JOURNAL OF BACTERIOLOGY, Nov. 1991, p. 7382–7386
0021-9193/91/227382-05$02.00/0

Vol. 173, No. 22

NOTES

Characterization of the *groEL*-Like Genes in *Streptomyces albus*

P. MAZODIER, G. GUGLIELMI,* J. DAVIES, AND C. J. THOMPSON

Unité de Génie Microbiologique, Institut Pasteur, Paris 75015, France

Received 3 June 1991/Accepted 11 September 1991

Three GroEL-like heat shock proteins (HSP56, HSP58, and HSP18) have been observed in *Streptomyces albus* (G. Guglielmi, P. Mazodier, C. J. Thompson, and J. Davies, J. Bacteriol. 173:7374–7381, 1991). Here we report the cloning and complete nucleotide sequence of *groEL1*, which encodes HSP18 and HSP58, and *groEL2*, which encodes HSP56. Both nucleotide sequences predicted proteins of 56,680 Da that were 70% identical. The 5′ nucleotide sequence of *groEL1* coded for a protein corresponding to HSP18 that may be a processed gene product. At least two *groEL*-like genes were present in all 12 *Streptomyces* species tested; they were not closely linked in the genome. *groEL1*, but not *groEL2*, was adjacent to a *groES*-like gene.

Pleiotropic genetic phenotypes and biochemical properties have demonstrated that the heat shock proteins of the GroEL family have physiologically diverse functions (for reviews, see references 13, 17, and 20). GroEL was originally named for the role it plays in bacteriophage virion assembly (7, 28), although its more general function was suggested by the fact that it is essential for viability of *Escherichia coli* at all temperatures (5) and that overproduction of GroEL suppresses defects in various enzymes (24). These apparently disparate phenotypes are presumed to be indirect and a reflection of the primary activity of GroEL in stabilizing intermediates in protein folding pathways that may lead to the export or assembly of enzyme complexes (12). Apparently, GroEL cannot provide this function for all proteins; SecB, DnaK, and the trigger factor seem to play similar roles but have different substrate specificities (1, 12). In spite of these diverse functions, which overlap with those of other proteins controlling protein renaturation, only one *groEL* gene has been reported so far in the numerous eubacteria that have been studied. Our observation of three GroEL-like proteins (HSP18, HSP56, and HSP58) in *Streptomyces albus* (8) implied that there are multiple copies of *groEL*-like genes in this organism. Although the amino acid sequence of the N-terminal region of HSP18 was similar to that of GroEL from various bacteria, its size (18 kDa) and basic charge were very different from those of all other GroEL proteins described to date, which, like HSP56 and HSP58, are acidic and have apparent molecular masses ranging from 55 to 65 kDa.

Cloning of two *S. albus groEL*-like genes. *Xho*I fragments (1.2 and 1.8 kb) that contained different *S. albus groEL*-like genes (8) were partially purified from genomic DNA by agarose gel electrophoresis and cloned into the *Sal*I site of pUC19 (25). Plasmid pGM1, containing a 1.8-kb insert, was isolated by screening with a *groEL1*-specific probe (OL1 [8]); pGM2, containing a 1.2 kb insert, was isolated by screening with a probe corresponding to the 5′ region of the *Mycobacterium leprae groEL* gene (15). DNA inserts in pGM1 and pGM2 were then used as probes to screen a genomic bank generated with cosmid pHC79 (11). Recombinant cosmids were obtained that hybridized to the 1.8-kb (pGM3) or 1.2-kb (pGM4) *Xho*I fragment. The fragments of pGM3 and pGM4 that corresponded to the region within and adjacent to the *Xho*I fragments were determined by Southern blot hybridization and subcloned into M13 for dideoxy sequencing reactions (19).

Sequence of *groEL1*. The nucleotide sequence of the region of pGM3 that corresponded to OL1, HSP18, and HSP58 is shown in Fig. 1. Codon usage analysis (data not shown) predicted a functional open reading frame (*groEL1*) that encoded an acidic protein (a pI of 4.60 was predicted by using the University of Wisconsin sequence analysis program Isoelectric) of 56,681 Da whose N-terminal amino acids corresponded to Edman degradation products of the N terminus of HSP58 (10 residues analyzed) and N-terminal and internal V8-generated peptides of HSP18 (48 of 49 residues analyzed). Although these amino acid sequence identities were unexpected, they provided strong evidence that HSP18 was the product of the *groEL1* gene. The hybridization data cited above (8) indicated that *S. albus* has only two *groEL*-like genes; nucleotide sequence analysis described below showed that the second gene did not encode HSP18. Nevertheless, the nucleotide sequence of *groEL1* (which has been determined from two independent clones, pGM1 and pGM3) had no in-frame stop codon that could give rise to an 18-kDa peptide. Since the predicted N-terminal 21-kDa polypeptide of GroEL (HSP18 migrates as a 21-kDa band under some denaturing conditions [8]) is acidic (estimated from the first 201 amino acids of *groEL1*; pI 4.43), the basic charge of HSP18 (8) also cannot be explained directly from the primary sequence of *groEL1*.

The synthesis of a truncated basic protein (HSP18) could be explained by ribosomal frameshifting at a site within *groEL1*. Frameshifting can generate unexpected alternative gene products (2) that are sometimes functional. For example (4, 6, 23), in the *E. coli dnaX* gene that encodes DNA

* Corresponding author.

