Expression of the *Caulobacter* Heat Shock Gene *dnaK* Is Developmentally Controlled during Growth at Normal Temperatures

SUELY LOPES GOMES,¹ JAMES W. GOBER,² AND LUCILLE SHAPIRO^{2*}

Departamento de Bioquimica, Instituto de Quimica, Universidade de Sao Paulo, Sao Paulo, S.P., 01498 Brazil,¹ and Department of Developmental Biology, Stanford University School of Medicine, Stanford, California 94305-5427²

Received 11 December 1989/Accepted 1 March 1990

Caulobacter crescentus has a single dnaK gene that is highly homologous to the hsp70 family of heat shock genes. Analysis of the cloned and sequenced dnaK gene has shown that the deduced amino acid sequence could encode a protein of 67.6 kilodaltons that is 68% identical to the DnaK protein of Escherichia coli and 49% identical to the Drosophila and human hsp70 protein family. A partial open reading frame 165 base pairs 3' to the end of *dnak* encodes a peptide of 190 amino acids that is 59% identical to DnaJ of E. coli. Northern blot analysis revealed a single 4.0-kilobase mRNA homologous to the cloned fragment. Since the dnaK coding region is 1.89 kilobases, dnaK and dnaJ may be transcribed as a polycistronic message. S1 mapping and primer extension experiments showed that transcription initiated at two sites 5' to the dnaK coding sequence. A single start site of transcription was identified during heat shock at 42°C, and the predicted promoter sequence conformed to the consensus heat shock promoters of E. coli. At normal growth temperature (30°C), a different start site was identified 3' to the heat shock start site that conformed to the E. coli sigma 70 promoter consensus sequence. S1 protection assays and analysis of expression of the dnaK gene fused to the lux transcription reporter gene showed that expression of *dnaK* is temporally controlled under normal physiological conditions and that transcription occurs just before the initiation of DNA replication. Thus, in both human cells (I. K. L. Milarski and R. I. Morimoto, Proc. Natl. Acad. Sci. USA 83:9517-9521, 1986) and in a simple bacterium, the transcription of a hsp70 gene is temporally controlled as a function of the cell cycle under normal growth conditions.

The heat shock response is an apparently universal phenomenon by which a cell subjected to a sudden increase in temperature induces the synthesis of a small number of highly conserved proteins, the heat shock proteins (hsps) (20, 21, 26). The high degree of conservation of the hsps suggests that these proteins are involved in important processes within the cell.

The best studied among these proteins is the family of polypeptides of M_r 70,000. Eucaryotic organisms typically contain several hsp70-related genes (17, 25, 34). Regulation of these genes is complex; some hsp70 genes are expressed only after cells are stressed, whereas others are expressed constitutively or are developmentally regulated. In contrast, Escherichia coli encodes only one hsp70 protein, the product of the dnaK gene (4). DnaK is an abundant protein even at normal temperatures (reported to compose 1.4% of the cellular protein), and its synthesis is induced by heat and other forms of stress (15). It is required for the initiation of lambda phage DNA replication (19), is essential for cellular growth, and appears to be required for the initiation of DNA replication at high temperatures (28, 31). We have previously shown that the DnaK homolog from Caulobacter crescentus is under complex regulation. In addition to increased synthesis after brief exposure to heat (13, 29) or peroxide (29), under physiologic growth conditions DnaK appears to be synthesized at a defined time in the cell cycle, just before the initiation of DNA replication (13). This response is similar to that observed with the human hsp70 protein, which has been shown in HeLa cells to rapidly increase synthesis during entry into the S phase and to decline in the late-S and G₂ phases (23). Furthermore, the human hsp70 protein is localized to the nucleus in the S phase. At each cell division in C. crescentus, only the progeny stalked cell is able to initiate DNA replication (8), and DnaK synthesized in the predivisional cell is selectively localized to this cell upon division (29).

To determine whether a single *dnaK* gene is responsible for both the stress response and the response to temporal control signals, we cloned *dnaK*; it occurred in a single copy on the genome, and we identified both a conserved heat shock promoter and a conserved $E\sigma^{70}$ promoter in tandem 5' to the start of the *dnaK* coding sequence. Analysis of *dnaK* transcription in vivo showed that the gene is only transcribed from the heat shock promoter at high temperatures but is transcribed from both promoters in a temporal manner at normal growth temperatures.

MATERIALS AND METHODS

Bacterial strains and plasmids. The synchronizable widetype strain *C. crescentus* CB15N, used in all experiments, was grown in PYE medium or in modified minimal M2 glucose medium (6). Cultures were synchronized by the method of Evinger and Agabian (12). *E. coli* HB101 was grown in LB medium (24). Plasmid B8, containing the *Drosophila* hsp70 gene cloned into pBR322, was obtained from E. Craig. Plasmid pUCD615, containing the promoterless *luxCDABE* cassette, was obtained from C. I. Kado (16). Plasmid pRK290.20R, a derivative of the broad-host-range plasmid pRK290 (9), contains a multiple restriction site cassette inserted in the *Eco*RI site.

Cloning of the *dnaK* gene. C. crescentus genomic DNA was digested with the restriction enzyme BamHI, transferred to nitrocellulose, and hybridized at low stringency ($6 \times$ SSC [1 \times SSC is 0.15 M NaCl plus 0.015 M sodium citrate]- $5 \times$

^{*} Corresponding author.



FIG. 1. Map position of the *C. crescentus dnaK* gene and a restriction map of a 3.5-kb *Bam*HI fragment containing *dnaK* and a portion of *dnaJ* (pGR24). A subclone, pJG42, containing 968 bp of the *dnaK* 5' region and 370 bp of *dnaK* coding sequence fused to a promoterless *lux* operon is shown below the restriction map. An 1,100-bp *Bam*HI-XhoI fragment used for S1 nuclease protection assays is also shown. Restriction enzymes: B, *Bam*HI; C, *Cla*I; P, *Pst*I; R, *Rsa*I; S, *SaI*I; X, *XhoI*.

Denhardt solution–0.1% sodium dodecyl sulfate at 55°C; washes in $5 \times SSC$ at 37°C) to the *Drosophila* hsp70 probe in the B8 plasmid (18). A single band of hybridization in the region corresponding to 3.5 kilobases (kb) was observed. *C. crescentus* genomic DNA was then digested with *Bam*HI and size fractionated by agarose gel electrophoresis. The DNA fragments from 2 to 4 kb in size were then electroeluted and ligated to *Bam*HI-digested pBR322 to create a partial genomic library. After transformation into *E. coli*, the resulting partial library was screened by probing Southern blots of digests of miniplasmid preparations or by colony hybridization with a nick-translated *Drosophila* hsp70 probe in the B8 plasmid. Both methods yielded the same 3.5-kb *Bam*HI fragment.

DNA sequencing. DNA sequence analysis of the entire 3.5-kb cloned fragment was done by the procedure of Sanger et al. (32). Various restriction fragments were subcloned into M13mp18 and M13mp19, and their complete sequences were determined on both strands after deletion subclones were prepared by using the cyclone system from International Biotechnologies. A restriction map of the clone is shown in Fig. 1, and the sequence is shown in Fig. 2.

Nuclease S1 and primer extension mapping. The probe for S1 mapping experiments was a 1.1-kb BamHI-XhoI fragment isolated from pGR24 (Fig. 1). S1 nuclease assays were performed as described by Amemiya et al. (1) with the following modifications. Hybridizations were carried out at 90°C for 10 min, followed by transfer to 56°C for 3 h, with 50 µg of C. crescentus total RNA. Nuclease S1 digestion was performed with 100 U per sample at 37°C for 30 min. Primer extension mapping (3) was done with a synthetic oligonucleotide (24-mer) that was complementary to nucleotides 433 to 456 of the dnaK coding sequence (Fig. 2). The oligonucleotide was obtained from Operon Technologies, Alameda, Calif. The oligonucleotide was 5' end labeled with γ -³²P and hybridized to 50 µg of C. crescentus total RNA. Annealing was carried out in 25 µl of 0.1 M piperazine-N,N'-bis(2ethanesulfonic acid) (pH 6.4)-1.0 M NaCl-5 mM EDTA at 45°C for 12 h. The samples were precipitated with ethanol and suspended in 49 µl of 50 mM Tris hydrochloride (pH 8.5)–2 mM dithiothreitol–5 mM MgCl₂–40 mM KCl–0.2 mM deoxynucleoside triphosphates and RNasin (1 U; Boeringer Mannheim Biochemicals, Indianapolis, Ind.). The oligonucleotide was extended at 42°C with 20 U of avian myeloblastosis virus reverse transcriptase (Life Sciences, St. Petersburg, Fla.) for 90 min. The unlabeled oligomer served as a primer for the sequencing ladder, and sequencing was performed by the dideoxynucleotide-chain termination method (32) with [α -³⁵S]ATP.

