Acta Biochimica Polonica

Vol. 47 No. 1/2000

209–214

QUARTERLY

Communication

Cloning, expression, and crystallization of Cpn60 proteins from *Thermococcus litoralis*✪

Jerzy Osipiuk[1]☆✉, Mahalingam Sriram [1], Xuhong Mai[2], Michael W.W. Adams[2] and Andrzej Joachimiak[1]

[1]*Argonne National Laboratory, Biosciences Division, 9700 S. Cass Ave., Argonne, IL 60439, U.S.A.;* [2]*University of Georgia, Department of Biochemistry & Molecular Biology, Athens, GA 30602, U.S.A*

Received: 6 October, 1999; accepted: 27 January, 2000

Key words: archaea, *Thermococcus litoralis*, chaperonin, heat-shock protein

Two genes of the extreme thermophilic archaeon *Thermococcus litoralis* homologous to those that code for Cpn60 chaperonins were cloned and expressed in *Escherichia coli*. Each of the Cpn60 subunits as well as the entire Cpn60 complex crystallize in a variety of morphological forms. The best crystals diffract to 3.6 Å resolution at room temperature and belong to the space group I422 with unit cell parameters a = b = 193.5 Å, c = 204.2 Å.

Cpn60 chaperonins are the only known proteins that are induced by heat stress in hyperthermophilic archaea [1, 2]. These 60-kDa proteins, which show sequence similarity to eukaryotic cytosolic TCP-1 protein complexes, form double-ring structures composed of 16 or 18 subunits [1–3]. However, in contrast to eukaryotic TCP-1 chaperonin, where 7–9 different subunits form an active complex, a single chaperonin gene has been found, e.g. in *Methanococcus jannaschii* [4], and in others two or three different *cpn60* genes have been described (e.g. in *Thermoplasma acidophilum* [5] or *Sulfolobus shibatae*, respectively [6]). Only limited information is available on the biological activities of these proteins and their role in protein folding [3, 7–9], but in some aspects they appear to differ from bacte-

✪This project was supported by the U.S. Department of Energy, Office of Health and Environmental Research, under Contract W-31-109-Eng-38.

☆Present address: EURx, Jaśkowa Dolina 29, 80-286 Gdańsk, Poland.

✉To whom correspondence should be addressed; tel: (48 58) 345 2155; fax: (48 58) 341 7423.

Abbreviations: kDa, kilodaltons; bp, basepairs, IPTG, isopropyl-thiogalactopyranoside; ORF, open reading frame; PCR, polymerase chain reaction.

rial and cytosolic eukaryotic chaperonins. The ring structure of archaeal chaperonins shows very high thermal stability and resistance to denaturants [8], and they appear to function as general thermotolerance factors [9]. In marked contrast, the eukaryotic TCP-1 shows high specificity toward tubulin and actin [10]. All archaeal chaperonins have an ATPase activity similar to other chaperonins [1, 2], but the ring structure appears to dissociate in response to ATP hydrolysis [10]. They bind casein and heat-denatured proteins and the chaperonin from *Sulfolobus solfataricus* was shown to refold inactivated enzymes [11]. Recently, the crystal structure of the archaeal chaperonin from *Thermoplasma acidophilum* was determined, showing structural similarities to the bacterial GroEL chaperonin [12]. A detailed understanding of archaeal chaperonin structure and function is obviously important for understanding the functional and evolutionary aspects of TCP-1 proteins in eukaryotic cells.

In this report, we describe the cloning, protein expression, and crystallization of two subunits of Cpn60 chaperonin from the archaeon *T. litoralis* which can grow up to 97°C [13]. The Cpn60 crystals obtained are the first reported for hyperthermophiles which are considered to grow optimally above 80°C. We compare the gene and amino-acid sequences of *T. litoralis cpn60* genes, examine their relationship to protein sequences of other known archaeal Cpn60 chaperonins, and analyze preliminary crystallographic data of the protein crystals.

Two different approaches were used to isolate the two chaperonin genes from *T. litoralis* DNA. First, we purified the chaperonin directly from *T. litoralis* cells and determined the amino-acid sequence of protein fragments after trypsine digestion. Based on partial amino-acid sequences, degenerate primers for PCR reaction of *T. litoralis* genomic DNA were designed. One PCR product having the predicted length was found; it was used to prepare a radioactive probe for Southern blotting analysis and to screen a recombinant *T. litoralis* DNA library in a pBluescript vector. A clone containing a 2.5 kb *Bgl*II/*Eco*RI DNA fragment covered the entire α subunit gene and was used for DNA sequencing.

