JOURNAL OF BACTERIOLOGY, Feb. 1997, p. 1153–1164 Vol. 179, No. 4
0021-9193/97/$04.00+0

The *dnaK* Operon of *Bacillus subtilis* Is Heptacistronic

GEORG HOMUTH,[1] SHOKO MASUDA,[2] AXEL MOGK,[1] YASUO KOBAYASHI,[2] AND WOLFGANG SCHUMANN[1]*

Institute of Genetics, University of Bayreuth, D-95440 Bayreuth, Germany,[1] and Faculty of Agriculture, Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183, Japan[2]

Received 7 October 1996/Accepted 26 November 1996

In 1992, we described the cloning and sequencing of the *dnaK* locus of *Bacillus subtilis* which, together with transcriptional studies, implied a tetracistronic structure of the operon consisting of the genes *hrcA*, *grpE*, *dnaK*, and *dnaJ*. We have repeated the Northern blot analysis, this time using riboprobes instead of oligonucleotides, and have detected a heat-inducible 8-kb transcript, suggesting the existence of additional heat shock genes downstream of *dnaJ*. Cloning and sequencing of that region revealed the existence of three novel heat shock genes named *orf35*, *orf28*, and *orf50*, extending the tetra- into a heptacistronic operon. This is now the largest *dnaK* operon to be described to date. The three new genes are transcribed as a part of the entire *dnaK* operon (8.0-kb heptacistronic heat-inducible transcript) and as part of a suboperon starting at an internal vegetative promoter immediately upstream of *dnaJ* (4.3-kb tetracistronic non-heat-inducible transcript). In addition, the Northern blot analysis detected several processing products of these two primary transcripts. To demonstrate the existence of the internal promoter, a DNA fragment containing this putative promoter structure was inserted upstream of a promoterless *bgaB* gene, resulting in the synthesis of β-galactosidase. Challenging this transcriptional fusion with various stress factors did not result in the activation of this promoter. To assign a biological function to the three novel genes, they have each been inactivated by the insertion of a *cat* cassette. All of the mutants were viable, and furthermore, these genes are (i) not essential for growth at high temperatures, (ii) not involved in the regulation of the heat shock response, and (iii) sporulation proficient. Blocking transcription of the suboperon from the upstream heat-inducible promoter did not impair growth and viability at high temperatures.

The heat shock response occurs when cells growing at a low temperature are shifted to a higher temperature and results in the induction of a subset of proteins called heat shock proteins. This response to temperature upshift is universal among prokaryotes and eukaryotes. The heat shock family of proteins has recently become central to the study of correct folding of nascent polypeptides, assembly of protein complexes, and translocation of proteins into organelles (13). Therefore, some heat shock proteins are widely recognized as molecular chaperones that play crucial roles even under physiological conditions (10, 11).

Escherichia coli has some 31 heat shock genes whose expression is regulated positively at the transcriptional level by the σ^{32} polypeptide, the *rpoH* gene product. σ^{32} is extremely unstable in vivo, with a half-life of ≈1 min at normal temperatures (for recent reviews, see references 3, 27, and 49).

In contrast to *E. coli*, there are at least three classes of heat shock genes in *Bacillus subtilis* (for recent reviews, see references 16 and 40). Class I heat shock genes constitute the CIRCE regulon, and these genes are negatively regulated by a repressor (38, 47). Class II heat shock genes are positively regulated by the alternative sigma-B factor and form the sigma-B regulon. Heat shock genes which do not belong to either class I or class II have been assigned to class III, whose members include *lon* (33), *clpC* (24), *ftsH* (8), and *htpG* (37). The regulatory mechanism of these genes is largely unknown.

Currently, the CIRCE regulon consists of two operons, the *dnaK* and *groESL* operons. Cloning, sequencing, and transcriptional analyses of the *dnaK* operon revealed four genes in the order *hrcA-grpE-dnaK-dnaJ*, where *hrcA* (formerly *orf39*) encodes the repressor of the CIRCE regulon (47), while the products of the remaining genes constitute the *dnaK* chaperone machine (46). Since a typical transcriptional terminator sequence was not detected downstream of the *dnaK* operon, we speculated that this operon might contain additional genes (46). To check for additional genes, we first used oligonucleotides complementary to the *dnaJ* transcript but failed to detect a band in Northern blots (19). We have now repeated these experiments, this time using a *dnaJ* riboprobe, and succeeded in demonstrating a heat-inducible transcript of about 8 kb. This strongly suggests the existence of additional genes downstream of *dnaJ* as part of the *dnaK* operon. This was confirmed by cloning and sequencing the region downstream of *dnaJ*, leading to the detection of three novel genes within the *dnaK* operon. Transcriptional analyses revealed an internal vegetative promoter from which the four distal genes were transcribed constitutively. Inactivation of the three novel genes revealed no obvious phenotype under various growth conditions.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. The bacterial strains and plasmids used in this study are given in Table 1. M13 phage vectors were used to prepare randomly overlapping libraries for shotgun sequencing. Bacteria were routinely grown aerobically at 37°C in Luria broth. Ampicillin (200 μg ml^{-1}) was included for all plasmid-bearing *E. coli* strains. Chloramphenicol and neomycin were added at concentrations of 5 and 10 μg ml^{-1}, respectively.

DNA manipulations and analysis. Plasmid DNA was purified on columns (QIAGEN, Hilden, Germany). PCR products containing the three distal genes of the *dnaK* operon were generated with *Taq* DNA polymerase as specified by the manufacturer (Eurobio, Raunheim, Germany). Chromosomal DNA of *B. subtilis* was used as the template, and *Bam*HI recognition sequences were included at the 5′ ends of the PCR primers. PCR products were purified with the QIAGEN PCR purification kit. Cloning procedures were carried out by standard

* Corresponding author. Mailing address: Institute of Genetics, University of Bayreuth, D-95440 Bayreuth, Germany. Phone: 49-921-552708. Fax: 49-921-2710. E-mail: wolfgang.schumann@uni-bayreuth.de.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Relevant genotype or host	Relevant phenotype or marker(s)	Property(ies)	Reference or source
Strains				
E. coli				
DH10B			Host for detecting chimeric plasmids	35
JM105			Host for M13 propagation	35
B. subtilis				
1012	*leuA8 metB5 trpC2 hsdrM1*			34
BT01	*hrcA*::*cat*	Cm^r	Derivative of 1012	39
BT02	*dnaK*::*cat*	Cm^r	Derivative of 1012	39
AS01	Δ*hrcA*	Cm^r	Derivative of 1012	38
JH642				6
AM01	*amyE*::*cat*	Cm^r	Derivative of 1012	7
AM03	*dnaK*::*bgaB* operon fusion at *amyE*	Neo^r	Derivative of 1012	7
AM04	*ctc*::*bgaB* operon fusion at *amyE*	Neo^r	Derivative of 1012	7
AM05	*clpC*::*bgaB* operon fusion at *amyE*	Neo^r	Derivative of 1012	7
AM13	*dnaJ*::*bgaB* operon fusion inserted at *amyE*	Neo^r	Derivative of 1012	This work
GH01	*orf35*::*cat*	Cm^r	Derivative of 1012	This work
GH02	*orf28*::*cat*	Cm^r	Derivative of 1012	This work
GH03	*orf50*::*cat*	Cm^r	Derivative of 1012	This work
GH04	*′dnaK-cat-dnaJ′*	Cm^r	*cat* cassette inserted between the potential terminator and the promoter P_{A2}	This work
Plasmids				
pBluescript SK^+ II	*E. coli*	Ap^r	Cloning and riboprobe vector	Stratagene
pBlue*Sal*I*Sma*IΔ	*E. coli*	Ap^r	pBluescript SK^+ II with deletion of the polylinker region between and including *Sal*I and *Sma*I	This work
pUC18Cm-S	*E. coli*	Ap^r Cm^r	*cat* cassette of pC194 cloned into pUC18	20
pSED377	*E. coli*	Ap^r	840-bp internal *Bgl*II fragment of *dnaK* inserted into pSPT18	44
pMWC862	*E. coli*	Ap^r Cm^r	Derivative of pJH101 containing *dnaK* through the 5′ end of *dnaJ*	45
p*dnaJ*	*E. coli*	Ap^r	pBluescript SK^+ II with PCR-amplified *dnaJ* inserted at *Bam*HI site	This work
p*orf35*	*E. coli*	Ap^r	pBlue*Sal*I*Sma*IΔ with PCR-amplified *orf35* at *Bam*HI site	This work
p*orf35*::*cat*	*E. coli*	Ap^r Cm^r	*cat* cassette at *Sma*I site of p*orf35*	This work
p*orf28*	*E. coli*	Ap^r	pBlue*Sal*I*Sma*IΔ with PCR-amplified *orf28* at *Bam*HI site	This work
p*orf28*::*cat*	*E. coli*	Ap^r Cm^r	*cat* cassette at *Hin*dIII site of p*orf28*	This work
p*orf50*	*E. coli*	Ap^r	pBluescript SK^+ II with PCR-amplified *orf50* at *Eco*RI site	This work
p*orf50*::*cat*	*E. coli*	Ap^r Cm^r	*cat* cassette at *Stu*I site of p*orf50*	This work
p′*dnaK-dnaJ′*	*E. coli*	Ap^r	1,656-bp *Eag*I fragment from pMWC862 cloned into pBlue*Sal*I*Sma*IΔ	This work
p′*dnaK-cat-dnaJ′*	*E. coli*	Ap^r Cm^r	1,348-bp *cat* cassette inserted between *dnaK* and *dnaJ*	This work
pBgaB	*E. coli*	Ap^r Neo^r	Promoter-probe vector containing the *bgaB* gene	28
p*PdnaJ-bgaB*	*E. coli*	Ap^r Neo^r	180-bp *dnaK-dnaJ* intercistronic region cloned into pBgaB	This work

methods (35). Double-stranded sequencing of plasmids and of single-stranded recombinant M13 DNA was performed by the dideoxynucleotide chain termination method (36). All nucleotide sequences were determined with the 373A DNA sequencer (Applied Biosystems), the *Taq* Dye Primer Cycle Sequencing kit, and the *Taq* DyeDeoxy Terminator Cycle Sequencing kit. We have determined the sequences of both strands. Shotgun sequencing methods were described previously (41).

