

## BIOCHEMISTRY

# Mechanism of the allosteric activation of the ClpP protease machinery by substrates and active-site inhibitors

Jan Felix<sup>1</sup>, Katharina Weinhäupl<sup>1</sup>, Christophe Chipot<sup>2,3,4</sup>, François Dehez<sup>2,3</sup>, Audrey Hessel<sup>1</sup>, Diego F. Gauto<sup>1</sup>, Cecile Morlot<sup>1</sup>, Olga Abian<sup>5,6,7,8</sup>, Irina Gutsche<sup>1</sup>, Adrian Velazquez-Campoy<sup>5,6,7,9</sup>, Paul Schanda<sup>1\*</sup>, Hugo Fraga<sup>1,10,11\*</sup>

Coordinated conformational transitions in oligomeric enzymatic complexes modulate function in response to substrates and play a crucial role in enzyme inhibition and activation. Caseinolytic protease (ClpP) is a tetradecameric complex, which has emerged as a drug target against multiple pathogenic bacteria. Activation of different ClpPs by inhibitors has been independently reported from drug development efforts, but no rationale for inhibitor-induced activation has been hitherto proposed. Using an integrated approach that includes x-ray crystallography, solid- and solution-state nuclear magnetic resonance, molecular dynamics simulations, and isothermal titration calorimetry, we show that the proteasome inhibitor bortezomib binds to the ClpP active-site serine, mimicking a peptide substrate, and induces a concerted allosteric activation of the complex. The bortezomib-activated conformation also exhibits a higher affinity for its cognate unfoldase ClpX. We propose a universal allosteric mechanism, where substrate binding to a single subunit locks ClpP into an active conformation optimized for chaperone association and protein processive degradation.

## INTRODUCTION

Among the most exciting antibacterial drug targets that have emerged in the past decade is the caseinolytic protease (ClpP) (1). While mostly studied in *Escherichia coli* (EcClpP), ClpP is present in a wide range of bacteria, as well as in mitochondria and chloroplasts. Functionally, ClpP can be defined as a cylindrical canonical serine protease composed of two stacked rings, each formed by seven subunits that encompass a large (300 kDa) proteolytic core (2). This architecture restricts access to the 14 proteolytic active sites located inside the barrel and is key for ClpP function. Although ClpP is able to rapidly hydrolyze peptides, degradation of large proteins requires the presence of an AAA+ adenosine triphosphatase (ATPase) complex, such as ClpA or ClpX in *E. coli* or ClpC in other species. These unfoldases provide substrate specificity not only by recognizing, unfolding, and translocating their globular protein substrates but also by opening the apical gate region of ClpP, which is otherwise blocked by flexible N-terminal tails.

ClpP protease and its associated chaperones are essential for the survival or virulence of several bacteria including pathogens such as *Mycobacterium tuberculosis* (*Mtb*) (3), *Staphylococcus aureus* (*Sa*) (4), and *Listeria monocytogenes* (*Lm*) (5). In addition, human mito-

chondrial ClpP was associated to human acute myeloid leukemia and obesity (6, 7). Recently, ClpP from *Plasmodium falciparum* (PfClpP), the parasite responsible for malaria infection, has also been proposed as a promising drug target for the control of this disease (8).

The fact that ClpP is a suitable Achilles' heel for these organisms is further reinforced by the recent discovery of natural antibiotics that target ClpP or its co-chaperones. Efficient bacterial killing could be accomplished either by opening the axial pore of ClpP and activating unregulated cell proteolysis (9), by blocking ClpP-ATPase interaction (10), or by specifically inhibiting its chaperones, ClpX (4) or ClpC1 (11).

Despite recent advances in the field, the study and development of drugs targeting ClpP have been thwarted by the fact that ClpPs from different species show notable functional differences. For example, *Mtb* contains two clpP genes, *clpP1* and *clpP2*, both of which are essential for viability and infectivity (1, 3). Although both genes encode serine proteases that form heptameric rings, first attempts to express and characterize isolated *Mtb*ClpP1 and *Mtb*ClpP2 in *E. coli* yielded complexes that lacked proteolytic activity (12, 13). The study of *Mtb*ClpP1P2 was not possible until the finding that a small activator, derived from classical serine protease inhibitor peptide aldehydes (3), was required to promote the formation of an active complex composed of one heptameric ring from ClpP1 and another from ClpP2 (3). Oddly enough, the small activators were found at the protein active site, and it is not clear how their presence at the active site can promote activation rather than classic competitive inhibition (14, 15). Furthermore, several structures obtained by x-ray crystallography and solution nuclear magnetic resonance (NMR) studies have suggested that ClpPs, while keeping a tetradecamer organization, can adopt several conformations (16, 17). Currently, at least three conformations have been reported: an extended state (catalytically active), a compressed state (catalytically inactive), and a compact state (catalytically inactive) (17–19). In addition, while some ClpPs are inherently active, others are purified as non-active forms (3, 12, 13, 15, 20–22).

Here, we report a study of ClpP from *Thermus thermophilus* (TtClpP), which is, as described for other ClpPs, purified in an inactive

Copyright © 2019  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
NonCommercial  
License 4.0 (CC BY-NC).

<sup>1</sup>Institut de Biologie Structurale, Université Grenoble Alpes, CEA, CNRS, IBS, 71 Avenue des Martyrs, F-38044 Grenoble, France. <sup>2</sup>LPCT, UMR 7019 Université de Lorraine CNRS, Vandoeuvre-les-Nancy F-54500, France. <sup>3</sup>Laboratoire International Associé CNRS and University of Illinois at Urbana-Champaign, Vandoeuvre-les-Nancy F-54506, France. <sup>4</sup>Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, IL 61801, USA. <sup>5</sup>Institute of Biocomputation and Physics of Complex Systems (BIFI), Joint Units IQFR-CSIC-BIFI and GBsC-CSIC-BIFI, and Department of Biochemistry and Molecular and Cell Biology, Universidad de Zaragoza, 50018 Zaragoza, Spain. <sup>6</sup>Aragon Institute for Health Research (IIS Aragon), 50009 Zaragoza, Spain. <sup>7</sup>Biomedical Research Networking Centre for Liver and Digestive Diseases (CIBERehd), Madrid, Spain. <sup>8</sup>Aragon Health Sciences Institute (IACS), 50009 Zaragoza, Spain. <sup>9</sup>Fundacion ARAID, Government of Aragon, 50018 Zaragoza, Spain. <sup>10</sup>Departamento de Biomedicina, Faculdade de Medicina da Universidade do Porto, Porto, Portugal. <sup>11</sup>I3S, Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal.

\*Corresponding author. Email: hfraga@med.up.pt (H.F.); paul.schanda@ibs.fr (P.S.)

conformation but can be in vitro activated by small molecules. Using an integrated approach including multiple biochemical assays, x-ray crystallography, magic angle spinning (MAS) NMR and solution-state NMR, isothermal titration calorimetry (ITC), and molecular dynamics (MD) simulations, we show that bortezomib, a boronic acid previously identified as an inhibitor of *Mtb*ClpP1P2, induces a concerted conformational change in ClpP, resulting in complex activation. On the basis of our data, and building further upon previously established results, we propose a universal mechanism for ClpP activation.

## RESULTS

### Bortezomib in low concentrations activates *Tt*ClpP peptidase activity

Although cellular protein quality control is likely to be particularly challenging for organisms living at high temperatures, ClpP from extremophiles remains modestly documented. *T. thermophilus* (*Tt*) is a Gram-negative eubacterium used in a wide range of biotechnological applications and as a model organism for genetic manipulation, structural genomics, and systems biology. The bacterium has an optimal growth temperature of about 65°C. We obtained *Tt*ClpP by recombinant expression and purified it as a canonical tetradecamer of around 300 kDa, as monitored by size exclusion chromatography (SEC) and multi-angle laser light scattering (MALLS) (fig. S1). This tetradecamer could be reorganized into a smaller heptameric species when ammonium sulfate (300 mM) was included in the elution buffer, likely as a result of the disruption of inter-ring salt bridges (fig. S1A). As anticipated, *Tt*ClpP displayed enhanced thermostability compared to previously studied ClpPs. Using the single tryptophan of the *Tt*ClpP monomers (14 per tetradecamer) as a probe, the melting point ( $T_M$ ) of ClpP was determined as 84°C (fig. S1C), 30°C above the value reported for *S. aureus* ClpP (*Sa*ClpP) (23). The increased stability is also reflected in an elevated optimal catalytic temperature. When we measured the degradation of the green fluorescent protein GFP<sub>srrA</sub> by *Tt*ClpP in association with the cognate *Tt*ClpX ATPase, we found maximal rates at 60°C (the highest temperature we could achieve with the fluorimeter equipment; fig. S2, A and B). However, although *Tt*ClpP is competent to degrade folded substrates in association with *Tt*ClpX, it is a rather inefficient peptidase even at 60°C. In comparison to *Ec*ClpP at 37°C, it has a 200-fold lower specific activity in degradation experiments with the short tripeptide substrate PKMamc (proline-lysine-methionine-7-amino-4-methylcoumarin; at a working concentration of 100 μM).

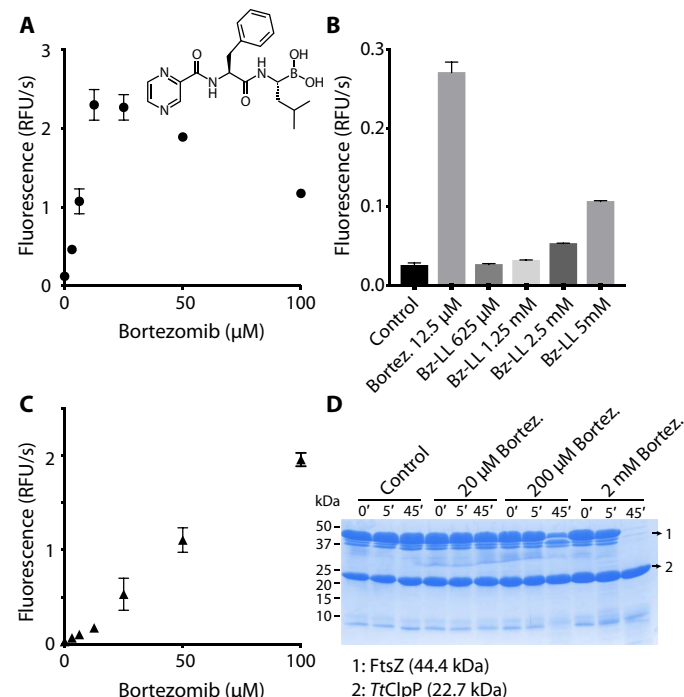
During attempts to identify inhibitors of *Tt*ClpP that could be used for further structural biology studies, we observed that bortezomib did not inhibit, but actually stimulated, *Tt*ClpP peptidase activity at concentrations from 3 to 100 μM (Fig. 1A and fig. S3A). Bortezomib is an N-protected dipeptide with a boronic acid instead of a carboxylic acid that forms adducts with activated threonines or serines (24). It was developed as a proteasome inhibitor targeting specifically the chymotrypsin-like sites of this complex with nanomolar affinity while exhibiting limited reactivity to the other active sites of the complex. We selected this molecule because it was recently identified in an in vivo cell screen as an inhibitor of GFP<sub>srrA</sub> degradation by the *Mtb*ClpXP1P2 complex and was proposed as a potential antituberculosis candidate (25). The bortezomib dependence of the *Tt*ClpP peptidase activity displays a bell shape, with maximum activation at 12.5 μM (about 10- to 20-fold activation), while higher drug concentrations reduced the enzymatic activity (Fig. 1A). A time-lag phase for the activation of *Tt*ClpP was consistently observed and likely resulted from the neces-

sary fluorimeter temperature equilibrium because the assays were executed at 45° to 60°C. Bortezomib activation derives from intrinsic properties of *Tt*ClpP, and when tested against *Ec*ClpP peptidase activity, it behaved as a powerful inhibitor with an apparent  $IC_{50}$  (median inhibitory concentration) of 1.6 μM (fig. S3B).

Although paradoxical, activation of some ClpP complexes by inhibitors has already been reported (3, 22, 26–28). Akopian and colleagues (3) have shown that the *Mtb*ClpP complex (*Mtb*ClpP1P2), otherwise inactive, is activated by small peptide aldehydes. The most potent of these activators in *Mtb*, an N-terminally blocked dileucine (Bz-LL), was also capable to activate *Tt*ClpP, albeit with lower efficiency (Fig. 1B). Conversely, bortezomib could also activate the peptidase activity of the *Mtb*ClpP1P2 complex. We applied a set of biochemical, biophysical, and structural studies to understand the mechanistic and functional foundation of this inhibitor-based activation.

### Bortezomib activates *Tt*ClpP unfolded protein degradation

After establishing the bortezomib-induced activation of *Tt*ClpP for degradation of small peptides, we investigated its potential to stimulate protein degradation. Contrary to small peptides, which can



**Fig. 1. Bortezomib activates *Tt*ClpP for peptide and intrinsically disordered protein degradation.** (A) *Tt*ClpP (1 μM complex) peptidase activity was measured with the substrate PKMamc (100 μM) in the presence of bortezomib at the indicated concentrations. As the peptide is cleaved, 7-amino-4-methylcoumarin is released, resulting in an increase in the measured fluorescence. Initial rates are plotted as a function of the bortezomib concentration. RFU, relative fluorescence units. (B) The activating effect of bortezomib is compared with the one observed with Bz-LL, a previously described activator of *Mtb*ClpP1P2. (C) The degradation of the unfolded protein substrate FITC-casein by *Tt*ClpP (0.1 μM complex) was measured in the presence of the indicated bortezomib concentrations. Initial degradation rates following temperature equilibration were plotted as a function of the bortezomib concentration. (D) *Tt*ClpP (1 μM complex) can degrade the intrinsically disordered *E. coli* FtsZ. The degradation of FtsZ by *Tt*ClpP was monitored by SDS-gel electrophoresis. While no degradation was observed with apo*Tt*ClpP, degradation of FtsZ was observed in the presence of 200 μM and 2 mM bortezomib.

independently diffuse into the proteolytic chamber, the basal degradation of proteins by ClpP is low because the flexible N-terminal tails of ClpP act as a barrier to protein entry into the chamber. Our assay used as proteolytic target fluorescein isothiocyanate-labeled casein (FITC-casein), a proline-rich protein lacking stable secondary structures. Protease-catalyzed hydrolysis of FITC-casein leads to highly fluorescent dye-labeled peptides. Similarly to the peptidase activity, addition of bortezomib did not result in inhibition of FITC-casein degradation, but instead, a striking activation was observed (Fig. 1C and fig. S4A). Furthermore, contrary to what was observed with the peptidase activity, concentration dependence of this effect was linear and no reduction of the activation was observed up to 100  $\mu\text{M}$  (Fig. 1C). In addition to casein, we tested whether bortezomib could also promote the degradation of other substrates. The bacterial guanosine triphosphatase (GTPase) FtsZ forms a cytokinetic ring at mid-cell, recruits the division machinery, and coordinates membrane and peptidoglycan cell wall invagination. An FtsZ monomer contains an intrinsically disordered C-terminal linker, and the degradation of FtsZ by ClpP activated by the natural antibiotic acyldepsipeptides (ADEPs) has been proposed to trigger growth inhibition (29). We therefore tested the effect of bortezomib ClpP-catalyzed proteolysis on FtsZ and found that FtsZ was, even at low temperature for *Tt*ClpP (37°C), degraded by ClpP when in the presence of bortezomib (Fig. 1D) (29). These experiments establish that bortezomib acts as a *Tt*ClpP activator for the degradation of different disordered protein substrates or small peptides.

