Title
Understanding the mechanisms of substrate processing by bacterial AAA+/protease Complexes

Permalink
https://escholarship.org/uc/item/30s700p0

Author
Lopez, Kyle Eric

Publication Date
2022

Peer reviewed|Thesis/dissertation

Understanding substrate processing mechanisms by bacterial AAA+/protease complexes

by
Kyle Eric Lopez

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
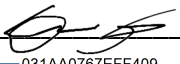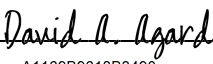DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

_____   Daniel R Southworth
031AA0767EFF409...
                                                                                          Chair

DocuSigned by:
Michelle Arkin   Michelle Arkin
_____
DocuSigned by: ...34...
David A. Agard   David A. Agard
_____
A1169B9613B3490...

_____

_____
                                                                              Committee Members

Dedication

My work over the past five years in graduate school is described in this dissertation. This work would not be possible without the encouragement of my late father Eric Edward Lopez. He did not get the chance to see me navigate undergrad let alone the rigor of graduate school. I know wherever he is, he is proud and I dedicate this work to him.

Acknowledgments

I must start by acknowledging Dr. Dan Southworth who advised me throughout my PhD. Dan was my number one advocate during my time at UCSF and ensured I was given the proper mentorship to succeed. Dan always pushed me to my full potential which led to much of my success during graduate school. I know I will always be able to look to Dan for advice and support even after I leave the lab.

I thank my lab mates who were the main reason I joined the Southworth lab. The Friday lunches at Parkside were the highlight of my time in graduate school. Thanks to Dr. Alex Rizo and Dr. Eric Tse for teaching me all I know about cryoEM, from data collection to processing. I also thank Dr. Greg Merz who is an honorary member of the Southworth lab and was always good for a distraction from work.

Next, I want to thank my undergraduate mentors. Dr. Bill Montfort at the University of Arizona taught me how to think like an experimental scientist and introduced me to the concept of looking at molecules. Dr. Andrzej Weichsel was a member of Bill's group who taught me how to purify and crystallize proteins and to him, I am forever grateful. Additionally, I would like to thank Dr. Doug Rees at Caltech who advised me over a summer studying ABC transporters and provided guidance as if I were his student.

My scientific career would not be where it is without the support of the University of Arizona MARC program. This is the single most important program in the nation to introduce minority students to the wonders of academic research. Without this program, I would not be writing this dissertation today. I want to specifically thank the MARC program administrator Cindy Neal who ensured all my cohort and countless others did the research they wanted to do. I would also like to thank Dr. Megan McEvoy now at UCLA who keeps in touch and is invested

always makes me proud that at least one of us has a cool job. Although she is not a scientist in the academic sense, she always listened to me vent about failed experiments and issues I have with science, and I appreciate that. As I continue to grow as a person and scientist, I know that development will be heavily influenced by her care and love, thank you.

To everyone who has encouraged me over the years, your contributions to this dissertation are not overlooked.

Understanding substrate processing mechanisms by bacterial AAA+/protease complexes

Kyle Eric Lopez

Abstract

ClpAP is a bacterial AAA+/protease complex responsible for regulated protein degradation of various substrates. ClpA undergoes large conformational changes coupled to ATP hydrolysis to unfold substrates targeted for degradation. The unfolded substrates are fed into the enclosed ClpP proteolytic cavity where they are degraded. This system in some drug-resistant bacteria such as *Mycobacterium tuberculosis* is a compelling target for new antibiotics, as our current stock is rendered useless by ever-evolving bacterial strains. To understand these systems, my thesis started with an investigation into the mechanisms of how *Escherichia coli* ClpAP couples large conformational changes with proteolysis. With cryoEM structures of the ClpAP complex bound to a native substrate, I propose a model where processive substrate translocation by ClpA results in a rotation relative to ClpP. Furthermore, I explore the interactions of the adaptor protein ClpS that delivers N-degron substrates to ClpAP. Again, cryoEM provides useful insight into how ClpS interacts with ClpA and how the substrate is transferred between the two proteins. Finally, I examine the *M. tuberculosis* ClpP1P2 protease bound to a small molecule activator derived from a class of natural product inhibitor molecules. This activator stimulates ClpXP1P2 activity by mimicking a peptide agonist and binds in the active site, which aligns the catalytic residues. From my thesis work, I uncovered the mechanism for how bacterial Clp proteins process various substrates, how adaptor proteins alter these processes and propose a novel stimulation mechanism for AAA/protease complexes.

Table of Contents

List of Figures

Chapter 3

Chapter 3 supplemental figures

List of Tables

Introduction

Protein homeostasis is maintained by a series of molecular chaperones and is essential for cell viability[1,2]. Errors in this process lead to various diseases such as neurodegenerative diseases and cancer. Understanding the mechanisms of all the key players is essential to developing and improving treatments for these diseases. Protein degradation is a crucial component of the proteostasis network and without it, many unfolded or non-functional proteins accumulate and hinder cellular processes. In bacteria, regulated protein degradation is performed by specific ATPases associated with various cellular activities (AAA+)/protease complexes, many of which are not present in eukaryotes[3-5]. This lack of overlap makes these bacterial protease systems a compelling target for new antibiotics that either inhibit regulated protein degradation or stimulate activity. Thus, the development of improved chemical modulators requires a better understanding of how these AAA+/protease work mechanistically.

AAA+ complexes are hexameric assemblies that couple ATP hydrolysis at protomer interfaces to large conformational changes that lead to substrate translocation down the central channel[6,7]. Before recent advancements in cryo-electron microscopy (cryoEM), most structural information on AAA+ complexes was limited to monomeric and single-domain crystal structures which provide little information about the conformational dynamics and mechanism of translocation[8,9]. Recent CryoEM structures from our lab and others of substrate bound AAA+ hexamers revealed these machines engage the polypeptide substrate in a spiral conformation along the channel, and ATP hydrolysis leads to large conformational changes that drive substrate release and re-binding during translocation steps[10-13]. The AAA+ complexes explored in my

thesis work are in the bacterial Hsp100 or Clp family and interact with the protease ClpP through

IGL loops motifs that dock into hydrophobic clefts at the interface of ClpP protomers[14-20].

In the bactieral Clp system, folded substrates with specific degradation tags are

recognized by ClpA and substrate translocation is coupled to ATP hydrolysis and large

conformational changes (Fig. I.1 a). ClpA interacts with the protease ClpP through IGL loop

motifs that dock into hydrophobic pockets at the top of ClpP. The substrate is degraded by ClpP

in an enclosed catalytic cavity formed by two stacked ClpP heptamer rings[21,22]. The protease

active sites are inaccessible to a folded substrate in the absence of ClpA, but when ClpA is

present the cavity entrance is opened by the interaction with IGL motifs and unfolded substrate is

fed into the cavity by ClpA[23,24]. How ClpA couples large conformational changes required for

substrate translocation, while remaining in contact with ClpP is unknown. Additionally, is the

hexamer-heptamer symmetry mismatch between ClpA and ClpP important to couple unfolding

to proteolysis? Structural details obtained via cryoEM reveal ClpA undergoes conserved AAA+

conformational changes to translocate substrate and conformational plasticity of the IGL loops

allow ClpA to undergo these changes while remaining in contact with ClpP[25].

Most substrates involved in bacterial degradation pathways are specifically targeted for

degradation through recognition of an N-terminal tag that by specific adaptor proteins, which

bind and deliver the tagged substrate to the AAA+ unfoldase via contacts with the AAA+ N-

terminal domain (NTD)[26,27]. The N-degron pathway is conserved from bacteria to eukaryotes and

is an essential proteolytic pathway. Although this pathway is essential, little is known about the

physiological function of the pathway, and many associated molecular mechanisms are poorly

understood[28]. The bacterial adaptor protein ClpS delivers N-degron substrates to the AAA+

ClpA in E. coli (Fig. I.1 b)[29]. The N-degron consists of either a lysine, phenylalanine, tyrosine or

tryptophan added to the N-terminal of targeted substrates, the degron is then recognized by the globular, ClpS substrate binding domain[30,31]. Once bound to the substrate, ClpS affinity is increased for the ClpA NTD and the unstructured ClpS N-terminal extension (NTE) is translocated through the ClpA channel (Fig. I.1 b)[32-34]. However, the mechanisms for the high-affinity complex formation and substrate transfer are not understood. Here I report cryoEM structures of the ClpAPS complex bound to an N-degron tagged GFP substrate. These structures reveal the translocation mechanism observed in the ClpAP complex is conserved between different substrates and low-resolution electron density reveals some insight into the ClpA-NTD and ClpS interaction.

In 2005, small molecule natural product acyldepsipeptides (ADEPs) were discovered to kill multiple strains of multi-drug resistant bacteria and specifically target the bacterial Clp system[35]. ADEPs mimick the ClpA IGL motif-ClpP interaction and inhibit complex formation[36,37]. While this class of small molecules does not kill every strain of bacteria including E. coli, variations of the molecule show potency in multi-drug resistant strains of bacteria such as *Mycobacterium tuberculosis*[38-40]. MtClpP1P2 is distinct from its *E. coli* homolog because it forms a hetero-complex with MtClpP1 and MtClpP2 heptameric rings stacked to form the tetradecamer. The MtClpP2 ring is the only heptamer able to bind to the regulatory AAA+ component, while MtClpP1 is responsible for most of the catalytic activity. The Schmitz and Sello groups report a fragment of the ADEP scaffold that stimulates *M. tuberculosis* ClpXP1P2 proteolytic activity[41-43]. From this work, it is difficult to determine the stimulation mechanism, however, with atomic detail from cryoEM structures, we observe the fragment binds to both MtClpP1 and MtClpP2 hydrophobic pockets and the active sites.
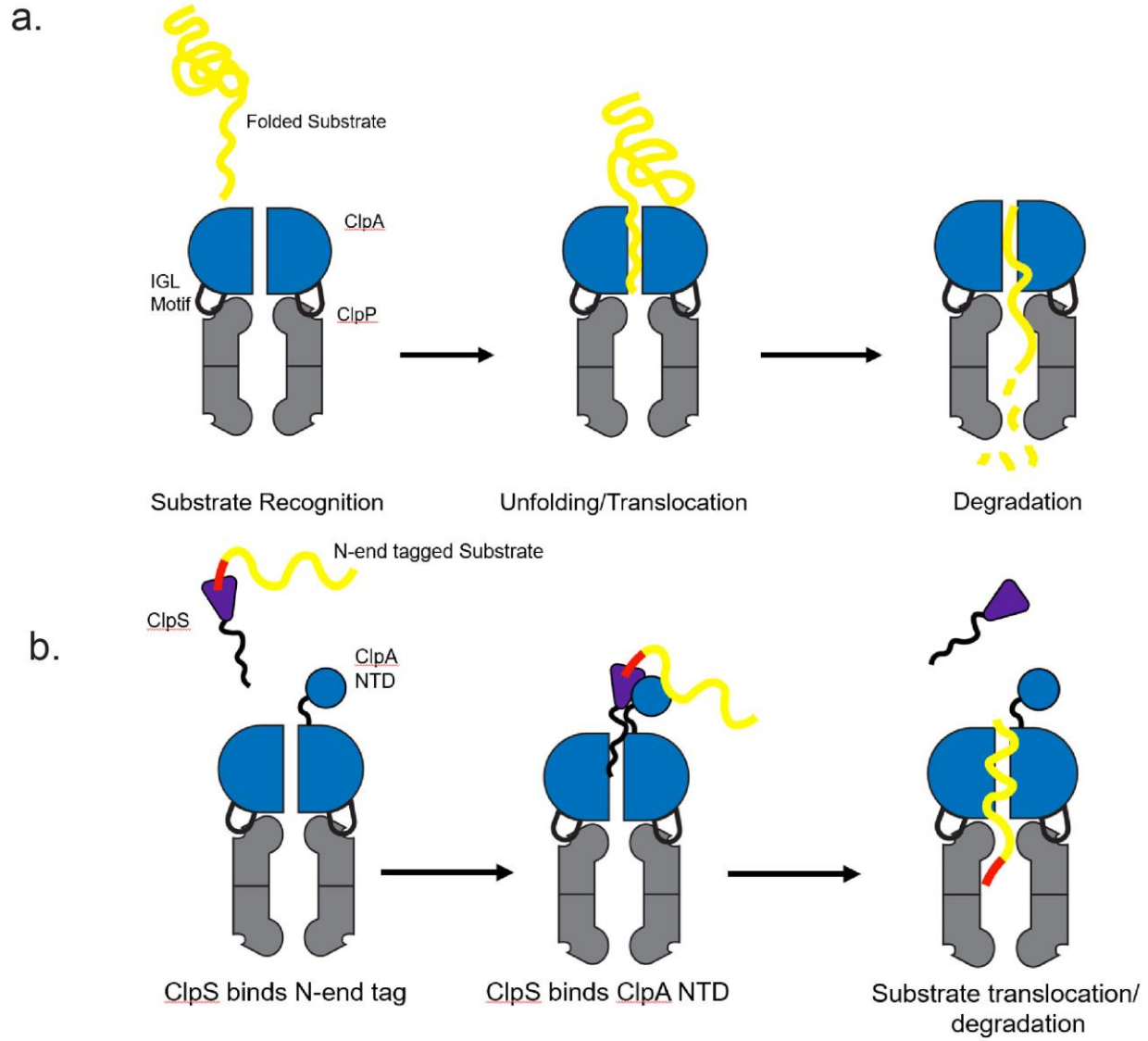
I.1 Model for ClpAP and ClpAPS substrate processing mechanismsa Folded substrate delivery to ClpA followed by translocation and degradation at ClpP. ClpA is blue, substrate is yellow and ClpP is grey. **b** Substrate delivery to ClpA via the adaptor ClpS (purple). ClpS binds the N-degron tag (red), then binds to the ClpA N-terminal domain (blue circle). Substrate is then transferred, translocated and degraded.

References

1.   Balchin, D., Hayer-Hartl, M. & Hartl, F.U. In vivo aspects of protein folding and quality control. *Science* **353**, aac4354 (2016).

2.   Hartl, F.U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324-32 (2011).

3.   Sauer, R.T. et al. Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell* **119**, 9-18 (2004).

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

5.   Olivares, A.O., Baker, T.A. & Sauer, R.T. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat Rev Microbiol* **14**, 33-44 (2016).

6.   Duran, E.C., Weaver, C.L. & Lucius, A.L. Comparative Analysis of the Structure and Function of AAA+ Motors ClpA, ClpB, and Hsp104: Common Threads and Disparate Functions. *Front Mol Biosci* **4**, 54 (2017).

7.   Shorter, J. & Southworth, D.R. Spiraling in Control: Structures and Mechanisms of the Hsp104 Disaggregase. *Cold Spring Harb Perspect Biol* **11**(2019).

8.   Guo, F., Maurizi, M.R., Esser, L. & Xia, D. Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. *J Biol Chem* **277**, 46743-52 (2002).

9.   Xia, D., Esser, L., Singh, S.K., Guo, F. & Maurizi, M.R. Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone. *J Struct Biol* **146**, 166-79 (2004).

10.   Deville, C. et al. Structural pathway of regulated substrate transfer andthreading through an Hsp100 disaggregase. *Sci Adv* **3**, e1701726 (2017).

11.   Gates, S.N. et al. Ratchet-like polypeptidetranslocation mechanism of theAAA+
        disaggregase Hsp104. *Science* **357**, 273-279 (2017).

12.   Rizo, A.N. et al. Structural basis for substrate gripping and translocation by the ClpB
        AAA+ disaggregase. *Nat Commun* **10**, 2393 (2019).

13.   Yu, H. et al. ATP hydrolysis-coupled peptide translocation mechanism of
        Mycobacterium tuberculosis ClpB. *Proc Natl Acad Sci U S A* **115**, E9560-E9569 (2018).

14.   Wickner, S. et al. A molecular chaperone, ClpA, functions like DnaK and DnaJ. *Proc
        Natl Acad Sci U S A* **91**, 12218-12222 (1994).

15.   Kessel, M. et al. Homology in Structural Organization Between E. coli ClpAP Protease
        and the Eukaryotic 26 S Proteasome. *J Mol Biol* **250**, 587-594 (1995).

16.   Hoskins, J.R., Pak, M., Maurizi, M.R. & Wickner, S. The role of the ClpA chaperone in
        proteolysis by ClpAP. *Proc. Natl. Acad. Sci. USA* **95**, 12135-12140 (1998).

17.   Ortega, J., Lee, H.S., Maurizi, M.R. & Steven, A.C. ClpA and ClpX ATPases bind
        simultaneously to opposite ends of ClpP peptidase to form active hybrid complexes. *J
        Struct Biol* **146**, 217-26 (2004).

18.   Hinnerwisch, J., Reid, B.G., Fenton, W.A. & Horwich, A.L. Roles of the N-domains of
        the ClpA unfoldase in binding substrate proteins and in stable complex formation with
        the ClpP protease. *J Biol Chem* **280**, 40838-44 (2005).

19.   Effantin, G., Ishikawa, T., De Donatis, G.M., Maurizi, M.R. & Steven, A.C. Local and
        global mobility in the ClpA AAA+ chaperone detected by cryo-electron microscopy:
        functional connotations. *Structure* **18**, 553-62 (2010).

20.   Kim, Y. et al. Molecular determinants of complex formation between Clp/Hsp100
        ATPases and the ClpP peptidase. *Nature* **8**, 230-233 (2001).

21.   Thompson, M.W., Singh, S.K. & Maurizi, M.R. Processive degradation of proteins by the

       ATP-dependent Clp protease from Escherichia coli. Requirement for the multiple array of

       active sites in ClpP but not ATP hydrolysis. *Journal of Biological Chemistry* **269**, 18209-

       18215 (1994).

22.   Wang, J., Hartling, J.A. & Flanagan, J.M. The Structure of ClpP at 2.3 A˚ Resolution

Suggests a Model for ATP-Dependent Proteolysis. *Cell* **91**, 447-456 (1997).

23.   Effantin, G., Maurizi, M.R. & Steven, A.C. Binding of the ClpA unfoldase opens the

       axial gate of ClpP peptidase. *J Biol Chem* **285**, 14834-40 (2010).

24.   Jennings, L.D., Bohon, J., Chance, M.R. & Licht, S. The ClpP N-Terminus Coordinates

       Substrate Access with Protease Active Site Reactivity. *Biochemistry* **47**, 11031-11040

       (2008).

25.   Lopez, K.E. et al. Conformational plasticity of the ClpAP AAA+ protease couples protein

       unfolding and proteolysis. *Nat Struct Mol Biol* **27**, 406-416 (2020).

26.   Erbse, A. et al. ClpS is an essential component of the N-end rule pathway in Escherichia

       coli. *Nature* **439**, 753-6 (2006).

27.   Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T. & Baker, T.A. The

       molecular basis of N-end rule recognition. *Mol Cell* **32**, 406-14 (2008).

28.   Dougan, D.A., Truscott, K.N. & Zeth, K. The bacterial N-end rule pathway: expect the

       unexpected. *Mol Microbiol* **76**, 545-58 (2010).

29.   Mogk, A., Schmidt, R. & Bukau, B. The N-end rule pathway for regulated proteolysis:

       prokaryotic and eukaryotic strategies. *Trends Cell Biol* **17**, 165-72 (2007).

30.   Zeth, K. et al. Structural analysis of the adaptor protein ClpS in complex with the N-

       terminal domain of ClpA. *Nat Struct Biol* **9**, 906-11 (2002).

31. Roman-Hernandez, G., Grant, R.A., Sauer, R.T. & Baker, T.A. Molecular basis of substrate selection by the N-endrule adaptor protein ClpS. *Proc Natl Acad Sci U S A* **106**(2009).

32. De Donatis, G.M., Singh, S.K., Viswanathan, S. & Maurizi, M.R. A single ClpS monomer is sufficient to direct the activity of the ClpA hexamer. *J Biol Chem* **285**, 8771-81 (2010).

33. Roman-Hernandez, G., Hou, J.Y., Grant, R.A., Sauer, R.T. & Baker, T.A. The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. *Mol Cell* **43**, 217-28 (2011).

34. Rivera-Rivera, I., Roman-Hernandez, G., Sauer, R.T. & Baker, T.A. Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. *Proc Natl Acad Sci U S A* **111**, E3853-9 (2014).

35. Brotz-Oesterhelt, H. et al. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* **11**, 1082-7 (2005).

36. Kirstein, J. et al. The antibiotic ADEP reprogrammes ClpP,switching it from a regulated to anuncontrolled protease. *EMBO Mol Med* **1**, 37-49 (2008).

37. Lee, B.G. et al. Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat Struct Mol Biol* **17**, 471-8 (2010).

38. Ollinger, J., O'Malley, T., Kesicki, E.A., Odingo, J. & Parish, T. Validation of the essential ClpP protease in Mycobacterium tuberculosis as a novel drug target. *J Bacteriol* **194**, 663-8 (2012).

39. Brotz-Oesterhelt, H. & Sass, P. Bacterial caseinolytic proteases as novel targets for antibacterial treatment. *Int J Med Microbiol* **304**, 23-30 (2014).

40.   Malik, I.T. & Brotz-Oesterhelt, H. Conformational control of the bacterial Clp protease by natural product antibiotics. *Nat Prod Rep* **34**, 815-831 (2017).

41.   Carney, D.W. et al. A simple fragment of cyclic acyldepsipeptides is necessary and sufficient for ClpP activation and antibacterial activity. *Chembiochem* **15**, 2216-20 (2014).

42.   Carney, D.W., Schmitz, K.R., Truong, J.V., Sauer, R.T. & Sello, J.K. Restriction of the conformational dynamics of the cyclic acyldepsipeptide antibiotics improves their antibacterial activity. *J Am Chem Soc* **136**, 1922-9 (2014).

43.   Schmitz, K.R. et al. Acyldepsipeptide Antibiotics and a Bioactive Fragment Thereof Differentially Perturb Mycobacterium tuberculosis ClpXP1P2 Activity in vitro. *ACS Chem Biol* (2020).

Chapter 1

Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis

Kyle E. Lopez[1,2,6], Alexandrea N. Rizo[2,3,6], Eric Tse[2], Jiabei Lin[4], Nathaniel W. Scull[5], Aye C. Thwin[2], Aaron L. Lucius[5], James Shorter[4], and Daniel R. Southworth[2]

Affiliations:

[1]Graduate Program in Biophysics, University of California, San Francisco, San Francisco, CA, USA.

[2]Department of Biochemistry and Biophysics, Institute for Neurodegenerative Diseases, University of California, San Francisco, San Francisco, CA, USA

[3]Graduate Program in Chemical Biology, University of Michigan, Ann Arbor, MI, USA

[4]Department of Biochemistry and Biophysics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA

[5]Department of Chemistry, The University of Alabama at Birmingham, Birmingham, AL, USA

[6]These authors contributed to the work equally

Contributions

The work in this chapter is published as Lopez and Rizo *et al.* in the journal *Nature Structural and Molecular Biology* (2020). K.E.L. and A.N.R. carried out all experiments, refinement and modeling procedures for structure determination; developed figures; and wrote and edited the manuscript. E.T. operated the Krios microscope and helped with data collection. J.B.L. performed biochemical substrate-binding experiments. N.W.S. expressed and purified protein components. A.C.T. performed degradation assays. A.L.L. and J.S. wrote and edited the manuscript. D.R.S. designed and supervised the project and wrote and edited the manuscript.

Abstract

The ClpAP complex is a conserved bacterial protease that unfolds and degrades proteins targeted for destruction. Two ClpA AAA+ hexamer rings power substrate unfolding and translocation into the ClpP proteolytic chamber. Here, we determined high-resolution structures of wildtype ClpAP undergoing active unfolding and proteolysis. A spiral of pore loop-substrate contacts span both ClpA AAA+ domains. Protomers at the spiral seam undergo nucleotide-specific rearrangements supporting substrate translocation. IGL loops extend flexibly to bind the planar, heptameric ClpP surface with the empty, symmetry-mismatched IGL pocket maintained at the seam. Three different structures identify a binding-pocket switch by the IGL loop of the lowest-positioned protomer involving release and re-engagement with the clockwise pocket. This switch is coupled to a ClpA rotation and a network of conformational changes across the seam, indicating ClpA can rotate around the ClpP apical surface during processive steps of translocation and proteolysis.

Background

The Hsp100 (Clp) family of proteins, widely present in bacteria and eukaryotes, function as protein unfoldases and disaggregases[1,2]. Some family members can assemble into large proteolytic machines homologous to the 26S proteasome and serve critical roles in targeted protein degradation and quality control[3-7]. Proteolysis requires substrate recognition and ATP hydrolysis-driven unfolding by a AAA+ Hsp100 complex, which unfolds and translocates the substrate into a proteolytic chamber[8-12]. The highly conserved serine protease, ClpP forms this chamber as a double ring of heptamers[13,14], which partner with 1-2 ClpX or ClpA AAA+ hexamers in bacteria, assembling into single and double-capped complexes[15-17]. To promote client degradation, ClpXP and ClpAP are aided by SspB[18,19] and ClpS[20,21], specificity adaptors that promote recognition of substrates including those containing the ssrA degron[22,23] and N-end rule substrates[24], respectively. Other substrates, such as the RepA DNA-binding protein, recognized by ClpA, are remodeled or degraded in support of specific cellular functions[3,25]. Hsp100 interactions with ClpP involve a hexamer-heptamer symmetry-mismatch, which is conserved among proteolytic machines that include the 20S core particle[3,6]. Contacts are mediated by IGF/L-motif loops in ClpX or ClpA and hydrophobic binding pockets on the apical surface of ClpP[6,26]. Engagement of these loops triggers an open-gate conformational change of adjacent N-terminal loops on ClpP, facilitating substrate transfer to proteolytic sites[27-29]. Indeed, the acyldepsipeptide class of antibiotics (ADEPs) compete for binding to these pockets and stabilize an open-gate conformation, thereby converting ClpP to an uncontrolled, general protease[30-33]. How these Hsp100-ClpP interactions are coordinated during active unfolding and translocation is unknown.

ClpA contains two nucleotide-binding AAA+ domains (D1 and D2) per protomer which power unfolding[34]. Structures of related disaggregases, Hsp104 and ClpB, identify the substrate-bound hexamer adopts a right-handed spiral in which conserved, Tyr-bearing pore loops across both AAA+ domains contact and stabilize the polypeptide substrate via backbone interactions spaced every two amino acids[35-38]. Distinct substrate-bound states reveal a ratchet-like mechanism defined by the spiral arrangement, in which the ATP hydrolysis cycle drives substrate release at the lower position and re-binding to the topmost position along the substrate[1,36,39]. A similar spiral architecture and array of substrate contacts has now been identified for many AAA+ machines, supporting a universal rotary translocation mechanism[40-43]. However, for this Hsp100 family it is unclear how the dynamic substrate translocation steps are coupled to proteolysis, or how interactions are maintained at the hexamer:heptamer interface during processive steps of unfolding.

Here, we sought to determine the structural basis for coupled protein unfolding and proteolysis by the ClpAP complex. Using ATP and a RepA-tagged GFP substrate we determined cryo-EM structures of intact, wildtype ClpAP from *E. coli* to ~3.0 Å resolution which reveal three distinct substrate translocation states. Comparison of these state reveals the ClpP-connecting IGL loop of the protomer in lowest substrate-bound position undergoes release and rebinding to the clockwise pocket on ClpP. This IGL switch movement coincides with a ClpA rotation that is supported by conformational plasticity of 5 IGL loops which are bound to the apical surface of ClpP. Nucleotide-specific rearrangements in the AAA+ domains are identified which support a two amino acid-step rotary translocation cycle. Remarkably, these conformational changes are connected across three protomers at the spiral seam revealing an 80 Å allosteric communication path that appears to coordinate IGL-loop switching with substrate

translocation. Together, these results reveal a model in which IGL-loop rearrangements enable ClpA to rotate its position on ClpP consecutively with substrate translocation steps thereby coupling substrate unfolding with ClpP activity.

Results

Architecture of active, substrate-bound ClpAP

Structures of wildtype ClpAP undergoing active substrate unfolding and proteolysis were desired in order to capture functional states. RepA-GFP constructs are proteolyzed by ClpAP and can be used to monitor unfolding by ClpA[10,44,45]. Therefore, RepA-GFP containing the first 25 residues of RepA (RepA[1-25]-GFP) was tested for proteolysis and complex formation (Supplementary Fig. 1.S1 a-c). While the slowly-hydrolysable analog, ATPγS, enables stable formation of AAA+ complexes containing translocated substrates[36,37], the reduced hydrolysis impairs function[10] and may limit the ClpAP conformational cycle required for unfolding and proteolysis. Indeed, substantial degradation of RepA[1-25]-GFP occurs within 15 minutes in the presence of saturating (10 mM) ATP while little degradation is observed with ATPγS (Supplementary Fig. 1.1S a). Therefore, in order to achieve active ClpAP for cryo-EM, incubations were carried out initially with ATPγS to promote assembly and then, following purification of substrate-bound complexes, 10mM ATP was added to initiate unfolding and proteolysis prior to vitrification. Assembly with ATPγS and mixtures with ATP have been previously established to support ClpA function[44,46] and we identify robust degradation occurs under these ATPγS-ATP conditions, indicating ClpAP is active prior to vitrification (Supplementary Fig. 1.S1 b,c).