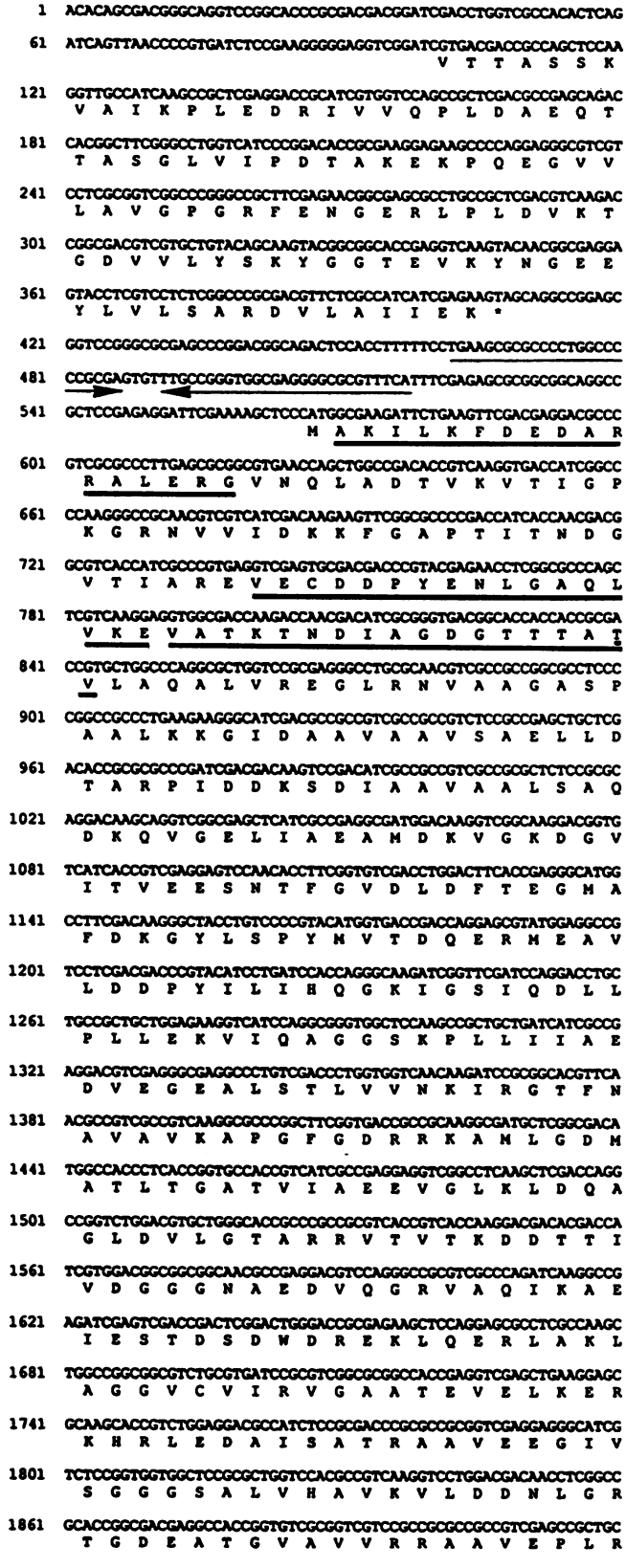

FIG. 1. Nucleotide sequence of *groEL1* and *groES*. Arrows indicate imperfect inverted repeat sequences. Underlining indicates experimental amino acid sequence data (8). The double-underlined T (threonine) residue indicates the only inconsistency with the experimentally determined peptide sequence (P2) (8).

polymerase III, the in-frame translational stop signal encodes a 71-kDa polymerase subunit. A −1 translational frameshift occurs in 50% of the transcripts at a site within the gene; as a result, ribosomes terminate early in the new frame to yield a 52-kDa protein. Both the 52- and 71-kDa products are thought to have physiological roles. In the case of *groEL1*, which can be translated as a full-sized protein of 58 kDa (8), a +2 or −1 frameshift would generate a stop codon to produce a 21-kDa basic protein. Alternatively, posttranslational cleavage and modification of GroEL1 are possible. We are carrying out C-terminal amino acid analysis of HSP18 to distinguish between these possibilities.

Codon usage analyses predict that the region upstream of *groEL1* contains a functional open reading frame encoding a 10,948-Da protein that is homologous to GroES found in a variety of bacteria, including *E. coli*, *Coxiella* spp., *Legionella* spp., cyanobacteria, *Mycobacterium tuberculosis*, and *M. bovis* (data not shown). The *groES*-like gene is separated from the *groEL*-like gene by a long inverted repeat sequence that contains a motif (TTTCAT[16 N]TTGCCG) that is also present upstream of the *M. tuberculosis*, *M. bovis*, and *M. leprae groEL*-like genes (3).

Sequence of *groEL2*. The nucleotide sequence of the region of pGM4 that hybridized to the *M. leprae* probe (Fig. 2) showed that it was different from *groEL1*. Codon usage analysis (data not shown) predicted a functional open reading frame that encoded an acidic protein (predicted pI, 4.61; GroEL2) of 56,682 Da whose N-terminal amino acids corresponded to the 10 N-terminal residues of HSP56 (8). Inspection of the sequence within 200 bp of *groEL2* did not reveal a *groES*-like gene like that near *groEL1*.

Similarities between GroEL1, GroEL2, and eubacterial GroEL proteins. Although GroEL1 and GroEL2 were 70% identical, there were significant differences. GroEL1 did not contain the Gly-Gly-Met motif found near the C terminus of these procaryotic proteins and in all reported eucaryotic HSP56-58 proteins (with the exception of chloroplast rubisco-binding protein). Instead, the C terminus of GroEL1 contained the motif Gly-His-Gly-His-Gly-His-Ser-His.

