

STRUCTURE NOTE

Open and compressed conformations of *Francisella tularensis* ClpP

Laura Díaz-Sáez, Genady Pankov, and William N. Hunter*

Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dundee, United Kingdom

ABSTRACT

Caseinolytic proteases are large oligomeric assemblies responsible for maintaining protein homeostasis in bacteria and in so doing influence a wide range of biological processes. The functional assembly involves three chaperones together with the oligomeric caseinolytic protease catalytic subunit P (ClpP). This protease represents a potential target for therapeutic intervention in pathogenic bacteria. Here, we detail an efficient protocol for production of recombinant ClpP from *Francisella tularensis*, and the structural characterization of three crystal forms which grow under similar conditions. One crystal form reveals a compressed state of the ClpP tetradecamer and two forms an open state. A comparison of the two types of structure infers that differences at the enzyme active site result from a conformational change involving a highly localized disorder-order transition of a β -strand α -helix combination. This transition occurs at a subunit-subunit interface. Our study may now underpin future efforts in a structure-based approach to target ClpP for inhibitor or activator development.

Proteins 2017; 85:188–194.

© 2016 The Authors Proteins: Structure, Function, and Bioinformatics Published by Wiley Periodicals, Inc.

Key words: protease; active site; conformational change; X-ray crystallography.

INTRODUCTION

The ATP-dependent caseinolytic proteases (Clp) maintain protein homeostasis in bacteria, especially under conditions of stress.¹ These proteases, by virtue of regulating protein levels, have pleiotropic effects and indirectly modulate gene expression, cell motility and division, the removal of stress-damaged proteins and are also implicated in pathogenesis by promoting virulence factor expression.^{1,2}

The functional Clp protease is a large multi-protein complex, with an assembly of 14 subunits of the promiscuous serine protease ClpP at the core. This ClpP tetradecamer, a dimer of heptamers each arranged in a ring, displays an overall cylindrical shape forming an internal chamber where proteolysis of unfolded proteins occurs. The catalytic core is surrounded by accessory proteins, AAA⁺ (ATPases Associated with diverse cellular Activities) chaperones termed ClpA, ClpX and ClpC, that regulate the presentation of substrates into the active sites.^{1,3}

Two axial pores are formed by the complex and these allow peptide substrates access to the active sites when an open conformation is present.⁴ The products are released through pores in the equatorial plane that are generated after a conformational change to a compressed state.^{4,5} ClpP, which is highly conserved across species,

Additional Supporting Information may be found in the online version of this article.

Abbreviations: AAA, ATPases Associated with diverse cellular Activities; ASA, accessible surface area

Grant sponsor: Defense Science and Technology Laboratory (Dstl, U.K.); Grant sponsor: The Wellcome Trust; grant number: grant 094090.

*Correspondence to: William N. Hunter, Division of Biological Chemistry and Drug Discovery, University of Dundee, Dow Street, Dundee DD1 5EH, United Kingdom. E-mail: w.n.hunter@dundee.ac.uk

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Received 25 July 2016; Revised 13 October 2016; Accepted 24 October 2016

Published online 1 November 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.25197

has been the subject of a number of crystallographic studies.⁶ The open and compressed conformations of ClpP have been characterized for the enzymes from *Bacillus subtilis* (BsClpP),⁵ *Mycobacterium tuberculosis*,⁷ and *Staphylococcus aureus*.^{4,8}