Construction of *dnaK-lux* **transcription fusion.** A *dnaK-lux* transcription fusion was made by inserting the 1.35-kb *BamHI-PstI* fragment from pGR24, containing the entire upstream region plus 370 base pairs (bp) of the coding region of the *dnaK* gene in front of the promoterless *lux* genes in pUCD615 (16) (Fig. 1). The *dnaK-lux* transcription fusion was then subcloned into pRK290.20R, and the resulting recombinant plasmid, pJG42, was mated into *C. crescentus*.

Biochemical techniques. Standard methods were used for DNA end labeling, nick translation, and Northern blots (22). Cellular RNA was purified from C. crescentus cells as described by Amemiya et al. (2). Luciferase activity was assayed in C. crescentus cultures containing plasmid-borne lux genes. To assay the activity of luciferase at defined time points in the cell cycle, synchronized cultures were incubated at 32°C, and then samples were shifted to 25°C for 10 min at the times indicated in Fig. 6B. Because luciferase is irreversibly denatured at temperatures greater than 30°C, this method provides a means of assaying the new synthesis of luciferase over a short period of time. After incubation at 25°C for 10 min, each sample of the culture (optical density at 660 nm, 0.3 to 1.0) was placed in a scintillation vial, and n-decylaldehyde (2% emulsion) was added to a final concentration of 0.002%. Emitted light was measured immediately over a period of 6 s in a Packard scintillation counter (model 3330). Counts obtained on the ³H channel were converted into quanta emitted by using a known radioactive standard for calibration (33).

RESULTS

Isolation and sequence analysis of the *dnak* gene from *C*. *crescentus*. A single segment of DNA in the *C*. *crescentus* genome was identified by hybridization of the *Drosophila* hsp70 probe B8 (18) to genomic DNA blots. DNA isolated from the region of the gel that hybridized to the probe was used to construct a *Bam*HI plasmid library in pBR322, as described in Materials and Methods. Two recombinant plasmids that hybridized to the *Drosophila* probe were isolated and shown to have identical restriction maps. The restriction map of the insert in pGR24 is shown in Fig. 1.

The existence of a physical map of the C. crescentus genome generated by the ordering of DraI restriction fragments with respect to the genomic map (10) allowed us to determine the general map position of the dnaK homolog. We used the cloned fragment as a probe of Southern blots of DraI-digested chromosomal DNA separated by pulsed-field gel electrophoresis (10). The cloned DNA hybridized to a single DraI restriction fragment of 305 kb (Fig. 1). Its relative position was determined by integrating in both orientations a plasmid that contained the dnaK gene. This homologous integration introduced another DraI site, and comparison of the restriction fragments generated by insertion in both orientations allowed the placement of dnaK near the junction with the adjacent 105-kb fragment (Fig. 1).

To confirm the identity of the cloned fragment as the dnaK gene, the fragment was cleaved with various restriction

