1016/j.cell.2016.08.070). With the permission of Prof. Kapoor, I have included relevant studies done by the others in the 2.2 Background section as they explain the rationale of my work. In 2.3 Results section, I focus mostly on work done by myself unless otherwise noted in the figure legend. In addition, I included Appendix 2 as a summary of collaboration between myself, several colleagues in the Kapoor lab and Tri-Institutional Therapeutic Discovery Institute on the development of antifungal reagents based on our discovery.

2.1 Introduction

Cell-permeable chemical inhibitors can be powerful tools for examining dynamic cellular processes, such as ribosome biogenesis, as the functions of target proteins can be blocked within minutes. Currently, the only known chemical inhibitor that directly targets eukaryotic ribosome assembly factors is diazaborine, a compound that is active only at ~0.4 mM in *S. cerevisiae* (Loibl et al., 2014), a concentration at which selective target inhibition may be difficult to achieve. Moreover, since diazaborine blocks cytoplasmic steps (i.e. pre-60S maturation) of ribosome biogenesis, we lack chemical probes for the distinct assembly steps that occur in the nucleolus and nucleus. Lamotrigine is another chemical inhibitor of ribosome assembly factors that has been recently described (Stokes et al., 2014). However, this compound has been shown to only block ribosome biogenesis in *E. coli* at low temperatures (Stokes et al., 2014).

In this study, we identify ribozinoindoles (or Rbins), as potent, reversible and specific inhibitors of Mdn1. Systematic genetic analyses of Rbins' sensitivity and resistance in fission yeast, along with biochemical characterization of Mdn1's ATPase activity, indicate that Rbins directly and specifically inhibit Mdn1 function *in vitro* and in cells.

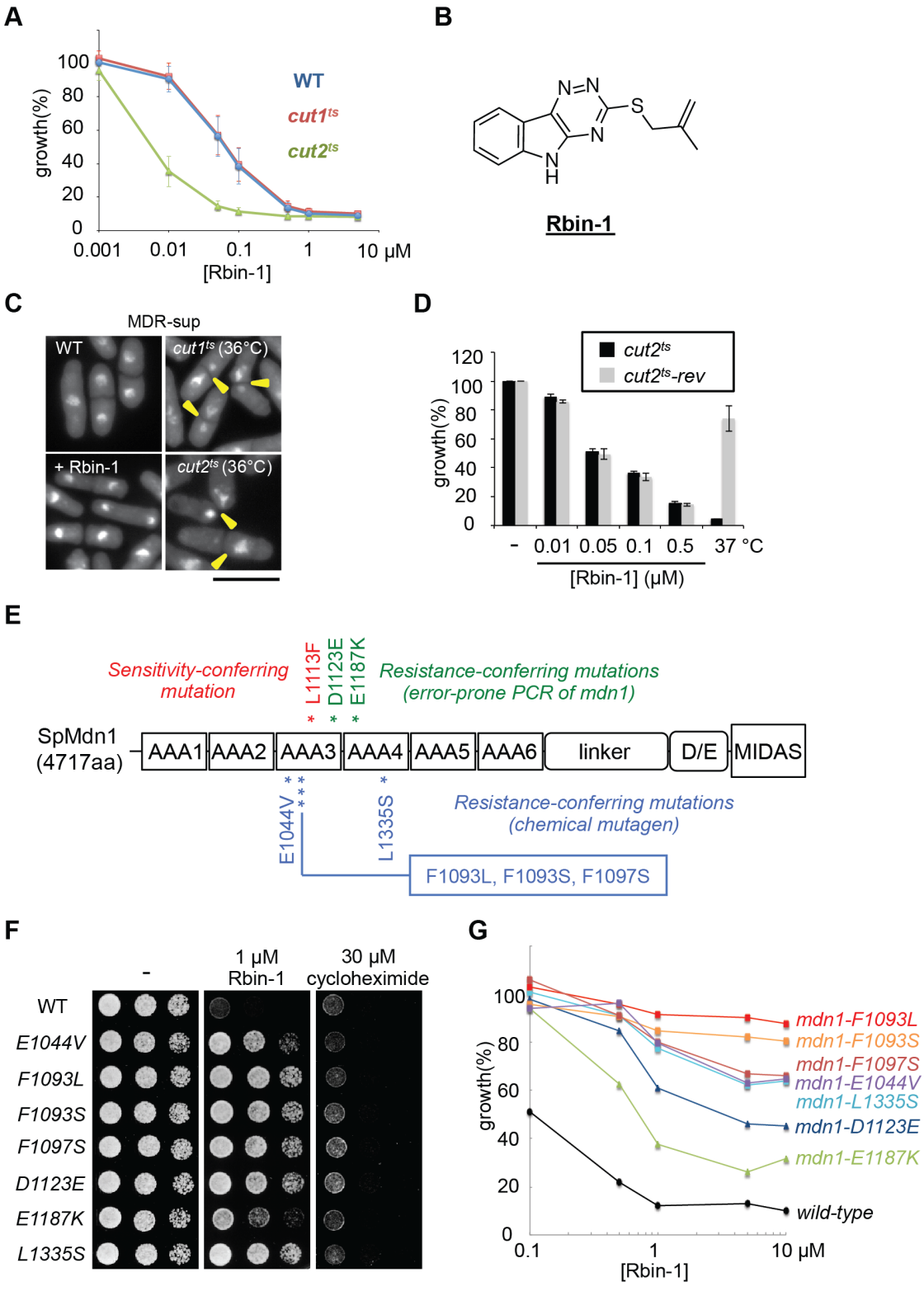
2.2 Background: Discovery of a chemical inhibitor of Mdn1, Rbin-1

To identify cell-permeable chemical probes of essential cellular processes we have developed fission yeast as a model system that allows us to efficiently combine genetic and chemical approaches (Aoi et al., 2014; Kawashima et al., 2012). In particular, we have generated fission yeast strains (named 'MDR-sup' strains) lacking critical factors for multi-drug resistance (or MDR) and have used them for chemical screens that mimic synthetic lethal genetic screens (Kawashima et al., 2012, 2013). We hypothesized that compounds that reveal enhanced toxicity to strains with a particular mutation, compared to a control strain, are likely to be more selective for a single protein target. Consistent with this hypothesis, our use of this strategy identified a selective inhibitor for Aurora kinase, a key cell cycle regulator (Kawashima et al., 2013). From a similar 'chemical synthetic lethal' screen carried out with a 10,353-member library of diverse chemicals we identified a triazinoindole-based heterocycle, which we named ribozinoindole-1 (or Rbin-1), that was more toxic to the MDR-sup cells that contained a mutation in *cut2,* compared to those with a wild-type *cut2*$^{+}$ or a mutation in *cut1* (Figures 2-1A and B). Both

Cut1 and Cut2 are essential proteins needed for faithful chromosome segregation (Funabiki et al., 1996). Subsequent analyses indicated that Rbin-1's 'synthetic lethality' was not due to a chromosome mis-segregation ('*cut*') phenotype (Figure 2-1C). In addition, we found that reverting the *cut2* mutation to wild-type *cut2*$^{+}$ did not suppress Rbin-1 sensitivity (Figure 2-1D). Together, these data indicate that the increased sensitivity to the chemical inhibitor was due to other 'background' mutations that were likely introduced when the mutant strain was first generated (Hirano et al., 1986).

To examine the mode-of-action of Rbin-1 we back-crossed the *cut2* mutant MDR-sup strain with wild-type *cut2*$^{+}$ MDR-sup and then analyzed the entire genome for mutations. Sequencing analysis identified 4 genes (SPCC737.08, SPAC14C4.01c, SPAC26A3.01, and SPAC26A3.05) bearing mutations, among which we found that a point mutation (Leu1113Phe) in the *mdn1* (SPCC737.08) gene confers sensitivity to Rbin-1 in engineered fission yeast strain, suggesting that this residue is the background mutation that increases the sensitivity of fission yeast to Rbin-1.

Figure 2-1. Characterization of Rbin-1, a potent chemical inhibitor of cell growth. (A) MDR-sup (WT), MDR-sup *cut1-22* (*cut1^{ts}*), and MDR-sup *cut2-364* (*cut2^{ts}*) cells were incubated with different concentrations of Rbin-1 (17 h, 29°C). Growth (%) is presented relative to DMSO-treated cells (mean±SD, n=3 independent experiments). (B) Chemical structure of Rbin-1. (C) Nuclear morphology of MDR-sup cells, MDR-sup *cut1-22* (*cut1^{ts}*) and MDR-sup *cut2-364* (*cut2^{ts}*) cells at 36°C (restrictive temperature), and MDR-sup cells treated with 10 μM Rbin-1 (+Rbin-1) were examined. Representative images are shown. Cells with '*cut*' phenotype, observed in temperature-sensitive *cut1* or *cut2* mutants, are highlighted (yellow arrowheads). Scale bars, 10 μm. (D) *cut2-364* (*cut2^{ts}*) and a *cut2^{ts}-rev* strain (with the *cut2-364* mutation reverted to wild-type *cut2^{+}*) were incubated (17 h) with indicated concentrations of Rbin-1 at 29°C or untreated at 37°C. Growth (%) is presented relative to DMSO-treated cells incubated at 29°C (mean±range, n=2 independent experiments). (E) Domain organization of the *S. pombe* Mdn1 protein (SpMdn1). The AAA domains, linker, Asp/Glu-rich region and MIDAS domain are highlighted. Locations of a Rbin-1-sensitivity-conferring mutation (red), Rbin-1-resistance-conferring mutations obtained by error-prone PCR of *mdn1* gene (green), and Rbin-1-resistance-conferring mutations obtained using a chemical mutagen (blue) are indicated. (F) Serial dilutions of the indicated *mdn1* mutants were spotted onto YE4S plate, or YE4S plate containing indicated compounds, and incubated at 29°C for two days. (G) Wild-type and indicated *mdn1* mutants were incubated for 17 h at 29°C with indicated concentrations of Rbin-1. Growth (%) is presented relative to DMSO-treated cells.


A
100
80
60
40
20
0
growth(%)
WT
cut1ts
cut2ts
0.001
0.01
0.1
1
10 μM
[Rbin-1]
B
N=N
S
N
N
H
Rbin-1
C
MDR-sup
WT
cut1ts (36°C)
+ Rbin-1
cut2ts (36°C)
D
cut2ts
cut2ts-rev
120
100
80
60
40
20
0
growth(%)
-
0.01
0.05
0.1
0.5
37 °C
[Rbin-1] (μM)
E
Sensitivity-conferring mutation
L1113F
D1123E
E1187K
Resistance-conferring mutations (error-prone PCR of mdn1)
SpMdn1 (4717aa)
AAA1
AAA2
AAA3
AAA4
AAA5
AAA6
linker
D/E
MIDAS
E1044V
L1335S
Resistance-conferring mutations (chemical mutagen)
F1093L, F1093S, F1097S
F
-
1 μM Rbin-1
30 μM cycloheximide
WT
E1044V
F1093L
F1093S
F1097S
D1123E
E1187K
L1335S
G
100
80
60
40
20
0
growth(%)
mdn1-F1093L
mdn1-F1093S
mdn1-F1097S
mdn1-E1044V
mdn1-L1335S
mdn1-D1123E
mdn1-E1187K
wild-type
0.1
1
10 μM
[Rbin-1]

To further address the mode of action of Rbin-1, we chemically mutagenized the MDR-sup fission yeast cells and isolated resistant clones that can grow in the presence of Rbin-1. Each of the 13 Rbin-1 resistant clones contained one of five different point mutations in the *mdn1* gene (Figures 2-1E and F). In separate experiments using a PCR-based random mutagenesis strategy, we isolated two additional *mdn1* mutations that confer Rbin-1-resistance (Figures 2-1E and F). We next confirmed, using overexpression on plasmids in MDR-sup cells and by replacement of endogenous *mdn1* in wild-type cells, that all the mutations we identified were sufficient to confer resistance to Rbin-1 (Figures 2-1F and G). Importantly, these mutations did not confer resistance to cycloheximide, another chemically unrelated inhibitor (Figure 2-1F). Interestingly, all seven of these resistance-conferring mutations and the sensitivity-conferring mutation clustered around two (AAA3 and AAA4) of the six AAA domains in Mdn1 (Figure 2-1E). These data are consistent with a hypothesis that Rbin-1 target Mdn1 or another protein along the same pathway of Mdn1 function in cells so that mutations in *mdn1* can affect the compound activities.

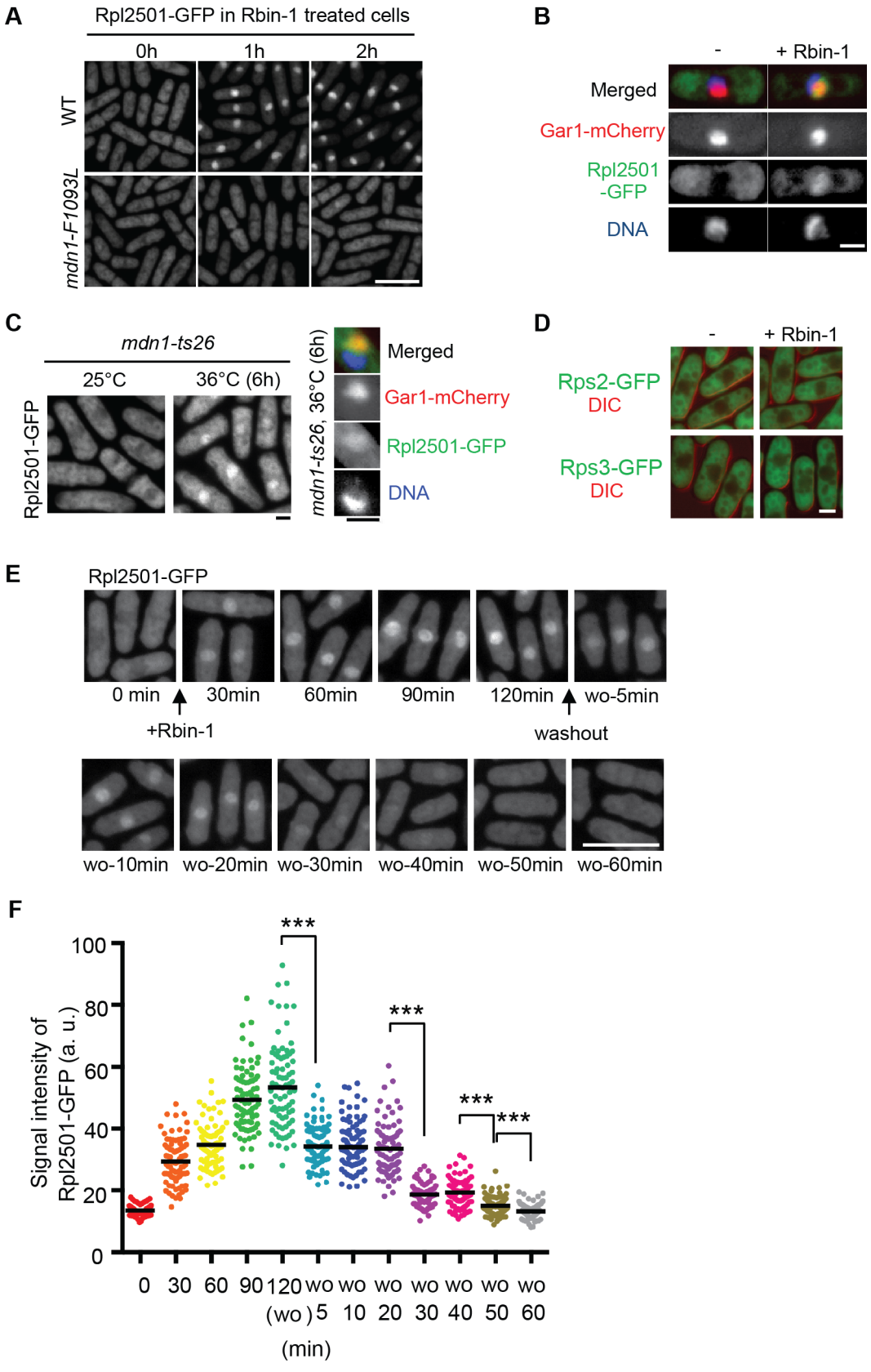
As Mdn1 has been linked to remodeling of precursors of 60S ribosome subunits in budding yeast before, we next examined if Rbin-1 indeed perturb ribosome biogenesis in fission yeast cells. We found that nuclear export of the pre-60S subunit, which can be tracked using Rpl2501-GFP, a homolog of budding yeast Rpl25 (Hurt et al., 1999), was blocked in Rbin-1-treated cells

(Figures 2-2A and B). Importantly, this defect was not observed in Rbin-1-treated *mdn1-F1093L* mutant cells (Figure 2-2A). In addition, this defect was also observed at the restrictive temperature in a *mdn1* temperature-sensitive mutant strain (Figure 2-2C). In contrast, the nuclear export of the pre-40S subunits that were tracked using two different reporters, Rps2-GFP and Rps3-GFP (Galani et al., 2004), was not blocked in Rbin-1-treated cells (Figure 2-2D), suggesting that Rbin-1 perturb specific steps of ribosome biogenesis that leads to block of the nuclear export of pre-60S particles but not pre-40S particles. Further, by comparing Rpl2501-GFP with Gar1-mCherry (nucleolus marker) and DAPI (DNA marker), we found that pre-60S particles accumulated in the nucleolus upon Mdn1 inhibition (Figures 2-2B and C). This inhibition can be observed within 30 mins after addition of Rbin-1 and reached a maximum level in 90-120 min (Figures 2-2E and F). Relief from Rbin-1 treatment revealed stepwise reductions in the nucleolar signal, with Rpl2501-GFP levels reducing substantially (~47%) within 5 min of washout, then another ~37% between 20-30 min, and ~15% between 40-60 min (Figure 2-2F), suggesting three distinct ribosome intermediates were trapped upon the addition of Rbin-1.

The data showing that Rbin-1 treatment phenocopies the loss of function of Mdn1 in the temperature-sensitive mutation strain are consistent with a hypothesis that Rbin-1 directly inhibit the activity of Mdn1 thus suppresses its ribosome biogenesis function and the cellular growth.

Considering that Mdn1 is an AAA ATPase, it is possible that Rbin-1 can inhibit the ATPase activity of Mdn1.

Figure 2-2 Rbin-1 causes blockage of nuclear export of pre-60S particles and phenocopies the temperature-sensitive-mdn1 strain at non-permissive temperate in fission yeast. (A) Rpl2501-GFP signals in wild-type (WT) and *mdn1-F1093L* cells with or without Rbin-1 treatment (1 μM) were examined. Scale bars, 10 μm. (B) Rpl2501-GFP signals (green) were compared with nucleolar marker Gar1-mCherry (red) and DNA (blue). Scale bars, 2 μm. (C) Rpl2501-GFP signals in *mdn1-ts26* cells at 25°C or 36°C for 6 hours were examined. Also shown are images of Rpl2501-GFP, Gar1-mCherry (red), and DNA stained with DAPI (blue). Scale bars, 2 μm. (D) The localization of markers for the 40S pre-ribosome, Rps2-GFP and Rps3-GFP (green, overlaid on the corresponding DIC image, red), was examined in the presence or absence of Rbin-1 (1 μM). Scale bars, 2 μm. (E, Rpl2501-GFP localization was examined in asynchronous cells (0 min), 1 μM Rbin-1-treated cells (30-120 min), and in cells after washing out Rbin-1 (wo-5-60 min) at 29°C. Representative images are shown in (E). Scale bars, 10 μm. The signal intensity of Rpl2501-GFP in the nucleolus was measured in (F). Averages are indicated by black bars (n=80 cells). Asterisks indicate unpaired two-tailed Student's t-test significance value: ***P<0.001.


A
Rpl2501-GFP in Rbin-1 treated cells
0h
1h
2h
WT
mdn1-F1093L
B
-
+ Rbin-1
Merged
Gar1-mCherry
Rpl2501
-GFP
DNA
C
mdn1-ts26
25°C
36°C (6h)
Rpl2501-GFP
mdn1-ts26, 36°C (6h)
Merged
Gar1-mCherry
Rpl2501-GFP
DNA
D
-
+ Rbin-1
Rps2-GFP
DIC
Rps3-GFP
DIC
E
Rpl2501-GFP
0 min
30min
60min
90min
120min
wo-5min
+Rbin-1
washout
wo-10min
wo-20min
wo-30min
wo-40min
wo-50min
wo-60min
F
Signal intensity of Rpl2501-GFP (a. u.)
100
80
60
40
20
0
***
***
***
***
0
30
60
90
120
(wo)
wo 5
wo 10
wo 20
wo 30
wo 40
wo 50
wo 60
(min)

2.3 Results

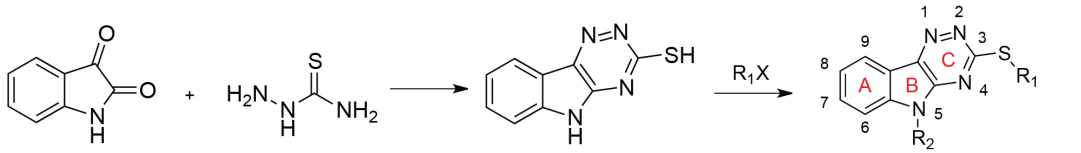
2.3.1 Structure-activity-relationship analysis of Rbin-1

We next optimized the chemical inhibitor's potency by generating a focused collection of analogs. Seventeen compounds were synthesized by condensing isatin (or derivatives) with hydrazinecarbothioamide, and further elaborated using halogenated aliphatic and aromatic moieties (Figure 2-3). This procedure yielded Rbin-1 analogs in >95% purity and 7-65% overall yield (see 2.5 Method section). Dose-dependent analyses were used to determine the potency of these analogs in growth assays (Figure 2-3B). Our studies reveal that alkyl substituents at the 3-thiol position result in compounds with activities that are reduced compared to that of Rbin-1 (Analogs 2-5 in Figure 2-3B). A benzyl moiety at the 3-thiol position increased potency but reduced solubility (Rbin-3 or Analog 6 in Figure 2-2B) while a pyridine moiety abolishes the activity (Analog 7). Larger aromatic substitutions at this position reduced compound activity (Rbin-5 or Analogs 8 in Figure 2-3B). Modification of the benzyl moiety by nitro, methoxy or iodine leads to inactive compounds (Analogs 9-11), while the ortho-fluoride benzyl Rbin has very similar activity as Rbin-3 (Analogs 6 and 12). A homo-propargyl modification at the indole nitrogen (R_2 position) also resulted in an inactive analog (Rbin-6 or Analog 21 in Figure 2-3B). We found that the most striking changes in activity resulted from modifications of the A-ring (Figure 2-3A). In particular, an analog (Rbin-2 or Analog 15 in Figure 2-3B) with a bromine substituent at position-7 was ~10-fold more active than Rbin-1 (GI_{50}=14 ± 1 nM (Rbin-2); 136 ± 7 nM (Rbin-1),

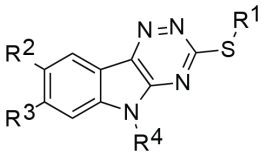
n=4, mean±SD). In contrast, an iodine substituent at position-8 suppressed activity (Rbin-7 or Analogs 22 and 23 in Figure 2-3B). Combinations of the bromine substituent at position-7 and benzyl or alkyl substituents on the thiol atom did not yield more active compounds (Analogs 16-17).

Figure 2-3. Structure-activity-relationship analysis of Rbin-1 analogs. (A) Synthesis scheme for RBin analogs. (B) Structure-activity relationship of Rbin analogs. Compounds were added to wild-type cells in diluted culture (in logarithmic phase, OD_{600} ~ 0.01) and the mixture was incubated for ~18 h at 29°C till OD_{600}~1 for DMSO control. Half maximum growth inhibition (GI_{50}, mean±SD, n=3 independent experiments) was determined by fitting relative growth to a four-parameter sigmoidal dose-response curve in PRISM. For compounds that inhibited growth by less than 50% at 500 nM and 2500 nM, we indicate GI_{50} to be ”≥ 500 nM” and ”≥ 2500 nM” respectively. For a few compounds, dose-dependent inhibition could not be readily fit to standard equations, likely due to limited compound solubility (Note: *: The number of analogs were assigned in the published manuscript and analogs I synthesized, except Analog 21, are shown; **: Analogs 21, synthesized by Dr. Anupam Patgari, is shown since it was used later in the biochemical assay as a negative control; ***: The analysis of the activities were carried out by myself and Dr. Patgari together).

A



B



Rbin1 analogs(*)	R^1	R^2	R^3	R^4	GI_{50} (nM)***
Rbin-1 or Analog 1		H	H	H	136 ± 7
Analog 2		H	H	H	≥ 500
Analog 3		H	H	H	≥ 500
Analog 5		H	H	H	≥ 500
Rbin-3 or Analog 6		H	H	H	40 ± 7
Analog 7		H	H	H	≥ 2500
Rbin-5 or Analog 8		H	H	H	≥ 2500
Analog 9		H	H	H	n.d.
Analog 10		H	H	H	≥ 2500
Analog 11		H	H	H	≥ 2500
Rbin-4 or Analog 12		H	H	H	28 ± 6
Rbin-2 or Analog 15		H	Br	H	14 ± 1
Analog 16		H	Br	H	≥ 500
Analog 17		H	Br	H	≥ 2500
Rbin-6 or Analog 21(**)		H	H		≥ 2500
Rbin-7 or Analog 22		I	H	H	≥ 2500
Analog 23		I	H	H	≥ 2500

I next confirmed that Rbin-2 activity against the strain expressing the resistance-conferring mutation (*mdn1-F1093L*) was substantially reduced, while activity against the strain expressing the sensitivity-conferring mutation (*mdn1-L1113F*) was increased relative to the wild-type strain (Figure 2-4). Together, these tests identify Rbin-2 as a potent Rbin-1 analog that is likely to have a similar mechanism of action.

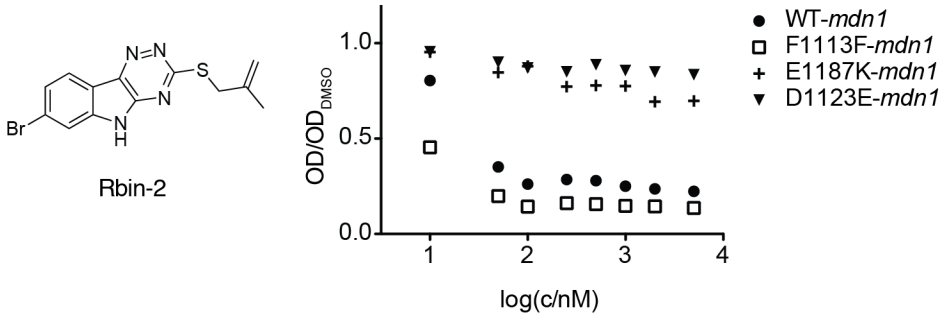


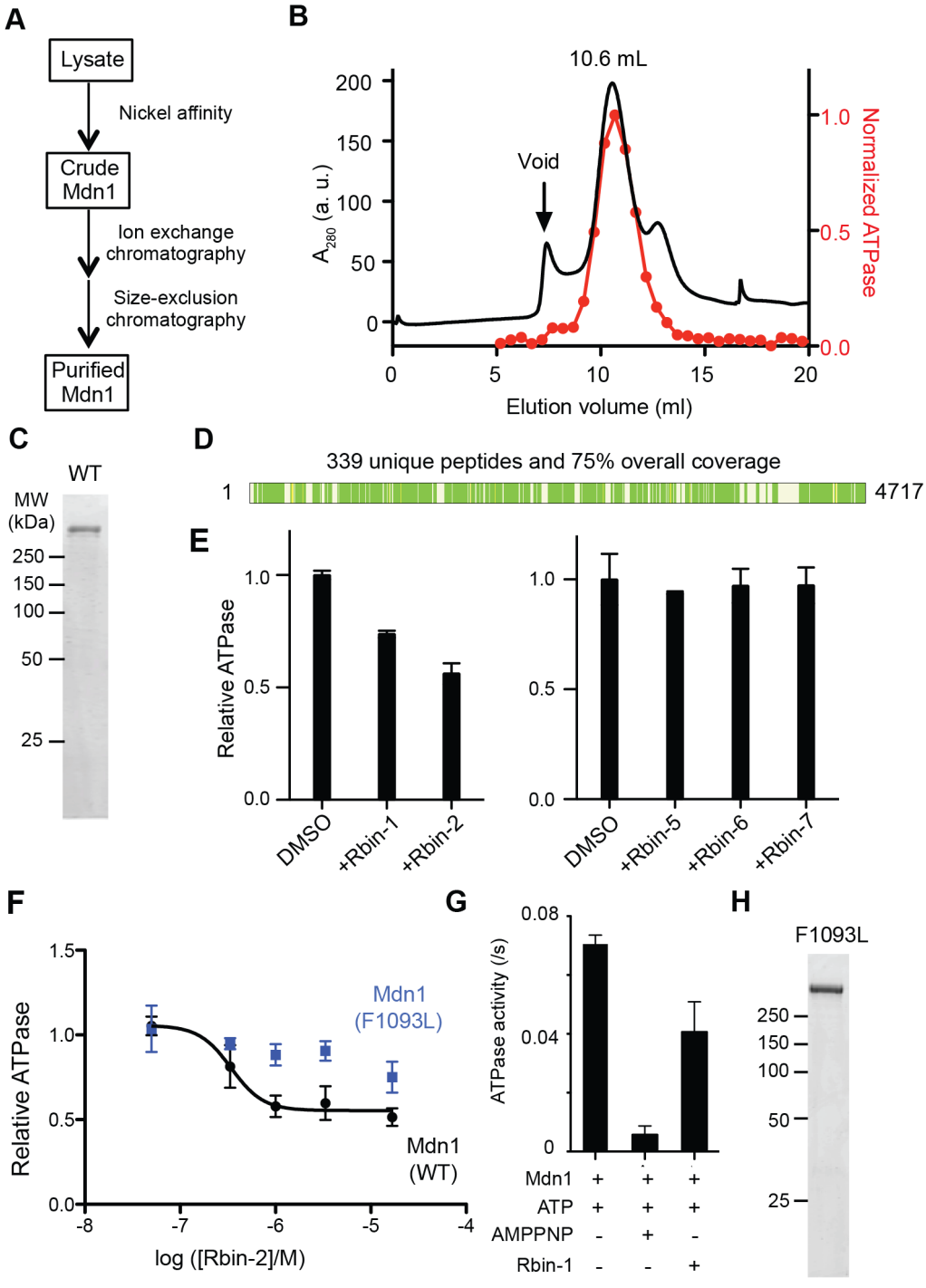
Figure 2-4. Growth inhibition of wild-type or mutant *S. pombe* strains by Rbin-2. The growth of wild-type and mutant S. pombe strains in the presence of different concentrations of Rbin-2 were analyzed by the OD values. The normalized growth relative to the DMSO control is plotted. Rbin-2 is more toxic to the Rbin-1-sensitive strain (with L1113F mutation in *mdn1*) and significantly less toxic in Rbin-1-resistant strains (with E1187K or D1123E mutation in *mdn1*).

2.3.2 Biochemical characterization of Mdn1 *in vitro*

To date, only *S. cerevisiae* Rea1 has been expressed and purified as a recombinant form. However, its ATPase activity has not been measured. *S. pombe* Mdn1 has not been purified nor characterized. To analyze *S. pombe* Mdn1's ATPase activity and examine inhibition by Rbin-1 we need to purify recombinant Mdn1 for *in vitro* biochemical assays. I employed the insect cell

system to generate recombinant Mdn1. After significant optimization, I developed a multi-step strategy, which included nickel affinity, ion-exchange and size-exclusion chromatography (Figure 2-5A). In the final purification step, Mdn1 eluted from a size-exclusion column at a volume (10.6 ml, Void: 7.5 ml) consistent with its large size (M.W. ~542 kDa for hexahistidine-tagged Mdn1) and elongated shape based on analyses of Rea1 (Figures 2-5B and C) (Nissan et al., 2004; Ulbrich et al., 2009). A typical purification yielded ~0.1 mg of recombinant Mdn1 per 3L cell culture. Mass spectrometry confirmed that the single band resolved by SDS-PAGE was Mdn1, as we were able to identify 339 unique peptides that covered >70% of the Mdn1 sequence (Figure 2-5D). In particular, the identified peptides confirmed the presence of ~80 amino acids at Mdn1's C-terminus. These data, along with the N-terminal tag-based isolation, indicate that we are able to purify full-length recombinant Mdn1.

Figure 2-5. Rbins inhibit Mdn1's ATPase activity *in vitro*. (A) Schematic for strategy used to purify recombinant full-length Mdn1. (B) The size-exclusion chromatography profile for Mdn1 (black trace). The activity of different fractions from this chromatography step was analyzed using a radioactive ATPase assay ([MgATP]= 0.1 mM). Activity was normalized to the most active fraction (red dots: each fraction, trace: interpolation). (C) SDS-PAGE analysis of purified full-length Mdn1 (WT) (Coomassie stain). (D) Peptides identified by mass spectrometry of the purified protein are indicated (green bars, schematic generated using Proteome Discoverer 1.4, Thermo Scientific). (E) Mdn1's activity in the presence of Rbin analogs (1 μM) was tested using an NADH-coupled ATPase assay ([MgATP]= 1 mM). The average relative activity and range (bars) are indicated (n=2 independent experiments). (F) Dose-dependent inhibition of the steady-state ATPase activity of wild-type Mdn1 and resistant F1093L Mdn1 by Rbin-2 (mean±SD, n=4 independent experiments). An apparent EC_{50} was estimated for the inhibition of wild-type Mdn1 using a sigmoidal dose-response curve. Using similar equations we were unable to properly fit the small decrease in activity of F1093L-Mdn1 across the concentration range tested. Unpaired T-tests (without Welch's correction) of the measured activity for WT and F1093L-Mdn1, at the highest three inhibitor concentrations tested, indicate that the differences in values are statistically significant (p= 0.0020 (at 16.7 μM), p= 0.0021 (at 3.33 μM) and p= 0.0019 (at 1.0 μM)). (G) The ATPase activity of Mdn1 in the presence Rbin-1 (1 μM) or AMP-PNP (2 mM) using a radioactive ATPase assay ([MgATP] = 0.1 mM). (H) SDS-PAGE analysis of purified full-length F1093L-Mdn1 (Coomassie stain).


