Volume 14 Number 24 1986 Nucleic Acids Research

# D protein of miniF plasmid acts as a repressor of transcription and as a site-specific resolvase

D.Lane*, R.de Feyter, M.Kennedy[1], S-H.Phua[2] and D.Semon[3]

Department of Cell Biology, University of Auckland, Auckland, New Zealand and [3]Biogen, S.A., Geneva, Switzerland

Received 25 September 1986; Revised and Accepted 24 November 1986

## ABSTRACT

Two activities of the D protein of the miniF plasmid have been found. Divergent promoters in ori-1 ("primary" replicative origin) of miniF are both repressed in cells which produce D protein. The mobilization of plasmids containing the ori-1 region by the F conjugation system is also repressed by D protein. In the former case D appears to act as a transcriptional repressor, whereas in the latter case D protein acts by resolving cointegrates of F and the mobilized plasmid. D protein resolves dimers whose monomer units contain the rfsF sequence needed for recA-independent, site-specific recombination of F. The nucleotide sequence of the D gene was determined. The D gene region contains two oppositely-oriented open reading frames which have the same reading phase and substantially overlap. Transposon insertion mutants were used to show that the gene for D protein occupies the top-strand (left-to-right) open reading frame.

## INTRODUCTION

The region of the E. coli plasmid F required for stable maintenance, called miniF, codes for seven known polypeptides. The functions of six of these are known: A and B are needed for partition, C for control of the pif operon and for replication from ori-1, E for replication from both ori-1 and ori-2, and H and G for preventing growth of plasmid-negative segregants (1,2). The seventh, D protein, has been reported to function in replication from ori-1 (3) and to improve the stability of miniF replicating from ori-2 (4). Although the evidence for these roles of D protein is not compelling, it is difficult to believe that D protein is not involved in some way in a plasmid maintenance activity since its gene is surrounded by other plasmid maintenance determinants. We have found that it plays a role in two processes close to or within the ori-1 sequence.

Ori-1 is the replication origin identified by electron microscopical analysis of autonomous miniF replicative intermediates (5). However it is dispensible, and plasmids lacking it replicate with equal efficiency from a second origin, ori-2 (6,7). Recent work has demonstrated that ori-1 promotes the replication of DNA polymerase I-dependent vectors in polA mutant cells (8,9). In a series of experiments conducted

Table 1 Bacterial strains and plasmids

| Strain | Description | Source/Reference |
|---|---|---|
| N100 | galK2, pro, recA3, rpsL | B. Rak |
| PB1132 | galT12, lacZ32, thr, recA1 | This laboratory |
| PB1150 | Δgal, his, recA1, rpsL | This laboratory |
| Plasmid | | |
| pKO100 | $rep_{322}$, $bla^+$, galK; promoter-cloning vector | (16) |
| pLG338 | $rep_{pSC101}$, $tet^r$, $kan^r$; medium copy vector | (32) |
| pNZ829 | pUC8 + 41.85-43.14Fkb, ori-$1^+$ | (10) |
| pNZ115 | pACYC184 + miniF (40.3-49.4Fkb), $tet^r$ | (33) |
| pNZ116 | pBR322 + miniF (40.3-49.4Fkb), $bla^+$, $tet^r$ | (34) |
| pNZ655 | pNZ116 deleted for 40.45-43.0Fkb (BamHI) | (29) |
| pNZ124 | pACYC184 + 40.45-43.0Fkb, $cam^r$, $C^+$; dimer when in strain N100 | (9) |
| pNZ122 | pACYC177 + 45.15-46.0Fkb, $bla^+$ $E^+$ | This laboratory |
| pPB052 | pNZ116::Tn5 (at 42.85Fkb), $G^+$ $D^+$ | " " |
| pPB056 | " 43.39 , $G^-$ $D^+$ | " " |
| pPB058 | " 43.48 , $G^-$ $D^-$ | " " |
| pPB067 | " 43.74 , $G^+$ $D^-$ | " " |
| pNZ352 | pNZ115::Tn5 (at 45.70Fkb), $G^+$ $D^+$ $E^-$ | (29,22) |
| pNZ361 | pNZ115::Tn3 (at 43.37Fkb), $G^+$ $D^+$ | " |
| pNZ362 | " 43.43 , $G^-$ $D^+$ | " |
| pNZ363 | " 43.52 , $G^-$ $D^-$ | " |
| pNZ364 | " 44.27 , $G^+$ $D^-$ | " |
| pNZ365 | " 44.32 , $G^+$ $D^-$ | " |
| pNZ366 | " 44.50 , $G^+$ $D^+$ | " |
| pNZ825 | pLG338 + 43.5-44.9Fkb, $tet^r$, $D^+$: D gene inserted into kan gene and controlled by the kan promoter of Tn903 | This laboratory |
| pBK57 | $rep_{pSC101}$ + 43.0-49.4Fkb, $kan^r$, $(A, B, E, H, G)^+$ | (11) |
| F8-1 | F'-$gal^+$ | (15) |
| F8-20, -21, -24 | $rep_{ts}$ mutants of F8-1 | " |
| F8-4-8, -12 | $rep_{ts}$ mutants of F8-4 (F'-$gal^+$) | " |

in this laboratory it was found that the ability of ori-1 to rescue vector plasmids depends on a 190bp sequence between the NcoI (42.88Fkb) and HpaI (43.07Fkb) sites, and on the C, E and G proteins of miniF (10).