Amplification of the region downstream of the *dnaJ* locus by long PCR. The region downstream of *dnaJ* was amplified by long PCR (15 s at 95°C and 12 min at 65°C [16 cycles], 15 s at 95°C and 12 min [increment = 15 s per cycle] at 65°C [12 cycles], and 10 min at 72°C [one cycle]) with *rTth* DNA polymerase XL (Perkin-Elmer Cetus) for direct sequencing and 0.1 μM (each) primers DNJD (5′-CCTGACGAACAGGAAATGAG-3′) and RPUD (5′-TCCTCATCCTCG GTC-3′).

Data handling and computer analysis. DNA sequence assembly was performed with the ATSQ program (Software Development Co., Ltd.). The compiled sequence was further analyzed for the locations of possible open reading frames (ORFs) with the GENETYX-MAC 6.2.4 program (Software Development Co., Ltd.) and use of the SwissProt Protein Sequence Database. Some sequence analysis was performed with the PC/GENE package (IntelliGenetics) for IBM personal computers and Heidelberg Unix Sequence Analysis Resources.

Construction of chromosomal insertion mutants. The three genes *orf35* (*yqeT*), *orf28* (*yqeU*), and *orf50* (*yqeV*) were generated by PCR with synthetic primers with *Bam*HI restriction sites on their 5′ ends. The products were cleaved by *Bam*HI; *orf50* was inserted into pBluescript SK^+ II, while *orf35* and *orf28* were ligated into a modified version of this vector, pBlue*Sal*I*Sma*IΔ. This version of pBluescript SK^+ II was obtained by digestion with *Sma*I and *Sal*I and isolation of the larger fragment, followed by fill-in of the *Sal*I end and self-ligation. This step removed unique restriction sites which occur in *orf35* and *orf28*, allowing these sites to be used in a second cloning step involving the insertion of a *cat* cassette. The three recombinant plasmids were designated p*orf35*, p*orf28*, and p*orf50*. Insertion mutants were obtained by ligating a *cat* cassette near the 5′ ends

of the three genes. To accomplish this goal, the *cat* cassette was recovered from pUC18Cm-S as a 1,348-bp *Bam*HI-*Hin*dIII fragment, blunt ended, and ligated into the three plasmids after linearization with *Sma*I (p*orf35*); *Hin*dIII (p*orf28*), followed by fill-in the 5′ protruding ends; or *Stu*I (p*orf50*). The transcription of the *dnaK* operon and *cat* genes occurs in the same direction in p*orf28*::*cat* and p*orf50*::*cat*, while they were transcribed in the opposite direction in the case of p*orf35*::*cat*. To obtain a mutant carrying a *cat* cassette at a *Sma*I site located between the potential rho-independent terminator downstream of *dnaK* and the promoter upstream of *dnaJ*, plasmid pMWC862 was digested with *Eag*I, and a 1,656-bp fragment containing the 3′ end of *dnaK*, the 5′ end of *dnaJ*, and the complete intercistronic sequence was recovered and cloned into *Eag*I-linearized pBlue*Sal*I*Sma*IΔ. The resulting plasmid, p′*dnaK-dnaJ*′, was linearized with *Sma*I, followed by insertion of the *cat* cassette. In the resulting plasmid, p′*dnaK-cat-dnaJ*′, the *cat* gene is transcribed in the same direction as the genes of the *dnaK* operon. To integrate the various *cat* cassettes into their respective chromosomal loci, competent cells of *B. subtilis* 1012 were transformed separately with the four plasmids. Transformants were selected on plates containing chloramphenicol, and chromosomal DNA of recombinants was screened either by PCR or by PCR and Southern blotting for replacement recombination at the three genes or at the intercistronic region (data not shown). In each case, one of the recombinants (GH01 = *orf35*::*cat*, GH02 = *orf28*::*cat*, GH03 = *orf50*::*cat*, and GH04 = *′dnaK-cat-dnaJ′*) was kept for further studies.

Construction and analysis of a *bgaB* fusion. To demonstrate transcription starting at the internal vegetative promoter P_{A2} (see Fig. 3) and to monitor expression under different stress conditions, this promoter region (nucleotides 4 to 183; see Fig. 1A) was generated by PCR, and *Bam*HI and *Eco*RI sites were introduced at the 5′ and 3′ ends, respectively, and inserted into the promoter-probe vector pBgaB (28) cut with *Eco*RI and *Bam*HI. This plasmid is a pBR322 derivative and contains the promoterless *bgaB* gene of *Bacillus stearothermophilus* encoding a heat-stable β-galactosidase (17) and a *neo* cassette sandwiched between *amyE*-front and *amyE*-back. The absence of replication in *B. subtilis* allows integration of any DNA sequence at the *amyE* locus of the *B. subtilis* chromosome. The ligation mixture was transformed into *E. coli* DH10B, and transformants carrying recombinant plasmids were detected on X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) indicator plates (60 μg of X-Gal/ml); the correct DNA sequence of the insert was verified by DNA sequencing (pP_{dnaJ}-*bgaB*). Subsequently, the transcriptional fusion was recombined at the *amyE*::*cat* locus of AM01; we selected for neomycin resistance and screened for the loss of the *cat* gene. This procedure ensures true replacement of the *cat* marker by the operon fusion (double crossover event) instead of integration of the whole plasmid (single crossover event). One strain (AM13) was kept for further studies. The BgaB activity was determined as previously described (18). One unit of BgaB enzyme cleaves 1 nmol of ONPG (*o*-nitrophenyl-β-D-galactopyranoside) per min.

Western immunoblot analyses. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was performed according to the method of Laemmli (25), and immunoblot analysis was performed as described elsewhere (19). Polyclonal antibodies were raised against GrpE and DnaJ as described for HrcA (39). Anti-GrpE antibodies were used at a dilution of 1:250, and those raised against DnaJ were used at a dilution of 1:1,000.

Analysis of transcription by slot blot and Northern (RNA) blot hybridization. Total RNA of *B. subtilis* 1012 and the different mutant strains was isolated by the acid-phenol method previously described (44), with the following modifications. The samples were extracted twice with acid phenol–chloroform–isoamy alcohol (25:24:1, vol/vol/vol) and once with chloroform-isoamyl alcohol (24:1, vol/vol); the extraction step with diethyl ether was omitted. Decreasing amounts of total RNA were transferred onto a positively charged nylon membrane (Amersham, Braunschweig, Germany) by slot blotting. After being baked at 120°C for 1 h, the filters were prehybridized, hybridized under stringent conditions at 68°C in 50% formamide and 5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) with digoxigenin (DIG)-labelled RNA probes, and washed, and hybridization signals were detected according to the manufacturer's instructions (Boehringer, Mannheim, Germany) with Fuji RX films. The hybridization signals of the luminographs were quantified with the WinCam software version 2.1 (Cybertech, Berlin, Germany). The induction ratios were calculated by dividing the volume of the signals obtained from RNA isolated from stressed cells by the volume of the signals obtained from RNA isolated from the controls (37°C, exponential growth). For Northern blot analysis, 10- and 5-μg samples of total RNA were separated under denaturing conditions in 0.6% agarose–2.1 M formaldehyde–MOPS (morpholinepropanesulfonic acid) gels, stained with ethidium bromide, and transferred to a Hybond-N membrane (Amersham) by vacuum blotting. Hybridization and detection were carried out as described for slot blots. The hybridizations specific for *dnaK* were carried out with DIG-labelled RNA synthesized in vitro with T7 RNA polymerase from *Hin*dIII-linearized plasmid pSED377 (44). RNA probes specific for *dnaJ* were obtained by in vitro transcription with T7 RNA polymerase from *Eco*RI-linearized plasmid p*dnaJ*. This plasmid contains the full-length *dnaJ* gene amplified by PCR with primers carrying *Bam*HI restriction sites at both ends. The PCR product was cleaved with *Bam*HI and inserted into *Bam*HI-linearized pBluescript SK^+ II. To obtain RNA probes specific for *orf35*, *orf28*, and *orf50*, plasmids p*orf35*, p*orf28*, and p*orf50*, respectively, were used as templates for in vitro transcription according to the manufacturer's instructions (DIG RNA Labelling kit, Boehringer). In the case of p*orf35*, *Eco*RI was used for linearization of the template; p*orf28* was linearized with *Xba*I; and p*orf50* was linearized with *Cla*I. DIG-labelled RNA probes were synthesized with T3 RNA polymerase (p*orf35* and p*orf50*) or T7 RNA polymerase (p*orf28*).

Nucleotide sequence accession number. The EMBL/GenBank/DDBJ accession number for the nucleotide sequence reported here is D83717.

