JOURNAL OF BACTERIOLOGY, May 1996, p. 2994–2997 Vol. 178, No. 10
0021-9193/96/$04.00+0

A Novel Member of the *cspA* Family of Genes That Is Induced by Cold Shock in *Escherichia coli*

KYOKO NAKASHIMA,[1]* KENGO KANAMARU,[2] TAKESHI MIZUNO,[3] AND KOKI HORIKOSHI[1]

Deep-Sea Microorganisms Research Group, Japan Marine Science and Technology Center, Yokosuka 237,[1] National Institute of Genetics, Genetic Stock Research Center, Mishima, Shizuoka-ken 411,[2] and Laboratory of Molecular Microbiology, School of Agriculture, Nagoya University, Chikusa-ku, Nagoya 464,[3] Japan

Received 30 October 1995/Accepted 12 March 1996

Escherichia coli **contains a major cold shock protein, CspA (or CS7.4), whose production is predominantly induced at low temperatures. This bacterium is known to possess five additional genes, each encoding a protein highly similar to CspA (referred to as the CspA family). Here we identified a gene that encodes a cold-shock-inducible analog of CspA and CspB. This newly cloned *cspG* gene is located at 22 min on the *E. coli* genetic map, apart from the other *cspA* family genes. Its gene product (70 amino acids) is 73 and 77% identical to CspA (70 amino acids) and CspB (71 amino acids), respectively. Analyses of a *cspG-lacZ* transcriptional fusion and Northern (RNA) hybridization revealed that *cspG* is a low-temperature-responsive gene. Its low-temperature-inducible promoters were determined, and the results indicated that the *cspG* sequence is highly similar to both the *cspA* and *cspB* sequences not only in the coding regions but also in the 5′-upstream noncoding regions surrounding their own promoters.**

In *Escherichia coli*, temperature shifts from the normal range (20 to 37°C) to a temperature of above 40°C or below 20°C elicit pronounced physiological changes in growing cells. The best-characterized one is the so-called heat shock response (3). The heat shock proteins have become central to the study of the correct folding of nascent and/or unfolded polypeptides (1). Studies have also been initiated to elucidate the effects of a downshift in growth temperature on *E. coli* physiology. After the downshift, like heat shock proteins, a set of cold shock proteins (CSPs) were found to be produced at rates higher than those at the normal temperatures (7). They include several known proteins, namely, NusA, IF2, polynucleotide phosphorylase, RecA, H-NS, and the α subunit of DNA gyrase. Beside these CSPs, recently much attention has focused on the newly discovered CspA family of proteins (5).

The CspA protein was originally discovered as a major CSP (also called CS7.4), which was transiently induced for up to 13% of total cellular protein synthesis upon the downshift from 37 to 10°C (2, 7). Five other members of the CspA family in *E. coli* have been reported so far (5). It should be emphasized, however, that only CspA and CspB are cold shock inducible; the others are not (10, 18). Interestingly, the proteins of the CspA family resemble significantly the cold shock domain of eukaryotic Y-box (or CCAAT motif)-binding transcription factors (e.g., YB1 of humans) (17). Nevertheless, no clear picture of the physiological function(s) of the CspA family has yet emerged. Here we provide evidence that the *cspA* family genes are more redundant in *E. coli* than previously thought.

