

Cloning, Nucleotide Sequencing, and Expression of the *Azospirillum brasilense lon* Gene: Involvement in Iron Uptake

ELENA MORI, MONICA FULCHIERI, CRISTINA INDORATO, RENATO FANI,
AND MARCO BAZZICALUPO*

Department of Animal Biology and Genetics, University of Florence, Florence,
I-50125 Italy

Received 6 October 1995/Accepted 12 April 1996

The *lon* gene of *Escherichia coli* encodes the Lon (La) protease, which is associated with cellular protein degradation. A *lon* gene homolog from *Azospirillum brasilense*, a nitrogen-fixing soil bacterium which lives in association with the roots of cereal grasses, was cloned and characterized. The nucleotide sequence of the *A. brasilense lon* gene was determined. It contains an open reading frame that encodes a protein of 810 amino acids with a predicted molecular mass of about 90 kDa. The deduced amino acid sequence showed a high level of homology with the sequences of all the known *lon* gene products. An open reading frame homologous to the *E. coli clpX* gene was found in front of the *lon* gene. Transcriptional analysis showed that the *lon* gene of *A. brasilense* is induced by heat shock and that the mRNA is monocistronic. An *A. brasilense* mutant, with Tn5 inserted in the *lon* gene, was shown to be defective in iron uptake and failed to express two membrane proteins that are induced by iron starvation in the parental strain.

ATP-dependent proteases are known to be involved in the regulation of the amounts of several proteins (14). In addition, many damaged or abnormal proteins are subjected to ATP-dependent proteolytic degradation. During various stresses, increasing amounts of misfolded and damaged proteins may accumulate; hence, the ATP-dependent proteases are also important during stress (14, 21). The *lon* gene of *Escherichia coli* encodes the ATP-dependent protease Lon (or La) (6, 12, 32, 33). The *lon* gene belongs to the *E. coli* heat shock regulon, whose transcription is increased several times upon heat induction through the action of the heat-shock-specific sigma factor σ^{32} (20, 35).

The biological functions of Lon have not yet been fully clarified. In *E. coli* it catalyzes the turnover of some proteins, the best studied being SulA and RcsA (14, 26, 27), which are involved in cell septation and capsule synthesis, respectively. In addition, Lon promotes the degradation of abnormal proteins (21). Lon's preference for abnormal protein substrates, together with its induction by heat and other stresses, led to the hypothesis, not yet proved, however, that Lon can help to eliminate stress-denatured proteins (21).

Since *E. coli* Lon protease plays such an important, but not completely understood, role in protein metabolism and in the regulation of various cellular processes, it is of interest to examine whether the enzyme and its function are present in other organisms. From *Bacillus brevis*, a gene homologous to the *E. coli lon* gene was cloned and characterized (16). The *B. brevis lon* gene product did not seem to be heat shock induced; however, it appeared to play a role similar to that of *E. coli* Lon in the degradation of abnormal and short-lived regulatory proteins (16). In *Bacillus subtilis*, a *lon*-homologous gene is induced by heat shock and other stresses (22). Interestingly, two genes (*lonV* and *lonD*) homologous to the *E. coli lon* gene have

been identified in the gram-negative bacterium *Myxococcus xanthus* (10, 29, 30). In contrast to that of *E. coli*, the *M. xanthus lonV* gene was shown to be essential for cell growth, since a null mutant could not be isolated (29). A gene homologous to *E. coli lon* has recently been cloned from the yeast *Saccharomyces cerevisiae*; its product is mitochondrially localized and plays an essential role in protein degradation in mitochondrial homeostasis (28). The identification of a human cDNA which encodes a Lon-like protein (31) further supports the notion that proteases similar to Lon may play central roles in almost all organisms.

From the associative nitrogen-fixing bacterium *Azospirillum brasilense* SPF94, we previously isolated, after Tn5 mutagenesis, mutants with impaired abilities to form orange haloes on Chrome-Azurool-S (CAS) plates (18), a phenotype that indicates some defect in the ability to take up iron (25). When cells of one of these mutants, referred to as SPFS5, were grown in iron-free medium, they showed a reduced growth rate and a smaller final yield compared with those of the parental strain while they produced catechol, a putative siderophore, at a level comparable to that of the parental strain; moreover, they were less able to take up ^{55}Fe from the medium (18). Southern blot analysis demonstrated that the mutant had a single Tn5 insertion in a 13-kb *EcoRI* fragment (18). In this paper we describe the further characterization of the mutant SPFS5, as well as the cloning, nucleotide sequencing, and transcriptional control of a gene of *A. brasilense* homologous to the *E. coli lon* gene, which is presumably involved in iron uptake and heat shock response.

MATERIALS AND METHODS

Bacteria, plasmids, media, and general microbiological methods. Strains and plasmids used are listed in Table 1. *A. brasilense* strains were grown at 33°C (unless otherwise indicated below) in Luria broth (23) or in MSP medium (5), in which FeCl_3 was replaced with Fe-EDTA (50 μM) or, under conditions of iron starvation, with 2-2'-dipyridyl (70 μM). *E. coli* was grown at 37°C in Luria broth. Transformation was performed by the CaCl_2 method as described previously (23). Conjugation was carried out as previously described (8). CAS plates for the detection of iron uptake were prepared as previously described (25).

Heat shock and UV sensitivity. The heat shock treatment was achieved by rapidly shifting 0.5 ml of MSP-grown, log-phase cultures to a water bath at

* Corresponding author. Mailing address: Department of Animal Biology and Genetics, University of Florence, Via Romana 17, Florence, I-50125 Italy. Phone: 39-55-222498. Fax: 39-55-222565. Electronic mail address: marcobazzi@dbag.unifi.it.

