Analysis of the AAA+ chaperone clpB gene and stress-response expression in the halophilic methanogenic archaeon *Methanohalophilus portucalensis*

Chao-Jen Shih and Mei-Chin Lai

Department of Life Sciences, National Chung Hsing University, Taichung, Taiwan.

ClpB is a member of the protein-disaggregating chaperone machinery belonging to the AAA+ superfamily. This paper describes a new *clpB* gene from the halophilic methanoarchaeon *Methanohalophilus portucalensis*, which has not been reported previously in Archaea. The partial sequence of *clpB* was identified from the investigation of the salt-stress response of *Me. portucalensis* by differential-display RT-PCR (DDRT-PCR). Furthermore, the complete *clpB* sequence (2610 nt) and its upstream genes encoding the type I chaperonin GroEL/ES were obtained through inverse PCR, Southern hybridization and sequencing. The G+C ratio of *clpB* is 49.6 mol%. The predicted ClpB polypeptide contains 869 aa and possesses a long central domain and a predicted distinctly discontinuous coiled-coil motif separating two nucleotide-binding domains (NBD1 and NBD2). NBD1 has a single Walker A and two Walker B motifs and NBD2 has only one of each Walker motif, a characteristic of HSP100 proteins. Two repeated Clp amino-terminal domain motifs (ClpN) were identified in ClpB. The putative amino acid sequence shared 75.6% identity with the predicted *clpB* homologue annotated as ATPase AAA-2 of *Methanococcoides burtonii* DSM 6242. Preliminary phylogenetic analysis clustered *Me. portucalensis* ClpB (MpClpB) with the low G+C Gram-positive bacteria. Stress response analysis of *clpB* by Northern blotting showed up to 1.5-fold increased transcription levels in response to both salt up-shock (from 2.1 to 3.1 M NaCl) and down-shock (from 2.1 to 0.9 M NaCl). Both *clpB* and *groEL/ES* transcript levels increased when the temperature was shifted from 37°C to 55°C. Under heat stress *clpB* transcription was repressed by the addition of the osmolyte betaine (1 mM). In conclusion, a novel AAA+ chaperone *clpB* gene from a halophilic methanogen that responded to the fluctuations in temperature, salt concentration and betaine has been identified and analysed for the first time.

**INTRODUCTION**

Molecular chaperones are ubiquitous in bacteria and eukaryotes. They can help nascent polypeptides to fold correctly and reactivate aggregated proteins which were damaged by stress (Lee & Tsai, 2005). One would expect that molecular chaperones also function in archaea that can survive in extreme environments, and help in their survival and adaptation under stresses such as temperature, pH, salt and pressure.

Molecular chaperones of all organisms can be classified into four types: prefoldins, the molecular chaperone machine (composed of DnaK, DnaJ and GrpE) and group I and II chaperonins (Conway de Macario et al., 2003). Prefoldins are composed of an α-subunit and a β-subunit, and their major function is to work with chaperonins to facilitate the correct folding of cellular proteins (Martin et al., 2004). Prefoldin genes can be found in all the archaeal genomes that have been sequenced (Macario et al., 2004). It has been verified that the mRNA levels of the prefoldin α-subunit gene of *Methanocaldococcus jannaschii* is upregulated during heat-shock stress (Boonyaratankonkit et al., 2005). The molecular chaperone machine is considered to be formed by Hsp70 (DnaK), Hsp40 (DnaJ) and a nucleotide-exchange factor (e.g. GrpE in prokaryotes and BAG-1 in eukaryotes); it catalyses the refolding of denatured proteins and assists the folding of...
newly synthesized proteins, preventing aggregation. Heat-shock proteins (HSPs) have been identified and characterized from various bacteria and archaea. Among archaea, Methanocarcina acetivorans has all versions of chaperonins and HSPs (Galagan et al., 2002; Conway de Macario et al., 2003) and has features in common with bacterial and eukaryotic heat-shock systems. Surprisingly, the major chaperone classes Hsp100/Hsp83 are absent from the genomes of the hyperthermophilic archaea, although they are present in mesophilic archaea (Laksanalamai et al., 2004). Group I chaperonins, such as GroEL and GroES, catalyse correct polypeptide folding. It has been verified that the group I chaperonins exist in archaea after whole-genome sequence analysis of Msa. acetivorans was performed (Galagan et al., 2002) and it has been conjectured that the bacterial chaperonin genes were obtained by lateral gene transfer (Depenmeier et al., 2002; Klunker et al., 2003), but this has been disputed (Macario et al., 2006). Group II chaperonins, which mediate protein folding, such as thermosome subunits and Hsp60, exist exclusively in eukaryotes and archaea and so far have not been found in bacteria (Laksanalamai et al., 2004). Among the extreme thermophiles, group II chaperonins have been investigated in Thermoplasma spp. (thermosomes), Sulfolobus spp. (archaeosomes) and Pyrococcus spp. (rosettasomes) (Kagawa et al., 1995; Quaite-Randall et al., 1995; Waldmann et al., 1995).

