JOURNAL OF BACTERIOLOGY, Jan. 1997, p. 487–495 Vol. 179, No. 2
0021-9193/97/$04.00+0

Cloning and Characterization of Two *groESL* Operons of *Rhodobacter sphaeroides*: Transcriptional Regulation of the Heat-Induced *groESL* Operon

W. THEODORE LEE, KATHERINE C. TERLESKY, AND F. ROBERT TABITA*

Department of Microbiology, The Ohio State University, Columbus, Ohio

Received 12 September 1996/Accepted 1 November 1996

The nonsulfur purple bacterium *Rhodobacter sphaeroides* was found to contain two *groESL* operons. The $groESL_1$ heat shock operon was cloned from a genomic library, and a 2.8-kb DNA fragment was sequenced and found to contain the *groES* and *groEL* genes. The deduced amino acid sequences of $GroEL_1$ (cpn60) and $GroES_1$ (cpn10) were in agreement with N-terminal sequences previously obtained for the isolated proteins (K. C. Terlesky and F. R. Tabita, Biochemistry 30:8181–8186, 1991). These sequences show a high degree of similarity to *groESL* genes isolated from other bacteria. Northern analysis indicated that the $groESL_1$ genes were expressed as part of a 2.2-kb polycistronic transcript that is induced 13-fold after heat shock. Transcript size was not affected by heat shock; however, the amount of transcript was induced to its greatest extent 15 to 30 min after a 40°C heat shock, from an initial temperature of 28°C, and remained elevated up to 120 min. The *R. sphaeroides* $groESL_1$ operon contains a putative hairpin loop at the start of the transcript that is present in other bacterial heat shock genes. Primer extension of the message showed that the transcription start site is at the start of this conserved hairpin loop. In this region were also found putative −35 and −10 sequences that are conserved upstream from other bacterial heat shock genes. Transcription of the $groESL_1$ genes was unexpectedly low under photoautotrophic growth conditions. Thus far, it has not been possible to construct a $groESL_1$ deletion strain, perhaps indicating that these genes are essential for growth. A second operon ($groESL_2$) was also cloned from *R. sphaeroides*, using a $groEL_1$ gene fragment as a probe; however, no transcript was observed for this operon under several different growth conditions. A $groESL_2$ deletion strain was constructed, but there was no detectable change in the phenotype of this strain compared to the parental strain.

The *groESL* genes were initially identified in *Escherichia coli*, where their products were shown to be required for the proper assembly and maturation of bacteriophage (11). Subsequently, the *groESL* genes were determined to encode proteins that are synthesized at elevated levels after cells are shocked with heat or other forms of stress. The finding that the GroEL protein showed homology to a chloroplast protein that appeared to bind and assist in the assembly of ribulose bisphosphate carboxylase/oxygenase (RubisCO) (18) eventually led to the unequivocal demonstration that GroEL and GroES are necessary for the proper folding of RubisCO from unfolded monomers in vitro and in vivo (14, 15). Eventually, it was recognized that these and related proteins act as molecular chaperones (6) for the productive folding and, perhaps, eventual assembly of several proteins in all types of cells (49). Presumably, these genes and their products are highly expressed after stress to assist the refolding of necessary proteins that may become denatured or unstable under adverse conditions. The *groES* gene encodes a protein with a molecular weight of 10,000, often referred to as chaperonin 10 (cpn10, hsp10, and GroES). cpn60 (also referred to as hsp60 or GroEL), encoded by the *groEL* gene, is a protein with a molecular weight of 60,000 that forms tetradecameric complexes, the structure of which has recently been described (2). These cpn60 complexes possess weak ATPase activity (19, 20) that is modulated by cpn10 (3). The cpn60 protein binds the unfolded substrate protein, which is released in a folded state, following hydrolysis of ATP. In many instances, cpn10 mediates the release of the folded protein (27). Not surprisingly, the *groESL* genes are essential for viability in *E. coli* (8) since there appears to be only a single copy of these genes in this organism.

As described above, the interactions and function of chaperonins in mediating the folding of model proteins have been most thoroughly described for RubisCO (39), the key enzyme of the Calvin-Benson-Bassham reductive pentose phosphate pathway of CO_2 fixation (37). This enzyme, which catalyzes the actual fixation of CO_2 in this pathway, requires chaperonins 10 and 60 for the folding and subsequent assembly of RubisCO both in vivo and in vitro (14, 15). In the purple nonsulfur bacterium *Rhodobacter sphaeroides*, used as a model to study the regulation and biochemistry of carbon fixation in prokaryotes (37), there are two RubisCO enzymes, the biosynthesis of which is both differentially and interdependently regulated, depending on the growth conditions (12, 37). Previously, it was shown that cpn10 and cpn60 synthesis parallel the regulated production of RubisCO under different growth conditions (38), suggesting important links among chaperonin synthesis, RubisCO control, and RubisCO function. Moreover, the fact that cpn60 synthesis is regulated when *R. sphaeroides* is grown with different sources of carbon (38) suggests that this chaperonin may exhibit some specificity with respect to its substrate protein(s). Indeed, there is some precedence for this since the levels of synthesis of two proteins involved in carbon metabolism in *E. coli*, 6-phosphogluconate dehydratase and aldolase, were both shown to be affected by changes in the cellular level of cpn 60 and cpn10 (21). Furthermore, both *Rhizobium meliloti* and *Bradyrhizobium japonicum*, organisms within the alpha subdivision of gram-negative bacteria, which also includes *R. sphaeroides*, contain more than one *groE*

* Corresponding author. Mailing address: Department of Microbiology, The Ohio State University, 484 W. 12th Avenue, Columbus, OH 43210-1292. Phone: (614) 292-4297. Fax: (614) 292-6337. E-mail: rtabita@magnus.acs.ohio-state.edu.

TABLE 1. Plasmids used in this study

Plasmid(s)	Relevant characteristics	Source or reference
pK18, pK19	Km^r, pUC derivatives	30
pUC18	Am^r	46
pKT8.0	pK18 containing 8.0-kb *Sal*I fragment containing entire *groESL*$_1$ operon	This study
pG2BB	pK18 containing 7.0-kb *Bam*HI fragment containing entire *groESL*$_2$ operon	This study
pG2BgBm	pK18 containing 4.0-kb *Bam*HI-*Bgl*II fragment of pG2BB	This study
pUCG2	pUC18 containing 4.0-kb insert of pG2BgBm as *Eco*RI-*Hin*dIII fragment	This study
pUCG2ΔS	pUCG2 digested with *Sph-Xba*I (in polylinker) and filled in	This study
pUC1318K	pUC1318 with kanamycin resistance gene	7
pG2ΔKn	pUCG2ΔS lacking 1.62-kb *Sal-Pst*I fragment with Km^r gene cartridge inserted	This study
pJP5603Ω	pKP5603 containing Ω fragment	H. Xu; this laboratory
pJPG2ΔKn	pJP5603Ω with 4.4-kb *Kpn*I fragment of pG2ΔKn inserted	This study

operon (10, 31). While the physiological role of multiple chaperonin genes has not been completely elucidated, a study of the *groE* operons in *B. japonicum* has indicated that one of these (*groESL*$_3$) is specifically expressed under conditions in which the nitrogen fixation genes are induced (10).

To further our study of chaperonin-mediated RubisCO folding and the regulation of chaperonin synthesis (38), we have isolated two distinct *groESL* operons from *R. sphaeroides*.

(A preliminary report of this study appeared recently [24]).

MATERIALS AND METHODS

Materials. Sequenase and Taquence were obtained from U.S. Biochemical, Cleveland, Ohio. Restriction enzymes, DNA polymerase I, and T4 DNA ligase were from Gibco Bethesda Research Laboratories, Gaithersburg, Md. ^{32}P- and ^{35}S-labeled nucleotides were from Dupont, NEN Research Products, Boston, Mass. The Sequitherm cycle sequencing kit used was from Epicentre Technologies, Madison, Wis. The random primed labeling kit and avian myeloblastosis virus reverse transcriptase were from Boehringer Mannheim, Indianapolis, Ind. Plasmid pRJ7928 (10), containing *B. japonicum groESL*$_2$, was a gift from H. M. Fisher. Sequence analysis and comparisons were performed with the Wisconsin Sequence Analysis package from the Genetics Computer Group, Madison, Wis.; PC/Gene from IntelliGenetics, Inc., Mountain View, Calif.; and the MacVector and AssemblyLIGN software programs from Kodak Scientific Imaging Systems, New Haven, Conn.

Bacterial strains and plasmids. *R. sphaeroides* HR is a derivative of ATCC 17023 but is capable of growth at 42°C (44). The media and growth conditions used were as previously described (38). The plasmid vectors and constructs used in this study and their relevant characteristics are listed in Table 1.

Isolation of DNA fragments. Southern hybridization and colony blotting of *R. sphaeroides* genomic DNA, using the *B. japonicum groESL*$_2$ DNA as a probe, were done with 50% formamide–4× SSC (1× SSC is 150 mM NaCl plus 15 mM sodium citrate)–20 mM Tris-HCl (pH 7.6)–5% dextran sulfate–0.1% sodium dodecyl sulfate (SDS) at 40°C and washing in 2× SSC–1% SDS at 45°C. Southern hybridization to detect additional *groE* genes of the *R. sphaeroides* genome was performed with the 1.0-kb *Sph*I-*Sma*I restriction fragment of the heat-induced *groEL*$_1$ gene. These Southern blots, and subsequent colony blots of cloned DNA fragments, were hybridized in 6× SSC–2× Denhardt solution (0.01% Ficoll, 0.01% polyvinylpyrrolidone, 0.01% bovine serum albumin) containing 0.1% SDS and 100 μg of denatured salmon sperm DNA per ml. The blots were hybridized at 60°C overnight and then washed three or four times in 5× SSC at 60°C.

