Archaeal grpE: Transcription in Two Different Morphologic Stages of Methanosarcina mazei and Comparison with dnaK and dnaJ

EVERLY CONWAY DE MACARIO, MANUEL CLARENS, and ALBERTO J. L. MACARIO
Wadsworth Center for Laboratories and Research, New York State Department of Health, and Department of Biomedical Sciences, School of Public Health, The University at Albany, Albany, New York

Received 15 August 1994/Accepted 24 November 1994

Transcription of the heat shock gene grpE was studied in two different morphologic stages of the archaeon Methanosarcina mazei S-6 that differ in resistance to physical and chemical traumas: single cells and packets. While single cells are directly exposed to environmental changes, such as temperature elevations, cells in packets are surrounded by intercellular and peripheral material that keeps them together in a globular structure which can reach several millimeters in diameter. grpE transcript levels determined by Northern (RNA) blotting peaked after a 15-min heat shock in single cells. In contrast, the highest transcript levels in packets were observed after the longest heat shock tested, 60 min. The same response profiles were demonstrated by primer extension experiments and S1 nuclease analysis. A comparison of the grpE response to heat shock with those of dnaK and dnaJ showed that the grpE transcript level was the most increased, closely followed by that of the dnaK transcript, with that of the dnaJ gene being the least augmented. Transcription of grpE started at the same site under normal and heat shock temperatures, and the transcript was consistently ~700 bases long. Codon usage patterns revealed that the three archaeal genes use most codons and have the same codon preference for 61% of the amino acids.

A gene encoding a GrpE homolog has recently been cloned from the genome of the archaeon Methanosarcina mazei S-6 and sequenced (6). It is the first example of a grpE heat shock gene found within the domain Archaea. The Archaea are organisms phylogenetically distinct from Bacteria (eubacteria) and Eucarya (eucaryotes) (41). However, Archaea resemble eucaryotes in several molecular biological features more than they resemble bacteria and are closer to eucaryotes than bacteria (15, 28, 30, 31, 41). A case in point is gene regulation, since there are data indicating that Archaea use a transcription regulation machinery similar to that of eucaryotes (10, 24, 28, 33). Whether this is the case also for archaeal heat shock genes has not been established. To address this question, it is necessary to first determine the expression pattern of the known archaeal heat shock genes and then proceed to analyze the regulation mechanisms at the molecular level.

Among the Archaea, M. mazei is particularly interesting because it is one of a small group of organisms, the methanosaecinae (22, 23), that undergo a growth cycle with morphologic conversions (3, 25, 32). The cycle includes a unicellular form or stage, named single cells, and two multicellular forms, called lamina and packets. The latter are globular structures with a zonal heterogeneity and intercellular connective material that extends the outer surface of the packets and wraps the cells together. This connective material is composed chiefly of a heteropolysaccharide (14, 16) which confers to the packets resistance to traumas of various types: mechanical, physical (e.g., heat), and chemical (unpublished observations). This resistance is in contrast to the relative fragility of lamina (25) and especially to the fragility of single cells (9). Conversion of one form into another is induced by environmental changes, among which changes in catabolic substrate and divalent cations are the best studied (3, 5, 12, 22, 23, 25, 35, 42).

May be postulated that packets are a form of resistance phenotype which are shielded from the environment by the intercellular and peripheral material and by the overall three-dimensional structure, with a central zone in which the cells are covered by layers of more-peripheral cells. In contrast, single cells are more directly exposed to the environment and thus are more dependent on a prompt protective reaction mounted by heat shock or stress genes, whose role in the cell’s response to stressors and survival is well documented (8, 29, 36, 44). In this work, we have compared the expression of grpE with that of the adjacent heat shock genes dnaK and dnaJ (19, 20) in response to heat shocks of various durations in single cells and packets.

