Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates

Xue Fei[1,*], Tristan A Bell[1,*], Sarah R Barkow[2],

Tania A Baker[1], and Robert T Sauer[1]

Departments of Biology[1] and Chemistry[2]

Massachusetts Institute of Technology

Cambridge, MA, USA 02139

[*] These authors contributed equally to this work

Corresponding author: R.T Sauer; bobsauer@mit.edu

Running title: *ClpXP-substrate recognition*

Keywords: AAA+ protease; degron; cryo-EM structures; tmRNA tagging; degradation specificity; recognition complex; intermediate complex

ABSTRACT

When ribosomes fail to complete normal translation, all cells have mechanisms to ensure degradation of the resulting partial proteins to safeguard proteome integrity. In *E. coli* and other eubacteria, the tmRNA system rescues stalled ribosomes and adds an ssrA tag or degron to the C-terminus of the incomplete protein, which directs degradation by the AAA+ ClpXP protease. Here, we present cryo-EM structures of ClpXP bound to the ssrA degron. C-terminal residues of the ssrA degron initially bind in the top of an otherwise closed ClpX axial channel and subsequently move deeper into an open channel. For short-degron protein substrates, we show that unfolding can occur directly from the initial closed-channel complex. For longer-degron substrates, our studies illuminate how ClpXP transitions from specific recognition into a nonspecific unfolding and translocation machine. Many AAA+ proteases and protein-remodeling motors are likely to employ similar multistep recognition and engagement strategies.

Introduction

ClpXP and related AAA+ proteases maintain cellular health by degrading incomplete, damaged, or unneeded proteins in a process that must be specific to avoid destruction of essential intracellular proteins (Sauer and Baker, 2011). In bacteria and eukaryotic organelles, AAA+ proteases typically recognize substrates via short N- or C-terminal peptide sequences. *Escherichia coli* ClpXP, for example, degrades proteins bearing a C-terminal degron called the ssrA tag that is added during tmRNA-rescue of stalled ribosomes (Keiler et al., 1996; Keiler, 2015). During rescue, tmRNA binds in the empty A-site of a stalled ribosome, adds a charged alanine to the nascent polypeptide in a tRNA-like reaction, replaces the original mRNA with a short open reading frame that directs translation of the remaining residues of the ssrA degron, and finally recruits translation termination factors via a stop codon. A different bacterial mechanism, which is similar to eukaryotic systems, adds alanine tails to the nascent polypeptide during the ribosome-rescue reaction (Buskirk and Green, 2017; Lytvynenko et al., 2019).

The sequence of the *E. coli* ssrA tag is AANDENYALAA-COO^- (Keiler et al., 1996). The terminal Ala-Ala-COO^- dipeptide of this degron is the most important element for ClpXP degradation (Flynn et al., 2001), and related degrons ending in Ala-Ala target other cellular proteins to ClpXP (Flynn et al., 2003; Neher et al., 2006; Lytvynenko et al., 2019). The ssrA tag initially binds in the axial channel of the hexameric AAA+ ClpX ring, where the pore-1, pore-2, and RKH loops contribute to recognition (Siddiqui et al., 2004; Farrell et al., 2007; Martin et al., 2008a; Iosefson et al., 2015). Subsequent mechanical reactions requiring ATP hydrolysis unfold adjacent regions of native protein structure

and then translocate the denatured polypeptide through the channel and into the degradation chamber of the double-ring $ClpP_{14}$ peptidase for proteolysis (Sauer and Baker; 2011; Olivares et al., 2018). Here, we establish the molecular basis of the recognition step in targeted ClpXP degradation of ssrA-tagged proteins. We also characterize subsequent unfolding/translocation steps that initiate processive degradation.

Results

Structure determination

For cryo-EM, we used an ssrA-tagged green fluorescent protein substrate (GFP-G_3YG_9S<u>ENYALAA</u>; ssrA residues underlined), a single-chain *E. coli* $ClpX^{\Delta N}$ pseudohexamer (Martin et al., 2005), and *E. coli* ClpP. Approximately 15 s before vitrification, we combined the GFP substrate and ATP with a mixture of the $ClpX^{\Delta N}$ variant, $ClpP_{14}$, and ATPγS. Classification and three-dimensional reconstruction of EM images containing one $ClpX^{\Delta N}$ pseudohexamer and one ClpP tetradecamer generated density maps of ClpX bound to one heptameric ring of $ClpP_{14}$ and 7-10 residues of the ssrA degron in two distinct conformations at resolutions of 3.1-3.2 Å (Figures 1A, 1B; Table 1). As seen in previous cryo-EM structures (Fei et al., 2020; Ripstein et al., 2020), subunits of the $ClpX^{\Delta N}$ hexamer formed a shallow spiral (labeled ABCDEF from the top to the bottom of the spiral in the clockwise direction), which docked asymmetrically with a flat $ClpP_7$ ring. The pore of the cis ClpP ring contacting ClpX was open, as expected (Fei et al., 2020; Ripstein et al., 2020), but the pore of the trans ClpP ring was closed in both new structures. The ssrA degron bound in the top of the ClpX channel in a

structure we call the recognition complex and moved ~25 Å or 6 residues deeper into the channel in a structure we call the intermediate complex (Figure 1B).

