JOURNAL OF BACTERIOLOGY, May 1992, p. 3300–3310
0021-9193/92/103300-11$02.00/0

Vol. 174, No. 10

Cloning, Sequencing, and Molecular Analysis of the *dnaK* Locus from *Bacillus subtilis*

MICHAEL WETZSTEIN,[1] UWE VÖLKER,[2] JÜRGEN DEDIO,[1] SABINE LÖBAU,[1] ULRICH ZUBER,[1] MICHAEL SCHIESSWOHL,[1] CHRISTIAN HERGET,[1] MICHAEL HECKER,[2] AND WOLFGANG SCHUMANN[1]*

Lehrstuhl für Genetik, Universität Bayreuth, W-8580 Bayreuth,[1] and Institut für Mikrobiologie, Universität Greifswald, 0-2200 Greifswald,[2] Germany

Received 10 December 1991/Accepted 4 March 1992

By using an internal part of the *dnaK* gene from *Bacillus megaterium* as a probe, a 5.2-kb *Hin*dIII fragment of chromosomal DNA of *Bacillus subtilis* was cloned. Downstream sequences were isolated by in vivo chromosome walking. Sequencing of 5,085 bp revealed four open reading frames in the order *orf39-grpE-dnaK-dnaJ*. *orf39* encodes a 39-kDa polypeptide of unknown biological function with no noticeable homology to any other protein within the data bases. Alignment of the GrpE protein with those of three other bacterial species revealed a low overall homology, but a higher homology restricted to two regions which might be involved in interactions with other proteins. Alignment of the DnaK protein with six bacterial DnaK polypeptides revealed that a contiguous region of 24 amino acids is absent from the DnaK proteins of all known gram-positive species. Primer extension studies revealed three potential transcription start sites, two preceding *orf39* (S1 and S2) and a third one in front of *grpE* (S3). S2 and S3 were activated at a high temperature. Northern (RNA) analysis led to the detection of three mRNA species of 4.9, 2.6, and 1.5 kb. RNA dot blot experiments revealed an at-least-fivefold increase in the amount of specific mRNA from 0 to 5 min postinduction and then a rapid decrease. A transcriptional fusion between *dnaK* and the *amyL* reporter gene exhibited a slight increase in α-amylase activity after heat induction. A 9-bp inverted repeat was detected in front of the coding region of *orf39*. This inverted repeat is present in a number of other heat shock operons in other microorganisms ranging from cyanobacteria to mycobacteria. The biological property of this inverted repeat as a putative key element in the induction of heat shock genes is discussed. The *dnaK* locus was mapped at about 223° on the *B. subtilis* genetic map.

The heat shock response is a homeostatic mechanism that enables cells to survive a variety of environmental stresses. It is characterized by the increased synthesis of a group of evolutionarily conserved proteins, heat shock proteins (HSPs), and is a universal feature of both prokaryotic and eukaryotic cells (35). When *Escherichia coli* cells are shifted to a high temperature, the synthesis of a set of about 20 HSPs transiently increases and then declines rapidly to steady-state levels characteristic of the new ambient temperature (44).

The highly conserved HSPs perform similar functions in all organisms. One of these functions is the well-established regulation of protein-protein interactions by the chaperonins (69). One of the most abundant HSPs, HSP70, is highly conserved in evolution. It is found in such diverse organisms as *E. coli*, *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and *Homo sapiens* (14, 35).

The *dnaK* gene of *E. coli* was originally discovered because mutations in it blocked bacteriophage lambda DNA replication at all temperatures (22, 23). Subsequently, the *dnaK* gene product was shown to be essential for *E. coli* viability at high and low temperatures (22, 30, 46, 51, 52), and genetic evidence indicates functions for DnaK after heat shock in the synthesis of RNA and DNA and in cell division (30, 46, 53).

Interestingly, very little work has been done on the heat shock response in *Bacillus subtilis*. Some descriptive studies indicate that about 20 proteins were synthesized with an increased rate after a temperature upshift (49, 59). A recent publication described the induction of as many as 60 proteins after a temperature upshift from 37 to 48°C (40). Arnosti et al. (4) identified 26 HSPs and showed that antibodies raised against the *E. coli* HSPs DnaK, GroEL, and Lon cross-react with HSPs from *B. subtilis*. However, there is no information on the regulation of the heat shock response or on the function of the HSPs in *B. subtilis*. Clearly, such studies are of great importance in assessing the role of this response in cellular physiology.

Our approach to this problem has been to clone and to sequence the most prominent heat shock genes and to use these genes as probes to study both the regulation and function of HSPs. Recently, part of the *dnaK* sequence was reported (26). We published its complete sequence (62) and that of the *grpE* gene (63). Here, we describe the isolation and the analysis of transcription of four heat shock genes which constitute the *dnaK* locus.

MATERIALS AND METHODS

Bacterial strains, plasmids, and bacteriophage. The bacterial strains and plasmids used in this study are listed in Table 1. Transducing phage PBS1 has been described previously (32). *B. subtilis* and *E. coli* strains were routinely grown aerobically at 37°C in Luria-Bertani (LB) medium. For the induction of the heat shock regulon, bacteria were shifted from 37 to 48°C. When necessary, media were supplemented with ampicillin, chloramphenicol, and kanamycin at final concentrations of 100, 5, and 25 μg/ml, respectively. *E. coli*

* Corresponding author.

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Characteristics	Source or reference
E. coli RR1	*ara-14 galK2 lacY1 mlt-1 proA2 rpsL20 supE44 xyl-5* $r_B^- m_B^-$	54
B. subtilis		
MB11	*hisH2 metB10 lys-3 trpC2*	48
YB886	*amyE metB5 trpC2 xin-1*	40
CHA5	YB886 carrying chromosomal insertion of pCHA5	This work
JDA1	YB886 carrying chromosomal insertion of pJDA1	This work
JDA2	YB886 carrying chromosomal insertion of pJDA2	This work
JDB1	YB886 carrying chromosomal insertion of pJDB1	This work
JDB2	YB886 carrying chromosomal insertion of pJDB2	This work
Plasmid		
pUC18	Ap^r	66
pJH101	Ap^r Tc^r Cm^r	19
pUC8+K	pUC8 containing *B. megaterium dnaK*	60
pWD3	Ap^r Cm^r promoterless *amyL*	64
pACYC177	Ap^r Km^r	10
pDG271	Ap^r Cm^r	3
pDG271-B	*Eco*RI site of pDG271 replaced by *Bgl*II site	This work
pCHA5	pACYC177 containing 0.4-kb *Hin*dIII (*dnaK-dnaJ*) and 2.7-kb *Bgl*II (*cat*)	This work
pCHA7	pACYC177 derivatives carrying *dnaJ*	This work
pJDA1	pHJ101 containing 0.8-kb *Bgl*II of *dnaK*	This work
pJDA2	pJH101 containing 0.8-kb *Bgl*II of *dnaK*	This work
pJDB1	pWD3 carrying 0.8-kb *Bgl*II of *dnaK*	This work
pJDB2	pWD3 carrying 0.8-kb *Bgl*II of *dnaK*	This work
pMWD25, pMWB5, pMWD60	pUC18 containing a 5.2-kb *Hin*dIII fragment of *dnaK*	This work
pMWC862	pJH101 containing 3′ end of *dnaK* and most of *dnaJ*	This work

and *B. subtilis* were transformed as described previously (2, 54).

Enzymes and chemicals. Restriction enzymes, T4 DNA ligase, and the Klenow fragment were obtained from Bethesda Research Laboratories (BRL); calf intestinal alkaline phosphatase and T4 polynucleotide kinase was obtained from Boehringer GmbH, Mannheim, Germany. RNasin ribonuclease inhibitor was purchased from Promega. DNA probes were labeled with [^{32}P]dATP (specific activity, 3,000 Ci/mmol; Amersham Corp.) by the random priming technique of Feinberg and Vogelstein (18).

DNA isolation and cloning of the *dnaK* locus. *B. subtilis* DNA was isolated as described previously (16). Total DNA from *B. subtilis* MB11 was digested with *Hin*dIII; 5- to 7-kb fragments were eluted from a 0.8% agarose gel, ligated to *Hin*dIII-digested pUC18, and transformed into *E. coli* RR1. Plasmid screening indicated that 80% of the colonies contained plasmids with 5- to 6-kb inserts.