A
Lysate
Nickel affinity
Crude Mdn1
Ion exchange chromatography
Size-exclusion chromatography
Purified Mdn1
B
10.6 mL
Void
A_{280} (a. u.)
Normalized ATPase
Elution volume (ml)
C
WT
MW (kDa)
250
150
100
50
25
D
339 unique peptides and 75% overall coverage
1
4717
E
Relative ATPase
DMSO
+Rbin-1
+Rbin-2
+Rbin-5
+Rbin-6
+Rbin-7
F
Relative ATPase
Mdn1 (F1093L)
Mdn1 (WT)
log ([Rbin-2]/M)
G
ATPase activity (/s)
Mdn1
ATP
AMPPNP
Rbin-1
H
F1093L

I then analyzed the ATPase activities of different fractions eluted from size-exclusion chromatography utilizing a radioactive-ATP based assay. The levels of ATPase activities closely matched the UV profile of the elution peak of Mdn1 (Figure 2-5B, red trace). We next examined the steady-state ATPase activity of Mdn1 using an NADH-coupled enzyme assay. The specific activity of Mdn1 was ~1.0 s^{-1} at 1 mM ATP. These data indicate that Mdn1 is an ATPase with specific activity comparable to that of other AAA+ proteins, such as VPS4 and dynein (Kon et al., 2004; Yang et al., 2015).

We next examined the inhibition of Mdn1's steady-state ATPase activity by Rbin-1 and analogs. Two of the active analogs (Rbin-1 and Rbin-2) inhibited the ATPase activity by ~40% at 1 μM (Figure 2-5E). Under similar conditions, analogs, i.e. Rbin-5, 6 and 7, that were inactive in cell-based assays did not suppress ATPase activity (Figure 2-5E). As the amounts of Mdn1 were limiting, we focused on the dose-dependent inhibition of Mdn1 activity by Rbin-2, the most potent analog. While increasing concentrations of Rbin-2 resulted in greater inhibition, the dose-dependent inhibition saturated at ~50%, consistent with inhibition of a subset of the ATPase domains in Mdn1. For the observed dose-response curve, we can estimate an apparent EC_{50} (~0.3 μM, Figure 2-5F). For comparison, we find that almost complete inhibition of the ATP hydrolysis (in a radioactive assay) can be achieved by a non-hydrolyzable ATP analog (AMPPNP, 2 mM), while Rbin-1 (1 μM) only inhibited activity by ~40% (Figure 2-5G) in the radioactive-ATP assay. It is

likely that Rbin analogs only suppresses a subset of the ATPase sites in Mdn1, while the cellular function of Mdn1 requires activities of all ATPase sites so the partial inhibition *in vitro* leads to full inhibition of the growth of yeast.

To establish that Mdn1 is Rbins' direct target we generated full-length recombinant Mdn1 containing a mutation (F1093L) that confers resistance to the Rbins in cell-based assays (hereafter, F1093L-Mdn1, Figure 2-5H). We find that F1093L-Mdn1 is an active enzyme whose ATPase activity (1.5 ± 0.2 ATP s^{-1}) in the steady-state assay is comparable to that of wild-type Mdn1. Importantly, Rbin-2 does not strongly suppress the ATPase activity of the F1093L-Mdn1 (Figure 2-5F). Together, these data establish what we refer to as 'gold standard' proof that the physiological target of Rbins is Mdn1, as the same single-point mutation can suppress inhibition by Rbins in both cell-based and *in vitro* biochemical activity assays (Wacker et al., 2012).

2.4 Discussion

Here, we report the discovery of ribozinoindoles, potent and selective chemical inhibitors for Midasin, an essential AAA+ protein required for eukaryotic ribosome biogenesis. Matched inhibitor-sensitive and inhibitor-resistant cells allow systematic analyses of dose-dependent target-specific effects of the chemical probe, addressing a major potential limitation of the use of chemical inhibitors to examine cellular mechanisms. Further, the fast timescale of Mdn1 inactivation and activation by Rbins is well suited to dissect

its role in coordinating the spatial and temporal dynamics of ribosome biogenesis.

Importantly, the timescale of Midasin inactivation and activation by Rbin-1 (less than 30 min) is significantly faster than what we can achieve using genetic approaches (~3-6 h using our temperature-sensitive allele). In particular, the temporal control over target function that can be achieved with these chemical probes is evident from studies of Rpl2501 localization, which are altered in live cells within 30 min of inhibitor washout (Figure 2-2E). Analyses of the timescales over which different phenotypes accumulate following chemical inhibitor treatments can guide interpretations and help develop models for target protein function in a dynamic, ordered, multi-step cellular process, such as eukaryotic ribosome biogenesis. It is likely that Rbins, which allow acute and reversible inhibition of Mdn1, will help dissect these fast assembly dynamics.

Mdn1 and the closely related motor protein dynein are the only known members of the AAA+ protein family that have all six AAA domains on a single polypeptide, a feature that likely allows functional specialization of each of the multiple AAA domains (Carter, 2013). In the case of dynein, which has been extensively studied over the last couple of decades, two of its ATPase sites contribute to the overall ATP hydrolysis and directional transport activity (Carter, 2013). As Rbin-resistance and -sensitivity conferring mutations cluster at one ATP binding site in Mdn1, we favor the possibility that Rbins act through inhibiting the activity of the ATPase site formed by AAA3 and AAA4

domains. Additional structural and biochemical studies will be needed to test this hypothesis.

Our discovery of a specific and potent chemical inhibitor of Midasin, along with the findings that dynein and p97/VCP can be selectively inhibited by cell-permeable low molecular weight compounds (Anderson et al., 2015; Chou et al., 2011; Firestone et al., 2012), indicate that drug-like inhibitors can be developed for AAA+ proteins. These discoveries lay the foundation for developing pharmacophore models for rationally designing chemical inhibitors for AAA+ proteins for which we lack chemical probes, such as proteins involved in DNA replication, mitochondrial function or cytoskeleton organization.

2.5 Methods

Methods related to the generation of *S. pombe* strains can be found in Kawashima *et al.* 2016.

SAR analysis of Rbin analogs using a growth assay of pombe.

S. pombe strains were grown to exponential phase and diluted to OD_{600} ~ 0.001. 1 µL DMSO or DMSO solution of Rbins (1000x) were mixed with 1 mL diluted culture. The cells were placed in a shaker (220 rpm) for 18 hours at 29 °C till OD_{600} of DMSO control reached ~1. Relative growth was calculated by dividing the measured OD at a specific concentration by the OD for the DMSO control. Half maximum growth inhibition (GI_{50}) was determined by

fitting relative growth to a four-parameter sigmoidal dose-response curve in PRISM.

Purification of recombinant *S. pombe* Mdn1

The cDNA encoding fission yeast Mdn1 was cloned into pFastBac HTC vector (Invitrogen). We used the Bac-to-Bac system (Invitrogen) to generate recombinant baculovirus. High Five cells (Life Technologies) were grown in Sf-900 II SFM (Life Technologies 10902-096) with 1X Antibiotic-Antimycotic (Life Technologies) to ~2.5 million/mL and then infected (1:80 dilution of P2 virus). The cells were harvested 60 h after infection. All of the following steps were done on ice or at 4 °C. The cells were lysed by sonication in an equal volume of lysis buffer (50 mM Tris [pH 7.5], 400 mM NaCl, 20 mM imidazole, 1 mM $MgCl_2$, 5 µM 2-mercaptoethanol, 200 µM ATP, 3 U/mL benzonase, 1X Roche complete protease inhibitor without EDTA, 10% glycerol). The homogenized lysate was then centrifuged at 55,000 rpm for 1 h. The supernatant was incubated with Ni-NTA beads (Qiagen) for 40 min. The beads were extensively washed using Washing buffer (50 mM Tris [pH 7.5], 400 mM NaCl, 20 mM imidazole, 1 mM $MgCl_2$, 5 µM 2-mercaptoethanol, 200 µM ATP, 10% glycerol). The protein was then eluted by high-imidazole buffer (20 mM Tris [pH 7.5], 120 mM NaCl, 300 mM imidazole, 1 mM $MgCl_2$, 5 µM 2-mercaptoethanol, 200 µM ATP). The eluted fraction was filtered and loaded onto a Mono Q column 5/50 GL (GE Healthcare Life Sciences). The protein was eluted around 400 mM NaCl and the fractions were collected and analyzed by SDS-PAGE. The relevant fractions were pooled and then concentrated using Amicon Ultra-4

Centrifugal Filter Units. The concentrated sample was then loaded onto a Superose 6, 10/300 GL column (GE Healthcare Life Sciences) using FPLC SEC buffer (20 mM Tris [pH 7.5], 150 mM NaCl, 1 mM $MgCl_2$, 1mM EGTA and 1 μM DTT. Peak fractions were collected and ATPase assay was carried out directly with fresh protein. A typical yield of these preparations is ~ 0.1 mg/ 3 L culture. Freshly purified Mdn1 was used immediately for EM studies.

Radioactive ATPase Assay

Radioactive γ-P32-ATP (PerkinElmer, BLU002Z250UC) was added to 600 μM MgATP (pH 7.0) solutions at volume ratios of 1:1000-1:300, depending on the lifetime of the radioactive reagent. The total volume of each reaction was 12 μL, including 6 μL of protein from size exclusion chromatography fractions (final concentration ~ 0-50 nM for different fractions, peak fractions were used for Rbin-1 and AMPPNP inhibition experiment in Figure 2-5), 4 μL FPLC SEC buffer with 3 mM Na_2SO_4 and 2 μL MgATP (final concentration = 100 μM). The reactions were then incubated at room temperature for 30~60 min before quenching with 12 μL 0.2 M EDTA. 1 μL from each reaction mixture was spotted on TLC PEI cellulose F plates (Millipore, 105579). The TLC buffer contained 0.15 M formic acid and 0.15 M lithium chloride. The TLC plates were then imaged by the Typhoon Scanner 9400 (GE Healthcare Life Sciences). ImageJ was used to calculate the densitometric ratio of the spots corresponding to radioactive free phosphate and ATP to calculate the percent of ATP hydrolyzed.

NADH-coupled steady-state ATPase assay

For each steady-state ATPase reaction, the final volume was 30 µL. The final concentration of protein was 50 nM. The FPLC SEC buffer was used to dilute the sample. NADH (N7410), phosphoenol pyruvic acid monopotassium salt (P7127), D-lactic dehydrogenase (L3888) and pyruvate kinase (ammonium sulfate suspension, P1506) were added to final concentration of 140 µM, 1.25 mM, 40 U/mL and 80 U/mL. 1 µL DMSO/compound solution was added and mixed. 3 µL 10 mM MgATP (pH = 7.0), was added to make the final total volume of 30 µL and final concentration of MgATP at 1 mM. Time course of fluorescence decrease due to NADH oxidation was measured using Synergy NEO Microplate Reader. The fluorescence values were plotted against time and fit by linear regression. The slopes of these lines were used to calculate the ATPase rate. The relative activities were obtained by dividing the measured activity by activity in DMSO control (no compound). The EC_{50} value for the inhibition of wild-type Mdn1 was determined by fitting the average values for relative activity at each inhibitor concentration (n=4 independent measurements) to a four-parameter sigmoidal dose-response curve using PRISM.

Chemical Synthesis

Materials and instrumentation.

Materials were purchased from Sigma-Aldrich and used without purification unless otherwise noted. Reactions were run in capped round bottom flasks stirred with Teflon®-coated magnetic stir bars. Moisture- and air-sensitive reactions were performed in flame-dried round bottom flasks, fitted

with rubber septa or glass gas adapters, under a positive pressure of nitrogen. Moisture- and air-sensitive liquids or solutions were transferred via nitrogen-flushed syringes. As necessary, solutions were deoxygenated by bubbling with nitrogen using a gas dispersion tube. Evaporation of solvents was accomplished by rotary evaporation using a Büchi rotary evaporator, equipped with a dry ice-acetone condenser, at 5-75 mm Hg at temperatures between 35°C and 50°C. Experiments were monitored by thin layer chromatography (TLC) or liquid chromatography-mass spectrometry (LC-MS). The maintenance of 30 °C to 150 °C reaction temperatures was accomplished by the use of an oil bath. Products obtained as solids or high boiling oils were dried under vacuum (~1 mm Hg).

Compounds were purified using silica gel column chromatography, as necessary. Analytical TLC was performed using Whatman 250 micron aluminum backed UV F254 precoated silica gel flexible plates. Subsequent to elution, ultraviolet illumination at 254 nm allowed for visualization of UV active materials. Staining with basic potassium permanganate solution allowed for further visualization.

Proton nuclear magnetic resonance spectra (^{1}H NMR) were recorded on Bruker DPX 400 MHz nuclear magnetic resonance spectrometer. Chemical shifts for ^{1}H NMR spectra are reported as δ in units of parts per million (ppm) relative to tetramethylsilane (δ 0.0) using the residual solvent signal as an internal standard or tetramethylsilane itself: dimethylsulfoxide-*d*6 (δ 2.50, quintet) and deuterium oxide-*d*2 (δ 4.80, singlet). Multiplicities are given as s

(singlet), d (doublet), t (triplet), or m (multiplet). Coupling constants are reported as a *J* value in Hertz (Hz). The number of protons (n) for a given resonance is indicated by nH.

Liquid chromatography-mass spectral analyses were obtained using a Waters MicroMassZQ mass spectrometer, with an electron spray ionization (ESI) probe, connected to a Waters 2795 HT Separation Module Alliance HT HPLC system running MassLynx (V4.0). The system used a Waters 996 Photodiode Array Detector set to 254 nm for peak detection, and a Symmetry® C18 (3.5 micron) 2.1 x 50 mm column for separation (mobile phase for positive mode: solvent A: water with 0.1% formic acid, solvent B: acetonitrile; mobile phase for negative mode: solvent A: water with 0.1% morpholine, solvent B: acetonitrile). Values are reported in units of mass to charge (*m/z*)

Synthesis of Rbin-1 analogs.

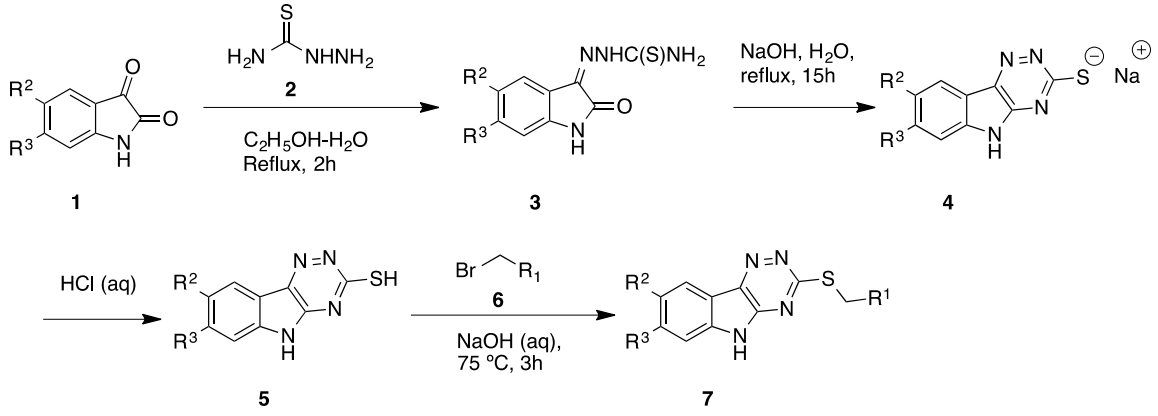


Figure 2-6. Synthetic scheme of Rbin-1 analogs.

The synthesis method has been adapted from previous studies (Ivachtchenko et al., 2002). Briefly, a solution of thiosemicarbazide **2** (3.38 mmol, 1 equivalent) in boiling water was added with stirring to a solution of isatin or substituted isatin **1** (3.38 mmol, 1 equivalent) in boiling ethanol. The resulting mixture was refluxed for 2h to form a precipitate of thiosemicarbazone **3**, which was isolated by filtration and washed with ethanol. Thiosemicarbazone **3**, was then transferred to a solution of sodium hydroxide (3.38 mmol, 1equivalent) in water and refluxed overnight. The reaction mixture was then cooled to room temperature and the pH of the solution was adjusted to 3 with dilute aqueous HCl (1:3, v/v) to get precipitates of **5**. Triazinoindole **5** was then recrystallized from ethanol or washed with hot ethanol to get pure **5** and then dried in air. Thiosubstitution of **5** was achieved by the addition of the alkyl bromide **6** (0.335 mmol, 0.67 equivalents) to a stirring solution of **5** (0.5 mmol, 1 equivalent) in 0.1 N NaOH (aqueous) and then refluxing the resulting mixture for 2h. The precipitated product was filtered, washed several times with water followed by hot ethanol to get pure **7**.

Characterizations of the Rbin analogs

All chemical structures and naming of analogs are listed in Figure 2-3B. Note: I only included characterizations of sixteen compounds synthesized by myself.

Analog 1 or Rbin-1 (ZC-3)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 1.84 (s, 3H), 3.98 (s, 2H), 4.91 (s, 1H), 5.1 (s, 1H), 7.42 (t, *J*= 7.38 Hz, 1H), 7.56 (d, *J*= 8.0

Hz, 1H), 7.68 (t, *J*= 7.54 Hz, 1H), 8.29 (d, *J*= 7.64 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 257.33, found: 257.57.

Analog 2 (ZC-9)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 3.96 (d, *J*= 6.68 Hz, 2H), 5.16 (d, *J*= 9.84 Hz, 1H), 5.39 (d, *J*= 16.8 Hz, 1H), 6.04 (m, 1H), 7.43 (t, *J*= 7.46 Hz, 1H), 7.52 (d, *J*= 8.0 Hz, 1H), 7.68 (t, *J*= 7.58 Hz, 1H), 8.30 (d, *J*= 8.0 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 243.30, found: 243.55.

Analog 3 (ZC11)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 1.03 (t, *J*= 7.28 Hz, 3H), 1.77 (m, 2H), 3.24 (t, *J*= 7.16 Hz, 2H), 7.43 (t, *J*= 7.44 Hz, 1H), 7.56 (d, *J*= 8.04 Hz, 1H), 7.68 (t, *J*= 7.6 Hz, 1H), 8.29 (d, *J*= 7.68 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 245.32, found: 245.57.

Analog 5 (ZC-51)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 0.37 (m, *J*= 4.81 Hz, 2H), 0.58 (m, *J*= 5.9 Hz, 2H), 1.24 (m, 1H), 3.23 (d, *J*= 7.12 Hz, 2H), 7.43 (t, *J*= 7.43 Hz, 1H), 7.56 (d, *J*= 8.08 Hz, 1H), 7.68 (t, *J*= 7.48 Hz, 1H), 8.30 (d, *J*= 7.7 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 257.33, found: 257.59.

Analog 6 or Rbin-3 (ZC-28)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.56 (s, 2H), 7.25 (t, *J*= 7.16 Hz, 1H), 7.32 (t, *J*= 7.22 Hz, 2H), 7.42 (t, *J*= 7.42 Hz, 1H), 7.47-7.53 (m, 2H), 7.58 (d, *J*= 8.04 Hz, 1H), 7.70 (t, *J*= 7.6 Hz, 1H), 8.30 (d, *J*= 7.64 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 293.36, found: 293.62.

Analog 7 (ZC-6)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.56 (s, 2H), 7.43 (t, *J*= 7.6 Hz, 1H), 7.52 (d, *J*= 4.96 Hz, 2H), 7.56 (d, *J*= 8.08 Hz, 1H), 7.68 (t, *J*= 7.62 Hz, 1H), 8.29 (d, *J*= 7.72 Hz, 1H), 8.50 (d, *J*= 3.72 Hz, 2H). ESI-MS (m/z) M+H]+ Calculated: 294.35, found: 294.59.

Analog 8 or Rbin-5 (ZC-10)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.74 (s, 2H), 7.43 (t, *J*= 7.54 Hz, 1H), 7.49 (t, *J*= 3.69 Hz, 2H), 7.58 (d, *J*= 8.08 Hz, 1H), 7.6 (m, 2H), 7.88 (d, *J*= 8.08 Hz, 3H), 8.04 (s, 1H), 8.30 (d, *J*= 7.6 Hz, 1H). ESI-MS (m/z) M+H]+ Calculated: 243.42, found: 243.65.

Analog 9 (ZC-5)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 3.73 (s, 3H), 4.53 (s, 2H), 6.83 (d, *J*= 7.2 Hz, 1H), 7.08 (d, *J*= 7.12 Hz, 2H), 7.24 (t, *J*= 7.9 Hz, 1H), 7.43 (t, *J*= 7.48 Hz, 1H), 7.58 (d, *J*= 8.04 Hz, 1H), 7.69 (t, *J*= 7.7 Hz, 1H), 8.31 (d, *J*= 7.6 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 323.39, found: 323.67.

Analog 10 (ZC-41)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.86 (s, 2H), 7.43 (t, *J*= 7.52 Hz, 1H), 7.56 (m, 2H), 7.7 (t, *J* = 7.72, 2H), 7.90 (d, *J*= 7.72 Hz, 1H), 8.05 (d, *J*= 8.12 Hz, 1H), 8.30 (d, *J*= 7.68 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 338.36, found: 338.61.

Analog 11 (ZC-7)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.69 (s, 2H), 7.44 (t, *J*= 7.6 Hz, 1H), 7.5 (d, *J*= 8.0 Hz, 1H), 7.69 (t, *J*= 7.6 Hz, 1H), 7.80 (d, *J*= 8.16 Hz, 2H), 8.17 (d, *J*= 8.16 Hz, 2H), 8.30 (d, *J*= 7.68 Hz, 1H). ESI-MS (m/z) M+H]+ Calculated: 338.36, found: 338.60.

Analog 12 or Rbin-4 (ZC-30)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.59 (s, 2H), 7.16 (t, *J*= 7.44 Hz, 1H), 7.22 (t, *J*= 9.28 Hz, 1H), 7.33 (m, 1H), 7.43 (t, *J*= 7.48 Hz, 1H), 7.58 (d, *J*= 8.08 Hz, 1H), 7.64 (t, *J*= 7.62 Hz, 1H), 7.69 (t, *J*= 7.64 Hz, 1H), 8.31 (d, *J*= 7.72 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 311.35, found: 311.57.

Analog 15 or Rbin-2 (ZC-32)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 1.84 (s, 3H), 3.98 (s, 2H), 4.91 (s, 1H), 5.12 (s, 1H), 7.58 (d, *J*= 8.28 Hz, 1H), 7.74 (s, 1H),

8.24 (d, *J*= 8.4 Hz, 1H). ESI-MS (m/z) [M]+ Calculated: 335.22, found: 335.45 and [M+2H]+ Calculated: 337.22, found: 337.48 (bromine pattern).

Analog 16 (ZC-33)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.56 (s, 2H), 7.26 (t, *J* = 7.24 Hz, 1H), 7.33 (t, *J* = 7.36 Hz, 2H), 7.51 (d, *J*= 7.36 Hz, 2H), 7.59 (d, *J*= 7.24 Hz, 1H), 7.75 (s, 1H), 8.24 (d, *J*= 8.32 Hz, 1H). ESI-MS (m/z) [M]+ Calculated: 371.26, found: 371.49 and [M+2H]+ Calculated: 373.26, found: 373.48 (bromine pattern).

Analog 17 (ZC-52)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 0.37 (d, *J* = 4.12 Hz, 2H), 0.58 (d, *J*= 7.2 Hz, 2H), 1.24 (m, 1H), 3.23 (d, *J*= 7.04 Hz, 2H), 7.58 (d, *J*= 8.12 Hz, 1H), 7.73 (s, 1H), 8.24 (d, *J*= 8.0 Hz, 1H). ESI-MS (m/z) [M]+ Calculated: 335.22, found: 335.52 and [M+2H]+ Calculated: 337.22, found: 337.55 (bromine pattern).

Analog 21 or Rbin-6 (AP-38)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 2.78 (m, 3H), 4.0 (s, 2H), 4.55 (t, *J* = 6.66, 2H), 4.91 (s, 1H), 5.14 (s, 1H), 7.48 (t, *J* = 7.46 Hz, 1H), 7.75 (t, *J* = 7.44 Hz, 1H), 7.90 (d, *J* = 8,24 Hz, 1H), 8.34 (d, *J* = 7.64 Hz, 1H). ESI-MS (m/z) [M+H]+ Calculated: 309.40, found: 309.65.

Analog 22 or Rbin-7 (ZC-21)**:** ^{1}H NMR (DMSO) δ 1.85 (s, 3H), 3.99 (s, 2H), 4.91 (s, 1H), 5.12 (s, 1H), 7.42 (d, *J*= 8.4 Hz, 1H), 7.96 (d, *J*= 8.4 Hz, 1H), 8.59 (s, 1H). ESI-MS (m/z) [M+H]+ Calculated: 383.22, found: 383.51.

Analog 23 (ZC-25)**:** ^{1}H NMR (400 MHz, DMSO-d_6) δ 4.56 (s, 2H), 7.25 (t, *J* = 7.04 Hz, 1H), 7.33 (t, *J* = 7.32 Hz, 2H), 7.42 (d, *J*= 8.4 Hz, 1H), 7.51 (d, *J* = 7.36 Hz, 2H), 7.96 (d, *J*= 8.56 Hz, 1H), 8.60 (s, 1H). ESI-MS (m/z) [M+H]+ Calculated: 419.26, found: 419.48.

CHAPTER 3

NEGATIVE-STAIN EM STUDIES OF MDN1

Note to readers: the results discussed below arose from a collaborative effort between myself and several colleagues in the Walz Lab at The Rockefeller University. In particular, Dr. Hiroshi Suzuki helped me to collect the negative-stain EM micrographs in this chapter.

3.1 Introduction

Members in the AAA protein family feature structurally conserved AAA domains that assemble to functional oligomers (Erzberger and Berger, 2006). ATP binding and hydrolysis at the ATPase sites, formed by the neighboring AAA domains, can drive the conformational changes of individual AAA domains and thus the oligomers. The conformational changes of the functional oligomers could exert mechanical force to their substrates. Therefore, understanding the nucleotide-specific conformations of AAA proteins are key to elucidate their cellular functions.

Mdn1 is required to remodel precursors of 60S subunit of ribosomes in the cell. Previous studies identified two ribosome assembly factors, Rsa4 and Ytm1, interact directly with the C-terminal MIDAS domain in Mdn1. The remodeling function of Mdn1 likely requires conformational coupling between the N-terminal AAA domains and the MIDAS domain. However, it has been difficult to build models for the coupling mechanism as these two domains are separated by more than 2000 aa without known functions.

Representative 2D averages of negative-stain EM images of Rea1 (Mdn1 ortholog in *S. cerevisiae*) revealed that Rea1 forms a ‘ring’ structure

with an elongated ‘tail’ and these two parts could adopt different relative orientations. The ring structure is likely formed by the AAA domains and the tail is comprised of the C-terminal domains. It was assumed that the MIDAS domain lies at the tip of the elongated tail. The heterogeneity of orientations of the tail relative to the ring was interpreted as a result from the lever-arm motions of the tail during Mdn1’s ATPase cycle. However, careful statistical analysis of all particles, rather than a few representative averages, is required to deduce the distributions of conformational states of Mdn1 particles.

In addition, previous negative-stain studies of Rea1 were performed in the absence of additional nucleotides and the different 2D class averages were interpreted to be various nucleotide-specific conformational states stalled during the purification (Ulbrich et al., 2009). However, the heterogeneity of the 2D projections of Rea1 can be caused by the different viewing angles of the particles. Therefore the assignment of specific conformations of Rea1 to certain nucleotide states remains an open question.

3.2 Results: Analysis of S. pombe Mdn1 in the presence of different nucleotides, analogs and Rbin-2

To characterize how the conformations of Mdn1 couple to ATP hydrolysis we used negative-stain EM to analyze its structure in the presence of different additives, including 1 mM ATP, 1 mM ATP with1 μM Rbin-2, 1 mM AMPPNP and 1 mM ADP (Figure 3-1). Representative 2D class averages of Mdn1 were generated using Iterative Stable Alignment and Clustering (ISAC)

(Yang et al., 2012). These 2D averages revealed that *S. pombe* Mdn1 is comprised of a 'ring' and an extended 'tail', reminiscent of what was observed in the negative-stain EM studies of *S. cerevisiae* Rea1 (Ulbrich et al., 2009). In different 2D averages, the orientations of the tail relative to the AAA ring appear to be variable in all conditions. Remarkably, a subset of the averages showed six discrete domains within the ring structure (the blue box in Figure 3-1B), consistent with the presence of six AAA domains in hexameric AAA proteins (Erzberger and Berger, 2006). In addition, there appear to be another 'ring' at the tip of the elongated tail in some 2D averages, resulting in a dumbbell-like overall shape. Another striking feature observed is that in the presence of ADP only, Mdn1's AAA ring can adopt an 'open ring' conformation with a gap observed in a subset of the 2D averages (Figure 3-1D). This gap locates right next to where the elongated tail extends from the ring, most likely between the N-terminal AAA1 domain and the AAA6 domain that connects to the C-terminal tail of Mdn1. This conformation would be consistent with a hypothesis that loss of ATP binding in the AAA6 ATPase site in Mdn1 causes the opening, reminiscent of the opening observed in the AAA1 ATPase site of cytoplasmic dynein when crystallized in the absence of nucleotide (Bhabha et al., 2014; Carter et al., 2011). The possible interpretations of this gap are further discussed below (3.3 Discussion).