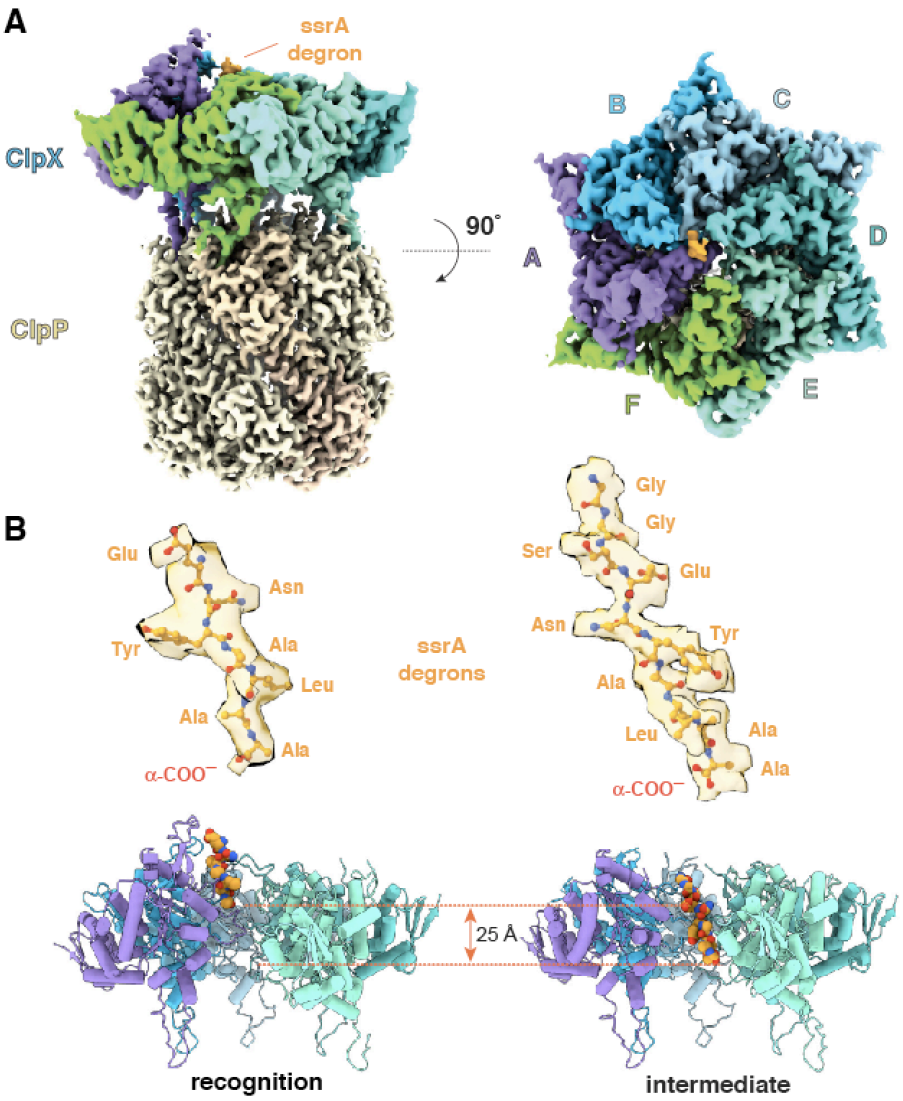


Figure 1
ClpXP complexes with ssrA degrons.
(**A**) Side and top views of density from the recognition complex. (**B**) The upper portion of the panel shows transparent density for the ssrA degron (stick representation) in the recognition and intermediate complexes; the lower portion shows the offset positions of the ssrA degron (space-filling representation) in the channel of ClpX (cartoon representation) in both complexes after removal of subunit F. In this and all subsequent figures, ClpX is colored blue, green, or purple; ClpP is yellow/gold; and substrate is orange.

Name	ClpP	ClpXP-ssrA Recognition complex	ClpXP-ssrA Intermediate complex
PDB ID	6WR2	6WRF	6WSG
EMDB ID	EMD-21875	EMD-21882	EMD-21892
	Data collection/processing		
Microscope	Talos Arctica		
Camera	K3		
Magnification	45000X		
Voltage (kV)	200		
Total electron dose ($e^-/Å^2$)	53		
Defocus range (μm)	-1.2 to -2.5		
Pixel size (Å)	0.435		
Micrographs collected	4525		
Final particles	344069	139817	130240
Symmetry	C1	C1	C1
Resolution Å (FSC 0.143)	2.8	3.1	3.2
	Model composition		
Non-hydrogen atoms	21310	26835	25932
Protein residues	2729	3450	3334
Ligands	0	6	5
	Refinement		
Map-model CC	0.82	0.77	0.74
RMSD bond lengths (Å)	0.015	0.003	0.011
RMSD bond angles (degrees)	1.16	0.73	1.13
	Validation		
MolProbity score	0.89	0.94	0.96
Clash score	1.5	1.9	1.7
C-beta deviations	0	0	0
Rotamer outliers (%)	0	0.07	0
Ramachandran favored (%)	98.9	99	99.3
Ramachandran disallowed (%)	0	0	0

Table 1
Cryo-EM data collection, processing, model building, and validation statistics.

Determinants of degron recognition

In the recognition complex, the ssrA degron bound high in the ClpX channel. Access to the lower channel was blocked by a previously unvisualized conformation of the pore-2 loop of ClpX subunit A (Figures 1B, 2A, 2B). Thr[199] and Val[202] side chains from this pore-2 loop contacted the Cβ methyl and α-COO⁻ of the C-terminal alanine of the ssrA degron. Tyr[153] and Val[154] side-chain or main-chain groups from pore-1 loops (subunits A/B) or Arg[228] and His[230] side chains from RKH loops (subunits B/C/D/F) made additional contacts with the Ala-Ala-COO⁻ or neighboring degron regions (Figures 2A, 2B, 2C; movie S1). As discussed below, multiple experiments support the role of this structure in specific substrate recognition.