A comparison of GroEL1, GroEL2 and other eubacterial GroEL proteins is shown in Fig. 3. With this alignment, the percent amino acid identity for all pairwise combinations was calculated and presented as a dendrogram (Fig. 4). This analysis revealed that, although GroEL1 clearly belonged to the actinomycete group, it was distinguishable from all other members, which formed a distinct family.

Multiple *groEL*-like genes in streptomycete genomes. Southern blots of genomic DNA isolated from 12 different *Streptomyces* spp. (*S. albus* and *S. coelicolor* A3 [2] [John Innes

```
   1  GGCAGGCAGGCCGTCCGTCGCGGGCGCCAGCACGGTCCGGAGTATCCACCCTCCCCAGAC

  61  AGAGTCCGGGGGACCCCCAGTCCTGGGAGGACCACTTCACATGGCCAAGATCATCGCGTT
                                                M  A  K  I  I  A  F

 121  CGACGAGGAGGCGCGGCGCGGTCTCGAGCGCGGGATGAACCAGCTCGCCGACGCCGTCAA
       D  E  E  A  R  R  G  L  E  R  G  M  N  Q  L  A  D  A  V  K

 181  GGTCACCCTTGGCCCCAAGGGCCGCAACGTCGTCCTCGAGAAGAAGTGGGGCGCCCCCAC
       V  T  L  G  P  K  G  R  N  V  V  L  E  K  K  W  G  A  P  T

 241  GATCACCAACGATGGTGTCTCCATCGCCAAGGAGATCGAGCTCGAAGACCCGTACGAGAA
       I  T  N  D  G  V  S  I  A  K  E  I  E  L  E  D  P  Y  E  K

 301  GATCGGCGCCGAGCTGGTCAAGGAGGTCGCGAAGAAGACGGACGACGTCGCCGGTGACGG
       I  G  A  E  L  V  K  E  V  A  K  K  T  D  D  V  A  G  D  G

 361  CACGACGACCGCGACCGTCCTGGCCCAGGCGCTCGTCCGCGAGGGTCTGCGCAACGTCGC
       T  T  T  A  T  V  L  A  Q  A  L  V  R  E  G  L  R  N  V  A

 421  CGCGGGTGCCAACCCGATGGCCCTGAAGCGCGGCATCGAGAAGGCCGTCGAGGCCGTCTC
       A  G  A  N  P  M  A  L  K  R  G  I  E  K  A  V  E  A  V  S

 481  CTCCGCCCTGCTGGAGCAGGCGAAGGACGTGGAGACCAAGGAGCAGATCGCCTCCACCGC
       S  A  L  L  E  Q  A  K  D  V  E  T  K  E  Q  I  A  S  T  A

 541  CTCCATCTCCGCCGCCGACACCCAGATCGGCGAGCTGATCGCCGAGGCCATGGACAAGGT
       S  I  S  A  A  D  T  Q  I  G  E  L  I  A  E  A  M  D  K  V

 601  CGGCAAGGAAGGCGTCATCACCGTCGAGGAGTCCCAGACCTTCGGTCTGGAGCTGGAGCT
       G  K  E  G  V  I  T  V  E  E  S  Q  T  F  G  L  E  L  E  L

 661  CACCGAGGGTATGCGCTTCGACAAGGGCTACATCTCGGCGTACTTCGCCACCGACATGGA
       T  E  G  M  R  F  D  K  G  Y  I  S  A  Y  F  A  T  D  M  E

 721  GCGGATGGAGGCGTCGCTCGACGACCCGTACATCCTCATCGTCAACTCCAAGATCGGCAA
       R  M  E  A  S  L  D  D  P  Y  I  L  I  V  N  S  K  I  G  N

 781  CGTGAAGGACCTGCTCCCGCTGCTGGAGAAGGTCATGCAGTCGGGCAAGCCGCTGCTGAT
       V  K  D  L  L  P  L  L  E  K  V  M  Q  S  G  K  P  L  L  I

 841  CATCGCCGAGGACGTCGAGGGCGAGGCCCTGTCGACCCTGGTGGTCAACAAGATCCGTGG
       I  A  E  D  V  E  G  E  A  L  S  T  L  V  V  N  K  I  R  G

 901  CACCTTCAAGTCCGTCGCCGTCAAGGCCCCGGGCTTCGGCGACCGCCGCAAGGCCATGCT
       T  F  K  S  V  A  V  K  A  P  G  F  G  D  R  R  K  A  M  L

 961  CGGTGACATCGCCATCCTCACGGGCGGCACGGTCATCTCCGAGGAGGTCGGCCTCAAGCT
       G  D  I  A  I  L  T  G  G  T  V  I  S  E  E  V  G  L  K  L

1021  GGAGAACGCCGGTCTCGACCTGCTGGGCCGCGCCCGCAAGGTCGTCATCACCAAGGACGA
       E  N  A  G  L  D  L  L  G  R  A  R  K  V  V  I  T  K  D  E

1081  GACCACGATCGTCGACGGTGCCGGCGACACGGACCAGGTCAACGGCCGCGTCGCGCAGAT
       T  T  I  V  D  G  A  G  D  T  D  Q  V  N  G  R  V  A  Q  I

1141  CCGCGCCGAGATCGAGAACAGCGACTCGGACTACGACCGCGAGAAGCTCCAGGAGCGTCT
       R  A  E  I  E  N  S  D  S  D  Y  D  R  E  K  L  Q  E  R  L

1201  GGCGAACGTGGCCGGCGGCGTGGCCGTCATCAAGGCCGGTGCCGCCACCGAGGTCGAGCT
       A  N  V  A  G  G  V  A  V  I  K  A  G  A  A  T  E  V  E  L

1261  GAAGGAGCGCAAGCACCGCATCGAGGACGCGGTGCGCAACGCCAAGGCCGCCGTCGAGGA
       K  E  R  K  H  R  I  E  D  A  V  R  N  A  K  A  A  V  E  E

1321  GGGCATCGTCGCCGGTGGTGGCGTGGCTCTGCTCCAGGCTTCCTCGGTCTTCGAGAAGCT
       G  I  V  A  G  G  G  V  A  L  L  Q  A  S  S  V  F  E  K  L

1381  CGAGCTCGAGGGCGACGAGGCCACCGGTGCCGCCGCCGTCAAGCTGGCCCTGGAGGCCCC
       E  L  E  G  D  E  A  T  G  A  A  A  V  K  L  A  L  E  A  P

1441  GCTCAAGCAGATCGCCGTCAACGGTGGCCTGGAGGGCGGCGTGGTCGTGGAGAAGGTGCG
       L  K  Q  I  A  V  N  G  G  L  E  G  G  V  V  V  E  K  V  R

1501  CAACCTCTCCGTGGGCCACGGCCTGAACGCCGCGACCGGGCAGTACGTCGACATGATCGC
       N  L  S  V  G  H  G  L  N  A  A  T  G  Q  Y  V  D  M  I  A

1561  CGAGGGCATCCTCGACCCGGCCAAGGTGACGCGTTCCGCGCTCCAGAACGCCGCGTCGAT
       E  G  I  L  D  P  A  K  V  T  R  S  A  L  Q  N  A  A  S  I

1621  CGCCGCCCTCTTCCTCACCACCGAGGCCGTCATCGCCGACAAGCCGGAGAAGGCCGCCGC
       A  A  L  F  L  T  T  E  A  V  I  A  D  K  P  E  K  A  A  A

1681  CGCTGCCCCGGGTGGCATGCCCGGCGGTGACATGGACTTCTGACCGGACCCGCCTTCCGG
       A  A  P  G  G  M  P  G  G  D  M  D  F  *

1741  TCGGCCCGTCCGGCACCACGCCCACCGGGGCGGTGTCCCTCTTCCGAGGGATGCCGCCCC
                          ---------------------->    <----------------

1801  GGTGGCGTTGGTGC  1814
```

