Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, expression, purification and preliminary X-ray crystallographic studies of *Escherichia coli* Hsp100 ClpB nucleotide-binding domain 1 (NBD1)

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 20 March 2001
Accepted 1 May 2001

Escherichia coli Hsp100 ClpB plays critical roles in multi-chaperone systems in cell physiology. After being activated 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 refold the non-native polypeptides. ClpB contains two nucleotide-binding domains with Walker A and B motifs within their primary sequences. Therefore, ClpB can be classified as a member of the large ATPase family known as ATPases associated with various cellular activities (AAAs). The mechanisms by which the ClpB acts as a molecular chaperone to disaggregate denatured polypeptides are unknown. To investigate how the nucleotide-binding domain participates in ClpB chaperone activity, we have cloned and crystallized ClpB nucleotide-binding domain 1 (NBD1). The ClpB NBD1 crystals diffract to 1.80 Å using a synchrotron X-ray source and belong to the space group $P2_12_12_1$, with unit-cell parameters $a = 38.41$, $b = 65.48$, $c = 79.13$ Å. Structure determination by the MAD method is under way.

1. Introduction

The heat-shock protein 100 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 cells to high temperature (reviewed in Schirmer *et al.*, 1996; Wawrzynow *et al.*, 1996). *E. coli* ClpA was the first identified member of the Hsp100 family (Goldberg, 1992; Gottesman & Maurizi, 1992). ClpA can directly bind the free end of the double-ring serine protease ClpP to promote protein degradation (Hoskins *et al.*, 1998). ClpA has been hypothesized to interact with recognition-peptide sequences and to then unfold the stable protein for subsequent protein digestion by the protease ClpP (Gottesman *et al.*, 1997; Larsen & Finley, 1997; Weber-Ban *et al.*, 1999). ClpX and ClpY, like ClpA, associate with proteases and play roles in protein unfolding and degradation (Schirmer *et al.*, 1996). In contrast, other Hsp100 members such as *E. coli* ClpB and *Saccharomyces cerevisiae* Hsp104 do not associate with proteases. Recently, they 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 Hsp100 ClpB can cooperate with the Hsp70 molecular-chaperone system (including DnaK, DnaJ and GrpE) in *E. coli* to form multi-chaperone systems to solublize protein aggregates and refold them into active states efficiently (Zolkiewski, 1999; Goloubinoff *et al.*, 1999). ClpB binds denatured protein aggregates and unfolds them by ATP hydrolysis. This exposes the hydrophobic side chains and allows subsequent DnaK binding and refolding of the polypeptides (Goloubinoff *et al.*, 1999). The molecular mechanism by which ClpB recognizes, binds and disaggregates denatured polypeptides is currently unknown.

All Hsp100 proteins contain either one or two nucleotide-binding domains as predicted by sequence analysis (Schirmer *et al.*, 1996; Patel & Latterich, 1998). They have moderate basal ATPase activities. The intrinsic ATPase activities can be stimulated by specific protein or peptide binding (Wawrzynow *et al.*, 1996). Hsp100 proteins belong to a large family of Mg-dependent ATPases known as ATPases associated with various cellular activities (AAAs; reviewed in Patel & Latterich, 1998). The hallmark of the AAA protein family is an AAA domain (about 200–250 amino-acid residues) containing Walker A and B motifs (Patel & Latterich, 1998). *E. coli* Hsp100 ClpB contains two AAA domains within its primary sequence, nucleotide-binding domain 1 (NBD1) and nucleotide-binding domain 2 (NBD2) (Schirmer *et al.*, 1996; Barnett *et al.*, 2000). NBD1 and NBD2 presumably contri-

bute to the Hsp100 protein ATPase activities at different levels as suggested by mutagenesis in the Walker A and B regions (Schirmer *et al.*, 1996). In ClpB, NBD1 shares about 24% primary sequence identity with NBD2 when the two domains are aligned. It remains unclear why ClpB needs two AAA domains to achieve ATPase activity for its molecular-chaperone ability.