This ori-1 region lies next to a sequence required for a recA-independent

site-specific recombination activity which resolves F or miniF dimers and permits F-mediated mobilization of certain plasmids (11,12,13). The cross-over site has been mapped to a short interval between the arms of an inverted repeat, rfsF, and the sequence next to this region is also involved in enhancing recombination (13). The dimer resolution effect of the recombination reaction is presumed to promote stable plasmid inheritance in a manner similar to that proposed for the loxP/cre system of lysogenic P1 phage (14).

While engaged in two distinct lines of enquiry, we made observations which suggested that D protein can act as a transcriptional repressor in ori-1 and as a resolvase at rfsF, and in this paper we present the evidence for these conclusions. The resolvase activity of D protein has also been observed in the course of an extensive study of the site-specific recombination reaction carried out by M. Malamy's group (M. Malamy, personal communication).

## MATERIALS AND METHODS

E. coli K12 strains, and plasmids used in this work, are described in Table 1. The sites of Tn3 and Tn5 insertions in the ccd region of miniF are shown in Fig.1. Media and growth conditions were standard, and have been described previously (9,15). All experiments involving recombinant DNA, as well as sequencing by the chain termination technique, used the standard methodology employed previously in our laboratory (9).

Promoter strength was measured by assaying galactokinase activity in strains in which galK expression is controlled by the promoters in question, as described by McKenney et al (ref.6). Site-specific recombination activity was measured by determining the frequency of mobilization of rfsF$^+$ plasmids to recipient cells by F-prime plasmids, as described previously (9,12). Minicell analysis was carried out by standard procedures (17,18).

Transcriptional start sites were determined by the S1-nuclease mapping procedure of Berk and Sharp (ref.19). Radioactive probes were prepared by using single-stranded templates of the cloned miniF promoter fragments inserted in M13 vectors and the standard 17mer sequencing primer to direct DNA-polymerase I (Klenow)-catalysed synthesis of complementary strands: the radioactive strands were isolated by restriction endonuclease digestion, heat-denaturation, gel electrophoresis and electroelution. The probes were annealed with total bacterial RNA which had been isolated by the hot phenol method (20), and the nucleic acids were then subjected to S1-nuclease digestion, polyacrylamide-urea gel electrophoresis, and radioautography.

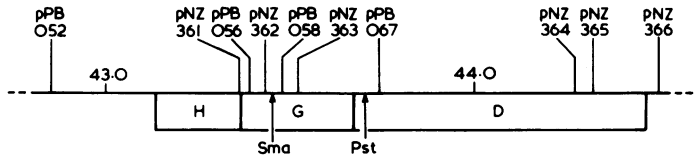


Fig.1: Sites of insertion of Tn3 (pNZ prefix) or Tn5 (pPB prefix) in plasmids listed in Table 1. The Tn3 insertions were originally isolated in miniF (22), and the miniF::Tn3 EcoRI fragments were then cloned in pACYC184 (29). Numbers on the line are F kilobase co-ordinates.

## RESULTS

### Promoters in the ori-1 region

The 9.1kb miniF fragment (f5) was digested with Sau3A and the fragments were inserted into the BamH1 site of the vector pKO100. The BamH1 site in this vector precedes a galK structural gene which is not expressed due to the absence of a promoter. Plasmids containing miniF promoters were obtained by selecting $Ap^{R}Gal^{+}$ transformants of the galK strain, N100, on MacConkey-galactose agar. The promoter fragments were mapped approximately by hybridization analysis with radioactive miniF subfragments, and then precisely by restriction analysis and, after subcloning in M13 vectors, by nucleotide sequencing. One promoter, $P_{705}$, was found to lie in the minimal ori-1 region. It is oriented so that transcription proceeds towards the C gene of miniF (Fig.2).

In an attempt to detect a promoter with the opposite orientation, the NcoI-HpaI fragment which contains ori-1 was inserted into the SmaI site of pKO100, and plasmids containing miniF promoters were selected as before. Of ten plasmids containing the NcoI-HpaI fragment, one had its insert oriented so that galK expression depends on transcription away from the C gene end of miniF. The promoter responsible was termed $P_{723}$.

The positions of the divergent $P_{705}$ and $P_{723}$ promoters were determined by S1 nuclease analysis. Total RNAs prepared from N100/pNZ705 or N100/pNZ723 were annealed with $^{32}P$-labelled single-stranded probe DNAs, and the nucleic acids were digested with S1 nuclease and analysed by polyacrylamide-urea gel electrophoresis and radioautography (Fig.3). In the case of $P_{705}$ three adjacent bands were observed with the central one being the strongest, suggesting that the promoter is located just upstream from this region. The start site corresponding to $P_{723}$ is less clear-cut because an extended set of bands is evident. Nevertheless both transcription start site regions are located a few bases 3' to promoter-like sequences, and these sequences are identified by the TARGSEARCH program (21) as the only ones in the ori-1 region with a score indicating significant homology with the consensus promoter ($P_{705}$, 45%;

