Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of an ATP-independent Lon-like protease and its complexes with covalent inhibitors

Jiahn-Haur Liao,[a] Kentaro Ihara,[b]‡ Chiao-I Kuo,[a] Kai-Fa Huang,[a] Soichi Wakatsuki,[b]§ Shih-Hsiung Wu[a,c]* and Chung-I Chang[a,c]*

[a]Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan, [b]Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan, and [c]Institute of Biochemical Sciences, College of Life Science, National Taiwan University, Taipei 10617, Taiwan

‡ Present address: Cell Biology, Graduate School of Medicine, Kyoto University, Kyoto 606-8501, Japan.
§ Present address: SLAC National Accelerator Laboratory, CA 94025 and School of Medicine, Stanford University, Beckman Center B105, 279 Campus Drive, Stanford, CA 94305, USA.

Correspondence e-mail: shwu@gate.sinica.edu.tw, chungi@gate.sinica.edu.tw

Received 26 February 2013
Accepted 25 March 2013

PDB References: MtaLonC, 4fw9; complex with bortezomib, 4fwd; complex with lactacystin, 4fwg; complex with MG262, 4fwh

The Lon proteases are a unique family of chambered proteases with a built-in AAA+ (ATPases associated with diverse cellular activities) module. Here, crystal structures of a unique member of the Lon family with no intrinsic ATPase activity in the proteolytically active form are reported both alone and in complexes with three covalent inhibitors: two peptidomimetics and one derived from a natural product. This work reveals the unique architectural features of an ATP-independent Lon that selectively degrades unfolded protein substrates. Importantly, these results provide mechanistic insights into the recognition of inhibitors and polypeptide substrates within the conserved proteolytic chamber, which may aid the development of specific Lon-protease inhibitors.

1. Introduction

Many aspects of life rely on the function of proteins, most of which are only marginally stable. Protein quality-control machinery is implemented in all living organisms to assist in either the degradation or the refolding of defective proteins, the deposition of which may lead to the loss or dysfunction of cells (Wickner *et al.*, 1999). ATP-dependent proteases are part of the protein quality-control machinery responsible for degrading misfolded and abnormal proteins in cells. Similar to the eukaryotic proteasome and the bacterial HslUV, they form oligomeric assemblies with a hollow chamber in which the proteolytic sites reside; the sites are only accessible through a narrow axial pore (Bochtler *et al.*, 2000; Groll *et al.*, 1997). The enzymatic operation consists of substrate recognition, threading of unfolded substrate polypeptides through the pore to the chamber, often in an ATP-dependent manner, and proteolysis of the substrates inside the chamber (Sauer & Baker, 2011).

The Lon family of proteases contain an AAA+ (ATPases associated with diverse cellular activities) module and a unique C-terminal protease domain in a single polypeptide chain of 700–1000 amino acids (Rotanova *et al.*, 2006). They form a functional assembly as a homo-oligomer similar to FtsH, with the proteolytic chamber encased by the fused AAA+ and protease rings (Suno *et al.*, 2006; Bieniossek *et al.*, 2006). Lon proteases are ubiquitously found in bacteria, archaea and the organelles of eukaryotes, including mitochondria, peroxisomes and chloroplasts. Lon proteases not only play a role in maintaining proteostasis, but are also involved in cellular and metabolic regulation by degrading specific protein substrates (Granot *et al.*, 2007; Mizusawa & Gottesman, 1983). A species-specific Lon-insertion domain

(LID) fused to the AAA+ module is characteristic of Lons. For example, bacterial and eukaryotic Lons (LonA) have a LID at the N-terminus of the AAA+ module (known as the LAN), whereas archaeal Lons (LonB) have a LID inserted within the AAA+ module in a series of transmembrane segments known as the membrane-anchoring region (MA; Rotanova *et al.*, 2004; Fig. 1*a*). Recently, we and others have characterized LonC, a new type of Lon-like protease with no ATPase activity (Maehara *et al.*, 2007; Liao *et al.*, 2012) previously identified in many bacterial species (Iyer *et al.*, 2004; Rotanova *et al.*, 2004). LonC contains a C-terminal Lon protease domain and a LID embedded within an AAA-like module similar to archaeal LonB, except that the LID is predicted to contain coiled-coil regions rather than transmembrane segments (Fig. 1*a*; Iyer *et al.*, 2004). Bacterial LonA has emerged as a novel target for development of antibiotic compounds because it confers virulence to many pathogenic bacteria (Frase *et al.*, 2006; Frase & Lee, 2007; Lee & Suzuki, 2008). Human mitochondrial LonA (huLon) selectively degrades oxidized proteins (Bota & Davies, 2002). Recently, it has been demonstrated that huLon is up-regulated in cancer cells and that its knockdown leads to lymphoma cell death; therefore, huLon may represent a new anticancer drug target (Bernstein *et al.*, 2012).

To date, crystal structures are available for isolated domains of Lon (Botos *et al.*, 2005; Li *et al.*, 2010; Botos, Melnikov, Cherry, Khalatova *et al.*, 2004; Botos, Melnikov, Cherry, Tropea *et al.*, 2004; Duman & Löwe, 2010; García-Nafría *et al.*, 2010; Im *et al.*, 2004; Li *et al.*, 2005) and for truncated AAA-protease dual-domain constructs without the LID (Duman & Löwe, 2010; Cha *et al.*, 2010). Notably, the latter structures have provided significant insights into the overall architectural features of the sequestered proteolytic chamber of Lon. However, in all of these structures the proteolytic sites were either modified to make them inactive (Botos, Melnikov, Cherry, Tropea *et al.*, 2004; Cha *et al.*, 2010) or were in a blocked conformation (Botos *et al.*, 2005; Botos, Melnikov, Cherry, Tropea *et al.*, 2004; Duman & Löwe, 2010; García-Nafría *et al.*, 2010; Im *et al.*, 2004; Li *et al.*, 2005). Consequently, the structural basis of how substrates or inhibitors are recognized by the proteolytic active sites within the proteolytic chamber has remained elusive. We have recently characterized