ClpB forms a ring-shaped structure in solution in the presence of ATP by self-oligomerization, while the ring-shaped structure of ClpB dissociates in the absence of ATP as observed by gel-filtration, chemical crosslinking and cryo-EM experiments (Barnett *et al.*, 2000; Kim *et al.*, 2000). The ClpB ring-shaped structure consists of a homohexamer or homoheptamer (Barnett *et al.*, 2000; Kim *et al.*, 2000). It was hypothesized that the C-terminal fragments in Hsp100 proteins form 'sensor and substrate discrimination' (SSD) domains that play roles in recognizing and binding protein substrates (Smith *et al.*, 1999). However, recent mutagenesis data show that the C-terminal domains of Hsp100 proteins may be involved in protein oligomerization (Barnett *et al.*, 2000). This proposition is also supported by the crystal structure of Hsp100 ClpY (Bochtler *et al.*, 2000; Sousa *et al.*, 2000).

Crystal structures of Hsp100 ClpYs (also called HslU) from *E. coli* and *Haemophilus influenzae* complexed with their protease partner HslV have been determined (Bochtler *et al.*, 2000; Sousa *et al.*, 2000). ClpY forms a homohexamer in the crystal structure. The N-terminal ATPase domain of ClpY resembles the similar folding topology of other AAA proteins and the C-terminal domain makes extensive contact with the N-terminal ATPase domain from the adjacent ClpY monomer to constitute the hexamer. An insertion domain within the N-terminal domain of ClpB protrudes towards the outside of the molecule (Bochtler *et al.*, 2000; Sousa *et al.*, 2000). No structural information is available for the other type of Hsp100 proteins, such as ClpB, which functions as a molecular chaperone.

2. Experiments

2.1. Cloning, expression and purification of *E. coli* ClpB nucleotide-binding domain 1 (NBD1)

The ClpB nucleotide-binding domain 1 (NBD1) has been determined to be within amino-acid residues 159–353 of the full-length ClpB (Li & Sha, 2001). The ClpB NBD1 was cloned into the vector pet15b (Novagen). The PCR products were digested by the restriction endonucleases *Nde*I and *Bam*HI following the instructions (New England BioLabs). The inserts were then ligated into the digested pet15b vector by T4 ligase. The nucleotide sequence was then confirmed by DNA sequencing. The plasmid encoding the ClpB NBD1 was transformed into *E. coli* strain BL21(DE3) for protein expression. The calculated molecular weight is 21 233 Da for the ClpB NBD1 monomer.

10 ml LB medium containing 50 µg ml^{-1} ampicillin was inoculated using the 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 containing 50 µg ml^{-1} ampicillin. 0.5 ml of 1 *M* IPTG was added to 1 l medium to induce protein expression when the OD_{600} of the medium reached 0.6. The cells were harvested 3 h after induction. The *E. coli* cells from 1 l medium were pelleted down by centrifugation and resuspended in 100 ml 100 m*M* Tris buffer pH 7.9, 150 m*M* NaCl. The cell were lysed by sonication at 277 K for 30 s. The debris and insoluble materials were pelleted down at 15 000 rev min^{-1} using a Beckman JA20 rotor. Since the ClpB NBD1 was histidine-tagged, it could be relatively easily purified with a metal-chelating column. The supernatant was pumped through an Ni-charged column with about 10 ml resin. 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 the contaminating proteins. The bound protein was then eluted with 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 of 10 m*M* HEPES buffer pH 7.2, 50 m*M* NaCl. The typical yield of soluble ClpB NBD1 (~95% pure from SDS–PAGE analysis) from 1 l culture is ~20 mg. The N-terminal histidine tag of NBD1 was then digested by thrombin treatment. One unit of thrombin (Sigma) was utilized per milligram of ClpB NBD1 protein. The digestion took place for 12 h at room temperature and was stopped by the addition of PMSF to a final concentration of 0.2 m*M*. 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 ClpB NBD1 is shown to be ~20 kDa based on the elution time of the protein peak, indicating that ClpB NBD1 is a monomer in solution.

