JOURNAL OF BACTERIOLOGY, May 1993, p. 2613–2624
0021-9193/93/092613-12$02.00/0

Vol. 175, No. 9

The *Escherichia coli* Heat Shock Gene *htpY*: Mutational Analysis, Cloning, Sequencing, and Transcriptional Regulation

DOMINIQUE MISSIAKAS,* COSTA GEORGOPOULOS, AND SATISH RAINA

Département de Biochimie Médicale, Centre Médical Universitaire, 1 Rue Michel-Servet, 1211 Genève 4, Switzerland, and Department of Cellular, Viral and Molecular Biology, University of Utah School of Medicine, Salt Lake City, Utah 84132*

Received 22 December 1992/Accepted 1 March 1993

We have identified a new heat shock gene, designated *htpY*, located 700 bp upstream of the *dnaK dnaJ* operon. We cloned it and showed that it is transcribed clockwise vis-à-vis the *Escherichia coli* genetic map, in the same direction as the *dnaK dnaJ* operon. The *htpY* gene encodes a 21,193-Da polypeptide. Promoter mapping experiments and Northern (RNA) analysis showed that the *htpY* gene belongs to the classical heat shock gene family, because the transcription from its major promoter is under the positive control of the *rpoH* gene product (σ^{32}) and resembles canonical Eσ^{32}-transcribed consensus promoter sequences. This conclusion has been strengthened by the construction and analysis of a p*htpY-lacZ* promoter fusion. Despite the fact that *htpY* null bacteria are viable, the expression of various Eσ^{32} heat shock promoters is significantly decreased, suggesting that HtpY plays an important role in the regulation of the heat shock response. Consistent with this interpretation, overproduction of the HtpY protein results in a generalized increase of the heat shock response in *E. coli*.

The heat shock response is one of the most highly conserved biological responses to environmental changes. Its purpose is to allow the cell to deal with the immediate and long-term consequences of otherwise injurious growth conditions. As an example, in *Escherichia coli*, a sudden upshift in temperature from 30°C to 42°C or above induces a transient acceleration in the synthesis of approximately 20 proteins (13, 24). This phenomenon is under the positive regulation of the *rpoH* (*htpR*) gene product, σ^{32}, an alternate sigma subunit of the RNA polymerase enzyme (E). Recently, a new heat shock regulon, under the positive control of the Eσ^{E} (σ^{24}) holoenzyme, has been discovered (7, 13, 37).

The most abundant heat shock proteins under Eσ^{32} regulation have been functionally and physiologically characterized. Some of them, e.g., DnaK, DnaJ, GroEL, GroES, and GrpE, were originally characterized as essential for bacteriophage λ propagation (11). It is now well known that they act as molecular chaperones, helping to ensure proper protein folding and trafficking (1, 11, 12). The products of the *groES* and *groEL* genes in particular are essential for *E. coli* growth at all temperatures (9).

Besides the major molecular chaperones (i.e., DnaK and GroEL), other heat shock proteins have been identified, including the essential sigma factor RpoD (σ^{70}), the Lon and ClpP proteases, HtpG (the eukaryotic equivalent of Hsp90) (13, 32), and others involved in different physiological pathways, e.g., the *lysU* (35) and *htrM* gene products (30). Despite this broad characterization of heat shock proteins, not all of them have been identified. Indeed, two-dimensional gel electrophoresis of a heat-shock bacterial sample shows a number of spots corresponding to as yet unidentified heat shock proteins.

The mechanistic details of heat shock regulation are critical to our understanding of the heat shock response. The regulatory gene *rpoH* is subjected to a variety of controls at both the transcriptional and posttranscriptional levels (13). For example, transcription of *rpoH* at 50°C occurs only from the p3 promoter (8). The regulation of transcription from the p3 promoter is not fully understood. Moreover, the gene(s) involved in its transcriptional regulation, presumably encoding the σ^{E} (σ^{24}) polypeptide, has not been identified so far. At the posttranscriptional level, it is well documented that mutations in the *dnaK*, *dnaJ*, or *grpE* gene result in constitutively elevated levels of heat shock proteins, their products playing a crucial role in the intracellular destabilization of the σ^{32} polypeptide (13).

Here, we report on the characterization of a new heat shock gene, *htpY*. The *htpY* gene is located 700 bp upstream of the *dnaK dnaJ* operon. It was cloned on the basis of our observations that a region located upstream of the *dnaK dnaJ* operon encodes a function antagonistic to that of DnaK, in terms of heat shock regulation. The *htpY* gene product was shown to positively mediate heat shock gene expression because (i) *htpY* null mutations result in a decrease in heat shock gene expression and (ii) overproduction of the HtpY protein leads to elevated levels of heat shock gene expression. These observations are interesting because it was recently reported that certain *dnaK* mutant bacteria presumably carry an additional mutation or mutations which map upstream of the *dnaK dnaJ* operon. A part of the temperature-sensitive phenotype of such *dnaK* mutant bacteria was ascribed to a mutation(s) in a putative gene, *htgA*, described as essential for bacterial growth at high temperatures (6). Although we show that the *htpY* and *htgA* genes are identical in terms of overall map location, null mutations in the *htpY* gene allow us to conclude that it is neither essential nor required for growth at high temperatures in many wild-type *E. coli* strains examined. Because we show that the *htpY* gene is heat shock regulated, specifically by the σ^{32} transcription factor, we use the conventional nomencla-

* Corresponding author.

TABLE 1. Bacterial strains and plasmids used in the study

Bacterial strains and plasmids	Relevant characteristic(s)[a]	Reference or source
Bacteria		
W3110	Wild type	Laboratory collection
MC4100	F⁻ *araD139* Δ(*argF-lac*)U169	3
MC1000	F⁻ *araD139* Δ(*lac*)X74	Laboratory collection
DH5α	F⁻ Φ80 Δ*lacZ* M15 *recA1*	Bethesda Research Laboratory
JC7623	AB1157 *recB21 recC22 sbcB15 sbcC201*	26
DM136	MC4100 *zab*::ΩKan^r inserted at the *Nco*I site	This work
DM407	MC4100 *htpY117*::ΩKan^r inserted at the *Eco*O109 site	This work
DM487	MC4100 Δ*htpY*(−271-959)::ΩKan^r; substitution of the *Nco*I-*Nco*I fragment	This work
DM78	MC4100 Φ(p*htpY-lacZ*) carrying the *Ava*I-*Ava*I fragment containing the *htpY* promoter region	This work
DM249	MC4100 Φ(p*htpY-lacZ*) carrying the *Fsp*I-*Apa*I fragment containing the *htpY* promoter region	This work
SR407	*zhf*::Tn*10 rpoH*$^+$ (KY1621 derivative constructed by P1 transduction)	29
KY1621	Δ*rpoH30*::Kan^r	42
SR381	MC4100 Δ*htrC*(21-575)::ΩTet^r	29
SR1208	MC4100 Φ(p*groE-lacZ*)	S. Raina, unpublished work
SR1421	MC4100 *lon*::λp*lac*Mu53	S. Raina, unpublished work
SR1418	MC4100 *htpG*::λp*lac*Mu53	S. Raina, unpublished work
SR1458	MC4100 Φ(p*htrA-lacZ*)	S. Raina, unpublished work
SR1588	MC4100 Φ(p*htrL-lacZ*)	21
SR1720	MC4100 Φ(p3*rpoH-lacZ*)	S. Raina, unpublished work
DM495	SR1208 Δ*htpY*(−271-959)::ΩKan^r	This work
DM496	SR1588 Δ*htpY*(−271-959)::ΩKan^r	This work
DM502	SR1720 Δ*htpY*(−271-959)::ΩKan^r	This work
DM504	SR1458 Δ*htpY*(−271-959)::ΩKan^r	This work
Plasmids		
pWSK29/30	Low-copy vector Amp^r	39
pBluescript SK and KS	High-copy cloning vector Amp^r	Stratagene
pDM38/39	pWSK29 with a 7-kb *Bam*HI fragment carrying *htpY dnaK dnaJ*; in opposite orientation with respect to T7 promoter	This work
pDM42	pSK carrying the 3.6-kb *Bam*HI-*Eco*RI *htpY*$^+$ fragment	This work
pDM49/50	pSK carrying the 2.2-kb *Bam*HI-*Bgl*II *htpY*$^+$ fragment; in opposite orientation with respect to T7 promoter	This work
pDM76	PRS528 (p*htpY-lacZ*) carrying the *Ava*I-*Ava*I fragment containing the *htpY* promoter	This work
pDM83/84	pSK carrying the 950-bp *Nco*I-*Bgl*II *htpY*$^+$ fragment; in opposite orientation with respect to T7 promoter	This work
pDM92	pDM42 ΩKan^r inserted at the *Nco*I site 271 bp upstream of *htpY* transcriptional start site	This work
pDM150	pSK carrying the 1.5-kb *Nco*I-*Mlu*I ORF-div$^+$ fragment	This work
PDM251	PRS528 (p*htpY-lacZ*) carrying the *Fsp*I-*Apa*I fragment containing the *htpY* promoter	This work
pDM335	pDM42 ΩKan^r inserted at the *Eco*O109 site 117 bp downstream of *htpY* transcriptional start site	This work
pDM462	pDM42 Δ*htpY*::ΩKan^r missing the 1,230-bp DNA fragment between the two *Nco*I sites	This work
pRS528	Promoter probe vector	31
pSR1232	pTTQ Kan^r carrying 1.3-kb *Eco*RV fragment containing the promoterless *rpoH* gene	This work

[a] The *htpY* allele numbers refer to nucleotide positions of the corresponding insertion relative to the start point of *htpY* transcription.

ture *htp*, meaning that this gene encodes a heat shock protein.

