ClpXP, an ATP-powered unfolding and protein-degradation machine

Citation

As Published
http://dx.doi.org/10.1016/j.bbamcr.2011.06.007

Publisher
Elsevier

Version
Final published version

Accessed
Fri Jun 30 21:23:54 EDT 2017

Citable Link
http://hdl.handle.net/1721.1/84570

Terms of Use
Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

Detailed Terms

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
ClpXP, an ATP-powered unfolding and protein-degradation machine

Tania A. Baker, a,b Robert T. Sauer a,*

a Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
b Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

A R T I C L E   I N F O
Article history:
Received 30 April 2011
Received in revised form 10 June 2011
Accepted 15 June 2011
Available online 27 June 2011

Keywords:
Targeted degradation
Energy-dependent proteolysis
Protein unfolding
Protein translocation

A B S T R A C T
ClpXP is a AAA+ protease that uses the energy of ATP binding and hydrolysis to perform mechanical work during targeted protein degradation within cells. ClpXP consists of hexamers of an AAA+ ATPase (ClpX) and a tetradecameric peptidase (ClpP). Asymmetric ClpX hexamers bind unstructured peptide tags in protein substrates, unfold stable tertiary structure in the substrate, and then translocate the unfolded polypeptide chain into an internal proteolytic compartment in ClpP. Here, we review our present understanding of ClpXP structure and function, as revealed by two decades of biochemical and biophysical studies. This article is part of a Special Issue entitled: AAA ATPases: Structure and function.

1. ClpXP is an archetypal AAA+ proteolytic machine
AAA+ enzymes utilize the energy of ATP binding and hydrolysis to perform the mechanical work required to power numerous biological reactions and processes [1]. An important subfamily of AAA+ machines function in ATP-dependent protein degradation in cells ranging from bacteria to humans [2]. Here, we review the structure, biological function, and molecular properties of ClpXP, a relatively simple and well-characterized AAA+ protease, which serves as paradigm for other ATP-dependent proteases, including ClpAP, ClpCP, HslUV, Lon, FtsH, PAN/20S, and the 26S proteasome.

ClpXP consists of two distinct proteins, a AAA+ ATPase called ClpX and a peptidase called ClpP. ClpX recognizes unstructured peptide sequences (called tags or degrons) in protein substrates, proceeds to unfold stable tertiary structure in the protein, and then spoils or translocates the unfolded polypeptide chain into a sequestered proteolytic compartment in ClpP for degradation into small peptide fragments (Fig. 1). Irrespective of the biological source, ClpX is active as a ring hexamer, whereas ClpP is active as a 14-subunit self-compartmentalized peptidase. In this review, we focus principally on ClpXP from Escherichia coli, which has been characterized extensively and was the first ClpX-family protease to be isolated and studied.

2. Discovery and early studies

The early history of ClpXP is intertwined with that of ClpAP, a different AAA+ protease that was initially purified in the late 1980s [3–5]. ClpAP consists of an ATPase (ClpA) and a separate peptidase (ClpP). By itself, ClpP digests small peptides but has no significant activity against proteins. Importantly, degradation of protein substrates requires ClpA, ClpP, and ATP hydrolysis. In the early 1990s, several groups discovered that ClpP could also function with a different AAA+ ATPase, ClpX, to carry out ATP-dependent proteolysis on substrates such as λO, a phage replication protein [6,7]. The specificities of ClpXP and ClpAP differed, suggesting that the AAA+ components of these proteases were responsible for recognizing substrates. As we discuss below, this principle is now well established. Indeed, a few years later, ClpX alone was purified as an enzyme that removed MuA transposase from DNA following site-specific recombination [8,9]. Thus, ClpX can function as an ATP-dependent disassembly chaperone in the absence of ClpP. Occasionally, ClpX is present but ClpP is absent in an organism, suggesting that remodeling is the main function of ClpX in these species.

3. ClpX structure and function

The biochemical functions of ClpX include binding substrates, adaptors, and ClpP, protein unfolding, and polypeptide translocation. Unfolding and translocation require ATP binding and hydrolysis to power the changes in enzyme conformation that drive these mechanical processes. Moreover, for ClpX to productively bind ClpP and some substrates, ATP binding but not hydrolysis is required [10–14].

3.1. Domain and hexamer structures

Subunits of ClpX contain a family-specific N-terminal domain and a AAA+ module, consisting of large and small AAA+ domains (Fig. 2).
The N domain folds independently of the AAA+ module as a dimer (Fig. 2B), with each subunit stabilized by coordination of a zinc atom [15–17]. The large and small AAA+ domains function together in hexameric rings, although the orientation between these domains can vary substantially in different subunits [18]. In the conformation shown in Fig. 2C, ATP or ADP binds in a cleft between the large and small AAA+ domains [18,19]. This cleft is largely formed by conserved sequence motifs (Walker A, Walker B, arginine finger, sensor-II arginine, etc.) that define the AAA+ superfamily [18–21]. Both full-length ClpX and ClpXΔN, which consists only of the large and small AAA+ domains, form ring hexamers [10,18,22]. Importantly, ClpXΔN binds ClpP with near wild-type affinity and supports ATP-dependent degradation of some native protein substrates at rates similar to wild-type ClpXP, demonstrating that the AAA+ domains perform the

Fig. 1. Cartoon model of substrate recognition and degradation by the ClpXP protease. In an initial recognition step, a peptide tag in a protein substrate binds in the axial pore of the ClpX hexamer. In subsequent ATP-dependent steps, ClpX unfolds the substrate and translocates the unfolded polypeptide into the degradation chamber of ClpP for proteolysis, where it is cleaved into small peptide fragments.

Adapted with permission from ref. [113].

Fig. 2. Domain structure of ClpX. (A) Arrangement of domains and characteristic functional motifs with respect to the linear sequence are shown for E. coli ClpX. Motifs are colored blue for ssrA-tag binding, orange for ATP binding and hydrolysis, or purple for ClpP binding. The pore-2 loop is also involved in ClpP binding. (B) Structure of the N-domain dimer (1OVX) [16]. Spheres represent zinc atoms. (C) Structure of a AAA+ module in a single ClpX subunit (3HWS) [18]. Nucleotide binds in the cleft between the large and small AAA+ domains. Motif colors correspond to those in panel A.
mechanical functions of ClpX [15,22–24]. N-domain dimers, which are flexibly tethered to the AAA+ ring of ClpX, are needed for recognition of adaptors and some substrates and contribute to hexamer stability [15,22].

