The EMBO Journal vol.12 no.7 pp.2901 – 2912, 1993

One member of a *groESL*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes

H.M.Fischer, M.Babst, T.Kaspar, G.Acuña[1], F.Arigoni[2] and H.Hennecke

Mikrobiologisches Institut, Eidgenössische Technische Hochschule, ETH-Zentrum, Schmelzbergstrasse 7, CH-8092 Zürich, Switzerland
[1]Present address: Institut für Toxikologie, Eidgenössische Technische Hochschule, Schorenstrasse 16, CH-8603 Schwerzenbach, Switzerland
[2]Present address: Institut de Biologie Physico-Chimique, 13 Rue Pierre et Marie Curie, F-75005 Paris, France

Communicated by J.Schell

This report is concerned with the structural characterization and genetic regulation of new bacterial *groES* and *groEL* chaperonin genes, and presents two novelties. The first is the discovery that the nitrogen fixing soybean root nodule bacterium, *Bradyrhizobium japonicum*, unlike all other prokaryotes investigated so far, possesses a multigene family consisting of five very similar, though not identical, *groESL*-like genes. The second novelty relates to the finding that these five homologues are expressed to different degrees and, in particular, that one family member (namely $groESL_3$) is induced by a mechanism that does not involve the well-known heat shock response. By contrast, the $groESL_3$ genes are co-regulated together with symbiotic nitrogen fixation genes, in that they are activated by the nitrogen fixation regulatory protein NifA at low oxygen conditions and transcribed from a $-24/-12$ promoter by the σ^{54} RNA polymerase. Two other members of the *groESL* gene family are apparently expressed constitutively at different levels, and yet another one is strongly induced by high temperature. As an attractive hypothesis it follows that *B.japonicum* may modulate its cellular contents of GroES- and GroEL-like chaperonins in response to specific environmental conditions and physiological needs.
Key words: gene regulation/*nifA*/oxygen control/phylogenetic tree/symbiosis

Introduction

Many prokaryotic and eukaryotic proteins require the assistance of molecular chaperones in order to become properly folded and/or assembled as oligomeric structures. It has been proposed that the essential function of chaperones is to prevent the formation of incorrect structures by inhibiting unproductive folding pathways (Ellis, 1990). In addition, chaperones play an important role in minimizing cellular damage caused by stress conditions such as heat shock. Accordingly, the majority of the ubiquitous chaperones belong to one of three families of highly conserved heat shock proteins, namely the Hsp90, Hsp70 and Hsp60 protein families (for recent reviews see Ellis and van der Vies, 1991; Gething and Sambrook, 1992; Hartl *et al.*, 1992). The latter family, also termed chaperonins (Hemmingsen *et al.*, 1988), includes a class of sequence-related proteins found in a large number of bacterial species as well as in mitochondria and plastids (Georgopoulos and Ang, 1990; Zeilstra-Ryalls *et al.*, 1991). The GroEL protein of *Escherichia coli* is probably the best-characterized chaperonin. The 57.3 kDa GroEL protein is encoded by the promoter-distal gene of the *E.coli groESL* operon which is under dual genetic control. Basal constitutive expression at low temperatures originates from a σ^{70}-dependent promoter, whereas at physiological temperatures (>30°C) transcription primarily starts from a σ^{32}-type promoter (heat shock promoter) located upstream of the σ^{70} promoter (Zhou *et al.*, 1988; Georgopoulos and Ang, 1990). The *groES* and *groEL* gene products are essential for growth of *E.coli* at all temperatures (Fayet *et al.*, 1989), and temperature-sensitive *groESL* mutants show a pleiotropic phenotype (Gething and Sambrook, 1992 and references therein). The functional GroESL complex consists of 14 GroEL subunits arranged in two stacked heptameric rings and a ring of seven GroES subunits bound to one end of the GroEL double ring. Substrate proteins to be folded probably bind to the central cavity of the GroEL cylinder, and it is thought that GroES couples the ATPase activity of GroEL with the protein folding mechanism (Langer *et al.*, 1992 and references therein).

From recent reports it became evident that chaperonin-mediated protein folding or assembly may also be involved in the complex regulation of nitrogen fixation by prokaryotic organisms. For example, it was shown that in the free-living diazotrophic bacterium *Klebsiella pneumoniae*, a GroEL-like protein might be involved in the control of nitrogenase biogenesis (Govezensky *et al.*, 1991). In *Rhizobium meliloti*, the root nodule symbiont of alfalfa, a GroEL-like protein is required for full activity of NodD3, a transcriptional activator of nodulation genes (Fisher and Long, 1992).

In the soybean symbiont *Bradyrhizobium japonicum*, conditions of low oxygen tension present in root nodules or in anaerobic bacterial cultures are critical for the transcriptional activation of different sets of genes involved in nitrogen fixation (Hennecke, 1990). The *B.japonicum* NifA protein, an oxygen-responsive transcriptional activator, plays a key role in this process (Fischer and Hennecke, 1987; Fischer *et al.*, 1988). Under low oxygen conditions, NifA binds to an upstream activator sequence and interacts with the RNA polymerase$-\sigma^{54}$ complex bound at $-24/-12$-type promoters thereby catalysing open complex formation and initiating transcription (Morett and Buck, 1989; Morett *et al.*, 1991). Accordingly, a *B.japonicum nifA* mutant is unable to fix nitrogen. There is now cumulative evidence that the regulatory scope of NifA extends beyond the functions related directly to nitrogen fixation. This was first observed in a *B.japonicum nifA* mutant that was not only defective in nitrogen fixation but also elicited nodules with a necrotic structure reminiscent of a hypersensitive

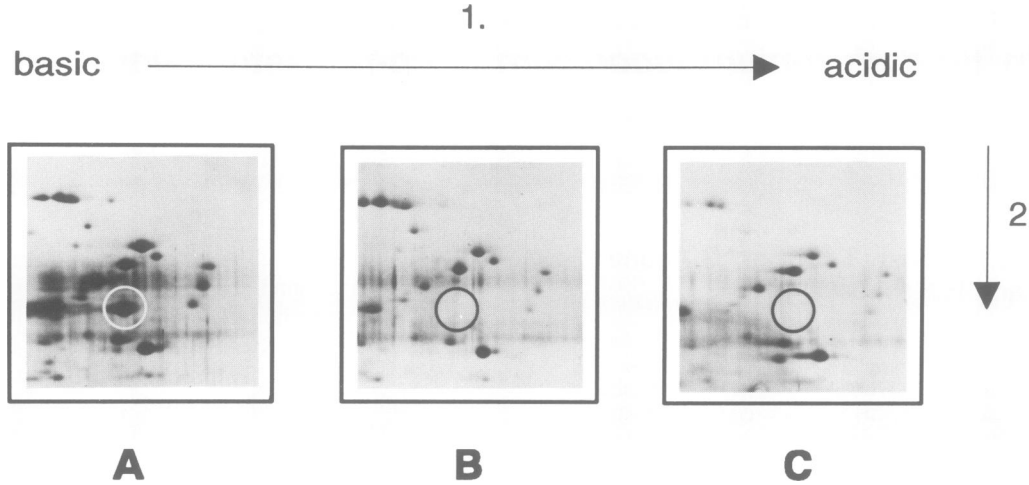


Fig. 1. Relevant sections of silver-stained, two-dimensional gels of lysates from anaerobically grown cells of *B.japonicum* 110*spc*4 (wild-type; **A**), A9 (*nifA*$^-$; **B**) and 8541 (*groEL*$_3$$^-$; **C**). The arrows indicate the direction of the isoelectric focusing (first dimension) and the SDS−PAGE (second dimension) steps. In panel A the 58 kDa GroEL$_3$ protein whose synthesis is induced under anaerobic conditions in the wild-type is circled in white, whereas the black circles in panels B and C mark its absence in extracts of *nifA*$^-$ and *groEL*$_3$$^-$ strains, respectively.

response in incompatible plant−pathogen interactions (Fischer *et al.*, 1986; Studer *et al.*, 1987; Parniske *et al.*, 1991). Since then, various (brady-)rhizobial species have been shown to harbor non-*nif* and non-*fix* genes whose expression is controlled by NifA (Hirsch and Smith, 1987; Hawkins and Johnston, 1988; Martin *et al.*, 1988; Murphy *et al.*; 1988; Sanjuan and Olivares, 1989). The pleiotropic phenotype of a *B.japonicum nifA* mutant was visualized by two-dimensional gel electrophoresis of protein extracts of microaerobically grown *B.japonicum nifA* mutant cells (Fischer *et al.*, 1986). In these extracts at least eight proteins were missing as compared with the wild type (Regensburger *et al.*, 1986). Among these were the previously identified NifH protein (nitrogenase component II apoprotein; Hahn *et al.*, 1984) and an unknown 58 kDa protein, both present as prominent spots in wild-type extracts. We exploited this system in the context of our efforts to identify new targets for NifA-mediated control. Here we describe the identification of the 58 kDa protein as a *B.japonicum* GroEL homologue which is encoded in one member of a family of five differentially regulated *B.japonicum groESL* operons. Interestingly, its synthesis is co-regulated at the transcriptional level with the process of symbiotic nitrogen fixation via NifA and the σ factor RpoN (σ^{54}). Two additional *groESL* genes are expressed constitutively, and a third homologue is heat shock-inducible.

Results

Identification of a GroEL-like protein whose synthesis is NifA-dependent

B.japonicum synthesizes a number of proteins specifically under microaerobic or anaerobic growth conditions in a NifA-dependent manner, as shown previously by two-dimensional gel electrophoresis of crude cell extracts (Fischer *et al.*, 1986). Among these we observed a prominent, well resolved protein with an apparent molecular mass of 58 kDa (Figure 1A). This protein was isolated by electroelution from two-dimensional gels (see Materials and methods), and ~150 pmol were used to determine its N-terminal amino acid sequence (Figure 2, second line). Four positions within the sequenced stretch of 29 amino acids could not be determined with certainty, and two positions yielded ambiguous results. No N-terminal methionine was found. Attempts to identify the *B.japonicum* DNA region that encoded the 58 kDa protein by using synthetic oligonucleotide hybridization probes deduced from the available N-terminal protein sequence were unsuccessful. However, homology searches in the MIPSX protein sequence databank revealed that the N-terminus of the 58 kDa protein showed a strong similarity to the N-terminus of the *E.coli* GroEL protein (M_r 57 259; Hemmingsen *et al.*, 1988) and to other members of the so-called GroEL protein family (Figure 2). Nineteen out of the 29 amino acids determined from the *B.japonicum* 58 kDa protein were identical as in the N-terminus of the *E.coli* GroEL protein (66% identity). This finding allowed us to use more specific probes in the search for the gene encoding the 58 kDa protein.