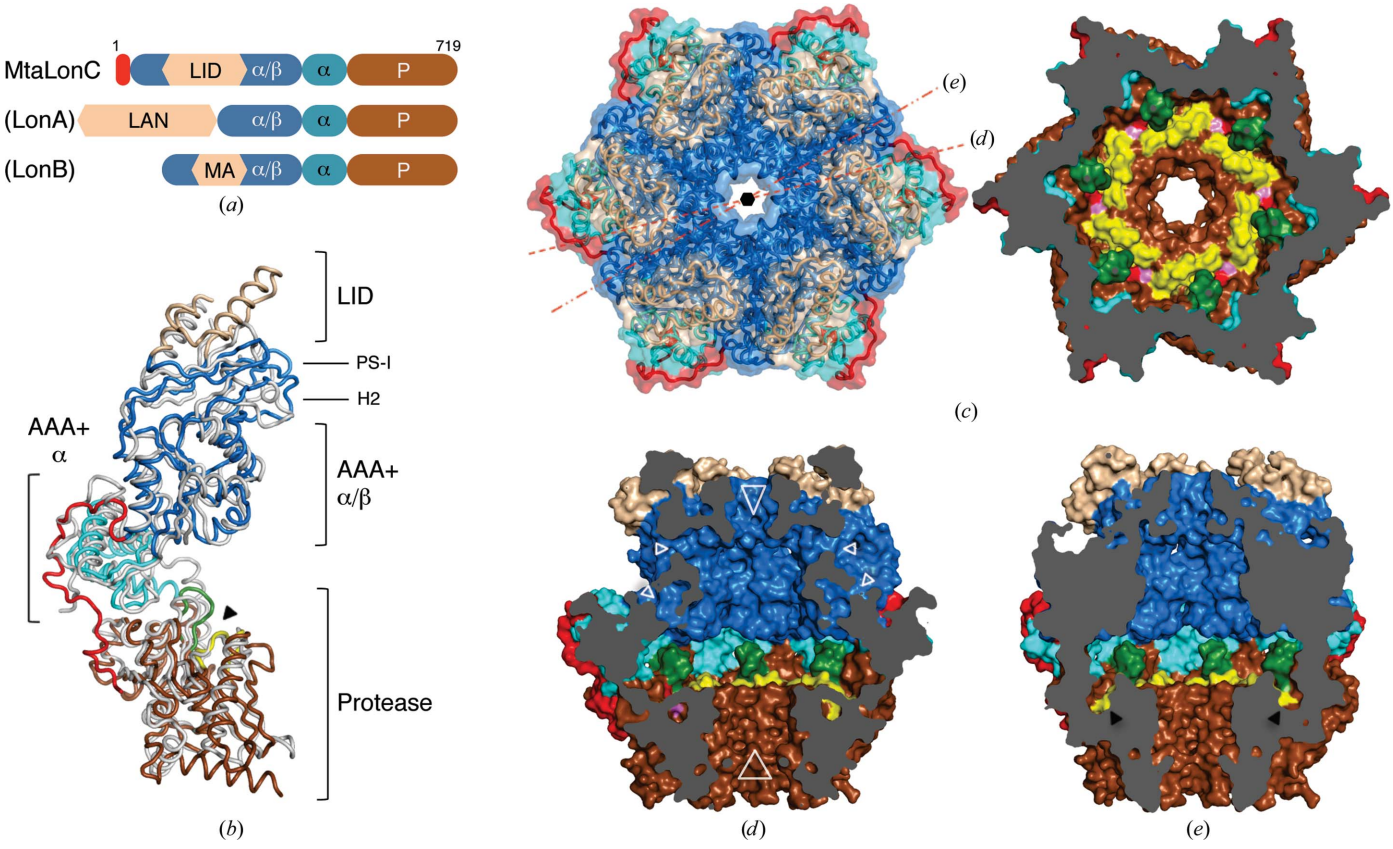

Figure 1
Overall structure and proteolytic chamber of MtaLonC. (*a*) Domain organization in the primary structures of the MtaLonC, LonA and LonB proteases. The N-terminal coil of MtaLonC is coloured red. (*b*) The MtaLonC monomer as a tube diagram, with similar colouring as in (*a*), superimposed with TonLonB (PDB entry 3k1j, chain *A*; grey). The substrate-binding groove between the $\beta1$–$\beta2$ hairpin (green) and the $\beta4$–$\alpha3$ loop (yellow) is indicated by the arrowhead. (*c*) View of the top axial pore (13 Å in diameter) of MtaLonC in ribbon representation with a semi-transparent surface (left); domains are coloured as in (*b*). Dashed lines represent the directions of the cutting planes to produce the views in (*d*) and (*e*). A cutaway top view is shown on the right, revealing the six proteolytic sites and the bottom axial opening (15 Å in diameter) formed by six protease domains. The catalytic dyad, noncatalytic Asp581, $\beta1$–$\beta2$ hairpin and $\beta4$–$\alpha3$ loop (see text) are coloured red, magenta, green and yellow, respectively. (*d*, *e*) Two cutaway side views of the proteolytic chamber. Open arrowheads indicate openings to the surface; black arrowheads mark the substrate-binding grooves.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell. All crystals belonged to space group *P*6.