DNA sequencing. DNA sequencing was performed by the dideoxy-chain termination method (33) to sequence double-stranded DNA templates, with Sequenase. Plasmid DNA was obtained from standard alkaline lysis preparations (32). To resolve compressions due to the high G-C content of *R. sphaeroides* DNA, both deaza-dGTP and dITP were used. In addition, some sequence data was obtained by using either Taquence or the Sequitherm kit. The entire sequence was read on both strands to ensure that all compressions were resolved. Calculation of the free energy of hydrolysis of putative hairpin loop structures was determined with the HAIRPIN program of PC/Gene.

RNA isolation and analysis. Total cellular RNA was isolated from cells (50 ml per preparation) grown to mid-log phase (A_{660} of 0.6 to 0.8). The cells were transferred to 0.1 volume of stop solution (50 mM Tris-HCl [pH 8.0], 20 mM EDTA [pH 8.0], 20 mM NaN_3) containing rifampin (1.5 mg/ml) and centrifuged at 3,290 × *g* for 10 min. The cell pellet was resuspended in 2 ml of boiling lysis buffer (100 mM Tris-HCl [pH 8.0], 2 mM EDTA [pH 8.0], 2% SDS) and kept in a boiling water bath for 5 min, after which 35 μl of 4 M potassium acetate was added. This was followed by a 5-min incubation on ice. The cell suspension was then centrifuged in a microcentrifuge for 10 min; the resulting supernatant was phenol extracted once and ethanol precipitated. The precipitated RNA samples were subsequently resuspended in 3 ml of water to which 1.5 g of cesium chloride was added. The RNA-cesium chloride solution was then layered over a 1.5-ml cushion of 5.7 M cesium chloride in an ultracentrifuge tube. The samples were centrifuged at 34,000 rpm for 19 h at 15°C in a Beckman SW50.1 rotor. The precipitate was then resuspended in water and ethanol precipitated. After resuspension of the RNA in water, the samples were used for Northern analysis (32) and primer extension analysis. Quantitation of bands on Northern blots was performed with an Instant Imager (Packard, Meriden, Conn.). Northern hybridizations were standardized by using a 16S rRNA gene probe as a control. This probe was prepared from genomic DNA after amplification of the specific DNA by PCR with 5′ and 3′ oligonucleotides based on the reported sequence (5). The amplified DNA was subsequently subcloned into plasmid pK18 and used for the 16S rRNA probe. Primer extension analysis was performed by a modification of the method of Triezenberg (40). A 19-mer oligonucleotide (5′ ACGGTCATG CAGCGGTTTG 3′), complementary to the +535 to +553 region of the *R. sphaeroides groES* coding sequence, was kinase treated with 50 μCi of [γ^{32}P]ATP, ethanol precipitated three times, and then hybridized to 100 to 200 μg of total *R. sphaeroides* RNA. The primer extension reaction was performed with avian myeloblastosis virus reverse transcriptase.

Deletion mutagenesis of *groESL*$_2$. Digestion of pUCG2ΔS with *Sal*I and *Pst*I removes a 1.62-kb fragment within the *groESL*$_2$ genes. The blunt-ended *Hin*dIII kanamycin resistance cassette from pUC1318K was inserted into the blunt-ended *Sal*I-*Pst*I site. A 4.2-kb *Kpn*I fragment containing the deletion construct was ligated into pJP5603Ω. Selection for the *R. sphaeroides groESL*$_2$ deletion strain was made by mating pJPG2ΔKn, via *E. coli* S17-1 λ*pir*, into *R. sphaeroides* HR on nonselective PYE (44) plates and then plating cell mixtures from overnight incubations onto Omerod's minimal malate plates containing kanamycin. Double recombinants were scored as those Km^r transconjugants that were Sp^s, indicating the loss of plasmid vector sequences (29). Southern analysis verified that the expected homologous recombination had taken place.

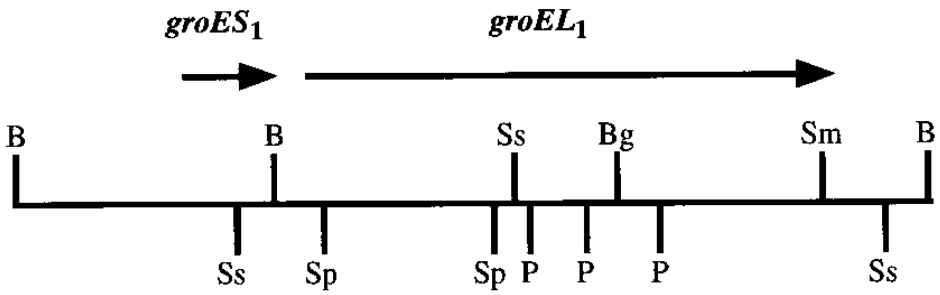
Nucleotide sequence accession numbers. The nucleotide sequence of the *R. sphaeroides groESL*$_1$ operon has been deposited in the GenBank library under accession number U37369; the nucleotide sequence of the *groESL*$_2$ operon is under accession number U66831.

RESULTS

Cloning and sequencing of the *groESL* operons. Southern blots of restriction enzyme digests of *R. sphaeroides* genomic DNA probed with a 1.7-kb *Apa*I-*Bst*I fragment of the *B. japonicum groESL*$_2$ operon exhibited a single *Sal*I restriction fragment of approximately 8.0 kb. *Sal*I-digested genomic DNA, ranging between 6 and 9 kb, was isolated and ligated to plasmid pK18, and a plasmid containing an 8.0-kb *Sal*I insert was isolated after screening of this partial library with the *B. japonicum groESL*$_2$ probe. The approximate location of the *groESL* genes within the 8-kb restriction fragment was determined by Southern analysis; subsequent sequencing verified that the *R. sphaeroides groESL*$_1$ operon had been cloned. The *R. sphaeroides groESL*$_1$ operon was localized to a 2.8-kb restriction fragment (Fig. 1A) and subsequently completely sequenced with additional plasmids and primers. The DNA sequence (Fig. 1B) of the *groESL*$_1$ operon revealed two open reading frames with a high degree of similarity to other bacterial *groESL* operons.

Southern blots of *R. sphaeroides* genomic DNA, probed with a 1.0-kb *Sph*I-*Sma*I *groEL*$_1$ fragment, showed the presence of a second putative *groESL* operon (Fig. 2). A 2.3-kb *Kpn*I restriction fragment was isolated from a library of *R. sphaeroides* genomic DNA. Partial DNA sequencing of this restriction fragment indicated that it contained most of a second *groEL*

A.



