JOURNAL OF BACTERIOLOGY, Oct. 1995, p. 5952–5958 Vol. 177, No. 20
0021-9193/95/$04.00+0

The *dnaKJ* Operon of *Agrobacterium tumefaciens*: Transcriptional Analysis and Evidence for a New Heat Shock Promoter

GIL SEGAL AND ELIORA Z. RON*

Department of Molecular Microbiology and Biotechnology, The George S. Wise Faculty of Life Sciences, Tel-Aviv University, 69978 Tel-Aviv, Israel

Received 5 June 1995/Accepted 14 August 1995

The *dnaKJ* operon of *Agrobacterium tumefaciens* was cloned and sequenced and was found to be highly homologous to previously analyzed *dnaKJ* operons. Transcription of this operon in *A. tumefaciens* was stimulated by heat shock as well as by exposure to ethanol and hydrogen peroxide. There were two transcripts representing the *dnaKJ* operon: one containing the *dnaK* and *dnaJ* genes and the second containing only the *dnaK* gene. Primer extension analysis indicated that transcription started from the same site in heat-shocked cells and in untreated cells. The upstream regulatory region of the *dnaKJ* operon of *A. tumefaciens* does not contain the highly conserved inverted repeat sequence previously found in the *groESL* operon of this bacterium, as well as in many other *groE* and *dnaK* operons. Sequence analysis of the promoter region of several *groESL* and *dnaK* operons from α-purple proteobacteria indicates the existence of a putative promoter sequence different from the known consensus promoter sequences recognized by the *Escherichia coli* vegetative or heat shock sigma factor. This promoter may constitute the heat shock promoter of these α-purple proteobacteria.

The heat shock response is a widespread phenomenon that was found in all living cells examined (11). It is characterized by the induction of several proteins, some of which are highly conserved in evolution, especially those encoded by the *groEL* (*hsp60*) and *dnaK* (*hsp70*) genes (5, 14, 53). Several of the heat shock proteins are also induced by other stress conditions such as exposure to ethanol, heavy metals, and hydrogen peroxide (30, 34, 45).

In *Escherichia coli* the heat shock response is mediated by the positive regulator protein sigma-32. This sigma factor recognizes a promoter sequence different from that of the vegetative sigma factor (sigma-70) and in this way specifically transcribes heat shock genes (7, 52).

Recently, it was observed that gram-positive bacteria and several gram-negative bacteria contain a highly conserved inverted repeat sequence in the upstream regulatory region of *groESL* and *dnaK* operons (1, 6, 10, 15, 21, 27–29, 31, 32, 38, 40, 44, 47, 48). In *Bacillus subtilis* (54) and *Lactococcus lactis* (44), it was shown that this inverted repeat is involved in the heat shock response and that these heat shock operons contain a promoter sequence recognized by the vegetative sigma factor. In *B. subtilis*, other heat shock-activated genes contain a promoter sequence recognized by the SigB sigma factor, which is known as a general stress factor (49).

In a previous paper (40), we showed that the *groESL* operon of *Agrobacterium tumefaciens* contains the highly conserved inverted repeat sequence and that the operon is heat shock activated. Here, we present results from the cloning and sequencing of the *dnaKJ* operon of *A. tumefaciens*. The operon was shown to be heat shock activated, although it does not contain the inverted repeat sequence found in the *groESL* operon of this bacterium. The results presented in this paper and sequence analysis of heat shock promoters from other *groESL* and *dnaK* operons from α-purple proteobacteria suggest the existence of a new putative promoter consensus sequence which is different from those of the vegetative and heat shock promoters of *E. coli.*

* Corresponding author. Phone: 972-3-6409379. Fax: 972-3-6429377. Electronic mail address: eliora@ccsg.tau.ac.il (Internet) or Ron@taunivm (Bitnet).

MATERIALS AND METHODS

Bacterial strains and growth conditions. *A. tumefaciens* C58 (ATCC 33970) was used for preparation of DNA and RNA and was cultivated as previously described (40). *E. coli* MC1022 [*araD139* Δ(*ara leu*)*7697* Δ(*lacZ*)M15 *galU galK strA*] (8) was used for transformations. *E. coli* 71-18 (F′ *lacI*^{q} Δ*lacZ*)M15 *proAB*/Δ*lac-proAB thi supE*) (51) was used for M13 manipulations. Heat shock was achieved by transfer of exponentially growing cells from 25 to 42°C. The other stress condition was 4% (vol/vol) ethanol or 2 mM H_2O_2.

RNA and DNA manipulations. RNA was prepared from 50-ml cultures of *A. tumefaciens* cells and was manipulated as described previously (40, 41). Each lane in the Northern (RNA) hybridizations and primer extension analysis contained 10 μg of RNA. The level of activation with the different stresses was analyzed by Northern hybridizations and calculated by using the Fuji BAS1000 PhosphorImager. Chromosomal DNA for PCR was prepared as previously described (39).

PCR conditions. For PCR mixtures that included degenerate primers, PCR with 30 cycles of 94°C for 1 min, 50°C for 0.5 min, and 72°C for 2 min was performed with 100-μl reaction mixtures, with the buffer supplied with the enzyme, 200 μM (each) deoxynucleoside triphosphates, 1 μg of chromosomal DNA, 50 pmol (each) of the primers, and 2 U of Bioprobe Systems (Montreuil, France) *Taq* DNA polymerase. For PCR mixtures that did not include degenerate primers, the same reaction conditions were used, except for annealing at 60°C. The PCR products were gel purified and cloned into *Hin*dII-digested pUC18 (51).

Probes and primers. The following probes were used for Northern hybridization. The *dnaK* probe was a 676-bp fragment generated from plasmid pGS-AK-04 (this work) by PCR with the KF2 and KR1 primers (positions 1051 to 1727 in the sequence). The *dnaJ* probe was a 225-bp fragment generated by an *Fnu*4HI digest of pGS-AK-06 (positions 2585 to 2810 in the sequence [this work]). The *groEL* probe was a 418-bp fragment generated by a *Pst*I digest of pGS-AG-3a (positions 2585 to 2810 in the sequence [40]). The primers used for PCR and for primer extension are listed in Table 1.

DNA sequencing. The sequence was determined by the dideoxy chain termination method (37) with the Sequenase II sequencing kit from United States Biochemicals.