TABLE 1. Strains and plasmids used

Strain or plasmid	Relevant feature(s)	Origin or reference
Strain		
<i>A. brasilense</i> SPF94	Parental strain, <i>rif</i>	This laboratory (8)
<i>A. brasilense</i> SPFS5	Tn5 induced, iron uptake defect, <i>lon</i>	This laboratory (18)
<i>A. brasilense</i> SPFS55	SPFS5(pAF371), the phenotype is that of the parental strain	This work
<i>A. brasilense</i> SPFS56	SPFS5(pAF371), the phenotype is that of the parental strain	This work
<i>E. coli</i> DH5 α	Cloning strain	15
Plasmid		
pUC18	Cloning vector	34
pAF300	Conjugative plasmid suitable for cloning in <i>A. brasilense</i>	8
pAF371	pAF300 derivative carrying a 3.2-kb <i>Pst</i> I- <i>Dra</i> I fragment containing the <i>lon</i> gene of <i>A. brasilense</i>	This work
pAFC1	pGEM7 (Promega, Madison, Wis.) derivative carrying a 900-bp <i>Sph</i> I- <i>Sa</i> II fragment (PC1) containing part of the 3' moiety of the <i>A. brasilense lon</i> gene (Fig. 2)	This work
pAFS1	pUC18 with a 4.3-kb <i>Sa</i> II fragment of SPFS5 containing fractions of the <i>lon</i> gene and of Tn5	This work
pAFS2	pUC18 with the 3.4-kb <i>Sa</i> II fragment of SPF94 containing a fraction of the <i>lon</i> gene	This work
pCEK1	pUC18 with a 5.5-kb <i>Pst</i> I fragment of SPF94 containing the entire <i>lon</i> gene	This work

different temperatures and for the times reported in Results. The cell suspension was then incubated for 15 min at 33°C and diluted for plate counting. Bacteria grown in Luria broth, washed, and resuspended in MSP medium up to 2×10^7 cells per ml were assayed for resistance to UV light. One milliliter was spread in an empty petri dish and irradiated in the dark under a Philips 15-W UV germicidal lamp at a distance of 40 cm. After treatment, cells were kept in the dark for 1 h and then diluted for plate counting.

Membrane proteins. *A. brasilense* cells were grown in MSP without iron up to the end of the log phase, diluted to an optical density at 590 nm (OD₅₉₀) of 0.2 in MSP medium with iron (50 μ M) or with 2-2'-dipyridyl (70 μ M) and left to grow for 15 to 20 h. Membrane proteins from 50-ml cultures were extracted essentially as previously described (3, 17). Protein concentrations were measured by the Bradford method with the protein assay reagent (Pierce Chemical Co., Rockford, Ill.). Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with 11% acrylamide and Tris-glycine SDS buffer at pH 8.3 and run at room temperature for 17 h at 15 mA as described previously (17). High molecular weight standard (Sigma Chemical Co., St. Louis, Mo.) was used as the marker.

Analysis of transcription. Total RNAs of the *A. brasilense* strains were isolated from cells (grown as reported in Results) by CsCl centrifugation as described previously (19). RNA concentrations were estimated spectrophotometrically, and their integrity was checked on agarose gel.

Northern (RNA) blotting was carried out as previously described (23); samples were run on 1% agarose, blotted on nylon membranes (Hybond; Amersham, Little Chalfont, United Kingdom), and hybridized to a digoxigenin-labeled RNA probe as instructed by the manufacturer (Boehringer, Mannheim, Germany). Detection of hybridization signals was carried out by the chemiluminescence method described in the Boehringer manual of DNA labeling and detection. An RNA probe for *A. brasilense lon* was synthesized in vitro with SP6 RNA polymerase, with a digoxigenin RNA labeling kit (Boehringer), from linearized plasmid pAFC1. Primer extension was performed according to the method of Ausubel et al. (2) with 20 μ g of RNA and 2 pmol (about 5 μ Ci) of primer, and the extended primer was end labeled at the 5' end with T4 polynucleotide kinase and [γ -³²P]ATP. Extension was carried out for 30 min at 48°C with Moloney murine leukemia virus reverse transcriptase (New England Biolabs, Beverly, Mass.). The primer oligonucleotide sequence was 5'-TTCTTCTGGGTGACGAG-3', corresponding to positions 1010 to 994 of the complete sequence reported in Fig. 3. The same primer was used for a sequencing reaction of plasmid pAFS2 (see Results) DNA, which was run on the same gel.

Other methods. Plasmid DNA preparation, purification of DNA from agarose gel, Southern blotting, and restriction enzyme analysis were performed as described previously (23). Chromosomal DNA was isolated as previously described (11). DNA sequences were determined according to the method of Sanger et al. (24). The sequences were analyzed with MacVector (IBI, New Haven, Conn.) computer software.

Nucleotide sequence accession number. The nucleotide sequence data reported in this paper appear in the GenBank database under accession number U35611.

RESULTS

Membrane proteins of *A. brasilense* strains. The *A. brasilense* mutant SPFS5, impaired in its ability to take up iron from the

medium, showed, in a preliminary characterization, an unmodified production of siderophore molecules (18). For this reason we assumed that the phenotype of this mutant was to be ascribed to some defect in the transport of the ferrisiderophore inside the cell. In *A. brasilense* at least four membrane proteins were found to be specifically induced by iron starvation and were therefore considered responsible for iron transport (3). We extracted the membrane proteins of the parental strain SPF94 and the mutant SPFS5, which were grown in iron-supplemented medium or in medium deprived of iron by the addition of the chelating agent 2-2'-dipyridyl, as described in Materials and Methods. The results of the SDS-PAGE of the membrane proteins are shown in Fig. 1. In the parental strain, four or five proteins were apparently induced by iron deprivation. The mutant strain SPFS5 showed a different protein pattern; in particular, the protein of about 80 kDa, which was most conspicuously induced by iron deprivation in the parental strain, was clearly missing in the mutant. Other modifications were also noticeable; for example, another protein of about 70 kDa repressed by iron in the parental strain was absent in the mutant cells.

Cloning and nucleotide sequence of the *A. brasilense lon* gene. Chromosomal DNA from *A. brasilense* mutant SPFS5 was digested with *Sa*II and probed by Southern blotting with the 4.6-kb *Not*I fragment of Tn5, leading to the identification of two bands of about 4.3 and 5.4 kb, respectively. These bands,

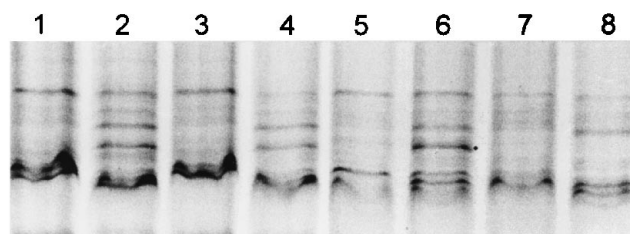


FIG. 1. SDS-PAGE of the membrane proteins of *A. brasilense* strains. Lanes 1 to 2, SPFS55; lanes 3 to 4, SPFS56; lanes 5 to 6, SPF94; lanes 7 to 8, SPFS5. Lanes 1, 3, 5, and 7 were grown in iron-supplemented medium, and lanes 2, 4, 6, and 8 were grown in iron-deprived medium. The dot on lane 6 indicates the 80-kDa protein expressed in strain SPF94 after iron deprivation that is missing in strain SPFS5. Each sample contained 25 μ g of protein.