Overall, all four types of molecular chaperone systems have been identified in mesophilic archaea, especially Msa. acetivorans, but not in all archaea (Laksanalamai et al., 2004; Conway de Macario et al., 2003).

The Clp/Hsp100 family, a class within the AAA+ (ATPases associated with various cellular activities) superfamily (Dougan et al., 2002; Neuwald et al., 1999; Schirmer et al., 1996), is involved in a variety of functions, such as protein quality control (refolding, disaggregation and degradation), membrane fusion and vesicular transport, DNA replication and repair, and cytoskeleton regulation (Ogura & Wilkinson, 2001). Clp/Hsp100 proteins play an essential role as members of both the chaperone and the protease machinery (Schirmer et al., 1996). The Clp/Hsp100 family includes eight subfamilies (ClpA, B, C, D, M, N, X and Y) (Schirmer et al., 1996). ClpA and ClpX are associated with the ClpP protease (Singh et al., 2000, 2001) and are involved in protein degradation and quality control. In contrast, ClpB/Hsp104 does not associate with a cellular protease (Woo et al., 1992), but participates in the multi-chaperone system that efficiently inhibits and reverses protein aggregation (Ben-Zvi & Goloubinoff, 2001; Zolkiowski, 1999). Bacterial ClpB and its eukaryotic orthologues, plant Hsp101 and yeast Hsp104, are heat-shock-induced, and are essential proteins for the stress response (Lee et al., 2004; Lee & Tsai, 2005; Queitsch et al., 2000; Sanchez & Lindquist, 1990; Squires et al., 1991).

Among archaea, ClpA/B homologues are predicted to be encoded by only a few methanogen genomes to date, including Methanospirillum hungatei YP_502751 (optimum growth temperature 37 °C), Methanococcales burtonii YP_566814 (23.4 °C) and the distant ClpB-type homologue of Methanothermobacter thermautotrophicus AAB84790 (65–70 °C) which could have been acquired by horizontal or lateral gene transfer from bacteria (Neuwald et al., 1999; Laksanalamai et al., 2004).

The strictly anaerobic Methanohalophilus portucalensis strain FDF1 can grow over a range of external NaCl concentrations from 1.2 to 2.9 M (Mathrani & Boone, 1985; Boone et al., 1993; Lai et al., 1991). To counter the changing osmotic stress, the cell could transport betaine or de novo synthesize betaine, β-glutamine and N-acetyl-β-lysine as compatible solutes to respond to salt concentration of the external environment (Lai et al., 1991, 1999, 2000, 2006). In addition to the accumulation of organic solutes, the internal K+ concentration also increased from 0.6 to 1.1 M (Lai et al., 1991). In order to elucidate the molecular mechanisms of salt adaptation and stress response of the halophilic methanoarchaeon Meth. portucalensis, differential-display RT-PCR (DDRT-PCR) was used. The differential expression of genes of Meth. portucalensis in response to hyper-saline stress (3.1 M NaCl) and hypo-saline stress (0.9 M NaCl) was studied and a clpB gene was identified and characterized. In this report, the AAA+ chaperone clpB gene of Meth. portucalensis is described for the first time as is its stress-response expression.

**METHODS**

**Organisms, vector and growth conditions.** Meth. portucalensis strain FDF1 (=OCMS9) was isolated from the solar saltern of Figueria da Foz, Portugal, by Mathrani & Boone (1985) and provided by R. Mah, School of Public Health, University of California, Los Angeles, USA (Boone et al., 1993; Lai et al., 1991). The cells were routinely incubated at 37 °C in defined medium that contained 2.1 M NaCl 1 and 40 mM trimethylamine as the sole carbon and energy source (Lai et al., 1991). Sterile medium was prepared under a N2/CO2 atmosphere (4:1) by a modification of the Hungate technique (Balc et al., 1979). The medium was anaerobically dispensed into serum bottles, which were then sealed with butyl rubber stoppers and aluminium crimp closures and autoclaved at 121 °C for 20 min. The methanogenic substrate and Na2S, H2O2 were prepared under the N2 atmosphere and were added to the sterile medium at 40 mM and 1 mM respectively just prior to cell inoculation. Sealed serum bottles were inoculated with a 5% volume of late-exponential phase culture containing Na2S2O3 and measuring the OD540.

The pGEM-T easy vector (Promega) was used to clone the DDRT-PCR products. The tRBA cloning vector (Yeastern Biotech) was used to clone the MpcldB gene. The pUC18 cloning vector was used for cloning the 5′ end and region upstream of Mplcb. Escherichia coli strains DH5α and JM101 were used as the hosts. E. coli was grown in Luria–Bertani (LB) medium or on solid medium at 37 °C (Sambrook & Russell, 2001). Ampicillin was added to the medium at a concentration of 100 μg ml−1.

