Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Mia Shin¹⁻³, Edmond R. Watson¹, Scott J. Novick⁴, Patrick Griffin⁴, R. Luke Wiseman^{2*}, Gabriel C. Lander^{1*}

¹Department of Integrative Structural and Computational Biology, Scripps Research, La Jolla, CA 92037

²Department of Molecular Medicine, Scripps Research, La Jolla, CA 92037

³Skaggs Graduate School of Chemical and Biological Sciences, Scripps Research, La Jolla, CA 92037

⁴Department of Molecular Medicine, The Scripps Research Institute, Jupiter, FL 33458

Abstract

The human mitochondrial AAA+ protein LONP1 is a critical quality control protease involved in regulating diverse aspects of mitochondrial biology including proteostasis, electron transport chain activity, and mitochondrial transcription. As such, genetic or aging-associated imbalances in LONP1 activity are implicated in the pathologic mitochondrial dysfunction associated with numerous human diseases. Despite this importance, the molecular basis for LONP1-dependent proteolytic activity remains poorly defined. Here, we solved cryo-electron microscopy structures of human LONP1 to reveal the molecular mechanism of substrate proteolysis. We show that, like bacterial Lon, human LONP1 adopts both an open and closed spiral staircase orientation dictated by the presence of substrate and nucleotide. However, unlike bacterial Lon, human LONP1 contains a second spiral staircase within its ATPase domain that engages substrate to increase interactions with the translocating peptide as it transits into the proteolytic chamber for proteolysis. Further, we show that substrate-bound LONP1 includes a second level of regulation at the proteolytic active site, wherein autoinhibition of the active site is only relieved by the presence of a peptide substrate. Ultimately, our results define a structural basis for human LONP1 proteolytic activation and activity, establishing a molecular framework to understand the critical importance of this protease for mitochondrial regulation in health and disease.

Introduction

Mitochondria are the site of essential cellular functions including oxidative phosphorylation, apoptotic signaling, calcium regulation, and iron-sulfur cluster biogenesis. The maintenance of mitochondrial protein homeostasis (or proteostasis) is critical for mammalian cell viability^{1,2}. To prevent the accumulation of misfolded, aggregated, or damaged proteins, mitochondria evolved a network of quality control proteases that function to regulate mitochondrial proteostasis in the presence and absence of cellular stimuli^{1,3}. This includes the ATP-dependent protease LONP1 – a protease that is conserved throughout evolution from bacteria to humans^{1,3-5}. In mammals, LONP1 is a primary quality control protease responsible for regulating proteostasis and function within the mitochondrial matrix through diverse mechanisms including the degradation of misfolded and oxidatively damaged proteins, components of the electron transport chain, and mtDNA regulatory factors such as TFAM and POLG⁴⁻¹⁰. The importance of LON in regulating mitochondria is underscored by the fact that homozygous deletion of *Lonp1* in mice is embryonic lethal¹¹. Further, imbalances in LONP1 activity are implicated in mitochondrial dysfunction associated with diverse pathologic conditions including organismal aging, cancer, and numerous other human diseases^{1,4,11-14}. Mutations in *LONP1* have also been identified to be causatively associated with cerebral, ocular, dental, auricular, skeletal (CODAS) syndrome, a multi-system developmental disorder consisting

of a wide array of clinical manifestations including hypotonia, ptosis, motor delay, hearing loss, postnatal cataracts, and skeletal abnormalities¹⁵⁻¹⁷. Despite the central importance of LONP1 for regulating mitochondria in health and disease, the structural basis for human LONP1 proteolytic activation and activity remains poorly defined.

The LONP1 protease consists of an N-terminal domain involved in substrate recognition and oligomerization, a AAA+ (ATPases associated with diverse cellular activities) domain that powers substrate translocation, and a C-terminal serine protease domain involved in substrate proteolysis^{1,4,18}. Numerous biophysical approaches have been employed to gain insight into the macromolecular structure and function of bacterial Lon¹⁹⁻²⁷, and recent cryo-electron microscopy (cryo-EM) studies of bacterial Lon in both substrate-free and substrate-bound conformations revealed the structural basis for substrate processing and protease activation²⁸. However, few structural studies have been directed towards understanding the molecular mechanism of protease activation and substrate processing by human LONP1. A prior crystal structure of the isolated protease domain of human LONP1 (PDB:2X36) revealed the structural conservation between human and bacterial Lon protease domains, suggesting similar mechanisms of action²⁹. Further, a low resolution cryo-EM structure of human LONP1 showed the ATPase and protease domains organized to form a pseudo six-fold symmetric chamber, while the N-terminal domains are arranged as a trimer of dimers atop the AAA+ domains³⁰. Given the complex

*Correspondence to: wiseman@scripps.edu, glander@scripps.edu

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Figure 1. Architectures of substrate-bound and substrate-free human mitochondrial LONP1 protease. **a.** Lateral view of the substrate-bound human LONP1 atomic model in a left-handed spiral configuration (center) flanked by axial exterior views of the ATPase (left) and protease (right) domain rings. Each subunit of the homo-hexamer is assigned a distinct color depending on its position in the spiral staircase, with the protease domains colored a lighter hue than the ATPase domains. The same coloring is used throughout the figure. The cryo-EM density of the substrate is shown as a solid isosurface colored orange. **b.** The relative height of individual protomers within the substrate-bound conformer are shown by aligning all subunits to a common view, and accentuated by dashed lines shown above the NTD^{3H}. **c.** The same orthogonal views in (**a**) are shown for the substrate-free human LONP1 atomic model. Each subunit of the homo-hexamer is assigned a distinct color depending on its position in the spiral staircase, **d.** The individual protomers of substrate-free human LONP1 lined up side-by-side, showing the descending positions of the domains in the structure.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

proteolytic demands of the mitochondrial matrix, we aimed to define the structural and mechanistic differences from bacterial Lon that have evolved to improve the activity or regulation of this protease for mitochondrial regulation.

To address this deficiency, we solved cryo-EM structures of human LONP1 in the presence and absence of a translocating substrate. We show that in the absence of substrate, LONP1 adopts an open, left-handed spiral conformation that is bound to ADP and has its protease active sites organized in an auto-inhibited conformation. In contrast, in the presence of substrate, the AAA+ domains of human LONP1 adopt a right-handed spiral configuration similar to that observed for many other AAA+ proteases^{18,31,32}. Surprisingly, and unlike other structures of AAA+ proteases, including bacterial Lon²⁸, we found that the protease domains in substrate-bound human LONP1 adopt an asymmetric arrangement wherein the protease active sites remain auto-inhibited. This asymmetric arrangement of the protease domains could be relieved by administration of the small molecule LONP1 inhibitor bortezomib, which covalently engages the proteolytic active site and gives rise to the six-fold symmetric arrangement of active protease domains, similar to that observed for bacterial Lon²⁸. This bortezomib-induced release of LONP1 protease auto-inhibition suggests a second level of regulation for protease activation that has evolved to tune LONP1-dependent regulation of mitochondrial proteostasis and function.

Results

Substrate-bound and substrate-free structures of human LONP1

Mature human LONP1 lacking its N-terminal mitochondrial targeting sequence was incubated with saturating amounts of ATP γ S (1 mM), a slowly hydrolyzing ATP analog, to stabilize the complex for structural studies. Single-particle cryo-EM analyses resulted in reconstructions of two distinct conformations of the essential protease: one structure containing a peptide substrate trapped in its central channel and the other devoid of substrate (**Fig. 1** and **Supplementary Fig. 1**). The presence of substrate in a subset of the complexes was attributed to co-purification of an endogenous protein substrate, as has been observed previously in cryo-EM studies of numerous other AAA+ proteins^{18,31-37}. In both structures of human LONP1, residues 420-785, comprising a portion of the N-terminal helical domain (NTD^{3H}), the AAA+ domain consisting of the small and large ATPase subdomains, and the protease domain, were well-resolved (**Fig. 1**); however the entirety of the 300-residue N-terminal domain was not resolvable in either structure, likely due to the inherent flexibility of this region.

The substrate-free human LONP1 reconstruction identified in our dataset was resolved to 3.4 Å resolution, which was sufficient for atomic modeling for five of the six LONP1 protomers (**Supplementary Fig. 1** and **Supplementary Table 1**). The organization is reminiscent of the left-handed, open lockwasher configurations previously observed for substrate-

free bacterial Lon^{26,28}, although human LONP1 appears to oligomerize with a slightly smaller helical pitch than the bacterial forms (**Fig. 1c,d** and **Supplementary Fig. 2**). The uppermost subunit of substrate-free human LONP hexamer demonstrates substantially more flexibility than its bacterial counterpart, as it is only discernible at low resolution in a subset of the particles (**Supplementary Fig. 1**), suggesting that oligomerization of the N-terminal domains into a trimer-of-dimers³⁰ may play a role in maintaining the hexamer's oligomeric state in the absence of substrate. Due to the low resolution of this uppermost subunit, a copy of the middle protomer from the five-subunit spiral was rigid body fit into the reconstruction for representational purposes. Similar to previous substrate-free structures of bacterial Lon^{26,28}, the subunits of human LONP1 in the absence of substrate are bound to ADP and all protease subdomains adopt an auto-inhibited conformation (**Supplementary Fig. 3**). This substrate-free structure likely represents a conformation important for substrate engagement and release, as suggested for bacterial Lon²⁸. According to this mechanism, the topmost subunit of the left-handed spiral is the first to undergo nucleotide exchange and engage substrate, followed by subsequent nucleotide and substrate engagement events that lead to a reorganization of the AAA+ domains within the hexamer into an active state that is competent for substrate translocation. The similarity of the substrate-free human and bacterial structures indicates that this model for substrate engagement is conserved.

The substrate-bound LONP1 structure was resolved to 3.2 Å resolution, which enabled atomic model building and refinement (**Supplementary Fig. 1** and **Supplementary Table 1**). The structural conservation of the ATPase domains from bacteria to human is considerable, as the protomers of the human LONP1 are in a nearly identical organization as their counterpart in *Y. pestis* Lon, with RMSDs between 0.8-0.9 Å (**Supplementary Fig. 4a**). The overall organization of these ATPase domains in the substrate-bound human LONP1 generally resembles that of previously determined substrate-bound AAA+ proteins, wherein the ATPases form a closed spiral staircase around a centrally positioned substrate peptide.