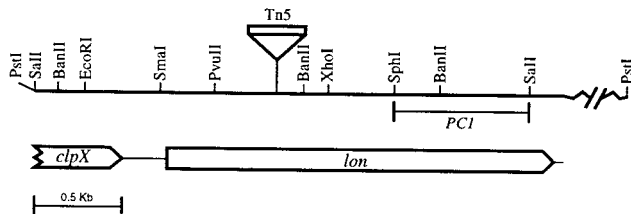


FIG. 2. Restriction map of the DNA fragment containing part of the *clpX* gene and the entire *lon* gene of *A. brasilense*. *PC1* indicates the sequence used as the probe in the Northern blot analysis. Pointed boxes beneath the restriction map indicate the ORFs and the direction of translation. The *PstI* site on the right is about 2 kb from the *SalI* site.

containing the DNA sequences surrounding Tn5, were isolated from agarose gel, ligated to pUC18, and introduced by transformation into *E. coli* DH5 α , and transformants for kanamycin resistance were selected. One of the transformants was found to harbor a plasmid, referred to as pAFS1, with the 4.3-kb *SalI* fragment, as demonstrated by Southern blotting. The labeling of this fragment and its use as the probe permitted the identification of a homologous 3.4-kb *SalI* fragment in the chromosomal DNA of parental strain SPF94. This latter fragment was purified and cloned into pUC18, thus generating the recombinant plasmid pAFS2. Southern blotting (not shown) demonstrated that the restriction map of pAFS2 was consistent with that of the SPF94 chromosomal DNA (Fig. 2). The nucleotide sequence of the 3.4-kb DNA fragment was determined by both subcloning and primer walking (Fig. 3).

The analysis of the sequence revealed the presence of two truncated open reading frames (ORFs) oriented in the same direction (Fig. 2 and 3). The longest ORF, starting with an ATG codon at position 826, encoded a protein with a high degree of sequence homology with the known Lon proteins. In order to obtain the nucleotide sequence of the entire *A. brasilense lon* gene, a *PstI* fragment of 5.5 kb containing the entire insert of pAFS2 and sequences downstream of *lon* (Fig. 2) was identified by Southern blotting of *PstI* digestion of the total DNA of SPF94, purified from the gel, and cloned into pUC18, thus giving the recombinant plasmid pCEK1. The regions flanking the *SalI* sites were sequenced to provide the entire putative *A. brasilense lon* gene. It was found to be 2,433 bp long (Fig. 3), a little longer than its *E. coli* homolog of 2,349 bp. The G+C content of the gene is about 63%, similar to that of other *A. brasilense* genes. Therefore, the codon usage is strongly biased, with most codons ending with G or C. The Tn5 insertion site in mutant SPFS5 was localized between codons

232 and 233 of the *lon* gene (corresponding to nucleotide 1521 of Fig. 3) by sequencing plasmid pAFS1 with an outward-bound Tn5 primer. The ATG codon at position 834 is preceded by a potential ribosome binding site extending from nucleotides 814 to 818. The search for sequences homologous to the *E. coli* σ^{70} promoter and to the iron box (7) or to the heat shock consensus promoter (20, 35) did not give any significant result.

About 550 nucleotides of the carboxy-terminal region of an ORF were identified 250 nucleotides upstream of the *lon* gene (Fig. 2 and 3). This ORF codes for a protein that shows a high degree of sequence homology with the *E. coli clpX* gene product (13).

Analysis of the amino acid sequence of the *A. brasilense Lon* protein. The *A. brasilense lon* gene encoded a protein of 810 amino acids with a predicted molecular mass of 90,135 Da. Examination of the deduced amino acid sequence revealed a high degree of homology with the other Lon proteins, with identities ranging from 33 to 60% and similarities ranging from 45 to 78% (Table 2). The sequence homology extended over almost the full length of the Lon proteins, except for the N-terminal amino acids of human and *S. cerevisiae* Lon, which represent a potential mitochondrial matrix targeting sequence.

The search for an ATP-binding site in the *A. brasilense Lon* sequence revealed the presence of two segments (A and B) which have previously been shown by Walker (referred to in reference 6) to be characteristic of many ATP-binding proteins. The two sequences, GPPGVGKT (segment A) and KAKSSNP LFLLD (segment B), were localized in *A. brasilense Lon* at positions 361 to 368 and 417 to 428. These locations are similar to those found in the *E. coli Lon* protein, in which the two sequences are located at positions 355 to 362 and 411 to 422, respectively. The multialignments of all the amino acid sequences revealed a highly conserved region in segment A, whereas the consensus in segment B was less restrictive, permitting various amino acid substitutions. A serine residue at position 679 in the *E. coli Lon* was proposed to be the active proteolytic site on the basis of site-directed mutagenesis experiments (1). We found a stretch of eight amino acids (PKD GPSAG), containing the putative active-site serine, present in all known Lon sequences, at position 680 to 687 of the *A. brasilense Lon* protein.

