doi:10.1016/j.jmb.2010.06.030 J. Mol. Biol. (2010) xx, xxx–xxx

JMB

Available online at www.sciencedirect.com
ScienceDirect

ELSEVIER

Crystal Structures of *Bacillus subtilis* Lon Protease

Ramona E. Duman and Jan Löwe*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK

Received 16 March 2010; received in revised form 14 June 2010; accepted 15 June 2010

Lon ATP-dependent proteases are key components of the protein quality control systems of bacterial cells and eukaryotic organelles. Eubacterial Lon proteases contain an N-terminal domain, an ATPase domain, and a protease domain, all in one polypeptide chain. The N-terminal domain is thought to be involved in substrate recognition, the ATPase domain in substrate unfolding and translocation into the protease chamber, and the protease domain in the hydrolysis of polypeptides into small peptide fragments. Like other AAA+ ATPases and self-compartmentalising proteases, Lon functions as an oligomeric complex, although the subunit stoichiometry is currently unclear. Here, we present crystal structures of truncated versions of Lon protease from *Bacillus subtilis* (BsLon), which reveal previously unknown architectural features of Lon complexes. Our analytical ultracentrifugation and electron microscopy show different oligomerisation of Lon proteases from two different bacterial species, *Aquifex aeolicus* and *B. subtilis*. The structure of BsLon-AP shows a hexameric complex consisting of a small part of the N-terminal domain, the ATPase, and protease domains. The structure shows the approximate arrangement of the three functional domains of Lon. It also reveals a resemblance between the architecture of Lon proteases and the bacterial proteasome-like protease HslUV. Our second structure, BsLon-N, represents the first 209 amino acids of the N-terminal domain of BsLon and consists of a globular domain, similar in structure to the *E. coli* Lon N-terminal domain, and an additional four-helix bundle, which is part of a predicted coiled-coil region. An unexpected dimeric interaction between BsLon-N monomers reveals the possibility that Lon complexes may be stabilised by coiled-coil interactions between neighbouring N-terminal domains. Together, BsLon-N and BsLon-AP are 36 amino acids short of offering a complete picture of a full-length Lon protease.

Keywords: self-compartmentalising protease; size-exclusion protease; serine protease; HslUV; proteasome

Edited by R. Huber

Introduction

Lon proteases are members of a network of degradation machines, responsible for maintaining protein quality control inside cells. Whilst the paradigm for cytosolic protein degradation in eukaryotic organisms is embodied by the proteasome/ubiquitin pathway, eubacteria make use of a variety of different proteases, such as Clp, Hsl, and Lon, which share some functional and architectural features with the proteasome. Whilst the complete and functionally relevant structure of Lon complexes is still unknown, Clp and Hsl proteases have a barrel-like architecture,[1] reminiscent of the proteasome, consisting of stacked oligomeric rings. The presence of a proteolytic core of two protease rings, which sequesters the catalytic active sites on the inside of the complex, is the reason that these protease complexes have been labelled 'self-compartmentalising'.[2] The operation of both bacterial and eukaryotic self-compartmentalising protein degradation machines consists of an ATP-driven sequence of events: substrate recognition and engagement are followed by unfolding and threading of the polypeptide chain through the inner channel of the complex, towards the protease active site, where polypeptide products ranging from 10 to

*Corresponding author. E-mail address: jyl@mrc-lmb.cam.ac.uk.

Abbreviations used: PDB, Protein Data Bank; SAD, single-wavelength anomalous diffraction; NCS, non-crystallographic symmetry; EDTA, ethylenediaminetetraacetic acid.

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

20 amino acids are created.[3] These discrete functions are facilitated by the presence of separate modules, which are responsible for substrate recognition, ATP hydrolysis, and proteolysis.

Lon was identified as one of the key components of the ATP-dependent protein degradation machinery in bacterial cell-free extracts nearly 30 years ago[4,5] and it was the first ATP-dependent protease to be biochemically characterised.[6] Lon homologues are present in both eubacteria and archaea, as well as in eukaryotic organelles.[7] In addition to their quality-control roles, which ensure the removal of up to 50% of aberrant proteins from *E. coli* cells,[6] Lon proteases are responsible for regulating physiological events through the targeted degradation of specific protein substrates, such as the cell division inhibitor SulA.[8]

Lon proteases, as well as Clp, Hsl, and proteasomal ATPases, are members of the large AAA+ (*A*TPases *a*ssociated with various cellular *a*ctivities) family of proteins.[9] Despite their diversity of function, AAA+ proteins seem to be involved mostly in generating ATP-driven conformational changes that are used to create mechanical force. They are generally hexameric ring structures and translocate substrates (DNA, RNA, or protein for different family members) through a central channel. Their ATPase modules share significant sequence homology[10] and typically consist of a larger α/β sub-domain and a smaller α sub-domain, also called the 'substrate sensor and discriminatory' domain.[11,12] The nucleotide binding site lies at the interface between the two sub-domains. Whilst Clp and Hsl complexes consist of separate ATPase and protease subunits that assemble into stacked rings,[13,14] Lon proteases accommodate the ATPase and the protease functions on the same polypeptide chain.

Based on sequence homology and domain architecture, Lon proteases are classified as either Lon A, a typical representative of which is *Escherichia coli* Lon protease, or Lon B, mainly found in archaea[15] (Fig. 1a). As previously revealed by sequence alignments (Supplementary Fig. S1) and limited proteolysis experiments, the Lon polypeptide chain consists of several independently folded domains.[16,17] Members of the Lon A family have three functional domains: an N-terminal, an ATPase, and a protease domain. The N-terminal domain covers about a third of the full-length Lon and is thought to be involved in substrate binding.[18] As already mentioned, the ATPase domain consists of two sub-domains, α and α/β, found in all AAA+ proteins.[11] The α/β sub-domain contains the ATP-binding and hydrolysing Walker A and B sequence motifs, as well as the 'sensor 1' motif. The α sub-domain carries another nucleotide binding sequence called 'sensor 2' (Fig. 1a), characterised by the presence of a conserved arginine residue.[10] Sequence alignments and experiments investigating substrate binding by the isolated α ATPase sub-domain of Lon have revealed that it is homologous to the substrate sensor and discriminatory domain of Clp ATPases and is likewise involved in substrate binding.[12] Finally, the protease domain of Lon is a 150- to 200-amino-acid-long, highly conserved region located at the C terminus. The identity of the catalytic serine residue was inferred from early discoveries that typical serine protease inhibitors block the peptidase activity of *E. coli* Lon[6] and then confirmed by sequence comparisons and site-directed mutagenesis studies.[19] Crystal structures of Lon protease domains show that Lon employs a Ser-Lys catalytic dyad,[20] as opposed to the canonical Ser-His-Asp triad used by trypsin and chymotrypsin.[21]

Lon B proteases lack the large N-terminal domain present in Lon A (Fig. 1a); however, they contain a hydrophobic transmembrane region inserted within the ATPase domain, between the Walker A and Walker B sequence motifs. Lon B ATPases thus somewhat resemble the domain architecture of the related bacterial HslU ATPases, which have a domain inserted between their Walker A and Walker B motifs.[14] Archaea lack the membrane-bound ATP-dependent protease FtsH, and thus, Lon proteases of the Lon B type fulfil their roles in these organisms.[22] Some bacteria, such as *Thermotoga maritima*, *Pseudomonas aeruginosa*, and *Bacillus subtilis*, express both Lon A and Lon B.[22]

Like all self-compartmentalising proteases, Lon can only function as an oligomer, although the subunit stoichiometry of Lon proteases is still undetermined or controversial. Based on early gel-filtration and analytical ultracentrifugation work on *E. coli* Lon, it was assumed for a long time that Lon proteases are tetramers.[23] Recent

Fig. 1. Characterisation of domain composition and oligomeric assembly of Lon proteases. (a) Domain composition for Lon A (top), Lon B (middle), and *B. subtilis* Lon (BsLon, the subject of this study, bottom). BsLon is a cytosolic Lon A protease. A COILS prediction for BsLon shows that amino acids ~170 to ~260 are highly likely to form a coiled coil, a feature that is conserved in all Lon A proteases. The two constructs crystallised in this study, BsLon-N and BsLon-AP, are indicated with the start and end amino acids. (b) Sedimentation velocity profiles for *A. aeolicus* Lon, showing the predominance of a 9.1-S species, which most likely corresponds to a hexamer. (c) Sedimentation equilibrium data: apparent molecular weight plotted against the AaLon concentration. Confirming the sedimentation velocity data, the calculated molecular weight, at all three concentrations tested, is close to a hexamer. (d) Sedimentation velocity profiles for *B. subtilis* Lon, revealing a mixture of monomers and small oligomeric species. (e) Sedimentation equilibrium data: apparent molecular weight plotted against the BsLon concentration. Average molecular mass for the mixture of different oligomeric species ranges between 140 and 170 kDa. (f) Left: negatively stained electron micrograph of AaLon. Right: Coomassie-stained SDS-PAGE gel: lanes from left to right: molecular weight marker, BsLon-N, BsLon-AP, BsLon, and AaLon.

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

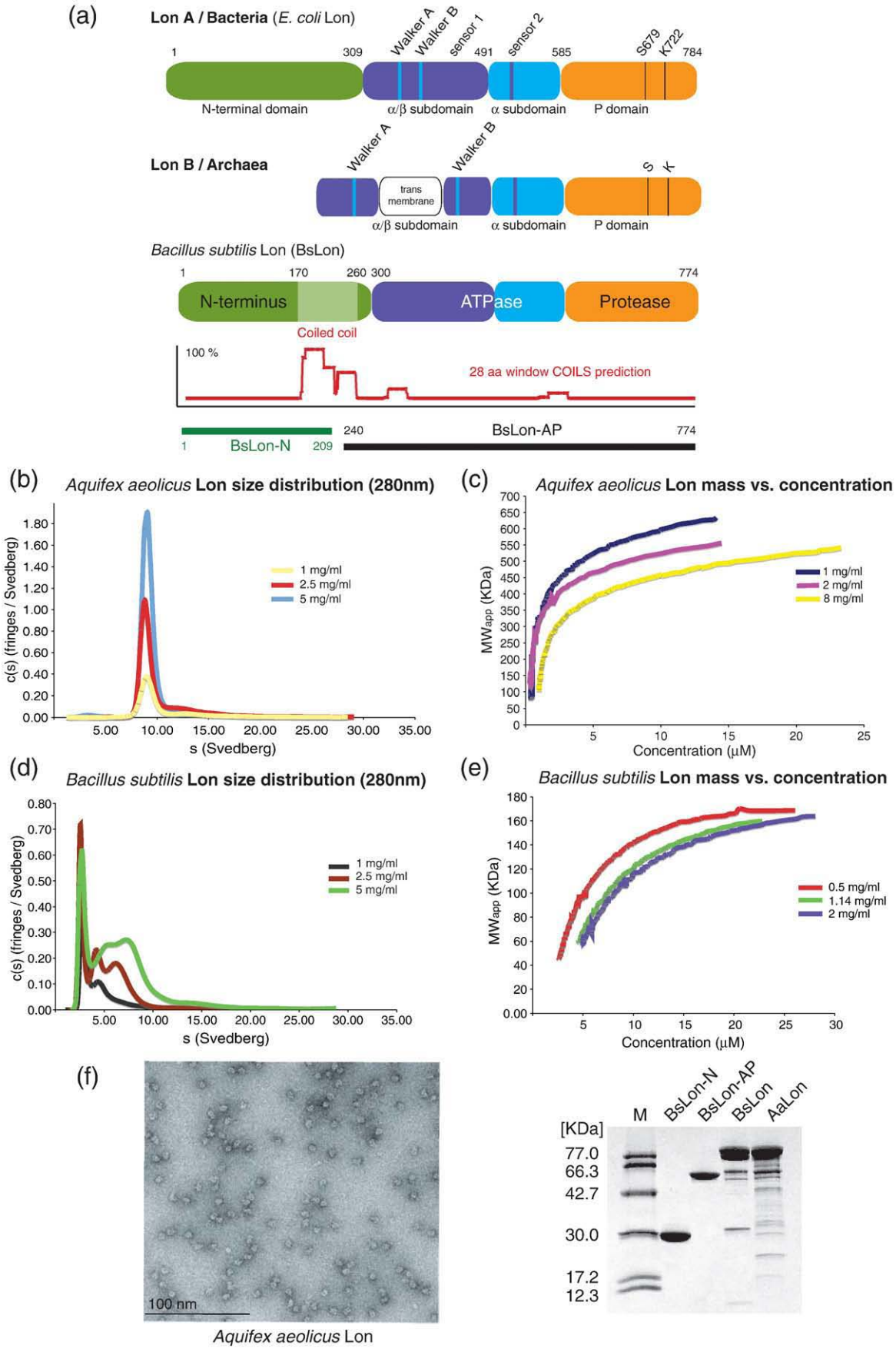

Fig. 1 (*legend on previous page*)