However, our substrate-bound human LONP1 exhibits several notable structural differences from the previously determined substrate-bound *Y. pestis* Lon. Whereas the AAA+ ring of bacterial Lon consists of four descending ATPase domains in its spiral staircase with two ascending 'seam' subunits (**Supplementary Fig. 5**), our substrate-bound structure of human LONP1 contains five descending ATPase domain subunits with a single 'seam' subunit transitioning between the highest and lowest subunits (**Fig. 1a,b**). This conformation is consistent with substrate-bound structures of AAA+ protein translocases involved in a wide range of biological functions from bacteria to mammals^{18,32-37}. Accordingly, the four topmost subunits of the staircase are bound to ATP γ S (hereafter referred to as ATP for simplicity) while the lowest contains ADP in its nucleotide binding site (**Fig. 1a,b** and **Supplementary Fig. 6**). Density corresponding to an ADP

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Figure 2. LONP1 employs two pore loop aromatic residues to facilitate substrate translocation. **a.** A twelve-residue polyalanine chain is modeled into the substrate density found in the substrate-bound human LONP1 structure, shown in a transparent orange surface representation. Y565 from pore-loop 1 (labeled P1) is shown using stick representations. Y565 residues from ATP1-4 and ADP1 subunits show intercalating, zipper-like interactions with substrate. Y599 from pore loop 2 (labeled P2) also shows intercalating interactions with substrate in the ATP1 (purple) and ATP2 (dark blue) subunits. **b.** Sequence alignment between LONP1 homologs highlighting evolutionary conservation within pore loop 1 and Walker B motifs, as well as the evolutionary integration of a hydrophobic residue into pore loop 2 in metazoans. **c.** Introducing pore loop mutations Y565A or Y599A shows a decrease in the degradation rate of a model substrate, FITC-casein. The hydrolysis-blocking Walker B mutation E591A is included as a control. Error bars show standard deviation from 3 independent replicates. ***indicates $p < 0.001$ relative to wildtype calculated by ANOVA. **d.** Basal-level (filled bars) and casein-stimulated (hatched bar) ATPase rates are shown for the same mutations as in (c). Error bars show standard deviation from 3 independent replicates. **indicates $p < 0.01$ and ***indicates $p < 0.001$ relative to wildtype calculated by ANOVA.

molecule is also observed in the ‘seam’ subunit. To maintain consistency with other AAA+ proteins, we will name the ATP-bound subunits, starting with the topmost subunit that is bound to ATP, ATP1-4, the lowest subunit in the spiral staircase that is ADP-bound will be referred to as ADP1, and the ‘seam’ subunit as ADP2 (Fig. 1b and Supplementary Fig. 6).

Multiple pore loop interactions facilitate substrate translocation in human LONP1

The distinct configuration of the ATPase domains observed in human LONP1, relative to bacterial Lon, notably influences interactions with the engaged substrate. AAA+ proteins engage translocating substrates through pore loop residues within the ATPase domain that extend toward the central channel of the ATPase ring¹⁸. Previous structures of substrate-engaged AAA+ proteins have shown that aromatic residues in the ATPase pore loops intercalate into the backbone of an engaged peptide substrate every two amino acids to facilitate substrate translocation¹⁸. The ATPases in our human

LONP1 reconstruction encircle a density corresponding to a 12-residue peptide substrate, which, surprisingly, is nearly twice the length of engaged substrate visualized in *Y. pestis* Lon²⁸. This observation provides an opportunity to examine the evolutionary differences in substrate translocation by Lon proteases in bacteria and human.

In human LONP1, the substrate-interacting Y565 in pore loop 1 from all four ATP-bound subunits, ATP1-4, as well as the ADP-bound subunit ADP1 at the bottom of the spiral staircase, are all observed engaging with the translocating substrate (Fig. 2a,b). This is in contrast to the substrate-interacting arrangement previously shown for bacterial Lon, where only the three uppermost ATP-bound subunits were tightly engaged with substrate, while the fourth subunit within the spiral only weakly interacts with substrate²⁸. Instead, the substrate interactions we observe in human LONP1 more closely resemble other substrate-bound structures of AAA+ proteins than those in bacterial Lon^{18,31,32}.

Notably, in addition to the pore loop 1 aromatic residue,

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Figure 3. LONP1 in the substrate-bound conformation contains an autoinhibited protease site. **a.** Surface representation of substrate-bound LONP1 showing a break in the protease domains between the highest subunit, ATP1, and its neighboring subunit, ATP2, highlighted using a dashed box. **b.** The protease domains from ATP1 and ATP2 are shown using a ribbon representation, with the auto-inhibited active sites, denoted by dashed boxes. **c,d.** Details views of the regions denoted in **(b)** are shown using a ribbon representation within a mesh representation of the EM map. A loop containing the catalytic serine (S855) is folded into a 3_{10} helix and an aspartic acid (D842) prevents the formation of the serine (S855)-lysine (K898) catalytic dyad. **e.** An alignment of the isolated protease domains from all six subunits in substrate-bound LONP1 shows that all protease domains adopt an auto-inhibited conformation.

we identified a second pore loop aromatic residue, Y599, within the AAA+ domains of human LONP1 that also engages the translocating substrate (**Fig. 2a**). This second pore loop residue assembles into a spiral that parallels the pore loop 1 aromatic, with pore loop 2 from ATP1 and ATP2 engaged with the trapped substrate (**Fig. 2a**). While the remaining subunits do not appear to engage substrate in our structure of human LONP1, the spiraling organization of these residues and their proximity to translocating substrates suggest that these subunits may facilitate substrate guidance through the central channel and into the proteolytic chamber. Interestingly, the identified interactions between the pore loop 2 aromatics of ATP1 and ATP2 and the translocating substrate extend the overall interactions between human LONP1 and substrate deeper into the central chamber. The presence of a methionine, which cannot establish intercalating interactions with substrate, at the equivalent residue in *Y. pestis* Lon (M432), is likely responsible for a lack of observable pore loop 2 interactions in its structure²⁸ (**Fig. 2b** and **Supplementary Fig. 7**).

Consistent with an important role for the pore loop 2

residue Y599 in substrate translocation in human LONP1, recombinant LONP1 containing a Y599A mutation exhibited changes in both substrate proteolysis and ATP hydrolysis. LONP1^{Y599A} showed a 66% decrease in the degradation rate of the model substrate FITC-casein (**Fig. 2c**) while showing a 36% increase in substrate-induced ATPase hydrolysis (**Fig. 2d**). This decreased degradation coupled with increased ATP consumption suggests that the Y599A mutant is a less efficient translocase compared to the wildtype enzyme, with increased likelihood of substrate “slippage” and unproductive ATP hydrolysis, as previously proposed for other AAA+ proteins³⁸. Similar results were observed with the pore loop 1 mutant LONP1^{Y565A}, indicating that these two pore loop residues are primarily important for substrate processing. Thus, inclusion of an aromatic residue, capable of intercalating into the backbone of an incoming substrate, into the pore loop 2 of human LONP may have evolved to increase interactions with the substrate to improve translocation and degradation efficiency, as well as to impart a degree of substrate specificity, within the mitochondria. Other AAA+ translocases have been shown to utilize multiple pore loops to facilitate translocation,

where it has been speculated that the integration of additional aromatic-containing pore loops into the translocation mechanism increases the ATPase motor's 'grip' on an incoming substrate^{18,31,32,34,36,39}.

Substrate-bound human LONP1 adopts an auto-inhibited protease conformation

Another striking structural divergence of our substrate-bound structure of human LONP1 from bacterial Lon is in the organization of the protease domains. Previous results from bacterial Lon showed that substrate binding in four ATPase subunits and one ATP hydrolysis event promotes a rearrangement of the protease domains into a six-fold symmetric configuration that releases the proteolytic active site from an auto-inhibited conformation, inducing the formation of a binding groove to position substrates for proteolysis²⁸. Surprisingly, despite having a peptide substrate tightly bound within the channel, the protease domains of substrate-bound human LONP1 do not adopt a six-fold symmetric configuration (**Figs. 1a and 3a**). Instead, the protease domains of the substrate bound hexamers follow the shallow spiraling trajectory of the AAA+ domains, with an opening in the proteolytic ring between ATP1 and ATP2, the two uppermost subunits of the spiral staircase (**Figs. 1b and 3a**). Further, we observe all six protease active sites to be in an auto-inhibited configuration, wherein a loop containing the catalytic serine residue (S855) is folded into a 3_{10} helix, sterically blocking access of substrates to the catalytic active site, while simultaneously an aspartic acid (D842) prevents the formation of the serine (S855)-lysine (K898) catalytic dyad, hindering substrate degradation (**Fig. 3b-e**).

Previous results for bacterial Lon and other AAA+ proteases suggested that a flexible interdomain linker allows for large scale conformational rearrangements of the AAA+ domains independent of the symmetric protease to facilitate substrate translocation^{28,33,40}. It was proposed that when the protease domains were positioned laterally alongside one another underneath the translocation-competent ATPase ring, the catalytic serine-containing loop would extend toward the neighboring protease domain to stabilize inter-subunit interactions, resulting in the formation of the six-fold symmetric protease ring with six active proteolytic sites^{28,41}. However, even though several of the protease domains in the substrate-bound human LONP1 structure position alongside one another in an identical fashion as the those observed in the substrate-bound bacterial Lon structure, the catalytic loop remains in an auto-inhibited conformation (**Supplementary Fig. 8**). Thus, our structure of substrate-bound human LONP1 suggests that the combination of a translocating substrate and lateral positioning of protease domains is insufficient to produce the six-fold symmetric activated protease conformation observed in bacterial Lon, indicating other factors must be involved in generating a proteolytically active enzyme.

Bortezomib relieves the protease auto-inhibition in substrate-bound human LONP1

One potential explanation for the lingering auto-inhibition observed in the protease active site of substrate-bound human LONP1 might be that, unlike bacterial Lon, substrate must be present to release the auto-inhibition of the protease active sites. To test this prediction, we determined a cryo-EM structure of substrate-bound human LONP1 incubated with saturating amounts of ATP γ S as before, but further incubated with a ten-fold molar excess of bortezomib, a covalent, peptidomimetic inhibitor of LON that engages the proteolytic active site^{6,25}. This complex, which we term LONP1^{Bz} to indicate the addition of bortezomib, was resolved to 3.7 Å resolution (**Supplementary Fig. 9** and **Supplementary Table 1**). Densities for the bortezomib molecules were clearly visible in each of the six active sites, positioned within the substrate binding groove and covalently attached to the catalytic serine (**Fig. 4a,d**).

Strikingly, bortezomib association with the protease active sites led to a dramatic reorganization of the LONP1 subunits, resulting in a conformation that strongly resembles the substrate-bound bacterial Lon. The LONP1^{Bz} conformation contains four ATP γ S-bound AAA+ domains arranged in a spiral, with two ADP-bound AAA+ domains adopting displaced 'seam' orientations (**Fig. 4a-c** and **Supplementary Fig. 10**). Despite differences in the spiral orientation of substrate-bound human LON and LONP1^{Bz}, both structures show a similar engagement of a 12-residue translocating peptide by pore loop 1 aromatics (**Fig. 4b**). Intriguingly, we were unable to determine the interactions between pore loop 2 aromatics and substrate in our LONP1^{Bz} structure, as these loops were not as well-ordered as they were in the substrate-bound LONP1 structure, suggesting a dissociation of substrate from these pore loop 2 residues occurs prior or after substrate association with the protease domain.

Importantly, the protease domains in LONP1^{Bz} are assembled into a six-fold symmetric organization. Further, the presence of bortezomib relieved the auto-inhibition of the protease active site in our substrate bound LONP1, allowing proper assembly of the proteolytic active site, as previously observed for the substrate-bound bacterial Lon (**Fig. 4d** and **Supplementary Fig. 11**). This suggests that the presence of a substrate in the protease active site induces rearrangement of the serine-containing catalytic loop, such that it extends toward the neighboring protease domain to allow formation of the proteolytically active S855-K898 catalytic dyad. Previous reports suggest that this catalytic loop configuration is stabilized by the highly conserved residues V809, P854, and E884 from the adjacent protease domain (**Fig. 4d** and **Supplementary Fig. 7**)^{25,28}. Consistent with an important role in protease activation, mutating these residues to alanine severely impairs human LONP1 proteolytic activity, while only minimally affecting ATP hydrolysis (**Fig. 4e-f**).