**Chromosomal DNA and RNA extraction.** Chromosomal DNA of Meth. portucalensis was isolated and purified by a modification of the methods of Jarrell et al. (1992) and Johnson (1985). Cells (500 ml)
were harvested at late-exponential phase by centrifugation and lysed with lysis buffer (10 mM Tris, 1 mM EDTA, 2% SDS and 100 μg protease K ml⁻¹). After phenol/chloroform extraction, nucleic acids were precipitated with 2-propanol and collected with a glass rod. The salt contained in the DNA was dissolved in 70% cold ethanol and nucleic acids were dissolved in 0.1 x standard saline citrate buffer (SSC). After treatment with RNase (50 μg ml⁻¹), the extraction and washes were repeated to purify the chromosomal DNA, which was dissolved in 0.1 x SSC.

To prepare RNA, mid-exponential-phase cells (50 ml, OD₅₄₀ ~0.6–0.8, ~10⁸ cells) were harvested and lysed with 1 ml of RareRNA (Genepure Technology) and incubated for 5 min at room temperature. After the addition of 0.3 ml chloroform, the mixture was shaken vigorously for 15 s, further incubated on ice for 5 min, and centrifuged at 12000g for 4 ℃ for 10 min. The aqueous phase, containing RNAs, was transferred into a new tube and the nucleic acids were precipitated with 2-propanol. The RNA pellet was dried by Speed Vac (SC100, Savant) then dissolved in 20 μl DEPC-treated water. Total RNA concentration was determined from A₂₆₀.

Cloning the clpB gene of Mei. portucalensis. Total RNA (5 μg) from control (cells grown in medium containing 2.1 M NaCl) and experimental samples (cells grown in medium containing 0.9 or 3.1 M NaCl) was subjected to reverse transcription in a GeneAmp PCR System 2400 (Perkin Elmer). The cDNA synthesis reactions were carried out for 1 h at 42 ℃ by using ImProm-II reverse transcriptase (Promega). The reaction was stopped by incubation at 95 ℃ for 5 min. The cDNAs were further amplified by PCR with two synthesized arbitrary primers P1 (5'-GCCAAAGGCTCAG-3') and P2 (5'-GGGTCATGCG-3') and Taq DNA polymerase (Takara). The resulting products were run in a 1.5% TAE/agarose gel, and the differentially expressed fragments were excised from the gel and purified using the GENECLEAN III (BIO 101 Systems, Q-BIOgene).

To test the relationship of clpB transcription with heat shock, mid-exponential-phase cells incubated at 37 ℃ were shifted to 15, 20, 25, 30, 33, 40, 45, 50 and 55 ℃. For a detailed time-course study measurements were obtained at 5, 10, 15, 30, 60, and 120 min. The effects of the osmolyte betaine on clpB transcription under temperature and salt stress were examined by using two parallel sets of mid-exponential-phase cultures (OD₅₄₀ 0.7) grown at 2.1 M NaCl, 37 ℃. Cultures were tested with salt/osmotic stress and temperature stress as described above, with or without the addition of betaine (1 mM). More than two different culture batches were tested for each stress response.
groEL-F: 5'-AGGTGGGCTGAGTAAAC-3' and groEL-R (5'-CTTGGGGTGGTCTCTG-3'), and the clpB-2.6 probe were labelled with the DIG DNA labelling kit (Roche Applied Science). The test was run in triplicate with the same batch of cultures.

**DNA and amino acid sequence analyses.** Sequence comparison of ClpB was performed by using information from the NCBI web site (http://www.ncbi.nlm.nih.gov/BLAST/). Conserved domain and motif analyses were carried out by using the SMART web site (http://smart.embl-heidelberg.de). The ClpB/Hsp100 amino acid sequences were obtained from the database of the NCBI web site (http://www.ncbi.nlm.nih.gov). Multiple sequence alignments were constructed by using the SDSC Biology Workbench website (http://workbench.sdsc.edu) with the CLUSTALW program (Thompson et al., 1994). Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 3.1 (Kumar et al., 2004). The coiled-coil structures were analysed using the COILS program from ISREC, Switzerland (http://www.ch.embnet.org/software/COILS_form.html). Sequences were run with the weighted MTIDK matrix.

