Stony Brook University

Establishment of an *in vitro* system for the biochemical characterization of an engineered human mitochondrial i-AAA protease

A Dissertation Presented

by

HUI SHI

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

December 2016

Stony Brook University

The Graduate School

HUI SHI

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Steven Glynn – Dissertation Advisor
Assistant Professor, Department of Biochemistry and Cell Biology

Miguel Garcia-Diaz – Chairperson of Defense
Associate Professor, Department of Pharmacological Sciences

Wali Karzai – Committee Member
Associate Professor, Department of Biochemistry and Cell Biology

Ed Luk – Committee Member
Assistant Professor, Department of Biochemistry and Cell Biology

Peter Chien – Outside Member
Associate Professor, Department of Biochemistry and Molecular Biology
University of Massachusetts Amherst

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Establishment of an *in vitro* system for the biochemical characterization of an engineered human mitochondrial i-AAA protease

by

HUI SHI

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

2016

YME1L (Yeast Mitochondrial Escape 1-like) is a mitochondrial inner membrane AAA+ protease. It carries a AAA+ ATPase module and a metallopeptidase domain belonging to the M41 family on the same polypeptide chain and forms a homohexamer as an active enzyme. Human YME1L plays important roles in mitochondrial quality control, maintaining mitochondrial homeostasis, and participating in mitochondrial biogenesis. YME1L degrades or processes substrates at the expense of ATP. Although cellular functions of YME1L were investigated and several substrates have been identified by a number of *in vivo* studies, detailed mechanisms of substrate recognition and degradation are still poorly understood due to a lack of an appropriate *in vitro* assay system. The transmembrane segment of YME1L poses a challenge of obtaining the protease via recombinant expression. I overcame this problem by fusing a 32-residue polypeptide that hexamerizes in solution to the N-terminus of the catalytic domains of YME1L. I showed that the resultant construct (hexYME1L) is soluble and active in hydrolyzing ATP and degrading protein substrates. Using this engineered hexYME1L system, I performed Michaelis–Menten kinetics characterization of ATP hydrolysis by the YME1L enzyme, which shows comparable kinetics parameters to the ones reported for other AAA+ proteases. By performing *in vitro* protein degradation assay using the reconstituted hexYME1L, I found that degradation of protein substrates by YME1L requires a recognizable signal called degron that can be located at the N-terminus, C-terminus, or internal of a substrate. Simply unfolding a substrate is not a sufficient condition of its degradation by YME1L. I also found that YME1L degrades protein substrates in a processive manner. In collaboration with Anthony Rampello in our lab, we found that YME1L is capable of degrading proteins with substantial stability. We also found that degron sequence

influences substrate degradation by YME1L. These studies show that the engineered YME1L hexamer is suitable for *in vitro* biochemical investigation, and suggest that this approach is applicable to other similar proteases within the same family.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

μl - microliter

μM - micromolar

AAA - ATPases associated with various cellular activities

ADP - Adenosine diphosphate

ATP - Adenosine triphosphate

BME - β-mercaptoethanol

CD - Circular dichroism

cDNA - Complementary DNA

CK - Creatine kinase

DNA - Deoxyribonucleic acid

DTT - Dithiothreitol

EDTA - Ethylenediaminetetraacetic acid

GFP - Green fluorescent protein

GSH - Glutathione

GST - Glutathione S-transferase

GTP - Guanosine-5'-triphosphate

GuHCl - Guanidinium chloride

IPTG - Isopropyl β-D-1-thiogalactopyranoside

kb - Kilobase

kbp - Kilobase pair

LB - Luria-Bertani medium

LDH - Lactate dehydrogenase

MBP - Maltose-Binding Protein

mDHFR - Murine dihydrofolate reductase

ml - Milliliter

mM - Millimolar

MTX - Methotrexate

NADH - Nicotinamide adenine dinucleotide

OD - Optical density

PCR - Polymerase chain reaction

PEP - Phosphoenolpyruvate

PK - Pyruvate kinase

PMSF - Phenylmethylsulfonyl fluoride

RFU - Relative fluorescence units

rRNA - Ribosomal ribonucleic acid

SDS-PAGE - Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SEC - Size-exclusion chromatography

TCEP - Tris(2-carboxyethyl)phosphine

TEV protease - Tobacco Etch Virus nuclear-inclusion-a endopeptidase

tRNA - Transfer ribonucleic acid

Acknowledgments

I would like to dedicate this thesis to my parents. I'm so grateful to them. They always listened to me and encouraged me. Their words made me feel confident again to battle with the difficulties that I faced. It would never be possible for me to accomplish my graduate study without their support.

I would like to thank my advisor, Dr. Steven Glynn. I feel so lucky to have the opportunity to work in his laboratory and get trained to do good science. His sharp mind and enthusiasm in science affected me so much, making me excited in doing research. He is always supportive and helpful. I benefit so much from him for my project. I'm grateful for his patience for answering whatever question I had. I cannot remember how many times I knocked on his door to ask questions about crystallography, especially structure determination. I definitely owe him a lot.

I would like to thank my committee members: Dr. Wali Karzai, Dr. Miguel Garcia-Diaz, Dr. Ed Luk and Dr. Peter Chien. Their constructive suggestions were so helpful for improving my research. I would like to thank Dr. Karzai for allowing me to join their lab group meeting. I benefit a lot from the discussion we had in the meeting and the thoughts we shared.

I would like to thank members in our lab, Bojian Ding and Anthony Rampello for their cooperation in the research. I would like to thank our previous lab members, Caitlin Pozmanter and Luke McGoldrick for their help. I would like to thank Dr. Steven Smith and members in his lab for sharing their equipment.

Also, I would like to thank all my friends who made my life more enjoyable. I'm particularly grateful to Neha Puri, for the detailed discussion in all aspects about the experiments and her help besides lab work.

Chapter 1 – Introduction

Mitochondria are essential organelles that perform constant dynamic fusion and fission (1-3), and are critical for the viability of cells and for cells to implement their functions normally, largely due to the ability to produce the energy currency, ATP. In addition, mitochondria play vital roles in other cellular activities, such as lipid metabolism (4), calcium signaling (5), autophagy (6), apoptosis (7), and biogenesis of iron-sulfur clusters (8-10). Since mitochondria are involved in a wide range of cellular activities, malfunction of mitochondria induces different deleterious outcomes, including cardiovascular disorders (11), type 1 and type 2 diabetes (11), neurodegenerative diseases (12), and cancer (13, 14). Indeed, several drug development strategies targeting mitochondrial functions have been suggested and applied (15, 16), one of which is to combat cancer by promoting apoptosis of tumor cells.

In order for these activities to be fulfilled and to avoid abnormalities, proteins located in all parts of mitochondria need to cooperate and to be controlled precisely under a surveillance system.

1.1 Mitochondrial structure, proteome and quality control

Mitochondrial structure and proteome

Mitochondria contain phospholipid bilayer double membranes and are organized in a compartmentalized architecture, which contains four major sections (Figure 1.1): the outer membrane (OM), the inner membrane (IM), the intermembrane space (IMS) enclosed by the OM and the IM, and the matrix (M), whose boundary is defined by the IM. A portion of the inner membrane termed cristae, invaginates into the matrix, thus increasing the surface area of the inner membrane, where a series of electron transport events take place that results in ATP generation, driven by the proton gradient built up during the process (17, 18). The separation of mitochondria into specific regions allows different biochemical reactions that require a specific environment to proceed under their optimal conditions without interference by other components in the organelle.

The outer membrane sets a border for mitochondria to be separated from the cytosolic environment. Interestingly, this membrane structure comprises a low ratio of protein and

cardiolipin (19, 20). Furthermore, porin proteins, which originated from Gram-negative bacteria, give rise to the permeability of the outer membrane, allowing the exchange of small hydrophilic molecules of up 6,000 Daltons between the intermembrane space of mitochondria and the cytosol (20-23). The outer membrane thus enables the communication of mitochondria with its surroundings. Mitochondrial porins are integral β-barrel proteins in the outer membrane and are also named as voltage-dependent anion channels (VDAC) (24-26). VDACs are regulated by metabolites and proteins to control the permeability of the outer membrane (24, 27).

The inner membrane of mitochondria is rich in protein and impermeable. As a result, channels in the inner membrane are needed to transport molecules such as ions between the matrix and the intermembrane space (28). Proteins including translocase of the inner membrane (TIM) complexes are present in the inner membrane and facilitate the translocation of nucleus-encoded proteins into the matrix (17, 28-30). Five complexes that constitute the respiratory chain responsible for ATP production are located at cristae, a characteristic structure of the inner membrane (17, 18). Electrons are transferred from the initial donor, NADH, to the final acceptor, oxygen molecules, through a series of redox reactions catalyzed by complex I - IV. During this process, protons are pumped from the matrix into the intermembrane space, inducing the formation of a proton gradient across the inner membrane. ATP molecules are then synthesized from ADP and inorganic phosphate by complex V to convert energy from the proton gradient into the chemical energy stored in ATP (18).

The intermembrane space and the matrix are aqueous environments with the function of metabolite storage and exchange as well as protein folding and degradation, and are connected to other cellular processes such as apoptosis (31). The matrix contains the mitochondrial genome that is transcribed by the mitochondrial RNA polymerase in the presence of other transcription factors (e.g. Transcription Factor A, Mitochondrial, TFAM), and translated on mitochondrial ribosomes. The mitochondrial genome is a compact and double-stranded DNA molecule of a size about 16.6 kbp. The whole genome encodes 13 proteins that are constituents of electron transport chain complex I, complex III, complex IV, and complex V, 22 mitochondrial tRNAs, and 2 rRNAs, with no intron sequences (32).

Although the mitochondrial genome encodes only 13 proteins, the proteome of mitochondria contains more than 1,000 members in mammals. Not surprisingly, a vast majority of mitochondrial proteome are nucleus-encoded, translated in the cytosol and imported into one of

the four compartments of mitochondria (33, 34). Experiments extensively characterizing the mitochondria proteome from human heart identified proteins involved in various cellular processes including oxidative phosphorylation (OXPHOS), cell signaling, ion channels, gene expression, metabolism, apoptosis, immune reaction, redox reactions, and protein degradation and translocation (33). A bioinformatics study further revealed the complexity of mitochondrial functions and the protein network (35).

In order for the proteins encoded by the nuclear genome to reach their correct locations in the mitochondria, mechanisms have evolved to ensure precise protein import into the mitochondria (36, 37). Proteins synthesized in the cytosol are imported through four separate pathways. Protein complexes such as translocase of the outer membrane (TOM) and translocase of the inner membrane (TIM) are essential components of these protein import pathways (36). Surprisingly, recent studies have suggested that certain intermembrane space proteins, when reduced and destabilized, can be exported back into the cytosol through Tom40 (38). Thus, a more extensive regulatory network of protein import and export allows mitochondria to maintain their homeostasis.

Mitochondrial quality control

In order to achieve proper function, mitochondria must coordinate the import of proteins from the cytosol and expression of the mitochondrial genome to achieve the correct stoichiometry of supramolecular complexes, for example, the respiratory chain. Unassembled components of the respiratory chain result in an imbalance in protein distribution and need to be cleared from the inner membrane. Mitochondria are constantly facing stress from the surrounding environment, but the major threats come from the OXPHOS process within mitochondria, during which ATP is generated and the harmful highly reactive byproducts, reactive oxygen species (ROS), are formed (39-41). Electrons can leak from the path of the electron transfer chain, mainly from complex I and complex III, and those electrons are captured by molecular oxygen to produce the superoxide anion (41), which is the primary form of mitochondrial ROS. Superoxide anion can further react by dismutation to give rise to hydrogen peroxide, another kind of mitochondrial ROS (42). Although ROS at physiological level is important for normal cell functions in signaling pathways such as cell growth and inflammation reaction (43, 44), over-produced ROS beyond the clearance capacity of mitochondria can be deleterious to cells, resulting in damage to mitochondria DNA, lipid molecules, and proteins, which have been linked to cancers (45). Obviously, due to the broad

range of cellular functions in which mitochondria are involved, any abnormality in mitochondria biogenesis and the failure to counteract internal challenges imposed by ROS will interfere with normal cellular processes and induce the onset of various diseases.

Fortunately, evolution has enabled a sophisticated quality control network to emerge so that mitochondrial homeostasis is under surveillance and catastrophic results are avoided. In general, the quality control systems in mitochondria can be divided into five levels (46), including ROS scavenging, repairing the damaged molecules, degradation of proteins that cannot be recovered, mitochondrial fission and fusion, and mitophagy. These quality control mechanisms are organized in a hierarchical yet integrated and interactive network (46).

Among these quality control systems, the most important one is probably the pool of mitochondrial proteases (47). Although the proteases generally function by degrading substrates, the modes of protein processing and the forms of final products resulted from protein degradation are distinct. A critical duty of the proteases is to remove mitochondrial proteins that are damaged to a level beyond the capability of the repair systems, which would otherwise lead to mitochondrial dysfunction. These proteases also play important roles in regulating mitochondrial gene expression, complex assembly, and dynamics (47, 48). There are 25 proteases identified to be located in human mitochondria, and most of the homologs of corresponding enzymes have been found in yeast (47). As summarized in the comprehensive reviews (47, 48), mitochondrial proteases can be grouped into several categories, depending on their locations and functions. Substrates of mitochondrial proteases are either processed into a shorter mature form or degraded destructively into small peptides. Proteases thus operate as processing enzymes or quality control enzymes with degradation activity. However, certain proteases, for example, the mitochondrial inner membrane AAA+ proteases (i-AAA and m-AAA), possess both of the two functions (49).

Mitochondrial processing peptidase (MPP), mitochondrial intermediate peptidase (MIP), and inner membrane peptidase (IMP) are involved in protein processing for maturation (48, 50). To remove damaged and unfolded proteins caused by ROS or respond to the stress from the inside or the outside of mitochondria, AAA+ proteases such as Lon, i-AAA, and m-AAA are utilized. Oligopeptidases in the intermembrane space (Mop112 and Prd1) digest the peptides resulted from proteolytic cleavage by proteases performing the quality control function to generate individual amino acids (48). Recently, experimental and bioinformatics studies showed the regulation of mitochondrial proteins by the ubiquitin-proteasome degradation pathway (51-53). Proteins were

found to be ubiquitinated, and components of the ubiquitin-proteasome system (UPS) were found to be localized to the mitochondria (52). Importantly, one yeast protein, Dma1p, was identified to be an E3 ligase of the UPS system in the matrix (53). Members of the mitochondrial proteome enclosed by the mitochondrial membrane, especially those residing in the matrix, are inaccessible by UPS in the cytosol. The identification of Dma1p as an E3 ligase thus expands our view of the regulation mechanisms of mitochondria, providing evidence of the existence of mitochondrial UPS.

The highly complicated but interactive organization of quality control supplies a vital surveillance system for mitochondria and allows cells to function properly (46). Within this system, proteases play central roles in integrating the functional pathways by either processing or degrading proteins related to specific cell activities. One specific family, named AAA+ protease, is of particular importance among the mitochondrial proteases and have prompted detailed study.

1.2 AAA+ ATPases

AAA+ proteases belong to a larger superfamily of AAA+ ATPase proteins. As the name AAA+ (ATPases Associated with diverse cellular Activities) implies, members of this superfamily hydrolyze ATP molecules and harvest the chemical energy stored in the high-energy phosphate-phosphate bonds of ATP. The energy is then converted to other forms as needed for a variety of cellular processes. Important roles that AAA+ ATPases play are exemplified by protein degradation (54), DNA replication (55), transcription (56), and peroxisome biogenesis (57). Although the core structure is highly conserved, significant variations in sequences and structures exit among the AAA+ ATPases.

1.2.1 Classification of AAA+ ATPases

Attempts have been made to classify the AAA+ ATPase superfamily, which contains a large number of protein members with diverse functions (58-62). Apart from the classical AAA ATPase family, other proteins such as helicases, proteases, transcriptional activators, and proteins taking part in initiation of DNA replication form separate families. All these protein families combined with the central AAA family to establish the AAA+ ATPase superfamily. Based on the emergence of specific elements, Lupas and Martin (59) built a phylogenetic classification system, within which the AAA+ ATPase are divided into seven families. Members of the AAA+

superfamily have a characteristic C-terminal α-helical subdomain that is used as a standard to claim the membership of this superfamily (62). By applying this criterion, Ammelburg and colleagues (62) suggested the removal of some proteins previously assigned as AAA+ proteins from the superfamily. A conserved motif called "second region of homology" (SRH) distinguishes the classical AAA ATPases from others in the superfamily (59).

1.2.2 Structure of the AAA+ module

The AAA+ module is a domain of about 250 amino acids, and forms a hexamer to efficiently hydrolyze ATP (Figure 1.2a). Generally, a single AAA+ module is composed of two subdomains, an N-terminal large subdomain of Rossmann fold comprising α-helices and β-strands, and a small domain at the C-terminus constructed from all α-helical elements (Figure 1.2b). The core of the N-terminal subdomain is a parallel β-sheet made up of five β-strands, with α-helices flanking both sides of the sheet. The parallel strands are connected by one or more helices, and are arranged in a 51432 configuration. Within this conserved module, several notable sequence motifs have been discovered and characterized (Figure 1.3) (63-65). Starting from the N-terminus of the module, there exist two Walker motifs, SRH, and sensor 2. The two Walker motifs, Walker A and Walker B are required for ATP binding and hydrolysis, respectively. The consensus sequence of the Walker A motif is denoted as GXXXXGK[T/S], within which the lysine residue is critical for binding of nucleotide. Mutation of this lysine residue results in the lack of ATP binding. The glutamic acid residue in the Walker B motif with a characteristic sequence of hhhhDE, where h stands for hydrophobic amino acid, is responsible for the catalysis of ATP hydrolysis. Replacing the glutamate with a glutamine severely reduces the ability of hydrolyzing ATP, yet has no effect on ATP binding. An additional consensus sequence termed SRH is characteristic of classical AAA proteins. Within SRH, a conserved polar residue, e.g. threonine or serine, is the sensor of the γ-phosphate of ATP (sensor 1) bound on the same subunit and mutation of this sensor 1 residue to an alanine or methionine down regulates the ATPase activity of AAA proteins (66). Moreover, so-called arginine fingers in the SRH are thought to mediate the communication between adjacent subunits, and contribute to the formation of nucleotide binding pocket of a neighboring subunit upon formation of a hexamer (63). A conserved arginine at the N-terminus of the small subdomain is called sensor 2, as it senses the γ-phosphate of ATP bound on its own subunit. The sensor 2

arginine also interacts with the N-terminal large subdomain of a neighbouring subunit. This conserved arginine at the sensor 2 position is thought to mediate the interaction between subunits, although it is not present in the classical AAA module (63).

In a hexameric structure, ATP binds within the cleft formed by the N-terminal large subdomain and the α-helical small subdomain, interacting with key residues present in the conserved elements. Through coordinated actions, conformational changes or domain movements induced by ATP hydrolysis are transmitted from one subunit to its neighbor, and the chemical energy generated is converted into a mechanical form for AAA+ ATPases to perform their functions (64).

1.3 AAA+ Proteases

One important family of the AAA+ ATPases is the AAA+ protease. At the cost of ATP consumption, AAA+ proteases not only degrade incorrectly folded or damaged proteins, but also remove proteins to initiate cell signaling pathways (54). Examples of AAA+ proteases have been found in all kingdoms. The serine peptidase ClpP has several AAA+ ATPase partners (ClpX, ClpA, and ClpC), with which ClpP combines to form a competent protease, respectively. Similarly, the heat shock protein HslU with ATPase activity and the compartmental protease HslV are found in bacteria where they assemble to form a functional AAA+ protease (67). The proteasome is present ubiquitously in eukaryotes and archaea, and some prokaryotic species such as *Mycobacterium tuberculosis* also possess a similar system (68-70). Members of the Lon protease family spread over prokaryotes and archaea, and are also found in membrane organelles inside the eukaryotic cells, for example, mitochondria and peroxisomes (71). The FtsH protease is the prototype of the FtsH family and is localized in the inner membrane of bacteria (72). Other two types of AAA+ protease belonging to the FtsH family were also discovered, namely i-AAA (Yme1 in yeast, YME1L in human) (73) and m-AAA (Yta10/12 in yeast, AFG3L2 and paraplegin in human) (74-76).

1.3.1 Structure of AAA+ proteases

AAA+ proteases can be grouped into two classes based on whether the AAA+ module and the peptidase domain reside on the same polypeptide chain or not (54) (Figure 1.4). The case of

two catalytic domains present on a single polypeptide has been found in the Lon and FtsH families (named group 1 hereafter). Whereas, in the case of Clp family, HslUV, and proteasome (named group 2 hereafter), the two functional modules are separated on two polypeptides. A common feature of the AAA+ proteases is the need of hexamerization in order to carry out ATP hydrolysis effectively (77). Although the peptidase domain in group 2 can be non-hexameric (the HslV peptidase is a dodecamer with two hexameric rings stacking together), the AAA+ module is a hexamer. In addition to the domains with catalytic activities, these proteases also carry an extra fragment called the N-domain at the N-terminus of the protein, with HslU as the only exception. As a unique feature, HslU contains an intermediate domain (I domain) between the two Walker motifs in the large subdomain of the AAA+ module. The inner membrane-anchored protease FtsH contains two transmembrane helices, whereas LonB in the Lon family contains one transmembrane span. Special situations are seen in the ClpA and ClpC AAA+ ATPases, where two AAA+ modules stack upon each other (Figure 1.4). The two AAA+ ATPase domains are usually named D1 and D2 module, respectively. ClpA interacts with ClpP via loops in the D2 domain (78). Although D1 and D2 can function independently and substrates with low stability can be processed efficiently by ClpAP variants with only D1 or D2 being active, both of the two AAA+ modules were shown to be required for the process of stable proteins (79). In some cases such as ClpXP and the 26S proteasome, a symmetry mismatch between the peptidases and their corresponding AAA+ regulatory partners is observed. While the ClpP peptidase is a tetradecamer made up of two heptameric rings, the core 20S proteolytic domain of the proteasome contains four homoheptamer rings stacking together that are comprised of two different types of proteins, resulting in a structure of 28 subunits. The hexameric ClpX/ClpA/ClpC ATPases stack upon ClpP to generate an active protease, respectively (54). In the case of proteasome, the 20S core peptidase has distinct ATPase regulators in different organisms. In eukaryotes, the ATPase hexamer is formed by six distinct gene products (Rpt1-6), each with ATP hydrolysis activity (69). The counterparts of Rpt1-6 particle are PAN in archaea (68, 69) and the Mpa ATPase in prokaryotes (70, 80), respectively.

1.3.2 Mechanism of protein degradation by AAA+ proteases

A major function of AAA+ proteases is to degrade proteins that are incorrectly translated, unfolded, and damaged as a result of various stress conditions. Figure 1.5 shows the general mechanism of protein degradation by a AAA+ protease (54). Usually, proteases form hexamers

that sequester the proteolytically active residues inside a hollow cylinder, where proteolysis takes place. Sitting on top of the peptidase is the hexameric AAA+ ATPase, with a central pore of limited size. A conserved loop with a consensus ΦVG motif (Φ represents the aromatic amino acid) between the well-studied Walker A and Walker B motifs projects into the pore, reducing the available space through which a polypeptide can pass in order to reach the cleavage site within the degradation chamber. This configuration ensures that peptidases digest their substrates in a AAA+ ATPase regulated way, and that the active residue responsible for catalyzing peptide bond cleavage does not interact with a substrate freely and cause nonspecific protein degradation.

Proteins to be degraded by AAA+ proteases usually carry a special sequence called a degron that can be recognized by the protease (54). Protein degradation is initiated upon engagement of the degron sequence by the protease. Cycling through the different stages of ATP hydrolysis induces conformational changes within the ATPase domain. Subsequently, displacement of the conserved loops in contact with degron in the central pore of AAA+ module transmits the information of structural variation to the substrate degron. The conformational changes allow the protease to exert a mechanical force on the substrate, leading to its unfolding and translocation into the proteolytic chamber, providing a linkage between ATP hydrolysis and protein degradation. This mechanism is supported by several biochemical and structural studies conducted on ClpXP (81, 82), FtsH (83-85), paraplegin (86), and 26S proteasome (87).

The requirement of degron enables specific substrates to be selected from a pool of proteins (54). Degrons can bear specific sequences, and attaching degrons to model proteins leads to the degradation of those proteins by proteases (88). The position of degrons is not limited within a substrate, and they are found to be located at the N-terminus, C-terminus or an internal segment of a protein (89). In addition to promoting the initial binding of substrates to proteases, degrons have been shown to regulate protease activities. Allosteric regulatory roles of degrons on Lon were proposed, which states that distinct Lon confirmations with different protein degradation rates and energy utilization efficiencies are stabilized by various degrons. Different operation modes of Lon caused by degron regulation thus enables Lon to accommodate the need of particular proteolysis activity under diverse conditions (90). Examples of degrons include ssrA (91, 92), sul20 (93) and β20 (94). Other recognition signals needed for protein degradation are also found in living organisms, for example, the simplest N-end rule signals (95), the ubiquitin/ubiquitin-like proteasome systems in prokaryotes (70, 80, 96), archaea (68), and eukaryotes (68).

1.4 Mitochondrial AAA+ proteases

Four AAA+ proteases have been found within mitochondria. Lon clears misfolded proteins caused by ROS damage in the matrix (97, 98). The two membrane-anchored proteases, i-AAA and m-AAA, function in the inner membrane, the matrix, and the intermembrane space (99). Compared to other well-studied AAA+ proteases, substrates of mitochondrial ClpXP are only recently described, including NOA1, a protein that is critical for synthesizing mitochondrial proteins (100). By partially unfolding the ALA synthase, which is responsible for producing the first precursor of heme, 5-aminolevulinic acid (ALA), mitochondrial ClpX stimulates ALA synthase by facilitating incorporation of the cofactor (101). This mechanism seems to be conserved across the lower and higher eukaryotes. Proteomic studies identified substrates of mitochondrial ClpXP that are involved in energy metabolism (102) and revealed the protective role of ClpXP against oxidative stress (103). These four AAA+ proteases are central components of maintaining mitochondrial homeostasis.

1.4.1 Mitochondrial AAA+ proteases in the inner membrane

i-AAA and m-AAA are the two types of mitochondrial membrane-anchored, compartmentalized AAA+ protease belonging to the FtsH family (49) (Figure 1.6). Both proteases contain an N-terminal domain, followed by one or two transmembrane segments and a AAA+ domain. At the C-terminus of the proteases, a peptidase domain belonging to the M41 metalloprotease family carries out the degradation of protein substrates. Both proteases contain features of the AAA+ module, as discussed above. Besides, a consensus HEXXH sequence within the peptidase domain characterizes the feature of zinc-binding motif in proteins of peptidase activity. The two histidine residues coordinate a zinc ion, with a remote C-terminal aspartate residue providing a third ligand. The divalent zinc ion activates a water molecule for the cleavage of peptide bonds. Similar to FtsH, the prototype of the family, i-AAA and m-AAA form hexamers, sequestering proteolytically active sites in the peptidase compartment. Protein unfolding and translocation depend on the ATPase domain, taking ATP as energy supply.

While both proteases share the same domain organization, there are some differences between the two. i-AAA is a homohexamer comprising YME1L subunits in human (104), and Yme1 subunits in yeast (105). However, different forms of m-AAA exist in different organisms.

A heterohexamer formed by Yta10 and Yta12 subunits constitutes the m-AAA in yeast (106). The corresponding m-AAA subunits are found in humans, where AFG3L2 (homolog of Yta12) forms a heterohexamer with paraplegin (homolog of Yta10). Interestingly, AFG3L2 itself is able to hexamerize, resulting in an active homohexamer (107). i-AAA harbors one transmembrane span, and its catalytic domains reside in the intermembrane space of mitochondria, with the N-terminal domain on the opposite side of the inner membrane. In contrast, m-AAA passes the inner membrane twice, projecting its N-domain and two active domains into the matrix. As a result, the i-AAA and m-AAA nomenclature comes from this opposite orientations of the ATPase and peptidase domains. One possible reason for nature to evolve two membrane-located AAA+ proteases with opposite configurations is to allow the degradation of proteins in different compartments of the mitochondria, including the matrix, the inner membrane, and the intermembrane space. This is supported by the evidences showing substrate degradation by the two proteases in all mitochondrial compartments (108-110). These two proteases play an important role in mitochondrial quality control. Dysfunction of the enzymes results in various diseases, including neurodegenerative diseases (e.g. hereditary spastic paraplegia (111)). Recent research showed the relationship of Yme1 and Barth syndrome, an X-linked disease (112, 113).

1.4.2 i-AAA

The i-AAA protease, Yme1, was discovered by a genetic screen, from which mutations of nucleus-encode genes resulting in increased mitochondrial DNA escape were isolated (114). And it was observed later that inhibiting Yme1 activity by either mutation or deletion results in a defect of yeast growth in non-fermentable media at high temperature (73). A human homolog, YME1L, was shown to be able to complement yeast Yme1 (104). Since then, a number of *in vivo* studies were conducted to explore cellular functions of YME1L/Yme1, focusing on the substrates of the enzyme and the cellular processes they are associated with.

Substrates

Unassembled subunit of the inner membrane integrated cytochrome c oxidase, Cox2, is cleared by Yme1 to avoid its accumulation in the inner membrane (115). A similar case has been observed for the TIM chaperones, Tim9/Tim10, in yeast (116, 117). An intermembrane space

protein, external NADH-ubiquinone oxidoreductase 1 (Nde1), when fused to a hemagglutinin epitope C-terminally which results in Nde1 destabilization but does not disturb its localization, is degraded in a Yme1-dependent manner (118). Taz1p mutants related to Barth syndrome were found to be degraded by Yme1 in the yeast model (112). In addition to its quality control function, Yme1 also participates in protein turnover at a physiological level, as seen in the case of phosphatidylserine decarboxylase 1 (Psd1) (119), Ups1/Ups2 (120) in yeast and PRELI (121) in human. Using the model organism *Saccharomyces cerevisiae*, researchers demonstrated that a mitochondrial outer membrane protein Atg23 needs to be processed by Yme1 at the C-terminus before it can interact with Atg11 to promote mitophagy (122). Impeding this processing by Yme1 results in the deficiency of mitophagy. Importantly, OPA1 (Mgm1 in yeast) required for mitochondrial fusion is processed by YME1L/Yme1 and OMA1 cooperatively, suggesting a role of YMEL1/Yme1 in regulating mitochondrial dynamics (123-128). Misregulation of OPA1 processing has been linked to the onset of heart disease (129). The role of YME1L in regulating mitochondrial protein import is evidenced by the degradation of TIM17A, a component of TIM23 complex, under stress conditions (130). Interestingly, OMA1 and YME1L show a reciprocal degradation pattern resulted from different challenges (131). Depolarization of the inner membrane of mitochondria results in OMA1 degradation by YME1L, whereas mitochondria depolarization and ATP depletion induces degradation of YME1L. An intriguing experiment showed the activity of Yme1 independent of its proteolytic function, in which polynucleotide phosphorylase (PNPase) is transported into the intermembrane space through translocation by Yme1 without being degraded (132).

N-domain and the transmembrane segment

The N-domains of AAA+ proteases in the FtsH family share no obvious sequence homology. By the bioinformatics analysis, N-domains (including all sequences of the proteases except for the AAA+ module and the peptidase domain) from different AAA+ metalloproteases were sorted into three main groups (133). Interestingly, YME1L/Yme1 and their homologs form two of the three main groups, evolutionally distant to FtsH and m-AAA. The periplasmic fragment of *E. coli* FtsH exists as a monomer in solution. However, a hexameric model can be generated by symmetry operations using the crystal structure of this fragment (133). Surprisingly, although N-domains are much less conserved in sequences compared to the catalytic domains of AAA+

metalloproteases, the structures of periplasmic fragment from different members in the FtsH family display high similarity. In addition, residues at the subunit interfaces in the hexameric model of periplasmic fragment of *E. coli* FtsH show conservation as compared to the hexamer of corresponding segment of m-AAA subunit (133, 134). This may indicate the role in oligomerization of this periplasmic fragment. In contrast, experiment results from *E. coli* FtsH showed that the second transmembrane segment is responsible for oligomerization. FtsH truncation without the first transmembrane segment and the periplasmic fragment still maintains the ATPase activity and is capable of degrading substrates, comparable to that of wild type enzymes (135). Nonetheless, it is possible that the periplasmic fragment contributes additionally to the stability of FtsH hexamer. It is also possible that this fragment plays a role in recruitment of physiological substrates of FtsH that have recognition sequences in the periplasm. The roles of N-domain in oligomerization and substrate recognition have also been suggested in other AAA+ proteases, such as ClpXP (136, 137) and ClpAP (138). Based on the conserved functions of FtsH family members and the structural conservation characteristics of periplasmic fragment, it is reasonable to speculate that the N-domain the transmembrane segment of YME1L may have similar functions in oligomerization and substrate engagement.

Substrate recognition and degradation

As discussed above, a substrate is generally recognized by a AAA+ protease through a specific sequence called degron. In prokaryotes, degrons are usually stretches of short peptide fragment of 10 – 20 residues. Well studied degrons of this type include ssrA (91, 92), sul20 (93) and β20 (94). Degrons can be recognized in different ways by the AAA+ proteases, through the pore loop, the N-domain, or corresponding adaptors (54). ClpXP recognizes the C-terminus of ssrA, and substitution of the last two alanine residues in the ssrA tag by aspartate ($ssrA^{DD}$) prevents substrates with $ssrA^{DD}$ attached from being degraded by the enzyme (91). Similar to the ssrA tag, the C-terminal fragment of LpxC about 20-amino acid long is important for the degradation of LpxC by FtsH (139). This fragment contains a conserved motif of non-polar residues and is highly specific for substrate selection by FtsH. Substitution of the last four amino acids at the C-terminus results in LpxC degradation by other proteases rather than FtsH. In the case of Lon protease, a motif containing hydrophobic residues (FPLF) in UmuD is critical for substrate degradation (140). A similar motif (WRFAWFP) is also found in the β20 degron to be important for substrate

degradation by Lon (94). Similarly, a degron sequence was identified in YfgM, a transmembrane protein degraded by FtsH under stress condition (141). Interestingly, other proteins carrying an N-terminal YfgM-like degron were shown to be degraded by FtsH (142). This YfgM-like motif may represent a type of degron that mediates the degradation of certain substrates by protease members in the FtsH family. Several experiment results suggested some criteria of substrate selection by the i-AAA protease. Yme1 seems to recognize the unfolded mDHFR (143), which can be achieved by sensing the folding state as in the case of Lon (94), or by engagement of the recognition sequence in mDHFR that becomes accessible when mDHFR is unfolded. The latter mode of substrate selection by revealing recognition signal when proteins do not exist in their native folding/assembly states may explain the degradation of unassembled Cox 2 by Yme1 (115), and the degradation of Ups1/Ups2 by Yme1 without formation of complexes with Mdm35, respectively (120). The human homolog of Ups1/Ups2, PRELI, was also reported to be degraded by human YME1L, which complexes with TRIAP1, a human homolog of Mdm35 (121). In addition, a minimum length requirement of about 20 amino acid residues was reported for substrates to be degraded by the i-AAA and m-AAA proteases (109). However, no defined degron sequences have been reported for substrate selection by YME1L. Considering the existence of versatile substrates of YME1L, it is possible that YME1L recognizes its substrates through specific sequences.