Cloning of five groEL-homologous regions

Five hybridizing *Bam*HI fragments of 3.7, 4.5, 6, 11 and 13 kb size were detected in Southern blot hybridization experiments with *Bam*HI-digested genomic DNA of *B.japonicum*, using an *E.coli groEL*-internal DNA fragment as probe (see Materials and methods). Essentially the same hybridization pattern was observed when oligonucleotide EL1 or EL3 was used as a probe (data not shown). This suggested the presence of multiple *groEL*-homologous regions in the *B.japonicum* genome. In order to clone these DNA regions, the genomic *Bam*HI DNA fragments in the respective size ranges of interest were isolated from preparative agarose gels and used for the construction of partial DNA libraries in *E.coli* DH5α. Colony hybridization screening of these libraries with the oligonucleotide EL1 as probe led to the identification of five recombinant plasmids (pRJ7931, pRJ7928, pRJ7932, pRJ7934 and pRJ7980) which represent the five *groEL*-homologous regions originally observed in the Southern blot hybridizations of genomic DNA (Figure 3A). Subsequent restriction analyses, further hybridization experiments, DNA sequence analyses and comparison of the deduced amino acid sequences with

	1 10 20 30
***E. coli* GroEL**	A A K D ***V K F G*** N D ***A R*** V K ***M L R G V*** N V ***L A D A V*** K ***V T*** . . .
Bj 58 kDa protein	X ***A*** X E ***V*** R/***K*** ***F G*** V N ***A R*** D R ***M L R G V*** D X ***L A*** D/N ***A V*** X ***V T*** . . .
Bj GroEL$_1$	M A A K E V K F S T D A R D R V L R G V D T L A N A V K V T . . .
Bj GroEL$_2$	M S A K E V K F G V D A R D R M L R G V D I L H N A V K V T . . .
Bj GroEL$_3$	M S A K E V K F G V N A R D R M L R G V D I L A N A V Q V T . . .
Bj GroEL$_4$	M A A K E V K F S V D A R D K I V R G V D I L A N X V K V T . . .
Bj GroEL$_5$	M A A K D V K F S G D A R E R M L R G V D I L A N A V K V T . . .

Fig. 2. Comparison of the N-terminal amino acid sequence of the *B.japonicum* 58 kDa protein with that of *E.coli* GroEL (Hemmingsen *et al.*, 1988) and the predicted N-termini of the *B.japonicum* GroEL$_1$ to GroEL$_5$ proteins. Identical residues of *E.coli* GroEL and the 58 kDa *B.japonicum* protein are marked in bold face italics. Underlined residues in the deduced *B.japonicum* GroEL sequences indicate positions in which the amino acids deviate from those of the sequenced N-terminus of the 58 kDa protein.

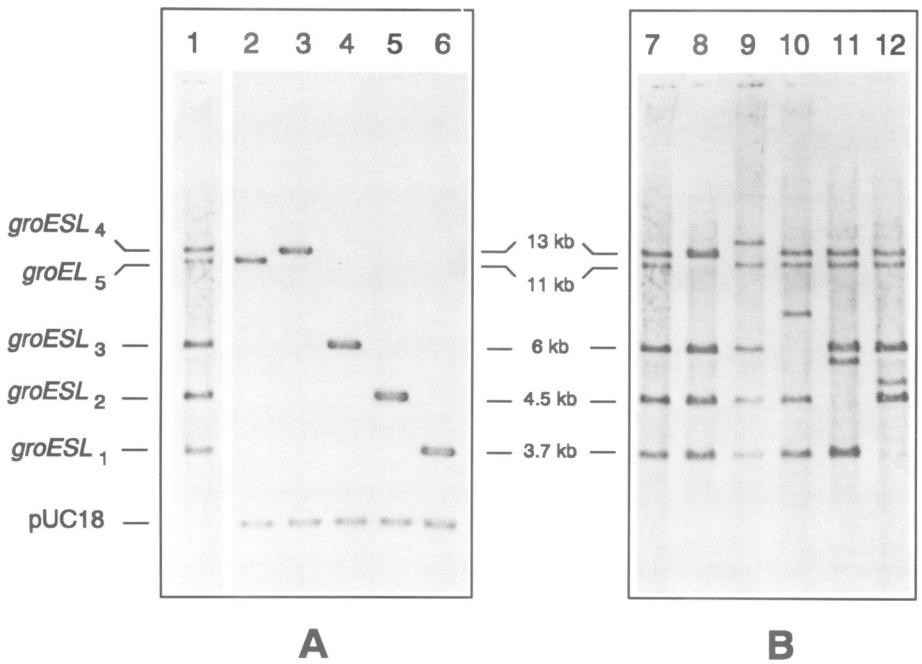


Fig. 3. Southern blot hybridizations of plasmids containing the cloned *groEL* genes of *B.japonicum* (**A**) and of genomic DNA isolated from *B.japonicum groEL* mutants (**B**). A 521 bp, *groEL*$_3$-internal DNA fragment that had been labelled by PCR with digoxigenin-11-dUTP, was used as a probe in both hybridization experiments. The *Bam*HI-digested DNA samples were as follows: chromosomal DNA of wild-type *B.japonicum* (lanes 1 and 7), pRJ7980 (*groEL*$_5$; lane 2), pRJ7934 (*groESL*$_4$; lane 3), pRJ7932 (*groESL*$_3$; lane 4), pRJ7928 (*groESL*$_2$; lane 5), pRJ7931 (*groESL*$_1$; lane 6), chromosomal DNA of *B.japonicum* mutant strain 7996 (*groEL*$_5^-$; lane 8), 8528 (*groEL*$_4^-$; lane 9), 8541 (*groEL*$_3^-$; lane 10), 7942 (*groEL*$_2^-$; lane 11) and 7940 (*groEL*$_1^-$; lane 12). The sizes of the *Bam*HI fragments containing the five *groEL* homologues are indicated between panels A and B. Hybridization to pUC18 in panel A originated from PCR-labelled vector sequences present in the probe. The weakly hybridizing band of ~3.7 kb in lane 12 originated from non-specific hybridization to a *Bam*HI fragment of about the same size as that containing *groESL*$_1$.

those of the *E.coli* GroEL protein led to the localization of five *groEL*-homologous genes (*groEL*$_1$ to *groEL*$_5$) on the cloned *Bam*HI DNA fragments as depicted in Figure 4. Genes homologous to *E.coli groES* were then detected immediately upstream of the *groEL* genes in four of the five *Bam*HI fragments (Figure 4). The four *groES*- and *groEL*-like genes cloned on plasmids pRJ7931, pRJ7928, pRJ7932 and pRJ7934 are probably organized in operons. The ATG start codon of the fifth *groEL*-like gene was located only 85 bp away from the left *Bam*HI site at the very end of the 11 kb *Bam*HI fragment cloned in plasmid pRJ7980 (Figure 4). By partial DNA sequence analysis of a cloned overlapping fragment, another *groES*-like gene (*groES*$_5$) was found which terminated 140 bp upstream of *groEL*$_5$, i.e. outside of the cloned DNA in pRJ7980 (data not shown).

In the course of mutagenizing *groESL*$_3$ and *groESL*$_4$ (see below), an antibiotic resistance cassette containing recognition sites for the rarely cutting restriction endonucleases *Pme*I, *Swa*I and *Pac*I (Kündig *et al.*, 1993) was introduced into these loci (Figure 4). This facilitated a physical mapping of *groESL*$_3$ and *groESL*$_4$ on the *B.japonicum* chromosome. Interestingly, *groESL*$_3$ was found to be located at map position 84° within a cluster of genes required for symbiotic nitrogen fixation, whereas *groESL*$_4$ mapped far outside of this region at map position 242°.

DNA sequence analysis of groESL$_2$ *and groESL*$_3$

The complete nucleotide sequences of the *groESL*$_2$ and *groESL*$_3$ regions on plasmids pRJ7928 and pRJ7932 were established (Figures 5 and 6). The *groES*$_2$ gene extending from position 352 (ATG) to 664 (TAG) was followed by the *groEL*$_2$ gene which started at position 738 (ATG) and ended at position 2388 (TAA) (Figure 5). Similarly, *groES*$_3$

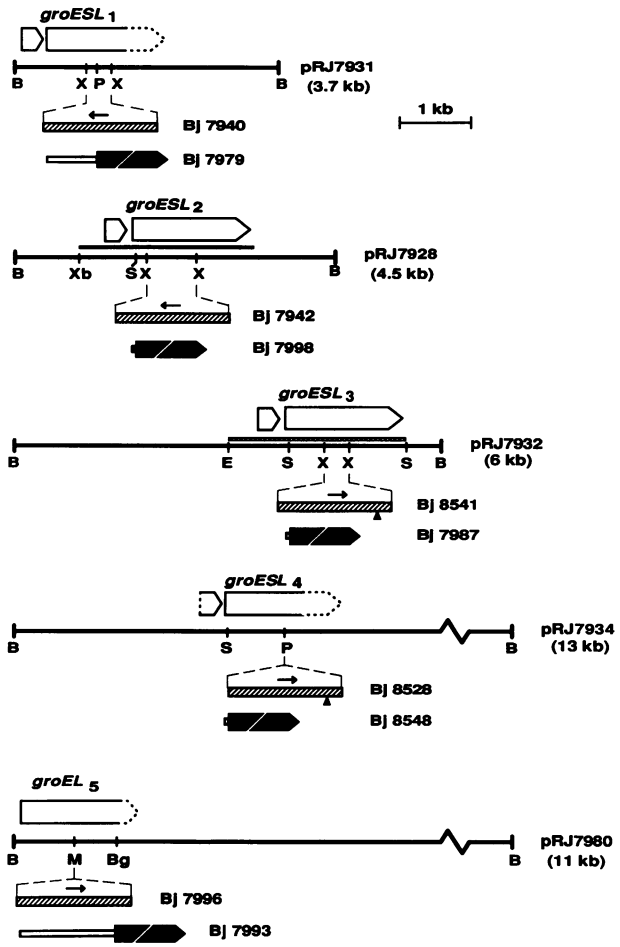


Fig. 4. Physical maps of the *Bam*HI inserts of the recombinant plasmids pRJ7931, pRJ7928, pRJ7932, pRJ7934 and pRJ7980. The location of the *groESL* genes as deduced from partial ($groESL_1$, $groESL_4$, $groEL_5$) or complete ($groESL_2$, $groESL_3$) nucleotide sequence analyses is indicated by open arrows. The 5′ region of $groES_4$ and the 3′ regions of $groEL_1$, $groEL_4$ and $groEL_5$ (broken lines) were not sequenced and the extension of these genes is drawn by analogy with $groESL_2$ and $groESL_3$. The cross-hatched bars below $groESL_2$ and $groESL_3$ mark the DNA regions whose sequences are shown in Figures 5 and 6, respectively. The structures of the *groEL* mutations are shown below the wild-type maps. The hatched bars with the small horizontal arrow refer to the inserted *aphII* (Km^r) cassettes and their orientation. Vertical filled arrowheads indicate the location on *aphII*-PSP (Kündig *et al.*, 1993) of recognition sites for the rarely cutting restriction endonucleases *Pme*I, *Swa*I and *Pac*I which were used for physical mapping of $groESL_3$ and $groESL_4$ on the *B.japonicum* chromosome. Open bars fused to black arrows symbolize the translational *groEL′-′lacZ* fusions. The numbers to the right refer to the respective strain designations. Only the relevant restriction sites are shown, and they are abbreviated as follows: B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; M, *Msc*I; P, *Pst*I, S, *Sal*I; X, *Xho*I; Xb, *Xba*I.