In reference-free 2D class averages, side and top views of ClpP particles double-capped with ClpA predominate suggesting the presence of two rings of different diameters, indicative of

ClpAP (Fig. 1.1 a and Supplementary Fig. 1.S1 d). While two ClpA hexamers were identified, one typically showed well-resolved features compared to the other, indicating preferred alignment likely due to flexibility across the double-capped complex. 3D classification yielded three distinct ClpAP conformations which refined to high-resolution (2.7-3.2 Å), hereafter referred to as the Engaged-1 (ClpAP^{Eng1}), Disengaged (ClpAP^{Dis}) and Engaged-2 (ClpAP^{Eng2}) states based on the binding states of the IGL loops (Supplementary Fig. 1.S1 e, f). As with 2D analysis, one ClpA hexamer showed improved features over the other. Therefore, the final models included one ClpA hexamer and two ClpP heptamers. In all states the D1 and D2 AAA+ rings of the ClpA hexamer adopt a right-handed spiral with the D2 ring contacting the planar, heptameric surface of ClpP via the IGL loops (residues 611-623) (Fig. 1.1 b). ClpA is comprised of protomers P1-P6 with P1 at the lowest and P5 at the highest position of the spiral, while P6 is asymmetric and positioned at the seam interface (Fig. 1.1 b). This architecture is similar to related ClpB and Hsp104 double-ring disaggregases in their substrate-bound states[35-37]. Resolution is the highest for ClpP (~2.5 Å), while ClpA is more variable (~2.5-4.5 Å for ClpAP^{Eng1}, ~3-6 Å for ClpAP^{Dis}, and ~3-6 Å for ClpAP^{Eng2}), with the spiral seam protomers (P1, P5 and P6) at lower resolutions due to their flexibility, based on local resolution estimates (Supplementary Fig. 1.S1 g-i). The high-resolution of the maps permitted accurate atomic models to be built for the full ClpAP complex (Fig. 1c, Supplementary Fig. 1j-k; Supplementary Table 1.1). Density for the flexible N-terminal (NT) domain of ClpA (residues 1-168) was not well-resolved, and thus was not modeled.

In all three states density corresponding to an unfolded polypeptide substrate is identified along the ~80 Å-length of the ClpA channel, spanning the D1 and D2 domains (Fig. 1.1 d, 1.2 a-c). The density is well-resolved and modeled as a 24-residue poly-Ala chain (Fig. 1.1 d).

Additional substrate density is not observed in the ClpP pore or chamber potentially due to flexibility and the absence of substrate-interacting residues, such as the pore loops in ClpA. Notably, in low-passed filtered maps of the final reconstruction, globular density at the entrance to the ClpA channel is visible at a reduced threshold that approximately corresponds to a GFP molecule (Supplementary Fig. 1.S1 j). These data along with the SEC and proteolysis analysis (Supplementary Fig. 1.S1 a-c) indicate that the wildtype ClpAP structures determined here are captured with a bound RepA-GFP substrate under active conditions using ATP, and thus represent conformational states that are associated with processive cycles of translocation and proteolysis.

ClpAP structures reveal IGL-loop switches to engage ClpP symmetry-mismatched pocket

Following multiple rounds of 3D classification three distinct conformations of substrate-bound ClpAP refined to high resolution (Fig. 1.2 a-c, Supplementary Fig. 1.S1 e-f). The major conformational differences involved the ClpA seam interface and include changes in substrate interactions and nucleotide states (discussed below), and changes in the IGL loops and position of ClpA. No substantial conformational differences are identified for ClpP between the different states (RMSD < 1 Å). In the ClpAP^{Eng1} structure, well-resolved density for the IGL loops from all 6 ClpA protomers is identified in the pockets around the ClpP apical surface (Fig. 1.2 a, d). One remaining empty pocket on ClpP, which results from the symmetry mismatch of the heptamer, is positioned at the ClpA spiral seam between protomers P1 and P6 (Fig. 1.2 a, d). In the ClpAP^{Dis} structure, density for the IGL loop of protomer P1, which is at the lowest position along the substrate, is no longer observed in the ClpP pocket, resulting in two neighboring empty pockets at the ClpA seam (Fig. 1.2 b, d). Remarkably, in the ClpAP^{Eng2} structure, density for the P1 IGL loop is observed in the clockwise adjacent pocket, revealing that the loop has switched

17

position in comparison with ClpAP[Eng1], and the empty, symmetry-mismatched pocket now resides between protomers P1 and P2 compared to P6 and P1 for the ClpAP[Eng1] (Figure 1.2 c, d). Difference maps of the ClpA-P interface region further validate the position of the P1 IGL loop in these structures (Supplementary Fig. 1.S2). If these structures represented a mix of states then the difference maps would show positive density in both IGL pockets. Importantly, however, positive density for the IGL loop only appears from P1 in the correct ClpP pocket corresponding to the ClpAP[Eng1] or ClpAP[Eng2] states, thereby verifying that these structures represent distinct states of the P1 IGL loop. Notably, for ClpAP[Dis], density for the P1 IGL loop is not observed in either pocket, indicating this loop is indeed unbound from ClpP and in an intermediate state (Supplementary Fig. 1.S2).

The ClpA channel and bound polypeptide substrate are substantially offset, between ~14° and 16°, from the ClpP pore in the different structures (Fig. 1.2 a-c). Upon alignment of the structures, ClpA is identified to be in three distinct positions relative to ClpP. These differences appear to occur through a pivot across ClpP and clockwise rotation around the substrate channel axis which coincides with the binding site-switch of the P1 IGL loop (Fig. 1.2 a-c). Going from the Engaged-1 to Disengaged states, ClpA pivots towards the P5-P6 side of the hexamer, shifting by approximately 10 Å across ClpP. From the Disengaged to Engaged-2 states ClpA rotates clockwise, resetting the orientation of the channel relative to ClpA but with an overall rotation of ~10° compared to ClpAP[Eng1]. The ClpA rotation is visualized in a morph between these states, revealing how protomers P4-P6 tilt towards ClpP, compressing the interface in this region and then expand through a clockwise rotation around the axial channel in the ClpAP[Eng2] state. Since these changes coincide with disengagement of the P1 IGL loop, this rotation of ClpA relative to ClpP likely facilitates IGL-loop switching.

In addition to these structures, we determined structures of ATPγS-stabilized ClpAP bound to RepA[1-25]-GFP in which ATP was not added prior to vitrification (Supplementary Fig. 1.S3). Following similar data classification and refinement procedures, we determined two ClpAP structures at 3.0 and 3.1 Å resolution which match the ClpAP[Eng1] and ClpAP[Dis] states described above (Supplementary Fig. 1.S3 a-g). Notably, the ClpAP[Eng2] state was unable to be classified as a distinct conformation despite similar-sized datasets. This could be due to changes in the conformational equilibrium resulting from the ATPγS-stabilized conditions compared to active conditions with ATP. Nonetheless, these structures further establish that P1 IGL loop undergoes engaged and disengaged conformational changes under conditions in which substrate binding and processing occurs.

IGL-loop plasticity enables ClpP engagement by the ClpA spiral

Previous crystal structures of ClpA were unable to resolve the IGL loops due to flexibility, but biochemical data for ClpX IGF loops suggest that they make static interactions with ClpP and all 6 IGF loops are required for optimal activity[47]. In the ClpAP structures, density for the ClpA IGL loops is well-defined, enabling atomic modeling for nearly all loop residues in each pocket (Supplementary Fig. 1.S4 a). The IGL-loop region extends from residues N606 and T637 in the base of the D2 large subdomain as two short α-helices. Residues 616-620 form the flexible loop, which extends into the hydrophobic binding pocket on ClpP, resulting in ~600 Å^2 of buried surface area compared to the empty pocket (Fig. 1.3 a, left). The IGL-loop binding pocket is formed by the interface of two ClpP protomers and includes α-helices B and C from one protomer and a 3-strand β-sheet (strands 1, 2, and 3) and the C-terminal (CT) strand from the adjacent protomer (Fig. 1.3a). The loop residues I617, G618, L619, and I620 bind a hydrophobic region in the pocket comprised of A52, L48, F49, and F82 in α-helices of one protomer and L23,

Y60, Y62, I90, M92, F112, L114, L189 in the adjacent protomer (Fig. 1.3 a, middle). Additional electrostatic contacts likely stabilize the loop as well, including R192 in the CT strand, and E26, which appear to interact with H621 and R614 and Q622, respectively (Fig. 1.3 a, right).

While the IGL loops all make identical contacts with ClpP, flexibility of the connecting helices (residues 608-615 and 624-635) enables the loops to extend from ClpA in a number of orientations around the hexamer and between the different states (Supplementary Fig. 1.S4 b). The largest changes occur with the P1 loop, which switches binding pockets on ClpP between the three states, as discussed above (Fig. 1.3 b-d). The loop is largely well-resolved in the ClpA^{Eng1} and ClpA^{Eng2} states, however residues 609-624 were unable to be modeled for ClpA^{Dis} due to weak density in the unbound, disengaged conformation. By comparison of the P1 IGL loop between the different states, the binding-pocket switch is identified to result from the overall clockwise rotation of ClpA (Fig. 1.2 a-c) and a large, 80° rotation of the loop around residues T604 and T637 in the connecting helices (Fig. 1.3 c,d). Surprisingly, the P5 IGL loop is also identified to contract and extend between the states through a partial unfolding of both connecting helices (residues 609-613 and 614-629) (Fig. 1.3 e). In the Engaged-1 state, the loop is extended by ~5 Å compared to the Disengaged state, whereas in the Engaged-2 state the P5 loop is partially extended by ~ 3 Å. Notably, this loop extension is only observed at the P5 position and appears to correlate with the orientation of ClpA in the different states. Overall, these results reveal a remarkable conformational plasticity of the IGL loops that likely functions to support strong ClpP interaction around the variable hexamer-heptamer interface and enable the binding-pocket switch movement of the P1 loop.

ClpP structure and N-terminal gating

The flexible N-terminal loop residues of ClpP (1-18) form a pore on the apical surface that functions as a substrate gate that is allosterically controlled by engagement of the adjacent IGF/L-binding pockets by ClpX/A or ADEP compounds[33]. In all three ClpAP structures, the ClpP NT loops from each protomer are well-resolved and adopt an extended configuration resulting in an open gate conformation that is positioned adjacent the ClpA translocation channel, ~30 Å away from where substrate is resolved (Fig. 1.4 a and Supplementary Fig. 1.S5 a,b). This is distinct from crystal structures showing the NT loops adopt an asymmetric open-gate arrangement[48], but similar to ADEP-bound structures where all the loops are in an extended conformation[31,33]. Additionally, no contact is observed between the NT loops and ClpA (Fig. 1.4 a), which may be distinct compared to ClpXP, in which NT loops have been identified to contact the ClpX pore-2 loops[47].

We identify two specific interactions: one across the ClpP NT loops and one with an adjacent helix A in the IGF/L pocket, which have not been previously characterized and appear to stabilize the open gate conformation (Fig. 1.4 b). A salt-bridge contact between residues R15 in one loop and E14 in the clockwise loop is identified in each protomer (Fig. 1.4 b and Supplementary Fig. 1.S5 c). Additionally, a potential salt-bridge contact involving E8 and K25 is also observed which may additionally stabilize the loop orientation (Fig. 1.4 b and Supplementary Fig. 1.S5 d). Notably, K25 is located in a helix that comprises part of the hydrophobic, IGL-binding pocket (Fig. 1.4 b). Thus, this interaction may be involved in the allosteric gating mechanism.

In an initial dataset of ATPγS-stabilized ClpAP, we identified a population of single-capped complexes which resolved into one 3D class (Fig. 1.4 c and Supplementary Fig. 1.S5 e,f),

enabling us to characterize the open and closed-gate conformations in one structure. While the resolution of the NT loops was not sufficient to model the closed conformation, at lower threshold values, density for the loops on the unbound end of ClpP appears to extend ~8 Å from ClpP, while density for the ClpA-bound end NT loops extends ~16 Å (Fig. 1.4 d). Additionally, the pore diameter is identified to be ~25 Å for the ClpA-bound end of ClpP, which is substantially wider compared to the unbound end, which is ~15 Å (Fig. 1.4 e). Thus, we identify the NT loop gating mechanism is specifically triggered by engagement of the cis-bound ClpA IGL loops, which may allosterically regulate the NT loops through salt bridges that stabilize the extended loop arrangement.

ClpA structure contacts and translocation states

To improve the resolution of the ClpA pore loop interactions and the seam protomers, particle subtraction and focused refinement of the ClpA hexamer was performed (Supplementary Fig. 1.S6 a). This resulted in an estimated resolution of 3.0Å and 3.1Å and 3.4Å for the ClpA^{Eng-1}, ClpA^{Dis} and ClpA^{Eng-2} focused maps, respectively (Supplementary Fig. 1.S6 b). While the overall resolution did not increase compared to the full map containing ClpP, improvements in the map density for the seam protomers and substrate contacts is observed, particularly for the Engaged-1 state (Supplementary Fig. 1.S6 c-e). Nonetheless, the seam protomers remain at a lower resolution (~3.5-6 Å) compared to the rest of the map, due to their flexibility. Models were further refined using these maps in order to characterize the substrate interactions and conformational changes between the states. Similar to other AAA+ structures, the conserved Tyr-pore loops in the D1 and D2 of ClpA extend into the channel and form a double spiral of substrate interactions spaced every two amino acids along a 24 amino acid-long polypeptide (Fig. 1.5 a). For all states, the D1 stabilizes a 9-residue segment through direct contact by Y259

from protomers P1-P4, which intercalates between the substrate side chains and contacts the backbone (Fig. 1.5 b, Supplementary Fig. 1.S6 f). The conserved flanking residues, K258 and R260, extend laterally to make electrostatic contacts with the upper and lower adjacent pore loops (D262 and E264), similar to ClpB D1[37,38]. Notably, in the ClpAP[Eng-1] structure the P5 and P6 D1 pore loops are disconnected from the substrate, with Y259 positioned ~18 Å and ~17 Å away, respectively (Fig. 1.5 a and Supplementary Fig. 1.S6 g). This 4-bound, 2-unbound configuration of the D1 pore loops is distinct from previous structures of ClpB and Hsp104[36-38]. The D2 similarly shows well-defined pore loop-substrate contacts for protomers P1-P4 in both states (Fig. 1.5 c, Supplementary Fig. 1.S6 f). These interactions stabilize a longer, 11 residue polypeptide segment and are primarily mediated by Y540 and V541, which form a Y-shaped clamp around the substrate backbone. Additional, pore-2 loops[49,50], conserved in ClpB and Hsp104[36-38], are present in both the D1 (residues: 292-302) and D2 (residues: 613-625), and line the channel, likely making additional contributions to stabilizing the polypeptide. Notably, residues E526, R527 and H528 from protomers P1-P5 contact the substrate and together form an "exit pore" which is adjacent the ClpP gating loops and thus may server to facilitate transfer to the ClpP chamber (Supplementary Fig. 1.S6 h,i).

As with previous Hsp100 structures[36-38], protomers P2-P4 show no substantial conformational changes between the states. Therefore, in order to compare conformational changes of the seam protomers (P1, P5 and P6), protomer P3 was used for alignments of the ClpA hexamer. The largest changes occur for these protomers between ClpAP[Eng-1] and ClpAP[Dis], and between ClpAP[Eng-1] and ClpAP[Eng-2], (RMSD ≈ 5.1 Å and 3.5 Å, respectively), while changes between ClpAP[Dis] and ClpAP[Eng-2] are more modest (RMSD ≈ 2.3 Å). For simplification, comparisons between ClpAP[Eng-1] and ClpAP[Eng-2] are shown (Fig. 1.5 b, c). Overall, the pore

loops for P5 and P6 shift closer to the polypeptide substrate and move up the channel axis going

from ClpAP^{Eng-1} to the ClpAP^{Dis} and ClpAP^{Eng-2} states, while the position of the P1 pore loop

does not change relative to the substrate (Fig. 1.5 b, c). Notably, the P5 pore loop moves up by

~4 Å and towards the substrate by ~8 Å (Supplementary Fig. 1.S6 g). This positions P5 Y259

adjacent the substrate two residues above the P4 Y259 position, however direct contact is not

identified. The largest changes occur with D2 pore loop of protomer P6, which moves up the

channel axis by ~7 Å, corresponding to a two-residue shift in the substrate position, but remains

unbound to substrate in all three states (Fig. 1.5 c). Together these changes reveal protomer

movements up substrate axis and appear on-path to a translocation step through engagement of

the next contact site along the substrate by the D1 in protomer P5 (Supplementary Fig 1.S6 j). In

order to identify how these changes are connected to the IGL loop movement, the C-α deviation

between the three states was mapped onto the hexamer model (Figure 1.5 d). As expected, the

IGL loops of the seam protomers show the greatest variability while protomers P2-P4 show little

change. Importantly, distinct regions of variability are identified at the spiral seam, revealing a

path of conformational changes that extend from the C- to N-termini across P1, P6 and P5,

respectively. These changes occur in the IGL loop and D2 small subdomain of P1, the D2 large

subdomain of P6 and the D1 large subdomain of P5 (Fig. 1.5 e). Remarkably, these changes

reveal an 80 Å-long allosteric communication path which appears to connect IGL-loop

movement in P1 to translocation steps that occur in P5 and P6.

Nucleotide states support hydrolysis-driven translocation

Similar to Hsp104 and ClpB, ATP hydrolysis activity in D1 and D2 is required for ClpA substrate translocation steps[51]. All three structures show well-resolved nucleotide pockets and the nucleotide state of each pocket was assessed based on the density of ATP and the position of the trans-activating Arg-finger residues (R339-R340 in the D1 and R643 in D2) (Supplementary Fig. 1.S7). For the substrate bound protomers P3 and P4, the D1 and D2 nucleotide pockets are identical across the three states and in an ATP, active configuration (Fig. 1.6 a and Supplementary Fig. 1.S7 a). However, for the seam protomers P1, P5 and P6 the nucleotide states vary, largely supporting models for hydrolysis-driven translocation previously described[36-38] (Fig. 1.6 a and Supplementary Fig. 1.S7 b). The P5-D1 appears to switch from an ADP state in ClpA^{Eng-1} to an ATP state in ClpA^{Dis} and ClpA^{Eng-2}, indicating nucleotide exchange may occur between these states. Notably, this coincides with the conformational changes that bring the P5-D1 pore-loop towards with the next contact position along the substrate after P4 (Figure 1.5 b), supporting current models in which a translocation step occurs upon ATP re-binding[1]. Conversely, the P5-D2 is in an ATP state and bound to substrate in all three structures (Fig. 1.6 a and Supplementary Fig. 1.S7 b). Protomer P6, which is at the spiral seam and unbound to substrate, is in a post-hydrolysis, ADP state for both the D1 and D2 across all three structures. For protomer P1, which is at the lowest substrate-contact position and undergoes IGL-loop switching between the states, the D1 appears bound to ATP in ClpA^{Eng-1} and bound to ADP in ClpA^{Dis} and ClpA^{Eng-2}, indicating hydrolysis likely occurs between these states. However, the P1-D2 appears bound to ADP all states. Notably, the clockwise P2-D2, which is activated by the P1 Arg finger, appears to switch from an ATP to an ADP state between ClpA^{Dis} and ClpA^{Eng-2} structures.

Together, the changes in nucleotide states between the three structures indicate ATP hydrolysis occurs at the spiral seam and proceeds counter-clockwise around the hexamer, supporting the rotary substrate translocation cycle in which protomers toward low position in the spiral (P1 and P2) undergo ATP hydrolysis and substrate release, then re-bind substrate at the top position (P5) with ATP binding[1] (Fig. 1.6 a and Supplementary Fig. 1.S7 a). Based on the different D1-D2 nucleotide states within protomers P1, P2 and P5, hydrolysis may be asynchronous, likely initiating in the D2 ring based on the ATP-ADP change identified for P2 between the ClpA^{Dis} and ClpA^{Eng-2} structures. This finding is similar to what is identified for ClpB[37] and indicates the D1 and D2 regulate distinct steps of translocation and coordination with ClpP. Surprisingly, certain conformational changes, including release of substrate, movement of the P1 IGL loop, and changes in P6, do not appear to directly correlate with changes in the cis nucleotide pocket. Allosteric communication and distinct functional roles have been described for the D1 and D2 of ClpB[52,53]. Thus, hydrolysis at adjacent sites, either across the D1 and D2 or between protomers connected by the Arg finger, may allosterically drive the conformational changes identified in the different structures. Indeed, the P1 IGL loop switching may be supported by hydrolysis at P1-D1 during disengagement (ClpA^{Eng-1} to ClpA^{Dis}) and at P2-D2 during engagement of the next pocket (ClpA^{Dis} to ClpA^{Eng-2}) (Fig. 1.6 a and Supplementary Fig. 1.S7 b).

Discussion

Structures of conserved substrate-bound AAA+ translocases, including the double-ring disaggregases ClpB and Hsp104, have revealed a spiral staircase of pore-loop:substrate contacts and a dynamic translocation mechanism involving ATP hydrolysis-driven substrate release and rebinding by the protomers at the spiral seam of the hexamer[36-38,54]. For AAA+ proteases such as

ClpX and ClpA it has been unclear how these conformational changes could occur with an attached heptameric protease during processive substrate translocation. To understand the mechanism of coupled protein unfolding and proteolysis, we determined structures of the wildtype *E. coli* ClpAP complex bound to a RepA-GFP substrate during active proteolysis with ATP. Three distinct structures, ClpAP[Eng-1], ClpAP[Dis], and ClpAP[Eng-2], were determined and reveal a dynamic ClpA-P interface in which the connecting IGL loops undergo large conformational changes that enable the ClpA hexamer to rotate on the ClpP apical surface during substrate translocation steps (Fig. 1.2). Most notably, the IGL loop of the protomer in the lowest substrate-bound site (P1) is identified in three different positions that together reveal a clockwise binding-pocket switch movement. This IGL loop movement appears coordinated with conformational changes associated with the substrate translocation steps, revealing allosteric communication path across the seam protomers which connects ClpP interactions with the pore loop-substrate contacts (Fig. 1.5 d).

For processive unfolding and proteolysis ClpA is expected to remain bound to ClpP during multiple translocation cycles[2,55,56]. Therefore, based on the three structures, we propose a rotary model for the ClpAP mechanism in which hexamer-heptamer symmetry mismatch is continually maintained with an empty IGL binding pocket aligned at the spiral seam of ClpA. During substrate translocation, the IGL loop of the adjacent protomer (P1) at the lowest substrate contact site, disengages from ClpP (ClpAP[Eng-1] to ClpAP[Dis], step 1) then re-binds to the clockwise empty pocket (ClpAP[Dis] to ClpAP[Eng2], step 2) in a manner that is regulated by ATP hydrolysis and conformational changes associated with substrate release and re-binding (Fig. 1.6 b). For a processive cycle, we propose that these steps could continue with IGL loop switching at each translocation step, rotating the position of empty IGL pocket around ClpP with the spiral

seam (Fig. 1.6 b). This would result in a slow rotation of the ClpA relative to translocation, in which the hexamer would shift by one clockwise binding position on the ClpP apical surface per 6 substrate translocation steps down the axial channel. This is based on established two amino acid-step translocation models involving consecutive hydrolysis around the ring[1,39,57]. Other models involving larger translocation step sizes[58] or alternate hydrolysis mechanisms[59] would likely confer different coordination with IGL loop switching. Nonetheless, by undergoing theses conformational changes, ClpA can maintain the spiral arrangement of substrate contacts and tight interaction with ClpP continuously during unfolding and proteolysis. This model suggests a functional significance of the hexamer:heptamer mismatch is that the 7th binding pocket on ClpP is available for the IGL loops to sequentially switch position at the spiral seam, enabling processivity without altering other contact arrangements with ClpP. This processivity model could be substrate-specific and possibly more critical for proteolysis of stable, folded substrates compared to more labile structures.

While other mechanisms may support substrate translocation and proteolysis by ClpAP, we note that IGL-loop switching between the same sites, stochastically or counterclockwise would result in an offset between the empty IGL pocket and the spiral seam of ClpA hexamer, given the AAA+ rotary-type mechanisms previously proposed[36,40,43]. None of these potential configurations of ClpAP were observed in any of the 3D classes for our ATPγS and ATP datasets. Additionally, recent structures of the related ClpXP complex bound to substrate identify conformations which are similar to the Engaged-1 and Disengaged states determined here and a complimentary rotary mechanism is proposed[60]. The additional Engaged-2 state structure determined here further supports these models by identifying that the P1 IGL loop indeed switches position and ClpA rotates clockwise relative to ClpP with an apparent substrate

translocation step. The discovery of this additional state in our study may have resulted from the use of WT enzyme and ATP, allowing an additional active state of substrate translocation to be captured.

We identify IGL-loop interactions with the ClpP hydrophobic pockets are identical at all positions, while flexibility of the helices that connect the loops to the D2 base of ClpA enables substantial variability in the ClpA position relative to ClpP. This flexibility is likely critical for maintaining ClpP binding during ratcheting conformational changes associated with substrate translocation and the rotations in ClpA between the different states (Fig. 1.2 a-c). Furthermore, the extension and unfolding of the P5 IGL-loop helices in the Engaged-1 state is striking and may also provide energetic constraints that could facilitate release and clockwise switch of the P1 IGL loop during the conformational change to the Disengaged and Engaged-2 states (Fig. 1.3 c).

The pore-loop spacing along the substrate and conformational changes between the Engaged and Disengaged states are consistent with a two amino-acid step translocation mechanism, similar to previous studies[35-37,43], but is smaller than step sizes reported for ClpA and ClpX by single molecule[58,61] and transient state kinetic methods[62]. However, recent studies with ClpB identify rapid and consecutive translocation modes in which this fundamental step size is not resolvable due to the high translocation speed[63]. The nucleotide states of ClpA are similar to those identified in related AAA+ complexes[35-37,43] in which the substrate-bound protomers are bound to ATP (P2-P4) and the protomers towards the seam to substrate appear to be in a post-hydrolysis ADP or apo state (Fig. 1.6 a and Supplementary Fig. 1.S6). Thus, we propose a similar consecutive translocation mechanism in which hydrolysis by protomers in the lower position along the substrate (P1-P2) triggers substrate release while ATP re-binding facilitates

binding to the next position along the substrate[1,36]. Notably, with this clockwise cycle, hydrolysis may occur first in the D2 ring and precede substrate release, based on the apparent change nucleotide state in P2 for ClpAP[Dis] and ClpAP[Eng-2] structures. Similar changes are also observed in structures of ClpB[37]. Finally, we identify a path of conformational variability between the states which connects the P1 IGL loop with the P5 pore loops and substrate contacts (Figure 1.5 d). These conformational changes are likely allosterically coupled to neighboring D1-D2 nucleotide states, thereby enabling IGL-loop switching to be hydrolysis driven and coordinated with substrate translocation.

A key discovery of this work is that for processive cycles of unfolding, ClpA may rotate around the apical surface of ClpP in coordination with the substrate translocation steps, thereby maintaining spiral seam aligned with hexamer:heptamer mismatch (Fig. 1.6 b). Thus, while binding by IGF/L loops is well-understood to trigger gate-opening in ClpP, the conformational plasticity and asymmetric binding interactions we identify reveal new insight into how these loops facilitate allosteric regulation between ClpA and ClpP[62,64]. Indeed, the IGF/L loops enable the AAA+ hexamer to simultaneously engage the ClpP planar surface and undergo dynamic conformational changes. A number of proteolytic machines operate as hexamer:heptamer assemblies[3,4] and recent structures of ClpXP reveal its IGF loops are similarly flexible, supporting a conserved rotary mechanism[60,65]. Notably, assembly of the eukaryotic Rpt and archaeal PAN AAA+ with its respective 20S core, a threonine protease with a heptameric ring, involves interaction with flexible C-terminal HbYX motifs and gate-opening of the 20S[66,67]. While the HbYX interactions are distinct and likely operate differently during translocation, recent structures reveal a conserved spiral staircase arrangement of 26S[41,68] and PAN[69] bound to substrates and a sequential rotation of the ATPase ring has been proposed for PAN[69]. For the Clp

protease system the symmetry mismatch likely serves a critical role in processivity by presenting

an available binding pocket for IGF/L loop movement during consecutive translocation steps,

thereby utilizing binding asymmetry to coordinate the rotary ATPase cycle and translocation

directionality by the AAA+ domains with substrate transfer and proteolysis by ClpP.