To find the chaperonin β subunit gene, we designed 26- to 34-mer PCR degenerate primers based on highly conserved regions of archaeal chaperonins and modified for codon frequency in *T. litoralis*. The initial PCR products were digested with the *Hin*dIII restriction enzyme to distinguish between the two chaperonin genes and were separated on agarose gels. Purified DNA fragments were then subcloned to a pBluescript vector and sequenced to confirm homology to chaperonin genes and to exclude clones of the α subunit. The subcloned DNA of the β subunit gene fragment was also used as a PCR template to obtain a DNA probe for Southern blot analysis. The results indicated that the 2.5 kb *Eco*RI/*Bgl*II restriction fragment contained the entire coding sequence. These two restriction enzymes were used to prepare a *T. litoralis* genomic library that served as a template for two PCR reactions. For each reaction, we used one of the pBluescript universal primers, T3 or T7, and one of two primers based on the partial DNA sequence of the β subunit found earlier with degenerate PCR primers. PCR reactions with those two pairs of primers produced two DNA fragments covering the 5′ and 3′ regions of the gene and overlapping the middle position of the gene. For DNA sequencing, these PCR products were again subcloned into the pBluescript vector. To minimize errors resulting from the PCR reactions, the Pfu DNA polymerase was used, which shows ten times higher fidelity than *T. aquaticus* polymerase, and the PCR cloning and sequencing procedures were repeated to confirm reproducibility.

Sequencing of DNA was done using an ABI automatic sequencer at the DNA Sequencing Facility of the University of Chicago, and chromatogram patterns were analyzed using Seqman program (DNASTAR Inc., Madison,

WI, U.S.A.). DNA sequences for both subunits are available from the Genome Sequence DataBase (GSDB) under accession numbers GSDB:S:76146 and GSDB:S:1113612.

Upstream of the open reading frames (ORFs), we localized putative promoter regions for both genes. The archaeal boxA sequence TTTATA, corresponding to the consensus archaeal promoter TTTA(T/A)A [14], was found 96 bp upstream of the α subunit gene and 51 bp upstream of the β subunit gene putative transcription starts (Fig. 1). Putative transcription termination sequences, TTTCTCTTCTT, identical for both genes, were localized 11 and 1 bp downstream of the stop signal in the α and β genes, respectively. This sequence resembles the termination sequence (TTTTT(C/T)T) found for the chaperonin genes in *S. shibatae* [15].

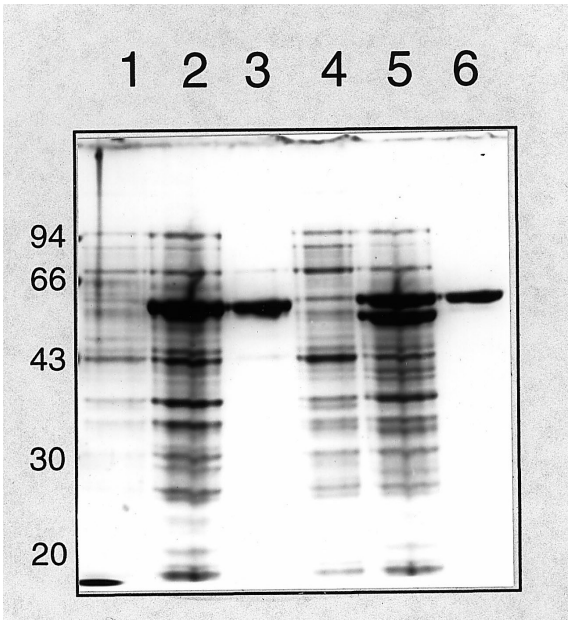

Figure 1. Expression of *T. litoralis* Cpn60 proteins in *E. coli*.

E. coli BL21(DE3) LysS strains carrying plasmids overproducing single subunits of *T. litoralis* Cpn60 proteins were grown in 2xYT medium at 37°C and induced with IPTG. Samples were denatured and run on a 10% polyacrylamide gel in the presence of SDS. Lane 1, uninduced *E. coli* cells carrying plasmid overproducing Cpn60 α subunit; lane 2, IPTG-induced *E. coli* cells carrying plasmid overproducing Cpn60 α subunit; lane 3, purified Cpn60 α subunit; lane 4, uninduced *E. coli* cells carrying plasmid overproducing Cpn60 β subunit; lane 5, IPTG-induced *E. coli* cells carrying plasmid overproducing Cpn60 β subunit; lane 6, purified Cpn60 β subunit.