MATERIALS AND METHODS

Cultures and media. M. mazei S-6 was grown in 100-ml serum bottles containing 50 ml of S6-2 alpha medium at 37°C and 1-atm (1 atm = 101.29 kPa) N2-CO2 (80:20) (3, 5, 12, 22, 23). For packets, the medium also contained 250 mM methanol, while 4.9 mM Mg2+, 27 mM Ca2+, and 40 mM trimethylamine were added for single cells (3, 25).

S6-2 alpha medium was composed of 0.4 g of KH2PO4 per liter, 0.1 g of MgCl2 · 6H2O per liter, 0.04 g of CaCl2 · 2H2O per liter, 1.0 g of NH4Cl per liter, 0.25 g of L-cysteine HCl per liter, 0.25 g of Na2S · 9H2O per liter, 2.0 g of yeast extract per liter, 2.0 g of Trypticase per liter, 3.8 g of NaHCO3 per liter, 10 ml of trace mineral solution, and 1.0 ml of 0.1% resazurin. The pH was 7.0. The trace mineral solution contained the following (in grams per liter): nitritotriacetic acid, 1.5; MgCl2 · 6H2O, 3.0; MnCl2 · 4H2O, 0.5; NaCl, 1; FeCl3 · 6H2O, 0.1; CoCl2 · 6H2O, 0.1; CuCl2 · 2H2O, 0.01; ZnCl2, 0.1; AlCl3, 0.01; Na2MoO4 · 2H2O, 0.01; and NiCl2 · 6H2O, 0.02.

Cultures for heat shock experiments were placed in a 45°C water bath and incubated with gentle mixing every 5 min.

RNA isolation. Samples (25 ml) of single-cell or packet cultures were pelleted by centrifugation (600 × g) at 4°C for 10 min and resuspended in 1.0 ml of denaturing solution containing 4.0 M guanidine isothiocyanate, 0.75 M sodium citrate, 10% N-lauryl sarcosine, and 0.1% 2-mercaptoethanol (1, 34). In addition, for packets, four sonications of 10 s each were done to obtain proper cell
ing was carried out by the dideoxynucleotide-chain termination method with α-32P-dATP (1, 34).

Mapping of the transcription initiation site by S1 nuclease analysis. Mapping the 5′ end of the grpE transcript was carried out by S1 nuclease analysis (1, 34). Twenty picomoles of a 30-mer synthetic oligonucleotide encompassing bases 49 to 78 downstream of the ATG start codon (Fig. 1) was 5′ end labelled with T4 polynucleotide kinase (New England Biolabs, Beverly, Mass.). This DNA was then annealed at 40°C for 15 min to the original 11-mer double-stranded-DNA clone that was extended with DNA polymerase (Klenow fragment; New England Biolabs) for 30 min at 37°C. This product was then digested with the enzyme ScaI (New England Biolabs), which cut 161 bases upstream of the grpE ATG start codon, yielding a fragment of 239 bases (Fig. 1). The probe was precipitated and rehydrated under alkaline conditions and run in a low-agarose denaturing gel containing 50 mM NaOH and 0.1 mM EDTA (pH 8.0) for 2.5 h at 30 V. The band containing the probe was excised from the gel and purified by standard phenol extraction. Radiolabelled, single-stranded probe was hybridized to 25 µg of total RNA in 80% formamide–40 µM piperezine-N,N′-bis(2-ethanesulfonic acid) (PIPES) (pH 6.4)–0.4 M NaCl–1 mM EDTA (pH 8.0) at 30°C for 18 h. The remaining single-stranded fragments were digested with S1 nuclease (Promega) for 60 min at 30°C. After ethanol precipitation, samples were resuspended in 3.0 µl of 1× TE buffer (10 mM Tris-Cl, 1 mM EDTA; pH 7.4)–3.0 µl of loading dye containing 98% formamide, 10 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue. The entire 6.0-µl sample was loaded and electrophoresed in a 6% polyacrylamide–7 M urea sequencing gel for 1 h at 80 W. The gel was exposed to X-Omat AR film (Eastman Kodak Company, Rochester, N.Y.) for 16 h and developed.