FIG. 2. Nucleotide sequence of *groEL2*. The arrows indicate the position of an imperfect inverted repeat.

```
S. sp. 7942     MAKRII-YNENARRALERGIDILAEAVAVTLGPKGRNVVLEKKFGAPQIINDGVTIAKEI
E. coli         MAAKDVKFGNDARVKMLRGVNVLADAVKVTLGPKGRNVVLDKSFGAPTITKDGVSVAREI
C. burneti      MAAKVLKFSHEVLHAMSRGVEVLANAVKVTLGPKGRNVVLDKSFGAPTITKDGVSVAKEI
L. pneumono.    MA-KELRFGDDARLQMLAGVNALADAVQVTMGPRGRNVVLEKSYGAPTVTKDGVSVAKEI
M. bovis        MA-KTIAYDEEARRGLERGLNALADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEI
M. tubercul.    MA-KTIAYDEEARRGLERGLNALADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEI
M. leprae       MA-KTIAYDEEARRGLERGLNSLADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEI
S. ALBUS-56     MA-KIIAFDEEARRGLERGMNQLADAVKVTLGPKGRNVVLEKKWGAPTITNDGVSIAKEI
S. ALBUS-58     MA-KILKFDEDARRALERGVNQLADTVKVTIGPKGRNVVIDKKFGAPTITNDGVTIAREV
CONSERVED       MA . . .....    .  G.. LA..V VT.GP.GRNVV..K..GAP ...DGV..A.E.

S. sp. 7942     ELEDHIENTGVALIRQAASKTNDAAGDGTTTATVLAHAVVKEGLRNVAAGANAILLKRGI
E. coli         ELEDKFENMGAQMVKEVASKANDAAGDGTTTATVLAQAIITEGLKAVAAGMNPMDLKRGI
C. burneti      ELEDKFENMGAQMVKEVASRTSDDAGDGTTTATVLAQAILVEGIKAVIAGMNPMDLKRGI
L. pneumono.    EFEERFMNMGAQMVKEVASKTSDTAGDGTTTATVLARSILVEGHKAVAAGMNPMDLKRGI
M. bovis        ELEDPYEKIGAELVKEVAKKTDDVAGDGTTTATVLAQALVREGLRNVAAGANPLGLKRGI
M. tubercul.    ELEDPYEKIGAELVKEVAKKTDDVAGDGTTTATVLAQALVREGLRNVAAGANPLGLKRGI
M. leprae       ELEDPYEKIGAELVKEVAKKTDDVAGDGTTTATVLAQALVKEGLRNVAAGANPLGLKRGI
S. ALBUS-56     ELEDPYEKIGAELVKEVAKKTDDVAGDGTTTATVLAQALVREGLRNVAAGANPMALKRGI
S. ALBUS-58     ECDDPYENLGAQLVKEVATKTNDIAGDGTTTATVLAQALVREGLRNVAAGASPAALKKGI
CONSERVED       E ..    . G.......A....D AGDGTTTATVLA.... EG ..V AG ..  LK.GI

S. sp. 7942     DKATNFLVEQIKSHARPVEDSKSIAQVGAISAGNDFEVGQMIADAMDKVGKEGVISLEEG
E. coli         DKAVTAAVEELKALSVPCSDSKAIAQVGTISANSDETVGKLIAEAMDKVGKEGVITVEDG
C. burneti      DKAVTAAVAELKKISKPCKDQKAIAQVGTISANSDKSIGDIIAEAMEKVGKEGVITVEDG
L. pneumono.    DKAVLAVTKKLQAMSKPCKDSKAIAQVGTISANSDEAIGAIIAEAMEKVGKEGVITVEDG
M. bovis        EKAVEKVTETLLKGAKEVETKEQIAATAAISAG-DQSIGDLIAEAMDKVGNEGVITVEES
M. tubercul.    EKAVEKVTETLLKGAKEVETKEQIAATAAISAG-DQSIGDLIAEAMDKVGNEGVITVEES
M. leprae       EKAVDKVTETLLKDAKEVETKEQIAATAAISAG-DQSIGDLIAEAMDKVGNEGVITVEES
S. ALBUS-56     EKAVEAVSSALLEQAKDVETKEQIASTASISAA-DTQIGELIAEAMDKVGKEGVITVEES
S. ALBUS-58     DAAVAAVSAELLDTARPIDDKSDIAAVAALSAQ-DKQVGELIAEAMDKVGKDGVITVEES
CONSERVED       . A.        .   .    .. . IA ....SA D  .F .IA.AM.KVG..GVI..E..