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank, and DDBJ nucleotide sequence databases under accession number X87113.

TABLE 1. Primers used in this study

Primer	Sequence[a]	Sequence position[b]
KF1	ATYGGNATYGAYCTNGGNAC	640–659
KF2	GTNCCNGCNTAYTTYAAYGA	1051–1070
KR1	GCNACNACYTCRTCNGGRTT	1727–1708
KR2	TCRAANGTNACYTCDATYTG	2042–2023
JR1	TCRTANGCNGCNCGYTTYTG	2819–2800
IR1	TCCTTGGTGGCCTGACGCTG	1094–1075
IR2	CTGAAGAACGACAAGCTCGC	1366–1385
IR3	GATACGCAGAACATCGAGAC	1130–1111
IR4	TGTTCGGCAAGGAGCCGCAC	1679–1698
PE2	TGAAGCGAGCTGTCTGAACC	613–594

[a] Sequences are presented from 5′ to 3′. N, G/A/T/C; R, G/A; Y, T/C; D, G/A/T.
[b] Positions in the *A. tumefaciens* sequence, as presented in Fig. 2.

RESULTS

Cloning of the *dnaKJ* operon of *A. tumefaciens*. The *dnaKJ* operon of *A. tumefaciens* was cloned by using PCR with degenerate primers designed according to known sequences of *dnaK* genes (20). The cloning involved three stages as follows and is summarized in Fig. 1.

(i) Amplification of an internal fragment of the *dnaK* gene with degenerate primers. By using the degenerate primers KF1 and KR2, a weak band with the expected size of 1.4 kb was amplified. This product was used for two nested PCRs (25): one with the KF1 and KR1 degenerated primers and the second with the KF2 and KR2 degenerated primers. Two fragments with sizes of about 1 kb each were amplified and cloned; these fragments had a 0.7-kb overlap between them, corresponding to the sequence between the KF2 and the KR1 degenerated primers. The plasmid containing the KF1-KR1 (1,087-bp) fragment was named pGS-AK-04, and the plasmid containing the KF2-KR2 (989-bp) fragment was named pGS-AK-05.

(ii) Amplification of a fragment containing the C-terminal end of the *dnaK* gene, the N-terminal part of the *dnaJ* gene, and the region between them. By using the JR1 degenerate primer (located at the N-terminal conserved region of known *dnaJ* genes) and the IR4 primer (located in the region that was cloned in stage i), a fragment with a size of 1,140 bp was cloned, and the plasmid was named pGS-AK-06.

(iii) Amplification of the upstream region of the *dnaK* gene by using the inverse-PCR technique (43). A 2.4-kb *Hin*dIII fragment was found by Southern hybridization to contain the upstream region of the *dnaK* gene (data not shown). This fragment was amplified with the IR2 and IR3 primers, and then a nested PCR was performed with the IR1 and IR4 primers (Fig. 1). A 1.8-kb fragment was amplified and then digested with *Hin*dIII and *Bss*HII, and a 722-bp fragment was purified from the gel and cloned. The 722-bp fragment contained the upstream region of the *dnaKJ* operon. The plasmid was named pGS-AK-10.

Sequence analysis of the *dnaKJ* genes. The nucleotide sequences of the *dnaK* gene and part of the adjacent *dnaJ* gene and the deduced sequences of the corresponding proteins are shown in Fig. 2. There is a high degree of homology of both sequences with previously analyzed *dnaK* and *dnaJ* genes, but no such homology was found in the upstream region of the *dnaK* gene. The DnaK protein from *A. tumefaciens* shows 83% identity and 90% similarity to the *Brucella ovis* (9) protein and 67% identity and 83% similarity to the *E. coli* (3) protein. In several gram-positive bacteria, the *grpE* gene was shown to be located upstream of the *dnaK* gene, with an operon structure of *orfA-grpE-dnaK-dnaJ* (29, 48). In *A. tumefaciens*, like in *E. coli* (3, 4), and in *Caulobacter crescentus* (19), the *dnaK* operon organization does not include the *grpE* gene upstream of the *dnaK* gene, and an operon structure of *dnaK-dnaJ* was found.

mRNA analysis of the *dnaKJ* operon under heat shock conditions. The in vivo transcripts of the *dnaK* and *dnaJ* genes were detected by Northern analysis. Total RNA was prepared from cells of *A. tumefaciens* growing exponentially at 25°C and at different times after transfer to 42°C. The RNA was probed with a radioactively labeled 676-bp internal fragment of the *dnaK* gene and a 225-bp fragment of the N-terminal end of *dnaJ*. The Northern analysis with the *dnaK* probe (Fig. 3) revealed heat shock activation of two bands with sizes of 2 and 4.1 kb. The 4.1-kb band also appeared with the *dnaJ* probe (Fig. 3), but the 2-kb band was restricted to the *dnaK* gene. The *dnaKJ* operon had a very low level of transcription at 25°C: the 2-kb band was observed only after long exposure, and the 4.1-kb band could not be detected at that temperature. The heat shock activation of the operon resulted in a large increase in the mRNA level of the 2-kb band and a small increase in that of the 4.1-kb band. Both bands reached their maximum level after 20 min, followed by a decrease after 30 min. The existence of the two transcripts was probably due to a transcription terminator located between the *dnaK* and *dnaJ* genes (marked in Fig. 2) that stops most of the transcripts from transcribing the *dnaJ* gene. The 2-kb band was the dominant band of the operon at low and at high temperatures; its level was about 10 times higher than that of the 4.1-kb band during heat shock.

Comparison of levels of activation of the *groESL* and *dnaKJ* operons by different stress conditions. So far, two operons of *A. tumefaciens* were shown to be heat shock activated: the *groESL* operon (40), which contains the conserved inverted repeat sequence that was shown to be involved in the heat shock response in gram-positive bacteria (44, 54), and the *dnaKJ* operon, which does not contain this inverted repeat sequence.

