

# Purification, Cloning, and Preliminary Characterization of a *Spiroplasma citri* Ribosomal Protein with DNA Binding Capacity\*

(Received for publication, February 18, 1998, and in revised form, May 26, 1998)

Loïck Le Dantec†‡, Michel Castroviejo¶, Joseph M. Bové‡, and Colette Saillard‡¶

From the ‡Laboratoire de Biologie Cellulaire et Moléculaire, Institut National de la Recherche Agronomique and Université Victor Segalen Bordeaux 2, 33883 Villenave d'Ornon Cedex and the ¶Institut de Biochimie et Génétique Cellulaires and Centre National de la Recherche Scientifique, 33077 Bordeaux Cedex, France

The *rpsB-tsfx* operon of *Spiroplasma citri* encodes ribosomal protein S2 and elongation factor Ts, two components of the translational apparatus, and an unidentified X protein. A potential DNA-binding site (a 20-base pair (bp) inverted repeat sequence) is located at the 3' end of *rpsB*. Southwestern analysis of *S. citri* proteins, with a 30-bp double-stranded oligonucleotide probe (IRS), containing the 20-bp inverted repeat sequence and the genomic flanking sequences, detected an IRS-binding protein of 46 kDa (P46). P46 protein, which displays preferential affinity for the IRS, was purified from *S. citri* by a combination of affinity and gel filtration chromatographies. The native form of P46 seems to be homomultimeric as estimated by SDS-polyacrylamide gel electrophoresis analysis and gel filtration. A 3.5-kilobase pair *S. citri* DNA fragment comprising the P46 gene and flanking sequences was cloned and sequenced. Sequence analysis of this DNA fragment indicated that the P46 gene is located within the S10-*spc* operon of *S. citri* at the position of the gene coding for ribosomal protein L29 in the known S10-*spc* operons. The similarity between the N-terminal domain of P46 and the L29 ribosomal protein family and the presence of a 46-kDa IRS-binding protein in *S. citri* ribosomes indicated that P46 is the L29 ribosomal protein of *S. citri*. We suggest that P46 is a bifunctional protein with an L29 N-terminal domain and a C-terminal domain involved in IRS binding.

*Spiroplasma*s are wall-free bacteria belonging to the class Mollicutes, a group of microorganisms phylogenetically related to Gram-positive bacteria with low guanine + cytosine contents (1). Sequence analysis (70) of a 6.8-kbp<sup>1</sup> DNA fragment (GenBank™ accession number AF012877) of the phytopathogenic mollicute *Spiroplasma citri* (2) made it possible to identify eight putative ORFs that encode ribosomal protein S2, elonga-

tion factor Ts, spiralin, 6-phosphofructokinase, pyruvate kinase, and three unidentified proteins (A, B, and X) (Fig. 1). Ribosomal protein S2 and the translational elongation factor Ts (Ef-Ts), respectively, encoded by *rpsB* and *tsf* genes, are both components of the translational apparatus in prokaryotes. These genes are adjacent in *Escherichia coli* (3, 4) and *Bacillus subtilis* (5), whereas in the genome of the two mollicutes *Mycoplasma genitalium* (6) and *Mycoplasma pneumoniae* (7) they reside at different locations, and thus each of them may constitute monocistronic transcriptional units or be part of two different polycistronic operons.

The organization and relative orientation of *rpsB* and *tsf* of *S. citri* (Fig. 1) are analogous to those reported for *E. coli* (3, 4) and *B. subtilis* (5). In *E. coli*, *rpsB* and *tsf* form a single transcriptional unit, and an attenuation mechanism was proposed to explain the 1 to 3 ratio of Ef-Ts to S2 (3). In *B. subtilis* a potential terminator is found between *rpsB* and *tsf*. In *S. citri*, the absence of a rho-independent termination signal in the spacer region between *rpsB* and *tsf*, and between *tsf* and *x*, indicates that *rpsB*, *tsf*, and *X* might represent a single transcriptional unit. Transcriptional analyses of *rpsB*, *tsf*, and *x* genes of *S. citri* have recently revealed two different transcripts (70), one corresponding to the *rpsB/tsfx* operon and the second to *rpsB* alone. These results suggested that a regulatory mechanism may act at the transcriptional level at the spacer region between *rpsB* and *tsf*. The only "regulatory" like sequence found in the *rpsB-tsfx* region is an inverted repeat sequence at the 3' end of *rpsB* (Fig. 1). This inverted repeat sequence is 20 bp long and represents two turns of helical DNA. It could be a binding site for a regulatory DNA-binding protein (8, 9).

In the study reported here, we have purified, by a combination of affinity and gel filtration chromatographies, a 46-kDa protein (P46) that displays preferential binding to a 30-bp double-stranded oligonucleotide containing the inverted repeat sequence. The gene coding for P46 has been cloned and sequenced. Sequence analysis has revealed significant similarities between the N-terminal part of P46 and the L29 eubacterial ribosomal protein family. Surprisingly, the *S. citri* protein is much larger than its eubacterial homologs, and the C-terminal domain of P46 shows significant similarities with the DNA-binding histone H1-like proteins found in some bacterial species. These results suggest a bifunctional role for P46. The protein could act as a ribosomal protein but also as a DNA-binding protein with a potential regulatory function.

## EXPERIMENTAL PROCEDURES

**Preparation of Crude Extracts**—Crude extracts were prepared from *S. citri* strain RSA2 (ATCC 27556) cultivated in 5 to 10 liters of SP4 medium (10) at 32 °C. Cells from exponential growth phase were harvested by centrifugation at 17,000 × *g* for 40 min, washed with phosphate-buffered saline containing 70 g/liter sorbitol, and frozen at -20 °C overnight. Cells were then thawed on ice, resuspended in 10 to 30 ml of binding buffer BB (10 mM HEPES, pH 7.9, 50 mM KCl, 1 mM

\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF031160.

‡ Supported by the Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche.

¶ To whom correspondence should be addressed: Laboratoire de Biologie Cellulaire et Moléculaire, Institut National de la Recherche Agronomique, Domaine de la Grande Ferrade, BP 81, 33883 Villenave d'Ornon Cedex, France. Tel.: 33 5 56 84 31 52; Fax: 33 5 56 84 31 59; E-mail: saillard@bordeaux.inra.fr.

<sup>1</sup> The abbreviations used are: kbp, kilobase pair(s); bp, base pair(s); FPLC, fast protein liquid chromatography; ORF, open reading frame; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; Ef-Ts, elongation factor Ts.

EDTA, 1 mM dithiothreitol, 12.5% glycerol) containing 1 mM phenylmethylsulfonyl fluoride, and disrupted by sonication (5/10-s pulses, 50 watts, 0 °C, 1 min sonication, and 1 min on ice alternatively, 8 to 10 times with a Vibracell sonicator). Mixture was cleared by centrifugation at 12,000 × *g* for 20 min; the pellet was discarded, and the supernatant was directly loaded onto heparin chromatography column or frozen at -70 °C for storage. The protein concentration of the supernatant, assayed by the Bradford procedure (11) using the Bio-Rad Protein Assay Kit with bovine serum albumin as a standard, was in the range of 1–10 mg/ml.

**Southwestern Blot Analysis of Protein-DNA Interactions**—The method used for Southwestern analysis is a modification of previously described "protein blotting" method (12). Crude or purified proteins were resolved by SDS-PAGE on 10% polyacrylamide gel using standard procedures (13). After electrophoresis, proteins were electroblotted (14) to nitrocellulose membranes (C extra, Amersham Pharmacia Biotech) for 45 min at 8 watts in transfer buffer (25 mM Tris base, 190 mM glycine, and 15% methanol) using a semi-dry transfer apparatus (Fast-blot Biometra). Membranes were dried 15 min at 37 °C and blocked for 60 min at 26 °C in buffer S (BB buffer containing 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, and 0.02% bovine serum albumin). Membranes were incubated in 10 ml of S buffer containing 10<sup>6</sup> cpm/ml end-labeled specific IRS probe. This probe is the synthetic double-stranded oligonucleotide (5'-GGATTTGCTTTACCAAAAGCAAAAAGCTG-3' and 3'-CCTAAACGAAATGGTTTTTCGTTTTTCGAC-5') containing the inverted repeat sequence (Fig. 1). It was end-labeled with [ $\gamma$ -<sup>32</sup>P]ATP (ICN Biochemicals) using T4 polynucleotide kinase and purified using a Sephadex G-25 column (15).

Incubations were at 26 °C for 60 min in screw-capped glass tubes with gentle rotation in rotating hybridization oven (Robbins Scientific), followed by three washes of 15 min in S buffer. The blots were then exposed 1–12 h for autoradiography.

Preferential affinity of P46 for the IRS sequence was examined by competition experiments; 15 min prior to the addition of the radiolabeled IRS, blots were incubated with a 200-fold molar excess of unlabeled IRS (specific competitor) or with a 200-fold molar excess of an unlabeled double-stranded oligonucleotide that does not contain the inverted repeat sequence (5'-GGTTATTAAGTTATCGATTT-3' and 3'-CCAATAATTCAATAGCTAAA-5') (nonspecific competitor). After addition of the radiolabeled IRS to these hybridization mixtures containing the competitors, the incubation was for 1 h as described above.

**Heparin-Agarose Affinity Chromatography**—Column operations were carried out by using a Bio-Pilot Scale FPLC system (Amersham Pharmacia Biotech). About 70 mg of total proteins were loaded onto HiTrap heparin column (1 ml, Amersham Pharmacia Biotech) equilibrated at a flow rate of 1 ml/min with ice-cold BB buffer. The column was washed with ice-cold BB buffer until the 280-nm absorbance reading returned to the base line, and proteins were eluted with a 20-ml linear gradient of 50 mM to 1 M KCl in BB buffer. Fractions of 0.3 ml were collected, and 10–50- $\mu$ l aliquots from each fraction were analyzed on SDS-PAGE gels by Coomassie Blue staining and tested for the presence of IRS-binding proteins by Southwestern blot analyses. Fractions containing IRS-binding proteins were loaded on a gel filtration chromatography column without freezing. Remaining fractions were stored frozen at -70 °C.

**Gel Filtration Chromatography**—Fractions (200  $\mu$ l) collected from the affinity chromatography containing IRS-binding proteins were loaded onto an FPLC Superose 12 column (Amersham Pharmacia Biotech) equilibrated with running buffer (BB buffer containing 200 mM KCl) at a flow rate of 0.2 ml/min. The column was eluted at 0.2 ml/min with running buffer; and 0.2-ml fractions were collected, and 10–100- $\mu$ l aliquots of each fraction were analyzed on SDS-PAGE gels by Coomassie Blue staining and tested for the presence of IRS-binding proteins by Southwestern analysis. Column calibration was performed using blue dextran 2000 (for void volume), catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa) (low molecular weight gel filtration calibration kit from Amersham Pharmacia Biotech).

**Sequencing of the Purified P46 Protein**—P46 containing fraction, recovered from gel filtration chromatography, was resolved on a 10% SDS-polyacrylamide gel. P46 protein was visualized by an overnight amido black staining (16). P46 band was excised and brought to near dryness under vacuum. P46 microsequencing was made by the "Laboratoire de Microsequencage des Proteines" (Institut Pasteur, Paris, France). After proteolytic digestion with endopeptidase Lys-C, three peptides, designated P1, P2, and P3, were sequenced on an Applied Biosystems 473A protein sequencer.

**General Methods in Molecular Biology**—Standard procedures in mo-

lecular biology were used for preparation of plasmid DNA, restriction enzyme digestion, DNA agarose gel electrophoresis, DNA ligation, and transformation of *E. coli* (15).

**Cloning and Sequencing of the P46 Gene**—To amplify *S. citri* genomic DNA, two pairs of degenerated primers were synthesized based on amino acid sequences of two peptides (P1 and P2) derived from purified P46. The codon usage of *S. citri*, based on previously sequenced genes (17), was used to reduce degeneration of the primers. PCR amplification was performed for 40 cycles in a DNA thermal cycler (Perkin-Elmer), using 40 cycles, each of 1 min at 92 °C, 1 min at 37 °C, 1 min at 72 °C. The primer pair, P46-2, 5'-GAAAA(C/T)AC(A/T)GC(A/T)AT(A/T)AA(C/T)GT(A/T)AA-3' (a sense oligonucleotide corresponding to the C-terminal end of internal peptide sequence P2, KNSGENTAINVK), and P46-1c, 5'-TT(C/T)CA(A/G)TT(A/T)GT(A/C/T)CC(A/G)TA-3' (an antisense oligonucleotide corresponding to the N-terminal end of internal peptide sequence P1, KEYTYGTNWK), was found to drive the amplification of a 0.7-kbp DNA fragment. This fragment was cloned into the pTAG vector (R & D Systems) that has complementary T overhangs, suitable for ligation of PCR fragments. This fragment was sequenced using T7 sequencing kit (Amersham Pharmacia Biotech) together with pTAG SEQ 5' and pTAG SEQ 3' primers (R & D Systems). The insert was excised from the vector, separated on agarose gel, purified by the GeneClean Kit (Bio 101, La Jolla, CA), random primer-labeled with [ $\alpha$ -<sup>32</sup>P]ATP (ICN Biochemicals), purified on a Sephadex G-50 column (15), and used as a probe in Southern blot experiments (18) with *S. citri* restricted genomic DNA. A 3.5-kbp *EcoRI* DNA fragment was detected. The probe was then used to screen an *S. citri* genomic library, constructed in *E. coli* XL1 Blue by cloning *EcoRI*-restricted DNA fragments into plasmid pBS+ (Stratagene Cloning Systems). One positive clone containing the 3.5-kbp *EcoRI* insert was selected. Insert was sequenced using the T7 sequencing kit (Amersham Pharmacia Biotech) or the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech).

**Sequence Analysis**—Sequences were analyzed using the Wisconsin Package (version 9.0) of software programs from Genetics Computer Group (GCG, Madison, WI) (19). Potential ORFs were examined by codon bias analysis (20) from Sequaid II software with the codon frequency table of *S. citri* based on previously sequenced genes (17). The proteins deduced from the ORFs were submitted for BLASTP (21) analysis against GenBank™ (National Center for Biotechnology Information at the National Library of Medicine, National Institutes of Health, Bethesda). Potentially homologous proteins were compared using GAP from GCG Package. The ProDom protein domain families data base<sup>2</sup> (22, 23) was used for analysis of domain arrangement of proteins. Multiple alignments were computed with MultAlign (24).

***S. citri* Ribosomes Purification**—Enriched ribosomal fraction was obtained using a modification of previously described methods (25, 26). *S. citri* cells were ground with alumina and suspended in buffer A (10 mM Tris-HCl, pH 7.8, 10 mM magnesium acetate, 60 mM NH<sub>4</sub>Cl, 6 mM 2-mercaptoethanol) containing DNase I at 2  $\mu$ g/ml. The extract was centrifuged at 22,000 × *g* for 30 min to remove cell debris and then at 33,000 × *g* for 30 min to obtain the S-30 fraction. The S-30 fraction was centrifuged at 105,000 × *g* for 2 h to sediment the ribosomes. The supernatant solution (S-105) was aspirated, and the ribosomes were suspended in buffer A and were washed by centrifuging again at 105,000 × *g* for 2 h. The supernatant fluid was decanted and discarded, and the ribosomes were suspended in buffer A. Enriched ribosomal fraction was aliquoted and stored at -70 °C or used directly in Southwestern experiments. Final protein concentration was 0.7 mg/ml in the ribosome fraction and 1.5 mg/ml in the S-105 fraction. 50- $\mu$ l aliquots of ribosome fraction and S-105 fraction were tested for the presence of IRS-binding proteins by Southwestern analysis (see above).