To confirm and further characterize the dynamics of protease domain configurations induced by bortezomib, we per-

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Figure 4. Bortezomib induces a six-fold symmetrization of the LONP1 protease domain and promotes activation of the proteolytic active site **a**. Two external axial views of the bortezomib- and substrate-bound human LONP1^{Bz} atomic model (left) is shown next to a cutaway lateral view of LONP1, using the same representation and coloring as in Fig. 1. Bortezomib (blue) bound within the protease active sites are shown using a sphere representation, with one bortezomib molecule denoted by a dashed box. **b**. A twelve-residue polyaniline chain is modeled into the substrate density within the central channel of the LONP1^{Bz} structure, shown as in Fig. 2. The pore loop 1 Y565 residues from subunits ATP1-4 maintain intercalating, zipper-like interactions with substrate; however, Y565 from ADP1 in LONP1^{Bz} is positioned further from substrate as compared to our substrate-bound structure (see Fig. 2). **c**. When the subunits are positioned side-by-side, the protease domains are aligned along a single plane. **d**. Bortezomib covalently attached to S855 in ATP1 is shown using a ribbon representation within a mesh representation of the cryo-EM map, where the loop containing the catalytic serine S855 is observed extended toward the neighboring protease domain and interacting with conserved residues V809, P854, and E882 from the neighboring protease domain, show in blue. **e, f**. Graphs showing rates of FITC-casein degradation (**e**) or ATPase rate (**f**) for LONP1 harboring alanine mutations in the conserved residues V809, P854, and E882 involved in stabilizing the catalytic loop in the active conformation of the protease active site. Error bars show standard deviation from 3 independent replicates. ***indicates $p < 0.001$ relative to wildtype calculated by ANOVA. **g**. A single protease domain is shown as a worm representation from the substrate-bound LONP1 structure in the absence (left) and presence (right) of bortezomib. Regions showing a decrease in D₂O exchange measured by HDX-MS upon bortezomib binding are colored red in the auto-inhibited state (signifying greater exchange) and blue in the bortezomib-bound state (signifying decreased exchange).

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Figure 5. Model for protease activation in LONP1 **a.** Substrate is engaged by the open state conformer (left), causing a transition to a translocation-competent organization of the ATPases (middle). Despite substrate engagement by the pore loops (purple) in the central channel, the protease active sites remain in an auto-inhibited conformation. This auto-inhibition is only relieved upon interaction with substrate (right), at which point the catalytic loop (red) extends away from the catalytic residues to enable substrate cleavage. **b.** Close-up views of the proteolytic active sites of substrate-free, substrate-bound, and proteolytically active LONP1 enzymes. In the auto-inhibited substrate-free and substrate-bound states, a 3_{10} helix sterically hinders substrates from entering the proteolytic active site for degradation in addition to preventing the formation of the catalytic dyad between S855 and K898. However, in the presence of substrate (represented by bortezomib in orange in the right panel) in the proteolytic active site, auto-inhibition releases as the catalytic serine-containing loop extends towards the neighboring subunit, the catalytic dyad is formed, and the protease domains activate for proteolytic cleavage.

formed hydrogen-deuterium exchange mass spectrometry (HDX-MS) in the presence and absence of bortezomib (**Fig. 4f** and **Supplementary Fig. 12**). Consistent with our structures of substrate-bound LONP1 and LONP1^{Bz}, HDX-MS shows that the primary effect of bortezomib is to induce conformational remodeling of the protease domain. Notably, peptides comprising residues within the two helices flanking the catalytic S855-containing loop and the loop itself (peptides 803-820, 836-859, and 885-909) show significant stabilization upon addition of bortezomib (**Fig. 4g** and **Supplementary Fig. 12**), reflecting the importance of these residues in maintaining the active form of proteolytic active site. These results, along with our structures of substrate-bound LONP1 and LONP1^{Bz}, support a model whereby engagement of substrate at the protease active site promotes conformational remodeling that includes inter-subunit stabilization of the auto-inhibitory loop region to promote proteolytic activation of human LONP1.

Discussion

Here, we determine structures for human LONP1 in three different states: substrate-free, substrate-bound, and both substrate- and bortezomib-bound. These structures allowed us to establish a molecular framework to define the proteolytic activation of human LONP1. Surprisingly, despite being highly conserved, our structures identify key differences between human and bacterial Lon that likely reflect evolutionary adaptations important for mitochondrial proteostasis regulation. We show that human LONP1 integrates a second pore loop residue that allows increased interactions with the translocating peptide. These increased interactions likely provide the human protease enhanced ‘grip’ on substrates, potentially improving its ability to degrade damaged or non-native proteins within the mitochondrial environment, such as oxidatively-modified proteins previously shown to be LONP1 substrates^{10,12}.

Further, we show that, unlike bacterial Lon, substrate engagement by the AAA+ domains of human LONP1 is not suf-

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

ficient to promote conformational rearrangements within the protease domain required for relief of protease auto-inhibition. Instead, formation of a substrate-binding cleft at the protease active site appears to be linked to substrate interaction with the active site, suggesting a second level of regulation for human LONP1 that could function to prevent the aberrant degradation of proteins important for mitochondrial function (Fig. 5). Indeed, given the diversity of substrates processed by human LONP1 within the mitochondrial matrix, substrate-induced activation of individual protease active sites would provide a powerfully effective means by which to increase substrate selectivity and prevent aberrant degradation of mis-targeted substrates. It is thus possible that all six protease active sites of human LONP1 are rarely, if ever, simultaneously activated for proteolytic activity. Such a system would endow human LONP1 with a level of proteolytic plasticity to enable enzymatic activity to be specifically tuned to the types and quantities of substrates present within the matrix, which continuously fluctuate in response to cellular queues.

Lastly, our structures also provide a structural framework for defining the pathologic and potentially therapeutic implications of LONP1 in human disease. For example, mapping known CODAS syndrome-associated mutations of human LONP1 onto our structure of proteolytically active LONP1^{Bz} highlights their position at interfaces between ATPase domains (Supplementary Fig. 13). This suggests that these mutations likely impact the tight subunit coordination required for ATP hydrolysis and substrate translocation. Consistent with this, CODAS mutations were previously shown to reduce LONP1 ATPase and proteolytic activities¹⁷. Given that the diversity of CODAS mutations, as well as the unique combination of mutations present in heterozygous individuals, result in a wide array of phenotypes in affected individuals, our atomic models of human LONP1 provide the structural basis for future analyses of CODAS patient mutations required to understand the diverse etiology of this disease. Further, our structures provide a molecular framework to facilitate the development of small molecule strategies to modulate LONP1 activity to mitigate pathologic mitochondrial dysfunctions associated with imbalances in LONP1 proteolysis^{14,42}.

Materials and Methods

Protein expression and purification

A gene encoding an N-terminally His₆-tagged human LONP1 with methionine 115 (the N-terminal residue of mature mitochondrial-matrix localized LONP1) as its first residue was inserted into a pET20b *E. coli* expression vector. The pET20b-His₆-LONP1 plasmid was transformed into the Rosetta 2(DE3)pLysS *E. coli* strain for recombinant protein expression. Cells were grown in 3% tryptone, 2% yeast extract, and 1% MOPS pH 7.2 with 100 µg/ml ampicillin and 25 µg/ml chloramphenicol and cultured at 37°C shaking at 220 rpm to an optical density (OD₆₀₀) of 0.8. Protein overexpression was induced with 1 mM IPTG for 16 hours at 16°C. Cells were harvested by centrifugation at 5,000 x g, resuspended

in buffer A (0.2M NaCl, 25mM Tris pH 7.5, 20% glycerol, and 2mM β-mercaptoethanol), flash-frozen and then stored at -80°C for later use. Upon thawing, cells were lysed by sonication. Cell lysate was cleared by centrifugation at 15,000 rpm and then added to TALON Cobalt resin equilibrated in Buffer A and loaded onto a gravity-flow column to allow His₆-LONP1 binding. The column was washed with buffer A containing 20 mM imidazole to remove unbound proteins. Bound proteins were eluted in 1.2 column volumes of buffer A containing 300 mM imidazole. Eluted proteins were concentrated and loaded onto a Superose 6 size exclusion column equilibrated with buffer A. Fractions containing His₆-LONP1 were pooled, concentrated, and flash-frozen for storage at -80°C. The pET20b-His₆-LONP1 plasmid served as a template for generating LONP1 mutants using the quick-change site directed mutagenesis approach. Each LONP1 mutant protein was expressed and purified as described above.

In vitro proteolysis assay

Purified human LONP1 (0.1 µM) was incubated for 15 minutes at 37°C in LONP1 activity buffer (50 mM Tris-HCl pH 8, 100 mM KCl, 10 mM MgCl₂, 1 mM DTT, and 10% Glycerol) in the presence or absence of 2.5 mM ATP. FITC-Casein (0.8 µM; Sigma-Aldrich) was added to initiate the degradation reaction. An increase of fluorescence (excitation 485 nm, emission 535 nm) resulting from free FITC molecules was monitored using a TECAN plate reader. Three biological replicates were performed for each LONP1 mutant and the data were fit to a straight line from which the slope was extracted to calculate the rate of substrate degradation. GraphPad Prism software was used for data analysis. Mean and standard error of mean (SEM) was calculated by performing column statistics. ANOVA was used to calculate p-value, as indicated.

In vitro ATP hydrolysis assay

In vitro ATP hydrolysis was carried out in LONP1 activity buffer (50 mM Tris-HCl pH 8, 100 mM KCl, 10 mM MgCl₂, 10% Glycerol, and 1 mM DTT) in the presence or absence of 2.5 mM ATP. Casein (9.7 µM) was added to the reaction mixture when measuring substrate-induced ATPase activity. Reaction components and purified LONP1 (0.1 µM) were incubated separately at 37°C for five minutes. LONP1 was added to initiate ATP hydrolysis and the amount of free inorganic phosphate at each timepoint was measured by adding Malachite Green Working Reagent (Sigma-Aldrich) and incubating components for 30 minutes at room temperature for color development before measuring absorbance at 620 nm (OD₆₂₀). Three biological replicates were performed for each LONP1 mutant in the presence and absence of substrate. Data were fit to a straight line and the slope were extracted to calculate ATP hydrolysis rates. GraphPad Prism software was used for data analysis. Mean and standard error of mean (SEM) was calculated by performing column statistics. ANOVA was used to calculate p-value, as indicated.

Hydrogen-deuterium exchange (HDX) detected by mass spectrometry (MS)

Differential HDX-MS experiments were conducted as previously described with a few modifications⁴³. Peptides were identified using tandem MS (MS/MS) with an Orbitrap mass spectrometer (Q Exactive, ThermoFisher). Product ion spectra were acquired in data-dependent mode with the top five most abundant ions selected for the product ion analysis per scan event. The MS/MS data files were submitted to Mascot (Matrix Science) for peptide identification. Peptides included in the HDX analysis peptide set had a MASCOT score greater than 20 and the MS/MS spectra were verified by manual inspection. The MASCOT search was repeated against a decoy (reverse) sequence and ambiguous identifications were ruled out and not included in the HDX peptide set.