Clp proteases have been investigated as antibacterial drug targets and some inhibitors do appear to attenuate the virulence of *S. aureus*⁹ and *Listeria monocytogenes*.¹⁰ In addition, compounds that activate the complex can enhance unspecific proteolysis and therefore, also have the potential to act as antibiotics.¹¹ We have identified ClpP as a potential drug target in the facultative intracellular bacterium *Francisella tularensis*, based on similar considerations applied to targets in *Pseudomonas aeruginosa*.¹² *F. tularensis* is a highly virulent Gram-negative aerobe classified as a category (A) bio-warfare agent by the Centers for Disease Control and Prevention (<http://www.cdc.gov>). We initiated a study of *F. tularensis* ClpP (*FtClpP*) to lay the foundation for early stage drug discovery¹³ targeting this enzyme. A structure of apo-*FtClpP*, in an open conformation at a resolution of 2.3 Å, was available in the Protein Data Bank (code: 3P2L) but there is no associated publication and only limited experimental details are available. Here we present the efficient production of a recombinant source of *FtClpP*, the crystal structure of a compressed form of the enzyme at 2.8 Å resolution and two crystal forms that reveal the structure of an open state at 1.9 and 1.7 Å resolution.

PDB codes: *FtClpP*:I 5G1Q, *FtClpP*:II 5G1R and *FtClpP*:III 5G1S.

MATERIALS AND METHODS

Recombinant protein production

Synthetic DNA, codon optimized for expression in *Escherichia coli* K12, encoding *FtClpP* (UniProt: Q5NH47) was purchased (GeneWiz). This gene was cloned into a modified pET15b vector (Novagen) that adds a hexahistidine tag (His₆-tag) followed by a TEV (tobacco etch virus) protease cleavage site at the N-terminus of the product. The restriction enzymes used were BamHI (at the 5' end) and NdeI (at the 3' end of the DNA strand). The construct was sequenced as a check. Recombinant expression was carried out in *E. coli* BL21(DE3). The bacteria were cultured in 1 L of LB broth containing 50 µg mL⁻¹ carbenicillin at 37°C until an O.D. of 0.6 at λ = 600 nm was achieved then the temperature was decreased to 20°C. Gene expression was induced by adding 1 mM IPTG and continued overnight. Cultures were centrifuged (4,000 g at 4°C for 10 minutes) to harvest the cells. The sample was then resuspended in buffer A (20 mM Tris-HCl pH 7.4 and 200 mM NaCl) containing an EDTA-free protease inhibitor cocktail tablet (Roche). Cell disruption was carried out using a French press (20 psi) followed by centrifugation (40,000g

for 40 minutes at 4°C). The soluble proteins were loaded on a 5 mL HisTrap HP column (GE Healthcare), loaded with Ni²⁺, to perform affinity chromatography. The buffers employed for *FtClpP* purification were buffer A and buffer B (buffer A with 0.8M imidazole). A fraction at 20% buffer B contained the desired product *FtClpP*. The sample was incubated with His₆-tagged TEV protease (1 mg per 20 mg protein) at 4°C overnight while performing dialysis into buffer A to remove the imidazole. A second affinity purification step was then performed under the same conditions but this time to separate the His₆-tag and TEV protease from the desired product. Size exclusion chromatography was then pursued with *FtClpP* with a HiLoad 16/60 superdex 200 prep grade column (GE Healthcare) and buffer A. The yield was 10 mg *FtClpP* per litre of bacterial culture and a high level of protein purity was confirmed using SDS-PAGE and MALDI-TOF-MS analyses. Protein concentration was determined using the predicted ε(*FtClpP*) = 8940 M⁻¹ cm⁻¹ calculated with PROTPARAM.¹⁴

Crystallization

FtClpP samples were concentrated to 20 mg mL⁻¹ and various commercially available crystallization conditions were tested using 1:1 and 1:2 protein solution to reservoir ratios (final drop volumes of 0.2 and 0.3 µL, respectively) in sitting drop 96-well plates. Orthorhombic crystals were obtained from The Classics Suite (Molecular Dimensions) condition B4 (0.2M NaCl, 30% v/v MPD (2-methyl-2,4-pentanediol) and 0.1M sodium acetate pH 4.6) and optimized using 1:1 protein solution to reservoir ratio in a final drop volume of 4 µL with the hanging drop method.