## RESULTS

We describe here the purification, preliminary characterization, and identification of an *S. citri* protein that displays preferential binding affinity for a 30-bp oligonucleotide containing the 20-bp inverted repeat sequence present at the 3'-OH end of *rpsB* and preceding *tsf* (Fig. 1). The 30-bp sequence will be designated IRS.

**Heparin-Agarose Affinity Chromatography**—Crude extract of *S. citri* cells obtained by sonication in binding buffer BB, as described under "Experimental Procedures," was applied onto

<sup>2</sup> Available on the WWW at <http://protein.toulouse.inra.fr/prodom.html>.

EDTA, 1 mM dithiothreitol, 12.5% glycerol) containing 1 mM phenylmethylsulfonyl fluoride, and disrupted by sonication (5/10-s pulses, 50 watts, 0 °C, 1 min sonication, and 1 min on ice alternatively, 8 to 10 times with a Vibracell sonicator). Mixture was cleared by centrifugation at  $12,000 \times g$  for 20 min; the pellet was discarded, and the supernatant was directly loaded onto heparin chromatography column or frozen at  $-70$  °C for storage. The protein concentration of the supernatant, assayed by the Bradford procedure (11) using the Bio-Rad Protein Assay Kit with bovine serum albumin as a standard, was in the range of 1–10 mg/ml.

**Southwestern Blot Analysis of Protein-DNA Interactions**—The method used for Southwestern analysis is a modification of previously described "protein blotting" method (12). Crude or purified proteins were resolved by SDS-PAGE on 10% polyacrylamide gel using standard procedures (13). After electrophoresis, proteins were electroblotted (14) to nitrocellulose membranes (C extra, Amersham Pharmacia Biotech) for 45 min at 8 watts in transfer buffer (25 mM Tris base, 190 mM glycine, and 15% methanol) using a semi-dry transfer apparatus (Fast-blot Biometra). Membranes were dried 15 min at 37 °C and blocked for 60 min at 26 °C in buffer S (BB buffer containing 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, and 0.02% bovine serum albumin). Membranes were incubated in 10 ml of S buffer containing  $10^6$  cpm/ml end-labeled specific IRS probe. This probe is the synthetic double-stranded oligonucleotide (5'-GGATTTGCTTTACCAAAAGCAAAAAGCTG-3' and 3'-CCTAAACGAAATGGTTTTTCGTTTTTCGAC-5') containing the inverted repeat sequence (Fig. 1). It was end-labeled with [ $\gamma$ - $^{32}$ P]ATP (ICN Biochemicals) using T4 polynucleotide kinase and purified using a Sephadex G-25 column (15).

Incubations were at 26 °C for 60 min in screw-capped glass tubes with gentle rotation in rotating hybridization oven (Robbins Scientific), followed by three washes of 15 min in S buffer. The blots were then exposed 1–12 h for autoradiography.

Preferential affinity of P46 for the IRS sequence was examined by competition experiments; 15 min prior to the addition of the radiolabeled IRS, blots were incubated with a 200-fold molar excess of unlabeled IRS (specific competitor) or with a 200-fold molar excess of an unlabeled double-stranded oligonucleotide that does not contain the inverted repeat sequence (5'-GGTTATTAAGTTATCGATTT-3' and 3'-CCAATAATTCAATAGCTAAA-5') (nonspecific competitor). After addition of the radiolabeled IRS to these hybridization mixtures containing the competitors, the incubation was for 1 h as described above.

**Heparin-Agarose Affinity Chromatography**—Column operations were carried out by using a Bio-Pilot Scale FPLC system (Amersham Pharmacia Biotech). About 70 mg of total proteins were loaded onto HiTrap heparin column (1 ml, Amersham Pharmacia Biotech) equilibrated at a flow rate of 1 ml/min with ice-cold BB buffer. The column was washed with ice-cold BB buffer until the 280-nm absorbance reading returned to the base line, and proteins were eluted with a 20-ml linear gradient of 50 mM to 1 M KCl in BB buffer. Fractions of 0.3 ml were collected, and 10–50- $\mu$ l aliquots from each fraction were analyzed on SDS-PAGE gels by Coomassie Blue staining and tested for the presence of IRS-binding proteins by Southwestern blot analyses. Fractions containing IRS-binding proteins were loaded on a gel filtration chromatography column without freezing. Remaining fractions were stored frozen at  $-70$  °C.

**Gel Filtration Chromatography**—Fractions (200  $\mu$ l) collected from the affinity chromatography containing IRS-binding proteins were loaded onto an FPLC Superose 12 column (Amersham Pharmacia Biotech) equilibrated with running buffer (BB buffer containing 200 mM KCl) at a flow rate of 0.2 ml/min. The column was eluted at 0.2 ml/min with running buffer; and 0.2-ml fractions were collected, and 10–100- $\mu$ l aliquots of each fraction were analyzed on SDS-PAGE gels by Coomassie Blue staining and tested for the presence of IRS-binding proteins by Southwestern analysis. Column calibration was performed using blue dextran 2000 (for void volume), catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa) (low molecular weight gel filtration calibration kit from Amersham Pharmacia Biotech).

**Sequencing of the Purified P46 Protein**—P46 containing fraction, recovered from gel filtration chromatography, was resolved on a 10% SDS-polyacrylamide gel. P46 protein was visualized by an overnight amido black staining (16). P46 band was excised and brought to near dryness under vacuum. P46 microsequencing was made by the "Laboratoire de Microsequencage des Proteines" (Institut Pasteur, Paris, France). After proteolytic digestion with endopeptidase Lys-C, three peptides, designated P1, P2, and P3, were sequenced on an Applied Biosystems 473A protein sequencer.

**General Methods in Molecular Biology**—Standard procedures in mo-

lecular biology were used for preparation of plasmid DNA, restriction enzyme digestion, DNA agarose gel electrophoresis, DNA ligation, and transformation of *E. coli* (15).

**Cloning and Sequencing of the P46 Gene**—To amplify *S. citri* genomic DNA, two pairs of degenerated primers were synthesized based on amino acid sequences of two peptides (P1 and P2) derived from purified P46. The codon usage of *S. citri*, based on previously sequenced genes (17), was used to reduce degeneration of the primers. PCR amplification was performed for 40 cycles in a DNA thermal cycler (Perkin-Elmer), using 40 cycles, each of 1 min at 92 °C, 1 min at 37 °C, 1 min at 72 °C. The primer pair, P46-2, 5'-GAAAA(C/T)AC(A/T)GC(A/T)AT(A/T)AA(C/T)GT(A/T)AA-3' (a sense oligonucleotide corresponding to the C-terminal end of internal peptide sequence P2, KNSGENTAINVK), and P46-1c, 5'-TT(C/T)CA(A/G)TT(A/T)GT(A/C/T)CC(A/G)TA-3' (an antisense oligonucleotide corresponding to the N-terminal end of internal peptide sequence P1, KEYTYGTNWK), was found to drive the amplification of a 0.7-kbp DNA fragment. This fragment was cloned into the pTAG vector (R & D Systems) that has complementary T overhangs, suitable for ligation of PCR fragments. This fragment was sequenced using T7 sequencing kit (Amersham Pharmacia Biotech) together with pTAG SEQ 5' and pTAG SEQ 3' primers (R & D Systems). The insert was excised from the vector, separated on agarose gel, purified by the GeneClean Kit (Bio 101, La Jolla, CA), random primer-labeled with [ $\alpha$ - $^{32}$ P]ATP (ICN Biochemicals), purified on a Sephadex G-50 column (15), and used as a probe in Southern blot experiments (18) with *S. citri* restricted genomic DNA. A 3.5-kbp *EcoRI* DNA fragment was detected. The probe was then used to screen an *S. citri* genomic library, constructed in *E. coli* XL1 Blue by cloning *EcoRI*-restricted DNA fragments into plasmid pBS+ (Stratagene Cloning Systems). One positive clone containing the 3.5-kbp *EcoRI* insert was selected. Insert was sequenced using the T7 sequencing kit (Amersham Pharmacia Biotech) or the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech).

**Sequence Analysis**—Sequences were analyzed using the Wisconsin Package (version 9.0) of software programs from Genetics Computer Group (GCG, Madison, WI) (19). Potential ORFs were examined by codon bias analysis (20) from Sequaid II software with the codon frequency table of *S. citri* based on previously sequenced genes (17). The proteins deduced from the ORFs were submitted for BLASTP (21) analysis against GenBank™ (National Center for Biotechnology Information at the National Library of Medicine, National Institutes of Health, Bethesda). Potentially homologous proteins were compared using GAP from GCG Package. The ProDom protein domain families data base<sup>2</sup> (22, 23) was used for analysis of domain arrangement of proteins. Multiple alignments were computed with MultAlign (24).

**S. citri Ribosomes Purification**—Enriched ribosomal fraction was obtained using a modification of previously described methods (25, 26). *S. citri* cells were ground with alumina and suspended in buffer A (10 mM Tris-HCl, pH 7.8, 10 mM magnesium acetate, 60 mM  $\text{NH}_4\text{Cl}$ , 6 mM 2-mercaptoethanol) containing DNase I at 2  $\mu\text{g/ml}$ . The extract was centrifuged at  $22,000 \times g$  for 30 min to remove cell debris and then at  $33,000 \times g$  for 30 min to obtain the S-30 fraction. The S-30 fraction was centrifuged at  $105,000 \times g$  for 2 h to sediment the ribosomes. The supernatant solution (S-105) was aspirated, and the ribosomes were suspended in buffer A and were washed by centrifuging again at  $105,000 \times g$  for 2 h. The supernatant fluid was decanted and discarded, and the ribosomes were suspended in buffer A. Enriched ribosomal fraction was aliquoted and stored at  $-70$  °C or used directly in Southwestern experiments. Final protein concentration was 0.7 mg/ml in the ribosome fraction and 1.5 mg/ml in the S-105 fraction. 50- $\mu$ l aliquots of ribosome fraction and S-105 fraction were tested for the presence of IRS-binding proteins by Southwestern analysis (see above).

## RESULTS

We describe here the purification, preliminary characterization, and identification of an *S. citri* protein that displays preferential binding affinity for a 30-bp oligonucleotide containing the 20-bp inverted repeat sequence present at the 3'-OH end of *rpsB* and preceding *tsf* (Fig. 1). The 30-bp sequence will be designated IRS.

**Heparin-Agarose Affinity Chromatography**—Crude extract of *S. citri* cells obtained by sonication in binding buffer BB, as described under "Experimental Procedures," was applied onto

<sup>2</sup> Available on the WWW at <http://protein.toulouse.inra.fr/prodom.html>.

# Purification, Cloning, and Preliminary Characterization of a *Spiroplasma citri* Ribosomal Protein with DNA Binding Capacity\*

(Received for publication, February 18, 1998, and in revised form, May 26, 1998)

Loïck Le Dantec<sup>‡§</sup>, Michel Castroviejo<sup>¶</sup>, Joseph M. Bové<sup>‡</sup>, and Colette Saillard<sup>‡¶</sup>From the <sup>‡</sup>Laboratoire de Biologie Cellulaire et Moléculaire, Institut National de la Recherche Agronomique and Université Victor Segalen Bordeaux 2, 33883 Villenave d'Ornon Cedex and the <sup>¶</sup>Institut de Biochimie et Génétique Cellulaires and Centre National de la Recherche Scientifique, 33077 Bordeaux Cedex, France

The *rpsB*-*tsf*-*x* operon of *Spiroplasma citri* encodes ribosomal protein S2 and elongation factor Ts, two components of the translational apparatus, and an unidentified X protein. A potential DNA-binding site (a 20-base pair (bp) inverted repeat sequence) is located at the 3' end of *rpsB*. Southwestern analysis of *S. citri* proteins, with a 30-bp double-stranded oligonucleotide probe (IRS), containing the 20-bp inverted repeat sequence and the genomic flanking sequences, detected an IRS-binding protein of 46 kDa (P46). P46 protein, which displays preferential affinity for the IRS, was purified from *S. citri* by a combination of affinity and gel filtration chromatographies. The native form of P46 seems to be homomultimeric as estimated by SDS-polyacrylamide gel electrophoresis analysis and gel filtration. A 3.5-kilobase pair *S. citri* DNA fragment comprising the P46 gene and flanking sequences was cloned and sequenced. Sequence analysis of this DNA fragment indicated that the P46 gene is located within the S10-*spc* operon of *S. citri* at the position of the gene coding for ribosomal protein L29 in the known S10-*spc* operons. The similarity between the N-terminal domain of P46 and the L29 ribosomal protein family and the presence of a 46-kDa IRS-binding protein in *S. citri* ribosomes indicated that P46 is the L29 ribosomal protein of *S. citri*. We suggest that P46 is a bifunctional protein with an L29 N-terminal domain and a C-terminal domain involved in IRS binding.

*Spiroplasmas* are wall-free bacteria belonging to the class Mollicutes, a group of microorganisms phylogenetically related to Gram-positive bacteria with low guanine + cytosine contents (1). Sequence analysis (70) of a 6.8-kbp<sup>1</sup> DNA fragment (GenBank<sup>TM</sup> accession number AF012877) of the phytopathogenic mollicute *Spiroplasma citri* (2) made it possible to identify eight putative ORFs that encode ribosomal protein S2, elonga-

tion factor Ts, spiralin, 6-phosphofructokinase, pyruvate kinase, and three unidentified proteins (A, B, and X) (Fig. 1). Ribosomal protein S2 and the translational elongation factor Ts (Ef-Ts), respectively, encoded by *rpsB* and *tsf* genes, are both components of the translational apparatus in prokaryotes. These genes are adjacent in *Escherichia coli* (3, 4) and *Bacillus subtilis* (5), whereas in the genome of the two mollicutes *Mycoplasma genitalium* (6) and *Mycoplasma pneumoniae* (7) they reside at different locations, and thus each of them may constitute monocistronic transcriptional units or be part of two different polycistronic operons.

The organization and relative orientation of *rpsB* and *tsf* of *S. citri* (Fig. 1) are analogous to those reported for *E. coli* (3, 4) and *B. subtilis* (5). In *E. coli*, *rpsB* and *tsf* form a single transcriptional unit, and an attenuation mechanism was proposed to explain the 1 to 3 ratio of Ef-Ts to S2 (3). In *B. subtilis* a potential terminator is found between *rpsB* and *tsf*. In *S. citri*, the absence of a rho-independent termination signal in the spacer region between *rpsB* and *tsf*, and between *tsf* and *x*, indicates that *rpsB*, *tsf*, and *X* might represent a single transcriptional unit. Transcriptional analyses of *rpsB*, *tsf*, and *x* genes of *S. citri* have recently revealed two different transcripts (70), one corresponding to the *rpsB*/*tsf*/*x* operon and the second to *rpsB* alone. These results suggested that a regulatory mechanism may act at the transcriptional level at the spacer region between *rpsB* and *tsf*. The only "regulatory" like sequence found in the *rpsB*-*tsf* region is an inverted repeat sequence at the 3' end of *rpsB* (Fig. 1). This inverted repeat sequence is 20 bp long and represents two turns of helical DNA. It could be a binding site for a regulatory DNA-binding protein (8, 9).