	SeMet MtaLonC (apo)	MtaLonC–bortezomib	MtaLonC–lactacystin	MtaLonC–MG262
Data collection				
Wavelength (Å)	0.98038			
Resolution range (Å)	200–2.00 (2.03–2.00)	50–2.02 (2.09–2.02)	20–1.94 (2.01–1.94)	50–2.10 (2.18–2.10)
Unit-cell parameters (Å)				
$a = b$	115.7	115.9	116.0	116.0
c	136.8	135.9	135.1	136.1
Total observations	1604929	498869	643407	603203
Unique reflections	70004 (3514)	65537 (6268)	75772 (7603)	59267 (5673)
Multiplicity	22.9 (22.8)	7.6 (5.6)	8.5 (7.0)	10.2 (9.3)
Completeness (%)	99.9 (100.0)	97.1 (93.3)	99.5 (100.0)	97.9 (94.6)
$\langle I/\sigma(I)\rangle$	43.1 (6.3)	24.6 (2.2)	13.4 (2.7)	40.1 (4.4)
$R_{\rm merge}$ (%)	9.1 (67.0)	7.5 (72.9)	16.0 (78.4)	9.1 (85.2)
Refinement				
Resolution (Å)	50–2.00 (2.05–2.00)	30–2.03 (2.08–2.03)	20–1.99 (2.04–1.99)	30–2.19 (2.25–2.19)
Reflections [$>0\sigma(F)$], working/test set	62887/3531	58229/3308	63099/3479	47307/2609
R factor/$R_{\rm free}$	0.183/0.218	0.170/0.195	0.229/0.259	0.174/0.227
R.m.s.d., bond lengths (Å)	0.027	0.029	0.030	0.025
R.m.s.d., angles (°)	2.5	2.7	2.7	2.4
No. of atoms				
Protein	4633	4514	4575	4471
Substrate/inhibitor	NA	28	15	35
Phosphate ion	5	5	5	5
Water	405	395	583	364
Average B factor (Å^2)				
Protein	42.2	40.4	35.9	43.9
Substrate/inhibitor	NA	49.1	24.3	40.4
Phosphate ion	30.7	31.8	24.6	35.5
Water	46.0	46.0	46.9	50.4
Ramachandran plot, residues in (%)				
Most favoured regions	90.9	89.9	91.2	89.8
Additionally allowed regions	8.3	9.3	8.4	9.6
Generously allowed regions	0.8	0.8	0.4	0.6
Disallowed regions	0	0	0	0
Disordered regions (residue Nos.)	109–217, 713–726	109–218, 227–228, 233–236, 245–248, 286–289, 342, 713–726	111–218, 242–245, 286–289, 711–726	102–217, 228, 233–236, 245–249, 286–289, 245–249, 286–289, 342, 245–249, 286–289, 713–726
PDB code	4fw9	4fwd	4fwg	4fwh

MtaLonC, a soluble hexameric Lon-like protease from *Meiothermus taiwanensis* with an intrinsically inactive ATPase module and a selective proteolytic activity against protein substrates that are either unfolded or unstructured (Liao *et al.*, 2012). In the absence of ATP and Mg^{2+}, which are required for the proteolytic activity of LonA/B, MtaLonC is able to cleave F-β20-Q, a fluorogenic polypeptide substrate derived from an internal segment of β-galactosidase (Gur & Sauer, 2008; Liao *et al.*, 2012). Here, we show that the cleavage of the polypeptide substrate by MtaLonC can be inhibited by the proteasome inhibitors bortezomib (Velcade), MG262 and lactacystin (Supplementary Fig. S1[1]); the latter two have previously been shown to inhibit LonA (Bayot *et al.*, 2008; Frase *et al.*, 2006; Frase & Lee, 2007). To obtain structural insights into degradation of the substrate in the Lon chamber, we have determined a set of structures of MtaLonC in a proteolytically active form both alone and in complexes with each of the three covalent inhibitors.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: DZ5280). Services for accessing this material are described at the back of the journal.

2. Materials and methods

2.1. Molecular cloning, protein expression and purification

cDNA encoding full-length MtaLonC (amino acids 1–719) was cloned into the bacterial expression vector pET21a (Novagen) with a C-terminal 6×His tag. The plasmid was transformed into *Escherichia coli* strain BL21 (DE3) cells (Promega). The cells were grown in LB medium and protein expression was induced with 1 m*M* isopropyl β-D-1-thio-galactopyranoside (IPTG) overnight at 293 K. After harvesting by centrifugation, the cells were resuspended in lysis/binding buffer (50 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl) and ruptured using a high-pressure homogenizer (Avestin). After centrifugation, the supernatant was mixed with nickel-chelate affinity resin (Ni–NTA, Qiagen). The resin was washed with binding buffer containing 20 m*M* imidazole and the protein was eluted with binding buffer containing 250 m*M* imidazole. Fractions containing MtaLonC were pooled and the protein was further purified on a Superose 6 column (GE Healthcare) equilibrated in 20 m*M* HEPES pH 7.5 with 1 m*M* DTT. The double mutant construct of MtaLonC was made by PCR-based mutagenesis, sequence verified, and the protein

prepared using the same purification procedure. Wild-type MtaLonC contains only four internal methionine residues; double Leu91Met and Ile359Met mutations were therefore engineered to facilitate *de novo* phasing with selenomethionine-substituted protein. *E. coli* strain B834 (DE3) cells (Novagen) were used and were grown in SelenoMet medium (Molecular Dimensions) following the manufacturer's instructions.

2.2. Protein crystallization

Crystallization of native or SeMet-derivatized MtaLonC was carried out at 295 K by the sitting-drop vapour-diffusion technique. MtaLonC formed hexagonal plate-like crystals using a reservoir consisting of 10% 2-propanol, 100 m*M* dipotassium phosphate, 100 m*M* sodium citrate pH 4.6. Complexes with bortezomib (Millennium Pharmaceuticals), MG262 (EMD Millipore) and *clasto*-lactacystin β-lactone (Enzo Life Sciences) were obtained by soaking apo crystals in solutions of the compounds dissolved in dimethyl sulfoxide (DMSO) at concentrations of 0.5 or 1 m*M* overnight. All crystals were harvested and were briefly immersed into mother liquor containing 30% glycerol prior to data collection.

2.3. Structure determination

X-ray diffraction data were collected from an SeMet-substituted crystal of MtaLonC-L91M/I359M at the peak wavelength on beamline AR-NE3A at the Photon Factory (PF) and were processed using *HKL*-2000 (Otwinowski & Minor, 1997) (Table 1). The structure was determined by single-wavelength anomalous dispersion (SAD) using *SOLVE*/*RESOLVE* (Terwilliger, 2003; Terwilliger & Berendzen, 1999). All seven selenium sites (including the first Met residue) were located in the asymmetric unit, which contained one MtaLonC monomer, in space group *P*6 and an initial structure model was built after density modification. Using reflection data to full resolution, an initial model with ~400 residues was autobuilt using *ARP*/*wARP* (Langer *et al.*, 2008). Iterative rounds of manual model building in *Coot* (Emsley & Cowtan, 2004) and refinement in *REFMAC*5 (Murshudov *et al.*, 2011) resulted in a 2.0 Å resolution model of MtaLonC with reasonable stereochemistry and crystallographic validation parameters (Table 1).

