
The nucleotide sequence surrounding the replication origin of the cop3 mutant of the bacteriocinogenic plasmid Clo DF13

A.R.Stuitje, E.Veltkamp, J.Maat[†] and H.L.Heyneker*

Dep. Genetics, Biological Lab., Free University, De Boelelaan 1087, Amsterdam-Buitenveldert, Netherlands, *Genentech. Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080, and [†]Dep. Biochemistry and Biophysics, University of California, San Francisco, CA 94143, USA

Received 19 February 1980

ABSTRACT

The nucleotide sequence from about 100 base-pairs downstream to about 600 base pairs upstream the CloDF13 replication origin has been determined. A comparison of this sequence with the corresponding ColE1 origin sequence reveals that:

- . The sequence at the origin of replication is conserved.
- . There are large differences in the nucleotide sequence downstream the replication origin, whereas there is a large homology in the region of about 410 base-pairs upstream the replication origin.
- . This conserved region might code for a largely homologous basic, arginine rich polypeptide of about 45 amino-acids, for both ColE1 and CloDF13.
- . Although there are large differences in the primary structure of the region coding for the 100 nucleotide RNA, the secondary structure of this region seems to be conserved.

INTRODUCTION

Previous studies have shown that the bacteriocinogenic plasmid CloDF13 replicates unidirectionally from a unique origin, which is located at 2.8% on the CloDF13 physical map (1). In our study on the control of CloDF13 replication, we used both conditional (2,3) and nonconditional (4) plasmid copy number (cop) mutants. It was shown that the region between 1.8% and 11.5% of the plasmid genome contains sufficient information for autonomous plasmid DNA replication. In addition this region also contains information for the control of the plasmid copy number since the nonconditional cop 3 mutation has been mapped within this part of the plasmid (3). Recently we were able to show that the temperature sensitive cop 1 ts mutation also resides within this region (Stuitje *et al.* manuscript in preparation).

In this paper we initiate a study on the molecular basis of the control of CloDF13 plasmid replication. For this purpose we intend to compare the

base sequences of the wild type CloDF13 origin region to that of the different copy number mutants. In this study we have determined the base sequence of the 1.8-9% region of the CloDF13 -cop 3 mutant. A comparison between this sequence and the sequence of the corresponding region of plasmid ColE1 is given. A possible role of particular sequences in the control of CloDF13 replication is discussed.

MATERIALS AND METHODS

Materials

Hpa I endonuclease was a generous gift of Mrs. M. Lupker-Wille. Sau 96 was a gift of Dr. J. van Embden. All other restriction endonucleases were obtained from New England Biolabs. T4 kinase and DNA polymerase I were obtained from Boehringer Mannheim (FRG). Piperidine, Hydrazine and dimethylsulfate were purchased from Serva. All other materials used have been described elsewhere (5,6,7).

Isolation of plasmid DNA

The CloDF13 miniplasmids used in this study, were derived from the CloDF13-cop3 plasmid as described previously (1,3). For the isolation of the plasmid DNA, Escherichia coli cells harbouring the plasmid were grown in BHI medium. Plasmid DNA was isolated from cleared lysates which were prepared from stationary phase cells, as described by Veltkamp et al., (8).

Isolation of restriction fragments

Restriction fragments produced by cleavage of about 2-20 μ M plasmid DNA with an endonuclease of interest, were separated on 5% polyacrylamide slabgels (400x200x0,4 mm). Electrophoresis was carried out at appr. 15 V/cm for about 1.5-4 hr in 90 mM Tris-borate pH 8.3, 1 mM EDTA. DNA bands were visualized by staining with ethidium bromide and irradiation with long wave U.V. light. The DNA was recovered from the gel by O/N elution at 37°C with 500 mM Ammonium acetate, 10 mM Magnesium acetate, 0.1% SDS, 0.1 mM EDTA, without crushing the gel (9). Ethidium bromide was removed by extraction with n-butanol, and the DNA was precipitated twice. This approach enables rapid isolation of DNA fragments within the size range of 20-500 bp.

Terminal labelling of restriction fragments

For sequencing procedures, DNA fragments were routinely labelled at the 5' termini with γ - 32 P-dATP and T_4 polynucleotide kinase as described previously (7) 3' end labelling of HinFI fragments was carried out by complementary strand synthesis, with DNA polymerase I and α - 32 P-dNTP's as described by Donelson and Wu (10).