RESULTS

Amplification and sequencing of the region downstream of *dnaJ*. The region downstream of *dnaJ* was amplified by long PCR directly from the chromosome (see Materials and Methods). The DNA sequence was determined by a strategy based on a combination of PCR and random sequencing, i.e., preparation of a random library, amplification of inserts, and filling of the gaps by region-specific primers (31). To avoid errors in nucleotide sequencing, at least two independently prepared PCR fragments were sequenced in both directions, and we used sequence outputs only from regions where base peaks were well separated. The region is 4,637 bp in length, but only the DNA sequences belonging to the *dnaK* operon are given in Fig. 1B.

Features of the three distal genes of the *dnaK* operon. The *dnaJ*-proximal gene called *orf35* (*yqeT*) has a size of 936 bp (Fig. 1B, nucleotides 85 to 1020), encoding a putative polypeptide of 311 amino acids with a predicted molecular mass of approximately 35 kDa. We compared the amino acid sequence of this and the other putative ORFs with known amino acid sequences present in data banks by using the Lipman-Pearson programs (26). The ORF35 protein shows strong homology to the ORF35 protein of *Staphylococcus aureus* (63% similarity, 45% identity) (32) and to ORFB of *Clostridium acetobutylicum* (62% similarity, 42% identity) (1). It seems that the organization of the *dnaK* operon in both organisms is quite similar to that in *B. subtilis*. In all three cases, the genes of the *dnaK* operon are arranged in the order *hrcA-grpE-dnaK-dnaJ-orf35*/*orfB*. In *S. aureus* and *C. acetobutylicum*, the region downstream of *orf35*/*orfB* has not yet been sequenced. Furthermore, the *B. subtilis* ORF35 exhibits significant homology to the ribosomal protein L11 methyltransferase encoded by *prmA* of *E. coli* (57% similarity, 32.5% identity) (42) and to the homologous PrmA protein of *Haemophilus influenzae* (54% similarity, 32% identity) (12).

The second gene, *orf28* (*yqeU*), with a size of 771 bp (Fig. 1B, nucleotides 1043 to 1813), encodes a potential polypeptide of 256 amino acids with a predicted molecular mass of 28 kDa. It exhibits significant homology to the 26.9-kDa protein YggJ from *E. coli* (53% similarity, 32% identity) (14, 23) and to the homologous protein of *H. influenzae* (55% similarity, 28% identity) (12). In both cases, the function of the YggJ protein is unknown. The third gene, *orf50* (*yqeV*), has a size of 1,356 bp (Fig. 1B, nucleotides 1813 to 3168) and encodes a potential protein of 451 amino acid residues with a predicted molecular mass of 50 kDa. The stop codon of *orf28* and the start codon of *orf50* overlap by 1 bp, suggesting translational coupling. The *orf50* gene product shows significant homology (54% similarity, 33.5% identity) to a *Pseudomonas aeruginosa* protein (22) and an *H. influenzae* protein deduced from ORFs (55% similarity, 33% identity) (12). The function of these proteins is unknown.

Putative ribosome-binding sites were identified in front of all three genes (Fig. 1B, marked by asterisks), located 4, 9, and 5 nucleotides upstream from the start codons of *orf35*, *orf28*, and *orf50*, respectively. These sites consist of 8 (*orf35*), 7 (*orf28*), and 9 (*orf50*) nucleotides complementary to the 3′ end of *B. subtilis* 16S rRNA (29).

The *dnaK* operon of *B. subtilis* is heptacistronic and transcribed from two independent promoters. Inspection of the

A

```
TAAGTTCTTTTTAGTGTCAGCCCCGCTTCCGGAGCTGACCGAAAAGAACACATTTCATAA      60
dnaK
                >>>>>>>>>>>>>>>>>  2  <<<<<<<<<<<<<<<<<
TCTGATTCAATGATTAGAAAGTCAAAGTCAGGCATCTCTTGGCTTTGACTTTTTTTCTTG     120

 SmaI             -35          18 bp       -10
CCCGGGATAAAAGGAATTGAAAAATCATAATTCAAAATGATACAATCTAATTTATGTGAG     180

        *   ****
AAATTCGGGAGAGTGAAGCGAGATG
                     dnaJ
```

B

```
dnaJ
CCTGACGAACAGGAAATGAGTTTCTTTGACAAGGTAAAACGCGCGTTTAAAGGCGATTAA      60
P  D  E  Q  E  M  S  F  F  D  K  V  K  R  A  F  K  G  D  *

          *****  **  *    orf35
TTATTGGATGAGGAGTTGGTAGTTTTGAAATGGTCCGAATTAAGCATTCACACAACACAT     120
                        M  K  W  S  E  L  S  I  H  T  T  H

GAAGCGGTCGAACCTATCTCAAATATATTGCATGAAGCTGGTGCAAGTGGGGTTGTGATA     180
E  A  V  E  P  I  S  N  I  L  H  E  A  G  A  S  G  V  V  I

GAGGACCCGCTTGATTTAATTAAAGAACGTGAGAATGTGTACGGGGAAATCTACCAGCTC     240
E  D  P  L  D  L  I  K  E  R  E  N  V  Y  G  E  I  Y  Q  L

GACCCCAATGATTACCCAGATGAGGGTGTCATTGTCAAAGCATATCTGCCGGTTAACAGT     300
D  P  N  D  Y  P  D  E  G  V  I  V  K  A  Y  L  P  V  N  S

TTTCTGGGCGAAACTGTAGACGGCATCAAAGAAACGATTAATAATCTCCTTCTTTACAAT     360
F  L  G  E  T  V  D  G  I  K  E  T  I  N  N  L  L  L  Y  N

ATTGATTTGGGCAGAAACCACATCACTATTTCTGAAGTAAATGAAGAAGAGTGGGCGACT     420
I  D  L  G  R  N  H  I  T  I  S  E  V  N  E  E  E  W  A  T

GCCTGGAAAAAGTATTATCATCCTGTGAAAATTTCTGAAAAGTTTACAATTGTGCCGACG     480
A  W  K  K  Y  Y  H  P  V  K  I  S  E  K  F  T  I  V  P  T
                                                   SmaI
TGGGAGGAATATACGCCGGTCCATACTGATGAACTGATTATTGAAATGGACCCCGGGAATG     540
W  E  E  Y  T  P  V  H  T  D  E  L  I  I  E  M  D  P  G  M

GCTTTCGGCACAGGGACACATCCGACAACCGTACTCTGCATTCAGGCGCATGAACGCTTT     600
A  F  G  T  G  T  H  P  T  T  V  L  C  I  Q  A  H  E  R  F

                                                >>>>>>>>>>>>
GTGCAAAAGGGTGATAAGGTGATTGATGTCGGCACTGGTTCGGGAATTTTAAGTATTGCG     660
V  Q  K  G  D  K  V  I  D  V  G  T  G  S  G  I  L  S  I  A

 3 <<<<<<<<<<<<
GCTGCAATGCTTGAGGCTGAATCGGTTCATGCCTACGATCTTGATCCCGTGGCTGTAGAA     720
A  A  M  L  E  A  E  S  V  H  A  Y  D  L  D  P  V  A  V  E

AGTGCACGCCTCAATCTCAAGCTGAACAAAGTCAGCGATATTGCTCAAGTGAAGCAAAAC     780
S  A  R  L  N  L  K  L  N  K  V  S  D  I  A  Q  V  K  Q  N

AATTTGTTAGACGGGATTGAAGGGGAACATGATGTCATTGTGGCCAACATTTTGGCTGAA     840
N  L  L  D  G  I  E  G  E  H  D  V  I  V  A  N  I  L  A  E
```

FIG. 1. (A) Nucleotide sequence of the intercistronic region between *dnaK* and *dnaJ*. Shown are the stop codon of *dnaK*, the start codon of *dnaJ* (both in boldface letters), the *Sma*I recognition site (in boldface and italic letters) used to insert a *cat* cassette, a potential rho-independent terminator (numbered 2 within the arrowheads above the sequence), a putative σ^A-dependent promoter (boldface letters and indicated by −35 and −10), and the potential Shine-Dalgarno sequence of *dnaJ* (asterisks above the sequence). (B) Nucleotide sequence of a 3,207-bp region encoding the three distal genes of the *dnaK* operon. Nucleotides are numbered from the 5′ end of the nontranscribed strand, and the predicted amino acid sequences of the three genes *orf35*, *orf28*, and *orf50* are given in single-letter code below the DNA sequence. Indicated are the potential Shine-Dalgarno sequences (asterisks above the sequence), inverted repeats (numbered 3 to 6 within the arrowheads above the sequence), three restriction enzyme recognition sites (boldface and italic letters) used to construct insertion mutants, and the start and stop codons (boldface letters). The numbered arrowheads correspond to those shown in Fig. 3 as stem-loop structures.