For HDX-MS analysis, LONP1 (10 μ M) was incubated with the respective ligands at a 1:10 protein-to-ligand molar ratio for 1 h at room temperature. Next, 5 μ l of sample was diluted into 20 μ l D₂O buffer (50 mM Tris-HCl, pH 8; 75 mM KCl; 10 mM MgCl₂) and incubated for various time points (0, 10, 60, 300, and 900 s) at 4°C. The deuterium exchange was then slowed by mixing with 25 μ l of cold (4°C) 0.1M Sodium Phosphate, 50mM TCEP. Quenched samples were immediately injected into the HDX platform. Upon injection, samples were passed through an immobilized pepsin column (2mm $\times$ 2cm) at 200 μ l min⁻¹ and the digested peptides were captured on a 2mm $\times$ 1cm C₈ trap column (Agilent) and desalted. Peptides were separated across a 2.1mm $\times$ 5cm C₁₈ column (1.9 μ l Hypersil Gold, ThermoFisher) with a linear gradient of 4% - 40% CH₃CN and 0.3% formic acid, over 5 min. Sample handling, protein digestion and peptide separation were conducted at 4°C. Mass spectrometric data were acquired using an Orbitrap mass spectrometer (Exactive, ThermoFisher). HDX analyses were performed in triplicate, with single preparations of each protein ligand complex. The intensity weighted mean m/z centroid value of each peptide envelope was calculated and subsequently converted into a percentage of deuterium incorporation. This is accomplished determining the observed averages of the undeuterated and fully deuterated spectra and using the conventional formula described elsewhere⁴⁴. Statistical significance for the differential HDX data is determined by an unpaired t-test for each time point, a procedure that is integrated into the HDX Workbench software⁴⁵. Corrections for back-exchange were made on the basis of an estimated 70% deuterium recovery, and accounting for the known 80% deuterium content of the deuterium exchange buffer.

For data rendering, HDX data from all overlapping peptides were consolidated to individual amino acid values using a residue averaging approach. Briefly, for each residue, the deuterium incorporation values and peptide lengths from all overlapping peptides were assembled. A weighting function was applied in which shorter peptides were weighted more heavily and longer peptides were weighted less. Each of the weighted deuterium incorporation values were then averaged to produce a single value for each amino acid. The

initial two residues of each peptide, as well as prolines, were omitted from the calculations. This approach is similar to that previously described⁴⁶. HDX analyses were performed in triplicate, with single preparations of each purified protein/complex. Statistical significance for the differential HDX data is determined by t test for each time point, and is integrated into the HDX Workbench software⁴⁵.

Sample preparation for electron microscopy

Wildtype human LONP1 was diluted to a concentration of 2.5 mg/ml in 50 mM Tris pH 8, 75 mM KCl, 10 mM MgCl₂, 1 mM TCEP, and 1 mM ATP γ S. 4 μ l of the sample were applied onto 300 mesh R1.2/1.3 UltrAuFoil Holey Gold Films (Quantifoil) that were plasma cleaned prior to sample application for 7 seconds using a Solarus plasma cleaner (Gatan, Inc.) with a 75% nitrogen, 25% oxygen atmosphere at 15W. Excess sample was blotted away for 4 s using Whatman No. 1 filter paper and vitrified by plunge freezing into a liquid ethane slurry cooled by liquid nitrogen using a manual plunger in a 4° C cold room whose humidity was raised to 95% using a humidifier. For bortezomib-bound LONP1, LONP1 was similarly diluted to a concentration of 2.5 mg/ml in 50 mM Tris pH 8, 75 mM KCl, 10 mM MgCl₂, 1 mM TCEP, and 1 mM ATP γ S with the addition of 10-fold molar excess bortezomib. Samples were prepared for cryo-EM analyses using the same procedures used for the wildtype sample.

Electron microscopy data acquisition

For wildtype human LONP1, cryo-EM data were collected on a Thermo-Fisher Talos Arctica transmission electron microscope operating at 200 keV using parallel illumination conditions⁴⁷. Micrographs were acquired using a Gatan K2 Summit direct electron detector, operated in electron counting mode applying a total electron exposure of 50 e⁻/Å² as a 114-frame dose-fractionated movie during a 11.4s exposure (**Supplementary Fig. 1a**). The Leginon data collection software⁴⁸ was used to collect 2914 micrographs at 36,000x nominal magnification (1.15 Å/pixel at the specimen level) with a nominal defocus set to -1.5 μ m. Variation from the nominally set defocus due to a 5% tilt in the stage gave rise to a defocus range (this tilt was not intentional and required service to correct). Stage movement was used to target the center of sixteen 1.2 μ m holes for focusing, and an image shift was used to acquire high magnification images in the center of each of the sixteen targeted holes.

Similarly for bortezomib-bound LONP1, cryo-EM data were collected on a Thermo-Fisher Talos Arctica transmission electron microscope operating at 200 keV using parallel illumination conditions⁴⁷. Micrographs were acquired using a Gatan K2 Summit direct electron detector, operated in electron counting mode applying a total electron exposure of 50 e⁻/Å² as a 52-frame dose-fractionated movie during a 10.4 s exposure (**Supplementary Fig. 1a**). The Leginon data collection software⁴⁸ was used to collect 4776 micrographs at 36,000x nominal magnification (1.15 Å/pixel at the specimen level) with a nominal defocus of -1.2 μ m (a single value due

to the tilted stage). Images for bortezomib-bound LONP1 were collected using a similar strategy for wildtype LONP1. Micrograph frames were aligned using MotionCor2⁴⁹ and CTF parameters were estimated with CTFFind4⁵⁰ in real-time during data acquisition to monitor image quality using the Appion image processing environment⁵¹.

Image processing for the wildtype human LONP1 + ATP γ S

A small dataset of 40,000 particles were picked using a difference of gaussians picker⁵², which was used for reference-free 2D classification in Appion49. Representative views were selected as templates for template-based particle picking using FindEM⁵³, which yielded 940,396 putative particle selections. All subsequent processing was performed in RELION 3.1⁵⁴. Particle coordinates were extracted at 3.45 Å/pixel from the motion-corrected/dose-weighted micrographs with a box size of 80 pixels, and 2D classified into 200 classes. Based on this 2D analysis, 564,930 particles belonging to classes displaying details corresponding to secondary structural elements were selected for further processing in 3D (examples shown in **Supplementary Fig. 1b**). A low-resolution negative stain reconstruction of human LONP1 was used as an initial model for 3D classification of the particles (3 classes, T=4, 25 iterations). One 3D class containing 52% of the particles resembled an open lockwasher (hereafter referred to as substrate-free), while the other two classes resembled the previously determined substrate-bound Lon complex²⁸ (**Supplementary Fig. 1b**). The alignment parameters were used to extract 293,886 and 271,044 centered, full-size particles corresponding to the substrate-free and substrate-bound complexes, respectively (1.15 Å/pixel, box size 288 pixels). The substrate-free and substrate-bound 3D classes were scaled to the appropriate pixel size and used as initial models for 3D auto-refinement of these particles, resulting in reconstructions at reported resolutions of 4.9 and 4.7 Å according to FSC of half maps at 0.143, respectively. The image shifts imposed to acquire each group of 16 exposures during data collection were used to generate 16 optics groups for CTF refinement⁵⁵ (per-particle defocus, per-micrograph astigmatism, and beam-tilt estimation). CTF refinement followed by 3D auto-refinement using a soft-edged 3D mask was repeated three times, ultimately resulting in reconstructions with reported resolutions of 3.4 Å and 3.2 Å for the substrate-free and substrate-bound complexes, respectively (**Supplementary Fig. 1b**).

Notably, the density for one of the six subunits in the substrate-free reconstruction was almost non-existent, which prompted us to perform a 3D classification without alignment on the particles contributing this reconstruction (3 classes, T=15, 50 iterations). One of the three classes contained stronger density corresponding to the sixth subunit, and the subset of 45,300 particles contributing to this reconstruction were selected for 3D auto-refinement. The resulting six-subunit structure was determined to have a resolution of 3.6 Å, although the sixth subunit remained insufficiently resolved for *ab initio* modeling (colored blue in **Supplementary Fig.**

1b).

3D classification without alignment was used to identify the particles containing the highest resolution structural data contained in both the substrate-free and substrate-bound particle stacks. 3D masks encompassing the five subunits of the substrate-free complex, and all six subunits for the substrate-bound complex were used for these classifications (3 classes, T=15, 50 iterations). 83,162 particles from the two highest-resolution classes of the substrate-free complex, and 38,130 particles from the highest resolution class of the substrate-bound complex were selected for a final masked 3D auto-refinement using local angular and translational searches (0.9 degrees, 2 pixel search range with a 0.5 pixel step). The resulting reconstructions showed a marginal qualitative improvement in density, although the reported resolutions were unchanged at 3.4 and 3.2 Å for the substrate-free and substrate-bound, respectively (**Supplementary Fig. 1c**).

Image processing for the wildtype human LONP1 + ATP γ S + bortezomib

Micrograph frames were aligned using MotionCor2⁴⁹ and CTF parameters were estimated with gCTF⁵⁶ using the RELION 3.1 processing environment⁵⁴. 2D classes representing orthogonal views of the human LONP1 complex were used for template-based particle selection in RELION. 2,736,565 particles were selected from micrographs and extracted using a box size of 256 pixels. The particle stack was binned by a factor of 4 for reference-free 2D alignment and classification (**Supplementary Fig. 9a**), and 1,539,141 particles were selected for initial processing. A low-resolution negative stain reconstruction of human LONP1 was used as an initial model for 3D refinement of particles. These particles refined to a reported resolution of 9.5 Å as estimated by FSC at a cutoff of 0.143. The particles from this reconstruction were then sorted by classification with alignment into five classes. The particles from this reconstruction were then sorted by classification with alignment into five classes, two of which, accounting for 92% of particles, displayed high-resolution features that did not contain anisotropic resolution artifacts. The 1,418,681 particles from these classes were merged and refined to 5.0 Å resolution. 3D classification without alignment was performed to increase resolution of the reconstruction. 1,142,531 particles were selected for further refinement. The X and Y image shifts applied during data acquisition were used to group images for beam tilt estimation. Refining with local defocus and beam tilt estimation within CTF Refinement in RELION 3.1 improved the overall reported resolution of our reconstruction of the substrate-bound human LONP1 to 4.1 Å by FSC at 0.143. Several more rounds of 3D classification without alignment were performed to continue increasing resolution of the reconstruction. 532,298 particles were selected for the final 3D refinement and Bayesian Polishing, resulting in a final reconstruction of bortezomib-bound human LONP1 with an overall reported resolution of 3.7 Å by FSC at 0.143 (**Supplementary Fig. 9c**). The cryo-EM density of the two seam subunits was poorly resolved in this reconstruction.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

tion, so a soft-edged 3D mask was generated to encompass the step subunits for focused refinement. Focused refinement improved the resolution in this region by 0.4 Å as shown by local resolution estimation in RELION 3.1 (**Supplementary Fig. 9d**).

Atomic model building and refinement

Model building and refinement were similarly performed for the substrate-bound and bortezomib-bound LONP1 structures. A homology model of human LONP1 was generated using a previously resolved cryo-EM structure of bacterial LONP1 as a starting model using SWISS-MODEL⁵⁷. This initial model was split into ATPase and protease domains and rigid body docked into the density of each of the subunits using UCSF Chimera⁵⁸. Real-space refinement of the docked structures and *ab initio* model building were performed in COOT⁵⁹. The flexible linker regions were modeled *ab initio* and a twelve-amino acid poly-alanine peptide as well as ATP, ADP, magnesium cofactor, and bortezomib molecules were built into the density corresponding to substrate, nucleotide, and drug, respectively. Further refinement of the full hexameric atomic model was performed using one round each of morphing and simulated annealing in addition to five real-space refinement macrocycles with atomic displacement parameters, secondary structure restraints, local grid searches, non-crystallographic symmetry, Ramachandran restraints, and global minimization in PHENIX⁶⁰. One round of geometry minimization with Ramachandran and rotamer restraints was used to minimize clash scores, followed by a final round of real-space refinement in PHENIX.