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

electron microscopy data show that negatively stained images of *E. coli* Lon particles appear hexameric,[24] although surprisingly, electron cryo-microscopy images of *S. cerevisiae* Lon particles suggest that mitochondrial Lon has 7-fold symmetry.[25] These studies show that Lon particles from *E. coli* and *S. cerevisiae* are indeed ring-like, with diameters between 12 and 15 nm, and assemble only in the presence of magnesium ions. Analytical ultracentrifugation data collected on Lon from *Mycobacterium smegmatis* indicate that Lon exists as a mixture of oligomers of different subunit stoichiometry and confirm the dependence of complex formation on magnesium.[26] Isolated Lon protease domains from two organisms crystallise as 6-fold assemblies.[20,27]

Complete structures of self-compartmentalising proteases are important because they reveal the mechanism by which the protease sites are shielded such that unwanted proteolysis of any cellular targets is prevented. So far, attempts to solve the oligomeric structure of a full-length Lon complex have only yielded structures of individual domains. A fragment consisting of amino acids 8–117 of the N-terminal domain of *E. coli* Lon,[28] the α subdomain of the *E. coli* AAA+ module,[29] and the protease domain from four different organisms have been successfully crystallised and their structures have been solved.[20,27,30,31]

In this study, we present two crystal structures of truncated forms of Lon from *B. subtilis*. BsLon-AP (*A*TPase and *p*rotease domains) is the first structure of a Lon fragment consisting of both the AAA+ and the protease domains, revealing their arrangement within a Lon monomer. BsLon-AP crystallises as a hexameric complex in which subunits follow a helical arrangement, previously seen with other related AAA+ ATPases.[32,33] BsLon-N (N-terminal domain) contains two-thirds of the N-terminal domain from the same protein and shows a domain-swapped dimer in the asymmetric unit. In addition to the globular domain already known from the *E. coli* structure,[28] it reveals part of the predicted coiled-coil region (Fig. 1a), represented by amino acids 120 to 209. In order to address the uncertainty regarding subunit stoichiometry, we investigate the oligomeric assembly of Lon complexes using analytical ultracentrifugation and electron microscopy. These results show a distinct behaviour of Lon proteases from *Aquifex aeolicus*, which forms mainly hexamers, and *B. subtilis*, which exists as a mixture of small oligomeric species in solution.

Results

Analytical ultracentrifugation and electron microscopy of *B. subtilis* and *A. aeolicus* Lon

Sedimentation velocity and equilibrium experiments were conducted on Lon from *B. subtilis* (BsLon) and the hyperthermophilic eubacterium *A. aeolicus* (AaLon) (Fig. 1b to e). Sedimentation velocity data for AaLon indicate that it has a sedimentation coefficient of 9.1 S (Fig. 1b). The size distribution profiles derived from the AaLon velocity data show that more than 90% of AaLon in solution belongs to this 9.1-S species, although there is a small tail corresponding to larger oligomeric species or aggregates. The three different concentrations tested gave identical results, suggesting that there is no concentration-dependent association in the range tested. The sedimentation equilibrium data for AaLon were best fitted with a mixture of two independent species of molecular masses of 545.7 kDa and 1.14 MDa, with residuals below 0.02. The molecular mass of an AaLon monomer is 90.8 kDa; hence, the 545.7 kDa presumably corresponds to a hexameric species. The calculated average molecular mass for the three different concentrations ranges between 500 and 600 kDa (Fig. 1c), which confirms the velocity results, showing the prevalence of the oligomeric species of 9.1 S, corresponding to the hexameric species.

BsLon sediments as a mixture of oligomers (Fig. 1d). At all three concentrations tested, most of BsLon sediments as a small species of 2.7 S; however, the amount of larger material increases with concentration, indicating concentration-dependent association. Sedimentation equilibrium data for BsLon were fitted with a model containing a mixture of three independent species (71.8 , 210.9, and 304.9 kDa, with residuals below 0.01). The theoretical molecular mass for a BsLon monomer is 87.4 kDa. Due to the abundance of larger oligomeric species, the calculated average molecular mass at concentrations 0.5, 1.14, and 2 mg/ml ranges between 140 and 170 kDa (Fig. 1e).

The analytical ultracentrifugation results show that, whilst AaLon exists mainly as hexameric species in solution, BsLon appears to be a mixture of monomeric and larger oligomeric species, with increasing amounts of larger oligomers present at larger concentrations. These results are in agreement with negatively stained electron microscopy images of AaLon and BsLon, showing that AaLon forms round particles of homogenous size (Fig. 1f), whilst BsLon appears not to form any large oligomers (data not shown).

Structure of the N-terminal domain of BsLon

Structure determination and structure of the monomer

The 2.6-Å structure of the N-terminal domain of *B. subtilis* Lon (BsLon-N) was solved by molecular replacement, using the *E. coli* Lon N-terminal domain structure [Protein Data Bank (PDB) ID: 2ANE[28]] as a search model. The BsLon-N structure comprises the first 209 of the 300 amino acids of the N-terminal domain of BsLon. Residues 1–3, 118, and 119 have been omitted from the final model, as there was no electron density for them. Crystallographic data and refinement statistics are presented in Tables 1 and 2.

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

Table 1. Crystallographic data of *B. subtilis* Lon

Crystal	λ (Å)	Resolution (Å)	$I/\sigma I$[a]	R_m[b] (%)	Multiplicity[c]	Completeness (%)[d]
Construct: protease La/Lon (*B. subtilis*) N-terminal domain (BsLon-N) GI:496557, residues 1–209 (α-chymotrypsin proteolysis product)						
$P2_12_12_1$ (a=43.9 Å, b=81.8 Å, c=115.1 Å)						
NATI	0.9795	2.6	11.0 (3.6)	0.073 (0.307)	3.5 (3.6)	98.6 (99.4)
Construct: protease La/Lon (*B. subtilis*) ATPase/protease domains (BsLon-AP) GI:496557, residues 240–774—LEHHHHHH						
$P6_5$ (a=b=91.1 Å, c=143.5 Å)						
	0.9764	3.2	9.5 (3.4)	0.105 (0.452)	5.4 (5.5)	99.2 (99.2)
$P2_12_12_1$ (a=149.3 Å, b=160.4 Å, c=199.0 Å)						
TABR	1.2546	5.5	15.2 (3.8)	0.088 (0.631)	12.4 (11.4)	99.1 (98.0)
NATI	0.9537	3.7	6.1 (1.8)	0.107 (0.619)	3.4 (3.3)	98.6 (97.3)
$I2$ (a=110.5 Å, b=142.0 Å, c=190.0 Å, β=90.0°)						
	0.9762	4.0	5.1 (2.5)	0.160 (0.491)	3.6 (3.7)	99.6 (99.6)
$P2_1$ (a=103.4 Å, b=127.4 Å, c=149.0 Å, β=100.5°)						
	0.8856	3.4	8.0 (2.3)	0.089 (0.578)	3.5 (3.6)	99.7 (100.0)
$P22_12_1$(a=104.9 Å, b=175.9 Å, c=237.9 Å)						
	0.9764	4.2	9.7 (2.8)	0.114 (0.641)	6.2 (6.4)	99.8 (100.0)

[a] Signal-to-noise ratio for merged intensities.
[b] R_m: $\sum_h\sum_i | I(h,i)-I(h) | / \sum_h\sum_i I(h,i)$ where $I(h,i)$ are symmetry-related intensities and $I(h)$ is the mean intensity of the reflection with unique index h.
[c] Multiplicity for unique reflections (anomalous multiplicity in parentheses).
[d] Completeness for unique reflections. Highest-resolution bins are in parentheses. The BsLon-AP structure was solved using the $P2_12_12_1$ TABR SAD data set (CC-anom: 0.94 to 5.5 Å) for phase determination and phase extension using the NATI data set with sevenfold averaging.

The structure of BsLon-N consists of two distinct regions, connected by an extended loop of 10 amino acids: a compact β-sheet rich, globular domain (amino acids 4–114), which is closely related to the structure of the *E. coli* N-terminal domain (PDB ID: 2ANE; amino acids 8–117), and an additional α-helical domain (amino acids 124–209), consisting of a four-helix bundle (Fig. 2a). The globular domain is anchored via an ~ 27-Å-long, straight linker to a 33-amino-acid-long helix at nearly 90° from the linker.

Table 2. Refinement statistics

	$P2_12_12_1$ crystal form, N-terminal domain	$P2_1$ crystal form, C-terminal domains
Model	2 chains/ASU Residues 4–118, 121–209 36 water molecules	6 chains/ASU Residues 246–382, 403–427, 437–770 6 ADP molecules 0 waters
Diffraction data	NATI, 100.0–2.6 Å, all data	NATI, 30.0–3.4 Å, all data
R-factor, R_{free}[a]	0.191 (0.254), 0.269 (0.358)	0.265 (0.380), 0.313 (0.45)
B-factors[b] (Å^2)	49.8, 12.1	194.6, 26.2
Geometry[c]	0.008 Å, 1.081°	0.016 Å, 1.672°
Ramachandran[d] (%)	99.6/0.0	97.5/0.0
PDB ID	3M65	3M6A

[a] Five percent of reflections were randomly selected for determination of R_{free}, prior to any refinement. R-factors for the highest-resolution bins are given in parentheses.
[b] Temperature factors averaged for all atoms and rmsd of temperature factors between bonded atoms.
[c] rmsds from ideal geometry for bond lengths and restraint angles.
[d] Percentage of residues in the "most favoured and additionally allowed region" of the Ramachandran plot (PROCHECK) and percentage of outliers. Data were collected on beamlines ID23eh1 and ID29 (European Synchrotron Radiation Facility, Grenoble, France) and on beamlines I03 and I02 (Diamond Light Source, Harwell, UK).

This long helix forms a four-helix bundle with three smaller helices at the C terminus of BsLon-N.

The chain break in the BsLon-N structure, after amino acid Glu117, coincides with the C terminus of the *E. coli* structure (PDB ID: 2ANE). Like the BsLon-N fragment, the *E. coli* Lon N-terminal fragment was generated by limited proteolysis, suggesting that this linker region must be flexible and exposed in both organisms. The first 117 residues in the N termini from EcLon and BsLon, which share 49.14% sequence identity, align very well structurally with an rmsd of 0.98 Å (Fig. 2c, left).

The BsLon-N dimer

The asymmetric unit of the BsLon-N crystals contains a tightly packed, domain-swapped dimer (Fig. 2b): the N-terminal globular domain of one monomer is next to the three C-terminal (C-terminal of BsLon-N) helices of the other monomer. The monomers cross each other, such that the C-terminal three helices of monomer A are packed between the globular domain and the long helix of monomer B. The long helix from each monomer lies parallel with the equivalent helix of the other monomer and to the non-crystallographic 2-fold axis of the dimer. Out of the 206 residues, 77 are involved in the dimer interface, which consists of both hydrophilic and hydrophobic contacts, established between two of the three C-terminal helices of one monomer and the long helix and globular domain of the other monomer.