Additionally, adaptor proteins may play roles in substrate degradation by the YME1L protease. Examples of prokaryotic adaptors have been reported, such as SspB for ClpXP (144, 145), ClpS for ClpAP (146), and MecA for ClpCP(147). SspB and ClpS interact with their corresponding protease and degrons on the substrate, increasing the effective concentrations of substrates on the enzyme. MecA is important for the oligomerization of ClpC to form an active enzyme. Previous studies suggested that two inner membrane proteins, Mgr1p and Mgr3p, may serve as Yme1 adaptors (148, 149). Mgr1p or Mgr3p alone can bind unfolded DHFR, and both proteins are required for association of unfolded DHFR with Yme1 at a maximal extent (149).

FtsH seems to possess only weak unfoldase activity (150), which can also be inferred from the experimental results of Yme1 (109, 143). FtsH only degrades proteins with low stability, but is incapable of degrading stable proteins such as GFP (150). In consistent, Yme1 degrades destabilized mDHFR rather than wild type mDHFR (109, 143).

Although the cryo-electron microscopy structure of m-AAA has been reported (151), as well as the crystal structure of the AAA+ module of paraplegin (86), crystal structures of the hexameric FtsH protease containing both the ATPase domain and the peptidase domain (83-85) at higher resolution provide more valuable information about models of substrate degradation by enzymes in this family. Structural comparison between the apo-form of FtsH (83) and the one complexed with ADP in all six subunits (84) indicates that domain movements occur in the AAA+ module, induced by different nucleotide-binding states. In the ADP-bound structure, AAA+ modules are rearranged to be more compact, compared to that in the apo structure. Moreover, four phenylalanine residues in the pore loop of the ADP-bound FtsH hexamer move inward, representing the behavior of substrate pulling by the pore loop. Alternatively, the structure of FtsH from *Thermus thermophilus* suggests a different operation mode (85). The FtsH hexamer adopts a configuration of alternate closed and open conformations. Adjacent subunits function as an active unit, cycling through the steps of ATP hydrolysis. The proteolytically active site is inaccessible in the closed-form subunit, and the phenylalanine residue in the pore loop of the closed-form subunit directs the substrate peptide to the active site in the peptidase domain of the adjacent open-form subunit for degradation. A beta-hairpin structure in the peptidase domain in the open-form subunit precludes the direct access of catalytic site of protein degradation without the guidance of pore loop in the closed-form subunit. Unfortunately, no structural data of i-AAA are available till now. It would be interesting and important to determine the structures of i-AAA in different nucleotide-binding states. Such structural information will shed light on mechanisms of substrate degradation. In addition, these structural data will help to map substrate binding sites on i-AAA more precisely, such as the surface exposed helices on the AAA+ module (NH) and the peptidase domain (CH) that are important for substrate binding (152).

Despite the extensive characterization on substrates and cellular functions of YME1L/Yme1, detailed mechanisms of how this enzyme can select diverse protein substrates seemly unrelated are still poorly understood. This is largely due to the unavailability of an *in vitro* system, which allows us to dissect biochemical reaction steps without the interference by other mitochondrial proteases. Also, it is of importance to investigate whether YME1L/Yme1 is capable of applying significant forces to unfold stable proteins, considering that forces of certain strength are required when degrading membrane integrated proteins by YME1L/Yme1 (115), which requires prior dislocation of substrate proteins from the inner membrane.

In this thesis, in order to establish an *in vitro* system for studying YME1L biochemically, I reconstructed YME1L by fusing a 32-residue polypeptide to the N-terminus of the catalytic domains for inducing hexamerization, circumventing the solubility issue imposed by the transmembrane segment of YME1L when expressing the protein using a heterologous system. The resultant construct (hexYME1L) is active in solution with the capabilities of hydrolyzing ATP and degrading protein substrates. Michaelis–Menten kinetics measurement of hexYME1L shows that YME1L hydrolyzes ATP with comparable kinetics parameters as previously studied AAA+ proteases. I found that, in contrast to earlier findings, YME1L is able to degrade stable protein substrates, and that a degron sequence is necessary for mediating degradation. Degron position is not restricted, it can localized to the N-terminus, C-terminus or internal of the substrate to be degraded by the YME1L protease. I also found that degradation of proteins by YME1L proceeds in a processively way. In collaboration with Anthony Rampello in our lab, the model degron β-20 was further characterized in detail to show the preference of amino acid identities by YME1L. Degron length also contributes to protein degradation kinetics. Finally, I solved the crystal structure of the ATPase domain of i-AAA from *Myceliophthora thermophila*. The structure shows common features observed in other AAA+ modules. In sum, this engineering strategy is suitable for biochemical study of YME1L *in vitro*, allowing us to investigate individual steps in proteins degradation. More importantly, application of this approach to other membrane AAA+ protease could be appreciated.

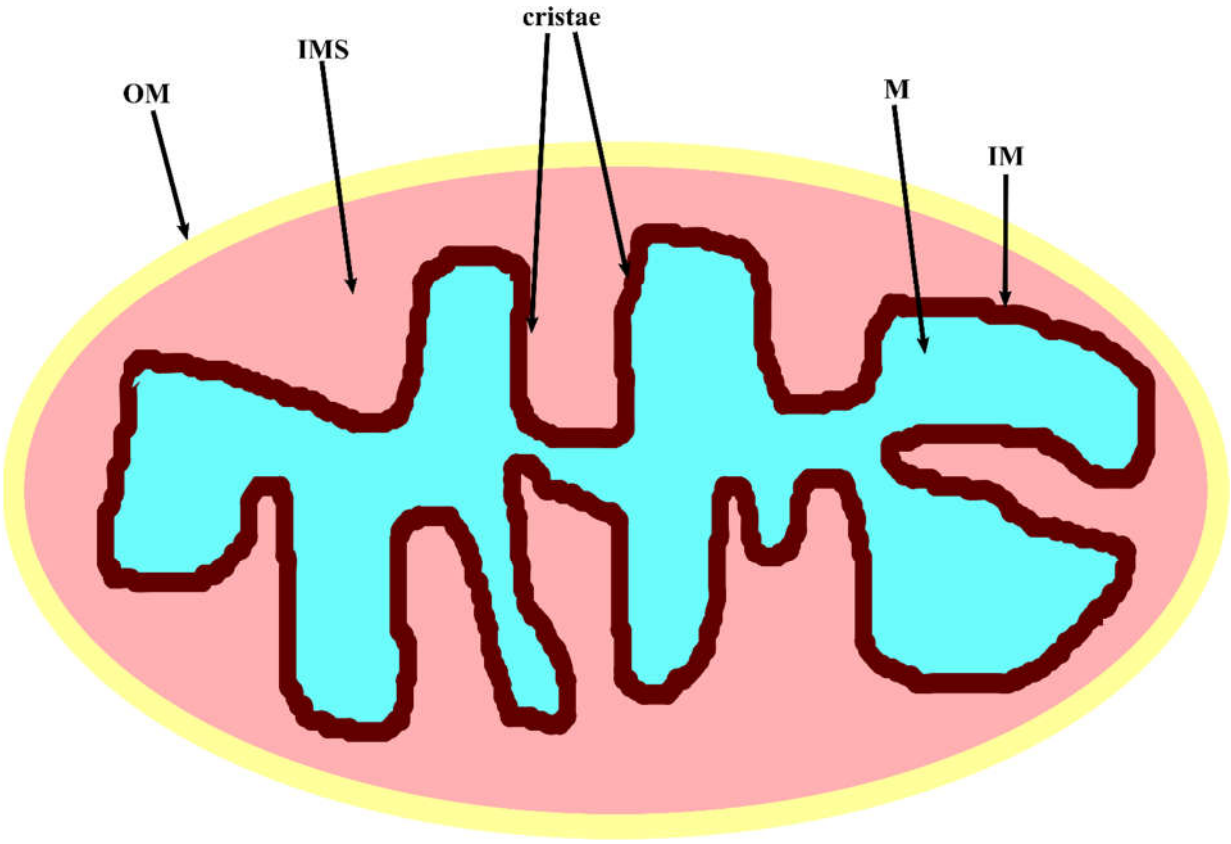


Figure 1.1 The structure of mitochondria.

A mitochondrion has a double membrane-enclosed structure. The outer membrane (OM) separates the organelle from the cytosol. The inner membrane (IM) circles the innermost aqueous environment, the matrix (M). Between the outer membrane and the inner membrane is an aqueous environment called the intermembrane space (IMS). A structure called cristae are formed from invagination of the inner membrane into the matrix.

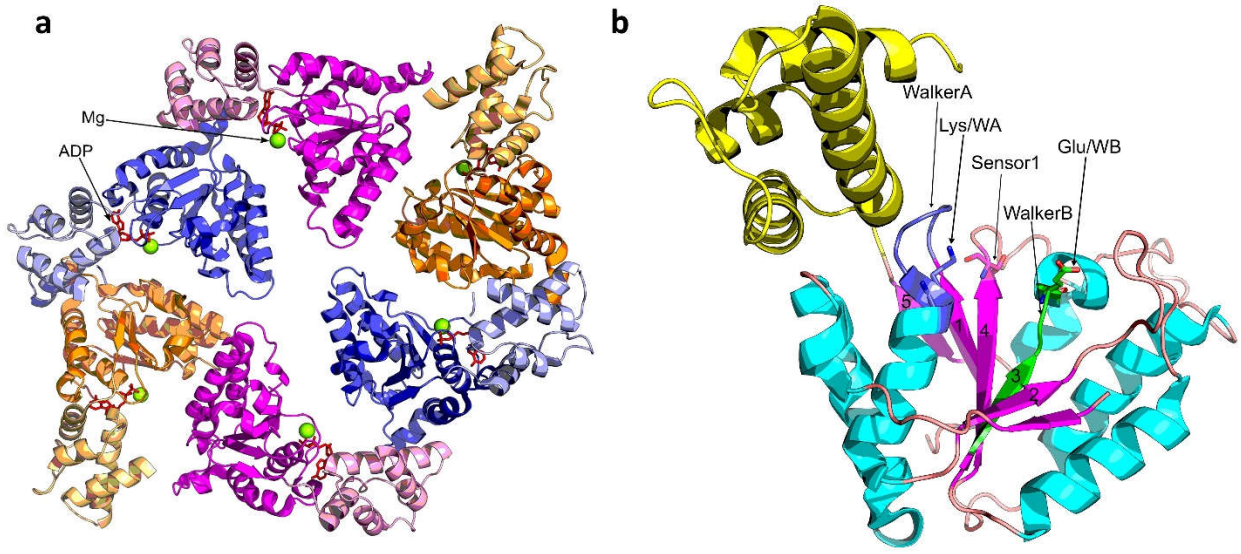


Figure 1.2 Structure of the AAA+ module.

a. Hexameric AAA+ module of FtsH from *Thermotoga maritima* (PDB code: 2CE7, residues: 151-402). Each of the six subunits is represented in one color, with two subunits at opposite positions shown in the same color. Large subdomains are colored in dark and the corresponding small subdomains are colored in the light version, respectively. ATP molecules are shown as red sticks and magnesium ions are represented as green spheres, respectively. **b.** Structure of AAA+ monomer of FtsH from *Escherichia coli* (PDB code: 1LV7). Large subdomain is comprised of α-helices (cyan) and β-sheet (pink). The five β-strands constituting the core β-sheet are organized in an order of 51432. The small subdomain is composed of α-helices colored in yellow. Conserved Walker A motif is colored in blue and Walker B motif is depicted in green color, respectively. Key residues are shown as stick representation, including lysine in the Walker A motif (Lys/WA), glutamate in the Walker B motif (Glu/WB), and sensor1, respectively.

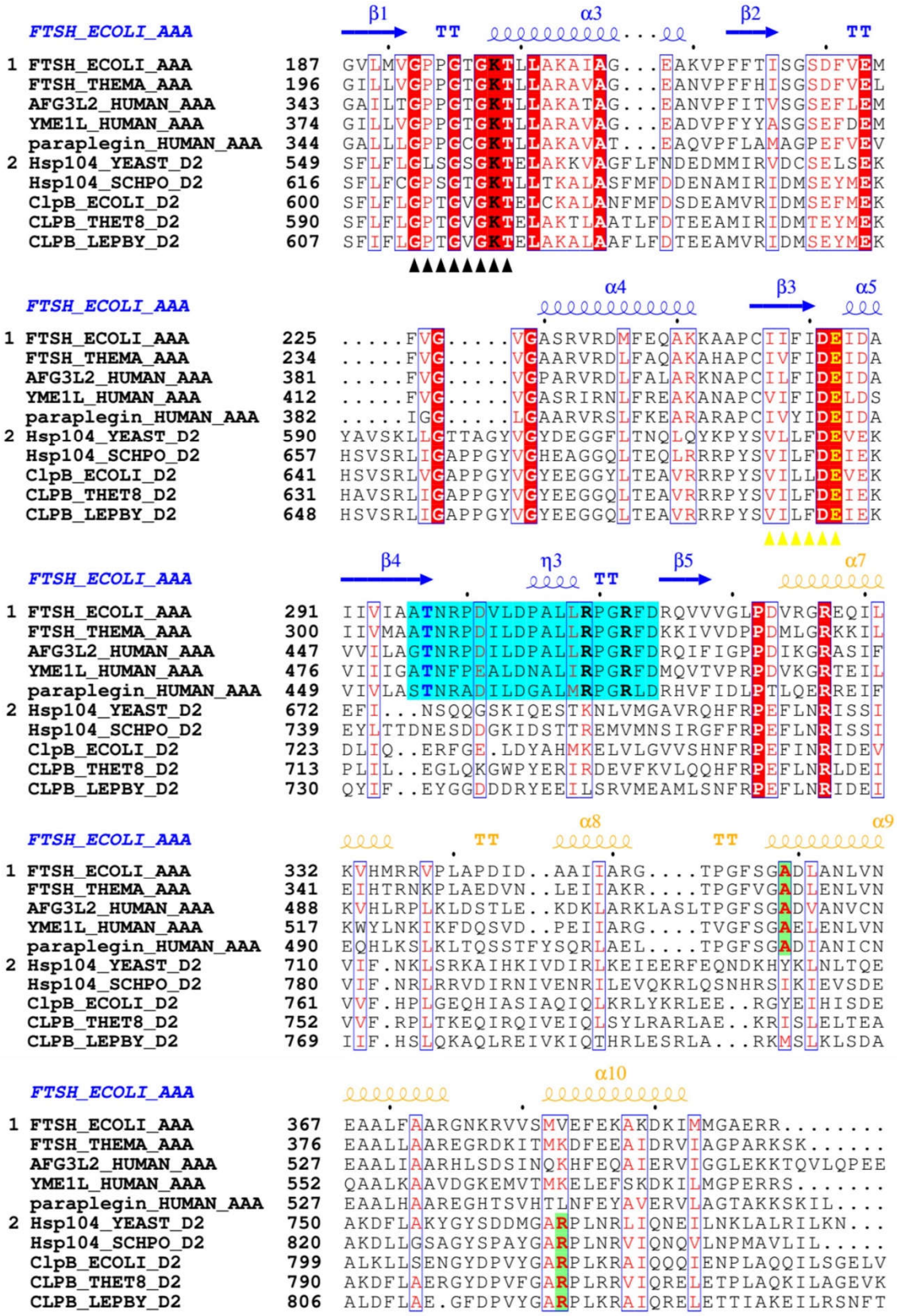

β1 TT α3 β2 TT
FTSH_ECOLI_AAA
1 FTSH_ECOLI_AAA 187 GVLMVGPPGTGKTLLAKAIAG...EAKVPFFTISGSDFVEM
FTSH_THEMA_AAA 196 GILLVGPPGTGKTLLARAVAG...EANVPFFHISGSDFVEL
AFG3L2_HUMAN_AAA 343 GAILTGPPGTGKTLLAKATAG...EANVPFITVSGSEFLEM
YME1L_HUMAN_AAA 374 GILLVGPPGTGKTLLARAVAG...EADVPFYYASGSEFDEM
paraplegin_HUMAN_AAA 344 GALLLGPPGCGKTLLAKAVAT...EAQVPFLAMAGPEFVEV
2 Hsp104_YEAST_D2 549 SFLFLGLSGSGKTELAKKVAGFLFNDEDMMIRVDCSELSEK
Hsp104_SCHPO_D2 616 SFLFCGPSGTGKTLLTKALASFMFDDENAMIRIDMSEYMEK
ClpB_ECOLI_D2 600 SFLFLGPTGVGKTELCKALANFMFDSDEAMVRIDMSEFMEK
CLPB_THET8_D2 590 SFLFLGPTGVGKTELAKTLAATLFDTEEAMIRIDMTEYMEK
CLPB_LEPBY_D2 607 SFIFLGPTGVGKTELAKALAAFLFDTEEAMVRIDMSEYMEK
α4 β3 α5
FTSH_ECOLI_AAA
1 FTSH_ECOLI_AAA 225 .....FVG.....VGASRVRDMFEQAKKAAPCIIFIDEIDA
FTSH_THEMA_AAA 234 .....FVG.....VGAARVRDLFAQAKAHAPCIVFIDEIDA
AFG3L2_HUMAN_AAA 381 .....FVG.....VGPARVRDLFALARKNAPCILFIDEIDA
YME1L_HUMAN_AAA 412 .....FVG.....VGASRIRNLFREAKANAPCVIFIDELDS
paraplegin_HUMAN_AAA 382 .....IGG.....LGAARVRSLFKEARARAPCIVYIDEIDA
2 Hsp104_YEAST_D2 590 YAVSKLLGTTAGYVGYDEGGFLTNQLQYKPYSVLLFDEVEK
Hsp104_SCHPO_D2 657 HSVSRLIGAPPGYVGHEAGGQLTEQLRRRPYSVILFDEIEK
ClpB_ECOLI_D2 641 HSVSRLVGAPPGYVGYEEGGYLTEAVRRRPYSVILLDEVEK
CLPB_THET8_D2 631 HAVSRLIGAPPGYVGYEEGGQLTEAVRRRPYSVILFDEIEK
CLPB_LEPBY_D2 648 HSVSRLIGAPPGYVGYEEGGQLTEAVRRRPYSVILFDEIEK
β4 η3 TT β5 α7
FTSH_ECOLI_AAA
1 FTSH_ECOLI_AAA 291 IIVIAATNRPDVLDPALLRPGRFDRQVVVGLPDVRGREQIL
FTSH_THEMA_AAA 300 IIVMAATNRPDILDPALLRPGRFDKKIVVDPPDMLGRKKIL
AFG3L2_HUMAN_AAA 447 VVILAGTNRPDILDPALLRPGRFDRQIFIGPPDIKGRASIF
YME1L_HUMAN_AAA 476 VIIIGATNFPEALDNALIRPGRFDMQVTVPRPDVKGRTEIL
paraplegin_HUMAN_AAA 449 VIVLASTNRADILDGALMRPGRLDRHVFIDLPTLQERREIF
2 Hsp104_YEAST_D2 672 EFI...NSQQGSKIQESTKNLVMGAVRQHFRPEFLNRISSI
Hsp104_SCHPO_D2 739 EYLTTDNESDDGKIDSTTREMVMNSIRGFFRPEFLNRISSI
ClpB_ECOLI_D2 723 DLIQ..ERFGE.LDYAHMKELVLGVVSHNFRPEFINRIDEV
CLPB_THET8_D2 713 PLIL..EGLQKGWPYERIRDEVFKVLQQHFRPEFLNRLDEI
CLPB_LEPBY_D2 730 QYIF..EYGGDDDRYEEILSRVMEAMLSNFRPEFLNRIDEI
α8 TT TT α9
FTSH_ECOLI_AAA
1 FTSH_ECOLI_AAA 332 KVHMRRVPLAPDID..AAIIARG....TPGFSGADLANLVN
FTSH_THEMA_AAA 341 EIHTRNKPLAEDVN..LEIIAKR....TPGFVGADLENLVN
AFG3L2_HUMAN_AAA 488 KVHLRPLKLDSTLE..KDKLARKLASLTPGFSGADVANVCN
YME1L_HUMAN_AAA 517 KWYLNKIKFDQSVD..PEIIARG....TVGFSGAELENLVN
paraplegin_HUMAN_AAA 490 EQHLKSLKLTQSSTFYSQRLAEL....TPGFSGADIANICN
2 Hsp104_YEAST_D2 710 VIF.NKLSRKAIHKIVDIRLKEIEERFEQNDKHYKLNLTQE
Hsp104_SCHPO_D2 780 VIF.NRLRRVDIRNIVENRILEVQKRLQSNHRSIKIEVSDE
ClpB_ECOLI_D2 761 VVF.HPLGEQHIASIAQIQLKRLYKRLEE..RGYEIHISDE
CLPB_THET8_D2 752 VVF.RPLTKEQIRQIVEIQLSYLRARLAE..KRISLELTEA
CLPB_LEPBY_D2 769 IIF.HSLQKAQLREIVKIQTHRLESRLA...RKMSLKLSDA
α10
FTSH_ECOLI_AAA
1 FTSH_ECOLI_AAA 367 EAALFAARGNKRVVSMVEFEKAKDKIMMGAERR.....
FTSH_THEMA_AAA 376 EAALLAAREGRDKITMKDFEEAIDRVIAGPARKSK.....
AFG3L2_HUMAN_AAA 527 EAALIAARHLSDSINQKHFEQAIERVIGGLEKKTQVLQPEE
YME1L_HUMAN_AAA 552 QAALKAAVDGKEMVTMKELEFSKDKILMGPERRS.....
paraplegin_HUMAN_AAA 527 EAALHAAREGHTSVHTLNFEYAVERVLAGTAKKSKIL....
2 Hsp104_YEAST_D2 750 AKDFLAKYGYSDDMGARPLNRLIQNEILNKLALRILKN...
Hsp104_SCHPO_D2 820 AKDLLGSAGYSPAYGARPLNRVIQNQVLNPMAVLIL.....
ClpB_ECOLI_D2 799 ALKLLSENGYDPVYGARPLKRAIQQQIENPLAQQILSGELV
CLPB_THET8_D2 790 AKDFLAERGYDPVFGARPLRRVIQRELETPLAQKILAGEVK
CLPB_LEPBY_D2 806 ALDFLAE.GFDPVYGARPLKRAIQRELETTIAKEILRSNFT

Figure 1.3 Multiple sequence alignment of AAA+ modules.

Multiple sequence alignment shows the conserved motifs and residues within the AAA+ module. Ten sequences are grouped into two classes, with sequences 1-5 being group 1 and sequences 6-10 being group 2, respectively. On top of the aligned sequences, secondary structures contained in the PDB file of the AAA+ module of *E. coli* FtsH (PDB code: 1LV7) are indicated. Secondary structures comprising the large subdomain are colored in blue and those constituting the small subdomain are colored in yellow, respectively. Black triangles under the aligned sequences indicate the conserved Walker A motif, while the yellow ones imply Walker B motif, respectively. The lysine residues in the Walker A motif are bolded in black, and the glutamate residues in the Walker B motif are bolded in yellow, respectively. Residues contained in the second region of homology (SRH) of the classical AAA+ module (group 1) are shown in the cyan box, with arginine fingers bolded in black and sensor 1 bolded in blue, respectively. The sensor 2 arginine residues in group 2 are indicated in Green-fluo box and bolded in red, while at the corresponding position, alanine residues are present (bolded in red in a Green-fluo box) in the classical AAA+ module (group 1). Sequence alignment was implemented in the MultAlin webserver (153), and the figure was prepared with ESPript 3.0 webserver (154).

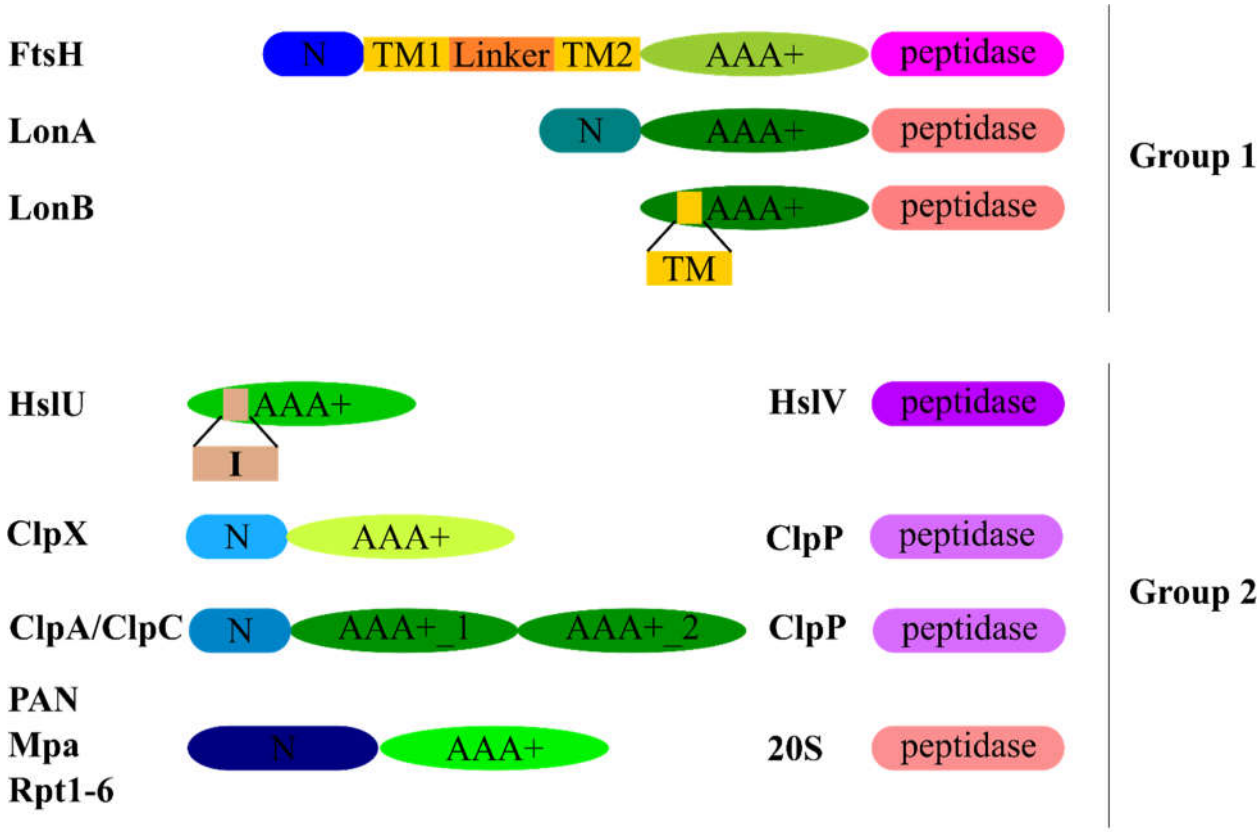


Figure 1.4 Domain organization of AAA+ proteases.

The AAA+ proteases are grouped into two categories. Group 1 contains FtsH, LonA, and LonB, having the AAA+ ATPase domain and peptidase domain exist on the same peptide chain, respectively. The AAA+ ATPase domain and peptidase domain are present on different peptide chains in the AAA+ proteases belonging to group 2. Group 2 contains HslUV, ClpXP, ClpAP, ClpCP, PAN/20S (archaea), Mpa/20S (prokaryotes), Rpt1-6/20S (eukaryotes). Two AAA+ modules (AAA+_1 and AAA+_2) are found in ClpA and ClpC, respectively. Extra elements are indicated as N (N-domain), TM (transmembrane segment), linker, and I (intermediate domain), respectively. AAA+ proteases are classified according to the review (54).

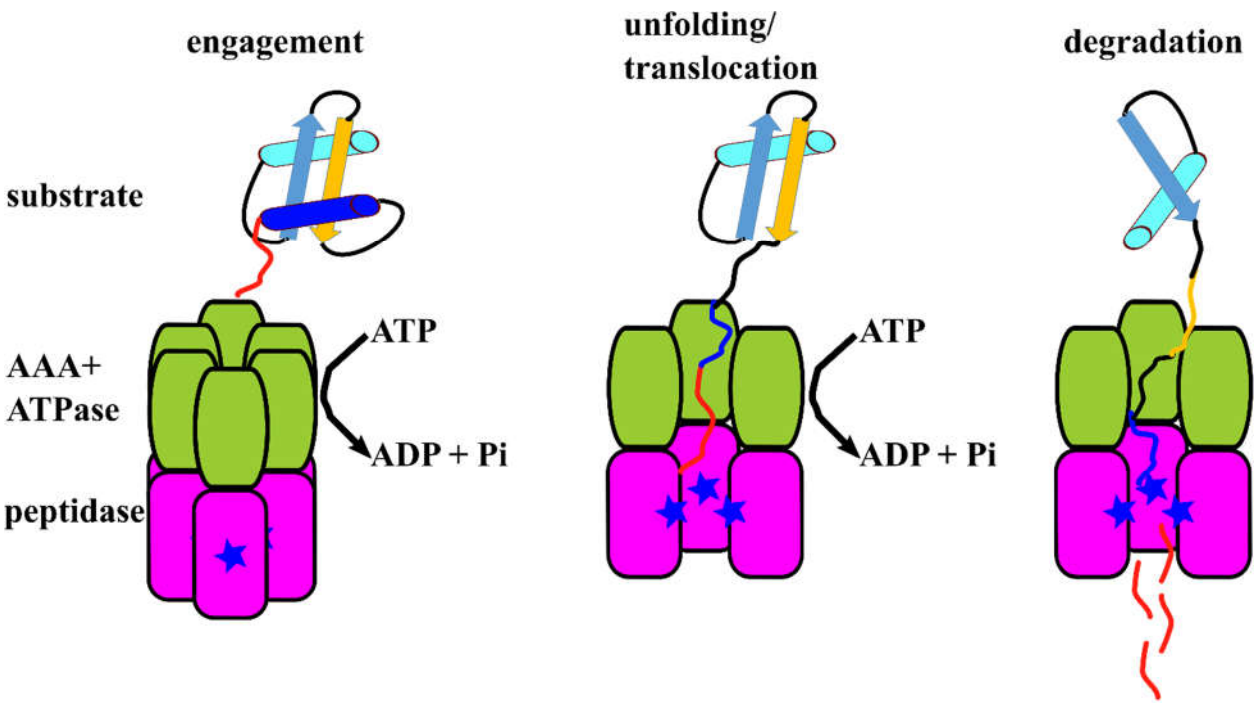


Figure 1.5 General mechanism of substrate degradation by a AAA+ protease.
Degradation of a native substrate by a AAA+ protease can be divided into several steps, engagement, unfolding, translocation, and degradation. A substrate is engaged by the AAA+ ATPase on the protease through an accessible degron present on the substrate (red fragment). AAA+ ATPase unfolds the native substrate and translocates the unstructured polypeptide through the central pore of the AAA+ module into the peptidase chamber. ATP molecules are hydrolyzed during unfolding and translocation. Denatured peptide chain is cleaved into fragments of oligopeptide by the peptidase. Blue star indicates the active site inside the peptidase chamber. Only 3 subunits of the protease are shown for clarity.

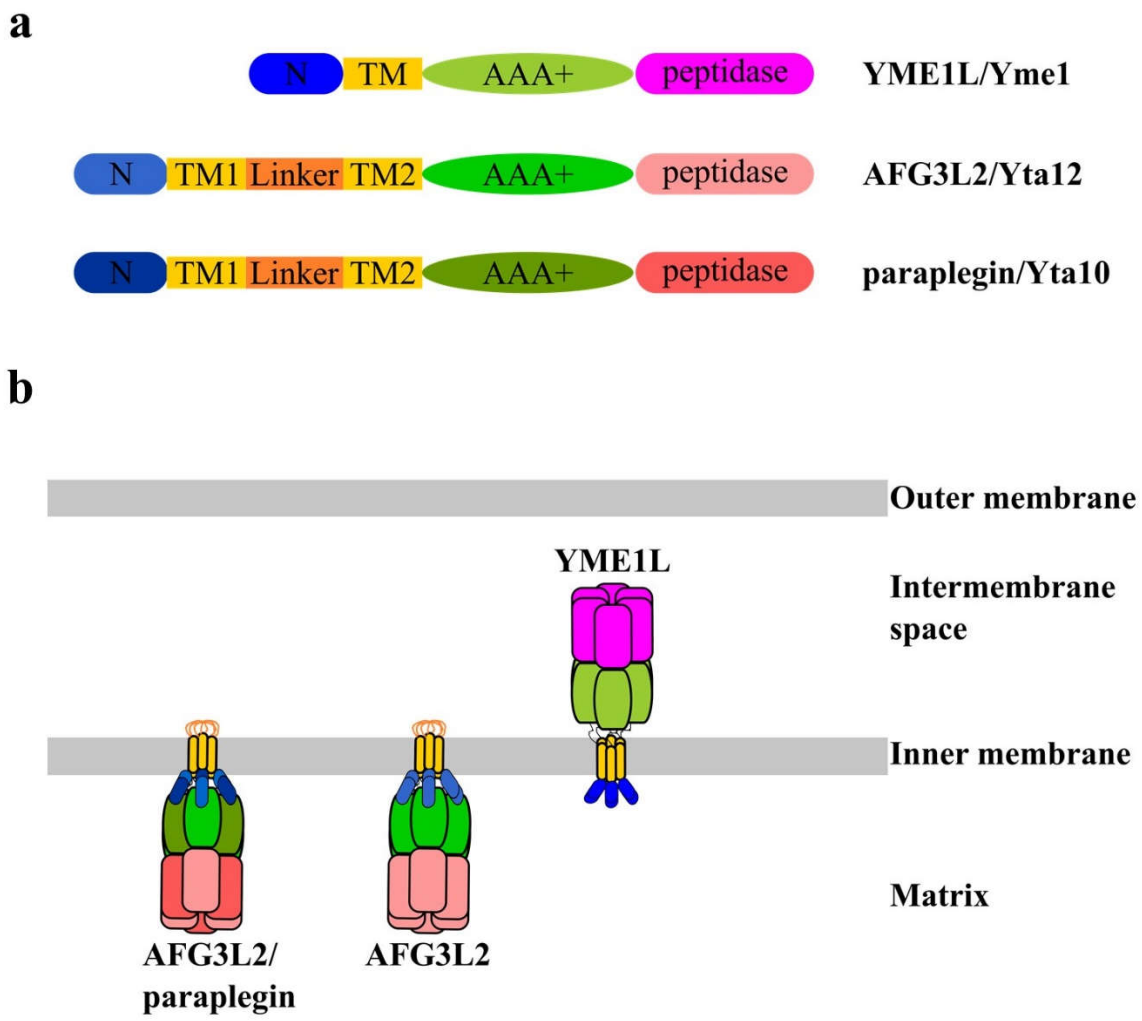


Figure 1.6 Mitochondrial AAA+ proteases in the inner membrane.

a. Domain organization of i-AAA (YME1L homohexamer in human and Yme1 homohexamer in yeast) and m-AAA (AFG3L2 homohexamer and AFG3L2/paraplegin heterohexamer in human, and Yta10/Yta12 heterohexamer in yeast). i-AAA contains an N-terminal domain (N), followed by a transmembrane segment (TM). The ATPase domain (AAA+) and the peptidase domain of i-AAA are located at the C-terminus of the protein. Similarly, m-AAA contains all four domains, with the only difference of two transmembrane spans (TM1 and TM2) present. **b.** Domain orientation of i-AAA and m-AAA in the inner membrane of mitochondria. The ATPase domain and peptidase domain of i-AAA reside in the intermembrane space, while the catalytic domains of m-AAA face the matrix. Colors are matched for each domain between the two panels.

Chapter 2 – Materials & Methods

The preparation of GFP variant constructs, purification of GFP variants and degradation of GFP variants were performed by Anthony Rampello in Dr. Steven Glynn laboratory at Stony Brook University.