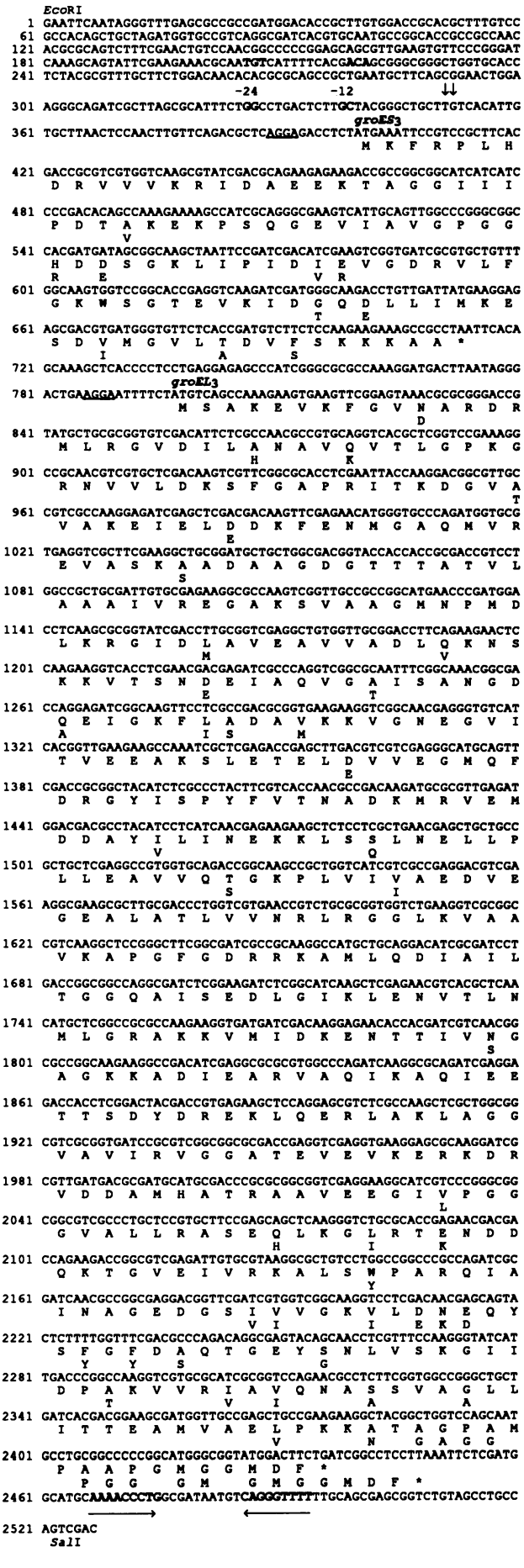
and $groEL_3$ spanned from positions 400 (ATG) to 712 (TAA) and 797 (ATG) to 2435 (TGA), respectively (Figure 6). Both genes, $groES_2$ and $groES_3$, encoded proteins of 104 amino acids (M_r 11 296 and 11 240, respectively), whereas the predicted gene products of $groEL_2$ and $groEL_3$ consisted of 550 and 546 amino acids with apparent molecular weights of 58 273 and 57 785, respectively. All four genes were preceded by presumptive ribosome binding sites. Putative transcription termination signals were identified downstream of both *groEL* genes by the presence of inverted repeats potentially able to form stem−loop structures. The coding sequences of the two

```
    XbaI
 1  TCTAGACTGCCGTTGGACGCGGCATCATTGTACCCGAACAGGCTAATCCAGACGCGCGCA
61  ACCCTGCCCGGCTAGCCATTTGGATGGCAAAGCCAAACAACAACCTGGGCAGACGGTACA
121 AGCCCGCCTAAAGTTTCCTCAGCGGTTCCTCCGACCAGCCCAGACTTGATTCCAGGGGCT
181 CCTTTAGGTGGCTCTGAATGGGCCCTCCGCGGACCGGTGAGGACGACGGATTTTCGAAAT
241 CCTTCGCGGCTTTGTCTTGTTTGCCGACCAAGCCGGCTTATTTTCGGAGCGGTCAGCACT
                                                    groES2
301 CGCTAGCTTCGACTGCTATTTCAAAAAGCGACTAGAAGATAGGAGGAGAGCATGAAATTC
                                                      M  K  F
361 CGTCCGCTTCACGACCGGGTCGTGGTCAAGCGCATCGACGCGGAAGAGAAGACCGCTGGC
     R  P  L  H  D  R  V  V  V  K  R  I  D  A  E  E  K  T  A  G
421 GGCATCATCATTCCCGACACGGTCAAGGAGAAACCCTCCCAAGGTGAAGTCATCGCCGTT
     G  I  I  I  P  D  T  V  K  E  K  P  S  Q  G  E  V  I  A  V
481 GGCCCCGGCGGTCGTGACGAAAGCGGCAAGCTGATCCCGATCGACGTCCGGGTCGGTGAC
     G  P  G  G  R  D  E  S  G  K  L  I  P  I  D  V  R  V  G  D
541 CGCGTATTGTTCGGCAAGTGGTCCGGCACCGAAGTCAAGATCGACACCCAGGAGCTGTTG
     R  V  L  F  G  K  W  S  G  T  E  V  K  I  D  T  Q  E  L  L
601 ATCATGAAGGAGAGCGACATCATGGGCGTTCTCGCCGACGTGTCTTCCAAGAAGAAGGCC
     I  M  K  E  S  D  I  M  G  V  L  A  D  V  S  S  K  K  K  A
661 GCCTAGTCCACCGCTCAAGCGCCTGTCGAGCGCCTCATCAAGGGTGAAACCAAGTCAAAA
     A  *                groEL2
721 ACTCAGGGAAACCCACTATGTCAGCAAAAGAAGTCAAATTCGGAGTCGATGCCCGCGACC
                        M  S  A  K  E  V  K  F  G  V  D  A  R  D
781 GCATGCTGCGCGGTGTCGACATCCTCCACAATGCGGTAAAAGTGACGCTCGGACCGAAGG
     R  M  L  R  G  V  D  I  L  H  N  A  V  K  V  T  L  G  P  K
841 GCCGCAACGTCGTCCTCGACAAGTCGTTCGGCGCTCCCCGTATTACCAAGGATGGCGTCA
     G  R  N  V  V  L  D  K  S  F  G  A  P  R  I  T  K  D  G  V
901 CCGTCGCCAAGGAGATCGAGCTCGAGGACAAGTTCGAGAACATGGGCGCCCAGATGGTGC
     T  V  A  K  E  I  E  L  E  D  K  F  E  N  M  G  A  Q  M  V
961 GCGAAGTCGCCTCCAAGTCCGCAGACGCGGCTGGCGACGGCACCACCACCGCGACCGTGC
     R  E  V  A  S  K  S  A  D  A  A  G  D  G  T  T  T  A  T  V
1021 TCGCCGCTGCGATCGTCCGTGAAGGCGCCAAGTCGGTTGCCGCCGGCATGAACCCGATGG
     L  A  A  A  I  V  R  E  G  A  K  S  V  A  A  G  M  N  P  M
1081 ATCTGAAGCGCGGCATCGACATGGCTGTGGAAGCCGTGGTCGCCGACCTCGTCAAAAACT
     D  L  K  R  G  I  D  M  A  V  E  A  V  V  A  D  L  V  K  N
1141 CCAAGAAGGTCACCTCGAACGAGGAAATCGCCCAGGTCGGCACGATTTCAGCCAATGGCG
     S  K  K  V  T  S  N  E  E  I  A  Q  V  G  T  I  S  A  N  G
1201 ATGCGGAGATCGGCAAGTTCATCTCCGACGCCATGAAGAAGGTCGGCAACGAGGGTGTCA
     D  A  E  I  G  K  F  I  S  D  A  M  K  K  V  G  N  E  G  V
1261 TCACCGTTGAGGAAGCCAAGTCGCTCGAGACCGAACTCGAAGTTGTCGAGGGCATGCAGT
     I  T  V  E  E  A  K  S  L  E  T  E  L  E  V  V  E  G  M  Q
1321 TCGACCGCGGCTATATCTCGCCCTACTTCGTCACCAACGCCGATAAGATGCGCGTTGAAA
     F  D  R  G  Y  I  S  P  Y  F  V  T  N  A  D  K  M  R  V  E
1381 TGGATGATGCCTACGTCCTGATTAATGAGAAAAAGCTATCTCAGTTGAATGAACTACTTC
     M  D  D  A  Y  V  L  I  N  E  K  K  L  S  Q  L  N  E  L  L
1441 CCTTGTTGGAAGCCGTGGTGCAGAGCGGCAAGCCGCTTGTGATTATCGCCGAAGACGTCG
     P  L  L  E  A  V  V  Q  S  G  K  P  L  V  I  I  A  E  D  V
1501 AAGGCGAAGCGCTCGCCACCCTCGTCGTCAACCGTCTCCGCGGCGGCCTGAAGGTCGCGG
     E  G  E  A  L  A  T  L  V  V  N  R  L  R  G  G  L  K  V  A
1561 CCGTCAAGGCTCCGGGCTTCGGCGATCGCCGCAAGGCCATGCTCCAGGACATCGCGATCC
     A  V  K  A  P  G  F  G  D  R  R  K  A  M  L  Q  D  I  A  I
1621 TAACCGGCGGCCAGGCGATCTCGGAAGACCTCGGCATCAAGCTCGAGAACGTCACGCTCA
     L  T  G  G  Q  A  I  S  E  D  L  G  I  K  L  E  N  V  T  L
1681 ATATGCTCGGTCGCGCCAAGAAGGTGATGATCGACAAGGAGAACACCACGATCGTCAGCG
     N  M  L  G  R  A  K  K  V  M  I  D  K  E  N  T  T  I  V  S
1741 GCGCCGGCAAGAAGGCCGACATCGAAGCGCGCGTGGCGCAAATCAAGGCGCAGATCGAGG
     G  A  G  K  K  A  D  I  E  A  R  V  A  Q  I  K  A  Q  I  E
1801 AGACCACCTCGGACTACGACCGTGAGAAGCTGCAGGAGCGTCTTGCCAAGCTCGCGGGCG
     E  T  T  S  D  Y  D  R  E  K  L  Q  E  R  L  A  K  L  A  G
1861 GCGTCGCGGTGATCCGCGTCGGCGGCGCGACTGAGGTCGAGGTGAAGGAGCGCAAGGATC
     G  V  A  V  I  R  V  G  G  A  T  E  V  E  V  K  E  R  K  D
1921 GCGTGGACGATGCGATGCATGCGACCCGCGCGGCGGTTGAAGAAGGCATCCTGCCAGGCG
     R  V  D  D  A  M  H  A  T  R  A  A  V  E  E  G  I  L  P  G
1981 GCGGCGTTGCGCTACTGCGTGCCTCCGAGCACCTCAAGGGCATTCGCACCAAAAATGACG
     G  G  V  A  L  L  R  A  S  E  H  L  K  G  I  R  T  K  N  D
2041 ACCAGAAGACCGGGGTCGAGATCGTGCGCAAGGCGTTGTCCTACCCGGCCCGCCAGATCG
     D  Q  K  T  G  V  E  I  V  R  K  A  L  S  Y  P  A  R  Q  I
2101 CAATCAACGCGGGCGAGGACGGCTCGGTCATCGTCGGCAAGATCCTCGAAAAGGATCAGT
     A  I  N  A  G  E  D  G  S  V  I  V  G  K  I  L  E  K  D  Q
2161 ACTCGTACGGCTACGACTCGCAGACCGGCGAATACGGCAACCTGGTCTCCAAGGGTATCA
     Y  S  Y  G  Y  D  S  Q  T  G  E  Y  G  N  L  V  S  K  G  I
2221 TCGATCCGACTAAGGTGGTTCGAGTGGCGATCCAGAACGCGGCGTCCGTCGCAGCGCTCC
     I  D  P  T  K  V  V  R  V  A  I  Q  N  A  A  S  V  A  A  L
2281 TGATCACCACCGAAGCTATGGTGGCGGAAGTGCCGAAGAAGAACACCGGTGCGGGCGGCA
     L  I  T  T  E  A  M  V  A  E  V  P  K  K  N  T  G  A  G  G
2341 TGCCTCCGGGCGGCGGCGGCATGGGTGGCATGGGGGGTATGGACTTCTAAATCCGACATT
     M  P  P  G  G  G  G  M  G  G  M  G  G  M  D  F  *
2401 CAGAGCTCAACAATAAGGCTCGGCAGCGACGCCGAGCCTTTGCT
```

Fig. 5. Nucleotide sequence of the *B.japonicum* $groESL_2$ operon. The shown 2444 bp sequence starts at an *Xba*I restriction site upstream of $groES_2$ (cf. Figure 4). Presumptive ribosome binding sites are underlined. The bold-face nucleotides underlined with inverted horizontal arrows indicate the 10 bp stem of a putative rho-independent transcription terminator downstream of $groEL_2$. This sequence has been submitted to the EMBL data library and assigned accession number Z22604.

groES as well as of the two *groEL* genes were highly conserved (85% identical nucleotides), whereas the similarity of the sequenced DNA upstream and downstream of these two *groESL* operons and in the *groES−groEL* intergenic regions was low (≤30% identical nucleotides). Accordingly, the two GroES proteins as well as the two GroEL proteins were extremely homologous (90 and 91% identical amino acids, respectively) as illustrated in Figure 6. A comparison of the *B.japonicum* $GroES_2$ and $GroES_3$ proteins with *E.coli* GroES revealed similarities of 50 and 46% identical amino acids, respectively. Similarly, the amino acid sequences of the *B.japonicum* $GroEL_2$ and $GroEL_3$ proteins exhibited ~65% identity with *E.coli* GroEL.