The structures of the large and small AAA+ domains were initially determined in ADP-bound subunits of Helicobacter pylori ClpXΔN, which formed continuous spirals in the crystal lattice [19]. Subsequent crystal structures of E. coli ClpXΔN variants, in which some subunits were covalently linked, established the detailed conformation of ring hexamers (Fig. 3A and B) [18]. In nucleotide-free (apo) and nucleotide-bound ClpXΔN hexamers, the conformations of the large and small AAA+ domains are essentially invariant and similar to those observed in the spiral structure. However, the hexamer structures contain two major classes of subunits that differ in domain–domain orientation. In four “loadable” or L subunits, the large and small AAA+ domains are oriented in nucleotide-binding conformations (Fig. 2C) similar to the one observed in the spiral structure. Strikingly, however, in two “unloadable” or U subunits, rotations of ~80° between the large and small AAA+ domains result in movements as large as 30 Å that destroy the nucleotide-binding site. The combination of nucleotide-binding competent and non-competent ClpX subunits in an L-L-U-L-L-U pattern results in a highly

---

**Fig. 3.** ClpX hexamer structures (3HWS). (A) Face view (substrate side) colored by subunits. Nucleotide loadable (L) and unloadable (U) subunits are marked. (B) Side view (substrate side up) colored by subunit. (C) Face view (substrate side) colored by rigid-body units. (D) Side view (substrate side up) colored by rigid-body units. (E) Close up of nucleotide-binding site and hinge between two rigid-body units. In each panel, the polypeptide backbone of the hinge region and ADP are shown in CPK representation, whereas most of the protein is shown in cartoon representation.
asymmetric hexameric ring (Fig. 3A and B) [18]. As we discuss below, there is strong experimental support that ClpX rings function as asymmetric hexamers, which can bind a maximum of four molecules of ATP and/or ADP.

When viewed from the top of a ClpX ring, each small AAA+ domain packs against the large AAA+ domain of the clockwise neighboring subunit in an essentially invariant fashion [18]. Each of these rigid-body interactions buries ~2000 Å² of surface area, providing the major subunit–subunit interactions that stabilize the ClpX ring. Thus, the ring can be viewed as six rigid-body units connected by a short hinge region between the large and small AAA+ domains of each subunit (Fig. 3C and D). As a consequence, the ring conformation is determined by the invariant geometry of each rigid body and by the detailed structure of the intervening hinge regions. For the loadable subunits, ATP/ADP binding and hinge conformation are related because nucleotide contacts are made by both AAA+ domains and the hinge, and the energetic coupling between these contacts depends on the precise hinge structure (Fig. 3E). Moreover, because the ClpX ring is topologically closed, changes in the orientations of the large and small domains in a single subunit, caused by ATP binding and/or hydrolysis, could easily propagate around the entire ring to help drive substrate unfolding and/or translocation.

4. Recognition of protein substrates

4.1. Direct recognition

ClpX recognizes protein substrates by binding to short unstructured peptide sequences, which are called degradation tags, degrons, or recognition signals. For example, C-terminal residues of the bacteriophage Mu repressor, the MuA transposase, and the RecN repair protein target these proteins to ClpX, and transfer of these residues to other proteins transfers susceptibility to ClpXP degradation [25–27]. Similarly, ClpXP degrades proteins that contain the E. coli ssrA-tag sequence at their C terminus [28]. When ribosomes stall during protein synthesis in eubacteria, the ssrA tag is added to the incomplete nascent protein by the tmRNA system, ensuring degradation of these molecules [29,30]. Importantly, appending the ssrA tag to a wide variety of model proteins via cloning has allowed biochemical and biophysical investigations of the influence of substrate structure and stability on the rate of ClpXP degradation [11,31–36]. The E. coli ssrA tag is 11 residues in length, but just two C-terminal alanines and the negatively charged α-carboxyl group (AA-COOH) are the principal determinants of ClpX recognition [11,37]. As we discuss below, other residues in the ssrA tag mediate binding to the SspB adaptors, which aids in the delivery of ssrA-tagged proteins for ClpXP degradation [37,38].

In proteomic studies, endogenous E. coli substrates were initially trapped in the proteolytic chambers of inactive ClpXP proteases in vivo and were then purified and identified in vitro [27,39]. Five classes of ClpXP degradation tags were established by bioinformatics analysis of the trapped sequences, peptide-blotting experiments, tag-transfer experiments, and mutational studies (Fig. 4). Two classes of C-terminal targeting signals share homology with the ssrA-tag and MuA-tag sequences, respectively. One class of N-terminal degrons was similar to a sequence that targets the λO protein for ClpXP degradation [40], whereas another N-terminal class included signal sequences for protein secretion [39]. Apparently, failure of normal secretion results in cytoplasmic degradation by ClpXP.

Some degradation tags bind in the axial pore of ClpX. For example, the ssrA tag can be crosslinked to the pore-1 and pore-2 loops within the axial channel of the ClpX ring [41]. Similarly, mutations in both of these pore loops increase κd and decrease apparent ClpXP affinity for ssrA-tagged substrates [24,42,43]. A set of flexible ClpX loops, which surround the entrance pore and contain an RKH sequence, also play important roles in tag recognition (Fig. 5A). The importance of these loops in substrate recognition is dramatically illustrated by the fact that human ClpX contains the same GYVG loop as E. coli ClpX, fails to recognize the ssrA tag, but a variant with transplanted pore-2 and RKKH loops from E. coli ClpX acquires this activity (Fig. 3B) [41]. Moreover, mutations that reduce the positive charge of the RKKH motif weaken binding of E. coli ClpX to the negatively charged α-carboxylate of the ssrA tag (Fig. 5C and D) [44]. Strikingly, the same mutations cause better binding to substrates with positively charged degradation tags (Fig. 5C and D), including the N-terminal λO recognition tag and the C-terminal MuA tag [44]. These results suggest that the substrate specificity of wild-type ClpX is an evolutionary compromise that allows recognition of many different types of substrates but has not been optimized for any single protein or class of substrates. Compromises of this type are probably common for proteases and chaperones that interact with a large number of radically different substrate and client proteins. More than 100 intracellular substrates have been identified for ClpXP [27,39].