MATERIALS AND METHODS

Bacterial strains and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1.

Media. M9 minimal medium was prepared as described by Maniatis et al. (19) and was supplemented with glucose (0.2%), thiamine (2 μg/ml), $MgSO_4$ (1 mM), $MgCl_2$ (3 mM), $CaCl_2$ (0.1 mM), and $FeCl_3$ (0.3 μM). For [^{35}S]methionine labeling experiments, the M9 high-sulfur medium, supplemented with a mixture of defined amino acids, was used as previously described (34). When necessary, the media were supplemented with ampicillin (100 μg/ml), tetracycline (15 μg/ml), or kanamycin (50 μg/ml).

Enzymes. Restriction enzymes were purchased from Bethesda Research Laboratories, Inc. Avian myeloblastosis virus reverse transcriptase, sequenase enzyme, and T4 polynucleotide kinase were purchased from Stratagene, U.S. Biochemical Corporation, and International Biotechnologies Inc., respectively.

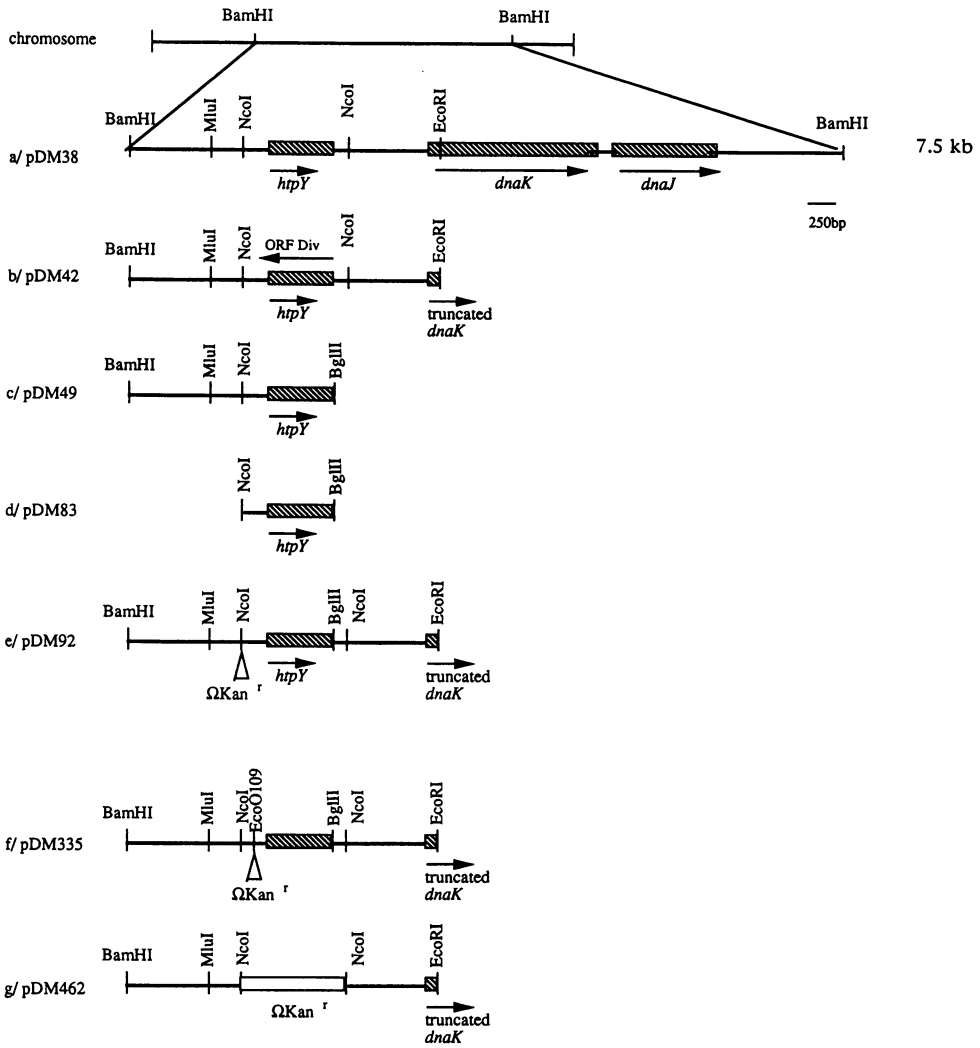

FIG. 1. Restriction map of the *htpY* gene and construction of *htpY* insertion and deletion derivatives. (a) Location of the *htpY* gene on the physical map of *E. coli*. (b through d) Clones carrying an intact *htpY* gene. (e) Map of pDM92 with the insertion of the ΩKan^r cassette upstream of the *htpY* gene, at the *Nco*I site. (f) Map of pDM335 with the insertion of the ΩKan^r cassette at the *Eco*O109 site, which lies between the +1 transcription start site and the putative translational start site. (g) Map of pDM462, which carries an ΩKan^r deletion substitution between the two *Nco*I sites in the *htpY* gene.

Cloning of the *htpY* gene. A 7.5-kb *Bam*HI fragment carrying the entire *dnaK dnaJ* operon and surrounding DNA regions was cloned from bacteriophage λ101 of the Kohara *E. coli* DNA library (15) into the low-copy vectors pWSK29 and pWSK30 (39), resulting in plasmids pDM38 and pDM39, respectively. DNA fragments from the upstream region of the *dnaK dnaJ* operon were derived from pDM38 and subcloned into pBluescript or the low-copy vectors pWSK29 and pWSK30 with standard DNA manipulation techniques (19) (Fig. 1). These subclones were subsequently characterized according to their ability to express the *htpY* gene product.

Construction of *htpY*::ΩKan^r insertions and deletions in the *htpY* gene. Three insertional mutations were constructed by inserting an ΩKan^r cassette (10) by in vitro DNA manipulation of the *htpY*$^+$-bearing plasmid pDM42 (Fig. 1). The first one resulted in plasmid pDM92, made by inserting the ΩKan^r cassette at the *Nco*I site located 271 bp upstream of the *htpY* transcriptional start site. The second construct, pDM335, was made by first removing the *Eco*O109 site in the polylinker and then inserting the ΩKan^r cassette at the *Eco*O109 site, located within the transcriptional unit of the *htpY*$^+$ gene, 117 bp downstream of the transcriptional start site (Fig. 1). The third construct, a complete deletion of the *htpY* gene, was achieved by removing 1,230 bp of the nucleotide sequence located between the two *Nco*I sites and inserting an ΩKan^r marker at this position, resulting in plasmid pDM462 (Fig. 1). Because the ΩKan^r element has transcriptional terminators at either end (10), the last two constructs, pDM335 and pM462, should not produce any HtpY protein, unless a minor or fortuitous promoter (in the case of pDM335) can be used to initiate transcription at some internal start site. Plasmid pDM462 does not possess any coding sequence of either the *htpY* open reading frame (ORF) or the divergent ORF (ORF-div). These ΩKan^r insertion plasmids were linearized by digestion with the appropriate restriction enzyme and transformed into a *recB recC sbcB15* mutant strain (JC7623), selecting for Kan^r transfor-

mants. This takes advantage of the fact that such strains tend to lose plasmids at a high frequency (26), thus selecting for homologous recombination events into the chromosome. The transformants were first tested for ampicillin sensitivity to eliminate Kan^r candidates carrying intact plasmids. The fact that the mutant alleles were successfully transferred onto the chromosome was verified by P1 transduction into other bacterial strain backgrounds, such as MC4100, MC1000, W3110, and the *ΔdnaK52* and *dnaJ259* mutant backgrounds. Southern DNA hybridizations were also used to verify that the mutant alleles had replaced the wild-type allele on the chromosome.

RNA isolation and Northern (RNA) blots. RNA isolation and Northern blotting were performed as follows. RNA was isolated by the hot phenol extraction procedure (19). If necessary, host DNA was eliminated from the RNA preparations by being treated for 15 min at 37°C with RNase-free DNase (RQ1; Stratagene) followed by ethanol precipitation. RNA samples (4 to 6 μg) were analyzed by the Northern blot technique as described by Colman (4) with the Hybond-N membrane (Amersham Corp.). ^{32}P-labeled, nick-translated *htpY* DNA probes were prepared as described by Maniatis et al. (19). RNA blots were stained with methylene blue to verify the equivalent amounts of RNA loaded.

Primer extension of RNA transcripts. A synthetic 24-mer oligonucleotide complementary to the *htpY* coding region from nucleotides (nt) 712 to 690 and a 27-mer oligonucleotide complementary to the ORF-div coding region from nt 1154 to 1180 were purified on a 20% polyacrylamide gel before use. Approximately 10 ng of ^{32}P-end-labeled primer and 10 μg of RNA (50 to 100 μg in the case of primer extension analysis of ORF-div), isolated from either wild-type bacterium MC4100 or from its isogenic derivatives, were coprecipitated and resuspended in 10 μl of 50 mM Tris-HCl (pH 8.3)–100 mM KCl. The mixtures were denatured at 95°C for 30 s and immediately annealed at 55°C for 20 min. The annealed primers were extended by avian myeloblastosis virus reverse transcriptase, essentially as described previously (19). The primer extension products were electrophoresed on the same gel as the dideoxy sequencing reactions by using the same primer.