In order to attribute one of the cloned *groEL* genes to the



58 kDa protein originally isolated from two-dimensional gels, the nucleotide sequences of the 5′ ends of *groEL*$_1$, *groEL*$_4$ and *groEL*$_5$ were determined (data not shown) in addition to the complete sequences of *groESL*$_2$ and *groESL*$_3$. A comparison of the N-terminal amino acid sequences predicted from the *groEL*$_1$ to *groEL*$_5$ 5′ ends with that of the 58 kDa protein revealed that only the protein sequence deduced from *groEL*$_3$ did not conflict (Figure 2). This result suggested that the 58 kDa protein is encoded by *groEL*$_3$.

Transcriptional mapping of the groESL$_3$ *promoter*

A well conserved −24/−12-type promoter was identified ~70 bp upstream of the coding region for *groES*$_3$ (T_{325}GGCCT-N_5-TTGCT_{340}; Figure 6). Moreover, a putative binding site for the transcriptional activator protein NifA was present at an appropriate distance of ~120 bp upstream of the −24/−12 promoter (T_{206}GT-N_{10}-ACA; Figure 6). Finally, the sequence motif AAGCAA-N_4-GTA on the coding strand between the NifA binding site and the core consensus promoter sequence (corresponding to positions 255−243 of the non-coding strand in Figure 6) deviated in only two positions from the consensus binding site for IHF (WATCAA-N_4-TTR; Craig and Nash, 1984). The simultaneous presence of these promoter elements, which are characteristic for NifA-regulated genes involved in symbiotic nitrogen fixation, strongly suggested that expression of *groESL*$_3$ in *B.japonicum* is dependent on NifA and the σ^{54} RNA polymerase.

The transcription start site of the *groESL*$_3$ operon was determined by primer extension experiments using two different 32mer oligonucleotides as primers (primers 702 and 703; see Materials and methods). Anaerobically grown cultures of the *B.japonicum* strains 7987 and A7987, which contain a *groEL*$_3$′-′*lacZ* fusion in the chromosome, were used as sources for the isolation of template RNA. Markedly increased levels of primer extension products were obtained with this type of RNA as compared with RNA isolated from the parental strains lacking the *groEL*$_3$′-′*lacZ* fusion (wild-type; A9) (data not shown). It appeared as if the *groESL*$_3$′-′*lacZ* hybrid mRNA was more stable than the genuine *groESL*$_3$ mRNA. As shown in Figure 7, both primers yielded a major extension product with the same 3′ end, and the second primer also gave rise to a product which was one nucleotide longer. These signals were observed only with RNA from strain 7987 but not with RNA isolated from the *nifA* mutant A7987, and they originated from transcripts ending with an A or a C residue that corresponded to T_{351} and G_{352} on the non-coding strand shown in Figure 6. We concluded that the −24/−12-type

Fig. 6. Nucleotide sequence of the 2527 bp *Eco*RI−*Sal*I fragment harbouring *groESL*$_3$ (cf. Figure 4). Presumptive ribosome binding sites are underlined. The bold-face nucleotides underlined with inverted horizontal arrows indicate the 9 bp stem of a putative rho-independent terminator downstream of *groEL*$_3$. Upstream of *groES*$_3$ a potential binding site for NifA (TGT-10bp-ACA) as well as the most conserved nucleotides of a −24/−12-type promoter (GG-10bp-GC) are shown in bold-face letters. Vertical arrows mark the transcription start sites of *groESL*$_3$ as determined by the primer extension experiment shown in Figure 7. The deduced amino acid sequences of $GroES_3$ and $GroEL_3$ are presented in the first line below the nucleotide sequence, whereas the second line shows those residues of $GroES_2$ and $GroEL_2$ which differ from $GroES_3$ and $GroEL_3$, respectively. This sequence has been submitted to the EMBL data library and assigned accession number Z22603.

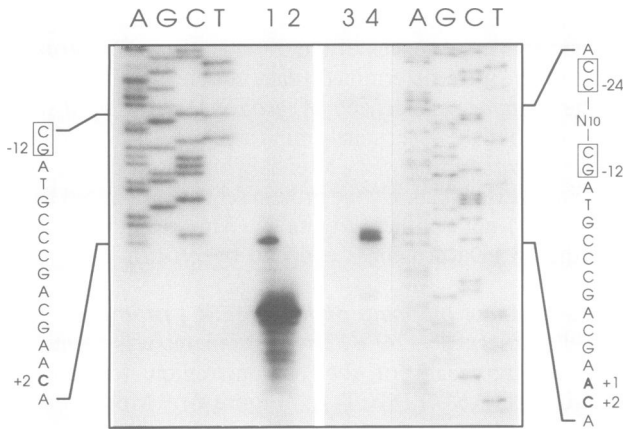


Fig. 7. Determination of the transcription start site of the *groESL*$_3$ operon. Total RNA was isolated from anaerobically grown cells of *B.japonicum* strains 7987 (lanes 1 and 4) and A7987 (lanes 2 and 3). The extension reactions were performed as described in Materials and methods using the primers 703 (lanes 1 and 2) and 702 (lanes 3 and 4). The same primers were used for the sequencing reactions shown in the left (703) and right (702) parts. The relevant sequence of the *groESL*$_3$ coding strand is shown in both margins, and the positions corresponding to the major extension products obtained with the two primers are marked with '+2' and '+1' (G_{352} and T_{351} of the non-coding strand shown in Figure 6). The dominant signals in lanes 1 and 2 originated from unextended primer 703. All samples were run side by side on the same gel. The exposure times were 15 h for the left sequence ladder and lanes 1 and 2, 7 h for lanes 3 and 4, and 4 days for the right sequence ladder.

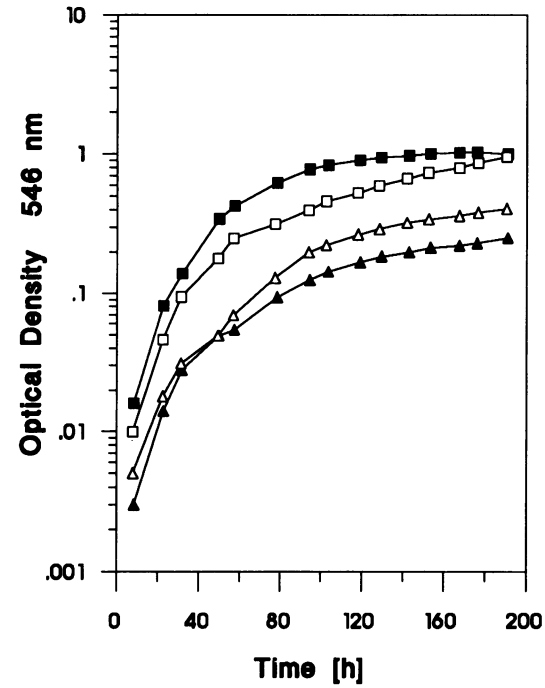


Fig. 8. Growth of the *B.japonicum* wild-type (rectangles) and the *groEL*$_4$ mutant 8528 (triangles) in minimal medium under aerobic (filled symbols) and microaerobic conditions (open symbols). The curves are based on mean values from four parallel cultures. The size of the standard deviations were smaller than the size of the symbols.

groESL$_3$ promoter as defined above was functional and directed the synthesis of a NifA-dependent transcript.

Construction and phenotypic analysis of groEL mutants

The five *groEL* genes were mutagenized individually as described in Materials and methods and as illustrated in Figure 4. For the mutagenesis of *groEL*$_1$ to *groEL*$_3$, *groEL*-internal *Xho*I fragments were replaced by a 1.6 kb *Xho*I fragment containing the *aphII* gene of Tn*5*. The *groEL*$_4$ and *groEL*$_5$ genes were mutagenized by inserting the *aphII* cassette into a *Pst*I site and an *Msc*I site, respectively. The genomic DNA structure of the *B.japonicum* mutant strains 7940 (*groEL*$_1^-$), 7942 (*groEL*$_2^-$), 8541 (*groEL*$_3^-$), 8528 (*groEL*$_4^-$) and 7996 (*groEL*$_5^-$), which resulted after marker replacement, was verified by Southern blot hybridization as shown in Figure 3B. It was thus ascertained that each individual mutant strain was affected only in the destined *groEL* gene but not in any of the remaining four *groEL* homologues.

The symbiotic properties of the *groEL* mutants were tested in plant infection tests. None of the mutant strains differed from the wild-type in the ability to nodulate soybean roots and to fix nitrogen under symbiotic conditions (data not shown). This indicated that none of the five *groEL* gene products was absolutely essential for symbiosis but that they might functionally replace each other. The mutants were further analysed for their behaviour under different growth conditions. Aerobic growth of the mutants in rich medium (PSY medium) was not affected (data not shown). However, aerobic growth of strain 8528 (*groEL*$_4^-$) in a defined minimal medium was impaired (Figure 8). This effect was specific for the *groEL*$_4$ mutant and was not observed with any of the other *groEL* mutants (data not shown). Interestingly, growth of *B.japonicum* 8528 in this medium was better under microaerobic conditions (0.56% oxygen) than in aerobic conditions, whereas the growth of the wild-type and that of the other four *groEL* mutants (data not shown) was attenuated by the switch from aerobiosis to microaerobiosis. It is possible that under aerobic conditions the reduced level of total cellular GroEL protein in the *groEL*$_4$ mutant became growth limiting, and that this limitation was partially overcome by an increased *groESL*$_3$ expression under microaerobic conditions (see below).

Crude extracts of the five *groEL* mutants were analysed by two-dimensional polyacrylamide gel electrophoresis in order to assign the originally isolated 58 kDa protein unambiguously to any of the cloned *groEL* genes. As shown in Figure 1 (panel C), extracts from anaerobically grown cells of the *groEL*$_3$ mutant 8541 lacked the 58 kDa protein whose synthesis is induced in anaerobically grown wild-type cells in a NifA-dependent manner (Figure 1, panels A and B). Analogous extracts from all other *groEL* mutants gave rise to protein patterns that were indistinguishable from that of the wild-type (data not shown). This result is fully consistent with the implication made above by the N-terminal amino acid sequence comparisons, i.e. the 58 kDa protein can be definitively regarded as the product of the *B.japonicum groEL*$_3$ gene.