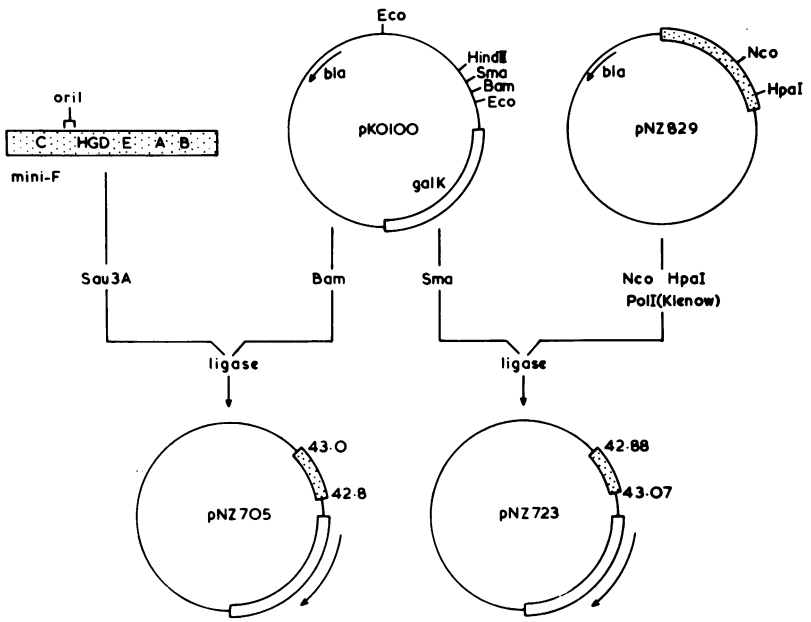


Fig.2: Isolation of promoters from the ori-1 region.

$P_{723}$ 46%). These scores imply weak promoter activity, as also suggested by the faint protection of probe DNA seen in the radioautographs of Fig.3. This conclusion is confirmed by the results of assays of galactokinase activity (below).

The sequence of the promoter region is shown at the bottom of Fig.3. The -10 and -35 regions are separated by 16bp ($P_{705}$) and 20bp ($P_{723}$), and show only moderate homology with the canonical optimum promoter sequence. The transcriptional start site of $P_{705}$ is much closer to the -10 region (2bp) than is the case with most promoters An alternative $P_{705}$ promoter at a greater distance from the transcription start site is also shown: however the -10 sequence of this promoter lacks the "invariant T" and is separated from the start site by 12bp.

Regulation of the ori-1 promoters

To discover whether the activity of $P_{705}$ and $P_{723}$ is affected by other miniF genes, we introduced plasmids containing various portions of miniF DNA into N100 strains carrying pNZ705 and pNZ723, and then assayed galactokinase activities in extracts of log-phase cells. The results (Table 2) show that these promoters are of about equal strength, and direct the transcription of galK at rates 4 to 5-fold higher than the background level. This indicates very weak promoter activity. Under the same conditions we found the following promoters, described by McKenney et al (ref.16)

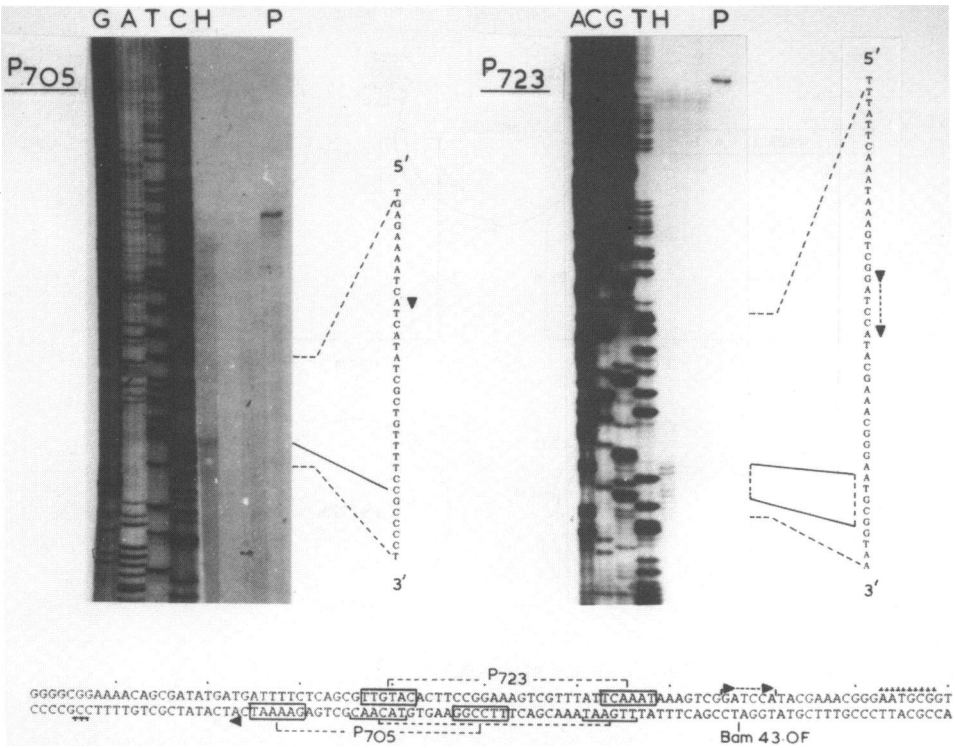


Fig.3: S1-nuclease mapping of transcriptional start sites in ori-1. Letters above the radioautographs: G,A,T,C - dideoxynucleotide sequencing reactions for the relevant ori-1 strand template; H - hybrid produced by annealing total cellular RNA with radioactive probe (pNZ705 insert, Fig.2, cleaved from M13 template with HindIII; pNZ723 insert cleaved from template with EcoRI) and digestion with S1-nuclease (Boehringer: 300U, 30 min, 37°C); P - probe fragment. Start-sites are shown as arrowheads on sequences to the right, positioned 17 bases above the hybrid band to correct for removal of the 17 base primer. The -35 and -10 regions of the promoters are shown boxed in the portion of ori-1 sequence below, along with start sites (▶) and hybrid positions (⊥⊤). For $P_{705}$ a possible alternative promoter is shown with -35 and -10 regions underlined.

as strong, moderate and weak, to yield specific activities of: $P_{gal}$, 254; $P_{c17}$, 180; and $P_{482}$, 67 units per ml of cells at $A_{650}$ = 1. The data in Table 2 show that transcription from $P_{705}$ and $P_{723}$ is repressed when miniF DNA, represented by pNZ115 or pBK57, is also present. The plasmids pNZ124 and pNZ122, which contain the C and E genes respectively, have no effect on the activity of these promoters, but pNZ825, which contains the D gene downstream from the neomycin phosphotransferase promoter of Tn903, strongly represses promoter activity.