Isolation of a gene whose expression is induced at low temperatures for growth. The promoter-cloning vector pJACK4 has the kanamycin-resistant (*neo*) and ampicillin-resistant (*ampC*) genes, but the *ampC* gene lacks its own promoter, thereby cells carrying this plasmid are sensitive to ampicillin (4). A mixture of *Eco*RI-digested chromosomal DNA segments from *E. coli* CSH26 was randomly inserted into the unique *Eco*RI site immediately upstream of the promoterless *ampC* gene (Fig. 1). We screened transformants that grew on plates containing ampicillin (50 μg/ml) at 15°C but failed to do so at 37°C. These transformants were suspected to carry a pJAC4 derivative which most likely contains a chromosomal DNA segment that exhibits an ability to transcribe into the *ampC* gene only at low temperatures. Among the few candidates isolated, we characterized in detail one in which a 3.8-kb *Eco*RI fragment was carried at the expected site of pJAC4 (plasmid pJACK001) (Fig. 1). Then, a larger chromosomal DNA segment encompassing the gene preceding *ampC* in pJACK001 was cloned from the Kohara λ phage library by using a portion of the *Eco*RI fragment as a probe. Two positive contiguous clones, λ4H11 and λ2F1, at around 22 min on the genetic map (8) were identified. A 4.8-kb *Eco*RV fragment was then cloned from λ2F1 on a plasmid, since this fragment was expected to carry the *Eco*RI site in question (Fig. 1). The nucleotide sequence was then determined for the region around the *Eco*RI site (Fig. 2A).

Finding a novel member of the CspA family of proteins. The nucleotide sequence revealed that the *Eco*RI site is within an open reading frame (Fig. 2A) which is oriented in the same direction as that of the *ampC* gene on the original plasmid, as expected (Fig. 1). A low-temperature-inducible promoter was thus presumed to be located in front of this open reading frame. A computer-aided search revealed that its deduced amino acid sequence highly similar to the members of the CspA family of proteins (Fig. 2B). So far, six genes have been reported for this family, namely, *cspA* (79 min), *cspB* (35 min), *cspC* (40 min), *cspD* (19 min), *cspE* (14 min), and *cspF* (map position unknown) (5, 18). Their gene products are relatively small (69 to 74 amino acids) and are highly similar to each other. An alignment of those amino acid sequences with the one newly determined in this study clearly indicated that the clone encodes a member of the CspA family of proteins, as far as the primary structure is concerned. This gene was thus named *cspG*, and its gene product comprising 70 amino acid residues is 73 and 77% identical to CspA and CspB of *E. coli*, respectively.

* Corresponding author. Mailing address: c/o Dr. Takeshi Mizuno, Laboratory of Molecular Microbiology, School of Agriculture, Nagoya University, Chikusa-ku, Nagoya 464, Japan. Phone: (81)-52-789-4089. Fax: (81)-52-789-4091.

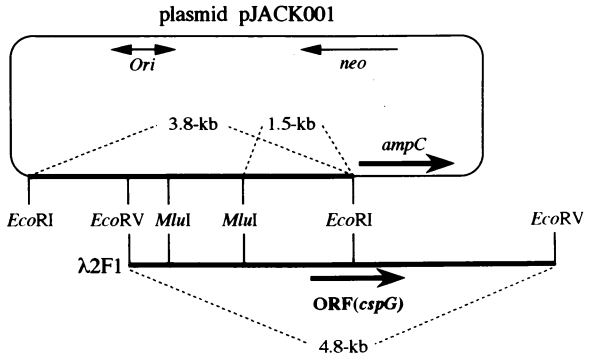

FIG. 1. Schematic representation of the structure of the recombinant plasmid pJACK001. The original plasmid, pJACK4, used as a cloning vector, contains the promoterless *ampC* gene (4). The 3.8-kb *Eco*RI fragment was cloned into the unique *Eco*RI site immediately upstream of the *ampC* gene on pJACK4. The 4.8-kb *Eco*RV fragment was then isolated from λ2F1 of the Kohara λ phage library. On pJACK001, the N-terminal (or 5′) half of an open reading frame (*cspG*) was revealed to be located in front of the *ampC* gene (Fig. 2).