Sequencing procedures

3' or 5' one end labelled restriction fragments were routinely sequenced, using the Maxam and Gilbert chemical degradation procedure (9). In special cases 5' end labelled fragments were sequenced with the procedure described by Maat and Smith (5), which involves nicktranslation in the presence of dideoxynucleoside triphosphates (ddNTP's). In order to sequence across certain restriction sites (see Figure 1) we have used the primed synthesis sequencing procedure as described by Sanger et al. (6). For this purpose single stranded template DNA was prepared by exonuclease III (Biolabs.) digestion of Bam HI or HpaI generated linear pEV35 duplex DNA: 3-4 μ M digested pEV35 DNA, was extracted with phenol and precipitated twice:

The exo III reaction was carried out at a DNA concentration of about 0.2 μ M/ μ l in a reaction mixture containing 66 mM Tris-HCl pH 8.0, 0.66 mM MgCl₂, 1.0 mM 2-mercapthoethanol, for 2 hours at 37°C (21). The reaction was stopped by heating the mixture for 10 min. at 100°C. Primer DNA was prepared by cleaving 10 μ M pEV₂₂ with HhaI endonuclease. Suitable fragments (see Figure 40-130 bp) were isolated as described above. The priming reaction was carried out by heating a primer/template (3:1) mixture (66 mM Tris-HCl pH 8.0, 0.66 mM MgCl₂, 1.0 mM 2-mercapthoethanol, 0.05 μ M Template/ μ l) for 3 min. at 100°C, and subsequently incubating this mixture for 1 hr. at 65°C.

Primed synthesis reactions using Klenow DNA polymerase I (Boehringer), dideoxynucleosides triphosphates and α - 32 P-dATP were carried out as detailed by Sanger et al. (6).

Sequencing gels were prepared as described previously (5)

RESULTS

A fine cleavage map of the 0-15% part of the CloDF13 genome

Previous studies have shown that both the origin of replication and

sequences involved in the control of CloDF13 replication are located within the 1.8-11.5% part of the plasmid genome (1,3). To determine the base sequence of the origin region, we have constructed a fine cleavage map, comprizing this part of the CloDF13 genome. For this purpose we have used deletion mutant pEV35, which was constructed from a CloDF13-Tn901 insertion mutant, by in vitro deletion of Bam HI restriction fragments (3). Figure 1b shows the physical map of pEV35 as well as the deleted parts of Clo DF13 and Tn901.

This plasmid contains single HpaI and BamHI cleavage sites, located at resp. 0% and 15%. To construct a fine cleavage map of the 0-15% region of the Clo DF13 genome, pEV35 linearized with either HpaI or BamHI was 5' end labelled and subsequently digested with resp. BamHI or HpaI. In this way, one end labelled 0-15% fragments were obtained. Restriction cleavage sites were mapped by using the partial digestion method described by Smith *et al.* (21). Figure 1a shows the cleavage map of this 0-15% part of the Clo DF13 genome.

The nucleotide sequence of the 1.8%-9% region of the Clo DF13 copy mutant Clo DF13-cop 3

To determine the base sequence of the Clo DF13-cop3 origin region, we have constructed a small Clo DF13-cop3 deletion mutant, pEV22. This deletion mutant was obtained from a Clo DF13-cop3:Tn901 insertion mutant pJN 67 by *in vitro* deletion of HaeII restriction fragments (7). The plasmid pJN 67 contains a Tn901, transposon inserted between the origin of replication and the immunity gene (1,3), at 11.5% on the Clo DF13 physical map.

Figure 1C shows the physical map of pEV22 as well as the deleted parts of pJN 67. pEV22 only contains the Clo DF13 region between 1.8% and 11.5%, as well as the region of transposon Tn901 between 0-45% comprizing the β -lactamase gene. Cleavage with HaeII endonuclease generates two fragments, a small 300 bp Hae II origin fragment containing the 1.8-5% part of Clo DF13, and a large fragment containing the Clo DF13 region between 5% and 11.5%, as well as the 45% part of the Tn901. Previous studies have shown that the cop3 mutation is preserved on this plasmid. (1,3) pEV22 is present in *E.coli* cells to the extend of about 200 copies per cell and therefore, plasmid DNA is easily isolated without amplification with chloramphenicol. pEV22 restriction fragments were prepared and sequenced as described in materials and methods. Figure 1A shows the DNA strands