Plasmids containing the desired fragments were identified by probing Southern blots of *Hin*dIII-digested recombinant plasmids with the ^{32}P-labeled 0.8-kb *Dra*I fragment of *Bacillus megaterium dnaK* (pMWB5, pMWD25, and pMWD60). For Southern blot hybridizations (54), DNA was bound to nylon membrane (Biodyne A; Pall Corp.). Hybridizations and washes were done at 60 and 65°C, respectively. Sequencing revealed that the 5.2-kb *Hin*dIII fragment did not contain the complete *dnaK* gene. The remaining part was isolated by in vivo chromosome walking as follows. A 0.8-kb internal *Bgl*II fragment of the *dnaK* gene was ligated into the unique *Bam*HI site of the integrating vector pJH101, resulting in pJDA1 and pJDA2 containing the *Bgl*II fragment in both possible orientations. Both recombinant plasmids were separately transformed into *B. subtilis* YB886, and Cm^r transformants were selected. Southern hybridizations were used to confirm that the Cm^r transformants were the result of a crossover between the plasmids and the chromosome at the *dnaK* locus. This procedure yielded strains JDA1 (Ap^r marker of the vector proximal to the 5′ end of *dnaK*) and JDA2 (Ap^r marker of the vector distal to the 5′ end of *dnaK*). Chromosomal DNA of JDA1 was cleaved with *Eco*RI, ligated, and transformed into *E. coli* RR1. One of the Ap^r-conferring plasmids (pMWC862) was used to sequence the 3′ end of *dnaK* and most of *dnaJ*.

To sequence the complete *dnaJ* gene, another plasmid was recovered from chromosomal DNA of *B. subtilis* by in vivo chromosome walking. Plasmid pACYC177 received a 0.4-kb *Hin*dIII fragment, which contains the 3′ end of *dnaK* and the 5′ end of *dnaJ* (see Fig. 2), and a 2.7-kb *Bgl*II fragment isolated from pDG271-B, which contains the *cat* gene. This recombinant integration vector, pCHA5, was transformed into YB886, and Cm^r transformants were selected. Southern hybridizations confirmed the integration at the correct site. Chromosomal DNA of one of these transformants (CHA5) was cut with *Nar*I, ligated, and transformed into *E. coli* RR1. One of the Ap^r-conferring plasmids (pCHA7) was selected for further studies.

Determination of DNA sequences. Most restriction fragments to be sequenced were ligated into the appropriate sites of M13 vectors. DNA sequencing was performed by the dideoxy-chain termination method of Sanger et al. (55), by using [α-^{35}S]dATP and the M13 universal primer (39). Double-stranded DNA sequencing was carried out by the method of Heinrich (27) with plasmid DNA as the template. The sequencing reactions were performed with the Sequenase kit (version 2.0) of U.S. Biochemical Corp.

Isolation of total RNA. Total RNA was isolated by a modification of the procedure of Ambulos et al. (1). Briefly, cell lysis, phenol, and ether extractions were carried out as described previously, and after the ethanol precipitation the contaminating DNA was removed with the aid of Quiagen-tip 100 cartridges (Quiagen, Düsseldorf, Germany).

Primer extension analysis. Synthetic oligonucleotides complementary to the region immediately downstream of the initiation codons of *orf39* (see Fig. 3, nucleotide positions 300 to 316), *grpE* (positions 1377 to 1393) (63), and *dnaJ* (positions 3988 to 4004) were 5′ end labeled with ^{32}P and used as primers. Labeled primer (1 pmol) was annealed to 5 μg of cellular RNA in 10 μl of hybridization buffer (10 mM Tris-HCl [pH 7.9], 1 mM EDTA, 250 mM KCl) at 55°C for 60 min. Subsequently, 25 μl of reaction buffer (10 mM Tris-HCl [pH 8.7], 10 mM $MgCl_2$, 5 μM dithiothreitol, 0.4 mM dNTPs, 1 U of RNase inhibitor per μl) was added. The oligonucleotide primers were extended with 20 U of avian myeloblastosis virus reverse transcriptase (Stratagene) at 37°C for 30

min. The samples were precipitated with ethanol, vacuum dried, and resuspended in 10 μl of stop mixture (Pharmacia). After heat denaturation, 1-μl samples were immediately loaded onto 6% polyacrylamide–urea sequencing gels for electrophoresis. Dideoxynucleotide sequencing reactions (27) using the same primer and an appropriate plasmid DNA template were run in parallel to allow determination of the end points of the extension products.

Analysis of mRNA synthesis by dot blotting. RNA was diluted in serial dilution steps in 10× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), denatured at 90°C for 10 min, and spotted onto a nitrocellulose filter soaked with 10× SSC. After baking at 80°C for 2 h, the filters were prehybridized for 2 h and hybridized overnight in hybridization buffer (50% formamide, 5× SSC, 50 mM phosphate buffer [pH 6.5], 5× Denhardt solution, 200 μg of salmon sperm DNA per ml) at 42°C. The 0.8-kb *Bgl*II fragment containing an internal part of the *dnaK* gene was used as a probe for hybridization. The nitrocellulose filters were washed twice at room temperature for 5 min and twice at 68°C for 15 min in 0.1× SSC–0.1% sodium dodecyl sulfate and exposed to X-ray films.

Northern analysis. Equal amounts of total RNA (15 μg each) were separated under denaturing conditions in a 2% agarose–2.1 M formaldehyde–MOPS [morpholinepropanesulfonic acid] gel, stained with ethidium bromide, and blotted onto nitrocellulose (Schleicher & Schuell) with 20× SSC by using a vacuum blotter. After baking at 80°C for 2 h, filters were prehybridized, hybridized, and washed as described for dot blotting. The same radioactively labeled DNA fragment was used as a probe for the hybridization. RNA molecular weight markers (Boehringer Mannheim) of 7.4, 5.3, 2.8, 1.9, and 1.6 kb were used.

Construction of a chromosomal *dnaK-amyL* transcriptional fusion. A transcriptional fusion between *dnaK* and the *amyL* reporter gene was constructed by first inserting the internal 0.8-kb *Bgl*II fragment of *dnaK* (see Fig. 2) from pMWD25 into the unique *Bam*HI site of pWD3. Recombinant plasmids containing this fragment in both possible orientations were isolated. In pJDB1, the *dnaK* reading frame is fused to that of *amyL*, whereas pJDB2 received this fragment in the opposite orientation. Next, each plasmid was transformed into YB886, and Cm^r transformants were isolated. Insertion of the plasmids within the *dnaK* gene was verified by Southern blotting. α-Amylase was assayed as described previously (45).

Mapping of the *dnaK* locus. Mapping of the *dnaK* locus was performed by PBS1 transduction (29) with a lysate from *B. subtilis* JDA1 and JDA2. Cm^r transductants were scored for linkage to several loci from the set of reference strains of Dedonder et al. (15). This arrangement has been confirmed by Southern blotting.

Both strains had a chloramphenicol acetyltransferase (*cat*) gene integrated at the site of *dnaK* in the chromosome. The strains were constructed by integration of plasmids pJDA1 and pJDA2 at the *dnaK* locus (see above).

DNA and protein sequence analysis. Sequence analysis was performed by using the PC/GENE package (IntelliGenetics) for the IBM personal computer and HUSAR (Heidelberg Unix Sequence Analysis Resources).

Nucleotide sequence accession number. Sequence data of the complete *dnaK* locus have been assigned GenBank data base accession no. M84964.

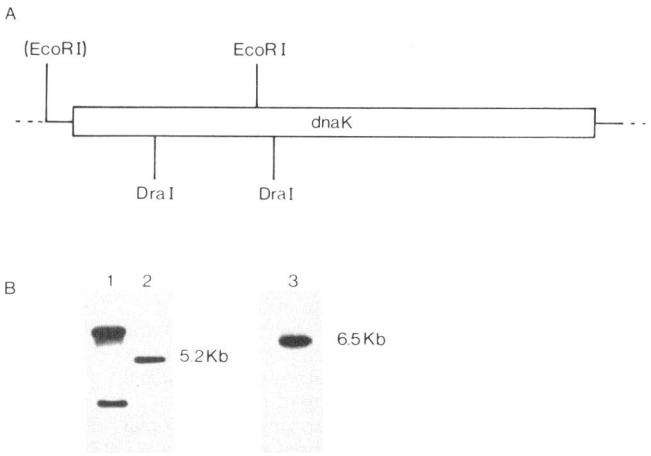

FIG. 1. Identification of the *dnaK* genes of *B. subtilis* and of *B. licheniformis*. (A) Schematic representation of the *dnaK* gene of *B. megaterium*; (B) Southern blot analysis, revealing two bands after digestion with *Eco*RI (lane 1, *B. megaterium*) or one band after digestion with *Hin*dIII (lane 2, *B. subtilis*; lane 3, *B. licheniformis*).