Analysis of transcription. Since inactivation of the *lon* gene in *A. brasilense* mutant SPFS5 led to impairment of iron uptake, we decided to study the expression of this gene in order to determine in which condition its transcription was induced and whether it was an iron- or a temperature-responsive gene or both. For this purpose, total RNAs were extracted, as de-

TABLE 2. S_{AB} values calculated for the deduced amino acid sequences of Lon proteins from various species

Species ^a	S_{AB} value for species ^b :							
	Ab	Ec	MxV	MxD	Bs	Bb	Hs	Sc
Ab		0.57	0.60	0.46	0.54	0.54	0.33	0.33
Ec	0.77		0.54	0.43	0.40	0.52	0.33	0.31
MxV	0.78	0.74		0.48	0.54	0.54	0.37	0.32
MxD	0.70	0.56	0.72		0.48	0.47	0.36	0.34
Bs	0.76	0.62	0.76	0.71		0.70	0.36	0.31
Bb	0.66	0.76	0.77	0.67	0.88		0.36	0.33
Hs	0.57	0.55	0.51	0.59	0.62	0.60		0.40
Sc	0.45	0.42	0.44	0.44	0.43	0.44	0.53	

^a Ab, *A. brasilense*; Ec, *E. coli*; MxV and MxD, *M. xanthus*; Bs, *B. subtilis*; Bb, *B. brevis*; Hs, *Homo sapiens*; Sc, *S. cerevisiae*.

^b The S_{AB} value for two proteins, A and B, is $2 \times (\text{number of similar or identical amino acids in A and B}) / (\text{number of total amino acids in A} + \text{number of total amino acids in B})$. The top right triangular section shows amino acid identities; the lower left triangular section shows amino acid similarities.

GTCGACACGAGCAACATCCTGTTTCATCTGCGGCGGCGCCTTCGCGGGGCTGGACAAGATCATCGCGCAGCGCGCAAGGGCACGTCGATCGGCT 94
V D T S N I L F I C G G A F A G L D K I I A Q R G K G T S I G F
TCGGCGCTGATGTGCGCGGCCGAGGAGCGTTCGACCGCGCACATCCTCGGTGAGGTCGAGCCGAGGATCTGCTGAAATTCGGCCTGATCCC 188
G A D V R G P D E R S T G D I L R E V E P E D L L K F G L I P
CGAGTTTACCGGTCTGCGCGGTGTCGCCACCCTGTCCGACTCGACGAAACGGCTCTGGTCGAGATCCTCACCAGCCGAAGAATGCCCTG 282
E F I G R L P V V A T L S D L D E T A L V E I L T K P K N A L
GTCAAGCAGTACCAGCGCTGTTCGAGATGGAGGACGTCCTGCTGGAATTCCTCGACGACGCCCTGCGCACCATCTCCCAACAGCCATCCAGC 376
V K Q Y Q R L F E M E D V R L E F S D D A L R T I S H K A I Q R
GGAAGACCGGCGCACGCGCGCTGCGGTGATCATGGAGTCGATCCTGCTCGACCCGATGTCGATCTGCCCGGCTGTCGGGGGTGGAAGCAT 470
K T G A R G L R S I M E S I L L D P M F D L P G L S G V E S I
CCTGGTCAACAAGGAGGTGGTCGAGGGGCGCGCAAGCCGCTTACGTCACCGCCGAAACCGCGGGGAGCAGCGGACCGGGTGTGAGTTC 564
L V N K E V V E G R A K P L Y V H A E R R G E Q Q A P G A *
GTCACGGCCGATCAGCATCGGGTACGAGGAAGGGCGCCTTTGGCGCCCTTTGTCTTCCGCTTTTGCAGACATCGAGGGCGCCAGCGGGCC 658
GCGGGCGTGGATCGGCATGTCCGACCCGCATACGGAATGGGGTTCGCGCCTTGAAGGGGGCTATGCGTGTGCCACGTTAGAAAAAGTGC 752
CGGTTTTGATCCCTGTGCTCATCCCCGAGGGGGGCTTCCGAGGCCATCCCGGGCTgacggCGTCCCAATGAAAGAGGCCCAATCAATG 846
M K E A Q S M
TTCGAAATCCCTCGTGTGCCCTCTATCCGGTCCCGCGCTCCGCGACATCGTGGTTTTCCCCACATGATCGTGCCTCTTTTCGTCGGCCGTCG 940
F E I P R G A L Y P V P P L R D I V V F P H M I V P L F V G R
AGAAATCCGTCGCGCCCTGGAAGACGTGATGAAGGACGATAAGCAGATCCTTTCGTCACCCAGAAGAACCGCGCGAGGACGATCCGACGCC 1034
E K S V R A L E D V M K D D K Q I L L V T Q K N A A Q D D P T P
GGCCGATATCTACAGCGTCGCGCACCGTGGACCGTGTTCGACGTCGTAAGCTGCCCGACGGAACGGTGAAGTGTCTGTTGAGGGCGGCCAA 1128
A D I Y S V G T V G T V L Q L L K L P D G T V K V L V E G G Q
CGCGCGTCCATCACCAGTTCGCGGAGACGAGGATTTCTCCAGGCCACCGCCTCGTCGAGGAGAAGTTCGGGAAAGCCAGGAAGTTC 1222
R A S I T K F A E N E D F F Q A H A D L V E E K V G E S A C T
AGGCGCTGGGACGCGCGGTCTGCTCGCAATTCGAGCAGTACATCAAGCTGAACAAGAAGATCCCGCGGAGGTCTGCTCGATCAACCAGAT 1136
E A L G R A V V S Q F E Q Y I K L N K K I P P E V L V S I N Q I