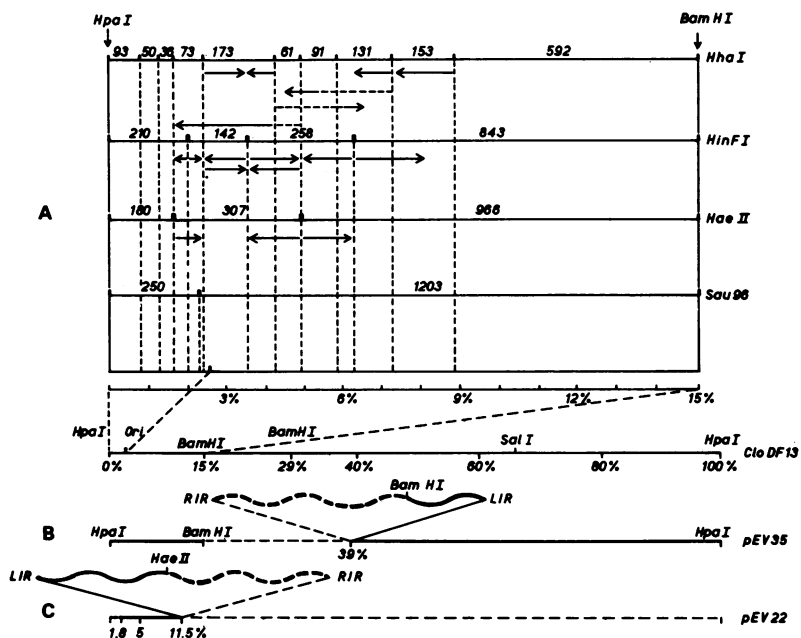


Figure 1. A) Fine cleavage maps of the 0-15% part of plasmid Clo DF13. The arrows show the part of the restriction fragments sequenced. The direction of the arrows indicate which strands were sequenced; they are aligned in the 5' to 3' direction. B) Physical map of plasmid pEV35, the dashed lines indicate those parts of Clo DF13 and Tn901 which have been deleted *in vitro* with BamHI endonuclease. C) Physical map of plasmid pEV22, the dashed lines indicate those parts of CloDF13 and Tn901 which have been deleted *in vitro* with HaeII endonuclease.

sequenced by this approach and the extent of overlapping sequences obtained. However by using this approach we were not able to sequence across certain essential restriction sites. (see figure 1A). For this purpose we have used the primed synthesis method as described by Sanger et al. (6), which involves the synthesis of a radioactive copy of the single-stranded target sequence with DNA polymerase I and a restriction fragment, which is used as a primer.

In this study we have prepared single stranded template DNA by exonuclease III digestion of linear duplex DNA. For this purpose we used pEV35 DNA linearized at either the HpaI or BamHI site (see above, and also figure 1B). Exonuclease III digestion thus generates a single stranded origin region.

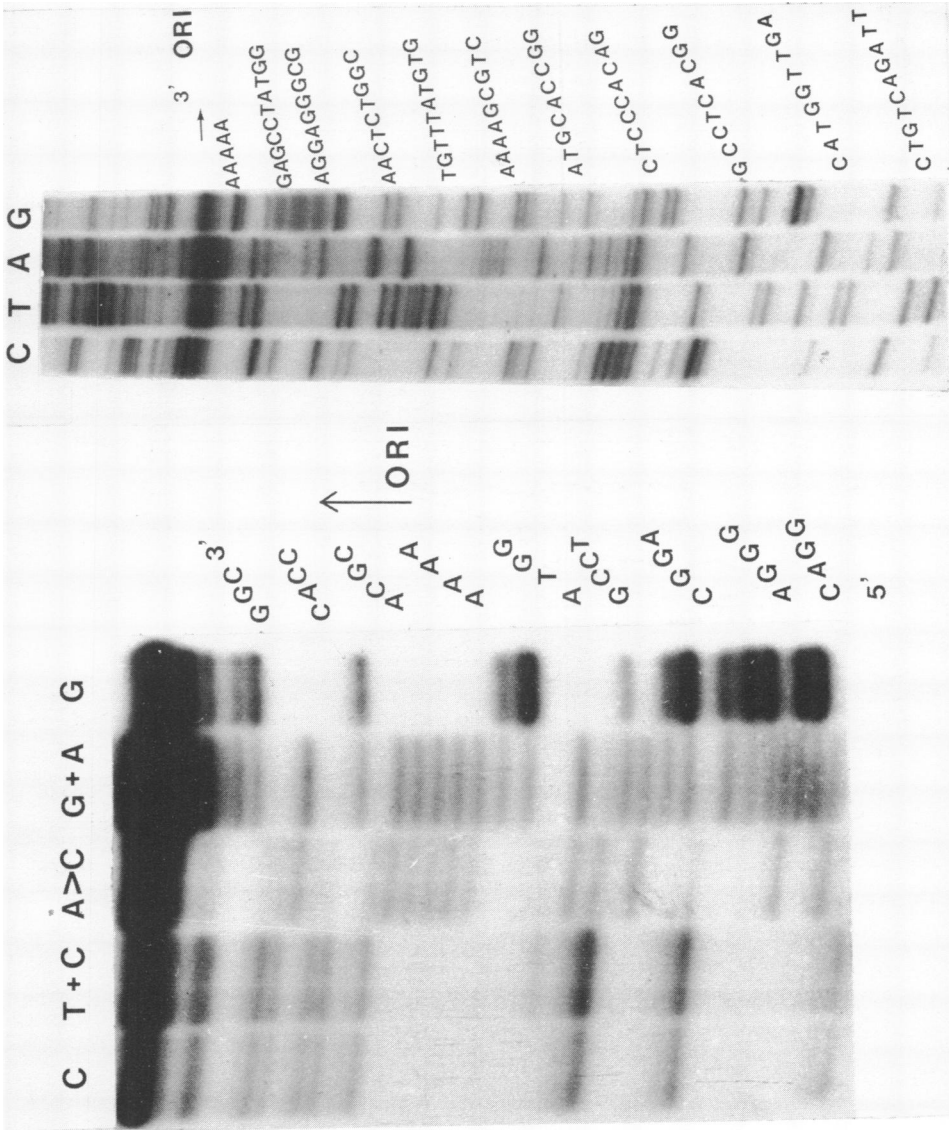


Figure 2. Result of respectively the Gilbert and Maxam procedure (left) and the primed synthesis DNA sequencing method (right), applied to the same part of the + strand, overlapping the replication origin (see text)

Subsequently HhaI origin fragments (see figure 1A), isolated from pEV22 were annealed to this partially single stranded template, and the reactions were carried out as described in materials and methods.