RESULTS

Hybridization of the *dnaK* homolog of *B. megaterium* to chromosomal DNAs of *B. subtilis* and *Bacillus licheniformis*. Sussmann and Setlow (60) published the sequence of the *dnaK* gene of *B. megaterium*. To verify whether this gene will allow isolation of its *B. subtilis* homolog, the internal 0.4-kb *Dra*I fragment (Fig. 1A) was isolated, radioactively labeled, hybridized to a chromosomal DNA blot, and washed under stringent conditions. The Southern blot of chromosomal DNA of three different *Bacillus* species revealed the presence of one or two bands (Fig. 1B).

DNA of *B. megaterium* cleaved with *Eco*RI exhibited two hybridizing bands (Fig. 1B, lane 1) because of the *Eco*RI site within the *Dra*I probe (Fig. 1A). Furthermore, the probe recognized one band in the *Hin*dIII digests of the chromosomal DNAs of *B. subtilis* (5.2 kb) and *B. licheniformis* (6.5 kb) (Fig. 1B, lanes 2 and 3, respectively). This result indicates the presence of at least one *dnaK* gene within these fragments.

Molecular cloning of the *dnaK* locus. Next, we cloned the *dnaK* homolog from the chromosomal DNA of *B. subtilis* as described in Materials and Methods. Screening of individual clones revealed three clones that gave clear hybridization signals (pMWD25, pMWB5, and pMWD60).

Sequencing of part of the insert of pMWD25 revealed the presence of the *dnaK* homolog near one end, but the gene was incomplete, missing its 3′ end. Cloning of this part was accomplished by in vivo chromosome walking as described in Materials and Methods. One of these recombinant plasmids (pMWC862), carrying about 2.2 kb of chromosomal DNA, was isolated, and part of the insert was sequenced. It revealed the 3′ end of the *dnaK* gene and most of another open reading frame (ORF) to be homologous with the *dnaJ* gene of *E. coli* (see below).

Since the insert did not carry the complete *dnaJ* gene, the low-copy-number vector pACYC177 was used as an integration vector. Chromosomal DNA of one integrant was isolated, cleaved with *Nar*I, ligated, and transformed into *E. coli* RR1. All of the transformants contained plasmids of the same size. One of these plasmids (pCHA7) was chosen for further analyses.

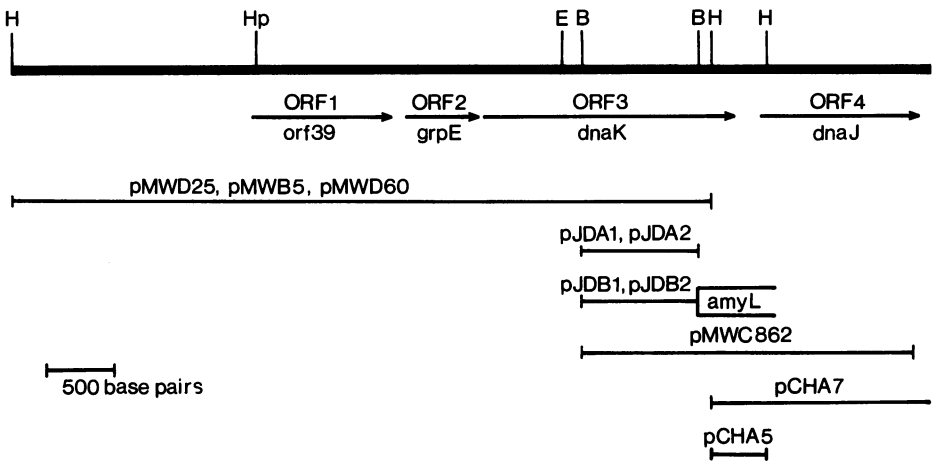

FIG. 2. Restriction map and plasmids of the *dnaK* locus. Inserts of the various plasmids described in the text are shown. Locations of the identified genes are indicated with arrows. B, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; Hp, *Hpa*I.

Sequencing and features of the *dnaK* locus. By using M13 subclones and synthetic oligonucleotides, the sequence of the *dnaK* locus was determined. A total of 5,085 bp was sequenced, revealing the presence of four contiguous ORFs (Fig. 2 and 3).

ORF1 is 1,029 nucleotides long, beginning with an AUG start codon at position 242 and ending with a UAA stop codon at position 1271. It encodes a peptide of 343 amino acids with a predicted molecular mass of 39,027 Da. By using standard computer homology searches, we were unable to identify other significant protein sequences in the GenBank, EMBL, and NBRF data bases. We tentatively designated this gene *orf39*.

ORF2 is 561 nucleotides long, beginning with an AUG start codon at position 1345 and ending with a UAA stop codon at position 1906. The start codon of ORF2 is 75 nucleotides from the last stop codon of ORF1. ORF2 encodes a peptide of 187 amino acids with a predicted molecular mass of 21,683 Da. A homology search revealed similarity with *grpE* of *E. coli* (see below). Therefore, we designated ORF2 *grpE*.

ORF3 is 1,833 nucleotides long, beginning with a GUG start codon at position 1932 and ending with a UAA stop codon at position 3765. Alternative start codons GUG, UUG, and AUU are frequently used in prokaryote genes; all are single-base changes from AUG (47). The start codon of ORF3 is 23 nucleotides from the stop codon of ORF2. ORF3 codes for a protein of 611 amino acids with a predicted molecular mass of 65,841 Da. As already mentioned, this protein is highly homologous to *dnaK* of *E. coli* and other bacterial species (see below), and therefore, ORF3 is called *dnaK*. To check whether the N terminus of the DnaK protein as deduced from its ORF agrees with that of the purified protein, DnaK was isolated from a two-dimensional gel and its N-terminal amino acid sequence was determined on an Applied Biosystems 373 protein sequencing machine. The first seven amino acid residues, SKVIGID, are in full agreement with those deduced from the nucleotide sequence (62).

ORF4 is 1,116 bp long, starts with an AUG codon at position 3967, and ends with a UAA stop codon at position 5085. It encodes a polypeptide of 372 amino acids with a predicted molecular mass of 40,847 Da. The deduced amino acid sequence is homologous to DnaJ of *E. coli*. Therefore, ORF4 was designated *dnaJ*. The start codon of *dnaJ* is 199 nucleotides downstream from the stop codon of *dnaK*.

Sequence similarities between GrpE, DnaK, and those of other bacterial species. The alignment of the *B. subtilis* GrpE amino acid sequence with that of three other bacterial species is shown in Fig. 4A. The overall identity among these proteins is low; the identity between the *E. coli* GrpE protein and the others ranges from 9.3% (*Mycobacterium tuberculosis*) to 30.5% (*B. subtilis*). A relaxation to include conservative amino acid changes increased similarity from 23.4% (*M. tuberculosis*) to 46.5% (*B. subtilis*).

Closer inspection of the sequences revealed two regions with higher homology. Region I (*E. coli* amino acid residues 61 to 106) exhibits from 34.8% (*M. tuberculosis*) to 54.3% (*Chlamydia trachomatis*) to 65.2% (*B. subtilis*) homology. Region II (amino acid residues 151 to 192) shows homology from 26.2% (*M. tuberculosis*) to 52.4% (*C. trachomatis* and *B. subtilis*).

In contrast, a comparative amino acid analysis of the 65.8-kDa DnaK protein of *B. subtilis* revealed marked sequence similarity to the DnaK proteins of other bacterial species. Overall, the amino acid identity between *B. subtilis* DnaK and the others ranges from 59.4% (*E. coli*) to 91.2% (*B. megaterium*). A relaxation to include conservative amino acid changes increased similarity from 71.6% (*E. coli*) to 92.9% (*B. megaterium*). The homology declined substantially towards the carboxy terminus (data not shown). This divergence is a feature shown by all members of the HSP70 family.

