Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary X-ray studies of ClpX from *Helicobacter pylori*

Dong Young Kim, Chun Ai Wu, Dong Ryoung Kim, Sung Chul Ha, Young-Hyun Han and Kyeong Kyu Kim*

Department of Molecular Cell Biology, Center for Molecular Medicine, SBRI, Sungkyunkwan University School of Medicine, Suwon 440-746, South Korea

Correspondence e-mail: kkim@med.skku.ac.kr

Received 18 January 2003
Accepted 30 June 2003

ClpX, a member of the HSP (heat-shock protein) 100 family, functions as a molecular chaperone and is a regulatory subunit of the ClpXP protease. To understand the chaperone and regulatory mechanisms of ClpX, *Helicobacter pylori* ClpX has been overexpressed in *Escherichia coli* and crystallized at 295 K using $(NH_4)_2HPO_4$ as precipitant. X-ray diffraction data have been collected to 2.6 Å resolution using a synchrotron-radiation source. The crystals belong to the hexagonal space group $P6_5$ or $P6_1$, with unit-cell parameters $a = b = 78.52$ (04), $c = 131.51$ (09) Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. The crystallographic asymmetric unit contains one molecule of ClpX, with a corresponding V_M of 2.78 Å^3 Da^{-1} and a solvent content of 55.8%.

1. Introduction

The protein remodelling and degradation carried out by the molecular-chaperone and protease families of enzymes have important physiological functions in living organisms. Three families of ATP-dependent proteases (protease La, the two-component ClpXP or ClpAP protease and FtsH) have been identified in bacteria. They are responsible for most intracellular protein degradation in bacteria, contributing to protein-quality control and modulating the intracellular concentration of important global regulatory proteins (Gottesman & Maurizi, 1992; Schirmer *et al.*, 1996; Suzuki *et al.*, 1997; Kang *et al.*, 2002).

Bacterial ClpX, which is the regulatory component of the ClpXP complex, belongs to the Clp/HSP 100 ATPase family, as does the bacterial ClpA in the ClpAP complex (Gottesman *et al.*, 1993; Wojtkowiak *et al.*, 1993). ClpX shows ATPase activity and exists as a hexameric complex (Grimaud *et al.*, 1998; Singh *et al.*, 2001). Each subunit is composed of three domains: a Cys -cluster domain and an ATPase domain at the amino-terminus and an SSD (sensor and substrate discrimination) domain at the carboxyl-terminus. The ATPase domain of ClpX contains a tripeptide [LIV]-G-[FL] motif that is responsible for ClpP recognition and is unique to ClpX and ClpA (Kim *et al.*, 2001). The SSD domain is involved in substrate-binding specificity (Smith *et al.*, 1999). The amino-terminus Cys-cluster domain is known to contribute to substrate binding or to perform a gating function affecting substrate access to other binding sites. However, it is considered that the Cys cluster is not necessary for the proteolytic activity of ClpXP (Singh *et al.*, 2001).

Unlike ClpX, ClpA has two ATPase domains. A recent crystal structure of *Escherichia coli* ClpA reveals that the two ATPase domains play non-equivalent roles (Guo *et al.*, 2002). The first ATPase domain, D1, is involved in hexameric assembly and the second ATPase domain, D2, possesses a stronger ATPase activity. In this respect, the crystal structure of ClpX is interesting in that its single ATPase domain has both ClpA D1 and D2 activities. Furthermore, the sequence identity between ClpX and ClpA D2 is only 19.3%, suggesting that ClpX should have a different structural architecture from that of ClpA, although it is expected that ClpX will have a similar fold to ClpA. In addition, in the reported crystal structure of *E. coli* ClpA, functionally important regions including the tripeptide motif (ClpP-binding loop) were not visible (Guo *et al.*, 2002). Although the ClpP-binding loop is proposed to be involved in ClpP binding by hydrophobic interaction (Kim *et al.*, 2001), it is still not clear how this loop makes contacts with ClpP and how the ClpX hexamer makes an asymmetric assembly with the ClpP heptamer as its structural features have never been reported. Therefore, the study of crystal structure of ClpX is required in order to understand the roles of the ATPase domain and the tripeptide motif. As the first step in this approach, we describe here the overexpression, purification, crystallization and

preliminary crystallographic data of ClpX from *H. pylori* lacking the N-terminal Cys-cluster domain.