In figure 1A the Hha fragments which were used as a primer are indicated by the dotted lines, whereas the solid lines represent the part of the DNA strands sequenced by this approach. Figure 2, shows an example of respectively the Maxam and Gilbert procedure and the primed synthesis method, applied to the same part of the + strand*, overlapping the origin of replication.

The DNA sequence of the origin region is shown in figure 3. The upper-line of this figure represents that part of the Clo DF13 - strand between the HaeII site at 1.8% and the HhaI site at 9%, comprising 705 base pairs. The lower sequence represents the corresponding sequence of the Col E1 origin region, as determined by Oka et al (11). The sequences are aligned in such a way, that maximum homology was obtained.

DISCUSSION

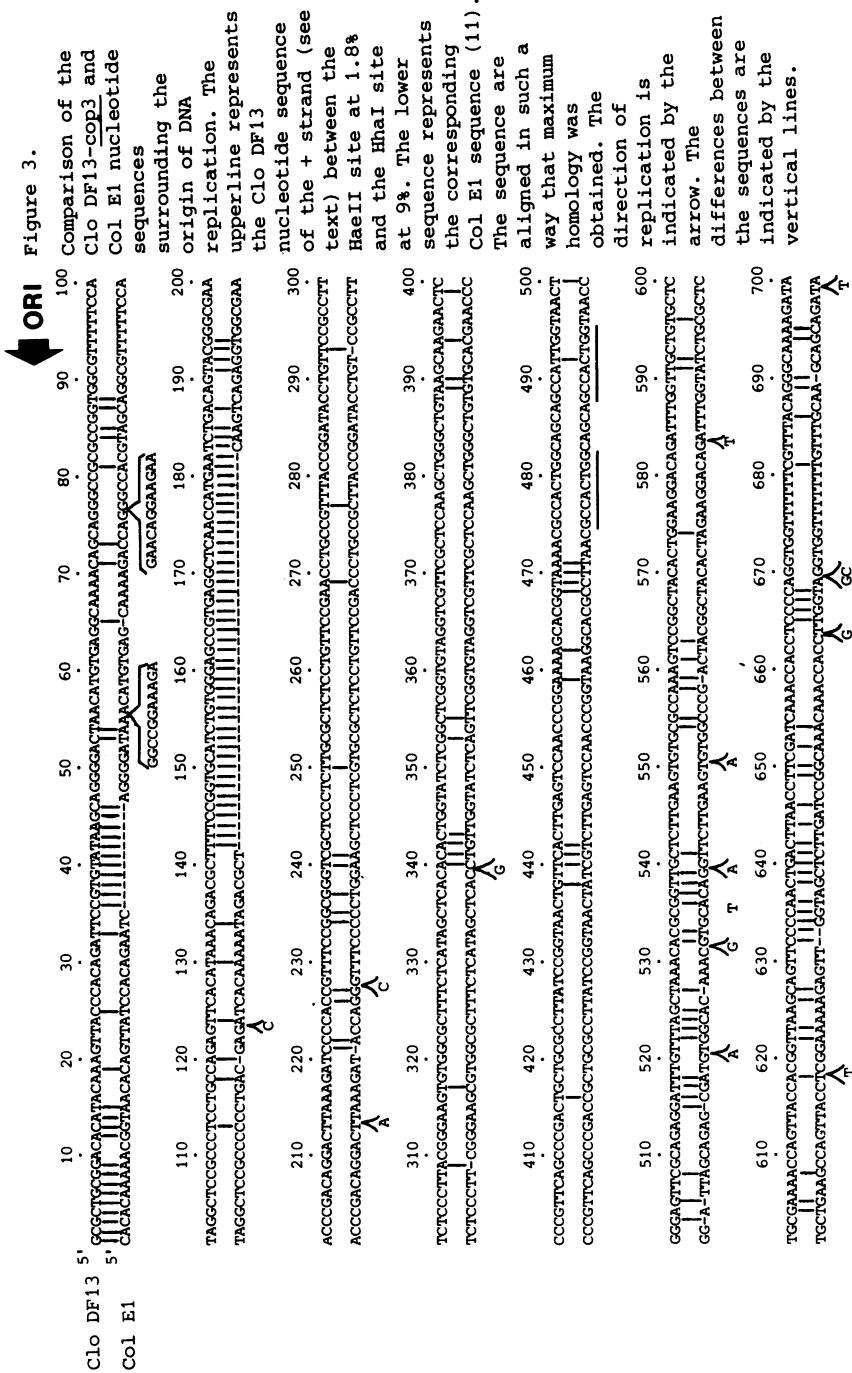
Both Col E1 and Clo DF13 belong to a group of plasmids with a relaxed replication control. As a result these plasmids are present in about 10-20 copies per cell (12,13). The replication of both Col E1 and Clo DF13 depends on the presence of the same host encoded DNA gene products, whereas their replication continues in absence of de novo protein synthesis (14). Furthermore Clo DF13 and Col E1 share large homologies in respect of their overall genetic organization (see figure 4), however, the gene products are not mutually exchangeable. These data suggest that Clo DF13 and Col E1 share homology with regard to the organization of base sequences within the origin region.