A novel feature of the DnaK protein of *B. subtilis* which is also common to other gram-positive bacteria examined so far is the presence of a large deletion of 24 amino acid residues near the N terminus (Fig. 4B). Since this deletion is not present in DnaK from gram-negative bacteria and from any other species, this finding indicates that it might be a unique and distinguishing characterization of the gram-positive bacteria.

Features of the noncoding regions. Probable ribosome-binding sites for *orf39* and *dnaJ*, marked with an asterisk in Fig. 3, are 4 and 5 nucleotides upstream from the start codons of ORF1 and ORF4, respectively. These sites consist of 10 (ORF1) and 5 (ORF4) nucleotides complementary to the 3′ end of *B. subtilis* 16S rRNA (41).

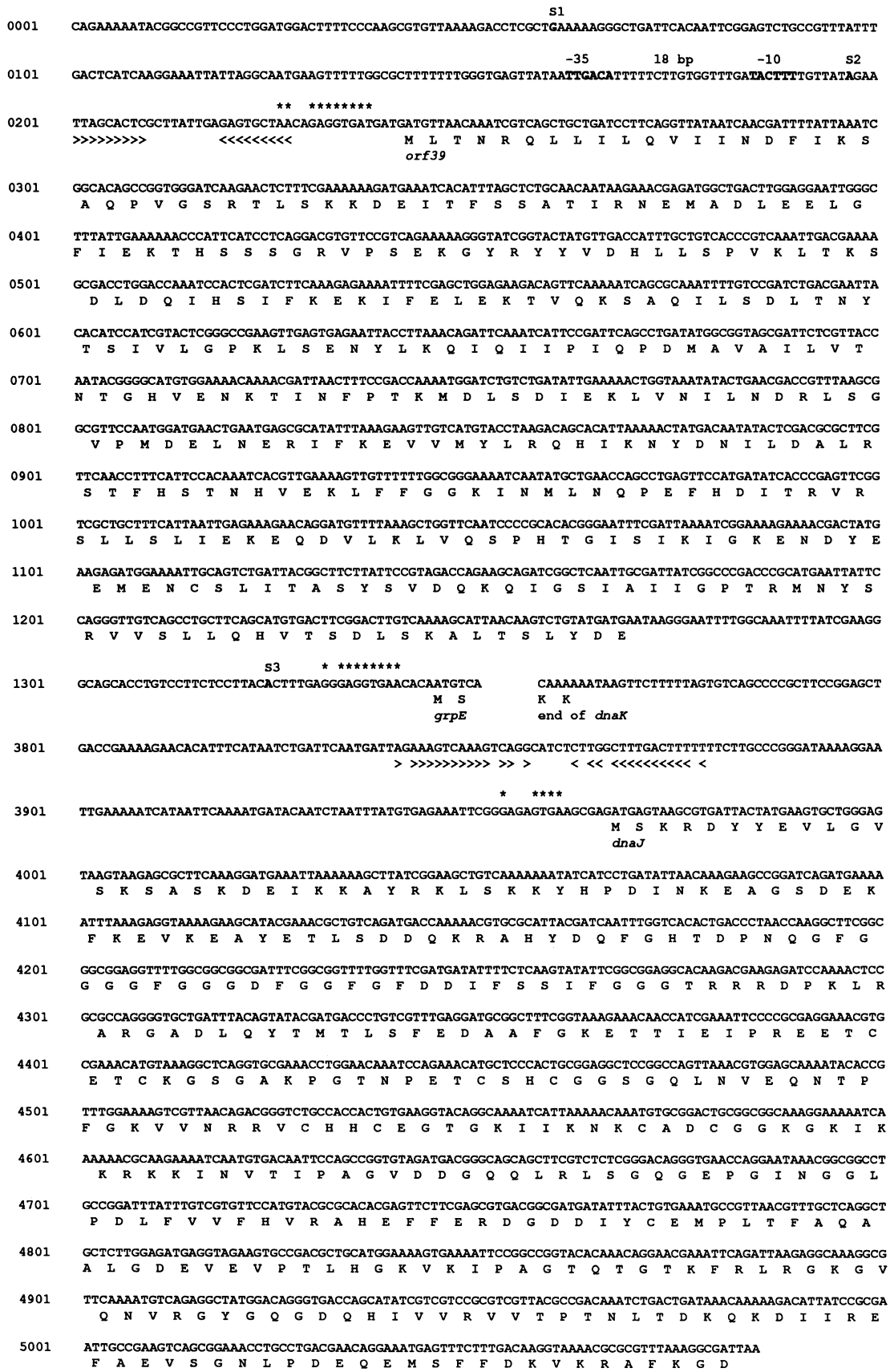
S1
-35
18 bp
-10
S2
** ********
orf39
S3
* ********
grpE
end of dnaK
dnaJ

FIG. 3. DNA sequence of *orf39* and of *dnaJ* of the *dnaK* locus and the deduced amino acid sequence. The sequences of *grpE* (63) and *dnaK* (62) have already been published and are not shown here. Indicated are the putative Shine-Dalgarno sequences (asterisks above the sequence), one potential sigma-43-dependent promoter (bold letters), and inverted repeats (arrowheads below the sequence). The 5′ ends of mRNAs (S1 though S3) as identified by primer extension are marked by bold letters.

Primer extension studies were used to identify the transcription start site(s) for the four ORFs. Total RNA was prepared from bacterial cells grown at 37°C or at different times after heat shock at 48°C. Synthetic oligonucleotides derived from the coding regions of *orf39*, *grpE*, and *dnaJ* were used as primers. The analysis revealed three potential transcription start points. The two major start points are located 45 nucleotides upstream of the translation initiation codon of *orf39* (designated S2; Fig. 3 and 5A) and 21 nucleotides upstream of the AUG initiation codon of *grpE* (S3; Fig. 3 and 5B). Only very few transcription starts originated from a weaker start point 179 nucleotides upstream of the AUG initiation start codon of *orf39* (S1 [Fig. 3]; data not shown). The number of transcripts initiated at S2 and S3 increased with the temperature shift from 37 to 48°C. The level of transcription from S1 remained unchanged and low (data not shown).

The transcriptional start point S2 is preceded by sequences which resemble the consensus sequence of promoters recognized by the vegetative sigma factor of *B. subtilis*, sigma-43. In contrast, S1 and S3 are not preceded by sequences resembling any known promoter. Heat-sensitive transcription from S2 might also be initiated by sigma-43. However, we found a palindromic structure downstream of S2 consisting of a 9-bp repeat separated by 9 bp (Fig. 3, nucleotides 201 to 227). This palindromic structure could be involved in the temperature-sensitive regulation of transcription initiation at S2.

An inverted repeat, extending from position 3840 to position 3878 (Fig. 3) could form a potential stem-loop structure and thereby act as a weak terminator ($\Delta G_0 = -13$ kcal).

Synthesis of mRNA. The in vivo transcripts of the *dnaK* locus were detected by dot blot and Northern analyses with total RNA prepared either from exponentially growing cells of *B. subtilis* or from bacteria at different times after a heat shock at 48°C. The dot blot experiments revealed a rapid and at-least-fivefold increase in the mRNA level of *orf39* (Fig. 6) after the temperature upshift. Five minutes after the heat shock, the amount of mRNA reached a maximum, and thereafter, it declined rapidly to the steady-state level.

The Northern analysis was performed by using a radioactively labeled *Bgl*II fragment of *dnaK* as the probe. The 0.8-kb *dnaK* fragment hybridized to three mRNA species 4.9, 2.6, and 1.5 kb long (Fig. 7). The 4.9-kb RNA could comprise the whole operon, whereas the 2.6-kb RNA could be the transcript of genes *grpE* and *dnaK*. The smallest transcript could be initiated at S2 and extend into *dnaK*. However, only the 2.6-kb mRNA was clearly detectable during exponential growth, and its intensity increased after a temperature upshift (see Discussion).