An electrostatic surface potential map of a BsLon-N monomer reveals three potential hydrophobic patches: on the underside of the globular domain (labelled A in Fig. 2a), on the inside of the long helix, at its N-terminal end (B), and on the inside of two of the three C-terminal helices (C). All the amino acids

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

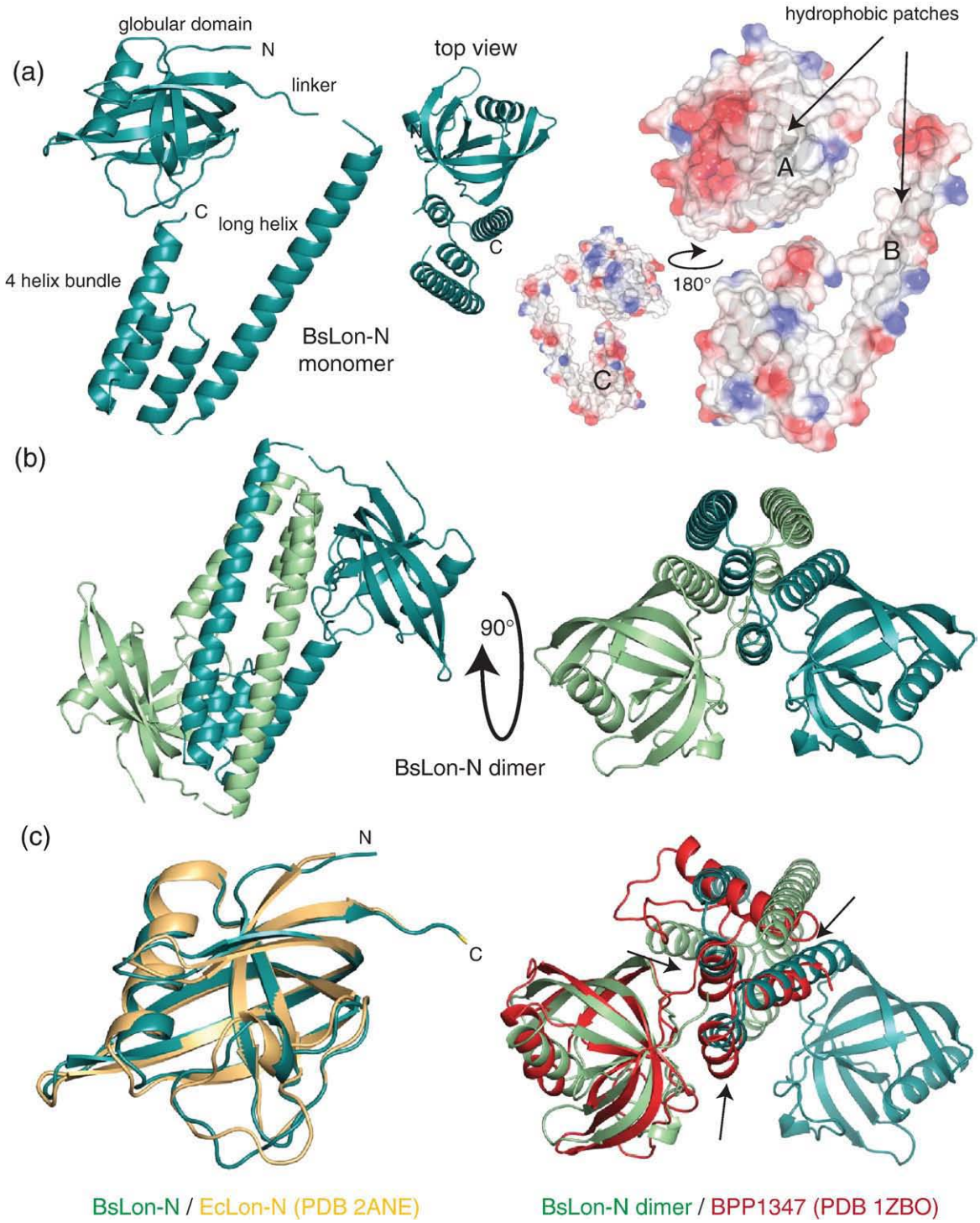

Fig. 2. Crystal structure of BsLon-N. (a) Left: side and top views of a BsLon-N monomer. The globular domain is connected to the four-helix bundle by an extended loop, 'the linker', which forms a 90° angle with the long helix. Right: electrostatic surface potential maps of a BsLon monomer, showing hydrophobic patches A (on the globular domain), B, and C (on the four-helix bundle). (b) Side (left) and top (right) views of the BsLon-N domain-swapped dimer. The dimeric interface consists mainly of contacts between the helical bundles of the two monomers. (c) Left: structural alignment of a BsLon-N globular domain (green) and *E. coli* N-terminal domain (yellow, PDB ID: 2ANE). Right: structural alignment of the BsLon-N dimer (green) and BPP1347 (red, PDB ID: 1ZBO). BPP1347 has a globular domain with a fold similar to that of the BsLon-N globular domain and a four-helix bundle. The BsLon-N dimer is shown to highlight the surprisingly similar topology of three helices in BPP1347 to the last three helices of the other BsLon-N monomer, marked by black arrows.

that form these patches are included in the dimer interface between the two domain-swapped monomers. The hydrophobic character of patch A is given mainly by residues forming two loops (Gly16, Leu17, Leu18, and Val19 on one hand and Ile66, Phe67, and Val69 on the other) and a β-strand

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

including residues Met23, Val24, Leu25, and Leu27. Patch B includes residues Leu128, Met129, Leu132, Leu133, and Phe136, whilst patch C consists of amino acids Met163, Ile166, and Val167 on one helix and Leu191, Val194, Ile195, and Ile198 on the other helix. In the context of the dimer, patches B and C from different monomers shield each other from the solvent.

Interestingly, the last helix of the monomer structure is part of a region of BsLon (residues ~170 to ~260) with a high likelihood of forming a coiled coil, according to the program COILS[34] (Fig. 1a), and hence gives an indication of the direction of this normally much longer helical segment.

A DALI[35] search against the PDB for proteins with structural similarities to BsLon-N reveals only one significant hit, as was found before[28] using the *E. coli* N-terminal domain structure. The structurally related protein BPP1347 (PDB ID: 1ZBO) from *Bordetella parapertussis* unfortunately has no known function. BPP1347 has a globular domain with an identical fold to BsLon-N and a structural alignment between BsLon-N and BPP1347 gives an rmsd of 2.61 Å. As for BsLon-N, the globular domain of BPP1347 is connected via an extended loop to a four-helix bundle. The extended loop is of similar length and in the same position as that of BsLon-N, whilst the four-helix bundle is shorter. The most striking feature of BPP1347 is that the topology of its three C-terminal helices, marked with arrows in Fig. 2c, right, coincides with that of the C-terminal helices of the other monomer in the BsLon-N asymmetric unit.

Structure of BsLon-AP

Structure determination

Full-length BsLon was initially crystallised in several different conditions. Crystals appeared within 2 to 3 weeks and diffracted only to ~9 Å. The protein composition of these initial crystals was analysed by SDS-PAGE and revealed that the crystals contained a proteolysis fragment of BsLon with an apparent molecular mass of ~55 kDa. N-terminal sequencing of this fragment showed that it started with amino acid Gly240. Western blot analysis using antibodies against the histidine tag confirmed that this fragment, labelled BsLon-AP, contained the C-terminal hexa-histidine tag of BsLon. BsLon-AP, consisting of amino acids 240–774 of BsLon, was subsequently cloned and expressed.

The structure of BsLon-AP was solved by tantalum single-wavelength anomalous diffraction (SAD) at 7 Å, using a data set from a tantalum bromide cluster-soaked crystal of the $P2_12_12_1$ form (Table 1). Solvent flattening led to the identification of seven monomers in the asymmetric unit. Poly-alanine models were manually fit into the density of each of the seven monomers. This enabled the use of 7-fold non-crystallographic symmetry (NCS) averaging to gradually extend the phase information to the 3.4 Å resolution of the native data set. The structure of the $P2_1$ crystal form, detailed in this article, was solved by molecular replacement, using the density of a monomer from the phase-extended map of the $P2_12_12_1$ form as a search model. The solution revealed six monomers in the $P2_1$ asymmetric unit, arranged in an open, helical configuration. Multi-crystal averaging, using three separate domains and two crystal forms, and NCS averaging were employed to further extend phases to 3.4 Å. The existing structures of the *E. coli* protease domain (PDB ID: 1RR9[20]) and the ATPase α subdomain (PDB ID: 1QZM[29]) were used as templates in building the model for this low-resolution structure.

The BsLon-AP monomer

The BsLon-AP monomer (Thr246 to Val770) consists of four distinct domains. Three of these, the C-terminal protease domain and the two ATPase sub-domains, form a three-lobed structure, whilst the fourth, the N-terminal domain, protrudes from the side of the ATPase domain (Fig. 3a, left).

The first three helices at the N terminus of BsLon-AP (Thr246 to Leu299) form a helical bundle and represent the carboxy-terminal end of the 300-amino-acid N-terminal domain of full-length BsLon (Fig. 1a). An extended loop of 10 amino acids separates these three helices from the ATPase domain (Asp310 to Lys579). The C-terminal protease domain (Gln592 to Val770) is linked to the ATPase domain via a flexible loop (Arg580 to Asp591). Two internal flexible loops, corresponding to amino acids Val383 to Gly402 and Ser428 to Pro434, had very poorly defined electron density and are therefore missing from the final model.

The ATPase domain of BsLon-AP closely resembles AAA+ modules of other related ATP-dependent proteases and consists of two distinct sub-domains: a smaller, mainly α-helical domain (Thr493 to Lys579) and a larger domain (Asp310 to Tyr492), which resembles a Rossmann fold. Typically, the two sub-domains of an ATPase module are referred to as the small 'α domain' and the larger 'α/β domain'. The two sub-domains are positioned relative to each other such that a deep groove is formed at their interface. The groove between the two ATPase sub-domains harbours the nucleotide binding site. In our BsLon-AP structure, the nucleotide binding site of each subunit contains an ADP molecule (Fig. 3c).

The protease domain is compact and rich in β-strands and superimposes well onto the existing structure of the *E. coli* protease domain (PDB ID: 1RR9[20]). Amino acids Val593 to Glu772 of BsLon-AP align to amino acids Val595 to Glu774 of the EcLon protease domain with an rmsd of 1.42 Å (Supplementary Fig. S3). Both structures represent inactive mutants since the catalytic serine was replaced by alanine (677 in BsLon and 679 in EcLon).

