Volume 17 Number 7 1989 Nucleic Acids Research

Purification and characterization of RepA, a protein involved in the copy number control of plasmid pLS1

Gloria H.del Solar, Adela G.de la Campa, José Pérez-Martín, Theodora Choli[1] and Manuel Espinosa*

Centro de Investigaciones Biológicas, CSIC, Velázquez 144, E-28006 Madrid, Spain and [1]Max-Planck Institut für Molekulare Genetik, Ihnestrasse 73, D-1000 Berlin 33, FRG

Received January 26, 1989; Revised and Accepted March 1, 1989

ABSTRACT

The promiscuous streptococcal plasmid pLS1 encodes for the 5.1 kDa RepA protein, involved in the regulation of the plasmid copy number. Synthesis of RepA was observed both in *Bacillus subtilis* minicells and in an *Escherichia coli* expression system. From this system, the protein has been purified and it appears to be a dimer of identical subunits. The amino acid sequence of RepA has been determined. RepA shows the αhelix-turn-αhelix motif typical of many DNA-binding proteins and it shares homology with a number of repressors, specially with the TrfB repressor encoded by the broad-host-range plasmid RK2. DNase I footprinting revealed that the RepA target is located in the region of the promoter for the *repA* and *repB* genes. *Trans*-complementation analysis showed that *in vivo*, RepA behaves as a repressor by regulating the plasmid copy number. We propose that the regulatory role of RepA is by limitation of the synthesis of the initiator protein RepB.

INTRODUCTION

Promiscuous plasmids are usually restricted to hosts of the same type, Gram-positive or Gram-negative (1). However, a few examples of plasmids that are able to surpass the so-called barrier between both types of eubacteria have been described (2,3). Among them, the best characterized replicon so far is the streptococcal multicopy plasmid pLS1, which is capable of autonomous replication in *Streptococcus pneumoniae*, *B.subtilis* and *E.coli* (3). This 4408-bp plasmid is unusual because it shows features typical to both Gram-positive and Gram-negative replicons. Plasmid pLS1 replicates by assymmetric rolling circle (4), as other small staphylococcal plasmids (5). On the other hand, like many Gram-negative replicons, pLS1 has three iterons that are included in its *ori*(+) (6). Regulation of pLS1 replication involves two genes: *repB* that codes for the site-specific initiator protein RepB (de la Campa et al., in preparation); and *repA*, placed upstream of *repB*, transcribed in the same mRNA and that specifies for a small protein RepA. RepA binds to and induces DNA bending in a plasmid region that includes the *repAB* promoter and the three iterons, and we have suggested that the protein could act as a repressor involved in the regulation of the synthesis of RepB (4,6,7).

Regulation of gene expression can be performed through a variety of mechanisms. One of them, best characterized in bacteriophages, is exerted through the interaction of a repressor protein with its operator target(s) (8). Several instances of phage- and of *E.coli*-encoded repressor proteins are very well characterized (8−12). However, information on plasmid-encoded repressor proteins is not so well documented (1). In the narrow host-range plasmid R1, a repressor protein, CopB, is involved in the control of the transcription from the *repA* promoter to synthesize the mRNA for the replication initiator protein RepA (13−16). In the case of the broad host-range plasmid RK2 two proteins, KorB and TrfB,

have been shown to act independently (and postulated to cooperate) in the regulation of transcription from the *trfA* and *trfB* promoters to limit the rate of synthesis of the replication initiator protein TrfA (1,17, and references therein). However, no information on repressor proteins that exert an inhibitor-target regulation is so far available for the small, multicopy Gram-positive plasmids (18,19).

Here we report the characterization of the pLS1-encoded RepA protein and its synthesis in two hosts: *B.subtilis* minicells (under the *repA* own transcription/translation signals) and *E.coli* (under the control of a phage T7 promoter for the T7 RNA polymerase). The protein has been hiperproduced, purified and its amino acid sequence determined. RepA shares homology with a variety of repressors in a domain that shows the αhelix-turn-αhelix motif found in many DNA-binding proteins (20). In addition to this, the amino terminal region of RepA shows a high degree of homology with regions of some regulatory proteins, specially with the RK2-repressor TrfB (17). By DNase I footprinting, the target of RepA has been located at the *repAB* promoter. Cloning of the wild type and of a *repA* mutant (*cop*7) genes into plasmid pC194 allowed us to perform *trans*-complementation analysis in *S.pneumoniae*. We observed a RepA-conditioned reduction in the number of copies of the pLS1*cop7* mutant, indicating that RepA acts as a repressor for the regulation of pLS1 copy number. We propose that RepA-repression is exerted by limitation of the synthesis of the initiator protein RepB.

MATERIALS AND METHODS

Bacterial strains and plasmids.

The *E.coli* strains used were C600 (*thr*-1 *thi*-1 *leu*-6 *lacY*1 *fhuA*21 *supE*44), BL21 (F^- *hsd*R *gal*) and its isogenic strain BL21DE3 the two latter kindly provided by F.W.Studier. The BL21DE3 strain contains a single copy of the gene for phage T7 RNA polymerase in the chromosome under the control of the inducible *lac*UV5 promoter (21). The minicell-producing strain was *B.subtilis* CU403 (*divIVB*1 *thyA thyB metB*). *S.pneumoniae* 708 (*end*-1 *exo*-2 *trt*-1 *hex*-4 *malM*594) was employed for the preparation of pLS1-based plasmids and for the complementation assays. Plamids used were: pLS1 and pLS5 (3), the deleted derivative pLS1Δ 24 (6), the high copy number mutant pLS1*cop*7 (4, this work), the expression vector pET5 (21,22) and the constructions described in this work. Plasmid pET5 (a gift of F.W.Studier) is a pBR322-based transcription vector in which the phage T7 RNA polymerase ϕ*10* promoter has been cloned (21,22).

Transformations, DNA preparations and DNA manipulations.

Procedures for the transformation of the three bacterial species used have been described (3,4,23). Construction and analysis of recombinant plasmids were performed either in the C600 or BL21 strains (for the pET5-based plasmids) or in *S.pneumoniae* (for the pC194-based recombinants). To induce the expression of cloned genes under the control of the phage T7 ϕ*10* promoter, various plasmids (see Fig.1) were transferred to the BL21DE3 strain. Selection for ampicillin in these strains was of 200 μg/ml to avoid the overgrowth of plasmid-free cells (21,22). Preparations of pure plasmid DNA and enzymatic treatments of DNA were done as reported (4). Sequencing of the *cop*7 mutation was performed from plasmid pJS3 (3) at S.A.Lacks' laboratory. The mutation was transferred to plasmid pLS1 by exchange of the fragment *Pst*IB and its presence was confirmed by DNA sequencing. DNA sequence determinations were performed by the chemical method of Maxam and Gilbert (24).