Mapping of the 3′ end of the grpE transcript. The termination site was mapped by S1 nuclease analysis (1). A synthetic 62-mer oligonucleotide complementary to the 3′ end of the mRNA was used (Fig. 1). The oligonucleotide was unlabelled and its 5′ end was labelled with T4 polynucleotide kinase (New England Biolabs) for 30 min at 37°C. Hybridization was done with 0.3 µg of labelled probe at 37°C for 18 h, using 25 µg of total RNA from cells grown at 37°C and from cells heated at 45°C for 60 min. The preparation was treated with 300 U of S1 nuclease (Promega) at 37°C for 15 min. The protected fragments were resolved by electrophoresis on a 10% polyacrylamide–7 M urea gel.

Dot blotting and phosphor image analysis. Dot blotting of the RNA was performed according to standard procedures (1, 34). Briefly, 10 µg of total RNA per sample was denatured for 15 min at 65°C in formamide hybridization solution, diluted to a final concentration of 10× SSC solution, and then dot blotted by filtration through an N-Hybond nylon membrane (Amersham) using a manifold apparatus with a filtration template for dot blots (Minifold I; Schleicher & Schuell, Keene, N.H.). The nylon membrane was washed at 37°C and cross-linked via irradiation at a 234-nm wavelength for 2 min in a UV fluorescence analysis cabinet (Spectronic CX-20). Hybridization was done by adding random-primed 32P-labelled PCR-generated probes (1.10 × 106 cpm/ml). For grpE, the probe prepared for the Northern blotting experiments was used (Fig. 1). For dnaK transcripts, a 1.448-kb PCR-generated probe was prepared by using two 18-mers, one from bases 12 to 29 and the other from bases 143 to 1460 on the dnaK gene. For dnaJ transcripts, an 882-bp PCR-generated probe was prepared by using two 18-mers, bases 25 through 42 and 890 through 907 on the dnaJ gene. The hybridization procedure was done in the same manner as for Northern blotting. The hybridization was done successively with the radiolabelled grpE, dnaK, and dnaJ probes with intervening stripings. Striping after each hybridization was done according to the manufacturer’s instructions (Amersham). Briefly, the membranes were boiled for 15 min in 0.1% SDS solution and then washed to ensure that the probe had been completely removed. Computation of the intensity of the dot blots was done with a PhosphorImager apparatus (model 425; Molecular Dynamics, Sunnyvale, Calif.) (13).

Codon usage and amino acid sequence analyses. Comparative analyses of codon usage and amino acid sequences were done as previously described (6).

Nucleotide sequence accession number. The nucleotide sequences of the grpE gene and its flanking regions have been submitted to EMBL-GenBank under accession no. X74553.

RESULTS

grpE transcript in the two M. mazei S-6 morphologic stages, single cells and packets. Nine independent experiments with RNA from single cells and a probe for grpE (Fig. 1) showed a positive hybridization band corresponding to a monocistronic transcript. The range of values was between 650 bases (once) and 750 bases (twice), with the most frequent result being 700 bases (four experiments, the results of one of which are shown in Fig. 2, top panel). Two independent experiments with packets RNA gave values of 800 and 700 bases (Fig. 2, bottom panel). This hybridization band was barely detectable in RNA from single cells and packets grown at the optimal growth
temperature of 37°C and increased after heat shock at 45°C. The greatest increase in single cells was caused by a 15-min heat shock (compared with 30 and 60 min) (Fig. 2, top panel). In contrast, the transcript level in packets peaked after a 60-min heat shock (Fig. 2, bottom panel).