Three complex structures of inhibitors covalently bound to wild-type MtaLonC were determined from data sets collected on beamlines 13B1 (bortezomib) and 13C1 (MG262) at NSRRC and BL44XU at SPring-8 (lactacystin). All data were processed using *HKL*-2000. Using the structure of apo MtaLonC as the search model, clear molecular replacement solutions were found using the program *Phaser* (McCoy *et al.*, 2007). The models of the inhibitor compounds were built using the programs *PRODRG* (Schüttelkopf & van Aalten, 2004) and *Coot* and were fitted to the electron-density map after the protein models had been completed in *Coot*. Crystallographic and refinement statistics are listed in Table 1. All structure figures were prepared with *PyMOL* (v.1.3; Schrödinger).

2.4. Enzymatic assay

The fluorescence peptide assays were performed with the F-β20-Q peptide {*ortho*-aminobenzoic acid (Abz)-QLRSLN-

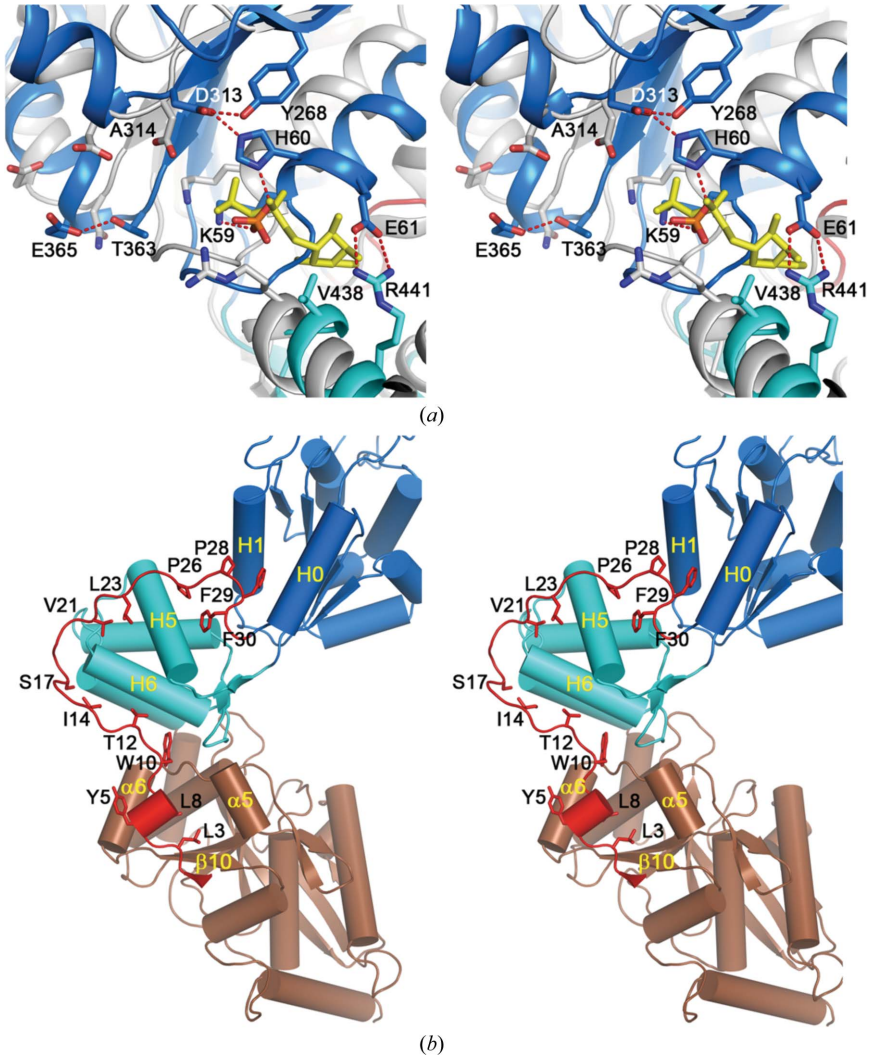

Figure 2
Structure of the catalytically inactive AAA-like module. (*a*) Ribbon diagram of the α/β-subdomain. Residues that are likely to affect the ATPase activity are shown as sticks and are labelled. Hydrogen bonds are indicated by dotted lines. For comparison, TonLonB (PDB entry 3k1j, chain *A*; grey) is superimposed, with bound ADP (yellow) and the corresponding AAA+ motifs shown in stick representation. (*b*) The N-terminal coil (red) tethering the AAA+ subdomains (blue and light blue) to the protease domain (brown), with interacting residues shown in stick representation.

GEWRFAWFPAPEAV[Tyr(3-NO_2)]A} flanked by a fluorophore (Abz) and quencher (nitrotyrosine) pair (Gur & Sauer, 2008; Shanghai Mocell Biotech, People's Republic of China). Reaction mixtures containing the β20 peptide (5 μM) in 20 mM HEPES pH 7.5, 1 mM DTT, 100 mM NaCl with or without MtaLonC (0.4 μM hexamer) were used as positive and negative controls. Samples in the same buffer containing the β20 peptide and MtaLonC pre-incubated with 0.2 mM bortezomib, MG262 or lactacystin for 30 min were used in the experiment. The reactions were carried out at 328 K for the periods indicated. Fluorescence spectra were recorded using a Jobin Yvon Fluorolog-3 fluorescence spectrophotometer. Excitation was at 320 nm and emission was measured at 420 nm. IC_{50} values were determined as described by Frase *et al.* (2006). All experiments were performed in triplicate.

3. Results and discussion

The crystal structure of full-length apo MtaLonC was determined by single-wavelength anomalous dispersion and the model was refined to 2.0 Å resolution (R_{work} and R_{free} of 18.3 and 21.8%, respectively; Table 1). Each monomer has the α/β-subdomain of the AAA-like module at the top, an α-subdomain in the central portion and the Lon protease domain at the bottom; the overall structure is similar to that of TonLonB (Cha *et al.*, 2010; r.m.s.d. of 3.1 Å; PDB entry 3k1j, chain *A*; Fig. 1*b*). MtaLonC forms a symmetrical barrel assembled from six crescent-shaped monomers, but unlike TonLonB its hollow chamber is accessible from two distinct open pores (>13 Å diameter) positioned along the sixfold axis and also from small peripheral openings (Figs. 1*c* and 1*d*). Six proteolytic sites with deep grooves are evenly distributed along a circular gallery in the central section (Figs. 1*d* and 1*e*). The α/β-core is modified by three AAA inserts with chelatase-like topology (Supplementary Figs. S2 and S3; Iyer *et al.*, 2004). The top AAA+ pore is formed by two flexible loops from the PS-I and H2 inserts, both of which lack any large aromatic/hydrophobic residues; these pore-loop residues form a gated portal in TonLonB and are critical for ATP-dependent substrate access to the chamber (Cha *et al.*, 2010). The bottom opening is flanked by six carboxylic ends of the C-terminal helices of the protease domains; it is distinct from that in TonLonB, which is rather restricted (Cha *et al.*, 2010). The open axial pores lacking the critical aromatic residues explain how MtaLonC uses a diffusion-based mechanism to

