Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, expression, purification and preliminary X-ray crystallographic studies of *Escherichia coli* Hsp100 ClpB N-terminal domain

Jingzhi Li and Bingdong Sha*

Department of Cell Biology, Center for Biophysical Sciences and Engineering, University of Alabama at Birmingham, Birmingham, AL 35294-0005, USA

Correspondence e-mail: sha@cmc.uab.edu

Received 16 July 2001
Accepted 15 October 2001

Escherichia coli Hsp100 ClpB plays critical roles in multi-chaperone systems in cell physiology. After the ATPase activity is stimulated by protein or peptide binding, ClpB disaggregates denatured polypeptides by employing ATP hydrolysis and allows other molecular chaperones such as Hsp70 DnaK and Hsp40 DnaJ to more efficiently refold the non-native polypeptides. The mechanisms by which the ClpB acts as a molecular chaperone to disaggregate non-native polypeptides are unknown. The N-terminal domain of ClpB has been proposed to interact with non-native polypeptides. To investigate whether the N-terminal domain participates in polypeptide recognition and binding or modulates the activity of ClpB, the ClpB N-terminal domain has been cloned, purified and crystallized. The ClpB N-terminal domain crystals diffract to 1.95 Å using a synchrotron X-ray source and belong to the space group $P1$, with unit-cell parameters $a = 50.2$, $b = 52.6$, $c = 56.8$ Å, $\alpha = 90.5$, $\beta = 111.8$, $\gamma = 107.1^\circ$. Structure determination by the multiple anomalous dispersion (MAD) method is under way.

1. Introduction

The heat-shock protein 100 (HSP100) family covers a wide range of proteins involving various cellular functions such as degradation of misfolded proteins, refolding of aggregated polypeptides, transcription regulation and tolerance of high temperature by cells (reviewed in Schirmer *et al.*, 1996; Wawrzynow *et al.*, 1996). The Hsp100/Clp members have been divided into two subfamilies. The class I members have two ATPase domains, while the class II family contain only one ATPase domain (Schirmer *et al.*, 1996). *E. coli* ClpB belongs to the class I Hsp100 subfamily. Recently, members of this subfamily were identified as functioning as molecular chaperones to disaggregate denatured polypeptides (Glover & Lindquist, 1998; Goloubinoff *et al.*, 1999). ClpB can be induced *in vivo* by heat shock or other stress factors and plays essential roles at elevated temperature in *E. coli* (Schirmer *et al.*, 1996). It has been reported that ClpB can cooperate with the Hsp70 molecular chaperone system (including DnaK, DnaJ and GrpE) in *E. coli* to form multi-chaperone systems to more efficiently solubilize protein aggregates and refold them efficiently into active states (Zolkiewski, 1999; Goloubinoff *et al.*, 1999). ClpB binds denatured protein aggregates and unfolds them while catalyzing ATP hydrolysis. This may expose the hydrophobic side chains and allow the subsequent DnaK binding and refolding of the polypeptides (Goloubinoff *et al.*, 1999).

It remains unclear which domain of ClpB is responsible for interacting with denatured polypeptides. It was hypothesized that the C-terminal fragment in Hsp100 proteins forms a 'sensor and substrate discrimination' (SSD) domain to bind non-native polypeptide, a PDZ-like domain that plays roles in recognizing and binding protein substrates (Levchenko *et al.*, 1997; Smith *et al.*, 1999). However, recent mutagenesis data show that the C-terminal domains of Hsp100 proteins may be involved in protein oligomerization rather than peptide binding (Barnett *et al.*, 2000). This proposition is also supported by the crystal structure of Hsp100 ClpY (Bochtler *et al.*, 2000; Sousa *et al.*, 2000).

The N-terminal domain of ClpB has also been proposed to function as the peptide-binding domain (Park *et al.*, 1993; Barnett *et al.*, 2000). ClpB has an N-terminal domain that contains about 150 amino-acid residues. This domain contains two repeat R regions which are about 70 amino-acid residues long (Lo *et al.*, 2001). ClpB without the N-terminal domain can self-associate and has intrinsic ATPase activity. However, the ATPase activity of the deletion mutant shows a much weaker response to the stimuli of the denatured proteins. The deletion mutant does not have

any refolding ability for denatured luciferase (Barnett *et al.*, 2000). The data suggests that the N-terminal region of ClpB is involved in interactions with non-native peptides (Park *et al.*, 1993; Barnett *et al.*, 2000).