In the study reported here, we have purified, by a combination of affinity and gel filtration chromatographies, a 46-kDa protein (P46) that displays preferential binding to a 30-bp double-stranded oligonucleotide containing the inverted repeat sequence. The gene coding for P46 has been cloned and sequenced. Sequence analysis has revealed significant similarities between the N-terminal part of P46 and the L29 eubacterial ribosomal protein family. Surprisingly, the *S. citri* protein is much larger than its eubacterial homologs, and the C-terminal domain of P46 shows significant similarities with the DNA-binding histone H1-like proteins found in some bacterial species. These results suggest a bifunctional role for P46. The protein could act as a ribosomal protein but also as a DNA-binding protein with a potential regulatory function.

## EXPERIMENTAL PROCEDURES

**Preparation of Crude Extracts**—Crude extracts were prepared from *S. citri* strain RSA2 (ATCC 27556) cultivated in 5 to 10 liters of SP4 medium (10) at 32 °C. Cells from exponential growth phase were harvested by centrifugation at 17,000 × *g* for 40 min, washed with phosphate-buffered saline containing 70 g/liter sorbitol, and frozen at -20 °C overnight. Cells were then thawed on ice, resuspended in 10 to 30 ml of binding buffer BB (10 mM HEPES, pH 7.9, 50 mM KCl, 1 mM

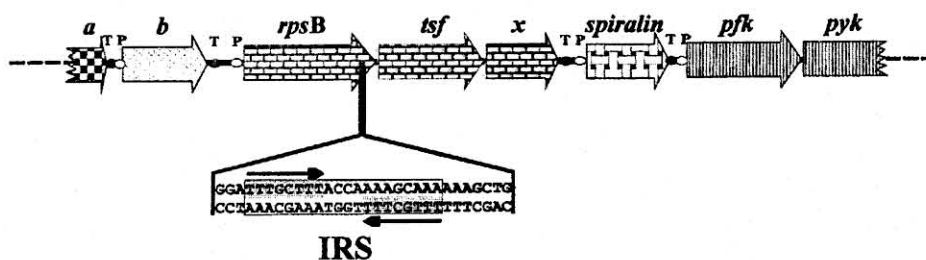
\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank<sup>TM</sup>/EBI Data Bank with accession number(s) AF031160.

§ Supported by the Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche.

¶ To whom correspondence should be addressed: Laboratoire de Biologie Cellulaire et Moléculaire, Institut National de la Recherche Agronomique, Domaine de la Grande Ferrade, BP 81, 33883 Villenave d'Ornon Cedex, France. Tel.: 33 5 56 84 31 52; Fax: 33 5 56 84 31 59; E-mail: saillard@bordeaux.inra.fr.

<sup>1</sup> The abbreviations used are: kbp, kilobase pair(s); bp, base pair(s); FPLC, fast protein liquid chromatography; ORF, open reading frame; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; Ef-Ts, elongation factor Ts.

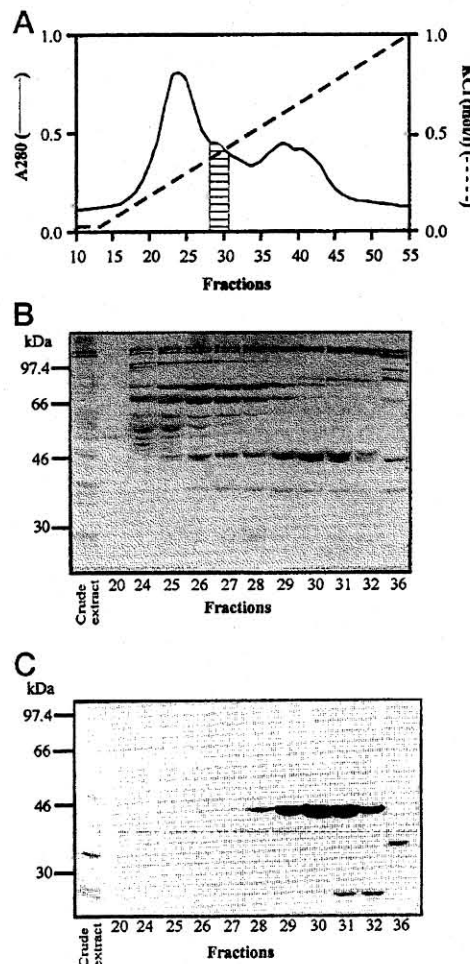


**FIG. 1. Organization of genes on the chromosomal DNA of *S. citri* in the spiralin gene region.** Translational products of ORFs *a*, *b*, and *x* are not significantly similar to any other proteins in the data banks. *rpsB*, *tsf*, *spiralin*, *pfk*, and *pyk* code, respectively, for ribosomal protein S2, elongation factor Ts (Ef-Ts), spiralin, 6-phosphofructokinase, and pyruvate kinase. The 30-bp sequence, designated IRS, used as a probe in Southwestern experiments is presented at the bottom. The 20-bp inverted repeat sequence is boxed. Potential promoters (P) and terminators (T) are represented by open and solid circles, respectively.

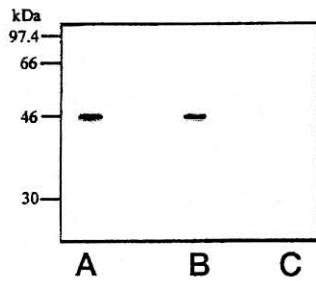
a heparin-agarose affinity column. Bound proteins were eluted with a linear gradient of 0.05–1 M KCl. Protein concentration was monitored by absorbance at 280 nm (Fig. 2A). Aliquots of the eluted fractions were first analyzed by SDS-PAGE electrophoresis and Coomassie Blue staining (Fig. 2B). Detection of IRS-binding proteins was carried out by Southwestern blots (12, 27, 28). In short, proteins were resolved on SDS-PAGE, transferred on a nitrocellulose sheet, and incubated with radiolabeled IRS. After washing, IRS-binding proteins were detected by autoradiography. Southwestern analysis of eluted fractions detected a predominant IRS-binding protein of 46 kDa (P46) (Fig. 2C) eluted at approximately 0.37–0.42 M KCl as indicated in Fig. 2A. Two other minor IRS-binding protein species of approximately 25 and 38 kDa were detected in fractions 31–32 and 36, respectively. Part of the binding activity was lost after freezing/thawing cycles; therefore active fractions were used for further experiments immediately after preparation.

**Preferential Affinity of P46 for IRS**—To determine if the P46 protein displays some preferential affinity for the IRS sequence, we have used Southwestern competition assay. Aliquots of fraction F29 containing P46 protein eluted from the heparin column (Fig. 2) were resolved on three separate tracks (Fig. 3) by 10% SDS-PAGE and subjected to Southwestern analysis. Membranes were incubated with radiolabeled IRS (Fig. 3 lane A) or with radiolabeled IRS in the presence of cold competitors (Fig. 3, lanes B and C). Binding of radiolabeled IRS to P46 was not significantly reduced by a 200-fold molar excess of a non-labeled double-stranded oligonucleotide (Fig. 3, lane B), which does not contain the inverted repeat sequence, as compared with the IRS binding in the absence of competitor (Fig. 3, lane A). Competition by a 200-fold molar excess of non-labeled IRS reduced dramatically, or made undetectable, the binding of radiolabeled IRS (Fig. 3, lane C). These results strongly suggested that P46 exhibits preferential binding affinity for IRS.

**Gel Filtration Chromatography**—In a second step toward P46 purification, 200- $\mu$ l aliquot of fraction F29 containing P46 protein retrieved from the heparin-agarose affinity column (Fig. 2) was loaded onto an FPLC Superose 12 column. Proteins were eluted at 0.2 ml/min with running buffer. Protein concentration was monitored by absorbance at 280 nm (Fig. 4A). Aliquots of eluted fractions were analyzed by SDS-PAGE, and proteins were stained with Coomassie Blue (Fig. 4B). IRS binding proteins were analyzed by Southwestern blots (Fig. 4C). Almost all of the P46 IRS binding activity was present in fractions 26–28, and, as judged by SDS-PAGE and Coomassie Blue staining, P46 was practically the only protein seen in these fractions. Comparison of the elution time of P46 from the Superose column with that of molecular mass markers including catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25



**FIG. 2. Fractionation of *S. citri* IRS-binding proteins by heparin-agarose affinity chromatography.** A, elution profile of *S. citri* proteins from heparin-agarose column. Crude extracts (approximately 70 mg of total protein) were loaded onto the pre-equilibrated heparin column (see "Experimental Procedures"). After washing, proteins were eluted with a linear gradient from 50 to 1000 mM KCl. Eluted proteins were monitored by absorbance at 280 nm (A280). Fractions (300  $\mu$ l) were collected. The hatched area on protein profile represents the P46 containing fractions. B, SDS-PAGE analysis of fractions eluted from the heparin-agarose column. 50- $\mu$ l aliquots of eluted fractions from the heparin-agarose column were resolved on 10% denaturing SDS-polyacrylamide gel, and proteins were detected by Coomassie Blue staining. C, Southwestern analysis of proteins eluted from the heparin-agarose column. 50- $\mu$ l aliquots of eluted fractions from the heparin-agarose column were resolved on 10% denaturing SDS-polyacrylamide gel and electroblotted onto nitrocellulose membrane. The blot was incubated with radiolabeled IRS oligonucleotide. Details of the procedure are described under "Experimental Procedures."



**FIG. 3. Effect of competitors on P46 affinity for IRS.** 10  $\mu$ l (4  $\mu$ g) (tracks A–C) of P46 containing fraction 29 eluted from the heparin-agarose column (see Fig. 2) were resolved by 10% SDS-PAGE and electroblotted onto a nitrocellulose membrane. Tracks A–C were cut and treated separately. Control track A was directly incubated, for 1 h, with radiolabeled IRS. Tracks B and C were preincubated for 15 min before the addition of the radiolabeled IRS probe, with a 200-fold molar excess of an unlabeled double-stranded oligonucleotide competitor that does not contain the IRS (track B) or with a 200-fold molar excess of unlabeled IRS (track C) (see “Experimental Procedures”). After addition of the radiolabeled IRS to the hybridization mixtures containing the competitors, incubation was conducted as for the control track A. After washing, the membranes were subjected to autoradiography.

kDa), and ribonuclease A (13.7 kDa) indicated that P46 eluted at a position consistent with a molecular mass of approximately 180 kDa. This suggests that the native form of P46 is a homomultimer.

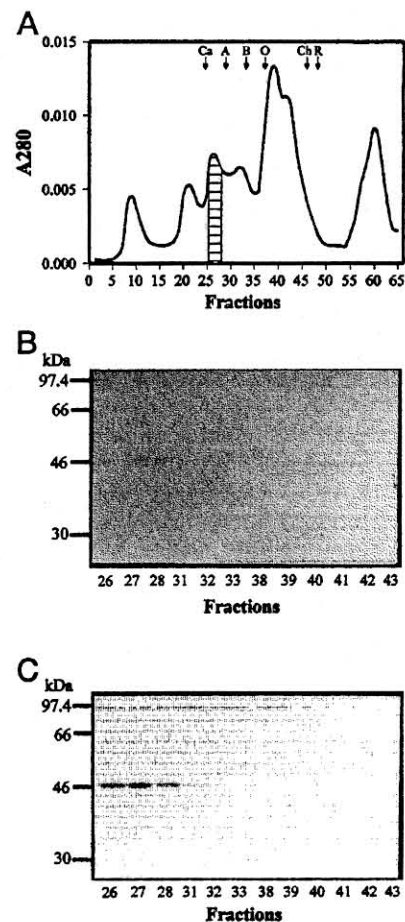
**Identification of the Gene Coding Protein P46**—To characterize further P46 we cloned its gene. First, P46 was partially sequenced. To that purpose an aliquot of fraction F27 containing P46 from the gel filtration purification step (Fig. 4) was resolved by 10% SDS-PAGE, and proteins were stained with amido black (16). The P46 band was excised from the gel and submitted to endoproteolysis with Endo Lys-C. The proteolytic products were separated by reversed phase high performance liquid chromatography and three peptides (P1, P2, and P3) were sequenced. None of these sequences exhibited homology to any known protein sequence available in the data bases. Amino acid sequences of peptides P1 (KEYTYGTNWK) and P2 (KNSGENTAINVK) were used to design two pairs of degenerated primers in order to amplify by polymerase chain reaction (PCR) part of the gene of the P46 protein. An amplified DNA fragment of 700 bp was generated using the P462/P461c primer pair (see “Experimental Procedures”), cloned into pTAG vector, and sequenced. Sequence analysis of the cloned DNA fragment revealed that the amino acid sequence predicted from the amplified product contained the N-terminal sequence of the P1 peptide and the entire sequence of the P3 peptide (KIDLELTK), confirming that this amplified product was generated from the P46 gene.

Cloning of the full-length P46 gene was undertaken to gain further sequence information. A 3.5-kbp *Eco*RI *S. citri* DNA fragment was detected by Southern blot analysis of restricted genomic DNA using the 700-bp DNA fragment as a probe. Screening of an *Eco*RI *S. citri* genomic library with this probe yielded a clone containing a 3.5-kbp insert which was sequenced. Sequence analysis (details are described under “Experimental Procedures”) of the 3.5-kbp insert gave the following results.

Authenticity of the insert was confirmed by perfect matches with the sequence of the three peptides (P1, P2, and P3) derived from purified P46 (Fig. 5).

The molecular mass deduced from the ORF of the cloned P46 gene was 37 kDa, whereas its apparent molecular mass in SDS-PAGE was 46 kDa.

Sequence similarities between deduced amino acid sequences of the insert and protein data bases (see Table I) showed that the gene coding for the P46 protein was located



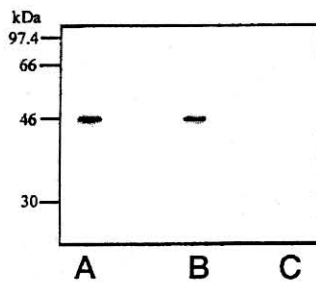
**FIG. 4. Purification of P46 by gel filtration chromatography.** A, elution profile of proteins of fraction F29. 200  $\mu$ l of fraction F29 eluted from the heparin-agarose column (Fig. 2) were loaded onto the pre-equilibrated Superose 12 column (see “Experimental Procedures”). The proteins were eluted from the column at 0.2 ml/min, and eluted proteins were monitored by absorbance at 280 nm ( $A_{280}$ ). The hatched area represents the P46 containing fractions. Ca, A, B, O, Ch, and R indicate the positions of the molecular mass markers catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa), respectively. B, SDS-PAGE analysis of fractions eluted from the Superose 12 column. 100- $\mu$ l aliquots of eluted fractions from the Superose 12 column were resolved by 10% SDS-PAGE, and proteins were detected by Coomassie Blue staining. C, Southwestern analysis of the fractions eluted from the Superose 12 column. 100- $\mu$ l aliquots of eluted fractions from the Superose 12 column were assayed for IRS binding activity using Southwestern analysis.

within a ribosomal protein operon corresponding to the S10 and *spc* operons of *E. coli* (29–31), *B. subtilis* (32–34), *Mycoplasma capricolum* (35), *M. genitalium* (6), and *M. pneumoniae* (7, 36) (see Fig. 6).