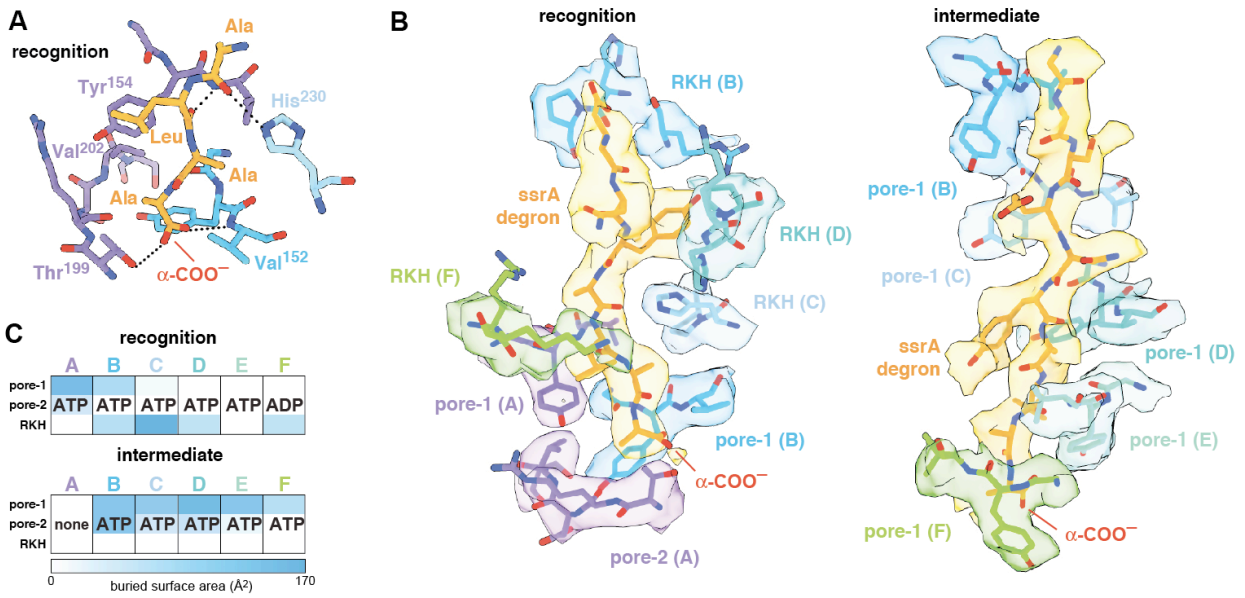


Figure 2
ClpX-degron interactions.
(**A**) Key ClpX residues (blue or purple) and the C-terminal segment of the ssrA degron (gold) in the recognition complex are shown in stick representation. Dashed lines indicate hydrogen bonds. (**B**) ClpX-degron contacts in the recognition and intermediate complexes. ClpX or degron

residues are shown in stick representation with semi-transparent density. (**C**) Subunit nucleotide state and degron/pore-loop interactions in the recognition and intermediate complexes.

First, we found that mutation of the penultimate or ultimate C-terminal alanines of the ssrA tag in 29-residue peptide substrates increased K_M for ClpXP degradation but had little effect on V_{max} (Figure 3A). A prior study showed that mutation of the antepenultimate residue of the ssrA tag (leucine) or the residue two amino acids upstream (tyrosine) also increases K_M modestly (Flynn et al., 2001), which can be rationalized based upon recognition-complex contacts. Second, previously characterized Y153A, V154F, and R228A ClpX mutations increase K_M for ClpP degradation of ssrA-tagged substrates 50-fold or more (Siddiqui et al., 2004; Farrell et al., 2007; Martin et al., 2008a; Iosefson et al., 2015). Based on the recognition-complex structure, we constructed new T199A, T199S, T199V, V202A, and H230A variants. In assays of GFP-ssrA degradation, the T199A, T199V, V202A, and H230A mutations also caused large increases in K_M (Figure 3B). By contrast, K_M for degradation by the conservative T199S variant increased only ~4-fold (Figure 3A), supporting a key role for a hydrogen bond between the side-chain hydroxyl of Thr^{199} in ClpX and the α-carboxylate of the ssrA degron (Figure 2A). Third, compared to wild-type ClpXP, the R228A variant displays reduced specificity for an ssrA-tagged substrate and increased specificity for an N-degron substrate (Farrell et al., 2007). Finally, human ClpXP has leucines at positions corresponding to Thr^{199} and His^{230} in *E. coli* ClpX and does not degrade ssrA-tagged substrates, but a human hybrid containing transplanted pore-2 and RKH loops from the *E. coli* enzyme acquires this activity (Martin et al., 2008a).