Analysis of a chromosomal transcriptional *dnaK-amyL* fusion. Strain JDB2 carries in its chromosome a transcriptional fusion in which a promoterless α-amylase gene has been fused to the *dnaK* gene at the *Bgl*II site at position 2656 (Fig. 2). By this construction, the 89 C-terminal amino acid residues have been replaced by 5 novel amino acid residues, thereby altering the C terminus of the DnaK protein. Strain JDB1 carries the reporter gene in the opposite orientation

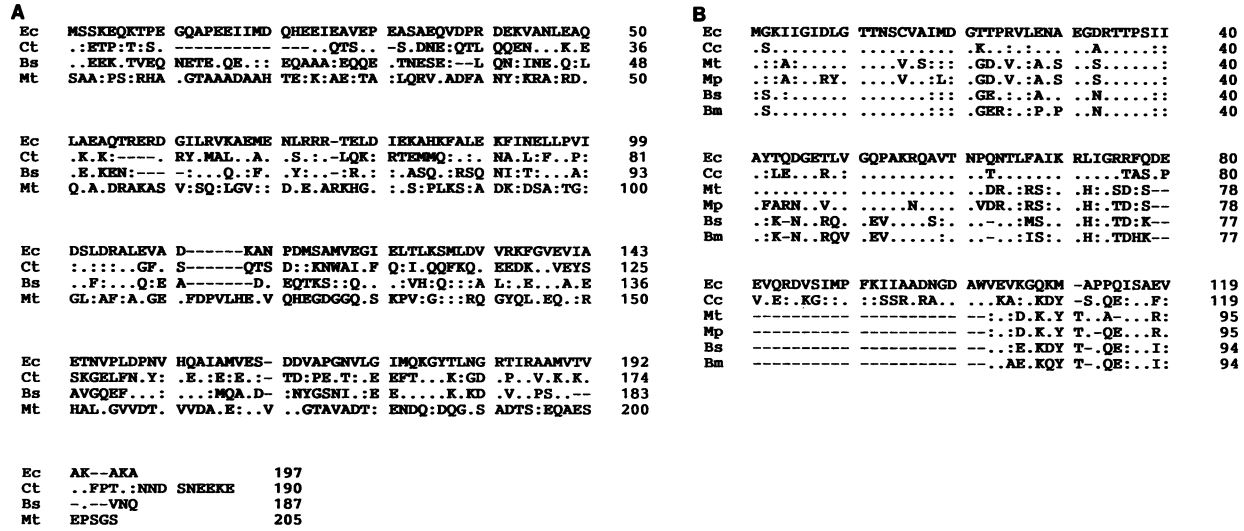

FIG. 4. Comparison of the deduced amino acid sequences encoded by genes *grpE* and *dnaK* from different bacterial species. (A) Complete amino acid sequences of GrpE from *E. coli* (Ec) (36), *C. trachomatis* (Ct) (17), *B. subtilis* (Bs) (63), and *M. tuberculosis* (Mt) (68). (B) N-terminal amino acid sequence of DnaK from Ec (8), *Caulobacter crescentus* (Cc) (24), Mt (67, 68), *M. paratuberculosis* (Mp) (58), *M. leprae* (Ml) (21), Bs (62), and *B. megaterium* (Bm) (60). Amino acids are given in the standard one-letter code. Amino acids identical with those of *E. coli* are indicated by a point, and conservative replacements are indicated by a colon. Conservative replacements are defined as being within the following groups: D and E; H, K, and R; S and T; Q and N; A and G; F, Y, and W; and I, L, V, and M (37). Gaps, indicated by dashes, are introduced in order to obtain a maximum fit.

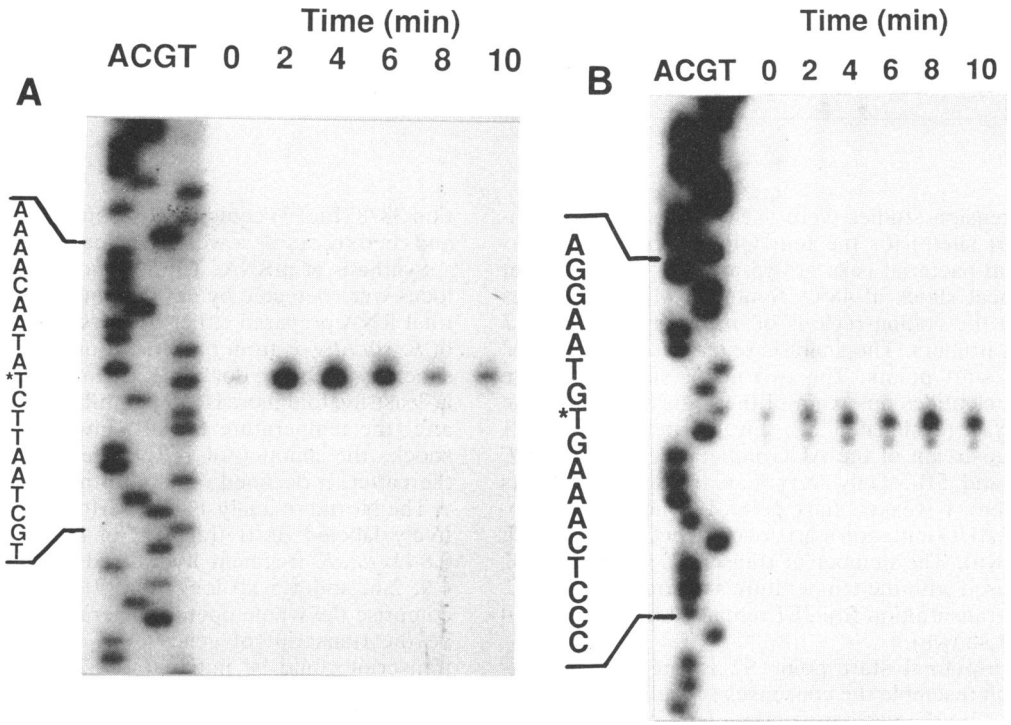

FIG. 5. Location of start sites for transcripts in the *dnaK* locus. The autoradiogram of a 6% polyacrylamide sequencing gel used to analyze the primer extension products is shown. Equal amounts of RNA isolated before (0 min) or at 2, 4, 6, 8, and 10 min after heat shock (48°C) were hybridized with primers complementary to the 5′ region of *orf39* (A) and *grpE* (B) and extended with avian myeloblastosis virus reverse transcriptase as described in Materials and Methods. Lanes A, C, G, and T show the dideoxy-sequencing ladder obtained with the same primers.

relative to the *dnaK* gene, thereby preventing its expression from promoters of the *dnaK* locus. Both chromosomal arrangements have been confirmed by Southern blotting (data not shown).

At a low temperature, the reporter gene was expressed, resulting in about 33 U (Fig. 7). After a temperature upshift, α-amylase activity increased to about 45 U. The control strain JDB1 exhibited a lower but significant activity (about 12 U), most probably due to expression from the plasmid promoter(s). This activity could not be increased by a temperature upshift (Fig. 8). From these results, we conclude that the *dnaK* gene is preceded by at least one promoter exhibiting temperature dependence.

Location of *dnaK* on the *B. subtilis* chromosome. The position of *dnaK* on the chromosome was determined by phage PBS1-mediated transduction, by using the chloramphenicol resistance gene in the *dnaK*::*cat* mutant (strains JDA1 and JDA2) as a selectable marker. *dnaK*::*cat* was 84% (JDA1) and 81% (JDA2) cotransduced with the auxotrophic mutation *aroD*, and it was 6% (JDA1) and 7% (JDA2) cotransduced with the *lys* marker. These results indicate that the order of genes is *aroD-dnaK-lys* and that the position of *dnaK* is at about 223° on the genetic map of Henner and Hoch (28).

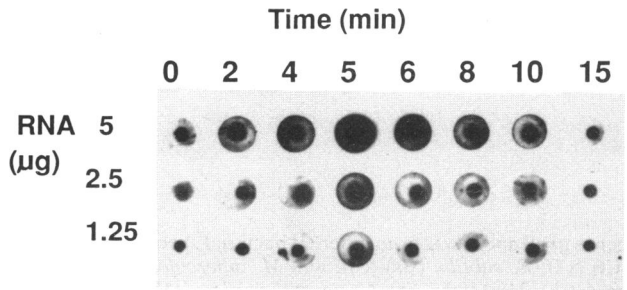

FIG. 6. Concentration of mRNA of the *dnaK* locus. Dot blot analyses of RNA isolated before (0 min) and at different times after heat shock (2, 4, 6, 8, 10, and 15 min). A radioactively labeled DNA fragment containing *dnaK* was used as the probe.

