
Structure and regulation of gene expression of a Clo DF13 plasmid DNA region involved in plasmid segregation and incompatibility

Peter J.M.van den Elzen, Marcel J.J.Hakkaart, Arnold J.van Putten, Hanneke H.B.Walters, Eduard Velkamp and H.John J.Nijkamp

Department of Genetics, Biological Laboratory, Vrije Universiteit, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands

Received 23 September 1983; Revised and Accepted 17 November 1983

ABSTRACT

The bacteriocinogenic plasmid Clo DF13 contains genetic information involved in the accurate partitioning of the plasmid (parA and parB) as well as in incompatibility phenomena (incA, B, C and D). In this paper we report on the primary structure and regulation of gene expression of the 29% - 50% part of Clo DF13, containing the DNA regions incA, incB and parB as well as genes K and L.

According to the results of our DNA sequence analysis, mapping of transposon insertions, RNA blotting and S1 mapping experiments, we conclude that:

a) genes K and L are transcribed as one operon; transcription of this operon is initiated at a promoter (P_2) located at 32.5% and proceeds in a clockwise direction.

b) treatment of cells with mitomycin-C, significantly enhances transcription from P_2 , although this promoter is probably not directly repressed by lexA protein.

c) Termination of transcription of this operon occurs between genes K and L, as well as distal to gene L.

The possible role of gene products and/or sites, located within the 29-50% DNA region, in plasmid incompatibility and segregation is discussed.

INTRODUCTION

The bacteriocinogenic plasmid Clo DF13 (9,600 bp; 1,2) is maintained in Escherichia coli at approximately 10 copies per cell (3). Copy control mutants of Clo DF13 have been isolated (4-7), such as the non-conditional mutant Clo DF13-cop3 with a plasmid copy number of about 70 (8). Clo DF13 encodes at least 9 proteins; the location of the genes encoding these proteins, and the functions of their gene products are presented in figure 1.

According to the functions encoded, Clo DF13 can be divided into at least 3 DNA regions, namely regions involved in: (a) initiation and regulation of DNA replication (1.8-9%; 6, 7), (b) bacteriocinogenity, immunity, lysis and transport of the bacteriocin (9-32%; 2, 9-14) and (c) mobilization (67-100%; 15, 16). For a detailed review, see reference 7. Beside the genetic information involved in the functions mentioned above,

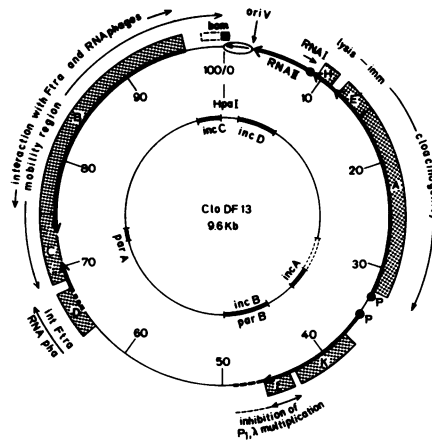


Figure 1. Genetic and functional map of bacteriocinogenic plasmid Clo DF13. The data presented in the outer circle are reviewed by Veltkamp and Stuitje (7). The inner circle shows the location of the DNA regions involved in incompatibility (incA, incB, incC and incD; 20) and stability (parA and parB; 18). For further explanation, see the text.

the Clo DF13 genome contains genetic information necessary for the stable maintenance of the plasmid (17). Recently Hakkaart, *et al.* localized 2 DNA regions on the Clo DF13 genome, parA and parB, which are involved in accurate partitioning of Clo DF13 plasmid molecules among daughter cells (see fig. 1; 18). Upon deletion of at least one of these par regions, mutants of Clo DF13 with an increased copynumber appeared to be unstable. Furthermore, these par⁻ plasmids are isolated from *E. coli* mainly as multimeric molecules. Another phenomenon involved in plasmid maintenance is plasmid incompatibility (19). Recently, four different regions involved in this incompatibility phenomenon (incA through incD, fig. 1) were mapped on the Clo DF13 plasmid genome (20).

It is remarkable, that one Clo DF13 region (between 45% and 50%) appears to be involved in both, plasmid incompatibility (incB; 20) and partitioning (parB; 18). Van de Pol *et al.* (21) demonstrated that the Clo DF13 DNA region between 35% and 50% encodes at least two different proteins, namely protein K (m.w., 21 kD) and protein L (m.w., 10.5 kD). Furthermore, they presented evidence that the product of gene L inhibits the multiplication of double stranded DNA phages, like P1 and λ .

In this paper, we report on the primary structure and the regulation of gene expression of an operon, located between 32.5% and 49%, comprising

Table 1. Bacterial plasmids

Plasmid ^a	Transposon	Insertion of transposon map position ^b	orientation ^c	Affected gene/function	Source of reference
Clo DF13- <u>cop3</u>	-	-	-	-	3, 5, 6
pVC22	Tn <u>5</u>	32.5	ND	<u>incA</u>	20
pJN52	Tn <u>901</u>	36.5	L	-	21, 24, 25
pJN66	Tn <u>901</u>	37.0	R	K	"
pJN50	Tn <u>901</u>	38.5	L	K	"
pJN63	Tn <u>901</u>	42	L	K	"
pJN03	Tn <u>901</u>	44.5	R	ND	24, 25
pJN56	Tn <u>901</u>	45	L	L, Dpi	21, 24, 25

^aPlasmids pVC22 and pJN52, 66, 63, 56 are derivatives of plasmid Clo DF13-cop3. pJN03 is a wild-type plasmid.

^bThe sites of integration of transposons Tn901 and Tn5 are given in percentages of the Clo DF13 genetic map, as determined previously (see references) and/or during this study.

^cThe orientation of the Tn901 insertion is indicated by capital L or R. L means that the left inverted repeat is orientated clockwise, whereas R indicates the opposite direction (24).

ND, not determined; Dpi, DNA phage interaction, stands for the inhibition of multiplication of DNA phages (21).

genes K and L. The promoter of this operon (located at 32.5%) can be activated in vivo by the addition of mitomycin-C. DNA sequence analysis of the DNA region between 29% and 50%, as well as mapping of transposon integration sites, enabled us to define more precisely genetic information involved in plasmid incompatibility (incA, incB) and partitioning (parB) as well as DNA phage interaction.

MATERIALS AND METHODS

Bacterial strains and plasmids

The plasmids used in this study are listed in table 1. The minicell producing strain P678-54 (22) was used as host strain for all Clo DF13 plasmids.

Media, Chemicals and Enzymes

Brain Heart Infusion medium (Oxoid Ltd., London) and Liquid Culture (LC) medium were normally used for growth in liquid and on agar plates. Except for Bam HI, all restriction enzymes used in this study were purchased from Boehringer Mannheim (FRG) and New England Biolabs. Bam HI was kindly

provided by C. Pols and was isolated according to Greene et al. (23). T4 polynucleotide kinase, E. coli DNA polymerase I, the "Klenow" fragment of DNA polymerase I and Calf Intestine Alkaline Phosphatase (CIAP) were supplied by Boehringer, Mannheim (FRG) S1 nuclease was purchased from Sigma Chemical Co. α - 32 P-dNTPs (350 Ci/mmol) and γ - 32 P-ATP were obtained from the Radiochemical Centre, Amersham, England. Piperazine-1,4-bis (2-ethane sulfonic acid) (PIPES) was supplied by Merck, Darmstadt (FRG).