The complete sequence of the D gene region, given later in the paper,

Table 2

| Coresident plasmid - phenotype: | C | H | G | D | E | A | B | Galactokinase sp. act. (U/ml cells at $A_{650}$ = 1) in cells containing: pNZ705 | pNZ723 |
|---|---|---|---|---|---|---|---|---|---|
| None | | | | | | | | 21.7 | 24.7 |
| pNZ115 | + | + | + | + | + | + | + | 5.3 | 10.3 |
| pBK57 | + | + | + | + | + | + | + | < 0.1 | |
| pLG338 | | | | | | | | 25.1 | 21.9 |
| pNZ124 | + | | | | | | | 23.6 | 25.7 |
| pNZ122 | | | | | + | | | 20.5 | |
| pNZ825 | | | | + | | | | 0.5 | < 0.1 |
| pNZ352 | + | + | + | + | - | + | + | 9.4 | |
| pNZ364 | | + | + | - | + | + | + | 35.2 | |
| pNZ361 | | + | + | + | + | + | + | 3.4 | |
| pNZ362 | | + | - | + | + | + | + | 6.2 | |
| pNZ363 | | + | - | - | + | + | + | 39.5 | |
| pNZ366 | | + | + | + | + | + | + | 8.4 | |

Galactokinase assay of activity of promoters $P_{705}$ and $P_{723}$. N100 strains carrying pNZ705 and pNZ723 were transformed with the plasmids indicated on the left. Transformants were grown to log-phase, and extracts were prepared and assayed for galactokinase. The unit is that defined by McKenney et al (ref.16). GalK specific activity of N100 containing pKO100 (4.9 u/$A_{650}$) has been subtracted to yield the figures shown which are the averages of at least three independent assays (average standard deviation = 0.34 x sp.act. figure). MiniF genes in the coresident plasmid are given as + = wild-type, - = mutant, blank = absent.

revealed that the D open reading frame in the top strand of miniF is mirrored in the bottom strand by an open reading frame of similar length, which we term D'. Therefore it was necessary to determine which of the potential coding sequences is translated into protein and thus responsible for the transcriptional repressor phenotype we had found. A series of miniF mutants having mapped insertions of Tn3 (22) were exploited for this purpose. The miniF::Tn3 EcoRI fragments were recloned in the pACYC184 vector (see Fig.1 and Table 1), and the polypeptide coding properties were examined by minicell analysis. The results (Fig.4) showed that insertion of Tn3 at 44.27 or 44.32 Fkb, near the 3' end of the top-strand coding sequence, caused the disappearance of the wild-type 27 kilodalton polypeptide and the appearance of 22kd and 24kd truncated peptides respectively, thus suggesting that most of the coding sequence precedes the insertion sites in these mutants and so lies leftward of them, in the top strand. Moreover, an insertion in the G gene,

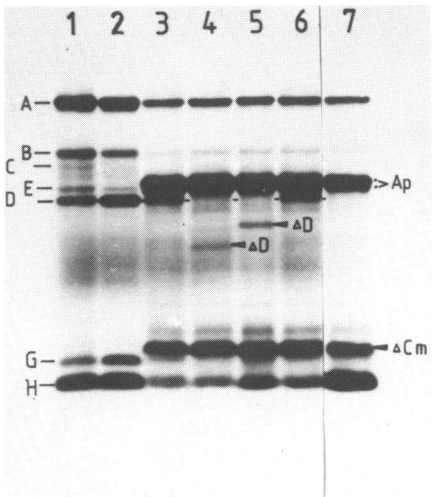


Fig.4: Minicell analysis of D gene expression. Polypeptides in minicells purified from the P678-54 strain containing: 1 - pNZ116, 2 - pNZ655, 3 - pNZ361, 4 - pNZ364, 5 - pNZ365, 6 - pNZ366, 7 - pNZ363: were labelled with $^{35}$S-methionine and subjected to SDS-PAGE and radioautography. A-H - positions of miniF polypeptides; ΔD - truncated D peptides; Ap - precursor and mature β-lactamase; ΔCm - truncated chloramphenicol transacetylase.

at 43.52 Fkb in pNZ363, results in failure to produce any peptide from the D gene region, presumably as a result of polarity (see Discussion). The simplest interpretation is that the wild-type 27kd polypeptide made by minicells is the product of the D gene (e.g. pNZ352) which is read left-to-right, i.e. as top strand sequence. This conclusion is consistent with the results of a separate maxicell analysis of proteins produced by Tn*5* insertion mutants (R. de Feyter, unpublished results). It is notable in these experiments that insertions in the H-G-D region frequently cause quantitative changes in the expression of the H and G (*ccd*) genes which are not easily explained. Reduced amounts of H and G proteins are evident for the D$^-$ mutants in Fig.4. We shall not be concerned with this question here.