CGAAGAGCCGGAAAGCTGGCGGACACCGCTCCACCTCGCGCTGAAGATTCGGAAGAACGAGCAGCTTCTGGAATGGCCACCGGTTTCG 1410
E E P G K L A D T V A S H L A L K I P E K Q Q L L E C A T V S
GAGCGGTTGGAGCGGTCTACGCTTCATGGAAGGCGAAATCGGTGTCCTCAAGTGGAAAAGCGCATCCGCAACCGCGTCAAGCGGCGATGG 1504
E R L E R V Y A F M E G E I G V L Q V E K R I R N R V K R Q M
AGAAGACCCAGCGGACTACTACCTCAACGAACAGCTCAAGGCGATCCAGAAGGAACTCGCGGAGACCGGAGGACCGCCGTGACGATCGGCCGA 1598
E K T Q R E A Y Y L N E Q L K A I Q K E L G E T E D G R D E S A E
GCTGGAGAGAAGATCAACAAGACCGCTTCTCAAGGAAGCCCGCACAAGGCCCTGGCCGAGCTGAAGAAGCTCGCTCCATGAGCCCGATG 1692
L E E K I N K T R F S K E A R D K A L A E L K K L R S M S P M
TCGGCCGAGGCGAGGTTGGTGCACAATCTGGACTGGATGCTCCATTCCGTTGGAAGAAGCGCACCAAGGTGAAGAAGGACCTGAAGCTCG 1786
S A E A T V V R N Y L D W M L S I P W K K R T K V K K D L K L
CTCAGAAGATCCTCGACCGCCACTACGGCTGGAGAAGTCAAGGAACGCATCCTCGATATCTTCGGTCCAGAACCAGCATGAACAAGT 1880
A Q K I L D A D H Y G L E K V K E R I L E Y L R V Q N R M N K V
CAAGGGCCGATCCAGAGCTCGTCGGGCGCCCGCGTGGCAAGACCTCGCTCGGCAAGTCGATCGCCAGTCCACCGCCGCAATTTCTGTG 1974
K G P I Q S L V G P P G V G K T S L G K S I A K S T G R N F V
CGCATGCTGCTCGCGCGGTGCGGACGAGGCTGAGGTCGCGCCACCGGACCTACATCGGTCGATGCCCGGCAAGGTCATCCAGGGCA 2068
R M S L G G V R D E A E V R G H R R T Y I G S M P G K V I Q G
TGAAGAGCGGAAGTCCCAACCCGCTGTTCTGCTGGATGAGATCGACAAGCTCGCGCCGACTGCGCGGCGACCCGCTCGTCGGCCCTGCT 2162
M K K A K S S N P L F L L D E I D K L G A D W R G D P S A L L
TGAGGTTCTCGATCCCGAGCAGAACCGGCACTTTCAACGACATTATCTGGAGGTCGATTACGACCTGTCGAGCTGATGTTCTGTCGACGGCC 2256
E V L D P E Q N G T F N D H Y L E V D Y D L S D V M F V C T A
AACACGATGCGCATGCCGACCGCTGCTGGACCGCATGGAGATCATCCGCTCGCCGGCTATACCGAGGATGAAAAGGTCTGAGATTTCCAAGC 2350
N T M R Y P Q P L L D R M E I I R V A Y T E D E K V E I S K
GCCACCTGATCGAAAAGCAGGTGAAGCCAACCGCTTGAAGAAGGGCAGTTCGCCATTTCCGACGACGCGCTGCGGACCTGATCGCCTATTA 2444
R H L I E K Q V E A N G L K K G E F A I S D D A L R D L I R Y Y
CACGCGGAAGCCGCGTCCGACGCTGGAGCGGAGATCGCCAACTCTGCGCAAGCCGTAAGGAGATCCTGATGAAGGGCTCCGCGCCG 2538
T R E A G V R S L E R E I A N L C R K A V K E I L M K G S A G
GCCAAGTTCGGTACGCGCGGAACTGGACAAGTACGCTGGCGTCCGCGCTTCCACTTCGCGAGGCGGAGCTGGAGATCTGGTGGCG 2632
A K V S V T R R N L D K Y A G V R R F H F G E A E L E D L V G
TCACCACCGGGCTGGCTGGACGGAAGTGGAGGCGAGCTGCTGATCGAGGCGGTGACGCTGCCCGCAAGGGCGGGTACCACCCGG 2726
V T T G L A W T E V G G E L L S I E A V S L P G F G R V T T T G
CAAGTGGGCGACGTCATGAAGGAGTCCGTCAGGCGGCGAGGACTACGTAAGTCCGCGGCGCCCGCTTCGGCATCAAGCCGACGCTTC 2820
K L G D V M K E S V Q A A E S Y V K S R A T A F G I K P T L F
GAGAAGCGGGACATCCACGTCACGTCGCCGAAGGCGCCACTCCGAAGGACGGCCCTCGCGGGCGTGGCGATGATCACCTCGATCGTTTCGG 2914
E K R D I H V H V P E G A T P K D G P S A G V A M I T S I V S
TCCTGACCGGTATCGCGTCCGCAAGGACGTGGCGATGACCGGTGAGATCACCCCTGCGCGCGGGTGTCTCCGATCGCGGTCTGAAGGAGAA 3008
V L T G I A V R K D V A M T G E I T L R G R V L P I G G L K E K
GCTGCTGCTGCTGCTGCGCGCGCCCTCAAGCATGTGCTGATCCGCAAGGACACGAGAAGGATCTCGCGGATCCCGGACACGTCGAAGCGC 3102
L L A A L R G G L K H V L I P K D N E K D L A E I P D N V K R
GGGCTGGAGATCCCAGTACGACCGTCCGACGCTCCTGAAGCACGCGCTGGTTCGGGAAGTCGAGCCATCGAGTGAAGGAGCGCGAAG 3196
G L T E I I P V S T V D V L K H A L V R E V E P I E W K E P E
CGGTCGAGCCGCGGTCCGCAAGCGCAGACTGACGCGGAGGCGAGGTGCTTCGTCACGCTCCAGCTTGGCGTTTAAAGACCCCTAAAGCG 3290
A V E P A V A K P Q T D G G G E V L R H *
ATACCATGCTTCGATCGCCCGCGCTGCAATTGACGGCGATGCGCGTTTATCGTTTGAATGCGTGTGATGTCATTGAGTGCCTTTCTTCCA 3384
TTTTCTTTCGCGACTGAGAAGGGGGAAGTAAAGTGAACAAGAACGA 3440