S. sp. 7942     KSMTTELEVTEGMRFDKGYISPYFATDTERMEAVFDEPFILITDKKIGLVQDLVPVLEQV
E. coli         TGLQDELDVVEGMQFDRGYLSPYFINKPETGAVELESPFILLADKKISNIREMLPVLEAV
C. burneti      SGLENALEVVEGMQFDRGYLSPYFINNQQNMSAELENPFILLVDKKISNIRELIPLLENV
L. pneumono.    NGLENELSVVEGMQLI-AVHSPYFINNQQNMSCELEHPFILLVDKKVSSIREMLSVLEGV
M. bovis        NTFGLQLELTEGMRFDKGYISGYFVTDPERQEAVLEDPYILLVSSKVSTVKDLLPLLEKV
M. tubercul.    NTFGLQLELTEGMRFDKGYISGYFVTDPERQEAVLEDPYILLVSSKVSTVKDLLPLLEKV
M. leprae       NTFGLQLELTEGMRFDKGYISGYFVTDAERQEAVLEEPYILLVSSKVSTVKDLLPLLEKV
S. ALBUS-56     QTFGLELELTEGMRFDKGYISAYFATDMERMEASLDDPYILIVNSKIGNVKDLLPLLEKV
S. ALBUS-58     NTFGVDLDFTEGMAFDKGYLSPYMVTDQERMEAVLDDPYILIHQGKIGSIQDLLPLLEKV
CONSERVED        ..  .L. .EGM .  .  S Y. ..  .   .  .. P.IL.  K.. .......LE V

S. sp. 7942     ARAG--RPLVIIAEDIEKEALATLVVNRLRGVLNVAAVKAPGFGDRRKAMLEDIAVLTGG
E. coli         AKAG--KPLLIIAEDVEGEALATAVVNTIRGIVKVAAVKAPGFGDRRKAMLQDIATLTGG
C. burneti      AKSG--RPLLVIAEDIEGEALATLVVNNIRGVVKVAAVKAPGFGDRRKAMLQDIAVLTGG
L. pneumono.    AKSG--RPLLIIAEDVEGEALATLVVNNMRGIVKVCAVKAPGFGDRRKAMLQDIAILTKG
M. bovis        IGAG--KPLLIIAEDVEGEALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLQDMAILTGG
M. tubercul.    IGAG--KPLLIIAEDVEGEALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLQDMAILTGG
M. leprae       IQAG--KSLLIIAEDVEGEALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLQDMAILTGA
S. ALBUS-56     MQSG--KPLLIIAEDVEGEALSTLVVNKIRGTFKSVAVKAPGFGDRRKAMLGDIAILTGG
S. ALBUS-58     IQAGGSKPLLIIAEDVEGEALSTLVVNKIRGTFNAVAVKAPGFGDRRKAMLGDMATLTGA
CONSERVED         .G   ..L..IAED.E EAL.T VVN .RG. .  AVKAPGFGDRRKAML G.A.LT .

S. sp. 7942     QLITEDAARKLDTTKLDQLGKARRITITKDNTTIV-AEGNEAAVKARVDQIRRQIEETES
E. coli         TVISEEIGMELEKATLEDLGQAKRVVINKDTTTIIDGVGEEAAIQGRVAQIRQQIEEATS
C. burneti      KVISEEVGLSLEAASLDDLGSAKRVVVTKDDTTIIDGSGDAGDIKNRVEQIRKEIENSSS
L. pneumono.    QVISEEIGKSLEGATLEDLGSAKRIVVTKENTTIIDGEGKATEINARITQIRAQMEETTS
M. bovis        QVISEEVGLTLENADLSLLGKARKVVVTKDETTIVEGAGDTDAIAGRVAQIRQEIENSDS
M. tubercul.    QVISEEVGLTLENADLSLLGKARKVVVTKDETTIVEGAGDTDAIAGRVAQIRQEIENSDS
M. leprae       QVISEEVGLTLENTDLSLLGKARKVVMTKDETTIVEGAGDTDAIAGRVAQIRTEIENSDS
S. ALBUS-56     TVISEEVGLKLENAGLDLLGRARKVVITKDETTIVDGAGDTDQVNGRVAQIRAEIENSDS
S. ALBUS-58     TVIAEEVGLKLDQAGLDVLGTARRVTVTKDDTTIVDGGGNAEDVQGRVAQIKAEIESTDS
CONSERVED        .I.E. . .L. . L. LG A......K..TTI. . G..... .R..QI. ..E...S

S. sp. 7942     SYDKEKLQERLAKLSGGVAVVKVGAATETEMKDRKLRLEDAINATKAAVEEGIVPGGGTT
E. coli         DYDREKLQERVAKLAGGVAVIKVGAATEVEMKEKKARVEDALHATRAAVEEGVVAGGGVA