	Ala Gly Val Leu Gln Gly Asp Val Lys Asp Val Leu Leu Leu Asp Val Thr Pro Leu Thr GCC GGG GTG CTG CAA GGC GAC GTC ANG GAC GTG CTG CTG CTG GAC GTG ACC CTG ACC 1560
90 120 150	Leu Gly Ile Asp Thr Leu Gly Gly Val Phe Thr Pro Leu Ile Glu Arg Asn Thr Thr Ile
GGCCCGGAATCCAGATTCCCCCGTCGTGGGGGGGGGTTATGGCCTTGCCGGGGGGCCCCTATCCCCGATATCCCGGCTTCG 180 210	CTG GGC ATC GAC ACC CTG GGC GGT GTG TTC ACC CCG CTG ATC GAA CGC AAC ACC ACC ATC 1620 1650
ANGCOGGGTTCAACGGGCTCGACAAACGCACTACGCCCCCCGTTTCGATTTCGCCGGAGCGCGCGC	Fro Thr Lys Arg Ser Gin Thr Phe Ser Thr Ale Asp Asp Asm Gin Ser Ale Vel Thr Ile CCG ACC AAG CGC TCG CAG ACC TTC TCC ACC GCC GAC GAC AAC CAG TCG GCC GTG ACG ATC 1680
CTTGANCOGTAGGAGCGAAGCGAACGGAGCGAAGAAAAACCGGGGGGCGCTTCTGAGAAGGCCCTTGTCACCG 330 360 390	Arg Val Phe Gin Giy Giu Arg Ala Met Ala Gin Asp Asn Lys Met Gin Giy Gin Phe Asp CGG GTG TTC CAG GGC GAA CGC GGG ATG GGG CAG GAC AAC AAG ATG CAG GGC CAG TTC GAG
REAL Net Ser Lys Ile Ile Gly Ile Asp Leu Gly Thr Thr CCTCCCCCTAAAGGAGCGACCGACGAGGAGGAGGAGGAGGAGGAGGAGG	1740 1770
	Leu Val Gly Ile Pro Pro Ala Pro Arg Gly Val Pro Gin Ile Glu Val Thr Phe Asp Ile CTG GTC GGC ATC CCC CCG GGC CCG CGG GTG CCG CAG ATC GAG GTC ACG TTC GAC ATC 1800
Ash ser cys val ala lie Met Asp Gly Lys Thr Pro Lys val lie Glu Ash Ala Glu Gly AAC TOS TGC GTC GCC ATG ANG GGC ANG ACC COS ANG GTG ATC GAG AAC GCC GAG GGC 480 510	Asp Ale Asm Gly Ile Vel Gin Vel His Ale Lys Asp Lys Ale Thr Asm Lys Glu Gin Ser
XhoI Ala Arg Thr Thr Pro Ser Val Val Ala Phe Leu Glu Asp Gly Glu Arg Leu Ile Gly Gln	GAC GCC AAC GGC ATT GTC CAG GTC CAC GCC AAG GAC AAG GCG ACC AAC AAG GAG CAG TCG 1860 1890
GUT GAC ACC COS TOS GTC GTC GCC TTC <u>CTC GAG</u> GAC GGT GAG CGC CTG ATC GGC CAG 540 570	Ile Arg Ile Gin Ala Asn Giy Giy Leu Ser Asp Ser Asp Ile Giu Arg Net Val Lys Giu Arc CGC Arc CAG GCC AAT GGC GGC CTT TCG GAC AGC GAC ATC GAG CCC ATG GTC AAG GAA
Pro Ala Lys Arg Gin Ala Val Thr Asn Pro Thr Asn Thr Leu Phe Ala Ile Lys Arg Leu CCG GCC AMG CGC CMG GCC GTC ACC AAC CCG ACC AMC ACG CTG TTC GCG ATC AMG CGC CTG	1920 Ala Glu Ala Asn Lys Ala Ala Asn Glu Lys Lys Lys Gln Leu Val Glu Ala Lys Asn Gln
530 Ile Gly Ary Thr Ala Ser Asp Pro Val Val Glu Lys Asp Lys Gly Met Val Pro Tyr Arg	GCC GNG GCC ANC NÃG GCT GCG GAC GAN ANG ANG ANG CAG CTG GTC GAG GCC ANG ANC CAG 1980 2010
ATT GGC CGC ACC GCC AGC GAC CCG GTG GTC GAG AAG GAC AAG GGC ATG GTG CCC TAC AGA 660 690	Gly Glu Ala Ile Leu His Ser Thr Glu Lys Ala Leu Ala Glu His Gly Asp Lys Val Gly
Ser Ser Ary Ala Ary Ala Gly Asp Ala Trp Val Lys Ala His Gly Lys Asp Tyr Ser Pro TCG TCA AGG GCC CGA GCC GGA GAC GCC TGG GTC AAG GCC CAC GGC AAG GAC TAC AGC CCG	2040 2070
720 750 750 Pati	Ala Ala Glu Lys Thr Ala Ile Glu Thr Gly Ile Thr Glu Leu Lys Thr Ala Leu Glu Gly GCG GCT GAG AAG ACC GCG ATC GAA ACC GGG ATC ACC GAG CTG AAG ACC GCC CTG GAA GGC 2100 2130
CH GIN VAL SET ALE PHE ILE LAU GIN DYE HET DYE GIN ALE ALE GIN ALE HIS LAU GIY CAG GAA GTC TOG GCC TTC ATT CTG CAG ANG ATG ANG GAA GCC GCC GAC GCC CAC CTT GGC 780 810	Glu Asp Val Glu Ala Ile Gln Ala Lys Thr Gln Ala Leu Ile Gln Ala Ser Net Lys Leu
Glu Pro Val Thr Lys Ala Val Ile Thr Val Pro Ala Tyr Phe Asn Asp Ala Gln Arg Gln	GAA GAC GTC GAG GCC ATC CAG GCC AAG ACC CAG GCC CTT ATC CAG GCT TCG ATG AAG CTG 2150 2190 Part
ENG COG GTG ACC ANG GOG GTC ATC ACG GTT COG GCC TAC TTT ANT GAC GCT CAG GGT CAG 840 870	Gly Glu Ala Met Tyr Ala Ala Gln Gln Gly Ser Ala Glu Gly Gly Asp Ala Lys Ala Asp GGC GAG GCC AT <u>G TAC</u> GCC GCC CAG CAA GGC TCG GCC GAG GGC GAC GCC AAG GCT GAC
Ala Thr Lys Asp Ala Gly Lys Ile Ala Gly Leu Glu Val Leu Arg Ile Ile Asn Glu Pro- GCC ACC AMG GAC GCC GGC AMG ATC GCC GGC CTG GAA GTC CTG CGC ATC ATC AAC GAG CCG	2220 2250 Sali Sali
	Ann Giv Val Val Ann Ala Glu Phe Glu Glu Val Ann Ann Ann Lvs Pro Ala Ala End
900 930 Thr Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asp Asp Gly Lys Lys Ile Ala Val Tyr	Asp Gly Val <u>Val Amp</u> Ala Glu Phe Glu Glu <u>Val Amp</u> Asp Asm Lys Pro Ala Ala End GAC GGC GTG <u>GTC GAC</u> GGC GAG TTC GAG GAA <u>GTC GAC</u> GAC AAC AAG CGC GGC GCT TAA GGG 2210 2310
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr Acg GCC GCT GCG CTG GCC TAC GGT CTC GAC ANG AAC GAC GGC ANG ANG ATC GCC GTC TAT 960 990	Asp Gly Val <u>Val Am</u> Ala Glu Phe Glu Glu <u>Val Am</u> Asp Asm Lys Pro Ala Ala End GAC GGC GTC <u>GTC GAC</u> GCC GAG TTC GAG GAA <u>GTC GAC</u> CAC AAG CAG CGG GGC GTA AGG 2280 2310 AGTCATGCGCGCCCCGGAACATTTCGCCACAGGATCGTTCCGGCGCCGCGGCACCTTCTGCCTCCCGGCCCATGAGGGCC 2310 230 230 240
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr Acg GCC GCT GCG CTC GCC TAC GGC ANG AAC GAC GAC GAG AAG ATC GCC GTC TAT 960 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu Fac CTG GGC GGC AAC TATC GAG CTG TGC ATC CTG GAA ATC GGC GGAT GGC GTC TTC GAG	Asp Gly Val Val Am Ala Glu Fhe Glu Glu Yal Am Asp Ash Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAG GAC GTC GAC GAC GAC GAC GCG GGG GTC GTC GAC 2310 AGTCATGGGGGCGCGGGAACATTTGGCCACAGGATGGTCGGGGGGGG
900 930 Thr Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr AcG GCC GCT GCG CTC GCC TAC GAC AAG AAC GAC GAC GAC AAG ATC GCC GTC TAT 900 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu GAC CTG GGC GGC AAC ATC GAC GTG TGC ATC CTG GAA ATC GGC GAT GGC GTC TC GAG 1020 1050 1050 1050 1050 1050 1050 1050	Asp Gly Val Val Amp Ala Glu Phe Glu Glu Yal Amp Asp Asm Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAG TTC GAG GAA GAC AAC AAR GCC GCG GCT TAA GCC 2210 2310 2310 AGTCATGCGCGCCCCCGGGAACATTTCGCCACAGGATGGTCGGCGCGGACCTTCTGCCCCCCCGGGACCATTCGCCGCGCCCCAGAGGCGC 2340 2370 2400 TTGCCTGGGGCCTTGGAGCTATCCCCCATTCCCCCGTGTAAAAGGCCACTTTCGGGGCCTTTGGGGCCCCAGAGGCGACC 2430 2450 2450 Clai Het Arg Asp Tyr Tyr Glu Lie Leg Gly Val Thr Arg Thr Lie Asp Glu Ala Gly Leg
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr Acc GCC GCT GCC TAC GCT CTC GAC ANG AAC GAC GCC AAG ANG ATC GCC GTC TAT 960 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu GAC CTG GGT GGC GGC ATC GTC GAC ATC CTG GAA ATC GGC GAT GGC GTC TTC GAG 1020 Sali Val Lys Sar Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Ile Val GTG AA <u>G TYC AA</u> C GGC GAC ACC TTC CTG GGC GGC GAG GAC TTC GAC GTC GTC 1060	Asp Gly Val Yal Amp Ala Glu Phe Glu Glu Yal Amp Asp Asm Lys Pro Ala Ala End GAC GGC GTC GTC GAC GCC GGC GCC GCC GAC GAC GAC GAC GA
900 930 Thr Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr AcG GCC GCT GCG CTG GCC TAC GGC AGG AAG AAC GAC GGC AGG AAG ATC GCC GTC TAT 960 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu Acc CTG GGC GGC AGC ACC TTC GAC GTG TCG ATC CTG GAA ATC GGC GAC GGC GGC GGC GGC 1020 5all Val Lys Sar Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Ile Val 1080 1080 1110 5all Tyr Leu Ala Asp Glu Phe Lys Lys Glu Gln Gly Yal Asp Leu Arg Lys Asp Lys Leu 6GC TAC CTG GCC GAC GAG TTC AMG AMG GAG GAG GCTC GCC AAG GAC TTC 6GC GAC GGC GGC AGC ANG CTC	Asp Gly Val Yal Amp Ala Glu Phe Glu Glu Yal Amp Asp Asm Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAM GAC GAC GAC GAC GAC GAC GCG GCT TA GCG 2310 AGTCATGGGGCGCCGGGAACATTTCGCCACAGGATCGTTCGGGGGGGG