To create an initial model for substrate-free LONP1, individual subunits from the substrate-bound human LONP1 atomic model were docked into the density of each of the five well-resolved subunits of substrate-free LONP1 using UCSF Chimera⁵⁶. Real-space refinement of the docked structures and *ab initio* model building were performed in COOT⁵⁷. Residues from the model in were trimmed to C α unless there was clearly discernible EM density for the C β atom. Further refinement of the model was performed using one round each of morphing and simulated annealing in addition to five real-space refinement macrocycles with atomic displacement parameters, secondary structure restraints, local grid searches, non-crystallographic symmetry, Ramachandran restraints, and global minimization in PHENIX⁶⁰. ADP molecules were built into the density corresponding to nucleotide into the five well-resolved subunits. One round of geometry minimization with Ramachandran and rotamer restraints was used to minimize clash scores followed by a final round of real-space refinement in PHENIX. Due to the low resolution of the sixth subunit (blue in **Supplementary Fig. 1b**) atomic coordinates for this subunit are not included in the deposited atomic model. However, since secondary structural elements were visible for this subunit in the reconstruction, a copy of subunit B from our five-subunit substrate-free model was docked into the density for the sixth subunit for interpretation and figure-making.

UCSF Chimera⁵⁸ and ChimeraX⁶¹ were used to interpret

the EM reconstructions and atomic models, as well as to generate figures.

Data availability

Maps for the substrate-free, substrate-bound and LONP1^{Bz} complexes were deposited in the Electron Microscopy Data Bank under accession IDs EMD-23019, EMD-23020, and EMD-23013, respectively. Corresponding atomic models were deposited in the Protein Data Bank under accession IDs 7KSL, 7KSM, and 7KRZ, respectively.

Acknowledgements

We thank J.C. Ducom at Scripps Research High Performance Computing and C. Bowman at Scripps Research for computational support, as well as B. Anderson at the Scripps Research Electron Microscopy Facility for microscopy support. We thank Nigel Moriarty for assistance in refinement of the LONP1 bortezomib-bound structure. Further we thank Carolyn Suzuki (Rutgers) and members of the Lander and Wiseman labs at Scripps Research for insightful and helpful discussions related to this work. M.S. is supported by the National Science Foundation Graduate Research Fellowship Program. G.C.L. is supported by the National Institutes of Health (NIH) AG067594 and an Amgen Young Investigator Award. R.L.W. is supported by the NIH NS095892. G.C.L. and R.L.W. are supported by NIH AG061697. Computational analyses of EM data were performed using shared instrumentation funded by NIH S10OD021634 to G.C.L.

References:

- Quiros, P. M., Langer, T. & Lopez-Otin, C. New roles for mitochondrial proteases in health, ageing and disease. *Nat Rev Mol Cell Biol* **16**, 345-359, doi:10.1038/nrm3984 (2015).
- Song, J., Herrmann, J. M. & Becker, T. Quality control of the mitochondrial proteome. *Nat Rev Mol Cell Biol*, doi:10.1038/s41580-020-00300-2 (2020).
- Lebeau, J., Rainbolt, T. K. & Wiseman, R. L. Coordinating Mitochondrial Biology Through the Stress-Responsive Regulation of Mitochondrial Proteases. *Int Rev Cell Mol Biol* **340**, 79-128, doi:10.1016/bs.ircmb.2018.05.003 (2018).
- Pinti, M., Gibellini, L., Liu, Y., Xu, S., Lu, B. & Cossarizza, A. Mitochondrial Lon protease at the crossroads of oxidative stress, ageing and cancer. *Cell Mol Life Sci* **72**, 4807-4824, doi:10.1007/s00018-015-2039-3 (2015).
- Venkatesh, S., Lee, J., Singh, K., Lee, I. & Suzuki, C. K. Multitasking in the mitochondrion by the ATP-dependent Lon protease. *Biochim Biophys Acta* **1823**, 56-66, doi:10.1016/j.bbamcr.2011.11.003 (2012).
- Lu, B., Lee, J., Nie, X., Li, M., Morozov, Y. I., Venkatesh, S., Bogenhagen, D. F., Temiakov, D. & Suzuki, C. K. Phosphorylation of human TFAM in mitochondria impairs DNA binding and promotes degradation by the AAA+ Lon protease. *Mol Cell* **49**, 121-132, doi:10.1016/j.molcel.2012.10.023 (2013).
- Fu, G. K. & Markovitz, D. M. The human LON protease binds to mitochondrial promoters in a single-stranded, site-specific, strand-specific manner. *Biochemistry* **37**, 1905-1909, doi:10.1021/bi970928c (1998).

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

8. Fukuda, R., Zhang, H., Kim, J. W., Shimoda, L., Dang, C. V. & Semenza, G. L. HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. *Cell* **129**, 111-122, doi:10.1016/j.cell.2007.01.047 (2007).
9. Matsushima, Y., Goto, Y. & Kaguni, L. S. Mitochondrial Lon protease regulates mitochondrial DNA copy number and transcription by selective degradation of mitochondrial transcription factor A (TFAM). *Proc Natl Acad Sci U S A* **107**, 18410-18415, doi:10.1073/pnas.1008924107 (2010).
10. Bota, D. A. & Davies, K. J. Lon protease preferentially degrades oxidized mitochondrial aconitase by an ATP-stimulated mechanism. *Nat Cell Biol* **4**, 674-680, doi:10.1038/ncb836 (2002).
11. Quiros, P. M., Espanol, Y., Acin-Perez, R., Rodriguez, F., Barcena, C., Watanabe, K., Calvo, E., Loureiro, M., Fernandez-Garcia, M. S., Fueyo, A., Vazquez, J., Enriquez, J. A. & Lopez-Otin, C. ATP-dependent Lon protease controls tumor bioenergetics by reprogramming mitochondrial activity. *Cell Rep* **8**, 542-556, doi:10.1016/j.celrep.2014.06.018 (2014).
12. Ngo, J. K. & Davies, K. J. Importance of the lon protease in mitochondrial maintenance and the significance of declining lon in aging. *Ann N Y Acad Sci* **1119**, 78-87, doi:10.1196/annals.1404.015 (2007).
13. Lu, B. Mitochondrial Lon Protease and Cancer. *Adv Exp Med Biol* **1038**, 173-182, doi:10.1007/978-981-10-6674-0_12 (2017).
14. Bota, D. A. & Davies, K. J. Mitochondrial Lon protease in human disease and aging: Including an etiologic classification of Lon-related diseases and disorders. *Free Radic Biol Med* **100**, 188-198, doi:10.1016/j.freeradbiomed.2016.06.031 (2016).
15. Dikoglu, E., Alfaiz, A., Gorna, M., Bertola, D., Chae, J. H., Cho, T. J., Derbent, M., Alanay, Y., Guran, T., Kim, O. H., Llerena, J. C., Jr., Yamamoto, G., Superti-Furga, G., Reymond, A., Xenarios, I., Stevenson, B., Campos-Xavier, B., Bonafe, L., Superti-Furga, A. & Unger, S. Mutations in LONP1, a mitochondrial matrix protease, cause CODAS syndrome. *Am J Med Genet A* **167**, 1501-1509, doi:10.1002/ajmg.a.37029 (2015).
16. Inui, T., Anzai, M., Takezawa, Y., Endo, W., Kakisaka, Y., Kikuchi, A., Onuma, A., Kure, S., Nishino, I., Ohba, C., Saitsu, H., Matsumoto, N. & Haginoya, K. A novel mutation in the proteolytic domain of LONP1 causes atypical CODAS syndrome. *J Hum Genet* **62**, 653-655, doi:10.1038/jhg.2017.11 (2017).
17. Strauss, K. A., Jinks, R. N., Puffenberger, E. G., Venkatesh, S., Singh, K., Cheng, I., Mikita, N., Thilagavathi, J., Lee, J., Sarafianos, S., Benkert, A., Koehler, A., Zhu, A., Trovillion, V., McGlinchey, M., Morlet, T., Deardorff, M., Innes, A. M., Prasad, C., Chudley, A. E., Lee, I. N. & Suzuki, C. K. CODAS syndrome is associated with mutations of LONP1, encoding mitochondrial AAA+ Lon protease. *Am J Hum Genet* **96**, 121-135, doi:10.1016/j.ajhg.2014.12.003 (2015).
18. Puchades, C., Sandate, C. R. & Lander, G. C. The molecular principles governing the activity and functional diversity of AAA+ proteins. *Nat Rev Mol Cell Biol* **21**, 43-58, doi:10.1038/s41580-019-0183-6 (2020).
19. Sauer, R. T. & Baker, T. A. AAA+ proteases: ATP-fueled machines of protein destruction. *Annu Rev Biochem* **80**, 587-612, doi:10.1146/annurev-biochem-060408-172623 (2011).
20. Gur, E., Vishkautzan, M. & Sauer, R. T. Protein unfolding and degradation by the AAA+ Lon protease. *Protein Sci* **21**, 268-278, doi:10.1002/pro.2013 (2012).
21. Vieux, E. F., Wohlever, M. L., Chen, J. Z., Sauer, R. T. & Baker, T. A. Distinct quaternary structures of the AAA+ Lon protease control substrate degradation. *Proc Natl Acad Sci U S A* **110**, E2002-2008, doi:10.1073/pnas.1307066110 (2013).
22. Gur, E. & Sauer, R. T. Recognition of misfolded proteins by Lon, a AAA(+) protease. *Genes Dev* **22**, 2267-2277, doi:10.1101/gad.1670908 (2008).
23. Cha, S. S., An, Y. J., Lee, C. R., Lee, H. S., Kim, Y. G., Kim, S. J., Kwon, K. K., De Donatis, G. M., Lee, J. H., Maurizi, M. R. & Kang, S. G. Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J* **29**, 3520-3530, doi:10.1038/emboj.2010.226 (2010).
24. Lin, C. C., Su, S. C., Su, M. Y., Liang, P. H., Feng, C. C., Wu, S. H. & Chang, C. I. Structural Insights into the Allosteric Operation of the Lon AAA+ Protease. *Structure* **24**, 667-675, doi:10.1016/j.str.2016.03.001 (2016).
25. Su, S. C., Lin, C. C., Tai, H. C., Chang, M. Y., Ho, M. R., Babu, C. S., Liao, J. H., Wu, S. H., Chang, Y. C., Lim, C. & Chang, C. I. Structural Basis for the Magnesium-Dependent Activation and Hexamerization of the Lon AAA+ Protease. *Structure* **24**, 676-686, doi:10.1016/j.str.2016.03.003 (2016).
26. Botos, I., Lountos, G. T., Wu, W., Cherry, S., Ghirlando, R., Kudzhav, A. M., Rotanova, T. V., de Val, N., Tropea, J. E., Gustchina, A. & Wlodawer, A. Cryo-EM structure of substrate-free E. coli Lon protease provides insights into the dynamics of Lon machinery. *Current Research in Structural Biology* **1**, 13-20, doi:10.1016/j.crstbi.2019.10.001 (2019).
27. Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulova, F., Dauter, Z., Maurizi, M. R., Rotanova, T. V., Wlodawer, A. & Gustchina, A. The catalytic domain of Escherichia coli Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J Biol Chem* **279**, 8140-8148, doi:10.1074/jbc.M312243200 (2004).
28. Shin, M., Puchades, C., Asmita, A., Puri, N., Adjei, E., Wiseman, R. L., Karzai, A. W. & Lander, G. C. Structural basis for distinct operational modes and protease activation in AAA+ protease Lon. *Sci Adv* **6**, eaba8404, doi:10.1126/sciadv.aba8404 (2020).
29. Garcia-Nafria, J., Ondrovicova, G., Blagova, E., Levdivikov, V. M., Bauer, J. A., Suzuki, C. K., Kutejova, E., Wilkinson, A. J. & Wilson, K. S. Structure of the catalytic domain of the human mitochondrial Lon protease: proposed relation of oligomer formation and activity. *Protein Sci* **19**, 987-999, doi:10.1002/pro.376 (2010).
30. Kereiche, S., Kovacic, L., Bednar, J., Pevala, V., Kunova, N., Ondrovicova, G., Bauer, J., Ambro, L., Bellova, J., Kutejova, E. & Raska, I. The N-terminal domain plays a crucial role in the structure of a full-length human mitochondrial Lon protease. *Sci Rep* **6**, 33631, doi:10.1038/srep33631 (2016).
31. Ripstein, Z. A., Vahidi, S., Houry, W. A., Rubinstein, J. L. & Kay, L. E. A processive rotary mechanism couples substrate unfolding and proteolysis in the ClpXP degradation machinery. *Elife* **9**, doi:10.7554/eLife.52158 (2020).
32. Lopez, K. E., Rizo, A. N., Tse, E., Lin, J., Scull, N. W., Thwin, A. C., Lucius, A. L., Shorter, J. & Southworth, D. R. Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. *Nat Struct Mol Biol* **27**, 406-416, doi:10.1038/s41594-020-0409-5 (2020).
33. Puchades, C., Rampello, A. J., Shin, M., Giuliano, C. J., Wiseman, R. L., Glynn, S. E. & Lander, G. C. Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. *Science* **358**, doi:10.1126/science.aao0464 (2017).