2. Materials and methods

2.1. Protein expression and purification

H. pylori ClpX (gene number HP1374) lacking the amino-terminal Cys-cluster domain (truncated ClpX, corresponding to residues 71–446) was expressed and purified. The N-terminal Cys-cluster domain has been removed to facilitate the crystallization of the remaining ATPase and SSD domains of ClpX. The gene encoding truncated ClpX was amplified from genomic DNA by the polymerase chain reaction (PCR) using forward (5′-CTAGCGCTCATATGGAAGAGTTTTTACTCT-3′) and reverse (5′-CTAGGAATTCTCAAGGAAGAATTTTAGAATGC-3′) oligonucleotide primers. The PCR products were digested by *Nde*I and *Eco*RI and inserted into *Nde*I/*Eco*RI-digested expression vector pET-22b (Novagen). The plasmid encoding truncated ClpX was transformed into *E. coli* strain BL21(DE3). The cells were grown in Luria–Bertani (LB) medium containing 50 μg ml^{-1} ampicillin and were induced with 1 m*M* isopropyl β-D-thiogalactopyranoside (IPTG) at 288 K. After overnight induction, the cells were harvested by centrifugation at 4200*g* (Sorvall GS3 rotor) for 10 min at 277 K. The cell pellet was resuspended in buffer *A* (20 m*M* Tris–HCl pH 7.0) and disrupted by sonication. The homogenate was clarified by centrifugation at 27 000*g* (Beckman JA 25.50 rotor) for 15 min at 277 K and further clarified by ultracentrifugation at 120 000*g* (40 000 rev min^{-1}; Beckman Type 70 Ti rotor) for 40 min at 277 K. The supernatant was loaded onto a Q-Sepharose anion-exchange column (Amersham-Pharmacia) which was equilibrated with buffer *B* (20 m*M* Tris–HCl pH 7.0 and 0.1 *M* NaCl). The protein was eluted with a linear gradient of 0.1–0.5 *M* NaCl in buffer *B*. The fractions containing truncated ClpX were pooled and loaded onto a HiLoad 16/60 Superdex 200 prep-grade gel-filtration column (Amersham-Pharmacia) previously equilibrated with buffer *A*. The protein was concentrated to 10 mg ml^{-1} for crystallization purposes. The purity of the truncated ClpX was checked by SDS–PAGE.

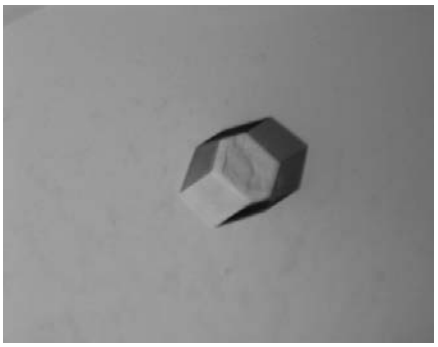

Figure 1
A crystal of recombinant ClpX. Crystal dimensions are about 0.3 × 0.2 × 0.2 mm.

2.2. Crystallization and X-ray diffraction experiment

Crystallization was performed by the hanging-drop vapour-diffusion method at 295 K using 24-well tissue-culture plates. Each hanging drop, prepared by mixing 1 μl reservoir solution and 1 μl protein solution (10 mg ml^{-1}), was equilibrated against 0.5 ml reservoir solution. Initial crystallization conditions were tested using Wizard Crystal Screen kits (Emerald Structures). Hexagonal crystals were grown using $(NH_4)_2HPO_4$ as precipitant (Fig. 1). Crystals suitable for diffraction experiments were obtained from protein solution containing 100 m*M* imidazole pH 8.0, 1.0 *M* $(NH_4)_2HPO_4$ and 0.2 *M* NaCl. When crystals were transferred directly into 26% glycerol solution, they cracked. Therefore, crystals were transferred through a series of cryoprotectant solutions of increasing glycerol concentration (2–26 %) before being flash-frozen in a cold nitrogen stream. X-ray diffraction data were collected at 100 K with a MacScience 2030b area detector at beamline 6B of the Pohang Accelerator Laboratory, South Korea. The wavelength of the synchrotron radiation was 1.127 Å. A total of 46 frames of 1.3° oscillation were measured with the crystal-to-detector distance set to 300 mm. Data were processed and integrated using *DENZO* and scaled using *SCALEPACK* from the *HKL* program suite (Otwinowski & Minor, 1997).

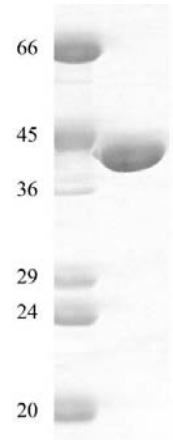

Figure 2
SDS–PAGE analysis of purified truncated ClpX. Left lane, molecular-weight markers in kDa; right lane, purified truncated ClpX.

Table 1
X-ray data-collection and processing statistics.