500 510 510 510 510 510 510 510	Asp Gly Val Yal Amp Ala Glu Phe Glu Glu Yal Amp Asp Asm Lys Pro Ala Ala End GAC GGC GTC GTC GAC GCC GAG GTC GAG GAC GTC GAC GAC AAC AAG CGC GCG GCT TAA GCC 2310 AGTCATGCGCGCCCCCGGAACATTTCGCCACAGGATCGTTCGGCGCCCGCGCCCCCCCC
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr Acc GC GCT GCG CTG GCC TAC GGT CTC GAC ANG AAC GAC GCA GGA ANG ATC GCC GTC TAT 960 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu GAC CTG GGT GGC GCC ACC TTC GAC GTG TCG ATC CTG GAA ATC GGC GAC GGC GCC GCC 1020 Sall Val Lys Sar Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Ile Val 1050 Sall 1060 Sall 1010 Sall 1010 Sall 1020 Sall	Asp Gly Val Yal Amp Ala Glu Phe Glu Glu Yal Amp Asp Asm Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAG GAC GAC GAC GAC GAC GAC GGC GCG GCT TAA GCG 22300 AGTCATGCGCCCCCGGAACAATTTCGCCACAGGATCGTTCGGCGCCCGCAACACTTCGGCCCCCGGCACAGGCCCACAGGCCCACAGGCCCACACGCCCACAGGCCCACACGCCCCACGCCCCCC
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr AcG GCC GCT GCG CTC GCC TAC GGT CTC GAC ANG AAC GAC GGC AGG ANG AAC GCC GTC TAT 900 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu Acc CTG GGC GGC ACC ACC TTC GAC GTG TCC AG AATC GGC GAT GGC GTC CTC GAG 1020 Sall Val Lys Sar Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Ile Val 1050 1050 Sall 1060 1010 10	Asp Gly Val Yal Am Ala Glu Phe Glu Glu Yal Am Asp Asm Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAG GAC GAC GAC GAC GAC GAC GA
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr Acc GCC GCT GCC CTC GCC TAC GGT CTC GAC ANG AAC GAC GCC ANG AAG ATC GCC GTC TAT 900 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu acc TCG GGT GGC GCC ACC ACC TCC GGC GTG TCC GAC ATC GGC GAC GTC GTC TCC GAG 1020 Val Lys Sar Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Ile Val 1050	Asp Giy Val Yal Amp Ala Glu Phe Glu Glu Yal Amp Asp Asm Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAG GTC GAG GAA GTC GAC GAC GCC GCG GCT TA A GCG 2310 AGTCATGGGGCGCGCGGAAACHTTGGGCAACAGATGGTCGGGGCGGCGGCACCACTTGGGCGCTGGGGCCCACAGAGGGCC 2400 TTGCCTGGGGCTTGGAGCTATCCCACTTGCGCGTGTAAAGGCCACTTGGGGGCCCCACAGAGGCGG 2400 TTGCCTGGGGCTTGGAGCTATCCCACTTGCGGCGCTGTGGAAGGCCGGCC
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr AcG GCC GCT GCG CTC GCC TAC GGT CTC GAC ANG AAC GAC GCC AGG ANG AAC GCC GTC TAT 960 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu GAC CTG GGT GGC GGC ACC ATC GAC TTC GAC GTC CTG GAA ATC GGC GAC GAT GCC GTC TCC GAC 1020	Asp Gly Val Yal Am Ala Glu Phe Glu Glu Yal Am Asp Asm Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAG GTC GAG GAA GTC GAC GAC GAC GGC GCC TA GAC GAC GCC GCC GCC AT GAC GAC GTC GAC GAC GCC GCC GCC AT GAC GAC GTC GAC GAC GCC GCC GCC AT GAC GAC GAC GCC GAC GCC GCC AT GAC GAC GAC GCC GCC GAC GGC CCC GCC GCC
900 910 910 910 910 910 910 910	Asp Giy Val Yal Amp Ala Glu Phe Glu Glu Yal Amp Asp Asp Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAG GCC GAC GAC GAC GAC GAC GCC GAG GCT TAA GCC 2310 ASTCATEGGCCCCCCGGAAACHTTCGCCACAGGACGTTCGGCGCGCGCGCCCCCCCCCC
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Ile Ala Val Tyr AcG GCC GCT GCG CTC GCC TAC GCC AAG ANG AAC GAC GCC AAG AAG ATC GCC GTC TAT 900 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu acc CTG GGT GGC GCC ACC GTC GTG TCG ATC CTG GAA ATC GGC GAT GGC GTC TTC GAG 1020 Val Lys Sat Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Tile Val 1050 Val Lys Sat Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Tile Val 1050 Val Lys Sat Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Tile Val 1050 Val Lys Sat Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Tile Val 1050 Sali Asp Tyr Leu Ala Asp Glu Phe Lys Lys Glu Gln Gly Yal Asp Leu Arg Lys Asp Lys Leu 1110 Sali Asp Tyr Leu Ala Asp Glu Phe Lys Lys Glu Gln Gly Yal Asp Leu Arg Lys Asp Lys Leu 1170 11	Asp Gly Val Yal Amp Ala Glu Phe Glu Glu Yal Amp Amp Amp Lys Pro Ala Ala End GAC GGC GTG GTC GAC GCC GAG GTC GAG GAA GTC GAC GAC AAC AAG CAG CCC GCG GCT TAA GCG 2310 AGTCATGCGCCCCCGGAACATTTCGCCACAGGATCGTTCGGCGCCCGCAACCACTTCGGCCCCCGGCCACAGGCCCACACGATCGCCCCCCACAGGCCCCACAGGCCCCACAGGCCCCACACGCCCCACAGGCCCCACACGCCCCACAGGCCCCACAGGCCCCACAGGCCCCACACGCCCCACAGGCCCCACAGGCCCCACAGGCCCCACAGGCCCCACAGGCCCCACACGCCCCACAGGCCCCCC
900 930 Thr Als Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys Lys Lis Lis Ala Val Tyr AcG GCC GCT GCC GTC GAC AAG AAC GAC GGC AGG AAG AAC ATC TAT 990 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu GAC CTG GGC GGC GGC ACC ACC TTC GAC GTG TCG AATC CTG GAA ATC GCC GAT GGC GTC TTC GAG 1020 Sall Val Lys Sar Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Tle Val GTG AAG ACC TTC CTG GGC GGC GGC GGC AGC ACC CTC CTG GGC GGC GGC GGC GAC TTC CTG GGC GGC GGC GAC TTC CTG GGC GGC GGC GAC TTC CTG GGC GCC CAC GTC GTC 1110 Map Tyr Leu Ala Asp Glu Phe Lys Lys Glu Gln Gly Yal Asp Leu Arg Lys Asp Lys Lau GAC TTC GTC GGC GGC GGC GAC ATC GTC GGC AGC GTC CAG GGC CAC CTC GGC GGC GGC GAC TTC GGC CAG GGC CAG GCC GGC GAC AGC GTC GGC GGC CAG GCC CAG CAG	Asp Gly Val Yal Am Ala Glu Phe Glu Glu Yal Am Asp Asm Lys Pro Ala Ala End GAC GGC GTG GTC GAC GAC GAC GAG GAA GTC GAC GAC ACA AM COM 2310 AGTCATGGGGCGCGGGAACATTTCGCCACAGGATCGTTCGGCGCGCGGACCCTCTGGCCCCCGGCACAGAGCCCACAGGCCCACAGGCCCACAGGCCCACAGGCCCCACAGGCCCCCC
900 930 Thr Ala Ala Ala Leu Ala Tyr Gly Leu Asp Lys Asn Asp Gly Lys Lys I ala Val Tyr AcG GCC GCT GCG CTC GCC TAC GGC AGC AAG AAG AC GCC AGG AAG AC TC GCC TTC TAC 900 Asp Leu Gly Gly Gly Thr Phe Asp Val Ser Ile Leu Glu Ile Gly Asp Gly Val Phe Glu Cac CrG GGT GGC GCC ACC GTC GCG CTG CTC GAC AAS ACC GCC GAT GCC GTC TTC GAG 1020 1020 Val Lys Sar Thr Asm Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Arg Tile Val 1050 105	Asp Gly Val Yal Am Ala Glu Phe Glu Glu Yal Am Asp Asm Lys Pro Ala Ala End GAC GGC GTG GTC AM GCC GAG GCC AM GAA GTC GAC GAC AC AM GCC GCG GCT TAA GCG 2310 AGTCATGGGGCGCGGGAACATTTCGCCACAGGATCGTTCGGCGCGCGGAACCTTCGGCGCCCAGGCGCACAGGCCAGACGTCGG 2310 TTGCCTGGGGCCGGGAACATTTCGCCACTGCCGTTGTAAAGGCGCCACAGCCTCGGCCCAGAGGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGACGCCAGAGCCCAGACGCCGC
900 930 Thr Als Als Als Lew Als Tyr Gly Lew Asp Lys Asm Asp Gly Lys Lys Iys Ile Als Val Tyr Acc GCC GCC GCC GCC GCC GCC GCC GCC GCC G	Asp Giy Val Yal Amp Ala Glu Phe Glu Glu Yal Amp Amp Amp Lys Pro Ala Ala End GAC GGC GTG GTC GAC GAC CAG GTC GAG GAC GTC GAC GAC CAC ACA MA CAG 2310 ASTCATEGGCCCCCCGGAAACHTTCGCCACTGCGACGCGCGCGCGCGCCCCCACTTCGCCCCCCGGCCCCACAGCCCCACAGGCCCCACAGCCCCACAGCCCCACAGCCCCACAGCCCCACAGCCCCACAGCCCCACAGCCCCACAGCCCCACAGCCCCACAGCCCCACAGCCCCCC