Construction of *htpY-lacZ* fusions. To clone the promoter region of the *htpY* gene, the *Ava*I-*Ava*I fragment of plasmid pDM49 was first treated with T4 DNA polymerase to make the ends blunt and then was ligated into the *Sma*I site of the *lacZ* operon fusion vector pRS528 (31) with T4 DNA ligase; this resulted in plasmid pDM76. A smaller promoter-containing fragment, encoding only 655 bp of sequence from nt 35 to 690 upstream of the *htpY* coding region, was isolated by polymerase chain reaction (14). The polymerase chain reaction-amplified DNA was first cloned into pSK and verified by sequencing before being transferred to pRS528 (resulting in plasmid pDM251). The Lac$^+$ constructs pDM76 (p*htpY-lacZ*) and pDM251 (p*htpY-lacZ*) were first verified in strain DH5α and then were subsequently used to transform MC4100 to the Lac$^+$ phenotype. To construct single-copy (p*htpY-lacZ*) fusions in the chromosome, we used the bacteriophage λ vector λRS45 (31) to transfer the fusions by homologous recombination into the Δ*lac*$^-$*recA*$^+$ host MC4100. The λRS45 (p*htpY-lacZ*) lysogens were first picked up as blue colonies and then were checked by inducing the prophage and assaying for the presence of Lac$^+$ bacteriophage.

Assay of β-galactosidase. β-Galactosidase activities were measured by the method of Miller (20). The bacterial cultures were grown overnight in either M9 minimal or Luria-Bertani rich medium at 30°C, diluted 1:100, and allowed to reach an optical density of 0.2 at 600 nm. Aliquots of the cultures were subsequently incubated at 30°C or shifted to 42 or 50°C. Protein synthesis was inhibited by the addition of chloramphenicol (30 μg/ml), and 50- to 100-μl samples were added to Z buffer. Sodium dodecyl sulfate (SDS; 0.1%) and chloroform were used to disrupt the cell membrane. Samples were assayed in duplicate, and the data presented here are the average of six independent experiments.

Protein labeling. Cultures of strains carrying either plasmid pBluescript SK$^+$ or KS$^+$ containing the *htpY* gene cloned under the T7 RNA polymerase control were used. Expression of the T7 RNA polymerase was induced by addition of isopropyl-β-D-thiogalactopyranoside (0.5 mM). After a 30-min incubation, cells were treated with rifampin (200 μg/ml) for another 15 min and then labeled with [^{35}S]methionine (50 μCi/ml) for 10 min. The proteins were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) (12.5% acrylamide).

Sequencing of the *htpY* gene. To sequence the *htpY* gene, we used the T7 DNA polymerase Sequenase (U.S. Biochemical Corporation), following the manufacturer's procedures. Nested sets of exonuclease-derived *htpY* subclones on the high-copy vector pBluescript (Stratagene) were used as templates for sequencing reactions. Whenever necessary, synthetic oligomers were synthesized as primers to confirm the DNA sequence of *htpY*.

Construction of an inducible *rpoH*-containing plasmid. We placed the σ^{32} gene under the regulation of a tightly regulated promoter by cloning the 1.3-kb *Eco*RV fragment from pFN97 (25) into a vector carrying *lacI*^{q} (pTTQ18 and pTTQ19; Amersham). The vectors pTTQ18 and pTTQ19 were first modified by replacing the ampicillin resistance fragment with a kanamycin resistance cassette from pHP45Ω-Kan^r (10). The resulting plasmid, pSR1232, carrying the minimal σ^{32} gene lacking its own promoter, was tested first for induction of σ^{32} gene expression from the p*tac* promoter and then was used for subsequent analyses.

Construction of fusions to heat shock promoters. Transcriptional fusions to a number of heat shock promoters regulated by Eσ^{32} (i.e., those of *lon*, *htpG*, and *groE*) were constructed. Transcriptional fusions to the Eσ^E (σ^{24})-regulated promoters p*htrA* of the *htrA* gene and p3 of the *rpoH* gene were also constructed. All of these fusions were placed as a single copy on the chromosome by using the λRS45 vector (31), except for the *lon* and *htpG* fusions, which were isolated with λMu53 (2). The detailed construction of these fusions will be described elsewhere. However, it is important to point out that the p3*rpoH-lacZ* fusion does not contain any of the other promoters of the *rpoH* gene, but it does contain the two DnaA boxes which are located upstream of the *rpoH* p3 promoter (38).

Computer analysis of sequence data. The predicted nucleic acid and amino acid sequences of the *htpY* gene and the HtpY protein were compared by the FASTA algorithm of Pearson and Lipman (28) against release 22 of SWISS-PROT and release 72 of GenPept.

Nucleotide sequence accession number. The GenBank accession number of the *htpY* gene is L03720.

RESULTS

It is known that DnaK negatively regulates the expression of heat shock genes in *E. coli* (34). Thus, the overexpression of DnaK leads to repression of the heat shock response. While cloning the entire *dnaK dnaJ* operon along with its

flanking sequences, we observed that the presence of cloned multiple copies of DNA region upstream of *dnaK* resulted in elevated levels of the heat shock proteins. We localized the sequence encoding this function by subcloning and analyzing the effects of mutations of the cloned fragments. We found that this function can be ascribed to a gene located 700 bp upstream of the *dnaK* gene and have designated it *htpY* (high-temperature protein) because its transcription is induced at high temperatures.

Cloning of the *htpY* gene. To clone the putative *htpY* gene, we first cloned a 7.5-kb *Bam*HI fragment from bacteriophage λ101 of the Kohara *E. coli* DNA library into the low-copy vectors pWSK29 and pWSK30 (39), resulting in plasmids pDM38 and pDM39, respectively. This cloned fragment was expected to carry the entire *dnaK dnaJ* operon along with its flanking sequences. The authenticity of the pDM38 and pDM39 clones was verified by their ability to restore both bacteriophage λ growth as well as bacterial growth at 43°C in either *dnaK* or *dnaJ* mutant bacteria. To clone the DNA region upstream of the *dnaK dnaJ* operon, we cloned into the high-copy vector pSK a 3.6-kb *Bam*HI-*Eco*RI fragment (pDM42) or the smaller *Bam*HI-*Bgl*II (pDM49) and *Nco*I-*Bgl*II (pDM83) fragments (Fig. 1). These constructs were then used to identify the gene product(s) encoded by the DNAs in an inducible T7 promoter expression system (33).

The *htpY* gene encodes a 21-kDa polypeptide. Expression of either the 2.2-kb *Bam*HI-*Bgl*II fragment (pDM49) or the 950-nt *Nco*I-*Bgl*II fragment (pDM83) under the exclusive T7 promoter expression system revealed the synthesis of an approximately 21-kDa protein common to all of the plasmid constructs (Fig. 2A). From the known orientation of the cloned fragments in these constructs, it was concluded that this 21-kDa species is encoded by a transcript which is synthesized in the *Nco*I-to-*Bgl*II direction. Thus, transcription is in the clockwise direction relative to the *E. coli* map, in the same orientation as the *dnaK dnaJ* operon. To determine whether mRNA encoding the 21-kDa protein is indeed being expressed in vivo from an endogenous promoter, we expressed the clones pDM49 and pDM83 in both the maxicell system and a coupled transcription-translation *E. coli* S30 extract system. In all cases, a unique 21-kDa protein was synthesized (Fig. 2B), although much lower levels of the protein were observed under these conditions than with the T7 expression system. A final confirmation that this 21-kDa species is indeed encoded by the cloned DNAs was the inability of clones carrying an ΩKan^r cassette inserted at the *Eco*O109 site of pDM42 to produce the protein. Subsequent DNA sequence analysis verified the presence of an ORF corresponding to the 21-kDa polypeptide.

Sequencing of the *htpY* gene. Plasmid pDM42 or its derived subclones in pBluescript vectors served as templates for DNA sequence analysis. Sequencing of a 1.5-kb *Mlu*I-*Nco*I DNA fragment revealed the presence of two divergent overlapping ORFs. One ORF, corresponding to the *htpY* gene, spans 588 nt, starting at nt 603 and ending at nt 1190. It encodes a protein with a predicted molecular mass of 21,193 Da (Fig. 3). The second ORF, spanning 714 nt, starts at nt 1234 and ends at nt 521. We refer to this ORF as ORF-div, which encodes a putative protein with a predicted molecular mass of 26,648 Da.

The *htpY* ORF is characterized by a rare initiation codon, CTG, that is preceded by a weak Shine-Dalgarno ribosome-binding site sequence. The next potential initiation codon found in the ORF is an ATG at nt 697, which is associated with a better Shine-Dalgarno-like consensus sequence.

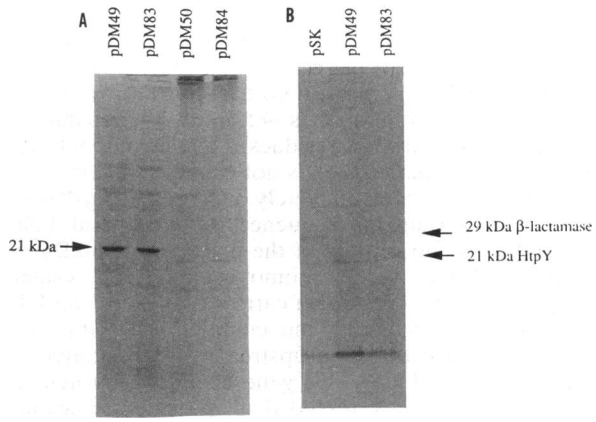

FIG. 2. Identification of the *htpY* gene product. (A) Cultures of BL21(DE3) carrying plasmid pDM49 or pDM50, which contain the same 2.2-kb *Bam*HI-*Bgl*II *htpY*$^+$ DNA fragment but in opposite orientations with respect to the T7 promoter of Bluescript vector, and pDM83 or pDM84, carrying 950-bp *Nco*I-*Bgl*II *htpY*$^+$ (also in opposite orientations with respect to the T7 promoter of Bluescript vector), were grown at 37°C in M9 minimal medium. T7 polymerase induction and protein labeling were done as described in Materials and Methods. (B) Coupled in vitro transcription-translation assay done with *E. coli* S30 extracts (Promega) to label the HtpY protein from its own promoter. Equivalent amounts of plasmid DNA pBluescript SK$^+$ plasmid, pDM49, or pDM83 were added to the same amounts of S30 extracts and incubated at 37°C. Proteins were labeled with [^{35}S]methionine (50 μCi/ml) for 30 min and resolved by SDS-PAGE (12.5% acrylamide). An autoradiogram of the dried gel is shown. The arrow indicates the position of the HtpY protein. Approximate molecular masses were estimated with low-molecular-mass markers from Bethesda Research Laboratories, which were loaded on the same gel.