No transcriptional signal sequences were found between the genes for ribosomal proteins S17 and L14, and therefore it appears that in *S. citri*, as in *M. capricolum* (35), *M. genitalium* (6), *M. pneumoniae* (36), and *B. subtilis* (34), operons S10 and *spc* are fused into one operon (Fig. 6), whereas in *E. coli* they are not (29, 37).

Two of the intergenic regions, between the genes for S3 and L16, P46 and S17, have overlapping translational stop/start codons, in which the third nucleotide of the stop codon TAA is the first nucleotide of the start codon ATG of the gene following. This kind of overlapping was also reported in *E. coli* (30, 31) and in *M. capricolum* (35) between the genes for L4 and L23, L16 and L29, and L29 and S17.

The gene coding for P46 is located within the S10-*spc* operon



**FIG. 3. Effect of competitors on P46 affinity for IRS.** 10  $\mu$ l (4  $\mu$ g) (tracks A–C) of P46 containing fraction 29 eluted from the heparin-agarose column (see Fig. 2) were resolved by 10% SDS-PAGE and electroblotted onto a nitrocellulose membrane. Tracks A–C were cut and treated separately. Control track A was directly incubated, for 1 h, with radiolabeled IRS. Tracks B and C were preincubated for 15 min before the addition of the radiolabeled IRS probe, with a 200-fold molar excess of an unlabeled double-stranded oligonucleotide competitor that does not contain the IRS (track B) or with a 200-fold molar excess of unlabeled IRS (track C) (see "Experimental Procedures"). After addition of the radiolabeled IRS to the hybridization mixtures containing the competitors, incubation was conducted as for the control track A. After washing, the membranes were subjected to autoradiography.

kDa), and ribonuclease A (13.7 kDa) indicated that P46 eluted at a position consistent with a molecular mass of approximately 180 kDa. This suggest that the native form of P46 is a homomultimer.

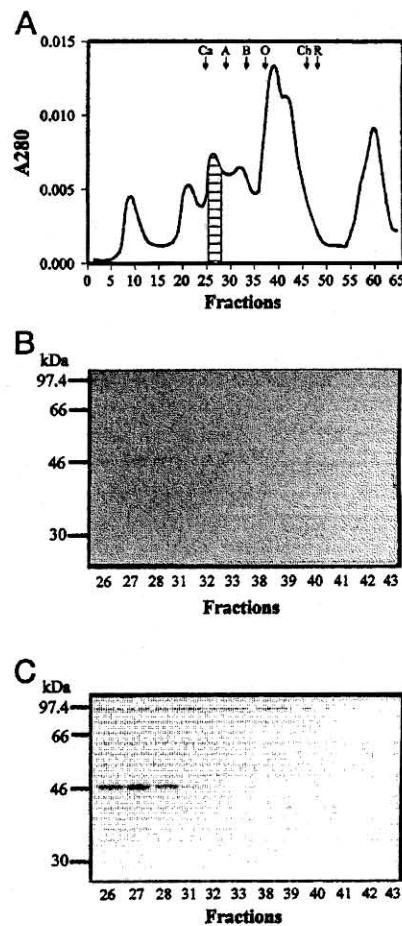
**Identification of the Gene Coding Protein P46**—To characterize further P46 we cloned its gene. First, P46 was partially sequenced. To that purpose an aliquot of fraction F27 containing P46 from the gel filtration purification step (Fig. 4) was resolved by 10% SDS-PAGE, and proteins were stained with amido black (16). The P46 band was excised from the gel and submitted to endoproteolysis with Endo Lys-C. The proteolytic products were separated by reversed phase high performance liquid chromatography and three peptides (P1, P2, and P3) were sequenced. None of these sequences exhibited homology to any known protein sequence available in the data bases. Amino acid sequences of peptides P1 (KEYTYGTNWK) and P2 (KNSGENTAINVK) were used to design two pairs of degenerated primers in order to amplify by polymerase chain reaction (PCR) part of the gene of the P46 protein. An amplified DNA fragment of 700 bp was generated using the P462/P461c primer pair (see "Experimental Procedures"), cloned into pTAG vector, and sequenced. Sequence analysis of the cloned DNA fragment revealed that the amino acid sequence predicted from the amplified product contained the N-terminal sequence of the P1 peptide and the entire sequence of the P3 peptide (KIDLELTK), confirming that this amplified product was generated from the P46 gene.

Cloning of the full-length P46 gene was undertaken to gain further sequence information. A 3.5-kbp *EcoRI* *S. citri* DNA fragment was detected by Southern blot analysis of restricted genomic DNA using the 700-bp DNA fragment as a probe. Screening of an *EcoRI* *S. citri* genomic library with this probe yielded a clone containing a 3.5-kbp insert which was sequenced. Sequence analysis (details are described under "Experimental Procedures") of the 3.5-kbp insert gave the following results.

Authenticity of the insert was confirmed by perfect matches with the sequence of the three peptides (P1, P2, and P3) derived from purified P46 (Fig. 5).

The molecular mass deduced from the ORF of the cloned P46 gene was 37 kDa, whereas its apparent molecular mass in SDS-PAGE was 46 kDa.

Sequence similarities between deduced amino acid sequences of the insert and protein data bases (see Table I) showed that the gene coding for the P46 protein was located



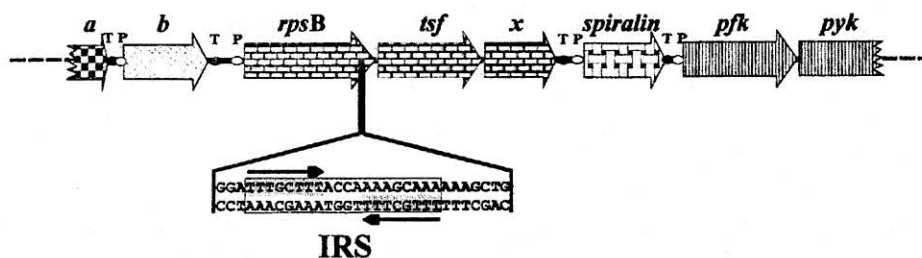
**FIG. 4. Purification of P46 by gel filtration chromatography.** A, elution profile of proteins from the gel filtration chromatography of P46 containing fraction F29. 200  $\mu$ l of fraction F29 eluted from the heparin-agarose column (Fig. 2) were loaded onto the pre-equilibrated Superose 12 column (see "Experimental Procedures"). The proteins were eluted from the column at 0.2 ml/min, and eluted proteins were monitored by absorbance at 280 nm (A280). The hatched area represents the P46 containing fractions. Ca, A, B, O, Ch, and R indicate the positions of the molecular mass markers catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa), respectively. B, SDS-PAGE analysis of fractions eluted from the Superose 12 column. 100- $\mu$ l aliquots of eluted fractions from the Superose 12 column were resolved by 10% SDS-PAGE, and proteins were detected by Coomassie Blue staining. C, Southwestern analysis of the fractions eluted from the Superose 12 column. 100- $\mu$ l aliquots of eluted fractions from the Superose 12 column were assayed for IRS binding activity using Southwestern analysis.

within a ribosomal protein operon corresponding to the S10 and *spc* operons of *E. coli* (29–31), *B. subtilis* (32–34), *Mycoplasma capricolum* (35), *M. genitalium* (6), and *M. pneumoniae* (7, 36) (see Fig. 6).

No transcriptional signal sequences were found between the genes for ribosomal proteins S17 and L14, and therefore it appears that in *S. citri*, as in *M. capricolum* (35), *M. genitalium* (6), *M. pneumoniae* (36), and *B. subtilis* (34), operons S10 and *spc* are fused into one operon (Fig. 6), whereas in *E. coli* they are not (29, 37).

Two of the intergenic regions, between the genes for S3 and L16, P46 and S17, have overlapping translational stop/start codons, in which the third nucleotide of the stop codon TAA is the first nucleotide of the start codon ATG of the gene following. This kind of overlapping was also reported in *E. coli* (30, 31) and in *M. capricolum* (35) between the genes for L4 and L23, L16 and L29, and L29 and S17.

The gene coding for P46 is located within the S10-*spc* operon

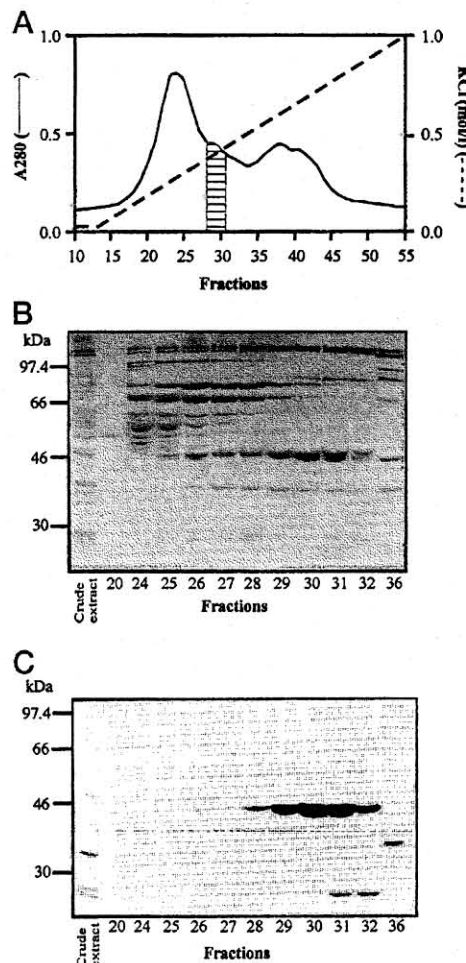


**FIG. 1. Organization of genes on the chromosomal DNA of *S. citri* in the spiralin gene region.** Translational products of ORFs *a*, *b*, and *x* are not significantly similar to any other proteins in the data banks. *rpsB*, *tsf*, *spiralin*, *pfk*, and *pyk* code, respectively, for ribosomal protein S2, elongation factor Ts (Ef-Ts), spiralin, 6-phosphofructokinase, and pyruvate kinase. The 30-bp sequence, designated IRS, used as a probe in Southwestern experiments is presented at the bottom. The 20-bp inverted repeat sequence is boxed. Potential promoters (P) and terminators (T) are represented by open and solid circles, respectively.

a heparin-agarose affinity column. Bound proteins were eluted with a linear gradient of 0.05–1 M KCl. Protein concentration was monitored by absorbance at 280 nm (Fig. 2A). Aliquots of the eluted fractions were first analyzed by SDS-PAGE electrophoresis and Coomassie Blue staining (Fig. 2B). Detection of IRS-binding proteins was carried out by Southwestern blots (12, 27, 28). In short, proteins were resolved on SDS-PAGE, transferred on a nitrocellulose sheet, and incubated with radiolabeled IRS. After washing, IRS-binding proteins were detected by autoradiography. Southwestern analysis of eluted fractions detected a predominant IRS-binding protein of 46 kDa (P46) (Fig. 2C) eluted at approximately 0.37–0.42 M KCl as indicated in Fig. 2A. Two other minor IRS-binding protein species of approximately 25 and 38 kDa were detected in fractions 31–32 and 36, respectively. Part of the binding activity was lost after freezing/thawing cycles; therefore active fractions were used for further experiments immediately after preparation.

**Preferential Affinity of P46 for IRS**—To determine if the P46 protein displays some preferential affinity for the IRS sequence, we have used Southwestern competition assay. Aliquots of fraction F29 containing P46 protein eluted from the heparin column (Fig. 2) were resolved on three separate tracks (Fig. 3) by 10% SDS-PAGE and subjected to Southwestern analysis. Membranes were incubated with radiolabeled IRS (Fig. 3 lane A) or with radiolabeled IRS in the presence of cold competitors (Fig. 3, lanes B and C). Binding of radiolabeled IRS to P46 was not significantly reduced by a 200-fold molar excess of a non-labeled double-stranded oligonucleotide (Fig. 3, lane B), which does not contain the inverted repeat sequence, as compared with the IRS binding in the absence of competitor (Fig. 3, lane A). Competition by a 200-fold molar excess of non-labeled IRS reduced dramatically, or made undetectable, the binding of radiolabeled IRS (Fig. 3, lane C). These results strongly suggested that P46 exhibits preferential binding affinity for IRS.

**Gel Filtration Chromatography**—In a second step toward P46 purification, 200- $\mu$ l aliquot of fraction F29 containing P46 protein retrieved from the heparin-agarose affinity column (Fig. 2) was loaded onto an FPLC Superose 12 column. Proteins were eluted at 0.2 ml/min with running buffer. Protein concentration was monitored by absorbance at 280 nm (Fig. 4A). Aliquots of eluted fractions were analyzed by SDS-PAGE, and proteins were stained with Coomassie Blue (Fig. 4B). IRS binding proteins were analyzed by Southwestern blots (Fig. 4C). Almost all of the P46 IRS binding activity was present in fractions 26–28, and, as judged by SDS-PAGE and Coomassie Blue staining, P46 was practically the only protein seen in these fractions. Comparison of the elution time of P46 from the Superose column with that of molecular mass markers including catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25



**FIG. 2. Fractionation of *S. citri* IRS-binding proteins by heparin-agarose affinity chromatography.** A, elution profile of *S. citri* proteins from heparin-agarose column. Crude extracts (approximately 70 mg of total protein) were loaded onto the pre-equilibrated heparin column (see "Experimental Procedures"). After washing, proteins were eluted with a linear gradient from 50 to 1000 mM KCl. Eluted proteins were monitored by absorbance at 280 nm (A280). Fractions (300  $\mu$ l) were collected. The hatched area on protein profile represents the P46 containing fractions. B, SDS-PAGE analysis of fractions eluted from the heparin-agarose column. 50- $\mu$ l aliquots of eluted fractions from the heparin-agarose column were resolved on 10% denaturing SDS-polyacrylamide gel, and proteins were detected by Coomassie Blue staining. C, Southwestern analysis of proteins eluted from the heparin-agarose column. 50- $\mu$ l aliquots of eluted fractions from the heparin-agarose column were resolved on 10% denaturing SDS-polyacrylamide gel and electroblotted onto nitrocellulose membrane. The blot was incubated with radiolabeled IRS oligonucleotide. Details of the procedure are described under "Experimental Procedures."