Isolation of plasmid DNA, restriction enzyme analysis and preparation of DNA restriction fragments

DNA of Clo DF13 and Clo DF13 transposon mutants was isolated as described previously (10). Analytical and preparative digestions of plasmid DNA with restriction endonucleases were performed in reaction mixtures recommended by New England Biolabs. Analytical digests were run on 1-2% agarose gels in 20 mM Tris-HCl pH 7.6, 10 mM NaAc, 1 mM EDTA, or on 3 mm 5% polyacrylamide gels (acryl: bisacryl, 40:1.36; 26). Preparative electrophoresis was carried out on 5% polyacrylamide gels (acryl:bisacryl, 38:2, 27), after which the DNA fragments were recovered from the gel by electro elution (10).

DNA sequence analysis procedures

DNA sequence analysis was carried out with the chemical degradation procedure of Maxam and Gilbert (27). DNA fragments were labeled at the 5'-end or the 3'-end with T4 polynucleotide kinase or "Klenow" polymerase respectively. In order to dephosphorylate the 5'-ends of DNA restriction fragments we incubated the DNA fragment during 30 min at 45°C in 10 mM Tris-HCl pH 9.5, 1 mM spermidine, 0.1 mM EDTA with 0.5 unit calves intestinal alkaline phosphatase (CIAP). CIAP was inactivated at 75°C during 20 minutes, after which kinase buffer, T4 polynucleotide kinase and γ - 32 P-ATP were added directly. Gel electrophoresis to separate single end labeled fragments or single strands, as well as elution of DNA and base modification reactions were carried out according to Maxam and Gilbert (27). Eight, twelve and twenty percent sequence gels were prepared as described previously (28).

Isolation of RNA, electrophoresis, transfer to DBM-paper, and hybridization

RNA was isolated from logarithmically growing cells as described previously (2). In order to prepare RNA blots, 650 μ g RNA was layered on a 2.5% agarose gel (0.2 x 13 x 24 cm) containing 5 mM methylmercuric hydroxide (29). Electrophoresis was carried out for about 7.5 h at 100 V,

after which the gel was stained with ethidium bromide. Preparation of the gel and the diazobenzyloxymethyl (DBM) paper for transfer of the RNA was done as described previously (2). The RNA blots were hybridized with nick-translated (30) Clo DF13 fragments, washed and autoradiographed according to Wahl *et al.* (31).

Mapping of in vivo transcripts by S1 nuclease protection of DNA-RNA hybrids

5'- and 3'-ends of *in vivo* synthesized RNAs were mapped essentially as described by Berk and Sharp (32). For this purpose we made use of DNA restriction fragments, labeled at their 5'-ends or at their 3'-ends as described above. RNA (25-200 μ g), isolated from logarithmically growing cells harbouring Clo DF13 (2), was precipitated with approximately 20 pmol DNA probe and washed twice with 70% alcohol. The dried pellet was dissolved in 15-20 μ l S1-hybridization buffer (80% formamide; 40 mM PIPES, pH 6.4; 1 mM EDTA; 0.4 M NaCl). The RNA-DNA mixture was denatured at 85°C during 15 min, after which hybridization was carried out during 16 h at 53°C. The hybridization was stopped by the addition of 10 volumes of cold S1 incubation buffer (0.25 M NaCl, 3 mM NaAc, pH 4.6, 1 mM ZnSO₄, 20 μ g/ml sonicated salmon sperm DNA) with 1000 units/ml S1 nuclease. After 1 h incubation at 37°C the S1-resistant hybrids were precipitated by the addition of ethanol and 1 μ g/100 μ l tRNA. After washing, the pellets were dissolved in 5 μ l 10 mM Tris-HCl pH 7.0, 1 mM EDTA, 0.1% SDS and centrifuged during 5 min at 12,000 x g. 4.5 microliter of the hybrids was mixed with 4.5 μ l 100% formamide (deionized), heated for 5 min at 100°C and loaded on 5% or 8% sequence gels.

RESULTS

The nucleotide sequence of the Clo DF13 DNA region comprising genes K and L

The DNA region between 29% and 50%, comprising the genes K and L, also contains part of the genetic information involved in incompatibility and stability (20). Hakkaart *et al.*, localized the *inca* region between 29% and 38%, whereas the *parB* and *incB* functions were assigned to the DNA region between 45% and 50% (see fig. 1; 20). To get insight in the genetic organization of this part of the plasmid genome, we analyzed the DNA sequence between 29% (BamHI site) and 50% (SmaI site) of the Clo DF13 genome. The strategy applied to analyse the DNA sequence by the chemical degradation method of Maxam and Gilbert is represented in figure 2; both DNA strands have been sequenced. The resulting nucleotide sequence is shown in figure 3.

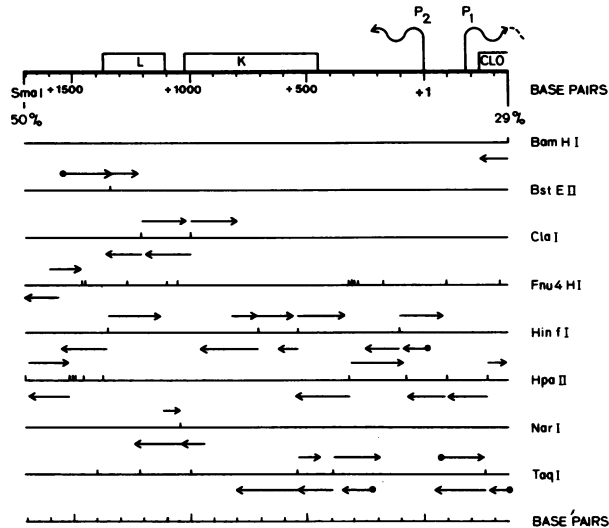


Figure 2. Sequencing strategy for the Clo DF13 DNA region between 29% and 50%. Indicated are: the N-terminal part of the cloacin gene, as well as the cloacin promoter (P_1 ; 10), the genes K, L and promoter P_2 . The direction and the length of the arrows indicate which strands were sequenced as well as the number of nucleotides read from the gels. Usually the label was introduced at the 5'-end, unless indicated by ●.

Codon analysis of the entire DNA sequence between 29% and 50% (fig. 4), revealed the presence of several "open" reading frames of more than 100 nucleotides, in both directions. From previous studies it is known that gene K is transcribed in a clockwise direction (21). Evidence was also presented that gene L is located distal to gene K, probably in the same operon (see fig. 1; 21). From figure 4A it can be concluded that only 2 "open" frames (designated K and L) might represent genes K and L, according to their location, direction of transcription and the predicted molecular weight of the encoded proteins. The molecular weights of proteins K (188 aminoacids; 21,305 D) and L (87 aminoacids; 9,895 D), as predicted from the DNA sequence, fit very well with the molecular weights estimated from SDS-polyacrylamide gelelectrophoresis (21 kD and 10.5 kD respectively; 21). The "open" frames N_1 , N_2 and N_3 might represent structural genes. It is unlikely that the "open frames" N_4 - N_{10} represent structural genes, since they are not preceded by Shine and Dalgarno sequences (35). Furthermore, in that case, transcription should proceed in a direction opposite to that of genes K and L (see also below).