The ability of a number of Tn*3* and Tn*5* insertion mutants to repress $P_{705}$ and $P_{723}$ promoter activity was tested (Table 2). Those plasmids which carried an intact, expressed D gene were capable of repression, whereas those with insertions in the D gene (pNZ364) were not. Most informatively, plasmids such as pNZ363 with insertions in the 3'-end of G gene, to the left of D, failed to repress, whereas those with insertions further leftward (pNZ361, pNZ362) or on the right of the D gene region (pNZ366) repressed $P_{705}$ as effectively as wild-type miniF DNA. These results are consistent only with a top strand reading frame as a source of repressor, and confirm

Table 3

| | Mobilization frequency ($Cm^RStr^R$ recipients per $Gal^+Str^R$ recipient, $x10^5$) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Experiment 1 | Experiment 2 | | | | | | Experiment 3 |
| | | +pLG338 | +pNZ278 | +pNZ825 | +pNZ858 | -pNZ825 | +pNZ825 | |
| F8 | 2.1 | 1.6 | 0.62 | 0.01 | 0.81 | 0.76 | < 0.01 | |
| F8-20 | 3.5 | | | | | | | |
| F8-21 | 2.4 | | | | | | | |
| F8-24 | 4.5 | | | | | | | |
| F8-4-7 | 16.0 | - | 24 | 16 | 24 | 2.3 | 2.3 | |
| F8-4-12 | 3.1 | | | | | | | |

Mobilization of rfsF$^+$ vector (pNZ124) by F'-gal$^+$. PB1132 strains carrying pNZ124 and each of the F-primes listed were mated with PB1150, with selection for $Gal^+Str^r$ (F'gal$^+$ transfer) and $Cm^RStr^r$ (pNZ124 moblilization) exconjugants. In experiments 2 and 3 an additional plasmid, as indicated at the top of each column, was present in each donor. Mobilization frequencies vary between experiments because of differences in the fertility of donor cultures caused by minor alterations in growth history.

the identity of the repressor as the 27kd polypeptide D made by minicells containing miniF DNA.

Effect of miniF genes on site-specific recombination

The formation of cointegrates by site-specific recombination is readily measured by determining the frequency with which vector plasmids carrying the recombination site are mobilized into recipient cells by F conjugation. In view of the involvement of G protein in ori-1 activity (10), we thought that F mutants lacking G protein might also show altered mobilization ability. We tested this possibility by measuring the frequency with which pNZ124 (rfsF$^+$, $Cm^R$) is transferred to recipient cells by a number of F rep-ts plasmids. Each of these mutant F-primes is expected to be $G^-$, i.e. $Ccd^-$, because the division of cells containing them is not affected by the cessation of F replication during growth at 42°C (23,24). We have demonstrated this to be the case for one plasmid, F8-21 (S. Francis, unpublished data). The results (Table 3 experiment 1) show that most of the plasmids mobilized pNZ124 with about the same frequency as wild-type F'-gal$^+$ (F8). The one exception, F8-4-7, showed about an eight-fold higher mobilization frequency than wild-type F8. Clearly, mutations in the G gene do not in general have a significant effect on site-specific recombination. However the high mobilization frequency of F8-4-7 suggested that some function, mutant in this plasmid, is needed to control the recombination reaction. Because the F rep-ts plasmids were selected from a nitrosoguanidine-mutagenized

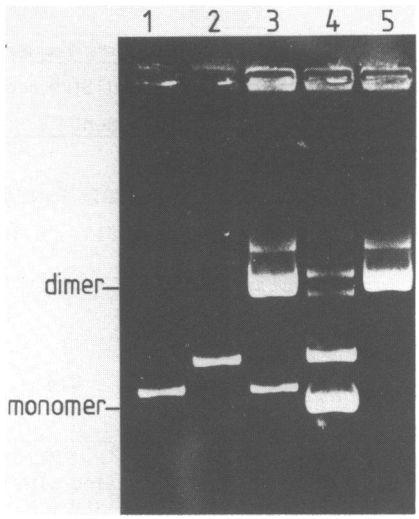


Fig.5: Agarose gel electrophoresis of plasmids in cleared lysates from N100 (recA) strains containing: 1 - pLG338, 2 - pNZ825 ($D^+$), 3 - pNZ124 + pLG338, 4 - pNZ124 + pNZ825, 5 - pNZ124

population to have mutations in the rep region and, unwittingly, the ccd region, it seemed possible that a sequence in this region might carry the mutation responsible for the raised mobilization frequency.

Therefore we tried to identify the presumed recombination control determinant by complementation of the F8-4-7 mutant allele with discrete miniF genes. Plasmids consisting of pLG338 with the G and H genes (pNZ858), the D gene (pNZ825), and the E gene (pNZ278) were introduced into mobilizing strains carrying F8-4-7 and the wild-type control plasmid, F8, and the frequencies of pNZ124 mobilization were measured. The results (Table 3 experiments 2 and 3) showed that none of the cloned miniF genes affected mobilization by F8-4-7. But surprisingly, mobilization by F8 was drastically reduced in the presence of pNZ825, which carries the D gene.

This result could be explained if D protein normally acts to promote resolution of cointegrates; thus the cloned D gene would resolve F8/pNZ124 cointegrates so rapidly that mobilization is prevented. The lack of effect of pNZ825 on mobilization by F8-4-7 suggests that this F-prime may have a mutation in the rfsF site.