FIG. 3. Nucleotide and deduced amino acid sequences of the *A. brasiliense lon* gene. The numbers on the right are nucleotide positions. The termination codons of the *lon* and *clpX* genes are indicated by asterisks. The putative ribosome binding site is indicated by boldface lowercase letters. The transcription start sites are indicated by arrows. Amino acids are represented by the single-letter code. The putative ATP-binding sites are in boldface type and double underlined, whereas the proteolytic site (PKDGPSAG) is underlined.

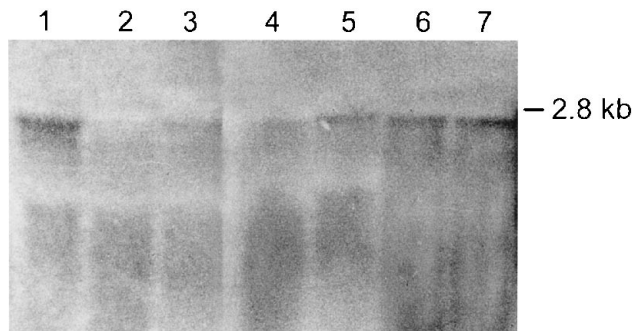


FIG. 4. Northern blot analysis of *A. brasilense* SPF94 total RNA hybridized to the *PC1* probe (Fig. 2). RNA was extracted from cells treated as follows: lane 1, 45°C for 5 min; lane 2, 45°C for 40 min; lane 3, 48°C for 5 min; lane 4, uninduced control kept at 33°C; lane 5, 45°C for 5 min; lane 6, 45°C for 10 min; lane 7, 45°C for 20 min. The 2.8-kb marker, not shown, was the digoxigenin-labeled (Boehringer) RNA molecular mass marker II.

scribed in Materials and Methods, from cultures of parental or mutant strains under the following conditions: (i) growth at 33°C in MSP medium with iron up to an OD_{550} of 0.4 and then a shift to 45°C for 5, 10, 20, or 40 min, or a shift to 48°C for 5 min; (ii) growth at 33°C in MSP medium without iron and with 70 μ M 2-2'-dipyridyl from an OD_{550} of 0.2 to an OD_{550} of 0.4; and (iii) growth at 33°C in MSP medium with iron up to an OD_{550} of 0.4 as a control. Northern blotting of RNA hybridized to the *A. brasilense lon*-specific probe *PC1* (Fig. 2) demonstrated the presence of a distinct band of about 2,400 nucleotides only in the RNA extracted from the SPF94 cells exposed to 45°C for 5, 10, or 20 min (Fig. 4). This result suggests that the transcription of this gene, like that of its *E. coli* homolog, was triggered by heat shock and that the level of its induction rose rapidly and lasted for little more than 20 min. RNA from parental strain cells without heat shock or grown in iron-deprived medium did not show any specific discrete signal, not did RNA from SPFS5 cells under any condition (not shown). The heat-shock at 48°C caused the appearance of a very faint RNA band. However, weak signals, visible only as smears on the Northern blot membrane, were present in all the parental cell RNA preparations, indicating a low level of constitutive transcription of the *lon* gene, which was also confirmed by RNA dot blot experiments (not shown). The size of the *lon* RNA molecule found in the heat-shocked cells (about 2.4 kb) suggests that the transcript was a monocistronic one, probably as large as the coding sequence (2,433 nucleotides) and therefore not including the preceding putative *clpX* gene.

In order to confirm the results obtained by Northern blot analysis and to try to localize the promoter of the gene, we used total RNA of the 5-min heat-induced parental cells to analyze the transcription initiation of the *lon* gene by primer extension, as described in Materials and Methods. The results obtained, shown in Fig. 5, reveal the presence of two putative start sites, one located at a G in position 749 and the other at a G in position 790, respectively 77 and 36 nucleotides upstream of the putative ATG start codon. RNA prepared from cells grown at 33°C also showed the same putative transcription initiation sites, although the signal was less intense (not shown).

Heat shock resistance and UV sensitivity. The *A. brasilense lon* gene was clearly induced by heat shock, and for this reason it is probably involved in some way in the heat shock response. In order to test whether one of the functions of this gene is the protection of cells from heat damage, we compared the resis-

tance of parental and *lon* mutant strains to high-temperature shock. The results of the experiment are reported in Table 3. Unexpectedly we found that, in the experimental condition used, the mutant SPFS5 was more resistant to heat shock than the parental strain SPF94. This result indicated a certain involvement of the *A. brasilense lon* gene product in the cellular response to heat shock. To test whether the mutation in the *lon* gene of *A. brasilense*, like the *E. coli lon* mutation, is responsible for the pleiotropic phenotype of increased sensitivity to UV irradiation, with consequent formation of filamentous cells, we compared the resistance of parental and *lon* mutant strains to UV as described in Materials and Methods. The results obtained indicated that the two strains were equally sensitive to different doses of UV. Moreover, elongated forms were not observed among the UV-treated cells (data not shown). As a further indication of the different functions of the *A. brasilense lon* gene compared with those of the *E. coli* gene, we did not observe the mucoid phenotype in the colonies of strain SPFS5.

To ascertain whether the phenotype of the mutant strain SPFS5 is actually the consequence of disruption by *Tn5* of the *lon* gene, we used the cloned wild-type *A. brasilense lon* gene to complement the mutation of SPFS5. A 3.2-kb *PstI-DraI* DNA fragment exactly containing the entire *lon* sequence was cloned in the broad-host-range vector pAF300 (8), giving the plasmid pAF371. This plasmid was transferred by transformation into *E. coli* DH5 α and then conjugated into *A. brasilense* SPFS5. Exconjugants were checked for the presence of plasmid pAF371, and two of them, referred to as SPFS55 and SPFS56, were assayed for their phenotypes, with respect to their iron-uptake abilities (the blue CAS plates were used), membrane protein patterns, and heat shock responses. The results obtained, reported in Fig. 1 and 6 and Table 3, clearly indicated that SPFS5 mutants carrying the plasmid with an intact copy of the *lon* gene resembled the parental strain.

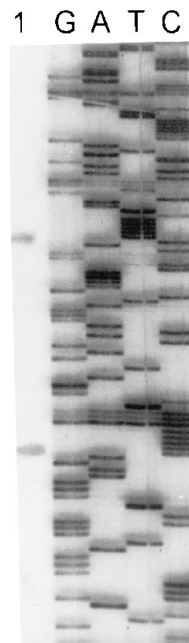


FIG. 5. Locations of potential transcription start sites by primer extension of the *lon* gene of *A. brasilense*. The primer used corresponds to positions 1010 to 994 of the sequence reported in Fig. 3. Lane 1, primer extension product; lanes C, A, T, and G, products of sequencing reactions with the same primer used for the extension. Note that the sequence is complementary to that shown in Fig. 3.

TABLE 3. Survival of *A. brasilense* strains to heat shock

Strain	Temperature (°C)	% Survival after heat shock for time (min) ^a		
		5	10	15
SPF94	49	92	61	36
SPFS5	49	101	81	84
SPF94	51	40	ND	27
SPFS5	51	113	ND	94
SPFS55	51	32	ND	ND
SPFS56	51	35	ND	ND
SPF94	52	23	3	ND
SPFS5	52	85	53	ND

^a Survival is expressed as a percentage of the CFU of the control kept at 33°C. One hundred percent corresponds to 5×10^7 to 10×10^7 CFU/ml. Data are means of three replicates. ND, not done.