Although the N domain of ClpX is dispensable for degradation of ssrA-tagged proteins, it plays a critical role in recognition of some substrates. For example, removal of the N domain completely abrogates degradation of the λO protein [15,22]. The N domain is also required for degradation of the UmuD’ subunit of UmuD-D’, a translesion DNA polymerase [45,46]. In this case, tethering interactions between the N domain and sequences in the UmuD subunit facilitate engagement of a tag in the UmuD’ subunit by the pore of the ClpX ring. The N domain plays an interesting role in ClpX interactions with the MuA transposase. Phage Mu replication requires ClpX disassembly of transpososomes, which consist of a MuA tetramer bound to recombined DNA [8,9,47]. MuA has a C-terminal tag that targets monomers for degradation by ClpXP and is also required for tetramer disassembly by ClpXP [26]. ClpXΔN supports ClpP degradation of MuA monomers at near wild-type rates, consistent with recognition of the MuA tag by the ClpX pore [48]. Strikingly, however, ClpXΔN fails to support Mu replication in vivo [15] and displays a dramatic defect in transpososome disassembly in vitro, caused by poor tetramer recognition [48]. Together, these results suggest that transpososome disassembly requires engagement of the MuA C-terminal tag by the axial pore of ClpX, in addition to recognition of additional MuA sequences that are only exposed in the DNA-bound tetramer by the N domain of ClpX [48].

Fig. 4. C-terminal and N-terminal sequence tags that target substrates for ClpXP degradation [39]. Red dots – α-carboxyl group. Blue dots – α-amino group of mature protein.
4.2. Adaptor-mediated substrate recognition

Two adaptor proteins, SspB and RssB, work in concert with ClpXP to modulate proteolysis of specific substrates. SspB helps deliver ssrA-tagged proteins and additional substrates for ClpXP degradation [38,49], whereas RssB is a two-component response regulator that controls degradation of σ5′, the E. coli stationary-phase transcription factor [50,51]. Because the molecular mechanisms of SspB delivery are better understood, we focus on this system below.

SspB was initially purified as an activity that specifically enhanced ClpXP degradation of ssrA-tagged proteins, predominantly by lowering $K_m$ [38]. Several features of this adaptor allow it to tether substrates to ClpX and thus to deliver them for degradation. First, each subunit of the SspB dimer contains a groove that binds to an N-terminal sequence of the ssrA tag, adjacent to the C-terminal segment recognized by ClpX (Fig. 6A) [12,37,52,53]. Second, each SspB subunit has a highly flexible C-terminal tail that ends in a sequence that docks with the N-terminal domain of ClpX [15,17,54–57]. Both tails of the SspB dimer are needed for high-affinity tethering of ssrA-tagged substrates to ClpX and for efficient stimulation of ClpXP degradation [56], indicating that SspB makes multivalent contacts with a ClpX hexamer during substrate delivery.

An ssrA-tagged substrate can be bound to SspB and simultaneously engaged by ClpX. For example, after covalently attaching the tag to SspB via an engineered disulfide bond, this complex was found to bind ClpX much more tightly than SspB alone or the tag alone [58]. Thus, the C-terminal residues of the ssrA tag must contact ClpX at the same time that its N-terminal residues bind SspB. The individual binary contacts that stabilize the ternary complex of ClpX, SspB, and an ssrA-tagged substrate are relatively weak. However, avidity or effective-concentration effects ensure that a combination of these contacts results in very strong binding. For example, each SspB tail binds to an N domain with ~20 μM affinity, the ssrA tag binds to the ClpX pore with an affinity of ~1 μM, and the combination of all three contacts results in ~15 nM binding of ClpX to the SspB-substrate complex [12,56–58]. Given the normal intracellular concentrations of ClpXP and SspB (100–150 nM), this tight affinity would ensure efficient degradation of ssrA-tagged substrates present in only a few copies per cell.

How is this stable delivery complex disrupted to begin substrate degradation? Strikingly, when the tag of GFP-ssrA was disulfide bonded to SspB, ClpXP degraded both GFP and the covalently attached ssrA tag with a 10-fold lower $K_m$ than the same protein with a λO tag [44]. (D) AKH ClpXP degrades the λO-tagged substrate with a 30-fold lower $K_m$ than the ssrA-tagged substrate [44]. Adapted with permission from ref. [44].

In addition to ssrA-tagged substrates, proteomic experiments revealed that SspB also stimulates ClpXP degradation of N-RseA, an N-terminal fragment of a transmembrane protein [49]. During the E. coli envelope-stress response, RseA is cleaved by membrane proteases, releasing a complex of N-RseA and the σE transcription factor from the membrane [59]. ClpXP, aided by SspB, then degrades the N-RseA portion of this complex, freeing σE to activate transcription of stress genes [60]. Like ssrA-tagged proteins, N-RseA has a C-terminal AA-clip degron that targets it to ClpX [49]. Surprisingly, however, the
Adapted with permission from ref.[58].

both the adapter and substrate are degraded, requiring concurrent translocation of 3 responsible for degrading most substrates produced by tmRNA-

Thus, SspB can deliver very different types of substrates for ClpXP common contact when the two cocrystal structures are compared orientation opposite that of the ssrA tag and makes only a single portion of N-RseA that binds SspB shows no sequence similarity with the region of the ssrA tag that binds SspB. Moreover, the segment of N-RseA that binds SspB interacts with the peptide groove in an orientation opposite that of the ssrA tag and makes only a single

At low substrate concentrations, SspB enhances the rate of ClpXP proteolysis of ssrA-tagged proteins or RseA – 10-fold but can stimulate degradation of model substrates with inherently weak degradation tags to a far greater extent [38,49,62,63]. For example, mutating the C-terminal residues of the ssrA tag to bind more weakly to the ClpX pore resulted in weak-αKαα substrates that were degraded ~100-fold faster in the presence than absence of SspB [62]. Indeed, such substrates accumulated in E. coli when SspB levels were very low, but were degraded rapidly upon induction of high-level SspB expression (Fig. 7). Similar systems, using SspB stimulation of ClpXP degradation of substrates with inherently weak degradation tags, have been successfully used in Bacillus subtilis, Mycobacterium smegmatis, and Mycobacterium tuberculosis as a means of interrogating function by rapidly removing specific proteins from the cell [64,65] and are likely to work in any bacterium that contains ClpXP. Simple tethering of substrates to ClpXP raises their effective concentration relative to the pore and is sufficient for adaptor function, as synthetic split-adaptors that associate in the presence of rapamycin mimic substrate delivery by wild-type SspB and can be used to direct biological degradation [66].