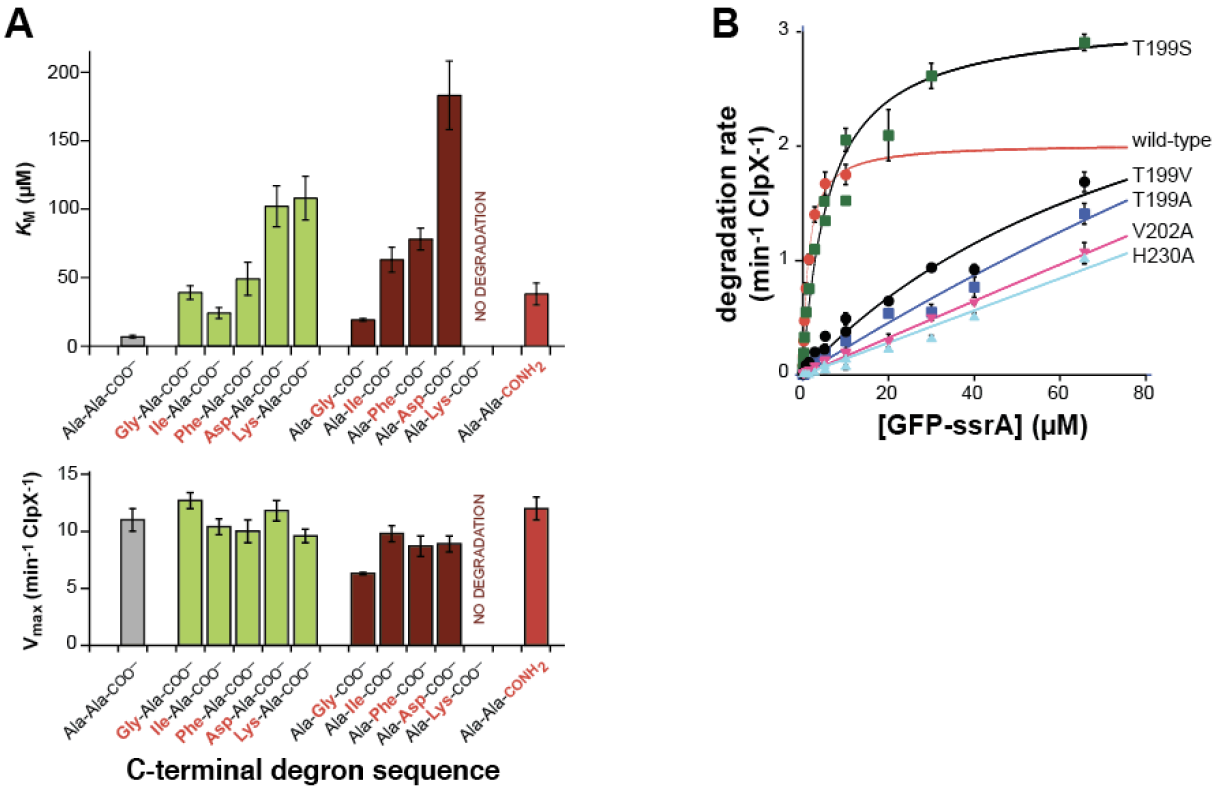


Figure 3
Mutations affecting recognition-complex contacts increase K_M for ClpXP degradation.
(**A**) Steady-state K_M (top) and V_{max} (bottom) parameters (means ± SD; n = 3) for ClpXP degradation of synthetic peptide substrates containing an aminobenzoic-acid (ABZ) fluorophore and nitrotyrosine (Y^{NO2}) quencher (wild-type sequence ABZ-FAPHMALVPY^{NO2}GYGGKKLAANDENYALAA; gray bar) with substitutions for the penultimate residue (green bars), C-terminal residue (dark red bars), or α-carboxylate (red bar). (**B**) Degradation of GFP-ssrA by ClpX$^{\Delta N}$/ClpP and variants. Data are means ± SD (n = 3 to 6), and lines are fits to the Michaelis-Menten equation. Fitted K_M values (μM) were 1.3 ± 0.2 (wild type), 6.3 ± 1.7 (T199S), and 89 ± 34 μM (T199V). Fitted V_{max} values (min^{-1} ClpX^{-1}) were 2.0 ± 0.1 (wild type), 3.1 ± 0.2 (T199S), and 3.7 ± 2.0 (T199V). For the T199A, V202A, and H230A variants, unconstrained fits had huge error limits, and fits with V_{max} constrained to ≥ 4 min^{-1} ClpX^{-1} gave K_M values from 140 to 240 μM.

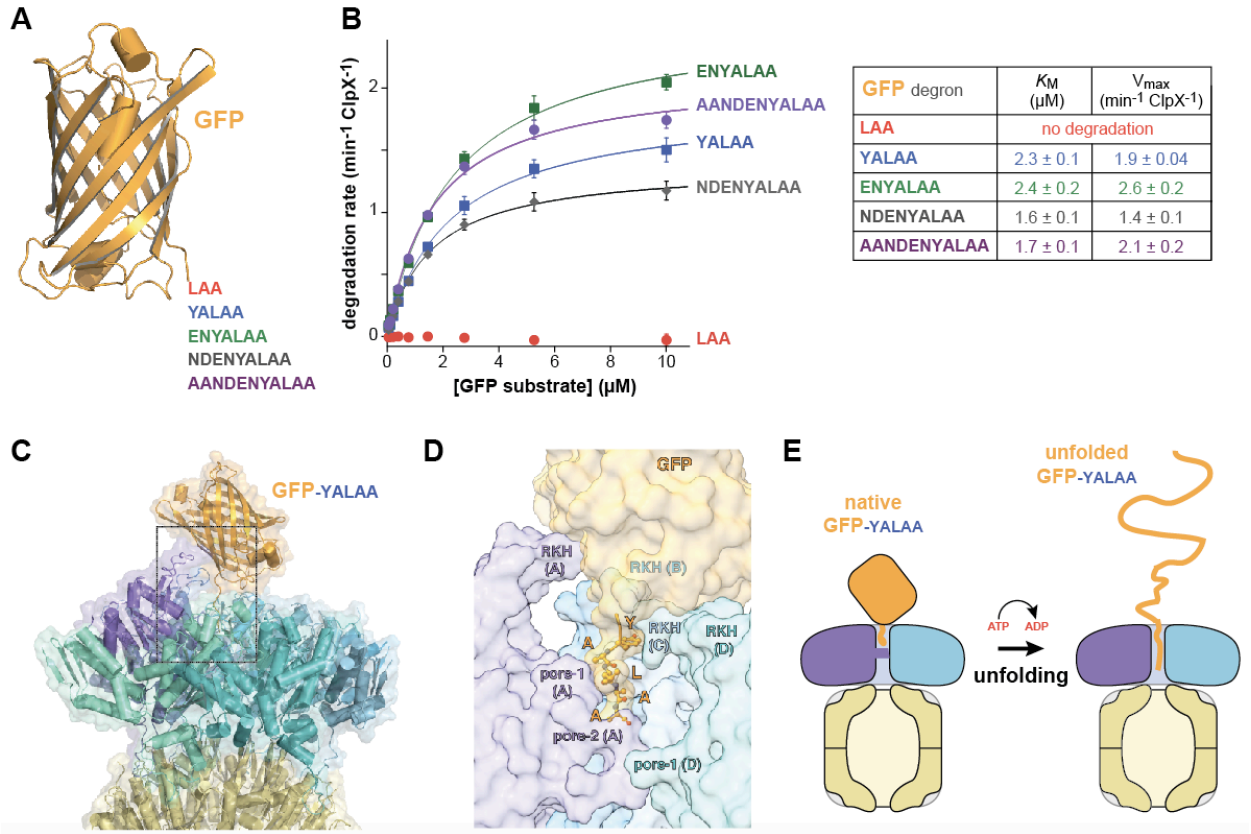
The intermediate complex resembles translocation complexes

In the intermediate complex, five ClpX pore-1 loops and four pore-2 loops packed against degron side chains with a periodicity of two residues (Figures 2B, 2C, movie S2). This arrangement of pore loops interacting with polypeptide in the channel has been observed previously in ClpXP complexes with other substrates and in different AAA+

proteases and protein-remodeling machines (Fei et al., 2020; Ripstein et al., 2020; Puchades et al., 2020). These enzyme structures are thought to reflect snapshots during non-specific translocation. Indeed, ClpXP translocates a variety of sequences, including polymeric tracts of glycine, proline, lysine, arginine, glutamate, and glutamine (Barkow et al., 2009), and the specific contacts observed in the recognition complex were absent in the intermediate complex. Non-specific translocation allows ClpXP to degrade any protein after degron recognition and unfolding of attached native structure (Sauer and Baker, 2011). Prior to recognition, our results suggest that the ClpX channel is occluded, thereby preventing non-specific binding and degron-independent degradation.