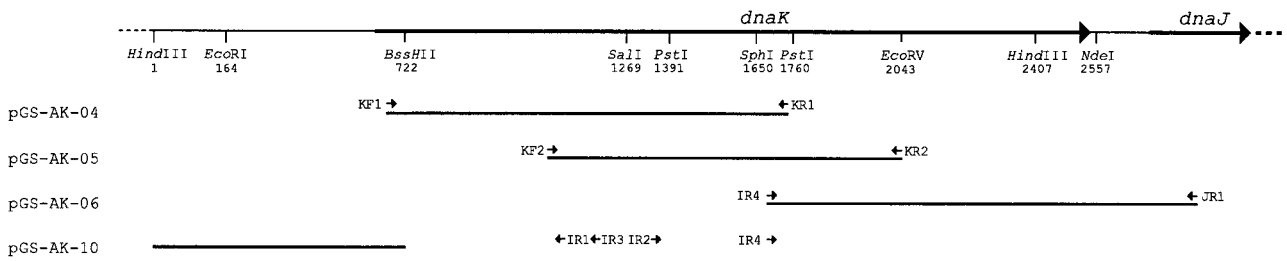

FIG. 1. Restriction map and plasmids of the *dnaKJ* locus. Inserts of the four plasmids described in the text are shown. The primers used in the different cloning stages are indicated with small arrows. The locations of the *dnaK* and *dnaJ* genes are indicated with large arrows.

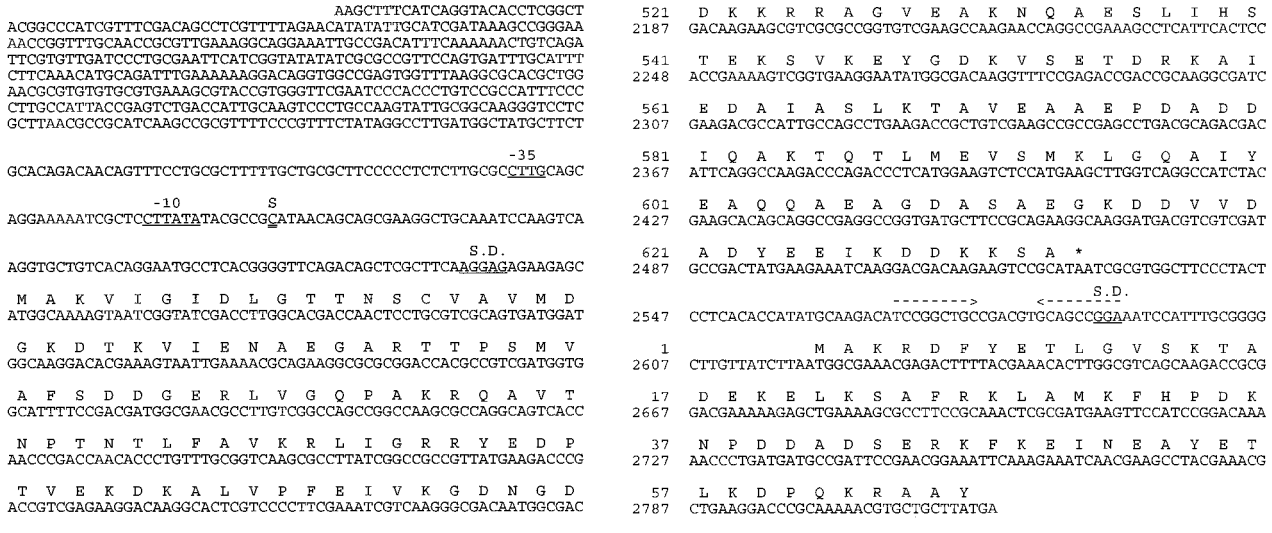

FIG. 2. Nucleotide sequence of the *dnaKJ* operon of *A. tumefaciens*. The sequence was determined by the dideoxy chain termination method as described in Materials and Methods. The deduced amino acid sequences of the DnaK protein and part of the DnaJ protein are shown above the DNA sequence (single-letter code). The translation stop codon is marked by an asterisk. The putative ribosome binding site is indicated as S.D. (Shine-Dalgarno). Arrows above the DNA sequence indicate the putative terminator sequence. The putative −10 and −35 regions are underlined. The transcript start point is marked S.

We compared the stress responses of these two operons under different stress conditions: heat shock, exposure to ethanol, and exposure to hydrogen peroxide (oxidative stress). The mRNA of the *groESL* operon was detected by using a 418-bp probe from the *groESL* gene. The mRNA of the *dnaKJ* operon was detected by using a 676-bp probe from the *dnaK* gene. The transcription level for each operon at 25°C was set as 1, and the results were analyzed as induction ratios.

The major difference between the two operons was the level of transcription at 25°C, at which the level of the *groESL* mRNA was remarkably higher than the level of the *dnaKJ* mRNA. The results presented in Fig. 4 show that at 42°C, the *dnaKJ* operon had a higher activation level than the *groESL* operon. The two operons had different kinetics of activation. The *groESL* operon reached its maximum level of activation after 5 min, and the *dnaKJ* operon reached its maximum level of activation after 20 min. The activation levels after exposure to ethanol were about the same for the two operons, and both had a similar kinetics of activation (maximum after 10 min). On the other hand, after exposure to hydrogen peroxide, the *dnaKJ* operon was activated, but the transcription level of the *groESL* operon decreased. These results indicate that the regulation of the two operons is complicated and is probably mediated by more than one regulator.

Determination of the transcription start site of the *dnaKJ* operon. The transcription start site of the *dnaKJ* operon was determined by primer extension analysis with the PE2 primer complementary to the 5′ end of the *dnaK* gene (Fig. 5). The transcription start site was in the same position at 25°C and under heat shock conditions and was located 92 bases upstream of the DnaK first methionine, with a C at the 5′ end of the mRNA. The conserved inverted repeat sequence found at the transcription start site of the *groESL* operon of *A. tumefaciens* and in many other *groESL* and *dnaK* operons was not found in the *dnaK* operon of *A. tumefaciens*, either at the transcription start site or upstream of it.