Synthesis of proteins in minicells.
Preparation of *B.subtilis* minicells was as reported (25). Plasmid-encoded proteins in the minicells were analyzed by labelling with [^{35}S]methionine, [^{3}H]valine or [^{3}H]lysine followed by electrophoresis in 5% to 25% SDS−polyacrylamide gradient gels, as described (3,25). Under these conditions, the products of the pLS1-encoded *tet* and *repA* genes were visualized, but not the product of the *repB* gene.
Gene expression under the phage T7 ϕ10 promoter.
Before preparation of extracts, *E.coli* BL21DE3 cultures carrying plasmids with the ϕ10 promoter were routinely checked to calculate the percentage of cells with a functional system and of plasmid-free cells (21). Various plasmid constructions and several conditions for induction of the T7 RNA polymerase with isopropyl-β-D-thiogalactopyranoside (IPTG) were tested. After analysis of the proteins synthesized in low-scale cultures, the plasmid chosen was pLS19 (Fig.1) and the conditions to get the highest yield of RepA were: incubation at 37°C, induction with 0.5 mM IPTG for 7 min followed by addition of 200 μg/ml rifampicin (freshly prepared at 10 mg/ml in dimethyl sulfoxide) and subsequent incubation for 3 h. At this time, 2 ml-culture samples were withdrawn, labelled with [^{35}S]methionine for 10 min and disrupted, as described (3). The proteins were separated by SDS−polyacrylamide gel electrophoresis (SDS−PAGE) in 10% to 20% gradient gels. Determination of the size of small polypeptides was done by SDS−urea−PAGE (15% acrylamide, following the protocol recommended by BRL in Focus 6:3). The labelled proteins were analyzed by fluorography. Protein concentration was determined by the use of the BCA protein assay (Pierce, USA).
Determination of the native molecular weight of RepA.
The native molecular weight of RepA was determined by sedimentation in glycerol gradients and by gel filtration in agarose. The first procedure was performed essentially as described (26), but the gradients were run at 4°C for 20 h. Gel filtration was done on agarose (Bio Rad A−0.5m, 200−400 mesh) in a 1.2cm×37cm column, with a sample volume of 300 ml and a flow rate of 6 ml/h, as reported (10).
Amino acid sequence determination.
Automatic sequencing of the intact protein was done in a pulsed gas−liquid phase sequencer (27) (Applied Biosystems, mod. 477A) with on-line detection of the PTH amino acids by HPLC using a gradient system (Applied Biosystems PTH-analyzer, mod. 120A). The protein was dissolved in 20 μl 100% TFA and spread into a glass fiber disk (GF) which was precycled with 2 mg Polybrene.
DNase I footprinting.
Plasmid pLS1Δ 24 has a clockwise 570-bp deletion including coordinates 4240 to 401 of pLS1 (6). pLS1Δ 24 DNA (8.5 μg) was digested with *Hin*fI, the fragments were 5'-end labelled, digested with *Eco*RI, *Hin*dIII and *Nco*I and the resulting ten fragments were separated by non-denaturing 5% PAGE as described (4). The 302 bp *Nco*I−*Hin*fI fragment (coordinates 4221 to 685 of pLS1; containing the *repAB* promoter) was eluted, ethanol-precipitated and recovered at the specific activity of 10^7 cpm/mg. Each reaction (50 μl) contained 5 ng of the labelled fragment, 2 mM $MnCl_2$, 100 mM KCl, 20 mM Tris HCl (pH 8.0), 1 mM EDTA, 5 mM DTT, 5% ethylene glycol. In the RepA-containing reactions, two amounts (1 μg and 0.1 μg) of the protein were used. The mixtures were incubated 15 min at 20°C and then, 0.007 units of DNase I (2.7 units/mg, RNase- and protease-free, Cooper Biomedical, USA) were added. Incubation proceeded for another 5 min and the reactions were stopped by the addition of 25 μl of a 2 M ammonium acetate, 0.15

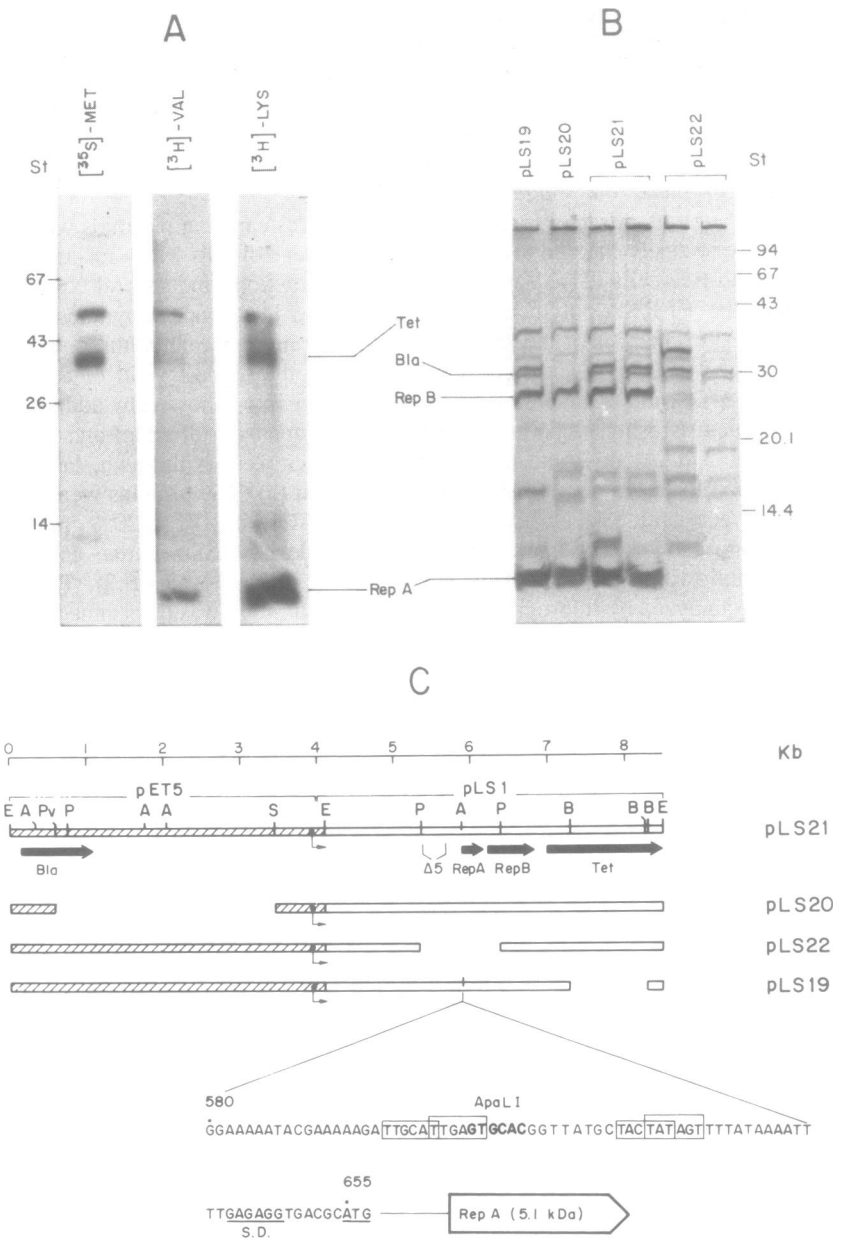

Figure 1. Expression of *repA*. **A:** Gene expression of *repA* under its own transcription/translation signals detected by SDS−PAGE and fluorography of labelled pLS1-encoded proteins synthesized in *B.subtilis* minicells. Amounts applied corresponded to about 8 μg of total cell protein. The position of unlabelled molecular weight standards (bovine serum albumin, ovoalbumin, chymotrypsinogen and lysozyme) is indicated on the left. **B:** Expression under the control of the phage T7 $\phi 10$ promoter in *E.coli* detected by SDS-PAGE and fluorography of labelled proteins synthesized in IPTG-induced cultures of *E.coli* BL21DE3 harbouring the indicated plasmids. Only the

M EDTA, 0.8 M Na-acetate solution containing 100 μg/ml of sonicated calf thymus DNA and 400 μg/ml tRNA. DNA was ethanol-precipitated, washed and suspended in 18 μl of sequencing loading buffer (24). A 6 μl sample (about 17,000 cpm) was loaded into 0.4 mm-thick 8% polyacrylamide sequencing gels (24) and run together with the sequencing reactions of the same fragment. Electrophoresis was conducted at 2000 V and was stopped 28 min after the bromophenol blue left the gel.