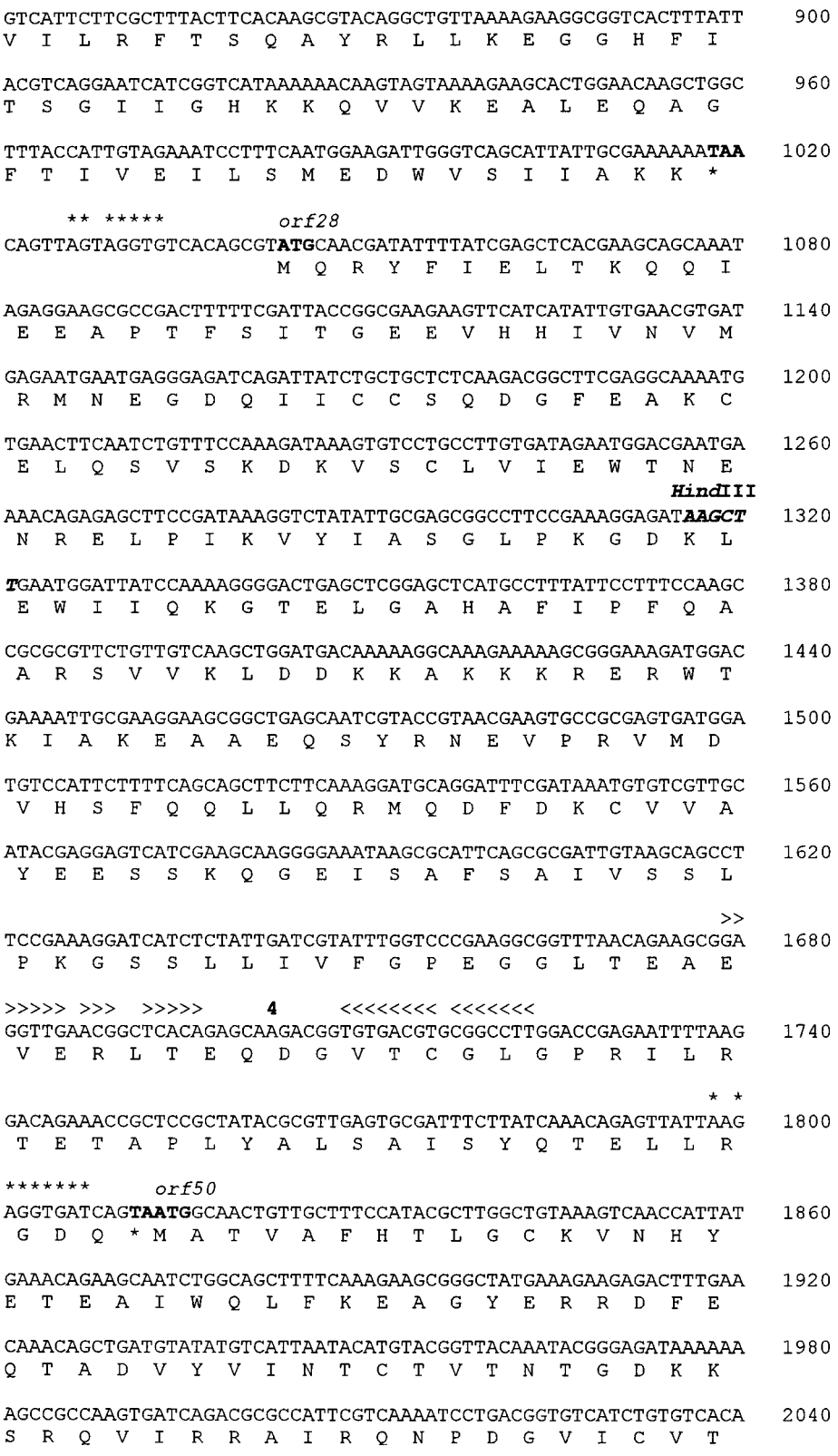

FIG. 1—*Continued.*

```
GGCTGCTACGCGCAAACCTCACCTGCTGAGATTATGGCGATCCCCGGTGTCGACATTGTA   2100
G  C  Y  A  Q  T  S  P  A  E  I  M  A  I  P  G  V  D  I  V

GTCGGAACACAGGACCGTGAAAAAATGCTTGGCTACATTGACCAATACCGAGAAGAACGC   2160
V  G  T  Q  D  R  E  K  M  L  G  Y  I  D  Q  Y  R  E  E  R

                          >> >>>>>>>>   >>>>>>     5     <<<<<
CAGCCGATTAATGGCGTCAGCAACATCATGAAAGCAAGAGTATATGAAGAGCTTGATGTG   2220
Q  P  I  N  G  V  S  N  I  M  K  A  R  V  Y  E  E  L  D  V

< <<<<<<<< <<
CCTGCTTTCACGGACCGCACAAGAGCGTCATTGAAAATACAGGAGGGCTGCAATAATTTC   2280
P  A  F  T  D  R  T  R  A  S  L  K  I  Q  E  G  C  N  N  F
                                 StuI
TGCACATTCTGTATCATTCCGTGGGCGAGAGGCCTGCTCCGCTCCCGCGACCCTGAAGAA   2340
C  T  F  C  I  I  P  W  A  R  G  L  L  R  S  R  D  P  E  E

GTCATTAAACAAGCGCAGCAGCTTGTCGATGCCGGATATAAGGAAATTGTTCTTACGGGT   2400
V  I  K  Q  A  Q  Q  L  V  D  A  G  Y  K  E  I  V  L  T  G

ATTCATACAGGCGGCTACGGCGAAGATATGAAAGATTACAACTTCGCGAAGCTTTTGAGC   2460
I  H  T  G  G  Y  G  E  D  M  K  D  Y  N  F  A  K  L  L  S

GAATTAGACACCCGTGTGGAAGGCGTGAAACGGATCAGAATTTCTTCAATCGAAGCCAGC   2520
E  L  D  T  R  V  E  G  V  K  R  I  R  I  S  S  I  E  A  S

CAAATCACTGACGAAGTGATCGAGGTCTTAGACCGCTCAGATAAGATAGTGAATCATTTG   2580
Q  I  T  D  E  V  I  E  V  L  D  R  S  D  K  I  V  N  H  L

CATATTCCTATCCAATCCGGATCAAACACAGTGCTTAAACGGATGAGACGCAAATATACA   2640
H  I  P  I  Q  S  G  S  N  T  V  L  K  R  M  R  R  K  Y  T

ATGGAATTCTTTGCAGACCGTTTAAATAAATTGAAAAAAGCACTGCCAGGCTTAGCTGTT   2700
M  E  F  F  A  D  R  L  N  K  L  K  K  A  L  P  G  L  A  V

ACTTCTGATGTCATCGTCGGGTTCCCTGGAGAAACAGAAGAAGAATTTATGGAAACCTAT   2760
T  S  D  V  I  V  G  F  P  G  E  T  E  E  E  F  M  E  T  Y

AACTTTATTAAAGAACATAAATTCTCTGAGCTGCATGTTTTCCCTTACAGTAAGCGTACA   2820
N  F  I  K  E  H  K  F  S  E  L  H  V  F  P  Y  S  K  R  T

GGAACACCAGCTGCACGAATGGAAGACCAAGTGGATGAAAACGTGAAAAACGAACGTGTG   2880
G  T  P  A  A  R  M  E  D  Q  V  D  E  N  V  K  N  E  R  V

CACCGCCTGATTGCTCTTTCTGACCAGCTTGCAAAAGAATACGCTTCTCAGTATGAAAAT   2940
H  R  L  I  A  L  S  D  Q  L  A  K  E  Y  A  S  Q  Y  E  N

GAAGTGCTTGAGATTATCCCTGAGGAAGCGTTTAAGGAAACAGAAGAAGAAAATATGTTT   3000
E  V  L  E  I  I  P  E  E  A  F  K  E  T  E  E  E  N  M  F

GTCGGTTACACAGACAATTATATGAAGGTTGTTTTCAAAGGTACGGAAGATATGATCGGC   3060
V  G  Y  T  D  N  Y  M  K  V  V  F  K  G  T  E  D  M  I  G

AAAATCGTAAAAGTCAAAATCCTAAAAGCAGGTTATCCTTACAACGAAGGACAGTTTGTC   3120
K  I  V  K  V  K  I  L  K  A  G  Y  P  Y  N  E  G  Q  F  V

                                                  >>>>>>>>>
CGTGTGGTTGAGGATGAAATAACTGAACACATGCGTTTGTCTTCTTAACCTAAAAAATCG   3180
R  V  V  E  D  E  I  T  E  H  M  R  L  S  S  *

>> >>>> 6<<<< <<<<<<<<<<
GTGCATTAAAATGTACCGATTTTTTTA
```

FIG. 1—*Continued.*

DNA sequence suggested that the three distal genes downstream of *dnaJ* are part of the *dnaK* operon, because they are not separated by large noncoding regions and because of the apparent absence of rho-independent terminators. To address this issue and to test for heat inducibility, Northern blotting was performed with total RNA prepared from wild-type strain 1012 and hybridized with DIG-labelled riboprobes representing genes *dnaK* through *orf50* (Fig. 2A through E). All blots