C. burneti      DYDKEKLQERLAKLAGGVAVIKVGAATEVEMKEKKARVEDALHATRAAVEEGVVPGGGVA
L. pneumono.    DYDREKLQERVAKLAGGVAVIKVGAATEVEMKEKKARVEDALHATRAAVEEGIVAGGGVA
M. bovis        DYDREKLQERLAKLAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVT
M. tubercul.    DYDREKLQERLAKLAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVT
M. leprae       DYDREKLQERLAKLAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVT
S. ALBUS-56     DYDREKLQERLANVAGGVAVIKAGAATEVELKERKHRIEDAVRNAKAAVEEGIVAGGGVA
S. ALBUS-58     DNDREKLQERLAKLAGGVCVIRVGAATEVELKERKHRLEDAISATRAAVEEGIVSGGGSA
CONSERVED       ..D.EKLQER.A...GGV V...GAATE.E.K..K R.EDA. ...AAVEEG.V.GGG .

S. sp. 7942     LAHLAPQLEEWATANLSGEELTGAQIV--ALTARLKRIAENAGLNGAVISERVKELPFDE
E. coli         LIRVASKLAD--LRGQNEDQNVGIKVALRAMEAPLRQIVLNCGEEPSVVANTV-KGGDGN
C. burneti      LIRVLKSLDS--VEVENEDQRVGVEIARRAMAYPLSQIVKNTGVQAAVVADKVLNHKDVN
L. pneumono.    LIRAQKALDS--LKGDNDDQNMGINILRRAIESPMRQIVTNAGYEASVVVNKVAEHKD-N
M. bovis        LLQAAPTLDE---LKLEGDEATGANIVKVALEAPLKQIAFNSGLEPGVVAEKVRNLPAGH
M. tubercul.    LLQAAPTLDE---LKLEGDEATGANIVKVALEAPLKQIAFNSGLEPGVVAEKVRNLPAGH
M. leprae       LLQAAPALDK---LKLTGDEATGANIVKVALEAPLKQIAFNSGMEPGVVAEKVRNLSVGH
S. ALBUS-56     LLQASSVFEK---LELEGDEATGAAAVKLALEAPLKQIAVNGGLEGGVVVEKVRNLSVGH
S. ALBUS-58     LVHAVKVLDD--NLGRTGDEATGVAVVRRAAVEPLRWIAENAGLEGYVITTKVAELDKGQ
CONSERVED       L .    ...      .... G     A    ... I. N G . V. . V .      .

S. sp. 7942     -GYDASNNQFVNMFTAGIVDPAKVTRSALMNAASIAAMVLTTECIVVDKPEPKEKAPAGA
E. coli         YGYNAATEEYGNMIDMGILDPTKVTRSALQYAASVAGLMITTECMVTDLPKNDAADL---
C. burneti      YGYNAATGEYGDMIEMGILDPTKVTRTALQNAASIAGLMITTECMVTEAPKKKEESMPGG
L. pneumono.    YGFNAATGEYGDMVEMGILDPTKVTRMALQNAASVASLMLTTECMVADLPKKEEGV--GA
M. bovis        -GLNAQTGVYEDLLAAGVADPVKVTRSALQNAASIAGLFLTTEAVVADKPEKEKASVPG-
M. tubercul.    -GLNAQTGVYEDLLAAGVADPVKVTRSALQNAASIAGLFLTTEAVVADKPEKEKASVPG-
M. leprae       -GLNAATGEYEDLLKAGVADPVKVTRSALQNAASIAGLFLTTEAVVADKPEKTAAPASDP
S. ALBUS-56     -GLNAATGQYVDMIAEGILDPAKVTRSALQNAASIAALFLTTEAVIADKPEKAAAAAPGG
S. ALBUS-58     -GFNAATGEYGDLVKAGVIDPVKVTRSALENAASIASLLLTTETLVVEKPAEEEPEAGHG
CONSERVED        G .A .. . ..   G. DP.KVTR AL. AAS.A.. .TTE .... P

S. sp. 7942     G------GGMGDFD--Y
E. coli         GAAGGMGGMGGMGGMM
C. burneti      GDMGGMGGMGGMGGMM
L. pneumono.    GDMGGMGGMGGMGGMM
M. bovis        G------GDMGGMD--F
M. tubercul.    G------GDMGGMD--F
M. leprae       T------GGMGGMD--F
S. ALBUS-56     M------PG-GDMD--F
S. ALBUS-58     H------GH-----S--H
CONSERVED
```