Dependence of substrate unfolding on degron length

Can a stable protein substrate, like GFP, be unfolded directly from the recognition complex or is substrate engagement by additional pore-1 and pore-2 loops deeper in the ClpX channel, as a consequence of one or more translocation steps, required to allow mechanical substrate denaturation? If direct unfolding from the recognition complex is possible, then a degron of ~5 residues, which is the number of ssrA-tag residues interacting with ClpX in the recognition complex, should be sufficient for degradation. By this model, longer degrons should also support degradation, but shorter degrons should not because they cannot make contacts needed for recognition. To test this model, we constructed substrates with degrons of 3, 5, 7, 9, or 11 residues following the last structured residue of GFP (Figure 4A) and assayed degradation.



GFP degron	K_M (μM)	V_{max} (min^{-1} ClpX^{-1})
LAA	no degradation	
YALAA	2.3 ± 0.1	1.9 ± 0.04
ENYALAA	2.4 ± 0.2	2.6 ± 0.2
NDENYALAA	1.6 ± 0.1	1.4 ± 0.1
AANDENYALAA	1.7 ± 0.1	2.1 ± 0.2

Figure 4
Degradation dependence on degron length.
(**A**) Cartoon of GFP (pdb 1EMA; ref 24) with degron tails of 3, 5, 7, 9, and 11 residues. (**B**) Plots of steady-state ClpX$^{\Delta N}$/ClpP degradation rates (means ± SD; n=3) as a function of substrate concentration. Lines are fits to the Michaelis-Menten equation. Fitted K_M and V_{max} values represent averages of three independent determinations ± SD. (**C**) Model of GFP-YALAA bound to ClpXP (cartoon and semi-transparent surface representation) created by aligning the YALAA of GFP to the same sequence in the recognition complex and then minimizing clashes both manually and computationally. (**D**) Close-up view of the ClpX-GFP-YALAA model (subunits E and F removed), corresponding to the dashed rectangle in panel C, with the YALAA shown in ball-and-stick representation. (**E**) Cartoon in which one power stroke unfolds GFP by translocating the YALAA degron six residues deeper into the channel of ClpX.

Strikingly, GFP-LAA was not degraded, whereas GFP-YALAA and substrates with longer tags were robustly degraded (Figure 4B). Modeling revealed that the native barrel of GFP-YALAA docked snugly with the top of the AAA+ ClpX ring, with the YALAA in the same position as in the recognition complex (Figures 4C, 4D; movie S3). By contrast,

the tag of GFP-LAA was too short to allow formation of recognition-complex contacts without severe steric clashes with ClpX. It might be argued that the C-terminal β-strand of GFP-YALAA, which is ~20 residues in length, unfolds to allow additional C-terminal residues of the substrate to bind deeper in the channel of ClpX. However, global GFP unfolding occurs with a half-life of years (Kim et al., 2000), and the C-terminal β-strand remains stably associated even when it is non-covalently attached to the remaining native structure (Nager et al., 2011). The inability of ClpXP to degrade GFP-LAA also argues against a model for degradation in which the C-terminal β-strand of GFP spontaneously denatures. We conclude that a power stroke initiated directly from the recognition complex can unfold GFP-YALAA (Figure 4E).

In the recognition complex, subunits ABCDE of ClpX contained ATP/ATPγS and subunit F contained ADP. In the intermediate complex, however, subunit A was nucleotide free and the remaining subunits were ATP/ATPγS bound (Figure 2C). Thus, it is plausible that ATP hydrolysis in subunit A of the recognition complex drives an unfolding power stroke and conversion to the intermediate complex. Movie S4 shows a morph illustrating the transition between these states. The 6-residue offset in the position of the ssrA tag between the two complexes (Figure 1B) is consistent with optical-trap experiments that show a fundamental ClpXP translocation step of 5-8 amino acids (Aubin-Tam et al., 2011; Maillard et al., 2011). Hence, for sufficiently short tags, a single power stroke from the recognition complex could move the tag six residues deeper into the channel, unfolding native structure N-terminal to the tag in the process (Figures 4E and 5, bottom pathway). For longer tags, including the 20-residue degron used for our cryo-EM studies,

continued translocation after the first power stroke would eventually bring the native protein in contact with the top of the ClpX channel, where additional power strokes could apply an unfolding force (Figure 5, top pathway).