However, utilization of this ATG as the initiation codon predicts the expression of a protein (161 amino acid residues [17,625 Da]) smaller than that observed from plasmids pDM49 and pDM83. It is most likely that the CTG codon at nt 603 represents the *htpY* initiation codon. However, cor-

```
      MluI
1     ACGCGTGCGCCTTTCAGCGTGCCGAAGAACATAAACAGCGTAAATACGCCCCACAGACCCAGGTAGACACCA
73    AGGAACTGTGCATTTGGCGCATCGGTCAGACCCAGTTTCGGCATCAGCAGAATCGCAACCAGCGTCAGCCAG
145   AAAGAACCGTAAGAGGTGAATGCGGTTAAACCGAAAGTGTTGCCTTTTTTGTACTCCAGCAGACCAGCAAAA
                                  NcoI
217   ATTTGCGCGATGCCGCCGTAGAAAATGCCCATGGCAAGAATAATACCGTCCAGAGCGAAATAACCCACGTTG
289   TGCAGGTTAAGCAGAATGGTGGTCATGCCGAAGCCCATCAGGCCCAGCGGTGCCGGATTAGCCAACTTAGTG
361   TTGCCCATAATTCCTCAAAAATCATCATCGAATGAATGGTGAAATAATTTCCCTGAATAACTGTAGTGTTTT
433   CAGGGCGCGGCATAATAATCAGCCAGTGGGGCAGTGTCTACGATCTTTTGAGGGGAAAATGAAAATTTTCCC
                p'phs
505   CGGTTTCCGGTATCAGACCTGAGTGGCGCTAACCATCCGGCGCAGGCAGGCGATTTGCAGTACGGCTGGAAT
                                        EcoO109
577   CGTCACGCGATAGGCGCTGCCGCTGACCGCTTTAACCCCATTTAGTGCCGCACCTACAGGGCCTCCCAGCCC
                                  M  T  A  L  T  P  F  S  A  A  P  T  G  P  P  S  P   17
649   CGCGCCGCGCAGCAAACCATGCCCAAGTACGCTCATTGCTGCGTGGGTGCGTAAAATGCGGGTCAGTTGGCT
        A  P  R  S  K  P  C  P  S  T  L  I  A  A  W  V  R  K  M  R  V  S  W  L   41
721   GGAAAGCAAATGCGACACACCTTTTGCCAATAATTTGTCTTTCATCAGCAGCGGCAGCAGCTCTTCCAGCTC
        E  S  K  C  D  T  P  F  A  N  N  L  S  F  I  S  S  G  S  S  S  S  S  S   65
793   ATTCACCCTGGCATCGACCGCGTGCAGAAACTCCTGCTTATGTTCCTCGTCCATTTTCTTCCAGGTATTACG
        F  T  L  A  S  T  A  C  R  N  S  C  L  C  S  S  S  I  F  F  Q  V  L  R   89
865   CAGAAATTGTTCCAGTAACTGTTGCTCAATTTCAAACGTAGACATCTCTTTGTCGGCTTTCAGCTTCAATCG
        R  N  C  S  S  N  C  C  S  I  S  N  V  D  I  S  L  S  A  F  S  F  N  R   113
937   CTTTGAAACATCGAGCAAAATGGCCCGATACAATTTACCGTGTCCGCGCAGTTTGTTGGCGATACTATCGCC
        F  E  T  S  S  K  M  A  R  Y  N  L  P  C  P  R  S  L  L  A  I  L  S  P   137
1009  ACCAAAATGCTGTAATTCTCCGGCAATCAGCTGCCAGTTGCGGCGATGTTGCTCGGGATGCCCTTCCATCGA
        P  K  C  C  N  S  P  A  I  S  C  Q  L  R  R  C  C  S  G  C  P  S  I  D   161
1081  TTTAAACAGTTCGTTGCGCATCAGTACGCTGGAGAGGCGAGTTTTGCCTTTTTCATTATGGGTGAGCAATCG
        L  N  S  S  L  R  I  S  T  L  E  R  R  V  L  P  F  S  L  W  V  S  N  R   185
                                                BglII
1153  GGCGAAATTTGCCAACTGTTCCTCACTACAATGCTGAAGAAAATCCAGATCTGAATCATTCAGGTAATTAAC
        A  K  F  A  N  C  S  S  L  Q  C                                          196
1225  ATTCATTTTTTGTGGCTTCTATATTCTGGCGTTAGTCGTCGCCGATAATTTTCAGCGTGGCCATATCCGATG
1297  AGTTCACCGTATGACCCGAAAAGGTGATTTTTGAGACGCAGCGTTTATTGTCGTTATCGCTGTTAATGTTGA
1369  TCCAGTCAGTGGTTTGCCCTTCTTTTATTTCTGAAGGAATATCAGGCTCTGACTGGCGCTACGGGCGGCTTT
                                  NcoI
1441  GAAATAAACCGATGCACCGCTTAACTGTAAATCGCCATGG
```

FIG. 3. Nucleotide sequence of the *htpY* gene and its flanking regions. The heat shock transcriptional start site is indicated by phs (located 81 nt upstream of the putative translational start site CTG). p′ is the transcriptional initiation site not regulated by Eσ^{32}.

roborative evidence from direct sequencing of the amino terminal residue of HtpY to exactly establish the initiation codon is needed.

The predicted HtpY protein is characterized by an unusually high abundance of serines (42 out of 196 residues) and cysteines (17 out of 196 residues). The functional significance of such a composition is not clear at present.

For ORF-div, which completely overlaps the *htpY* gene, a potential Shine-Dalgarno sequence is found at nt 1240 to 1245. However, no protein of the expected 26.6-kDa mass was detected under the T7 promoter expression system by using plasmid pDM150, which carries the 1.5-kb *Nco*I-*Mlu*I DNA fragment containing the complete ORF-div coding sequences along with 300 nt upstream of the putative ATG initiation codon. Hence, this gene encodes a protein very weakly expressed, if expressed at all. We cannot rule out the possibility of some posttranscriptional or posttranslational modifications which might be contributing to the inability to see the ORF-div gene product.

Phenotypic characterization of *htpY*::ΩKan^r derivatives. The three ΩKan^r insertions constructed in vitro, resulting in plasmids pDM92, pDM335, and pDM462, were first transferred to the chromosome in a single copy as described in Materials and Methods and resulted in strains DM136, DM407, and DM487, respectively. Subsequently, all three insertions were transduced into a number of genetic backgrounds. Out of these, DM136, constructed by transferring an ΩKan^r insertion, served as a control, because the ΩKan insertion lay outside the *htpY* transcriptional unit (see Fig. 1). The introduction of either of the two *htpY*::ΩKan^r alleles into backgrounds such as MC4100, MC1000, and W3110 did not result in major differences in growth and colony-forming ability at temperatures from 22 to 44°C compared with the isogenic wild type, except in the MC1000 background. In the MC1000 background, the *htpY*::ΩKan^r mutants did not form colonies at temperatures above 43.5°C. These *htpY* null strains were also tested for ability to support bacteriophage λ growth, and again, no major differences from their isogenic parents in terms of either the ability to form plaques or phage yields were observed.

The bacterial growth results were surprising, because Dean and James (6) reported that the *htgA* gene located upstream of the *dnaK dnaJ* operon was required for bacterial growth at high temperatures. They concluded that certain strains carrying mutant *dnaK* alleles, such as *dnaK756* and Δ*dnaK52*, simultaneously carry a mutation in this upstream gene. Hence, part of the temperature-sensitive phenotype of those mutant bacteria was ascribed to a mutated *htgA* gene. Because the *htpY* gene was also found to be located upstream of the *dnaK dnaJ* operon, we examined the possibility that *htpY*::ΩKan insertions may exhibit phenotype similar to that ascribed to a putative mutation(s) in the *htgA* gene. We constructed double mutants with *htpY*::ΩKan^r and either *dnaK756* or *dnaJ259*. We did not observe any additive effects on bacterial growth upon introducing the *htpY* null mutation into either a *dnaK756* or a *dnaJ259* mutant background. Both λ sensitivity and the temperature-sensitive phenotype of bacteria carrying these double mutations were indistinguishable from those carrying either the *dnaK756* or *dnaJ259* mutation alone.