FIG. 2. DNA sequence of the *C. crescentus dnaK* gene and a portion of the adjacent *dnaJ* gene. All of the sequence was determined on both strands and across all restriction sites of the insert in pGR24 shown in Fig. 1. The sequence complementary to the synthetic oligonucleotide (24-mer) is underlined. The inverted repeats in the DNA sequence are indicated by horizontal arrows.

enzymes and subcloned into M13mp18 and M13mp19, and the DNA sequence was determined by the dideoxy-chain termination method (Fig. 2). DNA sequence analysis revealed an open reading frame coding for a protein that was highly homologous to *E. coli* DnaK (Fig. 3). A comparison of the predicted amino acid sequence revealed 68% identity with the *E. coli* DnaK protein (4) and 49% identity with the *Drosophila* hsp70 protein (18). The *dnaK* gene could encode a protein of 67.6 kilodaltons. A second open reading frame that did not end in the clone was detected downstream of the dnaK gene (Fig. 2). The TAA stop codon for the dnaK gene was separated from the ATG start codon of the open reading frame by 165 bp. The predicted sequence of the possible 190 amino acids encoded by the open reading frame had 59% identity to the DnaJ protein (5, 27) of *E. coli* (Fig. 3).

Effect of heat shock on the in vivo transcription of *dnaK*. The in vivo transcript of the *dnaK* gene was detected by Northern blot analysis with total RNA prepared from C.

A Dask

<u>C.crescentus</u> <u>E. coli</u> <u>Drosophila</u>	M * *	S G P	K * -	I * A	I * *	G * *	I * *	D * *	L * *	G # #	T * *	T *	N + Y	s * *	C * *	V * *	λ + G	I * V	М * Y	D * Q	G * H	K T G	T * K	P * V	K R E	V * I	I L N	Е * λ	N * Y
<u>C.crescentus</u> <u>E.coli</u> <u>Drosophila</u>	A * D	E * Q	G * *	A D N	R * *	T * *	T * *	P * *	S * *	V I Y	V I *	A *	F Y *	L T T	E Q -	D * *	G * S	E * *	R T *	L * *	I V N	G * *	Q + E	P * *	λ * *	K *	R * N	Q * *	A * V
<u>C.crescentus</u> <u>E.coli</u> <u>Drosophila</u>	V * A	Т * М	N * *	P * *	T Q R	N * *	T * *	L * V	F * *	А * D	Ι * λ	K * *	R * *	L * *	I * *	G * *	R * *	T R K	A F Y	S Q D	D * *	P E *	V E K	V * I	E Q A	K R E	D * *	K V M	G S K
<u>C.crescentus</u> <u>E.coli</u> <u>Drosophila</u>	M I H	V M W	P *	Y F F	R K K	S I V	- - v	S I *	R A D	λ * G	R D G	X N K	G * P	D * K	A * I	W * G	V *	K E E	A V Y	H K K	G * *	- - E	- - s	K Q #	D K R	Y M F	S A A	P * *	Q P E
<u>C.crescentus</u> <u>E.coli</u> Drosophila	E Q #	V I I	s * *	λ * S	p E M	I V V	L * *	Q K T	K * *	M * *	K * *	EK*	A T T	λ * *	E * *	A D +	H Y Y	L * *	G * *	E * *	P * S	v * I	Т * *	K E D	λ * *	V * *	I * *	T *	V * *
<u>C.crescentus</u> <u>E.coli</u> Drosophila	P * *	λ * *	¥ *	F *	N * *	D + +	λ * S	Q * *	R * *	Q * *	λ * *	T * *	K * *	D * *	λ * *	G * *	K R H	I * *	λ * *	G * *	L * *	e * N	V * *	L K *	R * *	I * *	I * *	N * *	E * *
<u>C.crescentus</u> <u>E.coli</u> <u>Drosophila</u>	P * *	T * *	λ * *	λ * *	λ * *	L * *	λ * *	¥ *	G * *	L * *	D * *	K * *	n G +	D T L	- K	G # #	K N E	K R R	- T N	I * V	λ * L	V * I	¥ * F	D * *	L * *	G * *	G * *	G * *	т * *
<u>C.crescentus</u> <u>E.coli</u> <u>Drosophila</u>	F * *	D * *	V 1 *	s * *	I * *	L I *	E * T	I * *	D D	- E E	G V *	D * -	G * -	V E -	K S	- T L	F * *	E * *	V * *	K L R	S A *	T * *	N * A	G * *	D * *	T * *	F H H	L * *	G * *
<u>C.crescentus</u> <u>E.coli</u> Drosophila	G * *	E * *	D # #	F *	D * *	l S N	R * *	I L L	V I *	D N T	¥ * H	L * *	λ ∨ *	D E E	E * *	F + +	K * *	K + R	E D K	Q # Y	G * K	V I K	D * *	L * *	R * *	k N S	D * N	K P P	L * R
<u>C.crescentus</u> <u>E.coli</u> Drosophila	λ * *	L M *	Q # R	R * *	L * *	RK t	E * T	λ * *	- A A	E * *	K * R	λ * *	K * *	K I R	E * T	L * *	S * *	S * *	T A S	A Q T	Q * E	Y T A	e d t	V * I	N * E	L * -	P * -	F Y -	I * *
<u>C.crescentus</u> <u>E.coli</u> <u>Drosophila</u>	S T D	M A A	N D L	λ * F	s t e	G * *	P * Q	L K D	H + F	L M Y	N * T	I * -	K * *	L V V	S T *	R * *	λ * *	K * R	L * F	E * *	λ S E	L * *	V * C	D E A	D * N	L * *	I V F	A N R	R * N
<u>C.crescentus</u> <u>E.coli</u> <u>Drosophila</u>	T S *	I + L	GEQ	P * *	C L V	EK +	Q V K	λ * *	L * *	K Q N	D * *	λ + +	G * K	L * M	K S D	K ∀ *	S * G	D * Q	I * *	D * H	e D D	V * I	I * V	L * *	V * *	G * *	G * *	M Q S	S T T
<u>C.crescentus</u> <u>E.coli</u> Drosophila	R * *	M * I	P * *	K M +	V * *	Q # #	Q K S	λ K L	V * L	Q A *	D E E	F * *	P + +	- - H	G * *	R K K	e * N	P * L	H R N	K * L	G D S	V * I	N * *	P * *	D * *	E * *	V A A	V * *	λ * *
<u>C.crescentus</u> <u>E. coli</u> Drosophila	L I Y	G * *	λ * *	λ * *	V * *	Q * *	λ G *	G * A	V * I	L * *	Q T S	G * *	D * *	- Q	- - s	- G	- K	V # I	K * Q	D * *	V * *	L * *	L * *	L * *	D * *	V * *	Т + Д	P * *	և * *
<u>C.crescentus</u> <u>E. coli</u> Drosophila	T S S	L * *	G * *	I * *	D E E	т *	L M A	G * *	G * *	V * *	F H H	T * *	P T K	L * *	I * *	Ε λ *	R K *	N *	T + C	T * R	I * *	P * *	Т * С	K * *	R H Q	S * T	Q + K	T V *	F * *
<u>C.crescentus</u> <u>E. coli</u> Drosophila	s * *	T * *	А * Y	D E S	D * *	N * *	Q # #	s * P	λ * G	V * *	T * S	I * *	R H Q	V * *	F L Y	Q * E	G * *	E * *	R * *	λ K *	M R *	А * Т	Q A K	D * *	N * *	K * N	M S A	Q L L	G * *
<u>C.crescentus</u> <u>E. coli</u> <u>Drosophila</u>	Q * T	F *	D N *	L * *	V D S	G * *	I * *	P N *	P * *	λ * *	P * *	R * *	G * *	V M *	P * *	Q * *	I * *	E * *	V * *	T * *	F * *	D * *	I * L	D * *	λ * *	N D *	G * *	I * *	V L L
<u>C.crescentus</u> <u>E. coli</u> Drosophila	Q H N	V * *	H S S	λ * *	K * *	D + E	K * M	a N S	T S *	N G G	K * *	Е * Х	Q * K	s K N	I * *	R T T	I * *	Q K K	А * N	n S D	G S K	G * *	- R	L * *	S N *	D E Q	S D A	D E E	I * *
<u>C.crescentus</u> <u>E. coli</u> <u>Drosophila</u>	E Q D	R K *	M * *	V * *	K R N	E D *	λ * *	E * *	λ * K	N * Y	K X X	A E D	λ * Ε	D * *	ER+	K * *	k F H	K E R	QE *	L * R	V * I	e Q T	A T S	K R R	N * *	Q * A	G * L	Ε λ *	A H S
<u>C.crescentus</u> <u>E. coli</u> Drosophila	I L Y	L * V	H * F	s * N	T * V	E R K	K * Q	A Q S	L V V	a E E	e * Q	H A A	- - P	G * A	D * G	K * *	V L L	G P D	λ * Ε	À D ★	e D D	K * *	T * N	λ * S	I * V	E * L	T S D	G A K	I L C
<u>C.crescentus</u> <u>E. coli</u> <u>Drosophila</u>	t * N	E A +	L * T	K E I	T * R	λ * ₩	L * *	ë K D	G * S	e + N	D * T	V K T	E A A	λ * Ε	I * K	QEE	Ē	F	- D	А * Н	K * *	T M M	Q * E	A E E	L * *	1 A -	Q # -	A V -	S * T
<u>C.crescentus</u> <u>E. coli</u> <u>Drosophila</u>	M Q R	K * H	L * C	g M S	e * P	ī	А * М	M Q T	Q Q X	A Q M	- H H	λ \$ Q	Q * *	Q + G	G Q A	S T G	λ * *	Е - Л	G * *	G A *	- P	- - G	D + A	A * N	K S C	- A G	- N Q	N Q	λ • •
<u>C.crescentus</u> <u>E. coli</u> <u>Drosophila</u>	- K G	D * G	D * -	G D -	V * F	V * G	D * G	λ * Υ	e * S	- - G	- - P	- - T	F * V	E * *	E * *	V * *	D K	D * -	N K	K * -	P - -	х - -	х -	en * *	d				