Perusal of the corresponding sequences, shown in figure 3 allows some speculation about the function of the base sequences within this region. Some of the features of this base sequence are discussed below:

a) The origin of replication

Tomizawa et al. (15), showed that the initiation of Col E1 replication in vitro, or the transition point from primer RNA to DNA, occurs within a

* Clo DF13 replicates unidirectionally, therefore the DNA strands can be distinguished ; the DNA strand synthesized discontinuously is arbitrarily denoted as the - strand.



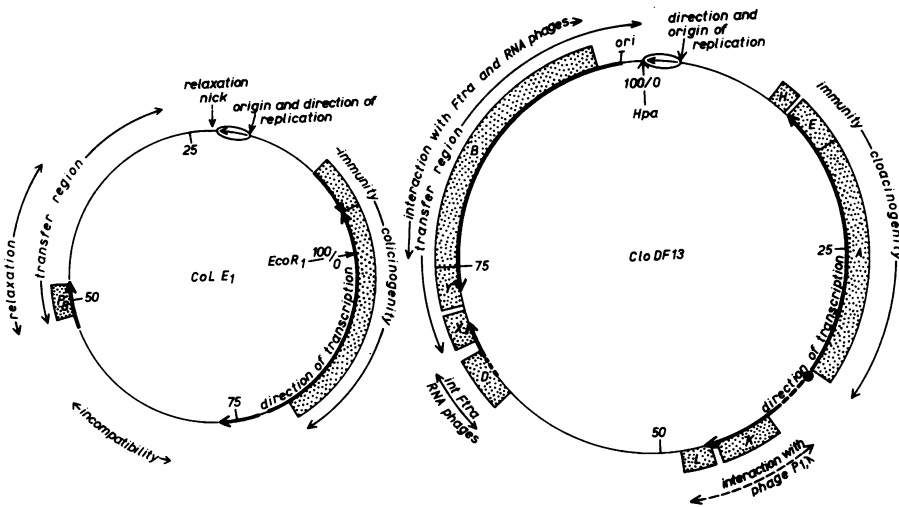


Figure 4. Comparison of genetic and functional maps of the non-conjugative plasmids Col E1 and Clo DF13. For references see Inselburg *et al.* (22), Dougan *et al.* (23), Collins *et al.* (24), Andreoli *et al.* (25), van de Pol *et al.* (26,27,28) and Stuitje *et al.* (1,3).

region of the Col E1 genome, which consists of a sequence of five A residues. The corresponding sequence of five T residues, within the Col E1 H-strand (15) is indicated in figure 3. Furthermore this figure shows that for Clo DF13 the same sequence is present at a position which highly agrees with the site of the replication origin, as determined by electron-microscopic analysis of *in vivo* replicating molecules (1). Bolivar *et al.* (16) also showed that *in vivo* replication of pMB1 (a Col E1 type plasmid) initiates at this sequence. Moreover a comparison of the Clo DF13 and Col E1 sequences surrounding the origin, reveals a large divergence in the nucleotide sequences downstream the replication origin (i.e. in the direction) of DNA chain growth) whereas there is a remarkable conservation of the nucleotide sequence, within a region of approximately 410 base pairs, upstream the replication origin (with exception of the ± 40 base pairs insertion, see below). This observation is consistent with the results of Oka *et al.* (11), which showed that in case of Col E1, a region of about 430 base pairs downstream the replication origin contains sufficient information for autonomous plasmid replication. We therefore suggest that the sequences which are indispensable for the plasmid DNA replication are

located within this region.

b) Presence of a 40 base pair insertion near the replication origin of Clo DF13

As shown in figure 3 a continuous stretch of bases seem to be inserted at about 45 base pairs upstream the replication origin of Clo DF13, when compared to Col E1. Whether this sequence plays a role in the control of plasmid replication remains unclear, however recently we were able to show that the cop3 mutation, which results in an increased plasmid copy number, is not located within this sequence (Stuitje et al., manuscript in preparation).