***htpY*::ΩKan^r mutants exhibit reduced levels of heat shock proteins.** When an *htpY*::ΩKan^r allele was introduced into an otherwise wild-type background, a mild overall reduction (two- to threefold) in the levels of heat shock gene expression was observed (Fig. 4). Accordingly, the introduction of the *htpY*::ΩKan^r allele resulted in a decline in β-galactosidase expression from the *groE* heat shock promoter (Table 2). When the *htpY*::ΩKan^r allele was transferred into bacteria containing a transcriptional fusion to a housekeeping gene, such as p*htrL-lacZ* (Table 2), no differences in β-galactosidase expression between the mutant and the isogenic wild-type strains were observed. Thus, it appears that the *htpY* null mutation uniquely affects transcription from heat shock promoters. Because transcription from heat shock promoters is subjected to regulation by $E\sigma^{32}$, these results

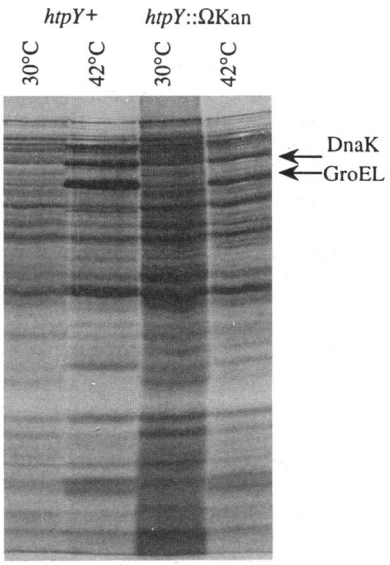

FIG. 4. Reduced level of heat shock proteins in *htpY* null mutants. Isogenic *htpY*$^+$ and *htpY*::ΩKan bacterial cultures were grown at 30°C in M9 minimal medium. Equivalent amounts of exponentially growing cultures were either directly labeled with [^{35}S]methionine (50 μCi/ml) at 30°C for 5 min or labeled at 42°C after a 5-min preincubation at 42°C. To approximate the reduction in the level of heat shock proteins, twice the amount of labeled extracts was loaded for samples made from *htpY* null mutant bacteria compared with those from *htpY*$^+$ bacterial extracts. An autoradiogram of the dried gel is shown.

TABLE 2. Effect of the *htpY* null mutation on the expression of various heat shock promoters fused to the *lacZ* gene

Promoter	β-Galactosidase activity (Miller U) at temp	
	30°C	42°C
$E\sigma^{32}$ dependent		
SR1208 = MC4100 Φ(p*groE-lacZ*)	888 ± 27	1,151 ± 34
DM495 = SR1208 *htpY*::ΩKan^r	298 ± 15	371 ± 22
$E\sigma^E$ (σ^{24}) dependent		
SR1458 = MC4100 Φ(p*htrA-lacZ*)	116 ± 6	148 ± 7
DM504 = SR1458 *htpY*::ΩKan^r	110 ± 6	126 ± 7
SR1720 = MC4100 Φ(p3*rpoH-lacZ*)	135 ± 7	140 ± 9
DM502 = SR1720 *htpY*::ΩKan^r	80 ± 4	90 ± 6
Control[a]		
SR1588 = MC4100 Φ(p*htrL-lacZ*)	229 ± 14	245 ± 15
DM496 = SR1588 *htpY*::ΩKan^r	234 ± 9	243 ± 11

[a] The non-heat shock promoter is equivalent to the $E\sigma^{70}$-dependent promoters.

TABLE 3. Effect of the overproduction of the HtpY protein on the expression of different heat shock promoters fused to the *lacZ* gene

Promoter	β-Galactosidase activity (Miller U) at temp	
	30°C	42°C
$E\sigma^{32}$ dependent		
SR1421 (= MC4100 *lon*::λp*lac*Mu53)/pSK	560 ± 22	858 ± 28
SR1421/pDM83	858 ± 29	1,112 ± 32
SR1418 (= MC4100 *htpG*::λp*lac*Mu53)/pSK	396 ± 17	607 ± 24
SR1418/pDM83	515 ± 21	897 ± 24
SR1208 [= MC4100 Φ(p*groE-lacZ*)]/pSK	717 ± 26	1,212 ± 31
SR1208/pDM83	1,154 ± 34	1,549 ± 47
$E\sigma^{E}$ (σ^{24}) dependent		
SR1458 [= MC4100 Φ(p*htrA-lacZ*)]/pSK	123 ± 6	187 ± 9
SR1458/pDM83	255 ± 14	360 ± 17
SR1720 [= MC4100 Φ(p3*rpoH-lacZ*)]/pSK	111 ± 7	120 ± 7
SR1720/pDM83	125 ± 6	150 ± 9
Control[a]		
SR1588 [= MC4100 Φ(p*htrL-lacZ*)]/pSK	230 ± 12	240 ± 14
SR1588/pDM83	239 ± 7	223 ± 9

[a] The non-heat shock promoter is equivalent to the σ^{70}-dependent promoters.

argue for a reduction or alteration in either the levels and/or the activity of σ^{32} in an *htpY* null mutant background.

We tested whether the transcription of the heat shock regulon transcribed by $E\sigma^{E}$ (σ^{24}) (7) was also affected by transducing the *htpY*::ΩKan^{r} mutation into strains carrying a *lacZ* fusion to the p3 promoter of *rpoH* or to the promoter of the *htrA* (*degP*) gene. It has been previously shown that the p3 promoter of *rpoH* is regulated by $E\sigma^{E}$ (σ^{24}) (7) and that *htrA* possesses a very similar promoter (7, 18). Our results show that the reduction in expression was less severe from $E\sigma^{E}$ (σ^{24})-regulated promoters than that observed from $E\sigma^{32}$-dependent promoter fusions (Table 2). In fact, the p3*rpoH-lacZ*-directed β-galactosidase synthesis was reduced by only 40 to 50%, while the expression from *phtrA* was more or less unaltered.

Overexpression of *htpY* induces the heat shock response. The *htpY* gene is weakly expressed, yet its presence on a multicopy plasmid led to certain interesting phenotypes such as the constitutive elevated synthesis of the major *E. coli* heat shock proteins.

To determine quantitatively whether the extent of heat shock gene induction was significant, we used three promoter fusion constructs transcribed by $E\sigma^{32}$, p*lon-lacZ*, p*groE-lacZ*, and p*htpG-lacZ*. These strains were transformed with the plasmid pDM83 (minimal *htpY*$^{+}$ clone) or with the vector alone (Table 3). In all cases, the levels of β-galactosidase induction were higher in strains carrying plasmid pDM83. We further tested the induction of other heat shock genes which are transcribed by $E\sigma^{E}$ (σ^{24}) by using *lacZ* promoter fusions p3*rpoH-lacZ* and p*htrA-lacZ*. Interestingly, significant induction was seen from the *htrA* promoter-driven *lacZ* expression; however, no induction was observed with the p3*rpoH-lacZ* construct. Hence, it

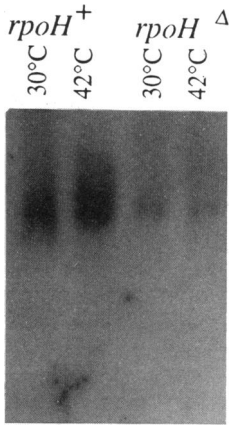

FIG. 5. Northern RNA analysis of *htpY* transcripts. RNA was extracted from isogenic *rpoH*$^{+}$ and Δ*rpoH* (KY1621) bacteria grown at 30°C or shifted to 42°C for 10 min before extraction. Approximately 4 to 6 μg of total RNA per lane was analyzed by the Northern blot technique and probed with 15 ng of ^{32}P-end-labeled oligonucleotide probe complementary to the fragment from nt 712 to 690, which is located 90 nt from the putative translational start site CTG.

appears that HtpY overproduction somehow leads to the induction of expression from $E\sigma^{32}$- and, to some extent, $E\sigma^{E}$ (σ^{24})-dependent promoters.

Transcriptional regulation of the *htpY* gene. As shown above, the *htpY* gene is located proximal to the *dnaK dnaJ* heat shock operon. It has been reported that in *Bacillus subtilis* (40), two genes, one homologous to the *E. coli* heat shock gene *grpE* and the other, ORF39, with no previous identity, are located immediately upstream of the *dnaK dnaJ* genes, all four genes constituting an operon. A similar arrangement exists in *Clostridium acetobutylicum* (23). Because of these observations, we examined the possibility of *htpY* transcription being subjected to heat shock regulation. Thus, we first determined the levels of *htpY*-specific mRNA at different temperatures by Northern analysis and further confirmed our results by the analysis of transcriptional fusions of the *htpY* promoter to the *lacZ* gene and by mapping the transcriptional start sites.

(i) Northern analysis of RNA. To test directly the effect of the temperature upshift on *htpY* transcription, we compared the levels of *htpY*-specific mRNA under heat shock and non-heat shock conditions. Figure 5 shows that an upshift in temperature from 30°C to 42°C resulted in significant induction of *htpY* transcription. We further examined whether this temperature-dependent induction was mediated by $E\sigma^{32}$ RNA polymerase holoenzyme, because it is known that induction of the classical heat shock regulon is exclusively under the transcriptional regulation of $E\sigma^{32}$ (24). Hence, RNAs isolated from isogenic *rpoH*$^{+}$ and Δ*rpoH* null mutant bacterial strains were probed for the presence of *htpY*-specific mRNA levels. Interestingly, the lower amounts of *htpY*-specific mRNA were observed in RNA prepared from Δ*rpoH* mutant bacteria at 30 or 42°C (Fig. 5). Furthermore, no induction of *htpY*-specific mRNA was observed in a strain carrying the Δ*rpoH* mutation at 42°C. These results argue that the *htpY* transcriptional regulation resembles that of the classical heat shock genes. However, it is quite clear that *htpY* transcription is not abolished in strains carrying the Δ*rpoH* mutation, and this gene may have another minor promoter which is transcribed in an $E\sigma^{32}$-independent manner. These results were further confirmed by measuring the

TABLE 4. β-Galactosidase activity of *lacZ* operon fusions to the *htpY* promoter[a]

Fusion	β-Galactosidase activity (Miller U) at temp		
	30°C	42°C	50°C
MC4100 (λRS74-p*lac*)	28 ± 2	24 ± 2	27 ± 2
SR1421 = MC4100 *lon*::λp*lac*Mu53	310 ± 12	627 ± 19	640 ± 21
DM249 = MC4100 Φ(p*htpY-lacZ*)	304 ± 9	381 ± 11	426 ± 14
DM249 Δ*rpoH* Φ(p*htpY-lacZ*)	102 ± 9	87 ± 6	73 ± 5
SR1588 = MC4100 Φ(p*htrL-lacZ*)	227 ± 14	243 ± 15	183 ± 12

[a] β-Galactosidase activity was measured after a 15-min incubation after a shift from 30°C to 42 or 50°C.