Crystallographic analyses

Crystals were flash-cooled directly from drops and diffracted to 2.8 Å resolution in-house. These samples were stored for use at the synchrotron. In one case, data were collected in-house using a Rigaku M007HF X-ray generator with a Saturn 944HG+ CCD detector, a 10° θ offset to compensate for a large mosaic spread of 0.9°, 10 seconds of exposure per 0.5° oscillation covering 115°. Data integration and analysis were carried out using iMOSFLM¹⁵ and AIMLESS.¹⁶ The crystal displayed space group *P*₂₁₂₁₂ and unit cell dimensions of *a* = 113.5, *b* = 125.9, *c* = 96.95 Å. The Matthews coefficient calculated with the CCP4i suite¹⁷ suggested seven monomers in the asymmetric unit (*V*_m 2.2 Å³ Da⁻¹ and bulk solvent about 45%). The *FtClpP* structure with PDB code 3P2L (2.3 Å resolution, space group *P*₂₁₂₁₂ and unit cell dimensions *a* = 120.52, *b* = 128.82, *c* = 98.03 Å), also with seven subunits in the asymmetric unit, provided a model of a single subunit, without residues 128–149, for molecular replacement calculations in PhaserMR.¹⁸ This structure is termed *FtClpP*:I and

Table I
Crystallographic Statistics

Structure	<i>FtClpP</i> :I	<i>FtClpP</i> :II	<i>FtClpP</i> :III
PDB code	5G1Q	5G1R	5G1S
Wavelength (Å)	1.5418	0.9799	0.9799
Space group	$P2_12_12$	$P2_12_12$	$P2_12_1$
Unit cell dimension <i>a</i> , <i>b</i> , <i>c</i> (Å)	113.5 125.9 96.9	117.9 128.2 98.3	98.2 127.9 353.8
Resolution range ^a (Å)	63.61–2.84	47.04–1.90	48.80–1.70
No. reflections	90264	838386	2604010
Unique reflections	32741	117661	460688
Completeness (%)	97.8 (98.8)	100 (100)	94.8 (70.2)
R_{merge}^b	0.098 (0.344)	0.085 (0.460)	0.084 (0.423)
Redundancy	2.8 (2.0)	7.1 (7.1)	5.7(2.5)
$\langle I/\sigma(I) \rangle$	5.9 (2.2)	18.5 (4.4)	9.3 (1.4)
Wilson <i>B</i> (Å ²)	15.6	10.8	9.8
$R_{\text{work}}^c/R_{\text{free}}^d$	0.249/0.326	0.201/0.269	0.204/0.242
R_{free} reflections (%)	4.99	4.98	4.98
DPI ^e (Å)	0.52	0.17	0.13
R.m.s.d. bond lengths (Å)/angles (°)	0.0079/1.243	0.018/1.875	0.008/1.285
Protein atoms/residues (Average <i>B</i> -factors, Å ²)	9337/1217 (37.6)	11663/2423 (20.9)	33173/8031 (14.6)
Water molecules (Average <i>B</i> -factors, Å ²)	–	1085 (25.8)	4011 (23.8)
Ligands (Average <i>B</i> -factors, Å ²)	–	5 MPD (47.1) 7 acetate (37.7)	32 MPD (40.1) 21 acetate (24.5)
Ramachandran analyses			
Outliers (%)	0.8	0.2	0.04
Favoured regions (%)	93	96.9	97.8
Rotamer outliers (%)	3.6	2.6	1.9