However, other experiments gave controversial results. *In vivo* experiments showed that the N-terminal truncated form of ClpB had a similar capacity to confer thermotolerance to that of the full-length ClpB protein (Clarke & Eriksson, 2000). This suggests that the N-terminal domain of ClpB might not be essential for the thermotolerance activity in ClpB. In yeast, two homologues of ClpB are expressed: a 104 kDa protein (Hsp104) existing mainly in the cytosol and a 78 kDa protein (Hsp78) mainly existing in the mitochondria. The Hsp78 also has no N-terminal domain of ClpB (Leonhardt *et al.*, 1993); however, Hsp78 can prevent protein aggregation and disaggregate existing protein aggregates (Moczko *et al.*, 1995; Schmitt *et al.*, 1996).

Little work has been performed on the structural studies of *E. coli* Hsp100 ClpB. The crystal structure of the ClpB N-terminal domain may reveal the mechanisms by which Hsp100 proteins recognize and unfold aggregated polypeptide.

2. Experimental

2.1. Cloning, expression and purification of *E. coli* ClpB N-terminal domain

The ClpB N-terminal domain (ClpBN) has been determined to be within amino-acid residues 1–148 of the full-length ClpB by limited proteolysis (unpublished work). The *ClpBN* gene was then cloned into vector pet15b (Novagen). The PCR products were digested using restriction endonucleases *Nde*I and *Bam*HI by following the instructions (New England BioLabs). The inserts were then ligated into the digested pet15b vector using T4 ligase. The nucleotide sequence was then confirmed by DNA sequencing. The plasmid encoding the ClpBN was transformed into *E. coli* strain BL21(DE3) for protein expression. The calculated molecular weight is 15 765 Da for the ClpBN monomer.

10 ml of LB medium with 50 µg ml^{-1} ampicillin was inoculated using transformed *E. coli* stocks. The cells were allowed to grow at 310 K in a shaker for 12 h. 10 ml of LB medium was then used to inoculate 1 l LB medium with 50 µg ml^{-1} ampicillin. When the OD_{600} of the medium reached 0.6, 0.5 ml of 1 *M* IPTG was added to 1 l of medium to induce protein expression. The cells were harvested 3 h after induction. The *E. coli* cells from 1 l of medium were pelleted down by centrifugation and resuspended in 100 ml 100 m*M* Tris buffer pH 7.9, 150 m*M* NaCl. The cells were lysed by sonication at 277 K. The debris and insoluble materials were pelleted at 15 000 rev min^{-1} using a Beckman JA20 rotor. Since the ClpBN was histidine tagged, it could relatively easily be purified by a metal-chelating column. The supernatant was pumped through the Ni-charged column with about 10 ml of resins. The column was thoroughly washed with 50 m*M* Tris buffer pH 7.9, 0.5 *M* NaCl and 50 m*M* imidazole to remove contaminating proteins. The bound protein was then eluted by 50 m*M* Tris buffer pH 7.9, 0.5 *M* NaCl and 200 m*M* imidazole. The eluted protein was dialyzed against 2 l 10 m*M* HEPES buffer pH 7.2, 50 m*M* NaCl. The typical yield of soluble ClpBN (∼95% pure from SDS–PAGE analysis) from 1 l culture is ∼30 mg. The N-terminal histidine tag of ClpBN was then digested by thrombin treatment. One unit of thrombin (Sigma) was utilized for 1 mg ClpBN protein. Digestion took place for 12 h at room temperature and was stopped by the addition of 0.2 m*M* PMSF. The protein was further purified on a Superdex 200 gel-filtration column (Pharmacia) mounted on an AKTA HPLC system (Pharmacia) to remove thrombin and digested peptides. The apparent molecular weight of the ClpBN is shown to be ∼15 kDa based on the elution time of the protein peak, indicating that the ClpBN is a monomer in solution.

2.2. Crystallization, data collection and processing of the ClpBN

The ClpBN protein was concentrated to 30 mg ml^{-1} in 10 m*M* HEPES buffer pH 7.5, 50 m*M* NaCl and subjected to crystallization trials. Small needle-shaped crystals (0.2 × 0.01 × 0.01 mm) were obtained by the hanging-drop vapor-diffusion method using Linbro plates at 277 K. The well solution consisted of 1 ml 100 m*M* Tris buffer pH 7.5, 12% PEG 20K, 0.2% *n*-octyl-β-D-glucopyranoside. The crystals grew to full size (0.5 × 0.05 × 0.05 mm) overnight after microseeding. The hanging drops consisted of 2 µl protein solution plus 2 µl well solution. To examine whether the crystals are composed of ClpBN, several crystals were collected, washed extensively and then dissolved in water. The sample was then subjected to SDS–PAGE analysis. The resulting gel showed that the crystals contained only ClpBN. To solve the crystal structure by the MAD method, selenomethionyl ClpBN was produced following earlier protocols (Sha *et al.*, 2000). The plasmid encoding ClpBN was transformed into *E. coli* strain B834(DE3), a methionine auxotroph. The cells were grown in medium which consisted of 2×M9 plus glucose minimal medium supplemented with 19 amino acids (not including methionine) at 40 µg ml^{-1}, seleno-L-methionine at 40 µg ml^{-1} and vitamins (thiamine, riboflavin and pyridoxine) at 1 µg ml^{-1}. The preparation of ClpBN seleno-L-methionine crystals was similar to that used to grow the native crystals.