FIG. 3. Alignment of eubacterial GroELs from *Synechococcus* sp. (*S. sp.*) strain PCC 7942 (27), *E. coli* (9), *Coxiella burnetii* (26), *Legionella pneumophila* (18), *M. tuberculosis* (21), and *M. bovis* (22). Alignment was done by using Clustal programs (10). Conserved amino acids are indicated; dots show conservation of similar amino acids. A block of conserved amino acids, a portion of which was used to construct a degenerated probe (see Fig. 5), is underlined.

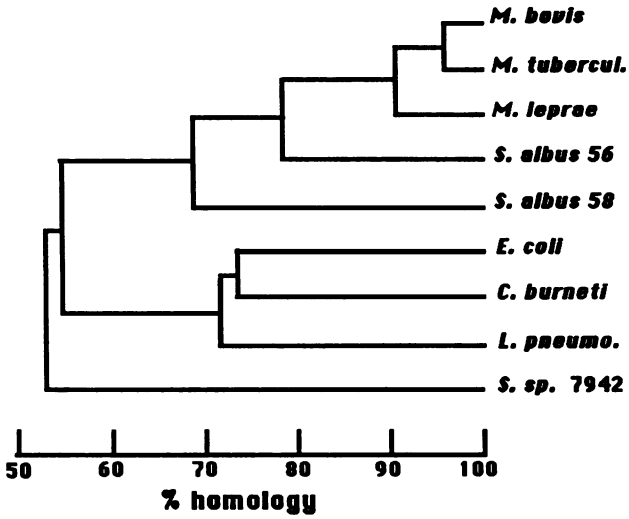

FIG. 4. Dendrogram representation of the similarities between eubacterial GroEL proteins. Analyses were carried out on the alignment shown in Fig. 3 by using CLUSTAL programs (10). Abbreviations and references are as described in the legend to Fig. 3.

Culture Collection]; *S. citricolor*, *S. colombiensis*, *S. endus*, *S. polychromogenes*, *S. rimosus*, and *S. spectabilis* (culture collection of Elizabeth Wellington]; *S. fradiae* ATCC 10745; *S. griseus* [ETH, SS-1198]; *S. kanamyceticus* ATCC 12853; *S. parvulus* ATCC 12434; and *S. viridochromogenes* GCC 40736) were hybridized with an oligonucleotide probe representing a sequence that is highly conserved among actinomycete *groEL* genes (Fig. 5; not all species are shown). Two strong bands of hybridization were observed with high-stringency conditions (60°C, 0.5× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate]) in every *Streptomyces* sp. tested. One band was observed in *Mycobacterium smegmatis* and no bands were observed in *E. coli* under these conditions of hybridization.

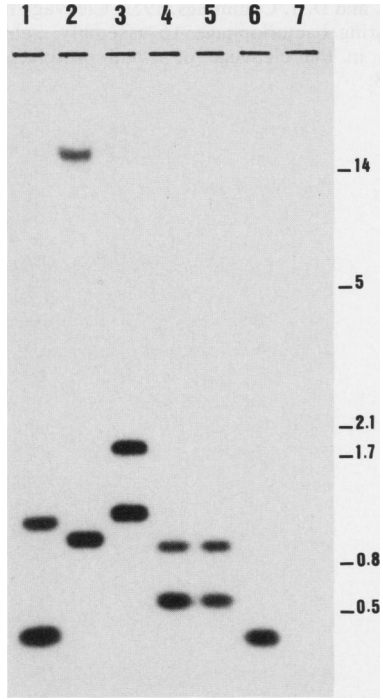

FIG. 5. Detection of two *groEL*-like genes in actinomycetes. Restriction endonuclease-digested genomic DNA of *S. viridochromogenes* (lane 1; *Xho*I), *S. parvulus* (lane 2; *Xho*I), *S. albus* (lane 3; *Xho*I), *S. lividans* (lane 4; *Xho*I), *S. coelicolor* (lane 5; *Xho*I), *M. smegmatis* (lane 6; *Pst*I), and *E. coli* (lane 7; *Pst*I) were screened with an OL3, a degenerate 30-mer oligonucleotide probe (GGCTTCGG[CT]GACCG CCGCAAGGC[CG]ATGCTC) conserved among actinomycete *groEL* genes (a portion of a block of conserved amino acids, underlined in Fig. 3). The positions and sizes (in kilobases) of molecular size markers (bacteriophage lambda cleaved with *Pst*I) are indicated by arrows.

The observation that the probe hybridized to two restriction fragments after digestion with either *Xho*I (Fig. 5) or *Sst*I (data not shown) strongly suggested that the two genes are not closely linked on the streptomycete chromosome. This was confirmed by using Southern blots (data not shown), which indicated that neither cosmid pGM3 nor pGM4 contained both genes. Finally, digestion of the *S. albus* genome by *Ssp*I followed by pulsed-field electrophoresis, blotting, and hybridization showed that the two genes were present on different 700- and 200-kb fragments (data not shown).

Streptomyces spp. are gram-positive mycelial eubacteria that are well known for their ability to produce antibiotics, secrete proteins, and differentiate during starvation; such conditions probably involve stress heat shock proteins (8). GroEL synthesis is induced in stationary-phase cultures of *E. coli* (16) starved for carbon, phosphate, or nitrogen (14). The observation that most *Streptomyces* species have two *groEL*-like genes that express three gene products supports the notion that multiple GroEL proteins (probably comparable to those observed in *S. albus*) fulfill physiologically distinct and important roles. Specialized functions may be required for substrate (protein or possibly peptide antibiotic) recognition, subcellular localization, or physiological changes associated with differentiation.