represents the closed or compressed form of the tetradecamer assembly. Diffraction data were also collected at ESRF beamline BM30-A from two samples, both of which displayed low mosaic spread (0.2°). High-resolution data were processed and analyzed using XDS¹⁹ and AIMLESS. These crystal forms displayed similar morphology and the same orthorhombic symmetry, space group ($P2_12_12$ or an alternative setting) but two distinct forms are evident with unit cell dimensions of $a = 117.90$, $b = 128.20$, $c = 98.26$ Å and $a = 98.49$, $b = 128.23$, $c = 355.18$. These structures are named *FtClpP*:II and *FtClpP*:III and represent open forms of the protein assembly. The Matthews coefficients (V_m of 2.4 Å³ Da⁻¹ and bulk solvent about 50%) suggested the presence of 7 and 21 subunits in the asymmetric units in *FtClpP*:II and *FtClpP*:III, respectively. *FtClpP*:II was solved using PhaserMR with subunit A from the 3P2L model in a similar fashion as employed for *FtClpP*:I. In the case of *FtClpP*:III, the search model was the refined *FtClpP*:II heptamer. The general refinement procedure that was applied in all three cases started with a round of rigid body refinement using REFMAC²⁰ followed by cycles of restrained refinement. Electron density and difference maps were inspected, and modifications to the models were made with COOT.²¹ Refinement of *FtClpP*:I was terminated when we were satisfied that the reliable parts of the polypeptide had been modeled and that disordered residues were excluded. No attempts were made to include solvent in the refinement given the modest resolution. In the case of the high resolution *FtClpP*:II and III structures the refinement extended to

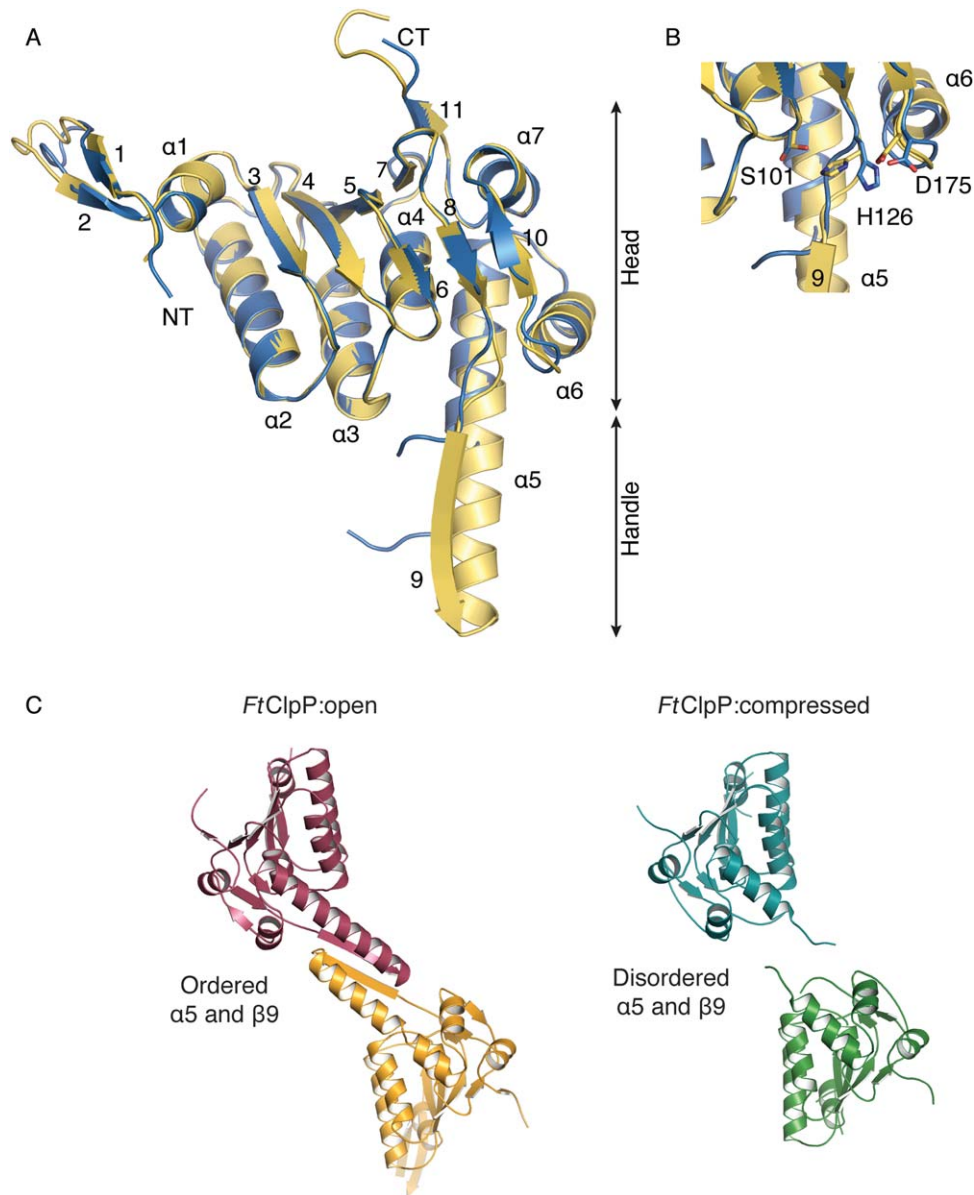
including rotamers, water molecules and ligands. Tight non-crystallographic symmetry restraints were used at the beginning of all refinements then removed gradually. The refinements were ended when inspection of the maps suggested no more changes were justified, and no significant changes in the R_{free} and R_{work} were observed. Crystallographic statistics are presented in Table I.