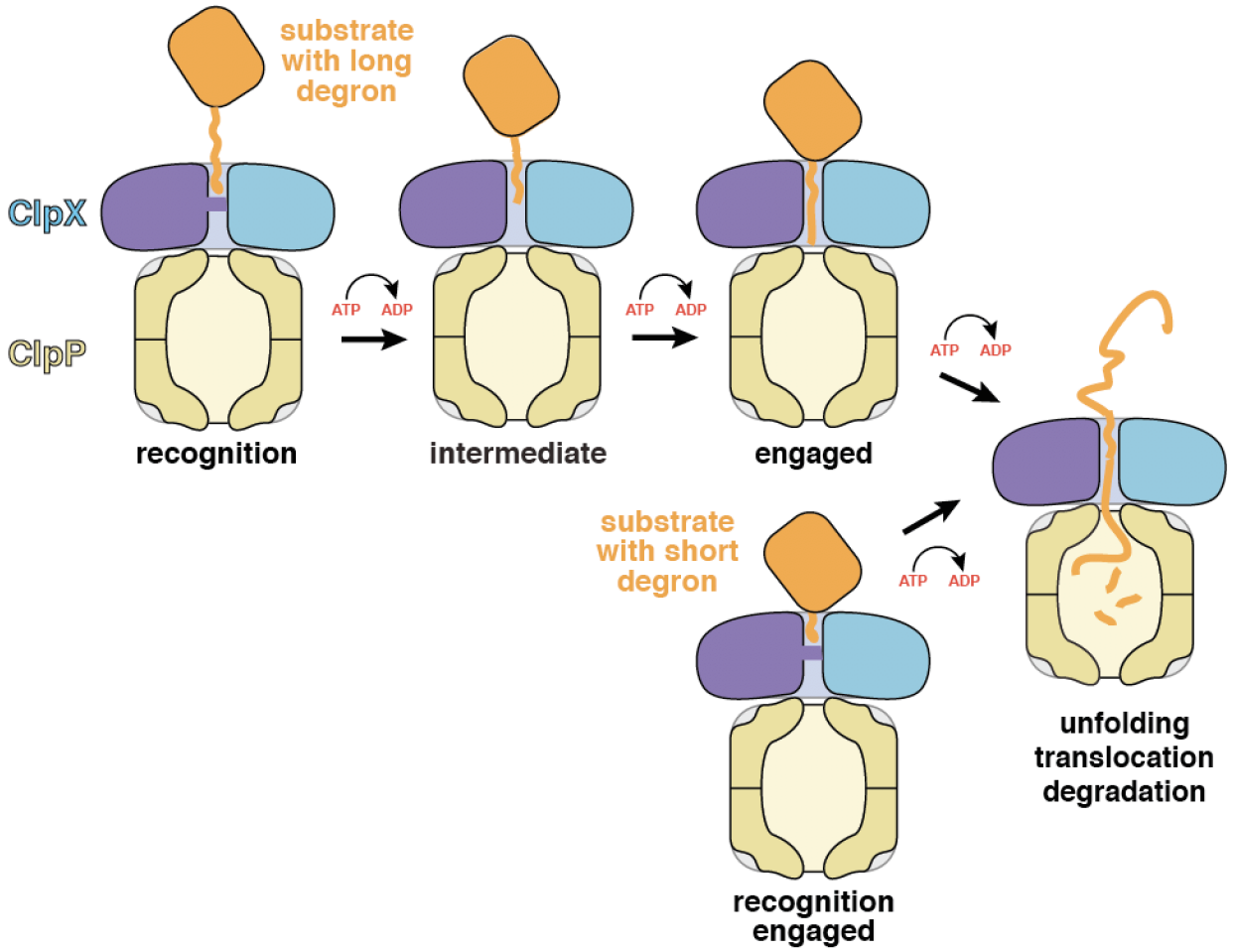


Figure 5
Models for substrate recognition, engagement, and unfolding by ClpXP.
Top. A substrate with a relatively long degron (~20 residues) is recognized and subsequent ATP-dependent power strokes then move the degron deeper into the ClpX channel in the intermediate complex, and then the engaged complex, from which unfolding occurs. Bottom. A substrate with a short degron (~5 residues) forms a recognition complex that is engaged and can therefore carry out direct ATP-dependent unfolding.

Discussion

Our cryo-EM structures of the recognition and intermediate complexes represent the first instances in which ClpX interactions with specific side chains of a protein substrate have been visualized. The general properties of the intermediate complex are similar to previously determined structures of ClpXP translocation complexes in the sense that the axial channel is open and five pore-1 loops and multiple pore-2 loops contact every two residues of the substrate polypeptide in the channel (Fei et al., 2020; Ripstein et al., 2020). This structural feature has been widely observed in AAA+ proteases and protein-remodeling machines, suggesting that these diverse molecular machines employ a common mechanism of substrate translocation (for review, see Puchades et al., 2020).

Strikingly, the recognition complex of ClpX bound to the ssrA degron is unique compared to previously determined structures. For example, all previous structures have open channels, whereas the axial channel of the AAA+ ring of ClpX in the recognition complex is closed by the pore-2 loop of subunit A, which makes specific contacts with the C-terminal residue of the ssrA tag. Moreover, only the pore-1 loops of the top two subunits in the ClpX spiral contact the substrate degron in the recognition complex, as opposed to contacts between five pore-1 loops and substrates in other known structures. The contacts we see between ClpX and the ssrA degron in the recognition complex explain multiple biochemical results, including ones that show that the C-terminal Ala-Ala dipeptide of the ssrA tag is most important for ClpXP degradation (Flynn et al., 2001; Figure 3A) and others that demonstrate that the side chains of six ClpX residues in the pore-1 loop (Tyr^{153}, Val^{154}), pore-2 loop (Thr^{199}, Val^{202}), and RKH

loop (Arg^{228}, His^{230}) play critical roles both in substrate binding and in substrate specificity (Siddiqui et al., 2004; Farrell et al., 2007; Martin et al., 2008a; Iosefson et al., 2015; Figure 3B).

In addition to providing a specific binding site for the ssrA degron, the closed axial channel of the recognition complex would preclude the non-specific binding of substrate polypeptides that is observed in the open channels of the intermediate complex and related structures. Non-specific binding in the intermediate complex explains how ClpXP efficiently translocates and degrades diverse polypeptide sequences (Barkow et al., 2009), including unnatural amino acids, but could also result in inappropriate degradation of proteins that initially bound in an open channel. Our structures suggest a simple solution to this problem. Only following specific recognition of the ssrA degron in the closed-pore conformation does the pore-2 loop of subunit A move to permit passage and interaction of deeper channel residues with any polypeptide sequence.

Structures, by themselves, can suggest but do not establish order in a kinetic pathway. However, a recent study of the kinetics of ClpXP association with a substrate bearing a 20-residue ssrA degron similar to the one studied here provides evidence for three sequentially occupied substrate-bound conformations (Saunders et al., 2020). The first two are likely to correspond to our recognition and intermediate complexes. Conversion of the first to the second complex depends on the rate of ATP hydrolysis, as does formation of the third kinetically defined state, which is probably similar to the fully engaged complex depicted in Figure 5. Prior ClpXP complexes showing the native

portion of a protein substrate contacting the top of the central channel with an attached peptide filling the channel (Fei et al., 2020) provide structural evidence for this fully engaged state. Our results indicate that ClpX can unfold short-degron substrates directly from the closed-channel recognition complex, in which just two pore-1 loops contact the degron, and long-degron substrates from a subsequent open-channel engaged complex in which five pore-1 loops engage the substrate (Fig. 5). Although unfolding in these two cases must be somewhat different in terms of the detailed mechanism, substrate residues in the upper part of the channel are gripped most tightly during ClpXP unfolding of substrates with long degrons (Bell et al., 2019), supporting a model in which ClpX-substrate contacts near the top of the channel play key roles in both unfolding mechanisms.