Expression and regulation of groEL$_1$ *through groEL*$_5$

In order to study the expression and regulation of the various *groEL* genes, translational *lacZ* fusions were constructed as described in Materials and methods and as illustrated in Figure 4. Since it was known from previous studies that the regulation of plasmid-borne *lacZ* fusions may differ from that of chromosomally located fusions (Gubler and Hennecke, 1988; Thöny *et al.*, 1989), all of the *groEL'-'lacZ*

Table I. Expression of chromosomally integrated *B.japonicum groEL'-'lacZ* fusions in cells grown under different conditions

Strain	Relevant genotype	β-Galactosidase activity (U)[a]		
		Aerobic[b,c]	Anaerobic[b,d]	Symbiotic[e]
Bj 7979	$groEL_1$'-'*lacZ*	0.9 ± 0.2	0.4 ± 0.3	1.9 ± 0.1
Bj 7998	$groEL_2$'-'*lacZ*	74.8 ± 13.5	83.6 ± 12.6	9.8 ± 3.1
Bj 7987	$groEL_3$'-'*lacZ*	4.2 ± 1.0	1880.0 ± 349.0	890.0 ± 72.0
Bj 8548	$groEL_4$'-'*lacZ*	292.7 ± 45.4	334.4 ± 67.2	30.4 ± 3.4
Bj 7993	$groEL_5$'-'*lacZ*	1.6 ± 0.3	9.9 ± 1.7	9.8 ± 1.6
Bj A7979	$groEL_1$'-'*lacZ nifA*$^-$	1.1 ± 0.1	0.8 ± 0.6	ND
Bj A7998	$groEL_2$'-'*lacZ nifA*$^-$	74.9 ± 9.8	88.6 ± 24.6	ND
Bj A7987	$groEL_3$'-'*lacZ nifA*$^-$	4.5 ± 1.2	10.7 ± 4.7	10.3 ± 2.7
Bj A8548	$groEL_4$'-'*lacZ nifA*$^-$	274.3 ± 68.0	223.9 ± 79.8	ND
Bj A7993	$groEL_5$'-'*lacZ nifA*$^-$	1.9 ± 0.4	9.8 ± 1.6	ND
Bj N7979	$groEL_1$'-'*lacZ* $rpoN_{1/2}^-$	0.9 ± 0.1	1.4 ± 1.3	ND
Bj N7998	$groEL_2$'-'*lacZ* $rpoN_{1/2}^-$	46.7 ± 13.1	83.6 ± 12.9	ND
Bj N7987	$groEL_3$'-'*lacZ* $rpoN_{1/2}^-$	2.6 ± 1.0	9.0 ± 4.2	18.0 ± 0.9
Bj N8548	$groEL_4$'-'*lacZ* $rpoN_{1/2}^-$	190.9 ± 30.7	200.0 ± 62.9	ND
Bj N7993	$groEL_5$'-'*lacZ* $rpoN_{1/2}^-$	1.4 ± 0.3	24.2 ± 9.8	ND

[a]The wild-type strain without any *lacZ* fusion showed <2 Miller units under all growth conditions.
[b]A total of six cultures were assayed in duplicate in three independent experiments.
[c]Aerobic cultures were grown in PSY medium containing 100 μg/ml spectinomycin.
[d]Anaerobic cultures were grown in YEM medium containing 10 mM KNO_3, 2 mM L-glutamine and 100 μg/ml spectinomycin.
[e]Bacteroid suspensions obtained from individual pools of all nodules from two soybean plants were assayed in duplicate. ND, not determined.

fusions were integrated by recombination into the chromosome at their homologous sites. The *groEL'-'lacZ* fusions were integrated into the *B.japonicum* wild-type, the *nifA* mutant A9, and the $rpoN_{1/2}$ double mutant N50−97 resulting in a set of 15 different strains. Cells of each strain were grown either under aerobic, anaerobic or symbiotic conditions, and the levels of β-galactosidase activity were determined (Table I).

The five *groEL* homologues could be divided into three classes, based on the results from the expression studies. (i) The $groEL_1$ and $groEL_5$ genes were expressed not at all or only very weakly in all backgrounds under the conditions tested. Expression of $groEL_5$'-'*lacZ* under anaerobic and symbiotic conditions was slightly higher than in aerobic conditions. If this difference is significant at all, it certainly does not involve NifA or RpoN, because the small increase was also observed in the respective *nifA* or *rpoN* mutant strains. (ii) The $groEL_2$ and $groEL_4$ genes were expressed at significant levels under all growth conditions, irrespective of the mutant background. Generally, the levels of β-galactosidase activity derived from the $groEL_4$'-'*lacZ* fusion were 2- to 4-fold higher than those obtained with the $groEL_2$'-'*lacZ* fusion. (iii) $GroEL_3$ gene expression was dramatically induced in anaerobically grown cells and in bacteroids (~450- and 200-fold, respectively), and this induction was strictly dependent on *nifA* and *rpoN*. Residual plant tissue contaminants in the bacteroid preparations, which may interfere with optical density measurements, make it difficult to compare directly the absolute values of β-galactosidase activities in free-living cells with those in bacteroids. Nevertheless, it is interesting to note that the expression levels of $groEL_2$ and $groEL_4$ in bacteroids were ~10-fold lower than in anaerobically grown cells, whereas this reduction was only ~2-fold for $groEL_3$.

Expression of the *groEL'-'lacZ* fusions integrated in the wild-type chromosome was also studied under heat shock conditions as described in Table II. A shift from 28°C to 39°C resulted in a >25-fold induction of $groEL_1$'-'*lacZ* expression (strain 7979) within 200 min. This effect was specific for $groESL_1$ and indicated that this operon is under heat shock control. The levels of β-galactosidase activities given in Tables I and II are not directly comparable because the *B.japonicum* cultures were assayed in different physiological states in the respective experiments.

Table II. Effect of heat shock on the expression of chromosomally integrated *groEL'-'lacZ* fusions in aerobically grown cultures

Strain	Relevant genotype	β-Galactosidase activity (U)[a]	
		28°C	39°C
Bj 7979	$groEL_1$'-'*lacZ*	1.7 ± 0.0	46.5 ± 5.7
Bj 7998	$groEL_2$'-'*lacZ*	9.3 ± 0.4	8.4 ± 0.5
Bj 7987	$groEL_3$'-'*lacZ*	4.4 ± 0.4	4.3 ± 0.3
Bj 8548	$groEL_4$'-'*lacZ*	225.9 ± 5.7	271.5 ± 4.0
Bj 7993	$groEL_5$'-'*lacZ*	5.9 ± 0.7	3.8 ± 0.1

[a]After growth in PSY medium at 28°C to an OD_{600nm} of ~0.3, the cultures were split and grown for an additional 3 h and 20 min at 28°C and 39°C before β-galactosidase activities were determined. Two parallel cultures of each strain were assayed in duplicate.

Discussion

This is a follow-up study based on our previous observation that a number of *B.japonicum* proteins whose synthesis depends on the symbiotic regulatory protein NifA can be visualized by two-dimensional gel electrophoresis (Fischer *et al.*, 1986). A prominent 58 kDa protein of this group was chosen for further characterization, and we report here its identification as a GroEL-like protein ($GroEL_3$). Considering the ubiquity of GroEL-like proteins in nature, the existence of an analogous protein in *B.japonicum* is no surprise. It is astounding, however, that it belongs to a class of proteins whose synthesis is NifA-dependent and whose gene is transcribed from a σ^{54} promoter. In view of the

possible requirement of GroEL for NifA folding and nitrogenase assembly as reported recently for the *K.pneumoniae nif* system (Goveszensky *et al.*, 1991), it appears sensible for *B.japonicum* to co-regulate the expression of a critical chaperonin together with the *nif* system. In addition, it may also be possible that *B.japonicum* chaperonins are involved in the translocation of proteins from the endosymbiotic bacteroids into the peribacteroid space during symbiosis. Secretion of a set of symbiosis-specific proteins was in fact shown in the symbiosis between pea and *Rhizobium leguminosarum* biovar *viciae* (Katinakis *et al.*, 1988). The presence of large amounts of a protein that cross-reacted with anti-*Tetrahymena* hsp58 and anti-*E.coli* GroEL immunoglobulins was previously reported in bacteroids of *B.japonicum* (Choi *et al.*, 1991) and in nitrogen-fixing heterocysts of *Anabaena* PCC 7120, respectively (Jäger and Bergman, 1990). Similarly, prokaryotic endosymbionts present in specialized cells of aphid species synthesize *de novo* essentially only one protein (symbionin) that is structurally and functionally homologous to GroEL of *E.coli* (Hara *et al.*, 1990; Kakeda and Ishikawa, 1991; Ohtaka *et al.*, 1992). In the latter example it was proposed that the chaperonin might be involved in the assembly of polypeptides imported from the host cell into the endosymbiont. Choi *et al.* (1991) speculated that intracellular symbionts such as bacteroids are stressed by oxygen and nutrient limitation or host defence mechanisms, which may result in the induction of heat shock genes. Our results with *B.japonicum* now show for the first time that heat shock is not the only signal for induction of a GroEL-like protein: the signal for induction of $GroEL_3$ synthesis is oxygen limitation, and this regulation is mediated by the oxygen-responsive NifA protein.

The finding that the gene for the 58 kDa GroEL protein is a member of a *groESL* multigene family is also unexpected. For at least two reasons we believe that the five *groESL* operons harbour functional genes rather than pseudogenes. First, the high degree of sequence conservation within the $groESL_2$ and $groESL_3$ coding regions as opposed to the flanking or intervening non-coding regions suggests that there is selective pressure for functional gene products. The same conclusion can be drawn from the available sequence information of the remaining three *groESL* regions (data not shown). Second, the expression studies with the *groEL'-'lacZ* fusions demonstrate that all five *groEL* homologues are transcribed and translated.

It is conceivable that *groESL* gene products are also required under conditions when NifA is not active (aerobiosis), i.e. when the $groESL_3$ operon is not expressed. In this context the existence of other *groESL*-homologous genes such as $groESL_2$ and $groESL_4$, which are expressed constitutively, appears quite plausible, although it is not obvious why two operons should be required to account for a basal level of chaperonin synthesis. A third mode of expression within the *groESL* multigene family is realized in the case of the $groESL_1$ operon, which is under heat shock control. This type of modular gene regulation is entirely different from *groESL* regulation in many other microorganisms. In *E.coli* (Zhou *et al.*, 1988) or *Bacillus subtilis* (Li and Wong, 1992; Schmidt *et al.*, 1992), for example, the heat shock induction of chaperonin synthesis results from the enhanced transcription of one and the same *groESL* operon that is also used for expression under non-inducing conditions. Expression of *B.japonicum* $groESL_5$ is not yet understood completely. The data in Table I suggest that $groESL_5$ expression is regulated independently of NifA by the cellular oxygen conditions albeit rather weakly. In summary, it is attractive to speculate that the presence of five differentially regulated *groESL* operons uniquely endows *B.japonicum* with a maximal flexibility to modulate its cellular contents of chaperonins in response to varying environmental conditions and physiological needs.

The structural redundancy of the *B.japonicum groESL* genes is also reflected at the functional level. Null mutations in *groES* or *groEL* are lethal for *E.coli* (Fayet *et al.*, 1989). By contrast, a mutation in *B.japonicum* $groEL_4$, which is the predominantly expressed homologue under aerobic conditions, did not affect growth, at least in rich medium. However, growth of this mutant in minimal medium was significantly impaired. The intensive biosynthetic activity under these growth conditions might be followed by an increased demand for chaperonins, so that the level of GroEL protein becomes limiting in the $groEL_4$ mutant. This idea is supported by the partial release of the growth defect when the cells are grown in microaerobiosis, a condition that leads to strong induction of $groESL_3$ expression.

It was intriguing to learn that a mutation in $groEL_3$ did not affect symbiotic nitrogen fixation even though this gene appeared to contribute ~90% to the total *groEL* gene expression measurable under symbiotic conditions. If the GroEL chaperonins were to play an essential role in symbiosis, we must then conclude that the individual gene products can functionally replace each other, and the total level of chaperonins expressed concurrently from $groEL_2$, $groEL_4$ and possibly $groEL_5$ must be sufficient to compensate for the lack of GroEL3 protein in the $groEL_3$ mutant. The extremely high amino acid sequence similarity between all five GroEL proteins corroborates this idea.