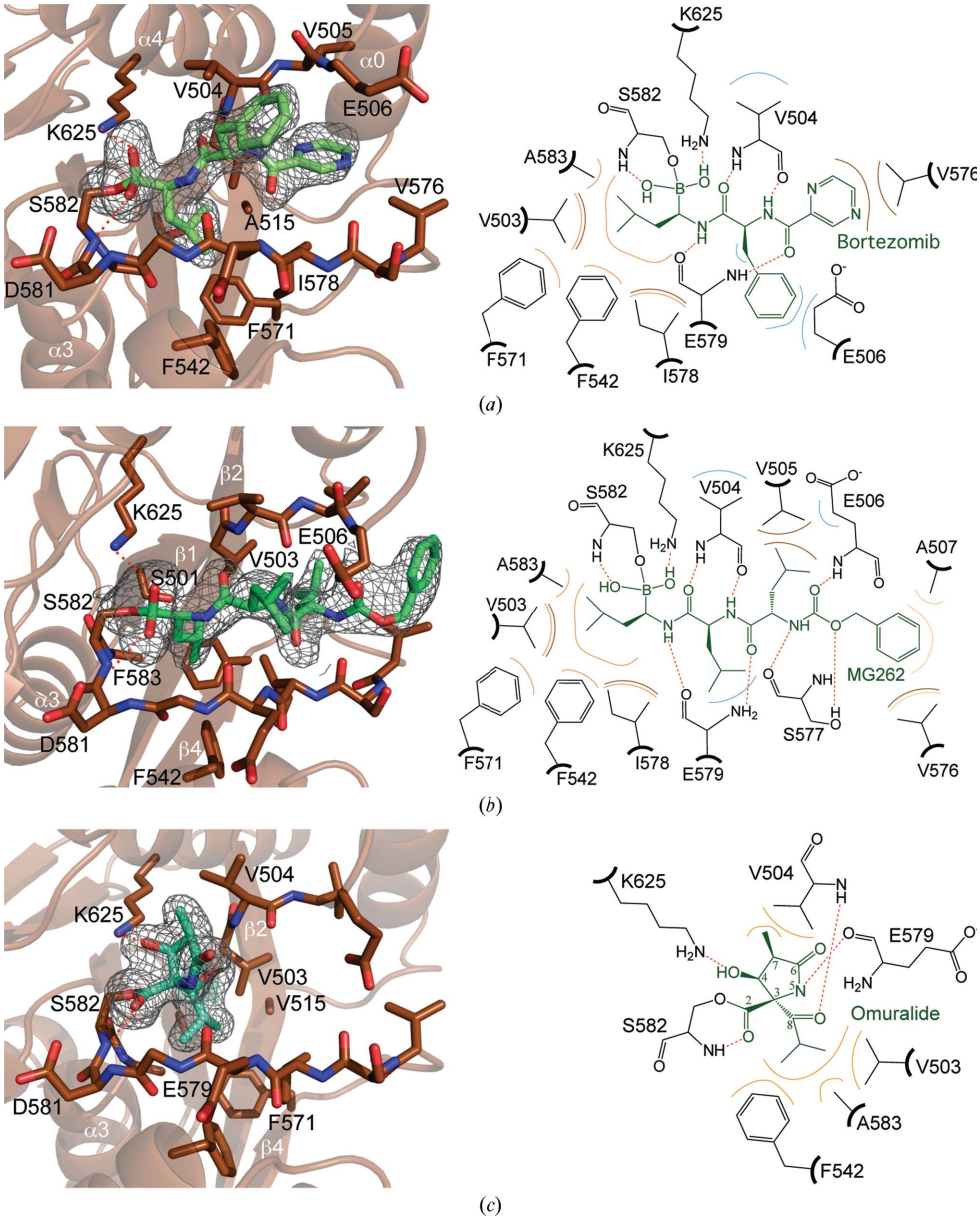

Figure 3
The substrate-binding groove with covalently bound inhibitors. MtaLonC is shown in a semi-transparent ribbon representation; bound bortezomib (*a*), MG262 (*b*) and lactacystin (in hydrolyzed form; omuralide) (*c*) are shown as stick models with a simulated-annealing $F_o - F_c$ OMIT map contoured at the 2.6σ level for the compounds. Groove residues are shown in stick representation. Also shown are schematic views of each of the bound inhibitors highlighting critical interactions (right panels), in which hydrophobic interactions and polar contacts are shown as solid curves and red dotted lines, respectively. Interacting residues are matched by curves with the same colours.

allow only unfolded/unstructured substrate polypeptides to access the degradation chamber in an ATP-independent way, as shown previously (Liao *et al.*, 2012).

The LID emerges from Ins1 of the α/β-core (Cha *et al.*, 2010), locating at the peripheral side of the top axial pore (Fig. 1*c*). Although the integrity of the full-length protein was confirmed by SDS–PAGE of dissolved crystals, a significant portion of the LID (residues 109–217) is disordered. In the crystal, MtaLonC hexamers pack into stable planar layers perpendicular to the *c* axis *via* side-by-side interactions (Supplementary Fig. S4*a*). The disordered LID appears to mediate a labile packing contact with the protease rings of the hexamers from the adjacent layer, since the undefined domains give rise to an ∼20 Å gap between the planar layers (Supplementary Fig. S4*b*). Labile interactions between crystallographic layers involving large disordered domains have been observed previously in an HslU crystal structure (Trame & McKay, 2001).