**RESULTS**

**Meh. portucalensis clpB cloning and analyses**

By DDRT-PCR of *Meh. portucalensis* mRNA from 0.9, 2.1 and 3.1 M NaCl cultures, fourteen differentially expressed cDNA fragments were obtained from the 0.9 or 3.1 M NaCl cultures (data not shown). Two of the differentially expressed genes from the 3.1 M NaCl culture shared high identity (70% and 74%) with the genes encoding transmembrane protein (MttP) and monomethylamine methyltransferase (MtmB), which participate in the unique energy-acquiring methanogenesis of methanogens. Three putative stringent-response-related genes were also identified by DDRT-PCR as the ribosomal protein L29, which was down-regulated in 0.9 M NaCl culture, and the translation elongation factor 1A and the lysyl tRNA synthetase, which were up-regulated in 0.9 M NaCl culture. The predicted genes for cobalamin biosynthesis protein (CbiD) (upregulated in 3.1 M NaCl culture), methenyl-H4MPT cyclohydrolase (upregulated in 0.9 M NaCl culture) and sulfite reductase (upregulated in 0.9 M NaCl culture) were also identified to be involved in the salt-stress response. Predictably, three known stress-response gene homologues were identified. One was the DNA mismatch repair protein Muts (upregulated in 3.1 M NaCl culture), the second was the universal stress protein USPA (upregulated in 0.9 M NaCl culture) and the third the differentially expressed cDNA fragment (1326 bp) that was obtained from the 3.1 M NaCl culture and had a deduced amino acid sequence that shared 55% identity with the putative ClpB protein of *Thermosynechococcus elongatus* BP-1.

The 1326 bp cDNA fragment, designated clpB-1.3, possessed a nucleotide-binding domain (NBD), which was similar to NBD2, and also possessed the conserved Walker motifs (Walker A and Walker B) typical of a bacterial-type ClpB chaperone (Schirmer et al., 1996). By using the clpB-1.3 as template, inverse PCR was performed and an extended 2042 bp partial clpB gene (clpB-2.0) with the downstream sequence was obtained. There was a 652 bp overlap between clpB-1.3 and clpB-2.0 and all these sequences were 100% identical. However, inverse PCR could not complete the 5' end of clpB so a labelled clpB-2.0 DNA probe was generated and the upstream and 5' end of clpB gene (total 1986 bp) was obtained through Southern hybridization and was further cloned and sequenced. There was a 1153 bp overlap between clpB-2.0 and this fragment and all sequences were 100% identical. Combining the results of DDRT-PCR, inverse PCR and Southern blotting, the complete clpB gene (2610 bp) and its upstream and downstream adjacent genes were obtained. The full length clpB gene (2610 bp) was amplified from the chromosomal DNA of *Meh. portucalensis* for further overexpression and Northern investigation by using the specific primers ClpBoeF and ClpBoeR designed as described in Methods. The sequence of PCR-amplified full-length clpB-2.6 was 100% identical with sequences obtained from the combination of DDRT-PCR, inverse PCR and Southern blotting.

The G+C content of clpB from *Meh. portucalensis* is 49.6 mol%, which is close to the G+C content of its genome (44 mol%). This is similar to the G+C contents of the *Mec. burtonii* ATPase AAA-2 gene (45.4 mol%), *M. hungatei* ATPase AAA-2 gene (47.5 mol%) and *Met. thermoautotrophicus* ClpA/B gene (47.4 mol%), which are close to the G+C contents of their genomes at 40, 45 and 49 mol%, respectively. Analysis of the promoter region of *Meh. portucalensis* clpB indicated that the predicted TATA box (5'-TATAAA-3') was located at the −37 to −42 position from the translation start codon and the BRE motif 5'-AAA-3' preceded the TATA box. The Shine–Dalgarno sequence 5'-GGAGG-3' was located six nucleotides upstream from the start codon (Fig. 1). Sequence analysis indicated that groEL and groES were located upstream of clpB in the same direction and CDP-diacylglycerol-serine O-phosphatidyltransferase and phosphatidylserine decarboxylase genes were located downstream in the opposite direction (Fig. 1). The deduced amino acid sequence of the complete groEL (1611 bp designed as MpgroEL) and the complete groES (276 bp designed as MpgroES) that we obtained shared 74% identity with the GroEL of *Methanosarcina barkeri* str. Fusaro (AAZ70491) and 71% identity with the GroES of *Msa. barkeri* str. Fusaro (AAZ70490). The complete CDP-diacylglycerol-serine O-phosphatidyltransferase gene shared 56% amino acid identity with the CDP-diacylglycerol-serine O-phosphatidyltransferase (YP_566813) of *Mec. burtonii* DSM 6242, and the partial phosphatidylserine decarboxylase-related gene shared 67% amino acid identity with the phosphatidylserine decarboxylase related protein (YP_566812) of *Mec. burtonii* DSM 6242.