### ClpX enhances peptidase activity of *Tt*ClpP

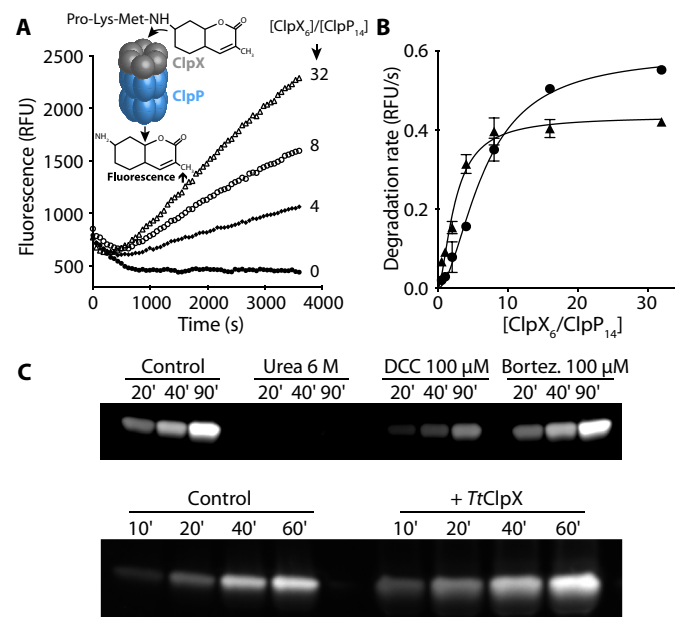
Because ClpP acts in association with specific chaperones in vivo (ClpX, ClpA, or ClpC), we proceeded by assessing the effect of *Tt*ClpX association on *Tt*ClpP peptidase activity. When we measured the degradation of PKMamc by *Tt*ClpP in the presence of *Tt*ClpX, we observed a marked increase in peptidase activity that was dependent on *Tt*ClpX concentration (Fig. 2, A and B). Similarly to *Tt*ClpP activation by bortezomib, a lag phase was observed, which could result either from temperature equilibrium or from oligomerization of *Tt*ClpX hexamers from monomers in the presence of adenosine 5'-triphosphate (ATP).

Because bortezomib and ClpX are both able to activate *Tt*ClpP, we tested whether the effects of bortezomib and ClpX were synergistic. Bortezomib (10  $\mu\text{M}$ ) did not induce further activation but instead acted as an inhibitor (fig. S4B) in the presence of ClpX. However, in the presence of bortezomib, a marked feature of the ClpX activation curve was a certain level of cooperativity, i.e., less *Tt*ClpX concentration was required to achieve maximal activity, suggesting that bortezomib binding can modulate the affinity of ClpP for ClpX (Fig. 2B). In the presence of bortezomib, half-maximal activation was achieved with a *Tt*ClpX<sub>6</sub>/ClpP<sub>14</sub> ratio of 2.4, corresponding to 120 nM *Tt*ClpX<sub>6</sub>. This value for *Tt*ClpX-*Tt*ClpP association is on the same order of magnitude as the one reported for *E. coli* using an indirect ATPase assay (27).

The large activation of ClpP-catalyzed degradation of small peptides by ClpX was unexpected, because auxiliary chaperones are required for unfolding, translocation, and opening of the N-terminal pores facilitating entry of protein substrates into the ClpP proteolytic chamber, while small peptides can just freely diffuse into the proteolytic chamber. The hydrolysis rate of small peptides (dipeptides) in *Ec*ClpP has been independently shown to be unaffected by ClpX association (30, 31). Activation by *Ec*ClpX is, however, essential for bigger peptides and proteins, and ClpX activation increases as a function of peptide molecular weight. ClpX binding has also been shown to stimulate active-site modification of ClpP by fluorophosphates, but only to the level predicted from faster

diffusion of the inhibitor into the ClpP chamber (30). Nonetheless, it was recently proposed that, in addition to the opening of the ClpP gate, chaperones can also allosterically activate the ClpP catalytic sites by inducing a shift of ClpP from the inactive compressed state, in which the catalytic triad is distorted, to the extended state, in which the catalytic residues are positioned for enzymatic action (23).

Because the small molecular weight of the substrate used in our assays is not an obstacle to its free diffusion into the proteolytic chamber (molecular weight of 574 Da) (30), we tested whether *Tt*ClpX binding to *Tt*ClpP could affect the protease active sites. To confirm the effect of co-chaperone or bortezomib on the active site, we used an assay based on fluorophosphonate (FP) conjugated tetramethylrhodamine (TAMRA-FP, molecular weight of 680 Da), which specifically and covalently labels activated Ser active sites with a fluorescent dye. TAMRA-FP can be used to quantify the reactivity of a given protease active site, e.g., if the catalytic triad is in the correct geometry for proper serine activation. The presence of an activated Ser can thus be detected using fluorescent gel imaging. We incubated *Tt*ClpP and TAMRA-FP with either urea, bortezomib, dichloroisocoumarin (DCC; an inhibitor of serine proteases), or *Tt*ClpX. While urea (by unfolding the complex) and DCC (by irreversibly blocking the active site) reduced the amount of TAMRA bound to ClpP, ClpX association resulted in a significant increase of active-site acetylation—a clear indication that ClpX association activates the ClpP active site. Although bortezomib also leads to an activation of ClpP, we failed to observe any increase of TAMRA-FP labeling (Fig. 2C). This could result from competition between the two inhibitors



**Fig. 2. *Tt*ClpX activates *Tt*ClpP peptidase activity.** (A) *Tt*ClpP (0.1  $\mu\text{M}$  complex) peptidase activity was measured in the presence of *Tt*ClpX at the indicated ClpX<sub>6</sub>/ClpP<sub>14</sub> ratios. (B) The activation of *Tt*ClpP peptidase activity in the absence (circles) or presence (triangles) of bortezomib (10  $\mu\text{M}$ ) is plotted as a function of the molar ClpX<sub>6</sub>/ClpP<sub>14</sub> ratio. Half-maximal peptidase activity was obtained at  $6.5 \pm 0.5$  and  $2.4 \pm 0.3$  ClpX<sub>6</sub>/ClpP<sub>14</sub> ratio for the apo and bortezomib-loaded *Tt*ClpP, respectively. (C) Activation of the *Tt*ClpP active site by bortezomib (100  $\mu\text{M}$ ) and *Tt*ClpX (1  $\mu\text{M}$ ), monitored by labeling with TAMRA-FP. TAMRA-FP (5  $\mu\text{M}$ ) was incubated with *Tt*ClpP (1 mg/ml), and aliquots of the reaction mixture were removed at the indicated time points. Labeling of the active-site serines is reflected by the increase in fluorescence observed.



(bortezomib and TAMRA-FP). Therefore, we cannot exclude that TAMRA-FP binding also results in *TtClpP* activation. However, contrary to bortezomib, TAMRA-FP forms irreversible adducts with *TtClpP*, thereby preventing activity measurements.

### Bortezomib binds to the serine of ClpP active site

To understand bortezomib's mechanism of action, we examined its binding site using NMR spectroscopy. With a molecular weight of 300 kDa, NMR studies of *TtClpP* are challenging. In solution, the slow tumbling leads to severe line broadening, to the extent that standard solution NMR studies are not feasible. The most sensitive route for observing solution NMR signals of proteins of the size of ClpP is the specific labeling of methyl groups in an otherwise deuterated background (32). Although being a very powerful approach for probing interactions and dynamics, methyl-directed NMR is able to observe only the subset of methyl-bearing amino acid types, and the sequence-specific assignment of NMR resonances is challenging, generally requiring multiple mutants.

MAS solid-state NMR does not suffer from this inherent size limitation and is able to detect all amino acid types. We have used samples of *TtClpP* that were either obtained from sedimenting ClpP from solution into MAS NMR sample tubes (1.3-mm rotors) or obtained by addition of methyl-pentane-diol, an oft-used precipitant in crystallography. Spectra acquired from these two preparations are very similar. We have used proton-detected three-dimensional (3D) and 4D MAS NMR approaches (33) to obtain the sequence-specific resonance assignment for ( $^1\text{H}^{\text{N}}$ ,  $^{15}\text{N}$ ,  $^{13}\text{C}\alpha$ ,  $^{13}\text{CO}$ ) nuclei of 97 residues spread throughout the molecule, including the active site and the central helix  $\alpha\text{E}$ , which undergoes most changes in the compressed-to-extended transition. The N-terminal gate region was not assigned, presumably due to its structural heterogeneity and dynamics, which lead to inefficient NMR coherence transfers and line broadening (16). Figure 3A shows the residue-wise secondary structure obtained from these MAS NMR resonance assignments. As expected, the location of secondary structures matches well the information inferred from crystallography (see below).

3D spectra correlating intra-residue  $^1\text{H}$ - $^{13}\text{C}\alpha$ - $^{15}\text{N}$  frequencies (referred to as hCANH) of sedimented *TtClpP* with and without bortezomib are reported in Fig. 3 (B and C). The observed changes are modest, which is likely linked to the relatively low affinity of bortezomib, making it difficult to saturate the protein in the highly concentrated sedimented sample. Nevertheless, chemical shift perturbations (CSPs) are observed for a number of residues in the vicinity of the active site, including Val<sup>70</sup>, Gly<sup>73</sup>, Tyr<sup>77</sup>, and the catalytic Ser<sup>97</sup>, as well as residues in the neighboring helix  $\alpha\text{E}$  (Fig. 3D). These data reveal that bortezomib interacts with the active site, and suggest that the interaction induces also small changes in the structure and the dynamics of helix  $\alpha\text{E}$ , which is involved in mediating the compressed-to-extended transition.

To independently confirm the direct involvement of the active site for bortezomib binding in solution, we have performed additional NMR experiments, making use of perdeuterated, Ile- $\delta$ 1- $\text{CH}_3$ -labeled *TtClpP*. Spectra of apo *TtClpP*, and of the latter in the presence of 2 mM bortezomib, are shown in Fig. 3E (left). Although we did not attempt to assign the resonances of the 18 isoleucines per ClpP subunit, the clear differences in the spectra without or with bortezomib unambiguously confirms binding. In a control experiment performed with the catalytic S97A mutant, no spectral changes were observed, corroborating that the bortezomib interaction requires the presence of the catalytic serine (Fig. 3E, right). The absence of any substantial

CSP in the S97A mutant demonstrates that all spectral changes are a result of the direct binding of bortezomib to the active site and its associated conformational changes—not from nonspecific binding to other parts of the protein.

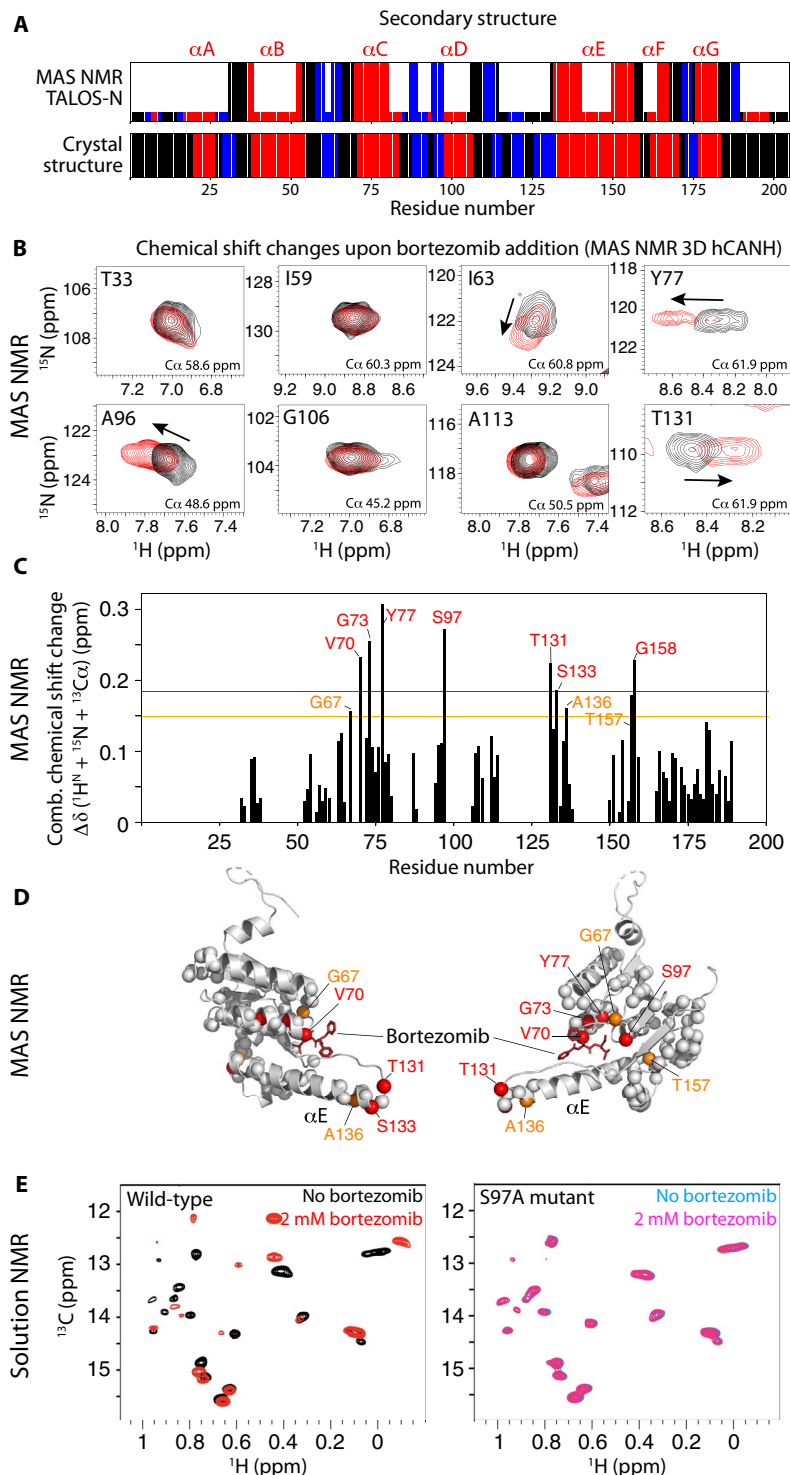
Together, MAS NMR and solution NMR provide direct evidence that bortezomib binds to the catalytic site. CSPs outside the active site (particularly in helix  $\alpha\text{E}$ ) suggest that binding may lead to further structural changes, possibly connected to the observed *TtClpP* activation.

### Structural basis for bortezomib activation

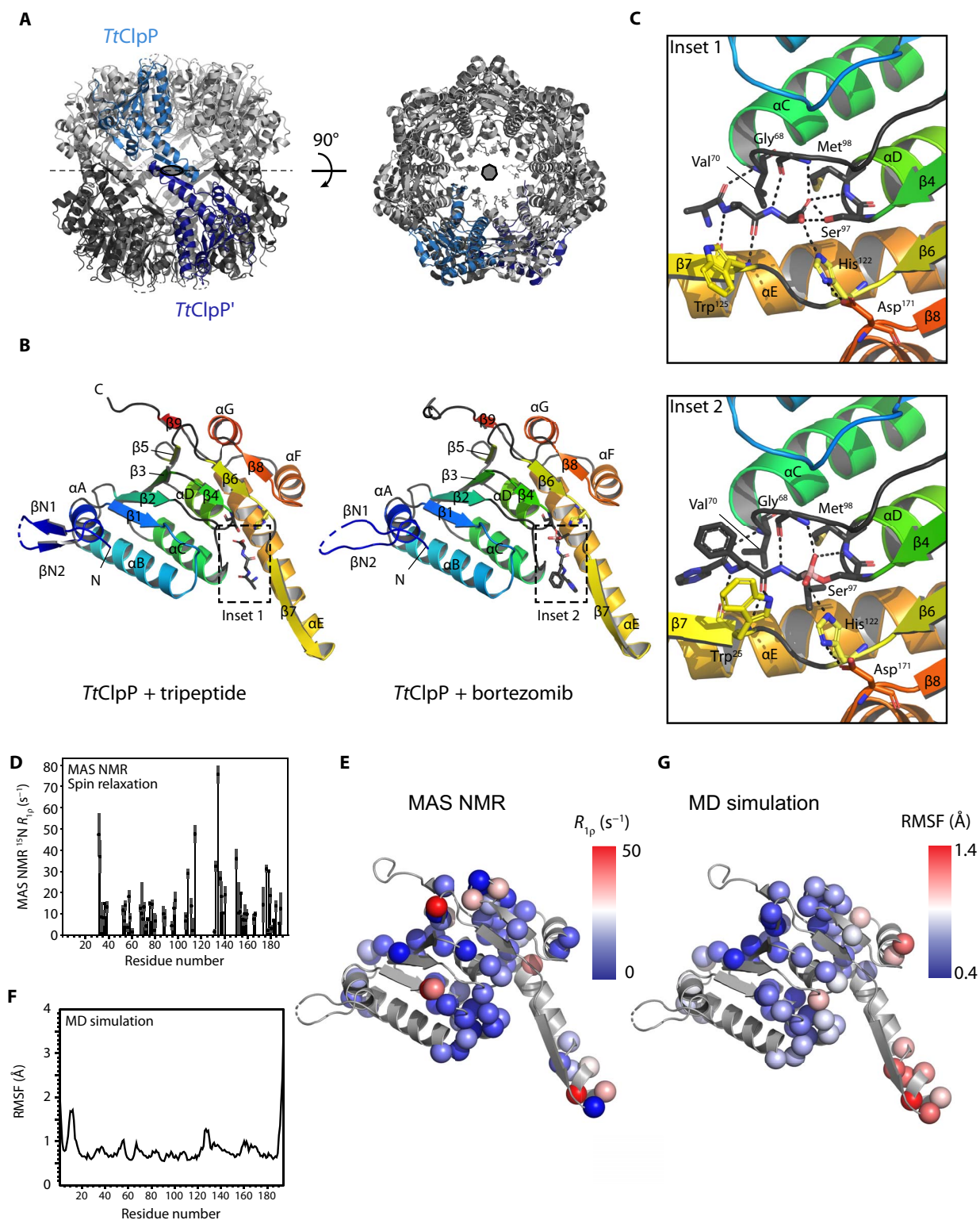
To understand the structural details of the *TtClpP*-bortezomib interaction initially characterized by NMR, we used x-ray crystallography and determined the crystal structures of *TtClpP* in the absence and presence of bortezomib at a resolution of 1.95 and 2.7 Å, respectively (Fig. 4).