B.

```
                                                                                                    GGATCCTTCGCCATCGCCTCCTTGAGCCGCGCCTCCTCCTCGGGGC   (46)
CTGCTTCCACGAGCAGCGCGGTGGCCTGCTCGATCCGCGGCCGGAGCCGGTCGAGCAGGGCCGCCTGTCGCGGGTCTCCCAGATGATAGGTTCCGGCCACCGTCACCTCGCGTCCCTCGC  (166)
GCGTTGCCTGCCAGAAGTTGCCTTCGGAGAAGGGCTGTCCTGCGGCCACGGCCTCCAGTTCCGCTCGGTCGGGCTGGGGAAGGGCCGCGATGAGGTCCGTGCCGGTGCAGGCGGCGTGGG  (286)
CCAGCGGCGCGGAGGCGAGGAAGAAGGCACAGATGATCGAGGTGAGGCGGGTCATGCTCAGGTCCTTGCGACGGCAGGCAGAAGGTGGCACGCCACGCCTCCCGCGCCAAGGCGGAGCGT  (406)
GCATTTTGATGCATCCCCCCGTTGACAGGTGCCGGACGCTCTCATATCTCTCGCGTCGTTGGCACTCGCCTCGGGTGAGTGCTAATCACATTCAACCTGGAACCTAGGGAGTGTTCTCAG  (526)

ATGGCTTTCAAACCGCTGCATGACCGTGTGCTGGTCCGCCGCGTCCAGAGCGACGAAAAGACCAAGGGCGGTCTGATCATCCCCGATACCGCCAAGGAAAAACCGGCTGAAGGCGAAGTC  (646)
 M  A  F  K  P  L  H  D  R  V  L  V  R  R  V  Q  S  D  E  K  T  K  G  G  L  I  I  P  D  T  A  K  E  K  P  A  E  G  E  V

GTGTCCTGCCGCCGAAGGCGCCCGCAAGGATTCGGGCGAGCTCATCGCCATGTCGGTGAAGGCGGGCGACCGCGTGCTGTTCGGCAAATGGTCGGGCACCGAAGTCACCATCGACGGTGCC  (766)
 V  S  C  G  E  G  A  R  K  D  S  G  E  L  I  A  M  S  V  K  A  G  D  R  V  L  F  G  K  W  S  G  T  E  V  T  I  D  G  A

GAGCTGCTCATCATGAAGGAAAGCGACATCCTGGGGATCCTCAGCTGATCCCGAGCTGACCGCGGCTTGGTCCCGGCCATGCGCCAGGGGCTTCGAGCCGTCCCCGAGACGATCGCATCA  (886)
 E  L  L  I  M  K  E  S  D  I  L  G  I  L  S  *

ACCAGACAAACAGGAGCAACAACATGGCTGCCAAGGACGTCAAGTTCGACACCGATGCCCGCGACCGCATGCTGCGTGGCGTGAACATCCTCGCCGATGCGGTGAAGGTCACGCTGGGCC  (1006)
                    M  A  A  K  D  V  K  F  D  T  D  A  R  D  R  M  L  R  G  V  N  I  L  A  D  A  V  K  V  T  L  G

CGAAAGGCCGCAACGTCGTGATCGACAAGTCGTTCGGCGCGCCGCGCATCACCAAGGACGGTGTGTCGGTCGCCAAGGAGATCGAACTCTCCGACAAGTTCGAGAACATGGGCGCCCAGA  (1126)
P  K  G  R  N  V  V  I  D  K  S  F  G  A  P  R  I  T  K  D  G  V  S  V  A  K  E  I  E  L  S  D  K  F  E  N  M  G  A  Q

TGGTGAAGGAAGTCGCTTCCCGCACCAACGACGAGGCGGGTGACGGCACCACCACCGCCACCGTGCTCGCCCAGGCGATCATCAAGGAAGGCCTCAAGGCCGTCGCCGCCGGCATGAACC  (1246)
M  V  K  E  V  A  S  R  T  N  D  E  A  G  D  G  T  T  T  A  T  V  L  A  Q  A  I  I  K  E  G  L  K  A  V  A  A  G  M  N

CGATGGACCTCAAGCGCGGCATCGACCTCGCGACCTCGAAAGTCGTCGAGGCGATCAAGGCCGCCGCCCGTCCGGTGAACGACTCGCACGAAGTCGCTCAGGTCGGCACGATCTCGGCCA  (1366)
P  M  D  L  K  R  G  I  D  L  A  T  S  K  V  V  E  A  I  K  A  A  A  R  P  V  N  D  S  H  E  V  A  Q  V  G  T  I  S  A

ACGGCGAAGCGCAGATCGGCCGCTTCATCGCTGACGCGATGCAGAAGGTCGGCAACGAGGGTGTCATCACCGTCGAAGAGAACAAGGGCCTCGAGACCGAAGTCGAAGTCGTCGAAGGCA  (1486)
N  G  E  A  Q  I  G  R  F  I  A  D  A  M  Q  K  V  G  N  E  G  V  I  T  V  E  E  N  K  G  L  E  T  E  V  E  V  V  E  G

TGCAGTTCGACCGCGGCTACCTCTCGCCCTACTTCGTCACCAACGCCGACAAGATGACGGCCGAGCTCGACGACGTCTACATCCTGCTCCACGAGAAAAAACTCTCGTCGCTGCAGCCGA  (1606)
M  Q  F  D  R  G  Y  L  S  P  Y  F  V  T  N  A  D  K  M  T  A  E  L  D  D  V  Y  I  L  L  H  E  K  K  L  S  S  L  Q  P

TGGTCCCGCTGCTCGAGGCCGTGATCCAGTCGCAGAAGCCGCTGCTGATCATCGCCGAGGACGTGGAAGGCGAAGCCCTCGCCACGCTCGTGGTCAACAAGCTGCGCGGTGGCCTGAAGA  (1726)
M  V  P  L  L  E  A  V  I  Q  S  Q  K  P  L  L  I  I  A  E  D  V  E  G  E  A  L  A  T  L  V  V  N  K  L  R  G  G  L  K

TCGCTGCCGTCAAGGCTCCGGGCTTCGGTGACCGTCGCAAGGCCATGCTGCAGGACATCGCGATCCTGACCGGCGGCCAGGTGATCTCGGAAGACCTCGGCATGAAGCTCGAGAATGTCA  (1846)
I  A  A  V  K  A  P  G  F  G  D  R  R  K  A  M  L  Q  D  I  A  I  L  T  G  G  Q  V  I  S  E  D  L  G  M  K  L  E  N  V

CCATCGACATGCTCGGCCGCGCCAAGAAGATCTCGATCAACAAGGACAACACCACGATCGTGGACGGCAACGGCGACAAGGCCGAGATCGACGCGCGCGTGGCCCAGATCCGCAACCAGA  (1966)
T  I  D  M  L  G  R  A  K  K  I  S  I  N  K  D  N  T  T  I  V  D  G  N  G  D  K  A  E  I  D  A  R  V  A  Q  I  R  N  Q

TCGAGGAAACCTCGTCCGACTACGACCGCGAGAAGCTGCAGGAGCGGGTCGCGAAACTGGCGGGCGGCGTGGCCGTCATCCGCGTCGGCGGCATGACCGAAGTCGAAGTGAAAGAGCGCA  (2086)
I  E  E  T  S  S  D  Y  D  R  E  K  L  Q  E  R  V  A  K  L  A  G  G  V  A  V  I  R  V  G  G  M  T  E  V  E  V  K  E  R

AGGACCGCGTCGATGACGCCCTGAACGCGACCCGTGCGGCCGTGCAGGAAGGCATCGTCGTCGGCGGCGGCGTGGCCCTGATCCAGGGCGGCAAGGCGCTCGACGGCCTGACCGGCGAGA  (2206)
K  D  R  V  D  D  A  L  N  A  T  R  A  A  V  Q  E  G  I  V  V  G  G  G  V  A  L  I  Q  G  G  K  A  L  D  G  L  T  G  E

ACCCCGACCAGAACGCGGGCATCACCATCGTGCGTCGCGCGCTGGAAGCTCCGCTGCGCCAGATCGCCCAGAACGCGGGCGTGGACGGTTCGGTCGTGGCCGGCAAGGTGCGCGAGTCGA  (2326)
N  P  D  Q  N  A  G  I  T  I  V  R  R  A  L  E  A  P  L  R  Q  I  A  Q  N  A  G  V  D  G  S  V  V  A  G  K  V  R  E  S

ACGAGAAGTCCTTCGGCTTCAACGCCCAGACCGAAGAATATGGCGACATGTTCAAGTTCGGCGTGATCGACCCGGCCAAGGTGGTTCGCACCGCTCTGGAAGACGCGGCCTCGGTCGCTT  (2446)
N  E  K  S  F  G  F  N  A  Q  T  E  E  Y  G  D  M  F  K  F  G  V  I  D  P  A  K  V  V  R  T  A  L  E  D  A  A  S  V  A

CGCTGCTCATCACCACCGAAGCCATGATCGCCGACAAGCCGGAGCCGAAATCTCCGGCCGGCGGCCCGGGCATGGGCGGCATGGGCGGCATGGACGGCATGATGTAATCCATCCACCGCA  (2566)
S  L  L  I  T  T  E  A  M  I  A  D  K  P  E  P  K  S  P  A  G  G  P  G  M  G  G  M  G  G  M  D  G  M  M  *

TGTCGGTCGGATCGAAAGGGGGCCTCCGGGCCCCCTTTCTCATGCGCATCCGCCGGTCAGACCGAGGGACGCCTGCACAGGAACGGGGCGCGGATCAGGCCGAGGGACGCATCCGCCGGA  (2686)

ACGGGTCGAGGATCACGCCGAGCTCGGTCATGCGTTGCCCTTCCGGTCAGGGGGCGGGGTGACGCTCCTCCGTGCCATTCGCGCTTCGCGAATGGCGGGCCGCCGCCGCGTCAGGCTGCAA  (2806)

TGGGCAGGGAAGGGGCTATTGCAGGATCTCGTCGGGATCC  (2846)
```

FIG. 1. Restriction map and nucleotide sequence of the *R. sphaeroides groESL*$_1$ operon. (A) Restriction map of the heat induced *groESL*$_1$ operon. The 2.8-kbp fragment was sequenced on both strands by using subclones and oligonucleotides. Abbreviations: B, *Bam*HI; Bg, *Bgl*II; P, *Pst*I; Sm, *Sma*I; Sp, *Sph*I; Ss, *SstI*. (B) Nucleotide sequence of the *R. sphaeroides groESL*$_1$ operon. The deduced amino acid sequences of the cpn10 and cpn60 proteins are listed below the DNA sequence. The underlined sequences represent potential −10 and −35 promoter sequences; the putative hairpin loop structures are denoted by the arrows.

gene, as well as some 3′ untranslated sequence. This *Kpn*I restriction fragment was used to isolate a 7.0-kb *Bam*HI restriction fragment that encoded the entire operon (Fig. 3A). Sequencing indicated the presence of both *groES* and *groEL* genes (Fig. 3B). However, it appears that the *groES*$_2$ gene is a pseudogene. There is homology to the *groES*$_1$ gene product in two of the *groES*$_2$ reading frames (Fig. 3B), and the codon usage is very poor for *R. sphaeroides*. In addition, the deduced