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

34. Gates, S. N., Yokom, A. L., Lin, J., Jackrel, M. E., Rizo, A. N., Kendsersky, N. M., Buell, C. E., Sweeny, E. A., Mack, K. L., Chuang, E., Torrente, M. P., Su, M., Shorter, J. & Southworth, D. R. Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. *Science* **357**, 273-279, doi:10.1126/science.aan1052 (2017).
35. de la Pena, A. H., Goodall, E. A., Gates, S. N., Lander, G. C. & Martin, A. Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-driven translocation. *Science* **362**, doi:10.1126/science.aav0725 (2018).
36. Rizo, A. N., Lin, J., Gates, S. N., Tse, E., Bart, S. M., Castellano, L. M., DiMaio, F., Shorter, J. & Southworth, D. R. Structural basis for substrate gripping and translocation by the ClpB AAA+ disaggregase. *Nat Commun* **10**, 2393, doi:10.1038/s41467-019-10150-y (2019).
37. Cooney, I., Han, H., Stewart, M. G., Carson, R. H., Hansen, D. T., Iwasa, J. H., Price, J. C., Hill, C. P. & Shen, P. S. Structure of the Cdc48 segregase in the act of unfolding an authentic substrate. *Science* **365**, 502-505, doi:10.1126/science.aax0486 (2019).
38. Bell, T. A., Baker, T. A. & Sauer, R. T. Interactions between a subset of substrate side chains and AAA+ motor pore loops determine grip during protein unfolding. *Elife* **8**, doi:10.7554/eLife.46808 (2019).
39. Puchades, C., Ding, B., Song, A., Wiseman, R. L., Lander, G. C. & Glynn, S. E. Unique Structural Features of the Mitochondrial AAA+ Protease AFG3L2 Reveal the Molecular Basis for Activity in Health and Disease. *Mol Cell* **75**, 1073-1085 e1076, doi:10.1016/j.molcel.2019.06.016 (2019).
40. Bieniossek, C., Niederhauser, B. & Baumann, U. M. The crystal structure of apo-FtsH reveals domain movements necessary for substrate unfolding and translocation. *Proc Natl Acad Sci U S A* **106**, 21579-21584, doi:10.1073/pnas.0910708106 (2009).
41. Zhang, K., Li, S., Hsieh, K.-Y., Su, S.-C., Pintilie, G. D., Chiu, W. & Chang, C.-I. Molecular basis for the ATPase-powered substrate translocation by the Lon AAA+ protease. *bioRxiv*, doi:10.1101/2020.04.29.068361 (2020).
42. Wong, K. S. & Houry, W. A. Recent Advances in Targeting Human Mitochondrial AAA+ Proteases to Develop Novel Cancer Therapeutics. *Adv Exp Med Biol* **1158**, 119-142, doi:10.1007/978-981-13-8367-0_8 (2019).
43. Chalmers, M. J., Busby, S. A., Pascal, B. D., He, Y., Hendrickson, C. L., Marshall, A. G. & Griffin, P. R. Probing protein ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. *Anal Chem* **78**, 1005-1014, doi:10.1021/ac051294f (2006).
44. Zhang, Z. & Smith, D. L. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci* **2**, 522-531, doi:10.1002/pro.5560020404 (1993).
45. Pascal, B. D., Willis, S., Lauer, J. L., Landgraf, R. R., West, G. M., Marciano, D., Novick, S., Goswami, D., Chalmers, M. J. & Griffin, P. R. HDX workbench: software for the analysis of H/D exchange MS data. *J Am Soc Mass Spectrom* **23**, 1512-1521, doi:10.1007/s13361-012-0419-6 (2012).
46. Keppel, T. R. & Weis, D. D. Mapping residual structure in intrinsically disordered proteins at residue resolution using millisecond hydrogen/deuterium exchange and residue averaging. *J Am Soc Mass Spectrom* **26**, 547-554, doi:10.1007/s13361-014-1033-6 (2015).
47. Herzik, M. A., Jr., Wu, M. & Lander, G. C. Achieving better-than-3-A resolution by single-particle cryo-EM at 200 keV. *Nat Methods* **14**, 1075-1078, doi:10.1038/nmeth.4461 (2017).
48. Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C. S. & Carragher, B. Automated molecular microscopy: the new Legimon system. *J Struct Biol* **151**, 41-60, doi:10.1016/j.jsb.2005.03.010 (2005).
49. Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y. & Agard, D. A. MotionCorr2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).
50. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221, doi:10.1016/j.jsb.2015.08.008 (2015).
51. Lander, G. C., Stagg, S. M., Voss, N. R., Cheng, A., Fellmann, D., Pulokas, J., Yoshioka, C., Irving, C., Mulder, A., Lau, P. W., Lyumkis, D., Potter, C. S. & Carragher, B. Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J Struct Biol* **166**, 95-102, doi:10.1016/j.jsb.2009.01.002 (2009).
52. Voss, N. R., Yoshioka, C. K., Radermacher, M., Potter, C. S. & Carragher, B. DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J Struct Biol* **166**, 205-213, doi:10.1016/j.jsb.2009.01.004 (2009).
53. Roseman, A. M. FindEM - a fast, efficient program for automatic selection of particles from electron micrographs. *J Struct Biol* **145**, 91-99, doi:10.1016/j.jsb.2003.11.007 (2004).
54. Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E. & Scheres, S. H. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, doi:10.7554/eLife.42166 (2018).
55. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253-267, doi:10.1107/S2052252520000081 (2020).
56. Zhang, K. Gctf: Real-time CTF determination and correction. *J Struct Biol* **193**, 1-12, doi:10.1016/j.jsb.2015.11.003 (2016).
57. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R. & Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303, doi:10.1093/nar/gky427 (2018).
58. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).
59. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, doi:10.1107/S0907444904019158 (2004).
60. Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A. & Adams, P. D. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544, doi:10.1107/S2059798318006551 (2018).
61. Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H. & Ferrin, T. E. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25, doi:10.1002/pro.3235 (2018).
62. Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Potter, C. S., Carragher, B. & Lyumkis, D. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods* **14**, 793-796, doi:10.1038/nmeth.4347 (2017).

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

63. Thompson, J. D., Higgins, D. G., & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research*, **22**, 4673–4680, doi.org/10.1093/nar/22.22.4673 (1994)
64. Barad, B. A., Echols, N., Wang, R. Y., Cheng, Y., DiMaio, F., Adams, P. D., & Fraser, J. S. EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. *Nature methods*, **12**, 943–946, doi.org/10.1038/nmeth.3541 (2015)

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 1. Cryo-EM structure determination of the human LONP1 substrate-free and substrate-bound complexes.

a. Representative micrograph from cryo-EM data collection. **b.** Cryo-EM data processing scheme followed using RELION 3.1 software⁵⁴ to obtain the final 3D reconstructions of substrate-bound and substrate-free human LONP1, including classification revealing a subset of substrate-free particles where the sixth subunit is visible (colored blue). The final maps were used for atomic model building and refinement. **c-d.** 3-Dimensional Fourier Shell Correlation (3DFSC)⁶² of the final reconstruction of substrate-bound (**c**) and substrate-free (**d**) human LONP1 reporting a global resolution of 3.2 Å and 3.4 Å at FSC=0.143, respectively. Both reconstructions showed a high level of sphericity, as calculated by the 3DFSC (0.983 out of 1 for the substrate-bound and 0.980 out of 1 for the substrate-free). **e-f.** Final substrate-bound (**e**) and substrate-free (**f**) reconstructions filtered and colored by local resolution, as calculated using RELION. The final EM densities are resolved to a range of resolutions. The substrate-bound complex is mostly resolved to 3.1 Å resolution at the core of the complex, but > 4.0 Å in more flexible regions, such as the ‘seam’ subunit. In the substrate-free reconstruction, the protease domains of the central subunits are mostly resolved to 3.2 Å resolution, while the peripheral subunits and most of the ATPase domains are resolved to > 4.0 Å. **g-h.** Euler angle distribution plots of the 38,130 and 83,162 particles used in the final reconstruction of the substrate-bound and substrate-free structures, shown from two orthogonal directions.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 2. Comparison of substrate-free Lon protease structures from *Y. pestis* and *H. sapiens*. *a.* Axial (left) and lateral (right) views of the substrate-free human LONP1 atomic model (**a**) next to the *Y. pestis* Lon substrate-free atomic model²⁸ (**b**). Axial views of the ATPase and protease domain rings are shown from the exterior of each ring. Subunits are colored as in **Fig. 1**. *c.* Individual subunits from the substrate-free human LONP1 structure (grey) were aligned using a secondary structure-based alignment and subsequently aligned to a subunit from the substrate-free *Y. pestis* structure (light blue). These alignments reveal structural conservation between substrate-free states Lon homologs. *d.* Individual protomers of substrate-free bacterial Lon (top) and substrate-free human LONP1 (bottom) in the intermediate state lined alongside one another relative to the protease domain produced by orienting all the protease domains to a common view. The subunits of the substrate-free *Y. pestis* structure possess a steeper left-handed helical pitch than in the human substrate-free state.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 4. ATPase domains of human LONP1 have a nearly identical organization as *Y. pestis* Lon. The AAA+ of the substrate-bound LONP1 conformer are aligned to the corresponding AAA+ domains from substrate-bound *Y. pestis* Lon²⁸. All six subunits of the ATPase overlay with RMSDs <1.2 Å.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 5. Comparison of substrate-bound Lon protease structures from *Y. pestis* and *H. sapiens*. **a.** Cutaway views of the substrate-bound human LONP1 and substrate-bound *Y. pestis* Lon atomic models²⁸. Subunits are colored as in **Fig. 1**. Notably, human LONP1 has a 12-residue substrate trapped in its central channel while *Y. pestis* contains 7 residues, likely due to additional substrate interactions evolved in human LONP1, namely, a tyrosine pore-loop 2. **b.** Axial views of the protease rings emphasize the asymmetry of the substrate-bound LONP1 (left), which has a slight opening between the protease subunits of ATP1 and ATP2, compared to the six-fold symmetric organization of the substrate-bound *Y. pestis* Lon (right). **c.** Individual protomers of substrate-bound bacterial Lon aligned side-by-side relative to the protease domain produced by orienting all the proteases to a common view. Subunits from substrate-bound human LONP1 are shown colored as in (a) and shown with transparency on top of the subunits of bacterial Lon (colored gray). While the spiraling ATPases sit atop a planar proteolytic ring in bacterial Lon, both AAA+ and protease domains in substrate-bound human LONP1 are asymmetric.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 6. Substrate-bound human LONP1 shows nucleotide densities in the nucleotide-binding pocket. The cryo-EM density, shown as an isosurface mesh contoured at sigma=4.0, in the vicinity of the nucleotide binding pocket was of sufficient quality to assign the nucleotide state of each of the subunits. ATP1, ATP2, ATP3, and ATP4 subunits correspond to the ATP γ S used to determine this structure coordinated by a magnesium cofactor. The nucleotide density in the ADP1 and ADP2 subunits likely corresponded to ADP molecules.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