Figures
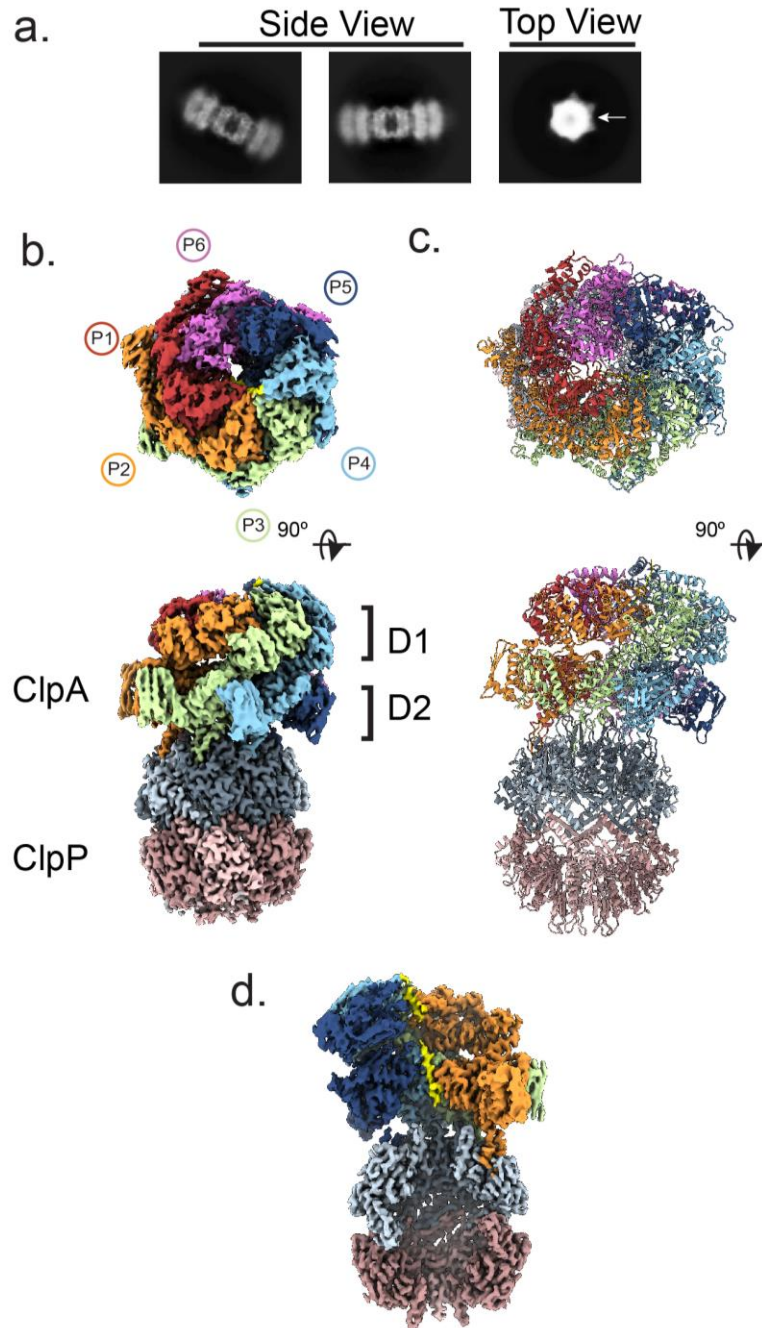
Figure 1.1 Architecture of the substrate-bound ClpAP complex a, Side and top view 2D class averages of double-capped ClpAP. Rings corresponding to ClpA (arrow) and ClpP rings are identified in the top views. **b**, Top and side views of the final ClpAPEng1 map. **c**, Model of ClpAPEng1. ClpA is colored by individual protomers, as indicated. **d**, Channel view showing substrate peptide bound to ClpA (yellow).
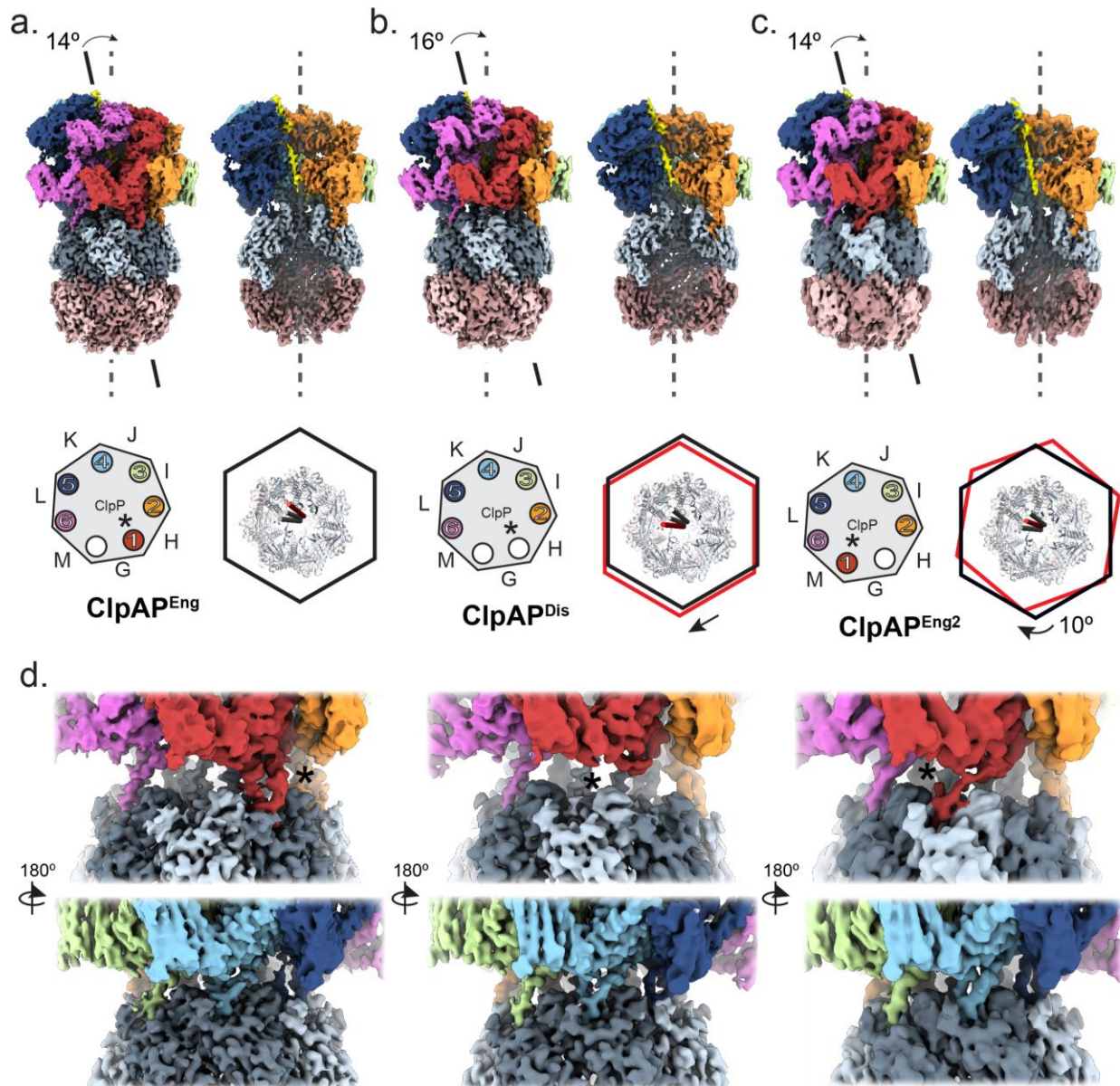
Figure 1.2 Three distinct structures of ClpAP showing IGL loop rearrangement. a–c, Cryo-EM maps for ClpAPEng1 (**a**), ClpAPDis (**b**) and ClpAPEng2 (**c**) showing the degree offset (arrow) of the ClpA channel axis (solid line) and substrate position (yellow density) compared to the ClpP pore and proteolytic chamber (dashed line). Schematic (lower left) shows occupancy of the ClpA IGL loops (circles, colored and numbered by protomer) around the ClpA hexamer, with the empty IGL pockets (white circles) and ClpA protomers indicated (letters) for the different states. Schematic (lower right) shows top view of ClpP with ClpA as a hexagon overlay (red, current state; black, previous state), and colored cylinders indicating substrate positions (red, current state). **d**, Cryo-EM density of the ClpA–P interface showing IGL loop interaction with ClpP in ClpAPEng1 (left), ClpAPDis (center) and ClpAPEng2 (right).
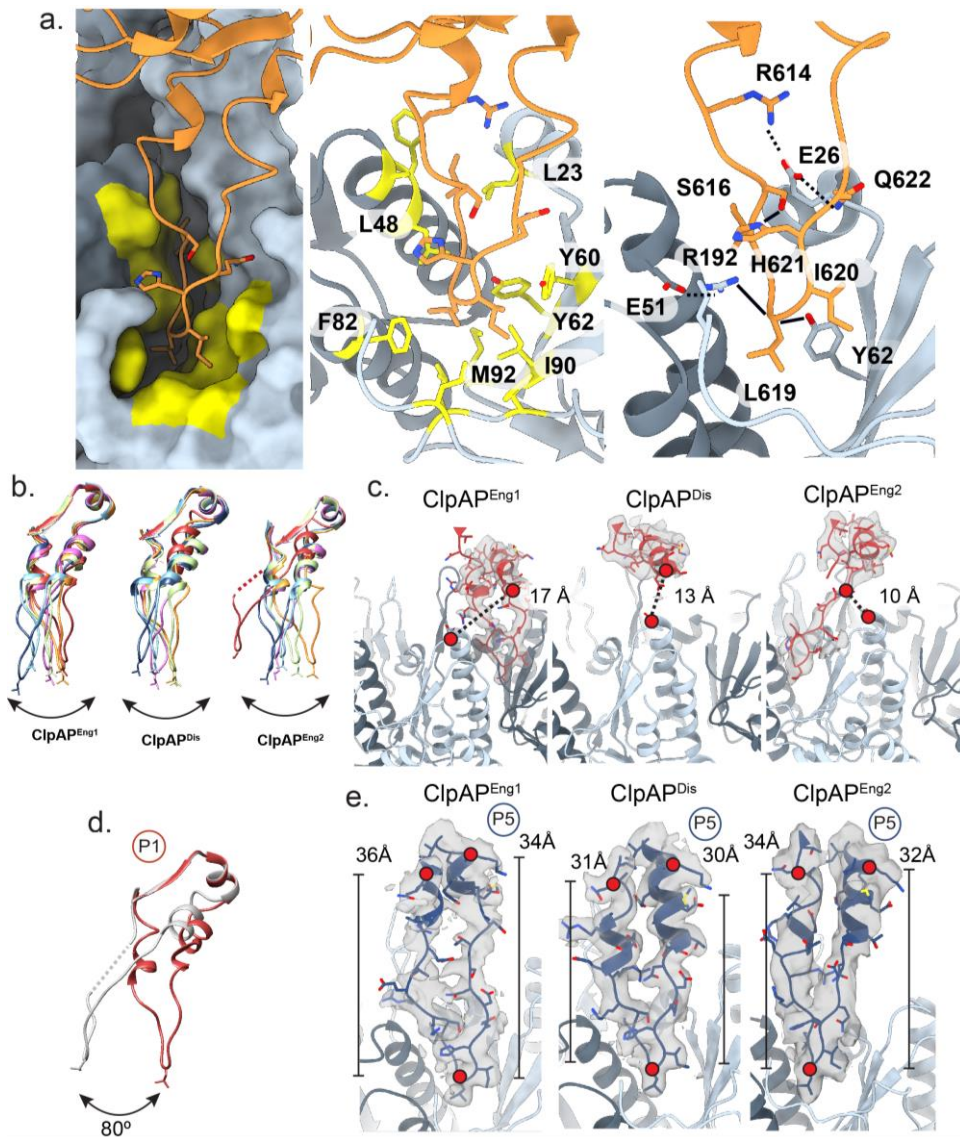
Figure 1.3 IGL loop interactions and conformational flexibility. a, Representative view of a bound IGL loop (orange, ribbon view) positioned in the binding pocket of ClpP, shown in surface view with hydrophobic residues colored in yellow (left), and shown in ribbon views with hydrophobic interactions (middle) and electrostatic contacts (right) labeled. **b**, Overlay of IGL loops (colored by protomer) of ClpAPEng1 (left), ClpAPDis (middle) and ClpAPEng2 (right). IGL loops are aligned to connecting residues 638–649. Dotted line represents missing residues not present in the density. **c**, Map and model showing that P1 IGL loop density extends into the IGL pocket for ClpAPEng1 (left) and ClpAPEng2 (right) but is disengaged for ClpAPDis (middle), contacting the adjacent apical ClpP surface (right). The distances between ClpP E67 and ClpA–P1 S625 in the three states are shown to indicate the shift in position of the P1 IGL loop relative to ClpP. **d**, Overlay of IGL loops of P1 for ClpAPEng1 (red) and ClpAPEng2 (gray). **e**, Map and model of the P5 IGL loop for ClpAPEng1 (left), ClpAPDis (middle) and ClpAPEng2 (right) showing extended and compact conformations, respectively, on the basis of distance measurements between loop residues 605–619 and 633–619 (red dots).
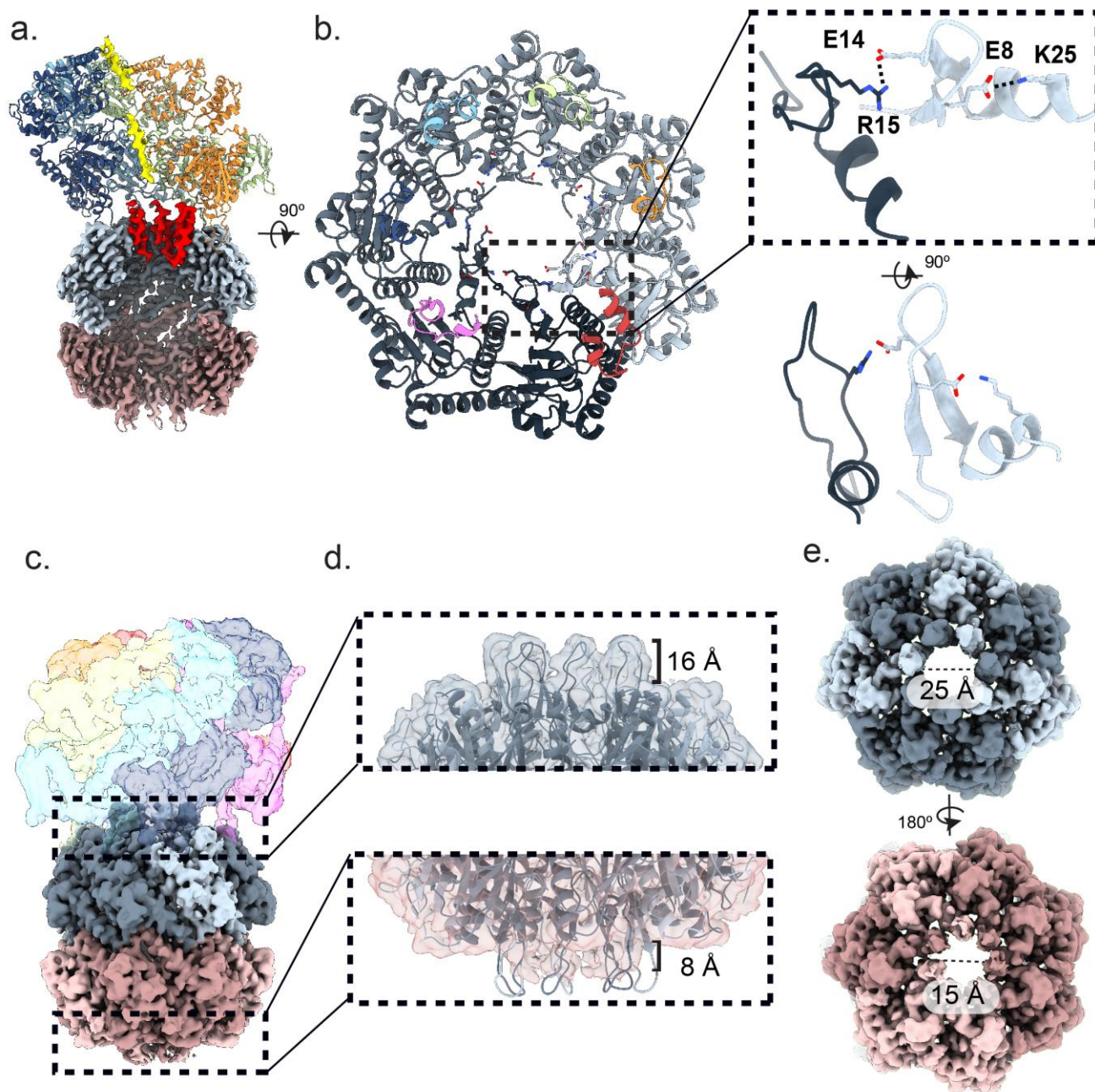
Figure 1.4 Structure of ClpP and NT gating loops. a, Channel view of ClpAP highlighting the ClpP NT gating loops (red) relative to substrate density (yellow). **b**, Top view (left) of ClpP NT loops with ClpA IGL loops (colored by protomer). To the right is an expanded view of an NT loop pair with cis (E8–K25) and trans (R15–E14) salt-bridge contacts. **c**, Side-view map of single-capped ClpAP complex. **d**, Expanded views of the ClpP pore for the ClpA-bound and unbound surfaces showing open- and closed-gate conformations, respectively. The open-gate conformation was modeled into both sites to show differences compared to the closed-gate density. **e**, Top views showing ClpP pore diameter for the open- (top) and closed- (bottom) gate conformations.

35
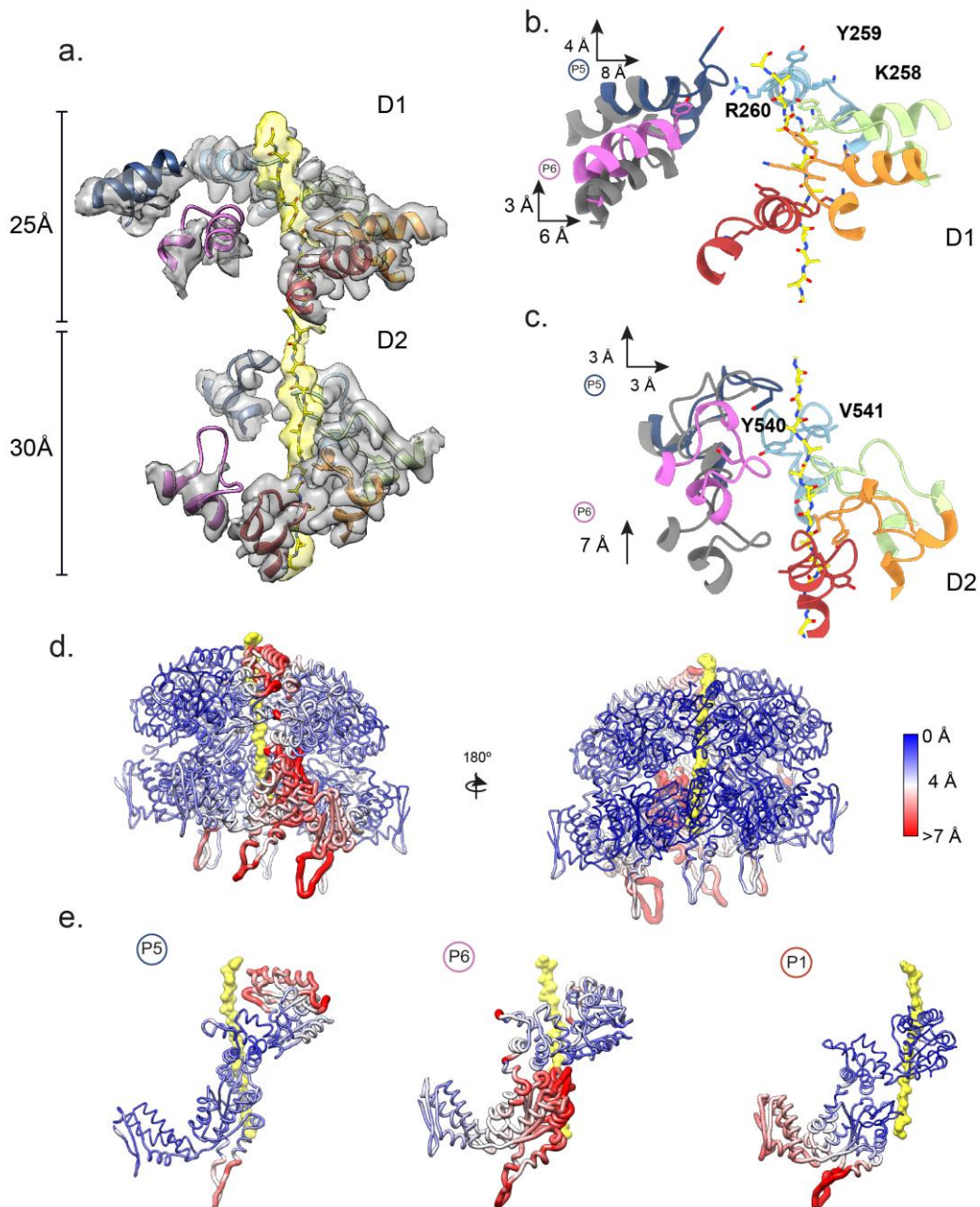
Figure 1.5 ClpA pore loop-substrate contacts and translocation states a, Segmented map and model of the substrate-bound P1–P6 pore loops, colored by protomer, with substrate (yellow) for ClpAPEng1. Distances shown indicate length of substrate interactions for the D1 and D2. b,c, Model of the D1 (**b**) and D2 (**c**) pore loops and substrate for ClpAPEng2 (colored by protomer) and overlayed with ClpAPEng1 (gray). Substratecontacting residues are indicated and shifts in the position of the pore loop protomers, P5 and P6, between states are shown. **d**, ClpAPEng2 model is displayed showing Cα r.m.s.d. between the three states, determined by alignment to protomer P3. Large changes (>7 Å) are indicated in red with wider tubes, intermediate changes (~6.0 Å) are colored in white and small/no changes are colored in blue. **e**, Individual seam protomers shown with Cα r.m.s.d. mapped as in d for P5 (left), P6 (middle) and P1 (right).

36
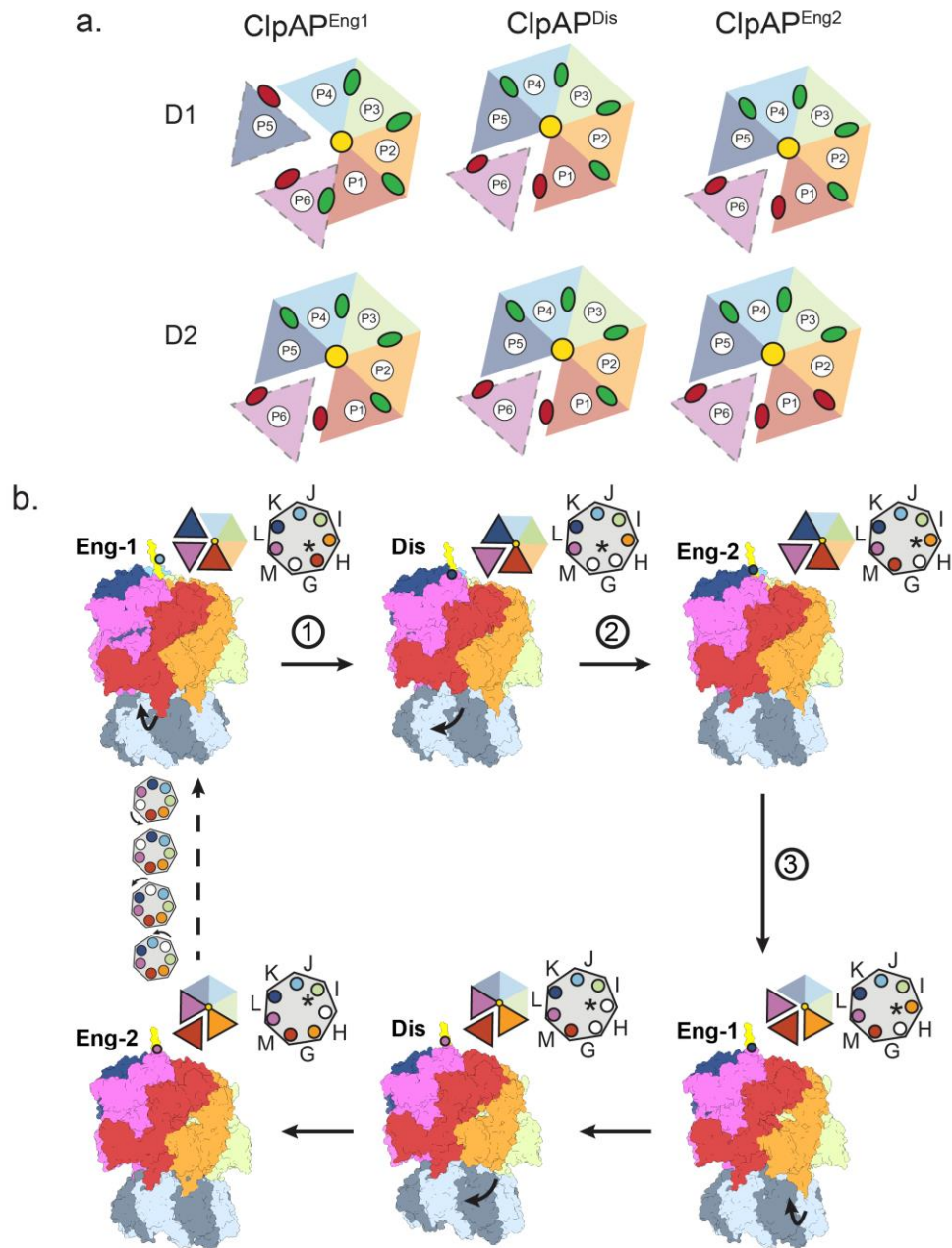
Figure 1.6 Nucleotide states and ClpA rotation model for processive unfolding and proteolysis a, Schematic showing nucleotide states and substrate contact for D1 and D2 of ClpAPEng1 (left), ClpAPDis (middle) and ClpAPEng2 (right), determined on the basis of difference maps (Extended Data Fig. 7). Protomer nucleotide states are denoted by colored circles (green for ATP and red for ADP). **b**, Model for ClpAP processive substrate translocation cycle. Two translocation steps are depicted and coupled to IGL loop disengagement (steps 1 and 4) and engagement to the next clockwise IGL pocket on ClpP (steps 2 and 5), indicated by arrows. Top view schematics show the rotary cycle of substrate binding by ClpA (left) and occupancy of ClpP IGL pockets (right). The protomer at the lowest substrate-contact site, which releases the IGL loop, is indicated by an asterisk (*).

Methods

Purification and analysis of ClpA, ClpP, and RepA(1-25)-GFP

ClpA and ClpP were purified as previously described[62,70]. RepA 1-25 protein was expressed with a C-terminal His6-tag construct from the pDS56/RBSII plasmid. Transformed BL21 cells were inoculated in LB media with 100 ug/mL Ampicillin and grown at 37°C to OD600nm = ~0.6–0.8. The cell culture was induced with 1 mM IPTG for ~4 h at 30°C. Cell pellet was resuspended in 40 mM HEPES pH 7.4, 2 mM β-mercaptoethanol, and 10% glycerol with protease inhibitors (EDTA-free) (Roche) and then lysed by sonication. Following centrifugation (16,000 x g, 20 min, 30°C), the supernatant was applied to a Ni-NTA column (GE Healthcare) followed by a gradient elution from 20 mM imidazole to 500 mM imidazole. Purity was verified by SDS-PAGE and fractions were combined and concentrated into a storage buffer (40 mM HEPES pH 7.4, 500 mM KCl, 20 mM MgCl2, 10% glycerol (v/v), and 2 mM β-mercaptoethanol).

The RepA(1-25)-GFP degradation assay (Supplementary Fig 1.S1 a-b) was performed in triplicate and consisted of 6 mM ClpA, 7 mM ClpP or ClpP-S98A, 1 mM RepA(1-25)-GFP and 2 mM nucleotide incubated in buffer at 20° for15 min containing 50 mM Tris-HCl pH 7.5, 150 mM KCl, 10 mM MgCl$_2$, and 1 mM DTT. For the assay with spiked nucleotide, 10 mM nucleotide was added after the initial incubation. Aliquots of the reaction were separated from the reaction at the specified time points and quenched in 2% SDS buffer, heated for 10 min and ran onto an acrylamide gel. The bands were visualized using silver staining (Sigma-Aldrich). Size exclusion chromatography (SEC) analysis and purification was performed by incubating 36 mM ClpA, 42 mM ClpP, 30 mM RepA(1-25)-GFP and 2 mM ATPgS in 50 mM Tris-HCl pH 7.5, 150 mM KCl, 10 mM MgCl$_2$, 1 mM DTT for 15 minutes at 20°. The complex incubation

reaction was then injected onto a Superose 6 Increase 3.2/300 column (GE Healthcare) and the eluted peaks were analyzed using SDS-PAGE.

Cryo-EM data collection and processing

The fraction corresponding to the largest molecular weight complex from SEC (Supplementary Fig. 1.S1 b) was isolated and incubated with 1 mM ATPgS. Before freezing, proper dilutions were made and 10 mM ATP was added to the dilution. After a 30 s. incubation, a 3.5 uL drop was applied to glow discharged holey carbon grid (R 1.2/1.3; Quantifoil), in which sample was then blotted for 2.5 s. at 4° and 100% humidity with Whatman No. 1 filter paper before being plunge frozen liquid ethane using a vitrobot (Thermo Fischer Scientific). The sample was then imaged on a Titan Krios TEM (Thermo Fischer Scientific) operated at 300 keV and equipped with a Gatan BioQuantum imaging energy filter using a 20eV zero loss energy slit (Gatan Inc). Movies were acquired in super-resolution mode on a K3 direct electron detector (Gatan Inc.) at a calibrated magnification of 58,600X corresponding to a pixel size of 0.4265 Å/pixel. A defocus range of .8 to 1.2 μm was used with a total exposure time of 2 seconds fractionated into 0.2s subframes for a total dose of 68 $e^-/Å^2$ at a dose rate of 25 e^-/pixel/s. Movies were subsequently corrected for drift using MotionCor2 (10.1038/nmeth.4193) and were Fourier-cropped by a factor of 2 to a final pixel size of 0.853 Å/pixel.