D protein promotes resolution at rfsF

To test the tentative conclusion arrived at above, that the effect of D protein on mobilization was by resolution of cointegrates, we made use of a dimer of pNZ124 which is maintained stably in the strain N100. The presence of D protein

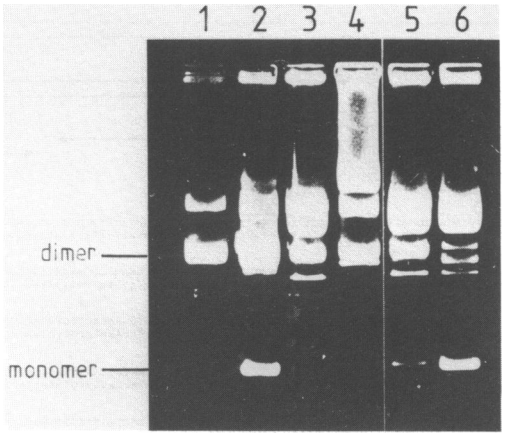


Fig.6: Agarose gel electrophoresis of plasmids in cleared lysates from N100 (recA) strains containing pNZ124 plus: 1 - no other plasmid, 2 - pNZ116, 3 - pPB067, 4 - pPB058, 5 - pPB056, 6 - pPB052. In lane 2, pNZ116 (13.8kb) has run with nearly the same mobility as pNZ124 dimer (12.6kb). In lane 3, 5 and 6 the major band is pNZ116::Tn5 monomer. Dimer forms of the added plasmids are present in lanes 2 and 4, and a tetramer form of pNZ124 is evident in lane 1. Mobility of some species has been distorted by overloading which was necessary to show the nonomer form in lane 5 while keeping sample size constant.

should lead to conversion of the dimer to monomer pNZ124 units. Plasmid DNA was extracted from strains carrying pNZ124 and either pNZ825 or its parent vector, pLG338, and analyzed by electrophoresis in agarose gels. The presence of pNZ825 does indeed lead to conversion of dimer to monomer (Fig.5), thus confirming that D protein is itself the resolvase or, at least, that it stimulates the resolvase activity of another enzyme. In contrast, pNZ124 dimers were not resolved in the presence of the cloned E gene (pNZ278) or H and G genes (pNZ858) (results not shown).

We also examined resolvase expression by a number of plasmids having Tn5 inserted in and around the D gene. The results in Fig.6 show that monomer pNZ124 is generated by wild-type miniF (pNZ116) but not by mutants having Tn5 inserted in the D gene (pPB067) or in the C-terminal portion of the G gene (pPB058). In contrast, insertion in the N-terminal portion of the G gene (pPB056) allows some resolution to proceed, and insertion leftward of the major HGD operon promoter (pPB052) allows resolution at wild type levels. It is notable that pPB058 which fails to make D protein is itself found predominantly in the dimer form. The fact that D gene-proximal insertions are polar on resolution activity strongly implies that the resolvase is the product of a gene transcribed rightwards, i.e. D protein itself.

Sequence of the D gene

Because it appeared likely that D protein binds to DNA to effect its

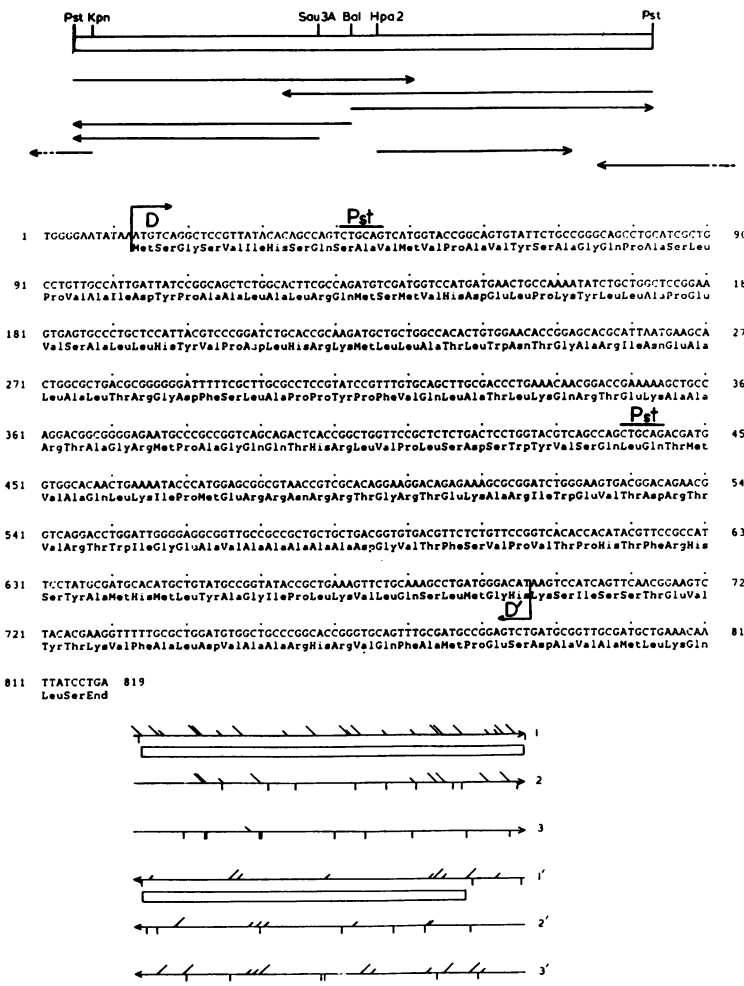


Fig.7: Sequence of D gene. The extent of sequence obtained from various single-stranded templates is shown at the top. The new sequence obtained is that between the two Pst sites indicated; the flanking sequence is included for completeness. D and D' indicate the start of the long open reading frames shown in the analysis of start (↘ATG, ↘GTG) and stop (⊤) signals below.

repressor and resolvase activities, we determined the nucleotide sequence of the D gene in order to search for the DNA-binding domains which are found as helix-turn-helix structures in well-studied DNA binding proteins. The sequences of the N-terminal 33bp and the C-terminal 380bp had been determined previously (25,26).