The ORFs code for two proteins, of 544 and 548 amino acids, that have very similar predicted molecular mass, 58,770 and 58,852 Da. In the β subunit, we found two successive methionine codons at the N-terminus of the coding sequence, both of which could serve as the starting codon for translation. It remains to be determined if the use of both codons serves some as yet unknown regulatory role. In this report, because of sequence similarity to other archaeal chaperonins, we assumed that the first methionine is the starting point for translation. The amino-acid sequence of the α and β subunits is very similar (82.4% identity). Some differences are found at the N-terminus, but the most striking ones are at the C-terminus. The α subunit C-terminal sequence is rich in glutamic acid and resembles the acidic C-terminal regulatory motif of Hsp70 chaperones [16]. In contrast, the β subunit C-terminal sequence is rich in glycine and methionine and shares the MGG motif with bacterial chaperonins [17]. This sequence has been shown to influence the stability of the double-ring structure of the *E. coli* GroEL chaperonin [18].

We compared the *T. litoralis* protein sequences to other known archaeal chaperonins using the BLAST WWW server at National Center for Biotechnology Information at NIH (U.S.A.). All archaeal proteins are highly homologous. The *T. litoralis* chaperonins bear the closest similarity to the thermophilic factor *of Pyrococcus horikoshii* [19] and the chaperonin α subunit of *Thermococcus* sp. [20] for *T. litoralis* α and β subunits, respectively. These proteins show similarity close to 86% and identity to about 80%, based on the GAP program from the GCG Sequence Analysis Package (Genetics Computer Group, University of Wisconsin, Madison, WI, U.S.A.).

Expression and purification of separate subunits of *T. litoralis* Cpn60 protein were performed as described for the isolation of DnaK protein from *Thermus thermophilus* bacterium

[21]. Briefly, protein expression was achieved by using pET21a expression vector cut with *Nde*I and *Eco*RI enzymes and *E. coli* BL21-(DE3)pLysS cells as a bacterial host. After 4 h induction with 1 mM isopropyl-thiogalactopyranoside (IPTG), cells were harvested and opened by sonication. Purification of the proteins utilized heat denaturation of the host proteins at 65°C for 15 min and protein chromatography on Fast Q-Sepharose and FPLC Mono-Q columns. In the case of the β subunit of Cpn60, size exclusion chromatography could be used for further purification.

Both subunits of *T. litoralis* Cpn60 protein were efficiently expressed (Fig. 1, lanes 1, 2, 4 and 5). After purification, we obtained proteins which were at least 90% pure as judged by SDS/polyacrylamide electrophoresis (Fig. 1, lanes 3 and 6). Expression of the β subunit was accompanied by the induction of a 55 kDa protein (Fig. 1, lane 4) which, most likely, was a truncated β subunit resulting from protein translation starting at methionine codon placed 141 bases downstream from the starting methionine codon. This artifact was easily separated from the full-length form of the β subunit by the purification procedure described above. Both subunits alone were able to form oligomeric complexes having a size that corresponded to the double-ring forms of the active Cpn60 protein based on electrophoresis in nondenaturing conditions (results not shown). However in the case of the α subunit, we observed a much smaller fraction of oligomeric forms.

All tested proteins i.e., Cpn60 protein purified directly from *T. litoralis* cells, the protein reconstituted from two recombinant proteins, and the separate subunits, crystallized at room temperature in a variety of morphological forms from methylpentanediol (MPD) in the presence of divalent cations (Mg^{2+}, Ca^{2+}, Ba^{2+}) in Bis-Tris-propane buffer, pH 7.2. The best crystals were obtained from protein purified directly for *T. litoralis* cells (Fig. 2). They diffract to 3.6 Å resolution at room temperature using CuKα (λ = 1.54 Å) X-rays generated from a Rigaku RU-200 rotating anode operating at 50 kV and 94 mA. Data were processed and reduced with the HKL suite of programs [22] up to the resolution of 4.05 Å (Table 1). The decrease in resolution is due to crystal decay which also prevented collection of a full data set. The crystals belong to the space group I422 with unit cell parameters a = b = 193.5 Å, c = 204.2 Å. The same type of crystal symmetry was determined for *T. acidophilum* thermosome crystals grown in different conditions of acetate buffer pH 5.6 in the presence of ammonium sulfate [12]. In our case,