Analyses of the crystallographic models

The models were analyzed using MolProbity.²² The secondary structure was inspected with DSSP²³ and the search for structural homologues and comparisons were performed with the DALI server.²⁴ Analysis of the quaternary structure and the interface surface areas for the two major conformations observed was carried out with PISA.²⁵ The average *B*-factors of each subunit were determined using the program BAVERAGE.¹⁷ Amino acid sequences were analyzed with Clustal Omega.²⁶ Structural alignments and molecular images were prepared using PyMol (www.pymol.org) and ALINE.²⁷

RESULTS AND DISCUSSION

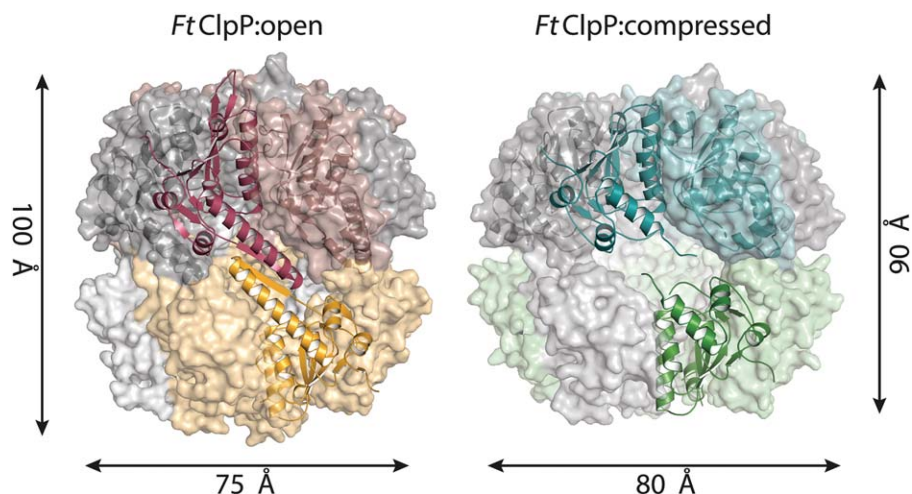
An efficient *E. coli* expression system and a robust purification protocol for *FtClpP* was established that yielded about 10 mg of enzyme per liter of bacterial culture. Analysis by SDS-PAGE and mass spectrometry (data not shown) indicated that the samples were of high purity and of the correct mass. Size exclusion chromatography, in the presence and absence of 10% glycerol

**Figure 1**

A. The ClpP subunit. An overlay ribbon diagram of the open form (yellow) with the closed form (cyan). NT and CT mark the N-terminal and C-terminal residues. The assigned secondary structure of *FtClpP* is represented as cylinders (α -helices) and arrows (β -strands). B. A close up view and showing the Ser-His-Asp catalytic triad residues in the active sites of open (yellow) and compressed forms (cyan). C. A side-by-side comparison of a pair of subunits in the open (left) and compressed (right) conformations. Note the order-disorder difference that involves the handle regions of the subunit at a subunit-subunit interface.