	N-Terminal 3-Helix Bundle
LON(P0A9M0)_Ecoli	KAEAELOKLKMMSPMSAEATVVRGYIDMVMQVPWNARSVKKDLRQAQEIILDTDHYGLER
LON(Q8D154)_Ypestis	KTEAELOKLKMMSPMSAEATVVRGYIDMVMQVPWNARSVKKDLVKAQEVLTDTDHYGLER
LONM(PIM1-P36775)_Yeast	IFDDEITKLSTLETSMSEFGVIRNYLDWLTSPWKGHSKEQYSIPRAKKILDEDEHYGMVD
LONM(O44952)_Celegans	VINEEKTKLQFLDPHSSEFSVTRNYLEWLTSPVWGLTSPENRRLSVAKKALDEGHYGMKD
LONM(Q7KUT2)_Drosophila	VIDEELTKLNFLESHSSEFNVTNRNYLDWLTSLPWGVISTENLCLCEKATETLNDDHYGMED
LONM(P36776)_Human	VVDEELSKLGLLDNHSSEFNVTNRNYLDWLTSPWGYKSNENLDLARAQAVLEEDHYGMED
	: * ** :. : * * : * : * : * : * : * : * : * : * : * : * : * : * : * : *
	WalkerA
LON(P0A9M0)_Ecoli	VKDRILEYLAVQSRVNKIKGPILCLVGPFGVGTSLGQSIAKATGRKYVRMALGGVRDEA
LON(Q8D154)_Ypestis	VKDRILEYLAVQSRVSKIKGPILCLVGPFGVGTSLGQSIAKATGRQYVRMALGGVRDEA
LONM(PIM1-P36775)_Yeast	VKDRILEYIAVGKLLGKVDGKIICFVGPFGVGTSLGQSIARALNRKFFRFVSGGMTDVA
LONM(O44952)_Celegans	VKERIMEFIIVNLLRKSIGGKILCFHGPFGVGTSLGQSIARALNRKFFRFVSGGMTDVA
LONM(Q7KUT2)_Drosophila	IKKRILEYIAVSSSLKGSTQKILCFHGPFGVGTSLGQSIARALNRKFFRFVSGGMTDVA
LONM(P36776)_Human	VKKRILEYIAVSQLRGSTQKILCFHGPFGVGTSLGQSIARALNRKFFRFVSGGMTDVA
	: * : *
	Pore-loop 1
LON(P0A9M0)_Ecoli	EIRGHRRTYIGBMPGKLIQKMAKVGKPNLFLIDEIDKMS-SDMRGDPASALLEVLDPED
LON(Q8D154)_Ypestis	EIRGHRRTYIGBMPGKLIQKMAKVGKPNLFLIDEIDKMA-SDMRGDPASALLEVLDPED
LONM(PIM1-P36775)_Yeast	EIKGHRRTYIGALPGRVVQALKKQCTQNPILLIDEIDKMGHIGDPSAALLEVLDPED
LONM(O44952)_Celegans	EIKGHRRTYVVGAMPKMIQCMKKVTENPLVLIDEVDKIGGAGFHGDPASALLEVLDPED
LONM(Q7KUT2)_Drosophila	EIKGHRRTYVVGAMPKLIQCLKKTENPLVLIDEVDKIG-KGYQGDPSALLEVLDPED
LONM(P36776)_Human	EIKGHRRTYVVGAMPKLIQCLKKTENPLVLIDEVDKIG-RGYQGDPSALLEVLDPED
	** : *
	WalkerB
	Loop-2
LON(P0A9M0)_Ecoli	NVAFSDHYLEVDYDLSVDFVATSNM-NIPAPLLDRMEVIRLSGYTEDEKLNIAKRHL
LON(Q8D154)_Ypestis	NVAFNDHYLEVDYDLSVDFVATSNM-NIPAPLLDRMEVIRLSGYTEDEKLNIAKHLL
LONM(PIM1-P36775)_Yeast	NNSFLDNYLDIPIDLSKVLFCVTANSLTIPRLLDRMEVIELTGVAEDKVKIAEQYLV
LONM(O44952)_Celegans	NANFNDHFLDVPVDLSRVLFICTANESKIPGLDRMEMIDVSGYLAEKVEIAHQHLI
LONM(Q7KUT2)_Drosophila	NANFLDHYLDVPVDLSRVLFICTANVIDTIPPLDRMELIEMSGYVAEETAIARQYLM
LONM(P36776)_Human	NANFLDHYLDVPVDLSRVLFICTANVIDTIPPLDRMEMINVSQYVQAEKLAIAERYLV
	* * :
	PS1βH
	p.Ser631Tyr
	S1
	Arg finger
LON(P0A9M0)_Ecoli	NVAFSDHYLEVDYDLSVDFVATSNM-NIPAPLLDRMEVIRLSGYTEDEKLNIAKRHL
LON(Q8D154)_Ypestis	NVAFNDHYLEVDYDLSVDFVATSNM-NIPAPLLDRMEVIRLSGYTEDEKLNIAKHLL
LONM(PIM1-P36775)_Yeast	NNSFLDNYLDIPIDLSKVLFCVTANSLTIPRLLDRMEVIELTGVAEDKVKIAEQYLV
LONM(O44952)_Celegans	NANFNDHFLDVPVDLSRVLFICTANESKIPGLDRMEMIDVSGYLAEKVEIAHQHLI
LONM(Q7KUT2)_Drosophila	NANFLDHYLDVPVDLSRVLFICTANVIDTIPPLDRMELIEMSGYVAEETAIARQYLM
LONM(P36776)_Human	NANFLDHYLDVPVDLSRVLFICTANVIDTIPPLDRMEMINVSQYVQAEKLAIAERYLV
	* * :
	p.Pro676Ser
	S2
	p.Arg721Gly
LON(P0A9M0)_Ecoli	PKQIERNALKKGELTVDDSAIIGIIRYYTREAGVRLEREISKLCRKAVKQLLLD----
LON(Q8D154)_Ypestis	PKQFERNALKKGELTIDDSAIMSIRYYTREAGVRLEREISKLCRKAVKQLLLD----
LONM(PIM1-P36775)_Yeast	PSAKKSAGLENSHVDMDTAITLMKYCYRESGVRNLKKHIEKIYRKAALQVVKLSIED
LONM(O44952)_Celegans	PDLRKDTSLATEQLKIEDSALEELIKHYCRESGVRNLQQHIERIFRKAALQIAEQNDE
LONM(Q7KUT2)_Drosophila	PQAMKDCGLTDKHINISDALNMLIRSYCRESGVRNLQKHIEKVIKRVAFRVVKK----
LONM(P36776)_Human	PQARALCGLDESKAKLSSDVLTLIKQYCRESGVRNLQKQVEKVLKSAKIVSG----
	* . : *
	p.Ala724Val
LON(P0A9M0)_Ecoli	-----KSLKHIEINGDN
LON(Q8D154)_Ypestis	-----KTVKHIEINGDN
LONM(PIM1-P36775)_Yeast	SPTSSADSKPKESVSSEEKAENNAKSSSEKTKDNNSEKTSDDIEALKTSEKINVSISQKN
LONM(O44952)_Celegans	EPAEKATTITENSEAE----P-----ITSTSSADCKSSAEQIVVCTEN
LONM(Q7KUT2)_Drosophila	-----EGEHFPVNADN
LONM(P36776)_Human	-----EAESVEVTPEN
	. : *
	Inter-domain Linker
LON(P0A9M0)_Ecoli	LHDYLGVRQFDYGRADNENRVGQVTGLAWTEVGGDLLTIETACV-----PGKGK
LON(Q8D154)_Ypestis	LKDFLGVQKVDYGRADTENRVGQVTGLAWTEVGGDLLTIETACV-----PGKGK
LONM(PIM1-P36775)_Yeast	LKDYGPPVYTTDRLYETTPPGVVMGLAWTNMGGCSLYVESVLEQPLH-----NCKKDG
LONM(O44952)_Celegans	LQKFVGRPKFTSDRMYEVTTPPGVIMGLAWTAMGGSALYIETVLKRPD-----LTNDKDG
LONM(Q7KUT2)_Drosophila	LTTFGLKQIFSSDRMYATTPPGVVMGLAWTAMGGSALYIETSSRRHIRQAKTDPNTVAGS
LONM(P36776)_Human	LQDFVGRPKVFTVERMYDVTTPPGVVMGLAWTAMGGSTLFTVETSLRRPQD-----KDAKGDKDG
	* : : * * . * : * : * : * : * : * : *
	Catalytic Loop
LON(P0A9M0)_Ecoli	LYTGTSLGVMQESIQAAALTVVRARAELGINPDFYEKRDIVHVHPGATPKDGSAGIA
LON(Q8D154)_Ypestis	LYTGTSLGVMQESIQAAALTVVRARADKLGINPDFYEKRDIVHVHPGATPKDGSAGIA
LONM(PIM1-P36775)_Yeast	FERTGQLGDMKESRLAYSFAKMYLAQKFENRFEKASIHLCPEGATPKDGSAGVT
LONM(O44952)_Celegans	IETTGNLGDVMKESVRTALTVAKGLAREQPDNKFDDKAHIIHVHPGATPKDGSAGVT
LONM(Q7KUT2)_Drosophila	LHITGNLGDVMKESQIALTVARNFLYSLEPNLFLQEQHIIHLHVHPGATPKDGSAGIT
LONM(P36776)_Human	LEVVTGQLGVMKESARIAYTFARAFMQHAPANDYLVTSIHHLHVHPGATPKDGSAGCT
	: * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *
	Protease Active Site
LON(P0A9M0)_Ecoli	MCTALVSLTGNPVRADVAMTGEITLRGQVLPIGGLKEKLLAAHRGGIKTVLIPFENKRD
LON(Q8D154)_Ypestis	MCTALVSLTGNPVRADVAMTGEITLRGLVLPIGGLKEKLLAAHRGGIKVVLIPDDNKRD
LONM(PIM1-P36775)_Yeast	MATSFSLALNKSIDPTVAMTGEITLRGQVLPIGGLKEKLLAAHRGGIKTVLIPFENKRD
LONM(O44952)_Celegans	LVSLLSLALDRPVVQDMAMTGEISLTGKVLPGVGIKELVIAARRVGAQRVFLPNENRRD
LONM(Q7KUT2)_Drosophila	IITALVSLATGKPVVRQDIAMTGEVSLGKVLPGVGIKELVIAARRSGVNCVLLPVDNKDD
LONM(P36776)_Human	IVTALLSLAMGRPVVRQNLAMTGEVSLTGKILPGVGIKELVIAAKRAGVTCIVLPAENKDD
	: : : : * . . : : * : * : * : * : * : * : * : * : * : * : * : * : * : * : *
LON(P0A9M0)_Ecoli	LEEIPDNVIADLDIHPVKRIIEVLTALQNEPSGMQVVTAK-----
LON(Q8D154)_Ypestis	LEEIPDNVIADLDIHPVKRIIDVLAIALEHPAFGAQPVAPK-----
LONM(PIM1-P36775)_Yeast	WEELPDNVKGELEPLAADWYNDIFQKLFKDVNTKEGNSVWKAEF-EILDAK-KEKD----
LONM(O44952)_Celegans	FDDLPEFMKSELDIRFVSHYDELYEHLFQ-----
LONM(Q7KUT2)_Drosophila	FEELPTYITDGLVHFATTYEDVYKIAFTDVTETTTNNVEEQEPLQKLSAAAAKSETWP
LONM(P36776)_Human	FYDLAAFIETGEVHFVEHYREIFDIAFPDEQAEALAVR-----
	: : : * : . : : * : *