Table 1
Statistics of the native and Se-Met data sets for ClpB NBD1.

Values in parentheses are for the highest resolution shell.

	Native crystals	Se-Met crystals
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)		
a	38.41	38.37
b	65.48	65.18
c	79.13	79.05
Resolution (Å)	1.8	2.0
No. of observations	83265	36866
Unique reflections	19159	13559
Mosaicity	0.350	0.331
Data completeness (%)	99.4 (96.3)	96.3 (85.1)
Average $I/\sigma(I)$	32.6 (11.9)	26.7 (9.9)
R_{sym} (%)	3.9 (14.2)	5.9 (13.0)

2.2. Crystallization, data collection and processing of the ClpB NBD1

The ClpB NBD1 protein was concentrated to 20 mg ml^{-1} in 10 m*M* HEPES buffer pH 7.5, 200 m*M* KCl, 10 m*M* $MgCl_2$ and subjected to crystallization trials. Initial crystallization screening with Hampton Research crystallization kits I and II did not produce protein crystals. We then carried out a larger screening by trying further conditions manually. Large needle-shaped crystals (0.5 × 0.05 × 0.05 mm) were finally obtained by the hanging-drop vapor-diffusion method using Linbro plates at room temperature. The crystals grew to full size overnight. The well solution consisted of 1 ml 100 m*M* Tris buffer pH 7.5, 2.25 *M* ammonium sulfate, 10 m*M* $MgCl_2$ and 5 m*M* adenylylimidodiphosphate (AMPPNP). The hanging drops consisted of 2 µl protein solution plus 2 µl well solution. The crystals of ClpB NBD1 can be stored in 100 m*M* Tris buffer pH 7.5, 2.5 *M* ammonium sulfate, 10 m*M* $MgCl_2$ and 5 m*M* AMPPNP for several days without any damage. To examine whether the crystals are composed of ClpB NBD1, several crystals were collected, washed several times and then dissolved in water. The sample was then subjected to SDS–PAGE analysis. The results showed that the crystals contained only ClpB NBD1.

The NBD1 crystals were then taken to BioCars beamline BMC14 at the APS for data collection. The crystals are sensitive to X-ray radiation and had to be frozen in order to collect a full data set from a single crystal. The crystal was transferred into cryoprotectant consisting of 100 m*M* Tris buffer pH 7.5, 2.5 *M* ammonium sulfate, 10 m*M* $MgCl_2$, 5 m*M* AMPPNP and 20% glycerol for several minutes before being flash-frozen at 100 K in a nitrogen-gas stream. The ClpB NBD1 crystal diffracted to

1.80 Å using 1.00 Å X-ray radiation. The data were collected using a Quantum 4 CCD detector (ADSC). During data collection, the crystal-to-detector distance was kept at 150 mm. The oscillation angle for the crystal was 1.5°. 80 images were collected and processed using *DENZO* and *SCALEPACK* (Minor, 1993; Otwinowski, 1993). The statistics of the data set are shown in Table 1. The V_M value of the native crystals is 2.25 Å^3 Da^{-1} for one molecule per asymmetric unit.

The initial structure-determination trials failed using the molecular-replacement method with several AAA structures as search models. It is possible that ClpB NBD1 has some distinct structure features from other AAA molecules. We therefore plan to utilize the MAD method to solve the ClpB NBD1 structure. Selenomethionyl ClpB NBD1 has been produced following the standard protocols (Sha *et al.*, 2000). The plasmid encoding NBD1 was transformed into *E. coli* strain B834(DE3) which is a methionine auxotroph. The cells were grown in a medium consisting 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 purification and crystallization of the Se-Met ClpB NBD1 were carried out using the same conditions as for the wild type. The selenomethionyl NBD1 crystals diffract to 2.0 Å at APS beamline BMC14 (Table 1). The crystal structure determination of ClpB NBD1 by the MAD method is under way.