The occurrence of multiple homologues of genes is becoming quite a familiar phenomenon in *B.japonicum*. Previously we reported on the existence of two highly conserved and differentially regulated *rpoN* genes (Kullik *et al.*, 1991) and on two homologues of the *nodD* gene (Göttfert *et al.*, 1992). In addition, Anthamatten *et al.* (1992) proposed the presence of two *fixK*-like genes in *B.japonicum*. Possibly, the large size (8.7 Mbp) of the *B.japonicum* chromosome (Kündig *et al.*, 1993) reflects this redundancy of individual genes. As far as chaperonins are concerned, it is interesting to note that two *groEL*-like genes have been found in *Streptomyces albus*, *Mycobacterium leprae* and *Synechocystis* sp. PCC 6803, but only one of them is linked to a *groES*-like gene and no differential regulation is known (Guglielmi *et al.*, 1991; Mazodier *et al.*, 1991; Rinke de Wit *et al.*, 1992; Lehel *et al.*, 1993). *R.meliloti* also appears to have two *groEL*-like genes as inferred from the deposition of two *R.meliloti groEL* nucleotide sequences in the GenBank database (E.Rusanganwa and R.S.Gupta, unpublished; see also Figure 9). The unique *B.japonicum groESL* multigene family consisting of five members is to some extent reminiscent of the essential *SSA* subfamily of at least nine *HSP70*-like genes present in yeast (Lindquist and Craig, 1988 and references therein). The four members of the *SSA* subfamily are highly similar at the nucleotide sequence level, and each of the gene products can functionally substitute at least partially for the absence of the other three. Moreover, they are differentially regulated in response to the growth

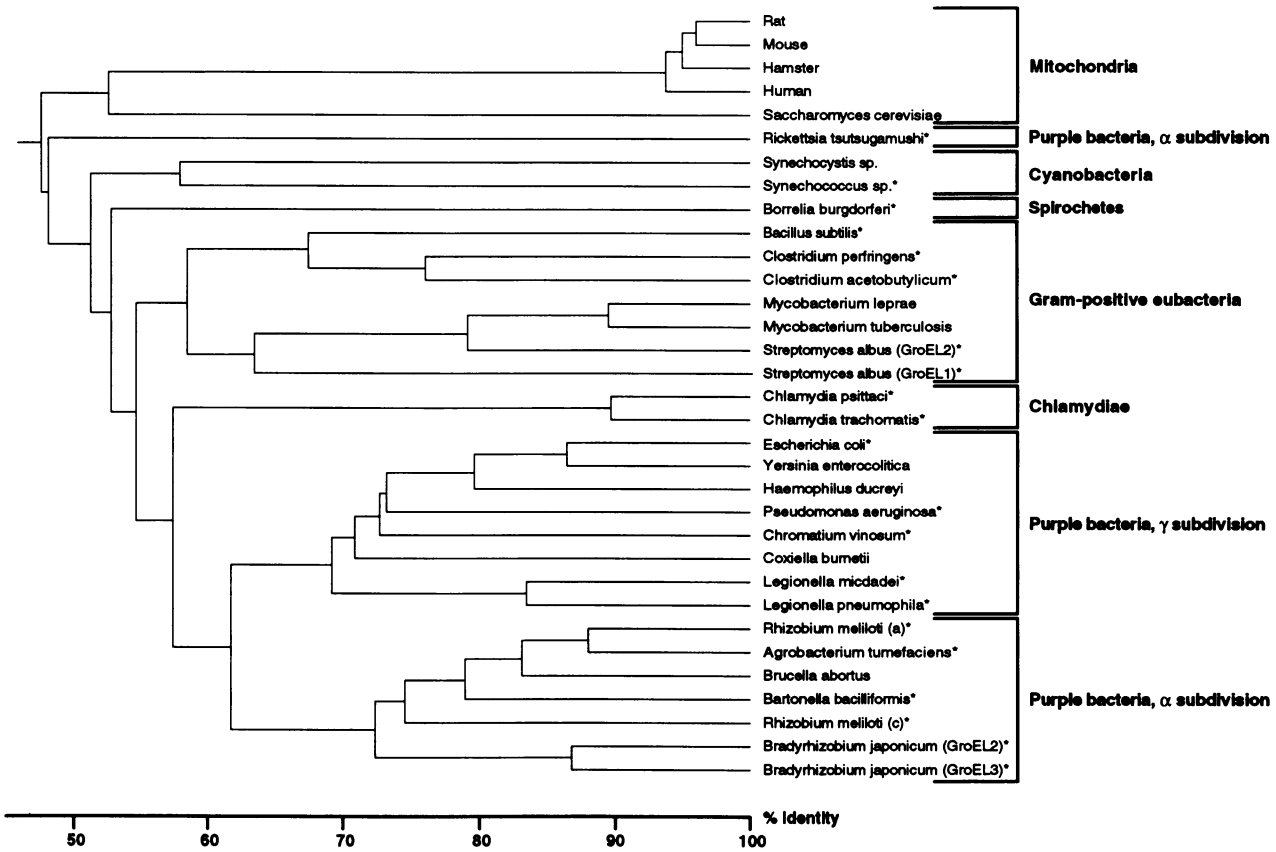


Fig. 9. Phylogenetic tree based on the similarities between the amino acid sequences of 27 eubacterial GroEL proteins and five mitochondrial Hsp60 proteins. The origin of the eukaryotic Hsp60 proteins and relevant eubacterial phyla including some of their subdivisions as defined by Woese (1987) are indicated in the right margin. Asterisks mark those eubacterial genera or species which were originally classified by Woese (1987) in the respective phylogenetic group. The indicated identity values (%) are calculated based on 568 positions (including gaps) which were compared in the alignment of all sequences. References or EMBL database accession numbers (release 33) for the Hsp60 and GroEL chaperonin sequences are as follows: rat (Peralta *et al.*, 1990), mouse (Lotscher and Allison, 1990), Chinese hamster (Picketts *et al.*, 1989), human (Jindal *et al.*, 1989), *Saccharomyces cerevisiae* (Johnson *et al.*, 1989; Reading *et al.*, 1989), *Rickettsia tsutsugamushi* (Stover *et al.*, 1990), *Synechocystis* sp. PCC 6803 (Chitnis and Nelson, 1991), *Synechococcus* sp. PCC 7942 (Webb *et al.*, 1990), *Borrelia burgdorferi* (Shanafelt *et al.*, 1991), *Bacillus subtilis* (Li and Wong, 1992; Schmidt *et al.*, 1992), *Clostridium perfringens* (Rusanganwa *et al.*, 1992), *Clostridium acetobutylicum* (Narberhaus and Bahl, 1992), *Mycobacterium leprae* (Mehra *et al.*, 1986), *Mycobacterium tuberculosis* (Shinnick, 1987), *Streptomyces albus* $GroEL_1$ and $GroEL_2$ (Mazodier *et al.*, 1991), *Chlamydia psittaci* (Morrison *et al.*, 1989), *Chlamydia trachomatis* (Morrison *et al.*, 1990), *Escherichia coli* (Hemmingsen *et al.*, 1988), *Yersinia enterocolitica* (Haefner and Roggenkamp, unpublished; X68526), *Haemophilus ducreyi* (Parsons *et al.*, 1992), *Pseudomonas aeruginosa* (Sipos *et al.*, 1991), *Chromatium vinosum* (Ferreyra *et al.*, unpublished; M99443), *Coxiella burnetii* (Vodkin and Williams, 1988), *Legionella micdadei* (Hindersson *et al.*, 1991), *Legionella pneumophila* (Sampson *et al.*, 1990), *Rhizobium meliloti* GroELa and GroELc (Rusanganwa and Gupta, unpublished; M94192, M94191), *Agrobacterium tumefaciens* (Segal and Ron, unpublished; X68263), *Brucella abortus* (Gor and Mayfield, 1992), *Bartonella bacilliformis* (Xu *et al.*, unpublished; M98257).

temperature and mutations in individual *SSA* genes were shown to influence expression of other members of this subfamily. In the light of the latter finding it would be of interest in future work to test whether or not a mutation in a given *B.japonicum groEL* gene affects expression of any of the remaining *groEL* genes.

The large number of available GroEL (Hsp60) amino acid sequences from many different organisms prompted us to exploit this information for an analysis of the phylogenetic relationship between them (Figure 9). Strikingly, the GroEL-based construction of a phylogenetic tree with 25 eubacterial species included in this study almost perfectly reflects the eubacterial phyla and subdivisions as defined by Woese (1987) who compared 16S rRNA sequences. Only the GroEL amino acid sequence of *Rickettsia tsutsugamushi* diverges quite significantly from the other members of the α subdivision of the purple bacteria. Interestingly, however, this representative of the rickettsia comes to lie closer to the branch formed by eukaryotes, which is in good agreement with the postulate that eukaryotic mitochondria originated from a member of the α subdivision of the purple bacteria (Yang *et al.*, 1985; Woese, 1987). Using fewer GroEL homologues for comparison, Gupta *et al.* (1989) also suggested a phylogenetic link between mitochondria and purple bacteria, but they did not include representatives of the α subdivision in their study. In summary, we are convinced that the ubiquity of GroEL (Hsp60)-like proteins and the moderate evolutionary divergence of their primary sequences make them very useful 'molecular chronometers' (Woese, 1987) for phylogenetic analyses.

Materials and methods

Bacterial strains and plasmids

The bacterial strains and plasmids used in this work are listed in Table III.

Media and growth of cells

For routine growth of *E.coli* cells, Luria—Bertani (LB) medium was used (Miller, 1972); for plasmid selection, it contained the following concentrations

Table III. Bacterial strains and plasmids

Strain or plasmid	Relevant phenotype or genotype	Reference or origin
Strains		
E.coli		
DH5α	*supE44 ΔlacU169* (Ψ80*lacZ*ΔM15) *hsdR17 recA1 gyrA96 thi-1 relA1*	BRL, Gaithersburg, MD
S17-1	Sm^r Sp^r *hsdR* RP4-2 *kan*::Tn*7 tet*::Mu, integrated in the chromosome	Simon *et al.* (1983)
JM101	*supE thi* Δ(*lac-proAB*) F'[*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ* ΔM15]	Messing (1983)
JM103	Sm^r *supE thi strA* Δ(*lac-proAB*) F'[*traD*36 *proAB*$^+$ *lacI*^{q} *lacZ* ΔM15]	Messing (1983)
B.japonicum		
110*spc*4	Sp^r (wild-type)	Regensburger and Hennecke (1983)
A9	Sp^r Km^r *nifA*::*aphII*	Fischer *et al.* (1986)
N50-97	Sp^r Km^r Sm^r *rpoN*$_1$::*aphII rpoN*$_2$::Ω	Kullik *et al.* (1991)
7940	Sp^r Km^r *groEL*$_1$::*aphII*	This work
7942	Sp^r Km^r *groEL*$_2$::*aphII*	This work
8541	Sp^r Km^r *groEL*$_3$::*aphII*-PSP	This work
8528	Sp^r Km^r *groEL*$_4$::*aphII*-PSP	This work
7996	Sp^r Km^r *groEL*$_5$::*aphII*-PSP	This work
7979[a]	Sp^r Tc^r *groEL*$_1$'-'*lacZ* chromosomally integrated	This work
7998[a]	Sp^r Tc^r *groEL*$_2$'-'*lacZ* chromosomally integrated	This work
7987[a]	Sp^r Tc^r *groEL*$_3$'-'*lacZ* chromosomally integrated	This work
8548[a]	Sp^r Tc^r *groEL*$_4$'-'*lacZ* chromosomally integrated	This work
7993[a]	Sp^r Tc^r *groEL*$_5$'-'*lacZ* chromosomally integrated	This work
Plasmids		
M13mp18		Norrander *et al.* (1983)
M13mp19		Norrander *et al.* (1983)
M13BM20		Boehringer, Mannheim, Germany
M13BM21		Boehringer, Mannheim, Germany
pUC18	Ap^r	Norrander *et al.* (1983)
pSUP202	Ap^r Cm^r Tc^r *oriT* from RP4	Simon *et al.* (1983)
pSUP202-P	Cm^r Tc^r (pSUP202) *Pac*I linker in *Nde*I site	Kündig *et al.* (1993)
pSUP202pol4	Tc^r (pSUP202) part of polylinker from pBluescript II KS+ between *Eco*RI and *Pst*I	H.M.Fischer, unpublished
pNM481	Ap^r (pUC8) *'lacZ*	Minton (1984)
pNM482	Ap^r (pUC8) *'lacZ*	Minton (1984)
pUC-4-KIXX	Ap^r Km^r (pUC4) *lacZ'*::*aphII*	Pharmacia LKB, Uppsala, Sweden
pUC-4-KIXX-PSP	Ap^r Km^r (pUC-4-KIXX) *Pme*I-*Swa*I-*Pac*I linker in *Sma*I site of *ble*	Kündig *et al.* (1993)
pND5	Ap^r Tc^r (pBR325) *E.coli groESL*	Jenkins *et al.* (1986)
pRJ7931	Ap^r (pUC18) *B.japonicum groESL*$_1$	This work
pRJ7928	Ap^r (pUC18) *B.japonicum groESL*$_2$	This work
pRJ7932	Ap^r (pUC18) *B.japonicum groESL*$_3$	This work
pRJ7934	Ap^r (pUC18) *B.japonicum groESL*$_4$	This work
pRJ7980	Ap^r (pUC18) *B.japonicum groEL*$_5$	This work