A total of ~18,000 micrographs were collected over two different datasets. The two datasets were processed separately and then were combined at the end. All the data-processing was performed in cryosparc2[71]. For particle picking, templates were generated from 100 particles, in which only side-views were selected. After inspecting the particles picked, approximately 1.6 million particles were extracted. Two rounds of 2D classification were performed to remove contamination and junk particles, which amounted to ~54% of the dataset.

A five-class ab-initio reconstruction was performed from the particle set and was used for initial classification steps.

To identify different conformations, heterogenous refinement was performed with 4 different classes (Supplementary Fig. 1.S1 f). Following this first round, maps showing high resolution features, which accounted for ~54% of the 739,000 particles going into 3D, were kept and grouped together. Another round of heterogenous refinement with 5 different classes was then performed. Following this second round, two unique states, $ClpAP^{Eng1}$ (24%, ~176,000 particles) and $ClpAP^{Dis}$ (24%, ~176,000 particles), were identified. The $ClpAP^{Dis}$ particles underwent another 5 class heterogenous refinement to further identify any more conformations. Following this third round, two unique states, $ClpAP^{Dis}$ (8%, 58,000 particles) and $ClpAP^{Eng2}$ (5%, 40,000 particles), were identified. Particles associated with each unique class were combined and homogenous refinement was performed separately on each state. To better improve the resolution of the mobile protomers following Non-Uniform refinement, the particles from each state underwent particle subtraction. Particle subtraction was performed in which the bottom half of ClpP was subtracted. A local-refinement was then performed, in which the fulcrum position was set to the center of ClpA. The same procedure was completed on all the states.

The final resolution of $ClpAP^{Eng1}$ was 2.8Å, $ClpAP^{Dis}$ was 3.2Å, and $ClpAP^{Eng2}$ was 3.4Å (Supplementary Fig. 1.S6 b). After completing local CTF refinement on of the final refinement runs the resolutions were improved to 2.7Å for $ClpAP^{Eng1}$, 3.0 Å for $ClpAP^{Dis}$ and 3.2 Å for $ClpAP^{Eng2}$ (Supplementary Fig. 1.S1 e).

Molecular modeling

An initial model for ClpA was obtain by using a ClpB structure (pdb 5ofo)[35] and generated in SWISS-MODEL[72] and the initial model for ClpP was taken directly from a ClpP crystal structure (pdb 1yg6)[48] previously solved. Both initial models were docked into the EM maps using the UCSF chimera's function *fit in map*[73]. Initial refinement was performed using Phenix[74] with 1 round of simulated annealing and morphing and 5 rounds of real-space refinement that included minimization_global, rigid_body, adp, local_grid_search, secondary structural restraints and non-crystallographic symmetry (NCS) restraints. The resulting model then underwent real space refinement in Coot[75]. Nucleotides were added in manually using Coot and real space refinement using cif files generated for ADP and ATPgS in Phenix eLBOW[76].

Density for the ClpA focus refinement was higher quality than the full map, therefore was used to model individual protomers using Rosetta Comparitive Modeling (RosettaCM)[77,78]. The structures for ClpA (pdb 1r6b)[79], Hsp104 (pdb 5d4w and 5vjh)[36], ClpB BAP form (pdb 5og1)[35] and PTEX (pdb 6e10)[42] were determined as homology models with HHpred[80] and used to constrain model refinement in Rosetta CM with *template_weight=0* and the initial model with *template_weight=1*. The lowest energy models were examined by eye to ensure the model fit into the density, the protomer was placed into the context of the whole structure and the Rosetta Relax protocol was run on the full complex.

Rosetta Enumerative Sampling (Rosetta ES) was used to de novo build in the IGL loops and NT loops for each protomer[81]. The ClpA residues 612 to 628 were deleted from each protomer and Rosetta ES was run to rebuild the loops with a beamwidth of 32. The resulting model with rebuilt IGL loops was added into the full model and the Rosetta Relax protocol was

run. Residues 16 to 32 from ClpP were deleted from each protomer and the same RosettaES

parameters were used to build in the NT loops, followed by the Rosetta Relax protocol.

References

1. Shorter, J. & Southworth, D.R. Spiraling in Control: Structures and Mechanisms of the Hsp104 Disaggregase. *Cold Spring Harb Perspect Biol* (2019).

2. Duran, E.C., Weaver, C.L. & Lucius, A.L. Comparative Analysis of the Structure and Function of AAA+ Motors ClpA, ClpB, and Hsp104: Common Threads and Disparate Functions. *Front Mol Biosci* **4**, 54 (2017).

3. Olivares, A.O., Baker, T.A. & Sauer, R.T. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat Rev Microbiol* **14**, 33-44 (2016).

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

5. Sauer, R.T. et al. Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell* **119**, 9-18 (2004).

6. Kessel, M. et al. Homology in structural organization between E. coli ClpAP protease and the eukaryotic 26 S proteasome. *J Mol Biol* **250**, 587-94 (1995).

7. Sousa, M.C. et al. Crystal and solution structures of an HslUV protease-chaperone complex. *Cell* **103**, 633-43 (2000).

8. Baker, T.A. & Sauer, R.T. ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim Biophys Acta* **1823**, 15-28 (2012).

9. Reid, B.G., Fenton, W.A., Horwich, A.L. & Weber-Ban, E.U. ClpA mediates directional translocation of substrate proteins into the ClpP protease. *Proc Natl Acad Sci U S A* **98**, 3768-72 (2001).

10. Hoskins, J.R., Pak, M., Maurizi, M.R. & Wickner, S. The role of the ClpA chaperone in proteolysis by ClpAP. *Proc Natl Acad Sci U S A* **95**, 12135-40 (1998).

11. Weber-Ban, E.U., Reid, B.G., Miranker, A.D. & Horwich, A.L. Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. *Nature* **401**, 90-3 (1999).

12. Ishikawa, T. et al. Translocation pathway of protein substrates in ClpAP protease. *Proc Natl Acad Sci U S A* **98**, 4328-33 (2001).

13. Wang, J., Hartling, J.A. & Flanagan, J.M. The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**, 447-56 (1997).

14. Yu, A.Y. & Houry, W.A. ClpP: a distinctive family of cylindrical energy-dependent serine proteases. *FEBS Lett* **581**, 3749-57 (2007).

15. Grimaud, R., Kessel, M., Beuron, F., Steven, A.C. & Maurizi, M.R. Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP. *J Biol Chem* **273**, 12476-81 (1998).

16. Ortega, J., Lee, H.S., Maurizi, M.R. & Steven, A.C. ClpA and ClpX ATPases bind simultaneously to opposite ends of ClpP peptidase to form active hybrid complexes. *J Struct Biol* **146**, 217-26 (2004).

17. Ortega, J., Singh, S.K., Ishikawa, T., Maurizi, M.R. & Steven, A.C. Visualization of substrate binding and translocation by the ATP-dependent protease, ClpXP. *Mol Cell* **6**, 1515-21 (2000).

18. Levchenko, I., Seidel, M., Sauer, R.T. & Baker, T.A. A specificity-enhancing factor for the ClpXP degradation machine. *Science* **289**, 2354-6 (2000).

19. Bolon, D.N., Grant, R.A., Baker, T.A. & Sauer, R.T. Nucleotide-dependent substrate handoff from the SspB adaptor to the AAA+ ClpXP protease. *Mol Cell* **16**, 343-50 (2004).

20.   Erbse, A. et al. ClpS is an essential component of the N-end rule pathway in Escherichia coli. *Nature* **439**, 753-6 (2006).

21.   Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T. & Baker, T.A. The molecular basis of N-end rule recognition. *Mol Cell* **32**, 406-14 (2008).

22.   Keiler, K.C., Waller, P.R. & Sauer, R.T. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990-3 (1996).

23.   Gottesman, S., Roche, E., Zhou, Y. & Sauer, R.T. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev* **12**, 1338-47 (1998).

24.   Mogk, A., Schmidt, R. & Bukau, B. The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol* **17**, 165-72 (2007).

25.   Wickner, S. et al. A molecular chaperone, ClpA, functions like DnaK and DnaJ. *Proc Natl Acad Sci U S A* **91**, 12218-22 (1994).

26.   Kim, Y.I. et al. Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nat Struct Biol* **8**, 230-3 (2001).

27.   Effantin, G., Maurizi, M.R. & Steven, A.C. Binding of the ClpA unfoldase opens the axial gate of ClpP peptidase. *J Biol Chem* **285**, 14834-40 (2010).

28.   Bewley, M.C., Graziano, V., Griffin, K. & Flanagan, J.M. Turned on for degradation: ATPase-independent degradation by ClpP. *J Struct Biol* **165**, 118-25 (2009).

29.   Jennings, L.D., Bohon, J., Chance, M.R. & Licht, S. The ClpP N-terminus coordinates substrate access with protease active site reactivity. *Biochemistry* **47**, 11031-40 (2008).

30.   Brotz-Oesterhelt, H. et al. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* **11**, 1082-7 (2005).

31. Lee, B.G. et al. Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat Struct Mol Biol* **17**, 471-8 (2010).

32. Gersch, M. et al. AAA+ chaperones and acyldepsipeptides activate the ClpP protease via conformational control. *Nat Commun* **6**, 6320 (2015).

33. Li, D.H. et al. Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. *Chem Biol* **17**, 959-69 (2010).

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

35. Deville, C. et al. Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase. *Sci Adv* **3**, e1701726 (2017).

36. Gates, S.N. et al. Ratchet-like polypeptide translocation mechanism of the AAA+ disaggregase Hsp104. *Science* **357**, 273-279 (2017).

37. Rizo, A.N. et al. Structural basis for substrate gripping and translocation by the ClpB AAA+ disaggregase. *Nat Commun* **10**, 2393 (2019).

38. Yu, H. et al. ATP hydrolysis-coupled peptide translocation mechanism of Mycobacterium tuberculosis ClpB. *Proc Natl Acad Sci U S A* **115**, E9560-E9569 (2018).

39. Gates, S.N. & Martin, A. Stairway to translocation: AAA+ motor structures reveal the mechanisms of ATP-dependent substrate translocation. *Protein Sci* **29**, 407-419 (2020).

40. Han, H., Monroe, N., Sundquist, W.I., Shen, P.S. & Hill, C.P. The AAA ATPase Vps4 binds ESCRT-III substrates through a repeating array of dipeptide-binding pockets. *Elife* **6**(2017).

41.   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**(2018).

42.   Ho, C.M. et al. Malaria parasite translocon structure and mechanism of effector export. *Nature* **561**, 70-75 (2018).

43.   Puchades, C. et al. Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. *Science* **358**(2017).

44.   Hoskins, J.R., Kim, S.Y. & Wickner, S. Substrate recognition by the ClpA chaperone component of ClpAP protease. *J Biol Chem* **275**, 35361-7 (2000).

45.   Hoskins, J.R. & Wickner, S. Two peptide sequences can function cooperatively to facilitate binding and unfolding by ClpA and degradation by ClpAP. *Proc Natl Acad Sci U S A* **103**, 909-14 (2006).

46.   Kress, W., Mutschler, H. & Weber-Ban, E. Assembly pathway of an AAA+ protein: tracking ClpA and ClpAP complex formation in real time. *Biochemistry* **46**, 6183-93 (2007).

47.   Martin, A., Baker, T.A. & Sauer, R.T. Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. *Mol Cell* **27**, 41-52 (2007).

48.   Bewley, M.C., Graziano, V., Griffin, K. & Flanagan, J.M. The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *J Struct Biol* **153**, 113-28 (2006).

49.   Hinnerwisch, J., Fenton, W.A., Furtak, K.J., Farr, G.W. & Horwich, A.L. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. *Cell* **121**, 1029-41 (2005).

50. Martin, A., Baker, T.A. & Sauer, R.T. Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol Cell* **29**, 441-50 (2008).

51. Kress, W., Mutschler, H. & Weber-Ban, E. Both ATPase domains of ClpA are critical for processing of stable protein structures. *J Biol Chem* **284**, 31441-52 (2009).

52. Mogk, A. et al. Roles of individual domains and conserved motifs of the AAA+ chaperone ClpB in oligomerization, ATP hydrolysis, and chaperone activity. *J Biol Chem* **278**, 17615-24 (2003).

53. Fernandez-Higuero, J.A. et al. Allosteric communication between the nucleotide binding domains of caseinolytic peptidase B. *J Biol Chem* **286**, 25547-55 (2011).

54. DeSantis, M.E. et al. Operational plasticity enables hsp104 to disaggregate diverse amyloid and nonamyloid clients. *Cell* **151**, 778-93 (2012).

55. Lee, C., Schwartz, M.P., Prakash, S., Iwakura, M. & Matouschek, A. ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol Cell* **7**, 627-37 (2001).

56. Olivares, A.O., Kotamarthi, H.C., Stein, B.J., Sauer, R.T. & Baker, T.A. Effect of directional pulling on mechanical protein degradation by ATP-dependent proteolytic machines. *Proc Natl Acad Sci U S A* (2017).

57. Lyubimov, A.Y., Strycharska, M. & Berger, J.M. The nuts and bolts of ring-translocase structure and mechanism. *Curr Opin Struct Biol* **21**, 240-8 (2011).

58. Aubin-Tam, M.E., Olivares, A.O., Sauer, R.T., Baker, T.A. & Lang, M.J. Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. *Cell* **145**, 257-67 (2011).

59.	Cordova, J.C. et al. Stochastic but highly coordinated protein unfolding and translocation by the ClpXP proteolytic machine. *Cell* **158**, 647-58 (2014).

60.	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**(2020).

61.	Olivares, A.O., Nager, A.R., Iosefson, O., Sauer, R.T. & Baker, T.A. Mechanochemical basis of protein degradation by a double-ring AAA+ machine. *Nat Struct Mol Biol* **21**, 871-5 (2014).

62.	Miller, J.M., Lin, J., Li, T. & Lucius, A.L. E. coli ClpA catalyzed polypeptide translocation is allosterically controlled by the protease ClpP. *J Mol Biol* **425**, 2795-812 (2013).

63.	Avellaneda, M.J. et al. Processive extrusion of polypeptide loops by a Hsp100 disaggregase. *Nature* (2020).

64.	Rajendar, B. & Lucius, A.L. Molecular mechanism of polypeptide translocation catalyzed by the Escherichia coli ClpA protein translocase. *J Mol Biol* **399**, 665-79 (2010).

65.	Gatsogiannis, C., Balogh, D., Merino, F., Sieber, S.A. & Raunser, S. Cryo-EM structure of the ClpXP protein degradation machinery. *Nat Struct Mol Biol* **26**, 946-954 (2019).

66.	Rabl, J. et al. Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases. *Mol Cell* **30**, 360-8 (2008).

67.	Smith, D.M. et al. Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry. *Mol Cell* **27**, 731-44 (2007).

68.   Dong, Y. et al. Cryo-EM structures and dynamics of substrate-engaged human 26S proteasome. *Nature* **565**, 49-55 (2019).

69.   Majumder, P. et al. Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle. *Proc Natl Acad Sci U S A* **116**, 534-539 (2019).

70.   Veronese, P.K., Stafford, R.P. & Lucius, A.L. The Escherichia coli ClpA molecular chaperone self-assembles into tetramers. *Biochemistry* **48**, 9221-33 (2009).

71.   Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).

72.   Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303 (2018).

73.   Pettersen, E.F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).

74.   Afonine, P.V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544 (2018).

75.   Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).

76.   Moriarty, N.W., Grosse-Kunstleve, R.W. & Adams, P.D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* **65**, 1074-80 (2009).

77.   DiMaio, F. et al. Atomic-accuracy models from 4.5-A cryo-electron microscopy data with density-guided iterative local refinement. *Nat Methods* **12**, 361-365 (2015).

78.   Song, Y. et al. High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735-42 (2013).

79. Xia, D., Esser, L., Singh, S.K., Guo, F. & Maurizi, M.R. Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone. *J Struct Biol* **146**, 166-79 (2004).

80. Zimmermann, L. et al. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* **430**, 2237-2243 (2018).

81. Frenz, B., Walls, A.C., Egelman, E.H., Veesler, D. & DiMaio, F. RosettaES: a sampling strategy enabling automated interpretation of difficult cryo-EM maps. *Nat Methods* **14**, 797-800 (2017).

Supplementary Information

Table 1.1: Cryo-EM data collection, refinement and validation statistics of ClpAP Engaged 1, Disengaged, Engaged 2 and ClpAP-ATPgammaS