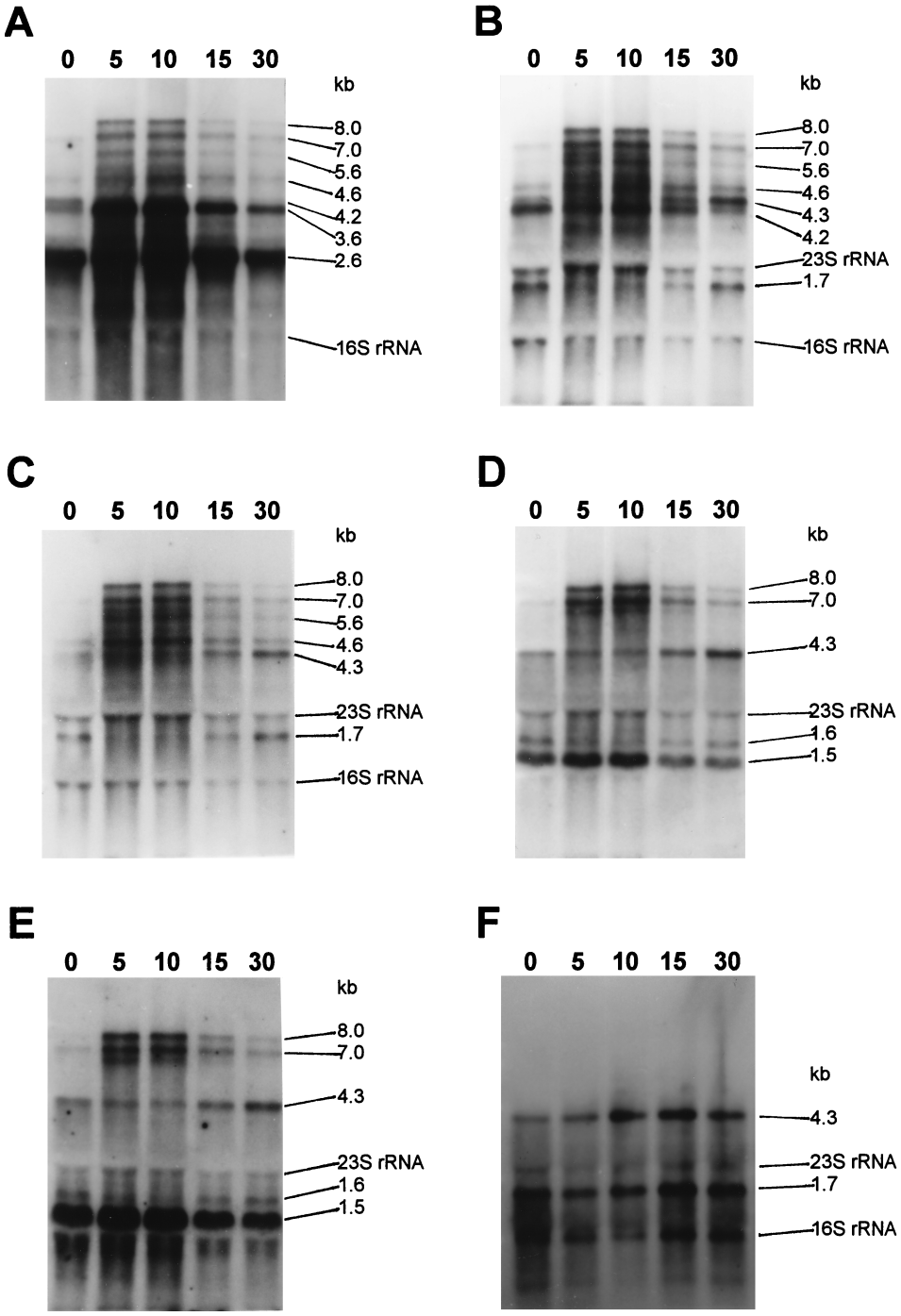

FIG. 2. Northern blot analyses. RNA was isolated from *B. subtilis* wild-type 1012 (A to E) and *dnaK*::*cat* BT02 (F) grown at 37°C in Luria broth (0 min) and at the indicated times after a temperature shift to 48°C. Ten micrograms of RNA was applied per lane, and the filters were hybridized to gene-specific riboprobes *dnaK* (A), *dnaJ* (B), *orf35* (C), *orf28* (D), *orf50* (E), and *dnaJ* (F). Transcript size was determined by comparison with RNA size markers (Gibco BRL, Eggenstein, Germany).

resolved bands of different lengths which represent primary transcripts, cleavage products, and additional bands due to hybridization to degradation products which were trapped by the rRNAs (see Discussion). The transcriptional organization of the complete *dnaK* operon, as deduced from the Northern blots and from additional experiments reported below, and the classification of the different bands are shown diagramatically in Fig. 3.

With the *dnaK* riboprobe, two large transcripts of 8.0 and 7.0 kb, barely detectable at 37°C and transiently induced after thermal upshock, were resolved (Fig. 2A). Whereas the 8.0-kb transcript is likely to represent the complete heptacistronic mRNA, the 7.0-kb transcript could be a cleavage product of the primary transcript lacking *hrcA*. If this assumption is correct, this transcript should not be detectable with an *hrcA*-specific riboprobe, and this was shown to be the case. Instead,

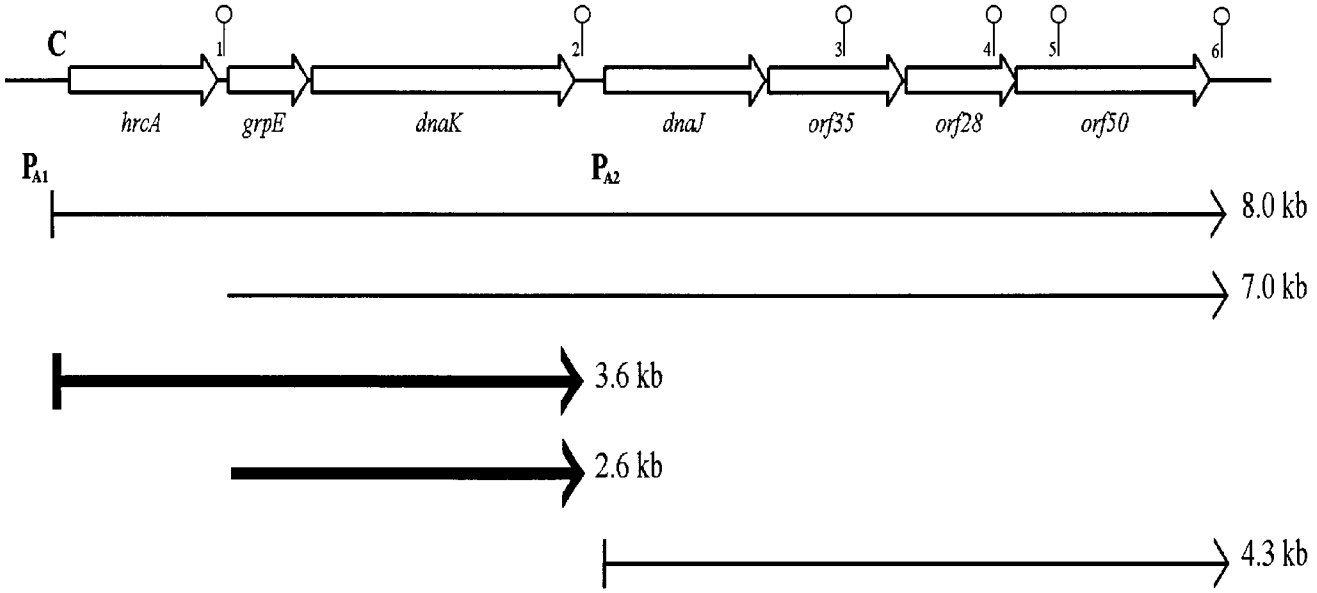

FIG. 3. Genetic map of the *dnaK* operon. The lengths of the various transcripts, as deduced from the Northern blot experiments, are indicated. P_{A1}, σ^A-dependent promoter negatively regulated by HrcA; P_{A2}, internal σ^A-dependent promoter not subject to heat regulation; C, CIRCE element (binding site for the HrcA repressor). Potential stem-loop structures numbered 1 to 6 are indicated. Primary transcripts are drawn as arrows with vertical bars, and processing products are drawn as arrows without vertical bars; the thickness of the arrow reflects the abundance of the transcripts.

a 1-kb transcript representing the monocistronic *hrcA* mRNA was demonstrated (19). Bands of 5.6, 4.6 and 4.2 kb, exhibiting heat inducibility, probably represent cleavage products of the 8.0- and/or 7.0-kb transcript. Analysis of the DNA sequence of the *dnaK* operon revealed several inverted repeats which could form potential stem-loop structures and thereby serve as cleavage and/or transcription termination sites. These are shown in Fig. 3 as stem-loop structures above the genes and are labelled 1 through 6. We suggest that stem-loop structures 2 and 6 serve as rho-independent transcription terminators (they occur outside the structural genes), whereas stem-loop structure 1 represents a potential processing site and stem-loop structures 3, 4, and 5 constitute degradation sites (they are situated within structural genes). Therefore, the 4.6-kb transcript could represent a cleavage product extending from stem-loop structure 1 through stem-loop structure 3. The heat-inducible bands of 3.6 and 2.6 kb represent the tricistronic transcript *hrcA* through *dnaK* initiated at P_{A1} and terminated at stem-loop structure 2 (3.6-kb band) and a cleavage product lacking *hrcA* (2.6-kb band) (Fig. 3).

Total RNA was also hybridized to the *dnaJ*-specific riboprobe. As expected, the transcripts of 8.0 kb (full-length transcript) and 7.0, 5.6, and 4.2 kb (cleavage products) were resolved (Fig. 2B). In addition, a new band of 4.3 kb appeared, which slightly increased in density 30 min after heat induction. We conclude that this transcript was initiated at the internal promoter P_{A2} and represents a tetracistronic transcript encompassing the four distal genes *dnaJ* through *orf50*. The transcript at 1.7 kb, not induced by heat, could represent a cleavage product of the 4.3-kb transcript ending at stem-loop structure 3 (Fig. 3).