4.3. Control of cellular protein levels by degradation

ClpXP degradation contributes to control of cellular protein levels in many ways. For quality control of protein synthesis, ClpXP is responsible for degrading most substrates produced by tmRNA-mediated addition of the ssrA tag during failed ribosomal translation [28–30,68,69]. This problem is significant, as approximately 1 of every 200 protein-synthesis events in E. coli terminates with ssrA tagging [67]. Once tagging occurs, there is sufficient intracellular ClpXP (~100 molecules) and SspB (~120 molecules) to ensure rapid degradation of the tagged proteins [68]. For example, when the N-terminal domain of λ repressor, which normally has a half-life of many hours, was modified by intracellular ssrA tagging, the half-life of the tagged protein was ~0.5 min in cells containing SspB and ~5 min in cells lacking SspB [38]. Because of the limited number of ClpXP proteases in cells, however, the capacity for degradation of ssrA-tagged substrates can be easily overwhelmed by over production of a genetically tagged protein. Indeed, most of the ssrA-tagged substrates that our lab uses for biochemical studies are over expressed and purified from cells containing wild-type ClpX, ClpP, and SspB.

Certain stress-response proteins contain degradation tags that result in efficient ClpXP degradation and need to be synthesized at high levels to maintain significant cellular concentrations. For example, E. coli RecN is a repair protein with an AA-COOH degron that targets it to ClpXP [27]. When DNA damage induces the SOS transcriptional response, high levels of recN mRNA and protein are synthesized, allowing RecN to accumulate despite constitutive ClpX degradation. Once DNA-damage triggered SOS transcription ceases, however, ClpXP degradation rapidly reduces RecN levels in the cell. This intrinsic instability of cellular stress-response proteins appears to be a common theme, allowing the proteome to rapidly return to the pre-stressed state and ensuring degradation of proteins that are important for fighting stress but are also deleterious or mildly toxic [27,45,46,70]. For ribosomes and many large macromolecular complexes, it appears be common that degrons recognized by ClpXP are exposed only in unassembled subunits [27,39]. When degrons are inaccessible in a native structure or buried in a protein–protein interface, it is straightforward to imagine how unfolding or post-translational modifications could be used to reveal these cryptic degrons. For
example, ClpXP only degrades Fnr repressor under aerobic conditions in which oxidation destroys an iron–sulfur cluster that is important for dimer stability, thereby promoting dissociation and access to degradation signals in the monomer [71]. Fnr monomers contain two sequences that target it for degradation by ClpXP. Similarly, among the ClpXP substrates identified by proteomics, ~25% contained at least two predicted degrons per polypeptide [27,39]. In some of these proteins, it is likely that efficient recognition by ClpXP requires binding to both degradation signals, allowing proteolysis to be controlled by binding or folding events that modulate the exposure of these sequence tags.

5. ClpP structure and function

The only positive biological function of ClpP is to serve as the proteolytic component of ClpXP, ClpAP, or related AAA+ proteases. In this role, ClpP simply needs to bind its partner ATPase and to cleave any polypeptide that is translocated into its proteolytic chamber. The size of the resulting peptide fragments must be small enough to exit the chamber and subsequently be degraded by exopeptidases to free amino acids. In addition, the free ClpP enzyme must not function as an efficient pro tease. The importance of this “negative role” is highlighted by the biological activity of acyldepsipeptide antibiotics, which kill bacteria by binding ClpP and facilitating degradation of nascent chains and other unfolded polypeptides [72,73].

E. coli ClpP is expressed as a proenzyme [74]. Autoproteolysis subsequently removes a 14-residue N-terminal propeptide to generate the mature sequence of 193 amino acids. The three-dimensional structure of ClpP, observed initially at low resolution by electron microscopy (EM) [75,76] and shortly thereafter at atomic resolution by X-ray crystallography [77], provided an elegant and straightforward explanation for its for its inactivity against protein substrates but its ability to degrade small peptides. Specifically, the proteolytic active sites of ClpP reside within a barrel-shaped chamber, formed by the face-to-face stacking of two heptameric ClpP rings (Fig. 8A). Small axial pores in each ClpP ring provide an opening into the proteolytic chamber. The dimensions of these pores in free ClpP (Fig. 8B) are large enough to allow small peptides to enter the chamber but small enough to exclude folded proteins and to dramatically slow the entry of large unfolded polypeptides [77].

Early studies revealed that the peptidase activity of ClpP was inhibited by diisopropylfluorophosphate [5,78,79], which covalently modifies the active site of serine proteases. In addition, the first crystal structure of E. coli ClpP showed that each subunit harbored a classical Ser–His–Asp catalytic triad and oxyanion hole, with conformations expected for a functional serine protease [77]. Subsequent structures of ClpP from E. coli and H. pylori were solved with a bound peptide inhibitor or product [80,81]. In these structures, the P1 side chain of the substrate mimic (representing the position immediately N-terminal to the cleaved peptide bond) binds in a spacious hydrophobic pocket, and the peptide backbone of residues P1–P4 forms β-sheet hydrogen bonds with ClpP (Fig. 8C). Although a wide variety of chemically dissimilar side chains can serve as P1 residues, non-polar residues are preferred [78,82]. For example, ~80% of products produced by ClpXP degradation of GFP-srrA result from cleavage after Leu, Gly, Met, Ala, and Tyr [82]. Moreover, in peptide substrates that are cleaved very rapidly by ClpP alone, non-polar side chains are strongly preferred at the P1 position with some additional preferences after Leu, Gly, Met, Ala, and Tyr [82]. Moreover, in peptide substrates that are cleaved very rapidly by ClpP alone, non-polar side chains are strongly preferred at the P1 position with some additional preferences after Leu, Gly, Met, Ala, and Tyr [82].