Determination of integration sites of transposons Tn901 and Tn5

In order to confirm the location of genes K and L within the DNA sequence, and to get more insight in the location and nature of incA, incB and parB, we determined the integration sites of several transposon mutants (see table 1). It is known that some of these mutants are defective in the synthesis of protein K (pJN50, 66 and 63) or protein L (pJN56) (21).

The integration sites of pJN63 and pJN50 were mapped by restriction enzyme analysis on 5% polyacrylamide gels, using the restriction enzymes HinfI, BglI and ClaI. All other integration sites were determined by DNA sequence analysis. For this purpose, we made use of restriction sites located within the inverted repeats of Tn5 (39) or Tn901. The inverted repeats of Tn901 are completely homologous with the inverted repeats of Tn3 (P. van den Elzen, unpublished results; 41). As an example, we present the determination of the integration site of Tn901 within plasmid pJN56 (fig. 5). In this case the DdeI restriction sites, located at 16 nucleotides from both ends of Tn901, were labeled at the 5'-end. As can be concluded from the 20% sequence gel, shown in figure 5, a duplication of 5 Clo DF13 nucleotides adjacent to the Tn901 inverted repeats has occurred. A similar 5 nucleotide duplication was observed in case of other Clo DF13::Tn901 mutants and was also reported for integrations of the related transposon Tn3 (40). In case of the Clo DF13::Tn5 insertion mutant pVC22, a 9 nucleotides duplication was found which is consistent with the observation of Auerswald et al. (39).

Based on the finemapping of the Tn901 integration sites within plasmids pJN56, 63, 60, 66 and 52 and their effects on the synthesis of proteins K and L (fig. 4; 21) we infer that the "open" frames K and L indeed encode the corresponding proteins. With respect to plasmid pJN66, we observed that integration of Tn901, in case of this mutant, has occurred proximal to the proposed structural gene K. However analysis of proteins encoded by this mutant in minicells, revealed that pJN66 does not synthesize protein K (21). This observation can be explained by polar effects of this transposon insertion, since transcription of the Tn901 specified β -lactamase gene in this case proceeds in opposite direction of that of gene K (21).

Transcription of the Clo DF13 DNA region comprising genes K and L

To study the transcription of the Clo DF13 DNA region encoding proteins K and L, we analyzed the length and relative quantity of RNAs transcribed from this region by RNA blotting experiments. Furthermore, analysis of the location of promoter and terminator sites, as well as analysis of the

29 %^{5'} -GGATCCGTACCACCCTGATGAGGGCCACCTTTACCAGACGTACCGTTACATTCCCAGTGGCTGACCACCATTATGTTCCAGACCT
^{3'} -CCTAGGCATGGTGGTGGACTACTCCGGTGAAGTGGTCTGCATGGCAAGTGAAGGGTGAACGGACTGGTGGTAAATACAGGGCTGGA
AspThrGlyGlyGlySerSerProGlyGlyLysGlySerThrGlyAsnValAsnGlySerAlaGlnGlyGlyAsnHisGlyLeuGly

GAATTACCCGGACCTGCACCGTCTCCACCCTCATATCGTTTTCTCTTTGACATTTAAACAATAAATTAACCAAAAATACTGTATATACACAGTA
 CTTAATGGGCCCTGGAGTGCAGGAGGTGGTGAATAGCAAAAAGGAGAACTGTAATTTTTTETTTAATTTTTGGTTTTATGACATATATATGTCAT
SerAsnGlyProGlyArgGlyAspGlyGlySerMet

TGATTACTACCTTACTCTCCAGTTTTCCAGGTGTCAATATATGAGATTAGTCTCATCTTCAGGCACAAAAAGCGCCGGAGCCCTGTAAAGCAGATC
 ACTAATAATGGAAAGGAGGCTCAAAAGTCCACAGTTATATCTCTAATCAAGACTAGAAGTCCGTGTTTTTCGCGGGCTCGCGGCACATTCGCTCAG

GCCTCTTTCCCGAAAATTTTCATGGCAACAATTTTATTTGAGATATTCATCACACTACTTTTTATATGGTATAGATGACCTGAAAGCGAAAT
 CGCAGAAAAGGGCTTTAAAGAGTACCGTTGTTTAAAAATAAAGCTCTAAGTACTGTGAATGAAAATAAACCATATCTACATGGACTTCGCTTTA
CI +1 pVC22

GGCATGAGGAGAGGACTCGGGTAAACGTGTCGGGAACTTGAGAGCCCGGAGGCCACTATGACGCACCTGAAACTGACTCCCGTGAAGAAAAGCGC
 CCGTACTCTCTCTGGACGCCAATTTGACAAGCCCTTGGACTCTCGGGCTCCGGTGATAGTGCGTGGACTTTGACTGAGGGGACTCTTTTTCGCG
 +100

CTGGCGGACGCTTTGTCGCTCAGGCACCTGGCCATGGGCTGAAAAGGAAACAACAGCTGCCAAAGCGGTAACAAACAGTGGCGGAAAGGCTGCTTC
 GACCGCTGGGAAACAGCGAGTCCGTGACCGGTACCCGGACTTTCCCTGTTTTGCGACGCTTTCCGCATTTTGTGTCACGCCCTTTCCACGAAG
 +200

Vai → N3
 TGCCATTACAGGACGAACGTGAAGCCAGCGGGCGCAGGAAATGCCTGGGAGAAATACGAGCGGCGGAAACACCTTACCTGGCAGCCGACGCGCC
 CACGGTAAGTCTGCTGCATCTCGGTCGCCCGCTCTTACGACGCCCTTTATGCTCCGCCCTTTGGTGGAAAGTGGACCTGCGCTCGCGGC

N2 → ***
 CCGGTGAAAACGATTTAAGGTTCTACTTTACGGTGTATCCACTTCGGAAGCATTTTAGGGATATGGTATTCGAAAATCATGCTTTCGGCAATCCAAATG
 GGCCACTTTTGCATAATCCAAAGTGAAGTCCACATAGGTGAAGCCTTCTGAAAAATCCCTATACCATATAGCTTTTAGTACGAAAGCCGTTAGGTTAC

GGTTTATACGATTTAAAAGGAAAAACATGGCAAAAAGGACACCTGTAAGTCTCGTTTCTATCATCTTTTCCACTCTGGATGTGATGACAGAAA
 CCAAATATGCTAAATTTTTCCCTTTTTGACTGTTTTCTGTTCAATCATCGAGACAAAAGATAGTGAAGAAAGTGAAGACCTACACTCGGCTTT
pJN66