Expression of *cspG* is induced in response to cold shock. Considering our rationale for the cloning of *cspG*, its expression appears to be induced at low temperatures. However, it was crucial to ask if the expression of *cspG* is indeed induced in response to a downshift in temperature. Thus, a 1.5-kb *Mlu*I-*Eco*RI fragment presumably containing a promoter of *cspG* (Fig. 1) was connected to a promoterless *lacZ* gene on a plasmid (14). The resultant plasmid (pCBK001) was introduced into CSH26, which was cultivated in Luria broth at 37°C. At this temperature, the steady-state level of expression of the *cspG-lacZ* gene was very low, as monitored by β-galactosidase activity (Fig. 3A). Upon the downshift from 37 to 15°C at the logarithmic-growth phase, however, the β-galactosidase activity increased rapidly, suggesting that the *cspG-lacZ* transcription was triggered in response to the temperature change. Furthermore, it was observed that when cells were grown at various temperatures, the steady-state levels of β-galactosidase activities were higher when cells were grown at lower temperatures (Fig. 3B). We then carried out Northern (RNA) hybridization analysis of the *cspG* mRNA, with an appropriate synthetic oligonucleotide as a probe (Fig. 2). *E. coli* cells were grown at 37°C to the mid-logarithmic-growth phase and then incubated further at 15°C. Total RNA fractions were isolated at intervals and then subjected to Northern hybridization analysis (Fig. 3C). The results showed that the *cspG* mRNA accumulated to a maximal level after 1 h in response to cold shock and then decreased to a basal level afterwards. From these results, we conclude that *cspG* is a low-temperature-inducible gene and that the accumulation of the *cspG* mRNA in response to cold shock appears to be transient.

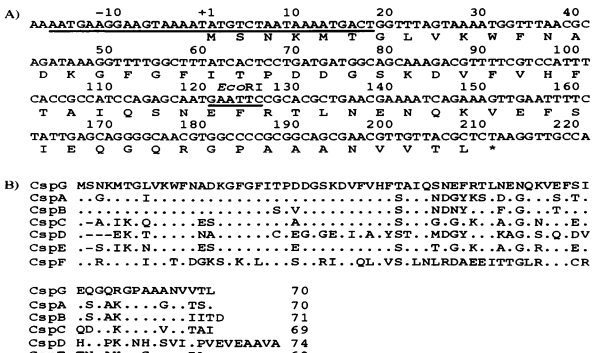

FIG. 2. The determined nucleotide and deduced amino acid sequences of *cspG* and comparison of its amino acid sequence with those of the CspA family of proteins. (A) About 1 kb of nucleotide sequence encompassing the *cspG* gene was determined (EMBL/GenBank/DDBJ accession no. D63344; note that only the relevant portion is shown). The 35-mer oligonucleotide used in this study as a probe and primer for Northern hybridization and primer extension analyses, respectively, was designed to be complementary to the underlined sequences. (B) The amino acid sequences of seven members of the CspA family of proteins are aligned. Identical residues are shown by dots, whereas gaps are indicated by dashes.