Transcription initiation site. To map the initiation site of the transcript identified by Northern blotting, primer extension experiments using a primer complementary to nucleotides 57 through 77 in the grpE coding region were performed (Fig. 1). The same site, a C located 44 bases upstream of the translation start codon, was found for single cells and packets (Fig. 3). The results also showed that there was a low level of transcription in non-heat-shocked single cells and that the level increased after heat shock. This increase was higher after a heat shock of 30 min than after a heat shock of 60 min. In packets, the transcript was clearly present only after a heat shock of 60 min. These temporal patterns paralleled those observed by Northern blotting.

To corroborate the results of the primer extension experiments, S1 nuclease analysis was applied, using the probe shown in Fig. 1. By this method, the transcription initiation site was located at an A 4 bases downstream of the C identified as the initiation site by primer extension (Fig. 4 and 5).

Transcription termination site. S1 analysis showed one major and two lighter bands (Fig. 6) corresponding to termination sites located on bases 42, 36, and 21, counting from the third base (not included) of the grpE translation stop codon (Fig. 5).

grpE response to heat shock in comparison with the dnaK and dnaJ responses. To compare the responses of the three genes to heat shocks of various durations, quantitative dot blotting was used along with hybridization measurements by

### Figures

**FIG. 2.** Northern blotting analysis. Total RNA (10 μg per lane) from single cells (top panel) or packets (bottom panel) grown at 37°C (lane A) or heat shocked at 45°C for 15, 30, and 60 min (lanes B through D, respectively) was electrophoresed in a 1.2% agarose-formaldehyde gel. Hybridization was done with a radiolabelled PCR-generated probe of 679 bases (Fig. 1). The approximate transcript size (in kilobases) is indicated on the right.

**FIG. 3.** Mapping of the transcription initiation site by primer extension. A radiolabelled oligonucleotide primer complementary to bases 57 through 77 within the grpE coding region (Fig. 1) was used with 10 μg of total RNA from single cells (lanes 1 to 3) or packets (lanes 4 to 6) per test. Single cells and packets had been grown at 37°C (lanes 1 and 4) or heat shocked at 45°C for 30 (lanes 2 and 5) or 60 (lanes 3 and 6) min. The primer-extended products were electrophoresed in a 6% acrylamide sequencing gel in parallel with the products of a sequencing reaction that was done with the same primer and the dideoxy-chain termination method (lanes G, A, T, and C). These lanes show the complementary (antisense) strand sequence. The coding (sense) strand sequence and the initiation site (asterisk) are shown on the left.

**FIG. 4.** Mapping of the transcription initiation site by S1 nuclease analysis. Lanes: A, ladder, with marker sizes shown on the left; B and C, protected products obtained with the probe shown in Fig. 1 and total RNA from single cells grown at 37°C or from cells heat shocked at 45°C for 30 min, respectively. The length of the major protected product is 118 bases, starting on the grpE coding region (78 bp downstream of the translation start codon’s first base), which maps the initiation site at the A located 40 bp upstream of the translation start codon (Fig. 5). The other, minor band corresponds to a protected product that would map the site 4 bases downstream.

**FIG. 5.** Nucleotide sequence of the 5’- and 3’-flanking regions of grpE. The transcription initiation sites demonstrated by primer extension (bases −44 and by S1 nuclease analysis (base −40) are indicated (open and closed triangle, respectively). Also shown are the putative promoter (bases −76 to −64) and ribosome-binding site (bases −14 to −9) (underlined), with bases that are identical to those in the respective consensus sequence for methanogens (asterisks) and positions within the promoter at which there is no base preference (4) (vertical lines) indicated, and the translation start (overlined) and stop (underlined) codons and the transcription termination sites determined by S1 nuclease analysis (arrows).
means of storage phosphor technology. RNA was extracted from non-heat-shocked packets grown at 37°C and also from packets which had been heat shocked at 45°C for 15, 30, and 60 min. The results of two independent experiments are displayed in Fig. 7. The transcript levels of the three genes augmented in parallel with the length of the heat shock. However, the increase of the \textit{grpE} and \textit{dnaK} transcript levels was faster than that of the \textit{dnaJ} transcript; the increase in the \textit{dnaJ} transcript level was clear only after a heat shock of 30 min. At this time, the factor of increase for the \textit{dnaJ} transcript was approximately 2.5. In contrast, a similar factor of increase for the \textit{dnaK} and \textit{grpE} transcripts was already reached after a heat shock of 15 min.