[a]The same *groEL'-'lacZ* fusions were also integrated into the chromosome of *B.japonicum* A9 and N50−97. The resulting strains were given the same numbers preceded by 'A' (A9 derivatives) or 'N' (N50−97 derivatives); see Table I.

of antibiotics (in μg/ml): ampicillin, 200; chloramphenicol, 20; kanamycin, 30; streptomycin, 50; and tetracycline, 10. Aerobic cultures of *B.japonicum* were grown in PSY medium (Regensburger and Hennecke, 1983), whereas YEM medium supplemented with 10 mM KNO_3 and 2 mM L-glutamine was used for anaerobic *B.japonicum* cultures (Daniel and Appleby, 1972). Anaerobic cultures (10 ml volume) were grown under argon in 100 ml serum bottles. The growth characteristics of *B.japonicum groEL* mutants were determined in the derepression minimal medium described by Regensburger *et al.* (1986). The concentrations of antibiotics in *B.japonicum* cultures were as follows (in μg/ml): spectinomycin, 100; kanamycin, 100; streptomycin, 100; tetracycline, 100 (solid media) or 50 (liquid media).

DNA manipulations

Recombinant DNA work was done using standard protocols (Sambrook *et al.*, 1989). To determine nucleotide sequences, appropriate DNA fragments were cloned into suitable bacteriophage M13 vectors, and single-stranded DNA derived therefrom was sequenced by the chain termination method (Sanger *et al.*, 1977). The nucleotide sequences of the 5' ends of *groEL*$_1$, *groEL*$_4$ and *groEL*$_5$ were determined by double-strand DNA sequence analysis using suitable oligonucleotides as primers. For computer-assisted DNA and protein sequence analyses the software package (release 6.2) of the UWGCG (Genetics Computer Group of the University of Wisconsin, Madison WI) and the program CLUSTAL (Higgins and Sharp, 1988) were used. GroEL and Hsp60 amino acid sequences used for sequence alignments were retrieved from the EMBL (release 33) and Swissprot (release 24.0) databases, respectively.

Two-dimensional gel electrophoresis

Two-dimensional gel electrophoresis of extracts from anaerobically grown *B.japonicum* cells was done as described by Scott *et al.* (1979) and O'Farrell (1975) with modifications given by Regensburger *et al.* (1986). Non-radioactively labelled proteins were visualized by Coomassie Blue staining or rapid silver staining (Bloom *et al.*, 1987).

Isolation and N-terminal sequencing of the 58 kDa protein

The 58 kDa *B.japonicum* protein was electroeluted from 12 two-dimensional gels using a Bio-Rad electroelution apparatus (model 442) provided with membranes of a molecular weight cut-off between 12 000 and 15 000 (Hunkapiller *et al.*, 1983). Electroelution was done in 50 mM ammonium bicarbonate, 0.1% SDS for 4 h at ~8 mA/membrane. After concentration of the eluate by evaporation in a vacuum centrifuge, the sample was extracted with cold methanol in order to remove most of the Coomassie stain. Approximately 150 pmol of the isolated protein were used for N-terminal sequencing with a gas-phase protein sequencer (model 407A, Applied Biosystems, Foster City, CA, USA).

DNA probes and hybridization conditions

A 1286 bp *Bst*BI−*Bam*HI fragment isolated from plasmid pND5 (Jenkins *et al.*, 1986) was radioactively labelled by nick translation and used as an *E.coli groEL*-internal probe in heterologous hybridizations. The sequences of the 17mer degenerate oligonucleotides EL1 and EL3 were derived from highly conserved GroEL domains and adapted to the *B.japonicum* codon preference (Ramseier and Göttfert, 1991). Oligonucleotide EL1 (5′-GA$^{G}/_{A}$GG$^{C}/_{I}$ATGCA$^{G}/_{A}$TT$^{C}/_{T}$GA-3′; 16 variants) is based on positions 1041−1057 of the non-coding strand of *E.coli groEL* which correspond to amino acids Glu191 to Asp196 (Hemmingsen *et al.*, 1988). Oligonucleotide EL3 (5′-TT$^{C}/_{T}$TC$^{G}/_{I}$CG$^{G}/_{A}$TC$^{G}/_{A}$TA$^{G}/_{A}$ TC-3′; 32 variants) is derived from the *E.coli groEL* coding strand positions 1561−1545 (amino acids Lys364 to Asp359). For probing the initial Southern blots of *B.japonicum* genomic DNA as well as the *E.coli* colony libraries, the oligonucleotides were end-labelled with [^{32}P]ATP using polynucleotide kinase. In subsequent hybridization experiments, a 521 bp *B.japonicum* *groEL*$_3$-internal DNA fragment was used. It was synthesized by PCR using EL1 and EL3 as primers, and labelled non-radioactively by the incorporation of digoxigenin-11-dUTP. Hybridizations with oligonucleotides EL1 and EL3 were performed at 40−42°C in 6 × SSC (1 × SSC is 150 mM NaCl plus 15 mM sodium citrate), 0.5% SDS, 20 mM sodium phosphate buffer (pH 6.5), 5 × Denhardt's solution (Denhardt, 1966), 200 μg/ml sonicated salmon sperm DNA, followed by three washes in 6 × SSC at increasing temperatures varying from 40 to 50°C. For hybridizations with the *E.coli* '*groEL*' or *B.japonicum* '*groEL*$_3$' probes, the salt conditions were changed to 5 × SSC, and the temperature for both hybridizations and washes was shifted to 58−62°C.

Transcript mapping

The transcriptional start site of the *groESL*$_3$ operon was mapped in a primer extension experiment using two different 32mer oligonucleotides complementary to sequence positions 359−390 (primer 703) and 457−488 (primer 702) (see Figure 6) (Sambrook *et al.*, 1989). Template RNA was isolated from anaerobically grown *B.japonicum* strains 7987 and A7987 (Kullik *et al.*, 1989), and 400 units of SuperScript reverse transcriptase (Life Technologies, Gaithersburg, MD) were used for the extension reactions which were performed for 1 h at 42°C.

Construction of B.japonicum groEL mutant strains

To construct the *groEL* mutations a 1.6 kb *aphII* cassette (Km^{r}) isolated from pUC-4-KIXX or pUC-4-KIXX-PSP was inserted into all of the five *B.japonicum groEL* genes using suitable restriction sites (see Figure 4). In the *B.japonicum groEL* mutants 7940 (*groEL*$_1^-$), 7942 (*groEL*$_2^-$) and 8541 (*groEL*$_3^-$), *groEL*-internal *Xho*I fragments were replaced by the *aphII* cassette present on a 1.6 kb *Xho*I fragment. Mutants 8528 (*groEL*$_4^-$) and 7996 (*groEL*$_5^-$) were constructed by inserting the *aphII* cassette as a 1.6 kb *Sma*I fragment into a *groEL*$_4$-internal *Pst*I site (made blunt with bacteriophage T4 DNA polymerase) and a *groEL*$_5$-internal *Msc*I site, respectively. Appropriate DNA fragments containing the mutated *groEL* genes were cloned into the vector pSUP202pol4 and mobilized into *B.japonicum* 110*spc*4 as described previously (Hahn and Hennecke, 1984). Co-integrate-containing exconjugants (resulting from single cross-over) were distinguished from true marker exchange mutants (resulting from double cross-over) by their tetracycline resistance provided by the vector pSUP202pol4. The correct genomic structure of all mutant strains was confirmed by Southern blot analyses of genomic DNAs.

Construction of chromosomally integrated groEL′-′lacZ fusions

Translational *lacZ* fusions to all five *groEL* genes were constructed by making use of *groEL*-internal restriction sites (see Figure 4) and the *lacZ* fusion vectors pNM481 (*groEL*$_5$) and pNM482 (*groEL*$_1$ to *groEL*$_4$). The *groEL*$_2$, *groEL*$_3$ and *groEL*$_4$ genes were fused at homologous *Sal*I sites corresponding to Val20 in the respective GroEL proteins. The fusion to *groEL*$_1$ was made at a central *Pst*I site and that to *groEL*$_5$ at a *Bgl*II site in the 3′ region. The *lacZ* fusions including sufficiently long stretches of *B.japonicum* upstream DNA were cloned into pSUP202pol4 and conjugated into *B.japonicum* 110*spc*4 (wild-type), A9 (*nifA*$^-$) as well as N50−97 (*rpoN*$_{1/2}^-$). By plating the exconjugants on tetracycline-containing plates, those clones were selected which contained the entire *lacZ* fusion plasmids integrated in the *B.japonicum* chromosome at the homologous position (single cross-over). The genomic structure of all strains containing *groEL′-′lacZ* fusions was verified by Southern blot analyses of genomic DNAs.

β-Galactosidase assays

B.japonicum cells to be assayed for β-galactosidase activity were grown aerobically in PSY medium or anaerobically in supplemented YEM medium for 4 and 7 days, respectively. Spectinomycin was the only antibiotic used in these cultures. Determination of β-galactosidase activities in free-living cells and in bacteroids was done as described (Miller, 1972; Thöny *et al.*, 1987; Gubler and Hennecke, 1988).

Plant infection test

The symbiotic phenotype of the *B.japonicum groEL* mutants was determined in a soybean plant infection test as described previously (Hahn and Hennecke, 1984; Göttfert *et al.*, 1990).

Acknowledgements

We thank P.Brouwer for skilful technical assistance. S.Fischer and Mrs Galaboff are greatly acknowledged for their contributions to the N-terminal protein microsequencing. This work was supported by grants from the Swiss National Foundation for Scientific Research, the Federal Institute of Technology, Zürich, and the Human Frontier Science Program.