The overall structure of *TtClpP* is essentially identical to the “extended” state (34–37) of ClpP described for other species [root mean square deviation (RMSD) to *EcClpP* tetradecamer [Protein Data Bank (PDB): 6MT6] = 2.115 Å for 1451 aligned C $\alpha$  atoms and RMSD to *EcClpP* monomer = 0.73 Å for 176 aligned C $\alpha$  atoms], which is in agreement with its high sequence similarity to ClpP orthologs (% identity to *EcClpP* = 58.25). The structure of one *TtClpP* monomer is presented in Fig. 4B and depicts an  $\alpha/\beta$ -fold composed of two domains: a head domain comprising six repeats of  $\alpha/\beta$  units (helices A to D, F, and G and strands 1 to 5, 7, and 8) and a handle domain (helix E and strand 6). The N terminus of *TtClpP* forms a short antiparallel  $\beta$ -hairpin that protrudes from the apical surface of the barrel. This N-terminal  $\beta$ -hairpin is constituted of residues 5 to 8 and residues 15 to 18 of *TtClpP*, with missing electron density for residues 9 to 14. While the overall monomer structure of ClpP is highly conserved among the structures solved to date, a high degree of variability exists in the conformation of the N-terminal regions. Several ClpP crystal structures feature N-terminal  $\beta$ -hairpin structures of varying lengths and degrees of flexibility, some of which appear to be stabilized by crystal-packing contacts (38). Solution NMR studies have demonstrated that the N termini of ClpP are structurally heterogeneous and adopt multiple conformations (16, 39), which might explain the missing density near the end of the N-terminal  $\beta$ -hairpins in the *TtClpP* structure.

Unexpectedly, in the *TtClpP* crystals without added bortezomib, we consistently observe a continuous stretch of electron density near the active-site serine (Ser<sup>97</sup>) of each *TtClpP* monomer, which, based on its shape, suggests the presence of a tripeptide (Fig. 4C and fig. S5). Because no peptide was added, the nature of this molecule was unknown. We, therefore, built it in the electron density as a tri-L-alanine (see also omit and polder maps presented in fig. S5). The presence of short peptides in *Helicobacter pylori* ClpP structures resulting from proteolysis during crystallization has been previously described (40). While in that particular case the observed tri- and tetrapeptides are caused by proteolysis of added heptapeptides before crystallization, the origin of the tripeptides observed in our *TtClpP* crystals remains unclear. Matrix-assisted laser desorption/ionization–time-of-flight (MALDI-TOF) measurements exclude the presence of these peptides in fresh purified *TtClpP*. We speculate that they are the result of autoprolysis of unfolded or partially unfolded *TtClpP* in the crystallization drops followed by the formation of *TtClpP*:peptide crystals. The serendipitous presence of peptides near the active-site Ser<sup>97</sup> provides valuable insights into how *TtClpP* binds its substrates (Fig. 4C, inset 1). Apart from the active-site Ser<sup>97</sup>, the terminal carboxy group of the peptide (residue P3) is further stabilized by hydrogen bonds



**Fig. 3. Bortezomib binding to TtClpP investigated by MAS NMR and solution NMR.** (A) Secondary structures (loop, black;  $\alpha$ -helix, red;  $\beta$ -strand, blue) as a function of sequence number obtained from the resonance assignments obtained by MAS NMR and analysis with TALOS-N (24) (tall bars). For residues for which no assignment has been obtained, the small bars are obtained by TALOS-N from a database approach. The bottom panel shows the comparison with the secondary structure obtained from the crystal structure (see Fig. 5). (B) Zooms onto 2D  $^1\text{H}$ - $^{15}\text{N}$  excerpts from 3D hCANH MAS NMR correlation spectra of TtClpP sedimented in the presence of 10 mM bortezomib (red) or without bortezomib (black). Chemical shift changes are indicated with arrows. Note that the CSPs are relatively small, presumably due to the difficulty of saturating the binding site in a sedimented sample, given the high  $K_d$  value, resulting in incompletely occupied binding sites. ppm, parts per million. (C) Combined CSPs upon addition of bortezomib from the added peak shifts in  $^1\text{H}$ ,  $^{13}\text{C}\alpha$ , and  $^{15}\text{N}$  dimensions obtained from the spectra shown in (B). (D) Location of residues with substantial CSPs mapped onto the x-ray structure of the TtClpP:bortezomib complex obtained from the data shown in (C). (E) Solution-state NMR spectra of wild-type (left) and S97A (right) TtClpP in the absence and presence of bortezomib, showing large spectra changes for the former due to binding, but no effect on the mutant.



**Fig. 4. Structure and dynamics of *TtClpP* from x-ray crystallography, MAS NMR, and MD simulations.** (A) Side and top views of the *TtClpP* 14-mer. One *TtClpP* monomer per heptameric ring (light and dark gray) is highlighted in light and dark blue, respectively. (B) Cartoon representation of the *TtClpP* monomer in peptide-bound (left) and bortezomib-bound (right) states. Helices are named by letters, and strands are indicated by numbers. A zoom of the ligands present in the active sites (dashed boxes) is shown in (C). (C) Substrate-binding pocket of *TtClpP*. The residues involved in the binding to the model peptide (inset 1) and bortezomib (inset 2) are shown as sticks. (D and E) Residue-wise MAS NMR amide <sup>15</sup>N  $R_{1\alpha}$  relaxation rate constants. High  $R_{1\alpha}$  rate constants point to enhanced nanosecond-to-millisecond motions and are found primarily in loop regions and in helix  $\alpha E$ , as shown in (C). (F and G) MD-derived root mean square fluctuations (RMSFs) over the 1- $\mu$ s-long MD trajectory of the assembled 14-mer ClpP.

with the Ne2 atom of His<sup>122</sup> (part of the ClpP catalytic triad) and the backbone amide groups of Met<sup>98</sup> and Gly<sup>68</sup>. The Nδ1 atom of His<sup>122</sup> on its turn forms an additional hydrogen bond with Asp<sup>171</sup>, the third residue of the ClpP catalytic triad. The amino groups of peptide residues P2 and P3 are stabilized by backbone interactions with Gly<sup>68</sup> and Trp<sup>125</sup>, respectively, while the carbonyl groups are stabilized by interactions with amino groups of Trp<sup>125</sup> and Val<sup>70</sup> (Fig. 4C, inset 1). The geometry and distances between the catalytic triad residues (Ser<sup>97</sup>, His<sup>122</sup>, and Asp<sup>171</sup>) are consistent with a functional *Tt*ClpP catalytic triad because their distances are in the range of the ones observed in classical serine proteases and active ClpPs (40). Figure 4B displays the 2.7-Å resolution structure of the bortezomib-bound state (see also fig. S5). Similarly to the peptides present in the *Tt*ClpP:peptide structure, bortezomib tightly fits between β-strand 6 and a small β-turn formed by Gly<sup>68</sup> and Val<sup>70</sup>, thereby forming an antiparallel β-sheet. This conformation is characterized by a hydrogen bond network almost identical to the one observed in the *Tt*ClpP:peptide interaction (Fig. 4C, inset 2). Unlike the *Tt*ClpP:peptide complex, where Ser<sup>97</sup> is forming hydrogen bonds with the terminal carboxy group of the peptide, bortezomib seems to be covalently bound to the active-site Ser<sup>97</sup> via its boron atom. Other significant differences are observed in the conformation of Trp<sup>125</sup>, which is displaced toward bortezomib and forms an aromatic interaction with the bortezomib phenylalanyl group. In addition, the nearby Thr<sup>131</sup> residue, which in peptide-bound *Tt*ClpP interacts with Arg<sup>170</sup> of a neighboring ClpP monomer, now forms a hydrogen bond with Gln<sup>123</sup> of an opposing ClpP monomer in the other heptameric ring. Despite the rearrangement of the Trp<sup>125</sup> side chain, the catalytic triad geometry is again consistent with a fully active protease.

To probe the flexibility of *Tt*ClpP, we used MAS NMR spin relaxation experiments and MD simulations. Site-specific amide <sup>15</sup>N *R*<sub>1ρ</sub> experiments probe the local amplitudes and time scales of backbone motion and are particularly sensitive to nanosecond-to-millisecond motions (41). The sites with highest <sup>15</sup>N *R*<sub>1ρ</sub> rate constants are found not only in the loop regions but also at the tip of helix αE (Fig. 4, D and E). One-microsecond-long MD simulations of the *Tt*ClpP tetradecamer in the apo and peptide-bound states corroborate this observation, albeit with a greater flexibility at the tip of the helix (Fig. 4, F and G). Enhanced flexibility in this helix has been reported also in *Ec*ClpP by methyl-directed solution-state NMR and suggested by other methodologies (16–19, 34, 35). Because of the lack of reliable parametrization of the boronic acid moiety of bortezomib, we focused on simulating tripeptides bound to the active site. It is noteworthy that over the length of three independent simulations of ClpP initially loaded with tri-alanine at all active sites, only a fourth of the peptides remain associated to their designated active sites (fig. S6), consistent with the idea that the products of *Tt*ClpP are necessarily weak binders, which ought to be released rapidly to leave the catalytic sites ready for the next degradation reaction.

### Bortezomib induces the transition from a tense state to a relaxed state with higher affinity for the substrate

Having identified the bortezomib-binding site of *Tt*ClpP by NMR and x-ray crystallography, we turned to ITC to characterize thermodynamically the interaction at different temperatures ranging from 25° to 45°C. At 25°C, far from the ideal catalytic temperature of *Tt*ClpP (65°C), only a moderate affinity and symmetric binding isotherm was observed. The binding enthalpy was negative and rather small (see Table 1), thus indicating that binding was mainly entropy-driven. However, when the temperature was increased to 35° and 45°C,

considerable changes were observed in the binding isotherm. The asymmetry of the curve and the appearance of two different phases in the binding isotherm along the titration, within the low ligand saturation region, were clear indicators of positive cooperativity (Fig. 5A). Because of the architecture of *Tt*ClpP, comprising two heptameric rings forming a tetradecameric assembly, the ITC data were fitted to the Monod-Wyman-Changeux (MWC) model (Fig. 5B), which considers two intrinsically noncooperative conformations, with low and high binding affinity, denoted as tense (T) state and relaxed (R) state, respectively. In the MWC model, ligand binding shifts the equilibrium toward the high-affinity conformation (R) in a concerted cooperative way (i.e., all subunits change their conformation simultaneously; Table 1). This model has been previously transposed to other enzymatic systems, where activation at low ligand concentration and inactivation at high ligand concentration were also observed (42).

Fitting of the ITC data revealed that in the absence of the ligand, 76% of *Tt*ClpP exists in the T state, which has a lower affinity for bortezomib ( $K_{d,T} = K_T^{-1} = 40 \mu\text{M}$ ). Bortezomib binding induces a concerted transition from the T state to the R state, with a higher affinity for substrate ( $K_{d,R} = K_R^{-1} = 15 \mu\text{M}$ ). A model in which we assumed that there is no ligand binding to the subunits in the T conformation ( $K_T = 0$ ,  $\Delta H_T = 0$ ) can be ruled out based on significantly worse fitting. The higher affinity of the ligand for the R conformation drives the conversion of T subunits into R subunits. In addition, the compulsory concerted conformational change within a given protein oligomer (i.e., all subunits in a given protein oligomer must undergo the conformational change at once, even if not all of them are occupied by ligand) will further promote the conversion of T subunits into R subunits, thereby increasing the total fraction of R subunits. The Hill coefficient [maximal slope in the  $\log(n_{LB}/(14 - n_{LB}))$  versus  $\log[\text{ligand}]$  plot] of *Tt*ClpP is 1.3, which is rather small for a tetradecameric protein but was expected considering the small value for the conformational equilibrium constant  $\gamma$ .

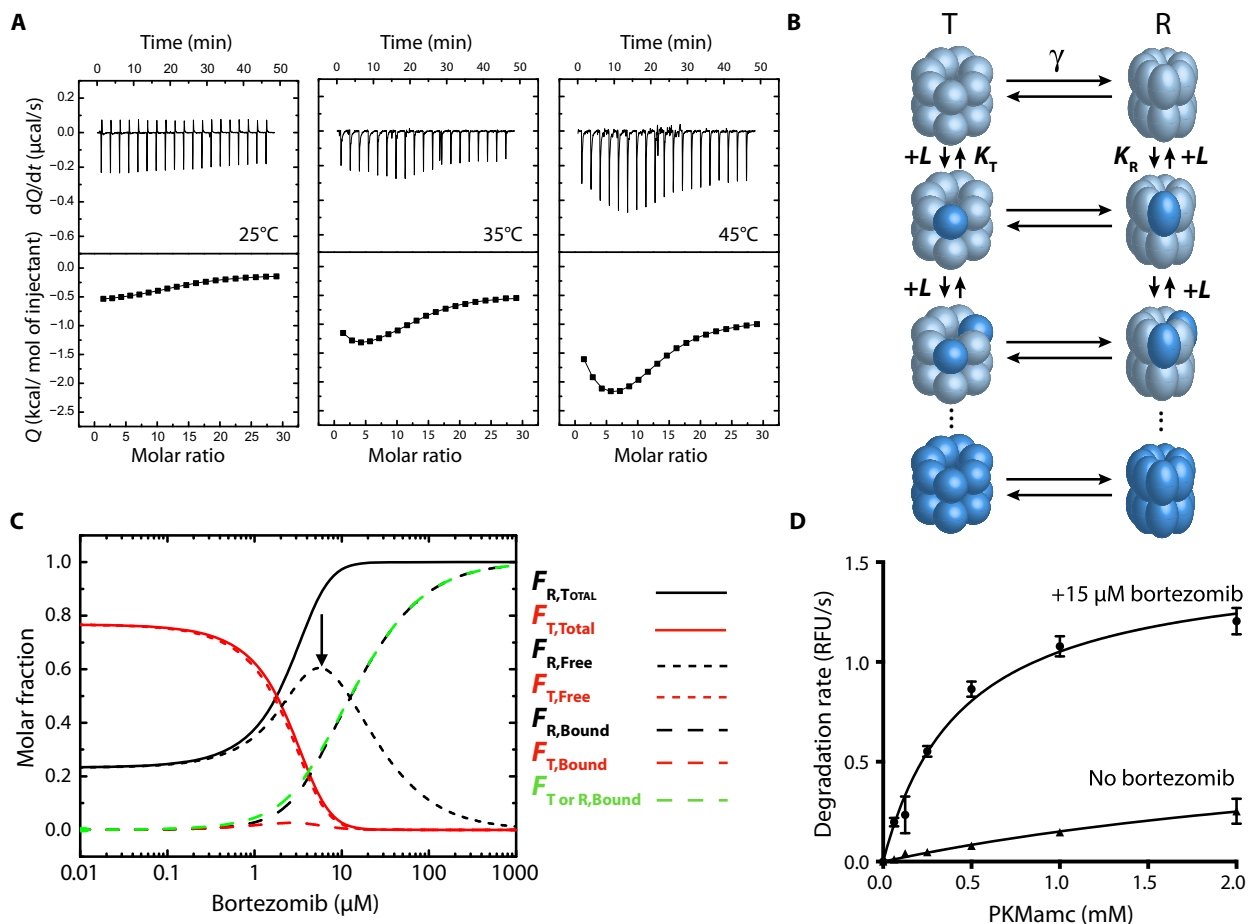
Although the MWC model was used 40 years ago for the calorimetric study of the interaction of trout hemoglobin with carbon monoxide using a gas-liquid reaction calorimeter (43), the present work represents the first detailed mathematical description and implementation of the MWC model in ITC, for which the experimental methodology and the mathematical formalism are not the same as those outlined in the aforementioned seminal work.