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

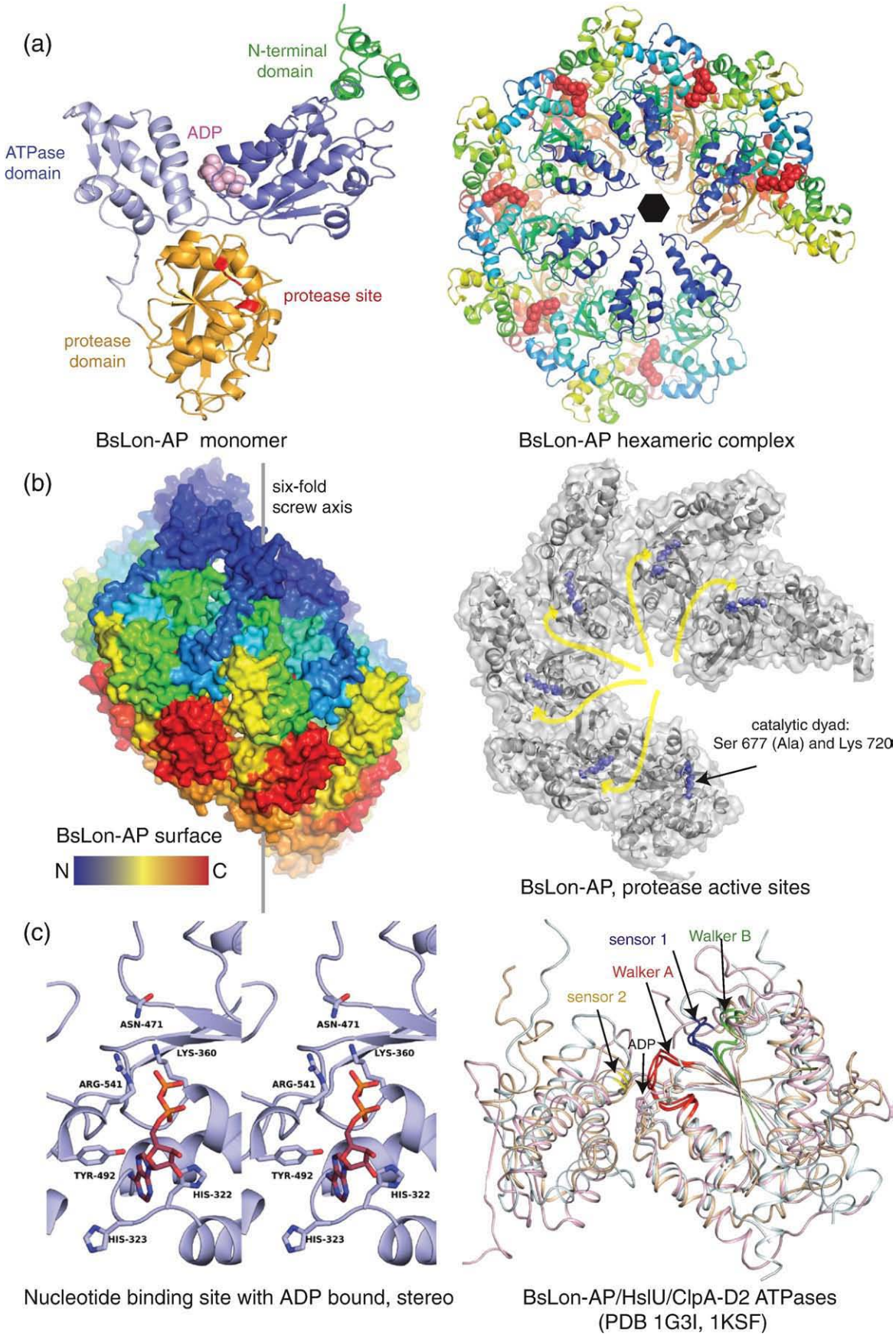

Fig. 3 (*legend on next page*)

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

The nucleotide binding site of BsLon: Identification of conserved nucleotide binding motifs from structural alignment with other ATPases

In AAA+ proteins, the nucleotide binding site is always found at the interface between the two sub-domains of the AAA+ module. The regions of the ATPase domain involved in nucleotide binding and hydrolysis, known as the 'Walker A and B' and 'sensor 1 and 2', are well conserved among Lon ATP-dependent proteases and share significant sequence similarity with equivalent regions in other AAA+ proteins.[10] Whilst the 'Walker A and B' sequences are universally conserved and can easily be identified through sequence alignments, the other nucleotide binding motifs, 'sensors 1 and 2' and the additional 'arginine finger', present considerably more sequence variability and can be more reliably identified through structural alignments with other ATPases. The role of these conserved motifs and their contribution to the ATPase and the protease activities were investigated in a site-directed mutagenesis study on Lon from *Thermoplasma acidophilum*.[36]

HslU and ClpA are bacterial ATPases closely related to Lon. Whilst HslU has only one nucleotide binding domain, ClpA has two (D1 and D2), of which D2 resembles the BsLon ATPase domain the closest. A structural alignment of the ATPase domain of BsLon-AP, HslU (PDB ID: 1G3I[37]), and ClpA-D2 (PDB ID: 1KSF[32]) is used to establish the location of some of the nucleotide binding motifs in BsLon and assess the degree of structural similarity of these regions between related bacterial ATP-dependent proteases (Fig. 3c). In BsLon-AP, amino acids Gly354 to Lys360 make up the Walker A 'P loop' (sequence GPPGVGK in BsLon, Supplementary Fig. S2). The Walker B sequence, responsible for magnesium coordination and nucleotide hydrolysis, corresponds to amino acids Leu419 to Lys425 of BsLon (sequence LLDEIDK).

In addition to the Walker sequences, three other motifs, known as 'sensor 1', 'sensor 2', and the 'arginine finger', are also well conserved in AAA+ proteins. The 'sensor 1' motif is characterised by the presence of a well-conserved threonine or asparagine, generally oriented such that its side chain is part of a hydrogen-bonding network that positions a nucleophile water close to the γ phosphate of ATP.[38] In the D2 ATPase domain of *E. coli* ClpA, this residue is Asn606[39] and its position coincides with that of Asn471 of BsLon. In most AAA+ proteins, 'sensor 2' is an arginine that can hydrogen bond to the terminal phosphate of ATP and is involved in conformational changes between the two ATPase sub-domains, as a result of ATP hydrolysis.[10] Arg702 and Arg326 fulfil this role in ClpA-D2[32] and HslU, respectively.[40] Their positions superimpose well on that of Arg541 of BsLon-AP, confirming the role inferred for this amino acid from sequence alignments.[34]

The location of the nucleotide binding site, in the groove between the two AAA+ sub-domains, makes it accessible to residues from the neighbouring ATPase domain, in the context of the assembled Lon complex. Inter-subunit communication, upon nucleotide binding and hydrolysis, relies on the presence of a very well conserved 'arginine finger', located on the surface of the α/β sub-domain, which reaches inside the nucleotide binding site of the neighbouring subunit to detect the presence and the status of the bound nucleotide. The Arg fingers of ClpA-D2 (Arg643) and HslU (Arg326) align well with Arg483 of BsLon, confirming its role as the 'arginine finger' of BsLon.

Quaternary structure of BsLon-AP

The six BsLon-AP monomers in the asymmetric unit form an open ring, being arranged in a helical pattern around an imperfect, unclosed 6-fold NCS axis. A top view of the hexameric complex, with the N termini of the monomers facing up, shows a compact arrangement of the monomers, facilitated by the complementary curvatures of the ATPase domain: the convex outer surface of an ATPase subunit packs against the concave inner surface of the neighbouring ATPase subunit (Fig. 3a, right, and b, left). This packing positions the ATPase α/β sub-domains towards the core of the complex, with the α sub-domains facing the outside of the ring.

The three N-terminal helices of all BsLon-AP monomers are orientated radially towards the core of the complex and partially occlude the central channel. This is due to the orientation of the three N-terminal helices, at roughly right angles from the rest of the molecule (Fig. 3a, left). Viewed from the side, the BsLon-AP complex appears to have a conical, washer-like shape, due to the reduced size of the protease domains, situated at the C-terminal end of

Fig. 3. Crystal structure of BsLon-AP. (a) Left: BsLon-AP monomer with bound ADP shown as spheres (pink). Each domain is shown in a different colour. The ATPase and the protease domains form a three-lobed structure with the nucleotide binding site in the centre. Ala677 and Lys720 of the protease domain catalytic dyad are shown in red, as sticks. Right: top view of the BsLon-AP hexamer in the asymmetric unit, in rainbow representation (blue to red: N to C). Six ADP molecules are shown as red spheres. (b) Left: side view, surface representation of the BsLon-AP hexamer, showing the helical arrangement of monomers. Right: top view of BsLon-AP protease active sites; shown are the protease and α ATPase sub-domains in grey; the α/β ATPase sub-domains were removed for clarity. The catalytic dyad residues are shown in violet, as spheres. The channels that allow access to the catalytic residues are marked with yellow arrows. (c) Left, stereo image of the nucleotide binding site with ADP shown as sticks. Amino acids that interact with the bound nucleotide are shown as sticks. These include His322, His323, Lys360, Tyr492, Arg541 (sensor 2), and Asn471 (sensor 1). Right: structural alignment of ATPase domains from BsLon-AP (light blue), HslU (dark yellow, PDB ID: 1G3I), and ClpA-D2 (pink, PDB ID: 1KSF); associated nucleotides are shown in identical colours. Conserved nucleotide binding motifs are highlighted: red, Walker A; green, Walker B; blue, sensor 1; yellow, sensor 2.

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

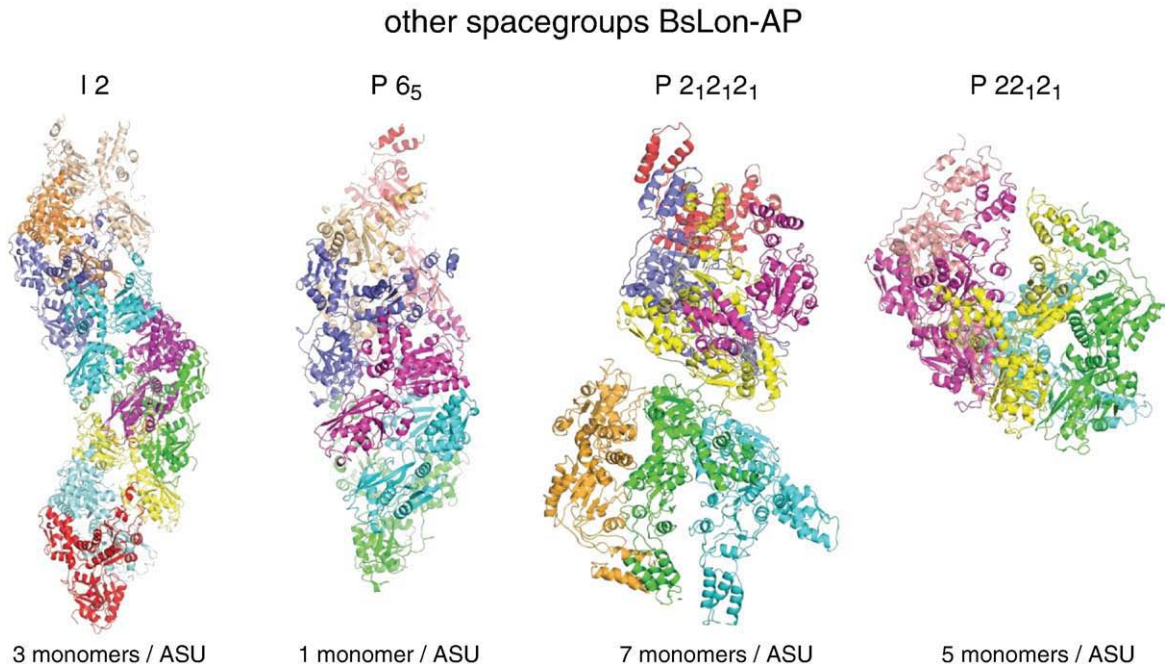

Fig. 4. Alternative space groups for BsLon-AP. Each BsLon-AP monomer is coloured differently. Despite different compositions of the asymmetric unit, all observed space groups of BsLon-AP show a unique but helical arrangement of monomers. These additional crystal forms were not refined and are only shown as solutions coming from molecular replacement calculations using the BsLon-AP monomer structure.

the complex, relative to the size of the ATPase domains, which occupy the middle of the complex (Fig. 3b, left).

The protease active sites of the six BsLon-AP monomers are not accessible from the outside of the complex. A top view of the complex, in which part of the α/β sub-domain of each monomer was removed for clarity, reveals that the catalytic dyad (Ser677, replaced here by Ala, and Lys720) faces the inside of the complex (Fig. 3b, right). In the BsLon-AP structure, the protease active sites are not directly open to the inside of the complex either and seem to be accessible via channels between the subunits, marked in yellow in Fig. 3b, right. The inward-facing protease active sites represent a hallmark of self-compartmentalising proteases, leading to the sequestration of the active sites and prevention of unwanted and unregulated proteolysis.