β-galactosidase expression from the p*htpY-lacZ* promoter fusion and by the primer extension experiments.

(ii) Analysis of p*htpY-lacZ* transcriptional fusions. Two *lacZ* transcriptional fusions containing the *htpY* promoter (DM78 and DM249) were used to quantitatively study the temperature dependence of *htpY* transcription described above. We observed an induction of β-galactosidase activity after a shift from either 30°C to 42 or 50°C (Table 4; data for DM249 are shown). Moreover, continued synthesis of β-galactosidase activity was observed at 50°C after a temperature shift from 30°C. As controls, we used a p*lon-lacZ* heat shock gene fusion (SR1421) and a fusion to a normal housekeeping gene, p*htrL-lacZ* (Table 4). The possible mediation of this induction by the *rpoH* gene product was verified by transducing a Δ*rpoH* Kan mutation into DM249; in such bacteria, the β-galactosidase synthesis was reduced at least four- to fivefold (Table 5). Alternatively, when the *rpoH* gene product (σ^{32}) under the regulation of an inducible promoter in pSR1232 was provided in *trans*, β-galactosidase synthesis was substantially enhanced after the addition of isopropyl-β-D-thiogalactopyranoside (Table 4). However, we observed that the level of induction of β-galactosidase synthesis in SR1208 carrying the p*groE-lacZ* fusion is higher than that observed with DM249 bearing the p*htpY-lacZ* fusion construct. This result may be due to differences in promoter strength. For example, the spacing between the −10 and −35 regions of the *htpY* gene promoter is unusually longer. Also, as shown by primer extension results, *htpY* transcription is initiated from two sites, and only one of them is σ^{32} dependent. However, it is quite clear that induction of the *htpY* promoter is significantly increased after a temperature upshift and that this induction is $E\sigma^{32}$ dependent, because it is severely reduced in a Δ*rpoH* mutant strain and induction of the *rpoH* gene product (σ^{32}) leads to a concomitant increase in β-galactosidase synthesis from the p*htpY-lacZ* fusions.

(iii) Mapping of the *htpY* promoters. A primer extension analysis was carried out with RNA isolated from wild-type bacteria under a range of growth temperatures. The results of this analysis showed that *htpY* mRNA has multiple 5′ ends located 80 to 82 nt upstream of the putative CTG initiation codon (Fig. 6A). At 30°C, the *htpY* transcription is initiated mostly from a distal site located at nt 82, while at 42 and 50°C, the transcription is initiated mostly from the site located at nt 80 and 81 (Fig. 6A). The −10 and −35 consensus boxes of the second start site at nt 81 present a striking homology to the promoter consensus sequences of $E\sigma^{32}$-transcribed genes (Fig. 7). The homology at the −10 region is 6 of 6 nt, whereas the −35 region shares 5 of 6 nt with the corresponding $E\sigma^{32}$ consensus sequence (Fig. 7, phs). To differentiate the site used by $E\sigma^{32}$ for transcription initiation, we analyzed RNA isolated from isogenic *rpoH*$^+$ and/or Δ*rpoH* strains. This analysis clearly showed that the transcription from nt 81 requires the functional *rpoH* gene product and thus represents the heat shock promoter designated phs (Ghs in Fig. 6). To the contrary, the transcription start from nt 82 (the more distal one, p′, i.e., A′ in Fig. 6B) was unaffected in the absence of *rpoH* gene product and thus does not represent the heat shock promoter.

Because the transcripts initiated from the proximal start site phs were more abundant after a temperature upshift, we further show that such transcripts are also constitutively induced in an *htrC* mutant background (Fig. 6A). Such a background has been previously shown to express elevated levels of heat shock proteins at permissive temperatures (29). In addition, the *htpY* gene transcripts accumulate at 50°C (note the increase in the intensity of lower bands of cDNA product made from RNA extracted from 50°C compared with that from 30°C samples; Fig. 6A). It is known that, at 50°C, the only genes which are actively transcribed are either $E\sigma^{32}$- or $E\sigma^{E}$ (σ^{24})-regulated genes. Taken together, these results provide good evidence that the major part of the *htpY* transcription is positively regulated by σ^{32}.

Mapping of ORF-div transcripts. As described above, a divergent ORF, ORF-div, completely overlaps the *htpY* gene. Our attempts to demonstrate expression of ORF-div by using the powerful T7 RNA polymerase expression system remained unsuccessful. Thus, it is quite likely that ORF-div may be weakly expressed in vivo, if expressed at all. To analyze the possibility of ORF-div gene expression, we examined total cellular RNA isolated from either wild-type bacteria or from isogenic bacterial strains carrying mutations in various global regulatory genes, such as *rpoS* (encoding KatF [16]), *rpoH*, or *lrp* (encoding the leucine response protein [36]) (Fig. 8). Primer extension analysis of such RNA preparations with a synthetic antisense oligonucleotide complementary to the sequence located between nt 1154 and 1180 revealed the presence of two weak transcriptional start sites located at nt 1280 and 1286 (Fig. 3 and 8) and few minor longer transcripts located at least 300 nt upstream of these sites. To detect these transcripts, we used a large excess of RNA (50 to 100 μg per reaction). Because the promoter usage was the same in all genetic backgrounds examined, it is assumed that expression of ORF-div is not controlled by any of the transcriptional factors mentioned above. In fact, the −10 and the −35 regions corresponding to these transcriptional start sites (located at nt 1280 and 1286) of ORF-div do resemble those of genes transcribed by the housekeeping form of RNA polymerase holoenzyme $E\sigma^{70}$. Thus, it is concluded that ORF-div may be transcribed in

TABLE 5. Expression of p*htpY-lacZ* with induction of the *rpoH* gene

Fusion	β-Galactosidase activity (Miller U) at time (min)		
	0	10[a]	20[a]
MC4100 (λRS74-p*lac*)/ pSR1232	8.8 ± 0.5	8.8 ± 0.5	9.1 ± 0.5
MC4100 Φ(p*htpY-lacZ*)/ pSR1232	272 ± 12	373 ± 17	406 ± 17
MC4100 Φ(p*groE-lacZ*)/ pSR1232	600 ± 23	1,276 ± 31	1,497 ± 34

[a] Isopropyl-β-D-thiogalactopyranoside (5 mM) added.

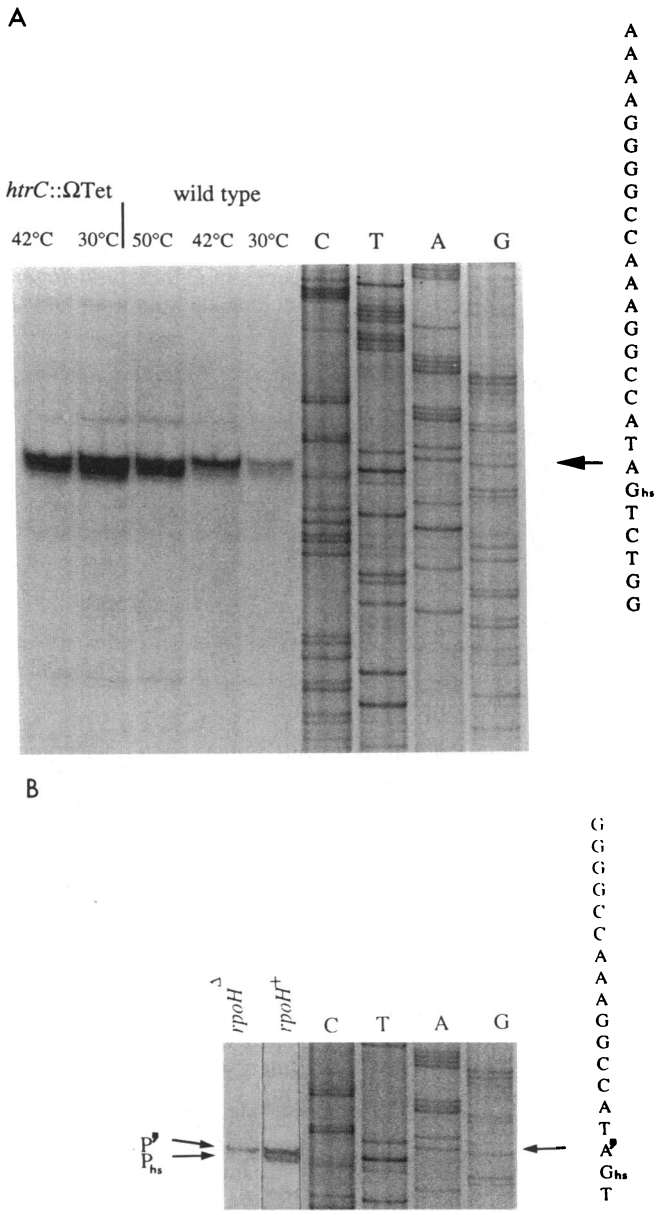

FIG. 6. Transcriptional regulation of the *htpY* gene. (A) Mapping of 5′ termini of *htpY* transcripts. Primer extension reactions of total cellular RNA hybridized to ^{32}P-end-labeled DNA oligonucleotide probe complementary to nt 712 to 690 of the *htpY* sense strand. RNA was extracted from *htpY*$^+$ bacteria grown at 30°C or shifted to either 42 or 50°C for 10 min. Similarly, RNA was extracted from *htrC* mutant bacteria, which constitutively overproduce heat shock proteins, and analyzed for *htpY*-specific transcripts. Lanes G, A, T, and C correspond to the dideoxy sequencing reactions carried out with the same oligonucleotide as the primer. (B) Mapping of 5′ termini of *htpY* transcripts with RNA extracted from *rpoH*$^+$ and Δ*rpoH* strains under the same conditions as panel A. Ghs, position of the putative Eσ^{32}-dependent transcriptional start site. A′ corresponds to the position of the Eσ^{32}-independent transcriptional start site.

vivo, albeit at very low levels. Because some longer transcripts were also observed, we do not rule out the possibility of a transcriptional coupling with an upstream ORF (41), which is transcribed in the same orientation as ORF-div.