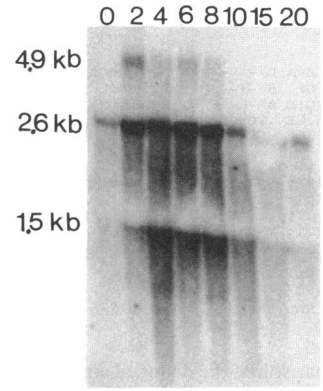

FIG. 7. Northern blot analysis. RNA was isolated from *B. subtilis* YB886 at 37°C (0 min) and at 2, 4, 6, 8, 10, 15, and 20 min after a temperature shift to 48°C. The filters were hybridized with a radioactively labeled DNA fragment containing *dnaK*. For details, see Materials and Methods.

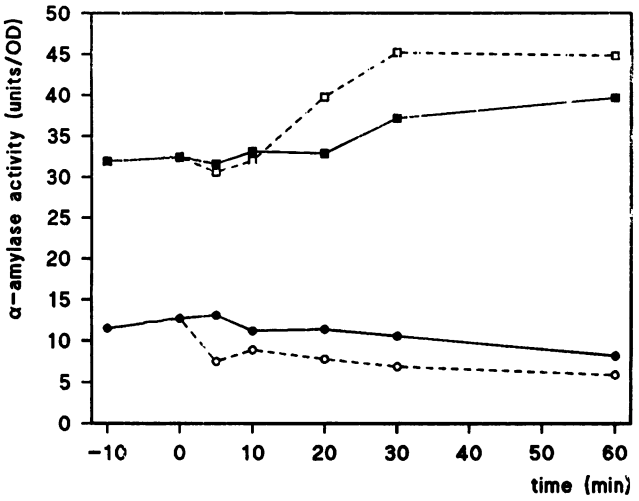

FIG. 8. α-Amylase activity of a transcriptional *dnaK-amyL* fusion. Strains JDB2 and JDB1 were grown at 30°C to mid-log phase in LB medium. Then the cultures were divided into two halves at time zero and grown further at 30 and 48°C, respectively. At the indicated times, samples were taken and α-amylase was assayed as described previously (45). Strain JDB2 grown at 30°C (●) and 48°C (○); strain JDB1 grown at 30°C (■) and 48°C (□).

DISCUSSION

We have cloned, sequenced, and characterized the *dnaK* locus of *B. subtilis*. This locus encodes four heat shock genes, and the genomic organization is *orf39-grpE-dnaK-dnaJ*. This represents the first identified complete heat shock operon within *Bacillus* spp. The genomic organization of these heat shock genes differs from that in *E. coli* in two aspects. First, the *grpE* gene is part of the *dnaK-dnaJ* operon. In *E. coli*, these three genes form two independent transcriptional units (50, 52). Second, there is a fourth potential heat shock gene, *orf39*, of still unknown biological function. By using an internal part of *orf39* as a gene probe, Southern blotting against chromosomal DNA of *E. coli* and the Kohara library of ordered lambda clones (33) failed to detect any specific hybridization signal (70). From these results, we conclude that either the DNA homology between *orf39* and the homologous gene in *E. coli* is not sufficient to yield a detectable hybridization signal or that this gene does not exist in the *E. coli* chromosome.

In *M. tuberculosis*, the *grpE* gene is closely linked to *dnaK* and *dnaJ*, as deduced from the nucleotide sequence (68). In *Clostridium acetobutylicum*, the same genomic organization has been found (43). Furthermore, the *orf39* homologous gene, *orfA*, is 31.6% identical and 59.5% homologous at the amino acid level (6). Therefore, the question arises as to whether an *orf39-orfA* homologous gene can be found within all gram-positive bacteria.

The alignment of GrpE proteins from four bacterial species revealed a lower overall homology among these proteins, as has been described for other HSPs such as GroEL or DnaK (see below). Within the GrpE polypeptides, there are two regions of increased homology. Amino acid residues within these regions could be involved in interaction with other proteins. One candidate is the DnaK protein, which has been shown in *E. coli* to functionally interact with GrpE both in vitro and in vivo (31, 71). Another candidate could be the DnaJ protein.

Alignment of the DnaK proteins of various bacterial species revealed a substantial overall homology, which declines towards the C-terminal end. Since the publication of the partial amino acid sequence of the *M. tuberculosis* DnaK protein, sequences of two more proteins from *Mycobacterium paratuberculosis* and *Mycobacterium leprae* have been published (21, 58, 67). Alignment of the C termini of the three proteins revealed significant homology between the *M. leprae* and *M. paratuberculosis* DnaK proteins but no homology at all between these two and the *M. tuberculosis* polypeptide. Inspection of the three reading frames deduced from the DNA sequence of the *M. tuberculosis* gene (68) at the point of dissimilarity showed that the homology can be restored by switching to another reading frame. Most probably, a sequencing error resulted in a frameshift mutation (data not shown).

Most interestingly, the DnaK protein of *B. subtilis*, those of the other gram-positive species shown in Fig. 4B, and those of *C. acetobutylicum* (43) have a deletion of 24 contiguous amino acids near the N-terminal end. Flaherty et al. (20) have reported the three-dimensional crystal structure of the N-terminal 386 amino acids of bovine HSC70, which retains ATPase activity. The crystal structure of the fragment revealed the presence of two distinct lobes of approximately equal sizes with a deep cleft in between and very similar tertiary structures. The 24 amino acids (*E. coli* residues 78 to 102; Fig. 4B) corresponding to the gap are located on the outside of lobe 1 and appear as an appendage for which the biological function is unknown. Experiments are in progress to delete the coding region for these 24 amino acid residues from the *E. coli dnaK* gene and to analyze the deletion derivative for its biological and biochemical activities.

Analyses of transcription of the *dnaK* locus revealed three mRNAs of various lengths and three potential transcription start sites. The 4.9-kb mRNA, clearly detected only after a temperature upshift, could correspond by length to the transcript(s) initiated at S1 and/or S2 and extending from *orf39* through *dnaJ*. Our failure to demonstrate this RNA species during Northern analysis could be based on its low abundance. Alternatively, transcripts initiated at S1 and/or S2 could be terminated prematurely. There was a significant increase in the amount of 4.9-kb mRNA upon heat induction (Fig. 7). We conclude that these transcripts were initiated at S2 rather than S1, since only S2 is activated after heat induction (Fig. 5A). The 2.6-kb mRNA could be initiated at S3, and it corresponds by length to genes *grpE* and *dnaK*. The 1.5-kb mRNA most probably constitutes a specific degradation product of the 4.9- and/or 2.6-kb transcript(s). These tentative conclusions have to be confirmed by mapping the 5′ end of the 1.5-kb RNA and the 3′ ends of all three mRNAs.

Expression of the *dnaK* locus at 37°C and after heat induction was examined by three different methods. Analysis of the amount of mRNA by dot blot using two different gene probes revealed an increase in the amount of mRNA between 0 and 5 min after induction and then a decrease to the uninduced level within 10 min (Fig. 6). Primer extension carried out with mRNA prepared before and at different times after heat induction exhibited a similar increase and decrease in the amount of specific mRNA (Fig. 5). This result clearly shows that the heat shock response is regulated at the level of transcription, as has been reported for all other organisms. Analysis of a transcriptional *dnaK-amyL* fusion resulted in an increase in the amount of α-amylase activity after a temperature upshift (Fig. 8). It should be mentioned that the construction of the chromosomal *dnaK-amyL* tran-