When riboprobes specific for *orf35*, *orf28*, and *orf50* were used, hybridization bands comparable to those obtained with the *dnaJ*-specific riboprobe were resolved (Fig. 2C through E). In all four Northern blots, the heat-inducible 8.0- and 7.0-kb transcripts and the 4.3-kb transcript which exhibited a slight increase 30 min after heat induction were revealed. The amounts of the putative cleavage products decreased when riboprobes downstream of *dnaJ* were used to detect transcripts (Fig. 2, compare panels B through E). This behavior underlines our interpretation that these additional bands represent cleaved products rather than primary transcripts or nonspecific hybridization products.

The three distal genes of the *dnaK* operon are heat inducible. The Northern blot experiments strongly suggested that the three distal genes *orf35*, *orf28*, and *orf50* are heat inducible, since the amount of the two large mRNAs (8 and 7 kb) was transiently enhanced after heat shock. To verify this finding and to calculate the extent of heat induction, slot blot analyses with DIG-labelled riboprobes specific for the three distal genes were carried out. Furthermore, riboprobes specific for *dnaJ* and *dnaK* mRNA were used.

As shown in Fig. 4, all four genes downstream of *dnaK* were clearly heat inducible. However, while transcription of *dnaK* was transiently induced six- to sevenfold, the four downstream genes were activated only three- to fourfold (Fig. 4). This difference is likely to be due to a large portion of the heat-induced transcription products starting at the CIRCE-controlled promoter P_{A1} upstream of *hrcA* (Fig. 3) and terminating at the potential rho-independent terminator between *dnaK* and *dnaJ*. Furthermore, the basal level of transcription starting at the constitutive P_{A2} promoter is increased in relation to the transcription initiated at the repressor-regulated P_{A1} promoter (compare Fig. 2A and B, $t = 0$). The results of the slot blot analysis are therefore in good agreement with the model of transcriptional organization of the *dnaK* operon derived from the sequence data and the Northern blot analyses.

The four distal genes of the *dnaK* operon are transcribed constitutively from an internal promoter preceding *dnaJ*. Northern blot analyses led to the identification of a 4.3-kb mRNA which was not inducible by heat and which could be detected only by using riboprobes with specificity for the four distal genes. Since the length of this transcript corresponds to the length of these genes, we assumed the presence of an

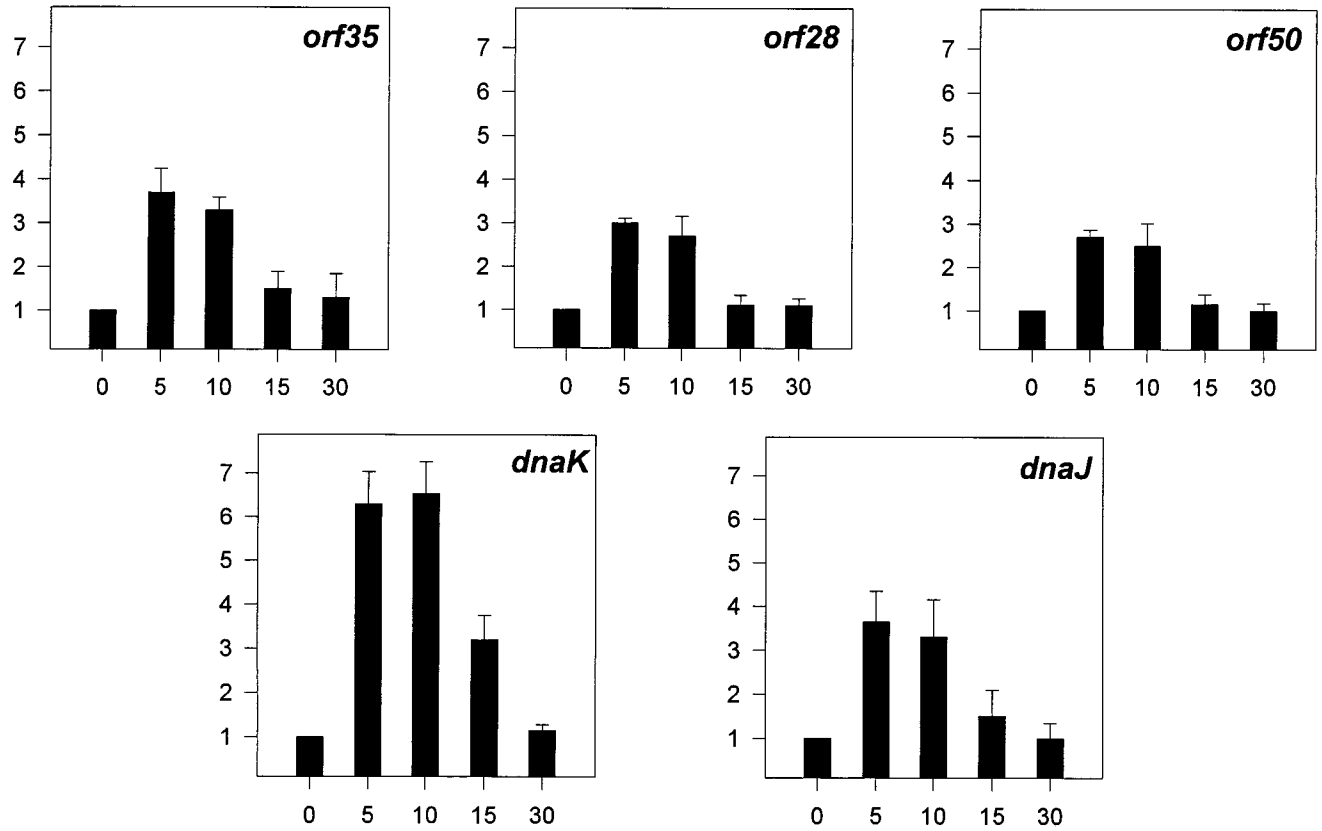

FIG. 4. Schematic representation of the changes in mRNA level of different genes of the *dnaK* operon before (0 min) and at different times after (5, 10, 15, and 30 min) heat shock to 48°C. Luminographs of three different experiments were quantified as described in Materials and Methods. Reproducible induction ratios are shown.

internal promoter downstream of the rho-independent terminator located between *dnaK* and *dnaJ*. Inspection of the DNA sequence within that region identified a sequence with homology to the consensus sequence of vegetative promoters recognized by σ^A (15) (Fig. 1A). In both the −35 and the −10 regions, 5 of 6 nucleotides correspond to the consensus sequence of σ^A-dependent promoters (15), and the spacing between these two regions is 18 bp instead of the optimal 17 bp. To provide additional support for this hypothesis, three different experiments were carried out.

In one experiment, total RNA from the *dnaK*::*cat* insertion mutant BT02 was analyzed in a Northern blot with the *dnaJ*-specific riboprobe. The *cat* gene within *dnaK* is expressed from its own promoter and flanked by two rho-independent terminators (21). In this mutant, the transcription of *dnaJ* from the P_{A1} promoter will be interrupted, due to the polar effect of the *cat* cassette, whereas the postulated vegetative promoter P_{A2} upstream of *dnaJ* should not be affected. This turned out to be the case. While the longer transcripts were absent, the 4.3-kb mRNA was still present and only slightly induced by heat, as described already for the other Northern blots (Fig. 2F). A similar result was found for the mutant strain GH04 (*dnaK-cat-dnaJ*) (data not shown).

In a second experiment, we checked for the presence of GrpE and DnaJ proteins in two mutant *hrcA* strains. In one strain (AS01), *hrcA* had been deleted from the chromosome, whereas in the other strain (BT01), *hrcA* had been inactivated by a *cat* cassette, resulting in a polar effect on the expression of the downstream genes (39). Whole-cell extracts from wild-type strain 1012 and from the two mutant strains were resolved by sodium dodecyl sulfate-polyacrylamide gel electrophoresis, blotted for Western assay, and then treated with antibodies raised against purified GrpE and DnaJ. Whereas the amount of both proteins increased after temperature upshift in the wild-type strain, they were synthesized at a high constitutive rate in the Δ*hrcA* deletion strain, due to the deregulation at P_{A1} (Fig. 5). In the *hrcA*::*cat* mutant, the GrpE protein was not detectable, while DnaJ was present, albeit at reduced amounts compared to the uninduced wild-type strain, and could not be enhanced after thermal shock (Fig. 5). These data confirm those obtained with Northern blot analyses and are in accord with the presence of a second, non-heat-inducible promoter upstream of *dnaJ*.

Further evidence for the existence of this promoter came from experiments in which a 180-bp DNA fragment containing the putative promoter sequence was generated by PCR, fused at the transcriptional level to the reporter gene *bgaB*, and integrated at the *amyE* locus, resulting in strain AM13. This strain exhibited a β-galactosidase activity of about 50 U, whereas the control strain with the promoterless *bgaB* gene produced less than 1 U. These results clearly demonstrate that the 180-bp fragment contains a promoter. Does P_{A2} represent a constitutive promoter, or will it be activated under certain circumstances, e.g., by stress factors? To examine this possibility, strain AM13 was challenged with different stresses. Neither heat, ethanol, oxygen limitation, nor oxidative or osmotic stress led to the induction of the transcriptional fusion (data not shown).