Comparison of the upstream regulatory regions of *dnaKJ* and *groESL* operons from α-purple proteobacteria. The promoter region of the *A. tumefaciens dnaKJ* operon did not appear to have the expected homology with the known *E. coli* vegetative or heat shock promoter. In order to find out if a promoter sequence can be identified at the *dnaK* regulatory

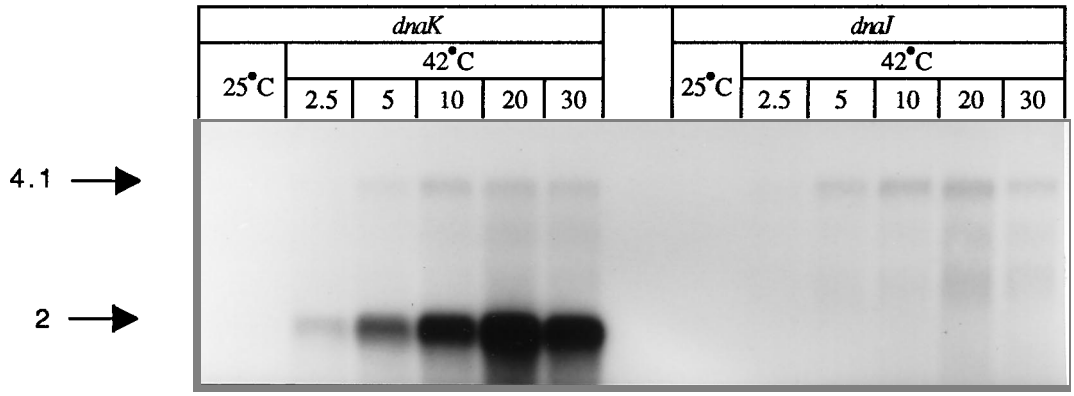

FIG. 3. Northern hybridization of the *dnaK* and *dnaJ* mRNAs. Total RNA of *A. tumefaciens* was isolated before (25°C) and at different time points after (2.5, 5, 10, 20, and 30 min) heat shock to 42°C. The mRNA was probed with a ^{32}P-labeled 676-bp internal fragment of the *dnaK* gene and with a ^{32}P-labeled 225-bp fragment of the *dnaJ* gene. The sizes (in kilobases) of the bands reflecting the *dnaK* (2 kb) and *dnaKJ* (4.1 kb) mRNAs are marked.

region, nine sequences of known heat shock operons (*dnaKJ* and *groESL*) from bacteria that belong to the α-purple proteobacteria (33, 50) were compared (Fig. 6A). The transcription start sites were determined for three of the nine operons (*groESL* and *dnaKJ* of *A. tumefaciens* [reference 40 and this work]) and the *dnaKJ* operon of *C. crescentus* [19]) and were found to be the same in vegetative growth and during heat shock. (In *C. crescentus*, a second transcription start site was found, but it was not heat shock activated [19].) The six *groESL* operons that were compared contain the conserved inverted repeat sequence, and the sequences were aligned according to it. The *dnaK* operons were aligned according to the transcription start site and homology of the promoter region. As can be seen in Fig. 6A, a consensus promoter sequence can be deduced from the nine sequences, with the sequence CTTG (17 to 18 bp) CYTAT-T--G. The putative −35 region was conserved in all nine sequences, and the putative −10 region was less conserved, but two bases (A--T) were found to be conserved in all of the sequences. In the three first bases of the −10 region, a CYT sequence was found as the consensus sequence and the changes observed were only other pyrimidines.

The putative consensus promoter sequence deduced from Fig. 6A was compared with the *E. coli* vegetative and heat shock promoter consensus sequences (Fig. 6B). Homology was found between the consensus sequences at the −35 region; the TTG sequence that appears in both of the *E. coli* consensus sequences also appears in the putative promoter consensus sequence. The rest of the sequence found at −35 of the *E. coli* consensus promoter sequences does not appear in the putative consensus sequence found in Fig. 6A. The putative −10 region contains an AT-rich sequence, as in the two *E. coli* consensus sequences, but differs from the *E. coli* consensus promoter sequences.

DISCUSSION

The *dnaK* gene and part of the *dnaJ* gene of *A. tumefaciens* were cloned and sequenced and show high levels of similarity to previously described *dnaK* and *dnaJ* genes and proteins. The *dnaKJ* operon was found, by Northern analysis and primer extension, to be transcriptionally activated by heat shock. The transcription of the *dnaKJ* operon of *A. tumefaciens* was initiated from the same position under normal and heat shock conditions. The *dnaKJ* operon of *A. tumefaciens* was found to contain two transcripts: a 2-kb band that contains only the *dnaK* gene and a 4.1-kb band that contains the *dnaK* and *dnaJ* genes. The 2-kb band was the major band under heat shock conditions and was the only band visible (at a very low level) at normal growth temperature. The appearance of the two bands probably results from a transcription terminator located between the *dnaK* and *dnaJ* genes that prevents most of the transcripts from transcribing the *dnaJ* gene.

So far, two operons of *A. tumefaciens* were shown to be heat shock activated: the *groESL* operon (40) and the *dnaKJ* operon (this work). The responses of these two operons under different stress conditions were determined. The major difference between the *dnaKJ* and the *groESL* operons in *A. tumefaciens* was the level of transcription under normal growth conditions, in which the *dnaK* operon had a very low level of transcription compared with the *groESL* operon. The level of temperature-induced activation of the *dnaK* operon appeared to be much higher than that of the *groESL* operon (Fig. 4). However,

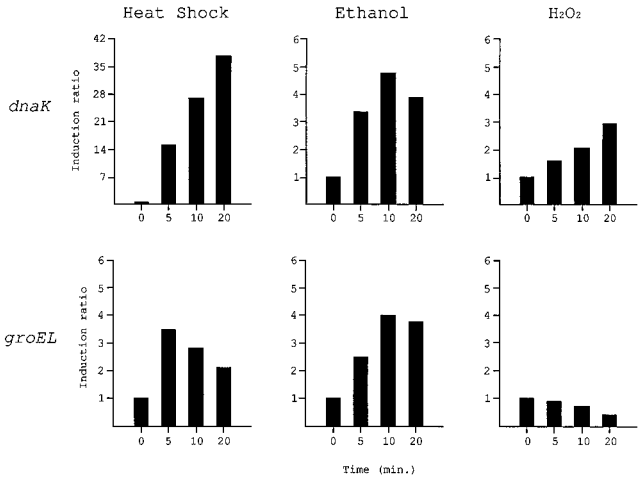

FIG. 4. Schematic representation of the increase in the *dnaK* and *groEL* mRNA levels after different stresses. *A. tumefaciens* cells were exposed to stresses as described in Materials and Methods. Northern hybridizations to total RNA prepared from cells before (0) and at different time points after (5, 10, and 20 min) exposure to the stress were performed. The mRNA was probed with a ^{32}P-labeled 676-bp internal fragment of the *dnaK* gene and with a ^{32}P-labeled 418-bp internal fragment of the *groEL* gene. The hybridization signals were quantified with a Fuji BAS1000 PhosphorImager. The induction ratios of the mRNAs are shown. The scale of the *dnaK* operon under heat shock conditions is different from the scale of the others.