FIG. 3. (A) Comparison of the predicted amino acid sequence of C. crescentus DnaK with E. coli DnaK (4) and Drosophila hsp70 (18). Those amino acids identical in both sequences are indicated with an asterisk. Gaps in the alignment are marked by hyphens within the sequences. (B) Comparison of the predicted amino acid sequence of the open reading frame adjacent to the dnaK gene with the amino acid sequence of E. coli DnaJ (5, 27).

crescentus either grown at 30°C or heat shocked at 42°C for the times shown in Fig. 4A. The RNA was electrophoretically separated on formaldehyde agarose gels, transferred onto nylon filters, and hybridized to a ³²P-labeled 3.5-kb *Bam*HI fragment from pGR24. The 3.5-kb fragment hybridized to a 4.0-kb mRNA that was present at low levels in cells growing at normal temperatures (30°C) and was transiently induced approximately 10- to 20-fold during heat shock at 42°C, reaching a peak at 15 min. The size of the *dnaK* gene alone is 1.89 kb. The large size of the transcript and its increased expression upon heat shock, together with the DNA sequence data, suggest that the *dnaK* and *dnaJ* genes of *C. crescentus* could be cotranscribed in a manner similar to that found in *E. coli* (30, 35).

The effect of heat shock on the expression of the dnaK gene was also determined by S1 nuclease protection assays. A probe was constructed by end labeling the BamHI-XhoI fragment shown in Fig. 1. This fragment was hybridized to RNA isolated from cells grown at normal temperatures (30°C) or from cells heat shocked at the times indicated in



FIG. 4. Effect of heat shock on in vivo transcription of the *dnaK* region. (A) Northern blot analysis of *C. crescentus* RNA with the entire 3.5-kb insert from pGR24 as a probe. Total RNA (5 μ g), isolated from cultures grown at 30°C (lane 1) and after 5, 10, 15, and 20 min of exposure to 42°C (lanes 2 through 5, respectively), was subjected to Northern blot analysis as described in Materials and Methods. The major transcript at 4 kb is indicated. (B) Nuclease S1 protection assays of *dnaK* transcripts. The RNA (50 μ g) was from cells growing at 30°C (lane 1) and from cultures shifted to 42°C for 10, 15, and 20 min (lanes 2 through 4, respectively). The 5' end-labeled DNA probe was the 1.1-kb *Bam*HI-XhoI fragment shown in Fig. 1. The heat shock-protected RNA fragment of 320 bases and the 280-base protected fragment from cells grown at 30°C are indicated.

Fig. 4B. Nuclease S1-resistant fragments were separated on urea-polyacrylamide gels. RNA isolated from cells grown at normal temperatures yielded a 280-nucleotide protected fragment (Fig. 4B), suggesting that the 5' terminus could lie approximately 150 nucleotides upstream of the *dnaK* translation start. A faint fragment of 320 nucleotides was also detected in cells grown at 30°C. Upon heat shock, however, this 320-nucleotide protected fragment transiently increased in amount with time of exposure to 42° C and peaked at 15 min. The heat shock transcript mapped to a region 190 nucleotides upstream of the *dnaK* translation start site.

Identification of the dnaK transcription initiation sites by primer extension. A 24-nucleotide synthetic oligomer complementary to nucleotides 433 to 456 of the dnaK coding sequence shown in Fig. 2 was end labeled with ³²P and hybridized to RNA isolated either from heat-shocked cells or from cells grown at normal temperature. The hybrids were extended by reverse transcriptase, and the products were resolved in urea-polyacrylamide gels run next to a sequence (ladder). A single extension fragment was obtained with heat-shocked RNA (Fig. 5C, lane 1). The migration of the fragments indicated an initiation site at either the cytidine or guanosine residue at position 236 or 237, respectively (Fig. 5B). Two extension fragments were observed with template RNA from cells grown at normal temperatures (Fig. 5C, lane 2). The less abundant fragment comigrated with the transcript obtained with heat-shocked RNA. The predominant extension fragment corresponded to the smaller transcript detected by S1 protection assays (Fig. 4B). The initiation site of this smaller fragment was at the cytidine or guanosine residue at position 279 or 280, respectively (Fig. 5B).

Sequences 5' to the two transcription start sites were compared with *E. coli* consensus promoter sequences located at positions -10 and -35 (Table 1). Both the sequences at positions -10 and -35 5' to the heat shock transcript (P1) were a close match to the tandem *E. coli* dnaK P1 and P2 heat shock promoters. However, the sequences 5' to the start of the transcription of the *C.* crescentus P2 promoter conform, at the -35 region, to the consensus $E\sigma^{70}$ promoter, and the -10 region more closely resembles the $E\sigma^{70} - 10$ region than it does the $E\sigma^{32} - 10$ region.

Temporal control of *dnak* expression. The time of appearance of the *dnak* transcript was measured in synchronous



FIG. 5. (A) Schematic diagram of the dnaK dnaJ region. (B) Sequence of the 5' region of both the dnaK gene and the adjacent dnaJ gene. Features of the promoter regions P1 and P2 are indicated. (C) Primer extension mapping of the dnaK gene transcriptional start sites. An oligonucleotide complementary to nucleotides 433 to 456 within the dnaK coding sequence (underlined in Fig. 2) was synthesized and 5' end labeled with $[\gamma^{-32}P]ATP$. This probe was hybridized for 12 h at 45°C to 50 µg of total RNA from heat-shocked cells (lane 1) or from cells grown at 30°C at 0.85 cell cycle division unit (lane 2). Due to the very strong signal obtained with heat-shocked cells, lane 1 was underexposed. The hybrids were then extended by using reverse transcriptase as described in Materials and Methods. The sequencing ladder was generated by using the same 24-mer as a primer and M13mp18 containing the 1.1-kb BamHI-XhoI fragment (coding strand). The sequence shown is the complement. The potential transcription start sites are indicated by arrows.