|	ClpAP^{Eng1} (EMD-21519, PDB 6W1Z)	ClpAP^{Dis} (EMD-21520, PDB 6W20)	ClpAP^{Eng2} (EMD-21521, PDB 6W21)	Focus ClpAP^{Eng1} (EMD-21522, PDB 6W22)	Focus ClpAP^{Dis} (EMD-21523, PDB 6W23)	Focus ClpAP^{Eng2} (EMD-21524, PDB 6W24)	ClpAP–ATPgS disengaged (EMD-20845, PDB 6UQE)	ClpAP–ATPgS engaged (EMD-20851, PDB 6UQO)
Data collection and processing								
Microscope and camera	Titan Krios, K3	Titan Krios, K3	Titan Krios, K3	Titan Krios, K3	Titan Krios, K3	Titan Krios, K3	Titan Krios, K3	Titan Krios, K3
Magnification	58,600	58,600	58,600	58,600	58,600	58,600	58,600	58,600
Voltage (kV)	300	300	300	300	300	300	300	300
Data acquisition software	Serial EM	Serial EM	Serial EM	Serial EM	Serial EM	Serial EM	Serial EM	Serial EM
Exposure navigation	Image shift	Image shift	Image shift	Image shift	Image shift	Image shift	Image shift	Image shift
Electron exposure (e$^-$/Å^2)	68	68	68	68	68	68	68	68
Defocus range (μm)	0.8–1.2	0.8–1.2	0.8–1.2	0.8–1.2	0.8–1.2	0.8–1.2	1.2–2	1.2–2
Pixel size (Å)	0.853	0.853	0.853	0.853	0.853	0.853	0.853	0.853
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1	C1
Initial particle images (no.)	739,000	739,000	739,000	739,000	739,000	739,000	1,800,000	1,800,000
Final particle images (no.)	176,232	57,848	39,177	176,232	57,848	39,177	314,000	169,000
Map resolution (Å)	2.7	3.0	3.2	3.0	3.1	3.4	3.0	3.1
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2–10	2–10	2–10	2.5–10	2.5–10	2.5–10	2–10	2–10
Refinement								
Model resolution (Å)	3.0	3.1	3.2	2.8	3.0	3.3	3.1	3.1
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Map sharpening B factor (Å^2)	−73.0	−68.5	−60.8	−80.3	−65.6	−55.8	−112.9	−103.2
Model composition								
Nonhydrogen atoms	48,556	48,252	48,346	27,542	27,238	27,332	48,402	48,522
Protein residues	6,180	6,134	6,147	3,492	3,446	3,459	6,161	6,174
Ligands	12	12	12	12	12	12	12	12
B factors (Å^2)								
Protein	33.4	125.4	177.4	0.5	153.0	210.8	33.7	33.4
Ligand	20.0	151.4	210.6	20.0	151.4	210.6	44.1	20.0
R.m.s. deviations								
Bond lengths (Å)	0.03	0.01	0.011	0.026	0.014	0.013	0.025	0.026
Bond angles (°)	1.76	0.60	0.58	1.81	0.95	0.61	1.82	1.84
Validation								
MolProbity score	1.72	1.96	1.78	1.62	2.26	1.9	1.37	1.25
Clashscore	9.43	13.39	13.92	8.42	15.9	14.4	3.52	2.58
Poor rotamers (%)	0.90	0.20	0.10	0.55	0.03	0.10	0.37	0.08
Ramachandran plot								
Favored (%)	96.58	96.81	96.48	97.07	95.59	97.31	96.55	96.72
Allowed (%)	3.11	3.16	3.20	2.47	4.35	2.51	2.83	2.84
Disallowed (%)	0.31	0.03	0.32	0.46	0.06	0.18	0.62	0.44
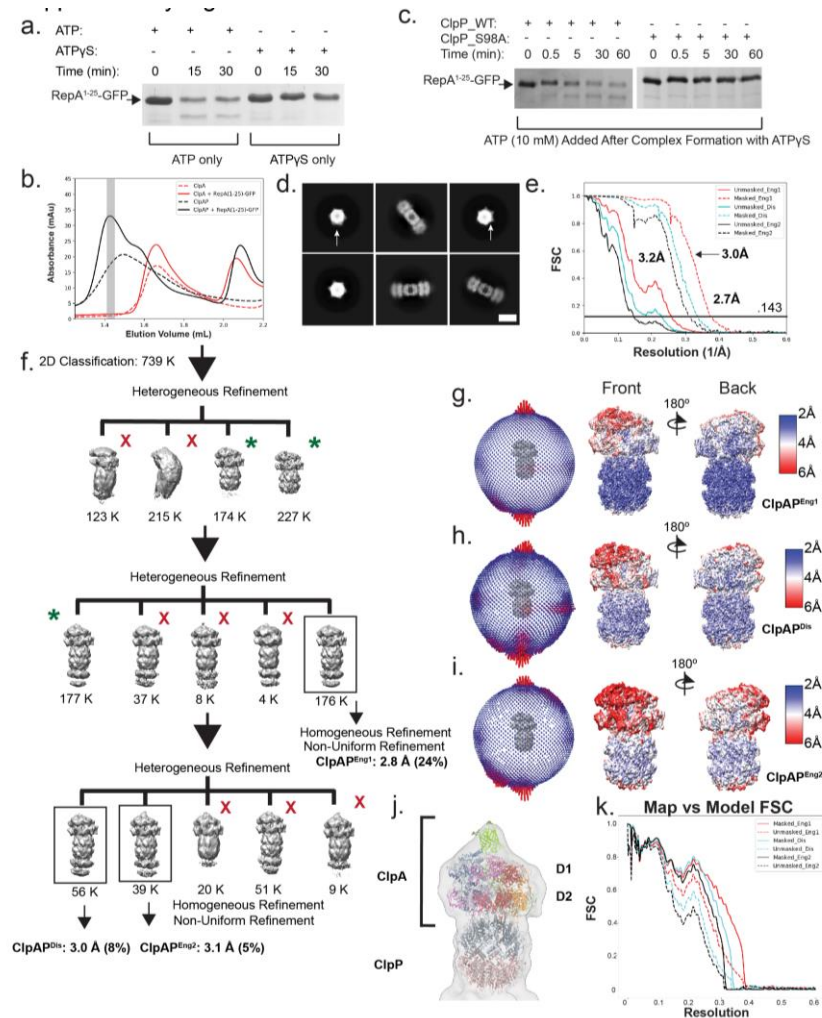
Supplemental Figure 1.S1 ClpAP complex formation with RepA(1-25)-GFP and cryoEM data analysis
a RepA1–25-GFP degradation assay in the presence of either ATPγS or ATP along with ClpA and ClpP. The assay was performed at 20°. Arrow represents RepA degradation product. **b** Size exclusion chromatography (SEC) trace of the components and formed ClpAP complex following incubation with RepA1–25-GFP and ATPyS. The 280 absorbance traces are shown for ClpA alone (red, dashed), ClpA with RepA1–25-GFP (red, solid), ClpAP alone (black, dashed) and ClpAP with RepA1–25-GFP (black, solid). **c** RepA1–25-GFP degradation assay in the presence of ATPγS with both ClpP WT and ClpP_S98A. ATP was spiked into the reaction at 10 mM after the initial complex formation for 15 min was completed with ATPγS. The zero-time point is before spiking ATP into the reaction. The assay was performed at 20°. **d** Reference-free 2D class averages of ClpAP bound to RepA1–25-GFP. The scale bar equals 125 Å. **e** Gold standard FSC-curves for the final refinement of ClpAPEng1(red), ClpAPDis(cyan), ClpAPEng2(black) of the ClpAP-RepA(1-25)-GFP complex. **f** 3D classification scheme used to identify the two different states in the ClpAP-RepA1–25-GFP dataset. Green asterisk represents the classes in which the particles were pooled together for further classification and refinement. The local resolution map of ClpAPEng-1 (**g**), ClpAPDis (**h**) and ClpAPEng-2 (**i**). **j** Low-pass filtered map showing globular density docked with GFP (PDB 1GFL) and additional N-terminal ClpA density (NTD). **k** Map vs. Model FSC of ClpAPEng1(red), ClpAPDis(cyan), ClpAPEng2 (black) of the ClpAP-RepA(1-25)-GFP complex following atomic modeling in Rosetta. Uncropped gel images are available as source data online.
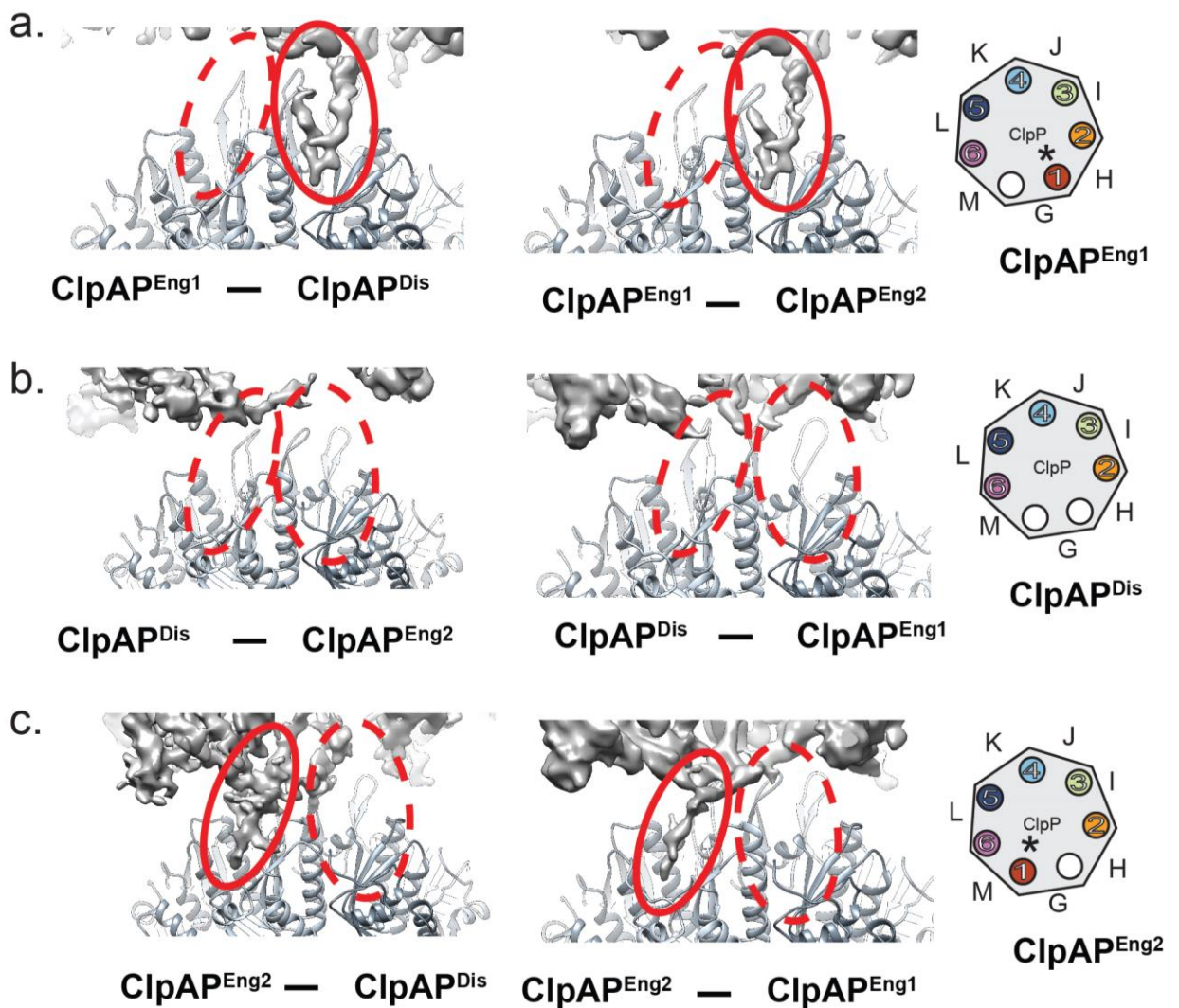
Supplemental Figure 1.S2 Difference maps of the ClpAP interface
Difference maps of the cryo-EM maps of **a** ClpAPEng1 vs. ClpAPDis and ClpAPEng-2, **b** ClpAPDis vs. ClpAPEng-1 and ClpAPEng-2, **c** ClpAPEng-2 vs. ClpAPDis and ClpAPEng-1. The IGL pockets are encompassed by red circle, open pocket (dashed) and occupied pocket (solid). Schematic (right) shows occupancy of the ClpA IGL-loops (circles, colored and numbered by protomer) around the ClpA hexamer, with the empty IGL pockets (white circles) and ClpA protomers indicated (letters) for the different states. Asterisk represents the IGL-loop that is engaging in that state.
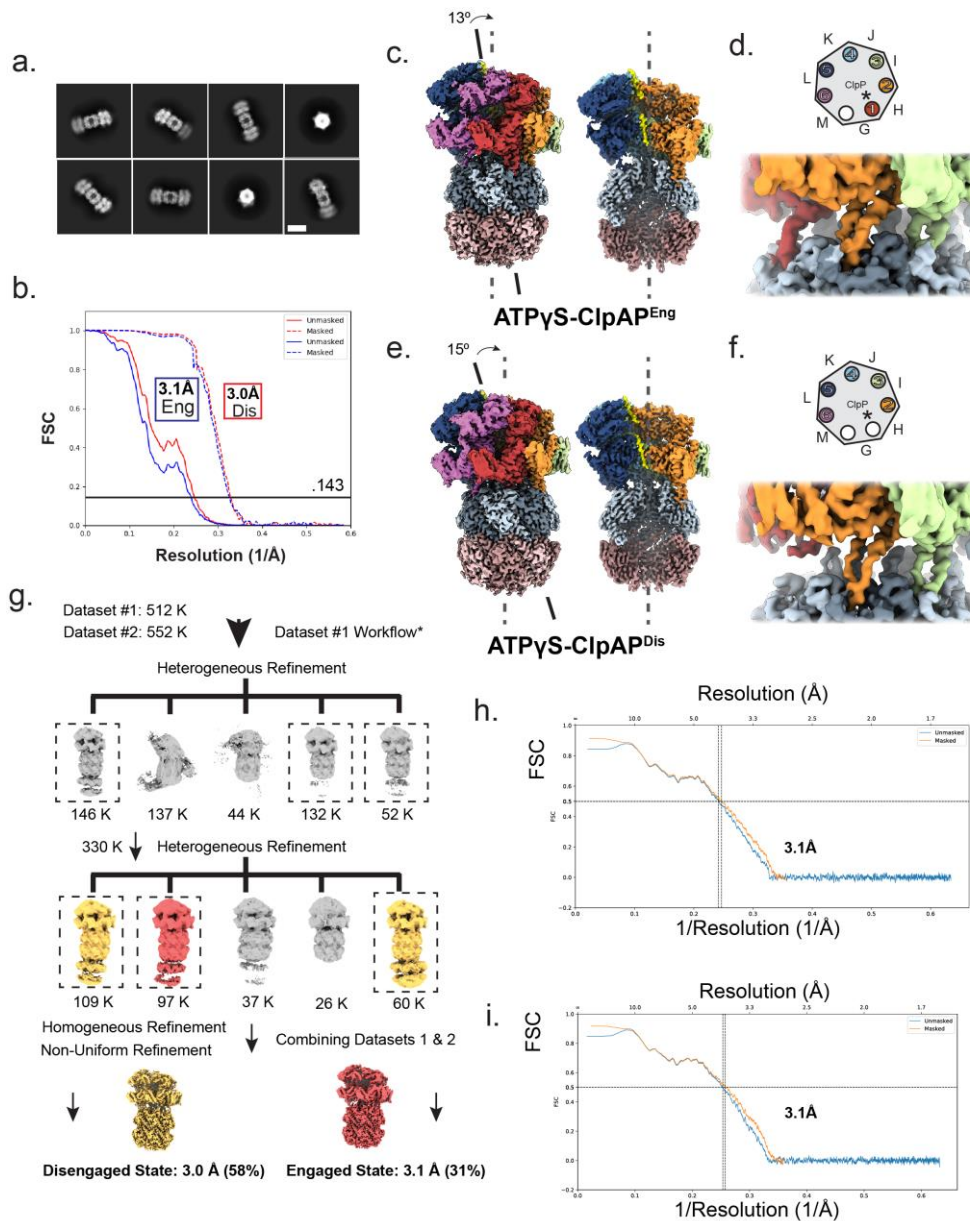
Supplementary Figure 1.S3 ATPgammaS-ClpAP cryoEM data analysis. a Reference-free 2D class averages of ClpAP-γS bound to RepA1–25-GFP. The scale bar equals 125 Å. **b** Gold standard FSC-curves for the final refinement of ATPγS-ClpAPEng (blue) and ATPγS-ClpAPDis (red) of the ClpAP-RepA(1-25)-GFP complex. ATPγS-ClpAPEng1 (**c**) and ATPγS-ClpAPDis (**d**) cryo-EM maps showing degree offset (arrow) of the ClpA channel axis (solid line) and substrate position (yellow density) compared to the ClpP pore and proteolytic chamber (dashed line). Schematic (below,left) shows occupancy of the ClpA IGL-loops (circles, colored and numbered by protomer) around the ClpA hexamer, with the empty IGL pockets (white circles) and ClpA protomers indicated (letters) for the different states. **e** 3D classification scheme used to identify the two different states in the ATPγS-ClpAP-RepA1–25-GFP dataset. Dotted boxes represent the classes in which the particles were pooled together for further classification and refinement. The maps for ClpAPEng (red) and ClpAPDis (yellow) are colored accordingly. Map vs. Model FSC of ATPγS-ClpAPEng(**f**) and ATPγS-ClpAPDis(**g**) following atomic modeling in Rosetta.
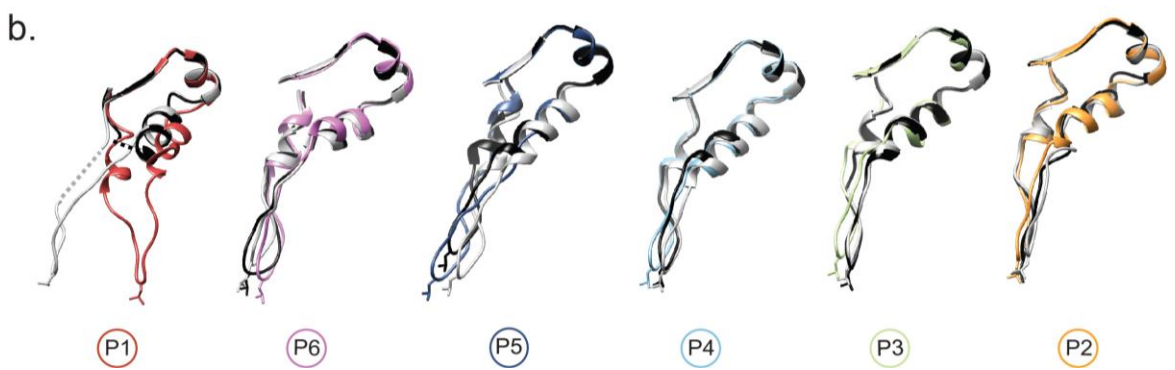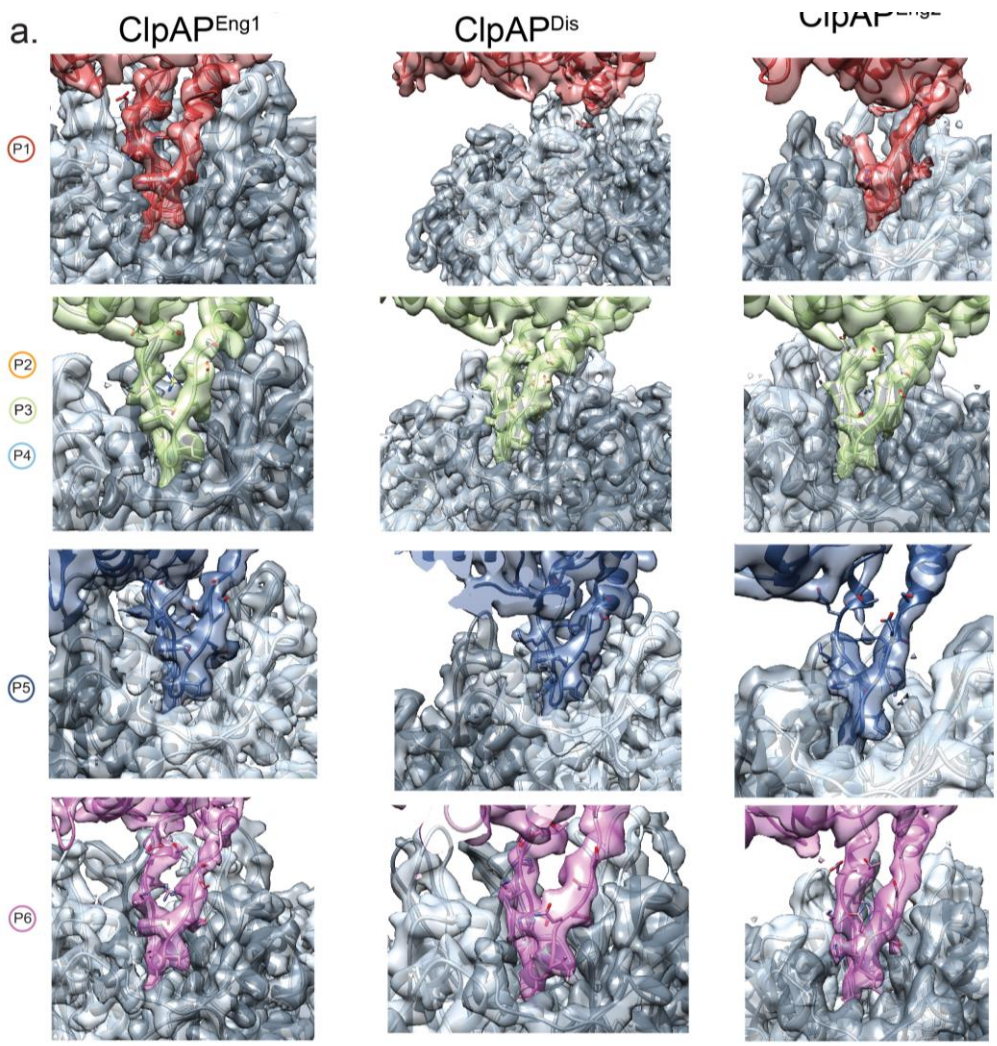
Supplementary Figure 1.S4 Comparison of IGL loops between the different states a EM map and model of the IGL-loop in the hydrophobic pocket of P1 (top), P2-P4 (middle, top), P5 (middle, bottom) and P6 (bottom) for ClpAPEng1(left), ClpAPDis (middle) and ClpAPEng2 (right). **b** Overlay of IGL-loops of ClpAPEng1 (colored by protomer) vs. ClpAPDis (black) vs. ClpAPEng2 (grey) laid out after alignment to the residues (638-649) above the IGL-loop. The dotted loop in P1 represents the missing loop in ClpAPDis and ClpAPEng-2.
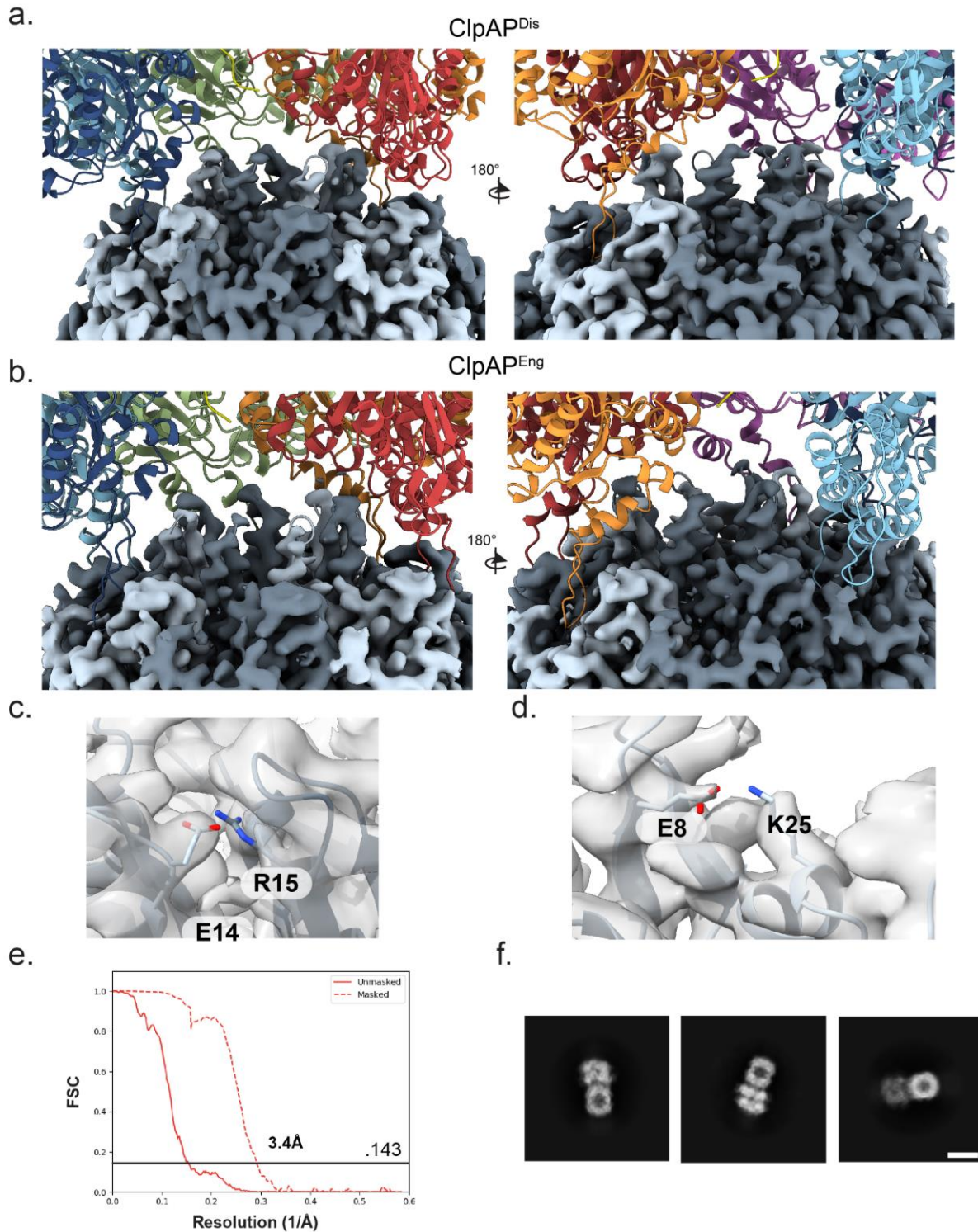
Supplementary Figure 1.S5 Single capped ClpAP structure and ClpP N-terminal loop interactions a Map of the ClpP N-terminal gating loops and the model for ClpA with substrate for ClpAPDis (**b**) ClpAPEng. Map and model view of ClpP residues E14 and R15 (**c**) and E8 and K25 (**d**). e Gold standard FSC curve and (**f**) 2D reference-free class averages of the single capped ClpAP structure.
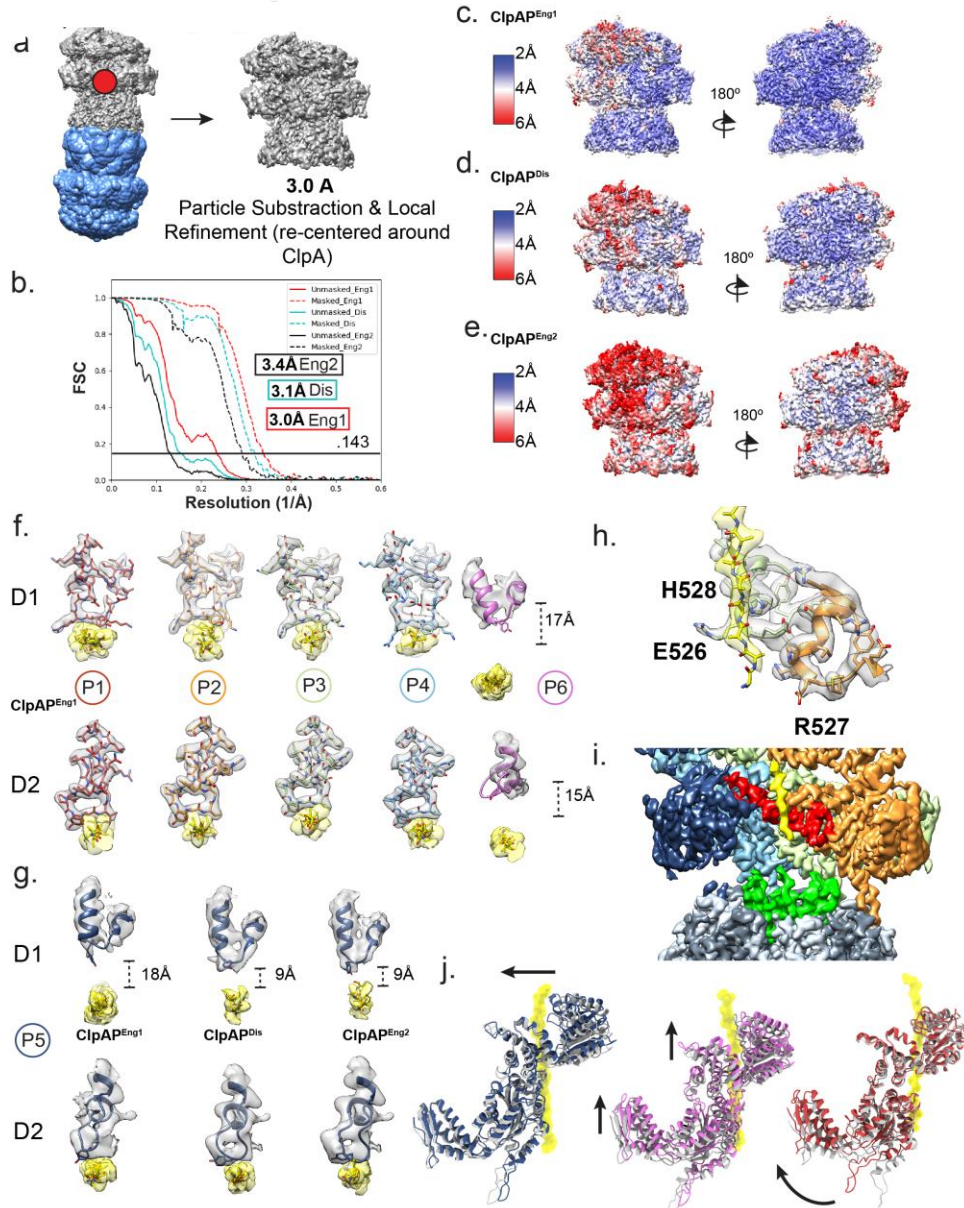
Supplementary Figure 1.S6 Particle substraction and focus refinement of ClpAP Engaged1, Engaged2 and Disengaged a EM map with mask (grey) used for particle subtraction of ClpA. Red dot represents the point in which particles were shifted to. **b** Gold standard FSC curve of both focus maps for ClpAPEng1 (red), ClpAPDis (cyan), and ClpAPEng2 (black). The local resolution map of ClpAPEng1 (**c**), ClpAPDis (**d**) and ClpAPEng2 (**e**). **f** EM map and model of each Tyr-containing pore loop in ClpAPEng1 for both D1 (top) and D2 (bottom), the substrate channel density is colored yellow. **g** EM map and model of each Tyr-containing pore loop in P5 for ClpAPEng1 (left), ClpAPDis (middle), and ClpAPEng2 (right) for both D1 (top) and D2 (bottom), the substrate channel density is colored yellow. The distance between the Tyr and the substrate is represented by dotted line. **h** EM map and model of ClpAPEng1 (colored by protomer) with the D2 secondary pore loops residues interacting with substrate. **i** ClpAPEng1 EM map colored by protomer with D2 secondary pore loops (red) and ClpP NTD-loops (green). **j** Overlay of the seam protomers P5 (left), P1 (middle), and P6 (right) for ClpAPEng1 (grey) and ClpAPEng2 (colored) showing conformational shifts (arrows) supporting translocation step.

58
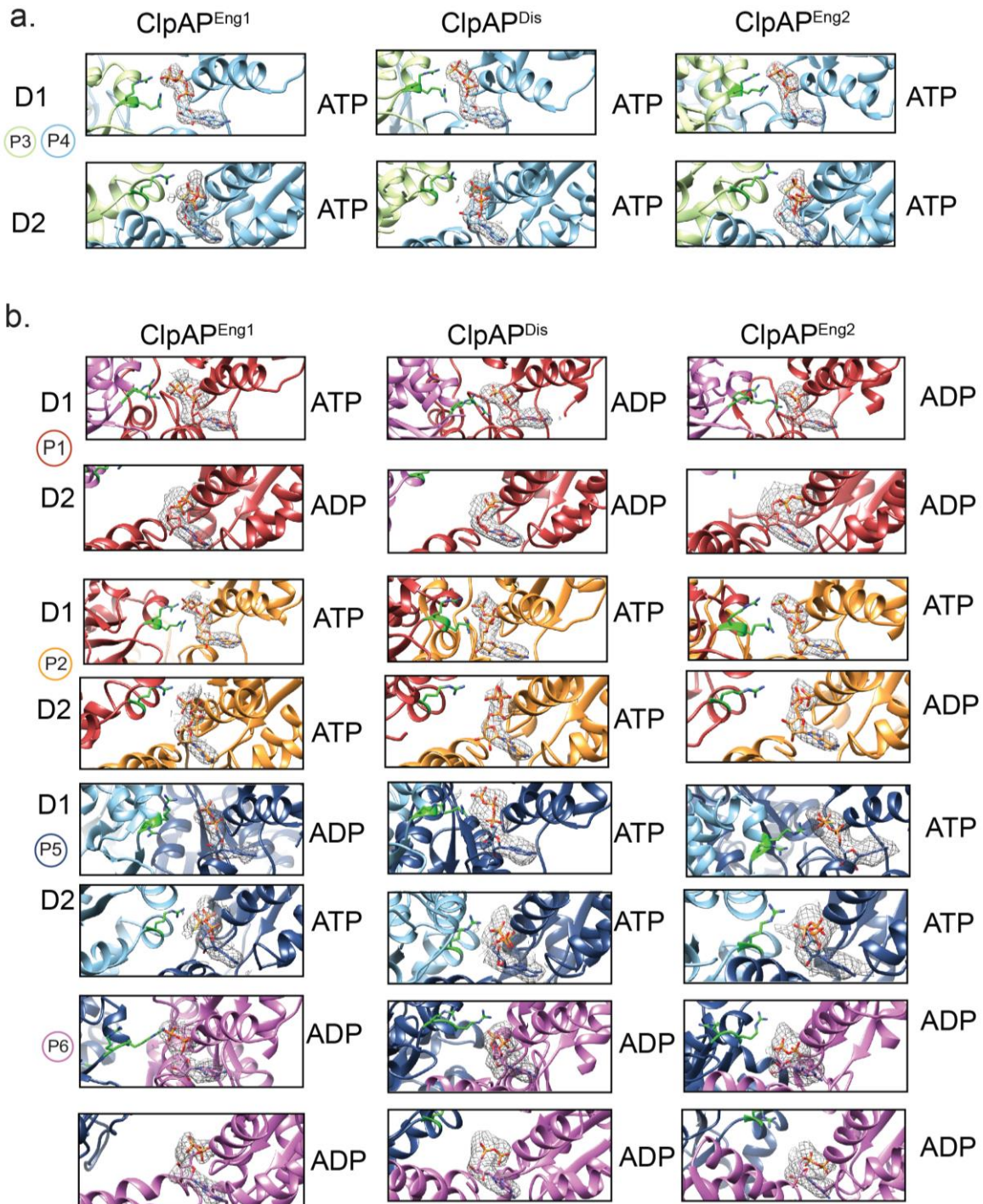
Supplementary Figure 1.S7 Nucleotide State of ClpAP Engaged1, Engaged2, and Disengaged a Difference map density for P4 D1 and D2 ATP with Arg finger residues displayed in green. There are no differences between P3 and P4, therefore P3 ATP density is not shown. **b** Difference map density for P1, P2, P5 and P6 for both D1 and D2 and Arg finger residues colored green.

Acknowledgments

Chapter 2

Structural determination of ClpAP bound to the adaptor protein ClpS and an N-degron substrate

Kyle E. Lopez[1,2], Eric Tse[2], Zachary G. Cuny[3], Aaron L. Lucius[3], and Daniel R. Soutworth[2]

Affiliations:

[1]Graduate Program in Biophysics, University of California, San Francisco, San Francisco, CA, USA.

[2]Department of Biochemistry and Biophysics, Institute for Neurodegenerative Diseases, University of California, San Francisco, San Francisco, CA, USA

[3]Department of Chemistry, The University of Alabama at Birmingham, Birmingham, AL, USA

Contributions

The work in this chapter is unpublished at the time of submission. K.E.L. carried out all experiments, refinement and modeling procedures for structure determination; developed figures; and wrote and edited the manuscript. E.T. operated the Krios microscope and helped with data collection. Z.G.C. expressed and purified protein components in the lab of A.L.L. D.R.S. designed and supervised the project and edited the chapter.

Abstract

ClpAP is a AAA+/protease complex responsible for regulated protein degradation in bacteria and interacts with specific substrates. The N-end rule degron (N-degron) is an essential bacterial degron which ClpA alone cannot recognize and requires the adaptor protein ClpS to deliver N-degron substrates. ClpS delivers N-degron substrates to ClpA via the ClpA N-terminal domains (NTD) and it forms a high-affinity delivery complex, however, the mechanism for substrate transfer is unknown. Here I present structures of ClpAPS in complex with an N-degron GFP construct (N-GFP) and show the substrate is processed by the same mechanism as RepA. Extra density at the top of the ClpA channel is a different shape and size than the previously reported ClpAP-RepA structure and is likely ClpS. Focus refinement of the top-most region of the ClpA channel and the extra density reveals low-resolution electron density for the ClpA NTD and ClpS. This work is still ongoing, but from these data, the NTD interacts with ClpS irrespective of the protomer register on the substrate and it is a highly dynamic interaction. Further work must be done to capture high-resolution structures of the high-affinity delivery complex to precisely propose a mechanism for substrate delivery and hand-off.

Introduction

Regulated protein degradation is essential for cellular function and bacteria use several pathways for this[1]. One major pathway is N-end rule degradation which relies on substrates that are tagged with an N-degron[2]. This degron is added to the N-terminal of the substrate and contains a Leucine, Tyrosine, Phenylalanine, or Tryptophan residue at the N-terminal of the tag. N-degron tagged substrates in *Escherichia coli* are unfolded and degraded by the AAA+/protease complex ClpAP[3]. The exact physiological purpose of the N-degron pathway is not understood, but molecular mechanisms involved in this pathway may provide insight into this problem.

Hsp100 (Clp) proteins, primarily ClpA and ClpX, are responsible for regulated protein degradation in *E coli*[4-7]. These hexameric AAA+ complexes are the regulatory component for the protease ClpP and interact at the apical interface via IGL/F loop motifs[8]. Additionally, ClpA and ClpX couple ATP hydrolysis to large conformational changes that lead to substrate translocation[7,9,10]. Structures of the ClpAP and ClpXP complex show the interaction with a substrate in atomic detail and describe a model where large conformational changes are accommodated by IGL/F motif plasticity and flexibility[11-14]. The loops stretch, compress, and disengage ClpP hydrophobic clefts, and in concert with AAA+ conformational changes results in a rotation of ClpA/X relative to ClpP[11,12].

N-degron tagged substrates interact weakly with ClpA and will not be translocated and degraded in the absence of the adaptor protein ClpS[15]. ClpS contains a globular domain with a hydrophobic pocket that binds the N-degron, then ClpS binds the flexible ClpA N-terminal domain (NTD) with increased affinity compared to apo ClpS[16-18]. Once ClpS interacts with the NTD, the ClpS non-structure N-terminal extension (NTE) is partially translocated and the high

delivery N-degron delivery complex is stabilized[19-21]. The substrate is then transferred to the ClpA channel, translocated, and degraded by ClpP.

While structures for the ClpAP complex are informative, they were determined in the presence of RepA which requires no adaptor proteins for delivery[11]. Therefore, it is unclear if a generalized translocation mechanism is used for all substrates, or if there are differences between different substrate classes. Furthermore, the substrate hand-off mechanism between ClpS and ClpA is unknown, and structures of the high-affinity delivery complex will elucidate this mechanism. In this chapter, I describe cryoEM structures of the ClpAP complex bound to ClpS and an N-degron tagged GFP substrate. These structures reveal the same conformations as ClpAP-RepA in chapter 1, which indicates the translocation mechanism for RepA is similar to an N-end rule substrate. Additionally, we observe density at the top of the ClpA channel which is not seen in our previous structure with RepA.

Results

Purified ClpS and N-degron GFP form stable complex with ClpAP

The complex was incubated with ATPgammaS for 15 minutes and then run on an analytical size exclusion chromatography column (SEC) to determine the extent of complex formation. ClpAP in the absence of substrate or ClpS yields a sizing trace that contains a broad peak that is a mixture of ClpAP complex and ClpA hexamer (Fig. 2.1 a, dashed red trace). N-GFP is incubated with ClpAP results in little change in the sizing trace, which indicates it is not able to form a stable complex with ClpAP (Fig. 2.1 a, dashed black trace). On the other hand, ClpS incubated with ClpAP leads to an increase in complex peak intensity which is similar to the

65

sizing trace from the ClpAP/RepA from the previous chapter (Fig. 2.1 a, solid red trace). Finally, ClpS, N-GFP, and ClpAP incubated together dramatically increase the corresponding complex peak and indicate the high affinity delivery complex is formed and stable (Fig. 2.1 a, solid black trace).

To ensure N-GFP is transferred to ClpA, translocated, and degraded by ClpP, we tested proteolytic activity in the presence of ATP. In the absence of ClpS, the N-GFP band intensity does not decrease, which suggests N-GFP is not degraded by ClpAP under these conditions (Fig. 2.1 b). However, it is previously reported that ClpAP degrades ClpS, but it is not essential for N-degron substrates degradation[19]. These data suggest ClpS is not degraded during the 1-hour incubation. When N-GFP and ClpS are present, the band for N-GFP begins to disappear, which indicates the complex is active, and N-GFP is transferred and degraded.