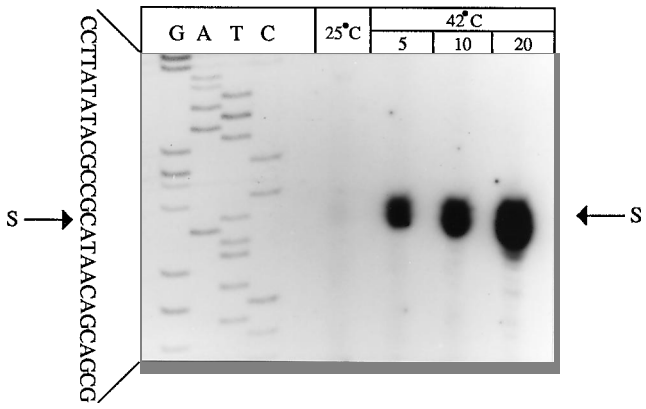

FIG. 5. Mapping of the 5′ end of the *dnaKJ* operon by primer extension analysis. The PE2 primer complementary to the 5′ end of the *dnaK* gene was hybridized with total RNA from *A. tumefaciens* isolated before (25°C) and at different time points after (5, 10 and 20 min) heat shock to 42°C. The primer extension products were analyzed on a sequencing gel. G, A, T, and C are products of the sequencing reaction obtained by using the same primer. The sequence presented is that of the sense strand. The arrow at S points the base representing the 5′ end of the mRNA.

because it was previously shown that in the *groESL* mRNA there is a heat shock-dependent cleavage (41), the aspect of mRNA stability has to be considered. In this case, it is possible that the two operons—*groESL* and *dnaKJ*—undergo the same activation but a different mRNA processing. When other stress conditions were examined, the response observed was different from that observed during temperature shift. After exposure to ethanol, the two operons had a similar induction ratio and a similar kinetics of activation. However, after exposure to hydrogen peroxide, the *dnaK* operon was activated, but the transcription of the *groESL* operon decreased. The difference observed after exposure to hydrogen peroxide clearly demonstrates that at least one factor activates only the *dnaK* operon, and it seems that the two operons are subjected to regulation by more than one regulator.

In previous studies, a conserved inverted repeat sequence was found at the transcription start site of the *groESL* operon of *A. tumefaciens* (40) and in many other *groESL* and *dnaK* operons (1, 6, 10, 15, 21, 27–29, 31, 32, 38, 44, 47, 48). This inverted repeat sequence was found to be involved in heat shock regulation in two species of gram-positive bacteria (44, 54). The *dnaK* operon of *A. tumefaciens* does not contain the inverted repeat sequence near its transcription start site or upstream of it. However, both the *groESL* and *dnaKJ* operons of *A. tumefaciens* were found to be heat shock activated. This observation can be explained by assuming that the inverted repeat sequence is responsible for the heat shock activation of the *groESL* operon and that the *dnaK* operon is activated by a different heat shock control system. An alternative explanation would assume that the inverted repeat sequence is not part of the heat shock activation system in *A. tumefaciens* and another system regulates the heat shock activation of both operons. It should be noted that there is another case—in the high-GC gram-positive bacteria—in which the *groESL* operon and the *groEL* gene contain the conserved inverted repeat sequence (13) and the *dnaK* operon does not (6). However, in this system, only the *groESL* operon and the *groEL* gene were studied and found to be heat shock activated (13); there is no information on the heat shock response of the *dnaK* operons.

In search of putative regulatory elements for the heat shock operons of *A. tumefaciens*, the upstream region of nine *dnaKJ* and *groESL* operons from bacteria that belong to the α-purple proteobacteria were compared (Fig. 6A). From this comparison, a putative consensus promoter sequence could be located, with the sequence CTTG (17 to 18 bp) CYTAT-T--G. The four bases in the −35 region (CTTG) and two bases at the −10 region (A--T) were found to be conserved in all nine operons. The putative promoter consensus sequence found in Fig. 6A has some degree of homology with the *E. coli* vegetative and heat shock promoter consensus sequences (Fig. 6B). However, the *E. coli* vegetative and heat shock consensus promoter sequences also share a certain degree of homology, and several of the conserved bases are also present in the putative consensus promoter sequence found in Fig. 6A. The major difference between the two *E. coli* promoter consensus sequences is the presence of the nine cytidines (five at the −35 region and four at the −10 region) found only in the heat shock promoter consensus sequence (52). Only two cytidines (one at −35 and one at −10) were found in the putative promoter consensus sequence described in Fig. 6A, which is missing the unique

A.

```
CcK - GTGGGAGGGGTTATGGC**CTTG**CGTGGCGGCCC-CTATCCCCC**ATAT**CC**G**GCTTCGAAGCGGGGTTGACGGGCTCGTCAACTTCGCACA
BoK - TTGTCTGGCTCTTGTGA**CTTG**CAGTGAAAAACTGCCTTT**CTTATA**TAC**G**CCTCGCACAGGGCTTTGATGACAGGCCGGAAACCGGTCC
AtK - TTCCCCCTCTCTTGCGC**CTTG**CAGCAGGAAAAATCGCTCC**TTATAT**ACG**G**CCGCATAACAGCAGCGAAGGCTGCAAATCCAAGTCAAGG
AtE - TTTTCCGGCAGCGCGTT**CTTG**ACTCCGGGAGAGCGCCTCC**TTAT**CTCGG**C**TCTGCTGGCACTCCAACAAGGGGAGTGCTAACACCCAT
BaE - CCTGAACGGCAATCCTC**CTTG**ACAAAAAAACATGCGGCTTC**TATC**TCAAGAACCAGGTTAGCACTCGGACGAATAGAGTGCTAACAGC
BjE - CTTCGTCGCGGCTTTGT**CTTG**TTTGCCGACCA-AGCCGGC**TTATTT**TCGGAGCGGTCAGCACTCGCTAGCTTCGACTGCTATTTCAAA
CcE - ACTTCGCGGCGCGCGTC**CTTG**ACTCCCCGAGGGCCGGCGC**CTA**ACTAACTCACCAGCGTCGTTGGCACTCGAATGGAGCGACTGCTAA
RmE - AAATCTCGGAGCCTATC**CTTG**ACTCATTTCACGGCCAGC**CTTATC**TCGGCTGGCGCTACGACTCGCTTGATATGAGTGCTAACACCCA
ZmE - AAATTTTTGAAGGATAG**CTTG**ACGTCCCGAAA-TACTAC**CCCATATGGG**GGACATTGGCACTTCGTAAGGGAGAGTGCCAGCAGGAAT

CONSENSUS -            CTTG                 CYTAT-T--G
```