**Codon usage analysis.** The codon usage pattern of the \textit{M. mazei} \textit{grpE} gene is displayed in Fig. 8, along with the patterns of the other two studied genes of the \textit{dnaK} locus of this archaeon. The three genes show the same codon preference for 11 of 18 amino acids (61%); Trp and Met were excluded because the former is not present in these genes’ products and Met has only one possible codon. The same preferences were observed when \textit{grpE} and \textit{dnaK} were compared, and the agreement increased to 72 and 78% for \textit{dnaK-dnaJ} and \textit{grpE-dnaJ}, respectively. The three genes use all codons with only five exceptions for \textit{grpE} and \textit{dnaJ} and six exceptions for \textit{dnaK} (which is considerably longer than the other two genes); CUA (Leu) and UGG (Trp) are not used by any of these genes; and 8 of 16 cases (50%) of missing codons involve a codon with A in the third position, while 6 of 16 (38%) involve a codon with G in this position (the other two missing codons end in U or C).

**DISCUSSION**

The \textit{grpE} transcript in \textit{M. mazei} S-6 at the optimal growth temperature as well as after heat shock was found to be \(\sim 700\) bases long, which corresponds to a monocistronic message. This was demonstrated by Northern blotting and by mapping the transcription initiation and termination sites.

Transcription, as demonstrated by primer extension, would start at a pyrimidine-pyrimidine (TC) dinucleotide. We examined 47 transcription initiation sites reported in the literature for 30 non-heat shock genes in 11 different methanogens and found a pyrimidine-pyrimidine dinucleotide 13 times (the most frequently found dinucleotide, occurring 17 times, was a pyrimidine-purine dinucleotide). A TC dinucleotide like that of the \textit{M. mazei} \textit{grpE} was found twice, the most frequent dinucleotides being TT (eight times), TG (seven times), and AT (six times). The information available for the \textit{grpE} gene is limited to bacteria, and the dinucleotides found at the transcription initiation sites are CA (pyrimidine-purine) for \textit{Bacillus subtilis} (40) and AA (purine-purine) for \textit{Escherichia coli} (18).

The transcription initiation sites determined by primer extension for heat-shocked and non-heat-shocked cells are the same, which is reminiscent of what has been reported for \textit{E. coli} \textit{grpE} (18). The site determined by S1 nuclease analysis is close to that determined by primer extension, within the expected variations of the methods. The sizes of the \textit{grpE} transcripts detected by Northern blotting and the transcription initiation sites are the same in single cells and packets. How-
FIG. 8. Codon usage patterns for *M. mazei* grpE, dnaK, and dnaJ. The fractions of codons with G or C (open bars) and A or U (solid bars) in the third position are indicated. The codons possible for each amino acid are indicated along the horizontal axis. Single letters at the bottom indicate amino acids, which are in alphabetical order for the amino acids with two, four, or six possible codons (from left to right).
ever, in packets, a longer heat shock is necessary to induce levels of gene expression comparable to those observed in single cells. Furthermore, the dnaK transcript increase after heat shock in single cells was twice as great as that in packets (data not shown). The mechanism responsible for the differences between single cells and packets remains to be elucidated. Intrinsic properties, distinctive of the cells of the single-cell stage and the cells in the packets, might contribute to the differences observed between the two morphologies in the promptness and levels of expression of grpE and dnaK. Such distinctive mechanisms are suggested by the observation that dnaJ expression levels are similar in packets (Fig. 7) and single cells (unpublished data), while expression of grpE and dnaK is more pronounced in single cells. Also, the possibility that cells in the center of the packets are less accessible to heat ought to be explored. It can be postulated that single cells, without a protective envelope, have the dnaK and grpE genes ready to respond quickly at all times, on the alert, as it were. In contrast, packets would have these genes down-regulated to preclude rapid activation. In this way, cells in packets would avoid diversion of energy and nutrients by heat shock genes from the pools required for other functions, such as cell division and synthesis of the extracellular protective material.