The MWC model provides a rationale for *Tt*ClpP activation by bortezomib. Activation of *Tt*ClpP by the inhibitor bortezomib at

**Table 1. Thermodynamic parameters of bortezomib binding from ITC and fitting to the MWC model.**  $K_R$  and  $K_T$  refer to the bortezomib association constants for the relaxed and tense state.  $\Delta H_R$ ,  $\Delta H_T$ , and  $\Delta H_\gamma$  refer to the binding enthalpy of the R state, T state, and conformational enthalpy between R and T conformations.  $\gamma$  is the equilibrium constant between R and T conformations in the absence of the ligand.  $N$  is the fraction of active or binding-competent protein. Relative errors in equilibrium constants are 30%. Absolute error in enthalpies is 0.3 kcal/mol.

T (°C)	N	$K_R$ (M <sup>-1</sup> )	$\Delta H_R$ (kcal/mol)	$K_T$ (M <sup>-1</sup> )	$\Delta H_T$ (kcal/mol)	$\gamma$	$\Delta H_\gamma$ (kcal/mol)
25°C	0.97	$4.5 \times 10^4$	-0.5	No apparent cooperativity			
35°C	0.97	$8.0 \times 10^4$	-0.9	$2.8 \times 10^4$	-0.1	3.3	-0.3
45°C	1.00	$6.5 \times 10^4$	-1.5	$2.5 \times 10^4$	-0.3	3.2	-0.2





**Fig. 5. Cooperative bortezomib binding detected by ITC experiments.** (A) Calorimetric titrations for the interaction of *Tt*ClpP (10  $\mu$ M in the calorimetric cell) with bortezomib (1.4 mM in the injecting syringe) in 50 mM Hepes (pH 7.6) and 50 mM NaCl. Experiments were performed at three different temperatures: 25 $^{\circ}$ C, 35 $^{\circ}$ C, and 45 $^{\circ}$ C. Thermograms (thermal power as a function of time) are displayed in the upper plots, and binding isotherms (ligand-normalized heat effects per injection as a function of the molar ratio,  $[L]_t/[P]_t$ ) are displayed in the lower plots. The binding isotherms were analyzed with the MWC model for *Tt*ClpP, consisting of 14 identical subunits, each one containing a single ligand-binding site. Nonlinear least squares regression analysis allows the determination of the following binding parameters (see Table 1): association constants for the R and T states ( $K_R$ ,  $K_T$ ), binding enthalpies to the R and T states ( $\Delta H_R$ ,  $\Delta H_T$ ), conformational equilibrium constant and conformational enthalpy change between states R and T ( $\gamma$ ,  $\Delta H_\gamma$ ), and fraction of active protein ( $N$ ). (B) MWC model for a 14-mer oligomeric protein. The protein can populate only two conformational states in equilibrium (with equilibrium constant  $\gamma$ ): all subunits in R (relaxed, ellipsoidal shape) conformation or all subunits in T (tense, spherical shape) conformation. Subunits have a single ligand-binding site, exhibiting the R conformation a higher binding affinity ( $K_R > K_T$ ). Ligand binding occurs through an independent (noncooperative) fashion within an oligomer (ligand-free subunits in light blue and ligand-bound subunits in dark blue). T conformation is favored at low ligand concentration; the higher ligand-binding affinity for R conformation promotes a highly cooperative compulsory concerted conformational change driven by ligand binding and involving all subunits within an oligomer, displacing the equilibrium toward the R conformation. (C) Molar fraction of the different protein species (total, ligand-free, and ligand-bound R and T conformations) as a function of ligand concentration: total fraction of subunits in R conformation (continuous black line), total fraction of subunits in T conformation (continuous red line), fraction of subunits in ligand-bound R conformation (dashed black line), fraction of subunits in ligand-bound T conformation (dashed red line), fraction of ligand-free subunits in R conformation (dotted black line, highlighted with an arrow), and fraction of ligand-free subunits in T conformation (dotted red line). In addition, the fraction of ligand-bound subunits in either R or T conformation is shown (dashed green line). It is obvious that the contribution of subunits in T conformation to the ligand binding is very small (dashed red line). At very low ligand concentration, 77% of the protein subunits are in T conformation and 23% are in R conformation, according to the value of the equilibrium constant  $\gamma$  equal to 3.3. At low ligand concentration, the total fraction of subunits in R conformation increases, while the total fraction of subunits in T conformation decreases due to the T $\rightarrow$ R conversion. However, both the fractions of ligand-bound subunits in R and T conformation increase due to ligand binding (although the increment in ligand-bound subunits in T conformation is negligible). At high ligand concentration, the total fraction of subunits in R conformation further increases, while that of total subunits in T conformation decreases, and the fraction of ligand-bound subunits in R conformation further increases due to ligand binding, but the fraction of ligand-bound subunits in T conformation decreases due to the T $\rightarrow$ R conversion. The fraction of ligand-free subunits in R conformation dominates this region, with a maximal population of 60%. This is due to the concerted conversion of subunits within a given protein oligomer (all subunits within an oligomer undergo the conformational change T $\rightarrow$ R, although not all of the subunits bind a ligand). At very high ligand concentration, the fraction of ligand-bound subunits in R conformation dominates the conformational ensemble, reaching a maximal population of 100%. The R $\leftrightarrow$ T equilibrium shows a switchover or crossover point (R and T are equally populated) at around 2  $\mu$ M free bortezomib concentration. (D) *Tt*ClpP (1  $\mu$ M complex) peptidase activity was measured as a function of PKMamc concentration in the presence (15  $\mu$ M) and absence of bortezomib.



moderate concentrations occurs as a consequence of the concerted conformational change within the protein tetradecamer, which makes the total fraction of subunits in the enzymatically active R conformation ( $F_{R,Total}$ ) much larger than the fraction of ligand-bound subunits in R conformation ( $F_{R,Bound}$ ), i.e., there is no proportionality between  $F_{R,Total}$  and  $F_{R,Bound}$  (Fig. 5C). At intermediate bortezomib concentration, the fraction of inhibitor-free (active) subunits in an R conformation is maximal and predominates in the speciation distribution (Fig. 5C). Thus, the extents of cooperativity for binding and cooperativity for activity are different [a phenomenon previously reported for allosteric proteins in (44)]. If the conformational changes were not concerted, then the activity of *TtClpP* would continuously decrease with increasing bortezomib concentration, with no maximal activity at intermediate bortezomib concentration. The increase in the fraction of inhibitor-free R subunits at moderate inhibitor concentration is the molecular event underlying the increase in activity. Similarly, the decrease in the fraction of inhibitor-free R subunits at high inhibitor concentration is the molecular event underlying the decrease in activity at high inhibitor concentration.

The ITC-derived model predicts accurately the activity data. According to the peptide degradation activity measurements (Fig. 1A), the maximal *TtClpP* activity is found around 12.5  $\mu\text{M}$  of total bortezomib. The ITC-derived MWC model predicts that the maximum fraction of inhibitor-free *TtClpP* in the R conformation ( $F_{R,Free}$ ) should be around 6  $\mu\text{M}$  of free bortezomib (Fig. 5C). Considering the population of bound *TtClpP* at 6  $\mu\text{M}$  free bortezomib concentration ( $F_{T_{orR,Bound}} = 0.3$ ; Fig. 5C), and taking into account the *TtClpP* subunit concentration of 10  $\mu\text{M}$  (used in our activity measurements), one can calculate that the concentration of bortezomib bound to *TtClpP* is 3  $\mu\text{M}$ . Consequently, the total concentration of bortezomib for achieving 6  $\mu\text{M}$  free bortezomib (i.e., the concentration for maximal bortezomib-free subunits in R conformation) is around 9  $\mu\text{M}$ , in excellent agreement with the activity measurements reporting a maximum activity for 12.5  $\mu\text{M}$  bortezomib concentration (Fig. 1A).

Furthermore, to verify that the bortezomib-induced R state has a higher affinity for the substrate, we measured peptide hydrolysis as a function of substrate concentration in the apo state and in the presence of nonsaturating concentrations of bortezomib (15  $\mu\text{M}$ ). While there are differences between  $K_d$ , a thermodynamic equilibrium constant, versus a purely kinetic constant as  $K_M$ , in the presence of bortezomib, *TtClpP* displayed an apparent  $K_M$  around 400  $\mu\text{M}$ . In its absence, no indications of substrate saturation were observed (Fig. 5D), reflecting the major differences between the bortezomib-bound and apo states. Put together, the activity measurements and ITC data reveal that activation of *TtClpP* by an inhibitor is achieved by shifting the conformational population toward a higher-affinity R state.

### Off-equilibrium MD simulations provide rationale for stabilization of extended state

Activity measurements, ITC data, x-ray structures, and NMR spectroscopic data have established that, by binding to a subset of *TtClpP* active sites, bortezomib and peptide substrates favor a catalytically more active particle by keeping both the bound and free subunits in an active state, with the catalytic triad poised for catalysis. We attempted to bridge the functional and thermodynamic information from biochemistry and ITC with the structural view by performing *in silico* experiments. In a nutshell, in these MD simulations, we investigated the energetics of the transition from the extended (active) to the compressed (inactive) state and interrogated whether the pres-

ence of a molecule bound to the active site alters the relative stability of the two states. Near-equilibrium MD experiments were performed, in which the two heptameric rings were slowly brought from the extended state to the compressed state using the centers of mass of the two rings as single variable for the steered MD protocol (see Materials and Methods). Figure 6 (A and B) displays the force required for this extended-to-compressed transition, obtained from five independent *in silico* experiments, as a function of the distance of the two half-rings. Figure 6B shows the work (i.e., the integrated force) for this process. The present simulations have been performed for both the apo state (gray) and a substoichiometrically peptide-loaded state in which only 2 or 3 of the 14 binding sites were occupied with tri-alanine peptide. During the compression process, none of the initially bound peptides were expelled from their target catalytic sites, and all the secondary structure elements were preserved (Fig. 6, C to E). Compressing ClpP from the extended conformation requires a stronger force in the presence of a substoichiometric amount of alanine tripeptides compared to the empty protein (Fig. 6, A and B).

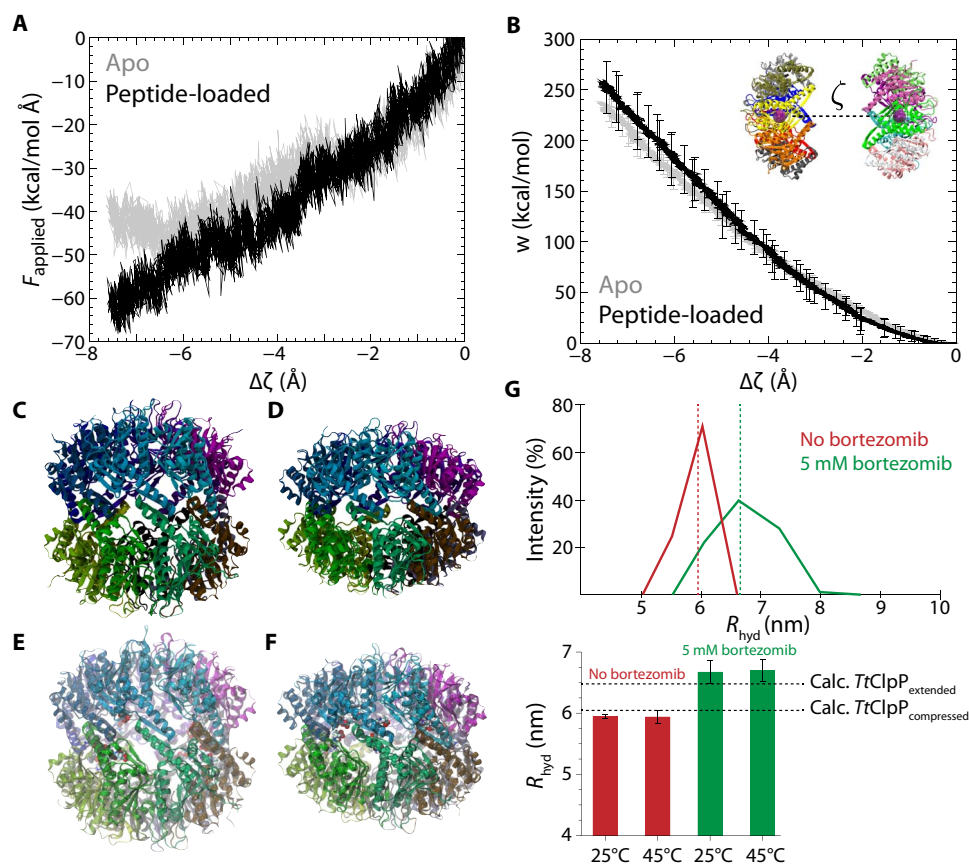
Together, these data provide a rationale for the enhanced stability of the extended conformation in the presence of a ligand: Even substoichiometric amounts of the ligand, i.e., only partially occupied sites, lead to an overall stabilization of all subunits in the active state. The simulations were performed with peptides rather than bortezomib due to the lack of reliable force field parameters for boronic acid and reveal the substrate-induced stabilization of the active extended state. It can be expected that this stability difference of extended over compressed state is even more pronounced with the covalently bound bortezomib.

The MD simulations support the idea that bound ligand, even at substoichiometric amounts, favors a stable extended form, which we believe is the R state modeled in the ITC data, while the apo state is a compressed state. We attempted to validate the compressed-to-extended state transition observed from the MD experiments by performing dynamic light scattering (DLS). Figure 6G shows the hydrodynamic radii ( $R_{hyd}$ ) of *TtClpP* in the absence and presence of 5 mM bortezomib, revealing that addition of bortezomib causes an increase of approximately 8 Å in  $R_{hyd}$  (or 16 Å in diameter) of *TtClpP*. The experimentally obtained hydrodynamic radii for *TtClpP* without and with 5 mM bortezomib (5.95 and 6.67 nm, respectively) are in excellent agreement with the theoretical hydrodynamic radii obtained for compressed and extended states of ClpP, supporting the view that the T-to-R transition modeled from the ITC data corresponds to a compressed-to-extended transition.

### DISCUSSION

Activation of different ClpPs by previously identified inhibitors has been independently reported (3, 22), but to date, no mechanism for these paradoxical results has been proposed. Bortezomib, just like N-blocked peptide aldehyde *Mtb* activators reported previously, does not follow the canonical definition for an allosteric activator, i.e., they bind to the protease active site (14, 15) and their effect is not universal to all ClpP homologs. For instance, we show here that although bortezomib activates *TtClpP* (up to 100  $\mu\text{M}$ ), it is, in the same concentration range, a good inhibitor of *EcClpP*. This result is rather puzzling, considering the high sequence and structural similarity between the two proteins (58% sequence identity).