2.1 Construct preparation

2.1.1 Human YME1L

The coding sequence of ATPase domain and peptidase domain (residues 317-773 of isoform 1) of human YME1L (UniProt Entry: Q96TA2), termed YME1L-AP, was amplified using PCR and constructed into a 2G-T vector (MacroLab, Addgene ID: 29707), a 2M-T vector (MacroLab, Addgene ID: 29708), and a 2S-T vector (MacroLab, Addgene ID: 29711), according to the Ligation independent cloning (LIC) protocol, respectively. The resultant construct includes a His_6-GST tag fused to the N-terminus of YME1L-AP.

To generate a hexameric homo-oligomer of YME1L-AP, hexYME1L, a codon optimized DNA sequence encoding a 32-residue polypeptide (ccHex (155) : GELKAIAQELKAIAKEL KAIA WELKAIAQGAG) was synthesized by Genscript, PCR amplified, and inserted into a modified 2G-T vector in our lab. Compared to YME1L-AP, the recombinant hexYME1L contains an extra sequence (GELKAIAQELKAIAKELKAIAWELKAIAQGAGGSGSYFQSNA) between His_6-GST and the catalytic domains of YME1L, with the underlined 10-residue sequence functioning as a linker separating cc-hex and YME1L-AP. The DNA sequence used to generate the modified 2G-T vector in our lab is inserted into the SspI cutting site in the original 2G-T vector. The DNA sequence is as follows: GGTAGCGGTAGT**GGCGAACTGAAAGCGATCGCG CAGGAGCTGAAGGCGATTGCCAAAGAATTAAAAGCCATCGCATGGGAGTTGAA AGCTATCGCTCAGGGCGCGGGG**TACTTCCAATCCAAT. Bolded sequence indicates the sequence encoding ccHex, and the flanking two short sequences are codons for linker amino acids.

Truncated constructs of YME1L were prepared by PCR amplification using YMEL1-AP as a template, then the amplicons were inserted into a 2G-T or 2M-T(MacroLab, Addgene ID: 29708), respectively. Constructs with a 2G-T plasmid backbone contain a His_6-GST fragment

fused to its N-terminus, respectively, whereas the 2M-T plasmid gives rise to an N-terminal His_6-MBP fusion protein. All the truncated YME1L constructs generated are summarized in Table 2.1.

The ATPase inactive variant of hexYME1L ^{E439Q} was prepared following "Round-the-horn" site-directed mutagenesis protocol (http://openwetware.org/wiki/'Round-the-horn_site-directed_mutagenesis), and the hexYME1L construct was used as a template.

2.1.2 i-AAA from *Myceliophthora thermophila*

i-AAA protease from *Myceliophthora thermophila* (UniProt Entry: G2QPI5) is termed IMT. To produce different truncations of IMT, cDNA sequences corresponding to the designated fragments were PCR-amplified and inserted into a 2G-T or 2M-T vector, respectively, following LIC protocol. Fusion proteins were attached to the N-termini of recombinant proteins similar to the truncations of YME1L, depending on the vectors used to produce the plasmids. Summary of truncations of IMT is shown in Table 2.2.

2.1.3 Substrates

2.1.3.1 I27 and variants

Constructs of I27 and certain variants (I27-β20, $I27^{CD}{}_{int}$β20, and β20-λCI-N) (94) were a generous gift provided by Dr. Robert Sauer (MIT). To create the mDHFR-I27-β20 chimeric construct, an mDHFR-sul20C plasmid, provided by Dr. Robert Sauer (MIT), was used as a template. The DNA sequence of C-terminal sul20C was initially exchanged by the coding sequence of β20 using PCR. Then a DNA sequence encoding I27 and a peptide (GSGS) which functions as a linker and is attached to the immediate N-terminus of I27, was PCR-amplified and inserted between mDHFR and β20 sequences.

$I27^{CD}$-17A and $I27^{CD}$-17B were constructed by "Round-the-horn" PCR, with the cDNA sequences of the C-terminal fragment (residues: 136-171) of human TIM17A (UniProt Entry: Q99595) and C-terminus (residues: 136-172) of human TIM17B (UniProt Entry: O60830) attached to the C-terminus of $I27^{CD}$, respectively. $I27^{CD}$ was produced by PCR using $I27^{CD}{}_{int}$β20 as a template, and the coding sequences of C-termini of human TIM17A and TIM17B were provided by Dr. R. L. Wiseman (The Scripps Research Institute).

2.1.3.2 Ups1/2-Mdm35 and PRELI/TRIAP1

Coding sequences of Ups1 (UniProt Entry: Q05776) and Mdm35 (UniProt Entry: O60200) from *Saccharomyces cerevisiae* were PCR amplified from yeast genomic DNA and subcloned into a pACYCDuet-1 (Novagen) vector together. Mdm35 was firstly cloned into the pACYCDuet-1 vector using NdeI/XhoI (New England Biolabs) restriction enzymes, then the DNA sequence of Ups1 was inserted into Mdm35-pACYCDuet-1 plasmid between the EcoRI/NotI cloning sites. This construct results in N-terminally His_6 tagged Ups1 and untagged Mdm35.

Ups2 (UniProt Entry: P35200)/Mdm35 were constructed using the same protocol as that used for Ups1/Mdm35, with Ups2 cloned between the EcoRI/NotI restriction sites. Human PRELI (UniProt Entry: Q9Y255) and TRIAP1 (UniProt Entry: O43715) were PCR amplified and cloned into the pACYCDuet-1 vector. TRIAP1 amplicon and pACYCDuet-1 were treated with NdeI/XhoI restriction enzymes, ligated and transformed into DH5α for plasmid reproduction. The resulting TRIAP1-pACYCDuet-1 plasmid was further employed to incorporate PRELI, utilizing the EcoRI/NotI cloning sites.

2.1.3.3 GFP variants

Plasmids containing the coding sequences of cp7-^{SF}GFP-β20, cp6-^{SF}GFP-β20, β20-cp7-^{SF}GFP, cp7-^{SF}GFP were generously provided by Dr. Robert Sauer (MIT) (88), and cp7-^{SF}GFP variants containing indicated degrons were produced by PCR, with coding sequences of degrons attached to the appropriate positions.

2.2 Protein expression and purification

All protein purification steps were carried out at 4 °C unless specified.

2.2.1 Human YME1L

Plasmids containing YME1L-AP (2G-T) or hexYME1L (modified 2G-T) were transformed into *E. coli* BL21 (DE3), separately, to allow expression under control by T7 polymerase and T7 promoter. Inoculation into 50 ml LB medium from corresponding glycerol stocks was carried out in the early morning and cells were grown at 37 °C until OD_{600} reached

around 0.2 – 0.3. Cells were then sub-cultured into flasks containing 1 liter LB medium each, with a ratio of 1:125. After the growth of cells at 37 °C reached a value of OD_{600} around 0.3, the temperature was brought down to 16 °C to allow additional growth until OD_{600} = 0.6. IPTG was added at a final concentration of 0.5 mM to induce protein expression at 16 °C for 16 hours. Cells were pelleted and washed in buffer A (25 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% glycerol, 0.1 mM EDTA, 10 mM $MgCl_2$).

After resuspension in buffer A supplemented with 10 mM BME and 1 mM PMSF, cells were lysed by sonication in the cold room. The cell lysate was spun down at 15, 000 rpm (Beckman Coulter, JA-20 rotor) for 30 minutes. Clarified supernatant was applied to Glutathione Superflow Agarose (Pierce) resins manually packed in a gravity column, which was equilibrated in buffer B (buffer A supplemented with 10 mM BME). Resins were then washed extensively in buffer B to remove unbound proteins. 10 mM glutathione was added to buffer B to elute target protein from the resins.

Fractions were pooled and the His_6-GST tag was removed from the fusion protein by TEV protease digestion overnight, with a ratio of 1 mg TEV / 50 mg fusion protein. Samples were applied to a gravity column containing Ni-NTA agarose (Thermo Scientific) equilibrated with buffer A supplemented with 10 mM imidazole in the absence of EDTA to remove the cleaved His_6-GST tag and TEV protease. YME1L-AP or hexYME1L containing fractions flowed through Ni-NTA agarose resin were collected and concentrated using an Amicon centrifugal filter unit (Merck Millipore) with a molecular weight cut-off of 30 kDa or 100 kDa, respectively. An SEC step in a Superose 6 10/300 GL column (GE Healthcare), equilibrated with buffer C (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 10% glycerol, 0.1 mM EDTA, 5 mM $MgCl_2$, and 1 mM DTT), was employed for final polishing of purification. Protein fractions corresponding to the correct oligomeric state were pooled and concentrated using an Amicon centrifugal filter unit with an appropriate molecular weight cut-off to a final concentration around 10 mg/ml, then flash-frozen in liquid nitrogen and stored at -80 °C. hexYME1L ^{E439Q} was expressed and purified using the same protocol as that of hexYME1L.

Expression and purification of truncations of YME1L were similar to YME1L-AP, except that Ni-NTA agarose resins instead of glutathione agarose resins were used as the first purification step for truncated constructs carrying 2M-T vectors. Proteins were subjected to SEC polishing for final purification in buffer C without EDTA and $MgCl_2$.

2.2.2 i-AAA from *Myceliophthora thermophila*

IMT constructs were expressed and purified by similar protocols as those for truncated constructs of YME1L. Strategies of initial purification were chosen depending on the tag fused to certain construct.

2.2.3 Substrates

2.2.3.1 I27 and variants

Plasmids harboring I27 and I27-β20 coding sequences were transformed into *E. coli* BL21 (DE3) for protein expression, separately. Cells were initially grown at 37 °C until OD_{600} = 0.6, then protein expression was induced by application of IPTG at a final concentration of 0.5 mM at 18 °C for 16 hours. Cells were harvested and re-suspended in buffer D (20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole, 10% glycerol, 1 mM PMSF). The cell lysate was clarified by centrifugation after sonication for cell lysis. Unbound proteins were removed by washing resins with buffer D supplemented with imidazole to a final concentration of 50 mM. 500 mM imidazole in buffer D was applied to elute proteins of interest. Proteins were further purified by SEC in a HiLoad 16/600 Superdex 200 column (GE Healthcare) in buffer E (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 10% glycerol). Peak fractions containing target proteins at positions of correct retention volume were pooled, concentrated, and flash-frozen prior to storage at -80 °C.

$I27^{CD}{}_{int}$β20, $I27^{CD}$-17A, and $I27^{CD}$-17B were expressed in *E. coli* BL21 (DE3). Initial culture was grown at 37 °C, and 0.5 mM IPTG was added when OD_{600} = 0.6 to allow induction of protein expression at 37 °C for 3 hours. Cells were harvested and lysed by sonication in buffer F (20 mM Tris-HCl pH 8.0). All three I27 variants were expressed as inclusion body and were pelleted by centrifugation at 15, 000 rpm for 30 minutes (Beckman Coulter, JA-20 rotor). Insoluble I27 variants were washed twice in buffer F then dissolved in buffer F with the addition of 6 M GuHCl. Target protein was incubated with Ni-NTA agarose resin followed by washing in buffer F containing 10 mM imidazole then eluted with 500 mM imidazole supplemented. Eluted fractions were subjected to a HiTrap Desalting column (GE Healthcare) for buffer exchange into 10 mM Tris-HCl pH 8.0, 100 mM NaCl. Proteins were concentrated, flash-frozen in liquid nitrogen and stored at -80 °C.

2.2.3.2 Ups1/Mdm35, Ups2/Mdm35, and PRELI/TRIAP1

pACYCDuet-1 plasmids containing yeast Ups1/Mdm35, Ups2/Mdm35, or human PRELI/TRIAP1 coding sequences were transformed into *E. coli* Origami 2(DE3) cells, separately. Cells were grown at 37 °C to OD_{600} = 1.2 and induced with 0.5 mM IPTG to allow co-expression for 5 hours at 30 °C. Cells were lysed by sonication in buffer G (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM PMSF and 40 mM imidazole). After clarifying by centrifugation, target proteins were bound to Ni-NTA agarose resin and unbound proteins were eliminated by applying buffer G. Protein complexes were eluted in buffer G containing 500 mM imidazole, and loaded onto a Superdex 75 10/300 GL column (GE Healthcare) equilibrated with buffer G in the absence of PMSF and imidazole for final purification. Target protein complex eluted at expected volume was pooled, concentrated, flash-frozen in liquid nitrogen and placed at -80 °C for long-term storage.

2.2.3.3 GFP variants and mDHFR-I27-β20

E. coli BL21 (DE3) cells containing plasmids that harbor coding sequences of cp6/cp7 variants with different degrons and mDHFR-I27-β20 were grown at 37 °C, separately. Proteins were expressed upon addition of 1 mM IPTG at 16 °C for 16 hours. After lysis by sonication and clarifying by centrifugation in buffer H (25 mM Tris-HCl pH8.0, 300 mM NaCl, 10% glycerol, 10 mM BME, and 10 mM imidazole), proteins were bound to Ni-NTA agarose resins. Target proteins were eluted following washing the resins thoroughly and purification was proceeded on a HiLoad 16/600 Superdex 200 column in buffer E supplemented with 1 mM DTT. Peak fractions containing proteins of interest were pooled, concentrated, flash-frozen in liquid nitrogen and stored at -80 °C.

2.3 Biochemical assays

2.3.1 ATPase assay

An established protocol of coupled assay (156) was employed to characterize ATPase activity of YME1L. A SpectraMax M5 plate reader (Molecular Devices) was used for the assay with a kinetics mode selected. Consumption of NADH was followed by monitoring the loss of absorbance at 340 nm for 20 minutes with an interval of 10 seconds between reads. Reaction was carried out in a 384-well clear/flat-bottom plate at 37 °C (Corning) in buffer I (25 mM HEPES-

KOH pH 8.0, 10% glycerol, 5 mM $MgCl_2$, 1 mM NADH, 21.5 U/ml LDH) supplemented with components for ATP regeneration (5 mM ATP, 7.5 mM PEP, and 18.8 U/ml pyruvate kinase PK) in a total volume of 30 µl, in the presence of 1 µM hexYME1L. Effects of substrates on ATPase activity at different concentrations of hexYME1L were assayed with 20 µM substrates included, respectively. The ATPase rate was calculated according to the formula: total ATPase rate = [rate of A340 nm signal loss (milliunit per minute)] / [pathlength (in cm) * 6.23 * hexYME1L concentration (in µM)]

2.3.2 Protein Degradation assay

Protein degradation was performed in buffer J (25 mM HEPES-KOH pH 8.0, 100 mM KCl, 10 % glycerol, 1 mM DTT, 10 mM $MgCl_2$, 25 µM $ZnCl_2$) at 37 °C, including an ATP regeneration system (5 mM ATP, 20 mM PEP and 18.75 U/ml PK). hexYME1L at a final concentration of 1 µM or 2 µM was introduced in the reaction, as indicated.

Michaelis–Menten kinetics characterization of GFP variants degradation was conducted in a 384-well black/flat-bottom plate (Corning) in a total volume of 30 µl at indicated substrate concentrations, implemented in a SpectraMax M5 plate reader. A kinetics mode was selected and the fluorescence signal (excitation = 467 nm, emission = 511 nm, auto cut-off = 495 nm) was recorded for 1 hour with a 10-second interval to quantify the degradation of GFP proteins. To calculate the rate of GFP variants degradation, the initial RFU was converted according to the concentration of GFP variants used in the reaction to correlate the RFU value and 1 µM GFP variant. Then the rate of protein substrates consumption was calculated, using [GFPs/(unit time * hexYME1L molecule)] as the rate of GFP degradation.

To quantify the degradation of non-fluorescent proteins and to visualize GFP variants degradation, proteins were resolved by SDS-PAGE then stained with Coomassie Brilliant Blue R-250 (Bio-Rad). Reactions were carried out in the presence of 20 µM substrate (10 µM Ups1/Mdm35) in a total volume of 100 µl, with the inclusion of 0.1 mg/ml CK for normalization of quantification. Aliquots were taken at indicated time points, mixed with 4X SDS-PAGE loading buffer, heated at 90 °C for 10 minutes and subjected to SDS-PAGE. Gels were scanned using Gel Logic 212 Pro (Carestream Health), quantified by ImageJ (157) and CK was used as a loading control to normalized the quantification.

The chimeric protein, mDHFR-I27-β20 (10 μM), degradation was assayed with MTX present (100 μM) or absent. Quantification of protein bands stained with Coomassie Brilliant Blue R-250 on SDS-PAGE was implemented by ImageJ (157) and the intensities were normalized to that of hexYME1L to minimize loading errors. The processivity ratio was calculated as previously described according to the definition (158).

2.4 Analytical size exclusion chromatography

200 μl of a total amount of 1 mg hexYME1L in the presence or absence of nucleotide was loaded onto a Superose 6 Increase 10/300 GL (GE Healthcare) column at 4 °C. Purified hexYME1L in the absence of EDTA or ATP analogue was incubated on ice for 1 hour, then injected onto the column equilibrated in buffer J (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM TCEP). EDTA-treated hexYME1L was incubated in buffer J supplemented with 5 mM EDTA on ice for 1 hour prior to loading onto the column, which was equilibrated in buffer J with the addition of 5 mM EDTA. To allow ATPγS to bind to hexYME1L, additional 10 mM ATPγS, and 10 mM $MgCl_2$ were added to EDTA-treated hexYME1L and 1-hour incubation on ice was allowed before applying to the column. 100 μM $MgCl_2$ and 100 μM ATPγS were included in the buffer for SEC analysis.

2.5 I27/I27-β20 carboxymethylation

Carboxymethylated I27/I27-β20 (^{CM}I27/^{CM}I27-β20) were produced similarly as previously described (158). Briefly, I27/I27-β20 were unfolded in a buffer containing 0.6M Tris-HCl pH 8.6, 6M GuHCl, 5mM DTT, for 3 hours at room temperature. Cysteine residues in the reduced form were then chemically modified by carboxymethylation upon the addition of freshly prepared iodoacetic acid with a molar ratio of 100-fold excess and incubated for 2 hours in the dark. Following carboxymethylation, proteins were loaded onto a HiTrap Desalting column (GE Healthcare) for buffer exchange into 5 mM sodium phosphate pH 8.0.

2.6 Circular Dichroism

Circular dichroism characterization of native and carboxymethylated I27/I27-β20 was carried out at 25 °C and a protein concentration of 20 μM in 5 mM sodium phosphate pH 8.0. A

quartz with a path length of 1 mm containing samples was placed in an Olis RSM CD spectrophotometer (Olis). Raw data of ellipticity were taken from 190 nm to 310 nm, with total 240 data points. Three scans were performed three times and the values were averaged. 1 mg/ml camphorsulfonic acid dissolved in ddH_2O was used at the beginning of the CD measurement as a standard to monitor the status of the CD instrument. Molar ellipticity was calculated from raw data with a unit of degs * cm^2 * mol^{-1}.

2.7 UV-VIS spectroscopy

Proteins at indicated concentrations (I27/I27-β20, ^{CM}I27/^{CM}I27-β20) were placed in a quartz cuvette with a path length of 1 cm, individually. Protein solution was resuspended well using pipettes. The absorbance signal was recorded for wavelengths from 200 nm to 550 nm at 37 °C using a SpectraMax M5 plate reader, with a step size of 1 nm. Absorbance signal of wavelengths ≥ 320 nm was examined to monitor protein aggregation (159).

2.8 X-ray crystallography

2.8.1 Crystallization

Crystallization was carried out using the hanging-drop vapor diffusion method by mixing an equal volume of protein and reservoir solution at 20 °C. 7mg/ml number 11 construct of YME1L (YME1L-11) was incubated with 5 mM ADP/$MgCl_2$ or 5 mM ATPγS/$MgCl_2$ for 1 hour on ice, respectively. YME1L-11 were then screened against commercially available sparse screening kits by mixing 1 µl of the protein with 1 µl of reservoir solution.

7.5 mg/ml of number 4 construct of IMT (IMT-4) and 7 mg/ml of number 9 construct of IMT (IMT-9) were allowed for a 1-hour incubation with 5 mM ADP/$MgCl_2$ on ice, respectively. Crystals of IMT-4 were grown by mixing 1 µl of IMT-4 with 1 µl of reservoir solution containing 0.2M L-Proline, 0.1 M HEPES pH 7.4, 9% PEG 3350. 30-35% glycerol were included in reservoir solution as a cryoprotectant for IMT-4 crystals. IMT-9 crystals were obtained by mixing 1 µl of the protein with 1 µl of reservoir solution containing 0.2M Ammonium Citrate pH7.3, 14% PEG 3350. Crystals were cryoprotected in reservoir solution with the addition of 25% glycerol. After cryoprotected, all crystals were flash-frozen in liquid nitrogen.

2.8.2 Data collection and structure determination

Diffraction data were collected at the BL14-1 beamline of the Stanford Synchrotron Radiation Lightsource (SSRL, Stanford, California). IMT-4 crystals did not give rise to a good diffraction quality. Diffraction data of IMT-9 crystals were collected to 2.45 Å at a wavelength of 1.18076 Å using a CCD MAR325 detector in oscillation mode with an oscillation angle of 0.5° and an exposure time of 3.0 seconds at 100 K. A total number of 460 diffraction images were collected.

Data were indexed, integrated, and scaled using HKL2000 (160), Molecular replacement was implemented in Phaser (161) for calculating the initial phase with an FtsH structure (PDB code: 1LV7) prepared by phenix.sculptor (161) as a search model guided by sequence alignment. Model completion was initially performed using phenix.autobuild (161), and further refinement was carried out by iterations of manual model building in COOT (162) and phenix.refine (161) with NCS restraints applied. TLS refinement was included at the later stage of refinement with the definition of TLS group determined by TLSMD implemented in PHENIX (161), and waters were added when refinement was close to complete under the guidance of difference maps and examination of the surrounding chemical environment. Model quality was validated by MolProbity (163). Structure figures were prepared using PyMOL (The PyMOL Molecular Graphics System, Version 1.8.4.0 Schrödinger, LLC.)

Table 2.1 Deletion constructs of the human YME1L protein.

AP stands for constructs containing ATPase and peptidase domain; number stands for the numbering of constructs generated. Fragment indicates the starting and ending positions of corresponding construct. Constructs are cloned into 2G-T or 2M-T vector, with corresponding GST or MBP fusion tag at the N-terminus. Columns of Ext.Coff., Mw, and pI show the extinction coefficient, molecular weight and isoelectric point of the corresponding construct after cleavage of fusion tag, respectively.

Domain	number	Fragment	Vector	Ext.Coff.	Mw	pI
AP	1	331-765	2G-T	24300	48521	6.07
	2	331-767	2G-T	24300	48707	5.95
	3	331-773	2G-T	24300	49460	6.21
	4	335-765	2G-T	24300	48033	5.95
	5	335-767	2G-T	24300	48219	5.84
	6	335-773	2G-T	24300	48973	6.08
	7	344-765	2G-T	24300	47021	5.88
	8	344-767	2G-T	24300	47207	5.77
	9	344-773	2G-T	24300	47961	6.01
AAA+	10	331-580	2M-T	10930	27783	5.63
	11	331-585	2M-T	10930	28409	5.89
	12	335-580	2M-T	10930	27297	5.45
	13	335-585	2M-T	10930	27922	5.65
	14	344-580	2M-T	10930	26284	5.31
	15	344-585	2M-T	10930	26910	5.49

Table 2.2 Deletion constructs of the i-AAA protein from Myceliophthora thermophila.

AP stands for constructs containing ATPase and peptidase domain; number stands for the numbering of constructs generated. Fragment indicates the starting and ending positions of corresponding construct. Constructs are cloned into 2G-T or 2M-T vector, with corresponding GST or MBP fusion tag at the N-terminus. Columns of Ext.Coff., Mw, and pI show the extinction coefficient, molecular weight and isoelectric point of the corresponding construct after cleavage of fusion tag, respectively.

Domain	number	Fragment	Vector	Ext.Coff.	Mw	pI
AP	1	238-670	2GT	30700	47592	7.31
	2	238-680	2GT	30700	48718	7.89
	3	241-670	2GT	30700	47265	6.90
	4	241-680	2GT	30700	48391	7.31
	5	249-670	2GT	30580	46635	6.78
	6	249-680	2GT	30580	47761	7.14
AAA+	7	238-483	2MT	9650	26914	9.25
	8	238-488	2MT	9650	27486	9.36
	9	241-483	2MT	9650	26587	9.12
	10	241-488	2MT	9650	27159	9.25
	11	249-483	2MT	9530	25685	9.22
	12	249-488	2MT	9530	26257	9.34

Chapter 3 – Establishment of an *in vitro* system and the biochemical characterization of YME1L

Most results in this chapter have been published in the research article: **Shi H**, Rampello AJ, Glynn SE. Engineered AAA+ proteases reveal principles of proteolysis at the mitochondrial inner membrane. Nature Communications 7: 13301. doi:10.1038/ncomms13301

The preparation of GFP variant constructs, purification of GFP variants and degradation of GFP variants (Figure 3.12b-d, Figure 3.13, Table 3.1, and Table 3.2) were performed by Anthony Rampello in the Steven Glynn laboratory at Stony Brook University.

3.1 Reconstruction of an active hexameric i-AAA enzyme

3.1.1 The catalytic core domains of YME1L do not assemble into active hexamers

As mentioned previously, the lack of an *in vitro* system for exploring YME1L biochemically greatly impeded our understanding of this important protease. In order to obtain reasonable amount of the YME1L protein for *in vitro* assay and structural studies, a prokaryotic expression system was utilized. The catalytic domains of humanYME1L range from residues 317 to 773, containing the AAA+ ATPase module and the M41 peptidase domain (UniProt ID: Q96TA2). This fragment of YME1L was sub-cloned and was termed YME1L-AP (Figure 1.8a). When expressed in the *E. coli* strain BL21 (DE3), good yields of YME1L-AP proteins with different N-terminal fusion tags were obtained upon induction by IPTG (Figure 3.1a). Comparing the relative expression levels and ease of purification of these proteins, the His_6-GST-fused YME1L-AP construct was chosen for use in future experiments. After multiple purification steps and removal of the His_6-GST tag by cleavage with TEV protease, the purity of final YMEL1-AP protein was greater than 95% as assayed by SDS-PAGE (Figure 3.1b). As AAA+ modules are typically required to hexamerize in order to carry out ATP hydrolysis (54, 64), an oligomeric form of YME1L-AP was expected to be observed. However, YME1L-AP eluted approximately as a monomer as calculated from the elution volume in size exclusion chromatography (Figure 3.2). Previous studies have shown that AAA+ proteins can form oligomers in a concentration and/or nucleotide-dependent manner, as in the case of ClpB (164-166). This transition among different

oligomeric forms could also exist in YME1L-AP. Therefore, the monomeric state observed in size exclusion chromatography might be a result of the dilution of YME1L-AP when migrating on the column.

A protocol of three-step coupled enzyme reactions (156) was used to test whether YME1L-AP is competent to hydrolyze ATP (Figure 3.3). Disappointedly, YME1L-AP did not show any detectable activity of ATP hydrolysis, even at high concentrations of ATP or enzyme. Although previous studies showed that the cytosolic fragment of the YME1L homolog, FtsH, from a thermophilic bacteria is hexameric and active (85), it is very likely that YME1L-AP exists as a monomer and is inactive in hydrolyzing ATP in solution regardless of the concentration of YME1L-AP or nucleotide, which is consistent with previous results from yeast Yme1 (143).

3.1.2 hexYME1L forms a hexamer *in vitro* and is able to hydrolyze ATP and degrade casein.

The inability of YME1L-AP to form a hexamer in solution likely accounts for its lack of ATPase activity, which may be caused by the removal of the N-terminal region (the N-domain and the transmembrane segment) preceding YME1L-AP. Indeed, it has been suggested that the second transmembrane segment of *E. coli* FtsH is important for its oligomeric state and proteolytic activity (135), which was further supported by *in vivo* (167) and *in vitro* (168) experiments. We sought ways to modify the YME1L-AP construct to promote its oligomerization, by replacing the transmembrane sequence with a soluble counterpart. Although fusing leucine zipper or transmembrane spans from other mitochondrial inner membrane proteins to the N-terminus of cytosolic fragment induce oligomerization of FtsH (167, 168), the methods do not guarantee the formation of correct hexamers. After testing different approaches, a rationally designed peptide comprised of 32 residues was found to satisfy our purpose (155). The so-called cc-hex peptide homo-hexamerizes in solution (Figure 2f from reference (155)) (Figure 3.4a) with a size of 47.8 Å * 33.5 Å * 32.2 Å as measured for in PyMOL. We appended the cc-hex sequence to the N-terminus of YME1L-AP, with a linker sequence separating the cc-hex and YME1L-AP, and the resulting construct was named hexYME1L (Figure 3.4b). The hexYME1L protein was purified to a comparable purity to that of YME1L-AP (Figure 3.5a). As expected, the elution peak of hexYME1L shifted to a position corresponding to a higher molecular weight by size exclusion chromatography, compared to that of YME1L-AP (Figure 3.5b). The calculated apparent molecular weight of hexYME1L is 295.6 kDa, close to the theoretical hexamer molecular weight

of 333.0 kDa, indicating that hexYME1L exists in a hexamerization state in solution. Taking the above-mentioned nucleotide-dependent oligomerization into account, we next examined whether the hexamerization of hexYME1L depends on incorporation of nucleotides, although no nucleotide is supplied when purifying hexYME1L and hexamerization of cc-hex does not rely on nucleotide. Unsurprisingly, no obvious change was seen on the size exclusion chromatography profiles of hexYME1L upon addition of the non-hydrolysable analog ATPγS, or when treated with a high concentration of EDTA to chelate magnesium ions and interrupt binding of any endogenously bound nucleotides (Figure 3.5c). Taken together, it can be concluded that hexYME1L exists as an oligomer in solution, most likely as a hexamer, and the oligomerization behavior depends on the addition of cc-hex, rather than nucleotide.

The formation of a correct hexamer is a prerequisite for high ATPase activities of AAA proteins (64, 135, 165). The assembly of monomers into a hexamers results in the formation of interfaces between adjacent monomers, allowing information to be passed between neighbouring subunits to coordinate movements and ATP hydrolysis within the hexameric ring (54, 64). We then used the same ATPase assay (Figure 3.3) (156) to confirm that the reengineered hexYME1L is functional in efficiently hydrolyzing ATP. After detection of ATPase activity of hexYME1L, the reaction conditions of the assay were optimized against pH and salt type and enzyme concentration. hexYME1L is most active at pH 8.0 and without salt present in the buffer (Figure 3.6a). This is not surprising, as previous research has suggested that activities of AAA proteins are lower at high salt concentration due to the destabilization of hexamers (164, 166). ATPase activity at pH values higher than 8.0 or lower than 7.0 were also tested, but the results showed lower ATP hydrolysis activity. The optimal conditions were then used to measure steady-state kinetics of hexYME1L ATPase activity, and the data fit well to the Hill version of the Michaelis–Menten equation (R^2 = 0.9832) (Figure 3.6b). hexYME1L exhibits comparable activity (Vmax = 42.4 ATPs * min^{-1} * $hexYME1L^{-1}$, $K_{1/2}$ = 1.4 mM) to those determined for purified FtsH (135) and 26S proteasome (169). The value of the Hill coefficient is 2.4, suggesting a positive cooperativity in ATP hydrolysis by hexYME1L.

As a AAA+ protease, YME1L catalyzes the degradation or processing of protein substrates (49). AAA+ proteases unfold a stable substrate, followed by translocation of the unstructured polypeptide into the peptidase chamber where it is digested into small peptide fragments (54). In order to demonstrate that hexYME1L is able to degrade proteins, a general substrate of proteases,

β-casein, was employed. Because of the unstructured property of β-casein, the initial unfolding step in the degradation process is not involved. β-casein was degraded by hexYME1L in the presence of ATP, whereas hexYME1L did not show any protease activity against β-casein when ATP is absent or ATPγS is provided instead of ATP (Figure 3.7a-c, e). To rule out the possibility that the observed casein degradation is a result of contaminating enzymes, we substituted the conserved glutamate (position 439) in the Walker B motif with a glutamine (hexYME1L^{E439Q}). The glutamate is the catalytic residue for ATP hydrolysis, activating the water molecule bound near the ATP, which then carries out nucleophilic attack to break the terminal phosphate bond of the ATP molecule. This E439Q substitution severely impairs the ATPase activity of hexYME1L, yet still allows ATP to bind to hexYME1L. No degradation of β-casein by hexYME1L^{E439Q} was observed within the 1-hour reaction (Figure 3.7d, e). In addition, no appreciable ATP hydrolysis was seen for hexYME1L^{E439Q}. The rate of β-casein degradation by hexYME1L was calculated to be 0.49 ± 0.08 molecules min^{-1} hexYME1L^{-1}, based on data points within the initial 20 minutes of the reaction (Figure 3.7f). In summary, these results indicate that hexYME1L is an active enzyme in solution, holding both of the ATPase and protease activities, and that hexYME1L is able to translocate and degrade unfolded substrates in an ATP-dependent manner.

3.2 Degradation of protein substrates by YME1L requires a degron sequence

Similar to hexYME1L, FtsH from *Thermos thermophiles* has been shown to digest α-casein, as well as various unstructured model substrates at high temperature (65 °C) with no apparent strictly conserved motif (170). Also, *E. coli* FtsH was shown to degrade its natural substrate σ^{32} both *in vivo* and *in vitro* (66, 135). Considering the wide range of substrates of YME1L, it is reasonable to speculate that there exists some criteria for YME1L to select proteins to be degraded or processed. The selection can be achieved by either recognition of accessible degrons (54, 171) or simply engaging exposed unfolded polypeptide (170). Some well-characterized degrons of other AAA+ proteases include ssrA (54, 172), sul20C and β20 (89, 90, 94).

To distinguish between these models of substrates selection by YME1L, we took advantage of a model substrate, the I27 domain of human titin (173). Degradation reactions were performed on four substrates: the native protein with a folded structure (I27), the folded protein with a C-terminal β20 tag (I27-β20) (β20 sequence: QLRSLNGEWRFAWFPAPEAV) (94), a denatured I27 protein generated by carboxymethylation of cysteine reisdues (^{CM}I27), and the unfolded I27

carrying a β20 tag at the C-terminus (^{CM}I27-β20) (Figure 3.8). Carboxymethylation of the two cysteines that fold into the interior of I27 irreversibly unfold the protein but retain the solubility of I27 (173). Successful disruption of the native structure of I27 proteins was verified by circular dichroism spectroscopy (Figure 3.9a). In order to rule out the possibility of aggregation of proteins as a result of denaturation, UV-VIS spectra for the proteins were taken at a wavelength range of 200 – 550 nm (Figure 3.9b). Particularly, no discernable absorbance can be recognized above 320 nm (Figure 3.9c), even at concentrations about 15 fold of those used in the degradation reactions for ^{CM}I27 and ^{CM}I27-β20 (Figure 3.9d), indicating the non-aggregated status of these two proteins in solution (89). hexYME1L did not degrade I27 and ^{CM}I27 (Figure 3.8a and 3.8b). Nevertheless, I27-β20 and ^{CM}I27-β20 carrying a β20 degron at the carboxyl terminus were rapidly degraded in the presence of ATP (Figure 3.8c and 3.8d), regardless of the folding state. However, hexYME1L degraded both proteins with a degron attached, the unstructured substrate was degraded faster than the folded one. The degradation rate of ^{CM}I27-β20 was determined to be 0.31 ± 0.020 molecules min^{-1} hexYME1L^{-1}, and that of I27-β20 is 0.24 ± 0.003 molecules min^{-1} hexYME1L^{-1}. This correlates with the general mechanism that an unfolding step is required for a native substrate before it can be translocated into the peptidase chamber for degradation (Figure 1.7). It has been presented that Lon, a AAA+ protease, degrades substrate bearing an internal β20 degron (94). Whether the change in position of β20 in a substrate affects its degradation by hexYME1L is an interesting question to explore. We used the same construct (I27$^{CD}_{int}$β20) (94) to examine this problem. I27$^{CD}_{int}$β20 carries a β20 tag at the interior of I27, close to the N-terminal portion. In addition, the two cysteines buried inside the I27 protein were replaced by aspartic acid. Similar to carboxymethylation, the substitution unfolds the protein yet preserves the solubility. Compared to the unstructured ^{CM}I27-β20, I27$^{CD}_{int}$β20 was degraded by hexYME1L at a much slower rate, suggesting a role that position of a degron plays in substrate degradation by hexYME1L.