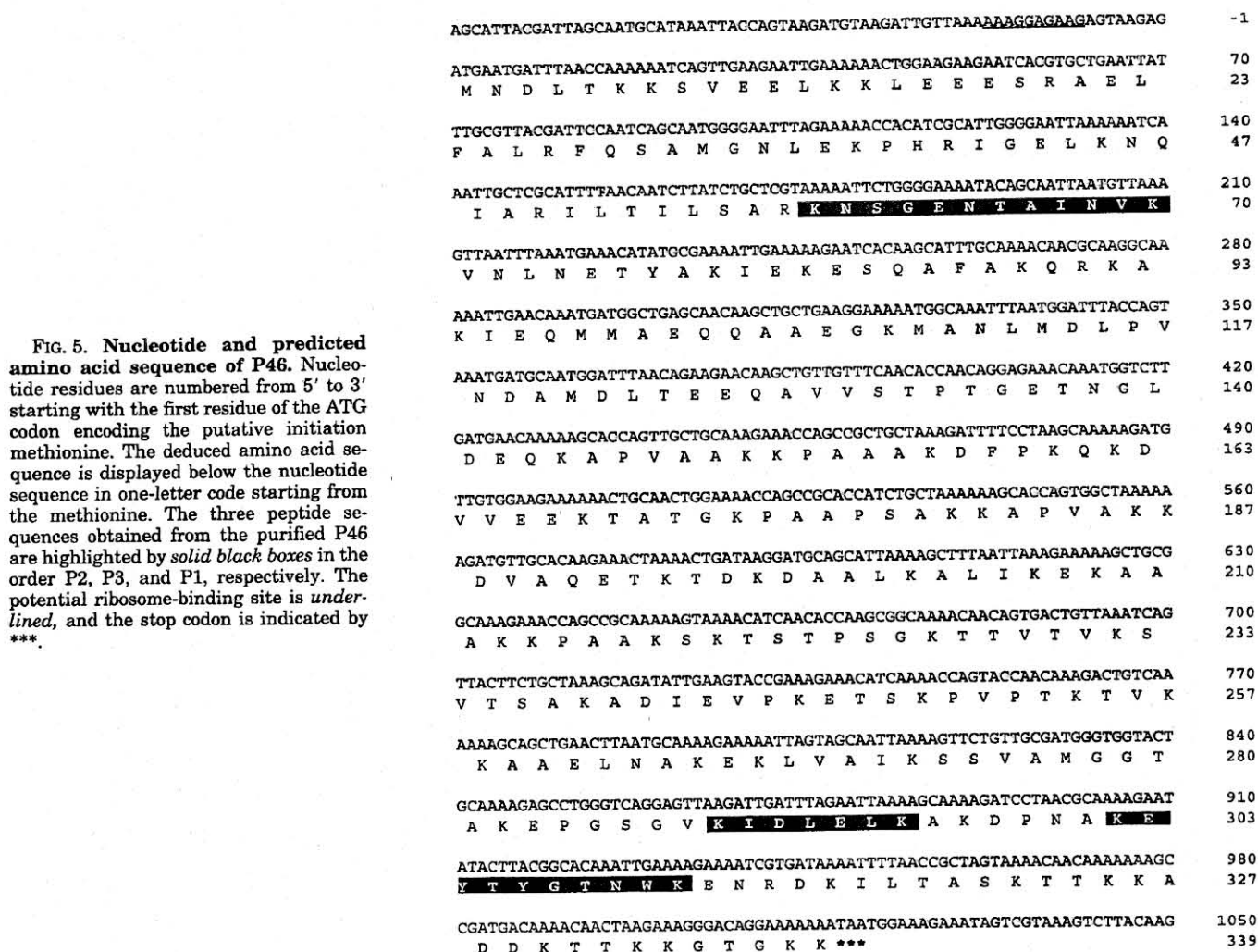


FIG. 5. Nucleotide and predicted amino acid sequence of P46. Nucleotide residues are numbered from 5' to 3' starting with the first residue of the ATG codon encoding the putative initiation methionine. The deduced amino acid sequence is displayed below the nucleotide sequence in one-letter code starting from the methionine. The three peptide sequences obtained from the purified P46 are highlighted by solid black boxes in the order P2, P3, and P1, respectively. The potential ribosome-binding site is underlined, and the stop codon is indicated by \*\*\*.

TABLE I  
ORFs and putative proteins of the 3.5-kbp *S. citri* DNA insert, containing the P46 gene (ORF V) and homologous genes and proteins from eubacteria

ORF	Location on spiroplasmal DNA		Size of putative spiroplasmal protein (no. amino acids)	Homologous protein (origin)	Size (no. amino acids)	Identity %	Corresponding gene	Accession number
	First nucleotide	Last nucleotide						
I	1 <sup>a</sup>	58	>19	Ribosomal protein S19 ( <i>M. capricolum</i> )	88	89.5	<i>rpsS</i>	p10132
II	90	426	112	Ribosomal protein L22 ( <i>M. capricolum</i> )	111	62.2	<i>rpIV</i>	p10139
III	448	1204	252	Ribosomal protein S3 ( <i>M. capricolum</i> )	232	62	<i>rpsC</i>	p02353
IV	1206	1617	137	Ribosomal protein L16 ( <i>M. capricolum</i> )	137	78.1	<i>rpIP</i>	p02415
V	1623	2640	339	Ribosomal protein L29 ( <i>M. capricolum</i> )	138	40.6	<i>rpmC</i>	p10142
VI	2642	2897	85	Ribosomal protein S17 ( <i>M. capricolum</i> )	85	75.3	<i>rpsQ</i>	p10131
VII	2921	3287	122	Ribosomal protein L14 ( <i>M. capricolum</i> )	122	74.6	<i>rpIN</i>	p10137
VIII	3301	3475 <sup>a</sup>	>58	Ribosomal protein L24 ( <i>M. capricolum</i> )	108	51.7	<i>rpIX</i>	p10141

<sup>a</sup> The genes are truncated.

between the *rpIP* and *rpsQ* genes that code, respectively, for ribosomal proteins L16 and S17. As shown in Fig. 6, comparison with similar operons in *E. coli*, *B. subtilis*, and several mollicutes revealed that P46 gene is at the position of ribosomal protein L29.

Only the N-terminal sequence of P46 shares similarities with the sequences of L29 ribosomal proteins in protein data bases. Domain analysis of P46 using the Prodom protein domain families data base (22, 23) revealed a unique feature of this protein among other known L29 proteins. P46, which is much larger than other L29 proteins, could be arranged into three domains (Fig. 7A). The first domain (domain I, from position 1-57 of the P46 sequence) matches the Prodom domain 1463

(Prodom 34.1) which is a common domain of the prokaryotic L29 family (Fig. 7B). The second domain (domain II, from position 58-137 of P46) matches the Prodom domain 24734 (Prodom 34.1) which is the C-terminal sequence of the *M. capricolum* L29 ribosomal protein (Fig. 7C). The third domain, ranging from position 138-339 on the P46 sequence shares no significant similarities with any Prodom domains.

However, BlastP analysis of this domain revealed significant similarities with DNA-binding histone H1-like proteins. It exhibits 28.6% amino acid identity to *Pseudomonas aeruginosa* transcriptional regulatory protein AlgP (38) also termed *algR3* (39), 31.3% amino acid identity to *Bordetella pertussis* histone H1 homolog BpH1 (40), and 26.3% amino acid identity to

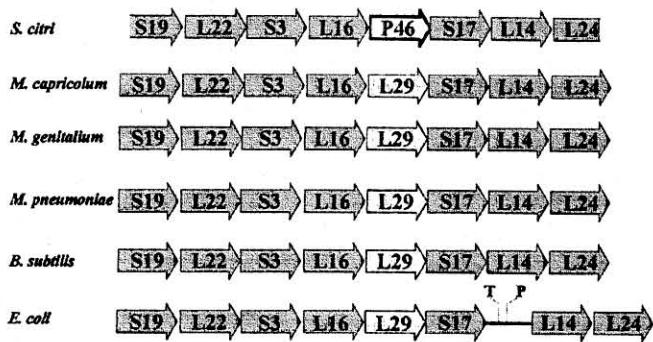


FIG. 6. Comparison of the organization of the *S. citri* ribosomal protein gene cluster with the S10 and *spc* operons of *E. coli*, *B. subtilis*, and several mollicutes. Transcription initiation (*P*) and termination (*T*) signals are indicated. Drawing is not to scale.

*Chlamydia trachomatis* histone H1 homolog Hc2 (41). Histone H1-like proteins exhibit sequence similarities to eucaryotic histones H1 and are involved in bacterial nucleoid organization and/or transcriptional regulation through their interactions with bacterial DNA. These observations suggested that the C-terminal domain (domain III, Fig. 7A) of P46 is implicated in DNA-protein interaction.

**Localization of P46**—The above results suggested that P46 is the L29 ribosomal protein of *S. citri*. To confirm this finding, we have analyzed *S. citri* ribosomal proteins. *S. citri* ribosomes were partially purified by differential ultracentrifugation (25, 26). Proteins from the supernatant (S-105) and the washed ribosomes were resolved on SDS-PAGE and examined by Southwestern analysis (Fig. 8). A 46-kDa IRS-binding protein was detected in both the S-105 supernatant (Fig. 8, lane 1) and in the washed ribosome fraction but in much larger quantities in the latter (Fig. 8, lane 2). These results indicated that P46 is present in the ribosomal fraction. Three other IRS-binding protein species of approximately 25, 26, and 38 kDa were detected in the ribosome fraction (Fig. 8, lane 2). These proteins are probably the DNA-binding proteins detected in crude extracts of spiroplasmal cells and in the eluted fractions of the heparin-agarose affinity column (Fig. 2C). The binding affinity of these proteins has not been investigated.

#### DISCUSSION

Despite the large amount of work on the regulation of ribosome synthesis (see Ref. 42 for a review), molecular mechanisms of regulation of some ribosomal protein operons have not been elucidated completely. This is the case of the *E. coli* *rpsB-tsfc* operon for which an attenuation mechanism was proposed but has not been demonstrated (3). In *S. citri*, *rpsB*, *tsfc*, and  $\alpha$  represent the three genes of a single transcriptional unit, and a potential regulatory sequence (a 20-bp inverted repeat sequence) was found at the 3' end of the *rpsB* gene (70). In the study reported in this paper we have detected, by Southwestern blot analysis, a protein (P46) that binds with preferential affinity to a 30-bp oligonucleotide (IRS) containing the 20-bp inverted repeat sequence. We have partially purified the P46 protein, cloned, and sequenced its gene.

The deduced amino acid sequence from the cloned gene indicated that the N-terminal end of P46 shared significant similarity with the L29 ribosomal protein family. In addition, sequence analysis of the whole cloned DNA insert carrying the P46 gene indicated that the P46 gene is located within the S10-*spc* operon, and more precisely between the *rpsL* and *rpsQ* genes coding, respectively, for ribosomal proteins L16 and S17 (Fig. 6). Thus, the P46 gene is located at the position of the L29 gene in the previously described S10-*spc* operons and represents a ribosomal protein gene. Indeed, Southwestern analysis

of *S. citri* ribosomal proteins revealed the presence of the 46-kDa IRS-binding protein (Fig. 8). These findings suggested that P46 is the L29 ribosomal protein of *S. citri*.

The native form of P46 has an apparent molecular mass of approximately 180 kDa, as estimated by gel filtration chromatography. This represent about four times the value of 46 kDa determined by SDS-PAGE and five to six times the theoretical molecular weight of P46 (36,559). Hence, the native form of P46 should be homomultimeric. The multimeric form of P46 is seen as that existing out of the ribosome (extraribosomal form) but not within the ribosome. Indeed, the only multi-copy ribosomal protein found in eubacterial ribosomes is the L7/12 ribosomal protein with four copies per ribosome (43, 44). The extraribosomal form of L7/L12 is a stable dimer (45–47). We have no experimental proof that P46 binds the IRS as a multimer; as in our Southwestern analyses, samples containing P46 were submitted to an SDS denaturing step before electroblotting. However, it is known that renaturation to the native conformation during electrotransfer of the protein from the gel to the cellulose membrane does occur (27).

The apparent molecular mass of P46 observed in SDS-PAGE (46 kDa) is larger than that calculated from the amino acid sequence (36,559 Da). Abnormal electrophoretic migration in SDS-PAGE has been described for other ribosomal proteins for which the apparent molecular mass in SDS gels is up to 30% higher than that calculated from the primary sequence (48). The molecular mass of P46 (apparent or calculated) is larger than that of all other eubacterial L29 proteins. P46 could be divided into three domains. The N-terminal domain of P46 (domain I) showed significant similarity with the L29 ribosomal protein family (60% identity with the N-terminal domain of *M. capricolum* L29 ribosomal protein) (Fig. 7, A and B). The internal domain of P46 (domain II) matched the C-terminal domain of *M. capricolum* L29 ribosomal protein (Fig. 7, A and C). The C-terminal domain of P46 (domain III) (Fig. 7A) shared significant similarities with the histone H1-like proteins found in some bacterial species such as Hc2 from *C. trachomatis* (41) and AlgP (*algR3*) (38, 39) from *P. aeruginosa*. The *C. trachomatis* Hc2 histone H1-like protein, initially identified by Southwestern blotting of chlamydial lysates (49), has been implicated in DNA binding, nucleoid compaction, and *in vitro* transcription/translation repression (41, 50, 51). The AlgP (*algR3*) histone H1-like protein is a DNA-binding protein involved in the transcriptional activation of *algD*, a necessary step for the establishment of mucoidy in *P. aeruginosa* (38, 39, 52). Repeated tetrapeptides like KPAA and variants are found in eukaryotic H1 histones and AlgP. Such repeated sequences appear to be crucial for DNA binding by AlgP (38). Repeats of such KPAA motifs and variants are also found, to a lower extent, within the C-terminal domain of P46. With these observations, it is tempting to associate the third domain of P46 with the IRS-binding property of the protein. We have made some preliminary experiments in order to confirm this hypothesis. The presence of a unique asparaginyl-glycyl peptide bond between position 138 and 139 of P46, *i.e.* between domain II and III (Fig. 7A), should allow cleavage of P46 at this position with hydroxylamine (53–57). This cleavage should generate a 138-amino acid peptide (domains I and II) and a 201-amino acid peptide (domain III). Keeping in mind the altered electrophoretic mobility of the P46, the 201-amino acid peptide should behave on SDS-PAGE as a polypeptide of apparent molecular mass of 28 kDa. Southwestern analysis of peptides generated by the action of hydroxylamine on P46 has indeed revealed only one IRS-binding peptide of approximately 29 kDa.<sup>3</sup> Altogether,

<sup>3</sup> L. Le Dantec, C. Saillard, and J. M. Bové, unpublished results.

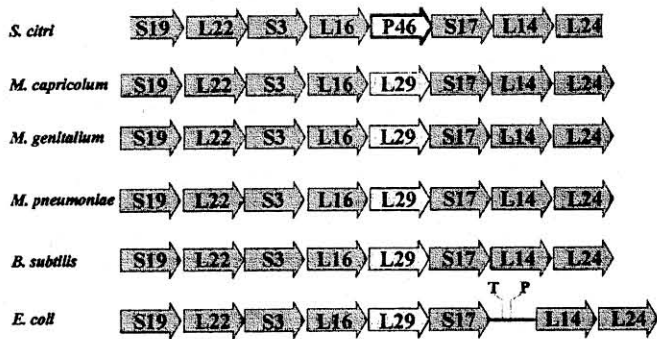


FIG. 6. Comparison of the organization of the *S. citri* ribosomal protein gene cluster with the S10 and *spc* operons of *E. coli*, *B. subtilis*, and several mollicutes. Transcription initiation (P) and termination (T) signals are indicated. Drawing is not to scale.