CryoEM structures of the ClpAPS/N-GFP substrate

In contrast to the previous ClpAP structure, the ClpAPS/N-GFP complex cryoEM samples are not isolated from SEC fractions and ATP is not added after incubation with ATPgammaS. Micrographs and 2D class averages show a double capped complex as previously observed and weaker density appears at the top of the complex in the 2D averages (Fig. 2.2 a,b). The 3D classification reveals 3 structural conformations that are identical to the structures reported in chapter 1 (Fig. 2.2 b)[11]. The IGL loops occupy the same pockets in each of the 3 states as before and further supports the rotary mechanism proposed in chapter 1[11].

In 2D class averages and low-pass filtered cryoEM maps, there is extra density at the top of the ClpA pore, and it is globular and connected to the substrate density in the ClpA channel (Fig. 2.1 a, Fig. 2.3 a). Compared to a low pass filtered map from the structure reported in chapter 1, the density is distinct, which indicates it is ClpS, not GFP. The rigid body docked

ClpS crystal structure (PDB 1LZW) into the extra density is encompassed by the density entirely[16]. To obtain a higher resolution map of this region, focus refinement was performed on the top of the ClpA pore and the extra density.

To filter the particles that did not contain the extra density, skip-alignment 3D classification was performed in Relion with a mask containing the area of interest (Fig. 2.4 a). After the classification, ~87% of the particles contained the extra density and were used for subsequent processing. The focused refinement reveals additional density that is connected to the top of the ClpA NBD1, which is likely the ClpA NTD (Fig. 2.4 a). The crystal structure for the ClpS-ClpA NTD (PDB 1LZW) fits into this extra density, where ClpS is in the top blob and the NTD is the density that protrudes from NBD1 (Fig. 2.4 b)[16].

Discussion and future directions

Despite the recent structures of the ClpAP and ClpAPS complexes, there is little known about the N-degron substrate transfer mechanism. The data presented in this chapter attempts to structurally characterize the high-affinity delivery state. I confirmed that individual purified components incubated with ATPgammaS stably form the wild type ClpAPS/N-GFP complex and, with ATP, the complex actively degrades N-GFP.

The structure of the WT ClpAPS/N-GFP complex reveals identical conformational states to the ClpAP/RepA-GFP complex and additional density at the top of ClpA. While these data only show the 3 previously observed structural states (Fig. 2.1 c), structural work done by Sora and colleagues reveals 6 conformations of the wild type ClpAPS/N-GFP complex[22]. Their work reveals the density in the channel is the ClpS NTE, and they obtain a state where there is only density contacting the D1 tyrosine pore loops and not the D2. Biochemical data suggest the D1

pore loops are essential for ClpS/N-degron substrate binding and the D2 are essential for degradation[23]. Their data with NTE interacting with D1 alone confirm this biochemistry and those particles are likely those in the high-affinity delivery complex. However, the other confirmations are similar to our structures reported in Chapter 1 and support the hypothesis that the NTE and N-degron substrates are processed by the same mechanism as the RepA-GFP substrate.

Furthermore, focused refinement reveals extra density for both the ClpA NTD and the ClpS globular domain. In the overall complex structures, the low pass filtered maps show only a small blob of density connected to the density in the ClpA channel, but no apparent density for the NTD. However, in the focused map this extra density is revealed, and this is likely because the NTD binds to ClpS irrespective of what its ATP or substrate binding states are. The mask for focus refinement only incorporates the topmost part of the ClpA NBD1 and the substrate, and therefore the particles during refinement are not aligned to the D1 and D2. Instead, they are aligned to the top of the channel, ClpS, and the NTD. The orientation of the ClpS-ClpA NTD crystal structure fits in the density quite well, however, the resolution of the focus map is not sufficient to determine any atomic details of the ClpS-NTD orientation or if any other contacts are made between ClpS and ClpA.

Overall, the data are promising and reveal some details of the ClpAPS high-affinity delivery complex. The current cryoEM workflow shows that focus refinement pulls out more details in the NTD/ClpS region of the ClpAPS complex, but more particles are required for refinements to reach high-resolution. Additionally, recent biochemical data suggest that D2 pore loop mutations strengthen the affinity of ClpS for ClpA and may be used to further stabilize this complex for cryoEM.

Figures
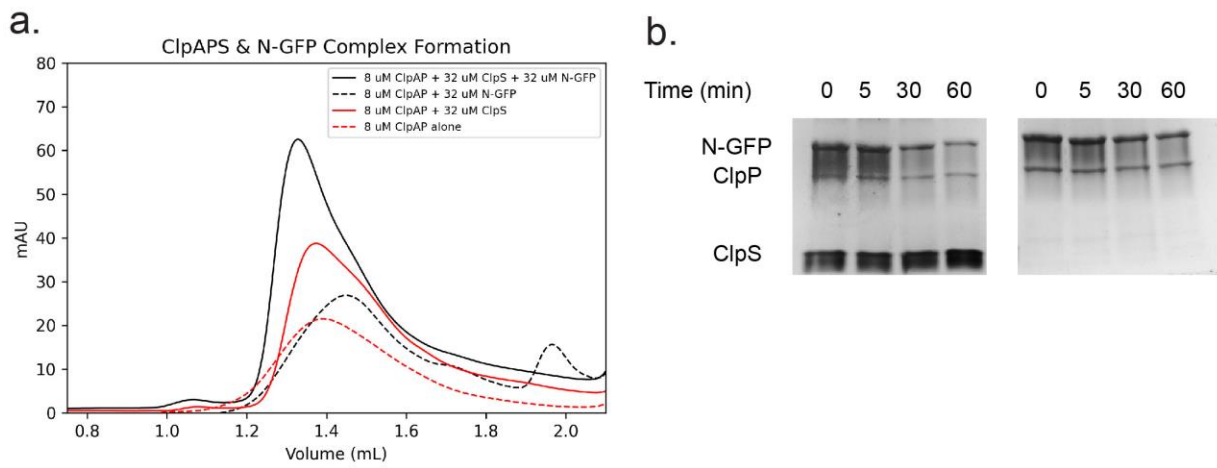
Figure 2.1 Active ClpAPS/N-GFP complex formation a Analytical size exclusion chromatography traces for complex incubation in the presence of ATPgammaS. The complexes shown are ClpAP alone (dashed red), ClpAPS (solid red), ClpAP/N-GFP (dashed black), and ClpAPS/N-GFP (solid black). **b** Silver-stained SDS-PAGE of proteolysis time course over 60 mins. The left gel is incubated with ClpAPS/N-NGFP and the right is ClpAP/N-GFP.
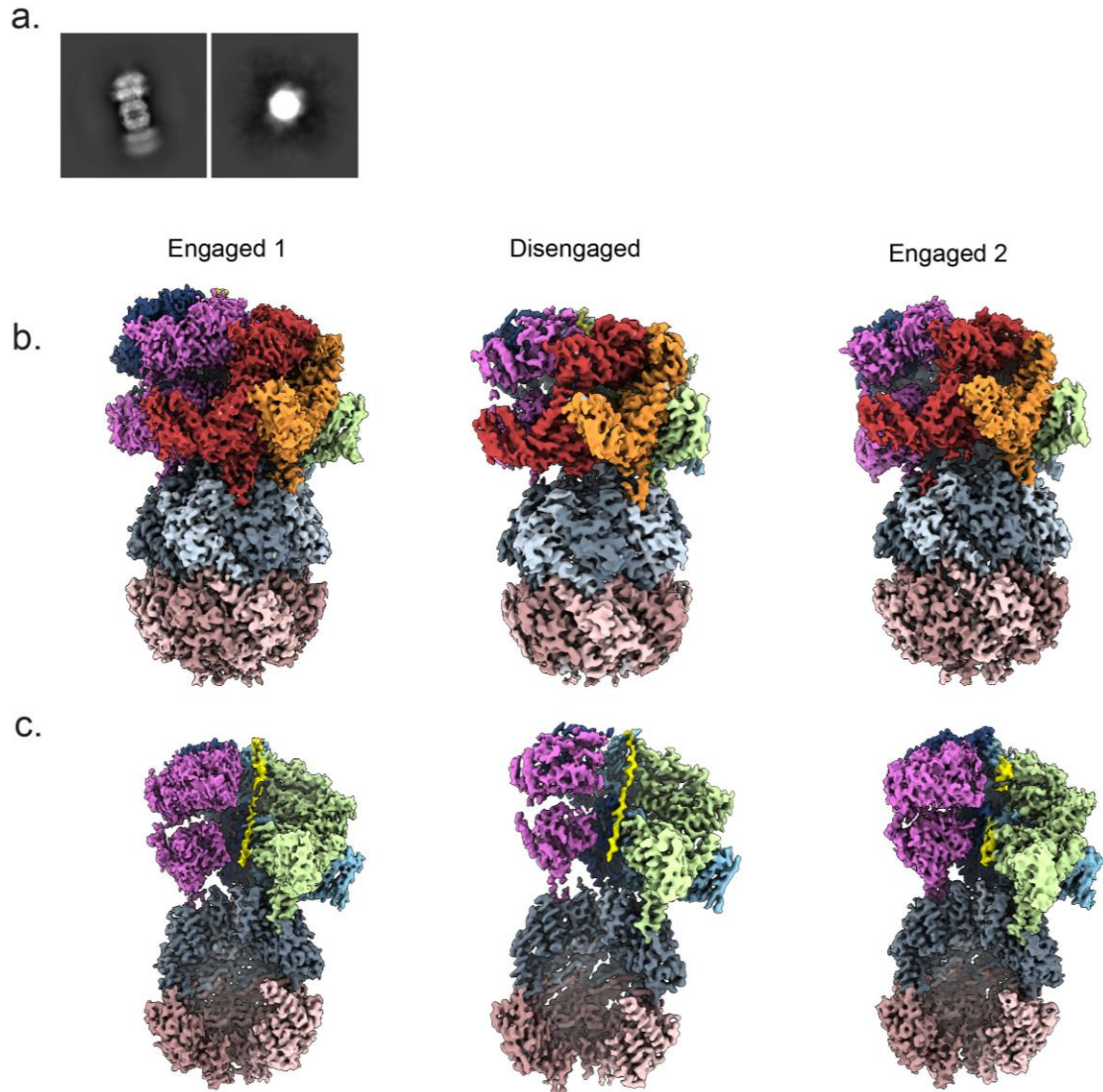
Figure 2.2 Structures of the ClpAPS/N-GFP complex a Reference-free 2D class averages from cryoEM data collection on the ClpAPS/N-GFP complex. The top view is shown on the right and the side view on the left. **b** CryoEM electron density for the Engaged1, Disengaged and Engaged2 states of ClpAP. **c** Channel view of the above structural states with substrate colored in yellow.
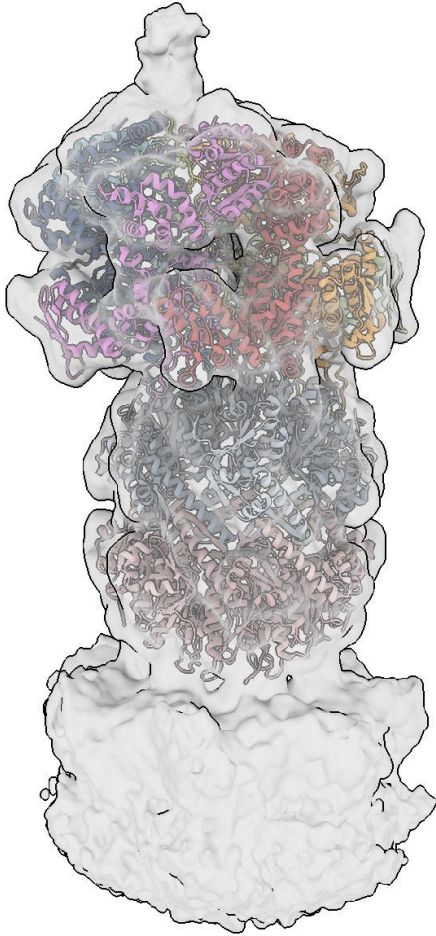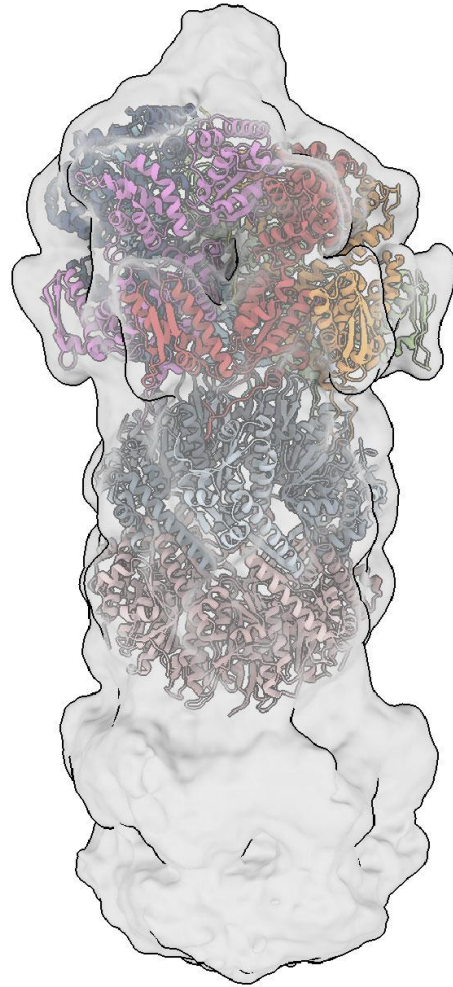
Figure 2.3 Substrate density comparison at ClpA pore entrance a Low pass filtered density for ClpAPS/N-GFP with the model fit into the density. **b** Low pass filtered density for ClpAP-RepA engaged 1 state with model fit into density.
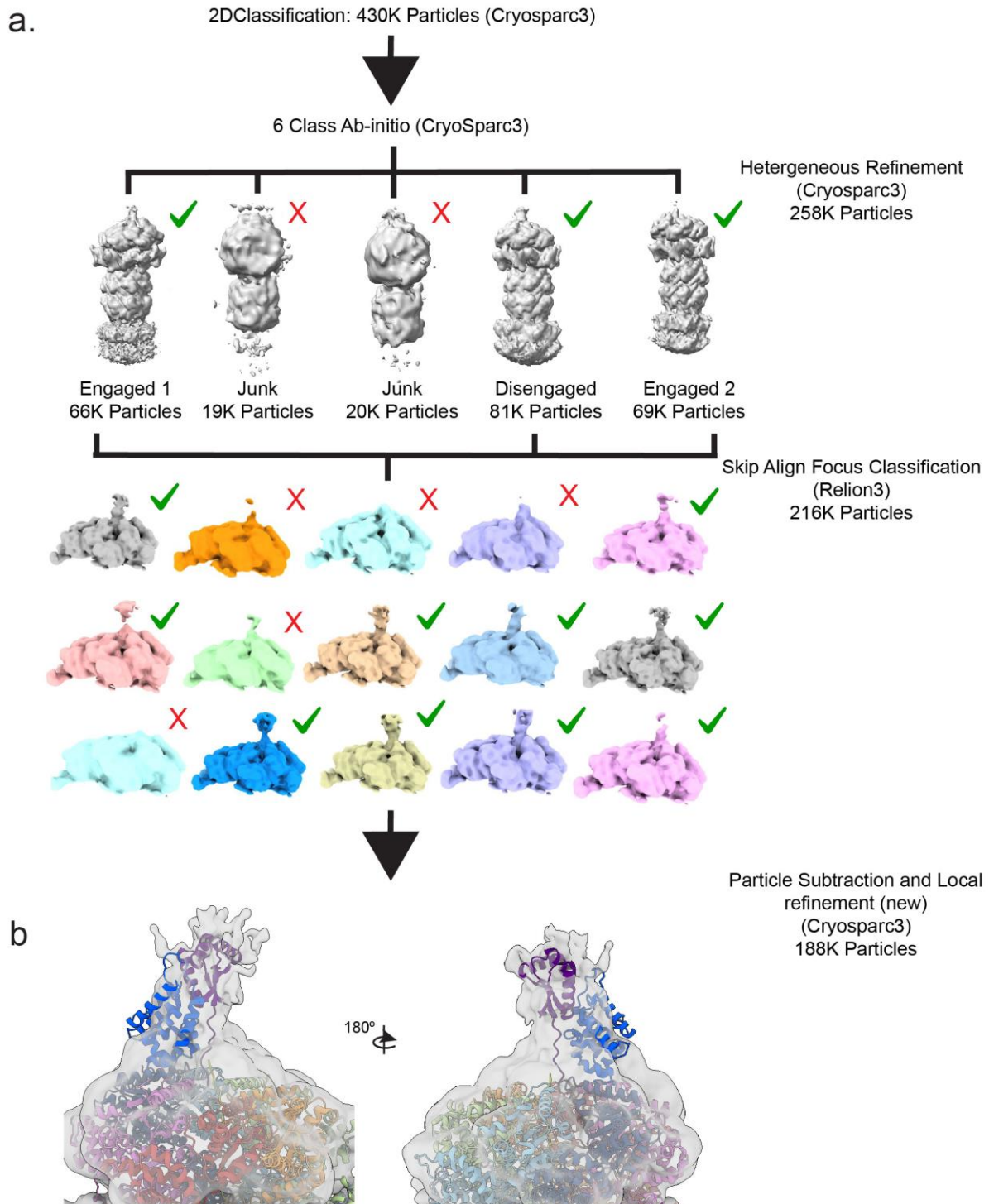
Figure 2.4 ClpAPS/N-GFP refinement strategy to improve ClpS/NTD resolution
a Refinement flow chart for ClpAPS/N-GFP cryoEM data processing. Heterogeneous refinement yields 3 classes that were combined and used for skip alignment classification in relion3. The best classes are indicated with a green checkmark and particles in those classes were used for particle subtraction and local refinement in cryosparc3. **b** Density from local refinement with the crystal structure (PDB 1LWZ) for ClpS/ClpA-NTD interaction rigid body fit into density at the top of ClpA channel[16].

72

Methods

Expression, purification, and analysis of ClpA, ClpP, ClpS, and N-GFP

All proteins were expressed in *E. coli* BL21(DE3) pLysS cells, and ClpA and ClpP were purified as previously described. Full-length ClpS plasmid (genescript) was expressed with an N-terminal His_6 tag that was adjacent to a SUMO cleavage site to ensure the proper N-terminal ClpS residue was present. Transformed cells were grown in Terrific Broth (TB) until OD=1.0 was expression was induced with 1 mM IPTG overnight at 25ºC. Cells were lysed in buffer S (20 mM Hepes pH 7.5, 150 mM KCl), 2 mM BME, 0.2 mM PMSF, and EDTA free protease inhibitor pellets (Roche), and after 1 hr incubation was run through an Avestin emulsiflex 3 times. The lysate was then spun in an ultracentrifuge for 30 mins and the supernatant was run onto a pre-equilibrated Ni-NTA column (Cytiva) with buffer S, 10% glycerol, 15 mM imidazole, and 2 mM BME, washed with 40 mM imidazole, then eluted on a gradient from 40 mM imidazole to 500 mM Imidazole. Eluted fractions were concentrated with 10 kDa cutoff spin concentrators (Corning Life Sciences) and 200:1 ClpS:SenP1SUMO protease was incubated at 4ºC overnight. Simultaneously, the ClpS/SenP1 incubation was injected into a Slide-A-Lyzer 10 kDa dialysis cassette (Thermo Scientific Pierce) in dialysis buffer (20 mM Hepes Ph 7.5, 50 mM KCl, 5 mM BME, 10% glycerol). The dialysate was then added to pre-equilibrated Ni-NTA beads (Thermo Scientific Pierce) and incubated at 4ºC for 1 hr. Incubated beads were applied to a gravity flow column and the flow-through was collected and applied to a pre-equilibrated Mono-Q column (Cytiva) in the same buffer. The sample was then eluted on a KCl gradient from

50 mM KCl to 1 M KCl and eluted peak corresponding to ClpS was concentrated, buffer was exchanged into Buffer S, 2 mM DTT, and 10% glycerol.

N-end rule GFP (nGFP) contained an N-terminal His_6 tag followed by a SUMO cleavage site and the N-end rule tag (YLFVQEL) to ensure the N-terminal residue is tyrosine. Expression and purification were done the same as ClpS, one adjustment is that an S200 column (Cytiva) was used in place of monoQ.

Protein degradation assay

The N-GFP degradation assay (Fig. 2.1 a) was performed in duplicate and consisted of 0.4 uM ClpA, 0.4 uM ClpP, 2.5 uM ClpS, 8 uM N-GFP, 4 mM ATP, 16 mM Creatine phosphate, and 0.32 mg/mL Creatine kinase (for ATP regeneration) incubated in buffer at 20° for15 min containing 50 mM Tris-HCl pH 7.5, 150 mM KCl, 10 mM $MgCl_2$, and 1 mM DTT. Aliquots of the reaction were separated from the reaction at the specified time points and quenched in 2% SDS buffer, heated for 10 min and ran onto an acrylamide gel. The bands were visualized using silver staining (Sigma-Aldrich).

CryoEM data collection and processing

1.5 uM ClpA, 1.5 uM ClpP, 9 uM ClpS, 9 uL N-GFP a 3.5 uL, and 4 mM ATPgammaS were incubated for 10 minutes then a 3.5 uL drop was applied to glow discharged holey carbon grid (R 1.2/1.3; Quantifoil), in which sample was then blotted for 2.5 s. at 4° and 100% humidity with Whatman No. 1 filter paper before being plunge frozen liquid ethane using a vitrobot (Thermo Fischer Scientific). The sample was then imaged on a Titan Krios TEM (Thermo Fischer Scientific) operated at 300 keV and equipped with a Gatan BioQuantum imaging energy filter using a 20eV zero loss energy slit (Gatan Inc). Movies were acquired in super-resolution

mode on a K3 direct electron detector (Gatan Inc.) at a calibrated magnification of 58,600X

corresponding to a pixel size of 0.4265 Å/pixel. A defocus range of .8 to 1.2 μm was used with a

total exposure time of 2 seconds fractionated into 0.2s subframes for a total dose of 68 e^-/Å^2 at a

dose rate of 25 e^-/pixel/s. Movies were subsequently corrected for drift using MotionCor2[24]

(10.1038/nmeth.4193) and were Fourier-cropped by a factor of 2 to a final pixel size of 0.853

Å/pixel.

A total of ~5,000 micrographs were collected and Initial and downstream data-processing

was performed in cryosparc3[25] and Relion3[26]. For particle picking, templates were generated in

cryosparc3 from 100 particles, in which top and side-views were selected. After inspecting the

particles picked, approximately 950,000 particles were extracted with a 576 Å box size. One

round of 2D classification was performed to remove contamination and junk particles, which

amounted to ~55% of the dataset. A six-class ab-initio reconstruction was performed from the

particle set (~430,000 particles) and was used for initial classification steps.

To identify different conformations, heterogenous refinement was performed with 5

different classes (Fig. 2.4 b) in cryosparc3. Following this first round, maps showing high

resolution features, ~216,000 particles were kept and grouped together and the 3 conformational

states were observed. The Engaged1 state has ~66,000 particles, the Disengaged state has

~81,000 particles, and the Engaged2 state has ~69,000 particles. The 3 classes were pooled for

focus refinement to improve the substrate density.

To improve the resolution of the ClpS density, local refinement (new) in cryosparc3 was

used with the pooled particles and a mask that encompasses the top of the ClpA channel and

entrance (Fig. 2.4 a). The resolution did not improve for the ClpS density, so skip alignment

classification in Relion3 was performed with the same mask, aligned particles from the local

refinement job, and a T parameter of 20. This resulted in ~28,000 particles that lacked the ClpS density and ~188,000 particles that contained good ClpS density. The particles that lacked density were removed from the data set. The ~188,000 particles were then reimported into cryosparc3, and the inverse mask used for local refinement and skip align classification was used to subtract the ClpA bottom half and ClpP regions from the particles. The subtracted particles then underwent local refinement (new) in cryosparc3 using the previous mask. This resulted in a 2.8 Å map, however, the density for the ClpS and NTD showed no high resolution features. In the future, more data will be collected and more classification techniques will be employed to obtain better resolution of this region.

References

1.   Gottesman, S. Proteolysis in bacterial regulatory circuits. *Annu Rev Cell Dev Biol* **19**, 565-87 (2003).

2.   Mogk, A., Schmidt, R. & Bukau, B. The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol* **17**, 165-72 (2007).

3.   Varshavsky, A. The N-end Rule. *Cell* **69**, 725-735 (1992).

4.   Wickner, S. et al. A molecular chaperone, ClpA, functions like DnaK and DnaJ. *Proc Natl Acad Sci U S A* **91**, 12218-12222 (1994).

5.   Grimaud, R., Kessel, M., Beuron, F., Steven, A.C. & Maurizi, M.R. Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP. *J Biol Chem* **273**, 12476-81 (1998).

6.   Singh, S.K., Grimaud, R., Hoskins, J.R., Wickner, S. & Maurizi, M.R. Unfolding and internalization of proteins by theATP-dependent proteases ClpXP and ClpAP. *Proc Natl Acad Sci U S A* **97**, 8898-8903 (2000).

7.   Hoskins, J.R., Singh, S.K., Maurizi, M.R. & Wickner, S. Protein binding and unfolding by the chaperone ClpAand degradation by the protease ClpAP. *Proc Natl Acad Sci U S A* **97**, 8892-8897 (2000).

8.   Kim, Y. et al. Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nature* **8**, 230-233 (2001).

9.   Gates, S.N. et al. Ratchet-like polypeptidetranslocation mechanism of theAAA+ disaggregase Hsp104. *Science* **357**, 273-279 (2017).

10.   Hinnerwisch, J., Fenton, W.A., Furtak, K.J., Farr, G.W. & Horwich, A.L. Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation. *Cell* **121**, 1029-41 (2005).

11.   Lopez, K.E. et al. Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. *Nat Struct Mol Biol* **27**, 406-416 (2020).

12.   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**(2020).

13.   Fei, X., Bell, T.A., Barkow, S.R., Baker, T.A. & Sauer, R.T. Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. *Elife* **9**(2020).

14.   Fei, X. et al. Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. *Elife* **9**(2020).

15.   Dougan, D.A., Reid, B.G., Horwich, A.L. & Bukau, B. ClpS, a Substrate Modulator of the ClpAP Machine. *Molecular Cell* **9**, 673-683 (2002).

16.   Zeth, K. et al. Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. *Nat Struct Biol* **9**, 906-11 (2002).

17.   Hou, J.Y., Sauer, R.T. & Baker, T.A. Distinct structural elements of the adaptor ClpS are required for regulating degradation by ClpAP. *Nat Struct Mol Biol* **15**, 288-94 (2008).

18.   Roman-Hernandez, G., Grant, R.A., Sauer, R.T. & Baker, T.A. Molecular basis of substrate selection by the N-endrule adaptor protein ClpS. *Proc Natl Acad Sci U S A* **106**(2009).

19.   Roman-Hernandez, G., Hou, J.Y., Grant, R.A., Sauer, R.T. & Baker, T.A. The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. *Mol Cell* **43**, 217-28 (2011).

20.   Rivera-Rivera, I., Roman-Hernandez, G., Sauer, R.T. & Baker, T.A. Remodeling of a delivery complex allows ClpS-mediated degradation of N-degron substrates. *Proc Natl Acad Sci U S A* **111**, E3853-9 (2014).

21.   Torres-Delgado, A., Kotamarthi, H.C., Sauer, R.T. & Baker, T.A. The Intrinsically Disordered N-terminal Extension of the ClpS Adaptor Reprograms Its Partner AAA+ ClpAP Protease. *J Mol Biol* **432**, 4908-4921 (2020).

22.   Kim, S., Fei, X., Sauer, R.T. & Baker, T.A. AAA+ protease-adaptor structures reveal altered conformations and ring specialization. *bioRxiv* (2022).

23.   Zuromski, K.L., Kim, S., Sauer, R.T. & Baker, T.A. Division of labor between the pore-1 loops of the D1 and D2 AAA+ rings coordinates substrate selectivity of the ClpAP protease. *J Biol Chem* **297**, 101407 (2021).

24.   Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).

25.   Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).

26.   Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**(2018).

Acknowledgments

Chapter 3

Structural basis of an Acyldepsipeptide fragment stimulation of *Mycobacterium tuberculosis* protease ClpP1P2 activity

Kyle E. Lopez[1,2], Eric Tse[2], Cady Burnside[3,4], Monika Prorok[3,4], Fan Fei[5,6], Jason Sello[6], Karl R. Schmitz[4], and Daniel R. Southworth[2]

Affiliations:

[1]Graduate Program in Biophysics, University of California, San Francisco, San Francisco, CA, USA.

[2]Department of Biochemistry and Biophysics, Institute for Neurodegenerative Diseases, University of California, San Francisco, San Francisco, CA, USA

[3]Graduate Program in Chemistry, University of Delaware, …

[4]Department of Biological Sciences, University of Delaware…

[5]Graduate Program in Chemistry, Brown University, Providence, RI, USA

[6]Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA, USA

Contributions

The work in this chapter is not published at the time of submission, but the manuscript will be submitted for publication at a later date. K.E.L. carried out all cryoEM refinement and modeling procedures for structure determination; developed figures; and wrote and edited the manuscript. E.T. operated the Krios microscope and helped with data collection. M.P and C.B. performed all biochemical experiments, protein expression, and purification. F.F. synthesized all ADEPs, agonist, and ADEP fragments. K.R.S. and J.S. wrote and edited the manuscript. D.R.S. designed and supervised the project and wrote and edited the manuscript.