The proteolytic chamber of ClpP is roughly spherical (diameter ~50 Å) and, in principle, could accommodate several hundred residues of an unstructured substrate [77]. Two factors ensure efficient cleavage of polypeptides that enter the chamber. First, the local concentration of active sites is very high (~350 mM). Thus, simply proximity drives active-site binding and cleavage of even weakly interacting sequences. Second, each of the 14 active sites in the ClpP chamber is ~25 Å from three additional active sites (two in neighboring subunits in the same ring and one in the opposite ring), a distance that could be spanned by ~8 substrate residues in an extended conformation. Thus, a polypeptide in the degradation chamber could simultaneously bind to multiple active sites, increasing avidity and tandem-cleavage events. The fastest reported ClpP cleavage of a peptide substrate occurs at a rate of ~10,000 min⁻¹ ClpP₄ [78]. This cleavage rate combined with an average product size of ~10 residues would be more than sufficient to prevent accumulation of substantial quantities of unfolded substrate in the ClpP chamber, as the ClpX translocation rate is substantially slower [36].

As noted above, the molecular architecture of ClpP limits the activity of the isolated enzyme to degradation of small peptides, and ClpP alone cleaves unstructured polypeptides larger than 20–30 residues very slowly [78,83]. This size restriction is enforced by the size of the axial channel, determined in part by residues 1–21 of mature ClpP (Fig. 8B), which form stem-loop structures that comprise the channel walls [77,80,81,84–87]. Although these stem-loop structures are visible in only some ClpP crystal structures, mutation of these residues allows degradation of larger peptides and unfolded proteins [83,88]. Importantly, structures of ClpP in complex with acyldepsipeptides revealed conformational changes that resulted in a much wider axial channel (~20 Å; Fig. 8D), explaining why these antibiotics result in ClpP degradation of large unfolded polypeptides [89,90].

ClpP orthologs are widespread throughout eubacteria and in the mitochondria and chloroplasts of many eukaryotic cells [91]. The active form of all ClpP orthologs is a double-ring tetradecamer. Single heptameric rings, which are likely to be assembly/disassembly intermediates, are observed under some conditions [92,93]. For example, the principal solution species formed by human ClpP is a single ring [93], although double-ring tetradecamers are observed in a crystal structure and in complexes with its ClpX partner [84,93]. By itself, human ClpP cleaves small peptides poorly, suggesting that inactive heptamers equilibrate with active tetradecamers, which in turn are stabilized by ClpX binding. The inactivity of single rings, despite exposure of the proteolytic sites, would prevent rogue degradation until the double-ring cage assembles and binds ClpX or ClpA.

ClpP tetradecamers adopt inactive as well as active conformations. For example, in some crystal structures of variants or orthologs, the active sites are malformed, part of a helix that forms the ring–ring interface is disordered, and the distance between axial pores is ~10 Å shorter than in “active” structures (Fig. 8E) [77,80,84–87,91,94]. Ligand-mediated positive cooperativity is observed in cleavage and activation experiments [79,80,89,90], suggesting that binding stabilizes a higher affinity conformation of ClpP. These changes may be similar to the smaller conformational differences observed between structures of “active” ClpP, ClpP bound to a peptide inhibitor, and ClpP bound to acyldepsipeptides. The ring–ring interface of ClpP also exchanges between distinct conformations in solution NMR experiments, potentially similar to those in the crystallographic expanded and compact structures [95]. This result led to the intriguing proposal that the egress of peptide products from the ClpP chamber does not involve the axial pores but rather occurs via windows that open transiently near the ring–ring interface [95].

6. Interaction of ClpX and ClpP

EM initially showed that hexameric ClpX rings stack coaxially onto one or both ends of the ClpP barrel to form singly- or doubly-capped structures [10]. In these complexes, the axial pores of the ClpX rings align with the pores in the ClpP rings, providing a route into the proteolytic chamber. Consistently, electron density for translocating substrates in the pore was visualized in some EM studies [96,97], and cysteines in a translocating substrate were shown to form disulfide...
crosslinks with cysteines introduced into the ClpX pore [24]. In doubly-capped ClpX₆•ClpP₁₄•ClpX₆ complexes, translocation at any given time appears to occur from only one of the two ClpX rings [97]. This result indicates that the two ClpX rings must coordinate their activities through ClpP.

A ClpX ring has six subunits and a ClpP ring has seven subunits, making a symmetry mismatch inevitable. Because high-resolution structures of ClpXP have not been solved, current models are based upon EM structures, modeling using crystal structures, and mutational and biochemical studies. The ClpXP complex appears to be stabilized by one set of interactions involving loops near the axial pores of each ring and another set between peripheral structural elements in the ClpX and ClpP rings (Fig. 9A). For example, the importance of axial contacts is indicated by destabilization of ClpXP complexes containing deletions or point mutations in the pore-2 loop of ClpX or in the N-terminal stem-loop residues of ClpP, by crosslinking studies, and by double-mutant experiments [24,84,85,88,98].

The peripheral ClpXP interactions involve docking of conserved sequences in surface ClpX loops with clefts on the faces of the ClpP barrel. Bioinformatic searches initially identified homologous tripeptides (IGF in E. coli ClpX) in a surface loop that was present in all ClpP-binding AAA+ ATPases [99]. IGF-loop mutations eliminate or weaken binding to ClpP without affecting other ClpX functions [13,99]. Moreover, chymotrypsin cleaves unbound ClpX hexamers immediately

---

**Fig. 8.** Structures of ClpP. (A) ClpP (1YG6) contains 14 subunits, arranged as two stacked heptameric rings [77,86]. Interactions between residues 125–146 (blue) in the handle region help stabilize the 14-mer. (B) The axial pore in free ClpP (1YG6) is very narrow, allowing entry of only small peptides into the internal proteolytic chamber. The pore dimensions are established by residues 1–21 (red), which form stem-loop structures. (C) Active site of a ClpP subunit (magenta; ZZ2) with a bound peptide product (cyan) [81]. Residues of the Ser-His-Asp catalytic triad are labeled and the oxyanion hole is marked. The P1 side chain of the substrate (phenylalanine) sits in a hydrophobic cavity. (D) Acyldepsipeptide binding increases the size of the axial pore (3MT6), allowing degradation of unfolded proteins by ClpP alone [73,89,90]. (E) ClpP can exist in a more compact structure (3HLN) in which the active-site residues assume non-functional conformations and a portion of the handle region is disordered [91,94].
after the IGF sequence but does not cleave this site in the ClpX complex [22]. Finally, engineered ClpX hexamers with just five IGF loops bind ClpP with ~40-fold reduced affinity and variants with four IGF loops fail to bind [24]. The IGF loops are disordered in crystal structures of E. coli ClpX hexamers and have varied lengths in orthologs, suggesting a high degree of loop flexibility that could facilitate asymmetric connections between the mismatched rings of ClpX and ClpP [18,99]. The role of hydrophobic clefts on the faces of the ClpP14 barrel as IGF-docking sites [86,99]. Acyldepsipeptides, which contain a phenylalanine residue, also bind in one of the ClpP clefts (surface representation) [89]. (C) Model of the ClpX IGF peptide binding in the ClpP cleft in a manner analogous to ADEP1.