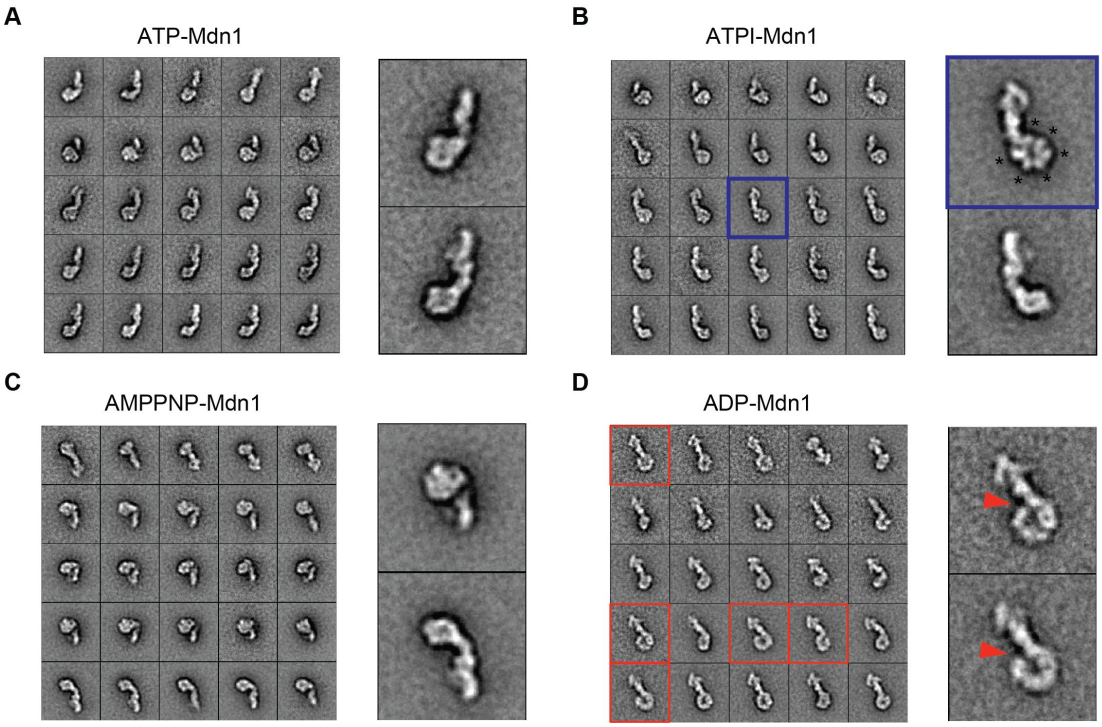


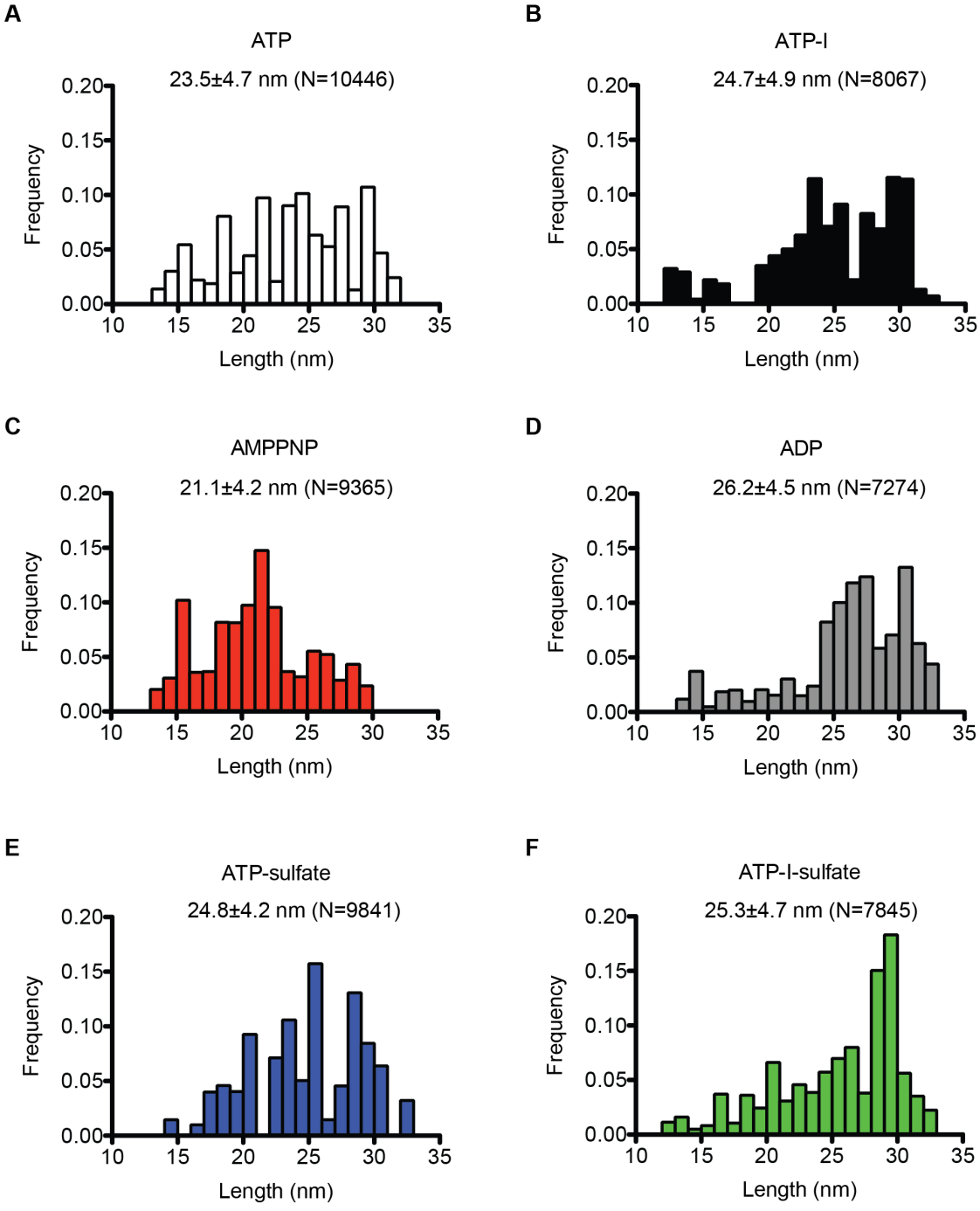
Figure 3-1. Representative 2D class averages of negative-stain EM images of Mdn1 in the presence of different additives. (A-D) Representative 2D class averages of Mdn1 in the presence of 1 mM ATP, hereafter named as ATP-Mdn1 (A), 1 mM ATP and 1 μM Rbin-2, hereafter named as ATPI-Mdn1 (B), 1 mM AMPPNP hereafter named as AMPPNP-Mdn1 (C) and 1 mM ADP, hereafter named as ADP-Mdn1 (D) resulted from Iterative Stable Alignment and Clustering (ISAC). The box size is 460x460 Å. Zoom-in views are also shown for two averages in each condition. In (B), a 2D average image of ATPI-Mdn1 with six discrete domains (marked by asterisks) in the ring structure is highlighted (blue square). In (D), averages of ADP-Mdn1 with partially opened AAA ring are highlighted (red squares) and the gap within the AAA ring is indicated by red triangles in the zoom-in views.

In order to analyze conformation distributions of all Mdn1 particles on grids, 2D classification was performed for all datasets using SPIDER (System for Processing Image Data from Electron Microscopy Related Fields) (Shaikh

et al., 2008). The lengths of the Mdn1 in the 2D averages were used to approximate the lengths of all particles in the corresponding classes. The distributions of lengths of Mdn1 were then plotted in histograms (Figures 3-2A to D). Overall, there are substantial variations of the lengths of the particles in all conditions (12-33 nm). In particular, Mdn1 has a broad distribution of lengths in the presence of ATP only (23.5±4.7 nm). The variation is consistent with the different viewing angles of different conformations when active ATP hydrolysis is allowed in AAA proteins. In comparison, the lengths of Mdn1 in the presence of AMPPNP and ADP appear to be shorter and longer, respectively, compared to ATP-Mdn1 (Figures 3-2C and D, 21.1±4.2 nm and 26.2±4.5 nm, respectively), suggesting the additions of AMPPNP and ADP may change the conformations of Mdn1 or the distribution of orientations of Mdn1 on the EM grids. However, even if AMPPNP and ADP can potentially trap the conformation of Mdn1 at specific states, the variations of the length distributions in these two conditions suggest that the orientation heterogeneity of Mdn1 on grids along can lead to large uncertainties of the lengths. In addition, the length distribution of Mdn1 in the presence of ATP and Rbin-2 (Figure 3-2B) does not appear to be significantly different compared to the sample without Rbin-2 (Figure 3-2A). I later found that Rbin-2 only inhibits the ATPase activity of Mdn1 in the presence of additional sulfates (2 mM). Therefore, I analyzed the distribution of lengths of Mdn1 in the presence of ATP and sulfate (2 mM), with or without Rbin-2 (Figures 3-2E and F). Variations of lengths in both conditions are similar to other conditions (similar standard deviations). However, there appears to be an enrichment of extended views of Mdn1 in the presence of Rbin-1 (Figure 3-2E). Together, these data suggest that the additions of nucleotides (ADP), nucleotide analogs

(AMPPNP) or the chemical inhibitor (ATP-Rbin-2-sulfate) to Mdn1 affect the 2D averages of Mdn1 on EM grids.

Figure 3-2. Length distributions of Mdn1 particles in the presence of different additives. (A-D) Length distributions of Mdn1 in the presence of 1 mM ATP (A), 1 mM ATP and 1 μM Rbin-1 (B), 1 mM AMPPNP (C) and 1 mM ADP (D). All particles are classified into 100 classes (k=100) and the length of the protein in each 2D average was used to approximate the lengths of all particles in that class. The mean and standard deviation of all particles in each dataset are also noted. The number of particles in each dataset was also shown. Note: all particles were manually picked from the micrographs without the bias of shapes except that substantially larger entities, most likely formed by more than two Mdn1 molecules, are excluded. (D-E) Length distributions of Mdn1 in the presence of 1 mM ATP and 2 mM sulfate (D), and 1 mM ATP, 2 mM sulfate and 1 μM Rbin-1 (E).


A
ATP
23.5±4.7 nm (N=10446)
B
ATP-I
24.7±4.9 nm (N=8067)
C
AMPPNP
21.1±4.2 nm (N=9365)
D
ADP
26.2±4.5 nm (N=7274)
E
ATP-sulfate
24.8±4.2 nm (N=9841)
F
ATP-I-sulfate
25.3±4.7 nm (N=7845)
Frequency
Length (nm)
0.00
0.05
0.10
0.15
0.20
10
15
20
25
30
35

3.3 Discussion

Together, the negative-stain EM analyses reveal the heterogeneity of *S. pombe* Mdn1's 2D projections on grids. This heterogeneity can result from the convolution of two relevant variations: the variations of Mdn1 conformations and the variations of Mdn1 orientations on the EM grids, which are determined by the properties of specific conformations. The differences in the length distributions of Mdn1 particles in the presence of various additives indicate that these additives change the conformations or other important biochemical aspects of Mdn1 that affects the orientation of Mdn1 on the EM grids. If not, the orientation heterogeneity and the appearance of Mdn1 on grids should remain similar. This is consistent with other AAA proteins that association with AMPPNP, ADP and ATPase inhibitors can potentially favor specific nucleotide states or other properties of Mdn1, e.g. changes the dynamics of residues nearby, which affect the orientations of Mdn1.

The limitation of the two-dimensional analysis is that the lengths of 2D averages cannot be directly correlated with the lengths of their corresponding conformations in three dimensions. In particular, the variations observed in AMPPNP-Mdn1 dataset suggests that there are large uncertainties of the orientations even though AMPPNP can inhibit the hydrolysis of ATP by Mdn1 (as shown in Figure 2-5G) and potentially trap specific conformations. It is possible that a particle with short 'length' in 2D projections corresponds to an extended conformation of Mdn1 that favors a specific orientation on the EM grid so that it looks compact in the projection view. Therefore, we cannot directly correlate the 2D projections and the conformations of Mdn1.

S. pombe Mdn1 has a ring-tail configuration similar to Rea1 (the *S. cerevisiae* ortholog). Although differences in the AAA ring cannot be discerned

in most cases, I observed that in a subset of 2D averages of Mdn1 in the presence of ADP only, there exists a gap within the AAA ring, most likely between AAA1 and AAA6 domains, reminiscent of the opening observed in other AAA proteins in the absence of nucleotide (Carter et al., 2011). These results suggest that the loss of ATP in this site can cause the opening for a subset of Mdn1 molecules, while the other Mdn1 molecules still have a closed AAA ring, most likely due to other unknown redundancy of the interactions between AAA1 and AAA6 domains. Interestingly, previous cell-based assay indicates that a mutant Mdn1 construct, which has its lysine residue in Walker A motif of AAA6 mutated to alanine, can partially rescue the loss of function of Mdn1 in cell, consistent with the hypothesis that ATP binding in AAA6 only partially contributes to the integrity of the closed-ring conformation of AAA domains. However, it is worthy to note that the conformation of wild-type Mdn1 in the presence of ADP alone may only be relevant *in vitro* as high concentration of ATP (2-5 mM) likely exists in the cellular environment.

3.4 Methods

Biochemical preparation of full-length *S. pombe* Mdn1.

The full-length *S. pombe* Mdn1 was prepared as described in Method 2.5 and used for grids apart from ADP-Mdn1. For ADP-Mdn1, the buffer for the size-exclusion chromatography contains 1 mM ADP instead of 50 μM ATP.

Negative-Stain EM Sample Preparation and Processing.

The grids for negative-stain were prepared by following a published protocol (Ohi et al., 2004). Freshly purified protein was mixed with nucleotides,

nucleotide analogs or inhibitor right before the grid preparation. The samples were diluted to ~0.01 mg/mL and stained with 0.7% uranyl formate (PFALTZ&BAUER, Inc. U01000). The images were then collected at room temperature with a Philips CM10 electron microscope equipped with a tungsten filament operating at 100 kV. All images were recorded on an AMT XR16L-ActiveVu charge-coupled device camera (Woburn, MA, USA) using a defocus of approximately -1.5 µm and a nominal magnification of 52000x.

Manual particle picking was performed using Boxer in EMAN and windowed into 216 x 216-pixel images (Tang et al., 2007). The particle images were then reduced to 64 x 64-pixel boxes. 2D classification was then performed using iterative stable alignment and clustering (ISAC) algorithm (Yang et al., 2012). Five iterations were performed using a 'threshold' of 0.07 and an 'initial number of particles per group' as 100.

System for processing image data from electron microscopy related fields (SPIDER) was also used to analyze the same datasets. In particular, all particles in each dataset are classified to 100 classes (k=100) using reference-free alignment for 10 iterations. The 2D class averages of the last iteration were then analyzed in Photoshop. The maximum 1D dimensions of the averages were assigned to the lengths of all particles in the class.

CHAPTER 4

Cryo-EM studies of *S. pombe* Mdn1 in the presence of AMPPNP

Note to readers: the results discussed below arose from a collaborative effort between myself and several colleagues in the Kapoor Lab and the Walz lab at The Rockefeller University. A closely related version is in press in Cell (DOI:10.1016/j.cell.2018.09.015). In particular, Dr. Hiroshi Suzuki provided help in building the models in this chapter.

4.1 Introduction

While the negative-stain studies of *S. pombe* Mdn1 constitute a step toward the understanding of its domain architecture, the limited resolution does not allow us to assign individual domains for this large protein, such as the MIDAS domain, let alone understand the coupling between domains. In addition, the heterogeneity of 2D averages of negative-stain images can result from the combination of both conformational heterogeneity and orientational heterogeneity on grids. Therefore, high-resolution three-dimensional structural analysis of Mdn1 is required to understand Mdn1's function during ribosome biogenesis.

4.2 Cryo-EM structure of Mdn1 in the presence of AMPPNP

Full-length S. pombe Mdn1 has the N domain, six AAA, linker, D/E-rich and MIDAS domains in a single polypeptide (aa 1-4717, Figure 4-1A). To reduce the potential conformational heterogeneity of Mdn1, the full-length *S. pombe* Mdn1 was incubated with AMPPNP (1 mM), a nucleotide analog that inhibits the ATPase activity of recombinant Mdn1 (Figure 2-5G) and analyzed using single-particle cryo-EM (Appendix Figure 4-1). Initial class averages

indicated that Mdn1 adopts an elongated structure with a hexagonal 'ring' (Figure 4-1B, filled arrowhead), an arrangement of AAA domains typical for the proteins in the superfamily (Erzberger and Berger, 2006; Hanson and Whiteheart, 2005), connected to an elongated, hook-shaped 'tail' (Figure 4-1B, empty arrowhead), similar to the negative-stain averages of Mdn1 (Figure 3-1).

Three-dimensional (3D) classification of Mdn1 particles indicated that the ring and tail densities varied slightly in their relative orientations, likely due to flexibility at a single site in the tail, which we named 'hinge point' (Figure 4-1C, red arrowheads, see Appendix Figure 4-1B for details). Therefore, we masked the densities for the ring and tail regions, above and below the hinge point, in Mdn1 and processed them separately. The final 3D maps were refined to an estimated resolution of 4.0 Å for the ring region, and 5.9 Å for the tail region (Figure 4-1C, Appendix Figures 4-1C and 4-2A). At 4.0 Å resolution, α-helices and individual β-strands can be resolved in Mdn1's ring region (Figure 4-1C and Appendix Figure 4-2C). We were also able to trace the backbone for the N-terminal portion, which included the N and AAA domains and the tail up to the hinge point (Figure 4-1C). A pseudo-atomic model for the AAA domains was built based on known AAA protein folds (Erzberger and Berger, 2006). The remaining tail density was modeled with polyalanine helices. Hereafter, we refer to this model as AMPPNP-Mdn1.

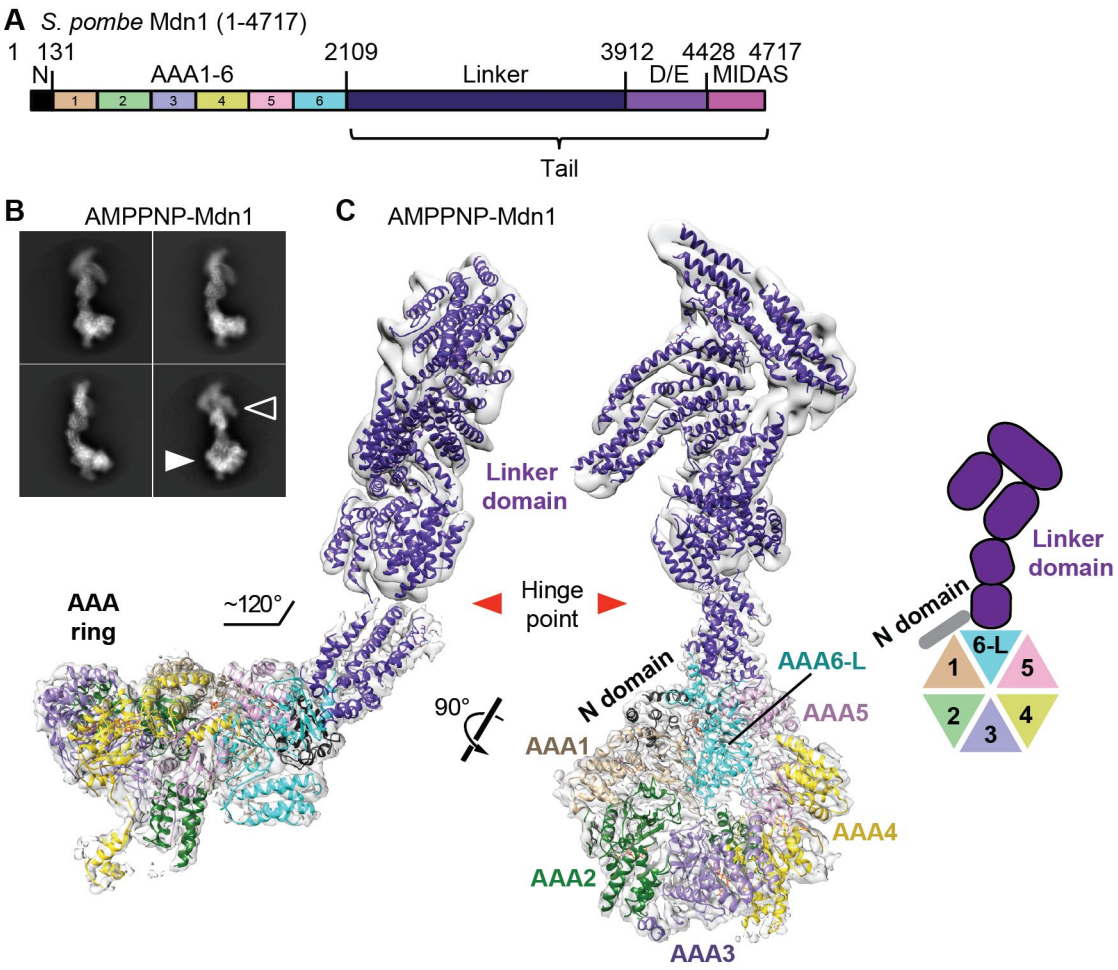
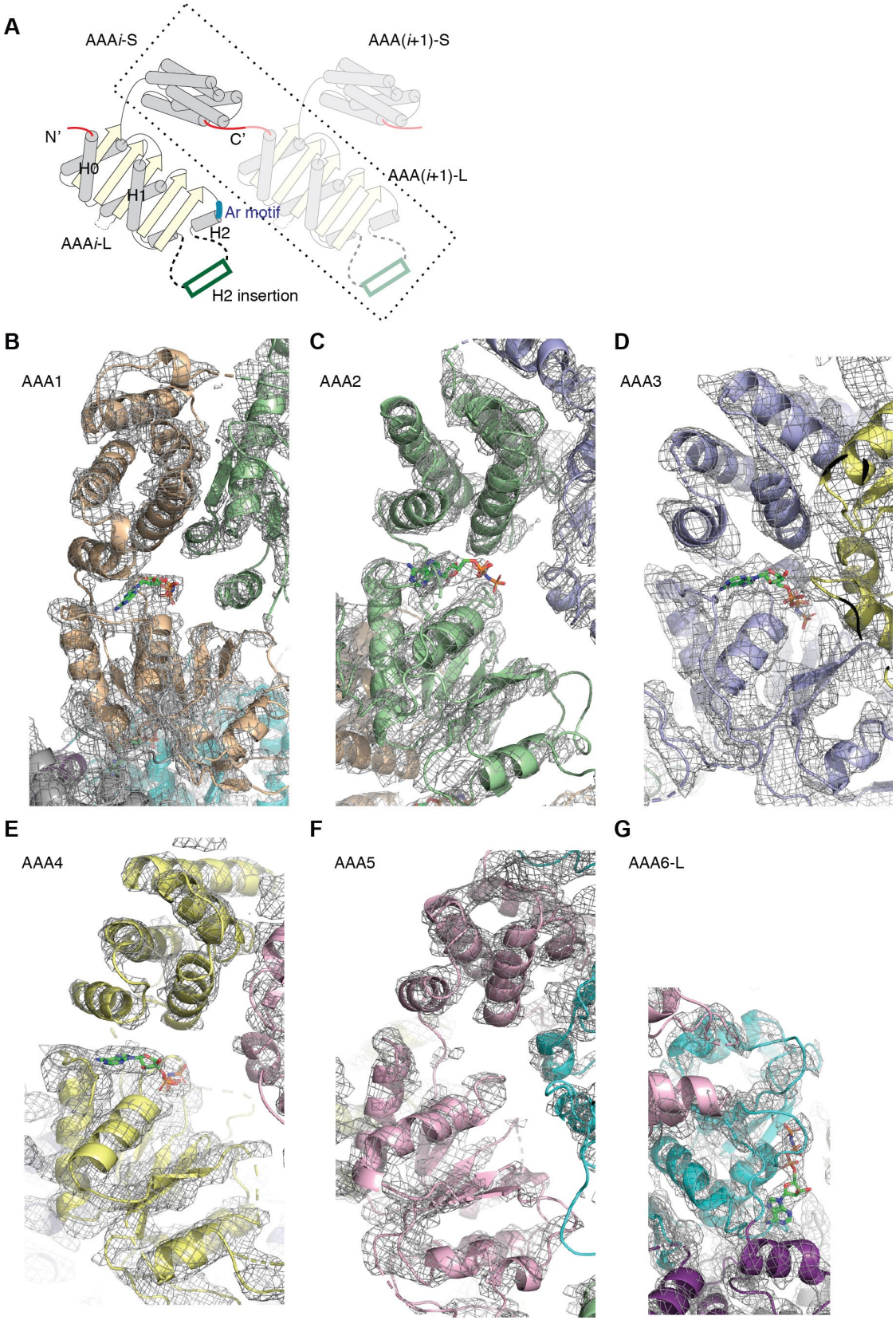


Figure 4-1. Overall architecture of full-length Mdn1 in the presence of AMPPNP. (A) Schematic for the domain organization of *S. pombe* Mdn1 (aa 1-4717). Residue numbers at selected domain boundaries are indicated. (B) Selected class averages of vitrified Mdn1 in the presence of AMPPNP (1 mM). The AAA ring (filled arrowhead) and the hook-shaped tail (open arrowhead) are indicated. (C) Two views of the Mdn1 density map (grey surface) with the model that was built into the map (ribbon representation, colored as in (A)) (hereafter, AMPPNP-Mdn1). The model and a schematic for the N, AAA and linker domains are shown. The hinge point (red arrowhead) in the tail region, where the density was split and processed separately, is also indicated in the model.

The AAA ring of AMPPNP-Mdn1 has an outer diameter of ~140 Å, an inner diameter of ~30 Å and a height of ~65 Å. The elongated tail region is composed of five bundles of α-helices. The tail first extends from the AAA ring for ~20 nm but then turns back to form a hook-like structure. The angle between the base of the tail and the plane of the AAA ring is ~120° (Figure 4-1C). The model for the hook-shaped tail accounts for ~1400 amino acids, which includes ~50% of the C-terminal residues after the AAA domains (~2600 aa, Figure 4-1A), suggesting that a portion of the C-terminal tail domain may be disordered.

We next examined the organization of the six AAA domains in Mdn1. In general, each AAA domain is comprised of a large N-terminal subdomain (AAA-L) and a small C-terminal subdomain (AAA-S) (Figure 4-2A). Six AAA-L and five AAA-S subdomains could be readily identified in the cryo-EM map as they adopt characteristic folds of AAA domains (Figures 4-2B to G) (Erzberger and Berger, 2006).

Figure 4-2. Density map of individual AAA domains in AMPPNP-Mdn1 with models. (A) Schematic of secondary-structure elements of two sequential AAA domains and the arrangement of their subdomains. α-Helices (grey cylinders) and β-strands (yellow arrows) are shown. The N and C termini of the first AAA domain (solid red linkers) and those of the second AAA domain (faded red linkers) are shown. The dashed box indicates rigid body i, which is formed by subdomains AAA*i*-S and AAA(*i+1*)-L. Helices 0 (H0), 1 (H1) and 2 (H2) are the first three helices in the large subdomain. The H2 insertions, which are present in all six AAA-L domains of Mdn1 but have different lengths (15-120 aa), are represented as green rectangles. The aromatic motif at the beginning of H2 helix is also labeled (blue bar). These motifs are also labeled in the alignment of Mdn1 orthologs' sequences in Figure 1-4. (B-G) The density map of each AAA domain (grey mesh) is shown with the built model (ribbon representation). The nucleotides (stick representation) bound between the large and small subdomains of AAA1-4 and next to the AAA6-L subdomain are shown. All nucleotide densities are modeled as AMPPNP molecules. Note that AAA5 shows no density for a nucleotide.


A
AAA*i*-S
AAA(*i*+1)-S
N'
C'
H0
H1
AAA(*i*+1)-L
Ar motif
H2
AAA*i*-L
H2 insertion
B
AAA1
C
AAA2
D
AAA3
E
AAA4
F
AAA5
G
AAA6-L

In the orientation shown in the left panel of Figure 4-3A, the AAA-L domains are at the lower face of the ring, while the AAA-S domains lie above. The polypeptide chain following the C terminus of AAA6-L, which was expected to form the AAA6-S domain, lacks structural similarity to other AAA-S domains within Mdn1. We find that this region of the protein extends out of the ring and forms the base of the linker domain (Figure 4-3A). The N, AAA1-L and AAA6-L domains form a structural motif, which we name the 'ring-linker junction', that establishes the orientation of the hook-shaped linker domain relative to the AAA ring (Figure 4-3A). The formation of this ring-linker junction covers multiple interfaces between the N, AAA1-L, AAA6-L and linker domains and also the nucleotide bound in the AAA6 ATPase site.

Sequence alignments reveal that Mdn1 belongs to a subfamily of AAA proteins that possess insertions following helix H2 in the AAA-L subdomain (called H2 insertion, green box in Figure 4-2A and sequence alignment in Figure 1-4) (Erzberger and Berger, 2006). For three AAA-L subdomains (AAA1-L, AAA3-L, and AAA5-L), these insertions are <20 amino acids, while the other AAA-L subdomains (AAA2-L, AAA4-L and AAA6-L) have longer insertions (60-120 aa, Figure 4-3A and sequence alignment in Figure 1-4). We find that all the long H2 insertions project out from the same face of the AAA ring, opposite to that from which the tail extends (Figure 4-3A). The AAA2 H2 insertion (120 aa) forms a four-helix bundle that is proximal to the central hole of the AAA ring (Figure 4-3A). The AAA4 H2 insertion (60 aa) contains a bundle of three helices. Finally, the AAA6 H2 insertion (~100 aa) is also

largely helical and adjoins the AAA6-L and AAA1-L domains (Figure 4-3A), possibly functioning as a bridging element.

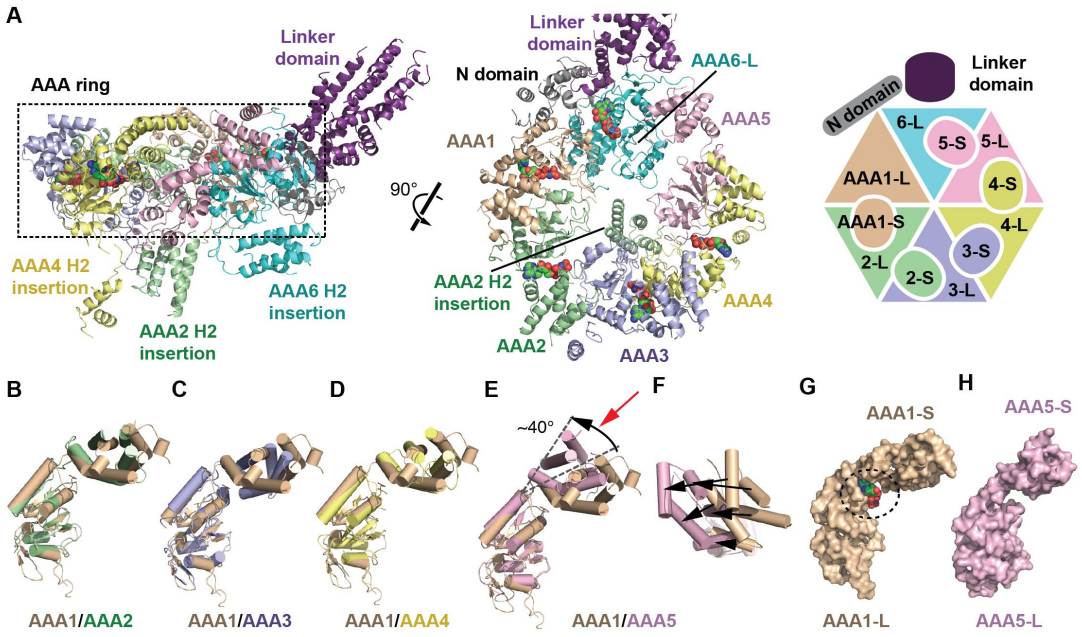


Figure 4-3. Asymmetric arrangement of six AAA domains in Mdn1. (A) Two views of the model for the six AAA domains and part of the linker domain, along with a schematic, are shown. The AAA ring core (dashed rectangle) is highlighted. (B-E) Alignment of the AAA2-5 domains with the AAA1 domain using their large subdomains. In (E), an ~40º difference between the orientations of the small subdomains in AAA5 and AAA1 is apparent. (F) AAA1 and AAA5 small subdomains viewed from the direction indicated by the red arrow in (E). Differences in the positions of the five equivalent α-helices (black arrows) are highlighted. (G and H) Surface representations of the AAA1 (G) and AAA5 (H) domains. In (G), the nucleotide (sphere representation) is also highlighted.

We note that the AAA1 and AAA6 in *S. pombe* Mdn1 lack the “Asp-Glu” Walker B motif that is required for ATP hydrolysis in AAA proteins (sequence alignment in Figure 1-4) (Erzberger and Berger, 2006), suggesting that

ATPase activity in Mdn1 is restricted to its AAA2-5 ATPase sites. Consistent with this hypothesis, cell-based studies have shown that key residues in the Walker A, Walker B and arginine finger motifs in the AAA2-5 ATPase sites are all required for Mdn1 function (Kawashima et al., 2016). We next compared the relative orientations of the five AAA-L and associated AAA-S subdomains (Figures 4-3B to E). AAA6 is not included as it lacks a typical AAA-S subdomain. The small and large subdomains in AAA2 to AAA4 show a similar arrangement as seen in AAA1 (Figures 4-3B to D). In AAA5, however, the small subdomain is rotated by ~40° relative to its large subdomain, causing substantial displacements of the five α-helices in AAA5-S (Figures 4-3E and F). This difference in relative orientation results in an opening between the small and large subdomains in AAA5 (Figures 4-3G and H).

Interestingly, we can assign densities corresponding to nucleotides in the AAA1-AAA4 and AAA6 ATPase sites but not in the AAA5 site (Figures 4-2B to G). As the cell-based studies indicate that ATP hydrolysis at the AAA5 ATPase site is needed for Mdn1 function, we expected that AAA5 contributes to Mdn1's overall ATPase activity. To test this we expressed and purified a construct with a Walker B mutation in Mdn1's AAA5 domain (E1637Q). Equivalent mutations have been used to disrupt ATP hydrolysis but not ATP binding in AAA proteins (Bhabha et al., 2014; Martin et al., 2005). The AAA5 mutant protein eluted from a size-exclusion column at 10.6 mL, similar to wild-type Mdn1 (elution volume: 10.6 mL) (Figures 4-4A and B). Negative-stain EM analysis revealed that the AAA5 mutant adopts an extended conformation,

with the ring and tail regions appearing similar to those observed for wild-type Mdn1 (Figure 4-4C). We found that the ATPase activity of the mutant construct is 0.15±0.02 ATP per second (Figure 4-4D), which is ~15% of what we measured under similar conditions for wild-type Mdn1 (Kawashima et al., 2016). Addition of Rbin-1 (1 μM) inhibited ~40% of the ATPase activity of the mutant, similar to what we observed for wild-type Mdn1 at this inhibitor concentration (Kawashima et al., 2016). These data are consistent with Rbin-1 targeting a site in Mdn1 other than the AAA5 ATPase site and with the observed ATPase activity being mainly due to the mutant Mdn1 construct. Together, our findings suggest that the AAA5 site contributes to Mdn1's overall ATPase activity and that the ATPase sites are structurally asymmetric and likely function asynchronously.

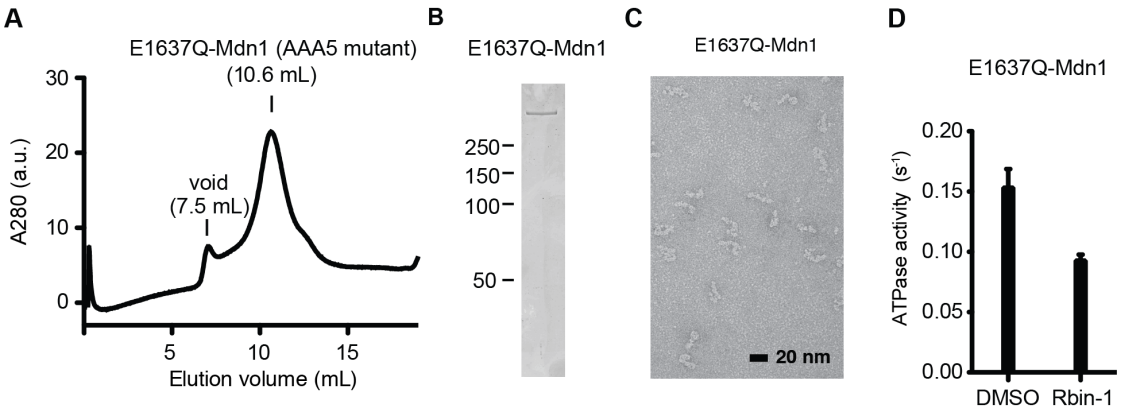


Figure 4-4 Characterization of AAA5 Walker B mutant Mdn1 (E1637Q-Mdn1). (A) Size-exclusion chromatography profile for E1637Q-Mdn1 (AAA5 mutant) (elution volume: 10.6 mL, void 7.5 mL). (B) SDS-PAGE analysis of purified E1637Q-Mdn1 (Coomassie blue stain). (C) Negative-stain EM image of Mdn1 carrying the E1637Q mutation in the Walker B motif of AAA5. Scale bar: 20 nm. (D) ATPase activities of Mdn1 in the presence of solvent control (DMSO) and Rbin-1 (1 μM, dissolved in DMSO) were measured using an NADH-coupled ATPase assay ([MgATP] = 1 mM).