Abstract

The *Mycobacteria tuberculosis* ClpP1P2 protease complex is an essential protein for virulence and acyldepsipeptides (ADEP) kill *M tuberculosis* by blocking the AAA+ ClpX interaction. Fragments lacking parts of the macrocycle stimulate MtbClpXP1P2 activity rather than inhibit and here we report atomic-resolution structures of MtbClpP1P2 bound to an ADEP fragment in the absence of full ADEP and dipeptide agonist. The fragment mimics the ClpX LGF loop and binds in the ClpP2 hydrophobic cleft and surprisingly binds in the homologous ClpP1 cleft which ClpX and ADEPs fail to bind. This opens both the ClpP1 and ClpP2 pores to substrate entry and exit. Additionally, the fragment binds in the protease active site, mimics the agonist, and pushes the complex to an active conformation. Finally, we use the structural information to rationally design better MtClpP1P2 stimulators from the original fragment. With this data, we propose a model of stimulation where weak affinity for the apical, ClpP2 hydrophobic clefts allow ClpX to compete away fragment, and the fragment is available to bind to the active site and stimulate activity as a more potent agonist molecule.

Background

Bacterial infections and antibiotic-resistant bacteria strains are one of the most pressing public health threats in the modern age[1,2]. Most antibiotics target DNA replication, transcription, and translation which are prone to mutations that keep cells virulent and resistant to these small molecules[3]. *Mycobacterium tuberculosis* causes the deadly respiratory illness tuberculosis, and several strains are resistant to multiple potent antibiotics[4-6]. To better handle the public health threat posed by *M. tuberculosis* and other multi-drug resistant bacteria, new antibiotics that target other cell processes must be developed.

Bacterial caseinolytic protease (ClpP) is a tetradecameric, serine protease regulated by AAA+ unfoldase proteins. ClpP contains the canonical serine protease Ser-His-Asp catalytic triad at the active site and the enclosed cavity consists of 14 active sites[7-9]. While not an essential gene in bacteria such as E. coli, ClpP is essential for M. tuberculosis cell survival and virulence and other bacteria responsible for infectious disease[7,10-14]. In M. tuberculosis, the ClpP complex consists of the ClpP1 and ClpP2 heptameric rings, stacked atop each other to create a catalytic cavity in the center. ClpP1 is responsible for a majority of the proteolytic activity, however, it is unable to interact with the regulatory AAA+ proteins ClpX and ClpC1. ClpP2 on the other hand contributes little to catalytic activity, but interfaces with the hexameric AAA+ complex and allows unfolded substrate to enter the central cavity[15].

On its own, ClpP1P2 is unable to degrade folded substrates and relies on the AAA+ unfoldases ClpX and ClpC1[16]. Structures of E. coli ClpXP and ClpAP demonstrate how the unfoldases couple protein unfolding to proteolysis and how the AAA component interacts with ClpP in atomic detail[17-20]. In M. tuberculosis ClpP1P2, ClpX interacts with hydrophobic pockets at the apical ClpP2 surface through flexible LGF motifs similar to its E Coli homolog, however,

84

the analogous pocket in ClpP1 is not large enough to accommodate this motif[15,21,22]. Because this

system is unique to bacteria and is required for mycobacterium virulence, it is a great target for

new antibiotics[23-28].

Small molecule acyldepsipeptides (ADEP) are natural products that inhibit ClpXP1P2

function by mimicking the LGF motif and binding in the ClpP2 hydrophobic cleft[24,29,30]. By

inhibiting this essential activity, ADEPs successfully kill M. tuberculosis in culture, which

makes them an interesting compound for antibiotic treatments[31]. When docked in the

hydrophobic pocket, ADEPs open the ClpP2 N-terminal gate which allows degradation of small

peptides and unstructured substrates, however, they do not bind to the analogous pocket in

ClpP1[22,32,33]. Additionally, a small peptide agonist that binds in the ClpP1P2 active sites, aligns

the catalytic triad and is essential for catalytic activity[8,33,34]. In an attempt to develop more potent

inhibitors, Schmitz and colleagues fragmented the ADEP scaffold and observed stimulated

activity with a particular fragment. While the fragment stimulated activity, the binding site and

mechanism are unclear[28,35,36].

Here we report high-resolution structures that indicate several binding locations of the

ADEP fragment and propose a model for the stimulation mechanism. With our structural

observations, we further improve upon the fragment by increasing the stimulatory effects.

Results

ADEP fragment opens both the MtClpP1 and MtClpP2 pores

We used CryoEM to structurally characterize the MtClpP1P2 complex bound to an

ADEP fragment and determined a high-resolution structure. The fragment molecule retains most

of the chemical groups from the original ADEP, however, the macrocycle is broken and

introduces flexibility into the molecule and lower binding affinity (Fig. 3.1 a,b)[36]. Micrographs and reference-free 2-dimensional (2D) class averages of MtClpP1P2 particles bound to fragment show both top and side views, and there are 2 distinct classes of the side orientation (Fig. 3.1 b and Supplemental Fig. 3.S1 b,c). The two views show a single tetradecamer complex and the other shows 2 ClpP2 heptamer rings interacting via the beta-hairpin motif that gates the ClpP2 pore entrance. Refinement of the single complex class yields electron density for the compact, inactive complex, and thus does not represent the active, fragment-stimulated state (Supplemental Fig. 3.S1 a). For reconstructions of the active complex, the refinement box encompasses a single tetradecamer, although some density appears for the second tetradecamer, and we determined a 2.8 Å structure (Fig. 3.1 c and Supplemental Fig. 3.S1 d). This atomic resolution is sufficient to unambiguously identify the fragment binding site and the specific interactions it makes with ClpP1P2.

The first obvious structural features are the ClpP2 beta hairpins that serve as a gate to the ClpP2 pore entrance and allow the substrate to enter the cavity when in the upward position (Fig. 3.1 d). However, these same gates are not present on the ClpP1 ring since ClpX does not bind to this ring and thus no substrate enters from this side of the complex[32,37,38]. For the ClpP2 rings to assume this conformation, one of three things must happen: 1) ClpX or ClpC1 interacts with ClpP2, 2) an ADEP is bound in the ClpP2 hydrophobic cleft, or 3) a canonical dipeptide agonist is bound in the active site[33,37].

ADEPs bind to the MtClpP2 hydrophobic pocket, mimic MtClpX or MtClpC1 LGF loop binding and stabilize the N-terminal beta hairpins in an open, upward conformation, which allows the substrate to enter the cavity. There are no known ADEPs that interact with the MtClpP1 pocket because the pocket is shallower and cannot accommodate ADEP or LGF motif

binding. Our structure contains ADEP fragments in both the MtClpP2 and MtClpP1 hydrophobic

clefts (Fig. 3.2 a,b).

The fragment bound in the MtClpP2 pocket binds in the same conformation as a typical

ADEP (Fig. 3.2 a), but the macrocycle region is missing some density (Supplemental Fig. 3.S2

a). Residues L61, Q101, T73, Y95, and Y75 appear to interact with the fragment similarly to

ADEP, however, R97 has poor side chain density, suggesting it does not interact with the

fragment (Supplemental Fig. 3.2 b). The missing interactions may explain why the fragment does

not bind tightly to ClpP1P2 compared to ADEP and ClpX. Previous data suggest mutating

residues Y75 and Y95 indeed diminish fragment binding at the ClpP2 cleft[15,22,36].

More interestingly, the fragment occupies the MtClpP1 hydrophobic cleft in our

structure, and similar to the MtClpP2 pocket density, the diflourophenylalanine, and acyl chain

density are present in the pocket, but the density for the entire macrocycle is missing (Fig. 3.2 b

and Supplemental Fig. 3.S2 b). The flexibility in the broken macrocycle likely allows the

fragment to bind in the cleft without steric clash with Y91 and Y113 that normally stick out into

the pocket (Fig. 3.2 b). Additionally, the cleft expands in the presence of fragment compared to

the apo pocket in the ADEP structure (Fig. 3.2 c). Although we observe fragment binding in the

MtClpP1 pocket, mutagenesis experiments with the combined mutations S61A, Y63V, L83A,

and Y91V that enable ADEP binding suggest this binding site does not contribute to the

observed stimulation[22,36].

The MtClpP2 pocket opens with a 24 Å diameter, similar to previously reported distances

with ADEP or agonist bound[19,22,33]. However, the MtClpP1 pore diameter opens from 10 Å in

the apo structure to 20 Å in the fragment structure with and without ADEPs. With the pore

doubled in size, this may facilitate more efficient diffusion of degraded peptides out of the

cavity, thus clearing up space for more substrate to get cleaved. However, with nothing bound in the MtClpP1 cleft and agonist bound in the active site, the pore is the same diameter as our fragment bound structure[22]. The fragment may bind to the active site which opens up the MtClpP1 cleft for fragment binding, which further stabilizes the active conformation.

MtClpP1P2 bound to ADEP fragment adopts an active conformation

Additionally, the ClpP2 residues D184 and R185, and the ClpP1 residues R171 and D170 serve as an oligomerization sensors and form salt bridges that stabilize the active conformation[29,33,39,40]. In our structure, we observe clear density for the sensor that forms stable salt bridges and show our complex is in the active conformation (Fig. 3.3 a,b, and Supplemental Fig. 3.S3 a). The root mean squared deviation (RMSD) between the oligomerization sensor in our structure and the active, Bz-LL bound structure is 1.3 Å. Meanwhile, the RMSD for the sensor in our structure and the inactive, ADEP bound structure is 3.5 Å. These data show the fragment is sufficient to stabilize the oligomerization sensor and stabilize the active MtClpP1P2 conformation.

In addition to the sensor region, the beta-sheet handle region at the interface between ClpP1 and ClpP2 monomers indicates an aligned catalytic triad and active complex[41]. In ADEP bound and apo structures, this region is disordered and explains why ADEPs in the absence of agonist are unable to activate the MtClpP1P2 complex (Fig. 3.3 c-f)[22,33]. On the other hand, structures of MtClpP1P2 bound to the agonist Bz-LL show this region ordered and the catalytic triad aligned for active proteolysis[22,33] (Fig. 3.3 f). The local resolution of the handle region in the reported structure is comparable to the global resolution (2.8 Å) and therefore the handle is well ordered and stably formed in the structure (Fig. 3.3 c and Supplemental Fig. 3.S3 b). The

fact that the handle is ordered in the structure with fragment alone suggests the fragment is sufficient to activate the complex.

ADEP fragment binds in both protease active sites

The handle region stability is an indicator of the status of the catalytic triad, and when it is ordered the triad is aligned for active proteolysis. In the MtClpP1P2 structure bound to fragment, this handle region is ordered, and the catalytic residues are aligned (Fig. 3.4 a and Supplemental Fig. 3.S4 b). There is no reported compound other than dipeptide agonist and covalent inhibitors that stabilize the handle and now this ADEP fragment. This resulted in the observation that the fragment is bound to the MtClpP2 active site and occupies the same space as a previously reported structure with an agonist bound (Fig. 3.4 a). The weak density makes it difficult to model in the fragment and suggests a weak interaction occurs at the MtClpP2 active site. Indeed, the dipeptide agonist also weakly binds to this active site and may explain why the MtClpP2 active site contributes little to the overall catalytic activity.

Additionally, a much stronger density is observed in the MtClpP1 active site and most of the fragment fits into the electron density. The fragment appears to bind in the active site slightly different than the agonist, where the N-acyl chain fits into a hydrophobic pocket that is typically occupied by an aromatic ring in agonist bound structures. It is likely the increased electronegativity of the fluorine atoms in the difluorophenylalanine moiety prevents binding in this pocket and as a result, the difluorophenylalanine appears to point towards the cavity and has weak density due to flexibility.

To support our structural observations, we tested the peptidolysis activity of MtClpP1P2 in the presence of a fragment without an agonist. While ADEP alone does not stimulate activity, the fragment alone is sufficient for activation, and activity is further stimulated when the agonist

is included. While we expect fragment stimulation to occur similarly regardless of agonist presence, it is likely further stimulated by the agonist because the fragment still binds to the ClpP2 and ClpP1 hydrophobic clefts. Indeed, when ADEP and fragment are present, agonist does not appear to affect activity at all, which supports our hypothesis that the fragment acts as a more potent agonist molecule (Fig. 3.4 c). These data support the model where fragment stimulates activity by mimicking agonist and aligns the catalytic residues.

Discussion

The inhibition mechanism of ADEPs is well understood since their discovery in 2005 and is of great interest for antibiotic uses, especially in *M tuberculosis*[24]. Schmitz and colleagues developed a fragmented ADEP molecule where the macrocycle is broken, and the molecule stimulates the activity of the MtbClpXP1P2 complex[28,34,36]. The structures presented here show this fragment binds in the ClpP2 and ClpP1 hydrophobic clefts, in addition to both ClpP1 and ClpP2 protease active sites. Biochemical data support the proposed stimulation mechanism from the structural data and show the fragment alone is sufficient to stimulate peptidolytic activity. Furthermore, when chemical groups are altered based on structural observations, we effectively increase or decrease stimulation.

The high-resolution structure shows the ClpP2 pore widening similar to previous ADEP bound structures which allows small peptide substrates to diffuse into the catalytic cavity. However, the fragment is more flexible than the ADEP due to the broken macrocycle and the fragment does not make as many interactions with the pocket as ADEPs. This explains why the fragment has a weaker affinity for the ClpP1P2 complex compared to ADEPs and since it does

not inhibit ClpXP1P2 proteolytic activity it likely has a weaker affinity compared to ClpX LGF loops (Fig. 3.5).

Additionally, the fragment binds in the ClpP1 hydrophobic cleft which ADEPs and ClpX LGF loops are unable to access. Morphs and RMSD values between apo ClpP1 cleft show the pocket widens to accommodate the fragment, and the flexibility in the broken macrocycle allow the molecule to avoid steric clashes experienced by ClpX and ADEPs. When bound in this pocket, the ClpP1 pore opens similar to previous the BZ-LL bound structure and is wide open compared to the apo structure[22]. It is possible that when this pore is open, cleaved product diffuses out of the cavity more rapidly than when it is closed. This decreases the concentration of product in the cavity and pushes the reaction forward, but since product release is not a rate-limiting step in this reaction, this is likely a minor effect on the proteolytic simulation.

The fragment also binds in both the Clp1 and ClpP2 active sites, although the density is weak for the ClpP2 site. From previous data and structures, most agonists weakly interact with the ClpP2 active site and it doesn't significantly contribute to overall proteolytic activity[15]. It is more likely the interaction with the ClpP1 active site is what contributes to the stimulatory effects of the fragment. The fragment appears to make other interactions in proximity to the active site that differ from the agonist and likely make this molecule a more potent agonist. These structural and biochemical observations lead us to propose a model where ClpX competes away fragment from the ClpP2 cleft and fragment binds in the active site and ClpP1 cleft. The interactions in the active site align the catalytic triad and allow rapid substrate degradation to occur and the interaction in the ClpP1 cleft opens the pore for substrate diffusion out of the cavity. Atomic details of the fragment interactions allow the rational design of new fragment molecules that decreases MtClpP2 cleft affinity and stimulate activity, while fragment

simultaneously enhances protease active site interactions. Further studies must be done to test the

antibiotic efficacy of these new fragments, but the detailed structural information allows us to

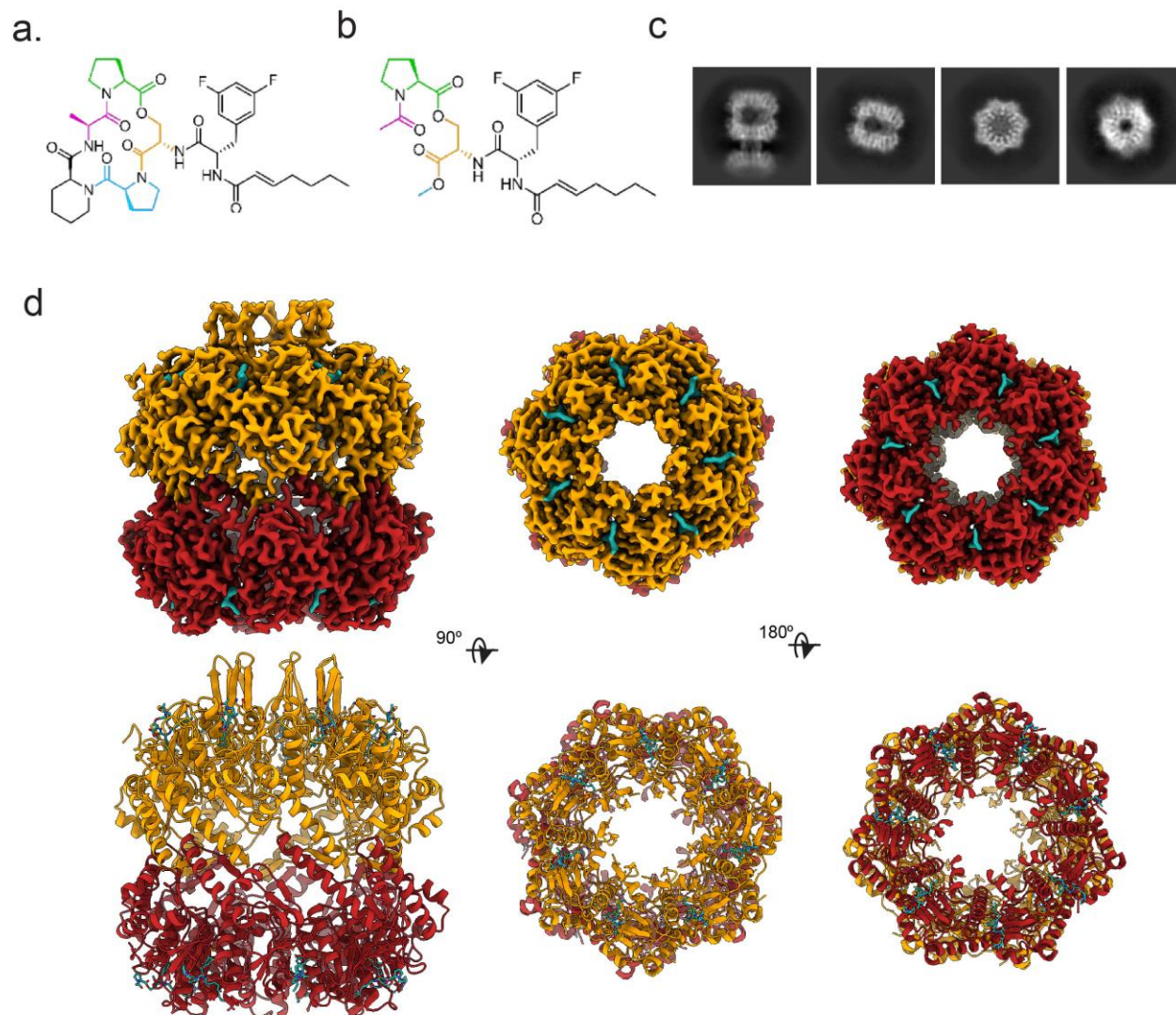change the molecule to fit various needs with no effect on desired outcomes.

Figures

Figure 3.1 Structure of ADEP, ADEP fragment, and MtClpP1P2 a Molecular structure of ADEP fragment and **b**, ADEP fragment. **c**, Representative reference-free 2D class averages which show side views where the first shows a double MtClpP1P2 complex and the second, a single, inactive complex. The averages also show 2 top views to the right. **d**, The electron density (above) and model (below) for the MtClpP1P2 complex bound to fragment with MtClpP2 colored orange, MtClpP1 colored red, and ADEP fragment colored teal.
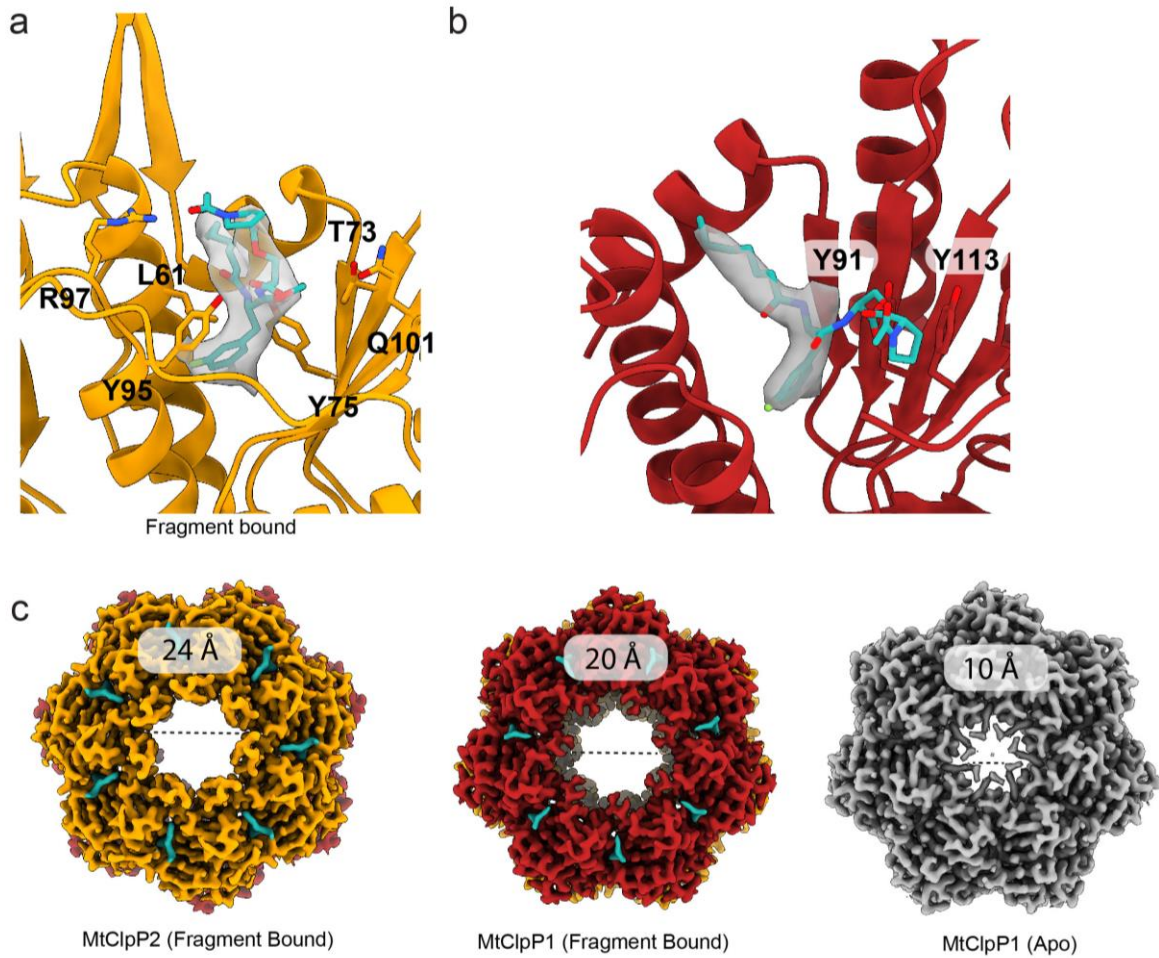
93

Figure 3.2 ADEP fragment and hydrophobic cleft interactions a Model for the MtClpP2 hydrophobic cleft and **b**, MtClpP1 with important residues displayed as sticks and labeled. The ADEP fragment electron density is displayed with the fragment stick model shown. **c**, Electron density for the MtClpP2 (left), MtClpP1 (middle), and an apo MtClpP1 (right) pore with diameter measured and labeled.
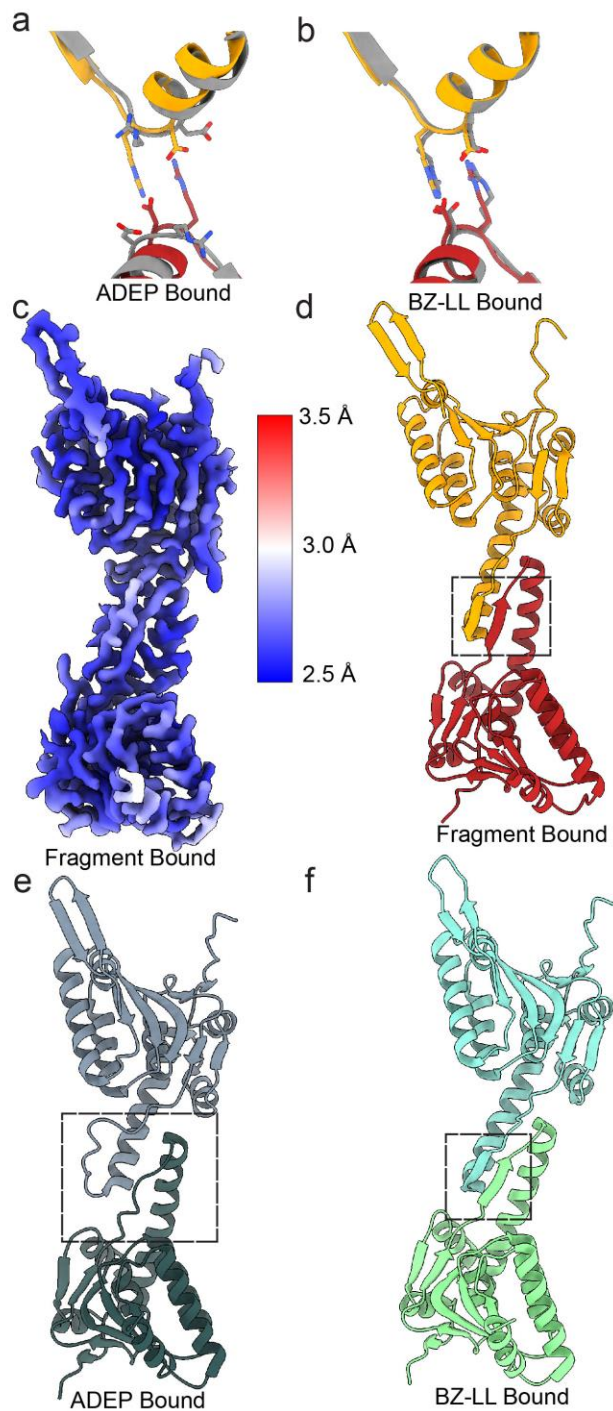
Figure 3.3 The active MtClpP1P2 complex a Model for the oligomerization sensor compared to ADEP bound structure (grey) and **b**, an agonist (BZ-LL) structure (grey). **c**, Local resolution for an interaction MtClpP1 and MtClpP2 monomer with resolution values displayed on the right, where blue indicates high-resolution regions and red represents low resolution. **d**, Molecular model of an MtClpP1 and MtClpP2 monomer for our fragment bound structure, **e**, an ADEP bound structure (PDB 6VGN), and **f**, a BZ-LL bound structure (PDB 5DZK)[38] with the beta-sheet handle region encompassed by the dashed-line box.
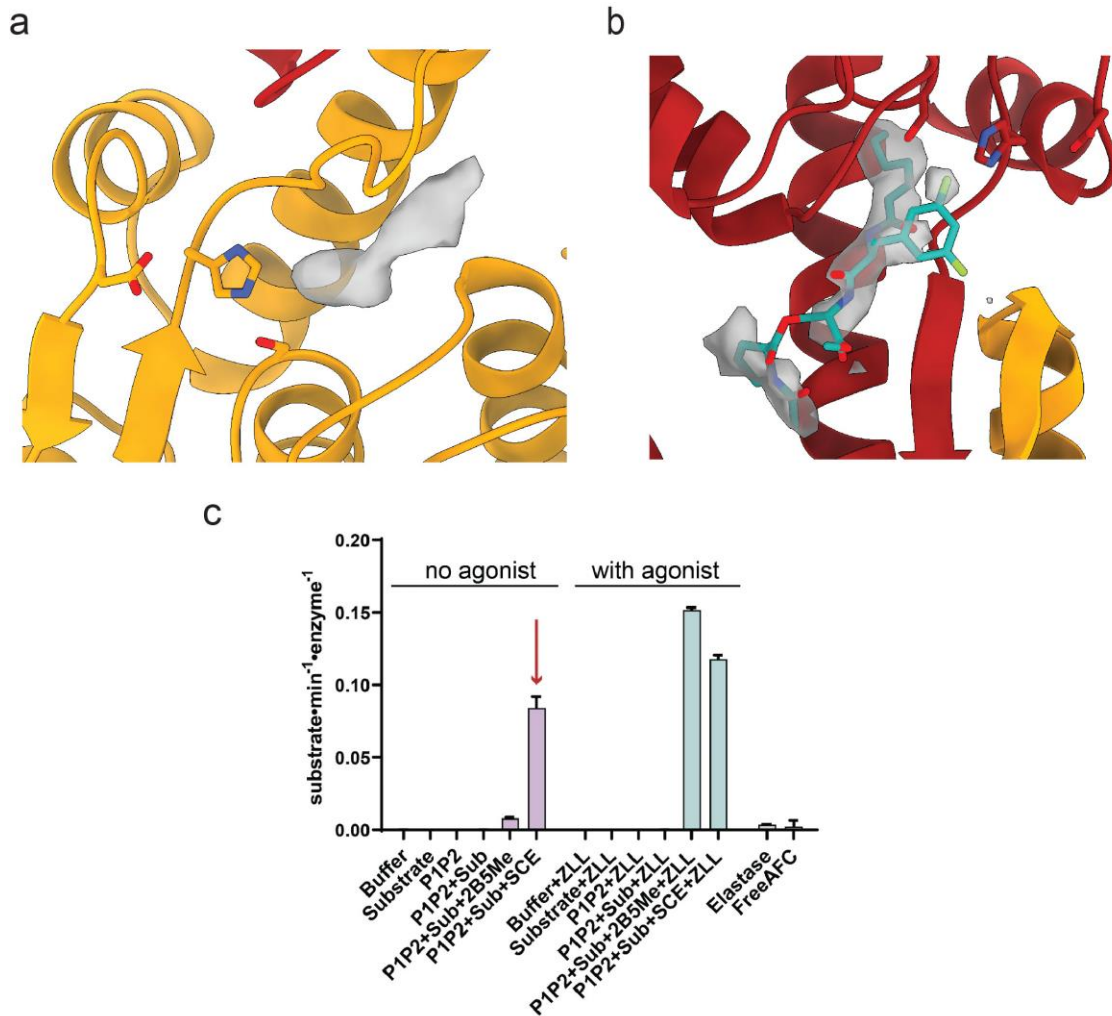
Figure 3.4 The agonist bound in both the MtClpP1 and MtClpP2 active sites a Molecular model displayed for the MtClpP2 and **b**, MtClpP1 active sites with catalytic triad represented as sticks. For the (**a**) MtClpP2 active site, extra electron density corresponding to the ADEP fragment is shown in grey, and **b**, for the MtClpP1, fragment density is also displayed and the fragment is modeled into the density. **c**, Peptidolysis assay where the left curves are run without agonist and the right curves with agonist present. The conditions corresponding to our structure (MtClpP1P2 and fragment) are shown with the red arrow.
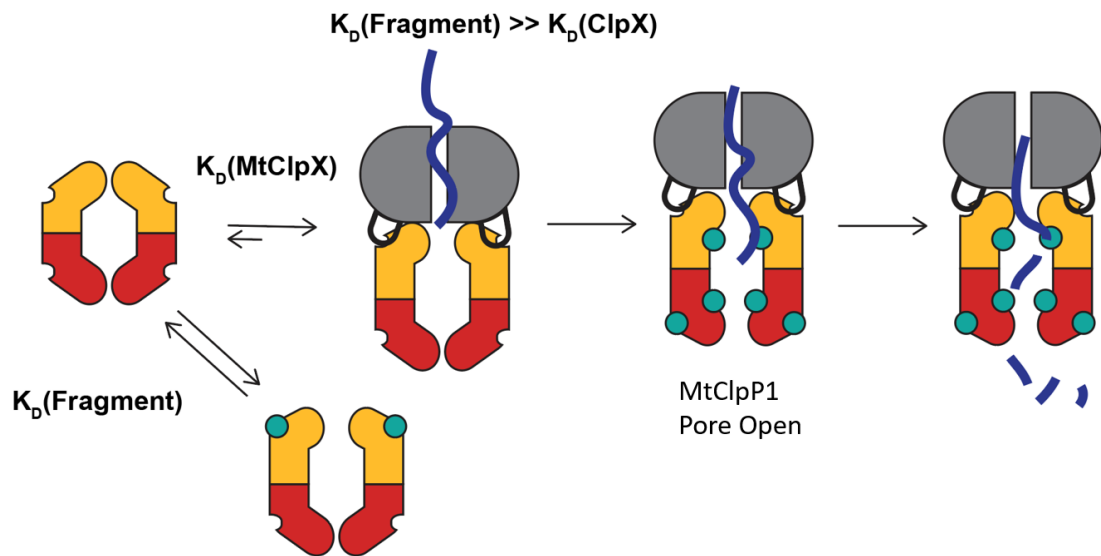
Figure 3.5 Fragment stimulation mechanism model MtClpP1P2 and fragment are colored as previously mentioned and MtClpX is colored in grey with LGF loops displayed as loops. Apo state shows both pores in the closed conformation. ClpX and fragment are depicted competing for the MtClpP2 hydrophobic cleft, but ClpX has a tighter interaction than the fragment. Unfolded substrate (blue) is fed into the cavity where fragment binds in both active sites, as well as the MtClpP1 hydrophobic cleft. The substrate is degraded when the fragment is bound in the active site and diffuses our the open, MtClpP1 pore.