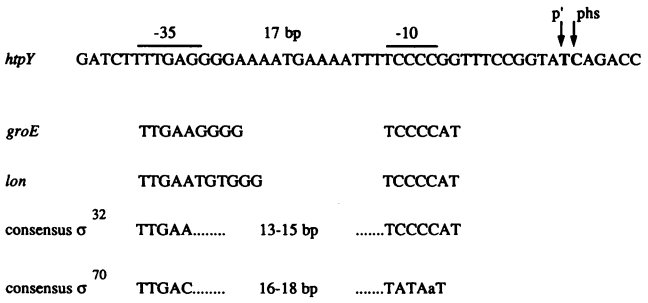

FIG. 7. Alignment of heat shock promoter consensus sequences and comparison with the *htpY* promoter sequence.

Homology with other genes and proteins. A comparison of the HtpY protein sequence with sequences in the GenPept and SWISS-PROT data bases (releases 72 and 22, respectively) showed no significant homology to any other known protein sequence. At the nucleotide sequence level, however, 24% homology between the *htpY* gene and ORF39 of *B. subtilis* was found.

During the preparation of this article, the *E. coli* genome sequence project (41) published the sequence of the 0- to 2.4-min region of the *E. coli* chromosome, which includes the sequences of *htpY* and ORF-div. We observed seven nucleotide differences with our sequence, specifically at nt 37, 509, 640, 650, 999, 1362, and 1458 (Fig. 3). An important change is the one at nt 509, because it lies in the −10 box of the *htpY* gene promoter region. This difference is clearly shown in Fig. 6A. Some of the other changes would disrupt the predicted protein coding capacity of both the *htpY* gene and ORF-div. For example, the difference at nt 999 would result in an *htpY* ORF only 116 amino acid residues long.

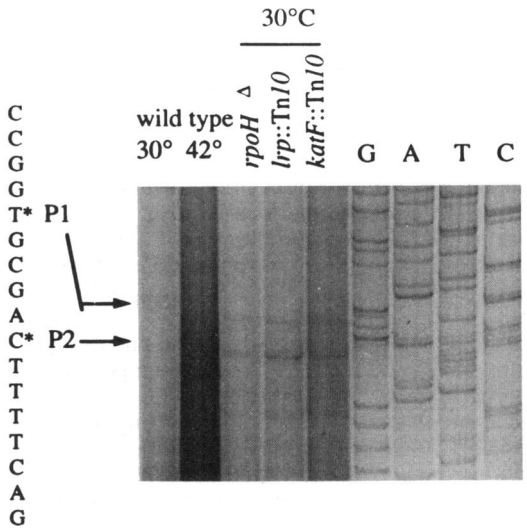

FIG. 8. Mapping of 5′ termini of ORF-div transcripts. Primer extension reactions were performed with total cellular RNA extracted from wild-type bacteria grown at 30°C or shifted to 42°C and from isogenic bacterial constructs carrying mutations in *rpoS*, *rpoH*, or *lrp* genes and grown at 30°C. The end-labeled primer was complementary to the fragment from nt 1154 to 1180, which is located 71 nt from the ATG translational start site of the ORF-div. Lanes G, A, T, and C correspond to the dideoxy sequencing reactions carried out with the same oligonucleotide as the primer.

This would be contrary to the results obtained for the HtpY protein with either the T7 promoter system or the in vitro S30 extract system, both of which showed the presence of an approximately 21-kDa protein.

DISCUSSION

While cloning regions upstream of the *dnaK dnaJ* operon, we identified a new heat shock gene, *htpY*, which, when present on a high-copy vector, was responsible for an elevated heat shock response in *E. coli*. Hence, this upstream region seemed to encode a function antagonistic to that of DnaK, which downregulates the heat shock response. To study this putative function, we cloned the minimal sequences carrying the function and replaced the wild-type *htpY* gene on the *E. coli* chromosome with null constructs. All of the phenotypes observed were assigned to a region located 700 bp upstream of the *dnaK dnaJ* operon, at the 0.2-min region of the *E. coli* chromosome. We have shown that the transcription of $E\sigma^{32}$-dependent heat shock promoters is increased when a multicopy *htpY*$^+$ plasmid is brought into an otherwise wild-type strain. Consistent with such results, *htpY* null mutants showed a significantly decreased level of $E\sigma^{32}$-regulated heat shock genes. It is possible that the HtpY protein somehow regulates either transcription of the *rpoH* gene or the activity of its gene product, σ^{32}. It is known that *rpoH* gene expression is regulated at the transcriptional (13) as well as posttranscriptional levels (22). In addition, the σ^{32} protein itself is extremely unstable (13). Thus, there are a variety of levels at which the HtpY protein could potentially participate in σ^{32} regulation. Perhaps the most interesting finding of this work is that HtpY plays a role opposite that of DnaK, DnaJ, and GrpE in terms of heat shock regulation. The positive role of HtpY may be important for *E. coli* physiology, antagonizing the negative roles of DnaK, DnaJ, and GrpE, thus fine-tuning the heat shock response.

We also tested whether the overexpression of HtpY protein had a similar effect on the other heat shock regulon, $E\sigma^E$ (σ^{24}). It was found that the expression of the *htrA* promoter was significantly increased, while the transcription from the p3 promoter of *rpoH* was marginally affected. Because the gene encoding the σ^E (σ^{24}) polypeptide has not been identified yet, the level of participation of the *htpY* gene product in $E\sigma^E$ (σ^{24}) transcription regulation remains unknown. It is unlikely that *htpY* itself encodes σ^E (σ^{24}) because one would expect that a σ^E (σ^{24}) null mutant would exhibit a temperature-sensitive phenotype. This prediction is based on the fact that $E\sigma^E$ (σ^{24}) is the only holoenzyme capable of transcribing the *htrA* gene (7) and *htrA* null mutants exhibit a temperature-sensitive phenotype (17). Our *htpY* null construct does not exhibit a temperature-sensitive phenotype in most *E. coli* strains tested, except in MC1000 in which the phenotype was exhibited only at temperatures above 43.5°C. Also, one would expect that a null mutation in the putative σ^E (σ^{24}) gene would abolish transcription from the p3 promoter of the *rpoH* gene and the *htrA* promoter, resulting in a temperature-sensitive phenotype. Because in the *htpY*::ΩKan bacteria only the transcription from the p3 promoter was affected without a decline in the transcription from the *htrA* promoter, direct involvement of HtpY in the transcription of $E\sigma^E$-transcribed promoters is ruled out. The differences observed in the relative effect of either HtpY overproduction or introduction of an *htpY*::ΩKan insertion into the p3*rpoH* versus *htrA* promoters are quite unexpected. One possible explanation could be the involvement of additional factors like DnaA protein in the case of the p3*rpoH* promoter. It has been shown previously that DnaA protein negatively regulates the expression of the *rpoH* gene, which is consistent with the presence of two DnaA boxes in close vicinity to the p3*rpoH* promoter region (38). The p3*rpoH-lacZ* fusion construct used in the present study indeed contains both DnaA boxes. Besides, the HtpY protein seems to only marginally affect this putative $E\sigma^E$ regulon.

There are recent reports describing a contiguous genetic organization of four heat shock genes in *C. acetobutylicum* (23) and *B. subtilis* (40), which include the *dnaK* and *dnaJ* genes. The fourth gene, identified in *B. subtilis*, is a new heat shock gene, ORF39, predicted to encode a 39-kDa protein and to be cotranscribed with the *dnaK dnaJ* operon (40). The function of this protein is unknown. Interestingly, the DNA sequence of *htpY* exhibits 24% homology to portions of the *B. subtilis* ORF39, although no homology at the protein level can be detected. The *htpY* gene may represent an analogous situation in *E. coli*, although it is not cotranscribed with the *dnaK dnaJ* heat shock operon.

The *htpY* gene is transcribed from two overlapping promoters. Interestingly, the transcription sites corresponding to these promoters are spaced only 1 nt apart. Our promoter mapping experiments show that the sequence of the −10 and −35 regions corresponding to the major promoter is similar to the consensus promoter sequence recognized by the heat shock-specific RNA polymerase holoenzyme $E\sigma^{32}$ (5). Accordingly, we found that *rpoH* null mutant bacteria accumulate reduced levels of *htpY* transcripts compared with *rpoH*$^+$ bacteria. *htpY* gene transcription is induced upon an upshift in temperature from 30°C to 42°C which continues at 50°C, similar to classical heat shock genes (for example, *dnaK* [8, 29]), and the overexpression of σ^{32} results in an increase in *htpY* gene expression. Finally, like other heat shock genes, transcription of the *htpY* gene is elevated in some mutant backgrounds, such as *htrC* (29), which show enhanced transcription of σ^{32}-regulated genes. All of these lines of evidence clearly demonstrate that the *htpY* gene is regulated in the same manner as other classical heat shock genes and is transcribed by $E\sigma^{32}$ RNA polymerase.