DISCUSSION

In this study we cloned and sequenced the *lon* gene of the nitrogen-fixing bacterium *A. brasilense*, starting from a Tn5-induced mutant that had previously been selected for its altered ability to take up iron from the medium. Before our study, some genes homologous to the *E. coli lon* gene from bacteria, *S. cerevisiae*, and humans (see the introduction for references) had been characterized. The amino acid sequence of *A. brasilense Lon* is highly homologous to the Lon sequences of other species, in particular with that of *M. xanthus* (they have over 60% identity in their amino acid sequences). This degree of sequence homology is higher than that generally found between the sequenced genes of *Azospirillum* spp. and their homologous counterparts in other bacterial species. A close examination of the aligned sequences showed regions with extremely high levels of homology which contain the major identified domains of the *E. coli* protein: the ATPase and proteolytic domains. This finding also confirms the similarity in the basic functions of these serine proteases.

The *lon* genes characterized so far, in spite of their overall homology, were not found to have the same physiological role in different species. Moreover, these functions have not been fully clarified in any of the organisms considered, including *E. coli* (14, 21). The phenotype of the *A. brasilense lon* mutant appeared to be different from those of other species, showing normal growth (therefore, *lon* is not an essential gene) and being affected in its iron uptake and heat shock response. As with *B. brevis* (16) and *B. subtilis* (22), features which are characteristic of *E. coli lon* mutants, such as mucoidy, filamentous growth, and increased sensitivity to UV (14), were not observed

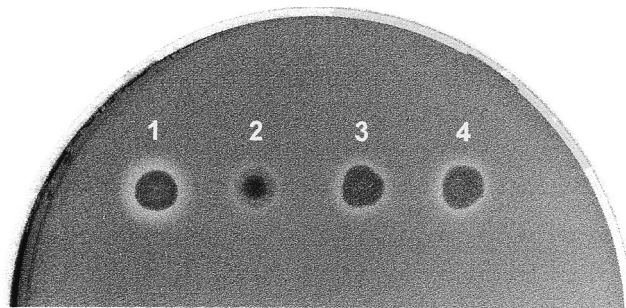


FIG. 6. CAS plate with colonies of *A. brasilense* strains. Colonies: 1, SPF94; 2, SPFS5; 3, SPFS55; 4, SPFS56. The plates were incubated for 3 days at 33°C. The diameters of the bright haloes indicate the amount of iron taken up by the cells.

in the *A. brasilense* mutant. This suggests that the specific (14) roles of Lon may be species specific. Our data showed that the *lon* gene is involved in the synthesis (or in the regulation of the synthesis) of one or more membrane proteins induced by iron deprivation. A difference of more than 10 kDa between estimated molecular masses should exclude the possibility that the 80-kDa membrane protein, induced by iron deprivation and missing in the mutant strain, is Lon.

A previous work demonstrated an inhibitory effect of temperature (growth at 42°C) on iron uptake in *A. brasilense* RG (4). In that case, temperature inhibited siderophore production while leaving iron transport and iron-regulated membrane proteins unaffected. From these data it is not possible to conclude that the *lon* gene is involved in the switch-off of siderophore production by temperature; however, the data suggest that a linkage exists between temperature and iron uptake, as has also been reported for other bacterial species (9). It is not known whether the *lon* homologous genes in other species are related to iron uptake, as, to our knowledge, this phenotype has not been looked for in other *lon* mutants. On the other hand the fact that *lon* transcription in *A. brasilense* is not induced either by iron starvation or by its abundance could signify an aspecific involvement of this gene in iron uptake.

Concerning heat shock response, the reported results clearly demonstrated that the transcription of the *A. brasilense lon* gene was induced by heat shock in a manner that closely resembles the induction of other heat shock genes, including *E. coli* and *B. subtilis lon*. In *B. brevis* and *M. xanthus*, however, *lon* is not a heat shock gene. Moreover, transcriptional start sites, mapped by primer extension, did not help much in the identification of promoter sequences homologous to those of known heat shock genes of *E. coli*. On the other hand, this is the first *A. brasilense* heat shock gene sequenced so far and it is not possible to infer any consensus sequence for promoters of these genes in this organism. Unfortunately, the mere fact that the *A. brasilense lon* gene is induced by heat shock does not give information about its specific role. An intriguing feature of the *A. brasilense lon* mutant is its particular resistance to heat shock, as revealed by a rate of survival greater than that of the parental strain after exposure to high temperature. This was unexpected, as we assumed that the *lon* gene product, like its *E. coli* counterpart, is induced as a response to heat shock stress for the sake of preventing an accumulation of denatured proteins, thus allowing cells to recover more easily from the shock. One possible explanation of this result is that, in the experimental conditions used, *A. brasilense* parental cells died after heat treatment also because the highly degradative action of Lon protease (which is missing in the mutant) could have destroyed some essential proteins that were still functioning.

The results reported in this paper add a new fragment to the complex and expanding picture of the functions of the family of Lon-like serine proteases, in that they indicate a certain involvement in iron uptake. The knowledge of the iron uptake system in *A. brasilense* is unfortunately still too preliminary to try to detect at what level the action of Lon occurs. We can conclude only that the *lon* defect of strain SPFS5 caused the loss of some iron-induced membrane proteins, indicating that, if the mutated gene is of the regulatory type, it should probably code for a repressor.

ACKNOWLEDGMENTS

We are grateful to G. Mastromei for a critical reading of the manuscript and for help in the production of the figures. The contribution of M. T. Ceccherini is also acknowledged.

This work was supported by the National Research Council of Italy, Special Project RAISA, subproject no. 2, paper no. 2754.