Trans-complementation analysis.

The *repA* gene was cloned into the unique *Hind*III site of a copy number mutant of pC194 (Ballester et al., in preparation), by digestion of pLS1 DNA with *Alu*I and *Bgl*I. The cohesive ends were eroded with nuclease S1 and made blunt, the DNA fragments ligated and used to transform competent *S.pneumoniae* cells. Selection for chloramphenicol resistance and screening of appropriate recombinants allowed the isolation and identification of a plasmid, termed pCGA3. This plasmid contains a pLS1 insert spanning from coordinates 505 to 804 in which the three iterons and the entire *repA* gene (with its own promoter) are included. The copy number of pCGA3 is about 200 per genome equivalent, as measured by densitometric scanning (28). This plasmid showed full structural and generational stability. To construct pCGA3*cop*7, the *Apa*LI−*Sty*I fragment of the cloned pLS1-wild type region was changed for the corresponding *cop*7 fragment. The complementation tests were carried out by transformation of competent cells harbouring resident plasmids and by selection for the donor marker. Although the pneumococcal strain used was rec^+ and donor and resident plasmids shared 300 bp homology, few instances of recombination between them were observed. Analysis of plasmid content in the transformants was performed by preparation of crude extracts from exponentially growing cultures at 4×10^8 viable cells/ml and agarose gel electrophoresis (4). The number of copies by genome equivalent were determined as above (28).

Densitometric scannings and computer work.

Soft laser densitometric scannings were performed in a LKB Ultroscan 2202 coupled to an Apple II computer. Search for homologies of RepA with the P.I.R. data bank resource was done with the use of the DNASTAR computer programs (DNASTAR, Inc., UK). Prediction of secondary structure of RepA was performed with the aid of the program PEPLOT (29), included in the UWGCG computer programs. In addition, the program PLOTCHOUFASMAN (developed by B.Jameson, S.Modrow, M.Moth and H.Wolf of the Max von Pettenkofer Institute of the University of Munich) was employed for the graphic output in the prediction of the secondary structure of RepA.

best conditions for expression are shown. For pLS19, IPTG was added 7 min before rifampicin and the total incubation time was of 3 h. For pLS20 and pLS21, IPTG and rifampicin were added at the same time and the incubation times were of 3 h (pLS20 and left track of pLS21) or of 1 h (right track of pLS21). For pLS22, IPTG was added 7 min before (left track) or simultaneously to (right track) the addition of rifampicin and the total incubation time was of 3 h. Amounts applied were corrected to load approximately 10 μg of total cell protein. The position of the products of genes *tet*, *bla*, *repB* and *repA* is indicated. The position of the unlabelled molecular weight standards (phosphorylase b, bovine serum albumin, ovoalbumin, carbonic anhydrase, trypsin inhibitor and α-lactalbumin) is indicated on the right. **C:** (upper) Maps of the pET5-based plasmids constructed in this work. Black dot and arrow indicates the position of the *ϕ10* promoter. The deletion in pLS5 is indicated. Only the pertinent gene products (filled arrows) and restriction sites are shown (E:*Eco*RI; A:*Apa*LI; Pv:*Pvu*I; P:*Pst*I; S:*Sal*I; B:*Bcl*I). Derivatives of plasmid pLS21 were constructed as follows: pLS20 by removal of the *Pvu*I−*Sal*I fragment of the pET5 moiety; pLS22 by deletion of the *Pst*IB fragment of the pLS1 moiety and pLS19 by deletion of the *Bcl*I fragments of pLS1. (lower) Nucleotide sequence of the pLS1 region pertinent for this work. The two possible overlapping promoters, the *Apa*LI site, the putative ribosomal binding site (S.D.) and the initiation codon of RepA are indicated.