to assist solubility indicated a high level of purity and an oligomer, a single species of approximate mass 290 kDa (Supporting Information Fig. S1A). Subsequently when the sample was applied to a native gel, two species become evident with approximate masses of 160 kDa and 290 kDa (Supporting Information Fig. S1B). This suggests the presence of heptameric and tetradodecameric species in the material used for crystallization experiments. Access to sufficient quantities of material supported a crystallization screen and led to the

identification of reproducible crystallization conditions. Three crystal forms of *FtClpP* have been characterized and these correspond to a closed or compressed conformation (termed *FtClpP*:I) and two structures of an open state (*FtClpP*:II and *FtClpP*:III), determined at 2.8 Å, 1.9 Å and 1.7 Å resolution, respectively. The structures were solved by molecular replacement (Table I). Each subunit (comprising an average of 175 residues in the case of the compressed conformation and 190 in the open forms) presents the characteristic ClpP family fold consisting of

**Figure 2**

Surface and cartoon representation of the *FtClpP*:III (open form; red, yellow and grey, left) and *FtClpP*:I (compressed form; green, turquoise and grey, right) tetradecamers. Heptamers are represented with darker colors. A pair of subunits are represented in cartoon ribbon format.

eleven β -strands and seven α -helices [Fig. 1(A)]. In all structures, the subunits are labeled alphabetically. The β -strands 1 and 2 are not ordered in subunit C of *FtClpP*:I or in subunits G, K, N, U from *FtClpP*:III. In the majority of the subunits of *FtClpP*:II and III, electron density and a chemical environment consistent with the presence of acetate was noted in a pocket close to the catalytic triad. This likely occurs due to the use of an acetate buffer in the crystallization conditions. In subunits E, K, L, and O of *FtClpP*:III the electron density was suggestive of some larger entity but we could not unambiguously identify this so acetate was retained in these sites.

The amino acid sequence of *FtClpP* together with the assigned secondary structure is shown in Supporting Information Figure S2. The polypeptide folds into two subdomains, which are termed the head and handle regions [Fig. 1(A)]. The handle is an extension formed by $\alpha 5$ and $\beta 9$. The active site and catalytic triad, comprising Ser101-His126-Asp175, are located in an area between the head and handle regions [Fig. 1(B)]. This is where the link between $\beta 6$ and $\alpha 4$, the C-terminal segment of $\beta 8$ and N-terminal region of $\beta 10$ are positioned close to each other.

A search for bacterial ClpP sequences (excluding *Francisella* sp.) in the UniProt database (<http://www.uniprot.org/>) revealed 405 reviewed/curated entries. These amino acid sequences were aligned and displayed a range of 40–70% identity with *FtClpP* extending to strict conservation of the three catalytic residues. Strong conservation of amino acid types was noted in elements of secondary structure with the exception of the C-terminal region of $\alpha 5$. These features are evident in the amino acid alignment of *FtClpP* with three orthologues for which structures have been reported (*BsClpP*, *E. coli* ClpP and *M. tuberculosis* ClpP) as

depicted in Supporting Information Figure S2. The sequence identity between *FtClpP* and *BsClpP* is 45% and about 60% with *EcClpP* and *MtClpP*. Extending from this level of sequence conservation we note a high level of structural similarity is maintained between *FtClpP* and *BsClpP* (PDB code: 3MT6), *EcClpP* (PDB code: 3TT7) and *MtClpP* (PDB code: 5DZK) which becomes evident from the least-squares overlay of all C α positions in the open form asymmetric units. The r.m.s.d. values are 0.5 ± 0.1 , 0.7 ± 0.1 , 1.3 ± 0.1 Å, respectively.