*Chlamydia trachomatis* histone H1 homolog Hc2 (41). Histone H1-like proteins exhibit sequence similarities to eucaryotic histones H1 and are involved in bacterial nucleoid organization and/or transcriptional regulation through their interactions with bacterial DNA. These observations suggested that the C-terminal domain (domain III, Fig. 7A) of P46 is implicated in DNA-protein interaction.

**Localization of P46**—The above results suggested that P46 is the L29 ribosomal protein of *S. citri*. To confirm this finding, we have analyzed *S. citri* ribosomal proteins. *S. citri* ribosomes were partially purified by differential ultracentrifugation (25, 26). Proteins from the supernatant (S-105) and the washed ribosomes were resolved on SDS-PAGE and examined by Southwestern analysis (Fig. 8). A 46-kDa IRS-binding protein was detected in both the S-105 supernatant (Fig. 8, lane 1) and in the washed ribosome fraction but in much larger quantities in the latter (Fig. 8, lane 2). These results indicated that P46 is present in the ribosomal fraction. Three other IRS-binding protein species of approximately 25, 26, and 38 kDa were detected in the ribosome fraction (Fig. 8, lane 2). These proteins are probably the DNA-binding proteins detected in crude extracts of spiroplasmal cells and in the eluted fractions of the heparin-agarose affinity column (Fig. 2C). The binding affinity of these proteins has not been investigated.

#### DISCUSSION

Despite the large amount of work on the regulation of ribosome synthesis (see Ref. 42 for a review), molecular mechanisms of regulation of some ribosomal protein operons have not been elucidated completely. This is the case of the *E. coli* *rpsB-tsfc* operon for which an attenuation mechanism was proposed but has not been demonstrated (3). In *S. citri*, *rpsB*, *tsfc*, and  $\alpha$  represent the three genes of a single transcriptional unit, and a potential regulatory sequence (a 20-bp inverted repeat sequence) was found at the 3' end of the *rpsB* gene (70). In the study reported in this paper we have detected, by Southwestern blot analysis, a protein (P46) that binds with preferential affinity to a 30-bp oligonucleotide (IRS) containing the 20-bp inverted repeat sequence. We have partially purified the P46 protein, cloned, and sequenced its gene.

The deduced amino acid sequence from the cloned gene indicated that the N-terminal end of P46 shared significant similarity with the L29 ribosomal protein family. In addition, sequence analysis of the whole cloned DNA insert carrying the P46 gene indicated that the P46 gene is located within the S10-*spc* operon, and more precisely between the *rpsL* and *rpsQ* genes coding, respectively, for ribosomal proteins L16 and S17 (Fig. 6). Thus, the P46 gene is located at the position of the L29 gene in the previously described S10-*spc* operons and represents a ribosomal protein gene. Indeed, Southwestern analysis

of *S. citri* ribosomal proteins revealed the presence of the 46-kDa IRS-binding protein (Fig. 8). These findings suggested that P46 is the L29 ribosomal protein of *S. citri*.

The native form of P46 has an apparent molecular mass of approximately 180 kDa, as estimated by gel filtration chromatography. This represents about four times the value of 46 kDa determined by SDS-PAGE and five to six times the theoretical molecular weight of P46 (36,559). Hence, the native form of P46 should be homomultimeric. The multimeric form of P46 is seen as that existing out of the ribosome (extraribosomal form) but not within the ribosome. Indeed, the only multi-copy ribosomal protein found in eubacterial ribosomes is the L7/L12 ribosomal protein with four copies per ribosome (43, 44). The extraribosomal form of L7/L12 is a stable dimer (45–47). We have no experimental proof that P46 binds the IRS as a multimer; as in our Southwestern analyses, samples containing P46 were submitted to an SDS denaturing step before electroblotting. However, it is known that renaturation to the native conformation during electrotransfer of the protein from the gel to the cellulose membrane does occur (27).

The apparent molecular mass of P46 observed in SDS-PAGE (46 kDa) is larger than that calculated from the amino acid sequence (36,559 Da). Abnormal electrophoretic migration in SDS-PAGE has been described for other ribosomal proteins for which the apparent molecular mass in SDS gels is up to 30% higher than that calculated from the primary sequence (48). The molecular mass of P46 (apparent or calculated) is larger than that of all other eubacterial L29 proteins. P46 could be divided into three domains. The N-terminal domain of P46 (domain I) showed significant similarity with the L29 ribosomal protein family (60% identity with the N-terminal domain of *M. capricolum* L29 ribosomal protein) (Fig. 7, A and B). The internal domain of P46 (domain II) matched the C-terminal domain of *M. capricolum* L29 ribosomal protein (Fig. 7, A and C). The C-terminal domain of P46 (domain III) (Fig. 7A) shared significant similarities with the histone H1-like proteins found in some bacterial species such as Hc2 from *C. trachomatis* (41) and AlgP (*algR3*) (38, 39) from *P. aeruginosa*. The *C. trachomatis* Hc2 histone H1-like protein, initially identified by Southwestern blotting of chlamydial lysates (49), has been implicated in DNA binding, nucleoid compaction, and *in vitro* transcription/translation repression (41, 50, 51). The AlgP (*algR3*) histone H1-like protein is a DNA-binding protein involved in the transcriptional activation of *algD*, a necessary step for the establishment of mucoidy in *P. aeruginosa* (38, 39, 52). Repeated tetrapeptides like KPAA and variants are found in eukaryotic H1 histones and AlgP. Such repeated sequences appear to be crucial for DNA binding by AlgP (38). Repeats of such KPAA motifs and variants are also found, to a lower extent, within the C-terminal domain of P46. With these observations, it is tempting to associate the third domain of P46 with the IRS-binding property of the protein. We have made some preliminary experiments in order to confirm this hypothesis. The presence of a unique asparaginyl-glycyl peptide bond between position 138 and 139 of P46, *i.e.* between domain II and III (Fig. 7A), should allow cleavage of P46 at this position with hydroxylamine (53–57). This cleavage should generate a 138-amino acid peptide (domains I and II) and a 201-amino acid peptide (domain III). Keeping in mind the altered electrophoretic mobility of the P46, the 201-amino acid peptide should behave on SDS-PAGE as a polypeptide of apparent molecular mass of 28 kDa. Southwestern analysis of peptides generated by the action of hydroxylamine on P46 has indeed revealed only one IRS-binding peptide of approximately 29 kDa.<sup>3</sup> Altogether,

<sup>3</sup> L. Le Dantec, C. Saillard, and J. M. Bové, unpublished results.

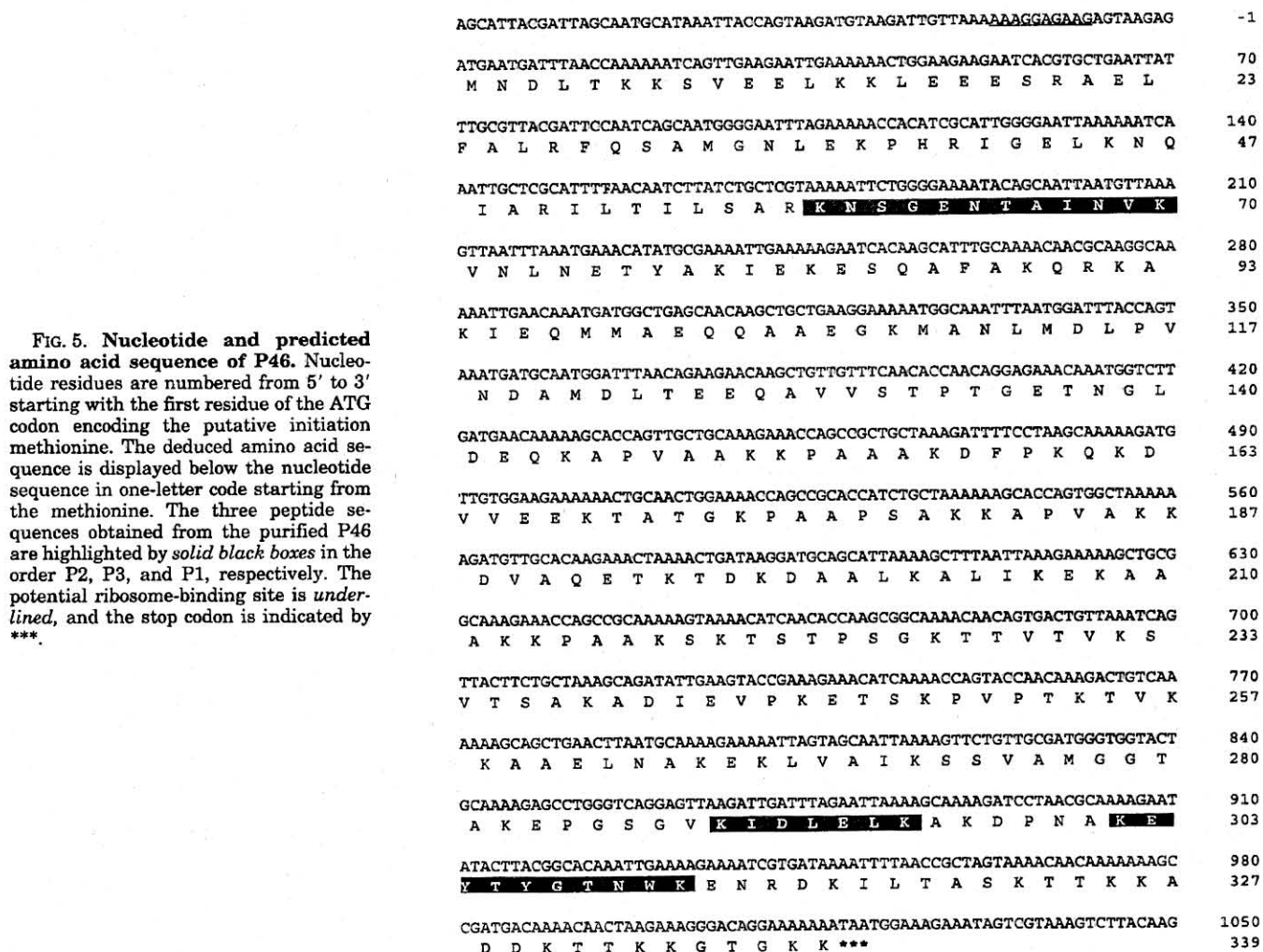


Fig. 5. Nucleotide and predicted amino acid sequence of P46. Nucleotide residues are numbered from 5' to 3' starting with the first residue of the ATG codon encoding the putative initiation methionine. The deduced amino acid sequence is displayed below the nucleotide sequence in one-letter code starting from the methionine. The three peptide sequences obtained from the purified P46 are highlighted by solid black boxes in the order P2, P3, and P1, respectively. The potential ribosome-binding site is underlined, and the stop codon is indicated by \*\*\*.

TABLE I  
 ORFs and putative proteins of the 3.5-kbp *S. citri* DNA insert, containing the P46 gene (ORF V) and homologous genes and proteins from eubacteria

ORF	Location on Spiroplasmal DNA		Size of putative Spiroplasmal protein (no. amino acids)	Homologous protein (origin)	Size (no. amino acids)	Identity	Corresponding gene	Accession number
	First nucleotide	Last nucleotide						
I	1 <sup>a</sup>	58	>19	Ribosomal protein S19 ( <i>M. capricolum</i> )	88	89.5	<i>rpsS</i>	p10132
II	90	426	112	Ribosomal protein L22 ( <i>M. capricolum</i> )	111	62.2	<i>rpIV</i>	p10139
III	448	1204	252	Ribosomal protein S3 ( <i>M. capricolum</i> )	232	62	<i>rpsC</i>	p02353
IV	1206	1617	137	Ribosomal protein L16 ( <i>M. capricolum</i> )	137	78.1	<i>rplP</i>	p02415
V	1623	2640	339	Ribosomal protein L29 ( <i>M. capricolum</i> )	138	40.6	<i>rpmC</i>	p10142
VI	2642	2897	85	Ribosomal protein S17 ( <i>M. capricolum</i> )	85	75.3	<i>rpsQ</i>	p10131
VII	2921	3287	122	Ribosomal protein L14 ( <i>M. capricolum</i> )	122	74.6	<i>rplN</i>	p10137
VIII	3301	3475 <sup>a</sup>	>58	Ribosomal protein L24 ( <i>M. capricolum</i> )	108	51.7	<i>rplX</i>	p10141

<sup>a</sup> The genes are truncated.

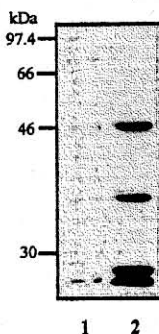
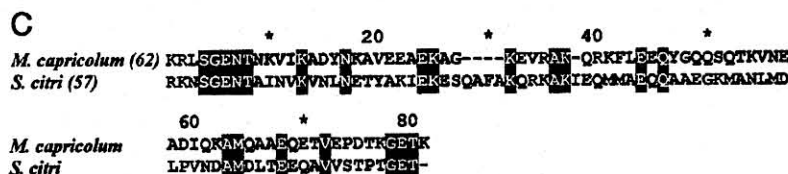
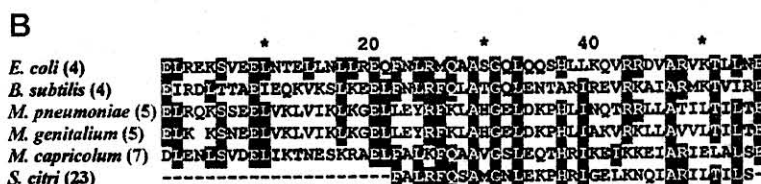
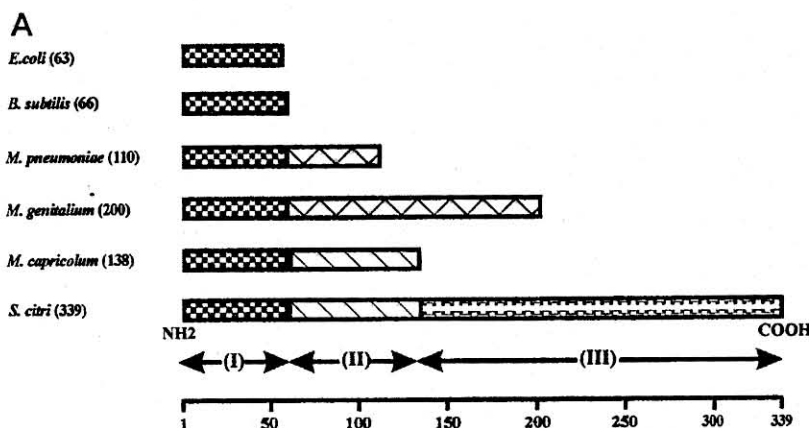
between the *rplP* and *rpsQ* genes that code, respectively, for ribosomal proteins L16 and S17. As shown in Fig. 6, comparison with similar operons in *E. coli*, *B. subtilis*, and several mollicutes revealed that P46 gene is at the position of ribosomal protein L29.