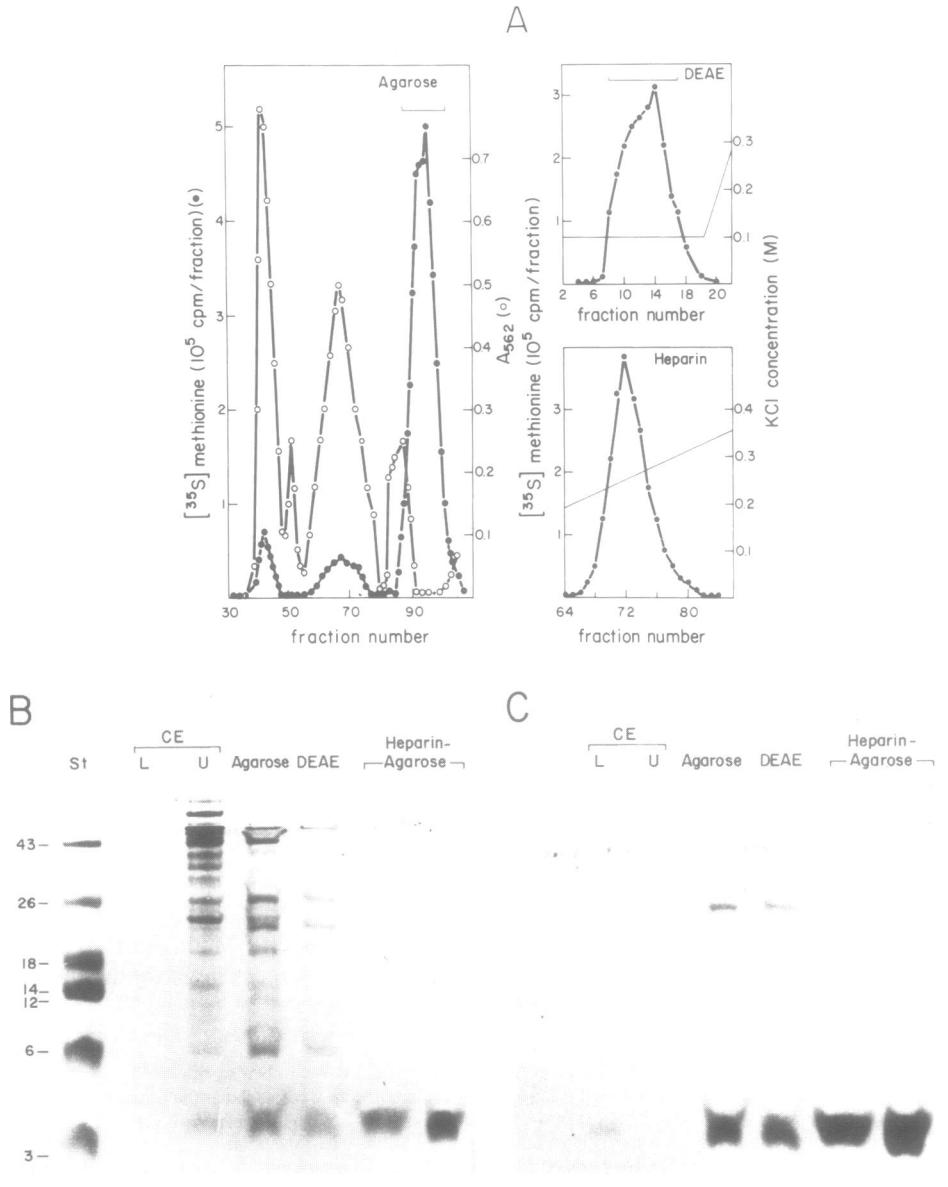

Figure 2. Purification of RepA. **A:** Agarose, DEAE Sephacell and Heparin−agarose fractionations. Appropriate fractions (horizontal bar) were pooled and applied to the columns, as indicated in the text. **B:** Proteins revealed by Coomassie Blue staining after SDS−urea−PAGE. St: molecular weight standards (ovoalbumin, α-chymotripsinogen, β-lactoglobulin, lysozyme, cytochrome C, bovine trypsine inhibitor and insulin). **C:** fluorography of the same gel. CE: crude extracts, labelled (L) or unlabelled (U).

RESULTS AND DISCUSSION

Expression of repA

To demonstrate that pLS1 codes for RepA and to show the RepA synthesis, differential labelling of proteins in *B.subtilis* minicells was performed with three different amino acids. As shown in Fig.1A, synthesis of RepA was only visualized when tritiated amino acids were used (instead of the commonly used [^{35}S]methionine) in spite of the low content of Val residues (4%) of RepA as compared to the Met (9%) or Lys (16%) content. We ascribe such a difference to the 70-fold intensification obtained by the fluorographic detection of ^{3}H over ^{35}S. The results show that RepA is synthesized in *B.subtilis* from pLS1, under its own transcription and translation signals. The *repAB* mRNA start point has been located by S1 mapping (not shown) and two possible overlapping promoters could be ascribed (Fig.1C). We favour the promoter situated more upstream because it shows the consensus spacer of 17 nucleotides (30) instead of the short 15-nucleotides spacer of the downstream promoter. If this is the case, the *repAB* promoter would correspond to the one previously proposed (3).

The pET5 vector/BL21DE3 host cloning system allows the specific and efficient synthesis (and the labelling) of proteins encoded by plasmid-borne genes under the control of the T7 ϕ*10* promoter (21,22). We could only clone in pET5 the pLS1- derivatives with the 332-bp deletion of pLS5 (3), probably because of the reported homology between this pLS1 region and the pBR322 origin (31). To unequivocally identify the pLS1-encoded gene products, different plasmid constructions were made (Fig. 1C). Plasmid pLS21 is a composite between pET5 and pLS5. Deletion of the *bla* gene of pLS21 and identification of the *repB* gene product were achieved by the construction of pLS20. In addition, identification of the *repA* and *repB* gene products was possible by construction of pLS22, in which the *Pst*IB fragment of pLS1 was removed. IPTG-induced cells carrying pLS20 or pLS21 showed high levels of synthesis of RepA and RepB, whereas no detectable levels of these proteins were observed for cells harbouring pLS22 (Fig.1B). Reorientation of the pLS5 moiety (from pLS20) in relation to the ϕ*10* promoter abolished the synthesis of pLS1-coded proteins (results not shown). The product of the *tet* gene of pLS1 could not be detected, perhaps because it is a membrane-bound protein (3). Since the cells did not grow well when the *tet* gene was present, it seemed that its product was toxic to the host (not shown). Consequently, pLS19 (in which the *tet* gene was deleted) was the best vector for high expression levels of RepA. Once the conditions of induction with this plasmid were optimized, we could obtain a high yield of RepA, amounting to about 5% of the total cell protein.

Purification of RepA

Purification of RepA from induced cultures of *E.coli* BL21DE3 carrying pLS19 was achieved as follows (Fig.2). *Step 1:* One litre of an exponentially growing culture (3×10^8 cell/ml) was induced under the conditions described in Methods. The cells were collected by centrifugation, washed with 100 ml of buffer A (500 mM NaCl, 20 mM Tris HCl (pH 7.5), 1 mM EDTA, 1mM DTT, 5% ethylene glycol) and suspended in 10 ml of the same buffer. The cell paste was passed twice through a chilled French-pressure cell at 20,000 psi. The lysate was centrifuged at 15,000 $\times$ g 20 min and divided into 2.5-ml aliquots. *Step 2*: A 2.5-ml sample of crude extracts in buffer A (15.4 mg of protein) was mixed with 500 μl of [^{35}S] methionine-labelled extract (12,000 cpm/μml), applied to a 1.6cm$\times$84cm agarose column (Bio-Rad A$-$0.5m, 200$-$400 mesh) and the flow rate was set at 18 ml/h. Fractions (2 ml) were collected, their radioactivity and protein concentration

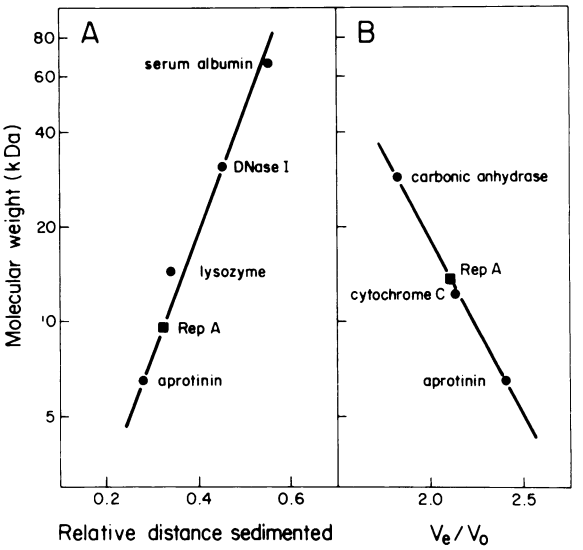

Figure 3. Native conformation of RepA. Sedimentation in glycerol gradients (**A**) or agarose gel filtration (**B**) of RepA and of protein standards. Dextran Blue was used to determine the void volume (V_0) and its value was compared to the protein elution volumes (V_e). RepA was positioned by [^{35}S] radioactivity determination.

measured, and the presence of RepA was analyzed by SDS−PAGE and fluorography. *Step 3*: Peak fractions (Fig.2A) were pooled and dyalized against buffer B (20 mM Tris (pH 8.0), 1 mM EDTA, 5 mM DTT, 5% ethylene glycol) containing 100 mM KCl. The sample was applied to a 0.9cm×13cm DEAE Sephacell (Pharmacia) column. The flow rate was set at 14 ml/h and 4-ml fractions were collected. RepA did not bind to the matrix in this salt concentration, which resulted in a high degree of purification of the protein (Fig.2B). *Step 4*: Peak fractions containing RepA (Fig.2A) were pooled, dialyzed against buffer B supplemented with 50 mM KCl and applied to a 1.4cm×4cm heparin−agarose column (Affi-Gel, Bio-Rad). The flow rate was set at 17 ml/h and 0.87-ml fractions were collected. RepA eluted at 250 mM KCl in a 50 to 500 mM KCl gradient. The protein seemed to be pure at this stage, as judged by SDS−urea−PAGE and fluorography (Fig.2). In these gels, the apparent molecular weight of RepA was of 5 kDa, in agreement with the predicted size deduced from the nucleotide sequence (5112 Da). The final yield of the purified RepA was 1.6 mg per litre of culture. The protein was concentrated by dialysis against buffer B (without ethylene glycol), supplemented with 50 mM KCl and 50% glycerol and stored at −25°C for several months without any appreciable loss of activity, as judged by its ability to bind to DNA and to protect the *Apa*LI site from digestion by this enzyme (31).