**Sequence analyses and characteristics of AAA+ chaperone ClpB from *Meh. portucalensis***

The predicted polypeptide of ClpB from *Meh. portucalensis* (MpClpB) contains 869 aa with a calculated molecular
mass of 98,070.24 Da and a calculated isoelectric point of 5.09. Compared to the other 16 ClpB/Hsp100/Hsp101/Hsp104 proteins from bacteria and eukaryotes (Table 1), the molecular mass of \( M_p \text{ClpB} \) is smaller than that of proteins from plants and yeast but larger than that from bacteria and protozoa. Except for the \( L_d \text{Hsp100} \) from \textit{Leishmania donovani} (pI 7.43), the calculated isoelectric points of all ClpBs were acidic. \( M_p \text{ClpB} \) contains 39.59% hydrophobic, 16.80% basic, 18.41% acidic and 25.20% neutral amino acids. The identity of the amino acid sequence between \( M_p \text{ClpB} \) and the \( M_b \text{ATPase} \), a ClpB homologue we identified from the genome database of \textit{Mec. burtonii} (Saunders et al., 2003), is 75.6% (Table 1). Results of a preliminary study of the phylogenetic relationships among ClpB/Hsp100 proteins within the three domains are shown in Fig. 2. The phylogenetic tree revealed that \( M_p \text{ClpB} \) from \textit{Meh. portucalensis} and its archaeal homologue \textit{Mec. burtonii} \( M_b \text{ATPase} \) as well as \textit{Mes. hungatei} \( M_h \text{ATPase} \) are grouped with bacterial ClpBs and distant from the eukaryotic ones. Methanoarchaeal ClpBs (\( M_p \text{ClpB}, \ M_b \text{ATPase} \) and \( M_h \text{ATPase} \)) are more closely related to those from the low-G+C Gram-positive bacteria such as \textit{Bacillus cereus} (\( B_c \text{ClpB} \)) than they are to those from the high-G+C Gram-positive bacteria such as \textit{Corynebacterium glutamicum} (\( C_g \text{ClpB} \)) (Fig. 2).

Analysis of the amino acid sequence of \( M_p \text{ClpB} \) revealed that there were two typical Clp signature motifs, i.e. the highly conserved NBD1 and NBD2; furthermore, NBD1 had a single Walker A (GXXGXGKT) and two Walker B (hhhhDE) motifs and NBD2 had only one of each Walker motif (Barnett et al., 2000; Schirmer et al., 1996; Squires & Squires, 1992) (Fig. 3a and Supplementary Figure S1). Two repeated Clp amino-terminal domain motifs (ClpN), which are typical of ClpA and ClpB and most ClpC proteins (Barnett et al., 2000), were also identified in \( M_p \text{ClpB} \) (Fig. 3a and Supplementary Figure S1). A long middle region (M) separating the two ATP-binding sites, characteristic of ClpB proteins, was also detected in \( M_p \text{ClpB} \) and was 124 aa in length (~130 aa for other ClpBs). For all known ClpB proteins, this M region contains a predicted coiled-coil motif (Celerin et al., 1998; Lee et al., 2003) formed by a bundle of \( \alpha \)-helices that are wound into a superhelix (Lupas, 1996; Martin et al., 2004). Using the COILS program, two major peaks with probability greater than 0.5 of forming a coiled-coil structure were

\[ \text{Fig. 1. Gene order in the clpB locus and clpB regulatory regions in Methanohalophilus portucalensis. DNA sequences of the promoter region and the region separating MpgroEL and MpclpB are also shown. Apparent transcription regulatory elements, as well as start and stop codons, are indicated with bars. Arrows indicate the major inverted repeat of the groEL gene. Solid lines indicate the completely sequenced genes. Dashed lines indicated incompletely sequenced genes. The lower two gene organizations show the homologous genes of Methanococcoides burtonii (MbpClpB) and Methanospirillum hungatei (Mhcpn10, MpgroEL and MhClpB).} \]
predicted. The first peak falls within the N-terminal domain (amino acids 92–119, score 0.76), and the second peak coincides with the position of the middle region (amino acids 411–444, score 1.00; 465–492, score 0.61; 507–534, score 0.517) (Fig. 3b). The middle region in \( \text{Mp} \) ClpB contains three segments of periodic seven-residue repeats a-b-c-d-e-f-g, in which the first position (a) and the fourth (d) are occupied by predominantly hydrophobic amino acids, whereas the neighbouring positions (g and e) are occupied by charged residues (Fig. 3c), which is similar to the middle region of \( \text{Mb} \) ATPase (data not shown).