Nucleotide sequence accession numbers. The nucleotide sequence data shown in Fig. 1 and 2 have been submitted to GenBank under accession numbers M76657 and M76658, respectively.

We thank E. Wellington for kindly providing streptomyces strains, G. Platas for pulsed-field electrophoretic analyses, David Holmes, Carlos Martin, and Pascale Servant for providing actinomycete chromosomal DNA, and David Holmes and Peter Kaiser for critical reading of the manuscript.

This work was supported by the Institut Pasteur and the Ministère de la Recherche et de la Technologie (CHVP program).

REFERENCES

1. **Altman, E., C. A. Kumamoto, and S. D. Emr.** 1991. Heat-shock proteins can substitute for SecB function during protein export in *Escherichia coli*. EMBO J. **10**:239–245.
2. **Atkins, J. F., R. B. Weiss, and R. F. Gesteland.** 1990. Ribosome gymnastics-degree of difficulty 9.5, style 10.0. Cell **62**:413–423.
3. **Baird, P. N., L. M. C. Hall, and A. R. M. Coates.** 1989. Cloning and sequence analysis of the 10 kDa antigen gene of *Mycobacterium tuberculosis*. J. Gen. Microbiol. **135**:931–939.
4. **Blinkowa, A. L., and J. R. Walker.** 1990. Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III γ subunit from within the τ subunit reading frame. Nucleic Acids Res. **18**:1725–1729.
5. **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.
6. **Flower, A. M., and C. S. McHenry.** 1990. The γ subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. Proc. Natl. Acad. Sci. USA **87**:3713–3717.

7. **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.
8. **Guglielmi, G., P. Mazodier, C. J. Thompson, and J. Davies.** 1991. A survey of the heat shock response in four *Streptomyces* species reveals two *groEL*-like genes and three GroEL-like proteins in *Streptomyces albus*. J. Bacteriol. **173**:7374–7381.
9. **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 (London) **333**:330–334.
10. **Higgins, D. G., and P. M. Sharp.** 1988. CLUSTAL: a package for performing multiple sequence alignments on a microcomputer. Gene **73**:237–244.
11. **Hohn, B., and J. Collins.** 1980. A small cosmid for efficient cloning of large DNA fragments. Gene **11**:291–298.
12. **Kumamoto, C. A.** 1991. Molecular chaperones and protein translocation across the *Escherichia coli* inner membrane. Mol. Microbiol. **5**:19–22.
13. **Lindquist, S., and E. A. Craig.** 1988. The heat-shock proteins. Annu. Rev. Genet. **22**:631–677.
14. **Matin, A.** 1991. The molecular basis of carbon-starvation-induced general resistance in *Escherichia coli*. Mol. Microbiol. **5**:19–22.
15. **Mehra, V., D. Sweetser, and R. A. Young.** 1986. Efficient mapping of protein antigenic determinants. Proc. Natl. Acad. Sci. USA **83**:7013–7017.
16. **Neidhardt, F. C., T. A. Phillips, R. A. VanBogelen, M. W. Smith, Y. Georgalis, and A. R. Subramanian.** 1981. Identity of the B56.5 protein, the A-protein, and the *groE* gene product of *Escherichia coli*. J. Bacteriol. **145**:513–520.
17. **Neidhardt, F. C., and R. A. VanBogelen.** 1987. The heat shock response, p. 1334–1345. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
18. **Sampson, J. S., S. P. O'Connor, B. P. Holloway, B. B. Plikaytis, G. M. Carlone, and L. W. Mayer.** 1990. Nucleotide sequence of *hptB*, the *Legionella pneumophila* gene encoding the 58-kilodalton (kDa) common antigen, formerly designated the 60-kDa common antigen. Infect. Immun. **58**:3154–3157.
19. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
20. **Schlesinger, M. J.** 1990. Heat shock proteins. J. Biol. Chem. **265**:12111–12114.
21. **Shinnick, T. M.** 1987. The 65-kilodalton antigen of *Mycobacterium tuberculosis*. J. Bacteriol. **169**:1080–1088.
22. **Thole, J. E. R., W. J. Keulen, A. H. Kolk, D. G. Groothuis, L. G. Berwald, R. H. Tiesjema, and J. D. A. van Embden.** 1987. Characterization, sequence determination, and immunogenicity of a 64-kilodalton protein of *Mycobacterium bovis* BCG expressed in *Escherichia coli* K-12. Infect. Immun. **55**:1466–1475.
23. **Tsuchihashi, Z., and A. Kornberg.** 1990. Translational frameshifting generates the γ subunit of DNA polymerase holoenzyme. Proc. Natl. Acad. Sci. USA **87**:2516–2520.
24. **Van Dyk, T. K., A. A. Gatenby, and R. A. LaRossa.** 1989. Demonstration by genetic suppression of interaction of GroE products with many proteins. Nature (London) **342**:451–453.
25. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
26. **Vodkin, M. H., and J. C. Williams.** 1988. A heat shock operon in *Coxiella burnetii* produces a major antigen homologous to a protein in both *Mycobacteria* and *Escherichia coli*. J. Bacteriol. **170**:1227–1234.
27. **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.
28. **Zweig, M., and D. J. Cummings.** 1973. Cleavage of head and tail protein during bacteriophage T5 assembly: selective host involvement in the cleavage of a tail protein. J. Mol. Biol. **80**:505–518.