Native molecular weight of RepA

The native molecular weight of the protein was determined by sedimentation through a 15% to 35% glycerol gradient and by gel filtration through agarose and comparison to appropriate molecular weight markers (Fig.3). We could not use gel filtration on Sephadex because RepA was retained by this matrix, as expected for a basic protein of small size. From the glycerol gradients (in buffer B, but without ethylene glycol and with 50 mM KCl), RepA seemed to have a molecular weight of about 9,600 (Fig.3A). From the agarose

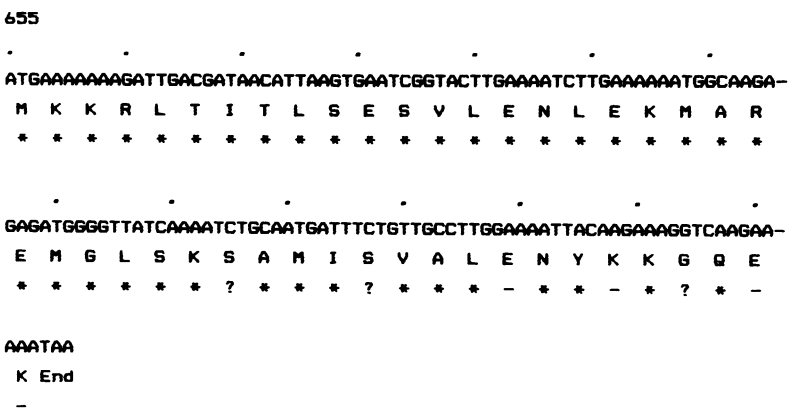

Figure 4. Amino acid sequence determination of RepA. The coding sequence for RepA (coordinates 655 to 792 of pLS1) and the predicted residues from the DNA sequence are shown. The amino acid sequence of the intact protein was determined by automated protein sequencer analysis. Amino acid identities are indicated by asterisks; question mark indicates possibility of identity; dash indicates no clear result.

column (in buffer B supplemented with 200 mM KCl), RepA eluted at a position corresponding to a protein with a molecular weight of 13,400 (Fig.3B). Similar results were obtained for elutions carried out at 50 or at 500 mM KCl (not shown). An indication of monomerization of RepA was obtained when the elution was performed at 2 M NaCl (not shown). We may conclude that, under the experimental conditions employed, RepA seemed to be synthesized as a dimer of identical subunits.

Amino acid sequence of RepA

Determination of the amino acid sequence of RepA (rather than determination of its amino terminus) was important to discard any cross contamination with other low-molecular weight peptides and to determine if the protein was subjected to post-transcriptional processing. Since RepA would be the first described repressor protein encoded by a Gram-positive replicon, it was worth unequivocally characterizing the protein. As shown in Figure 4, the amino acid sequence of RepA coincided with that predicted from the DNA sequence, although the identification of three residues was uncertain (S29, S33 and G42) and other four could not be determined (residues E37, K40, E44 and K45). We believe that the unambiguous characterization of RepA would prove important for future studies on its secondary and tertiary structures.

Characteristics of RepA

Predictions of the secondary structure and characteristics of RepA were achieved by computer-assisted programs. The isoelectric point of RepA was predicted to be 9.54, due to its 9 strongly basic residues. Predictions of the secondary structure of RepA by the PEPLOT program (29) showed the existence of a clear αhelix-turn-αhelix motif, typical of many DNA-binding proteins (20). This motif extends from residues 12 to 24 (αhelix-1), and from residues 29 to 37 (αhelix-2). From residues 5 to 9, a clear βsheet structure is observable. A relatively high hydrophobic moment was found between residues 12 to 22, included in the αhelix-1 of RepA (Fig.5A). This hydrophobic moment predicts that one part of this region faces the interior of the protein and the other faces the exterior (28). A flexible segment was also predicted to be between residues 25 to 28 (the turn motif).