Only the N-terminal sequence of P46 shares similarities with the sequences of L29 ribosomal proteins in protein data bases. Domain analysis of P46 using the Prodom protein domain families data base (22, 23) revealed a unique feature of this protein among other known L29 proteins. P46, which is much larger than other L29 proteins, could be arranged into three domains (Fig. 7A). The first domain (domain I, from position 1-57 of the P46 sequence) matches the Prodom domain 1463

(Prodom 34.1) which is a common domain of the prokaryotic L29 family (Fig. 7B). The second domain (domain II, from position 58-137 of P46) matches the Prodom domain 24734 (Prodom 34.1) which is the C-terminal sequence of the *M. capricolum* L29 ribosomal protein (Fig. 7C). The third domain, ranging from position 138-339 on the P46 sequence shares no significant similarities with any Prodom domains.

However, BlastP analysis of this domain revealed significant similarities with DNA-binding histone H1-like proteins. It exhibits 28.6% amino acid identity to *Pseudomonas aeruginosa* transcriptional regulatory protein AlgP (38) also termed *algr3* (39), 31.3% amino acid identity to *Bordetella pertussis* histone H1 homolog BpH1 (40), and 26.3% amino acid identity to

**FIG. 7. Comparison of P46 with several eubacterial ribosomal proteins L29.** A, schematic representation of primary structure and potential domain organization of P46 in comparison with those of several eubacterial ribosomal protein L29s. Numbers in parentheses after the names of eubacterial species refer to the length, in amino acid residues, of the protein. Scale at the bottom is in amino acids. I, II, and III represent the potential N-terminal, internal, and C-terminal domains of P46, respectively. Boxes with similar internal motifs represent the same domain in the Prodom protein domain families data base. B, alignment of amino acid sequence of domain I of P46 with that of five L29 ribosomal proteins from several eubacterial species. Numbers in parentheses after the names of eubacterial species refer to the number of residues preceding the first residue shown in the alignment. Conserved residues that are present in all or most of these proteins are indicated by the black areas. Hyphens represent gaps introduced for maximal alignment. C, alignment of amino acid sequence of domain II of P46 with that of *M. capricolum* ribosomal protein L29.



**FIG. 8. Southwestern analysis of *S. citri* ribosomal proteins.** The *S. citri* ribosome purification procedure is described under "Experimental Procedures." 50- $\mu$ l aliquots of the supernatant (S-105) fraction (75  $\mu$ g) (lane 1) and washed ribosome fraction (35  $\mu$ g) (lane 2) were resolved on SDS-PAGE and assayed for IRS binding activity using Southwestern analysis.

these data support the role of the C-terminal domain of P46 in DNA-protein interaction.

Thus P46 could be a bifunctional protein. The N-terminal domain has a ribosomal function as it has high homology with the L29 ribosomal protein family. The C-terminal domain is seen as the one involved in IRS binding and might have regulatory function at the IRS of the genomic DNA. The bifunctional nature and the presence of DNA binding motifs in some ribosomal proteins has been reported and has led to speculations on the origin of the ribosomal proteins (58–64). It is interesting to note that in *B. subtilis* the HPB12-L24 protein has been described as a bifunctional ribosomal protein (L24) with histone-like properties and DNA binding activity (63, 65, 66).

Mollicutes are the smallest and simplest self-replicating organisms, and the currently dominating hypothesis is that they have evolved by degenerative (regressive) evolution from

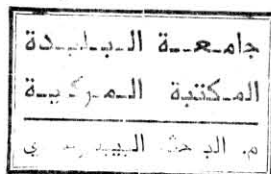
Gram-positive bacteria with low guanine + cytosine genomes (1, 67, 68). In term of evolution strategy, mollicutes may have concentrated two different functions in a single gene during genome size reduction. Evidence for a single gene affording dual enzymatic function (malate/lactate dehydrogenase) has been described in *M. genitalium* by Cordwell and co-workers (69). Similarly, analysis of the complete sequence of the *M. genitalium* genome led Fraser and co-workers (6) to state that some *M. genitalium* proteins may have become adapted to perform more than one function. P46 has most likely a ribosomal protein function by its N-terminal L29 domain and a putative regulatory function on the IRS of the *rpsB/tsf/x* operon by its C-terminal domain. S2 and Ef-Ts are components of the translational machinery, and P46 may play an interconnecting role in the coordinated regulation of the components of the translational apparatus. However, the mechanism by which the binding of P46 on the IRS could influence the transcription of the *rpsB/tsf/x* operon is still unknown. *In vitro* studies of the transcriptional regulation of the *rpsB/tsf/x* operon with and without the presence of the P46 protein may help understand the extraribosomal function of P46.

REFERENCES

1. Weisburg, W. G., Tully, J. G., Rose, D. L., Petzel, J. P., Oyaizu, H., Yang, D., Mandelco, L., Sechrest, J., Lawrence, T. G., Van Etten, J., Maniloff, J., and Woese, C. R. (1989) *J. Bacteriol.* 171, 6455–6467
2. Saglio, P., Lhospital, M., Lafleche, D., Dupont, G., Bové, J. M., Tully, J. G., Freundt, E. A. (1973) *Int. J. Syst. Bacteriol.* 23, 191–204
3. An, G., Bendiaak, D. S., Mamelak, L. A., and Friesen, J. D. (1981) *Nucleic Acids Res.* 9, 4163–4172
4. Bendiaak, D. S., and Friesen, J. D. (1981) *Mol. Gen. Genet.* 181, 356–362
5. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessières, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S.-K., Codani, J.-J., Connerton, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Dusterhöft, A., Ehrlich, S. D., Emmerson, P. T., Entian, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.-Y., Glaser, P., Goffeau, A., Golightly,

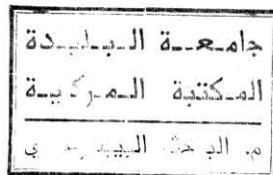
- E. J., Grandi, G., Guiseppi, G., Guy, B. J., Haga, K., Haiech, J., Harwood, C. R., Hénaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.-F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.-M., Levine, A., Liu, H., Masuda, S., Mauel, C., Médigue, C., Medina, N., Meilado, R. P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.-H., Parro, V., Pohl, T. M., Portetelle, D., Porwollik, S., Prescott, A. M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S. J., Serror, P., Shin, B.-S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzenegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.-F., Zumstein, E., Yoshikawa, H., and Danchin, A. (1997) *Nature* **390**, 249-256
6. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Merrick, J. M., Tomb, J. F., Dougherty, B. A., Bott, K. F., Hu, P. C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchinson, C. A., III, and Venter, J. C. (1995) *Science* **270**, 397-403
7. Himmelreich, R., Hilbert, H., Plagens, H., Pirkel, E., Li, B. C., and Herrmann, R. (1996) *Nucleic Acids Res.* **24**, 4420-4449
8. Pabo, C. O., and Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53**, 293-321
9. Harrison, S. C., and Aggarwal, A. K. (1990) *Annu. Rev. Biochem.* **59**, 933-969
10. Tully, J. G., Whitcomb, R. F., Clark, H. F., and Williamson, D. L. (1977) *Science* **195**, 892-894
11. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-254
12. Bowen, B., Steinberg, J., Laemmli, U. K., and Weintraub, H. (1980) *Nucleic Acids Res.* **8**, 1-21
13. Laemmli, U. K. (1970) *Nature* **227**, 680-685
14. Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4350-4354
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Wilson, C. (1979) *Anal. Biochem.* **96**, 263-278
17. Bové, J. M. (1993) *Clin. Infect. Dis.* **17**, 10-31
18. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503-517
19. Devereux, J., Haeblerli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395
20. Staden, R., and McLachlan, A. D. (1982) *Nucleic Acids Res.* **10**, 141-156
21. Altshul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403-410
22. Sonnhammer, E. L. L., and Kahn, D. (1994) *Protein Sci.* **3**, 482-492
23. Gouzy, J., Corpet, F., and Kahn, D. (1996) *Trends Biochem. Sci.* **21**, 493
24. Corpet, F. (1988) *Nucleic Acids Res.* **16**, 10881-10890
25. Tissières, A., Watson, J. D., Schlessinger, D., and Hollingworth, B. R. (1959) *J. Mol. Biol.* **1**, 221-233
26. Matthaei, J. H., and Nirenberg, M. W. (1961) *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1580-1588
27. Miskimins, W. K., Roberts, M. P., McClelland, A., and Ruddle, F. H. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 6741-6744
28. Michael, N., Spector, D., Mavromara-Nazos, P., Kristie, T. M., and Roizman, B. (1988) *Science* **239**, 1531-1534
29. Post, L. E., Arfsten, F. R., and Nomura, M. (1978) *Cell* **15**, 215-229
30. Cerretti, D. P., Dean, D., Davis, G. R., Bedwell, D. M., and Nomura, M. (1983) *Nucleic Acids Res.* **11**, 2599-2616
31. Zurawski, G., and Zurawski, S. M. (1985) *Nucleic Acids Res.* **13**, 4521-4526
32. Henkin, T. M., Moon, S. H., Mattheakis, L. C., and Nomura, M. (1989) *Nucleic Acids Res.* **17**, 7469-7486
33. Yoshikawa, H., and Doi, R. H., (1990) *Nucleic Acids Res.* **18**, 1647
34. Li, X., Lindahl, L., Sha, Y., and Zengel, J. M. (1997) *J. Bacteriol.* **179**, 7046-7054
35. Ohkubo, S., Muto, A., Kawauchi, Y., Yamao, F., and Osawa, S. (1987) *Mol. Gen. Genet.* **210**, 314-322
36. Hilbert, H., Himmelreich, R., Plagens, H., and Herrmann, R. (1996) *Nucleic Acids Res.* **24**, 628-639
37. Lindahl, L., Sor, F., Archer, R. H., Nomura, M., and Zengel, J. M. (1990) *Biochim. Biophys. Acta* **1050**, 337-342
38. Deretic, V., and Konyecsni, W. M. (1990) *J. Bacteriol.* **172**, 5544-5554
39. Kato, J., Misra, T. K., and Chakrabarty, A. M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2887-2891
40. Scariato, V., Arico, B., Goyard, S., Ricci, S., Manetti, R., Prugnola, A., Manetti, R., Polverino-De-Laureto, P., Ullmann, A., and Rappuoli, R. (1995) *Mol. Microbiol.* **15**, 871-881
41. Brickman, T. J., Barry, C. E., III, and Hackstadt, T. (1993) *J. Bacteriol.* **175**, 4274-4281
42. Zengel, J. M., and Lindahl, L. (1994) *Prog. Nucleic Acids Res. Mol. Biol.* **47**, 331-370
43. Hardy, S. J. S. (1976) *Mol. Gen. Genet.* **140**, 253-279
44. Subramanian, A. R. (1975) *J. Mol. Biol.* **95**, 1-8
45. Möller, W., Groene, A., Terhorst, C., and Amons, R. (1972) *Eur. J. Biochem.* **25**, 5-12
46. Gudkov, A. T., Behlke, J., Vtiurin, N. N., and Lim, V. I. (1977) *FEBS Lett.* **82**, 125-129
47. Gudkov, A. T., and Behlke, J. (1978) *Eur. J. Biochem.* **90**, 309-312
48. Pedersen, S., Bloch, P. L., Reeh, S., and Neidhardt, F. C. (1978) *Cell* **14**, 179-190
49. Wagar, E. A., and Stevens, R. S. (1988) *Infect. Immun.* **56**, 1678-1684
50. Perara, E., Ganem, D., and Engel, J. N. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2125-2129
51. Pedersen, L. B., Birkelund, S., and Christiansen, G. (1996) *Mol. Microbiol.* **20**, 295-311
52. Konyecsni, W. M., and Deretic, V. (1990) *J. Bacteriol.* **172**, 2511-2520
53. Bornstein, P., and Balian, G. (1970) *J. Biol. Chem.* **245**, 4854-4856
54. Bornstein, P., and Balian, G. (1977) *Methods Enzymol.* **114**, 132-144
55. Saris, C. J. M., van Eenbergen, J., Jenks, B. G., and Bloemers, H. P. J. (1983) *Anal. Biochem.* **132**, 54-67
56. Changchien, L.-M., and Craven G. C. (1986) *Nucleic Acids Res.* **14**, 1957-1966
57. Myers, D. A., and Vilemex C. L. (1988) *J. Biol. Chem.* **263**, 17122-17127
58. Wool, I. G. (1993) in *The Translational Apparatus* (Nierhaus, K. H., Franceschi, F., Subramanian, A. R., Erdmann, V. A., and Wittmann-Liebold, B., eds) pp. 727-737, Plenum Publishing Corp., New York
59. Gryaznova, O. I., Davydova, N. L., Gongadze, G. M., Jonsson, E. H., Garber, M. B., and Liljas, A. (1996) *Biochimie (Paris)* **78**, 915-919
60. Oberto, J., Bonnefoy, E., Mouray, E., Pellegrini, O., Wilkström, P. M., and Rouvière-Yaniv, J. (1996) *Mol. Microbiol.* **19**, 1319-1330
61. Wool, I. G. (1996) in *Ribosomal RNA and Group I Introns* (Green, R., and Shroeder, R., eds) pp. 153-178, R. G. Landes Co., Austin, TX
62. Wool, I. G. (1996) *Trends Biochem. Sci.* **21**, 164-165
63. Zouine, M., Beloin, C., Deneubourg, A.-M., Hirschbein, L., and Le Hégarat F. (1996) *FEMS Microbiol. Lett.* **145**, 41-48
64. Büttner, K., Pitch, A., Neubauer, P., Schmid, R., Bahl, H., and Hecker, M. (1997) *J. Basic Microbiol.* **37**, 3-9
65. Salti, V., Le Hégarat, F., and Hirschbein, L. (1989) *Biochim. Biophys. Acta* **1009**, 161-167
66. Arnold-Schultz-Gahmen, B., Salti Montesanto, V., Nguyen, J., Hirschbein, L., and Le Hégarat, F. (1994) *J. Bacteriol.* **176**, 50-60
67. Woese, C. R., Maniloff, J., and Zablén, L. B. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 494-498
68. Rogers, M. J., Simmons, J., Walker, R. T., Weisburg, W. G., Woese, C. R., Tanner, R. S., Robinson, I. M., Stahl, D. A., Olsen, G., Leach, R. H., and Maniloff, J. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 1160-1164
69. Cordwell, S. J., Basseal, D. J., Pollack, J. D., and Humphery-Smith, I. (1997) *Gene (Amst.)* **195**, 113-120
70. Le Dantec, L., Bové, J. M., and Saillard C. (1998) *Curr. Microbiol.*, in press

N° 44/04  
 Agno.