Table 1. Percentage identity of \( \text{Meh. portucalensis} \) ClpB with bacterial ClpB and eukaryotic Hsp100/Hsp101/Hsp104

<table>
<thead>
<tr>
<th>Organism</th>
<th>Protein</th>
<th>Length (aa)</th>
<th>pI</th>
<th>Identity with MpClpB (%)</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methanohalophilus portucalensis FDF1</td>
<td>MpClpB</td>
<td>869</td>
<td>5.09</td>
<td>100</td>
<td>DQ387867</td>
</tr>
<tr>
<td>Methanococccoides burtonii DSM6242</td>
<td>MbATPase</td>
<td>869</td>
<td>4.97</td>
<td>75.6</td>
<td>YP_566814</td>
</tr>
<tr>
<td>Methanospirillum hungatei IF-1</td>
<td>MbATPase</td>
<td>865</td>
<td>5.47</td>
<td>51.0</td>
<td>YP_502751</td>
</tr>
<tr>
<td>Methanothermobacter thermoautotrophicus Delta H</td>
<td>MClpA/B</td>
<td>616</td>
<td>4.91</td>
<td>30.0</td>
<td>AAB84790</td>
</tr>
<tr>
<td>Bacillus cereus ATCC 10989</td>
<td>BcClpB</td>
<td>866</td>
<td>5.19</td>
<td>57.8</td>
<td>Q73BY1</td>
</tr>
<tr>
<td>Lactococcus lactis subsp. cremoris</td>
<td>LClpB</td>
<td>867</td>
<td>4.86</td>
<td>53.1</td>
<td>O68185</td>
</tr>
<tr>
<td><em>Thermus thermophilus</em> HB8</td>
<td>TCClpB</td>
<td>854</td>
<td>5.53</td>
<td>56.4</td>
<td>BAA81745</td>
</tr>
<tr>
<td>Corynebacterium glutamicum</td>
<td>CgClpB</td>
<td>852</td>
<td>4.81</td>
<td>52.5</td>
<td>P53532</td>
</tr>
<tr>
<td>Mycobacterium tuberculosis H37Rv</td>
<td>MClpB</td>
<td>848</td>
<td>5.04</td>
<td>53.4</td>
<td>P63288</td>
</tr>
<tr>
<td>Haemophilus influenzae Rd KW20</td>
<td>HClpB</td>
<td>856</td>
<td>5.69</td>
<td>53.4</td>
<td>P44403</td>
</tr>
<tr>
<td><em>Escherichia coli</em> K12</td>
<td>EcClpB</td>
<td>857</td>
<td>5.18</td>
<td>51.8</td>
<td>P63284</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em> X2180</td>
<td>ScHsp104</td>
<td>908</td>
<td>5.14</td>
<td>41.5</td>
<td>P31539</td>
</tr>
<tr>
<td><em>Trypanosoma brucei brucei</em></td>
<td>THSp100</td>
<td>868</td>
<td>6.16</td>
<td>46.4</td>
<td>P31543</td>
</tr>
<tr>
<td><em>Leishmania donovani</em></td>
<td>LdHsp100</td>
<td>867</td>
<td>7.43</td>
<td>47.6</td>
<td>CAB08073</td>
</tr>
<tr>
<td>Zea mays</td>
<td>ZmHsp101</td>
<td>912</td>
<td>5.76</td>
<td>49.5</td>
<td>AAD25223</td>
</tr>
<tr>
<td>Glycine max</td>
<td>GmHsp101</td>
<td>911</td>
<td>5.77</td>
<td>48.4</td>
<td>AAA66338</td>
</tr>
<tr>
<td><em>Arabidopsis thaliana</em> Columbia</td>
<td>AtHsp101</td>
<td>911</td>
<td>5.73</td>
<td>48.6</td>
<td>P42730</td>
</tr>
</tbody>
</table>

Analysis of the other 16 ClpB/Hsp100/Hsp101/Hsp104 proteins (Fig. 2) revealed that almost all of them could form a coiled-coil structure in the middle region, with one exception, *Met. thermoautotrophicus* MtClpA/B, which does not contain the middle region.

**Table 1. Percentage identity of *Meh. portucalensis* ClpB with bacterial ClpB and eukaryotic Hsp100/Hsp101/Hsp104**

![Fig. 2. Phylogenetic relationship of ClpB/Hsp100/Hsp101/Hsp104 proteins. Bootstrap values are shown at nodes (percentage of 1000 replicates). Bar, 0.1 aa substitutions per 100 aa.](http://mic.sgmjournals.org)
In addition to the coiled-coil structure in the middle region, our analysis of MpClpB showed another coiled-coil structure located in front of the middle region (the first peak in Fig. 3b). This first peak was also detected in MbaATPase, BcClpB, LcClpB, AtHsp101, GmHsp101, ZmHsp101, TbClpB and CgClpB. Moreover, a third peak behind the middle region was detected only in LcClpB (data not shown). Comparing the positions of the first peaks for coiled-coil structure of these Clp proteins, those from MpClpB, MbaATPase, BcClpB, AtHsp101 and TbClpB were located in the N-terminal region, while those from LcClpB, GmHsp101, ZmHsp101 and CgClpB were located in NBD1. The distinctly discontinuous coiled coils are formed in the middle region of some ClpB proteins such as MpClpB, MbATPase, BcClpB, TbClpB, CgClpB and ScHsp104; however, only MpClpB and MbaATPase generate three distinct segments. The Tyr653 residue of E. coli ClpB, which is potentially involved with the disaggregation process as a substrate-binding site or driving translocation, is conserved in MpClpB as Tyr657. A highly conserved serine residue, Ser409, within a putative phosphorylation domain adjacent to the predicted coiled-coil region (Celerin et al., 1998), was also found in MpClpB (Fig. 3a and Supplementary Figure S1).