We show in the present contribution that ClpP activation by inhibitors derives from the intrinsically cooperative character of the



**Fig. 6. In silico pulling experiment of *TtClpP* from the extended state to the compressed state.** (A and B) Profile of the force (A) and work (B) required for compressing the two rings toward each other as a function of the distance  $\zeta$  between the rings (shown in the inset, with exaggerated distance for better visibility). Three independent 200-ns simulations were performed for apo *TtClpP* or *TtClpP*, in which 2 or 3 of the 14 binding sites were occupied with trialanine peptide (see Materials and Methods). (C and D) Conformations of *TtClpP* extended (C) and *SaClpP* compressed (PDB: 3QWD) (D) states observed in crystal structures. (E) Starting conformation of *TtClpP* in the steered MD simulations (shown as sphere representation). (F) Final conformation of the compressed state at the conclusion of the steered MD. Note that the peptides are still bound and that the structure is in good overall agreement with the compressed state from x-ray diffraction shown in (D). (G) DLS analysis of *TtClpP* (3  $\mu$ M) supplemented without or with 5 mM bortezomib. The upper graph shows one representative curve of three replicates measured at 25°C for *TtClpP* without (red curve) or with 5 mM bortezomib (green curve), with the relative intensity (in %) plotted versus particle size [hydrodynamic radius,  $R_{hyd}$  (in nm)]. The average of the three replicates is indicated using a dotted line. The lower graph depicts average hydrodynamic radii ( $R_{hyd}$ ) calculated from three independent experiments measured at 25° and 45°C (0 mM bortezomib, 25°C:  $5.946 \pm 0.036$  nm; 0 mM bortezomib, 45°C:  $5.940 \pm 0.102$  nm; 5 mM bortezomib, 25°C:  $6.673 \pm 0.192$  nm; 5 mM bortezomib, 45°C:  $6.702 \pm 0.176$  nm). Theoretical calculated hydrodynamic radii of compressed *TtClpP* (6.05 nm) and the extended *TtClpP*:bortezomib complex (6.47 nm) are indicated using a dotted black line (see also Materials and Methods).

ClpP tetradecameric structure. Bortezomib, although blocking a number of active sites, can overall increase the catalytic rate by promoting a concerted transition in the complex, stabilizing the active-state conformation. In line with this observation, the near-equilibrium MD simulations performed in our study suggest that compression of the extended-active conformation of *TtClpP* is hampered when a fraction of the catalytic sites is occupied by unspecific tripeptide substrates. In the presence of a stronger specific binder such as bortezomib, it could be reasonably anticipated that compression toward the inactive conformation is even more hindered. A comparable mechanism likely underlies *MtbClpP* activation, because it has been shown that the peptide activator would sterically clash with a non-active ClpP1 complex, thereby forcing the complex into an active conformation (22).

The similarity between the bortezomib *TtClpP* structure and the structure of *TtClpP* with peptides in the active site strongly suggests that our findings reflect an inherent substrate-dependent activation mechanism of the enzyme. Through this mechanism, the presence

of the substrate in the ClpP active site promotes the activation of the remaining subunits, allowing efficient and processive protein degradation. Although we were unable to obtain a crystal structure of apo *TtClpP*, the DLS experiments performed here are consistent with a bortezomib-induced transition from a compressed or compacted state to an extended *TtClpP* conformation. Addition of bortezomib altered chemical shifts in MAS NMR experiments, possibly pointing also toward the changes in the equatorial region of ClpP, expected for a compressed/extended transition (Fig. 3D).

Our results also support the previously described bidirectional cross-talk between chaperone binding and protease active sites (14, 22, 23, 26–28), because bortezomib is no longer an activator in the presence of the co-chaperone *TtClpX* and *TtClpP* partially saturated with bortezomib binds strongly to *TtClpX*. Significant differences were observed between the activation of *TtClpP* peptidase activity and proteolytic activity by bortezomib. While peptidase activation displayed a bell-shaped curve, protease activation showed a

linear rise in the range of the used bortezomib concentration. Although the process underlying bortezomib activation in the two activities should be common, important differences exist between the two assays. Protein degradation is strictly dependent on substrate access to the proteolytic chamber, because larger substrates cannot diffuse freely toward the proteolytic chamber. Our results showing protein degradation activated by bortezomib demonstrate that modulation of protein entry can also be achieved via binding to the ClpP active site and not exclusively by ClpX or ADEP binding. Protein access to the proteolytic chamber leads to high local concentrations of the substrate present at the ClpP active sites. Although bortezomib activates protein degradation at low concentrations, it must compete with high concentrations of substrate in the proteolytic activity assay, thereby rationalizing the requirement for higher concentrations of drug to achieve activation. The range of bortezomib concentrations used in our proteolytic activity assay likely corresponds to the rising part of the bell curve monitored in the peptidase activity assay and also explains why several ClpP inhibitors are 10-fold more potent in peptidase assays compared with protease assays (45).

It is unclear why the action of bortezomib and activators is not uniformly conserved between different ClpPs. Differences in enzymatic activity are generally more difficult to decode than major structural differences because allosteric pathways often result from subtle changes in conformational dynamics that affect population distributions, which are not readily observed from comparisons of structures. For example, single mutations or truncations in *EcClpP* N-terminal have been shown to induce reversible inactivation of ClpP (16), and using an indirect methodology based on the association of an *EcClpP* E14W mutant with *EcClpA*, the equilibrium constant between an *EcClpP* active state (R) and inactive state (T) was previously quantified as  $K_{eq} = 7.5$  (37°C) (46), i.e., a value similar to the one we describe here for *TtClpP*  $K_{eq} = 3.3$  (45°C). This finding is particularly interesting, as it suggests that in the absence of mutations, *EcClpP* preferentially populates active states, while *TtClpP* is preferentially inactive (16). The energy barrier between the two states is, however, rather small, with an energy gap of 0.7 kcal/mol between the R and the T conformations in *TtClpP*. Given the high sequence identity between *EcClpP* and *TtClpP*, the key to the difference in the equilibria between active and non-active states must necessarily result from small sequence changes. *EcClpP* and *SaClpP* contain a conserved Q131XT/S133 motif at the tip of the turn between helix  $\alpha E$  and the  $\beta$ -strand 6 in the handle domain (fig. S7). Here, Gln<sup>131</sup> engages in hydrogen bonds with the  $\eta 1$  and  $\eta 2$  nitrogens of Arg<sup>170</sup> on a neighboring subunit of the same heptamer ring (intra-ring) and with the backbone oxygen of Glu<sup>169</sup> and  $\epsilon 1$  oxygen of Gln<sup>123</sup> on a monomer of the other ring (inter-ring). The side chain of the conserved Thr<sup>133</sup> forms hydrogen bonds with the nitrogen atoms of Gln<sup>123</sup> and Lys<sup>146</sup> (inter-ring). Gln<sup>123</sup> is part of the conserved His-Gln-Pro (HQP) motif, which includes the catalytic triad His<sup>122</sup>, while Arg<sup>170</sup> is the vicinal residue of the catalytic triad Asp<sup>171</sup>. It is clear that this extended hydrogen bond network restricts the degrees of freedom of the aforementioned catalytic triad residues. It is noteworthy that these residues are conserved neither in *TtClpP* nor in *MtbClpP1*. In *TtClpP*, Gln<sup>131</sup> is replaced by a threonine (Thr<sup>131</sup>), which interacts either with the  $\epsilon 1$  nitrogen of Arg<sup>170</sup> (intra-ring) or with Gln<sup>123</sup> (inter-ring) in the *TtClpP*:peptide or *TtClpP*:bortezomib structures, respectively. Ser<sup>133</sup> of *TtClpP* contacts Gln<sup>123</sup> and Lys<sup>146</sup> (inter-ring) in a similar way as *EcClpP* Thr<sup>133</sup> (fig. S7). The replacement of the conserved glutamine Q131 by the shorter threonine in *TtClpP* or serine in *MtbClpP1* leads to modifications in the hydrogen network that normally sustains the

extended helix  $\alpha E$ . Structural and MD studies in *SaClpP* have shown that these key residues are essential for keeping the long helix  $\alpha E$  in a straight conformation. In *TtClpP*, the exclusive hydrogen bond acceptor character and shorter side chain of Thr<sup>131</sup> does not allow the formation of a network that simultaneously coordinates two of the members of the catalytic triad. This reduced H-bonding capacity is expected to result in an increased flexibility in that region and likely to an equilibrium shift toward nonfunctional conformations, supported by the flexibility measurements reported here by MAS NMR and MD simulations (Fig. 5). We propose that bortezomib binding through preferential stabilization of the active site (R form) shifts the entire population toward a fully active enzyme.

A different hydrogen bond network is also observed for *MtbClpP1*. In this case, Gln<sup>131</sup> is replaced by serine, while Thr<sup>133</sup> is replaced by alanine. When purified alone, ClpP1 is an inactive enzyme, and the x-ray structure of the ClpP1 tetradecamer shows that the catalytic triad is not in the active state (13). However, in complex with ClpP2, ClpP1 has been shown to be fully active, likely because *MtbClpP2*, which contains the conserved QFT motif at the tip of the helix  $\alpha E$ , can stabilize ClpP1 by forming a similar network, as has been observed in *EcClpP* (3).

Cooperativity allows multimeric systems to adapt rapidly to environmental changes by more efficiently converting a biochemical input, i.e., the substrate/ligand concentration, into a biochemical response. Cooperative response to incoming substrates has been reported for other machineries of the protein quality control system. Because protein degradation is key to cell homeostasis, several levels of regulation have evolved to prevent uncontrolled proteolysis that would be deleterious to the cell. ClpP, as well as other complexes like the 20S proteasome, hides their active sites inside an inaccessible catalytic chamber to prevent uncontrolled proteolysis. This architecture is fine-tuned by ClpP-specific co-chaperones, which add another layer of regulation in protease control, either by “opening” the axial pores of the cylinder and translocating unfolded substrates into the chamber or by directly activating the ClpP active sites (2). Several studies have reported that ClpX is able to revert the conformational inhibition of ClpP induced by chemical compounds (14) or even by mutations known to promote ClpP inactive states (23). The fact that one of these inactive-state promoting mutations, namely, the Asp<sup>172</sup>Asn *SaClpP* mutant, corresponds to the native form of *LmClpP1* and several other ClpPs suggests that this mechanism could be relevant for in vivo ClpP regulation in these species (47). Our data support a similar mechanism of conformational control for *TtClpP*. Our observation that *TtClpX* and *EcClpX* (data not shown) can activate *TtClpP* further suggests that this enzyme is conformationally attenuated, i.e., in the absence of the chaperone, it is in a latent inactive state. The nonphysiological activation observed with bortezomib is both a consequence of the inherent cooperativity of the protease and the fact that, in the absence of activators, its predominant state is the non-active T state (Fig. 5C). The small molecule-induced conformational shift opens new possibilities for drug development, particularly as not only peptide degradation but also protein degradation is stimulated, as we report here (Fig. 1C).

In summary, using a host of biochemical, biophysical, and structural approaches, we have proposed a mechanism for the activation of ClpP by inhibitors with a key element, namely, a concerted activating conformational change in the supramolecular ClpP tetradecameric structure elicited by the binding of inhibitors at subsaturating concentrations. Hints to a similar allosteric behavior have been reported for the proteasome in the presence of bortezomib (47). While bortezomib inhibits efficiently the chymotrypsin site of the proteasome,



it has been observed that it induces a paradoxical activation of the trypsin site of the complex (24).

## MATERIALS AND METHODS

### Cloning and mutagenesis

*TtClpP* DNA was directly amplified from *T. thermophilus* HB8 DNA and cloned in to a pET41c (Novagen) expression vector using *Ned I* and *Xho I* restriction enzymes. The expressed gene contained the additional residues LEHHHHHHHH at its C terminus to allow affinity chromatography purification. Full-length ClpX DNA was first cloned into a pET28 vector using *Ned I* and *Hind III* restriction enzymes, resulting in residues MGSSHHHHHSSGLVPRGSH added to the N terminus. While full-length ClpX was insoluble, similar to ClpX orthologs from other species, removal of its N-terminal domain (residues 1 to 54) allowed expression of a soluble ClpX construct ( $\Delta TtClpX$ ) (48). Removal of the *TtClpX* N terminus was performed by GeneCust using *Ned I* and *Hind III* as restriction enzymes. *TtClpP* and  $\Delta TtClpX$  were expressed in *E. coli* BL21RIL cells following overnight expression at 20°C after induction with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). A *TtClpP* mutant containing an S97A mutation was constructed using the QuikChange II Site-Directed Mutagenesis Kit from Thermo Fisher Scientific.

### Protein purification

GFPssrA was purified as previously described (49). Casein FITC from bovine milk was obtained from Sigma (C0528) and loaded onto a G25 column (GE Healthcare) to remove free fluorescein (49). *TtClpP* and  $\Delta TtClpX$  were purified using native NiNTA (nickel-charged nitrilotriacetic acid) affinity chromatography followed by SEC using a 16/600 S200 200 pg Superdex column and a final SEC buffer containing 100 mM tris (pH 8.5), 100 mM NaCl, and 5% glycerol. U- $[^2\text{H}, ^{15}\text{N}]$ , Ile- $\delta$ 1- $[^{13}\text{CH}_3]$  *TtClpP* was expressed as described (33). When expressed in  $\text{D}_2\text{O}$ , *TtClpP* was insoluble and the protein was recovered from *E. coli* inclusion bodies using denaturing conditions. Briefly, cells were harvested and resuspended in denaturing buffer containing 8 M urea followed by sonication and centrifugation to remove cell debris. The resulting supernatant was loaded onto a NiNTA column and washed several times with denaturing buffer. The column was then equilibrated in refolding buffer [100 mM sodium phosphate (NaPi) (pH 8), 5% glycerol, and 100 mM NaCl] followed by elution with elution buffer [100 mM NaPi (pH 8), 5% glycerol, 100 mM NaCl, and 400 mM imidazole]. The eluted fractions were loaded into a 16/600 S200 200 pg Superdex column, and the fractions corresponding to native ClpP were pulled. *EcClpP* was purified as described in (50). FtsZ was a gift of T. Hosek and was purified as previously described (50).

### Biochemical assays

Concentrations of ClpP and ClpX were calculated using the molecular weight of a tetradecamer and a hexamer, respectively. PKMamc degradation by *TtClpP* and *EcClpP* was measured as described (45). Briefly, *TtClpP* in Hepes buffer [50 mM Hepes (pH 7.6) and 50 mM NaCl] was mixed with indicated concentrations of PKMamc and bortezomib/Bz-LL [or respective dimethyl sulfoxide (DMSO) volume] at 60°C (or indicated temperature), and fluorescence (340-nm excitation, 462-nm emission) was measured at defined time intervals. FITC-casein degradation by *TtClpP* was measured as described (45); in a typical assay, FITC-casein (0.15  $\mu\text{M}$ ) was mixed with *TtClpP*<sub>14</sub> (0.1  $\mu\text{M}$ ) and fluorescence increase (440-nm excitation, 509-nm emis-

sion, 495-nm cutoff) was monitored. GFPssrA (0.5  $\mu\text{M}$ ) degradation by *TtClpX*<sub>6</sub> (0.2  $\mu\text{M}$ ) and *TtClpP*<sub>14</sub> (0.1  $\mu\text{M}$ ) in the presence of ATP (10 mM) was monitored, taking advantage of the intrinsic fluorescence of GFPssrA (440-nm excitation, 509-nm emission, 495-nm cutoff) at the indicated temperatures. FtsZ degradation by *TtClpP* was measured at 37°C. Briefly, *TtClpP*<sub>14</sub> (1  $\mu\text{M}$ ) was mixed with purified FtsZ (20  $\mu\text{M}$ ) in the presence of the indicated concentrations of bortezomib. Aliquots were removed at the indicated time points, and the samples were loaded in a 15% SDS agarose gel. It is important to note the substantial differences in *TtClpP* concentration used for the different techniques, with 10-fold more protease being used for the FtsZ protein degradation study compared with FITC-casein degradation, as a consequence of the lower sensitivity of SDS-polyacrylamide gel electrophoresis staining versus fluorescein fluorescence methods. *TtClpP* reaction with ActivX TAMRA-FP was executed according to the manufacturer's instructions (Thermo Fisher Scientific). *TtClpP*<sub>14</sub> (1 mg/ml) was incubated in the indicated conditions with 1  $\mu\text{l}$  of TAMRA-FR reagent (final concentration of 5  $\mu\text{M}$ ). At defined time intervals, aliquots were removed and loaded in a 4 to 20% SDS gel. Gels were visualized in a Bio-Rad ChemiDoc XRS system. *TtClpP* (20  $\mu\text{M}$ ) unfolding was measured in a Varian Cary Eclipse spectrofluorimeter monitoring tryptophan intrinsic fluorescence (280-nm excitation, 350-nm emission). *TtClpP*<sub>14</sub> analytical SEC was performed using a Superdex 200 10/300 GL column.