The multistep mechanism proposed here for protein degradation by ClpXP may be used by related AAA+ proteases that recognize unstructured degrons, as it nicely resolves how unfolding motors make initiation highly specific (and thus limited to appropriate protein targets) but then allow nonspecific unfolding and translocation for subsequent degradation and/or remodeling. Whether other AAA+ enzymes also use closed axial channels for degron recognition remains to be determined. More generally, AAA+ proteases and protein-remodeling machines must deal with problems similar to those encountered by multisubunit enzymes in transcription, translation, DNA replication, and protein secretion, which also require transitions from specific recognition conformations to complexes that utilize the chemical energy of ATP or GTP hydrolysis to processively and non-specifically move along their polymeric substrates.

Materials and methods

Proteins and peptides

His_6-tagged variants of *E. coli* ClpX, $ClpX^{\Delta N}$ (residues 62-424), single-chain $ClpX^{\Delta N}$ (units connected by six-residue linkers), full-length ClpX, and ClpP were purified as described (Kim et al. 2000; Bell et al., 2019). Peptide substrates were synthesized as described (Barkow et al., 2009), and mutant variants of $ClpX^{\Delta N}$ or GFP were constructed by PCR as described (Stinson et al., 2013).

Cryo-EM

We combined single-chain $ClpX^{\Delta N}$ (4 μM), ClpP (1.8 μM), and ATPγS (2.5 mM) in buffer containing 20 mM HEPES (pH 7.5), 100 mM KCl, and 25 mM $MgCl_2$. After 5 min at 25 °C, we added an equal volume of GFP-G_3YG_9SENYALAA (20 μM) and ATP (2.5 mM) in the same buffer, and applied 3 μL to glow discharged EM grids (Quantifoil R2/2 400 mesh). The GFP portion of this substrate consisted of residues 1-229, which includes only the C-terminal amino acids observed in a crystal structure (Ormö et al., 1996). An ATPγS/ATP mixture was used to allow rapid substrate binding, while slowing translocation and unfolding (Burton et al., 2003; Martin et al. 2008b; Saunders et al., 2020). Samples were blotted and vitrified within 15 s of final mixing. 4525 micrographs were collected using a Talos Arctica with a Gatan K3 direct electron detector (super-resolution pixel 0.435 Å; uncalibrated magnification X45000). EM images contained complexes in which $ClpP_{14}$ was bound to one (~60%) or two (~40%) $ClpX^{\Delta N}$ hexamers. After three rounds of 2D classification, 344,069 complexes with one $ClpX^{\Delta N}$ hexamer

were selected for 3D reconstruction. Using a 40-Å low-pass filtered ClpP map as the search model (EMDB: EMD-20434, Fei et al., 2020) but no mask, 3D auto-refinement without symmetry (C1) yielded a ClpXP map at 3.5-Å resolution. CTF-refinement and particle polishing improved the overall resolution of this map to 2.8 Å. After focused classification on ClpX without alignment, multiple runs with different class numbers converged to two major classes, which were later named the recognition complex (3.1 Å resolution) and intermediate complex (3.2 Å resolution). Data analysis and reconstruction was performed within the Relion 3.0.8 pipeline (Zivanov et al., 2018). We docked structures of ClpP (pdb 6PPE) or ClpXP-substrate complexes (pdb 6PP5 and 6PP7) into EM maps using Chimera (Pettersen et al., 2004), rigid-body refined ClpX domains using Coot (Emsley and Cowtan, 2004), and performed real-space refinement using PHENIX (Adams et al., 2010). The ssrA degron was first modeled and refined as polyalanine, and specific side chains were modeled and refined subsequently.

Degradation assays

Peptide degradation by ClpXP (0.3 μM full-length $ClpX_6$; 0.8 μM $ClpP_{14}$) was assayed by changes in fluorescence (excitation/emission: 320/420 nm) at 30 °C in 25 mM HEPES [pH 7.6], 200 mM KCl, 5 mM $MgCl_2$, 0.032% NP-40, 10% glycerol with 4 mM ATP, 16 mM creatine phosphate, and 0.32 mg/ml creatine phosphokinase (Barkow et al., 2009). Degradation of ssrA-tagged GFP substrates by wild-type or variant $ClpX^{\Delta N}$ hexamers (0.1-0.3 μM) and $ClpP_{14}$ (0.3-0.9 μM) was assayed by changes in fluorescence (excitation/emission: 467/511 nm for low substrate concentrations or

420/540 nm for high substrate concentrations) under similar conditions but at 37 °C, without NP-40, and with 5 mM ATP (Kim et al., 2000).

To model GFP-YALAA binding to ClpXP, we extended the C-terminal β-strand of GFP (pdb 1EMA; residues 2-229) with the sequence YALAA and aligned this sequence with the YALAA of recognition-complex degron using the PyMOL Molecular Graphics System, v. 2.0 (Schrödinger, LLC). Analysis in MolProbity (Williams et al., 2018) revealed multiple clashes between GFP and the RKH loops of ClpX, which were minimized manually in Coot (Emsley and Cowtan, 2004) by rotation of the residue 2-229 segment of GFP relative to the rest of the complex. To minimize clashes further, we created a map of the modeled complex and optimized geometry using the Calculate F(model) and Real-space refinement utilities, respectively in PHENIX (Adams et al., 2010). The final model had a MolProbity score of 1.0 (100th percentile) and one minor clash (0.44 Å) between GFP and ClpX.

Acknolwedgements

Supported by NIH grant GM-101988. We thank C. Xu, K. Song, and K. Lee at the Electron Microscopy Facility at the U. Mass. Medical School for advice and data collection.