Transcriptional stimulation of clpB by salt stress

The effect of salt stress on transcription of Meh. portucalensis clpB was investigated by Northern hybridization using clpB-2.6 as probe. Signals for 16S rRNA of Meh. portucalensis were used as an internal control. The relative transcript levels were calculated from the specific value of the clpB signal divided by that of the 16S rRNA signal. In cells grown on defined medium with 2.1 M NaCl, the optimal concentration for growth, a single band was detected (Fig. 4a). This indicated that the clpB gene was transcribed constitutively without salt shock and we used this transcript level as the basal transcription level to compare the effect of hypersaline or hyposaline stress. With an extended time frame (from 1 to 2 h), the transcription level of clpB was slightly increased. This is due to the dramatic changes in culture density, with an over twofold dilution of mid-exponential-phase cultures with fresh medium. The dilution stress as well as oxygen introduced during the addition of medium may stimulate the transcription of clpB. Under salt up-shock from 2.1 M NaCl to 3.1 M NaCl, the clpB transcript level was increased 1.57-fold within 1 h and 1.52-fold after 2 h (Fig. 4). Similarly, under salt down-shock from 2.1 M NaCl to...
0.9 M NaCl, the clpB transcript level was increased 1.03-fold with 1 h and 1.60-fold at 2 h (Fig. 4). The transcript level of clpB was increased at both salt up- and down-shock, indicating that the expression of the molecular chaperone clpB may be required for cell survival and adaptation while encountering salt stresses.

**Transcriptional stimulation of MpclpB and MpgroEL/ES by temperature stress**

Bacterial ClpB and its eukaryotic orthologues, plant Hsp101 and yeast Hsp104, are heat-shock induced and are essential proteins of the stress response (Queitsch et al., 2000; Sanchez & Lindquist, 1990; Squires et al., 1991; Lee et al., 2004; Lee & Tsai, 2005). To investigate whether this archaeal ClpB gene from a halophilic methanogen also responds to heat shock as do those of bacterial and eukaryotic cells, mid-exponential-phase cultures of *Meh. portucalensis* grown optimally at 37 °C were incubated at 45 °C and Northern hybridization was performed with clpB-2.6 as the probe. As shown in Fig. 5, the transcript level of clpB gradually increased and reached the maximal sevenfold increase within 60 min after heat shock. This result demonstrated that the *Meh. portucalensis* clpB is involved in heat-shock response as are the clpB genes from bacteria and eukaryotes. The transcription level of clpB in response to heat shock reached a maximum after 1 h then gradually decreased to 4.2-fold at 2 h. The effect of temperature (from 37 °C to 55 °C) on groEL/ES transcription was also tested with a culture of *Meh. portucalensis* grown at the optimal temperature (37 °C) as a control. The groEL and groES genes were co-transcribed as a 1.9 kb transcript (Fig. 6a). The transcription levels of groEL/ES were increased upon increasing the temperature, by 2.2-fold at 45 °C and by 2.6-fold at 55 °C (Fig. 6a). This result indicated that, as with *Meh. portucalensis* clpB (Fig. 6b), groEL/ES also respond to heat shock.

The effects of temperature (from 15 °C to 55 °C) on clpB transcription were also tested with a culture of *Meh.
portucalensis grown at the optimal temperature 37 °C as control. The transcript levels of clpB were increased by raising the temperature: by fivefold at 45 °C and by 10-fold at 55 °C (Fig. 5d). Under cold shock, clpB transcript levels increased fourfold at around 30 °C but there were no obvious effects at 25 °C or below (data not shown). These results suggest that transcription of clpB gene was stimulated under both heat and cold shock for Meh. portucalensis, but not at the optimal temperature range (37–40 °C), or temperatures below 25 °C.

Addition of the osmolyte betaine rescues the temperature stress effect and reduces clpB transcription

The halophilic methanoarchaeon Meh. portucalensis strain FDF1 can transport the osmolyte betaine through a high-affinity specific betaine transporter (Lai et al., 2000). The effect of betaine on clpB transcription was tested by Northern hybridization. During heat stress, the clpB transcript level decreased 2.5-, 0.6- and 1.4-fold at 30 °C, 37 °C and 55 °C, respectively, after the addition of betaine (1 mM) (Fig. 5e). These results demonstrate that the addition of betaine could rescue the temperature-stress effect and reduce clpB transcription.