TABLE 2. Consensus inverted repeat sequences in different bacterial species

Organism	Gene or protein	Inverted Repeat	Source or reference
B. subtilis	*orf39*	TTAGCACTC-N_9-GAGTGCTAA	This work
B. subtilis	*groESL*	TTAGCACTC-N_9-GAGTGCTAA	56
C. acetobutylicum	*orfA*	TTAGCACTC-N_9-GAGTGCTAA	43
C. acetobutylicum	*groESL*	TTAGCACTC-N_9-GAGTGCTAA	5
Synechococcus sp.	*groESL*	TTAGCACTC-N_9-GAGTGCTAA	61
Synechococcus sp.	*urf3*, *urf4*	TTAGCACTC-N_9-GAGTGCTAA	13
Synechocystis sp.	*cpn-60*	TTAGCACTC-N_9-GAGTGCTAA	11
M. tuberculosis	10-kDa AG[a]	TcAGCACTC-N_9-GAGTGCTAc	7
M. tuberculosis	65-kDa AG	cTtGCACTC-N_9-GAGTGCTAA	57
M. bovis BCG	MPB57	cTAGCACTC-N_9-GAGTGCTAg	65
M. leprae	65-kDa AG	cTgGCACTC-N_9-GAGTGCcAg	38
C. psittaci	*hypA*, *hypB*	gTAGCACTt-N_9-aAGTGCTAA	42
Consensus		TTAGCACTC-N_9-GAGTGCTAA	

[a] AG, antigen.

scriptional fusion resulted in the formation of a slightly truncated version of DnaK lacking its 89 C-terminal amino acids. Furthermore, by the integration of a plasmid, the *dnaK* gene was separated from the *dnaJ* gene. Therefore, we cannot exclude the possibility that, by these in vitro manipulations, the enzymatic activities of DnaK and/or the expression of *dnaJ* is impaired. In *E. coli*, it could be shown that a truncated DnaK protein lacking 74 amino acid residues of its carboxyl terminus showed all the enzymatic activities attributed to the complete DnaK protein (9).

Induction of the heat shock response is regulated at the level of transcription. In *E. coli*, the heat shock genes are positively regulated by the *rpoH* gene, whose product, a 32-kDa sigma factor, directs core RNA polymerase to recognize the promoters for heat shock genes (12, 25, 34). In *B. subtilis*, only one transcription start site is preceded by a promoter sequence recognized by sigma-43, the vegetative sigma factor. No nucleotide sequences similar to those recognized by other sigma factors could be identified in front of S1 and S3. In addition, there are no indications that alternate transcription start points are activated upon heat induction as shown for *Caulobacter crescentus* (24). At present, we cannot exclude the possibility that a heat shock promoter is overlapping with the vegetative promoter. The failure to detect a promoter-like sequence upstream of S3 could indicate that S3 is not a transcription start site but instead a processing site. This interpretation is strengthened by the finding that the kinetics of primer extension indicate a shift in the maximum mRNA synthesis from 4 to 8 min postinduction (Fig. 5A and B).

Most interestingly, we detected a 9-bp inverted repeat separated by a 9-bp spacer located between the transcription start point and the putative start codon (Fig. 3). The same inverted repeat has been found in front of the *groESL* of *B. subtilis* (56) and the *dnaK* and *groESL* operons of *C. acetobutylicum* (5, 43). A data base search for additional genes preceded by this inverted repeat led to the discovery of some more examples, which are listed in Table 2. In most cases, this inverted repeat can be found in front of heat shock genes. This finding strongly suggests that this inverted repeat is involved in the regulation of the heat shock response by, e.g., stabilizing the mRNA against degradation or acting as a binding site for a DNA or RNA binding protein. Baird et al. already described an inverted repeat preceding the 10-kDa antigen of *M. tuberculosis* and pointed to the possibility that this inverted repeat might be involved in the regulation of the heat shock response (7). Experiments are in progress to alter the DNA sequence of one half of the inverted repeat and to analyze the thus-obtained mutants with respect to heat shock induction.

ACKNOWLEDGMENTS

We thank D. J. Henner, D. McConnell, G. Rapoport, and P. Setlow for providing strains and plasmids. We are grateful to H. Bahl for sharing unpublished results on the *dnaK* locus of *C. acetobutylicum*.

This work was supported by grants from the BMFT to W.S., from the Fonds der Chemischen Industrie to M.H., and from the Alexander von Humboldt-Stiftung to U.V.

REFERENCES

1. **Ambulos, Jr., N. P., E. J. Duvall, and P. S. Lovett.** 1987. Method for blot hybridization analysis of mRNA molecules from *Bacillus subtilis*. Gene **51**:281–286.
2. **Anagnostopoulos, C., and J. Spizizen.** 1961. Requirements for transformation in *Bacillus subtilis*. J. Bacteriol. **81**:741–746.
3. **Antoniewsky, C., B. Savelli, and P. Stragier.** 1990. The *spo*IIJ gene, which regulates early developmental steps in *Bacillus subtilis*, belongs to a class of environmentally responsive genes. J. Bacteriol. **172**:86–93.
4. **Arnosti, D. N., V. L. Singer, and M. J. Chamberlin.** 1986. Characterization of heat shock in *Bacillus subtilis*. J. Bacteriol. **168**:1243–1249.
5. **Bahl, H.** Unpublished data.
6. **Bahl, H., and W. Schumann.** Unpublished data.
7. **Baird, P. N., L. M. C. Hall, and A. R. M. Coates.** 1989. Cloning and sequence analysis of the 10 kDa antigen gene of *Mycobacterium tuberculosis*. J. Gen. Microbiol. **135**:931–939.
8. **Bardwell, J. C. A., and E. A. Craig.** 1984. Major heat shock gene of *Drosophila* and the *Escherichia coli* heat-inducible *dnaK* gene are homologous. Proc. Natl. Acad. Sci. USA **81**:848–852.
9. **Cegielska, A., and C. Georgopoulos.** 1989. Functional domains of the *Escherichia coli dnaK* heat shock protein as revealed by mutational analysis. J. Biol. Chem. **264**:21122–21130.
10. **Chang, A. C. Y. C., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**: 1141–1156.
11. **Chitnis, P. R., and N. Nelson.** 1991. Molecular cloning of the genes encoding two chaperone proteins of the cyanobacterium *Synechocystis* sp. PCC 6803. J. Biol. Chem. **266**:58–65.
12. **Cowing, D. W., J. C. A. Bardwell, E. A. Craig, C. Woolford, R. W. Hendrix, and C. A. Gross.** Consensus sequence for