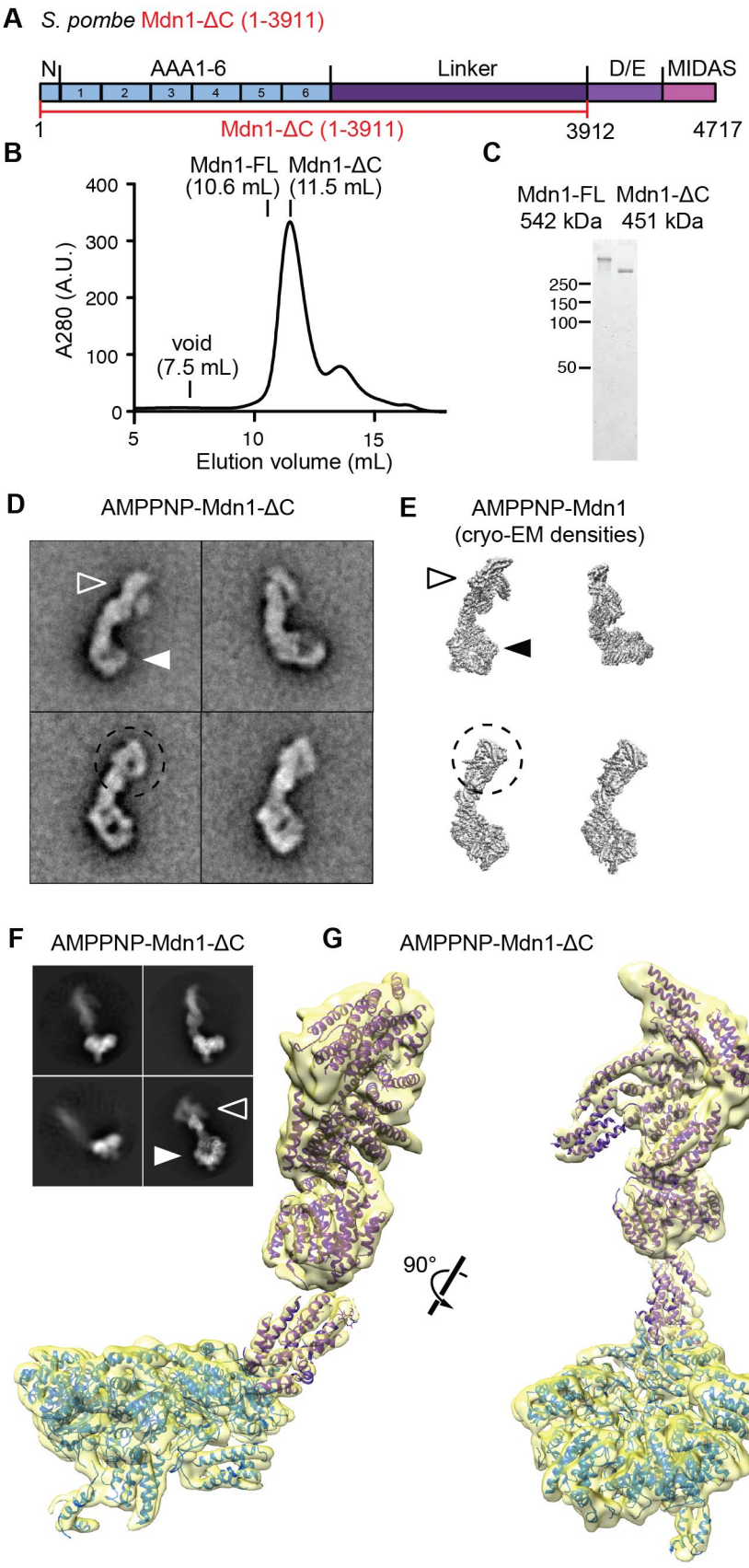
4.3 Cryo-EM structure of Mdn1-*Δ*C in the presence of AMPPNP

Although sequences of individual domains of N-terminal portion (aa 1-2109) of Mdn1 can be assigned to specific densities in the ring region of AMPPNP-Mdn1 map, the identities of sequences represented by the tail region of the same map are still unclear. In particular, only ~1400 aa is required to explain the densities of helical bundles in the hook-shaped tail (purple model in Figure 4-1C), indicating more than 1000 aa of the C-terminal sequence (aa 2110-4717) is not observed.

Secondary-structure prediction algorithms suggest that the ~500 aa-long D/E-rich domain, which follows the linker domain, likely contains only two

α-helices (formed by ~30 amino acids) and no β-strands (see secondary structure predictions in Figure 1-4). To test if the D/E-rich domain is indeed not observed in the map, I generated and purified an Mdn1 construct that lacked the D/E-rich and MIDAS domains (Mdn1-*Δ*C, aa 1-3911) (Figures 4-5A to C). I first used negative-stain EM to analyze the conformation of this truncated Mdn1 in the presence of AMPPNP. Two-dimensional class averages revealed a ring with a hook-shaped tail (Figure 4-5D), similar to the 2D class averages of the full-length Mdn1. Furthermore, the three-dimensional density map of AMPPNP-Mdn1 can be oriented to recapitulate these 2D projections of Mdn1-*Δ*C (Figure 4-5E). I then used single-particle cryo-EM to determine the structure of Mdn1-*Δ*C in the presence of AMPPNP. Maps at a resolution of 6.5 Å for the region including the AAA ring and 7.4 Å for the rest of the protein were obtained (Figures 4-5F and G, and Appendix Figure 4-3). The AMPPNP-Mdn1-*Δ*C map reveals a hexagonal ring and hook-shaped tail similar to the AMPPNP-Mdn1 map (Figure 4-5G). In addition, the model for full-length Mdn1 could be placed into the density map of Mdn1-*Δ*C as a single rigid body without further adjustment (Figure 4-5G). Together, these data indicate that the hook-shaped tail is formed by the linker domain alone (Figure 4-5G) and also confirm that the D/E-rich and MIDAS domains are not observed for the full-length Mdn1 map (Figure 4-1C). Further, these data also suggest that removal of the C-terminal D/E-rich and MIDAS domains does not affect the overall structure of the AAA and linker domains in the presence of AMPPNP.

Figure 4-5. Purification and Cryo-EM analysis of Mdn1-ΔC in the presence of AMPPNP. (A) Schematic of the Mdn1-ΔC construct that lacks the C-terminal D/E-rich and MIDAS domains (amino acids 1-3911). (B) Size-exclusion chromatography profile for Mdn1-ΔC (elution volume: 11.5 mL). The elution volumes of the void peak (7.5 mL) and full-length Mdn1 (Mdn1-FL, 10.6 mL) are also indicated for comparison. (C) Peak fractions from the size-exclusion chromatography of Mdn1-ΔC and Mdn1-FL were analyzed by SDS-PAGE (Coomassie stain). (D) Representative 2D class averages of Mdn1-ΔC in the presence of AMPPNP. The AAA ring (filled arrowhead) and the hook-shaped tail (open arrowhead) are indicated. (E) The cryo-EM map for AMPPNP-Mdn1 is oriented at similar viewing angles as seen in the 2D averages in (D). The AAA ring (filled arrowhead) and the hook-shaped tail (open arrowhead) are indicated. Note that the hook-shaped tail appears to be a ring (dashed oval) from certain viewing angles in both (D) and (E). (F) Selected class averages of vitrified Mdn1-ΔC in the presence of AMPPNP (1 mM). Note that the hook-shaped tail is discernible in some class averages (open arrowhead). (G) Two views of the EM density map of Mdn1-ΔC (yellow surface). The Mdn1 model for full-length Mdn1 (the same model shown in Figure 4-1C, blue for the ring region and purple for the tail region) could be placed into the Mdn1-ΔC map without further refinement.


A
S. pombe Mdn1-ΔC (1-3911)
N
AAA1-6
Linker
D/E
MIDAS
1
2
3
4
5
6
1
Mdn1-ΔC (1-3911)
3912
4717
B
Mdn1-FL (10.6 mL)
Mdn1-ΔC (11.5 mL)
A280 (A.U.)
400
300
200
100
0
void (7.5 mL)
5
10
15
Elution volume (mL)
C
Mdn1-FL 542 kDa
Mdn1-ΔC 451 kDa
250
150
100
50
D
AMPPNP-Mdn1-ΔC
E
AMPPNP-Mdn1 (cryo-EM densities)
F
AMPPNP-Mdn1-ΔC
G
AMPPNP-Mdn1-ΔC
90°

4.4 Discussion

In summary, cryo-EM structures of full-length and truncated S. pombe Mdn1 in the presence of AMPPNP provided high-resolution structural insights for Mdn1 in three dimensions. In particular, the pseudo-atomic model for the N-terminal portion of Mdn1 revealed an asymmetric AAA ring with H2 insertions extending from one side of the AAA ring, opposite to the side to which the tail of helical bundles extends. The models also revealed the domain architecture of the C-terminal portion of Mdn1.

S. pombe Mdn1 has Walker A motifs in all of its six AAA ATPase sites. In the presence of 1 mM AMPPNP, densities corresponding to nucleotides are observed in five out of six interfaces between neighboring AAA domains, leaving AAA5 site empty. Still, the hydrolysis of AAA5 appears to be important for the overall ATPase cycle of Mdn1. A single-point Walker B mutation (E to Q) in AAA5 reduces the ATPase activity of Mdn1 to ~15% of the wild-type Mdn1. As Rbin-1 inhibits ~40% of the wild-type and AAA5 mutant Mdn1's ATPase activities, it is likely that Rbin-1 blocks the activity at sites other than AAA5. Therefore, we suggest that AAA5 alone does not contribute to ~85% of Mdn1's overall ATPase activity. Studies of AAA proteins have revealed functional coupling between different AAA ATPase sites (Bhabha et al., 2014; DeWitt et al., 2015). We propose that similar to dynein's AAA3 domain, which can allosterically regulate the ATPase activity at AAA1, the nucleotide states of Mdn1's AAA5 domain modulate the ATPase activity at other AAA sites (Bhabha et al., 2014; DeWitt et al., 2015).

One unexpected structural feature of Mdn1 is the lack of AAA6-S subdomain in Mdn1, in contrast to other AAA proteins. However, the AAA ring appears to be closed at the interface between AAA1 and AAA6 domains in the ring-linker junction. In particular, the N-domain and the H2 insertion of AAA6 interact with both AAA1-L and AAA6-L subdomains, functioning as the bridges to close the AAA ring. The presence of nucleotide within the AAA6 ATPase site may also favor the closed interface, as nucleotide binding in the AAA1 ATPase site in dynein leads to closed interface between the AAA1 and AAA2 domains. These analyses are also consistent with the observation that the loss of the ATP binding in the AAA6 site only lead to open ring conformation in a subset of the molecules in the negative-stain analysis of ADP-Mdn1 sample due to the redundancy of interaction at this interface.

The linker domain in *S. pombe* Mdn1 (aa 2110-3911), which contains 1802 residues without good homology to well-studied proteins, forms multiple helical bundles. Our alanine-helices model of the hook-shaped tail contains 1334 residues, suggesting there are 468 amino acids of the linker domain missing in our model. Secondary structure prediction of the linker domain in *S. pombe* Mdn1 suggests there are 500 out of 1802 residues (aa 2110-3911) without well-defined secondary structures between multiple α-helices (Figure 1-4), consistent with a hypothesis that a substantial portion of the missing residues are involved in the formation of the loops that are not modeled between the ~50 α-helices. These analyses also suggest the linker domain is unlikely to form other large and well-defined structure in addition to the hook-

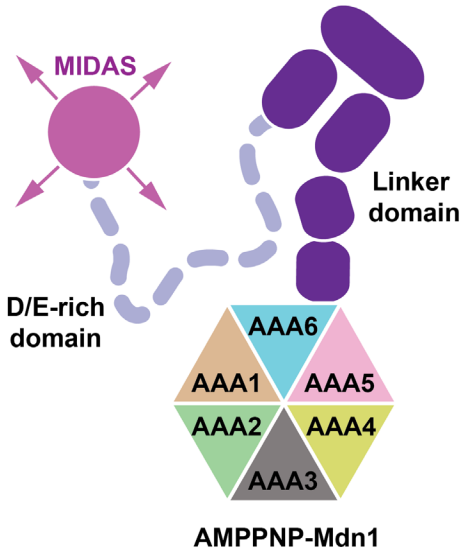
shaped tail and this information is helpful to assign unknown densities in other states (see Chapter 5). In addition, sequence comparisons of Mdn1 orthologs reveal that the conservation of the linker domain is only moderate (Table 4-1). In particular, conserved hydrophobic residues are arranged three or four residues apart (i, i+3 or i+4), consistent with the helical bundle structure (see sequence alignment in Figure 1-4). These analyses suggest that a substantial part of the linker domain may function as a structural unit by forming helical bundles. However, this cannot exclude that certain regions of the linker domain may still be important for specific functions of Mdn1.

Table 4-1. Comparisons of individual domains in *S. pombe* Mdn1 and *S. cerevisiae* Mdn1 (Rea1).

	S. pombe Mdn1	*S. cerevisiae* Mdn1	Sequence Identities
AAA	1979 aa	2030 aa	49%
Linker	1802 aa	1745 aa	18%
D/E-rich	517 aa	566 aa	24%
MIDAS	289 aa	284 aa	55%

Our results also suggest that the tether between the MIDAS and linker domains, i.e. the D/E-rich domain, adopt multiple conformations in the full-length Mdn1 in the presence of AMPPNP. The flexibility of D/E-rich domain is consistent with the secondary structure prediction based on the amino acid composition. In particular, the abundant negatively charged residues (Asp/Glu for 182 out of the 516 aa, >35%) would disfavor compact and well-defined folds, resulting in a flexible polypeptide. Interestingly, although the amino acids

composition of D/E-rich domains of Mdn1 orthologs are similar (>30% are Asp/Glu), the sequence identity of this domain is only moderate (Table 4-1), suggesting that evolution constrained the bias of the amino acid composition but not the exact positions of the negatively-charged residues. This is consistent with an idea that the D/E-rich domain has evolved to be a flexible polypeptide.



The flexibility of ~500 aa D/E-rich domain would disfavor the previously proposed lever-arm mechanism that is reminiscent of the powerstrokes of motor proteins like myosin or dynein (Bhabha et al., 2014; Spudich, 2001) since the flexible domain may not transmit force efficiently from the N-terminal AAA ring to the C-terminal MIDAS domain. In addition, the alternative hypothesis, i.e. the conformation communication between the AAA ring and the MIDAS domain through the entire tail in Mdn1, similar to the conformation coupling between the AAA domains and the microtubule-binding domain

through the structured coiled-coil stalk in dynein, is also unlikely due to the flexibility of the D/E-rich domain.

4.5 Methods

Biochemical preparation of full-length *S. pombe* Mdn1 and Mdn1-*Δ*C.

The full-length *S. pombe* Mdn1 and Mdn1-*Δ*C were prepared as described in Method 2.5.

Negative-Stain EM Sample Preparation and Processing.

The grids of Mdn1-*Δ*C were prepared and the micrographs were collected as described in Method 3.4. The data was processed using ISAC as in Method 3.4.

NADH-coupled steady-state ATPase assay.

For each steady-state ATPase reaction, the final volume was 25 µL. The final concentration of E1637Q-Mdn1 (AAA5 Walker B mutant) was ~40 nM. Sodium sulfate, NADH (Sigma N7410), phosphoenol pyruvic acid monopotassium salt (Sigma P7127), D-lactic dehydrogenase (Sigma L3888) and pyruvate kinase (ammonium sulfate suspension, Sigma P1506) were added to final concentrations of 2 mM, 200 µM, 1.0 mM, 30 U/mL and 30 U/mL, respectively. The presence of 2 mM sodium sulfate enhances inhibition of Mdn1's ATPase activity by Rbins. 1 µl DMSO or compound dissolved in DMSO was added and mixed. 5 µl 5 mM MgATP (pH = 7.0), was added to make the final total volume

to be 25 µl and the final concentration of MgATP to be 1 mM. Time course of fluorescence (440 nm) decrease due to NADH oxidation was measured using a Synergy NEO Microplate Reader. The fluorescence values were plotted against time and fit by linear regression. The slopes of these lines were used to calculate the ATPase rate.

Cryo-EM sample preparation and data collection.

To prepare cryo-EM grids of full-length Mdn1 with AMPPNP and Mdn1-ΔC with AMPPNP, the purified protein was concentrated to ~0.15 mg/mL using centrifugal filters (MW cutoff: 50 kDa; Millipore) and mixed with 10x AMPPNP stock solution (final concentration: 1 mM AMPPNP). To minimize the potential hydrolysis of AMPPNP, AMPPNP was added to Mdn1 immediately before the grids were frozen (<10 min). Samples were applied to Quantifoil R1.2/1.3 400 mesh Au holey carbon grids (Quantifoil), blotted and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). Cryo-EM data were collected on a 300-kV Titan Krios electron microscope (Thermo Fisher Scientific) equipped with a K2 Summit detector at a nominal magnification of 22,500x in super-resolution counting mode. After binning over 2 x 2 pixels, the calibrated pixel size was 1.3 Å on the specimen level. Exposures of 15 s were dose-fractionated into 50 frames with a dose rate of 10 electrons pixel^{-1} s^{-1}, resulting in a total dose of 88.8 electrons Å^{-2}. Defocus values ranging from -1.4 to -3.0 µm were used. Cryo-EM data collection statistics can be found in Table 4-1.

Cryo-EM data processing.

The data processing is summarized in Table 4-2. All movie frames were corrected with a gain reference collected in the same EM session, and specimen movement was corrected using MotionCor2 with dose weighting (Zheng et al., 2017). Images showing substantial ice contamination, abnormal background, thick ice or low contrast were discarded. The contrast transfer function (CTF) was estimated using CTFFIND4 (Rohou and Grigorieff, 2015). Approximately 10,000 particles were manually picked, subjected to 2D classification, and the resulting averages were used as templates for automatic particle picking using Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/). In particular, six representative 2D averages were rotated every 30° over 360° to generate seventy-two 2D views and then used as templates in Gautomatch. Particles selected for all micrographs were visually inspected, and poorly selected particles were manually adjusted or discarded. The final datasets for AMPPNP-Mdn1 and AMPPNP-Mdn1-ΔC contained 398,829 and 217,779 particles, respectively.

For AMPPNP-Mdn1, an initial pilot dataset containing 60,632 particles (box size: 300 x 300 pixels) was subjected to 2D classification using RELION 1.4 (Appendix Figure 4-1B). Classes that produced poor averages were discarded. Subsequent 3D classification into five classes was carried out in RELION 2.0, using the ATPI-Mdn1 map filtered to 25 Å as the initial reference. The resulting class averages differed slightly in the angle between the ring region (up to the hinge in the tail) and the rest of the tail. The class average showing the clearest structural features was further refined, resulting in a

density map at ~8.0 Å resolution (FSC = 0.143). This map was used to generate generous masks for the entire protein, the ring region, and the tail region. The full dataset containing 398,829 particles was processed using RELION 2.0 (Appendix Figure 4-1C) (See Table 4-2 for statistics for processing the complete dataset). All particles after 2D clean-up (313,336 particles) were first refined against the final map obtained from the pilot dataset filtered to 25 Å resolution using the mask for the entire protein. Then, 3D classifications into six classes for the ring region and eight classes for the tail region were performed separately, using the local search option and the respective masks. In both cases, the best-resolved class accounts for ~30% of the total number of particles. Refinement of the best class for the ring region (containing 87,580 particles) produced a map at ~4.4 Å resolution, and refinement of the best class for the tail region (containing 94,649 particles) yielded a map at ~5.3 Å resolution (FSC = 0.143). We observed that the FSC curve for the tail region approached the 0.143 level (at ~5.9 Å) but then increased abruptly again before crossing the 0.143 level at a later point (at ~5.3 Å) (Appendix Figure 4-2A). This map was low-pass filtered to 5.9 Å and used in figures. The map of the ring region after post-processing was used as the reference for refinement of all particles after 2D clean-up (313,336 particles) in FREALIGN (Grigorieff, 2016), yielding the final map at a resolution of ~4.0 Å (FSC = 0.143) (Appendix Figure 4-2A). Local-resolution maps of the ring and tail regions of AMPPNP-Mdn1 were calculated by ResMap in RELION 2.0 (Appendix Figures 4-2B and F). These local-resolution maps represent

relative differences in resolution across the maps but the absolute values may not be exact.

Table 4-2. Cryo-EM Data Collection and Reconstruction Statistics of AMPPNP-Mdn1 and AMPPNP-Mdn1-*Δ*C.

	AMPPNP-Mdn1 (ring)	AMPPNP-Mdn1 (tail)	AMPPNP-Mdn1-ΔC (ring)	AMPPNP-Mdn1-ΔC (tail)
Data collection				
Microscope		Titan Krios		
Detector		K2 Summit		
Pixel size (Å)		1.3		
Voltage (kV)		300		
Total electron exposure (e^{-}Å^{-2})		88.8		
Defocus range (μm)		-1.4 to -3.0		
Reconstruction (RELION)				
Particle number	87,580	94,649	21,164	22,921
Box size (pixels)	300	300	300	300
Resolution (Å) (FSC = 0.143)	4.4	5.3	6.1	7.4
B-factor (Å^2)	-171	-211	-161	-178
Reconstruction (FREALIGN)				
Particle number	313,336			
Box size (pixels)	300			
Resolution (Å) (FSC = 0.143)	4.0			
B-factor (Å^2)	-175			

The AMPPNP-Mdn1-*Δ*C dataset was processed similarly as the full dataset of AMPPNP-Mdn1 in RELION 2.0 (Appendix Figure 4-3) (See Table 4-2 for statistics for processing the complete dataset). All particles after autopicking (217,779 particles, box size: 300 x 300 pixels) were subjected to 2D classification. Classes generating poor averages were discarded, and the remaining 175,881 particles were refined to the AMPPNP-Mdn1 map filtered to 25 Å resolution using the mask for the entire protein. 3D classifications into five classes using the local search option were carried out separately for the ring and tail regions with the respective masks. Refinement of the best class

for the ring region (containing 26,164 particles) produced a map at ~6.1 Å resolution, and refinement of the best class for the tail region (containing 22,921 particles) yielded a map at ~7.4 Å resolution (FSC = 0.143) (Appendix Figure 4-3F). We observed the FSC curve for the ring region approached the 0.143 level (at ~6.5 Å) but then increased abruptly again before crossing the 0.143 level at a later point (at ~6.1 Å) (Appendix Figure 4-3F). This map was low-pass filtered to 6.5 Å and used in figures. Local-resolution maps of the ring and tail regions of AMPPNP-Mdn1-ΔC were calculated by ResMap in RELION 2.0 (Appendix Figure 4-3G). These local-resolution maps represent relative differences in resolution across the maps but the absolute values may not be exact.

Model building and refinement.

The structural model for AMPPNP-Mdn1 was built in three steps: first the AAA domains were built without the flexible loops and H2 insertions, then, the linker domain, and finally the N domain, loops and H2 insertions. Phyre2 was used to calculate initial homology models of the AAA domains (Kelley et al., 2015). Flexible loops, as well as the H2 insertions, were removed from the homology models, and the AAA domains were split into their large and small subdomains. These subdomains were fit into the ring region of the ~4.0 Å resolution AMPPNP-Mdn1 map using Chimera (Pettersen et al., 2004). Since most side chains were not resolved in the cryo-EM density map, we removed all side chains beyond the β-carbons using the phenix.pdbtools command in PHENIX (Adams et al., 2010). The β-carbons and sequence information was

preserved to constrain further refinement using the phenix.real_space_refine command in PHENIX with morphing, simulated annealing and secondary structure restrains.

For the density of the hook-shaped tail in AMPPNP-Mdn1, α-helical segments were first identified using the phenix.find_helices_strands command in PHENIX. Polyalanine α-helices were built into the density, which accounted for most, if not all, of the tail density. Guided by the secondary structure predictions from SPIDER2 (Yang et al., 2017) and PSIPRED (Buchan et al., 2013), we were able to trace the backbone and to establish the connectivity of the α-helices from the C-terminal end of the AAA6-L domain up to the hinge. Due to the ambiguity in the helical registry, we only built a polyalanine model for the traced backbone. For the cryo-EM map of the rest of the tail, i.e., the C-terminal density after the hinge, which was masked and refined separately and resolved at ~5.9 Å resolution, we used the identified polyalanine helices without connecting loops to present the overall architecture of the helical bundles (as shown in Figure 4-1C).

For the N domain, loops and H2 insertions, we first combined the models of the AAA domains (the backbone model with β-carbons) and the linker domain up to the hinge (the polyalanine model) and then focused on unassigned densities for the ring region of AMPPNP-Mdn1 map. *Ab initio* model building for the N domain, loops and H2 insertions was carried out in COOT (Emsley et al., 2010), guided by secondary-structure predictions. For regions in which bulky side chains were resolved as bumps (as shown in

Appendix 4-2C), we used these residues as anchor points during model building. Further refinement using the phenix.real_space_refine command in PHENIX with secondary structure restrains and manual adjustments in COOT were performed iteratively, yielding our final model for the N-terminal density. For model validation, the final model for the N-terminal region (up to the hinge) was refined against one of the independent half maps (work) of the corresponding ~4.0 Å resolution map. FSC curves were then calculated between the refined model and half map 1 (work), half map 2 (free) as well as the combined map (Appendix Figure 4-2E).

CHAPTER 5

CRYO-EM STUDIES OF MDN1 IN THE PRESENCE OF ATP AND RBIN-1

Note to readers: the results discussed below arose from a collaborative effort between myself and several colleagues in the Kapoor Lab and the Walz lab at The Rockefeller University and Kawashima lab at The University of Tokyo. A closely related version is in press in Cell (DOI:10.1016/j.cell.2018.09.015). In particular, Dr. Hiroshi Suzuki and Dr. Frank Dimaio provided help in building the models in this chapter. Dr. Shigehiro Kawashima and Dr. Yuki Kobayashi performed the yeast complementation assay.

5.1 Introduction

The structures of AMPPNP-Mdn1 revealed the domain organization of this large protein but could not resolve the position of the MIDAS domain, which is known to interact with other ribosome assembly factors, such as Rsa4 and Ytm1. Therefore, the coupling between the AAA and the MIDAS domains, or in other words, the coupling between the ATPase cycle of Mdn1 and Rsa4, remains an open question. In order to understand Mdn1's remodeling function during ribosome biogenesis, we set out to examine other potential nucleotide states of Mdn1.

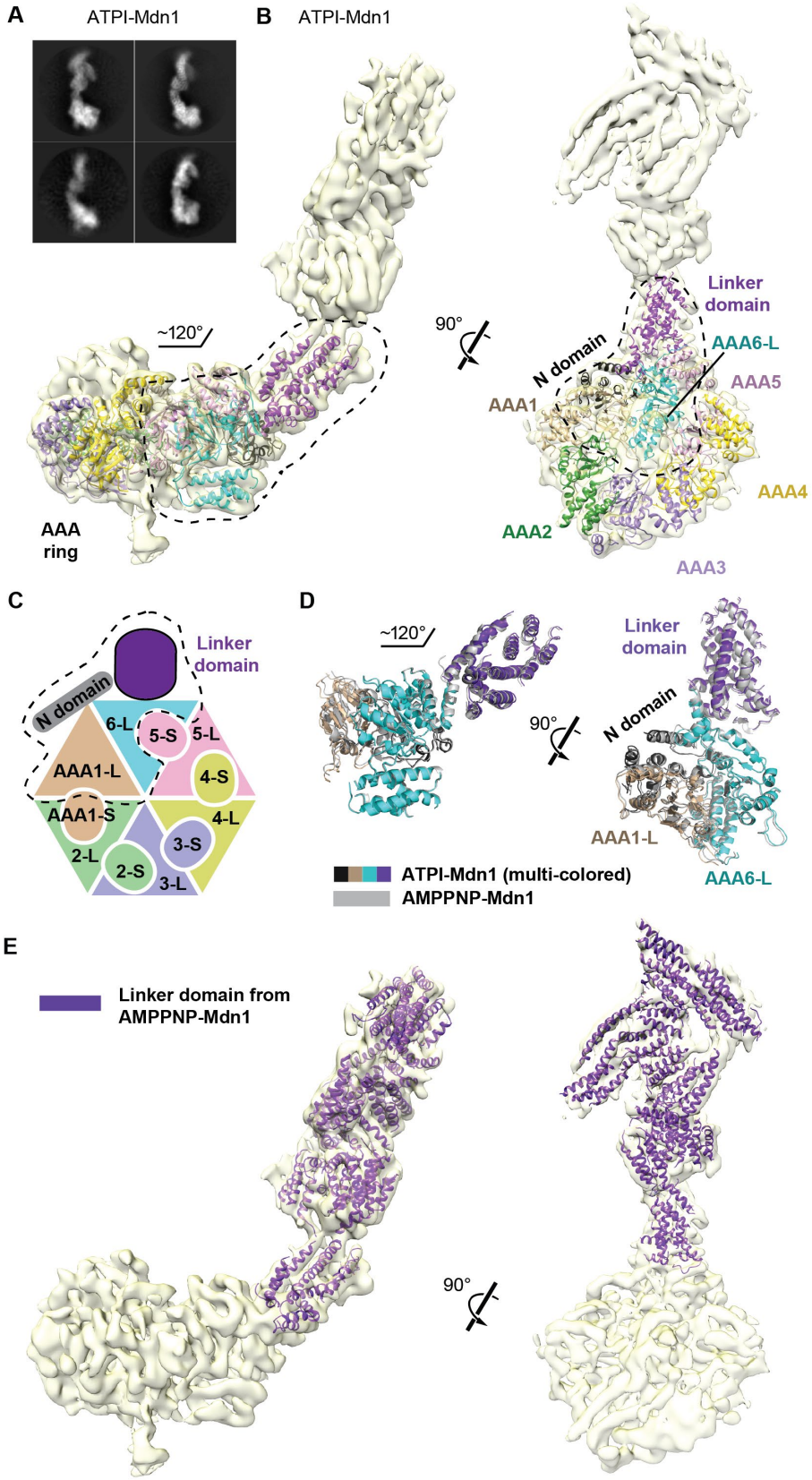
5.2 Cryo-EM structure of Mdn1 in the presence of ATP and Rbin-1

We next examined the potential conformational changes of Mdn1 during ATP hydrolysis by determining the cryo-EM map of Mdn1 in the presence of ATP (1 mM) and Rbin-1 (1 μM) (Kawashima et al., 2016). Two-dimensional class averages revealed an elongated structure similar to that of AMPPNP-Mdn1 (Figure 5-1A). 3D classification of the entire molecule yielded a class with well-defined density for both the ring and tail regions. The final

density map was refined to a resolution of 7.7 Å (Figure 5-1B and Appendix Figure 5-1, hereafter named ATPI-Mdn1), with α-helices and β-sheets, but not individual β-strands, readily distinguishable.