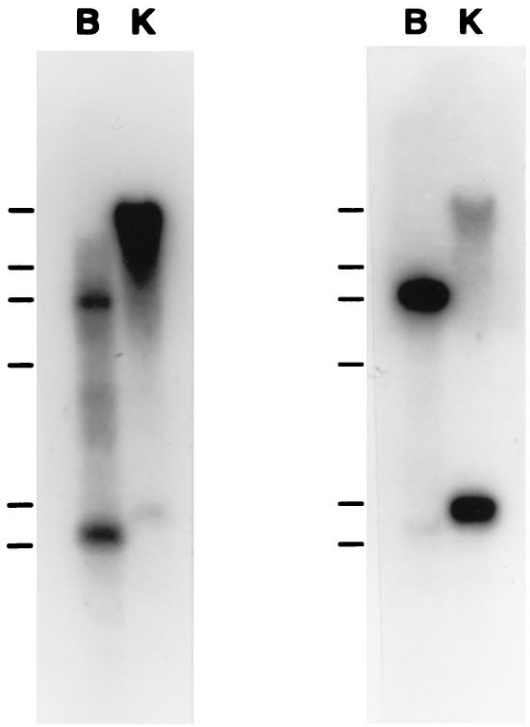


FIG. 2. Presence of two *groESL* operons in *R. sphaeroides*. Genomic DNA was digested to completion with *Bam*HI (B) and *Kpn*I (K), separated by agarose gel electrophoresis, blotted onto a nylon membrane, and probed with $groEL_1$ (left) and $groEL_2$ (right). The molecular size standards are (from the top) 23.1, 9.4, 6.6, 4.4, 2.3, and 2.0 kb.

protein would have a very high isoelectric point (pI) of 11.6. Finally, no $groESL_2$ transcript was seen following Northern analysis of RNA from cells grown photolithoautotrophically and photoheterotrophically at 28°C and heat shocked at 40°C. The $groEL_2$ gene product has a calculated molecular weight of 56,789 and a pI of 5.07 and uses the typical amino acid codons for *R. sphaeroides*. Use of either *R. sphaeroides* heat shock $groEL_1$ (1.0-kb *Sal*I-*Sph*I fragment) or $groEL_2$ (900-bp *Sph*I-*Pst*I fragment) as a probe gave no indication of additional *groEL* sequences (Fig. 2).

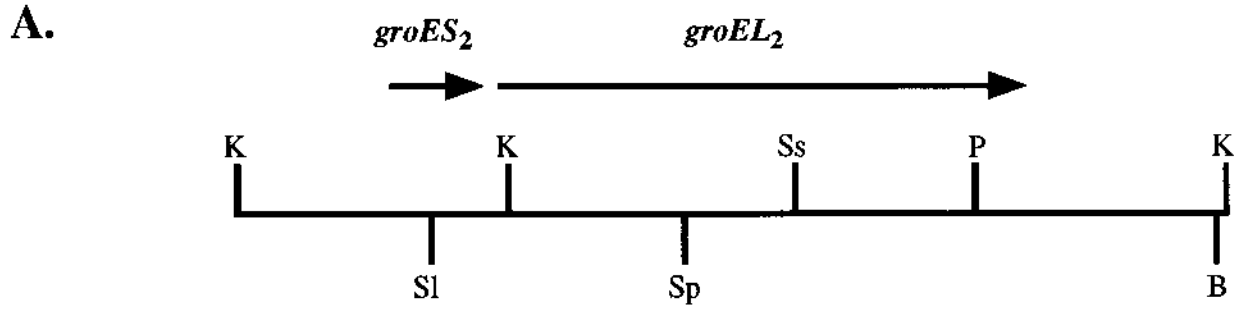
The deduced amino acid sequence indicated that $groES_1$ encodes a protein of 95 amino acids with a calculated molecular weight of 10,196 and a pI of 5.23. The $groEL_1$ gene encodes a protein of 547 amino acids with a calculated molecular weight of 57,946 and a pI of 4.81. The $groES_1$ and $groEL_1$ genes were expressed in *E. coli* and yielded high levels of recombinant protein (results not shown). There are 96 bp between the stop codon of $groES_1$ and the start codon of the $groEL_1$ gene. Three potential stem-loop structures were identified which may play a role in the regulation of transcription. A putative hairpin loop structure 5′ to the start site of the $groES_1$ gene has an estimated free energy of formation of −8.2 kcal (1 kcal = 4.184 kJ)/mol (Fig. 4A). This hairpin loop is conserved in many other bacterial heat shock genes. A putative hairpin or stem-loop structure, beginning 27 bp past the stop codon of $groEL_1$, with a calculated free energy of formation of −30.0 kcal/mol, is followed by another potential hairpin loop (free energy of formation of −17.2 kcal/mol). The two hairpin loop structures located 3′ to the $groEL_1$ gene may be involved in transcription termination, although no further studies have been performed to substantiate this.

The upstream *R. sphaeroides* hairpin sequence differs from other known conserved sequences in that it contains a G instead of an A at the third nucleotide on the 5′ branch of the hairpin (Fig. 4A). The G-T base pairing of the *R. sphaeroides* hairpin structure results in a 4-kcal/mol difference in the free energy of formation of this structure compared with bacterial hairpins that retain the conserved A-T base pairing. There are also conserved −10 and −35 upstream sequences, which are also found associated with several other bacterial *groESL* genes (Fig. 4B). Primer extension studies indicated that the transcript begins at the 5′ end of the conserved hairpin loop (Fig. 4C). A minor transcription start site is also seen 3′ to the hairpin loop. It is plausible that the upstream promoter region may have homology to both σ^{70} and σ^{32} consensus sequences (Fig. 4D).

Comparison of chaperonin amino acid sequences. The deduced amino acid sequence of *R. sphaeroides* cpn10 and cpn60 shows considerable similarity to sequences previously reported from other organisms. Alignment of cpn60 sequences from all divisions of eubacteria, as well as eucaryotic mitochondria and chloroplasts (data not shown), was in agreement with an alignment previously published (17), with *R. sphaeroides* cpn60 most closely related to cpn60s from other members of the alpha subdivision of purple bacteria. Finally, the deduced sequences of the two proteins of the *R. sphaeroides groEL* genes showed 72.5% identity.

Transcription of the *groESL* operon. Northern analysis of total RNA from photoheterotrophically cultured cells indicated that a 2.2-kb $groESL_1$ transcript was present that did not change in size after heat shock (cells were grown at 28°C and transferred to a 40°C illuminated water bath). However, levels of $groESL_1$-specific mRNA were enhanced 13-fold upon heat shock (Fig. 5), with maximum increases occurring between 15 and 30 min after heat shock. Changes in 16S rRNA transcripts were not substantial during heat shock; thus, this parameter provides an excellent internal standard in these experiments. The results indicated that transcript levels remained 10-fold elevated up to 120 min of heat shock; however, in other experiments there was a slight variance in $groESL_1$ transcripts from 60 to 120 min of heat shock. The $groESL_1$ operon was also induced by heat shock when cells were cultured under photoautotrophic and chemoheterotrophic growth conditions (data not shown). There was four times as much $groESL_1$ transcript in photoheterotrophically grown cells as in cells grown under photoautotrophic conditions (Fig. 6). The levels of 16S rRNA were very similar in these cells (data not shown). This response differs from the dramatic enhancement of *cbbL* transcription under photoautotrophic compared to photoheterotrophic growth conditions, as described here (Fig. 6) and in a previous report (13). Based on the amount of total RNA applied to gels, photoheterotrophically grown cells had about twice as much $groESL_1$ transcript as cells grown chemoheterotrophically; however, when standardized to experimentally determined 16S rRNA transcript levels, $groESL_1$ transcript levels were the same under both heterotrophic growth conditions. Both the size of the message and the experimentally determined start of transcription indicated that the $groES_1$ and $groEL_1$ genes were cotranscribed. In addition, the same transcript hybridized to either a $groEL_1$ or a $groES_1$ probe (results not shown). Finally, primer extension of RNA isolated from cells that had been heat shocked for 15 min exhibited exactly the same transcript start site (Fig. 4C), in agreement with Northern analyses that indicated no difference in the size of the transcript before and after heat shock.

Deletion analysis of *groESL* operons. Several attempts were made to delete the $groESL_1$ operon without success. Several matings were done at 30°C and at room temperature on both