Methods

MtClpP1 and MtClpP2 expression and purification

The MtClpP1 and MtClpP2 constructs used were expressed and purified as previously described[36].

Enzymatic assays

ClpP concentrations are reported in 14-mer equivalents. All in vitro kinetic assays were carried out at 30 C in a SpectraMax M5 microplate reader (Molecular Devices) in 40 uL reactions under initial velocity conditions. The peptidase activity of ClpP1P2, in the presence or absence of ligands as indicated in figure legends, was assayed in PD buffer (25 mM HEPES, 100 mM KCl, 5 mM MgCl2, 10% (wt/vol) glycerol, 10% (wt/vol) DMSO, 1 mM DTT, 0.1 mM EDTA) in the presence of 50 uM BZ-LL peptide agonist (Sigma). Glycerol was included to improve behavior of ClpP1 and ClpP2. 10% DMSO was included to ensure solubility of peptide components, particularly the BZ-LL agonist. Cleavage of fluorogenic tetrapeptide (Z-YVAD-AFC; 50 µM; Enzo Life Sciences) was followed by increase in 505 nm 7-amino-4-trifluoromethylcoumarin fluorescence upon 400 nm excitation.

CryoEM data collection and processing

MtClpP1 and MtClpP2 were incubated at equimolar amounts (7 uM) for 15 mins and concentrated with a spin filter to ~ 1 uM. After concentration, 100 uM fragment in 100% DMSO was incubated with the concentrated MtClpP1P2. After 10 min incubation, a 2.5 uL drop was applied to glow discharged gold holey carbon grid (R 1.2/1.3; Quantifoil), in which sample was then double blotted for 15 s each blot at 4°C and 100% humidity with Whatman No. 1 filter paper before being plunge frozen liquid ethane using a vitrobot (Thermo Fischer Scientific). The

sample was then imaged on a Titan Krios TEM (Thermo Fischer Scientific) operated at 300 keV and equipped with a Gatan BioQuantum imaging energy filter using a 20eV zero loss energy slit (Gatan Inc). Movies were acquired in super-resolution mode on a K3 direct electron detector (Gatan Inc.) at a calibrated magnification of 58,600X corresponding to a pixel size of 0.4265 Å/pixel. A defocus range of .8 to 1.2 µm was used with a total exposure time of 2 seconds fractionated into 0.2s subframes for a total dose of 68 e$^-$/Å^2 at a dose rate of 25 e$^-$/pixel/s. Movies were subsequently corrected for drift using MotionCor2 (10.1038/nmeth.4193) and were Fourier-cropped by a factor of 2 to a final pixel size of 0.853 Å/pixel[42].

A total of ~7200 micrographs were collected over a single collection. All the data-processing was performed in cryosparc3[43]. For particle picking, templates were generated from 100 particles, in which top and side views were selected. After inspecting the particles picked, approximately 770,00 particles were extracted. One round of 2D classification was performed to remove contamination and junk particles, which amounted to ~530,000 "good particles" or ~69% of the dataset. A five-class ab-initio reconstruction with no symmetry enforced was performed from the particle set and was used for initial classification steps.

To further separate active MtClpP1P2 complex from inactive conformations, heterogenous refinement with no symmetry enforced was performed with 4 different classes (Supplementary Fig 3.S1 a). The first class represents an inactive conformation where the N-terminal beta-hairpins are not in the upward conformation (66,000 particles, ~24%). Two classes with active conformations appeared identical and contained 166,000 particles, which is ~60% of the particles going into the refinement. Finally, the last class contained the active conformation (41,000 particles, ~15%), however, it lacked high resolution features and was not included for subsequent processing. The particles from classes 2 and 3 were combined and homogeneous

refinement with no symmetry enforced was performed, which resulted in a 3.2 Å resolution electron density map. To improve the resolution, homogeneous refinement with C7 symmetry enforced was performed, which resulted in a final resolution of 2.8 Å.

Molecular modeling

The starting model came from the MtClpP1P2 crystal structure bound to agonist and ADEP (pdb 4U0G) and for initial modeling, all ligands were removed from the model[22]. The initial model was a rigid body fit into the electron density map using UCSF ChimeraX's *fit-in map* tool[44,45]. Initial refinement was performed using Phenix with 1 round of simulated annealing and morphing and 5 rounds of real-space refinement that included minimization_global, rigid_body, adp, local_grid_search, secondary structural restraints, and non-crystallographic symmetry (NCS) restraints[46]. The ADEP fragment PDB and cif file were generated in Phenix eLBOW with the following SMILES code provided[47]:

O=C(OC[C@@H](C(OC)=O)NC([C@@H](NC(/C=C/CCCC)=O)CC1=CC(F)=CC(F)=C1)=O)[C@@H]2CCCN2C(C)=O. The fragment was fit into density using Phenix LigandFit, where the model for the asymmetric unit (1 MtClpP1 and 1 MtClpP2) and a zoned density map were used[48,49]. The parameters provided were 4 ligand copies in the asymmetric unit and 6 ligand conformers and the best fits were chosen. The model was then refined using real space refine in Coot to improve the model and validation parameters[50].

References

1.   Boucher, H.W. et al. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clin Infect Dis* **48**, 1-12 (2009).

2.   Cox, G. & Wright, G.D. Intrinsic antibiotic resistance: mechanisms, origins, challenges and solutions. *Int J Med Microbiol* **303**, 287-92 (2013).

3.   Brown, E.D. & Wright, G.D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336-43 (2016).

4.   Dorman, S.E. & Chaisson, R.E. From magic bullets back to the Magic Mountains: the rise of extensively drug-resistant tuberculosis. *Nat Med* **13**, 295-298 (2007).

5.   Koul, A., Arnoult, E., Lounis, N., Guillemont, J. & Andries, K. The challenge of new drug discovery for tuberculosis. *Nature* **469**, 483-90 (2011).

6.   Lienhardt, C. et al. Global tuberculosis control: lessons learnt and future prospects. *Nat Rev Microbiol* **10**, 407-16 (2012).

7.   Raju, R.M. et al. Mycobacterium tuberculosis ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. *PLoS Pathog* **8**, e1002511 (2012).

8.   Akopian, T. et al. The active ClpP protease from M. tuberculosis is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO J* **31**, 1529-41 (2012).

9.   Yu, A.Y. & Houry, W.A. ClpP: a distinctive family of cylindrical energy-dependent serine proteases. *FEBS Lett* **581**, 3749-57 (2007).

10.   Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S. & Berche, P. The ClpP serine protease is essential for the intracellularparasitism and virulence ofListeria monocytogenes. *Mol Microbiol* **35**, 1268-1294 (2000).

11. Kwon, H.Y. et al. Effect of heat shock and mutations in ClpL and ClpP on virulence gene expression in Streptococcus pneumoniae. *Infect Immun* **71**, 3757-65 (2003).

12. Kwon, H.Y. et al. The ClpP protease of Streptococcus pneumoniae modulates virulence gene expression and protects against fatal pneumococcal challenge. *Infect Immun* **72**, 5646-53 (2004).

13. Frees, D., Sorensen, K. & Ingmer, H. Global virulence regulation in Staphylococcus aureus: pinpointing the roles of ClpP and ClpX in the sar/agr regulatory network. *Infect Immun* **73**, 8100-8 (2005).

14. Ollinger, J., O'Malley, T., Kesicki, E.A., Odingo, J. & Parish, T. Validation of the essential ClpP protease in Mycobacterium tuberculosis as a novel drug target. *J Bacteriol* **194**, 663-8 (2012).

15. Leodolter, J., Warweg, J. & Weber-Ban, E. The Mycobacterium tuberculosis ClpP1P2 Protease Interacts Asymmetrically with Its ATPase Partners ClpX and ClpC1. *PLoS One* **10**, e0125345 (2015).

16. Schmitz, K.R. & Sauer, R.T. Substrate delivery by the AAA+ ClpX and ClpC1 unfoldases activates the mycobacterial ClpP1P2 peptidase. *Mol Microbiol* **93**, 617-28 (2014).

17. Fei, X., Bell, T.A., Barkow, S.R., Baker, T.A. & Sauer, R.T. Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. *Elife* **9**(2020).

18. Fei, X. et al. Structures of the ATP-fueled ClpXP proteolytic machine bound to protein substrate. *Elife* **9**(2020).

19. Lopez, K.E. et al. Conformational plasticity of the ClpAP AAA+ protease couples protein unfolding and proteolysis. *Nat Struct Mol Biol* **27**, 406-416 (2020).

20. 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**(2020).

21. Kim, Y. et al. Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase. *Nature* **8**, 230-233 (2001).

22. Schmitz, K.R., Carney, D.W., Sello, J.K. & Sauer, R.T. Crystal structure of Mycobacterium tuberculosis ClpP1P2 suggests a model for peptidase activation by AAA+ partner binding and substrate delivery. *Proc Natl Acad Sci U S A* **111**, E4587-95 (2014).

23. Binepal, G. et al. Development of Antibiotics That Dysregulate the Neisserial ClpP Protease. *ACS Infect Dis* **6**, 3224-3236 (2020).

24. Brotz-Oesterhelt, H. et al. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* **11**, 1082-7 (2005).

25. Moreno-Cinos, C. et al. ClpP Protease, a Promising Antimicrobial Target. *Int J Mol Sci* **20**(2019).

26. Lakemeyer, M. et al. Tailored Peptide Phenyl Esters Block ClpXP Proteolysis by an Unusual Breakdown into a Heptamer-Hexamer Assembly. *Angew Chem Int Ed Engl* **58**, 7127-7132 (2019).

27. Lavey, N.P., Coker, J.A., Ruben, E.A. & Duerfeldt, A.S. Sclerotiamide: The First Non-Peptide-Based Natural Product Activator of Bacterial Caseinolytic Protease P. *J Nat Prod* **79**, 1193-7 (2016).

28. Compton, C.L., Schmitz, K.R., Sauer, R.T. & Sello, J.K. Antibacterial activity of and resistance to small molecule inhibitors of the ClpP peptidase. *ACS Chem Biol* **8**, 2669-77 (2013).

29. Gersch, M. et al. AAA+ chaperones and acyldepsipeptides activate the ClpP protease via conformational control. *Nat Commun* **6**, 6320 (2015).

30. Kirstein, J. et al. The antibiotic ADEP reprogrammes ClpP,switching it from a regulated to anuncontrolled protease. *EMBO Mol Med* **1**, 37-49 (2008).

31. Famulla, K. et al. Acyldepsipeptide antibiotics kill mycobacteria by preventing the physiological functions of the ClpP1P2 protease. *Mol Microbiol* **101**, 194-209 (2016).

32. Lee, B.G. et al. Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat Struct Mol Biol* **17**, 471-8 (2010).

33. Vahidi, S. et al. An allosteric switch regulates Mycobacterium tuberculosis ClpP1P2 protease function as established by cryo-EM and methyl-TROSY NMR. *Proc Natl Acad Sci U S A* **117**, 5895-5906 (2020).

34. Carney, D.W. et al. A simple fragment of cyclic acyldepsipeptides is necessary and sufficient for ClpP activation and antibacterial activity. *Chembiochem* **15**, 2216-20 (2014).

35. Carney, D.W., Schmitz, K.R., Truong, J.V., Sauer, R.T. & Sello, J.K. Restriction of the conformational dynamics of the cyclic acyldepsipeptide antibiotics improves their antibacterial activity. *J Am Chem Soc* **136**, 1922-9 (2014).

36. Schmitz, K.R. et al. Acyldepsipeptide Antibiotics and a Bioactive Fragment Thereof Differentially Perturb Mycobacterium tuberculosis ClpXP1P2 Activity in vitro. *ACS Chem Biol* (2020).

37.   Jennings, L.D., Bohon, J., Chance, M.R. & Licht, S. The ClpP N-Terminus Coordinates Substrate Access with Protease Active Site Reactivity. *Biochemistry* **47**, 11031-11040 (2008).

38.   Li, M. et al. Structure and Functional Properties of the Active Form of the Proteolytic Complex, ClpP1P2, from Mycobacterium tuberculosis. *J Biol Chem* **291**, 7465-76 (2016).

39.   Zeiler, E. et al. Structural and functional insights into caseinolytic proteases reveal an unprecedented regulation principle of their catalytic triad. *Proc Natl Acad Sci U S A* **110**, 11302-7 (2013).

40.   Gersch, M., List, A., Groll, M. & Sieber, S.A. Insights into structural network responsible for oligomerization and activity of bacterial virulence regulator caseinolytic protease P (ClpP) protein. *J Biol Chem* **287**, 9484-94 (2012).

41.   Szyk, A. & Maurizi, M.R. Crystal structure at 1.9A of E. coli ClpP with a peptide covalently bound at the active site. *J Struct Biol* **156**, 165-74 (2006).

42.   Zheng, S.Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).

43.   Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).

44.   Goddard, T.D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).

45.   Pettersen, E.F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82 (2021).

46. Afonine, P.V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallographica Section D* **74**, 531-544 (2018).

47. Moriarty, N.W., Grosse-Kunstleve, R.W. & Adams, P.D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallographica Section D* **65**, 1074-1080 (2009).

48. Terwilliger, T.C., Klei, H., Adams, P.D., Moriarty, N.W. & Cohn, J.D. Automated ligand fitting by core-fragment fitting and extension into density. *Acta Crystallographica Section D* **62**, 915-922 (2006).

49. Terwilliger, T.C., Adams, P.D., Moriarty, N.W. & Cohn, J.D. Ligand identification using electron-density map correlations. *Acta Crystallographica Section D* **63**, 101-107 (2007).

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

Supplementary Information
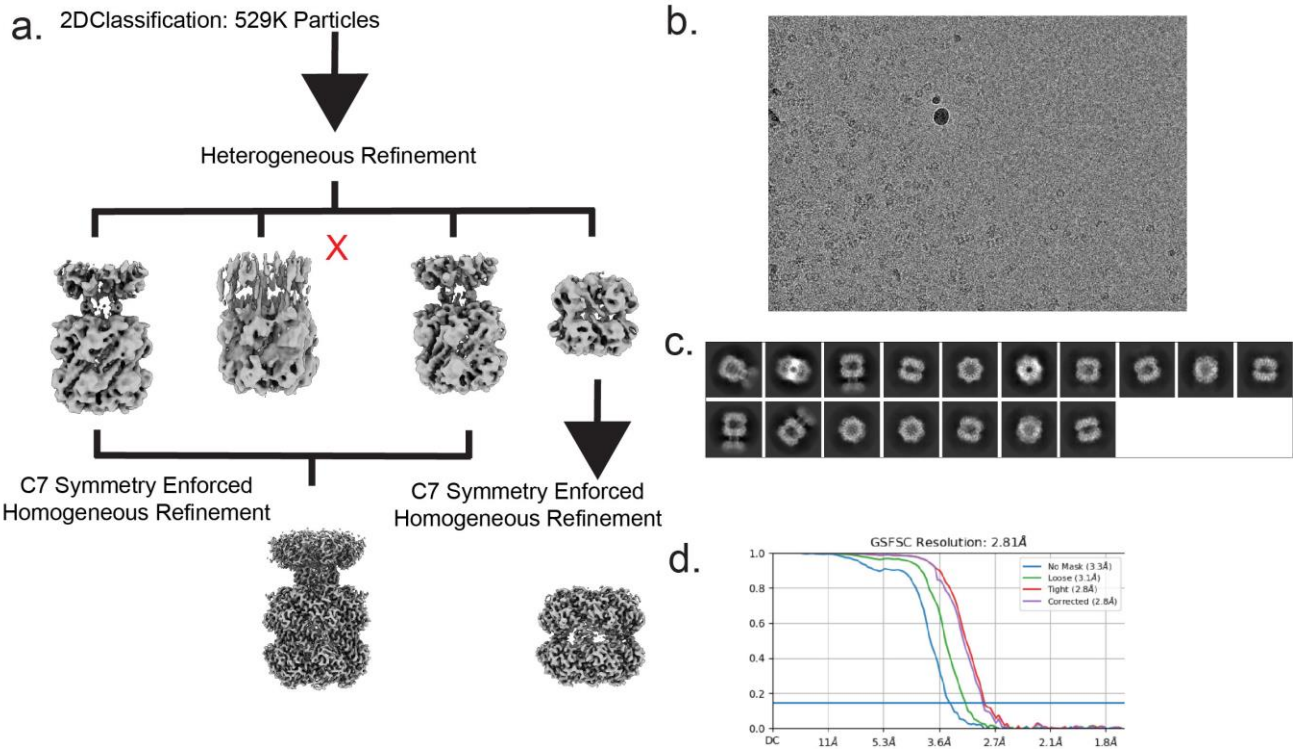
Supplemental Figure 3.S1 MtClpP1P2 cryoEM data collection and refinement
a 3D classification scheme to identify active, fragment-bound MtClpP1P2 structure. The red X indicates junk particles that were not used for further refinement. **b** Representative micrograph for dataset for reported structure. **c** Reference-free 2D class averages of the dataset used for processing. **d** Gold standard FSC-curves for the C7 refinement of MtClpP1P2 generated in cryosparc3.
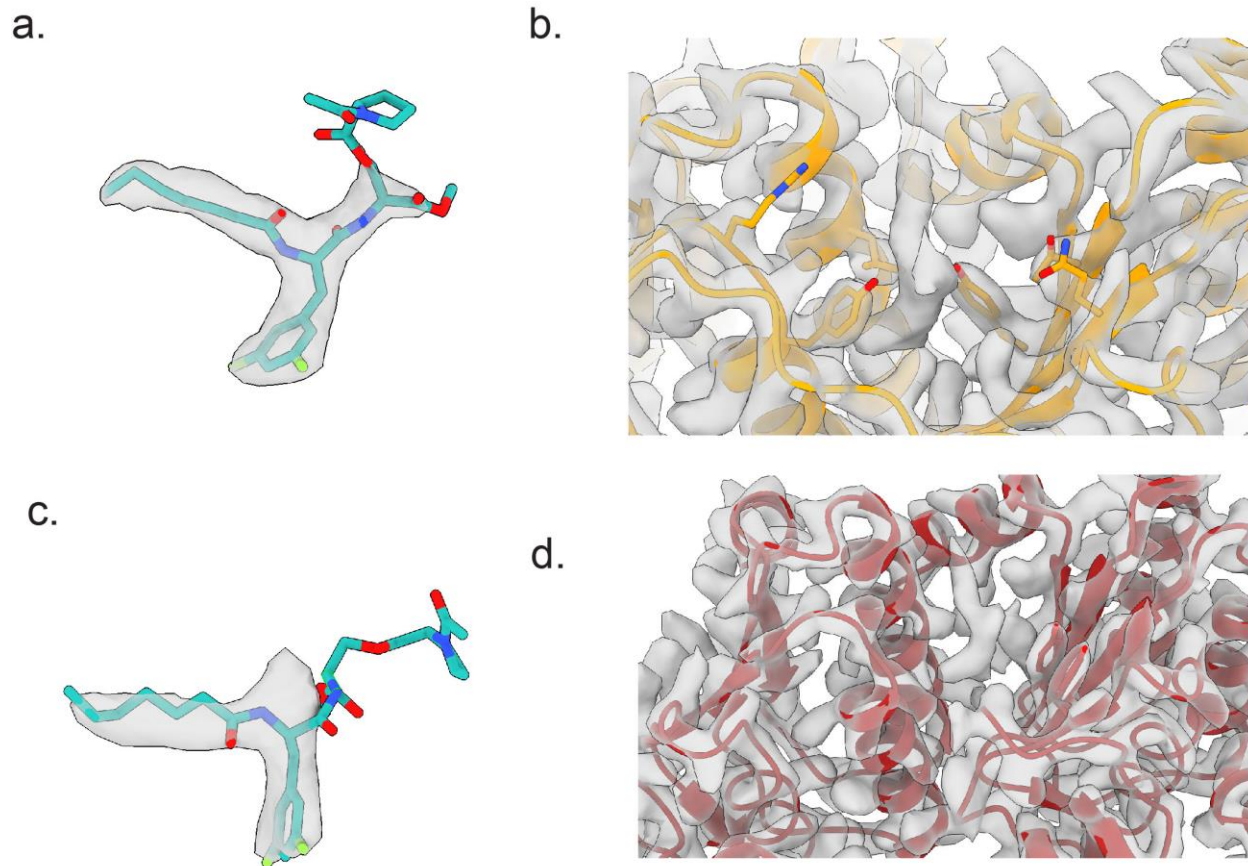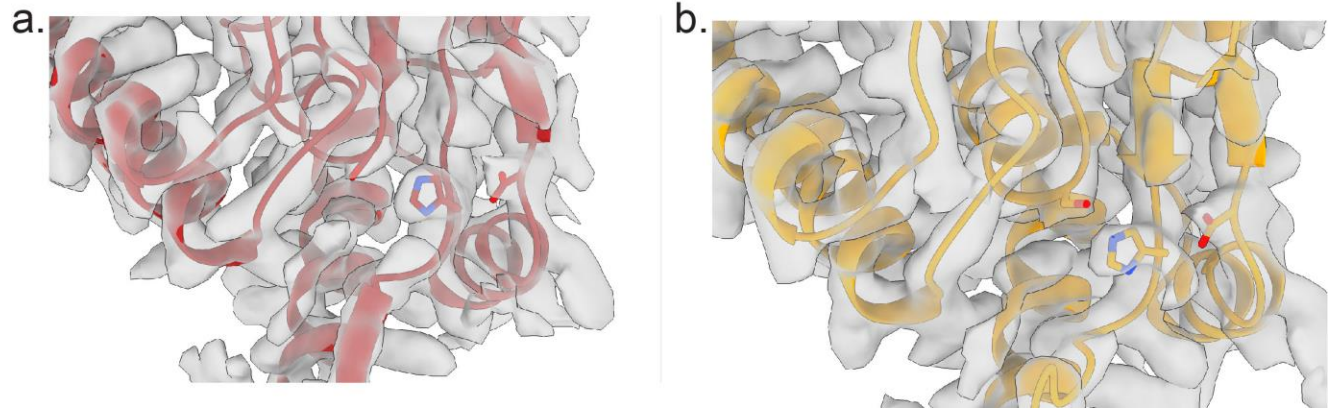
Supplemental Figure 3.S2 Fragment and MtClpP1P2 hydrophobic cleft density
a Density for the fragment in the MtClpP2 pocket and the MtClpP1 pocket (**c**). **b** Density for the MtClpP2 and MtClpP1 (**d**) hydrophobic cleft and with relevant residues represented as sticks. The fragment model is not fit into the density here.
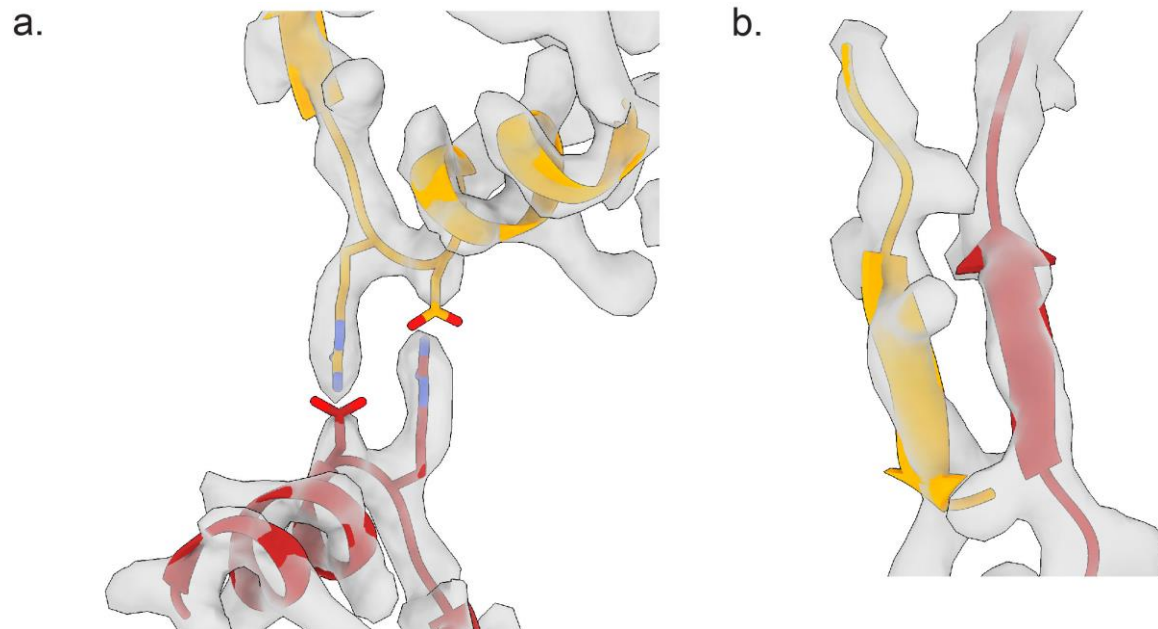
Supplemental Figure 3.S3 CryoEM density for MtClpP1P2 oligomerization sensor and handle region
a Density for the MtClpP1P2 handle region with relevant residues represented as sticks. **b** Density for the MtClpP1P2 beta handle region.

Supplemental Figure 3.S4 CryoEM density for MtClpP1P2 active sites
CryoEM density for the MtClpP1 (**a**) and MtClpP2 (**b**) active sites with the catalytic triad represented by sticks.
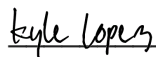
Acknowledgments

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

Kyle Lopez

—5129A2E83AA041A...   Author Signature

5/31/2022

Date

112