Fig. 9. Interaction of ClpX and ClpP. (A) The ClpX:ClpP complex is stabilized by peripheral interactions between the IGF loops of ClpX and hydrophobic clefts on ClpP, as well as by axial interactions between the pore-2 loops of ClpX and the N-terminal stem-loop of ClpP. Adapted with permission from ref. [24]. (B) Structure of an acyldepsipeptide (ADEP1; stick representation; 3MT6) bound in one of the ClpP clefts (surface representation) [89]. (C) Model of the ClpX IGF peptide binding in the ClpP cleft in a manner analogous to ADEP1.

7. Nucleotide transactions

7.1. Steady-state ATP hydrolysis

The steady-state kinetic parameters for ClpX hydrolysis of ATP change as a function of the presence of ClpP and/or protein substrates [99,101]. Strikingly, however, changes in $V_{max}$ are linearly correlated with changes in $K_m$ as expected if most ATP molecules that bind ClpX leave the enzyme by hydrolysis and subsequent dissociation of ADP [101]. This model predicts that $K_m$ for ATP hydrolysis should be much weaker than $K_m$ for ATP binding. In fact, $K_m$ values range from 80–400 μM, whereas $K_m$ for the hydrolysis-defective E185Q ClpX mutant is ~1 μM [14,101]. Moreover, because ClpX hydrolyzes ATP/γS at ~1/20 the rate of ATP, this model also explains why $K_m$ for ATP/γS hydrolysis (~10 μM) is substantially lower than $K_m$ for ATP hydrolysis [101].

At saturating concentrations of ATP, the hydrolysis rate per ClpX hexamer ranges from 100–600 min$^{-1}$, depending on the presence of ClpP and/or protein substrates [34,101]. From the perspective of energy conservation in a cell, it may seem somewhat odd that ClpX and ClpXP hydrolyze ATP quite readily even when protein substrates are absent. Nevertheless, ClpX orthologs show similar behavior as do most other AAA+ unfoldases and proteases. If these molecular machines are always working to degrade proteins in the cell, however, then there would be little evolutionary pressure for limiting the ATP-hydrolysis activity of substrate-free enzymes.

7.2. Asymmetric nucleotide interactions

As noted above, only four subunits in the ClpX hexamer can bind nucleotide in crystal structures [18]. biochemical studies strongly support the conclusion that ATP binding to the hexamer is also highly asymmetric in solution. For example, the hydrolysis-defective E185Q variant of ClpX binds a maximum of four ATPs per hexamer [14]. The binding-competent sites in E185Q ClpX also differ in the kinetics of ATP dissociation, with some sites releasing ATP with a half-life of ~5 s and others displaying a substantially longer half-life [14]. Binding of the ssaA tag to the axial pore of ClpX is thermodynamically linked to the binding of ATP and Mg$^{2+}$, and ~30-fold less ATP is required to support ClpX binding to a complex of SspB and an ssaA peptide than to support binding to the peptide alone. This difference correlates with the finding that the SspB-peptide complex binds ClpX much more tightly than the free ssaA peptide [58]. In these experiments, complex or ssaA-tag binding increases in a strongly cooperative manner as the ATP concentration is raised (Hill constant ~2.6) [14]. Both ATP binding and the rate of hydrolysis also increase cooperatively with ATP concentration (Hill constants 1.4–1.6) [14].
Because Hill constants define the minimum number of ligand-bound sites needed to support activity, fewer ATP-bound ClpX subunits are likely to be needed to allow hydrolysis than to bind ssrA-tagged substrates. Interestingly, a combination of ADP and ATP binding also facilitates ssrA-tag binding. For example, when E185Q ClpX hexamers are mixed with sub-saturating concentrations of ATP and an ssrA-tag peptide, addition of excess ADP causes a transient increase in ssrA-peptide binding followed by eventual dissociation [14]. This result indicates that hexamers with bound ATP and ADP are competent for tag binding.

Comparing the crystal structures of apo and ATP-bound ClpX hexamers shows some changes in the axial pore. However, the low resolution of the structures and significant disorder of the pore loops limit useful conclusions about potential conformational changes in this region [18]. When tryptophans were introduced into the pore-1 loop (to generate V154W ClpX), addition of ATP and Mg\(^{2+}\) resulted in a red shift in fluorescence, consistent with an increase in solvent accessibility in this region of the enzyme [14].

### 7.3. Lessons from single-chain hexamers

Because ClpX is a homo-hexamer, any mutation affects all six subunits. To allow studies of hexamers with different numbers of active subunits, Martin et al. [23] engineered single-chain molecules in which two, three, or six ClpX\(^{\Delta N}\) coding sequences were genetically linked to produce subunits covalently connected by flexible linkers (Figs. 10A and 10B). Importantly, these molecules formed pseudo hexamers (Fig. 10C) with the same ATP-dependent unfolding and degradation activities as unlinked ClpX\(^{\Delta N}\). Different numbers of wild-type subunits were then replaced with mutant E185Q or R370K variants to eliminate the ability of those subunits to hydrolyze ATP. Remarkably, a pseudo hexamer with just two opposed wild-type subunits degraded protein substrates at ~30% of the wild-type rate and with the same thermodynamic efficiency, calculated in terms of ATP consumed per substrate degraded (Fig. 10D) [23]. Moreover, a construct with just a single ATP-hydrolysis competent subunit displayed slow but highly specific degradation activity. Thus, protein unfolding and translocation can be powered by ATP hydrolysis in a subset of ClpX subunits, and hydrolysis in just one subunit is sufficient to drive the conformational changes in the enzyme necessary for these mechanical activities.