3. Discussion

We have cloned and crystallized the *E. coli* Hsp100 ClpB nucleotide-binding domain 1 (NBD1). The native crystals diffract to 1.80 Å at a synchrotron X-ray source.

The crystal structures of the Hsp100 protein ClpY from *E. coli* and *H. influenzae* have been determined (Bochtler *et al.*, 2000; Sousa *et al.*, 2000). However, ClpB plays distinct roles from ClpY in cell biology and these two proteins have different domain structures. Firstly, ClpB functions as a molecular chaperone by unfolding denatured polypeptides and does not associate with any proteases. ClpY, on the other hand, binds directly to its partner protease to play roles in misfolded protein degradation. ClpB and ClpY cannot substitute for each other in their cellular functions (Schirmer *et al.*, 1996). Secondly, ClpB has two AAA domains within its primary sequence while ClpY has only one. In ClpB, mutations in each AAA domain will compromise the intrinsic ATPase activity (Schirmer *et al.*, 1996). Sequence alignment among Hsp100 proteins shows that the ATPase domain of ClpY is homologous to ClpB nucleotide-binding domain 2 (NBD2) rather than NBD1 (Schirmer *et al.*, 1996). Thirdly, ClpB NBD1 forms a monomer in solution as suggested by gel-filtration experiments. It is likely that NBD1 is not involved in ClpB oligomerization. In ClpB, NBD2 and the C-terminal domain may be involved in protein oligomerization and NBD1 may not play an important role (Li & Sha, 2001). ClpY forms a hexamer in crystal structure. Based on the crystal structure, both the N-terminal AAA domain and the C-terminal domain of ClpY are involved in protein oligomerization (Bochtler *et al.*, 2000; Sousa *et al.*, 2000). All in all, it is likely that ClpB constitutes a different protein structure from ClpY.

Little work has been performed on the structural studies of *E. coli* Hsp100 ClpB. The crystal structure of ClpB NBD1 will represent the first domain structure from this type of Hsp100 proteins. It may reveal the mechanism by which this AAA domain contributes to ClpB chaperone activity. Moreover, based on the crystal structure, we may be able to propose a hypothesis on how Hsp100 proteins recognize and unfold aggregated polypeptides.

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 at 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.
Glover, J. R. & Lindquist, S. (1998). *Cell*, **94**, 73–82.
Goldberg, A. L. (1992). *Eur. J. Biochem.* **203**, 9–23.
Goloubinoff, P., Mogk, A., Zvi, A. P. B., Tomoyasu, T. & Bukau, B. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 13732–13737.
Gottesman, S. & Maurizi, M. R. (1992). *Microbiol. Rev.* **56**, 592–621.
Gottesman, S., Maurizi, M. R. & Wickner, S. (1997). *Cell*, **91**, 435–438.
Hoskins, J. R., Pak, M., Maurizi, M. R. & Wickner, S. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 12135–12140.
Kim, K. I., Cheong, G., Park, S., Ha, J., Woo, K. M., Choi, S. J. & Chung, C. H. (2000). *J. Mol. Biol.* **303**, 655–666.
Larsen, C. N. & Finley, D. (1997). *Cell*, **91**, 431–434.
Li, J. & Sha, B. D. (2001). Submitted.
Minor, W. (1993). *XDISPLAYF Program.* Purdue University, West Lafayette, IN, USA.
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.
Patel, S. & Latterich, M. (1998). *Trends Cell Biol.* **8**, 65–71.
Schirmer, E. C., Glover, J. R., Singer, M. A. & Lindquist, S. (1996). *Trends Biochem. Sci.* **21**, 289–296.
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.
Wawryznow, A., Banecki, B. & Zylicz, M. (1996). *Mol. Microbiol.* **21**, 895–899.
Weber-Ban, E. U., Reid, B. G., Miranker, A. D. & Horwich, A. L. (1999). *Nature (London)*, **401**, 90–93.
Zolkiewski, M. (1999). *J. Biol. Chem.* **274**, 28083–28086.