Values in square brackets refer to the highest resolution shell.

Space group	$P6_1$ or $P6_5$
Unit-cell parameters (Å, °)	$a = b = 78.52$ (04), $c = 131.51$ (09), $\alpha = \beta = 90$, $\gamma = 120$
Resolution range (Å)	30–2.6 [2.69–2.60]
No. of unique reflections	13273 [1369]
Data completeness (%)	93.7 [96.3]
R_{sym}† (%)	5.9 [32.6]
$I/\sigma(I)$	22.5 [2.6]

† $R_{sym} = \sum_{hkl} |I - \langle I \rangle| / \sum_{hkl} I$.

3. Results and discussion

The protein comprising amino-acid residues 71–446 of *H. pylori* ClpX was overexpressed and purified with an approximate yield of 10 mg of homogenous protein from a 1 l culture. The protein is more than 95% pure as judged from SDS–PAGE analysis (Fig. 2). The molecular weight of truncated ClpX is about 42 000 Da from SDS–PAGE, which is in agreement with the calculated molecular weight of 42 169 Da. *H. pylori* truncated ClpX exists as a hexamer in buffer *A* as determined by gel-filtration and dynamic light-scattering experiments (data not shown). In gel filtration, truncated ClpX eluted between ferritin (440 kDa) and catalase (233 kDa), suggesting it exists as a hexamer; the calculated molecular weight of the truncated ClpX hexamer is 253.0 kDa. The crystals grew to dimensions of 0.3 × 0.2 × 0.2 mm within 1 d (Fig. 1). The diffraction data from a cryogenic crystal were collected to 2.6 Å resolution at 100 K using a synchrotron-radiation source. A total of 40 304 measured reflections were merged to 13 273 unique reflections with an R_{merge} (on intensity) of 5.9%. The merged data set is 93.7% complete to 2.6 Å resolution. The crystals belong to the hexagonal space group $P6_5$ or $P6_1$, with unit-cell parameters a = 78.515, c = 131.514 Å. Assuming one protein per crystallographic asymmetric unit and based on a molecular weight of 42 169 Da, the value of the crystal-packing parameter V_M is 2.78 Å^3 Da^{-1}, with a corresponding solvent content of 55.8%, which falls within the range commonly observed for protein crystals (Matthews, 1968). Data statistics for the synchrotron data collection and processing are summarized in Table 1. Using these data, we are in the process of solving the structure by multi-

wavelength anomalous dispersion (MAD) methods.

We thank Dr H. S. Lee and beamline scientists for assistance during data collection at beamline 6B at Pohang Accelerator Laboratory, South Korea. This work was supported by grant No. FG-3-5-01 of the 21C Frontier Functional Human Genome Project from the Ministry of Science and Technology of Korea. Experiments at PLS were supported in part by MOST and POSCO.

References

Guo, F., Maurizi, M. R., Esser, L. & Xia, D. (2002). *J. Biol. Chem.* **277**, 46743–46752.
Gottesman, S., Clark, W. P., de Crecy-Lagard, V. & Maurizi, M. R. (1993). *J. Biol. Chem.* **268**, 22618–22626.
Gottesman, S. & Maurizi, M. R. (1992). *Microbiol. Rev.* **56**, 592–621.
Grimaud, R., Kessel, M., Beuron, F., Steven, A. C. & Maurizi, M. R. (1998). *J. Biol. Chem.* **273**, 12476–12481.
Kang, S. G., Ortega, J., Singh, S. K., Wang, N., Huang, N., Steven, A. C. & Maurizi, M. R. (2002). *J. Biol. Chem.* **277**, 21095–21102.
Kim, Y. I., Levchenko, I., Fraczkowska, K., Woodruff, R. V., Sauer, R. T. & Baker, T. A. (2001). *Nature Struct. Biol.* **8**, 230–233.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Schirmer, E. C., Glover, J. R., Singer, M. A. & Lindquist, S. (1996). *Trends Biochem. Sci.* **21**, 289–296.
Singh, S. K., Rozycki, J., Ortega, J., Ishikawa, T., Lo, J., Steven, A. C. & Maurizi, M. R. (2001). *J. Biol. Chem.* **276**, 29420–29429.
Smith, C. K., Baker, T. A. & Sauer, R. T. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 6678–6682.
Suzuki, C. K., Rep, M., van Dijl, J. M., Suda, K., Grivell, L. A. & Schatz, G. (1997). *Trends Biochem. Sci.* **22**, 118–223.
Wojtkowiak, D., Georgopoulos, C. & Zylicz, M. (1993). *J. Biol. Chem.* **268**, 22609–22617.