Notably, in the case of I27, ^{CM}I27, and I27$^{CD}_{int}$β20 degradation, hexYME1L was found to be auto-degraded to some extent when degradation reactions were allowed to proceed up to 3 hours (Figure 3.8a-b, 3.8e). However, auto-degradation of hexYME1L was not significant within one hour when substrates with fast degradation rate, for example, I27-β20 and ^{CM}I27-β20 are present.

To summarize, these results imply that for YME1L to degrade proteins, a suitable degron sequence is required, and that the position of the degron affects substrate degradation by the YME1L enzyme.

3.3 YME1L processively degrades substrates

Protein degradation in a processive mode has been seen in AAA+ proteases, which states that proteases unfold and degrade substrates from the degron end (48, 89, 174-176). This mechanism is distinct from one where any accessible unstructured region is translocated then cleaved following binding of the substrates on the proteases. In order to differentiate these two models, a previously described protein chimera containing a C-terminal β20 degron (mDHFR-I27-β20) was employed (89, 158). mDHFR-I27-β20 was fully degraded by hexYME1Lwithin one hour. However, in the presence of methotrexate (MTX), a proteolysis intermediate accumulated during the course of degradation (Figure 3.10a). Methotrexate is a small molecule that stabilizes mDHFR and prevents it from unfolding when being incorporated into the protein (89, 158). The accumulation of a proteolytic intermediate with a size of mDHFR in the presence of MTX suggests the processive mode of substrate degradation by hexYME1L. The mDHFR domain in the chimeric protein is processively unfolded, translocated and degraded from the C-terminus. When mDHFR is stabilized by MTX, which results in its resistance to unfolding by hexYME1L, a partially degraded product corresponding to the MTX-stabilized mDHFR is released from the enzyme. The intermediate accumulated was quantified, and a processivity ratio defined previously (158) was determined to be ~6 (Figure 3.10b).

The processive model of protein degradation by AAA+ proteases is also indicated by the elevated ATPase activity in the presence of substrates. Substrate-stimulated ATPase rate implies the interaction between the pore loops of AAA+ proteases and the substrates being translocated. Interestingly, such stimulation was detected upon the supplementation of I27 variants (Figure 3.11a). However, the observed auto-degradation of hexYME1L may potentially complicate the calculation of stimulated ATPase rate. We determined the amount of lost enzyme as less than 6% within one hour in the presence of I27-β20 or ^{CM}I27-β20. The ATPase rates of hexYME1L increased significantly upon addition of 20 μM I27-β20 or ^{CM}I27-β20 to the reaction at a hexYME1L concentration of 0.25 μM (Figure 3.11b). No significant increment of ATPase activity could be detected in the presence of I27 or ^{CM}I27. Also, when assayed at higher concentrations of hexYME1L, the stimulation of ATPase activity was less obvious, especially in the case of the unfolded ^{CM}I27-β20 (Figure 3.11c). These results reveal that upon the engagement of substrates through degrons, degradation by YME1L is a processive course from the degron end.

3.4 YME1L is able to degrade substrates with certain stabilities

Degradation of the folded I27 domain and the native mDHFR protein (as shown by the degradation of entire chimeric mDHFR-I27-β20 protein) by hexYME1L requires an unfolding force to be applied on the stable substrates. This is in contrast to previous results that show that FtsH, the bacterial prototype of the YME1L i-AAA protease, does not possess a robust unfolding power (150), and that Yme1, yeast homolog of YME1L, degrades destabilized proteins but not the native ones (109, 143). To substantiate our finding of the capability of hexYME1L to unfold stable proteins followed by substrate translocation and degradation, we investigated the degradation of proteins with characterized stabilities (Figure 3.12). As has been shown above, the β20 tag is able to convert a non-substrate of hexYME1L (I27 or ^{CM}I27) to a protein that can be degraded by hexYME1L (I27- β20 or ^{CM}I27-β20). The β20 degron was then appended to λCI-N (the N-terminal domain of the λCI protein) and different circularly permuted GFP proteins (cp6-^{SF}GFP and cp7-^{SF}GFP) of distinct local stabilities (88). λCI-N is stable (T_m = 54 °C) and almost completely folded at 37 °C (177). To a higher extent, the stabilities of various circularly permuted GFP proteins were determined to have the lowest T_m value of 68.3°C (178). The degradation of N-terminally β20-tagged λCI-N (Figure 3.12a) in an ATP-dependent manner strongly suggests that hexYME1L applies unfolding force to a stable substrate to initial the degradation process. In addition, the result demonstrates that hexYME1L can also engage an N-terminal degron and allow degradation to proceed towards the carboxyl terminus.

As shown in Figure 3.12b-d, cp7-^{SF}GFPs carrying a β20 tag at either terminus were degraded by hexYME1L, although the degradation of N-terminally tagged protein (β20-cp7-^{SF}GFP) seemed to be slower than that of C-terminally tagged version (cp7-^{SF}GFP-β20), which is consistent with results from the ClpXP protease (179). Again, this supports that the β20 tag can be located at the N-terminus of a substrate to allow it to be degraded by hexYME1L. In contrast, cp7-^{SF}GFP without a degron attached was not cleaved (Figure 3.12b and 3.12c). Interestingly, cp7-^{SF}GFP with a C-terminal β20 degron (cp7-^{SF}GFP-β20) was degraded at a faster rate than cp6-^{SF}GFP-β20. Despite the common C-terminal β20 sequence, the two proteins were reported to have distinct unfolding pathways and local stabilities (88, 179), which explains the discrepancy of degradation rates of the two GFP variants. Michaelis-Menten kinetics characterization of the apparent best substrate, cp7-^{SF}GFP-β20, was carried out in a plate reader by measuring the loss of

GFP fluorescent signal emitted at 511 nm (Figure 3.12d). The calculated maximal degradation rate is 0.11 GFP molecules min^{-1} hexYMEL1^{-1}. This rate is about 7-fold slower than the reported degradation rate of cp7-^{SF}GFP, which carries a preferred C-terminal Sul20C degron, by the Lon protease from *E. coli* (88). The K_M value of cp7-^{SF}GFP-β20 degradation by hexYME1L was determined to be 7.1 μM, about 2.5 fold higher than that of cp7-^{SF}GFP-Sul20C degradation by Lon (88). The different degrons (β20 and Sul20C) in these experiments may contribute to the variation in kinetics of substrate degradation. However, hexYME1L seems to be intrinsically weaker in unfolding cp7-^{SF}GFP, compared to *E. coli* Lon. Degradation of β20-cp7-^{SF}GFP by Lon at a faster rate (1.2 molecules min^{-1} Lon^{-1}) substantiates this hypothesis. Indeed, it has been suggested from earlier experiments that different proteases have distinct capability of unfolding stable proteins (158). We conclude that although YME1L is a protease of seemly weaker unfolding ability, it does unfold and degrade substrates with substantial stability, as can be seen in the examples of λCI-N and GFP variants. The effect of the location of β20 degron on cp7-^{SF}GFP may be a result of the different unfolding kinetics from N- or C- terminus.

3.5 Substrate degradation by YME1L is affected by degron sequence

As has been shown for other AAA+ protease, a degron is necessary for inducing substrate degradation. Alterations in degron sequences result in changes in proteins degradation, as exemplified by the well-characterized ssrA and β20 degrons (81, 94, 180). The requirement of degrons with specific sequences or motifs is critical for AAA+ proteases to execute their physiological functions. Considering that there are a variety of YME1L substrates of different functions, such requirement poses a criterion by which correct proteins are selected from a crowded environment full of other molecules.

To better understand the impact of degron sequences on substrate degradation by YME1L, we investigated cp7-^{SF}GFP degradation with several know degrons attached to the C-terminus (Figure 3.13 and Table 3.1). cp7-^{SF}GFP was chosen due to its fastest apparent degradation rate among the substrates tested (Figure 3.12). Firstly, the set of degrons containing the first five degrons in Table 3.1 (β20, ssrA, sul20C, β10, and ^{ext}ssrA) were examined (Figure 3.13a). Although ssrA and sul20C have been demonstrated as good degrons for ClpXP and Lon, respectively, hexYME1L appeared to prefer the β20 degron over ssrA (11 residues: AANDENYALAA) or sul20C (20 residues: ASSHATRQLSGLKIHSNLYH). Degradation of cp7-^{SF}GFP with either ssrA

or sul20C attached to the C-terminus is slower than that of β20-tagged cp7-^{SF}GFP. Kinetic parameters of the ssrA tag (k_{deg} = 0.031±0.004 molecules min^{-1} hexYME1L^{-1}, K_M = 12.0±4.0 μM) and the sul20C tag (k_{deg} = 0.036±0.005 molecules min^{-1} hexYME1L^{-1}, K_M = 15.2±5.0 μM) reveal slower degradation of the two proteins, compared to that of cp7-^{SF}GFP-β20 (k_{deg} = 0.110±0.003 molecules min^{-1} hexYME1L^{-1}) (Figure 3.13b and 3.13d). Despite having the same length of 20-residue unstructured fragment, cp7-^{SF}GFP-sul20C was degraded at a rate about 3-fold slower than that of cp7-^{SF}GFP-β20, indicating that amino acid composition in a degron sequence affects substrate degradation by hexYME1L. The degradation rate in the case of 11-residue ssrA-tagged cp7-^{SF}GFP is about four-fold lower than that of cp7-^{SF}GFP-β20 (Figure 3.13b and 3.13d). This result can be interpreted in two ways. It is possible that the 11-residue long ssrA tag does not allow optimal association of the substrate with hexYME1L. This is reasonable since studies on FtsH have suggested that a hydrophobic terminus of 20 or more residues are necessary for degrading protein substrates *in vivo* (181). Alternatively, it can be explained by the distinct amino acid identities between ssrA and β20.

In order to further probe the two possibilities, we constructed a truncated version of the β20 tag that lacks the N-terminal 10 residues. The shortened tag was termed β10 (FAWFPAPEAV). Surprisingly, this cp7-^{SF}GFP-β10 variant was degraded almost at the same rate (k_{deg} = 0.105±0.015 molecules min^{-1} hexYME1L^{-1}) as cp7-^{SF}GFP-β20, with an increased K_M value (K_M = 12.4±1.9 μM) (Figure 3.13b and 3.13d). This result suggests that the 10 residues at the N-terminus of β20 degron does not contribute to the substrate engagement by hexYME1L at a significant level. Also, a short degron containing only 10 residues was able to intermediate substrate recognition. An extended ssrA tag (^{ext}ssrA, QLRSLNGEAANDENYALAA) comprising the 10-residue N-terminal portion of β20 was introduced to the C-terminus of cp7-^{SF}GFP, resulting the cp7-^{SF}GFP-^{ext}ssrA construct. Compared to the ssrA-tagged cp7-^{SF}GFP, this modified cp7-^{SF}GFP-^{ext}ssrA was degraded at a much faster rate (k_{deg} = 0.106±0.007 molecules min^{-1} hexYME1L^{-1}) than ssrA but with a K_M value (K_M = 20.4±3.1 μM) about 3-fold higher than that of cp7-^{SF}GFP-β20 (Figure 3.13b and 3.13d). The results from the first set of 5 degrons indicate that YME1L distinguishes amino acid composition of degrons for substrate selection, with additional contribution from the length of degrons to the kinetic parameters of degradation.

A second set of modified β20 degrons were subsequently constructed, comprising the last four members in Table 3.1 (β5, βF, βF2, and β5F). β20 was chosen as for further characterization

of degron recognition as it was the fastest degraded of the three initial degrons tested (β20, sul20C and ssrA) tested. All four degrons were attached to the C-terminus of cp7-^{SF}GFP. A truncated version bearing only the C-terminal five amino acids (β5: APEAV) resulted in an unmeasurable k_{deg} and K_M as cp7-^{SF}GFP-β5 could not saturate the enzyme (Figure 3.13c), even at the high concentrations. This indicates that the recognition of β5 by hexYME1L is too weak to be assayed. The poor association of β5 can be explained by the requirement of a longer degron sequence than 5 residues. Or it can be interpreted as that the 5 amino acids removed from β10 to generate β5 are the key residues involved in substrate engagement, considering that the degradation rate of β10 is almost the same as that of β20. We then produced a construct with a C-terminal fusion tag consisting of these 5 residues potentially critical for substrate recognition (βF: FAWFP). The cp7-^{SF}GFP-βF protein showed slow but measurable degradation rate (k_{deg} = 0.106±0.007 molecules min^{-1} hexYME1L^{-1}), comparable to cp7-^{SF}GFP-sul20C (Figure 3.13b-d). This result suggests that degron having a 5-residue length is sufficient to induce substrate binding to hexYME1L, and possibly, this sequence is primarily responsible for the recognition of substrate by hexYME1L. Introduction of another copy of βF generated an extended degron, βF2 (FAWFPFAWFP) resulted in a degradation rate almost identical to that of cp7-^{SF}GFP-βF, but with an affinity about 15-fold higher (K_M = 1.3±0.4 μM). The apparent high affinity indicates the strong association of cp7-^{SF}GFP-βF2 on hexYME1L, impeding its unfolding and translocation by the hexYME1L enzyme. As has been shown above, the position of β20 degron plays a role in substrate degradation by YME1L, it would be interesting to determine if this is the same situation for the finer recognition element, the crucial FAWFP pentapeptide. We swapped the position of β5 and βF, resulting in a β5F tag (APEAVFAWFP). This variant exhibited 3-fold lower degradation rate than β10 but close to that of βF. This result supports the crucial function of the FAWFP motif in substrate recognition by YME1L, Also, registering this central motif at the right position in the degron seems to be a contributing factor for substrate engagement.

3.6 hexYME1L is able to degrade endogenous substrates

In order to show that reengineered hexYME1L not only degrades model substrates, but also degrades physiological substrates, we carried out protein degradation using several identified endogenous substrates. Members from the Ups1/PRELI-like family have been shown to be degraded by the i-AAA proteases (120, 121). Specifically, the intermembrane space proteins, Ups1

and Ups2 in yeast participate in mitochondrial lipid regulation (182, 183). Both of the proteins are intrinsically disordered and degraded by Yme1. Another metalloprotease in yeast, Atp23, also degrades Ups1. However, binding of both proteins to Mdm35 results in their stabilization (120). The counterpart of the Ups1 and Ups2 proteins in human, PRELI, was reported to be accumulated in the presence of an ATPase deficient variant of YME1L. However, in cells containing wild type YME1L, rapid degradation was observed under stress conditions (121). The human homolog of Mdm35, TRIAP1, contains the conserved Cx_9C motifs (121).

As reported, expression of Ups1, Ups2 or PRELI alone did not yield soluble proteins (184, 185). However, when co-expressing Ups proteins with their binding partners (Up1/Mdm35, Ups2/Mdm35, PRELI/TRIAP1), only the Ups1/Mdm35 complex gives soluble protein in a non-aggregated state. We then tested the degradation of Ups1/Mdm35 complex using our *in vitro* system. Intriguingly, Ups1 was degraded even when complexed with Mdm35 (Figure 3.14a). Although there seemed to be some loss of Ups1 in the absence of ATP (Figure 3.14b), the degradation of Ups1 by hexYME1L with ATP is significantly faster (Figure 3.14c). The degradation of Ups1 by hexYME1L *in vitro* in the presence of Mdm35 contradicts with the stabilization effect on Ups1 by Mdm35 *in vivo* (120). However, this can be explained by the possibility that Ups1/Mdm35 complex undergoes an association/dissociation equilibrium, and the dissociated Ups1 from Mdm35 exposes the recognition sequence due to its intrinsic instability. hexYME1L is then able to engage Ups1 through the recognition motif in the Ups1 sequence. This association/dissociation equilibrium behavior can also explain the observed slow degradation rate of Ups1 by YME1L *in vitro*, which is inconsistent with the rapid turnover of Ups1 *in vivo* (120).

Other endogenous substrates have been examined are the two variants of TIM17 proteins, TIM17A and TIM17B, which are integral inner membrane proteins. Either of the two proteins can be incorporated into a TIM23 translocase complex, resulting in the formation of different TIM23 complexes (36). Structural analysis of TIM23 translocase suggests that c-terminal tails of the TIM17 proteins face the intermembrane space (186, 187). A sequence alignment of TIM17A and TIM17B reveals the highly conserved characteristic of these two proteins (Figure 3.15a). The main differences are found at the C-terminal tails with a size about 30 residues. Interestingly, TIM17A, but not TIM17B, was discovered to be degraded by YME1L under stress condition (130). Therefore, the intermembrane space exposed C-terminal tails of the TIM17 proteins may serve as degrons to allow degradation by the YME1L protease. We attached the C-terminal tails from

TIM17A (36 residues) and TIM17B (37 residues) to the C-terminus of the unstructured $I27^{CD}$ protein ($I27^{CD}$-17A and $I27^{CD}$-17B). In contrast to the *in vivo* results (130), both $I27^{CD}$-17A and $I27^{CD}$-17B were degraded in an ATP-dependent manner (Figure 3.15b-e). However, $I27^{CD}$-17A was degraded significantly faster than $I27^{CD}$-17B by hexYME1L (Figure 3.15f). These observations indicate that the C-terminal sequences from the TIM17 proteins are possible degrons for mediating substrate engagement. The detected faster degradation rate of TIM17A may indicate that YME1L selects substrate through accessible recognition sequences. Additional contributions from the surrounding environment *in vivo* may be involved for the substrate selection by YME1L, which may explain the distinct behaviors of TIMA and TIM17B in degradation *in vivo*.

Taken together, these results verify the feasibility of this *in vitro* system in the analysis of degradation of physiological substrates, which would benefit the characterization of degrons from endogenous proteins.

3.7 Conclusion

The human i-AAA protease, YME1L, is a mitochondrial inner membrane protein with a single transmembrane segment composed of highly hydrophobic residues. Together with its homolog, m-AAA protease, YME1L maintains the mitochondrial homeostasis in the inner membrane, as well as the intermembrane space and the matrix (49). Results presented in this chapter show that: 1, a *de novo* designed 32-residue sequence containing a heptapeptide repeats (155) can be used to successfully generate a homo-hexamerized YME1L, hexYME1L. hexYME1L is soluble and capable of hydrolyzing ATP and protein substrates. 2, Protein degradation by YME1L requires an accessible degron sequence with preferred sequence and a certain length. Simply unfolding a structured protein lacking a degron sequence does not result in its degradation by the YME1L protease. YME1L processively degrades a substrate from the degron terminus to the other. 3, YME1L displays unfolding power at somewhat significant level, as evidenced by the degradation of stable proteins that maintain their native structure at 37 °C. These substrates include λCI-N, and circularly permuted GFP variants. 4, The *in vitro* degradation assay system can be used for investigating degradation of physiological substrates.

The cc-hex peptide was employed in this study to substitute the transmembrane span of YME1L. The transmembrane spans of FtsH and m-AAA were shown to be important for the

oligomerization of the proteases and play roles in their proteolysis function (135, 188, 189). However, as shown in this study, hexYME1L is able to degrade protein substrates in solution. Yet, the hexamerized cc-hex in hexYME1L contains no homologous sequence as the transmembrane segment of YME1L. This suggests that the transmembrane segment of YME1L is not essential in substrate degradation, and that it may mainly function as a membrane-tethering element and drive oligomerization of YME1L *in vivo*, playing roles structurally. Whereas, the possibility remains that the transmembrane domain of YME1L participates in the degradation of endogenous proteins that are anchored in the inner membrane.

The observed positive cooperativity (Hill coefficient h = 2.4) of ATP hydrolysis suggests the communication between contiguous subunits resulted from inter-subunit interactions. In turn, this indicates the correct assembly of hexamer with appropriate formed interfaces between adjacent subunits. Although we cannot completely rule out the possibility that the artificial cc-hex may restrict the motion of hexYME1L, we suspect this concern is unlikely. Because a native YME1L sequence of a size about 20 amino acids between cc-hex and the core of folded AAA+ module of hexYME1L is predicted to be unstructured.

YME1L displays preference of degrons. Degrons can be as short as 5-residue long for proteins to be degraded *in vitro*, as shown in this study. We identified a critical pentapeptide motif (FAWFP) in the β20 degron. This motif is the primary element for inducing substrate engagement on hexYME1L, and its correct registration on the enzyme contributes greatly to the kinetics of degradation. Consistent with our results, Gur and colleagues reported that the identity and placement of a similar WRFAWFP motif are a determinant for substrate binding on Lon (94). Similarly, the hydrophobic patch was found in mDHFR (LPWPPL, position 23-38), which becomes accessible when mDHFR is unfolded. Considering that YME1L degrades unassembled subunits of inner membrane complexes and proteins in the absence of binding partners in the intermembrane space, which are reminiscent of behavior of substrate degradation by Lon, it is reasonable to speculate that this FAWFP motif may represent a class of degron for mitochondrial protein degradation. However, a search of proteins in the intermembrane space of mitochondria using the FAWFP motif did not result in any hits. Interestingly, when searching using the motif in a degenerated form (F-h-h-F, h stands for hydrophobic), 21 human proteins in the intermembrane space were identified (Table 3.2). Within the list, two substrates of YME1L confirmed previously are found: the inner membrane protease OMA1 (FVVF, position 206-209) (131), and the

Ups1/PRELI-like family member in human PRELI (FAAF, position 19-22) (121). The stabilization effect of Mdm35 on Ups1 in yeast is largely mediated by the hydrophobic interactions (185, 190), preventing Ups1 from being recognized by Yme1. However, in the *in vitro* assays in this study, degradation of Ups1 may be explained by the slow dissociation of Ups1 from Mdm35 in the association-dissociation equilibrium, revealing the yet to be identified recognition signal in Ups1. For YME1L/Yme1 substrates that do not have FAWFP or F-h-h-F motifs, it is highly likely that these proteins contain different degrons or recognition sequences.

Although YME1L exhibits weaker activities than the robust AAA+ proteases ClpXP and Lon, as suggested by the degradation of cp7-^{SF}GFP proteins, YME1L is able to unfold these proteins with substantial stabilities. The processivity ratio about 6 determined in the case of mDHFR-I27-β20 degradation suggests the preference of mDHFR unfolding over substrate release from the enzyme in the absence of MTX. In combination with the observed stimulation of ATPase activity, it is suggested that YME1L degrades substrates in a processive mode.

In summary, the engineered hexYME1L enables *in vitro* biochemical characterization of this important transmembrane protease for the first time. By using this system, we determined some biochemical parameters of YME1L and unraveled the substrate recognition mechanism by YME1L. We expect a broader application of this strategy to other AAA+ proteases that are challenged by the similar issue of an unavailable *in vitro* system.

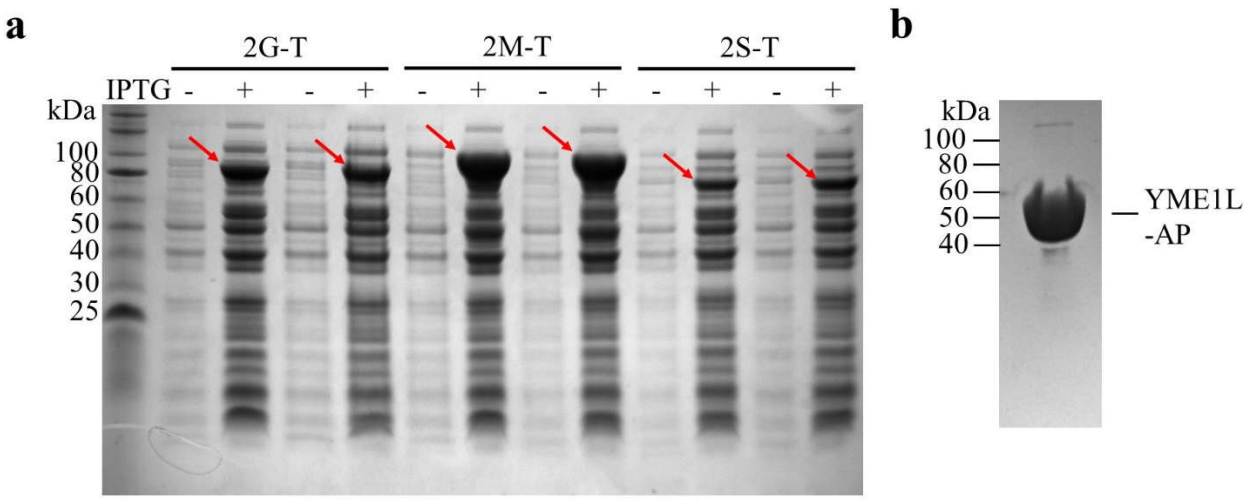


Figure 3.1 Expression and purification of YME1L-AP.

a. SDS-PAGE gels showing expression of YMEL1-AP proteins with different tags fused at the N-terminus in *E. coli* BL21 (DE3), respectively. Coding sequences of YME1L-AP were inserted into 2G-T, 2M-T, or 2S-T vectors, resulting in fusion proteins carrying an N-terminal His_6-GST, His_6-MBP, or His_6-SUMO tag, correspondingly. Recombinant proteins were overexpressed (red arrows) upon addition of IPTG at a final concentration of 1 mM. Theoretical molecular weights of different fusion proteins are: 79.0 kDa (His_6-GST-YME1L-AP), 94.1 kDa (His_6-MBP-YME1L-AP), and 70.0 kDa (His_6-SUMO-YME1L-AP) **b.** Final purity of YME1L-AP (theoretical molecular weight 51.0 kDa without N-terminal fusion tag) analyzed by SDS-PAGE.

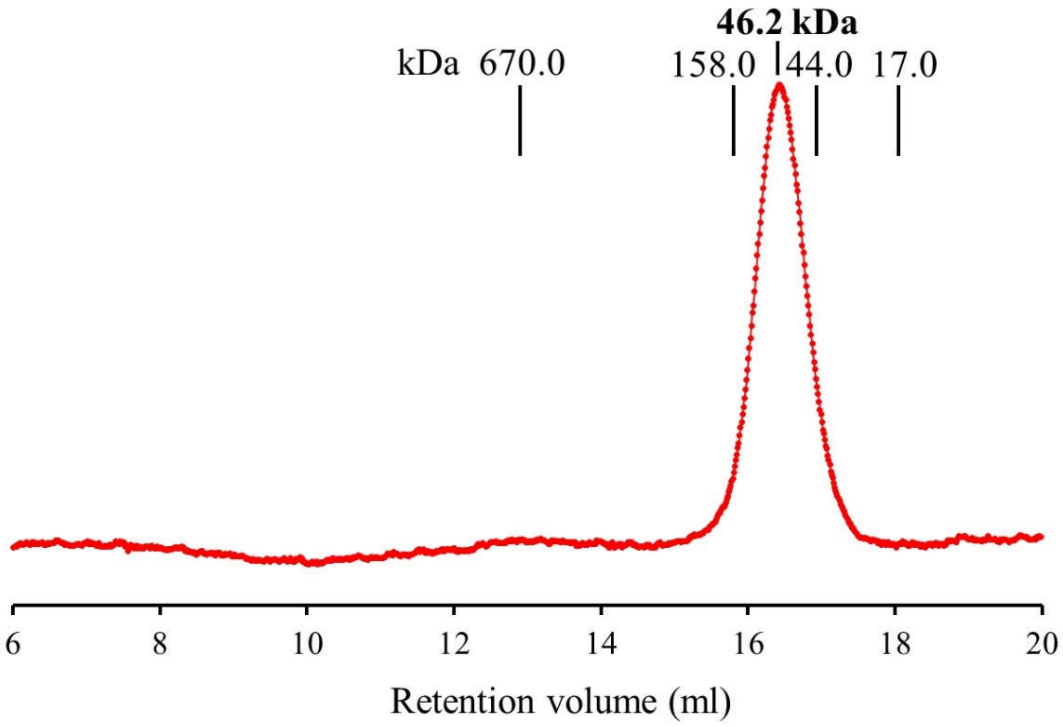


Figure 3.2 Size exclusion profile of YME1L-AP.

YME1L-AP was eluted at a retention volume of 16.43 ml on a Superose 6 Increase 10/300 GL column, corresponding to an apparent molecular weight of 46.2 kDa (theoretical molecular weight 51.0 kDa). Elution positions of molecular standards of known molecular weight are indicated on the profile.

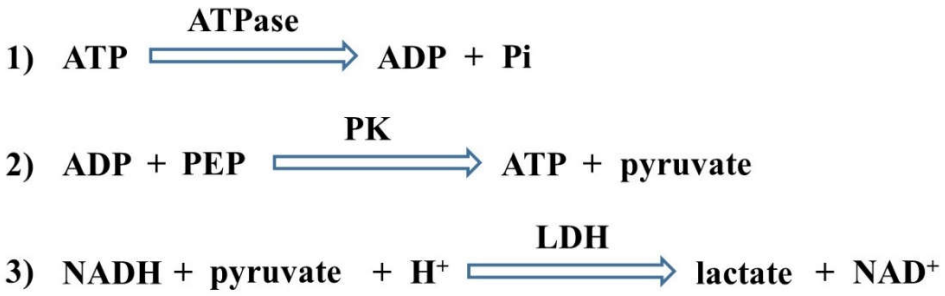


Figure 3.3 Coupled enzymatic reactions in ATPase assay.

Figure showing three coupled enzymatic reactions in the ATPase activity test. Firstly, ATP is hydrolyzed by the enzyme to be assayed, producing ADP and inorganic phosphate. Then ATP molecules are regenerated from the ADP produced at step 1 by an ATP regeneration system comprising PK and PEP, supplementing the ATP consumed by ATPases and keeping the ATP concentration to be constant. At the final step, pyruvate from step 2 is converted to lactate, catalyzed by LDH. At the same time, NAD^+ is formed resulted from the oxidation of NADH. NADH has an absorbance signal at 340 nm, whereas the oxidized NAD^+ does not absorb at this wavelength. Thus, A_{340nm} signal is monitored to follow the consumption rate of NADH, which can be directly correlated with the ATP hydrolysis rate.

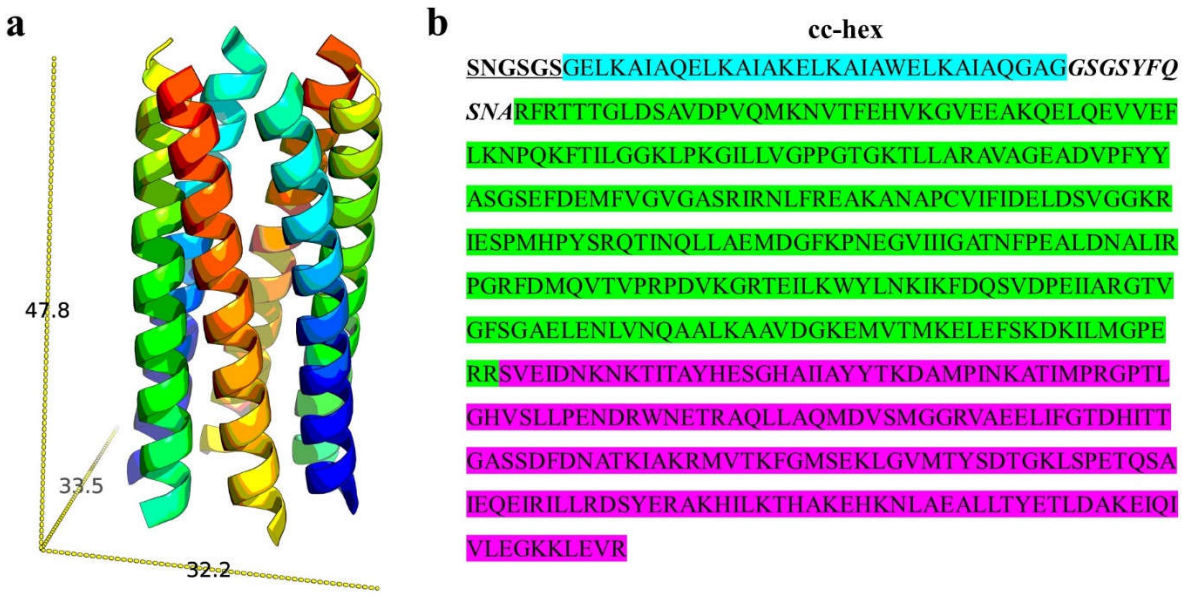


Figure 3.4 Constructing the hexameric hexYME1L.

a. Dimensions of the cc-hex hexamer (PDB code: 3R3K) measured using Draw_Protein_Dimensions.py script implemented PyMOL. Peptides are represented in a color spectrum. **b.** Full protein sequence of hexYME1L. The ATPase module of hexYME1L is highlighted in green and peptidase domain is highlighted in pink. Amino acids of cc-hex are highlighted in cyan. The linker between cc-hex and catalytic domains of YME1L is bolded and italicized. The six N-terminal residues in bold and underlined are generated by TEV digestion.

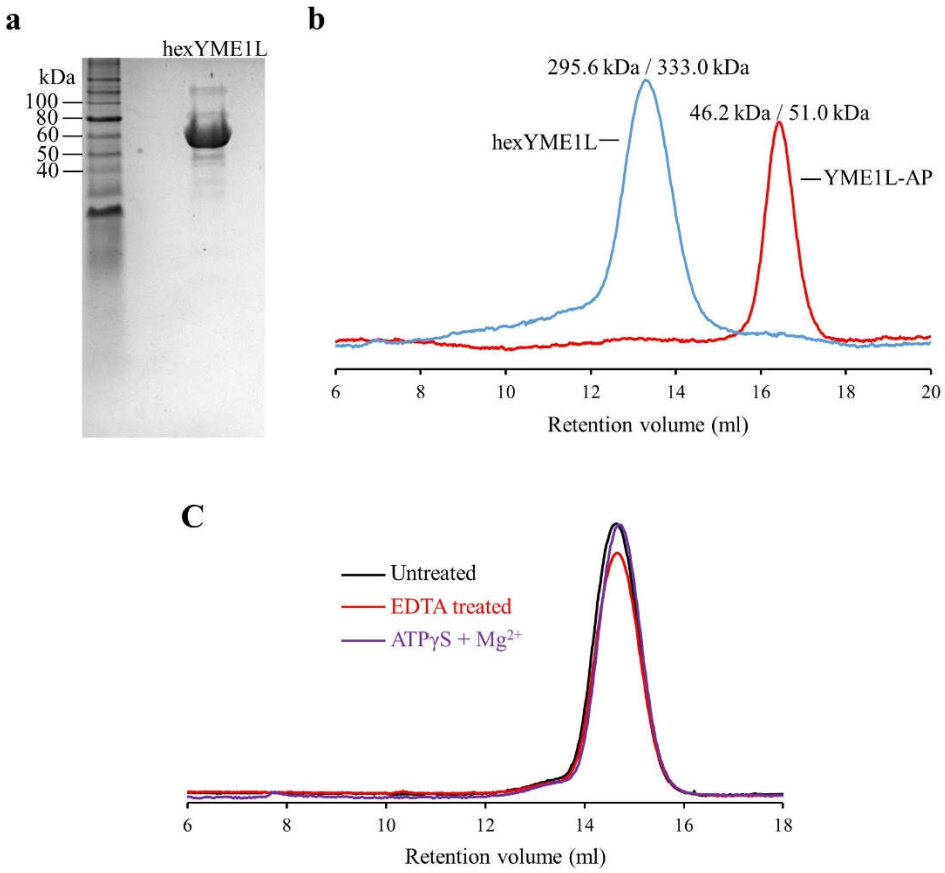


Figure 3.5 hexYME1L is a hexamer in solution.

a. Final purity of hexYME1L (theoretical monomer molecular weight 55.5 kDa without N-terminal fusion tag) analyzed by SDS-PAGE. **b.** Overlay of size exclusion profiles of YME1L-AP and hexYME1L on a Superose 6 Increase 10/300 GL column. Values of apparent Mw / theoretical Mw are indicated for both of the proteins. **c.** Overlay of size exclusion profiles of hexYME1L on a Superose 6 Increase 10/300 GL column, with no treatment (black), after one hour of 5 mM EDTA incubation (red), or one more hour treatment after addition of 10 mM ATPγS and $MgCl_2$ (purple).