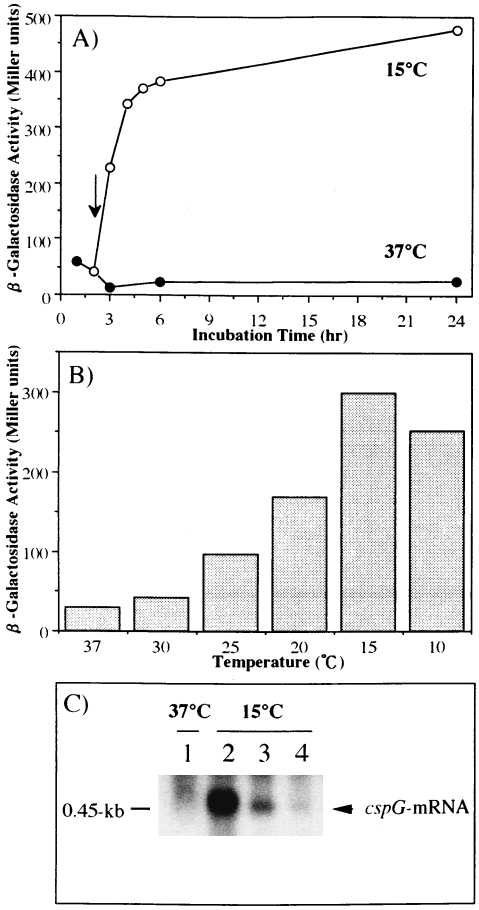

FIG. 3. β-Galactosidase activity expressed by cells carrying a *cspG-lacZ* transcriptional fusion gene and Northern hybridization analysis of the *cspG* mRNA. (A) Strain CSH26 [(Δ*lac-pro*) *ara thi* F⁻] containing pCBK001 carrying the *cspG-lacZ* fusion gene was grown in Luria broth at 37°C overnight. Cells were inoculated into the same fresh medium, and then the incubation temperature was changed from 37 to 15°C at the time indicated by the arrow. At intervals, the β-galactosidase activities expressed by these cells were measured (for clarity, the corresponding growth curves are not shown). (B) The same cells carrying pCBK001 were grown in Luria broth at 37°C until the logarithmic-growth phase and then incubated at the indicated temperatures for 4 h. The β-galactosidase activities expressed by these cells were then measured (11). (C) Total RNA fractions were prepared from cells grown under the following conditions: 37°C (lane 1) and 15°C for 1 (lane 2), 2 (lane 3), and 4 (lane 4) h. Samples were subjected to Northern hybridization analysis with the *cspG*-specific probe (Fig. 2).

Inspection of the structure of the *cspG* promoter. The total RNAs used in Fig. 3C were also subjected to primer extension analysis in order to identify the transcription start site(s) for the *cspG* mRNA (Fig. 4). Three transcripts with different start sites were found on the sequencing gel to be induced specifically in response to cold shock (P1, P2, and P3) (Fig. 4). Of

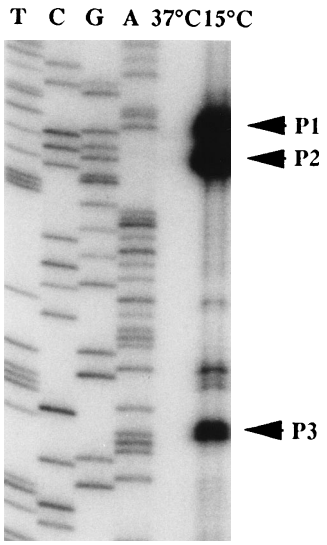

FIG. 4. Determination of the *cspG* transcription start sites. The set of total RNA fractions used in Fig. 3C were also subjected to primer extension analysis with the primer specific for *cspG* mRNA (Fig. 2). Cells were grown at 37°C and then treated further at 15°C for 1 h. The primer extension products were resolved on a sequencing gel alongside a sequencing ladder. The positions of three transcripts are shown on the right.

these, two major ones were found to start from AG (P1; positions −162 and −161) and GT (P2; positions −156 and −155), whereas a relatively minor one was found to start from AA (P3; positions −117 and −116) (the ATG initiation codon was considered +1). These results suggested that the *cspG* gene has multiple cold shock promoters upstream of its coding sequence. Their features are summarized in Fig. 5 by comparison with certain *csp* promoters (particularly with those of *cspA* and *cspB*). This inspection revealed that the *cspG* sequence is highly similar to both the *cspA* and *cspB* sequences not only in the coding regions but also in the 5′-upstream noncoding sequences. In particular, the relative position and nucleotide sequence of the major *cspG* promoter (P2) are well matched with those determined previously for the *cspA* promoter (15). Both of these transcription start sites are located about 160 nucleotides upstream of their ATG initiation codons, indicating that both have relatively long untranslated regions. Although the transcription start site for the *cspB* gene has not yet been determined, this gene most likely also has a promoter structure very similar to those of the former two (Fig. 4). It should be emphasized that these characteristics were not noticeable when sequences of other members (*cspC*, *cspD*, and *cspE*) were inspected.