B.

```
E. coli SIGMA-70 -              TTGACA     <17>        TATAAT
CONSENSUS        -             CTTG      <17/18>     CYTAT-T--G
E. coli SIGMA-32 -      TCTC-CCCTTGAA     <13/14>    CCCCAT-AT
```

FIG. 6. (A) Comparison between the promoter regions of three *dnaKJ* operons and six *groESL* operons from bacteria that belong to the α-purple proteobacteria. The *dnaKJ* operons are abbreviated as follows: CcK, *C. crescentus* (19); BoK, *B. ovis* (9); AtK, *A. tumefaciens* (this work). The *groESL* operons are abbreviated as follows: AtE, *A. tumefaciens* (40); BaE, *Brucella abortus* (22); BjE, *Bradyrhizobium japonicum* (16); CcE, *C. crescentus* (accession no. L41394); RmE, *Rhizobium meliloti* (36); ZmE, *Zymomonas mobilis* (2). Bases that appear in seven or more of the nine sequences are marked by boldface letters and are considered the consensus sequence. The experimentally determined transcription start sites are double underlined, and the sequences of the inverted repeat found in the *groESL* operons are single underlined. The letter Y in the consensus sequence marks pyrimidines (C or T). (B) Comparison between the *E. coli* vegetative sigma factor promoter consensus sequence (*E. coli* SIGMA-70) (12), the *E. coli* heat shock sigma factor promoter consensus sequence (*E. coli* SIGMA-32) (52), and the putative promoter consensus sequence that was described above (A).

sequences of the *E. coli* heat shock promoter consensus sequence.

The promoter sequence recognized by the *A. tumefaciens* vegetative sigma factor is not known because of insufficient data. However, there are reasons to suppose that it will recognize a promoter sequence similar to the one recognized by the *E. coli* vegetative sigma factor. It is known that in evolutionarily distant bacteria, such as *E. coli* and *B. subtilis* (33, 50), the vegetative sigma factor recognizes the same promoter sequence (12), and *A. tumefaciens* is more closely related to *E. coli* than to *B. subtilis* (33, 50). In addition, several amino acids involved in promoter recognition were identified in the *E. coli* vegetative sigma factor (17, 42, 46). Two arginine residues located at region 4 (at positions 584 and 588) were found to be involved in the recognition of the promoter −35 region, and a glutamic acid residue and a threonine residue located at region 2.4 (at positions 437 and 440, respectively) were found to be involved in the recognition of the promoter −10 region. Region 4 and region 2.4 are highly conserved among vegetative sigma factors (23) and were also found to be conserved in the *A. tumefaciens* vegetative sigma factor (39), which contains all four of the amino acids mentioned above in locations identical to those in the *E. coli* vegetative sigma factor. Therefore, it is most likely that the *A. tumefaciens* vegetative sigma factor will recognize a promoter sequence similar to the one recognized by the *E. coli* vegetative sigma factor.

There were two previous reports indicating that *dnaK* operons from bacteria that belong to the α-purple proteobacteria were heat shock activated in *E. coli*. A *lacZ* fusion of the *dnaK* operon from *Zymomonas mobilis* was thermoregulated in *E. coli* (26), and the *dnaK* operon of *B. ovis* was expressed in a heat shock-dependent manner in *E. coli* (9). In both cases, the transcription start site of the operon was not determined, and there is no evidence that transcription in *E. coli* occurred from the same promoter as that in the original bacteria. In another paper (24), the *E. coli dnaK* gene was introduced into *A. tumefaciens* and was shown to be heat shock activated from the same transcription start site as in *E. coli*. Because the putative consensus promoter sequence presented in Fig. 6A has some degree of homology with the *E. coli* heat shock and vegetative promoters (Fig. 6B), it is possible that the *E. coli dnaK* promoter was recognized by the *A. tumefaciens* alternative sigma factor and was heat shock activated.

Genes coding for sigma factors homologous to the *E. coli* heat shock sigma factor (sigma-32) were found in *Citrobacter freundii* (18) and in *Pseudomonas aeruginosa* (35), both belonging to the γ-purple proteobacteria (33, 50). In contrast, a gene homologous to this sigma factor could not be found in *A. tumefaciens* by low-stringency Southern hybridization with the *E. coli htpR* gene (40), and the corresponding protein could not be found with antibodies against the *E. coli* HtpR protein (24). In *B. subtilis* and other low-GC gram-positive bacteria, the *dnaK* and *groESL* operons that contain the evolutionarily conserved inverted repeat sequence are transcribed at normal growth temperature and are activated under heat shock by the vegetative sigma factor (21, 28, 29, 38, 48). Other heat shock genes in *B. subtilis* were found to be recognized by an alternative sigma factor—SigB. This sigma factor recognizes promoter consensus sequences different from those recognized by the vegetative sigma factor (49) and does not belong to the sigma-32 family of bacterial sigma factors (23).

From the analysis of the promoter region and the heat shock activation of the operons, it is suggested that the α-purple proteobacteria contain an alternative sigma factor for heat shock (or stress). This alternative sigma factor is expected to recognize the putative promoter consensus sequence that was found in the upstream region of the *groESL* operons (that contain the evolutionarily conserved inverted repeat sequence), as well as the *dnaKJ* operons (that do not contain the evolutionarily conserved inverted repeat sequence) from this division.