Single-chain studies also highlight the importance of communication between ClpX subunits. For example, the activities of pseudo-hexamers composed of linked trimers containing wild-type (W), E185Q (E), and R370K (R) subunits change depending on their arrangement within the hexamer [23]. Indeed, the RWE/RWE variant is 8-fold more active in degradation and 3-fold more active in ATP hydrolysis than the WRE/WRE isomer (Fig. 10D). Recall that E185Q subunits in homo-hexamers can adopt ATP-dependent conformations that support ClpP and ssrA-tag binding, whereas R370K subunits do not adopt these conformations. Thus, the activity of a wild-type subunit appears to be enhanced if its clockwise neighbor (the one that donates the large subunit of the shared rigid-body unit) can also undergo nucleotide-dependent conformational changes. E185Q subunits can probably adopt loadable (compatible with ATP binding) or unloadable (incompatible with binding) conformations, whereas R370K subunits may be better suited to the unloadable conformation because of poorer binding to ATP. By this model, ATP-bound ClpX subunits that lie between a nucleotide-free unloadable subunit and a nucleotide-bound loadable subunit probably have the highest hydrolysis activity.

### 7.4. Unresolved questions about the ATPase cycle

Major questions with respect to the ATPase cycle of ClpX remain unanswered. For example, once four ATPs bind to ClpX, are one, two, three, or four of these nucleotides hydrolyzed before ADP dissociation and ATP rebinding occur? Do subunits with loadable nucleotide-bound conformations convert to unloadable nucleotide-free conformations during the ATPase cycle and vice versa? Does ATP hydrolysis in a ClpX hexamer follow any pattern in terms of the order in which specific subunits fire? Models that demand a strict geometric progression (e.g., a subunit fires only after its immediate clockwise neighbor fires) seem to be inconsistent with the observation of functional single-chain hexamers with radically different arrangements of inactive and active subunits needed to support activity.
subunits [23]. One could argue that the inactive subunits in these hexamers obscure the wild-type pattern because ATP can be removed from an inactive subunit by dissociation rather than hydrolysis. However, this model requires that dissociation from inactive subunits be faster than the overall rate of ATP hydrolysis. In E185Q homo-hexamers, the fastest rate of ATP dissociation is ~8 min⁻¹, which is far too slow to account for the WWE/WWE ATP-turnover rate of 95 min⁻¹ [14,23]. Thus, one would also need to argue that dissociation from E185Q subunits in WWE/WWE occurs faster than in homo-hexamers.

It is also possible that ATP hydrolysis in wild-type ClpX does not follow a regular pattern but is instead probabilistic. For example, after hydrolysis in one subunit, any of the remaining three ATP-bound subunits might fire with distinct probabilities or the same subunit might rebind ATP and fire again, depending upon structural factors and/or interactions with neighboring subunits or a protein substrate. A recent paper argues that probabilistic models are incompatible with communication between subunits [102]. We note, however, that any equilibrium process is inherently probabilistic. The MWC model of allostery, for example, relies upon subunit–subunit communication, but any given oligomer in a population has a fixed probability of assuming the tense or relaxed conformation depending upon the intrinsic equilibrium between these states, ligand binding, and other interactions. [103].

8. Translocation and unfolding

Polypeptide translocation is the fundamental mechanical activity of ClpX. Translocation is required to spool unfolded substrates through the axial pore and into the ClpP chamber for degradation. Moreover, translocation of a peptide tag attached to a native protein drives unfolding as ClpX works to pull a large object through its narrow pore.

8.1. Pore machinery used for translocation and unfolding

As discussed above, cycles of ATP binding and hydrolysis drive rigid-body movements in the ClpX AAA+ ring. The pore-1 (GYVG) loops in the axial channel of the hexamer are responsible for linking these motions to pulling on the substrate and thus to translocation and unfolding. For example, the Y154F mutation in this loop decreases Vₘₐₓ for degradation of an unfolded substrate ~10-fold, supporting a role in translocation [42]. Furthermore, the Y153A pore-loop mutation eliminated substrate degradation completely, but it was not possible to determine whether this defect involved recognition, unfolding, or translocation.

To understand the Y153A defect in greater depth, this mutation was introduced into just the R subunits, just the W subunits, or just the E subunits of RWE/RWE linked ClpX hexamers [43]. Compared to the parental enzyme, the rates of ATP hydrolysis increased 2-fold or more for each mutant, whether the pore-loop substitutions were in active subunits or inactive subunits. These effects on enzyme “motor speed” indicate that the pore-loop machinery is tightly linked to the ATP-hydrolysis cycle. To adjust for these mutation-induced changes in motor speed, rates of degradation of unfolded or folded substrates can be divided by the working ATPase rate. Strikingly, Y153A substitutions in W or E subunits increased the energetic cost of degradation ~3-fold for unfolded substrate and 10-fold or more for native substrates [43], whereas mutations in the R subunits increased the degradation cost only for folded substrates. This increased cost of degradation suggests that some power strokes of the mutant ClpX enzyme fail to perform useful mechanical work because the mutant pore loops do not grip the substrate polypeptide tightly, do not transmit force efficiently, fail to move the substrate, or allow the polypeptide to slip following a translocation step. Moreover, these effects of GYVG pore-loop mutations on the process of translocation and unfolding strongly support a translocation-induced mechanism of protein denaturation.

8.2. Translocation direction, determinants, speed, and step size

There is no compulsory direction of ClpXP degradation, which can initiate at either the N- or C-terminus of a substrate and proceed to the opposite end [28,32,36,39,40,104]. Moreover, translocation does not require recognition by ClpX of specific side chains or peptide-bond spacings, as sequences with homopolymeric stretches of proline, glycine, positively or negative charged amino acids, aromatic residues, hydrophobic amino acids, α-amino acids, and unnatural residues with additional methylenes between successive amide bonds are all translocated at reasonable rates [105]. These results suggest that ClpX grips translocating substrates largely using van der Waal’s contacts and/or interactions with scattered carbonyl oxygens along the polypeptide chain. Poly-proline cannot form a β-strand or α-helix, and most other sequences would be unstable in the conformations that poly-proline does adopt. Thus, different segments of unfolded polypeptides are likely to be translocated in different conformations, and the step size for translocation may also vary with sequence.