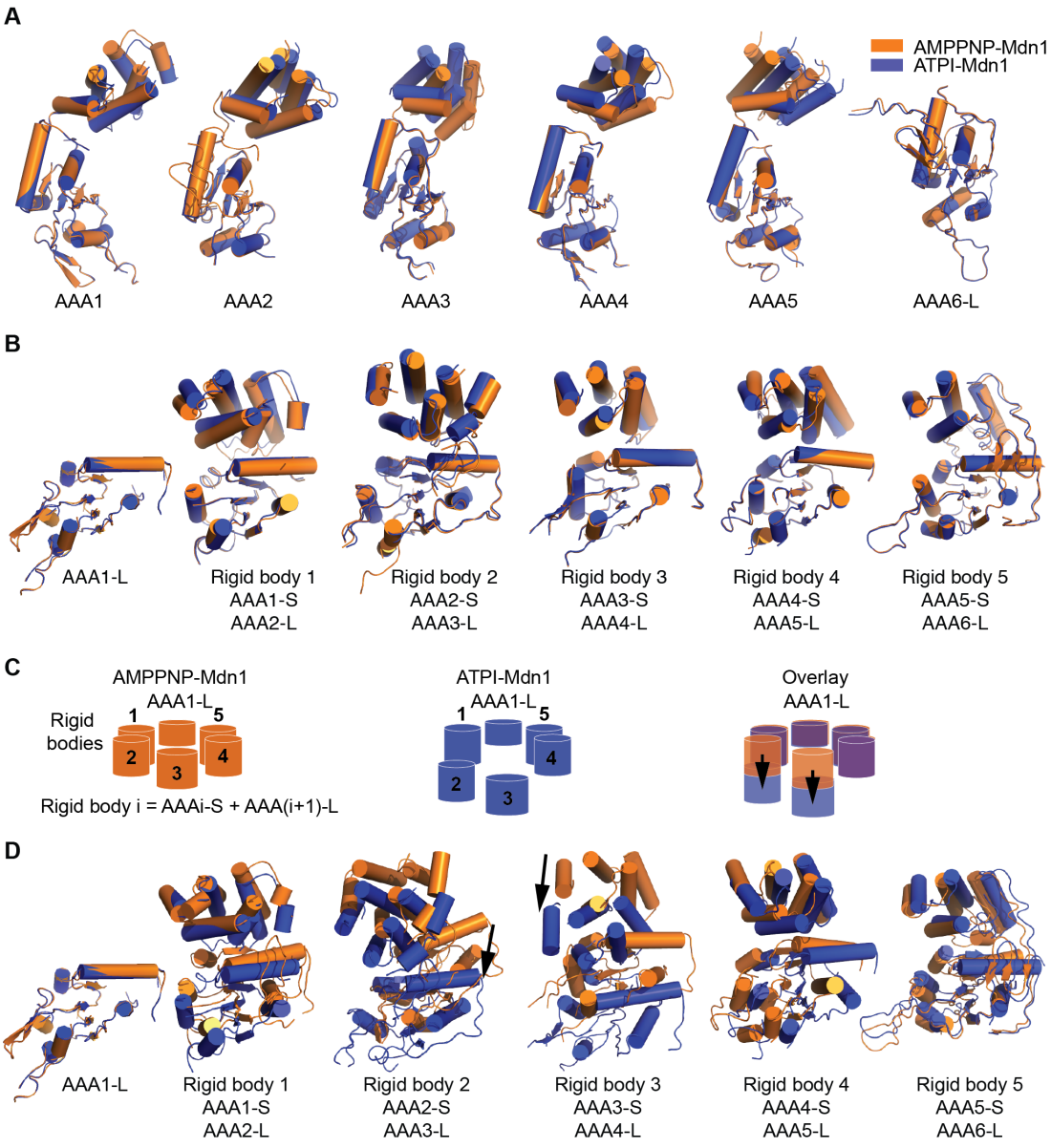
We found that the maps revealed similar overall organizations for ATPI- and AMPPNP-Mdn1 (Figures 5-1A to C). For the N-terminal region of ATPI-Mdn1, up to the hinge in the linker domain, we could build a backbone model using the AMPPNP-Mdn1 model (multicolored ribbons in Figure 5-1B). Nucleotides, Rbin-1 and the H2 insertions of AAA2 and AAA4 could not be readily assigned. Interestingly, ATPI-Mdn1 also has its tail oriented ~120° relative to the AAA ring (left panel in Figure 5-1B), as observed in AMPPNP-Mdn1 (Figure 4-1C). Overlaying the 'ring-linker junction' from these two structures also did not reveal substantial differences (Figure 5-1D). In addition, the poly-alanine model for the entire hook-shaped tail from AMPPNP-Mdn1 could be placed into the ATPI-Mdn1 density as a single rigid body without further adjustment (Figure 5-1E).

Figure 5-1. Overall architecture of Mdn1 in the presence of ATP and Rbin-1. (A) Selected class averages of vitrified Mdn1 in the presence of ATP (1 mM) and Rbin-1 (1 μM), a chemical inhibitor of Mdn1 (hereafter, ATPI-Mdn1). (B) Two views of the density map (yellow surface) with a Cα-backbone model for the region of Mdn1 N-terminal of the hinge point (ribbon colored as in Figures 4-1A). This model does not include the H2 insertions of AAA2 and AAA4. The ring-linker junction, comprised of N domain, AAA1-L, AAA6-L and linker domain, is outlined (black dashed lines). (C) Schematic highlighting the ring-linker junction (dashed line) is shown. (D) Two views of the overlay of the ring-linker junctions in AMPPNP-Mdn1 (grey ribbon) and ATPI-Mdn1 (multi-colored ribbon, colored as in (C)). (E) Two views of the ATPI-Mdn1 map (yellow surface) with the polyalanine-helix model for the entire linker domain (purple ribbon representation) generated based on the AMPPNP-Mdn1 map. A density not accounted for by the polyalanine-helix model is highlighted (dashed circle).


A
ATPI-Mdn1
B
ATPI-Mdn1
~120°
AAA
ring
90°
Linker
domain
N domain
AAA6-L
AAA5
AAA1
AAA4
AAA2
AAA3
C
Linker
domain
N domain
6-L
5-S
5-L
AAA1-L
4-S
AAA1-S
4-L
2-L
3-S
2-S
3-L
D
~120°
90°
Linker
domain
N domain
AAA1-L
AAA6-L
ATPI-Mdn1 (multi-colored)
AMPPNP-Mdn1
E
Linker domain from
AMPPNP-Mdn1
90°

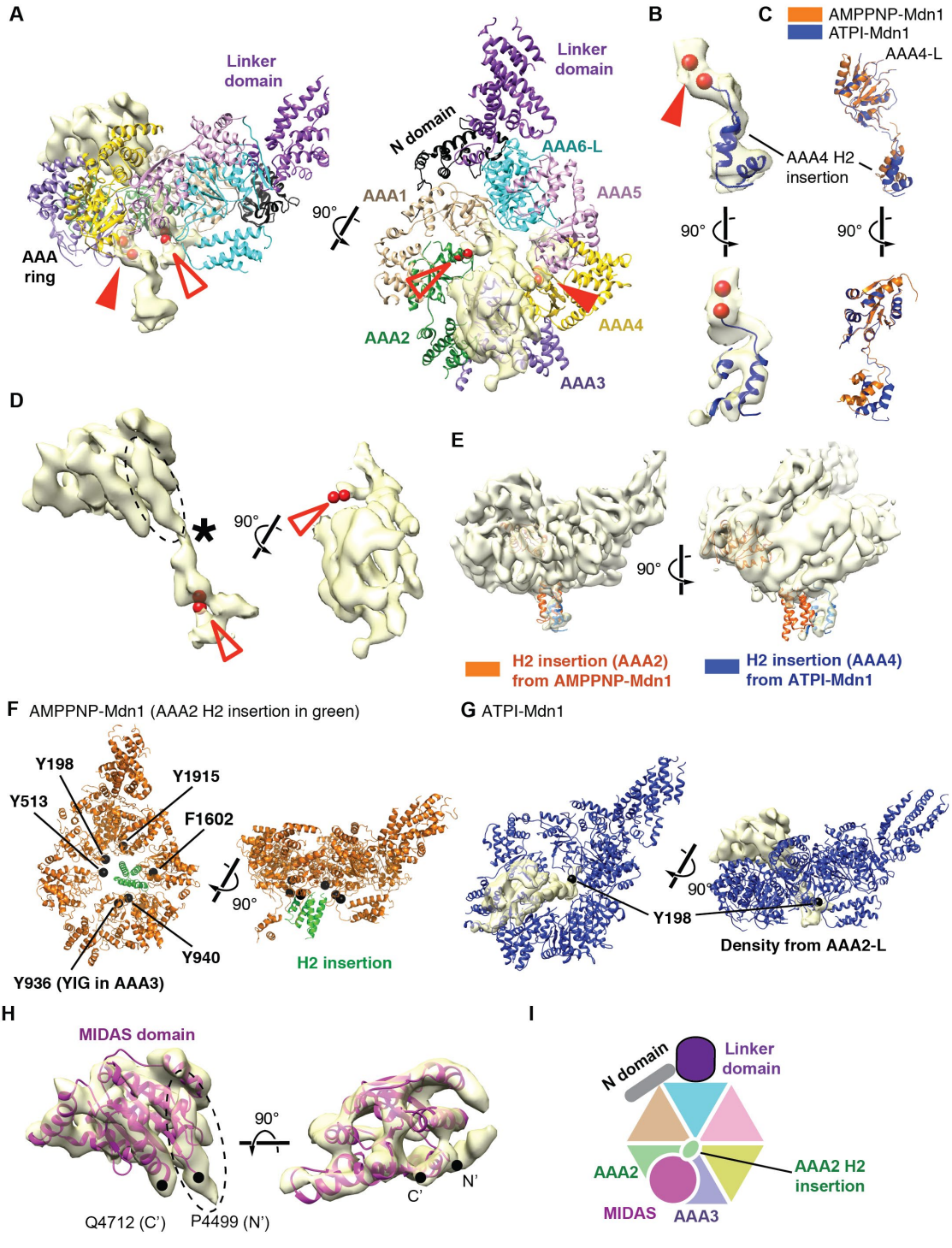
Pairwise comparisons of individual AAA domains reveal that the conformations of AAA1 and AAA6 are similar between AMPPNP- and ATPI-Mdn1 (Figure 5-2A). This may also contribute to the similar orientation of the tail relative to the AAA ring. In contrast, substantial differences are found across the AAA2-AAA5 domains in these two structures (Figure 5-2A). To examine these changes we focused on the structural units called 'rigid bodies' ('rigid body i' is defined as the combination of the AAAi-S and AAA(i+1)-L subdomains; as shown in Figure 5-2A), which during conformational changes in AAA proteins can move as single structural units (see comparisons in Figure 5-2B) (Glynn et al., 2009; Matyskiela et al., 2013). Rigid bodies 1, 4 and 5 rotate slightly but are essentially on the same plane as the AAA1-L subdomain and occupy similar positions in AMPPNP- and ATPI-Mdn1 (Figures 5-2C and D). In contrast, large displacements (>15 Å) of rigid bodies 2 and 3 between the two states are observed (Figures 5-2C and D). Together, these data suggest that local conformational changes of Mdn1's AAA ring are linked to ATPase activity at a subset of sites.

Figure 5-2. Comparisons of AAA domains in AMPPNP-Mdn1 and ATPI-Mdn1. (A) Alignment of the individual AAA domains (AAA1-AAA5 and AAA6-L) in AMPPNP-Mdn1 (orange) and ATPI-Mdn1 (blue). The models were aligned using the large subdomains. (B) Alignment of individual rigid bodies and AAA1-L in AMPPNP-Mdn1 (orange) and ATPI-Mdn1 (blue). The models were aligned based on the large subdomains. Note that there is no substantial conformational difference within individual rigid bodies between the two states, suggesting these rigid bodies move together during conformational changes. (C) Schematics of the five rigid bodies (cylinders) and AAA1L (shorter cylinder) in AMPPNP-Mdn1 (orange) and ATPI-Mdn1 (blue) and an overlay of them aligned using AAA1-L. The displacements of rigid bodies 2 and 3 (black arrows) are also indicated in the overlay. (D) The arrangement of individual rigid bodies in AMPPNP-Mdn1 (orange ribbon) and ATPI-Mdn1 (blue ribbon) when the two models are aligned using their AAA1-L subdomain. The displacements of rigid bodies 2 and 3 (black arrows) are indicated.


A
AMPPNP-Mdn1
ATPI-Mdn1
AAA1
AAA2
AAA3
AAA4
AAA5
AAA6-L
B
AAA1-L
Rigid body 1
AAA1-S
AAA2-L
Rigid body 2
AAA2-S
AAA3-L
Rigid body 3
AAA3-S
AAA4-L
Rigid body 4
AAA4-S
AAA5-L
Rigid body 5
AAA5-S
AAA6-L
C
AMPPNP-Mdn1
AAA1-L
Rigid bodies
1
2
3
4
5
Rigid body i = AAAi-S + AAA(i+1)-L
ATPI-Mdn1
AAA1-L
Overlay
AAA1-L
D
AAA1-L
Rigid body 1
AAA1-S
AAA2-L
Rigid body 2
AAA2-S
AAA3-L
Rigid body 3
AAA3-S
AAA4-L
Rigid body 4
AAA4-S
AAA5-L
Rigid body 5
AAA5-S
AAA6-L

We next examined the densities in the ATPI-Mdn1 map that could not be readily assigned using the AMPPNP-Mdn1 model (Figure 5-3A). First, on the face of the AAA ring, opposite to the one from which the tail extends, there is unassigned density that adjoins AAA4-L (filled red arrowheads in Figures 5-3A). Further, this density is directly connected to the residues flanking the AAA4 H2 insertion (Figures 5-3A, filled red arrowheads) that initially could not be modeled. We find that the three-helix bundle modeled for the AAA4 H2 insertion in the AMPPNP-Mdn1 structure fits into this density (Figure 5-3B). The relative orientation of the AAA4-L subdomain and its H2 insertion differs only slightly between the AMPPNP- and ATPI-Mdn1 structures (Figure 5-3C), suggesting that the AAA4 H2 insertion moves together with the AAA4-L subdomain during the conformational changes of Mdn1.

Figure 5-3. Occlusion of the central hole of the AAA ring and docking of the C-terminal MIDAS domain onto the N-terminal AAA ring in ATPI-Mdn1. (A) Two views of the unassigned density (yellow surface) and the initial ATPI-Mdn1 model (colored as in Figure 4-1A). The red spheres indicate the residues flanking the H2 insertions of AAA4 (filled red arrowheads) and AAA2 (empty red arrowheads). (B) Two views of the unassigned densities (yellow surface) that extend from the residues flanking the AAA4 H2 insertion (red spheres, top panel has the same view as in the left panel of (A)). The model for AAA4 H2 insertion in AMPPNP-Mdn1 was fitted into the density. (C) Alignment of the models of AAA4-L with its H2 insertion from the structures of AMPPNP-Mdn1 (orange ribbon) and ATPI-Mdn1 (blue ribbon). (D) Two views of the unassigned densities (yellow surface) that extend from the residues flanking the AAA2 H2 insertion (red spheres, views as in (A)). A constriction in this density (asterisk) is also highlighted. (E) Potential steric clashes between the H2 insertions of AAA2 and AAA4 in ATPI-Mdn1 if the AAA2 H2 insertion adopts the same conformation as in AMPPNP-Mdn1. The model for AAA2-L with its H2 insertion from AMPPNP-Mdn1 (orange ribbon) was aligned to the model for the AAA2-L subdomain. The model for the AAA4 H2 insertion from the ATPI-Mdn1 structure is shown in blue. (F) Two views of the AAA ring in the model for AMPPNP-Mdn1 (orange ribbon), with AAA2 H2 insertion (green ribbon) highlighted. Conserved aromatic residues (black spheres) near the beginning of the H2 helix in Mdn1's AAA domains (also labeled in Figure 1-4). (G) Model for the AAA ring (blue ribbon) of ATPI-Mdn1 and density representing the AAA2 H2 insertion and the MIDAS domain (yellow surface). The conserved Tyr-198 residue is in direct contact with the density assigned as the AAA2 H2 insertion. (H) Two views of the homology model for the MIDAS domain (magenta ribbon, aa Pro4499-Gln4712) that was fit into the density shown in (B) (yellow surface). Only the density above the constriction in (D) is shown. The N-terminal end of the model (N') connects to unassigned densities (dashed circles in (D) and (F)). (I) Schematic of ATPI-Mdn1 model highlighting the occlusion of AAA ring by AAA2 H2 insertion and the docking of MIDAS domain onto the AAA2 and AAA3 domains is shown.


A
Linker domain
AAA ring
N domain
AAA6-L
AAA1
AAA5
AAA2
AAA4
AAA3
90°
B
C
AMPPNP-Mdn1
ATPI-Mdn1
AAA4-L
AAA4 H2 insertion
90°
90°
D
90°
E
90°
H2 insertion (AAA2) from AMPPNP-Mdn1
H2 insertion (AAA4) from ATPI-Mdn1
F AMPPNP-Mdn1 (AAA2 H2 insertion in green)
Y198
Y1915
Y513
F1602
Y940
Y936 (YIG in AAA3)
90°
H2 insertion
G ATPI-Mdn1
90°
Y198
Density from AAA2-L
H
MIDAS domain
90°
Q4712 (C')
P4499 (N')
C'
N'
I
N domain
Linker domain
AAA2
AAA2 H2 insertion
MIDAS
AAA3

An additional density that extends from the same face of the AAA ring as the AAA4 H2 insertion was assigned to the AAA2 H2 insertion, as it connects to residues flanking the H2 insertion of AAA2 (Figures 5-3A and D, open red arrowheads). This density extends into the central hole of the AAA ring. The additional unassigned density that passes through the ring is not likely to be part of the AAA2 H2 insertion, as it is separated from the density representing the H2 insertion by a constriction (Figure 5-3D, asterisk). In AMPPNP-Mdn1, the AAA2 H2 insertion forms a four-helix bundle and covers only part of the central hole (Figure 4-3A). This four-helix bundle at this position in ATPI-Mdn1 would clash with the AAA4 H2 insertion density (Figure 5-3E). It is possible that the displacement of the AAA4-L domain during the ATPase cycle leads to the reorganization of the AAA2 H2 insertion. The occlusion of the central hole in the AAA ring is reminiscent of substrate binding in AAA unfoldases (Monroe et al., 2017; Puchades et al., 2017; Ripstein et al., 2017).

In AAA unfoldases, conserved aromatic residues in an ‘Ar-Φ-Gly’ tripeptide (Ar: aromatic residue; Φ: hydrophobic residue, named ‘pore loop’) interact with substrates inside the central hole of the AAA ring (Monroe et al., 2017; Puchades et al., 2017). Interestingly, Mdn1’s AAA3 domain possesses a conserved ‘Tyr-Ile/Leu-Gly’ motif. The other AAA domains in Mdn1 lack this motif but are enriched in glycine, aromatic and hydrophobic residues at corresponding sites near the beginning of the H2 helix (Figure 1-4). All of these aromatic residues lie proximal to the central hole in AMPPNP-Mdn1 and

only Tyr-198 in AAA1 could be modeled in our ATPI-Mdn1 map (Figures 5-3F and G). In ATPI-Mdn1, Tyr-198 in the AAA1 domain is in direct contact with the density that we assigned as the H2 insertion of AAA2 inside the central hole (Figure 5-3G), suggesting that this residue may contribute to interactions with and stabilization of the sequence that occludes the central hole.

Finally, the unassigned density on the same face as the hook-shaped tail positioned above rigid bodies 2 and 3 is sufficiently resolved to reveal well-defined secondary-structure elements (Figures 5-3A and D). As substantial portions of the well-defined structural domains in Mdn1, including the N domain, AAA domains and linker domain, are already accounted for the map (Figures 5-1E and 5-3A), the D/E-rich and MIDAS domains are the only candidate contributors to this unassigned density that encloses a volume larger than a typical ~150 aa AAA-L subdomain in the same map. Furthermore, the flexibility of D/E-rich is likely independent of the nucleotide states of the AAA domains as the flexibility results from the highly biased amino acid composition of the domain (>35% Asp/Glu), leaving the MIDAS domain (aa 4381-4717) as the potential identity of this unassigned density. Interestingly, the unassigned density possesses features that were similar to those observed in MIDAS domains of integrin receptors (Song et al., 2005; Xiong et al., 2002). Indeed, a homology model for the MIDAS domain, which consists of a central β-sheet surrounded by seven α-helices, could be unambiguously fit into this density (Figure 5-3H) (Song et al., 2005). Together, these data indicate that, when the central hole is blocked, the C-terminal

MIDAS domain, located ~2000 aa from the N-terminal AAA domains, docks onto the AAA ring (Figure 5-3I).

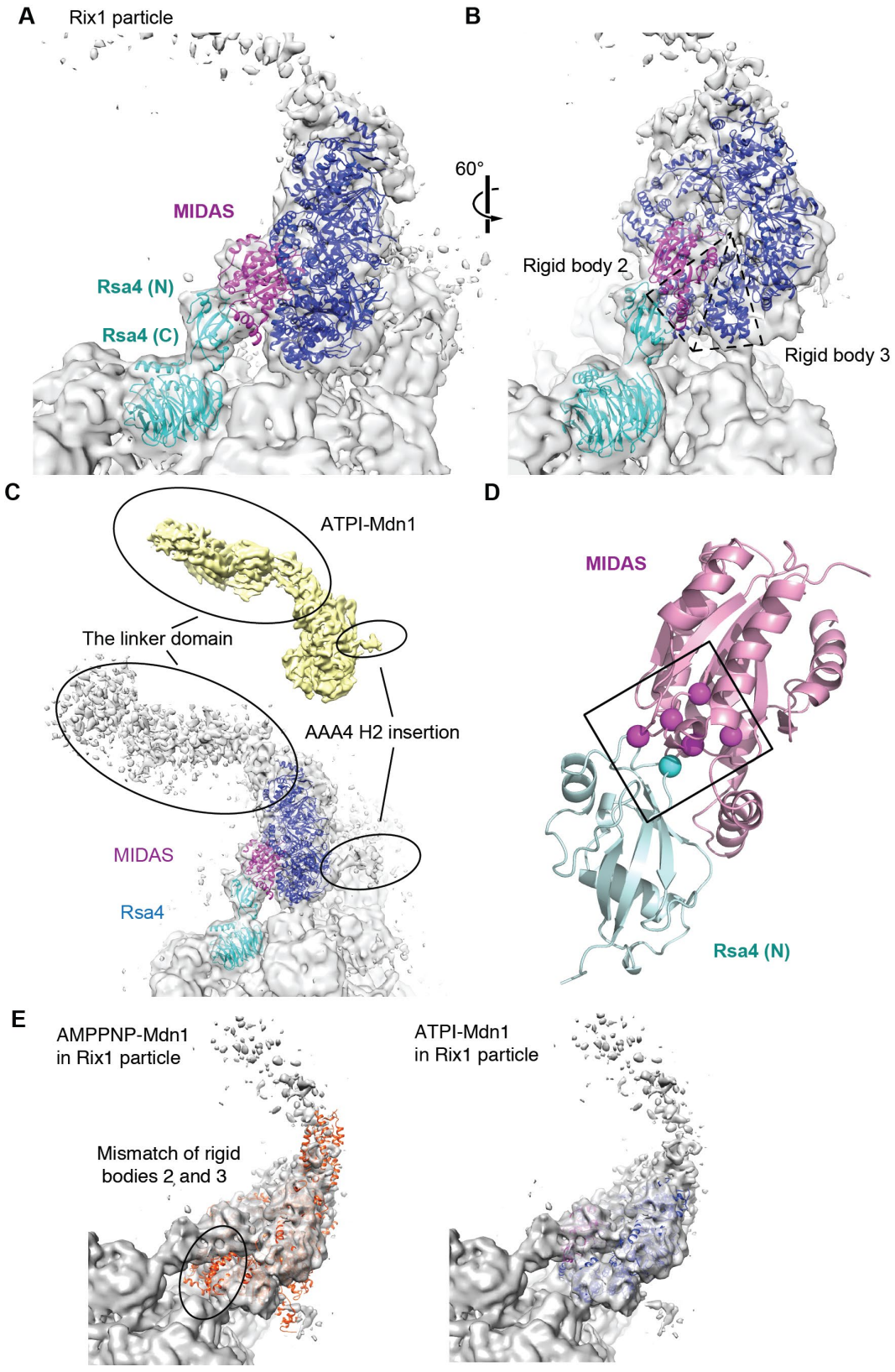
5.3 ATPI-Mdn1 can be docked into cryo-EM maps of the Rix1 particle

Mdn1 is required for the removal of assembly factors, including the removal of Rsa4 from the Rix1 particle, a precursor of the 60S subunit (Kressler et al., 2012; Matsuo et al., 2014; Ulbrich et al., 2009). Recent studies have characterized the structures of the *S. cerevisiae* Rix1 and Nog2 particles (Barrio-Garcia et al., 2016; Wu et al., 2016). These structures indicate that Rea1, the *S. cerevisiae* Mdn1 ortholog, binds the core of the pre-60S particle through a limited portion of the AAA ring's outer edge. The tail of Rea1, for which only weak density is observed, extends away from the Rix1 particle (Barrio-Garcia et al., 2016). We find that our ATPI-Mdn1 model fits well into the Rea1 density (Figures 5-4A to C). Remarkably, the ATPI-Mdn1 MIDAS domain occupies density in the Rix1-particle map that was previously unassigned (Barrio-Garcia et al., 2016; Wu et al., 2016). Importantly, this density that we now assign as the MIDAS domain, is adjacent to the density corresponding to the ribosome assembly factor Rsa4 in the Rix1-particle model (Barrio-Garcia et al., 2016; Wu et al., 2016). In this arrangement, residues that have been shown to be critical for Mdn1-Rsa4 interaction localize to the interface between the MIDAS and Rsa4 domains (Figure 5-4D) (Ulbrich et al., 2009). When contoured at a lower threshold level, the Rix1-particle map also reveals densities that correspond to the linker domain and

the AAA4 H2 insertion in our ATPI-Mdn1 structure (Figure 5-4C). These analyses reveal that our ATPI-Mdn1 structure mirrors the conformation of Mdn1 in the Rix1 particle, in which the MIDAS domain can directly interact with both the AAA domains and Rsa4. This model suggests that inhibition of Mdn1 by Rbin-1 would not preclude the assembly of the Rix1 particle but may block the release of Rsa4, which is consistent with our findings that acute inhibition of Mdn1 by Rbins can result in an enrichment of Rsa4 on Rix1 particles purified from fission yeast cells (Kawashima et al., 2016).

Interestingly, our model also reveals that ATPI-Mdn1 makes direct contact with the core of the Rix1 particle through its rigid bodies 2 and 3 (Figures 5-4A and B), which we find undergo substantial displacements relative to their positions in AMPPNP-Mdn1 to create a binding site for the MIDAS domain. On the other hand, the AMPPNP-Mdn1 model could not be docked into the Rix1-particle map. In particular, there is a clear mismatch between AMPPNP-Mdn1's AAA ring and the density proximal to the core of the Rix1 particle (Figure 5-4E), suggesting that Mdn1 associates with the Rix1 particle in a nucleotide state-specific manner.

Figure 5-4. Mdn1 interacts with other components of the Rix1 particles in a nucleotide state-specific manner. (A and B) Two views of the ATPI-Mdn1 model and the Rsa4 crystal structure (cyan, *C. thermophilum*, PDB: 4WJS) placed into the density map of the Rix1 particle (EMDB: 3199). For clarity, only the AAA ring (blue) and MIDAS domain (magenta) of ATPI-Mdn1 are shown. Details of the model fitting are described in the Methods section. Rigid bodies 2 and 3 of Mdn1's AAA ring (dashed triangles) are indicated in (B). (C) A zoomed-out view of (A) and the map of ATPI-Mdn1 (yellow surface). A lower threshold compared to (A) was used to contour the map of the Rix1 particle. The densities in the Rix1 particle that likely correspond to the linker domain and the AAA4 H2 insertion are circled. (D) The model for the MIDAS domain (magenta) and the N-terminal ubiquitin-like domain of Rsa4 (cyan) shown in the same orientation as (B). The Cα carbons of residues (spheres) implicated in the interaction between the MIDAS domain (Asp-4511, Ser-4513, Ser-4515, Thr-4582 and Asp-4619 in *S. pombe* Mdn1, also labeled in Supplementary Data Item 1) and Rsa4 (Glu-117 in *C. thermophilum* Rsa4, equivalent to Glu-114 in *S. cerevisiae* Rsa4) are highlighted. (E) Docking of the AMPPNP- and ATPI-Mdn1 models generated in this work into the density map of the Rix1 particle (EMDB: 3199). Rigid bodies 2 and 3 of Mdn1's AAA ring (black oval) are indicated in AMPPNP-Mdn1.

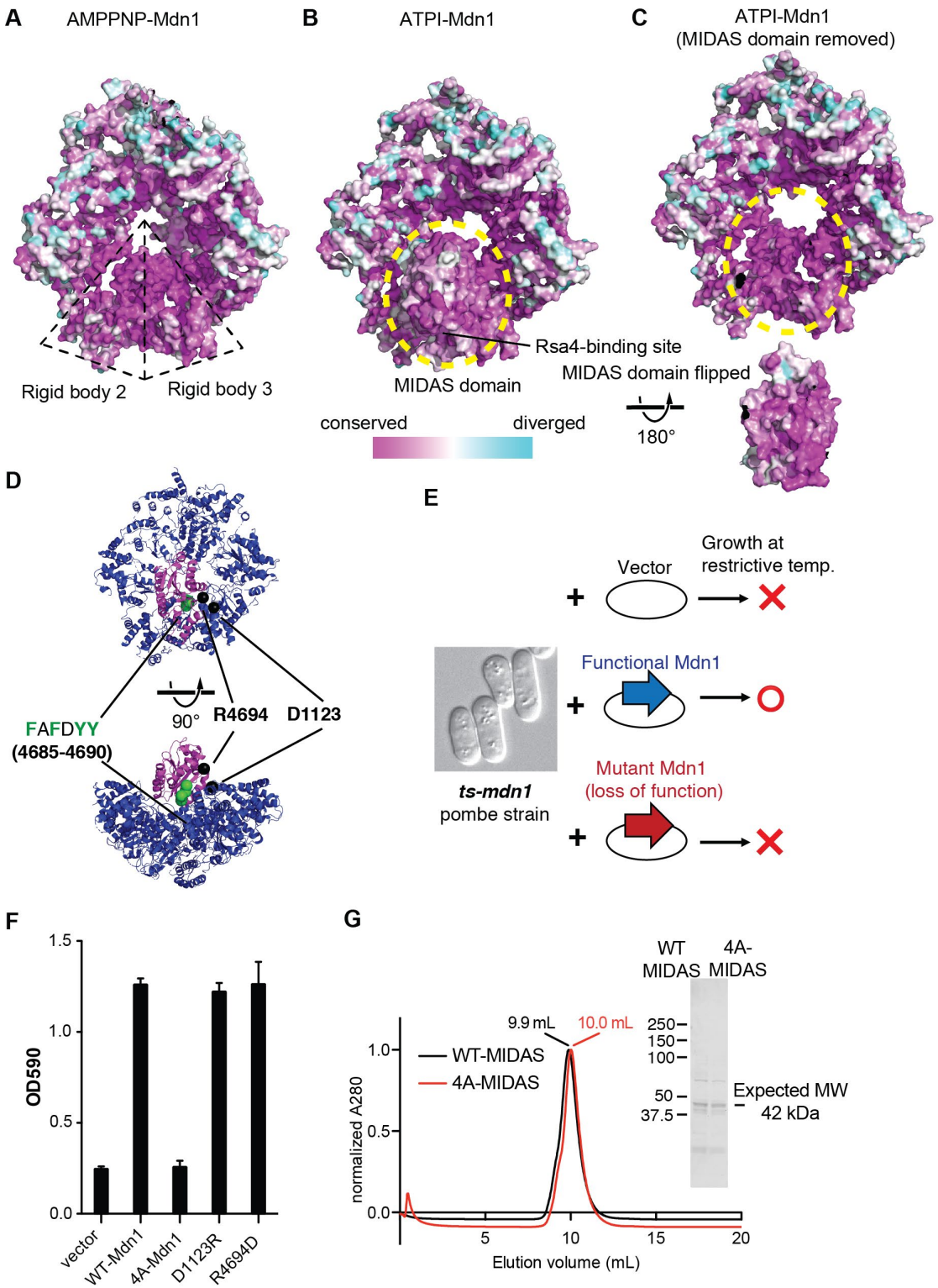

A
Rix1 particle
MIDAS
Rsa4 (N)
Rsa4 (C)
B
60°
Rigid body 2
Rigid body 3
C
ATPI-Mdn1
The linker domain
AAA4 H2 insertion
MIDAS
Rsa4
D
MIDAS
Rsa4 (N)
E
AMPPNP-Mdn1
in Rix1 particle
Mismatch of rigid
bodies 2 and 3
ATPI-Mdn1
in Rix1 particle

We then analyzed the conservation of the primary sequence among Mdn1 orthologs from yeast to human and mapped the conservation scores of individual residues onto the surface models of both AMPPNP- and ATPI-Mdn1 using ConSurf (Figures 5-5A to C) (Ashkenazy et al., 2016). Interestingly, AMPPNP-Mdn1 has a highly conserved surface above rigid bodies 2 and 3 in the AAA ring (Figure 5-5A), which is covered by the MIDAS domain in ATPI-Mdn1 (yellow dashed circle in Figure 5-5B). The surface of the MIDAS domain is overall conserved, including the region facing the AAA ring (Figure 5-5C) and the Rsa4-binding site (Figure 5-5B). The conservation level of the outer circumference of the AAA ring is variable, with higher conservation scores for residues near rigid bodies 2 and 3 that are in direct contact with the core of the Rix1 particle (Figures 5-5A and B).

To test the functional significance of the residues at the interface between the MIDAS domain and the AAA ring we used fission yeast genetics. In particular, we used a '*ts-mdn1* complementation assay', in which different Mdn1 mutants are overexpressed in a temperature-sensitive *mdn1* mutant strain and the growth of cells at the restrictive temperature is measured (Figure 5-5D-F) (Kawashima et al., 2016). We focused on a conserved loop in the MIDAS domain that is enriched in aromatic residues and is buried when the MIDAS domain docks onto the AAA ring. We generated a mutant construct (aa 4685-4690: FAFDYY to AAADAA, hereafter, named 4A-Mdn1) with the goal to disrupt this interaction (Figure 5-5D). As controls, we also tested mutant constructs in which residues at the edge of the same docking interface

were mutated to reverse their charges (D1123R-Mdn1 and R4694D-Mdn1). The cell-based assay indicated that 4A-Mdn1 cannot complement Mdn1 function in cells at the restrictive temperature, while the D1123R and R4694D mutants can (Figure 5-5F). We also generated recombinant wild-type MIDAS (aa 4381-4717) and mutant MIDAS (aa 4381-4717, FAFDYY to AAADAA, named 4A-MIDAS hereafter) domain constructs and found that both proteins eluted from a size-exclusion column at similar volumes (9.9 and 10.0 mL) (Figure 5-5G), suggesting that mutations in the aromatic loop in the MIDAS domain do not affect its overall folding. Together, these data are consistent with our model that the intramolecular docking of MIDAS on the AAA ring is required for Mdn1 function.