BsLon-AP has a tendency to crystallise as helix

BsLon-AP crystallised easily in a variety of different conditions, producing several crystal forms. The crystals diffracted to around 3.5 to 5 Å, and the corresponding structures, belonging to four different space groups, were solved by molecular replacement, using the monomer of the initial $P2_1$ structure as a search model. These new crystal forms have different numbers of molecules per asymmetric unit, ranging from three to seven; however, they are all characterised by a helical arrangement of the monomers (Fig. 4), with the helical rise and angles of rotation around the screw axis varying between the five crystal forms. Two of these, $P6_5$ and $I2$, present a considerably steeper helical rise, compared to that seen on the $P2_1$ structure. These results indicate that BsLon-AP has a natural tendency to crystallise in a helical pattern. A superposition of monomers from each of the five crystal forms presented here, with respect to the protease domains, shows that they overlay well and that the ATPase domains are in identical orientations (Supplementary Fig. S5). Thus, the three-lobed architecture of the ATPase and protease domains of BsLon-AP appears to be rigid.

Fig. 5. (a) Left: superposition of the six monomers in the BsLon-AP asymmetric unit. Side view showing positional variation of the N-terminal helices. Right: diagram showing the extent of the full-length BsLon (774 amino acids) covered by the structures of BsLon-N (amino acids 4–209) and BsLon-AP (amino acids 246–770). (b) Comparison between the architecture of the BsLon-AP complex (left) and that of an $HslU_6V_6$ complex (right, PDB ID: 1G3I). The two ATP-dependent proteases share similar subunit architecture and arrangement within their characteristic complexes, despite having different protease domain folds. The linker between the protease and the ATPase domains in BsLon-AP has very similar topology to the linker between protease HslV and ATPase HslU. (c) Model of a closed-ring hexameric BsLon-AP (N-terminal helices have been removed because they clash), showing the expected arrangement of ATPase and protease domains, within a functional Lon complex. The model was generated by structural alignment of the BsLon-AP protease domains to the protease domains in the hexameric structure of *E. coli* Lon P domain (PDB ID: 1RR9).

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

Model of BsLon-AP

In order to investigate how the BsLon-AP complex might look like when forming a closed ring, in the absence of the helical arrangement, we generated a model based on the hexameric structure of the *E. coli* Lon protease domain. This was done by structurally aligning the protease domain of each BsLon-AP

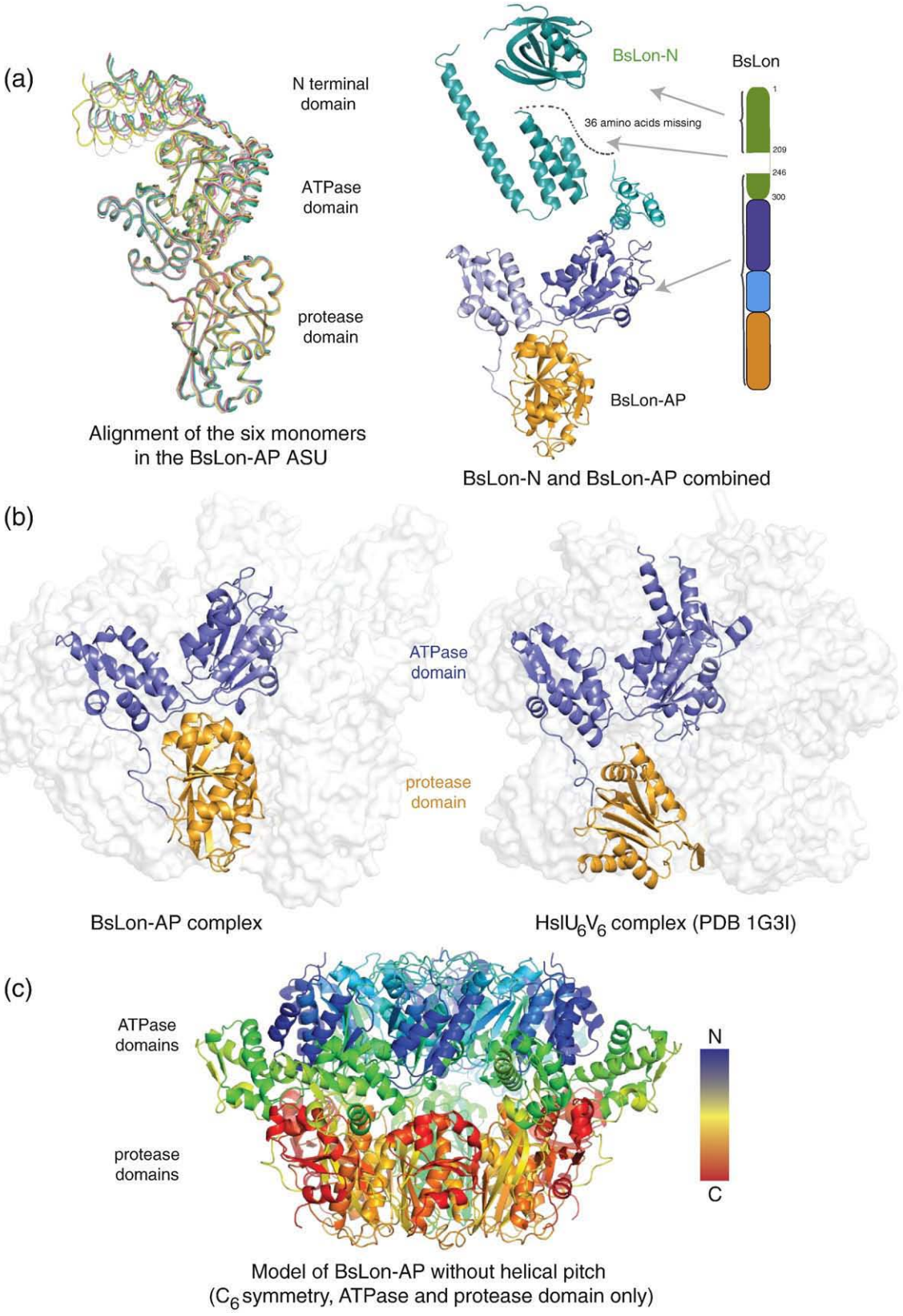

Fig. 5 (*legend on previous page*)

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

subunit to each protease domain of the *E. coli* structure (PDB ID: 1RR9[20]) (Fig. 5c). The orientations of the three BsLon-AP N-terminal helices were incompatible with this model as they clashed with other regions of the model. Upon removal of these three helices from the model, no major steric clashes were detected (with the exception of a few internal, flexible loops). The ATPase domains in the model have very similar interfaces to those found in the BsLon-AP structure, although the packing of the subunits is tighter. The packing of the ATPase domains seen in the BsLon-AP model is very similar to the arrangement seen in the $HslU_6$ ATPase complex,[41] where each α/β sub-domain fits more snugly into the groove between the sub-domains of neighbouring ATPase subunits. Top and bottom views of the model are shown in Supplementary Fig. S6. A side view of the model reveals that, except for the helicity of the BsLon-AP structure, the model and the structure have the same overall conical shape, given by the reduced size of the protease domain ring, compared to the ATPase domain ring (Fig. 5c).

Discussion

Lon and FtsH proteases share a characteristic that distinguishes them from other bacterial ATP-dependent proteases: they harbour their ATP hydrolysis and proteolysis functions within the same polypeptide chain. In addition to the ATPase and the protease domains, cytosolic Lons have a large N-terminal domain, putatively responsible for substrate binding. Coordination of these three different activities, substrate recognition, ATP hydrolysis, and proteolysis, requires the presence of three different functional modules and their arrangement within the same protein seems to show considerable flexibility, as attested here by the positions of the BsLon-AP N-terminal helices. This flexibility might be the reason that Lon has resisted crystallisation for nearly 30 years.

BsLon-N: Implications of the monomer structure

BsLon-N consists of two distinct regions: a globular domain, very similar to the existing structure of the *E. coli* Lon protease N-terminal domain, and an additional four-helix bundle, which includes a short helix marking the start of the coiled-coil segment of BsLon. The geometry of these two domains is intriguing: the right angle between the linker and the long helix of the helical bundle may be a result of the crystal packing since the linker is partly disordered and will not determine the position of the domain. We suggest that, in solution, the globular domain may be able to undergo large movements, independent of the helical bundle. This is, however, counterintuitive to the presence of a hydrophobic surface on the globular domain, which, in solution, would require interaction with another hydrophobic surface (which is covered in our structure by the domain swap). Lon proteases are responsible for degrading misfolded proteins, and this function is believed to rely on their ability to recognise hydrophobic regions that are exposed on the surface of such proteins.[42] The substrate-binding features of Lon proteases should therefore include hydrophobic surfaces that can interact with those on substrates. It is tempting to suggest that the hydrophobic patch on the globular domain of BsLon-N could have a substrate-binding role. However, one could also speculate that these hydrophobic surfaces exist because they form unknown interfaces in the full-length structure.

Functional significance of the dimer

BsLon-N is a structure of a truncated BsLon-N-terminal domain and, as such, the extent to which it represents a functional dimer is uncertain. There are two possible explanations for the packing of the monomers seen in the BsLon-N structure: it may represent a genuine dimerisation of N termini of BsLon or it could be a result of crystal packing and/or the truncation. The dimer interface is extensive and formed mostly by contacts between the helical bundles of the two monomers. Some of these interactions are of a coiled-coil nature.

The extent of the hydrophobic interactions established by the helical bundle and the tightness of the molecular packing of the BsLon-N dimer lend support to a functional relevance of the dimer. In addition to that, the apparent molecular mass derived from the gel-filtration profile of BsLon-N is around 53 kDa, close to the theoretical size of a dimer (47.8 kDa). Coiled-coil interactions are very stable and can often contribute to oligomerisation of complexes. For instance, the N-terminal domains of archaeal and actinobacterial proteasomal ATPases consist of single helices that dimerise by coiled-coil interactions, leading to stable dimers that further assemble into hexamers.[43] The prevalence of Lon dimers over monomers in solution has already been shown by cross-linking experiments on *Brevibacillus thermoruber* Lon, leading to the hypothesis that dimers are the initiators of oligomerisation, which occurs through a dimer–tertramer–hexamer progression.[44] These results may indicate that a dimerisation domain is present somewhere within Lon.

Assuming that the BsLon-N dimer is functionally relevant, however, the anti-parallel arrangement of the monomers is incompatible with the formation of a hexamer, which is the most expected oligomeric state of Lon proteases, based on existing electron microscopy and analytical ultracentrifugation data. The domain-swap arrangement seen in the BsLon-N dimer could be compatible with a head-to-head dimerisation of hexamers, leading to the formation of a dodecamer. The analytical ultracentrifugation results presented in this study reveal that AaLon exists as a mixture of hexameric and larger, perhaps dodecameric species, with the hexameric one being the most prevalent. Although the existence of

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

dodecameric species is not refuted by these results, our electron microscopy images of AaLon show a homogenous sample and fail to reveal the presence of larger species that could correspond to dodecamers (Fig. 1f). Other ATP-dependent proteases, such as the Clp and Hsl complexes, function as stacked oligomeric rings; however, there is no proof that this architecture is also characteristic of Lon proteases. In other words, the BsLon-N dimer seems only functionally relevant as long as Lon proteases exist as dodecameric species, for which there is no evidence, as yet.

The alternative explanation for the dimerisation of BsLon-N monomers is that it may represent a fortuitous arrangement, which facilitates crystal packing by shielding hydrophobic surfaces that are exposed due to the truncation of the N-terminal domain. In the absence of further data, we cannot speculate as to which of these two possibilities (BsLon-N dimer: genuine or artifact) is the most likely.