FtClpP forms a heptamer by a lateral association of subunits that involves interaction between β -strands (3 and 4) and two loops (between $\beta 5$ – $\beta 6$, and $\beta 7$ – $\beta 8$) with α -helices ($\alpha 2$, $\alpha 3$, $\alpha 4$, and $\alpha 5$) of an adjacent subunit. Two heptamers associate to create the tetradecamer exploiting primarily hydrophobic interactions between two $\alpha 5$ helices that are aligned antiparallel to each other, and with hydrogen bonds between the main chain functional groups on a pairing of $\beta 9$ strands [Fig. 1(C)]. In both *FtClpP*:I and II structures, a heptamer constitutes the asymmetric unit and the physiologically relevant tetradecamer is created across a crystallographic two-fold axis of symmetry. In *FtClpP*:III there are 21 subunits in the asymmetric unit and these are arranged to form a complete tetradecamer and one heptamer, which again is positioned at a two-fold axis that generates another tetradecamer. It is intriguing that the purified sample displayed both heptamer and tetradecamer species when analyzed in a native gel suggesting that distinct conformational states might be present and this may explain why we obtained open and compressed oligomer structures.

There is a high similarity between *FtClpP*:II and III, the open form structures. The overlay of C α positions in the tetradecamers gives an r.m.s.d. of 0.16 Å, so the later

is primarily used for further analysis given that it has been determined to a slightly better resolution. In the case of *FtClpP*:III, $35 \pm 2\%$ of a subunit accessible surface area (ASA) is involved in tetradecamer formation and $43 \pm 1\%$ in *FtClpP*:I. The buried surface areas for the tetradecamer assemblies are approximately 70% and 55%, respectively. These values indicate the presence of stable macromolecular assemblies²⁵ but that in the closed form, *FtClpP*:III buries a significantly greater surface area of the oligomeric assembly.

An overlay of C α positions of individual subunits in *FtClpP*:III (r.m.s.d. average of 0.38 ± 0.04 Å using all C α atoms common to each subunit) indicates a high degree of similarity. The same observation applies to the *FtClpP*:I subunits with an average r.m.s.d. 0.21 ± 0.03 Å. Performing pairwise overlays between subunits of *FtClpP*:I and *FtClpP*:III (r.m.s.d. 0.38 ± 0.04 Å) indicate minor differences, which primarily reside in strands $\beta 1$ and $\beta 2$ and at the C-terminus of the polypeptide. Even when heptamers are compared, there are no major structural differences (r.m.s.d. of 0.78 Å). In Figure 1(A) the overlay of a subunit in each form is presented and the overall close similarity is evident. However, the overlay of the open and closed tetradecamer assemblies results in an r.m.s.d. of 4.29 Å which quantifies the large, and obvious structural difference between open and closed forms of *FtClpP* [Figs. 1(C) and 2, Supporting Information Fig. S3].

As in other ClpP structures, it appears that the conformational change between open and closed forms is associated with an order/disorder transition of the “handle” subdomain [Fig. 1(A,C)]. In particular for *BsClpP* there is a change driven by the distortion/straightening of $\alpha 5$.²⁸ In *FtClpP* we observe something different. The handle in the open form, the $\alpha 5$ - $\beta 9$ segment, is well ordered but in the closed form we did not observe interpretable density with disorder evident. This results in a rearrangement of the subunit positions with respect to each other and in the loss of the active site groove organization [Figs. 1(C) and 2, Supporting Information Fig. S4]). There are changes at the active site, and equatorial pores are created where the reaction products might be released. The consequences of conformational changes involving the handle in *FtClpP* however are similar to those previously described, for example, for *BsClpP*.^{5,28}

ACKNOWLEDGMENTS

The authors thank Helen Atkins, Paul Fyfe and Phil Ireland for advice and support, and the European Synchrotron Radiation Facility for beam time.