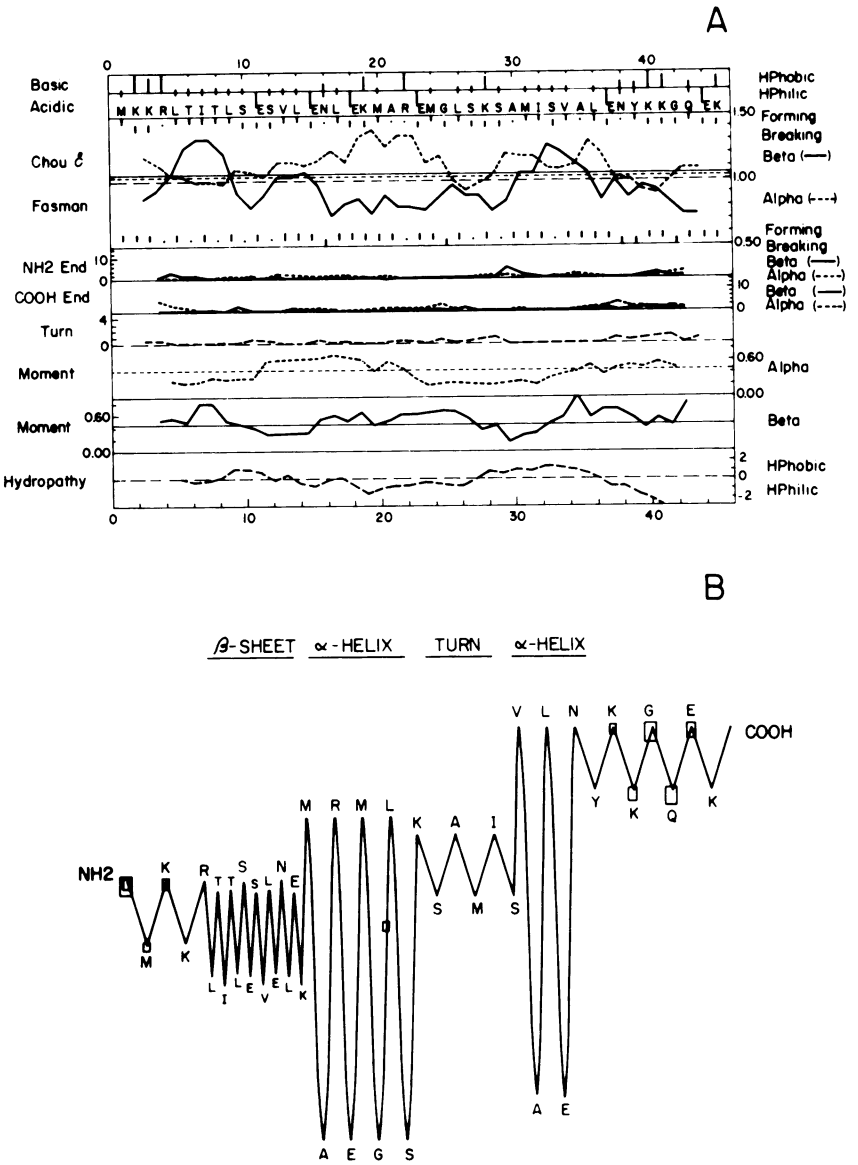

Figure 5. Computer-derived predictions on secondary structure of RepA. **A:** Drawing of the output of the PEPLOT program according to ref. 29. The upper part shows the sequence and the sequence schematic of RepA with the characteristics of the residues indicated on the left and right parts. The Chou and Fasman (38) alpha and beta conformational potentials, end-predictions and reverse turn potentials are indicated below. The last panels depict the alpha and beta hydrophobic moments and the hydropathy profile based on the Kyte and Doolitle predictions (39). **B:** Drawing of the output of the PLOTCHOUFASMAN computer program. The residues are indicated and, in the upper part, the predicted regions for αhelix, βsheet or turn motifs are underlined. The amino- and carboxi-termini of RepA could have a random coil structure. Small boxes indicate the most probable antigenic sites, which were confirmed by other computer programs.

	Position	α-helix								turn		α-helix									
		1				5					10					15					20
RepA	17	L	E	K	M	A	R	E	M	G	L	S	K	S	A	M	I	S	V	A	L
Lambda cII	26	T	*	*	T	*	E	A	V	*	V	D	*	*	Q	I	S	R	W	K	R
Lambda Cro	16	Q	T	*	T	*	K	D	L	*	V	Y	Q	*	*	I	N	K	A	I	H
Lambda cI	34	Q	*	S	V	*	D	K	*	*	M	G	Q	*	G	V	G	A	L	F	N
galR	4	I	K	D	V	*	*	L	A	*	V	*	V	A	T	V	S	R	*	I	N
LacI	6	*	Y	D	V	*	E	Y	A	*	V	*	Y	Q	T	V	S	R	*	V	N
φ105R	21	Q	V	Q	L	*	E	K	A	N	*	*	R	*	Y	L	A	D	I	E	R
RK2 KorB	171	K	G	D	I	*	K	*	I	*	K	*	P	A	F	I	T	Q	H	V	T
RK2 TrfB	35	Q	A	T	F	*	T	S	L	*	*	T	R	G	*	V	S	Q	A	V	H

RepA (1,19)	M	K	K	R	L	T	I	T	L	S	E	S	V	L	E	N	L	E	K
TrfB (1,19)	*	*	*	*	*	*	E	[S]	Q	F	[Q]	E	A	[I]	[Q]	G	*	*	V
P22 antirep (72,90)	*	*	*	*	*	M	G	C	[I]	[P]	L	K	K	*	[N]	G	W	L	F
R1 CopB (35,53)	A	[R]	*	*	*	*	H	K	E	I	K	V	F	[V]	K	*	P	L	*
434 Cro (9,27)	K	*	[R]	*	[I]	A	[L]	K	[M]	T	[Q]	[T]	E	*	A	T	K	A	G
Lambda cIII (6,24)	A	G	W	P	[V]	[A]	G	C	P	*	*	*	[L]	*	*	R	[I]	T	[R]

Figure 6. (upper) Comparison of the αhelix-turn-αhelix motif of RepA with those of various repressors and with lambda Cro. Position of first residue and relative residue positions of the DNA-binding domain in the proteins are indicated. (lower) Comparison of the amino terminus of RepA with homologous regions of various regulatory proteins. Amino acid positions are indicated in brackets. * indicates amino acid identities; boxed are the functional homologous residues.

The use of the PLOTCHOUFASMAN program helped to depict the αhelix-turn-αhelix motif in RepA, although the predicted structures seemed to be slightly displaced as compared to the PEPLOT predictions (Fig.5B). The most clear antigenic sites were predicted to be at residues 41 to 43, which was confirmed by a number of different computer programs.

In addition, the αhelix-turn-αhelix motif of RepA (Fig.6) fulfils the stereochemical requirements proposed by Ohlendorf et al. (32). At the relative positions 5 and 9, the conserved A and G residues are located. Residues at the relative positions 4 (M), 8 (M), 10 (L) and 15 (M) are hydrophobic; at positions 3 (K) and 7 (E) hydrophilic residues are placed. Accordingly, we could predict that residues at the relative positions 1, 2, 6, 11−13, 16, 17, and 20 would be involved in the interaction with the DNA when the protein is bound (32). The general structure of RepA at its DNA-binding motif seems to be in agreement with that of the well characterized repressor and Cro proteins from phages lambda and 434. By homology with those proteins, we could define the αhelix-1 of RepA from residues 17 to 24 and the αhelix-2 from residues 28 to 36. When we applied the systematic method of Dodd and Egan (33), the score of RepA for residues at relative positions 1 to 20 (Fig.6) was 1174, close to the 1141 and 1161 values reported for the plasmid RK2 proteins KorB and TrfB, respectively (17). Homologies with a number of repressors (as well as with Cro proteins from phages lambda and 434) were found at the αhelix-turn-αhelix motif (Fig.6). Search for homologies between RepA and the P.I.R. data base gave the highest score for residues 340 to 384 of the human involucrin (33% identity in a 36 amino acids overlap; not shown). The second highest score was obtained with the RK2 repressor TrfB (25% identity in a 32 amino acids overlap). The strongest homology is located at the amino termini of RepA and TrfB, as was also the case for lambda cIII and

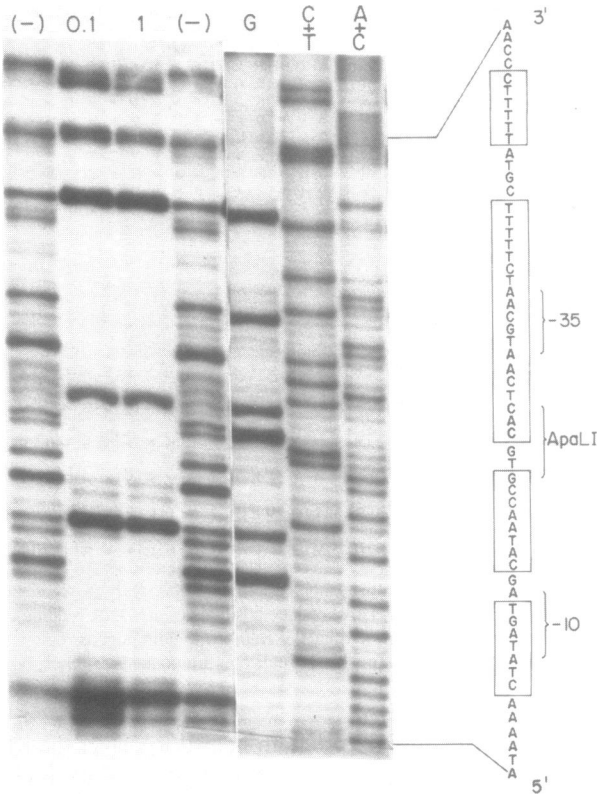

Figure 7. DNase I protection of the RepA-target DNA by purified RepA protein. Two amounts (0.1 and 1 μg) of protein were used; (−) indicates control samples without RepA. Tracks G, CT and AC of sequenced RepA target DNA are also shown. The nucleotide sequence of the labelled non-coding strand of pLS1 and the RepA-protected sites (boxed) are indicated on the right, as well as the position of the *repAB* promoter and the *Apa*LI restriction site (coordinate 607).