B.

```
                                                                GTACCATCGTGAAGGCCGGTCCCGTTTCTTCTGGTGCATTCTGGGTCATCCTCG   (54)
GGGCCATCGTCATCCTGGTACTGATGATACTCGGCGGGTGGTGGGGGTCAGCCACCTGAACCCTTCTCCGGCGCCGACCTTCCTCCCGCGCTCAGTCCGAAGGTGTCGCATCATTGGGTC  (174)
GTCGATATTCCGGGGTTATCATCTCCCGGCAAGACAGGCGCCGGTGGGCGCAGCAAGCAGAGCATTTGACTGTCGGGCGACGGGATGTCTTCGCCTGGATCAGCTTTGCGCCCCACCTCG  (294)
GGCTTGTGCGCGCCGACGGCCGATCTGCTGGCTCCTGCCCCGCTCGCCCGATGTGGCGAGCCCAGATCGAAAATGGCCTCCGAGACTGCTTTGGATCAGCCCCGCATGGCGAGCTTTCTG  (414)
TCAGGAGTGAGTGTGCGCGCGACCCGGCTGCTGTCCTTGTCAGGGAGGCTCGGGAGCTAACCATGCGCGAACAAACGGCTACGGGCTCCTGTCGGACCGCGCGGGCTCCTCCTGTCTTCC  (534)

ATGGACGATAGGCCTGCTGCCGAGCCCTGGGCCGTGGCCGGTTCACCGCATTTGGGGCGAGGGCGAACGAGGGGCACTGTCGTGACGTCTGATACTCCAACCGCAAATCGAGTGAAGCTA  (654)
 M  D  D  R  P  A  A  E  P  W  A  V  A  G  S  P  H  L  G  R  G  R  T  R  G  T  V  V  T  S  D  T  P  T  A  N  R  V  K  L
  W  T  I  G  L  L  P  S  P  G  P  W  P  V  H  R  I  W  G  E  G  E  R  G  A  L  S  *  R  L  I  L  Q  P  Q  I  E  *  S  *

AGCTCGGGTCGACACCGTGGGATCGGCCGAAGACGTCAGGACGGTGAAGGCGTGGCGCCCGATGAGAAAGCGGGCGACCGTATCCTGTTTGAAAGATCGTCCGGCGTGCAGATCGAGTGC  (774)
 S  S  G  R  H  R  G  I  G  R  R  R  Q  D  G  E  G  V  A  P  D  E  K  A  G  D  R  I  L  F  E  R  S  S  G  V  Q  I  E  C
  A  R  V  D  T  V  G  S  A  E  D  V  R  T  V  K  A  W  R  P  M  R  K  R  A  T  V  S  C  L  K  D  R  P  A  C  R  S  S  A

GACCGGCGAACACCCTCTGATCATGAAGGAGGCCGACATTCTCGGCGTGATCGCCTGAGCCCGTAAACCTTCCCTGCAATCTCAGCCTCAGGAACCCGCGAATGTCCGCTAAAGAAGTTC  (894)
 D  R  R  T  P  S  D  H  E  G  G  R  H  S  R  R  D  R  L  S  P  *                                    M  S  A  K  E  V
  T  G  E  H  P  L  I  M  K  E  A  D  I  L  G  V  I  A  *

GATTCGGTACCGACGCCCGCGGCCGTATGCTGAAAGGCATCAACACGCTGGCCGATACGGTCAAGATCACCCTCGGCCCGAAGGGCAGAAATGTCATTCTCGACACGTCCTATGGCGCGC  (1014)
 R  F  G  T  D  A  R  G  R  M  L  K  G  I  N  T  L  A  D  T  V  K  I  T  L  G  P  K  G  R  N  V  I  L  D  T  S  Y  G  A

CCCGTATCACCAAGGACGGCGTGACCGTCGCCAGGGAGATCGAGTTGTCGGACCGTTTCGAGAACGTGGGCGCCCAGATGGTGAAGGAAGTGGCCAGCCGCACCAACGAGGAGGCCGGCG  (1134)
 P  R  I  T  K  D  G  V  T  V  A  R  E  I  E  L  S  D  R  F  E  N  V  G  A  Q  M  V  K  E  V  A  S  R  T  N  E  E  A  G

ACGGCACGACCACGGCAACGGTTCTGGCACAGGCCATTGCCAAGGAAGGCATGAAGGCGGTCGCGGCCGGCATGAATCCGATGGATCTGAAGCGGGGGATCGACCGCGCGGTGGCCATCG  (1254)
 D  G  T  T  T  A  T  V  L  A  Q  A  I  A  K  E  G  M  K  A  V  A  A  G  M  N  P  M  D  L  K  R  G  I  D  R  A  V  A  I

TCATAGCCGAGATCAGGTCCATGTCCCGACCGGTCGGAGACAGCGCCGAAATTGCGAAGGTCGGGGCCCTATCGGCCAATGGCGAGGCGGCGATCGGCCGTCAGATCGCCCATGCAATGG  (1374)
 V  I  A  E  I  R  S  M  S  R  P  V  G  D  S  A  E  I  A  K  V  G  A  L  S  A  N  G  E  A  A  I  G  R  Q  I  A  D  A  M

CGAAGGTCGGCACTGCGGGGGTGATCAAGGTAGAGAATAAGGGCCTCGAGACCGAAACCGAGGTGGTGGAAGGCATGCAGTTCGACCGCGGTTACCTGAGCCCCTATTTCATCACTCATG  (1494)
 A  K  V  G  T  A  G  V  I  K  V  E  N  K  G  L  E  T  E  T  E  V  V  E  G  M  Q  F  D  R  G  Y  L  S  P  Y  F  I  T  H

CGCAGAAGATGGTCGTGGAATTGGACGACTGCGCCATCCTGCTCCACGAAGGCAAGCTGACCTCGCTTGCTTCGATGGTGCCGTTGCTGGAAGCGGTCGTTCAGGCAGAAAAGCAACTTC  (1614)
 A  Q  K  M  V  V  E  L  D  D  C  A  I  L  L  H  E  G  K  L  T  S  L  A  S  M  V  P  L  L  E  A  V  V  Q  A  E  K  Q  L

TCGTCGTGGCCGAGGACGTCGAGGGCGAGGCTCTGACGACCCTGGTGGTCAACAAGCTGCGCGGCGGCGGTCTCAAGGTTGCGGCAGCAAAGGCACCCGGCTTCGGGGATGGGCGTGCGGCGA  (1734)
 L  V  V  A  E  D  V  E  G  E  A  L  T  T  L  V  V  N  K  L  R  G  G  L  K  V  A  A  A  K  A  P  G  F  G  D  G  R  A  A

TGCTGCAGGATCTTGCAGTCCTGACGGGCGCGCACCTCATCTCGGCCGAGCTCGGGACGAAGCTGGAAACCGTCACGCTGGACATGCTCGGTTTTGCGAAGAAGGTTGTCCTGACGAAGG  (1854)
 M  L  E  D  L  A  V  L  T  G  A  H  L  I  S  A  E  L  G  T  K  L  E  T  V  T  L  D  M  L  G  F  A  K  K  V  V  L  T  K

ACAGCACCATCCTCATCGACAGTGCCGGCGACAAGGCGGCGATCGCTTCACGCATCGGCCAGATCCGCAACCAGATCGAGGATACGACGTCGGCCTACAACAAGGAGAAGTTGCAGGAAC  (1974)
 D  S  T  I  L  I  D  S  A  G  D  K  A  A  I  A  S  R  I  G  Q  I  R  N  Q  I  E  D  T  T  S  A  Y  N  K  E  K  L  Q  E

GTCTGGCGAGACTCGCGGGCCGGAGTTGCTGTCATTCGGGTCGGGGGTGCCACCGAAATCGAAGTGAAGGAGCGCAGGGACAGGGTGGAGGACACCCTCAACGCCACCCCCGCCCCCGGTAC  (2094)
 R  L  A  R  L  A  G  G  V  A  V  I  R  V  G  G  A  T  E  I  E  V  K  E  R  R  D  R  V  E  D  T  L  N  A  T  R  A  A  V

AGGAGGGCGTCGTGCCCGGCGGCGGGGCCGCACTGATTCATGCGGGCAAGGCTCTGGCGGGCTTGAAGGGCGACAATCCCGATCAGGATGCAGGCATCAAGATCATCCGCCCCGCAATCC  (2214)
 Q  E  G  V  V  P  G  G  G  A  A  L  I  H  A  G  K  A  L  A  G  L  K  G  D  N  P  D  Q  D  A  G  I  K  I  I  R  R  A  I

AGGCGCCGCTGCGCCAGATCGCCGACAACGCCGGGATCGACGGATCGGTCGTCGCCGGCAAGGTCATCGAGAACGACAGCGCAACCTTTGGCTTCGATGCGCAGTTGGAGACGTACGGCG  (2334)
 Q  A  P  L  R  Q  I  A  D  N  A  G  I  D  G  S  V  V  A  G  K  V  I  E  N  D  S  A  T  F  G  F  D  A  Q  L  E  T  Y  G

ATATGCTGCAGGCCGGCATCATCGATCCGACCAAGGTCGTGCGGATAGCATTGGAGGACGCCGCATCGATCGGCAGGCCTGCTGATCACGACGGAGGTGATCATCGCGCATAAGCCGGAAA  (2454)
 D  M  L  Q  A  G  I  I  D  P  T  K  V  V  R  I  A  L  E  D  A  A  S  I  A  G  L  L  I  T  T  E  V  I  I  A  H  K  P  E

GGGCGAACGAATGTCCCAGATGGATGAGATGGGCGGCATGATGTGAGGTTCGCCCGCTGCGTCGATCCCTGATGTCTGCATTTGCACTCAGGCAATCCGCTCCTCCCCGAACTGATCGC  (2574)
 R  G  E  R  M  S  Q  M  D  E  M  G  G  M  M  *

CAGAATCGCGTTACAGTCCCACCAAGACATAATCCGAAGGAAAAGCACGATGTCCAAAGGCGACAAGACACGCGGCAATCGCGAAGCTAAGAAGCCCAAGAAAATCAAGGAAAAGGTGGT  (2694)
GGAAACGGCAAGCTACACCAAGGGAGAGACGCCGCTGACCATCGCGGGCAAGAAGAGCCCCCGCGTCAAGAGGTAACTCGGGACCACTCTGGCGCGCTAGAATGGCTCGGGCGTGTCGGC  (2814)
GCAAGCGCTACTTGTCCGCGGCACATTTCAGGCGGCAGCCGTCGAAAGGAGGACTGTGCTGTGACCTATGATCACTTCTTCCTTCCAACAGTCTTGGAAGCTCTAGTCGTGCCATGATGA  (2934)
CGCCCACGGAAGCTGGCCACGCCGCGCGCTTTGACGCCGGGTTCAAACGCCTGTCGCAGCGCGACGACATGAAACGGCGGCTGACGAGCAATCTCCGGCATCGTTGGTCTCGGCACGCCA  (3054)
ACGCAGTTGGCCCCGCTGCTTAAAATGATGTGGTGCGGAAGGCAGATCCGGATCCGACACCAGGACATGTCTGAAGGAGGTACC                                    (3138)
```

FIG. 3. Restriction map and nucleotide sequence of the *R. sphaeroides groESL*$_2$ operon. (A) Restriction map of the *groESL*$_2$ operon. Abbreviations are the same as in Fig. 1A and as follows: K, *Kpn*I, Sl, *Sal*I. (B) Sequence of the *groESL*$_2$ operon. The deduced amino acid sequences of the putative cpn10 and cpn60 proteins are listed below the DNA sequence. The *groESL*$_2$ gene has homology to the *groES*$_1$ gene product in two of its reading frames. The regions with homology to the *groES*$_1$ gene product are underlined.