It has previously been shown that LonC has no ATPase activity (Liao *et al.*, 2012). Several unique structural features of MtaLonC support this lack of ATPase activity. Firstly, Lys59 of the Walker A motif (G*xxx*GKS/T, where *x* is any residue) is followed by His60. This Lys–His pair engages a phosphate ion from the crystallization buffer, with the side chain of His60 occupying the critical magnesium-binding site (Fig. 2*a*). Together with Tyr268, His60 forms hydrogen bonds to the carboxylate group of the conserved Asp313 of the Walker B motif ($\varphi\varphi\varphi\varphi$DE, where φ is a hydrophobic residue). Notably, the catalytic glutamate of this motif that is essential for ATP hydrolysis is missing and is replaced by Ala314 (Fig. 2*a* and Supplementary Fig. S2). The residues that are critical for sensing the hydrolysis of bound nucleotides, sensor II (normally Asn) and the Arg finger, are missing and are replaced by Val438 and Glu378, respectively. Sensor I (Thr363) is tethered by a hydrogen bond to Glu365. Finally, the structure reveals extensive interdomain interactions which may prohibit rotation between the two subdomains (Wang *et al.*, 2001); they include a salt bridge formed between Glu61 from the α/β-subdomain and Arg441 from the α-subdomain (Fig. 2*a*) and a distinct LonC-specific hydrophobic N-terminal coil that tethers the two subdomains together (Fig. 2*b*).

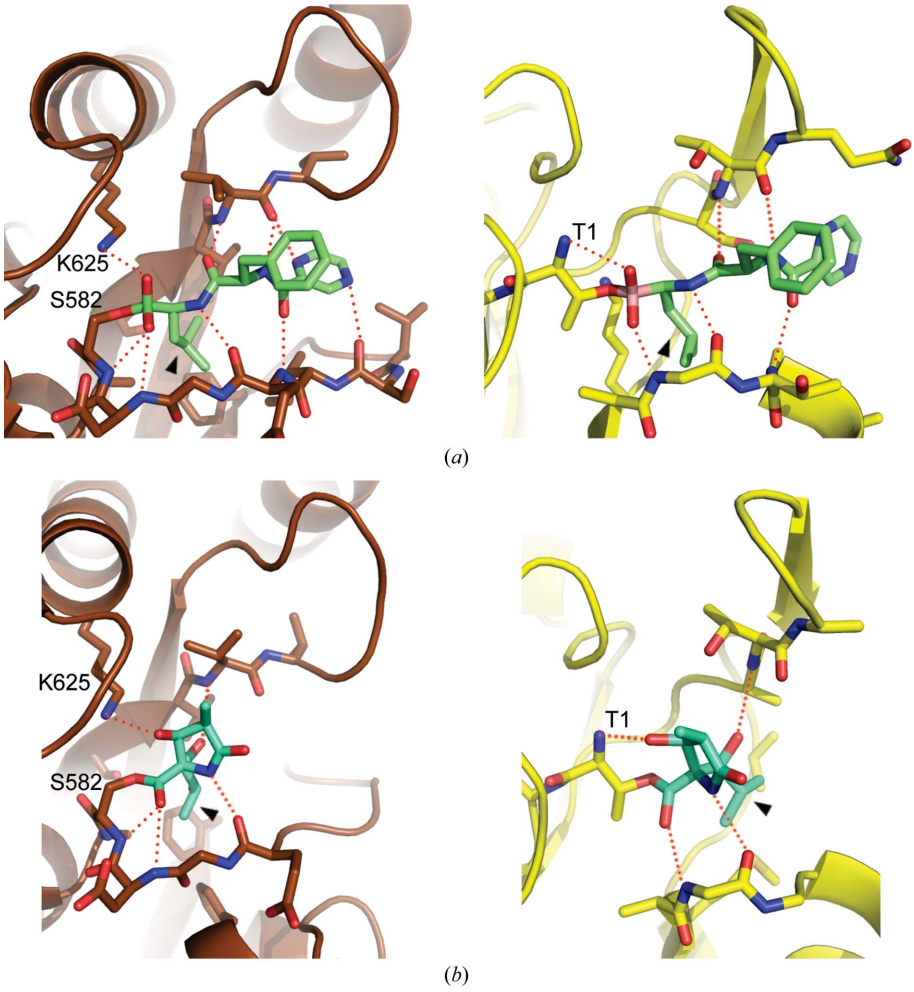

Figure 4
Structural comparison of bortezomib (*a*) and lactacystin (*b*) bound to proteolytic sites in MtaLonC (brown) and in subunits $\beta1$ and $\beta5$, respectively, of yeast 20S proteasome (yellow). Hydrogen bonds are depicted as red dotted lines. The hydrophobic groups of the inhibitors that may be modified to improve the binding specificity are indicated by arrowheads.

To characterize the protease active site in molecular detail, we determined the structures of complexes with the covalently bound peptidomimetics bortezomib and MG262, as well as with the natural product lactacystin, at resolutions of 2.1 Å or higher (Table 1). In each structure, the inhibitors are well defined in the electron-density maps, indicating that all six active sites in the hexameric chamber were fully occupied (Fig. 3). All of the inhibitors bind covalently to Ser582 O^{γ}, confirming the proposed role of this catalytic residue in covalent adduct formation (Frase *et al.*, 2006; Frase & Lee, 2007). In the structures, all three compounds sit in a deep groove demarcated by the extended loop (residues 571–583) connecting strand $\beta4$ and helix $\alpha3$ and by the hairpin loop (residues 503–512) linking strands $\beta1$ and $\beta2$, forming a contiguous hydrophobic S1–S3 pocket which is lined by the residues Ser501, Leu502, Val503, Val504, Val505, Ala515, Phe542, Phe571, Val576, Ile578 and Ala583. This prominent groove was blocked in all of the previously reported structures of isolated protease domains (Supplementary Fig. S5; Botos, Melnikov, Cherry, Tropea *et al.*, 2004; Botos *et al.*, 2005; Duman & Löwe, 2010; García-Nafría *et al.*, 2010; Im *et al.*, 2004; Li *et al.*, 2005). The catalytic Lys625–Ser582 dyad is located at an open end of the