LysGlyAsnGlyAsnAlaGluSerThrGlnIleThrLysAsnAspIleThrThrGlnLeuLysLeuAspLysThrGlnGlnProAspAspIleGluLys
 AAAGGGAATGCAACGCTGAATCGACACAATCACTAAAACATATAACAACCTCACTTAAGTATGATAAAAGCAACAACCTGACGATATAGAAAAGA
 TTTCCCTACCGTTGCGACTAGCTGTTTAAAGTATTTTGCATATTTGTCAGTGAATCAATCTATTTTTCGTTGTGGACTCCTATATCTTTCT

LysGlyAsnGlyLeuGlnAlaAsnSerThrProGluLysIleLysLysPheAsnGluLeuLysAspAspProLysIleLeuArgGlySerValValSer
 AAATAAAAAGAAATCCAGCCCACTCCACCCTGAAAAATAAAAAAATCAATGAATTAAGAGTACGACCTCAAAATCTCCAGAGGTGGTGTGTTCT
 TTTATTTCTTAGTCCGGTGGAGTGGGACTTTTTATTTTTTAAAGTACTTAATTTCTACTGCTGGGATTTTAGGATGCTCTACGCCAACAAAG
 +700

GluIleLysLysLysLeuProLeuLeuValAspGluAlaThrLeuMetThrAspValSerThrAspGlyGlyThrPheSerThrLysTyrValThrLys
 AGAAATTAAGAAAAACTACCCTATTGGTTGACGAAAGCCACATTAATGACAGATGTATCACTGACGAGGACTACATCTCATACAAAATATGTAACCAA
 TCTTTAATCTTTTTGATGGTGATAACCACTGCTCGGTGAATTAAGTGTCTACATAGTTGACTGCCCTCATGTAAAGATGTTTATACATTGGTTT
 +800

GlyIleSerAlaSerThrMetGluSerAspValThrLysAspAlaMetGlnLysAsnIleLysAsnSerThrCysSerAspAlaArgLeuLysVal
 GGGATTTACGCCAGCAGATGGAAGTGAATGTGGAAAGATGCCATGCAAAAAAACAATAAAATAGCTACTGCTGACGATGCCAGCTAAAAGTAT
 CCTAAAGTCGGTCTGCTACCTTTCACTACACACCTTTCTACGGTACGTTTTTTGATTTTTTATCGATGACGAGACTGCTACGGTCTGATTTTCATA
 +900

PheArgGluLeuPheProGluGlyValIleTyrAsnTyrTyrLeuSerAspLysLeuIleTyrThrTyrLysAlaLeuProSerIleCysSer
 TCAGAGAGCTATTTCGAAAGGTGTCATTTACAAGTACTACTTAAGTGAACAAATTAATCTACACATAACAAGGCCTTCCATCGATTTGTTCCCTAAAGTAT
 AGTCTCGATAAAGGACTTCCACAGTAAGTCTGATGAATCACTGTTTTAATTAGATGTGATGTTGCTGCTGAAGTGAAGTAAACAAGGATTTCA
 +1000

TTACCATAACGTCGCAAGCTGGCGCCCGCCAGCGGGCGCTTTTCTATATCTGTCCATAAAATTCAGAAAAGGAAAGCTAAATGGCTGCATATACCGCGA
 AATGGTATTGACGGTTCGACCGCGGGGTCGCCGCGAAAAGATATAGACAGGATATTAAGTCTTTTCCCTTTCGATTTACCAGCTATATGGTCGCT
 1069 +1100

ThrAlaLysPheLeuGlnSerValSerSerAspGlnThrTrpTrpLysSerArgTyrAsnProGlyAspSerValProHisSerGlyIleTyrArgCys
 AACAGCGAAGTTTTACAAGCGTCAGTAGCAGTCAAAATGGTGGAAAGTACGATATAATCCAGGAGACGCGTTCCTCCATCTGGCATTATCGATGT
 TTGTCGCTCAAAAATGTTTCGACAGTCTGCTAGTTTGTACCACTTCAGTGTATATTAGTGTATGTTGCTGCAAGGGGTGAAGCCGTAATAGCTACA
 +1200

ThrValCysGlyLysLeuIleThrSerAsnGlyAsnAspProPheProGlnAsnHisHisGlnHisSerGlnGlnAlaIleLysTrpGlnLeu
 ACAGTCTCGGAAAAGAGATCACCAGTACCGCAACCGCTTTCCGCGCAAAACACCTACAGCAGCTGACGCAACAAGCAATCAATGGCAACTGA
 TGTCAGACGCCTTTTCTAGTGGTCATTTGCCGTGCTGGGCAAGGGCGCGTTTTGGTGGTATGCTGTGACGCTGTTGCTGCTAGTTTACCGTTGACT
 +1300

IleValArgThrAspThrLysGlyAspArgPheGlyIleLys
 TTGTCAGAACGGACCTAAAGGTGACCGTTTTGGAATCAAATGATTGAGAGCCATCCGGAGCTTTTTTGTACTGAACTCGAAGATATTTACCATAAC
 AACAGTCTGCTGTTGTTCCACTGGCAAGCCCTAGTTTTACTAAGTCTCGGTAGGCCTGAAAAAACAATGGACTTGGAGCTCTATAAATGGTATTG
 +1400

ATCCCTTATCGGTACCGCCAGTAAGGCACCTGGCGGCCTTCCGGGGCTACGGGATAAAACAAGCAACAACCCACCGGAAAAACCGCGGAAACGGCC
 TAGGGAAATAGCCATGGCGGTTCATTCGGTGACCGCCGCAAAGCGCCGATGCCCTATTTTGTTCGTGTTGGGTGGGCCCTTTTGGCGGCCCTTGGCGGG
 KpnI BglI +1500

GGTAAATTTTATCCTAATGCATGCCTATGCACGCAACTGCATAGCATATGTTTTGACGTAAATGGATGCCGTAAGCCCTTTTTTACCGACTTTTGTGCT
 CATTAAAAATAGGATTACGTACGGATACGTGCGTTGACGTATCGTATACAAAACGTCATTTTACCTACGGCATTCGGGAAAAAATGGCTGAAAGACGA
 +1600

TCATCAGCCTGACTGACACGACCATTCTTTTCAGACCGCCCTGCTAACTTCTTCATATACCGCATTAAACCGGG-3' 50 %
 AGTAGTCGGACTGACTGTCTGGTATAAGAAAGTCTGGCGGGGACGATTGAAGAAGTATAETGGCCTAATTGGGCC-5'
 +1700

Figure 3. The nucleotide sequence of the 29%-50% DNA region of Clo DF13. The bases are numbered from the transcription initiation site (+ 1) of promoter P₂. The aminoacids of the cloacin gene, gene K and gene L are presented in Italics. The "open" frames N₁, N₂ and N₃ are indicated by (Met or Val → N) and by their stopcodon (N → ***). Marked are: P₁, cloacin promoter; Pribnow boxes ("10"; 33, 34); RNA polymerase recognition sites ("35"; 33); potential ribosome binding sites according to Shine and Dalgarno (SD; 35); SOS box, presumed *lexA* binding site (SOS; 10); potential interaction sites for cAMP-cAMP Receptor Protein Complex (CI and CII; 10, 36-38), terminator sequence (T; 33); recognition sites for restriction endonucleases BglI and KpnI. The integration sites of transposons Tn5 or Tn901 are indicated by the Clo DF13 nucleotides, duplicated upon insertion (39, 40). Wavy arrows indicate 5'-ends or 3'-ends of transcripts. Straight arrows indicate symmetric elements.

effect of mitomycin-C on initiation of transcription, were performed using the S1-nuclease protection method of Berk and Sharp (32).