c) The 100 nucleotide RNA cistron

Upstream the replication origin there is a second region within the Clo DF13 sequence (position + 500-670), which diverges largely from the corresponding Col E1 sequence. It has been shown that, in case of Col E1, this region is transcribed by RNA polymerase both in vivo and in vitro (17,18). Moreover, Morita and Oka, have determined the sequence of the approximately 100 nucleotide transcription product (17). In figure 3, this noncoding RNA molecule is transcribed from left to right. Transcription initiates at position 580, while the terminator region is located at position 684-686 (17) in this figure. Recently van den Elzen et al. (manuscript in preparation) showed that the corresponding region of Clo DF13 also codes for an approximately 100 nucleotide RNA in vivo. Figure 3 shows that although the primary structure of the corresponding region differ largely (about 30%), the base sequences of the RNA polymerase recognition (position 542-546) and binding (position 568-574) sites as well as the terminator region at the other end, seem to be conserved (17). In fact figure 5 shows that these corresponding RNA molecules can be folded into a similar secondary structures. Backman et al. (20) suggested that this 100 nucleotide RNA is required for replication since in pMB1-pSC201 co-integrates, the Col E1 type replication is eliminated when the pSC101 derivative is inserted at the AluI site (position 637) within this RNA coding region, whereas the insertion of a small EcoRI linker fragment (net insertion 10 bp) at the same site, is tolerated. The observation of Heffron et al. (33) and Conrad and Campbell (32), that an insertion of a

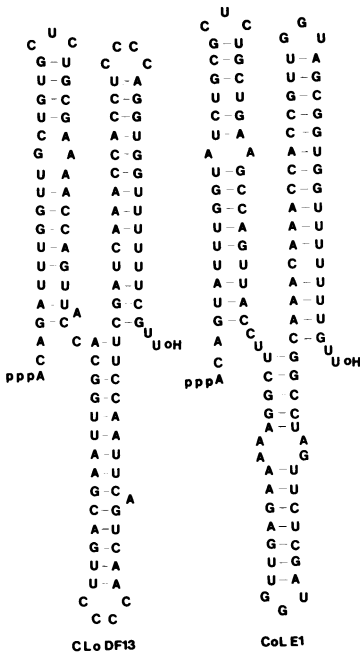


Figure 5.

A possible secondary structure for the 100 nucleotide RNA transcript.

bp EcoRI linker fragment, somewhere within the 100 nucleotide RNA coding region, caused an increase in plasmid copy number, also implicates that the 100 nucleotide RNA is involved in Col E1 replication and control of plasmid copy number. Backman *et al.* (20) proposed a model for the replication of Col E1 type plasmids in which this RNA is processed, and the product of this processing serves as a primer for elongation by DNA polymerase I at the origin of replication, since this RNA contains a region homologous to the sequence at the origin of replication (position 620-626). Supporting this model, Conrad and Campbell (32) showed that RNA'se III, an enzyme which preferentially cleaves double stranded RNA, is required for Col E1 replication *in vitro*.

However, perusal of the sequence of the 100 nucleotide RNA coding region of plasmid Clo DF13, shows that this region in contrast to Col E1, has no homology with the sequence at the origin of replication. Furthermore the results obtained by Oka *et al.* (11), showing that the 436 base pair Col E1 HaeIII fragment located between position 78-555 (figure 3) contains sufficient information to permit replication, implicates that the 100 nucleotide RNA is not required for replication. However, preliminary sequence data of the wild type Clo DF13 origin region, revealed that the

cop3 mutant contains at least one mutation within the 100 nucleotide RNA coding region. Further sequence analysis of the wild type plasmid and that of different nonconditional and conditional copy number mutants, may therefore help to resolve the contradictory results obtained by Oka et al. on the one side, and Backman et al. (20), Heffron et al. (33) and Conrad et al. (32) on the other hand.

d) Codon analysis

The codon analysis of the Clo DF13 sequence between 1.8% and 9% is shown in figure 6. Initiation and termination codons present on each DNA strand, were plotted on the corresponding reading frames. Long and short angled lines represent the initiation codons AUG and GUG respectively, whereas the vertical lines represent the termination codons UAG, UAA and UGA. The boxes indicate those parts of the sequence that might code for