434 Cro (Fig.6). This would suggest a functional role for the amino terminal domain of these proteins apart from (or in conjunction with) the role of the DNA-binding domain. Recently, it has been reported (34) that the amino-terminal arm of the lambda repressor wraps around the DNA and makes contacts in the major groove on the back side of the DNA. Homologies were also found between the amino terminus of RepA and regions of the plasmid R1-repressor CopB and of the phage P22 antirepressor (Fig.6). No homology was found between RepA and the ORF D of the staphylococcal plasmid pNS1 (35). This plasmid drew our attention because its ORF D is placed upstream of the postulated replication initiator protein, showing a genetic arrangement which resembles that of pLS1.

Characterization of the nature of the *cop*7 mutation, at the level of its nucleotide sequence, demonstrated that the mutation consisted in a transversion C to A at coordinate 743. The result of this single-base change is the substitution of the A30 residue for an E residue. The copy number of pLS1*cop*7 was about 160 as compared to the 25 copies of the wild type plasmid. We have not yet purified the mutant protein, but prediction of its secondary structure by the above mentioned programs did not show any substantial alteration of it

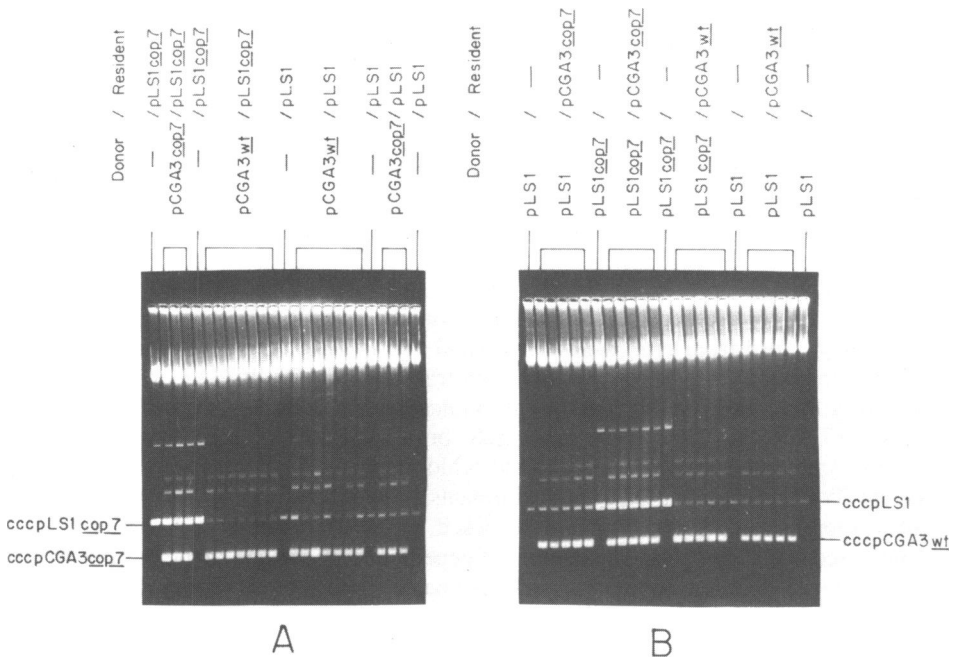

Figure 8. Effect of *repA* product on pLS1 copy number analyzed by *trans* complementation Tests I (**A**) and II (**B**). Plasmid-bearing pneumococcal cells were transformed with the indicated donor plasmid DNAs and selection was applied for the incoming plasmid. Crude extracts of various clones were prepared and electrophoresed in 1.2% agarose gels. Homoplasmid strains used as controls are indicated by (−). Chromosomal DNA migrates at the upper part of the gels and the positions of the covalently closed circular (ccc) monomers of pLS1 and pC194-replicons are indicated. Note a high amount of multimers in the strains harbouring pC194-derived plasmids.

(as compared to the wild type RepA) except for an increase in its average hydrophilicity and a decrease in its isoelectric point. Nevertheless, since the mutation is placed within the αhelix-2 of the protein and it resulted in the change of an hydrophobic for a strong acidic residue, we may assume that the *cop*7 mutation should affect the affinity of the RepA7 mutant protein for its target. Based on previous comparisons (36), it appeared that the *repA* codon usage was more similar to that of *B.subtilis* than to *E.coli*, although the unfrequently used codon UUG (L, 14% for *B.subtilis* and 9% for *E.coli*) was used at the frequency of 33%. At present it is still premature to assign this difference to a special feature of the codon usage of streptococcal genes.

Determination of the target of RepA

The target of RepA within pLS1 DNA has been defined to be in the vicinity of the *Apa*LI single-site (7). To refine the binding site of RepA, DNase I footprintings were performed by using two amounts of protein which were bound to the labelled non-coding strand of pLS1. The results (Fig.7) clearly showed four protected regions located between coordinates 581 to 628. These regions covered the *repAB* promoter, in which the *Apa*LI site is located. No further protected regions were observed and since the three iterons are included in the labelled fragment, we may conclude that the plasmid three direct repeats (coordinates 534 to 566) are placed outside the target of RepA, confirming previous assumptions (6,31).

Computer search for the presence of secondary structures or direct repeats in the protected region showed the possibility of generation of a weak hairpin, centered at coordinate 579 and in which the −35 region of the *repAB* promoter is included (not shown). Whether this structure is the actual target of RepA or if a non evident operator structure is present in this region is a question under current investigation. The binding of RepA at the *repAB* promoter, in conjunction with the homologies between RepA and the plasmid-coded transcriptional repressors TrfB and CopB, suggested to us that the role of RepA would be at the transcriptional level, by controlling its own synthesis and thus limiting the synthesis of RepB.

In vivo activity of RepA

Two kinds of complementation tests, based on the quantitative incompatibility tests developed by Nordström et al. (37) were carried out: in Test I, competent pneumococcal cells carrying either pLS1 or pLS1*cop*7 were transformed with pCGA3 or pCGA3*cop*7 DNA (Fig.8A). In Test II, competent *S.pneumoniae* cells harbouring pCGA3 or pCGA3*cop*7 were transformed with either pLS1 or pLS1*cop*7 DNA (Fig.8B). In both cases, selection was done for the donor plasmid (chloramphenicol or tetracycline resistence, respectively). After plating of the transformants, several colonies of the same size and incubated for the same period of time were picked, grown and crude extracts were prepared and analyzed in agarose gels. The number of generations (**g**) for one single pneumococcal transformant to give rise to a colony was estimated to be 27 (37). From a colony to a full grown culture the value of **g** was estimated to be 10, based on the number of viable cells per colony, the incubation time, the doubling time of the cultures and the final number of viable cells. Thus, the **g** value for all cases was estimated to be 37. The results of these complementation analyses (Fig.8) can be summarized as follows: when the wild type RepA was supplied in *trans*, the number of copies of the wild type plasmid pLS1 was not significantly affected, but the copy number of the mutant pLS1*cop*7 was severely reduced. This effect was not due to a permanent alteration of the pLS1*cop*7 plasmid because transformation of the plasmid-free strain with the DNA of these 'low' copy number plasmids fully restored the *cop*7 phenotype (not shown). No effect was observed when the mutant protein RepA7 was given in *trans* nor when the vector pC194 was used. The RepA-conditioned reduction in the copy number of pLS1*cop*7 seemed to be due to a secondary inhibitor−target interaction rather than to incompatibility caused by a primary inhibitor (18). This is because no incompatibility towards pLS1 was observable and because no further reduction in the copy number of pLS1*cop*7 was detected after 60 generations without selective pressure. It would thus seem that once the equilibrium in the heteroplasmid population was reached, no further changes occurred. Furthermore, the results show that the mutant RepA7 is impaired in its ability to control the plasmid copy number, as expected from the nature of its mutation. The *cop*7 mutation leads to a 5-fold increase in the plasmid copy number without detectable unstabilization, similarly to the *cop*B mutants of plasmid R1 (15). Deletions affecting RepA showed that the protein is not essential for plasmid maintenance (not shown).