RNA blotting experiments

RNA's, isolated from *E. coli* cells, harbouring Clo DF13 were run on a 2.5% agarose gel containing methylmercuric hydroxide. After electrophoresis and transfer of the RNAs to DBM-paper, strips were hybridized with Clo DF13 DNA or DNA fragments (specific activity of 20-30 x 10⁶ dpm/pmol). The individual strips were autoradiographed during 1-2 days. The result of this experiment is shown in figure 6. From the hybridization pattern of Clo DF13 specific RNAs with Clo DF13 DNA fragments, it can be observed that the DNA regions located between 0% and 47% of the plasmid genome (fig. 6, lanes a-d) are transcribed relatively actively. The 47-70% DNA fragment only weakly hybridizes with 2 RNAs (lane e) whereas transcription of the mobility region of Clo DF13 seems to be repressed (lanes f, g). The RNAs encoded by the DNA regions comprising the origin of DNA replication and the bacteriocin operon have been the subject of earlier investigations (2, 7, 9, 10, 42) and will not be discussed in this paper.

It can be observed in figure 5 (compare lanes b, d and e) that the 35%-47% BglI fragment specifically hybridizes with several RNAs varying in length from about 400 till 1400 nucleotides. A comparison of lanes b, d and

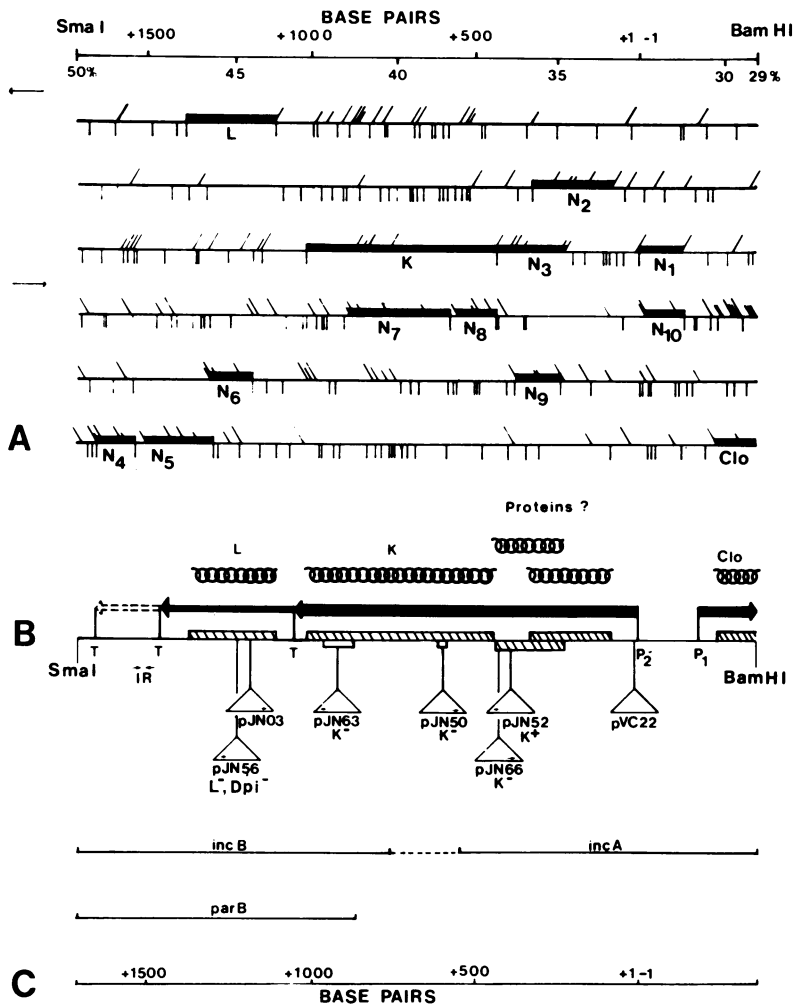


Figure 4. A. Codon analysis of the Clo DF13 DNA region between 29% and 50%. Initiation and termination codons for translation are plotted on each frame, in both directions. Long and short angled-lines represent ATG and GTG startcodons respectively; vertical lines represent translation stopcodons. Indicated are: the genes encoding cloacin DF13 (Clo) and the proteins K and L (K, L); reading frames longer than 100 nucleotides (N₁ through N₁₀).

B. Mapping of integration sites of transposons Tn5 (pVC22) and Tn901 (pJN plasmids) by DNA sequence analysis or restriction enzyme digestions. In the latter case \square indicates the maximum deviation of our localization. Transcripts: as localized by S1 mapping (see fig. 7) are represented by thick arrows. Symbols used are: P₁ and P₂ promoter P₁ and P₂, T, terminator; IR, inverted repeat of 13 nucleotides; Dpi: DNA phage interaction.

C. Location of DNA regions involved in stability (parB) and incompatibility (incA + B), according to Hakkaart *et al.* (18, 20).

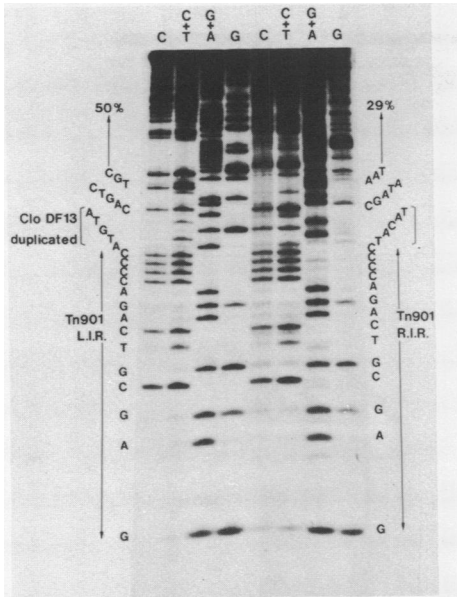


Figure 5. Determination of the integration site of transposon Tn901 within plasmid pJN56 by DNA sequence analysis on a 20% sequence gel. The DNA sequence of restriction fragments containing part of the left and right inverted repeats of Tn901 (LIR and RIR) is read from the DdeI recognition sites present in both inverted repeats, in 5' → 3' direction. The 5 Clo DF13 nucleotides, duplicated upon integration of Tn901, and the direction within the Clo DF13 sequence (29% or 50%) in which the autoradiograph is read, are indicated.

e, after prolonged exposure, revealed that at least the 1200, 1250 and 1400 nucleotide RNAs hybridize weakly with the 30-35% DNA region (not shown). We infer that the coding regions of these RNAs are almost entirely confined to the 35-47% DNA region, but extend beyond the BglI site at 35%. From the RNA blotting