### NMR experiments

MAS solid-state NMR experiments were performed on a Bruker Avance III spectrometer with a 1.3-mm MAS NMR probe tuned to  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^2\text{H}$  frequencies at a MAS frequency of 55 kHz and an effective sample temperature of 30°C. Backbone resonance assignment was achieved with proton-detected triple-resonance experiments described earlier (33). The assignment was obtained for 97 residues; many more spin systems were observable in the spectra, but spectral overlap challenged further assignment, which likely requires additional high-dimensional experiments. Given that in the present context we were interested in probing interactions and dynamics, the current extent of assignment was sufficient here.

For the measurement of CSPs upon addition of bortezomib, we prepared two samples of U- $[^2\text{H}, ^{13}\text{C}, ^{15}\text{N}]$ -labeled ClpP in  $\text{H}_2\text{O}$ -based buffer [100 mM tris (pH 8.5), 5% glycerol, and 100 mM NaCl] containing either 10 mM bortezomib and 5% DMSO or only 5% DMSO. In both cases, the protein (10 mg/ml) was sedimented into a 1.3-mm Bruker MAS NMR rotor, and hCANH spectra were collected in ca. 2.5 days per spectrum. The peak positions in the two spectra were extracted for all assigned backbone sites using CCPN (Collaborative Computing Project for NMR) software (51).

Spin relaxation measurements ( $^{15}\text{N}$   $R_{1\rho}$ ) were performed with proton-detected 3D hCANH experiments. The 3D approach was chosen to resolve sites that would be overlapped in 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra. The experiment comprised a  $^{15}\text{N}$  spin-lock element of 15-kHz radio-frequency field strength and delays of 2 and 55 ms. Total experimental times were 3 days per 3D experiment. All spectra were acquired with TopSpin v3.5 and analyzed with CCPN. Relaxation rate constants were obtained using in-house written Python scripts using the peak intensities extracted from the 3D spectra, and error bars were determined using Monte Carlo analysis based on the spectral noise level (determined in CCPN).

Solution-state NMR experiments were performed on a Bruker Avance III spectrometer, equipped with cryogenically cooled TCI



probeheads, operating at a magnetic field strength corresponding to  $^1\text{H}$  Larmor frequencies of 700 MHz. The sample temperature was set to 60°C, and 2D SOFAST-methyl-TROSY NMR experiments were recorded with an adjusted duration depending on the final concentration of the proteins (experimental time ranging from 10 to 120 min per sample) (52). The angle of the proton excitation pulse was set to 30°, and the recycling delay was optimized to 0.6 s to achieve the highest sensitivity (51).  $\text{U}-[^2\text{H}, ^{15}\text{N}], \text{Ile-}\delta 1-[^{13}\text{CH}_3]\text{TtClpP}$  and  $\text{U}-[^2\text{H}, ^{15}\text{N}], \text{Ile-}\delta 1-[^{13}\text{CH}_3]\text{S97ATtClpP}$  (250  $\mu\text{M}$ ) were mixed with 2 mM bortezomib (or DMSO equivalent) in tris buffer [100 mM tris (pH 8.5), 100 mM NaCl, and 5% glycerol].

### Crystallization of TtClpP and the TtClpP:bortezomib complex

Crystallization trials were set up manually in a sitting drop vapor diffusion setup using 24-well sitting drop plates (Hampton Research) with drop volumes of 1  $\mu\text{l}$  (0.5  $\mu\text{l}$  of protein solution + 0.5  $\mu\text{l}$  of reservoir solution). Purified TtClpP was concentrated to 9 mg/ml. Approximately 1 week after setting up crystal plates at 293 K, large cube-shaped crystals were obtained in a condition containing 0.1 M sodium acetate (pH 4.8) and 30% polyethylene glycol 400 (PEG-400).

To obtain crystals of the TtClpP:bortezomib complex before setting up crystal plates, 5  $\mu\text{l}$  of a 100 mM bortezomib solution dissolved in DMSO was added to 45  $\mu\text{l}$  of TtClpP solution (9 mg/ml), resulting in a final concentration of 10 mM bortezomib and 8.1 mg/ml of TtClpP. Large brick-shaped crystals of the TtClpP:bortezomib complex were obtained at 293 K in the same condition as TtClpP [0.1 M sodium acetate (pH 4.8) and 30% PEG-400]. TtClpP crystals were scooped, transferred to a solution containing 32% PEG-400 and 15% glycerol, and subsequently flash-cooled in liquid nitrogen. Crystals of the TtClpP:bortezomib complex were scooped directly from the drops and flash-cooled in liquid nitrogen. The diffraction data of TtClpP (1.95 Å) and the TtClpP:bortezomib complex (2.70 Å) were collected at the ID30A-MASSIF (TtClpP) and ID23-1 MX (TtClpP:bortezomib) beamlines at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. The obtained datasets were processed using the XDS package (53).

### Structure determination of TtClpP and the TtClpP:bortezomib complex

Diffraction data for TtClpP were processed in space group C2 ( $a = 105.98$  Å,  $b = 162.79$  Å,  $c = 107.95$  Å,  $\alpha = \gamma = 90^\circ$ ,  $\beta = 116.34^\circ$ ). The structure was solved by maximum-likelihood molecular replacement (MR) using Phaser (54) in Phenix (55). The x-ray structure of EcClpP in complex with ADEP1 was used as a search model [PDB: 3KTI; (56)]. Before running Phaser, side chains of the search model were trimmed using the Schwarzenbacher (57) side-chain pruning method in Sculptor (Phenix). The resulting MR solution was refined in Phenix using positional (XYZ) and real-space refinement with noncrystallographic symmetry (NCS) restraints, occupancy refinement, individual ADP (atomic displacement parameter) refinement with TLS (translation liberation screw), and optimized x-ray/stereochemistry and x-ray/ADP weights. Multiple rounds of refinement in Phenix were followed by additional manual model building carried out in Coot (58).

Diffraction data for TtClpP:bortezomib were processed in space group C222<sub>1</sub> ( $a = 135.14$  Å,  $b = 168.74$  Å,  $c = 166.08$  Å,  $\alpha = \beta = \gamma = 90^\circ$ ). The structure was solved by maximum-likelihood MR using Phaser in Phenix, taking the refined 1.95-Å TtClpP structure as a search model. Refinement of the TtClpP:bortezomib structure in Phenix was carried out using the same strategy as for TtClpP (positional and real-space

refinement with NCS restraints, occupancy refinement, individual ADP refinement with TLS, and optimized x-ray/stereochemistry and x-ray/ADP weights, followed by additional manual model building in Coot).

### Isothermal titration calorimetry

The interaction between TtClpP and bortezomib was assessed by ITC using an Auto-iTC200 microcalorimeter (MicroCal-Malvern Panalytical, Malvern, UK). Calorimetric titrations were performed with a 1.4 mM bortezomib solution in the injecting syringe and a 10  $\mu\text{M}$  ClpP solution in the calorimetric cell in 50 mM Hepes (pH 7.6) and 50 mM NaCl. All solutions were properly degassed to avoid bubble formation during stirring. In each titration, a sequence of 19 2- $\mu\text{l}$  injections was programmed, with reference power of 10  $\mu\text{cal/s}$ , initial delay of 60 s, spacing between injections of 150 s, and stirring speed of 750 rpm.

The heat evolved after each ligand injection was obtained from the integral of the calorimetric signal. Experiments were performed in replicates, and data were analyzed using in-house developed software implemented in Origin 7.0 (OriginLab, Northampton, MA). The binding isotherms (ligand-normalized heat as a function of the molar ratio) were analyzed considering the MWC model for ClpP, an oligomeric macromolecule consisting of 14 identical subunits containing a single ligand-binding site each (52). Nonlinear least squares regression allows the determination of the following binding parameters: equilibrium constants ( $K_R$ ,  $K_T$ ,  $\gamma$ ), enthalpies ( $\Delta H_R$ ,  $\Delta H_T$ ,  $\Delta H_\gamma$ ), and fraction of active protein ( $N$ ). A fitting routine was implemented in Origin 7.0 (OriginLab).

### MWC model implemented in ITC

The MWC model considers an oligomeric macromolecule consisting of  $n$  identical subunits containing a single ligand-binding site each. Those subunits may adopt two conformations, R (relaxed) and T (tense), with different ligand-binding affinities ( $K_R > K_T$ ). Within a given oligomer, all subunits exhibit the same conformation and ligand binding occurs to any subunit independently. However, although T state predominates initially, as a consequence of the higher ligand-binding affinity for R conformation, ligand binding will shift the conformational equilibrium by indirectly eliciting a concerted conformational change of all subunits toward R conformation within the same oligomer at once, resulting in a cooperative binding behavior. The MWC model can only reproduce positive cooperativity, while the Koshland-Nemethy-Filmer (KNF) model can reproduce both negative and positive cooperativities. Both models, MWC and KNF, represent reduced and limited versions of a general allosteric model (44, 59) that takes into account all possible conformational states (e.g., mixed conformations within a protein oligomer) and all possible liganded states, but, in practice, they are considerably more useful and manageable than the general allosteric model.

The binding polynomial for a macromolecule with  $n$  subunits that can exist, all at once within a given oligomer, in two different conformations, R and T, and each subunit with a single ligand-binding site, is given by

$$Z = \sum_{i=0}^n \frac{[PL_i]}{[P]} = \sum_{i=0}^n \left( \frac{[RL_i]}{[R]} + \frac{[TL_i]}{[T]} \right) \quad (1)$$

where  $PL_i$  represents the protein complex with  $i$  ligand molecules bound, and  $RL_i$  and  $TL_i$  refer to the complexes of each oligomeric conformational state with  $i$  binding sites occupied by ligand molecules. In terms of

site-specific binding parameters, the MWC binding polynomial is written as follows

$$\begin{aligned} Z &= (1 + K_R[L])^n + \gamma(1 + K_T[L])^n \\ &= \sum_{i=0}^n \binom{n}{i} (K_R^i + \gamma K_T^i) [L]^i \end{aligned} \quad (2)$$

where  $K_R$  and  $K_T$  are the site-specific microscopic intrinsic association constants for a binding site in a protein subunit in the R and T states, respectively, with  $K_R > K_T$ , and  $\gamma$  is the equilibrium constant for the conformational equilibrium between the T and R oligomers ( $\gamma = [T_n]/[R_n]$ ). Because initially the oligomer in T conformation predominates, the equilibrium constant  $\gamma$  is larger than 1. Each term represents the binding subpolynomial considering  $n$  independent ligand-binding sites in each conformational state, R and T.

Valuable information can be extracted from the binding polynomial. In particular, the molar fraction of each liganded species

$$\begin{aligned} F_{RL_i} &= \frac{\binom{n}{i} K_R^i [L]^i}{Z} \\ F_{TL_i} &= \frac{\binom{n}{i} \gamma K_T^i [L]^i}{Z} \end{aligned} \quad (3)$$

The two first derivatives of the binding polynomial relative to ligand concentration and temperature provide two fundamental quantities, the average number of ligand molecules bound per oligomeric macromolecule

$$n_{LB} = \langle i \rangle = \frac{\partial \ln Z}{\partial \ln [L]} = \frac{n}{Z} \left( K_R [L] (1 + K_R [L])^{n-1} + \gamma K_T [L] (1 + K_T [L])^{n-1} \right) \quad (4)$$

that can be expressed more conveniently as

$$\begin{aligned} n_{LB} = \langle i \rangle &= \frac{\partial \ln Z}{\partial \ln [L]} = \frac{1}{Z} \left( \sum_{i=0}^n \binom{n}{i} i (K_R^i + \gamma K_T^i) [L]^i \right) \\ &= \sum_{i=0}^n i F_{RL_i} + \sum_{i=0}^n i F_{TL_i} \end{aligned} \quad (5)$$

and the average molar excess binding enthalpy

$$\begin{aligned} \langle \Delta H \rangle &= RT^2 \frac{\partial \ln Z}{\partial T} = \frac{1}{Z} \left( n K_R [L] (1 + K_R [L])^{n-1} \Delta H_R \right. \\ &\quad \left. + \gamma (1 + K_T [L])^n \Delta H_\gamma + n \gamma K_T [L] (1 + K_T [L])^{n-1} \Delta H_T \right) \end{aligned} \quad (6)$$

that can also be expressed more conveniently as

$$\begin{aligned} \langle \Delta H \rangle &= RT^2 \frac{\partial \ln Z}{\partial T} \\ &= \frac{1}{Z} \left( \sum_{i=0}^n \binom{n}{i} \left( K_R^i [L]^i i \Delta H_R + \gamma K_T^i [L]^i (\Delta H_\gamma + i \Delta H_T) \right) \right) \\ &= \sum_{i=0}^n F_{RL_i} i \Delta H_R + \sum_{i=0}^n F_{TL_i} (\Delta H_\gamma + i \Delta H_T) \end{aligned} \quad (7)$$

where  $\Delta H_R$  and  $\Delta H_T$  are the site-specific microscopic intrinsic ligand-binding enthalpies for a binding site in a protein subunit in the R and T states, respectively, and  $\Delta H_\gamma$  is the enthalpy associated to the concerted conformational change between the R and T states.

The binding equations corresponding to the binding equilibrium derive from the mass conservation and the chemical equilibrium

$$\begin{aligned} [P]_T &= [P]Z \\ [L]_T &= [L] + [P]_T n_{LB} = [L] + [P]_T \frac{\partial \ln Z}{\partial \ln [L]} \end{aligned} \quad (8)$$

The last equation can be transformed into a  $(n + 1)$ th-degree polynomial equation in  $[L]$  (in this case, a 15th degree polynomial equation) with coefficients that are functions of  $K_R$ ,  $K_T$ ,  $\gamma$ ,  $[P]_T$ , and  $[L]_T$  and that can be solved numerically (e.g., Newton-Raphson method) for the unknown  $[L]$ . Equation 8 must be solved for each experimental point in the calorimetric titration (i.e., each ligand injection  $j$ ), for which the total concentrations of protein and ligand after each injection  $j$  are calculated as follows

$$\begin{aligned} [P]_{T,j} &= [P]_0 \left( 1 - \frac{v}{V_0} \right)^j \\ [L]_{T,j} &= [L]_0 \left( 1 - \left( 1 - \frac{v}{V_0} \right)^j \right) \end{aligned} \quad (9)$$

where  $[P]_0$  is the initial macromolecule concentration in the cell,  $[L]_0$  is the ligand concentration in the syringe,  $V_0$  is the calorimetric cell volume, and  $v$  is the injection volume. Once the free ligand concentration is known, the concentration of each complex after each injection can be calculated (subscript  $j$  omitted for the sake of clarity)

$$\begin{aligned} [RL_i] &= [P]_T F_{RL_i} = [P]_T \frac{\binom{n}{i} K_R^i [L]^i}{Z} \\ [TL_i] &= [P]_T F_{TL_i} = [P]_T \frac{\binom{n}{i} \gamma K_T^i [L]^i}{Z} \end{aligned} \quad (10)$$

The heat effect,  $q_j$ , associated with each injection  $j$  is calculated, considering that it reflects the change in the average excess molar binding enthalpy or in the concentration of all complexes in the calorimetric cell between injection  $j$  and  $j - 1$

$$\begin{aligned} q_j &= V_0 \left( [P]_{T,j} \langle \Delta H \rangle_j - [P]_{T,j-1} \langle \Delta H \rangle_{j-1} \left( 1 - \frac{v}{V_0} \right) \right) \\ &= V_0 \left( [P]_{T,j} \frac{1}{(1 + K_R [L]_j)^n + \gamma (1 + K_T [L]_j)^n} \left( n K_R [L]_j (1 + K_R [L]_j)^{n-1} \right. \right. \\ &\quad \left. \left. \Delta H_R + \gamma (1 + K_T [L]_j)^n \Delta H_\gamma + n \gamma K_T [L]_j (1 + K_T [L]_j)^{n-1} \Delta H_T \right) - \right. \\ &\quad \left. [P]_{T,j-1} \frac{1}{(1 + K_R [L]_{j-1})^n + \gamma (1 + K_T [L]_{j-1})^n} \left( n K_R [L]_{j-1} (1 + K_R [L]_{j-1})^{n-1} \right. \right. \\ &\quad \left. \left. \Delta H_R + \gamma (1 + K_T [L]_{j-1})^n \Delta H_\gamma + n \gamma K_T [L]_{j-1} (1 + K_T [L]_{j-1})^{n-1} \Delta H_T \right) \left( 1 - \frac{v}{V_0} \right) \right) \\ &= V_0 \left( [P]_{T,j} \sum_{i=0}^n \left( F_{RL_i,j} i \Delta H_R + F_{TL_i,j} (\Delta H_\gamma + i \Delta H_T) \right) - \right. \\ &\quad \left. [P]_{T,j-1} \sum_{i=0}^n \left( F_{RL_i,j-1} i \Delta H_R + F_{TL_i,j-1} (\Delta H_\gamma + i \Delta H_T) \right) \left( 1 - \frac{v}{V_0} \right) \right) \\ &= V_0 \left( \sum_{i=0}^n \left( i \Delta H_R \left( [RL_i]_j - [RL_i]_{j-1} \left( 1 - \frac{v}{V_0} \right) \right) + \right. \right. \\ &\quad \left. \left. (\Delta H_\gamma + i \Delta H_T) \left( [TL_i]_j - [TL_i]_{j-1} \left( 1 - \frac{v}{V_0} \right) \right) \right) \right) \end{aligned} \quad (11)$$

Last,  $q_j$  is normalized by the amount of the ligand injected during each injection, and an adjustable parameter  $q_d$  accounting for the background injection heat (due to solution mismatch, turbulence, etc.) is also included

$$Q_j = \frac{q_j}{v[L]_0} + q_d \quad (12)$$

In addition, a normalizing parameter  $N$  is included in Eq. 9 by multiplying  $[P]_0$  to account for the active or binding-competent fraction of macromolecule (percentage of protein able to bind ligand).