ClpXP can also initiate degradation at internal substrate sequences and degrade disulfide-bonded substrates, necessitating concurrent translocation of two or more polypeptides through the pore (Fig. 6B) [33,58,106]. How does the ClpX hexamer pull on radically different sequences in single or multiple polypeptides? The pore in the crystal structure of the ClpX hexamer is almost completely closed and would need to expand to grip substrates with large side chains or multiple chains [18,105]. In some sense, this might occur analogously to the unhinging of the jaws of a snake, which allow it to swallow large prey. For example, the hinge between the large and small AAA+ domains of unfoldable subunits and the flanking helix of the small domain could unravel to allow the pore to widen to accommodate larger substrates and then refold to allow the pore to reclose and maintain intimate contacts with smaller substrates (Fig. 11) [18].

Single-molecule and biochemical studies demonstrate that ClpX and ClpXP take translocation steps of 5–8 amino acids, with an average rate of ~1800 residues min⁻¹ [36,107–109]. Translocation is highly processive. For example, translocation of an unfolded filamin domain (~100 residues) occurs in ~15 sequential steps without detectable slipping in single molecule studies, except at high resisting loads or low ATP concentrations [36]. Comparing the average rates of ATP hydrolysis and translocation indicates that ~1 ATP is hydrolyzed per step. Translocation slows as the ATP concentration is reduced, as ATPₚ럴 is added, or as the resisting force increases. Nevertheless, translocation steps of ~10 Å are still observed against resisting forces of 20 pN, demonstrating that ClpXP can perform a minimum of 3 kcal/mol of mechanical work per power stroke. The low force-dependence of translocation also suggests that a chemical step rather than a force-dependent step is rate limiting under normal conditions. The ATP-fueled conformational changes that drive translocation and unfolding would be expected to be force-dependent, suggesting that these motions are not rate limiting for ATP hydrolysis.

8.3. Unfolding activity, costs, and limits

ClpXP has a robust ability to unfold native proteins with appropriate degradation tags, including Thermus thermophilus RNase H⁺ (ΔGₚᵤₐₜ –12 kcal/mol), GFP (half-life of years for solution unfolding), and the I27 domain of human titin (AFM denaturation requires ~200 pN of force) [11,34,35,110,111]. Importantly, ClpXP degrades native titinIIᴵᴰˢᵗ₋sSᵣA –16-fold slower than an unfolded variant, demonstrating that unfolding is rate limiting for degradation of the folded substrate [34]. In addition, ClpXP degrades N-terminally tagged titinIIᴵᴰˢᵗ –4-fold faster than C-terminally tagged titinIIᴵᴰˢᵗ [112], suggesting that pulling on the N terminus and nearby structural elements results in faster unfolding. Matouschek and colleagues [32] originally proposed the idea that AAA⁺ proteases promote unfolding by pulling on and disrupting local elements of structure [32]. If the
stability of local structure is paramount in resisting ClpXP pulling, then neither the global stabilities nor the solution unfolding kinetics of substrate proteins are expected to correlate with the maximal rates at which they are degraded by ClpXP, a result which is frequently observed [35].

Several experiments show that ClpXP successfully unfolds stable protein domains only after numerous failed attempts. (i) An average of ~500 ATPs are hydrolyzed by ClpXP during the time needed to unfold a single native titin[27] molecule [34], indicating that most power strokes do not result in denaturation. (ii) Single-molecule traces of ClpXP degradation of a substrate with multiple filamin domains show dwells between completion of translocation of one domain and unfolding of the next in which the substrate-bound enzyme shows no detectable movement (Fig. 12) [36]. During these dwell periods, ClpX continues to hydrolyze ATP as it attempts to unfold the next domain. Moreover, the pre-unfolding dwell times for any specific domain are exponentially distributed and can vary from a few to several hundred seconds in different experiments. In combination, these results indicate that the proximal local structure of the target domain resists unfolding during most ClpXP power strokes. Occasionally, however, a power stroke will coincide with transient thermal destabilization of this local structure, allowing productive unfolding. When unfolding of a filamin domain does occur in single-molecule experiments, the reaction is complete within a few milliseconds and is generally highly cooperative [36]. This time is far smaller than the average time between ATP-hydrolysis events, providing strong evidence that unfolding, although stochastic, eventually results from a single ATP-dependent power stroke.

Some proteins completely resist ClpXP unfolding. For example, ClpXP cannot degrade methotrexate-bound mouse DHFR or degrade GFP with an N-terminal degradation tag at detectable rates [32,106]. ClpXP unfolding and degradation of GFP from the C-terminus also requires a minimum threshold rate of ATP hydrolysis, suggesting that multiple coordinated hydrolysis events are needed before global unfolding of this substrate occurs [107]. This result would be expected if one power stroke extracted only the C-terminal strand of this β-barrel substrate, requiring rapid ATP-dependent translocation and subsequent pulling on the remaining structure to guarantee global unfolding. When a protein resists the unfolding force of a ClpXP power stroke, two outcomes are possible. In some cases, the enzyme maintains its grip on the substrate and simply tries again [36]. In other instances, the substrate dissociates upon failed unfolding [32,104]. If dissociation occurs before proteolysis of the degradation tag, then ClpXP can rebind and eventually degrade the substrate. Thus, ClpXP is not committed to degrade an engaged protein that stubbornly resists unfolding, allowing it to preferentially degrade the most easily unfolded proteins in a mixture of substrates [104]. However, dissociation after tag cleavage can generate partially processed proteins that are no longer substrates [32,104]. In the context of a cell, ClpXp and its substrates co-evolve. Therefore, it is unclear if natural substrates are ever processed by ClpXP or generally require as much ATP hydrolysis for unfolding and degradation as is observed for “hyper-stable” model substrates in vitro. Nevertheless, ClpXP can unfold and degrade an enormous assortment of proteins with a wide range of structures and stabilities, allowing it to function in protein quality control and numerous regulatory circuits in bacteria and eukaryotic organelles.

**Conflict of interest statement**

The authors declare no conflict of interest.

**Acknowledgements**

We thank the members of our labs and all of those who have contributed to our present understanding of ClpXP. Work in our labs is supported in part from grants from the NIH (AI-15706, AI-16892, GM-
References