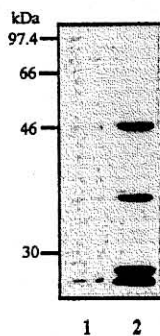
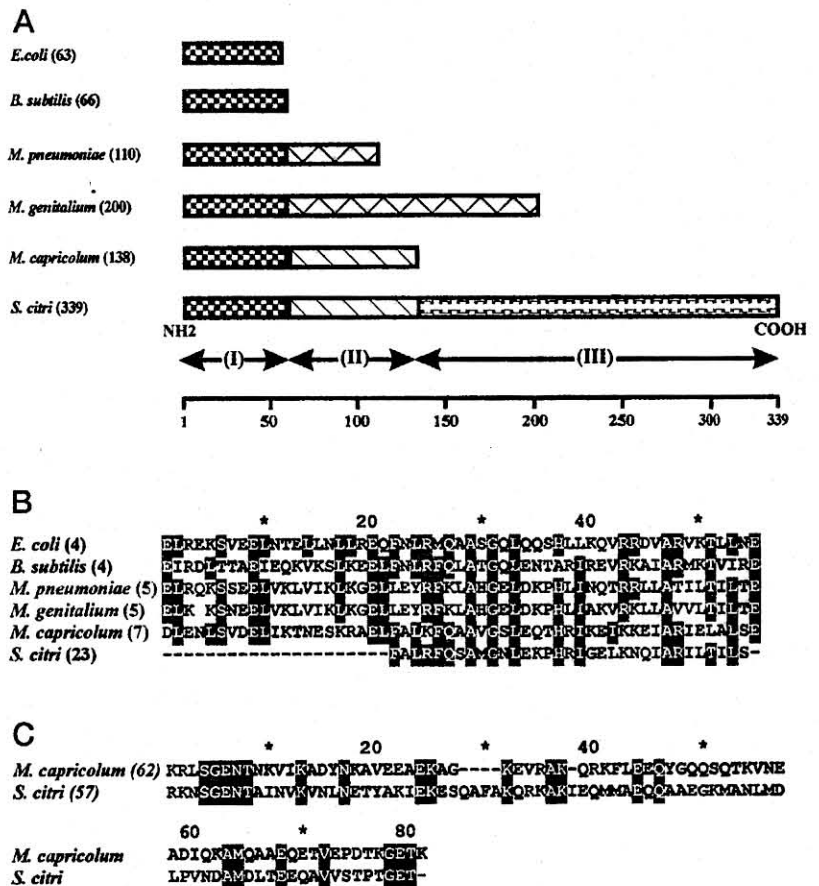


- E. J., Grandi, G., Guiseppi, G., Guy, B. J., Haga, K., Haiech, J., Harwood, C. R., Hénaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.-F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerr-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningsstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.-M., Levine, A., Liu, H., Masuda, S., Mauël, C., Médigue, C., Medina, N., Meilado, R. P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.-H., Parro, V., Pohl, T. M., Portetelle, D., Porwollik, S., Prescott, A. M., Presecan, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S. J., Serror, P., Shin, B.-S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Tamakoshi, A., Tanaka, T., Terpstra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarrotti, A., Viari, A., Wambutt, R., Wedler, H., Wedler, H., Weitznegger, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.-F., Zumstein, E., Yoshikawa, H., and Danchin, A. (1997) *Nature* **390**, 249-256
6. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Merrick, J. M., Tomb, J. F., Dougherty, B. A., Bott, K. F., Hu, P. C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchinson, C. A., III, and Venter, J. C. (1995) *Science* **270**, 397-403
7. Himmelreich, R., Hilbert, H., Plagens, H., Pirkel, E., Li, B. C., and Herrmann, R. (1996) *Nucleic Acids Res.* **24**, 4420-4449
8. Pabo, C. O., and Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53**, 293-321
9. Harrison, S. C., and Aggarwal, A. K. (1990) *Annu. Rev. Biochem.* **59**, 933-969
10. Tully, J. G., Whitcomb, R. F., Clark, H. F., and Williamson, D. L. (1977) *Science* **195**, 892-894
11. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248-254
12. Bowen, B., Steinberg, J., Laemmli, U. K., and Weintraub, H. (1980) *Nucleic Acids Res.* **8**, 1-21
13. Laemmli, U. K. (1970) *Nature* **227**, 680-685
14. Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4350-4354
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Wilson, C. (1979) *Anal. Biochem.* **96**, 263-278
17. Bové, J. M. (1993) *Clin. Infect. Dis.* **17**, 10-31
18. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503-517
19. Devereux, J., Haeblerli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395
20. Staden, R., and McLachlan, A. D. (1982) *Nucleic Acids Res.* **10**, 141-156
21. Altshul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403-410
22. Sonnhammer, E. L. L., and Kahn, D. (1994) *Protein Sci.* **3**, 482-492
23. Gouzy, J., Corpet, F., and Kahn, D. (1996) *Trends Biochem. Sci.* **21**, 493
24. Corpet, F. (1988) *Nucleic Acids Res.* **16**, 10881-10890
25. Tissières, A., Watson, J. D., Schlessinger, D., and Hollingworth, B. R. (1959) *J. Mol. Biol.* **1**, 221-233
26. Matthaei, J. H., and Nirenberg, M. W. (1961) *Proc. Natl. Acad. Sci. U. S. A.* **47**, 1580-1588
27. Miskimins, W. K., Roberts, M. P., McClelland, A., and Ruddle, F. H. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 6741-6744
28. Michael, N., Spector, D., Mavromara-Nazos, P., Kristie, T. M., and Roizman, B. (1988) *Science* **239**, 1531-1534
29. Post, L. E., Arfsten, F. R., and Nomura, M. (1978) *Cell* **15**, 215-229
30. Cerretti, D. P., Dean, D., Davis, G. R., Bedwell, D. M., and Nomura, M. (1983) *Nucleic Acids Res.* **11**, 2599-2616
31. Zurawski, G., and Zurawski, S. M. (1985) *Nucleic Acids Res.* **13**, 4521-4526
32. Henkin, T. M., Moon, S. H., Mattheakis, L. C., and Nomura, M. (1989) *Nucleic Acids Res.* **17**, 7469-7486
33. Yoshikawa, H., and Doi, R. H., (1990) *Nucleic Acids Res.* **18**, 1647
34. Li, X., Lindahl, L., Sha, Y., and Zengel, J. M. (1997) *J. Bacteriol.* **179**, 7046-7054
35. Ohkubo, S., Muto, A., Kawauchi, Y., Yamao, F., and Osawa, S. (1987) *Mol. Gen. Genet.* **210**, 314-322
36. Hilbert, H., Himmelreich, R., Plagens, H., and Herrmann, R. (1996) *Nucleic Acids Res.* **24**, 628-639
37. Lindahl, L., Sor, F., Archer, R. H., Nomura, M., and Zengel, J. M. (1990) *Biochim. Biophys. Acta* **1050**, 337-342
38. Deretic, V., and Konyecsni, W. M. (1990) *J. Bacteriol.* **172**, 5544-5554
39. Kato, J., Misra, T. K., and Chakrabarty, A. M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2887-2891
40. Scarlato, V., Arico, B., Goyard, S., Ricci, S., Manetti, R., Prugnola, A., Manetti, R., Polverino-De-Lauroto, P., Ullmann, A., and Rappuoli, R. (1995) *Mol. Microbiol.* **15**, 871-881
41. Brickman, T. J., Barry, C. E., III, and Hackstadt, T. (1993) *J. Bacteriol.* **175**, 4274-4281
42. Zengel, J. M., and Lindahl, L. (1994) *Prog. Nucleic Acids Res. Mol. Biol.* **47**, 331-370
43. Hardy, S. J. S. (1975) *Mol. Gen. Genet.* **140**, 253-279
44. Subramanian, A. R. (1975) *J. Mol. Biol.* **95**, 1-8
45. Möller, W., Groene, A., Terhorst, C., and Amons, R. (1972) *Eur. J. Biochem.* **25**, 5-12
46. Gudkov, A. T., Behlke, J., Vtiurin, N. N., and Lim, V. I. (1977) *FEBS Lett.* **82**, 125-129
47. Gudkov, A. T., and Behlke, J. (1978) *Eur. J. Biochem.* **90**, 309-312
48. Pedersen, S., Bloch, P. L., Reeh, S., and Neidhardt, F. C. (1978) *Cell* **14**, 179-190
49. Wagar, E. A., and Stevens, R. S. (1988) *Infect. Immun.* **56**, 1678-1684
50. Perara, E., Ganem, D., and Engel, J. N. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2125-2129
51. Pedersen, L. B., Birkelund, S., and Christiansen, G. (1996) *Mol. Microbiol.* **20**, 295-311
52. Konyecsni, W. M., and Deretic, V. (1990) *J. Bacteriol.* **172**, 2511-2520
53. Bornstein, P., and Balian, G. (1970) *J. Biol. Chem.* **245**, 4854-4856
54. Bornstein, P., and Balian, G. (1977) *Methods Enzymol.* **114**, 132-144
55. Saris, C. J. M., van Eenbergen, J., Jenks, B. G., and Bloemers, H. P. J. (1983) *Anal. Biochem.* **132**, 54-67
56. Changchien, L.-M., and Craven G. C. (1986) *Nucleic Acids Res.* **14**, 1957-1966
57. Myers, D. A., and Vilemez C. L. (1988) *J. Biol. Chem.* **263**, 17122-17127
58. Wool, I. G. (1993) in *The Translational Apparatus* (Nierhaus, K. H., Franceschi, F., Subramanian, A. R., Erdmann, V. A., and Wittmann-Liebold, B., eds) pp. 727-737, Plenum Publishing Corp., New York
59. Gryaznova, O. I., Davydova, N. L., Gongadze, G. M., Jonsson, B. H., Garber, M. B., and Liljas, A. (1996) *Biochimie (Paris)* **78**, 915-919
60. Oberto, J., Bonnefoy, E., Mouray, E., Pellegrini, O., Wilkström, P. M., and Rouviere-Yaniv, J. (1996) *Mol. Microbiol.* **19**, 1319-1330
61. Wool, I. G. (1996) in *Ribosomal RNA and Group I Introns* (Green, R., and Schroeder, R., eds) pp. 153-178, R. G. Landes Co., Austin, TX
62. Wool, I. G. (1996) *Trends Biochem. Sci.* **21**, 164-165
63. Zouine, M., Beloin, C., Deneubourg, A.-M., Hirschbein, L., and Le Hégarat F. (1996) *FEMS Microbiol. Lett.* **145**, 41-48
64. Büttner, K., Pitch, A., Neubauer, P., Schmid, R., Bahl, H., and Hecker, M. (1997) *J. Basic Microbiol.* **37**, 3-9
65. Salti, V., Le Hégarat, F., and Hirschbein, L. (1989) *Biochim. Biophys. Acta* **1009**, 161-167
66. Arnold-Schultz-Gahmen, B., Salti Montesanto, V., Nguyen, J., Hirschbein, L., and Le Hégarat, F. (1994) *J. Bacteriol.* **176**, 50-60
67. Woese, C. R., Maniloff, J., and Zablen, L. B. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 494-498
68. Rogers, M. J., Simmons, J., Walker, R. T., Weisburg, W. G., Woese, C. R., Tanner, R. S., Robinson, I. M., Stahl, D. A., Olsen, G., Leach, R. H., and Maniloff, J. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 1160-1164
69. Cordwell, S. J., Basseal, D. J., Pollack, J. D., and Humphery-Smith, I. (1997) *Gene (Amst.)* **195**, 113-120
70. Le Dantec, L., Bové, J. M., and Saillard C. (1998) *Curr. Microbiol.*, in press

N° 44/04  
 Agno.



**FIG. 7. Comparison of P46 with several eubacterial ribosomal proteins L29.** A, schematic representation of primary structure and potential domain organization of P46 in comparison with those of several eubacterial ribosomal protein L29s. Numbers in parentheses after the names of eubacterial species refer to the length, in amino acid residues, of the protein. Scale at the bottom is in amino acids. I, II, and III represent the potential N-terminal, internal, and C-terminal domains of P46, respectively. Boxes with similar internal motifs represent the same domain in the Prodom protein domain families data base. B, alignment of amino acid sequence of domain I of P46 with that of five L29 ribosomal proteins from several eubacterial species. Numbers in parentheses after the names of eubacterial species refer to the number of residues preceding the first residue shown in the alignment. Conserved residues that are present in all or most of these proteins are indicated by the black areas. Hyphens represent gaps introduced for maximal alignment. C, alignment of amino acid sequence of domain II of P46 with that of *M. capricolum* ribosomal protein L29.



**FIG. 8. Southwestern analysis of *S. citri* ribosomal proteins.** The *S. citri* ribosome purification procedure is described under "Experimental Procedures." 50- $\mu$ l aliquots of the supernatant (S-105) fraction (75  $\mu$ g) (lane 1) and washed ribosome fraction (35  $\mu$ g) (lane 2) were resolved on SDS-PAGE and assayed for IRS binding activity using Southwestern analysis.

these data support the role of the C-terminal domain of P46 in DNA-protein interaction.

Thus P46 could be a bifunctional protein. The N-terminal domain has a ribosomal function as it has high homology with the L29 ribosomal protein family. The C-terminal domain is seen as the one involved in IRS binding and might have regulatory function at the IRS of the genomic DNA. The bifunctional nature and the presence of DNA binding motifs in some ribosomal proteins has been reported and has led to speculations on the origin of the ribosomal proteins (58-64). It is interesting to note that in *B. subtilis* the HPB12-L24 protein has been described as a bifunctional ribosomal protein (L24) with histone-like properties and DNA binding activity (63, 65, 66).

Mollicutes are the smallest and simplest self-replicating organisms, and the currently dominating hypothesis is that they have evolved by degenerative (regressive) evolution from

Gram-positive bacteria with low guanine + cytosine genomes (1, 67, 68). In term of evolution strategy, mollicutes may have concentrated two different functions in a single gene during genome size reduction. Evidence for a single gene affording dual enzymatic function (malate/lactate dehydrogenase) has been described in *M. genitalium* by Cordwell and co-workers (69). Similarly, analysis of the complete sequence of the *M. genitalium* genome led Fraser and co-workers (6) to state that some *M. genitalium* proteins may have become adapted to perform more than one function. P46 has most likely a ribosomal protein function by its N-terminal L29 domain and a putative regulatory function on the IRS of the *rpsB/tsf/x* operon by its C-terminal domain. S2 and Ef-Ts are components of the translational machinery, and P46 may play an interconnecting role in the coordinated regulation of the components of the translational apparatus. However, the mechanism by which the binding of P46 on the IRS could influence the transcription of the *rpsB/tsf/x* operon is still unknown. *In vitro* studies of the transcriptional regulation of the *rpsB/tsf/x* operon with and without the presence of the P46 protein may help understand the extraribosomal function of P46.

REFERENCES

1. Weisburg, W. G., Tully, J. G., Rose, D. L., Petzel, J. P., Oyaizu, H., Yang, D., Mandelco, L., Sechrest, J., Lawrence, T. G., Van Etten, J., Maniloff, J., and Woese, C. R. (1989) *J. Bacteriol.* 171, 6455-6467
2. Saglio, P., Lhospital, M., Lafliche, D., Dupont, G., Bové, J. M., Tully, J. G., Freundt, E. A. (1973) *Int. J. Syst. Bacteriol.* 23, 191-204
3. An, G., Bendiak, D. S., Mamelak, L. A., and Friesen, J. D. (1981) *Nucleic Acids Res.* 9, 4163-4172
4. Bendiak, D. S., and Friesen, J. D. (1981) *Mol. Gen. Genet.* 181, 356-362
5. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessières, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S.-K., Codani, J.-J., Connerton, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Dusterhöft, A., Ehrlich, S. D., Emmerson, P. T., Entian, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Furma, S., Galizzi, A., Galleron, N., Ghim, S.-Y., Glaser, P., Goffeau, A., Golightly,