From the results presented here, we may conclude that RepA acts as a modulator of the copy number of plasmid pLS1 in a secondary control mechanism. It is worth noting that regulation of plasmid copy number by *trans*-acting proteins does not seem to be a common plasmid feature (1). A putative countertranscribed RNA II has been postulated

to negatively control the pLS1 copy number, most likely being the main replication control element of the plasmid (4,6).

ACKNOWLEDGEMENTS

Thanks are due to J.C.Alonso for providing accession to the programs PEPLOT and PLOTCHOUFASMAN. We are grateful to S.A.Lacks and S.Ballester for their help at the primary characterization of the *cop*7 mutation, to P.López and Prof.B.Wittmann-Liebold for discussions and advices, to M.T.Alda and A.Hurtado for their technical help and to W.M.Newton for corrections of the manuscript. Research at CIB supported by Grant BIO88-0449 of the CICYT. The agreement between CSIC and Max-Planck Gesellschaft supported one of us (M.E.) for stays in Berlin.

*To whom correspondence should be addressed

REFERENCES

1. Thomas,C.M. (1988) Biochim. Biophys. Acta 949, 253-263.
2. Kok,J., van der Vossen,J.M.B. and Venema,G. (1984) Appl. Environ. Microbiol. 48, 726-731.
3. Lacks,S.A., López,P., Greenberg,B. and Espinosa,M. (1986) J. Mol. Biol. 192, 753-765.
4. del Solar,G.H., Puyet,A. and Espinosa,M. (1987) Nucl. Acids Res. 15, 5561-5580.
5. te Riele,H., Michel,B. and Ehrlich,S.D. (1986) EMBO J. 5, 631-637.
6. Puyet,A., del Solar,G.H. and Espinosa,M. (1987) Nucl. Acids Res. 16, 115-133.
7. Pérez-Martín,J., del Solar,G.H., de la Campa,A.G. and Espinosa,M. (1989) EMBO J., submitted.
8. Patshne,M. (1986) A genetic switch. Cell Press & Blackwell Scientific Publications. Cambridge and Palo Alto.
9. Meyer,B.J., Kleid,D.G. and Ptashne,M. (1975) Proc. Natl. Acad. Sci. USA 72, 4785-4789.
10. Poteete,A.R., Hehir,K. and Sauer,R.T. (1986) Biochemistry 25, 251-256.
11. van Kaer,L., van Montagu,M. and Dhaese,P. (1987) J. Mol. Biol. 197, 55-67.
12. Lehming,N., Sartorius,J., Oehler,S., von Wilcken-Bergmann,B. and Müller-Hill,B. (1988). Proc. Natl. Acad. Sci. USA 85, 7947-7951.
13. Nordstöm,K., Molin,S. and Light,J. (1984) Plasmid 12, 71-90.
14. Rownd,R.H., Womble,D.D., Dong,X., Luckow,V.A. and Wu, R.P. (1985) In Helinski,D.R., Cohen,S.N., Clewell,D.B., Jackson,D.A. and Hollaender,A. (eds), Plasmids in Bacteria. Plenum Press, New York, pp. 335-354.
15. Molin,S., Stougaard,P., Light,J., Nordtröm,M. and Nordström,K. (1981) Mol. Gen. Genet. 181, 123-130.
16. Riise,E. and Molin,S. (1986) Plasmid 15, 163-171.
17. Theophilus,B.D.M. and Thomas, C.M. (1987) Nucl. Acids Res. 15, 7443-7450.
18. Novick,R.P. (1987) Microbiol. Rev. 51, 381-395.
19. Projan,S.J. and Novick,R. (1988) Plasmid 19, 203-221.
20. Pabo,C.O. and Sauer,R.T. (1984) Ann. Rev. Biochem. 53, 293-321.
21. Studier F.W. and Moffatt,B.A. (1986) J. Mol. Biol. 189, 113-130.
22. Rosenberg,A.H., Lade,B.N., Chui,D-S., Lin,S-W., Dunn,J.J. and Studier,F.W. (1987) Gene 56, 125-135.
23. López,P., Espinosa,M. and Lacks,S.A. (1984) Mol. Gen. Genet. 195, 402-410.
24. Maxam,A.M. and Gilbert,W. (1980) In Grossman,L. and Moldave,K. (eds.), Methods in Enzymology, Academic Press, New York, vol.65, pp.499-560.
25. Espinosa,M., López,P. and Lacks,S.A. (1984) Gene 28, 301-310.
26. de la Campa,A.G., Springhorn,S.S., Kale,P. and Lacks,S.A. (1988) J. Biol. Chem. 263, 14696-14702.
27. Hewick,R.H., Hunkapillar,M.W., Hrod,L.E. and Dreyer,W.J. (1981) J. Biol. Chem. 256, 7990-7997.
28. Projan,S.J., Carleton,S. and Novick,R.P. (1983) Plasmid 9, 182-190.
29. Gribskov,M., Burgess,R.R. and Devereux,J. (1986) Nucl. Acids Res. 14, 327-334.
30. Rosenberg,M. and Court,D. (1979) Annu. Rev. Genet. 13, 319-353.
31. Pérez-Martín,J., del Solar,G.H., de la Campa,A.G. and Espinosa,M. (1988) Nucl. Acids Res. 16, 9113-9126.
32. Ohlendorf,D.H., Anderson,W.F. and Matthews,B.W. (1983) J. Mol. Evol. 19, 109-114.
33. Dodd,I.B. and Egan,B.J. (1987) J. Mol. Biol. 194, 557-564.
34. Jordan,S.R. and Pabo,C.O. (1988) Science 242, 893-899

35. Noguchi,N., Aoki,T., Sasatsu,M., Kono,M., Shishido,K. and Ando,T. (1986) FEMS Microbiol. Lett. 37, 283–288.
36. Piggot,P.J. and Hoch,J.A. (1985) Microbiol. Rev. 49, 158–179.
37. Nordström,K., Molin,S. and Aagaard-Hansen,H. (1980) Plasmid 4, 332–349.
38. Chou,P.Y. and Fasman,G.D. (1978) Adv. Enz. 47, 45–147.
39. Kyte,J. and Doolitle, R.F. (1982) J. Mol. Biol. 157, 105–132.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.