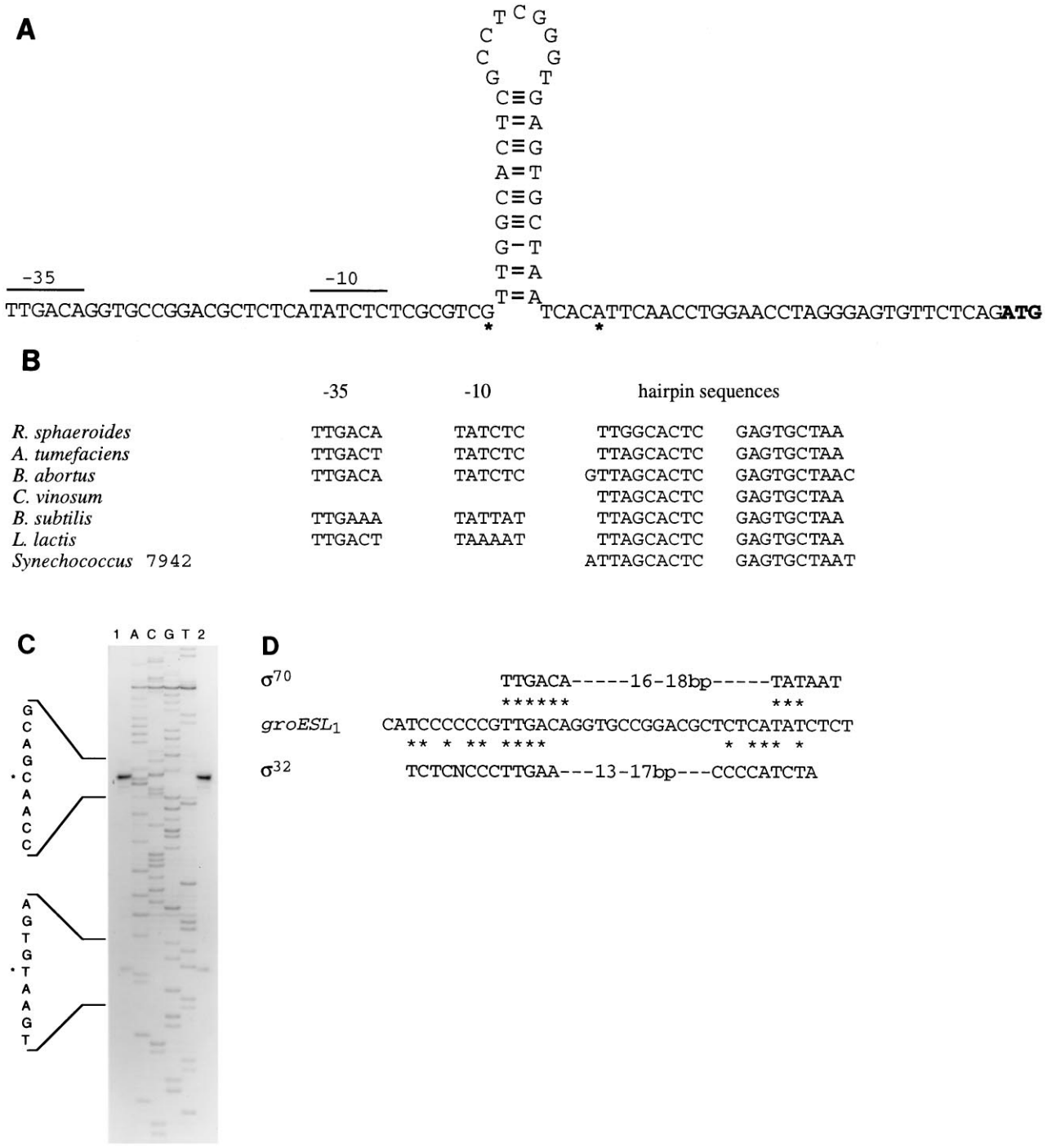


FIG. 4. Initiation of transcription of the *groESL*$_1$ operon at the 5′ end of a conserved hairpin loop. (A) Proposed structure of putative hairpin loop located upstream of *groES*$_1$. The asterisks show the start site of transcription as determined by primer extension of total RNA isolated from cells grown photoheterotrophically, and the start codon of *groES* is in boldface. (B) Conservation of the hairpin loop sequence and −10 and −35 sequences in the diverse eubacteria *A. tumefaciens* (35), *Brucella abortus* (16), *Chromatium vinosum* (9), *B. subtilis* (25, 34), *Lactobacillus lactis* (22), and *Synechococcus* sp. strain PCC 7942 (45). (C) Primer extension of the *groESL*$_1$ operon from total cellular RNA of *R. sphaeroides*. The major transcription start site is at the G site at the 5′ end of the hairpin loop; a minor transcription start site (A) is located 5 bp from the loop. (D) Homology of the putative promoter sequence to the *groESL*$_1$ operon to conserved portions of both the *E. coli* σ^{70} and σ^{32} consensus sequences (48). Identical residues are marked by asterisks.

rich and minimal media, and no colonies containing a double recombination knockout of the *groESL*$_1$ genes were isolated. More than 2,000 colonies were screened in an attempt to isolate a deletion of the *groESL*$_1$ operon. These results may indicate that the *groESL*$_1$ genes are essential for growth in *R. sphaeroides*. On the other hand, the *groESL*$_2$ genes were successfully deleted and the constructed strain grew identically to the parental strain under photoautotrophic, photoheterotro-

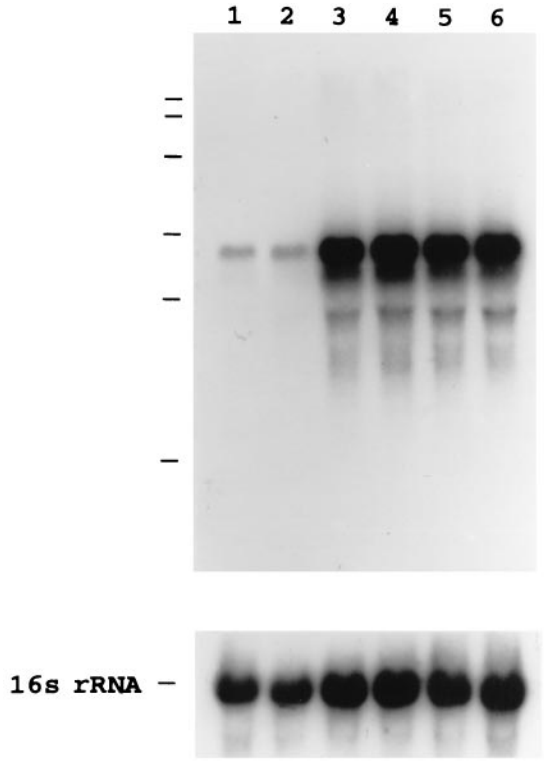


FIG. 5. Induction of the $groESL_1$ operon following heat shock at 40°C. Total cellular RNA (20 μg per lane) was separated by using a 1.5% agarose gel and blotted onto nylon membranes by standard protocols. The autoradiogram at the top represents a Northern blot probed with a 1.0-kb *Sph*I-*Sma*I $groEL_1$ fragment. RNA from cells harvested before heat shock (28°C) (lane 1) and RNAs from cells harvested at 5, 15, 30, 60, and 120 min, respectively, after heat shock at 40°C (lanes 2 to 6) were analyzed. Only the $groESL_1$ transcript was detected with this probe. At the bottom is the same blot rehybridized with an *R. sphaeroides* 16S rRNA probe. The molecular size standards are (from the top) 9.49, 7.46, 4.40, 2.37, 1.35, and 0.24 kb.