Figure 5-5. Analysis of the intramolecular docking of MIDAS domain on the AAA ring using a cell-based assay. (A-C) Residue conservation mapped onto the surfaces of (A) AMPPNP-Mdn1, (B) ATPI-Mdn1, and (C) ATPI-Mdn1 without the MIDAS domain and the surface of the MIDAS domain that interacts with the AAA ring. The region where the MIDAS domain docks onto the AAA ring are highlighted (yellow dashed ovals) in (B) and (C). (D) Two orthogonal views of the ATPI-Mdn1 model showing the Cα atoms of the aromatic residues of the FAFDYY sequence (aa 4685-4690, green spheres) and the Cα atoms of residues D1123 and R4694 (black spheres). (E) Schematic of the yeast complementation assay using a *ts*-mdn1 strain. (F) The OD_{590} values of exponentially growing cell cultures of ts-mdn1 strains overexpressing wild-type (WT) Mdn1 or the indicated mutants (4A-Mdn1: aa 4685-4690, 'FAFDYY' to 'AAADAA', D1123R-Mdn1 and R4694D-Mdn1) of full-length Mdn1 (n=3, mean ± SD). The data for the vector control are also shown. (G) Overlay of the size-exclusion chromatography profiles for the wild-type MIDAS domain (aa 4381-4717) and mutant 4A-MIDAS domain (aa 4381-4717, loop 4685-4690 mutation: FAFDYY to AAADAA). The elution volumes of both proteins are indicated for comparison. The SDS-PAGE analysis (Coomassie blue stain) of WT-MIDAS and 4A-MIDAS are shown in the inset. The major bands match the expected molecular weight of the MIDAS domain.


A
AMPPNP-Mdn1
Rigid body 2
Rigid body 3
B
ATPI-Mdn1
Rsa4-binding site
MIDAS domain
conserved
diverged
C
ATPI-Mdn1
(MIDAS domain removed)
MIDAS domain flipped
180°
D
90°
R4694
D1123
FAFDYY
(4685-4690)
E
Vector
Growth at
restrictive temp.
Functional Mdn1
ts-mdn1
pombe strain
Mutant Mdn1
(loss of function)
F
OD590
1.5
1.0
0.5
0.0
vector
WT-Mdn1
4A-Mdn1
D1123R
R4694D
G
WT
MIDAS
4A-
MIDAS
9.9 mL
10.0 mL
WT-MIDAS
4A-MIDAS
normalized A280
1.0
0.5
0.0
5
10
15
20
Elution volume (mL)
250
150
100
50
37.5
Expected MW
42 kDa

5.4 Discussion

Together, our single-particle cryo-EM studies reveal two distinct conformations of Mdn1. These structural models, along with biochemical and cell-based analyses, suggest how conformational changes in Mdn1 contribute to ribosome biogenesis.

5.4.1 Insights into the Conformational Dynamics of Mdn1's C-terminal Tail

Our studies shed light on the conformational dynamics of Mdn1's C-terminal tail, comprised of the linker, D/E-rich and MIDAS domains. The linker domain adopts similar conformations in all three of our Mdn1 density maps (Figures 4-1C, 4-5G and 5-1B) and also has a similar orientation relative to the AAA ring in Rea1 in the published map (~10 Å local resolution) of the budding yeast Rix1 particle (Figure 5-4C). In our datasets, different 3D classes reveal a hinge point in the linker domain that results in small changes in orientation (<10°, Appendix Figures 4-1B and 5-1B). We did not see any 3D classes in which the linker bends substantially. In addition, the D/E-rich domain was not resolved in any of the aforementioned structures, suggesting its nature of flexibility. We propose that Mdn1's linker domain serves as a structural extension from the AAA ring, similar to a fishing pole, to which the MIDAS domain is tethered through the D/E-rich motif. The flexibility of the D/E-rich domain could allow intramolecular docking of the MIDAS domain on the AAA ring in Mdn1 (see the model in Figure 5-6A), allowing their direct coupling.

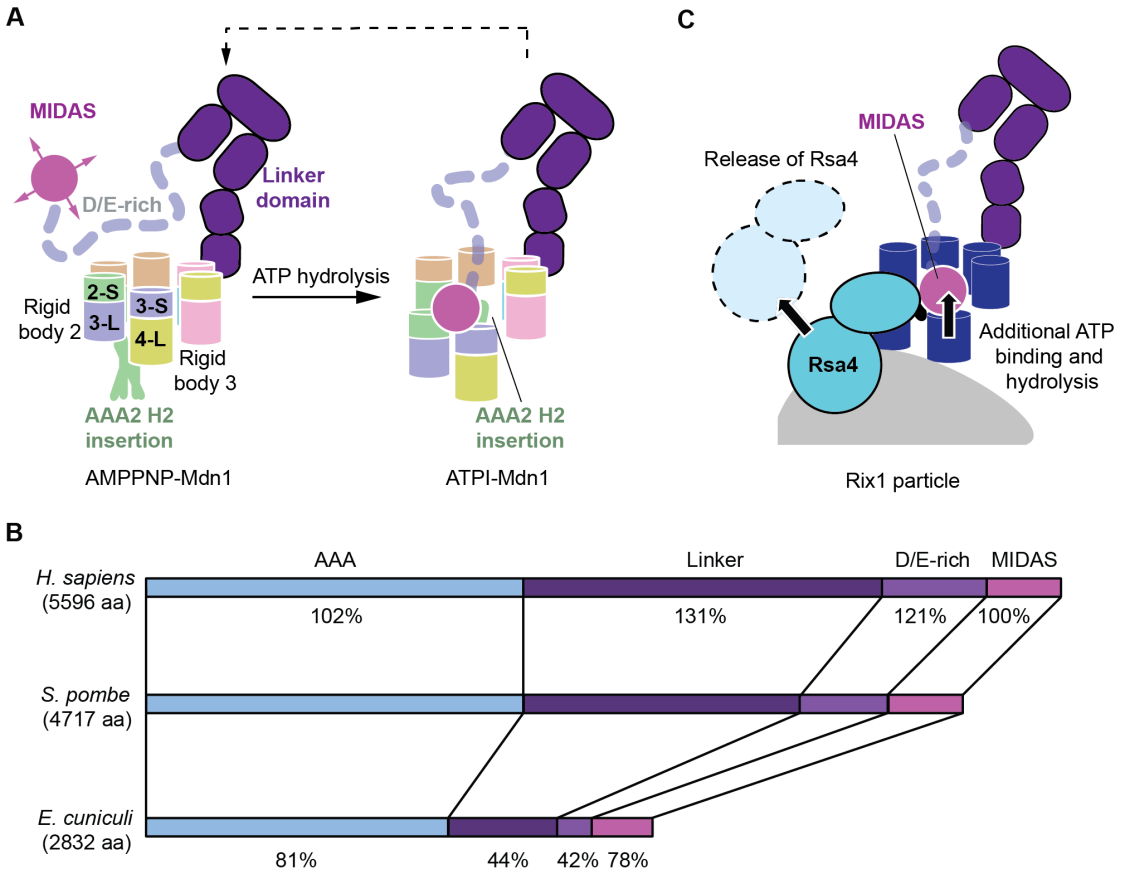


Figure 5-6. Model for the function of Mdn1 during ribosome biogenesis. (A) Schematics of two distinct Mdn1 conformations observed in this study. The linker, D/E-rich and MIDAS domains are indicated. Individual AAA domains are colored as in Figure 4-1A. The magenta arrows indicate the undefined location of the MIDAS domain in AMPPNP-Mdn1. Additional ATP binding and hydrolysis by Mdn1 can lead to additional conformational changes (dashed arrow). (B) Comparison of lengths of individual domains in Mdn1 orthologs from *H. sapiens*, *S. pombe* and *E. cuniculi*. The relative lengths of individual domains (normalized to the length of the corresponding domains in *S. pombe* Mdn1) are listed for *H. sapiens* and *E. cuniculi* Mdn1. (C) Mdn1 adopts a conformation similar to ATPI-Mdn1 in the Rix1 particle. Additional ATP binding and hydrolysis can lead to local conformational changes in the AAA ring and displace the MIDAS domain to release Rsa4 from the Rix1 particle.

Interestingly, alignment of Mdn1 orthologs reveals the evolutionary constraints on individual domains in Mdn1. As mentioned previously, the linker and D/E-rich domains are conserved moderately compared to the AAA and MIDAS domains (sequence identities: ~20% vs ~50%, see Table 4-1). In addition, *H. sapiens* Mdn1 has ~20-30% longer linker and D/E-rich domains compared to *S. pombe* Mdn1, while the lengths of the AAA and MIDAS domains are very similar (<2% difference) (Figure 5-6B). On the other hand, a much smaller Mdn1 ortholog is found in the eukaryotic parasite *E. cuniculi* (2832 aa), in which the primary sequence suggests that the lengths of the linker and D/E-rich domains are reduced by ~60%, compared to *S. pombe* Mdn1, while the lengths of the AAA and MIDAS domains are reduced by only ~20% (Figure 5-6B) (Garbarino and Gibbons, 2002). These comparisons suggest that the linker and D/E-rich domains in Mdn1 are much more prone to genetic drift and tolerant to reductions in length, while there are more stringent constraints on the residues in the AAA and MIDAS domains, which are involved in specific intramolecular or intermolecular interactions critical for Mdn1 function (Figure 5-5). Our analysis also reveals that the length of the linker domain correlates with that of the D/E-rich domain, indicating that a longer linker domain may be compensated by a longer D/E-rich domain serving as a tether, and suggests that our proposed intramolecular docking mechanism for Mdn1 is likely conserved across eukaryotes.

5.4.2 Mdn1 Remodels Precursors of the 60S Ribosome

Our structures also suggest a simple model for how nucleotide-dependent conformational changes of Mdn1 contribute to ribosome biogenesis. In solution, Mdn1 could bind ATP and adopt conformations similar to that of AMPPNP-Mdn1. In this state, the MIDAS domain would sample different positions but not stably dock onto the AAA ring. ATP hydrolysis within the AAA2-AAA5 domains would drive displacements of rigid bodies 2 and 3 (Figure 5-6A). These movements would displace AAA4-L and its H2 insertion, resulting in a steric clash with Mdn1's AAA2 H2 insertion, causing its reorganization such that it inserts into the central hole of the AAA ring. Together, these changes would create a site for the C-terminal MIDAS domain to dock onto the AAA ring (Figure 5-6A). This conformation is capable of stabilizing an interaction network formed by the core of the Rix1 particle, Rsa4 and Mdn1's MIDAS domain and rigid bodies 2 and 3 (Figure 5-6C). It is also possible that Rsa4 and the pre-60S particle core can stimulate Mdn1's activity and catalyze these conformational transitions. As the final step, additional ATP binding, hydrolysis or nucleotide release could reorganize these dynamic regions in Mdn1 and promote the breakdown of the interaction network in the Rix1 particle to release Rsa4 and Mdn1 (Figure 5-6C).

Our proposed model is consistent with previous biochemical assays showing that addition of ATP, but not AMPPNP, to affinity-purified Rix1 particle triggers the release of Mdn1 and Rsa4. It has also been shown that a Walker A lysine-to-alanine substitution at the AAA3 ATPase site does not affect the

association of Mdn1 with the Rix1 particle but blocks the release of Mdn1/Rsa4 upon ATP addition (Barrio-Garcia et al., 2016). Interestingly, our data show that Mdn1 bound to the Rix1 particle adopts a conformation corresponding to ATPI-Mdn1, which has a wider opening in AAA3 compared to what is observed in AMPPNP-Mdn1 (Figure 5-2A), consistent with AAA3 being in the apo state. In addition, our data suggest that rigid bodies 2 and 3 in Mdn1, which flank the AAA3 ATPase site, would likely be displaced in the ATPase cycle. We speculate that the lack of ATP binding or hydrolysis in the AAA3 ATPase site of the AAA3 Walker A mutant restricts conformational changes of rigid bodies 2 and 3 needed for the release of Rsa4. In addition, we can use our structural model for Mdn1 to map the single-point mutations that alter the potency of Rbin-1 in fission yeast cells (Figure 5-7) (Kawashima et al., 2016). Interestingly, seven different resistance-conferring mutations and one sensitivity-conferring mutation all lie in the AAA3-S and AAA4-L subdomains, suggesting that Rbin-1 directly or allosterically affects the conformational transitions of rigid body 3 in Mdn1.

Previous studies have revealed that the H2 insertion of AAA2 in Mdn1 interacts with the C-terminal sequence of Rix1 in a Rix1-Ipi1-Ipi3 complex and this interaction is required for the recruitment of Mdn1 to the Rix1 particle (Barrio-Garcia et al., 2016). Currently, it is not known whether this interaction exists in the Rix1 particle or is only established before the assembly of the Rix1 particle. Our structures reveal the dynamic nature of this H2 insertion in

Mdn1 and suggest that its displacement could be coupled to its binding state with the C-terminal sequence of Rix1.

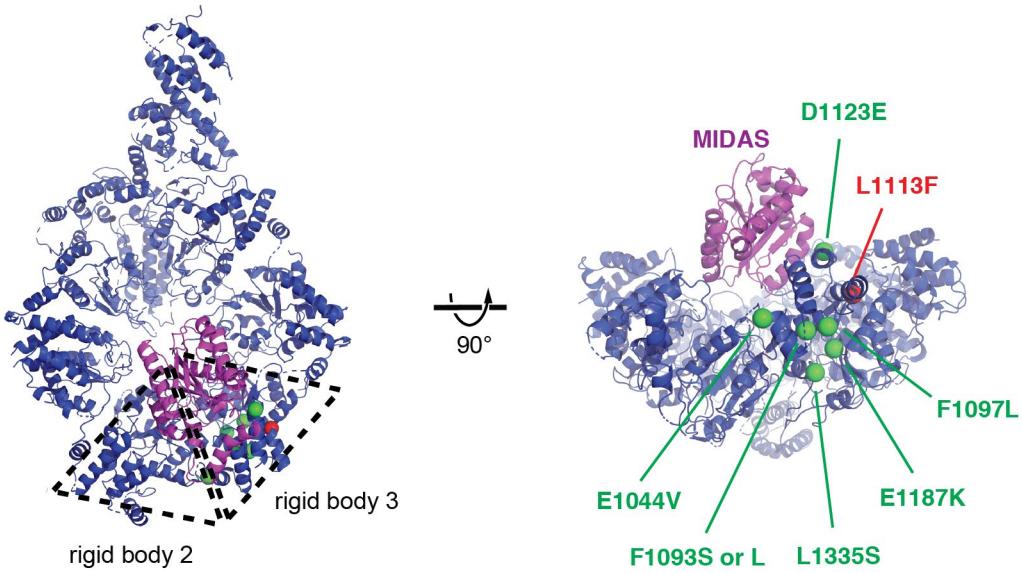


Figure 5-7. Resistance- and sensitivity-conferring mutations to Rbin activities in *S. pombe* cells mapped to ATPI-Mdn1 model. The α carbons of residues that can suppress or sensitize the activity of Rbin-1 in cells are highlighted in the ATPI-Mdn1 model (green and red spheres, respectively). Rigid bodies 2 and 3 are outlined (dashed triangles).

In addition to remodeling the Rix1 particle, previous studies suggest that in *S. cerevisiae*, Mdn1 also releases Ytm1 from nucleolar pre-60S intermediates. This remodeling step requires the interaction between Mdn1's MIDAS domain and the N-terminal domain of Ytm1 (Bassler et al., 2010; Kater et al., 2017). A recent cryo-EM structure of the Nsa1 particle reveals the position of Ytm1 on the pre-ribosomes. Alignment of the Nsa1 particle to the Rix1 particle, using common motifs, indicates that Ytm1 binds to the core of the pre-60S particle at essentially the opposite side from where Rsa4 binds

(separation ~25 nm) (Barrio-Garcia et al., 2016; Kater et al., 2017). However, Rea1 and the N-terminal domain of Ytm1 are not resolved in the Nsa1 particle. Therefore, at this stage, we do not know the precise interactions needed to remodel the Nsa1 particle or other nucleolar pre-ribosomes. It is likely that the MIDAS domain would dock on the AAA ring and then interact with the N-terminal domain of Ytm1, similar to the AAA ring-MIDAS-Rsa4 configuration in the Rix1 particle. This would suggest that Mdn1's AAA ring has to be recruited proximal to Ytm1 in the nucleolar pre-60S particle. Further studies are needed to test this hypothesis.

5.4.3 Comparisons of Mdn1 to dynein and other AAA unfoldases

While the overall conformational dynamics of Mdn1 can be considered to be unlike what has been described for dynein, some common principles are apparent. Dynein, like Mdn1, has a pseudo-hexameric AAA ring and >10 nm protrusions extending from the ring, all within a single polypeptide (Schmidt and Carter, 2016). The ATPase activity of AAA1 in dynein is thought to drive local conformational changes of the AAA ring, and to create two alternative intramolecular binding sites for its cargo-binding domain (Schmidt and Carter, 2016). Our structural data reveal that similar to dynein, Mdn1 also couples nucleotide hydrolysis at a subset of ATPase sites to create an intramolecular docking site for its MIDAS domain.

Mdn1 also shares similarities with AAA unfoldases, such as Yme1 and Vps4. These unfoldases translocate different protein substrates through the

central hole of the AAA ring to unfold them for further processing (e.g., proteolysis) (Monroe et al., 2017; Puchades et al., 2017). In the case of Mdn1, the central hole in its AAA ring is also occupied by a polypeptide whose insertion appears to be nucleotide state-dependent. However, unique to Mdn1, the inserted motif does not derive from a different substrate protein but is a segment of the same polypeptide. We posit that the occlusion of the central hole by the H2 insertion of AAA2 serves an auto-regulatory function. Together, our data reveal how conformational dynamics required for the functions of different AAA proteins can be combined in Mdn1, an unusually large protein, to remodel ribosome precursors.

5.5 Methods

Biochemical preparation of full-length *S. pombe* Mdn1.

The full-length *S. pombe* Mdn1 was prepared as described in Method 2.5.

Recombinant wild-type MIDAS domain (aa 4381-4717) or 4A-MIDAS domain (aa 4685-4690: FAFDYY to AAADAA) were cloned into pET28 vector and overexpressed in *E. coli*. The cells were lysed by French press in lysis buffer (20 mL lysis buffer per 1 L culture; 50 mM Tris [pH 7.5], 150 mM NaCl, 20 mM imidazole, 1 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 3 U/mL benzonase, 1X Roche complete protease inhibitor without EDTA and 10% glycerol). All of the following steps were carried out on ice or at 4 °C. The lysate was centrifuged at 55,000 rpm for 30 min. The supernatant was incubated with Ni-

NTA beads (QIAGEN) for 40 min. The beads were washed using washing buffer (50 mM Tris [pH 7.5], 150 mM NaCl, 20 mM imidazole, 1 mM $MgCl_2$, 5 mM 2-mercaptoethanol and 10% glycerol) and then eluted with 300 mM imidazole in the washing buffer. The eluted protein was concentrated using Amicon Ultra-4 Centrifugal Filter Units (Millipore UFC805008) and loaded on a Superdex 75 16/60 column (GE Healthcare Life Sciences) in size-exclusion buffer (20 mM HEPES [pH 7.5], 150 mM NaCl, 1 mM $MgCl_2$ and 2 mM 2-mercaptoethanol). The eluted fractions were analyzed by SDS-PAGE and fractions corresponding to expected molecular weight (~42 kDa) were concentrated and loaded on a Superdex 75 10/300 column (GE Healthcare Life Sciences) using the same size-exclusion buffer.

Cryo-EM sample preparation and data collection.

To prepare cryo-EM grids of full-length Mdn1 with ATP and Rbin-1, the purified protein was concentrated to ~0.15 mg/mL using centrifugal filters (MW cutoff: 50 kDa; Millipore) and mixed with 10x (ATP+Rbin-1) stock solution (final concentration: 1 mM ATP, 1 μM Rbin-1, 2 mM Na_2SO_4 and 0.2% DMSO). Samples were applied to Quantifoil R1.2/1.3 400 mesh Cu holey carbon grids (Quantifoil) covered with a homemade thin carbon film, and then blotted and frozen in liquid ethane with a Cryoplunge 3 system (Gatan). Cryo-EM data were collected on a 200-kV Talos Arctica electron microscope (Thermo Fisher Scientific) equipped with a K2 Summit detector (Gatan) at a nominal magnification of 28,000x in super-resolution counting mode. After binning over 2 x 2 pixels, the calibrated pixel size was 1.5 Å on the specimen

level. Exposures of 15 s were dose-fractionated into 50 frames with a dose rate of 10 electrons pixel^{-1} s^{-1}, resulting in a total dose of 66.7 electrons Å^{-2}. Defocus values ranging from -2.0 to -4.0 µm were used.

Cryo-EM data processing.

All movie frames were corrected with a gain reference collected in the same EM session, and specimen movement was corrected using MotionCor2 with dose weighting (Zheng et al., 2017). Images showing substantial ice contamination, abnormal background, thick ice or low contrast were discarded. The contrast transfer function (CTF) was estimated using CTFFIND4 (Rohou and Grigorieff, 2015). Approximately 10,000 particles were manually picked, subjected to 2D classification, and the resulting averages were used as templates for automatic particle picking using either RELION 1.4 for the pilot dataset of ATPI-Mdn1 only (Scheres, 2012) or Gautomatch for all other datasets (http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/). In particular, six representative 2D averages were rotated every 30° over 360° to generate seventy-two 2D views and then used as templates in Gautomatch. Particles selected for all micrographs were visually inspected, and poorly selected particles were manually adjusted or discarded. The final dataset for ATPI-Mdn1 contained 247,841 particles, respectively.

For ATPI-Mdn1 dataset, an initial pilot dataset containing 80,152 particles (box size: 256 x 256 pixels) was subjected to 2D classification using RELION 1.4 (Appendix Figure 5-1B). Classes that produced poor averages were discarded. Averages showing representative views were used to

generate an initial density map with the VIPER command implemented in SPARX (Hohn et al., 2007). The particles after 2D clean-up were then subjected to 3D classification into eight classes using the VIPER map as an initial reference. Subsequent refinement and post-processing of the best class yielded a map at 17.5-Å resolution (FSC = 0.143). A mask big enough to cover all 3D classes was then generated and used to process the complete dataset. For the complete dataset, all particles (247,841 particles) were subjected to 2D classification in RELION 2.0 and poor averages were discarded (See Table 5-1 for statistics for processing the complete dataset) (Kimanius et al., 2016). The remaining 210,668 particles were refined to the final map from the pilot dataset filtered to 25-Å resolution with the mask. A subsequent 3D classification into eight classes without alignment yielded one class (containing 54,557 particles) that showed better-resolved structural features. Refinement of this class yielded the final density map at a resolution of 7.7 Å (FSC = 0.143, Appendix Figure 5-1D). Local-resolution map of ATPI-Mdn1 was calculated by ResMap in RELION 2.0 (Appendix Figure 5-1E). These local resolution maps represent relative differences in resolution across the maps but that the absolute values may not be exact.

Table 5-1. Cryo-EM Data Collection and Reconstruction Statistics of ATPI-Mdn1.

	ATPI-Mdn1
Data collection	
Microscope	Talos Arctica
Detector	K2 Summit
Pixel size (Å)	1.5
Voltage (kV)	200
Total electron exposure (e^{-}Å^{-2})	66.7
Defocus range (μm)	-2.0 to -4.0
Reconstruction (RELION)	
Particle number	54,557
Box size (pixels)	256
Resolution (Å) (FSC = 0.143)	7.7
B-factor (Å^2)	-476

Model building and refinement.

To interpret the map for ATPI-Mdn1, the AMPPNP-Mdn1 model was divided into the N domain, the individual AAA subdomains and the linker domain before the hinge. Using Chimera, the models for these domains were individually fit as rigid bodies into the ATPI-Mdn1 density map. The entire AAA6-L subdomain, including its H2 insertion, and the H2 insertion of AAA4 from AMPPNP-Mdn1 fit well into the ATPI-Mdn1 map. The model for the H2 insertions of AAA2 in AMPPNP-Mdn1 could not be readily fit and was therefore not modeled in the ATPI-Mdn1 map. The models for the different domains were then combined, and rigid-body refinement was performed using the phenix.real_space_refine command in PHENIX with secondary structure restraints.

Models for the MIDAS domain were built using RosettaCM (Song et al., 2013). Homology modeling was carried out using seven templates (PDB: 1AUQ, 1IJB, 1MF7, 1SHU, 4HQO, 5HGJ, and 5HJ2) identified by hhsearch (Soding, 2005). For each template, a “partial thread” was generated, in which residue identities were changed to the target sequence but the backbone was kept fixed. Residues not present in each template were not modeled. To determine the overall orientation of these models within the target density in the ATPI-Mdn1 map, a rotational search was carried out in Rosetta around this density region, using a modified version of a previously developed fragment-docking method (Wang et al., 2015). For each of the seven templates, the rotational search identified a single orientation with much better density agreement than any other. Rosetta-based homology modeling was then carried out, combining pieces from the seven homologous structures and refining the models with the Rosetta all-atom energy function that was augmented with a term assessing the agreement of the model to the density. In total, 1500 models were generated, and the models with the lowest overall energy were selected. The five lowest-energy models were quite well converged (inter-structure root-mean-square deviations < 1.5 Å), giving fairly high confidence in the final models.

Fitting of the ATPI-Mdn1 model into the Rix1-particle map.

As shown in Figures 5-5A to C, the model for ATPI-Mdn1 and the crystal structure of Rsa4 (cyan, *C. thermophilum*, PDB: 4WJS) were docked into the published density map of the Rix1 particle (gray surface, EMDB:

3199). The docking of Rsa4 was guided by the reported model for the Rix1 particle (PDB: 5FL8) that was based on the same map (EMDB: 3199). A slight adjustment was necessary to prevent inappropriate fitting of the Rsa4 structure into a previously unknown density, which we assigned, based on our model in the current work, as the MIDAS domain of Mdn1.

Calculation of surface conservation with the ConSurf server.

The Clustal Omega server was used to align the Mdn1 orthologs from sixteen species (*Schizosaccharomyces pombe* Mdn1, Uniprot: O94248; *Saccharomyces cerevisiae* Mdn1, Uniprot: Q12019; *Candida albicans* Mdn1, Uniprot: A0A1D8PL61; *Komagataella phaffii* Mdn1, Uniprot: F2QMK3; *Arabidopsis thaliana* Mdn1, Uniprot: A0A1P8AUY4; *Dictyostelium discoideum* Mdn1, Uniprot: Q869L3; *Caenorhabditis elegans* Mdn1, NCBI: NP_001263748.1; *Drosophila melanogaster* Mdn1, NCBI: NP_001097279.2; *Danio rerio* Mdn1, NCBI: NP_003200751.2; *Xenopus tropicalis* Mdn1, NCBI: NP_002940362.2; *Anolis carolinensis* Mdn1, Uniprot: G1KPU9; *Equus caballus* Mdn1, Uniprot: F6XJW2; *Bos taurus* Mdn1, Uniprot: E1BC24; *Canis lupus* Mdn1, Uniprot: J9NZY7; *Mus musculus* Mdn1, Uniprot: A2ANY6; *Homo sapiens* Mdn1, Uniprot: Q9NU22). The alignment was then analyzed with the ConSurf server to determine the conservation values for the individual residues (Ashkenazy et al., 2016). The output PDB file, which contains the conservation scores as B-factor values, was then displayed in PYMOL.

CHAPTER 6

CONCLUSION AND FUTURE DIRECTIONS

6.1 Conclusion

Previous work done by my colleagues in the Kapoor laboratory discovered Rbin-1 as the first chemical inhibitor of eukaryotic ribosome biogenesis in *S. pombe*. Chemical genetics studies of *S. pombe* revealed that single-point mutations on *mdn1*, a gene that encodes the largest AAA protein (4717 aa) in the yeast genome, could increase or decrease the sensitivities of the cells to Rbin-1.

Here, I established the overexpression and purification strategies to obtain recombinant full-length Mdn1 at high purity. Biochemical characterization of recombinant wild-type Mdn1 and its mutant revealed that they are both active ATPases *in vitro*. In addition, Rbin-1 can inhibit the ATPase activity of the wild-type Mdn1 but not the mutant protein that carries the resistance-conferring single-point mutation, indicating Mdn1 is the physiological target of Rbin-1 in cells.

In addition, I performed structure-activity-relationship analysis of Rbin-1 by synthesizing and testing ~20 analogs in cells. Rbin-2 was discovered to be a 10-fold more potent compound compared to Rbin-1 in the cell-based assay.

I then solved three Cryo-EM structures of *S. pombe* Mdn1 at two different nucleotide states. These structures provided the first pseudo-atomic models of Mdn1 and revealed that the N-terminal AAA domains and the C-terminal substrate-interacting motif, the MIDAS domain, are coupled by directly interacting with each other at specific nucleotide state. These models

also suggest a mechanism of how Mdn1 applies force to its client protein Rsa4 in pre-60S particles.

Together, my chemical biology, biochemical and structural studies of Mdn1 provided new tools to investigate the function of Mdn1 and unveiled an unexpected molecular mechanism for how this unusually large AAA protein contributes to ribosome biogenesis in eukaryotes.

6.2 Future directions

In the following section, I propose to (i) analyze the intermediates of ribosome biogenesis trapped by Rbin-1 in cells to study other steps of ribosome biogenesis that require Mdn1 function, (ii) develop chemical inhibitors of human Mdn1 to dissect of role of ribosome biogenesis in cellular transformation processes and (iii) perform further structural and biophysical analysis of Mdn1 to understand its mechanochemical cycle.

6.2.1 Using Rbins as chemical probes to trap specific ribosome intermediates in cells.

Ribosome biogenesis is a multistep process that involves more than 200 assembly factors. Each assembly factor binds and unbinds to the precursors of ribosomes under tight spatial and temporal regulations, leading to a large number of intermediates of ribosomes with distinct protein compositions. Specific ribosome intermediates can be separated by performing pull-down purification targeting certain assembly factors with

affinity tags. However, this method may bias the purification of non-transient intermediates. Currently, it is not known how many ribosome intermediates exist throughout the whole biogenesis process. In particular, although the function of Mdn1 has been implicated in other steps of ribosome biogenesis in addition to remodeling of Rix1 particles, the exact intermediates with Mdn1 bound have not been isolated and characterized so far.

Interestingly, the addition of Rbin-1 in *S. pombe* cells appears to synchronize the precursors of 60S ribosomal subunits at three specific steps (see the three-step decreases of Rpl2501-GFP signal in the nucleolus after wash out of Rbin-1 in *S. pombe* cells, Figure 2-2F). Therefore, I propose to use Rbin-1 as a perturbation tool to study these specific ribosome intermediates, whose further processing most likely require the function of Mdn1. Using the chemical inhibitors, the endogenous populations of ribosomal intermediates can be enriched. Biochemical fractionation of the nucleus of Rbin-1-treated cells can be performed and sucrose gradient can be used to separate the trapped ribosome intermediates. Mass spectrometry analysis of these fractions will reveal the protein composition. Careful selection of the target proteins for affinity tags can then lead to strategies to separate these distinct intermediates. Further biochemical and structural characterization using single-particle cryo-EM will reveal the architectures of these ribosomal subunits that are relevant to Mdn1 function. These studies can reveal how Mdn1 contribute to other remodeling steps of ribosome biogenesis.

6.2.2 Development of chemical inhibitor targeting human Mdn1.

An enlarged nucleolus has been used as a prognostic marker for tumor development for more than 200 years (Ruggero, 2012). The nucleolus is also

where transcription of ribosomal DNA, i.e. the first step of ribosome biogenesis, occurs. Recent studies suggest that deregulations of oncoproteins and tumor suppressors that occur frequently in a variety of human cancers could lead to not only the loss of control in the cell cycle progression, but also enhanced ribosome biogenesis (Montanaro et al., 2008). For instance, c-Myc and Rb have been shown to regulate the activities of both RNA polymerase I and III that are both responsible for the transcription of rDNA (Cavanaugh et al., 1995; Gjidoda and Henry, 2013; Gomez-Roman et al., 2006; Hirsch et al., 2004). However, whether the cellular transformation process requires extra ribosomes to meet the high demands for protein synthesis or this upregulation of ribosome biogenesis is a bystander effect caused by deregulations of important cancer regulators remains an open question. Currently, it has been difficult to dissect the role of ribosome biogenesis in the cellular transformation process, partially due to the lack of tools to selectively suppress the assembly of ribosomes, which is only one of the pathways affected by proteins like c-Myc or Rb. In addition, recent studies revealed that >20 known chemotherapeutic drugs affect ribosome biogenesis at different stages according to the analysis rRNA processing (Burger et al., 2010), raising the question on how much of the therapeutic effect of these drugs is resulted from the suppression of ribosome biogenesis and whether targeting ribosome biogenesis in cancer cells would be a viable therapeutic approach (Pelletier et al., 2018). Therefore, I propose to generate and purify human Mdn1 and screen for selective cell-permeable inhibitors based on *in vitro* ATPase assays and the available library of Rbin analogs (~100 compounds). These experiments may provide new tools to dissect the role of ribosome biogenesis in the pathogenesis of cancer and it is also interesting to use the inhibitors to

study the biogenesis pathway in human. These studies may also provide new insights into the evolution of ribosome biogenesis in different eukaryotes.

6.2.3 Structural and biophysical studies of Mdn1.

The structures of Mdn1 at two different nucleotide states reported in this dissertation provided two snapshots of Mdn1 during its ATPase cycle. These data revealed the domain organizations of Mdn1 and which domains have to undergo substantial conformational changes between these two states. However, other conformational intermediates and the contributions of individual ATPase sites to the conformational changes are not clear. Therefore, I propose to perform single-particle cryo-EM studies of E1637Q-Mdn1 that cannot hydrolyze ATP in AAA5 ATPase site. Even though there is no nucleotide density observed in the AAA5 ATPase site in AMPPNP-Mdn1 map, the E1637Q-Mdn1 hydrolyzes ATP at a significantly slower rate compared to the wild-type Mdn1 (~15%), suggesting ATP hydrolysis in AAA5, and ATP binding as a prerequisite, are both important for specific steps of the ATPase cycle. Therefore, the structures of this mutant Mdn1 will likely reveal conformations of Mdn1 with ATP bound in AAA5 ATPase site. These studies may provide insights into the role of ATP hydrolysis in AAA5 ATPase site in the mechanochemical cycle of Mdn1.