Implications. *E. coli* contains at least seven *csp* genes, including the newly identified *cspG*. Among them, the *cspA*, *cspB*, and *cspG* genes are actually cold shock inducible. Therefore, CspA, CspB, and CspG may constitute a subfamily, whose functions might be closely related to each other but distinctive from others, with regard to *E. coli* physiology. In any case, only very limited information is available for understanding their molecular functions. Several lines of evidence suggest that the CspA family of proteins may be nucleic acid (DNA and/or RNA)-binding proteins (5, 9). In this context, CspA has been implicated in the transcriptional regulation of two cold shock genes, namely, *hns* and *gyrA* (6, 9). Another feature of the proteins of the CspA family has also emerged from an apparently unlinked line of experimentation (18). Both the *cspC* and *cspE* genes were identified as multicopy suppressors for a temperature-sensitive *mukB* mutation (*mukB106* allele) which produces normal-sized but anucleate cells. The tertiary structure of CspA has been determined recently by both X-ray crystallography and nuclear magnetic resonance (12, 13). A homolog of CSPs has also been identified in *Bacillus subtilis* (16). The insertional inactivation mutant (*cspB*::*cat*) of this *B. subtilis* gene exhibits slightly decreased viability after cycles of freezing and thawing. In spite of these and other intensive studies (5, 6, 9, 12, 13, 16, 18), it is difficult to ascribe physiological functions to the CspA family of proteins in the absence of genetic studies, including the isolation of mutants for each *csp* gene. In this context, our finding in this study is crucial because *E. coli* was demonstrated to contain at least three redundant cold-shock-inducible *csp* genes (i.e., *cspA*, *cspB*, and *cspG*), which produce highly analogous proteins (Fig. 2B). We would be not surprised if a *cspA* null mutant (even a *cspA cspB* double null mutant) yields no noticeable phenotypes. The isolation of a triple null mutant should be considered.

Although the mechanism of regulation of cold shock gene expression is also of interest, little is known about this particular aspect for the CspA family of genes (5). In this regard, a comparative study of the *cis*-acting regions of *cspA*, *cspB*, and *cspG* should shed light on the molecular mechanism underlying cold induction at the level of transcription and/or translation, since it is most likely that these three genes are regulated through a common mechanism in response to cold shock. In fact, we have shown that all of these genes have extensive similarities in their putative *cis*-acting and noncoding sequences, including a canonical promoter and a long untranslated sequence. The *cis*-acting sequences particularly conserved among them (Fig. 5) may give us a clue to address the relevant issues.

Nucleotide sequence accession number. The nucleotide sequence of the *cspG* gene has been deposited in the EMBL, GenBank, and DDBJ databases under accession number D63344.

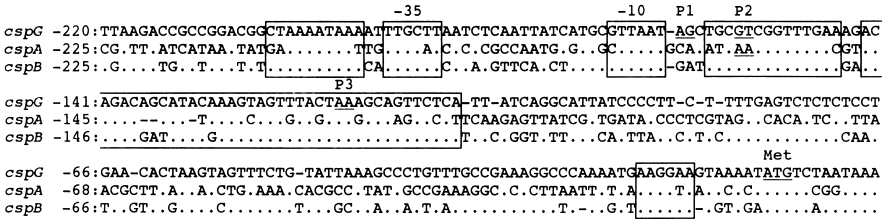

FIG. 5. The nucleotide sequence of the noncoding region of the *cspG* gene and its comparison with those of the *cspA* and *cspB* genes. Nucleotides identical in these three genes are denoted by dots (for numbering, the first nucleotide of the corresponding ATG codon was considered +1). In order to align appropriately, some gaps (dashes) were introduced. Three transcription start sites (P1, P2, and P3) were determined for *cspG* by primer extension analysis with the *cspG*-specific primer. A pair of canonical promoter sequences (−35 and −10 regions) was assigned for the major P2 promoter. The previously determined transcription start site for *cspA* is also underlined (15) (it is located at the same site as is P2 of *cspG*). Note that the highly conserved sequences among these genes are boxed.