The sigma factor that recognizes the new putative promoter sequence is not known. It is probably different from the *E. coli* heat shock sigma factor (sigma-32), since no homolog for the *E. coli* heat shock sigma factor was found in *A. tumefaciens*. In addition, there are also differences between the *E. coli* (γ-purple proteobacteria) heat shock consensus promoter sequence and the putative consensus promoter sequence of the heat shock-activated operons from the α-purple proteobacteria. It is still possible that in *A. tumefaciens*, the heat shock promoters are recognized by a sigma factor functionally related to sigma-32 that recognizes the putative, different promoter.

ACKNOWLEDGMENTS

This work was supported by the Israel Science Foundation. G.S. was supported by the Clore Scholars program.

REFERENCES

1. **Ballard, S. A., R. P. A. M. Segers, N. Bleuminl-Pluym, J. Fyfe, S. Faine, and B. Adler.** 1993. Molecular analysis of the hsp (*groE*) operon of *Leptospira interrogans* serovar copenhageni. Mol. Microbiol. **8:**739–751.
2. **Barbosa, M. F. S., L. P. Yomano, and L. O. Ingram.** 1994. Cloning, sequencing and expression of the stress genes from the ethanol producing bacterium *Zymomonas mobilis*: the *groESL* operon. Gene **128:**51–57.
3. **Bardwell, J. C. A., and E. A. Craig.** 1984. Major heat shock gene of *Drosophila* and *Escherichia coli* heat-inducible *dnaK* gene are homologous. Proc. Natl. Acad. Sci. USA **81:**848–852.
4. **Bardwell, J. C. A., K. Tilly, E. A. Craig, J. King, M. Zylicz, and C. Georgopoulos.** 1986. The nucleotide sequence of the *Escherichia coli* K12 *dnaJ* gene. J. Biol. Chem. **261:**1782–1785.
5. **Boorstein, W., T. Ziegelhoffer, and E. A. Craig.** 1994. Molecular evolution of the HSP70 multigene family. J. Mol. Evol. **38:**1–17.
6. **Bucca, G., C. P. Smith, M. Alberti, G. Seidita, R. Passantino, and A. M. Puglia.** 1993. Cloning and sequencing of the *dnaK* region of *Streptomyces coelicolor* A3(2). Gene **130:**141–144.
7. **Bukau, B.** 1993. Regulation of the *Escherichia coli* heat-shock response. Mol. Microbiol. **9:**671–680.
8. **Casadaban, M. J., and S. M. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138:**179–207.
9. **Cellier, M. F. M., J. Tessier, M. Nicolas, J. P. Liautard, J. Marti, and J. S. Widada.** 1992. Cloning and characterization of the *Brucella ovis* heat shock protein DnaK functionally expressed in *Escherichia coli*. J. Bacteriol. **174:** 8036–8042.
10. **Chitnis, P. R., and N. Nelson.** 1991. Molecular cloning of the genes encoding two chaperon proteins of the cyanobacterium *Synechocystis* sp. PCC 6803. J. Biol. Chem. **266:**58–65.
11. **Craig, E. A.** 1985. The heat shock response. Crit. Rev. Biochem. **18:**239–280.
12. **Doi, R. H., and L.-F. Wang.** 1986. Multiple procaryotic ribonucleic acid polymerase sigma factors. Microbiol. Rev. **50:**227–243.
13. **Duchene, A. M., C. J. Thompson, and P. Mazodier.** 1994. Transcription analysis of *groEL* genes in *Streptomyces coelicolor* A3(2). Mol. Gen. Genet. **245:**61–68.
14. **Falah, M., and R. S. Gupta.** 1994. Cloning of the *hsp70* (*dnaK*) genes from *Rhizobium meliloti* and *Pseudomonas cepacia*: phylogenetic analysis of the mitochondrial origin based on a highly conserved protein sequence. J. Bacteriol. **176:**7748–7753.
15. **Ferreyra, R. G., F. C. Soncini, and A. M. Viale.** 1993. Cloning, characterization and functional expression in *Escherichia coli* of chaperonin (*groESL*) genes from the phototrophic sulfur bacterium *Chromatium vinosum*. J. Bacteriol. **175:**1514–1523.
16. **Fischer, H. M., M. Babst, T. Kaspar, G. Acuna, F. Arigoni, and H. Hennecke.** 1993. One member of a *groESL*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. EMBO J. **12:**2901–2912.
17. **Gardella, T.** 1989. A mutant *Escherichia coli* sigma-70 subunit of RNA polymerase with altered promoter specificity. J. Mol. Biol. **206:**579–590.
18. **Garvin, L. D., and S. C. Hardies.** 1989. Nucleotide sequence for the htpR gene from *Citrobacter freundii*. Nucleic Acids Res. **17:**4889.
19. **Gomes, S. L., J. W. Gober, and L. Shapiro.** 1990. Expression of the *Caulobacter* heat shock gene *dnaK* is developmentally controlled during growth at normal temperatures. J. Bacteriol. **172:**3051–3059.
20. **Gupta, R. S., and B. Singh.** 1992. Cloning of the HSP70 gene from *Halobac-*