In addition to the structural analysis, I propose to perform fluorescence resonance energy transfer (FRET)-based assays to examine the kinetics of conformational transitions of Mdn1. In particular, I am interested in the dynamics of the intramolecular docking of the MIDAS domain in Mdn1. Therefore, I propose to introduce HaloTag and SNAP tag to the N- and C-terminus of the full-length Mdn1 and then label them with different

fluorophores that are appropriate for the FRET assays. The kinetics of intramolecular docking of MIDAS domain on the AAA ring can then be analyzed by single-molecule FRET assay. In addition, similar experiments can be performed using E1637Q-Mdn1 and the comparisons between the results of wild-type and mutant Mdn1 will reveal how hydrolysis in AAA5 ATPase site contribute to the intramolecular docking of the MIDAS domain. These studies will reveal the kinetic aspects of the mechanochemical cycle of Mdn1 and also provide broader insights into the potential powerstroke of Mdn1 on pre-ribosomes.

APPENDICES

APPENDIX 2

Targeting pathogenic fungi using Rbin analogs

Note to readers: the results discussed below arose from a collaborative effort between myself and several colleagues in the Kapoor Lab at The Rockefeller University and the Tri-institutional Therapeutics Discovery Institute. With the permission of Prof. Kapoor, I have included studies I contributed and noted other persons' contribution to specific data in the Figure legends.

Appendix 2.1 Introduction

Invasive fungal infections are important causes of morbidity and mortality (Roemer and Krysan, 2014). In particular, the increasing number of patients affected by HIV/AIDS, primary immune deficiency and cancer chemotherapy, dictates a larger pool of vulnerable patients for pathogenic fungi. On the other hand, there are currently only three classes of antifungal drugs approved by FDA: azoles, polyenes and echinocandins (Roemer and Krysan, 2014). The emerging drug-resistant pathogenic fungi strains also highlight the unmet medical needs for the development of new therapeutic reagents.

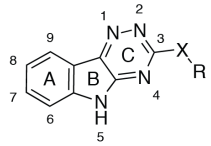
Our studies on Rbin analogs suggest that these cell-permeable compounds potently inhibit the growth of fission yeast by blocking the ribosome biogenesis, an essential process that has not been targeted before. Here, we aimed to optimize the chemical inhibitor by screening a broader range of Rbin analogs for a proof of concept study of using ribosome biogenesis inhibitor as antifungal reagents.

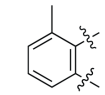
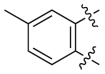
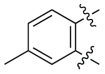
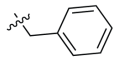
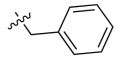
Appendix 2.2 Results

We first aimed to optimize Rbins' potency and potential pharmacological properties. Twenty compounds were designed and synthesized to analyze the Structure-Activity-Relationship of the derivatives based on the tri-cyclic triazinoindole scaffold (Appendix Figures 2-1A to C). Fixed doses (10, 1 and 0.1 μM) of compounds were tested in the fission growth assay in liquid culture (Appendix Figure 2-1D) and dose-dependent inhibition curves were analyzed to determine the potency of the active analogs (Appendix Figures 2-1A, B and E). We discovered that addition of a single methyl group at the A ring of Rbin-1 results in substantial fluctuations of potency, from ~20 fold more potent for the 7-methyl analog (Rbin-4469, GI_{50}= 7 ± 2 nM, n=6, mean ± SD) to ~2 fold less potent for the 6-methyl analog (Rbin-4481, GI_{50}=240 ± 80 nM, n=6, mean ± SD) when compared to Rbin-1 (reported GI_{50}=136 ± 7 nM, n=6, mean ± SD). These results, along with the previous data showing that addition of bromine or iodine substitution on the A ring dramatically change the potency of the analogs, are consistent with a hypothesis that the A ring is close to a well-defined pocket when Rbin analogs bind to their cellular target, so that activities of Rbin analogs in fission yeast are greatly affected by various modifications.

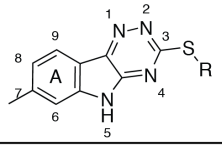
Appendix Figure 2-1. Structure-activity relationship of Rbin derivatives in *S. pombe*. (A-C) Twenty Rbin derivatives are grouped as (A) variations of A ring and thiol linkage, (B) variations of thiol substituents and (C) variations of the triazinoindole ring. Fixed doses of Rbin analogs were mixed with wild-type *S. pombe* cells and kept growing at 29 °C for 18 hours. Half maximum growth inhibition for active compounds (GI_{50}, mean ± SD, n = 6, two independent experiments) are summarized in (A) and (B). The corresponding data are shown in (D) for screening or (E) for dose-response curves. The GI_{50} is determined by fitting relative growth to a four-parameter sigmoidal dose-response curve in PRISM (dose-response curves and GI_{50} values of wild-type *S. pombe* strain (black) or the mutant strain carrying L1113F mutation in Mdn1 that is more sensitive to Rbin-1/Rbin-2 inhibition, mean ± SD, n = 6, two independent experiments). (F) Comparisons of growth inhibition of wild-type *S. pombe* strain and the mutant strain carrying E1187K mutation in Mdn1 that is more resistant to Rbin-1/Rbin-2 inhibition by the Rbin analogs at 1 μM (mean ± SD, n = 6, two independent experiments). Note: The compounds were provided by Tri-Institutional Therapeutic Discovery Institute. The experiments on analyzing the compound activities are performed by myself and Ashley C. Wang, a research assistant in the Kapoor laboratory together.

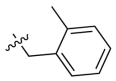
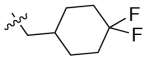
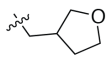
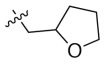
A



	A	X	R	GI_{50} (μM)
Rbin-4483		S		0.018 ± 0.004
Rbin-4467		S		0.15 ± 0.03
Rbin-4469		S		0.007 ± 0.002
Rbin-4481		S		0.24 ± 0.08
Rbin-4686		NH		> 10
Rbin-4541		NH		> 1
Rbin-4839		O		1.1 ± 0.3
Rbin-4542		O		0.47 ± 0.07

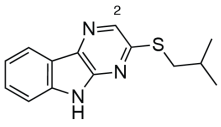
B



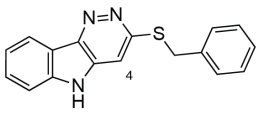
	R	GI_{50} (μM)
Rbin-5041		0.025 ± 0.006
Rbin-5138		0.069 ± 0.04
Rbin-6294		0.017 ± 0.002
Rbin-6311		0.0026 ± 0.0003
Rbin-6312		0.021 ± 0.003
Rbin-5032		0.18 ± 0.05
Rbin-5179		0.7 ± 0.1
Rbin-6299		0.10 ± 0.01

C

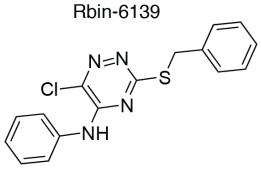
Rbin-6127



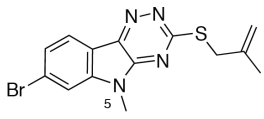
Rbin-6206

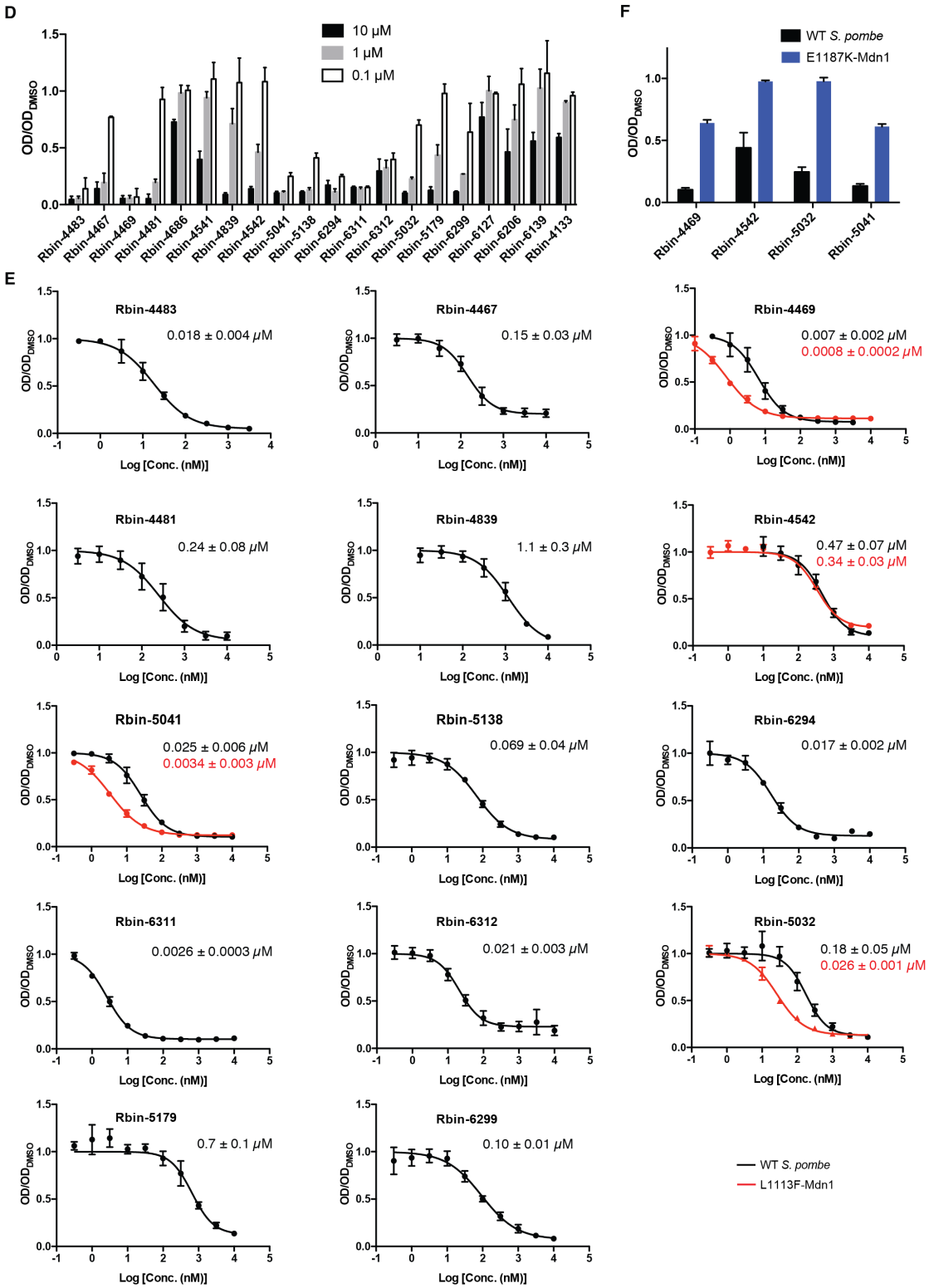


Rbin-6139



Rbin-4133



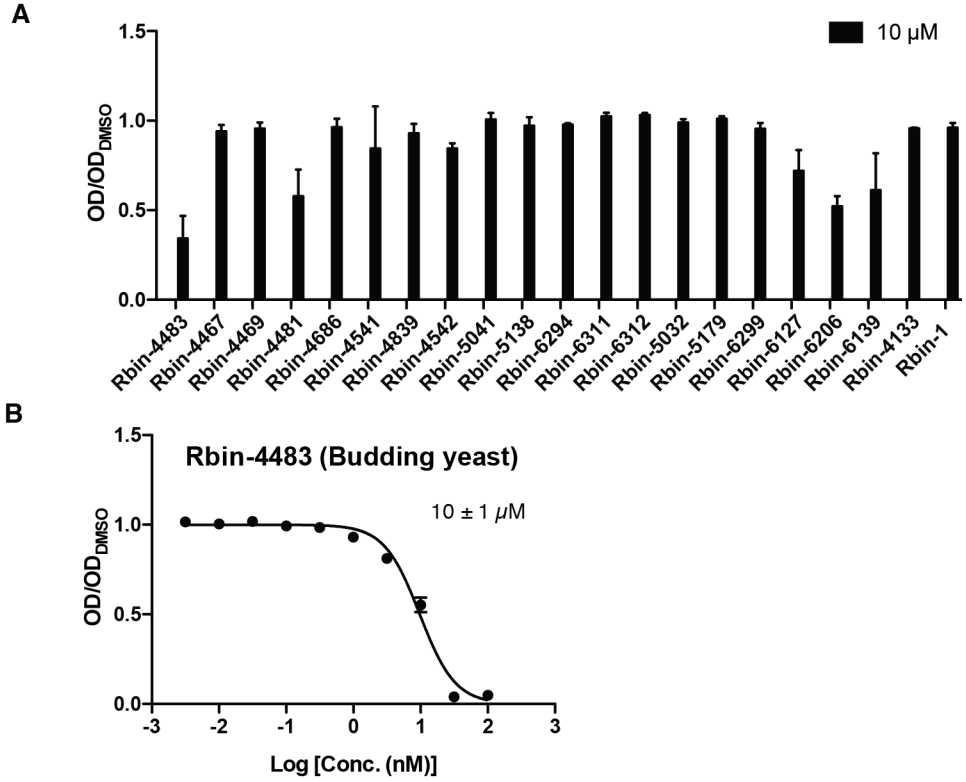

D
10 μM
1 μM
0.1 μM
OD/OD_DMSO
Rbin-4483
Rbin-4467
Rbin-4469
Rbin-4481
Rbin-4686
Rbin-4541
Rbin-4839
Rbin-4542
Rbin-5041
Rbin-5138
Rbin-6294
Rbin-6311
Rbin-6312
Rbin-5032
Rbin-5179
Rbin-6299
Rbin-6127
Rbin-6206
Rbin-6139
Rbin-4133
F
WT S. pombe
E1187K-Mdn1
Rbin-4469
Rbin-4542
Rbin-5032
Rbin-5041
E
Rbin-4483
0.018 ± 0.004 μM
Rbin-4467
0.15 ± 0.03 μM
Rbin-4469
0.007 ± 0.002 μM
0.0008 ± 0.0002 μM
Rbin-4481
0.24 ± 0.08 μM
Rbin-4839
1.1 ± 0.3 μM
Rbin-4542
0.47 ± 0.07 μM
0.34 ± 0.03 μM
Rbin-5041
0.025 ± 0.006 μM
0.0034 ± 0.003 μM
Rbin-5138
0.069 ± 0.04 μM
Rbin-6294
0.017 ± 0.002 μM
Rbin-6311
0.0026 ± 0.0003 μM
Rbin-6312
0.021 ± 0.003 μM
Rbin-5032
0.18 ± 0.05 μM
0.026 ± 0.001 μM
Rbin-5179
0.7 ± 0.1 μM
Rbin-6299
0.10 ± 0.01 μM
Log [Conc. (nM)]
WT S. pombe
L1113F-Mdn1

We next focused on addressing the potential liability of the sulfur linkage in the Rbin scaffold as oxidation of the sulfur ether to sulfoxide or sulfone can potentially make this substituent a good leaving group for aromatic nucleophilic substitution. Briefly, we discovered that substitution of the sulfur ether linker by a secondary amine linker 'NH' almost completely abolished the effect of growth inhibition in yeast (Rbin-4541 and Rbin-4686). On the other hand, the analogs with oxygen ether linker with isobutene substituent retain the potency (GI50: 1.1 ± 0.3 μM (Rbin-4839); 0.47 ± 0.7 μM (Rbin-4542); n=6, mean ± SD).

We also tried to change the sulfur substituents based on the 7-methyl analog in order to avoid potential oxidation of the unsaturated double bond and reduce the hydrophobicity of Rbin-1 (Appendix Figure 2-1B). Saturated alkyl chain or ring analogs possess weakened activities in the growth assay. In addition, analogs with more hydrophilic groups, such as difluoro cyclohexane and tetrahydrofuran (THF), are all potent inhibitors of yeast growth (GI50< 0.2 μM). These results suggest the thiol substituents can be either hydrophobic or hydrophilic for active Rbin analogs.

Analogs with modifications of the tricyclic core (triazinoindole) in Rbin-1 were also synthesized and tested in the growth assays (Appendix Figures 2-1C and D). In particular, two analogs with one fewer nitrogen atoms in the triazine ring abolished the activities (Rbin-6127 and Rbin-6206). Breakdown of the central five-membered ring (B ring as labeled in Appendix Figure 2-1A) also gave rise to inactive compounds (Rbin-6139). Adding a methyl group to the indole nitrogen on an active compound, Rbin-2 (reported in 2.3 Results), also leads to inactive compound (Rbin-4133). In sum, the SAR analysis of Rbin-1 analogs suggests that activities of Rbin in fission yeast require the

tricyclic triazinoindole core while the A ring and thiol substituents are amenable to changes. We also analyzed the effect of these twenty compounds on the growth of S. cerevisiae, another widely used model fungi organism in laboratories. Most compounds are not active against budding yeast except Rbin-4483 (GI_{50}: 10 ± 1 μM).

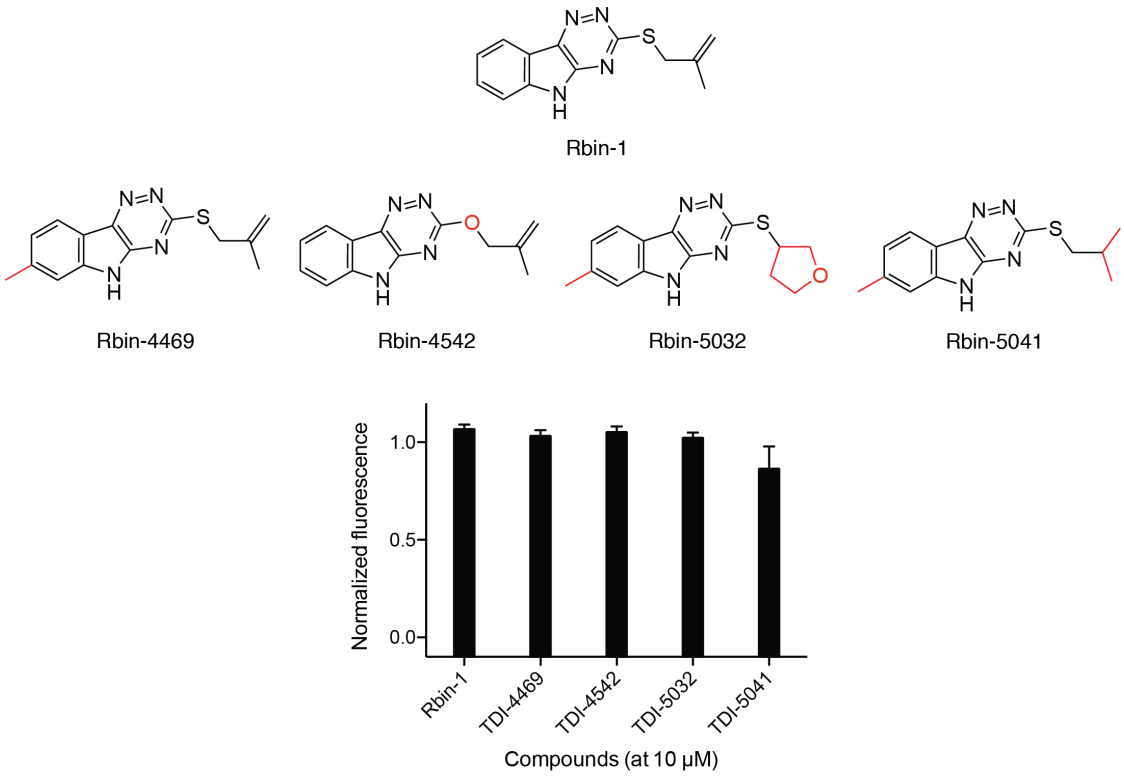


Appendix Figure 2-2. Structure-activity relationship of Rbin analogs in *S. cerevisiae*. (A) *S. cerevisiae* cells with deleted efflux pumps were incubated with Rbin analogs (10 μM) and kept growing at 29 °C for 18 hours. The relative OD_{600} values of cultures with compounds to cultures with DMSO are plotted (mean ± SD, n = 4, two independent experiments). (B) Half maximum growth inhibition for Rbin-4483 was determined by fitting relative growth to a four-parameter sigmoidal dose-response curve in PRISM (GI_{50}= 10 ± 1 μM, mean ± SD, n = 6, two independent experiments). (Note: The experiments on analyzing the compound activities are performed by Ashley Wang, a research assistant in the Kapoor laboratory).

Activities of the lead compound, Rbin-1 and four different chemotypes were further examined in the following studies: 1) Rbin-4469, which is the most potent compound in fission yeast; 2) Rbin-4542, which has an oxygen ether linker and a benzyl substituent; 3) Rbin-5032, which has a more hydrophilic THF substituent on the sulfur atom; 4) Rbin-5041, which has saturated isobutyl substituent on the sulfur atom.

First, I analyzed the toxicities of Rbin-1 and the four representative chemotypes in HCT116 cell line, a human colon cancer cell line (Appendix Figure 2-3). At 10 μM, all five compounds did not affect the growth of the human cell line compared to DMSO control (final DMSO: 1%), suggesting these compounds can potentially be used as anti-fungi reagents in mammalians.

The potential antifungal activities of the four representative chemotypes were then tested in a variety of pathogenic fungi including *Candida parapsilosis*, *Candida krusei*, *Candida albicans*, *Candida glabrata*, *Cryptococcus neoformans*, *Aspergillus fumigatus*, *Paecilomyces variotii*, *Fusarium* sp., *Rhizopus arrhizus*, *Scedosporium*, *Coccidioides* sp. and *Histoplasma capsulatum*. Interestingly, Rbin-4469 has MIC values <0.03 μg/mL and <0.25 μg/mL of against *Coccidioides sp.* and *Histoplasma capsulatum* according to liquid culture growth assays (Data done by others).



Appendix Figure 2-3. Toxicities of Rbin analogs in HCT116 cell line. (A) *S. cerevisiae* cells with deleted efflux pumps were incubated with Rbin analogs (10 μM) and kept growing at 29 °C for 18 hours. The relative OD_{600} values of cultures with compounds to cultures with DMSO are plotted (mean ± SD, n = 4, two independent experiments). (B) Half maximum growth inhibition for Rbin-4483 was determined by fitting relative growth to a four-parameter sigmoidal dose-response curve in PRISM (GI_{50}= 10 ± 1 μM, mean ± SD, n = 6, two independent experiments).

Appendix 2.3 Discussion

Our results suggesting activities of Rbin derivatives in *S. pombe* can be tuned by making different modifications in the substituents of the tri-cyclic triazinoindole core but not the core itself, suggesting the heteroatoms in the core are required to interact with their protein targets.

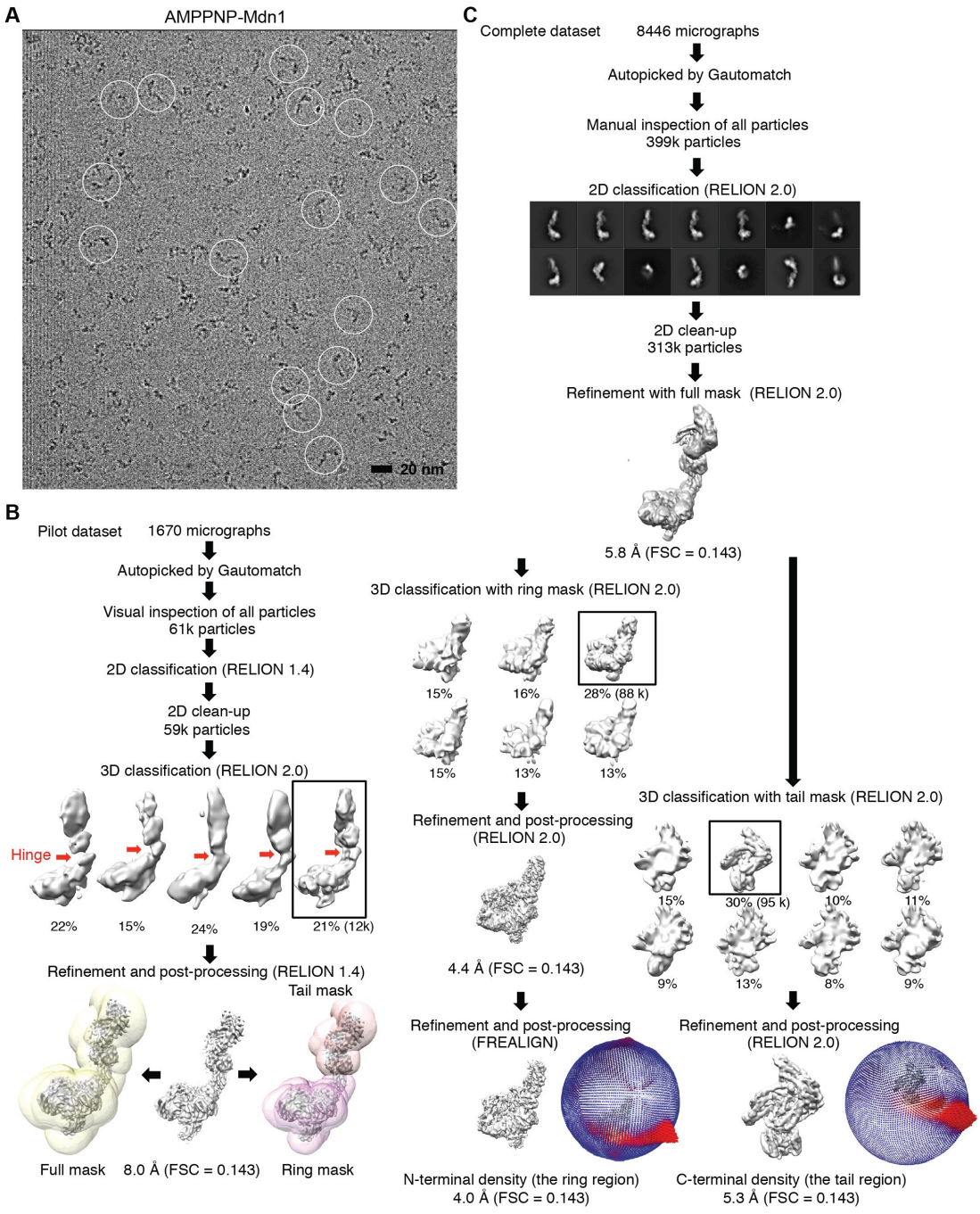
The lack of activities of Rbin derivatives in *S. cerevisiae* and human cells may result from th`e differences between *S. cerevisiae* and *H. sapiens* and *S. pombe* Mdn1 or difference in compound metabolism or efflux pumps in these different organisms.

The activities of Rbin-4469 against only Coccidiodies and Histoplasma cells are interesting as these two organisms are closely related in the phylogenic tree of fungi. Experiments using mouse model for *Coccidiodies* or *Histoplasma* infection are currently undergoing to investigate whether Rbin analogs indeed can be used to suppress the pathogenic fungi *in vivo*.

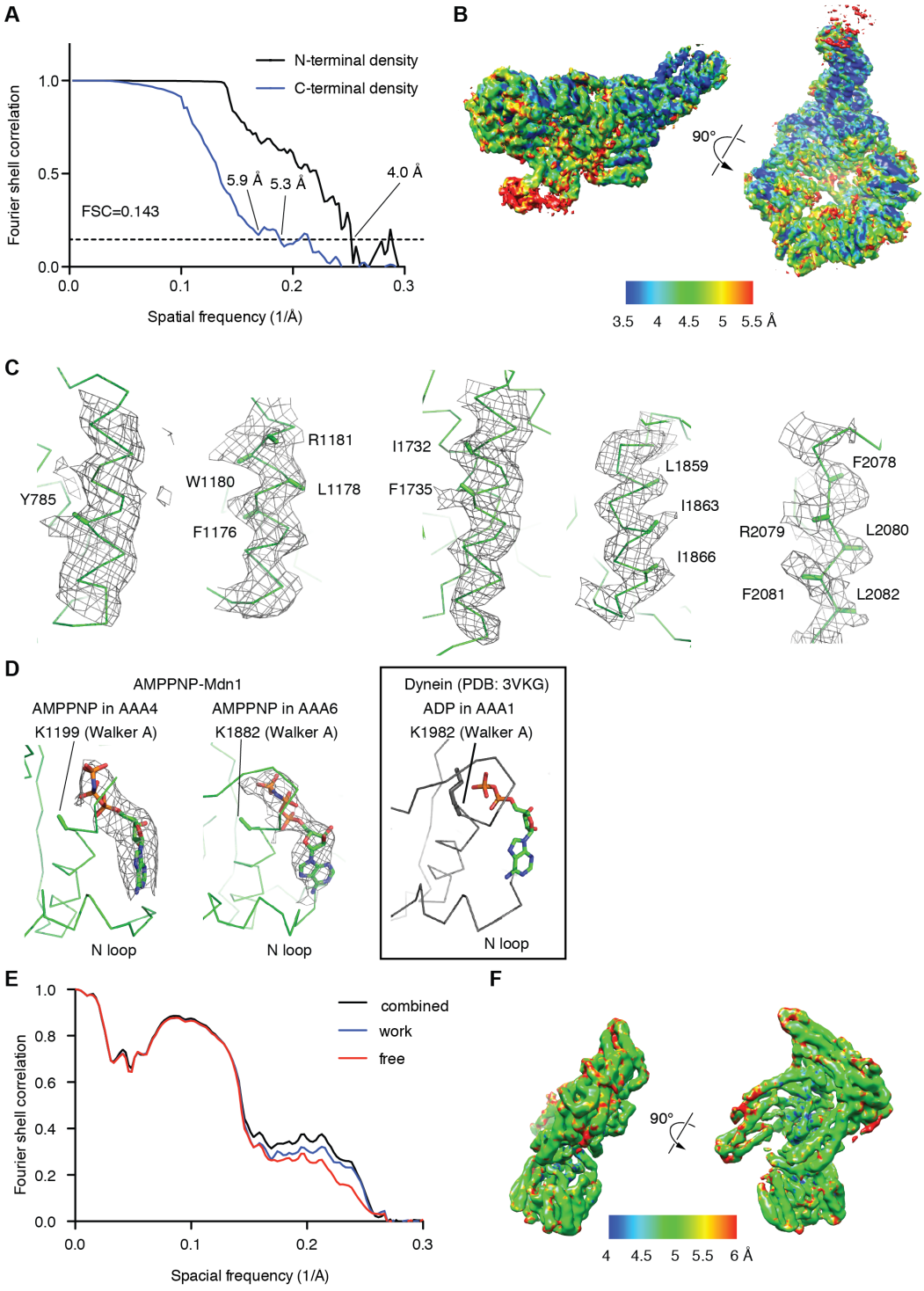
APPENDIX 4

EM processing of AMPPNP-Mdn1 and Mdn1-*Δ*C

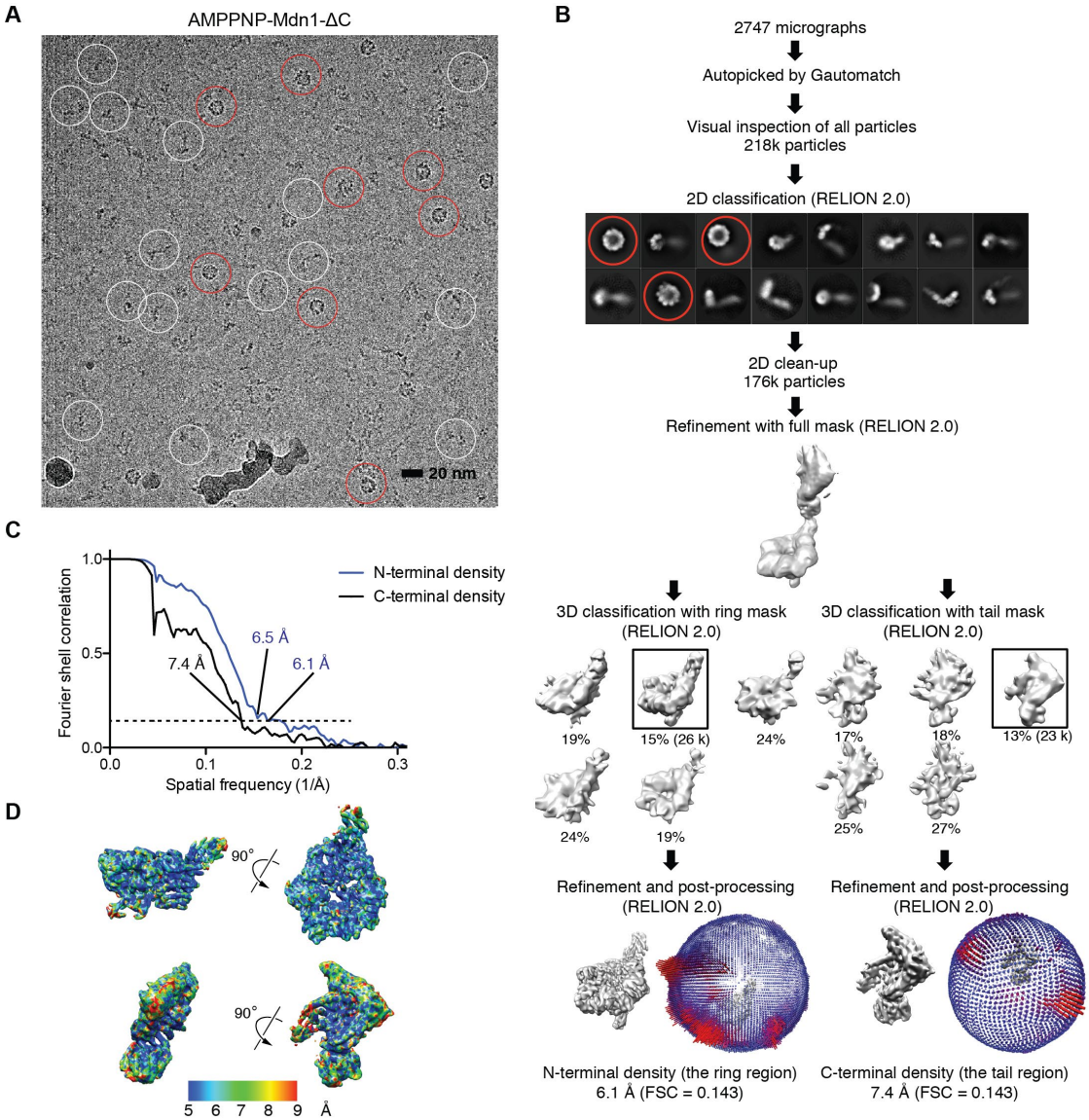
Appendix Figure 4-1. Processing of the Cryo-EM data for full-length Mdn1 in the presence of AMPPNP. (A) A cryo-EM image of AMPPNP-Mdn1 processed with MotionCor2. Some Mdn1 particles are marked (white circles). Scale bar: 20 nm. (B) The workflow used to process the pilot dataset. 3D classification revealed structural variability due to movement around a hinge point (red arrows). After refinement of the class showing the most structural detail, the final map was used to generate masks for the entire elongated molecule, as well as for the AAA ring region (N-terminal of the hinge point) and tail region alone (C-terminal of the hinge point). The density map was also used as the initial reference for processing the complete dataset. (C) The workflow used to process the complete dataset. The angular distributions of the particles in the final reconstructions are shown next to the corresponding maps. Note: the resolution of the final reconstruction is further discussed in Appendix Figure 4-2A.