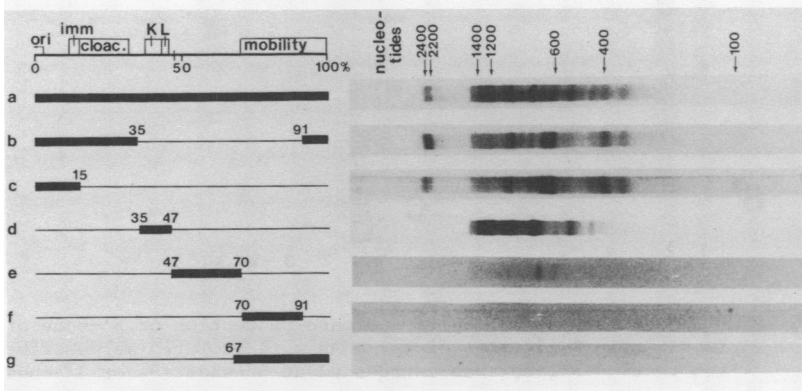


Figure 6. Hybridization patterns of RNA isolated from *E. coli*, harbouring Clo DF13-cop3. Isolation of total cellular RNA, electrophoresis, RNA transfer to DBM-paper and hybridization with ³²P nick-translated DNA were performed as described in Materials and Methods. Different part of the Clo DF13 genome were used as probes, such as total circular Clo DF13 DNA (lane a), Clo DF13 BglI fragments (lanes b, d, e and f).

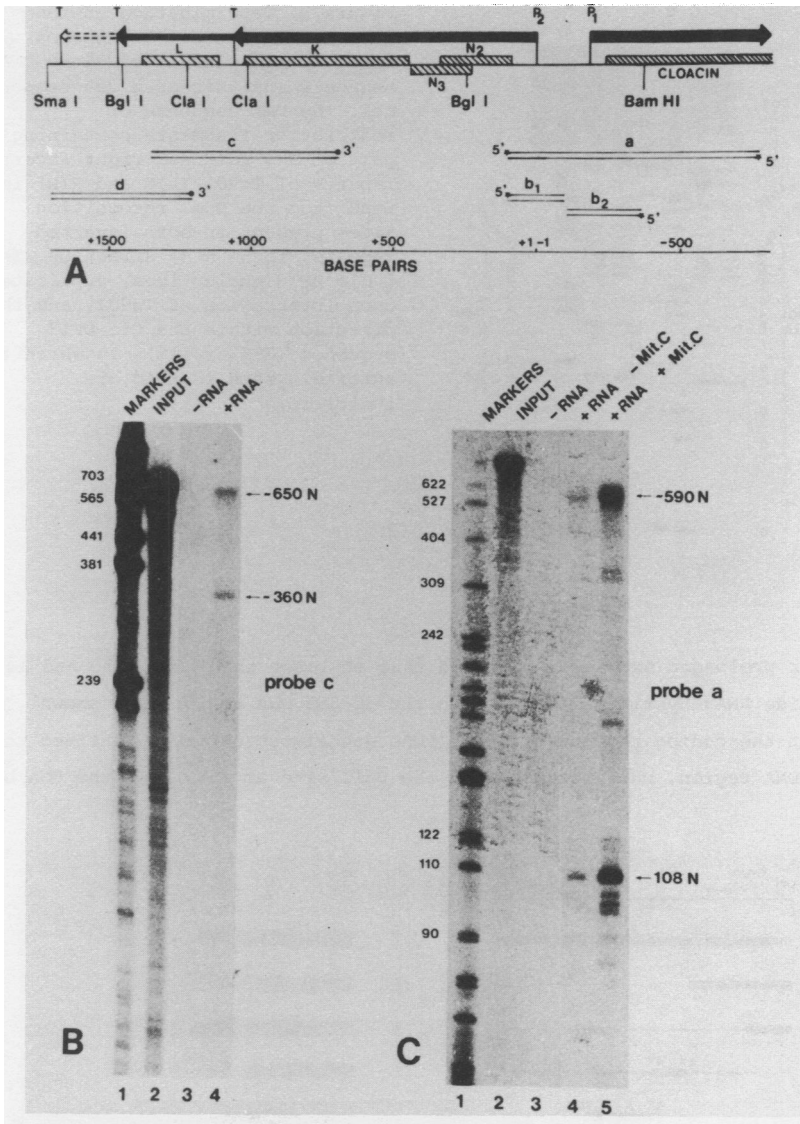


Figure 7. A. Strategy for mapping by S1 nuclease protection of 5'-ends of *in vivo* Clo DF13 transcripts. RNAs were isolated as described in the text. ^{32}P label was introduced at 5'-ends (*) or 3'-ends (●) of the DNA fragments used as probes to map promoter and terminators respectively.

B. Mapping of the terminator between genes K and L (see text). M13 mp8 (43) x *Taq*I was used as marker.

C. Mapping of promoter P₂ and analysis of the effect of mitomycin-C (Mit. C) on transcription from P₁ and P₂ (see text). pBR322 (44) x *Hpa*II was used as marker.

experiment, using the 35-47% DNA region as probe (fig. 6,d), we conclude that no transcript is observed longer than 1400 nucleotides. We presume that some of the smaller transcripts are the result of degradation or premature termination.

Since transcription of genes K and L proceeds in a clockwise direction, these data indicate that initiation of transcription occurs between 35% (BglI site) and 31% (the cloacin promoter). In vitro transcription experiments (unpublished results) pointed to the existence of a promoter (designated P_2), at about 240 nucleotides upstream from the BglI site (at approximately 32.4%; figures 3, 7A).

S₁ protection experiments

The strategy, to localize the 5'-ends and the 3'-ends of in vivo RNAs, isolated from logarithmically growing cells is represented in figure 7A.

In a first approach we localized promoter P_2 by the use of probes a and b_1 which were labeled at the 5' end. An example of an S_1 experiment using probe a, overlapping with RNAs transcribed from P_1 (the cloacin promoter) and, P_2 , is presented in figure 7C. Two major bands of about 590 and 108 nucleotides appear in lanes 4 and 5. These two DNA fragments are protected by RNAs transcribed from promoter P_1 (590 N) and P_2 (108 N) respectively. The minor bands of approximately 350 and 180 nucleotides probably are the result of S_1 nuclease sensitive AT-rich stretches within the DNA sequence of the cloacin operon. Based on the length of the protected DNA fragment we concluded that promoter P_2 is located at 32.5% (fig. 7A). This result was confirmed by an identical experiment using probes b_1 (data not shown). From these experiments and from a similar experiment using probe b_2 we also concluded that no transcriptional activity is present in the region between promoter P_1 and P_2 . Since treatment of cells with mitomycin-C results in enhanced transcription from promoter P_1 (the cloacin promoter), which is due to the recA dependent inactivation of the lexA repressor (10), we also determined the effect of mitomycin-C on transcription from promoter P_2 . For this purpose, we performed the S_1 nuclease protection experiments, shown in figure 7C (lanes 4 and 5), with equal amounts of RNA, isolated from a mitomycin-C induced culture (lane 5) and from a control culture (lane 4). Mitomycin-C was added 60 min prior to RNA isolation. The results, shown in figure 7C, demonstrate that the addition of mitomycin-C induces transcription from the cloacin promoter as reported before (10). Strikingly, transcription from promoter P_2 is also enhanced upon the addition of mitomycin-C. However, analysis of the DNA sequence adjacent to promoter P_2 did not reveal sequences homologous to the so-called SOS boxes (lexA binding sites) observed in the case of lexA repressible genes (45). The presence of such a SOS box,