Author contributions

X.F., T.A. Bell, S.R.B., and R.T.S. performed cryo-EM experiments, model building, or biochemical experiments. T.A. Baker and R.T.S. supervised research. All authors contributed to writing and/or revising the manuscript.

References

Adams, P.V., Afonine, G., Bunkoczi, V.B., Chen, I.W., Davis, N., Echols, J.J., Headd, L.W., Hung, G.J., Kapral, R.W., Grosse-Kunstleve, A.J., McCoy, N.W., Moriarty, R., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., and Zwart, P.H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst*. D **66**, 213–221.

Aubin-Tam, M.E., Olivares, A.O., Sauer, R.T., Baker, T.A., Lang, M.J. (2011). Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. *Cell* **145**, 257–267.

Barkow, S.R., Levchenko, I., Baker, T.A., and Sauer, R.T. (2009). Polypeptide translocation by the AAA+ ClpXP protease machine. *Chem. Biol*. **16**, 605–612.

Bell, T.A., Baker. T.A., and Sauer, R.T. (2019). Interactions between a subset of substrate side chains and AAA+ motor pore loops determine grip during protein unfolding. *eLife* **8**, e46808.

Burton, R.E., Baker, T.A., and Sauer, R.T. (2003). Energy-dependent degradation: Linkage between ClpX-catalyzed nucleotide hydrolysis and protein-substrate processing. *Protein Sci*. **12**, 893–902.

Buskirk, A.R., and Green, R. (2017). Ribosome pausing, arrest and rescue in bacteria and eukaryotes. *Philos. Trans. R. Soc. Lond. B Biol. Sci*. **372**,1716.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Cryst. D* **60**, 2126–2132.

Farrell, C.M., Baker, T.A., and Sauer, R.T. (2007). Altered specificity of a AAA+ protease. *Mol. Cell* **25**, 161–166).

Fei, X., Bell, T.A., Jenni, S. Stinson, B.M., Baker, T.A., Harrison, S.C., and Sauer, R.T. (2020). Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. *eLife* **9**, e52774.

Flynn, J.M., Levchenko, I., Seidel, M., Wickner, S.H., Sauer, R.T., and Baker, T.A. (2001). Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proc. Natl. Acad. Sci. USA* **98**, 10584–10589.

Flynn, J.M., Neher, S.B., Kim, Y.I., Sauer, R.T., and Baker, T.A. (2003). Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell* **11**, 671–683.

Iosefson, O., Nager, A.R., Baker, T.A., and Sauer, R.T. (2015). Coordinated gripping of substrate by subunits of a AAA+ proteolytic machine. *Nat. Chem. Biol*. **11**, 201–206.

Keiler, K.C. (2015). Mechanisms of ribosome rescue in bacteria. *Nat. Rev. Microbiol.* **13**, 285–297.

Keiler, K.C., Waller, P.R.H., and Sauer, R.T. (1996). Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990–993.

Kim, Y.I., Burton, R.E., Burton, B.M., Sauer, R.T., and Baker, T.A. (2000). Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Mol. Cell* **5**, 639–648.

Lytvynenko, I., Paternoga, H., Thrun, A., Balke, A., Müller, T.A., Chiang, C.H., Nagler, K., Tsaprailis, G., Anders, S., Bischofs, I., Maupin-Furlow, J.A., Spahn, C.M.T., and Joazeiro, C.A.P. (2019). Alanine tails signal proteolysis in bacterial ribosome-associated quality control. *Cell* **178**, 76–90.

Maillard, R.A., Chistol, G., Sen, G., Righini, M., Tan, J., Kaiser, C.M., Hodges, C., Martin, A., Bustamante, C. (2011). ClpX(P) generates mechanical force to unfold and translocate its protein substrates. *Cell* **45**, 459–469.

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

Martin, A., Baker, T.A., and Sauer, R.T. (2008b) Protein unfolding by a AAA+ protease is dependent on ATP-hydrolysis rates and substrate energy landscapes. *Nat. Struct. Mol. Biol*. **15**, 139–145.

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

Nager, A.R., Baker, T.A., and Sauer, R.T. (2011). Stepwise unfolding of a β barrel protein by the AAA+ ClpXP protease. *J. Mol. Biol*. **413**, 4–16.

Neher, S.B., Villén, J., Oakes, E.C., Bakalarski, C.E., Sauer, R.T., Gygi, S.P., and Baker, T.A. (2006). Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon. *Mol. Cell* **22**, 193–204.

Olivares, A.O., Baker, T.A., and Sauer, R.T. (2018). Mechanical protein unfolding and degradation. *Annu. Rev. Physiol*. **80**, 413–429.

Ormö, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., and Remington, S.J. (1996). Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science* **273**, 1392–1395.

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

Puchades, C., Sandate, C.R., and Lander, G.C. (2020). The molecular principles governing the activity and functional diversity of AAA+ proteins. *Nat. Rev. Mol. Cell. Biol*. **21**, 43–58.

Ripstein, Z.A., Vahidi, S., Houry, W.A., Rubinstein, J.L., and Kay, L.E. (2020). A processive rotary mechanism couples substrate unfolding and proteolysis in the ClpXP degradation machinery. *eLife*. **9**, e52158.

Sauer, R.T., and Baker, T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. *Annu. Rev. Biochem*. **80**, 587–612.

Saunders, R.A., Stinson, B.M., Baker, T.A, and Sauer, R.T. (2020). Multistep substrate binding and engagement by the AAA+ ClpXP protease. *BioRχiv* https://doi.org/10.1101/2020.05.04.076331.

Siddiqui, S.M., Sauer, R.T., and Baker, T.A. (2004). Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. *Genes Dev*. **18**, 369–374.

Stinson, B.M., Nager, A.R., Glynn, S.E., Schmitz, K.R., Baker, T.A., and Sauer, R.T. (2013). Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. *Cell* **153**, 628–639.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., Jain, S., Lewis, S.M., Arendall, W.B. 3rd, Snoeyink, J., Adams, P.D., Lovell, S.C., Richardson, J.S., and Richardson, D.C. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci*. **27**, 293–315.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166.