Supplementary Figure 7 (caption on next page)

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 7. Sequence alignment of Lon protease homologs. ClustalW⁶³ alignment of the Uniprot sequences of Lon homologs, including cytosolic Lon from *E. coli* and *Y. pestis*, mitochondrial Lon from yeast (Pim1), *C. elegans*, *Drosophila melanogaster*, and humans. This alignment suggests conservation across Lon proteins. Notably, we found several key residues to be strictly conserved in all sequences studied: the N-terminal 3-Helix Bundle (purple boxes around conserved residues studied in *Y. pestis*), P-loop (magenta), pore-loop 1 aromatic-hydrophobic residues (blue), the Walker B motif (teal), Loop-2 (green), pre-sensor 1 beta hairpin insertion (light green), sensor-1 (yellow), a trans-acting arginine finger (orange-yellow box), a cis-acting arginine finger present in sensor-2 (orange), an inter-domain linker containing a glycine residue at its N-terminus (red-orange), and a catalytic loop that is folded into a 3_{10} helix in the proteolytic active site in auto-inhibited Lon and extended in the activated conformation (salmon). Conserved residues involved in stabilizing the activated configuration of the protease domain are highlighted using filled red boxes (V809, P854, and E882). Mutations associated with CODAS syndrome are highlighted using cyan boxes: p.Ser631Tyr, p.Pro676Ser, p.Arg721Gly, and p.Ala724Val. Residues mutated in CODAS syndrome exhibit conservation across Lon homologs, suggesting their importance for Lon protease activity.

Supplementary Figure 8. Protease domains in the substrate-bound LONP1 structure are auto-inhibited. The protease domains of substrate-bound human LONP1 aligned to their corresponding subunits from the substrate-free *Y. pestis* Lon²⁸. All six subunits of the protease domain are in similar conformations, with RMSDs between 0.9-1.1 Å. Despite having substrate trapped in its central channel, the protease domain of semi-closed LONP1 is auto-inhibited by the catalytic loop forming a 3_{10} helix in the active site, similar to substrate-free forms of bacterial and human Lon.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 9 (caption on next page)

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 9. Cryo-EM structure determination of the human LONP1 bound to bortezomib. **a.** Representative micrograph from cryo-EM data collection. **b.** Cryo-EM data processing scheme followed using RELION 3.1 software⁵⁴ to obtain the final 3D reconstruction of substrate-bound human LONP1^{Bz}. Final steps included performing focused refinement with a soft mask over the ADP1 and ADP2 subunits to improve resolution of the seam subunits. The focused refinement map was used for atomic model building and refinement. **c.** 3DFSC⁶² of the final reconstruction of LONP1^{Bz} reporting a global resolution of 3.7 Å at FSC=0.143 and Sphericity of 0.986 out of 1. **d.** Final LONP1^{Bz} reconstruction colored by local resolution calculated using RELION, showing regions resolved to 3.5 Å at the core of the complex to worse than 4.1 Å in flexible regions. Focused refinement with a mask around ADP1 and ADP2 qualitatively improved the structural detail of these subunits, shown below. **e.** Euler angle distribution plot of the 532,298 particles used in the final reconstruction of the LONP1^{Bz} structure.

Supplementary Figure 10. LONP1 bound to substrate and bortezomib shows distinct nucleotide densities in the nucleotide-binding pocket. The cryo-EM density, shown as an isosurface mesh contoured at sigma=0.01, in the vicinity of the nucleotide binding pocket was of sufficient quality to unambiguously assign the nucleotide state of each of the subunits. ATP1, ATP2, ATP3, and ATP4 subunits correspond to the ATP γ S used to determine this structure coordinated by a magnesium cofactor. The nucleotide density in the ADP1 and ADP2 subunits corresponded to ADP molecules, as there is no apparent gamma density or magnesium.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 11. Substrate-bound LONP1^{Bz} and substrate-bound *Y. pestis* Lon protease both adopt similar proteolytically active conformations. **a,b.** Axial and lateral views of the LONP1^{Bz} atomic model (**a**) and the *Y. pestis* substrate-bound state atomic model²⁸ (**b**). **c.** Individual protomers of LONP1^{Bz} aligned side-by-side relative to the protease domain produced by orienting all the proteases to a common view. The respective subunits from substrate-bound bacterial Lon are shown using a transparent ribbon representation overlaid on the subunits of substrate-bound LONP1^{Bz}. In both cases, the spiraling ATPases sit atop a planar proteolytic ring. **d.** Protease domains from the LONP1^{Bz} conformer are aligned to their corresponding subunits from substrate-bound *Y. pestis* Lon. All six subunits of the protease domain are in similar conformations, with RMSDs between 0.7-0.8 Å. The protease domains of bortezomib-bound LONP1^{Bz} are activated by extending the catalytic loop, allowing catalytic dyad formation and formation of a substrate-binding groove similar to substrate-bound *Y. pestis* Lon. **e,f.** Secondary structure-based alignments between protease subunit pairs, ATP4 and ADP1 (**e**), and ADP1 and ADP2 (**f**), from substrate-bound LONP1 and LONP1^{Bz} are shown using a ribbon representation. Although the RMSDs for the subunit pair alignment are both 0.7 Å, the catalytic serine-containing loop (highlighted in pink) is divergent between the substrate-bound LONP1 and LONP1^{Bz} due to the proteases adopting auto-inhibited (light pink) and activated (dark pink) conformations, respectively.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 12. Hydrogen-deuterium exchange mass spectrometry shows a conformational change in the protease domains upon binding bortezomib. **a.** Amino acid sequence of human mitochondrial Lon. Colored boxes located underneath the amino acid sequence denote digested peptides detected in mass spectrometry experiments. Regions with no change in exchange upon the addition of Bortezomib have grey boxes while regions with reduced D₂O exchange upon the addition of Bortezomib are colored from blue to red based on the scale shown in panel **b**; D₂O exchange decreased in regions colored blue while regions colored in red depict where the greatest reduction in D₂O exchange occurred. Regions with most D₂O exchange included residues 803-814 and 815-820 colored yellow and red, respectively, belong to an alpha helix and flexible loop framing the catalytic loop the activated conformation of the protease active site. These regions likely experience a reduction in D₂O exchange due to protease symmetrization and ring closure upon the addition of Bortezomib. **b.** Peptides that were detected as having a change in D₂O exchange rate during hydrogen-deuterium exchange mass spectrometry are colored using the shown scale. **c.** An isolated subunit from **(b)** with a hatched box highlighting the catalytic serine-containing loop and labels for catalytic residues S855 and K898 showing reduction in D₂O exchange in proximity of the proteolytic active site in the presence of Bortezomib.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Supplementary Figure 13. Mutations associated with CODAS syndrome localize to inter-subunit interfaces. **a.** The locations of the following point mutations associated with CODAS syndrome are denoted on the structure of the substrate-bound human LONP1 with spheres: p.Ser631Tyr (purple), p.Pro676Ser (dark blue), p.Arg721Gly (light blue), and p.Ala724Val (green). **b.** A close-up of the nucleotide binding pocket formed between ATP1 (purple) and ATP2 (blue) subunits. Residues bearing mutations associated with CODAS Syndrome (S631, P676, R721, and A724) are located at the inter-subunit interfaces, potentially playing a role in nucleotide binding, hydrolysis, or associated allostery. These mutations likely perturb inter-subunit interactions required for ATP-driven substrate translocation.

Structures of the Human LONP1 Protease Reveal Regulatory Steps Involved in Protease Activation

Table S1. Cryo-EM data collection, refinement, and validation statistics

	Substrate-free LONP1	Substrate-bound LONP1	Substrate-bound LONP1 ^{Bz}
EMDB ID	23019	23020	23013
PDB ID	7KSL	7KSM	7KRZ
Data Collection			
Microscope	Talos Arctica	Talos Arctica	Talos Arctica
Camera	K2 Summit	K2 Summit	K2 Summit
Magnification (nominal/at detector)	36,000/43,478	36,000/43,478	36,000/43,478
Voltage (kV)	200	200	200
Data acquisition software	Leginon ⁴⁸	Leginon	Leginon
Exposure navigation	image shift to 16	image shift to 16	image shift to 16
Total electron exposure (e ⁻ /Å ²)	50	50	50
Exposure rate (e ⁻ /pixel/sec)	5.8	6.3	6.3
Frame length (ms)	100	200	200
Number of frames	114	52	52
Pixel size (Å/pixel)*	1.15	1.15	1.15
Defocus range (μm)	0.8 to 1.5	0.8 to 1.5	0.8 to 1.5
Micrographs collected (no.)	2,912	4,774	4,774
Reconstruction			
Total extracted picks (no.)	940,396	2,736,565	2,736,565
Particles used for 3D (no.)	564,930	1,539,141	1,539,141
Final Particles (no.)	83,162	38,130	532,298
Symmetry	C1	C1	C1
Resolution (global) (Å)			
FSC 0.5 (unmasked / masked)	4.4 / 3.8	4.2 / 3.6	4.2 / 3.9
FSC 0.143 (unmasked / masked)	3.8 / 3.4	3.6 / 3.2	3.9 / 3.7
Local resolution range (Å)	3.2 - 5.0	3.1 - 4.2	3.5 - 4.3
3DFSC Sphericity ⁶² (%)	98.0	98.3	98.6
Accuracy of rotations / translations	0.82 / 0.45	0.65 / 0.35	0.80 / 1.7
Map sharpening B-factor (Å ²)	-30.1	-7.1	-157.2
Model composition			
Protein residues	2,611	3,070	3,076
Ligands	5	10	16
Model Refinement*			
Refinement package	Phenix ⁶⁰	Phenix	Phenix
MapCC (volume/mask)	0.77/0.77	0.82/0.82	0.83/0.86
R.m.s. deviations			
Bond lengths (Å)	0.00	0.0	0.01
Bond angles (°)	0.66	0.61	0.74
Validation			
Map-to-model FSC 0.5	4.00	3.61	3.98
Ramachandran plot			
Favored (%)	95.39	95.97	93.29
Allowed (%)	4.61	4.03	6.71
Outliers (%)	0.10	0.0	0.0
MolProbity score	1.82	1.77	1.92
Clashscore	11.00	9.23	9.08
Poor rotamers (%)	0.00	0.04	0.00
C-beta deviations	0.0	0.0	0.0
CaBLAM outliers (%)	3.23	2.61	3.68
EMRinger score ⁶⁴	1.95	1.91	2.61