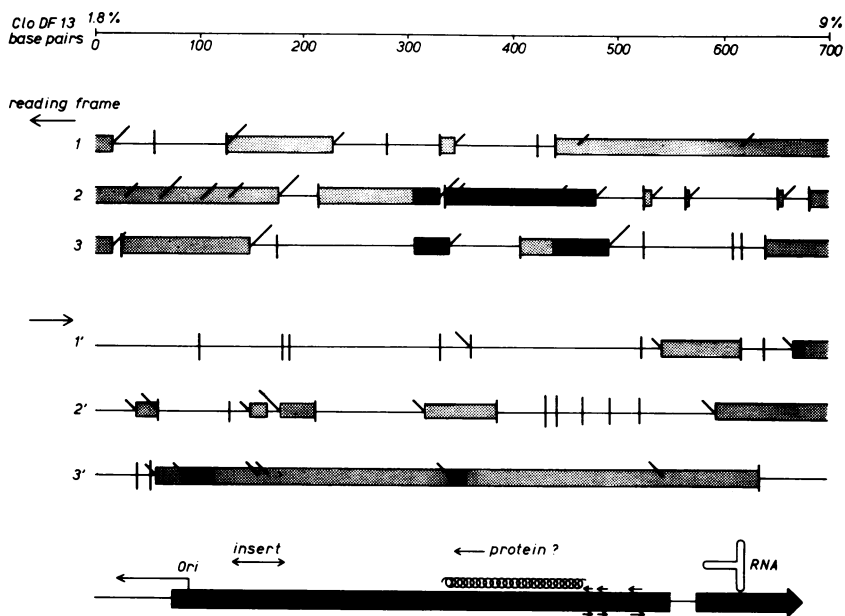


Figure 6.

Codon analysis of the Clo DF13-cop3 sequence, between 1.8 and 9% (see discussion). The lower part of this figure shows the corresponding Col E1 region, containing sufficient information to permit replication. The 100 nucleotide RNA is also indicated.

polypeptides. We have also looked for homology between these polypeptides and the polypeptides for which the corresponding Col E1 sequence might code. In figure 6 those coding sequences are indicated in black. The largest homologous coding sequence occurs between position 479-339 (reading frame 2). Translation of this coding sequence into an amino-acid sequence, reveals that we are dealing with a very arginine rich, basic protein (see figure 7). Interestingly this coding sequence is preceded by a large conserved symmetrical sequence, (22 base pairs, position 475-496.) figure 3 which in turn is composed of a smaller duplicated symmetrical sequence (underlined in figure 3). This sequence resembles known protein-DNA interaction sites. Staudenbauer (14) as well as Kahn and Helinsky (29) have demonstrated that no plasmid encoded proteins are required for the replication of plasmid Col E1. However, temperature sensitive copy number mutants have been isolated for both Col E1 and Clo DF13 (4,30) suggesting a possibly negatively controlled replication. In case of the temperature inducible copy number mutant Clo DF13-*cop1* ts, it was shown recently in our laboratory, that the kill phenotype at 42° C (4), depends on transcription initiated at the cloacin promotor located at approximate-

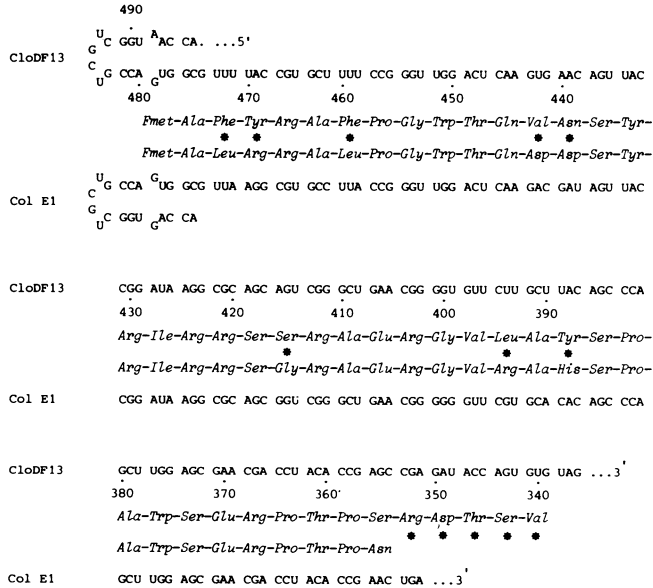


Figure 7. Comparison of the aminoacid sequence of a protein, possibly encoded by the origin region of plasmids Col E1 and Clo DF13. Differences between both sequences are indicated by an asterisk.

ly 32% (van den Elzen et al., manuscript submitted). In figure 3 this transcription proceeds from right to left. Deletion of the cloacin promoter, results in the loss of the kill phenotype, without effecting the temperature inducible copy number. However by in vitro recombination of restriction fragments, it was shown that the kill phenotype is restored by positioning the β -lactamase promoter of Tn901 in the same orientation. Furthermore we have located the cop1 ts mutation, within the 1,8-12% region of the plasmid (Stuitje et al., manuscript in preparation). These results suggest that Clo DF13 replication is possibly negatively controlled by a plasmid encoded protein. However it has been shown (31) that a single base substitution in a non-coding region of SV40 DNA can cause a temperature sensitive replication. Furthermore Conrad et al. (32) have demonstrated that a plasmid encoded, highly secondary structured 100 nucleotide non-coding RNA is involved in the control of Col E1 plasmid copy number.

In view of these observations it is still premature to conclude that both plasmid encoded RNA as well as protein are involved in the regulations of plasmid replication. Further sequence analysis of plasmid mutants with an altered plasmid copy number is therefore necessary to obtain a more profound insight in how the replication of relaxed plasmids is regulated.

ACKNOWLEDGEMENTS

We are grateful to Richard van Schaik and all the members of the department of Physiological Chemistry of the University of Leiden, for their advice and help in the course of this work. This investigation was supported by The Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for Advancement of Pure Research (Z.W.O.)