Diffraction data for the seleno-L-methionine ClpBN crystal were collected on the BioCars beamline BMD14 at APS. The crystal was flash-frozen at 100 K in a nitrogen-gas stream in a cryoprotectant consisting of 100 m*M* Tris buffer pH 7.5, 15% PEG 20K, 0.2% *n*-octyl-β-D-glucopyranoside and 15% glycerol. The seleno-L-methionyl ClpBN crystals diffracted to 1.95 Å. The data were collected using a Quantum 4 CCD detector (ADSC) at a crystal-to-detector distance of 160 mm. 90 images covering an oscillation range of 180° were collected and processed using *DENZO* and *SCALEPACK* (Minor, 1993; Otwinowski, 1993). The results show a mosaicity value of 0.356 for the crystal. The data set is 97.7% (93.9%) complete, with an $I/\sigma(I)$ of 24.4 (11.2). The R_{sym} for the data set is 5.0% (12.8%). Values in parentheses are for the outer resolution shell. Crystal structure determination of ClpBN using the MAD method is under way.

We wish to thank Dr C. Squires for providing us with the cDNA of *E. coli* ClpB. We are grateful to Drs Robert Henning, Keith Brister, Gary Navrotski and Zhong Ren of APS BioCars BMD-14 and BMC-14 for their help in data collection. This work was supported by NIH (R01 DK56203), HHMI, American Heart Association and NASA.

References

Barnett, M. E., Zolkiewska, A. & Zolkiewski, M. (2000). *J. Biol. Chem.* **275**, 37565–37571.
Bochtler, M., Hartmann, C., Song, H. K., Bourenkov, G. P., Bartunik, H, D. & Huber, R. (2000). *Nature (London)*, **403**, 800–805.
Clarke, A. K. & Eriksson, M. J. (2000). *J. Bacteriol.* **182**, 7092–7096.
Glover, J. R. & Lindquist, S. (1998). *Cell*, **94**, 73–82.
Goloubinoff, P., Mogk, A., Zvi, A. P. B., Tomoyasu, T. & Bukau, B. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 13732–13737.
Leonhardt, S. A., Fearson, K., Danese, P. N. & Mason, T. L. (1993). *Mol. Cell. Biol.* **13**, 6304–6013.
Levchenko, I., Smith, C. K., Walsh, N. P., Sauer, R. T. & Baker, T. A. (1997). *Cell*, **91**, 939–947.

Lo, H., Baker, T. A. & Sauer, R. T. (2001). *Protein Sci.* **10**, 551–559.
Minor, W. (1993). *XDISPLAYF Program*. Purdue University, West Lafayette, USA.
Moczko, M., Schonfisch, B., Voos, W., Pfanner, N. & Rassow, J. (1995). *J. Mol. Biol.* **254**, 538–543.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Park, S. K., Kim, K. I., Woo, K. M., Seol, J. H., Tanaka, K., Ichihara, A., Ha, D. B. & Chung, C. H. (1993). *J. Biol. Chem.* **268**, 20170–20174.
Schirmer, E. C., Glover, J. R., Singer, M. A. & Lindquist, S. (1996). *Trends Biochem. Sci.* **21**, 289–296.
Schmitt, M., Neupert, W. & Langer, T. (1996). *J. Cell. Biol.* **134**, 1375–1386.
Sha, B. D., Lee, S. & Cyr, D. M. (2000). *Structure*, **8**, 799–807.
Smith, C. K., Baker, T. A. & Sauer, R.T. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 6678–6682.
Sousa, M. C., Trame, C. B., Tsuruta, H., Wilbanks, S. M., Reddy, V. S. & McKay, D. B. (2000). *Cell*, **103**, 633–643.
Wawrzynow, A., Banecki, B. & Zylicz, M. (1996). *Mol. Microbiol.* **21**, 895–899.
Zolkiewski, M. (1999). *J. Biol. Chem.* **274**, 28083–28086.