Although this study proves the validity of the coiled-coil prediction for the N-terminal domain of Lon, it does not solve the ambiguity regarding the inter-molecular *versus* intra-molecular nature of the coiled coil. There is a long coiled-coil domain in ClpB, an ATPase, which, unlike its other Clp ATPase relatives, does not associate with a ClpP protease.[33] ClpB is a chaperone involved in dismantling large aggregates, a function that is critically dependent upon its 85-Å-long and mobile coiled-coil region. It is conceivable that, if not responsible for dimerisation, the coiled-coil region of Lon has the same role as that of ClpB. Interestingly, a substrate-binding and discriminatory role has been implied for the coiled-coil region of *E. coli* Lon, where a single amino acid change, Glu240 to Lys, abolishes recognition of a Lon substrate, RcsA, without affecting the degradation of others.[45]

BsLon-AP

BsLon-AP is the first crystal structure that reveals the relationship between the ATPase and the protease domains of Lon. Within a monomer of BsLon-AP, the two domains form a three-lobed structure in the centre of which lies the nucleotide binding site. This distribution of the three domains around the nucleotide pocket appears functionally significant because it places the protease domain in close proximity of the nucleotide binding site. This is likely to be relevant for the communication between the ATPase and the protease domains during substrate degradation. The observed domain arrangement is in agreement with fluorescence spectroscopy data showing that, upon ATP binding and hydrolysis, the protease domain catalytic residue undergoes conformational changes.[46] ATP hydrolysis causes changes in the relative positions of the two ATPase sub-domains,[41] which could be transmitted to the protease active site via the linker between the two sub-domains, below which the protease domain is situated. The long loop (Arg580 to Asp591), which physically connects the α sub-domain of the ATPase to the protease domain, could also mediate communication between the two functional modules.

The structural relationship between the protease and the ATPase domains of BsLon is remarkably similar to that between the protease (HslV) and the ATPase (HslU) of the HslUV complex (Fig. 5b). Within HslUV, the protease is also located underneath the nucleotide binding site of the ATPase, roughly between the two ATPase sub-domains. Moreover, the link between the ATPase and the protease of HslUV, extending from the C terminus of the α sub-domain, has the same topology as in BsLon-AP[40] (Fig. 5b). This seems even more striking given that HslV and BsLon have different protease folds. The absence of a crystal structure of a complex between the ClpP protease and a Clp ATPase prevents a direct comparison between BsLon and a Clp complex.

The N-terminal domain of BsLon-AP is very flexible

The three helices attached to the N terminus of the ATPase domain belong to the N-terminal domain of BsLon. In the BsLon-AP structure, they form a helical bundle orientated at roughly 90° from a vertical axis passing through the ATPase and the protease domains. A structural alignment of the six monomers in the asymmetric unit, with respect to their ATPase and protease domains, reveals variability in the positions of the N-terminal helical bundles (Fig. 5a, left). Significantly increased temperature factors for this region of the model also denote flexibility of the three helices. This is perhaps not surprising, given that they belong to a domain that was truncated in the BsLon-AP construct and are tethered to the ATPase domain via a long, extended loop that allows significant movement of the helices. It would be expected that, even in the context of the entire N-terminal domain, this loop, consisting of mainly charged and hydrophilic amino acids, would still allow considerable flexibility and relative movement of the two domains it links.

Tendency to crystallise as a helix

BsLon-AP has a tendency to crystallise as a helical assembly, proven by the fact that all five crystal forms presented in this article have such an arrangement of monomers in the asymmetric unit. The most obvious explanation for this behaviour is the absence of the first 239 amino acids of BsLon that may prevent the packing of the BsLon-AP subunits in the correct conformation. The presence of the three N-terminal helices could also be a trigger for the helical arrangement. Their orientation relative to the ATPase and protease domain is likely to be incompatible with the formation of a native complex, as there would be clashes with neighbouring subunits. With this in mind, we attempted the crystallisation of a BsLon truncation containing only the ATPase and the protease domains. This was unsuccessful, due to very poor solubility and unstable protein.

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

The helicity of BsLon-AP could arguably be the result of packing constraints imposed by the truncation; however, it has to be considered that several other related ATPases, such as ClpA,[32] ClpB,[33] and the AAA+ domain of FtsH,[47] also crystallise only as helical arrangements. This indicates that it is more likely to be an intrinsic characteristic of these related ATPases, rather than a specific feature of BsLon-AP, which favours the helical crystal packing.

An alternative explanation for the helical arrangement of the BsLon-AP monomers is the presence of ADP in the nucleotide binding pockets of the six subunits. BsLon-AP was crystallised in the presence of ATPγS; nevertheless, all six monomers in the asymmetric unit are ADP-bound. Lon proteases require ATP for their activity and ADP was shown to have an inhibitory effect upon proteolysis.[48] The ADP-bound conformation seen in the BsLon-AP structure may represent an ADP-inhibited form of Lon, which is not compatible with the formation of a hexameric closed ring, expected for the active Lon complex. Crystallisation of BsLon-AP was also attempted with ATP and AMPPNP; however, the crystals obtained were considerably smaller and weakly diffracting; optimisation of crystallisation conditions failed to improve their diffracting quality beyond 8–9 Å. Nevertheless, HslUV complexes have been crystallised in an ADP-bound form, and their general complex architecture is identical with the ATP-bound form, although conformational differences do exist on a smaller scale, at the level of the ATPase domain.[41]

Limited tryptic digestion experiments indicate that ATP binding causes conformational changes;[49] however, the extent of the structural differences between the nucleotide-free, ATP- and ADP-bound forms of Lon is yet unknown. Enzymatic studies have revealed that the six nucleotide binding sites in active Lon proteases are functionally non-equivalent: half of them have high affinity for ATP, whilst at the same time, the other half have low affinity.[50] These states are interchangeable, as the nucleotide binding sites are structurally identical. The two types of nucleotide affinities are associated with different rates of ATP hydrolysis, with the high-affinity sites being slow hydrolysing and the low-affinity ones being fast hydrolysing.[51] This implies that, in the process of ATP hydrolysis, the nucleotide binding sites of Lon have a mixed population of nucleotides. As a result, an active conformation of the Lon complex may be characterised by the presence of a mixture of nucleotide states, and therefore it may not be compatible with identical nucleotide states of the monomers.

BsLon-AP resembles $HslU_6V_6$

Although the helical arrangement of the hexameric BsLon-AP, in an open-ring conformation, is probably not the functional conformation of Lon, one can deduce from similarities to other known structures that the domain arrangement is conserved. The helix seen in BsLon-AP would rather represent a distorted view of the functional ring-shaped structure expected for a self-compartmentalising protease complex.

A comparison between the structures of the BsLon-AP and $HslU_6V_6$ complexes shows that, in addition to a conserved subunit architecture and structural relationship between the ATPase and the protease domains, the way the subunits assemble into the oligomeric structure is also very similar between the two complexes, as it can be seen in Fig. 5b. The similarity between the AAA+ modules of BsLon and HslU leads to a similarity of interfaces between neighbouring ATPase subunits. In both structures, the α/β sub-domain of the ATPase is nestled in the groove between the ATPase sub-domains of a neighbouring subunit and is located on the inside of the complex, whilst the α sub-domain faces the outside.

The hexameric conformation of the protease domains in BsLon-AP is similar to the *E. coli* Lon structure

Another piece of evidence in support of the near-functional conformation of the helical BsLon-AP structure comes from a comparison between a top view of the arrangement of the six protease domains in BsLon-AP and a similar view of the *E. coli* protease domain structure, solved in isolation[20] (Supplementary Fig. S4). Both structures are hexameric; however, the *E. coli* structure consists of a planar closed ring, as opposed to the open helical ring seen in BsLon-AP. In spite of this difference, the interface between neighbouring domains and their orientation in the complex is almost identical. Also, rather than facing the centre of the channel directly, the catalytic active site is accessible only via a channel between subunits, in both structures (Supupplementary Fig. S4).

The model of BsLon-AP, based on the NCS arrangement of the *E. coli* P domain, can also be used to support the near-native conformation of BsLon-AP. This model is a crude approximation of the expected architecture of the ATPase and protease domains of Lon in the context of a ring-shaped, hexameric complex and the fact that it allows a sensible packing of the ATPase subunits shows that the *E. coli* P domain structure and BsLon-AP are compatible. The fact that both structures are hexameric is in agreement with pre-existing evidence from electron microscopy and analytical ultracentrifugation studies, suggesting that Lon complexes are hexameric in general.

The 36-amino-acid gap

The two structures presented in this article come tantalisingly close to offering a complete picture of a full-length Lon monomer. The BsLon-N structure ends with amino acid 209, whilst BsLon-AP begins with amino acid 246, leaving a gap of 36 amino acids for which we have no structural information (Fig.

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

5a, right). These missing amino acids are predicted to be α-helical and part of the coiled-coil region of the N-terminal domain. It is conceivable that, in the context of the full-length BsLon, these remaining amino acids may fold in a way that would satisfy the hydrophobic requirements of the helical bundle at the C terminus of BsLon. The previously mentioned resemblance between BsLon-N and BPP1347 may be used to argue this point. The topology of the three C-terminal α helices of BPP1347 coincides with the topology of the three C-terminal helices of the other monomer in the BsLon-N dimer. In addition to this, electrostatic surface potential maps of a BsLon-N monomer reveal that, in the absence of a binding partner, three hydrophobic patches would be exposed. Both of these pieces of evidence point towards the same conclusion: in the context of full-length BsLon, the space between the globular domain and the long helix of BsLon-N has to be somehow filled. There are two ways in which this can be achieved: either by the dimerisation of BsLon-N termini, seen in the structure of BsLon-N, or by the folding of the remaining 36 amino acids in this space. These two possibilities are, of course, mutually exclusive, and the answer as to which is the correct one can only come from a structure containing the missing amino acids.

Materials and Methods

Cloning, protein expression, and purification

The gene encoding *B. subtilis* Lon (GI:496557) was amplified from genomic DNA by PCR, using a forward primer containing an NcoI restriction site and a reverse primer carrying an XhoI restriction site. The PCR product was digested with NcoI/XhoI prior to ligation with plasmid pET28a (Novagen), which had also been cut with NcoI/XhoI. The new plasmid, pETBsLon, was used as template for the amplification of the truncated fragment corresponding to the full-length N-terminal domain of BsLon, using the same forward primer, as per pETBsLon, and a reverse primer containing an XhoI restriction site. This gene fragment was cut with NcoI/XhoI and ligated into NcoI/XhoI-digested pET28a, resulting in plasmid pETBsLonNf, which encodes amino acids 1–297 of BsLon.

In order to change the catalytic Ser677 to Ala, we introduced a single-point mutation in the sequence of pETBsLon, using a QuikChange Mutagenesis Kit (Stratagene). The new plasmid $pETBsL_{S677A}$ was used as template for the PCR amplification of gene fragment BsLon-AP, encoding for residues 240–774 of BsLon, using a forward primer carrying an NcoI restriction site and the same reverse primer as per BsLon. This fragment was similarly cut with NcoI/XhoI and ligated with NcoI/XhoI-digested pET28a, resulting in plasmid $pETBsL239_{S677A}$.

The gene encoding *A. aeolicus* Lon (GI: 2982953) was amplified by PCR, using a forward primer containing an NdeI restriction site and a reverse primer with a BamHI restriction site. The PCR product was digested with NdeI/BamHI and ligated with pHis17 plasmid (Bruno Miroux, MRC-LMB, personal communication), which had been cut with the same restriction enzymes.