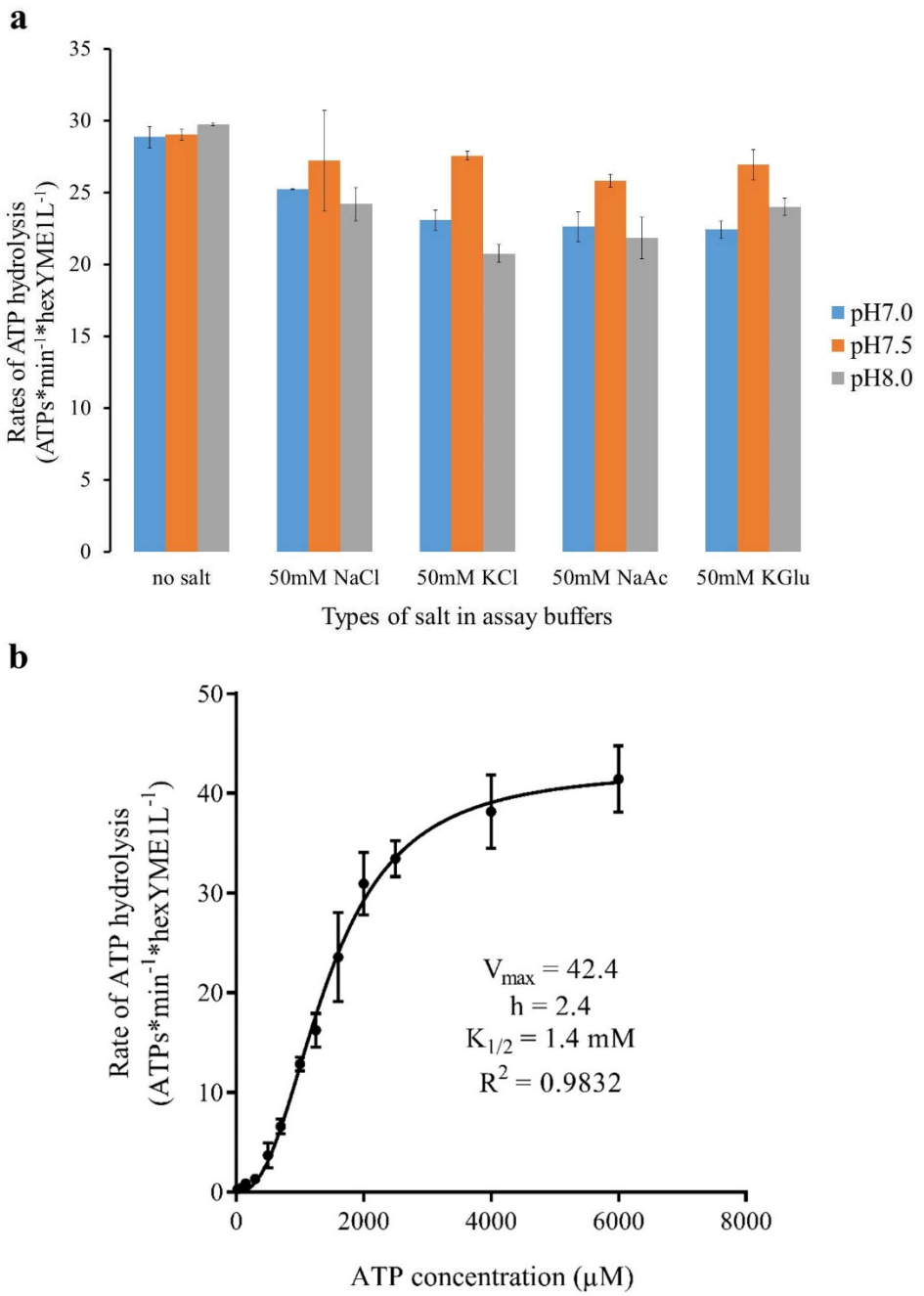


Figure 3.6 ATPase activity of hexYME1L.

a. Bar graph showing buffer optimization for ATPase assay of hexYME1L (1μM). Different salt types and concentrations were tested. An array of pH values were analyzed, in combination with variation of salt. Results from conditions with pH values higher than 8.0 or lower than 7.0 are not shown here. **b.** Rate of ATP hydrolysis by hexYME1L in the presence of increasing ATP concentration. Data were fitted well to Allosteric sigmoidal model ($R^2 = 0.9832$) with the formula: rate = Vmax / (($K_{1/2}$/[ATP])^h + 1), using GraphPad Prism version 7.02 for Windows, GraphPad Software, La Jolla California USA, "www.graphpad.com". Parameters of the ATPase activity of hexYME1L are Vmax = 42.4 ATPs * min^{-1} * hexYME1L^{-1}, $K_{1/2}$ = 1.4 mM, hill coefficient h = 2.4. Data are from independent experiments and represented as mean ± S.E.M. (n = 3)

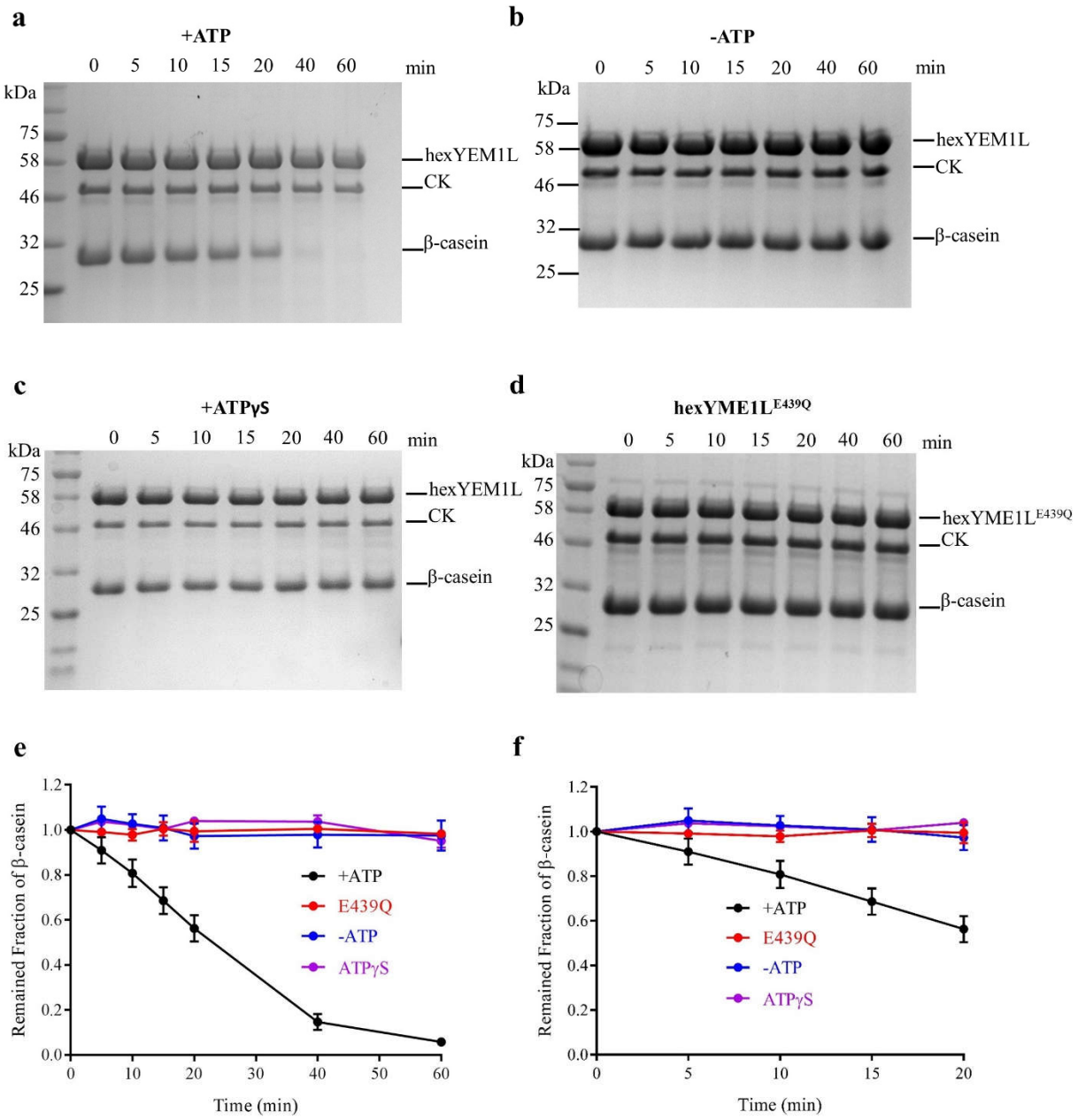


Figure 3.7 Degradation of β-casein by hexYME1L.

a-d. SDS-PAGE gels showing degradation of 20 μM β-casein by 1 μM hexYME1L. **a.** in the presence of ATP. **b.** in the absence of ATP. **c.** in the presence of ATPγS. **d.** hexYME1LE439Q with ATP supplemented. **e.** Quantification of the remaining β-casein for **a** (black), **b** (blue), **c** (purple) and **d** (red). **f.** Zoomed in presentation of the first 20 minutes of **e**. Quantification was normalized to the intensity of CK (0.1 mg/ml in the reaction) for each degradation reaction, respectively. The degradation rate of β-casein by hexYME1L in the presence of ATP (black) (0.49±0.08 molecules * min^{-1} * $hexYME1L^{-1}$) calculated from **f**. Data are from independent experiments and represented as mean ± S.E.M. (n = 3)

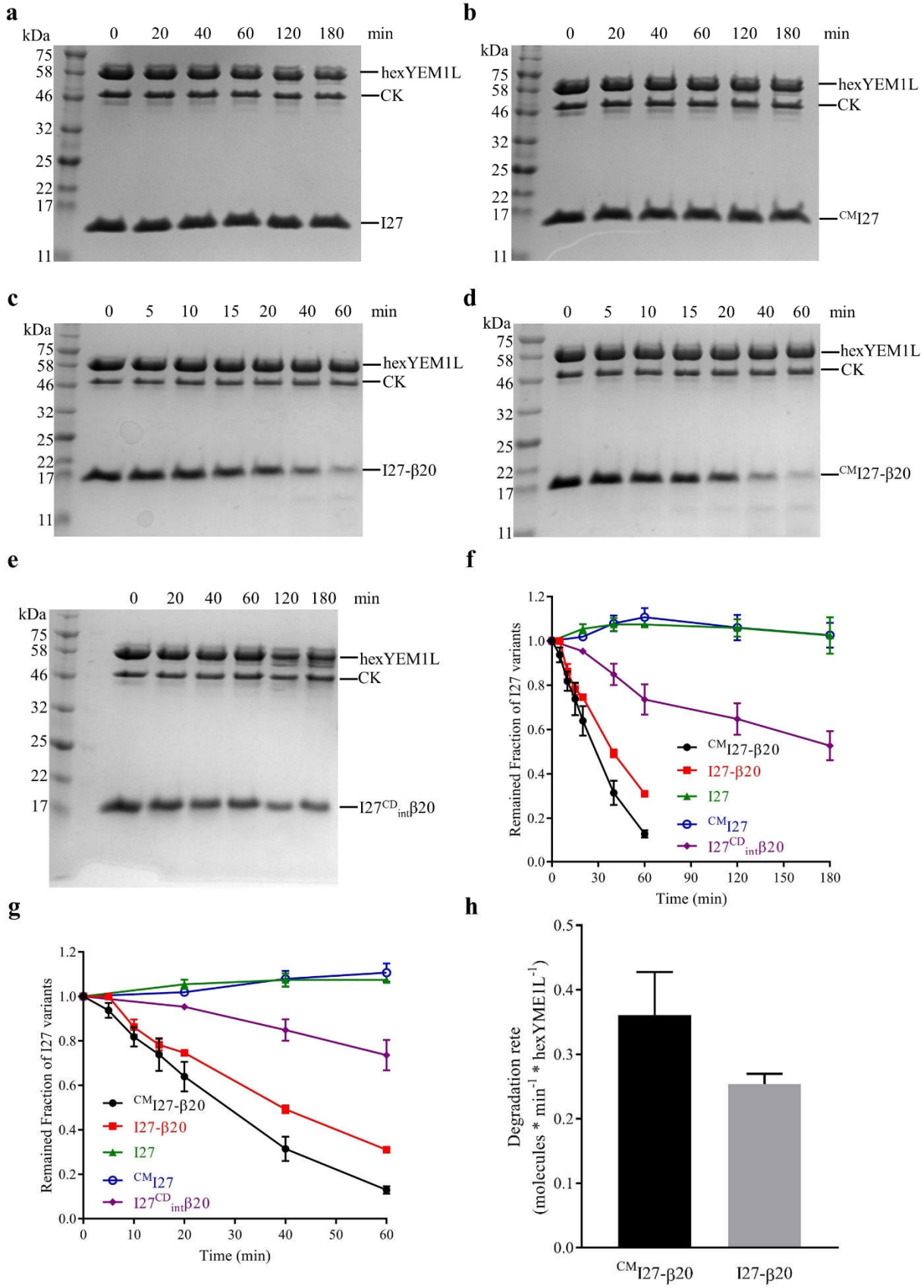

a
kDa
0 20 40 60 120 180 min
75
58
46
32
25
22
17
11
hexYEM1L
CK
I27
b
kDa
0 20 40 60 120 180 min
hexYEM1L
CK
CMI27
c
kDa
0 5 10 15 20 40 60 min
hexYEM1L
CK
I27-β20
d
kDa
0 5 10 15 20 40 60 min
hexYEM1L
CK
CMI27-β20
e
kDa
0 20 40 60 120 180 min
hexYEM1L
CK
I27CD intβ20
f
Remained Fraction of I27 variants
Time (min)
CMI27-β20
I27-β20
I27
CMI27
I27CD intβ20
g
Remained Fraction of I27 variants
Time (min)
CMI27-β20
I27-β20
I27
CMI27
I27CD intβ20
h
Degradation rete
(molecules * min-1 * hexYME1L-1)
CMI27-β20
I27-β20

Figure 3.8 Degradation of I27 variants by hexYME1L.

a-e. SDS-PAGE gels showing degradation of 20 μM I27 variants by 1 μM hexYME1L. **a.** I27 in the presence of ATP. **b.** ^{CM}I27 in the presence of ATP. **c.** I27-β20 in the presence of ATP. **d.** ^{CM}I27-β20 in the presence of ATP. **e.** I27$^{CD}_{int}$β20 in the presence of ATP. **f.** Quantification of the remaining I27 variants for **a** (green), **b** (blue), **c** (red), **d** (black), and **e** (purple). **g.** Zoomed in presentation of the first 60 minutes from **f**. Quantification was normalized to the intensity of CK (0.1 mg/ml in the reaction) for each degradation reaction, respectively. **h.** Degradation rates of I27-β20 and ^{CM}I27-β20 of calculated from **g**. Data are from independent experiments and represented as mean ± S.E.M. (n = 3)

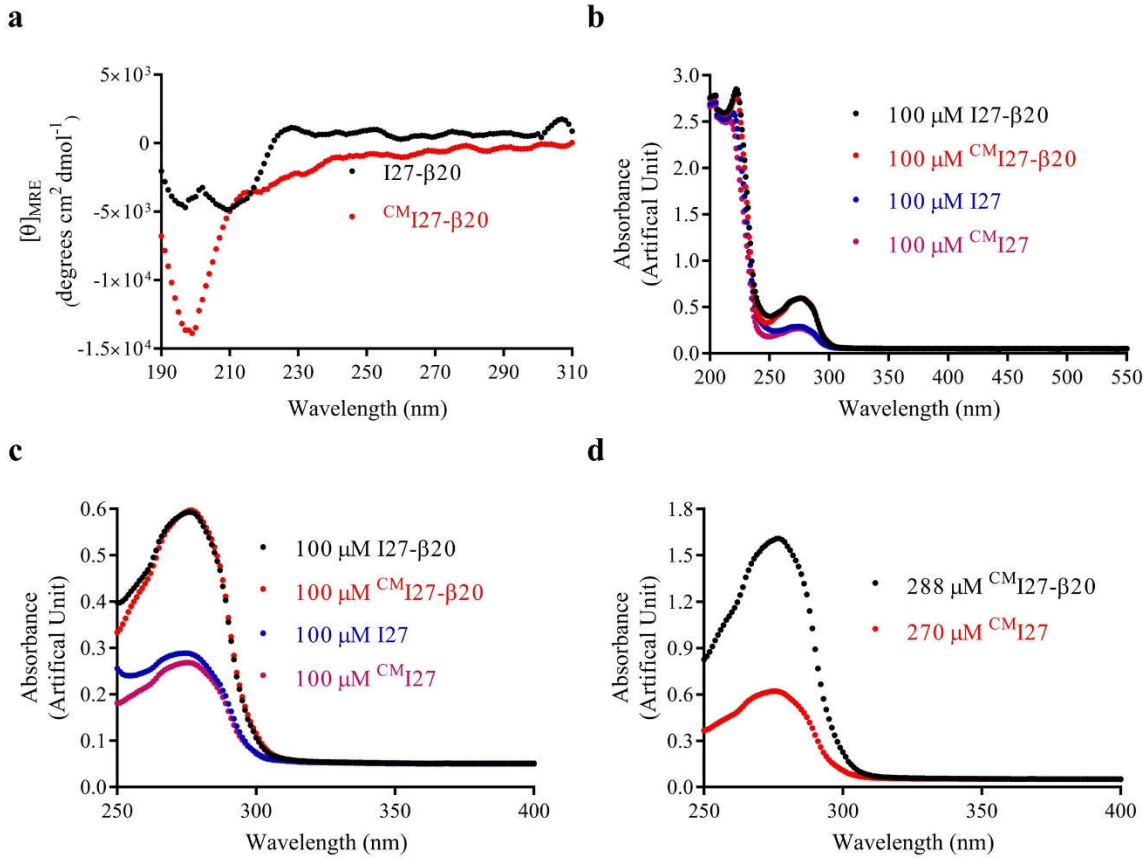


Figure 3.9 Circular dichroism and UV-VIS spectra of I27 variants.

a. Circular dichroism spectra of I27-β20 (black) and carboxymethylated ^{CM}I27-β20 **b.** UV-VIS light absorbance spectra (200 – 550 nm) of 100 μM I27 variants: I27-β20 (black), ^{CM}I27-β20 (red), I27 (blue), and ^{CM}I27 (purple). **c.** Zoomed in presentation of 250 – 400 nm light absorbance from **b**. **d.** Light absorbance spectra (250 – 400 nm) of I27 variants: 288 μM ^{CM}I27-β20 (black) and 270 μM ^{CM}I27 (red).

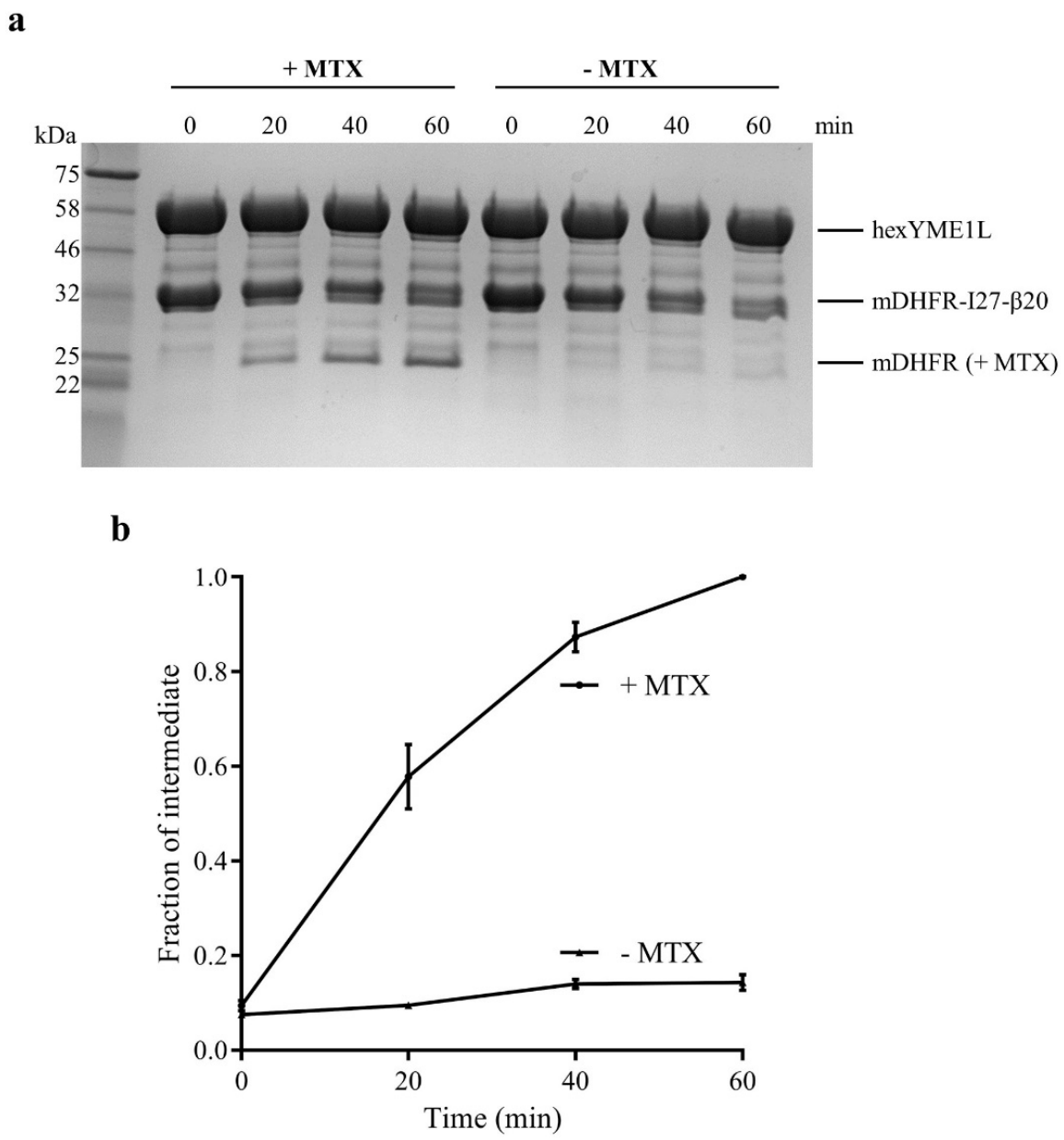


Figure 3.10 Degradation of mDHFR-I27-β20 by hexYME1L.

a. SDS-PAGE gels showing degradation of 10 μM mDHFR-I27-β20 by 1 μM hexYME1L in the presence or absence of MTX. mDHFR (+ MTX) indicates the MTX-stabilized mDHFR intermediate. **b.** Quantification of the intermediate appearing in the degradation reaction from **a**. Quantification was normalized to the intensity of hexYME1L for each degradation reaction, respectively. Data are from independent experiments and represented as mean ± S.E.M. (n = 3)

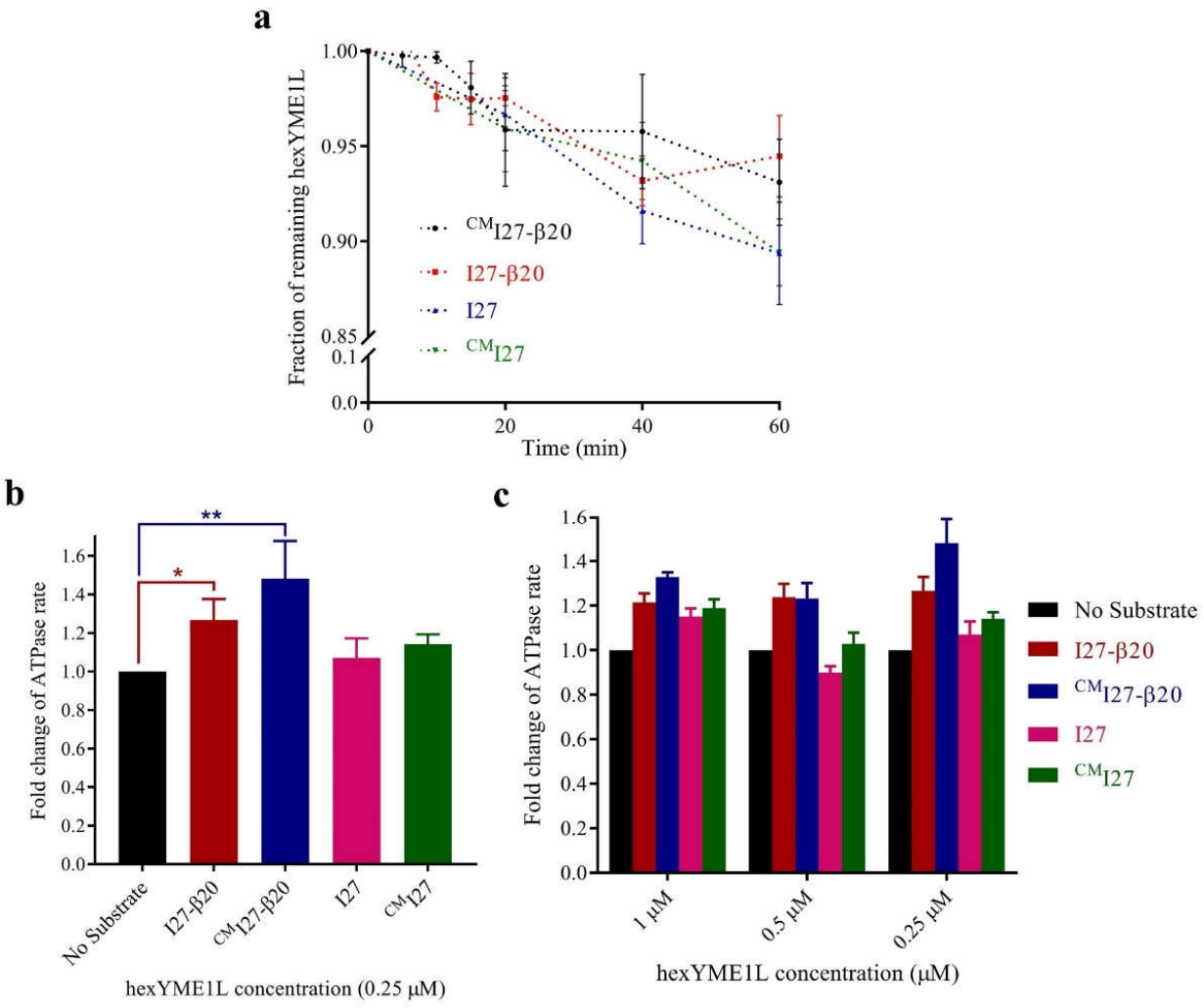


Figure 3.11 Stimulation of ATPase activity of hexYME1L.

a. Loss of hexYME1L over time in degradation reactions in the presence of different substrates: I27 (blue), ^{CM}I27 (green), I27-β20 (red), ^{CM}I27-β20 (black) **b.** ATPase rate of hexYME1L (0.25 μM) in the absence of substrate (black) and in the presence of 20 μM substrates, I27 (pink), ^{CM}I27 (green), I27-β20 (red), ^{CM}I27-β20 (blue). **c.** ATPase rate of hexYME1L at different concentrations (1 μM, 0.5 μM, and 0.25 μM) in the absence of substrate (black) and in the presence of 20 μM substrates, I27 (pink), ^{CM}I27 (green), I27-β20 (red), ^{CM}I27-β20 (blue). Non-paired one-way ANOVA analysis was carry out for **b**, comparing each group with substrates present to the "No substrate" control group using GraphPad Prism version 7.02 for Windows, GraphPad Software, La Jolla California USA, "www.graphpad.com". *P ≤ 0.05, **P ≤ 0.001. Data are from independent experiments and represented as mean ± S.E.M. (n = 3).

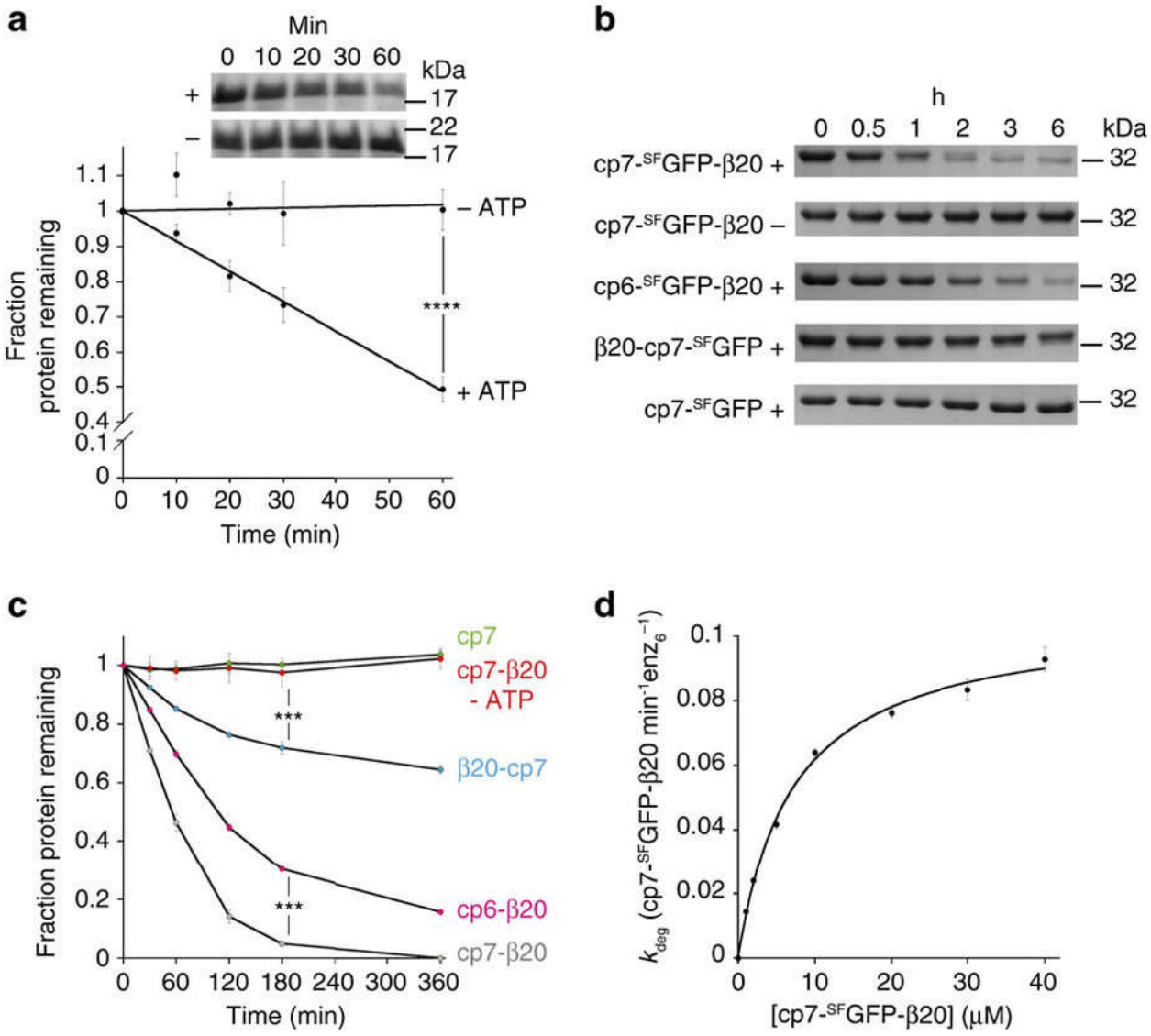


Figure 3.12 Unfolding and degradation of stable proteins by hexYME1L.

a. SDS–PAGE gels showing degradation of β20-λcI-N (20 μM) by hexYME1L (1 μM) in the presence (+) and absence (-) of ATP and a plot of the loss of β20-λcI-N over time. **b.** SDS–PAGE showing degradation of circularly-permuted variants of GFP (20 μM) by hexYME1L (1 μM) in the presence (+) and absence (-) of ATP. **c.** Loss of GFP variants over time from **b**. **d.** Rate of cp7-^{SF}GFP- β20 degradation against increasing substrate concentration displaying a nonlinear least-squares fitting to the Michaelis–Menten equation (k_{deg} = 0.11 molecules * min^{-1} * hexYME1L^{-1}; K_M = 7.1 μM). **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001 calculated from an unpaired Student's t-test (two-tailed). Data are from independent experiments and represented as mean ± S.E.M. (n = 3). (Figure adapted from the published paper: **Shi H**, Rampello AJ, Glynn SE. Engineered AAA+ proteases reveal principles of proteolysis at the mitochondrial inner membrane. Nature Communications 7: 13301. doi:10.1038/ncomms13301) (Data from Figure b-d contributed by **Anthony Rampello**)

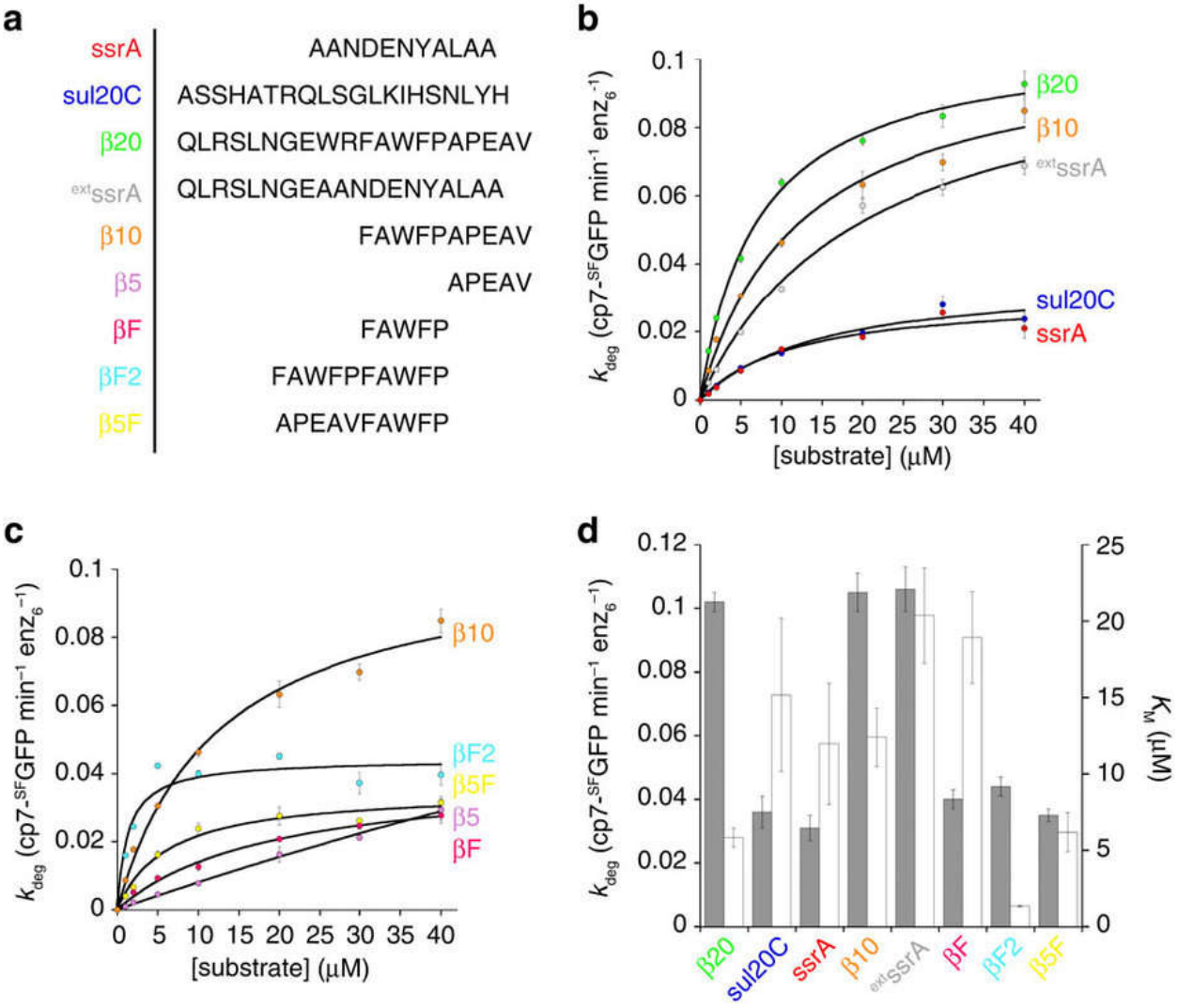


Figure 3.13 YME1L discriminates between degrons by sequence.