groove, where Lys625 N^{ζ} is likely to be deprotonated owing to the positive helical dipoles of helices α3 and α4. The peptidyl boronates bortezomib and MG262 form a tetrahedral adduct with Ser582 O^{γ}, with the boronate hydroxyl groups hydrogen bonded to Lys625 N^{ζ} and to Ser582 N that stabilizes the oxyanion hole (Figs. 3*a* and 3*b*). Lactacystin binds to Ser582 O^{γ} as an acyl-ester adduct in its water-hydrolyzed form, *clasto*-lactacystin β-lactone (omuralide), as the result of nucleophilic attack on the C2 carbonyl (Fig. 3*c*); the C4 hydroxyl resulting from the opening of the lactone ring forms a hydrogen bond to Lys625 N^{ζ}. Similar to the mechanism proposed for its inhibition of the proteasome, the C4 hydroxy group, which is fixed in position by the γ-lactam ring, may occupy the position adopted by the nucleophilic water molecule, thereby preventing cleavage of the ester bond (Groll, Huber *et al.*, 2006). The polar atoms of omuralide have been demonstrated to engage in hydrogen-bonding interactions with the active sites of the proteasome and chlamydial protease (Groll, Huber *et al.*, 2006; Huang *et al.*, 2008). Here, the C2 carbonyl O atom is in the oxyanion hole, interacting with Ser582 N; the N5 and the C6 carbonyl form hydrogen bonds to the backbone amide groups of Glu579 and Val504, respectively (Fig. 3*c*). Overall, these structures support a mechanism in which nucleophilic attack of Ser582 O^{γ} on the peptide carbonyl of the substrate is catalyzed by the deprotonated Lys625 N^{ζ}, which is positioned to accept the Ser582 O^{γ} proton directly. Peptide-bond cleavage occurs through a tetrahedral intermediate whose oxyanion is stabilized by Ser582 N. Finally, the deprotonated Lys625 N^{ζ} catalyzes the attack of a nucleophilic water molecule on the Ser582 O^{γ}–CO acyl-ester intermediate, thereby cleaving the ester bond.

The bound boronates engage in backbone hydrogen-bonding interactions with the β1–β2 hairpin and adopt a parallel β-sheet conformation with the extended β4–α3 loop (Figs. 3*a* and 3*b*). Notably, both the P1 leucine side chain of the boronates and the isopropyl group of omuralide at C8 protrude into the S1 side of the pocket surrounded by Val503, Ala583, Phe542 and Ile578 and by Ser501 and Phe571 from below. This S1 subpocket is more voluminous than all of the P1 groups studied here; therefore, it is a potential target for inhibitor development. The P3 leucine side chain of MG262 and the pyrazine moiety of bortezomib wedge into the unobstructed S3 side of the pocket. The P2 side chains of MG262 and bortezomib (leucine and phenylalanine, respectively) all point into empty space; nevertheless, they make van der Waals contacts with Val504 and Glu506 of the hairpin loop and modulate their side-chain conformations. The side chains of the P1/P3 residues are completely buried in the hydrophobic groove, a result that is consistent with their important roles in recognition by the protease, which favours nonpolar and disfavours charged or large polar P1/P3 residues (Nishii *et al.*, 2005).

Structural comparison of the bortezomib–MtaLonC and lactacystin–MtaLonC complexes with those of the same inhibitors covalently bound to the active-site threonine (T1) of the β1/β5 subunits of yeast 20S proteasome reveals an interesting overall similarity in the binding modes of both inhibitors. Both bortezomib and lactacystin engage in highly conserved hydrogen-bonding interactions with the backbone atoms of the proteolytic active sites of both proteases, which share remarkable structural similarity (Fig. 4; Groll, Berkers *et al.*, 2006; Groll *et al.*, 1997). The backbone atoms of bortezomib contribute most of the hydrogen bonds, although the pyrimidine ring forms an additional hydrogen bond to the MtaLonC active site (Fig. 4*a*). All polar atoms except for the C6 carbonyl O atom of lactacystin participate in hydrogen bonding to the active-site backbone atoms in MtaLonC and the proteasome (Fig. 4*b*). The most likely moieties of the two inhibitors that may be modified to improve the binding specificity to the two proteases are the P1 residue of bortezomib and the isopropyl group of lactacystin, as these two groups adopt a different rotamer conformation on binding to the distinct S1 pockets in the two enzyme active sites (Fig. 4).

In conclusion, the current work reveals the structural features of a unique member of the Lon proteases with no intrinsic ATPase activity which selectively degrades unfolded protein substrates in an ATP-independent fashion. Importantly, it provides the first snapshots of how the Lon protease active site engages substrate-peptide mimetics and a natural compound, defining a contiguous substrate-binding groove at each proteolytic site and revealing that the prominent S1 subpocket is critical for interaction with all of these inhibitors. Both bortezomib and lactacystin are widely used compounds both as reagents and as therapeutics. As uncovering the mechanism of the inhibition of the catalytic action of chambered proteases remains an important research field, these results will be helpful in developing improved inhibitors with increased selectivity and with a decrease in side effects such as peripheral neuropathy. Previously, obtaining diffracting crystals of LonA/B in a functional hexameric assembly has been difficult (Cha *et al.*, 2010; Duman & Löwe, 2010). High-resolution cocrystallographic analysis of substrate/inhibitor recognition by LonA/B was further hampered by the need to remove the catalytic dyad to avoid self-degradation during purification and the fact that uptake of substrates or inhibitors into the degradation chamber requires ATP, the hydrolysis of which by the hexameric Lon may introduce structural heterogeneity that is incompatible with crystal-soaking experiments. Our results suggest that the LonC proteases may be employed as valuable crystallographic tools for the future structure-based drug development of specific Lon inhibitors with therapeutic potential.

We thank Johann Deisenhofer and Jue Xiang Wang for critical reading of the manuscript. The structure coordinate file for the lactacystin–proteasome complex was kindly provided by Michael Groll. We are grateful to Yi-Hui Chen and Chin-Ling Ko of beamlines 13B1 and 13C1 at the National Synchrotron Radiation Research Center (NSRRC), Taiwan and Masato Yoshimura of BL44XU at SPring-8, Japan for assistance with synchrotron data collection. The laboratories of C-IC and S-HW were supported by Academia Sinica and National Science Council of Taiwan (NSC 101-2311-B-001-

007, NSC 100-2113-M-001-022-MY3 and NSC 101-2811-M-001-164).

References

Bayot, A., Basse, N., Lee, I., Gareil, M., Pirotte, B., Bulteau, A.-L., Friguet, B. & Reboud-Ravaux, M. (2008). *Biochimie*, **90**, 260–269.