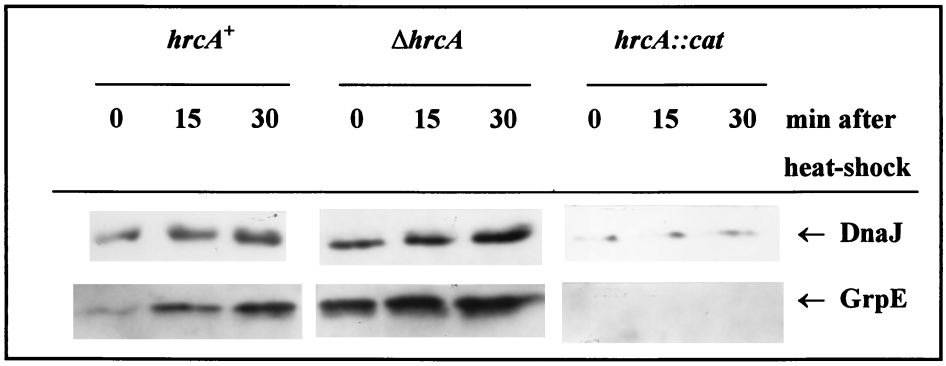

FIG. 5. Immunoblot analysis of proteins from wild-type and two *hrcA* mutant strains. Extracts were prepared before ($t = 0$) and at 15 and 30 min after temperature upshift. Each lane contained 10 μg of protein.

Mutations in *orf35*, *orf28*, or *orf50* do not impair growth of the cells after heat shock. In an attempt to identify the function of the three distal genes, we first determined whether they are essential for survival. The genes were separately generated by PCR, and a *cat* cassette was inserted near their 5′ ends. The resulting constructs were recombined into the bacterial chromosome in place of the corresponding wild-type genes. All three mutants were obtained, demonstrating that none of these genes is essential for survival.

Next, we asked whether one or more of the distal genes are required for growth at high temperatures. We have already shown that insertion of the *cat* cassette within *dnaK* (39) prevents growth of the mutant strains at temperatures above 50°C. The *B. subtilis* strains *orf35*::*cat*, *orf28*::*cat*, and *orf50*::*cat* were analyzed in a similar way, and no difference in their growth profile compared to that of the wild-type strain 1012 was revealed (data not shown). We conclude from these results that the products of the three distal genes are not required for growth at high temperatures.

Heat induction of *dnaJ* is not necessary for growth of *B. subtilis* at high temperatures. Our data revealed that *dnaJ* (and the three distal genes) is transcribed from two different promoters, namely, P_{A1} and P_{A2}. After heat induction, the amount of transcription through the potential rho-independent terminator immediately downstream of *dnaK* increased, and thereby, the total amounts of *dnaJ*-specific transcript and, as shown by immunoblotting (see Fig. 5), of DnaJ protein increased also. This result indicates that enhanced amounts of DnaJ protein might be necessary for the cells to survive high temperatures. To test this hypothesis, we analyzed strain GH04 for growth after thermal shock. In this mutant, a *cat* cassette blocks the readthrough into the four distal genes of transcription initiated at P_{A1}. Consequently, the basal level of DnaJ protein is reduced to about 50% of the wild-type level (i.e., contribution from P_{A1}), and furthermore, increased synthesis after heat shock is prevented (28). The growth curves indicate that there is no significant difference in the growth behavior between the wild-type and the mutant strains under heat shock conditions (data not shown). We therefore tentatively conclude that increased amounts of DnaJ protein are not necessary to cope with a heat shock.

DISCUSSION

A detailed analysis of the transcriptional organization of the *dnaK* operon by Northern analysis and slot blotting and by insertion of a *cat* cassette blocking transcription beyond the cassette led to the following transcriptional organization of the operon. The whole operon is preceded by the vegetative promoter P_{A1} recognized by σ^A (4, 48). Transcription starting at P_{A1} results in two primary transcripts of different lengths and amounts, an 8-kb transcript encompassing the whole operon and synthesized in very low amounts and large amounts of a 3.6-kb tricistronic transcript terminated at a rho-independent terminator immediately downstream of *dnaK* (Fig. 3). Both primary transcripts are unstable and are cleaved at one or more internal sites. One cleavage site has already been identified as being located between *hrcA* and *grpE* (39), as being responsible for extensive cleavage between both genes, and as resulting in processed products of 7 and 2.6 kb. In addition, there seem to be cleavage sites within the distal half of the 8-kb transcript, giving rise to several potential cleaved products. While the stem-loop structure between *hrcA* and *grpE* seems to be a processing site, those downstream of *dnaJ* are recognized as degradation sites, since they are located within the coding sequence of *orf*35, *orf28*, and *orf50*. If this interpretation is correct, what could be the biological function of cleavage between *hrcA* and *grpE*? Since only small amounts of the HrcA repressor are needed, in contrast to large amounts of GrpE and DnaK proteins, this requirement could be regulated by cleavage of the primary transcript and different half-lives of the processed products. This is supported by the observation that the 2.6-kb product can easily be demonstrated in Northern blots in contrast to the 1.0-kb product representing *hrcA* (19).

In general, cleavage at processing sites leads to a product with an enhanced half-life, while cleavage at a degradation site expedites breakdown of the RNA molecules by introduction of entry sites for exoribonucleases. In both cases, cleavage occurs within or in the vicinity of published stem-loop structures. RNase III seems to specifically recognize stem-loop structures and cleaves either on one or on both sides of the stem, usually within an internal unpaired region, as has been reported for cleavage of the 30S ribosomal precursor RNA of *E. coli* (2). RNase E and RNase K cleave close to published secondary structures in AU-rich sequences of, e.g., T4 gene 32 RNA (9) and the *ompA* 5′ untranslated region (5), respectively.

Most of the Northern blots shown here exhibit hybridization signals at the two positions where the 16S and 23S RNAs migrate. These bands occur due to the trapping of degradation products by the rRNAs rather than by unspecific hybridization. This interpretation is strengthened by the following observations. First, DIG-labelled oligonucleotides mostly failed to give hybridization signals at the positions of the rRNAs, and second, if the transcript to be detected migrates below the rRNA(s), the riboprobe fails to produce a signal (19).

In addition to P_{A1}, there is a second internal promoter, P_{A2},

located between *dnaK* and *dnaJ* (Fig. 3) and most probably also recognized by σ^A. Transcription initiated at this promoter seems to be constitutive and not to be influenced by treating cells with different stress factors. Although it is unlikely, however, we cannot exclude the possibility that regulatory sequences located outside the 180-bp fragment fused to the reporter gene are necessary for activation of P_{A2}.

As already mentioned, the CIRCE regulon is under negative control of *hrcA*. Since the synthesis of *dnaK* operon mRNA is transiently induced upon temperature upshift, the activity of the repressor has to be regulated in such a way to allow increased expression from P_{A1}. The protein(s) involved in this regulation have not yet been identified. Therefore, we asked whether any of the three distal genes might be involved in the regulation of HrcA activity. To examine this possibility, we made use of strain AM03 carrying a transcriptional fusion between the promoter P_{A1} and the reporter gene *bgaB* integrated at the *amyE* locus. The three mutations *orf35*::*cat*, *orf28*::*cat*, and *orf50*::*cat* were separately crossed into AM03, and the β-galactosidase activities of the resulting strains were measured before and after thermal upshock. Since neither the expression of the fusion at 37°C nor the induction behavior was altered in the mutants (data not shown), we conclude that none of the three distal genes are involved in the regulation of P_{A1}. Similar results have been obtained by analyzing class II and class III heat shock genes (data not shown). We also determined the sporulation frequencies with the three null mutants, and they turned out to be completely sporulation proficient (data not shown).

Only P_{A1} becomes active after temperature upshift, probably as a result of inactivation of the HrcA repressor. This results in a transient increase of the two primary transcripts and their processing products. Interestingly, elevated amounts of DnaJ and of the three distal gene products do not appear to be required for the cells to cope with the heat stress, since blocking readthrough transcription from P_{A1} into the four distal genes did not produce a phenotype. This raises the question about the biological consequences of transcriptionally fusing the four downstream genes to the 3′ end of the *dnaK* operon. If increased synthesis of the four distal genes is not required for the cells under heat stress conditions, why are they part of a heat shock operon? Are there stressful conditions other than that of heat which require their increased synthesis? It should be mentioned in this context that class I heat shock genes are characterized by their failure to be induced by other stress factors such as osmotic or oxidative stress or O_2 limitation (44). The biological significance of the heat inducibility of this operon remains an enigma at the moment.

The genomic organization of the *dnaK* operons of *B. stearothermophilus*, *C. acetobutylicum*, and *S. aureus* is comparable to that of *B. subtilis* (18, 30, 32). However, in the latter two species, there is no vegetative promoter located between *dnaK* and *dnaJ* (30, 32). Though Ohta and coworkers (32) deduced the presence of a promoter in *S. aureus* upstream of *dnaJ* on the basis of primer extension analysis, the detected signal is located in the terminator structure located between *dnaK* and *dnaJ* and could be explained by the interference of this secondary structure on the activity of the reverse transcriptase (32). As already mentioned, *dnaJ* is followed by *orf35* in both *C. acetobutylicum* and *S. aureus* (1, 32), and we suggest that in both cases it will also be followed by *orf28* and *orf50.*

What is known about the possible functions of the three distal ORFs? A search in the protein databases revealed significant homologies to proteins encoded by other bacterial species, but only in the case of ORF35 was homology to a protein with a known function found, namely, to the PrmA protein of *E. coli*. This protein was shown to be a methyltransferase responsible for the methylation of ribosomal protein L11 at different positions (42). In this context, it is interesting that the *prmA* gene of *E. coli* is dispensable, since its inactivation did not result in any phenotype (43). An alignment of the five PrmA proteins deduced from their ORFs shows a low overall homology and four regions of good homology, where three regions constitute a binding site for the methyl group donor *S*-adenosylmethionine, and a fourth region, designated the PrmA motif, is strongly conserved only in the five potential PrmA enzymes (data not shown). As all five proteins share this motif, this box must be necessary for their specific function, e.g., for the recognition of their methylation substrate.