a. Sequences of each degron fused to cp7-SFGFP and tested for ATP-dependent degradation by hexYME1L. **b.** Michaelis–Menten plots showing degradation of cp7-^{SF}GFP variants bearing C-terminal fusions of degrons previously shown to enhance degradation by AAA+ proteases (β20 (green) and sul20C (blue), Lon protease; ssrA (red), ClpXP protease; β10 (orange); ^{ext}ssrA (grey)). Lines are nonlinear least-squares fits to the Michealis–Menten equation. **c.** Michaelis–Menten plots of cp7-^{SF}GFP proteins bearing C-terminal fusions of multiple truncated variants of the β20 degron (β10 (orange); β5, (pink); βF (magenta); βF2 (cyan); β5F (yellow)). Lines are nonlinear least-squares fits to the Michealis–Menten equation. **d**. Maximal degradation rates (black columns) and K_M values (white columns) for each degron sequence fused to the C-terminus of cp7-^{SF}GFP. All degradation reactions contained hexYME1L (1 μM). Data are from independent experiments and represented as mean ± S.E.M. (n = 3). (Figure adapted from the published paper: **Shi H**, Rampello AJ, Glynn SE. Engineered AAA+ proteases reveal principles of proteolysis at the mitochondrial inner membrane. Nature Communications 7: 13301. doi:10.1038/ncomms13301) (Data contributed by **Anthony Rampello**)

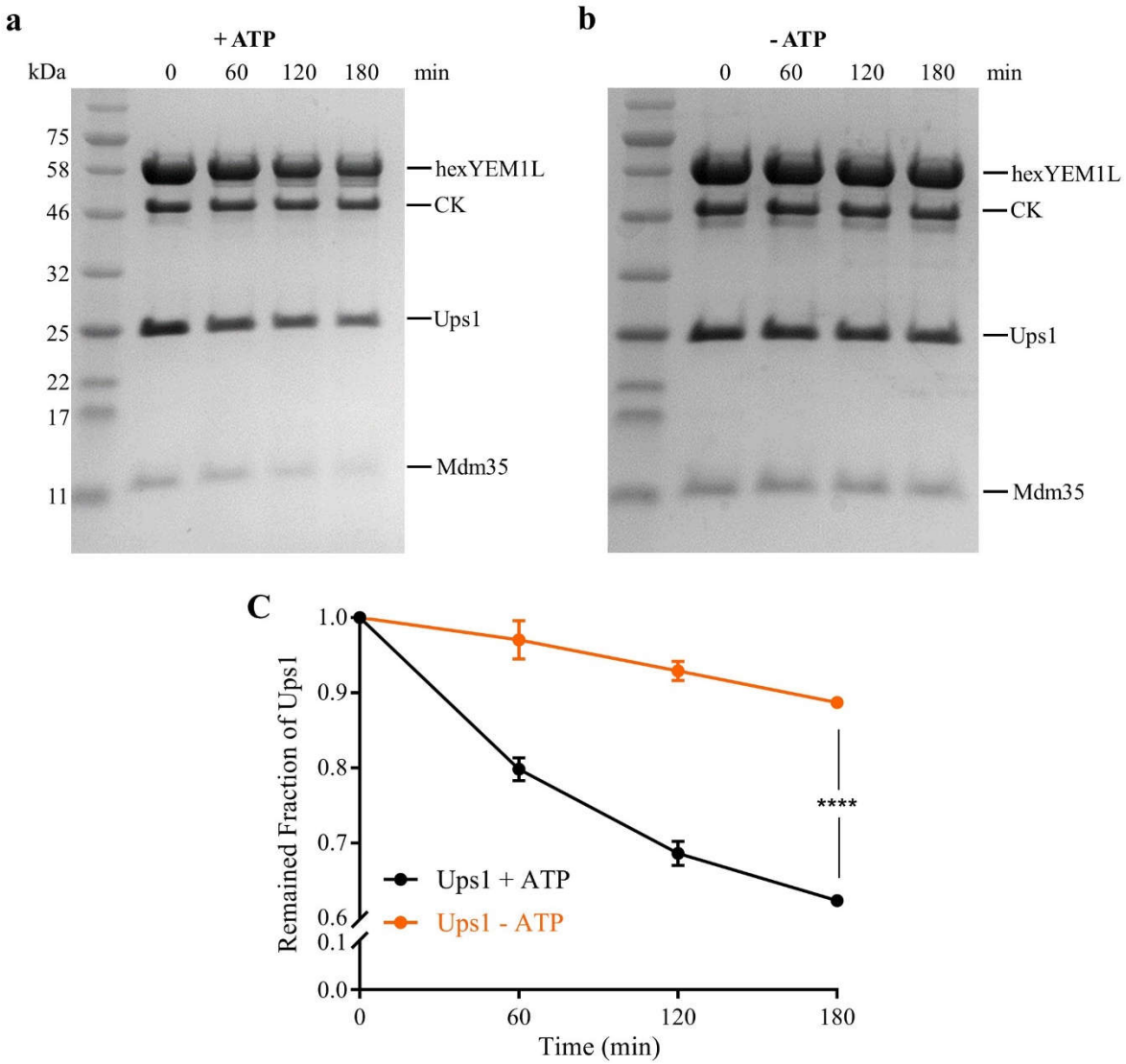


Figure 3.14 Degradation of Ups1/Mdm35 complex by hexYME1L.

a. SDS-PAGE gels showing degradation of 10 µM Ups1/Mdm35 complex by 1 µM hexYME1L in the presence of MTX. **b.** SDS-PAGE gels showing degradation of 10 µM Ups1/Mdm35 complex by 1 µM hexYME1L in the absence of MTX. **c.** Quantification of remaining Ups1 over time in the presence of ATP (black) from **a**, in the absence of ATP (orange) from **b**. Quantification was normalized to the intensity of CK (0.1 mg/ml in the reaction) for each degradation reaction, respectively. ****P ≤ 0.0001 calculated from an unpaired Student's t-test (two-tailed). Data are from independent experiments and represented as mean ± S.E.M. (n = 3).

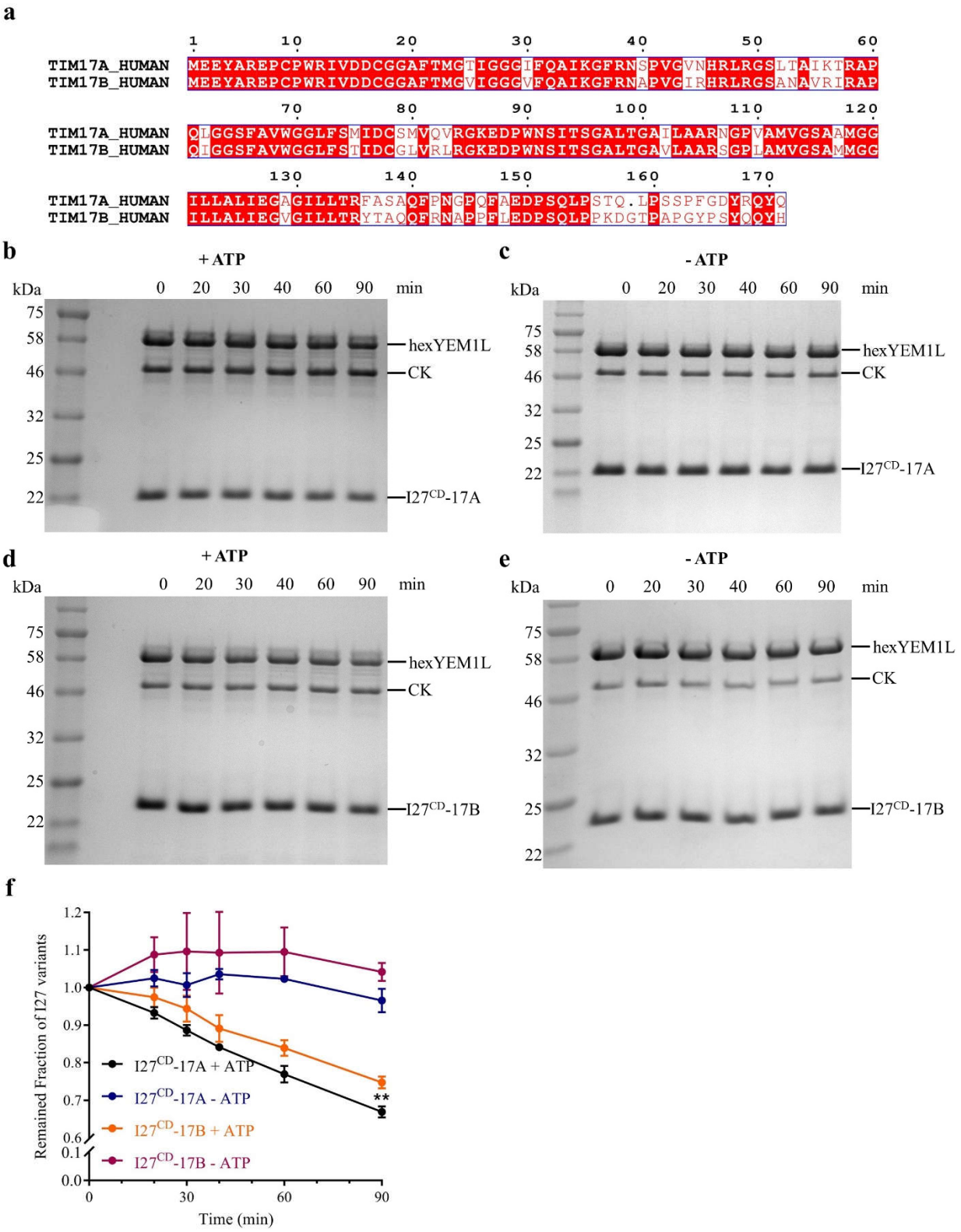

a
TIM17A_HUMAN
TIM17B_HUMAN
b
+ATP
kDa
0 20 30 40 60 90 min
75
58
46
32
25
22
hexYEM1L
CK
I27CD-17A
c
- ATP
d
+ATP
I27CD-17B
e
- ATP
f
Remained Fraction of I27 variants
I27CD-17A + ATP
I27CD-17A - ATP
I27CD-17B + ATP
I27CD-17B - ATP
**
Time (min)

Figure 3.15 Degradation of $I27^{CD}$-17A and $I27^{CD}$-17B by hexYME1L.

a. Sequence alignment of TIM17A and TIM17B, with **b-e.** SDS-PAGE gels showing degradation of 20 μM substrates by 1 μM hexYME1L. **b**. $I27^{CD}$-17A in the presence of ATP. **c**. $I27^{CD}$-17A in the absence of ATP. **d**. $I27^{CD}$-17B in the presence of ATP. **e**. $I27^{CD}$-17B in the absence of ATP. **f.** Quantification of remaining substrates over time from **b-e**, $I27^{CD}$-17A in the presence of ATP (black), $I27^{CD}$-17A in the absence of ATP (blue), $I27^{CD}$-17B in the presence of ATP (orange), $I27^{CD}$-17B in the absence of ATP (purple). Quantification was normalized to the intensity of CK (0.1 mg/ml in the reaction) for each degradation reaction, respectively. $^{**}P \leq 0.01$ calculated from an unpaired Student's t-test (two-tailed). Data are from independent experiments and represented as mean ± S.E.M. (n = 3).

Table 3.1 Degradation of GFP proteins bearing different degron sequences.

ND stands for not determined (Table adapted from the published paper: **Shi H**, Rampello AJ, Glynn SE. Engineered AAA+ proteases reveal principles of proteolysis at the mitochondrial inner membrane. Nature Communications 7: 13301. doi:10.1038/ncomms13301) (Data contributed by **Anthony Rampello**)

Degron	Number of residues	Sequence	k_{deg} (min^{-1}*hexYME1L^{-1})	K_M (μM)
β20	20	QLRSLNGEWRFAWFPAPEAV	0.110±0.003	7.1±0.6
ssrA	11	AANDENYALAA	0.031±0.004	12.0±4.0
sul20C	20	ASSHATRQLSGLKIHSNLYH	0.036±0.005	15.2±5.0
β10	10	FAWFPAPEAV	0.105±0.015	12.4±1.9
^{ext}ssrA	19	QLRSLNGEAANDENYALAA	0.106±0.007	20.4±3.1
β5	5	APEAV	ND	ND
βF	5	FAWFP	0.040±0.003	18.9±3.0
βF2	10	FAWFPFAWFP	0.044±0.003	1.3±0.4
β5F	10	APEAVFAWFP	0.035±0.002	6.1±0.6

Table 3.2 F-h-h-F motifs identified in IMS proteins.

(Table adapted from the published paper: **Shi H**, Rampello AJ, Glynn SE. Engineered AAA+ proteases reveal principles of proteolysis at the mitochondrial inner membrane. Nature Communications 7: 13301. doi:10.1038/ncomms13301) (Data contributed by **Anthony Rampello**)

Uniprot ID	Name	Sequence	Location
P13073	COX41	FIGF	110-113
P36551	HEM6	FGLF	405-408
Q8TB36	GDAP1	FLGF FMLF	260-263 337-340
Q96E52	OMA1	FVVF	206-209
O95831	AIFM1	FGGF	418-421
P00505	AATM	FAFF	239-242
Q9H078	CLPB	FLPF	568-571
Q9NRV9	HEBP1	FAVF	84-87
O43676	NDUB3	FAAF	75-78
Q6NUK1	SCMC1	FGGF	235-238
O43674	NDUB5	FGGF	29-32
O95169	NDUB8	FLAF FMIF	136-139 139-142
P00395	COX1	FWFF FLGF FPLF	235-238 282-285 397-400
Q9BPX6	MICU1	FALF	417-420
P03915	NU5M	FAGF	463-466
P03905	NU4M	FYIF	118-121
O14880	MGST3	FLFF	71-74
O14949	QCR8	FVVF	53-56
Q9Y255	PRELI	FAAF	19-22
Q14257	RCN2	FIAF	174-177
O95202	LETM1	FLVF	212-215

Chapter 4 – Structural basis for the AAA+ module of i-AAA from *Myceliophthora thermophila*

In order to elucidate the mechanism of i-AAA, we carried out crystallographic studies of i-AAA proteases from human (YME1L) (UniProt ID: Q96TA2) and *Myceliophthora thermophila* (IMT) (UniProt ID: G2QPI5). We sub-cloned different constructs spanning different regions of the proteins (Table 2.1 and Table 2.2) and screened against several commercially available crystallization kits. Crystallization conditions from initials hits were optimized to obtain diffraction-quality crystals. Two constructs, IMT-9 (residues: 241-483, referred as IMT-AAA hereafter) and YME1L-11(residues: 331-585, referred as YME1L-AAA hereafter), yielded promising crystals (Figure 4.1). However, crystals of YME1L-11 diffracted poorly. Data sets at high resolution were only obtained from crystals of IMT-AAA.

4.1 Overall structure of IMT-AAA

The crystal structure of IMT-AAA containing only the AAA+ module was determined to a resolution of 2.45 Å (Table 4.1), with dihedral angles of 97.47% of total amino acids falling in the most favoured region and that of additional 2.53% in the allowed region. In the final model, two monomers are present in one asymmetric unit (Figure 4.2a), each containing one ADP molecule. The structure contains residues 241-483 (residue numbering according to UniProt ID: G2QPI5) for both of the monomers, with extra N-terminal Ser-Asp-Ala amino acids in chain A and Asp-Ala in chain B, resulting from TEV enzyme digestion. In chain A, amino acids Gly273 and Tyr316 are missing. Residues 352-357 in a loop region are absent in chain B. Each IMT-AAA monomer is comprised of two canonical subdomains of AAA+ module, a large N-terminal α/β domain and a small α-helical bundle at the C-terminus. These two monomers can be superimposed on to each other with subtle conformational differences. Superimposition on the backbones of the large subdomain (residues: 248 - 406) results in an RMSD value of 0.515 Å and a rotation of the small domain (residues: 407 - 483) about 7.3° (Figure 4.2b).

4.2 Conserved features and the ADP-binding site

IMT-AAA shares sequences of high homology with other members from the FtsH family (Figure 4.3). The N-terminal large subdomain is comprised of a parallel β-sheet of five β-strands, with α-helices flanking on both sides (Figure 4.3 and Figure 4.4). Two conserved Walker motifs (Walker A and Walker B) and key residues for nucleotide binding and hydrolysis are found in the large subdomain (Figure 4.4). The lysine residue (Lys289 in IMT-AAA) in the Walker A motif is important for ATP binding, mutation of which abolishes the ATP-binding ability of AAA+ proteins (63, 64). Within the Walker B motif, a highly conserved glutamate residue (Glu343 in IMT-AAA) is the catalytically active residue for ATP hydrolysis. A conserved aromatic residue (Tyr316) present in the pore loop mediates unfolding and translocation of protein substrates. The characteristic conserved secondary region of homology (SRH) motif of classical AAA proteins, contains a polar residue (Asn386 in IMT-AAA) termed sensor 1. This sensor residue is involved in coordinating ATP hydrolysis (64, 65). In addition, two arginine residues located at the C-terminus of SRH (Arg397 and Arg400 in IMT-AAA), which are referred as arginine fingers, are thought to mediate ATP hydrolysis between adjacent subunits within the hexameric ATPase ring (63, 64).

Interactions between the bound ADP molecule and IMT-AAA largely involve the β-phosphate group of ADP (Figure 4.5) and the Walker A motif. The sidechain of the Walker A lysine interacts with the oxygen on the β-phosphate group. In the Walker A motif, Gly286, Gly288, Lys289 and Thr290 form hydrogen-bonds with oxygens on the β-phosphate group through their backbone nitrogen atoms, respectively. The hydroxyl side chain of Thr290 forms hydrogen bond with oxygen atoms on the α-phosphate group and β-phosphate group of ADP, respectively. Asp342 interacts with the β-phosphate group of ADP mediated by a water molecule. Leu291 hydrogen bonds with the α-phosphate group using its backbone nitrogen. His246 and His422 interact with the adenine portion of ADP via atoms on the main chain and side chain, respectively. A water molecule mediates the interaction between adenine and the nitrogen atom on Gly446. The base of ADP sits in a hydrophobic pocket formed by Val 245, Leu291, and Ile 418.

4.3 Hexameric model of IMT-AAA

The structure of IMT-AAA was subjected to the online DALI server (191) to search for similar proteins based on structural information. Most hits of smallest RMSD value are those FtsH structures (2CE7, 2CEA, RMSD 1.4 - 1.6 Å; 4WWO, 2R65, RMSD 1.6 - 2.0 Å). Some other structures, 3H4M (proteasome-activating nucleotidase) and 5C1B (transitional endoplasmic reticulum ATPase) are also found to have small RMSD values ($RMSD_{3H4M}$ = 1.6 Å, $RMSD_{5C1B}$ = 1.9 Å). A hexameric model was then built by superposition of IMT-AAA (chain B) onto FtsH (2CE7) based on the structural alignment of large subdomain (IMT-AAA: 248 - 406, FtsH: 166 - 327). Both the large and small subdomains of IMT-AAA overlay well with that of FtsH (Figure 4.6).

Similar to Phe234 in the FtsH structure (84), the aromatic residues in the pore loop of IMT-AAA (Tyr316) align in the central pore at three different layers, with the distances between opposing tyrosine residues decreasing when the aromatic residues getting close to the peptidase chamber (Figure 4.7a). Furthermore, the arginine fingers of IMT-AAA have almost identical conformations as those of FtsH (Arg400 in IMT-AAA and Arg320 in FtsH) (Figure 4.7b). These indicate that useful information can be extracted from the hexameric model of IMT-AAA. Interestingly, based on the comparison of available structures of AAA+ proteins, Zhang and colleagues (192) reported that a glutamate switch (a polar residue N/T/S in an octapeptide sequence) in AAA+ proteins, located between the Walker A and Walker B motifs, links the ATPase activity and ligand binding. The polar residue interacts with the Walker B glutamate residue when AAA+ proteins bind ATP, but no such interaction was observed when the proteins are complexed with ADP, such as PspF (ref. 196 figure 2b and 2f). However, when bound to ADP and DNA, the asparagine in the glutamate switch of ORC1 hydrogen-bonds with the Walker B glutamate residue. In contrast, this interaction is lost when ADPNP is incorporated into ORC2 (ref. 196 figure 2c and 2g). In the structure model of hexameric IMT-AAA, the glutamate in the Walker B motif (Glu343) interacts with the serine residues (Ser308 and Ser310) in the glutamate switch in the presence of ADP (Figure 4.8a). It would be interesting to determine if such interaction will disappear when IMT-AAA is co-crystallized with a non-hydrolyzable ATP analogue. An inter-subunit signaling motif (ISS) has been described previously (193). Mutations of the amino acids in this motif exhibit gain-of-function in the presence of substitution of Glu by Gln in the Walker B motif (Figure 4.3).

An interaction between the aspartate residue (D421 in Yta10 in yeast) in the ISS motif and the arginine finger (R450 in Yta10 in yeast) is thought to link the movement of arginine finger and ATP-binding site on the same subunit (193). In consistent, we observed the same interaction between the arginine finger (Arg400) and Asp371 (corresponding to D421 in Yta10 in yeast) (Figure 4.8b), which may have the same function.

In conclusion, the IMT-AAA structure was determined in complex with ADP, revealing the highly conserved the structure of classical AAA proteins. Further structural data of proteins containing both the ATPase domain and the peptidase domain, in the apo form and complexed with different ATP analogues, are of interest and importance to understand the function of i-AAA at the molecular basis.

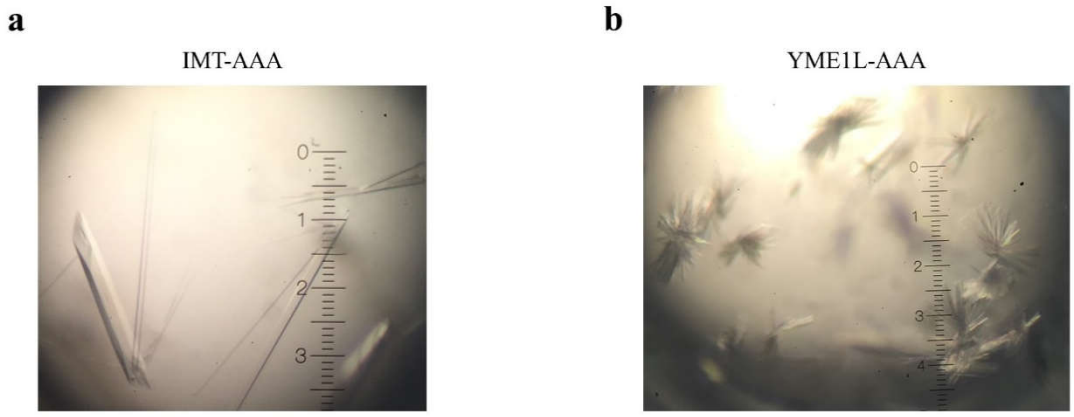


Figure 4.1 Crystals of IMT-AAA and YME1L-AAA.

Figures showing crystals grown after optimization of crystallization conditions at 20 °C. **a.** IMT-AAA (Ammonium citrate 0.2M, pH 7.3, PEG3350 14%). **b**. YME1L-AAA (0.1M Sodium acetate, pH4.6, PEG8000, 6%, 0.1 M magnesium acetate).

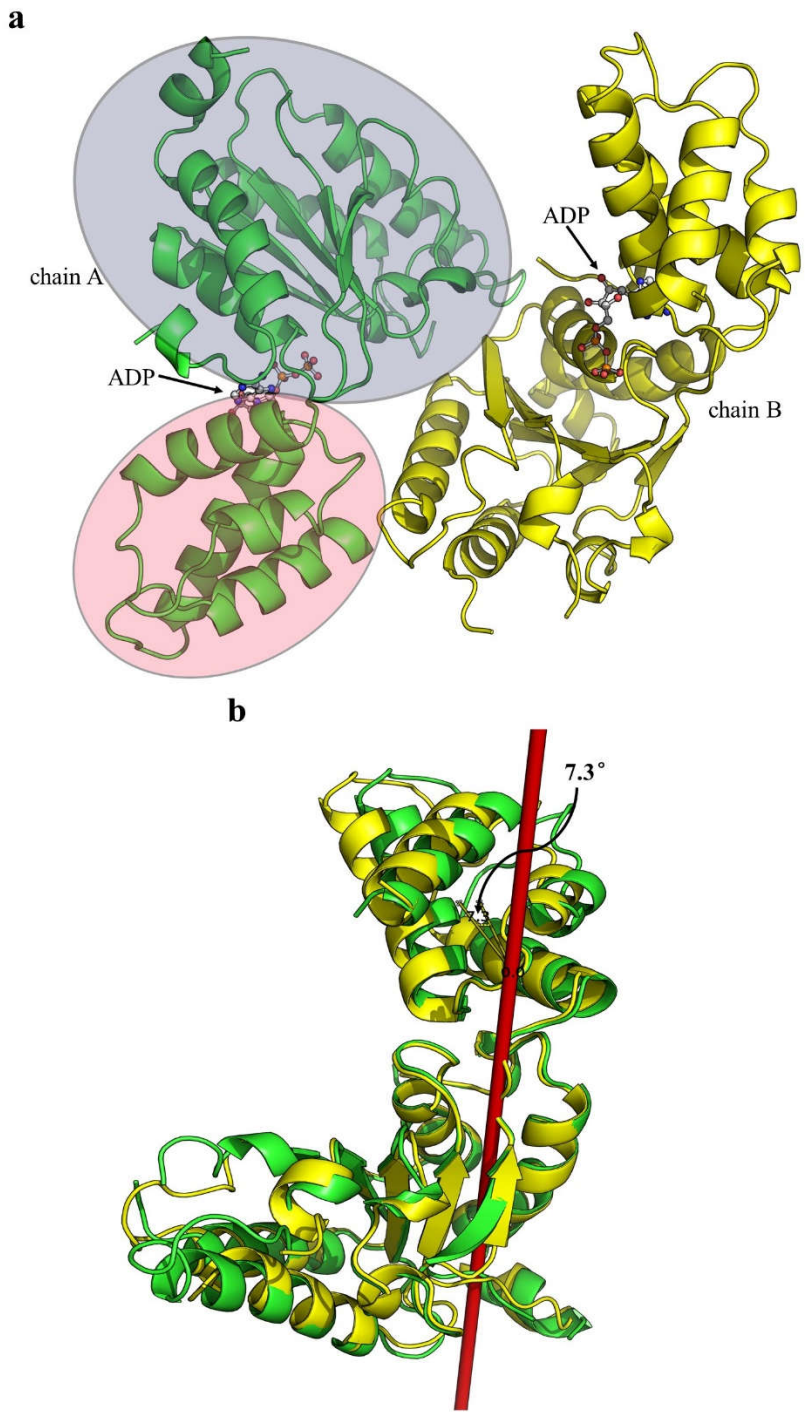
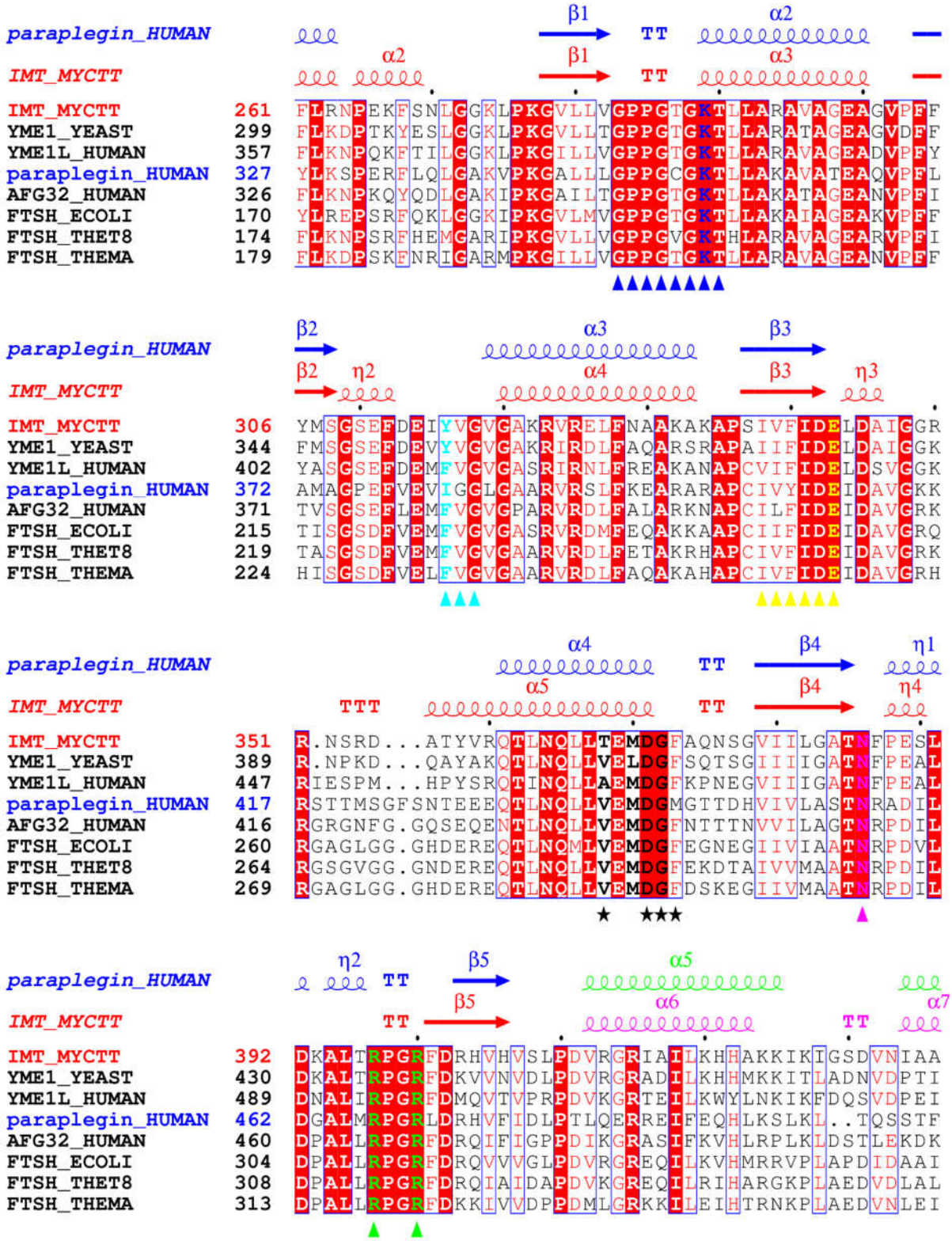


Figure 4.2 Overall structure of IMT-AAA.

a. Structure of one asymmetric unit (ASU) of IMT-AAA. Two monomers are present in one ASU. Chain A is colored in green and chain B is colored in yellow. ADP is represented as ball and stick. The N-terminal large subdomain of chain A is shaded in blue ellipse and the small subdomain is shaded in red ellipse. **b**. Superposition of large subdomain of chain A onto that of chain B results in a rotation of 7.3° of the small subdomain. Colors of different monomers are the same as shown in panel **a**.


paraplegin_HUMAN
IMT_MYCTT
IMT_MYCTT 261 FLRNPEKFSNLGGKLPKGVLLVGPPGTGKTLLARAVAGEAGVPFF
YME1_YEAST 299 FLKDPTKYESLGGKLPKGVLLTGPPGTGKTLLARATAGEAGVDFF
YME1L_HUMAN 357 FLKNPQKFTILGGKLPKGILLVGPPGTGKTLLARAVAGEADVPFY
paraplegin_HUMAN 327 YLKSPERFLQLGAKVPKGALLLGPPGCGKTLLAKAVATEAQVPFL
AFG32_HUMAN 326 FLKNPKQYQDLGAKIPKGAILTGPPGTGKTLLAKATAGEANVPFI
FTSH_ECOLI 170 YLREPSRFQKLGGKIPKGVLMVGPPGTGKTLLAKAIAGEAKVPFF
FTSH_THET8 174 FLKNPSRFHEMGARIPKGVLLVGPPGVGKTHLARAVAGEARVPFI
FTSH_THEMA 179 FLKDPSKFNRIGARMPKGILLVGPPGTGKTLLARAVAGEANVPFF
IMT_MYCTT 306 YMSGSEFDEIYVGVGAKRVRELFNAAKAKAPSIVFIDELDAIGGR
YME1_YEAST 344 FMSGSEFDEVYVGVGAKRIRDLFAQARSRAPAIIFIDELDAIGGK
YME1L_HUMAN 402 YASGSEFDEMFVGVGASRIRNLFREAKANAPCVIFIDELDSVGGK
paraplegin_HUMAN 372 AMAGPEFVEVIGGLGAARVRSLFKEARARAPCIVYIDEIDAVGKK
AFG32_HUMAN 371 TVSGSEFLEMFVGVGPARVRDLFALARKNAPCILFIDEIDAVGRK
FTSH_ECOLI 215 TISGSDFVEMFVGVGASRVRDMFEQAKKAAPCIIFIDEIDAVGRQ
FTSH_THET8 219 TASGSDFVEMFVGVGAARVRDLFETAKRHAPCIVFIDEIDAVGRK
FTSH_THEMA 224 HISGSDFVELFVGVGAARVRDLFAQAKAHAPCIVFIDEIDAVGRH
IMT_MYCTT 351 R.NSRD...ATYVRQTLNQLLTEMDGFAQNSGVIILGATNFPESL
YME1_YEAST 389 R.NPKD...QAYAKQTLNQLLVELDGFSQTSGIIIIGATNFPEAL
YME1L_HUMAN 447 RIESPM...HPYSRQTINQLLAEMDGFKPNEGVIIIGATNFPEAL
paraplegin_HUMAN 417 RSTTMSGFSNTEEEQTLNQLLVEMDGMGTTDHVIVLASTNRADIL
AFG32_HUMAN 416 RGRGNFG.GQSEQENTLNQLLVEMDGFNTTTNVVILAGTNRPDIL
FTSH_ECOLI 260 RGAGLGG.GHDEREQTLNQMLVEMDGFEGNEGIIVIAATNRPDVL
FTSH_THET8 264 RGSGVGG.GNDEREQTLNQLLVEMDGFEKDTAIVVMAATNRPDIL
FTSH_THEMA 269 RGAGLGG.GHDEREQTLNQLLVEMDGFDSKEGIIVMAATNRPDIL
IMT_MYCTT 392 DKALTRPGRFDRHVHVSLPDVRGRIAILKHHAKKIKIGSDVNIAA
YME1_YEAST 430 DKALTRPGRFDKVVNVDLPDVRGRADILKHHMKKITLADNVDPTI
YME1L_HUMAN 489 DNALIRPGRFDMQVTVPRPDVKGRTEILKWYLNKIKFDQSVDPEI
paraplegin_HUMAN 462 DGALMRPGRLDRHVFIDLPTLQERREIFEQHLKSLKL..TQSSTF
AFG32_HUMAN 460 DPALLRPGRFDRQIFIGPPDIKGRASIFKVHLRPLKLDSTLEKDK
FTSH_ECOLI 304 DPALLRPGRFDRQVVVGLPDVRGREQILKVHMRRVPLAPDIDAAI
FTSH_THET8 308 DPALLRPGRFDRQIAIDAPDVKGREQILRIHARGKPLAEDVDLAL
FTSH_THEMA 313 DPALLRPGRFDKKIVVDPPDMLGRKKILEIHTRNKPLAEDVNLEI

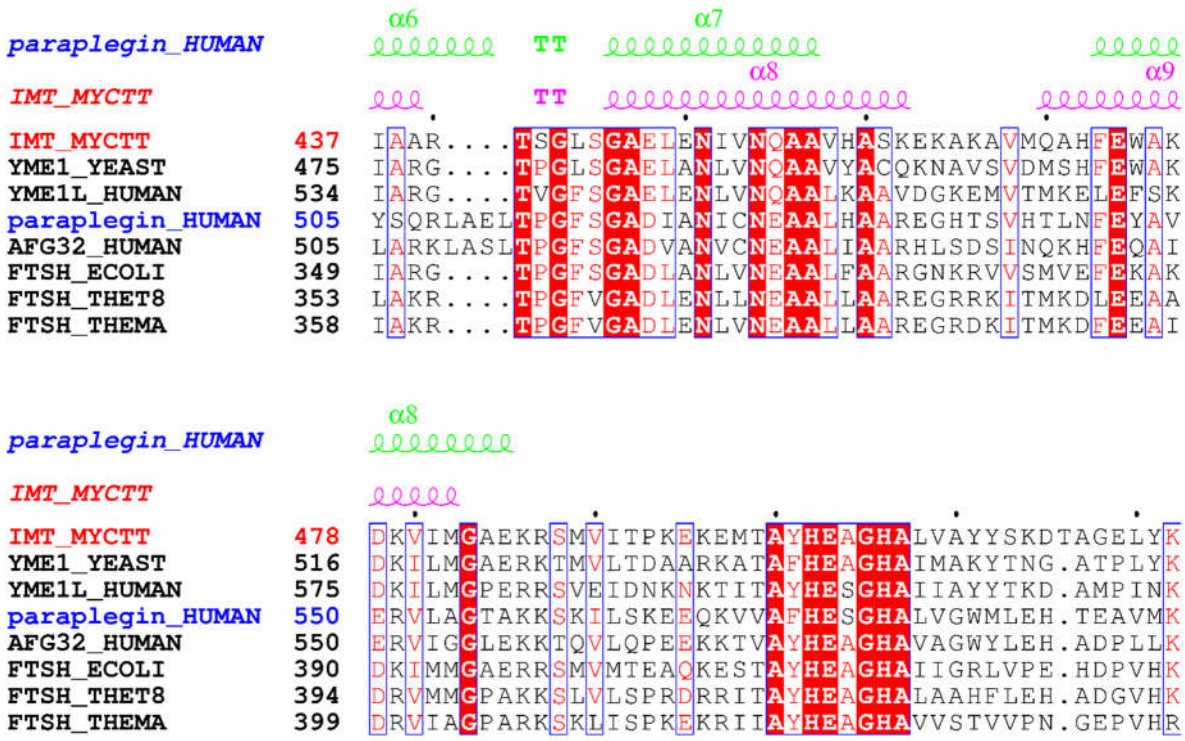


Figure 4.3 Multiple Sequence alignment of the AAA+ modules of i-AAA proteins.