REFERENCES

1. **Gething, M. J., and J. Sambrook.** 1992. Protein folding in the cell. Nature (London) **355:**33–45.
2. **Goldstein, J., N. S. Pollitt, and M. Inouye.** 1990. Major cold shock proteins of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **87:**283–287.
3. **Gross, C. A., D. B. Straus, J. W. Erickson, and T. Yura.** 1990. Function and regulation of heat shock proteins in *Escherichia coli*, p. 167–189. *In* R. Morimoto, A. Tissieres, and C. Georgopoulos (ed.), The biology of heat shock proteins and molecular chaperones. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
4. **Jaurin, B.** 1987. A promoter probe vector (pJAC4) that utilizes the *ampC* β-lactamase gene of *Escherichia coli*. Nucleic Acids Res. **15:**8567.
5. **Jones, P. G., and M. Inouye.** 1994. The cold-shock response—a hot topic. Mol. Microbiol. **11:**811–818.
6. **Jones, P. G., R. Krah, S. R. Tafuri, and A. P. Wolffe.** 1992. DNA gyrase, CS7.4, and the cold shock response in *Escherichia coli*. J. Bacteriol. **174:** 5798–5802.
7. **Jones, P. G., R. A. VanBogelen, and F. C. Neidhardt.** 1987. Induction of proteins in response to low temperature in *Escherichia coli*. J. Bacteriol. **169:**2092–2095.
8. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50:**495–508.
9. **La Teana, A., A. Brandi, M. Falconi, R. Spurio, C. L. Pon, and C. O. Gualerzi.** 1991. Identification of a cold shock transcriptional enhancer of the *Escherichia coli* gene encoding nucleoid protein H-NS. Proc. Natl. Acad. Sci. USA **88:**10907–10911.
10. **Lee, S. J., A. Xie, W. Jiang, J. Etchegaray, P. G. Jones, and M. Inouye.** 1994. Family of the major cold-shock protein, CspA (CS7.4), of *Escherichia coli*, whose members show a high sequence similarity with the eukaryotic Y-box binding proteins. Mol. Microbiol. **11:**833–839.
11. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 352–355. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
12. **Newkirk, K., W. Feng, W. Jiang, R. Tejero, S. D. Emerson, M. Inouye, and G. T. Montelione.** 1994. Solution NMR structure of the major cold shock protein (CspA) from *Escherichia coli*: identification of a binding epitope for DNA. Proc. Natl. Acad. Sci. USA **91:**5114–5118.
13. **Schindelin, H., W. Jiang, M. Inouye, and U. Heinemann.** 1994. Crystal structure of CspA, the major cold shock protein of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **91:**5119–5123.
14. **Schneider, K., and C. F. Beck.** 1986. Promoter-probe vectors for the analysis of divergently arranged promoters. Gene **42:**37–48.
15. **Tanabe, H., J. Goldstein, M. Yang, and M. Inouye.** 1992. Identification of the promoter region of the *Escherichia coli* major cold shock gene, *cspA*. J. Bacteriol. **174:**3867–3873.
16. **Willimsky, G., H. Bang, G. Fischer, and M. A. Marahiel.** 1992. Characterization of *cspB*, a *Bacillus subtilis* inducible cold shock gene affecting cell viability at low temperatures. J. Bacteriol. **174:**6326–6335.
17. **Wolffe, A. P., S. Tafuri, M. Ranjan, and M. Familari.** 1992. The Y-box factors: a family of nucleic acid binding proteins conserved from *Escherichia coli* to man. New Biol. **4:**290–298.
18. **Yamanaka, K., T. Mitani, T. Ogura, H. Niki, and S. Hiraga.** 1994. Cloning, sequencing, and characterization of multicopy suppressors of a *mukB* mutation in *Escherichia coli*. Mol. Microbiol. **13:**301–312.