The data suggest the organization depicted in Fig. 5 for the grpE gene of M. mazei S-6. A putative promoter is centered around a base that is 26 bp upstream of the transcription initiation site determined by primer extension, or 31 bp in front of the initiation site determined by S1 nuclease analysis. This promoter sequence resembles the consensus sequence generated for promoters in methanogens (4). No sequences resembling bacterial heat shock promoters (7) could be identified. A ribosome-binding sequence, also similar to the corresponding consensus obtained for methanogens (4), is located 8 bp upstream of the translation start codon. The coding region would encode a protein whose amino acid sequence has identities and conservative substitutions with regard to bacterial heat shock proteins that fall within the range observed among the latter (6) and between them and a recently described homolog from the mitochondria of Saccharomyces cerevisiae (2, 11, 17) (data not shown). The overall organization of the grpE chromosomal region in M. mazei is similar to those of B. subtilis (39, 40), Clostridium acetobutylicum (26), Borrelia burgdorferi (38), and Staphylococcus aureus (27). In these bacteria and in M. mazei, grpE is clustered with dnaK and dnaJ, located upstream of the former gene and preceded by another gene whose role is still incompletely understood. The archaeal dnaK locus, however, has distinctive characteristics: it is followed in the 3' direction by a trkA homolog (21), and the distance between grpE and the nearest gene on either flanking region is longer than the bacterial equivalents. In the bacterial clusters, the length of the intergenic region between grpE and dnaK ranges between 21 and 68 bp, while in M. mazei it is 431 bp. Likewise, the separation between M. mazei grpE and its preceding gene (orf147 [6a]) is 127 bp, whereas the equivalent regions in B. subtilis, S. aureus, C. acetobutylicum, and Chlamydia trachomatis (35) are 72, 32, 21, and 0 bp long, respectively.

Codon usage patterns revealed that the three archaean heat shock genes use practically all codons with few exceptions, which is in contrast to the homolog genes from the other two phylogenetic domains, Bacteria and Eucarya. The three archaean genes share the same codon preference for 61% of the amino acids. The agreement is even greater between dnaK, dnaJ (72%) and grpE-dnaJ (78%). These percentages are remarkable, as they contrast with those calculated for the agreements between the three M. mazei genes and pools of E. coli and C. trachomatis genes (35). The archaean genes prefer the same codons as E. coli and C. trachomatis genes for 39 and 17% of the amino acids, respectively. The agreement between the two bacterial gene pools is 33%, that between the three M. mazei heat shock genes and the trkA gene is 39%, and that between the heat shock genes and an S-layer gene recently described for this archaeon (43) is 28%. The significance of the codon usage patterns and the high degree of agreement in codon preferences distinctive of the three archaean heat shock genes in translation regulation, for example, remains to be determined.

ACKNOWLEDGMENTS

We thank Charles B. Dugan, Victor H. Simon, James S. Swab, and Fotios Koumpouras for assistance; Tim Moran, from the Molecular Genetics Core Facility of Wadsworth Center for Laboratories and Research, for oligonucleotides; the personnel of the Photography Unit for their art; Tracy Godfrey for word processing; and Charles J. Daniels for critical reading of the manuscript. This work was supported in part by grant 706-RIER-BEA-85 from NYSERDA and by the Lavoisier fellowship from the French government to M.C.

REFERENCES