Multiple sequence alignment shows the conserved motifs and residues within the AAA+ modules of i-AAA proteins. Eight sequences belonging to the FtsH family are aligned. On top of the aligned sequences, secondary structures contained in the PDB files of IMT-AAA and paraplegin (PDB code: 2QZ4) are indicated. Secondary structures comprising the large subdomain are colored in blue for paraplegin and in red for IMT-AAA, respectively. Those constituting the small subdomain of paraplegin and IMT-AAA are colored in green and magenta, respectively. Blue triangles under the aligned sequences indicate the conserved Walker A motif, while the yellow ones imply the Walker B motif, respectively. The lysine residues in the Walker A motif are bolded in blue, and the glutamate residues in the Walker B motif are bolded in yellow, respectively. Arginine fingers are bolded in green and sensor 1 residues are bolded in magenta, respectively. Black stars under the aligned sequences indicate the inter-subunit signaling motif (197). Sequence alignment was implemented in the MultAlin webserver (153), and the figure was prepared with ESPript 3.0 webserver (154).

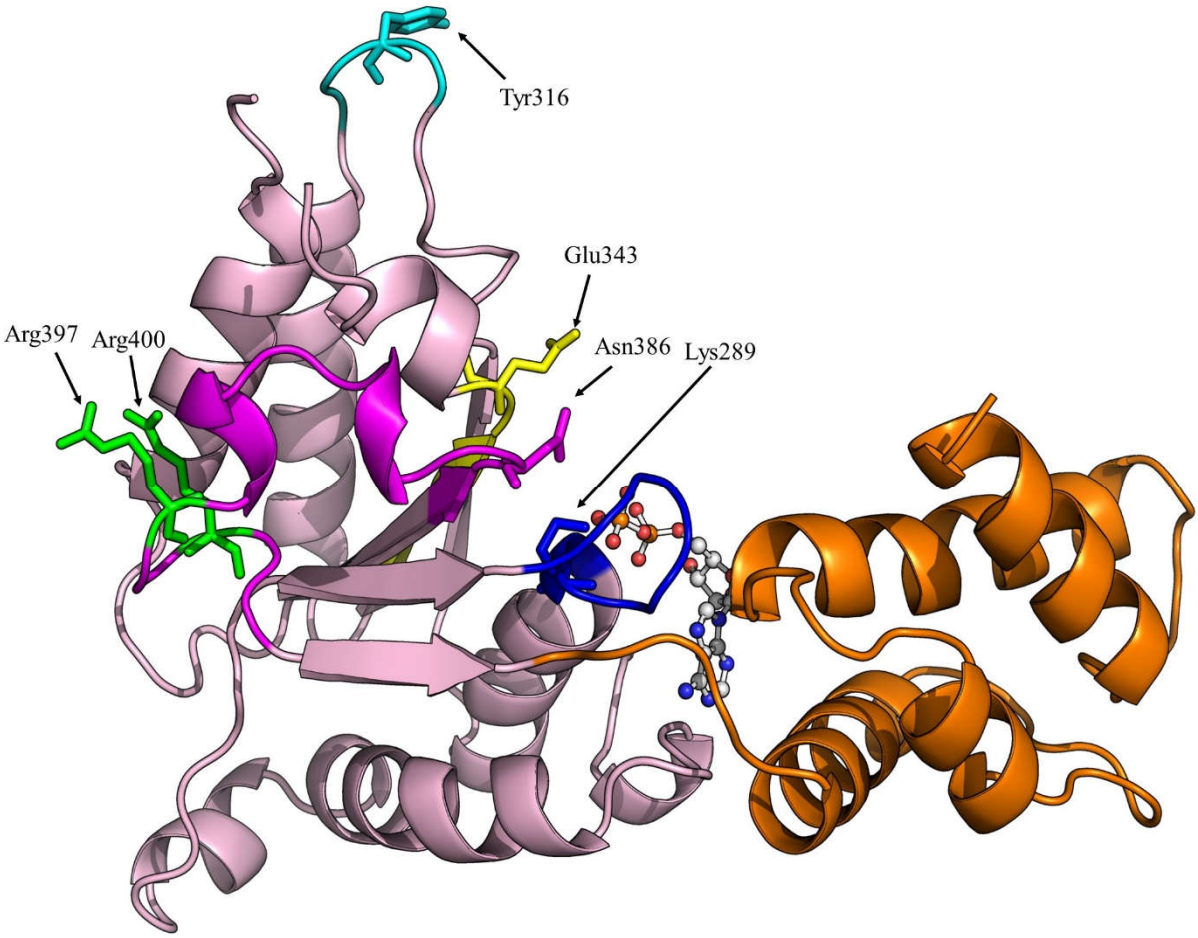


Figure 4.4 Conserved motifs and residues in IMT-AAA.

Cartoon representation showing the structure of an IMT-AAA monomer. The large subdomain is colored in light pink and the small subdomain is colored in orange. The Walker A motif is indicated in blue and the Walker B motif is shown in yellow. The pore loop is colored in cyan. Secondary region of homology (SRH) is colored in magenta. ADP is represented as ball-and-stick. Conserved residues are shown as stick. (Walker A Lys289: blue, pore loop Tyr316: cyan, Walker B Glu343: yellow, sensor 1 Asn386: magenta, Arginine fingers Arg397 and Arg400: green).

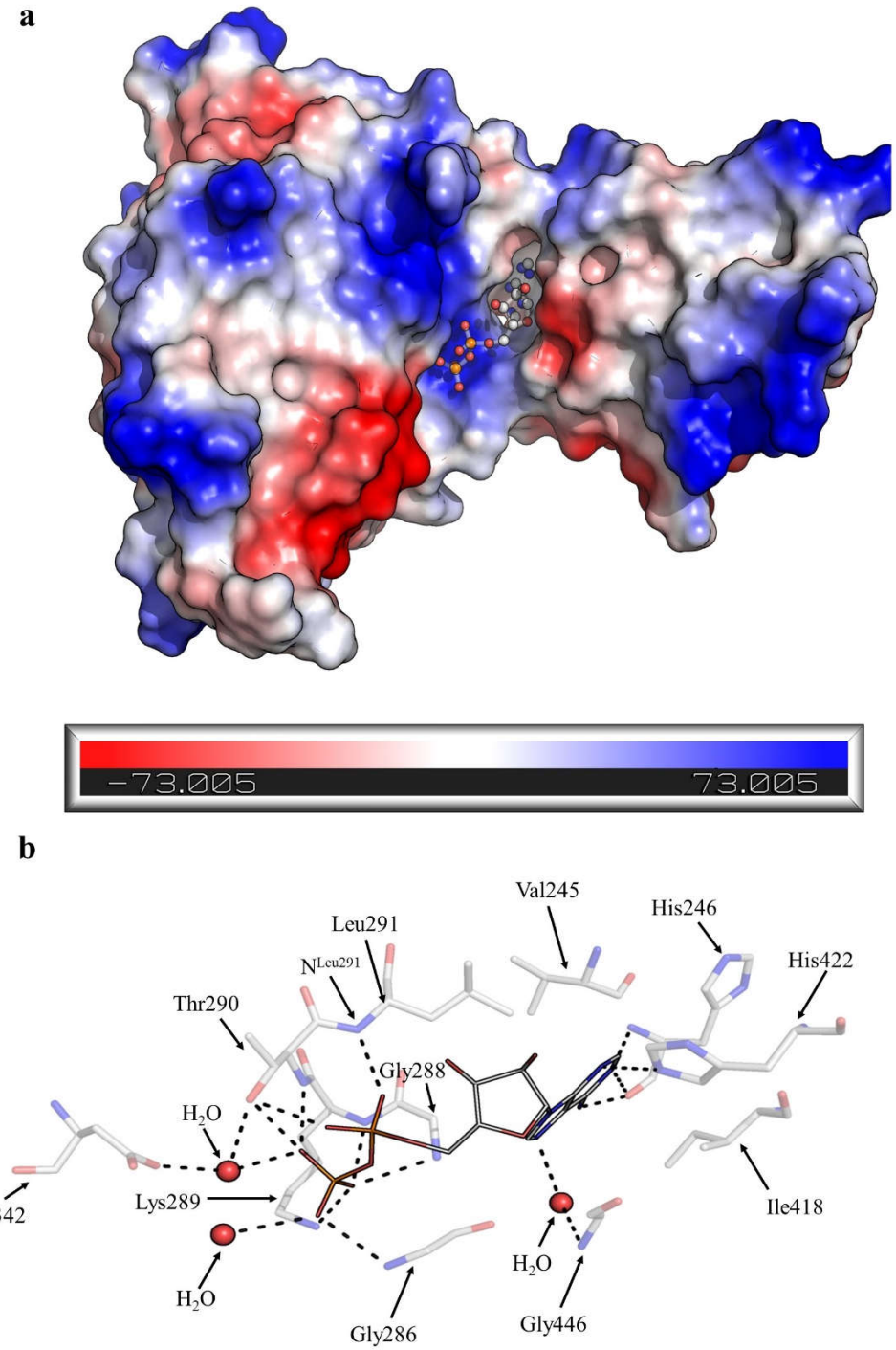


Figure 4.5 ADP-binding site on IMT-AAA.

a. Electrostatic potential representation showing the ADP-binding site on IMT-AAA. A cleft for ADP-binding is formed at the interface of the large and small subdomains. Red color indicates negative charge and positive charge is colored in blue. ADP molecule is represented as ball-and-stick. **b**. The interactions between ADP and residues on IMT-AAA. Water molecules are represented as spheres in red. ADP molecule is shown as line. ADP-interacting residues are shown as stick. Black dash indicates hydrogen bond.

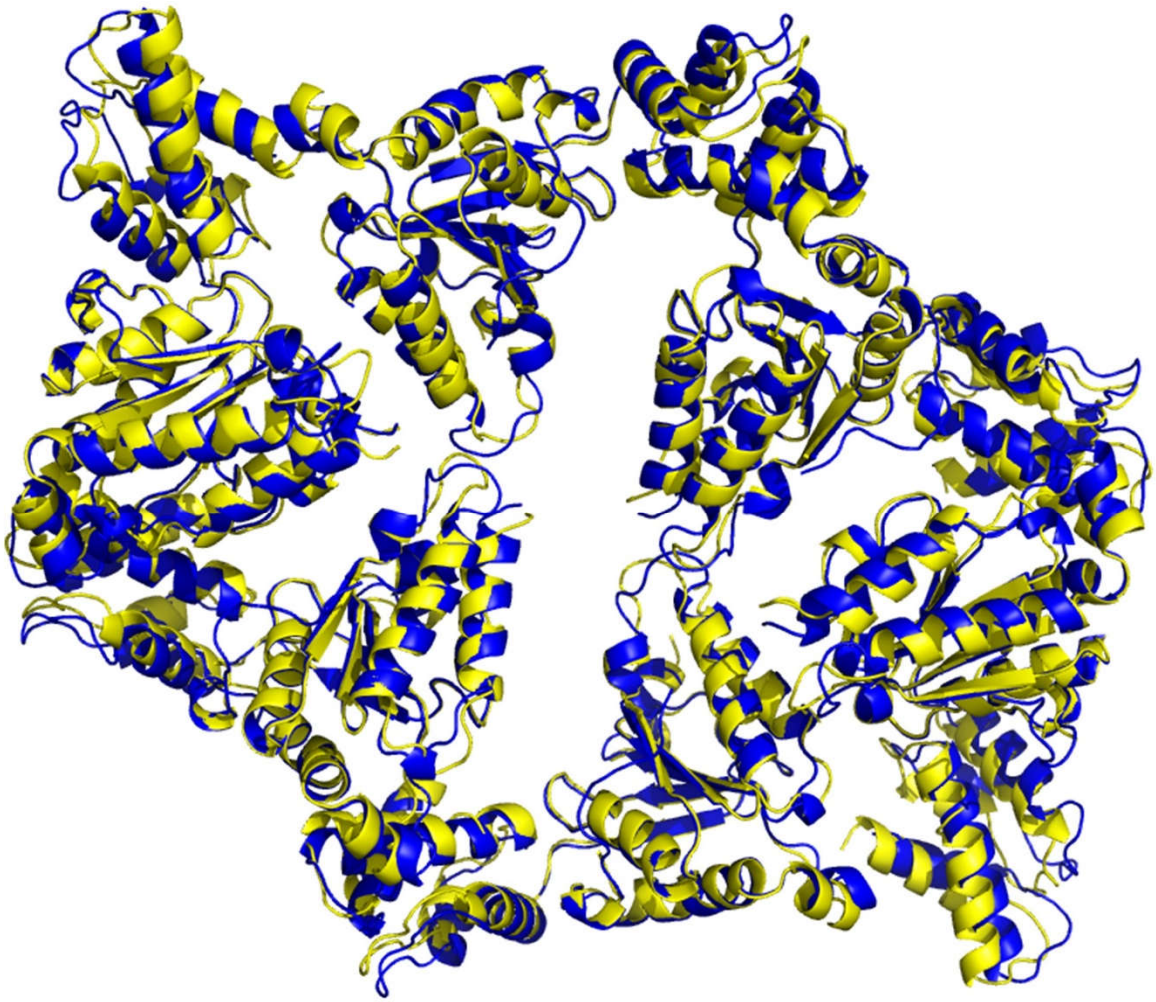


Figure 4.6 Hexameric model of IMT-AAA.

Superposition of IMT-AAA (yellow) on the crystal structure of FtsH (2CE7) (blue) generates a hexameric model of IMT-AAA. Superpose was based on structural alignment of fragments 248 - 408 of IMT-AAA and 166 - 327 of FtsH.

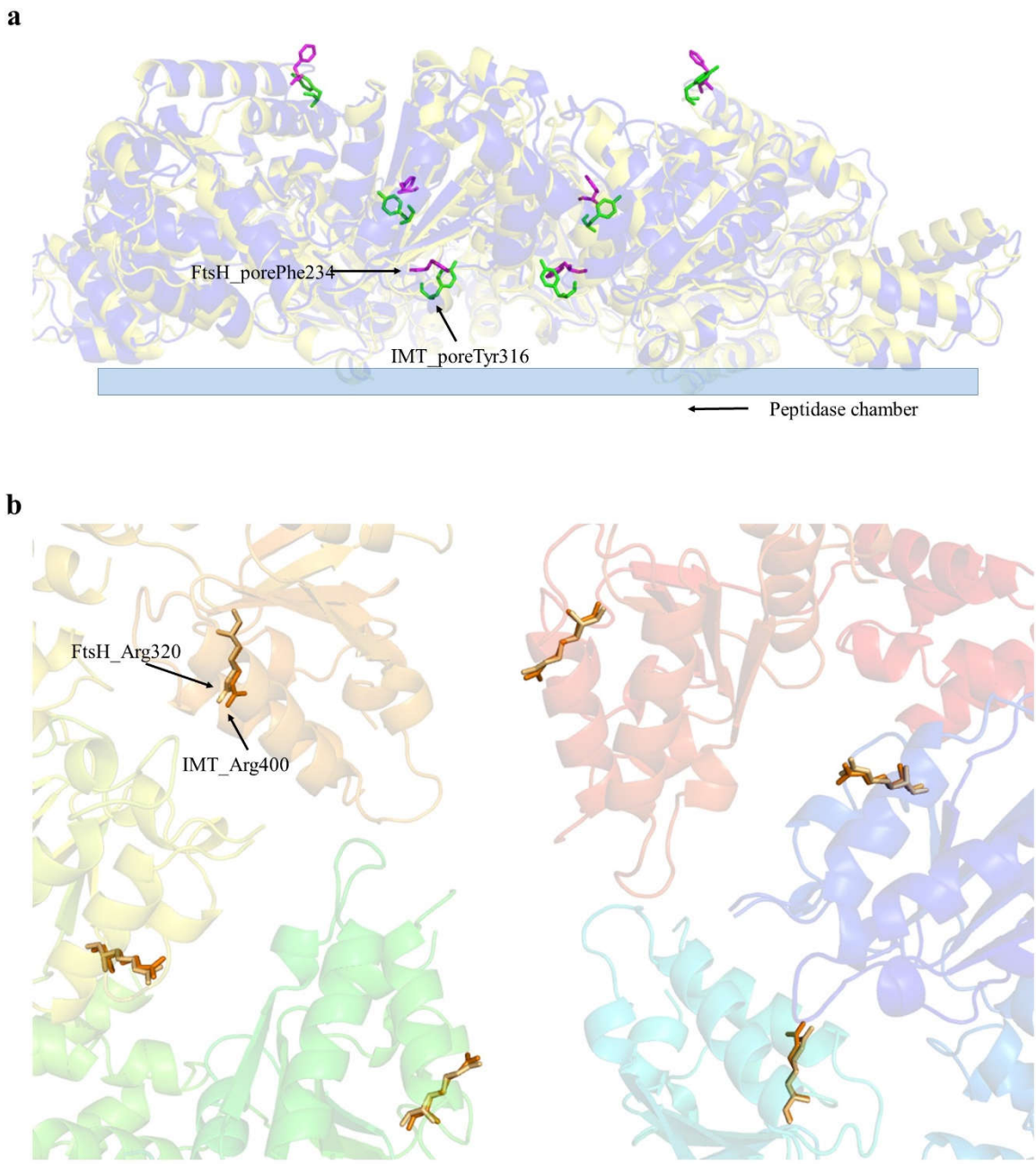


Figure 4.7 Pore loop residues and Arginine fingers in the hexameric structure.

a. Tyr316 residues in the pore loops of IMT-AAA show a similar three-layer organization aligning the central pore of the AAA+ module. Transparent cartoon representation of the hexameric model of IMT-AAA is colored in yellow and FtsH ATPase domain (2CE7) is colored in blue. Position of the peptidase domain of FtsH is as indicated. Tyr316 (IMT-AAA) is represented as green stick and Phe234 (FtsH) is shown as stick and colored in magenta. **b**. Arginine fingers of IMT-AAA (Arg400) indicated as orange stick adopt the same conformation as those of FtsH (Arg320, light orange stick). Transparent cartoon shows the hexameric model of IMT-AAA, with different subunits shown in different colors.

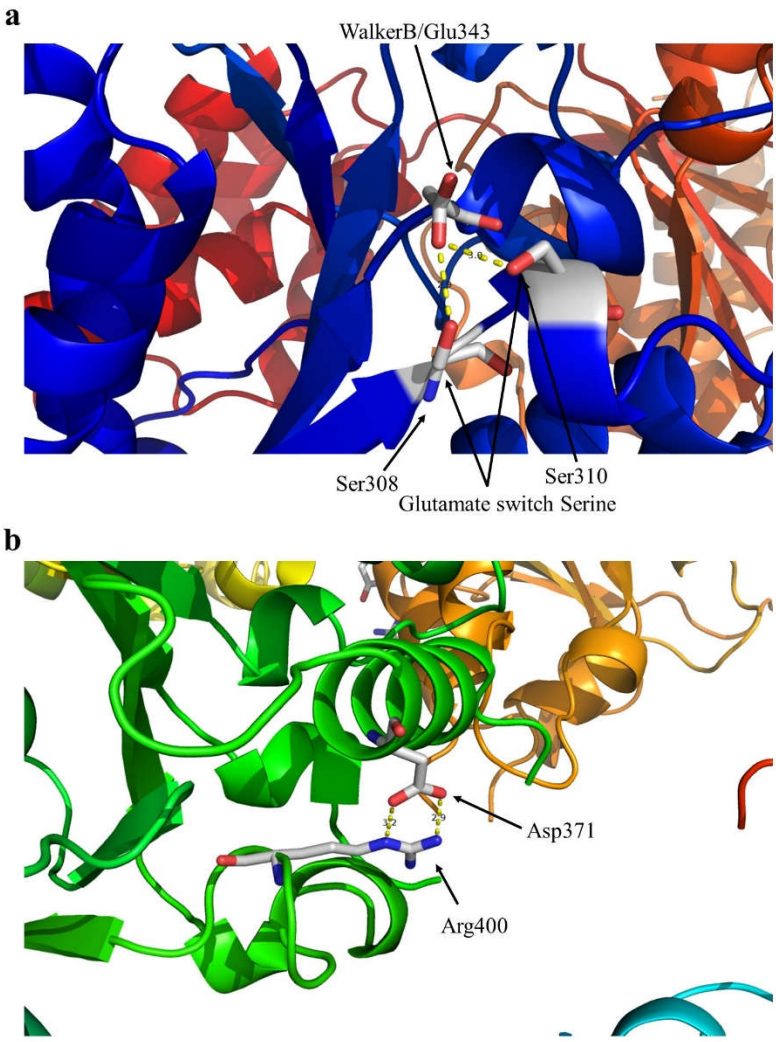


Figure 4.8 Glutamate switch and ISS of IMT-AAA.

a. Cartoon representation showing the position of polar residues in the glutamate switch motif (196) (Figure 4.3). The Walker B motif (Glu343) forms hydrogen-bonds with two serine residues (Ser308 and Ser310) in the glutamate switch, respectively. Interacting residues are shown as stick. **b**. Cartoon representation showing the position of Asp371 in the inter-subunit signaling (ISS) motif (197). Asp371 interacts with the arginine finger (Arg400) in the same subunit. Interacting residues are shown as stick.

Table 4.1 Data collection and refinement statistics.

Crystal	IMT-AAA
Data collection	
Wavelength (Å)	1.18
Temperature (K)	100
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	45.916, 86.624, 155.025
α, β, γ (°)	90, 90, 90
Resolution (Å)	50 - 2.45(2.538 - 2.45)[a]
R_{sym}	0.093(1.025)
$R_{meas,}$	0.099(1.092)
R_{pim}	0.032(0.370)
$I/\sigma(I)$	27.1(2.0)
Completeness (%)	100(100)
Redundancy	9.2(8.0)
Refinement	
Resolution (Å)	2.45
No. reflections	23503
R_{work}[b] / R_{free}[c]	0.196/0.237
No. atoms	
Total	3785
Protein	3654
ADP	54
Water	58
B factors	
Protein	60.69
ADP	54.13
Water	59.41
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.436
Ramachandran plot	
Most favoured (%)	97.47
Allowed (%)	2.53

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

[b] $R_{work} = \sum||F_{obs}|-|F_{calc}||/\sum|F_{obs}|$. F_{obs} and F_{calc} are amplitudes of observed and calculated structure-factor for dataset used in refinement, respectively.

[c] R_{free} is the R factor calculated on the 5.15% of data excluded from refinement.

Chapter 5 – Future Direction

The i-AAA protease is an essential component of the mitochondrial quality control system and contributes significantly to the maintenance of mitochondrial homeostasis. YME1L implements its function by regulating its substrates involved in various biological processes in mitochondria. Dysfunction of YME1L results in mitochondrial defects and is related to neurological disorders. A mutation in the mitochondrial presequence of YME1L was identified in people carrying a novel mitochondriopathy (194). This mutation impairs its maturation and subsequently precludes the formation of functional i-AAA, which finally leads to the abnormality of mitochondria. Despite the great advances in understanding the physiological roles of YME1L in detail, there still remain interesting questions to explore.

5.1 Different modes of substrate regulation by YME1L

Regulation of substrates by YME1L can be dependent or independent on its proteolytic activity. Destructive digestion of substrates has been seen in examples of the degradation of PRELI (121), and the degradation of TIM17A under stress (130). Processing of OPA1, a human inner membrane protein that is key to mitochondrial dynamics, is dependent on the peptidase activity of YME1L. However, this processing action does not cause complete degradation of OPA1. Instead, cleavage of OPA1 by YME1L results in a shortened version (127). Interestingly, polynucleotide phosphorylase (PNPase) in yeast is translocated into the mitochondrial intermembrane space by Yme1 from the TIM23 translocon, the yeast homolog of YME1L, independent of its peptidase activity (132). Similar chaperone-mimicking function is seen in the example of mitochondrial ClpXP, a AAA+ protease. ClpX regulates heme biogenesis by activating a key enzyme, without the requirement of ClpP (101). How does YME1L determine if a substrate needs to be processed or degraded completely? Is it the nature of the substrate sequence or the stability of a substrate that decides the destiny of a protein substrate? Is there any sequence preference of cleavage site? How does Yme1 translocate PNPase independent of its metallopeptidase domain? The translocation of PNPase by Yme1 may happen in the way similar to that of ClpXP. The ATPase module and the peptidase domain of ClpXP are separated apart into two different peptide chains. In contrast, Yme1 is a self-compartmental protease containing both of the two domains in a single polypeptide chain. The difference complicates the explanation for YME1L. Although the translocation of PNPase

depends on an ATPase-competent Yme1, it is interesting to determine whether this chaperone-like activity of Yme1 is dependent on the aromatic residues in the pore loop. Corresponding YME1L variant affecting substrate gripping can be utilized to investigate whether PNPase is translocated through the central pore of the AAA+ module. In addition, previously reported helices located at the AAA+ module (NH) and the peptidase domain (CH) that are responsible for substrate recognition (152) are also worthy of investigation of their roles in PNPase translocation.

5.2 Biochemical characterization of YME1L

Although several substrates have been identified for YME1L, no degron sequences from native substrate were reported. Using the cc-hex fused YME1L, hexYME1L, we found a conserved motif that is a promising YME1L degron candidate. Further experiments are needed for *in vitro* study of native substrate degradation and degron characterization, by virtue of the hexYME1L. An N-terminal helix (NH) and a C-terminal helix (CH) about 50-amino acid long each have been suggested to engage substrates for degradation (152). However, no further detail was given for substrate binding. The hexYME1L presented in this thesis can be applied in this regard, to pinpoint the key residue on NH and CH for substrate recognition. Furthermore, by taking the advantage of the variants of cc-hex, cc-Hex-H24 and cc-Hex-D24, inter subunit interaction of YME1L can be investigated. cc-Hex-H24 carries a histidine substitution at position 24, while an aspartate is introduced as substitution at the same position for cc-Hex-D24. These two variants can form hetero-hexamers in solution, with a version of alternating arrangement to be dominant. By using the cc-hex variants, a hetero-hexameric YME1L can possibly be created, carrying different residues at the same position. The information obtained from the study will greatly improve our understanding of YME1L in more detail. It would be interesting to prepare a construct of YME1L whose transmembrane segment is replaced by cc-hex. The construct containing the N-terminal region can be used to explore the effect of N-domain on substrate recognition and degradation.

5.3 Adaptor characterization

It has been suggested that Mgr1 and Mgr3 associate with i-AAA to form a supercomplex (148), and the two proteins were later reported to be adaptors of Yme1 (149). Similarly, adaptors have been identified for bacterial AAA+ proteases, such as SspB, which functions as an adaptor

of ClpXP. Whether Mgr1 and Mgr3 function in the same way as SspB or they utilize a different mechanism remains to be examined. Moreover, are there any more adaptors existing for YME1L? Further studies are needed to address these questions.

5.4 Autodegradation of hexYME1L

In this established *in vitro* system, hexYME1L is autodegraded during the course of degradation reaction. Autodegradation of hexYME1L seems to be more severe in the presence of poor substrates. To identify the potential sequence that induces autodegradation, series of YME1L truncations at either or both termini can be generated, with both ATPase activity and peptidase activity preserved yet autodegradation eliminated or greatly reduced. Considering the long unstructured fragment present between the cc-hex fragment and the N-terminus of the AAA+ module of YME1L, we expect that the signal responsible for YME1L autodegradation exists in this flexible segment. In addition to constructing YME1L truncations, amino acid substitutions can be introduced into the suspect regions important for autodegradation. This strategy may affect the structure and activities of YME1L to a less extent, compared to truncating YME1L.

5.5 Comparison of hexYME1L to the membrane-anchored YME1L

YME1L contains a single transmembrane segment and is anchored in the inner membrane. This cellular localization of YME1L restricts its position inside the mitochondria. The N-terminus of the AAA+ module of YME1L is separated from the transmembrane segment by a predicted unstructured sequence containing about 20 residues. The space between the inner membrane and the AAA+ module has a defined size, precluding free access of the enzyme. On the other hand, this distance between the ATPase domain and the inner membrane may explain the minimal length requirement of recognition sequence about 20 residues long (139).

There may be more than one recognition sequences present on the YME1L substrates. As two distant substrate-binding sites on Yme1 have been suggested (152). For certain substrates, the CH fragment initially binds the substrate, followed by involvement of the NH sequence in substrate binding. This hypothesis can be easily tested using the hexYME1L *in vitro* system, as no structural variation of the catalytic domains of YME1L and steric hindrance caused by the inner membrane are present in the soluble hexYME1L.

Considering the physical separation of the N-domain and the catalytic modules, it is impossible that the N-domain of YME1L contributes to the recruitment of soluble proteins in the intermembrane space, such as Ups1/Mdm35 tested in the hexYME1L system. However, a possibility still remains that endogenous transmembrane substrates with a matrix fragment present bind to the N-domain of YME1L, resulting in the subsequent degradation of proteins by YME1L.

Structural variation of endogenous membrane proteins under stress condition may result in the availability of sequences that can be recognized by the YME1L enzyme. This mechanism is supported by the degradation of TIM17A, an inner membrane protein that is a component of the TIM23 complex (130). However, the degradation of model proteins with the C-terminal fragment of TIM17B attached by hexYME1L in solution is in contrast with *in vivo* results, which showed no degradation of TIM17B by YME1L (130). In the cellular context, it is conceivable that additional molecules modulate the accessibility of degron sequences. Whereas, in the *in vitro* assays, such regulation mechanism does not exit. As a result, the C-terminal sequence of TIM17B may serve as a degron for mediating substrate degradation by hexYME1L in solution.

Although the hexYME1L system does not represent YME1L in the physiological context, it is still valuable for investigation of biochemical properties of YME1L, such as unfolding strength and substrate recruitment.

In conclusion, to understand the physiological function of YME1L at the molecular basis, *in vivo* and *in vitro* studies are necessary, using cellular, biochemical and structural approaches.

References

1. Youle RJ, van der Bliek AM. Mitochondrial fission, fusion, and stress. Science. 2012;337(6098):1062-5.

2. van der Bliek AM, Shen Q, Kawajiri S. Mechanisms of mitochondrial fission and fusion. Cold Spring Harb Perspect Biol. 2013;5(6).

3. Yin F, Cadenas E. Mitochondria: the cellular hub of the dynamic coordinated network. Antioxid Redox Signal. 2015;22(12):961-4.

4. Tatsuta T, Scharwey M, Langer T. Mitochondrial lipid trafficking. Trends Cell Biol. 2014;24(1):44-52.

5. Contreras L, Drago I, Zampese E, Pozzan T. Mitochondria: the calcium connection. Biochim Biophys Acta. 2010;1797(6-7):607-18.

6. Okamoto K, Kondo-Okamoto N. Mitochondria and autophagy: Critical interplay between the two homeostats. Bba-Gen Subjects. 2012;1820(5):595-600.

7. Wang C, Youle RJ. The role of mitochondria in apoptosis*. Annu Rev Genet. 2009;43:95-118.

8. Rouault TA. Biogenesis of iron-sulfur clusters in mammalian cells: new insights and relevance to human disease. Dis Model Mech. 2012;5(2):155-64.

9. Stehling O, Lill R. The role of mitochondria in cellular iron-sulfur protein biogenesis: mechanisms, connected processes, and diseases. Cold Spring Harb Perspect Biol. 2013;5(8):a011312.