ACKNOWLEDGMENTS

We are grateful to Colin Harwood for critical reading of the manuscript.

This work was supported by a Grant-in-Aid for Creative Basic Research "Human Genome Analysis" from the Ministry of Education, Science, and Culture, Japan, to Y.K. and by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie (W.S.).

REFERENCES

1. **Behrens, S., F. Narberhaus, and H. Bahl.** 1993. Cloning, nucleotide sequence and structural analysis of the *Clostridium acetobutylicum dnaJ* gene. FEMS Microbiol. Lett. **114:**53–60.
2. **Bram, R. J., R. A. Young, and J. A. Steitz.** 1980. The ribonuclease III site flanking 23S sequences in the 30S ribosomal precursor RNA of *Escherichia coli*. Cell **19:**393–401.
3. **Bukau, B.** 1993. Regulation of the heat-shock response. Mol. Microbiol. **9:**671–680.
4. **Chang, B.-Y., K.-Y. Chen, Y.-D. Wen, and C.-T. Liao.** 1994. The response of a *Bacillus subtilis* temperature-sensitive *sigA* mutant to heat stress. J. Bacteriol. **176:**3102–3110.
5. **Chen, L.-H., S. A. Emory, A. L. Bricker, P. Bouvet, and J. G. Belasco.** 1991. Structure and function of a bacterial mRNA stabilizer: analysis of the 5′ untranslated region of *ompA* mRNA. J. Bacteriol. **173:**4578–4586.
6. **Dean, D. R., J. A. Hoch, and A. I. Aronson.** 1977. Alteration of the *Bacillus subtilis* glutamine synthetase results in overproduction of the enzyme. J. Bacteriol. **131:**981–987.
7. **Deuerling, E., A. Mogk, C. Richter, M. Purucker, and W. Schumann.** The *ftsH* gene of *Bacillus subtilis* is involved in major cellular processes such as sporulation, stress adaptation and secretion. Mol. Microbiol., in press.
8. **Deuerling, E., B. Paeslack, and W. Schumann.** 1995. The *ftsH* gene of *Bacillus subtilis* is transiently induced after osmotic and temperature upshock. J. Bacteriol. **177:**4105–4112.
9. **Ehretsmann, C. P., A. J. Carpousis, and H. M. Krisch.** 1992. *Escherichia coli* RNase E has a role in the decay of bacteriophage T4 mRNA. Genes Dev. **6:**149–159.
10. **Ellis, J.** 1987. Proteins as molecular chaperones. Nature (London) **328:**378–379.
11. **Ellis, R. J., and S. M. Van der Vies.** 1991. Molecular chaperones. Annu. Rev. Biochem. **60:**321–347.
12. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, and E. Hedblom.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
13. **Georgopoulos, C.** 1992. The emergence of the chaperone machines. Trends Biochem. Sci. **17:**295–299.
14. **Gushima, H., E. Yasuda, E. Soeda, M. Yokota, M. Kondo, and A. Kimura.** 1984. Complete nucleotide sequence of the *E. coli* glutathione synthetase gsh-II. Nucleic Acids Res. **12:**9299–9307.
15. **Haldenwang, W. G.** 1995. The sigma factors of *Bacillus subtilis*. Microbiol. Rev. **59:**1–30.
16. **Hecker, M., W. Schumann, and U. Völker.** 1996. Heat-shock and general stress response in *Bacillus subtilis*. Mol. Microbiol. **19:**417–428.
17. **Hirata, H., T. Fukazawa, S. Negoro, and H. Okada.** 1986. Structure of a β-galactosidase gene of *Bacillus stearothermophilus*. J. Bacteriol. **166:**722–727.
18. **Hirata, H., S. Negoro, and H. Okada.** 1985. High production of thermostable β-galactosidase of *Bacillus stearothermophilus* in *Bacillus subtilis*. Appl. Environ. Microbiol. **49:**1547–1549.
19. **Homuth, G.** 1996. Unpublished data.
20. **Homuth, G., M. Heinemann, U. Zuber, and W. Schumann.** 1996. The genes *lepA* and *hemN* form a bicistronic operon in *Bacillus subtilis*. Microbiology **142:**1641–1649.

21. **Horinouchi, S., and B. Weisblum.** 1982. Nucleotide sequence and functional map of pC194, a plasmid that specifies inducible chloramphenicol resistance. J. Bacteriol. **150:**815–825.
22. **Hungerer, C., B. Troup, U. Römling, and D. Jahn.** 1995. Cloning, mapping and characterization of the *Pseudomonas aeruginosa hemL* gene. Mol. Gen. Genet. **248:**375–380.
23. **Jekel, M., and W. Wackernagel.** 1994. Location of the *endA* gene coding for endonuclease I on the physical map of the *Escherichia coli* K-12 chromosome. J. Bacteriol. **176:**1550–1551.
24. **Krüger, E., U. Völker, and M. Hecker.** 1994. Stress induction of *clpC* in *Bacillus subtilis* and its involvement in stress tolerance. J. Bacteriol. **176:** 3360–3367.
25. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227:**680–685.
26. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227:**1435–1441.
27. **Missiakas, D., S. Raina, and C. Georgopoulos.** 1996. Heat shock regulation, p. 481–502. *In* E. C. C. Lin and A. S. Lynch (ed.), Regulation of gene expression in *Escherichia coli*. Chapman and Hall, London, England.
28. **Mogk, A., R. Hayward, and W. Schumann.** 1996. Integrative vectors for constructing single-copy transcriptional fusions between *Bacillus subtilis* promoters and various reporter genes encoding heat-stable enzymes. Gene **182:**33–36.
29. **Moran, C. P., Jr., N. Lang, S. F. J. Le Grice, G. Lee, M. Stephens, A. L. Sonenshein, J. Pero, and R. Losick.** 1982. Nucleotide sequences that signal initiation of transcription and translation in *Bacillus subtilis*. Mol. Gen. Genet. **186:**339–346.
30. **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.
31. **Ogasawara, N., S. Nakai, and H. Yoshikawa.** 1994. Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. DNA Res. **1:**1–14.
32. **Ohta, T., K. Saito, M. Kuroda, K. Honda, H. Hirata, and H. Hayashi.** 1994. Molecular cloning of two new heat shock genes related to the *hsp70* genes in *Staphylococcus aureus*. J. Bacteriol. **176:**4779–4783.
33. **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.
34. **Saito, H., T. Shibata, and T. Ando.** 1979. Mapping of genes determining nonpermissiveness and host-specific restriction to bacteriophages in *Bacillus subtilis* Marburg. Mol. Gen. Genet. **170:**117–122.
35. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
36. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
37. **Schulz, A.** 1996. Unpublished data.
38. **Schulz, A., and W. Schumann.** 1996. *hrcA*, the first gene of the *Bacillus subtilis dnaK* operon, encodes a negative regulator of class I heat shock genes. J. Bacteriol. **178:**1088–1093.
39. **Schulz, A., B. Tzschaschel, and W. Schumann.** 1995. Isolation and analysis of mutants of the *dnaK* operon of *Bacillus subtilis*. Mol. Microbiol. **15:**421–429.
40. **Schumann, W.** 1996. The heat shock stimulon of *Bacillus subtilis*. Braz. J. Genet. **19:**387–398.
41. **Takemaru, K., M. Mizuno, T. Sato, M. Takeuchi, and Y. Kobayashi.** 1995. Complete nucleotide sequence of a *skin* element excised by DNA rearrangement during sporulation in *Bacillus subtilis*. Microbiology **141:**323–327.
42. **Vanet, A., J. A. Plumbridge, and J. Alix.** 1993. Cotranscription of two genes necessary for ribosomal protein L11 methylation (*prmA*) and pantothenate transport (*panF*) in *Escherichia coli* K-12. J. Bacteriol. **175:**7178–7188.
43. **Vanet, A., J. A. Plumbridge, M. Guérin, and J. Alix.** 1994. Ribosomal protein methylation in *Escherichia coli*: the gene *prmA*, encoding the ribosomal protein L11 methyltransferase, is dispensable. Mol. Microbiol. **14:**947–958.
44. **Völker, U., S. Engelmann, B. Maul, S. Riethdorf, A. Völker, R. Schmid, H. Mach, and M. Hecker.** 1994. Analysis of the induction of general stress proteins of *Bacillus subtilis*. Microbiology **140:**741–752.
45. **Wetzstein, M.** Unpublished data.
46. **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.
47. **Yuan, G., and S. Wong.** 1995. Isolation and characterization of *Bacillus subtilis* regulatory mutants: evidence for *orf39* in the *dnaK* operon as a repressor gene in regulating the expression of both *groE* and *dnaK*. J. Bacteriol. **177:**6462–6468.
48. **Yuan, G., S.-L. Wong, and S. L. Wong.** 1995. Regulation of *groE* expression in *Bacillus subtilis*: the involvement of the σ^A-like promoter and the roles of the inverted repeat sequence (CIRCE). J. Bacteriol. **177:**5427–5433.
49. **Yura, T., H. Nagai, and H. Mori.** 1993. Regulation of the heat-shock response in bacteria. Annu. Rev. Microbiol. **47:**321–350.