All proteins, AaLon, BsLon, BsLon-AP, and BsLon-Nf, had a hexa-histidine tag at their C termini and were expressed in BL21(AI) cells (Invitrogen). Following introduction of plasmids into cells by electroporation, cells were first grown overnight, at 37 °C, on six nutrient agar plates containing kanamycin (for BsLon, BsLon-AP, and BsLon-N) and ampicillin (for AaLon). All cells from the six agar plates were then transferred to 12 l of 2× TY medium containing 50 μg/ml of kanamycin, or 100 μg/ml of ampicillin, and grown at 37 °C. Protein synthesis was induced at OD_{600} (optical density at 600 nm) of 0.5 by the addition of L-arabinose (Carbosynth) to a final concentration of 0.2%, followed by overnight incubation at 25 °C. Cells were harvested the next day and the pellet was frozen in liquid nitrogen and kept at −80 °C, until needed.

For purification, the cells expressing BsLon or BsLon-AP were resuspended in 50 mM Tris–HCl, pH 8.5, and 300 mM NaCl with small amounts of DNase and lysed using a cell disruptor (Constant Systems). The lysate was cleared by centrifugation for 60 min at 40,000 rpm (25,171*g*) and the supernatant was loaded onto 2× 5 ml Ni-NTA HisTrap HP columns (GE Healthcare), which were then washed thoroughly with 50 mM Tris–HCl, pH 8.5, and 300 mM NaCl. The proteins were eluted with the same buffer containing 300 mM imidazole and subsequently diluted 10-fold with 20 mM Capso, pH 9.5, 1 mM ethylenediaminetetraacetic acid (EDTA), and 1 mM NaN_3, loaded onto 2×5 ml HiTrap Q columns (GE Healthcare) and eluted in a linear gradient of 0 to 1 M NaCl in the same buffer. This step was followed by concentration of proteins using Vivaspin 20 concentrators with a 30,000 molecular weight cutoff (Sartorius Stedim Biotech) and loading onto a HiPrep Sephacryl S300 gel-filtration column (GE Healthcare) equilibrated in 20 mM Tris–HCl, pH 8.5, 1 mM EDTA, 1 mM NaN_3 and 300 mM NaCl. BsLon-AP eluted as a single peak, at an apparent molecular mass of 139 kDa, and was concentrated to 20 mg/ml for use in crystallisation experiments. BsLon suffered from degradation during purification and eluted as several peaks. The most pure fractions were collected and concentrated to 20 mg/ml.

For purification of BsLon-Nf, the protein was precipitated from the lysate supernatant by addition of ammonium sulfate to a final concentration of 20% (saturation). This was done gradually, whilst the protein was incubated at 4 °C. The precipitated protein was spun at 15,000 rpm (17,590*g*) and subsequently resuspended in 250 ml of 20 mM Tris–HCl, pH 8.5, 1 mM EDTA, and 1 mM NaN_3. The protein concentration was measured using a NanoDrop ND-1000 Spectrophotometer (Thermo) and α-chymotrypsin (Sigma C3142) was added to a concentration equivalent to 500 times less than the measured concentration of BsLon-Nf. Incubation at 37 °C for 1 h was followed by loading of the protein solution onto 2×5 ml HiTrap Q HP and elution with a liner gradient of 0 to 1 M NaCl in 20 mM Tris–HCl, pH 8.5, 1 mM EDTA, and 1 mM NaN_3. The eluted protein, labelled BsLon-N, was concentrated and further loaded onto a HiPrep Sephacryl S200 gel-filtration column, equilibrated in 20 mM Tris–HCl, pH 8.5, 1 mM EDTA, 1 mM NaN_3, and 50 mM NaCl. BsLon-N eluted as a single peak, at an apparent molecular mass of 53 kDa, and was concentrated to 15 mg/ml and frozen. This limited proteolysis fragment had a molecular mass of 23.9 kDa, as shown by mass spectrometry analysis, and its N terminus coincided with that of BsLon, as confirmed by N-terminal sequencing.

Purification of AaLon followed a very similar protocol to that used for BsLon, the main difference being the buffer used for both HiTrap Q and Sephacryl S300 columns: 20 mM Tris–HCl, pH 7.5, 1 mM EDTA, and 1 mM NaN_3. AaLon

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

eluted as a single peak, at an apparent molecular mass of 579 kDa, and was concentrated to 10 mg/ml and frozen.

Western blot

Twenty crystals from BsLon crystallisation trials were dissolved in 40 μl of SDS-PAGE loading buffer and run on SDS-PAGE gel, transferred to a polyvinylidene fluoride membrane, and incubated with HisProbe horseradish peroxidase (Pierce). The Western blot was developed using SuperSignal West Pico Substrate (Pierce).

Crystallisation

Crystals of BsLon-AP and BsLon-N were grown by the sitting drop vapour diffusion technique at 19 °C. More than 1500 initial crystallisation conditions were screened using our in-house high-throughput facility[52] and an Innovadyne Screenmaker 96+8 crystallisation robot, set up to dispense drops containing 100 nl of mother liquor and 100 nl of protein.

BsLon-AP crystals grew in many different crystallisation conditions; the final optimised condition contained 100 mM sodium citrate, pH 5.6, 500 mM ammonium acetate, and 14% polyethylene glycol 4000. Crystals were cryo-protected in mother liquor containing 20% (v/v) ethylene glycol, before being flash-frozen. The best diffracting BsLon-N crystals were grown in 100 mM sodium citrate, pH 4.0, and 5% (w/v) polyethylene glycol 6000 and were cryo-protected using mother liquor containing 25% (v/v) ethylene glycol.

Structure determination

The BsLon-AP structure was solved using tantalum SAD and extensive phase extension with multi-crystal averaging. A tantalum bromide cluster ($Ta_6Br_{12}^{2+}$)-soaked crystal of the $P2_12_12_1$ crystal form was used for a tantalum SAD experiment at the L-III edge. This yielded a data set (TABR, Table 1) with very high anomalous correlation and enabled the location of sites with HYSS[53] at 7 Å resolution. Phases were obtained using Phaser 2.1's SAD module[54] and subsequent solvent flattening using DM[55] yielded a clear map, showing seven monomers. The densities of the seven monomers were filled by hand with secondary structural elements, enabling 7-fold averaging in DM. Phase extension required the use of separate masks for the two lobes of the ATPase domain and the protease domain, since they were found in slightly different orientations, and the higher-resolution isomorphous native data set (NATI, Table 1) reaching 3.5 Å. The resulting density was very clear and showed the disorganised arrangement of the seven subunits in the crystal. For solving the other crystal forms (Table 1), the density for a monomer was cut from the phase-extended map and used as a search model in Phaser 2.1. The structure was built for the $P2_1$ crystal form since it shows the washer-type hexamer. For building, a phase-extended electron density map was generated at 3.4 Å resolution by multi-crystal averaging using the original TABR phases, NATI, and the $P2_1$ data set, using the three masks mentioned above. For this, phases were truncated at 8 Å initially and then extended to 3.4 Å slowly in 1000 steps. The resulting map was buildable and a complete model of the hexamer was built manually using MAIN[56] and with the published structures of the ATPase domain and protease domains serving as guides. The structure was refined with PHENIX.refine[57] using tight NCS restraints and TLS. As expected at this resolution, temperature factors are high (no upper limit in PHENIX.refine) and especially so in the N-terminal part of the structure (three helices, residues 246–298) that is more flexible and more weakly defined than the rest of the structure, but clearly visible in kicked (0.5 Å) omit $2F_o-F_c$ difference maps. The Wilson *B*-factor of the data is 74.7 Å^2. The refined structure of the hexameric C-terminal domain (BsLon-AP) was deposited in the PDB with ID 3M6A.

The BsLon-N structure was solved by molecular replacement using an existing (but smaller) domain X-ray crystal structure of *E. coli* Lon N-terminal domain (PDB ID: 2ANE). The eight monomers in 2ANE were superimposed and used as an ensemble in Phaser 2.1, yielding a solution with two monomers per asymmetric unit. Because the search model only covers about half of the molecule 2-fold averaging with DM, using an expanded mask based on an initial $2F_o-F_c$ map was used to render the electron density map buildable. The model was then built manually using MAIN and refined with PHENIX.refine. The refined structure of BsLon-N was deposited in the PDB with ID 3M65.

Analytical ultracentrifugation

Sedimentation velocity and equilibrium runs were performed in a Beckman Optima XL-I analytical ultracentrifuge with an An-60 Ti rotor, using interference and absorbance at 280 nm. Sedimentation velocity experiments were run at 45,000 rpm in two-channel Epon charcoal-filled centerpieces, at 4 °C. Three different concentrations were used for both BsLon and AaLon: 1, 2.5, and 5 mg/ml. The data were analysed with the program Sedfit[58] and are presented as a continuous distribution *c*(*s*) [fringes/Svedberg] against the sedimentation coefficient [Svedberg]. Sedimentation equilibrium experiments were carried out in six-sector 12-mm path length cells, at 4 °C. The concentrations used for AaLon were 1, 2, and 8 mg/ml, whilst BsLon samples were run at 0.5, 1.14, and 2 mg/ml. The samples were centrifuged at 9000 rpm until they reached equilibrium, as judged by the lack of changes in subsequent scans. The data were analysed using the program Sedphat† and Ultraspin‡. Partial specific volumes were calculated from the amino acid composition of the proteins, using SEDNTERP,[59] also employed for the estimation of the density and viscosity of buffers.

Electron microscopy

Negative stain electron microscopy images were collected using a Tecnai 12 (FEI) electron microscope operating at 120 kV and a magnification of 50,000×. AaLon was diluted into 20 mM Tris–HCl, pH 7.5, 1 mM EDTA, and 1 mM NaN_3 to a concentration of 0.01 mg/ml and BsLon was diluted into 20 mM Tris–HCl, pH 8.5, 1 mM EDTA, and 1 mM NaN_3 to a final concentration of 0.1 mg/ml. For BsLon, higher concentrations of 0.5, 1, and 5 mg/ml were also tried, with identical results (data not shown). Carbon-coated grids were made hydrophilic by glow discharging for 15 s, using an Edwards Sputter Coater S150B. Within 15 min after glow discharging, 5 μl of samples was applied to the grids and incubated for 1 min. Excess sample was blotted off and grids were stained by application of three drops of 2% uranyl acetate.