REFERENCES

- Brötz-Oesterhelt H, Sass P. Bacterial caseinolytic proteases as novel targets for antibacterial treatment. *Int J Med Microbiol* 2014; 304: 23–30.
- Frees D, Gerth U, Ingmer H. Clp chaperones and proteases are central in stress survival, virulence and antibiotic resistance of *Staphylococcus aureus*. *Int J Med Microbiol* 2014;304:142–149.
- Alexopoulos JA, Guarné A, Ortega J. ClpP: a structurally dynamic protease regulated by AAA+ proteins. *J Struct Biol* 2012; 179:202–210.
- Geiger SR, Böttcher T, Sieber SA, Cramer P. A conformational switch underlies ClpP protease function. *Angew Chem Int Ed* 2011; 50:5749–5752.
- Lee B-G, Kim MK, Song HK. Structural insights into the conformational diversity of ClpP from *Bacillus subtilis*. *Mol Cells* 2011;32: 589–595.
- Yu AY, Houry WA. ClpP: a distinctive family of cylindrical energy-dependent serine proteases. *FEBS Lett* 2007; 581:3749–3757.
- Ingvarsson H, Maté MJ, Högbom M, Portnoi D, Benaroudj N, Alzari PM, Ortiz-Lombardía M, Unge T. Insights into the inter-ring plasticity of caseinolytic proteases from the X-ray structure of *Mycobacterium tuberculosis* ClpP1. *Acta Crystallogr D* 2007;63:249–259.
- Zhang J, Ye F, Lan L, Jiang H, Luo C, Yang CG. Structural switching of *Staphylococcus aureus* Clp protease: a key to understanding protease dynamics. *J Biol Chem* 2011;286:37590–37601.
- Böttcher T, Sieber SA. Beta-lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of *Staphylococcus aureus*. *J Am Chem Soc* 2008;130:14400–14401.
- Böttcher T, and, Sieber SA. Beta-lactones decrease the intracellular virulence of *Listeria monocytogenes* in macrophages. *ChemMedChem* 2009;4:1260–1263.
- Sowole MA, Alexopoulos JA, Cheng Y-Q, Ortega J, Konermann LJ. Activation of ClpP protease by ADEP antibiotics: insights from hydrogen exchange mass spectrometry. *J Mol Biol* 2013; 425:4508–4519.
- Moynie L, Schnell R, McMahon SA, Sandalova T, Boulkerou WA, Schmidberger JW, Alphey M, Cukier C, Duthie F, Kopec J, Liu H, Jacewicz A, Hunter WN, Naismith JH, Schneider G. The AEROPATH project targeting *Pseudomonas aeruginosa*: crystallographic studies for assessment of potential targets in early-stage drug discovery. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2013;69:25–34.
- Hunter WN. Structure-based ligand design and the promise held for antiprotozoan drug discovery. *J Biol Chem* 2009; 284:11749–11753.
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. EXPASY: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 2003; 31:3784–3788.
- Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 2011; 67: 271–281.
- Evans PR. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D* 2011; 67: 282–292.
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. *Acta Crystallogr D* 2011; 67:235–242.
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystal* 2007; 40:658–674.
- Kabsch W. XDS. *Acta Crystallogr D* 2010; 66:125–132.
- Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls NA, Winn MD, Long F, Vagin AA. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* 2011; 67:355–367.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D* 2010; 66:486–501.
- Chen VB, Arendall IIIWB, Headd JJ, Keedy DA, Immormino GJK, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom

- structure validation for macromolecular crystallography. *Acta Crystallogr D* 2010; 66:12–21.
23. Touw WG, Baakman C, Black J, te Beek TAH, Krieger E, Joosten RP, Vriend G. A series of PDB-related databanks for everyday needs. *Nucleic Acids Res* 2015; 43:364–368.
 24. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* 2010; 38:545–549.
 25. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 2007; 372:774–797.
 26. Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins D. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011; 7.
 27. Bond CS, Schüttelkopf AW. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D* 2009; 65:510–512.
 28. Gersch M, List A, Groll M, Sieber SA. Insights into structural network responsible for oligomerization and activity of bacterial virulence regulator caseinolytic protease P (ClpP) protein. *J Biol Chem* 2012; 287:9484–9494.
 29. Cruickshank DW. Remarks about protein structure precision. *Acta Crystallogr D Biol Crystallogr* 1999; 55:583–601.