overlapping with the cloacin promoter (P_1), was reported previously (fig. 3, 10). We were able to exclude the possibility of an overall effect of mitomycin-C on transcription of Clo DF13, since S_1 protection experiments demonstrated that e.g. the quantity of Clo DF13 pre-primer RNA (6,46) is not affected by the addition of mitomycin-C (data not shown). In order to localize transcription termination sites, we made use of DNA fragments labeled at one 3'-end (probes c and d, fig. 7A). From minicell experiments it is known that protein K is synthesized in larger quantities than protein L. (52). DNA sequence analysis of the region between genes K and L revealed a potential terminator structure, resembling ρ -independent types of terminators, located at position + 1050 (see fig. 3;33). To obtain evidence that attenuation of transcription occurs between genes K and L, and to map the terminator involved in this process, we made use of probe c (fig. 7). As can be concluded from figure 7B, lane 4, not only complete protection of probe c occurs, but also partial protection resulting in a 360 nucleotide DNA band. From this result, we conclude that termination partially occurs between genes K and L, approximately at position + 1069, distal to the terminator indicated in figures 3 and 7A. The total protection of probe c also indicates that transcription proceeds beyond this terminator into gene L. To investigate at which site(s) transcription terminates distal to gene L, and whether transcription proceeds beyond the SmaI site at 50%, we performed on S_1 protection experiment using probe d. As could be concluded from the length of the protected fragment, termination of transcription might occur approximately at position 1465. Proximal to position 1465 a terminator-like structure (33) can be observed in the DNA sequence (fig. 3). This termination site is located within the parB region as designated by Hakkaart *et al.* (18). Although several minor bands were observed in this S_1 protection experiment, due to S_1 sensitive AT-rich stretches within this region (see fig. 3), no protection of the entire probe d occurred, indicating that transcription probably does not proceed beyond the SmaI site.

From the results of the S_1 protection experiments presented in this section, we concluded that transcription of the operon comprising genes K and L initiates at promoter P_2 . Termination occurs partially at a terminator located between genes K and L, and within a region between gene L and the SmaI site at 50%.

DISCUSSION

In this paper, we present the nucleotide sequence of the Clo DF13 DNA region between 29% and 50%. Within this DNA sequence, we were able to identify

the genes K and L by transposon insertion mapping. Furthermore, we analyzed the mode of transcription of this DNA region by RNA blotting and S₁ nuclease protection experiments.

From the data obtained sofar, we conclude that:

- a) The genes K and L are located in one operon; transcription of this operon is initiated at 32.5% (promoter P₂), and proceeds in a clockwise direction on the Clo DF13 genetic map.
- b) Attenuation of transcription occurs between genes K and L. Transcription is terminated distal to gene L in a region involved in plasmid maintenance, designated parB, and does not proceed beyond the SmaI site, located at 50%.
- c) The genes K and L, as well as the potential genes N₂ and N₃, possess their own ribosome binding sites (35). The ribosome binding site of gene K shares about 60% homology with the optimal 47 nucleotide ribosome binding site predicted by Scherer et al. (47), and is therefor probably very efficient.

With respect to the regulation of transcription initiated at promoter P₂, three features can be distinguished:

- 1) Promoter P₂ can be activated by mitomycin-C
- 2) A SOS box is not present in the vicinity of this promoter, indicating that lexA protein is not directly involved. However, recently Brandsma et al. (52) described a situation in which a lexA binding site located between two promoters can be functional in both directions. Although promoter P₁ and P₂ are not overlapping, a similar role for the unique lexA binding site between both promoters is possible.
- 3) A potential binding site for cAMP-cAMP Receptor protein can be observed; this sequence (CI in fig. 3) shows a high degree of homology with consensus sequences published before (36-38). This cAMP-CRP binding site could be involved in negative regulation of transcription, as reported for the galP₂ promoter (48).

The elucidation of the nucleotide sequence of the DNA region between 29% and 50% as well as the localization of transposon integration sites within this region enabled us to define more precisely the genetic information involved in incompatibility (incA and incB), stability (parB) as well as the inhibition of the multiplication of DNA phages (20, 21). At this moment there is no definite proof that the function of incA is correlated with a particular site in the plasmid genome or with a plasmid encoded gene product. In this respect, it is striking that integration of TNS₅ within plasmid pVC22, which integration affects incompatibility (incA⁻;20), has occurred around the transcription initiation site of promoter P₂. Since integrations of

transposon Tn5 are highly polar (49), the possibility exists that this insertion has a polar effect on the expression of genetic information located downstream. The incA region is located at about 32% (fig. 4c; 20). If the incA phenotype is correlated with a particular gene product, the "open" frames N_2 and N_3 (fig. 4A) are the most likely candidates. Concerning the region between approximately 45% and 50% (SmaI), we know from previous (21,18) and recent (Hakkaart et al., manuscript submitted) experiments that this region is involved not only in the inhibition of the multiplication of DNA phages but also in plasmid incompatibility (incB) and partitioning (parB). Recently, evidence was obtained that transcription into the parB region is essential for its functioning in plasmid maintenance. We suggest that termination of transcription in the parB region at position 1460 might be a crucial feature in the expression of the parB function. In the DNA region distal to gene L, symmetric elements are present; for instance, we observed a 13 nucleotide inverted repeat, separated by 36 nucleotides, around position 1490 (fig. 3). The role of these symmetric elements in plasmid maintenance will be discussed in a forthcoming publication (Hakkaart et al., manuscript submitted).

Further experiments will be focussed on the role of cAMP-CRP complex, mitomycin-C and/or lexA on the expression of the operon containing genes K and L.

ACKNOWLEDGEMENTS

The authors are very grateful to Aart-Gerrit Jochemsen for his help with the S1 nuclease protection experiments, and to Peter van der Meer, who performed part of the DNA sequence analysis.