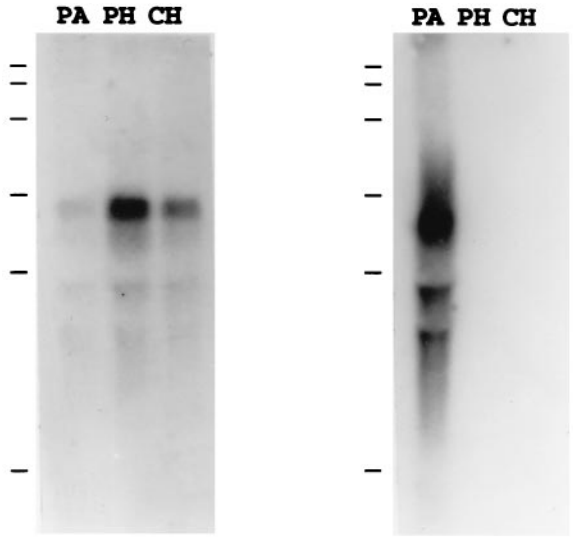


FIG. 6. Transcript levels under different growth conditions. Total cellular RNAs were isolated from *R. sphaeroides* grown photoautotrophically (PA), photoheterotrophically (PH), and chemoheterotrophically (CH). The RNA (approximately 20 μg per lane) was separated by using a 1.5% agarose gel, blotted onto a nylon membrane, and hybridized to a $groEL_1$ probe (left) and a $cbbL_1$ probe (right). The standards are the same as in Fig. 5.

phic, and chemoheterotrophic conditions. It was also capable of fixing nitrogen, as determined by growing strains in nitrogen-free growth media.

DISCUSSION

Nonsulfur purple bacteria catalyze several important metabolic processes, including nitrogen fixation (23, 26), hydrogen uptake (41), and autotrophic carbon dioxide assimilation (37). In each case, large quantities of biosynthetic energy must be expended to meet the requirement for the substantial levels of proteins and enzymes that are specific to these processes. These organisms also synthesize a copious intracellular membrane network, which houses the photosynthetic apparatus, in response to the levels of oxygen present in the culture (4). In all of the above scenarios, proper folding and subsequent assembly of large amounts of newly synthesized protein are required. In view of the well-established role of chaperonin proteins in assisting the folding of unfolded polypeptides after they are synthesized (49), studies on the role of such proteins in *R. sphaeroides* were initiated. In particular, we are interested in how these proteins influence carbon metabolism in this organism. The cloning of two *groESL* operons of *R. sphaeroides* reported here is thus a step towards understanding the role of chaperonin proteins in nonsulfur purple bacteria. In addition to their role in protein folding, chaperonins are also known to be involved in the heat shock response, as well as other stress-related processes (49). Presumably, the interaction with unfolded proteins is maximized during times of stress; thus, the requirement for up-regulating chaperonin protein synthesis seems apparent. RubisCO, which has a central role in the metabolism of *R. sphaeroides* (37), has been used as a model protein to gain an understanding of how chaperonin proteins influence or assist the folding pathway (14, 15, 39). In addition, the presence of multiple cpn60/cpn10 chaperonins in different bacteria, including two of the alpha subdivision of gram-negative bacteria, provides a rationale for considering the potential of specialized chaperonins interacting with different proteins. In *R. sphaeroides*, it has already been demonstrated that the level of at least one cpn60 is related to growth conditions that favor enhanced RubisCO synthesis (38). Thus, it is possible that a major function of cpn60 in *R. sphaeroides* is to facilitate the folding and assembly of the large quantities of RubisCO and other proteins involved in carbon fixation and subsequent metabolism. Indeed, a physical association of cpn60 and form I RubisCO of *R. sphaeroides* has recently been described (42) and the overexpression of both the $groEL_1$ and $groES_1$ genes in *E. coli* to yield copious quantities of recombinant cpn60 and cpn10 should greatly facilitate in vitro studies.

The $groESL_1$ genes described here represent a typical bacterial heat shock *groESL* operon. Like many bacterial heat shock *groESL* genes, $groES_1$ and $groEL_1$ are cotranscribed and possess a conserved upstream hairpin loop structure, as well as conserved upstream −10 and −35 sequences. The deduced amino acid sequences of cpn60 and cpn10 of *R. sphaeroides* agree with N-terminal sequences derived from purified proteins described earlier (38, 43) and are most closely related to similar sequences obtained from bacteria in the alpha subdivision. Within this subdivision are *B. japonicum* and *R. meliloti*, both of which possess multiple *groESL* operons. Other *groESL* genes from this subdivision have been cloned from *Brucella abortus* (16) and *Agrobacterium tumefaciens* (35). In these organisms, only a single heat shock *groESL* operon was cloned after screening with *E. coli groESL* probes. Our initial library screen with the *B. japonicum* $groESL_2$ probe also identified

only a single *groESL* gene. However, when genomic DNA was rescreened with a probe prepared from *R. sphaeroides* $groESL_1$, a second *groEL* gene was identified. Transcripts to the second *groEL* were not seen, even after heat shock. It also appears the $groES_2$ gene is a pseudogene.

The transcription start of the first *groESL* operon was determined by primer extension. Much like that of *A. tumefaciens* (35) and *Clostridium acetobutylicum* (28), transcription begins at the 5′ end of the conserved hairpin loop. The putative −35 and −10 sequences present upstream of the *R. sphaeroides* *groESL* genes are conserved in 5′ sequences of heat shock genes from several bacteria. The hairpin loop, which also seems to be highly conserved and has been referred to as CIRCE, for controlling inverted repeat of chaperone expression (50), has been postulated to be important in the heat shock response; however, its precise function has not been determined. In this context, the CIRCE of *Bacillus subtilis* appears to negatively regulate transcription under nonstress conditions (50). Another study further demonstrated that CIRCE is involved in transcript turnover, as transcripts which lacked the hairpin loop were much more stable (47). A recent paper indicated that the *groESL* operon was still induced by heat shock, even after most of the CIRCE was deleted (36). There are CIRCE-like sequences in three of the *B. japonicum* *groESL* operons; however, only one of these ($groESL_4$) is induced by heat shock (1). The CIRCE sequence has no apparent function for the other two operons. In addition, the $groESL_1$ operon of *B. japonicum* is induced by heat shock but lacks the CIRCE sequence (1). It should also be noted that the role of the conserved −10 and −35 sequences in the regulation of the heat-induced *groESL* operons containing the CIRCE sequence has not been determined. Analysis of the $groESL_1$ upstream region of *R. sphaeroides* indicates that it is plausible that this operon is under the control of two RNA polymerases, one of which (σ^{70}) is active under normal growth conditions and another (σ^{32}) transcribes the operon during heat shock or other stress (Fig. 5D). The determination of the role of this putative promoter sequence would be of great interest if there were indeed complex regulation of this operon by different RNA polymerases.

While chaperonins have been shown to be involved in assisting the folding and, perhaps, subsequent assembly of many proteins, the interaction of cpn60 and cpn10 with RubisCO has been most extensively studied (39). Indeed, in *R. sphaeroides*, the levels of cpn60 were greatest under growth conditions in which RubisCO synthesis was also maximized (38). Thus, the unexpectedly lower transcription of $groESL_1$ under photoautotrophic growth conditions might indicate that the products of $groEL_2$ (despite our inability to demonstrate a specific transcript), or other undiscovered chaperonin proteins, are maximally synthesized when CO_2 is the sole source of carbon. The lower level of transcription of the $groESL_1$ operon may indicate that the level of cpn60 protein is regulated posttranscriptionally under photoautotrophic conditions since Southern blotting experiments indicated that there were no other *groE* operons. In addition, the N-terminal sequence of cpn60 isolated from two different laboratories (38, 43) agrees with the sequence of the $groEL_1$ gene product reported here. Until a specific $groESL_2$ transcript or specific $cpn60_2$ protein can be demonstrated, the results presented here indicate that only the $groESL_1$ operon is expressed in *R. sphaeroides*, suggesting that the products of these genes play an important role in the folding and assembly of RubisCO and other proteins in this organism. Future studies will be directed at biochemical and genetic investigations of the role of each cpn60 in RubisCO folding, spurred by the isolation of a cpn60-form I RubisCO complex from extracts of *R. sphaeroides* (42). The recent demonstration that the intracellular levels of chaperonins influence the synthesis of proteins involved in carbon metabolism in *E. coli* (21) suggests that the correlation between RubisCO and cpn60 synthesis (38) may be of great importance in *R. sphaeroides.*

ACKNOWLEDGMENTS

We thank H. M. Fischer for the *B. japonicum* $groESL_2$ gene and Charles Daniels for use of the Packard Instant Imager. We also thank the reviewer who suggested the homology of the $groESL_1$ promoter and consensus sequences for both σ^{32} and σ^{70} from *E. coli.*

This work was supported by Public Health Service grant GM24497 from the National Institutes of Health.