Nonlinear least squares regression allows the determination of the binding parameters  $K_R$ ,  $K_T$ ,  $\gamma$ ,  $\Delta H_R$ ,  $\Delta H_T$ ,  $\Delta H_\gamma$ ,  $N$ , and  $q_d$ . A fitting routine was implemented in Origin 7.0 (OriginLab).

For the sake of simplicity, a non-normalized binding polynomial has been used in Eq. 2, where only the ligand-free R state has been taken as the reference state. A renormalized binding polynomial can be constructed taking the subensemble of ligand-free states, R and T, as the reference (ensemble) state, resulting in a standard binding polynomial with a leading term equal to 1 when grouping terms according to powers in  $[L]$ . The new binding polynomial is given by

$$Z^* = \frac{Z}{1 + \gamma} = \frac{1}{1 + \gamma} (1 + K_R[L])^n + \frac{\gamma}{1 + \gamma} (1 + K_T[L])^n \quad (13)$$

and the only difference in the subsequent development is that a different reference value for the excess average ligand-binding enthalpy is used

$$\langle \Delta H^* \rangle = \langle \Delta H \rangle - \frac{\gamma}{1 + \gamma} \Delta H_\gamma \quad (14)$$

The second term in the right-hand side of Eq. 14 after the minus sign is a constant value equal to the average excess enthalpy of the non-liganded fraction of protein (i.e., ligand-free R and T states)

$$\langle \Delta H^* \rangle = \langle \Delta H \rangle - \langle \Delta H \rangle_{R+T} \quad (15)$$

Thus, that term introduces a shift in the enthalpy scale, and an additional constant term in the average excess molar enthalpy makes no difference when calculating the heat effect for each injection, as can be easily proven

$$\begin{aligned} q_j &= V_0 \left( [P]_{T,j} \langle \Delta H \rangle_j - [P]_{T,j-1} \langle \Delta H \rangle_{j-1} \left( 1 - \frac{v}{V_0} \right) \right) \\ &= V_0 [P]_{T,j} (\langle \Delta H \rangle_j - \langle \Delta H \rangle_{j-1}) \\ &= V_0 [P]_{T,j} (\langle \Delta H^* \rangle_j - \langle \Delta H^* \rangle_{j-1}) \end{aligned} \quad (16)$$

A nested MWC cooperativity model for the behavior of ClpP in which each heptamer might undergo a specific concerted conformational change could have been applied to ClpP but that would have added three additional equilibrium constants and three additional enthalpy changes, which would result in overparameterization in the fitting function and correlation/dependency between fitting parameters. In addition, the KNF cooperativity model in which sequential conformational changes are occurring as the oligomer is be-

ing occupied by the ligand could have been applied to ClpP, but, again, more parameters should have been considered in the fitting function, resulting in overparameterization. Besides, the KNF model cannot reproduce the activation effect induced by bortezomib because, in the KNF model, the conformational change induced by the ligand is specifically restricted to those subunits binding the ligand, with no possibility of ligand-free subunits undergoing the activating conformational change. Thus, the MWC cooperativity model is the minimal model able to reproduce the behavior of ClpP.

The concentration (and the evolution along a titration) of the different protein species (R and T conformations) can be readily calculated from the previous equations. The total fractions of protein subunits in the conformation R or T,  $F_{R,T}$  and  $F_{T,T}$ , respectively, are given by

$$\begin{aligned} F_{R,T} &= \frac{(1 + K_R[L])^n}{Z} = \frac{1}{Z} \sum_{i=0}^n \binom{n}{i} K_R^i [L]^i = \sum_{i=0}^n F_{RL_i} \\ F_{T,T} &= \frac{\gamma(1 + K_T[L])^n}{Z} = \frac{\gamma}{Z} \sum_{i=0}^n \binom{n}{i} \gamma K_T^i [L]^i = \sum_{i=0}^n F_{TL_i} \end{aligned} \quad (17)$$

and the fractions of protein subunits in the conformation R or T that are bound to the ligand,  $F_{R,B}$  and  $F_{T,B}$ , respectively, are given by

$$\begin{aligned} F_{R,B} &= \frac{K_R[L](1 + K_R[L])^{n-1}}{Z} = \frac{1}{nZ} \sum_{i=0}^n \binom{n}{i} i K_R^i [L]^i = \frac{1}{n} \sum_{i=0}^n i F_{RL_i} \\ F_{T,B} &= \frac{\gamma K_T[L](1 + K_T[L])^{n-1}}{Z} = \frac{\gamma}{nZ} \sum_{i=0}^n \binom{n}{i} \gamma K_T^i [L]^i = \frac{1}{n} \sum_{i=0}^n i F_{TL_i} \end{aligned} \quad (18)$$

The fractions of ligand-free protein subunits in the conformation R or T,  $F_{R,F}$  and  $F_{T,F}$ , respectively, can be calculated as the following differences

$$\begin{aligned} F_{R,F} &= F_{R,T} - F_{R,B} = \frac{(1 + K_R[L])^{n-1}}{Z} = \sum_{i=0}^n F_{RL_i} - \frac{1}{n} \sum_{i=0}^n i F_{RL_i} \\ F_{T,F} &= F_{T,T} - F_{T,B} = \frac{\gamma(1 + K_T[L])^{n-1}}{Z} = \sum_{i=0}^n F_{TL_i} - \frac{1}{n} \sum_{i=0}^n i F_{TL_i} \end{aligned} \quad (19)$$

## MD simulations

### Computational assays

The 3D structure used in the simulations was obtained from *Tt*ClpP:peptide, which corresponds to an extended conformation of the protein, bereft of substrate. The missing N-terminal  $\beta$ -hairpin residues were added using Modeller 9.17 (60). Positioning of the seven loops was done concomitantly for each of the upper and the lower hemispheres of the protein. For the protein with bound peptide substrates, use was made of the *Tt*ClpP:peptide structure featuring stretches of electron density characteristic of short peptides, over which zwitterionic alanine tripeptides were superimposed. The computational assays consisted of the extended *Tt*ClpP, with or devoid of alanine tripeptides bound to the 14 serine active sites of the protein, immersed in a bath of 75,307 water molecules, with a 0.16 M NaCl concentration corresponding to a cubic simulation cell of length equal to 138 Å at thermodynamic equilibrium.

## MD protocol

All the MD simulations reported in this study were performed using NAMD 2.12 (61), with the CHARMM36 (62) force field for proteins and lipids, and the TIP3P water model. A Langevin thermostat with a damping coefficient of  $1 \text{ p s}^{-1}$  maintained the temperature at  $27^\circ\text{C}$ . The Langevin piston method was used to keep the computational assay at a nominal pressure of 1 atm. Covalent bonds involving hydrogen atoms were constrained to their equilibrium length by the Rattle algorithm. The Settle algorithm was used to constrain water molecules to their equilibrium geometry. Long-range electrostatic forces were evaluated using the particle mesh Ewald algorithm with a grid spacing of  $1.2 \text{ \AA}$ , while a smoothed  $12\text{-\AA}$  spherical cutoff was applied to truncate short-range van der Waals and electrostatic interactions. The r-RESPA multiple time-stepping algorithm combined with a mass-repartitioning scheme was used to integrate the equations of motion with an effective time step of 4 fs for short-range interactions and of 8 fs for long-range electrostatic interactions.

## Simulation design

The extended apo-*TtClpP* was probed in a  $1\text{-}\mu\text{s}$  MD simulation. In addition, the computational assay consisting of the protein with zwitterionic alanine tripeptides bound to its 14 serine active sites was examined in three independent,  $1\text{-}\mu\text{s}$  MD simulations, using distinct initial conditions. These simulations are aimed at exploring how the presence of the substrate affects the flexibility of *TtClpP*, alongside the relevance of the alanine tripeptide as a proxy for the peptide chain that binds to the Ser<sup>97</sup> active site.

## Near-equilibrium pulling experiments

To address how the occupancy of the serine active sites of *TtClpP* affects the compressibility of *TtClpP*, we have examined the conformational transition of the protein from its extended state to its compressed state by means of near-equilibrium pulling experiments. Toward this end, a steered MD protocol was used, pulling at constant velocity a stiff spring of force constant equal to  $100 \text{ kcal/mol \AA}$ . The transition coordinate was defined as the Euclidian distance separating the center of mass of the upper hemisphere to that of the lower hemisphere computed from the position of all  $\alpha$ -carbon atoms. The length of the transition coordinate was determined on the basis of apo-*TtClpP* and the structure of the compressed conformation (PDB: 3QWD), and amounts to  $8 \text{ \AA}$ . Three independent, 200-ns-long realizations were performed for the empty protein and three trajectories for ClpP containing either two (in one trajectory) or three (in the other trajectories) zwitterionic alanine tripeptides, pulling irreversibly the two lobes of the protein from an extended state to a compressed one.

## Multi-angle laser light scattering

SEC-MALLS experiments were conducted at room temperature on a high-performance liquid chromatography (HPLC) system consisting of a DGU-20 AD degasser, an LC-20 AD pump, a SIL20-AC<sub>HT</sub> autosampler, a CBM-20A communication interface, an SPD-M20A UV-Vis detector (Shimadzu, Kyoto, Japan), an XL-Therm column oven (WynSep, Sainte Foy d'Aigrefeuille, France), and static light scattering miniDAWN Treos, DLS DynaPro NANOSTAR, and refractive index Optilab rEX detectors. Data were analyzed using version 5.4.3.20 of the ASTRA software (Wyatt Technology, Santa Barbara, USA). *TtClpP* ( $1 \text{ mg/ml}$ ,  $100 \mu\text{l}$ ) was injected at  $0.5 \text{ ml/min}$  on a Superdex 200 10/300 GL column (GE Healthcare), equilibrated with a buffer containing  $100 \text{ mM}$  tris (pH 8.5),  $100 \text{ mM}$  NaCl, and 5% glycerol. Bovine serum

albumin at  $2 \text{ mg/ml}$  in phosphate-buffered saline (PBS) buffer was injected as a control in each experiment. The extinction coefficient and refractive index increments for the proteins were calculated from the amino acid sequences using the program SEDFIT.

## Dynamic light scattering

*TtClpP* ( $3 \mu\text{M}$ ) in a buffer containing  $100 \text{ mM}$  tris (pH 8.5),  $100 \text{ mM}$  NaCl, and 5% glycerol without or with  $5 \text{ mM}$  added bortezomib was measured using a DynaPro NANOSTAR (Wyatt, Santa Barbara, USA) instrument operated at  $25^\circ$  or  $45^\circ\text{C}$  and analyzed with the associated DYNAMICS software. Each experiment was performed in triplicate (three times for each of two *TtClpP* samples, both measured at  $25^\circ$  and  $45^\circ\text{C}$ ), and hydrodynamic radii were calculated as the average of these replicates. Theoretical hydrodynamic radii of compressed *TtClpP* [based on PDB: 3QWD (18)] and extended *TtClpP* (using the *TtClpP*:bortezomib structure), after modeling of missing residues in the N-terminal  $\beta$ -hairpins and the N and C termini using SWISS-MODEL (63), were calculated using Hydropro10 (64) using a solvent viscosity and a solvent density of  $0.0010387 \text{ poise}$  and  $1.0114 \text{ g/cm}^3$ , respectively (5% glycerol,  $25^\circ\text{C}$ ), or  $0.00068656 \text{ poise}$  and  $1.0043 \text{ g/cm}^3$ , respectively (5% glycerol,  $45^\circ\text{C}$ ).

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/5/9/eaaw3818/DC1>

Fig. S1. Biochemical characterization of *TtClpP*.

Fig. S2. Characterization of GFPsra degradation by *TtClpP*.

Fig. S3. Characterization of *TtClpP* peptidase activity in the absence and presence of bortezomib.

Fig. S4. Characterization of FITC-casein degradation by *TtClpP* in the absence and presence of bortezomib and/or *TtClpP*.

Fig. S5. Ligand validation by  $2\text{mF}_o - \text{DF}_c$ , omit, and polder maps.

Fig. S6. Distance RMSD of the 14 alanine tripeptide ligands with respect to their initial bound state in *TtClpP* for three independent,  $1\text{-}\mu\text{s}$ -long MD simulations.

Fig. S7. Analysis of structural differences between *TtClpP* and *EcClpP*.

Table S1. Crystallographic data collection and refinement statistics.

Reference (65)

## REFERENCES AND NOTES

- R. M. Raju, A. L. Goldberg, E. J. Rubin, Bacterial proteolytic complexes as therapeutic targets. *Nat. Rev. Drug Discov.* **11**, 777–789 (2012).
- J. Wang, J. A. Hartling, J. M. Flanagan, The structure of ClpP at  $2.3 \text{ \AA}$  resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**, 447–456 (1997).
- T. Akopian, O. Kandror, R. M. Raju, M. Unnikrishnan, E. J. Rubin, A. L. Goldberg, The active ClpP protease from *M. tuberculosis* is a complex composed of a heptameric ClpP1 and a ClpP2 ring. *EMBO J.* **31**, 1529–1541 (2012).
- C. Fetzter, V. S. Korotkov, R. Thänert, K. M. Lee, M. Neuenschwander, J. P. von Kries, E. Medina, S. A. Sieber, A chemical disruptor of the ClpX chaperone complex attenuates the virulence of multidrug-resistant *Staphylococcus aureus*. *Angew. Chem. Int. Ed. Engl.* **56**, 15746–15750 (2017).
- O. Gaillot, E. Pellegrini, S. Bregenholt, S. Nair, P. Berche, The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol. Microbiol.* **35**, 1286–1294 (2000).
- S. Bhaskaran, G. Pharaoh, R. Ranjit, A. Murphy, S. Matsuzaki, B. C. Nair, B. Forbes, S. Gispert, G. Auburger, K. M. Humphries, M. Kinter, T. M. Griffin, S. S. Deepa, Loss of mitochondrial protease ClpP protects mice from diet-induced obesity and insulin resistance. *EMBO Rep.* **19**, e45009 (2018).
- A. Cole, Z. Wang, E. Coyaud, V. Voisin, M. Gronda, Y. Jitkova, R. Mattson, R. Hurren, S. Babovic, N. Maclean, I. Restall, X. Wang, D. V. Jeyaraju, M. A. Sukhai, S. Prabha, S. Bashir, A. Ramakrishnan, E. Leung, Y. H. Qia, N. Zhang, K. R. Combes, T. Ketela, F. Lin, W. A. Houry, A. Aman, R. Al-awar, W. Zheng, E. Wienholds, C. J. Xu, J. Dick, J. C. Y. Wang, J. Moffat, M. D. Minden, C. J. Eaves, G. D. Bader, Z. Hao, S. M. Kornblau, B. Raught, A. D. Schimmer, Inhibition of the mitochondrial protease ClpP as a therapeutic strategy for human acute myeloid leukemia. *Cancer Cell* **27**, 864–876 (2015).