A
AMPPNP-Mdn1
20 nm
B
Pilot dataset 1670 micrographs
Autopicked by Gautomatch
Visual inspection of all particles
61k particles
2D classification (RELION 1.4)
2D clean-up
59k particles
3D classification (RELION 2.0)
Hinge
22%
15%
24%
19%
21% (12k)
Refinement and post-processing (RELION 1.4)
Tail mask
Full mask
8.0 Å (FSC = 0.143)
Ring mask
C
Complete dataset 8446 micrographs
Autopicked by Gautomatch
Manual inspection of all particles
399k particles
2D classification (RELION 2.0)
2D clean-up
313k particles
Refinement with full mask (RELION 2.0)
5.8 Å (FSC = 0.143)
3D classification with ring mask (RELION 2.0)
15%
16%
28% (88 k)
15%
13%
13%
Refinement and post-processing
(RELION 2.0)
4.4 Å (FSC = 0.143)
Refinement and post-processing
(FREALIGN)
N-terminal density (the ring region)
4.0 Å (FSC = 0.143)
3D classification with tail mask (RELION 2.0)
15%
30% (95 k)
10%
11%
9%
13%
8%
9%
Refinement and post-processing
(RELION 2.0)
C-terminal density (the tail region)
5.3 Å (FSC = 0.143)

Appendix Figure 4-2. Cryo-EM analysis of full-length Mdn1 in the presence of AMPPNP. (A) Gold-standard Fourier Shell Correlation (FSC) curves calculated between independently refined half maps (each half map was reconstructed using randomly selected 50% of the particles) for the N- and C-terminal regions of AMPPNP-Mdn1 relative to the hinge point. For the map of the C-terminal density, the resolution was estimated as ~5.9 Å, more conservatively than the value suggested by the FSC = 0.143 criterion (~5.3 Å). (B) Local-resolution map of the N-terminal density that includes the AAA ring calculated by ResMap in RELION 2.0. (C) Representative EM densities are shown (grey mesh) for the map of the N-terminal AMPPNP-Mdn1 region, together with the built model (green ribbons). The β-carbons are shown for selected side chains. (D) The densities corresponding to the nucleotides in AAA4 and AAA6 of AMPPNP-Mdn1 are shown (grey mesh), together with the modeled AMPPNP molecules (sticks). The β-carbons of the key lysine residues in the Walker A motifs in AAA4 of Mdn1 (K1199) and AAA1 of dynein (K1882) are shown. The alpha-carbons of residues in the N loops that are important for the interaction with the adenine ring of the nucleotides are also shown. For comparison, the box to the right shows ADP bound to AAA1 in *D. discoideum* dynein (PDB: 3VKG) (Kon et al., 2012). (E) Validation of the model for the N-terminal ring region of AMPPNP-Mdn1 (as also shown in Figure 4-1C) by cross-validation FSC curves: “work” (blue curve), refined model *versus* half map 1 used for refinement (work map); “free” (red curve), refined model *versus* half map 2 that was not used for refinement (free map); “combined” (black), refined model *versus* the combined final map. The similarity of the “work” and “free” curves suggests no substantial over-fitting. (F) Local-resolution map of the density of the C-terminal tail region of AMPPNP-Mdn1 calculated by ResMap in RELION 2.0.


A
N-terminal density
C-terminal density
Fourier shell correlation
1.0
0.5
0.0
FSC=0.143
5.9 Å
5.3 Å
4.0 Å
0.0
0.1
0.2
0.3
Spatial frequency (1/Å)
B
90°
3.5 4 4.5 5 5.5 Å
C
Y785
R1181
W1180
L1178
F1176
I1732
F1735
L1859
I1863
I1866
F2078
R2079
L2080
F2081
L2082
D
AMPPNP-Mdn1
AMPPNP in AAA4
K1199 (Walker A)
AMPPNP in AAA6
K1882 (Walker A)
Dynein (PDB: 3VKG)
ADP in AAA1
K1982 (Walker A)
N loop
N loop
N loop
E
combined
work
free
Fourier shell correlation
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.1
0.2
0.3
Spacial frequency (1/Å)
F
90°
4 4.5 5 5.5 6 Å

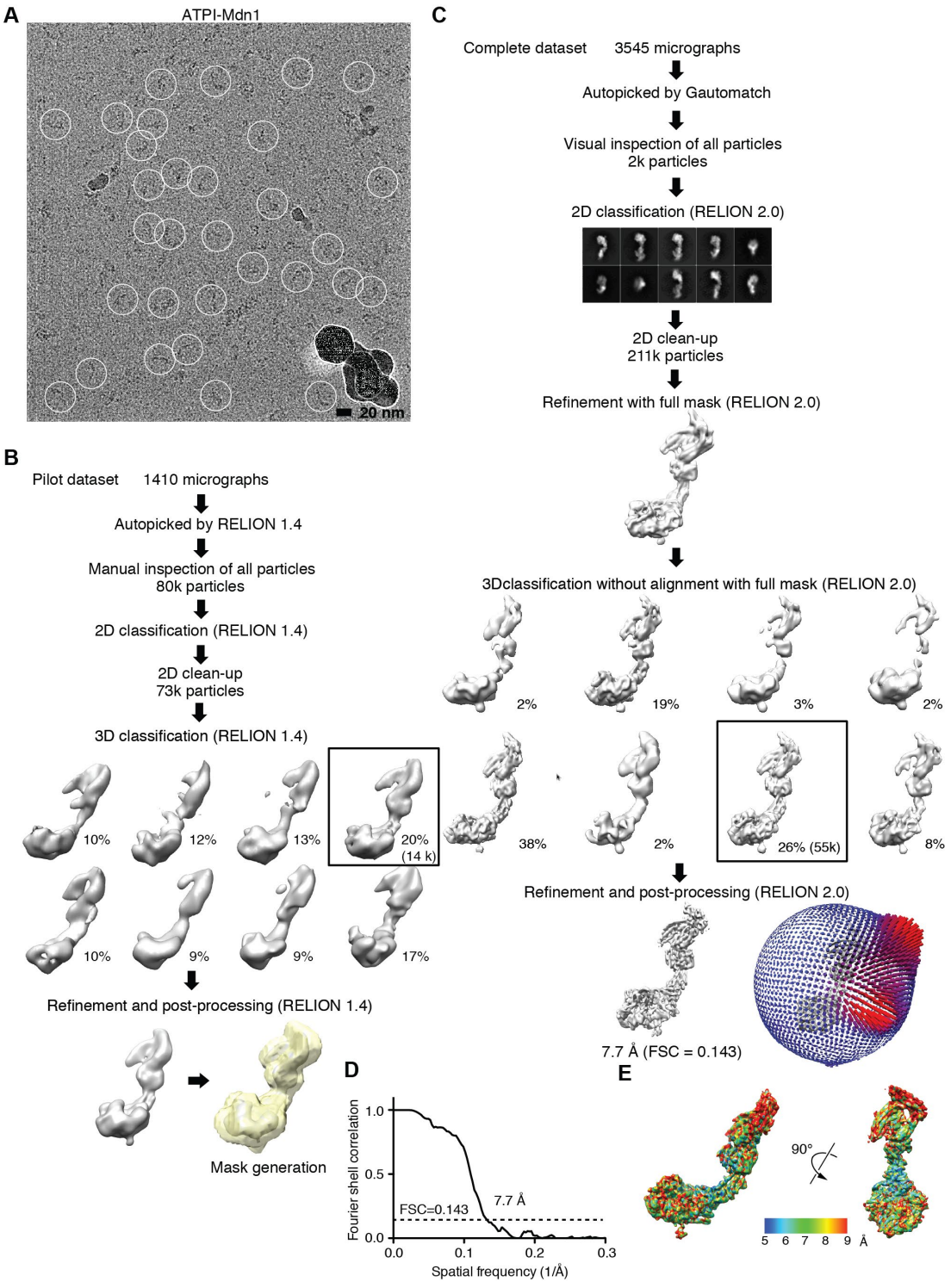
Appendix Figure 4-3. Cryo-EM analysis of Mdn1-ΔC in the presence of AMPPNP. (A) Schematic of the Mdn1-ΔC construct that lacks the C-terminal D/E-rich and MIDAS domains (aa 1-3911). The D/E-rich domain starts at aa 3912. (B) Size-exclusion chromatography profile for Mdn1-ΔC (elution volume: 11.5 mL). The elution volumes of the void peak (7.5 mL) and full-length Mdn1 (Mdn1-FL, 10.6 mL) are also indicated for comparison. (C) Peak fractions from the size-exclusion chromatography of Mdn1-ΔC and Mdn1-FL were analyzed by SDS-PAGE (Coomassie blue stain). (D) A cryo-EM image of AMPPNP-Mdn1-ΔC processed with MotionCor2. Some Mdn1-ΔC particles are marked (white circles). Contaminant particles are also marked (red circles). Scale bar: 20 nm. (E) Workflow used to process the dataset. Classes representing contaminant particles (red circles) were removed from the dataset. The angular distributions of the particles in the final reconstructions are shown next to the corresponding maps. Note: the resolution of the final reconstructions is further discussed in (F). (F) Gold-standard FSC curve calculated between independently refined half maps for the N- and C-terminal regions of AMPPNP-Mdn1-ΔC. For the map of the N-terminal density, the resolution was estimated as ~6.5 Å, more conservatively than the value suggested by the FSC = 0.143 criterion (~6.1 Å). (G) Local-resolution maps for the N- and C-terminal regions of Mdn1-ΔC calculated by ResMap in RELION 2.0. (H) Two views of the EM density map of Mdn1-ΔC (yellow surface). The Mdn1 model for full-length Mdn1 (ribbons, colored blue and purple for the AAA ring and hook-shaped tail regions, respectively) could be placed into the Mdn1-ΔC map without further refinement.


A
AMPPNP-Mdn1-ΔC
20 nm
B
2747 micrographs
Autopicked by Gautomatch
Visual inspection of all particles
218k particles
2D classification (RELION 2.0)
2D clean-up
176k particles
Refinement with full mask (RELION 2.0)
3D classification with ring mask
(RELION 2.0)
3D classification with tail mask
(RELION 2.0)
19%
15% (26 k)
24%
17%
18%
13% (23 k)
24%
19%
25%
27%
Refinement and post-processing
(RELION 2.0)
Refinement and post-processing
(RELION 2.0)
N-terminal density (the ring region)
6.1 Å (FSC = 0.143)
C-terminal density (the tail region)
7.4 Å (FSC = 0.143)
C
N-terminal density
C-terminal density
Fourier shell correlation
Spatial frequency (1/Å)
6.5 Å
7.4 Å
6.1 Å
D
90°
90°
5 6 7 8 9 Å

APPENDIX 5

EM processing of ATPI-Mdn1

Appendix Figure 5-1 Cryo-EM analysis of Mdn1 in the presence of ATP and Rbin-1. (A) A cryo-EM image of ATPI-Mdn1 processed with MotionCor2. Some Mdn1 particles are marked with white circles. Scale bar: 20 nm. (B) The workflow used to process the pilot dataset. (C) The workflow used to process the complete dataset. The angular distribution of the particles in the final reconstruction is shown next to the final map. (D) Gold-standard FSC curve calculated between independently refined half maps for ATPI-Mdn1. The resolution was estimated as ~7.7 Å (FSC = 0.143). (E) Local-resolution map calculated by ResMap in RELION 2.0.


A
ATPI-Mdn1
20 nm
B
Pilot dataset 1410 micrographs
Autopicked by RELION 1.4
Manual inspection of all particles
80k particles
2D classification (RELION 1.4)
2D clean-up
73k particles
3D classification (RELION 1.4)
10%
12%
13%
20%
(14 k)
10%
9%
9%
17%
Refinement and post-processing (RELION 1.4)
Mask generation
C
Complete dataset 3545 micrographs
Autopicked by Gautomatch
Visual inspection of all particles
2k particles
2D classification (RELION 2.0)
2D clean-up
211k particles
Refinement with full mask (RELION 2.0)
3Dclassification without alignment with full mask (RELION 2.0)
2%
19%
3%
2%
38%
2%
26% (55k)
8%
Refinement and post-processing (RELION 2.0)
7.7 Å (FSC = 0.143)
D
Fourier shell correlation
1.0
0.5
0.0
FSC=0.143
7.7 Å
0.0
0.1
0.2
0.3
Spatial frequency (1/Å)
E
90°
5 6 7 8 9 Å

BIBLIOGRAPHY

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D, Biological Crystallography *66*, 213-221.

Anderson, D.J., Le Moigne, R., Djakovic, S., Kumar, B., Rice, J., Wong, S., Wang, J., Yao, B., Valle, E., Kiss von Soly, S.*, et al.* (2015). Targeting the AAA ATPase p97 as an approach to treat cancer through disruption of protein homeostasis. Cancer Cell *28*, 653-665.

Aoi, Y., Sato, M., Sutani, T., Shirahige, K., Kapoor, T.M., and Kawashima, S.A. (2014). Dissecting the first and the second meiotic divisions using a marker-less drug-hypersensitive fission yeast. Cell Cycle *13*, 1327-1334.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., and Ben-Tal, N. (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Research *44*, W344-350.

Barrio-Garcia, C., Thoms, M., Flemming, D., Kater, L., Berninghausen, O., Bassler, J., Beckmann, R., and Hurt, E. (2016). Architecture of the Rix1-Rea1 checkpoint machinery during pre-60S-ribosome remodeling. Nature Structural & Molecular Biology *23*, 37-44.

Bassler, J., Kallas, M., Pertschy, B., Ulbrich, C., Thoms, M., and Hurt, E. (2010). The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. Molecular Cell *38*, 712-721.

Bassler, J., Paternoga, H., Holdermann, I., Thoms, M., Granneman, S., Barrio-Garcia, C., Nyarko, A., Lee, W., Stier, G., Clark, S.A.*, et al.* (2014). A network of assembly factors is involved in remodeling rRNA elements during preribosome maturation. The Journal of Cell Biology *207*, 481-498.

Bhabha, G., Cheng, H.C., Zhang, N., Moeller, A., Liao, M., Speir, J.A., Cheng, Y., and Vale, R.D. (2014). Allosteric communication in the dynein motor domain. Cell *159*, 857-868.

Bodnar, N., and Rapoport, T. (2017). Toward an understanding of the Cdc48/p97 ATPase. F1000Research *6*, 1318.

Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K., and Jones, D.T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Research *41*, W349-357.

Burger, K., Muhl, B., Harasim, T., Rohrmoser, M., Malamoussi, A., Orban, M., Kellner, M., Gruber-Eber, A., Kremmer, E., Holzel, M.*, et al.* (2010). Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. The Journal of Biological Chemistry *285*, 12416-12425.

Carter, A.P. (2013). Crystal clear insights into how the dynein motor moves. Journal of Cell Science *126*, 705-713.

Carter, A.P., Cho, C., Jin, L., and Vale, R.D. (2011). Crystal structure of the dynein motor domain. Science *331*, 1159-1165.

Cavanaugh, A.H., Hempel, W.M., Taylor, L.J., Rogalsky, V., Todorov, G., and Rothblum, L.I. (1995). Activity of RNA polymerase I transcription factor UBF blocked by Rb gene product. Nature *374*, 177-180.

Chaker-Margot, M., Barandun, J., Hunziker, M., and Klinge, S. (2017). Architecture of the yeast small subunit processome. Science *355*.

Chou, T.F., Brown, S.J., Minond, D., Nordin, B.E., Li, K., Jones, A.C., Chase, P., Porubsky, P.R., Stoltz, B.M., Schoenen, F.J.*, et al.* (2011). Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways. Proceedings of the National Academy of Sciences of the United States of America *108*, 4834-4839.

DeWitt, M.A., Cypranowska, C.A., Cleary, F.B., Belyy, V., and Yildiz, A. (2015). The AAA3 domain of cytoplasmic dynein acts as a switch to facilitate microtubule release. Nature Structural & Molecular Biology *22*, 73-80.

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

Erzberger, J.P., and Berger, J.M. (2006). Evolutionary relationships and structural mechanisms of AAA+ proteins. Annual Review of Biophysics and Biomolecular Structure *35*, 93-114.

Firestone, A.J., Weinger, J.S., Maldonado, M., Barlan, K., Langston, L.D., O'Donnell, M., Gelfand, V.I., Kapoor, T.M., and Chen, J.K. (2012). Small-molecule inhibitors of the AAA+ ATPase motor cytoplasmic dynein. Nature *484*, 125-129.

Freed, E.F., Bleichert, F., Dutca, L.M., and Baserga, S.J. (2010). When ribosomes go bad: diseases of ribosome biogenesis. Molecular BioSystems *6*, 481-493.

Funabiki, H., Kumada, K., and Yanagida, M. (1996). Fission yeast Cut1 and Cut2 are essential for sister chromatid separation, concentrate along the metaphase spindle and form large complexes. EMBO Journal *15*, 6617-6628.

Galani, K., Nissan, T.A., Petfalski, E., Tollervey, D., and Hurt, E. (2004). Rea1, a dynein-related nuclear AAA-ATPase, is involved in late rRNA processing and nuclear export of 60 S subunits. The Journal of Biological Chemistry *279*, 55411-55418.

Garbarino, J.E., and Gibbons, I.R. (2002). Expression and genomic analysis of midasin, a novel and highly conserved AAA protein distantly related to dynein. BMC Genomics *3*, 18.

Gasse, L., Flemming, D., and Hurt, E. (2015). Coordinated ribosomal ITS2 RNA processing by the Las1 complex integrating endonuclease, polynucleotide kinase, and exonuclease activities. Molecular Cell *60*, 808-815.
Gjidoda, A., and Henry, R.W. (2013). RNA polymerase III repression by the retinoblastoma tumor suppressor protein. Biochimica et Biophysica Acta *1829*, 385-392.

Glynn, S.E., Martin, A., Nager, A.R., Baker, T.A., and Sauer, R.T. (2009). Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. Cell *139*, 744-756.

Gomez-Roman, N., Felton-Edkins, Z.A., Kenneth, N.S., Goodfellow, S.J., Athineos, D., Zhang, J., Ramsbottom, B.A., Innes, F., Kantidakis, T., Kerr, E.R.*, et al.* (2006). Activation by c-Myc of transcription by RNA polymerases I, II and III. Biochemical Society Symposium, 141-154.

Grandi, P., Rybin, V., Bassler, J., Petfalski, E., Strauss, D., Marzioch, M., Schafer, T., Kuster, B., Tschochner, H., Tollervey, D.*, et al.* (2002). 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. Molecular Cell *10*, 105-115.

Grigorieff, N. (2016). Frealign: an exploratory tool for single-particle cryo-EM. Methods in Enzymology *579*, 191-226.

Hanson, P.I., and Whiteheart, S.W. (2005). AAA+ proteins: have engine, will work. Nature Reviews Molecular Cell Biology *6*, 519-529.

Hirano, T., Funahashi, S., Uemura, T., and Yanagida, M. (1986). Isolation and characterization of Schizosaccharomyces pombe cutmutants that block nuclear division but not cytokinesis. EMBO Journal *5*, 2973-2979.

Hirsch, H.A., Jawdekar, G.W., Lee, K.A., Gu, L., and Henry, R.W. (2004). Distinct mechanisms for repression of RNA polymerase III transcription by the retinoblastoma tumor suppressor protein. Molecular and Cellular Biology *24*, 5989-5999.

Hohn, M., Tang, G., Goodyear, G., Baldwin, P.R., Huang, Z., Penczek, P.A., Yang, C., Glaeser, R.M., Adams, P.D., and Ludtke, S.J. (2007). SPARX, a new environment for cryo-EM image processing. Journal of Structural Biology *157*, 47-55.

Hurt, E., Hannus, S., Schmelzl, B., Lau, D., Tollervey, D., and Simos, G. (1999). A novel in vivo assay reveals inhibition of ribosomal nuclear export in ran-cycle and nucleoporin mutants. The Journal of Cell Biology *144*, 389-401.

Ivachtchenko, A.V., Il'yin, A.P., Kobak, V.V., Zolotarev, D.A., Boksha, L.V., Trifilenkov, A.S., and Ugoleva, D.M. (2002). New scaffolds for combinatorial synthesis. 1. 5-sulfamoylisatins and their reactions with 1,2-diamines. Journal of Combinatorial Chemistry *4*, 419-428.

Kappel, L., Loibl, M., Zisser, G., Klein, I., Fruhmann, G., Gruber, C., Unterweger, S., Rechberger, G., Pertschy, B., and Bergler, H. (2012). Rlp24 activates the AAA-ATPase Drg1 to initiate cytoplasmic pre-60S maturation. The Journal of Cell Biology *199*, 771-782.

Kater, L., Thoms, M., Barrio-Garcia, C., Cheng, J., Ismail, S., Ahmed, Y.L., Bange, G., Kressler, D., Berninghausen, O., Sinning, I.*, et al.* (2017). Visualizing the Assembly Pathway of Nucleolar Pre-60S Ribosomes. Cell *171*, 1599-1610 e1514.

Kawashima, S.A., Chen, Z., Aoi, Y., Patgiri, A., Kobayashi, Y., Nurse, P., and Kapoor, T.M. (2016). Potent, Reversible, and Specific Chemical Inhibitors of Eukaryotic Ribosome Biogenesis. Cell *167*, 512-524 e514.

Kawashima, S.A., Takemoto, A., Nurse, P., and Kapoor, T.M. (2012). Analyzing fission yeast multidrug resistance mechanisms to develop a genetically tractable model system for chemical biology. Chemistry & Biology *19*, 893-901.

Kawashima, S.A., Takemoto, A., Nurse, P., and Kapoor, T.M. (2013). A chemical biology strategy to analyze rheostat-like protein kinase-dependent regulation. Chemistry & Biology *20*, 262-271.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols *10*, 845-858.

Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. eLife *5*.

Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S., and Ban, N. (2011). Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. Science *334*, 941-948.

Kon, T., Nishiura, M., Ohkura, R., Toyoshima, Y.Y., and Sutoh, K. (2004). Distinct functions of nucleotide-binding/hydrolysis sites in the four AAA modules of cytoplasmic dynein. Biochemistry *43*, 11266-11274.

Kon, T., Oyama, T., Shimo-Kon, R., Imamula, K., Shima, T., Sutoh, K., and Kurisu, G. (2012). The 2.8 Å crystal structure of the dynein motor domain. Nature *484*, 345-350.

Konikkat, S., and Woolford, J.L., Jr. (2017). Principles of 60S ribosomal subunit assembly emerging from recent studies in yeast. The Biochemical Journal *474*, 195-214.

Kornprobst, M., Turk, M., Kellner, N., Cheng, J., Flemming, D., Kos-Braun, I., Kos, M., Thoms, M., Berninghausen, O., Beckmann, R.*, et al.* (2016). Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome. Cell *166*, 380-393.

Kressler, D., Hurt, E., Bergler, H., and Bassler, J. (2012). The power of AAA-ATPases on the road of pre-60S ribosome maturation molecular machines that strip pre-ribosomal particles. Biochimica et Biophysica Acta *1823*, 92-100.

Kressler, D., Roser, D., Pertschy, B., and Hurt, E. (2008). The AAA ATPase Rix7 powers progression of ribosome biogenesis by stripping Nsa1 from pre-60S particles. The Journal of Cell Biology *181*, 935-944.

Lo, K.Y., Li, Z., Bussiere, C., Bresson, S., Marcotte, E.M., and Johnson, A.W. (2010). Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. Molecular Cell *39*, 196-208.

Loibl, M., Klein, I., Prattes, M., Schmidt, C., Kappel, L., Zisser, G., Gungl, A., Krieger, E., Pertschy, B., and Bergler, H. (2014). The drug diazaborine blocks ribosome biogenesis by inhibiting the AAA-ATPase Drg1. The Journal of Biological Chemistry *289*, 3913-3922.

Ma, C., Wu, S., Li, N., Chen, Y., Yan, K., Li, Z., Zheng, L., Lei, J., Woolford, J.L., Jr., and Gao, N. (2017). Structural snapshot of cytoplasmic pre-60S ribosomal particles bound by Nmd3, Lsg1, Tif6 and Reh1. Nature Structural & Molecular Biology *24*, 214-220.

Martin, A., Baker, T.A., and Sauer, R.T. (2005). Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. Nature *437*, 1115-1120.

Martin, A., Baker, T.A., and Sauer, R.T. (2008). Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. Molecular Cell *29*, 441-450.

Matsuo, Y., Granneman, S., Thoms, M., Manikas, R.G., Tollervey, D., and Hurt, E. (2014). Coupled GTPase and remodelling ATPase activities form a checkpoint for ribosome export. Nature *505*, 112-116.

Matyskiela, M.E., Lander, G.C., and Martin, A. (2013). Conformational switching of the 26S proteasome enables substrate degradation. Nature Structural & Molecular Biology *20*, 781-788.

Miller, O.L., Jr., and Beatty, B.R. (1969). Visualization of nucleolar genes. Science *164*, 955-957.

Monroe, N., Han, H., Shen, P.S., Sundquist, W.I., and Hill, C.P. (2017). Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. eLife *6*.

Montanaro, L., Trere, D., and Derenzini, M. (2008). Nucleolus, ribosomes, and cancer. The American Journal of Pathology *173*, 301-310.

Nissan, T.A., Bassler, J., Petfalski, E., Tollervey, D., and Hurt, E. (2002). 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. EMBO Journal *21*, 5539-5547.

Nissan, T.A., Galani, K., Maco, B., Tollervey, D., Aebi, U., and Hurt, E. (2004). A pre-ribosome with a tadpole-like structure functions in ATP-dependent maturation of 60S subunits. Molecular Cell *15*, 295-301.

Ohi, M., Li, Y., Cheng, Y., and Walz, T. (2004). Negative staining and image classification - powerful tools in modern electron microscopy. Biological Procedures online *6*, 23-34.

Pelletier, J., Thomas, G., and Volarevic, S. (2018). Ribosome biogenesis in cancer: new players and therapeutic avenues. Nature Reviews Cancer *18*, 51-63.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal of Computational Chemistry *25*, 1605-1612.

Puchades, C., Rampello, A.J., Shin, M., Giuliano, C.J., Wiseman, R.L., Glynn, S.E., and Lander, G.C. (2017). Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. Science *358*.

Reck-Peterson, S.L., Yildiz, A., Carter, A.P., Gennerich, A., Zhang, N., and Vale, R.D. (2006). Single-molecule analysis of dynein processivity and stepping behavior. Cell *126*, 335-348.

Ripstein, Z.A., Huang, R., Augustyniak, R., Kay, L.E., and Rubinstein, J.L. (2017). Structure of a AAA+ unfoldase in the process of unfolding substrate. eLife *6*.

Roemer, T., and Krysan, D.J. (2014). Antifungal drug development: challenges, unmet clinical needs, and new approaches. Cold Spring Harbor Perspectives in Medicine *4*.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: fast and accurate defocus estimation from electron micrographs. Journal of Structural Biology *192*, 216-221.

Ruggero, D. (2012). Revisiting the nucleolus: from marker to dynamic integrator of cancer signaling. Science Signaling *5*, pe38.

Sanghai, Z.A., Miller, L., Molloy, K.R., Barandun, J., Hunziker, M., Chaker-Margot, M., Wang, J., Chait, B.T., and Klinge, S. (2018). Modular assembly of the nucleolar pre-60S ribosomal subunit. Nature *556*, 126-129.

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. Journal of Structural Biology *180*, 519-530.
Schmidt, H., and Carter, A.P. (2016). Review: Structure and mechanism of the dynein motor ATPase. Biopolymers *105*, 557-567.

Schmidt, H., Zalyte, R., Urnavicius, L., and Carter, A.P. (2015). Structure of human cytoplasmic dynein-2 primed for its power stroke. Nature *518*, 435-438.
Shaikh, T.R., Gao, H., Baxter, W.T., Asturias, F.J., Boisset, N., Leith, A., and Frank, J. (2008). SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nature Protocols *3*, 1941-1974.

Soding, J. (2005). Protein homology detection by HMM-HMM comparison. Bioinformatics *21*, 951-960.

Song, G., Yang, Y., Liu, J.H., Casasnovas, J.M., Shimaoka, M., Springer, T.A., and Wang, J.H. (2005). An atomic resolution view of ICAM recognition in a complex between the binding domains of ICAM-3 and integrin alphaLbeta2. Proceedings of the National Academy of Sciences of the United States of America *102*, 3366-3371.

Song, Y., DiMaio, F., Wang, R.Y., Kim, D., Miles, C., Brunette, T., Thompson, J., and Baker, D. (2013). High-resolution comparative modeling with RosettaCM. Structure *21*, 1735-1742.

Spudich, J.A. (2001). The myosin swinging cross-bridge model. Nature Reviews Molecular Cell Biology *2*, 387-392.

Stokes, J.M., Davis, J.H., Mangat, C.S., Williamson, J.R., and Brown, E.D. (2014). Discovery of a small molecule that inhibits bacterial ribosome biogenesis. eLife *3*, e03574.

Sun, Q., Zhu, X., Qi, J., An, W., Lan, P., Tan, D., Chen, R., Wang, B., Zheng, S., Zhang, C*., et al.* (2017). Molecular architecture of the 90S small subunit pre-ribosome. eLife *6*.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. Journal of Structural Biology *157*, 38-46.

Thomson, E., Ferreira-Cerca, S., and Hurt, E. (2013). Eukaryotic ribosome biogenesis at a glance. Journal of Cell Science *126*, 4815-4821.

Tschochner, H., and Hurt, E. (2003). Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends in Cell Biology *13*, 255-263.

Ulbrich, C., Diepholz, M., Bassler, J., Kressler, D., Pertschy, B., Galani, K., Bottcher, B., and Hurt, E. (2009). Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. Cell *138*, 911-922.

Wacker, S.A., Houghtaling, B.R., Elemento, O., and Kapoor, T.M. (2012). Using transcriptome sequencing to identify mechanisms of drug action and resistance. Nature Chemical Biology *8*, 235-237.

Wang, R.Y., Kudryashev, M., Li, X., Egelman, E.H., Basler, M., Cheng, Y., Baker, D., and DiMaio, F. (2015). De novo protein structure determination from near-atomic-resolution cryo-EM maps. Nature Methods *12*, 335-338.

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

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.

Woolford, J.L., Jr., and Baserga, S.J. (2013). Ribosome biogenesis in the yeast Saccharomyces cerevisiae. Genetics *195*, 643-681.

Wu, S., Tutuncuoglu, B., Yan, K., Brown, H., Zhang, Y., Tan, D., Gamalinda, M., Yuan, Y., Li, Z., Jakovljevic, J.*, et al.* (2016). Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. Nature *534*, 133-137.

Xiong, J.P., Stehle, T., Zhang, R., Joachimiak, A., Frech, M., Goodman, S.L., and Arnaout, M.A. (2002). Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. Science *296*, 151-155.

Yang, B., Stjepanovic, G., Shen, Q., Martin, A., and Hurley, J.H. (2015). Vps4 disassembles an ESCRT-III filament by global unfolding and processive translocation. Nature Structural & Molecular Biology *22*, 492-498.

Yang, Y., Heffernan, R., Paliwal, K., Lyons, J., Dehzangi, A., Sharma, A., Wang, J., Sattar, A., and Zhou, Y. (2017). SPIDER2: A Package to predict secondary structure, accessible surface area, and main-chain torsional angles by deep neural networks. Methods in Molecular Biology *1484*, 55-63.

Yang, Z., Fang, J., Chittuluru, J., Asturias, F.J., and Penczek, P.A. (2012). Iterative stable alignment and clustering of 2D transmission electron microscope images. Structure *20*, 237-247.

Yusupova, G., and Yusupov, M. (2014). High-resolution structure of the eukaryotic 80S ribosome. Annual Review of Biochemistry *83*, 467-486.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nature Methods *14*, 331-332.