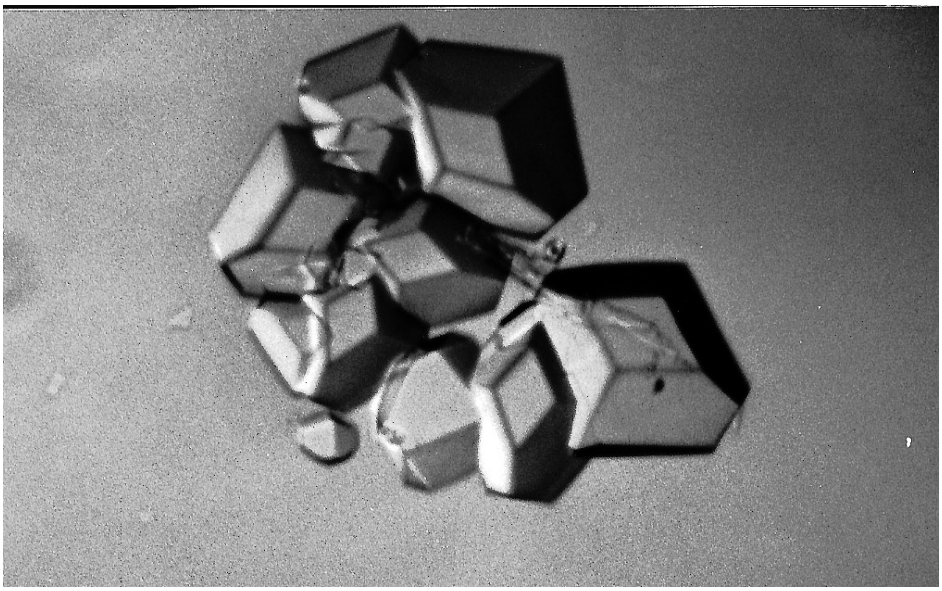

Figure 2. Crystals of *T. litoralis* Cpn60 protein.

The largest crystal have dimensions 0.4 mm × 0.4 mm × 0.4 mm.

Table 1. Summary of data collection and processing for *Thermococcus litoralis* Cpn60 crystals

	Data set 1	Data set 2	Two data sets scaled together
Resolution (Å)	9–4.05	12–4.55	12–4.05
Number of observations	70586	107873	182168
Number of unique reflections	16135	11389	16135
Space group	I422	I422	I422
Unit cell parameters	a = b = 193.7 Å, c = 204.0 Å, $\alpha = \beta = \gamma = 90^o$	a = b = 191.7 Å, c = 204.2 Å, $\alpha = \beta = \gamma = 90^o$	a = b = 193.5 Å, c = 204.2 Å, $\alpha = \beta = \gamma = 90^o$
Completeness (%)	69.5	98.6	91.8
χ^2	1.23	1.10	1.12
R_{merge} (%)	23.1	21.5	23.1
$I/\sigma(I)$	4.72	5.43	5.44

$R_{merge} = \Sigma_{hkl} \Sigma^N_{i=1} | < I_i^{hkl} > - I_i^{hkl} | / \Sigma_{hkl} \Sigma^N_{i=1} | I_i^{hkl} |$

$\chi^2 = \{\Sigma_{hkl} \Sigma^N_{i=1} (< I_i^{hkl} > - I_i^{hkl})^2 / [\sigma(I_i^{hkl})]^2\} \quad (N-1)/N$

however, the unit cell parameters a and b are 15% greater than those for *T. acidophilum* crystals resulting probably from different crystal packing. The diffraction pattern shows distinct 4-fold symmetry with two subunits in the asymmetric unit. This is consistent with the idea that archaeal Cpn60 proteins consist of two heterooligomeric rings suggested by Nitsch *et al.* [22] from electron cryo-microscopy studies and later confirmed by a crystallographic analysis [12]. In that model, each of the two rings consists of four α subunits and four β subunits. Each of the two subunits in the complex has the nearest neighbors of the other type.

The high amino-acid sequence similarity between *T. litoralis* and *T. acidophilum* chaperonins (65–69%, Table 2) should allow the determination of the structure of *T. litoralis* chaperonin by the molecular replacement method using the *T. acidophilum* structure as a model. Using our data, we were able to find a single solution using the CNS suite [24]. As in the case of the *T. acidophilum* structure [12], we observed a double-ring structure for the Cpn60 protein complex. Those results strongly suggest that better quality crystals diffracting to higher resolution will be sufficient to reveal the detailed structure of the *T. litoralis* Cpn60 complex.

Table 2. The percent of similarity/identity of amino-acid sequences of *Thermococcus litoralis* and *Thermoplasma acidophilum* chaperonins.