terium marismortui: relatedness of archaebacterial HSP70 to its eubacterial homologs and a model for the evolution of the HSP70 gene. J. Bacteriol. **174:**4594–4605.
21. **Li, M., and S.-L. Wong.** 1992. Cloning and characterization of the *groESL* operon from *Bacillus subtilis*. J. Bacteriol. **174:**3981–3992.
22. **Lin, J., L. G. Adams, and T. A. Ficht.** 1992. Characterization of the heat shock response in *Brucella abortus* and isolation of the genes encoding the GroE heat shock proteins. Infect. Immun. **60:**2425–2431.
23. **Lonetto, M., M. Gribskov, and C. A. Gross.** 1992. The σ^{70} family: sequence conservation and evolutionary relationships. J. Bacteriol. **174:**3843–3849.
24. **Mantis, N. J., and S. C. Winans.** 1992. Characterization of the *Agrobacterium tumefaciens* heat shock response: evidence for a σ^{32}-like sigma factor. J. Bacteriol. **174:**991–997.
25. **McPherson, M. J., K. M. Jones, and S. J. Gurr.** 1991. PCR with highly degenerate primers, p. 171–186. *In* M. J. McPherson, P. Quirke, and G. R. Taylor (ed.), PCR: a practical approach. IRL Press, Oxford.
26. **Michel, G. P. F.** 1993. Cloning and expression in *Escherichia coli* of the *dnaK* gene of *Zymomonas mobilis*. J. Bacteriol. **175:**3228–3231.
27. **Morrison, R. P., R. J. Belland, K. Lyng, and H. D. Caldwell.** 1989. Chlamydial disease pathogenesis. The 70-kD chlamydial hypersensitivity antigen is a stress response protein. J. Exp. Med. **170:**1271–1283.
28. **Narberhaus, F., and H. Bahl.** 1992. Cloning, sequencing, and molecular analysis of the *groESL* operon of *Clostridium acetobutylicum*. J. Bacteriol. **174:**3282–3289.
29. **Narberhaus, F., K. Giebeler, and H. Bahl.** 1992. Molecular characterization of the *dnaK* gene region of *Clostridium acetobutylicum*, including *grpE*, *dnaJ*, and a new heat shock gene. J. Bacteriol. **174:**3290–3299.
30. **Neidhardt, F. C., and R. A. VanBogelen.** 1987. Heat shock response, p. 1334–1357. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
31. **Nimura, K., H. Yoshikawa, and H. Takahashi.** 1994. Identification of *dnaK* multigene family in *Synechococcus* sp. PCC7942. Biochem. Biophys. Res. Commun. **201:**466–471.
32. **Nimura, K., H. Yoshikawa, and H. Takahashi.** 1994. Sequence analysis of the third *dnaK* homolog gene in *Synechococcus* sp. PCC7942. Biochem. Biophys. Res. Commun. **202:**848–854.
33. **Olsen, G. J., C. R. Woese, and R. Overbeek.** 1994. The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. **176:**1–6.
34. **Riethdorf, S., U. Völker, U. Gerth, A. Winkler, S. Engelmann, and M. Hecker.** 1994. Cloning, nucleotide sequence, and expression of the *Bacillus subtilis lon* gene. J. Bacteriol. **176:**6518–6527.
35. **Ronald, S., M. A. Farinha, B. J. Allan, and A. M. Kropinski.** 1992. Cloning and physical mapping of transcriptional regulatory (sigma) factors from *Pseudomonas aeruginosa*, p. 249–257. *In* E. Galli, S. Silver, and B. Witholt (ed.), *Pseudomonas* molecular biology and biotechnology. American Society for Microbiology, Washington, D.C.
36. **Rusanganwa, E., and R. S. Gupta.** 1993. Cloning and characterization of multiple *groEL* chaperonin-encoding genes in *Rhizobium meliloti*. Gene **126:** 67–75.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
38. **Schmidt, A., M. Schiesswohl, U. Völker, M. Hecker, and W. Schumann.** 1992. Cloning, sequencing, mapping and transcriptional analysis of the *groESL* operon from *Bacillus subtilis*. J. Bacteriol. **174:**3993–3999.
39. **Segal, G., and E. Z. Ron.** 1993. Cloning, sequencing, and transcriptional analysis of the gene coding for the vegetative sigma factor of *Agrobacterium tumefaciens*. J. Bacteriol. **175:**3026–3030.
40. **Segal, G., and E. Z. Ron.** 1993. Heat shock transcription of the *groESL* operon of *Agrobacterium tumefaciens* may involve a hairpin-loop structure. J. Bacteriol. **175:**3083–3088.
41. **Segal, G., and E. Z. Ron.** 1995. The *groESL* operon of *Agrobacterium tumefaciens*: evidence for heat shock-dependent mRNA cleavage. J. Bacteriol. **177:**750–757.
42. **Siegele, D. H., J. C. Hu, W. A. Walter, and C. A. Gross.** 1989. Altered promoter recognition by mutant forms of the sigma-70 subunit of *Escherichia coli* RNA polymerase. J. Mol. Biol. **206:**591–603.
43. **Silver, J.** 1991. Inverse polymerase chain reaction, p. 137–146. *In* M. J. McPherson, P. Quirke, and G. R. Taylor (ed.), PCR: a practical approach. IRL Press, Oxford.
44. **van Asseldonk, M., A. Simons, H. Visser, W. M. de Vos, and G. Simons.** 1993. Cloning, nucleotide sequence, and regulatory analysis of the *Lactococcus lactis dnaJ* gene. J. Bacteriol. **175:**1637–1644.
45. **VanBogelen, R. A., P. M. Kelley, and F. C. Neidhardt.** 1987. Differential induction of heat shock, SOS, and oxidative stress regulons and accumulation of nucleotides in *Escherichia coli*. J. Bacteriol. **169:**26–32.
46. **Waldburger, C., T. Gardella, R. Wong, and M. M. Susskind.** 1990. Changes in conserved region 2 of *Escherichia coli* sigma-70 affecting promoter recognition. J. Mol. Biol. **215:**267–276.
47. **Webb, R., K. J. Reddy, and L. A. Sherman.** 1990. Regulation and sequence of the *Synechococcus* sp. strain PCC 7942 *groESL* operon, encoding a cyanobacterial chaperonin. J. Bacteriol. **172:**5079–5088.
48. **Wetzstein, M., U. Völker, J. Dedio, S. Löbau, U. Zuber, M. Schiesswohl, C. Herget, M. Hecker, and W. Schumann.** 1992. Cloning, sequencing, and molecular analysis of the *dnaK* locus from *Bacillus subtilis*. J. Bacteriol. **174:**3300–3310.
49. **Wise, A. A., and C. W. Price.** 1995. Four additional genes in the *sigB* operon of *Bacillus subtilis* that control activity of the general stress factor σ^B in response to environmental signals. J. Bacteriol. **177:**123–133.
50. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51:**221–271.
51. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
52. **Yura, T., H. Nagai, and H. Mori.** 1993. Regulation of the heat-shock response in bacteria. Annu. Rev. Microbiol. **47:**321–350.
53. **Zeilstra-Ryalls, J., O. Fayet, and C. Georgopoulos.** 1991. The universally conserved GroE (Hsp60) chaperonins. Annu. Rev. Microbiol. **45:**301–325.
54. **Zuber, U., and W. Schumann.** 1994. CIRCE, a novel heat shock element involved in regulation of heat shock operon *dnaK* of *Bacillus subtilis*. J. Bacteriol. **176:**1359–1363.