It was previously concluded (6) that certain *dnaK* mutant bacteria, e.g., those carrying the *dnaK756* or Δ*dnaK52* alleles (27), have an additional mutation located in the putative gene *htgA* mapping upstream of the *dnaK dnaJ* operon. This conclusion was based on the inability of minimal clones carrying only the *dnaK* gene to complement growth in *dnaK* mutant bacteria at temperatures above 43°C, although they did restore λ sensitivity to the mutant bacteria. Thus, this argued for the existence of an additional gene (*htgA*) upstream of *dnaK*, probably required for bacterial growth at high temperatures. However, in the study described by Dean and James (6), neither the transcriptional organization nor the product of the putative *htgA* gene was described. On the basis of the restriction patterns of their DNA clones, we conclude that the *htgA* and the *htpY* genes are identical. However, unlike Dean and James, we showed that the *htpY* gene is not essential for *E. coli* growth at 43°C in most bacterial strain backgrounds tested.

Another potentially important finding of our studies is the presence of the divergently transcribed ORF-div, which spans the entire length of the *htpY* gene. This genetic organization may possess some biological significance in terms of *htpY* gene regulation. We showed that this divergently transcribed ORF-div is weakly expressed at best and is not regulated by either the *rpoH*, *rpoS*, or *lrp* regulatory

gene products. However, the phenotypes associated with *htpY*::ΩKan^r, such as the negative effect on heat shock gene transcription, cannot be ascribed to the disruption of ORF-div, because they were fully reversed when a minimal *htpY*$^+$ clone, pDM83, which does not carry the amino-terminal coding portion of ORF-div, was provided in *trans*. Because all *htpY* null alleles used in this study also disrupt ORF-div, it was not possible to examine the phenotypes associated with mutations in ORF-div alone. Further experiments are needed to answer such questions.

ACKNOWLEDGMENTS

We thank D. Ang for carefully reading the manuscript.

D.M. is a recipient of an EMBO fellowship. This work was supported by grants from the Fond National Scientifique Recherche Suisse FN 31-31129-91 and by NIH grant AI21039.

REFERENCES

1. **Ang, D., K. Liberek, D. Skowyra, M. Zylicz, and C. Georgopoulos.** 1991. Biological role and regulation of the universally conserved heat shock proteins. J. Biol. Chem. **266:**24233–24236.
2. **Bremer, E., T. J. Silhavy, and G. M. Weinstock.** 1985. Transposable λ *plac* Mu bacteriophages containing *lacZ* operon fusions and kanamycin resistance insertions in *Escherichia coli*. J. Bacteriol. **162:**1092–1099.
3. **Casadaban, M. J.** 1976. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. J. Mol. Biol. **104:**541–555.
4. **Colman, A.** 1984. Expression of exogenous DNA in Xenopus oocytes, p. 46–49. *In* B. D. Hames and S. J. Higgins (ed.), Transcription and translation: a practical approach. IRL Press, Washington, D.C.
5. **Cowing, D. W., J. C. A. Bardwell, E. A. Craig, C. Woolford, R. W. Hendrix, and C. A. Gross.** 1985. Consensus sequence for *Escherichia coli* heat-shock promoters. Proc. Natl. Acad. Sci. USA **80:**2679–2683.
6. **Dean, D. O., and R. James.** 1991. Identification of a gene, closely linked to *dnaK*, which is required for high-temperature growth of *Escherichia coli*. J. Gen. Microbiol. **137:**1271–1277.
7. **Erickson, J. W., and C. A. Gross.** 1989. Identification of the σ^E subunit of *Escherichia coli* RNA polymerase: a second alternate σ factor involved in high-temperature gene expression. Genes Dev. **3:**1462–1471.
8. **Erickson, J. W., V. Vaughn, W. A. Walter, F. C. Neidhardt, and C. A. Gross.** 1987. Regulation of the promoters and transcripts of *rpoH*, the *Escherichia coli* heat shock regulatory gene. Genes Dev. **1:**419–432.
9. **Fayet, O., T. Ziegelhoffer, and C. Georgopoulos.** 1989. The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. J. Bacteriol. **171:**1379–1385.
10. **Fellay, R., J. Frey, and H. Krisch.** 1987. Interposon mutagenesis of soil and water bacteria: a family of DNA fragments designed for *in vitro* insertional mutagenesis of Gram-negative bacteria. Gene **52:**147–154.
11. **Georgopoulos, C.** 1992. The emergence of the chaperone machines. Trends Biochem. Sci. **17:**295–299.
12. **Gething, M.-J., and J. Sambrook.** 1992. Protein folding in the cell. Nature (London) **355:**33–45.
13. **Gross, C. A., D. B. Straus, J. W. Erickson, and T. Yura.** 1990. The function and regulation of heat shock proteins in Escherichia coli, p. 166–190. *In* R. Morimoto, A. Tissieres, and C. Georgopoulos (ed.), Stress proteins in biology and medicine. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **Innis, M. A., D. H. Gelfand, J. J. Sninsky, and T. J. White.** 1990. PCR protocols: a guide to methods and applications. Academic Press, Inc., New York.
15. **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.
16. **Lange, R., and R. Hengge-Aronis.** 1991. Growth phase-regulated expression of *bolA* and morphology of stationary-phase *Escherichia coli* cells are controlled by the novel sigma factor σ^S. J. Bacteriol. **173:**4474–4481.
17. **Lipinska, B., O. Fayet, L. Baird, and C. Georgopoulos.** 1989. Identification, characterization, and mapping of the *Escherichia coli htrA* gene, whose product is essential for bacterial growth only at elevated temperatures. J. Bacteriol. **171:**1574–1584.
18. **Lipinska, B., S. Sharma, and C. Georgopoulos.** 1988. Sequence analysis and regulation of the *htrA* gene of *Escherichia coli*: a σ^{32}-independent mechanism of heat-inducible transcription. Nucleic Acids Res. **16:**10053–10067.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Missiakas, D., C. Georgopoulos, and S. Raina.** Submitted for publication.
22. **Nagai, H., H. Yuzawa, and T. Yura.** 1991. Interplay of two cis-acting mRNA regions in translational control of σ^{32} synthesis during the heat shock response of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **88:**10515–10519.
23. **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:**3290–3298.
24. **Neidhardt, F. C., and R. A. VanBogelen.** 1987. Heat shock response, p. 1334–1345. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
25. **Neidhardt, F. C., R. A. VanBogelen, and E. T. Lau.** 1983. Molecular cloning and expression of a gene that controls the high-temperature regulon of *Escherichia coli*. J. Bacteriol. **153:**597–603.
26. **Oden, K. L., L. C. De Veaux, C. R. T. Vibat, J. E. Cronan, Jr., and R. B. Gennis.** 1990. Genomic replacement in *Escherichia coli* K-12 using covalently closed circular plasmid DNA. Gene **96:**29–36.
27. **Paek, K.-H., and G. C. Walker.** 1987. *Escherichia coli dnaK* null mutants are inviable at high temperature. J. Bacteriol. **169:**283–290.
28. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
29. **Raina, S., and C. Georgopoulos.** 1990. A new *Escherichia coli* heat shock gene, *htrC*, whose product is essential for viability only at high temperatures. J. Bacteriol. **172:**3417–3426.
30. **Raina, S., and C. Georgopoulos.** 1991. The *htrM* gene, whose product is essential for *Escherichia coli* viability only at elevated temperatures, is identical to the *rfaD* gene. Nucleic Acids Res. **14:**3811–3819.
31. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy lac-based cloning vectors for protein and operon fusions. Gene **53:**85–96.
32. **Squires, C., and C. L. Squires.** 1992. The Clp proteins: proteolysis regulators or molecular chaperones? J. Bacteriol. **174:**1081–1085.
33. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
34. **Tilly, K., N. McKittrick, M. Zylicz, and C. Georgopoulos.** 1983. The *dnaK* protein modulates the heat-shock response of *Escherichia coli*. Cell **34:**641–646.
35. **VanBogelen, R. A., V. Vaughn, and F. C. Neidhardt.** 1983. Gene for heat-inducible lysyl-tRNA synthetase (*lysU*) maps near *cadA* in *Escherichia coli*. J. Bacteriol. **153:**1066–1068.
36. **Van der Woude, M. W., B. A. Braaten, and D. A. Low.** 1992. Evidence for global regulatory control of pilus expression in *Escherichia coli* by lrp and DNA methylation: model building based on analysis of *pap*. Mol. Microbiol. **6:**2429–2435.

37. **Wang, Q., and J. M. Kaguni.** 1989. A novel sigma factor is involved in expression of the *rpoH* gene of *Escherichia coli*. J. Bacteriol. **171**:4248–4253.
38. **Wang, Q., and J. M. Kaguni.** 1989. dnaA protein regulates transcription of the *rpoH* gene of *Escherichia coli*. J. Biol. Chem. **264**:7338–7344.
39. **Wang, R. F., and S. R. Kushner.** 1991. Construction of versatile low-copy number vectors for cloning, sequencing and gene expression in *Escherichia coli*. Gene **100**:195–199.
40. **Wetzstein, M., U. Volker, J. Dedio, S. Lobau, 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.
41. **Yura, T., H. Mori, H. Nagai, T. Nagata, A. Ishihama, N. Fujita, K. Isono, K. Mizobuchi, and A. Nakata.** 1992. Systematic sequencing of the *Escherichia coli* genome: analysis of the 0–2.4 min region. Nucleic Acids Res. **20**:3305–3308.
42. **Zhou, Y.-N., N. Kusukawa, J. W. Erickson, C. A. Gross, and T. Yura.** 1988. Isolation and characterization of *Escherichia coli* mutants that lack the heat shock sigma factor (σ^{32}). J. Bacteriol. **170**:3640–3649.