10. Lill R, Srinivasan V, Muhlenhoff U. The role of mitochondria in cytosolic-nuclear iron-sulfur protein biogenesis and in cellular iron regulation. Curr Opin Microbiol. 2014;22:111-9.

11. Vamecq J, Dessein AF, Fontaine M, Briand G, Porchet N, Latruffe N, et al. Mitochondrial Dysfunction and Lipid Homeostasis. Current Drug Metabolism. 2012;13(10):1388-400.

12. Nunnari J, Suomalainen A. Mitochondria: in sickness and in health. Cell. 2012;148(6):1145-59.

13. Vyas S, Zaganjor E, Haigis MC. Mitochondria and Cancer. Cell. 2016;166(3):555-66.

14. Zong WX, Rabinowitz JD, White E. Mitochondria and Cancer. Mol Cell. 2016;61(5):667-76.

15. Szewczyk A, Wojtczak L. Mitochondria as a pharmacological target. Pharmacol Rev. 2002;54(1):101-27.

16. Olszewska A, Szewczyk A. Mitochondria as a pharmacological target: Magnum overview. Iubmb Life. 2013;65(3):273-81.

17. Kuhlbrandt W. Structure and function of mitochondrial membrane protein complexes. BMC Biol. 2015;13:89.

18. Dudkina NV, Kouril R, Peters K, Braun HP, Boekema EJ. Structure and function of mitochondrial supercomplexes. Biochim Biophys Acta. 2010;1797(6-7):664-70.

19. Walther DM, Rapaport D. Biogenesis of mitochondrial outer membrane proteins. Biochim Biophys Acta. 2009;1793(1):42-51.

20. Pfanner N, Wiedemann N, Meisinger C, Lithgow T. Assembling the mitochondrial outer membrane. Nat Struct Mol Biol. 2004;11(11):1044-8.

21. Benz R. Porin from Bacterial and Mitochondrial Outer Membranes. Crc Cr Rev Bioch Mol. 1985;19(2):145-90.

22. Benz R. Biophysical Properties of Porin Pores from Mitochondrial Outer-Membrane of Eukaryotic Cells. Experientia. 1990;46(2):131-7.

23. Tommassen J. Assembly of outer-membrane proteins in bacteria and mitochondria. Microbiology. 2010;156(Pt 9):2587-96.

24. Shoshan-Barmatz V, Gincel D. The voltage-dependent anion channel: characterization, modulation, and role in mitochondrial function in cell life and death. Cell Biochem Biophys. 2003;39(3):279-92.

25. Hoogenboom BW, Suda K, Engel A, Fotiadis D. The supramolecular assemblies of voltage-dependent anion channels in the native membrane. J Mol Biol. 2007;370(2):246-55.

26. Misra R. Assembly of the beta-Barrel Outer Membrane Proteins in Gram-Negative Bacteria, Mitochondria, and Chloroplasts. ISRN Mol Biol. 2012;2012:708203.

27. Lemasters JJ, Holmuhamedov E. Voltage-dependent anion channel (VDAC) as mitochondrial governator - Thinking outside the box. Bba-Mol Basis Dis. 2006;1762(2):181-90.

28. Distler AM, Kerner J, Hoppel CL. Proteomics of mitochondrial inner and outer membranes. Proteomics. 2008;8(19):4066-82.

29. Becker T, Gebert M, Pfanner N, van der Laan M. Biogenesis of mitochondrial membrane proteins. Curr Opin Cell Biol. 2009;21(4):484-93.

30. Neupert W, Herrmann JM. Translocation of proteins into mitochondria. Annual Review of Biochemistry. 2007;76:723-49.

31. Herrmann JM, Riemer J. The intermembrane space of mitochondria. Antioxid Redox Signal. 2010;13(9):1341-58.

32. Taanman JW. The mitochondrial genome: structure, transcription, translation and replication. Biochim Biophys Acta. 1999;1410(2):103-23.

33. Taylor SW, Fahy E, Zhang B, Glenn GM, Warnock DE, Wiley S, et al. Characterization of the human heart mitochondrial proteome. Nat Biotechnol. 2003;21(3):281-6.

34. Calvo SE, Clauser KR, Mootha VK. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res. 2016;44(D1):D1251-7.

35. Yang JS, Kim J, Park S, Jeon J, Shin YE, Kim S. Spatial and functional organization of mitochondrial protein network. Sci Rep. 2013;3:1403.

36. Schmidt O, Pfanner N, Meisinger C. Mitochondrial protein import: from proteomics to functional mechanisms. Nat Rev Mol Cell Biol. 2010;11(9):655-67.

37. Dudek J, Rehling P, van der Laan M. Mitochondrial protein import: common principles and physiological networks. Biochim Biophys Acta. 2013;1833(2):274-85.

38. Bragoszewski P, Wasilewski M, Sakowska P, Gornicka A, Bottinger L, Qiu J, et al. Retro-translocation of mitochondrial intermembrane space proteins. Proc Natl Acad Sci U S A. 2015;112(25):7713-8.

39. St-Pierre J, Buckingham JA, Roebuck SJ, Brand MD. Topology of superoxide production from different sites in the mitochondrial electron transport chain. J Biol Chem. 2002;277(47):44784-90.

40. Murphy MP. How mitochondria produce reactive oxygen species. Biochem J. 2009;417:1-13.

41. Brand MD. The sites and topology of mitochondrial superoxide production. Exp Gerontol. 2010;45(7-8):466-72.

42. Turrens JF. Mitochondrial formation of reactive oxygen species. J Physiol. 2003;552(Pt 2):335-44.

43. Ray PD, Huang BW, Tsuji Y. Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. Cell Signal. 2012;24(5):981-90.

44. Finkel T. Signal transduction by reactive oxygen species. J Cell Biol. 2011;194(1):7-15.

45. Sabharwal SS, Schumacker PT. Mitochondrial ROS in cancer: initiators, amplifiers or an Achilles' heel? Nat Rev Cancer. 2014;14(11):709-21.

46. Fischer F, Hamann A, Osiewacz HD. Mitochondrial quality control: an integrated network of pathways. Trends in Biochemical Sciences. 2012;37(7):284-92.

47. Quiros PM, Langer T, Lopez-Otin C. New roles for mitochondrial proteases in health, ageing and disease. Nat Rev Mol Cell Biol. 2015;16(6):345-59.

48. Koppen M, Langer T. Protein degradation within mitochondria: Versatile activities of AAA proteases and other peptidases. Crit Rev Biochem Mol. 2007;42(3):221-42.

49. Gerdes F, Tatsuta T, Langer T. Mitochondrial AAA proteases--towards a molecular understanding of membrane-bound proteolytic machines. Biochim Biophys Acta. 2012;1823(1):49-55.

50. Gakh O, Cavadini P, Isaya G. Mitochondrial processing peptidases. Biochim Biophys Acta. 2002;1592(1):63-77.

51. Taylor EB, Rutter J. Mitochondrial quality control by the ubiquitin-proteasome system. Biochem Soc Trans. 2011;39(5):1509-13.

52. Lehmann G, Udasin RG, Ciechanover A. On the linkage between the ubiquitin-proteasome system and the mitochondria. Biochem Bioph Res Co. 2016;473(1):80-6.

53. Lehmann G, Ziv T, Braten O, Admon A, Udasin RG, Ciechanover A. Ubiquitination of specific mitochondrial matrix proteins. Biochem Biophys Res Commun. 2016;475(1):13-8.

54. Sauer RT, Baker TA. AAA+ proteases: ATP-fueled machines of protein destruction. Annu Rev Biochem. 2011;80:587-612.

55. Bochman ML, Schwacha A. The Mcm complex: unwinding the mechanism of a replicative helicase. Microbiol Mol Biol Rev. 2009;73(4):652-83.

56. Joly N, Zhang N, Buck M, Zhang X. Coupling AAA protein function to regulated gene expression. Biochim Biophys Acta. 2012;1823(1):108-16.

57. Grimm I, Erdmann R, Girzalsky W. Role of AAA(+)-proteins in peroxisome biogenesis and function. Biochim Biophys Acta. 2016;1863(5):828-37.

58. Neuwald AF, Aravind L, Spouge JL, Koonin EV. AAA(+): A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. Genome Res. 1999;9(1):27-43.

59. Lupas AN, Martin J. AAA proteins. Curr Opin Struct Biol. 2002;12(6):746-53.

60. Frickey T, Lupas AN. Phylogenetic analysis of AAA proteins. J Struct Biol. 2004;146(1-2):2-10.

61. Iyer LM, Leipe DD, Koonin EV, Aravind L. Evolutionary history and higher order classification of AAA+ ATPases. J Struct Biol. 2004;146(1-2):11-31.

62. Ammelburg M, Frickey T, Lupas AN. Classification of AAA+ proteins. J Struct Biol. 2006;156(1):2-11.

63. Ogura T, Whiteheart SW, Wilkinson AJ. Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. J Struct Biol. 2004;146(1-2):106-12.

64. Hanson PI, Whiteheart SW. AAA+ proteins: Have engine, will work. Nat Rev Mol Cell Bio. 2005;6(7):519-29.

65. Wendler P, Ciniawsky S, Kock M, Kube S. Structure and function of the AAA+ nucleotide binding pocket. Biochim Biophys Acta. 2012;1823(1):2-14.

66. Karata K, Inagawa T, Wilkinson AJ, Tatsuta T, Ogura T. Dissecting the role of a conserved motif (the second region of homology) in the AAA family of ATPases. Site-directed mutagenesis of the ATP-dependent protease FtsH. J Biol Chem. 1999;274(37):26225-32.

67. Raju RM, Goldberg AL, Rubin EJ. Bacterial proteolytic complexes as therapeutic targets. Nat Rev Drug Discov. 2012;11(10):777-89.

68. Maupin-Furlow J. Proteasomes and protein conjugation across domains of life. Nat Rev Microbiol. 2011;10(2):100-11.

69. Bar-Nun S, Glickman MH. Proteasomal AAA-ATPases: structure and function. Biochim Biophys Acta. 2012;1823(1):67-82.

70. Samanovic MI, Li H, Darwin KH. The pup-proteasome system of Mycobacterium tuberculosis. Subcell Biochem. 2013;66:267-95.

71. Lee I, Suzuki CK. Functional mechanics of the ATP-dependent Lon protease- lessons from endogenous protein and synthetic peptide substrates. Biochim Biophys Acta. 2008;1784(5):727-35.

72. Langklotz S, Baumann U, Narberhaus F. Structure and function of the bacterial AAA protease FtsH. Biochim Biophys Acta. 2012;1823(1):40-8.

73. Thorsness PE, White KH, Fox TD. Inactivation of YME1, a member of the ftsH-SEC18-PAS1-CDC48 family of putative ATPase-encoding genes, causes increased escape of DNA from mitochondria in Saccharomyces cerevisiae. Mol Cell Biol. 1993;13(9):5418-26.

74. Guelin E, Rep M, Grivell LA. Sequence of the AFG3 gene encoding a new member of the FtsH/Yme1/Tma subfamily of the AAA-protein family. Yeast. 1994;10(10):1389-94.

75. Tauer R, Mannhaupt G, Schnall R, Pajic A, Langer T, Feldmann H. Yta10p, a Member of a Novel Atpase Family in Yeast, Is Essential for Mitochondrial-Function. Febs Letters. 1994;353(2):197-200.

76. Tzagoloff A, Yue J, Jang J, Paul MF. A New Member of a Family of Atpases Is Essential for Assembly of Mitochondrial Respiratory-Chain and Atp-Synthetase Complexes in Saccharomyces-Cerevisiae. Journal of Biological Chemistry. 1994;269(42):26144-51.

77. Stahlberg H, Kutejova E, Suda K, Wolpensinger B, Lustig A, Schatz G, et al. Mitochondrial Lon of Saccharomyces cerevisiae is a ring-shaped protease with seven flexible subunits. P Natl Acad Sci USA. 1999;96(12):6787-90.

78. Kim YI, Levchenko I, Fraczkowska K, Woodruff RV, Sauer RT, Baker TA. Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. Nat Struct Biol. 2001;8(3):230-3.

79. Kress W, Mutschler H, Weber-Ban E. Both ATPase Domains of ClpA Are Critical for Processing of Stable Protein Structures. Journal of Biological Chemistry. 2009;284(45):31441-52.

80. Wang T, Darwin KH, Li H. Binding-induced folding of prokaryotic ubiquitin-like protein on the Mycobacterium proteasomal ATPase targets substrates for degradation. Nat Struct Mol Biol. 2010;17(11):1352-7.

81. Martin A, Baker TA, Sauer RT. Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. Nat Struct Mol Biol. 2008;15(11):1147-51.

82. Glynn SE, Martin A, Nager AR, Baker TA, Sauer RT. Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. Cell. 2009;139(4):744-56.

83. Bieniossek C, Niederhauser B, Baumann UM. The crystal structure of apo-FtsH reveals domain movements necessary for substrate unfolding and translocation. Proc Natl Acad Sci U S A. 2009;106(51):21579-84.

84. Bieniossek C, Schalch T, Bumann M, Meister M, Meier R, Baumann U. The molecular architecture of the metalloprotease FtsH. Proc Natl Acad Sci U S A. 2006;103(9):3066-71.

85. Suno R, Niwa H, Tsuchiya D, Zhang X, Yoshida M, Morikawa K. Structure of the whole cytosolic region of ATP-dependent protease FtsH. Mol Cell. 2006;22(5):575-85.

86. Karlberg T, van den Berg S, Hammarstrom M, Sagemark J, Johansson I, Holmberg-Schiavone L, et al. Crystal structure of the ATPase domain of the human AAA+ protein paraplegin/SPG7. PLoS One. 2009;4(10):e6975.

87. Matyskiela ME, Lander GC, Martin A. Conformational switching of the 26S proteasome enables substrate degradation. Nat Struct Mol Biol. 2013;20(7):781-8.

88. Wohlever ML, Nager AR, Baker TA, Sauer RT. Engineering fluorescent protein substrates for the AAA+ Lon protease. Protein Eng Des Sel. 2013;26(4):299-305.

89. Gur E, Vishkautzan M, Sauer RT. Protein unfolding and degradation by the AAA plus Lon protease. Protein Science. 2012;21(2):268-+.

90. Gur E, Sauer RT. Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine. Proc Natl Acad Sci U S A. 2009;106(44):18503-8.

91. Kim YI, Burton RE, Burton BM, Sauer RT, Baker TA. Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. Molecular Cell. 2000;5(4):639-48.

92. Flynn JM, Levchenko I, Seidel M, Wickner SH, Sauer RT, Baker TA. Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. P Natl Acad Sci USA. 2001;98(19):10584-9.

93. Higashitani A, Ishii Y, Kato Y, Koriuchi K. Functional dissection of a cell-division inhibitor, SulA, of Escherichia coli and its negative regulation by Lon. Mol Gen Genet. 1997;254(4):351-7.

94. Gur E, Sauer RT. Recognition of misfolded proteins by Lon, a AAA(+) protease. Genes Dev. 2008;22(16):2267-77.

95. Varshavsky A. The N-end rule pathway and regulation by proteolysis. Protein Science. 2011;20(8):1298-345.

96. Striebel F, Imkamp F, Ozcelik D, Weber-Ban E. Pupylation as a signal for proteasomal degradation in bacteria. Bba-Mol Cell Res. 2014;1843(1):103-13.

97. Bota DA, Davies KJA. Lon protease preferentially degrades oxidized mitochondrial aconitase by an ATP-stimulated mechanism. Nat Cell Biol. 2002;4(9):674-80.

98. Bayot A, Gareil M, Rogowska-Wrzesinska A, Roepstorff P, Friguet B, Bulteau AL. Identification of Novel Oxidized Protein Substrates and Physiological Partners of the Mitochondrial ATP-dependent Lon-like Protease Pim1. Journal of Biological Chemistry. 2010;285(15):11445-57.

99. Baker MJ, Tatsuta T, Langer T. Quality control of mitochondrial proteostasis. Cold Spring Harb Perspect Biol. 2011;3(7).

100. Al-Furoukh N, Kardon JR, Kruger M, Szibor M, Baker TA, Braun T. NOA1, a novel ClpXP substrate, takes an unexpected nuclear detour prior to mitochondrial import. PLoS One. 2014;9(7):e103141.

101. Kardon JR, Yien YY, Huston NC, Branco DS, Hildick-Smith GJ, Rhee KY, et al. Mitochondrial ClpX Activates a Key Enzyme for Heme Biosynthesis and Erythropoiesis. Cell. 2015;161(4):858-67.

102. Fischer F, Langer JD, Osiewacz HD. Identification of potential mitochondrial CLPXP protease interactors and substrates suggests its central role in energy metabolism. Sci Rep. 2015;5:18375.

103. Seo JH, Rivadeneira DB, Caino MC, Chae YC, Speicher DW, Tang HY, et al. The Mitochondrial Unfoldase-Peptidase Complex ClpXP Controls Bioenergetics Stress and Metastasis. PLoS Biol. 2016;14(7):e1002507.

104. Shah ZH, Hakkaart GA, Arku B, de Jong L, van der Spek H, Grivell LA, et al. The human homologue of the yeast mitochondrial AAA metalloprotease Yme1p complements a yeast yme1 disruptant. FEBS Lett. 2000;478(3):267-70.

105. Weber ER, Hanekamp T, Thorsness PE. Biochemical and functional analysis of the YME1 gene product, an ATP and zinc-dependent mitochondrial protease from S. cerevisiae. Mol Biol Cell. 1996;7(2):307-17.

106. Arlt H, Tauer R, Feldmann H, Neupert W, Langer T. The YTA10-12 complex, an AAA protease with chaperone-like activity in the inner membrane of mitochondria. Cell. 1996;85(6):875-85.

107. Atorino L, Silvestri L, Koppen M, Cassina L, Ballabio A, Marconi R, et al. Loss of m-AAA protease in mitochondria causes complex I deficiency and increased sensitivity to oxidative stress in hereditary spastic paraplegia. J Cell Biol. 2003;163(4):777-87.

108. Leonhard K, Herrmann JM, Stuart RA, Mannhaupt G, Neupert W, Langer T. AAA proteases with catalytic sites on opposite membrane surfaces comprise a proteolytic system for the ATP-dependent degradation of inner membrane proteins in mitochondria. Embo J. 1996;15(16):4218-29.

109. Leonhard K, Guiard B, Pellecchia G, Tzagoloff A, Neupert W, Langer T. Membrane protein degradation by AAA proteases in mitochondria: extraction of substrates from either membrane surface. Mol Cell. 2000;5(4):629-38.

110. Klanner C, Prokisch H, Langer T. MAP-1 and IAP-1, two novel AAA proteases with catalytic sites on opposite membrane surfaces in mitochondrial inner membrane of Neurospora crassa. Molecular Biology of the Cell. 2001;12(9):2858-69.

111. Casari G, De Fusco M, Ciarmatori S, Zeviani M, Mora M, Fernandez P, et al. Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. Cell. 1998;93(6):973-83.

112. Claypool SM, Whited K, Srijumnong S, Han XL, Koehler CM. Barth syndrome mutations that cause tafazzin complex lability. Journal of Cell Biology. 2011;192(3):447-62.

113. Gaspard GJ, McMaster CR. The Mitochondrial Quality Control Protein Yme1 Is Necessary to Prevent Defective Mitophagy in a Yeast Model of Barth Syndrome. Journal of Biological Chemistry. 2015;290(14):9284-98.

114. Thorsness PE, Fox TD. Nuclear Mutations in Saccharomyces-Cerevisiae That Affect the Escape of DNA from Mitochondria to the Nucleus. Genetics. 1993;134(1):21-8.

115. Nakai T, Yasuhara T, Fujiki Y, Ohashi A. Multiple genes, including a member of the AAA family, are essential for degradation of unassembled subunit 2 of cytochrome c oxidase in yeast mitochondria. Mol Cell Biol. 1995;15(8):4441-52.

116. Baker MJ, Mooga VP, Guiard B, Langer T, Ryan MT, Stojanovski D. Impaired folding of the mitochondrial small TIM chaperones induces clearance by the i-AAA protease. J Mol Biol. 2012;424(5):227-39.

117. Spiller MP, Guo L, Wang Q, Tran P, Lu H. Mitochondrial Tim9 protects Tim10 from degradation by the protease Yme1. Biosci Rep. 2015;35(3).

118. Augustin S, Nolden M, Muller S, Hardt O, Arnold I, Langer T. Characterization of peptides released from mitochondria: evidence for constant proteolysis and peptide efflux. J Biol Chem. 2005;280(4):2691-9.

119. Nebauer R, Schuiki I, Kulterer B, Trajanoski Z, Daum G. The phosphatidylethanolamine level of yeast mitochondria is affected by the mitochondrial components Oxa1p and Yme1p. Febs Journal. 2007;274(23):6180-90.

120. Potting C, Wilmes C, Engmann T, Osman C, Langer T. Regulation of mitochondrial phospholipids by Ups1/PRELI-like proteins depends on proteolysis and Mdm35. Embo J. 2010;29(17):2888-98.

121. Potting C, Tatsuta T, Konig T, Haag M, Wai T, Aaltonen MJ, et al. TRIAP1/PRELI complexes prevent apoptosis by mediating intramitochondrial transport of phosphatidic acid. Cell Metab. 2013;18(2):287-95.

122. Wang K, Jin M, Liu X, Klionsky DJ. Proteolytic processing of Atg32 by the mitochondrial i-AAA protease Yme1 regulates mitophagy. Autophagy. 2013;9(11):1828-36.

123. Griparic L, Kanazawa T, van der Bliek AM. Regulation of the mitochondrial dynamin-like protein Opa1 by proteolytic cleavage. J Cell Biol. 2007;178(5):757-64.

124. Song ZY, Chen HC, Fiket M, Alexander C, Chan DC. OPA1 processing controls mitochondrial fusion and is regulated by mRNA splicing, membrane potential, and Yme1L. Journal of Cell Biology. 2007;178(5):749-55.

125. Guillery O, Malka F, Landes T, Guillou E, Blackstone C, Lombes A, et al. Metalloprotease-mediated OPA1 processing is modulated by the mitochondria) membrane potential. Biol Cell. 2008;100(5):315-25.

126. Head B, Griparic L, Amiri M, Gandre-Babbe S, van der Bliek AM. Inducible proteolytic inactivation of OPA1 mediated by the OMA1 protease in mammalian cells. J Cell Biol. 2009;187(7):959-66.

127. Anand R, Wai T, Baker MJ, Kladt N, Schauss AC, Rugarli E, et al. The i-AAA protease YME1L and OMA1 cleave OPA1 to balance mitochondrial fusion and fission. J Cell Biol. 2014;204(6):919-29.

128. Mishra P, Carelli V, Manfredi G, Chan DC. Proteolytic cleavage of Opa1 stimulates mitochondrial inner membrane fusion and couples fusion to oxidative phosphorylation. Cell Metab. 2014;19(4):630-41.

129. Wai T, Garcia-Prieto J, Baker MJ, Merkwirth C, Benit P, Rustin P, et al. Imbalanced OPA1 processing and mitochondrial fragmentation cause heart failure in mice. Science. 2015;350(6265):aad0116.

130. Rainbolt TK, Atanassova N, Genereux JC, Wiseman RL. Stress-regulated translational attenuation adapts mitochondrial protein import through Tim17A degradation. Cell Metab. 2013;18(6):908-19.

131. Rainbolt TK, Lebeau J, Puchades C, Wiseman RL. Reciprocal Degradation of YME1L and OMA1 Adapts Mitochondrial Proteolytic Activity during Stress. Cell Rep. 2016;14(9):2041-9.

132. Rainey RN, Glavin JD, Chen HW, French SW, Teitell MA, Koehler CM. A new function in translocation for the mitochondrial i-AAA protease Yme1: Import of polynucleotide phosphorylase into the intermembrane space. Mol Cell Biol. 2006;26(22):8488-97.

133. Scharfenberg F, Serek-Heuberger J, Coles M, Hartmann MD, Habeck M, Martin J, et al. Structure and evolution of N-domains in AAA metalloproteases. J Mol Biol. 2015;427(4):910-23.

134. Ramelot TA, Yang Y, Sahu ID, Lee HW, Xiao R, Lorigan GA, et al. NMR structure and MD simulations of the AAA protease intermembrane space domain indicates peripheral membrane localization within the hexaoligomer. FEBS Lett. 2013;587(21):3522-8.

135. Makino S, Makino T, Abe K, Hashimoto J, Tatsuta T, Kitagawa M, et al. Second transmembrane segment of FtsH plays a role in its proteolytic activity and homo-oligomerization. FEBS Lett. 1999;460(3):554-8.

136. Donaldson LW, Wojtyra U, Houry WA. Solution structure of the dimeric zinc binding domain of the chaperone ClpX. Journal of Biological Chemistry. 2003;278(49):48991-6.

137. Chowdhury T, Chien P, Ebrahim S, Sauer RT, Baker TA. Versatile modes of peptide recognition by the ClpX N domain mediate alternative adaptor-binding specificities in different bacterial species. Protein Sci. 2010;19(2):242-54.

138. Erbse AH, Wagner JN, Truscott KN, Spall SK, Kirstein J, Zeth K, et al. Conserved residues in the N-domain of the AAA+ chaperone ClpA regulate substrate recognition and unfolding. FEBS J. 2008;275(7):1400-10.

139. Fuhrer F, Muller A, Baumann H, Langklotz S, Kutscher B, Narberhaus F. Sequence and length recognition of the C-terminal turnover element of LpxC, a soluble substrate of the membrane-bound FtsH protease. J Mol Biol. 2007;372(2):485-96.

140. Gonzalez M, Frank EG, Levine AS, Woodgate R. Lon-mediated proteolysis of the Escherichia coli UmuD mutagenesis protein: in vitro degradation and identification of residues required for proteolysis. Gene Dev. 1998;12(24):3889-99.

141. Bittner LM, Westphal K, Narberhaus F. Conditional Proteolysis of the Membrane Protein YfgM by the FtsH Protease Depends on a Novel N-terminal Degron. J Biol Chem. 2015;290(31):19367-78.

142. Arends J, Thomanek N, Kuhlmann K, Marcus K, Narberhaus F. In vivo trapping of FtsH substrates by label-free quantitative proteomics. Proteomics. 2016.

143. Leonhard K, Stiegler A, Neupert W, Langer T. Chaperone-like activity of the AAA domain of the yeast Yme1 AAA protease. Nature. 1999;398(6725):348-51.

144. Wah DA, Levchenko I, Rieckhof GE, Bolon DN, Baker TA, Sauer RT. Flexible linkers leash the substrate binding domain of SspB to a peptide module that stabilizes delivery complexes with the AAA plus ClpXP protease. Molecular Cell. 2003;12(2):355-63.

145. Dougan DA, Weber-Ban E, Bukau B. Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA plus protein ClpX. Molecular Cell. 2003;12(2):373-80.

146. Dougan DA, Reid BG, Horwich AL, Bukau B. ClpS, a substrate modulator of the ClpAP machine. Molecular Cell. 2002;9(3):673-83.

147. Kirstein J, Schlothauer T, Dougan DA, Lilie H, Tischendorf G, Mogk A, et al. Adaptor protein controlled oligomerization activates the AAA plus protein ClpC. Embo J. 2006;25(7):1481-91.

148. Dunn CD, Lee MS, Spencer FA, Jensen RE. A genomewide screen for petite-negative yeast strains yields a new subunit of the i-AAA protease complex. Mol Biol Cell. 2006;17(1):213-26.

149. Dunn CD, Tamura Y, Sesaki H, Jensen RE. Mgr3p and Mgr1p are adaptors for the mitochondrial i-AAA protease complex. Mol Biol Cell. 2008;19(12):5387-97.

150. Herman C, Prakash S, Lu CZ, Matouschek A, Gross CA. Lack of a robust unfoldase activity confers a unique level of substrate specificity to the universal AAA protease FtsH. Mol Cell. 2003;11(3):659-69.

151. Lee S, Augustin S, Tatsuta T, Gerdes F, Langer T, Tsai FT. Electron cryomicroscopy structure of a membrane-anchored mitochondrial AAA protease. J Biol Chem. 2011;286(6):4404-11.

152. Graef M, Seewald G, Langer T. Substrate recognition by AAA+ ATPases: distinct substrate binding modes in ATP-dependent protease Yme1 of the mitochondrial intermembrane space. Mol Cell Biol. 2007;27(7):2476-85.

153. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 1988;16(22):10881-90.

154. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014;42(Web Server issue):W320-4.

155. Zaccai NR, Chi B, Thomson AR, Boyle AL, Bartlett GJ, Bruning M, et al. A de novo peptide hexamer with a mutable channel. Nat Chem Biol. 2011;7(12):935-41.

156. Norby JG. Coupled assay of Na+,K+-ATPase activity. Methods Enzymol. 1988;156:116-9.

157. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-5.

158. Koodathingal P, Jaffe NE, Kraut DA, Prakash S, Fishbain S, Herman C, et al. ATP-dependent proteases differ substantially in their ability to unfold globular proteins. J Biol Chem. 2009;284(28):18674-84.

159. Kenan DJ, Strittmatter WJ, Burke JR. Phage display screening for peptides that inhibit polyglutamine aggregation. Method Enzymol. 2006;413:253-73.

160. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography, Pt A. 1997;276:307-26.

161. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D. 2010;66:213-21.

162. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D. 2004;60:2126-32.

163. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D. 2010;66:12-21.

164. Akoev V, Gogol EP, Barnett ME, Zolkiewski M. Nucleotide-induced switch in oligomerization of the AAA+ ATPase ClpB. Protein Sci. 2004;13(3):567-74.

165. Mogk A, Schlieker C, Strub C, Rist W, Weibezahn J, Bukau B. Roles of individual domains and conserved motifs of the AAA+ chaperone ClpB in oligomerization, ATP hydrolysis, and chaperone activity. J Biol Chem. 2003;278(20):17615-24.

166. Lin J, Lucius AL. Examination of the dynamic assembly equilibrium for E. coli ClpB. Proteins. 2015;83(11):2008-24.

167. Akiyama Y, Ito K. Roles of multimerization and membrane association in the proteolytic functions of FtsH (HflB). Embo J. 2000;19(15):3888-95.

168. Akiyama Y, Ito K. Roles of homooligomerization and membrane association in ATPase and proteolytic activities of FtsH in vitro. Biochemistry. 2001;40(25):7687-93.

169. Beckwith R, Estrin E, Worden EJ, Martin A. Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA+ unfoldase. Nat Struct Mol Biol. 2013;20(10):1164-72.

170. Asahara Y, Atsuta K, Motohashi K, Taguchi H, Yohda M, Yoshida M. FtsH recognizes proteins with unfolded structure and hydrolyzes the carboxyl side of hydrophobic residues. J Biochem. 2000;127(5):931-7.

171. Baker TA, Sauer RT. ClpXP, an ATP-powered unfolding and protein-degradation machine. Biochim Biophys Acta. 2012;1823(1):15-28.

172. Gur E, Sauer RT. Evolution of the ssrA degradation tag in Mycoplasma: specificity switch to a different protease. Proc Natl Acad Sci U S A. 2008;105(42):16113-8.

173. Kenniston JA, Baker TA, Fernandez JM, Sauer RT. Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. Cell. 2003;114(4):511-20.

174. Lee C, Schwartz MP, Prakash S, Iwakura M, Matouschek A. ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. Molecular Cell. 2001;7(3):627-37.

175. Shin Y, Davis JH, Brau RR, Martin A, Kenniston JA, Baker TA, et al. Single-molecule denaturation and degradation of proteins by the AAA plus ClpXP protease. P Natl Acad Sci USA. 2009;106(46):19340-5.

176. Cipriano DJ, Jung J, Vivona S, Fenn TD, Brunger AT, Bryant Z. Processive ATP-driven Substrate Disassembly by the N-Ethylmaleimide-sensitive Factor (NSF) Molecular Machine. Journal of Biological Chemistry. 2013;288(32):23436-45.

177. Parsell DA, Sauer RT. The Structural Stability of a Protein Is an Important Determinant of Its Proteolytic Susceptibility in Escherichia-Coli. Journal of Biological Chemistry. 1989;264(13):7590-5.

178. Topell S, Hennecke J, Glockshuber R. Circularly permuted variants of the green fluorescent protein. Febs Letters. 1999;457(2):283-9.

179. Nager AR, Baker TA, Sauer RT. Stepwise Unfolding of a beta Barrel Protein by the AAA plus ClpXP Protease. Journal of Molecular Biology. 2011;413(1):4-16.

180. Abo T, Inada T, Ogawa K, Aiba H. SsrA-mediated tagging and proteolysis of LacI and its role in the regulation of lac operon. Embo J. 2000;19(14):3762-9.

181. Chiba S, Akiyama Y, Mori H, Matsuo E, Ito K. Length recognition at the N-terminal tail for the initiation of FtsH-mediated proteolysis. Embo Rep. 2000;1(1):47-52.

182. Osman C, Haag M, Potting C, Rodenfels J, Dip PV, Wieland FT, et al. The genetic interactome of prohibitins: coordinated control of cardiolipin and phosphatidylethanolamine by conserved regulators in mitochondria. J Cell Biol. 2009;184(4):583-96.

183. Tamura Y, Endo T, Iijima M, Sesaki H. Ups1p and Ups2p antagonistically regulate cardiolipin metabolism in mitochondria. J Cell Biol. 2009;185(6):1029-45.

184. Miliara X, Garnett JA, Tatsuta T, Abid Ali F, Baldie H, Perez-Dorado I, et al. Structural insight into the TRIAP1/PRELI-like domain family of mitochondrial phospholipid transfer complexes. Embo Rep. 2015;16(7):824-35.

185. Yu F, He FY, Yao HY, Wang CY, Wang JC, Li JX, et al. Structural basis of intramitochondrial phosphatidic acid transport mediated by Ups1-Mdm35 complex. Embo Rep. 2015;16(7):813-23.

186. Ting SY, Schilke BA, Hayashi M, Craig EA. Architecture of the TIM23 Inner Mitochondrial Translocon and Interactions with the Matrix Import Motor. Journal of Biological Chemistry. 2014;289(41):28689-96.

187. Alder NN, Sutherland J, Buhring AI, Jensen RE, Johnson AE. Quaternary structure of the mitochondrial TIM23 complex reveals dynamic association between Tim23p and other subunits. Molecular Biology of the Cell. 2008;19(1):159-70.

188. Patel S, Latterich M. The AAA team: related ATPases with diverse functions. Trends Cell Biol. 1998;8(2):65-71.

189. Korbel D, Wurth S, Kaser M, Langer T. Membrane protein turnover by the m-AAA protease in mitochondria depends on the transmembrane domains of its subunits. Embo Rep. 2004;5(7):698-703.

190. Watanabe Y, Tamura Y, Kawano S, Endo T. Structural and mechanistic insights into phospholipid transfer by Ups1-Mdm35 in mitochondria. Nat Commun. 2015;6.

191. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 2010;38(Web Server issue):W545-9.

192. Zhang X, Wigley DB. The 'glutamate switch' provides a link between ATPase activity and ligand binding in AAA+ proteins. Nat Struct Mol Biol. 2008;15(11):1223-7.

193. Augustin S, Gerdes F, Lee S, Tsai FT, Langer T, Tatsuta T. An intersubunit signaling network coordinates ATP hydrolysis by m-AAA proteases. Mol Cell. 2009;35(5):574-85.

194. Hartmann B, Wai T, Hu H, MacVicar T, Musante L, Fischer-Zirnsak B, et al. Homozygous YME1L1 Mutation Causes Mitochondriopathy with Optic Atrophy and Mitochondrial Network Fragmentation. Elife. 2016;5.