Escherichia coli heat shock gene promoters. Proc. Natl. Acad. Sci. USA **82**:2679–2683.
13. **Cozens, A. L., and J. E. Walker.** 1987. The organization and sequence of the genes for ATP synthase subunits in the cyanobacterium *Synechococcus* 6301. J. Mol. Biol. **194**:359–383.
14. **Craig, E. A.** 1985. The heat shock response. Crit. Rev. Biochem. **18**:239–280.
15. **Dedonder, R. A., J.-A. Lepesant, J. Lepesant-Kejzlarova, A. Billault, M. Steinmetz, and F. Kunst.** 1977. Construction of a kit of reference strains for rapid genetic mapping in *Bacillus subtilis* 168. Appl. Environ. Microbiol. **33**:989–993.
16. **Dubnau, D., and R. Davidoff-Abelson.** 1971. Fate of transforming DNA following uptake by competent *Bacillus subtilis*. I. Formation and properties of the donor-recipient complex. J. Mol. Biol. **56**:209–221.
17. **Engel, J. N., J. Pollack, E. Perara, and D. Ganem.** 1990. Heat shock response of murine *Chlamydia trachomatis*. J. Bacteriol. **172**:6959–6972.
18. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
19. **Ferrari, F. A., A. Nguyen, D. Lang, and J. D. Hoch.** 1983. Construction and properties of an integrable plasmid for *Bacillus subtilis*. J. Bacteriol. **154**:1513–1515.
20. **Flaherty, K. M., C. DeLuca-Flaherty, and D. B. McKay.** 1990. Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. Nature (London) **346**:623–628.
21. **Garsia, R. J., L. Hellqvist, R. J. Booth, A. J. Radford, W. J. Britton, L. Astbury, R. J. Trent, and A. Basten.** 1989. Homology of the 70-kilodalton antigens from *Mycobacterium leprae* and *Mycobacterium bovis* with the *Mycobacterium tuberculosis* 71-kilodalton antigen and with conserved heat shock protein 70 of eucaryotes. Infect. Immun. **57**:204–212.
22. **Georgopoulos, C. P.** 1977. A new bacterial gene (*groPC*) which affects λ DNA replication. Mol. Gen. Genet. **151**:35–39.
23. **Georgopoulos, C. P., and I. Herskowitz.** 1971. *Escherichia coli* mutants blocked in lambda DNA synthesis, p. 553–564. *In* A. D. Hershey (ed.), The bacteriophage Lambda. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **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.
25. **Grossman, A. D., J. W. Erickson, and C. A. Gross.** 1984. The *htpR* gene product of *E. coli* is a sigma factor for heat-shock promoters. Cell **38**:383–390.
26. **Hearne, C. M., and D. J. Ellar.** 1989. Nucleotide sequence of a *Bacillus subtilis* gene homologous to the *dnaK* gene of *Escherichia coli*. Nucleic Acids Res. **17**:8373.
27. **Heinrich, P.** 1986. Guidelines for quick and simple plasmid sequencing. Boehringer GmbH, Mannheim, Germany.
28. **Henner, D. J., and J. A. Hoch.** 1980. The *Bacillus subtilis* chromosome. Microbiol. Rev. **44**:57–82.
29. **Hoch, J. A., M. Barat, and C. Anagnostopoulos.** 1967. Transformation and transduction in recombination-defective mutants of *Bacillus subtilis*. J. Bacteriol. **93**:1925–1937.
30. **Itikawa, H., and J.-I. Ryu.** 1979. Isolation and characterization of a temperature-sensitive *dnaK* mutant of *Escherichia coli* B. J. Bacteriol. **138**:339–344.
31. **Johnson, C., G. N. Chandrasekhar, and C. Georgopoulos.** 1989. *Escherichia coli* DnaK and GrpE heat shock proteins interact both in vivo and in vitro. J. Bacteriol. **171**:1590–1596.
32. **Karamata, D., and J. D. Gross.** 1970. Isolation and genetic analysis of temperature-sensitive mutants of *B. subtilis* defective in DNA synthesis. Mol. Gen. Genet. **108**:277–287.
33. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
34. **Landick, R., V. Vaughn, E. T. Lau, R. A. VanBogelen, J. W. Erickson, and F. C. Neidhardt.** 1984. Nucleotide sequence of the heat shock regulatory gene of *E. coli* suggests its protein product may be a transcription factor. Cell **38**:175–182.
35. **Lindquist, S., and E. A. Craig.** 1988. The heat-shock proteins. Annu. Rev. Genet. **22**:631–677.
36. **Lipinska, B., J. King, D. Ang, and C. Georgopoulos.** 1988. Sequence analysis and transcriptional regulation of the *Escherichia coli grpE* gene, encoding a heat shock protein. Nucleic Acids Res. **16**:7545–7562.
37. **Margolin, W., and M. M. Howe.** 1986. Localization and DNA sequence analysis of the C gene of bacteriophage Mu, the positive regulator of Mu late transcription. Nucleic Acids Res. **14**:4881–4897.
38. **Mehra, V., D. Sweetser, and R. A. Young.** 1986. Efficient mapping of protein antigenic determinants. Proc. Natl. Acad. Sci. USA **83**:7013–7017.
39. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
40. **Miller, B. S., T. E. Kennedy, and U. N. Streips.** 1991. Molecular characterization of specific heat shock proteins in *Bacillus subtilis*. Curr. Microbiol. **22**:231–236.
41. **Moran, Jr., C. P., N. Lang, S. F. J. LeGrice, 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.
42. **Morrisson, R. P., R. J. Belland, K. Lyng, and H. D. Caldwell.** 1989. Chlamydial disease pathogenesis. The 57-kD chlamydial hypersensitivity antigen is a stress response protein. J. Exp. Med. **170**:1271–1283.
43. **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. **174**:3290–3299.
44. **Neidhardt, F. C., R. A. VanBogelen, and V. Vaughn.** 1984. The genetics and regulation of heat-shock proteins. Annu. Rev. Genet. **18**:295–329.
45. **O'Kane, C., M. A. Stephens, and D. McConnell.** 1986. Integrable α-amylase plasmid for generating random transcriptional fusions in *Bacillus subtilis*. J. Bacteriol. **168**:973–981.
46. **Paek, K.-H., and G. C. Walker.** 1987. *Escherichia coli dnaK* null mutants are inviable at high temperature. J. Bacteriol. **169**:283–290.
47. **Parker, J.** 1989. Errors and alternatives in reading the universal genetic code. Microbiol. Rev. **53**:273–298.
48. **Puyet, A., H. Sandoval, P. Lopez, A. Aguilar, J. F. Martin, and M. Espinosa.** 1987. A simple medium for rapid regeneration of *Bacillus subtilis* protoplasts transformed with plasmid DNA. FEMS Microbiol. Lett. **40**:1–5.
49. **Richter, A., and M. Hecker.** 1986. Heat-shock proteins in *Bacillus subtilis*: a two-dimensional gel electrophoresis study. FEMS Microbiol. Lett. **36**:69–71.
50. **Saito, H., Y. Nakamura, and H. Uchida.** 1978. A transducing λ phage carrying *grpE*, a bacterial gene necessary for λ DNA replication, and two ribosomal protein genes, *rpsP* (S16) and *rplS* (L19). Mol. Gen. Genet. **165**:247–256.
51. **Saito, H., and H. Uchida.** 1977. Initiation of the DNA replication of bacteriophage lambda in *Escherichia coli* K12. J. Mol. Biol. **113**:1–25.
52. **Saito, H., and H. Uchida.** 1978. Organization and expression of the *dnaJ* and *dnaK* genes of *Escherichia coli* K12. Mol. Gen. Genet. **164**:1–8.
53. **Sakakibara, Y.** 1988. The *dnaK* gene of *Escherichia coli* functions in initiation of chromosome replication. J. Bacteriol. **170**:972–979.
54. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
55. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
56. **Schmidt, A., M. Schiesswohl, U. Völker, M. Hecker, and W. Schumann.** 1992. Submitted for publication.
57. **Shinnick, T. M.** 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis*. J. Bacteriol. **169**:1080–1088.
58. **Stevenson, K., N. F. Inglis, B. Rae, W. Donachie, and J. M. Sharp.** 1991. Complete nucleotide sequence of a gene encoding

the 70 kd heat shock protein of *Mycobacterium paratuberculosis*. Nucleic Acids Res. **19:**4552.

59. **Streips, U. N., and F. W. Polio.** 1985. Heat shock proteins in bacilli. J. Bacteriol. **162:**434–437.
60. **Sussman, M. D., and P. Setlow.** 1987. Nucleotide sequence of a *Bacillus megaterium* gene homologous to the *dnaK* gene of *Escherichia coli*. Nucleic Acids Res. **15:**3923.
61. **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.
62. **Wetzstein, M., Dedio, J., and W. Schumann.** 1990. Complete nucleotide sequence of the *Bacillus subtilis dnaK* gene. Nucleic Acids Res. **18:**2172.
63. **Wetzstein, M., and W. Schumann.** 1990. Nucleotide sequence of a *Bacillus subtilis* gene homologous to the *grpE* gene of *E. coli* located immediately upstream of the *dnaK* gene. Nucleic Acids Res. **18:**1289.
64. **Wood, H. E., M. T. Dawson, K. M. Devine, and D. J. McConnell.** 1990. Characterization of PBSX, a defective prophage of *Bacillus subtilis*. J. Bacteriol. **172:**2667–2674.
65. **Yamaguchi, R., K. Matsuo, A. Yamazaki, S. Nagai, K. Terasaka, and T. Yamada.** 1988. Immunogenetic protein MPB57 from *Mycobacterium bovis* BCG: molecular cloning, nucleotide sequence and expression. FEBS Lett. **240:**115–117.
66. **Yanish-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.
67. **Young, D., R. Lathigra, R. Hendrix, D. Sweetser, and R. A. Young.** 1988. Stress proteins are immune targets in leprosy and tuberculosis. Proc. Natl. Acad. Sci. USA **85:**4267–4270.
68. **Young, R. A.** Personal communication.
69. **Zeilstra-Ryalls, J., O. Fayet, and C. Georgopoulos.** 1991. The universally conserved GroE (HSP60) chaperonins. Annu. Rev. Microbiol. **45:**301–325.
70. **Zuber, U.** Unpublished results.
71. **Zylicz, M., D. Ang, and C. Georgopolous.** 1987. The *grpE* protein of *Escherichia coli*. Purification and properties. J. Biol. Chem. **262:**17437–17442.