The Pst fragment (43.7-44.1Fkb) was inserted in both orientations in the sequencing vector M13 mp8, and the resulting clones were used to obtain the bulk

| | |
|---|---|
| rfsF-L | A G A T C C G a A A |
| rfsF-R | A G A T C C G t A A |
| $P_{705}$-sense | c t t T C C G g A A |
| $P_{705}$-antisense | A c t T C C G g A A |
| $P_{723}$ | g G A T C C G a c t |

Fig.8: Sequence homologies in the rfsF and ori-1 regions, representing potential sites of interaction with D protein.

of the sequence data. Subclones were then prepared to refine and confirm the sequence, which is shown in Fig.7 together with an analysis of the translational start and stop signals in each reading frame. Two substantial open reading frames, one on each strand, are evident. They occupy the same reading phase, encode polypeptides of 29.6 and 26.4 kdaltons (top and bottom strand sequences respectively), and are preceded by sequences sufficiently purine-rich to serve as ribosome-binding sites. Nevertheless only for the top strand sequence is there evidence that the open reading frame is expressed as a functional polypeptide (D protein): because no evidence for the existence of a D' protein exists at present we shall not consider further the bottom strand open reading frame.

The deduced amino acid sequence gives a monomer molecular weight for the D protein of 29,627 daltons, slightly higher than that deduced from electrophoretic mobility data. Our examination of possible secondary structure components of the protein showed that although there are regions of substantial α-helix content none of these show strong homology at the amino acid sequence level with the helix-turn-helix domains common to most DNA binding proteins so far studied. The regions of the D protein which are important for repressor and resolvase activities remain a subject for further study.

## DISCUSSION

The complete D gene sequence reported here has revealed an overlapping, complementary open reading frame on the bottom strand of miniF. However, in addition to the evidence presented in this paper, there were prior reasons for believing that the D polypeptide is encoded by the top strand. Komai et al (ref.27) showed that minicells containing miniF DNA deleted rightwards of the 43.7 or 44.1Fkb co-ordinates made fused or truncated peptides, whereas leftward deletions from this point resulted in no trace of either D protein or derived peptides. Bex et al (ref.28) observed that insertion of Tn5 at 43.5Fkb decreased the amount of D protein made in minicells whereas an insertion near the H/G gene border, at 43.2kb, did not, and suggested on this basis that D is transcribed from a promoter lying between these two co-ordinates. We have found that the ability of our G::Tn3/Tn5 mutants to make

D protein is correlated with the position of the insertion: insertions rightward of 43.50 Fkb abolish D synthesis, insertions to the left of this point do not (Fig.4, and ref.29). It appears that D gene is normally transcribed from a promoter in this region, and probably also from the ccd promoter.

What these studies demonstrate is that it is the rightward-oriented open reading frame (D) that encodes the 26-29kd proteins seen in minicells. They do not prove that the biological activities which we have described in this paper result from the activity of this protein. We have shown, however, that these activities disappear when insertions 5' to the D and 3' to the D' reading frames are present. Because it is simpler to explain the effect of the insertions on the basis of polarity on downstream gene expression than as due to retroregulation or some other mechanism affecting upstream gene expression, we take these results as evidence that it is the product of the D gene (i.e. D protein), rather than of the D' "gene", that acts as repressor and resolvase.

Until now D protein has been the only miniF polypeptide for which no function had been found. The claim that it is needed for replication from ori-1 was based on mapping to within the D gene of Tn5 insertions which abolish ori-1 activity (3). From the description of the mapping strategy given we suspect that the insertion sites were not mapped with sufficient precision, for we have shown, using a collection of mutants with well-mapped insertions and deletions, that the nearby G gene, but not the D gene, is required for ori-1 activity (Caughey et al, in preparation). This suspicion has been confirmed by mapping of the Tn5 insertions carried out by M. Malamy (personal communication). Nevertheless the results reported in this paper show that D protein does interact with ori-1 DNA, presumably by direct binding, to repress transcription at divergent promoters. The function of this transcriptional repression is unknown, and it is not yet even clear whether the transcripts themselves are important in origin activity. If D protein does participate in replication, its role is inessential or can be played by another gene product. The galactokinase assay data of Table 2 show that the D protein produced by a plasmid (pBK57) with a 3-4 fold higher copy number than F is sufficient to repress severely the activity of $P_{705}$ on a high copy number vector, and that the levels of D protein achieved by expression from the Tn903 $P_{neo}$ promoter in pNZ825 completely abolish $P_{705}$ and $P_{723}$ activity. These observations, coupled with the inherent weakness of the promoters, make it likely that in wild-type F the activity of these ori-1 promoters is very low.

The observation that D protein promotes the resolution of cointegrates or dimers whose components each carry an rfsF site suggests a role for this protein in maintaining stable inheritance of miniF. The formation of dimers by homologous recombination would occasionally lead to segregation of plasmid-free cells. Provided

the monomers are reformed the partition apparatus will be presented with an adequate number of copies to ensure distribution of at least one to each daughter cell at division. We propose that D gene has this function. It is highly likely that the reaction which resolves miniF dimers takes place at the same site as that used for resolution of F:pBR322 cointegrates (rfsF; ref.30). The rfsF sequence consists of an imperfect inverted repeat of 10bp which flanks an 8bp spacer region containing the crossover site. Thus the structure of this region is similar to the loxP site of phage P1 where the cre protein acts to form and to resolve multimers (31). Whether D protein drives the rfsF reaction in both directions in an energy-independent manner, as Cre does with loxP, remains to be seen.