DISCUSSION

The complete Meh. portucalensis clpB was cloned, sequenced and analysed. The deduced amino acid sequence of clpB shares high identity with the known molecular chaperone ClpBs. Although stress-response genes and molecular chaperones among Archaea have long been discussed (Albers et al., 2000; Cannio et al., 2000; Macario et al., 1999, 2004; Maruyama & Furutani, 2000; Roberts, 2000; Scandurra et al., 2000), archaeal ClpBs have never been defined and published. Three archaeal ClpA/B homologues were identified through bioinformatic surveys, Mes. hungatei YP_50 2751, Mec. burtonii YP_566814 (Saunders et al., 2003) and the closest ClpA/B-type homologue of Met. thermoautotrophicus AAB84790 (Neuwald et al., 1999; Laksanalamai et al., 2004). The length of the thermophilic archaeon Met. thermoautotrophicus ClpA/B homologue, 616 aa, is shorter than all known ClpB/Hsp100 proteins and it possesses only one NBD, which does not conform to the typical ClpB features. According to the amino acid sequence identity comparison, the ClpA/B homologue of Met. thermoautotrophicus is distant from MpClpB and the known ClpB/Hsp100 proteins. In this study, we describe preliminary research on the archaeal clpB from the halophilic methanogen Meh. portucalensis strain FDF1. Based on the preliminary search for phylogenetic relationships, MpClpB and MbaATPase were clustered with homologues from bacteria and were closely related to those from the low-G+C Gram-positive bacteria, B. cereus and Lactococcus lactis, but not with those from the high-G+C Gram-positive organisms C. glutamicum and Mycobacterium tuberculosis (Fig. 2).

Amino acid sequence analysis revealed that MpClpB possesses the characteristics of the molecular chaperone ClpB, with two conserved nucleotide-binding domains, NBD1 and NBD2, which are separated by a long middle region that was predicted to form a coiled-coil structure. There were 6.5 % amino acid substitutions in the NBD1 and NBD2 of MpClpB compared with BcClpB. The substrate-binding site of EcClpB was reported to be located at the central pore of NBD1 and the conserved Tyr251 of EcClpB directly binds to the oligopeptides (Schlieker et al., 2004). Based on the amino acid alignment, a hydrophobic amino acid, Phe253, is suggested as the substrate-binding site in MpClpB, MbaATPase and BcClpB (Supplementary Figure S1). The middle region of MpClpB (124 residues) represents a predicted long coiled-coil structure. Although the overall identity among pairwise comparisons of the ClpB/Hsp100/Hsp101/Hsp104 sequences from different organisms is high, the percentage identity of pairwise comparisons within the predicted coiled-coil region is surprisingly low (data not shown), given the amount of conservation at the secondary structural level. The protein structure of MpClpB was predicted by using the SWISS-MODEL web server (http://swissmodel.expasy.org) and the comparison of the predicted MpClpB 3D structure with the known ClpB 3D structure from Thermus thermophilus (Lee et al., 2003) indicated that the 3D structure of this domain,
but not the sequence per se, is conserved. The long coiled-coil structure in the middle region in MpClpB contains three fragments of periodic seven-residue repeats (Fig. 3b), which is different from the continuous-type periodic seven-residue repeats among ClpB of bacteria and eukaryotes. This particular coiled-coil structure of archaeal ClpB needs further study.

The group I chaperonin genes groEL and groES occur in the genomes of Methanosarcina species and Mes. hungatei but they were not found in any of the other archaea whose genomes have been sequenced (Deppenmeier et al., 2002; Klunker et al., 2003; Macario et al., 2004). In this study, analysis of the upstream sequence of MpclpB revealed a complete GroEL gene (1611 bp) and a complete GroES gene (276 bp), whereas a complete CDP-diacylglycerol-serine O-phosphatidyltransferase gene (684 bp) and a partial phosphatidylserine decarboxylase gene were located downstream of MpclpB. This is the first finding showing that the group I chaperonin GroEL and GroES genes and the chaperone ClpB gene coexist in Archaea. Also, the gene order with groEL located in front of clpB is novel and has never been found in any other organisms (Fig. 1). In The. thermophilus, clpB is located behind the dnaK operon (Osipiuk & Joachimiak, 1997) and they form a chaperone network (Ben-Zvi & Goloubinoff, 2001; Zolkiewski, 1999).

Under heat shock at 45 and 55 °C, the transcript level of groEL is increased 1.55-fold in salt down-shock cultures (with 2.1 M NaCl). In contrast, the transcript level of MphpclpB was increased 1.55-fold in salt down-shock cultures (with 0.9 M NaCl) by the addition of betaine (data not shown). This result suggests that the addition of betaine at salt down-shock may actually be harmful and the transcription of clpB increased to rescue the effect. For temperature effects, the addition of 1 mM betaine slightly repressed the transcription of clpB at the tested temperatures (30, 37 and 55 °C). The interplay among osmolytes, chaperone networks and stresses is of interest for further investigation.

REFERENCES


Edited by: J. van der Oost