REFERENCES

1. **Babst, M., H. Hennecke, and H.-M. Fisher.** 1996. Two different mechanisms are involved in the heat-shock regulation of chaperonin gene expression in *Bradyrhizobium japonicum*. Mol. Microbiol. **19:**827–839.
2. **Braig, K., Z. Otwinowski, R. Hedge, D. C. Boisvert, A. Joachimiak, A. R. Horwich, and P. B. Sigler.** 1994. The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. Nature **371:**578–586.
3. **Chandrasekhar, G. N., K. Tilly, C. Woolford, R. Hendrix, and C. Georgopoulos.** 1986. Purification and properties of the groES morphogenetic protein of *Escherichia coli*. J. Biol. Chem. **261:**12414–12419.
4. **Drews, G., and J. R. Golecki.** 1995. Structure, molecular organization, and biosynthesis of membranes of purple bacteria, p. 231–257. *In* R. E. Blankenship, M. T. Madigan, and C. E. Bauer (ed.), Anoxygenic photosynthetic bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
5. **Dryden, S. C., and S. Kaplan.** 1990. Localization and structural analysis of the ribosomal RNA operons of *Rhodobacter sphaeroides*. Nucleic Acids Res. **18:**7267–7277.
6. **Ellis, R. J.** 1994. Chaperoning nascent proteins. Nature **370:**96–97.
7. **Falcone, D. L., and F. R. Tabita.** 1991. Expression of endogenous and foreign ribulose 1,5-bisphosphate carboxylase-oxygenase (RubisCO) genes in a RubisCO deletion mutant of *Rhodobacter sphaeroides*. J. Bacteriol. **173:** 2099–2108.
8. **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.
9. **Ferreyra, R. G., F. C. Soncini, and A. M. Viale.** 1993. Cloning, characterization, and functional expression in *Escherichia coli* of chaperonin (*groESL*) genes from the phototrophic sulfur bacterium *Chromatium vinosum*. J. Bacteriol. **175:**1514–1523.
10. **Fischer, H. M., M. Babst, T. Kaspar, G. Acuna, F. Arigoni, and H. Hennecke.** 1993. One member of the *groELS*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes. EMBO J. **12:**2901–2912.
11. **Georgopoulos, C. P., R. W. Hendrix, S. R. Casjens, and A. D. Kaiser.** 1973. Host participation in bacteriophage lambda head assembly. J. Mol. Biol. **76:**45–60.
12. **Gibson, J. L.** 1995. Genetic analysis of CO_2 fixation genes, p. 1107–1124. *In* R. E. Blankenship, M. T. Madigan, and C. E. Bauer (ed.), Anoxygenic photosynthetic bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
13. **Gibson, J. L., D. L. Falcone, and F. R. Tabita.** 1991. Nucleotide sequence, transcriptional analysis, and expression of genes encoded within the form I CO_2 fixation operon of *Rhodobacter sphaeroides*. J. Biol. Chem. **266:**14646–14653.
14. **Goloubinoff, P., A. A. Gatenby, and G. H. Lorimer.** 1989. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. Nature **337:**44–47.
15. **Goloubinoff, P., J. T. Christeller, A. A. Gatenby, and G. H. Lorimer.** 1989. Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. Nature **342:**884–889.
16. **Gor, D., and J. E. Mayfield.** 1992. Cloning and nucleotide sequence of the *Brucella abortus groE* operon. Biochim. Biophys. Acta **1130:**120–122.
17. **Gupta, R. S.** 1995. Evolution of the chaperonin families (Hsp60, Hsp10 and Tcp-1) of proteins and the origin of eukaryotic cells. Mol. Microbiol. **15:**1–11.
18. **Hemmingsen, S. M., C. Woolford, S. M. van der Vies, K. Tilly, D. T. Dennis, C. P. Georgopoulos, R. W. Hendrix, and R. J. Ellis.** 1988. Homologous plant and bacterial proteins chaperone oligomeric protein assembly. Nature **333:** 330–334.
19. **Hendrix, R. W.** 1979. Purification and properties of groE, a host protein involved in bacteriophage assembly. J. Mol. Biol. **129:**375–392.
20. **Ishihama, A., T. Ikeuchi, A. Matsumoto, and S. Yamamoto.** 1976. A novel

adenosine triphosphatase isolated from RNA polymerase preparations from *Escherichia coli*. II. Enzymatic properties and molecular structure. J. Biochem. **79:**927–936.
21. **Kanemori, M., H. Mori, and T. Yura.** 1994. Effects of reduced levels of GroE chaperones on protein metabolism: enhanced synthesis of heat shock proteins during steady-state growth of *Escherichia coli*. J. Bacteriol. **176:**4235–4242.
22. **Kim, S. G., and C. A. Batt.** 1993. Cloning and sequencing of the *Lactococcus lactis* subsp. *lactis groESL* operon. Gene **127:**121–126.
23. **Kranz, R. G., and P. J. Cullen.** 1995. Regulation of nitrogen fixation genes, p. 1191–1208. *In* R. E. Blankenship, M. T. Madigan, and C. E. Bauer (ed.), Anoxygenic photosynthetic bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
24. **Lee, W. T., K. C. Terlesky, and F. R. Tabita.** 1995. Heat-induced and non-heat-induced *groESL* operons from *Rhodobacter sphaeroides*, abstr. K143, p. 561. *In* Abstracts of the 95th General Meeting of the American Society for Microbiology 1995. American Society for Microbiology, Washington, D.C.
25. **Li, M., and S.-L. Wong.** 1992. Cloning and characterization of the *groESL* operon from *Bacillus subtilis*. J. Bacteriol. **174:**3981–3992.
26. **Ludden, P. W., and G. P. Roberts.** 1995. The biochemistry and genetics of nitrogen fixation by photosynthetic bacteria, p. 929–947. *In* R. E. Blankenship, M. T. Madigan, and C. E. Bauer (ed.), Anoxygenic photosynthetic bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
27. **Martin, J., M. Mayhew, T. Langer, and F. U. Hartl.** 1993. The reaction cycle of GroEL and GroES in chaperonin-assisted protein folding. Nature **366:**228–233.
28. **Narberhaus, F., and H. Bahl.** 1992. Cloning, sequencing, and molecular analysis of the *groESL* operon of *Clostridium acetobutylicum*. J. Bacteriol. **174:**3282–3289.
29. **Penfold, R. J., and J. M. Pemberton.** 1992. An improved suicide vector for construction of chromosomal insertion mutations in bacteria. Gene **118:**145–146.
30. **Pridemore, R. D.** 1987. New and versatile cloning vectors with kanamycin-resistance marker. Gene **56:**309–312.
31. **Rusanganwa, E., and R. S. Gupta.** 1993. Cloning and characterization of multiple *groEL* chaperonin-encoding genes in *Rhizobium meliloti*. Gene **126:**67–75.
32. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
34. **Schmidt, A., M. Schiesswohl, U. Völker, M. Hecker, and W. Schumann.** 1992. Cloning, sequencing, mapping, and transcriptional analysis of the *groESL* operon from *Bacillus subtilis*. J. Bacteriol. **174:**3993–3999.
35. **Segal, G., and E. Z. Ron.** 1993. Heat shock transcription of *groESL* operon of *Agrobacterium tumefaciens* may involve a hairpin-loop structure. J. Bacteriol. **175:**3083–3088.
36. **Segal, G., and E. Z. Ron.** 1996. Heat shock activation of the *groESL* operon of *Agrobacterium tumefaciens* and the regulatory roles of the inverted repeat. J. Bacteriol. **178:**3634–3640.
37. **Tabita, F. R.** 1995. The biochemistry and metabolic regulation of carbon metabolism and CO_2 fixation in purple bacteria, p. 885–914. *In* R. E. Blankenship, M. T. Madigan, and C. E. Bauer (ed.), Anoxygenic photosynthetic bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
38. **Terlesky, K. C., and F. R. Tabita.** 1991. Purification and characterization of the chaperonin 10 and chaperonin 60 proteins from *Rhodobacter sphaeroides*. Biochemistry **30:**8181–8186.
39. **Todd, M. J., P. V. Viitanen, and G. H. Lorimer.** 1994. Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. Science **265:**659–666.
40. **Triezenberg, S. J.** 1994. Primer extension, p. 4.8.1–4.8.5. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. John Wiley & Sons, Inc., New York, N.Y.
41. **Vignais, P. M., B. Toussaint, and A. Colbeau.** 1995. Regulation of hydrogenase gene expression, p. 1175–1190. *In* R. E. Blankenship, M. T. Madigan, and C. E. Bauer (ed.), Anoxygenic photosynthetic bacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
42. **Wang, X., and F. R. Tabita.** 1992. Interaction of inactivated and active ribulose 1,5-bisphosphate carboxylase/oxygenase of *Rhodobacter sphaeroides* with nucleotides and the chaperonin 60 (GroEL) protein. J. Bacteriol. **174:**3607–3611.
43. **Watson, G. M. F., N. H. Mann, G. A. MacDonald, and B. Dunbar.** 1990. Identification and characterization of the GroEL homologue in *Rhodobacter sphaeroides*. FEMS Microbiol. Lett. **72:**349–354.
44. **Weaver, K. E., and F. R. Tabita.** 1983. Isolation and partial characterization of *Rhodopseudomonas sphaeroides* mutants defective in the regulation of ribulose bisphosphate carboxylase/oxygenase. J. Bacteriol. **156:**507–515.
45. **Webb, R., K. J. Reddy, and L. A. Sherman.** 1990. Regulation and sequence of the *Synechococcus* sp. strain PCC 7942 *groESL* operon, encoding a cyanobacterial chaperonin. J. Bacteriol. **172:**5079–5088.
46. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
47. **Yuan, G., and S.-L. Wong.** 1995. Regulation of *groE* expression in *Bacillus subtilis*: the involvement of the σ^A-like promoter and the roles of the inverted repeat sequence (CIRCE). J. Bacteriol. **177:**5427–5433.
48. **Yura, T., N. Nagai, and H. Mori.** 1993. Regulation of the heat-shock response in bacteria. Annu. Rev. Microbiol. **47:**321–350.
49. **Zeilstra-Ryalls, J., O. Fayet, and C. Georgopoulos.** 1991. The universally conserved GroE (Hsp60) chaperonins. Annu. Rev. Microbiol. **45:**301–325.
50. **Zuber, U., and W. Schumann.** 1994. CIRCE, a novel heat shock element involved in regulation of heat shock operon *dnaK* of *Bacillus subtilis*. J. Bacteriol. **176:**1359–1363.