The rfsF repeat sequences are shown in Fig.8. It is striking that three homologous sequences are found in the region of the divergent promoters in ori-1, with the obvious implication that D binds at these sequences to repress transcription. We have not yet observed any effect of D protein, or its absence, on ori-1 replication, and the possibility that the putative D-binding sequences in ori-1 have functional significance must be tested by more direct experiments.

**ACKNOWLEDGEMENTS**

We thank Carolyn Aitken for capable technical assistance, Dave Saul for help with computerized sequence handling, and Julian Davies for the hospitality shown to D.L. at Biogen, S.A. where the sequencing work was begun. We also thank Michael Malamy for communicating his results to us before publication.

This research was supported by grants from the Medical Research Council of N.Z. and the Auckland University Research Grants Committee.

*To whom correspondence should be addressed

Present addresses: [1]Clinical School, University of Otago, Christchurch, New Zealand and [2]Department of Biology, McGill University, Montreal, PQ, Canada

**REFERENCES**

1. Kline, B.C. (1985) Plasmid 14, 1-16
2. Jaffe,A., Ogura,T. and Hiraga,S. (1985) J. Bacteriol. 163, 841-849
3. Tanimoto, K. and Iino, T. (1985) Mol. Gen. Genet. 198, 358-359
4. Ebbers, J. and Eichenlaub, R. (1981) J. Bacteriol. 147, 736-743
5. Eichenlaub, R., Figurski, D. and Helinski, D.R. (1977) Proc. Nat. Acad. Sci. USA 74, 1138-1141
6. Manis, J.J. and Kline, B.C. (1977) Mol. Gen. Genet. 152, 175-182
7. Eichenlaub, R., Wehlmann, H. and Ebbers, J. (1981) in Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids, Levy, S.B., Clowes, R.C. and Koenig, E.L. Eds., pp 327-336, Plenum, New York
8. Tanimoto, K. and Iino, T. (1984) Mol. Gen. Genet. 196, 59-63
9. Lane, D., Hill, D., Caughey, P. and Gunn, P. (1984) J. Mol. Biol. 180, 267-282

10. Caughey, P.A. (1986) PhD Thesis, University of Auckland
11. Manis, J. and Kline, B. (1978) Plasmid 1, 480-491
12. Kilbane, J.J. and Malamy, M.H. (1980) J. Mol. Biol. 143, 73-93
13. O'Connor, M.B. and Malamy, M.H. (1984) J. Mol. Biol. 175, 263-264
14. Austin, S., Ziese, M. and Sternberg, N. (1981) Cell 25, 729-736
15. Bergquist, P.L., Jamieson, A.F., Gardner, R.C. and Lane, D. (1980) Plasmid 3, 165-178
16. McKenney, K., Shimatake, H., Court, D., Schmeissner, J., Brady, C. and Rosenberg, M. (1981) in Gene Amplification and Analysis, vol.2: Analysis of Nucleic Acids by Enzymatic Methods, Chirikjian, J. and Paps, T. Eds., pp 383-415, Elsevier, New York
17. Dougan, G. and Sherratt, D. (1977) Mol. Gen. Genet. 151, 151-160
18. Phua, S-H., Bergquist, P.L. and Lane, H.E.D. (1982) Mol. Gen. Genet. 188, 353-355
19. Berk, A.J. and Sharp, P.A. (1977) Cell 12, 21-32
20. Aiba, H., Adhya, S. and de Crombrugghe, B. (1981) J. Biol. Chem. 256, 11905-11910
21. Mulligan, M.E., Hawley, D.K., Entriken, R. and McClure, W.R. (1984) Nuc. Acids Res. 12, 789-800
22. Bergquist, P.L., Downard, R.A., Caughey, P.A., Gardner, R.C. and Lane, H.E.D. (1981) J. Bacteriol. 147, 888-899
23. Gardner, R.C., Caughey, P.A., Lane, D. and Bergquist, P.L. (1980) Plasmid 3, 179-192
24. Ogura, T. and Hiraga, S. (1983) Cell 32, 351-360
25. Miki, T., Yashioka, K. and Horiuchi, T. (1984) J. Mol. Biol. 174, 605-625
26. Murotsu, T., Matsubara, K. Sugisaki, H. and Takanami, M. (1981) Gene 15, 257-271
27. Komai, N., Nishizawa, T., Hayakawa, Y., Murotsu, T. and Matsubara, K. (1982) Mol. Gen. Genet. 186, 193-203
28. Bex, F., Karoui, H., Rokeach, L., Dreze, P., Garcia, L. and Couturier, M. (1983) EMBO J. 2, 1853-1861
29. Phua, S-H. (1984) PhD Thesis, University of Auckland
30. O'Connor, M.B., Kilbane, J. and Malamy, M.H. (1986) J. Mol. Biol.
31. Hoess, R.H. and Abremski, K. (1985) J. Mol. Biol. 181, 351-362
32. Stoker, N.G., Fairweather, N.F. and Spratt, B.G. (1982) Gene 18, 335-341
33. Watson, L.A., Phua, S-H., Bergquist, P.L. and Lane, H.E.D. (1982) Gene 19, 173-178
34. Gardner, R.C., Malcolm, L., Bergquist, P.L. and Lane, H.E.D. (1982) Mol. Gen. Genet. 188, 345-352