REFERENCES

- Amerik, A. Y., V. K. Antonov, A. E. Gobalenya, S. A. Kotova, T. V. Rotanova, and E. V. Shimbarevich. 1991. Site-directed mutagenesis of La protease: a catalytically active serine residue. *FEBS Lett.* **287**:211–214.
- Ausubel, F. M., R. Brent, E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1988. *Current protocols in molecular biology*. John Wiley, New York.
- Bachhawat, A. K., and S. Gosh. 1987. Iron transport in *Azospirillum brasilense*: role of the siderophore spirilobactin. *J. Gen. Microbiol.* **133**:1759–1765.
- Bachhawat, A. K., and S. Gosh. 1989. Temperature inhibition of siderophore production in *Azospirillum brasilense*. *J. Bacteriol.* **171**:4092–4094.
- Bani, D., C. Barberio, M. Bazzicalupo, F. Favilli, E. Gallori, and M. Polsinelli. 1980. Isolation and characterization of glutamate synthase mutants of *Azospirillum brasilense*. *J. Gen. Microbiol.* **119**:239–244.
- Chin, D. T., S. A. Goff, T. Webster, T. Smith, and A. L. Goldberg. 1988. Sequence of the *lon* gene in *Escherichia coli*. *J. Biol. Chem.* **263**:11718–11728.
- de Lorenzo, V., S. Wee, M. Herrero, and J. B. Neilands. 1987. Operator sequences of the aerobactin operon of plasmid ColV-K30 binding the ferric uptake regulation (*fur*) repressor. *J. Bacteriol.* **169**:2624–2630.
- Fani, R., M. Bazzicalupo, F. Ricci, C. Schipani, and M. Polsinelli. 1988. A plasmid vector for the selection and study of transcription promoters in *Azospirillum brasilense*. *FEMS Microbiol. Lett.* **50**:271–276.
- Garibaldi, J. A. 1971. Influence of temperature on the iron metabolism of a fluorescent pseudomonad. *J. Bacteriol.* **105**:1036–1038.
- Gill, R. E., M. Karlok, and D. Benton. 1993. *Myxococcus xanthus* encodes an ATP-dependent protease which is required for developmental gene transcription and intercellular signaling. *J. Bacteriol.* **175**:4538–4544.
- Giovannetti, L., S. Ventura, M. Bazzicalupo, R. Fani, and R. Materassi. 1990. DNA restriction fingerprint analysis of the soil bacterium *Azospirillum*. *J. Gen. Microbiol.* **136**:1161–1166.
- Goldberg, A. L. 1994. ATP-dependent protease La (*Lon*) from *Escherichia coli*. *Methods Enzymol.* **244**:350–375.
- Gottesman, S., W. P. Clark, V. Cr  cy-Lagard, and M. R. Maurizi. 1993. *ClpX*, an alternative subunit for ATP-dependent *Clp* protease of *E. coli*. *J. Biol. Chem.* **268**:22618–22626.
- Gottesman, S., and M. R. Maurizi. 1992. Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol. Rev.* **56**:592–621.
- Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**:557–580.
- Ito, K., S. Udaka, and H. Yamagata. 1992. Cloning, characterization, and inactivation of the *Bacillus brevis lon* gene. *J. Bacteriol.* **174**:2281–2287.
- Lugtenberg, B., J. Meijers, R. Peters, P. Van Der Hock, and L. Van Alphen. 1975. Electrophoretic resolution of the major outer membrane protein of *E. coli* K12 into 4 bands. *FEBS Lett.* **58**:254–258.
- Mori, E., R. Fani, E. Gallori, O. Fantappi  , and M. Bazzicalupo. 1992. Mutants of *Azospirillum brasilense* altered in the uptake of iron. *Symbiosis* **13**:115–122.
- Mugnai, M., M. Bazzicalupo, R. Fani, E. Gallori, D. Paffetti, and R. Pas-torelli. 1994. Factors affecting nitrogen fixation and *nif* gene expression in *Azospirillum brasilense*. *FEMS Microbiol. Lett.* **120**:133–136.
- Neidhardt, F. C., and R. A. VanBogelen. 1987. Heat shock response, p. 1334–1345. In F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
- Parsell, D. A., and S. Lindquist. 1993. The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. *Annu. Rev. Genet.* **27**:437–496.
- Riethdorf, S., U. Volker, U. Gerth, A. Winkler, S. Engelmann, and M. Hecker. 1994. Cloning, nucleotide sequence and expression of the *Bacillus subtilis lon* gene. *J. Bacteriol.* **176**:6518–6527.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
- Schwyn, B., and J. B. Neilands. 1987. Universal chemical assay for the detection and determination of siderophores. *Anal. Biochem.* **160**:47–56.
- Stout, V., and S. Gottesman. 1991. Regulation of capsular polysaccharide synthesis in *Escherichia coli* K12. *Mol. Microbiol.* **5**:1599–1606.
- Stout, V., A. Torres-Cabassa, M. R. Maurizi, D. Gutnick, and S. Gottesman. 1991. RcsA, an unstable positive regulator of capsular polysaccharide synthesis. *J. Bacteriol.* **173**:1738–1747.
- Suzuki, C. K., K. Suda, N. Wang, and G. Schatz. 1994. Requirement for the yeast gene *Lon* in intramitochondrial proteolysis and maintenance of respiration. *Science* **264**:273–276.
- Tojo, N., S. Inouye, and T. Komano. 1993. Cloning and nucleotide sequence of the *Myxococcus xanthus lon* gene: indispensability of *lon* for vegetative growth. *J. Bacteriol.* **175**:2271–2277.
- Tojo, N., S. Inouye, and T. Komano. 1993. The *lonD* gene is homologous to the *lon* gene encoding an ATP-dependent protease and is essential for the development of *Myxococcus xanthus*. *J. Bacteriol.* **175**:4545–4549.
- Wang, N., S. Gottesman, M. C. Willingham, M. M. Gottesman, and M. R. Maurizi. 1993. A human mitochondrial ATP-dependent protease that is highly homologous to bacterial *Lon* protease. *Proc. Natl. Acad. Sci. USA* **90**:11247–11251.
- Waxman, L., and A. L. Goldberg. 1982. Protease La from *Escherichia coli* hydrolyzes ATP and proteins in a linked fashion. *Proc. Natl. Acad. Sci. USA* **79**:4883–4887.
- Waxman, L., and A. L. Goldberg. 1985. Protease La, the *lon* gene product, cleaves specific fluorogenic peptides in an ATP-dependent reaction. *J. Biol. Chem.* **260**:12022–12028.
- Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103–119.
- Yura, T., H. Nagai, and H. Mori. 1993. Regulation of the heat-shock response in bacteria. *Annu. Rev. Microbiol.* **47**:321–350.