8. S. Rathore, D. Sinha, M. Asad, T. Böttcher, F. Afrin, V. S. Chauhan, D. Gupta, S. A. Sieber, A. Mohammed, A cyanobacterial serine protease of *Plasmodium falciparum* is targeted to the apicoplast and plays an important role in its growth and development. *Mol. Microbiol.* **77**, 873–890 (2010).
9. H. Brötz-Oesterheld, D. Beyer, H.-P. Kroll, R. Endermann, C. Ladel, W. Schroeder, B. Hinzen, S. Raddatz, H. Paulsen, K. Henninger, J. E. Bandow, H.-G. Sahl, H. Labischinski, Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat. Med.* **11**, 1082–1087 (2005).
10. K. Famulla, P. Sass, I. Malik, T. Akopian, A. Kandror, M. Alber, B. Hinzen, H. Ruebsamen-Schaeff, R. Kalscheuer, A. L. Goldberg, H. Brötz-Oesterheld, Acyldepsipeptide antibiotics kill mycobacteria by preventing the physiological functions of the ClpP1P2 protease. *Mol. Microbiol.* **101**, 194–209 (2016).
11. K. Weinhäupl, M. Brennich, U. Kazmaier, J. Lelievre, L. Ballell, A. Goldberg, P. Schanda, H. Fraga, The antibiotic cyclomarin blocks arginine-phosphate-induced millisecond dynamics in the N-terminal domain of ClpC1 from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **293**, 8379–8393 (2018).
12. N. Benaroudj, B. Raynal, M. Miot, M. Ortiz-Lombardia, Assembly and proteolytic processing of mycobacterial ClpP1 and ClpP2. *BMC Biochem.* **12**, 61 (2011).
13. H. Ingvarsson, M. J. Maté, M. Högbom, D. Portnoi, N. Benaroudj, P. M. Alzari, M. Ortiz-Lombardia, T. Unge, Insights into the inter-ring plasticity of caseinolytic proteases from the X-ray structure of *Mycobacterium tuberculosis* ClpP1. *Acta Crystallogr. D* **63**, 249–259 (2007).
14. M. Li, O. Kandror, T. Akopian, P. Dharkar, A. Wlodawer, M. R. Maurizi, A. L. Goldberg, Structure and functional properties of the active form of the proteolytic complex, ClpP1P2, from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **291**, 7465–7476 (2016).
15. K. R. Schmitz, D. W. Carney, J. K. Sello, R. T. Sauer, 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–E4595 (2014).
16. S. Vahidi, Z. A. Ripstein, M. Bonomi, T. Yuwen, M. F. Mabanglo, J. B. Juravsky, K. Rizzolo, A. Velyvis, W. A. Houry, M. Vendruscolo, J. L. Rubinstein, L. E. Kay, Reversible inhibition of the ClpP protease via an N-terminal conformational switch. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E6447–E6456 (2018).
17. R. Sprangers, A. Gribun, P. M. Hwang, W. A. Houry, L. E. Kay, Quantitative NMR spectroscopy of supramolecular complexes: Dynamic side pores in ClpP are important for product release. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16678–16683 (2005).
18. S. R. Geiger, T. Böttcher, S. A. Sieber, P. Cramer, A conformational switch underlies ClpP protease function. *Angew. Chem. Int. Ed. Engl.* **50**, 5749–5752 (2011).
19. J. Zhang, F. Ye, L. Lan, H. Jiang, C. Luo, C. G. Yang, Structural switching of *Staphylococcus aureus* Clp protease. *J. Biol. Chem.* **286**, 37590–37601 (2011).
20. N. A. Wood, K. Y. Chung, A. M. Blocker, N. Rodrigues de Almeida, M. Conda-Sheridan, D. J. Fisher, S. P. Ouellette, Initial characterization of the two ClpP paralogs of *Chlamydia trachomatis* suggests unique functionality for each. *J. Bacteriol.* **201**, e00635–18 (2019).
21. M. Dahmen, M.-T. Vielberg, M. Groll, S. A. Sieber, Structure and mechanism of the caseinolytic protease ClpP1/2 heterocomplex from *Listeria monocytogenes*. *Angew. Chem. Int. Ed.* **54**, 3598–3602 (2015).
22. K. R. Schmitz, R. T. Sauer, Substrate delivery by the AAA+ ClpX and ClpC1 unfoldases activates the mycobacterial ClpP1P2 peptidase. *Mol. Microbiol.* **93**, 617–628 (2014).
23. M. Gersch, K. Famulla, M. Dahmen, C. Göbl, I. Malik, K. Richter, V. S. Korotkov, P. Sass, H. Rübsamen-Schaeff, T. Madl, H. Brötz-Oesterheld, S. A. Sieber, AAA+ chaperones and acyldepsipeptides activate the ClpP protease via conformational control. *Nat. Commun.* **6**, 6320 (2015).
24. M. Groll, C. R. Berkers, H. L. Ploegh, H. Ovaa, Crystal structure of the boronic acid-based proteasome inhibitor bortezomib in complex with the yeast 20S proteasome. *Structure* **14**, 451–456 (2006).
25. W. Moreira, S. Santhanakrishnan, B. W. Dymock, T. Dick, Bortezomib warhead-switch confers dual activity against mycobacterial caseinolytic protease and proteasome and selectivity against human proteasome. *Front. Microbiol.* **8**, 333–336 (2017).
26. D. Balogh, M. Dahmen, M. Stahl, M. Poreba, M. Gersch, M. Drag, S. A. Sieber, Insights into ClpXP proteolysis: Heterooligomerization and partial deactivation enhance chaperone affinity and substrate turnover in *Listeria monocytogenes*. *Chem. Sci.* **8**, 1592–1600 (2017).
27. S. A. Joshi, G. L. Hersch, T. A. Baker, R. T. Sauer, Communication between ClpX and ClpP during substrate processing and degradation. *Nat. Struct. Mol. Biol.* **11**, 404–411 (2004).
28. S. K. Singh, F. Guo, M. R. Maurizi, ClpA and ClpP remain associated during multiple rounds of ATP-dependent protein degradation by ClpAP protease. *Biochemistry* **38**, 14906–14915 (1999).
29. P. Sass, M. Josten, K. Famulla, G. Schiffer, H. G. Sahl, L. Hamoen, H. Brotz-Oesterheld, Antibiotic acyldepsipeptides activate ClpP peptidase to degrade the cell division protein FtsZ. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17474–17479 (2011).
30. M. E. Lee, T. A. Baker, R. T. Sauer, Control of substrate gating and translocation into ClpP by channel residues and ClpX binding. *J. Mol. Biol.* **399**, 707–718 (2010).
31. M. W. Thompson, S. K. Singh, M. R. Maurizi, 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. *J. Biol. Chem.* **269**, 18209–18215 (1994).
32. P. M. Vitali Tugarinov, A. Hwang, E. Jason, L. E. K. Ollerenshaw, Cross-Correlated Relaxation Enhanced <sup>1</sup>H–<sup>13</sup>C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.* **125**, 10420–10428 (2003).
33. H. Fraga, C.-A. Arnaud, D. F. Gauto, M. Audin, V. Kurauskas, P. Macek, C. Krichel, J.-Y. Guan, J. Boisbouvier, R. Sprangers, C. Breyton, P. Schanda, Solid-State NMR H-N-(C)-H and H-N-C-C 3D/4D correlation experiments for resonance assignment of large proteins. *ChemPhysChem* **18**, 2697–2703 (2017).
34. F. Ye, J. Zhang, H. Liu, R. Hilgenfeld, R. Zhang, X. Kong, L. Li, J. Lu, X. Zhang, D. Li, H. Jiang, C.-G. Yang, C. Luo, Helix unfolding/refolding characterizes the functional dynamics of *Staphylococcus aureus* Clp protease. *J. Biol. Chem.* **288**, 17643–17653 (2013).
35. M. S. Kimber, A. Y. H. Yu, M. Borg, E. Leung, H. S. Chan, W. A. Houry, Structural and theoretical studies indicate that the cylindrical protease ClpP samples extended and compact conformations. *Structure* **18**, 798–808 (2010).
36. D. H. S. Li, Y. S. Chung, M. Gloyd, E. Joseph, R. Ghirlando, G. D. Wright, Y.-Q. Cheng, M. R. Maurizi, A. Guarné, J. Ortega, 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**, 959–969 (2010).
37. A. Szyk, M. R. Maurizi, Crystal structure at 1.9 Å of *E. coli* ClpP with a peptide covalently bound at the active site. *J. Struct. Biol.* **156**, 165–174 (2006).
38. J. A. Alexopoulos, A. Guarné, J. Ortega, ClpP: A structurally dynamic protease regulated by AAA+ proteins. *J. Struct. Biol.* **179**, 202–210 (2012).
39. T. L. Religa, A. M. Ruschak, R. Rosenzweig, L. E. Kay, Site-directed methyl group labeling as an NMR probe of structure and dynamics in supramolecular protein systems: Applications to the proteasome and to the ClpP protease. *J. Am. Chem. Soc.* **133**, 9063–9068 (2011).
40. D. Y. Kim, K. K. Kim, The structural basis for the activation and peptide recognition of bacterial ClpP. *J. Mol. Biol.* **379**, 760–771 (2008).
41. P. Schanda, M. Ernst, Studying dynamics by magic-angle spinning solid-state NMR spectroscopy: Principles and applications to biomolecules. *Prog. Nucl. Magn. Reson. Spectrosc.* **96**, 1–46 (2016).
42. J. C. Gerhart, A. B. Pardee, The effect of the feedback inhibitor CTP, on subunit interactions in aspartate transcarbamylase. *Cold Spring Harb. Symp. Quant. Biol.* **28**, 491–496 (1963).
43. B. G. Barisais, S. J. Gill, Thermodynamic analysis of carbon monoxide binding by hemoglobin trout I. *Biophys. Chem.* **9**, 235–244 (1979).
44. J. Monod, J. Wyman, J.-P. Changeux, On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **12**, 88–118 (1965).
45. T. Akopian, O. Kandror, C. Tsu, J. H. Lai, W. Wu, Y. Liu, P. Zhao, A. Park, L. Wolf, L. R. Dick, E. J. Rubin, W. Bachovchin, A. L. Goldberg, Cleavage Specificity of *Mycobacterium tuberculosis* ClpP1P2 protease and identification of novel peptide substrates and boronate inhibitors with anti-bacterial activity. *J. Biol. Chem.* **290**, 11008–11020 (2015).
46. W. Kress, H. Mutschler, E. Weber-Ban, Assembly pathway of an AAA+ protein: Tracking ClpA and ClpAP complex formation in real time. *Biochemistry* **46**, 6183–6193 (2007).
47. E. Zeiler, A. List, F. Alte, M. Gersch, R. Wachtel, M. Poreba, M. Drag, M. Groll, S. A. Sieber, Structural and functional insights into caseinolytic proteases reveal an unprecedented regulation principle of their catalytic triad. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11302–11307 (2013).
48. U. A. Wojtyra, G. Thibault, A. Tuite, W. A. Houry, The N-terminal zinc binding domain of ClpX is a dimerization domain that modulates the chaperone function. *J. Biol. Chem.* **278**, 48981–48990 (2003).
49. D. M. Smith, H. Fraga, C. Reis, G. Kafri, A. L. Goldberg, ATP binds to proteasomal ATPases in pairs with distinct functional effects, implying an ordered reaction cycle. *Cell* **144**, 526–538 (2011).
50. E. Leung, A. Datti, M. Cossette, J. Goodreid, S. E. McCaw, M. Mah, A. Nakhamchik, K. Ogata, M. E. Bakouri, Y.-Q. Cheng, S. J. Wodak, B. T. Eger, E. F. Pai, J. Liu, S. Gray-Owen, R. A. Batey, W. A. Houry, Activators of cylindrical proteases as antimicrobials: Identification and development of small molecule activators of ClpP protease. *Chem. Biol.* **18**, 1167–1178 (2011).
51. W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **59**, 687–696 (2005).
52. C. Amero, P. Schanda, M. A. Durá, I. Ayala, D. Marion, B. Franzetti, B. Brutscher, J. Boisbouvier, Fast two-dimensional NMR spectroscopy of high molecular weight protein assemblies. *J. Am. Chem. Soc.* **131**, 3448–3449 (2009).
53. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
54. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
55. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner,

- R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *PHENIX: A comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
56. B.-G. Lee, E. Y. Park, K.-E. Lee, H. Jeon, K. H. Sung, H. Paulsen, H. Rübsamen-Schaeff, H. Brötz-Oesterhelt, H. K. Song, Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat. Struct. Mol. Biol.* **17**, 471–478 (2010).
57. R. Schwarzenbacher, A. Godzik, S. K. Grzechnik, L. Jaroszewski, The importance of alignment accuracy for molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 1229–1236 (2004).
58. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
59. J. Wyman, S. J. Gill, *Binding and Linkage: Functional Chemistry of Biological Macromolecules* (University Science Books, 1990).
60. A. Šali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
61. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
62. A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
63. A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, *SWISS-MODEL: Homology modelling of protein structures and complexes. Nucleic Acids Res.* **46**, W296–W303 (2018).
64. A. Ortega, D. Amorós, J. García de la Torre, Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101**, 892–898 (2011).
65. P. A. Karplus, K. Diederichs, Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).

#### Acknowledgments

**Funding:** This work used the platforms of the Grenoble Instruct center (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from INSTRUCT (“Innovative EM/NMR approach for the characterization of the drug target ClpP APPID: 301”), FRISBI (ANR-10-INSB-05-02), and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). We

thank A. Le Roi and C. Ebel for assistance with MALLS and DLS data acquisition. Special thanks to T. Hosek (Institute Biologie Structurale) for FtsZ and T. Akopian and A. Goldberg (Harvard School of Public Health, USA) for testing peptide substrates. Activation of *MtbClpP1P2* by bortezomib was first observed by T. Akopian and A. Goldberg. We thank J. Boissouvier (IBS Grenoble) for hosting H.F. in his team for a related (previously published) work. We thank the ESRF for beamtime at ID30A and ID23-1. **Funding:** This work was supported by Spanish Ministerio de Economía y Competitividad (BFU2016-78232-P) and Instituto de Salud Carlos III co-funded by European Union (P115/00663 and P118/00349, ERDF/ESF, “Investing in your future”). This work was financially supported by the European Research Council (ERC-Stg-2012-311318 to P.S.). J.F. is supported by an EMBO long-term post-doctoral fellowship (ALTF441-2017). **Author contributions:** J.F. prepared ClpP + inhibitor crystals, processed x-ray diffraction data, refined and analyzed crystal structures, performed MALLS and DLS experiments, prepared figures, and wrote the manuscript. K.W. performed solution NMR experiments, discussed results, and prepared figures. C.C. and F.D. designed, performed, and analyzed the MD simulations and prepared figures. A.H. performed site-directed mutagenesis and protein sample preparation. D.F.G. performed modeling and computational simulations proof of concepts and prepared figures. C.M. prepared ClpP crystals and collected x-ray diffraction data. O.A. performed and analyzed ITC experiments. I.G. designed research and discussed results. A.V.-C. performed and analyzed ITC experiments, prepared figures, discussed results, and wrote the manuscript. P.S. performed and analyzed NMR experiments, designed research, managed and prepared figures, discussed results, and wrote the manuscript. H.F. performed biochemical experiments and sample preparation for NMR and x-ray, managed and prepared figures, discussed results, and wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors. Coordinates and structure factors for *TtClpP* in complex with peptide or bortezomib have been deposited in the PDB with accession codes 6HWM and 6HWN, respectively.

Submitted 18 December 2018

Accepted 2 August 2019

Published 4 September 2019

10.1126/sciadv.aaw3818

**Citation:** J. Felix, K. Weinhäupl, C. Chipot, F. Dehez, A. Hessel, D. F. Gauto, C. Morlot, O. Abian, I. Gutsche, A. Velazquez-Campoy, P. Schanda, H. Fraga, Mechanism of the allosteric activation of the ClpP protease machinery by substrates and active-site inhibitors. *Sci. Adv.* **5**, eaaw3818 (2019).