The values are based on GAP program from the GCG Sequence Analysis Package (Genetics Computer Group, University of Wisconsin, Madison, WI, U.S.A.).

	T. litoralis β subunit	*T. acidophilum* α subunit	*T. acidophilum* β subunit
T. litoralis α subunit	88.0/82.4	69.6/61.7	68.3/57.9
T. litoralis β subunit	–	67.8/59.7	64.8/54.7
T. acidophilum α subunit	–	–	69.3/59.9

We thank Dr. Elsie Quaite-Randall and Dr. Martin Walsh for their helpful comments in preparation of this manuscript.

REFERENCES

1. Trent, J., Nimmesgern, E., Wall, J.S., Hartl, F.-U. & Horwich, A.L. (1991) *Nature* **354**, 490–493.

2. Phipps, B.M., Typke, D., Hegerl, R., Volker, S., Hoffmann, A., Stetter, K.O. & Baumeister, W. (1993) *Nature* **361**, 475–477.

3. Knapp, S., Schmidt-Krey, I., Hebert, H., Bergman, T., Jornvall, H. & Ladenstein, R. (1994) *J. Mol. Biol.* **242**, 397–407.

4. Bult, C.J. *et al.* (1996) *Science* **273**, 1058–1073.

5. Waldmann, T., Nimmesgern, E., Nitsch, M., Peters, J., Pfeifer, G., Muller, S., Kellermann, J., Engel, A., Hartl, F.-U. & Baumeister, W. (1995) *Eur. J. Biochem.* **227**, 848–856.

6. Archibald, J.M., Logsdon, J.M. & Doolittle, W.F. (1999) *Curr. Biol.* **9**, 1053–1056.

7. Kagawa, H.K., Osipiuk, J., Maltsev, N., Overbeek, R., Quaite-Randall, E., Joachimiak, A. & Trent, J.D. (1995) *J. Mol. Biol.* **253**, 712–725.

8. Quaite-Randall, E., Trent, J.D., Josephs, R. & Joachimiak, A. (1995) *J. Biol. Chem.* **270**, 28818–28824.

9. Trent, J.D., Osipiuk, J. & Pinkau, T. (1990) *J. Bacteriol.* **172**, 1478–1482.

10. Tian, G., Vainberg, I.E., Tap, W.D., Lewis, S.A. & Cowan, N.J. (1995) *Nature* **375**, 250–253.

11. Guagliardi, A., Cerchia, L., Bartolucci, S. & Rossi, M. (1994) *Protein Sci.* **3**, 1436–1443.

12. Ditzel, L., Lowe, J., Stock, D., Stetter, K.-O., Huber, H., Huber, R. & Steinbacher, S. (1998) *Cell* **93**, 125–138.

13. Neuner, A., Jannasch, H.W., Belkin, S. & Stetter, K.O. (1990) *Arch. Microbiol.* **153**, 205–207.

14. Reiter, W.-D., Palm, P. & Zillig, W. (1988) *Nucleic Acids Res.* **16**, 1–19.

15. Reiter, W.-D., Palm, P. & Zillig, W. (1988) *Nucleic Acids Res.* **16**, 2445–2459.

16. Freeman, B.C., Myers, M.P., Schumacher, R. & Morimoto, R. (1995) *EMBO J.* **14**, 2281–2292.

17. Hemmingsen, S.M., Woolfor, D.C., Van Der Vies, S.M., Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W. & Ellis, R.J. (1988) *Nature* **333**, 330–334.

18. Burnett, P.P., Dorwich, A.L. & Low, K.B. (1994) *J. Bacteriol.* **176**, 6980–6985.

19. Kawarabayasi, Y., *et al.* (1998) *DNA Res.* **5**, 147–155.

20. Yoshida, T., Yohda, M., Iida, T., Maruyama, T., Taguchi, H., Yazaki, K., Ohta, T., Odaka, M., Endo, I. & Kagawa, Y. (1997) *J. Mol. Biol.* **273**, 635–645.

21. Osipiuk, J. & Joachimiak, A. (1997) *Biochem. Biophys. Acta* **1353**, 253–265.

22. Otwinowski, Z. & Minor, W. (1995) *The HKL Manual*, Yale University Press, New Haven, CT.

23. Nitsch, M., Klumpp, M., Lupas, A. & Baumeister, W. (1997) *J. Mol. Biol.* **267**, 142–149.

24. Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T. & Warren, G.L. (1998) *Acta Cryst.* D **54**, 905–921.