Bernstein, S. H., Venkatesh, S., Li, M., Lee, J., Lu, B., Hilchey, S. P., Morse, K. M., Metcalfe, H. M., Skalska, J., Andreeff, M., Brookes, P. S. & Suzuki, C. K. (2012). *Blood*, **119**, 3321–3329.

Bieniossek, C., Schalch, T., Bumann, M., Meister, M., Meier, R. & Baumann, U. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 3066–3071.

Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D. & Huber, R. (2000). *Nature (London)*, **403**, 800–805.

Bota, D. A. & Davies, K. J. (2002). *Nature Cell Biol.* **4**, 674–680.

Botos, I., Melnikov, E. E., Cherry, S., Khalatova, A. G., Rasulova, F. S., Tropea, J. E., Maurizi, M. R., Rotanova, T. V., Gustchina, A. & Wlodawer, A. (2004). *J. Struct. Biol.* **146**, 113–122.

Botos, I., Melnikov, E. E., Cherry, S., Kozlov, S., Makhovskaya, O. V., Tropea, J. E., Gustchina, A., Rotanova, T. V. & Wlodawer, A. (2005). *J. Mol. Biol.* **351**, 144–157.

Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulova, F., Dauter, Z., Maurizi, M. R., Rotanova, T. V., Wlodawer, A. & Gustchina, A. (2004). *J. Biol. Chem.* **279**, 8140–8148.

Cha, S.-S., An, Y. J., Lee, C. R., Lee, H. S., Kim, Y.-G., Kim, S. J., Kwon, K. K., De Donatis, G. M., Lee, J.-H., Maurizi, M. R. & Kang, S. G. (2010). *EMBO J.* **29**, 3520–3530.

Duman, R. E. & Löwe, J. (2010). *J. Mol. Biol.* **401**, 653–670.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Frase, H., Hudak, J. & Lee, I. (2006). *Biochemistry*, **45**, 8264–8274.

Frase, H. & Lee, I. (2007). *Biochemistry*, **46**, 6647–6657.

García-Nafría, J., Ondrovicová, G., Blagova, E., Levdikov, V. M., Bauer, J. A., Suzuki, C. K., Kutejová, E., Wilkinson, A. J. & Wilson, K. S. (2010). *Protein Sci.* **19**, 987–999.

Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Granot, Z., Melamed-Book, N., Bahat, A. & Orly, J. (2007). *Mol. Cell. Endocrinol.* **265–266**, 51–58.

Groll, M., Berkers, C. R., Ploegh, H. L. & Ovaa, H. (2006). *Structure*, **14**, 451–456.

Groll, M., Ditzel, L., Löwe, J., Stock, D., Bochtler, M., Bartunik, H. D. & Huber, R. (1997). *Nature (London)*, **386**, 463–471.

Groll, M., Huber, R. & Potts, B. C. (2006). *J. Am. Chem. Soc.* **128**, 5136–5141.

Gur, E. & Sauer, R. T. (2008). *Genes Dev.* **22**, 2267–2277.

Huang, Z., Feng, Y., Chen, D., Wu, X., Huang, S., Wang, X., Xiao, X., Li, W., Huang, N., Gu, L., Zhong, G. & Chai, J. (2008). *Cell Host Microbe*, **4**, 529–542.

Im, Y. J., Na, Y., Kang, G. B., Rho, S.-H., Kim, M.-K., Lee, J. H., Chung, C. H. & Eom, S. H. (2004). *J. Biol. Chem.* **279**, 53451–53457.

Iyer, L. M., Leipe, D. D., Koonin, E. V. & Aravind, L. (2004). *J. Struct. Biol.* **146**, 11–31.

Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.

Lee, I. & Suzuki, C. K. (2008). *Biochim. Biophys. Acta*, **1784**, 727–735.

Li, M., Gustchina, A., Rasulova, F. S., Melnikov, E. E., Maurizi, M. R., Rotanova, T. V., Dauter, Z. & Wlodawer, A. (2010). *Acta Cryst.* D**66**, 865–873.

Li, M., Rasulova, F., Melnikov, E. E., Rotanova, T. V., Gustchina, A., Maurizi, M. R. & Wlodawer, A. (2005). *Protein Sci.* **14**, 2895–2900.

Liao, J.-H., Kuo, C.-I., Huang, Y-Y., Lin, Y.-C., Lin, Y.-C., Yang, C.-Y., Wu, W.-L., Chang, W.-H., Liaw, Y.-C., Lin, L.-H., Chang, C.-I. & Wu, S.-H. (2012). *PLoS One*, **7**, e40226.

Maehara, T., Hoshino, T. & Nakamura, A. (2007). *Extremophiles*, **12**, 285–296.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Mizusawa, S. & Gottesman, S. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 358–362.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nishii, W., Suzuki, T., Nakada, M., Kim, Y. T., Muramatsu, T. & Takahashi, K. (2005). *FEBS Lett.* **579**, 6846–6850.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rotanova, T. V., Botos, I., Melnikov, E. E., Rasulova, F., Gustchina, A., Maurizi, M. R. & Wlodawer, A. (2006). *Protein Sci.* **15**, 1815–1828.

Rotanova, T. V., Melnikov, E. E., Khalatova, A. G., Makhovskaya, O. V., Botos, I., Wlodawer, A. & Gustchina, A. (2004). *Eur. J. Biochem.* **271**, 4865–4871.

Sauer, R. T. & Baker, T. A. (2011). *Annu. Rev. Biochem.* **80**, 587–612.

Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). *Acta Cryst.* D**60**, 1355–1363.

Suno, R., Niwa, H., Tsuchiya, D., Zhang, X., Yoshida, M. & Morikawa, K. (2006). *Mol. Cells*, **22**, 575–585.

Terwilliger, T. C. (2003). *Acta Cryst.* D**59**, 38–44.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Trame, C. B. & McKay, D. B. (2001). *Acta Cryst.* D**57**, 1079–1090.

Wang, J., Song, J. J., Seong, I. S., Franklin, M. C., Kamtekar, S., Eom, S. H. & Chung, C. H. (2001). *Structure*, **9**, 1107–1116.

Wickner, S., Maurizi, M. R. & Gottesman, S. (1999). *Science*, **286**, 1888–1893.