REFERENCES

- (1) Stuitje, A.R., Veltkamp, E., Weyers, P.J. and Nijkamp, H.J.J. (1979). Nucl. Acid. Res. 6, 71-80.
- (2) Kool, A.J. and Nijkamp, H.J.J. (1974). J. Bacteriol. 120, 569-578.
- (3) Stuitje, A.R., Veltkamp, E., Elzen, P.J.M. van den and Nijkamp, H.J.J. (1978). Nucl. Acid. Res. 5, 1801-1820.
- (4) Andreoli, P.M., Brandsma, R.A., Veltkamp, E. and Nijkamp, H.J.J. (1978). J. Bacteriol. 135, 612-621.
- (5) Maat, J. and Smith, A.J.H. (1978). Nucl. Acid. Res. 5, 4537-4546.
- (6) Sanger, F., Nicklen, S. and Coulson, A.R. (1977). Proc. Nat. Acad. Sci. USA 74, 5463-5467.

-
- (7) Ormond, H. van, Maa, J., de Waard, A. and van der Eb, A.J. (1978) *Gene* 4, 309-378.
- (8) Veltkamp, E., Barendsen, W. and Nijkamp, H.J.J. (1974). *J. Bacteriol.* 118, 165-174.
- (9) Maxam, A.M. and Gilbert, W. (1977). *Proc. Nat. Acad. Sci. USA* 74, 5605-564.
- (10) Donelson, J.E. and Wu, R. (1972). *J. Biol. Chem.* 247, 4661-4668.
- (11) Oka, A., Nomura, N., Morita, M., Sugisaki, H., Sugimoto, K. and Takamami, M. (1979). *Molec. gen. Genet.*, 172, 151-159.
- (12) Bazaral, M. and Helinsky, D.R. 1970, *Biochemistry*, 9, 399-406.
- (13) Veltkamp, E. and Nijkamp, H.J.J. (1976). *Biochim. Biophys. Acta* 425, 356-357.
- (14) Staudenbauer, W.L. (1978). *Curr. Top. Microb. and Immunol.* 78, 201-237.
- (15) Tomizawa, J.J., Ohmori, H. and Bird, R.E. (1977). *Proc. Nat. Acad. Sci. USA* 74, 1865-1869.
- (16) Bolivar, F., Betlach, M.C., Heyneker, H.L., Shine, J., Rodriquez, R.L. and Boyer, H.W. (1977). *Proc. Nat. Acad. Sci. USA* 74, 5265-5269.
- (17) Morita, M. and Oka, A., (1979). *Eur. J. Biochem.* 97, 435-443.
- (18) Levine, A.D., Rupp, W.D. In: *Microbiology 1978* (D. Schlessinger, ed.), 136-166. Washington D.C.: Am. Soc. Microbiol. 1978.
- (19) Kolter, R. and Helinsky, D.R. (1979). *Ann. Rev. Genet.* in press.
- (20) Backman, K., Betlach, M., Boyer, H.W. and Yanofsky, S. (1978). *Cold Spring Harbor Symp. Quant. Biol.* 43, 69-76.
- (21) Smith, H.O. and Birnstiel, M.L. (1976). *Nucl. Acid. Res.* 3, 2387-2398.
- (22) Inselburg, J. (1977). *J. Bacteriol.* 132, 332-340.
- (23) Dougan, G., Saul, M., Warran, G. and Sherratt, D. (1978). *Mol. gen. Genet.* 158, 325-327.
- (24) Collins, J. (1979). *Gene*, 6, 29-42.
- (25) Andreoli, P.M., Overbeeke, N., Veltkamp, E., Embden, J.D.A. van, and Nijkamp, H.J.J. (1978). *Mol. gen. Genet.* 60, 1-11.
- (26) Pol, H. van de, Veltkamp, E. and Nijkamp, H.J.J. (1980). *J. Bact.* manuscript submitted.
- (27) Pol, H. van de, Veltkamp, E. and Nijkamp, H.J.J. (1980). *Mol. gen. Genet.* 168, 309-317.
- (28) Pol, H. van de, Veltkamp, E. and Nijkamp, H.J.J. (1980). *Mol. gen. Genet.* manuscript submitted.
- (29) Kahn, M. and Helinsky, D.R. (1978). *Proc. Nat. Acad. Sci. USA* 75, 2220-2224.
- (30) Grindley, N.D.F., Grindley, J.N. and Kelly, W.F. (1978). In "Microbiology 1978" ed. D. Schlessinger. publ. Am. Soc. Microbiol. pp. 71-73.
- (31) Shortle, D. and Nathans, D. (1979). *J. Mol. Biol.* 131, 801-817.
- (32) Conrad, S.E. and Campbell, J.L. (1979). *Cell.* 18, 61-71.
- (33) Heffron, F., So, M. and McCarthy, B. (1978). *Proc. Nat. Acad. Sci. USA* 75, 6012-6016.
-