FIG. 6. Expression of the *dnaK* gene as a function of the *C. crescentus* cell cycle. (A) Nuclease S1 protection assays of *dnaK* transcripts isolated from samples of synchronized CB15N cells at the cell division units indicated below the autoradiogram. RNA (50 μ g) was from cells heat shocked for 10 min (HS) or was isolated from cells at 0, 0.15, 0.25, 0.35, 0.5, 0.65, 0.85, or 1.0 cell division unit. The probe was the 1.1-kb *Bam*HI-XhoI fragment labeled at the 5' end. (B) Expression of luciferase from the *dnaK-lux* transcription fusion on plasmid pJG42 as a function of the cell cycle. Cultures of *C. crescentus* CB15N containing pJG42 were synchronized, and the synthesis of luciferase at the indicated division units of the cell cycle shown below the graph of activity was monitored by measuring bioluminescence as described in Materials and Methods. O.D., Opical density. One count per second is equivalent to 273 photons emitted per s.

cultures grown at 30°C with the S1 nuclease protection assay (Fig. 6A). Total RNA isolated at different times during the cell cycle was hybridized to a ³²P-labeled 1.1-kb *Bam*HI-*Xho*I probe (Fig. 1) and then treated with S1 nuclease. The 280-nucleotide protected fragment was present at high levels between 0.25 and 0.35 division units, coincident with the initiation of DNA replication. After a period of low transcript levels, the RNA fragment was again detected in large amounts near the end of the cell cycle, just before the initiation of DNA replication in the progeny stalked cell. The 320-nucleotide fragment, which was protected by RNA isolated from heat-shocked cells, was also detected in these

Promoter	-35 region	Spacing (bp)	-10 region			
$E\sigma^{32}$ consensus	T tC CcCTTGAA	13–15	CCCCATtT			
E. coli						
dnaK P1	TCTCCCCCTTGAT	14	CCCCATTT			
dnaK P2	TTGGGCAGTTGAA	13	CCCCTATT			
groE	TTTCCCCCTTGAA	13	CCCCATTT			
C. crescentus						
dnaK P1	TTATGGCCTTGCG	14	CCCCATAT			
dnaJ	GAGCCGCCTTGCC	15	TCCCACTT			
$E\sigma^{70}$ consensus	TTGACA	16–18	TATAAT			
C. crescentus dnaK P2	TTGACG	15	CACAAC			

TABLE 1. E. coli and C. crescentus promoter sequences

^a From Cowing et al. (7).

experiments, but at lower levels. At normal physiological temperatures the transcript from the heat shock promoter P1 was temporally controlled, analogous to the transcript from the P2 promoter.

To determine whether the sequences 5' to the start of the dnaK gene were sufficient to control its temporal expression, the 1.35-kb BamHI-PstI fragment containing the entire upstream region plus 370 bp of the coding region of the dnaK gene (Fig. 1) was subcloned into the transcription reporter vector pUCD615. This vector contains a promoterless lux operon from Vibrio fisheri (16). The expression of the lux genes fused to the dnaK promoter was monitored by measuring photon emission as described in Materials and Methods. Expression of the transcription fusion directed by the dnaK promoter was cell cycle regulated (Fig. 6B). The promoter region of dnaK appeared to initiate transcription of the chimeric gene during the swarmer to stalk transition and again in the predivisional cell just before cell division. In each instance the gene was expressed immediately before the initiation of DNA replication.

DISCUSSION

hsp70 genes are present in a wide variety of species. Genes related to Drosophila hsp70 have been isolated from organisms as diverse as E. coli, Saccharomyces cerevisiae, maize, chickens, and humans (21). A high degree of similarity exists among these genes and their encoded proteins. A comparison of the protein-coding region of the E. coli and Drosophila genes, for instance, reveals 57% identity at the nucleotide level (4). Using a Drosophila hsp70 gene probe, we isolated a fragment of C. crescentus DNA that appeared to encode the DnaK heat shock protein. Sequence analysis revealed that the predicted amino acid sequence of the C. crescentus DnaK protein was 68% identical to E. coli DnaK and 49% identical to Drosophila hsp70. Sequence analysis of the region downstream of the dnaK gene revealed a coding sequence for a peptide that was 59% identical to E. coli DnaJ.

The heat shock response in *E. coli* is regulated by the *rpoH* gene product, σ^{32} , which directs RNA polymerase to initiate transcription from heat shock promoters at all temperatures (7). The increased synthesis of heat shock proteins after a shift to higher temperatures is caused by increased transcription initiation from heat shock gene promoters by σ^{32} RNA polymerase. Recent experiments have demonstrated that the concentration of $E\sigma^{32}$ transiently increases after a shift to a high temperature, due to an increase in *rpoH*

mRNA level occurring mainly through a posttranscriptional mechanism (11, 14). Three heat-inducible promoters for the operon containing dnaK and dnaJ genes were identified both in vivo and in vitro in *E. coli*. These promoters are also used at normal temperatures. In *C. crescentus*, the dnaK gene also had two transcription start sites, but only one was heat inducible.

The promoter corresponding to the transcript that increased in abundance during the heat shock response, P1, had a -10 region that conforms to the consensus -10 region for heat shock promoters of *E. coli* (4), and its -35 region fits reasonably well with the consensus -35 region of *E. coli* heat shock promoters (Table 1). These results suggest that *C. crescentus* contains a minor sigma factor similar to σ^{32} of *E. coli*. A heat shock protein of 37 kilodaltons was previously detected in *C. crescentus* by immunoprecipitation with anti-RNA polymerase antibody (29). This protein is a candidate for the *C. crescentus* heat shock sigma factor but awaits biochemical analyses with the purified protein.

The P2 promoter, which appears to conform to the $E\sigma^{70}$ promoter and the σ^{70} promoters found in front of some C. crescentus genes, was active only during growth at physiologic temperature (30°C) but was under temporal control. All of the regulatory sequences required for temporal control of dnaK transcription appeared to be within the 820 bp 5' to the start of transcription. This conclusion is supported by the observation that this 5' regulatory region and 370 bp of dnaK coding sequence fused to a promoterless lux operon resulted in the same temporal expression of the lux activity as the expression of the chromosomal dnaK gene. The C. crescentus dnaK gene was transcribed just preceding the S phase during the transition from swarmer to stalked cells and again in late predivisional cells just before the initiation of DNA replication in the progeny stalked cell. These results confirm previous observations that the C. crescentus dnaK protein is synthesized at defined times in the cell cycle (13).

Transcription at the putative heat shock promoter under normal physiological conditions also varied as a function of the cell cycle, showing an increase at the transition from swarmer to stalked cells and again when cells were beginning to divide (0.85 to 1.0 division units). There are at least two possible explanations for this. In other systems the increase in transcription from heat shock promoters is mediated through increases in the level of σ^{32} . It may be that the comparable heat shock sigma factor gene is under cell cycle modulation and that other heat shock proteins are also increased at this point during the cell cycle. More likely, *cis*-acting sequences that contribute to temporally controlled modulation of gene expression affect the rate of dnaK transcription by both the constitutive but low levels of the heat shock sigma factor and the sigma 70 factor.