REFERENCES

- (1) Tieve, G.A., Stouthamer, A.H., Jansz, H.S., Zandberg, J., and Bruggen, E.E.J. van (1969) *Molec. gen. Genet.* 106, 48-65.
- (2) Elzen, P.J.M. van den, Konings, R.N.H., Veltkamp, E., and Nijkamp, H.J.J. (1980) *J. Bacteriol.* 144, 579-591.
- (3) Veltkamp, E., Barendsen, W., and Nijkamp, H.J.J. (1974) *J. Bacteriol.* 118, 165-174.
- (4) Andreoli, P.M., Brandsma, R.A., Veltkamp, E. and Nijkamp, H.J.J. (1978) *J. Bacteriol.* 135, 612-621.
- (5) Kool, A.J. and Nijkamp, H.J.J. (1974) *J. Bacteriol.* 120, 569-578.
- (6) Stuitje, A.R., Spelt, C.E., Veltkamp, E. and Nijkamp, H.J.J. (1981) *Nature* 290, 264-267.
- (7) Veltkamp, E. and Stuitje, A.R. (1981) *Plasmid* 5, 76-99.
- (8) Veltkamp, E. and Nijkamp, H.J.J. (1976) *Biochim. Biophys. Acta* 425, 356-367.
- (9) Elzen, P.J.M. van den, Gastra, W., Spelt, C.E., Graaf, F.K. de,

- Veltkamp, E. and Nijkamp, H.J.J. (1980) *Nucl. Acids Res.* 8, 4349-4363.
- (10) Elzen, P.J.M. van den, Maat, J., Walters, H.H.B., Veltkamp, E. and Nijkamp, H.J.J. (1982) *Nucl. Acids Res.* 10, 1913-1928.
- (11) Hakkaart, M.J.J., Veltkamp, E., and Nijkamp, H.J.J. (1981). *Molec. gen. Genet.* 183, 318-325.
- (12) Hakkaart, M.J.J., Veltkamp, E., and Nijkamp, H.J.J. (1981) *Molec. gen. Genet.* 1983, 326-332.
- (13) Tiel-Menkveld, G.J. van, Veltkamp, E., and Graaf, F.K. de (1981) *J. Bacteriol.* 146, 41-48.
- (14) Oudega, B., Stegehuis, F., Tiel-Menkveld, G.J. van, and Graaf, F.D. de (1982) *J. Bacteriol.* 150, 1115-1121.
- (15) Pol, H. van de, Veltkamp, E., and Nijkamp, H.J.J. (1978) *Molec. gen. Genet.* 160, 139-149.
- (16) Pol, H. van de (1980) Ph. D. Thesis, Amsterdam, The Netherlands.
- (17) Stuitje, A.R., Veltkamp, E., Elzen, P.J.M. van den, and Nijkamp, H.J.J. (1978) *Nucl. Acids Res.* 5, 1801-1820.
- (18) Hakkaart, M.J.J., Veltkamp, E., and Nijkamp, H.J.J. (1982) *Mol. gen. Genet.* 188, 338-344.
- (19) Novick, R.P., and Hppensteadt, F.C. (1978) *Plasmid* 1, 421-434.
- (20) Hakkaart, M.J.J., Wesseling, J.G., Veltkamp, E., and Nijkamp, H.J.J. (1982) *Molec. gen. Genet.* 186, 531-539.
- (21) Pol, H. van de, Veltkamp, E., Nijkamp, H.J.J. (1980) *Molec. gen. Genet.* 178, 535-540.
- (22) Adler, H.J., Fischer, W.D., Cohen, A., and Hardigree, A.A. van (1966) *Proc. Natl. Acad. Sci. USA* 57, 321-326.
- (23) Greene, P.J., Heyneker, H.J., Bolivar, F., Rodriguez, R.L., Betlach, M.C., Covarrubias, A.A., Backman, K., Russel, D.J., Tait, R., and Boyer, H.W. (1978) *Nucl. Acids Res.* 5, 2373-2380.
- (24) Emlden, J.D.A. van, Veltkamp, E., Stuitje, T., Andreoli, P.M., and Nijkamp, H.J.J. (1978) *Plasmid* 1, 204-217.
- (25) Andreoli, P.M., Overbeeke, N., Veltkamp, E., Emlden, J.D.A. van, Nijkamp, H.J.J. (1978) *Molec. gen. Genet.* 160, 1-11.
- (26) Stuitje, A.R., Veltkamp, E., Maat, J., and Heyneker, H. (1980) *Nucl. Acids Res.* 8, 1459-1473.
- (27) Maxam, A.M. and Gilbert, W. (1980) in *Methods in Enzymology* (Grossman, L. and Moldave, K. eds) Vol. 65, pp. 499-560.
- (28) Maat, J. and Smith, A.J.H. (1978) *Nucl. Acids Res.* 5, 4537-4546.
- (29) Bailey, J.M. and Davidson, N. (1976) *Anal. Biochem.* 70, 75-85.
- (30) Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) *J. Mol. Biol.* 113, 237-251.
- (31) Wahl, G.M., Stern, M., and Stark, G.R. (1979) *Proc. Natl. Acad. Sci. USA* 76, 3683-3688.
- (32) Berk, A.J., and Sharp, P.A. (1977) *Cell* 12, 721-732.
- (33) Rosenberg, M., and Court, D. (1979) *Ann. Rev. Genet.* 13, 319-353.
- (34) Pribnow, D. (1975) *J. Mol. Biol.* 99, 419-443.
- (35) Shine, J., and Dalgarno, L. (1975) *Nature* 254, 34-38.
- (36) Taniguchi, T., O'Neill, M., and Crombrughe, B. de (1979) *Proc. Natl. Acad. Sci. USA* 76, 5090-5094.
- (37) O'Neill, M.C., Amass, K., and Crombrughe, B. de (1981) *Proc. Natl. Acad. Sci. USA* 78, 2213-2217.
- (38) Queen, C., and Rosenberg, M. (1981) *Nucl. Acids Res.* 9, 3365-3377.
- (39) Auerswald, E.-A., Ludwig, G., and Schaller, H. (1980) *Cold Spring Harbor Symposium* 45, 107-113.
- (40) Chen-Pei D. Tu, Cohen, S.N. (1980) *Cell* 19, 151-160.
- (41) Heffron, F., McCarthy, B.J., Ohtsubo, H., and Ohtsubo, E. (1979) *Cell* 18, 1153-1163.
- (42) Elzen, P.J.M. van den, Walters, H.H.B., Veltkamp, E., and Nijkamp, H.J.J.

- (1982) *Nucleic Acids Res.* 11, 2465-2477.
- (43) Messing, J., Crea, R. and Seeburg, P.H. (1981) *Nucleic Acids Res.* 309-321.
- (44) Sutcliffe, J.G. (1979) *Cold Spring Harbor Symposium* 43, 77-90.
- (45) Little, J.W. and Mount, D.W. (1982) *Cell* 29, 11-22.
- (46) Tomizawa, J., and Itoh, T. (1981) *Proc. Natl. Acad. Sci.* 78, 6096-6100.
- (47) Scherer, G.F.E., Walkinshaw, M.D., Arnott, S. and Morré, D.J. (1980) *Nucl. Acids Res.* 8, 3895-3907.
- (48) Busby, S., Aiba, H. and Crombrughe, B. de (1982) *J. Mol. Biol.* 154, 211-227.
- (49) Rothstein, S.J., Jorgensen, R.A., Postle, K., and Reznikoff, W.S. (1980) *Cell* 19, 795-805.
- (50) Austin, S., Ziese, M., and Sternberg, N. (1981) *Cell* 25, 729-736.
- (51) Haess, R.H., Ziese, M., and Sternberg, N. (1982) *Proc. Natl. Acad. Sci. USA* 79, 3398-3402.
- (52) Van de Pol, H. (1980) The bacteriocinogenic plasmid Clo DF13: Mobilization and phage interaction. Academic thesis, Free University, Amsterdam, The Netherlands.
- (53) Brandsma, J.A., Bosch, D., Backendorf, C., and van de Putte, P. (1983) *Nature*, in press.