†http://www.analyticalultracentrifugation.com/sedphat/sedphat.htm

‡http://ultraspin.mrc-cpe.cam.ac.uk

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

Accession numbers

Coordinates and structure factors have been deposited in the PDB with accession numbers 3M65, for BsLon-N, and 3M6A, for BsLon-AP.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2010.06.030

References

1. Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A. *et al.* (2004). Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell*, **119**, 9–18.
2. Lupas, A., Flanagan, J. M., Tamura, T. & Baumeister, W. (1997). Self-compartmentalizing proteases. *Trends Biochem. Sci.* **22**, 399–404.
3. Baker, T. A. & Sauer, R. T. (2006). ATP-dependent proteases of bacteria: recognition logic and operating principles. *Trends Biochem. Sci.* **31**, 647–653.
4. Charette, M. F., Henderson, G. W. & Markovitz, A. (1981). ATP hydrolysis-dependent protease activity of the lon (capR) protein of *Escherichia col*i K-12. *Proc. Natl Acad. Sci. USA*, **78**, 4728–4732.
5. Murakami, K., Voellmy, R. & Goldberg, A. L. (1979). Protein degradation is stimulated by ATP in extracts of *Escherichia coli*. *J. Biol. Chem.* **254**, 8194–8200.
6. Chung, C. H. & Goldberg, A. L. (1981). The product of the *lon (capR)* gene in *Escherichia coli* is the ATP-dependent protease, protease La. *Proc. Natl Acad. Sci. USA*, **78**, 4931–4935.
7. Van Dyck, L., Pearce, D. A. & Sherman, F. (1994). PIM1 encodes a mitochondrial ATP-dependent protease that is required for mitochondrial function in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* **269**, 238–242.
8. Gottesman, S. & Maurizi, M. R. (1992). Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol. Rev.* **56**, 592–621.
9. Ogura, T. & Wilkinson, A. J. (2001). AAA+ superfamily ATPases: common structure—diverse function. *Genes Cells*, **6**, 575–597.
10. Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. (1999). AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**, 27–43.
11. Lupas, A. N. & Martin, J. (2002). AAA proteins. *Curr. Opin. Struct. Biol.* **12**, 746–753.
12. Smith, C. K., Baker, T. A. & Sauer, R. T. (1999). Lon and Clp family proteases and chaperones share homologous substrate-recognition domains. *Proc. Natl Acad. Sci. USA*, **96**, 6678–6682.
13. Beuron, F., Maurizi, M. R., Belnap, D. M., Kocsis, E., Booy, F. P., Kessel, M. & Steven, A. C. (1998). At sixes and sevens: characterization of the symmetry mismatch of the ClpAP chaperone-assisted protease. *J. Struct. Biol.* **123**, 248–259.
14. Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H. D. & Huber, R. (2000). The structures of HsIU and the ATP-dependent protease HsIU–HsIV. *Nature*, **403**, 800–805.
15. Rotanova, T. V., Melnikov, E. E., Khalatova, A. G., Makhovskaya, O. V., Botos, I., Wlodawer, A. & Gustchina, A. (2004). Classification of ATP-dependent proteases Lon and comparison of the active sites of their proteolytic domains. *Eur. J. Biochem.* **271**, 4865–4871.
16. Melnikov, E. E., Andrianova, A. G., Morozkin, A. D., Stepnov, A. A., Makhovskaya, O. V., Botos, I. *et al.* (2008). Limited proteolysis of *E. coli* ATP-dependent protease Lon—a unified view of the subunit architecture and characterization of isolated enzyme fragments. *Acta Biochim. Pol.* **55**, 281–296.
17. Rotanova, T. V., Botos, I., Melnikov, E. E., Rasulova, F., Gustchina, A., Maurizi, M. R. & Wlodawer, A. (2006). Slicing a protease: structural features of the ATP-dependent Lon proteases gleaned from investigations of isolated domains. *Protein Sci.* **15**, 1815–1828.
18. Roudiak, S. G. & Shrader, T. E. (1998). Functional role of the N-terminal region of the Lon protease from *Mycobacterium smegmatis*. *Biochemistry*, **37**, 11255–11263.
19. Amerik, A. Y., Antonov, V. K., Gorbalenya, A. E., Kotova, S. A., Rotanova, T. V. & Shimbarevich, E. V. (1991). Site-directed mutagenesis of La protease. A catalytically active serine residue. *FEBS Lett.* **287**, 211–214.
20. Botos, I., Melnikov, E. E., Cherry, S., Tropea, J. E., Khalatova, A. G., Rasulova, F. *et al.* (2004). The catalytic domain of *Escherichia coli* Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J. Biol. Chem.* **279**, 8140–8148.
21. Dodson, G. & Wlodawer, A. (1998). Catalytic triads and their relatives. *Trends Biochem. Sci.* **23**, 347–352.
22. Ward, D. E., Shockley, K. R., Chang, L. S., Levy, R. D., Michel, J. K., Conners, S. B. & Kelly, R. M. (2002). Proteolysis in hyperthermophilic microorganisms. *Archaea*, **1**, 63–74.
23. Goldberg, A. L., Moerschell, R. P., Chung, C. H. & Maurizi, M. R. (1994). ATP-dependent protease La (lon) from *Escherichia coli*. *Methods Enzymol.* **244**, 350–375.
24. Park, S. C., Jia, B., Yang, J. K., Van, D. L., Shao, Y. G., Han, S. W. *et al.* (2006). Oligomeric structure of the ATP-dependent protease La (Lon) of *Escherichia coli*. *Mol. Cell*, **21**, 129–134.
25. Stahlberg, H., Kutejova, E., Suda, K., Wolpensinger, B., Lustig, A., Schatz, G. *et al.* (1999). Mitochondrial Lon of *Saccharomyces cerevisiae* is a ring-shaped protease with seven flexible subunits. *Proc. Natl Acad. Sci. USA*, **96**, 6787–6790.
26. Rudyak, S. G., Brenowitz, M. & Shrader, T. E. (2001). Mg^{2+}-linked oligomerization modulates the catalytic activity of the Lon (La) protease from *Mycobacterium smegmatis*. *Biochemistry*, **40**, 9317–9323.
27. Botos, I., Melnikov, E. E., Cherry, S., Kozlov, S., Makhovskaya, O. V., Tropea, J. E. *et al.* (2005). Atomic-resolution crystal structure of the proteolytic domain of *Archaeoglobus fulgidus* lon reveals the conformational variability in the active sites of lon proteases. *J. Mol. Biol.* **351**, 144–157.
28. Li, M., Rasulova, F., Melnikov, E. E., Rotanova, T. V., Gustchina, A., Maurizi, M. R. & Wlodawer, A. (2005). Crystal structure of the N-terminal domain of *E. coli* Lon protease. *Protein Sci.* **14**, 2895–2900.
29. Botos, I., Melnikov, E. E., Cherry, S., Khalatova, A. G., Rasulova, F. S., Tropea, J. E. *et al.* (2004). Crystal structure of the AAA+ alpha domain of *E. coli* Lon protease at 1.9A resolution. *J. Struct. Biol.* **146**, 113–122.

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030

30. Im, Y. J., Na, Y., Kang, G. B., Rho, S. H., Kim, M. K., Lee, J. H. *et al.* (2004). The active site of a lon protease from *Methanococcus jannaschii* distinctly differs from the canonical catalytic Dyad of Lon proteases. *J. Biol. Chem.* **279**, 53451–53457.
31. Garcia-Nafria, J., Ondrovicova, G., Blagova, E., Levdikov, V. M., Bauer, J. A., Suzuki, C. K. *et al.* (2010). Structure of the catalytic domain of the human mitochondrial Lon protease: proposed relation of oligomer formation and activity. *Protein Sci.* **19**, 987–999.
32. Guo, F., Maurizi, M. R., Esser, L. & Xia, D. (2002). Crystal structure of ClpA, an Hsp100 chaperone and regulator of ClpAP protease. *J. Biol. Chem.* **277**, 46743–46752.
33. Lee, S., Sowa, M. E., Watanabe, Y. H., Sigler, P. B., Chiu, W., Yoshida, M. & Tsai, F. T. (2003). The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. *Cell*, **115**, 229–240.
34. Lupas, A., Van Dyke, M. & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science*, **252**, 1162–1164.
35. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. (2008). Searching protein structure databases with DaliLite v.3. *Bioinformatics*, **24**, 2780–2781.
36. Besche, H., Tamura, N., Tamura, T. & Zwickl, P. (2004). Mutational analysis of conserved AAA+ residues in the archaeal Lon protease from *Thermoplasma acidophilum*. *FEBS Lett.* **574**, 161–166.
37. Sousa, M. C., Trame, C. B., Tsuruta, H., Wilbanks, S. M., Reddy, V. S. & McKay, D. B. (2000). Crystal and solution structures of an HslUV protease–chaperone complex. *Cell*, **103**, 633–643.
38. Lenzen, C. U., Steinmann, D., Whiteheart, S. W. & Weis, W. I. (1998). Crystal structure of the hexamerization domain of *N*-ethylmaleimide-sensitive fusion protein. *Cell*, **94**, 525–536.
39. Hattendorf, D. A. & Lindquist, S. L. (2002). Cooperative kinetics of both Hsp104 ATPase domains and interdomain communication revealed by AAA sensor-1 mutants. *EMBO J.* **21**, 12–21.
40. Wang, J., Song, J. J., Franklin, M. C., Kamtekar, S., Im, Y. J., Rho, S. H. *et al.* (2001). Crystal structures of the HslVU peptidase–ATPase complex reveal an ATP-dependent proteolysis mechanism. *Structure*, **9**, 177–184.
41. Wang, J., Song, J. J., Seong, I. S., Franklin, M. C., Kamtekar, S., Eom, S. H. & Chung, C. H. (2001). Nucleotide-dependent conformational changes in a protease-associated ATPase HsIU. *Structure*, **9**, 1107–1116.
42. Gur, E. & Sauer, R. T. (2009). Degrons in protein substrates program the speed and operating efficiency of the AAA+ Lon proteolytic machine. *Proc. Natl Acad. Sci. USA*, **106**, 18503–18508.
43. Djuranovic, S., Hartmann, M. D., Habeck, M., Ursinus, A., Zwickl, P., Martin, J. *et al.* (2009). Structure and activity of the N-terminal substrate recognition domains in proteasomal ATPases. *Mol. Cell*, **34**, 580–590.
44. Lee, A. Y., Hsu, C. H. & Wu, S. H. (2004). Functional domains of *Brevibacillus thermoruber* lon protease for oligomerization and DNA binding: role of N-terminal and sensor and substrate discrimination domains. *J. Biol. Chem.* **279**, 34903–34912.
45. Ebel, W., Skinner, M. M., Dierksen, K. P., Scott, J. M. & Trempy, J. E. (1999). A conserved domain in *Escherichia coli* Lon protease is involved in substrate discriminator activity. *J. Bacteriol.* **181**, 2236–2243.
46. Patterson-Ward, J., Huang, J. & Lee, I. (2007). Detection and characterization of two ATP-dependent conformational changes in proteolytically inactive *Escherichia coli* Lon mutants by stopped flow kinetic techniques. *Biochemistry*, **46**, 13593–13605.
47. Niwa, H., Tsuchiya, D., Makyio, H., Yoshida, M. & Morikawa, K. (2002). Hexameric ring structure of the ATPase domain of the membrane-integrated metalloprotease FtsH from *Thermus thermophilus* HB8. *Structure*, **10**, 1415–1423.
48. Waxman, L. & Goldberg, A. L. (1985). Protease La, the lon gene product, cleaves specific fluorogenic peptides in an ATP-dependent reaction. *J. Biol. Chem.* **260**, 12022–12028.
49. Patterson, J., Vineyard, D., Thomas-Wohlever, J., Behshad, R., Burke, M. & Lee, I. (2004). Correlation of an adenine-specific conformational change with the ATP-dependent peptidase activity of *Escherichia coli* Lon. *Biochemistry*, **43**, 7432–7442.
50. Menon, A. S. & Goldberg, A. L. (1987). Binding of nucleotides to the ATP-dependent protease La from *Escherichia coli*. *J. Biol. Chem.* **262**, 14921–14928.
51. Vineyard, D., Patterson-Ward, J., Berdis, A. J. & Lee, I. (2005). Monitoring the timing of ATP hydrolysis with activation of peptide cleavage in *Escherichia coli* Lon by transient kinetics. *Biochemistry*, **44**, 1671–1682.
52. Stock, D., Perisic, O. & Löwe, J. (2005). Robotic nanolitre protein crystallisation at the MRC Laboratory of Molecular Biology. *Prog. Biophys. Mol. Biol.* **88**, 311–327.
53. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J. *et al.* (2008). Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* **426**, 419–435.
54. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
55. Collaborative Computational Project, No. 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763.
56. Turk, D. (1992). Weiterentwicklung eines Programms für Molekülgraphik und Elektronendichte-Manipulation und seine Anwendung auf verschiedene Protein Strukturaufklärungen. PhD Thesis, Technical University of Munich, Germany.
57. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W. *et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **58**, 1948–1954.
58. Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619.
59. Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). *Analytical Ultracentrifugation in Biochemistry and Polymer Science.* Royal Society of Chemistry, Cambridge, UK.

Please cite this article as: Duman, R. E., Löwe, J., Crystal Structures of *Bacillus subtilis* Lon Protease, *J. Mol. Biol.* (2010), doi:10.1016/j.jmb.2010.06.030