The function of DnaK in C. crescentus is not known, although there are a number of possible roles for this heat shock protein during the normal cell cycle. DnaK and other heat shock proteins might participate in global control of development in C. crescentus. This function may be mediated through stability and proteolysis of cell-type-specific proteins. Alternatively, the DnaK protein might be required for some aspect of C. crescentus DNA replication. It is known that the E. coli DnaK and DnaJ proteins are required for successful initiation of phage lambda DNA replication (19), and recently a mutant allele of dnaK has been shown to cause conditionally defective initiation of E. coli DNA replication (31). Maximal transcription of dnaK in C. crescentus reflects the time of initiation of new rounds of DNA replication; in fact, the DnaK protein specifically segregates to the progeny stalked cell, in which a new round of DNA replication occurs (29). These results are analogous to that seen with the cell-cycle-controlled transcription of the human hsp70 gene in HeLa cells (23). Thus, the expression of a highly conserved heat shock protein in such diverse cell types as a bacterial cell and a human cell shows parallel cell cycle control under normal physiological conditions.

ACKNOWLEDGMENTS

We thank E. Craig for generously providing the *Drosophila* hsp70 gene probe, F. Reinach for kindly transmitting his expertise on DNA sequence determination, and S. Reuter for technical assistance in the initial stage of this investigation.

This investigation was supported, in part, by grants to S.L.G. from Fundacao de Amparo a Pesquisa do Estado de Sao Paulo (FAPESP 87/1577-3), Financiadora de Estudos e Projetos (FINEP 43870160.00), Third World Academy of Sciences (RGBC89-17), and to L.S. by Public Health Service grant GM32506 from the National Institutes of Health and grant mv408 from the American Cancer Society. S.L.G. was partially supported by Conselho Nacional de Desenvolvimento Científico and Coordenadoria de Aperfeicoamento de Pessoal de Ensino Superior. J.W.G. is a Helen Hay Whitney Fellow.

LITERATURE CITED

- Amemiya, K., V. Bellofatto, L. Shapiro, and J. Feingold. 1986. in vitro and in vivo transcription initiation at a highly conserved promoter within a 16S ribosomal RNA. J. Mol. Biol. 187:1–14.
- Amemiya, K., B. Raboy, and L. Shapiro. 1980. Involvement of host RNA polymerase in the early transcription program of *Caulobacter crescentus* bacteriophage φCd1 DNA. Virology 104:109-114.
- Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1987. Current protocols in molecular biology, p. 4.8.1–4.8.3. John Wiley & Sons, Inc., New York.
- Bardwell, J. C. A., and E. A. Craig. 1984. Major heat shock gene of *Drosophila* and the *E. coli* heat inducible *dnaK* gene are homologous. Proc. Natl. Acad. Sci. USA 81:848-852.
- Bardwell, J. C. A., K. Tilly, E. Craig, J. King, M. Zylicz, and C. Georgopoulos. 1986. The nucleotide sequence of the *E. coli* K12 *dna J* gene. J. Biol. Chem. 261:1782–1785.
- Contreras, I., L. Shapiro, and S. Henry. 1987. Membrane phospholipid composition of *Caulobacter crescentus J. Bacte*riol. 135:1130–1136.
- Cowing, D. W., J. C. A. Bardwell, E. A. Craig, C. Woolford, R. W. Hendrix, and C. A. Gross. 1985. Consensus sequence for *E. coli* heat shock gene promoters. Proc. Natl. Acad. Sci. USA

82:2679-2683.

- 8. Degnan, S. T., and A. Newton. 1972. Chromosome replication during development in *Caubacter crescentus*. J. Mol. Biol. 64:671-680.
- Ditta, G., S. Stanfield, D. Corbin, and D. R. Helinski. 1980. Broad host range DNA cloning system for gram negative bacteria: construction of a gene bank of *Rhizobium meliloti*. Proc. Natl. Acad. Sci. USA 77:7347-7351.
- 10. Ely, B., and C. Gerardot. 1988. Use of pulsed field gradient gel electrophoresis to construct a physical map of the *Caulobacter crescentus* genome. Gene 68:323-330.
- 11. Erickson, J. W., V. Vaughn, W. A. Walter, F. C. Neidhart, and C. A. Gross. 1987. Regulation of the promoters and transcripts of *rpoH*, the *E. coli* heat shock regulatory gene. Genes Dev. 1:419-432.
- Evinger, M., and N. Agabian. 1977. Envelope-associated nucleoid from *Caulobacter crescentus* stalked and swarmer cells. J. Bacteriol. 132:294-301.
- Gomes, S. L., M. H. Juliani, J. C. C. Maia, and A. M. Silva. 1986. Heat shock protein synthesis during development in *Caulobacter crescentus*. J. Bacteriol. 168:923-930.
- Grossman, A. D., D. B. Straus, W. A. Walter, and C. A. Gross. 1987. σ³² synthesis can regulate the synthesis of heat shock proteins in *E. coli*. Genes Dev. 1:179–184.
- Herendeen, S. L., R. A. VanBogelen, and F. C. Neidhardt. 1979. Levels of major proteins of *E. coli* during growth at different temperatures. J. Bacteriol. 139:185-194.
- Hirooka, T., P. M. Rogowsky, and C. I. Kado. 1987. Characterization of the virE locus of Agrobacterium tumefaciens plasmid pTiC58. J. Bacteriol. 169:1529–1536.
- 17. Ingolia, T. D., and E. A. Craig. 1982. *Drosophila* gene related to the major heat shock-induced gene is transcribed at normal temperatures and not induced by heat shock. Proc. Natl. Acad. Sci. USA 79:525–529.
- Ingolia, T. D., E. A. Craig, and J. McCarthy. 1980. Sequence of three copies of the major *Drosophila* heat shock induced protein and their flanking regions. Cell 21:669–679.
- LeBowitz, J. H., M. Zylicz, C. Georgopoulos, and R. McMacken. 1985. Initiation of DNA replication on single stranded DNA templates catalyzed by purified replication proteins of bacteriophage lambda and *E. coli*. Proc. Natl. Acad. Sci. USA 82: 3988-3992.
- Lindquist, S. 1986. The heat-shock response. Annu. Rev. Biochem. 55:1115–1191.
- 21. Lindquist, S., and E. A. Craig. 1988. The heat shock proteins. Annu. Rev. Genet. 22:631-677.
- 22. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual, p. 454. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Milarski, I. K. L., and R. I. Morimoto. 1986. Expression of human HSP70 during the synthetic phase of the cell cycle. Proc. Natl. Acad. Sci. USA 83:9517-9521.
- 24. Miller, J. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Mues, G. I., T. Z. Munn, and J. D. Raese. 1986. A human gene family with sequence homology to *Drosophila melanogaster* Hsp70 heat shock gene. J. Biol. Chem. 261:874–877.
- Neidhardt, F. C., R. A. VanBogelen, and V. Vaughn. 1984. The genetics and regulation of heat shock proteins. Annu. Rev. Genet. 18:295-329.
- Ohki, M., F. Tamura, S. Nishimura, and H. Uchida. 1986. Nucleotide sequence of the *E. coli dnaJ* gene and purification of the gene product. J. Biol. Chem. 261:1778–1781.
- Paek, K.-H., and G. C. Walker. 1987. Escherichia coli dnaK null mutants are inviable at high temperature. J. Bacteriol. 169: 283-290.
- Reuter, S. H., and L. Shapiro. 1987. Asymmetric segregation of heat-shock proteins upon cell division in *Caulobacter crescentus*. J. Mol. Biol. 194:653-662.
- 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.
- 31. Sakakibara, Y. 1988. The dnaK gene of Escherichia coli func-

tions in initiation of chromosome replication. J. Bacteriol. 170:972-979.

- 32. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- 33. Schmetterer, G., C. P. Wolk, and J. Elhai. 1986. Expression of Luciferases from Vibrio harveyi and Vibrio fischeri in filamentous Cyanobacteria. J. Bacteriol. 167:411-414.
- 34. Werner-Washburne, M., D. E. Stone, and E. A. Craig. 1987. Complex interactions among members of an essential subfamily of hsp70 genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 7:2568-2577.
- 35. Yochem, J., H. Uchida, M. Sunshine, H. Saito, C. P. Georgopoulos, and M. Feiss. 1978. Genetic analysis of two genes, *dnaJ* and *dnaK*, necessary for *E. coli* and bacteriophage lambda DNA replication. Mol. Gen. Genet. 164:9–14.