References

Anthamatten,D., Scherb,B. and Hennecke,H. (1992) *J. Bacteriol.*, **174**, 2111−2120.
Bloom,H., Beier,H. and Gross,H.S. (1987) *Electrophoresis*, **8**, 93−99.
Chitnis,P.R. and Nelson,N. (1991) *J. Biol. Chem.*, **266**, 58−65.
Choi,E.Y., Ahn,G.S. and Jeon,K.W. (1991) *BioSystems*, **25**, 205−212.
Craig,N.L. and Nash,H.A. (1984) *Cell*, **39**, 707−716.
Daniel,R.M. and Appleby,C.A. (1972) *Biochim. Biophys. Acta*, **275**, 347−354.
Denhardt,D.T. (1966) *Biochem. Biophys. Res. Commun.*, **23**, 641−646.
Ellis,R.J. (1990) *Sem. Cell Biol.*, **1**, 1−9.
Ellis,R.J. and van der Vies,S. (1991) *Annu. Rev. Biochem.*, **60**, 321−347.
Fayet,O., Ziegelhoffer,T. and Georgopoulos,C. (1989) *J. Bacteriol.*, **171**, 1379−1385.
Fischer,H.M. and Hennecke,H. (1987) *Mol. Gen. Genet.*, **209**, 621−626.
Fischer,H.M., Alvarez-Morales,A. and Hennecke,H. (1986) *EMBO J.*, **5**, 1165−1173.
Fischer,H.M., Bruderer,T. and Hennecke,H. (1988) *Nucleic Acids Res.*, **16**, 2207−2224.
Fisher,R.F. and Long,S.R. (1992) *Nature*, **357**, 655−660.
Georgopoulos,C. and Ang,D. (1990) *Sem. Cell Biol.*, **1**, 19−25.
Gething,M.J. and Sambrook,J. (1992) *Nature*, **355**, 33−45.
Gor,D. and Mayfield,J.E. (1992) *Biochim. Biophys. Acta*, **1130**, 120−122.
Govezensky,D., Greener,T., Segal,G. and Zamir,A. (1991) *J. Bacteriol.*, **173**, 6339−6346.
Göttfert,G., Hitz,S. and Hennecke,H. (1990) *Mol. Plant−Microbe Interact.*, **3**, 308−316.
Göttfert,M., Holzhäuser,D., Bäni,D. and Hennecke,H. (1992) *Mol. Plant−Microbe Interact.*, **5**, 257−265.
Gubler,M. and Hennecke,H. (1988) *J. Bacteriol.*, **170**, 1205−1214.
Guglielmi,G., Mazodier,P., Thompson,C.J. and Davies,J. (1991) *J. Bacteriol.*, **173**, 7374−7381.
Gupta,R.S., Picketts,D.J. and Ahmad,S. (1989) *Biochem. Biophys. Res. Commun.*, **163**, 780−787.
Hahn,M. and Hennecke,H. (1984) *Mol. Gen. Genet.*, **193**, 46−52.
Hahn,M., Meyer,L., Studer,D., Regensburger,B. and Hennecke,H. (1984) *Plant Mol. Biol.*, **3**, 159−168.
Hara,E., Fukatsu,T., Kakeda,K., Kengaku,M., Ohtaka,C. and Ishikawa,H. (1990) *Symbiosis*, **8**, 271−283.
Hartl,F.U., Martin,J. and Neupert,W. (1992) *Annu. Rev. Biophys. Biomol. Struct.*, **21**, 293−322.
Hawkins,F.K.L. and Johnston,A.W.B. (1988) *Mol. Microbiol.*, **2**, 331−337.
Hemmingsen,S.M., Woolford,C., van der Vies,S., Tilly,K., Dennis,D.T., Georgopoulos,C.P., Hendrix,R.W. and Ellis,R.J. (1988) *Nature*, **333**, 330−334.
Hennecke,H. (1990) *FEBS Lett.*, **268**, 422−426.
Higgins,D.G. and Sharp,P.M. (1988) *Gene*, **73**, 237−244.
Hindersson,P., Hoiby,N. and Bangsborg,J. (1991) *FEMS Microbiol. Lett.*, **77**, 31−38.
Hirsch,A.M. and Smith,C.A. (1987) *J. Bacteriol.*, **169**, 1137−1146.
Hunkapillar,M., Lujan,E., Ostrander,F. and Hood,L.E. (1983) *Methods Enzymol.*, **91**, 227−236.
Jäger,K.M. and Bergman,B. (1990) *Planta*, **183**, 120−125.
Jenkins,A.J., March,J.B., Oliver,I.R. and Masters,M. (1986) *Mol. Gen. Genet.*, **202**, 446−454.
Jindal,S., Dudani,A.K., Singh,B., Harley,C.B. and Gupta,R.S. (1989) *Mol. Cell. Biol.*, **9**, 2279−2283.

Johnson,R.B., Fearon,K., Mason,T. and Jindal,S. (1989) *Gene*, **84**, 295–302.
Kakeda,K. and Ishikawa,H. (1991) *J. Biochem.*, **110**, 583–587.
Katinakis,P., Klein-Lankhorst,R.M., Louwerse,J., van Kammen,A. and van den Bos,R.C. (1988) *Plant Mol. Biol.*, **11**, 183–190.
Kullik,I., Hennecke,H. and Fischer,H.M. (1989) *Arch. Microbiol.*, **151**, 191–197.
Kullik,I., Fritsche,S., Knobel,H., Sanjuan,J., Hennecke,H. and Fischer,H.M. (1991) *J. Bacteriol.*, **173**, 1125–1138.
Kündig,C., Hennecke,H. and Göttfert,M. (1993) *J. Bacteriol.*, **175**, 613–622.
Langer,T., Pfeifer,G., Martin,J., Baumeister,W. and Hartl,F.U. (1992) *EMBO J.*, **11**, 4757–4765.
Lehel,C., Los,D., Wada,H., Gyorgyei,J. Horvath,I., Kovacs,E., Murata,N. and Vigh,L. (1993) *J. Biol. Chem.* **268**, 1799–1804.
Li,M. and Wong,S.L. (1992) *J. Bacteriol.*, **174**, 3981–3992.
Lindquist,S. and Craig,E.A. (1988) *Annu. Rev. Genet.* **22**, 631–677.
Lotscher,E. and Allison,J.P. (1990) *Nucleic Acids Res.*, **18**, 7153.
Martin,G.B., Chapman,K.A. and Chelm,B.K. (1988) *J. Bacteriol.*, **170**, 5452–5459.
Mazodier,P., Guglielmi,G., Davies,J. and Thompson,C.J. (1991) *J. Bacteriol.*, **173**, 7382–7386.
Mehra,V., Sweeter,D. and Young,R.A. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 7013–7017.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20–78.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Minton,N.P. (1984) *Gene*, **31**, 269–273.
Morett,E. and Buck,M. (1989) *J. Mol. Biol.*, **210**, 65–77.
Morett,E., Fischer,H.M. and Hennecke,H. (1991) *J. Bacteriol.*, **173**, 3478–3487.
Morrison,R.P., Belland,R.J., Lyng,K. and Caldwell,H.D. (1989) *J. Exp. Med.*, **170**, 1271–1283.
Morrison,R.P., Su,H., Lyng,K. and Yuan,Y. (1990) *Infect. Immun.*, **58**, 2701–2705.
Murphy,P.J., Heycke,N., Trenz,S.T., Ratet,P., de Bruijn,F.J. and Schell,J. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 9133–9137.
Narberhaus,F. and Bahl,H. (1992) *J. Bacteriol.*, **174**, 3282–3289.
Norrander,J., Kempe,T. and Messing,J. (1983) *Gene*, **26**, 101–106.
O'Farrell,P.H. (1975) *J. Biol. Chem.*, **230**, 4007–4021.
Ohtaka,C. Nakamura,H. and Ishikawa,H. (1992) *J. Bacteriol.*, **174**, 1869–1874.
Parniske,M., Fischer,H.M., Hennecke,H. and Werner,D. (1991) *Z. Naturforsch.*, **46c**, 318–320.
Parsons,L.M., Waring,A.L. and Shayegani,M. (1992) *Infect. Immun.*, **60**, 4111–4118.
Peralta,D., Hartman,D.J., McIntosh,A.M., Hoogenraad,N.J. and Hoj,P.B. (1990) *Nucleic Acids Res.*, **18**, 7162.
Picketts,D.J., Mayanil,C.S.K. and Gupta,R.S. (1989) *J. Biol. Chem.*, **264**, 12001–12008.
Ramseier,T.M. and Göttfert,M. (1991) *Arch. Microbiol.*, **156**, 270–276.
Reading,D.S., Hallberg,R.L. and Myers,A.M. (1989) *Nature*, **337**, 655–659.
Regensburger,B. and Hennecke,H. (1983) *Arch. Microbiol.*, **135**, 103–109.
Regensburger,B., Meyer,L., Filser,M., Weber,J., Studer,D., Lamb,J.W., Fischer,H.M., Hahn,M. and Hennecke,H. (1986) *Arch. Microbiol.*, **144**, 355–366.
Rinke de Wit,T.F., Bekelie,S., Osland,A., Miko,T.L., Hermans,P.W.M., van Soolingen,D., Drijfhout,J.W., Schöningh,R., Janson,A.A.M. and Thole,J.E.R. (1992) *Mol. Microbiol.*, **6**, 1995–2007.
Rusanganwa,E., Singh,B. and Gupta,R.S. (1992) *Biochim. Biophys. Acta*, **1130**, 90–94.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual.* Second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sampson,J.S., O'Connor,S.P., Holloway,B.P., Plikaytis,B.B., Carlone,G.M. and Mayer,L.W. (1990) *Infect. Immun.*, **58**, 3154–3157.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
Sanjuan,J. and Olivares,J. (1989) *J. Bacteriol.*, **171**, 4154–4161.
Schmidt,A., Schiesswohl,M., Völker,U., Hecker,M. and Schumann,W. (1992) *J. Bacteriol.*, **174**, 3993–3999.
Scott,D.B., Hennecke,H. and Lim,S.T. (1979) *Biochim. Biophys. Acta*, **565**, 365–378.
Shanafelt,M.-C., Hindersson,P., Soderberg,C., Mensi,N., Turck,C.W., Webb,D., Yssel,H. and Peltz,G. (1991) *J. Immunol.*, **146**, 3985–3992.
Shinnick,T.M. (1987) *J. Bacteriol.*, **169**, 1080–1088.
Simon,R., Priefer,U. and Pühler,A. (1983) In Pühler,A. (ed.), *Molecular Genetics of the Bacteria—Plant Interaction.* Springer, Berlin, pp. 98–106.
Sipos,A., Klocke,M. and Frosch,M. (1991) *Infect. Immun.*, **59**, 3219–3226.
Stover,C.K., Marana,D.P., Dasch,G.A. and Oaks,E.V. (1990) *Infect. Immun.*, **58**, 1360–1368.
Studer,D., Gloudemans,T., Franssen,H.J., Fischer,H.M., Bisseling,T. and Hennecke,H. (1987) *Eur. J. Cell Biol.*, **45**, 177–184.
Thöny,B., Fischer,H.M., Anthamatten,D., Bruderer,T. and Hennecke,H. (1987) *Nucleic Acids Res.*, **15**, 8479–8499.
Thöny,B., Anthamatten,D. and Hennecke,H. (1989) *J. Bacteriol.*, **171**, 4162–4169.
Vodkin,M.H. and Williams,J.C. (1988) *J. Bacteriol.*, **170**, 1227–1234.
Webb,R., Reddy,K.J. and Sherman,L.A. (1990) *J. Bacteriol.*, **172**, 5079–5088.
Woese,C.R. (1987) *Microbiol. Rev.*, **51**, 221–271.
Yang,D., Oyaizu,Y., Oyaizu,H., Olson,G.J. and Woese,C.R. (1985) *Proc. Natl Acad. Sci. USA*, **82**, 4443–4447.
Zeilstra-Ryalls,J., Fayet,O. and Georgopoulos,C. (1991) *Annu. Rev. Microbiol.*, **45**, 301–325.
Zhou,Y.N., Kusukawa,N., Erickson,J.W., Gross,C.A. and Yura,T. (1988) *J. Bacteriol.*, **170**, 3640–3649.

Received on March 22, 1993